European Statistical System handbook for quality and metadata reports

2020 edition

MANUALS AND GUIDELINES

eurostat

European Statistical System handbook for quality and metadata reports

2020 edition

Manuscript completed in February 2020.

Luxembourg: Publications Office of the European Union, 2020

Luxembourg: Publications Office of the European Union, 2020

Collection: Manuals and guidelines
Theme: General and regional statistics

PDF ISBN 978-92-76-09154-7 ISSN 2363-197X doi: 10.2785/666412 KS-GQ-19-006-EN-N
PRINT ISBN 978-92-76-09154-7 ISSN 2315-0815 doi: 10.2785/666412 KS-GQ-19-006-EN-N

FOREWORD

The *ESS Handbook for Quality and Metadata Reports (EHQMR)* is recognised as an ESS standard. It is included in the Catalogue of ESS standards, the collection of non-legislative normative documents underpinning the ESS. It is therefore a visible component of the ESS standardisation process, the importance of which goes well beyond making the current body of standards accessible.

Standardisation is important for modernising statistical production, in particular, to make statistical processes more efficient and robust, and to improve reporting on the quality of processes and output. Standards support the sharing of knowledge-based experiences and methodologies and, consequently, the sharing of tools, data, services and resources. The implementation of standard methods and tools also ultimately improves the comparability of statistical outputs, thus benefiting users of statistics.

As a standard, this publication encourages sharing approaches to European statistics and the spread of best practices, in particular for quality and metadata reporting. It aims to promote harmonised quality reporting across statistical processes and Member States, and thus to facilitate cross-comparisons of processes and outputs. The Handbook applies to Eurostat, national statistical institutes and other national authorities in their roles as producers, compilers and disseminators of European statistics. It supports compliance with the *European Statistics Code of Practice* by providing recommendations on how to prepare comprehensive quality reports for the full range of statistical processes and their outputs.

Quality reporting is a well-established area with a long tradition in the European Statistical System (ESS). The first ESS quality guidelines for standard quality reports were adopted in 2003, accompanied by the first *ESS Handbook on Quality Reports (EHQR)*, which was further extended in 2009 to cover all types of statistical processes and incorporate the standard ESS Quality and Performance Indicators. The *ESS Standard for Quality Reports Structure (ESQRS)*, a more detailed quality reporting structure, was launched in 2010.

At the same time, ESS metadata principles and practices were being developed. The *Euro-SDMX Metadata Structure (ESMS)* was set out in Commission recommendation 2009/498/EC, with the objective of harmonising reference metadata and facilitating exchange of such metadata within the ESS. While the ESQRS was aimed at designers of statistical processes and producers of statistics, the ESMS focused on users and statistical outputs.

In order to streamline and simplify quality reporting, the two standards – ESQRS and ESMS – were combined in the *Single Integrated Metadata Structure (SIMS)*, which was published in 2013. Minor updates were made in 2015 and SIMS 2.0 was adopted by the European Statistical System Committee. Its aim was to create an integrated and consistent quality and metadata reporting framework, harmonising and streamlining reporting across the statistical domains and countries.

This 2020 edition of the *ESS Handbook for Quality and Metadata Reports (EHQMR)* fully incorporates SIMS 2.0 and the two standards – ESQRS and ESMS – which are unified in SIMS. The document provides guidelines for producer reports and user reports within the overarching SIMS framework. Producer reports focus on quality aspects, while user reports focus on satisfying user needs for metadata (information about the data). However, both types of reports comprise metadata and both include quality metadata.

The Handbook includes many additional examples of reports. The focus has been on including real examples rather than artificial ones. Thus, while many examples are complete and well written, there are also some that are not so good and their limitations are highlighted. In addition, the Handbook contains new material on administrative data, big data, multi-source processes, etc.

The Handbook has been prepared under the guidance of Eurostat's Quality Team by two consultants, Michael Colledge and Jörgen Dalén, with extensive input from national statistical institutes and Eurostat units. I would like to thank them all as well as all the other colleagues in the ESS who have helped to prepare this publication.

Mariana Kotzeva

Director-General Eurostat

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

BOP	Balance of payments
CPI	Consumer price index
DESAP	Development of a Self-Assessment Programme (checklist)
EFQM	European Foundation for Quality Management
EFTA	European Free Trade Association
EGR	EuroGroups Register
EHQMR	ESS Handbook for Quality and Metadata Reports
EHQR	ESS Handbook for Quality Reports
ES CoP	European Statistics Code of Practice
ESMS	Euro-SDMX Metadata Structure
ESQRS	ESS Standard for Quality Reports Structure
ESS	European Statistical System
ESSC	European Statistical System Committee
ESS-MH	ESS Metadata Handler
EU	European Union
EU-SILC	EU Statistics on Income and Living Conditions
GSBPM	Generic Statistical Business Process Model
HBS	Household budget survey
HICP	Harmonised index of consumer prices
ICT	Information and Communications Technology
IMF	International Monetary Fund
NA	National accounts
NACE	EU statistical classification of economic activities
NSA	National statistical authority, meaning NSI or ONA
NSI	National statistical institute, the leading NSA in a country
OECD	Organisation for Economic Cooperation and Development
OOH	Owner-Occupied Housing
ONA	Other national authority (producing official statistics)
(P)	Concept/sub-concept for producer reports only
PPP	Purchasing power parity
QAF	Quality assurance framework
QPI	Quality and performance indicator
RADAR	Results, approaches, deploy, assess and refine
SBR	Statistical business register
SDMX	Statistical Data and Metadata Exchange
SIMS	Single Integrated Metadata Structure
(U)	Concept/sub-concept for user reports only

PART I CONTEXT

1 (Part I) Introduction

1.1 Overview

The general aim of the ***ESS Handbook for Quality and Metadata Reports (EHQMR)***, i.e., this document, is to provide guidelines for the preparation of producer and user reports for the full range of statistical processes and their outputs within Member States, EFTA countries and Eurostat. In this context,

- the term *statistical process* refers to a survey, administrative data process, or macro-aggregate compilation conducted by a national statistical authority or by Eurostat;
- the term *statistical output* refers to data that are disseminated together with the related services;
- the term *metadata* here refers to descriptions of the statistical process and the concepts underlying its outputs and their quality, more precisely called *reference metadata*;
- the term *national statistical authority (NSA)* refers to the *national statistical institute (NSI)* that plays the lead role in a *national statistical system (NSS)* or to any *other national authority (ONA)* that produces official statistics of relevance to the *European Statistical System (ESS);*
- the term *statistical authority* refers to an NSA or to Eurostat;
- the term *statistical organisation* is used as a synonym for statistical authority.

The EHQMR (also referred to as *the Handbook* where the context is clear) provides explicit guidelines for two types of report:

- a *producer report*, also (more precisely) called a *producer-oriented report*, and (less precisely) a *quality report* **-** comprising metadata, especially quality metadata, for use (1) within the NSA to record quality problems and improvements and (2) by Eurostat to review and summarise quality across NSAs;
- a *user report*, also (more precisely) called a *user-oriented report*, and (less precisely) a *metadata report* - comprising metadata, including quality metadata, that are intended for users of the statistical outputs, enabling them to assess whether the outputs are appropriate for the purposes they have in mind.

A producer report should be structured according to the *ESS Standard for Quality Reports Structure (ESQRS) V2.0*, meaning the report comprises descriptions of the concepts and sub-concepts specified in ESQRS V2.0.

A user report should be structured according to the *Euro-SDMX Metadata Structure (ESMS) V2.0,* meaning the report comprises descriptions of the concepts and sub-concepts specified in the ESMS V2.0*.*

ESQRS V2.0 and the ESMS V2.0 are subsets of the *Single Integrated Metadata System (SIMS) V2.0*. The majority of its 19 concepts and 80 sub-concepts are common to both ESQRS V2.0 and ESMS V2.0. Some belong only to ESQRS V2.0; others only to ESMS V2.0. This structure enables *once for all purposes reporting*, meaning that, for a particular statistical process, *for each concept and sub-concepts that is common to both structures*, exactly the same description can be used in a user report as in a producer report.

The Handbook can be viewed as the guidelines accompanying SIMS V2.0 and its constituent structures ESQRS V2.0 and ESMS V2.0*.* The SIMS V2.0 concepts, sub-concepts and their descriptions, exactly as in the SIMS structure, are its starting point. The Handbook includes revised guidelines, adds *examples of reports*; and, for selected concepts and sub-concepts, it provides additional *background information* and/or *further guidelines*.

The specific objectives of the Handbook are:

- to promote harmonised producer and user reporting for each type of statistical process and its outputs across NSAs, hence facilitating comparisons across Member States and EFTA countries;
- to promote harmonised producer and user reporting across statistical processes and their outputs within an NSA, hence facilitating comparisons across processes and outputs;
- to ensure that producer reports contain all the information required to facilitate identification of quality problems and potential improvements in statistical processes and their outputs; and
- to ensure that user reports contain all the information required by users to assess whether statistical outputs are fit for the purposes they have in mind.

1.2 Users and uses

The Handbook is addressed to:

NSA staff

1. preparing reports for dissemination to users;
2. preparing reports for internal purposes, especially quality assessment and improvement;
3. submitting user or producer reports to the corresponding Eurostat units;

Eurostat staff

4. preparing reports for users of European statistics;
5. preparing reports for internal purposes, especially quality improvement;
6. summarizing process and output quality across the Member States and EFTA countries (based on NSA reports) and reporting to stakeholders, for example, to the European Parliament or the Council; and
7. preparing statistical regulations or guidelines and wishing to incorporate material on quality and metadata reporting.

1.3 Changes from previous versions

SIMS V2.0 was formed by integrating and harmonising the original ESQRS and the ESMS structures so that all concepts in the constituent structures were included, did not overlap, and appeared once only. The Handbook is built on SIMS V2.0.

- It covers producer and user reports.
- It expands and supersedes the ESS Handbook for Quality Reports (EHQR) 2014 by fully incorporating all SIMS V2.0 concepts and by updating the guidelines that the EHQR provided.
- It supersedes the Single Integrated Metadata Structure and its Technical Manual - 2014 by referring to SIMS V2.0 and by updating the guidelines that the Manual provided.
- It supplements the current ESS Quality Glossary by including additional terms relating to metadata and quality.
- It refines the typology of statistical processes.
- It contains additional material on accuracy, administrative data processes, and big data.

1.4 Content of document

The remaining chapters of Part I summarise the basis on which the guidelines in Part II have been constructed.

- Chapter 2 describes the *ESS common quality framework*, in the context of which producer reports are prepared.
- Chapter 3 discusses the ways of describing statistical processes and their outputs; it covers SIMS V2.0 and the Generic Statistical Business Process Model (GSBPM) Version 5.1 and their relationship to one another. It provides the context for user reports.
- Chapter 4 classifies statistical processes into mutually exclusive types that need to be distinguished in reporting on accuracy and on some other concepts. It also introduces *big data*, not as a separate type of statistical process but rather as data source having particular characteristics including high volume, variability and velocity.
- Chapter 5 describes in more detail the various types of producer and user reports, including preparation of reports in accordance with sectoral regulations.
- Chapter 6 indicates the structure of the guidelines (in Part II) and the storage options for producer and user reports in the ESS Metadata Handler.

Part II is the core of the document. It comprises the guidelines for preparation of producer and user reports. The chapters are organised by SIMS V2.0 concept.

Part III contains copies of key reference documents, including the expanded ESS Quality and Metadata Glossary, SIMS V2.0 and its substructures, ESQRS V2.0 and ESMS V2.0, the ESS quality and performance indicators, sectoral regulations having quality reporting references, an introduction to big data, and a list of other key reference documents.

2 (Part I) ESS Common Quality Framework

2.1 Introduction

The Handbook is based on the *ESS Common Quality Framework (ESS CQF)*. The Preamble of the European Statistics Code of Practice (ES CoP) (2017) states that the *ESS CQF* is

> *composed of the European Statistics Code of Practice, the Quality Assurance Framework of the European Statistical System and the general quality management principles (such as continuous interaction with users, commitment of leadership, partnership, staff satisfaction, continuous improvement, integration and harmonisation).*

The ESS CQF complements the ESS legal framework, which is based on the Regulation (EC) No 223/2009 on European statistics. The ES CoP includes the *ESS Quality Declaration* which demonstrates quality awareness in the ESS and the commitment of its members to continuously developing, producing and disseminating high-quality European statistics and services in order to sustainably provide value to its users.

The ESS CQF incorporates general quality management concepts and principles, and it specialises them to the particular situation of statistical authorities in the ESS. Its implementation is underpinned by statistical regulations, and facilitated by quality related standards, guidelines, methods and tools, many of which are accessible via the Quality overview pages and via the dedicated sections of the various statistical domains of the Eurostat website.

This chapter describes the ESS CQF in detail, beginning with general quality management concepts and principles, then summarising the ES CoP and accompanying ESS Quality Assurance Framework (ESS QAF), then drawing attention to pertinent statistical regulations, and finally outlining relevant standards, guidelines, methods and tools. This is the context within which reports that focus on quality are prepared.

2.2 General quality management concepts and principles

2.2.1 GENERAL DEFINITIONS

The ES CoP does not formally define *quality* or quality related concepts. By implication, it depends upon the definitions in the ESS Quality Glossary and assumes that NSAs will adopt general quality management principles, whether through explicit statements or implicitly.

As there is potential for confusion between closely related terms such as *quality management* and *quality assurance*, all the quality related terms used in this Handbook are defined in the *ESS Quality and Metadata Management Glossary*, which is a revised and extended version of the ESS Glossary, included as Supplementary Document A in Part III.

The definitions of *general quality related concepts* are derived from the ISO 9000 family of standards, which are the most widely used quality standards in the world, in particular ISO 9000:2015 Quality management systems - Fundamentals and vocabulary. The definitions are expressed in general terms that apply to any organisation, not specifically a statistical authority. Each definition is accompanied by explanatory notes, indicating amongst other things, how the definition is applied in the ESS context.

For ease of reference, the key definitions are presented below with the pertinent notes.

***Quality* is the degree to which a set of inherent characteristics of an object fulfils requirements.**

- In the ESS context, the *object* may be a statistical product, service, process, system, methodology, organisation, resource, or input. *Characteristic* means distinguishing feature. *Inherent* means existing in the object, not assigned to it (such as a price). *Requirement* means need or expectation that is stated, generally implied or obligatory.
- Quality is a multi-faceted concept, The ES CoP lists five output principles, namely: relevance; accuracy; timeliness and punctuality; accessibility and clarity; and comparability and coherence.

***Quality management* comprises all the activities that an organisation uses to direct, control, and coordinate quality.**

- Quality management includes formulating a quality policy, setting quality objectives, quality planning, quality assurance, quality control, and quality improvement.

In the ESS context:

- quality management is defined (in SIMS) as the systems and frameworks in place within an organisation to manage the quality of statistical products and processes;
- quality management is deemed to cover a statistical authority as a whole, in contrast to quality assurance which focusses the core business of the authority, i.e. development, production and dissemination of statistics.
- In SIMS, and this Handbook, quality management includes quality assurance (S.11.1), quality assessment (S.11.2) and quality documentation (S.10.7).

A *quality management system* is a management system to direct and control an organisation with regard to quality.

- A quality management system (QMS) comprises a set of interrelated or interacting elements that an organisation uses to formulate quality policies and quality objectives and to establish the processes that are needed to ensure that policies are followed and objectives are achieved.
- In the ESS context, a distinction is made between a *general* QMS, which can apply to any organisation no matter what its core business and a *statistical* QMS, which applies exclusively to a statistical authority. The latter is more commonly referred to as a quality management framework, *quality assurance framework*, or simply a quality framework.

***Quality management principles* are principles on which a quality management system is based.**

- In the ESS context a distinction is made between *general quality management principles*, which are typically derived from a general QMS and associated with a statistical authority as a whole, and the (statistical) principles that are specified in the ES CoP and are associated with the core statistical environment, processes and outputs.

***Quality assurance* is the part of quality management focused on providing confidence that quality requirements are fulfilled.**

- In the ESS context, quality assurance focusses on the core business of a statistical authority, i.e. development, production and dissemination of statistics. It is an authority's guarantee that the products and services it offers meet the requirements for the statistical outputs. It is implemented via a *quality assurance framework*.

- It includes quality assessment.

A *quality assurance framework* is the set of procedures and systems that support quality assurance within an organisation.

- A quality assurance framework (QAF), sometimes referred to simply as a *quality framework*, covers the statistical outputs, the processes by which they are produced and the organisational environment within which the processes are conducted.
- The distinguishing characteristics of a QAF are that: it provides an umbrella for quality practices; it refers to a range of surveys/statistical processes or the entire statistical programme rather than a single survey/process. It covers all aspects of data processing and output, not just a single aspect, and it typically includes a template that can be used for quality assessment.

***Quality assessment* is the aspect of quality assurance that focuses on the extent to which statistical outputs and the processes that produced them meet quality requirements.**

- A quality report is a typical way of recording the results of a quality assessment.

2.2.2 GENERAL QUALITY MANAGEMENT PRINCIPLES

A commonly used expression of general quality management principles is provided by the ISO 9000:2015 Quality Management Systems - Fundamentals and vocabulary standard, as follows.

Principle 1 – Customer focus

Organisations depend on their customers and therefore should understand current and future customer needs, should meet customer requirements and strive to exceed customer expectations.

Principle 2 – Leadership

Leaders establish unity of purpose and direction of the organisation. They should create and maintain the internal environment in which people can become fully involved in achieving the organisation's objectives.

Principle 3 – Engagement of people

People at all levels are the essence of an organisation and their full involvement enables their abilities to be used for the organisation's benefit.

Principle 4 – Process approach

A desired result is achieved more efficiently when activities and related resources are managed as a process.

Principle 5 – Improvement

Improvement of the organisation's overall performance should be a permanent objective of the organisation.

Principle 6 – Evidence-based decision making

Effective decisions are based on the analysis of data and information.

Principle 7 – Relationship management

An organisation and its external providers (suppliers, contractors, service providers) are interdependent and a mutually beneficial relationship enhances the ability of both to create value.

Some NSAs may prefer the slightly different formulation presented in the European Foundation for Quality Management (EFQM) Excellence Model, which defines eight *fundamental concepts* of excellence:

- Succeeding through the talent of people;
- Sustaining outstanding results;
- Adding values for customers;
- Creating a sustainable future;
- Developing organisational capacity;
- Harnessing creativity and innovation;
- Leading with vision, inspiration and integrity;
- Managing with agility.

There are other formulations of general quality management principles, including Lean Six Sigma and Balanced Scorecard.

An NSA may, or may not, explicitly adopt and publicise a particular set of general quality management principles. In the former case, it may decide to seek quality certification for the organisation as a whole using a quality management standard. Several NSAs have sought and obtained quality certification according to ISO 9001: 2015. Others, including Eurostat, have used the EFQM Excellence Model.

2.3 ES Code of Practice and ESS Quality Assurance Framework

2.3.1 INTRODUCTORY REMARKS

The general quality management definitions and concepts in Section 2.2 are largely applicable to organisations in general and only lightly finetuned to the ESS context, for example by noting that the *organisations* under consideration are *statistical authorities* and *customers* are *users*. The ES CoP assumes these definitions and concepts and builds on them in the specific context of the ESS. It provides the framework for managing ESS statistical quality by setting out 16 principles for developing, producing and disseminating European statistics. The principles (which are reproduced in the following subsections for ease of reference) are in three groups:

- institutional environment (7 principles);
- statistical processes (4 principles); and
- statistical outputs (5 principles).

Each principle is accompanied by indicators that reflect good practice and show how compliance with the principle can be demonstrated.

Implementation of the ES CoP is facilitated by the ESS QAF, which provides methods for verifying conformance.

2.3.2 INSTITUTIONAL ENVIRONMENT PRINCIPLES

Institutional and organisational factors have a significant influence on the effectiveness and credibility of a statistical authority developing, producing and disseminating European Statistics.

- *Principle 1: Professional independence.* Professional independence of statistical authorities from other policy, regulatory or administrative departments and bodies, as well as from private sector operators, ensures the credibility of European Statistics.

  Principle 1bis: Coordination and Cooperation. National Statistical Institutes and Eurostat ensure the coordination of all activities for the development, production and dissemination of European statistics at the level of the national statistical system and the European Statistical System, respectively. Statistical authorities actively cooperate within the partnership of the European Statistical System, so as to ensure the development, production and dissemination of European statistics.

- *Principle 2: Mandate for Data Collection and Access to Data*. Statistical authorities have a clear legal mandate to collect and access information from multiple data sources for European statistical purposes. Administrations, enterprises and households, and the public at large may be compelled by law to allow access to or deliver data for European statistical purposes at the request of statistical authorities.

- *Principle 3: Adequacy of Resources.* The resources available to statistical authorities are sufficient to meet European Statistics requirements.

- *Principle 4: Commitment to Quality*. Statistical authorities are committed to quality. They systematically and regularly identify strengths and weaknesses to continuously improve process and output quality.

- *Principle 5: Statistical Confidentiality and Data Protection*. The privacy of data providers, the confidentiality of the information they provide, its use only for statistical purposes and the security of the data are absolutely guaranteed.

- *Principle 6: Impartiality and Objectivity*. Statistical authorities develop, produce and disseminate European Statistics respecting scientific independence and in an objective, professional and transparent manner in which all users are treated equitably.

2.3.3 STATISTICAL PROCESS PRINCIPLES

European and other international standards, guidelines and good practices are fully observed in the statistical processes used by the statistical authorities to develop, produce and disseminate European Statistics, while constantly striving for innovation. The credibility of the statistics is enhanced by a reputation for good management and efficiency.

- *Principle 7: Sound Methodology*. Sound methodology underpins quality statistics. This requires adequate tools, procedures and expertise.

- *Principle 8: Appropriate Statistical Procedures*. Appropriate statistical procedures, implemented throughout the statistical processes, underpin quality statistics.

- *Principle 9: Non-Excessive Burden on Respondents*. The burden is proportionate to the needs of the users and is not excessive for respondents. The statistical authorities monitor the response burden and sets targets for its reduction over time.

- *Principle 10: Cost Effectiveness*. Resources are used effectively.

2.3.4 STATISTICAL OUTPUT PRINCIPLES

Available statistics meet users' needs. Statistics comply with the European quality standards and serve the needs of European institutions, governments, research institutions, business concerns and the public generally.

- *Principle 11: Relevance*. European statistics meet the needs of users.

- *Principle 12: Accuracy and Reliability*. European Statistics accurately and reliably portray reality.

- *Principle 13: Timeliness and Punctuality*. European statistics are released in a timely and punctual manner.

- *Principle 14: Coherence and Comparability*. European statistics are consistent internally, over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different data sources.

- *Principle 15: Accessibility and Clarity*. European statistics are presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.

2.3.5 ESS QUALITY ASSURANCE FRAMEWORK

The *ESS Quality Assurance Framework (ESS QAF)* identifies possible activities, methods and tools that can provide guidance and evidence for the implementation of the ES CoP indicators. They are provided at process level and at institutional level so that support and evaluation can take place by statistical process or for the NSA as a whole. The document supports reviews of conformance with the ES CoP.

The previous version ESS QAF V1.2 was adopted by the ESSC in May 2015 and covered all the indicators associated with Principles 4-15 of the ES CoP (2011). The current version ESS QAF V2.0 adopted by the ESSC in May 2019 has been extended to cover the remaining principles and to reflect the changes introduced in the 2017 edition of the ES CoP.

2.4 Quality related regulations

2.4.1 REGULATION ON EUROPEAN STATISTICS

Extracts from the articles of Regulation (EC) No 223/2009 on European statistics amended by Regulation 2015/759 that are of particular relevance to ESS quality are reproduced below.

Article 11 European statistics Code of Practice

1. The Code of Practice shall aim at ensuring public trust in European statistics by establishing how European statistics are to be developed, produced and disseminated in conformity with the statistical principles as set out in Article 2(1) and best international statistical practice.
3. Member States and the Commission shall take all the necessary measures to maintain confidence in European statistics. To this effect, ‘Commitments on Confidence in Statistics’ (the Commitments) by Member States and by the Commission shall further aim to ensure public trust in European statistics and progress in the implementation of the statistical principles contained in the Code of Practice…
4. The Commitments by Member States shall be monitored regularly by the Commission on the basis of annual reports sent by Member States and shall be updated as necessary…

Article 12 Statistical quality

1. To guarantee the quality of results, European statistics shall be developed, produced and disseminated on the basis of uniform standards and of harmonised methods. In this respect, the following quality criteria shall apply:
   (a) ‘relevance’, which refers to the degree to which statistics meet current and potential needs of the users;
   (b) ‘accuracy’, which refers to the closeness of estimates to the unknown true values;
   (c) ‘timeliness’, which refers to the period between the availability of the information and the event or phenomenon it describes;
   (d) ‘punctuality’, which refers to the delay between the date of the release of the data and the target date (the date by which the data should have been delivered);
   (e) ‘accessibility’ and ‘clarity’, which refer to the conditions and modalities by which users can obtain, use and interpret data;
   (f) ‘comparability’, which refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time;
   (g) ‘coherence’, which refers to the adequacy of the data to be reliably combined in different ways and for various uses.
2. Specific quality requirements, such as target values and minimum standards for the production of statistics, may also be laid down in sectoral legislation.

   In order to ensure the uniform application of the quality criteria laid down in paragraph 1 to the data covered by sectoral legislation in specific statistical domains, the Commission shall adopt implementing acts laying down the modalities, structure and periodicity of quality reports covered by sectoral legislation. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 27(2).

3. Member States shall provide the Commission (Eurostat) with reports on the quality of data transmitted, including any concerns they have regarding the accuracy of the data. The Commission (Eurostat) shall assess the quality of the data transmitted, on the basis of appropriate analysis, and shall prepare and publish reports and communications on the quality of European statistics.
4. In the interest of transparency, the Commission (Eurostat) shall, where appropriate, make public its assessment of the quality of national contributions to European statistics.
5. Where sectoral legislation provides for fines in cases where Member States misrepresent statistical data, the Commission may, in accordance with the Treaties and such sectoral legislation, initiate and conduct investigations as necessary including, where appropriate, on-site inspections in order to establish whether such misrepresentation was serious and intentional or grossly negligent

2.4.2 COMPLIANCE WITH THE ES COP

Compliance with the ES CoP is regularly monitored through the ESS-wide exercise of peer reviews. Typically, a review starts with a national self-assessment questionnaire followed by the peer review using the ESS QAF. Improvement actions identified in the peer review are subsequently monitored and reported upon on an annual basis.

2.4.3 REGULATIONS CONCERNING QUALITY IN SPECIFIC STATISTICAL DOMAINS

The regulations that create the legal basis for the provision of European statistics in the various domains include quality related requirements. At a minimum they provide criteria relating to relevance by specifying the needs for European statistics. They may also set methodological standards or targets for accuracy, timeliness and comparability.

The document entitled Quality Reporting - Quality Requirements/Standards provides a comprehensive and recently updated list of all relevant regulations.. For ease of reference, it is reproduced as Supplementary Document D in Part III.

2.5 Other quality standards, guidelines, methods and tools

2.5.1 INTRODUCTORY REMARK

In addition to the ES CoP and the ESS QAF, the Eurostat quality web pages provide links to other ESS quality standards, guidelines, methods and tools that may be of use as the basis for preparing quality and metadata reports, as outlined in the following sections.

2.5.2 ESS QUALITY GLOSSARY

The ESS Quality Glossary, first published in 2003, was transferred to the Concepts and Definitions Database (CODED), where it is now available as a theme. It covers many of the technical terms used in discussing quality, providing a short definition of each term and indicating the source of the definition.

As previously noted, for the purposes of this Handbook, the Glossary has been revised and expanded to become the *ESS Quality and Metadata Glossary*. It is included as Supplementary Document A in Part III.

2.5.3 ESS QUALITY AND PERFORMANCE INDICATORS

The ESS Quality and Performance Indicators 2014 are a standard set of indicators covering significant aspects of quality and performance in a standardised way. They have been incorporated in the Single Integrated Metadata Structure (SIMS), as further discussed in Chapter 3.

The indicators are accompanied by guidelines which, for each indicator, give the definition, applicability, calculation formulae, target value, aggregation levels, interpretation, and references. The guidelines are referenced by SIMS but not part of SIMS.

For ease of reference the indicators and guidelines are reproduced as Supplementary Document C in of Part III.

2.5.4 OTHER ESS QUALITY GUIDELINES, METHODS AND TOOLS

The Handbook on Data Quality - Assessment Methods and Tools details the full range of methods for assessing process and output quality and the tools that support them.

The Handbook on improving quality by analysis of process variables describes a general approach and useful tools for identifying, measuring and analysing key variables associated with a statistical process.

The European Self-Assessment Checklist for Survey Managers (DESAP) enables the conduct of quick but systematic and comprehensive quality assessments of a statistical process (survey, census or administrative data process) and its outputs and identification of potential improvements. The documents are also available in electronic form: Electronic DESAP-E checklist and Electronic DESAP user guide and an abbreviated version is available as DESAP condensed.

Quality Indicators for the Generic Statistical Business Process Model (GSBPM) - For Statistics derived from Surveys and Administrative Data Sources provides a comprehensive set of quality indicators structured in accordance with the GSBPM.

In some individual statistical domains, domain-specific quality guidelines, methods and tools have been developed. The level and complexity of these vary from one domain to another. In principle, they all fit within the overall ESS common quality framework. For example, the European Committee on Monetary, Financial and Balance of Payment Statistics (CMFB) has defined three levels of quality assurance and reporting, as detailed in the ESS-ECB quality assurance framework for statistics underlying the Macroeconomic Imbalances Procedure (MIP) Scoreboard.

In addition to the documents referenced above, which are at European level, many NSAs have compiled and published quality frameworks, guidelines and reporting tools. They are too numerous to reference here.

3 (Part I)

Describing statistical processes and outputs

3.1 Introductory remarks

This chapter describes two standards that are of great significance in preparing user and/or producer reports. The first, and most important of the two, is the *Single Integrated Metadata Structure (SIMS*). It is a template facilitating harmonised and efficient preparation of both producer and user reports. As noted in Chapter 1 and detailed below, it is the union of two structures, namely the *ESS Standard for Quality Reports Structure (ESQRS)*, and the *Euro-SDMX Metadata Structure (ESMS)*. It fully satisfies the needs of users for metadata about the statistical outputs (and processes by which they were produced). At the same time, it supports the detail required for producer reports. It is essentially an output-oriented way of viewing a statistical process and is the chosen basis for user and producer reports in the ESS.

Copies of SIMS, ESQRS and ESMS are included within Supplementary Document B in Part III.

SIMS references the 16 ESS standard quality and performance indicators (QPIs). As previously noted, their descriptions and compilation formulae are included as Supplementary Document C in Part III.

The second standard is the *Generic Statistical Business Process Model (GSBPM)*. The most recent version is 5.1. Although it was developed by the UNECE it has been accepted as an ESS standard. It facilitates the harmonised design, development, implementation and evaluation of statistical processes across the various types of statistical process, and across countries. As its name suggests, it is a process-oriented way of viewing a statistical process. Whilst it is not the standard for ESS producer or user reports, it is the basis on which many countries now design and document their statistical processes.

Sections 3.2 and 3.3 describe SIMS and GSBPM respectively and Section 3.4 discusses their inter-relationship.

3.2 Single Integrated Metadata Structure

3.2.1 SIMS OBJECTIVES AND CREATION

The aims of the SIMS are to:

- streamline and harmonise user and producer reports in the ESS;
- decrease the reporting burden on NSAs by creating the framework that enables *once for all purposes reporting*, where concepts that are common to user and producer reports are reported upon once for both purposes;

- create an integrated and consistent reporting framework where the reports can be stored in a single database;
- create a flexible and up to date system where future extensions are possible by adding new concepts.

As previously noted, SIMS was formed by integrating and harmonising two reporting structures, namely the *Euro-SDMX Metadata Structure (ESMS)* and the *ESS Standard for Quality Reports Structure (ESQRS)* so that all concepts in these structures are included, appear once only, and are consistent with the statistical standards in the *Statistical Data and Metadata Exchange* (*SDMX) Content-oriented Guidelines*.

SIMS provides the basis for *user reports*, for *producer reports***,** and for *once for all purposes* reporting.

3.2.2 SIMS STRUCTURE

SIMS V2.0, which was introduced in 2015, is the latest version of SIMS. It is a superset of the ESQRS and the ESMS, comprising 19 concepts and 80 sub-concepts. The concepts and sub-concepts are units of knowledge created by a unique combination of characteristics. Many are within the list of the standard SDMX cross-domain concepts and are therefore fully SDMX compliant. This has benefits for implementation systems.

- 13 concepts and 48 sub-concepts are common to the ESQRS and the ESMS. These are the *common* (or *shared*) concepts and sub-concepts for which *once for all purposes reporting* is applicable.
- 24 sub-concepts belong only to the ESQRS. They are indicated by (P) in the tabular displays of concepts in Part II.
- 10 of these are associated with Accuracy, 4 with Coherence and comparability, 3 with Timeliness and punctuality, 3 with Accessibility and clarity, 2 with Statistical presentation, 1 with Data revision, and 1 with Relevance.
- 8 are standard quality and performance indicators.
- 6 concepts and 8 associated sub-concepts belong only to ESMS. They are indicated by (U) in the tabular displays of concepts in Part II.

All SIMS concepts and sub-concepts are listed in Supplementary Document B2 and are presented together with their definitions and guidelines in Supplementary Document B6 in Part III.

SIMS concept S.13 is called *Accuracy and reliability*. It would be more appropriately entitled *Accuracy* as the concept of reliability is included in *S.17 Data revision* and is reflected in the SIMS guidelines associated with that concept.

SIMS incorporates the 16 standard ESS Quality and Performance Indicators (QPIs) within the sub-concepts (for user reports) and as sub-concepts of their own (for producer reports). The aim of these QPIs is to measure key aspects of quality and performance in a standard way. For datasets that are submitted to the ESS-MH, the values of some of these indicators may sometimes be compiled or pre-filled by Eurostat.

The definitions and compilation methods for the QPIs are specified in the ESS Guidelines for the Implementation of the ESS Quality and Performance Indicators, which, for ease of reference, is reproduced in Part III Supplementary Document C.

Six standard QPIs are common to ESQRS and ESMS. These are the ones for which *once for all purposes reporting* is appropriate. Two other standard QPIs are included in both ESQRS and ESMS but with different compilation formulae. The remaining QPIs are included in ESQRS only.

The inclusion of the ESS standard indicators in SIMS in no way precludes use, in addition or instead, of other indicators that might be more pertinent to the particular statistical process being discussed. For example, as noted in Section 2.5.4, Quality Indicators for the Generic Statistical Business Process Model provides a comprehensive set of quality indicators for surveys and administrative data processes. Macro-aggregate compilation processes such as the National Accounts, typically have their particular sets of quality indicators.

3.2.3 ESQRS AND PRODUCER REPORTS

The ESQRS is for producer reports. It contains all the concepts and sub-concepts that are required:

- to fully document the essential characteristics of a statistical process and its outputs from a quality perspective;
- to reflect their quality; and
- to determine possible improvements to address quality concerns.

The ESQRS includes 13 SIMS concepts and 72 sub-concepts. They are listed in Supplementary Document B3. They are organised into 12 ESQRS concepts. The ESQRS concepts coincide with the SIMS concepts except that ESQRS 6 Accuracy and Reliability covers SIMS' S.13 Accuracy, S.17 Data Revision and S.18.6.1 Seasonal adjustment. (The last-mentioned inclusion seems a little unusual as seasonal adjustment is an analytical tool not an error correction or revision mechanism.)

As summarised in Table 3.1 below and detailed in Supplementary Document B5, the ESQRS ordering is quite different from SIMS, being more natural for the presentation of a producer report.

Table 3.1. ESQRS concepts and SIMS concepts

ESQRS Id	ESQRS Concept	SIMS Id	SIMS Concept
1	Contact	S.01	Contact
2	Statistical presentation	S.03	Statistical presentation
3	Statistical processing	S.18 (ex S18.6.1)	Statistical processing
4	Quality management	S.11	Quality management
5	Relevance	S.12	Relevance
6	Accuracy and reliability	S.13	Accuracy
		S.17	Data revision
		S.18.6.1	Seasonal adjustment
7	Timeliness and punctuality	S.14	Timeliness and punctuality
8	Coherence and comparability	S.15	Coherence and comparability
9	Accessibility and clarity	S.10	Accessibility and clarity
10	Cost and burden	S.16	Cost and burden
11	Confidentiality	S.07	Confidentiality
12	Comment	S.19	Comment

The ESQRS includes all 16 standard QPIs as separate sub-concepts, as indicated in Table 3.2. In this table, a compilation formula that is labelled:

- *specific to producer reports* implies that for the same QPI, there is a simplified formula for user reports;
- *only for producer reports* implies that the QPI is only for use in producer reports and is not included in user reports.

Table 3.2. Standard quality and performance indicators for producer reports

ESQRS	SIMS sub-concept	QPI code	QPI Name	Compilation formula
Relevance				
5.3.1	S.12.3.1	R1	Data completeness – rate	Specific to producer report
Accuracy and reliability				
6.2.1	S.13.2.1.	A1	Sampling error – indicators	Common to producer and user reports
6.3.1.1	S.13.3.1.1	A2	Over-coverage – rate	Only for producer report
6.3.1.2	S.13.3.1.2	A3	Common units – proportion	Only for producer report
6.3.3.1	S.13.3.3.1	A4	Unit non-response – rate	Common to producer and user reports
6.3.3.2	S.13.3.3.2	A5	Item non-response – rate	Common to producer and user reports
6.3.4.1	S.18.5.1	A7	Imputation – rate	Only for producer report
6.5	S.17.2.1	A6	Data revision – average size	Common to producer and user reports
Timeliness and punctuality				
7.1.1	S.14.1.1	TP1	Time lag – first results	Only for producer report
7.1.2	S.14.1.2	TP2	Time lag – final results	Common to producer and user reports
7.2.1	S.14.2.1	TP3	Punctuality – delivery and publication	Specific for producer report
Coherence and comparability				
8.1.1	S.15.1.1	CC1	Asymmetry for mirror flows statistics – coefficient	Only for producer report
8.2.1	S.15.2.1	CC2	Length of comparable time series	Common to producer and user reports
9.3.1	S.10.3.1	AC1	Data tables – consultations	Only for producer report
9.7.2	S.10.5.1	AC2	Metadata – consultations	Only for producer report
9.7.1	S.10.6.1	AC3	Metadata completeness – rate	Only for producer report

3.2.4 ESMS AND USER REPORTS

The ESMS is for user reports. It contains all the concepts and sub-concepts that are required by a user to understand the outputs, and how they were produced, in sufficient detail to determine if they are fit for the purpose the user has in mind.

The ESMS contains all 19 SIMS concepts and 56 sub-concepts. They are listed in Supplementary Document B4. They have the same numbering and ordering as SIMS., as detailed in Supplementary Document B5.

ESMS includes eight standard ESS QPIs within corresponding SIMS sub-concepts, as listed in Table 3.3 below. In this table, a compilation formula that is *specific to user report* implies that there is different formula for the same indicator for a producer report.

Table 3.3. Standard quality and performance indicators for user reports

Included in SIMS sub-concept	QPI Code	QPI Name	Compilation Formula
S.12.3	R1	Data completeness – rate	Specific to user report
S.13.2	A1	Sampling error – indicators	Common to producer and user reports
S.13.3	A4	Unit non-response – rate	Common to producer and user reports
S.13.3	A5	Item non-response – rate	Common to producer and user reports
S.14.1	TP2	Time lag - final results	Common to producer and user reports
S.14.2	TP3	Punctuality - delivery and publication	Specific to user report
S.15.2	CC2	Length of comparable time series	Common to producer and user reports
S.17.2	A6	Data revision – average size	Common to producer and user reports

3.2.5 FUTURE REVISION OF SIMS

The Handbook incorporates the concepts and definitions of SIMS V2.0 without change. As authorised by Eurostat, and with the aim of improving clarity, the Handbook includes revisions to, and extensions of, the SIMS guidelines in the SIMS Technical Manual.

It can be expected that SIMS will further evolve. This Handbook is a catalyst for such evolution. For example, improvements could include:

- changing the name of concept S.13 to Accuracy, or alternatively, merging S.17 with S.13 so that S.13 covers reliability as well as accuracy;
- better alignment of SIMS with GSBPM, as further discussed in Section 3.4.

Requests for future revision of SIMS concepts, definition or guidelines will be submitted to Eurostat. They will be analysed by a Task Force that will be set up periodically, comprising members of the Working Group on Quality in Statistics. The results of the work of the Task Force will be reviewed and approved by the Working Group and may require approval of the European Statistical System Committee (ESSC).

3.3 Generic Statistical Business Process Model

3.3.1 OVERVIEW OF GSBPM

As stated in a theme paper in Collaboration in Research and Methodology for Official Statistics:

> *the GSBPM is a means to describe statistics production in a general and process-oriented way. It is used both within and between statistical offices as a common basis for work with statistics production in different ways, such as quality, efficiency, standardisation, and process-orientation. It is used for all types of surveys, and "business" is not related to "business statistics" but refers to the statistical office, simply expressed.*

Thus, the aim of the GSBPM is to provide a standard structure that can be used to describe all types of statistical processes in a way that is harmonised within and across statistical authorities. Given that harmonisation across processes and authorities is a goal of this Handbook too, the GSBPM is ideally suited to be the standard that underpins the descriptions of statistical processes upon which producer and user reports are based.

The following paragraphs, which are extracted more or less verbatim from GSBPM V5.1, summarise its objectives and content. Some sentences are italicised to indicate their strong connection with quality and metadata reporting.

(Paragraph 1) The GSBPM…provides a standard framework and harmonised terminology to help statistical organisations to modernise their statistical production processes, as well as to share methods and components. *The GSBPM can also be used for integrating data and metadata standards, as a template for process documentation, for harmonizing statistical computing infrastructures, and to provide a framework for process quality assessment and improvement*.

(Paragraph 12) The GSBPM comprises three levels:

- Level 0, the statistical business process itself;
- Level 1, the eight phases of the statistical business process, which are: Specify Needs, Design, Build, Collect, Process, Analyse, Disseminate, and Evaluate;
- Level 2, the *sub-processes* within each phase – 44 sub-processes in all

(Paragraph 15). The GSBPM is intended to apply to all activities undertaken by producers of official statistics, at both the national and international levels, which result in data outputs.

(Paragraph 16) The model is designed to be applicable regardless of the data source, so it can be used for the description and quality assessment of processes based on surveys, censuses, administrative registers, and other non-statistical or mixed sources.

(Paragraph 17). Whilst typical statistical business processes include collection and processing of data to produce statistical outputs, the GSBPM also applies to cases where existing data are revised or time-series are re-calculated, either as a result of improved source data, or a change in methodology. In these cases, the input data can be original microdata and/or additional data, which are then processed and analysed to produce revised outputs. In such cases, it is likely that several sub-processes and possibly some phases (particularly the early ones) would be omitted. Similarly, the GSBPM can be applied to processes such as the compilation of National Accounts, and the typical processes in international statistical organisations that use secondary data from countries or other organisations.

3.3 2 GSBPM PHASES

Figure 3.1 indicates the eight phases and 44 sub-processes that comprise levels 2 and 3 of the GSBPM. The figure also indicates that there are *over-arching processes*, i.e., processes that span individual statistical processes and their phases. They include quality management, metadata management, data management, process data management, knowledge management and provider management

The eight phases are as follows.

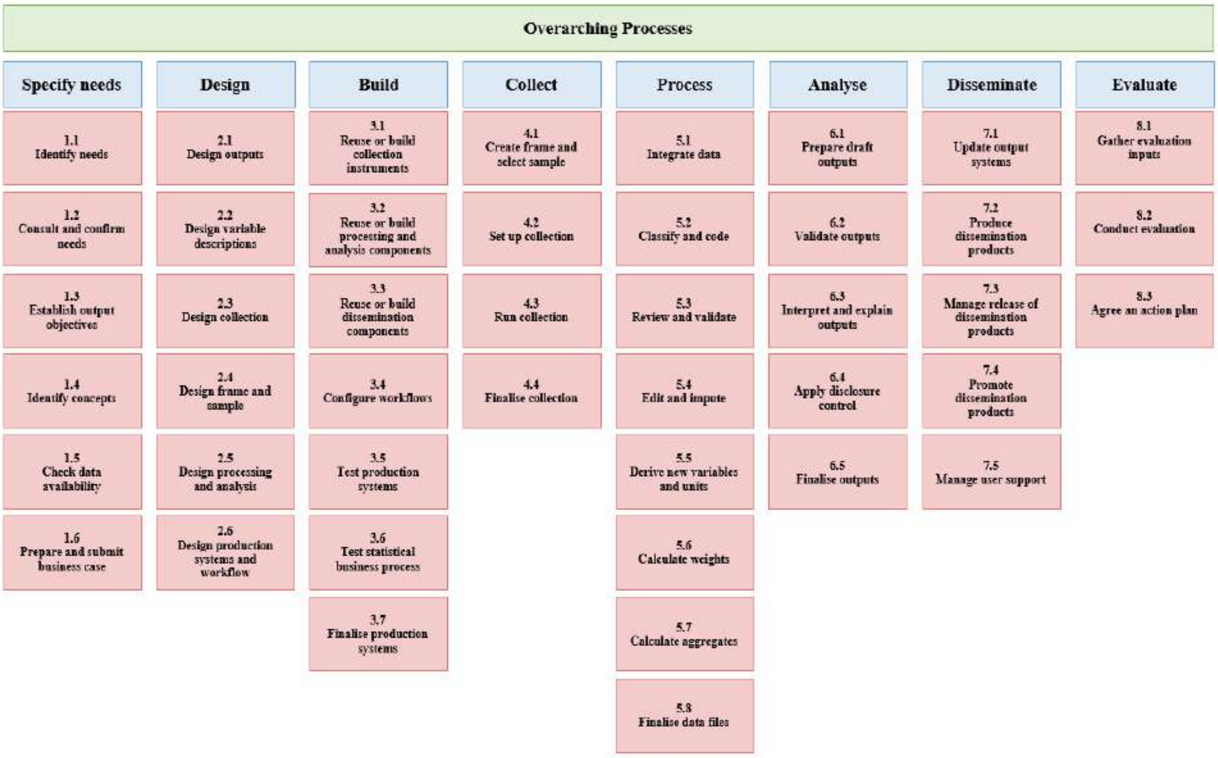
Phase 1: Specify needs

(Paragraph 34). This phase is triggered when a need for new statistics is identified, or feedback about current statistics initiates a review. It includes all activities associated with engaging stakeholders to identify their detailed statistical needs (current or future), proposing high level solution options and preparing business cases to meet these needs.

Phase 2: Design

(Paragraph 45). This phase describes the development and design activities, and any associated practical research work needed to define the statistical outputs, concepts, methodologies, collection instruments and operational processes. It includes all the design elements needed to define or refine the statistical products or services identified in the business case. This phase specifies all relevant metadata, ready for use later in the statistical business process, as well as quality assurance procedures. For statistical outputs produced on a regular basis, this phase usually occurs for the first iteration, and whenever improvement actions are identified in the Evaluate phase of a previous iteration.

Figure 3.1: GSBPM 5.1 phases and sub-processes



Phase 3: Build

(Paragraph 56). This phase builds and tests the production solution to the point where it is ready for use in the "live" environment. The outputs of the "Design" phase are assembled and configured in this phase to create the complete operational environment to run the process. New services are built by exception, created in response to gaps in the existing catalogue of services sourced from within the organisation and externally. These new services are constructed to be broadly reusable within the business architecture of the organisation where possible.

Phase 4: Collect

(Paragraph 66). This phase collects or gathers all necessary information (data, metadata and paradata), using different collection modes (e.g. acquisition, collection extraction, transfer), and loads them into the appropriate environment for further processing. Whilst it can include validation of data set formats, it does not include any transformations of the data themselves, as these are all done in the "Process" phase. For statistical outputs produced regularly, this phase occurs in each iteration.

Phase 5: Process

(Paragraph 75). This phase describes the processing of input data and their preparation for analysis. It is made up of sub-processes that integrate, classify, check, clean, and transform input data, so that they can be analysed and disseminated as statistical outputs. For statistical outputs produced regularly, this phase occurs in each iteration. The sub-processes in this phase can apply to data from both statistical and non-statistical sources (with the possible exception of sub-process 5.6 (Calculate weights), which is usually specific to survey data).

Phase 6: Analyse

(Paragraph 88). In this phase, statistical outputs are produced and examined in detail. It includes preparing statistical content (including commentary, technical notes, etc.), and ensuring outputs are “fit for purpose” prior to dissemination to users. This phase also includes the sub-processes and activities that enable statistical analysts to understand the statistics produced. The outputs of this phase could also be used as an input to other sub-processes (e.g. analysis of new sources as input to the “Design” phase). For statistical outputs produced regularly, this phase occurs in every iteration. The "Analyse" phase and sub-processes are generic for all statistical outputs, regardless of how the data were sourced.

Phase 7: Disseminate

(Paragraph 95). This phase manages the release of the statistical products to users. It includes all activities associated with assembling and releasing a range of static and dynamic products via a range of channels. These activities support users to access and use the outputs released by the statistical organisation. For statistical products produced regularly, this phase occurs in each iteration.

Phase 8: Evaluate

(Paragraph 103). This phase manages the evaluation of a specific instance of a statistical business process, as opposed to the more general over-arching process of statistical quality management... It can take place at the end of the instance of the process, but can also be done on an ongoing basis during the statistical production process. It relies on inputs gathered throughout the different phases. It includes evaluating the success of a specific instance of the statistical business process, drawing on a range of quantitative and qualitative inputs, and identifying and prioritising potential improvements..

3.4 Relationship between SIMS and GSBPM

SIMS and GSBPM are both templates designed to standardise the descriptions of statistical processes and their outputs. In other words, they have the same subject type, namely a statistical process and its outputs. However, they have different focuses as regards the information about a statistical process that they describe

- SIMS, being the union of ESMS and ESQRS, is a template for providing all the information a user may want to know about a statistical process and its outputs or a producer may want to know about the quality aspects.
- GSBPM is a template for a description of every aspect of a statistical process. It is not designed to focus on user needs for metadata, or on quality aspects, except in so far as these topics are covered in the User Needs and Evaluate Phases.

In summary, SIMS concentrates on particular aspects that are relevant to users and/or quality reporting whereas GSBPM gives description of a statistical process that has a more even degree of detail across the phases and sub-processes. Thus, SIMS and GSBPM cannot be expected to be expressed in terms of exactly the same set of concepts and sub-concepts, or to have exactly the same structure. Nevertheless, for simplicity in relating metadata for users and/or about quality (using SIMS) to general descriptions of the statistical process (using GSBPM) it is desirable SIMS and GSBPM should be harmonised to the fullest extent possible

A simplified view of the relationships between SIMS concepts/sub-concepts and GSBPM phases is shown in Table 3.3. Evidently there are ways in which SIMS and GSBPM could be brought more closely into alignment, and these will, no doubt, be explored in future revisions of the two standards.

Table 3.4 Relationships between SIMS concepts/sub-concepts and GSBPM phases

<table>
<tr><th colspan="2">SIMS concept/ sub-concept</th><th>Relationship to</th><th>GSBPM phase</th></tr>
<tr><td>S.12.1</td><td>User Needs</td><td>Subset of</td><td>Specify Needs</td></tr>
<tr><td>S.18.1</td><td>Source data</td><td rowspan="2">Subsets of</td><td rowspan="3">Design</td></tr>
<tr><td>S.18.2</td><td>Frequency of data collection</td></tr>
<tr><td rowspan="3">S.18.3</td><td rowspan="3">Data collection</td><td>Subset of</td></tr>
<tr><td>Subset of</td><td>Build</td></tr>
<tr><td>Covers</td><td>Collect</td></tr>
<tr><td>S.18.4</td><td>Data validation</td><td rowspan="3">Subsets of</td><td>Process</td></tr>
<tr><td>S.18.5</td><td>Data compilation</td><td rowspan="2">Analyse</td></tr>
<tr><td>S.18.6</td><td>Adjustment</td></tr>
<tr><td>S.07.1</td><td>Confidentiality policy</td><td>Subset of</td><td>Analyse</td></tr>
<tr><td>S.07.2</td><td>Confidentiality data treatment</td><td rowspan="4">Subsets of</td><td rowspan="4">Disseminate</td></tr>
<tr><td>S.08</td><td>Release policy</td></tr>
<tr><td>S.09</td><td>Frequency of dissemination</td></tr>
<tr><td>S.10</td><td>Accessibility and Clarity</td></tr>
<tr><td>S.12.2</td><td>User satisfaction</td><td rowspan="2">Subsets of</td><td rowspan="2">Evaluate</td></tr>
<tr><td>S.11.2</td><td>Quality assessment</td></tr>
</table>

4 (Part I)

Types of statistical process

4.1 Introduction

The types of statistical processes by which ESS statistics are produced are diverse. This diversity results in some complexity when it comes to user and producer reports since there are many different types of process and statistics for which to account.

- Direct collection of individual data from a sample of respondents for statistical purposes includes three types of process according to whether data are collected from all units, and if not, whether probability sampling or non-probability sampling is used. The corresponding statistical process types are: ***Probability Survey, Non-Probability Survey*** and ***Census Survey***.
- One type of process makes use of individual data that have been acquired for administrative purposes. This process is called an ***Administrative Data Process.***
- Two types of process make use of data from more than one of the above types of process. These processes are referred to as ***Multisource Process*** and ***Macro-aggregate Compilation Process.***

In many processes the target data item is clear-cut in the form of a population total, an average or a count, but in others it is a more abstract concept, like total production in monetary terms, or a price change. Thus, notwithstanding the overall desire to treat all statistical processes in essentially the same way, from the perspective of producer and user reports based on SIMS it is necessary to distinguish different types of statistical process for reporting of some concepts, in particular for reporting S.13 Accuracy, S.15 Coherence and Comparability, and S.18 Statistical Processing.

In this chapter, six mutually exclusive types of statistical process are distinguished, and each type is described in detail. The classification into types is based on data source and compilation method, as outlined above. There is still significant heterogeneity within each type and some special cases are discussed in Section 4.8. In addition, the notion of *big data* is described. It is not viewed as a different type of statistical process but rather as a characteristic of a data source, distinguished by its large volume, the distinctive ways in which the data are produced, their real-time or near real-time generation and their diverse structure.

In Part II of this handbook some guidelines vary according to type, notably those referring to reporting accuracy (S.13). Thus, it is recommended that each statistical process in the ESS system should be classified by type. This may be best done by the domain representatives at the ESS level except where the type is different across countries, in which case it is better done at the national level.

4.2 Probability survey

A survey based on probability sampling, involving direct collection of data from respondents, is the most analysed type of statistical process. There is a well-established body of theory regarding sample design and the types of error (sampling and non-sampling) that can occur. Inferences can be made from the sample about the population from which the sample is drawn. An inference normally uses sampling weights reflecting the inclusion probabilities of the sampling units.

The starting point for a probability survey is the *target population* for which a number of variables are to be measured at a certain point/period in time. The *survey population* is the list of units that is the closest approximation to the target population that can be created in practice. The list of units together with the data about them that are needed to sample them and to conduct the survey is the *survey frame*.

Where the target population differs from the *ideal population* from a user perspective, a relevance issue arises. An example of this situation is the use of so called cut-off sampling of businesses. Here the ideal population for important users may be *all businesses* (possibly in a certain industry) but the target population is *all businesses with more than X employees*.

The survey frame may be less than perfect in which case a coverage error occurs.

Not all respondents assigned for the survey may respond in which case non-response errors occur.

Responses from respondents may not be fully accurate resulting in measurement errors.

Errors may occur in the phase of recording the response variables leading to processing error.

Finally, the model used in compiling survey estimates may not reflect reality in which case there is a model assumption error.

Important subclasses of probability sampling include:

- Multi-stage sampling in which samples are drawn in two or more stages and for which frames are used in all stages. For example, a first stage may include a frame of geographical areas and in each area a frame of dwellings is constructed.
- Stratified samples with different probabilities for each stratum. Some strata may even be covered with probability one, constituting a census portion of the survey. This type of sampling design is common for business surveys, where it is desired to cover the largest enterprises in a take-all stratum.
- Samples based on land areas using geospatial data. Some agricultural surveys are of this kind.

4.3 Non-probability survey

Some surveys involve data collection from a set of sampling units but do not follow a rigorous probability design. The validity of inferences from such surveys depends, explicitly or implicitly, on model assumptions. A possible sub-classification of such surveys is([1]):

- *Quota surveys.* Here the sample is designed so that units are included in a way that ensures that sample frequencies agree with population frequencies for certain “background” variables for which population frequencies are known. Inferences from such a survey depend on the model assumption that there are no differences between the sample and the population other than those described by the known background variables.
- *Subjective cut-off selections* of largest, “most typical”, “most sold”, or the like units. This is akin to cut-off surveys, except that the inclusion criterion is not clearly defined and has a subjective component.
- *Voluntary surveys (opt-in surveys)*. Examples are web surveys where persons volunteer to

([1]) Other researchers have used other classifications, see e.g. Yang and Banamah (2013) or ILO (2009).

answer a questionnaire of some kind or to be included in a panel. In this case there is normally no plausible assumption that relates the sample to a population and it therefore has to be stated that "the sample only represents itself".

In the ESS system there are several examples of non-probability surveys([2]):

- Quota surveys are used in Germany for the Household Budget Survey.
- In the ESS Short Term Statistics (STS) construction price index, the selection of construction units does not follow a probabilistic design but rather a subjective cut-off approach.
- In the agricultural price index, respondents are often chosen on the basis of assessed importance or convenience, which may also be regarded as a subjective cut-off approach.
- In the Harmonised Index of Consumer Prices (HICP) quota sampling is used for certain products with many characteristics, such as package holidays. (The HICP itself contains a large mosaic of different subprocesses and sampling methods for different product groups.) For such price surveys, inferences rely on the implicit assumption that price changes do not systematically differ between sampled units and other units. The basis for such an assumption can be related to market forces, which normally tend to enforce a more or less uniform price development pattern for outlets selling the same product or similar products (although it is not necessarily always the case).

Surveys based on opt-in panels are becoming increasingly used for commercial purposes but are not yet much used in the ESS. There is no or little theory to back up estimates from such surveys. Ultimately, their use depends on an implicit assumption that the sample achieved is sufficiently representative of the target population.

4.4 Census survey

A *census survey* is defined as a statistical process that involves data collection from all the units in the survey frame**.** Since data are collected from responding units, it is appropriate to think of the process as a kind of survey - a *census survey*. The abbreviated term *census* is commonly used. However, *census* can also have a different meaning, as further discussed below.

As a census can be seen as a special case of a survey, the components of accuracy that are defined for a probability survey are also applicable to a census survey, except that sampling errors are zero by definition([3]).

It is worth emphasising that, although an attempt is made to reach all units in the frame, the coverage can be less than perfect. First, the frame may not cover the target population; second, data may not be obtained from all units.

If the objective of a census survey is to build a register for statistical purposes, to be maintained over a long period, it can be termed a *longitudinal census*.

A census survey may be repeated, but if so, it is usually at fairly long intervals, such as five or ten years.

The classical type of census survey is the *population census* or, frequently, the *population and housing census*. Here an attempt is made to include all the people living in a country and obtain basic data about them.

In many countries, population censuses are partly or wholly based on administrative data. In this case, according to the typology presented here, the statistical process is a census, but <u>not</u> a census survey. It is a government *administrative data process* (see Section 4.5) or a *multisource process* (see Section 4.6), depending on the particular source(s([4]) used.

([2]) For a longer presentation of non-probability techniques, see ILO (2004, 5.27-5.48)

([3]) If subsampling among non-respondents is used, sampling error could still exist.

([4]) The various types of population censuses are described in detail In UNECE (2015)

Economic censuses, usually confined to commercial businesses, are carried out in some countries for example, the USA, India, Japan and Indonesia, but rarely in Europe. They may be census surveys, or administrative data processes, or multisource processes.

Structural business surveys are typically stratified by size, with full coverage of the largest stratum and probability sampling in the other strata. These are classified as sample surveys.

The process of constructing and maintaining a statistical business register (SBR) typically uses administrative and survey data in combination. If statistics, for example counts of businesses by industry and size, are published directly from an SBR, the production process is classified as a multiple source survey not a census.

4.5 Administrative data process

An administrative data process produces statistics based on administrative data([5]) collected by organisations for administrative (regulatory, accounting, commercial or other non-statistical) purposes. The organisations may be governmental institutions or private companies or other non-governmental entities. No sampling is normally involved in this type of process. Data are produced by the internal processes of the organisations. These processes should be described and the quality of the data they produce should be assessed, just as for a survey.

An example is where statistical tabulations are produced from an administrative database maintained by the government agency responsible for higher education. However, it is important to make a distinction between this and the case where questionnaires are sent by the statistical authority to educational institutions asking for information on students, teachers, courses etc. This is considered to be a survey (possibly census) regardless of how, or from what, administrative data sources the responding institutions retrieve the information. The key point here is that the questionnaire and the subsequent data collection, including the definitions of the variables, is designed and performed by the statistical authority.

There are several subcategories within this category. One way to classify them is the following:

- *Public sector accounting data*. Statistics on central and local government finance and pensions are examples where complete information is drawn from public sector databases and assembled into statistical information. The quality issues in these cases could, for example, be coverage issues or inconsistencies in classification or periodization of certain expenditures.

- *Administrative registers.* Registers (administrative data with unique identifiers for population, business, cars, dwellings, electricity accounts, telephone accounts, water service accounts, etc.) are kept by government and other organisations for control, tax and other purposes. Statistics are sometimes based directly on the content of such registers at a certain point in time. Quality issues here may be coverage (including lags in registration) or definition of variables and consistencies in the input reporting to the register.

- *Event-reporting systems*. In these cases, a responsible administrative authority (police, hospitals, customs, etc.) reports an event, including a number of variables characterising the event. Three examples are crimes committed, road accidents and causes of death. Trade in goods with non-EU countries, as reported to Customs authorities when goods pass the EU borders, is another example. Some of these data may be *big data*, as further described in Section 4.8. Common quality issues may be underreported events and the classification of events.

Relevance issues are often important for administrative data. Within accuracy there are usually no random sampling errors whereas coverage and measurement issues could be crucial. The definitions of the various error sources differ to some extent from those in sampling surveys and are further

([5]) *Administrative data* refer to units and data derived from an administrative source. They are collected for the purposes of registration, transaction and record keeping, usually during the delivery of a service by the administrative source. They are not collected primarily for research or statistical purposes. They include *administrative register data* (with unique identifiers) and *administrative transaction data*.

discussed in Chapter 13.

4.6 Multisource process

A multisource process is a statistical process that makes use of data from more than one of the above types of process. (A macro-aggregate compilation process, which is defined as a different type below, is excluded from this definition.) Over time multisource processes are becoming increasingly common and new methodological developments have recently appeared.

Many multisource processes comprise a combination of surveys (probability as well as non-probability) and administrative data processes to arrive at the best estimates of variables. Sometimes the components cover different subpopulations (the *split-population* case) and sometimes they provide different variables for the same population (the *split-variable* case)

For multisource processes, the major quality issues are coverage/completeness of data, precise definitions, measurement errors, how the data are brought together, and coherence/comparability both within and across countries.

The Quality Guidelines for Multisource Statistics version 1.1, October 2019 present final quality guidelines for multisource statistics in which the following categorisation is proposed.

Table 4.1. Basic data configurations for integrating multisources (ESSnet Komuso)

Basic data configuration	Description
1	Complementary microdata sources
2	Overlapping microdata sources
3	Overlapping microdata sources with undercoverage
4	Microdata and macrodata
5	Only macrodata
6	Longitudinal data

De Waal et al (2019) further divide the basic data configuration 1 into the split-variable and the split-population cases. They provide a mathematical model for the latter case where there are several non-overlapping data sets together covering the entire target population and the only source of errors are classification errors.

An example of the split-population case is a statistical process collecting data on waste management. Waste is generated by households, businesses and government institutions. Each of these subpopulations of waste generators needs to be reached by different process designs. The end result comprises overall estimates of total waste generation, by type of waste. Example 4.1 illustrates this situation well.

Example 4.1 Data sources used in creation of waste statistics in France

[In this example all components of a split-population process are surveys. The surveys are presumably of different types, by probability as well as by non-probability]

Data from industry surveys by professional bodies.

Sample surveys of undertakings with at least ten employees.

Survey of large companies in the sector.

Survey of installations for the treatment of non-hazardous waste.

Survey on the production of waste and dredged material.

Survey of commercial waste and waste from the transport sector.

Surveys of the production of household waste.

A final remark here is that processes should, as far as possible, be broken down into *process homogeneous* units rather than be labelled as multisource. Entire subject domains, like culture or crime, consist of many statistical processes, some of which may be surveys, and some based on administrative data. Each of these processes should be reported separately rather than collectively as a multisource process. For example, crime statistics may include data from:

- crimes reported to the police (administrative data);
- sentences issued by courts (administrative data);
- inmates in prisons (administrative data); and
- victimisation surveys (probability survey).

Each of these should be reported separately.

4.7 Macro-aggregate compilation process

A macro-aggregate compilation process could be seen as a special case of a multisource process but it is defined here as a separate type with two distinguishing features: first the inputs are aggregates rather than microdata; and, second, it is typically organised according to an internationally recognised system with a common set of definitions and rules for compilation.

More precisely, a macro-aggregate compilation process is characterised by combining aggregate data from two or more sources in order to compile macro-aggregates in a particular area, such as the national accounts. Macro-aggregates often reflect economic concepts (such as production, consumption, investment, inflation, import and export) usually covering the whole nation.

The compilation of macro-aggregates is typically subject to harmonisation laid down in detailed manuals issued by international organisations like the United Nations, IMF, ILO, OECD and Eurostat itself. Thus, an important quality aspect of an aggregate compilation process is the extent to which the rules and recommendations in such manuals have been followed. Sometimes “A/B/C approaches” are described in the manuals where A solutions are best, B acceptable and C unacceptable. This categorisation can be used in a report.

Whether certain satellite accounts like agricultural accounts, forestry accounts and environmental accounts are best classified as macro-aggregate or as multisource (described below) is left open.

4.8 Special cases

Price indexes

There are several price indexes in the ESS. Due to their different designs they belong to different process types. What they have in common is that the target concept, namely *price change*, is more complex than the simple averages, sums or counts that are the target parameters of other types of statistics. Thus, they share a dimension of accuracy based on economic theory that is additional to the dimension associated with other process types. For example, the well-known notion of substitution bias in price indexes is entirely based on economic theory.

Surveys that mix probability and non-probability elements

For example, a probability design may be used in stage 1 but a non-probability design in stage 2. What matters here is whether the final sample can reasonably be analysed as a probability sample so that the typical error structure of a probability survey is applicable. The classification of such a survey is ultimately left to a decision by the producer.

An example of this is Producer Price Indexes, where, in some countries, there is a probability sample of businesses in stage 1 but a subjective sample of representative products (“typical”, “large value”, “lasts long” etc.) in stage 2.

Big data

The concept of what is conventionally called ***big data*** is new to statistical terminology. References to and use of big data started only in the twenty-first century. Big data are generally characterised by:

- their large volume;
- their diverse structure;
- their real-time or near real-time generation;
- the multidimensionality of the statistical unit of interest; and
- the ways by which the data are produced;

The first three items above are sometimes referred to as the three Vs - volume, variety and velocity.

With reference to sources, UNECE([6]) proposed a classification of big data sources based on how they are generated.

- *Human-sourced* data available mostly from social networks, blogs, internet searches, etc. where data are loosely structured and often ungoverned, for example, Facebook and Twitter.
- *Process-mediated* data available from the IT systems of (private and public) organisations where data is usually structured and stored in relational databases, for example, credit card transactions stored by banks, bank transfers, booking systems, web platforms such as AirBnB, and Uber.
- *Machine-generated* data captured by sensors and other machines used to measure and record events in the physical world, for example traffic sensors and web logs.

The use of big data for official statistics in the ESS is presently at an experimental stage. There are currently no examples of statistical processes that are completely, or even mainly, based on big data.

There are examples where big data are used as a secondary source in a multisource process. One such example is in production of consumer price indexes and the HICP, where so called *scanner data* (data registered at cash desks on prices and quantities of purchased goods in supermarkets and department stores) are increasingly used in combination with traditional price data in European countries. There is a consensus that, in this particular case, big data can lead to significant quality improvement in statistical output.

Other pilot projects include the use of web scraping, text mining and inference techniques for producing statistics on online job vacancies and businesses' characteristics. Additionally, other experimental projects, concern the use of data from smart energy meters, data from the automatic tracking system for ships (Automatic Identification System) and mobile phone data.

Supplementary Document E in Part III provides an extended discussion of big data and their quality aspects. However, in view of, first, the lack of extended experience in the production of statistics using big data, and, second, the diverse big data types, big data is not regarded as a separate process type in this Handbook and there are no explicit reporting guidelines for big data. For the time being, review and reporting of the quality of statistical processes using big data has to be on a case by case basis.

Modelling

In some processes, models are applied to directly collected or administrative data in order to produce estimates of a target variable. This adds an additional layer of complexity, especially when addressing the accuracy of the process, since the validity of the model is also a source of uncertainty. (Model assumption errors are treated in S.13.3.5.

Forecasts

Certain processes in the ESS system such as population projections result in *forecasts*. They are

([6]) Classification of Types of Big Data developed by the Task Team on Big Data, in June 2013, accessed at https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data different from statistical processes from an accuracy perspective, since the true value can only be established in the future. If statistics, as a concept, is seen as a quantitative description of reality, forecasts are not statistics and forecasting processes are not considered to be statistical processes. Hence, they are not covered by this Handbook, although most concepts other than accuracy are applicable.

5 (Part I) Types of reports

5.1 Introductory remarks

There is a wide range of types of reports according to the target audience (users or producers), the subject (indicator, statistical process, domain, institution, country), the level (national, European), the degree of detail (monitoring, structural), the reference period (every cycle, annual, periodic), and whether or not the report is in response to a domain specific regulation.

This chapter describes the various types of reports and indicates the extent to which they are covered by the guidelines presented in Part II of this document.

5.2 Subject of report

The subject of a report can be narrow or broad. It can vary from a specific indicator and the process that produces it, to the entire national statistical system. The range of possible subjects is illustrated by the rows in Table 5.1.

Table 5.1: Subject/level of reporting

Subject	National level	European level
Country	National statistical system	European statistical system
Organisation	National statistical authority	Eurostat
Statistical domain (e.g. health, agriculture)	Many/all statistical processes within a statistical domain	Many/all statistical processes within same statistical domain across all Member States and EFTA countries
Statistical process	Process and its outputs, as developed and conducted by a national statistical authority	Process and outputs across all Member States and EFTA countries
Subgroup within statistical process	Subgroup for which outputs are produced	European aggregates for the subgroup
Specific indicator(s)	Outputs in the form of single numbers or time series of such numbers	European aggregates of single numbers or time series of such numbers

The guidelines in this document are primarily aimed at reporting information about a *statistical* *process,* in other words the row shaded in grey in Table 5.2. However, they also provide guidance for reports with a more restricted subject (*subgroup* or *specific indicator*) or a broader subject (*statistical domain*).

For reports that have the *organisation* or *country* as the subject, the ESS Quality Assurance Framework and/or a general quality management system, such as the European Foundation for Quality Management (EFQM) or the ISO 9001: 2015 Quality Management Systems, provide more appropriate guidance.

5.3 National and European levels

As indicated by the columns in Table 5.1, reports may be at national level, or at European level. Reports produced for European level production processes and statistics are typically but not exclusively based on national level reports.

Two aspects of European level statistics stand out as distinct from national statistics and hence of special importance from the perspective of quality and user reports.

- European level statistics may include aggregations (averages, sums etc.) of national estimates applicable to a European entity (for example, EEA, Euro area, etc.). If so, the report will refer to these aggregations.
- European level statistics may include comparisons and contrasts of national outputs. If so, the report will refer to the comparability of outputs across Member States and EFTA countries.

Thus, there are four possible objectives of a European level report:

- to provide information on the content and quality of statistics aggregated to European level;
- to provide information on the collection, processing, content and quality of statistics collected at European level;
- to provide information on the content and quality of comparisons between national statistics; and
- to provide an overview of the quality of national outputs collectively.

5.4 Producer and user reports

Reports differ according to their intended audiences. As indicated in Section 1.1, the two types of report that have been explicitly identified and catered for by SIMS and its (ESQRS and ESMS) substructures are *producer reports* and *user reports.*

- As previously noted, a *producer report*, would be more precisely called a *producer-oriented report.* It is a report intended for use by an NSA to record quality problems and improvements and by Eurostat to review and summarise quality across NSAs. It is often referred to as a quality report (as reflected in the title of the Handbook). However, the term producer report is preferred in this Handbook because a quality report could equally well be aimed at users.
- A *user report* would be more precisely termed a *user-oriented report*. It is a report intended for users, not a report about users. In the past it was often referred to as a *metadata report* (as reflected in the title of this Handbook). However, this name is not used in the Handbook as it is somewhat ambiguous given that both user reports and producer reports are metadata.

