TRAINFORTRADE

Strengthening knowledge and skills through innovative approaches for sustainable economic development

Module 4

Metadata and Data Quality

Participants Manual

Training Course on International Merchandise Trade Statistics (IMTS)

UNCTAD TrainForTrade | UNCTAD Statistics | UNITED NATIONS | WORLD TRADE ORGANIZATION

TABLE OF CONTENTS

Forward

E-learning course on international merchandise trade statistics

International trade in goods are an increasingly important part of global commerce. International Merchandise Trade Statistics (IMTS) play a vital part in monitoring, analysing and projecting macroeconomic developments in individual economies and the world economy. IMTS are well developed and international standards have been defined in the IMTS 2010 to harmonize standards and practices. However, many measurement issues and comparability problems remain. Increasingly the important phenomenon of globalization, entailing the internationalization of production and sales and new forms of delivering goods and services to customers across countries, new developments in information and communication technologies and the growing importance of e-commerce requires new approaches and poses new methodological challenges.

The objective of this e-learning course on IMTS is to provide more easily accessible and rather complete training material for those involved in the collection, compilation, analysis and dissemination of IMTS.

The main goals are to enhance statisticians' ability to apply the most recent internationally agreed recommendations on IMTS, define best possible data sources, set up adequate collection systems, and enhance statistics compilation processes. Furthermore, the course would communicate the importance of quality, metadata, timely dissemination, and links to economic analysis and national policy objectives. The training would guide trainees on how to better use the internationally available guidance, especially the IMTS Concepts and Definitions 2010 and the related Compilers Manual.

Notes and acknowledgements

This course has been developed and produced jointly by:

- UNCTAD, Division on Globalization and Development Strategies, Development Statistics and Information Branch;
- UNCTAD, Division on Technology and Logistics, Knowledge Development Branch, Human Resources Development Section/TrainForTrade;
- WTO, Economic Research and Statistics Division, International Trade Statistics Section and
- UN Statistics Division.

It has also benefited from the close collaboration of the Inter-Agency Task Force on 'Statistics of International Trade Statistics'.

The training package has been drafted by Romesh Paul, with the collaboration of: Steve Mac Feely, Mark Assaf, Dominique Chantrel, Aylwin Zabula, Sanja Blazevic, Yumiko Mochizuki, Flavine Creppy, Andreas Maurer, Barbara D'Andrea, Mercedes Ninez Piezas-Jerbi, Christophe Degain, Markie Muryawan, and Nancy Snyder.

1. INTRODUCTION

International trade is defined, in the Balance of Payments as in the System of National Accounts, as the whole set of international transactions in items that are outcomes of production activities. The term "international merchandise trade statistics" (IMTS) refers to a specialized multipurpose domain of official statistics concerned with the provision of data on the movements of goods between countries and areas.

This module on 'Metadata and data quality' describes the IMTS metadata, monitoring its progress, scheduling of work, producing report on data quality and production of IMTS on a timely basis.

This module is based on the recommendations contained in chapter IX of 'IMTS: Concepts and Definitions 2010' (IMTS 2010), on data quality and metadata, as well as on chapter IX of 'IMTS: Compilers Manual' (IMTS 2010-CM), on data quality.

Data on international merchandise trade statistics are the end product of a complex process comprising many stages, from the collection and processing of basic records to the compilation and dissemination of official statistics. Recommendations on how to approach the issue of data quality in a systematic and comprehensive way are provided in section 2 below.

The objective of quality measurement of international merchandise trade statistics can be twofold: to provide producers with the appropriate information to monitor and further enhance data quality and to provide users with sufficient information to judge whether the data are adequate for their intended use (to judge their "fitness for use"). Quality is a multidimensional concept that is difficult to measure

2. IMTS 2010 RECOMMENDATIONS

IMTS 2010 new recommendations

Data quality and metadata (chap. IX)

43. Systematic approach to data quality: Follow a systematic approach to data quality and develop standards and related good practices covering the institutional arrangements, the statistical processes and outputs (the entire trade statistics programme) (para. 9.4). *New recommendation*
44. Standard for quality reports: Develop a standard for regular quality reports that cover the full range of statistical processes and outputs and are based on principles and standards (para. 9.5). *New recommendation*
45. Frequency of quality reports: Complete or update quality reports of international merchandise trade statistics at least every five years, or more frequently if significant methodological changes or changes in the data sources occur (para. 9.6). *New recommendation*
46. Contents of quality reports: Base quality reports on a set of quantitative and qualitative indicators for international merchandise trade statistics and on a checklist covering data collection, processing and dissemination in order to allow for an assessment of strengths and weaknesses in the statistical process and to identify possible quality improvement actions (para. 9.7). *New recommendation*
47. Dimensions of data quality: Take the following dimensions into account when developing a quality assessment framework: prerequisites of quality, relevance, credibility, accuracy, timeliness, methodological soundness, coherence and accessibility (para. 9.10). *New recommendation*
48. Quality indicators: Ensure that the defined quality indicators satisfy the following criteria: (a) they cover all dimensions of quality; (b) they are based on the consistent application of a sound methodology; (c) the indicators are easy to interpret by both internal and external users (para. 9.15). *New recommendation*
49. Cross-country data comparability: Countries are encouraged to periodically conduct bilateral and multilateral reconciliation studies or implement data exchanges (para. 9.18). *Updated encouragement*
50. Metadata categories: Cover at least the categories of metadata provided in para. 9.23 (para. 9.23). *New recommendation*
51. Metadata as high priority: View the development of metadata as a high priority and consider their dissemination an integral part of the dissemination of international merchandise trade statistics (para. 9.25). *New recommendation*

3. SYSTEMATIC APPROACH TO DATA QUALITY

Enhancing the data quality of international merchandise trade statistics

Enhancing data quality is a process covering all stages of the statistical production process and cuts across all issues covered in the IMTS 2010 recommendations. It 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.

