sdmx
Statistical Data and Metadata eXchange

SDMX CONTENT-ORIENTED GUIDELINES

ANNEX 4:

METADATA COMMON VOCABULARY

2009

METADATA COMMON VOCABULARY

Table of Contents

METADATA COMMON VOCABULARY (2009)

1. Access

Definition: The means and conditions under which data can be viewed or used.

Context: Access also refers to terms, copyright issues and confidentiality constraints related to how the data can be used. A particular case is access to microdata, which can be defined as the process of providing users which are external to the data provider with access to unit level data, usually under strictly controlled conditions.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Accessibility
Confidentiality
Data
Data provider
Microdata

2. Accessibility

Definition: The ease and conditions under which statistical information can be obtained.

Context: Accessibility refers to the availability of statistical information to the user. It includes the ease with which the existence of information can be ascertained, as well as 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.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Access
Clarity
Dissemination format
Integrity
Simultaneous release

3. Accounting conventions

Definition: The practical procedures, standards and other aspects used when compiling data from diverse sources under a common methodological framework.

Context: In the SDDS/DQAF framework, accounting conventions refer to descriptions of the types of prices used to value flows and stocks, or other units of measurements used for recording the phenomena being observed; the time

of recording of the flows and stocks or the time of recording of other phenomena that are measured, including the reference period employed; and the grossing/netting procedures that are used.

Accounting conventions may refer to whether the data are recorded on a cash/accrual or mixed accounting basis, the time of their recording and the reference period (fiscal or calendar year) employed. The description could also include how consistent the practices used are with internationally accepted standards - such as the Balance of Payments 5th Manual or SNA93 - or good practices.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Reference period
Special Data Dissemination Standard, SDDS
Time of recording
Valuation

4. Accuracy

Definition: Closeness of computations or estimates to the exact or true values that the statistics were intended to measure.

Context: The accuracy of statistical information is the degree to which the information correctly describes the phenomena. It is usually characterized in terms of error in statistical estimates and is often decomposed into bias (systematic error) and variance (random error) components. Accuracy can contain either measures of accuracy (numerical results of the methods for assessing the accuracy of data) or qualitative assessment indicators. It may also be described in terms of the major sources of error that potentially cause inaccuracy (e.g., coverage, sampling, non response, response error). Accuracy is associated with the "reliability" of the data, which is defined as the closeness of the initial estimated value to the subsequent estimated value.

This concept can be broken down into: Accuracy - overall (summary assessment); Accuracy - non-sampling error; Accuracy - sampling error.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Bias
Error of estimation
Precision
Quality
Reliability
Statistical error

5. Adjustment

Definition: 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.

Context: Adjustments may be associated with changes in definitions, exchange rates, prices, seasons and other factors. Adjustments are in particular applied to compile consistent time series, but the concept is also used for describing adjustments related to other types of data.

Adjustment can be distinguished from editing and imputation, in that before adjustment, the data are already of sufficient quality to be considered usable.

"Adjustment - coded" refers to the type of adjustment used, represented by a code, while "Adjustment - detail" refers to the textual description of the type of adjustment used.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Compilation practices
Seasonal adjustment
Special Data Dissemination Standard, SDDS

6. Administered item

Definition: Registry item for which administrative information is recorded in an administration record.

Context: In the ISO/IEC 11179 standard:

Administered item "classification" is the relationship where an Administered Item is classified based on a specified Classification Scheme.

Administered item "context" is the relationship that provides a Context for an Administered Item.

Administered item "identifier" is an identifier for an administered item.

"Administrative note" is any general note about the Administered item.

"Origin" is the source (document, project, discipline or model) for the Administered item.

(ISO/IEC International Standard 11179-3 "Information technology - Metadata registries - Part 3: Registry metamodel and basic attributes", February 2003)

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Administration record
Administrative data
Context
Creation date
Date of last change
ISO/IEC 11179
Metadata registry
Origin
Originator data identifier
Registration
Registry item

7. Administration record

Definition: Collection of administrative information for an administered item.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Administered item
ISO/IEC 11179

8. Administrative data

Definition: The set of units and data derived from an administrative source.

Context: A set of administrative data that can be searched by unit identifiers can be referred to as an administrative register.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administered item
Administrative source
Data

9. Administrative source

Definition: A data holding containing information collected and maintained for the purpose of implementing one or more administrative regulations.

Context: A wider definition of administrative sources, is used in the Eurostat Business Registers Recommendations Manual: a data holding containing information which is not primarily collected for statistical purposes.

The organisational unit responsible for maintaining one or more administrative sources is known as an administrative organisation.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administrative data
Data

10. Age

Definition: The length of time that a person has lived or a thing has existed.

Context: Age can be expressed as a number, e.g. 25 years old, or as a range, e.g. "between 25 and 29 years" or "6 to 11 months".

It is in general a coded concept.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

11. Agency

See "Organisation"

12. Aggregated data

Definition: The result of transforming unit level data into quantitative measures for a set of characteristics of a population.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Characteristic
Measure

13. Aggregation

Definition: A process that transforms microdata into aggregate-level information by using an aggregation function such as count, sum average, standard deviation, etc.

Context: Aggregation denotes the compounding of primary data into an aggregate, usually for the purpose of expressing them in a summary form. For example, national income and price index numbers are aggregative, in contrast to the income of an individual or the price of a single commodity.

With standard hierarchical classifications, statistics for related categories can be grouped or collated (aggregated) to provide a broader picture, or categories can be split (disaggregated) when finer details are required and made possible by the codes given to primary observations (United Nations Glossary of Classification Terms, http://unstats.un.org/unsd/class/family/glossary_short.htm).

A special case of aggregation is the grossing up of samples.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Compilation practices
Data set

Disaggregation
Estimation
Frequency
Microdata
Statistical processing

14. Analytical framework

Definition: The system of criteria and standards within which data are analysed.

Context: In the context of SDDS, "analytical framework" consists of the following components:

1. International / Supranational guidelines: identifies the standardized system of definitions and classifications applied for analytical purposes with regard to the conceptual framework and organization of the related transactions data. (e.g., national accounts System of National Accounts, U.N. 1993; government operations A Manual on Government Finance Statistics, IMF, 2001, balance of payments Balance of Payments Manual, IMF, 1993).

2. Specificities of national practice: describes how concepts, definitions, and classifications for the national data aggregates disseminated deviate from those contained in relevant international or regional standards and/or guidelines.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Classification
Concept
Definition
Special Data Dissemination Standard, SDDS

15. Analytical unit

Definition: Real or artificially constructed units, for which statistics are compiled.

Context: Analytical units are created by statisticians, often by splitting or combining observation units with the help of estimations and imputations in order to compile more detailed and more homogeneous statistics than is possible using data on observation units (United Nations, Introduction to ISIC Rev. 3 (International Standard Industrial Classification of All Economic Activities, Revision 3), para. 63). Analytical units can correspond therefore for example to enterprises, local units, kind-of-activity units (KAU), local kind-of-activity units (local KAU) as well as to units of homogeneous production (UHP) and local units of homogeneous production (local UHP).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Classification
Observation unit
Statistical unit

16. Area sampling

Definition: A method of sampling used when no complete frame of reference is available.

Context: In area sampling, the total area under investigation is divided into small sub-areas which are sampled at random or by some restricted random process. Each of the chosen sub-areas is then fully inspected and enumerated, and may form a frame for further sampling if desired.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Frame
Sampling

17. Attachment level

Definition: A property of attributes defining the object to which data or metadata are linked.

Context: In SDMX (and previously in Gesmes) for each attribute specified in a data structure, there is a definition of whether this attribute takes:

- a value for each observation in the data set

- a value for each time series in the data set

- a value for each sibling group in the data set

- a single value for the entire data set.

Some metadata concepts (e.g. frequency) may not be meaningful at the observation level, but only when applied to a higher level (e.g. to a time series of observations). Time, on the other hand, is meaningful at observation level, because every observation is associated with a specific point or period in time. Data structure definitions and Metadata structure definitions provide information about the level at which a particular concept descriptor is relevant: at observation level, time-series level, group level, dataset level or even agency level. This is known as the "attachment level" of the concept.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Attribute
GESMES/TS

18. Attribute

Definition: A characteristic of an object or entity.

Context: An entity is any concrete or abstract thing of interest, including associations

among things. A composite attribute is an attribute whose datatype is non-atomic. An attribute instance is a specific instance of an attribute. An attribute value is the value associated with an attribute instance (ISO/IEC International Standard 11179-3 "Information technology - Metadata registries (MDR)-Part 3: Registry metamodel and basic attributes", February 2003).

A data or metadata attribute is a statistical concept providing qualitative information about a specific statistical object such as a data set, observation, data provider, or dataflow. Concepts such as units, magnitude, currency of denomination, titles and methodological comments can be used as attributes in the context of an agreed data exchange.

A conditional attribute is permitted to take empty values. A mandatory attribute is an attribute which must take a value, otherwise the corresponding observation, which it refers to, is not considered as meaningful enough, e.g. with regard to the "status" of an observation or the units in which a whole time series is expressed.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms:
Attachment level
Class
Constraint
Data exchange
Data flow definition
Data structure definition
Entity
GESMES
GESMES/TS
ISO/IEC 11179
Metadata item
Metadata structure definition
Object
Statistical concept
Value domain

19. Base period

Definition: The period of time used as the base of an index number, or to which a constant series refers.

Context: The base period refers to the period when the published index is 100, or to which weights or base data refer to. It can be one single year (e.g. 1995=100) but it may be as short as one day or as long as a specified number of years. "Base period" may include an indication of the value of the series in the base period (usually 1 or 100).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Base weight
Compilation practices
Reference period
Special Data Dissemination Standard, SDDS
Weight period

20. Base weight

Definition: The weights of a weighting system for an index number computed according to the information relating to the base period instead, for example, of the current period.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Base period
Index number
Weight

21. Base year

See "Base period"

22. Basic data

See "Source data"

23. Benchmark

Definition: A recognised standard that forms the basis for comparison.

Context: In quality improvement lexicon, a benchmark is a best in class achievement. This achievement then becomes the reference point or recognized standard of excellence against which similar processes are measured.

Benchmark data is a recognised standard data set against which other data sets are compared.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

24. Benchmarking

Definition: Comparing data, metadata or processes against a recognised standard.

Context: 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, as annual benchmark data are available some time after quarterly data. Benchmarking does 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 (Maitland-Smith, F, "Use of Benchmark Data to Align or Derive Quarterly/Monthly Estimates", paper presented at the June 2002 meeting of the OECD Short-term Economic Statistics Working Party, Paris).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Interpolation

25. Bias

Definition: An effect which deprives a statistical result of representativeness by systematically distorting it, as distinct from a random error which may distort on any one occasion but balances out on the average.

Context: The bias of an estimator is the difference between its mathematical expectation and the true value it estimates. In the case it is zero, the estimator is said to be unbiased.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Accuracy

26. Bilateral exchange

Definition: Exchange of data and/or metadata between a sending organisation and a receiving organisation where all aspects of the exchange process are agreed between counterparties, including the mechanism for exchange of data and metadata, the formats, the frequency or schedule, and the mode used for communications regarding the exchange.

Context: Apart from bilateral exchange, the SDMX initiative identifies two other basic forms of exchange of statistics and metadata between organisations, i.e. multilateral exchange and data-sharing exchange.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data exchange
Data sharing
Multilateral exchange

27. Break

See "Time series break"

28. Category

Definition: The generic term for items at any level within a classification.

Context: Typical classification levels are: tabulation categories, sections, subsections, divisions, subdivisions, groups, subgroups, classes and subclasses. Classification categories are usually identified by codes (alphabetical or numerical) which provide both a unique identifier for each category and denote their place within the hierarchy. They contain elements which are subsets of the classification to which they belong, such as activities, products, types of occupations, types of education, etc.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Category Scheme
Classification
Structure

29. Category Scheme

Definition: Scheme made up of a hierarchy of categories, which may include any type of useful classification for the organization of data and metadata.

Context: In SDMX, the category scheme is an artefact for organising categories which themselves link to dataflow definition or metadataflow definition.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Category
Classification
Dataflow definition
Metadataflow definition

30. Census

Definition: A survey conducted on the full set of observation objects belonging to a given population or universe.

Context: A census is the complete enumeration of a population or groups at a point in time with respect to well defined characteristics: for example, Population, Production, Traffic on particular roads. In some connection the term is associated with the data collected rather than the extent of the collection so

that the term sample census has a distinct meaning. The partial enumeration resulting from a failure to cover the whole population, as distinct from a designed sample enquiry, may be referred to as an "incomplete census". (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003).

Specific types of censuses include:

Register-based censuses - the enumeration of a population through the use of data from one or more registers rather than a statistical survey.

Rolling census - the enumeration of population over a period of time using multiple waves of data collection, each of which covers a proportion of that population.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Observation
Sample
Statistical population
Survey

31. Chain index

Definition: An index number in which the value in any given period is related to a base in the previous period, as distinct from one which is related to a fixed base.

Context: The comparison of non-adjacent periods is usually made by multiplying consecutive values of the index numbers, which, as it were, form a chain from one period to another.

In practice chain index numbers are usually formed from weighted average of link-relatives, namely the values of magnitudes for a given period divided by the corresponding values in the previous period.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Index type

32. Characteristic

Definition: An abstraction of a property of an object or of a set of objects.

Context: Essential characteristic is a characteristic which is indispensable to understanding a concept [ISO 1087-1:2000]

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Aggregated data
Concept
Data
ISO/IEC 11179
Object
Statistical concept
Statistical subject-matter domain
Time series
Statistical variable

33. Civil Status

Definition: Legal, conjugal status of each individual in relation to the marriage laws or customs of the country.

Context: The civil status is often referred to as marital status and represented through codes of the respective code list.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

34. Clarity

Definition: The extent to which easily comprehensible metadata are available, where these metadata are necessary to give a full understanding of statistical data.

Context: 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 European Statistics Code of Practice, clarity is strictly associated to accessibility to form one single quality criteria: "accessibility and clarity".

Source: SDMX (2009)

Hyperlink:

Related Terms: Accessibility

35. Class

Definition: A description of a set of objects that share the same attributes, operations, methods, relationships, and semantics.

Context:

Source: [ISO/IEC 19501-1:2001, 2.5.2.9] in ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", September 2004

Hyperlink: http://metadata-standards.org/11179/#11179-1

Related Terms: Attribute
ISO/IEC 11179
Object

36. Classification

Definition: A set of discrete, exhaustive and mutually exclusive observations, which can be assigned to one or more variables to be measured in the collation and/or presentation of data.

Context: The Neuchâtel terminology model states that the term classification is normally used to denote one of the following concepts: a. The general idea of assigning statistical units to categories representing the values of a certain variable. b. The general concept of a structured list of mutually exclusive categories, each of which describes a possible value of the classification variable. Such a structured list may be linear or hierarchically structured. A linear classification is a list of categories, which are all at one and the same level (e.g. the ISO 3166 country code list, or a classification of marital status). In a hierarchical classification the categories are arranged in a tree-structure with two or more levels, where each level contains a set of mutually exclusive categories. The items of each level but the highest (most aggregated) are aggregated to the nearest higher level. In common usage the term classification often implies a hierarchical classification. c. One particular structured list of mutually exclusive categories, which is named, has a certain stability and normative status, and is valid for a given period of time (e.g. ISIC Rev. 1). d. One particular named set of several structured lists of mutually exclusive categories, which are consecutive over time and describe the possible values of the same variable (e.g. ISIC). The distinction between concepts c. and d. above, although seldom made explicit, is particularly crucial in any systematic register of classifications or in the development of a classification database. (Neuchâtel Group, "Neuchâtel Terminology: Classification database object types and their attributes - Version 2", September 2002).

The terms "classification" and "nomenclature" are often used interchangeably, despite the definition of a "nomenclature" being narrower than that of a "classification". The structure of classification can be either hierarchical or flat. Hierarchical classifications range from the broadest level (e.g. division) to the detailed level (e.g. class). Flat classifications (e.g. sex classification) are not hierarchical.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Analytical unit
Category
Category scheme
Classification system
Classification unit
Classification version
Disaggregation
Maintenance agency
Nomenclature
Observation

Observation unit
Standard classification
Statistical unit
Taxonomy

37. Classification system

Definition: Arrangement or division of objects into groups based on characteristics which the objects have in common.

Context: The definition entails the description of the classification being used and how this conforms to internationally agreed standards, guidelines, or good practices. It also refers to the description of deviations of the classification system compared to statistical standards, guidelines, or good practices, when relevant.

Examples of international classifications are ISCO (International Standard Classification of Occupations), ISIC (the United Nations International Standard Industrial Classification of All Economic Activities), NACE (Statistical Classification Of Economic Activities) and NUTS (Nomenclature of Territorial Units for statistics).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Classification

38. Classification unit

Definition: The basic unit to be classified in the classification.

Context: For instance, in an activity classification, the classification unit would be the establishment or enterprise, while in an occupational classification it will be the job.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.asp

Related Terms: Classification

39. Classification version

Definition: List of mutually exclusive categories representing the version-specific values of the classification variable.

Context: A classification version has a certain normative status and is valid for a given period of time. A new version of a classification differs in essential ways from the previous version. "Essential" changes are changes that alter the borders between categories, i.e. a statistical object/unit may belong to different categories in the new and the older version. "Border" changes may

be caused by creating or deleting categories, or moving a part of a category to another. The addition of case law, changes in explanatory notes or in the titles do not lead to a new version.

Source: Neuchatel Group, "Neuchatel Terminology: Classification database object types and their attributes - Version 2", September 2002

Hyperlink:

Related Terms: Classification
Institutional unit

40. Code

Definition: A language-independent set of letters, numbers or symbols that represent a concept whose meaning is described in a natural language.

Context: A code normally consists of one or more alphabetic, numeric or alpha/numeric characters.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Code list
Coding
Level

41. Code list

Definition: A predefined list from which some statistical coded concepts take their values.

Context:

Source: SDMX (2009)

Hyperlink:

Related Terms: Code
Coding
Concept scheme
Dimension
GESMES/TS
Keyword
Statistical concept
Structural definition

42. Coding

Definition: The process of converting verbal or textual information into codes representing classes within a classification scheme, to facilitate data processing, storage or dissemination.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Code
Code list
Coding error

43. Coding error

Definition: The assignment of an incorrect code to a data item

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Coding

44. Coherence

Definition: Adequacy of statistics to be combined in different ways and for various uses.

Context: When originating from different sources, and in particular from statistical surveys using different methodology, statistics are often not completely identical, but show differences in results due to different approaches, classifications and methodological standards. 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 from the same socio-economic domain, and between survey statistics and national accounts.

The concept of coherence is closely related to the concept of comparability between statistical domains. Both coherence and comparability refer to a data set with respect to another. The difference between the two is that comparability refers to comparisons between statistics based on usually unrelated statistical populations and coherence refers to comparisons between statistics for the same or largely similar populations.

Coherence can be generally broken down into "Coherence - cross domain" and "Coherence - internal".

Users should be aware that, in the Data Quality Assessment Framework of the International Monetary Fund, the term "consistency" is used for indicating "logical and numerical coherence". In that framework, "internal consistency" and "intersectoral and cross-domain consistency" can be mapped to "internal coherence" and "cross-domain coherence" respectively.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Comparability

Consistency
Data confrontation

45. Collection

See "Data collection"

46. Comment

Definition: Supplementary descriptive text which can be attached to data or metadata.

Context: In SDMX messages, a comment may contain a descriptive text which can be attached to an agency, provision agreement, dataflow, data set, group, time series, or observation.

A list of ID broken down by attachment level is therefore needed: COMMENT_AGENCY, COMMENT_AGR, COMMENT_DFL, COMMENT_DSET, COMMENT_GRP, COMMENT_TS, COMMENT_OBS.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

47. Comparability

Definition: The extent to which differences between statistics can be attributed to differences between the true values of the statistical characteristics.

Context: Comparability aims at measuring the impact of differences in applied statistical concepts and definitions on the comparison of statistics between geographical areas, non-geographical dimensions, or over time. Comparability of statistics, i.e. their usefulness in drawing comparisons and contrast 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 must convey such information that will help any interested party in evaluating comparability of the data, which is the result of a multitude of factors.

