
**Statistical Working Group
Statistical Guidelines**

**CL\_OBS\_STATUS
Code list for Observation Status**

**Version 2.1 - 15/9/2016**

#### CL\_OBS\_STATUS

**Name**: Code list for Observation Status.

**Description**: This code list provides coded information about the "status" of an observation (with respect to events such as the ones reflected in the codes composing the code list). The observation status provides information on 1) the quality of a value or 2) unusual or missing values.

The codes in this list are known to be heterogeneous, that is they represent different concepts. They are put together in one code list for practical and historical implementation reasons.

The concept to be used for CL\_OBS\_STATUS is OBS\_STATUS.

**See also**

* "Possible ways of implementing CL\_OBS\_STATUS"
* "Guidelines for the Creation and Management of SDMX Cross-Domain Code Lists"[[1]](#footnote-1)

**Established international standard(s) used as input for the code list**: None.

**Version**: 2.1, 15 September 2016

|  |  |
| --- | --- |
| **Recommended code values descriptions** | **Annotation** |
| **A** | Normal value | To be used as default value if no value is provided or when no special coded qualification is assumed. Usually, it can be assumed that the source agency assigns sufficient confidence to the provided observation and/or the value is not expected to be dramatically revised. |
| **B** | Time series break | Observations are characterised as such when different content exists or a different methodology has been applied to this observation as compared with the preceding one (the one given for the previous period). |
| **D** | Definition differs | Used to indicate slight deviations from the established methodology (footnote-type information); these divergences do not imply a break in time series.  |
| **E** | Estimated value | Observation obtained through an estimation methodology (e.g. to produce back-casts) or based on the use of a limited amount of data or ad hoc sampling and through additional calculations (e.g. to produce a value at an early stage of the production stage while not all data are available). It may also be used in case of experimental data (e.g. in the context of a pilot ahead of a full scale production process) or in case of data of (anticipated/assessed) low quality. If needed, additional information can be provided through free text using the COMMENT\_OBS attribute at the observation level or at a higher level. |
| **F** | Forecast value | Value deemed to assess the magnitude which a quantity will assume at some future point of time (as distinct from "estimated value" which attempts to assess the magnitude of an already existent quantity). |
| **G** | Experimental value | Data collected on the basis of definitions or (alternative) collection methods under development. Data not of guaranteed quality as normally expected from provider. |
| **I** | Imputed value (CCSA definition) | Observation imputed by international organisations to replace or fill gaps in national data series, in line with the recommendations of the United Nations Committee for the Coordination of Statistical Activities (CCSA). |
| **K** | Data included in another category | This code is used when data for a given category are missing and are included in another category, sub-total or total. Generally where code “K” is used there should be a corresponding code "W - Includes data from another category" assigned to the over-covered category. Implementers and data reporters should use the COMMENT\_OBS observation-level attribute to specify under which category the data are included.  |
| **W** | Includes data from another category | This code is used when data include another category, or go beyond the scope of the data collection and are therefore over-covered. Generally, where code "W" is used there should be a corresponding code "K - Data included in another category" assigned to the category which is under-covered. Implementers and data reporters should use the COMMENT\_OBS observation-level attribute to specify which additional data are included. |
| **O** | Missing value | This code is to be used when no breakdown is made between the reasons why data are missing. Data can be missing due to many reasons: data cannot exist, data exist but are not collected (e.g. because they are below a certain threshold or subject to a derogation clause), data are unreliable, etc. |
| **M** | Missing value; data cannot exist | Used to denote empty cells resulting from the impossibility to collect a statistical value (e.g. a particular education level or type of institution may be not applicable to a given country's education system).[[2]](#footnote-2) |
| **P** | Provisional value | An observation is characterised as "provisional" when the source agency – while it bases its calculations on its standard production methodology – considers that the data, almost certainly, are expected to be revised. |
| **S** | Strike and other special events | Special circumstances (e.g. strike) affecting the observation or causing a missing value. |
| **L** | Missing value; data exist but were not collected | Used, for example, when some data are not reported/disseminated because they are below a certain threshold. |
| **H** | Missing value; holiday or weekend | Used in some daily data flows. |
| **Q** | Missing value; suppressed | Used, for example, when data are suppressed due to statistical confidentiality considerations. |
| **J** | Derogation | Clause in an agreement (e.g. legal act, gentlemen’s agreement) stating that some provisions in the agreement are not to be implemented by designated parties; these derogations may affect the observation or cause a missing value. In general, derogations are limited in time. |
| **N** | Not significant | Used to indicate a value which is not a "real" zero (e.g. a result of 0.0004 rounded to zero). |
| **U** | Low reliability | This indicates existing observations, but for which the user should also be aware of the low quality assigned. |
| **V** | Unvalidated value | Observation as received from the respondent without further evaluation of data quality. |

**Default value**

 In some implementations the concept OBS\_STATUS may not have been defined as "mandatory" within the data structure definition (DSD). If this is the case, each single observation may not be necessarily accompanied by an OBS\_STATUS code value and, if no value has been associated to an observation, the default code value ("A" - normal) can be assumed.

**Changes compared to version 2.0**

The rationale for the revision of this code list was the request from international organizations to add two codes which are necessary for their data production processes indicating that:

* data are included in another category (under-coverage)
* a category includes data from another category (over-coverage)

The detail of the changes from version 2.0 to version 2.1 is presented below:

|  |  |  |
| --- | --- | --- |
| **K** | Data included in another category | This code is used when data for a given category are missing and are included in another category, sub-total or total. Generally where code “K” is used there should be a corresponding code "W - Includes data from another category" assigned to the over-covered category. Implementers and data reporters should use the COMMENT\_OBS observation-level attribute to specify under which category the data are included.  |
| **W** | Includes data from another category | This code is used when data include another category, or go beyond the scope of the data collection and are therefore over-covered. Generally, where code "W" is used there should be a corresponding code "K - Data included in another category" assigned to the category which is under-covered. Implementers and data reporters should use the COMMENT\_OBS observation-level attribute to specify which additional data are included. |

A very minor change was also carried out under code "E":

|  |  |  |
| --- | --- | --- |
| **E** | Estimated value | Observation obtained through an estimation methodology (e.g. to produce back-casts) or based on the use of a limited amount of data or ad hoc sampling and through additional calculations (e.g. to produce a value at an early stage of the production stage while not all data are available). It may also be used in case of experimental data (e.g. in the context of a pilot ahead of a full scale production process) or in case of data of (anticipated/assessed) low quality. If needed, additional information can be provided through free text using the COMMENT\_OBS attribute at the observation level or at a higher level. |

1. Documents available from: https://sdmx.org/?page\_id=4345 [↑](#footnote-ref-1)
2. This concept is sometimes referred to as "not applicable". However, it is important to note that "not applicable" as meaning "data cannot exist" is different from the concept of "not applicable" (represented with code "\_Z") as described in the list of SDMX generic codes presented in the "Guidelines for the Creation and Management of SDMX Cross-Domain Code Lists". Code "M – Missing value; data cannot exist" as mentioned here is used to characterise the impossibility for a statistical value to exist. As such it is a quality indicator used as an attribute to a statistical value. \_Z is a residual category in a code list to which statistical values are attached. [↑](#footnote-ref-2)