Systematic approach to data quality

Actions for enhancing data quality are often focused on individual elements of the data, such as reported commodity, value, quantity, quantity unit and trading partner (or combinations thereof), but must also address more general issues of coverage and comprehensiveness of recording. The systematic approach to data quality implies that all aspects of the entire trade statistics program be examined and evaluated against certain principles and standards in order to more effectively identify and implement appropriate action to further improve data quality. It is recommended that countries develop such standards and related good practices covering institutional arrangements, statistical processes and outputs.

Quality reporting

Further, it is recommended that countries develop a standard for regular quality reports that cover the full range of statistical processes and their outputs and use the above-mentioned principles and standards as its basis. Such reports can be either producer-oriented, with the aim of identifying strengths and weaknesses of the statistical process and leading to or containing the definition of quality improvement actions, or user-oriented, with the aim of keeping users informed on the methodology of statistical process and the quality of statistical output.

It is recommended that quality reports of international merchandise trade statistics be completed or updated at least every five years, or more frequently if significant methodological changes or changes in the data sources occur. For monitoring the quality of the processes and effectiveness of quality-improvement actions, reviews should be conducted more frequently.

It is recommended that countries base their quality reports on a set of quantitative and qualitative indicators for international merchandise trade statistics, as well as on a checklist covering data collection, processing and dissemination, in order to assess the strengths and weaknesses in the statistical process and to identify possible quality-improvement actions.

4. DIMENSIONS OF DATA QUALITY

Figure 1:Dimensions of quality

National statistical offices and other agencies responsible for the compilation and dissemination of international merchandise trade statistics may decide to implement one of the existing frameworks for quality assessment or develop on their basis national quality-assessment frameworks that best fit their country's practices and circumstances. It is recommended that the following dimensions of quality are taken into account while developing such frameworks: prerequisites of quality, relevance, credibility, accuracy, timeliness, methodological soundness, coherence and accessibility.

These dimensions may be briefly defined as follows:

(a) Prerequisites of quality. Prerequisites of quality refer to all institutional and organizational conditions that have an impact on the quality of international merchandise trade statistics. The elements within this dimension include the legal basis for the compilation of data; the adequacy of data-sharing and coordination among data-producing agencies; assurance of confidentiality; the adequacy of human, financial and technical resources for implementation of international merchandise trade statistics programmes and the implementation of measures to ensure their efficient use; balancing of the costs to data providers with user needs, especially where use of non-customs sources of data is considered; and quality awareness;

(b) Relevance. The relevance of international merchandise trade statistics reflects the degree to which they meet user needs;

(c) Credibility. The credibility of international merchandise trade statistics refers to the confidence that users place in those data, based on the image of the statistical office or agency that produces the data. The confidence of users is built over time. One important aspect of credibility is trust in the objectivity of the data, which implies that the data are perceived to be produced professionally in accordance with appropriate statistical standards, and that policies and practices are transparent. For example, data should not be manipulated, nor should their release be timed in response to political pressure;

(d) Accuracy. The accuracy of international merchandise trade statistics refers to the closeness of the disseminated statistics to the true (yet unknown) characteristics of the trade flows and can be assessed only indirectly. It has many facets and in practice there is no single aggregate for or overall measure of accuracy (see sect. 6 below for indicators of accuracy);

(e) Timeliness. The timeliness of international merchandise trade statistics reflects the time between the end of the reference period to which the data pertain and the date on which the data are released. Timeliness is closely tied to the existence of a publication (including revisions) schedule. A publication schedule may comprise a set of target release dates or may entail a commitment to release international merchandise trade data within a prescribed time period following their receipt. This factor usually involves a trade-off with respect to

accuracy. The improved timeliness of statistics enhances their relevance;

(f) Methodological soundness. Methodological soundness is a dimension that encompasses the application of the available international standards, guidelines and good practices in the production of international merchandise trade statistics.

(g) Coherence. The coherence of international merchandise trade statistics reflects the degree to which the data are logically connected and mutually consistent, as well as the degree to which they can be successfully brought together with other statistical information within a broad analytical framework and over time.

(h) Accessibility. The accessibility of international merchandise trade statistics refers to the ease with which they can be obtained from the statistical office, including the ease with which the existence of information can be ascertained, as well as the suitability of the form or the media of dissemination through which the information can be accessed. Accessibility also includes the availability of metadata and the existence of user-support services, and requires the development of an advance release calendar so that users will be informed in advance when and where data will be available and how to access them.

These dimensions of quality are overlapping and interconnected. Actions taken to address or modify one aspect of quality will tend to affect other aspects. For example, there may be a trade-off between aiming for the most accurate estimation of the value of a country's total exports and imports and providing that information in a timely manner.