The guidelines apply to each of the two types of report. For the *common* concepts and sub-concepts, the Guidelines make no distinction between producer reports and user reports. In other words, the SIMS definitions and guidelines, and the background and further guidelines in the Handbook, are identical for both types of report. This facilitates using the same descriptions of common concepts and sub-concepts in user report and producer reports, in accordance with the *once for all purposes* reporting approach.

5.5 Degree of detail

The major variation in degree of detail between reports is reflected in this Handbook as the distinction between producer (more detailed) reports and user (less detailed) reports. However, within each of these two groups there are subgroups. For example, producer reports may be addressed to methodologists, risk management experts, senior managers, or Eurostat, all with different needs. In the case of user reports, there are sophisticated users such as central banks and national accounts, and unsophisticated users, such as the general public, with quite different needs. Thus, whilst there is a single set of guidelines, the level of detail appropriate for a particular report may vary and is at the discretion of the report author.

More specifically, a producer report can range:

- from *short*, for example, simply including the standard quality and performance indicators and quality check results;
- to *more detailed*, for example a self-assessment, based on a quality checklist such as the. Development of a Self-Assessment Programme (DESAP);
- to *comprehensive*, covering all aspects of the statistical process and its outputs in full detail.

Likewise, a user report can be more or less detailed according the types of users to whom it is principally targeted.

The *further guidelines* for common concepts and sub-concepts in Part II of this document are aimed at meeting the needs of the most detailed form of report, whether a producer or a user report. Authors intending to produce less detailed reports should take this into account.

5.6 Reporting reference period and frequency

Reports may be prepared for every cycle (repetition, occasion) of a statistical process, or may be prepared less frequently and cover several cycles. Typically, the shorter the reference period of the process and/or the more frequent the report, the less detail is likely to be included.

- For concepts and sub-concepts that are common, or used only in user reports, the guidelines are aimed at the sort of report that will be produced once per year for a sub-annual or annual process, and on each occasion for a less frequent process. The aim is to fully satisfy users' needs.
- For sub-concepts that are covered only in producer reports, the guidelines are aimed at the sort of comprehensive report that will be produced annually or less frequently, or after major changes.

5.7 Domain specific regulations involving quality reporting

As noted in Section 2.4.3, in certain statistical domains, regulations may require the production of quality reports and may specify their structure, contents and periodicity. The structure and/or content of the report required by regulation may well differ from SIMS either because the regulation predates SIMS or because the authors of the regulation did not want to tie the regulation to a standard such as SIMS that could subsequently be updated.

An example of a domain specific regulation follows.

REGULATION (EU) 2018/1091 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 July 2018 on integrated farm statistics

Article 11 Quality

1. Member States shall take the necessary measures to ensure the quality of the transmitted data and metadata.
2. For the purposes of this Regulation, the quality criteria defined in Article 12(1) of Regulation (EC) No 223/2009 shall apply.
3. The Commission (Eurostat) shall assess the quality of the data and metadata transmitted.
4. For that purpose, Member States shall transmit a quality report describing the statistical process to the Commission (Eurostat), for each reference year covered by this Regulation, and in particular:

   (a) metadata describing the methodology used and how technical specifications were achieved by reference to those laid down by this Regulation;

   (b) information on compliance with the minimum requirements for the survey frames used, including in developing and updating them, as laid down in this Regulation;

   The Commission may adopt implementing acts setting out the practical arrangements for, and the contents of, the quality reports. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 17(2) and shall not impose significant additional burdens or costs on the Member States.
5. Member States shall inform the Commission (Eurostat) as soon as possible about any relevant information or change with regard to the implementation of this Regulation that could influence the quality of the data transmitted.
6. At the request of the Commission (Eurostat), Member States shall provide necessary additional clarification to evaluate the quality of the statistical information.

As another example, in the implementing regulation on the quality of data on national and regional accounts, there is the following reference.

> *As the information in the quality reports on national and regional accounts should be based on the European Statistical System standards on quality reporting published by the Commission (Eurostat), the Annex to this Regulation should be drawn up in line with those standards. Information about the ESA 2010 implementation already provided by Member States should be reused by the Commission and should not be requested in the quality reports*.

More generally, if the regulation reporting structure and content coincide with, or are a subset of, SIMS, then the producer or user report can be readily prepared using the SIMS structure and including all the information required by the regulation. If they do not coincide then the report should follow the SIMS structure, but with the addition of (possibly overlapping) content required by the regulation. Three possible solutions from most to least preferable are as follows:

- Map the domain-specific requirements to the appropriate SIMS concepts without changing anything.
- Map the domain-specific requirements to the appropriate SIMS concepts and add domain-specific guidelines.
- Add domain-specific sub-concepts to the existing SIMS concepts.

The quality reporting requirements associated with the Harmonised index of consumer prices (HICP) as described in the HICP regulation Article 9 provide a good example of the last mentioned.

Ideally over time, as regulations are revised, and new regulations are introduced, they will be brought into better alignment with SIMS. Conversely, future revisions of SIMS should take into account the purposes and contents of domain specific regulations.

5.8 Related documentation

A producer or user report is one type of documentation for a statistical process. Many other types of documentation are produced, and national practices differ widely. Some countries produce comprehensive technical reports. Others have standard operating procedures describing statistical methods and procedures in detail, for example, including details of editing rules and estimation formulae. When such documentation exists and is readily accessible, a producer or user report can refer to it and the information it contains need not be repeated in the body of the report. However, when such documentation is not available, information on methods and procedures has to be included in the report itself to provide the context necessary to understand the report.

6 (Part I)

Structure of guidelines and storage of reports

6.1 Structure of guidelines

The guidelines in Part II of this document are organised into chapters by SIMS concept in numerical order. As the concepts are of very unequal complexity, the chapters are of very unequal lengths, ranging from one page for concept *S.19 Comments* to more than 30 pages for concept *S.13 Accuracy and reliability*.

Within each chapter the sections follow a standard format:

- an initial table giving the *SIMS name, definition and guidelines* for the concept and/or its sub-concepts, including related standard quality and performance indicators;
- *background* information about the concept or sub-concept - for complex concepts/sub-concepts only;
- *further guidelines* for the concept and/or sub-concepts - for all but the very simplest concepts/sub-concepts;
- an *example*, or *examples* – those drawn from reports in the ESS Metadata Handler are labelled [ESS-MH].

Within this standard format there are three distinct variants reflecting the complexity of the concept and its sub-concepts.

- For the less complex concepts, the concept and its sub-concepts are covered in a single section.
- For the more complex concepts, the sub-concepts are divided into groups, and guidelines are presented for each group in a separate section.
- For the most complex concepts, in particular S.13 Accuracy, there is a further breakdown by type of statistical process within some sections.

In the tabular displays of concepts and sub-concepts at the beginning of each section:

- those concepts and sub-concepts that apply only to producer reports are marked (P); and
- those concepts and sub-concepts that apply only to user reports are marked (U).

6.2 ESS Metadata Handler

The ESS Metadata Handler (ESS-MH) has been developed by Eurostat to assist NSAs in compiling and storing producer and user reports.

It facilitates creation of reports structured according to SIMS.

It provides definitions and guidelines from the relevant sections of SIMS.

It provides storage, access and viewing facilities for producer and user reports.

It contains many producer and user reports, thus is a great source of examples.

PART II Guidelines

S.01 (Part II) Contact

SIMS	Concept name	Definition	Guidelines
S.01	Contact	Individual or organisational contact points for the data or metadata, including information on how to reach the contact points.	(Information relating to this concept is provided by reporting on its sub-concepts.)
This concept and its sub-concepts are included in both ESMS based (user) reports and in ESQRS based (producer) reports. In ESQRS based reports it is ESQRS Concept 1.			
S.01.1	Contact organisation	The name of the organisation of the contact points for the data or metadata.	Provide the full name (not just code name). of the organisation responsible for the process and outputs (data and metadata) that are the subject of the report.
S.01.2	Contact organisation unit	An addressable subdivision of an organisation.	Provide the full name of the organisational unit responsible. The name can include a unit number.
S.01.3	Contact name	The names of the contact points for the data or metadata.	Provide the first and last names of the contact point(s). If more than one name is provided, the main contact should be indicated. If the author of the report is different from the person(s) responsible for process and its outputs, provide this name also.
S.01.4	Contact person function	The area of technical responsibility of the contact, such as "methodology", "database management" or "dissemination".	Provide the title(s) and area(s) of responsibility of the person(s) indicated as contact(s), for example Senior Research Assistant, Economics Division.

SIMS	Concept name	Definition	Guidelines
S.01.5	Contact mail address	The postal address of the contact points for the data or metadata.	Provide the postal address(es) of the person(s) indicated as contacts.
S.01.6	Contact email address	E-mail address of the contact points for the data or metadata.	Provide the email address(es) of the person(s) indicated as contacts. The address(es) can be (an) individual e-mail address(es) or a mailbox in the organisation to which the person(s) has (have) access.
S.01.7	Contact phone number	The telephone number of the contact points for the data or metadata.	Provide the telephone number(s) of the person(s) indicated as contacts.
S.01.8	Contact fax number	Fax number of the contact points for the data or metadata.	Provide the fax number(s) of the person(s) indicated as contacts.

S.01 EXAMPLE

Example S.01-1: Foreign Affiliate Statistics Annual 2015

Statistical Office of the Slovak Republic [ESS-MH]

[This is a typical example including more or less all that is needed. It would have been useful to specify the function of the contact person. The contact person name and contact details have been suppressed to preserve privacy.]

1.1. Contact organisation: Statistical Office of the Slovak Republic

1.2. Contact organisation unit: Business Statistics Directorate, Department of Business Statistics Methodology and Synthesis

1.3. Contact name: xxxxxxxx

1.4. Contact person function: Department of Business Statistics Methodology and Synthesis

1.5. Contact mail address: Miletičova 3, 824 67 Bratislava, Slovak Republic

1.6. Contact email address: xxxx@statistics.sk

1.7. Contact phone number: +421 2 xxxxxxxx

1.8. Contact fax number: not available

S.02 (Part II) Metadata update

SIMS	Concept name	Definition	Guidelines
S.02	Metadata update (U)	The date on which the metadata element was inserted or modified in the database.	(Information relating to this concept is provided by reporting on its sub-concepts.)
This concept and its sub-concepts are included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).			
S.02.1	Metadata last certified (U)	Date of the latest certification provided by the domain manager to confirm that the metadata posted are still up-to-date, even if the content has not been amended.	Certification can be provided even if the metadata have not been amended since the previous certification. *European Level* Certification for European level metadata.
S.02.2	Metadata last posted (U)	Date of the latest dissemination of the metadata.	The date when the complete set of metadata was last disseminated as a block should be provided (manually, or automatically by the metadata system). *European level* Date refers to the European level metadata.
S.02.3	Metadata last update (U)	Date of last update of the content of the metadata.	The date when any metadata were last updated should be provided (manually, or automatically by the metadata system). *European level* Date refers to the European level metadata.

S.02 EXAMPLE

Example S.02-1 Accidents at work (ESAW, 2008 onwards)

Belgian Federal Agency for Professional Risks [ESS-MH]

[This is a typical example, indicating all that is needed.]

2.1 Metadata last certified: 11/11/2016

2.2 Metadata last posted: 30/06/2016

2.3 Metadata last updated: 11/11/2016

S.03 (Part II) Statistical presentation

SIMS	Concept name	Definition	Guidelines
S.03	Statistical presentation	Description of the disseminated data which can be displayed to users as tables, graphs or maps.	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and its sub-concepts are included in both ESMS based (user) reports and in ESQRS based (producer) reports. In ESQRS based reports it is ESQRS Concept 2

S.03.1-3 Data description, classification system and sector coverage

SIMS	Concept name	Definition	Guidelines
S.03.1	Data description	Main characteristics of the data set, referring to the data and indicators disseminated.	Describe briefly the main characteristics of the data in an easily and quickly understandable manner, referring to the main variables disseminated. More detailed descriptions of the variables are in S.03.4.
S.03.2	Classification system	Arrangement or division of objects into groups based on characteristics which the objects have in common.	List all classifications and breakdowns that are used in the data (with their detailed names) and provide links (if publicly available). Explain deviations, if any, from ESS or international standards. *European level* Provide an overview of national deviations from ESS and/or international standards.
S.03.3	Sector coverage	Main economic or other sectors	List the main economic or other sectors covered by the data and the size classes used, for

SIMS	Concept name	Definition	Guidelines
		covered by the statistics.	example, size classes based on number of employees. *European level* Provide a summary of differences in the main economic or other sectors covered by national data and the size classes used.

S.03.1-3 EXAMPLE

Example S.03.1-3-1 Production in industry 2016

Statistics Denmark [ESS-MH]

[This example includes all that is needed.]

3.1 Data description

The statistic provides a monthly estimate of the activity in the manufacturing industry shown by the level of production and turnover. The results are published on a monthly basis as indexes with a 2010 base year. Furthermore, the turnover index is split by domestic turnover and by export turnover. Both indexes are categorized into four industry sectors and 12 industries and seasonally adjusted data for both indexes are available as well.

3.2 Classification system

Industry activity is categorized according to the Danish industrial classifications, Dansk Branchekode 2007 (DB07), which is the national classification system based on the NACE rev. 2. A complete overview can be found on the DB07 site. Data sent to Eurostat is classified by the NACE rev. 2 industrial classification. The connection between the two classifications can be found at dst.dk, Danish industrial classifications. For aggregated index calculation purposes, the industries are categorized into 47 sub-industries (journal groups) based on the Danish classification and the NACE rev. 2. These sub-industries are not published.

3.3 Sector coverage

Manufacturing (C), mining and quarrying (B), and utility services (D+E). All letters and numbers in brackets relate to the Danish industrial classification (DB07). Industry C manufacturing are split into 12 groups (industry code in brackets):

- CA Manufacture of food products, beverages and tobacco (10 – 12)
- CB Textiles and leather products (13 – 15)
- CC Wood and paper products and printing (16 – 18)
- CDE Manufacture of chemicals and oil refineries etc. (19 – 20)
- CF Pharmaceuticals (21)
- CG Manufacture of plastic, glass and concrete (22 – 23)
- CH Basic metals and fabricated metal products (24 – 25)
- CI Manufacture of electronic components (26)
- CJ Electrical equipment (27)
- CK Manufacture of machinery (28)
- CL Transport equipment (29 – 30)
- CM Manufacture of furniture and other manufacturing (31 – 33)

The industries are also categorized into four sectors by the utilization of the sold goods and services: - capital goods - intermediate goods - durable consumer goods - non-durable consumer goods. The split of industries into these four sectors are listed in an annex.

S.03.4-6 Statistical concepts, definitions, units and populations

SIMS	Concept name	Definition	Guidelines
S.03.4	Statistical concepts and definitions	Statistical characteristics of statistical observations, variables.	Define and describe briefly the main statistical variables that have been observed or derived. Indicate their types. Indicate discrepancies, if any, from the ESS or international standards. Note that any difference between these variables and the variables desired by users is a relevance issue and is discussed in S.12. *European level* Summarise the national discrepancies from the ESS and/or international standards.
S.03.5	Statistical unit	Entity for which information is sought and for which statistics are ultimately compiled.	Define the type of statistical unit about which data are collected, e.g. enterprise, kind of activity unit, local unit, private household, dwelling, person, import transaction. If there is more than one type of unit, define each type. *European level* Summarise the differences in units used at national level.
S.03.6	Statistical population	The total membership or population or "universe" of a defined class of people, objects or events.	Define the *target population* of statistical units for which information is sought. The *survey (frame) population* of statistical units (which is the approximation to the target population used in practice) is described in S.18.1. The difference between *target population* and the *survey population* is a coverage issue and is discussed in S13.3 *European level* Summarise the differences in target populations used at national level.

S.03.4-6 EXAMPLES

Example S.03.4-6-1 Documentation of statistics for Foreign Owned Enterprises 2016. Statistics Denmark Memo

[In this example the statistical population referenced in Section 3.6 below is the target population.]

2.4 Statistical concepts and definitions

Number of employees: Persons on the payroll in full-time equivalent units.

Enterprise: Usually corresponding to the legal unit, e.g. limited-liability corporations, sole traders, partnerships, etc. In a few cases several legal units which are run as one entity are gathered into one enterprise.

Turnover: Turnover represents the net sales. Included are capitalised work performed by the firm for own purposes and all charges (transport, packaging, etc.) passed on to the customer. Excluded are reduction in prices, rebates, discounts, VAT and excise duties. Income classified as other operating income, financial income and extraordinary income in company accounts is also excluded from turnover.

The ultimate owner: The statistics are defining a company's ownership attached to the ultimate owner who has control over the company, ie, have the ability to determine a company's consult the general policy, if necessary by select a board. As a rule interpreted the controlling unit as the ultimate owner, directly or indirectly, more than 50 per cent. of equity or shareholders' voting rights. The ultimate owner must be understood in relation to the direct owner, since a company can immediately be owned (directly) from a country, even if it ultimately (ultimately) is the owner of another country.

2.5 Statistical unit

The unit in the statistics is enterprise. Usually corresponding to the legal unit, e.g. limited-liability corporations, sole traders, partnerships, etc. In a few cases several legal units which are run as one entity are gathered into one enterprise.

2.6 Statistical population

Danish and foreign-owned enterprises in Denmark.

Example S.03.4-6-2 EU Farm Structure Survey, Statistics Poland

[This example covers only units and population. Under Statistical Population both the *target population*, which is labelled 1), and the *survey population*, which is labelled 3), are presented. Explicit use of these terms would have made the exposition clearer. Strictly speaking survey population should be described in S.18.1.

Some of the detail provided in the original example has been omitted from the text below.]

Statistical unit

The national definition of the agricultural holding;

> *Agricultural holding is understood as a single unit, both technically and economically, which has a single management and which conducts agricultural activity*.

Agricultural activity (primary or secondary), according to the NACE. rev.2, includes

> *activities listed in section A, division 01, groups: - 01.1 - growing of non-perennial crops; 01.2 - growing of perennial crops; 01.3 - plant propagation; 01.4 - animal production (subgroup 01.49 is excluded, with the exception of the raising and breeding of ostriches, emu and rabbits as well as other fur animals); 01.5 - mixed farming; 01.6 - class 01.61 - support activities for crop production (maintaining good agricultural condition following environment protection standards).*

Statistical population

1) The number of holdings forming the entire universe of agricultural holdings in the country:

- 1 548 116 (Only the agricultural holdings meeting the below mentioned thresholds are considered as agricultural holdings).

2) The national survey coverage: the thresholds applied in the national survey and the geographical coverage:

- a natural person's agricultural holding…is covered in the survey if it has an agricultural land of 1 ha or more or if it has an agricultural land less than 1 ha (even without agricultural land) conducting special branches of agricultural activity or complying with the following physical thresholds….:or runs organic production….)
- a legal person's agricultural holding…in case agricultural activity is not primary - is covered in the survey if it has an agricultural land of 1 ha or more .or if it runs livestock production.

3) The number of holdings in the national survey coverage:

- According to the above mentioned definition and thresholds, the national survey covered 1 410 704 agricultural holdings.

S.03.7-9 Reference area, time coverage and base period

SIMS	Concept name	Definition	Guidelines
S.03.7	Reference area	The country or geographic area to which the measured statistical phenomenon relates.	Describe the country, the regions, the districts, or the other geographical aggregates, to which the data refer. Identify any specific exclusions in the data disseminated. If coverage includes overseas territories this should be stated, and they should be specified. European level Describe the geographical area covered by the data disseminated, e.g., EU Members states, EU regions, USA, Japan, or aggregates such as EU, EEA).
S.03.8	Time coverage	The length of time for which data are available.	State the time period(s) covered by the data, e.g. first quarter 2018, or quarters 2015-2018, or year 2018, or years 1985-2018. Note that any issues concerning comparability over time are discussed in S.15.
S.03.9	Base period	The period of time used as the base of an index number, or to which a constant series refers.	Note that this concept applies only to certain types of outputs, such as indexes, for which a base period is defined and used. State the base period, for example, year 2000. Indicate base period update time frame and date of next update.

S.03.7-9 EXAMPLE

Example S.03-7-9 Production in industry 2016, Statistics Denmark [ESS-MH]

[This example indicates all that is needed.]

3.7. Reference area

Denmark not including the Faroe Islands and Greenland. In some cases, turnover will include sales from goods not manufactured in Denmark; for instance, if goods are produced as part of contract work for other enterprises for a Danish manufacturer.

3.8. Coverage - Time

The statistics covers the period from 2000 onwards. Older time series are described under Comparability over time.

3.9. Base period

Point of reference for the indexes is the average production and turnover of 2010, which is shown as 100 in the indexes. The base year is updated every five years. The next update is scheduled to take place spring 2018 for the base year 2015.

S.04 (Part II) Unit of measure

SIMS	Concept name	Definition	Guidelines
S.04	Unit of measure (U)	The unit in which the data values are measured.	The data usually involves several units of measure depending upon the variables. Examples are: Euro, national currency, number of persons, and rate per 100,000 inhabitants. The magnitude (e.g., thousand, million) of numerical units should be included.

This concept is included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).

S.04 EXAMPLE

Example S.04-1 HICP 2012

Lithuanian Department of Statistics [ESS-MH]

[In the interests of clarity, the example has been slightly modified from the original.]

Following units are used:

- Index (actually unit-less), i.e. it is the ratio of the price of the basket in a given year to the price in the reference year multiplied by 100. However, the HICP can be thought of as the amount the average consumer would have to spend in a given year to buy the same basic goods and services that one would have to pay 100 monetary units for in the reference year);
- Percentage change relative to the same period in the previous year (rate);
- Percentage change relative to the previous period (rate);
- Proportion of the total expenditure (weight) expressed as per thousand.

S.05 (PART II) Reference period

SIMS	Concept name	Definition	Guidelines
S.05	Reference period (U)	The period of time or point in time to which the measured observation is intended to refer.	The value of a variable refers to a specific time period (for example, the last week of a month, a month, a fiscal year, a calendar year, or several calendar years), or to a point in time (for example, a specific day, or the last day of a month). The variables in a dataset may refer to more than one reference period. All reference periods should be stated. Note that the difference, if any, between the target reference period(s) and the actual reference period(s) is an accuracy issue and should be discussed in S.13.3. Note that if the survey population does not include all the units in the target population for the specified reference period, this is a coverage issue and should be discussed in S.13.3. *European level* Summarise differences in reference period across countries.
This concept is included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).			

S.05 EXAMPLE

Example S.05-1 Owner-Occupied Housing price index

Lithuanian Department of Statistics [ESS-MH]

[the following example is very short and not very informative. It would be improved by explaining in what sense the indices represented the quarter.]

The quarterly indices compiled represent the whole calendar quarter.

S.06 (PART II) Institutional mandate

SIMS	Concept name	Definition	Guidelines
S.06	Institutional mandate (U)	Law, set of rules or other formal set of instructions assigning responsibility as well as the authority to an organisation for the collection, processing, and dissemination of statistics.	(Information relating to this concept is provided by reporting on its sub-concepts.)
This concept and its sub-concepts are included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).			
S.06.1	Legal acts and other agreements (U)	Legal acts or other formal or informal agreements that assign responsibility as well as the authority to an agency for the collection, processing, and dissemination of statistics.	State the national legal acts and/or other reporting agreements, including EU legal acts, the implementation of EU directives. *European level* State the legal base or other agreement, for example, the EU legal act, or ESS Five-Year-Program, that underpins the reporting obligations on countries.
S.06.2	Data sharing (U)	Arrangements or procedures for data sharing and coordination between data producing agencies.	Describe the arrangements, procedures or agreements to facilitate data sharing and exchange between data producing agencies within the national statistical system. *European level* Describe the arrangements, procedures or agreements to facilitate data sharing and exchange between international data producing agencies, for example, a Eurostat data collection or production that is in common with the OECD or the UN.

S.06 EXAMPLES

Example S.06-1 Owner-Occupied Housing (OOH) Price Index

Lithuanian Department of Statistics [ESS-MH]

[The following illustrates a comprehensive description of legal acts and other agreements. Access to administrative data is recorded as “None”, because the Price Department has no direct access, but receives the data by query from Statistics Lithuania’s IT Department. Data sharing is recorded as “None” because there is no mandate for sharing prices on dwellings, only indexes are shared.]

6.1. Institutional mandate - legal acts and other agreements

Legal acts and other agreements - EU level

Regulation (EU) 2016/792 of the European Parliament and of the Council of 11 May 2016 on harmonised indices of consumer prices and the house price index, and repealing Council Regulation (EC) No 2494/95.

Commission Regulation (EU) No 93/2013 of 1 February 2013 laying down detailed rules for the implementation of Council Regulation (EC) No 2494/95 concerning harmonized indices of consumer prices, as regards establishing owner-occupied housing price indices.

Legal acts and other agreements - national level

There is no specific Lithuanian legislation for the production of the OOH price index.

The Law on Statistics of the Republic of Lithuania is the main national legal act regarding official statistics. Production of the OOH price index as well as other official statistics is included in the annual Official Statistics Work Programme. The OOH price index is calculated based on the methodology of the Lithuanian OOH price index production approved on 29 December of 2014 by Order No D-403 of the Director General of Statistics Lithuania.

Contract signed between Statistics Lithuania and the owner of administrative data on dwelling transactions – the state enterprise Centre of Registers.

Access to administrative data:

None.

6.2. Institutional mandate - data sharing

None.

Example S.06-2 Accidents at Work, 2008 onwards)

Belgian Federal Agency for Professional Risks [ESS-MH]

[This example indicates a minimal response. It is not clear why data sharing is “not applicable”.]

6.1. Institutional mandate - legal acts and other agreements

For accidents in the private sector the Belgian law of 10 April 1971 defines the relevant rules and procedures. Accidents in the public sector are regulated by the law of 4 July 1967. The obligation of data transmission is also included in these laws.

6.2. Institutional mandate - data sharing

Not applicable.

S.07 (Part II) Confidentiality

SIMS	Concept name	Definition	Guidelines
S.07	Confidentiality	A property of data indicating the extent to which their unauthorised disclosure could be prejudicial or harmful to the interest of the source or other relevant parties.	(Information relating to this concept is provided by reporting on its sub-concepts.)
This concept and its sub-concepts are included in both ESMS based (user) reports and in ESQRS based (producer) reports. In ESQRS based reports it is ESQRS Concept 11.			

S.07 Background

At the European level the legal framework and the conceptual framework for confidentiality is the Regulation 223/2009 on European statistics. The following paragraphs reproduce extracts from the Regulation and from the ES Code of Practice, and provide additional notes to ensure the context for reporting on confidentiality is well understood.

It is useful to distinguish between the protection of the statistical production process (secure IT environment, setting of access rights etc.) and protection of the statistical output, as these require different types of actions and usually also have different actors.

Regulation 223/2009 on European statistics

(From the Preamble)

(23) The confidential information which the national and Community statistical authorities collect for the production of European statistics should be protected, in order to gain and maintain the confidence of the parties responsible for providing that information. The confidentiality of data should satisfy the same principles in all the Member States.

(24) For that purpose, it is necessary to establish common principles and guidelines ensuring the confidentiality of data used for the production of European statistics and the access to those confidential data with due account for technical developments and the requirements of users in a democratic society.

(26) The research community should enjoy wider access to confidential data used for the development, production and dissemination of European statistics, for analysis in the interest

of scientific progress in Europe. Access to confidential data by researchers for scientific purposes should therefore be improved without compromising the high level of protection that confidential statistical data require.

(27) The use of confidential data for purposes that are not exclusively statistical, such as administrative, legal or tax purposes, or for the verification against the statistical units should be strictly prohibited.

(From Article 3 Definitions)

Confidential data means data that allow individual statistical units to be identified, either directly or indirectly, thereby disclosing individual information. To determine whether a statistical unit is identifiable, account shall be taken of all relevant means that might reasonably be used by a third party to identify the statistical unit.

Use for statistical purposes means the exclusive use for the development and production of statistical results and analyses (Article 3-8).

Direct identification means the identification of a statistical unit from its name or address, or from a publicly accessible identification number; indirect identification means the identification of a statistical unit by any other means than by way of direct identification (Articles 3-9,3-10).

(From Chapter V Statistical Confidentiality)

Article 20 Protection of confidential data

1. The following rules and measures shall apply to ensure that confidential data are exclusively used for statistical purposes and to prevent their unlawful disclosure.
2. Confidential data obtained exclusively for the production of European statistics shall be used by the NSIs and other national authorities and by the Commission (Eurostat) exclusively for statistical purposes unless the statistical unit has unambiguously given its consent to the use for any other purposes.
3. Statistical results which may make it possible to identify a statistical unit may be disseminated by the NSIs and other national authorities and the Commission (Eurostat) in the following exceptional cases:
   (a) where specific conditions and modalities are determined by an act of the European Parliament and of the Council (…) and the statistical results are amended in such a way that their dissemination does not prejudice statistical confidentiality whenever the statistical unit has so requested; or
   (b) where the statistical unit has unambiguously agreed to the disclosure of data.
4. Within their respective spheres of competence, the NSIs and other national authorities and the Commission (Eurostat) shall take all necessary regulatory, administrative, technical and organisational measures to ensure the physical and logical protection of confidential data (statistical disclosure control).

Article 21 Transmission of confidential data

1. Transmission of confidential data from an ESS authority…that collected the data to another ESS authority may take place provided that this transmission is necessary for the efficient development, production and dissemination of European statistics or for increasing the quality of European statistics.
2. Transmission of confidential data between an ESS authority that collected the data and an ESCB member may take place provided that this transmission is necessary for the efficient development, production and dissemination of European statistics or for increasing the quality of European statistics, within the respective spheres of competence of the ESS and the ESCB, and that this necessity has been justified.

4. National rules on statistical confidentiality shall not be invoked to prevent the transmission of confidential data under paragraphs 1 and 2 (…).
5. Confidential data transmitted in accordance with this Article shall be used exclusively for statistical purposes and only accessible to staff working in statistical activities within their

specific domain of work.

6. The provisions on statistical confidentiality provided for in this Regulation shall apply to all confidential data transmitted within the ESS and between the ESS and the ESCB.

Article 22 Protection of confidential data in the Commission (Eurostat)

1. Confidential data shall be accessible, subject to the exceptions laid down in paragraph 2, only to officials of the Commission (Eurostat) within their specific domain of work.

2. The Commission (Eurostat) may in exceptional cases grant access to confidential data to its other staff and to other natural persons working for the Commission (Eurostat) under contract within their specific domain of work

3. Persons having access to confidential data shall use these data exclusively for statistical purposes. They shall be subject to this restriction even after cessation of their functions.. .

Article 23 Access to confidential data for scientific purposes

Access to confidential data which only allow for indirect identification of the statistical units may be granted to researchers carrying out statistical analyses for scientific purposes by the Commission (Eurostat) or by the NSIs or other national authorities, within their respective spheres of competence. If the data have been transmitted to the Commission (Eurostat) the approval of the NSI or other national authority which provided the data is required.

Commission Regulation No 557/2013 provides implementation information.

European Statistics Code of Practice (2017) Principle 5: Statistical Confidentiality and Data Protection.

The privacy of data providers, the confidentiality of the information they provide, its use only for statistical purposes and the security of the data are absolutely guaranteed.

Indicator 5.1: Statistical confidentiality is guaranteed in law.

Indicator 5.2: Staff sign legal confidentiality commitments on appointment.

Indicator 5.3: Penalties are prescribed for any wilful breaches of statistical confidentiality.

Indicator 5.4: Guidelines and instructions are provided to staff on the protection of statistical confidentiality throughout the statistical processes. The confidentiality policy is made known to the public.

Indicator 5.5: The necessary regulatory, administrative, technical and organisational measures are in place to protect the security and integrity of statistical data and their transmission, in accordance with best practices, international standards, as well as European and national legislation.

Indicator 5.6: Strict protocols apply to external users accessing statistical microdata for research purposes.

Additional Notes

NSAs are responsible for the statistical confidentiality of their data. Although Eurostat provides guidelines, national methods are not fully harmonised. There are different national rules and traditions.

Confidential data are made available to Eurostat for statistical purposes, especially for the calculation of European aggregates, in most statistical domains. Ideally, confidential cells are tagged using a coding system such as SDMX flags to indicate the reasons why the data are confidential. This helps to optimise subsequent processing of the data and to avoid further unnecessary suppression. The use of confidentiality flags for reasons other than to ensure statistical confidentiality (e.g. to indicate poor quality or no national dissemination) should be avoided.

Microdata

As noted above, confidential microdata can be made available for research purposes under strict conditions. More information on microdata access is provided in Eurostat's Access to Microdata Web Page.

Microdata may also be fully protected against loss of confidentiality and made generally available as public use files, see Eurostat's Public Micro Data. Whilst this usually involves considerable loss of information, the resulting data may be useful for researchers not having access to more complete microdata sets, and for school and university students.

Aggregate Data

Publication of statistical data is more common at aggregate level rather than micro level. Combining data for the units in each cell reduces, but does not necessarily eliminate, the risk that some cells are confidential due to small numbers of units and/or dominant units, especially in business statistics. Protection of aggregate data involves:

- specification of what constitutes a confidential cell;
- determination of confidential cells (primary confidentiality);
- determination of cells or tables subject to residual disclosure (meaning that confidential values can be revealed by tables or bringing tables together – derivation or secondary confidentiality);
- redesign of tables, random perturbation of values in confidential cells, suppression of data for confidential cells, etc.

S.07.1 Confidentiality - policy

SIMS	Concept name	Definition	Guidelines
S.07.1	Confidentiality – policy	Legislative measures or other formal procedures which prevent unauthorised disclosure of data that identify a person or economic entity either directly or indirectly.	Describe all European or national legislation, or other formal requirements, that relate to confidentiality. Describe relevant policy (if any). Note that the existence of legislation and/or policy provides some assurance that methods necessary to assure confidentiality have been applied to the data. *European level* Summarise the commonalties and differences in national approaches to confidentiality policy.

S.07.1 FURTHER GUIDELINES

Note that confidentiality policy is likely to be for the NSA as a whole, not just for the statistical process that is the subject of the report. In describing policy, cover three aspects:

- law - state whether confidentiality is required by law, and, if so, whether survey staff are required to sign legally binding confidentiality commitments;
- general guidelines and coordination – state who can have access to confidential data and under what circumstances, including staff obligations and user access to microdata for research purposes; and
- security - state what security policies, if any, have been introduced with confidentiality in mind.

Note that the corresponding confidentiality provisions are reported under S.07.2, and the access procedures under S.10.4.

Check whether, as a matter of policy:

- direct identifiers are removed from data as early as possible in the statistical production process and replaced by quasi-identifiers known and used only for statistical purposes;
- access to data with direct identifiers is limited and justified.

It is not advisable to include in the user reports details about the methods of protection of confidentiality applied, for risk of possible reversal of this treatment and consequent disclosure of confidential information.

S.07.1 EXAMPLES

Example S.07.1-1 Confidentiality Statement, Statistics Lithuania

[This example illustrates a minimal response, but points to a more comprehensive description available through the Internet.]

In the process of statistical data collection, processing and analysis and dissemination of statistical information, Statistics Lithuania fully guarantees the confidentiality of the data submitted by respondents (households, enterprises, institutions, organisations and other statistical units), as defined in the Confidentiality Policy Guidelines of Statistics Lithuania (https://www.stat.gov.lt/en/konfidencialumo-uztikrinimas).

Example S.07.1-2 Census 2011 Round

Office for National Statistics, United Kingdom [ESS-MH]

[This example illustrates a very comprehensive response.]

The National Statistician and the Registrars General for Scotland and Northern Ireland published a joint agreement to adopt a common statistical disclosure control policy as part of the move towards seeking harmonised statistical outputs from the 2011 Census across the UK. The policy position is based on the principle for protecting confidentiality set out in the National Statistics Code of Practice.

The statutory requirement to provide personal census information is prescribed by the provisions of the Census Act 1920 and the Census (Northern Ireland) Act 1969 and in the associated Orders and Regulations made under these Acts. In return, the Acts – strengthened by the confidentiality provisions of the Statistics and Registration Service Act 2007 – also impose strict requirements on Office for National Statistics (ONS), the National Records of Scotland (NRS) and the Northern Ireland Statistics and Research Agency (NISRA) to protect the confidentiality of any such information collected.

The information collected in the 2011 Census is used solely for the production of statistics and statistical research. Usage complies fully with the Census Acts, the Statistics and Registration Service Act and the requirements of data protection and freedom of information legislation. There are legal penalties for the unlawful disclosure of personal information collected in the census.

In the longer term, census records have an archival value and completed census questionnaires become open for public inspection after 100 years when they are used for genealogical and historical research. In Northern Ireland, Census returns are closed indefinitely.

Privacy Impact Assessment

Privacy Impact Assessments (PIA) help to identify privacy risks, foresee problems and bring forward solutions. It is a process for evaluating a proposal to:

- identify its potential effects upon individual privacy and data protection compliance

- examine how any detrimental effects might be overcome, and
- ensure that new projects comply with the data protection principles.

The PIA assessment for England and Wales was published in November 2009.

The assessment for Scotland was published in January 2011.

The Assessment for Northern Ireland was published in May 2010.

S.07.2 Confidentiality – data treatment

SIMS	Concept name	Definition	Guidelines
S.07.2	Confidentiality - data treatment	Rules applied for treating the datasets to ensure statistical confidentiality and prevent unauthorised disclosure.	For aggregate outputs: • Provide the rules that define a *confidential cell*; • Describe the procedures for detecting confidential cells (primary confidentiality) and checking for residual disclosure (derivation or secondary confidentiality); • Describe the procedures for reducing the risk of disclosure by treating confidential cells, for example by perturbation, controlled rounding, cell suppression, or cell aggregation. For micro-level outputs: • Describe the procedures that are used in protecting confidentiality. *European level* Summarise the commonalties and differences in national approaches.

S.07.2 FURTHER GUIDELINES

If external users may access microdata for research purposes, describe the confidentiality provisions that are applied.

Describe the procedures for ensuring data security during data collection, processing, analysis and dissemination, thereby preventing unauthorised access.

S.07.2 EXAMPLES

Example 07.2-1 Farm Structure Survey, Department for Environment Food and Rural Affairs, UK. [ESS-MH]

[This example illustrates an informative response, but not covering all details.]

11.2. Confidentiality - data treatment

Results from all of our surveys are disseminated according to legislation and the United Kingdom's Code of Practice for Official Statistics. In any tabular publications, all cells where there are less than five contributors are to be suppressed (usually represented by #), although where there are zero contributors this is allowed. If a table contains both holding counts and a variable specific estimate (e.g. wheat area or number of pigs) both values must be suppressed. Further where tables have subtotals there is a need to suppress an additional record within the same group in the table to prevent users from deriving the suppressed data through simple differencing.

An additional level of protection is applied if the tables are for spatial scales of NUTS4 or finer. This additional level involves calculating the proportion of the cell total contributed by the highest contributing farm. Where this value exceeds 85%, the cell value is suppressed to protect the identity of this dominant contributor.

Example 07.2-2 Census 2011 Round

Office for National Statistics, United Kingdom [ESS-MH]

[This example illustrates a very comprehensive response. The first two paragraphs are illustrative of policy, the remainder of procedures.]

England and Wales

Only those people under the management and/or control of the UK Statistics Authority including those agents acting, or providing services, on its behalf for the purpose of the census, and researchers approved under the provisions of the Statistics and Registration Service Act 2007, have access to personal census information.

All members of the census organisations and outside agents providing services to the UK Statistics Authority were required to sign undertakings to ensure their awareness of their statutory confidentiality obligations. Any breaches of the law rendered them liable to prosecution.

In producing the standard statistical outputs, a number of procedures were implemented to prevent the release of information that identified characteristics about an individual person or household.

- modifying some of the data before the statistics were released by using a record swapping methodology
- restricting the number of output categories into which a variable may be classified, such as aggregated age groups
- where the number of people or households in an area fell below a minimum threshold, the statistical output – except for basic headcounts – were amalgamated with that for a sufficiently large enough neighbouring area.

In Northern Ireland and Scotland, similar access control and protection measures to those described for England & Wales were taken by NISRA and NRS within the framework of the relevant census legislation applying in each country: the Census (Northern Ireland) Act 1969 and the Census Act 1920 respectively.

Day-to-day and operational aspects were considered by all three statistical agencies (ONS, NISRA and NRS) as part of the Independent Information Assurance Review.

Information on data security measures employed in England and Wales can be found here:

Copies of the privacy impact assessment (PIA) can be downloaded from this page.

Information on data security measures employed in Scotland can be found here.

Independent Information Assurance Review

The Independent Information Assurance Review (IIAR) was conducted to provide an independent review of the protection to be applied to personal information gathered as part of the 2011 Census. Copies of the IIAR can be downloaded from this page.

The final report was published in June 2012.

S.08 (Part II) Release policy

SIMS	Concept name	Definition	Guidelines
S.08	Release policy (U)	Rules for disseminating statistical data to all interested parties.	(Information relating to this concept is provided by reporting on its sub-concepts.)
This concept and its sub-concepts are included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).			
S.08.1	Release calendar (U)	The schedule of statistical release dates.	State whether there is a release calendar for the statistical outputs from the process being reported, and if so, whether this calendar is publicly accessible.
S.08.2	Release calendar access (U)	Access to the release calendar information.	Describe how the release calendar can be accessed and, if possible, give a link or reference.
S.08.3	User access (U)	The policy for release of the data to users, the scope of dissemination, how users are informed that the data are being released, and whether the policy determines the dissemination of statistical data to all users.	Describe the general data release policy of the organisation. Describe the release policy applied to the outputs of the process being reported, highlighting any deviations from the general policy. Note that the effect of not having a release calendar, or whether releases have been in accordance with a release calendar, is reported in S.14.2. *European level* Summarise country adherence to the impartiality protocol linked to Principle 6 of the European Statistics Code of Practice, which requires those responsible for the statistical domain to make public any and all kinds of pre-release.)

S.08 EXAMPLES

Example S.08-1 Owner-Occupied Housing Price Index

Lithuanian Department of Statistics [ESS-MH]

[This example illustrates a succinct, complete response.]

8.1. Release calendar

Statistical information is published in accordance with an approved release calendar.

8.2. Release calendar access

The calendar is placed on the Official Statistics Portal of Statistics Lithuania.

8.3. Release policy - user access

The data is disseminated to all users through the Official Statistics Portal. The data are released simultaneously to all interested parties by issuing a news release on changes in housing prices and updating the information on the Database of Indicators on the Official Statistics Portal. At the same time the news release is also e-mailed to the media. The news release is issued in Lithuanian only.

The Database of Indicators is available in Lithuanian and English.

Example S.08-2 Census 2011 Round

Office for National Statistics, United Kingdom [ESS-MH]

[This example does not include anything on release calendar access, which should be reported here even if is subsequently referenced in S.14.2]

8.1. Release calendar

Data are made available 27 months after the end of the reference period (March 2014)

8.3. Release policy - user access

All three statistical agencies (Office for national Statistics, Northern Island Statistics and Research Agency and National Records of Scotland) conducted independent user consultation exercises to establish the information requirements from the 2011 Census. These requirements formed the basis of the standard outputs that would be produced by each of the UK Census Offices, who worked collaboratively to harmonise the outputs wherever possible. The three agencies worked on the principle that all standard outputs should be free to users at the point of delivery. Census data is available via the websites of each of the agencies.

Access to published data and metadata is free under the Open Government Licence.

Access to some microdata products is restricted to accredited researchers.

Non-standard tables can be commissioned from the respective statistical agencies for a fee that covers the cost of the additional processing necessary to generate the data. Once a commissioned table has been produced it is published on the relevant agency’s website.

S.09 (Part II) Frequency of dissemination

SIMS	Concept name	Definition	Guidelines
S.09	Frequency of dissemination (U)	The time interval at which the statistics are disseminated over a given time period.	State the frequency with which the data are disseminated, e.g. monthly, quarterly, yearly. The frequency can also be expressed by using a code from the harmonised ESS code list so long as this is considered to be easily understandable by users.

This concept is included in user (ESMS based) reports but not in producer (ESQRS based) reports. This is indicated by (U).

S.09 EXAMPLE

Example S.09-1 Owner-Occupied Housing Price Index

Lithuanian Department of Statistics [ESS-MH]

[This example illustrates a sufficient response.]

Quarterly for the Owner Occupied Household Price Index, and annual for the weights.

S.10 (Part II) Accessibility and clarity

SIMS	Concept name	Definition	Guidelines
S.10	Accessibility and clarity	The conditions and modalities by which users can access, use and interpret data.	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and its sub-concepts are included in ESQRS based (producer) reports' where it is ESQRS Concept 9. In ESMS based (user) reports the concept and 7 sub-concepts are included. The remaining 3 sub-concepts that are included only in producer reports are indicated by (P).

S.10 Background

Dissemination format

In the previous version of SIMS, "dissemination format" was included in the concept name. It refers to the media and the various mechanisms by which statistical output and associated metadata are disseminated to users. It includes a description of the various formats available, and where and how to get the information, for example via news release, printed publication, electronic publication and on-line database. The dissemination formats have a significant impact on accessibility and clarity.

Note that, in this context, "format" does not refer to electronic file formats such as DOC or XLS.

Corresponding Code of Practice principles

In the European Statistics Code of Practice (ES CoP), the notions of *accessibility* and *clarity* are combined in a single principle entitled *Accessibility and clarity* and stating that "European statistics should be presented in a clear and understandable form, disseminated in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance".

Accessibility

Accessibility is an attribute of statistical output describing the set of conditions and modalities by which users can obtain data and accompanying metadata.

Accessibility is reported by describing each of the various means of dissemination and how effective it is regarding ease of access to the data. It includes consideration of the cost of access and limitations set by confidentiality provisions.

Clarity

Clarity is sometimes referred to as *interpretability*. It is an attribute of statistical output describing the extent to which the data are accompanied by readily comprehensible metadata, including information on data quality, and the extent to which additional assistance is made available to users to help them understand the data.

Clarity is reported by describing the metadata that accompany the data, the relevance of these metadata and the ease with which they can be understood.

Classes of users

Accessibility and clarity include accommodating the needs of the various classes of users, including a few very important users, such as government departments, who need to be individually managed. A classification of users is discussed in S.12.1 User Needs. Here (in S.10) it is sufficient to classify users according to the frequency with which they access data and the depth of their interests into two classes:

- *occasional users* who typically prefer simple and clear presentation of data and accompanying metadata data so that they are easy to find and interpret; and
- *professional users* who typically prefer a database approach to dissemination so that they can select and download those data that are of interest to them for further data manipulation and analysis.

Assessment methods

User feedback is the best way to assess the accessibility and clarity. Questions on user experiences regarding ease of access to the data and their interpretation should be included when user satisfaction surveys are designed. The results of such surveys and any other user feedback should be reported.

Counts and movements in the numbers of subscribers to publications (paper or electronic), sales of publications and user consultation hits on websites are also useful indicators.

S.10.1-5 Dissemination mechanisms/formats

SIMS	Concept name	Definition	Guidelines
S.10.1	News release	Regular or ad-hoc press releases linked to the data.	List any regular or ad-hoc press releases linked to the data over the past year.
S.10.2	Publication	Regular or ad-hoc publications in which the data are made available to the public.	List the titles of any publications, including publisher, year, and links to on-line documents (if available). Provide number of subscriptions/purchases of each of the key paper reports.
S.10.3	On-line database	Information about on-line databases in which the disseminated data can be accessed.	Provide the domain name. Provide link to the on-line database (if any) and number of accesses in a recent period.
S.10.3.1	AC1. Data tables – consultations (P)	Number of consultations of data tables within a statistical domain for a	For producer reports only Provide values of Indicator AC1 by month/quarter/year. The indicator AC1 is defined in Supplementary Document C.

SIMS	Concept name	Definition	Guidelines
		given time period displayed in a graph.	
S.10.4	Micro-data access	Information on whether micro-data are also disseminated.	State whether the data are accessible in micro-data form, e.g. for researchers. If so, cross reference the micro-data confidentiality rules in S.7.
S.10.5	Other	References to the most important other data dissemination done.	Describe any other important dissemination mechanisms, for example policy papers, within outputs produced by other statistical processes. Summarise the accessibility and clarity of the data associated with the various dissemination formats, including relevant results from user surveys, and the effects of pricing policies and confidentiality provisions. Describe dissemination of data to Eurostat and other international organisations, and internal dissemination.
S.10.5. 1	AC 2. Metadata – consultations (P)	Number of metadata consultations within a statistical domain for a given time period.	For producer reports only, provide values of Indicator AC2 by month/quarter/year.

S.10.1-5 FURTHER GUIDELINES

Accessibility

For each dissemination mechanism:

- Describe the pricing policies and their likely effect on user access.
- Describe how(else) accessibility has been assessed (for example by user survey);
- Summarise the results of the assessment;
- Describe any changes that will be made to improve access in the near future.

Under S.10.3 describe the process of registration for on-line database access (if any) and comment on its ease of use.

Under S10.3.1 (producer reports only) analyse the values of standard QPI *AC1 - Number of consultations of data tables* and summarise the conclusions.

Under S.10.5, summarise the limitations to access that are the consequence of confidentiality provisions and how these have been explained to users with a cross reference to S.07 Confidentiality.

Clarity

For each dissemination mechanism:

- Describe how clarity has been assessed and the results of the assessment;
- Describe the changes that will be made to improve clarity.

Under S.10.5, describe any efforts that have been made to classify users by the level of detail of metadata they require and to match the metadata provided to these levels of needs.

Under S.10.5.1 (producer reports only), analyse the values of standard QPI AC2 Number of metadata consultations within a statistical domain and summarise the conclusions.

Other dissemination mechanisms

Under S.10.5:

- provide information regarding whether there is equal access for all users, or there is privileged access;
- provide information on user support for data access and interpretation;
- describe any data provided to Eurostat or other international organisations for example, IMF, OECD, or FAO, that are not already described under S.6.1 Legal acts and other agreements.
- describe internal provision of data to other statistical processes within the NSA that ultimately result in dissemination of these data, for example provision of data to the national accounts.

S.10.1-5 EXAMPLES

Example S.10.1-5-1 Occupancy of tourist accommodation establishments 2015, Italian National Institute of Statistics [ESS-MH]

[This example covers the dissemination mechanisms concisely at a minimal level of detail. Also note that, while the example contains links to the questionnaires, the questionnaires should be reported under S.18.3]

10.1. Dissemination format - News release

Capacity: See Annex I and Section 1 of Annual Report – http://www.istat.it/it/archivio/176210

Occupancy: See Annex I and Section 2 of Quarterly press releases of Flash Report http://www.istat.it/it/archivio/esercizi+ricettivi and Annual Report http://www.istat.it/it/archivio/193005

10.2. Dissemination format - Publications

Istat flagships:

Annuario statistico italiano (chapter 19 dedicated to tourism); Italia in cifre;

Noi Italia; Italian Historical Statistical Repository.

10.3. Dissemination format - online database

Data warehouse I.Stat: http://dati.istat.it/?lang=en.

10.4. Dissemination format - microdata access

Available upon request at the contact centre for all users.

10.5. Dissemination format - other

Other information collected and published at national level but not transmitted to Eurostat

- Capacity variables by sub-type of accommodation establishment (see point 3.2 Classification system).
- Arrivals and nights spent by sub-type of accommodation establishment (see point 3.2 "Classification system").
- Arrivals and nights spent by regions of residence for Italian residents.

Annex 1: Questionnaire (in English) - Occupancy; Questionnaire (in English) - Capacity

Example S.10.1-5-2 Service producer prices, 2014

Italian National Institute of Statistics [ESS-MH].

[This example also covers all dissemination mechanisms concisely at a minimal level of detail. The link provided to press release is no longer operational and thus has been excluded.]

10.1. Dissemination format - News release

No ministerial commentaries.

Press releases: (link)

10.2. Dissemination format - Publications

Name of national paper publications

- Release: “Comunicato stampa – Indici dei prezzi alla produzione dei servizi”.
- Release: “Annuario Statistico Italiano (ASI)”

Name of national electronic dissemination: Series are included in the ISTAT on-line database.

10.3. Dissemination format - online database

The on-line database is open and free. Statistics are organised by theme in a two-level hierarchical tree. Data are presented in aggregate form in multidimensional tables; acting on variables, reference periods and the arrangement of heads and sides users can create custom tables and graphs. The system can be searched by keyword, theme and region.

A wide range of standard metadata facilitates the retrieval and understanding of statistics by users.

More information available in the User Guide.

10.4. Dissemination format - microdata access

Before data dissemination, validated microdata are stored in the repository of the Institute, the Archive of validated microdata (ARMIDA). The repository was established with the main objective of preserving and documenting the data produced by Istat surveys and has subsequently supported the objective of disseminating data. The data archived in the repository supplies, in fact, the different channels for the dissemination of microdata (for internal use at the Institute through the Memorandum of access to microdata for internal users. The micro-data stored in the repository are also used to respond to requests of external users submitted to the Adele laboratory.

10.5. Dissemination format - other

Planned changes in national dissemination methods: None

S.10.6-7 Documentation on methodology and quality

SIMS	Concept name	Definition	Guidelines
S.10.6	Documentation on methodology	Descriptive text and references to methodological documents available.	List national reference metadata files, methodological papers, summary documents and handbooks relevant to the statistical process.

SIMS	Concept name	Definition	Guidelines
			For each item provide the title, publisher, year and link to on-line version (if any).
S.10.6.1	AC 3. Metadata completeness – rate (P)	The ratio of the number of metadata elements provided to the total number of metadata elements applicable.	*For producer reports only* Provide AC3: metadata completeness rate, noting that the “metadata elements” are the SIMS concepts and sub-concepts.
S.10.7	Quality documentation	Documentation on procedures applied for quality management and quality assessment.	List relevant quality related documents, for example, other quality reports, studies. Cross reference to descriptions of quality procedures in other chapters, especially S.13. *European level* Summarise availability of national quality reports.

S.10.6-7 FURTHER GUIDELINES

Documentation on methodology

For the most significant of methodological documents identified, provide a short summary of the content relevant to the statistical process being reported.

For producer reports only, analyse the values of standard QPI *AC 3. Metadata completeness – rate* and summarise the conclusions.

Documentation on quality

For each of the quality related documents that have been listed, provide a short summary of the content relevant to the statistical process being reported.

In cross referencing descriptions of quality procedures other chapters and accompany each cross reference with a summary of the pertinent quality aspects.

S.10.6-7 EXAMPLES

Example S.10.6-7-1 Turnover and volume of sales index, 2016

Statistics Netherlands

[This example illustrates a minimal response to “Documentation on methodology”. The link is not in English. The response to “Quality documentation” probably reflects a misunderstanding of what is to be reported.]

Documentation on methodology

Dissemination of documentation on methodology and sources used in preparing statistics: A short description is available via the Statistics Netherland website :

Quality documentation

Not available.

Example S.10.6-7-2 Urban Audit, 2013

Office for National Statistics, United Kingdom

[Although this example does not appear detailed, the UK Metadata Report to which it refers is very detailed indeed, as its introduction (copied immediately below) states.]

The main purpose of this report is to provide detailed metadata on the Urban Audit data for the United Kingdom that has been supplied to Eurostat. The report starts with some brief background on Urban Audit in the United Kingdom, its geography, and some consideration of the quality of the statistics provided. The main body of the report then takes each reference year in turn and provides metadata for each variable supplied. It should be noted that the scope of the report is limited to the data that had been supplied to Eurostat as at March 2015. Some additional datasets will be supplied once data become available – metadata for these datasets will be provided with the data.

10.6. Documentation on methodology

The definitions of the Methodological Manual on City Statistics have been followed for most variables. Additional information on the deviating methodology used in the collection is provided in the UK Metadata Report – 2013.

Annexes: UK metadata report - 2013.

10.7. Quality management - documentation

The quality assurance procedures detailed in the Methodological Manual on City Statistics have been applied. Additional information on quality is provided in Annex A - UK metadata report-2013.

S.11 (PART II) Quality management

SIMS	Concept Name	Definition	Guidelines
S.11	Quality management	Systems and frameworks in place within an organisation to manage the quality of statistical products and processes.	(Information relating to this concept is provided by reporting on its sub-concepts.)
This concept and its sub-concepts are included in both ESMS based (user) reports and in ESQRS based (producer) reports. In ESQRS based reports it is ESQRS Concept 4.			

S.11 Background

Definitions of key terms

As previously noted, all the quality related terms used in this Handbook are defined in the *ESS Quality and Metadata Reporting Glossary*, which is included as Supplementary Document A, and the key terms used in describing quality management are defined in Section 2.2 of Part I.

Quality assurance framework - benefits and components

Many NSAs have developed and implemented their own quality assurance framework (QAF) with the intention that it will support quality management by:

- providing a basis for creating and maintaining a quality culture;
- providing a systematic mechanism for ongoing identification of quality problems and possible actions for their resolution, from incremental improvements to full scale re-engineering;
- supporting continual quality improvement;
- stimulating interaction between staff throughout the organisation;
- containing reference material that is helpful in training;
- giving transparency to the processes by which quality is assured and reinforce the image of the organisation as a credible provider of good quality statistics;

- being the mechanism for exchange of ideas on quality assurance with other NSAs and international statistical organisations.

The components of a QAF typically include:

- organisational arrangements for managing quality – for example, with the aid of a quality unit, quality manager, quality committee;
- quality concepts, principles and dimensions – based on the ES CoP;
- quality guidelines – developed in house or borrowed from another NSA;
- programme for promotion of a quality culture – training programme to instill and maintain a quality culture throughout the organisation;
- quality and performance indicators (QPIs) – including the standard ESS QPIs and other QPIs needed to monitor statistical processes from identifying user needs to dissemination;
- quality monitoring and control – using QPIs, quality gates, statistical quality control;
- quality assessment programme and procedures – including self-assessment, peer review, external review/audit and any other evaluation programme;
- well defined relationship to other strategic directions and functions, such as methodology, metadata management, and risk management.

General quality management system

As noted in Chapter 2 (Part I), in addition to a QAF, some NSAs adopt, and adapt to their purposes, a general quality management system (QMS) that covers all aspects of quality management, not just the core statistical processes that are at the heart of a QAF. Sometimes the adapted QMS incorporates all the features of a QAF and there is no separate QAF.

The three most commonly used general QMSs are outlined in the following paragraphs.

European Foundation for Quality Management (EFQM) Excellence Model

The EFQM model comprises three integrated components:

- the *fundamental concepts of excellence* - the underlying principles that form the foundation for achieving sustainable excellence in any organisation;
- the assessment framework and management tool referred to as *RADAR* which is an acronym for *results, approaches, deploy, assess and refine*, reflecting a similar logic to the Deming Plan-Do-Check-Adjust (PDCA) cycle*;*
- the *criteria* – that provide a framework to help organisations to convert the fundamental concepts and RADAR thinking into practice.

ISO 9000, 9001 and 9004: Quality management systems

The ISO 9000 family of standards address the various aspects of quality management. The standards provide guidance and tools for organisations that want to ensure that their products and services consistently meet customer's requirements, and that quality is consistently improved. They can be used by any organisation, large or small, regardless of its field of activity, including NSAs.

- ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary provides the fundamental concepts, principles and vocabulary for a QMS and provides the foundation for the tor other standards in the family. As noted in Chapter 2 (Part I), it incorporates seven basic principles: customer focus, leadership, engagement of people, process approach, improvement, evidence-based decision making, and relationship management.
- ISO 9001:2015 Quality management systems builds on the seven quality management principles described in ISO 9000:2015: It sets out the criteria for a QMS. It is the only standard in the family with respect to which an organisation can be certified. There are over one million certified companies and organisations in over 170 countries, reflecting its perceived utility.

- ISO 9004:2018 Managing for the sustained success of an organization provides guidance to organisations to support the achievement of sustained success by a QMS.

Lean Six Sigma

Lean Six Sigma is a commercially developed standard. It does not have a particular owner. It combines two quality management methods.

- ***Lean*** focuses on streamlining both processes by eliminating waste while continuing to deliver value to customers. “Waste” is defined as any activity within a process that is not required to produce a product or provide a service to specification.
- ***Six Sigma*** is a method of efficiently solving a problem. Using Six Sigma reduces the number of defective products produced or services provided, resulting in increased revenue and greater customer satisfaction

Quality methods, tools and practices

As previously noted in Chapter 2, to support quality management implementation an NSA can draw on the set of quality tools developed by the ESS over the past 15 years and available at the Eurostat Quality Reporting webpage. It can also draw on statistical guidelines developed by individual European NSAs.

S.11.1 Quality assurance

SIMS	Concept name	Definition	Guidelines
S.11.1	Quality assurance	All systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the requirements for the statistical output.	Describe the procedures (such as use of a general quality management system based on EFQM or ISO 9000 series) to promote general quality management principles in the organisation. Describe the quality assurance framework used to implement statistical quality principles. Describe the quality assurance procedures specifically applied to the statistical process for which the report is being prepared, for example training courses, process monitoring, benchmarking, assessments, and use of best practices. Include descriptions of all forms of quality assessment procedures (such as user satisfaction survey, self-assessment, peer review, compliance monitoring, audit, labelling, certification) and when they most recently took place. Describe any ongoing or planned improvements in quality assurance procedures.

S.11.1 FURTHER GUIDELINES

The aim of S.11.1 is to describe the quality management and assurance procedures that are in place. It is **not** to present an assessment of the quality of the statistical process and its outputs, as this is summarised in S.11.2 and detailed when reporting on other quality related concepts, specifically including S.9, S.10, S.12, S.13, and S.14.

Note that many of the quality management and assurance procedures in place may apply to the NSA as a whole, not just for the statistical process that is the subject of the report.

Note that quality documentation (lists of documents and files) is reported in S.10.7 even though the content of this documentation is described here.

State whether the organisation has adopted a general quality management system (QMS). If it has:

- describe briefly the general QMS and how it has been adapted to, and implemented in, the organisation;
- if the general QMS based on ISO:9001 state whether the NSA (or some part of the NSA within which the statistical process being reported takes place) has been certified and if so when and with what results;
- if the general QMS is based on the EFQM, state the level the NSA has reached; and

State whether the organisation has defined and implemented a quality assurance framework (QAF) for its core statistical functions in addition to, or instead of, a general QMS.

At a level of detail appropriate for the report and without repeating anything that may have been included in describing a general QMS:

- summarise the organisational arrangements for managing quality, in particular whether there is a designated quality manager, a dedicated quality unit, and/or a senior level committee that takes responsibility for quality issues and improvements.
- state the origin of the quality concepts, principles and dimensions on which the QAF is based, for example, the ES CoP;
- indicate the source of the quality guidelines being used, whether they have been developed in house or obtained from an external source, for example, another NSA;
- refer to planned and actual quality improvements and to trade-offs between quality aspects;
- list the quality and performance indicators being used to monitor the statistical process and describe how they are used; and
- describe any other quality monitoring and control procedures, e.g., quality gates, statistical quality control.

S.11.1 EXAMPLES

Example S.11.1-1 Production in industry (Index of Production), Annual, 2016, Office for National Statistics, United Kingdom

[Note this example illustrates a minimal answer but with a link to more detail.]

Quality assurance

The ONS has developed Guidelines for Measuring Statistical Quality; these are based upon the European Statistical System (ESS) quality dimensions. More information can be found on the ONS website, see the following link for further details.

http://www.ons.gov.uk/ons/guide-method/method-quality/quality/guidelines-for-measuring-statistical-quality/index.html

Example S.11.1-2 Crop Production, Annual 2016, Department for Environment Food and Rural Affairs United Kingdom [ESS-MH]

[This example illustrates the elements of the corresponding quality assurance framework in the left hand column.]

Which kind of data validation measures are in place?	Comprehensive data checks, credibility checks, external challenge from the industry, consistency with weather conditions and market intelligence.
What do they target?	Range checks and completeness of survey returns. Response rate checks.
Are the data cross-validated against another dataset?	Yes
If yes, which kind of dataset?	Previous results (from same dataset)
Is there a quality management process in place for crop statistics?	Yes
If, yes, what are the components?	The Code of Practice for Official Statistics, assessment audits by the UK Statistics Authority, Department's quality strategy for statistics.
Is there a Quality Report available	Yes, various reports available
If yes, please provide a link	http://www.statisticsauthority.gov.uk/assessment/assessment/assessmer reports/assessment-report-22---assessment-of-agriculture-in-the-uk-and-selected-crop-and-livestock-statistics.pdf http://www.statisticsauthority.gov.uk/assessment/assessment/assessmer reports/confirmation-of-designation-letters/letter-of-confirmation-as-national-statistics---assessment-report-22.pdf
To which data source is it linked?	June Survey of agriculture and horticulture, Crop Production Survey, Stocks Surveys, Usage surveys, panel estimates for fruit and vegetables, Potato Council estimates.
Has a peer-review been carried out for crop statistics?	No.
If, yes, which were the main conclusions?	
What quality improvement measures are planned for the next 3 years?	Review of spreadsheet design. Simplification of data flows. Specific reviews under the Defra quality strategy. Areas for review not yet identified

Example S.11.1-3 Chemicals Regulation Division, Health and Safety Executive, United Kingdom

[Note. The point of including this example is to illustrate that it is easy to interpret S.11.1 as being a request to describe the quality of the statistical process and output. So, the response is an excellent description of possible errors and inaccuracies, but it belongs in S.13 Accuracy. What is required here is a description of the QAF which enabled such a precise assessment of accuracy to be obtained.]

Of the frame population of 98 companies, 5 companies were not surveyed as they had no product data in the United Kingdom Pesticide Guide (UKPG) 2016, the British Crop Production Council

(BCPC) confirmed that 10 companies had not sold products, withdrawn from the UK market since publication of the 2016 UKPG in January 2016, amalgamated or changed their name in that calendar year.

Of the 83 remaining units surveyed, 45 responded (5 of which reported combined data from 10 companies listed as separate in 2016 UKPG), with 2 outright refusals, 2 nil returns, 9 communicated that they made no sales but did not complete a return and 38 not responded despite 3 reminders. Of the 2 nil returns, and 9 who communicated no sales, 6 listed products that contained just one AS with no common AS across companies, and the remaining 5 listed amenity/domestic products - their quantities of AS would be small therefore BCPC did not consider this missing response to be significant. Of the 2 outright refusals, one is a multinational estimated to have approximately 8% share of the UK market by volume. This company is reported to have consistently refused to supply data so their nonresponse has no impact on year on year comparisons. BCPC has advised that the responding units (response rate = 55/83 = 66.26%) and the data provided accounts for more than 80% by volume of UK pesticides sales. Although a high response is no guarantee of data quality, when we account for factors such as the difference between responders and nonresponders, we feel confident with a response of 66.26% as a quality measure for a voluntary survey within a specialised sector….

S.11.2 Quality assessment

SIMS	Concept name	Definition	Guidelines
S.11.2	Quality assessment	Overall assessment of data quality, based on standard quality criteria.	Summarise the results of the most recent quality assessments and cross reference to the chapters in the report where the results are presented in more detail.

S.11.2 FURTHER GUIDELINES

Note that the *results* of quality assessment procedures are described here. The actual quality assessment procedures are an integral part of the quality assurance and are described in S.11.1.

State if the statistical process was subject to self-assessment during the reporting period, and if so

- what quality assessment tool(s) were used; and
- in which sections of this report are the results are presented.

State if the statistical process was subject to a peer review during the reporting period, and if so

- what quality assessment tool(s) were used; and
- in which sections of this report results are presented.

State if the statistical process was subject to an external review/audit during the reporting period, and if so

- what was the basis for the review/audit;
- what quality assessment methods and tools were used: and
- in which sections of this report results are presented.

S.11.2 EXAMPLE

Example S.11.2-1 Production in industry (Index of Production), 2016

Office for National Statistics, United Kingdom

[This example illustrates a minimal response, but it provides links to two quality related reports. The first link points to a description of the methods and improvements; the second to an assessment of compliance with the UK Code of Practice]

The work that the ONS is undertaking to monitor the quality of published statistics, and develop improvements to existing data can be found in the following link:

https://webarchive.nationalarchives.gov.uk/20160106003944/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/index-of-production/index.html

Furthermore, the UK Statistics Authority publishes a report on each assessment it carries out. The Index of Production is assessed as part of the Short Term Economic Output Indicators and the assessment can be found in the following link, report number 278.

https://www.statisticsauthority.gov.uk/wp-content/uploads/2015/11/images-assessmentreport278statisticsongdpiopandio_tcm97-437252.pdf

S.12 (PART II) Relevance

SIMS	Concept name	Definition	Guidelines
S.12	Relevance	The degree to which statistical information meet current and potential needs of the users.	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and its sub-concepts are included in both ESMS based (user) reports and in ESQRS based (producer) reports. In ESQRS based reports it is ESQRS Concept 5.

The slightly different treatment between ESMS reports and ESQRS reports of the quality and performance indicator RI, is indicated by (U) and (P).

S.12 Background

Relevance as a quality component has to do with the *output content* of a statistical process. Content issues include, target population, variables, definition of parameters (such as totals, means, counts or indexes) and periodicity of publishing results.

Relevance is the extent and degree to which the content of a process satisfies the needs of different users. It depends on whether all the statistics that are needed are produced and the extent to which concepts used (definitions, classifications etc.,) reflect user needs.

Descriptions of content – data and presentations - are provided in S.03 Statistical Presentations.

User needs and satisfaction also relate to other quality components such as accuracy, timeliness and comparability, which are treated in separate chapters. However, for practical reasons a comprehensive picture of user needs and satisfaction should normally be provided in one place in a report, covering all aspects of the process. The recommendation is that this be done under the heading of relevance.

S.12.1 User needs

SIMS	Concept name	Definition	Guidelines
S.12.1	User needs	Description of users and their respective needs with respect to the statistical data.	Provide: • a classification of users, also indicating their relative importance; • an indication of the uses for which users want the statistical outputs; • an assessment of the key outputs desired by different categories of users and any shortcomings in outputs for important users; • information on unmet user needs and any plans to satisfy them in the future; and • details regarding those quality components which do not meet user requirements.

S.12.1 BACKGROUND

There are varying needs of different classes of users, for example, government, business, academia, and the public at large. The NSA's challenge is to weigh and balance the (possibly conflicting) needs of current and potential users and to produce statistics that satisfy the most important, or majority of, these needs within given resource constraints.

In learning about user needs, one approach is to ask users directly about the statistics that the NSA produces through a user survey. Another approach, which provides indirect evidence, is to ascertain if there are processes in place to determine the uses of data and the views of their users.

The description of user needs should address all quality components. Such needs are expressed not only in terms of data content but also in terms of the degree of accuracy required, the timing, the dissemination arrangements, the metadata required for interpretation, and the relationship to other relevant statistical outputs. In other words, they cover the whole range of the output quality components.

Assessment of user needs is not trivial, first because there are many types of users, second, because there are many different uses for which the users want the outputs, and hence multiple output requirements, and finally because there are three aspects of needs – data, accompanying metadata, and support by the NSA.

S.12.1A Understanding and classifying users

S.12.1A FURTHER GUIDELINES

The first step is to assemble information about the *users* - who they are, how many they are, and how important they are individually and collectively from the perspective of the NSA. Based on information available from, for example, advisory committees, lists of paying recipients, and Internet accesses, the usual approach is to develop a *classification of users,* to estimate the number of each type, and identify the *key users,* i.e., the users that are sufficiently important to require individual consideration.

The second step is to determine the *needs* of each class of users in terms of ideal population and data content, and, in the case of key users, their individual needs. For users, acquiring output data is a means to an end, not an end in itself, and the uses to which these data are put are relevant and should be identified. Quite frequently, users may not fully understand what data they actually need or may not know what is available. By understanding the uses of data, the NSA is in a better position to determine the actual needs. Furthermore, these needs have to be interpreted in the statistical context in which they are to be addressed. The concepts, accuracy, timing, etc., have to be aligned with what can actually be delivered.

Obtaining information about user needs, and the uses for which data are needed, is typically accomplished through domain specific advisory committees, user groups, ad hoc focus groups, requests, complaints, user surveys (as discussed in Section S.12.2), and other user feedback.

The third step is to determine, in general terms, the *priorities* to be given to the key users and the various classes of users in satisfying their needs. For example, the needs of government policy makers may be set ahead of those of academic researchers. Some needs are important but transient. Some users may also be respondents and their requirements merit special consideration.

In summary, the report should contain a classification of users, complete names of key users, an indication of the uses for which users want the outputs, the priorities in satisfying their needs, and an account of how all this information was obtained and used.

S.12.1A EXAMPLES

Example S.12.1-1 Balance of Payments and Related Results Compilation 2011, Ireland Central Statistics Office Ireland, 2013, p. 13

[The example illustrates a description of the main users and their use of the data]

These statutory inquiries are conducted to meet the requirements of Regulation (EC) No 184/2005 of the European Parliament and of the Council of 12 January 2005 on community statistics concerning balance of payments, international trade in services and foreign direct investment (as amended by Regulation Nos 601/2006, 602/2006, 1137/2008 and 707/2009) and the ECB Guideline ECB/2004/15 (as amended by ECB Guideline ECB/2007/3 and recast in Guideline ECB/2011/23) on the statistical reporting requirements of the European Central Bank in the field of balance of payments and international investment position statistics.

