General items of quality

Data quality framework

Indicators of quality

Markie Muryawan
United Nations Statistics Division
© 2017

Data Quality
(IMTS 2010 Chapter IX)

- Four parts
  - Enhancing quality of international merchandise trade statistics
  - Measuring quality of international merchandise trade statistics
  - Quality measures and indicators
  - Cross-country data comparability
  - Metadata (this considered as part of quality)

Enhancing data quality

- Enhancing data quality is a process covering all stages of the statistical production process:
  - starts with validation of the data provided by the trader or broker when completing the customs declaration,
  - requires appropriate institutional arrangements, for example, in order to allow adequate access to different data sources and
  - implies the appropriate use of information and communication technology among many other relevant activities.

Measuring quality

- Most international organizations and many countries have developed definitions of quality, outlining the various dimensions (aspects) of quality and quality measurement, and integrated them into quality assessment frameworks
  - IMF Data Quality Assessment Framework (DQAF)
  - European Statistical System (ESS) adopted a Code of Practice
  - OECD quality measurement framework
  - United Nations National Quantity Assessment Framework (NQAF)

Quality dimensions

- It is recommended that the following dimensions of quality are taken into account while developing such frameworks:
  - Relevance (degree to which they meet the user needs)
  - Credibility (confidence that users place in those statistics)
  - Accuracy
  - Timeliness
  - Methodological soundness
  - Coherence
  - Accessibility

Measures (indicators) of quality

- When countries define the quality indicators for their international merchandise trade statistics, it is recommended that the indicators satisfy the following criteria:

    a) they cover all dimensions of quality

    b) are based on the consistent application of a sound methodology; and

    c) the indicators are easy to interpret both by internal and external users.

- It is recommended that countries maintain a balance between different dimensions of quality and the number of indicators
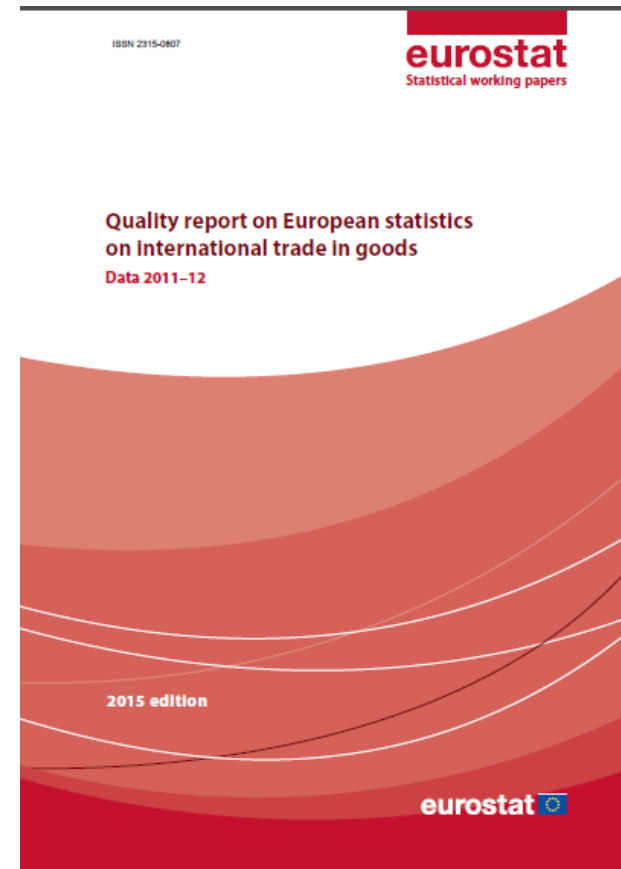
Examples IMTS quality reports

- EU-COMPASS MEDSTAT III 2013

- UK Quality Report 2011

- Eurostat Quality Report 2015

Template of Quality Report

- Based on existing guidelines and reports, template of quality report for IMTS was developed

- This template had been distributed and discussed in several events

- Countries are recommended to start the design of their quality report using this template
  - With addition or amendment, if necessary

Thank you

Please send your comments to
muryawan@un.org

Zero draft –IMTS Report of Quality

Version 20 Feb 2017

In order to measure quality of IMTS data, it is necessary to apply existing data quality frameworks developed by international organisations (e.g., IMF, UNSD, Eurostat, OECD) outlining the various dimensions or aspects of quality. These dimensions (Relevance, Methodological soundness, Accuracy, Timeliness, Punctuality, Accessibility, Clarity, Comparability, and Coherence) are measured by sets of indicator. However there is no internationally agreed data quality indicators related to IMTS. Therefore, template on quality reports including indicators of quality had been developed for further discussion.

The quality dimensions and indicators are described below.

