Improvements on the statistical quality practices at the Brazilian Institute of Geography and Statistics

Zélia Magalhães Bianchini*
Brazilian Institute of Geography and Statistics, Rio de Janeiro, Brazil – zelia.bianchini@ibge.gov.br

Maria Luiza Barcellos Zacharias
Brazilian Institute of Geography and Statistics, Rio de Janeiro, Brazil – maria.zacharias@ibge.gov.br

Abstract

In the last years, IBGE, the Brazilian Institute of Geography and Statistics, has been focusing on different aspects of the quality of statistical production, while taking into consideration the best practices and the principles of official statistics. The aim of this paper is to describe some of the statistical quality management practices implemented recently. Although a lot of improvements have been made, much still remains to be done. In this sense, a set of procedures to enhance the statistical quality management system is in progress and will be also presented.

Keywords: Statistical quality, Code or Practice, Metadata, Fundamental Principles of Official Statistics

1. Introduction

Several strategies for assessing and improving statistical quality have been continuously developed and implemented at the Brazilian Institute of Geography and Statistics (IBGE) in the last 8 years. The paper will present some of the initiatives that have been taken and others that are in progress in order to improve the statistical quality management at IBGE.

2. IBGE Code of Practice

With this publication released in 2013, IBGE made available a set of guidelines, principles and indicators on the best practices adopted by the Institute in the production of official statistics, taking as reference the Statistics Code of Practice for Latin America and the Caribbean.

The guidelines aim at standardizing professional procedures to foster best practices in statistics, since their application is crucial to establish institutional credibility and, consequently, trust in the information produced by the Institution. Being applied, those principles can be assessed by society by means of indicators that allow the monitoring and evaluation of the quality of IBGE's statistical production.

The IBGE Code of Practice[1] sets out 17 key principles and 80 indicators, concerning the institutional environment and coordination; statistical processes and statistical products, which provide a reference for measuring the implementation of the Code.

The publication of the code is also seen as an important step towards improving the governance and consistency of official statistics in Brazil. That is because it aims not only at fostering the quality statistics debate in the several agencies of the National System of Statistics - NSS, but also at providing information to the creation of a more comprehensive version of this document, which will

---

[1]Available at http://www.ibge.gov.br/home/disseminacao/eventos/missao/codigo_boas_praticas.shtm

deepen issues concerning the management of the NSS and the production of statistics by the institutions which are part of it. Specifically, the code is seen as a mechanism to introduce a common understanding of quality across the diverse compilers of official statistics in Brazil and align national practices with international standards. For this intended future edition of the NSS Code of Practice, IBGE will promote discussion forums with other institutions responsible for the production of official statistics in the country.

Subsequent to this publication, a set of guidelines and protocols have been disseminated to provide guidance and support compliance with the principles and quality indicators established by the Code of Practice.

3. Quality indicators

Assessing data quality is one of the core aspects of a statistical institute's work, highlighted by the Code of Practice. Its principles require an assessment of the various product quality components like relevance, accuracy, timeliness and punctuality, accessibility and clarity as well as comparability and coherence. The code at the same time requires systematic assessments of the processes, including the operations in place for data collection, editing, imputation and weighting as well as the dissemination of statistics.

To comply with these principles, IBGE has released in 2014 the publication "The Minimum set of standard quality indicators to be implemented at MERCOSUR", based on European guidelines, which will provide relevant inputs for the development of IBGE´s own statistical quality framework in a near future, using these indicators as tools to measure the quality of statistics products or processes from several aspects.

Further improvements consist on developing a quality assurance framework, a system to monitor the quality of IBGE statistical production implementing the set of indicators mentioned above.

4. Guidelines on the dissemination of results by the IBGE

One of IBGE´s main objective is to ensure that the information produced is made available on an impartial basis to honour citizens' entitlement to public information, according to the first of the Fundamental Principles of Official Statistics.

In order to disseminate this principle widely among its technicians and Brazilian citizens, in October 2014 IBGE published at its website the **Protocol on Impartial Access to IBGE data**[2], containing recommendations related to both oral and written dissemination of official statistics, ensuring the compliance to equal access and impartiality of information. This initiative is aligned with two principles of the IBGE Code of Practice, together with their respective indicators, that are: impartiality e objectivity (Principle 7) e Accessibility e Clarity (Principle 17).

In April 2015, IBGE established other important orientation guide to ensure uniform error management through the introduction of the "**Guideline on how to deal with publication errors**". It is important to set and to widely disseminate the procedures of dealing with publication errors, where incorrect data are published despite all the quality assurance measures taken.

---

[2]Available at
http://www.ibge.gov.br/home/disseminacao/eventos/missao/principios_fundamentais_orientacoes_divulgacoes.shtm

In accordance with the IBGE Code of Practice, in its principle 15 - Timeliness and punctuality, indicator 15.4, "Errors discovered in published statistics are corrected at the earliest possible date and publicised".

