

In the Construction of an International Agricultural Data Quality Assessment Framework (ADQAF)

Flavia Kyeyago Ouma, Uganda Bureau of Statistics, fwooma@yahoo.com

ESK Muwanga Zake (PHD), Uganda Bureau of Statistics, muwangazake@hotmail.com

Seth Mayinza , Uganda Bureau of Statistics , smayinza@hotmail.com

Abstract

The fact that there is still no ADQAF yet at global or the national level until now is not very surprising as the DQAF for the overall official statistics is still under development and not completed yet; the situation itself however is intolerable with the urgency and needs pressed by users for timely and better quality agricultural statistics as reflected in the Global strategy for Agricultural statistics. As specialized area and subset of the overall official statistics, data collection, compilation and dissemination for agricultural statistics are often carried out by Statisticians in the Ministries of Agriculture rather than the National Statistics Offices (NSOs) an ADQAF is supposed to be simpler and easier than the DQAF for the overall official statistics; and this ADQAF should be a pioneer rather than a lagger in the development of the overall DQAF.

In this paper, a framework for the ADQAF is proposed based on the principles guiding rules and methodology that have been developed for the overall DQAF and CountrySTAT. The proposed ADQAF will be tested from both the global and national perspectives, especially the previous experience in compiling agricultural metadata in different continents and countries. The field data and metadata collected for agricultural statistics in Uganda is used as a country case study to test out the proposed ADQAF. Such an ADQAF will be valuable and useful to provide a starting point for further discussion and future development in this area. Related issues , problems and difficulties will be discussed.

Key words: DQAF, ADQAF; Global Strategy for Agricultural Statistics; Data; Metadata; and ICAS-5

1.0 Introduction

Quality is generally accepted as “fitness for purpose” amongst most National Statistics Offices. Fitness for purpose implies an assessment of an output, with specific reference to its intended objectives or aims. Quality is therefore a multidimensional concept which does not only include the accuracy of statistics, but also stretches to include other aspects such as relevance and interpretability. Most commonly, the term "Data Quality" is described as data that is "Fit-for-use" (Wang & Strong, 1996), which implies that it is relative, as data considered appropriate for one use may not possess sufficient attributes for another use (Tayi & Ballou, 1998). Data quality is a multidimensional concept (Klein & Rossin, 1999) as data itself

is multidimensional (Fox & Levitin, 1996 ; Juran & Godfrey 1999) . Modern definitions of data quality have a wider frame of reference and many more attributes than the obvious characteristics of accuracy.

Redman (1999) comes to the following definition based on Joseph Juran (Juran & Godfrey 1999) “Data are of high quality if they are fit for their intended uses in operations, decision-making, and planning. Data are fit for use if they are free of defects and possess desired features (Redman, 2001)”. Tayi and Ballou (Tayi & Ballou, 1998) concur, noting that data considered appropriate for one use may not possess sufficient quality for another use, citing the trend towards multiple uses of data through data warehouses.

2.0 Data Quality Frameworks

To develop and implement a quality framework tailor-made to the objectives of the Agricultural statistical function, there is need for a stock taking exercise of the existing frameworks that have been developed within the statistical functions by National Statistical bodies and international organisations. However, due to differences in the institutional environments the models developed differ somewhat with regard to their stakeholders, definitions of quality and scope. A number of these frameworks need to be assessed in order to determine whether they contain elements that could serve the Agricultural statistical purposes. Significant work in the area of quality have defined various extensive frameworks to review systems within institutions and different working environments. The frameworks all seek to assess areas where poor quality workmanship or inefficiencies that may reduce the profitability of the working institution and established environment. At its most basic, a data quality framework is a tool for the assessment of data quality within an organisation (Wang, Strong & Guarasciou, 1996). The framework can go beyond the individual elements of data quality assessment, becoming integrated within the processes of the organisation. Willshire and Meyen (Willshire & Meyen , 1997) describe data quality frameworks as ‘a vehicle that an organisation can use to define a model of its data environment, identify relevant data quality attributes, analyse data quality attributes in their current or future context, and provide guidance for data quality improvement’. Eppler and Wittig (Eppler & Wittig , 2000) add that a framework should not only evaluate, but also provide a scheme to analyse and solve data quality problems with their proactive management

3.0 The Construction of the ADQAF and synergies with existing quality models

The construction of the ADQAF needs to be built using the existing quality frameworks that have been adapted into the statistical environment and operational features. The major focus of this paper will be the IMF data quality framework the Data Quality Assessment Framework and the CountryStat or FAO Stat Data Quality Framework that are analyzed below.