Developing a quality-assessment framework

Countries are also encouraged to develop their own quality-assessment frameworks based on the above-mentioned dimensions, taking into account the specific circumstances of their countries. The adoption of a quality-assessment framework offers the compiling agency a practical approach for a self-assessment of the statistical programme and for the assessment of whether the data provided meet the needs of users. Also, the publication of quality reports allows users to judge for themselves whether a data set meets their particular quality requirements.

Use of quality measures and indicators

The measurement of quality of any statistical data, including international merchandise trade statistics data, is not a simple task. Problems arise from the difficulties involved in quantifying the levels of individual dimensions and in aggregating the levels of all dimensions. Under these circumstances, deriving a single quantitative measure of quality is not possible. In the absence of such a single measure, countries are encouraged to use a system of quality measures and indicators following the recommendations contained in section 6 below.

5. QUALITY REPORTS

The purpose of quality reports is to provide both the producers and users of IMTS with a tool for assessing the quality of these statistics. A quality report is for both internal and public use. It provides a summary of the main quality indicators mentioned in section 6 below. The quality report should be updated regularly and should include quality trends.

Guidelines for the production of quality reports in the ESS

Within the European Statistical System (ESS), very strong efforts have been undertaken to develop a concept of quality and to implement it comprehensively. A central achievement was the adoption of the European Statistics Code of Practice, which provides a broad conceptual framework for viewing quality and sets standards for the institutional environment, statistical processes and statistical outputs. The ESS Standard for Quality Reports issued by Eurostat provides recommendations for preparing comprehensive quality reports for the full range of statistical processes and their outputs. The ESS Handbook for Quality Reports provides much more detailed guidelines and examples of quality reporting practices.

Specific objectives of the ESS guidelines

The specific objectives of the ESS guidelines contained in the ESS Standard for Quality Reports are:

a. To promote harmonized quality reporting across statistical processes and their outputs within a country and hence to facilitate comparisons across processes and outputs;

b. To promote harmonized quality reporting for similar statistical processes and outputs across countries and hence to facilitate comparisons across countries; and

c. To ensure that reports include all of the information required to facilitate identification of statistical process and output quality problems and potential improvements.

Structure of the ESS guidelines

The guidelines are organized by statistical output and process quality components, with the primary section headings being:

1. Introduction to the statistical process and its outputs (overview to provide context);
2. Relevance (output quality component);
3. Accuracy (output quality component);
4. Timeliness and punctuality (output quality components);
5. Accessibility and clarity (output quality components);
6. Coherence and comparability (output quality components);
7. Trade-offs between output quality components;
8. Assessment of user needs and perceptions (covering all aspects of output quality);
9. Performance, cost and respondent burden (process quality components) and
10. Confidentiality, transparency and security (process quality components).

User-oriented quality reports

User-oriented quality reports are keeping users informed about the methodology of the statistical process and the quality of statistical output. Many statistical agencies have adopted principles and standards for data quality and a data quality assessment framework which outlines the different dimensions of quality and their measurement. The quality assessment framework provides a general layout for the quality report; however, not all dimensions are equally relevant for users. User-oriented quality reports are often provided on an on-going basis as part of the metadata provided to users, and are updated regularly.

Producer-oriented quality reports

Producer-oriented quality reports aim at identifying strengths and weaknesses of the statistical process and lead to, or contain the definition of, quality improvement actions. Producer-oriented quality reports can be either motivated by internal interest or externally mandated. They can take the form of, for example, internal review, benchmarking (comparison with others) and audits. By their nature, producer-oriented quality reports are often produced for particular reasons, for example, to fulfil a specific external requirement or to deal with specific issues or problems.

Steps and guidelines for producing data quality reports

Some information about data quality is available in all offices involved in the compilation and dissemination of official trade statistics and its compilation often provides the starting point for a quality report or might already be considered to be a quality report. The following considerations (or steps) that are suggested for the production of a data quality report for merchandise trade statistics also appear to be relevant when a quality report already exists but is being reviewed:

a. Collection and review of existing standards, guidelines, requirements, practices, examples or past quality reports within the same office or other offices within the same country or in other countries, including regional and international organizations, as appropriate, in order to ensure that standards and best practices are followed;

b. Discussion and decision on the scope and type (purpose) and frequency of the quality report under consideration and on the available resources;

c. Assembly of a team and allocation of resources;

d. Elaboration and discussion of the detailed structure of the quality report;

e. Compilation of the required information: quality assessment and measurement;

f. Drafting of report;

g. Review of report;

h. Dissemination and communication of the report and its results.

Information on quality measurement

As an example, the following metadata on quality are provided to users of United Kingdom trade statistics:

a. Quality standards against which the quality is measured and

b. Assessment (including self-assessment) against some of these quality standards, including:

- quantitative assessment against indicators for the six output quality dimensions;
- qualitative assessment of own methods and adherence to EU legislation;
- channels for and results of post publication quality assurance.

5. QUALITY INDICATORS

Quality measures

Quality measures directly reflect a particular aspect of quality. For example, the time lag from the end of the reference period to the release of particular international merchandise trade statistics is a direct quality measure. However, in practice, quality measures can be difficult or costly to calculate. Instead, quality indicators may be used for the quality assessment.

