SDMX CONTENT-ORIENTED GUIDELINES:

METADATA COMMON VOCABULARY

(DRAFT MARCH 2006)

CONTENTS

1. SCOPE OF THE CONTENT-ORIENTED GUIDELINES 12
2. METADATA COMMON VOCABULARY 12
3. GLOSSARY 15

Accessibility 15
Accounting basis 15
Accounting conventions 15
Accuracy 16
Adjustment 16
Adjustment Methods 17
Administered item 17
Administration record 17
Administrative data 17
Administrative data collection 18
Administrative source 18
Agency 18
Aggregation 19
Aggregation Equation 19
Analytical framework 19
Analytical unit 19
Area sampling 20
Attachment level 20
Attribute 21
Availability 21
Base period 21
Base weight 22
Base year 22
Basic attribute 22
Benchmark 22
Benchmarking 23
Bias 23
Bilateral exchange 23
Break 24
Category 24
Category Scheme 24
Census 24
Chain index 25
Characteristic 25
Clarity 26
Class 26
Classification 26
Classification changes 27
Classification scheme 28
Classification unit 28
Co-ordination of samples 28
Code 29
Code list 29

Coding 30
Coding error 30
Coherence 30
Collection 30
Comparability 31
Compilation 31
Compilation practices 31
Compiling Agency 31
Completeness 32
Computation of lowest level indices 32
Computer Assisted Interviewing, CAI 32
Concept 32
Concept Scheme 33
Conceptual data model 33
Conceptual domain 34
Confidential data 34
Confidentiality 35
Consistency 35
Consolidation 35
Constraint 36
Contact 36
Context 37
Country identifier 37
Coverage 38
Coverage errors 38
Coverage ratio 38
Creation date 39
Cross-domain Concepts 39
Cut-off survey 39
Cut-off threshold 40
Data 40
Data analysis 41
Data attribute 41
Data capture 41
Data checking 41
Data collection 41
Data confrontation 42
Data consumer 42
Data dissemination 42
Data dissemination standards 43
Data editing 43
Data element 44
Data element concept 44
Data element derivation 45
Data exchange 45
Data exchange context 46
Data flow definition 46
Data identifier 46
Data interchange 47
Data item 47

196 Data model 47
197 Data presentation 48
198 Data processing 48
199 Data provider 48
200 Data provider series key 49
201 Data reconciliation 49
202 Data security 49
203 Data set 50
204 Data set identifier 50
205 Data sharing exchange 50
206 Data source 51
207 Data status upon release 51
208 Data structure definition 51
209 Datatype 52
210 Date 52
211 Date of last change 53
212 Definition 53
213 Derivation input 54
214 Derivation output 54
215 Derivation rule 54
216 Derived data element 54
217 Derived statistic 55
218 Dimension 55
219 Dimensionality 56
220 Disaggregation 56
221 Disclosure analysis 57
222 Dissemination format 57
223 Documentation 57
224 Domain groups 57
225 Dublin Core 58
226 EDIFACT 58
227 Electronic data interchange (EDI) 58
228 Entity 59
229 Error of estimation 59
230 Error of observation 59
231 Estimate 60
232 Estimation 60
233 Estimator 60
234 Expected value 61
235 Expression Node 61
236 Flag 61
237 Flow data series 62
238 Follow-up 62
239 Footnote 62
240 Frame 62
241 Frame error 63
242 Frequency 63
243 Gateway 64
244 Gateway exchange 64
245 General Data Dissemination System (GDDS) 64

Geographical coverage 64
GESMES 65
GESMES/CB 65
GESMES/TS 65
GESMES/TS data model 66
Glossary 66
Graphical data editing 67
Grossing/Netting 67
Guidelines 67
Hierarchy 68
Identifier 68
Imputation 68
Index number 69
Information 69
Information system 69
Inlier 69
Institutional framework 70
Institutional sector 70
Institutional unit 71
Integrity 71
Internal access 72
International code designator 72
International statistical standard 72
Interpolation 73
Interviewer error 73
ISO/IEC 11179 73
Item response rate 75
Key (time series or sibling group) 75
Key family 76
Key structure 76
Keyword 76
Language 76
Level 77
Levels of data 77
Longitudinal data 77
Macro editing 77
Maintenance Agency 78
Measure 78
Measurement error 78
Metadata 79
Metadata Attribute 79
Metadata dimension 79
Metadataflow definition 80
Metadata item 80
Metadata layer 80
Metadata object 81
Metadata registry 81
Metadata set 82
Metadata Structure Definition 82
Metamodel 82

Methodological soundness 83
Methodology 83
Micro editing 83
Ministerial commentary 83
Misclassification 84
Missing data 84
Model assumption error 84
Multilateral exchange 84
Name 85
Nature of the basic data 85
Nomenclature 85
Non-probability sample 86
Non-response 86
Non-response bias 86
Non-response error 86
Non-response rate 87
Non-sampling error 87
Not seasonally adjusted series 88
Number raised estimation 88
Object 88
Object class 89
Objectives 89
Observation 89
Observation confidentiality 90
Observation unit 90
Observation value 90
Ontology 90
Organisation 91
Organisation identifier 91
Organisation Role 92
Origin 92
Out-of-scope units 92
Outliers 92
Over-coverage 93
Period 93
Periodicity 93
Permissible value 94
Permitted value 94
Pre-break observation 95
Pre-Break Value 95
Precision 95
Preferred definition 95
Prerequisites of quality 96
Primary data 96
Primary source of statistical data 96
Probability sample 96
Processing error 97
Product 97
Professionalism 97
Property 98

Provider load 98
Provision Agreement 98
Public disclosure 99
Punctuality 99
Qualitative data 99
Quality 100
Quality (Eurostat context) 100
Quality (IMF context) 101
Quality (OECD context) 102
Quality control survey 102
Quality differences 102
Quality index 103
Quantitative data 103
Questionnaire 103
Questionnaire design 103
Ratio estimation 104
Recommended use of data 104
Record check 105
Record-keeping error 105
Recording of transactions 105
Reference document 105
Reference metadata 106
Reference period 106
Reference time 106
Refusal rate 107
Register 107
Registrar 107
Registration 108
Registration authority 108
Registry item 108
Registry metamodel 108
Related data reference 109
Related metadata reference 109
Relationship 109
Relative Standard error 110
Release calendar 110
Relevance 110
Reliability 111
Reporting unit 111
Respondent burden 111
Respondent load 111
Response errors 112
Response rate 112
Responsible organization 112
Revision policy 112
Sample 113
Sample design 114
Sample size 114
Sample survey 114
Sampling 115

Sampling error 115
Sampling fraction 115
Sampling frame 115
Sampling technique 115
Sampling unit 116
Schedule 116
Scope 116
SDMX-EDI 117
SDMX-ML 117
SDMX Registry 118
Seasonal adjustment 118
Secondary source of statistical data 119
Semantics 119
Serviceability 120
Sibling group 120
Simultaneous release 120
Source 121
Source data 121
Special Data Dissemination Standard (SDDS) 121
Special language 122
Standard Classification 122
Standard error 123
Statistical concept 123
Statistical Data and Metadata Exchange (SDMX) 124
Statistical error 124
Statistical indicator 124
Statistical macrodata 124
Statistical measure 125
Statistical message 125
Statistical metadata 125
Statistical metadata repository 126
Statistical metadata system 126
Statistical metainformation 127
Statistical metainformation system 127
Statistical methodology 127
Statistical microdata 128
Statistical population 128
Statistical processing 128
Statistical production 129
Statistical standard 129
Statistical subject-matter domain 130
Statistical unit 130
Stewardship 131
Stratification 131
Structural definition 131
Structural metadata 132
Structure 132
Study domain 133
Submission 133
Submitting organization 133

Supplementary data 134
Survey 134
Survey data collection 134
Survey design 134
Syntax 135
Target population 135
Taxonomy 135
Term 136
Terminological entry 136
Terminological system 136
Terminology 137
Thesaurus 137
Time coverage 137
Time of recording 137
Time Period 138
Time series 138
Time series breaks 138
Timeliness 139
Transparency 139
Trend 139
Trend estimates 140
True value 140
Type of data collection 140
Under-coverage 140
Unit non-response 141
Unit of measure 141
Unit response rate 141
Unit value 142
Unit value index 142
User needs (for statistics) 142
User satisfaction survey 142
Validation 143
Valuation 143
Value domain 143
Value item 144
Value meaning 144
Variable 145
Verification 145
Weight 146
XML 146
Year-to-date data 146
Sources 148

1. SCOPE OF THE CONTENT-ORIENTED GUIDELINES

This is the first draft release of the SDMX Content-Oriented Guidelines, consisting of:

- *SDMX Content-Oriented Guidelines: Cross-Domain Concepts;*
- *SDMX Content-Oriented Guidelines: Statistical Subject-Matter Domains;*
- *SDMX Content-Oriented Guidelines: Metadata Common Vocabulary.*

It is the intent of these guidelines to establish practices in the use of terminology, the structuring of data and metadata sets, and the classification of data and metadata to support the exchange of data and metadata. The content guidelines are designed to work within the specified SDMX technical framework to produce maximum interoperability in the exchange of data and metadata.

The intent of the SDMX content-oriented guidelines is to encourage reuse where possible across statistical domains in the following areas:

1. Concepts - as described in the "Cross-Domain Concepts" guideline
2. Classifications - as described in the "Statistical Subject-Matter Domains" guideline
3. Terminology - as described in the "Metadata Common Vocabulary" (MCV) guideline

2. METADATA COMMON VOCABULARY

The Metadata Common Vocabulary (MCV) is an SDMX repository which contains concepts[*] (and related definitions) to which terminology used in structural and reference metadata of international organisations and national data producing agencies may be mapped.

The MCV covers a selected range of metadata concepts:

(1) General metadata concepts, mostly derived from ISO, UNECE and UN documents, useful for providing a general context to metadata management;

(2) Metadata terms describing statistical methodologies (frequency, reference period, data collection, source, adjustment, etc.);

(3) Metadata for assessing quality (accuracy, timeliness, etc.), and

(4) Terms referring specifically to data and metadata exchange (terminology from the SDMX information model and from existing data structure definitions, etc.).

More specifically, the MCV provides:

- ISO/IEC 11179-compliant definitions for a wide range of statistical metadata terms, which may be used directly, or against which other terminology systems may be

[*] Note that the term “concept” is used throughout this section in a broad sense, rather than with the narrow definition provided in the SDMX technical standards.

mapped. This set of terms is inclusive of the terminology used within the SDMX Technical Standards;

- definitions for terms on which the SDMX cross-domain metadata concepts work is built. It plays an important role in providing a repository for the common set of metadata terms and their associated definitions that can be used to describe the collection, processing and dissemination of data; and

- other terminology used within the SDMX initiative.

The MCV is not intended to cover the whole range of statistical terminology, as this area is already covered by other general and domain-specific glossaries. The focus of the MCV is largely those terms that are normally used for building and understanding metadata systems. A metadata glossary is necessarily linked to a series of other subject-specific glossaries (on classifications, on data editing, on subject-matter statistical areas) or to more universal statistical glossaries such as Eurostat's CODED or the *OECD Glossary of Statistical Terms*. These more extensive glossaries also contain numerous terms and definitions relevant to specific statistical domains (such as prices, national accounts, external trade, etc.). The insertion within the MCV of some definitions derived from other glossaries should not be seen as a redundancy, but as a means of resolving the complex and interdisciplinary nature of metadata.

Agreement on and updates to the content-oriented guidelines containing and defining SDMX cross-domain metadata concepts imply updating the MCV to reflect the SDMX standards. In addition, since the cross-domain metadata concepts will be subject to revision and supplementation, the MCV will never be considered as complete or final as the need to include new terms, refine existing definitions and provide more context information will always arise.

A value added of the MCV is in the opportunity of having one single entry point for accessing a variety of terms, sometimes not available or hard to find on the Internet. In some cases, the MCV deliberately presents one definition and several context explanations for the same term, always quoting the respective source, sometimes providing additional explanations, other times highlighting peculiarities in how a certain definition is applied within a certain domain or geographical context. Users can live with different metadata models, as long as each concept is well identified and transparent to users. In other words, transparency is a pre-requisite for a correct interpretation (and for convergence) of the different statistical frameworks.

2.1. Point of departure and current status

The MCV project built on work already undertaken by several organisations, rather than confusing the situation by the development of a whole new set of definitions. Where possible, definitions have been drawn from existing international standards or from recommended statistical practices. Where standard definitions were not available or not satisfactory, suitable national definitions have been considered or new definitions formulated.

The MCV glossary is also available on the web through extensive statistical glossary databases such as CODED (Eurostat concepts and definitions database, section "Metadata terminology") or the OECD Glossary of Statistical Terms. Extractions will be available in suitable formats, such as HTML and XML.

The present MCV draft consists of about 380 terms. It presents the following "fields":

- term
- definition
- source
- related terms
- context

As mentioned above, the "context" field is used extensively throughout the glossary, sometimes providing additional explanations, other times highlighting peculiarities in how a certain definition is applied within a certain domain or geographical context.

3. GLOSSARY

Accessibility

The ease and the conditions with which statistical information can be obtained.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Accessibility refers to the availability of statistical information to the user (International Monetary Fund, "Data Quality Assessment Framework - DQAF - Glossary").

Accessibility 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. (Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition)

In SDMX, "Accessibility of Documentation" refers to the availability of documentation of various aspects of the data (sources and methods documents) and the content of such documentation.

Hyperlink

http://www.sdmx.org/

Related terms

Clarity
Integrity
Quality
Simultaneous release

Accounting basis

See "Accounting conventions"

Accounting conventions

Term capturing the practical aspects and conventions used when compiling data from diverse sources under a common methodological framework.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "Accounting Conventions" (also referred to as accounting "basis") refers 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.

Hyperlink

http://www.sdmx.org/

Related terms

Recording of transactions
Reference period

SDMX
Time of recording
Valuation

Accuracy

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

Source

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

Context

In SDMX, "Accuracy" refers to the provision of either measures of accuracy or precision (numerical results of the methods/processes for assessing the accuracy or precision of data) or qualitative assessment indicators. It may also be described in terms of the major sources of error that potentially cause inaccuracy. It includes providing the results of the assessment of source data for coverage, sampling error, response error and non-sampling error.

The accuracy of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. It is usually characterized in terms of error in statistical estimates and is traditionally decomposed into bias (systematic error) and variance (random error) components. It may also be described in terms of the major sources of error that potentially cause inaccuracy (e.g., coverage, sampling, non response, response). (Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 6-7, available at http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1)

Accuracy is the second quality component in the Eurostat Definition.

Hyperlink

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

Related terms

Error of estimation
Precision
Quality (Eurostat context)
Quality (IMF context)
Reliability
SDMX
Statistical error

Adjustment

The set of procedures employed to improve coverage/classification/timing/valuation of the data or to conform to an accounting/recording basis or address data quality differences in compiling specific data sets.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Items covered may include changes in positions during the reference period associated with transactions, exchange rate changes, price changes, seasonal adjustment and other adjustments.

Hyperlink

http://www.sdmx.org/

Related terms

Compilation practices
Revision policy
Seasonal adjustment
Special Data Dissemination Standard (SDDS)

Adjustment Methods

See "Adjustment"

Administered item

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

Source

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

Context

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)

Hyperlink

Related terms

Administration record
Context
Creation date
Data identifier
Date of last change
Effective date
ISO/IEC 11179
Metadata registry
Origin
Registration
Registry item
Stewardship
Submission

Administration record

Collection of administrative information for an administered item.