3.1 Data quality Assessment framework (DQAF)

The DQAF was developed by the IMF, in collaboration with the World Bank, as a methodology for assessing data quality that brings together best practices and internationally accepted concepts and definitions in statistics, including those of the United Nations Fundamental Principles of Official Statistics and the General Data Dissemination System (GDDS). It facilitates a comprehensive view of data quality, one that recognizes interrelations, including tradeoffs, among elements of quality and allows emphasis to vary across data categories and uses/users (DQAF Factsheet , 2003).

3.1.1 Content of the Framework

The DQAF’s coverage of governance, processes, and products is organized around a set of prerequisites and six dimensions of data quality—assurances of integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility. For each dimension, the DQAF identifies 3-5 elements of good practice, and for each element, several relevant indicators. (See Annex 1) Further, in a cascading structure, more detail and more concreteness tailored to the dataset are provided by focal issues and key points.

Table 1: The dimensions of the Data Quality Assessment Framework

Dimensions
1. Prerequisites of Quality <i>The essentials to be in place before quality is implemented</i>
2. Integrity <i>The principle of objectivity in the collection, processing, and dissemination of statistics is firmly adhered to.</i>
3. Methodological soundness <i>The methodological basis for the statistics follows internationally accepted standards, guidelines, or good practices.</i>
4. Accuracy and reliability <i>Source data and statistical techniques are sound and statistical outputs sufficiently portray reality.</i>
5. Serviceability <i>Statistics, with adequate periodicity and timeliness, are consistent and follow a predictable revisions policy.</i>
6. Accessibility <i>Data and metadata are easily available and assistance to users is adequate.</i>

3.2 CountrySTAT Framework

The CountrySTAT is a web-based system that serves as a statistical framework for the organization, harmonization and synchronization of data collection and dissemination in food and agriculture for analysis and policy-making. It is a customized country-specific version of the FAOSTAT. The CountrySTAT provides tools for data exchange, and promotes the use of data and metadata standards. It serves as the gateway for food and agricultural statistics, including metadata for member countries. Just as importantly, it served as a source of country-level data for FAOSTAT.

CountrySTAT is a scaled down version of the FAOSTAT Statistical System and provides a mechanism for countries to consolidate statistical data required for national food and agricultural policy analysis and development in one database. The CountrySTAT database includes detailed statistical metadata and data quality indicators into the same format as the FAOSTAT Statistical System and provides assessment at the source of the data level. The CountrySTAT statistical system provides a unique opportunity to review and enhance the whole statistical process (see Appendix 3) from the collection of data at national level until the dissemination of data and metadata to the users. Data is collected on an annual basis from these countries. A summary data quality assessment of data obtained from each country is to be completed regularly with feedback provided if necessary.

3.2.1 Content of the CountrySTAT Framework

FAO has included standard data quality dimensions and indicators where possible in the FAOSTAT Statistical System. A list of FAOSTAT quality indicators is included in Annex 2. These data quality indicators are calculated at key points the source of data, the country statistics office and the FAO statistics division where possible. Most data quality indicators include some objective measurement on the data domain or country performance. The FAO Statistical system supplements these objective measurements with a subjective review and adjustment where necessary

3.2.2 Data Quality Dimensions of the FAOSTAT / CountrySTAT

The CountrySTAT framework is organized into a set of six dimensions of data quality—Relevance and Completeness, Timeliness, Accessibility and Clarity, Comparability, Coherence and Subjective. For each dimension, the DQAF identifies 3-5 elements of good practice, and for each element, several relevant indicators. (See Annex 1) Further, in a cascading structure, more detail and more concreteness tailored to the dataset are provided by focal issues and key points.