Figure 2: Quality measures

Quality indicators

Quality indicators provide summarized quantitative or qualitative evidence about the quality of the data. They are generally defined with respect to some reference point and can assist in making different types of comparisons. When countries define the quality indicators for their international merchandise trade statistics, it is recommended that they ensure that the indicators satisfy the following criteria:

a. They cover all dimensions of quality as defined in section 4 above;

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

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

It is recommended that countries maintain a balance between the different dimensions of quality and the number of indicators. The objective of quality measurement is to have a limited set (minimum number) of indicators that can be used to measure and follow over time the quality of international merchandise trade statistics and to ensure that users are provided with a useful summary of overall quality, while not overburdening respondents with demands for unrealistic amounts of quality metadata.

Suggested quality measures and indicators

Table below presents a possible set of indicators (and measures) that countries may wish to consider using to measure the quality of international merchandise trade statistics. Their utilization provides users with a clear and up-to-date overview of the overall quality of international merchandise trade statistics.

Table 1: Suggested quality measures and indicators

Quality dimension	Quality measure and indicator
Relevance	1. Gaps between key user interests and compiled international, merchandise trade statistics in terms of concepts, coverage and detail. 2. Results of users' satisfaction surveys and meetings with user groups.
Accuracy	3. Application of reporting thresholds. 4. Under coverage (percentage of non-reporting due to thresholds, percentage of non-reporting due to non-response). 5. Characteristics and frequency of revisions (e.g., as percentage of total value). 6. Application of confidentiality and its impact. 7. Use of data validation techniques and their impact. 8. In the case of sample surveys-based international merchandise trade estimates, the accuracy can be measured using the following indicators: 9. Sampling errors. 10. Non-sampling errors: (a) Unit response rate; (b) Item response rate. 11. Number and average size of revisions of international merchandise trade data.
Timeliness	Time lag between the end of the reference period and the date of the first release (or the release of final results) of international merchandise trade data.
Methodological Soundness	Number and degrees of divergences from the relevant international statistical standards in concepts and measurement procedures that are used in the collection and compilation of international merchandise trade statistics (preferably in terms of the amount of data affected).
Coherence	1. Use of common concepts, classifications, data sources and methods. 2. Availability of appropriate bridging tables.
Accessibility	1. Number and types of means used for dissemination of international merchandise trade statistics. 2. Degree to which all detailed data sets are made available, as a percentage of total international merchandise trade statistics data sets produced. 3. Dissemination of complete metadata used.

7. BILATERAL TRADE ASYMMETRIES

Cross-country data comparability remains an important issue. Non-comparability is caused, inter alia, by differences in coverage, different methods for the treatment of certain goods (e.g., military goods, ship's stores, confidential data), value increases in intermediary countries, differences in classification of goods, time lags in reporting, differences in valuation, including CIF/FOB differences, currency conversion, methods of partner country attribution, and trade via third country intermediaries.

Such non-comparability may be substantially reduced by the adoption of the recommended IMTS 2010 concepts and definitions. Nevertheless, a certain amount of non-comparability will remain because of differences resulting from following the recommended approaches to valuation and partner country attribution for imports and exports, as well as variations in data sources, reporting errors, errors in data collection or the processing and forwarding of results, the use of fraudulent documents or the inability of traders to furnish accurate information.

Countries are encouraged, therefore, to periodically conduct bilateral and multilateral reconciliation studies or implement data exchanges so that their statistics can be made more accurate and useful for both national purposes and for international comparisons.

Goals of bilateral reconciliation studies

Often, it is assumed that exports of country A to country B should be equal to imports of country B from country A. This, however, is rarely the case. Reconciliation studies aim towards explaining these discrepancies between the bilateral imports and exports statistics of trading partners, i.e., by identifying conceptual and methodological differences in their respective types of data collection. However, even if both countries compile their data using the same methodological principles and international guidelines, their bilateral merchandise trade data will usually not match owing to multiple factors (see para. "Reasons for differences in data" below). The purpose of a reconciliation study may be limited to the identification of major differences in the statistics of two countries and the results will help each partner to better understand the bilateral trade flows. However, the reconciliation process may reveal systematic measurement errors and gaps, which should be corrected. Therefore, the goal of reconciliations studies might include assessing the causes of differences and making adjustments to various data components. Another, even more ambitious objective could be the harmonization of the conceptual framework of the two sets of statistics, which could lead to the revision of certain procedures and definitions and, in some cases, could suggest the use of alternative data sources, all of which is expected to improve the overall quality of foreign trade data. At the policy level, a reconciliation exercise will yield a common perception of the facts and can thus facilitate the development of bilateral economic negotiations and international cooperation.

Limitations of the reconciled data

Reconciliation studies might entail the creation of a reconciled data set. However, the reconciled data do not reflect any changes in the officially published trade figures of either partner country. Reconciliation adjustments normally include a series of estimates, which are not sufficiently precise to permit modifications to officially published data. For example, many countries' import data are valued on a CIF-type basis, that is, including insurance and freight charges, which must be removed during reconciliation, since the partner country's exports are usually valued on an FOB-type basis; however, estimates of insurance and freight charges are usually derived indirectly and do not necessarily reflect their true amount.

Basic procedures for bilateral reconciliation studies

In general, reconciliation may include the following activities:

a. Setting the objectives for the project and reaching agreement on the basic procedures;

b. Establishing a common conceptual framework for reconciliation purposes;

c. Converting officially published data to the common framework;

d. Examining the differences in data and methodology;

e. Making necessary data adjustments to achieve mutually agreed sets of trade figures; and

f. Formulation of conclusions of the reconciliation study.

