Data quality assessment and metadata

Workshop on Travel and Tourism Statistics for the CARICOM countries
Roseau, Dominica, 14 – 17 May 2013

United Nations Statistics Division
Statistics of International Trade in Services Section

Outline

Quality

- Relevance of statistical concepts
- Accuracy
- Timeliness
- Accessibility and clarity of information
- Comparability of statistics
- Coherence
- Completeness/coverage
- Cost and burden

Metadata

Data quality

Available resources to collect, analyze and store international trade in services statistics will make an effect on the quality of the data.

Data quality

Several statistical organizations and countries have developed definitions of quality, outlining the various dimensions (aspects) of quality and quality measurement and have integrated them into **quality assessment frameworks**

Examples of quality assessment frameworks:

European Statistical System (ESS) focuses on the statistical outputs and defines quality with reference to six criteria

IMF Data Quality Assessment Framework (DQAF) – holistic view of data quality, including governance of statistical system

OECD Quality Measurement Framework – takes the user's side to approach quality – uses seven dimensions

Data quality

No unique indicator of data quality – several criteria are used, like:

- Relevance of statistical concepts
- Accuracy
- Timeliness
- Accessibility and clarity of information
- Comparability of statistics
- Coherence
- Completeness/coverage
- Cost and burden

Relevance

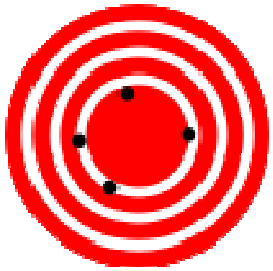
Relevance in statistics is assured when statistical concepts meet current and potential users' needs. Identification of the users and their expectations is a must.

Accuracy

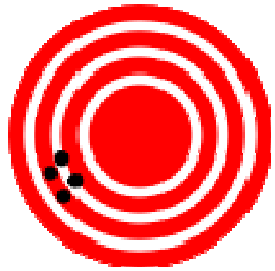
Accuracy is defined as the closeness between the computations or estimates and the (unknown) true population value.

Assessing the accuracy of an estimate involves analysing the total error associated with the estimate: bias (+/-) and standard deviation (when possible).

High **accuracy** but low **precision** (large sample error?)



High **precision** but low **accuracy** (biased estimate?)



Accuracy (cont.)

- **Sampling errors**: lack of accuracy due to observing only a sample instead of the whole population (quantifiable by the *standard error*)

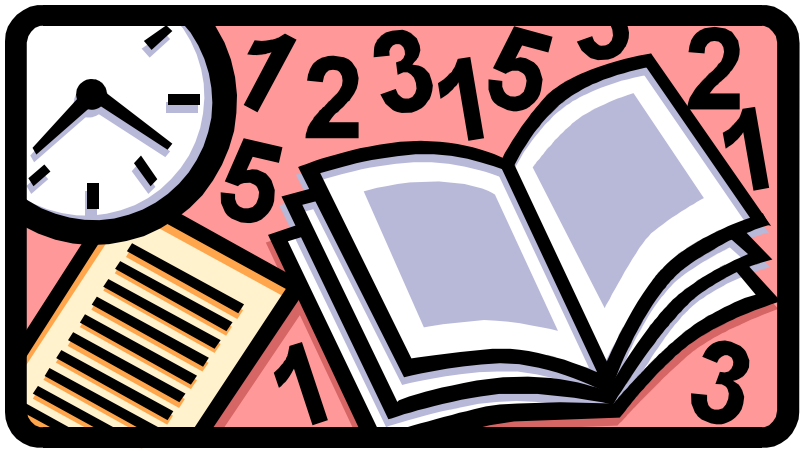
- **Non-sampling errors:**
  - Coverage errors (under- or over coverage)
  - Non-response errors (surveys)
  - Measurement errors
  - Processing errors
  - Model assumption errors

