ITEM 3.6

DRAFT COMMISSION IMPLEMENTING REGULATION ON QUALITY

MEETING OF EUROPEAN DIRECTORS OF SOCIAL STATISTICS

LUXEMBOURG 23-24 MAY 2019

BECH BUILDING - ROOM AMPERE
COMMISSION IMPLEMENTING REGULATION (EU) …/...

of XXX

implementing Regulation (EU) xxxx/xxxx of the European Parliament and of the Council establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples, as regards the modalities and the content of the quality reports

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Regulation (EU) XXX/YYYY of the European Parliament and of the Council of DD MMMM YYYY establishing a common framework for European statistics relating to persons and households, based on data at individual level collected from samples, and in particular Article 12(6) thereof,

Whereas:

1) Regulation (EU) XXX/YYYY establishes a common framework for European statistics relating to persons and households, based on data at individual level obtained from samples.

2) Pursuant Article 12(6) of Regulation (EU) XXX/YYYY, the Commission is empowered to adopt implementing acts specifying the modalities and content of the quality reports, including indications on the method for assessing compliance with precision requirements.

3) There is a need to define the general content of the quality reports common to the statistics covered by Regulation (EU) XXX/YYYY.

4) The measures set out in this Regulation are in accordance with the opinion of the European Statistical System Committee,

HAS ADOPTED THIS REGULATION:

Article 1

Subject matter
This Regulation lays down the modalities and the content of the quality reports, including indications of the methods for assessing compliance with precision requirements, on the data that Member States shall transmit to the Commission (Eurostat) in accordance with Regulation xxxx/xxxx.

**Article 2**

**Definitions**

For the purposes of this Regulation, the following operational definitions shall apply:

(a) 'proxy interview' means an interview with someone (e.g. parent, spouse) other than the person from whom information is being sought. There should be a set of rules specific to each survey that defines when proxy interviews can be accepted;

(b) 'non-response' means the failure of a sample survey to collect data for all data items in the survey questionnaire from all the population units designated for data collection. Non-response encompasses a wide variety of reasons for non-observation, such as "impossible to contact", "not at home", "unable to answer", "incapacity", "hard core refusal", "inaccessible" or "unreturned questionnaire". In the first two cases, a contact with the selected unit is never established;

(c) ‘sampling error’ means 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

(c') 'non-sampling error' means error in survey estimates which cannot be attributed to sampling fluctuations.

(d) 'unit non-response' means one type of non-response which occurs when no data are collected about a population unit designated for data collection;

(e) 'item non-response' means one type of non-response which occurs when no data are collected about a survey variable in relation to a designated population unit;

(f) 'substitution (of respondents)' means the replacement of an originally sampled unit by another unit. This might be an inter- or intra-household replacement;

(g) 'gross sample' also referred to as 'initial sample' means the set of population units initially selected from the sampling frame. The gross sample is made up of the eligible units (net sample and the non-response units) and the ineligible units (units in the sample not belonging to the target population);

(h) ‘eligible units’ means the set of population units selected from the sample frame and belonging to the target population;

(i) 'net' sample also referred to as 'achieved sample' means the set of population units selected from the sampling frame from which enough information have been obtained to include the unit in the survey estimations;

(j) 'imputation procedure' means a rule for entering a value for a specific data item where the response is missing, invalid or inconsistent.
Article 3
Quality reports

1. Member States shall submit to the Commission (Eurostat) quality reports, in respect to the data sets transmitted according to Regulation (EU) XXX/YYYY.

2. The quality reports shall contain quality-related data and metadata according to the quality criteria and statistical concepts laid down in the Annex to this Regulation. These reports shall specify also instances in which the quality criteria and the statistical concepts have not been complied with.

Article 4
Indications on the methods for assessing compliance with precision requirements

The Commission (Eurostat) shall assess the compliance with precision requirements of the data transmitted by the Member States based notably on:

a) the magnitude and frequency of the non-compliance and its impact on quality, in particular comparability, of the key indicators;

b) the possibility to correct the non-compliance immediately, and the fact that the Member States are effectively implementing the corrections;

c) the possibility to mitigate indirectly the non-compliance through in particular estimation techniques, and the fact that the Member States are implementing these mitigating measures;

d) the control that the Member States are exercising over the non-compliance, that can arise for reasons beyond the Member States control;

e) the persistence of the non-compliance over successive implementation of data collections;

f) the existence and effective implementation of a corrective action plan approved by the Commission (Eurostat), whose appreciation takes into account the duration needed to correct the non-compliance, in particular in case of panel data collections.

Article 5
Technical standards for the transmission of quality reports

1. For the purpose of supporting quality management and process documentation the transmission of quality reports shall be done by means of the technical standards put in place by the Commission (Eurostat).

2. To allow retrieving the data by electronic means, the data shall be provided to the Commission (Eurostat) through the single entry point.

