Documentation of official statistics and Quality reporting

Training course
"Quality management of official statistics"
Jerusalem, Israel 13 – 15 March 2018

Giorgia Simeoni
Istat, Division «Methods, quality and metadata»
Project manager «Methods and tools for reference metadata documentation and communication»
simeoni@istat.it

Istat

Contents

- Documentation and Quality reporting in ENSC CoP and QAF
- Importance of documentation for producers
- Definitions of main concepts (metadata, quality reporting)
- Quality reporting as a tool for assessment
- Standards for quality reporting

Istat

Documentation and Quality reporting in ENSC CoP

Principle 15: Accessibility and Clarity. Official 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.

- **Indicator 15.5:** Metadata are documented according to standardised metadata systems.
- **Indicator 15.6:** Users are kept informed about the methodology of statistical processes including the use of administrative data.
- **Indicator 15.7:** Users are kept informed about the quality of statistical outputs with respect to the quality criteria for official statistics.

Istat

Documentation and Quality reporting in ENSC SCoP

Indicator 4.3: Product quality is regularly monitored, assessed with regard to possible trade-offs, and reported according to the quality criteria for official statistics.

Indicator 6.4: Information on the methods and procedures used is publicly available.

Indicator 12.2: Sampling errors and non-sampling errors are measured and systematically documented according to the European and International standards.

Istat

Documentation and Quality reporting in QAF

Methods at institutional level: develop

infrastracture for documentation,

procedure to use quality reports for monitor product quality

At process/product level: produce documentation on

administrative data concepts,

methodology,

data collection,

coding,

editing and imputation,

repositories of data and metadata,

data transmission.

Istat

Why this stress on documentation?

Documentation is additional work, it is boring, it is time consuming, and the benefits are not immediate, but…

… it helps user in the interpretation of the statistics

… it assures traceability

… it helps sharing good practices

… knowledge is no more in the head of persons that could change job or retire

… it supports standardising terminology

… it supports comparability

… it increase accessibility an clarity

… it is a prerequisite for quality assessment and improvement

Istat

What should be documented?

- Description of the **contents of statistical outputs**:

  Documentation concepts and definitions adopted for variables, classifications, indicators disseminated
- Description of **statistical process**:

  Documentation of methods and tools applied to produce the statistical outputs, from the design (sampling design, questionnaire development…) till the dissemination
- Description of **product and process quality**

  Documentation of actions implemented to improve quality, qualitative and quantitative assessment by quality dimension (timeliness, accuracy)

In one word … **METADATA**!

Istat

What should be documented?

"The actual documentation should consist of metadata on the production process and the information content, quality measures and indicators concerning the product, and data on the producing organisation's strategies, policies and user relationships."

Summary Report from the Leadership Group (LEG) on Quality

Istat

Metadata

Metadata is data that ***defines*** and ***describes*** other data.

Source: OECD Glossary of statistical terms

Metadata: 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.

Source: Eurostat's Concepts and Definitions Database

Istat

Metadata

Metadata can be classified in several ways, but a traditional distinction is made between

- *structural metadata* and
- *reference metadata*.

Istat

Structural metadata

Structural metadata refers to metadata that act as ***identifiers*** and descriptors of the data.

Source: OECD Glossary of statistical terms

Structural metadata are needed to ***identify***, use, and process data matrixes and data cubes, e.g. names of columns or dimensions of statistical cubes. Structural metadata must be associated with the statistical data, otherwise it becomes impossible to identify, retrieve and navigate the data.

Source: SDMX Metadata Common vocabulary

Istat

Structural metadata

Structural metadata is 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.

Source: Eurostat's Concepts and Definitions Database

Istat

Reference metadata

Reference metadata describes the **contents** and the **quality** of the statistical data from a semantic point of view. It includes explanatory texts on the **context** of the statistical data, **methodologies** for data collection and data aggregation as well as **quality** and **dissemination** characteristics.

Source: Eurostat's Concepts and Definitions Database

Istat

Reference metadata

Reference metadata are metadata describing the **contents** and the **quality** of the statistical data.

Reference metadata should include all of the following:

a) - "**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;

b) - "**methodological**" metadata, describing methods used for the generation of the data (e.g. sampling, collection methods, editing processes);

c) - "**quality**" metadata, describing the different quality dimensions of the resulting statistics (e.g. timeliness, accuracy).