Table 2: Data Quality Dimensions of the FAOSTAT / CountrySTAT

Data Quality Dimensions	Assessment	
	Objective	Subjective
Relevance and completeness	Yes	Yes
Timeliness	Yes	Yes
Accessibility and Clarity	Yes	Yes
Comparability	Yes	Yes
Coherence	Yes	Yes
Subjective	Yes	Yes

Accuracy and reliability are the corner-stones of data quality. Without accuracy and reliability there is little further need to measure and monitor other dimensions of quality. Accuracy and reliability should be primarily assessed and monitored at the statistical element and item levels and summarized at the statistical domain level. Reporting on accuracy and reliability should be undertaken at both points q₂ and q₃ in the statistical dissemination process. The domain, element and item level accuracy assessments and reporting should then be included in the second essential component of data quality at dissemination – **Transparency and Metadata**. Without sufficient statistical metadata there is little understanding of the data item. Sufficient statistical metadata covering the standard items (concepts, definitions, source, etc) is required at both points q₂ and q₃ in the statistical dissemination process. Sufficient statistical metadata is also required at the domain and element levels for points q₂ and q₃ in the statistical dissemination process.

The third essential component is the use of **Standard Classifications**. When statistical data do not use

or correspond to standard (common or international) classifications, they are isolated and of little use outside the specific data domains. Thus, the use of standard classifications is essential for points q₂ and q₃ in the statistical process. It is also essential for statistical domains (and by definition elements and items within those domains) to apply standard classifications. The fourth essential data quality for the data is **Timeliness**. Data that does not reflect the reality of the situation that is being measured will likely have a negative impact on analysis and subsequent policy development that it is used for. Data does not need to come from the most recent data time-period, but it needs to be refreshed and available as appropriate. It is also essential for statistical domains (and by definition elements and items within those domains) to have data that is up-to-date. Having up-to-date data within a domain/element is as important as having current data at the dissemination point of national and international agencies (points q₂ and q₃) as at those data points it is used in a cross-sectional fashion and individual data that is out-of-date at the element level would negatively impact analysis and policy formulation.

The fifth essential component is the concept of **Comparable Data Domains**. Data can be isolated by

not using common or international classifications, but it can also be isolated by the lack of integration.

Integration can be in a physical form such as the data stored in formats which are not common or in databases that cannot communicate with other databases. Data can also be isolated in a conceptual context: that is data without an integrated framework.. The sixth essential component of data quality at dissemination is the concept of **Accessibility**. Without suitable methods of dissemination, data loses substantial quality. By integrating data whether in a physical form or a virtual form data also gains quality by becoming more accessible.

Data quality Stamp in FAOSTAT

One of the quality practices being developed by the FAO Statistics Division is a “Quality Stamp” to be attached to every dataset included in the FAOSTAT statistical system. The FAOSTAT Statistical Quality Stamp will inform users that the data has met the minimum requirements for FAOSTAT and has undergone various checks to ensure quality.

Key ingredients in statistical data integration - methodologies and data standards

The CountryStat system mainly relies on the inclusion of data from diverse sources to address the needs of organisations. By conforming to common data concepts, definitions, codes and methodologies.

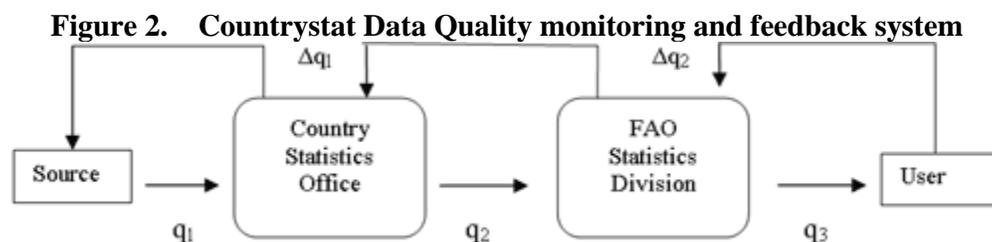
Metadata

Metadata enhances understanding of any given data item within the system by documenting its definition, history of its values, methodology used in its collection, national contacts, etc. Statistical metadata is one of the keys for the integration of national and international statistical data in CountryStat. This information is also useful to statisticians who compile, validate and analyze the data, as well as to the users, both internal and external to FAO, who access the data.. The statistical metadata system covers: concepts and definitions; classifications; symbols and units; explanatory notes; statistical methodology; data dissemination; data and metadata quality.