To maintain confidence in official statistics in such a case, it is important that IBGE reacts reasonably and in a uniform and transparent way. Therefore, the whole procedure of dealing with publication errors in IBGE is described in this guideline, which provides an error classification and the reaction to each type of publication errors considering the different forms of disseminating statistical data.

Other initiatives in line with standard procedures are in progress and soon two new publications will be disseminated: one addressing procedures for dealing with misuse of statistical data, and other one establishing IBGE practices concerning the revision of statistical data.

The Fundamental Principle 4 establishes that the statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

As provided for in IBGE Code of Practice, in Principle 1 – Institutional Independence, indicator 1.7, "The statistical authorities of IBGE, when appropriate, comment publicly on statistical issues or advise on public comments, including criticisms and misuses of statistics as far as considered suitable.

Types of misuse include: discarding unfavourable data, loaded questions, overgeneralization, biased samples, misreporting or misunderstanding of estimated error, false causality, proof of the null hypothesis, data dredging, data manipulation, class fallacies, etc.

The guide on misuse of statistical data aims to define the procedures to be adopted by IBGE when a misuse, a wrong interpretation or even an overreaction from the media, society or government is detected, in relation to the statistical information released.

On the other hand, the publication on the practices related to the revision of statistical data describes these procedures in relation to planning, time of implementation and purpose. The reasons for revising data may be, for instance, due to the inclusion of an additional data source or due to transition to a new base period.

At IBGE, each statistical operation adopts its own data revision policy consistent with the type of information it produces. So, for example, the consumer price index used as the official inflation index is not subject to revisions, for legal reasons. On the other hand, some statistical business surveys revise their initial estimates because further data becomes available, and new, more complete or additional data that was not available at the time of the first release of statistics are included. And these definitions should be made clear to all users.

5. Statistical Metadata Base

Metadata is an important part of the standard dissemination procedure for official statistics. Metadata are intended to help users and researchers understand what the data are measuring and how they have been created. Without a proper description of a survey's design and the methods used when collecting and processing the data, there is a significant risk that the user will misunderstand and even misuse them. Good documentation also reduces the amount of user support statistical staff must offer external users of their microdata. Metadata are also intended to help users assess the quality of data. (DUPRIEZ & BOYKO, 2010).

According to UN principles governing international statistical activities, good practices on transparency include documenting the concepts, definitions and classifications, as well as data

collection and processing procedures used and the quality assessments carried out and making this information publicly accessible.

The description of good metadata can be extracted from the UK Data Archive's "Good Practices in Data Documentation":

> "A crucial part of creating a good dataset with longlasting usability is ensuring that the data are easy to understand and analyze. This requires accompanying data description and documentation that is userfriendly, clear and detailed, yet comprehensive."
> (http://www.data archive.ac.uk)

In addition, the Fundamental Principle 9 foster the use of international standards, in order to promote comparability of data produced by different agencies. Comparability is an important dimension of quality. If data are not comparable, they lose a lot of their utility. It is also a key principle of work to modernise official statistics production and services, that the use of common standards improves efficiency, both within individual agencies, and within the official statistics "industry" as a whole.

According to OECD Principles and Guidelines for Access to Research Data (2007), "Technological and semantic interoperability is a key consideration in enabling and promoting international and interdisciplinary access to and use of research data. Access arrangements must pay due attention to the relevant international data documentation standards."

Consistent with these good practices, IBGE is undergoing a wide redesign of its corporate statistical metadata base. The intent of the metadata modernization initiative is to establish common semantic structures, robustness of systems and harmonization of survey questions, concepts and definition, in an user friendly environment. Furthermore, the metadata structure is being developed in alignment with international standards, in particular, ISO/IEC11179 (Metadata registries) and the Data Documentation Initiative – DDI standard.

The International Standard ISO/IEC 11179 describes the standardizing and registering of data elements to make data understandable and shareable, by providing a standard for describing semantic and syntactical characteristics of data (i.e. metadata) and for managing and administering these descriptions in a registry in order to make them conveniently accessible

The DDI is an international metadata standard designed to describe socioeconomic surveys, censuses, and other microdata collection activities. The DDI is dedicated to microdata documentation that enables documentation of even the most complex microdata files in a way that is simultaneously flexible and rigorous. It provides a straightforward means of recording and communicating all the salient characteristics of microdatasets. To take full advantage of web technology, DDI is defined in eXtensible Markup Language, or XML, a common tool to structure information to be shared on the Web and between software systems. XML not only provides a common language and facilitates metadata management but also is easy and cost effective to adopt as a technology. XML documents are stored in a standard text format, ensuring that anyone can read them and long term preservation

The use of these standards facilitate data communication between organisations and software systems and improve the quality of statistical documentation provided to users of data, by means of a structured framework for organising and disseminating information on content and structure of statistical information.