An individual study could be limited to the activities described in (a), (b) and (c) above. It is advised that a reconciliation study cover trade for at least a full year, and that a reconciliation table (which identifies all the additions and subtractions that need to be performed in order for the trade data of one partner to correspond to the trade data reported by the other) be prepared as one of the outputs.

Cooperation between trading partners

The success of reconciliation is dependent on the full cooperation of trading partners from the very beginning and on clear identification of procedures to be followed at all stages of the process, from the initial exchange of the required information to the mutual agreement on the final results. The agencies conducting the reconciliation study should examine not only the various organizational aspects of the proposed study but also its legal implications (for example, in certain cases, an exchange of data at the level of transactions may involve the issue of confidentiality). At the beginning of a reconciliation exercise, both parties have to agree on which data should be used as the benchmark for a specific category of goods. For example, in the past, import data were normally used as a benchmark for the comparison of most commodities. Import data were, in general, considered to be of better quality than export data because imports are reported in sufficient detail to allow customs to apply duties, taxes or other regulatory controls. However, for certain commodities and in some countries, export data were viewed as being more accurate for the same reasons.

A common conceptual framework and conversion of data to that framework

Establishing the common conceptual framework involves an exchange and comparison of methodologies and compiling practices, and adoption of the same definitions and classifications for use in the reconciliation study. Issues to be considered are, among others: what are the major conceptual differences; whether information is available on a country of origin/last known destination or on some other basis; whether there are significant differences in compilation procedures (such as suppression of confidential or low-value trade) which will affect bilateral comparability; and whether there are certain transactions (such as processing trade) for which streamlined reporting provisions exist that could affect comparability. The common framework serves as a practical working tool with which to facilitate comparison of data between the two countries: it does not replace official methodologies of the countries involved. The partners must also decide on such issues as the working currency for the study, and whether currency conversion should be carried out on a monthly or an annual basis (if exchange rates are fluctuating significantly, annual conversion could create additional discrepancies).

Reasons for differences in data and further guidance

Coverage. Specific goods or types of transactions may be defined differently by trading partners, and may be included in the trade statistics by one partner but excluded by the other (e.g., leased goods, military goods, goods imported or exported for or after repair). Further, countries usually have different provisions for the treatment of low-value shipments, which may be excluded from statistics, reported in less detail, or estimated instead of being compiled.

Trade systems. If one partner uses the special system of trade and the other uses the general system, goods moving between premises for customs warehousing and customs free zones of those countries will not be accounted for by the country using the special system. Where both countries use the special system of trade, goods moving between customs free zones will not be recorded by either country and will not affect their export and import totals. To facilitate reconciliation, countries should clearly define their statistical territories, specifying any particular inclusions and exclusions. For example, Puerto Rico and the United States Virgin Islands are part of the statistical territory of the United States; therefore, exports to/imports from those territories should be recorded as trade with the United States in any reconciliation study involving the United States.

Timing of recording. Many factors contribute to timing differences, including the time needed for the transportation from the exporting to the importing country, the time needed for the completion of the customs procedures and warehousing operations. Further, the filing and recording of various documents at different stages and their statistical recording might follow different practices. For example, in one country, a trade flow may be attributed to the time period in which the invoice is received in the importing country, while another country may attribute the transaction

to the time period in which the amounts owed to the customs administration are paid. Also, the deadline for reporting statistical information, the use of summary reporting, the definition of the reporting period and the procedures for handling late or incorrect records might be different in countries and therefore affect the time of recording. Such timing differences can be significant, particularly in the case of monthly data or where the level of trade in a given commodity has changed extensively (so that effects of timing differences between the study period and the preceding and succeeding periods are not equivalent).

Interpretation and application of the commodity classification. All trading countries have adopted the Harmonized System (HS) for commodity classification. Despite that significant achievement, there are differences in interpreting and applying the HS, both within the same country and among different countries. In order to reconcile trade in particular commodities, an analysis of uniformity of the HS application might be required. Differences and errors in classification normally affect only the distribution of the goods among different classes; however, they may sometimes lead to differences in total trade, which would be the case, for example, if different recording thresholds were applied to different commodities.

Valuation. IMTS 2010 (para 4.4) recommends that countries adopt the World Trade Organization Agreement on Customs Valuation as the basis for valuation of their international merchandise trade for statistical purposes (whether the country is a World Trade Organization member or not) for both imports and exports. Further, it is recommended that imports be recorded on a CIF-type basis and exports on a FOB-type basis. Therefore, CIF imports would exceed the counterpart export value by the value of international insurance and freight charges even if there were no other sources of difference. Where such charges have been included, a negative adjustment is made to remove them, for comparison with FOB export values. If the actual freight charges are not known, estimates may be derived from unit value differences or through other approaches, such as the application of general CIF/FOB ratios. The determination of the customs value of imported goods is regulated by the WTO Agreement on Customs Valuation.

However, there can be a multiplicity of other reasons for valuation differences, including undetected under and over-declaration of values, different value estimates for transactions without valuation such as charitable/relief shipments, barter trade or related party transactions and different views on the exclusion or inclusions of services. In all of these cases, compensating adjustments are needed if the differences are significant.

Currency conversion. Currency conversion practices for goods invoiced in foreign currencies may also cause discrepancies between one country's import value and the counterpart export value, particularly when the exchange rate between the partners fluctuates rapidly. Further, the procedures used in the reconciliation for expressing both sets of statistics in the same currency for comparison purposes can also create discrepancies.