Data Quality Integration

CountryStat data quality ultimately depends on the quality of the data reported by member countries and processed by FAO. The standard data quality evaluation and monitoring processes covers both data submitted by countries and data and then processed within FAO. On the basis of the quality descriptions, a list of process-oriented data quality indicators is already provided to every country and on the country websites.

A predefined format (prepared by FAOSTAT) was identified with minimum and standardized set of information items to be included. Data quality - including value added and performance - is monitored at three different points of the statistical process. Data quality problems identified at any point can be addressed directly, often in real time. Feedback loops provide a mechanism for improving data quality (see Figure 2).



Quality indicators are compatible with those used by other agencies in the International Statistical System, adjusted to the specifics of agricultural data. FAO follows closely the work being carried out by other international statistical offices on data quality monitoring.

Data dissemination

The CountryStat data portal (www.countrystat.org) is the most visible worldwide product offering users (policy formulators, decision-makers, statisticians, journalists, etc) a fast and reliable way to access different countries with statistical information in the areas of food and agriculture. CountryStat data and metadata can be downloaded in different formats for further use, and also provides users with their preferred tools for analysis. Key country indicators in the form of precompiled reports will provide a useful vehicle to assess a situation at a glance.

The minimum data requirements for CountryStat include:

- Global coverage; data provided on the CountryStat should cover both national and sub national
- Regular, committed sustainable data collection activity;
- Time-series data;
- Data quality assessment performed;
- Statistical metadata available; this meta data should provide for the source of data , title of tables, definitions of labels, measurement units recognized internationally, explanation of symbols, copyright information, contacts of information etc.
- Data is edited and validated.
- Additional metadata on the methods used for data collection, revision, calculation and estimation of the statistics
- Description of errors sources and accuracy of statistics
- The data should be comparable with data from all other counties on the countrystat website

Table 3: Comparison of in DQAF and the Countrystat Framework

DQAF	CountrySTAT Framework
1. <i>Prerequisites of Quality</i>	1. <i>Relevance and completeness</i>
2. <i>Integrity</i>	2. <i>Timeliness</i>
3. <i>Methodological soundness</i>	3. <i>Accessibility and Clarity</i>
4. <i>Accuracy and reliability</i>	4. <i>Comparability</i>
5. <i>Serviceability</i>	5. <i>Coherence</i>
6. <i>Accessibility</i>	6. <i>Subjective</i>

Note: The double arrow shows that there is a link between the DQAF dimension with those of the Countrystat Framework

There is a relationship between DQAF and Countrystat Framework, in that

- Both frameworks agree that for any data to be used it must be accompanied with metadata to ease understanding of data and that the data should be easily available to the users. Countrystat Framework ensures that clarity of data is put across for all the data provided
- Both frameworks compliment that Accuracy and reliability is covered to ensure source of data and statistical techniques are sound and outputs portray reality. Though the Countrystat framework that the data should be relevant and complete meaning that the sources of data , methodologies, assumptions, errors etc should be provided for to ensure completeness. While the DQAF brings out methodological soundness as a stand alone dimension that should provide for internationally acceptable standards, guidelines and good practices.
- Both frameworks advocates for timeliness, periodicity, consistence under serviceability for the DQAF and under Timeliness and Coherence for the Countrystat Framework. The data should therefore be timely, revisable and updated regularly.

However the frameworks have some components they do not agree with like:

- The issue of comparability in the Countrystat framework, it is a dimension that provides for the ability of the users to compare data across various countrystat systems of different countries. It is spelt out clearly and in addition it promotes for the dimension of subjectivity, where by the data is biased towards the promotion of agricultural growth and functionality.
- The DQAF has two components like Prerequisites that must be put in place before quality is adhered to like the legal issues, the structural aspects etc and Integrity which calls for objectivity in the procedures and steps of data production which should be firmly adhered to.

The CountrySTAT quality standards therefore provide building blocks for the construction of the ADQAF.