Quality dimensions		Indicators
1. RELEVANCE		**Relevance is the degree to which statistical outputs meet current and potential user needs. It depends on whether all the statistics that are needed are produced and the extent to which concepts used (definitions, classifications etc.) reflect user needs**
1.1.	Statistical information **not** currently available but needed by users	
		Number of missing variables NUMBER List of missing variables TEXT
		Number of missing variables according to the IMTS 2010 recommended data transmission protocol to UNSD NUMBER List of missing variables TEXT
1.2.	What are the main gaps between user needs and published trade statistics? TEXT	
1.3.	Consultation of users or user satisfaction surveys	
		Regular survey? YES/NO

		Main reasons for users' dissatisfaction TEXT
1.4.	Have you established a user database? YES/NO	

2. METHODOLOGICAL SOUNDNESS	Dimension that encompasses the application of the available international standards, guidelines and good practices in the production of data
2.1.	To which degree IMTS concepts and definitions of 2010 related to the coverage and methodological requirements are implemented in trade statistics? (e.g. WELL ADVANCED, PARTIAL COMPLIANCE, BASIC)
	Do you follow general guideline of IMTS scope (record all goods which add to or subtract from the stock of material resources of a country by entering (imports) or leaving (exports) its economic territory)? YES/NO
	Do you estimate trade below threshold if significant? YES/NO
	Do you use general trade system? YES/NO
	Do you compile exports/re-exports/imports/re-exports? YES/NO
	For valuation, do you compile Exports FOB, Imports CIF? YES/NO
	Do you collect or compile quantity information in WCO standards unit of quantity and net weight? YES/NO
	Do you record country of origin for imports? YES/NO
	Do you record country of consignment for imports? YES/NO
	Do you record country of last known destination for exports? YES/NO
	Do you compile trade data by mode of transport? YES/NO
	Do you compile trade data by customs procedure codes? YES/NO
	Do you have formal agreement or coordination mechanism with agencies providing basic data? YES/NO
	Do you publish quality report regularly? YES/NO
	Do you treat all users equally and disseminate data without preference to any national or international user group? YES/NO
	Do you compile and publish external trade indices? YES/NO
	Do you compile and publish seasonally adjusted data? YES/NO
	Do you conduct linking trade and business statistics? YES/NO
	Do you disseminate data in SDMX-IMTS format? YES/NO
2.2.	What are the main conceptual divergences with IMTS 2010? TEXT

3. ACCURACY – Coverage & Estimates		**The accuracy of statistical outputs in the general statistical sense is the degree of closeness estimates to the true values.**

3.1.	Measures taken in order to ensure the data exhaustiveness	
	Data Sources	Is customs declaration your main data source for trade statistics? YES/NO
		Do you use additional data sources to supplement customs declaration? YES/NO If yes, please explain for which type of goods you use non-customs data TEXT
	Correction by Customs	Are the post clearance amendments or changes in customs declarations regularly transmitted to the compiling agency? YES/NO
		If yes, do you systematically include such information in the statistical compilation process? YES/NO If no, please explain the reasons TEXT
	Availability of Customs Procedure Codes	Are all transactions with all customs procedures transmitted to you by Customs, allowing for full statistical control of relevant transactions? YES/NO
		Have you established a correlation table linking customs procedures to special (and/or to general) trade systems? YES/NO
3.2.	What is the share of electronic declarations in data collected by Customs (%of total trade):	
		Imports NUMBER
		Exports NUMBER
3.3.	Measures and practices for specific goods or movements of particular importance in your country (please add or amend, if necessary)	
	Oil & Gas, etc.	What is the data source? TEXT
		Do you publish trade data including oil and oil products? YES/NO
		On import, are you able to identify the country of origin? YES/NO
		On export, are you able to identify the country of destination? YES/NO
		Are you able to exclude transit trade? YES/NO/N.A.
		How do you compile the statistical value? TEXT What is the definition of statistical value for oil? TEXT
		Do you compare quantitative data with the data provided by other data sources (e.g. energy statistics, grid operators, administrative data, etc.) in order to ensure correct coverage? YES/NO
	Military goods	Are the goods intended for military use included in trade statistics (at least total monthly statistical value of exports and imports)? YES/NO
	Ships & Aircrafts	Do you record trade in ships and aircrafts? YES/NO
		What is the data source for recording trade in vessels and aircraft? TEXT

3.4.	Please list specific movements or goods which diverge from recommended coverage by IMTS (e.g., military goods, oil and oil products, etc. not included in the officially published trade statistics) TEXT		
3.5.	Share of collected data from Customs declarations (% from total value)		
		Imports NUMBER	
		Exports NUMBER	
3.6.	Share of data from other data sources (% from total value)		
		Imports NUMBER	
		Exports NUMBER	
3.7.	Share of estimated trade below the statistical threshold (% from total value)		
		Imports NUMBER	
		Exports NUMBER	
3.8.	Availability of quantity information and degree of quantity estimation		
	Number of 8-digit product codes without quantity	Imports NUMBER	
		Exports NUMBER	
	Share of records without quantity in total trade value (in %)	Imports NUMBER	
		Exports NUMBER	
	Share of estimated quantity in total trade value (in %)	Imports NUMBER	
		Exports NUMBER	
4. ACCURACY - Revision and thresholds	**The accuracy of statistical outputs in the general statistical sense is the degree of closeness estimates to the true values.**		
4.1.	Annual revision rate (trade values ratio of first release of annual to the latest release) (in %)		
		Imports NUMBER	
		Exports NUMBER	
4.2.	How often do you revise the data? MONTHLY/QUARTERLY/ANNUAL/NO REVISION		
4.3.	Do you compile revision data base which contains archives of previous revision? YES/NO		
4.4.	In how many months after the end of the reference period the final annual data become available? (e.g., in 10 month) NUMBER		
4.5.	Are corrections, amendments applied to the correct reference month? YES/NO		
4.6.	What is the statistical value threshold under which customs declarations are not processed in details? (in USD) NUMBER		