Source: SDMX Metadata Common vocabulary

Istat

eurostat

Important legal notice

v3.4.1-20170407-5840-PROD_EUROBASE
DATA-EXPLORER_PRODmanaged11

Explanatory texts (metadata) | Information | Download | Preview | Bookmark | Demo | Help | Login

Nights spent at tourist accommodation establishments - monthly data **[tour_occ_nim]**

Last update 08-02-2018

Table Customization show

TIME | GEO | Tourism indicator: Nights spent, total

Unit of measure: Number | Classification of economic activities - NACE Rev.2: Hotels; holiday and other short-stay accommodation; camping grou

GEO \ TIME	2016M01	2016M02	2016M03	2016M04	2016M05	2016M06	2016M07	2016M08	2016M09	2016M10
Belgium	1,580,235	2,117,115	2,575,269	2,920,417	3,345,618	3,022,995	5,160,596	5,083,503	3,175,838	3,123,535
Bulgaria	811,748	903,532	742,183	814,402	1,350,464	3,583,521	5,734,834	5,838,154	3,073,622	881,781
Czech Republic	2,664,561	3,078,926	3,400,017	3,222,920	4,114,441	4,303,686	7,346,645	7,111,196	4,588,389	3,854,640
Denmark	1,091,901	1,306,345	1,704,422	2,235,399	3,112,694	3,471,128	6,650,011	4,675,163	2,600,882	2,113,145
Germany (until 1990 former t	19,333,712	21,736,454	26,896,100	27,808,370	38,062,981	36,638,069	47,115,398	48,494,358	39,741,272	36,621,844
Estonia	332,606	326,326	392,503	424,579	515,136	653,758	938,930	799,710	480,827	503,603
Ireland	: (u)	: (u)	: (u)	: (u)	: (u)	: (u)	: (u)	: (u)	: (u)	: (u)
Greece	1,256,803 (e)	1,178,884 (e)	1,823,110 (e)	3,297,271 (e)	9,539,876 (e)	14,844,477 (e)	21,312,069 (e)	23,738,199 (e)	14,952,730 (e)	6,976,113 (e)
Spain	19,905,363	22,064,716	29,883,131	29,978,833	39,181,694	46,041,801	62,823,799	71,377,518	49,297,403	39,386,844
France	17,519,839	20,436,026	23,507,974	22,866,869	32,589,165	37,391,696	73,271,713	83,930,209	35,975,687	22,945,832
Croatia	386,479	419,303	839,528	1,667,191	4,256,988	8,940,168	22,799,348	25,420,248	9,641,602	2,261,069
Italy	15,743,677	17,004,048	21,144,458	21,266,784	30,994,676	45,071,352	72,084,413	83,434,333	43,811,269	24,225,497
Cyprus	218,431	280,506	474,197	855,595	1,737,872	2,079,629	2,407,604	2,465,873	2,111,845	1,846,069
Latvia	240,921	226,173	271,388	299,649	378,593	475,236	643,721	604,326	374,570	340,728
Lithuania	382,715	371,167	437,829	463,373	622,641	753,265	994,850	1,007,241	620,679	491,724
Luxembourg	134,874	139,400	166,530	217,277	340,466	358,697	473,476	406,572	236,357	200,986
Hungary	1,291,142	1,381,548	1,790,939	1,857,931	2,395,916	2,818,855	4,703,652	4,713,677	2,507,099	2,198,867
Malta	415,152	487,975	639,689	737,471	825,864	873,672	1,029,652	1,071,272	890,160	879,523
Netherlands	4,509,759	5,051,093	6,207,956	8,035,272	11,272,876	9,663,002	14,215,918	17,136,242	10,130,234	9,320,390
Austria	11,385,155	13,578,525	12,466,646	5,266,850	7,183,131	8,300,971	14,238,871	16,319,834	9,790,662	6,858,899
Poland	4,270,122	4,920,050	4,576,351	5,079,877	6,556,418	7,769,788	12,019,819	11,812,597	7,343,738	5,808,791
Portugal	2,521,570	3,015,320	4,262,210	4,711,178	5,846,340	6,702,761	8,700,442	10,675,315	7,377,161	5,848,413
Romania	1,145,214	1,247,724	1,236,508	1,471,539	1,907,962	2,413,192	3,877,574	4,370,892	2,580,584	1,946,044
Slovenia	596,357	627,781	626,500	634,164	800,010	971,451	1,605,226	1,867,197	1,021,997	770,229
Slovakia	867,544	1,036,310	967,003	893,533	1,076,391	1,220,286	1,834,989	1,895,764	1,275,541	1,124,791
Finland	1,302,613	1,488,419	1,697,070	1,410,108	1,349,952	1,971,370	3,004,626	2,194,726	1,645,655	1,432,206
Sweden	2,703,712	3,107,771	3,481,674	3,405,790	4,706,470	5,909,790	12,029,687	8,144,318	4,091,601	3,578,840
United Kingdom	12,898,814	13,838,053	18,428,081	24,073,363	28,869,003	29,209,324	36,649,767	42,852,663	28,423,847	23,707,590
Iceland	282,558	390,927	475,160	405,175	501,029	1,020,481	1,527,095	1,394,927	717,254	509,900
Liechtenstein	9,532	14,631	14,320	7,598	8,877	11,536	14,519	16,243	11,403	8,736
Norway	1,614,820	2,059,580	2,200,618	1,985,848	2,467,670	4,015,145	6,205,464	4,644,917	2,671,799	2,027,388
Switzerland	3,749,646	5,256,196	4,531,561	2,875,816	3,503,474	4,108,776	6,614,193	6,464,341	4,516,754	3,639,951
Montenegro	84,784	85,652	98,770	159,185	453,314	1,201,434	3,144,240	4,038,338	1,514,083	261,162
Former Yugoslav Republic of I	72,572	69,349	91,450	116,329	170,554	167,950	248,452	257,962	186,383	144,655
Serbia	474,690	545,134	512,063	554,454	707,041	682,260	861,339	967,151	674,978	597,937
Turkey										