As a result of its role in monitoring Ireland's economic performance, the Department of Finance is interested in all aspects of the balance of payments (BOP). The main focus of the Department of Enterprise, Trade and Employment is on industrial development in the manufacturing and services sectors. This Department and Forfás, an agency operating under its aegis and involved in attracting foreign direct investment to Ireland, are particularly interested in the direct investment aspects of the BOP, as well as in the data on merchandise and services. Data are also used by stockbrokers, analysts in the field of economic and social research as well as universities. The National Accounts Division also uses BOP results internally within the Central Statistical Office. The Central Statistical Office supplies data to international organisations such as the ECB, the European Commission (Eurostat), the IMF and the OECD.

Example 12.1-2: Measuring the Non-Observed Economy: A Handbook OECD et al, 2002, p. 98

[The example provides an exhaustive list of users.]

In the Handbook prepared by the OECD, IMF and other international organisations, there is a grouping of users under nine broad headings:

- national government – the national bank, and the ministries dealing with economic affairs, finance, treasury, industry, trade, employment, environment;
- regional and local governments;

- business community – individual large businesses and business associations;
- trade unions and non-governmental organisations;
- academia – universities, colleges, schools, research institutes, etc.;
- media – newspapers, radio and TV stations, magazines, etc.;
- general public;
- international organisations

Example S.12.1-3; Assessment of relevance, EU-Statistics on income and living conditions, Eurostat, 2013, p. 3-412.7

[The example provides a list of types of users.]

The relevance of an instrument has to be assessed in the light of the needs of its users. As for EU Statistics on Income and Living Conditions (EU-SILC) the main users are the following:

- Institutional users like the Directorate-General of Employment, Social Affairs and Inclusion and the Social Protection Committee, in charge of the monitoring of social protection and social inclusion, or other Commission services;
- Statistical users in Eurostat or in National Statistical Institutes to feed sectorial or transversal publications;
- Researchers having access to microdata; and
- End users – including the media - interested in living conditions and social cohesion in the EU.

The EU-SILC instrument is the main source for comparable indicators for monitoring and reporting on living conditions and social cohesion at the EU level. It has been moreover recognized by Heads of States and Governments as the data source for the Europe 2020 strategy headline target on poverty.

S.12.1B Assessment of outputs relative to user needs

S.12.1B FURTHER GUIDELINES

An assessment of the *key outputs* desired by different categories of users should be given and any shortcomings in outputs for important individual users and classes of users should be mentioned. Shortcomings could involve, for example, data items not being available, definitions that are inappropriate, insufficient breakdown of data into sub-domains, time series that are too short, or outputs that are too infrequent, for example quarterly instead of monthly. Not all user needs can be met, reasons being budgetary and/or technical.

Shortcomings in data content

Shortcomings in the target population and data content relative to the ideal population and content from a user perspective should be identified and summarised. For example, the ideal population for a user who wants to analyse the economic status of an industry is typically ***all*** enterprises that are active in that industry. For cost or response burden reasons, an NSA may want to limit the target population to the ***large and medium size*** enterprises in that industry, based on turnover, number of employees, or other criterion. The difference between all enterprises (the ideal population) and the set of large and medium size enterprises (the target population) is a relevance issue which should be reported together with some sort of assessment of the proportion (in terms of turnover, number of employees) of an ideal population that has been omitted.

Discrepancies in concepts

Discrepancies between *actual* definitions of statistical concepts and the definitions that would be *ideal from a user perspective* should be given. Concepts defined during the design and planning of the statistical process include target population, target definition of units, and aggregation formula. It is often the case that what is ideal differs between users and, if so, this should be noted. Sometimes it is possible to present results for more than one definition from the same microdata. More usually this is not possible, and a single definition has to be selected, in which case the motivation for the chosen definition should be given.

Any discrepancies between the definitions used and accepted ESS or international definitions should be clearly pointed out.

In **administrative data processes**, data item definitions are often a critical aspect since they are determined by administrative considerations and may not be ideal for statistical purposes. Important discrepancies in this regard should be described.

In **price indexes**, although defined in general terms by economic theory, the target of estimation is usually impossible to specify exactly and is even open to some controversy. A producer report should discuss important issues concerning the target of estimation and its relation to approaches and methods chosen, also relating these to recommendations in international manuals and legal documents in the ESS system. See Example S.12.1B-1 below.

For **macro-aggregate compilation processes**, a producer report needs to relate the chosen definitions and concepts to those recommended in regulations, international manuals or other general agreements. Important discrepancies should be mentioned.

For statistical processes such as **price indexes and national accounts**, recommended definitions and estimation methods are given in international handbooks and manuals. If the recommendations are followed, then, in terms of the primary uses of the outputs, there can be no relevance issue. For secondary uses there may well be relevance problems that need to be highlighted. Incidentally, failure to fully follow the recommendations may result in accuracy issues, which should be reported in Chapter S.13.

Numerical illustrations of the likely sensitivities of the results to the chosen definitions can be very informative and should be provided whenever possible. The basis for these illustrations could be sensitivity analyses or simulations. Such illustrations inform users of the risks of a *relevance problem* for their particular application, i.e., of a discrepancy between the definitions used and what they want.

Relationship to other quality components

Definitions also affect coherence and comparability (see Chapter S.15) and can be discussed instead, or as well, under that heading.

There is a grey zone between certain relevance problems and accuracy, as illustrated in the case of cut-off sampling (see S.13.2A)Summary

Highlight where concepts, data content and target population are less than ideal for key users and classes of user. The reasons for not being able to meet important user needs should be explained.

The report should cover all the points in the analysis above including descriptions of user needs, unmet user needs, the reasons why certain needs cannot be fully satisfied, and any plans to satisfy needs more completely in the future.

S.12.1B EXAMPLES

Example S.12.1B-1 Discussion on the purpose of HICP as a CPI, Eurostat, 2001, p. 36-37

[The example presents and discusses the target concept of a complex statistic.]

Relevance refers to the purpose of the HICP. As noted in Section 3.1. above the aim of the HICP is to measure inflation as distinct from the cost of living. It is therefore inappropriate to criticise the HICP from the latter perspective. However, a great deal has been said over the years about bias in CPIs without recognition of the fact that there is a limit to what can be said with any degree of certainty. Unless the target has been precisely defined, it is impossible to say by how much it has been missed. CPIs can be compared one with another, and it can be argued that certain differences should be removed, as has been done in the harmonization process, but there is no operational definition of the unbiased index by which to judge all other CPIs. Each CPI has been developed over a long period of time with the index compilers solving the operational problems in as consistent and coherent a way as possible. The actual conceptual framework for any CPI is thus embodied in its history. Meanwhile, efforts have been made to build alternative conceptual frameworks relying on economic and statistical theory. These ideas have influenced index design but have not, for the most part, determined actual operational practice.

The Treaty and the framework Regulation define the HICP. The Treaty requires a consumer price inflation index; the framework Regulation requires that it should be a Laspeyres-type index measuring the average change in the prices of goods and services available for purchase in the economic territory of the MSs. This definition was agreed, following the requirement of the Treaty, between Eurostat and the main users. As such, the definition constitutes a broad operational definition of ‘inflation’.

Example S.12.1B-2: Compiled variables in Short-term Business Statistics, Building Permits (411 and 412), Bulgaria, (Eurostat1, 2011, p. 6-7)

[The example provides a list of variables compiled and their purposes.]

Please indicate which variables are compiled for national and STS Regulation purposes.

Data item	For national purposes (X)	For STS Regulation (X)
Building permits: number of building permits	x	
Building permits: number of buildings	x	
Building permits: number of dwellings	x	x
Building permits: useful floor area	x	x
Building permits: alternative size measure (sq. m)	x	x

S.12.2 User satisfaction

SIMS	Concept name	Definition	Guidelines
S.12.2	User satisfaction	Measures to determine user satisfaction.	Describe how, and how often, the views and opinions of the users are collected, for example by user satisfaction surveys or other user consultations. State how often such investigations are conducted and when the most recent took place. Present the key results from the recent investigations. Present view of user satisfaction over time, in the form of a user satisfaction index if available.

S.12.2 FURTHER GUIDELINES

User satisfaction is the number one priority. The most comprehensive method of evaluation is a full-scale user satisfaction survey, conducted in accordance with normal survey best practices - drawing a representative sample of users from an appropriate frame (if available), designing and testing a suitable questionnaire, collecting, processing and analysing the results, etc.

Conducting a user satisfaction survey is not always affordable, particularly for small statistical processes where it would represent a significant share of the operation’s total budget. Other methods of assessment include analysis of publication sales, user comments, requests and complaints received, web site accesses, and feedback from advisory committees and focus groups.

The methods used for assessment and the measures taken to improve user satisfaction should be described.

If appropriate and available, the results of a user satisfaction survey, or analyses, broken down by the most important classes of users should be provided, with reference to more complete information elsewhere.

S.12.2 EXAMPLES

Example S.12.2-1 User Satisfaction Assessment for Euro-SICS database (Ladiray & Sartori, 2001,p. 647)

[This example illustrates provision of information about an evaluation of user satisfaction, including the method used.]

Eurostat conducts an evaluation of user satisfaction for the European Statistical Indicators Common Site (*Euro-SICS) database* containing Euro-zone short-term indicators. It is undertaken mainly through continuous dialogue with its two main users, Directorate-General for Economic and Financial Affairs and the European Central Bank. The January 2001 Quality Report noted that users requested “more indicators but less breakdowns”. This is obviously the type of information that helps give an idea of the relevance of the output and to orient future developments

Example S.12.2-2 Report on the EUROSTAT 2014 User Satisfaction Survey 2014, p. 2

[This example illustrates provision of information about a user satisfaction survey.]

Eurostat's mission is to be the leading provider of high-quality statistics on Europe. In order to measure the degree to which it meets its obligations towards its users, Eurostat carried out a general User Satisfaction Survey (USS) over the period of April – June 2014. It was based on the agreed model questionnaire for the European Statistical System and was designed to obtain a better knowledge about users, their needs and satisfaction with the services provided by Eurostat. The first survey of this kind was held in 2007 and then repeated in 2009, 2011, 2012 and 2013. The USS 2014 was, therefore, the sixth of a general nature.

The survey covered four main aspects:

- information on types of users and uses of European statistics;
- quality aspects;
- trust in European statistics; and
- dissemination of statistics.

Example S.12.2-3 User Satisfaction Assessment for the EuroGroups Register (EGR), the statistical business register of multinational enterprise groups in Europe [ESS-MH]

[This example gives results from a user consultation survey.]

In April 2017, an SBR user consultation survey was conducted by Eurostat. The survey addressed users and potential users of the national statistical business registers (SBRs) and the EuroGroups Register (EGR).

The survey results show that 35.8% of respondents regularly use information provided by the EGR in production of statistics.

The EGR data are used mostly in business statistics (44%) and globalisation (37%). The use of EGR in macroeconomic accounts is 7%.

S.12.3 Completeness

<table>
<tr><th>SIMS</th><th>Concept name</th><th>Definition</th><th>Summary Guidelines</th></tr>
<tr><td rowspan="2">S.12.3</td><td>Completeness</td><td>The extent to which all statistics that are needed are available.</td><td rowspan="3">Provide qualitative information on the extent to which content requirements in relevant legislation, regulations and guidelines are met. Where such requirements are not fully met, reasons for this should be provided.
Provide information on the extent to which user needs related to content are satisfied.
Provide values of indicator R1 Data completeness rate, for each required data item for each relevant regulation/ guideline at producer/user level of detail as appropriate.
In the case where the indicator refers to data sent to Eurostat, this indicator can be compiled by Eurostat.
European level
Summarise across countries the extent to which ESS requirements for data items are met</td></tr>
<tr><td>R1 Data completeness rate (U)</td><td rowspan="2">The ratio of the number of data cells provided to the number of data cells required by a regulation/ guideline.</td></tr>
<tr><td>S.12.3.1</td><td>R1 Data completeness rate (P)</td></tr>
</table>

S.12.3 FURTHER GUIDELINES

Completeness of statistics may relate to:

- requirements in regulations or guidelines, usually those at EU level; and/or
- other user needs, for example at national level.

Regulatory requirements

If certain data items and/or segments of the populations specified in applicable legislation, regulations or guidelines are not covered, the statistics are defined to be *incomplete*. The legislation, regulations or guidelines may be national, or ESS, or international.

The concept of completeness can be broadened to refer to a whole subject domain. For example, the *cultural statistics* domain includes many areas like book sales, libraries, cinemas, theatres, concerts etc. If one of these areas is not covered by statistics, then the statistics on culture can be said to be incomplete. This level of completeness would not normally be covered in a producer report for a single statistical process.

There is a distinction between *incomplete statistics* (to be reported here) and *undercoverage* (to be reported under S.13 Accuracy). Incomplete statistics means that certain data items or segments of the population specified in a legislation/regulation/guideline are not included, whereas undercoverage refers to a discrepancy resulting from the survey methodology between the target population and the survey population.

An explicit statement of the degree of completeness of the statistical outputs in terms of all applicable legislation, regulations and guidelines should be given, including plans for improvements in this respect. Where completeness is not 100%, a statement about plans for future improvements future should be provided.

Other user needs related to content

Where there are known user needs other than those addressed by regulations, an assessment of how they are met should be provided. For example, where EU harmonisation is a priority, the specific needs of important national users may not be fully satisfied.

In Chapter S.03, data descriptions are to be provided. In Chapter S.12 the data published could further be related to known user needs. An assessment regarding the key outputs (variables, definitions, periodicity etc.) desired by different categories of users should be given and any shortcomings in outputs for important users be mentioned. This could, for example, involve insufficient breakdown of data into sub-domains, time series that are too short, or quarterly instead of monthly publications. Not all user needs can be met, reasons being either budgetary or technical.

The report should include information on important unmet user needs, the reasons why certain needs cannot be fully satisfied, and any plans to satisfy needs more completely in the future. (This information could alternatively be provided under S.12.1.)

Numerical illustrations of the likely sensitivities of the results to the chosen definitions can be very informative. The basis for these illustrations could be sensitivity analyses or simulations. Such illustrations inform users of the risks of a ***relevance problem*** for their particular application, i.e., of a discrepancy between the definitions used and what the user wants. For example, various definitions of unemployment are preferred by different users and often give quite different results. These differences are quite easy to illustrate and also explain quantitatively.

R1. Data completeness rate

The R1 indicator is defined in Supplementary Document C. Two definitions are given, one for user reports the other to producer reports. In either case it should be noted that:

- not all output data items are of equal importance; thus, an appropriate scheme for weighting the items may improve the usefulness of this indicator but could reduce comparability across countries;
- the set of data items to be taken into account, and the weighting scheme, are at the discretion of the manager of the statistical process.

S.12.3 EXAMPLES

Example S.12.3-1 Relevance of Statistical Concepts in Slovenian Household Budget Survey (HBS) Arnež et al., 2008, p. 9-10

[This example gives detailed statistics on completeness but not using the standard QPI.]

Share of missing statistics

The share of missing statistics is 0.007 (3/457), considering all variables which should be submitted to Eurostat. The implementation of HBS is not governed by regulations of the European Commission. Therefore, Eurostat collects data provided in this questionnaire under a Gentlemen's Agreement, every 5 years. The document „Data transmission for the HBS round of the reference year 2005" as of the end of January 2004 lays down 457 variables which should be communicated to Eurostat. Of these, 430 are basic variables and 27 derived variables at the household level. In order to calculate derived variables at the household level, 16 basic and derived variables at the level of a member should be calculated, which are not to be submitted to Eurostat. Of the basic variables at the household level, there are only three which we cannot ensure: HD02 (furnishing of a rented dwelling), HD03 (type of dwelling; individual houses cannot be divided into two types); HD08.01 (the number of years spent in the present dwelling). The missing variables are included in the HBS questionnaire as from 2005 onwards; therefore all variables required will be provided in the future. On 15 June 2007, individual data at the household level for 2004 were communicated to Eurostat (on the basis of data collected in 2003, 2004 and 2005), and 25 tables for 2004, which included data for 2004 with the consumer price index, calculated according to the Eurostat reference year 2005. The small size of the sample is the reason that the HBS data is available only at the state level; tables for some requests are made simultaneously. In order to satisfy the needs of users as much as possible, we plan to elaborate additional standard tables considering their present demand.

S.13 (PART II) Accuracy and reliability

SIMS	Concept name	Definition	Guidelines
S.13	Accuracy and reliability	Accuracy of data is the closeness of computations or estimates to the exact or true values that the statistics were intended to measure. Reliability of the data, defined as the closeness of the initial estimated value to the subsequent estimated value.	(Information relating to accuracy is provided by reporting on S.13 sub-concepts. Information relating to Reliability is reported in S.17 Data Revision).

This concept and all its sub-concepts are included in ESQRS based (producer) reports, where it is ESQRS Concept 6.

The concept and some of its sub-concepts are included in ESMS based (user) reports. Those sub-concepts that are included but treated differently in ESMS are indicated by (U); those that are included only in ESQRS are indicated by (P).

S.13 Background

The concept of accuracy relates a numerical estimate to its true value according to an agreed definition. The closer the estimate is to its true value, the more accurate it is. The difference between the estimate and the true value is called the *error* of the estimate and error is thus a technical term to represent the degree of lack of accuracy. The error has a random component (variance) as well as a systematic component (bias). It is sometimes better to speak of uncertainty than error, when the term error risks to be confused with a mistake committed, which is a very different matter.

The concept of accuracy, thus defined, is universal across all types of statistics. However, the true value is rarely available and hence accuracy needs to be assessed in the form of indicators, quantitative or qualitative. Occasionally, a true value (or at least something close to a true value) is available at a later point in time and in these situations such an "ex-post" evaluation is very useful. A situation of this kind is a difference between a preliminary estimate and the final estimate of a variable, where the final estimate can be assumed to be closer to the true value than the preliminary estimate. (Also, see S.17 Data Revision.)

Probability survey

For surveys based on probability sampling a model with a decomposition for total survey error (TSE) has been established. This is not the case for the other types of statistical process.

The error components for a probability survey are as follows:

- *Sampling error*. That part of the difference between a population value and an estimate thereof, derived from a random sample, due to the fact that only a subset of the population has been enumerated.
- *Non-sampling error*. An error in survey estimates that cannot be attributed to sampling fluctuations, divided into:
  - *Coverage error*. Errors due to the divergence of the survey population from the target population.
  - *Measurement error*. Errors that occur during data collection and cause the collected values to be different from the true ones.
  - *Nonresponse errors*. Errors that result from the failure to get a response to some, or possibly all, of the questions.
  - *Processing error*. Errors in data processing subsequent to collection such as data entry, keying and editing.
  - *Model assumption errors.* Errors due to the domain specific models that are needed to define the target of estimation.

For probability surveys it is theoretically possible to measure total error of a certain estimate in terms of variance and bias. The variance is mainly the result of sampling. Certain measurement errors can also have random components that, in principle, contribute to the variance of the estimates but in practice these errors are not possible to separate from sampling errors. Bias is the net effect of all the other error sources (some probability designs may also result in biases that are mostly of minor importance).

Sampling theory provides techniques for giving an objective, scientific, measure of the random error affecting published estimates. Furthermore, sampling biases are normally zero or negligible so that the variance can be taken to represent total sampling error. The variability of an estimator around its expected value may be expressed by its variance, standard error, coefficient of variation (CV), or confidence interval.

As regards non-sampling errors, computation of the bias requires knowledge of the true population value and detailed knowledge of the survey processes. In practice, it is often possible to have an informed idea about whether the bias risk is upwards or downwards, but rarely possible to estimate its size well and some examples of this could be found in this chapter([7]).

The *total error* of an estimate relative to the unknown true population value is expressed as the *root mean square error (RMSE)*, defined as the square root of the sum of variance and the square of the bias. Although being the most relevant direct measurement of accuracy from a user point of view, the RMSE can rarely be estimated.

Other statistical processes

For other types of statistical process, there are no agreed definitions of error components. The accuracy report therefore needs to follow a unique structure in each case according to the specific error profile([8]) of the process. Nevertheless, the types of errors defined above can often be used, but it is important that the definitions are made clear in each case. In the following we give some suggestions.

Non-probability surveys are obviously also affected by sampling error although there is no elaborate theory that can be used for error estimation. It may be considered that all of sampling error is in fact

([7]) In S.13.3.1C coverage issues for censuses in the form of undercounts or overcounts are such cases. By subject-matter related knowledge there are often strong hypotheses on the direction of non-response bias. In price indexes, the direction of formula biases or quality adjustment biases are sometimes quite clear.

([8]) An error profile is the full set of errors of all types affecting an estimate. Brooks and Bailar (1978) define it as: “A single document that catalogs what is known about each of the component errors present in a given survey.”

sampling bias and no random error exists. Another, model-based, approach would be to assume that the sample is in fact “effectively random” in some sense and use sampling variances as estimates of the random error component. The difference between estimates from the sample obtained and the target population value could either be seen as sampling bias or as coverage error (bias). Nonresponse is not a meaningful concept since the set of respondents is not fixed in advance by the sampling procedure. Measurement and processing errors are of the same nature as for probability surveys.

For ***administrative data processes*** there are often coverage issues in relation to the target population, which can be referred to as coverage error (over- or undercoverage). For some variables there may be missing data that can be considered as (item) nonresponse. Measurement errors can also be present. No sampling error normally exists.

For ***macro-aggregate processes*** there are often sampling errors in probability sampling based components from which the aggregates are built. Coverage issues could refer to gaps in the primary statistics going into the aggregates. Model assumptions are often applied in order to fill these gaps and other deficiencies in primary statistics but there may be model assumption errors affecting the estimates.

Below, two examples are given, where particular error structures are used for administrative data processes.

Example S.13-1 Pension beneficiaries

Based on *Quality Report ESSPROS Pension beneficiaries, Eurostat (2011)*

[This example describes a particular error structure that is not defined in terms of the typology used in sampling surveys]

Statistics on pensions in EU countries use a variety of sources but administrative data is the main source for most countries.

Sampling errors are thus small and not described at length in existing producer reports. Instead accuracy issues are:

- Geographical coverage in that the whole country is sometimes not covered.
- Coverage of pension schemes. In any country there is a multitude of different pension schemes that are typically not covered by administrative data to 100 %.
- Various methods are used for estimating pensions where direct data are missing. Each country uses its own detailed methodology which is briefly described in producer reports.
- The producer report at EU level attempts to summarize the methods used by countries but a numerical error indicator is not considered possible to present.

A more complete report on accuracy would go further in analysing the estimation methods.

Example S.13-2 Quarterly financial accounts of General Government

From *quality report 2007-05-03 pursuant to Regulation (EC) No 501/2004 on quarterly financial accounts for general government*

[This example describes a particular error structure that is not defined in terms of the typology used in sampling surveys]

Statistical processes are administrative data (own accounts of government units or central databases) and surveys.

Main accuracy issues as stated in the report are:

- Coverage of government units;

- Coverage of financial instruments;
- Valuation of equity or securities – at market value (correct) or face value. Could be regarded as measurement error;
- Time of recording transactions;
- National practices with regard to consistency and plausibility checks.

S.13A TECHNIQUES FOR EVALUATING ACCURACY

For any particular type of statistical process there are unique opportunities for error checking and evaluation. This section gives a few examples, mostly to inspire producers to invent other methods, similar or not, that are suitable for evaluating their particular statistical processes. Creativity is certainly a virtue in this field.

Mirror statistics

The classical example of mirror statistics is for international trade in goods. In principle, country A's exports to country B over a certain period must equal country B's imports from country A. In practice, the comparison is blurred by factors such as valuation (i.e. whether freight and insurance is included or not), timing (arrival at B may be later than departure from A), classification differences and coverage errors due to different thresholds for inclusion. However, adjustments for these factors can usually be made so that the extent of the actual errors can be more or less accurately determined.

Another case where mirror statistics can be of use is for statistics on migration.

Mirror statistics can be used for detecting accuracy problems but not for estimating their size and are thus rather the beginning than the end of an analysis of accuracy. See also S.15.1 Geographical comparability.

Unexplained variation over time in event-reporting

In event-reporting statistics, there is normally some stability in reporting patterns from the relevant authorities (police, hospitals, customs, etc). Lags in reporting or failure to report by a particular local institution cause undercoverage. It is simple to keep track of the reports from each institution subject to the reporting obligation. If this is done irregularities in the number of reports give rise to suspicion that something is wrong and corrective action can be taken.

Reasonability arguments

In all statistics, subject-matter knowledge on what is possible and reasonable is a useful tool. Often, all that is needed is a creative use of common sense. A more intricate example of such an argument is used in price statistics as described in the following example.

A control statistic based on a reasonability argument.

Quality change is a phenomenon that affects accuracy in a Consumer Price Index. A tool for evaluating the risk for bias in the CPI is illustrated below.

For a certain product in a Consumer Price Index, a raw average price for all observations in any given month can be computed. The ratio between such average prices between two months could be called the raw price index, which will differ from the actual price index due to implicit or explicit quality adjustments. Now the following statistic can be calculated.

IQI=Implicit quality index =(raw price index)/(actual price index).

If the quality adjustments are correct and the IQI shows an increase of 10 %, then this implies that there has been a 10 % quality improvement in the product concerned. This could then be tested against the general consumer experience, which may for example be that quality improvements have occurred for high-tech goods (PCs, cars, TVs, stereos etc.) but not for non-technical goods such as clothing and household utensils.

S.13.1 Overall accuracy

SIMS	Concept name	Definition	Guidelines
S.13.1	Overall accuracy	Assessment of accuracy, linked to a certain data set or domain, which is summarising the various components.	Describe the main sources of random and systematic errors in the statistical outputs and provide a summary assessment of all errors with special focus on the impact on key estimates. The bias assessment can be in quantitative or qualitative terms, or both, and may be expressed as bias risk. It should reflect the producer's best current understanding (sign and order of magnitude) and include actions taken to reduce bias. *European level* Provide a summary picture of accuracy across countries. The emphasis placed on various types of errors should depend upon the error profile of the respective process. For repetitive processes, describe how accuracy is developing over time and what efforts are underway to improve accuracy from an ESS perspective.

S.13.1 FURTHER GUIDELINES FOR ALL PROCESSES

For any statistical process, the most important errors and error risks should be identified. A producer report should then treat each component of error or uncertainty according to its relative importance in terms of overall accuracy. Reporting should reflect the best knowledge and understanding of the producer, whether this knowledge can be expressed in quantitative terms or only in qualitative terms. Sometimes, a standardised error structure can be used but on other occasions the structure may have to be tailored to the particular statistical process.

Under overall accuracy the big picture should be provided, identifying the major error sources and summarising them. Where one type of error dominates the picture an indication of its size should be provided, where possible. It should be noted that the effects of error sources on final estimates are more important than the errors themselves.

S.13.1 EXAMPLES

Example S.13.1-1 QUALITY DECLARATION. National Accounts, quarterly and annual estimates Statistics Sweden Version 1 2018-09-13([9])

[This example illustrates a very comprehensive assessment of overall accuracy.]

Overall accuracy

The description of accuracy is limited to the accuracy of GDP in total.

The annual estimates, published after 21 months, are based on more complete and detailed statistics than the statistics available for quarterly estimates published two months after the end of each quarter. In simple terms, it can be said that the annual estimates determine the level of GDP and other aggregates in the national accounts system, while quarterly estimates are a way of distributing

([9]) See https://www.scb.se/contentassets/66e9dae3a5d94bf8b4c299ce25294348/nr0103_qd_2018.pdf for the complete quality declaration.

the results from the annual estimates on each quarter and for estimating current quarters.

The national accounts, both the annual and the quarterly, are based on a large number of primary statistical sources. Accuracy depends largely on the quality of the different sources and on the model assumptions used to estimate the target characteristics of the national accounts. In some cases these may differ significantly from the target characteristics of the primary statistics.

For some areas there are no recurring short-term statistics produced, or recurring statistics at all, necessitating reliance on model assumptions.

An aggregate measure of accuracy is not possible to compile due to the large number of sources, the model assumptions and the balancing between the estimates from the expenditure and production approaches in order to achieve one single GDP estimate.

GDP calculated from the production and expenditure approaches or sides are, in theory, identical. However, there is always a discrepancy between these calculations, which, as far as possible, should be based on separate sources. The approach that ends up being the highest or lowest varies over time. Part of the compilation process is to balance the accounting system so that the expenditure and production approaches result in the same estimate of GDP. The size of the discrepancy between the two approaches varies between years and different quarters.

Annual compilations

In the annual compilations, production and expenditure are broken down to approximately 400 product groups. The supply side comprising production and imports, is confronted with the use of each product group. The use consists of intermediate consumption, consumption, gross capital formation and exports. Due to the uncertainties in the calculations, larger or smaller discrepancies arise in the different product balances. By analysing these, some errors can be identified and corrected. The analysis also provides support for assessments on how supply and use should be adjusted.

The existence of these discrepancies reveals that there is uncertainty in estimating GDP. The analysis performed is expected to help reduce uncertainty, but the fact that assessments and automatic adjustments are needed to eliminate discrepancies mean that uncertainty remains in the balanced results.

Below is a summary of the total discrepancy between supply and use in three stages during the production process: Before the analysis, halfway through the analysis and before RAS-process. (RAS is an iterative proportional adaptation process to gradually reduce the product groups discrepancies to zero, as well as the overall discrepancy. It is a standard procedure within the national accounts to deal with the minor discrepancies remaining after major discrepancies have been analysed and corrected.)

Total discrepancy, supply minus use, as a percentage of GDP, current and previous year's price

	2011		2012		2013		2014		2015		2016	
Before analysis	2,1	2,2	2,0	1,8	-0,2	0,1	0,6	0,5	-0,2	-0,5	0,0	-0,1
Halfway	0,4	0,3	0,2	0,4	-0,4	-0,5	0,6	0,5	0,0	0,3	-0,1	-0,2
Before RAS	0,1	0,1	0,0	0,0	-0,2	-0,3	0,0	-0,1	-0,2	-0,2	0,0	0,0

Example S.13.1-2 Industrial producer prices 2017, Spain, [ESS-MH]

[This example illustrates a comprehensive assessment of overall accuracy.]

Overall accuracy

The accuracy is tackled at national, Community and market (euro and non-euro zone) levels by eliminating non-sampling errors as much as possible and studying and analysing revisions.

The main sources of error are non-response and overcoverage. There is no evidence that the response rate is distributed in a way that generates a bias in the index or its evolution. In addition, steps are taken to improve the updating of the sample, replacing, in a more agile way, units that have been deleted from the sample for new ones.

Due to the timeliness of the survey there is a non-response rate of 7% at the time of the initial release. The questionnaires keep being required three months after the end of the reference month. This way new questionnaires are recorded after the first publication of the results and the non-response rate decreases. Another consequence is that the published results are updated monthly including both new and edited data.

Since the cut-off sampling is used, best sampling method to be used taking into account the skewness of the distribution of the Industrial businesses size in Spain, the estimation is biased, although there is no information about how much.

Data editing consists on several phases throughout the process. The first one is micro editing and takes place during the data collection. The e-questionnaire contains workflows and several hard and soft edits. It makes possible that the data are cleaned by the respondent and there is no need to recontact. Finally the macro editing phase occurs, for the purpose of checking the results to be published.

During the whole data collection process the response rate is checked and attention is paid to get the data of the influential units.

S.13.1A FURTHER GUIDELINES FOR PROBABILITY SURVEY

For a probability survey the most straight-forward approach is to list the standard sources of error and point out which ones are significant for the survey at hand. For each of these the significance should be assessed according to the best knowledge of the producer, expressed in quantitative terms where possible.

S.13.1A EXAMPLES FOR PROBABILITY SURVEY

Example S.13.1A-1 Turnover in Industry, Hungarian Central Statistical Office ESS-MH

[This example illustrates a brief assessment of overall accuracy of a probability survey.]

Sampling error is available for the gross output value of industrial activities, without value added tax, including price subsidies at this moment. It is calculated for the strata and for some aggregates every month.

Although the response rate is fairly good the non-response is the main source of non-sampling errors. There are some enterprises which send late data, but their questionnaires can be taken into consideration over the revisions. Moreover, data are revised using SBS data, so the final version of the figures of industrial production is closed the one and half years after the beginning of the data collection.

The undercoverage of the survey population is rather small because the online connection between our Business Register and Registry Court assures a reliable survey frame. A good value about the rate of overcoverage can be calculated on the basis of Business Register. The rate of other errors (for example classification error) is noted over the data collection, and the statistician can list these cases and can determine the number of that.

S.13.1B FURTHER BACKGROUND AND GUIDELINES FOR NON-PROBABILITY SURVEY

Although probability sampling is a widely accepted norm for survey design, there are sometimes reasons for adopting non-probability designs. One reason is that the bias-protecting property of a probability design is in practice destroyed by high non-response or (less commonly) poor coverage. For example, in Household Budget Surveys, the willingness of a household to participate is often very low. Then a well-designed quota sample could, in the end, be more representative than a probability sample with, say, less than 50 % response rate.

In other cases, a probability sample is very difficult to perform in practice. This is often the case in price surveys, where a representative item (also called product offer in a CPI) is to be followed over time. In producer or construction price surveys the priced item is often selected on the basis of its perceived representativeness for a larger sales value. The degree of homogeneity of these sales values with regard to price changes is then at issue, as well as any systematic effects of the sampling mechanism. Although representativeness cannot be claimed in each case, it is often assumed and possible that the selection mechanism does not result in large aggregate biases.

Even though there is no theory to support quantitative error estimates for non-probability surveys, the representativeness of the sample should be assessed and the most important error components for the main estimates highlighted.

A possible starting point for an assessment is that, even for these surveys, there is a random sampling error. There may also be a bias that is due to the lack of full representativeness of the sample and this may well be larger than the random error.

The random component could be estimated, based on an assumption that the sample is effectively random, by using the same variance formulas that are used for a probability sample that reflects the assumed random structure of the sample. This could be considered as setting up a probability model for the sample. It is not a common practice today.

For assessing the possible bias, the key issue is the *representativeness* of the final sample. With a probability sample, representativeness is guaranteed by the probability sampling design (however, subject to coverage imperfection and nonresponse). Without this safeguard there must be some other mechanism that can provide assurance of representativeness. This may be in the form of a statistical model appropriate to the subject-matter field.

A common, implicit, model used in economic statistics is to assume that estimates of change are more accurate than estimates of level. The underlying basis for this assumption is that of a uniformity of market developments with respect to growth or price change. Obviously, such assumptions are imperfect and need to be evaluated where possible. Sometimes evaluations made in one country have some credibility also for other similarly developed countries and references to such evaluations in the same subject matter field are therefore useful. Still, evaluations are not common.

For quota surveys, representativeness depends critically on the effectiveness of the quota variables for capturing systematic variations in the target population. The inclusion mechanism in quota surveys (which is due both to the selection made by the interviewer and the acceptance of participation by the respondent) has to be uncorrelated with the target variables of the survey for a bias not to exist.

S.13.1C FURTHER BACKGROUND AND GUIDELINES FOR CENSUS SURVEY

The objective of a census survey is to collect data from all units for a specified target population. Three important categories of census survey are:

- population (and housing) census - the units are housing units, households and persons;
- economic census - the units are enterprises and local units (a producing unit of an enterprise with a physical address) or other intermediate units (kind-of-activity units, local kind of activity units.)
- agricultural census - the units can be of two kinds – agricultural businesses (farms) and/or land-based units.

For census surveys, there are, by definition, no sampling errors. Other survey errors are present as in sample surveys. However, the error profile of a census may be very different from that of a sample survey and may vary greatly depending on the type of census and type of approach used, especially since some countries use direct data collection whereas others use administrative data. This affects the relative emphasis that should be put in the report. For example, an often-mentioned error in population censuses is the so-called *over/undercount*. This type of error should be treated as a coverage error, as further discussed below.

A population census collects data for a large number of variables and there are often unique aspects of accuracy for each of them. Example 13.1C-1 shows how this can be treated.

The following quality aspects are of special importance for censuses that are based on direct data collection by extensive field work but coverage issues also occur for censuses based on administrative data.

Undercoverage and overcoverage

Undercoverage and overcoverage are also referred to as *undercount* and *overcount* or *double count* in the census context. The report should assess this potential source of error, i.e., that field procedures do not reach all target units or that they reach them twice.

A special, deliberate, case of not covering all units arises in the context of a *cut-off threshold*, which should be reported.

Measurement and nonresponse errors

Measurement and nonresponse errors may well be important. The same assessment and reporting principles apply as for sample surveys.

Processing errors

Processing errors in the form of data entry or coding errors can be of great importance in a census. Data entry errors may occur when the information is provided by respondents on paper and data are captured either manually or through an optical reading device. Coding is a further source of error. Over-editing is another issue. All these should be discussed.

Summary

At a minimum the most important error risks for the main estimates should be highlighted and the relative importance of coverage, nonresponse, measurement and processing errors should be stated. If appropriate, the error components can be described in more detail.

S.13.1C EXAMPLE FOR CENSUS SURVEY

Example S.13.1C-1 Population census 2011

INSEE, France, [ESS-MH]

[The example gives qualitative assessments for eight variables. In the actual report, 37 variables are assessed as this level of detail.]

14.1.1. Overall accuracy - Usual residence

There are no particular reasons for data unreliability for this topic

14.1.2. Overall accuracy - Sex

There are no particular reasons for data unreliability for this topic

14.1.3. Overall accuracy - Age

There are no particular reasons for data unreliability for this topic

14.1.4. Overall accuracy - Marital status

In some cases, the answers of the questioned persons may differ from the real legal situation. For example, a person who lives separately from their spouse but who is not yet divorced, and therefore legally married, may hesitate between "single", "married" and "divorced". Similarly, persons living together or in a registered partnership (PACS) may hesitate to declare themselves "married".

14.1.5. Overall accuracy - Family status

Added inaccuracy is linked to the complementary processing sampling rate: the results in municipalities with a population of 2,000 or less may be fragile.

14.1.6. Overall accuracy - Household status

Added inaccuracy is linked to the complementary processing sampling rate: the results in municipalities with a population of 2,000 or less may be fragile.

14.1.7. Overall accuracy - Current activity status

Unemployment is modelled to correspond to the ILO concept

European Regulation (EC) No 1201/2009 on population and housing censuses requires measurement of unemployment as defined by the ILO (International Labour Office). Because the French census is declarative, a model has been developed to approximate declared unemployment more closely to the ILO definition. The model is based on the Continuous Labour Force Survey (LFS), the reference in France for measuring unemployment as defined by the ILO, and which also provides the spontaneous declarations of those surveyed. Based on this model, probabilities for the conversion from declared unemployment to ILO unemployment have been determined by region and by sex. Additional information was also used, such as socio-professional category. Ultimately, unemployment modelled in this way approaches the ILO definition, without however claiming to reflect the full complexity of the ILO system.

Consequently, detailed data on current activity status should be considered as unreliable. Current activity status is reliable for region and sex level.

14.1.8. Overall accuracy - Occupation

The coding stage can generate some defects in terms of quality.

In census questionnaires persons fill out a title for their occupation. Several questions enable the coding of the occupation to be fine-tuned: establishment where the occupation is exercised, own-account or salaried worker, number of employees for own-account workers, function and occupational position of employees.

Based on these titles and additional information, the occupation and socio-professional category are then coded using automatic coding software. About one quarter of cases cannot be coded automatically: the software does not recognize the appellation of the occupation or finds inconsistencies between this appellation and the additional information. Coding is then done "manually" by agents who assign a code based on all the available information.

The French census codes occupations to the 2003 Nomenclature of Occupations and Socio-Professional Categories (PCS).

In most cases, occupations can also be coded in ISCO nomenclature. However, there is a lack of information on supervision (the fact of supervising another person's work).

This information thus had to be imputed in order to estimate persons classified in mode 1 of the ISCO nomenclature (for this reason, detailed data should be considered as unreliable).

For unemployed persons, the information on the previous occupation is very succinct and required estimations in order to switch to ISCO nomenclature.

This estimation was compared to the Continuous Labour Force Survey data to check its quality.

Added inaccuracy is linked to the complementary processing sampling rate: the results in municipalities with a population of 2,000 or less may be fragile...

S.13.1D FURTHER GUIDELINES FOR ADMINISTRATIVE DATA PROCESS

Processes using public sector accounting data

The main issues regarding accuracy should be presented. See Examples S.13-1 and S.13-2 above. Coverage issues are often critical. Other issues can be related to valuation (in case of monetary data) or periodisation.

Processes using registers

This type of process refers to the case where the statistical outputs are derived directly from a register, without additional data sources.

There is no sampling of the data obtained, thus sampling errors do not exist. The pertinent errors are as follows.

Coverage errors

Over- and undercoverage of the target population, using the register definition, should be assessed and reported. Lags in entering information into registers are crucial for understanding the coverage properties of a register. Evaluation approaches regarding these errors have much in common with those mentioned in connection with census surveys.

Nonresponse errors

Missing data items should be assessed and reported. The nature of missing variables in an administrative data set is similar to item non-response in surveys. (Unit nonresponse is normally not distinct from undercoverage.)

Measurement errors

For various reasons, a record may have erroneous values for one or more variables. The cause of the error may be that the value was erroneously provided in the first place or that a later change in the variable has not yet been recorded in the register. An example of the latter error would be that a business that should submit its quarterly VAT return fails to do that in time, which leads to an erroneous value for its turnover in that period.

The lag structure associated with register updating can be analysed in order to throw light on the latter aspect.

Processing errors

When registers are maintained by external agencies, two levels of data processing can be identified:

- processing performed at level of the data provider; and
- processing carried out by the NSA during data checking, integration of the data with other sources, and deriving statistics.

The quality report should provide a summary of the processing undertaken by the provider with the aim of indicating potential sources of error as well as covering errors occurring at the NSA level.

Differences in concepts

If the target statistical concepts differ from the register concepts, the effect on outputs of differences should be assessed for key indicators, quantitatively to the extent possible.

Sometimes, a model-based adjustment is used to bridge the differences between target and register concepts. The model should then be assessed with respect to how well the model assumptions are met.

Administrative event-reporting systems

The quality of data from an event-reporting system depends first and foremost on the completeness of the reporting system. The *rate of unrecorded events* is a key quality factor, although sometimes difficult to estimate. It is a special type of undercoverage error.

Errors in the classifying variables (such as type of crime, type of accident, or type of goods) can best be regarded as a processing error. Approaches to monitoring these errors are normally domain related. For example, in crime statistics there is an intricate system for coding main crimes and related crimes, counting crimes, etc., that depends upon principles and practices in the area of criminology. A similar situation occurs for cause of death statistics, where the coding procedures depend upon medical principles and practices.

S.13.1E FURTHER GUIDELINES FOR MULTISOURCE PROCESS

When dealing with output from a statistical process involving multiple data sources, the overall assessment has to take all possible error components into account, as well as possible coherence problems in combining data from the sources (see also S.15.4).

Error models for multisource processes are beginning to appear in the literature. It can be expected that such models have to limit themselves to certain special cases, for example according to the classification in Eurostat (2018a) above. De Waal et al (2019) present one such model dealing with a split-population situation with classification errors.

In most cases an assessment of overall accuracy will be a key part of the report. However, it is difficult to assess the accuracy because of the many sampling and non-sampling errors involved in the different sources being considered. An informed judgement as to which are the most significant errors or error risks is necessary and these errors should then be explained in the greatest detail possible. Where it is possible to describe major error sources in quantitative terms, this should be done.

S.13.1E EXAMPLE FOR MULTISOURCE PROCESS

Example S.13.1E-1 Remuneration of civil servants - key indicators (Art. 65)

Eurostat 2015 [ESS-MH]

[There are few examples of producer reports for multisource or administrative data processes in the ESS. Here, the relative accuracy of different estimates is addressed in qualitative terms.]

The precision of specific indicators is considered to increase with the level of aggregation. This means that the global specific indicator (EU average) will be more reliable, or precise, than the indicator for an individual Member State. Similarly, the overall average indicator for an individual Member State will be more reliable, or precise, than the indicator for "administrator-equivalents" which is one of the defined categories. Similarly, the indicator for "administrator-equivalents" will be more reliable than the indicator for individual occupation grades within that category.

The input data into the specific indicator calculation process comes from several sources, specifically, from special price surveys of remuneration of national civil servants, harmonised index of consumer prices, gross domestic product. This makes it impossible to calculate any meaningful, numerical measure of error margins for these indicators.

The precision of the Joint Index is considered to increase with the level of aggregation. This means that the index at the level of total household consumption will be more reliable, or precise, than the index for "food and non-alcoholic beverages" which is one of the sub-aggregates of final household consumption (12 main COICOP groups). Similarly, the index for "food and non-alcoholic beverages" will be more reliable than the index for "bread and cereals" which is one of the analytical categories within that COICOP group.

The input data into the index calculation process comes from several sources, specifically, from national data collection on consumer goods and services, internal Commission sources on staff numbers, and special surveys of household consumption expenditure. This makes it impossible to calculate any meaningful, numerical measure of error margins

S.13.1F FURTHER GUIDELINES FOR MACRO-AGGREGATE COMPILATION PROCESS

For macro-aggregates there is often a particular error structure defined in manuals or special legislation. The best approach may then be to report all errors under overall accuracy, where a general statement on the extent to which it has been possible to follow these recommendations could be given, covering, for example:

- whether A or B methods been followed;
- how different first estimates are from revised and final estimates;
- whether the informal economy has been included appropriately;
- the extent to which large parts of the aggregate have been estimated directly, as opposed to based on imputation.

In practice, the accuracy of a macro-aggregate is determined by the completeness, coherence and accuracy of all the components that enter into the aggregates. There is, however, no good approach for summarising these aspects.

Example S.13-1-1 above describes a producer report for National Accounts which focuses on the model assumptions used when aggregating primary statistics into its conceptual framework. The discrepancies resulting from estimation of GDP from the production and expenditure approaches are used.

Biemer et al (2017) propose a decomposition of error sources for the National Accounts into six components([10]):

- input data source error
- compilation error (data modelling)
- compilation error (processing)
- deflation/reflation error
- balancing error and
- revision error

Specialised guidelines for producer reporting are often produced for big compilation systems such as the National Accounts, Agricultural Accounts etc.

A detailed account of the accuracy problems according to either specialised guidelines or a structure

([10])This decomposition was used by Statistics Sweden 2012-17 for the purpose of the external evaluations of Accuracy. It is, however, no longer being applied, as Statistics Sweden has modified the approach in 2019. The National Accounts, as other statistical products, will now be evaluated based on the structure given in the Swedish quality concept i.e. Overall Accuracy, Sources of uncertainty (Frame coverage, Sampling, Non-response, Measurement, Data processing, Model assumptions), and Preliminary statistics compared to final statistics.

such as described above should be given.

There is a special exercise carried out at the European level regarding Quality assurance of statistics underlying the Macroeconomic Imbalances Procedure (MIP) Scoreboard (https://www.cmfb.org/main-topics/mip-quality). Annually, the quality of such economic statistics is assessed in special reports.

S.13.2 Sampling error

<table>
<tr><th>SIMS</th><th>Concept name</th><th>Definition</th><th>Guidelines</th></tr>
<tr><td rowspan="2">S.13.2</td><td>Sampling error</td><td>That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a subset of the population is enumerated.</td><td rowspan="3">State whether sampling error is relevant.
If probability sampling is used:
• for user reports, provide the range of variation of the A1 indicator among key variables at user report level of detail;
• for producer reports, provide the range of variation of the A1 indicator among key variables at producer report level of detail;
• indicate the impact of sampling error on the overall accuracy of the results;
• state how the calculation of sampling error is affected by adjustments for nonresponse, misclassifications and other sources of uncertainty, such as outlier treatment.
If non-probability sampling is used, provide an assessment of representativeness, a motivation for the invoked model for estimation and risk of sampling bias.
European level
If probability sampling is used:
• present sampling errors for key estimates across countries;
• indicate which country to country differences are significant and which are not;
• for a repetitive survey, describe at least broadly the trends in sampling error over time
• provide sampling errors for ESS level estimates.</td></tr>
<tr><td>A1. Sampling error indicators (U)</td><td>Measures of the random variation of an estimator due to sampling, at a level of detail appropriate for user reports.</td></tr>
<tr><td>S.13.2. 1</td><td>A1. Sampling error indicators (P)</td><td>Measures of the random variation of an estimator due to sampling, at a level of detail appropriate for producer reports.</td></tr>
</table>

S.13.2A FURTHER GUIDELINES FOR PROBABILITY SURVEY

For probability sampling, sampling theory provides techniques for the estimation of the expected value and variance of specific indicators over all possible samples. Therefore, the random variation due to sampling can be calculated. Furthermore, sampling biases are normally zero or negligible so that the variance can be taken to represent total sampling error (subject to complete response, as further discussed below in connection with nonresponse errors). Techniques for estimating variances can be found in textbooks or other scientific reports. Software is often available for doing the calculations.

The variability of an estimator around its expected value may be expressed by its variance, standard error, coefficient of variation (CV), or confidence interval. In this context, there are several presentational devices that can be used. (The standard sampling error indicator A1 is defined and explained in Supplementary Document C.)

The *standard error* is the square root of the variance of an estimator. Usually the standard error is not suitable for use by itself since its interpretation is not obvious to the average user.

The *coefficient of variation (CV)* is defined as the standard error divided by the expected value of the estimator. It is the standard error in relative (percentage) terms. It is the most suitable sampling error statistic for quantitative variables with large positive values, which are common in economic statistics. It is not recommended for proportions, for estimates that are expressed in percentage terms, or for changes, where it could easily be misunderstood. It is also not usable for estimates that can take on negative values such as profits, the net export/import value etc.

The *confidence interval* is defined as an interval that covers the true value with a certain probability. In most cases where it is reasonable to assume the estimator follows a normal distribution, the interval that results from taking ± 1.96 * estimated standard error from the point estimate results in a 95 % confidence interval. Taking instead ± 1.96 * estimated CV expresses the interval in percentage terms.

For key indicators the sampling error should be expressed as a confidence interval, since this is the most rigorous and clear way of demonstrating sampling variability.

For large sets of estimates in tables, confidence intervals often lead to a rather clumsy presentation and CVs, or CV intervals, are more natural to use. A CV interval indicates that the CV is in a certain range (say 5-10 %) of the estimate. Different ranges can be denoted with letters (e.g., A≤2 %, B= 2-5 %, C= 5-10%, D≥10 %). Use of ranges is also appropriate because estimates of sampling variability are not exact.

For business surveys, especially where large positive numbers (of production, turnover, export, etc.,) are targeted, estimated CVs are normally the best way to express sampling error. The size of the sampling error relates to the sample size for the domain to which the estimates relate, so, for a large table with many cells that would be overburdened with an estimated CV in every cell, they are instead best presented in a separate table.

Especially in business surveys, *outliers,* i.e., sample units with extreme values, can greatly influence the estimates and lead to major sampling errors. Outliers can be incorrect or correct but still considered non-representative for their stratum and their treatment involves judgement by the producer. The producer report should state clearly, whether, how and why outliers have received special treatment in the estimation process.

If imputations are used it should be stated how the calculation of sampling errors is affected.

In household surveys, results are often presented as proportions or percentages and it is not usually appropriate to present random sampling errors in the form of CVs. Confidence intervals are a better choice. It is sometimes possible to present simplified indicators of sampling errors, where a certain range of estimated proportions are associated with a certain level of sampling error according to the well-known formula variance = p(1-p)/n, where p is the proportion and n the sample size.

For example, in an opinion poll with a sample of n=1000 persons the estimate of support for a political party is 20 %. The variance of this estimate is V=0.2x0.8/1000=0.00016. The standard error is √V=√0.00016=0.01265=1.265 %. The confidence interval can thus be expressed as 20 % ± 1.96x1.265 % or 20% ± 2.48 % and the corresponding confidence interval approximately goes from 17.5 % to 22.5 %.

Where CV thresholds are included in regulations, a comparison between estimated CVs and the relevant thresholds should be included.

Some further technical points concerning the presentation of sampling errors are:

- Nonresponse should be taken into account, i.e., the sample size should be the effective sample, i.e., after deduction of the number of non-respondents.
- The original stratification should be applied, i.e., the sampling error should not be artificially reduced by first moving outliers to a special stratum.
- Variance estimation should be in accordance with the actual sampling and estimation method applied.

Sampling errors for estimates of change are of great importance, although sometimes more difficult to calculate due to non-independence between samples in adjacent periods([11]). An assumption of independence normally leads to an overestimate of the sampling error for a change (since the covariance term is actually negative). If this is the case a statement like "*The sampling error for the change between Q3 2017 and Q3 2018 is at most X"* is valid, where X is calculated under an assumption of independence.

Cut-off sampling

One type of sampling that is frequently applied in economic surveys and therefore needs special attention is the use of a *cut-off* threshold. Units (businesses, enterprises, establishments) below a certain size threshold, although belonging to the ideal population for important users, are not sampled at all. Such a procedure is referred to as *cut-off sampling*.

Technically this situation is similar to undercoverage (further discussed below under *coverage errors*) but with the distinctive feature that the cut-off is intentional and there is information about the excluded units, that enables some sort of assessment of contribution of the non-sampled portion of the population. Two of the reasons for a cut-off threshold are reduction of the response burden for small units and a relatively small contribution they make to the total estimate and to the total error (sampling and non-sampling).

It is recommended that the uncovered part below the threshold be treated as a limitation in the target population of the survey and thus as a relevance issue rather than as a sampling or a coverage issue.

Two different cut-off scenarios are possible. Either there is a census (take-all) above the threshold or there is a probability sample. In the first case the survey is a census survey with no sampling error. In the second case we have a probability survey, where sampling error is calculated on the basis of the sampling procedure used for the actual sampled population.

A cut-off threshold is often combined with probability sampling above the threshold and in this case can be called ***sampling with cut-off*** as opposed to ***census with cut-off*** where all units above the threshold are included.

If a model is employed for estimating the cut-off portion of the population, this should be stated and the validity of the model assessed, if possible in quantitative terms.

Summary

At a minimum, sampling errors should be indicated directly in tables and databases, using coefficients of variation, standard errors or confidence intervals. The suitability of each concept depends on the type of variable. Margins of error can be presented without going into detail about the statistical techniques used.

If more detail is appropriate, estimation and variance formulas should be provided together with an analysis of the adequacy of the sampling design for optimising accuracy and cost. Non-standard estimation procedures (such as those that are based on judgement), for example for outliers, should be highlighted. Sampling errors for domains of estimation (table cells) and, if possible, for estimates of change should be provided. This could be done in the form of approximate error margins instead of exact numbers.

([11]) Nordberg (2001) and Wood (2008) discuss this problem at a fairly general and technical level.

S.13.2A EXAMPLES FOR PROBABILITY SURVEY

Example S.13.2A-1: Presentation of CVs and design effect (Arnež et al., 2008, p. 12-14)

[This example is suitable for a producer report.]

Sampling errors can be expressed in different ways: in absolute form (se), relative form as a coefficient of variation (cv), or with confidence interval (estimation ± 1,96*se). The most frequently used is the coefficient of variation, which indicates the degree of precision to which the estimate ($\hat{x}$) is compared:

$cv(x) = \frac{se(\hat{x})}{\hat{x}}$ *100

If the coefficient of variation is small, this means small sampling variability with regard to the estimate. The coefficient of variation depends on the size of estimate, the number of units in the sample which are subject to the calculation of estimate, distribution of the sample for such variable, and on the application of auxiliary information in the estimation procedure.

The quality of sample designs is measured also by means of a design effect (deff). This is a general measure to compare the variance of simple random sample (SRS) with the variance of complex samples of equal size, where two variances are compared for the same variable.

deff= var($\hat{x}$)/var$_{SRS}$ ($\hat{x}$)

In general, stratification in comparison to SRS sampling decreases, while multi-stage sampling increases the sampling error.

deft = Sqrt (deff) means the factor which widens or narrows the confidence interval due to the sampling design in comparison to the sampling error which would result from the SRS sample. The following table (truncated from original text) provides examples of estimates and sampling errors for such estimates.

Estimates and errors of estimates for allocated assets (including own production), HBS 2004

Code	Description	Average per household	cv (%)	deff
	Allocated assets	4 118 459	1,4	1,4
	Consumption expenditure	3 627 955	1,3	1,4
.01	Food and non-alcoholic beverages	689 466	1,1	1,3
.02	Alcoholic beverages and tobacco	101 406	2,5	1,3
.03	Clothing and footwear	292 196	2,3	1,4
….	…	…	…	…

Example S.13.2A-2 Labour Force Survey, 2017, DESTATIS, Germany

[This example gives numbers only and is suitable for a less detailed user report]

	Number employed persons	Employment rate as a percentage of population	Number part-time employed persons	Number of unemployed persons	Unemployment rate as a percentage of labour force	Youth unemployment rate as percentage of labour force	Average actual hours work per week
Age group	20-64	20-64	20-64	15-74	15-74	15-24	20-64
CV	0.25	0.17	0.60	1.94	1.93	4.27	0.14
SE	00197	0.14	63682	30690	0.07	0.30	0.05
CI (95%)	39542484-39935256	79.2 - 79.74	0414814 - 10664448	1524294 - 1644597	3.51 - 3.78	6.39 - 7.56	35.84 - 36.03

Example S.13.*2A-3* Crop production, 2016 France [ESS-MH]

[This example gives numbers only and is suitable for a less ambitious user report]

Coefficient of variation (CV) for the area (at MS level)

	Arable lands survey	*FSS survey on vegetable and fruit areas*
Cereals for the production of grain (%)	between 0.4 and 10	
Dried pulses and protein crops (%)	around 3	
Root crops (in %)	around 10	
Oilseeds (in %)	around 1	
Plants harvested green from arable land (%)	0.5	
Total vegetables, melons and strawberries (%)		2,7
Fruit trees (%)		between 0.4 and 2
Nut trees (%)		1
Citrus fruit trees (%)		2

Example S.13.2A-4 Farm Structure Survey, INE, Spain, (ESS-MH, 2016)

[This example relates sampling errors to precision requirements]

The relative standard error of the amount of "breeding sows" errors for ES24 is equal to 7,45% and for ES51 is 5,5%.

The sample was obtained having into account the variable "breeding sows" in the precision requirements for each considered NUTS. The variability of that variable and the changes in time have originated that the final error exceed 5% in the resultant sample.

For NUTS1= ES2, the relative standard error of the amount of breeding sows is equal to 5,96%. This result is due to the ES24, which is a part of ES2.

In the case of "Area of citrus plantation" in NUTS2= ES62, the relative standard error is equal to 5,48%.

As the preceding case, the sample was obtained taking into account such variable in the precision requirements. Changes in time have originated that in the resultant sample the RSE exceed the 5%.

S.13.2B FURTHER GUIDELINES FOR NON-PROBABILITY SURVEY

When *non-probability sampling* is applied, random errors cannot be estimated without reference to a model of some kind. Furthermore, sampling biases may well be significant and need to be assessed as well. There are many types of non-probability sampling, each of which require their own evaluation depending on the situation at hand.

For some forms of non-probability sampling, for example those applied for price indexes, it may be reasonable to apply standard error estimators as if the sample were effectively random, or to use some other model-based approach. This approach has, however, to be complemented with a discussion of possible sampling bias and of possible limitations in the sampling model used. For example, it can often be stated whether (and why) the estimates of sampling error thus derived are "conservative" (i.e., upper limits) relative to the real errors.

The extent to which the survey approach conforms with recommended international practice should be stated. Is the sampling method used common international practice for this kind of survey? Where the approach differs from recommended practice in international manuals or the like, the reasons for this should be stated and a more detailed assessment of the method used is required.

A key issue with non-probability surveys is their representativeness. An assessment of representativity with respect to the estimates of the survey should be given referring to the type of inference that the survey data are used for and to supporting models or assumptions.

S.13.2B EXAMPLE FOR NON-PROBABILITY SURVEY

Example S.13.2B-1 Household Budget Survey (EVS, a quota survey)

DESTATIS, Germany, (ESS-MH, 2015)

[This example indicates how precision measures for a non-probability survey are computed. When it is said that "results are projected from the sample to the household reference population with known levels of precision", this means that a model has been applied, in which it is assumed that the sample is effectively random. It is recommended that a reference is made to a document, where this model and its validity is further described.]

Like in any sample survey, the statistics generated from the EVS data may be liable to errors which are inherent in the survey method used. Usually, a sample of households is selected in a way that the probability of a household being selected is known. Though EVS is a quota sample, the results are projected from the sample to the household reference population with known levels of precision, i.e. standard errors and confidence intervals for survey estimates can be constructed.

The EVS data are weighted. Sample weights are needed to correct for imperfections in the sample that might lead to bias and also to rectify other departures between the sample and the reference population. The design weights are calculated for each sampled household as the inverse of its probability of selection as part of the sample.

S.13.2C FURTHER GUIDELINES FOR OTHER TYPE OF STATISTICAL PROCESS

If there is an important sampling component in a multisource process, the sampling errors should be reported according to the above recommendations at a level of detail appropriate for their importance to the process as a whole.

Macro-aggregates are compiled from several statistical sources. If any of these are sample surveys an assessment of the significance of the sampling errors should be included.

S.13.2C EXAMPLE FOR OTHER TYPE OF STATISTICAL PROCESS

Example S.13.2C-1 QUALITY DECLARATION. National Accounts, quarterly and annual estimates Statistics Sweden Version 1 2018-09-13

[The example gives a brief assessment of the significance of sampling]

In the national accounts, for calculations in current prices, a large part of the statistics is based on surveys that do not use sampling. In cases where sampling is used (especially for quarterly statistics), access to information for sample allocation is limited, and some information is not relevant. This affects the efficiency of the sample. The impact on accuracy due to sampling is judged to be moderate for the quarterly estimates and small for the annual estimates.

Fixed price calculations include price indices estimated through sampling. The impact on the accuracy from sampling to the fixed prices estimates is judged to be moderate.

S.13.3 Non-sampling error

SIMS	Concept name	Definition	Guidelines
S.13.3	Non-sampling error	Error in survey estimates which cannot be attributed to sampling fluctuations.	Summarise the most important aspects of coverage, measurement, non-response, processing and model assumption errors. Discuss the corresponding bias risks and actions undertaken to reduce them. *European level* Provide a summary of the above across countries.
	A4. Unit non-response - rate (U)	The ratio of the number of units with no information or not usable information to the total number of in-scope (eligible) units, at a level of detail appropriate for a user report.	*For probability and census surveys*: • Report A4: unit non-response rates. *For repetitive surveys* • also briefly describe the trend for A4. (Note: for producer reports A4 is reported in S.13.3.3.1.)
	A5. Item non-response - rate (U)	The ratio of the in-scope (eligible) units that have not responded to a particular item and the in-scope units that are required to respond to that particular item, at a level of detail appropriate for a user report.	Report A5: item non-response rates for key variables. (Note: for producer reports A5 is reported in S.13.3.3.2.)

S.13.3 FURTHER GUIDELINES

Errors that play a significant role for the interpretation and use of the statistical process should be described. When this is the case the following guidelines are applicable. (For definitions and further details on the respective errors, see subsections S13.3.1-5 below.)

Non-sampling errors as a concept is mainly applicable to surveys. For other processes, some of the guidelines below are still applicable to a smaller or larger extent.

Coverage error

Coverage error (or frame error) is due to divergences between the survey population and the target population.

Two types of coverage error are distinguished:

- *Undercoverage*: there are target population units that are not accessible via the survey frame (e.g., persons without a phone will not be listed in a telephone catalogue);
- *Overcoverage*: there are units accessible via the survey frame that do not belong to the target population (e.g., deceased persons still listed in a telephone catalogue); this includes multiple listings (duplication) where target population units are present more than once in the survey frame (e.g., persons with two or more telephone connections).

Whereas undercoverage is often the more serious problem it is also more difficult to quantify and then has to be reported in qualitative terms only. Overcoverage is often revealed in the sampling process, for example when a first contact with a sampling unit fails because it no longer alive or active. Overcoverage can be reported quantitatively together with non-response errors.

See Eurostat (2018a) for more information on quantitative assessment of coverage errors.

Measurement error

A *measurement error* is an error that occurs during data collection and causes the recorded values of variables to be different from the true ones. The causes are commonly categorized as:

- *Survey instrument*: the form, questionnaire or measuring device used for data collection may lead to the recording of wrong values. Reasons may be the wording of the questions in the questionnaire, the order or context in which the questions are presented.
- *Respondent*: respondents may, consciously or unconsciously, give erroneous data due to confusion, ignorance, carelessness or dishonesty;
- *Interviewer*, due to inadequate training or surveillance, interviewer may inappropriately influence the answers

In other statistical processes than surveys a measurement error can be defined as an error in recorded values for individual units in the data sets used.

Where significant, measurement errors and actions taken to reduce them should be reported according to the producer's best knowledge.

Response error

A *nonresponse error* is relevant and significant mainly for probability surveys and censuses.

There are two types of nonresponse:

- *unit nonresponse,* which occurs when no data are collected from a unit in the sample; and
- *item nonresponse* which occurs when values for some but not all survey data items (variables) are obtained from a unit.

Accordingly, response (and nonresponse) is measured in terms of *response rates* of two kinds, broadly defined as follows:

- *unit response rate*: the ratio of the number of respondents (i.e., units for which data for some or all data items have been collected) to the total number of eligible units in the sample;

- *item response rate*: the ratio of the number of units that have provided data for a given data item to the in-scope units that are required to respond to that particular item.

Unit nonresponse rates should always be reported. Where significant, item nonresponse rates for important data items should also be reported.

Processing error

A processing error is defined as an error arising from the faulty implementation of correctly planned implementation methods. In practice, such errors are of mainly two kinds.

- Errors in microdata due to various problems in the transformation of data into computer readable formats.
- Mistakes in implementing procedures. Examples are programming mistakes, misprints in press releases etc. They are by definition unexpected and often only discovered after publication. Actions taken to eliminate or reduce the scope for mistakes should be reported where the risk is great. See also S.11 Quality Management and S.17 Revisions.

Model assumption error

Models are often applied in statistics. Sometimes the target of estimation relies on an abstract model defined by a subject matter discipline. In other cases, such as seasonal adjustment, the model is of a purely mathematical-statistical nature. Sometimes a model is applied in estimation in order to improve precision or to adjust for non-response. Where models play a significant role for interpreting the outputs, they should be described and their effects on the accuracy of results assessed.

S.13.3.1 Coverage error

SIMS	Concept name	Definition	Guidelines
S.13.3.1	Coverage error (P)	Divergence between the survey population and the target population.	Provide information on the frame and its sources. Provide an assessment, whenever possible quantitative, of overcoverage and undercoverage, including an evaluation of the bias risks associated with the latter. Describe actions taken for reduction of undercoverage and associated bias risks. *European level* Provide an overall picture of coverage across countries. This is often best done in the form of tables with the important coverage aspects, country by country.
S.13.3.1.1	A2. Overcoverage – rate (P)	The proportion of units accessible via the frame that do not belong to the target population.	*For probability surveys and censuses:* Report A2, Overcoverage – rate (for definition, see Supplementary Document C).
S.13.3.1.2	A3. Common units – proportion (P)	The proportion of units covered by both the survey and the administrative data in relation to the total number of units in the survey.	*For multiple source processes*, where one source is a survey and the other source(s) is (are) administrative: Report A3, Common units – proportion (for definition, see Supplementary Document C)

S.13.3.1A FURTHER GUIDELINES FOR PROBABILITY SURVEY

A *coverage error* (or *frame error*) is due to a divergence between the survey population and the target population resulting from an imperfect frame.

Two types of coverage error are distinguished:

- ***Undercoverage***: there are target population units that are not accessible via the survey frame (e.g., persons without a phone will not be listed in a telephone catalogue);
- ***Overcoverage***: there are units accessible via the frame that do not belong to the target population (e.g., deceased persons still listed in a telephone catalogue); this includes multiple listings (duplication) where target population units are present more than once in the frame survey (e.g., persons with two or more telephone connections).

Other sorts of frame deficiencies that can cause errors involve incorrect classification, contact and auxiliary information about the units included in the frame. Such deficiencies can also cause errors other than coverage errors. For example, wrong contact information (address, phone number) may result in nonresponse error, or if the size of a unit as recorded in the frame is smaller than its actual size, the sampling error may increase (sometimes dramatically for outliers).

Background information on the register or other frame sources should be provided including reference period, frequency and timing of frame updates, possible discrepancies between the survey frame and the target population. (This assists in understanding coverage errors and their effects.)

Overcoverage

Wrongly included units can be detected during the measurement process and are straightforward to handle in the estimation procedure. They result in increases in sampling error and survey costs.

Multiple listings, if detected, can be handled by statistical methods and also result in an increase of sampling error and cost but no significant biases. However, multiple listings of smaller units for which sampling rates are low are difficult to detect. If there is a significant risk of such errors, this should be reported.

Quantitative information on overcoverage including multiple listings is normally easy to obtain in sample surveys and censuses. This information should be included in the producer report in sufficient detail with respect to important sub-domains.

Undercoverage

Undercoverage cannot be detected in the measurement process and is the most serious type of coverage error. The resulting bias depends on the units outside the survey population but in the target population, and the differences between the characteristics of these units and the units in the survey population. Thus, a qualitative description of these units is a first step in assessing the under-coverage bias. Methods to detect undercoverage and assess its effects include, for example (i) when there is a time lag in registering frame units, a later frame version can provide information, and (ii) through comparisons with another frame or other external information. Where undercoverage is suspected to be significant, an assessment is always needed. As far as possible estimates of undercoverage (extent and effect) should be included in the producer report.

Undercoverage can be “defined away” by limiting the survey to what is covered by the survey frame. In this case, the coverage error is transformed into a relevance problem and should be treated under S.12.

Evaluation of coverage errors

Possible methods include the following.

Matching with a different register. The survey frame is matched with a control register that wholly or partly covers the same population as the frame. If the survey frame and control register are not both electronically stored, then matching can be done on a sample basis. If the control register is of superior quality, then errors in the frame can be directly assessed. Otherwise a reconciliation process, involving checking (a sample of) the non-matches is needed to determine the extent of errors in the survey frame.

Analysis of lag structure. Every frame for a repeating survey is updated with a certain lag according to the delay in registering births, deaths and changes of units. Thus, the frame will always, to a smaller or larger extent, have a less than perfect coverage at the time of use. The lag effect can be studied, for example, by matching two consecutive register versions and establishing which of the units in the latter version should, by definition, have been included in the former. Other approaches are also possible. Register errors can be studied in several consecutive versions. It may be possible to observe certain stability in error levels that can be assumed to continue into the future. The degree of under- or over-coverage as well as rates of change in contact data etc., can thereby be estimated. It is also possible to use this kind of information for a *model-based adjustment* of the estimates themselves.

It should also be noted that:

- overcoverage is best reported together with nonresponse in a coherent manner so that, for example, the treatment of units with unknown status is made clear. See S.13.3.3 on non-response;
- it is also possible to define rates of misclassification, incorrect contact details and multiple listings in straightforward ways. However, in most cases these indicators are not as important as the overcoverage rate;
- although the rate of undercoverage is the most important indicator it is not usually directly observable and thus not included in the set of standard quality and performance indicators.

Summary

Provide a full analysis of the coverage issues.

For undercoverage, provide an estimate based on a full evaluation, or a qualitative assessment based on the producer's best knowledge.

For overcoverage (including wrongly included units and multiple listings) provide quantitative estimates and their approximate effects on the main variables and subdomains.

S.13.3.1A EXAMPLES FOR PROBABILITY SURVEY

Example S.13.3.1A-1 Overcoverage errors (Slovenia: Standard Quality Report for the Monthly Survey on Turnover, New Orders and Value of Inventories in Industry, 2005) (Seljak & Katnič, 2006, p. 10)

[Shows the seasonal variation in overcoverage rates in two ways]

The table below shows data on inappropriate units in the sample, which is simultaneously the assessment of the share of overcoverage. *Inappropriate units* are those that we identify at the beginning of the year (January) when preparing the list of observation units They are included in the list although they do not belong there according to their activities. In the following months, the level of inappropriateness also takes account of the units that were appropriate when the selection for the survey was made, but then changed their activity or stopped operating during the year. The table presents the unweighted and weighted levels of overcoverage.

Weighted and unweighted levels of coverage

Level of over-coverage	Jan. 2005	Feb. 2005	Mar. 2005	Apr. 2005	May 2005	June 2005	July 2005	Aug. 2005	Sep. 2005	Oct. 2005	Nov. 2005	Dec. 2005	Average value
Unweighted	9.2%	9.3%	9.5%	9.3%	10.5%	10.6%	10.8%	11.3%	11.6%	11.9%	12.1%	12.5%	10.7%
Weighted	3.1%	3.2%	3.3%	3.4%	4.2%	4.3%	4.5%	4.8%	5.1%	5.5%	5.6%	5.9%	4.4%

Example S.13.3.1A-2 Coverage errors, Farm Structure Survey 2016, Statistics Poland, (ESS-MH)

[This example provides an estimate of undercoverage, presumably based on an analysis of the lag structure, and also estimates of multiple listings.]

Coverage error

1. Undercoverage errors

The under-coverage rate was estimated on the basis of an annual analysis of changes in the number of agricultural holdings in the frame. The analysis showed that every year about 3% of newly created agricultural holdings were not included in the surveys in given year.

2. Overcoverage errors

Over-coverage units were identified during data collection and appropriately coded. They were eliminated from the central set of data and the weights were corrected.