Source

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

Context

Hyperlink

Related terms

Administered item
ISO/IEC 11179

Administrative data

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

Source
OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Second Draft, Annex 2, Glossary, Paris, 2002

Context

Hyperlink
www.oecd.org/dataoecd/9/20/1963116.pdf

Related terms
Administrative data collection
Administrative source
Data source

Administrative data collection

The set of activities involved in the collection, processing, storage and dissemination of statistical data from one or more administrative sources. The equivalent of a survey but with the source of data being administrative records rather than direct contact with respondents.

Source
OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Second Draft, Annex 2, Glossary, Paris, 2002

Context
In this context, the administrative source is the register of units and data associated with an administrative regulation (or group of regulations) viewed as a source of statistical data.

Hyperlink
www.oecd.org/dataoecd/9/20/1963116.pdf

Related terms
Administrative data
Data collection
Data source

Administrative source

The organisational unit responsible for implementing an administrative regulation (or group of regulations) for which the corresponding register of units and the transactions are viewed as a source of statistical data.

Source
OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Second Draft, Annex 2, Glossary, Paris, 2002

Context

Hyperlink
www.oecd.org/dataoecd/9/20/1963116.pdf

Related terms
Accident at work
Administrative data
Data source

Agency

See Organisation

Aggregation

The combination of related categories, usually within a common branch of a hierarchy, to provide information at a broader level to that at which detailed observations are taken.

Source

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

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).

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, as contrasted with the income of an individual or the price of a single commodity (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003).

Hyperlink

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

Related terms

Compilation practices
Data set
Disaggregation
Grossing/Netting
Special Data Dissemination Standard (SDDS)

Aggregation Equation

See "Aggregation"

Analytical framework

An analytical framework describes the conceptual system of definitions and classifications of the related data.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink

http://www.sdmx.org/

Related terms

Special Data Dissemination Standard (SDDS)

Analytical unit

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

Source
Eurostat, CODED database

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).

Hyperlink

Related terms
Classification
Observation unit
Statistical unit

Area sampling

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

Context
The term may also be used (but is not recommended) as meaning the sampling of a domain to determine area, e.g. under a crop.

Hyperlink

Related terms
Frame
Sampling

Attachment level

A property of attributes in Gesmes/TS.

Source
GESMES/TS User Guide, Release 3.00, February, 2003; unpublished on paper

Context
For each attribute specified in a key family, it is defined whether this attribute takes:
- an independent value for each observation in the data set
- an independent value for each time series in the data set
- an independent value for each sibling group in the data set
- a single value for the entire data set.

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

Related terms
Attribute [Gesmes terminology]
GESMES/TS

Attribute

A characteristic of an object or entity.

Source

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

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).

Within SDMX, 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. Within the SDMX information model, attribute value is the value of an attribute, such as the instance of a coded or uncoded attribute in the context of a data structure.

Hyperlink

Related terms

Attachment level
Basic attribute
Class
Data exchange
Data structure definition
Entity
ISO/IEC 11179

Metadata item
Object
Statistical concept
Value domain

Availability

See "Accessibility"

Base period

The period of time data used as the base of an index number.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

This period is frequently one year (e.g. 1995=100) but it may be as short as one day or as long as the average of a group of years.

Under the SDDS, this refers to the period when the published index = 100, or the reference period to which the average level and/or constant price series data refer.

Hyperlink

http://www.sdmx.org/

Related terms

Base weight

Compilation practices
Reference period
Special Data Dissemination Standard (SDDS)
Weight period

Base weight

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.

Source

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

Context

Hyperlink

Related terms

Base period
Weight

Base year

See "Base period"

Basic attribute

An attribute of a metadata item commonly needed in its specification.

Source

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

Context

Categories of basic attributes:
- Identifying are attributes that are applicable for the identification of a data element.
- Definitional are attributes that describe the semantic aspects of a data element. These attributes may be derived by inheritance from characteristics of data element concepts, objects or entities.
- Relational are attributes that describe associations among data elements and/or associations between data elements and classification schemes, data element concepts, objects, entities.
- Representational are attributes that describe representational aspects of a data element.
- Administrative are attributes that describe management and control aspects of a data element.

A basic attribute that is applicable to all types of metadata item is a common attribute. (ISO/IEC International Standard 11179-3 "Information technology-Metadata registries (MDR)-Part 3: Registry metamodel and basic attributes", February 2003)

Hyperlink

Related terms

Attribute
ISO/IEC 11179

Benchmark

In the 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.

Source
United States Bureau of Census, "Glossary of Selected Abbreviations and Acronyms"; unpublished on paper

Context

Hyperlink
http://eire.census.gov/cgi-bin/ssd/Glossary

Benchmarking

Benchmarking refers to the case where there are two sources of data for the same target variable, with different frequencies, and is concerned with correcting inconsistencies between the different estimates, e.g. quarterly and annual estimates of value-added from different sources.

Source
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 Expert Group, Paris

Context
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 (benchmark: indicator ratio) is extrapolated forward to improve quarterly estimates for the most recent periods for which benchmark data are not yet available.

Hyperlink
www.oecd.org/std/meeting-papers

Related terms
Interpolation

Bias

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.

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

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.

Hyperlink

Bilateral exchange

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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Data exchange
Data sharing exchange
Multilateral exchange

Break

See "Time series break"

Category

Generic term for items at any level within a classification, typically tabulation categories, sections, subsections, divisions, subdivisions, groups, subgroups, classes and subclasses (UN Glossary Classification Terms):

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

Context
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.

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

Related terms
Category Scheme
Structure

Category Scheme

The descriptive information for an arrangement or division of categories into groups based on characteristics, which the objects have in common.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
The category scheme is an artefact for organising categories which themselves link to dataflow definition or metadataflow definition

Hyperlink
http://www.sdmx.org/

Related terms
Category
Characteristic

Census

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

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

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).

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

Related terms
Data collection
Statistical population

Chain index

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

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

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.

Hyperlink

Related terms
Index number

Characteristic

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

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

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

Hyperlink

Related terms
Category Scheme
Concept
Concept Scheme
Data
ISO/IEC 11179
Object
Statistical concept
Statistical subject-matter domain
Time series
Variable

Clarity

Clarity refers to the data's information environment: whether data are accompanied with appropriate metadata, illustrations such as graphs and maps, whether information on their quality is also available (including limitation in use) and the extent to which additional assistance is provided by National Statistical Institutes.

Source

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

Context

The fourth quality component in the Eurostat Definition of quality is "accessibility and clarity".

Hyperlink

Related terms

Accessibility
Quality (Eurostat context)

Class

A description of a set of objects that share the same attributes, operations, methods, relationships, and semantics [ISO/IEC 19501-1:2001, 2.5.2.9]

Source

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

Context

An association is a semantic relationship between two classes. An association is a type of relationship. [Adapted from ISO/IEC 19501-1:2001, 2.5.2.3]
An association class is an association that is also a class. It not only connects a set of classes, but also defines a set of features that belong to the relationship itself.
[Adapted from ISO/IEC 19501-1:2001, 2.5.2.4] (ISO/IEC International Standard 11179-3 "Information technology-Metadata registries (MDR) - Part 3: Registry metamodel and basic attributes", February 2003)

Hyperlink

Related terms

Attribute
ISO/IEC 11179
Object

Classification

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.

Source

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

Context

In SDMX, "Classification Systems" refer to a description of the classification systems being used and how they conform with internationally accepted standards guidelines, or good practices. It also refers to the description of deviations of classification systems compared to accepted statistical standards, guidelines, or good practices, when relevant.
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.

Examples of classification are NACE Rev. 1 (Statistical Classification Of Economic Activities), NUTS (Nomenclature of Territorial Units for statistics), and ISCO-88 (International Standard Classification of Occupations). ISIC is the United Nations International Standard Industrial Classification of All Economic Activities.
Version 2.0 of 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).

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

Related terms
Analytical unit
Classification changes
Classification scheme
Classification unit
Disaggregation
Maintenance Agency

Nomenclature
Observation
Observation unit
SDMX
Standard Classification
Statistical unit
Taxonomy

Classification changes

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 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 category.

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

Context
Changes in classifications and structure comprise changes in sector classification and structure of institutional units and changes in classification of assets and liabilities (Eurostat, "European System of Accounts - ESA 1995", Office for Official Publications of the European Communities, Luxembourg, 1996, par.6.29).

Hyperlink

Related terms
Classification
Institutional unit

Classification scheme

Information for an arrangement or division of objects into groups based on characteristics, which the objects have in common.

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

Context
Attributes of Classification scheme:
"Classification scheme type name" is the name of the type of Classification scheme
"Classification scheme administration record" is the Administration record for a Classification scheme.
"Classification scheme item" is an item of content in a classification scheme. This may be a node in
a taxonomy or ontology, a term in a thesaurus, etc.
Attributes of Classification scheme item:
Classification scheme item type name is the name of the type of the Classification scheme item
Classification scheme item value an instance of a Classification scheme item.
Classification scheme item relationship is the relationship among items within a Classification scheme. Such relations serve to assist navigation through a large number of Classification Scheme Items.
Classification scheme item relationship type description is a description of the type of relationship between
a Classification scheme item and one or more other Classification scheme items in a Classification scheme.
Classification scheme membership is the relationship of a Classification scheme with its items.
(ISO/IEC 11179-3 "Information technology - Metadata registries-Part 3: Registry metamodel and basic attributes", February 2003)

Hyperlink

Related terms
Classification
ISO/IEC 11179

Classification unit

The basic unit to be classified in the classification (e.g. in an activity classification this would be the establishment or enterprise, 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

Context

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

Related terms
Classification

Co-ordination of samples

Increasing the sample overlap for some surveys rather than drawing the samples independently is known as positive coordination. A positive coordination is often searched in repeated surveys

over time (panels) in order to obtain a better accuracy of statistics depending on correlated variables from two surveys. Reducing the overlap between samples for different surveys is known as negative coordination. A negative coordination is used in order to share more equally the response burden among responding units when statistics from surveys are not used together or are not correlated.

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 Paper 4, Office of Federal Statistical Policy Standards, 1978

Context

Hyperlink

Related terms

Sample

Code

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

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

Hyperlink

http://www.sdmx.org/

Related terms

Code list
Coding

Code list

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

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

Context

Each code list has the following properties: a) identifier (it provides a unique identification within the set of code lists specified by a structural definitions maintenance agency); b) name (also unique); c) description (a description of the purpose of the code list); and d) code value length (either an exact or a maximum number of characters and a type, i.e. numeric or alphanumeric).

Hyperlink

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

Related terms

Code
Coding
Dimension
GESMES/TS
Statistical concept
Structural definition

Coding

A technical procedure for converting verbal information into numbers or other symbols which can be more easily counted and tabulated.

Source

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

Context

Hyperlink

Related terms

Code
Code list
Coding error

Coding error

The assignment of an incorrect code to a survey response.

Source

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

Context

Hyperlink

Related terms

Coding

Coherence

Coherence of statistics is their adequacy to be reliably combined in different ways and for various uses.

Source

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

Context

The coherence of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys. Coherence does not necessarily imply full numerical consistency. (Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 7, http://www.statcan.ca:8096/bsolc/english/bsolc?catno=12-539-X&CHROPG=1).
Coherence is the sixth quality component in the Eurostat definition.

Hyperlink

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

Related terms

Comparability
Data confrontation
Quality
SDMX

Collection

See "Data collection"

Comparability

The extent to which differences between statistics from different geographical areas, non-geographical domains, or over time, can be attributed to differences between the true values of the statistics.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "Comparability" is closely associated with "Coherence", which is the adequacy of statistics to be reliably combined in different ways and for various uses. The use of standard concepts, classifications and target populations promotes coherence, as does the use of common methodology across surveys. Coherence does not necessarily imply full numerical consistency.

The comparability of statistics within and across statistical frameworks - and the ability to perform cross-checks and reconciliations - requires the dissemination of components underlying aggregate series, dissemination within a statistical framework, and/or the dissemination of related data that support and encourage users' ability to check and verify the quality of data.

The sources of distortion of comparability in statistics, increasing or reducing it, are mainly twofold:

- use of different concepts/definitions, or
- use of different measuring tools, compilation and presentation practices

Hyperlink

Related terms

Coherence
Quality
SDMX

Compilation

See "Statistical processing"

Compilation practices

See "Statistical processing"

Compiling Agency

Agency that compiled the data being reported.

Source

Status Report on the BIS-IMF-OECD-World Bank Joint External Debt Hub: Prepared by the SDMX Pilot Project Team, May 2005

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 is not sufficient, as one provider could disseminate the data compiled by different compiling agencies.

Hyperlink

Related terms

Concept
Dimension

Completeness

The extent to which all statistics that are needed are available

Source

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

Context

In the European context, completeness is the availability of statistics to meet the requirements of the European Statistical System.

Hyperlink

Related terms

Quality (Eurostat context)

Computation of lowest level indices

Methods used to combine the basic price observations to obtain the first level index (ratio of average prices, average of price relatives or geometric mean; long-term relative from base period vs. short-term relative; weighted or unweighted arithmetic or geometric average).

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Under the SDDS, in the context of labour market this would entail, e.g., weighted or unweighted ratio of average labour service prices, weighted or unweighted arithmetic or geometric average of labour service price relatives, long-term relative from base period versus short-term relative.

Hyperlink

http://www.sdmx.org/

Related terms

Compilation practices
Index number
Special Data Dissemination Standard (SDDS)

Computer Assisted Interviewing, CAI

The use of the computer during interviewing.

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

Context

Any contradictory data can be flagged by edit routines and the resultant 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 interview 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).

Hyperlink

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

Concept

A unit of knowledge created by a unique combination of characteristics [ISO 1087-1:2000, 3.2.1]

Source

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

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).

Hyperlink

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

Related terms

Characteristic
Compiling Agency
Data element concept
Definition
Dimension
ISO/IEC 11179
Key family
Metadata Structure Definition
Ontology
Statistical concept
Statistical subject-matter domain
Structural definition
Terminological system

Concept Scheme

The descriptive information for an arrangement or division of concepts into groups based on characteristics, which the objects have in common

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Within SDMX, a concept scheme is a maintained list of concepts that are used in key family and metadata structure definitions. There can be many such concept schemes. A core representation of the concept can be specified (e.g. a code list, or other representation such as date).

Hyperlink

http://www.sdmx.org/

Related terms

Characteristic [ISO terminology]
Object

Conceptual data model

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

Source

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

Context

Hyperlink

Related terms
Data model
ISO/IEC 11179

Conceptual domain

A set of valid value meanings.

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

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

Hyperlink

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

Confidential data

Data which are subject to confidentiality clauses.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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 primary 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.

Hyperlink
http://www.sdmx.org/

Related terms
Confidentiality

Confidentiality

A property of data, usually resulting from legislative measures, which prevents it from unauthorised disclosure.

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

Context
In SDMX, "Confidentiality" refers to the legislative measures or other formal provision which prevent unauthorised disclosure of data that identify a moral or physical person either directly or indirectly. Also refers to the procedures in place to prevent disclosure of confidential data, including rules applying to staff, aggregation rules when disseminating data, provision of unit records, etc.

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

Related terms
Confidential data
SDMX

Consistency

Consistency refers to logical and numerical coherence.

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

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 revisions 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".