In some quality frameworks, for instance in the European Statistical Code of Practice, comparability is strictly associated with the coherence of statistics.

The concept can be further broken down into:

(a) Comparability - geographical, referring to the degree of comparability between statistics measuring the same phenomenon for different geographical areas.

(b) Comparability over time, referring to the degree of comparability between two or more instances of data on the same phenomenon measured at different points in time.

(c) Comparability between domains, referring to the comparability between different survey results which target similar characteristics in different statistical domains.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Coherence
Statistical unit

48. Compilation

See "Data compilation"

49. Compilation practices

See "Data compilation"

50. Compiling agency

Definition: The organisation compiling the data being reported.

Context: The dimension is needed as two agencies might be compiling the exact same data but using different sources or concepts (the latter would be partially captured by the dimensions). The provider ID may be not sufficient, as one provider could disseminate the data compiled by different compiling agencies.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

51. Completeness

See "Relevance"

52. Computer Assisted Interviewing, CAI

Definition: The use of the computer during interviewing.

Context: Any contradictory data can be flagged by edit routines and the resulting data can be immediately adjusted by information from the respondent. An added benefit is that data capture (key-entry) is occurring at interview time.

CAI assists the interviewer in the wording of questions and tailors succeeding questions based on previous responses. CAI has been mainly used in "Computer-Assisted Telephone Interviews" (CATI) or "Computer-Assisted Personal Interviewing" (CAPI).

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://amrads.jrc.cec.eu.int/k-base/glossary/glossALL.htm

Related Terms:

53. Concept

Definition: A unit of knowledge created by a unique combination of characteristics.

Context: Concepts are abstract summaries, general notions, knowledge, etc., of a whole set of behaviours, attitudes or characteristics which are seen as having something in common. Concepts are used to assist in presenting/conveying precise meaning, categorising, interpreting, structuring and making sense of phenomena (such as classifications) ("United Nations Glossary of Classification Terms" prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper, available at: http://unstats.un.org/unsd/class/family/glossary_short.htm).

A semantic link among two or more concepts is a concept relationship. A description of the type of relationship among two or more concepts is a concept relationship type description. (ISO/IEC International Standard 11179-3 "Information technology-Metadata registries (MDR)-Part 3: Registry metamodel and basic attributes", February 2003).

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms:
Analytical framework
Characteristic
Data structure definition
Definition
Dimension
ISO/IEC 11179
Metadata structure definition
Ontology
Statistical concept
Statistical subject-matter domain
Structural definition
Terminological system

54. Concept Scheme

Definition: Maintained list of concepts that are used in data structure definitions or metadata structure definitions.

Context: Structural definitions of both data and reference metadata associate specific statistical concepts with their representations, whether textual, coded, etc. In SDMX version 2.0, these concepts are taken from a "concept scheme" which is maintained by a specific agency. Concept schemes group a set of concepts, provide their definitions and names, and allow for semantic relationships to be expressed, when some concepts are specifications of others. It is possible for a single concept scheme to be used both for data structures and metadata structures. A core representation of each concept can be specified (e.g. a code list, or other representations such as "date").

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Code list
Data structure definition
Metadata structure definiton
Reference metadata
SDMX, Statistical data and metadata exchange
Statistical concept
Structural definition

55. Conceptual data model

Definition: A data model that represents an abstract view of the real world. A conceptual model represents the human understanding of a system.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Data model
ISO/IEC 11179

56. Conceptual domain

Definition: Concept that expresses its valid instance meanings or description.

Context: In ISO/IEC 11179, the value meanings in a conceptual domain may either be enumerated or expressed via a description. Enumerated conceptual domain is a conceptual domain that is specified by a list of all its value meanings.

"Non-enumerated" is a conceptual domain that is not specified by a list of all valid value meanings."Non-enumerated conceptual domain description" is a description or specification of a rule, reference, or range for a set of all Value Meanings for the Conceptual.

"Conceptual domain relationship" is a relationship among two or more conceptual domains. "Conceptual domain relationship type description" is a description of the type of relationship among two or more Conceptual domains. "Conceptual domain representation" is a relationship between a Conceptual domain and a Value domain.

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: Data model
ISO/IEC 11179
Permissible value
Unit of measure
Value meaning

57. Confidential data

Definition: Data which are subject to confidentiality clauses.

Context: The data collected by many national statistical agencies are subject to national rules regarding confidentiality. The two main reasons for declaring data to be confidential are: a) too few units in a cell; b) dominance of one or two units in a cell. The limits of what constitutes "too few" or "dominance" vary between statistical domains.

In the European Union, confidential data is defined in Article 13 of Council Regulation No 322/97, as:

1. Data used by the national authorities and the Community authority for the production of Community statistics shall be considered confidential when they allow 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 the means that might reasonably be used by a third party to identify the said statistical unit.

2. By derogation from paragraph 1, data taken from sources which are available to the public and remain available to the public at the national authorities according to national legislation, shall not be considered confidential.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Confidentiality

58. Confidentiality

Definition: 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.

Context: Confidentiality refers to a property of data with respect to whether, for example, they are public or their disclosure is subject to restrictions. For instance, data allowing the identification of a physical or legal person, either directly or indirectly, may be characterised as confidential according to the relevant national or international legislation. Unauthorised disclosure of data that are restricted or confidential is not permitted and even legislative measures or other formal provisions may be used to prevent disclosure.

Often, there are procedures in place to prevent disclosure of restricted or confidential data, including rules applying to staff, aggregation rules when disseminating data, provision of unit records, etc.

Sensitivity (of information) is sometimes used as a synonym to confidentiality.

This concept can be broken down into: Confidentiality - policy; Confidentiality - status; Confidentiality - data treatment.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Access
Confidential data
Data security

59. Consistency

Definition: Logical and numerical coherence.

Context: An estimator is called consistent if it converges in probability to its estimand as sample increases (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003). Consistency over time, within datasets, and across datasets (often referred to as inter-sectoral consistency) are major aspects of consistency. In each, consistency in a looser sense carries the notion of "at least reconcilable". For example, if two series purporting to cover the same phenomena differ, the differences in time of recording, valuation, and coverage should be identified so that the series can be reconciled. Inconsistency over time refers to changes that lead to breaks in series stemming from, for example, changes in concepts, definitions, and methodology. Inconsistency within datasets may exist, for example, when two sides of an implied balancing statement-assets and liabilities or inflows and outflows-do not balance. Inconsistency across datasets may exist when, for example, exports and imports in the national accounts do not reconcile with exports and imports within the balance or payments. Within the IMF definition of quality, "consistency" is one of the elements of "serviceability".

Source: International Monetary Fund, "Data Quality Assessment Framework - DQAF - Glossary", unpublished

Hyperlink:

Related Terms: Coherence
Serviceability

60. Consolidation

Definition: The process that takes data from different systems, entities (and possibly formats) and combines that information to create a unified view.

Context: Consolidation refers to the elimination, both from uses and resources, of transactions which occur between units when the latter are grouped and to the elimination of reciprocal financial assets and liabilities

In the context of fiscal sector data, consolidation provides an indication of the methods used to combine data from separate central government accounts and funds to derive statistics for transactions between the entire central government and any other sector, exclusive of transactions between units within the same coverage of central government, as defined in the Government Finance Statistics Manual.: 1) whether all transactions between units of general government have been eliminated in consolidation is requested; 2) debt issues of one unit of central government that are held by another unit are reported on a consolidated or unconsolidated basis (e.g., central government securities held by the social security fund) are also specified.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Statistical processing

61. Constraint

Definition: Specification of what may be contained in a data or metadata set in terms of the content or, for data only, in terms of the set of key combinations to which specific attributes (defined by the data structure) may be attached.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Attribute

62. Contact

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

Context: "Contact" describes contact points for the data or metadata, including how to reach the contact points.
The main attributes of "contacts" are:
Contact mail address: The mailing address of the contact.
Contact name: The name of the contact.
Contact title: The name of the position held by the contact.
Contact email address: An e-mail address for correspondence with the contact.
Contect phone number: A telephone number for spoken correspondence.
Other detailed attributes (e.g. country code, city code, first name, family name, etc.) can be identified for the purpose of database storage.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: ISO/IEC 11179
SDMX, Statistical Data and Metadata Exchange

63. Context

Definition: Circumstance, purpose, and perspective under which an object is defined or used.

Context: A context description language is the identifier of the language used in the context description (ISO/IEC FCD 11179-3, "Registry Metamodel, Final Committee Draft", 2001). The administration record for a context is a context administration record. The textual description of the context is a context description. The identifier of the language used in the context description is a context description language identifier (ISO/IEC International Standard 11179-3 "Information technology-Metadata registries (MDR)-Part 3: Registry metamodel and basic attributes", February 2003).

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Administered item
ISO/IEC 11179

64. Cost and burden

Definition: Cost associated with the collection and production of a statistical product, and burden on respondents.

Context: The cost is associated with a statistical product and can be financial, human or time-related. It may consist of staff costs, data collection costs and other costs related to reporting obligations.

The burden is often measured by costs for the respondents (businesses, institutions, households, individuals) imposed by a statistical obligation. The overall burden of delivering the information depends on: a) the number of respondents; b) the average time required to provide the information, including time spent after receipt of the questionnaire ("recontact time"); and c) the hourly cost of a respondent's time.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

65. Counterpart reference area

Definition: The secondary area, as opposed to the reference area, to which the measured data is in relation.

Context: The "counterpart area" (also known as "vis-a-vis area") is related to statistics on foreign trade, migration or other domains. It determines, from the point of view of the reporting country, the corresponding area to which the economic or other flows are related to (for instance, in statistics on imports, the counterpart reference area is the area of origin of the goods).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Reference area

66. Country identifier

Definition: An identifier specifying a geopolitical area.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Identifier
ISO/IEC 11179

67. Coverage

Definition: The definition of the population that statistics aim to cover.

Context: The term "coverage" encompasses the descriptions of key dimensions delimiting the statistics produced, e.g. geographical, institutional, product, economic sector, industry, occupation, transaction, etc., as well as relevant exceptions and exclusions.

The term Coverage describes the scope of the data compiled, rather than the characteristics of the survey.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Coverage error
Coverage ratio
Data
Geographical coverage
Scope
Special Data Dissemination Standard, SDDS

68. Coverage error

Definition: Error caused by a failure to cover adequately all components of the population being studied, which results in differences between the target population and the sampling frame.

Context: Coverage errors include over-coverage, under-coverage and misclassification. Incomplete sampling frames often result in coverage errors.

Source: Statistical Office of the United Nations, "Handbook of Household Surveys, Revised Edition", (para. 8.3), Studies in Methods, Series F, No. 31, United Nations, New York, 1984

Hyperlink:

Related Terms: Coverage
Statistical population

69. Coverage ratio

Definition: A measure of the extent to which observations designated as primary to a particular category are undertaken by units primarily involved with the observations related to that category.

Context: In industry statistics, the coverage ratio is the output of goods and services characteristic of a particular industry in proportion to the total output of the same goods and services by the economy as a whole.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Coverage
Observation

70. Creation date

Definition: The date on which the record, data or metadata item was created.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administered item
Date
Date of last change
ISO/IEC 11179

71. Credibility

Definition: Confidence that users place in statistical products based simply on their image of the data producer, i.e., the brand image.

Context: Confidence by 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: OECD, "Quality Framework for OECD Statistics", Paris, June 2002 (http://www.oecd.org/document/43/0,2340,en_2649_34257_21571947_119820_1_1_1,00.html)

Hyperlink: http://www.oecd.org/document/43/0,2340,en_2649_34257_21571947_119820_1_1_1,00.htm

Related Terms:

72. Cross-domain Concepts

Definition: List of standard concepts, covering structural and reference metadata, which should be used in several statistical domains wherever possible to enhance possibilities of the exchange of data and metadata between organisations.

Context: Within SDMX, cross-domain concepts are envisaged to cover various elements describing statistical data and their quality. When exchanging statistics, institutions can select from a standard set of content-oriented concepts. The list of concepts and their definitions reflects recommended practices and can be the basis for mapping between internal systems when data and metadata are exchanged or shared between and among institutions.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Reference metadata
SDMX, Statistical Data and Metadata Exchange
Structural metadata

73. Currency

Definition: Monetary denomination of the object being measured.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

74. Cut-off survey

Definition: A survey in which all the entities falling above or below a threshold determined according to one or more characteristics of those entities are either included or excluded

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Cut-off threshold
Survey

75. Cut-off threshold

Definition: A threshold used to exclude entities from the target population (and hence from the frame).

Context: Entities may be excluded from the target population using a cut-off threshold for various reasons, for example:

- they contribute very little to the resulting statistics;

- the cost / burden of including them outweighs the benefits;

- they are out of scope (e.g. large enterprises in a survey of SMEs)

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Cut-off survey
Statistical population
Target population

76. Data

Definition: Characteristics or information, usually numerical, that are collected through observation.

Context: Data are the physical representation of information in a manner suitable for communication, interpretation, or processing by human beings or by automatic means (Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000).

Statistical data are data derived from either statistical or non-statistical sources, which are used in the process of producing statistical products.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Characteristic
Coverage
Data analysis
Data presentation
Metadata
Periodicity
Special Data Dissemination Standard, SDDS
Timeliness

Access
Administrative source

77. Data analysis

Definition: The process of transforming raw data into usable information, often presented in the form of a published analytical article, in order to add value to the statistical output.

Context:

Source: Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 70

Hyperlink: http://www.statcan.ca/english/freepub/12-539-XIE/12-539-XIE03001.pdf.

Related Terms: Data

78. Data attribute

See "Attribute"

79. Data capture

Definition: The process by which collected data are put into a machine-readable form.

Context: Elementary edit checks are often performed in sub-modules of the software that does data capture.

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms:

80. Data checking

Definition: Activity through which the correctness conditions of the data are verified.

Context: It also includes the specification of the type of the error or condition not met, and the qualification of the data and its division into the "error free" and "erroneous data". Data checking may be aimed at detecting error-free data or at detecting erroneous data.

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms: Data reconciliation
Graphical data editing
Macro editing

81. Data collection

Definition: Systematic process of gathering data for official statistics.

Context: There are a number of data collection methods used for official statistics, including computer-aided personal or telephone interview CAPI/CATI, mailed questionnaires, electronic or internet questionnaires and direct observation. The data collection may be exclusively for statistical purposes, or primarily for non-statistical purposes.

Descriptions of data collection methods should include the purpose for which the data were collected, the period the data refer to, the classifications and definitions used, and any constraints related to further use of these data.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

82. Data compilation

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

Context: "Data compilation" refers to the description of statistical procedures used for producing intermediate data and final statistical outputs. Data compilation covers, among other things, the use of weighting schemes, methods for imputing missing values or source data, statistical adjustment, balancing/cross-checking techniques and relevant characteristics of the specific methods applied.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

83. Data compiling agency

see "Compiling agency"

84. Data confrontation

Definition: The process of comparing data that has generally been derived from different surveys or other sources, especially those of different frequencies, in order to assess and possibly improve their coherency, and identify the reasons for any differences.

Context: Such data may not be coherent for a number of reasons including the use of different data item definitions, classifications, scope, reference period, etc.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Coherence
Data reconciliation

85. Data consumer

Definition: Organisation which uses data as input for further processing.

Context: An organisation can play a number of organisation roles. In the SDMX information model, three roles are identified at present: Data Provider; Data Consumer; Maintenance Agency. Data Consumer also embraces the activity of metadata provision.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data provider
Maintenance agency
Organisation role

86. Data dissemination agency

Definition: The organisation disseminating the data.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

87. Data editing

Definition: Activity aimed at detecting and correcting errors, logical inconsistencies and suspicious data.

Context: Editing techniques refer to a range of procedures and processes used for detecting and handling errors in data, also aiming at avoiding their future repetition. An "edit" is the correction of an error in data. An "editing rule" is

the specification of the conditions under which edits are made.

Examples of different techniques include the different approaches to editing such as micro-editing/ macro-editing, input/output editing, or to the various tools available for editing such as graphical editing, interactive editing, etc.

Edit types refer to the actual nature of edits applied to data during input or output processing. Examples include:

- logical edits, to ensure that two or more data items do not have contradictory values;

- consistency edits, to ensure that precise and correct arithmetic relationships exists between two or more data items;

- range edits, identifying whether or not a data item value falls inside a determined acceptable range;

- variance edits, looking for suspiciously high variances at the output edit stage.

Micro-editing and macro-editing may be distinguished in order to calculate rate of edits. Edit types may also refer to whether these edits are fatal or query type, i.e. whether they detect errors with certainty or point to suspicious data items.

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms:
Data validation
Graphical data editing
Macro editing
Micro editing

88. Data element

Definition: A unit of data for which the definition, identification, representation, and permissible values are specified by means of a set of attributes.

Context: “Data element administration record” is the Administration record for a Data element. “Data element precision” is the degree of specificity for a Data element. “Data element representation” is the relationship between a Data element and its Value domain. “Data element representation class” is the class of representation of a Data element. (ISO/IEC International Standard 11179-3 "Information technology - Metadata registries-Part 3: Registry metamodel and basic attributes", February 2003)

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms:
ISO/IEC 11179
Keyword
Related data reference
Thesaurus

89. Data exchange

Definition: The process of sending and receiving data.

Context: Data exchange should take place in such a manner that the information content or meaning assigned to the data is not altered during the transmission.

"Data exchange context" is the framework in which two or more partners agree to exchange one or more identified sets of data (exchanged time series, ETS) and related attributes, and use one or more data structure definitions to serve this requirement, possibly on the basis of a formal or informal agreement.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Attribute
Bilateral exchange
Data sharing
Data structure definition
Electronic data interchange, EDI
Gateway exchange
GESMES/TS
Multilateral exchange
Statistical message

90. Dataflow definition

Definition: A structure which describes, categorises and constrains the allowable content of a data set that providers will supply for different reference periods.

Context: In SDMX, data sets are reported or disseminated according to a data flow definition. The data flow definition identifies the data structure definition and may be associated with one or more subject matter domains. This facilitates the search for data according to organised category schemes.

A "data flow", in this context, is an abstract concept of the data sets, i.e. a structure without any data. While a data structure definition defines dimensions, attributes, measures and associated representation that comprise the valid structure of data and related metadata contained in a data set, the dataflow definition associates a data structure definition with one or more category. This gives a system the ability to state which data sets are to be reported for a given category and which data sets can be reported using the data structure definition. The dataflow definition may also have additional metadata attached, defining qualitative information and constraints on the use of the data structure definition, in terms of reporting periodicity or specifying the subset of codes to be used in a dimension.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Attribute
Category scheme
Metadataflow definition

91. Data identifier

See "Originator Data identifier"

92. Data imputation

See "Imputation"

93. Data integration

Definition: The process of combining data from two or more sources to produce statistical outputs.

Context: Data integration can be at the micro-level, where it is often referred to as matching, or at the macro-level.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

94. Data interchange

See "Data exchange"

95. Data model

Definition: A graphical and/or lexical representation of data, specifying their properties, structure and inter-relationships.

Context: The UN defines a data model as a user's conceptual design of a data set that describes the database entities and their relations to one another (United Nations Department of Economic and Social Affairs, "Handbook on Geographic Information Systems and Digital Mapping", Studies in Methods, Series F, No. 79, Annex VI - Glossary, New York, 2000)

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Conceptual data model
Conceptual domain
ISO/IEC 11179
Metamodel

96. Data presentation

Definition: Description of the disseminated data.

Context: Data can be displayed to users as tables, graphs or maps. According to the Fundamental Principles of Official Statistics, the choice of appropriate presentation methods should be made in accordance with professional considerations. Data presentation includes the description of the dataset disseminated with the main variables covered, the classifications and breakdowns used, the reference area, a summary information on the time period covered and, if applicable, the base period used.

This concept can be broken down into: "data presentation: description" and "data presentation: disseminated detail".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data
Data set
SDMX, Statistical Data and Metadata Exchange

97. Data processing

Definition: The operation performed on data in order to derive new information according to a given set of rules.

