**SCB-FS 2016:17**

 Printed

 28 September 2016

**Statistics Sweden’s regulations on quality in official statistics;**

decided on 14 June 2016.

Statistics Sweden prescribes the following with the support of Section 16, item 2 of the Official Statistics Ordinance (2001:100).

**Area of application**

**Section 1** These regulations contain provisions on quality in official statistics according to Section 3a of the Official Statistics Act (2001:99).

**Section 2** It follows from Section 3a of the Official Statistics Act (2001:99) that, in order to assure the quality of statistics with regard to its development, production, and dissemination, the quality criteria relevance, accuracy, timeliness, punctuality, accessibility and clarity, comparability, and coherence shall be applied.

**The quality concept and its main, sub-, and sub-subcomponents**

**Section 3**  In the development and production of official statistics and in its dissemination, including quality declarations according to Statistics Sweden’s regulations and general guidelines for the official release, etc. of official statistics (SCB-FS 2002:16), the quality of the statistics shall be described based on a quality concept consisting of five main components. These are linked to the quality criteria in Section 3a of the Official Statistics Act (2001:99).

The relation between the quality criteria and the quality concept’s main components should be as follows.

|  |  |
| --- | --- |
| **Quality criterion** | **Main component** |
| Relevance | *Relevance* |
| Accuracy | *Accuracy* |
| Timeliness | *Timeliness and punctuality* |
| Punctuality |
| Accessibility and clarity | *Accessibility and clarity* |
| Comparability | *Comparability and coherence* |
| Coherence |

The main components shall in turn be divided into subcomponents and, where applicable, into sub-subcomponents, according to the following.

|  |
| --- |
| **Main component: *Relevance*** |
| **Subcomponent** | **Sub-subcomponent** | **Described in** |
| *Purpose and information needs* |  | Section 7 |
|  | *Purpose of the statistics* | Section 9 |
|  | *User information needs* | Section 10 |
| *Content of the statistics* |  | Section 11 |
|  | *Unit and population* | Section 13 |
|  | *Variables* | Section 14 |
|  | *Statistical measures* | Section 15 |
|  | *Study domains* | Section 16 |
|  | *Reference times* | Section 17 |

|  |
| --- |
| **Main component: *Accuracy*** |
| **Subcomponent** | **Sub-subcomponent** | **Described in**  |
| *Overall accuracy* |  | Section 20 |
| *Sources of uncertainty* |  | Section 21 |
|  | *Sampling* | Section 23 |
|  | *Frame coverage* | Section 24 |
|  | *Measurement* | Section 25 |
|  | *Non-response* | Section 26 |
|  | *Data processing* | Section 27 |
|  | *Model assumptions* | Section 28 |
| *Preliminary statistics compared with final statistics* |  | Section 29 |

|  |
| --- |
| **Main component: *Timeliness and punctuality*** |
| **Subcomponent** | **Sub-subcomponent** | **Described in**  |
| *Production time* |  | Section 32 |
| *Frequency* |  | Section 33 |
| *Punctuality* |  | Section 34 |

|  |
| --- |
| **Main component: *Accessibility and clarity*** |
| **Subcomponent** | **Sub-subcomponent** | **Described in**  |
| *Access to the statistics*  |  | Section 37 |
| *Possibility of additional statistics* |  | Section 38 |
| *Presentation*  |  | Section 39 |
| *Documentation* |  | Section 40 |

|  |
| --- |
| **Main component: *Comparability and coherence***  |
| **Subcomponent** | **Sub-subcomponent** | **Described in**  |
| *Comparability over time* |  | Section 43 |
| *Comparability among groups* |  | Section 44 |
| *Other coherence* |  | Section 45 |
| *Numerical consistency* |  | Section 46 |