Available flags:
b break in time series | c confidential | d definition differs, see metadata
e estimated | f forecast | i see metadata (phased out)
n not significant | p provisional | r revised
s Eurostat estimate (phased out) | u low reliability | z not applicable

Special value:
: not available

eurostat

Occupancy of tourist accommodation establishments (tour_occ)

Reference Metadata in Euro SDMX Metadata Structure (ESMS)
Compiling agency: Eurostat, the statistical office of the European Union

Eurostat metadata

Reference metadata

1. Contact
2. Metadata update
3. Statistical presentation
4. Unit of measure
5. Reference Period
6. Institutional Mandate
7. Confidentiality
8. Release policy
9. Frequency of dissemination
10. Accessibility and clarity
11. Quality management
12. Relevance
13. Accuracy
14. Timeliness and punctuality
15. Coherence and comparability
16. Cost and Burden
17. Data revision
18. Statistical processing
19. Comment

Related Metadata
Annexes (including footnotes)

National metadata

National reference metadata
National metadata produced by countries and released by Eurostat

Belgium	Bulgaria	Czech Republic	Denmark
Germany	Estonia	Ireland	Spain
France	Croatia	Italy	Latvia
Lithuania	Luxembourg	Hungary	Malta
Netherlands	Austria	Poland	Portugal
Romania	Slovenia	Slovakia	Finland
Liechtenstein	Switzerland	Montenegro	Serbia

For any question on data and metadata, please contact: EUROPEAN STATISTICAL DATA SUPPORT

Download

1. Contact Top

1.1. Contact organisation	Eurostat, the statistical office of the European Union
1.2. Contact organisation unit	G3: Short-term Business Statistics and Tourism
1.5. Contact mail address	2920 Luxembourg LUXEMBOURG

2. Metadata update Top

2.1. Metadata last certified	08/06/2017
2.2. Metadata last posted	08/06/2017
2.3. Metadata last update	08/06/2017

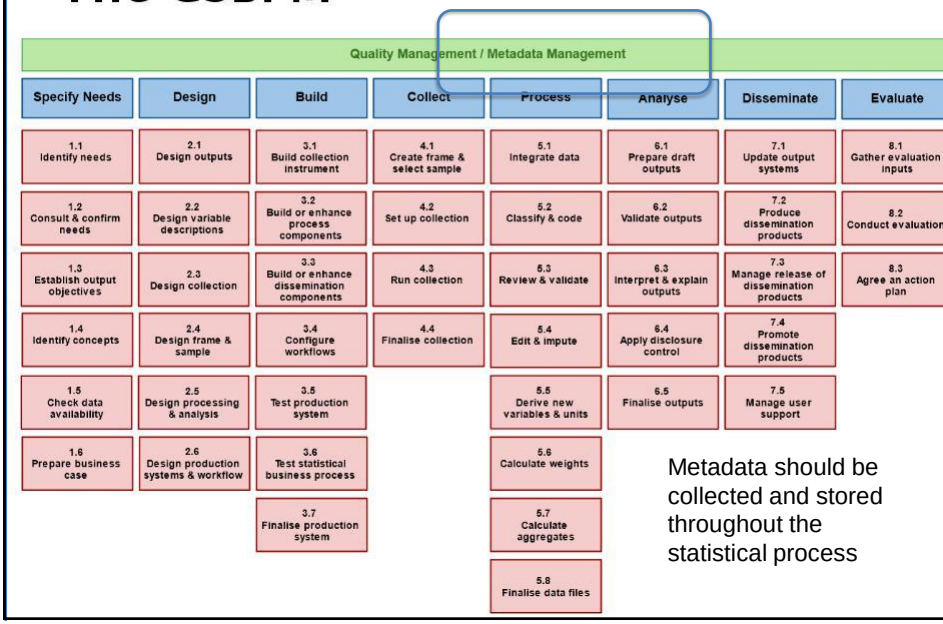
3. Statistical presentation Top

3.1. Data description

Accommodation statistics are a key part of the system of tourism statistics in the EU and have a long history of data collection. Annex I of the Regulation (EU) 692/2011 of the European Parliament and of the Council deals with accommodation statistics and includes 4 sections focusing on accommodation statistics of which sections 1 and 2 include the requirements concerning rented accommodation (capacity and occupancy respectively).
Data are collected by the competent national authorities of the Member States and are compiled according to a harmonised methodology established by EU regulations before transmission to Eurostat. Most of the time, data are collected via sample or census surveys. However, in a few cases data are compiled from a demand-side perspective (e.g. via visitor surveys or border surveys). Surveys on the occupancy of accommodation establishments are

http://ec.europa.eu/eurostat/cache/metadata/en/tour_occ_esms.htm

The GSBPM



Metadata should be collected and stored throughout the statistical process

Reference metadata and quality reporting

Information on output quality is included in the concept of reference metadata.