Context: The processing site refers to the organisation, institute, agency, etc, responsible for undertaking the collection, tabulation, manipulation and preparation of data and metadata output. The processing site may or may not also refer to the physical location(s) at which such activities are carried out. A processing system embodies both manual and automated systems used by agencies to despatch questionnaires, collect, compile, manipulate, analyse and disseminate data and metadata output. Such systems therefore cover all stages of the statistical processing cycle.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Compilation practices
Processing error
Statistical processing

98. Data provider

Definition: Organisation which produces data or metadata.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Access
Data provider series key
Metadata

99. Data provider series key

Definition: Identifier used by the data providers systems to distinguish between data series

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data provider
Identifier
Key
Time series

100. Data reconciliation

Definition: The process of adjusting data derived from two different sources to remove, or at least reduce, the impact of differences identified.

Context: Editing and reconcilation may involve fixing errors or adopting alternative sources and methods that are aimed at improving the process of reviewing or understanding data.

This may entail the reconciliation of stocks and transactions data; reconciliation of reported data with money and banking statistics, custodian data; differences with partner data or preshipment inspection data; the treatment of differences between GDP compiled for the production approach and GDP compiled from the expenditure approach. It is a special kind of editing done after initial compilation.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Compilation practices
Data checking
Data confrontation
Special Data Dissemination Standard, SDDS
Statistical processing

101. Data revision

Definition: Any change in a value of a statistic released to the public by an official statistical agency.

Context: Preliminary data are revised when more and better source data become available. "Data revision" describes the policy and practice for identifying the revision status of the data, as well as the availability of revision studies and analyses.

This concept can be broken down into: "Data revision - policy"; "Data revision - practice"; "Data revision - studies".

"Revision policy" refers to the policy aimed at ensuring the transparency of disseminated data. The general guidelines for handling data revisions applied by a data providing agency should be described.

"Revision practice" refers to documentation regarding the source data used and the way they are adjusted. It also describes the revision status of available data. Data may also be subject to regular or 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.

"Revision studies" refers to the information about available studies and analyses on data revision. These studies can contain mean revision and revision variance in estimates or other quantitative measures of the effects of revisions.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Compilation practices

102. Data security

Definition: The measures taken to prevent unauthorized access or use of data.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Confidentiality
Disclosure analysis

103. Data set

Definition: Any organised collection of data.

Context: Within SDMX, a data set can be understood as a collection of similar data, sharing a structure, which covers a fixed period of time. A data set is any permanently stored collection of information usually containing either case level data, aggregation of case level data, or statistical manipulations of either the case level or aggregated survey data, for multiple survey instances (United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section

3.3.7, page 14). The terms database and data set are often used interchangeably. A logical collection of values or database objects relating to a single subject (United Nations Department of Economic and Social Affairs, "Handbook on Geographic Information Systems and Digital Mapping", Studies in Methods, Series F, No. 79, Annex VI - Glossary, New York, 2000).

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms:
Aggregation
Data set identifier
Data presentation
Data structure definition
GESMES/CB
Measure
Sibling group
Time series

104. Data set identifier

Definition: Sequence of characters identifying the data set with which it is associated.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Data set
Identifier
Originator data identifier

105. Data sharing

Definition: Exchange of data and/or metadata in a situation involving the use of open, freely available data formats and where process patterns are known and standard.

Context: Apart from data-sharing, SDMX identifies two other basic forms of exchange of statistics and metadata between organisations, i.e. bilateral exchange and multilateral exchange.

In data sharing exchange, any organization or individual can use any counterparty's data and metadata (assuming they are permitted access to it). This model requires no bilateral agreement, but only requires that data and metadata providers and consumers adhere to the standards.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Bilateral exchange
Data exchange
Multilateral exchange

106. Data status

Definition: Whether the data initially disseminated to the public are final data or preliminary and therefore subject to revision.

Context: The data identified as meeting the standard may be preliminary and subject to revision and designated as such.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Special Data Dissemination Standard, SDDS

107. Data storage

Definition: The means of recording or archiving data so that they are available for future use.

Context: The Fundamental Principles of Official Statistics require that decisions on data storage are made in accordance with strictly professional considerations.

Source: SDMX (2009)

Hyperlink:

Related Terms:

108. Data structure definition

Definition: Set of structural metadata associated to a data set, which includes information about how concepts are associated with the measures, dimensions, and attributes of a data cube, along with information about the representation of data and related descriptive metadata.

Context: In order to exchange statistical information, a central institution has to agree with its partners about which statistical concepts are necessary for identifying the series (and for use as dimensions) and which statistical concepts are to be used as attributes. These definitions form the data structure definition (also called "key family" in Gesmes/TS). The coded statistical concepts assigned as dimensions in a key structure are called the dimensions of the data structure definition. In addition to the dimensions, each data structure definition assigns a set of statistical concepts that may qualify the statistical information at data set, sibling group, time series or observation level. The statistical concepts used in this way are called "attributes".

Each data structure definition has the following properties: a) identifier (providing a unique identification within an exchanged time series); b) name (also unique); and c) description (a description of the purpose and domain covered).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Attribute
Concept
Concept scheme
Data exchange
Data set
Maintenance agency
Statistical concept
Structural definition
Structural metadata
Structure

109. Data type

Definition: A set of distinct values characterized by properties of those values and by operations on those values. [ISO/IEC 11404:1996, 4.11]

Context: Datatype annotation is the specifying information to further define the Datatype.

Datatype description descriptive information to further clarify the Datatype.

Datatype name is a designation for the Datatype.

Datatype scheme reference is a reference identifying the source of the Datatype specification.

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179

110. Data update

Definition: The date on which the data element was inserted or modified in the database.

Context: The data update relates to the date which is registered in the production or the dissemination database of a data compiling organisation. The modification can imply several actions: data can be updated, verified and validated without change, or deleted. The data update does not necessarily imply that data are released.

The date of the data update may refer to the last update of the content, or to the latest verification without update of the content, or to the date of dissemination on the web. Correspondingly, this concept can be broken down into: data update - last update; data update - last verification; data update - last posted.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

111. Data validation

Definition: Process of monitoring the results of data compilation and ensuring the quality of the statistical results.

Context: Data validation describes methods and processes for assessing statistical data, and how the results of the assessments are monitored and made available to improve statistical processes.

All the controls made in terms of quality of the data to be published or already published are included in the validation process. Validation also takes into account the results of studies and analysis of revisions and how they are used to improve statistical processes. In this process, two dimensions can be distinguished: (i) validation before publication of the figures and (ii) validation after publication.

This concept can be further broken down into "Data validation: intermediate", "Data validation: output" and "Data validation: source".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data editing

112. Date

Definition: A time reference.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Creation date
Date of last change

113. Date of last change

Definition: The date the administered item was last changed.

Context:

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: Administered item
Creation date

Date
ISO/IEC 11179
Time of recording

114. Decimals

Definition: The number of digits of an observation to the right of a decimal point.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

115. Definition

Definition: A statement that represents the concept and the essential nature of the term.

Context: In classifications, this item refers to the explanation of the concepts encompassed in category description and often includes specific examples of what is and is not included in particular categories (United Nations Glossary of Classification Terms).

According to ISO/IEC International Standard 11179-3 (Registry metamodel and basic attributes) a definition is a "representation of a concept by a descriptive statement which serves to differentiate it from related concepts [ISO 1087-1:2000, 3.3.1].

Source: Dublin Core Metadata Initiative, DCMI Metadata Terms, 2008

Hyperlink: http://dublincore.org/documents/dcmi-terms

Related Terms: Concept
ISO/IEC 11179
Metadataflow definition
Analytical framework

116. Derived data element

Definition: Data element derived from other data elements using a mathematical, logical, or other type of transformation, e.g. arithmetic formula, composition, and aggregation.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:

117. Derived statistic

Definition: A statistic which is obtained by an arithmetical observation from the primary observations.

Context: In this sense, almost every statistic is "derived". The term is mainly used to denote descriptive statistical quantities obtained from data which are primary in the sense of being mere summaries of observations, e.g. population figures are primary and so are geographical areas, but population-per-square-mile is a derived quantity.

Source: Marriott, F.H.C for the ISI, "A Dictionary of Statistical Terms", 5th edition, Longman Scientific & Technical , New York , 1990

Hyperlink:

Related Terms: Observation

118. Dimension

Definition: A statistical concept used, in combination with other statistical concepts, to identify a statistical series or single observations.

Context: In the GESMES/TS context, "dimension" is a coded statistical concept used (most probably together with other coded statistical concepts) to identify a time series, e.g. a statistical concept indicating a certain economic activity or a geographical reference area. (European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper)

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Code list
Concept
Dimensionality
GESMES/TS
Key
Key structure
Statistical concept
Time series
Unit of measure

119. Dimensionality

Definition: An expression of measurement without units.

Context: A quantity is a value with an associated unit of measure. 32º Fahrenheit, 0º Celsius, $100 USD, and 10 reams (of paper) are quantities. Equivalence

between two units of measure is determined by the existence of a quantity preserving one-to-one correspondence between values measured in one unit of measure and values measured in the other unit of measure, independent of context, and where characterizing operations are the same. Equivalent units of measure in this sense have the same dimensionality. The equivalence defined here forms an equivalence relation on the set of all units of measure. Each equivalence class corresponds to a dimensionality. The units of measure "temperature in degrees Fahrenheit" and "temperature in degrees Celsius" have the same dimensionality, because for each value measured in degrees Fahrenheit there is a value measured in degrees Celsius with the same quantity, and vice-versa. The same operations may be performed on quantities in each unit of measure. Quantity preserving one-to-one correspondences are the well-known equations C° = (5/9)*(F° - 32) and F° = (9/5)*(C°) + 32.

In other words, the equivalence between two units of measure is determined by the existence of an invertible transformation of one set of units to the other. This means that two units of measure have the same dimensionality if there is a function that maps values in one unit of measure to values into the other.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Dimension
ISO/IEC 11179
Unit of measure

120. Disaggregation

Definition: The breakdown of observations, usually within a common branch of a hierarchy, to a more detailed level to that at which detailed observations are taken.

Context: With standard hierarchical classifications, statistics for related categories can be grouped or collated (aggregated) to provide a broader picture, or categories can be split (disaggregated) when finer details are required and made possible by the codes given to primary observations ("United Nations Glossary of Classification Terms"; prepared by the Expert Group on International Economic and Social Classifications).

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Aggregation
Classification
Compilation practices
Observation
Statistical processing

121. Disclosure analysis

Definition: The process of protecting the confidentiality of data. It involves limiting the amount of detailed information disseminated and/or masking data via noise addition, data swapping, generation of simulated or synthetic data, etc.

Context:

Source: United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section 3.3.17, page 28

Hyperlink: http://www.census.gov/srd/www/metadata/metada18.pdf

Related Terms: Data security

122. Dissemination

Definition: Distribution or transmission of statistical data to users.

Context: Dissemination covers all activities by statistical producers aiming at making data and metadata accessible to users.

For data dissemination, various release media are possible, such as electronic format including the internet, CD-ROM, paper publications, files available to authorised users or for public use; fax response to a special request, public speeches, press releases.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Organisation role
Statistical Data and Metadata Exchange, SDMX

123. Dissemination format

Definition: Media by which statistical data and metadata are disseminated.

Context: A statistical product is a combination of macro-data and metadata packaged and presented in a specific form for dissemination. "Dissemination format" refers to the various means of dissemination used for making the data available to the public. It includes a description of the various formats available, including where and how to get the information (for instance paper, electronic publications, on-line databases).

This concept can be further broken down into: Microdata access, News release, Online database, Publications and Other formats.

Micro-data access refers to the information on whether micro-data are also disseminated. It should be stated if micro-data are also disseminated, e.g. to researchers. Access conditions should be described in short.

News release refers to the regular or ad-hoc press releases linked to the

data. This concept covers press releases or other kind of similar releases linked to data or metadata.

On-line database refers to the Information about an on-line database in which the disseminated data can be accessed. Link to the on-line database where the data are available, with a summary identification of domain names as released on the website, as well as the related access conditions.

Publications refers to the regular or ad-hoc publications in which the data are made available to the public. References to the most important data dissemination done through paper or on-line publications, including a summary identification and information on availability of the publication means.

Other refers to the references to the most important other data dissemination done. Examples of other dissemination formats are analytical publications edited by policy users, etc. This concept also includes "supplementary data", i.e. data not routinely disseminated that are made available to users upon request.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Accessibility
Statistical Data and Metadata Exchange, SDMX

124. Dissemination standard

Definition: Guideline or legally-binding prescription developed by international organisations for guiding the dissemination of statistics to the public.

Context: The formulation of data and metadata dissemination standards entails a consultation process between international organisations and with member countries.

An example of a dissemination standard is the Special Data Dissemination Standard (SDDS) guiding members of the International Monetary Fund in the dissemination to the public of comprehensive, timely, accessible, and reliable economic and financial statistics.

Source: SDMX (2009)

Hyperlink:

Related Terms: General Data Dissemination System, GDDS
Special Data Dissemination Standard, SDDS

125. Documentation

Definition: Descriptive text used to define or describe an object, design, specification, instructions or procedure.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:

126. Documentation on methodology

Definition: Descriptive text and references to methodological documents available.

Context: "Documentation on methodology" refers to the availability of documentation related to various aspects of the data, such as methodological documents, summary notes or papers covering concepts, scope, classifications and statistical techniques.

This concept also includes the "Advance notice of changes in methodology", indicating whether the public is notified before a methodological change affects disseminated data and, if so, how long before.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Methodology

127. Domain groups

Definition: Groups comprised of international organisations and/or national agencies working, formally or informally, towards the development of international guidelines and recommendations relevant to one or more statistical subject matter domains.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Statistical subject-matter domain

128. Dublin core

Definition: The Dublin Core Metadata Initiative is an open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models.

Context:

Source: Dublin Core

Hyperlink: http://www.dublincore.org

Related Terms: Glossary
Methodology

129. EDIFACT

Definition: Electronic Data Interchange for Administration, Commerce and Transport.

Context: EDIFACT was prepared by UN/ECE Trade Division and adopted by ISO/TC 154. The UN/ECE has also prepared Message Design Guidelines which are included in the UN/ECE Trade Data Interchange Directory. The standard was published in 1988 and amended with very small changes in 1990.

Source: ISO International Standard 9735:1988, "Electronic data interchange for administration, commerce and transport (EDIFACT)", September 1996.

Hyperlink:

Related Terms: Electronic data interchange, EDI
GESMES
SDMX-EDI

130. Education level

Definition: The highest level of an educational programme the person has successfully completed.

Context: The highest level of an educational programme the person has successfully completed is also called "educational attainment of a person". At international level, the ISCED is the standard classification of educational attainment.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

131. Electronic data interchange, EDI

Definition: Electronic exchange of data usually in forms that are compatible so that software or a combination of individuals and software can put the data in a compatible form at the receiving end if necessary.

Context: EDI offers businesses the opportunity to retrieve information electronically from their internal systems and to forward that information to trade partners/suppliers/customers/government through a communications network. An example might be pulling data of one type of data base management system into a sequential format and then transferring the data to a second location where the data are stored in a format different from the originating data base management system.

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms: Data exchange
EDIFACT
GESMES

132. Embargo time

Definition: The exact time at which the data can be made available to the public.

Context: Usually, there is a time span between the finalisation of the production process of statistical data and the moment when the data produced is released and made available to the users. This time span is called "embargo time".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

133. Entity

Definition: Any concrete or abstract thing that exists, did exist, or might exist, including associations among these things e.g. a person, object, event, idea, process, etc.

Context: An entity exists whether data about it are available or not.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Attribute
ISO/IEC 11179
Observation unit
Ontology

134. Error of estimation

Definition: The difference between an estimated value and the true value of a parameter or, sometimes, of a value to be predicted.

Context: It is immediately associated with accuracy since accuracy is used to mean "the inverse of the total error, including bias and variance" (Kish L., "Survey Sampling", 1965). The larger the error, the lower the accuracy.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Accuracy
Estimate

135. Error of observation

Definition: Error arising from imperfections in the method of observing a quantity, whether due to instrumental or to human factors.

Context:

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Measurement error

136. Estimate

Definition: The particular value yielded by an estimator in a given set of circumstances.

Context: The expression is widely used to denote the rule by which such particular values are calculated. It seems preferable to use the word "estimator" for the rule of procedure, and "estimate" for the values to which it leads in particular cases.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Error of estimation
Estimator
Non-sampling error
Reliability
Trend estimates

137. Estimation

Definition: The process of producing an estimate.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Aggregation
Compilation practices
Estimator
Precision
Ratio estimation
Special Data Dissemination Standard, SDDS
Statistical processing

138. Estimator

Definition: A rule or method of estimating a parameter of a population.

Context: An estimator is usually expressed as a function of sample values and hence is a variable whose distribution is of great importance in assessing the reliability of the estimate to which it leads.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Estimate
Estimation

139. Flag

Definition: An attribute of a cell in a data set representing qualitative information about the value of that cell.

Context: Examples of qualitative information that can be exchanged via a flag are: "provisional value", "estimated value", "revised value", "forecast", "unreliable or uncertain data (see explanatory texts)", "break in series (see explanatory texts)", "more information in...".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Footnote
Quantitative data

140. Flow data series

Definition: Statistical series presented as flow data series are cumulated during the reference period, for example, passenger car registrations, where the figure for the reference period is the sum of daily registrations.

Context:

Source: Organisation for Economic Co-operation and Development (OECD), "Main Economic Indicators"

Hyperlink:

Related Terms: Metadata repository

141. Follow-up

Definition: A further attempt to obtain information from a reporting unit in a survey or field experiment because the initial attempt has failed or later information is available.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Non-response
Non-response error

142. Footnote

Definition: A note or other text located at the bottom of a page of text, manuscript, book or statistical tabulation that provides comment on or cites a reference for a designated part of the text or table.

Context: Attention is drawn to the footnote by means of a number, mark, etc, in the main body of the text. A footnote generally contains information that is related to but of lesser importance than the larger work in the main body of the text or statistical table. An endnote serves the same purpose as a footnote but is generally located at the end of the text or following the last statistical table.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Flag

143. Frame

Definition: A list, map or other specification of the units which define a population to be completely enumerated or sampled.

Context: The frame consists of previously available descriptions of the objects or material related to the physical field in the form of maps, lists, directories, etc., from which sampling units may be constructed and a set of sampling units selected (Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003). The frame may or may not contain information about the size or other supplementary information about the units, but should have enough details so that a unit, if included in the sample, may be located and taken up for inquiry. The nature of the frame exerts a considerable influence over the structure of a sample survey. It is rarely perfect, and may be inaccurate, incomplete, inadequately described, out of date or subject to some degree of duplication. Reasonable reliability in the frame is a desirable condition for the reliability of a sample survey based on it.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Area sampling
Frame error
Under-coverage

144. Frame error

Definition: Error caused by inherent limitations of input data, or by delays and errors in data acquisition and processing.

Context: Frame errors cover: - coverage errors - erroneous inclusions, omissions and duplications; - classification errors - units not classified, or misclassified by industry, geography or size; - contact errors - units with incomplete or incorrect contact data.

A frame error consists of a difference in the information presented in the frame and the information as it should be. There are various reasons for the differences between this image and the real world. The sources of information used to maintain and update the frame will generally contain irregularities of some sort. The frame may be subject to certain lags in the recording of real world events, or it may have gaps due to the lack of adequate sources for certain types of information. If these distortions of the real world are considered to be acceptable by users of the frame, they should not be considered to be errors. If they are not acceptable, procedures or sources need to be changed or improved.

Source: Lessler, J.T. and Kalsbeek, W.D. (1992), "Non Sampling Error in Survey", New York: John Wiley or US department of Commerce (1978), "Glossary of Non Sampling Error Terms: An Illustration of a Semantic Problem in Statistics", Statistical Policy Working Pape

Hyperlink:

Related Terms: Frame

145. Frequency

Definition: The time interval at which observations occur over a given time period.

Context: If a data series has a constant time interval between its observations, this interval determines the frequency of the series (e.g. monthly, quarterly, yearly). "Frequency" - also called "periodicity" - may refer to several stages in the production process, e.g. in data collection or in data dissemination. (e.g., a time series could be available at annual frequency but the underlying data are compiled monthly). Therefore, "Frequency" can be broken down into "Frequency - data collection" and "Frequency - data dissemination".