2.1 Multiple listings

Before the survey, 8.5% of potential duplicates were identified in the frame. They were eliminated during the data collection. The survey showed that approximately 1% of the real duplicates were in the frame.

Duplicates were treated as liquidated units therefore the weights were adjusted.

3. Misclassification errors

There were cases when out of date information on the agricultural area of the holding in the frame caused its classification to inadequate stratum. These farms were captured after the survey as outliers in a given stratum.

4. Contact errors

The questionnaire application allowed updating the address data. Approximately 7% of addresses and 15% of phone numbers were corrected.

5. Other relevant information, if any

Not available.

Overcoverage - rate

6.2%

Over-coverage rate was computed as the proportion of units from the sample which do not belong to the target population to the overall sample size.

S.13.3.1B FURTHER GUIDELINES FOR NON-PROBABILITY SURVEY

For a non-probability survey, the survey population and frame are often not well-defined. In this case, given there is no clear-cut distinction between a coverage error and a sampling error, it is more practical to treat them jointly as a problem of *representativeness,* as discussed in S.13.2.

In some cases, a probability design is used in a first sampling stage and a non-probability design in a second stage. In this case, first stage coverage errors should be reported according to S.13.3.1A.

S.13.3.1C FURTHER GUIDELINES FOR CENSUS SURVEY

For census surveys, coverage issues are, in principle, the same as for probability sampling surveys. However, in the census context, undercoverage and overcoverage are often referred to as *undercount* and *overcount* or *doublecount*. The producer report should assess this potential source of error, i.e., that field procedures do not enumerate all target units or result in multiple listings.

There is a fine distinction between coverage error and the definition of eligible units. For example, in a population census, the target population is often people who live permanently within the country's borders. Illegal immigrants may or may not be an undercoverage or a matter of relevance depending on the definition of target population adopted. Similarly, for an economic census, definitional issues regarding whether a business is active in a certain period may affect whether it is to be counted as part of undercoverage or not.

Coverage errors in censuses can, for example, be estimated through post-enumeration surveys based on a well-designed probability sample. Such surveys are not much used in Europe, however.

S.13.3.1C EXAMPLE FOR CENSUS SURVEY

Example S.13.3-2: Comparison of census undercount in US decennial censuses. Williams D., 2012, p. 10

[The American example provided here shows how an analysis of undercount could be presented.]

Table 2. Percentage Net Decennial Census Undercount by Race, as Estimated by Demographic Analysis, 1940 through 2000

	1940	1950	1960	1970	1980	1990	2000
Total population	5.4%	4.1%	3.1%	2.7%	1.2%	1.65%	0.12%
Black	8.4%	7.5%	6.6%	6.5%	4.5%	5.52%	2.78%
Non-Black	5.0%	3.8%	2.7%	2.2%	0.8%	1.08%	-0.29%

Sources: Estimates for 1940 through1980 are from J.G. Robinson, et al., "Estimates of Population Coverage in the 1990 United States Census Based on Demographic Analysis," *Journal of the American Statistical Association*, vol. 88 (September 1993), p. 1065, reprinted in U.S. Bureau of the Census, *Accuracy and Coverage Evaluation, Statement on the Feasibility of Using Statistical Methods to Improve the Accuracy of Census 2000*, June 2000 (unpublished document). Estimates for 1990 and 2000 are from U.S. Bureau of the Census, *Coverage Measurement from the Perspective of March 2001 Accuracy and Coverage Evaluation*, Census 2000 Topic Report no. 4 (Washington: U.S. Bureau of the Census, February 2004), p. 9.

Note: All estimates except one indicate net percentage undercounts of the total population or groups within the population. The exception, -0.29% for non-blacks in 2000, indicates a net overcount of this group.

Table 2 shows net percentage undercount estimates for the 1940 through 2000 censuses, as derived by demographic analysis. The last two columns of the table, for 1990 and 2000, reflect the revised DA estimates discussed above. The table indicates a decrease in the estimated net undercount rates for the total population, blacks, and non-blacks in every census year except 1990, when the rates increased for the overall population and the two groups within it. In each of the seven censuses, a differential undercount was noted: the estimated net rate was higher for blacks than for non-blacks.

S.13.3.1D FURTHER GUIDELINES FOR ADMINISTRATIVE DATA PROCESS

Overcoverage and undercoverage, including multiple listings, can be significant sources of error in administrative data processes. They can appear in many forms, often unique to a process. Thus, there is no standardised way to present them.

For *register-based data*, coverage errors should be treated in the same way as for censuses, since registers are essentially to be seen as census data with reference to a specific target population. A special issue to be addressed is lags in registration, or deregistration, of units, which may give rise to under-and overcoverage, respectively, at any fixed point in time.

For *event-reporting systems* the key accuracy issue is the undercoverage that results from some events going unrecorded. Special evaluation studies are typically needed to get a precise measure of *the rate of unrecorded events* but there is usually good qualitative knowledge of where this rate is serious. For example:

- Road accidents without personal injuries go unrecorded to a much greater extent than those with personal injuries.
- The rate of unrecorded crimes is bigger for certain crimes (e.g. rape) than for others (e.g. murder).

S.13.3.1D EXAMPLE FOR ADMINISTRATIVE DATA PROCESS

Example S.13.3.1D-1 Crime statistics, Sweden([12])

[This example provides a brief exposition of undercoverage.]

The greatest problem when drawing conclusions from crime statistics on real crimes committed is that far from all crimes enter statistics. Daily, a large number of crimes are committed that are unknown to the system of justice and for many types of crime real criminality is therefore considerably more abundant than what can be seen from official statistics.

There is a hidden criminality in addition to reported crime. The size of the hidden criminality is unknown and varies across types of crime but, for some crimes, is assumed to be so large that you can compare with the tip of an iceberg. The relation between the real and the reported number of crimes is called *the rate of unrecorded crimes(*[13]).

S.13.3.1E FURTHER GUIDELINES FOR MULTISOURCE PROCESS

Often sources are combined in order to cover a larger population. For example, in statistical processes that produce statistics on *culture* there are several sub-populations - museums, theatres, concerts, cinemas, libraries, book publishers etc. – for which data may be obtained from different sources.

It is also possible that the sources have different populations that cannot be combined for data collection purposes because they contain different types of units. For example, waste is created by households, businesses and government institutions (see example 4.1). These subpopulations of waste generators need to be reached by different data collection methods.

Each of the component sources has its own coverage issues, and they should be described at a level of detail commensurate with their impact on the statistics.

S.13.3.1F FURTHER GUIDELINES FOR MACRO-AGGREGATE COMPILATION PROCESS

In macro-aggregate compilation processes, coverage is often defined by international standards. Coverage issues are usually handled by model-based procedures for estimating gaps and are therefore of a different nature than in sample surveys.

Deviations from international standards with respect to coverage should be reported here if not already included in overall accuracy.

([12]) Translated from Swedish
([13]) In Swedish *mörkertal*, literally translated as “darkness number”.

S.13.3.2 Measurement error

SIMS	Concept name	Definition	Guidelines
S.13.3.2	Measurement error (P)	Measurement errors are errors that occur during data collection and cause recorded values of variables to be different from the true ones	The main sources of measurement error should be reported and assessed. Their description should be accompanied by any available analysis, otherwise by the producer's best knowledge. Where available and relevant describe: • identification and general assessment of the main sources of measurement error; • efforts made in questionnaire design and testing, information on interviewer training and other work on error prevention; • results of assessments based on comparisons with external data, re-interviews or experiments; • results of indirect analysis, for example, of the editing phase; and • actions taken to correct measurement errors. *European level* Where measurement errors are important as a single source of error provide a comparative summary across countries. Otherwise include them within overall accuracy in S.13.1.

S.13.3.2 FURTHER GUIDELINES

Measurement errors are errors that occur during data collection and cause the recorded values of variables to be different from the true ones. Their causes are commonly categorized as:

- *Survey instrument*: the form, questionnaire or measuring device used for data collection may lead to the recording of wrong values. Reasons may be the wording of the questions in the questionnaire, the order or context in which the questions are presented.
- *Respondent*: respondents may, consciously or unconsciously, give erroneous data due to confusion, ignorance, carelessness or dishonesty;
- *Interviewer*, due to inadequate training or surveillance, interviewer may inappropriately influence the answers

The term "measurement" here refers to measurement *at the unit level*, for example, the monthly income of a person or the annual turnover of a company. The result of a measurement may be viewed as comprising the true value plus an error term that is zero if the measurement is correct. This implies that a true value exists, which is sometimes subject to debate.

Measurement errors can be systematic or random. Random errors are often associated with the idea of replication, i.e., if the measurement process is repeated many times for the same unit under fixed conditions the registered measurement values will vary randomly whereas the systematic error will stay constant. The following simple model can be used to represent this fact for the registered value y_k:

$y_k=Y_k+B_k+e_k$,

where Y_k is the true value, B_k the systematic error and e_k the random error for unit k.

e_k has an average of 0 over repeated measurements whereas B_k is constant for a given unit.

More complex models can be obtained by splitting B and e according to the causes of error, e.g., questionnaire, respondent, collection method, or interviewer(14).

Thus, measurement errors may cause both bias and extra variability of statistical outputs. Bias is usually the main problem. The evaluation of measurement errors depends on the type of data at hand. The producer report should identify the main risks in terms of measurement error for the statistical process under consideration.

Respondent errors may be caused by a desire to appear socially acceptable, or in responses to sensitive questions. Where such factors are at play in the survey data, a specific discussion of possible resulting measurement errors is necessary.

Questionnaires used in the survey should be attached to the producer report as annexes (or as hyperlinks if they are large). The efforts made in designing and testing the questionnaires should be briefly described.

The report should also include a list of the characteristics (variables) that are likely to have the highest measurement errors, together with the possible reasons (complexity, sensitivity, unclear formulation of question, respondents' or interviewers' inability to provide accurate answers, etc).

Data editing identifies inconsistencies. They can be measurement errors but may also be the result of processing errors due to coding or data entry. Information from the data editing process should be included in the producer report, since it is indicative of the risk of measurement error. The failure rate of each edit rule can be calculated over the records to which the edit is applied. Clerical correction and/or automatic imputation are usually applied in order to remove inconsistencies in the data. The failure rates, therefore, are an indication of the quality of data collection and processing and not of the quality of the final data. The amount of detail on data editing in a producer report should be related to the importance of measurement errors in the survey in general and for the key indicators.

A recent work by Laitila et al (2017)(15) presents results from a selective editing approach for business data, where influential errors in micro data are corrected. This method also allows for an estimate of remaining errors after the selective editing.

A method for finding errors in economic data is to subject the data to accounting rules and reasonableness checks. These approaches are usually used in the editing stage in order to correct the data before final estimation.

Evaluation

When the risk of substantial bias is considered high, evaluation studies are needed. Respondent error can be assessed by a re-interview study in which the respondent is asked to provide the same data on a second occasion. If there is no memory effect, the two interviews may be considered independent and the difference between the responses is an indication of the size of the measurement error.

In order to assess instrument and/or interviewer effects, repeated measurements can be made with different instruments (e.g., alternative phrasing of questions) or different interviewers. Alternatively, an experiment can be carried out with subsamples being randomly allocated to different instruments and /or interviewers. This approach is mostly appropriate for surveys on attitudes/opinions or where memory effects are involved. Information on relevant aspects of interviewer training could also be included. The interviewer effect can also be estimated with the data from the survey (without a further re-interview on a subsample), if the allocation of units to interviewers was random (this is quite simple in CATI surveys) or carried out with the interpenetrating sampling technique.

For data of a factual nature, especially economic data, the potential for finding other databases with similar data is often good. Such databases may contain similar data with a time lag and can be used for evaluating earlier versions of the present statistical output. However, when comparing two sets of data, it is necessary to distinguish measurement errors from comparability issues, such as differences in definitions, with which they may be confounded.

Four groups of methods are applicable for evaluating errors at unit level. Such errors could have been

(14) Biemer and Stokes (1991) give an overview over many possible measurement models.

(15) Quantitative Measurement Errors in Partially Edited Business Survey Data. In Total Survey Error in Practice, Wiley, 2017.

generated in the measurement phase, the processing phase or they could have existed already in the survey frame.

Comparison with data from another source at unit level. This is a good way to obtain an overall quality check, covering all sorts of errors, provided the other source can be assumed to have accurate data for certain variables. The comparison is likely to be most informative if the sources have a common unit identification scheme, otherwise matching of units in the two sources will be required, as mentioned under coverage errors above. In making comparison care must be taken to distinguish actual errors from differences resulting from differences in definitions or measurement points in time.

Control at source /re-interview with superior method. Control at source means that the evaluator gets access to source data (company accounts or records kept at an agency etc.) A re-interview with a superior method may use an expert interviewer or face-to-face instead of mail interview. Another approach is to use the same interview method once again (but with a different interviewer) and use a reconciliation procedure (for example an expert panel) where different responses are obtained. Such methods capture all types of errors that have occurred during measurement and data entry, whether due to respondent, questionnaire, interviewer or data entry. They are best done for a random sample of units resulting in unbiased estimates of error.

Replication. Replication means that there are two or more observed values for a sampled survey unit. Such values can be obtained by different interviewers, from different respondents (answering for the same sampled unit) or simply by repeating the measurements after sufficient time for the respondents not to remember their initial responses. The differences between the measurement values can be used for learning how stable the measurement process is. Formal analyses of replication often assume that errors are independent between replications. This assumption is rarely fully met in practice.

The method is used for estimating the random variation due to measurement. Under some circumstances (for example if an expert interviewer or respondent is used) it can also provide some information on the systematic error (bias).

Effects of data editing. By comparing results from original and edited data the extent of initial measurement error can be deduced. Of course, this gives a minimum estimate of the error levels, if not all errors are detected in the editing process. Such analyses provide ideas for improving the measurement methods, but no information on the undetected measurement errors nor how they affect the statistical outputs.

More details for certain types of process

The description above applies to surveys of all kinds. For certain types of surveys, more specific guidelines are possible.

For *business surveys*, a critical issue is how certain economic and accounting concepts are explained and understood. Additionally, failure or inconsistencies in attributing economic flows to the right period (periodisation) may be an important source of measurement error.

In *price indexes*, quality change and quality adjustment when replacements of products are necessary result in measurement error due to imperfect procedures.

In *National Accounts* and similar compilations, measurement approaches are sometimes classified into A, B or C according to risk of error. Thus, the frequency of A/B/C methods may be seen as an indicator of measurement error.

S.13.3.2 EXAMPLES

Example S.13.3.2-1 Labour costs survey 2008, 2012 and 2016, Norway - NACE Rev. 2 activity ESS-MH

[This example describes specific measurement errors in a survey and ways to avoid or correct for them.]

Measurement errors mainly occur because the respondent misunderstands what is included in, and/or consequently report wrong, each column in the questionnaire, or because information requested is difficult to obtain. To avoid this, the questionnaire uses the most common book-keeping terms and commonly known aggregates of time and hours such as normal working hours, overtime, vacation and various types of absence as far as possible. If suspected erroneous, data can be corrected by asking respondents to update the questionnaire or obtain data from other sources such as administrative registers. In cases where none of the previous mentioned methods apply, related statistics were used to establish base levels or valid boundaries/extremes, and logical controls were used for further correction and/or imputation.

The respondents were asked to report the average number of employees throughout the year. To help generate this number, respondents were to fill in the number of employees for each month. In some cases, there were mismatch between the level of costs accumulated through the year and the number of employees. We have therefore cross-checked reported numbers of employees with the NAV State Register of Employers and Employees (EE-register) In cases where substantial discrepancy was revealed imputation/correction were made.

Example S.13.3.2-2 Farm structure survey, DESTATIS, Germany 2016 [ESS-MH]

[This example describes specific measurement errors in a survey and how they were addressed.]

6.3.2. Measurement error

Characteristics that caused high measurement errors

The primary reasons for missing or erroneous information in the 2016 Farm Structure Survey are the size of the questionnaire and different reference periods between variables. Furthermore, some questionnaire variables are considered sensitive by respondents (e.g. ownership and tenancy including rents (national purpose), and the manure management), which lessens response willingness. In addition, and in despite of the great care that was taken in preparing the questionnaire comprehension difficulties frequently occurred in the questionnaire sections soil cover, tillage methods, crop rotation and questions about other gainful activities, as the relatively large number of follow-up enquiries by farmers showed. In one Land there were technical difficulties of matching länder specific rural development measures and EU rural development measures.

All measurement errors were corrected – if recognised as such, for example through distinct deviations from previous year or experienced values – during data editing. Moreover a pretest was conducted with voluntary farmers to improve the questionnaire. In the context of the pretest, the performance and the usability (understanding / user-friendliness) of the online-questionnaire were tested.

Example S.13.3.2-3 Labour force survey 2017, France [ESS-MH]

[This example is a response to a questionnaire that illustrates how to account for measurement errors taking into account the specific features of the survey. It has been abbreviated.]

6.3.2 Measurement Errors

a) Proxy interviews

Are proxy interviews allowed for the AHM? (Y/N)

Y (but highly discouraged)

b) Testing of the questionnaire

Was the AHM questionnaire tested? (Y/N)

Y

If Yes: which methods were applied	Number of tested persons/respondents
Paper test	155
Capi test	273
General rehearsal	Scenarios with 6 households and 12 respondents

c)Training of interviewers

Features	***Brief comments***
Special training on the ad hoc module? (Y/N)	Y
Written instructions? (Y/N)	Y
Debriefing with interviewers? (Y/N)	N
Comments (other)	Written comments by the interviewers during field work

d) Description of problems experienced and solutions adopted. Before LFS fieldwork, i.e. pre-test, questionnaire design, translation, etc

General	Doubts about the measurement of economic dependency led to the introduction of specific French variables. Tests then were used mostly to test and correct these French variables
By variables	
MAINCLNT	Clients" can be ambiguous. The order of response items was changed for fluidity, and they were read to the respondents (to avoid, for instance, that they would count their precise number of clients when it's more than 10)
SEDIFFIC	Some items were felt as missing by respondents, especially "heavy tax burden". As a consequence, they chose a proximate answer, "heavy administrative charges". The wording was changed to avoid this.

S.13.3.3 Nonresponse error

SIMS	Concept name	Definition	Guidelines
S.13.3. 3	Nonresponse error (P)	Nonresponse errors occur when the survey fails to get a response to one, or possibly all, of the questions	Provide a qualitative assessment of unit nonresponse. Highlight the variables that are most subject to item nonresponse (e.g. associated with sensitive questions). Provide a qualitative assessment of the bias associated with nonresponse. Provide a breakdown of nonrespondents according to cause for nonresponse. Describe efforts to reduce nonresponse during data collection and follow-up. Describe treatment of nonresponse at the estimation stage, including response modelling. *European level* Provide a qualitative assessment of unit and item nonresponse across countries.
S.13.3. 3.1	A4. Unit nonresponse - rate (P)	The ratio of the number of units with no information or not usable information to the total number of in-scope (eligible) units, at a level of detail appropriate for a producer report.	Report A4: Unit nonresponse rate overall and at a level of detail appropriate for a producer report. *European level* Unit nonresponse rates across countries
S.13.3. 3.2	A5. Item nonresponse - rate (P)	The ratio of the in-scope (eligible) units which have not responded to a particular item to the in-scope units that are required to respond to that particular item, at a level of detail appropriate for a producer report.	Report A5: Item nonresponse rate for all variables *European level* Item nonresponse rates across countries

S.13.3.3 FURTHER GUIDELINES

Whilst it is, in principle, possible to speak of unit nonresponse for a non-probability survey or an administrative data process, it is not clearly distinguishable from undercoverage in this case. Item non-response, however, can be equated to missing data in other statistical processes. Thus, these guidelines are foremost aimed at probability surveys and census surveys. For other processes, some of the recommendations could still be applicable.

The difference between the statistics computed from the collected data and those that would be computed if there were no missing values is the nonresponse error.

There are two types of nonresponse:

- *unit nonresponse,* which occurs when no data are collected from a unit in the sample; and

- *item nonresponse* which occurs when values for some but not all survey data items (variables) are obtained from a unit.

Accordingly, response (and nonresponse) is measured in terms of *response rates* of two kinds, broadly defined as follows:

- *unit response rate*: the ratio of the number of respondents (i.e., units for which data for some or all data items have been collected) to the total number of eligible units in the sample – the exact definition of the indicator A4 is provided in Supplementary Document C;

- *item response rate*: the ratio of the number of units that have provided data for a given data item to the in-scope units that are required to respond to that particular item - the exact definition of the indicator A5 is provided in Supplementary Document C.

Other ratios are sometimes used instead of, or as well as, these ratios of counts. They are:

- *design-weighted response rate,* which sums the weights of the respondents using the sample design weights;

- *size-weighted response rate,* which sums the values of auxiliary variables multiplied with the design weights, instead of the design weights alone.

Measures of nonresponse

The sample can be divided into the following subsets:

- R: Responding units in the target population;
  - F: Units in R for which a full response is obtained;
  - P: Units in R for which only partial responses were obtained;
- N: Non-responding units that belong to the target population;
- U: Units with unknown target population status (either nonresponse or overcoverage);
- O: Units not belonging to the target population (overcoverage).

The number of sample units in each subset is denoted n_X, where X is one of the letters in the above list.

The total sample size $n = n_R + n_N + n_U + n_O$ and $n_R = n_F + n_P$

The design weight d_j of unit j in the sample is its inverse inclusion probability. For the size-weighted case x_j is the size of unit j.

For the units with unknown status, it is assumed that a proportion α is nonresponse. In practice, unless there are strong reasons to the contrary, it is recommended to set $\alpha=1$, which gives a conservative (upper) bound to the nonresponse rate.

Reasons for nonresponse

There are several possible reasons for nonresponse. The following is a list of such reasons.

- Failure of the data collector to locate/identify the sample unit;
- Failure to make contact with the sample unit;
- Refusal of the sample unit to participate;
- Inability of the sample unit to participate (e.g. ill health, absence, etc);
- Inability of the data collector and sample unit to communicate (e.g. language barriers);
- Accidental loss of the data/ questionnaire.

Since the risk for bias may be quite different for these types, the producer report should, as far as possible, provide a breakdown of nonresponse according to a classification of this kind.

Unit nonresponse reporting

Using this nomenclature, the various types of unit response rates are provided in Figure S.13.3-1 below

Figure S.13.3-1 Definitions of response and nonresponse Rates:

	Response rate	Nonresponse rate
Unweighted	$Rr_{uw} = \frac{n_R}{n_R + n_N + \alpha n_U}$	$NRr_{uw} = 1 - Rr_{uw}$
Design-weighted	$Rr_{dw} = \frac{\sum_R d_j}{\sum_R d_j + \sum_N d_j + \alpha \sum_U d_j}$	$NRr_{dw} = 1 - Rr_{dw}$
Size-weighted	$Rr_{sw} = \frac{\sum_R d_j x_j}{\sum_R d_j x_j + \sum_N d_j x_j + \alpha \sum_U d_j x_j}$	$NRr_{sw} = 1 - Rr_{sw}$

Where nonresponse exists, unit response rates thus defined should always be included in the producer report using the most relevant variants (unweighted, design-weighted or size- weighted) according to the judgement of the survey manager in each case. For business surveys, size-weighted nonresponse rates are normally the most relevant but it may also be informative to include several measures side by side.

The exact definition of response or nonresponse rates (formulae, etc.) should be included in the producer report along with the numerical information on the rates.

The rates should also be presented for important sub-domains.

A breakdown of the non-respondents into refusals, no contact and other causes is also informative.

In all definitions of response or nonresponse rates, sampling units identified as overcoverage should neither be included among the respondents nor among the non-respondents. However, it is often informative when presenting the nonresponse rates to also include overcoverage as a separate category.

The impact of nonresponse on the statistical outputs is likely an introduction of bias and an increase in sampling error. Sampling error increases simply because the available number of responses is reduced. Bias, which is the main problem with nonresponse, is introduced if non-respondents are not similar to respondents for all data items in all strata (whereas standard methods for handling nonresponse assume they are).

If monetary or other incentives are used to reward respondents, this may affect the relative propensity to respond in different subgroups and thereby introduce another kind of bias. If incentives are used a producer report should therefore assess their effects on nonresponse and bias.

Item nonresponse reporting

For item nonresponse rates there is basically a choice between two reporting approaches, which can also be used in parallel.

- If the focus is on a particular data item Y, response rates with regard to that item can be defined as in Figure S.13.3-1 above but with R defined as “responding to data item Y”. These rates are the most relevant ones for judging the accuracy of an estimate for data item Y and should be used for all key data items in a survey. They are referred to as *item Y response rates.*
- If the focus is on item response rates across all data items, then the rate of *full response for all data items* is of interest. In practice this is less frequently used.

In the cases of an item nonresponse, there is a choice of explicitly imputing, or not, the missing value. Practices regarding imputation should be included in the producer report together with an assessment of their impact on estimates and sampling errors.

Evaluating the effects of nonresponse

The technical treatment of nonresponse at the estimation stage (by imputation, re-weighting, or by exclusion) should also be clearly stated.

As discussed in Section S.13.2A, the increased sampling error due to nonresponse can and should be taken into account when computing CVs or confidence intervals. Efficient use of auxiliary information can sometimes improve precision considerably in the presence of nonresponse.

The remaining and more difficult issue is how to obtain information on nonresponse bias. Response rates provide an indication of the risk of bias, but the actual bias depends also (and mainly) on the average differences between the respondents and non-respondents with respect to survey variables. Normally there is some evidence, although rarely firm, on this matter, which should be included in the producer report in the form of a qualitative assessment. A basic approach is to compare the response and non-response strata with respect to any data that are available for both these strata. Three other approaches are outlined below.

Complementing with register data. This method assumes that there is a sufficiently strong correlation between a survey variable for which there is nonresponse and another variable in the survey frame or another register. This information can be utilised in various ways. For evaluation, one way is to compare the “estimate” of this other variable derived from the whole sample with that derived from the sample excluding nonrespondents. A small difference provides some indication of a small nonresponse bias for the survey variable as well. The better the correlation is between the two variables, the better, of course, is the judgement that can be made in this way.

Special data collections. These methods aim to show how the nonresponse error would change if it were possible to increase the response rate. The studies are done so that a higher response level is reached than the one achieved with normal effort. For example, more effort can be set aside for tracing, more effort by other staff for persuading refusers to respond, increased time for field work, allowing other collection forms, reducing response burden by concentration on fewer variables or by offering incentives to the respondent. The differences in estimates thus obtained will reflect not only nonresponse error but also measurement and random sampling errors.

Variations over response waves. The purpose of studying responses over response waves is to show how estimates change as a larger share of data collection is accomplished. Results are of interest when intending to publish flash estimates based on data obtained before a certain date. Another use arises in the context of a need (for budgetary or timeliness purposes) to reduce the target response rates and to be able to judge in advance the consequences of such a reduction.

A more controversial use of such studies is to draw conclusions about the remaining non-respondents based on those that responded in the last wave. Although such an approach can shed some light, further evidence is needed before drawing strong conclusions on bias.

S.13.3.3 EXAMPLES

Example S.13.3.3-1 Unit non-response in EU-Survey of Income and Living Conditions (EY-SILC), Eurostat, 2015 [ESS-MH]

[Here is an example illustrating decomposition of non-response rate according to data collection mode.]

The Commission Regulation (EC) No 28/2004 defined indicators aimed at measuring unit non-response in EU-SILC. They are respectively:

- Address contact rate (Ra): the ratio of the number of addresses successfully contacted to the number of valid addresses selected.
- Household response rate (Rh): the ratio of the number of household interviews completed (and accepted in the data base), to the number of eligible households at the contacted addresses.
- Individual response rate (Rp): the ratio of the number of personal interviews completed (and accepted in the data base), to the number of eligible individuals in completed households.

Non-response is cumulative at the three stages (address contact, household interview and personal interview), so that the overall non-response rates for households and individual interviews are defined, respectively, as follows:

- Overall household interview non-response rate: NRh = 1 – (Ra*Rh)
- Overall personal interview non-response rate: *NRp = 1 – (Ra*Rh*Rp)

Example S.13.3.3-2 Item non-response in Labour force survey, Luxembourg, Eurostat, 2017 [ESS-MH]

[This example shows item-non-response rates for a number of important variables.]

Variables	Share of item non-response in % (unweighted) - before imputation (i.e. blank answers out of total applicable answers)
MAINCLNT	28.13
WORKORG	29.68
REASSE	30.5
SEDIFFIC	30.73
REASNOEM	38.4
BPARTNER	27.66
PLANEMPL	29.08
JBSATISF	0.32
AUTONOMY	25.28
PREFSTAP	26.14
OBSTACSE	99.07

Example S.13.3.3-3 Labour force survey at ESS level, Quality report of the European Union, Labour force survey 2015, 2017 edition, Eurostat

[This is an ESS level example of presentation of non-response rates.]

Table 4.4: Rates of unit non-response by wave. Annual average 2015

	Total	Waves							
		1	2	3	4	5	6	7	8
Belgium	26.7	26.7							
Bulgaria	22.2	29.2	21.5	20.1	17.5				
Czech Republic	20.5	22.6	20.3	19.7	19.8	20.3			
Denmark	47.0	50.0	48.0	47.0	45.0				
Germany	3.4	3.4							
Estonia	28.1	48.4	23.7	18.4	11.0				
Ireland	25.1	25.2	24.6	24.4	24.6	24.8			
Greece	25.9	25.7	25.2	24.6	23.8	23.3	22.3		
Spain	12.4	16.1	11.8	11.7	11.9	12.1	12.3		
France	20.3	24.3	20.0	19.1	19.3	18.9	20.1		
Croatia	30.3	30.1	30.2	30.8	30.2				
Italy	12.5	28.5	5.4	5.3	3.7				
Cyprus	5.4	6.5	5.1	5.2	5.2	5.1	5.0		
Latvia	37.9	36.2	33.8	39.7	39.6				
Lithuania	20.3	28.7	22.9	18.6	15.7				
Luxembourg	48.0								
Hungary	17.2	34.2	26.0	15.4	11.6	9.0	7.5		
Malta	23.4	16.8	26.7	25.3	24.9				
Netherlands	45.5	42.6	38.4	9.1	4.6	3.7			
Austria	7.8	5.9	7.9	7.4	7.7	7.7			
Poland	34.9	39.7	35.7	32.7	31.4				
Portugal	15.4	11.1	15.3	15.8	17.3	17.2	15.1		
Romania	12.2	13.3	11.5	12.7	11.2				
Slovenia	21.3	32.5	23.4	13.9	12.8	7.9			
Slovakia	15.2	19.8	14.0	14.6	14.0	13.5			
Finland	29.2	29.2	28.2	28.9	29.8	29.8			
Sweden	40.1	42.4	40.8	40.5	39.9	40.0	39.5	38.9	38.3
United Kingdom	47.4	43.7	42.7	47.3	50.5	52.9			
Iceland	22.7	23.0	23.4	23.2	22.4	21.3			
Norway	20.3	21.2	20.4	20.7	20.5	20.3	20.7	20.0	18.1
Switzerland	18.3	38.0	7.8	9.3	4.5				
Former Yugoslav	5.2	8.3	4.5	4.3	3.6				

Rep of Macedonia

S.13.3.4 Processing error

SIMS	Concept name	Definition	Guidelines
S.13.3. 4	Processing error (P)	The error in final data collection process results arising from the faulty implementation of correctly planned implementation methods.	If processing errors are significant, identify the main issues regarding them. Present an analysis of processing errors, where available, otherwise a qualitative assessment. Report their extent, and impact on the outputs, of the most significant types of error. Include descriptions of linking and coding errors, if applicable. Where mistakes relating to programming or publishing have occurred, corrective measures taken as well as actions for avoiding them in the future should be reported. *European level* Provide a summary across countries of processing errors.

S.13.3.4 FURTHER GUIDELINES

The following guidelines refer to the processing of *individual data*, data about individual units, referred to here as *microdata,* from a survey or administrative data process.

The SIMS definition of processing error can be better expressed as follows. Between data collection and statistical analysis, microdata undergoes *processing*, defined to comprise data entry, data coding, data editing (checking and correction), imputation, derivation, aggregation and weighting. All errors introduced in all these stages are collectively referred to as *processing errors*, with the exception of errors that are a consequence of the weighting methods, which are considered to be *model assumption errors* and described in Section S.13.3.5.

Measurement errors and processing errors both refer to microdata and thus evaluation of either type of error tend to involve the other type.

A type of processing error that is especially important to evaluate and report is that associated with coding response data provided in free text format, for example when information on occupation or education are requested in a population census. Coding the economic activity of a business from a description provided by the business is another example. The quality of a coding operation depends in a complex way on the coding rules, how they are interpreted in practice and on the knowledge/skills of the coders and/or quality of the automated coding systems.

Processing errors affecting individual observations cause bias and variation in the resulting statistics, just as measurement errors do. The importance of micro-data processing errors varies greatly between different statistical processes and their treatment in a producer report needs to be proportional to their importance. When they are significant, their extent and impact on the results should be evaluated. If such an evaluation has been made it should be included in the producer report.

In multisource statistics imperfect record linkages (the process of associating records in one dataset to those in another dataset) could be a significant error source. Where this is the case, the it should be treated in the quality report.

A different type of processing error is mistakes in implementing procedures after finalising the micro-data. These could be errors in programming, errors in handling aggregate data in spreadsheets (for example in macroaggregate processes or price indexes) or misprints in press releases. They are by definition unexpected and often only discovered after publication. Where such mistakes have occurred or the risk is known to be great, actions taken to eliminate or reduce them should be reported. See also S.11 Quality Management and S.17 Data Revision.

Evaluation

Studies of effects of editing. The effects of editing are obtained by comparing edited and unedited data. By calculating the final estimates based on both data sets, the total net effect of editing can be measured. These effects can be broken down by unit in a so called top-down list, where the effects by unit are sorted in descending sequence and the most influential units can be seen. Such a list can serve several purposes. One is to check once more that the influential units have their correct values; another is to generate ideas for optimising the editing procedures.([16])

Studies of coding variation. In an independent coding control study the coding is done twice without the coders being allowed to see each other's results. In dependent coding the second coder has access to the first coder's proposals. Dependent coding gives, as expected, smaller variation between the coders.([17]) High coding variation is of course an indicator of a large potential processing error. Similar control studies can be conducted to evaluate processing errors deriving from other forms of treatment, for example data entry.

S.13.3.4 EXAMPLE

Example 13.3.4-1: Labour Force Survey (LFS) 2018 Sweden Quality Declaration: Processing Error. Statistics Sweden (2018a)

(Section 2.2.5 Data processing)

[This example presents results from checking studies as applied to various classifications.]

Because the telephone interviews are computer-assisted, the main part of data registration occurs at the same time as the data collection. The only other registration is connected with the coding of the variables industry (SNI2007), sector (INSEKT2014), occupation (SSYK2012, ISCO-08) and socioeconomic group (SEI).

These variables are chiefly coded during the interview via matching against a list of occupations or via information collected from the statement of earnings register. For observations where matching cannot be carried out directly during the interview, manual coding is carried out subsequently, based on open answers collected during the interview.

In general, coding errors give rise to incorrect classifications, which in turn give rise to errors in the statistics. Some groups may be slightly overestimated, for example with respect to the number of persons employed, at the expense of other groups that will be underestimated. The table below shows the results of the checking studies done regarding coding in the LFS.

([16]) For more information on editing procedures including quality aspects the reader is referred to the UN handbook in three volumes: *Statistical Data Editing (UN), Vol 1, Statistical Data Editing (UN), Vol 2 and Statistical Data Editing (UN), Vol 3*

([17]) Lyberg (1981) gives an extensive treatment of the topic of coding.

Results of control studies regarding coding in the LFS

Variable	Level	Proportion of correct coding (%)	Control year
Swedish occupational classification standard (SSYK12)	1	95.9	2016
ISCO-08	1	95.4	2016
SNI2007 (NACE Rev. 2)	2	64.4	2014
Swedish Socio-economic classification (SEI)	1	96.0	2016
Standard for institutional sector classification (INSEKT14)	1	99.3	2014

S.13.3.5 Model assumption error

SIMS	Concept name	Definition	Guidelines
S.13.3. 5	Model assumption error (P)	Error due to domain specific models needed to define the target of estimation.	Describe process specific models, for example, as needed to define the target of estimation itself. Provide an assessment of the validity of each model. (Descriptions of models used in treatment of specific sources of error should be presented in the section dealing with those errors.) *European level* Where different models are used across countries, provide a comparative overview and discuss their validity and the likely effects of the differences.

S.13.3.5 FURTHER GUIDELINES

The following guidelines apply to all types of statistical process.

Models are often applied in statistics. Sometimes the target of estimation relies on an abstract model defined by a subject matter discipline. In other cases, such as seasonal adjustment, the model is of a purely mathematical-statistical nature. Sometimes a model is applied in estimation in order to improve precision or to adjust for non-response.

In *model-assisted*(18) estimation, models are only used for the purpose of reducing sampling error as defined by the design-based paradigm. Sampling error calculated according to the relevant variance estimation formulas is sufficient and no separate discussion of model assumptions is needed in the producer report. If the basic design-based estimation is extended to adjust for non-sampling errors, such as non-response, a description should be provided (in the appropriate section).

Model-dependent estimation is a different matter. In this case there are no design-based estimators to use and the inference depends on the model, whose assumptions need to be critically checked. When model-dependent estimation is used as a remedy for a particular non-sampling error (like nonresponse or measurement error) the discussion of the model should be in the error section concerned.

In other cases, the *target of estimation is based on a model in a domain related science*. The target is in such cases not a simple function of observed values but a more abstract concept. Natural science models are used in environmental statistics, medical models for some parts of health statistics and

(18) In the sense of Särndal et al, (1992)

economic models for concepts in economic statistics such as productivity and inflation. In such cases, the model should be described in the producer report and its validity assessed.

European level

Where the target of estimation is model based, usually defined by a domain specific science, this model should be presented and its validity discussed. Sometimes this could be done by linking to other reports where such issues are discussed in greater detail and only provide a brief treatment in the producer report. Examples of such model based concepts are production, consumption and inflation (defined by economics), climate change (defined by natural sciences) or various targets related to the environment.

S.13.3.5 EXAMPLES

Example S.13.3.5-1 Healthy life years expectancy Eurostat yearbook 2011, Health

[This example illustrates the logic for developing a model-based indicator.]

Since life expectancy at birth is not able to fully answer this question, indicators of health expectancies, such as healthy life years (also called disability-free life expectancy) have been developed. These focus on the quality of life spent in a healthy state, rather than the quantity of life as measured by life expectancy. The calculation of the healthy life years indicator is based on a self-perceived question which aims to measure the extent of any limitations because of a health problem that may have affected respondents as regards activities they usually do (for at least six months).

Example S.13.3.5-2 Greenhouse Gas Emissions [ESS-MH]

[This example illustrates the description of a model-based indicator.]

The indicator measures all man-made emissions of the so called ‘Kyoto basket’ of greenhouse gases, including carbon dioxide (CO2), methane (CH4), nitrous oxide (N2O), and the so-called F-gases (hydrofluorocarbons, perfluorocarbons, nitrogen trifluoride (NF3) and sulphur hexafluoride (SF6)). Using each gas’ individual global warming potential (GWP), they are being integrated into a single indicator expressed in units of CO2 equivalents. Emissions data are submitted annually by the EU Member States and EFTA countries as part of the reporting under the United Nations Framework Convention on Climate Change (UNFCCC).

Example S.13.3.5-3 International trade in goods Statistics

User guide on European statistics on international trade in goods 2016 edition, p. 32 (Eurostat, 2016)

[This example illustrates the logic for preparing estimates of data that cannot be directly connected.]

Estimates need to be made for data that cannot be collected because the trader is below the exemption threshold or that have not yet been collected because the trader is late in supplying the data. The main information sources used by the Member States to estimate missing data in intra-EU trade are the VAT returns and VAT recapitulative statements (VIES data) sent by intra-EU traders to the national tax administration. Data collected through Intrastat declarations are used to allocate estimated total values by product and partner Member States. Under the Intrastat legislation, the estimates for any missing data — i.e. on trade below the Intrastat exemption thresholds, and/or due to missing Intrastat declarations — must be compiled at least by HS2 codes and partner Member State. Missing data are even estimated at a more detailed level — HS4 or CN8 codes — by some Member States).

Example S.13.3.5-4 Consumption of toxic chemicals by hazardousness - http://ec.europa.eu/eurostat/web/products-datasets/-/sdg_12_10

[This example illustrates the description of an indicator.]

The indicator measures the volume of aggregated consumption of toxic chemicals, expressed in million tonnes. The consumption of toxic chemicals is calculated as the sum of two production-related indicators ('production of toxic non-hazardous chemicals' and 'production of toxic chemicals hazardous to human health respectively to the environment''), including data from official statistics on international trade in goods.

S.14 (PART II) Timeliness and punctuality

SIMS	Concept name	Definition	Guidelines
S.14	Timeliness and punctuality	(Defined by its sub-concepts)	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and all its sub-concepts are included in ESQRS based (producer) reports, where it is ESQRS Concept 7.

In ESMS based (user) reports the concept is included but some sub-concepts are treated differently or excluded. Those treated differently are indicated by (U); those included only in ESQRS are indicated by (P).

S.14.1 Timeliness

SIMS	Concept Name	Definition	Guidelines
S.14.1	Timeliness	Length of time between data availability and the event or phenomenon the data describe.	Outline the reasons for the time lag. Outline efforts to reduce time lag in future. *European level* For reports only published at European level do the above. Otherwise summarise the above across countries.
	TP2. Time lag - final results (U)	The number of days (or weeks or months) from the last day of the reference period to the day of publication of final results.	*For user reports only* Explain the meaning of TP2 and provide its value for the most recent cycle, and the average over a past period, say three years, at a level of detail appropriate for users.

SIMS	Concept Name	Definition	Guidelines
S.14.1.1	TP1. Time lag - first results (P)	The number of days (or weeks or months) from the last day of the reference period to the day of publication of first results, at producer report level of detail	*For producer reports only* Explain and provide TP1 values for most recent cycle, and average, and maximum over a past period, say three years.
S.14.1.2	TP2. Time lag - final results (P)	The number of days (or weeks or months) from the last day of the reference period to the day of publication of complete and final results	*For producer reports only* Explain and provide TP2 values for most recent cycle, and average, and maximum over a past period, say three years at a level of detail appropriate for producers.

S.14.1 FURTHER GUIDELINES

The most common measure of timeliness (and reflected in indicators TP1 and TP2) is the *production time*, defined as the time interval from

- the end of the reference period (or point) to which the data refer; to
- the day of the release of statistics, in whatever format and by whatever medium they are first released.

The appropriate unit of time (day, week, month) depends upon the reference period, and is at the discretion of the report writer.

As production time is perceived by users as a *time lag,* the reasons for its length should be explained, especially when it could be considered by users to be very long. Efforts to reduce it should be detailed

For processes that are repeated annually or more frequently, the values can be averaged over a given number of process cycles or a time period, say three years. The number of cycles or the time period is at the discretion of the report writer.

Where a timeliness standard is specified in a domain-specific regulation, it should be used as a benchmark and the ratio of, or difference between, the actual production time and the specified standard production time should be stated.

For producer reports only

Some statistics are released in two or more several versions, for example preliminary and final versions, or preliminary, revised and final versions. In this case each release has its own production time. These releases should be distinguished and the timeliness for each presented separately. In particular, the values of TP1 (first release) and TP2 (final release) should be reported.

For processes that are repeated annually or more frequently, the maximum value(s) over the same period, which indicate(s) the worst recorded case(s), should also be reported.

S.14.1 EXAMPLES

Example 14.1-1 Foreign Affiliate Trade Statistics Annual 2016

Statistical Office of the Slovak Republic [ESS-MH]

[This example illustrates a minimal level of detail.]

Structural Business Statistics: 7-8 months after the reference year.

National Bank of Slovakia: 16-17 months after the reference year...

Administrative data: preliminary 8-9 months after the reference year; definitive data 16 months after the reference year.

Example 14.1-2 National Accounts Quality Report DESTATIS, Germany, 2018, p. 9

[This example illustrates a comprehensive description. It also indicates that results are not final until scheduled revisions are complete.]

The quarterly gross domestic product (GDP) is initially published in a GDP first release after about 45 days. This is followed by more detailed results in a press release published about 55 days after the end of the reference quarter (that is, for the first quarter of a year in May, for the second quarter in August, for the third quarter in November and for the fourth quarter in February). On those occasions, the previous results of the last few quarters – in August those of the last four years – are updated, too. The first annual result is published at a press conference in January, about 15 days after the end of the reference year. Although the legally binding European standards (t+70) thus are definitely more than met by German national accounts, the revisions caused by that are justifiable. However, there is a trade-off between timeliness and accuracy, that is, lower accuracy in the form of more need for revision is the price of more rapid calculation and earlier publication.

Generally, the last four years including the relevant quarters are revised in August of each year. The results of the earliest of the years become final at that status of calculation and need not be revised regularly any more. For example, the results of reference year 2008 became final in August 2012, subject to future major revisions. Such regular revisions are necessary to include into the national accounting system large-scale annual statistics whose results become available with some time lag from the end of the reference period. The results of these source statistics replace the data at the recent end of the series which was until then obtained partly through indicator-based calculations.

Example 14.1-3 Labour Force Survey Quality and Methodology Information Report

UK Office for National Statistics, 2011, p. 3 [ESS-MH]

[This example includes a component of timeliness, namely the data collection time, as well as the total production time. It also includes a comment on punctuality.]

For the LFS, the time lag between the delivery date of data and the end of the reference period is approximately 16 days, and the elapsed time between the end of the reference period and the publication date is approximately six weeks. Publication takes place strictly in accordance with published release dates for Labour Market Statistics, following the Code of Practice for Official Statistics. The publication date has never been missed.

Timeliness on a continuous survey such as the LFS should be carefully compared against surveys or administrative series which report on a point or only part of the reference period, particularly in regard to issues around discontinuities in the data (see the Labour Force Survey User Guide Volume 1: LFS Background and Methodology for guidance).

S.14.2 Punctuality

SIMS	Concept name	Definition	Guidelines
S.14.2	Punctuality	Time lag between the actual delivery of the data and the target date when it should have been delivered.	Report only for annual or more frequent releases. If a release schedule was made available to users and/or specified in a regulation; • provide TP3 (user formula), i.e., the percentage of releases delivered on time, based on scheduled release dates, over a specified period and/or set of outputs. • in the event of any non-punctual releases, explain the reasons and outline efforts to improve punctuality. In the absence of a release schedule, explain why there is no schedule and indicate what efforts will be made to make one available in the future. *European level* For outputs first published at European level, report as above. For outputs first published at country level, • state the agreed time frame for delivery of national data and the actual delivery dates; • summarise punctuality across countries.
	TP3. Punctuality - delivery and publication (U)	The percentage of release delivered on time.	
S.14.2.1	TP3. Punctuality - delivery and publication (P)	The number of days between the delivery/ release date of data and the target date on which they were scheduled for delivery/ release.	Explain the meaning of indicator TP3 with producer report calculation formula, i.e., the time lag between scheduled release date and actual release date. Provide the average value of TP3 for the most recent cycle. In the case where the indicator refers to data tables sent to Eurostat, the value of this indicator can be compiled by Eurostat For a repeating process, provide the average value of TP3 over a past period, say three years. *European level* For outputs that are first published at European Level, do as above.

S.14.2 FURTHER GUIDELINES

When there is no release schedule

If no release schedule is specified in a regulation and/or made available to users, the report should state that punctuality cannot be measured or further discussed.

When there is a release schedule

Note that the primary measure, i.e., the standard indicator TP3, has a completely different calculation formula for user and producer reports.

The specified period/set of outputs for TP3 (user report) is at the discretion of the report writer.

For producer reports only

The unit of measure for TP3 (producer report) is specified as *days* but is at the discretion of the report writer.

For processes that are repeated annually or more frequently, the time period, or number of process cycles over which averages are presented is at the discretion of the report writer.

Some statistics are released in several versions, for example preliminary, revised and final. In this case each release may have its own target release date. The releases should be distinguished and the punctuality of each separately reported.

S.14.2 EXAMPLES

Example 14.2-1 National Accounts Quality Report, DESTATIS, Germany, 2013, p. 9 [ESS-MH]

[This example illustrates a minimal but informative report on punctuality. It would be improved by a link to the schedule.]

The release dates to be reported to Eurostat and the IMF are indicated in the annual release calendar of the Federal Statistical Office for major economic indicators one year in advance. In the past, those deadlines were always met.

Example 14.2-2 Labour input in industry, number of persons employed, 2016 Statistics Sweden

[This example illustrates another minimal report on punctuality that would be improved by a link to the schedule.]

The statistics are always published on schedule. All deadlines for quarterly publishing are met.

Example 4.2-3 Labour Force Survey Quality and Methodology Information Report, UK Office for National Statistics, 2011, p. 3

[This example includes more details regarding punctuality.]

Publication takes place strictly in accordance with published release dates for Labour Market Statistics, following the Code of Practice for Official Statistics. The publication date has never been missed.

For more details on related releases, the UK National Statistics Publication Hub is available online and provides 12 months' advance notice of release dates. If there are any changes to the pre-announced release schedule, public attention will be drawn to the change and the reasons for the change will be explained fully at the same time, as set out in the Code of Practice for Official Statistics. See:

- http://www.ons.gov.uk/ons/guide-method/ons-independence/publication-hub/index.html
- http://www.ons.gov.uk/ons/guide-method/revisions/ons-compliance-statement/index.html

Example 4.2-4 Quality Report on National and Regional Accounts, 2018 Edition, Eurostat.

[This comprehensive example is from a European level producer report.]

6.2. Punctuality of ESA 2010 tables

Punctuality is calculated as the actual date of data delivery minus the scheduled date of transmission to Eurostat. It shows how many calendar days the first data transmission was behind the legal deadline. Figures 21 to 35 present in detail the information on punctuality for each national accounts domain for EU Member States as well as for Iceland, Norway and Switzerland.

The ESA 2010 Transmission Programme specifies the deadlines for Member States and EFTA countries' data deliveries. However, due to derogations, the transmission dates vary across countries. The analyses of punctuality in this section take this into account.

6.2.1. Quarterly data

As defined in the ESA 2010 Transmission Programme, Member States and EFTA countries must submit to Eurostat the following quarterly tables:

- national accounts main aggregates (Table 1), at t+2 and months,
- non-financial sector accounts (Table 801), at t+85 days,
- financial accounts of general government (Table 27), at t+85 days and at t+3 months,
- government debt (Maastricht debt) for general government (Table 28), at t+3 months.

The overall punctuality of quarterly national accounts was relatively high in 2016 as more than half of the countries submitted all mandatory quarterly accounts on time. 16 EU Member States (Belgium, the Czech Republic, Germany, Estonia, Spain, France, Latvia, Lithuania, the Netherlands, Austria, Portugal, Slovenia, Slovakia, Finland, Sweden and the United Kingdom) submitted data at or before the legal deadline. Four EU Member States (Germany, France, the Netherlands and Finland) submitted 90 % or more of their quarterly tables before the legal deadline.

S.15 Coherence and comparability

SIMS	Concept name	Definition	Guidelines
S.15	Coherence and comparability	Adequacy of statistics to be reliably combined in different ways and for various uses and the extent to which differences between statistics can be attributed to differences between the true values of the statistical characteristics.	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and all its sub-concepts are included in ESQRS based (producer) reports, where it is ESQRS Concept 8.

In ESMS based (user) reports the concept is included but some sub-concepts are treated differently or excluded. Those treated differently are indicated by (U); those included only in ESQRS are indicated by (P).

S.15 Background

S.15A TERMINOLOGY AND GENERAL EXPLANATIONS

European statistics should be coherent in the sense of being consistent internally and over time, and comparable between regions and countries. It should be possible to combine and make joint use of related data from different sources.

There is not an agreed, unified, definition of *coherence*. In different statistical domains the term is used in different ways. Generally speaking, the coherence of statistics reflects the degree to which they can be successfully combined within a broad analytic framework and over time. European statistics should be coherent in this sense.

A distinction between coherence and accuracy is necessary to make. Coherence refers to, and is measured in terms of, *design metadata* (i.e., concepts and methods) about the processes, whereas accuracy is measured and assessed in terms of *operational metadata* (sampling rates, data capture error rates, etc.) associated with the actual operations that produced the data. With this understanding, coherence may be assessed in terms of the design metadata, and accuracy in terms of operational metadata. Differences that purely result from sampling variability are not due to incoherence but to lack of accuracy.

Comparability can be seen as a special case of coherence when the focus is on comparisons between regions, countries, domains and over time.

The term *coherence* is usually used when assessing the extent to which the outputs from different statistical processes have the potential to be reliably used in combination, whereas *comparability* is used when assessing the extent to which outputs from (nominally) the same statistical process but for different time periods, different countries/regions and/or different domains have the potential to be reliably used for comparisons.

Incoherence and non-comparability can affect statistics originating from different sources. Causes may be:

- Differences in concepts. For example, a household could be defined in a number of ways with respect to the individuals who belong or not belong to it. An enterprise can be defined according to kind of activity, location or ownership.
- Differences in methods. For example, employment estimated from a household survey gives different results than when estimated from administrative data such as from an employment agency.

Either or both of these may be a result of changes in the statistical process(es) as they are modified over time. Modifications may occur for a whole variety of reasons – introducing improved questionnaires, methods, automation, new technology, more up to date classifications, or in response to changes in legislation, or as a result of contractions or expansions in budget and hence in sample size or follow-up capacity, etc. For example, when Finland changed the data collection medium of the Labour Force Survey from postal enquiries to personal interviewing in 1983 the result was an increase of 100,000 in the estimate of employed people.

There are several areas where the assessment of coherence is regularly conducted: between provisional and final statistics, between annual and short-term statistics, between statistics within the same socio-economic domain, and between national accounts and other domains.

Comparability aims at measuring the impact of differences in applied statistical concepts, definitions, measurement tools and procedures on comparisons of statistics across geographical areas, non-geographical dimensions, sectoral domains and over time. Comparability of statistics, i.e. their usefulness in drawing comparisons and contrasts among different populations, is a complex concept, difficult to assess in precise or absolute terms. In general terms, it means that statistics for different populations can be legitimately aggregated, compared and interpreted in relation to each other or against some common standard. Metadata in a producer report should convey such information that will help any interested party in evaluating comparability of the data, which is the result of a multitude of factors. Sometimes, it is possible to reconcile, wholly or partly, two estimates by quantifying effects of different concepts etc. and this is then of great value to users.

Consistency

The term *consistency* is closely associated, but not synonymous, with coherence. There are two types of consistency– logical consistency and numerical consistency.

Logical consistency requires that a statistical concept has one and only one definition in all areas of statistics that are subject to combination or comparison. This applies, for example, to such concepts as household, enterprise, employment and waste for which competing definitions are abundant. Where there are different definitions, statistics from different processes cannot be reliably combined or compared unless they can be reconciled, which is usually not possible.

Numerical consistency requires, for example, that, within a set of outputs for a statistical process, the numerical value for a whole is equal to the sum of those for its parts. Or that values for conceptually the same data item derived from different processes should be the same. When this is the case, coherence results and consistency is thus a sufficient condition for coherence. But also where there is not full numerical consistency, there can be coherence since the differences ("inconsistencies") can be explained by sampling variability or reconciled as being the result of different definitions whose effects can be quantified, thus enabling a coherent analytical approach to the data.

Logical consistency is thus an aspect of coherence, and numerical consistency depends upon both coherence and accuracy. The reason that logical and numerical consistency are not explicitly SIMS sub-concepts is because they are included within other SIMS sub-concepts. Nevertheless, the term consistency is frequently used in in the National Accounts, implying that statistical concepts should be logically consistent and there should be internal consistency, conceptually as well as numerically.

S.15B POSSIBLE DIFFERENCES IN CONCEPTS

Target population – units and coverage

The target populations may differ for two statistical processes, or for the same process over time, in a variety of different ways, as illustrated in the following examples.

- The definition of *economically active population* used in the labour force survey may differ from one country to another. In one country it might be all persons aged 16-65 who are employed or seek employment, in another country all persons aged 15-70 who are employed or seek employment.
- Monthly statistics of *industry* might include just manufacturing enterprises whereas another statistical output with the same name might include electricity, gas and water producers as well.
- An annual structural business survey might use an *enterprise* as its target statistical unit whereas a monthly production survey might use an establishment.

Geographical coverage

For example, rural areas might be included in one country's labour force survey and excluded from another's.

Reference period

For example:

- in a survey of employees, an enterprise might be asked for the number of full-time employees as of the first of the month, or as of third Monday in the month;
- an annual survey might refer to a *fiscal year*, another to a *calendar year*.

Data item definitions and classifications

As an example of a difference in definitions, the labour force survey definition of *unemployed person* might be:

- Any economically active person who does not work, is actively looking for a job and is available for employment during the survey; or
- Any economically active person who does not work, is actively looking for a job and is or will be available for employment in the period of up to two weeks after the survey's reference week

For example, persons waiting to start a new job are counted as *unemployed* in the EU standard Labour Force Survey but as *employed* in the US Current Population Survey. This has resulted in a difference of 0.23% in unemployment rate (Sorrentino, 2000).

Changes in classification schemes, in particular *revisions* in accordance with new versions of international standards, are a very common cause of coherence/comparability problems. An example would be adoption of the latest version of NACE in place of an older classification of economic activities.

In addition, even without a change in classification, the procedures for assigning classification codes may be different or change over time, for example with improved training of staff or the introduction of an automated or computer assisted schemes.

S.15C POSSIBLE DIFFERENCES IN METHODS

Survey population

The actual coverage of a survey depends upon the frame used for the survey. Possible examples of differences are as follows.

- A substantive difference would occur where one frame was based on value added tax, i.e., a source covering all enterprises paying VAT, whereas another survey frame was based on employment deductions, i.e., a source covering all enterprises with employees, who are subject to tax deductions.
- The legal requirements for VAT registration may change, resulting in more or fewer enterprises in survey frames.
- Surveys may be designed as cross sectional or longitudinal with significant difference in estimates of change as a result. Even within a longitudinal survey, panels or rotation patterns may change over time or between countries.
- Even without any nominal difference in statistical units, the procedures by which statistical units for large enterprises are actually delineated may differ or change over time in accordance with better training or new methods. For example, the procedures for treatment of the creation, amalgamation, merger, split, or cessation of an enterprise may change.
- Procedures for and timing of updating of survey frames result in the frames being more, or less, up to date.

Source(s) of data and sample design

An example of a difference might be that in one survey financial data for small enterprises are obtained from income tax data whereas in another they are obtained by direct survey.

Data collection, capture and editing

In one survey there might be intensive follow-up of non–response and consequential reduction of non-response rate to 10%, in another there might be no resources for follow-up, leading to a non-response rate of 40%, thus giving rise to a substantially increased probability of non-response bias. In a producer report for each survey this would probably be reported as an accuracy issue but for example in an EU publication covering all national surveys this would best be seen as a comparability issue, unless a full quantitative analysis of the respective biases can be provided.

Imputation and estimation

Different imputation practices may be applied for dealing with missing data items. For example, in one survey, zeroes might be imputed for missing financial items whereas in another survey non-zero values may be imputed based on the “nearest neighbouring” records.

Likewise, in dealing with missing records in an enterprise survey the corresponding enterprises might be assumed non-operational or might be assumed operational and similar to enterprises that have responded.

S.15D RELATIONSHIP BETWEEN COHERENCE/COMPARABILITY AND ACCURACY

As pointed out above *numerical inconsistency* between estimates can be caused by:

(i) sampling errors, response errors and/or other non-sampling errors; or

(ii) logical inconsistency, i.e., in concepts.

Whereas (i) are accuracy issues that should in principle be dealt with under S.13, (ii) are coherence issues. However, it is not always possible to fully disentangle these two factors and, if so, it is best to report both these factors under the heading of coherence/comparability.

The differences between *preliminary, revised and final* estimates generated by the same basic process relate to accuracy rather than coherence as there are no difference in concepts, only in data processing cut-off dates.

Where the error profiles of statistical processes are known and included within the descriptions of accuracy there is no need for further reference to them under coherence/ comparability, unless the errors cannot be explained in terms of accuracy alone. For example, suppose sampling error bounds are published for two values of the same data item for adjacent time periods indicating the range within which a movement from one period to the next may be due to chance alone and not reflect any actual change in the phenomenon being measured. If and only if the measured movement is larger than this, is there any point in discussing whether the movement is real or due to non-comparability.

Where the error profiles are not fully known and hence cannot be adequately reported under accuracy, there may still be a case for considering them as a possible cause of lack of coherence/ comparability. For example, if there is no assessment of non-response error then the assessment of coherence/comparability may include the possible consequences of differential non-response rates and patterns.

S.15E GENERAL ADVICE ON COHERENCE AND COMPARABILITY

Conceptual variations, such as deviations from relevant ESS legislation and other international standards, should primarily be dealt with under S.12.1 Relevance - User needs, where *details of definitions which differ from user requirements* are to be described.

Methodology is primarily presented in S.18 Statistical Processing and errors in S.13 Accuracy.

Hence the normal, and most logical, approach to presenting comparability and coherence issues is to refer to S.12, S.13 and S.18 for the details regarding concepts, errors and methods, respectively, and to limit the presentation in S.15 to the effects the differences in concepts and methods have on coherence and comparability.

Under each subcomponent of coherence and comparability, the report should make as clear as possible to what causes a given problem can be attributed. Ideally, the sources of incoherence/ non-comparability should be quantitatively decomposed by each possible source. If this is possible the corresponding sets of statistical outputs are said to be *reconcilable*. Although this is usually not fully attainable, the producer report should be as informative as possible with this goal in mind.

S.15.1 Geographical comparability

SIMS	Concept name	Definition	Guidelines
S.15.1	Comparability – geographical	The extent to which statistics are comparable between geographical areas.	Describe any problems of comparability between regions of the country. The reasons for the problems should be described and as well an assessment (preferably quantitative) of the possible effect on the output values. Give information on discrepancies from the ESS/ international concepts, definitions, with reference to other chapters for more details. *European level* Focus on factors that affect the comparability between countries. Analyse asymmetries in statistical mirror flows where possible.
S.15.1.1	CC1. Asymmetry for mirror flows statistics – coefficient	The difference or the absolute difference of inbound and outbound flows between a pair of countries divided by the average of these two values.	*For producer reports only* Provide measures of asymmetries for key variables.

S.15.1 FURTHER GUIDELINES

As regards geographical comparability – a producer report for a country has two different objectives:

- to describe problems in comparing different regions or other geographical entities within the country itself.
- to indicate where national concepts, definitions or methods may result in problems in comparing data with those of other EU countries. In this respect deviations from EU or international standards or guidelines should be pointed out. (Detailed descriptions should be given in S.12, S.13 or S.18.)

Add a description of the kind of problems that could arise in comparing different regions or other geographical entities within the country itself. Purchasing power between regions or countries could be an example to consider, the definition of urban areas another.

ESS level

Geographical comparability across countries assumes a key importance at ESS level.

Two broad categories of situation can be identified:

- where essentially the same statistical processes are used, for example, a labour force survey is designed in accordance with an ESS standard, and differences across countries are expected to be quite small; and
- where a different sort of statistical process is used, for example a direct survey in one country and register-based data in another. In such cases the differences are likely to be more profound.

A common situation is that a number of data points for each country is presented in ESS level tables. There should then be assessments as to whether these data are based on a coherent system of concepts and methods.

All sources of non-comparability concerning both concepts and methods should be considered. For a given statistical process, a few of these are likely to be of greater importance and for these a systematic exposition across countries should be provided. For the less important ones a shorter description is sufficient.

Comparability may be assessed in two different ways: pair-wise comparisons of the metadata across countries; and comparison of metadata for a country with a standard, in particular an ESS standard or, in its absence, an example of best practice from one of the NSAs. • A comparability matrix summarising by country the possible sources of non-comparability relative to a specified standard should be given.

Detailed reports on definitions, methods and errors across countries should primarily be provided under S.12, S.13 and S.18 whereas the effect of differences on comparability should be given here.

Mirror statistics

As previously noted in the section on the *Subject-dependent Techniques for Evaluation of Accuracy* (Section S.13.A), for certain selected statistical outputs from a country, notably in trade, balance of payments, migration and tourism, it may be possible to find counterpart statistical outputs in another country.

For example, the Dutch and Belgian statistical agencies may both produce statistics on migration between the two countries. If the Belgium estimate of emigration to the Netherlands in a particular year exceeds that of the Dutch estimate of immigration from Belgium for the same year by 10% then this could reflect lack of accuracy in the form of over counting in Belgium or undercounting in the Netherlands, and/or it may be the result of non-comparability of the Dutch and Belgium definitions of immigration, or emigration, or both.

The classical example of mirror statistics is for international trade in goods. In principle, country A's exports to country B over a certain period must equal country B's imports from country A. In practice, the comparison is blurred by factors such as valuation (i.e. whether freight and insurance are included or not), timing (arrival at B may be later than departure from A), and classification differences. However, adjustments for these factors can usually be made so that the extent of the actual errors can be more or less accurately determined.

CC1. Asymmetry for mirror flows statistics - coefficient

This standard indicator applies only to producer reports. See Supplementary Document C for its definition.

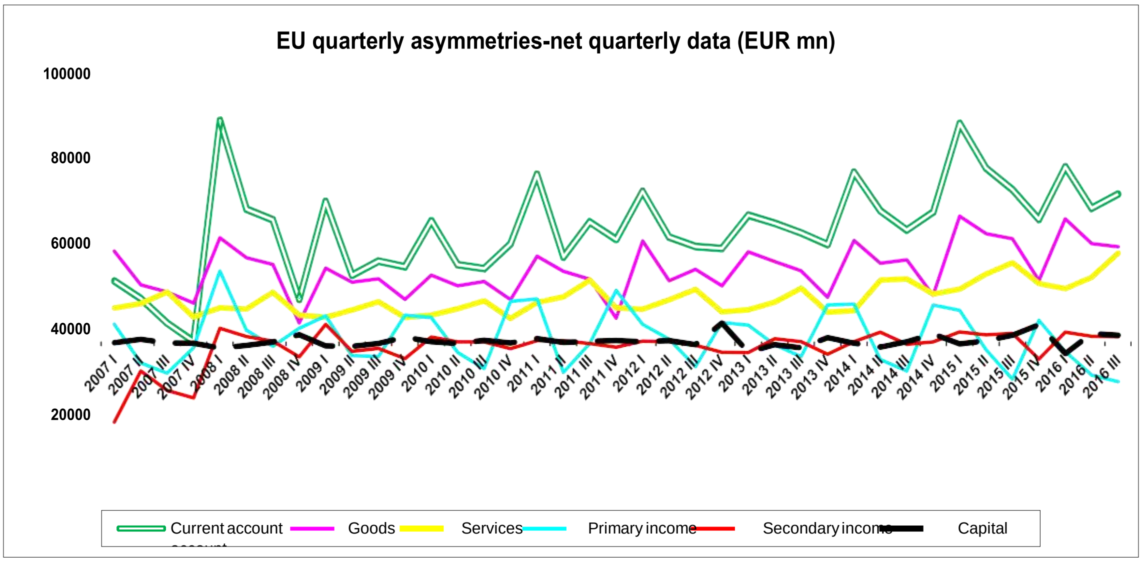
S.15.1 EXAMPLES

Example S.15.1-1 Mirror flows in Balance of Payments (Eurostat, 2018c)

[This example gives a detailed quantitative account of the BOP mirror flows.]

Figure 1 below shows total Intra-EU asymmetries based on quarterly BOP figures for periods from Q1 of 2007 until Q3 of 2016.

Figure 1 EU total asymmetries for main current and capital account items



Asymmetries for total current account mainly reflect fluctuations in asymmetries in trade in goods having positive imbalances (excess of recorded credits over debits). Asymmetries for services have been stable, also positive and lower than for goods. For primary and secondary income signs of imbalances have been changing; being quite low and without clear pattern for primary income and negative or around zero for secondary income. Current account asymmetries recorded a maximum value in Q1 of 2008 (3.6% of sum of credits and debits) and since then kept decreasing up to the first quarter of 2015 when a new peak was recorded. Starting from 2010 asymmetries have been around 1% of sum of credits and debits, with some growth in 2015 and 2016, for which data can be still considered as preliminary. Seasonal pattern can be observed with generally highest asymmetries’ values in the first quarters of each year.

Example S.15.1-2 National Accounts, EU level, [ESS-MH]

[This example explains geographical comparability in the National Accounts.]

The geographical comparability of national accounts in Member States and EFTA countries is ensured by the application of common definitions of the European System of Accounts (ESA) 2010. Worldwide geographical comparison is also possible as most non-European countries apply the SNA 2008 guidelines, and SNA 2008 is consistent with ESA 2010.

In so far as the countries correctly apply the ESA 2010 concepts, the country data published by Eurostat are comparable with each other. In addition, the euro area and EU aggregates are comparable with corresponding data for other economic areas that base their estimations on ESA 2010 or SNA 2008 methodology, like the United States. An exception concerns the euro area and EU estimates of the annual and quarterly national accounts main aggregates for the exports and imports of goods and services: they are not consolidated but presented on a 'gross' basis.

S. 15.2 Comparability over time

SIMS	Concept Name	Definition	Summary Guidelines
S.15.2	Comparability – over time	The extent to which statistics are comparable or reconcilable over time.	Provide information on possible limitations in the use of data for comparisons over time. Distinguish three broad possibilities: 1. There have been no changes, in which case this should be reported. 2. There have been some changes but not enough to warrant the designation of a break in series. 3. There have been sufficient changes to warrant the designation of a break in series. Provide values of CC2: Length of comparable time series at appropriate level of detail for user or producer report. The indicator CC2 is defined in Supplementary Document C.
	CC2. Length of comparable time series (U)	The number of reference periods in time series from last break.	
S.15.2.1	CC2. Length of comparable time series (P)		

S.15.2 FURTHER GUIDELINES

Comparability over time is a crucial quality aspect for all statistical outputs published on consecutive occasions. For many users, changes over time of economic or social phenomena are the most interesting aspects of the statistics, and comparability over time is essential if the data are to reflect the actual economic or social changes that occurred.

Regardless of whether statistics are directly published in time series form or whether the users have to construct their time series themselves from basic data, users need to be informed about possible limitations in the use of data for comparisons over time.

A producer report can be seen as intended also for advanced users, who often make complex combinations of data from different time periods for their analytical purposes. Therefore, a more detailed account of the differences in sources and methods used in different time periods will be of great benefit to the analyst.

Where there have been some changes but not enough to warrant the designation of a *break in series* the report should simply record the changes in the metadata describing the process.

Where there have been sufficient changes to warrant the designation of a *break in series*, users must be informed that there has been a break and provided with the information they need to deal with its consequences. The information provided may range from very complete to minimal depending upon the NSA resources available and the size of the break.

- The most comprehensive treatment is to carry forward both series for a period of time and/or to backcast the series, i.e., to convert the old series to what it would have been with the new approach by duplicating the measurement in one time period using the original and the revised definitions/methods.
- Another possibility is to provide the users with transition adjustment factors giving them the means of dealing with the break for example by doing their own backcasting.
- The least ambitious treatment is to simply describe the changes that have occurred and provide only qualitative assessments of their probable impact upon the estimates. Obviously, this is the least satisfactory from the user perspective.

S.15.2 EXAMPLE

Example S.15.2-1 National Accounts, EU level, [ESS-MH]

[This example illustrates a straight-forward declaration that there is full comparability over time.]

As the data for all reference periods are compiled according to the requirements of the ESA 2010, national accounts data are fully comparable over time. Also, in the case of fundamental changes to methods or classifications, revisions of long time series are performed, usually going far back into the past.

In principle, all country data and euro area and EU aggregates disseminated by Eurostat are comparable over time. However, in a limited number of cases comparability over time might be hampered. This is often the result when new sources for part of the time series are used. If possible, these cases are indicated in the database by marking them with a flag B (break).

S.15.3 Coherence – cross domain

SIMS	Concept Name	Definition	Summary Guidelines
S.15.3	Coherence- cross domain	The extent to which statistics are reconcilable with those obtained through other data sources or statistical domains.	An analysis of incoherence should be provided, where this is an issue of importance. Reporting under 15.3 is for coherence problems that are not reported under 15.3.1, 15.3.2 or 15.4
S.15.3.1	Coherence - subannual and annual statistics (P)	The extent to which statistics of different frequencies are reconcilable.	*For producer reports only*. Coherence between subannual and annual statistical outputs is a natural expectation but the statistical processes producing them are often quite different. Compare subannual and annual estimates and, eventually, describe reasons for lack of coherence between subannual and annual statistical outputs.
S.15.3.2	Coherence- National Accounts (P)	The extent to which statistics are reconcilable with National Accounts.	*For producer reports only.* Where relevant, the results of comparisons with the National Account framework and

SIMS	Concept Name	Definition	Summary Guidelines
			feedback from National Accounts with respect to coherence and accuracy problems should be reported and should be a trigger for further investigation.

S.15.3 FURTHER GUIDELINES

Cross domain coherence within a country refers to the consistency of outputs produced by different statistical processes within the country, for example, whether or not the employment produced by a labour force survey is reconcilable with the number of employees produced by an economic survey of enterprises.