Information on methodologies applied during the statistical process are related to process quality.

Information on concepts are connected to relevance, comparability and coherence of statistics.

Quality reporting and reference metadata are strictly connected

Istat

Definition of quality reporting

Quality reporting is the preparation and dissemination, on a regular or irregular basis, of reports conveying information about the quality of a statistical product or survey (Eurostat, 2003b).

A quality report provides information on the *main quality characteristics* of a product so that the user should be able to assess product quality. In the optimal case quality reports are based on *quality indicators* (Bergdahl et al., 2007)

Istat

Introduction to quality reporting

Main features of a quality report:

- Scope: a statistical product or statistical process
- Target recipients: users or producers of official statistics
- Purposes: to accompany official statistics dissemination, to support quality assessment
- Contents: quality dimensions, quality indicators
- Further characteristics: standard structure, different levels of detail tailored on recipient and purpose

Istat

User oriented quality reporting

- To communicate quality to users
- To report on strengths and limits in order to support a proper use of statistics
- To improve transparency

Concise, not much technical and focussed on the statistical output

Istat

Producer oriented quality reporting

- To support management monitoring
- To identify points for further improvements
- To produce reference documentation for quality assessment (e.g.: audit and self-assessment)

Detailed, comprehensive, supplemented with quality indicators and focussed more on the statistical process, even if following product quality dimensions

Istat

The importance of quality reporting

Quality reporting underpins quality assessment, which in turn is the starting point for quality improvements. Thus, *standards and guidelines* for effective quality reporting are an essential aspect of the quality management / assurance framework.

Source: ESS Handbook for Quality reports

Istat

Approaches to quality reporting

- NSIs and ONAs can define their own strategy for quality reporting and, eventually, their own standards.
- Standardising Quality reporting improves comparability and clarity (set a “minimum” level of documentation)

Istat

Examples of quality reporting at NSI level

Office for National Statistics — Information on quality and methodology

DESTATIS Statistisches Bundesamt — Quality reports

INE Instituto Nacional de Estadística — Standardised Methodological Report

REPUBLIC OF BULGARIA National statistical institute — Metadata and Methodology

Istat

Approaches to quality reporting

- International Organisations often standardise their requests in terms of reference metadata and quality reporting
- The European Statistical System standard for reference metadata and quality reporting is the **Single Integrated Metadata Structure** (**SIMS**) and its derived reporting structures **Euro SDMX Metadata Structure** (**ESMS**) and **ESS Standard for Quality Report Structure** (**ESQRS**)
- The International Monetary Fund (IMF) presents the metadata according to the **Data Quality Assessment Framework** (DQAF).

Istat

ESS Quality reporting standards

European Statistical Law

Regulation (EU) 2015/759 of 29 April 2015 amending Regulation (EC) No 223/2009 on European statistics.

Art. 12 (3): "Member States shall provide the Commission (Eurostat) with reports on the quality of data transmitted…"

The **Single Integrated Metadata Structure (SIMS)** version 2.0 is approved as **ESS standard** by the ESSC

SIMS can be considered the **standard** for quality reporting according to Article 12 of Regulation 223/2009 on European statistics

Istat

The Single Integrated Metadata Structure (SIMS)

SIMS is a dynamic and unique inventory of ESS standard reference metadata statistical concepts with definitions and reporting guidelines.

S.15	Accuracy and reliability	ACCURACY	Accuracy: closeness of computations or estimates to the exact or true values that the statistics were intended to measure. Reliability: closeness of the initial estimated value to the subsequent estimated value.	
S.15.1	Overall accuracy	ACCURACY_OVERALL	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 error 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. It should reflect the producer's best current understanding (sign and order of magnitude) including actions taken to reduce bias. Revision aspects should also be included here if considered relevant.
S.15.2	Sampling error / A1. Sampling errors - indicators for U	SAMPLING_ERR / SAMPLING_ERR_IND_U	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.	If probability sampling is used, the range of variation, among key variables, of the A1 indicator should be reported. It should be also stated if adjustments for non-response, misclassifications and other uncertainty sources such as outlier treatment are included. The calculation of sampling error could be also affected by imputation. This should be noted unless special methods have been applied to deal with this. If non-probability sampling is used, the responsible for the statistical domain should provide estimates of the accuracy, a motivation for the invoked model for this estimation, and brief discussion of sampling bias. / QPI: A1 Sampling errors - indicators for U, with different LEVEL OF DETAILS for U and P
S.15.2.1	*A1. Sampling errors - indicators for P*	*SAMPLING_ERR_IND_P*	Precision measures for estimating the random variation of an estimator due to sampling.	QPI: A1 Sampling errors - indicators for P, with different LEVEL OF DETAILS for U and P
S.15.3	Non-sampling error and *A4. Unit non-response - rate for U* and *A5. Item non-response - rate for U*	NONSAMPLING_ERR / UNIT_NONRESPONSE_RATE_U / ITEM_NONRESPONSE_RATE_U	Error in survey estimates which cannot be attributed to sampling fluctuations.	U: Provide a user-oriented summary of the (preferably quantitative) assessment of the non-sampling errors, non-response rates and the bias risks which are associated with them (coverage error: over/ undercoverage and multiple listings; measurement error: survey instrument, respondent and interviewer effect where relevant; nonresponse error: level of unit (non)response including causes and measures for nonresponse, level of item nonresponse for key variables; processing error: data editing, coding and imputation error where relevant; model assumption error: specific models used in estimation) and actions undertaken to reduce the different types of errors. P: Not to be reported, information to be included in the sub-concepts S.15.3.1-S.15.3.5. / QPI: A4 Unit non-response - rate for U, with different LEVEL OF DETAILS for U and P / QPI: A5. Item non-response - rate for U, with different LEVEL OF DETAILS for U and P
S.15.3.1	Coverage error	*COVERAGE_ERR*	Divergence between the frame population and the target population.	Some information on the register or other frame source should be reported upon (this assists in understanding coverage errors and their effects): reference period, frequency and timing of frame updates, updating actions, eventual discrepancies between the units reported in the frame and the target population unit, references to other documents on frame quality and effects of frame deficiencies on the outputs. Provide an assessment, whenever possible quantitative, on overcoverage and multiple listings, and on the extent of undercoverage. Report also an evaluation of the bias risks associated with the latter. Describe actions taken for reduction of undercoverage and associated bias risks.