For data messages, the frequency is represented through codes. Any additional detail needed (e.g. "weekly on Thursday") must be inserted as free text within "Frequency detail".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Aggregation
Compilation practices
Periodicity
Statistical Data and Metadata Exchange, SDMX

146. Gateway

Definition: An interface between some external source of information and a World Wide Web server. In this instance a gateway is a web enabled search mechanism which allows users to search a distributed network of directory nodes.

Context:

Source: Office for National Statistics (ONS), "National statistics, methods and quality report: Glossary of terms"; unpublished on paper

Hyperlink: http://www.statistics.gov.uk/methods_quality/data_annex.asp

Related Terms:

147. Gateway exchange

Definition: An organized set of bilateral exchanges, in which several data and metadata sending organizations or individuals agree to exchange the collected information with each other in a single, known format, and according to a single, known process.

Context: This pattern has the effect of reducing the burden of managing multiple bilateral exchanges (in data and metadata collection) across the sharing organizations/individuals. This is also a very common process pattern in the statistical area, where communities of institutions agree on ways to gain efficiencies within the scope of their collective responsibilities.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data exchange

148. General Data Dissemination System, GDDS

Definition: A structured process through which member countries of the International Monetary Fund commit voluntarily to improving the quality of the data produced and disseminated by their statistical systems over the long run to meet the needs of macroeconomic analysis.

Context:

Source: International Monetary Fund (IMF), "Guide to the General Data Dissemination System", 2002

Hyperlink: http://dsbb.imf.org/Applications/web/gdds/gddsguidelangs/

Related Terms: Dissemination standard
Special Data Dissemination Standard, SDDS

149. Geographical coverage

Definition: see "Reference area"

150. GESMES

Definition: GESMES (Generic Statistical Message) is a United Nations standard (EDIFACT message) allowing partner institutions to exchange statistical multidimentional arrays in a generic but standardised way. It has been designed by Expert Group 6 (Statistics) of the

Context:

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink: http://www.sdmx.org/Data/GesmesTS_rel3.pdf

Related Terms: Attribute
EDIFACT
Electronic data interchange, EDI
GESMES/CB
GESMES/TS
Statistical message

151. GESMES/CB

Definition: Message profile for data exchange used by the central banking community.

Context: GESMES/CB was developed within the central banking community to facilitate exchange of time-series data in an EDIFACT-syntax format. It was updated to support non-central-banking applications and revised, as GESMES/TS in 2003.

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink: http://www.ecb.int/stats/services/gesmes/html/index.en.html

Related Terms: Data set
GESMES
GESMES/TS

152. GESMES/TS

Definition: GESMES Time Series data exchange message. It is a message (a GESMES profile) allowing the exchange of statistical time series, related attributes and structural definitions using a standardised format.

Context: The GESMES/TS data model is a time-series data exchange model which allows to exchange and identify time series through a multidimensional key and various associated metadata.

GESMES/TS has been renamed from GESMES/CB to GESMES/TS in 2003, reflecting also the adoption of the message by a large statistical community, including the BIS, the ECB, Eurostat, the IMF and OECD.

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink: http://www.ecb.int/stats/services/gesmes/html/index.en.html

Related Terms:
Attachment level
Attribute
Code list
Data exchange
Dimension
GESMES
GESMES/CB
Maintenance agency
Sibling group
Statistical concept
Structural definition
Structural metadata

153. Glossary

Definition: An alphabetised list of terms with definitions.

Context: Glossaries are normally created by an organisation (or by a group of organisations) to reflect its needs or those of its customers.

A glossary commonly contains an explanation of words, concepts or terms that are usually listed in alphabetical order (Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000, available at: http://www.unece.org/stats/publications/53metadaterminology.pdf).

Examples of statistical glossary databases are Eurostat's CODED Glossary (available at http://forum.europa.eu.int/irc/dsis/coded/info/data/coded/en.htm) and the OECD Glossary of Statistical Terms (available at http://cs3-hq.oecd.org/scripts/stats/glossary/index.htm).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Dublin core
Keyword
Maintenance agency

154. Graphical data editing

Definition: Use of graphs to identify anomalies in data.

Context: While such graphical methods can employ paper, the more sophisticated use powerful interactive methods that interconnect groups of graphs automatically and retrieve detailed records for manual review and editing.

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms: Data checking
Data editing

155. Grossing / Netting

Definition: Form of consolidation used in presenting the data.

Context: Combinations in which all statistical items are shown for their full values are called "gross" recordings. Combinations whereby the values of some elementary items are offset against items on the other side of the account or which have an opposite sign are called "net" or consolidated recordings.

Individual units or sectors may have the same kind of transactions both as a use and as a resource (e.g., they both pay and receive interest) and the same kind of financial instrument as an asset and as a liability.

Examples of the application of this concept include gross versus net domestic product (GDP less consumption of fixed capital), and various consolidations across units in presentations of statements of operations and balance sheets for general government and for financial corporations, among others.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

156. Grossing up

Definition: Activity aimed at transforming, based on statistical methodology, micro-data from samples into aggregate-level information representative for the target population.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

157. Guidelines

Definition: Directions or principles used in the development, maintenance and application of rules.

Context: Guidelines may or may not be necessarily mandatory, but are provided as an aid to interpretation and use of rules.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

158. Hierarchy

Definition: Classification structure arranged in levels of detail from the broadest to the most detailed level. Each level of the classification is defined in terms of the categories at the next lower level of the classification.

Context: In SDMX this is known as a level based hierarchy. SDMX also has the concept of the value based hierarchy where the hierarchy of categories are not organised into formal levels.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Level
Structure

159. Identifier

Definition: A sequence of characters, capable of uniquely identifying that with which it is associated, within a specified context.

Context: A name should not be used as an identifier because it is not linguistically neutral.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Country identifier
Data provider series key
Data set identifier
Originator data identifier
ISO/IEC 11179
Organisation identifier

160. Imputation

Definition: Procedure for entering a value for a specific data item where the response is missing or unusable.

Context: Imputation is the process used to determine and assign replacement values for missing, invalid or inconsistent data. This can be done by changing some of the responses or assigning values when they are missing on the record being edited to ensure that estimates are of high quality and that a plausible, internally consistent record is created. (Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 41, available at: http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1)

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink:

Related Terms: Missing data

161. Index type

Definition: The type of index number used in the statistical production process.

Context: Index type refers to the various indices used in the statistical production process (Laspeyres, modified Laspeyres, Paasche, Value-Added, Fisher, Tornqvist or other indexes). Important features in the construction of an index number are its coverage, base period, weighting system and method of averaging statistical results.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Base weight
Chain index
Compilation practices
Weight period

162. Indicator

see "Statistical indicator"

163. Information

Definition: Knowledge concerning any objects such as facts, events, things, processes or ideas including concepts that within a certain context have a particular meaning.

Context:

Source: ISO/IEC 2382-1; 1992 - Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:

164. Information system

Definition: A system which supports decision-making concerning some piece of reality, the object system, by giving decision makers access to information concerning relevant aspects of the object system and its environment.

Context: A "statistical information system" is the information system oriented towards the collection, storage, transformation and distribution of statistical information.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:

165. Institutional mandate

Definition: 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

Context: It also includes arrangements or procedures to facilitate data sharing and coordination between data producing agencies.

This concept can be further broken down into: Institutional mandate - data sharing; Institutional mandate - legal acts and other agreements; Institutional mandate - respondent relations.

"Data sharing" refers to the arrangements or procedures for data sharing and coordination between data producing agencies. "Legal acts and other agreements" refers to the 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. "Respondent relations" refers to the measures to encourage statistical reporting and/or to sanction non-reporting.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

166. Institutional sector

Definition: An aggregation of institutional units on the basis of the type of producer and depending on their principal activity and function, which are considered to be indicative of their economic behaviour.

Context: A sector is divided into sub-sectors according to the criteria relevant to that sector; this permits a more precise description of the economic behaviour of the units.

Source: Eurostat, "European System of Accounts - ESA 1995", Office for Official Publications of the European Communities, Luxembourg, 1996, 2.17-2.18

Hyperlink:

Related Terms: Institutional unit

167. Institutional unit

Definition: The elementary economic decision-making centre characterised by uniformity of behaviour and decision-making autonomy in the exercise of its principal function. A resident unit is regarded as constituting an institutional unit if it has decision-making aut

Context: A resident unit is regarded as constituting an institutional unit if it has decision-making autonomy in respect of its principal function and either keeps a complete set of accounts or it would be possible and meaningful, from both an economic and legal viewpoint, to compile a complete set of accounts if they were required.

The need for aggregation means that it is impossible to consider individual institutional units separately; they must be combined into groups called institutional sectors or simply sectors, some of which are divided into sub-sectors.

The System of National Accounts 1993 states that "Institutional units are grouped together to form institutional sectors, on the basis of their principal functions, behaviour, and objectives". (United Nations, "System of National Account (SNA) 1993", par. 2.20).

Source: Eurostat, "European System of Accounts - ESA 1995", Office for Official Publications of the European Communities, Luxembourg, 1996, 2.12

Hyperlink:

Related Terms: Classification version
Institutional sector

168. Integrity

Definition: Values and related practices that maintain confidence in the eyes of users in the agency producing statistics and ultimately in the statistical product.

Context: Under the SDDS framework, "integrity" is the third of four dimensions of the standard (i.e., data, access, integrity, and quality) for which evidence of a subscribing member's observance of the standard can be obtained.

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.

One important aspect, in integrity, is the trust in the objectivity of statistics. It implies that professionalism should guide policies and practices and it is supported by ethical standards and by transparency of policies and practices.

Source: International Monetary Fund, "Data Quality Assessment Framework - DQAF - Glossary", unpublished

Hyperlink:

Related Terms: Accessibility
Ministerial commentary
Pre-release access
Professionalism
Revision policy
Special Data Dissemination Standard, SDDS

169. Internal access

see "Pre-release access"

170. International code designator

Definition: An identifier of an organization identification scheme.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: ISO/IEC 11179

171. International statistical standard

Definition: The comprehensive body of international statistical guidelines and recommendations that have been developed by international organisations working with national agencies.

Context: The formulation of international statistical standards necessarily entails an extensive process of consultation and discussion between international organisations and between international organisations and their member countries. The standards cover almost every field of statistical endeavour from data collection, processing and dissemination and almost very statistical subject. Such standards also include international statistical

classifications. The most comprehensive database of existing international statistical guidelines and recommendations is maintained on the United Nations Statistical Division website, the Methodological publications in statistics. This database also lists standards currently being developed by international organisations.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Statistical concept
Statistical standard

172. Interpolation

Definition: The use of a formula to estimate an intermediate data value.

Context: A common example is the quarterly estimation of output of non-profit institutions serving households (NPISH) from annual national accounts. A quarterly pattern for interpolation may be derived:

- from previous (discontinued) survey data;

- from proxy variables;

- as a smooth mathematical function.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Benchmarking

173. Interpretability

See "Clarity"

174. Interviewer error

Definition: Effects on respondents' answers stemming from the different ways that interviewers administer the same survey.

Context: Examples of these errors include the failure to read the question correctly (leading to response errors by the respondent), delivery of the question with an intonation that influences the respondent's choice of answer, and failure to record the respondent's answer correctly.

Source: Paul P. Biemer, Robert M.Groves, Lars E. Lyberg, Nancy A.Mathiowetz, Seymour Sudman, "Measurement errors in survey", John Wiley &Sons, 1991

Hyperlink:

Related Terms:

175. ISO/IEC 11179

Definition: International ISO/IEC Standard on metadata registries addressing the semantics of data, the representation of data, and the registration of the descriptions of data.

Context: ISO/IEC 11179 specifies the kind and quality of metadata necessary to describe data, and it specifies the management and administration of those metadata in a registry. The purposes of the standard are to promote a standard description of data, a common understanding of data across organizational elements and between organizations, re-use and standardization of data components.

ISO/IEC 11179 is a six-part standard:

Part 1 - Framework - Contains an overview of the standard and describes the basic concepts

Part 2 - Classification - Describes how to manage a classification scheme in a metadata registry

Part 3 - Registry metamodel and basic attributes - Provides the basic conceptual model, including the basic attributes and relationships, for a metadata registry

Part 4 - Formulation of data definitions - Rules and guidelines for forming quality definitions for data elements and their components

Part 5 - Naming and identification principles - Describes how to form conventions for naming data elements and their components

Part 6 - Registration - Specifies the roles and requirements for the registration process in an ISO/IEC 11179 metadata registry

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Administrative item
Administrative record
Attribute
Characteristics
Class
Concept
Conceptual data model
Conceptual domain
Contact
Context
Creation date
Data element
Data model
Data type
Date of last change
Definition
Dimensionality
Entity
Identifier
International code designator
Keyword
Language
Metadata

Metadata item
Metadata object
Metadata set
Metamodel
Name
Object
Object class
Organisation
Organisation identifier
Permissible value
Permitted value
Property
Reference document
Register
Registrar
Registration
Registration authority
Registry item
Registry metamodel
Related data reference
Related metadata reference
Relationship
Semantics
Special Data Dissemination Standard, SDDS
Submitting organization
Syntax
Taxonomy
Thesaurus
Unit of measure
Value domain
Value item
Value meaning

176. Item response rate

Definition: The ratio of the number of eligible units responding to an item to the number of responding units eligible to have responded to the item.

Context:

Source: Madow, W., Nisselson, H., and Olkin, I., "Incomplete Data in Sample Surveys", Academic Press, New York, 1983

Hyperlink:

Related Terms: Non-response rate
Refusal rate
Response rate

177. Key

Definition: Unique identification of a time series or sibling group within a data set.

Context: In the GESMES/TS data model, every time series takes a value for every dimension of the data structure definition ("key family") to which the series

belongs. The meaning attached to the value of one dimension is not permitted to depend upon the values of any other dimensions.

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink:

Related Terms: Data provider series key
Dimension
Key structure
Sibling group
Time series

178. Key family

See "Data structure definition"

179. Key structure

Definition: An ordered set of coded statistical concepts whose combination of values (dimension values) uniquely identifies each time series within a data set.

Context:

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink: http://www.sdmx.org/Data/GesmesTS_rel3.pdf

Related Terms: Dimension
Key
Statistical concept

180. Keyword

Definition: One or more significant words used for retrieval of any data or metadata element.

Context: In general terms, a keyword is a word used for linking to certain classified objects, for instance in a registry or in a website.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Code list
Data element
Glossary
ISO/IEC 11179

181. Language

Definition: A system of signs for communication, usually consisting of a vocabulary and rules [ISO 5127:2001, 1.1.2.01]

Context:

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Special Data Dissemination Standard, SDDS

182. Level

Definition: A group of codes which are characterised by homogeneous coding, and where the parent of each code in the group is at the same higher level of the hierarchy.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Code
Hierarchy

183. Macro editing

Definition: A procedure for tracking suspicious data by checking aggregates or applying statistical methods on all records or on a subset of them.

Context: A macro-edit detects individual errors by checks on aggregated data, or checks applied to the whole body of records. The checks are typically based on the models, either graphical or numerical formula based, that determine the impact of specific fields in individual records on the aggregate estimates. (Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000, available at http://www.unece.org/stats/publications/editingglossary.pdf

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, May 2002

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms: Data checking
Data editing
Micro editing

184. Macrodata

Definition: The result of a statistical transformation process in the form of aggregated information.

Context: A macrodata set is normally checked and if necessary adjusted, so that it is ready for dissemination as authoritative results of official statistics.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

185. Maintenance agency

Definition: The organisation or other expert body that maintains a domain-specific data or metadata structure definitions.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Classification
Data consumer
Data structure definition
GESMES/TS
Glossary
Structural definition

186. Matching

Definition: The linkage of micro-data from different sources based on common features present in those sources.

Context: If these common features include a common reference or identification number, the process can be referred to as exact matching. In exact matching there are two possible outcomes, either the records match exactly on the basis of the common identifier, or they don't. Where common identifiers are not present, or are of poor quality, the alternative is to use other variables common to the sources involved, but with values that are not necessarily unique to a particular record. In such cases, matching routines tend to rely on probabilities to determine which records match, this is referred to as probabilistic matching.

Source: SDMX (2009)

Hyperlink:

Related Terms: Data Integration

187. Measure

Definition: The phenomenon or phenomena to be measured in a data set.

Context: In a data set, the instance of a measure is often called an observation.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data set
Observation
Aggredated data

188. Measurement error

Definition: Error in reading, calculating or recording numerical value.

Context: Measurement errors occur when the response provided differs from the real value. Such errors may be attributable to the respondent, the interviewer, the questionnaire, the collection method or the respondent's record-keeping system. Errors may be random or they may result in a systematic bias if they are not random.

Measurement error in a survey response may result from respondents' confusion, ignorance, carelessness or dishonesty; error attributable to the interviewer, may be a consequence of poor or inadequate training, prior expectations regarding respondents' responses, or deliberate errors; and error attributable to the wording of the questions in the questionnaire, the order or context in which the questions are presented, and the method used to obtain the responses. Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 5)

Source: The Cambridge Dictionary of Statistics, B.S. Everitt, Cambridge University Press, 1998

Hyperlink:

Related Terms: Error of observation

189. Metadata

Definition: Data that defines and describes other data.

Context: For the ISO standard, metadata is defined as data that defines and describes other data and processes. This means that metadata are data that describe other data, and data become metadata when they are used in this way. This happens under particular circumstances and for particular purposes, as no data are always metadata. The set of circumstances and purposes (or perspective) for which some data are used as metadata is called the context. So, metadata are data about data in some context.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Data
Data provider
ISO/IEC 11179
Metadata layer
Metadata registry
Metadata set
Metadataflow definition
Statistical metadata
Statistical metadata system

190. Metadata attribute

See "Attribute"

191. Metadata dimension

Definition: The higher level of the metadata structure which, combined with more detailed elements, forms the basic framework under which data are described.

Context: A metadata framework is usually presented in a hierarchical structure where a small number of "dimensions" (e.g., data, access, integrity and quality in the SDDS format) is further broken down in a series of sub-elements (e.g., coverage, periodicity, and timeliness) to form the basic framework under which data are described.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Special Data Dissemination Standard, SDDS

192. Metadata item

Definition: An instance of a metadata object.

Context: A metadata item has associated attributes, as appropriate for the metadata object it instantiates. Each metadata item can have a distinct status: mandatory (always required), conditional (understood as required under certain specified conditions) and optional (permitted but not required).

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: Attribute
ISO/IEC 11179
Metadata object

Metadata set
Registry
Registry item
Related metadata reference

193. Metadata layer

Definition: A layer in the reference model for standardisation in statistics used to denote the set of attributes related to statistical metainformation.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:
Metadata
Metadata registry
Statistical metadata
Statistical metadata system
Statistical metainformation

194. Metadata object

Definition: An object type defined by a metamodel.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms:
ISO/IEC 11179
Metadata item
Metamodel
Object

195. Metadata registry

Definition: Information system for registering metadata.

Context: Within ISO/IEC 11179, a metadata registry is a database of metadata that supports the functionality of registration. Registration accomplishes three main goals: identification, provenance, and monitoring quality. Identification is accomplished by assigning a unique identifier (within the registry) to each object registered there. Provenance addresses the source of the metadata and the object described. Monitoring quality ensures that the metadata does the job it is designed to do. A metadata registry manages the semantics of data. Understanding data is fundamental to its design, harmonization, standardization, use, re-use, and interchange. The

underlying model is designed to capture all the basic components of the semantics of data, independent of any application or subject matter area. Registration also allows two or more administered items describing identical objects to be identified, and it will identify situations where similar or identical names are in use for administered items that are significantly different in one or more respects.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administered item
ISO/IEC 11179
Metadata
Metadata layer
Registry
Registry item
Registry metamodel
SDMX Registry
Statistical metadata
Submitting organization

196. Metadata repository

Definition: A logically central statistical repository that allows for querying, editing and managing of metadata.

Context: Such a system provides a mechanism for looking up information about statistical products as well as their design, development, and analysis. Too often metadata is scattered, incomplete or missing. Many times the only source for some information is from subject matter experts. The effective and efficient management of statistical metadata greatly increases the usefulness of statistical data. Since metadata is data, it can be stored and retrieved in a repository just as the data it describes is stored and retrieved in a database.