A report for a particular statistical process should identify and report on known coherence problems in relation to other statistical processes that either measure the same concept or use the outputs from the process for producing aggregate results.

Where possible, a quantitative analysis of any lack of coherence should be presented. The goal of such an analysis should be to reconcile different estimates in order to facilitate an analysis of the results. Where this has been done in a special report, a summary can be provided together with a reference to the report.

The analysis should include a description of the differences between the statistical outputs being reported and other related statistical outputs, for example including differences in concepts and definitions, statistical units, classifications, geographical breakdown, reference period, and correction methods. It can also touch on limitations on using data from other sources set by coherence considerations.

In the case of the national accounts, the focus should be on comparisons with data from specific sub-sectors where no (or small) conceptual/population differences are to be expected.

S.15.3 EXAMPLES

Example 15.3-1 Coherence between the Norwegian Structure of Earnings Survey and the Labour Force Survey (Lien et al, 2009, p. 17-19)

[This example illustrates a full analysis of coherence issues between two surveys, provided as a part of a producer report for the Structure of Earnings Survey (excerpts only).]

Coherence with the Labour Force Survey (LFS) third quarter 2006

The following is a short presentation and comparison of the Norwegian SES and the Norwegian LFS surveys. It points out basic differences that possibly could be the cause of differences between the surveys as they are observed in the following tables. Statistics from the LFS are based on published figures. It contains comparison of basic information on model assumption, sampling, units and purpose. Several basic aspects of the LFS and SES are compared.

Population and sampling units

LFS

- Population All individuals aged 15-74
- Sampling unit Families
- Analysis unit Individuals
- Reporting unit Individuals
- Frequency Quarterly

SES

- Population All enterprises with employees
- Sampling unit Enterprises (by industry)
- Analysis unit Employees
- Reporting unit Employee (enterprise)
- Frequency Annual

Variable definitions

LFS

- Employed Persons on sick leave included
- Working time Full-time - 37 hours or more, if not defined otherwise by the reporting unit.

SES

- Working time Full-time - 33 hours or more per week

Objective of the LFS and SES statistics

LFS

- Provide statistics on employed and unemployed and labour force participation

SES

- Provide statistics on the level and composition of earnings for all employees (wage and salary earners)

Tabular results and comparisons with the LFS

Labour Force Survey. Distribution of full-time employees by sex and industry. 3rd quarter 2006

Industry	Frequency (%)		
	Males and females	Males only	Females only
C Oil and gas extraction, mining	1.7	2.1	0.9
D Manufacturing	16.0	20.4	8.5
E Electricity supply	1.0	1.3	0.5
F Construction	10.1	15.5	1.0
G Wholesale and retail trade and H Hotels and restaurants	16.7	16.8	16.4
I Transport and communication	8.3	10.5	4.5
J Financial intermediation	3.1	2.7	3.8
K Real estate and business services	13.7	14.8	11.8
M Teaching staff, private education	8.6	5.5	13.7
N Health and social work	16.7	6.9	33.4
O Social and personal service activities	4.3	3.5	5.5
Total	100.0	100.0	100.0

(*Not all of the table is included*.)

For the tables that refer to distributions of full-time and part-time employees respectively by age, discrepancies are small.

The same factors mentioned above will also explain discrepancies between the tables that show the distribution of full-time employees by industry. In general, it seems that the distribution of employees by sex and industry and sex and age are very similar.

Example 15.3-2: Coherence of the Eurosystem Household Finance and Consumption Survey (HFCS), and the EU-SILC data on income (European Central Bank, 2013, p. 98-101)

[This example gives a comprehensive account of how income is treated in two EU surveys.]

EU-SILC provides a useful benchmark for comparing income data of the HFCS. Being a household survey, EU-SILC is conducted for similar purposes and uses data collection methods similar to those of the HFCS. It should be acknowledged, though, that the HFCS aims at maximising the efficiency of the estimates of the wealthiest households, while the main target of the EU-SILC is low income households. This leads to different sampling strategies in these surveys...

The definitions of household and the target population are identical in both surveys. However, in Italy the EU- SILC definition of private households ("Cohabitants related through marriage, kinship, affinity, patronage and affection") is different from the one used in other countries and in the HFCS...

Some differences in the data collection methods can be observed between EU-SILC and HFCS. In seven countries, the main data collection method was the Computer Assisted Personal Interview (CAPI) for both EU- SILC and the HFCS. In Finland, both surveys use Computer Assisted Telephone Interviews (CATI). ...

In the HFCS, the income concept is gross income, i.e. taxes, social contributions and other transfers paid by households are not deducted from the income totals. Consequently, comparisons with external sources should only be made to similar income concepts, and not to after-tax income (disposable income). Data from EU-SILC enables a comparison to a concept of gross income that is identical with the HFCS one, with the exception of income from private use of a company car that is not included in the HFCS. The table below shows the correspondence between individual income items collected in the two surveys. For most individual items, EU- SILC definitions were applied as such to the HFCS, although some differences that are explained in the table below remain. Data on social transfers in EU-SILC are collected in a more detailed manner, while financial income is more detailed in the HFCS.

Correspondence table – Household gross income in HFCS and EU-SILC

	EU-SILC	HFCS
Employee cash or near cash income	Employee income	Comment
Income from private use of company car		Not included in HFCS
Cash benefits or losses from self-employment	Self-employment income	
Old-age benefits Survivors' benefits Disability benefits	Income from public pensions	
Pension from individual private plans	Income from private and occupational pensions	
Unemployment benefits	Income from unemployment benefits	Severance and termination payments and redundancy compensation included in other income in the HFCS.

…

The table below provides a comparison of the median household gross income between HFCS and EU-SILC. The coherence between the figures is very good, especially taking into account some differences in definitions.

Country	Median gross income HFCS, €	Median gross income EU-SILC, €	HFCS, % of EU-SILC
Belgium	34,000	35,000	97%
Germany	33,000	33,000	100%
Greece	22,000	24,000	92%
Spain	25,000	26,000	96%
France	29,000	36,000	81%
Italy	26,000	31,000	84%
Cyprus	33,000	34,000	97%
Luxembourg	65,000	66,000	98%
Malta	22,000	22,000	100%
Netherlands	41,000	43,000	95%
Austria	32,000	41,000	78%
Portugal	15,000	17,000	88%
Slovenia	18,000	23,000	78%
Slovakia	11,000	12,000	92%
Finland	36,000	36,000	100%

S.15.3.1 Coherence – subannual and annual statistics

Coherence between subannual and annual statistical outputs is a natural expectation on the part of users and yet the statistical processes producing them are often quite different. Thus, reasons for lack of coherence need to be assessed and explained.

The starting point for assessing the likely magnitude of differences due to lack of coherence is to compare subannual and annual estimates.

- If both annual and subannual estimates measure levels, then annual aggregates can be constructed from subannual estimates and compared to totals from the annual series.
- If one or other of the series produces only growth rates not levels, then comparison can be made of year over year growth rates.

If the differences thereby observed cannot be fully explained in terms of sampling error or other measures of accuracy then their explanation requires assessment of the possible causes by metadata comparison, as for all forms of coherence assessment.

S.15.3.2 Coherence – national accounts

S.15.3.2 FURTHER GUIDELINES

As previously noted, the National Accounts compilation process can reveal lack of coherence in data received from its various source statistical processes, whether they be direct surveys, register based surveys or indexes. Feedback from the National Accounts on the degree of incoherence and the adjustments that had to be made in order to bring the accounts into balance are excellent indicators of the accuracy and/or coherence of the statistical outputs received. They should be reported and should be a trigger for further investigation.

S.15.3.2 EXAMPLES

Example 15.3.2-1 National Accounts, EU level, [ESS-MH]

[This example gives detailed information on NA coherence and consistency issues.]

Within the system of national accounts there is full consistency between the domains: annual and quarterly national accounts, government accounts, sector accounts, financial accounts, regional accounts, supply and use tables. However, in practice full consistency may not always be possible and temporary discrepancies might occur. They are usually the result of vintage differences.

Primary statistics like structural business statistics (SBS), short term statistics (STS) and labour force statistics (LFS) are widely used as input for national accounts. However, there is no full consistency between these statistical domains and national accounts. Main reasons are differences in concepts/definitions and in coverage. Balance of payments is also used as an important source for national accounts. The definitions and coverage of balance of payments, as defined in the BPM6 manual, are fully harmonised with those in ESA 2010. Therefore, balance of payments variables are in principle fully coherent with the corresponding national accounts variables.

Although in theory the national accounts data should be harmonised between all national accounts domains, in practice there will be discrepancies in the data of several countries as well as in the euro area and EU aggregates. The reasons are the following.

Quarterly data of most countries are usually aligned to annual data once a year; inconsistencies between quarters and the corresponding year might appear for some time before the alignment.

National data for different national accounts domains are compiled at different points in time, which is usually related to availability of sources.

Usually the discrepancies between the national accounts domains only concern the most recent reference period; previous reference periods tend to be coherent between domains.

Example 15.3.2.2 Coherence between the Gross Value Added calculation in the Annual Business Survey and National Accounts (UK Office for National Statistics, 2012, p. 60-63)

[This example explains the conceptual difference between Gross Value Added in the National Accounts and the Annual Business Survey. By giving the exact definitions of both, a tool for reconciliation of the difference is provided. It should be noted that the adjustments done by the National Accounts do not have the same weight in all industries.]

The Annual Business Survey (ABS) publishes an approximate measure of Gross Value Added at basic prices (aGVA).

Gross Value Added (GVA) at basic prices is output at basic prices minus intermediate consumption at purchaser prices. The basic price is the amount receivable by the producer from the purchaser for a unit of a good or service minus any tax payable plus any subsidy receivable on that unit.

There are differences between the ABS approximate measure of GVA and the measure published by National Accounts. National Accounts carry out scope adjustments, coverage adjustments, conceptual and value adjustments such as subtracting taxes and adding subsidies not included in the ABS measure, quality adjustments and coherence adjustments. The National Accounts estimate of GVA uses input from a number of sources, and covers the whole UK economy, whereas ABS does not include some parts of the agriculture and financial activities sectors, or public administration and defence. ABS total aGVA is two-thirds of the National Accounts whole economy GVA, because of these differences in scope, coverage and calculation.

No real (inflation-adjusted) estimates of regional GVA are published in the National Accounts, however, nominal (non-inflation-adjusted) regional GVA and approximate regional GVA at basic prices are published by Regional Accounts and ABS respectively.

The calculation of approximate GVA in the ABS

Approximate GVA is calculated as follows. The variables in bold are those published in the ABS statistical releases. Other variables are available on request from abs@ons.gsi.gov.uk.

aGVA = output at basic prices – intermediate consumption

= total turnover

+ movement in total stocks

+ work of a capital nature carried out by own staff

+ value of insurance claims received

+ other subsidies received

+ amounts paid in business rates

+ amounts paid in vehicle excise duty

- total purchases

- amounts received through the Work Programme (formerly the Welfare to Work Scheme)

- total net taxes (note: for service industries, this is total taxes, not total net taxes)

The National Accounts calculation of GVA

The official UK estimate of GVA published by National Accounts includes, in addition to the ABS variables:

- inclusion of own account work (i.e. work consumed by the producer, for example, farmers producing crops to feed their own animals, or computer software written in-house) and non-market output. These are conceptually out of scope of the ABS and are calculated from other survey data supplements to ABS data, other surveys and administrative data, to cover the whole economy. This includes public corporations from company accounts and data on the public sector
- adjustments to output to account for income in kind, own account computer software, work in progress, and, for total sales, the addition of taxes less subsidies on production
- an undercoverage adjustment to output, to account for the one per cent of businesses not covered by the IDBR in terms of economic activity
- adjustments to intermediate consumption, including the addition of insurance premium supplements and financial intermediation services indirectly measured (FISIM)

These additional components account for the differences between the published values of GVA and aGVA.

Figure 9.1 and Figure 9.2 below show the size of the components of the National Accounts estimations of output and intermediate consumption. ABS total sales contribute the largest component of total output (around 70 per cent in 2010). Other key components of total output include non-market output and own account output. ABS total purchases contribute the largest component of intermediate consumption (around 80 per cent in 2010).

Figure 9.1 Components of the National Accounts estimate of total output (whole economy)

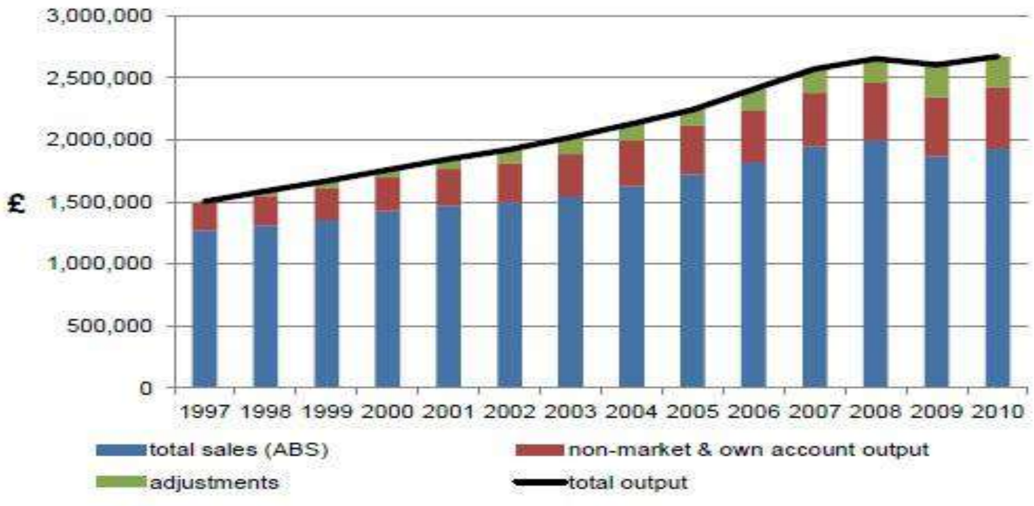
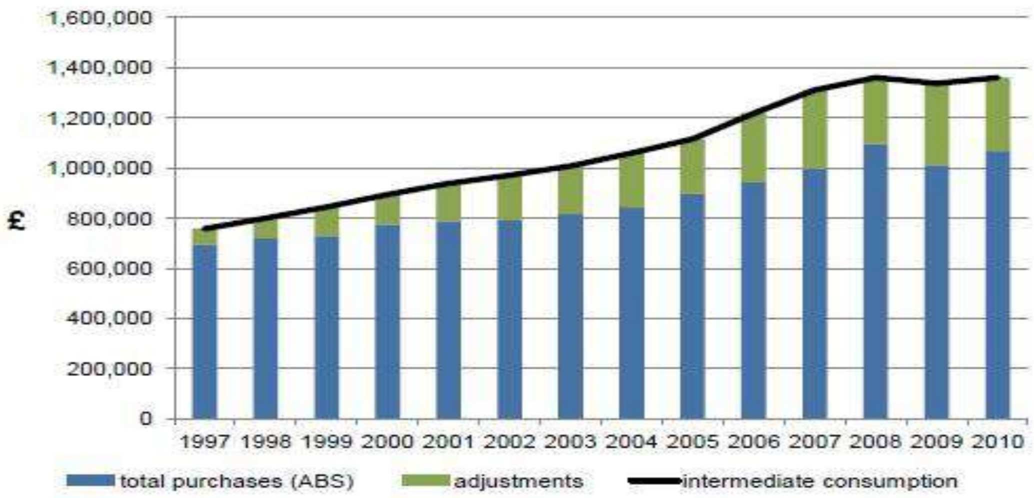


Figure 9.2 Components of the National Accounts estimation of intermediate consumption



S.15.4 Coherence – internal

SIMS	Concept Name	Definition	Summary Guidelines
S.15.4	Coherence – internal	The extent to which statistics are consistent within a given data set.	Each set of outputs should be internally consistent. If statistical outputs within the data set in question are not consistent, any resulting lack of coherence in the output of the statistical process itself should be stated as well as a brief explanation of the reasons for publishing such results.

S.15.4 FURTHER GUIDELINES

Based on a given statistical process, statistical outputs are published. Each set of outputs should be internally consistent, meaning that all the appropriate arithmetic and accounting identities should be observed. However, this is not always the case. One reason for this is that the process comprises data from different sources. Another possibility is that otherwise efficient estimation methods have this drawback. In these circumstances a brief explanation should be given to users and also be reflected in a producer report, with the reasons for publishing non-coherent results explained.

Note that internal consistency can refer to either (i) numerical consistency within one table or (ii) consistency between different tables or presentations that are derived from the same data set. In both cases, any lack of consistency should be explained.

Special attention should be given to coherence problems in multisource statistics, where outputs originating from different surveys and administrative data processes may have different definitions or be based on different methodologies.

Similarly, attention should be paid to coherence problems in the national accounts framework, which, by definition, represents a consistent set of accounts. In case of inconsistency, where balanced and consistent datasets are required, it is appropriate to follow up the guidance provided for the national accounts domain.

S.16 (PART II) Cost and burden

SIMS	Concept Name	Definition	Guidelines
S.16	Cost and burden	Cost associated with the collection and production of a statistical product and burden on respondents.	**Cost** Provide annual operational costs of the process, with breakdown by major cost component. Describe recent efforts to improve efficiency and comment on the extent to which information and communication technology is used. *European level* Describe recent initiatives and efforts to improve efficiency at the European level. **Burden** Provide an estimate of the respondent burden imposed by the process. Describe all the means taken to minimise burden. *European level* Describe recent initiatives and efforts to minimise burden at the European level.
This concept is included in both ESMS based (user) reports and in ESQRS based (producer) reports. In ESQRS based reports it is ESQRS Concept 10.			

S.16 Background

Measurement and monitoring of cost and burden is required in reviewing and improving performance and in maintaining user relations. In addition, whilst cost and respondent burden are not aspects of output quality, there is a balance between cost and burden on the one hand, and the output quality components on the other. Expressed differently, cost and burden are constraints on output quality. Thus, assessment of cost and burden is also required to underpin decisions regarding output quality.

Measuring cost

The capacity to calculate cost is required for efficient management in general, and for performance assessment in particular, for all types of statistical process. Cost benefit analyses help determine the appropriate balance between costs and benefits in terms of the output quality. The ESS Quality Assurance Framework (ESS QAF) presents methods at both the institutional and the statistical product/process level for measuring costs and the trade-off between quality and costs.

- *Full cost approach*. This involves combining the *direct costs* (that can be directly attributed to the process) and an appropriate proportion of *indirect costs, i.e., costs that are shared* with other statistical processes (for example, the costs of a statistical business register and standards management) and *overhead costs* (office space, utility bills etc).

- *Direct Cost Approach*. As indirect costs are more difficult to measure and allocate than direct costs, a simpler approach is to measure only the principal direct costs, typically based on the working days spent on the statistical process.

The appropriate choice of approach depends on the cost accounting system in the NSA. A full-cost approach is preferred if the data are available.

Eurostat has developed its *Guiding Principles for Cost Assessment Surveys* to underpin the conduct of surveys of statistical process production costs in NSAs using the full cost approach, and analysis of the cost of European statistics by product.

In conjunction with key users and Member State NSAs, Eurostat undertakes systematic *rolling reviews* of its statistical work. Their objectives are to investigate issues, such as whether the production processes are organised efficiently, what the costs to Eurostat, Member States and EFTA countries are, and whether the work could be done more efficiently. They make use of a range of assessment tools, including an assessment checklist, user surveys and partner surveys

Measuring Burden

Burden applies only to a survey, i.e., a statistical process involving direct collection of data from respondents about themselves. It does not apply to acquisition of administrative data or micro-level data from another survey, nor to access to, or compilation of macro-level data.

The requirement to measure cost and burden is exemplified in Regulation (EC) No 295/2008 of the European Parliament and of the Council of 11 March 2008 concerning Structural Business Statistics. Article 6 states *"Quality evaluation shall be carried out comparing the benefits of the availability of the data with the costs of collection and the burden on business, especially on small enterprises".*

The capacity to calculate burden is useful in the context of respondent relations. Completion of a questionnaire imposes a burden on respondents that has to be balanced against the benefits of the data thus obtained. In trying to maintain response rates and minimise response error, it is crucial that the organisation measures burden, keeps it to a minimum and assures respondents that it is doing so.

In some domains there are specific regulations, For example, Article 14 of *Council Regulation No 1165/98 of 19 May 1998 concerning Short-term Statistics (Council Regulation (EC), 1998)* states "

> *"The Commission shall (…) submit a Report (…) on the statistics compiled (…) and in particular on (…) the burden on business."*

The ESS Quality Assurance Framework (ESS QAF) presents methods at institutional and the statistical product/process level for justifying, analysing and measuring burden.

The usual measure of burden is the total number of hours spent by respondents in completing questionnaires for the survey during the course of a year.

- Typically, it is calculated by multiplying the number of completed questionnaires by an estimate of the average time required in staff-hours for the responding organisation to complete and submit its response, multiplied by the number of survey cycles during the year.

- Sometimes the *intended burden* is used, computed using the total sample size rather than the number of completed responses.

- The average time taken to complete the response may be estimated from a question on the questionnaire asking for the completion time, or may be observed as in the case of a personal interview, or may be derived as in the case of a web-based survey, or may simply be estimated by an in-house exercise.

Sometimes burden is measured as a cost to responding organisations, typically derived as the measurement of hours multiplied by a notional cost/hour for the respondents. This measure is not greatly used, first because it depends upon a, usually very flimsy, estimate of the hourly rate, second because it is more likely to inflame respondents.

Some efforts may also be made to define and measure *perceived* burden, i.e., the burden as perceived by respondents. This may include the clarity of the questions, the ease of navigating through the questionnaire, and the ease with which respondents can access the information required to answer the questions. Measurement of perceived burden can by undertaken using a small set of questions at the end of the questionnaire, or subsequently by follow-up contact with a sub-sample of the surveyed population.

S.16 FURTHER GUIDELINES

Reporting cost

Whilst users may not be particularly interested in the costs of production, they need to be convinced that resources are adequate and are efficiently used. Thus, at a minimum a report should:

- state the annual operational cost; and
- outline recent and current efforts to improve efficiency.

The report can provide more detail if this seems warranted. It can do any, or all, of the following:

- report the annual operational cost for the statistical process, with breakdown by major cost component, using the full-cost approach if the appropriate data are available, otherwise the direct cost approach;
- describe the procedures for internal assessment of efficiency and for independent external assessment of efficiency;
- describe recent and current efforts made to improve efficiency;
- describe the extent to which routine operations, in particular data capture, coding, validation and imputation, are automated through effective use of information and communications technology (ICT).

Reporting burden

Users may not be especially interested in respondent burden but they need to be aware that data collection by survey imposes a burden, and that this burden is has to be justified by the use made of the data collected. Thus, at a minimum, the report should:

- state the estimated annual respondent burden in hours;
- outline the use (if any) made of administrative data in order to reduce burden; and
- outline other recent and current efforts to reduce respondent burden.

The report can provide more detail if this seems warranted. It can cover any, or all, of the following:

- how respondent burden is calculated;
- whether the NSA has set a burden reduction target for the survey being reported, and, if so, what it is;
- whether the NSA has set a burden reduction target for a group of surveys, including the survey being reported, or for all surveys, and, if so, what it is;
- how the range and detail of data collected have been limited to what is necessary to meet the stated objectives of the survey;
- the administrative data sources considered in place of, or to supplement the survey, and why

they were not used for the data collected by the survey;

- (for surveys of businesses) the efforts made to ensure that data sought were readily available from the usual business accounts;
- the reasons for omission(s), if any, of significant reporting mechanisms such as mailout-mailback, web, computer assisted personal interviewing (CAPI), and/or computer assisted telephone interviewing (CATI);
- whether reporting burden on individual respondents was equitably spread to the extent possible by minimizing the overlap with other surveys, and if so how; and
- recent and current efforts made to reduce the burden associated with the survey.

S.16 EXAMPLES

Example S.16-1 Turnover in services, 2014

Office for National Statistics, United Kingdom [ESS-MH]

[This example illustrates minimal report.]

Cost to produce the services turnover is equivalent to that of the Index of services.

Number of units 26,000;

Cost (NSA) hours per year 2,932,205;

Burden (respondents) hours per year 107,756

Example S.16-2 Census 2011: England and Wales

Office for National Statistics, United Kingdom [ESS-MH]

[This example illustrates a comprehensive description of what was done to minimise facilitate reporting and thus minimise the burden. It does not include an estimate of the burden.]

Minimising the burden on respondents was a key consideration in the planning and design of the 2011 Census questionnaire. In dealing with requests for the inclusion of new questions, consideration was given to their number, complexity and acceptability and assessments made about the potential quality of response.

A Privacy Impact Assessment (PIA), published in 2009, examined a wide range of such issues and concluded that ‘The burden on each household in completing the census questionnaire is considered reasonable. For a typical household of four people the once-a-decade Census should take around 30-40 minutes.

To support people who did not speak English as a first language, a translation booklet was produced by combining the census questionnaire and information leaflet. This was translated into 56 languages for guidance although respondents still had to complete the English version. Printed copies were available on request from the census helpline (each language had its own dedicated phone number), the online help facility and from census collectors. In addition they could be viewed and downloaded from the census website.

For people with sensory impairments, a number of materials were available:

- audio cassette tape, CD and video and audio clips on the website;
- large print questionnaire and information leaflet;

- easy read information leaflet;
- Braille guidance booklet;
- British sign language DVD and clips on the website; and
- audio version of the information leaflet in the top ten most widely spoken languages.

Example S.16-3 Census 2011: England and Wales - Costs

Office for National Statistics, United Kingdom [ESS-MH]

[This example illustrates comprehensive coverage of costs.]

The 2011 Census in England and Wales cost approximately £482 million. This equates to less than £1 per person per year over the 10-year planning and operational cycle of the census.

While the 2011 Census cost around twice that of the 2001, around one third of the increase was due to inflation and the larger number of people to count. In real terms, this equates to around 35 per cent more than the last census.

A number of automated processes were introduced for the 2011 Census to reduce overall costs and increase effectiveness. These included providing the option to complete the questionnaire online and the use of a comprehensive address list allowing questionnaires to be delivered and returned by post. As a result, the number of enumerators required in 2011 (approximately 35,000) was around half of the number in 2001.

Census results underpin the planning of services and the allocation of resources to local communities. A project to assess the benefits of the census in England and Wales has been under way since 2012 and a number of case studies are published on the ONS website http://www.ons.gov.uk/ons/guide-method/census/2011/2011-census-benefits/index.html

Example S.16-4 Commercial and industrial waste survey (Wales only)

Department for Environment, Food and Rural Affairs (Defra), UK [ESS-MH]

[This example illustrates a succinct and informative report on burden.]

Type and total number of respondents

15,679 calls made, including calls which resulted in there being wrong number, no answer, contact busy, refusal to participate, call-back being required, or appointment being secured.

Actual number of respondents and time required for response

1,540 completed surveys, 1 hour (average).

Measures taken to minimise the burden

Trained interviewers, Clustering visits, Appointment reminders with details of information for the business to adequately prepare in advance of the survey visit, Electronic form, delivered using laptop, designed to collect information required with minimum questions.

Notes

The burden reported is only that which occurred in respect of the reference year (2012). Some of the data used for 2012 estimates has been extrapolated by applying business population data to waste factors gained from surveys carried out in previous reporting periods. In these cases, no further burden on respondents has been reported for 2012. Administrative returns completed by government officials or local authorities (e.g. WasteDataFlow) are assumed to incur zero additional burden for statistical purposes due to the data collected being essential for the operators themselves. Site returns for fully permitted treatment facilities and treatment facilities licensed under complex exemptions (Scotland and Northern Ireland) are assumed to incur zero additional burden for statistical purposes as they are required for regulatory activities.

S.17 Data revision

SIMS	Concept name	Definition	Guidelines
S.17	Data revision	Any change in a value of a statistic released to the public.	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and its sub-concepts are included in ESMS based (user) reports.

In ESQRS based (producer) reports the concept and its sub-concepts are included within ESQRS Concept 6 Accuracy and Reliability.

The difference in treatment between ESMS and ESQRS of quality and performance indicator A6 is indicated by (U) and (P).

S.17 Background

From a quality perspective, reliability is the result of data revision policies and practices. Thus, the information in S.17.2 Data Revision Practice should be regarded as being the *Reliability* complement of the information about *Accuracy* reported in S.13 Accuracy and Reliability. (Indeed, in ESQRS, which provides the format for producer reports, Data Revision Policy and Practice are reported in the section on Accuracy and Reliability.)

Data revisions should follow standard, well-established and transparent policies and procedures. Pre-announcements are desirable.

Note that the *size of a revision* of an item is defined as the difference between a later and an earlier estimate.

The ESS guidelines on revision policy for Principal European Economic Indicators (PEEIs) provide a very comprehensive description of the reasons for revision, and revision policies and methods.

The essential difference between what is reported under S.17.1 Data revision policy and S.17.2 Data revision practice is that

- the former refers to the *general* revision policies and procedures that the NSA has or might be expected to follow; and
- the latter refers to the *actual* policy and procedures applied to the outputs of the statistical

process that is the subject of the report.

S.17.1 Data revision - policy

SIMS	Concept name	Definition	Guidelines
S.17.1	Data revision – policy	Policy aimed at ensuring the transparency of disseminated data, whereby preliminary data are compiled that are later revised.	Describe the data revision policy applicable to data output from the statistical process being reported. In so far as they are relevant to the process being reported, summarise the general procedures for treatment of planned revisions, benchmark revisions, unplanned revisions, and revisions due to conceptual and/or methodological changes. *European level* Describe the data revision policy and procedures at European level.

S.17.1 FURTHER GUIDELINES

State whether the organisation has a *general data revision policy* covering all or many statistical processes. If it does, provide a link to a document detailing the policy.

Describe the actual data revision policy for the statistical process being reported in so far as it differs from, or extends, the general revision policy, and the reasons for any differences or extensions.

For any of the following situations that could be relevant to the statistical process being reported:

- describe the *general procedure* (if any) for handling a scheduled revision of the form *preliminary* to *final*, where the difference is due to additional data being obtained;
- describe the *general procedure* (if any) for handling a revision that is due to revision of a benchmark, for example, re-benchmarking quarterly production based on the results of an economic census;
- describe the *general procedure* (if any) or handling an unplanned revision required to correct a *mistake*, i.e., error found in the data after release;
- describe the *general procedure* (if any) for handling conceptual or methodological changes that cause changes in data values requiring revision of historical data, or a break in series.

S.17.1 EXAMPLES

Example S.17.1-1 Harmonised index of consumer prices (HICP), Monthly, 2017, Italian National Institute of Statistics

[This example contains a succinct description of the circumstances under which a revision can take place.]

HICP series, including back data, are revisable under the terms set in Commission Regulation (EC) No 1921/2001 of 28 September 2001. The published HICP data may be revised for mistakes, new or improved information, and changes in the system of harmonised rules.

The first data dissemination concerns provisional data for the latest month. These are confirmed or revised to the final figures within the second week of the following month. Other major revisions are generally released with explanatory notes in the press release. Methodological changes are explained with the first release of data affected by such changes.

Example S.17.1-2 Retail sales Index, Monthly, 2016

Office for National statistics, United Kingdom, [ESS-MH]

[This example contains a more comprehensive description of the circumstances under which a revision can take place.]

Retail Sales non-seasonally adjusted data is revised as needed. This typically occurs due to late data returns, updated respondent information, replacing adjustments with actual data, and reclassifications of respondents to the appropriate category either within or out of retail. Changes are not made to any non-seasonally adjusted data prior to 2001. Revisions to non-seasonally adjusted data will directly impact on the seasonally adjusted estimates.

Revisions and sampling variations are a consequence of the trade-off between timeliness and accuracy. All estimates are subject to statistical error which refers to the uncertainty inherent in any process or calculation that uses sampling, estimation or modelling. Estimates for the most recent month are provisional and subject to revision because of:

- late responses to the Monthly Business Survey - Retail Sales Index;
- revisions to seasonal adjustment factors which are re-estimated every month and reviewed annually;
- changes from the annual seasonal adjustment review; and
- annual updating of the business register that forms the basis for the sample for the RSI (usually occurring in January) and
- other methodological improvements.

Policy regarding revisions to Retail sales is available in the monthly release on the website.

Example S.17.1-3 National Accounts, 2017 EU level, [ESS-MH]

[This example illustrates a comprehensive description at European level.]

Data revision - policy

Macroeconomic statistics, such as national accounts, the balance of payments and the international investment position, are produced from a large variety of data sources. These data sources are reconciled using an approach based on an agreed set of international guidelines. The sources used to estimate macroeconomic aggregates are provided with varying degrees of timeliness, taking up to three years or more in the case of structural sources. As users need national and international data as fast as possible, particularly on certain key aggregates like gross domestic product (GDP), data are produced using the sources and related indicators that are more readily available. As more complete data are obtained from these sources in due course and the structural sources are made available, the statistics are updated to incorporate the new information.

Such revisions of macroeconomic statistics are necessary to improve quality, but can be inconvenient for users. To minimise this inconvenience, revisions should ideally be coordinated within one country, across different statistics, and then across countries. International comparability – and the compilation of EU and euro area aggregate statistics – is hampered when different revision policy schemes are applied in different countries. As the schedule of revision of national accounts and balance of payments statistics varies from country to country, this creates inconsistencies among different statistical domains.

A distinction should be made between 'routine' revisions and 'major' or 'benchmark' revisions.

- Routine revisions refer to the changes made to the economic data published initially and to its subsequent releases for a particular reference quarter or year.

- Benchmark revision is carried out at much longer time intervals. Its purpose is to incorporate the main new data sources and major changes in international statistical methodology (such as ESA 2010 or BPM6). In benchmark revision, many years are open for revision in order to create the longest possible consistent time series.

The European Statistical System (ESS) and the European System of Central Banks (ESCB) try to strike the right balance between incorporating the necessary statistical revisions and maintaining an acceptable degree of consistency across domains and countries. To this end, the two systems have worked together to draw up guidelines for a harmonised revision policy for macroeconomic statistics. National accounts data, like any other statistics, are subject to continuous revisions as new input data become available, and users require most up-to-date information.

In order to address users' concerns about the lack of coordination of revisions, National Statistical Offices and National Central Banks agreed to gradually implement a common harmonised European revision policy for national accounts and balance of payments statistics. This policy includes differentiated guidelines regarding the timing and depth of revisions of quarterly and annual data. It aims to improve adherence to the twofold principle of alignment between statistical domains at national level and coordinated alignment across countries at EU level.

In addition to source availability, national release calendars also influence and sometimes constrain the design of the common revision policy. This is because national release calendars are determined by national policy needs, as well as by the statistical reporting obligations that exist in the EU. The guidelines also need to be aligned with the schedule of the various administrative uses of national accounts figures, namely the Macroeconomic Imbalances Procedure in the context of the European Semester, gross national income (GNI) for EU budget purposes, and general government deficit and debt figures for fiscal policy.

Implementation is now underway and will take some time. The National Statistical Offices and the National Central Banks are not legally bound by this common policy, but voluntarily agree to it and commit to gradually implement it with the aim of delivering more consistent statistics to users. The level of adherence to the guidelines of Member States' revision policies will be monitored regularly.

S.17.2 Data revision - practice

SIMS	Concept name	Definition	Guidelines
S.17.2	Data revision - practice	Information on the data revision practice.	If there are no revisions to report for the statistical process that is the subject of the report, state this and close the reporting of this concept. Report the reasons and schedule for planned revisions (if any). Explain A6 Data revision - average size indicator. Report A6 at a level of detail appropriate for user or producer report. In the absence of data to compile A6 Data revision, make a qualitative assessment of the average size of planned revisions and their direction based on historical data. Describe the main reasons for unplanned revisions that have occurred, and the actions taken to prevent the need for such revisions in the future. *European level* Describe the planned and unplanned revisions at European level.
	A6. Data revision - average size (U)	The average over a time period of the revisions of a key item, for user report.	
S.17.2. 1	A6. Data revision - average size (P)	The average over a time period of the revisions of a key item, for producer report.	

SIMS	Concept name	Definition	Guidelines
			Report A6 aggregated over countries.

S.17.2 FURTHER GUIDELINES

Planned revisions

Report deviations from the schedule (if any) and why they occurred, to the extent that this is not reported under S.14.2.

Report Indicator A6 over the most recent period for which data are available, preferably at least a three year period. The choices of period and compilation formula are at the discretion of the author. For producer reports provide A6 for more indicators.

Unplanned revisions

Describe the main reasons for any other (i.e., unplanned) revisions that have occurred over the past three years or since the previous report (whichever is shorter), for example, due to the need to correct a mistake or take into account a change in methodology or in concepts.

In the case of a mistake, describe the mistake and measures taken to guard against making similar mistakes in the future.

S.17.2: EXAMPLES

Example S.17.2-1 Foreign Direct Investment and Lithuanian Direct Investment Abroad: Revision Procedure

[This example contains a comprehensive description of planned revisions.]

Statistics Lithuania carries out an exhaustive annual and a sample quarterly foreign direct investment (FDI) survey. Primary statistical data are provided for the Bank of Lithuania (BoL). The BoL, using additional statistical data sources, calculates aggregated statistical indicators and revises statistical information for the previous periods.

FDI statistical indicators are revised according to a statistical indicator revision calendar. A short-term revision of FDI statistical indicators is carried out every year. When needed, statistical indicators of previous quarters are also revised. After annual FDI statistics have been published in September, quarterly statistical information for the corresponding year is recalculated. After a change in the methodology in 2014 (Balance of Payments Manual, Sixth Edition), a long-term revision was performed. Annual and quarterly FDI statistics were recalculated from 2004.

Responsible division: Construction and Investment Statistics Division.

Reasons for revision

The main reason for short-term revision of quarterly statistical information is the supplementation and refinement of primary statistical data: respondents’ errors are corrected, respondents revise statistical data for the previous periods, statistical data are compared against other statistical data sources and revised respectively.

Methods

A long-term revision was carried out because of the coming into force of a new – sixth – edition of the Balance of Payments Manual, which changes the FDI calculation methodology.

Level of detail and periodicity

All statistical information of a corresponding period is revised: foreign direct investment and Lithuanian direct investment abroad by country, economic activity, and regional statistics. Previous quarters' data are revised every quarter, if needed.

Comparability

The time series of foreign direct investment and Lithuanian direct investment abroad are available from 1997 to 2004 and from 2004 up to date.

Revision calendar

An approved Calendar of Scheduled Revisions of Statistical Indicators for the current year is available on the Official Statistics Portal.

Dissemination of results

Revised FDI statistics are published in a joint SL and BoL press release, on the Official Statistics Portal, in the Eurostat's database, in the BoL database, in the BoL publication Monthly Bulletin, in SL publications (Economic and Social Development.

Example S.17.2-2 Job Vacancy Statistics, Quarterly, 2015

Italian National Institute of Statistics [ESS-MH]

[This example contains a comprehensive description of planned revisions.]

Each year in June, when the data for the first quarter of a year are published, revisions of the previous eight quarters are also disseminated.

Reasons for revisions:

- inclusion of late responses (i.e. arrived after the first publication of the quarterly figures);
- the revision of the auxiliary sources which are used for editing and imputation and calibration.

Furthermore, all data referring to periods preceding 2016Q1 (originally covering employees with the exclusion of managers) have been appropriately revised to avoid any break with to those from 2016Q1 onwards (covering also managers). The revised data have been transmitted together with those for 2016Q1. (See sheet "June 2016 revisions" in the excel file "Tables and graph" in the annex section for a comparison of figures for 2015 transmitted for the 70-day Regulation deadline with those disseminated in June 2016.)

Furthermore, the preliminary confidential estimates for the entire reference population, which are transmitted to Eurostat within 45 days from the end of the reference quarter, can be revised when the data for the Nace Rev. 2 sections for that quarter are produced for the 70-day regulation deadline, for the reasons indicated above. (See sheet "revisions 45-70 day deadlines" in the excel file "Tables and graph" in the annex section.)

Example S.17.2-3 National Accounts, 2017 EU level, [ESS-MH]

[This example illustrates a comprehensive description at European level.]

Data revision - practice

The harmonised policy applies to the national revision practices of national accounts and balance of payments statistics at both the annual and quarterly frequencies.

Between one and two months after the end of a given quarter, Member States publish the first ('flash') estimates of the main national accounts aggregates, including GDP. Around two months after that quarter, new estimates ('preliminary estimates') of these main aggregates are published, which may revise the flash estimates and contain some additional detail. Finally, around three months after the quarter, a complete set of quarterly national accounts are published, including institutional sector accounts, and quarterly balance of payments data. These estimates could subsequently be revised again in future quarters and years to align them with new annual data.

Quarterly estimates are usually revised retrospectively for up to four years, although the policy allows unlimited revisions in quarter 3.

Routine revisions for annual data. The sequence of publications/revisions regarding annual data, in calendar year t, is as follows:

- Quarter 1: First estimate of annual data for year t-1, usually corresponding to the sum of the quarterly figures released;
- Quarter 2: First possible revision of annual data for year t-1 to include revisions of quarters of t-1;
- Quarter 3: First estimate of annual data for year t-1 based on both annual and first available sources.
- Quarter 4: Exceptional revisions of annual data for year t-1 may be carried out to take into account changes following Excessive Deficit Procedure and Own Resources notifications (GNI Expert Group).

Annual estimates are usually revised retrospectively for up to four years, although the policy allows unlimited revisions in quarter 3. Some countries are obliged to publish the initial annual data on t-1 in quarter 2 instead of quarter 3.

Major or benchmark revisions

In 2014, all Member States disseminated revised data according to ESA 2010. The agreed guidelines specify that Member States should disseminate the results of the next benchmark revisions in 2019 and 2024 respectively. It is expected that most EU countries will be able to meet the 2019 target and that all EU countries will undertake the subsequent benchmark revision in 2024. Disseminating the results of a benchmark revision always involves revising all, or at least a large part of the time series.

While revisions should be seen as a process to progressively improve the quality of national accounts as e.g. better sources and/or methods become available, the availability of metadata on revisions is a key element for understanding national accounts data and revisions between subsequent releases.

Therefore, information on the main reasons for revisions and their nature (new source data available, new methods, etc.) as well as possibly quantitative and qualitative assessment on the average size of revisions and their direction based on historical data is required.

For some national accounts domains Eurostat publishes information on its publication of national data and on its revision policy for European aggregates on its website. For example, information on the revision policy for main aggregates can be found here.

S.18 (PART II) Statistical processing

SIMS	Concept name	Definition	Guidelines
S.18	Statistical processing	(Defined by its sub-concepts)	(Information relating to this concept is provided by reporting on its sub-concepts.)

This concept and all its sub-concepts are included in ESQRS based (producer) reports. The concept is ESQRS Concept 3. However Sub-concept S.18.5.1 is ESQRS Sub-concept 6.3.4.1 and Sub-concept S.18.6.1 is ESQRS Sub-concept 6.4.

This concept and all but the two sub-concepts referenced above are included in ESMS based (user) reports. The two concepts included only in ESQRS reports are indicated by (P).

S.18 Background

As noted in Section 5.8, a producer/user report is just one type of documentation for a statistical process. Other types of documentation are produced. In this respect national practices differ widely. Some countries produce technical reports describing statistical methods and operational procedures in detail, for example, including sampling procedures, estimation formulae, measurement methods and editing rules. When such technical documentation exists and is readily accessible, the producer/user report should simply summarise it and provide a link to it. However, when such documentation is not available, information on methods and procedures should be included in the report itself to provide context.

S.18 is the place where information about methods and procedures, covering the design, collect, process, and analyse phases (according to the GSBPM) of a statistical process can be presented. (Information about the dissemination phase can be presented in S.7, S.8, S.9 and S.10; information about the user needs and evaluation phases in S12.1 and S11.2 respectively.)

The level of detail provided should be commensurate with the particular purposes of the report.

S.18.1 Source data

SIMS	Concept name	Definition	Guidelines
S.18.1	Source data	Characteristics and components of the raw statistical data used for compiling statistical aggregates.	Indicate if the data are based on a survey, administrative data, multiple data sources, or macro-aggregates. In the event of multisource or macro-aggregate processes describe each source. For each survey source, report the survey population, cross referencing the description of the target population presented in S.03.6, and summarise the sample design. For each dataset from an administrative source, summarise the source, its primary purpose, and the most important data items acquired. *European level* Provide an overview of the sources used across countries.

S.18.1 FURTHER GUIDELINES

(The types of statistical process referenced below are defined and described in Chapter 4, Part I.)

All types of statistical process

Describe the design and development activities associated with the process, including any practical research work that was needed to define the statistical outputs, concepts, methods, collection instruments and/or operational procedures.

Describe the extent to which the design made use of relevant international and national standard definitions, classifications and methods.

Probability survey or census survey

In so far as they are not discussed elsewhere:

- describe methods by which the survey frame was obtained or created, including efforts to prevent overcoverage, undercoverage, and inaccurate unit classification and contact details.
- state the numbers of units in survey frame, in total, and for high level breakdowns.

Describe the sample design, including the statistical unit(s), target population, survey population, and survey frame.

For a probability survey, describe the stratification, allocation and sample selection procedures.

State the numbers of units in survey frame and sample, in total, and (for stratified surveys) in the most significant strata.

Non-probability survey

Describe the rationale for not using probability sampling and any assessment that was made of the sampling procedure that would support inferences about the population from the sample.

Describe the actual approach for the selection of sampling units.

For *quota surveys*, state the variables used for setting the quotas and describe how the interviewers were instructed and trained in selecting the final sample units.

For *subjective cut-off surveys*, state what variables and criteria were used to identify the largest or most important sampling units and describe the procedures used.

For *voluntary surveys*, describe the mechanism through which respondents were engaged in the survey, and provide a description of the final sample in terms of potentially important background variables, such as geographic location, sex, age, social status variables.

Administrative data process

Describe the source of the data, process by which the data were produced, and the process by which they were acquired by the NSA.

Describe the units and the population to which the data refer, the data content (variables and characteristics) and their relationships to the target statistical concepts.

Multisource process

Describe how the process is composed in terms of all its components. Describe each component in sufficient detail taking its overall importance for the outputs into account.

Macro-aggregate compilation process

Describe the source of each input dataset and process by which it was produced and acquired.

Describe the data content (variables and characteristics) and their relationships to any target statistical concepts.

S.18.1 EXAMPLES

Example S.18.1-1 House price index and value of housing transactions, 2017, DESTATIS, Germany [ESS-MH]

[This example contains an overview of sources, also data items and procedures.]

New (turnkey-ready) dwellings data and existing dwellings data are collected from the local Expert Committees for Property Valuation. The main tasks of these expert committees consist on the one hand in collecting data on real transaction prices (both cash and mortgage) and characteristics of buildings and dwellings, and on the other hand in estimating current market values of dwellings and land.

Type of data set: Transaction prices (both cash and mortgage) for new (turnkey-ready) dwellings and existing dwellings (census data in areas where possible).

Characteristics/Variables in the data set: Existing and new turnkey-ready dwellings: type of dwelling (single-, two-family house, freehold flat), type of house (free-standing, terraced, semidetached), type of construction (conventionally built, prefabricated), date of purchase, total purchase price, age of dwelling, size of plot of land, size of living area, proportionate price of plot of land, standard land value ('Bodenrichtwert'), furnishing/luxury elements (kitchen, sauna/swimming-pool, attic storey), car parking facilities, characteristics of location (state, district, municipality; general rating of location: simple/medium/good), number of rooms/floors.

House Price Index: The weight, i.e. the total expenditure (total revenue) in the residential property market is derived from data taken from GEWOS (Institute for City, Regional and Housing Research) as well as from DEGI (German Association for Real Estate Investment Funds).

Price index for (the purchase of) new dwellings: weight is derived from NA data (gross fixed capital formation according to housing). The additional mark-up for land is derived from data on revenues on residential building land, published by GEWOS.

Existing dwellings index: Houses and flats are weighted by expenditures of all the respective transactions (transaction totals) in the base year. The federal states are weighted by the population distribution according to the respective federal state.

Example S.18.1-2 Labour Force Survey 2016

National Institute of Statistics and Economic Studies, France [ESS-MH]

[This example contains a very detailed description of methods. Only the first third of the report is included below. There is further description of stratification, rotation, sample sizes and rates, use of subsamples (wave approach) and calculation of weights.]

Sample design

- The sample in Metropolitan France is a two stage stratified sample of dwellings. In overseas departments, the sample is a one stage stratified sample of dwellings. Dwellings are uniformly distributed over reference weeks of the year.

Sampling frame

- Since Q4 2011 the base used for the metropolitan sample is tax registers only. The sample is updated each year with new information and a sample of new dwellings is added.
- For overseas departments, the base is the French annual population census. Q4 of year N and Q1, Q2, Q3 of year N+1 are selected from the population census of year N-3.
- The sampling frame is updated each year.

Primary sampling unit

- Geographic sectors in Metropolitan France. No PSU in overseas departments

Final sampling unit

- Dwellings

First (and intermediate) stage sampling method

- For Metropolitan France, the sampling design consists in a selection of around 3200 geographic sectors with a stratified and balanced method. The stratification is carried out by NUTS2. The balanced sampling uses the following variables: age, incomes, type of dwelling, type of urban/rural areas, resident status (owner/tenant).
  - For overseas departments, the sample is composed of dwellings selected through a stratified systematic sampling (systematic sampling with equal inclusion probabilities, within geographic strata which form a partition of the territory).

Final stage sampling method

- The sample unit is the dwelling: in each sampled area, every person living in its main residence is surveyed.
- For Metropolitan France, the sectors are cut into 6 clusters of nearby dwellings, in such a way that there be around 20 main residences in each cluster. Inside the sectors, each cluster is randomly assigned a number between 1 and 6; this number determines when the cluster enters the sample, each cluster is interviewed 6 times and then replaced by another cluster of the same sector.
- For overseas departments, dwellings are directly selected within strata through a systematic sampling with equal inclusion probabilities. Strata sample sizes are proportional to the total numbers of main residences in the strata.

Example S.18.1-3 Gross nutrient balance 2016 [ESS-MH]

[This is a European level example illustrating the range of sources and methods used by NSAs. It covers S.18.3 and S.18.4 as well as S.18.1.]

Source data

The balance is the final outcome of a set of calculations provided by the countries. Countries use several data sources to estimate the balances. The basic data used include the consumption of fertilizers (tonnes), livestock population (1000 heads), manure imports, exports and treatment (tonnes), crop and fodder production (tonnes), crop residues (tonnes), seed production (tonnes), area of leguminous crops (1000 ha), utilized agricultural area (1000 ha). Countries may have used different types of data sources for these data. For instance some countries use estimates of the livestock population based on data from the livestock surveys or they have used other data sources like national registers on livestock. Data sources that are used available in Eurostat include: Crop Production Statistics (production and land-use), Livestock Statistics (livestock numbers), FSS (livestock numbers).

Countries have estimated coefficients based on measurements, scientific research, expert judgment, default values etc.

For the countries which don't provide data, estimates are calculated and published by Eurostat based on various available data sources, most importantly the Eurostat fertilizers, crop and livestock statistics, National inventory submissions to UNFCCC and CLRTAP, Fertilizers Europe and FAO database.

Frequency of data collection

Every second year. Next collection in 2019.

Data collection

The transmission file for the gross nitrogen balance consists of 27 interconnected excel worksheets and the transmission file for phosphorus balances of 20 interconnected excel worksheets. The transmission files are pre-described and discussed in advance of data collection with countries in the Working Group meeting on Agro-environmental statistics.

The methodology of the balances are described in the Eurostat/OECD Nutrient Budgets Handbook. The balance is estimated from total nutrient inputs minus total nutrient outputs. Inputs include Fertilisers, Manure, Seeds and planting material, biological nitrogen fixation and atmospheric nitrogen deposition. Outputs include removal of nutrients with the harvest of crops, harvest and grazing of fodder, removal and burning of crop residues and nitrogen emissions. The nutrient input or nutrient output is estimated by multiplying basic data of amounts with coefficients to convert the data into nutrient contents. Basic data (fertiliser consumption, crop production, livestock number, agricultural area) are mostly derived from agricultural statistics. Coefficients are mainly estimated by research institutes and can be based on models, statistical data, measured data as well as expert judgements. The final sheets of the transmission files calculate the results which are disseminated in Eurostat online database. There are no confidential data involved.

Data validation

Data validation includes checks with data from available data sources in Eurostat, Fertilizers Europe, FAO, GHG inventories, CLRTAP submissions, EMEP modelled depositions, data from other countries and sources, checks for internal coherence, aggregates, completeness etc.

Guidelines are described in the Handbook. Countries are expected to follow the Guidelines. In case countries deviate from the guidelines this should be clearly noted in the metadata file.

Example S.18.1-4 Income and Consumption Survey, DESTATIS, Germany [ESS-MH]

[This example contains an overview of the sample design.]

The survey is conducted every five years. As there is no legal requirement to take part in the survey, household participation is completely voluntary. Roughly 60,000 households, 14,000 of which are situated in the new Länder and Berlin, are covered by this sample survey in Germany. The EVS has been carried out in the former territory of the Federal Republic since 1962/63 and in the new Länder and Berlin since 1993.

The survey is based on quota sampling, that is, all households are selected and surveyed in accordance with a given quota plan. Using specific quota characteristics, the statistical population of households is broken down into groups. For each group the quota is specified as the number of households to be covered. Similarly to the micro census approach, the statistical population of households is initially subdivided among the Länder. As the next step, the households in each Land are grouped in accordance with the following quota characteristics: type of household, social status of the main income earner and household net income. The number of households to be surveyed is determined for each of the quota cells obtained in this way.

Example S.18.1-5 Harmonised index of consumer prices (HICP), 2017 Italian National Institute of Statistics (Istat) [ESS-MH]

[This example contains a very detailed description of the sample design and the sources.]

Prices are monthly collected through two distinct surveys carried out at central and territorial level:

- The survey carried out directly by Istat concerns prices of products that show no variability along national territory or are administered at national level (i.e. tobacco, magazine and other periodicals), that are technically too complex to be collected at territorial level because of quality adjustment issues (i.e. mobile phones, computers) or whose consumption is not strictly linked to the territorial areas (tourist services, some transport services where data collection is carried out by municipal offices of statistics (MOSs). In terms of weights, survey at central level covers the 22.1% of the 2017 HICP basket (137,500 prices per month);

- The survey at territorial level is carried out by the MOSs in 80 municipalities (out of a total of 110) which participate in the indices calculation of all representative items included in the basket and in other 16 municipalities participating in the survey for a subset of products which includes local tariffs (water supply, solid waste, sewerage collection, gas for domestic use, taxi, urban transport, car transfer ownership, canteens in schools, public day nursery, etc.) and some local services (building work, football matches, cinema, theatre shows, secondary school education, canteens in universities etc.) (501,900 prices per month, including almost 8,000 rents).

In addition, an administrative source is used, i.e. the data base of fuel prices of Ministry of Economic Development. The automotive fuels indices (the weight on the basket is 3.9%) are calculated using this data base, that collects prices for these products. 76,000 price quotes are monthly used to estimate inflation and they come from about 13,596 fuel stations on the territory, that is 69.3% of the ones active and present in the Ministry data base. The 13,596 fuel stations cover the entire national territory and they are located in different geographical areas as follows: 3,600 in the North-West; 3,200 in the North-East; 3,000 in the Centre; almost 2,400 in the South and about 1,400 in the Island.

Number of price observations per month: 706,500

Sample size: Prices are collected by territorial survey in more than 41,700 statistical units (including outlets, enterprises and institutions).

The selection of outlets is based on a non-probabilistic sampling: outlets with the highest sales revenues for different groups of products are sampled by MOSs which carry out the data collection at local level according to methods and standards established by ISTAT. The outlet selection is carried out directly by MOS, according to rules established by ISTAT and on the basis of information gathered from the Census, business plans, business registers and other available sources. The extent of the outlet sample varies town by town taking into account:

- relative weight of products in the basket;
- territory size of the municipality;
- demographic size of the town and population distribution on the municipal territory;
- type and distribution of outlets on municipal territory;
- turnover shares of large, small scale and traditional distribution;
- relevance with consumers;
- variability of prices;
- availability of products included in the basket (MOS carry out a preliminary survey in selected outlets to verify the availability of products whose prices have to be collected).

The list of outlets is updated once a year, usually in December.

Dwelling selection for rental survey is carried out directly by MOS taking into account the dwelling size, location, and ownership type (private or public).

Number of representative items at the lowest classification level

The aggregates of products at lowest classification level whose prices are monthly collected are:

- All-items 409
- 01 Food and non-alcoholic beverages 118
- 02 Alcoholic beverages, tobacco 11
- 03 Clothing and footwear 40
- 04 Housing, water, electricity, gas and other fuels 19
- 05 Furnishings, household equipment and routine maintenance of the house: 37
- 06 Health 23
- 07 Transport 35
- 08 Communications 12
- 09 Recreation and culture 57
- 10 Education 5
- 11 Restaurants and hotels 19
- 12 Miscellaneous goods and services 33

Many aggregates of products shown singly cover large areas and use sub samples (i.e. mobile phone services are one item but use a sample of about one hundred tariff plans). In the first stage, products selection is carried out by Istat using several sources: National Accounts and Household Budget Survey data; several outside sources and information from MOSs. In the second stage, the product-offers selection is made by price collectors in the field, in accordance with the 'the most sold' principle. Price collection covers both tightly and loosely specified products. Loose specifications are used if the prices within a product group are considered sufficiently homogenous. The product specification for aggregates products is defined for one year.

S.18.2 Frequency of data collection

SIMS	Concept name	Definition	Guidelines
S.18.2	Frequency of data collection	Frequency with which the source data are collected.	Indicate the frequency of data collection (e.g. monthly, quarterly, annually, or continuous).

S.18.2 EXAMPLE

Example S.18.2-1 Harmonised index of consumer prices (HICP)

National Institute of Statistics and Economic Studies, France

[This example provides a full description of this simple sub-concept.]

The collection is carried out throughout all four weeks of a given month. Although the prices of certain institutional products traditionally only change once a year, tracking is carried out throughout the year (waste collection; education).

However, there is one exception to these monthly price surveys: rents. The Rent Index is calculated on a monthly basis based on the results of household surveys, in particular the quarterly 'Rent and charges' survey for rents in the 'free' sector. As these surveys are quarterly, forecasts need to be integrated for intercalary months.

S.18.3: Data collection

SIMS	Concept Name	Definition	Guidelines
S.18.3	Data collection	Systematic process of gathering data for official statistics.	For each source of survey data: • describe the method(s) used to gather data from respondents; • annex or hyperlink the questionnaires(s). For each source of administrative data: • describe the acquisition process and how it was tested. For all sources: • describe the types of checks applied at the time of data entry. *European level* Provide a summary of the commonalities and differences in the collection methods, questionnaires and checks used in different countries.

S.18.3 FURTHER GUIDELINES

For a survey

Describe the testing of data collection procedures and systems, and the training of data collection staff.

Describe the data non-response follow-up procedures and schedule, and indicate the extent to which the schedule was followed.

Describe the procedures for entering the collected data into the input database for further processing, and the procedures for checking these data during entry.

For an administrative data process

Describe the procedures for checking the data prior to and during entry into the database.

S.18.3 EXAMPLES

Example S.18.3-1 Industrial producer prices 2016

Statistics Lithuania [ESS-MH]

[The example contains a minimal description of collection procedures. It includes a link to the questionnaire.]

Statistical data on the annual sales volume of representative industrial products in value terms are received from a statistical questionnaire KA-09 (annual). Specialists of Price Statistics Division collect data, perform control and correct the errors detected. Specialists of the regional data preparation divisions of Statistics Lithuania collect statistical data on prices for representative industrial products through statistical questionnaire KA-08 (monthly).

Reporting methods: electronic statistical data preparation and transfer system e-Statistics, by fax, email.

Example S.18.3-2 Income and Consumption Survey,

DESTATIS, Germany [ESS-MH]

[The example succinctly summarises the collection modes but does not cover all aspects of data collection.]

Regarding the survey process and design, the survey comprises four components. In addition to the basic socio-demographic and socio-economic data of households and individuals, the 'general information' component covers the housing situation and equipment of households with consumer durables. The households can supply the relevant information online. The questionnaire on financial and non-financial assets comprises questions about financial and real property as well as consumer credits and mortgage debt. The third component of the sample survey of income and expenditure is the household book. Each participating household records its income and expenditure in this book over a period of three months (rsp. one quarter). Recording is organised in such a way that the total of households participating in the survey is broken down into four quarters each of which records the relevant data over one quarter of the reference year. The final component is the 'detailed diary' with information about expenditure on food, beverages and tobacco. One in five households participating in the sample survey of income and expenditure lists, in a detailed manner, all its expenditure on food and beverages by quantity and price over a period of one month.

Example S.18.3-3 Farm Structure Survey, 2016, Statistics Austria

[This is a very comprehensive and well written example.]

1. Data collection modes

The Farm Structure Survey was held solely using an electronic questionnaire (e-Quest). The farmers were able to submit their return either directly at the computer after entering their user ID and password (direct respondents, CAWI) or during a personal interview by telephone using the same electronic questionnaire (CATI). About 60% of the questionnaires were returned by direct respondents. About 40 % of the respondents provided information by means of telephone interviews.

Those farmers, who took the help of the telephone interviewers to complete and submit their questionnaire, either called directly the free hotline or arranged an interview by sending a prepaid reply card back to Statistics Austria, filled in with their telephone number and availability (workday and time window). The competent interviewers opened the farmer’s survey form using the farmer’s access data and filled out the electronic questionnaire with the necessary information.

The survey took the form of a personalised electronic questionnaire, in which the name and the address of the holdings were already entered and only had to be checked and, if necessary, corrected. Detailed information material on how to use the electronic questionnaire and administer the Farm Structure Survey was sent directly to the respondents by post in October 2016. These consisted of an accompanying letter, a survey manual with a full description of the electronic questionnaire and a reply card (to arrange an interview) with a prepaid envelope.

A dedicated free hotline was set up by Statistics Austria to answer any questions that arose during the survey phase. In addition, queries could be sent by e-mail to Agrarstrukturerhebung@statistik.gv.at.

2. Data entry modes

Due to an exclusive use of the electronic questionnaire, no separate data collection was necessary. The data were entered by the respondents (CAWI) or by the telephone interviewers (Electronic data capture during telephone interview, see above). The data of only very few holdings, who submitted their questionnaire by mail or fax to Statistics Austria, had to be entered manually into the electronic questionnaire.

The IT department took the data directly from the questionnaire in tranches and imported it into a database.

3. Measures taken to increase response rates

The following measures were taken to increase response rates.

Maintaining up-to-date information in the Farm Register (or Agricultural and Forestry Register, AFR)

The Farm Register (or Agricultural and Forestry Register, AFR) on which the FSS is based is constantly being enhanced in terms of technical aspects and content maintenance as a result of increased updating options (administrative data, other registers etc.). In the run-up to the survey additional measures were taken to improve the up-to-dateness (e.g. adjustments with necrologies etc.).

Awareness campaign

Concerted articles were published in the relevant newspapers, journals and web pages to inform on the survey, its purposes and the importance of cooperation.

Trainings

Training of all hotline agents (permanent staff of Statistics Austria) engaged in the FSS.

Training of all telephone interviewers (temporary staff) engaged in the FSS.

Hotline-strategy

For the hotline staff and telephone interviewers at Statistics Austria, a list of frequently asked questions with regard to the motivation of respondents was compiled as a means of preparing them for questions from "difficult" respondents. Hotline agents were trained to convince respondents, who called and signalised to boycott the FSS, to give the information via direct telephone interview. Regular meetings of the hotline agents and the issue of newsletters facilitated information flows.

Telephone Interviews

Although the relatively low return quota of the reply cards (to arrange an interview) requires further improvement (only about 12,1% reply cards were returned), the farmers very willingly accepted the possibility to provide the information via telephone interview in the end. The low return quota of the reply cards had to be offset by intensified research for telephone contact information in the Farm Register (or Agricultural and Forestry Register, AFR), phone book and internet.

Reminders for overdue questionnaires

A graded series of reminders for overdue questionnaires was applied.

In former surveys the local authorities had to ensure that all questionnaires were returned and it was their responsibility to ask the holdings (either in person, over the telephone, in writing or by registered mail) to complete the questionnaire. Since the FSS 2013 the Austrian local authorities have not been directly involved. Now the holdings were targeted by telephone-interview procedures by staff at Statistics Austria.

About 3457 farmers, who had failed to complete the questionnaire on time or who could not be contacted by phone, were reminded and notified of the legal consequences via registered letter.

93 farmers insisted on their refusal and were ultimately reported by Statistics Austria to the relevant authorities in April 2017 to initiate administrative penal proceedings. Since Statistics Austria has no executive power to pursue administrative penal proceedings, information about these holdings had to be given to the relevant administrative districts that are responsible in Austria for conducting prosecutions. Normally a fine is imposed and a deadline is set for supplying the required information, i.e. payment of the fine does not release the farmer from the duty of supplying information; he/she must still provide the data in all cases.

After prosecution 24 farmers submitted their data properly – if very late in some cases – to Statistics Austria. 69 farmers did not cooperate at all. Their data had to be imputed.

S.18.4-5: Data validation and data compilation

SIMS	Concept name	Definition	Guidelines
S.18.4	Data validation	Process of monitoring the results of data compilation and ensuring the quality of statistical results.	Describe the procedures for checking and validating the source data and how the results are monitored and used. Describe the procedures for validating the aggregate output data (statistics) after compilation, including checking coverage and response rates, and comparing with data for previous cycles and with expectations. List other output datasets to which the data relate and outline the procedures for identifying inconsistencies between the output data and these other datasets. *European level* Provide a summary of the commonalities and differences in the validation methods used by countries.

SIMS	Concept name	Definition	Guidelines
S.18.5	Data compilation	Operations performed on data to derive new information according to a given set of rules.	Describe the procedures for imputation, the most common reasons for imputation and imputation rates within each of the main strata. Describe the likely impact of imputation. Describe the procedures to derive new variables and to calculate aggregates and complex statistics. Describe the procedures for adjustment for non-response and the corrections to the design weights to account for differences in response rates. Describe the calculation of design weights, including calibration (if used). Describe the procedures for combining input data from different sources.
S.18.5.1	A7. Imputation – rate (P)	The ratio of the number of replaced values to the total number of values for a given variable.	*For producer report only:* Provide values of indicator A7 Imputation – rate. The indicator A7 is defined in Supplementary Document C

S.18.4-5 FURTHER GUIDELINES

For a probability survey

Describe the non-response adjustment procedure and the corrections to the design weights to account for differences in response rates.

For a census survey

Describe the non-response adjustment procedure.

Describe the post-enumeration survey (if any).

For all types of survey and administrative data process

Describe the results of application of outlier detection and treatment procedures, including the kinds and numbers of outliers detected and the effects of their treatment on the estimates.

For a macroaggregate compilation process

Describe the procedures and methods used for combining and aggregating data according to the relevant framework for the macroaggregate in question. Intermediate calculation procedures such as process tables should be described or referred to.

Describe the model assumptions that are applied where primary data are missing or incomplete.

For all types of statistical process

Describe the procedures for comparing the data with previous cycles (if applicable) and cross reference to the results of the comparisons in S.15.2 Comparability over time

Describe the procedures for comparing the data with other relevant data, general expectations and other domain intelligence; cross reference to the results of the comparisons in S.15.3 Coherence - cross domain.

Describe model assumptions applied where primary statistics are missing or incomplete.

In each case above, describe how the results of investigations are used, for example, to further process the input data.

The ESS guidelines entitled Methodology for data validation 2.0 provide additional information about validation.

S.18.4-5 EXAMPLES

Example S.18.4-5-1 Industrial producer prices 2016

Statistics Lithuania [ESS-MH]

[The example illustrates a minimal account of data validation procedures.]

Statistical data control requirements are provided in a survey programming technical specification. Error protocol is formed from the errors detected, which contains error code, error text, error attribute indicating whether the error must be corrected or may be ignored. Errors may be logical and arithmetical; they may have been made by the respondent or during the data entry or processing stages.

To ensure statistical data quality, primary database check is run additionally (secondary editing, statistical data validation).

The error protocol, statistical data completeness and reliability are checked, links between indicators are analysed. Statistical data are corrected according to error types (errors that must be corrected or may be ignored).

Example S.18.4-5-2 Production in Industry 2016

Statistics Denmark [ESS-MH]

[The example illustrates a brief description of validation and compilation procedures.]

The online survey includes views of previous data for each enterprise. It is possible to edit data for the two previous reported months. When Statistic Denmark receives the monthly survey data, the data is checked for errors electronically by parameters set up to make sure the data does not vary too much from previous reported data. Too large variations between the monthly data will show in an error check- list for the data to be manually compared to previous reported data and data comparisons between the enterprise and similar enterprises. The enterprise is contacted if further information is needed. Missing data is imputed by using the last reported data. If an enterprise has not reported anything because the enterprise is new in the sample, missing data cannot be imputed and the enterprise data will not show in the sample until data is reported.

The level of non-response is carefully watched; the number of non-responses should decrease from the first published version of the monthly data towards the 2nd revision of the same data two months later where these data are assumed as final.

At KAU level, the data is compared to the quarterly data on turnover in the statistic manufacturers' sales of goods and services (the Danish Prodcom Statistics, OMS5 in StatBank Denmark).

Example 18.4-5-3 House Price Index and Value of Housing Transactions, 2017, DESTATIS, Germany [ESS-MH]

[The example illustrates brief description of compilation procedures.]

With regard to the House Price Indices (HPI, price indices for new and existing dwellings) chain indices are used (current reference year: 2015=100).

With regard to existing and new turnkey-ready dwellings the Jevons Index formula is used on the elementary level.

For existing and new turnkey-ready dwellings indices two different hedonic regression models have been applied (for single-family/two-family houses and for flats in apartment blocks). The regression models include the explanatory variables which have an effect on prices (price determining variables). These are: size of plot of land, size of living area, age of dwelling (derived from the year of construction), type of dwelling (free-standing detached house, semidetached/terraced house), cellar/basement (yes/no), standard land value ('Bodenrichtwert'), car parking facilities (number of spaces), district (locational information at NUTS 3 level), ground floor (flat on ground floor or higher), number of flats in block

Example 18.4-5-4 Farm Structure Survey, 2016, Statistics Austria

[This is a very comprehensive and well written example.]

Data validation

Edit rules/checks

Micro-level processing was carried out by means of extensive plausibility checks. The formal checks on the data involved a plausibility program containing about 196 plausibility rules, which again included all the controls of the Data Suppliers Manual. The types of checks performed were completeness checks, data validation, valid value checks, range checks, relational checks, arithmetic checks, ratio edits.

The plausibility rules made distinctions between the following types of error:

- Automatic errors - errors that could be automatically corrected using programmed instructions.
- Information errors - this mainly involved identifying input errors. Limit values were incorporated into the program for certain items in particular, e.g. to prevent entries being made in the wrong units of measurement (for example m²) in the case of specialised crops. If these limits were exceeded, this fact was reported. Processing staff then had to investigate or use their specialised knowledge to confirm that the data were correct or make the necessary corrections.
- Other errors - processing staff had to correct these, either by recalling/consulting the respondents or on the basis of their specialised knowledge.

Moreover, the nil returns were examined. If, for example, administrative information on the holding was available, the holding was surveyed again. This was done in close collaboration with staff dealing with the Farm Register (or Agricultural and Forestry Register, AFR), as the information from the nil returns (business closure, leasing, etc.) were used for updating the registers.

Tools used for data validation

Application for Plausibility checks

The data sets were checked for missing, incorrect or implausible information, using an extensive plausibility application. The program was developed in cooperation with the IT Dept. Care was taken to ensure that missing, incorrect and implausible entries were detected by the program and either highlighted or immediately corrected. The functionality of the plausibility program was first checked using fictitious holdings. The correction applications contained a number of deliberate errors in order to check whether the program would recognise and report them.

For the FSS 2016 about 81% of the questionnaires needed further checking due to information errors or "real" errors. This share was higher than in 2013 due to the variables concerning soil and manure management practices. For each holding, all errors and information errors were listed and categorised.

The errors detected (incorrect entries, missing or implausible data) had to be investigated and rectified by the processing team. Errors were eliminated and plausibility checks carried out directly via the application. Holdings for which errors still remain are flagged as incorrect and had to be processed once again. This process was repeated until the program detected no more errors or inconsistencies. The staff themselves could correct logical obvious errors. Frequent meetings of the staff facilitated information flows. Discussing the main issues arising from the work made it easier to standardise the criteria to solve similar situations.

Missing or incorrect entries were completed from other data sources wherever available (e.g. administrative data from IACS or ÖPUL, “total area information” from the social insurance for farmers) to avoid burdening the respondents. The forestry yearbook, containing the areas of Austria’s largest forestry holdings, was another means of checking data. If these sources were not exhaustive, individual items from the 2010 or 2013 Farm Structure Survey were used, wherever possible, to supplement and/or check the data. Where this did not provide clarity, individual holdings had to be contacted by telephone.