Partner-country attribution. It is recommended that imports be attributed to the country of origin and exports to the country of last known destination (see IMTS 2010, para. 6.25). This partner-country attribution can explain many differences between the statistics of trading partners in cases when goods move from the country of origin to the country of destination through third countries. For example, in the case where goods are produced in country A, sold and shipped to country B and afterwards resold and dispatched to country C, the trade statistics of country B will show exports to country C, but the statistics of country C will not attribute its imports to country B: they will indicate that goods were imported from country A.

Different application of rules of origin. If countries have different rules of origin, the trade flows will also be recorded differently. Consider the following example. Goods are produced in country C, imported by country A, undergo certain processing and are exported to country B. If countries A and B have different rules of origin, the processed goods dispatched from country A to country B may be considered (in country A) as a domestic export to country B but as an import from country C in country B (if the rules of origin adopted by country B do not recognize the processing in country A as origin-conferring). This example illustrates the necessity of developing harmonized rules of origin.

Partner attribution in the case of re-exports and re-imports. Consider the case of goods originating in country A, exported, and returning to country A from country B without being substantially transformed while abroad. Some countries record such goods as re-imports from country B, while others treat them as imports from themselves, following the recording of country of origin. In the latter case, there would be a discrepancy between exports of country B to country A, which would include those goods, and imports of country A from country B, which would not.

"Through trade" operations. With the lowering of tariffs, "through trade" operations are increasingly taking place.

Consider the following example. Goods are exported from country A to country B but are shipped through country C. Instead of being declared in transit, they are declared for home use in country C and then re-exported to country B. If the exporter in country A has properly reported the country of final destination (country B), such a practice will create a discrepancy between the export data of country A and the import data of country C, as well as between the export data of country C and the import data of country B. As more and more tariffs are reduced or eliminated, this reason for discrepancy in trade statistics is likely to increase.

Unknown final destination. In some cases, the country of destination may not be known at the time of export. For some products shipped by vessel, such as petroleum and some chemicals, the ship may sail before the goods are sold and be directed to the final destination en route. In reconciliation, those kinds of transactions should be identified and attributed to the appropriate partner country.

Confidentiality. The application of confidentiality on partner or commodity level automatically generates reporting asymmetries, which should be taken into account during reconciliation studies.

Other sources of discrepancy. A considerable discrepancy between import and export statistics may arise if the information on imports is more complete than the information on exports. Differences in data-collection procedures may also contribute noticeably to data divergences (e.g., export statistics compiled using sampling techniques might be quite different from imports data derived from customs records). Reporting errors may in some instances seriously affect the comparability of data sets as well. Another source of discrepancies can be simplified reporting, under which not all data items are provided.

Adjustments to data to achieve mutually agreed sets of trade figures.

The preparation of analytical tabulations comparing import and export data for various groupings and at various levels of details helps to identify and assess the disparities. Once the analytical tables are completed, a series of adjustments may be applied (usually either based on supplementary information or derived by a series of estimates) to align data as closely as possible. Depending on the reconciliation methodology and the procedures agreed upon, adjustments are applied at either the high-level aggregates or detailed product level. Adjustments at the high-level aggregates level include adjustments for differences in commodity coverage and trade system definition, varying procedures of valuation, insurance and freight, and timing; and underreporting, country definition, indirect trade, re-exports and re-imports. In some cases, it may be necessary to investigate discrepancies in transaction-level data and make use of information supplied by declarants, trade associations and other government agencies or obtained by means of special investigations. Classification adjustments may also be applicable, especially if items shown in chapters 98 and 99 of the HS are not included in the total trade. In such cases, they should be distributed at least to the chapter level and investigated for possible reclassification and inclusion. There may be cases where discrepancies are identified but remain unresolved because of the difficulty of establishing which data are more reliable for adjustment purposes without expending unreasonable amounts of time and resources. Depending upon the information available, it may or may not be possible to estimate the effect of every identified difference and agree on an appropriate adjustment. Difficulties in the preparation of adjustments may lead to further reconciliation activities, such as analysis of the differences at a more detailed commodity level and calculation of the residual adjustment (referred to as □other□) by subtracting the adjusted export value from the agreed upon adjusted import value.

Conclusions of the reconciliation study.

The trading partners must decide at what point to consider the study completed. They must also decide on how to present the results: whether to compute a □reconciled value for each direction of trade or simply to present an explanation of why the two data sets differ. The reconciliation study may conclude with a summary statement of its major results and a set of annexes detailing specific findings. It is unlikely that all significant discrepancies can be resolved. Although reconciliations between partner countries are usually unique for each set of countries, common kinds of major adjustments have typically been applied to arrive at reconciled trade flows.

Example 1

Reconciliations studies: Example of United States and China

Both countries established a statistical working group to examine bilateral merchandise trade owing to the unusually large and growing statistical discrepancies in their respective officially published data. The working group examined the years 2000, 2004 and 2006, focusing on trade flows from China to the United States owing to the larger discrepancies in that direction. The distinction was made between trade moving directly from China to the United States, and trade moving indirectly through intermediary countries. The working group identified factors that explained the majority of the bilateral statistical discrepancies: differences in statistical territory definitions, differences in timing of recording and inclusion of re-exports in exports statistics contribute to the discrepancy. For more details, see the Report on the statistical discrepancy of merchandise trade between the United States and China‖, October 2009, Department of Commerce and Office of the United States Trade Representative, United States of America, and Ministry of Commerce, People 's Republic of China.