Hyperlink

Related terms
Quality
Serviceability

Consolidation

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

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms
Consolidation (national accounts)

Consolidation (national accounts)

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

Context
Under the SDDS, in the context of fiscal sector data 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
United Nations, "System of National Account (SNA) 1993"

Hyperlink
http://unstats.un.org/unsd/sna1993/introduction.asp

Related terms
Consolidation

Constraint

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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms
Attribute
Data set
Key
Metadata set
Data structure definition

Contact

An instance of a role of an individual or an organization (or organization part or organization person) to whom an information item(s), a material object(s) and/or person(s) can be sent to or from in a specified context.

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

Context
In SDMX, "Contact" describes contact points for the data or metadata, including how to reach the contact points
Attributes of "contacts" (from ISO 11179) 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.
Electronic mail address: An e-mail address for correspondence with the Contact.
Phone number: A telephone number for spoken correspondence with the Contact.
Fax number: A facsimile number for correspondence with the Contact.
Contact information is the information that enables a Contact to be located or communicated with.

Hyperlink

Related terms
ISO/IEC 11179
SDMX
Stewardship
Submission

Context

The context is the circumstances, purpose, and perspective under which an object is defined or used.

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

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).

Hyperlink

Related terms
Administered item
ISO/IEC 11179

Country identifier

An identifier further specifying the geopolitical area associated with the language.

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

Context
1 Metamodel construct is: Attribute of Language Identification.
2 Use of three digit numeric codes from ISO 3166-1, with extensions if required, is recommended by ISO.

Related terms
Identifier

ISO/IEC 11179

Coverage

The population from which observations for a particular topic can be drawn.

Source

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

Context

An understanding of coverage is required to facilitate the comparison of data. Coverage issues are often explained through the use of tables showing linkages (e.g. part or full correspondence) and can also be used to explain the ratio of coverage. The rules and conventions of coverage are largely determined by concept definitions, scope rules, information requirements and, in the case of statistical collections and classifications, collection and counting units and the collection methodology (United Nations Glossary of Classification Terms).

Coverage is a term used in sampling in two senses: (1) to denote the scope of the material collected from the sample members (as distinct from the extent of the survey, which refers to the number of units included); (2) to mean the extent or area covered by the sampling as in expressions such as "50% coverage", which means that one-half of the population under discussion have been examined (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003).

Under the SDDS and in SDMX, the term "Coverage" encompasses the key features of the scope of the data disseminated (e.g., geographic, institutional, product, industry sector, occupation, transaction, etc. as well as relevant exceptions and exclusions), which a user must be aware of in order to use and interpret the data appropriately.

Hyperlink

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

Related terms

Coverage errors
Coverage ratio
Data
Scope
SDMX
Special Data Dissemination Standard (SDDS)

Coverage errors

Coverage errors arise from failure to cover adequately all components of the population being studied. 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

Context

Coverage errors are due to divergences between the target population and the frame. Coverage errors include over-coverage, under-coverage and misclassification.

Hyperlink

Related terms

Coverage
Statistical population

Coverage ratio

The coverage ratio measures 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. 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

Context

Hyperlink

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

Related terms

Coverage
Observation

Creation date

The date the Administered item was created.

Source

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

Context

Hyperlink

Related terms

Administered item
Date
Date of last change
ISO/IEC 11179

Cross-domain Concepts

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Within SDMX, cross-domain “metadata” concepts are envisaged to cover various aspects of the statistical data, including data quality. When exchanging statistics, institutions can select from a standard set of content-oriented concepts. The list of concepts and their definitions reflects good practices and can be the basis for mapping between internal systems when data and metadata are exchanged or shared between and among institutions.
Examples of common concepts are data source used, periodicity, population coverage and seasonal adjustments.

Hyperlink

http://www.sdmx.org/

Related terms

Reference metadata
SDMX
Structural metadata

Cut-off survey

A survey is a survey in which the sample excludes all units that are less than a specified size.

Source
OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Second Draft, Annex 2, Glossary, Paris, 2002

Context

Hyperlink
www.oecd.org/dataoecd/9/20/1963116.pdf

Related terms
Cut-off threshold
Survey

Cut-off threshold

A threshold used, mainly for cost or burden reasons, to exclude from the target population (hence from the frame) units contributing very little to the requested statistics, small businesses for instance.

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

Context

Hyperlink

Related terms
Cut-off survey
Statistical population
Target population

Data

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

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

Context
Data is 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 refers to data from a survey or administrative source used to produce statistics. (OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Annex 2, Glossary, Paris, 2002, available at http://www.oecd.org/dataoecd/9/20/1963116.pdf)

Hyperlink
http://www.oecd.org/dataoecd/9/20/1963116.pdf

Related terms
Characteristic
Coverage
Data presentation
Metadata
Periodicity
Special Data Dissemination Standard (SDDS)
Timeliness

Data analysis

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.

Source

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

Context

Hyperlink

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

Data attribute

See "Attribute"

Data capture

The process by which collected data are put in a machine-readable form.

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

Context

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

Hyperlink

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

Data checking

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

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

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.

Hyperlink

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

Related terms

Data reconciliation

Data collection

The process of gathering 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

Context

Data collection encompasses such concepts as: the type(s) of interview used for data collection (e.g. personal or by telephone, paper and pencil, facsimile, computer-aided personal or telephone interview (CAPI/CATI), or mailed questionnaires); the duration of the field work

(specify the dates); the period used for data collection; whether a permanent survey organisation exists or personnel for each survey round are recruited, etc. Data may be observed, measured, or collected by means of questioning, as in survey or census response.

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

Related terms
Administrative data collection
Census
Compilation practices
Observation
Organisation Role
Period
Special Data Dissemination Standard (SDDS)
Survey
Survey data collection
Type of data collection

Data confrontation

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 their coherency and the reasons for any differences identified. Data confrontation may also take place between statistics produced in different countries. Such processes may or may not attempt to quantify the impact of any differences identified.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Coherence
Data reconciliation

Data consumer

Organisation using data as input for further processing

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Organisation Role

Data dissemination

Dissemination is the release to users of information obtained through a statistical activity.

Source
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 19982003, page 67

Context
Data dissemination consists of distributing or transmitting statistical data to users. Various release media are possible; for example: 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.
In SDMX, "Supplementary Data" refers to a description of data not routinely disseminated that are made available to users upon request. It may include customized tabulations that can be provided (perhaps for a fee) to meet specific requests. Also include information on procedures for obtaining these supplementary data.

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

Related terms
Data Dissemination Standards
Data presentation
Dissemination format
Public disclosure
SDMX

Data dissemination standards

Standards to guide members of the International Monetary Fund in the dissemination to the public of their economic and financial data.

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

Context

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

Related terms
Data dissemination
General Data Dissemination System (GDDS)
Special Data Dissemination Standard (SDDS)

Data editing

Activity aimed at detecting and correcting errors (logical inconsistencies) in 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

Context
Editing techniques refers to a range of procedures and processes used for detecting and handling errors in data. 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. These include:
- validation edits - to check the validity of basic identification of classificatory items in unit data;
- logical edits - ensure that two or more data items do not have contradictory values;
- consistency edits - check to ensure that precise and correct arithmetic relationships exists between two or more data items;
- range edits - identify whether or not a data item value falls inside a determined acceptable range;

- variance edits - involve looking for suspiciously high variances at the output edit stage.
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.
Micro-editing and macro-editing may be distinguished in order to calculate rate of edits.

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

Related terms
Graphical data editing
Macro editing
Micro editing

Data element

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

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

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)

Hyperlink

Related terms
Data element concept
Derivation input
Derivation output
Derivation rule
ISO/IEC 11179
Keyword
Related data reference
Thesaurus

Data element concept

A concept that can be represented in the form of a data element, described independently of any particular representation.

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

Context
Data element concept administration record is the Administration record for a Data element concept.
Data element concept object class is the designation of an Object class for a Data element concept.
Object class qualifier is a qualifier of the Data element concept object class.
Data element concept property is the designation of a Property for a Data element concept.
Data element concept conceptual domain relationship is the relationship between a Data element concept and its Conceptual domain.
Data element concept expression is the relationship between a Data element and a Data element concept.
Data element concept relationship is the relationship among two or more Data element concepts.

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

Hyperlink

Related terms
Concept
Data element
ISO/IEC 11179

Data element derivation

Relationship among a data element which is derived, the rule controlling its derivation, and the data element(s) from which it is derived.

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

Context

Hyperlink

Related terms
Derivation input
Derivation output
Derivation rule
ISO/IEC 11179

Data exchange

The process of sending and receiving data in such a manner that the information content or meaning assigned to the data is not altered during the transmission.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
Data exchange context is the framework in which two or more partners agree to: exchange one or more identified sets of data and related attributes (Exchanged time series, ETS) use one or more key families to serve this requirement, possibly, on some business and implementation agreements. Exchanged time series (ETS) is a collection of data sets (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)

Hyperlink
http://www.sdmx.org/

Related terms
Attribute [Gesmes terminology]
Bilateral exchange
Data exchange context
Data sharing exchange
Electronic data interchange (EDI)
GESMES/TS
Key family
Multilateral exchange
Statistical message

Data exchange context

The framework in which two or more partners agree to: exchange one or more identified sets of data and related attributes (exchanged time series; ETS), and use one or more data structure definitions (key families) to serve this requirement, possibly, on some business and implementation agreements.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

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

Related terms

Attribute
Data exchange
Key family

Data flow definition

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, data sets are reported or disseminated according to a data flow definition. The data flow definition identifies the data structure definition (key family) 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.

A Data structure definition (Key Family) defines the 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 Key Family 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 Key Family definition. The Dataflow Definition may also have additional metadata attached, defining qualitative information and constraints on the use of the Key Family, in terms of reporting periodicity or specifying the subset of codes to be used in a dimension.

Hyperlink

Related terms

Attribute
Category
Code
Data flow
Data set
Data structure definition
Definition
Dimension
Measure
Metadata flow definition

Data identifier

The unique identifier for an administered item within a registration authority.

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

Context

Hyperlink

Related terms
Administered item
Data set identifier
Identifier
ISO/IEC 11179

Data interchange

See "Data exchange"

Data item

An occurrence of a data element.

Source
ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, 1999

Context
A data item is a specific sub-component of a data record. For instance, in a population census, specific data items might be last name, first name, sex, and age (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
Data element
ISO/IEC 11179
Variable

Data model

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

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

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)

Hyperlink

Related terms
Conceptual data model
Conceptual domain
GESMES/TS data model
ISO/IEC 11179
Metamodel

Data presentation

Description of the way the data are presented.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "Data Presentation" includes the description of the table contents, with their data breakdowns. It should also include summary information on units of measurement, time span covered, adjustments to data (e.g., seasonal adjustments for time series) and availability of textual analysis of current-period development with the dissemination of the data.

Hyperlink

Related terms

Data
Data set
Data dissemination
SDMX

Data processing

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

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

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.

Hyperlink

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

Related terms

Compilation practices
Processing error

Data provider

Organisation which produces data or metadata.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

Related terms

Data

Data provider series key
Metadata
Data source

Data provider series key

Identifier used by the data providers systems

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Data provider
Identifier
Key (time series or sibling group)
Time series

Data reconciliation

Frequently used as a synonym for Data confrontation. In the more active sense, the term implies the process of adjusting data derived from two different sources to remove or at least reduce the impact of differences identified in the process of data confrontation.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Editing and reconciliation may involve fixing errors or adopting alternative sources and methods that are aimed at improving the process of reviewing or understanding data.
(International Monetary Fund (IMF)," Quarterly National Accounts Manual", Washington D.C., 2001, available at:http://www.imf.org/external/pubs/ft/qna/2000/textbook)
Under the SDDS, this may entail-according to the data category under consideration-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.

Hyperlink

http://www.sdmx.org/

Related terms

Compilation practices
Data checking
Data confrontation
Special Data Dissemination Standard (SDDS)

Data security

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

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

Context

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

Data set

Any organised collection of 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

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).

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

Related terms
Aggregation
Data presentation
GESMES/CB
Key family
Measure
Sibling group
Time series

Data set identifier

See "Data identifier"

Data sharing exchange

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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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

Hyperlink
http://www.sdmx.org/

Related terms
Bilateral exchange

Data exchange
Multilateral exchange

Data source

A specific data set, metadata set, database or metadata repository from where data or metadata are available.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

The source of data is often used as a synonym for the term “data provider” However, in the context of SDMX the latter term refers to the organisation or individual from where statistics are obtained. The term "data source" refers to the characteristics and components of the raw statistical data used for compiling statistical aggregates. Sources can be distinguished, according to the modality of data collection, in: a) administrative (for data coming from administrative records); b) survey (for data coming from surveys for a specific sector or institutional unit).

Hyperlink

http://www.sdmx.org/

Related terms

Administrative data
Administrative data collection
Administrative source
Data provider
Organisation
Primary source of statistical data
Provision Agreement
Revision policy
Secondary source of statistical data
Survey data collection

Data status upon release

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Under the SDDS the data identified as meeting the standard may be preliminary and subject to revision and designated as such. On the other hand, estimates that are not based on data collected for the given reference period would not be considered preliminary and would not be in line with the SDDS specifications.

Hyperlink

http://www.sdmx.org/

Related terms

Revision policy
Special Data Dissemination Standard (SDDS)

Data structure definition

Set of structural metadata associated to a data set, which include 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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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 key family. 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).

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

Related terms
Attribute
Concept
Data exchange
Data exchange context
Data set
Dimension
GESMES/TS
Key (time series or sibling group)
Key family
Maintenance Agency
Measure
Statistical concept
Structural definition
Structural metadata

Datatype

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

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

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.

Hyperlink

Related terms
ISO/IEC 11179

Date

A time reference.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms
Creation date
Date of last change
Effective date

Date of last change

The date the administered item was last changed.

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

Context
In SDMX, "Date of Update" refers to the date on which the metadata element was inserted or modified in the database. It can be further detailed in: a) last update of content; b) last certified without update; c) last posted on web site.
In more general terms, this refers to the date on which either data or metadata were inserted or modified in the database.

Hyperlink

Related terms
Administered item
Creation date
Date
Effective date
ISO/IEC 11179
SDMX
Time of recording

Definition

A statement of the precise meaning of something.

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

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.
According to ISO/IEC International Standard 11179-3 "Information technology - Metadata registries-Part 3: Registry metamodel and basic attributes", February 2003, 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]".
Definition (of Administered Item) is the definition of an Administered item within a Context.
Definition source reference is a reference to the source from which the Definition is taken.
Definition text is the text of the Definition.

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

Related terms
Concept
ISO/IEC 11179

Derivation input

Relationship specifying the source Data element(s) for a Data element derivation.

Source

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

Context

Hyperlink

Related terms

Data element
Data element derivation
ISO/IEC 11179

Derivation output

Relationship denoting the result of a Data element.

Source

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

Context

Hyperlink

Related terms

Data element
Data element derivation
ISO/IEC 11179

Derivation rule

Derivation rule is the logical, mathematical, and/or other operations specifying derivation.

Source

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

Context

Derivation rule administration record is the Administration record for a Derivation rule.
Derivation rule application is the relationship specifying the Derivation rule for a Data element derivation.
Derivation rule specification is the text of a specification of Data element derivation.