Istat

SIMS 2.0

Item No	Concept name
S.1	**Contact**
S.1.1	Contact organisation
S.1.2	Contact organisation unit
S.1.3	Contact name
S.1.4	Contact person function
S.1.5	Contact mail address
S.1.6	Contact email address
S.1.7	Contact phone number
S.1.8	Contact fax number
S.2	**Metadata update**
S.2.1	Metadata last certified
S.2.2	Metadata last posted
S.2.3	Metadata last update
S.3	**Statistical presentation**
S.3.1	Data description
S.3.2	Classification system
S.3.3	Sector coverage
S.3.4	Statistical concepts and definitions
S.3.5	Statistical unit
S.3.6	Statistical population
S.3.7	Reference area
S.3.8	Time coverage
S.3.9	Base period
S.4	**Unit of measure**
S.5	**Reference period**
S.6	**Institutional mandate**
S.6.1	Legal acts and other agreements
S.6.2	Data sharing
S.7	**Confidentiality**
S.7.1	Confidentiality - policy
S.7.2	Confidentiality - data treatment
S.8	**Release policy**
S.8.1	Release calendar
S.8.2	Release calendar access
S.8.3	User access
S.9	**Frequency of dissemination**
S.10	**Accessibility and clarity**
S.10.1	News release
S.10.2	Publications
S.10.3	On-line database

Item No	Concept name
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
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

Item No	Concept name
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**

Istat

From SIMS to ESMS and ESQRS

SIMS is an inventory of concepts, it is not implemented directly but by means of the two reporting structures:

- Euro SDMX Metadata Structure (ESMS) recommended for *user-oriented quality reporting*
- ESS Standard for Quality Reports Structure (ESQRS) recommended for *producer-oriented quality reporting*

Istat

ESMS 2.0

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

Item No	Concept Name
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**
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

Item No	Concept Name
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**

Example:
http://ec.europa.eu/eurostat/cache/metadata/en/prom_esms.htm

Istat

ESQRS 2.0

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

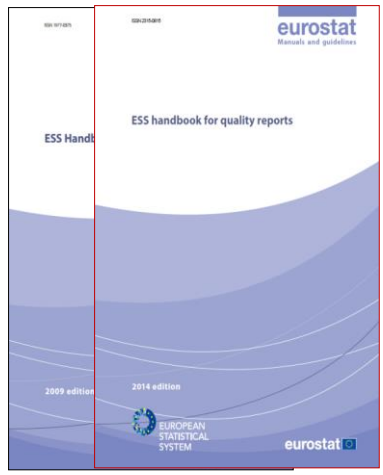
Item no	Concept Name
6	**Accuracy and reliability**
6.1	Accuracy - overall
6.2	Sampling error
6.2.1	Sampling error - indicators
6.3	Non-sampling error
6.3.1	Coverage error
6.3.1.1	Over-coverage - rate
6.3.1.2	Common units - proportion
6.3.2	Measurement error
6.3.3	Non response error
6.3.3.1	Unit non-response - 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

Item no	Concept Name
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**

Example: http://ec.europa.eu/eurostat/cache/metadata/EN/jvs_esqrs.htm

Istat

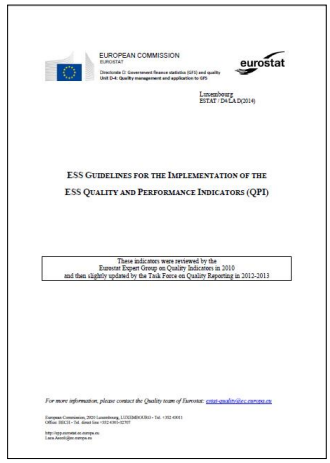
ESS Handbook for quality reports



The new version of the Handbook should be available in 2019.