Level of data validation

At the level of the Electronic questionnaire (respondent/telephone interviewer)

The questionnaire was designed so that certain data items were checked for plausibility and for completeness respectively while being entered or before the questionnaire could be returned, with the result that serious errors did not go undetected and were not accepted. In order not to overload the questionnaire application and make it unnecessarily difficult for the respondents to use, this immediate plausibility check had to be limited to the most important content. Preventive measures were also taken to avoid instances of individual questionnaire sheets inadvertently being "skipped over": the marker was required to enter "The entries on this page are complete" on every page of the questionnaire. The checks in the electronic questionnaire included completeness checks, valid value checks, relational checks.

At the level of the Application for Plausibility checks (processing team)

Formal checks on the data imported into the database involved the plausibility program mentioned above.

Data compilation

Methodology for determination of weights (extrapolation factors)

1. Design weights

Design weights were obtained by taking the inverse of selection probabilities of holdings.

2. Adjustment of weights for non-response

The response rate was ultimately 99.77%. The data of the 69 units which refused to fill in the questionnaire could be imputed by using administrative data or other data-sources (Internet etc.). Therefore, no re-weighting for non-response was necessary.

Non-response because of non-existence of holdings at the time of the data collection: No adaption of sample weights necessary.

3. Adjustment of weights to external data sources.

The weights were not adjusted to external sources.

4. Any other applied adjustment of weights

No other adjustment of weights

Example 18.4-5-5 QUALITY DECLARATION. National Accounts, quarterly and annual estimates, Statistics Sweden (2018b)

Model assumptions are made where there are no primary statistics that are directly adapted to the needs of national accounts. Below are some of the most important model assumptions:

1. Given that information on value added from the business sector is missing to a large extent for the quarterly estimates, value-added is assumed to follow the development of gross output (see section 1.2.2).
2. The calculations for own-homeowners are largely model dependent. This applies in particular to values in current prices, since the valuation of such living accommodation is a particularly difficult issue. Changes in volume are also very model-dependent, but in the short term this is less susceptible to model errors, as the estimations are based on the development of the housing stock, which is relatively slow-moving. The above-mentioned housing costs apply in the same way to the production (for own final use) of corresponding housing services in owned homes. The uncertainty due to model assumptions will therefore not be shown in the residual item.
3. The estimation of housing investments is based on the start-up statistics. By assuming a certain construction time for single-family and multi-family houses, the number of begun units is modelled for investments in the following quarters. The model implies uncertainty as construction times are assumed to change in the long term and also vary in the short term. There is a strong dependence between these calculations and the calculation of construction output, which is why the uncertainty due to modelling will only to a limited extent expressed in the residual item.
4. The production of R&D and own end-use software invested in enterprises is valued based on the production cost plus a model-based profit increase. Especially for software, the calculation of the production cost is also model-based. The calculations are made primarily on a yearly basis, while the quarterly calculations are based on the assumption that production changes are in line with the enterprises' market output. The uncertainty here will not be reflected in the residual item but will affects the production and expenditure side in the same way.
5. In accordance with the guidelines applicable to European national accounts, direct volume measures are used for parts of public production and consumption. One example is student hours for education. As student hours are a simplified expression for the production of education, some uncertainty arises in the volume estimates. In addition, quarterly calculations also use forecast models for certain volume measures.
6. In the current calculations of household consumption expenditure, the result largely based on change estimates from sales data for those industries supplying goods and services to households. The calculation is based on the assumption that the distribution in these industries' sales to households and other customer categories do not change. The resulting uncertainty applies primarily to the quarterly estimates.
7. For black and illegal production and expenditure, no recurring statistics exist which is why special calculations are made with longer intervals. For intermediate years and quarters, projections are made based on corresponding or related white and legal activities.

Model assumptions are also used in the primary statistics. For example, it is common practice for model-based estimates to be used for a part of the target population (such as small businesses).

S.18.6: Adjustment

SIMS	Concept Name	Definition	Guidelines
S.18.6	Adjustment	The set of procedures employed to modify statistical data to enable it to conform to national or international standards or to address data quality differences when compiling specific data sets.	Summarise seasonal adjustment procedures at a level of detail appropriate for a user report. Outline any other macro-level adjustment procedures applied to compiled estimates that are used to improve conformance with standards and/or to address quality concerns.
S.18.6.1	Seasonal adjustment (P)	The statistical technique used to remove the effects of seasonal calendar influences operating on a series.	*For producer report only* Detail the seasonal adjustment procedures including pre-treatment (macro outlier detection, calendar correction), model selection, adjustment tool; validation procedures and revision process.

S.18.6 FURTHER GUIDELINES

Seasonal adjustment

The 2015 ESS guidelines on seasonal adjustment present theoretical aspects and practical implementation issues in a user friendly and easy to read framework. They foster the transparency of seasonal adjustment practices by encouraging the documentation of all seasonal adjustment steps and the dissemination of metadata on seasonal adjustment.

In reporting seasonal adjustment procedures, list the time series that have been adjusted, and for each series:

- name the seasonal adjustment tool (software and version) for example JDEMETRA+ or RJDEMETRA and methods for example SEATS, ARIMA-X13, State Space;
- outline the pre-treatment procedures, including calendar effects corrected for, calendar used, type of macro level outliers detected and corrected, model selection and revision and decomposition scheme adopted;
- describe the quality measures and diagnostics used to validate the model, the results of the diagnosis and the procedures for revision of seasonal factors; and
- describe the approach for handling revision of seasonally adjusted data in combination (or not) with revision of raw data, and the timing for review of seasonal factors.

For producer reports, provide more details under S.18.6.1.

Other adjustment methods

Even if already mentioned elsewhere, summarise any other methods of adjustment, such as reconciliation, balancing, and consolidation methods.

In this context the ESS guidelines on temporal disaggregation, benchmarking and reconciliation give guidance when reporting on such methods.

- They provide the terminology and describe the methods that are widely used across the ESS to obtain temporal and spatial consistency in sets of time series in official statistics.

- They provide a classification of such methods and criteria for their use, including an A/B/C quality categorisation.
- They support producers and users of official statistics and complement other ESS guidelines.

S.18.6 EXAMPLES

Example S.18.6-1 Industrial producer prices 2016

Statistics Lithuania [ESS-MH]

[The example provides a short description of price adjustment procedures.]

The specialists of the Price Statistics Division are responsible for price adjustment. Primary data of prices representative products are adjusted because of changes in quality. If the quality of the replacement product significantly differs from that of the replaced one, the assessment of the impact of the change in quality on the increase or decrease in the price has to be made and price is recalculated. To maintain comparability between the price of the replaced and replacing product, the price of the replaced product in the previous month is adjusted by eliminating the impact of the change in quality. In order to reduce the number of items adjusted in terms of quality, products are grouped into product segments taking into account the purpose of use. The quality adjustment may be done by several methods: expert judgment, option pricing, bridged overlap, quantity adjustment.

Example S.18.6-2 Production in Industry 2016

Central Statistical Office, Hungary [ESS-MH]

[The example provides a succinct description of seasonal adjustment procedures.]

Calendar adjustment: When a new seasonal model is created at the beginning of the year, calendar adjustment is checked and if the examinations show one the effects is not significant this will not be set in. For the calendar adjustment the Central Statistical Office uses a country-specific (Hungarian) calendar.

Method used: TRAMO/SEATS

Software used: Demetra version 2.04

Re-identification, re-estimation: The re-identification of the model, outliers and calendar regressors are carried out at the beginning of every year irrespective of any revisions of previous data. The re-estimation of parameters is undertaken at the beginning of the year and during the year if necessary. After a revision, complete series are re-calculated and published.

Example S.18.6-3 Production in Industry 2016

Statistics Lithuania [ESS-MH]

[The example provides a more detailed description of seasonal adjustment procedures.]

Since 2003, seasonal and working days adjustments have been implemented using DEMETRA. Final seasonally and working day adjusted data is prepared at the end of the reference year and based on the valid data.

Software used is Demetra.

All adjustments are done once per year in the end of the year.

Calendar adjustment:

- country-specific (national holidays) calendar is used;
- moving holidays that are adjusted are Easter;
- leap year effects are adjusted for.
- calendar effects are adjusted by regARIMA.

Seasonal adjustment:

- model selection is automatic;
- the model and the respective parameters re-estimated annual (once per year);
- the seasonally adjusted time series revised once per year;
- seasonal adjustment decomposition is additive and log-additive.

Model used – models checked for adequacy.

Critical value for outlier detection is automatically chosen.

Indirect adjustment via components is used; aggregation level is 2-digit level residual seasonality is checked.

Example S.18.6-4 Labour Cost Index Quarterly, 2015

Italian National Institute of Statistics [ESS-MH]

[The following comprehensive description of seasonal adjustments procedures is accompanied by a table (not included here) giving the revision history of the seasonal adjustment model.)

The calendar and seasonally adjusted Italian LCI time series are produced according to an indirect approach (Ciammola, Tuzi, 2010 and 2015): the total labour cost aggregate by section, as well as the totals of each labour cost component are derived by summing up the related adjusted component series[1]. A hierarchical calculation sequence must be followed: firstly the total labour cost by sections is calculated from wages and other costs and secondly the sum by section for each of the three variables brings to the respective LCI totals.

The transition to the new seasonal adjustment approach has been stimulated by several drawbacks produced by the direct approach. This latter was, at a first stage, chosen to afford some of the complexities of the LCI system of time series. Firstly, the volatility of the hours worked at the denominator, that would have significantly benefited by the independent adjustment of the single time series. Furthermore, to prevent the aggregates by spurious seasonality due to the low quality of the ARIMA models estimated for some of the components, particularly the other costs series by definition more volatile than wages. On the other hand, the direct approach does not guarantee consistency between the adjusted aggregate and its components[2]. In the case of the Italian LCI indices, this problem was particularly noticeable in relation to the total labour cost aggregate, because of the number of the composing variables (only two) and the fact that although highly correlated, the components may be affected by different exogenous interventions (changes in regulations that involve only the other costs, occasional payments not subjected to social security contribution, etc.).

Since June 2012 working day and seasonal adjustment have been extended to the O to S sections. Consequently, all the time series provided to Eurostat are now delivered in raw, WDA and SA format (18 sections + 3 totals: B to N, O to S and B to S).

The adjustment of the Italian LCI time series is performed through a Reg-ARIMA model-based approach using TRAMO-SEATS (Linux 2010 version). Before the adjustment for seasonality, the series are pre-treated for calendar effects according to the ESS Guidelines on Seasonal

Adjustment (Eurostat, 2015); this pre-treatment is performed only for those series showing significant and plausible effects. A country specific calendar is used. The seasonal adjustment strategy is based on a partial concurrent adjustment approach that implies models, filters, outliers and calendar regressors to be re-identified once a year and the respective parameters and factors re-estimated every time a new or revised data becomes available. This approach is aimed at minimising the frequency of revisions while guaranteeing accurate seasonally adjusted data at any given time point. The identification of the new models is normally performed with the first quarter release, when most of the interventions implying revision on raw data are concentrated (see §2 and §3 '3.1 Source data' and '6.5 Data revision'). Each quarter, the appropriateness of the identified models and the results of the seasonal adjustment process are evaluated analysing the quality measures and diagnostics provided by TRAMO-SEATS. More in depth analysis on revisions and stability of the estimates are considered when models are re-identified once a year. In this step the diagnostic facilities available on JDemetra+ version 2.0.0 are also used.

Revisions are calculated as difference between the last release (Lt) and the previous release (Pt. In the table some summary measures are also calculated (MPE, MAPE and RMAPE).

Small revisions affect all the observed vintages. These revisions incorporate those already observed in the raw data, that are caused by the regular revision policies of OROS-LES and VELA-LES surveys and by the new QNA routine revision policy referring to the O to S series (see '6.5 Data revision' and '3.1 Source data' §2 and §3). Furthermore, in the June release, they incorporate the revision of the seasonal adjustment models. A remark must be done on the effects of the recent intervention policy on the Italian labour market (starting from the second quarter of 2015) that changed the trend of the last LCI observations, implying higher revisions in initial and last observations of the SA data.

More in general, revisions of the SA LCI data highly depend on the volatility of the hours worked, that implies a lower performance of the models used in the time series adjustment and on the higher revisions of the raw data on the O to S sections, due to the multiplicity of sources used to get these NA aggregates.

S.19 (PART II) Comment

SIMS	Concept Name	Definition	Guidelines
S.19	Comment	Supplementary descriptive text which can be attached to data or metadata.	Provide any information • that is pertinent to the report but does not fit under any of the other concepts; or • to repeat key issues; or • to make reference to annexes that might be attached to the report.

This concept is included in both ESMS based (user) reports and in ESQRS based (producer) reports.

PART III Supplementary Documents

A (PART III) Quality and metadata reporting glossary

A1 Introductory notes

Starting point and approach

The starting point in constructing the ESS Quality and Metadata Reporting Glossary was the set of terms and their definitions in the ESS Quality Glossary, of which the most recent version is recorded the ***Specialised Glossary -Quality Glossary***, which is a subset of ***Eurostat's Concepts and Definitions Database (CODED)***. Additional sources have been used in updating and augmenting these terms and definitions.

Some definitions have been drawn verbatim from a source, others are based on a source or combination of sources rather than being exactly as in a source.

Updating definitions

The definitions of all the terms in the ESS Quality Glossary were reviewed and revised to improve clarity and/or to update them to the most recent versions of the standards on which they are based. In particular, many definitions were updated to the latest versions of:

- European Statistics Code of Practice (2017);
- ISO 9000:2015 Quality Management Systems Fundamentals and Vocabulary.

No terms were removed. Not all terms in the Glossary are used in the EHQMR.

Adding terms and definitions

Terms were added because a common understanding of the terms is required for the Handbook, but the terms were not defined in ESS Quality Glossary.

- New metadata related terms include: attribute, characteristics, quality metadata, reference metadata, structural metadata:
- New statistical terms are: error profile, outlier, ideal population, target population, survey population, survey frame
- New quality related terms include: quality assurance framework

Sources

The sources used in constructing the ESS Quality and Metadata Reporting Glossary were the following.:

Eurostat's Concepts and Definitions Database (CODED):

(The following are all accessed through the CODED hyperlink.)

- Specialised Glossary - Quality Glossary;
- General Statistical Terminology;
- Metadata;
- Specialised Glossary - SDMX Glossary, 2016.

Other international statistical glossaries:

- OECD Glossary of Statistical Terms; (Downloadable version);
- ISI Multilingual Glossary of Statistical Terms;
- UN National Quality Assurance Framework (NQAF) Glossary.

International quality glossaries:

- ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary;
- ISO Translated into Plain English,Praxiom Research Group Limited.

International metadata glossaries:

- SDMX Glossary Version 2.0, October 2018.

A2 Glossary

ACCESSIBILITY

Accessibility is an attribute of statistics describing the set of conditions and modalities by which users can obtain statistical information.

Accessibility measures the ease and conditions with which statistical information can be obtained. It refers to the availability of statistical information to the user. It includes the suitability of the form or medium through which the information can be accessed. The cost of the information may also be an aspect of accessibility for some users.

Accessibility refers to the physical conditions in which users can obtain data: where to go, how to order, delivery time, clear pricing policy, convenient marketing conditions (copyright, etc.), availability of micro or macro data, various formats (paper, files, CD-ROM, Internet), etc.

The European Statistics Code of Practice Principle 15 refers to *Accessibility and Clarity* and states

> *European statistics should be presented in a clear and understandable form, disseminated in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance*.

SIMS defines S.10 Accessibility and clarity as *the conditions and modalities by which users can obtain, use and interpret data*.

Sources

European Statistics Code of Practice (2017), Eurostat.

National Quality Assurance Framework (NQAF) Glossary, UN.

ACCURACY

Accuracy is an attribute of statistics measuring the closeness of estimates to the unknown true values.

In the European Statistics Code of Practice, Principle 12 states *that European statistics must accurately and reliably portray reality*.

SIMS defines S.13 Accuracy as *closeness of computations or estimates to the unknown exact or true values that the statistics were intended to measure.*

Sources

European Statistics Code of Practice (2017), Eurostat.

ADEQUACY OF RESOURCES

Adequacy of resources is the characteristic of a statistical authority that enables it to meet statistical requirements.

Resources include human, financial and technical resources, which must be adequate both in magnitude and in quality.

In the European Statistics Code of Practice, Principle 3 states *the resources available to statistical authorities are sufficient to meet European Statistics requirements*.

Source

European Statistics Code of Practice (2017), Eurostat.

ADMINISTRATIVE DATA

Administrative data refer to units and data derived from an administrative source.

Administrative data are collected for regulatory, accounting, commercial or other non-statistical purposes, not primarily for statistical purposes. They are used for registration, transaction and record keeping, usually during the delivery of a service, by a government unit, company or other type of organisation that is the source.

Administrative data are typically of two types – an administrative register (with unique identifiers) and administrative transactions.

By way of contrast, statistical registers are registers created for statistical purposes. They are typically created by transforming and combining data from administrative and statistical sources.

Source

OECD Glossary of Statistical Terms

Eurostat Statistics Explained Glossary.

ADMINISTRATIVE DATA PROCESS

An administrative data process is a type of statistical process in which statistics are produced based on administrative data.

Note that administrative data may be collected for administrative (regulatory, accounting, commercial or other non-statistical) purposes by government organisations or companies, or non-profit organisations.

Source

This Handbook.

ADMINISTRATIVE DATA SOURCE

[Same as Administrative source.]

ADMINISTRATIVE SOURCE

An administrative source is an organisational unit responsible for implementing a regulation (or group of regulations), for which the corresponding register of units and transactions are viewed as a source of data by a statistical authority.

For the purposes of this Handbook, the organisational unit may be governmental, or may be commercial, or other type of non-government unit, and the regulation may be imposed by the government or may be an operating procedure of a commercial or other type of organisation.

This is an extension of the more usual definition of administrative source (which is confined to government organisations). It enables an *administrative data process* to include a statistical process that uses data from commercial or other non-government organisation.

Source

OECD Glossary of Statistical Terms

APPROPRIATE STATISTICAL PROCEDURE

An appropriate statistical procedure is one that underpins the production of quality statistics.

Principle 8 of the European Statistics Code of Practice states “Appropriate statistical procedures, implemented throughout the statistical processes, underpin quality statistics.”

Source

European Statistics Code of Practice (2017), Eurostat.

ATTRIBUTE

An attribute is a concept providing qualitative information about a specific object.

In the ESS context, the object is a statistical object.

The specific object in a data set can be a data set, observation, series key or partial key, and in a metadata set can be any object in the underlying information model.

Concepts such as unit of measure, magnitude, currency of denomination, and title are commonly specified as attributes in a data structure; methodological comments, quality statements are commonly specified as attributes in a metadata structure. They can be used as attributes in the context of an agreed data exchange.

The *attribute value* is the reported value in a data set or a metadata set such as a specific currency or a specific dissemination policy applicable to the object to which the attribute value is attached.

Source

SDMX Glossary Version 2.0, October 2018

BALANCED SCORECARD

The balanced scorecard is a strategic planning and management system that is used extensively in business and industry, government, and non-profit organisations worldwide to align business activities to the vision and strategy of the organisation, improve internal and external communications, and monitor organisation performance against strategic goals.

The balanced scorecard has evolved from its early use as a simple performance measurement framework to a full strategic planning and management system.

The balanced scorecard transforms an organisation's strategic plan from an attractive but passive document into the "marching orders" for the organisation on a daily basis. It provides a framework that not only provides performance measurements, but helps planners identify what should be done and measured. It enables executives to truly execute their strategies.

Source

Website of the Balanced Scorecard Institute

Balanced Scorecard.

BENCHMARKING

Benchmarking means comparing data, metadata or processes against a recognised standard.

Benchmarking may refer, for instance, to the case where there are two sources of data for the same target variable with different frequencies, e.g. quarterly and annual estimates of value-added from different sources.

Benchmarking is generally done retrospectively, for example annual benchmark data are available some time after quarterly data. Benchmarking may have a forward-looking element, however, in that the relationship between benchmark and indicator data is extrapolated forward to improve quarterly estimates for the most recent periods for which benchmark data are not yet available

Source

National Quality Assurance Frameworks (NQAF) Glossary, UN.

CERTIFICATION

Certification is an activity which assesses whether a particular product, service or process or system complies with requirements defined by a standard or other document containing criteria.

Certification is conducted by an external independent certification body. The result of the successful certification is the certificate awarded to the organisation by the certification body.

In the context of quality management within the ESS, certification may refers to certification of a statistical authority's quality management system or of selected outputs

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary

CHARACTERISTIC

A characteristic is a distinctive feature or property of something.

Characteristics can be inherent or assigned. An inherent characteristic exists in something or is a permanent feature of something, while an assigned characteristic is a feature that is attributed or attached to something.

A characteristic can be qualitative or quantitative.

Source

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary

ISO Translated into Plain English, Praxiom Research Group Limited.

CLARITY (OF STATISTICS)

Clarity is an attribute of statistics describing the extent to which the metadata necessary to give a full understanding of those statistics are readily available and are easily comprehensible.

Clarity is sometimes referred to as "interpretability". It refers to the data information environment: whether data are accompanied by appropriate metadata, including information on their quality, and the extent to which additional assistance is provided to users by data providers.

In the Principle 15 of the European Statistics Code of Practice, clarity is associated with accessibility. Principle 15 states

> *European Statistics are presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.*

SIMS defines S.10 Accessibility and clarity as *the conditions and modalities by which users can obtain, use and interpret data*.

Source

European Statistics Code of Practice (2017), Eurostat.

COHERENCE

Coherence is an attribute of statistics measuring the adequacy of the data to be reliably combined in different ways and for various uses.

According to Principle 14 of the European Statistics Code of Practice, European statistics *should be consistent internally, over time and comparable between regions and countries; it should be possible to combine and make joint use of related data from different sources.*

When originating from different sources, statistics reflect differences due to different approaches, concepts, classifications and methods.

There are several areas where the assessment of coherence is regularly conducted: between provisional and final statistics, between annual and short-term statistics, between statistics within the same socio-economic domain, and between survey statistics and national accounts.

The concept of coherence includes the concept of *comparability* as a special case. Comparability refers to the coherence of sets of statistics for nominally the same populations and data content, but at different points in time or in different regions.

In SIMS, S.15 Coherence is broken down into S.15.1 Comparability – geographical, S.15.2, Comparability – over time, S.15.3 Coherence - cross domain, and S.15.4 Coherence - internal.

In the Data Quality Assessment Framework (DQAF) of the International Monetary Fund, the term *consistency* is used to indicate "logical and numerical coherence". DQAF's "internal consistency" and "intersectoral and cross-domain consistency" can be mapped to "internal coherence" and "cross-domain coherence" respectively.

Sources

European Statistics Code of Practice (2017), Eurostat.

National Quality Assurance Framework (NQAF) Glossary, UN

Data Quality Assessment Framework, International Monetary Fund

COMMITMENT TO QUALITY

Commitment to quality is a characteristic of a statistical authority through which it systematically and regularly identifies strengths and weaknesses and continuously improves process and product quality.

Principle 4 of the European Statistics Code of Practice states that (European) *statistical authorities are committed to quality*.

Source

European Statistics Code of Practice (2017), Eurostat.

COMMON ASSESSMENT FRAMEWORK

The Common Assessment Framework (CAF) is a European quality management instrument for the public sector developed by the public sector.

The CAF is an easy-to-use, free tool to assist public-sector organisations across Europe in using quality management techniques to improve their performance.

The CAF is a total quality management (TQM) tool which is inspired by the major Total Quality models in general, and by the Excellence Model of the European Foundation for Quality Management (EFQM) in particular. It is especially designed for public-sector organisations, taking into account their characteristics.

The model is based on the premise that excellent results in organisational performance, citizens/customers, people and society are achieved through leadership driving strategy and planning, people, partnerships, resources and processes. It looks at the organisation from different angles at the same time; a holistic approach to organisation performance analysis.

Source

Common Assessment Framework, European Institute of Public Administration (EIPA)

COMPARABILITY

Comparability is the degree of coherence of datasets referring to nominally the same populations and data content, but at different points in time, or in different regions, or in different domains.

Comparability is a special case of coherence. As in the case of coherence, lack of comparability can be due to differences in statistical concepts, definitions, measurement tools or procedures.

In SIMS, comparability is broken down into:

- *S.15.1 Comparability – geographical*,- referring to the degree of comparability between statistics measuring the same phenomenon for different geographical areas;
- *S.15.2 Comparability over time*, referring to the degree of comparability between two or more instances of statistics on the same phenomenon measured at different points in time.

Sources

European Statistics Code of Practice (2017), Eurostat.

SDMX Glossary Version 2.0, October 2018

CONCEPTUAL METADATA

Conceptual metadata describe the concepts used, and their practical implementation in a statistical process and the statistics it produces.

Conceptual metadata are a type of reference metadata. They aid users in understanding what the statistics are measuring and, thus, their fitness for use.

Source(s)

SDMX Glossary Version 2.0, October 2018

CONFIDENTIALITY

Same as statistical confidentiality

CONSISTENCY

Consistency is an attribute of statistics that is closely associated, but not synonymous, with coherence. There are two types of consistency– logical consistency and numerical consistency.

Logical consistency requires that a statistical concept has one and only one definition in all areas of statistics that are subject to combination or comparison. The term is much used in the context of the National Accounts.

Numerical consistency requires, for example, that, within a set of outputs for a statistical process, the numerical value for a whole is equal to the sum of those for its parts. Or that values for conceptually the same data item derived from different processes should be the same.

Within the IMF's Data Quality Assessment Framework, consistency is one of the elements of *serviceability*.

In probability sampling theory an estimator is said to be consistent if it converges in probability to the true value as sample increases. This is a different meaning of the same term and not the one used in this Handbook.

Source(s)

National Quality Assurance Framework (NQAF),UN

COST EFFECTIVENESS

Cost effectiveness is a characteristic of a process, where the cost of producing the statistics is made as small as possible, while still achieving the desired output quality. Alternatively, the best possible output quality is achieved within a fixed budget.

Principle 10 of the European Statistics Code of Practice states simply *Resources are used effectively*.

Sometimes, but not in the ES CoP or in the EHQMR, cost-effectiveness is also taken to include minimising respondent burden, which can be seen as part of the cost of producing statistics. In the ES CoP, *Principle 9 Non-Excessive burden on respondents* is a separate principle.

Source(s)

European Statistics Code of Practice (2017), Eurostat

SDMX Glossary Version 2.0, October 2018

CREDIBILITY

Credibility is the confidence that users place in statistical products based simply on their image of the data producer, the statistical authority i.e., the brand image.

Credibility is an aspect of quality defined and used by the OECD. The elements leading to credibility are in incorporated in the ES CoP principles.

The confidence of users is built over time. One important aspect is trust in the objectivity of the data. This implies that the data are perceived to be produced professionally in accordance with appropriate statistical standards, and that policies and practices are transparent. For example, data are not manipulated, nor their release timed in response to political pressure.

Credibility is determined in part by the integrity of the production process. Principle 2 of the Fundamental Principles of Official Statistics states:

> *to retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.*

Source(s)

OECD Glossary of Statistical Terms.

EFQM EXCELLENCE MODEL

The EFQM Excellence Model allows people to understand the cause and effect relationships between what their organisation does and the results it achieves.

The Excellence Model is developed and maintained by the *European Foundation for Quality Management (EFQM)*.

The Excellence Model refers to eight fundamental concepts of excellence, which lay the foundation for achieving sustainable excellence in any organisation. They are:

- *Succeeding through the talent of people;*
- *Sustaining outstanding results;*
- *Adding values for customers;*
- *Creating a sustainable future;*
- *Developing organisational capacity;*
- *Harnessing creativity and innovation;*
- *Leading with vision, inspiration and integrity;* and
- *Managing with agility*

These fundamental concepts are effectively an alternative expression of the *ISO 9000 quality management principles*.

The Excellence Model provides a basis for certification using nine criteria, five of which are *enablers* (what an organisation does and how it does it) and four of which are *results* (what an organisation achieves).

- The enabler criteria are: *Leadership; People; Strategy; Partnerships and Resources: and Processes, Products and Services*.
- The results criteria are: *People results; Customer results; Society results; and Business results*.

The Excellence Model is the basis for the majority of national and regional quality awards in Europe.

Used as a tool for assessment, the Excellence Model delivers a picture of how well the organisation compares to similar or very different kinds of organisation. Used as a management model it can help define aspirations for the organisation's capability and performance.

Source

European Foundation for Quality Management (EFQM) Excellence Model

ERROR PROFILE

An error profile is a compilation that catalogues what is known about each of the component errors present in a given survey.

The error profile serves as a balanced summary of all sampling and non-sampling errors, from which to make an informed judgment on where to focus both the efforts on improving accuracy and the reporting of accuracy and its components.

Source

Brooks and Bailar (1978)

EUROPEAN FOUNDATION FOR QUALITY MANAGEMENT

EFQM is a not for profit membership foundation, based in Brussels, managing the *EFQM Excellence Model* and supporting organisations in Europe and beyond, reach sustainable excellence.

With its 30 years' experience, carefully designed portfolio of services and a network of 30.000 organisations from all sectors, of various sizes and degrees of maturity, EFQM is recognised as a key partner in achieving excellence.

Source(s)

European Foundation for Quality Management

EUROPEAN STATISTICS CODE OF PRACTICE

The European Statistics Code of Practice is the cornerstone of the common quality framework of the European Statistical System.

It provides a structure for supporting quality improvement in the ESS.

It is based on 16 principles arranged in three groups: *institutional environment, statistical processes statistical outputs*. A set of indicators of best practices and standards for each of the principles provides guidance and reference for reviewing its implementation.

It was first adopted in 2005 by the Statistical Programme Committee and revised by the European Statistical System Committee in 2011 and 2017.

It is a self-regulatory instrument.

Source(s)

European Statistics Code of Practice, Eurostat (pdf)

European Statistics Code of Practice (2017), Eurostat. (Web site)

FRAME

Same as survey frame.

IDEAL POPULATION

An ideal population reflects the needs of a user in terms of set of units to which the user wants the statistical outputs to refer.

Different users of the results of a statistical process may well have different ideal populations in mind. Thus, when deciding on the target population for a survey, the producer has to balance these possible ideal populations against one another with what can be achieved the available resources and time constraints.

The ideal population is an aspect of a user's ideal survey goal, which also includes the user's ideal definitions of variables etc.

Source

Särndal et al (1992)

IMPARTIALITY

Impartiality is an attribute of a statistical authority, or a process conducted by that authority, confirming that statistics are developed, produced and disseminated in a neutral manner, and that all users are treated equitably.

ES CoP Principle 6 Impartiality and Objectivity states that *Statistical authorities develop, produce and disseminate European Statistics respecting scientific independence and in an objective, professional and transparent manner in which all users are treated equitably*.

Source(s)

European Statistics Code of Practice, Eurostat.

National Quality Assurance Framework (NQAF),UN

INSTITUTIONAL ENVIRONMENT

The institutional environment comprises the organisational structures and rules that are the basis for producing statistics in the organisation.

The ES CoP defines seven principles referring to the institutional environment, namely:

- *professional independence;*
- *coordination and cooperation;*
- *mandate for data collection and access to data;*
- *adequacy of resources;*
- *commitment-to quality;*
- *statistical confidentiality and data protection; and*
- *impartiality and objectivity*.

Source(s)

European Statistics Code of Practice, Eurostat.

INTEGRITY

Integrity is the set of values and related practices of a statistical organisation that maintain confidence in the eyes of users regarding the organisation and, hence, of its statistical outputs.

Integrity refers to the description of the policy on the availability of the terms and conditions under which statistics are collected, processed, and disseminated. It also describes the policy of providing advanced notice of major changes in methodology, source data, and statistical techniques; the policy on internal governmental access to statistics prior to their release; the policy on statistical products' identification.

Integrity is the third of four dimensions of IMF's SDDS.

Integrity is not one of the ES CoP principles but the elements of integrity are included in the ES CoP principles relating to the institutional environment.

Source(s)

Data Quality Assessment Framework, IMF.

European Statistics Code of Practice, Eurostat

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION

The International Organization for Standardization (ISO} is the world's largest developer and publisher of international standards.

The ISO is an independent, non-governmental organization made up of members from the national standards bodies of 164 countries. Its members play a vital role in how it operate, meeting once a year for a General Assembly that decides its strategic objectives.

The Central Secretariat in Geneva, Switzerland, coordinates the ISO and runs day-to-day operations, overseen by the Secretary General.

Through its members, the ISO brings together experts to share knowledge and develop voluntary, consensus-based, market relevant international standards that support innovation and provide solutions to global challenges.

The standards of greatest interest from the perspective of quality management in the ESS are the *ISO 9000 Family -Quality Management Systems*.

Source

International Organization for Standardisation

ISO 9000 FAMILY – QUALITY MANAGEMENT SYSTEMS

ISO 9000 family of standards addresses the various aspects of quality management and provide guidance and tools for companies and organisations

The ISO 9000 family contains some of the ISO's best known standards. They are aimed at organisations who want to ensure that their products and services consistently meet customers' requirements, and that quality is consistently improved.

There are three key standards in the family that are not specific to any one industry and can be applied to organisations of any size

- ISO 9001:2015: Quality management systems - Requirements
- ISO 9000:2015: Quality management systems - Fundamentals and vocabulary
- ISO 9004:2018: Quality management systems – Managing for the sustained success of an organization

Source

International Organization for Standardisation

ISO 20252:2019: MARKET, OPINION AND SOCIAL RESEARCH – VOCABULARY AND SERVICE REQUIREMENTS

ISO 20252:2019 establishes terms, definitions and service requirements for service providers conducting market, opinion and social research, including insights and data analytics.

It sets out guidance and requirements relating to the way in which market research studies are planned, carried out, supervised, and reported to clients commissioning such projects.

It encourages consistency and transparency in the way surveys are carried out, and confidence in their results and in their providers.

Non-market research activities, such as direct marketing, are outside its scope.

Source

ISO 20252:2019 Market, opinion and social research -- Vocabulary and service requirements

ISO 9000:2015 QUALITY MANAGEMENT SYSTEMS -- FUNDAMENTALS AND VOCABULARY

ISO 9000:2015 describes the fundamental concepts and principles of quality management that are universally applicable any type of organisation.

It is addressed to:

- organizations seeking sustained success through the implementation of a quality management system;
- customers seeking confidence in an organization's ability to consistently provide products and services conforming to their requirements;
- organizations seeking confidence in their supply chain that their product and service requirements will be met;
- organizations and interested parties seeking to improve communication through a common understanding of the vocabulary used in quality management;
- organizations performing conformity assessments against the requirements of ISO 9001;
- providers of training, assessment or advice in quality management;
- developers of related standards.

In the context of the EHQMR, ISO 9000:2015 is the primary source of general *quality concepts* and *quality management principles*.

Source

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary

ISO 9001:2015 QUALITY MANAGEMENT SYSTEMS - REQUIREMENTS

ISO 9001:2015 sets out the criteria for a quality management system.

It is the only standard in the ISO 9000 family to which an organisation can be certified. However, certification is not a requirement for its use.

It can be used by any organisation, large or small, regardless of its field of activity. There are over one million companies and organizations in over 170 countries certified to ISO 9001.

It is based on the general quality management principles expressed in ISO 9000:2015.

Source

ISO 9001:2015: Quality management systems - Requirements

LABELLING QUALITY

Labelling quality means that a label relating to quality is attached to statistics or to the NSA that produced the statistics.

A label may indicate that the statistics meet a quality standard. It may indicate that the producer of the statistics meets a quality standard.

Users are likely to place higher trust in the statistics with a label, and in statistics from a producer with a label.

Labelling requires a procedure to guarantee that the message is appropriate.

Source(s)

Handbook on Data Quality Assessment Methods and Tools (DaTQAM), Eurostat

LEAN SIX SIGMA

Lean Six Sigma is a philosophy of improvement that drives customer satisfaction and performance by reducing variation, waste, and cycle time, while promoting the use of work standardization and flow.

Lean Six Sigma is a quality management approach that combines the strategies of *Lean* and *Six Sigma*. Lean principles help to reduce or eliminate process waste. Six Sigma focuses on variation - reduction in process.

Lean Six Sigma is fact-based and data-driven. It emphasises defect prevention rather than defect detection. It applies anywhere variation and waste exist. It involves all employees.

In the ESS context, defects are interpreted as errors and waste is interpreted as lack of cost effectiveness.

Source

Lean Six Sigma, GreyCampus

MANDATE FOR DATA COLLECTION AND ACCESS TO DATA

A mandate for data collection and access to data is an attribute of a statistical authority that has the legal power to collect information for statistical purposes.

ES CoP Principle 2 Mandate for Data collection and Access to Data states

> *Statistical authorities have a clear legal mandate to collect and access information from multiple data sources for European statistical purposes. Administrations, enterprises and households, and the public at large may be compelled by law to allow access to or deliver data for European statistical purposes at the request of statistical authorities.*

Having a mandate implies that a statistical authority (1) has the right to collect and access information/data, (2) may compel response to statistical surveys, and (3) is allowed by law to access administrative data, promptly and free of charge, and use them for statistical purposes.

Source

European Statistics Code of Practice, Eurostat

METADATA

Metadata are information that is needed to be able to use and interpret statistics. Metadata describe data by giving definitions of populations, objects, variables, the methodology and quality.

Metadata are divided int two broad types - *structural metadata* and *reference metadata*.

Source

Eurostat's Concepts and Definitions Database (CODED): General Statistical Terminology;

METHODOLOGICAL METADATA

Methodological metadata are metadata describing methods used in a statistical process for generation of the data

Methodological metadata are a type of *reference metadata*. They include, for example, sample design, collection methods, and editing processes.

Source

SDMX Glossary Version 2.0, October 2018

METHODOLOGICAL SOUNDNESS

Same as Sound Methodology

NON-EXCESSIVE BURDEN ON RESPONDENTS

Non-excessive burden on respondents means that the burden imposed by direct data collection on respondents is proportionate to the needs of the users and is not excessive for respondents.

ES CoP Principle 9 Non-excessive burden on respondents states:

> *The response burden is proportionate to the needs of the users and is not excessive for*

respondents. The statistical authorities monitor the response burden and set targets for its reduction over time.

Respondent burden does not refer to acquisition of administrative data.

Source

European Statistics Code of Practice, Eurostat

OBJECTIVITY

Objectivity is an attribute of a statistical authority confirming that statistics are developed, produced and disseminated in a systematic, reliable and unbiased manner.

Objectivity implies the use of professional and ethical standards, and that the policies and practices followed are transparent to users and survey respondents.

ES CoP Principle 6 Impartiality and Objectivity states:

> *Statistical authorities develop, produce and disseminate European Statistics respecting scientific independence and in an objective, professional and transparent manner in which all users are treated equitably*

Source

European Statistics Code of Practice, Eurostat

OBSERVATION UNIT

Unit about which information is obtained.

Usually, but not necessarily, the target statistical unit for a statistical process.

Also referred to in several examples as the *reporting unit*. The Handbook avoids using the term reporting unit as it may also refer to the unit *from which* information is obtained. For example, the reporting unit could be an accounting business that reports about the observation unit on behalf of the observation unit.

Source

This Handbook.

OPERATIONAL METADATA

Operational metadata are metadata that describe the expected or actual outcomes of a process using evaluable and operational metrics."

Operational metadata are a type of *reference metadata.* They include *quality metadata* and metadata measuring performance.

An alternative name is paradata.

Source

Based on definition of process metadata provided by the ESS net for data warehousing.

OUTLIER

Outliers at the micro level are sampling units with extreme values that are assessed not to be representative of the stratum to which they belong.

Outliers at an aggregate level are estimates with extreme values that disturb a certain analysis.

When applying standard sampling weights to outliers, estimates are believed to be distorted. It is important to recognise, however, that the distortion is, in principle, the result of a large sampling error and not a bias. Any special treatment of outliers in probability surveys results in biases to the estimates, which may, however, be smaller than the sampling error resulting from no treatment at all.

In this handbook outliers at aggregate levels are only referred to with regard to time series analysis.

Source

Lee (1995)

OUTPUT QUALITY

Same as Statistical Output Quality

PEER REVIEW

A peer review is an evaluation of the performance and/or quality of an output of an organisational unit by experts drawn from different but comparable organisational units.

In the ESS a peer review may be of a statistical process and outputs, by persons responsible for a different process within the NSA or by persons responsible for a similar process within another NSA. It may be of the NSA as a whole, by another NSA

Typically, a peer review is less formal than an audit and aims at assessing the quality in general rather than conformity to a quality standard.

Sources

National Quality Assurance Frameworks (NQAF) Glossary, UN

Handbook on Data Quality Assessment Methods and Tools, Eurostat

PRODUCT QUALITY

Same as Output Quality

PRE-RELEASE ACCESS

The pre-release access is the practice of giving certain individuals or organisations access to data under embargo before those data are released to the public.

Pre-release access should include transparent recording of persons or officials holding designated positions within the government (but outside the statistical system producing the data) who have pre-release access to the data and the reporting of the schedule according to which they receive access.

In the Data Quality Assessment Framework of the IMF, pre-release access is referred to as *internal access*.

Source

SDMX Glossary Version 2.0, October 2018

PREREQUISITES OF QUALITY

Prerequisites of quality are institutional conditions for the pursuit of data quality as defined in the Data Quality Assessment Framework (DQAF) if the IMF

The DQAF groups the indicators of this kind into four elements: legal and institutional environment, resources, relevance and other quality management. These elements and indicators are identified to reinforce the idea that data users, who often cannot replicate or otherwise verify data, must place their trust in the institutions that produce statistics and the people who staff them.

The term *prerequisites of quality* is not used In the ES CoP. The prerequisites are included within the eight principles referring to the institutional environment.

Source(s)

Data Quality Assessment Framework (DQAF), IMF

European Statistics Code of Practice, Eurostat

PROCESS APPROACH

A process approach is a management strategy requiring managers to manage and control the processes that make up their organisation, the interaction between these processes, and the inputs and outputs that tie these processes together.

A process approach involves systematic identification and management of the processes in the organisation and particularly of the interactions between such processes.

Applying a process approach implies:

- systematically defining the processes and sub-processes necessary to obtain a desired result;
- establishing clear responsibility and accountability for managing the key sub-processes;
- analysing and measuring of the capabilities of the key sub-processes;
- identifying the interfaces between the key sub-processes;
- identify resources, methods, and materials that will improve the keysub-processes.

Sources

ISO Translated into Plain English, Praxiom Research Group Limited.

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary

PROCESS DESCRIPTION

A process description is a document which describes a process.

A process description may contain:

- the name and the aim of the process;
- the process owner and operators;

- inputs (and the processes they come from); outputs (and the processes they go to);
- sub-processes that transform inputs into outputs;
- regulatives (internal, external) that characterise the regulated environment;
- resources that are used in the transformation;
- how the process is managed and improved (performance and quality indicators with target values);
- the way of monitoring, measurement, analysis; improvements;
- records stating results achieved or providing evidence of activities performed.

Source

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary

PROCESS QUALITY

See Statistical Process Quality

PROCESS VARIABLE

A process variable is an indicator of the quality of a process.

Process variables vary with each repetition of a process. Key process variables are the ones that are highly correlated with product characteristics, i.e. that best indicate the quality of the product.

Source(s)

Handbook on improving quality by analysis of process variables, Eurostat, "

PROCESS-BASED QUALITY MANAGEMENT SYSTEM

A process-based quality management system (QMS) uses a process approach to manage and control how its quality policy is implemented and how its quality objectives are achieved.

A process-based QMS is a network of interrelated and interconnected processes. Each process uses resources to transform inputs into outputs. Since the output of one process becomes the input of another process, processes interact and are interrelated by means of such input-output relationships. These process interactions create a single integrated process-based QMS.

Source(s)

ISO Translated into Plain English, Praxiom Research Group Limited.

PROFESSIONAL INDEPENDENCE

Professional independence is a characteristic of a statistical authority reflecting its capacity to develop, produce and disseminate statistics in an independent manner, free from pressure from political or interest groups or other national authorities or the Community.

ES CoP Principle 1 Professional independence states:

Professional independence of statistical authorities from other policy, regulatory or administrative *departments and bodies, as well as from private sector operators, ensures the credibility of European Statistics.*

Professional independence applies to the selection of techniques, definitions, methodologies and sources to be used, and the timing and content of all forms of dissemination.

It implies that the heads of the NSIs, of other NSAs (where appropriate) and of Eurostat have sufficiently high hierarchical standing to ensure senior level access to policy authorities and administrative public bodies, also that the procedures for their recruitment and appointment are transparent and based on professional criteria only, and that the reasons on the basis of which their incumbency can be terminated are specified in the legal framework and cannot include reasons compromising professional or scientific independence.

Source(s)

European Statistics Code of Practice, Eurostat.

PROFESSIONALISM

Professionalism is the set of standards, skills and abilities of a statistical authority that are suitable for producing statistics of good quality.

Principle 2 of the Fundamental Principles of Official Statistics states:

To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data

Professionalism implies that:

- statistics are produced on an impartial basis;
- choices of sources and statistical techniques as well as decisions about dissemination are informed solely by statistical considerations;
- recruitment and promotion of staff are based on relevant aptitude; and
- an NSA is entitled to comment on erroneous interpretation and misuse of statistics.

Source(s)

SDMX Glossary Version 2.0, October 2018.

PUNCTUALITY

Punctuality is an attribute of statistics measuring the delay between the date of the release of the statistics and the date by which the statistics were scheduled for release.

ES CoP Principle 13 states:

European statistics are released in a timely and punctual manner.

SIMS Concept S.14.2 defines punctuality as *the time lag between the actual delivery of the data and the target date when it should have been delivered*.

Source(s)

European Statistics Code of Practice, Eurostat.

QUALITY

Quality is the degree to which a set of inherent characteristics of an object fulfils requirements.

In the context of the ESS:

- The *object* may be a statistical output, service, process, system, methodology, organisation, resource, or input;
- *Characteristic* means distinguishing feature;
- *Inherent* means existing in the object, not assigned to it; and
- *Requirement* means need or expectation that is stated, generally implied, or obligatory.

The ES CoP defines 16 principles in three groups, namely *institutional environment, statistical processes,* and *statistical output*.

Quality of statistical output is a multi-faceted concept. The dimensions of quality that are most important depend on user perspectives, needs and priorities, which vary across groups of users.

Article 12 of Regulation (EC) No 223/2009 on European statistics amended by Regulation 2015/759 refers to seven quality criteria: relevance, accuracy, timeliness, punctuality, accessibility and clarity, comparability, and coherence.

Sources

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary, ISO, Geneva, 2015.

European Statistics Code of Practice (2017), Eurostat,

QUALITY ASSESSMENT

Quality assessment is the aspect of quality assurance that focuses on the extent to which statistical outputs and the processes that produced them meet quality requirements

Quality assessment may be quantitative or qualitative. It may be based on a standard assessment template.

A quality report is the typical output of quality assessment.

In the ESS context

> In SIMS, *quality assessment* (sub-concept S.13.2) is defined as *overall assessment of data quality, based on standard quality criteria.*

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary, ISO, Geneva, 2015.

QUALITY ASSURANCE

Quality assurance is the part of quality management focused on providing confidence that quality requirements are fulfilled.

Quality assurance is achieved by identifying what “quality" means in context, specifying methods by which its presence can be ensured, and specifying ways in which it can be assessed to verify satisfaction of requirements.

In the ESS context

> In SIMS, *quality assurance (sub-concept S.11.1)* is defined as *all systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the*

requirements for the statistical output.

Quality assurance focusses on the core business of a statistical organisation, i.e. development, production and dissemination of statistics. It is an organisation's guarantee that the products and services it offers meet the requirements for the statistical outputs. It comprises all the planned and systematic activities implemented that can be demonstrated to provide confidence that the statistical processes will meet the needs for the statistical outputs.

Quality assurance is implemented via a *quality assurance framework*.

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary, ISO, Geneva, 2015.

European Statistics Code of Practice (2017) Eurostat.

National Quality Assurance Frameworks (NQAF) Glossary, UN.

QUALITY ASSURANCE FRAMEWORK

A quality assurance framework (QAF) is management system for implementing quality assurance activities within an organisation.

In the ESS context

The distinguishing characteristics of a QAF are that it:

- provides an umbrella for quality practices;
- refers to a range of surveys/statistical processes or the entire statistical programme rather than a single survey/process; and
- covers all aspects of data processing and output, not just a single aspect.

It typically includes a template that can be used for quality assessment.

A quality assurance framework (QAF), sometimes referred to simply as a *Quality Framework*, covers the statistical outputs, the processes by which they are produced and the organisational environment within which the processes are conducted.

Sources

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary.

National Quality Assurance Frameworks (NQAF) Glossary, UNSD:

QUALITY AUDIT

A quality audit is a systematic, independent and documented process for obtaining evidence that is verifiable and relevant, and for evaluating it objectively to determine the extent to which the quality audit criteria are fulfilled.

Evidence can be in the form of records, statements of fact or other information.

Relevant here means pertinent to the quality audit criteria.

Criteria can be policies, procedures or requirements.

Source(s)

Handbook on Data Quality Assessment Methods and Tools (DatQAM), Eurostat- 2007.

QUALITY CONTROL

Quality control comprise detection, measurement and possible correction of the variability in the characteristics of output that are attributable to the production system.

Quality control is an aspect of quality management that focuses on fulfilling output quality requirements.

Source(s)

Glossary of The Knowledge Base on Statistical Data Editing, Economic Commission for Europe of the United Nations.

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary.

QUALITY CONTROL SURVEY

A quality control survey is a replicated survey which is carried out on a small scale by very experienced staff in order to obtain some "zero-default" results with which the actual results of the survey can be compared.

Source

Handbook on Data Quality Assessment Methods and Tools, Eurostat

QUALITY FRAMEWORK

Same as Quality Assurance Framework

QUALITY IMPROVEMENT

Quality improvement refers to an improvement in one or more quality components of a statistical process.

The perception of a quality improvement may depend on the user in the case where the change results in an improvement in one quality component but a worsening in another.

Source

This Handbook.

QUALITY INDEX

A quality index is a one-dimensional synthesis of information on quality, possibly calculated as a weighted mean of a number of available quality indicators.

Source

Assessment of Quality in Statistics", item 4.2C "Methodological documents, Glossary, Eurostat, Working Group Luxembourg, 2-3 October 2003.

QUALITY INDICATOR

A quality indicator is a quantitative or qualitative measure that give an indication of process or output quality.

Source

Handbook on Data Quality Assessment Methods and Tools (DaTQAM), Eurostat.

QUALITY MANAGEMENT

Quality management comprises all the activities that an organisation uses to direct, control, and coordinate quality.

Quality management includes formulating a quality policy, setting quality objectives, quality planning, quality assurance, and quality improvement. quality control.

Quality management is implemented via a *quality management system*.

In the ESS context

> Quality management refers to the application of a formalised system that documents the structure, responsibilities and procedures put in place for satisfying users, while continuing to improve the data production and dissemination process. It also includes how well resources are used in meeting requirements.
>
> Quality management is deemed to cover a statistical authority as a whole, in contrast to quality assurance which focusses the core business of the authority, i.e. development, production and dissemination of statistics.
>
> In SIMS, quality management (concept S.11) is defined *as the systems and frameworks in place within an organisation to manage the quality of statistical products and processes*.
>
> in SIMS, and in the Handbook, quality management includes *quality assurance (S.11.1), quality assessment (S.11.2) and quality documentation (S.10.7).*

Sources

ISO 9000:2015 Quality Management Systems Fundamentals and Vocabulary.

ISO Standards into Plain English. Praxiom Research Group Limited.

QUALITY MANAGEMENT PRINCIPLES

***Quality management principles* are principles on which a quality management system is based.**

In the ESS context a distinction is made between *(general) quality management principles*, which are typically derived from a general QMS and associated with a statistical authority as a whole, and the *(statistical) principles* that are specified in the ES CoP and are associated with the core statistical environment, processes and outputs.

Sources

ISO 9000:2015 Quality Management Systems Fundamentals and Vocabulary.

ISO Standards into Plain English. Praxiom Research Group Limited.

European Statistics Code of Practice, Eurostat.

QUALITY MANAGEMENT SYSTEM

A quality management system (QMS) is a management system to direct and control an organisation with regard to quality.

A QMS comprises a set of interrelated or interacting elements that organisations use to formulate quality policies and quality objectives and to establish the processes that are needed to ensure that policies are followed and objectives are achieved.

A QMS includes the policies, plans, practices, and the supporting infrastructure by which an organisation aims to reduce and eventually eliminate non-conformance to specifications, standards, and customer expectations in the most cost effective and efficient manner.

A *process-based* QMS uses a *process approach* to manage and control how quality policy is implemented and quality objectives are achieved. It is a network of interrelated and interconnected processes, each of which uses resources to transform inputs into outputs. Since the output of one process becomes the input of another process, processes interact and are interrelated by means of such input-output relationships, thus creating a single process-based QMS.

In the ESS context

A distinction is made between a *general* QMS, which can apply to any organisation no matter what its core business and a *statistical* QMS, which applies exclusively to a statistical authority. The latter is more commonly referred to as a quality management framework, *quality assurance framework*, or simply a quality framework.

Source(s)

ISO 9000:2015 Quality Management Systems Fundamentals and Vocabulary.

European Statistics Code of Practice, Eurostat.

QUALITY MANUAL

A quality manual documents an organisation's quality management system.

A quality manual can vary in detail and format to suit the size and complexity of an individual organization.

It can be a paper manual or an electronic manual.

Source(s)

ISO 9000:2015 Quality Management Systems Fundamentals and Vocabulary.

QUALITY METADATA

Quality metadata are metadata describing the various quality aspects of statistics and of the processes that produce them.

Quality metadata are a type of *reference metadata*.

The ESS Standard for Quality Reports Structure (ESQRS) is the ESS standard for presentation of quality metadata.

Source

Based on SDMX Glossary Version 2.0, October 2018.

QUALITY OBJECTIVE

A quality objective is a quality result that an organisation intends to achieve.

Quality objectives are based on, or derived from, an organization's quality policy and must be consistent with it. They are usually formulated at all relevant levels within the organization and for all relevant functions.

Source(s)

ISO 9000:2015 QUALITY MANAGEMENT SYSTEMS FUNDAMENTALS AND VOCABULARY

QUALITY PLAN

A quality plan is a document that is used to specify the procedures and resources that will be needed to carry out a project, perform a process, realize a product, or manage a contract.

A quality plan also specifies who will do what and when.

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary.

QUALITY PLANNING

Quality planning involves setting quality objectives and then specifying the operational processes and resources that will be needed to achieve those objectives.

Quality planning is a part of quality management.

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary.

QUALITY POLICY

A quality policy defines top management's commitment to quality.

A quality policy should describe an organisation's general quality orientation and clarify its basic intentions.

Quality policies should be used to generate quality objectives and should serve as a general framework for action.

Quality policies can be based on the ISO 9000 Quality Management Principles and should be consistent with the organisation's other policies.

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary.

ISO Translated into Plain English, Praxiom Research Group Limited.

QUALITY PROFILE

A quality profile is a user-oriented summary of the main quality features of statistics.

Quality profiles are developed and disseminated by Eurostat for structural (sustainable development) indicators, Euro indicators (PEEIs) and for the Europe 2020 indicators.

In line with the Eurostat quality concept, quality is defined along several dimensions. For the structural indicators, the quality profile aims at a quick overview on how far a structural indicator is deemed "fit for use" with regard to its key objectives. More information on quality of the indicators, including for some surveys, detailed producer reports and explanations of the applied concepts and methodologies are available under explanatory texts.

Source

Eurostat website.

QUALITY REPORT

A quality report is a report conveying information about the quality of a statistical product or process.

Source

Handbook on Data Quality Assessment Methods and Tools, Eurostat.

REFERENCE METADATA

Reference metadata are metadata describing the contents and the quality of statistical data.

Reference metadata are also called explanatory metadata. They include explanatory texts on the context of the statistical data, methodologies for data collection and data aggregation as well as quality and dissemination characteristics.

Reference metadata include:

- "conceptual" metadata, describing the concepts used and their practical implementation, allowing users to understand what the statistics are measuring and, thus, their fitness for use;
- "methodological" metadata, describing methods used for the generation of the data, for example, sample design, collection methods, editing processes;
- "quality" metadata, describing the different quality dimensions of the resulting statistics, for example, timeliness, accuracy.

Reference metadata do not define the actual structure of a dataset. Structural metadata do this.

In the ESS, the Single Integrated Management System (SIMS) is the standard for presenting reference metadata.

Sources

Based on SDMX Glossary Version 2.0, October 2018.

RELEVANCE

Relevance is an attribute of statistics measuring the degree to which statistics meet current and potential needs of the users.

ES CoP Principle 11 states: *European Statistics meet the needs of users.*

In SIMS S.12 Relevance is defined as *the degree to which statistical information meet current and potential needs of the users*. It is broken down into three sub-concepts:

- S.12.1: *Relevance - user needs*, comprising *a description of users and their respective needs* *with respect to the statistical data*;

- S.12.2 *Relevance - user satisfaction*, comprising *measures to determine user satisfaction*; and
- S.12.3 *Relevance – completeness*, measuring *the extent to which all statistics that are needed are available*.

Relevance is concerned with whether the available statistics shed light on the issues that are important to users. It depends upon the varying needs of users. An NSA's challenge is to weigh and balance the various, possibly conflicting, needs of current and potential users and to produce statistics that satisfy the most important needs within given resource constraints.

The measurement of the availability of the necessary statistics normally refers to datasets and compares the required datasets to the available ones.

In assessing relevance, one approach is to gauge relevance directly, by polling users about the statistics.

Indirect evidence of relevance may be found by ascertaining where there are processes in place to determine the uses of data and the views of their users or to use the data in-house for research and other analysis.

Sources

European Statistics Code of Practice, Eurostat.

National Quality Assurance Frameworks (NQAF) Glossary, UN.

RELIABILITY

Closeness of initially released statistics to the subsequently released statistics.

If initial values are consistently higher or lower than subsequent values, then there is evidence of bias in the initial values.

If the differences are random, but large, then consideration should be given to reassessing the timeliness/reliability trade-off that has been adopted.

Reliability depends upon *revision policy*.

ES CoP Principle 12: Accuracy and Reliability states. *European Statistics accurately and reliably portray reality*. This does not separate the notions of accuracy and reliability.

Sources

National Quality Assurance Frameworks (NQAF) Glossary, UN.

European Statistics Code of Practice, Eurostat.

REVISION POLICY

A policy or set of policies, aimed at ensuring the transparency of disseminated data whereby preliminary data are compiled that are later revised when more and better source data become available.

Providing users with documentation regarding the source data used and the way they are adjusted gives compilers with the possibility to incorporate new and more accurate information into estimates, thus improving their accuracy without introducing breaks in the time series.

Data may also be subject to ad hoc revisions as a result of the introduction of new classifications, compilation frameworks and methodologies which result in the compilation of historical data that replaces previously released data. Whether or not such changes constitute an actual "revision" or the compilation of a "new" series is a matter of judgment on the part of the statistical authority.

Source

Eurostat's Concepts and Definitions Database (CODED): General Statistical Terminology

ROLLING REVIEW

A rolling review is an in-depth assessment done by an external expert, including a user survey and a partner survey.

A rolling review is generally undertaken over time and across statistical programmes to assess their continuing relevance and other quality attributes.

A rolling review is likely to provide a broader assessment of a statistical process than a self-assessment.

Source

National Quality Assurance Frameworks (NQAF) Glossary, UN.

SAMPLING FRAME

Same as survey frame

Survey frame is the preferred term.

Source

Defined for this Handbook.

SELF-ASSESSMENT

A self-assessment is an assessment of an organisation's processes and/or outputs carried out by the organisation itself.

The assessment may be referenced to a model or framework.

It may be comprehensive, or it may be limited in scope. For example, completion of a DESAP questionnaire would be limited in scope unless it were accompanied by in depth review of the quality issues thereby revealed.

Source

Handbook on Data Quality Assessment Methods and Tools (DaTQaM), Eurostat.

SERVICEABILITY

Serviceability is the set of practical aspects describing how well the available data meet users' needs.

Serviceability is a term that captures the practical aspects of usability of data. The emphasis on "use" thus assumes that data are available. Thus, key aspects of usability are relevance, timeliness and frequency, consistency, and revision policy and practices.

Source(s)

Data Quality Assessment Framework, IMF.

SIX SIGMA

Six Sigma is a business management strategy that seeks to improve the quality of process outputs by identifying and removing the causes of defects (errors) and minimizing variability in manufacturing and business processes.

Six Sigma was originally developed by Motorola, USA in 1981. It uses a set of quality management methods, including statistical methods, and creates a special infrastructure of people within the organisation who are experts in these methods. Each Six Sigma project carried out within an organisation follows a defined sequence of steps and has quantified targets. These targets can be financial (cost reduction or profit increase) or whatever is critical to the customer of that process (cycle time, safety, delivery, etc.)

The term six sigma originated from terminology associated with statistical modelling of manufacturing processes. The maturity of a manufacturing process can be described by a *sigma rating* indicating the percentage of defect-free products it creates. A six-sigma process is one in which 99.99966% of the products manufactured are free of defects. Motorola set a goal of "six sigmas" for all of its manufacturing operations and this goal became a byword for the management and engineering practices used to achieve it.

It is often combined with LEAN to form Lean Six Sigma.

Source

Lean Six Sigma, GreyCampus.

SOUND METHODOLOGY

Sound methodology implies that the methodology used to compile statistics is in accordance with scientific principles and complies with the relevant European and other international standards, guidelines, and good practices.

Sound methodology is also referred to as *methodological soundness*.

ES CoP Principle 7 states:

> *sound methodology underpins quality statistics. This requires adequate tools, procedures and expertise.*

Having sound methodology includes constantly striving for innovation.

Source

National Quality Assurance Frameworks (NQAF) Glossary, UN.

European Statistics Code of Practice, Eurostat.

STATISTICAL CONFIDENTIALITY

Statistical confidentiality means that data related to individual units, obtained directly for statistical purposes or indirectly from administrative or other sources, are not disclosed and their use for non-statistical purposes is prohibited.

The unit may be a physical or legal person. It may be a business or other organisation.

ES CoP Principle 5 Statistical Confidentiality and Data Protection states

> *The privacy of data providers, the confidentiality of the information they provide, its use only for statistical purposes and the security of the data are absolutely guaranteed.*

Confidential is a property of a cell in an output dataset implying that its disclosure is subject to restrictions. In particular, a data value allowing the identification of a physical or legal person, either directly or indirectly, may be characterised as confidential.

SIMS S.07 defines Confidentiality as a *property of data indicating the extent to which their unauthorised disclosure could be prejudicial or harmful to the interest of the source or other relevant parties*. It subdivides confidentiality into:

S.07.1 Confidentiality – policy: *legislative measures or other formal procedures which prevent unauthorised disclosure of data that identify a person or economic entity either directly or indirectly*;

S.07.2 Confidentiality – data treatment: *rules applied for treating the data set to ensure statistical confidentiality and prevent unauthorised disclosure*.

Source(s)

European Statistics Code of Practice, Eurostat.

SDMX Glossary Version 2.0, October 2018.

STATISTICAL OUTPUT QUALITY

Statistical output quality is the degree to which available statistics meet users' needs.

The five ES CoP principles relating to statistical output are prefaced by a statement that:

> *Statistics comply with the European quality standards and serve the needs of European institutions, governments, research institutions, business concerns and the public generally*

Within the five principles the ES CoP indicates nine output quality components, namely relevance; accuracy, reliability, timeliness, punctuality, coherence, comparability, accessibility and clarity.

Source

European Statistics Code of Practice, Eurostat.

STATISTICAL PROCESS QUALITY

Process quality is the degree to which a set of inherent characteristics fulfils process requirements.

The ES CoP defines four principles that relate to process quality:

- sound methodology;
- appropriate statistical procedures;
- non-excessive burden on respondents; and
- cost effectiveness.

Source

European Statistics Code of Practice, Eurostat.

STATISTICAL BUSINESS PROCESS

(Same as Statistical Production Process.)

STATISTICAL PROCESS

(In the EHQMR, same as Statistical Production Process,)

STATISTICAL PRODUCTION PROCESS

A statistical production process is a complete set of interrelated or interacting sub-processes that transform inputs into statistics.

A statistical production process take place in the institutional environment (characterised by external or internal regulations) using necessary resources (human, financial, ICT).

In the EHQMR a statistical production process is referred to as a statistical process.

Sources

Generic Statistical Business Process Model v5.1, UNECE

National Quality Assurance Frameworks (NQAF) Glossary, UN

STATISTICAL UNIT

Entity for which information is sought and for which statistics are ultimately compiled.

For any particular statistical process, the statistical unit may be referred to as the *target statistical unit*.

Source

SIMS V2.0.

STRUCTURAL METADATA

Structural metadata are metadata that identify and describe data and reference metadata.

Structural metadata are needed and used to identify, formally describe or retrieve statistical data, such as dimension names, variable names, dictionaries, dataset technical descriptions, dataset locations, keywords for finding data etc. For example, structural metadata includes the titles of the variables and dimensions of statistical datasets, as well as the units employed, code lists (e.g. for territorial coding), data formats, potential value ranges, time dimensions, value ranges of flags, classifications used, etc

Structural metadata are needed to identify, use, and process data matrixes and data cubes, including. names of columns or dimensions of statistical cubes.

Sources

SDMX Glossary Version 2.0, October 2018.

Eurostat's Concepts and Definitions Database (CODED): General Statistical Terminology

SUPPORTED SELF-ASSESSMENT

Supported self-assessment is a special type of self-assessment, in which a quality expert from outside the production unit supports the assessment process.

In Eurostat, a supported self-assessment is conducted by a statistical production unit with the active participation of the Quality Unit.

Source

Handbook on Data Quality Assessment Methods and Tools (DaTQAM), Eurostat

SURVEY FRAME

The *survey frame* is the set of units in the survey population together with all the information required for sample selection and for contact of the selected units.

The survey frame is also referred to simply as the *frame,*

Source

Consistent with Särndal et al (1992).

SURVEY POPULATION

The survey population is the list of units that is the closest approximation to the target population that can be created in practice.

The survey population usually differs from the *target population* as a result of imperfections in the coverage that can be obtained in practice. The difference is referred to as *coverage error*.

The survey population comprises the set of units in the *survey frame*.

Source

Särndal et al (1992).

TARGET POPULATION

The target population is the set of units for which the producer decides that information is wanted.

The target population may be conceptually different from the *ideal population* sought by a user as the producer has to balance different user needs with each other and with practical constraints.

Source

Särndal et al (1992).

TIMELINESS

Timeliness is an attribute of statistics measuring the period between the availability of the information and the event or phenomenon it describes.

ES CoP Principle 13: Timeliness and Punctuality states: *European Statistics are released in a timely and punctual manner*.

SIMS concept S.14.1 Timeliness defines timeliness as *the length of time between data availability and the event or phenomenon they describe*.

Source

Regulation (EC) No 223/2009 on European statistics amended by Regulation 2015/759

European Statistics Code of Practice, Eurostat.

TOTAL QUALITY MANAGEMENT

Total quality management (TQM) is a management philosophy that is based quality management principles, that is driven by customer needs and expectations, and that aims to create a quality culture.

At the core of the TQM approach is the concept of continuous improvement, often illustrated by the Plan-Do-Check-Act cycle made popular by Deming.

Source(s)

ISO 9000:2015 Quality management systems -- Fundamentals and vocabulary.

National Quality Assurance Frameworks (NQAF) Glossary, UN.

TRANSPARENCY

The condition in which all policies and practices surrounding a statistical process are made known to the stakeholders, particularly the respondents and the users.

Transparency involves ensuring that respondents know the legal basis for data collection, the purposes for which the data are required, and the measures taken to protect confidentiality.

Source

National Quality Assurance Frameworks (NQAF) Glossary, UN.

USER SATISFACTION SURVEY

A user satisfaction survey is a survey which aims at assessing the satisfaction or the perception of the users, normally as a basis for improvement actions.

User satisfaction surveys can take several different forms for example using standardised questionnaires or through qualitative interviews.

In interpreting the results of such surveys it is important to weight the views of users according to their significance to the NSA and its objectives.

Source

Handbook on Data Quality Assessment Methods and Tools, Eurostat.

B (PART III) Single Integrated Metadata Structure

B1 Introductory remarks

For ease of reference:

- Supplementary Document B2 contains all SIMS V2.0 concepts (including sub-concepts) on a single page.
- Supplementary Document B3 contains all ESQRS V2.0 concepts (including sub-concepts) on a single page.
- Supplementary Document B4 contains the ESMS V2.0 concepts (including sub-concepts) on a single page.

Section B5 shows the relationships between SIMS, ESMS and ESQRS concepts and sub-concepts.

- Every ESQRS concept corresponds to SIMS concept.
- Every ESQRS sub-concept has a one to one relationship with a SIMS sub-concept.
- The ESMS and SIMS concepts are identical.
- Every ESMS sub-concept corresponds to a SIMS sub-concept.
- There are no SIMS sub-concepts that are not in ESMS or ESQRS.

Supplementary Document B6 brings the SIMS concepts and sub-concepts and their definitions together with the SIMS reporting guidelines presented in Part II. As noted in Section 1.1 of Part I, the SIMS definitions are exactly as in V2.0 but the corresponding guidelines have been reviewed and revised for this document.

The definitions and guidelines for ESMS and ESQRS concepts and sub-concepts are those associated with the corresponding SIMS concepts and sub-concepts, as presented in B5.

For the sub-concepts that are common to ESMS and ESQRS, and for the concepts without sub-concepts that are common to ESMS and ESQRS, the guidelines are identical and the *once for all purposes reporting* approach is applicable.