Hyperlink

Related terms

Data element
Data element derivation
ISO/IEC 11179

Derived data element

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

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

Context

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

Derived statistic

A derived statistic is obtained by an arithmetical observation from the primary observations. 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

Context

Hyperlink

Related terms
Observation

Dimension

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

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
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 in the other and the inverse of the function maps values in the second units back to values in the first.
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)

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

Related terms
Code list
Compiling Agency
Concept
Cube
Dimensionality
GESMES/TS
Key (time series or sibling group)
Key family

Key structure
Statistical concept
Time series
Unit of measure

Dimensionality

An expression of measurement without units.

Source

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

Context

A quantity is a value with an associated unit of measure. 32 Fahrenheit, 0 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.

Hyperlink

Related terms

Dimension
ISO/IEC 11179
Unit of measure

Disaggregation

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.

Source

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

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).

Hyperlink

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

Related terms

Aggregation
Classification
Compilation practices
Observation

Disclosure analysis

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.

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

Context

Hyperlink

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

Dissemination format

Media by which statistical data and /or metadata are disseminated to users.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "Dissemination Formats" refers to the various means of dissemination used for making the data available to the public. It would include a description of the various formats available, including where and how to get the information (for instance paper, electronic publications, on-line databases).

Under the SDDS, the concept of dissemination formats is divided into two categories: "hardcopy" and "electronic" publications, which detail the reference documents through which users may access the data described in the metadata and, where relevant, detailed components beyond the minimum prescribed.

Hyperlink

http://www.sdmx.org/

Related terms

Data dissemination

SDMX

Documentation

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

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

Context

Hyperlink

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

Domain groups

A domain group comprises 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.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms
Statistical subject-matter domain

Dublin Core

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.

Source
Dublin Core

Context

Hyperlink
http://www.dublincore.org

Related terms
Glossary

EDIFACT

Electronic Data Interchange for administration, commerce and transport. 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) Application level syntax rules, September 1996

Context

Hyperlink
http://www.nls.fi/ptk/standardisation/2.html

Related terms
EDI
Electronic data interchange (EDI)
GESMES
SDMX-EDI

Electronic data interchange (EDI)

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

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

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.

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

Related terms
Data exchange
EDIFACT
GESMES

Entity

A concrete or abstract thing including associations among these things e.g. a person, object, event, idea, process, etc.

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

Context
An entity exists whether data about it are available or not. [ISO/IEC 2382-17:1999, 17.02.05]

Hyperlink

Related terms
Attribute
ISO/IEC 11179
Observation unit
Ontology

Error of estimation

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

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

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.

Hyperlink

Related terms
Accuracy
Estimate

Error of observation

An error arising from imperfections in the method of observing a quantity, whether due to instrumental or to human factors.

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

Context

Hyperlink

Related terms

Estimate

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

Source

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

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.

Hyperlink

Related terms

Error of estimation
Estimator
Non-sampling error
Reliability [Quality reports]
Trend estimates

Estimation

Estimation is concerned with inference about the numerical value of unknown population values from incomplete data such as a sample. If a single figure is calculated for each unknown parameter the process is called "point estimation". If an interval is calculated within which the parameter is likely, in some sense, to lie, the process is called "interval estimation".

Source

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

Context

Sample survey data only relates to the units in the sample. Therefore the sample estimates need to be inflated to represent the whole population of interest. Estimation is the means by which this inflation occurs. The estimation process is also referred to as "grossing up".
Under the SDDS, estimation is of particular relevance to deriving missing data in the compilation of national accounts and consumer and producer aggregates (e.g., extrapolation of annual benchmark using value or volume changes from industrial surveys, use of fixed input/output ratios, etc, for current price for national accounts, and price imputations for consumer and producer prices) and the techniques applied in updating the ancillary information used in the estimation process.

Hyperlink

Related terms

Compilation practices
Estimator
Number raised estimation
Precision Ratio estimation
Special Data Dissemination Standard (SDDS)

Estimator

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

Source

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

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.

Hyperlink

Related terms
Estimate
Estimation

Expected value

The hypothetical average from the conceived replicates of the survey all conducted under the same essential conditions.

Source
Federal Committee on Statistical Methodology, "Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic", Statistical Policy Working Paper, December, 1978

Context

Hyperlink

Related terms

Expression Node

A node in a transformation scheme that is part of a hierarchy of nodes that together define or document an expression.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms

Flag

An attribute of a cell in a data set representing qualitative information on the value of that cell.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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...".

Hyperlink
http://www.sdmx.org/

Related terms
Footnote
Quantitative data

Flow data series

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.

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

Context

Hyperlink

Related terms
Statistical metadata repository
Stock data series

Follow-up

A further attempt to obtain information from an individual in a survey or field experiment because the initial attempt has failed or later information is available.

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

Context

Hyperlink

Related terms
Non-response
Non-response error

Footnote

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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Flag

Frame

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

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

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 of 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.

Hyperlink

Related terms

Area sampling
Frame error
Under-coverage

Frame error

Frame error may be caused by the inherent limitations of input data, or by delays and errors in data acquisition and processing.

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 Paper 4, Office of Federal Statistical Policy Standards, 1978

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.

Hyperlink

Related terms

Frame

Frequency

The rate at which something happens or is repeated.

Source

The Oxford Advanced Learners Dictionary, Oxford University Press

Context

If a time series has a constant time interval between its observations, this interval determines the frequency of the time series (e.g. monthly, quarterly, yearly).

In GESMES/TS, frequency is a dimension of the time series key. Frequency must be assigned as a dimension in every key family and it has to be the first dimension. ("GESMES/TS User Guide", Release 3)

In SDMX, "Frequency" is closely associated with "Periodicity" to form a single entity, named "frequency and periodicity". While frequency refers to the time interval between the observations of a time series. 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).

Hyperlink

Related terms
Periodicity
SDMX

Gateway

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.

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

Context

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

Related terms

Gateway exchange

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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms

General Data Dissemination System (GDDS)

The GDDS is a structured process through which Fund member countries 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.

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

Context

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

Related terms
Data Dissemination Standards
Special Data Dissemination Standard (SDDS)

Geographical coverage

The country or geographic area which is related to the measured economic phenomenon.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
This entity is also commonly called reporter.

Hyperlink
http://www.sdmx.org/

Related terms
Coverage
Reference area

GESMES

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

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

Context

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

Related terms
Attribute
EDIFACT [ISO terminology]
Electronic data interchange (EDI)
GESMES/CB
GESMES/TS
Statistical message

GESMES/CB

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

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

Context
The message has been renamed from GESMES/CB to GESMES/TS in 2003

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

Related terms
Data set
GESMES
GESMES/TS

GESMES/TS

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.

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

Context
The message 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.

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

Related terms
Attachment level
Attribute
Code list
Data exchange
Dimension
GESMES
GESMES/CB
GESMES/TS data model
Key family
Maintenance Agency
Sibling group
Statistical concept
Structural definition
Structural metadata

GESMES/TS data model

A time-series data exchange model which allows to exchange and identify time series through a multidimensional key and various associated metadata.

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

Context

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

Related terms
Data item
Data model
GESMES/TS
Key (time series or sibling group)

Glossary

An alphabetized list of terms with definitions often created by an organization to reflect its needs. Normally lacks hierarchical arrangement or cross references. Also known as a term list.

Source
Dublin Core

Context
A glossary also 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).

Hyperlink
http://www.dublincore.org

Related terms
Dublin Core
Maintenance Agency

Graphical data editing

Use of graphs to identify anomalies in 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

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.

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

Related terms
Data editing

Grossing/Netting

Combinations in which all elementary 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 recordings.

Source
United Nations, "System of National Account (SNA) 1993"

Context
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.

Hyperlink
http://unstats.un.org/unsd/sna1993/introduction.asp

Related terms
Aggregation
Consolidation

Guidelines

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

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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

Hyperlink
http://www.sdmx.org/

Hierarchy

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.

Source

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

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 is not organised into formal levels.

Hyperlink

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

Related terms

Structure

Identifier

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

Source

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

Context

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

Hyperlink

Related terms

Country identifier
Data identifier
Data Provider Series Key
ISO/IEC 11179
Organisation identifier

Imputation

Imputation is a procedure for entering a value for a specific data item where the response is missing or unusable.

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

Context

Imputation is the process used to determine and assign replacement values for missing, invalid or inconsistent data that have failed edits. This is 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)

Hyperlink

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

Related terms

Missing data

Index number

A quantity which shows by its variations the changes of a magnitude over time or space.

Source

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

Context

Index type refers to any of the various indices (e.g., Laspeyres, modified Laspeyres, Paasche, Value-Added, Fisher, Tornqvist, etc.) used in the statistical production process.
Important features in the construction of an index number are its coverage, base period, weighting system and method of averaging observations. A price index reflects an average of the proportionate changes in the prices of a specified set of goods and services between two periods of time (United Nations, "System of National Account (SNA) 1993", 16.14, available at http://unstats.un.org/unsd/sna1993/introduction.asp).

Hyperlink

Related terms

Chain index
Compilation practices
Computation of lowest level indices
Weight period

Information

Information is knowledge concerning any objects such as facts, events, things, processes or ideas including concepts that within a certain context have a particular meaning.

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

Context

Hyperlink

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

Information system

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

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

Context

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

Hyperlink

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

Inlier

A data value that lies in the interior of a statistical distribution and is in error. Because inliers are difficult to distinguish from good data values they are sometimes difficult to find and correct.

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

Context
A simple example of an inlier might be a value in a record reported in the wrong units, say degrees Fahrenheit instead of degrees Celsius.

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

Related terms
Outliers

Institutional framework

A set of rules used as the basis for producing statistics.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
In SDMX, "Institutional Framework" refers to a law or other formal provision that assign primary responsibility as well as the authority to an agency for the collection, processing, and dissemination of the statistics; it also includes arrangements or procedures to facilitate data sharing and coordination between data producing agencies ("reporting requirements").
In detail, the institutional framework for statistics may include information on:
- The legislation within which the statistical agency operates. Typically it is proclaimed in one or more statistics acts and in accompanying or supplementary government regulations covering a number of issues including the right to collect data; ensuring confidentiality of data collected, etc.
- The organisational structure of the statistical agency. Such structures include economic data collection, processing and analysis; social data collection, processing and analysis; national accounts, balance of payments and economic analysis.
- Planning framework under which major initiatives and statistical outputs envisaged for the future are envisaged.
- Quality consciousness and organisational culture

Hyperlink
http://www.sdmx.org/

Related terms
Integrity
Internal access
Ministerial commentary
SDMX

Institutional sector

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. 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

Context

Hyperlink

Related terms
Activity
Institutional unit

Institutional unit

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 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.

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

Context
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 (Eurostat, "European System of Accounts - ESA 1995", Office for Official Publications of the European Communities, Luxembourg, 1996, 2.12).
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)

Hyperlink

Related terms
Classification changes
Institutional sector

Integrity

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

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

Context
In SDMX, "Transparency" describes 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
Confidence by users is built over time. One important aspect 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.
Under the SDDS, "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.

Hyperlink

Related terms
Accessibility
Institutional framework
Internal access
Ministerial commentary
Quality
Professionalism

Revision policy
SDMX
Special Data Dissemination Standard (SDDS)

Internal access

Internal access refers to giving full transparency to any necessary pre-release access within government-as deemed appropriate by the government.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Under the SDDS, this entails the listing of persons or officials holding designated positions within the government, but outside the agency producing the data, who have pre-release access to the data and the reporting of the schedule according to which they receive access.

Hyperlink

http://www.sdmx.org/

Related terms

Institutional framework
Integrity
Ministerial commentary
Revision policy
Special Data Dissemination Standard (SDDS)

International code designator

An identifier of an organization identification scheme. [ISO/IEC 6523-1:1998, 3.8]

Source

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

Context

Hyperlink

Related terms

ISO/IEC 11179

International statistical standard

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Statistical concept
Statistical standard

Interpolation

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

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

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.

Hyperlink

Related terms
Benchmarking

Interviewer error

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

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

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.

Hyperlink

Related terms

ISO/IEC 11179

The International Standard ISO/IEC 11179 on metadata registries addresses the semantics of data, the representation of data, and the registration of the descriptions of data. ISO/IEC 11179 specifies the kind and quality of metadata necessary to describe data, and it specifies the management and administration of that metadata in a metadata registry (MDR).

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

Context
The purposes of the ISO/IEC 11179 standard are to promote the following:
- standard description of data
- common understanding of data across organizational elements and between organizations
- re-use and standardization of data over time, space, and applications
- harmonization and standardization of data within an organization and across organizations
- management of the components of data

- re-use of the components of data

ISO/IEC 11179 is 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

Hyperlink

Related terms

Administered item
Administration record
Attribute
Basic attribute
Characteristic
Class
Classification scheme
Concept
Conceptual data model
Conceptual domain
Contact
Context
Country identifier
Creation date
Data element
Data element concept
Data element derivation
Data identifier
Data item
Data model
Datatype
Date of last change
Definition
Derivation input
Derivation output
Derivation rule
Dimensionality
Effective date
Entity
Identifier
International code designator
Keyword
Language
Metadata
Metadata item
Metadata object
Metadata registry
Metadata set
Metamodel
Name
Object
Object class
Object class term
Organisation
Organisation identifier

Permissible value
Permitted value
Preferred definition
Property
Reference document
Register
Registrar
Registration
Registration authority
Registry item
Registry metamodel
Related data reference
Related metadata reference
Relationship
Responsible organization
Semantics
Special language
Stewardship
Submission
Submitting organization
Syntax
Taxonomy
Terminological entry
Terminological system
Thesaurus
Unit of measure
Value domain
Value item
Value meaning

Item response rate

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.

Source

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

Context

Hyperlink

Related terms

Non-response rate
Refusal rate
Response rate

Key (time series or sibling group)

The key uniquely defines a time series or sibling group within a data set.

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

Context

Every time series takes a value for every dimension of the 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.

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

Related terms
Data Provider Series Key
Dimension
GESMES/TS data model
Key family
Sibling group
Time series

Key family

See “Data Structure Definition”

Key structure

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

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

Context

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

Related terms
Dimension
Statistical concept

Keyword

One or more significant words used for retrieval of data elements.

Source
ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, 1999

Context

Hyperlink

Related terms
Data element
ISO/IEC 11179

Language

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

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

Context

Hyperlink

Related terms
ISO/IEC 11179
Special language

Level

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

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary (adapted from the Neuchâtel terminology)

Context

Related terms
Code
Hierarchy

Levels of data

Data expressed as levels are expressed in absolute terms (values, numbers, units) for a given period (month, quarter, year).

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

Context
Flow data for annual and quarterly levels may be presented as either the:
- sum of the component quarters or months, or
- average of the component months or quarters.

Stock data by definition only have one value for each frequency (i.e. year, quarter, month).

Hyperlink

Related terms

Longitudinal data

Data in which many units are observed over multiple time periods.

Source
U.S. Department of Labor, Bureau of Labor Statistics, "Glossary", unpublished on paper

Context

Hyperlink
http://stats.bls.gov/bls/glossary.htm

Related terms
Community statistics on income and living conditions (EU-SILC)

Macro editing

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

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

Context

A macro-edit detects individual errors by:
1) checks on aggregated data, or
2) 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

Hyperlink

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

Related terms

Data editing
Micro editing

Maintenance Agency

Organisation responsible for maintaining or updating artefacts such as statistical classifications, glossaries, data structure definitions (key families) and metadata structure definitions

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Classification
Data structure definition
GESMES/TS
Glossary
Key family
Structural definition

Measure

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

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

Hyperlink

http://www.sdmx.org/

Related terms

Data set
Key family
Observation

Measurement error

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. Such errors may be random or they may result in a systematic bias if they are not random.