Article 6
Entry into force

This Regulation shall enter into force on the twentieth day following that of its publication in the Official Journal of the European Union.

This Regulation shall be binding in its entirety and directly applicable in all Member States.
Done at Brussels,

For the Commission

The President
ANNEX

Quality criteria and statistical concepts

The quality report shall contain quality-related data and metadata according to the following quality criteria and statistical concepts.

In case that for a statistical operation a statistical concept is not relevant, the statistical concept would remain in the quality report with the indication ‘Not applicable’.

1. CONTACT:

*Individual or organisational contact points for the data or metadata, including information on how to reach the contact points.*

2. STATISTICAL PRESENTATION:

*Description of the disseminated data, which can be displayed to users as tables, graphs or maps.*

- Data description
- Classifications systems.
- Sector coverage.
- Statistical concepts and definitions, including the reference period.
- Statistical units.
- Statistical population.
- Reference area.
- Time coverage.

3. STATISTICAL PROCESSING:

*Operations performed on data to derive new information according to a given set of rules.*

- Source data and sampling frame.
- Sample design.
- Frequency of data collection.
- Data collection.
- Data validation.
- Data compilation.

4. QUALITY MANAGEMENT:

*Systems and frameworks in place within an organisation to manage the quality of statistical products and processes.*

- Quality assurance.
- Qualitative assessment.
5. RELEVANCE:

Degree to which statistical information meet current and potential needs of users such as fulfilment of user needs, user satisfaction and completeness of the data sent by the Member States.

- User needs.
- User satisfaction.
- Completeness.

6. ACCURACY AND RELIABILITY:

Accuracy: closeness of computations or estimates to the exact or true values that the statistics are intended to measure.

Reliability: closeness of the initial estimated value to the subsequent estimated value.

6.1. OVERALL ACCURACY

Assessment of accuracy, linked to a certain data set or domain, which is summarising the various components.

- Description of 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.

6.2. SAMPLING ERROR

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.

- Description of the methodology for calculating precision estimates.
- Precision measures of the estimates as detailed in the implementing acts specifying the technical specifications of the individual data sets.

6.3. NON-SAMPLING ERROR

Error in survey estimates which cannot be attributed to sampling fluctuations.

6.3.1. COVERAGE ERROR

Divergence between the frame population and the target population.

- Frequency and timing of frame updates.
- Errors due to the discrepancies between the sampling frame and the target population and sub-populations (over-coverage, under-coverage, misclassifications).

6.3.2. MEASUREMENT ERROR

Errors that occur during data collection and cause recorded values of variables to be different from the true ones.

- Description of efforts made in questionnaire design and testing.
- Description of interviewer training.
- Proxy interview rates.
6.3.3. **NON-RESPONSE ERROR**

Errors that occur when the survey fails to get a response to one, or possibly all, of the questions

- Characteristics of non-respondents.
- Unit and item non-response rates.
- Substitution rates.
- Gross sample size, number of eligible units and net sample size.

6.3.4. **PROCESSING ERROR**

The error in final data collection process results arising from the faulty implementation of correctly planned implementation methods.

- Description of quality checks and of data editing process.
- Description of imputation procedures.
- Imputation rates.

6.3.5. **MODEL ASSUMPTION ERROR**

Error due to domain specific models needed to define the target of estimation.

6.4. **SEASONAL ADJUSTMENT**

Statistical technique used to remove the effects of seasonal calendar influences operating on a series.

6.5. **DATA REVISION – POLICY**

Policy aimed at ensuring the transparency of disseminated data, whereby preliminary data are compiled that are later revised.

6.6. **DATA REVISION – PRACTICE**

Information on the data revision practice.

7. **TIMELINESS AND PUNCTUALITY:**

**Timeliness**: Length of time between data availability and the event or phenomenon the data describe.

**Punctuality**: Time lag between the actual delivery of the data and the target date when it should have been delivered.

- Date of the dissemination of national results.

8. **COHERENCE AND COMPARABILITY:**

**Coherence**: Adequacy of statistics to be reliably combined in different ways and for various uses

**Comparability**: Measurement of the impact of differences in applied statistical concepts, measurement tools and procedures where statistics are compared between geographical areas or over time.

- Comparability – geographical, including in particular how the requirements established in the domain-specific Regulation have been
fulfilled, including, when relevant, deviations from the European questionnaire and definitions.

- Comparability – over time
- Coherence – cross domain
- Coherence - internal

9. ACCESSIBILITY AND CLARITY:
Simplicity and ease, the conditions and modalities by which users can access, use and interpret statistics, with the appropriate sup por ting information and assistance.

- Dissemination formats.
- Documentation on methodology and quality.

10. COST AND BURDEN:
Cost associated with the collection and production of the statistical product and burden on respondents.

- Average time for answering the survey; if possible and relevant by data collection mode.

11. 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.

- Confidentiality policy
- Confidentiality – data treatment

12. COMMENT:
Supplementary descriptive text that can be attached to the quality report.