B2 Single Integrated Metadata Structure (ESMS) V2.0: Concepts

Item No	Concept name
S.01	**Contact**
S.01.1	Contact organisation
S.01.2	Contact organisation unit
S.01.3	Contact name
S.01.4	Contact person function
S.01.5	Contact mail address
S.01.6	Contact email address
S.01.7	Contact phone number
S.01.8	Contact fax number
S.02	**Metadata update**
S.02.1	Metadata last certified
S.02.2	Metadata last posted
S.02.3	Metadata last update
S.03	**Statistical presentation**
S.03.1	Data description
S.03.2	Classification system
S.03.3	Sector coverage
S.03.4	Statistical concepts and definitions
S.03.5	Statistical unit
S.03.6	Statistical population
S.03.7	Reference area
S.03.8	Time coverage
S.03.9	Base period
S.04	**Unit of measure**
S.05	**Reference period**
S.06	**Institutional mandate**
S.06.1	Legal acts and other agreements
S.06.2	Data sharing
S.07	**Confidentiality**
S.07.1	Confidentiality - policy
S.07.2	Confidentiality - data treatment
S0.8	**Release policy**
S.08.1	Release calendar
S.08.2	Release calendar access
S.08.3	User access
S.09	**Frequency of dissemination**
S.10	**Accessibility and clarity**
S.10.1	News release
S.10.2	Publications
S.10.3	On-line database
S.10.3.1	AC1. Data tables - consultations
S.10.4	Micro-data access
S.10.5	Other
S.10.5.1	AC 2. Metadata - consultations
S.10.6	Documentation on methodology
S.10.6.1	AC 3. Metadata completeness - rate
S.10.7	Quality documentation
S.11	**Quality management**
S.11.1	Quality assurance
S.11.2	Quality assessment

Item No	Concept name
S.12	**Relevance**
S.12.1	User needs
S.12.2	User satisfaction
S.12.3	Completeness and R1. Data completeness - rate for U
S.12.3.1	R1. Data completeness - rate for P
S.13	**Accuracy and reliability**
S.13.1	Overall accuracy
S.13.2	Sampling error and A1. Sampling errors - indicators for U
S.13.2.1	A1. Sampling errors - indicators for P
S.13.3	Non-sampling error and A4. Unit non-response - rate for U and A5. Item non-response - rate for U
S.13.3.1	Coverage error
S.13.3.1.1	A2. Over-coverage - rate
S.13.3.1.2	A3. Common units - proportion
S.13.3.2	Measurement error
S.13.3.3	Non response error
S.13.3.3.1	A4. Unit non-response - rate for P
S.13.3.3.2	A5. Item non-response - rate for P
S.13.3.4	Processing error
S.13.3.5	Model assumption error
S.14	**Timeliness and punctuality**
S.14.1	Timeliness and TP2. Time lag - final results for U
S.14.1.1	TP1. Time lag - first results for P
S.14.1.2	TP2. Time lag - final results for P
S.14.2	Punctuality and TP3. Punctuality - delivery and publication for U
S.14.2.1	TP3. Punctuality - delivery and publication for P
S.15	**Coherence and comparability**
S.15.1	Comparability - geographical
S.15.1.1	CC1. Asymmetry for mirror flows statistics - coefficient
S.15.2	Comparability - over time and CC2. Length of comparable time series for U
S.15.2.1	CC2. Length of comparable time series for P
S.15.3	Coherence- cross domain
S.15.3.1	Coherence - sub annual and annual statistics
S.15.3.2	Coherence- National Accounts
S.15.4	Coherence - internal
S.16	**Cost and burden**
S.17	**Data revision**
S.17.1	Data revision - policy
S.17.2	Data revision - practice and A6. Data revision – average size for U
S.17.2.1	A6. Data revision - average size for P
S.18	**Statistical processing**
S.18.1	Source data
S.18.2	Frequency of data collection
S.18.3	Data collection
S.18.4	Data validation
S.18.5	Data compilation
S.18.5.1	A7. Imputation - rate
S.18.6	Adjustment
S.18.6.1	Seasonal adjustment
S.19	**Comment**

B3 ESS Standard for Quality Reports (ESQRS) V2.0: Concepts

Item No	Concept name
1	**Contact**
1.1	Contact organisation
1.2	Contact organisation unit
1.3	Contact name
1.4	Contact person function
1.5	Contact mail address
1.6	Contact email address
1.7	Contact phone number
1.8	Contact fax number
2	**Statistical presentation**
2.1	Data description
2.2	Classification system
2.3	Sector coverage
2.4	Statistical concepts and definitions
2.5	Statistical unit
2.6	Statistical population
2.7	Reference area
2.8	Time coverage
2.9	Base period
3	**Statistical processing**
3.1	Source data
3.2	Frequency of data collection
3.3	Data collection
3.4	Data validation
3.5	Data compilation
3.6	Adjustment
4	**Quality management**
4.1	Quality assurance
4.2	Quality assessment
5	**Relevance**
5.1	User Needs
5.2	User Satisfaction
5.3	Completeness
5.3.1	Data completeness - rate
6	**Accuracy and reliability**
6.1	Accuracy - overall
6.2	Sampling error
6.2.1	Sampling error - indicators
6.3	Non-sampling error
6.3.1	Coverage error
6.3.1.1	Over-coverage - rate
6.3.1.2	Common units - proportion
6.3.2	Measurement error

Item No	Concept name
6	**Accuracy and reliability (continued)**
6.3.3	Non response error
6.3.3.1	Unit non-response - rate
6.3.3.2	Item non-response - rate
6.3.4	Processing error
6.3.4.1	Imputation – rate
6.3.5	Model assumption error
6.4	Seasonal adjustment
6.5	Data revision - policy
6.6	Data revision - practice
6.6.1	Data revision - average size
7	**Timeliness and punctuality**
7.1	Timeliness
7.1.1	Time lag - first result
7.1.2	Time lag - final result
7.2	Punctuality
7.2.1	Punctuality - delivery and publication
8	**Coherence and comparability**
8.1	Comparability - geographical
8.1.1	Asymmetry for mirror flow statistics - coefficient
8.2	Comparability - over time
8.2.1	Length of comparable time series
8.3	Coherence - cross domain
8.4	Coherence - sub annual and annual statistics
8.5	Coherence - National Accounts
8.6	Coherence - internal
9	**Accessibility and clarity**
9.1	News release
9.2	Publications
9.3	Online database
9.3.1	Data tables - consultations
9.4	Microdata access
9.5	Other
9.6	Documentation on methodology
9.7	Quality documentation
9.7.1	Metadata completeness - rate
9.7.2	Metadata - consultations
10	**Cost and Burden**
11	**Confidentiality**
11.1	Confidentiality - policy
11.2	Confidentiality - data treatment
12	**Comment**

B4 Euro-SDMX Metadata Structure V2: Concepts

Item No	Concept name
1	**Contact**
1.1	Contact organisation
1.2	Contact organisation unit
1.3	Contact name
1.4	Contact person function
1.5	Contact mail address
1.6	Contact email address
1.7	Contact phone number
1.8	Contact fax number
2	**Metadata update**
2.1	Metadata last certified
2.2	Metadata last posted
2.3	Metadata last update
3	**Statistical presentation**
3.1	Data description
3.2	Classification system
3.3	Sector coverage
3.4	Statistical concepts and definitions
3.5	Statistical unit
3.6	Statistical population
3.7	Reference area
3.8	Time coverage
3.9	Base period
4	**Unit of measure**
5	**Reference period**
6	**Institutional mandate**
6.1	Legal acts and other agreements
6.2	Data sharing
7	**Confidentiality**
7.1	Confidentiality - policy
7.2	Confidentiality - data treatment
8	**Release policy**
8.1	Release calendar
8.2	Release calendar access
8.3	User access
9	**Frequency of dissemination**

Item No	Concept name
10	**Accessibility and clarity**
10.1	News release
10.2	Publications
10.3	On-line database
10.4	Micro-data access
10.5	Other
10.6	Documentation on methodology
10.7	Quality documentation
11	**Quality management**
11.1	Quality assurance
11.2	Quality assessment
12	**Relevance**
12.1	User needs
12.2	User satisfaction
12.3	Completeness
13	**Accuracy and reliability**
13.1	Overall accuracy
13.2	Sampling error
13.3	Non-sampling error
14	**Timeliness and punctuality**
14.1	Timeliness
14.2	Punctuality
15	**Coherence and comparability**
15.1	Comparability - geographical
15.2	Comparability - over time
15.3	Coherence - cross domain
15.4	Coherence - internal
16	**Cost and burden**
17	**Data revision**
17.1	Data revision - policy
17.2	Data revision - practice
18	**Statistical processing**
18.1	Source data
18.2	Frequency of data collection
18.3	Data collection
18.4	Data validation
18.5	Data compilation
18.6	Adjustment
19	**Comment**

B5 SIMS, ESMS and ESQRS inter-relationships

EURO-SDMX Metadata Structure		Single Integrated Metadata Structure		ESS Standard for Quality Reports Structure	
1	Contact	S.01	Contact	1	Contact
1.1	Contact organisation	S.01.1	Contact organisation	1.1	Contact organisation
1.2	Contact organisation unit	S.01.2	Contact organisation unit	1.2	Contact organisation unit
1.3	Contact name	S.01.3	Contact name	1.3	Contact name
1.4	Contact person function	S.01.4	Contact person function	1.4	Contact person function
1.5	Contact mail address	S.01.5	Contact mail address	1.5	Contact mail address
1.6	Contact email address	S.01.6	Contact email address	1.6	Contact email address
1.7	Contact phone number	S.01.7	Contact phone number	1.7	Contact phone number
1.8	Contact fax number	S.01.8	Contact fax number	1.8	Contact fax number
2	Metadata update	S.02	Metadata update		
2.1	Metadata last certified	S.02.1	Metadata last certified		
2.2	Metadata last posted	S.02.2	Metadata last posted		
2.3	Metadata last update	S.02.3	Metadata last update		
3	Statistical presentation	S.03	Statistical presentation	2	Statistical presentation
3.1	Data description	S.03.1	Data description	2.1	Data description
3.2	Classification system	S.03.2	Classification system	2.2	Classification system
3.3	Sector coverage	S.03.3	Sector coverage	2.3	Sector coverage
3.4	Statistical concepts and definitions	S0.3.4	Statistical concepts and definitions	2.4	Statistical concepts and definitions
3.5	Statistical unit	S.03.5	Statistical unit	2.5	Statistical unit
3.6	Statistical population	S.03.6	Statistical population	2.6	Statistical population
3.7	Reference area	S.03.7	Reference area	2.7	Reference area
3.8	Time coverage	S.03.8	Time coverage	2.8	Time coverage
3.9	Base period	S.03.9	Base period	2.9	Base period
4	Unit of measure	S.04	Unit of measure		
5	Reference period	S.05	Reference period		
6	Institutional mandate	S.06	Institutional mandate		
6.1	Legal acts and other agreements	S.06.1	Legal acts and other agreements		
6.2	Data sharing	S06.2	Data sharing		
7	Confidentiality	S.07	Confidentiality	11	Confidentiality
7.1	Confidentiality - policy	S.07.1	Confidentiality - policy	11.1	Confidentiality - policy
7.2	Confidentiality - data treatment	S.07.2	Confidentiality - data treatment	11.2	Confidentiality - data treatment

EURO-SDMX Metadata Structure		Single Integrated Metadata Structure		ESS Standard for Quality Reports Structure	
8	Release policy	S.08	Release policy		
8.1	Release calendar	S.08.1	Release calendar		
8.2	Release calendar access	S.08.2	Release calendar access		
8.3	User access	S.08.3	User access		
9	Frequency of dissemination	S.09	Frequency of dissemination		
10	Accessibility and clarity	S.10	Accessibility and clarity	9	Accessibility and clarity
10.1	News release	S.10.1	News release	9.1	News release
10.2	Publications	S.10.2	Publications	9.2	Publication
10.3	On-line database	S.10.3	On-line database	9.3	On-line database
		S.10.3.1	AC1. Data tables - consultations	9.3.1	Data tables - consultations
10.4	Micro-data access	S.10.4	Micro-data access	9.4	Micro-data access
10.5	Other	S.10.5	Other	9.5	Other
		S.10.5.1	AC 2. Metadata - consultations	9.7.2	Metadata - consultations
10.6	Documentation on methodology	S.10.6	Documentation on methodology	9.6	Documentation on methodology
		S.10.6.1	AC 3. Metadata completeness - rate	9.7.1	Metadata completeness – rate
10.7	Quality documentation	S.10.7	Quality documentation	9.7	Quality documentation
11	Quality management	S.11	Quality management	4	Quality management
11.1	Quality assurance	S.11.1	Quality assurance	4.1	Quality assurance
11.2	Quality assessment	S.11.2	Quality assessment	4.2	Quality assessment
12	Relevance	S.12	Relevance	5	Relevance
12.1	User needs	S.12.1	User needs	5.1	User needs
12.2	User satisfaction	S.12.2	User satisfaction	5.2	User satisfaction
12.3	Completeness	S.12.3	Completeness and R1. Data completeness - rate for U	5.3	Completeness
		S.12.3.1	R1. Data completeness - rate for P	5.3.1	Data completeness - rate
13	Accuracy and reliability	S.13	Accuracy and reliability	6	Accuracy and reliability
13.1	Overall accuracy	S.13.1	Overall accuracy	6.1	Overall accuracy
13.2	Sampling error	S.13.2	Sampling error and A1. Sampling errors - indicators for U	6.2	Sampling error
		S.13.2.1	A1. Sampling errors - indicators for P	6.2.1	Sampling errors - indicators

EURO-SDMX Metadata Structure		Single Integrated Metadata Structure		ESS Standard for Quality Reports Structure	
13.3	Non-sampling error	S.13.3	Non-sampling error and A4. Unit non-response - rate for U and A5. Item non-response - rate for U	6.3	Non-sampling error
		S.13.3.1	Coverage error	6.3.1	Coverage error
		S.13.3.1.1	A2. Over-coverage - rate	6.3.1.1	Over-coverage - rate
		S.13.3.1.2	A3. Common units - proportion	6.3.1.2	Common units - proportion
		S.13.3.2	Measurement error	6.3.2	Measurement error
		S.13.3.3	Nonresponse error	6.3.3	Nonresponse error
		S.13.3.3.1	A4. Unit non-response - rate for P	6.3.3.1	Unit non-response - rate
		S.13.3.3.2	A5. Item non-response - rate for P	6.3.3.2	Item non-response - rate
		S.13.3.4	Processing error	6.3.4	Processing error
			(See S.18.5.1)	6.3.4.1	Imputation - rate
		S.13.3.5	Model assumption error	6.3.5	Model assumption error
			(See S.18.6.1)	6.4	Seasonal adjustment
14	Timeliness and punctuality	S.14	Timeliness and punctuality	7	Timeliness and punctuality
14.1	Timeliness	S.14.1	Timeliness and TP2. Time lag - final results for U	7.1	Timeliness
		S.14.1.1	TP1. Time lag - first results for P	7.1.1	Time lag - first results
		S.14.1.2	TP2. Time lag - final results for P	7.1.2	Time lag - final results
14.2	Punctuality	S.14.2	Punctuality and TP3. Punctuality - delivery and publication for U	7.2	Punctuality
		S.14.2.1	TP3. Punctuality - delivery and publication for P	7.2.1	Punctuality - delivery and publication
15	Coherence and comparability	S.15	Coherence and comparability	8	Coherence and comparability
15.1	Comparability - geographical	S.15.1	Comparability - geographical	8.1	Comparability - geographical
		S.15.1.1	CC1. Asymmetry for mirror flows statistics - coefficient	8.1.1	Asymmetry for mirror flows statistics - coefficient
15.2	Comparability - over time	S.15.2	Comparability - over time and *CC2. Length of* comparable time series for U	8.2	Comparability - over time
		S.15.2.1	CC2. Length of comparable time series for P	8.2.1	Length of comparable time series
15.3	Coherence - cross domain	S.15.3	Coherence- cross domain	8.3	Coherence- cross domain

EURO-SDMX Metadata Structure		Single Integrated Metadata Structure		ESS Standard for Quality Reports Structure	
		S.15.3.1	Coherence - sub annual and annual statistics	8.4	Coherence - sub annual and annual statistics
		S.15.3.2	Coherence- National Accounts	8.5	Coherence- National Accounts
15.4	Coherence - internal	S.15.4	Coherence - internal	8.6	Coherence - internal
16	Cost and burden	S.16	Cost and burden	10	Cost and burden
17	Data revision	S.17	Data revision		
17.1	Data revision - policy	S.17.1	Data revision - policy	6.5	Data revision - policy
17.2	Data revision - practice	S.17.2	Data revision - practice and A6. Data revision - average size for U	6.6	Data revision- practice
		S.17.2.1	A6. Data revision - average size for P	6.6.1	Data revision - average size
18	Statistical processing	S.18	Statistical processing	3	Statistical Processing
18.1	Source data	S.18.1	Source data	3.1	Source data
18.2	Frequency of data collection	S.18.2	Frequency of data collection	3.2	Frequency of data collection
18.3	Data collection	S.18.3	Data collection	3.3	Data collection
18.4	Data validation	S.18.4	Data validation	3.4	Data validation
18.5	Data compilation	S.18.5	Data compilation	3.5	Data compilation
		S.18.5.1	A7. Imputation - rate		(See 6.3.4.1)
18.6	Adjustment	S.18.6	Adjustment	3.6	Adjustment
		S.18.6.1	Seasonal adjustment		(See 6.4)
19	Comment	S.19	Comment	12	Comment

Legend

Common concepts in SIMS, ESMS and ESQRS
Common concepts in SIMS and ESMS
Common concepts in SIMS and ESQRS

B6 SIMS Concepts, definitions and guidelines

SIMS	Concept name	Definition	Guidelines
S.01	Contact	Individual or organisational contact points for the data or metadata, including information on how to reach the contact points.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.01.1	Contact organisation	The name of the organisation of the contact points for the data or metadata.	Provide the full name (not just code name). of organisation responsible for the process and outputs (data and metadata) that are the subject of the report.
S.01.2	Contact organisation unit	An addressable subdivision of an organisation.	Provide the full name of the organisational unit responsible. The name can include a unit number.
S.01.3	Contact name	The names of the contact points for the data or metadata.	Provide the first and last names of the contact point(s). If more than one name is provided, the main contact should be indicated. If the author of the report is different from the person(s) responsible for process and its outputs, provide this name also.
S.01.4	Contact person function	The area of technical responsibility of the contact, such as "methodology", "database management" or "dissemination".	Provide the title(s) and area(s) of responsibility of the person(s) indicated as contact(s), for example Senior Research Assistant, Economics Division.
S.01.5	Contact mail address	The postal address of the contact points for the data or metadata.	Provide the postal address(es) of the person(s) indicated as contacts.
S.01.6	Contact email address	E-mail address of the contact points for the data or metadata.	Provide the email address(es) of the person(s) indicated as contacts. The address can be an 1individual e-mail address or a mailbox for the organisation to which the person has access.
S.01.7	Contact phone number	The telephone number of the contact points for the data or metadata.	Provide the telephone number(s) of the person(s) indicated as contacts.

SIMS	Concept name	Definition	Guidelines
S.01.8	Contact fax number	Fax number of the contact points for the data or metadata.	Provide the fax number(s) of the person(s) indicated as contacts.

S.02	Metadata update (U)	The date on which the metadata element was inserted or modified in the database.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.02.1	Metadata last certified (U)	Date of the latest certification provided by the domain manager to confirm that the metadata posted are still up-to-date, even if the content has not been amended.	Certification can be provided even if the metadata have not been amended since the previous certification. *European Level* Certification for European level metadata.
S.02.2	Metadata last posted (U)	Date of the latest dissemination of the metadata.	The date when the complete set of metadata was last disseminated as a block should be provided (manually, or automatically by the metadata system). *European level* Date refers to the European level metadata.
S.02.3	Metadata last update (U)	Date of last update of the content of the metadata.	The date when any metadata were last updated should be provided (manually, or automatically by the metadata system). *European level* Date refers to the European level metadata.

S.03	Statistical presentation	Description of the disseminated data which can be displayed to users as tables, graphs or maps.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.03.1	Data description	Main characteristics of the data set, referring to the	Describe briefly the main characteristics of the data in an easily and quickly understandable manner, referring to the main variables disseminated. More detailed descriptions of the variables are in S.03.4.

SIMS	Concept name	Definition	Guidelines
		data and indicators disseminated.	
S.03.2	Classification system	Arrangement or division of objects into groups based on characteristics which the objects have in common.	List all classifications and breakdowns that are used in the data (with their detailed names) and provide links (if publicly available). Explain deviations, if any, from ESS or international standards. *European level* Provide an overview of national deviations from ESS and/or international standards.
S.03.3	Sector coverage	Main economic or other sectors covered by the statistics.	List the main economic or other sectors covered by the data and the size classes used, for example, size classes based on number of employees. *European level* Provide a summary of differences in the main economic or other sectors covered by national data and the size classes used.
S.03.4	Statistical concepts and definitions	Statistical characteristics of statistical observations, variables.	Define and describe briefly the main statistical variables that have been observed or derived. Indicate their types. Indicate discrepancies, if any, from the ESS or international standards. Note that any difference between these variables and the variables desired by users is a relevance issue and is discussed in S.12. *European level* Summarise the national discrepancies from the ESS and/or international standards.
S.3.5	Statistical unit	Entity for which information is sought and for which statistics are ultimately compiled.	Define the type of statistical unit about which data are collected, e.g. enterprise, kind of activity unit, local unit, private household, dwelling, person, import transaction. If there is more than one type of unit, define each type. *European level* Summarise the differences in units used at national level.
S.3.6	Statistical population	The total membership or population or "universe" of a defined class of people, objects or events.	Define the *target population* of statistical units for which information is sought. Note that a difference between the *target population* and the *ideal population* desired by users is a relevance issue and is discussed in S.12; and the difference between *target population* and the *survey population* is a coverage issue and is discussed in S13.3

SIMS	Concept name	Definition	Guidelines
			If there is more than one type of statistical population, define each type. *European level* Summarise the differences in statistical populations used at national level.
S.3.7	Reference area	The country or geographic area to which the measured statistical phenomenon relates.	Describe the country, the regions, the districts, or the other geographical aggregates, to which the data refer. Identify any specific exclusions in the data disseminated. If coverage includes overseas territories this should be stated, and they should be specified. *European level* Describe the geographical area covered by the data disseminated, e.g., EU Members states, EU regions, USA, Japan, or aggregates such as EU, EEA).
S.3.8	Time coverage	The length of time for which data are available.	State the time period(s) covered by the data, e.g. first quarter 2018, or quarters 2015-2018, or year 2018, or years 1985-2018. Note that any issues concerning comparability over time are discussed in S.15.
S.3.9	Base period	The period of time used as the base of an index number, or to which a constant series refers.	Note that this concept applies only to certain types of outputs, such as indexes. State the base period, for example, year 2000. Indicate base period update time frame and date of next update.
S.04	Unit of measure (U)	The unit in which the data values are measured.	The data usually involves several units of measure depending upon the variables. Examples are: Euro, national currency, number of persons, and rate per 100,000 inhabitants. The magnitude (e.g., thousand, million) of numerical units should be included.
S.05	Reference period (U)	The period of time or point in time to which the measured observation is intended to refer.	The value of a variable refers to a specific time period (for example, the last week of a month, a month, a fiscal year, a calendar year, or several calendar years), or to a point in time (for example, a specific day, or the last day of a month). The variables in a dataset may refer to more than one reference period. All reference periods should be stated

SIMS	Concept name	Definition	Guidelines
			Note that the difference, if any, between the target reference period(s) and the actual reference period(s) is an accuracy issue and should be discussed in S.13.3. Note that if the survey population does not include all the units in the target population for the specified reference period, this is a coverage issue and should be discussed in S.13.3. *European level* Summarise differences in reference period across countries.
S.06	Institutional mandate (U)	Law, set of rules or other formal set of instructions assigning responsibility as well as the authority to an organisation for the collection, processing, and dissemination of statistics.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.06.1	Legal acts and other agreements (U)	Legal acts or other formal or informal agreements that assign responsibility as well as the authority to an agency for the collection, processing, and dissemination of statistics.	State the national legal acts and/or other reporting agreements, including EU legal acts, the implementation of EU directives. *European level* State the legal base or other agreement, for example the EU legal act, or ESS Five-Year-Program, that underpins the reporting obligations on countries.)
S.06.2	Data sharing (U)	Arrangements or procedures for data sharing and coordination between data producing agencies.	Describe the arrangements, procedures or agreements related to data sharing and exchange between data producing agencies within the national statistical system, *European level* Describe the arrangements, procedures or agreements related to data sharing and exchange between international data producing agencies, for example, a Eurostat data collection or production that is in common with the OECD or the UN.

SIMS	Concept name	Definition	Guidelines
S.07	Confidentiality	A property of data indicating the extent to which their unauthorised disclosure could be prejudicial or harmful to the interest of the source or other relevant parties.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.07.1	Confidentiality – policy	Legislative measures or other formal procedures which prevent unauthorised disclosure of data that identify a person or economic entity either directly or indirectly.	Describe all European or national legislation, or other formal requirements, that relate to confidentiality. Describe relevant policy (if any). Note that the existence of legislation and/or policy provides some assurance that methods necessary to assure confidentiality have been applied to the data. *European level* Summarise the commonalties and differences in national approaches to confidentiality policy.
S.07.2	Confidentiality - data treatment	Rules applied for treating the datasets to ensure statistical confidentiality and prevent unauthorised disclosure.	*For aggregate outputs* Provide the rules that define a *confidential cell*. Describe the procedures for detecting confidential cells, including checking for residual disclosure. Describe the procedures for reducing the risk of disclosure by treating confidential cells, for example by perturbation, controlled rounding, cell suppression, or cell aggregation. *For micro-level outputs:* Describe the procedures that are used in protecting confidentiality. *European level* Summarise the commonalties and differences in national approaches.
S.08	Release policy (U)	Rules for disseminating statistical data to all interested parties.	(Information relating to this concept is provided by reporting on its sub-concepts.)

SIMS	Concept name	Definition	Guidelines
S.08.1	Release calendar (U)	The schedule of statistical release dates.	State whether there is a release calendar for the statistical outputs from the process being reported, and if so, whether this calendar is publicly accessible.
S.08.2	Release calendar access (U)	Access to the release calendar information.	Give a link or reference to the release calendar (if any).
S.08.3	User access (U)	The policy for release of the data to users, the scope of dissemination, how users are informed that the data are being released, and whether the policy determines the dissemination of statistical data to all users.	Describe the general data release policy of the organisation. Describe the release policy applied to the outputs of the process being reported, highlighting any deviations from the general policy. Note that the effect of not having a release calendar, or whether releases have been in accordance with a release calendar, is reported in S.14.2. *European level* Summarise country adherence to the impartiality protocol linked to Principle 6 of the European Statistics Code of Practice, which requires those responsible for the statistical domain to make public any and all kinds of pre-release.)
S.09	Frequency of dissemination (U)	The time interval at which the statistics are disseminated over a given time period.	State the frequency with which the data are disseminated, e.g. monthly, quarterly, yearly. The frequency can also be expressed by using a code from the harmonised ESS code list so long as this is considered to be easily understandable by users.
S.10	Accessibility and clarity	The conditions and modalities by which users can access, use and interpret data.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.10.1	News release	Regular or ad-hoc press releases linked to the data.	List any regular or ad-hoc press releases linked to the data over the past year.
S.10.2	Publication	Regular or ad-hoc publications in which the data are made available to the public.	List the titles of any publications, including publisher, year, and links to on-line documents (if available). Provide number of subscriptions/purchases of each of the key paper reports.

SIMS	Concept name	Definition	Guidelines
S.10.3	On-line database	Information about on-line databases in which the disseminated data can be accessed.	Provide the domain name and link to the on-line database (if any). Provide number of accesses to on-line databases.
S.10.3 .1	AC1. Data tables – consultations (P)	Number of consultations of data tables within a statistical domain for a given time period displayed in a graph.	*For producer reports only* Provide values of Indicator AC1 by month/quarter/year. The indicator AC1 is defined in Supplementary Document C.
S.10.4	Micro-data access	Information on whether micro-data are also disseminated.	State whether the data are accessible in micro-data form, e.g. for researchers. If so, cross reference the micro-data confidentiality rules in S.7.
S.10.5	Other	References to the most important other data dissemination done.	Describe any other important dissemination mechanisms, for example policy papers, within outputs produced by other statistical processes. Summarise the accessibility and clarity of the data associated with the various dissemination formats, and the effects of pricing policies and confidentiality provisions. Describe dissemination of data to Eurostat and other international organisations, and internal dissemination.
S.10.5 .1	AC2. Metadata – consultations (P)	Number of metadata consultations within a statistical domain for a given time period.	*For producer reports only*: provide values of Indicator AC2 by month/quarter/year.
S.10.6	Documentation on methodology	Descriptive text and references to methodological documents available.	List national reference metadata files, methodological papers, summary documents and handbooks relevant to the statistical process. For each item provide the title, publisher, year and link to on-line version (if any).
S.10.6 .1	AC3. Metadata completeness – rate (P)	The ratio of the number of metadata elements provided to the total number of metadata elements applicable.	*For producer reports only* Provide AC3: metadata completeness rate, noting that the “metadata elements” are the SIMS concepts and sub-concepts.

SIMS	Concept name	Definition	Guidelines
S.10.7	Quality documentation	Documentation on procedures applied for quality management and quality assessment.	List relevant quality related documents, for example, other quality reports, studies. Cross reference to descriptions of quality procedures in other chapters, especially S.13. *European level* Summarise availability of national quality reports.
S.11	Quality management	Systems and frameworks in place within an organisation to manage the quality of statistical products and processes.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.11.1	Quality assurance	All systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the requirements for the statistical output.	Describe the procedures (such as use of a general quality management system based on EFQM or ISO 9000 series) to promote general quality management principles in the organisation. Describe the quality assurance framework used to implement statistical quality principles. Describe the quality assurance procedures specifically applied to the statistical process for which the report is being prepared, for example training courses, benchmarking, assessments, and use of best practices. Include descriptions of all forms of quality assessment procedures (self-assessment, peer review, compliance monitoring, audit) and when they most recently took place. Describe any ongoing or planned improvements in quality assurance procedures.
S.11.2	Quality assessment	Overall assessment of data quality, based on standard quality criteria.	Summarise the results of the most recent quality assessments and cross reference to the chapters in the report where the results are presented in more detail.
S.12	Relevance	The degree to which statistical information meet current and potential needs of the users.	(Information relating to this concept is provided by reporting on its sub-concepts.)

SIMS	Concept name	Definition	Guidelines
S.12.1	User needs	Description of users and their respective needs with respect to the statistical data.	Provide: • a classification of users, also indicating their relative importance; • an indication of the uses for which users want the statistical outputs; • an assessment of the key outputs desired by different categories of users and any shortcomings in outputs for important users; • information on unmet user needs and any plans to satisfy them in the future; and details regarding those quality components which do not meet user requirements.
S.12.2	User satisfaction	Measures to determine user satisfaction.	Describe how, and how often, the views and opinions of the users are collected, for example by user satisfaction surveys or other user consultations. State how often such investigations are conducted and when the most recent took place. Present the key results from the recent investigations. Present view of user satisfaction over time, in the form of a user satisfaction index if available.
S.12.3	Completeness	The extent to which all statistics that are needed are available.	Provide qualitative information on the extent to which content requirements in relevant legislation, regulations and guidelines are met. Where such requirements are not fully met, reasons for this should be provided. Provide information on the extent to which user needs related to content are satisfied. Provide values of indicator R1 Data completeness rate, for each required data item for each relevant regulation/ guideline at producer/user level of detail as appropriate. In the case where the indicator refers to data sent to Eurostat, this indicator can be compiled by Eurostat. *European level* Summarise across countries the extent to which ESS requirements for data items are met
	R1 Data completeness rate (U)	The ratio of the number of data cells provided to the number of data cells required by a regulation/ guideline.	
S.12.3 .1	R1 Data completeness rate (P)		
S.13	Accuracy and reliability	Accuracy of data is the closeness of computations or estimates to the exact or true values that the	(Information relating to accuracy is provided by reporting on S.13 sub-concepts. Information on reliability is reported in S.17 Data Revision).

SIMS	Concept name	Definition	Guidelines
		statistics were intended to measure. Reliability of the data, defined as the closeness of the initial estimated value to the subsequent estimated value.	
S.13.1	Overall accuracy	Assessment of accuracy, linked to a certain data set or domain, which is summarising the various components.	Describe the main sources of random and systematic errors in the statistical outputs and provide a summary assessment of all errors with special focus on the impact on key estimates. The bias assessment can be in quantitative or qualitative terms, or both, and may be expressed as bias risk. It should reflect the producer's best current understanding (sign and order of magnitude) and include actions taken to reduce bias. *European level* Provide a summary picture of accuracy across countries. The emphasis placed on various types of errors should depend upon the error profile of the respective process. For repetitive processes, describe how accuracy is developing over time and what efforts are underway to improve accuracy from an ESS perspective.
S.13.2	Sampling error	That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a subset of the population is enumerated.	State whether sampling error is relevant. If probability sampling is used: • for user reports, provide the range of variation of the A1 indicator among key variables at user report level of detail; • for producer reports, provide the range of variation of the A1 indicator among key variables at producer report level of detail; • indicate the impact of sampling error on the overall accuracy of the results;
	A1. Sampling error indicators (U)	Measures of the random variation of an estimator due to sampling, at a level of detail appropriate for user reports.	• state how the calculation of sampling error is affected by imputation for nonresponse, misclassifications and other sources of uncertainty, such as outlier treatment. If non-probability sampling is used, provide an assessment of representativeness and risk of sampling bias. *European level* If probability sampling is used:
S.13.2.1	A1. Sampling error indicators (P)	Measures of the random variation of an estimator due to	• present sampling errors for key estimates across countries;

SIMS	Concept name	Definition	Guidelines
		sampling, at a level of detail appropriate for producer reports.	• indicate which country to country differences are significant and which are not; • for a repetitive survey, describe at least broadly the trends in sampling error over time provide sampling errors for ESS level estimates.
S.13.3	Non-sampling error	Error in survey estimates which cannot be attributed to sampling fluctuations.	Summarise the most important aspects of coverage, measurement, non-response, processing and model assumption errors. Discuss the corresponding bias risks and actions undertaken to reduce them. *European level* Provide a summary of the above across countries.
	A4. Unit non-response - rate (U)	The ratio of the number of units with no information or not usable information to the total number of in-scope (eligible) units, at a level of detail appropriate for a user report.	For probability and census surveys: report A4: unit non-response rates. For repetitive surveys: describe the trend for A4.
	A5. Item non-response - rate (U)	The ratio of the in-scope (eligible) units that have not responded to a particular item and the in-scope units that are required to respond to that particular item, at a level of detail appropriate for a user report.	Report A5: item non-response rates for key variables
S.13.3.1	Coverage error (P)	Divergence between the survey population and the target population.	Provide information on the frame and its sources. Provide an assessment, whenever possible quantitative, of overcoverage and undercoverage, including an evaluation of the bias risks associated with the latter. Describe actions taken for reduction of undercoverage and associated bias risks. *European level* Provide an overall picture of coverage across countries. This is often best done in the form of tables with the important coverage aspects, country by country.

SIMS	Concept name	Definition	Guidelines
S.13.3 .1.1	A2. Overcoverage – rate (P)	The proportion of units accessible via the frame that do not belong to the target population.	For probability surveys: report A2, Overcoverage – rate (for definition, see Supplementary Document C)
S.13.3 .1.2	A3. Common units – proportion (P)	The proportion of units covered by both the survey and the administrative data in relation to the total number of units in the survey.	For multisource processes where one source is a survey and the other source(s) is (are) administrative: • report A3, Common units – proportion (for definition, see Supplementary Document C)
S.13.3 .2	Measurement error (P)	Measurement errors are errors that occur during data collection and cause recorded values of variables to be different from the true ones	The main sources of measurement error should be reported and assessed. Their description should be accompanied by any available analysis, otherwise by the producer's best knowledge. Where available and relevant describe: • identification and general assessment of the main sources of measurement error; • efforts made in questionnaire design and testing, information on interviewer training and other work on error prevention; • results of assessments based on comparisons with external data, re-interviews or experiments; • results of indirect analysis, for example, of the editing phase; and • actions taken to correct measurement errors. *European level* Where measurement errors are important as a single source of error provide a comparative summary across countries. Otherwise include them within overall accuracy in S.13.1.
S.13.3 . 3	Nonresponse error (P)	Nonresponse errors occur when the survey fails to get a response to one, or possibly all, of the questions	Provide a qualitative assessment of unit nonresponse. Highlight the variables that are most subject to item nonresponse (e.g. associated with sensitive questions). Provide a qualitative assessment of the bias associated with nonresponse. Provide a breakdown of nonrespondents according to cause for nonresponse. Describe efforts to reduce nonresponse during data collection and follow-up. Describe treatment of nonresponse at the estimation stage, including response modelling.

SIMS	Concept name	Definition	Guidelines
			European level Provide a qualitative assessment of unit and item nonresponse across countries.
S.13.3. 3.1	A4. Unit nonresponse - rate (P)	The ratio of the number of units with no information or not usable information to the total number of in-scope (eligible) units, at a level of detail appropriate for a producer report.	Report A4: Unit nonresponse rate overall and at a level of detail appropriate for a producer report. *European level* Unit nonresponse rates across countries
S.13.3. 3.2	A5. Item nonresponse - rate (P)	The ratio of the in-scope (eligible) units which have not responded to a particular item to the in-scope units that are required to respond to that particular item, at a level of detail appropriate for a producer report.	Report A5: Item nonresponse rate for all variables. *European level* Item nonresponse rates across countries
S.13.3. 4	Processing error (P)	The error in final data collection process results arising from the faulty implementation of correctly planned implementation methods.	If processing errors are significant, identify the main issues regarding them. Present an analysis of processing errors, where available, otherwise a qualitative assessment. Report their extent, and impact on the outputs, of the most significant types of error. Include descriptions of linking and coding errors, if applicable. Where mistakes relating to programming or publishing have occurred, corrective measures taken as well as actions for avoiding them in the future should be reported. *European level* Provide a summary across countries of processing errors.
S.13.3. 5	Model assumption error (P)	Error due to domain specific models needed to define the target of estimation.	Describe process specific models, for example, as needed to define the target of estimation itself. Provide an assessment of the validity of each model. (Descriptions of models used in treatment of specific sources of error should be presented in the section dealing with those errors.)

SIMS	Concept name	Definition	Guidelines
			European level Where different models are used across countries, provide a comparative overview and discuss their validity and the likely effects of the differences.

SIMS	Concept name	Definition	Guidelines
S.14	Timeliness and punctuality	(Defined by its sub-concepts)	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.14.1	Timeliness	Length of time between data availability and the event or phenomenon the data describe.	Outline the reasons for the time lag. Outline efforts to reduce time lag in future. *European level* For reports only published at European level do the above. Otherwise summarise the above across countries.
	TP2. Time lag - final results (U)	The number of days (or weeks or months) from the last day of the reference period to the day of publication of final results.	*For user reports only* Explain the meaning of TP2 and provide its values for the most recent cycle, and the average over a past period, say three years at a level of detail appropriate for users.
S.14.1 .1	TP1. Time lag - first results (P)	The number of days (or weeks or months) from the last day of the reference period to the day of publication of first results, at producer report level of detail	*For producer reports only* Explain and provide TP1 values for most recent cycle, and average, and maximum over a past period, say three years.
S.14.1 .2	TP2. Time lag - final results (P)	The number of days (or weeks or months) from the last day of the reference period to the day of publication of complete and final results	*For producer reports only* Explain and provide TP2 values for most recent cycle, and average, and maximum over a past period, say three years at a level of detail appropriate for producers.
S.14.2	Punctuality	Time lag between the actual delivery of the data and the target date when it	Report only for annual or more frequent releases. If a release schedule was made available to users and/or specified in a regulation;

SIMS	Concept name	Definition	Guidelines
		should have been delivered.	• provide TP3 (user formula), i.e., the percentage of releases delivered on time, based on scheduled release dates, over a specified period and/or set of outputs. • in the event of any non-punctual releases, explain the reasons and outline efforts to improve punctuality. In the absence of a release schedule, explain why there is no schedule and indicate what efforts will be made to make one available in the future. *European level* For outputs first published at European level, report as above. For outputs first published at country level, • state the agreed time frame for delivery of national data and the actual delivery dates; summarise punctuality across countries.
	TP3. Punctuality - delivery and publication (U)	The percentage of release delivered on time.	
S.14.2.1	TP3. Punctuality - delivery and publication (P)	The number of days between the delivery/ release date of data and the target date on which they were scheduled for delivery/ release.	Explain the meaning of indicator TP3 with producer report calculation formula, i.e., the time lag between scheduled release date and actual release date. Provide the average value of TP3 for the most recent cycle. In the case where the indicator refers to data tables sent to Eurostat, the value of this indicator can be compiled by Eurostat. For a repeating process, provide the average value of TP3 over a past period, say three years. *European level* For outputs that are first published at European Level, do as above.
S.15	Coherence and comparability	Adequacy of statistics to be reliably combined in different ways and for various uses and the extent to which differences between statistics can be attributed to differences between the true values of the statistical characteristics.	(Information relating to this concept is provided by reporting on its sub-concepts.)

SIMS	Concept name	Definition	Guidelines
S.15.1	Comparability – geographical	The extent to which statistics are comparable between geographical areas.	Describe any problems of comparability between regions of the country. The reasons for the problems should be described and as well an assessment (preferably quantitative) of the possible effect on the output values. Give information on discrepancies from the ESS/ international concepts, definitions, with reference to other chapters for more details. *European level* Focus on factors that affect the comparability between countries. Analyse asymmetries in statistical mirror flows where possible.
S.15.1 .1	CC1. Asymmetry for mirror flows statistics – coefficient	The difference or the absolute difference of inbound and outbound flows between a pair of countries divided by the average of these two values.	For producer reports only: provide measures of asymmetries for key variables.
S.15.2	Comparability – over time	The extent to which statistics are comparable or reconcilable over time.	Provide information on possible limitations in the use of data for comparisons over time. Distinguish three broad possibilities: 4. There have been no changes, in which case this should be reported. 5. There have been some changes but not enough to warrant the designation of a break in series. 6. There have been sufficient changes to warrant the designation of a break in series. Provide values of CC2: Length of comparable time series at appropriate level of detail for user or producer report. The indicator CC2 is defined in Supplementary Document C.
	CC2. Length of comparable time series (U)	The number of reference periods in time series from last break.	
S.15.2 .1	CC2. Length of comparable time series (P)		
S.15.3	Coherence- cross domain	The extent to which statistics are reconcilable with those obtained through other data sources or statistical domains.	An analysis of incoherence should be provided, where this is an issue of importance. Reporting under 15.3 is for coherence problems that are not reported under 15.3.1, 15.3.2 or 15.4
S.15.3 .1	Coherence - subannual and annual statistics (P)	The extent to which statistics of different frequencies are reconcilable.	Coherence between subannual and annual statistical outputs is a natural expectation but the statistical processes producing them are often quite different. Compare subannual and annual estimates and, eventually, describe reasons for

SIMS	Concept name	Definition	Guidelines
			lack of coherence between subannual and annual statistical outputs.
S.15.3 .2	Coherence-National Accounts (P)	The extent to which statistics are reconcilable with National Accounts.	Where relevant, the results of comparisons with the National Account framework and feedback from National Accounts with respect to coherence and accuracy problems should be reported and should be a trigger for further investigation.
S.15.4	Coherence – internal	The extent to which statistics are consistent within a given data set.	Each set of outputs should be internally consistent. If statistical outputs within the data set in question are not consistent, any resulting lack of coherence in the output of the statistical process itself should be stated as well as a brief explanation of the reasons for publishing such results.
S.16	Cost and burden	Cost associated with the collection and production of a statistical product and burden on respondents.	**Cost** Provide annual operational costs of the process, with breakdown by major cost component. Describe recent efforts to improve efficiency and comment on the extent to which information and communication technology is used. *European level* Describe recent initiatives and efforts to improve efficiency at the European level. **Burden** Provide an estimate of the respondent burden imposed by the process. Describe all the means taken to minimise burden. *European level* Describe recent initiatives and efforts to minimise burden at the European level.
S.17	Data revision	Any change in a value of a statistic released to the public.	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.17.1	Data revision – policy	Policy aimed at ensuring the transparency of disseminated data, whereby preliminary data are compiled that are later revised.	Describe the data revision policy applicable to data output from the statistical process being reported. In so far as they are relevant to the process being reported, summarise the general procedures for treatment of planned revisions, benchmark revisions, unplanned revisions, and revisions due to conceptual and/or methodological changes. *European level*

SIMS	Concept name	Definition	Guidelines
			Describe the data revision policy and procedures at European level.
S.17.2	Data revision - practice	Information on the data revision practice	If there are no revisions to report for the statistical process that is the subject of the report, state this and close the reporting of this concept. Report the reasons and schedule for planned revisions (if any). Explain A6 Data revision - average size indicator. Report A6 at a level of detail appropriate for user or producer report. In the absence of data to compile A6 Data revision, make a qualitative assessment of the average size of planned revisions and their direction based on historical data. Describe the main reasons for unplanned revisions that have occurred, and the actions taken to prevent the need for such revisions in the future. *European level* Describe the planned and unplanned revisions at European level. Report A6 aggregated over countries.
	A6. Data revision - average size (U)	The average over a time period of the revisions of a key item, for user report.	
S.17.2.1	A6. Data revision - average size (P)	The average over a time period of the revisions of a key item, for producer report.	

SIMS	Concept name	Definition	Guidelines
S.18	Statistical processing	(Defined by its sub-concepts)	(Information relating to this concept is provided by reporting on its sub-concepts.)
S.18.1	Source data	Characteristics and components of the raw statistical data used for compiling statistical aggregates.	Indicate if the data are based on a survey, administrative data process, multisource process, or macro-aggregates. In the event of multisource or macro-aggregates, describe each data source and indicate how they are combined. For each survey dataset, summarise the sample design, cross referencing the descriptions of the target and survey populations, presented in S.03.6. For each administrative dataset, summarise the source, its primary purpose, and the most important data items acquired. *European level* Provide an overview of the sources used across countries.
S.18.2	Frequency of data collection	Frequency with which the source data are collected.	Indicate the frequency of data collection (e.g. monthly, quarterly, annually, or continuous).
S.18.3	Data collection	Systematic process of	For each survey data source:

SIMS	Concept name	Definition	Guidelines
		gathering data for official statistics.	• describe the method(s) used to gather data from respondents; • annex or hyperlink the questionnaires(s). For each administrative data source: • describe the acquisition process and how it was tested. For all sources: • describe the types of checks applied at the time of data entry. *European level* Provide a summary of the commonalities and differences in the collection methods, questionnaires and checks used in different countries.
S.18.4	Data validation	Process of monitoring the results of data compilation and ensuring the quality of statistical results.	Describe the procedures for checking and validating the source data and how the results are monitored and used. Describe the procedures for validating the aggregate output data (statistics) after compilation, including checking coverage and response rates, and comparing with data for previous cycles and with expectations. List other output datasets to which the data relate and outline the procedures for identifying inconsistencies between the output data and these other datasets. *European level* Provide a summary of the commonalities and differences in the validation methods used by countries.
S.18.5	Data compilation	Operations performed on data to derive new information according to a given set of rules.	Describe the procedures for imputation, the most common reasons for imputation and imputation rates within each of the main strata. Describe the likely impact of imputation. Describe the procedures to derive new variables and to calculate aggregates and complex statistics. Describe the procedures for adjustment for non-response and the corrections to the design weights to account for differences in response rates. Describe the calculation of design weights, including calibration (if used). Describe the procedures for combining input data from different sources.
S.18.5 .1	A7. Imputation – rate (P)	The ratio of the number of replaced values to the total number of	*For producer report only:*

SIMS	Concept name	Definition	Guidelines
		values for a given variable.	Provide values of indicator A7 Imputation – rate. The indicator A7 is defined in Supplementary Document C
S.18.6	Adjustment	The set of procedures employed to modify statistical data to enable it to conform to national or international standards or to address data quality differences when compiling specific data sets.	Summarise seasonal adjustment procedures at a level of detail appropriate for a user report. Outline any other macro-level adjustment procedures applied to compiled estimates that are used to improve conformance with standards and/or to address quality concerns.
S.18.6.1	Seasonal adjustment (P)	The statistical technique used to remove the effects of seasonal calendar influences operating on a series.	*For producer report only* Detail the seasonal adjustment procedures including pre-treatment (macro outlier detection, calendar correction), model selection, adjustment tool; validation procedures and revision process.
S.19	Comment	Supplementary descriptive text which can be attached to data or metadata.	Provide any information • that is pertinent to the report but does not fit under any of the other concepts; or • to repeat key issues; or • to make reference to annexes that might be attached to the report.

C (PART III)

Guidelines for ESS Quality and Performance Indicators

C1 Standard Quality and Performance Indicators

The ESS Guidelines for the Implementation of ESS Quality and Performance Indicators were prepared by the in charge of quality within Directorate D: *Government finance statistics (GFS) and quality* at EUROSTAT. They were reviewed by the Eurostat Expert Group on Quality Indicators in 2010 and then slightly updated by the Task Force on Quality Reporting in 2012-2013.

The indicators are as follows.

R1. Data completeness - rate

A1. Sampling error - indicators

A2. Over-coverage - rate

A3. Common units - proportion

A4. Unit non-response - rate

A5. Item non-response - rate

A6. Data revision - average size

A7. Imputation - rate

TP1. Time lag - first results

TP2. Time lag - final results

TP3. Punctuality - delivery and publication

CC1. Asymmetry for mirror flows statistics - coefficient

CC2. Length of comparable time series

AC1. Data tables – consultations

AC2. Metadata - consultations

AC3. Metadata completeness - rate

Some definitions, and some statements on the applicability of the QPIs were deemed to be incorrect or inappropriate and have been modified for this Handbook.

All QPIs apply to a “statistical process”, which can be a survey, an administrative data process, or a multisource process. Certain QPIs, notably A1, A2, A3, A4, A5 and A7, can be applied separately to each separate source in a multisource process.

.

C2 Guidelines for compilation of indicators

Name:	R1. Data completeness - rate
Definition:	The ratio of the number of data cells (entities to be specified by the Eurostat domain manager) provided to the number of data cells required by Eurostat or relevant. The ratio is computed for a chosen dataset and a given period.
Applicability:	The rate of available data is applicable: • to all statistical processes (including those using administrative data); • to users and producers, with different focus and calculation formulae. Computed only by Eurostat but recommended also for inclusion in national quality reports.
Calculation formulae:	For a specific key variable: For producers: $$R1_{PDR} = \frac{\# A_D^{rqd}}{\# D^{rqd}}$$ D^{rqd} in the denominator is the set of data cells required (i.e. excl. derogations/ confidentiality) and $\# A_D^{rqd}$ in the numerator is the corresponding subset of <u>available/provided</u> data cells. The notation # D means the number of elements in the set D (the cardinality). For users $$R1_U = \frac{\# A_D^{rel}}{\# D^{rel}}$$ D^{rel} in the denominator is the set of relevant data cells (full coverage, i.e. excl. only those entities for which the data wouldn't be relevant like e.g. fishing fleet in Hungary) and A_D^{rel} in the numerator is the corresponding subset of <u>available/ provided</u> data cells. The notation # D means the number of elements in the set D (the cardinality). The main difference between the two formulas lies in the selection of the denominators' datasets. Regarding the first formula, for **producers**, this set comprises the required data cells excluding derogations/confidentiality, since producers are interested in assessing the level of compliance with the requirements. On the other hand, for **users**, the formula gives the rate of provided data cells to the ones that are theoretically relevant, meaning that missing cells due to derogations/ confidentiality or any other reason for missing data are included here, leaving out only those cells for which data wouldn't be relevant like e.g. fishing fleet in Hungary.
Target value:	The target value for this indicator is 1 meaning that 100% of the required or relevant data cells are available.
Aggregation levels and principles:	The calculation is done, for a meaningful choice by the domain manager, at subject matter domain level. Aggregations are recommended at EU level for the user-oriented indicator. The number of data cells provided and the number of data cells required/relevant are aggregated separately, from which a ratio is then computed. The indicator shows to what extent statistics are available compared to what should be available.

Interpretation:	For producers: It can be used to evaluate the degree of compliance by a given Member State for a given dataset and period to be specified by the domain manager. For users: At EU level, it can be used to • identify whether important variables are missing for some individual Member State or alternatively • give users an overall measurement (aggregate across countries and/or key variables) of the availability of statistics.
Specific guidance:	The indicator should be accompanied by information about which variable are missing and the reasons for incompleteness as well as, where relevant, the impact of the missing data on the EU aggregate and plans for improving completeness in the future. Calculation would need intervention by the Eurostat domain manager at the initial stage (to define the key variables and the period to be monitored). Later on, the indicators should be calculated automatically. Both formulas are to be computed per key variable, nevertheless an aggregate for all variables can be calculated. **For producers:** This indicator forms part of Eurostat compliance monitoring, thus for producers it should be computed per Member State. **For users:** If certain relevant variables are not reported, the statistics are incomplete. This can be due to data not being collected or data being of low quality or confidential. For users an aggregate across countries for all the key variables could suffice.
References:	• ISO/IEC FDIS 11179-1 "Information technology – Metadata registries • – Part 1: Framework", March 2004 (according to the SDMX Metadata Common Vocabulary draft February. 2008).

Name:	A1. Sampling error - indicators
Definition:	The sampling error can be expressed: a) in relative terms, in which case the relative standard error or, synonymously, the coefficient of variation (CV) is used. (The standard error of the estimator $\hat{\theta}$ is the square root of its variance, $\sqrt{V(\hat{\theta})}$.) The estimated relative standard error (the estimated CV) is the estimated standard error of the estimator divided by the estimated value of the parameter, see calculation formulae below. b) in terms of confidence intervals, i.e. an interval that includes with a given level of confidence the true value of a parameter θ. The width of the interval is related to the standard error. The estimator should take into account the sampling design and should further integrate the effect on precision of adjustments for non-response, corrections for misclassifications, use of auxiliary information through calibration methods etc.
Applicability:	Sampling errors indicator are applicable: • to statistical processes based on probability samples or other sampling procedures allowing computation of such information. • to users and producers, with different level of details given. *Note:* This QPI should be separately applied to each separate source in a multisource process.
Calculation formulae:	Coefficient of variation: $$CV_e(\hat{\emptyset}) = \frac{\sqrt{\hat{V}(\hat{\emptyset})}}{\hat{\emptyset}}$$ Remark: The subscript "e" stands for estimate. Confidence interval, symmetric: $\left[\hat{\emptyset} - d; \hat{\emptyset} + d\right]$ or $\hat{\emptyset} \pm$ d The length of the interval, which is *2d*, depends on the confidence level (e.g. 95%), the assumptions convening the distribution of the estimator of the parameter, and the sampling error. In many cases d has the form below, where t depends on the distribution and the confidence level. d = t$\sqrt{\hat{V}(\hat{\emptyset})}$ In case of totals, means and ratios, formulas for aggregation of coefficients of variation at EU level can be found in the third reference below. The calculation formulae depend on the sampling design, the estimator, and the method chosen for estimating the variance $V(\hat{\emptyset})$.
Target value:	The smaller the CV, the standard error, and the width of the confidence interval, the more accurate is the estimator. Survey regulations may include specifications for precision thresholds at different population levels.

Name:	**A1. Sampling error - indicators**
Aggregation levels and principles:	The calculation is done for all statistics based on probability surveys or equivalent. Aggregations are possible at Member State and EU levels, depending on estimators and degree of harmonisation. The principle for computing the coefficient of variation of an aggregate depends on the method for aggregation of the estimator belonging to that variable
Interpretation:	The CV is a relative (dimensionless) measure of the precision of a statistical estimator, often expressed as a percentage. More specifically, it has the property of eliminating measurement units from precision measures and one of its roles is to make possible comparisons between precision of estimates of different indicators. However, this property has no value added in case of proportions (which are by definition dimensionless indicators). The CV is also inappropriate for variables that can take on negative values.
Specific guidance:	There are several precision measures which can be used to estimate the random variation of an estimator due to sampling, such as coefficients of variation, standard errors and confidence intervals. The coefficient of variation is suitable for quantitative variables with large positive values. It is not robust for percentages or changes and is not usable for data estimates of negative values, where they may be substituted by absolute measures of precision (standard errors or confidence intervals). The confidence interval is usually the precision measure preferred by data users. It is the clearest way of understanding and interpreting the sampling variability. Provision of confidence intervals is voluntary. The CV has the advantage of being dimensionless. The standard error or a confidence interval is sometimes preferable, as discussed.
Reference:	Variance estimation methods in the European Union, Monographs of official Statistics, 2002 edition.

Name	A2. Over-coverage - rate
Definition:	The rate of over-coverage is defined for probability surveys and is the proportion of units accessible via the frame that do not belong to the target population (are out-of-scope). The *target population* is the population for which inferences are made. The *survey frame* (or frames) is a device that permits access to population units. The *survey population* is the set of population units which can be accessed through the frame. The concept of a frame is mainly used for sample surveys but is also applicable for censuses and multisource processes involving probability surveys. Coverage deficiencies may be due to delays in reporting (typical for business statistics) and to errors in unit identification, classification, coding etc. The rate may be calculated either as un-weighted or as weighted to refer to the overall level (frame/population rather than sample). Units of unknown eligibility provide an inherent difficulty; see below.
Applicability:	The rate of over-coverage is applicable: • mainly to probability surveys • to producers. If the survey has more than one type of unit, a rate may be calculated for each type. If there is more than one frame or if over-coverage rates vary strongly between sub-populations, rates should be separated. *Note:* This QPI should be separately applied to each separate source in a multisource process.
Calculation formulae:	The over-coverage rate has three main versions written in one and the same formula as the weighted over-coverage rate *OCR* $OCR = \frac{\sum_O w_j + (1-\alpha)\sum_Q w_j}{\sum_O w_j + \sum_E w_j + \sum_Q w_j}$, where O is the set of out-of-scope units (over-coverage, resolved and not belonging to the target population), E is the set of in-scope units (resolved units belonging to the target population; eligible units), Q is the set of units of unknown eligibility. w_j weight of unit *j*, described below, α is the estimated proportion of cases of unknown eligibility that are actually eligible. It should be set to 1 unless there is strong evidence at country level for assuming otherwise. The three main cases are: Un-weighted rate: $w_j = 1$ Design-weighted rate: $w_j = d_j$ where basically $d_j = 1/\pi_j$, meaning that the design weight is the inverse of the selection probability. Size-weighted rate: $w_j = d_j x_j$ where x_j is the value of a variable x for unit j.
Calculation Formulae (continued):	The variable X, which is chosen subjectively, shows the size or importance of the units. The value should be known for all units. X is auxiliary information, often available in the frame. Examples are turnover for businesses and population for municipalities. For the over-coverage rate the un-weighted and the design-weighted alternatives are the ones mostly used, see Interpretation below. The weight d_j is a “raising” factor when unit *j* represents more than itself, else it is equal

Name	**A2. Over-coverage - rate**
	to one.
Target value:	The target value of this indicator is as much as possible close to 0.
Aggregation levels and principles:	National Level: the indicator is to be calculated for survey populations where meaningful, e.g. over industries. Then separate survey populations are treated as one survey population. ESS Level: the indicator can be aggregated across countries only where statistical production processes are fully harmonised. For the statistical processes involved, the separate survey populations are treated as one survey population. Where production processes differ across countries, lower and higher over-coverage rates can be shown to indicate the range.
Interpretation:	*Over-coverage*: there are units accessible via the frame, which do not belong to the target population (e.g., deceased persons still listed in a Population Register or no longer operating enterprises still in the Business Register). The un-weighted over-coverage rate gives the number of units that have been found not belonging to the target in proportion to the total number of observed units. The number refers to the sample, the census or the register population studied. The design-weighted over-coverage rate is an estimate for the survey population in comparison with the target population, based on the information at hand, usually a sample. The size-weighted over-coverage rate expresses the rate in terms of a chosen size variable, e.g. turnover in business statistics. (This case is less interesting for over-coverage than for non-response.)
Specific guidance:	-

Name:	**A3. Common units - proportion**
Definition:	The proportion of units covered by both the survey and data from an administrative source or sources in relation to the total number of units in the survey.
Applicability:	The proportion is applicable • to mixed statistical processes where some variables or data for some units come from survey data and others from administrative source(s) • to producers. *Note:* This QPI should be separately applied to each separate source in a multisource process.
Calculation formulae:	$$Ad = \frac{\text{No. of common units across survey data and admin. sources}}{\text{No. of unique units in survey data}}$$
Target value:	-
Aggregation levels and principles:	-
Interpretation:	The indicator is used when administrative data is combined with survey data in such a way that data on unit level are obtained from both the survey and one or more administrative sources (some variables come from the survey and other variables from the administrative data) or when data for part of the units come from survey data and for another part of the units from one or more administrative sources. The indicator provides an idea of completeness/coverage of the sources – to what extent units exist in both administrative data and survey data. This indicator does not apply if administrative data is used only to produce estimates without being combined with survey data.
Specific guidance:	Common units refer to those units that are included in the administrative data and the survey data. For the purpose of this indicator, the “unique units in survey data” in the denominator means that if a unit exists in more than one source it should only be counted once. If the survey is conducted for only a subset of the units in the administrative data (e.g. for larger enterprises only), this indicator should be calculated for the relevant subset. Linking errors should be detected and resolved before this indicator is calculated. If there are few common units due to the design of the statistical output (e.g. a combination of survey and administrative data), this should be explained.
References:	ESSNet use of administrative and accounts data in business statistics, WP6 Quality Indicators when using Administrative Data in Statistical Operations, November 2010.

Name:	**A4. Unit non-response – rate**
Definition:	The ratio of the number of units with no information or not usable information (non-response, etc.) to the total number of in-scope (eligible) units. The ratio can be weighted or un-weighted.
Applicability:	The unit non-response rate is applicable: • mainly to probability sampling • to users and producers, with different level of details given. *Note:* This QPI should be separately applied to each separate source in a multisource process.
Calculation formulae:	The non-response rate has three main versions written in one and the same formula as the weighted unit non-response rate *NRR* $NRR = 1 - \frac{\sum_R w_j}{\sum_R w_j + \sum_{NR} w_j + \alpha \sum_Q w_j}$, where R is the set of responding eligible units, NR is the set of non-responding eligible units Q is the set of selected units with unknown eligibility (un-resolved selected units), w_j is the weight of unit *j*, described below α is the estimated proportion of cases of unknown eligibility that are actually eligible. It should be set equal 1 unless there is strong evidence at country level for assuming otherwise. The three main cases are: The three main cases are: Un-weighted rate: $w_j = 1$ Design-weighted rate: $w_j = d_j$ where basically $d_j = 1/\pi_j$, meaning that the design weight is the inverse of the selection probability. Size-weighted rate: $w_j = d_j x_j$ where x_j is the value of a variable X for unit j. The variable X, which is chosen subjectively, shows the size or importance of the units. The value should be known for all units. X is auxiliary information, often available in the frame. Examples are turnover for businesses and population for municipalities. For the unit non-response rate all three alternatives are frequently used, see Interpretation below. The weight d_j is a “raising” factor when unit *j* represents more than itself, else it is equal to one.

Target value:	The target value for this indicator is as close to 0 as possible.
Aggregation levels and principles:	National Level: the indicator is to be calculated at statistical process level ESS Level: rather than aggregating this indicator over countries or to calculate a mean, lower and higher unit non-response rates can be shown by Eurostat for a given variable at statistical process level.
Interpretation:	Unit non-response occurs when no data about an eligible unit are recorded (or data are so few or so low in quality that they are deleted). The un-weighted unit non-response rate shows the result of the data collection in the sample (the units included), rather than an indirect measure of the potential bias associated with non-response. If α=1, it assumes that all the units with unknown eligibility are eligible, so it provides a conservative estimate of A4 with regard to other choices of α. The design-weighted unit non-response rate shows how well the data collection worked considering the population of interest. The size-weighted unit non-response rate would represent an indirect indicator of potential bias caused by non-response prior to any calibration adjustments. Note overall that the bias may be low even if the non-response rate is high, depending on the pattern of the non-responses and the possibilities to adjust successfully for non-response.
Specific guidance:	Non-response is a source of errors in survey statistics mainly for two reasons: • it reduces the number of responses and therefore the precision of the estimates (this may be particularly relevant when samples are used); • it might introduce bias. The size of bias depends on the non-response rate but also on the differences between the respondents and the non-respondents with respect to the variable of interest; furthermore on the strength of auxiliary information.
References:	U.S. Census Bureau Statistical Quality Standards, Reissued 2010. Trépanier, Julien, and Kovar. “Reporting Response Rates when Survey and Administrative Data are Combined.” *Proceedings of the Federal Committee on Statistical Methodology Research Conference 2005.*

Name:	A5. Item non-response - rate
Definition:	The item non-response rate for a given variable is defined as the (weighted) ratio between in-scope units that have not responded and in-scope units that are required to respond to the particular item.
Applicability	The item non-response rate is applicable: • mainly to probability samples (in other statistical processes the term missing data is more often used and could have different definitions); • to users and producers, for selected key variables or for variables with very high item non-response rates, and with different level of details given. If the survey has more than one unit type or data source, a rate may be calculated for each type or data source. If there is more than one frame, or if rates vary strongly between sub-populations, rates should (also) be calculated for separate sub-populations (or strata, groups). *Note:* This QPI should be separately applied to each separate source in a multisource process.
Calculation formulae:	The item non-response rate has three main versions written in one and the same formula as the weighted item non-response rate NRR_Y $NRR_Y = 1 - \frac{\sum_{R_Y} w_j}{\sum_{R_Y} w_j + \sum_{NR_Y} w_j}$, where R_Y is the set of responding eligible units, NR_Y is the set of non-responding eligible units w_j is the weight of unit j, described below The three main cases are: The three main cases are: Un-weighted rate: $w_j = 1$ Design-weighted rate: $w_j = d_j$ where basically $d_j = 1/\pi_j$, meaning that the design weight is the inverse of the selection probability. Size-weighted rate: $w_j = d_j x_j$ where x_j is the value of a variable X for unit j. The variable X, which is chosen subjectively, shows the size or importance of the units. The value should be known for all units. X is auxiliary information, often available in the frame. Examples are turnover for businesses and population for municipalities. The design weight may in the computation of final estimates be modified to correct for non-response, under-coverage etc. This design weight should be used if the rates are to apply to final estimates. The design-weighted rate is mainly used for samples surveys. The weight d j is a “raising” factor when unit j represents more than itself. Otherwise d_j is equal to one.
Target value:	The target value for this indicator is as close to 0 as possible.

Name:	**A5. Item non-response - rate**
Aggregation levels and principles:	National Level: the indicator is to be calculated at statistical process level for key variables and variables with low rates. ESS Level: rather than to aggregate this indicator over countries or to calculate a mean, lower and higher item non-response rates can be shown by Eurostat for a given variable at statistical process level.
Interpretation:	A high item non-response rate indicates difficulties in providing information, e.g. a sensitive question or unclear wording for social statistics or information not available in the accounting system for business statistics. The indicator is a proxy indicator of the possible bias caused by item non-response. In spite of the low item response rate, the bias may still be low, depending on causes, response pattern, and auxiliary information to adjust/impute.
Specific guidance	The un-weighted item non-response rate should be calculated before the data editing and imputation in order to measure the impact of item non-response for the key variables.
References	U.S. Census Bureau Statistical Quality Standards, Reissued 2010. Trépanier, Julien, and Kovar. “Reporting Response Rates when Survey and Administrative Data are Combined.” *Proceedings of the Federal Committee on Statistical Methodology Research Conference 2005.*

<table>
<tr><th>Name:</th><th>A6. Data revision - average size</th></tr>
<tr><td>Definition:</td><td>
The average over a time period of the revisions of a key indicator.
The “revision” is defined as the difference between a later and an earlier estimate of the key item.
The number of releases (K) of a key item (number of times it is published) is fixed and specified in the revision policy. Usually, revisions involve a time series: when publishing an estimate of the key indicator referring to time t, it is a common practice to release the revised version of the indicator referring to a set of previous periods.
In the following table this situation is illustrated for a revision analysis where the policy has K revisions and n reference periods are included in the analysis.
<table>
<tr><th></th><th colspan="9">Reference periods</th></tr>
<tr><td>Releases</td><td>1</td><td>…</td><td>t</td><td>…</td><td>n</td><td></td><td></td><td></td><td></td></tr>
<tr><td>1st release</td><td>X_{11}</td><td>…</td><td>X_{1t}</td><td>…</td><td>X_{1n}</td><td></td><td></td><td></td><td></td></tr>
<tr><td>…</td><td>…</td><td>…</td><td>…</td><td>…</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>kth release</td><td>X_{k1}</td><td>…</td><td>X_{kt}</td><td>…</td><td>X_{kn}</td><td>…</td><td>…</td><td>…</td><td>…</td></tr>
<tr><td>Kth and final release</td><td>X_{K1}</td><td>…</td><td>X_{Kt}</td><td>…</td><td>X_{Kn}</td><td></td><td></td><td></td><td></td></tr>
</table>
Different indicators can be derived by different ways of averaging the revisions for a time series (revisions can be averaged in absolute value or not, the indicator can be absolute or relative).
</td></tr>
<tr><td>Applicability:</td><td>
The average size of revisions is applicable:

to statistical processes where initial and subsequent (revised) estimates are published according to a revision policy (quarterly national accounts, short term statistics);
to users and producers, with different level of details given.

</td></tr>
<tr><td>Calculation formulae:</td><td>
With reference to the two-dimensional situation described in the definition there are several strategies to compute indicators: with or without sign, absolute or relative values, for specific pairs of revisions over time or over a sequence of revisions etc. The main suggestion here is to consider an average for a given revision step over a set of n successive reference/publication periods.
MAR (Mean Absolute Revision)
$MAR = \frac{1}{n}\sum_{t=1}^{n}|X_{Lt} - X_{Pt}|$, where
X_{Lt} is a later estimate for reference period t
X_{Pt} is an earlier estimate for reference period t;
n = No. of estimates (reference periods) in the time series taken into account. n ≥ 20 is recommended for quarterly estimates while n ≥ 30 is recommended for monthly estimates. The indicator is not recommended for annual estimates.
MAR provides an idea of the average size of a given revision step.
This indicator can alternatively be expressed in relative terms:
RMAR: Relative Mean Absolute Revision
</td></tr>
</table>

Name:	A6. Data revision - average size
	$RMAR = \frac{\sum_{t=1}^{n}\lvert X_{Lt}-X_{Pt}\rvert}{\sum_{t=1}^{n}\lvert X_{Lt}\rvert}$, In addition – at the level of Eurostat – and where the sign is interesting, there is the mean revision from Release *P* to Release *L* over the *n* reference periods: **MR (Mean Revision):** $MR = \frac{1}{n}\sum_{t=1}^{n}\lvert X_{Lt} - X_{Pt}\rvert$, Different combinations of *P* and *L* can be considered. For instance, OECD suggests to compare the following releases: Monthly data Quarterly data Release L Release P Release L Release P After 2 Months First After 5 Months First After 3 Months First After 1 Year After 5 Months After 3 Months After 2 Months After 1 Year First After 1 Year First After 2 Years First After 2 Years First Latest available First Latest available First After 2 Years After 1 Year After 2 Years After 1 Year
Target value:	
Aggregation levels and principles:	National Level: the indicator is to be calculated at statistical process level. ESS Level: the indicator is calculated on the revisions made on the EU aggregate/indicator.
Interpretation:	**MAR** provides an idea of the average size of a given revision step for a key item step over the time. The **RMAR** indicator normalises the MAR measure using the final estimates. It facilitates international comparisons and comparisons over time periods. When estimating growth rates this measure corrects the MAR for the size of growth and, so, takes account of the fact that revisions might be expected to be larger in periods of high growth than in periods of slow growth. Both MAR and RMAR indicators provide information on the stability of the estimates. They do not provide information on the direction of revisions, since the absolute values of revisions are considered. Such information is provided by **MR**. A positive sign means upwards revision (underestimation), and a negative sign indicates overestimation in the first case. MR sometimes is referred to as ‘average bias’, but a nonzero MR is not sufficient to establish whether the size of revisions is systematically biased in a given direction. To ascertain the presence of bias it has to be assessed whether MR is statistically different from zero (given no changes in definitions, methodologies, etc.).
Specific guidance:	Either MAR or RMAR should be presented under this indicator. In addition MR could also be calculated at EU-level.

Name:	**A7. Imputation - rate**
Definition:	Imputation is the process used to assign replacement values for missing, invalid or inconsistent data that have failed edits. This includes automatic and manual imputations; it excludes follow-up with respondents and the corresponding corrections (if applicable). Thus, imputation as defined above occurs after data collection, no matter from which source or mix of sources the data have been obtained, including administrative data. After imputation, the data file should normally only contain plausible and internally consistent data records. This indicator is influenced both by the item non-response and the editing process. It measures both the relative amount of imputed values and the relative influence on the final estimates from the imputation procedures. The unweighted imputation rate for a variable is the ratio of the number of imputed values to the total number of values requested for the variable. The weighted rate shows the relative contribution to a statistic from imputed values; typically a total for a quantitative variable. For a qualitative variable, the relative contribution is based on the number of units with an imputed value for the qualitative item.
Applicability	The imputation rate is applicable: • to all statistical processes with micro data; • to producers. *Note:* This QPI should be separately applied to each separate source in a multisource process.
Calculation formulae:	1. Unweighted on the statistical process and variable level: $$A7_{uw} = \frac{n_{AV}}{n_{AV} + n_{OV}}$$ n_{AV} and n_{OV} are the numbers of assigned values and observed values, respectively. 2. The contribution of imputed values is calculated in an analogous way, but weighted and with variable values. $$A7_{w} = \frac{\sum_{AV} w_j\, y_j}{\sum_{AV} w_j\, y_j + \sum_{OV} w_j\, y_j}$$ Here, *AV* and *OV* are the sets of units with assigned and observed values, respectively. In addition, w_j is the weight (normally the weight used for estimation takes into account the sample design as well as adjustment for unit nonresponse and final calibration) of unit j with value y_j. In case of a qualitative variable, y_j=1 if the *j*th unit has a given characteristic and 0 otherwise. When imputation is counted the following changes have to be considered i. imputation of a (non-blank) value for a missing item ii. imputation of a (non-blank) value to correct an observed invalid (non-blank) value iii. imputation of a blank value to correct an undue invalid (non- blank) response.:

Name:	A7. Imputation - rate
	The two main cases for the imputation rate are: Design-weighted rate: $w_j = d_j$ where basically $d_j = 1/\pi_j$, meaning that the design weight is the inverse of the selection probability. Size-weighted rate: $w_j = d_j x_j$ where x_j is the value of a variable x for unit j.
Target value:	A value equal or close to zero is desirable; imputation indicates missing and invalid values.
Aggregation levels and principles:	National Level: The calculation is done for key variables at statistical process level. ESS Level: Aggregations can be made at the level of EU on the basis of harmonised statistical production processes across Member States, considering this as a single statistical process. Alternatively, Eurostat can report lower and higher imputation rates for a given variable at statistical process level.
Interpretation:	The unweighted rate shows, for a particular variable, the proportion of units for which a value has been imputed due to the original value being a missing, implausible, or inconsistent value in comparison with the number of units with a value for this variable. Units with imputation of a blank value to correct an undue invalid (non-blank) response (type iii) have to be included in both numerator and denominator. The weighted rate shows, for a particular variable, the relative contribution of imputed values to the estimate of this item/variable. Obviously, this weighted indicator is meaningful when the objective of a survey is that of estimating the total amount or the average of a variable. When the objective of the estimation is that of estimating complex indices, the weighted indicator is not meaningful.
Specific guidance:	-
References:	Statistics Canada Quality Guidelines, Fifth Edition – October 2009

Name:	**TP1. Time lag - first results**
Definition:	*General definition:* The timeliness of statistical outputs is the length of time between the end of the event or phenomenon they describe and their availability. *Specific definition:* The number of days (or weeks or months) from the last day of the reference period to the day of publication of first results.
Applicability:	This indicator is applicable: • to all statistical processes with **preliminary data releases**; • to producers. TP1 is **not** applicable for statistical processes with only one, directly final, set of results/statistics – then only TP2 is used.
Calculation formulae:	$TP_1 = d_{frst} - d_{refp}$, where d_{frst} is the release date of first results; d_{refp} is the last day (date) of the reference period of the statistics *Measurement units*: datum format (calendar days; if the number of days is large, it may be converted into weeks or months) Instead of a period, the reference can also be a time point.
Target value:	The target values usually are fixed by legislation or gentlemen's agreement. Nevertheless, smaller values denote higher timeliness.
Aggregation levels and principles:	The calculation is done, for a meaningful choice, at subject matter domain level. It could refer to the current production round or be an average over a time period. Aggregations are possible at EU and domain (e.g. social statistics, business statistics) level.
Interpretation:	This indicator quantifies the gap between the release date of first results and the date of reference for the data. Comparisons could be made among statistical processes with the same periodicity.
Specific guidance	The reasons for possible long production times should be explained and efforts to improve the situation should be described. For annual statistics or where timeliness is measured in years rather than in days a sentence stating timeliness would be sufficient.