5. ACCURACY - Confidentiality		The accuracy of statistical outputs in the general statistical sense is the degree of closeness estimates to the true values.	
5.1.	Do you apply passive confidentiality? YES/NO		
5.2.	If you apply passive confidentiality, how the traders are informed about the possibility to hide the data at product level? TEXT		
5.3.	Impact of confidentiality in detailed data in figures		
	Number of affected 8-digit product codes	Imports NUMBER	
		Exports NUMBER	
	Number of affected individual trading partners	Imports NUMBER	
		Exports NUMBER	
	Share of confidential records in total trade value (in %)	Imports NUMBER	
		Exports NUMBER	

6. ACCURACY – Data validation		The accuracy of statistical outputs in the general statistical sense is the degree of closeness estimates to the true values.	
6.1.	Data Validation on input data		
		Do you check the exhaustiveness of declarations (i.e., that all compulsory boxes are filled in)? YES/NO if yes, please describe TEXT	
		Do you apply validity checks (e.g., country code, commodity code)? YES/NO if yes, describe TEXT	
		Do you apply credibility checks (e.g., certain customs regime can only be applied to specific transaction, unit values within acceptable range)? YES/NO if yes, please describe TEXT	
6.2.	After data validation on input data, what is the average share of erroneous data?		
		As percentage of the total number of declarations? NUMBER	
		As a percentage of the total value of trade? NUMBER	
6.3.	Data Validation on output data		
		Do you apply time series check on output data? YES/NO if yes, describe TEXT	
		Do you apply outlier detection checks? YES/NO if yes, please describe TEXT	
		Do you compare with other sources (fiscal data, mirror statistics)? YES/NO if yes, please describe TEXT	

		The timeliness of statistical outputs is the length of time between the event or phenomenon they describe and their availability. Punctuality is the time lag between the release date of data and the target date on which they were scheduled for release as announced in an official release calendar.
7.	**TIMELINESS AND PUNCTUALITY**	
7.1.	What is the time lag between the end of the reference period and the date of dissemination of the first [preliminary] results? (in calendar days)	
		Monthly data NUMBER
		Annual data NUMBER
7.2.	What is average time lag between the end of the reference period and the release date of the final data? (in months)	
		Monthly data NUMBER
		Annual data NUMBER
7.3.	Does data release calendar define precisely the dates for monthly publication of trade statistics? YES/NO If yes, what is an average delay for publication of monthly results (in no of days) NUMBER	
8.	**ACCESSIBILITY AND CLARITY**	The accessibility of statistical outputs is the measure of the ease with which users can obtain the data. The clarity of statistical outputs is the measure of the ease with which users can understand the data.
8.1.	Please mark data dissemination channels for trade statistics:	
		Paper publications YES/NO (please mention the type of publications) TEXT
		Interactive web dissemination tool YES/NO (please mention the level of details, e.g. product at 8 digit level*partner country*value*quantity*etc.) TEXT
		Add hock reports on request by users YES/NO
		Routine (monthly/annual) data submission to UNSD YES/NO
8.2.	Frequency of data dissemination (e.g. MONTHLY/QUARTERLY/SEMI-ANUUAL/ANNUAL/IRREGULAR)	
8.3.	Are the data provided free of charge to all users? YES/NO	
8.4.	What type of metadata is provided to the users?	
		Basic concepts are provided in paper publications on trade statistics YES/NO
		Basic concepts are provided on the website of the national statistical office YES/NO
		Detailed metadata are provided to the users on the website of the national statistical office YES/NO
9.	**COMPARABILITY AND COHERENCE**	Statistics are consistent internally, over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different sources.

9.1.	Which nomenclatures/classifications do you use in your trade statistics?	
		Commodity: TEXT
		Trading Partner: TEXT
		Quantity Unit: TEXT
		Customs Procedure Codes: TEXT
9.2.	Do you perform regular reconciliation study to reduce asymmetries with your trading partners? YES/NO	
9.3.	Please mentioned the latest (within the last 3 years) changes in nomenclatures, in definitions of data elements and their availability, changes in coverage of customs procedures, etc., which had an impact on the comparability of the time series of the most detailed data? TEXT	
9.4.	Do you provide assistance to BOP trade in goods compilers with regard to IMTS data adjustment? YES/NO	