



General items of quality

Data quality framework

Indicators of quality

Markie Muryawan
United Nations Statistics Division
© 2016

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

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:
 - Prerequisites of quality
 - Relevance (degree to which they meet the user needs)
 - Credibility (confidence that users place in those statistics)
 - Accuracy
 - Timeliness
 - Methodological soundness
 - Coherence
 - Accessibility

Measures 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 of quality indicators

Relevance

Quality dimensions	SUB-ITEMS			
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 according to the agreed data transmission protocol			
	List of variables			
1.2	Consultation of users or user satisfaction surveys			
	Regular survey? YES/NO			
	Reasons for users' dissatisfaction			
1.3	Have you established a user database? YES/NO			
1.4	What are the main gaps between user needs and published trade statistics?			

Accuracy - Coverage

2. ACCURACY - Coverage	The accuracy of statistical outputs in the general statistical sense is the degree of closeness estimates to the true values.		
2.1	Measures taken in order to ensure the data exhaustiveness		
2.2	Measures and practices for specific goods or movements		
2.3	Share of collected data from Customs declarations (% from total value)		
2.4	Share of data from other data sources (% from total value)		
2.5	Share of estimated trade below the statistical threshold (% from total value)		
2.6	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).		

Accuracy – Revision & thresholds

3. ACCURACY - Revision and thresholds			
3.1	Annual revision rate (first release of annual exports, imports /the latest release) in %		
3.2	How often do you revise the data? monthly/quarterly/annually/no corrections		
3.3	Do you compile revision data base? YES/NO		
3.4	In how many month after the end of the reference period the final annual data become available? (e.g., in 10 month)		
3.5	Are corrections, amendments applied to the correct reference month? YES/NO		
3.6	What is the statistical value threshold under which customs declarations are not processed in details? (for comparability purposes please provide in USD)		

Accuracy - Confidentiality

4. ACCURACY - Confidentiality						
4.1	Do you apply passive confidentiality? YES/NO					
4.2	If you apply passive confidentiality, how the traders are informed about the possibility to hide the data at product level?					
4.3	Confidentiality in detailed data - Impact in figures	Number of 8-digit product codes affected				
		Share of confidential records in total trade value				

Accuracy – Control Procedure

5. ACCURACY - Control procedures					
5.1	Do you check the exhaustiveness of declarations (i.e., that all compulsory boxes are filled in)? YES/NO (if YES, please provide description in "Additional comments" column.				
5.2	Do you apply validity checks (e.g., the validity of codes)? YES/NO (if YES, please provide description in "Additional comments" column.				
5.3	Do you apply credibility checks? YES/NO (if YES, please provide description in "Additional comments" column.				
5.4	Do you apply logical checks? (e.g. comparisons with historical data or other data sources or other statistics). YES/NO (if YES, please provide description in "Additional comments" column.				
5.5	Do you apply outlier detection checks? YES/NO (if YES, please provide description in "Additional comments" column.				
5.6	What is the share of electronic declarations in data collected by Customs:				
	as a percentage of the total number of declarations				
	as a percentage of the total value of trade				

Timeliness and Punctuality

6. TIMELINESS AND PUNCTUALITY	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.
6.1	What is the time lag between the end of the reference month and the date of dissemination of the first results? (in calendar days)
6.2	What is average time lag between the end of the reference month and the release date of the final results? (in months)
6.3	Does your national data release calendar defines precisely the dates for monthly publication of trade statistics? YES/NO
6.4	If YES, what is an average delay for publication of monthly results (in No of days)

Accessibility and Clarity

7. 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.
7.1	Please mark data dissemination channels for trade statistics:
	paper publications YES/NO (please mention the type of publications in additional comments)
	interactive web dissemination tool YES/NO (please mention the level of details, e.g. product at 8 digit level*partner country*value*quantity*etc.)
	add hock reports on request by users YES/NO
7.2	Data dissemination frequency (e.g. monthly / quarterly / half year / annually / irregular)
7.3	Are the data provided free of charge to all users? YES/NO
7.4	What type of metadata are 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
	other, please provide details

Comparability and Coherence

8. 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.

8.1 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, partially, basic)

8.2 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?

Examples IMTS quality reports

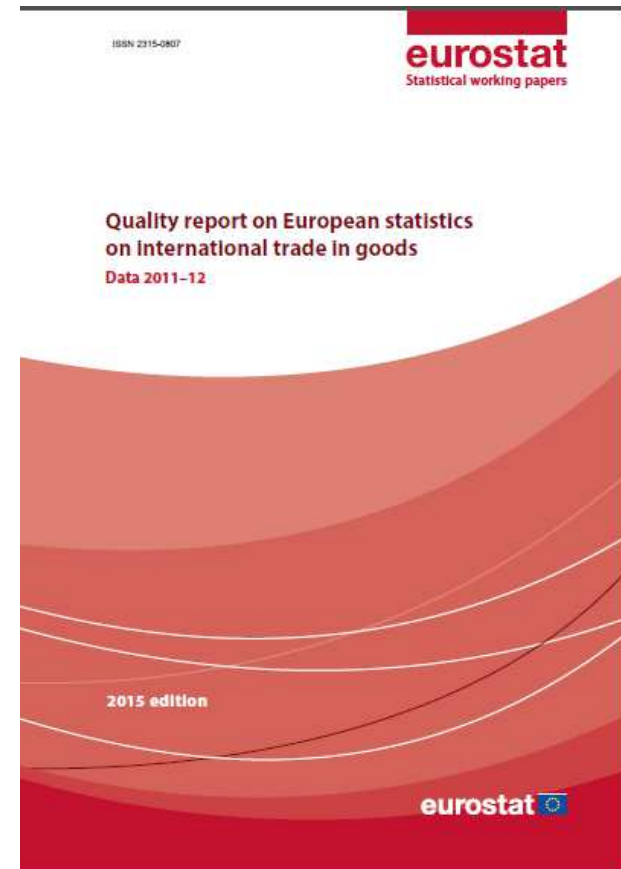
- EU-COMPASS MEDSTAT III 2013
- UK Quality Report 2011
- Eurostat Quality Report 2015



Quality Report UK Trade-in-Goods Statistics

October 2011

ECSM Trade Statistics
HM Revenue & Customs
Alexander House
21 Victoria Avenue
Southend-on-Sea
Essex SS99 1AA



Thank you

Please send your comments to
muryawan@un.org