There are many functions for which statistical metadata repositories are designed. Primarily, it is a standard tool for researchers and analysts to locate data and descriptions of surveys. Data dictionaries, record layouts, questionnaires, sample designs, and standard errors are the types of information that are directly available in such a repository. Less obviously, users can compare designs of different surveys and find common information collected by different surveys (United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section 3.4.5, pages 53, 54).

Source: Organisation for Economic Co-operation and Development (OECD), "Main Economic Indicators", monthly

Hyperlink:

Related Terms: Flow data series

197. Metadata set

Definition: Any collection of metadata.

Context:

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Metadata
Metadata item

198. Metadata structure definition

Definition: A collection of metadata concepts, structure and usage when used to collect or disseminate reference metadata.

Context: A reference metadata set also has a set of structural metadata which describes how it is organized. This metadata identifies what reference metadata concepts are being reported, how these concepts relate to each other (typically as hierarchies), what their presentational structure is, how they may be represented (as free text, as coded values, etc.), and with which formal object types they are associated.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Attribute
Concept
Concept scheme
Reference metadata
Structural metadata
Structure

199. Metadata update

Definition: The date on which the metadata element was inserted or modified in the database.

Context: The date of the metadata update may refer to the update of a whole metadata set or to the update of any single metadata item. The update can refer to the file update (with or without change in the content) or to the date on which the metadata have been posted on the web.

Correspondingly, this concept can be broken down into: Metadata update - last certified; Metadata update - last posted; Metadata update - last update.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

200. Metadataflow definition

Definition: A collection of metadata concepts, structure and usage when used to collect or disseminate reference metadata.

Context: A reference metadata set also has a set of structural metadata which describes how it is organized. This metadata identifies what reference metadata concepts are being reported, how these concepts relate to each other (typically as hierarchies), what their presentational structure is, how they may be represented (as free text, as coded values, etc.), and with which formal object types they are associated.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Category scheme
Data flow definition
Definition
Metadata

201. Metamodel

Definition: A data model that specifies one or more other data models.

Context: The metamodel provides a framework for understanding the important metadata that needs to be captured when describing data.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Data model
ISO/IEC 11179
Metadata object
Registry metamodel

202. Methodological soundness

Definition: The extent to which the methodology used to compile statistics complies with the relevant international standards, including the professional standards enshrined in the Fundamental Principles for Official Statistics.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

203. Methodology

Definition: A structured approach to solve a problem.

Context: A set of research methods and techniques applied to a particular field of study (Statistics Canada, Glossary, available at: http://www.statcan.ca/english/edu/power/glossary/gloss.htm).

Advance notice in major changes in methodology, in SDMX, refers to the policy on notifying the public of changes in methodology, indicating whether the public is notified before a methodological change affects disseminated data and, if so, how long before. (http://www.sdmx.org/)

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Documentation on methodology
Dublin core
Statistical methodology
Statistical subject-matter domain

204. Micro editing

Definition: An exhaustive check to find errors by inspecting each individual observation.

Context: Editing done at the record, or questionnaire level.

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, May 2002

Hyperlink:

Related Terms: Data editing
Macro editing

205. Microdata

Definition: Non-aggregated observations, or measurements of characteristics of individual units.

Context: Microdata set is the result of a survey instance or other data collection instance after unit-level editing and imputation and possible matching with other unit-level data. It organizes unit level data so that relationships between individual units and their characteristics can be identified, so as to allow all forms of aggregation.

Context: This is sometimes called an observation register (see, for instance, B. Sundgren: Guidelines for the modelling of statistical data and metadata, UN Geneva and New York, 1995).

Anonymised microdata are individual statistical records which have been modified in order to minimise, in accordance with current best practice, the

risk of identification of the statistical units to which they relate (European Commission Regulation (EC) No 831/2002 of 17 May 2002 implementing Council Regulation (EC) No 322/97 on Community Statistics, concerning access to confidential data for scientific purposes).

Longitudinal microdata sets combine unit-level data from succeeding data collection instances over multiple time periods. Related terms Longitudinal data.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Access
Aggregation

206. Ministerial commentary

Definition: Commentary on the released data provided by the government authority.

Context: Under the SDDS, this entails the identification of any such commentary so as to maintain the objectivity or freedom from political judgement of the official statistical data being disseminated.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Integrity
Pre-release access
Special Data Dissemination Standard, SDDS

207. Misclassification

Definition: Erroneous classification of a subject into a category in which the subject does not belong.

Context:

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms:

208. Missing data

Definition: Observations which were planned and are missing.

Context: Missing data in a survey may occur when there are no data whatsoever for a respondent (non-response) or when some variables for a respondent are unknown (item non-response) because of refusal to provide or failure to collect the response (ISI).

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Imputation
Non-response
Observation

209. Model assumption error

Definition: Error that occurs with the use of methods, such as calibration, generalised regression estimator, calculation based on full scope or constant scope, benchmarking, seasonal adjustment and other models not included in the preceding accuracy components, in o

Context:

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms:

210. Multilateral exchange

Definition: The exchange of statistics and / or metadata between a sending and several receiving organisations for a specific data flow where all parties agree on all aspects of the exchange (including the mechanism for exchange, the formats, the frequency or schedul

Context: This exchange process has the effect of reducing the burden of a sending organisation of managing multiple unique bilateral exchanges of statistics and / or metadata with several receiving organisations. This is also a very common exchange process in the statistical area, where communities of national and international institutes agree on ways to gain efficiencies within the scope of their collective responsibilities. Apart from Multilateral exchange, the SDMX initiative identifies two other basic forms of exchange of statistics and metadata between organisations, i.e. bilateral exchange and data-sharing exchange.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Bilateral exchange
Data exchange
Data sharing

211. Name

Definition: The designation of an object by a linguistic expression.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: ISO/IEC 11179

212. Nomenclature

Definition: A systematic naming of things or a system of names or terms for things. In classification, nomenclature involves a systemic naming of categories or items.

Context: The terms "classification" and "nomenclature" are often used interchangeably, despite the definition of a "classification" being broader than that of a "nomenclature". A nomenclature is essentially a convention for describing observations, whereas a classification structures and codifies the observations as well.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Classification

213. Non-probability sample

Definition: A sample in which the selection of units is based in factors other than random chance, e.g. convenience, prior experience or the judgement of a researcher.

Context:

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms: Probability sample

214. Non-response

Definition: A form of observation present in most surveys, which means failure to obtain a measurement on one or more study variables for one or more elements k selected for the survey.

Context: The term 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 element is never established.

There are two broad types of non-response: first, a sampled unit that is contacted may fail to respond: this represents "unit non-response"; second, the unit may respond to the questionnaire incompletely: this is referred to as "item non-response".

Non-response leads to an increase in variance as a result of a reduction in the actual size of the sample and the recourse to imputation. This produces a bias if the non-respondents have characteristics of interest that are different from those of the respondents. Furthermore, there is a risk of significantly underestimating the sampling error, if imputed data are treated as though they were observed data. (Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 59).

Source: Sarndal C.E., Swensson B., Wretman J., "Model assisted survey sampling", Springer - Verlag, New York, 1992

Hyperlink: http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1

Related Terms:
Follow-up
Missing data
Non-response error
Non-response rate
Observation
Refusal rate
Survey
Weight

215. Non-response bias

See "Non-response error"

216. Non-response error

Definition: Error that occurs when the survey fails to get a response to one, or possibly all, of the questions.

Context: Non-response errors result from a failure to collect complete information on all units in the selected sample. These are known as "unit non- response" and "item non-response".

Non-response errors affect survey results in two ways. First, the decrease in sample size or in the amount of information collected in response to a particular question results in larger standard errors. Second, and perhaps more important, a bias is introduced to the extent that non-respondents differ from respondents within a selected sample.

Non-response errors are determined by collecting any or all of the following: unit response rate, weighted unit response rate, item response rate, item coverage rate, refusal rate, distribution of reason for non response, comparison of data across contacts, link to administrative data for non-respondents, estimate of non-response bias (Statistical Policy Working Paper 15: Quality in Establishment Surveys, Office of Management and Budget, Washington D.C., July 1988, page 68).

Source: Statistics Canada, "Statistics Canada Quality Guidelines", 3rd edition, October 1998.

Hyperlink: http://www.statcan.ca/english/freepub/12-539-XIE/12-539-XIE.pdf

Related Terms: Follow-up
Non-response
Weight

217. Non-response rate

Definition: The proportion of non-response in a sample.

Context:

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Item response rate
Non-response
Refusal rate
Response rate
Sample

218. Non-sampling error

Definition: Error in sample estimates which cannot be attributed to sampling fluctuations.

Context: Non-sampling error may arise from many different sources such as defects in the sampling frame, faulty demarcation of sample units, defects in the selection of sample units, mistakes in the collection of data due to personal variations, misunderstanding, bias, negligence or dishonesty on the part of the investigator or of the interviewer, mistakes at the stage of the processing of the data, etc.

Non- sampling errors may be categorised as:

- Coverage errors (or frame errors) due to divergences between the target population and the frame population ;

- Measurement errors occurring during data collection.

- Nonresponse errors caused by no data collected for a population unit or for some survey variables.

- Processing errors due to errors introduced during data entry, data editing, sometimes coding and imputation.

- Model assumption errors.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Estimate
Sampling error

219. Object

Definition: Anything perceivable or conceivable.

Context: Objects may be material (e.g. an engine, a sheet of paper, a diamond), immaterial (e.g. a conversion ratio, a project plan) or imagined [adapted from ISO 1087-1:2000].

In object-oriented design or programming, an object is a concrete realisation of a class that consists of data and the operations associated with that data. An item that a user can manipulate as a single unit to perform a task.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Attribute
Characteristic
Class
ISO/IEC 11179
Metadata object
Object class
Ontology
Property

220. Object class

Definition: A set of ideas, abstractions, or things in the real world that can be identified with explicit boundaries and meaning and whose properties and behavior follow the same rules.

Context: Object class administration record is the Administration record for an Object class

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: ISO/IEC 11179
Object
Property

221. Objectives

Definition: The purposes for which information is required, stated within the context of the program, research problem or hypotheses that gave rise to the need for information.

Context:

Source: Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 11

Hyperlink: http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1

Related Terms:

222. Observation

Definition: The value, at a particular period, of a particular variable.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Census
Classification
Coverage ratio
Derived statistic
Disaggregation
Measure
Missing data
Non-response
Observation pre-break value
Observation unit
Observation value
Reference period
Statistical concept
Statistical variable
Time period
Time series

223. Observation confidentiality

See "Confidentiality"

224. Observation pre-break value

Definition: The observation, at a time series break period, that was calculated using the old methodology.

Context: At a time series break period, two observations may be recorded: the pre-break value produced on the basis of the old methodology and the post-break value, as measured by the new methodology. SDMX allows for a pre-break value in the case of a series break, where one would use the observation value to show the post-break value.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Observation
Time series
Time series break

225. Observation status

Definition: Information on the quality of a value or an unusual or missing value.

Context: This item is normally coded and uses codes providing information about the status of a value, with respect to events such as "break", "estimated value", "forecast", "missing value", or "provisional value". In some cases, there is more than one event that may have influenced the value (e.g. a break in methodology may be accompanied with the fact that an observation is an estimate).

A textual comment providing more detailed information on important events related to an observation can be added via the attribute "Comment".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

226. Observation unit

Definition: Those entities for which information is received.

Context: During the collection of data, this is the unit for which data is recorded. It should be noted that this may, or may not be, the same as the reporting unit.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Analytical unit
Classification
Entity
Observation
Statistical unit

227. Observation value

Definition: The value of a particular variable at a particular period.

Context: The observation value is the field which holds the data.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Observation

228. Occupation

Definition: Job or position held by an individual who performs a set of tasks and duties.

Context: Occupation refers to the type of work done during the reference period by the person employed (or the type of work done previously, if the person is unemployed), irrespective of the industry or the status in employment in which the person should be classified. Occupation is defined in terms of jobs or posts. "Job" is defined by the International Labour Organisation (ILO) as a set of tasks and duties executed, or meant to be executed, by one person. A set of jobs whose main tasks and duties are characterised by a high degree of similarity constitutes an occupation. Persons are classified by occupation through their relationship to a past, present or future job. The international standard for classification of occupations is the International Standard Classification of Occupations (ISCO). Therefore the concept is normally coded.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

229. Official statistics

Definition: Any statistical activity carried out within a national statistical system, or under the statistical programme of an intergovernmental organization.

Context: Official statistics are compiled in accordance with the Fundamental Principles for Official Statistics, the European Statistics Code of Practice or a similar authoritative framework ensuring minimum professional standards.

National statistical systems comprise the ensemble of statistical organisations and units within a country that jointly collect, process and disseminate official statistics on behalf of national government. They also include the mechanisms of interaction between suppliers, producers, users and other stakeholders.

The global system of official statistics comprises all national and international producers of official statistics.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

230. Ontology

Definition: A formal specification of a conceptualisation, i.e. the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them.

Context: Ontology it is a branch of metaphysics concerned with the nature and

relations of being.

In its general meaning, ontology is the study or concern about what kinds of things exist - what entities there are in the universe. It derives from the Greek onto (being) and logia (written or spoken discourse). In artificial intelligence, an ontology is, according to Tom Gruber, "the specification of conceptualizations, used to help programs and humans share knowledge." In this usage, an ontology is a set of concepts - such as things, events, and relations - that are specified in some way in order to create an agreed-upon vocabulary for exchanging information (http://www-ksl.stanford.edu/kst/what-is-an-ontology.html).

Source: United Nations Statistical Commission and Economic Commission For Europe

Conference Of European Statisticians, Statistical Standards And Studies - No. 53, "Terminology on Statistical Metadata", United Nations, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Concept
Entity
Object
Taxonomy

231. Organisation

Definition: A unique framework of authority within which a person or persons act, or are designated to act, towards some purpose.

Context: International organisations are entities established by formal political agreements between their members that have the status of international treaties; their existence is recognised by law in their member countries; they are not treated as resident institutional units of the countries in which they are located (United Nations, "System of National Account (SNA) 1993", par.4.164).

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Organisation identifier
Organisation role

232. Organisation identifier

Definition: The identifier assigned to an organization within an organization identification scheme, and unique within that scheme.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Identifier
ISO/IEC 11179
Organisation

233. Organisation role

Definition: The function or activities of an organisation, in statistical processes such as collection, processing and dissemination.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Data consumer
Dissemination
Organisation

234. Origin

Definition: The source (document, project, discipline or model) for the Administered item.

Context:

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: Administered item

235. Originator data identifier

Definition: The data identifier as found in the originating database.

Context: A unique identifier should enable data producers to recognise the database where the data were stored.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administered item
Data set identifier

236. Out-of-scope units

Definition: Units that should not be included in the sampling frame because they do not belong to the target population in the reference period.

Context: If enumerated, out of scope units cause over-coverage.

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms: Over-coverage
Scope

237. Over-coverage

Definition: Error due to the inclusion in the sample of elements that do not belong there.

Context: Over-coverage arises from the presence in the frame of units not belonging to the target population and of units belonging to the target population that appear in the frame more than once (Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003).

Source: United States Federal Committee on Statistical Methodology, "Statistical Policy Working Paper 4 - Glossary of Non-sampling Error Terms: An Illustration of a Semantic Problem in Statistics", 1978

Hyperlink: http://www.fcsm.gov/working-papers/sw4.html

Related Terms: Out-of-scope units
Under-coverage

238. Period

Definition: The time interval of single repetition of a varying quantity of a motion or phenomenon which repeats itself regularly.

Context: The period is the reciprocal of the frequency. More loosely, the expression is used to denote the time interval or average interval between identifiable points of recurrence, e.g. between peaks or troughs of the series (month, quarter, year,).

In GESMES/TS, a period is a time reference.

Source: McGraw-Hill Encyclopedia of Science and Technology

Hyperlink:

Related Terms: Reference period
Time period
Weight period

239. Periodicity

Definition: Frequency of compilation of the data.

Context: Periodicity refers to the frequency of compilation of the data (e.g., a time series could be available at annual frequency but the underlying data are compiled monthly, thus have a monthly periodicity). The periodicity of a particular data category is determined by several factors, including the ease of observation or compilation and the needs of analysis. Periodicity is usually expressed in terms of divisions of the calendar (e.g. monthly, quarterly).

Periodicity of original data refers to the frequency of compilation of data by the source agency, i.e. the national agency or international organisation that provided the information. This agency may or may not be the agency responsible for the original collection of the data.

Source: International Monetary Fund (IMF), "Guide to the Data Dissemination Standards, Module 1: The Special Data Dissemination Standard", Washington, May 1996

Hyperlink: http://dsbb.imf.org/Applications/web/gdds/gddsguidelangs/

Related Terms: Data
Frequency
Special Data Dissemination Standard, SDDS

240. Permissible value

Definition: Designation of a value meaning within a specific value domain.

Context: Permissible value meaning is an expression of a value meaning allowed in a specific value domain.

Attributes of Permissible value according to ISO/IEC 11179:

"Permissible value begin date" is the date this value became/becomes allowed in the Value domain. A Registration authority may determine whether this date is the date the value becomes valid in a registry or the date the value becomes part of the source domain or some other date.

"Permissible value end date" is the date this value became/becomes no longer allowed in the Value domain. A Registration authority may determine whether this date is the date the value becomes no longer valid in a registry or the date the value becomes no longer part of the source domain or some other date.

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: Conceptual domain
ISO/IEC 11179
Permitted value
Registration authority
Value domain
Value meaning

241. Permitted value

Definition: Representation of a value meaning in a specific value domain – the actual value.

Context:

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Permissible value
Value domain

242. Pre-break observation

See "Pre-break value"

243. Pre-break value

See "Observation Pre-break value"

244. Precision

Definition: The property of the set of measurements of being very reproducible or of an estimate of having small random error of estimation.

Context: Precision is to be contrasted with accuracy, which is the property of being close to a target or true value.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Accuracy
Estimation

245. Pre-release access

Definition: The practice of giving certain individuals or organisations access to data under embargo before those data are released to the public

Context: This entails the 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. Also called "internal access" in the Data Quality Assessment Framework of the IMF.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Integrity
Ministerial commentary
Revision policy
Special Data Dissemination Standard, SDDS

246. Prerequisites of quality

Definition: Institutional conditions for the pursuit of data quality.

Context: The Data Quality Assessment Framework (DQAF) groups the indicators of this kind into three elements: legal and institutional environment, resources, and quality awareness.

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. Typically, these pointers refer to the larger institution (called the "umbrella institution" in the DQAF) of which the compiling unit, such as a national accounts division or a balance of payments department, is a part. Further, these prerequisites typically influence more than one of the five dimensions in the DQAF.

Source: International Monetary Fund (IMF), "Data Quality Assessment Framework (DQAF) Glossary"

Hyperlink:

Related Terms:

247. Primary source

Definition: The organisation or individual responsible for the collection and aggregation of data from the initial supplier.

Context: For information derived from surveys or censuses, such data comprise unit record information about individual entities. For administrative data, the primary source is the agency responsible for the compilation of data from individual persons or organisations to meet administrative or regulatory requirements.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Secondary supplier

248. Probability sample

Definition: A sample selected by a method based on the theory of probability (random process), that is, by a method involving knowledge of the likelihood of any unit being selected.

Context:

Source: United Nations Statistics Division, "Handbook of Vital Statistics Systems and Methods, Volume 1: Legal, Organisational and Technical Aspects", Studies in Methods, Series F, No. 35, United Nations, New York, 1991

Hyperlink:

Related Terms: Non-probability sample
Sample

249. Processing error

Definition: The error in final survey results arising from the faulty implementation of correctly planned implementation methods.

Context: Sources of processing errors include all post-collection operations, as well as the printing of questionnaires. Most processing errors occur in data for individual units, although errors can also be introduced in the implementation of systems and estimates.

In survey data, for example, processing errors may include transcription errors, coding errors, data entry errors and errors of arithmetic in tabulation.

Source: United States Federal Committee on Statistical Methodology, "Statistical Policy Working Paper 15: Quality in Establishment Surveys", Washington D.C., July 1988, page 79

Hyperlink: http://www.fcsm.gov/working-papers/wp15.html

Related Terms: Data processing
Survey

250. Product

Definition: Goods and services exchanged and used for various purposes, as inputs in the production of other goods and services, as final consumption or for investment.