Source
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 59

Context
Measurement error includes the error in a survey response as a result of respondent confusion, ignorance, carelessness, or dishonesty; the error attributable to the interviewer, perhaps as 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.

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

Related terms

Metadata

Data that defines and describes other data.

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

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.

Hyperlink

Related terms
Data
ISO/IEC 11179
Metadata layer
Metadata registry
Statistical metadata
Statistical metadata system

Metadata Attribute

See attribute

Metadata dimension

The higher level of the metadata structure (e.g., data, access, integrity and quality in the SDDS format), which, combined with elements (e.g., coverage, periodicity, and timeliness) forms the basic framework under which data are described.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
The SDDS prescribes that subscribing members provide a summary description of methodology for each data category, including statements of major differences from international guidelines. The term "methodology" is used in the SDDS in a broad sense to cover the aspects of analytical framework, concepts, definitions, classifications, accounting conventions, sources of data, and compilation practices.

Hyperlink
http://www.sdmx.org/

Related terms
Special Data Dissemination Standard (SDDS)

Metadataflow definition

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

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
A “metadata flow”, in this context, is an abstract concept of the metadata sets, i.e. a structure without any actual metadata. A Metadataflow definition associates a Metadata structure definition with one or more category (possibly from different category schemes). This gives a system the ability to state which metadata sets are to be reported/disseminated for a given category, and which metadata sets can be reported using the Metadata structure definition.

Hyperlink

Related terms
Category
Data flow definition
Definition
Metadata
Metadata set

Metadata item

An instance of a metadata object.

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

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).

Hyperlink

Related terms
Attribute
ISO/IEC 11179
Metadata object
Registry
Registry item
Related metadata reference

Metadata layer

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

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

Context

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

Related terms
Metadata
Metadata registry
Statistical metadata
Statistical metadata system
Statistical metainformation

Metadata object

An object type defined by a metamodel.

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

Context

Hyperlink

Related terms
ISO/IEC 11179
Metadata item
Metamodel
Object

Metadata registry

Information system for registering metadata.

Source
ISO/IEC International Standard 11179-3 - Information technology - Metadata registries (MDR) - Part 1: Framework, March 2004.

Context
Within ISO/IEC International standard 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.

Hyperlink

Related terms
Administered item
ISO/IEC 11179
Metadata
Metadata layer
Registry
Registry item
Registry metamodel
SDMX registry
Statistical metadata

Submitting organization

Metadata set

A collection of metadata.

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

Context

Hyperlink

Related terms
ISO/IEC 11179

Metadata Structure Definition

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

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Concept
Maintenance Agency
Reference metadata
Structural metadata
Structure

Metamodel

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

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

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

Hyperlink

Related terms
Data model
ISO/IEC 11179
Metadata object
Registry metamodel

Methodological soundness

Methodological soundness refers to constructs and principles of accounting that are basic building blocks of macroeconomic data.

Source

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

Context

Hyperlink

Related terms

Quality (IMF context)

Methodology

A structured approach to solve a problem.

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

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).

Hyperlink

http://www.statcan.ca/english/edu/power/glossary/gloss.htm

Related terms

Statistical methodology

Statistical subject-matter domain

Micro editing

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

Source

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

Context

Editing done at the record, or questionnaire level.

Hyperlink

Related terms

Data editing

Macro editing

Ministerial commentary

Internal access refers to the practice of giving full transparency to any necessary pre-release access within government-as deemed appropriate by the government.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink

http://www.sdmx.org/

Related terms
Institutional framework
Integrity
Internal access
Revision policy
Special Data Dissemination Standard (SDDS)

Misclassification

When a subject is falsely classified into a category in which the subject does not belong. It may result from misreporting by study subjects, from the use of less than optimal measurement devices, or from random error.

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

Context

Hyperlink

Missing data

Observations which were planned and are missing.

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

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).

Hyperlink

Related terms
Imputation
Non-response
Observation

Model assumption error

Model assumption errors occur 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 order to calculate statistics or indexes.

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

Context

Hyperlink

Multilateral exchange

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 schedule, mode used for

communications and the actual content of the exchange). This exchange process is also known as gateway exchange.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink

http://www.sdmx.org/

Related terms

Bilateral exchange
Data exchange
Data sharing exchange

Name

The designation of an object by a linguistic expression

Source

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

Context

Hyperlink

Related terms

ISO/IEC 11179

Nature of the basic data

See "Source data"

Nomenclature

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.

Source

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

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.

Hyperlink

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

Related terms

Classification

Non-probability sample

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.

Source

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

Context

Hyperlink

Related terms

Probability sample

Non-response

A form of non observation present in most surveys. Non response means failure to obtain a measurement on one or more study variables for one or more elements k selected for the survey. The term encompasses a wide variety of reasons for non observation: "impossible to contact", "not at home", "unable to answer", "incapacity", "hard core refusal", "inaccessible", "unreturned questionnaire", and others. In the first two cases contact with the selected element is never established.

Source

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

Context

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

There are two 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".

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

Non-response bias

See "Non-response error"

Non-response error

Non-response errors occur when the survey fails to get a response to one, or possibly all, of the questions.

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

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).

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

Related terms
Follow-up
Non-response
Weight

Non-response rate

In sample surveys, the failure to obtain information from a designated individual for any reason (death, absence or refusal to reply) is often called a non-response and the proportion of such individuals of the sample aimed at is called the non-response rate.

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

Context
It would be better, however, to call this a "failure" rate or "non-achievement" rate and to confine "non-response" to those cases where the individual concerned is contacted but refuses to reply or is unable to do so for reasons such as deafness or illness.
Non-availability of information in other situations, e.g. arrival of the investigator for crop cutting experiments after harvesting, may also be termed non-response, or better, non-achievement.
When several items of information are to be collected for the same sample unit, it may so happen that information is not available for some of the items but available for others. The term non-response is usually not applied in such a situation; but incomplete response or incomplete achievement may be used.

Hyperlink

Related terms
Item response rate
Non-response
Refusal rate
Response rate
Sample

Non-sampling error

An error in sample estimates which cannot be attributed to sampling fluctuations.

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

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

Hyperlink

Related terms
Estimate
Sampling error

Not seasonally adjusted series

Data series not subject to the seasonal adjustment process. In other words, the effects of regular, or seasonal, patterns have not been removed from these series.

Source
U.S. Department of Labor, Bureau of Labor Statistics, "Glossary", unpublished on paper

Context

Hyperlink
http://stats.bls.gov/bls/glossary.htm

Related terms
Seasonal adjustment

Number raised estimation

The application of weights to the individual survey records. Number-raised weights are given by N/n (where N is the total number of units in the population for the stratum, and n is the number of responding units in the sample for that stratum). The weight assigned to each survey unit indicates the number of units in the target population that the survey unit is meant to represent. For example, a survey unit with a weight of 100 represents 100 units in the population. Using number-raised weights, each survey unit in a stratum is given the same weight. Number-raised weights can only be used to weight simple random samples.

Source
Australian Bureau of Statistics, Statistical Concepts Library, "Labour Statistics: Concepts, Sources and Methods", Chapter 16 - Overview of Survey Methods, Canberra, 2001

Context
The advantages of number-raised estimation are: it does not require auxiliary data; it is unbiased; and the accuracy of the estimates can be calculated relatively simply. However, number-raised estimation is not as accurate as some other methods.

Hyperlink
http://www.abs.gov.au/AUSSTATS/abs%40.nsf/7884593a92027766ca2568b5007b8617/93a0165bdf598509ca256aa000036c90!OpenDocument

Related terms
Estimation

Object

Anything perceivable or conceivable

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

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.

Hyperlink

Related terms

Attribute
Characteristic [ISO terminology]
Class
Concept Scheme
ISO/IEC 11179
Metadata object
Object class
Ontology
Property

Object class

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

Source

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

Context

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

Hyperlink

Related terms

ISO/IEC 11179
Object
Property

Objectives

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.

Source

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

Context

Hyperlink

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

Related terms

Observation

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms
Classification
Coverage ratio
Data collection
Derived statistic
Disaggregation
Measure
Missing data
Non-response
Observation unit
Pre-Break Value
Statistical concept
Time series
Variable

Observation confidentiality

See "Confidentiality"

Observation unit

Those entities on which information is received and statistics are compiled.

Source
Statistical Office of the United Nations, "International Standard Industrial Classification of all Economic Activities, Third Revision", Statistical Papers Series M No. 4, Rev. 3, United Nations, New York, 1990, para. 63

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.

Hyperlink

Related terms
Analytical unit
Classification
Entity
Observation
Statistical unit

Observation value

See "Observation"

Ontology

A formal specification of a conceptualization; 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.

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

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, 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).

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

Related terms
Concept
Entity
Object
Taxonomy

Organisation

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

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

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).
An organization name is a designation for the Organization

Hyperlink

Related terms
Agency
Data source
ISO/IEC 11179
Organisation identifier
Organisation Role
Responsible organization
Stewardship

Organisation identifier

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

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

Context

Hyperlink

Related terms
Identifier
ISO/IEC 11179
Organisation

Organisation Role

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

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Data collection
Data Consumer
Dissemination
Organisation

Origin

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

Source

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

Context

Hyperlink

Related terms

Administered item

Out-of-scope units

Units are units that should not be included in the sampling frame because they do not belong to the target population in the reference period. If enumerated, they cause over-coverage.

Source

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

Context

Hyperlink

Related terms

Over-coverage
Scope

Outliers

A data value that lies in the tail of the statistical distribution of a set of data values.

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

Context

The intuition is that outliers in the distribution of uncorrected (raw) data are more likely to be incorrect. Examples of outliers are data values that lie in the tails of the distributions of ratios of

two fields (ratio edits), weighted sums of fields (linear inequality edits), and Mahalanobis distributions (multivariate normal) or outlying points to point clouds of graphs.

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

Related terms
Inlier

Over-coverage

Errors which occur due to the inclusion in the sample of elements that do not belong there.

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

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).

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

Related terms
Out-of-scope units
Under-coverage

Period

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

Source
McGraw-Hill Encyclopedia of Science and Technology

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. (GESMES/TS User Guide", Release 3)

Hyperlink

Related terms
Data collection
Reference period
Weight period

Periodicity

Frequency of compilation of the data.

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

Context
In SDMX, "Periodicity" is closely associated with "Frequency" to form a single entity, named "Frequency and Periodicity". While frequency refers to the time interval between the observations of a time series. 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 from respondent or administrative sources.

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

Related terms
Data
Frequency
Release calendar
SDMX

Permissible value

An expression of a value meaning allowed in a specific value domain.

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

Context
Permissible value meaning is the relationship of a Value meaning from an Enumerated conceptual domain with a Permissible value from an Enumerated value domain.
Permissible value set is the set of Permissible values for an Enumerated value domain.
Attributes of Permissible value:
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.

Hyperlink

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

Permitted value

The use of a value as a Permissible Value in an Enumerated Value Domain.

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

Context

Hyperlink

Related terms
ISO/IEC 11179

Permissible value
Value domain

Pre-break observation

See Pre-break value

Pre-Break Value

The observation, at a time series break period, that was calculated using the "old" methodology, with the "observation" being calculated following the "new" methodology.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

SDMX allows for a pre-break value in the case of a series break (where you would use the observation value to show the post-break value)

Hyperlink

http://www.sdmx.org/

Related terms

Observation
Time series
Time series breaks

Precision

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

Source

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

Context

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

Hyperlink

Related terms

Accuracy
Estimation

Preferred definition

Preferred definition is an indicator that the definition text is a preferred definition for an Administered Item within a language.

Source

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

Context

Hyperlink

Related terms

ISO/IEC 11179

Prerequisites of quality

Prerequisites of Quality refer to overarching institutional conditions for the pursuit of data quality.

Source

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

Context

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.

The DQAF groups the indicators of this kind into three elements: legal and institutional environment, resources, and quality awareness.

Hyperlink

Related terms

Quality (IMF context)

Primary data

The most important inputs from among the universe of institutional, administrative, sample survey and/or census based information used in compiling statistical aggregates.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Basic statistical data
Secondary source of statistical data
Special Data Dissemination Standard (SDDS)

Primary source of statistical data

The organisation or individual responsible for the collection and aggregation of data from their.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

For information derived from surveys or censuses such data comprises 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.

Hyperlink

http://www.sdmx.org/

Related terms

Data source
Secondary source of statistical data

Probability sample

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.

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

Context

Hyperlink

Related terms
Non-probability sample
Sample

Processing error

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

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

Context
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 errors of transcription, errors of coding, errors of data entry and errors of arithmetic in tabulation (The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003).

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

Related terms
Data processing
Survey

Product

The representative groups of goods and/or services - and the varieties within them - used to compile the basic statistical data from which an index is derived.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
Under the SDDS, this point should include information on the approach used to select items, replace them when they become permanently unavailable, and introduce new products in the item structure prior to the next official weight update.

Hyperlink
http://www.sdmx.org/

Related terms

Professionalism

The standard, skill and ability suitable for producing statistics of good quality.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "Professionalism and Ethical Standards" 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 based 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.

Hyperlink

http://www.sdmx.org/

Related terms

Quality
SDMX

Property

A characteristic common to all members of an object class.

Source

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

Context

A property qualifier is a qualifier of the element concept property. Property administration record is the Administration record for a property.
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.

Hyperlink

Related terms

ISO/IEC 11179
Object
Object class

Provider load

The effort, in terms of time and cost, required for respondents to provide satisfactory answers to a survey.

Source

Australian Bureau of Statistics, Service Industries Statistics, "Glossary of Terms"; unpublished on paper

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.

Hyperlink

http://www.abs.gov.au/CA25670D007E9EA1/0/DB35F160E9383A1FCA256B650006C3D0?Open&Highlight=0,Glossary

Related terms

Provision Agreement

Arrangement within which the provider supplies data or metadata.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
The agreement may define the scope of the data or metadata that can be provided

Hyperlink
http://www.sdmx.org/

Related terms
Data source

Public disclosure

The act of making information or data readily accessible and available to all interested individuals and institutions. 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

Context

Hyperlink
http://www.imf.org/external/np/mae/mft/sup/part1.htm#appendix_III

Related terms
Data dissemination

Punctuality

Punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar or previously agreed among partners.

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

Context
In SDMX, "Timeliness and Punctuality" is a single entity. While timeliness refers to the lapse of time between the end of a reference period (or a reference date) and dissemination of the data, punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar or previously agreed among partners.

Hyperlink

Related terms
Quality
SDMX
Timeliness

Qualitative data

Data describing the attributes or properties that an object possesses. 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

Context

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

Related terms
Quantitative data

Quality

The totality of features and characteristics of a product or service that bear on its ability to satisfy stated or implied needs.

Source
ISO 8402:1994 Quality management and quality assurance Vocabulary, withdrawn and revised by ISO 9000:2000 Quality management systems -- Fundamentals and vocabulary, March 2004

Context
Several statistical organisations have developed complementary definitions of quality, outlining the various dimensions of quality (e.g. accuracy, timeliness, etc) and the quality of statistical processes.
In SDMX, "Quality Management" refers to processes in place to focus on quality, to monitor the quality of the statistical programs, to deal with quality considerations in planning the statistical programs. It also includes how well the resources meet the requirements of the program, and measures to ensure efficient use of resources (staffing, facilities, computing resources, and financing of statistical programs).