Uganda's Experience in the implementation of Countrystat

Uganda launched their countrystat website in 2010, this has been in the implementation phases since 2007. Training amongst the Ministries, Departments and statistical Agencies that collect Statistical data on how to use the Countrystat system was done in April, and several stakeholder meetings have been held in a bid to design steps through which the data is received from the institutions. The MDAs have appreciated the system and are able to download Agricultural data with ease. All the data currently posted on the Countrystat website is received by the Countrystat Secretariat (whose head quarters is Uganda Bureau of Statistics- UBOS) from the different Agricultural data providing agencies, the data is then checked for all the meta data requirements, that includes units, time, source of data, frequency of data, contacts for the data providers etc. The data is then uploaded on the website by the CountrySTAT administrator who manages the website and is a member of the CountrySTAT secretariat. All the uploaded data is checked for any missing quality requirements by the FAOSTAT team in FAO. If any data is missing the data is forwarded back to the secretariat for scrutiny and adjustment. Should the CountrySTAT secretariat not be able to correct the data set, its then sent back to the original institution to edit and the cycle goes on. The countrystat Uganda system is hosted at www.countrystat.org/uga and currently 32 data sets have been uploaded on the system. More data is uploaded as it is provided.

Currently Uganda has not provided for data on all the domains of the Countrystat system because of various reasons. The exercise does not come with a lot of ease, some challenges have been met along the way in both meeting the quality requirements and collecting data for Countrystat Challenges of meeting the CountryStat quality requirements.

- The sizes of the sample selection made in Uganda cannot provide for all the data required nationally and sub nationally, it depends on estimation. There is need for increasing the sample sizes
- The lack or poor documentation procedures some meta data is still not properly registered by the data collecting institutions leading to lack of meta data.
- A number of units of measurement used in Uganda are not recognized internationally leading to poor data conversions. These include bags/sacks, heaps, tins, mwigo, yard etc.
- Some of the major crops in Uganda are measured in seasons and it has proven very difficult to collect data seasonally over the years. For example cassava, maize, ground nuts, beans are grown in different seasons throughout the year.
- The constant change in administrative areas like the number of districts from 56 in 2002 to over 112 in 2010 cannot allow for making district estimates and thus meeting of quality requirements.

Challenges of finding data for the CountryStat Uganda website

Currently data is still missing in many areas and domains, because of various reasons:

- Uganda collects agricultural data but for which their indicators are not provided in the predefined indicators in the CountryStat system. For example the recent Livestock Census, the data collected does not have indicators like number of camels, donkeys, sheep in the country that feature in the current CountrySTAT system
- Because of the wars, poor record and data management techniques data for past years is not easy to come across hence creating gaps in the Agricultural data sets. Uganda lacks records on crop production as far back as 1960s. This makes our agricultural data hard to compare with other countries.
- Some data has only been collected at national level due to the limitation in sample size and financial implications of the data collection process hence there is no sub national data. For example data on crop production, area planted is for only a few selected administrative areas. The same goes for the Consumer Price Index, the figures are based on 6 regional areas where the data is collected and cannot really reflect the sub national figures.
- The delay in releasing data by the primary institutions gathering the information poses challenges in having data ready for immediate use on the CountryStat website

- Some institutions collecting agricultural data may not be willing to release it causing gaps.
- There is lack of harmonized agricultural indicators causing problems in data interpretation and clarity for agricultural data. In a bid to increase harmonization of indicators amongst the data providers UBOS in collaboration with these institutions has helped develop compendiums and meta data definition dictionary to be used across the National Statistical System (NSS).
- CountryStat has tried to bring some collaboration amongst the agricultural data collecting institutions but it is still challenge that needs to be addressed.

Having compared both DQAF and Countrystat Framework and looked at the challenges of implementing the CountrySTAT in Uganda, the proposed ADQAF can be built using all the combination of the six dimensions of Countrystat and the DQAF but it is critical that each participating institution should put in place prerequisites of quality and strictly adhere to them so as to guide the successful implementation the ADQAF.

Proposed dimensions of the ADQAF

DQAF	Explanation
1. Prerequisites of Quality	<i>The essentials to be in place before quality is implemented</i>
2. <i>Relevance and Completeness</i>	<i>The data that is provided should be relevant to the country , collected following the statistical standards, with complete information (metadata)</i>
3. Methodological soundness	
4. Accuracy , reliability and clarity	<i>Source data and statistical techniques are sound and statistical outputs sufficiently portray reality and are clear to understand</i>
5. Serviceability and coherence	<i>Statistics, with adequate periodicity and timeliness, are consistent and follow a predictable revisions policy.</i>
6. Comparability and Subjective	Statistics provided should be biased towards the Agricultural cause and comparable across countries and institutions
7. Accessibility	<i>Data and metadata are easily available and assistance to users is adequate.</i>