Context: Products can be both the subject of statistics on production, and the outcome of statistical processes, for example published results.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

251. Professionalism

Definition: The standard, skill and ability suitable for producing statistics of good quality.

Context: 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 (Fundamental Principles of Official Statistics, principle 2).

"Professionalism" describes the elements providing assurances that: statistics are produced on an impartial basis; elements providing assurances that the choices of sources and statistical techniques as well as decisions about dissemination are informed solely by statistical considerations; elements providing assurances that the recruitment and promotion of staff are based on relevant aptitude; elements providing assurances that the statistical entity is entitled to comment on erroneous interpretation and misuse of statistics, guidelines for staff behaviour and procedures used to make these guidelines known to staff; other practices that provide assurances of the independence, integrity, and accountability of the statistical agency.

This concept can be further broken down into: Professionalism - code of conduct; Professionalism - impartiality; Professionalism - methodology; Professionalism - statistical commentary.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Integrity
Statistical Data and Metadata Exchange, SDMX

252. Property

Definition: A characteristic common to all members of an object class.

Context: Within SDMX, an "attribute property" allows ad hoc simple metadata concepts, such as URL, to be specified for a metadata attribute, within the context of a metadata structure definition.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: ISO/IEC 11179
Object
Object class

253. Provider load

Definition: The effort, in terms of time and cost, required for respondents to provide satisfactory answers to a survey.

Context: This burden can lead to providers experiencing annoyance, anger, frustration,

etc., at being requested to participate, with escalation of these feelings generated by the complexity, length and/or frequency of surveys. The terms "respondent burden" and "respondent load" are also used to describe provider load.

Source: Australian Bureau of Statistics, Service Industries Statistics, "Glossary of Terms"; unpublished on paper

Hyperlink: http://www.abs.gov.au/CA25670D007E9EA1/0/DB35F160E9383A1FCA256B650006C3D0?Open&Highlight=0,Glossary

Related Terms:

254. Provision agreement

Definition: Arrangement within which the provider supplies data or metadata.

Context: The agreement may define the scope of the data or metadata that can be provided.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

255. Public disclosure

Definition: The act of making information or data readily accessible and available to all interested individuals and institutions.

Context: Some examples of the different forms that public disclosure may take include: verbal or written statements released to a public forum, to the news media, or to the general public; publication in an official bulletin, gazette, report, or stand-alone document; and information posted on a website.

Source: Code of Good Practices on Transparency in Monetary and Financial Policies, Part 1-Introduction; approved by the IMF Executive Board on July 24, 2000

Hyperlink: http://www.imf.org/external/np/mae/mft/sup/part1.htm#appendix_III

Related Terms:

256. Punctuality

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

Context: Punctuality may be calculated, for instance, with reference to target dates announced in an official release calendar, laid down by regulations or previously agreed among partners.

In quality assessment, punctuality is often associated with timeliness, which refers to the time lag between the end of the reference period and the release of data.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: SDMX, Statistical Data and Metadata Exchange
Timeliness

257. Qualitative data

Definition: Data describing the attributes or properties that an object possesses.

Context: The properties are categorized into classes that may be assigned numeric values. However, there is no significance to the data values themselves; they simply represent attributes of the object concerned.

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://amrads.jrc.cec.eu.int/k-base/glossary/glossALL.htm

Related Terms: Quantitative data

258. Quality

Definition: The degree to which a set of inherent characteristics fulfils requirements.

Context: Quality is a multi-faceted concept. The dimensions of quality that are considered most important depend on user perspectives, needs and priorities, which vary across groups of users. Several statistical organisations have developed lists of quality dimensions, which, for international organisations, are being harmonised under the leadership of the Committee for the Coordination of Statistical Activities (CCSA).

A generic list would include the following dimensions, all of which are defined elsewhere in the Metadata Common Vocabulary:
Relevance
Accuracy
Timeliness
Punctuality
Accessibility
Clarity / interpretability
Comparability
Coherence
Integrity
Credibility
Methodological soundness

Source: ISO 9000/2005: Quality Management and Quality Assurance Vocabulary

Hyperlink:

Related Terms: Accuracy
Quality differences
Quality management
Relevance
Reliability

259. Quality control survey

Definition: A survey usually carried out on a small scale by experienced staff in order to validate the results of a larger survey or statistics compiled from non-statistical sources.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

260. Quality differences

Definition: Differences in the various dimensions of data quality promulgated by international organisations and national agencies.

Context: Comparisons of these dimensions may be made for data between countries, for the same series over time or between the same series compiled by different agencies in the same country.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Quality

261. Quality index

Definition: A one-dimension synthetical information on quality, possibly calculated as a weighted mean of all available quality indicators.

Context:

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms:

262. Quality management

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

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 the resources meet the requirement.

This concept can be broken down into: "Quality management - quality assurance"; "Quality management - assessment"; "Quality management - documentation".

"Quality assurance" refers to all the planned and systematic activities implemented that can be demonstrated to provide confidence that the processes will fulfil the requirements for the statistical output. This includes the design of programmes for quality management, the description of planning process, scheduling of work, frequency of plan updates, and other organisational arrangements to support and maintain planning function.

"Quality assessment" contains the overall assessment of data quality, based on standard quality criteria. This may include the result of a scoring or grading process for quality. Scoring may be quantitative or qualitative.

"Quality documentation" contains documentation on methods and standards for assessing data quality, based on standard quality criteria such as relevance, accuracy and reliability, timeliness and punctuality, accessibility and clarity, comparability, and coherence.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Quality

263. Quantitative data

Definition: Data expressing a certain quantity, amount or range.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/editingglossary.pdf

Related Terms: Flag
Qualitative data

264. Questionnaire

Definition: A group or sequence of questions designed to elicit information on a subject, or sequence of subjects, from a reporting unit or from another producer of official statistics.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Questionnaire design

Schedule
Survey

265. Questionnaire design

Definition: The design (text, order, and conditions for skipping) of the questions used to obtain the data needed for the survey.

Context:

Source: United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section 3.3.17, page 26

Hyperlink: http://www.census.gov/srd/www/metadata/metada18.pdf

Related Terms: Questionnaire
Survey design

266. Ratio estimation

Definition: The use of known population totals for auxiliary variables to improve the weighting from sample values to population estimates.

Context: Ratio estimation operates by comparing the survey sample estimate for an auxiliary variable with the known populati The ratio weights are given by X/x (where X is the known population total for the auxiliary variable, and x is the corresponding estimate of the total based on all responding units in the sample). These weights assume that the population total for the variable of interest will be estimated by the sample equally as well (or poorly) as the population total for the auxiliary variable is estimated by the sample. Ratio estimation can be more accurate than number-raised estimation if the auxiliary variable is highly correlated with the variable of interest. However it is slightly biased, with the bias increasing for smaller sample sizes and where there is lower correlation between the auxiliary

Source: Australian Bureau of Statistics, Statistical Concepts Library, "Labour Statistics: Concepts, Sources and Methods", Chapter 16 - Overview of Survey Methods, Canberra, 2001

Hyperlink:

Related Terms: Estimation
Weight

267. Receiving organisation

Definition: Agency that receives data from one or more sending agencies during a data exchange.

Context: There can be more than one receiving agency for a single data exchange.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

268. Recommended use of data

Definition: Text that is intended to provide users with explicit information on the appropriate use of the statistics within the limitations imposed by the definition or main concepts, scope and coverage, collection methodology, etc.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

269. Record check

Definition: A study in which data on individual units, obtained by one method of data collection, are checked against data for the same units from available records, obtained by a different method of data collection.

Context: An example of record check is the comparison of ages as reported in a census with information from birth certificates.

Source: Lessler, J.T. and Kalsbeek, W.D. (1992), "Non Sampling Error in Survey", New York: John Wiley or US department of Commerce (1978), "Glossary of Non Sampling Error Terms: An Illustration of a Semantic Problem in Statistics", Statistical Policy Working Pape

Hyperlink:

Related Terms:

270. Recording basis

Definition: Processes and standards employed in calculating statistical aggregates.

Context: The recording of transactions covers a broad range of processes and accounting conventions, including types of valuation, prices, conversion rates, the accounting basis, units of measurement used in data collection, etc. It also refers to descriptions of the time of recording (e.g. cash or accrual basis) employed. The description may also include how consistent the practices used are with internationally accepted standards, guidelines, or good practices.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

271. Recording of transactions

Definition: A broad range of processes and standards employed in calculating statistical aggregates.

Context: The conventions include types of valuation, prices, conversion rates, the accounting basis, units of measurement used in data collection, etc.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

272. Record-keeping error

Definition: An error which arises from inaccuracy in the records used for responses.

Context:

Source: Lessler, J.T. and Kalsbeek, W.D. (1992), "Non Sampling Error in Survey", New York: John Wiley or US department of Commerce (1978), "Glossary of Non Sampling Error Terms: An Illustration of a Semantic Problem in Statistics", Statistical Policy Working Pape

Hyperlink:

Related Terms:

273. Reference area

Definition: The country or geographic area to which the measured statistical phenomenon relates.

Context: The concept refers to the country, geographical or political group of countries or regions within a country.

The concept is subject to a variety of hierarchies, as countries comprise territorial entities that are States (as understood by international law and practice), regions and other territorial entities that are not States (such as Hong Kong) but for which statistical data are produced internationally on a separate and independent basis.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Counterpart reference area
Geographical coverage

274. Reference document

Definition: A document that provides pertinent details for consultation about a subject.

Context: Attributes of Reference document:

Reference document identifier is an identifier for the Reference document

Reference document language identifier is the identifier of the natural or special language used in the Reference document

Reference document title is the title of the Reference document

Reference document type description is a description of the type of Reference document.

Reference documents can be publications (hardcopy, electronic), other databases (internal, external), Internet (Internet addresses), methodological references (for instance to summary metadata, detailed metadata, information on major changes).

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179

275. Reference metadata

Definition: Metadata describing the contents and the quality of the statistical data.

Context: Preferably, 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 (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Concept scheme
Cross-domain Concepts
Metadata structure definition
SDMX Registry
Statistical metadata
Structural metadata

276. Reference period

Definition: The period of time or point in time to which the measured observation is intended to refer.

Context: In many cases, the reference period and time period will be identical, but there are also cases where they are different. This can happen if data are not available for the target reference period, but are available for a time period which is judged to be sufficiently close. For example, the reference

period may be a calendar year, whereas data may only be available for a fiscal year. In such cases, “reference period” should refer to the target reference period rather than the actual time period of the data. The difference between target and actual reference period can be highlighted in a free text note.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Accounting conventions
Base period
Compilation practices
Observation
Period
Statistical processing
Time period

277. Reference time

See "Reference period"

278. Refusal rate

Definition: The proportion of observation units for which the reporting unit has been successfully contacted, but has refused to give the information sought.

Context: The proportion is usually and preferably calculated by dividing the number of refusals by the total number of the sample which was originally desired to achieve.

Observation and reporting units are often (but not always) the same entity.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Item response rate
Non-response
Non-response rate
Response rate

279. Register

Definition: Data store where registered items are recorded and managed.

Context: A register is a written and complete record containing regular entries of items and details on particular set of objects (Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000, http://www.unece.org/stats/publications/53metadaterminology.pdf).

A statistical register is a regularly updated list of units and their

characteristics to be used for statistical purposes. In some cases statistical registers may hold characteristics that are deemed by the appropriate legislation not to be subject to statistical confidentiality. In such cases, there may be non-statistical uses of data relating to these characteristics.

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Registration authority

280. Registrar

Definition: A representative of a Registration authority.

Context: Registrar contact is the contact information associated with a Registrar. Registrar identifier is an identifier for the Registrar.

Source: ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, March 2004

Hyperlink:

Related Terms: ISO/IEC 11179
Registration authority
SDMX Registry

281. Registration

Definition: The records relating to an administered item and its transactions with the registration authority with respect to a specific register.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administered item
ISO/IEC 11179
Registration authority
Registry
SDMX Registry

282. Registration authority

Definition: Organization responsible for maintaining a register.

Context: Registration authority identifier is an identifier assigned to a registration authority. Registration authority registrar is the relationship between a Registration Authority and a Registrar. A registration status is a designation of the status in the registration life-cycle of an administered item.

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Permissible value
Register
Registrar
Registration

283. Registry

Definition: An application which stores metadata for querying, and which can be used by any other application in the network with sufficient access privileges.

Context: A registry can be understood as the index of a database or metadata repository which is made up of all the data providers' data sets and reference metadata sets within a statistical community, distributed across the Internet or similar network. Note that the registry services are not concerned with the storage of data. The registry services concern themselves with providing visibility of the data and reference metadata, and information needed to access the data and reference metadata.

Querying: The registry has interfaces for querying the metadata it contains, so that applications and users can discover the existence of data sets and reference metadata sets, structural metadata, the providers/agencies associated with those objects, and the provider agreements which describe how the data and metadata are made available, and how they are categorized.

Subscription/Notification: It is possible to subscribe to specific objects in the registry, so that a notification will be sent to all subscribers whenever the registry objects are updated.

Registration (structural metadata submission): A registry service which allows users to inform the registry that data sets, reference metadata sets, structural metadata, or data provisioning information.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Metadata item
Metadata registry
Registration

284. Registry item

Definition: Metadata item recorded in a Registry.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Administered item
ISO/IEC 11179
Metadata item
Metadata registry

285. Registry metamodel

Definition: A metamodel specifying a metadata registry.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: ISO/IEC 11179
Metadata registry
Metamodel

286. Related data reference

Definition: A reference between a data element and any related data.

Context:

Source: ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, March 2004

Hyperlink:

Related Terms: Data element
ISO/IEC 11179

287. Related metadata reference

Definition: A reference from one metadata item to another.

Context: A Registration Authority could choose to use a Reference Document, an administrative note or an explanatory comment to record a related metadata reference.

Source: ISO/IEC Committee Draft 11179-3: 2007, Information technology - Metadata Registries (MDR) - Part 3: Registry Metamodel and basic attributes, August 2007

Hyperlink:

Related Terms: ISO/IEC 11179
Metadata item

288. Relationship

Definition: A connection among model elements.

Context: In ISO/IEC 11179-3, "Registry metamodel and basic attributes", a relationship is either an association or a generalization.

"Association" is a semantic relationship between two classes. "Generalization" is a relationship between a more general class (the parent) and a more specific class (the child) that is fully consistent with the first class (i.e. it has all of its attributes and relationships) and that adds additional information. A relationship between a Data element example and its Data element is called an "Exemplification".

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: ISO/IEC 11179

289. Release policy

Definition: Rules for disseminating statistical data to interested parties.

Context: Describes the policy for release of the data to the public, how the public is informed that the data are being released, and whether the data are disseminated to all interested parties at the same time.

This concept can be broken down into: Release policy - legal acts and other agreements; Release policy - policy commentary; Release policy - release calendar; Release policy - release calendar access; Release policy - transparency; Release policy - user access.

"Legal acts and other agreements" refers to the legal acts and other agreements pertaining to data access. It involves the description of the legal and institutional framework defining which users have access to what data, on what conditions, and on what time schedule. "Policy commentary" refers to the fact whether or not ministerial commentary is provided on the occasion of statistical release. This concept relates to policy commentary that other authorities might add to the data released to the public.

"Release calendar" refers to the schedule of statistical release dates. An advance release calendar is the schedule for release of data, which is publicly disseminated so as to provide prior notice of the precise release dates on which a national statistical agency, other national agency, or international organization undertakes to release specified statistical information to the public.

"Release calendar access" refers to the access to the release calendar information (how the calendar can be accessed).

"Transparency" refers to the dissemination of the release policy to the public. It concerns whether the description of the release policy is disseminated to the public and by what modality, but not the description of the release policy itself.

"User access" refers to the policy for release of the data to users, the scope of dissemination (eg, to the public, to selected users), how users are informed that the data are being released, and whether the policy provides

for the dissemination of statistical data to all users at the same time. It also describes the policy for briefing the press in advance of the release of the data.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Timeliness

290. Relevance

Definition: The degree to which statistics meet current and potential users' needs.

Context: Relevance is concerned with whether the available information sheds light on the issues that are important to users. Assessing relevance is subjective and depends upon the varying needs of users. The Agency's challenge is to weight and balance the conflicting needs of current and potential users to produce statistics that satisfy the most important needs within given resource constraints. In assessing relevance, one approach is to gauge relevance directly, by polling users about the data. 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. Relevance refers to the processes for monitoring the relevance and practical usefulness of existing statistics in meeting users" needs and how these processes impact the development of statistical programmes.

This concept can be broken down into: "Relevance - completeness"; "Relevance - user needs"; "Relevance - user satisfaction".

"Completeness" refers to the the extent to which all statistics that are needed are available. The measurement of the availability of the necessary statistics normally refers to data sets and compares the required data set to the available one.

"User Needs" refers to the description of users and their respective needs with respect to the statistical data. The main users (e.g. official authorities, the public or others) and user needs should be stated, e.g. official authorities with the needs for policy indicators, national users, etc.

"User Satisfaction" refers to the measure to determine user satisfaction. This concerns how well the disseminated statistics meet the expressed user needs. If user satisfaction surveys have been conducted, the domain manager should mention them. Otherwise, any other indication or measure to determine user satisfaction might be used.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Quality
Serviceability

291. Reliability

Definition: Closeness of the initial estimated value to the subsequent estimated value.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Accuracy
Estimate
Quality

292. Reporting agency

Definition: The organisation that supplies the data for a given instance of the statistics.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Reporting unit

293. Reporting unit

Definition: The unit that supplies the data for a given survey instance.

Context: If an accountant provides data for several businesses, each of which have been selected for a statistical survey, the accountant is the reporting unit, whereas the individual businesses are observation units.

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Reporting agency
Survey

294. Respondent burden

See "Provider load"

295. Respondent load

See "Provider load"

296. Response error

Definition: Error arising from the interviewing process.

Context: Response error can result from a number of circumstances, such as the following:

- inadequate concepts or questions;

- inadequate training;

- interviewer failures;

- respondent failures.

Response error may result from the failure of the respondent to report the correct value (respondent error), the failure of the interviewer to record the value reported correctly (interviewer error), or the failure of the instrument to measure the value correctly (instrument error). (United States Federal Committee on Statistical Methodology, "Statistical Policy Working Paper 15: Quality in Establishment Surveys", Washington D.C., July 1988, page 57)

Source: Statistical Office of the United Nations, "Handbook of Household Surveys, Revised Edition", (para. 8.6), Studies in Methods, Series F, No. 31, United Nations, New York, 1984

Hyperlink:

Related Terms:

297. Response rate

Definition: The number of observation units for which data have been received, as a proportion of the number of observation units for which data was sought.

Context: Usually response rate is expressed as a percentage. The response rate can also apply to individual questions.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Item response rate
Non-response rate
Refusal rate

298. Revision policy

See "Data revision"

299. Sample

Definition: A subset of a frame where elements are selected based on a process with a known probability of selection.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Census
Non-response rate
Probability sample
Sample design
Sample size
Sample survey
Sampling
Sampling fraction
Sampling technique
Schedule
Stratification

300. Sample design

Definition: Design that provides information on the target and final sample sizes, strata definitions and the sample selection methodology.

Context: The usage is not uniform as regards the precise meaning of this and similar terms like "sample plan", "survey design", "sampling plan" or "sampling design". These cover one or more parts constituting the entire planning of a sample survey inclusive of processing, etc. The term "sampling plan" may be restricted to mean all steps taken in selecting the sample; the term "sample design" cover in addition the method of estimation; and "survey design" may cover also other aspects of the survey, e.g. choice and training of interviewers, tabulation plans, etc. "Sample design" is sometimes used in a clearly defined sense, with reference to a given frame, as the set of rules or specifications for the drawing of a sample in an unequivocal manner (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003)

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, May 2002

Hyperlink:

Related Terms:
Sample
Survey design

301. Sample size

Definition: The number of observation units which are to be included in the sample.

Context: In the case of a multi-stage sample this number refers to the number of units at the final stage in the sampling.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Sample

Sampling
Sampling fraction

302. Sample survey

Definition: A survey which is carried out using a sampling method.

Context: In sample survey only a portion, and not the whole population is surveyed.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Sample
Survey

303. Sampling

Definition: The process of selecting a number of cases from all the cases in a particular group or universe.