It is expected to move IBGE toward greater quality and coherence in terminology used for metadata by means of harmonization. A Metadata Harmonization working group is being created at IBGE in

order to harmonize survey questions, concepts and definitions to ensure that the quality of official statistics meets user requirements, making it possible to compare data from different sources with confidence and to merge and match data more easily.

The harmonization strategy will bring many benefits to both producers and users of statistics, such as:

- Better comparability for cross-survey analysis and integration by making statistics comparable across time and across different sources, thereby creating an environment in which multiple data sources can be integrated;
- Improved data quality by reducing potential misunderstanding or confusion caused by the use of different questions from the same topic area;
- Increased efficiencies by avoiding "reinventing the wheel" by using pre-existing harmonized principles;
- Improved ease of use and comparability – for data users who require multiple datasets, having multiple concepts with multiple meanings could not only be arduous and time consuming to handle, but increase the possibility of mistakes or make results incomparable.

6. Governance and National Statistical Coordination

The Fundamental Principle 8 establishes that coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system. Nonetheless, according to Global Review (2013), this principle is among the least implemented of the Fundamental Principles of Official Statistics.

In Brazil, there is a legal framework that establishes a National System of Statistical Information and states that the responsibility of regulating and coordinating that System will be in charge of the National Statistical Office, i.e., IBGE. Although this coordination does not take place comprehensively, IBGE plays an important role in providing technical and methodological guidance on technical matters of statistics products and services to specific demands from other statistical producers.

Significant efforts have been made to increase the statistical use of administrative records, in order to improve the efficiency of the statistical production and to reduce the burden placed on the respondents to statistical surveys.

In order to promote the effective coordination and governance of the National Statistical System, the Social Statistics Committee was established, in December 2007, as a forum, coordinated by IBGE, consisting of the main producers of social statistics of the federal government aiming at the integration of the various databases derived from surveys and administrative records. The following guidelines and rules were set for guiding the statistical work of the members of the Committee:

- To ensure and to improve the comparability and accessibility of information;
- To encourage the adoption of concepts, classifications and methods for processing, archiving and presenting official statistical information;
- To fill in information gaps and expand the integration of data sources;
- To organize the demand for information and coordinate their production and use;
- To regulate and to facilitate access by various government agencies to existing databases, ensuring transparency, visibility and confidentiality;
- To adopt the Fundamental Principles of Official Statistics.

Similar efforts are being conducted to create the Economic Statistics Committee, congregating the main producers of economic and business statistics on the federal government level.

Another important initiative towards the coordination of the National Statistical Systems refers to the use of a common structure of metadata among the main producers. IBGE is implementing a training plan in order to develop skills and professional expertise on documentation of statistical operations and administrative records, using the standards adopted by IBGE Metadata Base.

The plan, in a near future, is to congregate all data and metadata bases of the members of the National Statistical System under a common TI and methodological infrastructure, in order to organize, regulate and harmonize the statistical production of the country. Specific standards should be established to preserve, capture, analysis and exchange of information by electronic means between IBGE and the main producers of the national statistical system. The development of such project is a challenge that will require strong interagency articulation while aiming at the efficiency and the benefits of having a national statistical information system that upholds quality, relevance, objectivity and accessibility of the statistical information.

7. Conclusion

IBGE has been continuously implementing initiatives to enhance the management of statistical quality, in order to improve the accuracy, comparability, accessibility and transparency of the statistical production process. Although much have been done recently, a lot of significant efforts are currently in progress and still being implemented to meet the relevant quality standards required and to improve the consistency and efficiency of the national statistical system, by reinforcing IBGE coordination role.

References:

Conjunto mínimo de indicadores padrão de qualidade a ser aplicado no MERCOSUL. Textos para discussão nº 52. Diretoria de Pesquisas, IBGE. 2014

Dupriez, Olivier and Ernie Boyko. Dissemination of Microdata Files. Formulating Policies and Procedures. International Household Survey Network, IHSN Working Paper No 005. 2010.

IHSN technical note on metadata standards – DRAFT. Available at http://www.ihsn.org/home/sites/default/files/resources/DDI_SDMX_IHSN_DRAFT.pdf

Implementation of the Fundamental Principles of Official Statistics Background Document. United Nations Statistics Division. 2013. Available at http://unstats.un.org/unsd/statcom/doc13/BG-FP.pdf.

ISO/IEC 11179, Information Technology -- Metadata registries (MDR). Available at http://metadata-stds.org/11179/

OECD Principles and Guidelines for Access to Research Data from Public Funding. Organisation for Economic Cooperation and Development (OECD). 2007. Available at http://www.oecd.org/dataoecd/9/61/38500813.pdf

Thomas, Jan; Shah, Palvi; Aquilina, Rebecca, Harmonising Standard Questions, Classifications and Concepts at the ONS. European Conference on Quality in Official Statistics. 2014

United Nations Fundamental Principles of Official Statistics – Implementation guidelines, 2015. Available at http://unstats.un.org/unsd/dnss/gp/Implementation_Guidelines_ FINAL_without_edit.pdf