In light of the above therefore for the Agricultural Data Quality Framework to be established it should be able to meet certain prerequisites. These may include

- The framework should be able to show how data quality of Agriculture statistical data can be improved by stating what actions that makes the biggest increase in data quality.
- The framework should be based on the great Agriculture statistical need for data quality.
- It should provide measures so that the change in data quality can be tracked over time.
- It should prioritize the dimensions from the establishing institutions' point of view.
- It should also allow for comparison of the actual level of data quality with the needed levels and .
- The framework should be time-efficient and point out where to focus the efforts, both for further investigations and for improvements of the Agricultural statistical data.
- The framework should take into consideration the users experience as a basis for measuring.
- The framework should have flexibility in the type of data it may require in the beginning and allows for data improvement depending on the capacities and data availability of the different institutions

Conclusion

Defining a Data Quality Framework for Agricultural statistics globally is a complex and multi-faceted issue made even more difficult by factors like the presence of agricultural statistical data from various data sources, data handling capacities in the various statistical entities, policy and financial issues. This paper attempts to evaluate the experience of countrystat in Uganda and suggests possible dimensions which can be applied to measure the quality of Agricultural data and hence the development of ADQAF.

References

- Klein, B. and D.F. Rossin, (1999). Data errors in neural network and linear regression models: An experimental comparison. *Data Quality*, 5(1): p. 25.
- Fox, C.J. and A.V. Levitin, (1996). Data and Data Quality. *Encyclopedia of Library and Information Science*, 57(Suppliment 20): p. 100-122.
- Juran, J.M. and A.B. Godfrey, (1999). *Juran's Quality Handbook*. 5 ed., New York: McGraw-Hill.
- Klein, B. and D.F. Rossin, (1999). Data errors in neural network and linear regression models: An experimental comparison. *Data Quality*, 5(1): p. 25
- Redman, T.C., *Data Quality* (2001). The Field Guide, Boston: Digital Press. 241.
- Tayi, G.K. and D.P. Ballou, (1998). Examining data quality. *Communications of the ACM*, 41(2): p. 54-57.
- Wang, R.Y., D.M. Strong, and L.M. Guarascio, (1996) Beyond Accuracy: What data quality means to data consumers. *Total Data Quality Management Programme*.
- Brown, S., (2002). Data quality: Relatively critical and critically relative. *DM Review*, 2002.
- Redman, T.C., (2001). *Data Quality. The Field Guide*. 2001, Boston: Digital Press. 241.
- Willshire, M.J. and D. Meyen, (1997). A process for improving data quality. *Data Quality*. 3(1).
- Eppler, M.J. and D. Witting, (2000). A Review of Information Quality Frameworks from the Last Ten Years. *Net Academy*.
- Wang, R.Y. & Strong, D.M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, Spring, 5–33.
- DQRS Factsheet (2003) <http://dsbb.imf.org/> accessed on 28th August 2010
- CountrySTAT*, 2004. Food and Agriculture Organization of the United Nations, <http://www.fao.org/faostat/countrystat_en.asp
- Robert Mayo (2004). Revising the Conceptual Statistical Frameworks, Underlying Statistical Methodologies, Database System and Processes, Statistics FAOSTAT2 Project Division, Food and Agriculture Organization of the United Nations <http://unstats.un.org/unsd/accsub/2004docs-CDQIO/1-FAO.pdf>
- Muwanga Zaake (2009) First panorama Report http://www.countrystat.org/country/UGA/documents/docs/Panorama_Report_I-UGA.pdf

Annex 1: DQAF—Generic Framework (July 2003)