Context: Refers to information on sample design, sample size, sample frame, sample updating, etc.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Area sampling
Sample
Sample size
Sampling error

304. Sampling error

Definition: 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 a subset of the population is enumerated.

Context: Sampling errors are distinct from errors due to imperfect selection, bias in response or estimation, errors of observation and recording, etc.

For probability sampling, the random variation due to sampling can be calculated. For non-probability sampling, random errors cannot be calculated without reference to some kind of model. The totality of sampling errors in all possible samples of the same size generates the sampling distribution of the statistic which is being used to estimate the parent value.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Non-sampling error
Sampling

305. Sampling fraction

Definition: The ratio of the sample size to the population size.

Context:

Source: Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 23

Hyperlink: http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1

Related Terms: Sample
Sample size

306. Sampling frame

See "Frame"

307. Sampling technique

Definition: The name or other identification of the specific process by which the entities of the sample have been selected.

Context:

Source: United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section 3.3.23, page 32

Hyperlink: http://www.census.gov/srd/www/metadata/metada18.pdf

Related Terms: Sample

308. Schedule

Definition: Synonym of questionnaire, in the theory of sample surveys.

Context: A schedule occurs in the specialized sense of a group, or sequence, of questions designed to elicit information upon a subject. Usually, it is completed by an investigator on the basis of information supplied by the particular member of the population chosen for inclusion in the sample, but sometimes it is completed by that member him- or herself, as in postal enquiries.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Questionnaire
Sample
Survey

309. Scope

Definition: The coverage or sphere of what is to be observed. It is the total membership or population of a defined set of people, objects or events.

Context: Scope describes the coverage of the statistics and how consistent this is with internationally accepted standards, guidelines, or recommended practices. The scope includes a description of target population, and geographic, sector, institutional, item, population, product, and other coverage.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Coverage
Out-of-scope units
Statistical population

310. SDMX, Statistical Data and Metadata Exchange

Definition: Set of technical standards and content-oriented guidelines, together with an IT architecture and tools, to be used for the efficient exchange and sharing of statistical data and metadata.

Context: Seven organisations (Bank of International Settlements, European Central Bank, Eurostat, International Monetary Fund, Organisation for Economic Co-operation and Development, United Nations Statistical Division and World Bank) act as sponsors of SDMX.

The first version of the SDMX technical standard (1.0, finalised in 2004) has been approved as an ISO standard (ISO/TS 17369:2005). SDMX V2.0 was also approved by the sponsors in 2005.

Public consultations on the content-oriented guidelines were conducted in 2006 and in 2008, with the aim of compiling guidelines on cross-domain concepts and code-lists, statistical subject-matter domains and the metadata common vocabulary. The first set of content-oriented guidelines was finalised at the end of 2008 to be published in January 2009.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Concept scheme
Cross-domain Concepts
Contact
Data presentation
Dissemnination format
Frequency
Professionalism
Punctuality

Registry
SDMX-EDI
SDMX-ML
SDMX Registry
Simultaneous release
Statistical processing

311. SDMX Registry

See "Registry"

312. SDMX-EDI

Definition: EDIFACT format for exchange of SDMX-structured data and metadata.

Context: The SDMX-EDI format is drawn from the GESMES/TS version 3.0 implementation guide, as published as a standard of the SDMX initiative. SDMX-EDI is a message designed for the exchange of statistical information between organisations in a platform independent manner.

The message implements a data exchange model (SDMX INFORMATION Model) which provides for the exchange of time series identified through a multi-dimensional key and a variety of associated metadata. It employs an appropriate GESMES profile and, for the version described in this Guide, the EDIFACT syntax. Though GESMES is a generic statistical data model which affords sufficient flexibility to describe syntactically virtually any statistical data model, SDMX-EDI has a fixed syntax. This allows partner institutions to design and to build the applications needed to "read" and "write" SDMX-EDI messages, avoiding intermediate files and special translators; the design of the read/write applications is further simplified by eliminating genericity which is not needed when exchanging time series data. Due to the fixed syntax, in most cases, the rules used in SDMX-EDI are stronger and more restrictive than those in generic GESMES. However, the current design allows the possibility of future enhancements and progressive generalisation, if this is needed, upon agreement of the parties involved.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: EDIFACT
SDMX-ML
Statistical Data and Metadata Exchange, SDMX

313. SDMX-ML

Definition: XML format for the exchange of SDMX-structured data and metadata.

Context: The SDMX package contains normative sections describing the use of the XML syntax in SDMX messages, and is accompanied by a set of XML schemas and sample XML document instances.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: SDMX-EDI
Statistical Data and Metadata Exchange, SDMX
XML

314. Seasonal adjustment

Definition: A statistical technique to remove the effects of seasonal calendar influences operating on a series.

Context: Seasonal effects usually reflect the influence of the seasons themselves either directly or through production series related to them, or social conventions. Other types of calendar variation occur as a result of influences such as number of days in the calendar period, the accounting or recording practices adopted or the incidence of moving holidays (such as Easter).

Seasonally adjusted series are data series that have been subjected to a seasonal adjustment process. In other words, the effects of regular, or seasonal, patterns have been removed from the series. Where appropriate, series are also corrected for calendar variations (e.g. workdays per month) and constrained for annual coherency.

A good description of seasonal adjustment methods entails the availability, publication, and level at which seasonal adjustment takes place, the methods used and an indication regarding which data series the methods are applied to (e.g. aggregate series derived from lower-level seasonally-adjusted series versus independently adjusted; adjusted at 1-digit SITC level using X-11 method and aggregated to totals; seasonal adjustment is conducted on four components of final expenditures after annual balancing and then aggregated to total GDP), and on consumer and producer price indexes.

Source: Australian Bureau of Statistics, "An Analytical Framework for Price Indexes in Australia: Glossary and References", Canberra, 1997

Hyperlink: http://www.abs.gov.au

Related Terms: Adjustment
Compilation practices
Statistical processing

315. Secondary supplier

Definition: The organisation or individual other than those responsible for the collection and aggregation of data from the initial source.

Context: Secondary sources may redistribute information received from the primary source either in their initial form or after some transformation including further aggregation, reclassification or other manipulation such as seasonal adjustment.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Primary source

316. Semantics

Definition: The branch of linguistic science which deals with the meaning of words.

Context:

Source: ISO/IEC CD 11179-5 "Information technology - Metadata registries (MDR) - Part 5 : Naming and identification principles", January 2004

Hyperlink:

Related Terms: ISO/IEC 11179
Syntax

317. Sending organisation

Definition: Agency that sends data to one or more receiving agencies during a data exchange.

Context: There can be more than one sending agency for a single data exchange.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

318. Serviceability

Definition: Practical aspects of how well the available data meet users' needs.

Context: 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: International Monetary Fund (IMF), "Data Quality Assessment Framework (DQAF) Glossary"

Hyperlink:

Related Terms: Consistency
Relevance

319. Sex

Definition: The state of being male or female.

Context: This concept is applied if data needs to be categorised by sex. The concept is in general coded, i.e. represented through a code list.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

320. Sibling group

Definition: A set of time series whose keys differ only in the value taken by the frequency dimension.

Context: Within an ETS (exchanged time series) in GESMES/TS, a sibling group is uniquely identified by a data set identifier combined with the sibling group key.

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink: http://www.sdmx.org/Data/GesmesTS_rel3.pdf

Related Terms: Data set
GESMES/TS
Key
Time series

321. Simultaneous release

Definition: The dissemination of statistical data to all interested parties at the same time.

Context: Simultaneous release describes the policy for release of the data to the public, how the public is informed that the data are being released, and whether the policy provides for the dissemination of statistical data to all interested parties at the same time. It also describes the policy for briefing the press in advance of the release of the data.

According to the Fundamental Principles of Official Statistics, the release has to be simultaneous for all users. It is recommended to pre-announce the date of release. Simultaneous release (to all interested parties) is an element of the principle of ready and equal access to official statistics by the public that strengthens transparency in data dissemination practices.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Accessibility
SDMX, Statistical Data and Metadata Exchange

322. Source

Definition: A specific data set, metadata set, database or metadata repository from where data or metadata are available.

Context: The term source is often used as a synonym for "data provider". However, in the context of SDMX, data provider is the organisation or individual from where statistics are obtained.

The term "source" refers to the origins, characteristics and components of the raw data used for compiling statistical aggregates. Sources can be distinguished according to the following classification:

- Statistical source - data collected exclusively for statistical purposes, including data from statistical registers, statistical surveys (including censuses), and other official statistical agencies, as well as data from combinations of these sources.

- Non-statistical source- data not primarily collected for statistical purposes, including data from administrative sources, commercial sources and automatic monitoring and recording devices.

- Mixed source- data from a combination of statistical and non-statistical sources.

An alternative approach is to divide sources into primary sources (the organisation responsible for the original data collection and aggregation) and secondary sources (organisations that re-distribute data received from primary sources either in their initial form or after some transformation).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

323. Source data

Definition: Characteristics and components of the raw statistical data used for compiling statistical aggregates.

Context: The type of source, including whether it is a statistical or non-statistical source, and any relevant characteristics (e.g. sample size for survey data, or characteristics of administrative data) should be mentioned.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

324. Special Data Dissemination Standard, SDDS

Definition: Standard for data dissemination established by the International Monetary Fund.

Context: The Special Data Dissemination Standard (SDDS) was established by the

International Monetary Fund (IMF) to guide members that have, or that might seek, access to international capital markets in the provision of their economic and financial data to the public. Subscription to the SDDS was opened in April 1996.

Source: International Monetary Fund (IMF), "Guide to the Data Dissemination Standards, Module 1: The Special Data Dissemination Standard", Washington, May 1996

Hyperlink: http://dsbb.imf.org/Applications/web/getpage/?pagename=legaltext

Related Terms:
Accounting conventions
Adjustment
Analytical framework
Base period
Coverage
Data
Data reconcilation
Data status
Dissemination
Dissemination standard
Estimation
General Data Dissemination System, GDDS
Integrity
ISO/IEC 11179
Language
Metadata dimension
Microdata
Ministerial commentary
Pre-release access
Periodicity
Terminology
Standard classification
Statistical processing
Quality

325. Standard classification

Definition: Classifications that follow prescribed rules and are generally recommended and accepted.

Context: Standard classifications aim to ensure that information is classified consistently regardless of the collection, source, point of time etc.

In the international context, standard classifications include ISIC, ISCO, CPC, NACE, etc. Many national statistical systems also have their own versions of standard classifications, which in the main are consistent with international standard classifications, though modified to meet national circumstances.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms:
Classification
Special Data Dissemination Standard, SDDS

326. Standard error

Definition: The positive square root of the variance of the sampling distribution of a statistic.

Context: It includes the precision with which the statistics estimates the relevant parameter as contrasted with the standard deviation that describes the variability of primary observations.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms:

327. Statistical activity

Definition: The collection, storage, transformation and distribution of statistical information.

Context: Non-statistical activities carried out by the staff of a statistical agency in support of the statistical business process, e.g. finance or human resource management, can be referred to as support activities.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

328. Statistical business process

Definition: The complete set of sub-processes needed to support statistical production.

Context: The statistical business process is described in Eurostat's Cycle de Vie des Données (CVD) model, and the Generic Statistical Business Process Model, being developed by the METIS group on statistical metadata.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

329. Statistical concept

Definition: A statistical characteristic of data.

Context: In SDMX, "Statistical concepts and definitions" refer to the internationally accepted statistical standards, guidelines, or recommended practices on which the concepts and definitions that are used for compiling the statistics

are based. It also refers to the description of deviations of the concepts and definitions from accepted statistical standards, guidelines, or good practices, when relevant. This should define the statistical concept under measure and the organisation of data, i.e. the type of variables included in the domain of study.

A statistical concept is either coded or uncoded. A coded statistical concept takes values from a code list of valid values. For example, a coded statistical concept called "reporting country" might be created, taking its values from the ISO list of country codes. A code list may supply the values of more than one statistical concept. An uncoded statistical concept takes its values as free form text (e.g. time series title).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Attribute
Characteristic
Code list
Concept
Concept scheme
Data structure definition
Dimension
GESMES/TS
International statistical standard
Key structure
Observation
Structural definition

330. Statistical error

Definition: The unknown difference between the retained value and the true value.

Context: It is immediately associated with accuracy since accuracy is used to mean "the inverse of the total error, including bias and variance" (Kish L., "Survey Sampling", John Wiley, New York 1965). The larger the error, the lower the accuracy.

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms: Accuracy

331. Statistical indicator

Definition: A data element that represents statistical data for a specified time, place, and other characteristics, and is corrected for at least one dimension (usually size) to allow for meaningful comparisons.

Context: A simple aggregation such as the number of accidents, total income or women Members of Parliament, is not in itself an indicator, as it is not comparable between populations. However, if these values are standardized, e.g. number of accidents per thousand of population, average

income, or women Members of Parliament as a percentage of the total, the result meets the criteria for an indicator.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

332. Statistical information

Definition: Aggregated or unit-level information obtained through statistical activities.

Context: Statistical information includes semantic descriptions (metadata) needed for interpretation of these data.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

333. Statistical macrodata

Definition: An observation data gained by a purposeful aggregation of statistical microdata conforming to statistical methodology.

Context: Macrodata is data derived from microdata by statistics on groups or aggregates, such as counts, means, or frequencies. (United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section 3.4.4, page 39).

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Statistical microdata

334. Statistical measure

Definition: A summary (means, mode, total, index, etc.) of the individual quantitative variable values for the statistical units in a specific group (study domains).

Context:

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms:

335. Statistical message

Definition: A message carrying statistical data.

Context: A statistical message is a predefined and agreed way of representing syntactically sets of statistical data, attributes and structural definitions which need to be exchanged between partners.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Data exchange
GESMES

336. Statistical metadata

Definition: Data about statistical data.

Context: Statistical metadata comprise data and other documentation that describe objects in a formalised way (Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000, http://www.unece.org/stats/publications/53metadaterminology.pdf)

They provide information on data and about processes of producing and using data. Statistical metadata describe statistical data and - to some extent - processes and tools involved in the production and usage of statistical data (UNECE, "Guidelines for the Modelling of Statistical Data and Metadata", 1995).

Statistical metadata can be classified in various ways, but there is a clear high-level distinction between the metadata needed to search for and display data (Structural metadata) and the metadata that give more information on definitions, methodologies, processes and quality (Reference metadata).

Source: United Nations Statistical Commission and Economic Commission for Europe of the United Nations (UNECE), "Guidelines for the Modeling of Statistical Data and Metadata", Conference of European Statisticians, Methodological material, United Nations, Geneva,

Hyperlink: http://www.unece.org/stats/publications/metadatamodeling.pdf

Related Terms: Metadata
Metadata layer
Metadata registry
Reference metadata
Statistical metadata system
Statistical metainformation
Structural metadata

337. Statistical metadata system

Definition: A data processing system that uses, stores and produces statistical metadata.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Metadata
Metadata layer
Statistical metadata
Statistical metainformation

338. Statistical metainformation

Definition: Knowledge of objects described by statistical metadata.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Metadata layer
Statistical metadata
Statistical metadata system
Statistical metainformation system

339. Statistical metainformation system

Definition: System which uses and produces statistical metadata and which fulfils its tasks by means of functions like "statistical metadata collection", "statistical metadata processing", "statistical metadata storage" and "statistical metadata dissemination".

Context: A metainformation system may be active or passive. An active metainformation system is physically integrated with the information system containing the data that the metadata in the metainformation system informs about. A passive metainformation system contains only references to data, not the data themselves (United Nations Economic Commission for Europe/United Nation Statistical Commision (UNECE/UNSC), "Guidelines for the Modelling of Statistical Data and Metadata", Conference of European Statisticians Methodological Material, Geneva, 1995, p. 4).

Source: United Nations Statistical Commission and Economic Commission for Europe of the United Nations (UNECE), "Guidelines for the Modeling of Statistical Data and Metadata", Conference of European Statisticians, Methodological material, United Nations, Geneva,

Hyperlink:

Related Terms: Statistical metainformation

340. Statistical methodology

Definition: Theory and methods of data collection, processing and analysis.

Context:

Source: Statistics Canada, Integrated Metadata Base, "Glossary", unpublished on paper

Hyperlink: http://www.statcan.ca/english/edu/power/toc/contents.htm

Related Terms: Methodology

341. Statistical microdata

Definition: An observation data collected on an individual object.

Context: Microdata is data on the characteristics of units of a population, such as individuals, households, or establishments, collected by a census, survey, or experiment. (United States Bureau of the Census, Software and Standards Management Branch, Systems Support Division, "Survey Design and Statistical Methodology Metadata", Washington D.C., August 1998, Section 3.4.4, page 39, at http://www.census.gov/srd/www/metadata/metada18.pdf).

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Statistical macrodata

342. Statistical organisation

Definition: A producer of official statistics.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

343. Statistical population

Definition: The total membership or population or "universe" of a defined class of people, objects or events.

Context: There are two types of population, viz., target population and survey population. A "target population" is the population outlined in the survey objects about which information is to be sought and a "survey population" is the population from which information is obtained in a survey. The target population is also known as the scope of the survey and the survey population as the coverage of the survey. For administrative data sources, the corresponding populations are the "target population", as defined by the relevant legislation and regulations, and the actual "client population" ("United Nations Glossary of Classification Terms" prepared by the Expert Group on International Economic and Social Classifications).

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms:
Census
Coverage error
Cut-off threshold
Scope
Stratification
True value

344. Statistical processing

See Data Compilation

Context:

Source:

Hyperlink:

Related Terms:
Aggregation
Compilation practices
Consolidation
Data processing
Data reconciliation
Disaggregation
Estimation
Index number
Reference period
Revision policy
Seasonal adjustment
Special Data Dissemination Standard, SDDS
Statistical Data and Metadata Exchange, SDMX

345. Statistical production

Definition: The activity that is carried out within statistical information system and aimed at producing statistics.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:

346. Statistical purpose

Definition: The use of data in a way that complies with the Fundamental Principles of Official Statistics, fits into one or more phases of the statistical business process, and contributes to the production of official statistics.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

347. Statistical standard

Definition: An agreed rule or guideline on how one or more parts of the statistical business process should be carried out, conforming with requirements for professionalism.

Context: Components of a standard include:
- definition(s)
- statistical units
- classification(s)
- coding process(es)
- questionnaire module(s)
- output categories

The use of statistical standards permits the repeated collection of statistics on a consistent basis. They also enable the integration of data over time and across different data sources, allowing the use of data beyond the immediate purpose for which it was produced. Standards also reduce the resource requirements associated with many aspects of survey development and maintenance (Statistics New Zealand," Classifications and Standards"; unpublished on paper)

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: International statistical standard

348. Statistical subject-matter domain

Definition: A statistical activity that has common characteristics with respect to concepts and methodologies for data collection, manipulation and transformation.

Context: Within SDMX, the list of Statistical Subject-Matter Domains (aligned to the UN/CES Classification of International Statistical Activities) is a standard reference list against which the categorisation schemes of various participants in exchange arrangements can be mapped to facilitate data and metadata exchange. This allows the identification of subject matter domain groups involved in the development of guidelines and recommendations relevant to one or more statistical domains. Each of these groups could define domain-specific data structure definitions, concepts, etc.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Characteristic
Concept
Domain groups
Methodology
Study domain

349. Statistical unit

Definition: Entity for which information is sought and for which statistics are ultimately compiled.

Context: The statistical unit is the object of a statistical survey and the bearer of statistical characteristics. These units can, in turn, be divided into observation units and analytical units.

Statistical units for economic statistics are defined on the basis of three criteria: 1) Legal, accounting or organisational criteria; 2) Geographical criteria; 3) Activity criteria.

Statistical units comprise the enterprise, enterprise group, kind-of-activity unit (KAU), local unit, establishment, homogeneous unit of production, persons, households, geographical areas, events etc.

Statistical units can be categorised into basic statistical units, i.e. those for which data is collected, and derived statistical units, i.e. those which are constructed during the statistical production process. A basic statistical unit is the most detailed level to which the obtained characteristics can be attached.