Hyperlink

Related terms
Quality (Eurostat context)
Quality (IMF context)
Quality (OECD context)
Quality differences
SDMX

Quality (Eurostat context)

Eurostat defines the quality of statistics with reference to six criteria:
1. Relevance: an inquiry is relevant if it meets users' needs. The identification of users and their expectations is therefore necessary. In the European context, domains for which statistics are available should reflect the needs and priorities expressed by the users of the European Statistical System (completeness).
2. Accuracy: accuracy is defined as the closeness between the estimated value and the (unknown) true value.
3. Timeliness and punctuality in disseminating results: most users want up-to-date figures which are published frequently and on time at pre-established dates.
4. Accessibility and clarity of the information: statistical data have most value when they are easily accessible by users, are available in the forms users desire and are adequately documented.
5. Comparability: statistics for a given characteristic have the greatest usefulness when they enable reliable comparisons of values taken by the characteristic across space and time. The comparability component stresses the comparison of the same statistics between countries in order to evaluate the meaning of aggregated statistics at the European level.
6. Coherence: when originating from a single source, statistics are coherent in that elementary concepts can be combined reliably in more complex ways. When originating from different sources, and in particular from statistical surveys of different frequencies, statistics are coherent

in so far as they are based on common definitions, classifications and methodological standards.

Source

Eurostat, "Assessment of quality in statistics - Definition of Quality in Statistics", Working Group, Luxembourg, October 2003

Context

Hyperlink

Related terms

Accessibility
Accuracy
Clarity
Coherence
Comparability
Completeness
Punctuality
Quality
Quality (IMF context)
Quality (OECD context)
Relevance
Timeliness

Quality (IMF context)

The dimensions of the IMF definition of "data quality" are:
- integrity;
- methodological soundness;
- accuracy and reliability;
- serviceability;
- accessibility.

There are a number of prerequisites for quality. These comprise:
- legal and institutional environment;
- resources;
- quality awareness.

Source

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

Context

Under the SDDS, "quality" is the fourth of the four dimensions that comprise the standard (i.e., data, access, integrity, and quality) for which evidence of a subscribing member's observance can be judged through monitorable proxies (the dissemination of documentation on the methodology and sources used and the dissemination of data that supports statistical cross-checks)

Hyperlink

Related terms

Accessibility
Accuracy
Consistency
Integrity
Methodological soundness
Prerequisites of quality
Quality
Quality (Eurostat context)
Quality (OECD context)
Reliability
Serviceability
Special Data Dissemination Standard (SDDS)

Quality (OECD context)

Quality is viewed as a multi-faceted concept. The quality characteristics of most importance depend on user perspectives, needs and priorities, which vary across groups of users. Given the work already done in the area of quality by several organisations, notably, Eurostat, IMF and Statistics Canada, the OECD was able to draw on their work and adapt it to the OECD. Thus quality is viewed in terms of seven dimensions, namely:

- relevance
- accuracy
- credibility
- timeliness and punctuality
- accessibility
- interpretability
- coherence.

Source

Organisation for Economic Co-operation and Development (OECD), "Quality Framework for OECD Statistics", Paris, June 2002

Context

Hyperlink

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

Related terms

Quality
Quality (Eurostat context)
Quality (IMF context)

Quality control survey

A replicated survey carried out on a small scale by very experienced staff in order to obtain some "zero-default" results with which the actual results of the survey can be compared.

Source

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

Context

Hyperlink

Related terms

Quality differences

Differences in the various dimensions of data quality promulgated by international organisations and national agencies. 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

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Quality

Quality index

One-dimension synthetical information on quality, possibly calculated as a weighted mean of all available quality indicators.

Source

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

Context

Hyperlink

Related terms

Quantitative data

Data expressing a certain quantity, amount or range. Usually, there are measurement units associated with the data, e.g. meters, in the case of the height of a person. It makes sense to set boundary limits to such data, and it is also meaningful to apply arithmetic operations to the 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

Context

Hyperlink

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

Related terms

Flag

Qualitative data

Questionnaire

A group or sequence of questions designed to elicit information upon a subject, or sequence of subjects, from an informant.

Source

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

Context

Hyperlink

Related terms

Questionnaire design

Schedule

Survey

Questionnaire design

The design (text, order, and conditions for skipping) of the questions used to obtain the data needed for the survey.

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

Context

Hyperlink

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

Related terms

Questionnaire

Survey design

Ratio estimation

Ratio estimation involves the use of known population totals for auxiliary variables to improve the weighting from sample values to population estimates. It operates by comparing the survey sample estimate for an auxiliary variable with the known population total for the same variable on the frame. The ratio of the sample estimate of the auxiliary variable to its population total on the frame is used to adjust the sample estimate for the variable of interest.

Source

Australian Bureau of Statistics, Statistical Concepts Library, "Labour Statistics: Concepts, Sources and Methods", Chapter 16 - Overview of Survey Methods, Canberra, 2001

Context

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 variable and the variable of interest.

Hyperlink

http://www.abs.gov.au/AUSSTATS/abs%40.nsf/7884593a92027766ca2568b5007b8617/93a0165bdf598509ca256aa000036c90!OpenDocument

Related terms

Estimation

Weight

Recommended use of data

The recommended use(s) of statistical data refers to text that is intended to provide users with explicit information on the appropriate use(s) of the statistics within the limitations imposed by the definition or main concepts, scope and coverage, collection methodology, etc.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Record check

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 (for example, comparison of ages as reported in censuses with information on ages 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 Paper 4, Office of Federal Statistical Policy Standards, 1978

Context

Hyperlink

Related terms

Record-keeping error

An error which arises from inaccuracy in the records used for responses.

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 Paper 4, Office of Federal Statistical Policy Standards, 1978

Context

Hyperlink

Related terms

Recording of transactions

The recording of transactions pertains to a broad range of processes and standards employed in calculating statistical aggregates. The conventions include types of valuation, prices, conversion rates, the accounting basis, units of measurement used in data collection, etc.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Accounting basis

Special Data Dissemination Standard (SDDS)

Reference document

A document that provides pertinent details for consultation about a subject.

Source

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

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).

Hyperlink

Related terms
ISO/IEC 11179

Reference metadata

Metadata describing the contents and the quality of the statistical data.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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).

Hyperlink
http://www.sdmx.org/

Related terms
Common Metadata Concepts
Metadata Structure Definition
Statistical metadata
Structural metadata

Reference period

The time period to which a variable refers.

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

Context
Statistical variables refer to specific times, which may be limited to a reference time point (e.g. a specific day) or a period (e.g. a month, calendar year or fiscal year).
(Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003)

Hyperlink

Related terms
Accounting basis
Base period
Compilation practices
Period

Reference time

See "Reference period"

Refusal rate

In the sampling of human populations, the proportion of individuals who, though successfully contacted, refuse to give the information sought. 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.

Source

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

Context

Hyperlink

Related terms

Item response rate
Non-response
Non-response rate
Response rate

Register

A set of files (paper, electronic, or a combination) containing the assigned data elements and the associated information.

Source

ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, 1999

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).

Hyperlink

Related terms

ISO/IEC 11179
Registration authority

Registrar

A representative of a Registration authority.

Source

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

Context

Registrar contact is the contact information associated with a Registrar.
Registrar identifier is an identifier for the Registrar.

Hyperlink

Related terms

ISO/IEC 11179
Registration authority

Registration

The relationship between an administered item and the registration authority.

Source

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

Context

Hyperlink

Related terms

Administered item
ISO/IEC 11179
Registration authority
Registry

Registration authority

Organization responsible for maintaining a register.

Source

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

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.

Hyperlink

Related terms

ISO/IEC 11179
Permissible value
Register
Registrar
Registration

Registry item

Metadata item recorded in a Metadata Registry.

Source

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

Context

Hyperlink

Related terms

Administered item
ISO/IEC 11179
Metadata item
Metadata registry

Registry metamodel

A metamodel specifying a metadata registry.

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

Context

Hyperlink

Related terms
ISO/IEC 11179
Metadata registry
Metamodel

Related data reference

A reference between a data element and any related data.

Source
ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, 1999

Context

Hyperlink

Related terms
Data element
ISO/IEC 11179

Related metadata reference

A reference from one metadata item to another.

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

Context
A Registration Authority could choose to use a Reference Document, an administrative note or an explanatory comment to record a related metadata reference.

Hyperlink

Related terms
ISO/IEC 11179
Metadata item

Relationship

A connection among model elements. [ISO/IEC 19501-1:2001,2.5.2.36]

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

Context
In ISO/IEC International Standard 11179-3 "Information technology - Metadata registries-Part 3: Registry metamodel and basic attributes", February 2003 a relationship is either an association or a generalization. [ISO/IEC 19501-1:2001, 2.5.2.36]
Association is a semantic relationship between two classes. [ISO/IEC 19501-1:2001, 2.5.2.3]
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. [ISO/IEC 19501-1:2001, 2.5.2.24]
A relationship between a Data element example and its Data element is called an exemplification.

Hyperlink

Related terms
ISO/IEC 11179

Relative Standard error

See "Coefficient of variation"

Release calendar

A statement on the schedule of release of data in terms of periodicity and timeliness

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
An advance release calendar provides a general statement on the schedule of 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. Such information may be provided for statistical releases in the coming week, month, quarter or year.
In SDMX, "Release Calendar" describes the policy regarding the release of statistics according to a preannounced schedule and its availability. It also contains the release calendar information.
Advance release calendar information is one of the requirements of the Special Data Dissemination Standards (SDDS). Such information is disseminated on the Internet on the IMF`s Data Standards Bulletin Board (DSBB) or on national websites

Hyperlink
http://www.sdmx.org/

Related terms
Periodicity
SDMX
Simultaneous release
Timeliness

Relevance

The degree to which statistical information meets the real needs of clients.

Source
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 64

Context
In SDMX, "Relevance" refers to the processes for monitoring the relevance and practical utility of existing statistics in meeting users' needs and how these processes inform the development of statistical programs.
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 weigh and balance the conflicting needs of current and potential users to produce a program that goes as far as possible in satisfying the most important needs within given resource constraints (Statistics Canada, "Statistics Canada Quality Guidelines). 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. The uses and users of a given dataset may change over time, and new needs may arise that require new data; thus, the best processes have a dynamic nature (International Monetary Found, "Data Quality Assessment Framework (DQAF) Glossary").

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

Related terms
Quality
SDMX
Serviceability

Reliability

Closeness of the initial estimated value(s) to the subsequent estimated value(s).

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

Context
The third element of the IMF definition of quality is "accuracy and reliability".

Hyperlink

Related terms
Accuracy
Estimate
Quality

Reporting unit

A reporting unit is a unit that supplies the data for a given survey instance.

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

Context
When, for a specific survey, the book keeping office completes questionnaires for each of the locations of a business, these locations are the reporting units (Statistics Netherlands, "Reference manual on Design and Implementation of Business Surveys", March 1995, page 16).

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

Related terms
Survey

Respondent burden

See "Provider load"

Respondent load

See "Provider load"

Response errors

Errors arising from the interviewing process. Such errors can result from a number of circumstances, such as the following:
- inadequate concepts or questions;
- inadequate training;
- interviewer failures;
- respondent failures.

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

Context

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)

Hyperlink

Related terms

Response rate

The number of respondents who complete a questionnaire compared to the number assigned, usually expressed as a percentage. The response rate can also apply to individual questions.

Source

Australian Government Initiative, Statistical Clearing House, "Glossary"

Context

Hyperlink

http://www.sch.abs.gov.au/SCH/A1610103.NSF/Glossary?OpenView

Related terms

Item response rate
Non-response rate
Refusal rate

Responsible organization

The organization or unit within an organization that is responsible for the contents of the mandatory attributes by which the data element is specified.

Source

ISO/IEC FCD 11179-6 Information technology - Metadata registries - Part 6: Registration, January 2004

Context

Hyperlink

Related terms

ISO/IEC 11179
Organisation

Revision policy

A policy or set of policies, aimed at ensuring the transparency of disseminated data whereby preliminary data are compiled that are later revised when more and better source data become available.

Source

International Monetary Fund (IMF)," Quarterly National Accounts Manual", Washington D.C., 2001

Context

In SDMX, "Revision Policy and Practice" describes the data revision policy, the policy and practice for identifying the revision status of available data, as well as the availability of revision studies and analyses.

Providing users with documentation regarding the source data used and the way they are adjusted gives compilers with the possibility to incorporate new and more accurate information into estimates, thus improving their accuracy without introducing breaks in the time series.

Data may also be subject to 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. Whether or not such changes constitute an actual "revision" or the compilation of a "new" series is a matter of judgment on the part of the statistical agency.

Under the requirements of the Special Data Dissemination Standard (SDDS), an organisation's revision policy for specific statistics is disseminated on the Internet on the IMF`s Dissemination Standards Bulletin Board (DSBB).

Hyperlink

http://www.imf.org/external/pubs/ft/qna/2000/textbook/

Related terms

Adjustment Methods
Compilation practices
Data source
Data status (upon release)
Integrity
Internal access
Ministerial commentary
SDMX
Time series breaks

Sample

A subset of a frame where elements are selected based on a randomised process with a known probability of selection.

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

Context

Hyperlink

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

Related terms

Co-ordination of samples
Non-response rate
Probability sample
Sample design
Sample size
Sample survey
Sampling
Sampling fraction
Sampling technique
Sampling unit
Schedule
Stratification

Sample design

The sample design provides information on the target and final sample sizes, strata definitions and the sample selection methodology.

Source

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

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)

Hyperlink

Related terms

Sample
Survey design

Sample size

The number of sampling units which are to be included in the sample. 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

Context

Hyperlink

Related terms

Sample
Sampling
Sampling fraction
Sampling unit

Sample survey

A survey which is carried out using a sampling method, i.e. in which a portion only, 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

Context

Hyperlink

Related terms

Sample
Survey

Sampling

The process of selecting a number of cases from all the cases in a particular group or universe.

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

Context

Hyperlink

Related terms

Area sampling
Sample
Sample size

Sampling error

That part of the difference between a population value and an estimate, derived from a random sample, which is due to the fact that only a sample of values is observed; as distinct from errors due to imperfect selection, bias in response or estimation, errors of observation and recording, etc. 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

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

Context

Hyperlink

Related terms

Non-sampling error

Sampling fraction

The ratio of the sample size to the population size.

Source

Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 23

Context

Hyperlink

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

Related terms

Sample
Sample size

Sampling frame

See Frame

Sampling technique

The name or other identification of the specific process by which the entities of the sample have been selected.

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

Context

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

Related terms
Sample

Sampling unit

One of the units into which an aggregate is divided for the purpose of sampling, each unit being regarded as individual and indivisible when the selection is made.

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

Context
The definition of unit may be made on some natural basis, e.g., household, persons, units of product, tickets, etc., or upon some arbitrary basis, e.g., areas defined by grid co-ordinates on a map.

Hyperlink

Related terms
Sample
Sample size

Schedule

In the theory of sample surveys, schedule is synonymous with questionnaire.

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

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.

Hyperlink

Related terms
Questionnaire
Sample
Survey

Scope

The coverage or sphere of what is to be observed. It is the total membership or population of a defined set of people, object or events.

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

Context

In SDMX, "Scope/Coverage" describes the coverage of the statistics and how consistent this is with internationally accepted standards, guidelines, or good practices. The scope/coverage includes a description of target population, and geographic, sector, institutional, item, population, product, and other coverage.