8. METADATA

Types of statistical data

Generally, statistical data consist of the following:

a. Microdata: data on the characteristics of individual transactions collected by customs or other sources (such as administrative records or surveys) or estimated;

b. Macrodata: data derived from microdata by grouping or aggregating them, such as total exports of goods classified in a particular HS subheading;

c. Metadata: data that define or describe the microdata, macrodata or other metadata.

Metadata

Metadata are "data about data" which enable and facilitate sharing, querying, understanding and using statistical data over the different stages of collection, compilation and dissemination, and at their various levels of aggregation (i.e., from microdata to macrodata). They encompass administrative facts about the data (who has created them and when), and definition of concepts applied, as well as description of how data were collected and processed before they were disseminated or stored in a database (see IMTS, para. 9.20). Not only are metadata important for users, but they also play a crucial role in the statistical production process, as common standards and definitions should be followed to the extent possible throughout all statistical domains in order to facilitate the linking and integration of statistical information.

Structural metadata. Structural metadata are identifiers and descriptors which are essential for discovering, organizing, retrieving and processing statistical data sets. They can be thought of as the "labels" that need to be associated to each data item in order for it to have any meaning at all. In the context of international merchandise trade statistics, structural metadata include items such as unit of measurement, time period, commodity code, identification of reporting and partner countries (country codes), identification of trade flow, etc. Structural metadata also include the information required to link data points across periods in order to obtain meaningful time series.

Reference metadata. Reference metadata are of a more general nature and "may refer to specific statistical data, to entire data collections or even to the institution that provides the data". More specifically, in the context of international merchandise trade statistics, reference metadata items include:

a. explanations of the concepts and definitions adopted and their practical implementation (e.g., coverage and valuation);

b. details on the methodologies used for the generation of the data (e.g., specification of data source(s), description of the sampling framework in the context of survey-based data, description of data validation and editing techniques, etc.); and

c. information describing the various quality dimensions of the resulting international merchandise trade statistics (e.g., relevance, accuracy, timeliness, methodological soundness, coherence and accessibility).

Statistical metadata

Statistical metadata describe or document statistical data, that is to say, microdata, macrodata or other metadata. They 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 compiled. There is a bidirectional relationship between metadata and quality. On the one hand, metadata describe the quality of statistics. On the other hand, metadata are themselves a quality component, which improves the availability and accessibility of statistical data.

Required metadata

Taking into account many types of users and uses for any given set of data, a broad spectrum of metadata requirements must be addressed. In particular, the responsible agency must make sufficient metadata available to enable the least and the most sophisticated users to readily assess data and their quality.

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

a. Reference metadata should be presented in the form of a detailed explanatory note describing the scope, coverage and quality of a data set. It should be made available electronically alongside the database or in a special publication;

b. Structural metadata should be presented as an integral part of the international merchandise trade statistics database that can be extracted together with any data item; it may be published as part of a statistical table.

Metadata and data quality

There is a bidirectional relationship between metadata and data quality. On the one hand, metadata provide details on the various quality dimensions of international merchandise trade statistics, as stated in the previous paragraph. On the other hand, the availability of adequate metadata to users is in itself a quality component through its role as an indicator of the accessibility of international merchandise trade statistics (see IMTS 2010, paras. 9.17 and 9.21). Compilers should aim to provide users with all the metadata required to understand both the strengths and the limitations of the international merchandise trade statistics they produce, documenting in a timely manner all methodological aspects underlying the data that are relevant for their proper use and interpretation (e.g., definitions, classifications, scope, confidentiality issues, etc.).

Purposes of international merchandise trade statistics metadata

The primary purpose of metadata is to help the users of international merchandise trade statistics to interpret, understand and analyse the data. International merchandise trade statistics metadata should help users to transform statistical data into information. Metadata is also crucial for conducting bilateral reconciliation studies, and allows countries to compare their practices and to learn from each other.

It is recommended that countries view the development of metadata as a high priority and consider their dissemination an integral part of the dissemination of international merchandise trade statistics. Moreover, it is recommended that this is done in compliance with the approach adopted by a given country to metadata across all areas of economic statistics. Countries are encouraged to take advantage of the metadata standards proposed by various international organizations while developing their metadata in general and trade metadata in particular.

Reference metadata items for international merchandise trade statistics

It is recommended that countries at least cover the following categories of reference metadata for their international merchandise trade statistics:

a. Legal framework and institutional arrangements (e.g., references to relevant laws and regulations, role of all institutions involved in compilation, etc.);

b. Underlying concepts and definitions (e.g., definition of statistical territory and trade system used, definition of statistical value, scope of international merchandise trade statistic and their relationship to national accounts and balance-of-payments statistics, etc.);

c. Description of data sources used and methodology of data integration (e.g., customs declarations and related customs records, non-customs administrative sources, enterprise surveys, and a brief description of how data obtained from these sources are merged to create an integrated data set);

d. Description of data-collection and data-processing procedures (e.g., frequency of data collection, description of specific procedures used for data collection, validation, editing, aggregation, etc.);

e. Estimation methods (e.g., estimation of value of trade below customs and statistical thresholds, CIF-FOB adjustments, estimation of quantities, etc.);

f. Dissemination policy (e.g., release and revision schedules);

g. Additional explanations and footnotes concerning the data, as required (e.g., explanatory notes on revisions, breaks in series, application of confidentiality rules, treatment of special categories of goods, etc.);

h. Quality reporting.