**Definitions**

**Section 4**  In this regulation, the following terms are used with the meanings stated below.

|  |  |
| --- | --- |
| **Term** | **Meaning** |
| Statistical measure | A computational rule used to summarise variable values. Such a rule may involve the counting of numbers or the computation of a sum, a mean value, or a median. |
| Statistical characteristic  | The value received by summarising individual variable values for units in a population or subpopulation using a statistical measure. Both units and variables are associated with a specific time. Units and variables are often, but not always, associated with the same point or period of time. |
| Estimation procedure | A computational algorithm used to estimate a statistical characteristic. |
| Statistical value | The estimated value of a statistical characteristic. |
| Characteristic of interest | The statistical characteristic associated with a need. Units, populations, and variables included in characteristics of interest are referred to as units of interest, populations of interest, and variables of interest. |
| Target characteristic | The statistical characteristic that the statistical agency has decided to estimate. Units, populations, and variables included in target characteristics are referred to as target units, target populations, and target variables. |
| Observation unit | A unit about which data is collected. |
| Observation variable | A variable for which values are collected. |
| Data source | A source – such as an individual, a register or a database – from which variable values are collected. |
| Frame procedure | The procedure leading to the determination of observation units, data sources, frame, and means of contacting data sources.Observation units and data sources may coincide, but not necessarily. |

**Main component: *Relevance***

**General provisions**

**Section 5** Relevance refers to how well statistics illuminate the issues that are of importance for users of the statistics. The agreement between target characteristics and characteristics of interest is an important element in the assessment of the relevance of statistics. Before the statistical agency defines the target characteristics to be estimated, the agency makes considerations regarding user information needs, quality requirements associated with the statistics, the cost of producing the statistics, and the response burden.

 **Section 6**  Section 3 sets out that the main component *Relevance* consists of the subcomponents *Purpose and information needs* and *Content of the statistics.*

**The subcomponent *Purpose and information needs* and its associated sub-subcomponents**

**Section 7**  The subcomponent *Purpose and information needs* refers to the statistical agency’s overarching considerations for the statistics. It involves the purpose, the information needs that the statistics are intended to meet, and the knowledge about the needs for statistical information.

 **Section 8** Section 3 sets out that the subcomponent *Purpose and information needs* consists of the sub-subcomponents *Purpose of the statistics* and *User information needs*.

 **Section 9**  *Purpose of the statistics* refers to the primary objective and area of use of the disseminated statistics.

 **Section 10** *User information needs* refers to the statistical agency’s knowledge about the need for statistics as well as main users and uses.

**The subcomponent *Content of the statistics* with associated sub-subcomponents**

**Section 11**  The subcomponent *Content of the statistics* refers to the target characteristics.

 **Section 12**  Section 3 sets out that the subcomponent consists of the sub-subcomponents *Unit and population, Variables*, *Statistical measures, Study domains,* and *Reference times.*

**Section 13**  *Unit and population* refers to the target population, with its associated target units and delineation. The sub-subcomponent also includes the relations between the target population and the populations of interest as well as between the target units and observation units.

 **Section 14**  *Variables* refers to the target variables and their relations to the variables of interest and the observation variables.

 **Section 15**  *Statistical measures* refers to the statistical measures used in the statistical target characteristics.

 **Section 16**  *Study domains* refers to the divisions or breakdowns into subpopulations used in the statistical target characteristics.

**Section 17**  *Reference times* refers to the points in time or periods of time associated with the statistical target characteristics.

**Main component: *Accuracy***

**General provisions**

**Section 18** Accuracy refers to how well a statistical value estimates its target characteristic, i.e. the expected deviation between the statistical value and the target characteristic. Accuracy is largely determined by the estimation procedure. The total deviation may be regarded as the total of all partial deviations, whether positive or negative, caused by various sources of uncertainty. Partial deviations may interact in various ways. Statistics may be disseminated using preliminary statistical values, once or more, followed by final statistical values disseminated at a later stage. Preliminary values are generally associated with greater uncertainty, which affects the total deviation from the target characteristics.

 **Section 19**  Section 3 sets out that the main component *Accuracy* consists of the subcomponents *Overall accuracy*, *Sources of uncertainty,* and *Preliminary statistics compared with final statistics.*

**The subcomponent *Overall accuracy***

**Section 20** *Overall accuracy* refers to the expected deviation of the statistical value from the target characteristic.

**The subcomponent *Sources of uncertainty with associated sub-subcomponents***

**Section 21** *Sources of uncertainty* refers to each source of uncertainty and its impact on the statistics.

**Section 22** Section 3 sets out that the subcomponent consists of the sub-subcomponents *Sampling*, *Frame coverage*, *Measurement*, *Non-response*, *Data processing,* and *Model assumptions*.

**Section 23** *Sampling* refers to the element of uncertainty caused by the fact that the statistics are based on observations of a sample of observation units.