Hyperlink

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

Related terms

Coverage
Out-of-scope units
SDMX
Statistical population

SDMX-EDI

EDIFACT format for exchange of SDMX-structured data and metadata.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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. (source, p. 10)

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.

Hyperlink

http://www.sdmx.org/

Related terms

EDIFACT
SDMX
SDMX-ML

SDMX-ML

XML format for the exchange of SDMX-structured data and metadata.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink

http://www.sdmx.org/

Related terms
SDMX
SDMX-EDI

SDMX Registry

An application which stores metadata for querying, and which can be used by any other application in the network with sufficient access privileges.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
The SDMX model of statistical exchange is based on a set of registry services which are not concerned with the storage of data or reference metadata, under the assumption that data and metadata live on the sites of data providers.
A registry can be understood as an index of data or metadata repositories of all the data providers within a statistical community, distributed across the Internet or similar networks. The registry services concern themselves with providing visibility to the data and reference metadata, and information needed for the access.
The SDMX registry provides the following services:
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.

Hyperlink
http://www.sdmx.org/

Related terms
Reference metadata
Structural metadata
Provider
Metadata registry
Registration
SDMX

Seasonal adjustment

A statistical technique to remove the effects of seasonal calendar influences operating on a series. 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).

Source
Australian Bureau of Statistics, "An Analytical Framework for Price Indexes in Australia: Glossary and References", Canberra, 1997

Context
Series are adjusted for seasonal variations and in some cases for calendar working days variations. When available, seasonally adjusted data are taken directly from national statistical sources; otherwise, the method used for de-seasonalisation is the standard X-11 ARIMA which was developed by the US Bureau of Census and incorporates general smoothing techniques and spectral analyses. (Further details may be found in Technical Paper No. 15 of the Bureau of the Census.). Where appropriate, series are also corrected for calendar variations (e.g. workdays per month) and constrained for annual coherency. (Organisation for Economic Co-

operation and Development (OECD), "The OECD Economic Outlook: Sources and Methods", available at www.oecd.org/eco/sources-and-methods)
Seasonal adjustment is normally done using off-the-shelf programs - most commonly worldwide by one of the programs in the X-11 family. Other programs in common use include the TRAMO-SEATS package developed by Bank of Spain and promoted by Eurostat and the German BV4 program (International Monetary Fund (IMF)," Quarterly National Accounts Manual", Washington D.C., 2001, para. 8.13).
Under the SDDS this 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.

Hyperlink
http://www.abs.gov.au/ausstats/abs%40.nsf/66f306f503e529a5ca25697e0017661f/ff4de83064a2e425ca25697e0018fd44!OpenDocument

Related terms
Adjustment Methods
Compilation practices
Not seasonally adjusted series
Special Data Dissemination Standard (SDDS)

Secondary source of statistical data

The organisation or individual other than those responsible for the collection and aggregation of data from their initial source. 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
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink
http://www.sdmx.org/

Related terms
Data source
Primary data
Primary source of statistical data

Semantics

The branch of linguistic science which deals with the meaning of words.

Source
ISO/IEC CD 11179-5 "Information technology - Metadata registries (MDR) - Part 5: Naming and identification principles", January 2003

Context

Hyperlink

Related terms
ISO/IEC 11179
Syntax

Serviceability

Serviceability refers to the practical aspects of how well the available data meet users' needs.

Source

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

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.

Hyperlink

Related terms

Consistency
Quality (IMF context)
Relevance

Sibling group

A set of time series whose keys differ only in the value taken by the frequency dimension.

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

Context

Within an ETS (exchanged time series), a sibling group is uniquely identified by a data set identifier combined with the sibling group key.

Hyperlink

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

Related terms

Data set
GESMES/TS
Key (time series or sibling group)
Time series

Simultaneous release

The dissemination of statistical data to all interested parties at the same time.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "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.
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.

Hyperlink

http://www.sdmx.org/

Related terms

Accessibility
Release calendar
SDMXSDMX

Source

See "Data source"

Source data

Data collected on a regular basis (by survey from respondents, or from administrative sources) by survey statisticians in the national statistical system to be edited, imputed, aggregated and/or used in the compilation and production of official statistics.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

IN SDMX, "Source Data" refers to a description of the data collection programs and their adequacy for the production of statistics, including meeting the requirements for methodological frameworks, scope, classifications systems, and basis for recording.
In the context of the SDDS, "nature of the basic data" includes whether data are compiled from administrative records (e.g., monetary and government data), surveys, censuses, or any combination of these. It should also briefly describe the means of data collection. In cases where sampling techniques are used, the sampling procedures should be briefly described.

Hyperlink

http://www.sdmx.org/

Related terms

Classification system
Data collection
Primary data
Scope
SDMX
Special Data Dissemination Standard (SDDS)

Special Data Dissemination Standard (SDDS)

"Special Data Dissemination Standard". The 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 early April 1996.

Source

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

Context

Hyperlink

Related terms

Accounting basis
Adjustment Methods
Aggregation
Analytical framework
Base period
Basic statistical data
Comparability
Compilation practices
Computation of lowest level indices
Coverage
Data
Data collection
Data Dissemination Standards

Data item
Data reconciliation
Data status (upon release)
Estimation
General Data Dissemination System (GDDS)
Integrity
Internal access
Metadata dimension (SDDS)
Ministerial commentary
Primary data
Quality (IMF context)
Recording of transactions
Seasonal adjustment
Simultaneous release
Standard Classification
Types of prices
Valuation
Verification

Special language

A language used in a subject field and characterized by the use of specific linguistic means of expression.

Source

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

Context

The specific linguistic means of expression always include subject-specific terminology and phraseology and also may cover stylistic or syntactic features. [ISO 1087-1:2000, 3.1.3]

Hyperlink

Related terms

ISO/IEC 11179
Language
Terminology

Standard Classification

Classifications that follow prescribed rules and are generally recommended and accepted. They aim to ensure that information is classified consistently regardless of the collection, source, point of time etc.

Source

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

Context

In the international context, standard classifications include ISIC Rev. 3, ISCO, CPC, NACE Rev 1, 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. Many of the international and national standard classifications are listed in the RAMON database of classifications, available at http://europa.eu.int/comm/eurostat/ramon

In the SDDS context, the criteria used to classify major economic transactions, industrial activities, commodities and services, consumption, data components, international transactions or the sectorization of accounts for collection and/or dissemination and whether these criteria are consistent with relevant international or regional standard classifications and/or guidelines and at what level. (e.g. revenue consists of tax and nontax revenue classified according to the GFSM; Expenditure by function is classified according to the SNA classification of the functions

of government (COFOG), consumption of products classified according to CPC and aggregation by COICOP or other standard system, etc.

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

Related terms
Classification
Special Data Dissemination Standard (SDDS)

Standard error

The positive square root of the variance of the sampling distribution of a statistic.

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

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.

Hyperlink

Statistical concept

A statistical characteristic of a time series or an observation.

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

Context
In SDMX, "Concepts and Definitions" refer to the internationally accepted statistical standards, guidelines, or good 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
Each 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).

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

Related terms
Attribute
Characteristic
Code list
Concept
Dimension
GESMES/TS
International statistical standard
Key family
Key structure
Observation
SDMX
Structural definition

Statistical Data and Metadata Exchange (SDMX)

A task force sponsored by BIS, ECB, Eurostat, IMF, OECD, UN and World Bank to address standardization of the exchange of statistical information.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Statistical error

The (unknown) difference between the retained value and the true value.

Source

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

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.

Hyperlink

Related terms

Accuracy

Statistical indicator

A data element that represents statistical data for a specified time, place, and other characteristics.

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

Context

Hyperlink

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

Related terms

Statistical macrodata

An observation data gained by a purposeful aggregation of statistical microdata conforming to statistical methodology.

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

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).

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

Related terms
Statistical microdata

Statistical measure

A summary (means, mode, total, index, etc.) of the individual quantitative variable values for the statistical units in a specific group (study domains).

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

Context

Hyperlink

Related terms

Statistical message

A message carrying statistical data.

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

Context
In the context of GESMES/TS, 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 (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).

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

Related terms
Data exchange
GESMES

Statistical metadata

Data about statistical data.

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, 1995

Context
Metadata comprises data and other documentation that describes 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)
Metadata provide information on data and about processes of producing and using data. Metadata describe statistical data and - to some extent - processes and tools involved in the production and usage of statistical data (UNECE, "Guidelines for the Modeling of Statistical Data and Metadata", 1995).

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

Statistical metadata repository

A logically central statistical metadata repository that allows for the query, editing, and managing of metadata. Such a system provides a mechanism for looking up information about statistical products as well as their design, development, and analysis.

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

Context
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).

Hyperlink

Related terms
Data series flow

Statistical metadata system

A data processing system that uses, stores and produces statistical metadata.

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

Context

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

Related terms
Metadata
Metadata layer
Statistical metadata
Statistical metainformation

Statistical metainformation

Knowledge of objects described by statistical metadata

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

Context

Hyperlink

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

Related terms

Metadata layer
Statistical metadata
Statistical metadata system
Statistical metainformation system

Statistical metainformation system

A system is a system which uses and produces statistical metadata, informing about statistical data, and which fulfils its tasks by means of functions like "statistical metadata collection", "statistical metadata processing", "statistical metadata storage", and "statistical metadata dissemination".

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, 1995

Context

An alternative definition of a statistical metainformation system is an information system for which the object is the statistical information system (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,
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 Commission (UNECE/UNSC), "Guidelines for the Modelling of Statistical Data and Metadata", Conference of European Statisticians Methodological Material, Geneva, 1995, p. 4).

Hyperlink

Related terms

Statistical metainformation

Statistical methodology

Theory and methods of data collection, processing and analysis.

Source

Statistics Canada, Integrated Metadata Base, "Glossary", unpublished on paper

Context

Hyperlink

http://www.statcan.ca/english/edu/power/toc/contents.htm

Related terms

Methodology

Statistical microdata

An observation data collected on an individual object - statistical unit.

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

Context

Microdata are 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).

Hyperlink

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

Related terms

Statistical macrodata

Statistical population

The total membership or population or "universe" of a defined class of people, objects or events.

Source

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

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 can be obtained in the 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 records 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).

Hyperlink

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

Related terms

Census
Coverage errors
Cut-off threshold
Scope
Stratification
True value

Statistical processing

The processes for manipulating or classifying statistical data into various categories with the object of producing statistics.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

In SDMX, "Statistical Processing" refers to a description of the data compilation and other statistical procedures to deal with intermediate data and statistical outputs (e.g., data adjustments and transformation, and statistical analysis). The items covered include, inter alia, weighting schemes, methods for imputing missing values or source data, statistical adjustment,

and balancing/cross-checking techniques and relevant characteristics of the specific approach/approaches applied.

Hyperlink
http://www.sdmx.org/

Related terms
Adjustment Methods
Aggregation
Compilation practices
Computation of lowest level indices
Consolidation
Data collection
Data processing
Data reconciliation
Disaggregation
Estimation
Index number
Reference period
Revision policy
SDMX
Seasonal adjustment
Special Data Dissemination Standard (SDDS)
Verification

Statistical production

The activity that is carried out within statistical information system and aimed at producing of statistics.

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

Context

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

Related terms

Statistical standard

A comprehensive set of guidelines for surveys and administrative sources collecting information on a particular topic.

Source
Statistics New Zealand," Classifications and Standards"; unpublished on paper

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.

Hyperlink
http://www.stats.govt.nz/domino/external/web/prod_serv.nsf/092edeb76ed5aa6bcc256afe0081d84e/35b11e7066c13db1cc256ca5006f44e4?OpenDocument

Related terms
International statistical standard

Statistical subject-matter domain

A statistical activity that has common characteristics with respect to concepts and methodologies for data collection, manipulation and transformation.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Characteristic
Concept
Domain
Domain groups
Methodology

Statistical unit

An object of statistical survey and the bearer of statistical characteristics. The statistical unit is the basic unit of statistical observation within a statistical survey.

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

Context
Statistical units are the entities for which information is sought and for which statistics are ultimately compiled. These units can, in turn, be divided into observation units and analytical units. The statistical units in ISIC Rev. 3 comprise the:
- enterprise;
- enterprise group;
- kind-of-activity unit (KAU);
- local unit;
- establishment;
- homogeneous unit of production.

(Statistical Office of the United Nations, "International Standard Industrial Classification of all Economic Activities, Third Revision", Statistical Papers Series M No. 4, Rev. 3, United Nations, New York, 1990, para. 63, 76).
Statistical units are defined on the basis of three criteria:
- Legal, accounting or organisational criteria;
- Geographical criteria;
- Activity criteria.

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

Related terms
Analytical unit
Classification
Observation unit

Stewardship

Stewardship (of metadata) is the responsibility for the maintenance of Administration Records applicable to one or more Administered Items.

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

Context
The responsibility for the registration of metadata may be different from the responsibility for stewardship of metadata.
Stewardship contact is the contact information associated with a Stewardship.

Hyperlink

Related terms
Administered item
Contact
ISO/IEC 11179
Organisation
Submission

Stratification

Stratification consists of dividing the population into subsets (called strata) within each of which an independent sample is selected.

Source
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 19982003, page 21

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).

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

Related terms
Sample
Statistical population

Structural definition

Statistical concepts, key families and code lists defined by a centre institution, (usually for the exchange of statistical information with its partners).

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

Context

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

Related terms
Code list
Concept
GESMES/TS
Key family
Maintenance Agency
Statistical concept
Structural metadata

Structural metadata

Metadata that act as identifiers and descriptors of the data.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

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.

Hyperlink
http://www.sdmx.org/

Related terms
Common Metadata Concepts
GESMES/TS
Key family
Metadata Structure Definition
Reference metadata
Statistical metadata
Structural definition

Structure

It provides the means for identifying relationships, usually hierarchical, between categories.

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

Context
A hierarchical classification is based on a tree structure where each set of its detailed categories are subsets of categories at the level about the one in which they contained.

Hyperlink

Related terms
Category
Hierarchy
Data structure definition
Metadata structure definition

Study domain

A major segment of the population for which separate statistics are needed.
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.

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

Context

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.

Hyperlink

Related terms

Statistical subject-matter domain

Submission

Submission (of an administered Item) is the relationship of an Administered item, a Contact, and an Organization involved in a submission of metadata.

Source

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

Context

Submission contact is the contact information associated with a submission.

Hyperlink

Related terms

Administered item
Contact
ISO/IEC 11179
Stewardship
Submitting organization

Submitting organization

The submitting organization is responsible for requesting that a new metadata item be registered in the registry.

Source

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

Context

Hyperlink

Related terms

ISO/IEC 11179

Metadata registry
Submission

Supplementary data

See "Data dissemination"

Survey

A 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.

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

Context

The term survey covers any activity that collects or acquires statistical data. Included are censuses, sample surveys, the collection of data from administrative records and derived statistical activities. (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)

Hyperlink

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

Related terms

Cut-off survey
Data collection
Non-response
Processing error
Questionnaire
Reporting unit
Sample survey
Schedule
Survey design

Survey data collection

An activity of the survey life cycle for gathering data from respondents and recording it for further processing.

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

Context

Hyperlink

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

Related terms

Data collection
Data source

Survey design

Survey design covers the delineation of all aspects of a survey from the establishment of a need for data to the production of final outputs (the microdata file, statistical series, and analysis).

Source
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 8

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).

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

Related terms
Questionnaire design
Sample design
Survey

Syntax

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.