Structural metadata items for international merchandise trade statistics

The following are typical structural metadata items that belong to the trade values and quantities in each international merchandise trade statistics dataset:

a. Reporting country: identification of the country that reported the given data set;

b. Reference period: identification of the specific month, quarter, year, etc.;

c. Trade flow: whether the data refer to exports, imports, re-exports, etc.;

d. Commodity code: identification of the commodity or commodity group;

e. Commodity description: stand-alone descriptors of the commodity codes at any level of aggregation;

f. Commodity classification: name of the classification used to report the data (e.g., HS2007, SITC Rev. 4, national classification, etc.);

g. Partner country or region: identification of the partner country or region;

h. Country or region classification (e.g., alpha-2 or alpha-3 ISO country codes; United Nations standard country or area codes for statistical use,255 etc.), including information on the country-composition of geographical regions;

i. Currency unit: identification of the currency unit (e.g., national currency, United States dollars, etc.) in which trade values are expressed;

j. Mode of transport;

k. Weight unit: identification of the weight unit in which trade quantity data (net weight data) are expressed;

l. Supplementary quantity unit: identification of the unit of measurement, by HS code, in which supplementary trade quantity data are expressed;

m. Valuation: identification of whether trade values are expressed in FOB or CIF terms;

n. Custom procedure code (or applicable transaction code).

Institutional arrangements for metadata compilation

To reduce the burden associated with trade metadata projects, it is good practice for compilers to closely cooperate with the specific units responsible for ensuring within the national statistical system that metadata are produced, that they adhere to a standard format and that they are properly maintained and updated.

Access to metadata

Compilers of international merchandise trade statistics should make every effort to ensure that users have ready access to metadata through multiple dissemination channels, both in printed and in electronic format (whereby Internet dissemination plays a key role). As a general rule, as metadata is considered as having a high public good component, their on-line dissemination should be free of charge, regardless of whether the international merchandise trade statistics they describe are disseminated for a fee according to the compiling organization's policies.

Example 2

Country example: Brazil

Metadata pertaining to Brazil's international merchandise trade statistics are available in Portuguese, English and Spanish through the online AliceWeb2 system (http://aliceweb2.mdic.gov.br/).

Concepts and definitions in respect of the main variables are given directly below:

a. *Export:* corresponds to a good shipped to the outside, without return;

b. *Import:* corresponds to the entry of a good originating from outside, without return;

c. *Commodity*: for the purpose of the classification of goods, Brazil has, since 1996, used the common classification of Mercosur (NCM), which is also used by other countries of the block (Argentina, Paraguay and Uruguay), which is based on the Harmonized Commodity Description and Coding System;

d. *Country of destination (for exports)*: for the purpose of dissemination of export statistics, the country of destination is the last country known at the time of the exportation, to which goods are to be delivered, irrespective of where they have been initially dispatched to and whether or not, on their way to the last country, they are subject to any commercial transactions or other operations that change their legal status;

e. *Country of origin (for imports)*: for the purpose of disseminating statistics on imports, the country of origin is the country where the agricultural products were grown, the minerals extracted or the manufactured goods produced in whole or in part. In the case of manufactured goods, the country of origin is where the last phase of processing was completed and the product adopted its final form (a concept defined by the Kyoto Convention);

f. *Economic bloc*: countries are grouped by economic blocks following the formation of geo-economic regions and international agreements. A country may be part of more than one economic bloc;

g. *State (unit of the federation)*: for the purpose of disseminating statistics on exports, the State is the unit of the Federation where agricultural products were grown, minerals were extracted or manufactured goods were produced in whole or in part. In the case of manufactures, "State"‖ is the State that has completed the last stage of the manufacturing process in which the product takes its final form (concept of origin). In the case of imports, "State"‖ is defined as the State of the tax residence of the importer;

h. *Mode of transport*: in the case of exports, this refers to the means of transportation used from the last place from which the goods are shipped abroad. In the case of imports, it refers to the means of transportation used at the first point of entry of goods into national territory. In accordance with the framework of the Mercosur countries, Brazil records the following modes of transport: sea, river, lake, air, postal consignments, mail or courier shipments, railway, road, pipelines, transmission line (cables) and self-propelled goods;

i. *Port*: for exports, this refers to the port or airport where the shipment of goods takes place, or the last place from which the good leaves the country. In the case of imports, it is the port or airport where the goods are unloaded or the first place at which the goods enter the country.

References

Eurostat, European Statistics Code of Practice for the national and community authorities, 2011.

Eurostat, ESS Handbook for Quality Reports, 2014.

United Nations, "International Merchandise Trade Statistics: Concepts and Definitions 2010", New York, 2011.

United Nations, "International Merchandise Trade Statistics: Compilers Manual, Revision 1 (IMTS 2010-CM)", New York, 2013.

HOW TO CONTACT US

Human Resources Development
Section/TrainForTrade
Knowledge Development Branch
Division on Technology and Logistics
United Nations Conference on Trade
and Development
Palais des Nations,
8-14, Av. de la Paix,
1211 Geneva 10, Switzerland

T: +41 22 917 5512
F: +41 22 917 0050
E-mail: trainfortrade@unctad.org
Web: tft.unctad.org

UNITED NATIONS
UNCTAD