Name:	TP2. Time lag - final results
Definition:	*General definition:* The timeliness of statistical outputs is the length of time between the end of the event or phenomenon they describe and their availability. *Specific definition:* The number of days (or weeks or months) from the last day of the reference period to the day of publication of complete and final results.
Applicability:	This indicator is applicable: • to all statistical processes; • to users and producers, with different level of details given.
Calculation formulae:	$TP_2 = d_{finl} - d_{refp}$, where d_{finl} is the release date of final results; d_{refp} is the last day (date) of the reference period of the statistics *Measurement units*: datum format (calendar days; if the number of days is large, it may be converted into weeks or months). Instead of a period, the reference can also be a time point.
Target value:	The target values usually are fixed by legislation or gentlemen's agreement. Nevertheless, smaller values denote higher timeliness.
Aggregation levels and principles:	The calculation is done, for a meaningful choice, at subject matter domain level. It could refer to the current production round or be an average over a time period. Aggregations are possible at EU and domain (e.g. social statistics, business statistics) level.
Interpretation:	This indicator quantifies the gap between the release date of the final results and the end of the reference period. Comparisons could be made among statistical processes with the same periodicity.
Specific guidance	The reasons for possible long production times should be explained and efforts to improve the situation should be described. To be further defined by subject matter domain, taking the revisions' policy into account, what could be considered by "final results". For annual statistics or where timeliness is measured in years rather than in days a sentence stating timeliness would be sufficient.

Name:	**TP3. Punctuality - delivery and publication**
Definition:	Punctuality is the time lag between the delivery/release date of data and the target date for delivery/release as agreed for delivery or announced in an official release calendar, laid down by Regulations or previously agreed among partners.
Applicability:	The punctuality of publication is applicable: • to all statistical processes with fixed/pre-announced release dates, • to users and producers, with different aspects and calculation formulae. Computed only by Eurostat but recommended also for inclusion in national quality reports.
Calculation formulae:	**For producers:** Punctuality of data delivery TP3 $TP_3 = d_{act} - d_{sch}$ where d_{act} is the actual date of the effective provision of the statistics, d_{sch} is the scheduled date of the effective provision of the statistics *Measurement units*: datum format (calendar days) **For users:** Rate of punctuality of data publication $TP3_R$ Relevant for a group of statistics/results $TP3_R$ is the proportion of datasets that have met the release calendar date within a group of datasets. $TP3_R = \frac{m_{pc}}{m_{pc}+m_{up}}$ where m_{pc} is the number of statistics/results that have been published on the date announced in the calendar or have been released earlier (punctual). m_{up} is the number of statistics/results that have not met the date announced in the calendar (unpunctual).
Target value:	The target value for TP3 is 0 meaning that there is no delay on the delivery/transmission of data. For $TP3_R$ the target value is 1 meaning that 100% of the items were published on the pre-fixed calendar date.
Aggregation levels and principles:	There are two aspects: - National data deliveries to Eurostat (producer-oriented), - Publication/release by Eurostat (user oriented), The calculation is done at statistical process level. Aggregations are to be made at EU-level over countries and over domains.
Interpretation:	The indicator **Punctuality of data delivery** quantifies the difference (time lag) between actual and target date. This should be interpreted according to the periodicity of the statistical process. The indicator **Rate of punctuality** of release ($TP3_R$) evaluates the punctuality of release of a group of particular datasets.

Name:	**TP3. Punctuality - delivery and publication**
Specific guidance	For producers: For compliance monitoring purposes Eurostat domain managers should monitor this indicator for individual countries. This information can be pre-filled by Eurostat as it is known when data are received from countries. Formula TP3 should be applied in this case. This indicator can be presented in table format for the different countries. The reasons for late or non-punctual delivery should be stated along with their effect on the statistical product, meaning that because of late data deliveries the quality assurance procedures for the whole product/series might not be completed. For users: Enough to compile this indicator as an aggregate at ESTAT level. Formula $TP3_R$ should be applied in this case. Some explanations should be given to users concerning non-punctual publication.

Name:	CC1. Asymmetry for mirror flows statistics - coefficient
Definition:	General definition: Discrepancies between data related to flows, e.g. for pairs of countries. Specific definition (a few versions are provided) Bilateral mirror statistics: The difference or the absolute difference of inbound and outbound flows between a pair of countries divided by the average of these two values. Comment Outbound and inbound flows should be considered to be any kind of flows specific to each subject matter domain (amounts of products traded, number of people visiting a country for tourism purposes, etc.)
Applicability:	The asymmetries for statistics mirror flows is applicable: - to domains in which mirror statistics (flows concerning trade, migration, tourism statistics, FATS, balance of payment etc) are available - to producers. Computed by Eurostat (pre-filled in quality report)
Calculation formulae	Bilateral mirror statistics: For each pair of countries, suppose: A – Country A B – Country B $CC1A_B = \frac{OF_{AB} - mIF_{AB}}{(OF_{AB} + mIF_{AB})/2}$ $CC1B_A = \frac{OF_{BA} - mIF_{BA}}{(OF_{BA} + mIF_{BA})/2}$ A joint measure can be obtained from the two differences in relation to an average flow (several possibilities, one is given below): $CC1_{AB} = \frac{\lvert OF_{AB} - mIF_{AB}\rvert + \lvert OF_{BA} - mIF_{BA}\rvert}{(OF_{AB} + mIF_{AB})/2 + (OF_{BA} + mIF_{BA})/2}$ OF_{AB} - outbound flow going from country A to country B, as published by country A mIF_{AB} – mirror inbound flow, as published by country B Definitions of variables with suffix BA follow accordingly. Multilateral mirror statistics: It is possible to define multilateral mirror statistics according to the same logic as above. There are several possibilities, however, and it is not meaningful to provide a particular one here.
Target value	The value of this indicator should be as close to zero as possible, since – at least in theory – the value of inbound and outbound flows between pairs of countries should match.
Aggregation levels and principles:	National Level: The calculation is done for key variables/sub-series to be selected by the Eurostat domain manager. ESS Level: Aggregations are possible at EU-level (see multilateral mirror statistics formulae). Alternatively, where e.g. not all information is available, lower and higher values of bilateral mirror statistics can be reported to indicate the range.

Name:	CC1. Asymmetry for mirror flows statistics - coefficient
Interpretation:	In domains where mirror statistics are available it is possible to assess geographical comparability measuring the discrepancies between inbound and outbound flows for pairs of countries. Mirror data can help checking the consistency of data reporting, of data, of the reporting process and the definitions used. Finally, they can help to estimate missing data. For the users the asymmetries indicators provide some indication of overall data credibility. There is perfect symmetry (outbound flows are equal to mirror inbound flows) when the coefficient is equal to zero. The more the coefficient diverges from zero, the more the asymmetry between outbound flows and mirror inbound flows becomes important.
Specific guidance:	$CC1A_B$ and $CC1B_A$ indicators can be negative or positive. Indicator $CC1_{AB}$ is always non-negative. Outbound flows from Member State A to Member State B, as reported by A, should be almost equal to inbound flows into B coming from A, as reported by B. Because some domains use a different valuation principle, inbound flows can be slightly different from outbound flows. Therefore, comparisons dealing with mirror statistics have to be made cautiously and should take into account the existence of these discrepancies. The asymmetry coefficient $CC1_{AB}$ is useful because it can be monitored over time. Indicators $CC1A_B$ and $CC1B_A$ can be either positive or negative and can be used to estimate if a country is globally declaring higher or lower level of flows compared with the mirror flows declared by its partner countries. Indicators $CC1A_B$ and $CC1B_A$ should be presented in a table (example statistics for international trade in goods).
References:	International trade in services statistics - Monitoring progress on implementation of the Manual and assessing data quality – OECD Eurostat Trade in services experts meeting 2005.

Name:	CC2. Length of comparable time series
Definition:	Number of reference periods in time series from last break. *Comment* Breaks in statistical time series may occur when there is a change in the definition of the parameter to be estimated (e.g. variable or population) or the methodology used for the estimation. Sometimes a break can be prevented, e.g. by linking.
Applicability:	The length of comparable time series is applicable: • to all statistical processes producing time-series; • to users and producers, with different level of details given. Computed only by Eurostat but recommended also for inclusion in national quality reports.
Calculation formula:	The reference periods are numbered. $CC2 = J_{last} - J_{first} + 1$ J_{last} is the number of the last reference period with disseminated statistics. J_{first} is the number of the first reference period with comparable statistics.
Target value:	A long time series may seem desirable, but it may be motivated to make changes, e.g. since reality motivates new concepts or to achieve coherence with other statistics.
Aggregation levels and principles:	The calculation is done at statistical process level. Aggregations are possible at MS, EU, and Domain (e.g. social statistics, business statistics) level. The indicator for the EU or domain level should be calculated by Eurostat considering the time series of the EU aggregate.
Interpretation:	If there has not been any break, the indicator is equal to the number of the time points in the time series.
Specific guidance:	The length of the series with comparable statistics is expressed as the number of time periods (points) in this series. It is counted from the first time period with statistics after the break onwards. The result does not depend on the length of the reference period. Only applicable for the statistical data disseminated in the sequence of regular time periods (points). If more than one series exist for one statistical process the domain manager should select the appropriate ones for calculation.

Name:	AC1. Data tables – consultations ([19])
Definition:	Number of consultations of data tables within a statistical domain for a given time period. By "number of consultations" it is meant number of data tables views, where multiples views in a single session count only once. Some information available through the monthly Monitoring report on Eurostat Electronic Dissemination and its excel files with detailed figures.
Applicability:	The number of consultations of data tables is applicable: • to all statistical processes using on-line data tables for dissemination of statistics; • to producers (Eurostat domain managers). Computed only by Eurostat but recommended also for inclusion in national quality reports.
Calculation formulae:	AC1 = *#CONS* where *#CONS* denotes the absolute number of elements in the set CONS (this is also called cardinality of the set). In this case CONS represents the consultations of a data table for a specific subject-matter domain. The frequency of collection of the figures for this indicator should be monthly. Remark: internal page views will be excluded.
Target value:	There is no immediate interpretation of low and high values of this indicator, and there is no particular target.
Aggregation levels and principles:	The calculation is done at statistical process level. Aggregation is possible at the following level: • Domains specific data tables. • Annual aggregation. The principle is to calculate the number of consultations of data tables by subject matter.
Interpretation:	This indicator should be carefully analysed and combined with other information that will complement the analysis. The indicator contributes to the assessment of users' demand of data (level of interest), for the assessment of the relevance of subject-matter domains. A ratio can be computed to give insight to the proportion of consultation of the data tables in question in comparison to the total number of consultations for all the domains.
Specific guidance:	An informative and straightforward way to represent the output of this indicator is by plotting the figures over time in a graph. In particular, it would be a graph where the horizontal (x) axis would represent months and the vertical (y) axis would represent the number of datasets consulted. It would be possible to monitor the interest of users for each dataset at the domain specific level. A graph of both the number of consultations of data tables and metadata files (AC2) would be interesting to display.

([19]) The indicator must be collected in collaboration with Unit D4 - Dissemination

Name:	AC2. Metadata – consultations [20]
Definition:	Number of metadata consultations within a statistical domain for a given time period. By "number of consultations" it is meant the number of times a published metadata file is viewed. Some information is available through the monthly Monitoring report on Eurostat Electronic Dissemination and its excel files with detailed figures.
Applicability	This indicator is applicable: • to all statistical processes; • to producers (Eurostat domain managers). Computed only by Eurostat.
Calculation formulae:	AC2 = the ESS-MH files consulted for a specific subject-matter domain for a given time period. Remark: internal page views are excluded.
Target value:	There is no immediate interpretation of low and high values of this indicator, and there is no particular target.
Aggregation levels and principles:	The calculation is done at statistical process level. Aggregation is possible at the following levels: • Domains specific ESS-MH files. • Annual aggregation. The principle is to calculate the number of consultations of ESS-MH files by subject matter domains.
Interpretation:	The indicator contributes to the assessment of users' demand of metadata (level of interest), for the assessment of the relevance of subject-matter domains. A ratio can be computed to give insight to the proportion of consultation of the ESS-MH files in question as a proportion of the total number of consultations for all the domains.
Specific guidance	An informative and straightforward way to represent the output of this indicator is by plotting the figures over time in a graph. In particular, it would be a graph where the horizontal (x) axis would represent months and the vertical (y) axis would represent the number of ESS-MH files consulted. It would be possible to monitor the interest of users for each ESS-MH file at the domain specific level. A graph of both the number of consultations of data tables (indicator AC1) and metadata files with a correspondence would be interesting to display, over time.

(20) The indicator must be collected in collaboration with Unit D4 - Dissemination

Name:	AC3. Metadata completeness - rate
Definition:	The ratio of the number of metadata elements provided to the total number of metadata elements applicable.
Applicability:	The rate of completeness of metadata is applicable: • to all statistical processes; • to producers (Eurostat domain managers). Computed only by Eurostat but recommended also for inclusion in national quality reports.
Calculation formulae:	$AC3_C = \frac{\#M_L}{\#L}$ L in the denominator is the set of applicable metadata elements under consideration and M_L in the numerator is the subset of L of available metadata elements. The notation $\# L$ means the number of elements in the set L (the cardinality). Letter C in the left-hand side of the formula stands for both EU and EFTA countries. The set L is obtained by calculation for a group of metadata elements as explained below over a geographical entity (MS or the EU+EFTA), a statistical domain, etc. There are three groups of metadata, described below together with a categorisation using the SIMSV2.0 concepts (only the main concepts are included in the following breakdown). Metadata about statistical outputs - concepts 2, 3, 4, 7.1, 8, 9; Metadata about statistical processes - concepts 5, 6, 7.2, 17,18; Metadata about quality: concepts 10-16 Computations are made separately for each of the three groups at national level and at ESS level.
Target value:	The target value is 1 meaning that 100% of metadata is available from what is required/applicable to the statistical process, or aggregate, in question.
Aggregation levels and principles:	The calculation is done at the level of ESS-MH files. Aggregations are possible at MS, EU, and Domain (e.g. social statistics, business statistics) level. The principle is to calculate the indicators as an un-weighted rate at the level of MS and EU for a statistical domain (social statistics, business statistics etc.).
Interpretation:	Each indicator shows to what extent metadata of a specific type is available compared to what should be available. This indicator should be carefully analysed since this rate only reflects the existing amount of metadata for a certain statistical process but not the quality of that information.

Specific guidance:	All the information is to be retrieved from ESS-MH files. In case the ESS-MH is empty for the different categories specified previously no calculation is needed but a descriptive text should be replaced. It should be taken into account what availability of metadata actually means.
References:	Euro SDMX Metadata Structure, version March 2009.

(PART III)

D Domain specific regulations involving quality reporting

D1 Introductory remarks

A list of ESS regulations that are relevant for quality reporting in the various domains is provided in Quality Requirements/Standards for Quality Reporting (Updated 2018). For ease or reference, the document is reproduced below with some additions and corrections, and with links to regulations.

D2 List of and links to regulations involving quality reporting

Domains and sub-domains	**Quality requirements/standards for Quality reporting**
Agriculture and fisheries	
Economic accounts for agriculture	
Integrated farm statistics	Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018 on integrated farm statistics, *and repealing* Regulations (EC) No 1166/2008 and (EU) No 1337/2011
Vineyard data collection	Regulation (EU) No 1337/2011 of the European Parliament and of the Council of 13 December 2011concerning European statistics on permanent crops and repealing Council Regulation (EEC) No 357/79 and Directive 2001/109/EC of the European Parliament and of the Council (Note that, in accordance with Article 20 of Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018, the above Regulation will be repealed as of 1 January 2022.)
Orchards	Regulation (EU) No 1337/2011 of the European Parliament and of the Council of 13 December 2011concerning European statistics on permanent

Domains and sub-domains	Quality requirements/standards for Quality reporting
	crops and repealing Council Regulation (EEC) No 357/79 and Directive 2001/109/EC of the European Parliament and of the Council (Note that, in accordance with Article 20 of Regulation (EU) 2018/1091 of the European Parliament and of the Council of 18 July 2018, the above Regulation will be repealed as of 1 January 2022.)
Annual crops	Regulation (EC) No **543/2009** of the European Parliament and of the Council of 18 June 2009 concerning crop statistics, *and repealing* Council Regulations (EEC) No 837/90 and (EEC) No 959/93
Pigs, bovine livestock and sheep and goats	Regulation (EC) No **1165/2008** of the European Parliament and of the Council of 19 November 2008 concerning livestock and meat statistics, *and repealing* Council Directives 93/23/EEC, 93/24/EEC and 93/25/EEC
Milk and milk products	
Submission of data on fishery products	Regulation (EC) No **1921/2006** of the European Parliament and of the Council of 18 December 2006 on the submission of statistical data on landings of fishery products in Member States, *and repealing* Council Regulation (EEC) No 1382/91.
Aquaculture	Regulation (EC) No **762/2008** of the European Parliament and of the Council of 9 July 2008 on the submission by Member States of statistics on aquaculture, *and repealing* Council Regulation (EC) No 788/96.
Pesticides	Regulation (EC) No **1185/2009** of the European Parliament and of the Council of 25 November 2009 concerning statistics on pesticides
International trade in goods and services	
Intrastat	Regulation (EC) No **638/2004** of the European Parliament and of the Council of 31 March 2004 on Community statistics relating to the trading of goods between Member States and repealing Council Regulation (EEC) No 3330/91, *as amended by* Regulation (EC) No **222/2009** of the European Parliament and of the Council of 11 March 2009 amending Regulation (EC) No 638/2004 on Community statistics relating to the trading of goods between Member States Commission Regulation (EU) No 1093/2013 of 4 November 2013 amending Regulation (EC) No 638/2004 of the European Parliament and of the Council and Commission Regulation (EC) No 1982/2004 as regards the simplification within the Intrastat system and the collection of Intrastat information Regulation (EU) No 659/2014 of the European Parliament and of the Council of 15 May 2014 amending Regulation (EC) No 638/2004 on Community statistics relating to trading of goods between Member States as regards conferring delegated and implementing powers on the Commission for the adoption of certain measures, the communication of

Domains and sub-domains	Quality requirements/standards for Quality reporting
	information by the customs administration, the exchange of confidential data between Member States and the definition of statistical value *and as implemented by* Commission Regulation (EC) No **1982/2004** of 18 November 2004 implementing Regulation (EC) No 638/2004 of the European Parliament and of the Council on Community statistics relating to the trading of goods between Member States and repealing Commission Regulations (EC) No 1901/2000 and (EEC) No 3590/92 *as amended by*Commission Regulation (EC) No 1915/2005 of 24 November 2005 amending Regulation (EC) No 1982/2004 with regard to the simplification of the recording of the quantity and specifications on particular movements of goods Commission Regulation (EU) No 91/2010 of 2 February 2010 amending Regulation (EC) No 1982/2004 implementing Regulation (EC) No 638/2004 of the European Parliament and of the Council on Community statistics relating to the trading of goods between Member States, as regards the list of goods excluded from statistics, the communication of information by the tax administration and quality assessment Commission Regulation (EU) No 96/2010 of 4 February 2010 amending Regulation (EC) No 1982/2004 implementing Regulation (EC) No 638/2004 of the European Parliament and of the Council on Community statistics relating to the trading of goods between Member States, as regards the simplification threshold, trade by business characteristics, specific goods and movements and nature of transaction codes Commission Regulation (EU) No 1093/2013 of 4 November 2013 amending Regulation (EC) No 638/2004 of the European Parliament and of the Council and Commission Regulation (EC) No 1982/2004 as regards the simplification within the Intrastat system and the collection of Intrastat information
Extrastat	Regulation (EC) No **471/2009** of the European Parliament and of the Council of 6 May 2009 on Community statistics relating to international trade in goods with non-member countries *and repealing* Council Regulation (EC) No 1172/95 *as amended by* Regulation (EU) 2016/1724 of the European Parliament and of the Council of 14 September 2016 amending Regulation (EC) No 471/2009 on Community statistics relating to external trade with non-member countries as regards conferring of delegated and implementing powers upon the Commission for the adoption of certain measures Commission Regulation (EU) 2016/2119 of 2 December 2016 amending Regulation (EC) No 471/2009 of the European Parliament and of the Council and Commission Regulation (EU) No 113/2010 as regards the adaptation of the list of customs procedures and the definition of the data *and as implemented by* • Commission Regulation (EU) No 92/2010 of 2 February 2010 implementing Regulation (EC) No 471/2009 of the European Parliament and of the Council on Community statistics relating to external trade with non-member countries, as regards data exchange between customs authorities and national statistical authorities, compilation of statistics and quality assessment *as amended by* Commission Implementing Regulation (EU) 2016/1253 of 29 July 2016 amending Regulation (EU) No 92/2010 as regards the data

Domains and sub-domains	Quality requirements/standards for Quality reporting
	exchange between customs authorities and national statistical authorities and the compilation of statistics • Commission Regulation (EU) No 113/2010 of 9 February 2010 implementing Regulation (EC) No 471/2009 of the European Parliament and of the Council on Community statistics relating to external trade with non-member countries, as regards trade coverage, definition of the data, compilation of statistics on trade by business characteristics and by invoicing currency, and specific goods or movements *as amended by* Commission Regulation (EU) No 1106/2012 of 27 November 2012 implementing Regulation (EC) No 471/2009 of the European Parliament and of the Council on Community statistics relating to external trade with non-member countries, as regards the update of the nomenclature of countries and territories

Economy and finance

GNI	Regulation (EU) 2019/516 of the European Parliament and of the Council of 19 March 2019 on the harmonisation of gross national income at market prices and repealing Council Directive 89/130/EEC, Euratom and Council Regulation (EC, Euratom) No 1287/2003 (GNI Regulation).
Excessive Deficit Procedure	Council Regulation (EC) No **479/2009** of 25 May 2009 on the application of the Protocol on the excessive deficit procedure annexed to the Treaty establishing the European Community, *as amended by:* Council Regulation (EU) No 679/2010 of 26 July 2010 amending Regulation (EC) No 479/2009 as regards the quality of statistical data in the context of the excessive deficit procedure.
European System of Accounts – ESA95	Regulation (EU) No 549/2013 of the European Parliament and of the Council of 21 May 2013 on the European system of national and regional accounts in the European Union, *as implemented by* Commission Implementing Regulation (EU) 2016/2304 of 19 December 2016 on the modalities, structure, periodicity and assessment indicators of the quality reports on data transmitted pursuant to Regulation (EU) No 549/2013 of the European Parliament and of the Council.
Prices	Regulation (EU) 2016/792 of the European Parliament and of the Council of 11 May 2016 on harmonised indices of consumer prices and the house price index, and repealing Council Regulation (EC) No 2494/95.
Balance of Payments (BOP)	Regulation (EC) No **184/2005** of the European Parliament and of the *Council* of 12 January 2005 on Community statistics concerning balance of payments, international trade in services and foreign direct investment, *as amended by* Regulation (EU) 2016/1013 of the European Parliament and of the Council of 8 June 2016 amending Regulation (EC) No 184/2005 on Community statistics concerning balance of payments, international trade in services and foreign direct investment *as implemented by*

Domains and sub-domains	Quality requirements/standards for Quality reporting
	Commission Regulation (EC) No **1055/2008** of 27 October 2008 implementing Regulation (EC) No 184/2005 of the European Parliament and of the Council, as regards quality criteria and quality reporting for balance of payments statistics, *as amended by* Commission Regulation (EU) No 1227/2010 of 20 December 2010 amending Regulation (EC) No 1055/2008 implementing Regulation (EC) No 184/2005 of the European Parliament and of the Council, as regards quality criteria and quality reporting for balance of payments statistics.
Purchasing Power Parities (PPP)	Regulation (EC) No **1445/2007** of the European Parliament and of the Council of 11 December 2007 establishing common rule for the provision of basic information on Purchasing Power Parities and for their calculation and dissemination, *as implemented and adapted by* Commission Regulation (EU) No 193/2011 of 28 February 2011 implementing Regulation (EC) No 1445/2007 of the European Parliament and of the Council as regards the system of quality control used for Purchasing Power Parities.
Energy and environment	
Energy	Regulation (EU) 2016/1952 of the European Parliament and of the Council of 26 October 2016 on European statistics on natural gas and electricity prices and repealing Directive 2008/92/EC, *as implemented by:* Regulation (EC) No **1099/2008** of the European Parliament and of the Council of 22 October 2008 on energy statistics.
Environment	Regulation (EC) No **2150/2002** of the European Parliament and of the Council of 25 November 2002 on waste statistics *as amended by:* Regulation (EC) No **221/2009** of the European Parliament and of the Council of 11 March 2009 amending Regulation (EC) No 2150/2002 on waste statistics, as regards the implementing powers conferred on the Commission; and Commission Regulation (EU) No 849/2010 of 27 September 2010 amending Regulation (EC) No 2150/2002 of the European Parliament and of the Council on waste statistics; *and as implemented by:* (Commission Regulation (EC) No **783/2005** of 24 May 2005 setting out the format for the transmission of results on waste statistics); Commission Regulation (EC) No **1445/2005** of 5 September 2005 defining the proper quality evaluation criteria and the contents of the quality reports for waste statistics for the purposes of Regulation (EC) No 2150/2002 of the European Parliament and of the Council.
Industry, services and tourism	
Structural Business Statistics (SBS)	Commission Regulation (EC) No **1618/1999** of 23 July 1999 concerning the criteria for the evaluation of quality of structural business statistics. Regulation (EC) No **295/2008** of the European Parliament and of the Council of 11 March 2008 concerning structural business statistics (recast),

Domains and sub-domains	Quality requirements/standards for Quality reporting
	as implemented by Commission Regulation (EU) No 275/2010 of 30 March 2010 implementing Regulation (EC) No 295/2008 of the European Parliament and of the Council, as regards the criteria for the evaluation of the quality of structural business statistics, *and as amended by* Commission Regulation (EU) No 446/2014 of 2 May 2014 amending Regulation (EC) No 295/2008 of the European Parliament and of the Council concerning structural business statistics, and Commission Regulations (EC) No 251/2009 and (EU) No 275/2010, as regards the series of data to be produced and the criteria for evaluation of the quality of structural business statistics.
Short-Term statistics (STS)	Council Regulation (EC) No **1165/98** of 19 May 1998 concerning short-term statistics, *as adapted to the Regulatory Procedure with Scrutiny by* Regulation (EC) No **596/2009** of the European Parliament and of the Council of 18 June 2009 adapting a number of instruments subject to the procedure referred to in Article 251 of the Treaty to Council Decision 1999/468/EC with regard to the regulatory procedure with scrutiny; Adaptation to the regulatory procedure with scrutiny — Part Four, *as amended by* Regulation (EC) No **1158/2005** of the European Parliament and of the Council of 6 July 2005 amending Council Regulation (EC) No 1165/98 concerning short-term statistics.
PRODCOM	Council Regulation (EEC) No 3924/91 of 19 December 1991 on the establishment of a Community survey of industrial production (Prodcom) *as amended by:* Regulation (EC) No 1893/2006 of the European Parliament and of the Council of 20 December 2006 establishing the statistical classification of economic activities NACE Revision 2 and amending Council Regulation (EEC) No 3037/90 as well as certain EC Regulations on specific statistical domains
Information Society	Regulation (EC) No **808/2004** of the European Parliament and of the Council of 21 April 2004 concerning Community statistics on the information society *as amended by* Regulation (EC) No **1006/2009** of the European Parliament and of the Council of 16 September 2009 amending Regulation (EC) No 808/2004 concerning Community statistics on the information society *and as implemented by* Commission Regulation (EC) No **1099/2005** of 13 July 2005 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EC) No **1031/2006** of 4 July 2006 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EC) No **847/2007** of 18 July 2007 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EC) No **1023/2009** of 29 October 2009 implementing Regulation (EC) No 808/2004 of the European Parliament

Domains and sub-domains	Quality requirements/standards for Quality reporting
	and of the Council concerning Community statistics on the information society Commission Regulation (EU) No 821/2010 of 17 September 2010 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EU) No 937/2011 of 21 September 2011 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EU) No 1046/2012 of 8 November 2012 implementing Regulation (EC) No 1059/2003 of the European Parliament and of the Council on the establishment of a common classification of territorial units for statistics (NUTS), as regards the transmission of the time series for the new regional breakdown Commission Regulation (EU) No 859/2013 of 5 September 2013 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EU) No 1196/2014 of 30 October 2014 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EU) 2015/2003 of 10 November 2015 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EU) 2016/2015 of 17 November 2016 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society Commission Regulation (EU) 2017/1515 of 31 August 2017 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society for the reference year 2018 Commission Regulation (EU) 2018/1798 of 21 November 2018 implementing Regulation (EC) No 808/2004 of the European Parliament and of the Council concerning Community statistics on the information society for the reference year 2019
Business Registers	Regulation (EC) No **177/2008** of the European Parliament and of the Council of 20 February 2008 establishing a common framework for business registers for statistical purposes and repealing Council Regulation (EEC) No 2186/93 *as implemented by* Commission Regulation (EU) No 1097/2010 of 26 November 2010 implementing Regulation (EC) No 177/2008 of the European Parliament and of the Council establishing a common framework for business registers for statistical purposes, as regards the exchange of confidential data between the Commission (Eurostat) and central banks
FATS	Regulation (EC) No **716/2007** of the European Parliament and of the Council of 20 June 2007 on Community statistics on the structure and activity of foreign affiliates *as implemented by*

Domains and sub-domains	Quality requirements/standards for Quality reporting
	Commission Regulation (EC) No **834/2009** of 11 September 2009 implementing Regulation (EC) No 716/2007 of the European Parliament and of the Council on Community statistics on the structure and activity of foreign affiliates, as regards the quality reports.
Tourism	Regulation (EU) No 692/2011 of the European Parliament and of the Council of 6 July 2011 concerning European statistics on tourism and repealing Council Directive 95/57/EC, *as implemented by* Commission Implementing Regulation (EU) No 1051/2011 of 20 October 2011 implementing Regulation (EU) No 692/2011 of the European Parliament and of the Council concerning European statistics on tourism, as regards the structure of the quality reports and the transmission of the data.
Coal and steel statistics	
	Regulation (EC) No **48/2004** of the European Parliament and of the Council of 5 December 2003 on the production of annual Community statistics on the steel industry for the reference years 2003-2009.
Population and social conditions	
Demography and Migration	Regulation (EC) No 862/2007 of the European Parliament and of the Council of 11 July 2007 on Community statistics on migration and international protection and repealing Council Regulation (EEC) No 311/76 on the compilation of statistics on foreign workers. Regulation (EU) No 1260/2013 of the European Parliament and of the Council of 20 November 2013 on European demographic statistics, *as implemented by* Commission Regulation (EU) No 205/2014 of 4 March 2014 laying down uniformed conditions for the implementation of Regulation (EU) No 1260/2013 of the European Parliament and the Council on European demographic statistics, as regards breakdowns of data, deadlines and data revisions.
Population and Housing Censuses	Regulation (EC) No 763/2008 of the European Parliament and of the Council of 9 July 2008 on population and housing censuses. OJ L 218, 13.8.2008, p. 14-20, *as implemented by:* Commission Regulation (EU) 2017/712 of 20 April 2017 establishing the reference year and the programme of the statistical data and metadata for population and housing censuses provided for by Regulation (EC) No 763/2008 of the European Parliament and of the Council; Commission implementing Regulation (EU) 2017/881 of 23 May 2017 implementing Regulation (EC) No 763/2008 of the European Parliament and of the Council on population and housing censuses, as regards the modalities and structure of the quality reports and the technical format for data transmission, and amending Regulation (EU) No 151/2010([21]).

([21]) Quality reporting requirements for the 2021 EU census are explained in detail in chapters 5 and 6 of Eurostat (2019a).

Domains and sub-domains	Quality requirements/standards for Quality reporting
Health and safety	Regulation (EC) No **1338/2008** of the European Parliament and of the Council of 16 December 2008 on Community statistics on public health and health and safety at work, *as implemented by* Commission Regulation (EU) No 141/2013 of 19 February 2013 implementing Regulation (EC) No 1338/2008 of the European Parliament and of the Council on Community statistics on public health and health and safety at work, as regards statistics based on the European Health Interview Survey (EHIS); Commission Regulation (EU) 2018/255 of 19 February 2018 implementing Regulation (EC) No 1338/2008 of the European Parliament and of the Council as regards statistics based on the European Health Interview Survey (EHIS).
Labour market	Council Regulation (EC) No **577/98** of 9 March 1998 on the organisation of a labour force sample survey in the Community, *as adapted to the Regulatory Procedure with Scrutiny by* Regulation (EC) No **596/2009** of the European Parliament and of the Council of 18 June 2009 adapting a number of instruments subject to the procedure referred to in Article 251 of the Treaty to Council Decision 1999/468/EC with regard to the regulatory procedure with scrutiny; Adaptation to the regulatory procedure with scrutiny — Part Four Council Regulation (EC) No **530/1999** of 9 March 1999 concerning structural statistics on earnings and on labour costs, *as adapted to the Regulatory Procedure with Scrutiny by* Regulation (EC) No **596/2009** of the European Parliament and of the Council of 18 June 2009 adapting a number of instruments subject to the procedure referred to in Article 251 of the Treaty to Council Decision 1999/468/EC with regard to the regulatory procedure with scrutiny Adaptation to the regulatory procedure with scrutiny — Part Four *and as implemented by* Commission Regulation (EC) No **452/2000** of 28 February 2000 implementing Council Regulation (EC) No 530/1999 concerning structural statistics on earnings and on labour costs as regards quality evaluation on labour costs statistics, Commission Regulation (EC) No **72/2002** of 16 January 2002 implementing Council Regulation (EC) No 530/1999 as regards quality evaluation of structural statistics on earnings, Commission Regulation (EC) No **698/2006** of 5 May 2006 implementing Council Regulation (EC) No 530/1999 as regards quality evaluation of structural statistics on labour costs and earnings. Regulation (EC) No **450/2003** of the European Parliament and of the Council of 27 February 2003 concerning the labour cost index, *as adapted to the Regulatory Procedure with Scrutiny by* Regulation (EC) No **596/2009** of the European Parliament and of the Council of 18 June 2009 adapting a number of instruments subject to the procedure referred to in Article 251 of the Treaty to Council Decision 1999/468/EC with regard to the regulatory procedure with scrutiny Adaptation to the regulatory procedure with scrutiny — Part Four, *and as implemented by* Commission Regulation (EC) No **1216/2003** of 7 July 2003 implementing Regulation (EC) No 450/2003 of the European Parliament and of the Council concerning the labour cost index.

Domains and sub-domains	Quality requirements/standards for Quality reporting
	Regulation (EC) No **453/2008** of the European Parliament and of the Council of 23 April 2008 on quarterly statistics on Community job vacancies, *as implemented by* Commission Regulation (EC) No **1062/2008** of 28 October 2008 implementing Regulation (EC) No 453/2008 of the European Parliament and of the Council on quarterly statistics on Community job vacancies, as regards seasonal adjustment procedures and quality reports.
Living conditions and social protection	Regulation (EC) No **1177/2003** of the European Parliament and of the Council of 16 June 2003 concerning Community statistics on income and living conditions, *as implemented by* Commission Regulation (EC) No **28/2004** of 5 January 2004 implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU-SILC) as regards the detailed content of intermediate and final quality reports. Regulation (EC) No **458/2007** of the European Parliament and of the Council of 25 April 2007 on the European system of integrated social protection statistics (ESSPROS), *as implemented by* Commission Regulation (EC) No **1322/2007** of 12 November 2007 implementing Regulation (EC) No 458/2007 of the European Parliament and of the Council of the European system of integrated social protection statistics (ESSPROS) as regards the appropriate formats for transmission, results to be transmitted and criteria for measuring quality for the ESSPROS core system and the module on pension beneficiaries, *as implemented by* Commission Regulation (EU) No 110/2011 of 8 February 2011 implementing Regulation (EC) No 458/2007 of the European Parliament and of the Council on the European system of integrated social protection statistics (ESSPROS) as regards the appropriate formats for the transmission of data, the results to be transmitted and the criteria for measuring quality for the ESSPROS module on net social protection benefits.

Domains and sub-domains	Quality requirements/standards for Quality reporting
Education, science and culture	Regulation (EC) No **1552/2005** of the European Parliament and of the Council of 7 September 2005 on statistics relating to vocational training in enterprises, *as adapted to the Regulatory Procedure with Scrutiny by* Regulation (EC) No **596/2009** of the European Parliament and of the Council of 18 June 2009 adapting a number of instruments subject to the procedure referred to in Article 251 of the Treaty to Council Decision 1999/468/EC with regard to the regulatory procedure with scrutiny Adaptation to the regulatory procedure with scrutiny — Part Four, *and as implemented by:* Commission Regulation (EC) No **198/2006** of 3 February 2006 implementing Regulation (EC) No 1552/2005 of the European Parliament and the Council on statistics relating to vocational training in enterprises; Commission Regulation (EU) No 822/2010 of 17 September 2010 amending Regulation (EC) No 198/2006 implementing Regulation (EC) No 1552/2005 of the European Parliament and of the Council on statistics relating to vocational training in enterprises, as regards the data to be collected, the sampling, precision and quality requirements; Commission Regulation (EU) No 1153/2014 of 29 October 2014 amending Regulation (EC) No 198/2006 as regards the data to be collected, and the sampling, precision and quality requirements. Regulation (EC) No **452/2008** of the European Parliament and of the Council of 23 April 2008 concerning the production and development of statistics on education and lifelong learning, *as implemented by:* Commission Regulation (EU) No 823/2010 of 17 September 2010 implementing Regulation (EC) No 452/2008 of the European Parliament and of the Council concerning the production and development of statistics on education and lifelong learning, as regards statistics on the participation of adults in lifelong learning; Commission Regulation (EU) No 88/2011 of 2 February 2011 implementing Regulation (EC) No 452/2008 of the European Parliament and of the Council concerning the production and development of statistics on education and lifelong learning, as regards statistics on education and training systems; Commission Regulation (EU) No 912/2013 of 23 September 2013 implementing Regulation (EC) No 452/2008 of the European Parliament and of the Council concerning the production and development of statistics on education and lifelong learning, as regards statistics on education and training systems; Commission Regulation (EU) No 1175/2014 of 30 October 2014 implementing Regulation (EC) No 452/2008 of the European Parliament and of the Council concerning the production and development of statistics on education and lifelong learning, as regards statistics on the participation of adults in lifelong learning and repealing Commission Regulation (EU) No 823/2010.
Transport	
Air	Regulation (EC) No 437/2003 of the European Parliament and of the Council of 27 February 2003 on statistical returns in respect of the carriage of passengers, freight and mail by air.

Domains and sub-domains	Quality requirements/standards for Quality reporting
Inlands waterways	Regulation (EU) 2018/974 of the European Parliament and of the Council of 4 July 2018 on statistics of goods transport by inland waterways.
Rail	Regulation (EU) 2018/643 of the European Parliament and of the Council of 18 April 2018 on rail transport statistics.
Road	Council Regulation (EC) No 1172/98 of 25 May 1998 on statistical returns in respect of the carriage of goods by road. Regulation (EU) No 70/2012 of the European Parliament and of the Council of 18 January 2012 on statistical returns in respect of the carriage of goods by road.
Sea	Directive 2009/42/EC of the European Parliament and of the Council of 6 May 2009 on statistical returns in respect of carriage of goods and passengers by sea.
Research and development	
Science technology, innovation	Decision No **1608/2003**/EC of the European Parliament and of the Council of 22 July 2003 concerning the production and development of Community statistics on science and technology, *as implemented by:* Commission Regulation (EC) No **753/2004** of 22 April 2004 implementing Decision No 1608/2003/EC of the European Parliament and of the Council as regards statistics on science and technology; Commission Regulation (EC) No **1450/2004** of 13 August 2004 implementing Decision No 1608/2003/EC of the European Parliament and of the Council concerning the production and development of Community statistics on innovation; and Commission Implementing Regulation (EU) No 995/2012 of 26 October 2012 laying down detailed rules for the implementation of Decision No 1608/2003/EC of the European Parliament and of the Council concerning the production and development of Community statistics on science and technology.

E (PART III) Introduction to big data

E1 Purpose of the document

The document aims at, first, providing state-of-the-art information and clarification with regard to relevant work on quality that has been accomplished in the context of using big data for producing statistics and, second, proposing a number of quality aspects that should be taken on board in a quality framework that would include the use of big data. Subsequently these issues could be taken into account in the Single Integrated Metadata Structure (SIMS) and the underlying reporting structures.

For the purposes of quality reporting, and in order to address the diversity of methods for the production of official statistics, the current handbook defines seven mutually exclusive, exhaustive types of statistical process. It does not treat big data as a statistical process but as a characteristic of diverse range of data sources that may be used for the production of official statistics.

Big data are characterised by the multidimensionality of the statistical unit of interest, the organic way in which they are produced, their real-time or near real-time generation and their diverse structure. Essentially, big data are high dimensional exhaust data left from the use of IT systems or captured by sensors. As there is no single definition, the term is only conventionally used. Moreover, big data calls for a paradigm shift in the types of processes to use, analyse and make inference from them, e.g. algorithmic inference, machine learning, etc.

Despite the tendency to refer to big data as if there were uniformity, there is a great range of data types and various sources that can be described as big data, for example, geo-positioning data of mobile devices, data of satellite images, natural language in social media posts, web-scraped data from enterprise websites, data of smart sensors. All of these sources can be used for producing statistics.

Recent experience with the big data pilots has demonstrated that the possible uses of big data are not associated with a single type of statistical process. For example, the statistical process associated with the use of geo-positioning data collected by mobile devices to produce human mobility statistics is very different from the one associated to the use of data on people's online activity for nowcasting statistical indicators. Also, big data are commonly combined with data from other sources in producing statistics.

Because of the diversity of big data sources and the very recent developments of processes to make use of them, there is presently no well-established quality framework for assessing or reporting quality of the processes using big data and their outputs ([22]). The following paragraphs in this document

([22]) ESSnet Big Data (2018) Report describing the quality aspects of Big Data for Official Statistics, accessed at: https://webgate.ec.europa.eu/fpfis/mwikis/essnetbigdata/images/5/56/WP8_Deliverable_8.2_Quality_aspects.pdf describe current attempts to address this problem.

E2 Past, present and future work on using big data sources

The ESS committed itself to exploring the potential of big data for producing official statistics by adopting the Scheveningen Memorandum (ESSC, 2013))([23]) in 2013 and the Big Data Action Plan and Roadmap (ESSC, 2014))([24]) in 2014.

To implement the plan, Eurostat launched initiatives, such as the ESSnet Big Data([25]) project, to explore the potential of big data and to address its challenges. Until now, the ESSnet Big Data project has developed several proofs-of-concept exploring various data and data sources for producing statistics. These have not yet been implemented in statistical processes producing official statistics. The intention is to start the implementation phase with developments of prototypes, drafting of guidelines and specifications for implementation within the time frame 2018-2020.

It is important that the development of adequate statistical processes and the identification of the necessary quality reporting elements (corresponding to the quality dimensions within a quality framework) should be under the common responsibility of the methodologists in the NSIs as well as under the respective subject matter experts and statistical production units. Further to the involvement of data scientists in exploiting the potential of using big data and producing proofs-of-concept, the active engagement of "frontline" statisticians is the essential prerequisite in all aspects. From that point of view, it is not yet possible to describe fully the statistical products, specify relevant statistical processes and identify quality reporting requirements grouping relevant data types and data sources in the context of using big data for official statistics. This part of the work is as well foreseen within the time frame 2018-2020.

Big data has drawn attention to the value of information and knowledge that one may draw from data. This becomes more prominent considering the extended use of smart technologies and of the Internet of Things (IoT) that will eventually take big data to a whole new level and change the data landscape. Data capturing and processing capabilities coupled with analytical and statistical capabilities will be embedded in the smart systems themselves. In addition, algorithms will handle huge amounts of (near) real-time data. Therefore, data originally "designed" to satisfy the operational requirements of smart systems, could be used for producing smart statistics.

Statistics are only useful when they are trusted. In order to build trust in smart statistics the data life-cycle needs to be auditable and transparent, with guarantees of accuracy and privacy protection by design. Trusted smart statistics is an extension of the current Eurostat work that aims at enhancing data innovation processes and production processes for official statistics.

It is expected that algorithmic processing will go far beyond traditional data processing methods. Therefore, quality should be examined in the wider context of smart technologies shaping the future of statistics emphasising the need to embed trust in smart statistics under principles for transposing algorithmic transparency and accountability in smart statistics. Work on smart statistics will start within the time 2018-2020 and extend to 2022 (Business case on "Smart Statistics & Big Data" 2018).

([23]) European Statistical System Committee (2013). Scheveningen Memorandum on ‘Big Data and Official Statistics’. Retrieved from http://www.cros-portal.eu/sites/default/files/SCHEVENINGEN_MEMORANDUM%20Final%20version.pdf

([24]) European Statistical System Committee (2014). Big Data Action Plan and Roadmap. Retrieved from http://www.cros-portal.eu/sites/default/files/ESSC%20doc%2022_8_2014_EN_Final%20with%20ESSC%20opinion.pdf

([25]) **ESSnet Big Data** has been a project within the European Statistical System (ESS) jointly undertaken by 22 partners. Its objective has been the integration of big data in the regular production of official statistics, through pilots (development of proofs-of-concepts) exploring the potential of selected big data sources and building concrete applications. ESSnet Big Data has started in 2016 and ran for 28 months until May 2018; it consisted of horizontal and content-oriented work package. https://ec.europa.eu/eurostat/cros/content/ess-big-data-action-plan-and-roadmap-10_en

E3 Relationship between big data and administrative data

There is a small overlap between administrative data and big data, but it is basically incorrect to say that big data is a special case of administrative data.

A first opinion on the relationship was expressed at the UNECE(26) level where it was made clear that one cannot consider administrative data "big" according to the classification (for the big data) that was then proposed. It was moreover suggested that "*Administrative data can become "big" when the velocity increases, e.g. using extensively administrative data where data is collected every day or every week instead of the usual once a year or once a month.*" The latter statement was, however, addressing only part of the problem.

In order to understand how big data and administrative data relate to each other it is important to distinguish the data source from the data itself, in particular the type of data they refer to.

Administrative data, according to Eurostat glossary(27), "*refer to the set of units and data derived from an administrative source*". However, in terms of data type it consists of registers(28).

In addition, *administrative source*, according to the glossary(29) in statistics explained "*is the register of units and data associated with an administrative regulation (or group of regulations), viewed as a source of statistical data*".

Therefore, administrative data are generated in a controlled environment (often based on requirements of legal or administrative nature) while the big data generation process is typically not controlled. Multiple data types and data sources add to the complexity of big data.

The big data phenomenon makes one realise that there are several types of data, e.g. multimedia (images, sound and video), network data (composed of nodes and links between nodes), natural language data, geo-positioning data, signal data (from sensors), online / web activity. What characterises these types of data is their extremely high dimensionality in relation to the statistical unit of analysis.

With reference to sources, there are several types of big data sources. UNECE, 2013(30) proposes a classification of big data sources based on how they are generated. Firstly, *human-sourced* information available mostly from social networks, blogs, internet searches, etc. where data are loosely structured and often ungoverned (e.g. Facebook and Twitter); Secondly, *process-mediated* data available from the IT systems of organisations (private or public), where data is usually structured and stored in relational databases (e.g. credit card transactions stored by banks, bank transfers, booking systems, web platforms such as AirBnB, Uber, etc.); Thirdly, *machine-generated* data captured by sensors and other machines used to measure and record events in the physical world (e.g. traffic sensors and web logs).

With these conceptual considerations, we can now understand that in terms of data sources big data is a superset of administrative data and not a special case, as administrative data are part of *process-mediated* data. In terms of data type, it should be clear that some types of big data are clearly not possible to be captured / generated by "a register of units and data associated with an administrative regulation". The small overlap exists when the register is very detailed (e.g. bank transfers) and, when considering the statistical unit of interest (e.g. individual person), it results in high dimensional data (e.g. for each person, the list of all transactions during one year).

(26) UNECE, 2013 What does "big data" mean for official statistics https://statswiki.unece.org/pages/viewpage.action?pageId=77170614&preview=/77170614/80805923/Big%20Data%20HLG%20Final%20Published%20Version.docx

(27) http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Administrative_data

(28) For a discussion about register data see Wallgren, Wallgren (2014) Register-based Statistics: Statistical Methods for Administrative Data

(29) http://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Administrative_source

(30) https://statswiki.unece.org/display/bigdata/Classification+of+Types+of+Big+Data

E4 Quality of statistics produced with use of big data

In the case of what conventionally is described as big data, the sheer volume of data and observations does not guarantee quality. On the contrary, the unwanted bias and noise in most big data sources are without a doubt some of the more complicated challenges for statisticians. There is already a significant body of work done and literature on quality assessment of statistics produced with the use of big data, even if there isn't still an established quality framework.

There are few specific proposals of quality assessment frameworks. The first proposal is a deliverable from the UNECE/HGLProject([31]), The Role of Big Data in the Modernisation of Statistical Production, and specifically describes the findings of the Big Data Quality Task Team. The UNECE framework is briefly presented in Section E5 below.

Another proposal is a report on big data, produced by the American Association for Public Opinion Research([32]). Among many other aspects, the report suggested the extension of the total survey error (TSE) and the development of a Big Data Total Error (BDTE) framework. The report identified the need to include sources of errors that are unique to big data (related to generic steps where errors may originate, i.e., Generation, ETL – Extraction, Transform and Load – Analysis) and can create substantial biases and uncertainties in statistical inference from big data.

The third proposal has been one of the initial efforts of Eurostat (Eurostat, 2014)([33]) to investigate accreditation procedures of quality frameworks in the context of using data from a wide range of secondary data sources. Among other aspects, it was pointed out that relevant data sources should be evaluated *ex ante* according to specific quality criteria.

In 2016, a "Comparative assessment of (the) three quality frameworks for statistics derived from big data: the cases of Wikipedia page views and Automatic Identification Systems"([34]) has been presented in the European Conference on Quality of Official Statistics in Madrid. Further advancement of this work titled "Conceptualising quality for big data"([35]) has been presented in the 2018 European conference on Quality in Official Statistics in Krakow, proposing explicit quality dimensions for a big data quality framework.

The ESSnet Big Data has also published its report on quality aspects([36]). The report identified the most important quality aspects in the context of the specific pilot projects and not necessarily on products declared as official statistics. The seven quality aspects as identified in a workshop that was organised in the framework of the project, sorted by the importance assigned during the workshop, were coverage, comparability over time, processing errors, process chain control, linkability, measurement errors, model errors and precision.

It is noted that Eurostat has published "An overview of methods for treating selectivity in big data sources"([37]), which covers definitional aspects, the particular case of the quality dimension accuracy, especially in what concerns inference and also the statistical process when big data sources are

(31) UNECE (2014) A Suggested Framework for the Quality of Big Data https://statswiki.unece.org/display/bigdata/2014+Project?preview=%2F108102944%2F108298642%2FBig+Data+Quality+Framework+-+final-+Jan08-2015.pdf, accessed on 16 June 2018.

(32) American Association for Public Opinion Research (AAPOR), 2015, Report on Big Data

(33) Eurostat (2014) Petrakos M et al., Analysis of methodologies for using the Internet for the collection of information society and other statistics, 2014, European Commission

Deliverable on the "Accreditation procedure for statistical data from non-official sources"

https://ec.europa.eu/eurostat/cros/system/files/D5_Accreditation%20procedure%20for%20statistical%20data%20from%20non-official%20sources_20140206_0.pdf

(34) Reis, et.al. (2016) Comparative assessment of three quality frameworks for statistics derived from big data

(35) Brancato, Di-Consiglio (2018) Conceptualising quality for big data

(36) ESSnet Big Data (2018) Report describing the quality aspects of Big Data for Official Statistics

(37) Beręsewicz M. et al. 2018, An overview of methods for treating selectivity in big data sources https://ec.europa.eu/eurostat/web/products-statistical-working-papers/-/KS-TC-18-004

used.

E5 Suggestions for quality of big data and statistics

Quality of statistical data or products is effectively assessed with reference to quality frameworks. These quality frameworks systematically address quality by referring to "quality dimensions". The following does not aim at proposing a new quality framework at this stage.

Table E5.1 Quality dimensions in the UNECE framework for the quality of big data

<table>
<tr><th></th><th>Input</th><th>Throughput</th><th>Output</th></tr>
<tr><td rowspan="2">Source</td><td>Institutional/business environment</td><td rowspan="14">System independence
Steady States
Quality Gates</td><td>Institutional/business environment</td></tr>
<tr><td>Privacy and security</td><td>Privacy and security</td></tr>
<tr><td rowspan="7">Metadata</td><td>Complexity</td><td>Complexity</td></tr>
<tr><td>Completeness</td><td>Accessibility and Clarity</td></tr>
<tr><td>Usability</td><td>Relevance</td></tr>
<tr><td>Linkability</td><td></td></tr>
<tr><td>Coherence – consistency</td><td></td></tr>
<tr><td>Validity</td><td></td></tr>
<tr><td>Time-related factors</td><td></td></tr>
<tr><td rowspan="5">Data</td><td>Accuracy and selectivity</td><td>Accuracy and selectivity</td></tr>
<tr><td>Linkability</td><td>Coherence – Linkability</td></tr>
<tr><td>Coherence-consistency</td><td>Coherence-consistency</td></tr>
<tr><td>Validity</td><td>Validity</td></tr>
<tr><td></td><td>Time-related factors</td></tr>
</table>

The UNECE quality framework([31]) (see Table E5.1, above) addresses quality issues following an *input - throughput - output* model for statistical production and adopts a hierarchical structure, as already suggested in the quality framework for administrative data, where quality dimensions are nested in three hyper-dimensions: *source, metadata and data*. The *data* hyperdimension relates to the data themselves, whereas the *source and metadata* hyper-dimensions relate to the conditions that govern the data supply and to the availability and kind of information available on the concepts and contents, respectively. The three hyper-dimensions and corresponding quality dimensions are considered in each phase of the *input - throughput - output* model.

As assessed in (Reis et al., 2016)([34]), in this quality framework, specific aspects related to big data are taken into account by considering a new quality dimension (in comparison to proposed frameworks for the quality of administrative data) for the complexity of the *input* data, and the need for new skills and new IT infrastructures.

Despite the different terminology, the *output* quality dimensions listed in the UNECE quality framework are related to the overarching quality criteria as defined in Article 12 of EU-Statistics Regulation 223([38]). With respect to the quality of results (output) the latter quality criteria refer to two additional criteria. "*Punctuality*" which refers to the delay between the date of the release of the data

([38]) Regulation (EC) No 223/2009 (11 March 2009) on European statistics https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32009R0223&from=EN, accessed on 21 June 2018 and the target date (the date by which the data should have been delivered) and "*comparability*" which refers to the measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas, sectoral domains or over time.

With reference to the "*input*" stage the earlier study of Eurostat(2014)[33] proposes an accreditation procedure which would guide statistical authorities in their selection of big data sources. The study differentiates the issue of complementary data collection methods such as scraping websites of enterprises and refers to the design of an accreditation procedure for (new) data sources. Practically this decision-making procedure is based on standard, desired and ideal properties. Namely, the procedures should be compliant with established principles of quality frameworks and consistent with quality assurance practices. Subsequently, procedures should be flexible, follow a stepwise approach and include an empirical assessment with real data. As the data source is not under the control of the statistical authority in question, a risk assessment should accompany the quality assessment. It is expected that certification and accreditation of new data sources become a mark of quality for official statistics and contribute to developing trusted statistics when using new data sources.

Based on the experiences in developing proofs-of-concepts and the existing streams of work mentioned above, the aim is complementary to the quality dimensions proposed by the UNECE to enumerate/describe relevant quality "aspects" that should be considered in the Single Integrated Metadata Structure (SIMS) and the future underlying quality reporting structures.

Selectivity is identified as an important sub-dimension of data accuracy when assessing quality of big data (*input* phase). An important factor to be considered is that big data include the registration of events (e.g. registration of mobile devices' geo-positioning) that indirectly refer to statistical units in the target population (e.g. persons). Therefore, as stated in the report describing the quality aspects of big data for official statistics, selectivity is strongly related to the Linkability of big data and to the comparability over time.

The UNECE's quality framework does not refer to explicit processes that are involved in the "*throughput* phase"([39]), nor lists a specific set of quality indicators for this phase (process and analyse stages). However, selectivity may become an issue in weighting procedures that usually take place in the *throughput* phase hence considered source of **processing errors**.

In more general terms "*A more detailed treatment of quality issues in this* [*throughput*-phase] *part of the business process needs to take into account the wide and expanding range of data sources and the uses to which they can be put. This involves an expansion of statistical quality control to a wider range of data sources and data types."*[31]

Another quality aspect is the **measurement error** i.e. error that cause recorded values of variables to be different from the true ones hence values included in big data may not be correctly measured (*input* phase). It is referred to as error of data collection and not the error in the source itself. Scraping a job vacancy when there is no job vacancy in a company, is not a measurement error but a coverage error. Similarly, if there is no transmission of Automated Identification System (AIS) data for vessel tracking due to unfavourable meteorological conditions or due to switched off AIS data transmitter, it should be considered as non-response.

The report on the quality of the pilot projects refers to numerous **measurement errors** examples. Web scraping of job vacancies on job portals faces measurement errors, that are mainly the result of scraping errors (scraper may download incorrect data from the web page), errors on the web page or incorrect data on the web page (e.g. employers may upload incorrect data).

Using AIS data, measurement errors may be technical errors and/or human errors. Technical errors are related to dynamic data such as position of ship, speed, course and rotation which comes from AIS device (sensors, cables and antenna). As AIS is a radio signal, technical errors can also arise due to meteorological or magnetic factors disturbing the transmission of the radio signal. These errors can affect every part of the message. Human errors are related to static (ship number, ship's name, call sign, type, length) or voyage data (draught, destination) which are manually entered in the AIS devices so therefore are a common cause of errors. Most of these errors are due to faulty or missing

([39]) (UNECE, 2014) Throughput – any point in the business process in which data is transformed, analysed or manipulated. This might also be referred to as 'process quality'. In GSBPM terms, it encapsulates the "process and analyse" stages of the business process.

input by the ship crews.

When using data from road sensors, the most important measurement error relates to measuring the number of vehicles per lane and the type of vehicles (at a given space and time).

Comparability over time is beyond the UNECE quality dimension of "time related factors" (metadata hyperdimension, *input* phase), that refers to factors such as "Timeliness", "Periodicity", and "Changes through time". This quality aspect is not an issue when conducting pilots and developing one off proofs-of-concepts. However, it becomes an issue the moment we attempt to introduce statistical processes for regularly produced statistical products. Technological changes as well as spreading of technology will affect comparability over time.

For example, "multiple counting" is the result of the same job ads appearing in an increasing number of platforms and websites. Subsequently, the job vacancy could be counted more times as they appear in more platforms. The issue of comparability would become more complex as in the course of time the job vacancy can be still visible -in some places- even if it was filled. In addition, as a result of "technological development" more job ads will appear online (including social media) because more companies will hire online.

In the case of smart electricity meters, although the sources remain the same, source of variation may be the continuously increasing number of installed meters, issues related to measuring consumption for businesses (intermediate consumption) and for households (final consumption), and the fact that electricity consumers become electricity producers.

Moreover, technological changes affect comparability over time between countries because technological evolution is not uniformly distributed across countries.

As it was mentioned above, **selectivity** is strongly related to the **Linkability** of big data, referring to linking or combining big data with other data sources. The issue of using the same statistical unit (e.g. geo-positioning data of a portable device vs geo-positioning of a person as reported), using common identifiers and the methods for data linking are important.

Fitting big data concepts into concepts relevant for statistical units and in general combining various data sources is expected to have consequences on all quality dimensions, the most significant ones are accuracy and comparability.

The quality dimensions in the UNECE framework for the quality of big data presented in Table E5.1 are structured in the three stages *"input", "throughput" and "output"*. With reference to "*throughput*", the variety of processing steps when using the wide range of different big data sources only allows quality indicators that describe general principles. The three principles enumerated by the framework are "System Independence", "Steady states" and "Quality gates". System independence refers to a data processing pipeline which results do not depend on the particular implementation in terms of software/hardware, but solely on the methods adopted. Steady states refer to intermediate datasets in the data processing pipeline which are stored and assessed. Potentially, these steady-states can feed the data processing pipelines of several different statistical products. Quality gates refer to intermediate datasets in the data processing pipeline for which quality is assessed. However, a detailed description of processing steps – that correspond to the "*throughput*" stage - and the respective processing errors can be found for each data source in the report describing the quality aspects of big data for official statistics([40]).

In addition, "*throughput*" can be conventionally used as the right stage to refer to the **process chain control.** The distinct importance of process chain in the overall quality comes from the fact that only data of sufficient quality (and quantity) should be used for solving a problem. This data however may come from intermediate process steps for which they have to meet specific standards. For example, quality checks were included in the processing of road sensor data used for the Road Traffic Intensity Statistics in the Netherlands. The process chain comprised the following steps for which quality standards had to be met: transformation and selection of data, the selection of reliable sensors, imputation of missing values and calibration. In addition, this quality aspect becomes more important with the involvement of multiple partners in a data-driven problem-solving framework. From that aspect process chain control is related to the data and process governance that may no longer be under the respective responsibility of the individual countries when preparing national official

([40]) ESSnet Big Data (2018) Report describing the quality aspects of Big Data for Official Statistics, p.32-44.

statistics.

In contrast to design-based statistical inference, the complexity of model-based and algorithmic inference lies primarily on the connection of the information of interest and the plethora of data from the big data sources. Whereas probabilities lie in the core of statistical models, machine learning algorithms are of ad hoc and heuristic nature. Therefore, estimating the quality of models is important. Particularly in machine learning, minimising the prediction error would be necessary. However, though the predictive accuracy of supervised methods can be estimated, in the case of using non-probabilistic sample of training data from the target population should not be trusted alone. Additional techniques can be exploited e.g. using feedback information, for improving the predictive accuracy.

Summaries of models that were used when developing proofs-of-concept using big data sources can be found in the report describing the quality aspects of big data for official statistics([41]).

As it was mentioned above in the context of future work on trusted smart statistics, quality should be examined for the automated decisions in smart technologies. These decisions that are part of the statistical operations are driven by algorithms of greater complexity than in the past. These algorithms should be considered under principles of transparency and accountability.

E6 Concluding remarks

The development of proofs-of-concept and prototypes for producing statistics using big data sources is important in order to illustrate sources of bias and errors and develop adequate error estimation models. It should be followed by an attempt to define an enhanced quality framework combing the advantages of the previously developed frameworks, with specific quality factors and quality indicators respectively.

Trust in official statistics requires among other factors (structural factors and reputation) compliance with standards and having sound methodological practices, robust statistical processes and quality outputs. The situation is not different when using big data even though it is considered necessary to revisit the issue of quality on a case by case basis.

Subsequently, the corresponding quality reporting elements (set of quality indicators) should find their place in reporting quality in the use of big data for statistical production.

([41]) ESSnet Big Data (2018) Report describing the quality aspects of Big Data for Official Statistics, p.49-53

F (PART III) References

Arnež, Marta; Hribar, Bernarda; Inthihar Stanka (2008), *Standard Quality Report for Household Budget Survey for the year 2004*, Statistical Office of the Republic of Slovenia

Biemer, P. and Stokes, L. (1991): *Approaches to Modelling Measurement Error,* in Biemer et al (1991)

Biemer, P. Groves, R.M. Lyberg, L.E. Mathiowetz, N.A. and Sudman, S., (eds.) (1991) *Measurement Errors in Surveys. John Wiley & Sons, New York*

Biemer, P., de Leeuw, E., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L. E., Tucker and N.,C., West, B.,T., (eds.), (2017): *Total Survey Error in Practice*. Wiley, New Jersey.

Brooks, C.A. and Bailar, B., A. (1978): *An Error Profile: Employment as Measured by the Current Population Survey*. Statistical Working Paper 3, Office of Federal, Statistical Policy and Standards, U.S. Department of Commerce.

Central Statistics Office Ireland (2013), *Standard Report on Methods and Quality for BOP and Related Results Compilation 2011*

DESTATIS, Germany (2018), *Quality Report on National Accounts*, Wiesbaden. https://www.destatis.de/EN/Methods/Quality/QualityReports/National-Accounts-Domestic-Product/national-accounts.pdf?__blob=publicationFile

European Central Bank (2013), *The Eurosystem Household Finance and Consumption Survey - Methodological Report for the First Wave*. Frankfurt: ECB, Statistics Paper Series No 1

Eurostat (2001), *Compendium of HICP reference documents (2/2001/B/5),* European Communities,

Eurostat (2007): *Quality report 2007-05-03 pursuant to Regulation (EC) No 501/2004 on quarterly financial accounts for general government* https://ec.europa.eu/eurostat/documents/1015035/2041369/QR-ON-QFAGG-OJ.pdf

Eurostat (2011): ESSPROS Quality Reports. Core System and Pension Beneficiaries https://ec.europa.eu/eurostat/documents/3888793/5855557/KS-RA-11-030-EN.PDF/0ae0e7f4-114a-4fe0-838d-18b9a8b73a7b?version=1.0

Eurostat (2015): ESS guidelines on seasonal adjustment https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-15-001

Eurostat (2016): *User guide on European statistics on international trade in goods 2016 edition* https://ec.europa.eu/eurostat/documents/3859598/7679615/KS-GQ-16-009-EN-N.pdf/073b853a-a4f4-4c55-aaba-162671544c78

Eurostat (2017): *Report on the Eurostat 2017 user satisfaction survey*. https://ec.europa.eu/eurostat/documents/64157/4375449/USS2017+Report/e6f2ad0d-ee22-4213-911c-22e8b1481296

Eurostat (2017): *Labour Force Survey at ESS level, Quality report of the European Union, Labour*

Force Survey 2015
https://ec.europa.eu/eurostat/search?p_auth=2Vxm0r7K&p_p_id=estatsearchportlet_WAR_estatsearchportlet&p_p_lifecycle=1&p_p_state=maximized&p_p_mode=view&_estatsearchportlet_WAR_estatsearchportlet_theme=empty&_estatsearchportlet_WAR_estatsearchportlet_action=search&_estatsearchportlet_WAR_estatsearchportlet_collection=empty&text=Labour+Force+Survey+at+ESS+level

Eurostat (2018c): *Balance of Payments Quality Report.*
https://ec.europa.eu/eurostat/search?p_auth=2Vxm0r7K&p_p_id=estatsearchportlet_WAR_estatsearchportlet&p_p_lifecycle=1&p_p_state=maximized&p_p_mode=view&_estatsearchportlet_WAR_estatsearchportlet_theme=empty&_estatsearchportlet_WAR_estatsearchportlet_action=search&_estatsearchportlet_WAR_estatsearchportlet_collection=empty&text=Balance+of+Payments+Quality+Report

Eurostat (2018d): *ESS guidelines on temporal disaggregation, benchmarking and reconciliation*
https://ec.europa.eu/eurostat/documents/3859598/9441376/KS-06-18-355-EN.pdf/fce32fc9-966f-4c13-9d20-8ce6ccf079b6

Eurostat (2019a): *EU legislation on the 2021 population and housing censuses – explanatory notes*
https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/KS-GQ-18-010

Eurostat (2019b): Quality report on European statistics on international trade in goods.
https://ec.europa.eu/eurostat/documents/7870049/9568307/KS-FT-19-002-EN-N.pdf/856f28e5-e9f6-4669-8be0-2a7aa5b1ee67

Eurostat (2019c): Quality Guidelines for Multisource Statistics. ESSnet KOMUSO.
https://ec.europa.eu/eurostat/cros/content/quality-guidelines-multisource-statistics-qgmss_en

ILO (2004): *Consumer Price Index Manual: Theory and Practice*, International Working Group on Price Statistics.

ILO (2009): *ILO School-to-work transition survey: A methodological guide*
https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/instructionalmaterial/wcms_140859.pdf

Ladiray, D. and Sartori, F. (2001), Q*uality in short-term statistics: the Euro-SICS database of Eurostat*,
https://ec.europa.eu/eurostat/documents/3888793/5810937/KS-AN-01-018-EN.PDF/a5f06157-35dc-48cd-a3f9-8071b90bce30

Laitila, T., Lindgren, K., Norberg, A. and Tongur, C. (2017) *Quantifying Measurement Errors in Partially Edited Business Survey Data*. In Biemer et al (2017)

Lee, H. (1995): *Outliers in Business Surveys*. In Cox, B.G., Binder, D.A., Chinnappa, B.N., Christiansson, A., Colledge, M.J. and Kott, P.S. (1995): *Business Survey Methods,* Wiley, New York.

Lien, H., Beyrer, S. and Lunde, H. (2009): *Quality Report on the Norwegian Structure of Earnings Survey 2006 Oslo-Kongsvinger*: Statistics Norway, Reports 2009/20

Lyberg, L. (1983): *The development of procedures for industry and occupation coding at Statistics Sweden*. Statistical Review 21, no. 5, 139-156

Nordberg, L. (2000): *On Variance Estimation for Measures of Change. When Samples are Coordinated by the Use of Permanent Random Numbers*, Journal of Official Statistics, 16, 363-378.

OECD et al (2002): *Measurement of the Non-Observed Economy: A Handbook*, OECD Publications Service https://www.oecd.org/sdd/na/1963116.pdf

Särndal, C.-E., Swensson, B. and Wretman, J. (1992): *Model assisted survey sampling*, Springer-Verlag, New York.

Seljak, R.; Katnič, N. (2006): *Standard Quality Report for the Monthly Survey on* Turnover, New Orders and Value of Inventories in Industry 2005, Statistical Office of the Republic of Slovenia

Statistics Sweden (2018a) Labour Force Survey Quality Declaration:
http://www.scb.se/contentassets/c12fd0d28d604529b2b4ffc2eb742fbe/am0401_kd_2018_pb_180906.pdf

Statistics Sweden (2019): Quality Declaration, National Accounts. https://www.scb.se/en/finding-statistics/statistics-by-subject-area/national-accounts/national-accounts/national-accounts-quarterly-and-annual-estimates/

UK Office for National Statistics (2011): *Labour Force Survey Quality and Methodology Information Report*

UK Office for National Statistics (2012): Coherence between the Gross Value Added calculation in the Annual Business Survey and National Accounts

UNECE (2015): *Recommendations for the 2020 Censuses of Population and Housing*. United Nations https://www.unece.org/publications/2020recomm.html

de Waal, T,. van Delden, A. and Scholtus, S. (2019): Quality Measures for Multisource Statistics. Statistical Journal of the IAOS 35:2 https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji180468?resultNumber=0&totalResults=6191&start=0&q=Quality+Measures+for+Multisource+Statistics&resultsPageSize=10&rows=10

Williams, J. (2012): T*he 2010 Decennial Census: Background and Issues Specialist in American National Government*, Congressional Research Service

Wood, J. (2008): *On the Covariance Between Related Horvitz-Thompson Estimators*, Journal of Official Statistics, 24, 53-78.

Yang, K. and Banamah, A. (2013): Quota Sampling as an Alternative to Probability Sampling? An Experimental Study, Sociological Research Online. https://journals.sagepub.com/doi/10.5153/sro.3199

Getting in touch with the EU

In person
All over the European Union there are hundreds of Europe Direct Information Centres. You can find the address of the centre nearest you at: https://europa.eu/contact

On the phone or by e-mail
Europe Direct is a service that answers your questions about the European Union. You can contact this service
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by electronic mail via: https://europa.eu/contact

Finding information about the EU

Online
Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu

EU Publications
You can download or order free and priced EU publications from EU Bookshop at: https://bookshop.europa.eu. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/contact)

EU law and related documents
For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex at: https://eur-lex.europa.eu

Open data from the EU
The EU Open Data Portal (https://data.europa.eu/euodp/en/data) provides access to datasets from the EU. Data can be downloaded and reused for free, both for commercial and non-commercial purposes.

KS-GQ-19-006-EN-N

European Statistical System handbook for quality and metadata reports

The ESS handbook for quality and metadata reports is recognised as an ESS standard and included in the catalogue of ESS standards thus representing a visible component of the ESS standardisation process. It updates the 2014 ESS handbook on quality reports (EHQR) and has been profoundly revised after the endorsement of SIMS V2.0 by the ESSC in November 2015. This publication fully incorporates SIMS V2.0 combining the ESS Standard for Quality Report Structure (ESQRS) and the Euro-SDMX Metadata Structure (ESMS), and follows the structure of SIMS V2.0. The handbook includes revised guidelines, many examples of existing reports, new material on administrative data, big data, multi-source processes, and other information.

For more information

https://ec.europa.eu/eurostat/

Publications Office of the European Union

EUROPEAN STATISTICAL SYSTEM

ISBN 978-92-76-09154-7