Timeliness

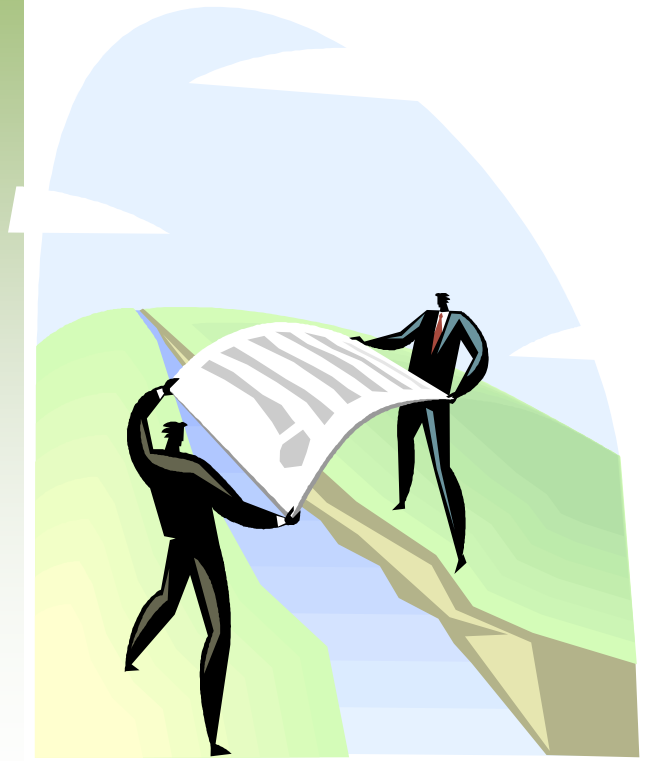
Users want the latest data that are published frequently and on time at pre-established dates.

Data

- ✓ Collection
- ✓ Editing
- ✓ Consolidation
- ✓ Dissemination



Accessibility and clarity of information



Statistical data are most valuable when they are:

- ✓ Easily accessible by users
- ✓ Available in the form users desire
- ✓ Adequately documented – accompanied by good **metadata**

Assistance in using and interpreting the statistics should also be forthcoming from the providers.

Comparability of statistics

Statistics for a given characteristic have the greatest usefulness when they enable reliable comparisons of values across geography and over time.

Providing comparable country data makes it possible for international organizations to publish regional and world totals.

Comparability of statistics (cont.)

For comparability the following are needed:

- ✓Common definitions
- ✓Common unit of measurement
- ✓Unified methodology
- ✓Timely submission of data to international organizations

Coherence

Coherence is the measure of the extent to which one set of statistical characteristics agrees with an other and can be used together (with each other) or as an alternative (to each other).

Completeness/coverage

The component of completeness reflects the extent to which the statistical system in place answers the users' needs and priorities by comparing all user demands with the availability of statistics.

Cost and burden

Although not measures of quality, they are positively correlated with quality.

Costs: office space, utility bills, staff-hours involved, funding of surveys, etc.

Response burden: simplest way to measure is the time spent by the respondent to provide information

A compromise between quality and cost and burden must be achieved

Metadata

- Statistical metadata facilitate sharing, querying and understanding of statistical data over the lifetime of the data. They also refer to any methodological descriptions on how data are collected and processed.

- Metadata is essential for the interpretation of statistical data.

Metadata

There is a bidirectional relationship between metadata and quality:

- ✓ Metadata describe the quality of statistics

- ✓ Metadata are themselves a quality component improving the availability and accessibility of statistical data

Metadata

As a minimum segmentation, the following two levels of metadata are recommended:

- Structural metadata presented as an integral part of the data tables – for example footnotes explaining the statistical output
- Reference metadata providing details on the content and quality of data – for example a description of data sources and statistical processes and estimations related to producing the statistics

Metadata

- Metadata provides a mechanism for comparing national practices in the compilation of statistics. This may help and encourage countries to implement international standards and to adopt the best practices.

- Better harmonization of approaches will also improve general quality of the data.