Quality Dimensions	Elements	Indicators
<p>1. Prerequisites of Quality</p>	<p>Legal and institutional environment—<i>The environment is supportive of statistics.</i></p> <p>Resources—<i>Resources are commensurate with needs of statistical programs.</i></p> <p>Relevance—<i>Statistics cover relevant information on the subject field.</i></p> <p>Other quality management—<i>Quality is a cornerstone of statistical work.</i></p>	<ul style="list-style-type: none"> ▪ The responsibility for collecting, processing, and disseminating the statistics is clearly specified. ▪ Data sharing and coordination among data-producing agencies are adequate. ▪ Individual reporters' data are to be kept confidential and used for statistical purposes only. ▪ Statistical reporting is ensured through legal mandate and/or measures to encourage response. ▪ Staff, facilities, computing resources, and financing are commensurate with statistical programs. ▪ Measures to ensure efficient use of resources are implemented. ▪ The relevance and practical utility of existing statistics in meeting users' needs are monitored. ▪ Processes are in place to focus on quality. ▪ Processes are in place to monitor the quality of the statistical program. ▪ Processes are in place to

		deal with quality considerations in planning the statistical program.
2. Integrity	<ul style="list-style-type: none"> ▪ Professionalism ▪ Transparency ▪ Ethical standards 	<ul style="list-style-type: none"> ▪ Statistics are produced on an impartial basis. ▪ Choices of sources and statistical techniques as well as decisions about dissemination are informed solely by statistical considerations. ▪ The appropriate statistical entity is entitled to comment on erroneous interpretation and misuse of statistics. ▪ The terms and conditions under which statistics are collected, processed, and disseminated are available to the public. ▪ Internal governmental access to statistics prior to their release is publicly identified. ▪ Products of statistical agencies/units are clearly identified as ▪ such. ▪ Advance notice is given of major changes in methodology, source data, and statistical techniques. ▪ Guidelines for staff behavior are in place and are well known to the staff.
3. Methodological	<ul style="list-style-type: none"> ▪ Concepts and definitions 	<ul style="list-style-type: none"> ▪ The overall structure in

<p>soundness</p> <p><i>The methodological basis for the statistics follows internationally accepted standards, guidelines, or good practices.</i></p>	<ul style="list-style-type: none"> ▪ Scope ▪ Classifications ▪ Basis for recording: accounting rules and valuation principles 	<p>terms of concepts and definitions follows internationally accepted standards, guidelines, or good practices.</p> <ul style="list-style-type: none"> ▪ The scope is broadly consistent with internationally accepted standards, guidelines, or good practices. ▪ Classification/sectorization systems used are broadly consistent with internationally accepted standards, guidelines, or good practices. ▪ 2.4.1 Market prices are used to value flows and stocks. ▪ 2.4.2 Recording is done on an accrual basis. ▪ 2.4.3 Grossing/netting procedures are broadly consistent with internationally accepted standards, guidelines, or good practices.
<p>4. Accuracy and reliability</p> <p><i>Source data and statistical techniques are sound and statistical outputs sufficiently portray reality.</i></p>	<ul style="list-style-type: none"> ▪ <u>Assessment of source data</u>— Data Sources are regularly assessed. ▪ <u>Statistical techniques</u>—Statistical techniques employed conform to sound statistical procedures. ▪ <u>Assessment and validation of intermediate data and statistical outputs</u>—Intermediate results 	<p>Source data—including censuses, sample surveys, and administrative records—are routinely assessed, e.g., for coverage, sample error, response error, and nonsampling error; the results of the assessments are</p>