Istat

ESS Quality and Performance Indicators


EUROPEAN COMMISSION
EUROSTAT

eurostat

Luxembourg
ESTAT / D4 LA D(2014)

ESS GUIDELINES FOR THE IMPLEMENTATION OF THE ESS QUALITY AND PERFORMANCE INDICATORS (QPI)

These indicators 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

For more information, please contact the Quality team of Eurostat: estat-quality@ec.europa.eu

- The "ESS Quality and Performance Indicators" represent the standard set of indicators shared at European level to be included in standard quality reports.
- Indeed, they are included in SIMS and in the derived reporting structures
- There are explicit "concepts" (fields, items) in ESQRS, while the indicators should be included in the related metadata concept/field/item in ESMS
- Some of them should be calculated at Eurostat level, others also at national level

Istat

IMF Special Data Dissemination Standard (SDDS)

SDDS purpose: to guide member countries in providing economic and financial data to the public

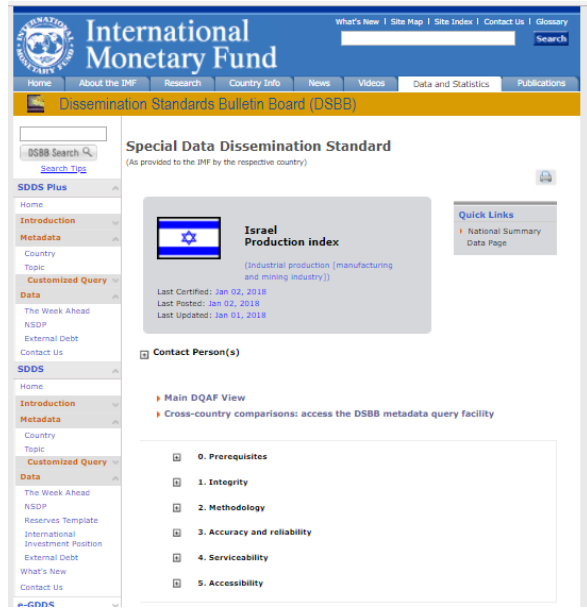
SDDS identifies four dimensions of data dissemination, for which best practice are provided that can be also monitored by users:

- Data coverage, periodicity (frequency), and timeliness;
- Access by the public;
- Integrity of the disseminated data;
- Quality of the disseminated data.

Metadata are requested together with data to SDDS subscribers. According to the DQAF template.