UNSD request for metadata:

Items requested:

- Contact info
- Agencies involved in data collection and processing
- Methodological framework followed
- Data sources
- Data dissemination
- Other information

UNSD metadata on SITS

UNSD Request for data on external trade in services. July 2011 ENGLISH

Contact information

Country:	CHINA	Contact person:	HU Hong	Tel:	00-86-10-68402093
Institution:	State Administration of Foreign Exchange	E-mail:	hu-hong@safe.gov.cn	Fax:	00-86-10-68402316

Metadata

Are other agencies involved in collecting and processing data on external trade in services in your country? *Please check all that apply*

The National Statistical Office:	NO
The Central/National Bank	YES
The Trade Ministry	NO
Other institutions (please specify):	the National Tourism Administration of The People's Republic of China,the Immigration Administration Department of the Ministry of Public Security ,General Administration of customs of the people's Republic of China

Which methodological framework is followed in your country? *Please check all that apply*

EBOPS - Manual on Statistics of International Trade in Services	NO
BPM5 - Balance of Payment Manual, 5th edition	YES
Eurostat recommendations	NO
Other methodological framework (please name methodology):	

Data sources: *Please check all that apply*

International transactions reporting system ITRS	YES
Enterprise survey	YES
Household survey	NO
Statistics on commodity trade	YES
Cross border visitors survey	YES
Partner countries' statistics	YES
Other data sources, please specify:	DIRECT REPORTING OF FINANCIAL INSTITUTIONS

Please briefly describe your *main data sources* for each of the compilation of the main EBOPS service items:

1. Transportation	Beginning in 1996, credit entries are derived from the ITRS. Debit entries are drawn from import statistics compiled by Customs and from information derived from the ITRS.
2. Travel	Data on travel credits are obtained from the National Tourism Administration (NTA). The NTA collects the data through sample surveys conducted by the National Bureau of Statistics. Data on travel debit are estimated on data from the Immigration Administration Department of the Ministry of Public Security and main partner economies' travel receipts from China.
3. Communications services	Derived from the ITRS.
4. Construction services	Derived from the ITRS.
5. Insurance services	Derived from the ITRS and estimate based on customs data.
6. Financial services	Derived from the ITRS and direct Reporting of Financial Institutions
7. Computer and information services	Derived from the ITRS.
8. Royalties and license fees	Derived from the ITRS.

- UNSD collect external trade in services metadata on SITS from countries as part of the data collection on SITS.

More information: http://unstats.un.org/unsd/tradeserv/datacollection.htm

Example:
The Bahamas Balance of Payments quarterly, 2011 and 2012

Explanation of terminology, content and Methodology of some statistical Indicators …

Table 7.1 Balance of Payments

(B$ Millions)