**Section 24** *Frame coverage* refers to the element of uncertainty caused by differences between the frame population, the set of units identified in the frame procedure, and the target population.

 **Section 25** *Measurement* refers to the element of uncertainty caused by differences between observed variable values and true variable values.

 **Section 26** *Non-response* refers to the element of uncertainty that arises when all variable values are missing for some of the planned observation units, referred to as unit non-response, or when individual variable values are missing for certain observation units, referred to as item non-response.

 **Section 27** *Data processing* refers to the element of uncertainty caused by data processing deficiencies.

**Section 28** *Model assumptions* refers to the element of uncertainty caused when assumptions in the models on which the statistics are wholly or partially based have not been fully met.

**The subcomponent *Preliminary statistics compared with final statistics***

**Section 29** *Preliminary statistics compared with final statistics* refers to information on the size and direction of the revisions made to the preliminary statistics before they become final.

**Main component: *Timeliness and punctuality***

**General provisions**

**Section 30** Timeliness and punctuality refers to the time lag and frequency required of recurring statistics for the latest available statistics to be regarded as current.

 **Section 31** According to Section 3, the main component *Timeliness and punctuality* consists of the subcomponents *Production time, Frequency,* and *Punctuality*.

**The subcomponent *Production time***

**Section 32** The *Production time* of statistics refers to the time lag between the end of the reference time and the point in time when the statistics are disseminated.

**The subcomponent *Frequency***

**Section 33** The *Frequency* of recurring statistics refers to the periodicity of the reference time of the statistics, the survey’s data collection, and the dissemination of the statistics. For recurring surveys, the time lag from the time of use to the last available reference time is not only affected by the production time but also by these periodicities.

**The subcomponent *Punctuality***

**Section 34** *Punctuality* refers to how well the actual point in time when statistics are made accessible corresponds to the pre-announced point in time.

**Main component: *Accessibility and clarity***

**General provisions**

**Section 35** Accessibility and clarity refers to the possibility of accessing statistics that have been or can be produced as well as information on the properties of the statistics.

**Section 36** Section 3 sets out that the main component *Accessibility and clarity* consists of the subcomponents *Access to the statistics*, *Possibility of additional statistic*s, *Presentation,* and *Documentation*.

**The subcomponent *Access to the statistics***

**Section 37** *Access to statistics* refers to the distribution channels used to relay statistics to users.

**The subcomponent *Possibility of additional statistics***

**Section 38** *Possibility of additional statistics* refers to the extent to which users may obtain access to statistics that have not been produced but can be produced based on existing data.

**The subcomponent *Presentation***

**Section 39** *Presentation* refers to how statistical information is presented and visualised, such as in texts, tables, and diagrams.

**The subcomponent *Documentation***

**Section 40** *Documentation* refers to the users’ possibilities to access information to facilitate the understanding and interpretation of statistics.

**Main component: *Comparability and coherence***

**General provisions**

**Section 41** Comparability refers to the possibility of making comparisons between various reference times or different groups. Coherence refers to the possibility of combining and jointly analysing various statistical values. In both cases, the possibility is governed by how similarly or differently the target characteristics are defined and by differences in survey methods. Comparability is a specific aspect of coherence.

**Section 42** Section 3 sets out that the main component *Comparability and coherence* consists of the subcomponents *Comparability over time*, *Comparability among groups*, *Other coherence,* and *Numerical consistency*.

**The subcomponent *Comparability over time***

**Section 43** *Comparability over time* refers to the extent to which statistical values related to different reference times can be compared. The statistical values may come from one or several surveys.

**The subcomponent *Comparability among groups***

**Section 44** *Comparability among groups* refers to the extent to which statistical values related to different groups can be compared. The statistical values may come from one or several surveys.

**The subcomponent *Other coherence***

**Section 45** Other coherence refers to the extent to which statistical values can be combined for other purposes than comparisons over time and among groups. The statistical values may come from one or several surveys.
 **The subcomponent *Numerical consistency***

**Section 46** Numerical consistency refers to whether statistical values relate to each other numerically in a way that is logically warranted by the subject matters that the statistics are intended to illuminate. The statistical values may come from one or several surveys.

This statute shall enter into force on 1 January 2017.

On behalf of Statistics Sweden

Stefan Lundgren

 Eva Nilsson