Istat

IMF metadata organised according to the DQAF



http://dsbb.imf.org/Pages/SDDS/DQAFBaseCollapsed.aspx?ctycode=ISR&catcode=IND003

Istat

Dimension	Element	Indicator
0.Prerequisites of quality	0.1Legal and institutional environment	0.1.1The responsibility for collecting, processing, and disseminating the statistics is clearly specified.
		0.1.2Data sharing and coordination among data-producing agencies are adequate.
		0.1.3Individual reporters' data are kept confidential and used for statistical purposes only.
		0.1.4Statistical reporting is ensured through legal mandate and/or measures to encourage response.
	0.2Resources	0.2.1Staff, facilities, computing resources, and financing are commensurate with statistical programs.
		0.2.2Measures to ensure efficient use of resources are implemented.
	0.3Relevance	0.3.1The relevance and practical utility of existing statistics in meeting users' needs are monitored.
	0.4Other quality management	0.4.1Processes are in place to focus on quality.
		0.4.2Processes are in place to monitor quality during the planning and implementation of the statistical program.
1.Assurances of integrity	1.1Institutional Integrity	1.1.1Statistics are produced on an impartial basis.
		1.1.2Choice of data sources and statistical techniques,as well as decisions about dissemination,are informed solely by statistical considerations.
		1.1.3The appropriate statistical entity is entitled to comment on erroneous interpretation and misuse of statistics.
	1.2Transparency	1.2.1The terms and conditions under which statistics are collected, processed, and disseminated are available to the public.
		1.2.2Internal governmental access to statistics prior to theirrelease is publicly identified.
		1.2.3Products of statistical agencies/unitsare clearly identified as such.
		1.2.4Advance notice is given of major changes in methodology, source data, and statistical techniques.
	1.3Ethical standards	1.3.1Guidelines for staff behavior are in place and are well known to the staff.
2.Methodological soundness	2.1Concepts and definitions	2.1.1The overall structure in terms of concepts and definitions follows internationally accepted standards, guidelines, or good practices.
	2.2Scope	2.2.1The scope is broadly consistent with internationally accepted standards, guidelines, or good practices.
	2.3Classification/sectorization	2.3.1Classification/sectorization systems used are broadly consistent with internationally accepted standards, guidelines, or good practices.
	2.4Basis for recording	2.4.1Market prices are used to value flows and stocks.
		2.4.2Recording is done on an accrual basis.
		2.4.3 Grossing/netting procedures are broadly consistent with internationally accepted standards, guidelines, or good practices.
3.Accuracy and reliability	3.1Source data	3.1.1Source data are obtained from comprehensive data collection programs that take into account country-specific conditions.
		3.1.2Source data reasonably approximate the definitions, scope, sectorization, classifications, valuation, and time of recording required.4
		3.1.3Source data are timely.
	3.2Assessment of source data	3.2.1Source data—including censuses, sample surveys,and administrative records—are routinely assessed, e.g., for coverage, sample error, response error, and nonsampling error; the results of the assessments are monitored and made available toguide statistical processes.
	3.3Statistical techniques	3.3.1Data compilation employs sound statistical techniques to deal with data sources.
		3.3.2Other statistical procedures (e.g., data adjustments and transformations, andstatistical analysis) employ sound statistical techniques.
	3.4Assessment and validation of intermediate data and statistical outputs	3.4.1Intermediate results are validated against other information,where applicable.
		3.4.2Statistical discrepancies in intermediate data are assessed and investigated.
		3.4.3Statistical discrepancies and other potential indicators of problems in statistical outputs are investigated.
	3.5Revision studies	3.5.1Studies and analyses of revisions and/or updates are carried out and used internally to inform statistical processes(see also 4.3.3.)
4.Serviceability	4.1Periodicity and timeliness	4.1.1Periodicity follows dissemination standards.
		4.1.2Timeliness follows dissemination standards.
	4.2Consistency	4.2.1Statistics are consistent within the dataset.
		4.2.2Statistics are consistentor reconcilable over a reasonable period of time.
		4.2.3Statistics are consistent or reconcilable with those obtained through other data sources and/or statistical frameworks.
	4.3Revision policy and practice	4.3.1Revisions and/or updates follow a regular and transparent schedule.
		4.3.2Preliminary and/or revised/updateddata are clearly identified.
		4.3.3Studies and analyses of revisions are made public (see also 3.5.1).
5.Accessibility	5.1Data accessibility	5.1.1Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons (layout and clarity of text, tables, and charts).
		5.1.2Dissemination media and format are adequate.
		5.1.3Statistics are released on a preannounced schedule.
		5.1.4Statistics are made available to all users at the same time.
		5.1.5Statistics not routinely disseminated are made available upon request.
	5.2Metadata accessibility	5.2.1Documentation on concepts, scope, classifications, basisof recording, data sources, and statistical techniques is available, and differences from internationally accepted standards, guidelines, or good practices are annotated.
		5.2.2Levels of detail are adapted to the needs of the intended audience.
	5.3Assistance to users	5.3.1Contact points are publicized.
		5.3.2Publications, documents, and other services, including information on any charges, are widely available

DQAF

ADDITIONAL SLIDES

Istat

Conceptual and methodological metadata

Item No	Concept name
S.1	Contact
S.1.1	Contact organisation
S.1.2	Contact organisation unit
S.1.3	Contact name
S.1.4	Contact person function
S.1.5	Contact mail address
S.1.6	Contact email address
S.1.7	Contact phone number
S.1.8	Contact fax number
S.2	Metadata update
S.2.1	Metadata last certified
S.2.2	Metadata last posted
S.2.3	Metadata last update
S.3	Statistical presentation
S.3.1	Data description
S.3.2	Classification system
S.3.3	Sector coverage
S.3.4	Statistical concepts and definitions
S.3.5	Statistical unit
S.3.6	Statistical population
S.3.7	Reference area
S.3.8	Time coverage
S.3.9	Base period
S.4	Unit of measure
S.5	Reference period
S.6	Institutional mandate
S.6.1	Legal acts and other agreements
S.6.2	Data sharing
S.7	Confidentiality
S.7.1	Confidentiality - policy
S.7.2	Confidentiality - data treatment
S.8	Release policy
S.8.1	Release calendar
S.8.2	Release calendar access
S.8.3	User access
S.9	Frequency of dissemination
S.10	Accessibility and clarity
S.10.1	News release
S.10.2	Publications
S.10.3	On-line database

Item No	Concept name
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
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

Item No	Concept name
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

Istat

Conceptual and methodological metadata

S.1	Contact
S.1.1	Contact organisation
S.1.2	Contact organisation unit
S.1.3	Contact name
S.1.4	Contact person function
S.1.5	Contact mail address
S.1.6	Contact email address
S.1.7	Contact phone number
S.1.8	Contact fax number
S.2	Metadata update
S.2.1	Metadata last certified
S.2.2	Metadata last posted
S.2.3	Metadata last update

«Demographic» information on the reference metadata provided

Istat

Conceptual and methodological metadata

S.3	**Statistical presentation**
S.3.1	Data description
S.3.2	Classification system
S.3.3	Sector coverage
S.3.4	Statistical concepts and definitions
S.3.5	Statistical unit
S.3.6	Statistical population
S.3.7	Reference area
S.3.8	Time coverage
S.3.9	Base period
S.4	**Unit of measure**
S.5	**Reference period**

Description of data to which QR is referred:
- Unit
- Variables
- Classifications
- Time and space reference

Close to «structural metadata»