	2011 Qtr.IIIp		2011 Qtr.IVp		2012 Qtr.Ip		2012 Qtr.IIp		2012 Qtr.IIIp	
	Credit	Debit	Credit	Debit	Credit	Debit	Credit	Debit	Credit	Debit
1. CURRENT ACCOUNT	**856.6**	**1,191.2**	**902.9**	**1,248.7**	**979.4**	**1,395.0**	**1,034.3**	**1,235.4**	**961.9**	**1,366.0**
A. Goods & Services	**820.9**	**1,101.0**	**865.2**	**1,136.0**	**929.6**	**1,293.5**	**970.7**	**1,121.2**	**884.5**	**1,232.6**
a. Goods	**216.0**	**785.7**	**211.3**	**819.5**	**236.1**	**902.5**	**230.1**	**807.8**	**254.2**	**785.7**
1. Merchandise	135.4	779.5	130.7	815.2	143.4	896.3	119.2	801.7	116.6	779.5
i. Oil Trade (local Consumption)	0.0	221.7	0.0	191.8	0.0	225.6	0.0	237.4	0.0	274.4
ii. Non-Oil Merchandise	135.4	557.9	130.7	623.4	143.4	670.7	119.2	564.3	116.6	505.2
2. Goods procured in port by carrier	80.6	6.2	80.6	4.3	92.7	6.2	110.8	6.1	137.6	6.2
b. Services	**604.9**	**315.3**	**654.0**	**316.5**	**693.5**	**391.0**	**740.6**	**313.4**	**630.3**	**446.9**
1. Transportation	34.7	90.9	32.3	92.1	31.4	107.3	31.6	99.5	30.3	90.4
i. Passenger Services	6.3	24.6	2.3	36.8	6.3	25.0	6.3	26.7	6.3	23.2
ii. Air and Sea Freight Services	0.0	44.1	0.0	48.4	0.0	60.4	0.0	50.8	0.0	45.2
iii. Port & Airport Charges	28.5	22.2	30.1	6.9	25.1	21.9	25.4	22.0	24.0	22.0
2. Travel	525.4	76.1	575.6	60.6	602.1	55.8	640.6	57.4	533.9	94.1
3. Insurance Services	0.0	44.3	0.0	32.9	0.0	44.9	0.0	31.5	0.0	71.3
i. Freight Insurance	0.0	4.9	0.0	5.4	0.0	6.7	0.0	5.6	0.0	5.0
ii. Non-Merchandise Insurance	0.0	39.4	0.0	27.5	0.0	38.2	0.0	25.8	0.0	66.2
4. Construction Services	0.0	15.1	0.0	55.7	0.0	90.3	0.0	27.7	0.0	88.3
5. Royalty and License Fees	0.0	4.0	0.0	6.1	0.0	3.3	0.0	2.6	0.0	4.2
6. Offshore companies local expenses	28.0	0.0	26.8	0.0	40.2	0.0	47.7	0.0	44.4	0.0
7. Other Services	5.0	73.6	9.7	68.6	8.4	87.4	10.1	89.4	9.6	94.3
8. Government Services	11.7	11.4	9.6	0.5	11.4	2.1	10.5	5.4	12.2	4.4
i. Resident government	1.4	11.4	1.2	0.5	1.4	2.1	1.8	5.4	1.5	4.4
ii. Foreign government	10.4	0.0	8.4	0.0	9.9	0.0	8.7	0.0	10.7	0.0
B. Income	**11.9**	**63.9**	**12.8**	**87.1**	**14.4**	**70.8**	**11.4**	**82.4**	**11.7**	**98.6**
a. Compensation of Employees	**0.0**	**13.9**	**0.0**	**11.6**	**0.0**	**12.6**	**0.0**	**13.6**	**0.0**	**10.5**
1. Labour Income	0.0	13.9	0.0	11.6	0.0	12.6	0.0	13.6	0.0	10.5
b. Investment Income	**11.9**	**49.9**	**12.8**	**75.5**	**14.4**	**58.2**	**11.4**	**68.8**	**11.7**	**88.1**

Source: The Central Bank of the Bahamas

Example:
The Bahamas Balance of Payments quarterly, 2011 and 2012

SECTION 7 **INTERNATIONAL TRADE AND PAYMENTS**

Table 7.1 **Balance of Payments**

The table format is based on the IMF's standard format for reporting balance of payments statistics. The information on oil trade is supplied by oil companies and is desegregated into oil imported for domestic consumption and that for bunkering of foreign ships and aircraft. Oil that is imported for trans-shipment or refining and subsequently re-exported is excluded from the trade account since no change of ownership occurs.

Interest, Dividends and Profits: data for banks exclude transactions relative to offshore activities.

Import and Export: data differ from those published by the Department of Statistics owing, inter alia, to some erratic movement in the series compiled by this department. For the years 1978-1987:3, the Central Bank's estimates for imports were based on import duties to which a multiplier of 3.75 was applied.

Since 1987:4 - 1989, imports have been compiled as a percentage of tourism expenditure. Exports have been estimated from the data supplied by offshore exporting companies and applying a multiplier of 2.22.

Thereafter, the Bank has reverted to using imports and exports data from the Department of Statistics. In the absence of timely data, the Bank estimates exports from previous years' information and imports are obtained from The Bahamas Customs Department.

Travel: debit is based on Exchange Control approvals for purchases of foreign currency.

For a detailed exposition of the components of the table, please see the article, 'An overview of Bahamas Balance of Payments 1973-1979', Quarterly Review, March 1977.

The non-oil imports data for 1985 have been revised upwards to reflect the impact of the sharp upward adjustment in tourist expenditure, as reported by the Ministry of Tourism. The revisions by the ministry reflect the new methodology employed in calculating tourist expenditure by using an average per visit measurement instead of a per diem concept. The adjustments to non-oil imports were designed to maintain the historical relationship which exits between tourism receipts and imports. Accordingly, the freight and insurance data were also revised.