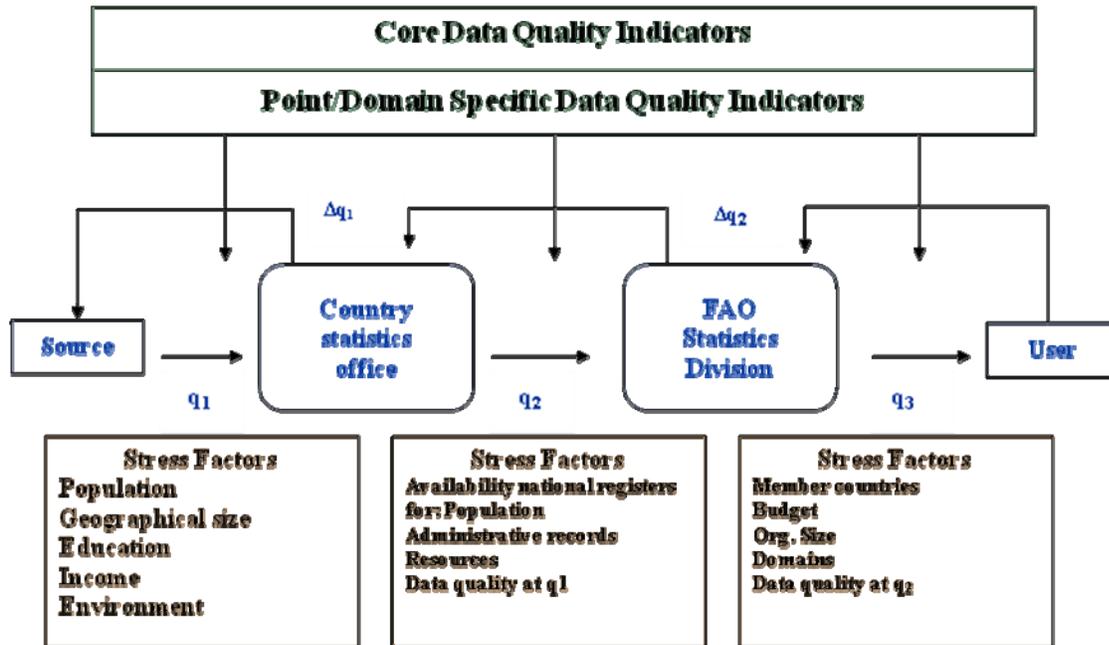
	<p>and statistical outputs are regularly assessed and validated.</p> <ul style="list-style-type: none"> ▪ <u>Revisions</u>, as a gauge of reliability, are tracked and mined for the information they may provide 	<p>monitored and made available to guide statistical processes.</p> <ul style="list-style-type: none"> ▪ Data compilation employs sound statistical techniques to deal with data sources. ▪ Other statistical procedures (e.g., data adjustments and transformations, and statistical analysis) employ sound statistical techniques. <p>Intermediate results are validated against other information where applicable.</p> <ul style="list-style-type: none"> ▪ Statistical discrepancies in intermediate data are assessed and investigated. ▪ Statistical discrepancies and other potential indicators or problems in statistical outputs are investigated. ▪ Studies and analyses of revisions are carried out routinely and used internally to inform statistical processes
<p>5. Serviceability <i>Statistics, with adequate periodicity and timeliness, are consistent and follow a predictable revisions policy.</i></p>	<ul style="list-style-type: none"> ▪ Periodicity and timeliness—Periodicity and timeliness follow internationally accepted dissemination standards. ▪ Consistency—Statistics are 	<ul style="list-style-type: none"> ▪ Periodicity follows dissemination standards. ▪ Timeliness follows dissemination standards. ▪ Statistics are consistent

	<p>consistent within the dataset, over time, and with major datasets.</p> <ul style="list-style-type: none"> ▪ Revision policy and practice—Data revisions follow a regular and publicized procedure. 	<p>within the dataset.</p> <ul style="list-style-type: none"> ▪ Statistics are consistent or reconcilable over a reasonable period of time. ▪ Statistics are consistent or reconcilable with those obtained through other data sources and/or statistical frameworks. ▪ Revisions follow a regular and transparent schedule. ▪ Preliminary and/or revised data are clearly identified. ▪ Studies and analyses of revisions are made public (see also 3.5.1).
<p>6. Accessibility <i>Data and metadata are easily available and assistance to users is adequate.</i></p>	<ul style="list-style-type: none"> ▪ <u>Data accessibility</u>: Statistics should be presented in clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis ▪ <u>Metadata accessibility</u>—Up-to-date and pertinent metadata are made available. ▪ <u>Assistance to users</u>—Prompt and knowledgeable support service is available. 	<p>Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons (layout and clarity of text, tables, and charts).</p> <ul style="list-style-type: none"> ▪ Dissemination media and format are adequate. ▪ Statistics are released on a preannounced schedule. ▪ Statistics are made available to all users at the same time. ▪ Statistics not routinely disseminated are made available upon request. ▪ Documentation on concepts, scope, classifications, basis of

		<p>recording, data sources, and statistical techniques is available, and differences from internationally accepted standards, guidelines, or good practices are annotated.</p> <ul style="list-style-type: none">▪ Levels of detail are adapted to the needs of the intended audience.▪ Contact points for each subject field are publicized.▪ Catalogs of publications, documents, and other services, including information on any changes, are widely available.
--	--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Annex 2: FAO Statistical Data Quality Framework

FAO Statistical Data Quality Framework - stress factors and data quality indicators at points q1, q2 and q3



Annex 3: FAOSTAT Data Flow