In other statistical domains, statistical units can include persons, households, geographical areas, events etc. Statistical units can be categorised into basic statistical units, i.e. those for which data is collected, and derived statistical units, i.e. those which are constructed during the statistical production process.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Analytical unit
Classification
Comparability
Observation unit

350. Statistical variable

Definition: A characteristic of a unit being observed that may assume more than one of a set of values.

Context: A variable in the mathematical sense, i.e. a quantity which may take any one of specified set of values. It is convenient to apply the same word to denote non-measurable characteristics, e.g., 'sex' is a variable in this sense since any human individual may take one of two 'values', male or female. It is useful, but far from being the general practice, to distinguish between a variable as so defined and a random variable (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003).

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink: http://unstats.un.org/unsd/class/family/glossary_short.htm

Related Terms: Characteristic
Observation

351. Stratification

Definition: A sampling procedure in which the population is divided into homogeneous subgroups or strata and the selection of samples is done independently in each stratum.

Context: The division of a population into parts is known as strata, especially for the purpose of drawing a sample, an assigned proportion of the sample then being selected from each stratum. The process of stratification may be undertaken on a geographical basis, e.g. by dividing up the sampled area into sub-areas on a map; or by reference to some other quality of the population, e.g. by dividing the persons in a town into strata according to sex or into three strata according to whether they belong to upper, middle or lower income groups.

The term stratum is sometimes used to denote any division of the population for which a separate estimate is desired, i.e. in the sense of a domain of study. It is also used sometimes to denote any division of the population for which neither separate estimates nor actual separate sample selection is made. (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003).

Source: Statistics Canada, Canada's National Statistical Agency, online glossary of statistical terms and definitions

Hyperlink: http://www.statcan.ca/english/freepub/81-004-XIE/def/stratdef.htm

Related Terms: Sample
Statistical population

352. Structural definition

Definition: Statistical concepts, key families and code lists defined by a centre institution (usually for the exchange of statistical information with its partners).

Context:

Source: European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper

Hyperlink: http://www.ecb.int/stats/services/gesmes/html/index.en.html

Related Terms:
Code list
Concept
Concept scheme
Data structure definition
GESMES/TS
Maintenance agency
Statistical concept
Structural metadata

353. Structural metadata

Definition: Metadata that act as identifiers and descriptors of the data.

Context: 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 (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:
Cross-domain Concepts
Data structure definition
GESMES/TS
Metadata structure definition
Reference metadata
SDMX Registry
Statistical metadata
Structural definition

354. Structure

Definition: A hierarchical classification for identifying relationships between categories.

Context: A hierarchical classification is based on a tree structure where each set of its detailed categories are subsets of categories at the level above the one

in which they are contained.

Source: United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper

Hyperlink:

Related Terms: Category
Data structure definition
Hierarchy
Metadata structure definition

355. Study domain

Definition: A segment of the population for which separate statistics are needed.

Context: A study domain could consist of a geographical area such as a region or major population centre. It could also comprise a specified population category, such as a major national or ethnic group. The number of domains has an important bearing on the size and distribution of the sample.

Normally statistics are presented for different sub-groups of the population, so called study domains. These study domains can be geographical as well as non-geographical. Often these sub-groups are according to some classification (e.g. territorial units, economic activity etc.) (Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003).

In the course of tabulation, data may actually be provided for many population segments; however, a study domain would be a segment identified in the overall statistical plan as one for which a certain level of detail and certain data reliability were required. The study domains chosen may coincide with the strata adopted for stratified sampling or may cut across them.

Source: Statistical Office of the United Nations, "Handbook of Household Surveys, Revised Edition", (paras. 4.6, 4.7), Studies in Methods, Series F, No. 31, United Nations, New York, 1984

Hyperlink:

Related Terms: Statistical subject-matter domain

356. Submitting organisation

Definition: The organization responsible for requesting that a new metadata item be registered in the registry.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: ISO/IEC 11179
Metadata registry

357. Survey

Definition: An investigation about the characteristics of a given population by means of collecting data from a sample of that population and estimating their characteristics through the systematic use of statistical methodology.

Context: Included are:

- censuses, which attempt to collect data from all members of a population;

- sample surveys, in which data are collected from a (usually random) sample of population members.

Surveys can be unique in time or repeated with regular or irregular periodicity. A single wave of a repeated survey is called survey instance.

A wider definition under which the term survey covers any activity that collects or acquires statistical data (including censuses, sample surveys, the collection of data from administrative records and derived statistical activities) has also been proposed. (see Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 7, available at http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1)

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:
Census
Cut-off survey
Non-response
Processing error
Questionnaire
Reporting unit
Sample survey
Schedule
Survey design

358. Survey data collection

Definition: An activity of the survey life cycle for gathering data from respondents and recording it for further processing.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms:

359. Survey design

Definition: All the aspects of a survey from the establishment of a need for data to the production of final outputs.

Context: The survey design addresses the following issues: what statistics are produced, for which population, when, and with what accuracy; what data are to be collected for which units of the population of interest, and what are the methods by which those data are to be collected and processed to produce the required statistics. Operational, organisational and administrative issues are usually addressed (Lessler, J.T. and Kalsbeek, W.D., "Non Sampling Error in Survey", John Wiley, New York, 1992 or US Department of Commerce, "Glossary of Non Sampling Error Terms: An Illustration of a Semantic Problem in Statistics", Statistical Policy Working Paper 4, Office of Federal Statistical Policy Standards, 1978).

Source: Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 8

Hyperlink: http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1

Related Terms: Questionnaire design
Sample design
Survey

360. Syntax

Definition: The relationships among characters or groups of characters, independent of their meanings or the manner of their interpretation and use; the structure of expressions in a language, and the rules governing the structure of a language.

Context:

Source: ISO/IEC CD 11179-5 "Information technology - Metadata registries (MDR) - Part 5 : Naming and identification principles", January 2004

Hyperlink:

Related Terms: ISO/IEC 11179
Semantics

361. Target population

Definition: The set of elements about which information is wanted and estimates are required. Practical considerations may dictate that some units are excluded (e.g., institutionalized individuals, the homeless, or those that are not be possible to access without incurring excessive cost).

Context:

Source: Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 17

Hyperlink: http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1

Related Terms: Cut-off threshold
Under-coverage

362. Taxonomy

Definition: Classification of things according to a presumed relationship among types and subtypes.

Context: The term "taxonomy" comes from the Greek τάξις, taxis (meaning 'order', 'arrangement') and νόμος, nomos ('law' or 'science'). Taxonomic schemes are composed of taxonomic units that are arranged frequently in a hierarchical structure. Typically they are related by subtype-supertype relationships also called parent-child relationships.

Within SDMX, a reporting taxonomy is a scheme which defines the composition structure of a data or metadata report where each component is described by an independent data flow definition.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Classification
ISO/IEC 11179
Ontology

363. Term

Definition: A designation of a defined concept in a special language by a linguistic expression.

Context: A term is a word or phrase used to designate a concept (Terminology on Statistical Metadata, Conference of European Statisticians Statistical Standards and Studies, No. 53, UNECE, Geneva 2000, http://www.unece.org/stats/publications/53metadaterminology.pdf).

Source: ISO International Standard 1087-1:2000, Terminology work -- Vocabulary -- Part 1: Theory and application, November 2004

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Terminology
Thesaurus
Terminological entry
Terminological system

364. Terminological entry

Definition: An entry containing information on a specific terminological unit within a context or subject field.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Term
Terminology
Thesaurus
Terminological system

365. Terminological system

Definition: A concept system with designations for each concept.

Context:

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Concept
Term
Terminology
Terminological entry

366. Terminology

Definition: A set of terms.

Context:

Source: Economic Commission for Europe of the United Nations (UNECE), "Terminology on Statistical Metadata", Conference of European Statisticians Statistical Standards and Studies, No. 53, Geneva, 2000

Hyperlink: http://www.unece.org/stats/publications/53metadaterminology.pdf

Related Terms: Special Data Dissemination Standard, SDDS
Term
Terminological system
Terminology

367. Thesaurus

Definition: Structured list of expressions intended to represent in unambiguous way the conceptual content of a documentary system and the queries addressed to that system.

Context: A statistical thesaurus may assist in locating an existing data element.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Data element
ISO/IEC 11179
Term

368. Time coverage

Definition: The length of time, e.g. years, for which data are collected.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

369. Time format

Definition: Technical format in which time is represented for the measured phenomenon.

Context: The technical time format and its related code list are part of the technical standards for SDMX-EDI and SDMX-XML.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:

370. Time of recording

Definition: The date the item was recorded in a dissemination medium.

Context: The time of recording may be the date the item was first recorded or the date an existing item was amended. In National Accounts, time of recording pertains to the issues involved in deciding whether to record a transaction with regard to when the claim arises (accrual) or when it is to be paid (cash).

See: United Nations, "System of National Accounts (SNA) 1993" and International Monetary Found, "Balance of Payments Manual (BPM)", Washington D.C., 1993.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Accounting conventions
Date of last change

371. Time period

Definition: The period of time or point in time to which the measured observation refers.

Context: The measurement represented by each observation corresponds to a specific point in time (e.g. a single day) or a period (e.g. a month, a fiscal year, or a calendar year). This is used as a time stamp and is of particular importance for time series data. In cases where the actual time period of the data differs from the target reference period, “time period” refers to the actual period.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms:
Observation
Period
Reference period
Time series

372. Time series

Definition: A set of ordered observations on a quantitative characteristic of an individual or collective phenomenon taken at different points of time.

Context: Although it is not essential, it is common for these points to be equidistant in time. The essential quality of the series is the order of the observations according to the time variable, as distinct from those which are not ordered at all, e.g. in a random sample chosen simultaneously or are ordered to their internal properties, e.g. a set arranged in order of magnitude.

In GESMES/TS, a time series is a time-ordered vector of observations. A time series is uniquely defined with in a data set by its key. (European Central Bank (ECB), Bank for International Settlement (BIS), Eurostat, International Monetary Fund (IMF), Organisation for Economic Co-operation and Development (OECD), "GESMES/TS User Guide", Release 3.00, February, 2003; unpublished on paper available at http://www.sdmx.org/Data/GesmesTS_rel3.pdf)

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink: http://www.sdmx.org/Data/GesmesTS_rel3.pdf

Related Terms:
Characteristic
Data provider series key
Data set
Dimension
Key
Observation
Observation pre-break value
Sibling group
Time period
Time series break
Trend

373. Time series break

Definition: Break occurring when there is a change in the standards for defining and observing a variable over time.

Context: A break may be the result of a single change or the combination of multiple changes at any one point in time of observation of the variable. The specific causes of breaks in a statistical time series include changes in: classifications used, definitions of the variable, coverage; etc.

Statistical agencies and users of time series data for economic research to analyse and interpret economic and social events and conditions attach very high importance to the continuity and consistency of data over time. However, it should be emphasised that the occurrence of time series break may not necessarily jeopardise the reliability of a time series. Statistical agencies frequently apply a number of techniques to ensure the continuity of a time series. Finally, the impact of a time series break is often a matter of judgement on the part of the user and depends on the use(s) to which the data are put.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Observation pre-break value
Revision policy
Time series

374. Timeliness

Definition: Length of time between data availability and the event or phenomenon they describe.

Context: Timeliness refers to the speed of data availability, whether for dissemination or for further processing, and it is measured with respect to the time lag between the end of the reference period and the release of data. Timeliness is a crucial element of data quality: adequate timeliness corresponds to a situation where policy-makers can take informed decisions in time for achieving the targeted results. In quality assessment, timeliness is often associated with punctuality, which refers to the time lag between the release date of data and the target date announced in some official release calendar.

Timeliness can be further broken down into "Timeliness - output" and "Timeliness - source data".

"Timeliness - output" refers to the the lapse of time between the end of a reference period (or a reference date) and the release of a version of the data: provisional, preliminary, or final results. This reflects many factors, including some that are related to institutional arrangements, such as the preparation of accompanying commentary and printing. Usually, data are not released immediately at the end of the period they refer to, since data collection, data processing and data dissemination work needs to be performed.

"Timeliness - source data" refers to the time between the end of a reference period (or a reference date) and actual receipt of the data by the compiling agency. Compared to the parent concept - timeliness - this concept only

covers the time period between the end of the reference period and the receipt of the data by the data compiling agency. This time period is determined by factors such as delays accomodating the institutional arrangements for data transmission.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org

Related Terms: Data
Punctuality
Quality
Release policy

375. Title

Definition: Textual label used as identification of a statistical object.

Context: The title is a short name describing and identifying a statistical object it is attached to.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

376. Transparency

See "Integrity"

377. Trend

Definition: A long-term movement in an ordered series, say a time series, which may be regarded, together with the oscillation and random component, as generating the observed values.

Context: In time series analysis, a given time series can be decomposed into: a) a cyclical component; b) a trend component; c) a seasonal component; d) an irregular component.

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Time series
Trend estimates

378. Trend estimates

Definition: Estimates derived from seasonally adjusted estimates via an averaging process which attempts to remove the irregular component of the time series. This allows the underlying direction of a time series to be identified.

Context:

Source: Australian Bureau of Statistics, "An Analytical Framework for Price Indexes in Australia: Glossary and References", Canberra, 1997

Hyperlink: http://www.abs.gov.au

Related Terms: Estimate
Trend

379. True value

Definition: The actual population value that would be obtained with perfect measuring instruments and without committing any error of any type, both in collecting the primary data and in carrying out mathematical operations.

Context:

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms: Statistical population

380. Type of data collection

Definition: Main process used in the collection of statistical data by the primary source of the data.

Context: The type of data collection refers to the main process used in the collection of statistical data by the primary source of the data, those commonly used being survey data collection and administrative data collection. Each of these broad types may be further broken down on the basis of some characteristic, e.g. the nature of the data provider (enterprise / household) or exhaustiveness (sample survey, complete enumeration census).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

381. Under-coverage

Definition: Failure to include required units in the frame, which results in .the absence of information for those units.

Context: Under-coverage should not be confused with non-response.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Frame
Over-coverage
Target population

382. Unit multiplier

Definition: Exponent in base 10 specified so that multiplying the observation numeric values by 10^UNIT_MULT gives a value expressed in the unit of measure.

Context: In some data bases, it is referred to as SCALE, MAGNITUDE or POWER, e.g. "UM=6" means that observations are in millions.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

383. Unit non-response

See "Non-response error"

384. Unit of measure

Definition: The unit in which the data values are measured.

Context: The unit of measure is a quantity or increment by which something is counted or described, such as kg, mm, °C, °F, monetary units such as Euro or US dollar, simple number counts or index numbers. The unit of measure has a type (e.g. currency) and, in connection with the unit multiplier, provides the level of detail for the value of the variable (e.g. Euro, 1000 Euro).

For data messages, the concept is always represented by codes. Any additional detail needed must be inserted as free text within "unit of measure detail".

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Conceptual domain
Dimension
Dimensionality
ISO/IEC 11179

385. Unit response rate

Definition: The percentage of an eligible sample for whom information is obtained.

Context: For a survey, the numerator of the formula is the number of responses. The denominator is the total sample size minus non-eligible respondents, i.e. minus those not meeting the criteria for a potential respondent as defined for that particular study.

The weighted response rate calculates the ratio using the inverse probability of inclusion in the sample as a weight for each unit. In some occasions a value that reflects the importance of the unit is also used as a weighting factor (like size of workforce for establishment).

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

386. Unit value

Definition: Expenditures or value of production of an item divided by the quantity.

Context:

Source: United Nations Department of Economic and Social Development - Statistical Division, Handbook of the International Comparison Programme, Studies in Methods, Series F, No. 62, New York, 1992, Glossary

Hyperlink:

Related Terms:

387. User

Definition: Recipient of statistical information, who transforms it into knowledge needed for decision making or research.

Context: User needs refer to data and metadata requirements of persons or organisations to meet a particular use or set of uses. User needs may be specified in terms of the quality dimensions promulgated by international organisations or national agencies.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms:

388. User satisfaction survey

Definition: A statistical survey aiming to assess the satisfaction of users of statistics.

Context:

Source: Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003

Hyperlink:

Related Terms:

389. Validation

See "Data Validation"

390. Valuation

Definition: The definition of the price per unit, for goods and services flows and asset stocks.

Context: Standard national accounts valuations include the basic price (what the seller receives) and the purchaser's price (what the purchaser pays). The purchaser's price is the basic price, plus taxes less subsidies on products, plus invoiced transportation and insurance services, plus distribution margin. Other valuation bases may be used in other contexts. International trade in goods considers the free on board (fob) price and cost-insurance-freight price, among others.

The concept refers to valuation rules used for recording flows and stocks, including how consistent the practices used are with internationally accepted standards, guidelines, or good practices.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Accounting conventions

391. Value domain

Definition: A set of permissible values.

Context: In the context of ISO 11179, a domain is the set of possible data values of an attribute. A "data value" is an element of a value domain.

The value domain provides representation, but has no implication as to what data element concept the values may be associated with, or what the values mean. The permissible values may either be enumerated or expressed via a description.

Enumerated value domain is a value domain that is specified by a list of all its permissible values.

Non-enumerated value domain is a value domain that is specified by a description rather than a list of all Permissible values.

Non-enumerated value domain description is a description or specification of a rule, reference, or range for a set of all permissible values for the value domain.

Value domain representation class is the class of representation of a value domain. Attributes of value domain:

Value domain administration record is the administration record for a value domain.

Value domain datatype is the datatype used in a value domain.

Value domain format is a template for the structure of the presentation of the Value(s) e.g. - YYYY-MM-DD for a date.

Value domain maximum character quantity is the maximum number of characters to represent the Data Element value and is applicable only to character datatypes.

Value domain relationship is a relationship among two or more Value domains.

Value domain unit of measure is the unit of measure used in a value domain. (ISO/IEC 11179-3 "Information technology - Metadata registries- Part 3: Registry metamodel and basic attributes", February 2003)

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Attribute
ISO/IEC 11179
Permissible value
Permitted value
Value item
Value meaning

392. Value item

See "Permitted value"

393. Value meaning

Definition: The meaning or semantic content of a value.

Context: Given a permissible value, representation of its value meaning shall be independent of (and shall not constrain) the representation of its corresponding value.

The representation of value meanings in a registry shall be independent of (and shall not constrain) their representation in any corresponding value domain.

Value meaning set is the relationship between a conceptual domain and a set of value meanings.

Attributes of value meaning:

Value meaning begin date is the effective date of this value meaning in the conceptual domain. A registration authority may determine whether this date is the date the value meaning becomes valid in a registry or the date the value meaning becomes part of the source domain or some other date.

Value meaning description is a description of a value meaning.

Value meaning end date is the date this value meaning became/becomes invalid. A registration authority may determine whether this date is the date the value meaning becomes no longer valid in a registry or the date the value meaning becomes no longer part of the source domain or some other date.

Value meaning identifier is the unique identifier for a value meaning.

Source: ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004

Hyperlink:

Related Terms: Conceptual domain
ISO/IEC 11179
Permissible value
Value domain
Value item

394. Verification

Definition: Principal methods to review, audit, or verify the accuracy of the disseminated data.

Context: Verification methods (e.g., internal review, statistical confidence tests, internal audit, audit by outside accountants, cross-checks with other macroeconomic accounts, etc.) may entail the reconciliation of stocks and transactions data, reconciliation of reported data with money and banking statistics, custodian data, differences with partner data or preshipment inspection data, the treatment of differences between GDP compiled for the production approach and GDP compiled from the expenditure approach.

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Compilation practices

395. Weight

Definition: The importance of an object in relation to a set of objects to which it belongs.

Context: A weight is represented as a numerical coefficient attached to an observation, frequently by multiplication, in order that it shall assume a desired degree of importance in a function of all the observations of the set.

Reweighting consists of raising the original weights for the respondent values when estimates are computed. Reweighting concerns mainly unit non-response. It may also be used to increase precision through the use of auxiliary information. Standard methods include post-stratification, calibration and response propensity modelling (Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003).

Source: The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003

Hyperlink:

Related Terms: Base weight
Non-response
Non-response error
Ratio estimation
Weight period

396. Weight period

Definition: The period that provides the weights for an index number.

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: Base period
Index type
Period
Weight

397. XML

Definition: eXtensible Mark-up Language

Context:

Source: SDMX (2009)

Hyperlink: http://www.sdmx.org/

Related Terms: SDMX-ML