Source
ISO/IEC CD 11179-5 "Information technology - Metadata registries (MDR) - Part 5: Naming and identification principles", January 2003

Context

Hyperlink

Related terms
ISO/IEC 11179
Semantics

Target population

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).

Source
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003, page 17

Context

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

Related terms
Cut-off threshold
Under-coverage

Taxonomy

The classification according to presumed natural relationships among types and their subtypes.

Source
ISO/IEC International Standard 11179, Part 1, Framework for the specification and standardization of data elements, 1999

Context
Within SDMX, a “reporting taxonomy” is a scheme which defines the composition structure of a data report where each component is described by an independent data flow definition

Hyperlink

Related terms
Classification
ISO/IEC 11179
Ontology

Term

A designation of a defined concept in a special language by a linguistic expression.

Source
ISO International Standard 1087-1:2000 Terminology work -- Vocabulary -- Part 1: Theory and application, November 2004

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).

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

Related terms
Terminology
Thesaurus

Terminological entry

An entry containing information on terminological units for a specific Administered item within a Context (subject field).

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

Context

Hyperlink

Related terms
ISO/IEC 11179

Terminological system

A concept system with designations for each concept.

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

Context

Hyperlink

Related terms
Concept
ISO/IEC 11179

Terminology

A set of terms.

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

Context

Hyperlink

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

Related terms

Special language
Term

Thesaurus

A tool that associates related terms, and thesaurus terms assist in locating an existing data element.

Source

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

Context

Hyperlink

Related terms

Data element
ISO/IEC 11179
Term

Time coverage

The length of time, e.g. years, for which data are collected.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Time of recording

The date the item was recorded in a dissemination medium. This may be the date the item was first recorded or the date an existing item was amended.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

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.

The time of recording for a transaction is governed by the principle of accrual accounting.

Hyperlink
http://www.sdmx.org/

Related terms
Accounting basis
Date of last change

Time Period

See "Period"

Time series

A set of ordered observations on a quantitative characteristic of an individual or collective phenomenon taken at different points of time.

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

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, within 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)

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

Related terms
Characteristic
Data Provider Series Key
Data set
Dimension
Key (time series or sibling group)
Observation
Pre-Break Value
Sibling group
Time series breaks
Trend

Time series breaks

Breaks occurred when there is a change in the standards for defining and observing a variable over time. Such changes 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.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
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.

Hyperlink
http://www.sdmx.org/

Related terms
Pre-Break Value
Revision policy
Time series

Timeliness

Speed of dissemination of the data - i.e., the lapse of time between the end of a reference period (or a reference date) and dissemination of the data.

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

Context
In SDMX, "Timeliness and Punctuality" is a single entity. Timeliness refers to the speed of dissemination of the data - i.e., the lapse of time between the end of a reference period (or a reference date) and dissemination of the data. It reflects many factors, including some that are related to institutional arrangements, such as the preparation of accompanying commentary and printing. Punctuality refers to the possible time lag existing between the actual delivery date of data and the target date when it should have been delivered, for instance, with reference to dates announced in some official release calendar or previously agreed among partners.

Hyperlink

Related terms
Data
Punctuality
Quality
Release calendar
SDMX

Transparency

See "Integrity"

Trend

A long-term movement in an ordered series, which may be regarded, together with the oscillation and random component, as generating the observed values.

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

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.

Hyperlink

Related terms
Time series
Trend estimates

Trend estimates

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.

Source

Australian Bureau of Statistics, "An Analytical Framework for Price Indexes in Australia: Glossary and References", Canberra, 1997

Context

Hyperlink

http://www.abs.gov.au/ausstats/abs%40.nsf/66f306f503e529a5ca25697e0017661f/ff4de83064a2e425ca25697e0018fd44!OpenDocument

Related terms

Estimate
Trend

True value

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.

Source

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

Context

Hyperlink

Related terms

Statistical population

Type of data collection

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

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Data collection

Under-coverage

Under-coverage results from the omission from the frame of units belonging to the target population.

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

Context

Hyperlink

Related terms
Frame
Over-coverage
Target population

Unit non-response

See "Non-response error"

Unit of measure

The actual unit in which the associated values are measured.

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

Context
The dimensionality of the associated conceptual domain must be appropriate for the specified unit of measure.
ISO 31-0:1982 specifies a system of physical measurement (the International System of Units, SI). Physical
measurement is only one type of measurement. Value measurement is another type of measurement. ISO/IEC International Standard 11179-3 allows the use of any appropriate system of measurement.
Attributes of a unit of measure:
Unit of measure name is the name of a unit of measure.
Unit of measure precision is the degree of specificity for a unit of measure. Expressed as a number of decimal places to be used in any associated data element values. To be used as a default if no precision is specified on the data element itself.

Hyperlink

Related terms
Conceptual domain
Dimension
Dimensionality
ISO/IEC 11179

Unit response rate

The ratio, expressed in percentage of the number of interviews to the number of eligible units in the sample.

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

Context
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).

Hyperlink

Unit value

Expenditures or value of production of an item is divided by the quantity.

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

Context

Hyperlink

Related terms

Unit value index

Unit value index

The percentage of an eligible sample for which information is obtained.

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 Paper 4, Office of Federal Statistical Policy Standards, 1978

Context

For an interview survey, the numerator of the formula is the number of interviews. 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. Weighted response rate is more useful. In addition to this, another non-response rate using the following weights may be computed: the sampling weights times any existing important variable from the sampling frame (turnover or size). The latter figure illustrates an impact of non-response.

Hyperlink

Related terms

Unit value

User needs (for statistics)

User needs refer to the data and metadata requirements of persons or organisations to meet a particular use or set of uses. Such needs may be specified in terms of the quality dimensions promulgated by international organisations or national agencies.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

User satisfaction survey

A statistical survey aiming to assess the satisfaction of users of statistics.

Source

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

Context

Hyperlink

Related terms

Validation

A continuous monitoring of the process of compilation and of the results of this process.

Source

Eurostat, "Handbook on Quarterly National Accounts", Luxembourg, 1999

Context

In SDMX, "Validation" describes methods and processes for routinely assessing source data – including censuses, sample surveys, and administrative records – and how the results of the assessments are monitored and made available to guide statistical processes. It also describes how intermediate results are validated against other information where applicable, how statistical discrepancies in intermediate data are assessed and investigated and how statistical discrepancies and other potential indicators or problems in statistical outputs are investigated. 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 includes the results of studies and analysis of revisions and how they are used to inform the statistical processes. In this process, two dimensions can be distinguished: (i) validation before publication of the figures and (ii) validation after publication.

Hyperlink

Related terms

SDMX

Valuation

Under the SDDS, valuation refers to such items as: 1) the exchange rates or conversion factors used to convert foreign-currency-denominated assets and liabilities into the national currency equivalent. 2) whether transactions are recorded at market prices, face or nominal value vs. issue or discounted prices; 3) adjustments to convert CIF imports to an FOB basis; 4) adjustments to align data with balance of payments concepts as well as the frequency of any revaluation (e.g., daily, monthly, quarterly, yearly).

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Accounting basis

Special Data Dissemination Standard (SDDS)

Value domain

A value domain is a set of permissible values.

Source

ISO/IEC FCD 11179-1 "Information technology - Metadata registers - Part 1: Framework", May 2003

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. "Enumerated value domain" is a value domain that is specified by a list of all its permissible values.

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.

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)

Hyperlink

Related terms

Attribute

ISO/IEC 11179

Permissible value

Permitted value

Value item

Value item

A representation of a value meaning in a specific value domain - the actual value.

Source

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

Context

Hyperlink

Related terms

ISO/IEC 11179

Value domain

Value meaning

Value meaning

The meaning or semantic content of a value.

Source

ISO/IEC FCD 11179-1 "Information technology - Metadata registers - Part 1: Framework", May 2003

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.

Hyperlink

Related terms
Conceptual domain
ISO/IEC 11179
Permissible value
Value item

Variable

A characteristic of a unit being observed that may assume more than one of a set of values to which a numerical measure or a category from a classification can be assigned (e.g. income, age, weight, etc. and "occupation", "industry", "disease" etc).

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

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).

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

Related terms
Characteristic
Data item
Observation

Verification

Principal methods to review, audit, or verify the accuracy of the disseminated data (e.g., internal review, statistical confidence tests, internal audit, audit by outside accountants, cross-checks with other macroeconomic accounts, etc.).

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context
Under the SDDS, this may entail-according to the data category under consideration-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.

Hyperlink
http://www.sdmx.org/

Related terms
Compilation practices
Special Data Dissemination Standard (SDDS)

Weight

The importance of an object in relation to a set of objects to which it belongs; 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.

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

Context
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).

Hyperlink

Related terms
Base weight
Non-response
Non-response error
Ratio estimation
Weight period

Weight period

The period that provides the weights for an index number.

Source
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

Related terms
Base period
Index number
Period
Weight

XML

eXtensible Mark-up Language

Year-to-date data

Data expressed in cumulative terms from the beginning of the year; sometimes referred to as cumulative data.

Source

Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary

Context

Hyperlink

http://www.sdmx.org/

Related terms

Sources

Australian Bureau of Statistics, "An Analytical Framework for Price Indexes in Australia: Glossary and References", Canberra, 1997
Australian Bureau of Statistics, Service Industries Statistics, "Glossary of Terms"; unpublished on paper
Australian Bureau of Statistics, Statistical Concepts Library, "Labour Statistics: Concepts, Sources and Methods", Chapter 16 - Overview of Survey Methods, Canberra, 2001
Australian Government Initiative, Statistical Clearing House , "Glossary"
Code of Good Practices on Transparency in Monetary and Financial Policies, Part 1- Introduction; approved by the IMF Executive Board on July 24, 2000
Dublin Core
Economic Commission for Europe of the United Nations (UNECE), "Glossary of Terms on Statistical Data Editing", Conference of European Statisticians Methodological material, Geneva, 2000
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
Eurostat
Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, May 2002
Eurostat, "Assessment of Quality in Statistics: Glossary", Working Group, Luxembourg, October 2003
Eurostat, "Assessment of quality in statistics - Definition of Quality in Statistics", Working Group, Luxembourg, October 2003
Eurostat, "European System of Accounts - ESA 1995", Office for Official Publications of the European Communities, Luxembourg, 1996, 2.12
Eurostat, "European System of Accounts - ESA 1995", Office for Official Publications of the European Communities, Luxembourg, 1996, 2.17-2.18
Eurostat, "Handbook on Quarterly National Accounts", Luxembourg, 1999
Federal Committee on Statistical Methodology, "Glossary of Nonsampling Error Terms: An Illustration of a Semantic Problem in Statistic", Statistical Policy Working Paper, December, 1978
GESMES/TS User Guide, Release 3.00, February, 2003; unpublished on paper
ISO 8402:1994 “Quality management and quality assurance – Vocabulary”, withdrawn and revised by ISO 9000:2000 “Quality management systems -- Fundamentals and vocabulary”, March 2004
ISO International Standard 1087-1:2000 “Terminology work -- Vocabulary -- Part 1: Theory and application”, November 2004
ISO International Standard 9735:1988 “Electronic data interchange for administration, commerce and transport (EDIFACT) – Application level syntax rule s”, September 1996
ISO/IEC 11179-3 "Information technology - Metadata registries - Part 3: Registry metamodel and basic attributes", February 2003
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
ISO/IEC CD 11179-5 "Information technology - Metadata registries (MDR) - Part 5 : Naming and identification principles", January 2003
ISO/IEC FCD 11179-1 "Information technology - Metadata registers - Part 1: Framework", May 2003
ISO/IEC FCD 11179-6 “Information technology - Metadata registries - Part 6: Registration”,

January 2004
ISO/IEC FDIS 11179-1 "Information technology - Metadata registries - Part 1: Framework", March 2004
ISO/IEC International Standard 11179, Part 1, “Framework for the specification and standardization of data elements”, 1999
ISO/IEC International Standard 11179-3 "Information technology - Metadata registries (MDR) - Part 3: Registry metamodel and basic attributes", February 2003
International Monetary Fund (IMF), "Data Quality Assessment Framework (DQAF) Glossary"
International Monetary Fund (IMF), "Guide to the Data Dissemination Standards, Module 1: The Special Data Dissemination Standard", Washington, May 199 6
International Monetary Fund (IMF), "Guide to the General Data Dissemination System", 2002
International Monetary Fund (IMF)," Quarterly National Accounts Manual", Washington D.C., 2001
International Monetary Fund, "Data Quality Assessment Framework - DQAF - Glossary", unpublished
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 Paper 4, Office of Federal Statistical Policy Standards, 1978
Madow, W., Nisselson, H., and Olkin, I., "Incomplete Data in Sample Surveys", Academic Press, New York, 1983
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 Expert Group, Paris
Marriott, F.H.C for the ISI, "A Dictionary of Statistical Terms", 5th edition, Longman Scientific & Technical , New York , 1990
McGraw-Hill Encyclopedia of Science and Technology
Neuchâtel Group, "Neuchâtel Terminology: Classification database object types and their attributes - Version 2", September 2002
OECD, IMF, ILO, Interstate Statistical Committee of the Commonwealth of Independent States, "Measuring the Non-Observed Economy: A Handbook", Second Draft, Annex 2, Glossary, Paris, 2002
Office for National Statistics (ONS), "National statistics, methods and quality report: Glossary of terms"; unpublished on paper
Organisation for Economic Co-operation and Development (OECD), "Main Economic Indicators", monthly
Organisation for Economic Co-operation and Development (OECD), "Quality Framework for OECD Statistics", Paris, June 2002
Paul P. Biemer, Robert M.Groves, Lars E. Lyberg, Nancy A.Mathiowetz, Seymour Sudman, "Measurement errors in survey", John Wiley &Sons, 1991
Sarndal C.E., Swensson B., Wretman J., "Model assisted survey sampling", Springer - Verlag, New York, 1992
Statistical Data and Metadata Exchange (SDMX) - BIS, ECB, Eurostat, IBRD, IMF, OECD and UNSD - Metadata Common Vocabulary
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
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
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
Statistical Office of the United Nations, "International Standard Industrial Classification of all Economic Activities, Third Revision", Statistical Papers Series M No. 4, Rev. 3, United Nations, New York, 1990, para. 63

Statistics Canada, "Statistics Canada Quality Guidelines", 3rd edition, October 1998.
Statistics Canada, "Statistics Canada Quality Guidelines", 4th edition, October 2003
Statistics Canada, Integrated Metadata Base, "Glossary", unpublished on paper
Statistics New Zealand," Classifications and Standards"; unpublished on paper
Status Report on the BIS-IMF-OECD-World Bank Joint External Debt Hub: Prepared by the SDMX Pilot Project Team, May 2005
The International Statistical Institute, "The Oxford Dictionary of Statistical Terms", edited by Yadolah Dodge, Oxford University Press, 2003
The Oxford Advanced Learner's Dictionary, Oxford University Press
U.S. Department of Labor, Bureau of Labor Statistics, "Glossary", unpublished on paper
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
United Nations Glossary of Classification Terms; prepared by the Expert Group on International Economic and Social Classifications, unpublished on paper
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
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, 1995
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
United States Bureau of Census, "Glossary of Selected Abbreviations and Acronyms"; unpublished on paper
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
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
United States Federal Committee on Statistical Methodology, "Statistical Policy Working Paper 15: Quality in Establishment Surveys", Washington D.C., July 1988, page 79
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

7507