Istat

Conceptual and methodological metadata

S.6	**Institutional mandate**
S.6.1	Legal acts and other agreements
S.6.2	Data sharing
S.7	**Confidentiality**
S.7.1	Confidentiality - policy
S.7.2	Confidentiality - data treatment
S.8	**Release policy**
S.8.1	Release calendar
S.8.2	Release calendar access
S.8.3	User access
S.9	**Frequency of dissemination**

- «User-oriented» information mainly at «institutional level»

Istat

Conceptual and methodological metadata

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

- Description of the statistical process
- Not particularly detailed
- Born to document Eurostat processes

Istat

Quality metadata

Item No	Concept name
S.1	**Contact**
S.1.1	Contact organisation
S.1.2	Contact organisation unit
S.1.3	Contact name
S.1.4	Contact person function
S.1.5	Contact mail address
S.1.6	Contact email address
S.1.7	Contact phone number
S.1.8	Contact fax number
S.2	**Metadata update**
S.2.1	Metadata last certified
S.2.2	Metadata last posted
S.2.3	Metadata last update
S.3	**Statistical presentation**
S.3.1	Data description
S.3.2	Classification system
S.3.3	Sector coverage
S.3.4	Statistical concepts and definitions
S.3.5	Statistical unit
S.3.6	Statistical population
S.3.7	Reference area
S.3.8	Time coverage
S.3.9	Base period
S.4	**Unit of measure**
S.5	**Reference period**
S.6	**Institutional mandate**
S.6.1	Legal acts and other agreements
S.6.2	Data sharing
S.7	**Confidentiality**
S.7.1	Confidentiality - policy
S.7.2	Confidentiality - data treatment
S.8	**Release policy**
S.8.1	Release calendar
S.8.2	Release calendar access
S.8.3	User access
S.9	**Frequency of dissemination**
S.10	Accessibility and clarity
S.10.1	News release
S.10.2	Publications
S.10.3	On-line database

Item No	Concept name
S.10.3.1	AC1. Data tables - consultations
S.10.4	Micro-data access
S.10.5	Other
S.10.6	Documentation on methodology
S.10.6.1	AC 3. Metadata completeness - rate
S.11	Quality management
S.11.1	Quality assurance
S.11.2	Quality assessment
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.13	Accuracy and reliability
S.13.1	Overall accuracy
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.13.3.3.2	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.14.2	Punctuality and TP3. Punctuality - delivery and publication for U
S.15	Coherence and comparability
S.15.1	Comparability - geographical
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

Item No	Concept name
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**

Quality metadata

S.11	Quality management
S.11.1	Quality assurance
S.11.2	Quality assessment

- «Introduction to quality metadata»
- Quality policy at Institutional level
- Overall quality assessment on the product/process to which the report is referred to

Istat

Quality metadata

11.2	Quality assessment
12	Relevance
12.1	User needs
12.2	User satisfaction
12.3	Completeness

- User/producer dialogue:
  - Who are the users, users met and unmet needs
  - User satisfaction surveys
- Completeness dimension included in Relevance

Istat

Quality metadata

6	Accuracy and reliability
6.1	Accuracy - overall
6.2	Sampling error
6.3	Non-sampling error
6.3.1	Coverage error
6.3.2	Measurement error
6.3.3	Non response error
6.3.4	Processing error
6.3.5	Model assumption error
6.4	Seasonal adjustment
6.5	Data revision - policy
6.6	Data revision - practice

- Accuracy: closeness between estimate and true value of a popolation parameter
- The difference is the error, that can be classified in sampling and non-sampling
- Main sources of non-sampling errors

- Reliability and data revision
- Revision policy at institutional level and at process level

Istat

Quality metadata

Item No	Concept Name
14	Timeliness and punctuality
14.1	Timeliness
14.2	Punctuality

- Apart from the indicators on timeliness and punctuality, a description of the time dedicated to the different phases ofthe production process can be provided.

Istat

Quality metadata

8	Coherence and comparability
8.1	Comparability - geographical
8.2	Comparability - over time
8.3	Coherence - cross domain
8.4	Coherence - sub annual and annual statistics
8.5	Coherence - National Accounts
8.6	Coherence - internal

- Space and time comparability
- Coherence with the more relevant domains
- Focus on the definitions of concepts and adopted tools and methods

Istat

Quality metadata

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

- Focus on the ways in which data are disseminated
- Links to other methodological documents and quality metadata

Istat

Quality metadata

S.16	Cost and burden

- Quality constraints!

Istat

ESS Quality and Performance Indicators

Relevance
- Data completeness - rate

Accuracy and reliability
- Sampling error - indicators
- Over-coverage - rate
- Common units - proportion
- Unit non-response - rate
- Item non-response - rate
- Imputation - rate
- Data revision - average size

Timeliness and punctuality
- Time lag - first results
- Time lag - final results
- Punctuality - delivery and publication

Comparability and coherence
- Asymmetry for mirror flows statistics - coefficient
- Length of comparable time series

Accessibility and clarity
- Data tables - consultations
- Metadata - consultations
- Metadata completeness - rate

Istat