

Mapping potential administrative data for statistics purposes – Statistics Portugal approach within Public Administration

Magda Ribeiro¹, Maria João Zilhão², Paulo Saraiva³

¹ *Statistics Portugal, Lisbon, Portugal; magda.ribeiro@ine.pt*

² *Statistics Portugal, Lisbon, Portugal; mjoao.zilhao@ine.pt*

³ *Statistics Portugal, Lisbon, Portugal; paulo.saraiva@ine.pt*

Abstract

The use of administrative data for statistical purposes has assumed particular importance in the context of the production of official statistics, without jeopardizing its quality, as a way to reduce costs and decrease the burden on respondents. It is a part of a strategically change which Statistics Portugal (SP) has been following, and for which is indispensable the cooperation of different public entities which owns such kind of data.

This paper describes the approach taken by SP to: (i) identify some of the constraints associated with obtaining data through administrative sources, even though the legal existing framework for statistical production; (ii) summarize the results of a mapping exercise of administrative data sources with potential to be used for statistical purposes, regardless of its usage (eg. appropriation of administrative data for replacing or simplifying survey data collection; and/or for data validation purposes, already collected by other means); and, (iii) report on the administrative sources that are already being used by SP, focusing on economic statistics, stressing on the success factors and on the constraints of its usage in relation to the cooperation amongst institutions and legal frameworks.

A strong focus will be put on the mapping exercise exploring the approach and the means developed for it; the relation with this mapping exercise with all activities at SP production classification activities system; and finally the results of its potential usage classified by a short or long term perspective of its real appropriation, taking into account the diligences that have to be taken in order to assure the adequate quality criteria for statistics proposes.

Keywords: Administrative data; potential administrative data usage; quality and administrative data.

1. Framework for accessing Administrative data in Statistics Portugal

The use of administrative data for statistical purposes may allow greater efficiency of production processes, in particular by reducing production costs without jeopardizing the quality of statistics and reducing the effort and burden on respondents.

This strategic line is expressed in various domains and strategic documentation both at national and European levels, by which Statistics Portugal (SP) is guided, namely:

- The Portuguese National Statistical System Law (Law n.º 22/2008, 13th may)¹.
- The General Guidelines for Official Statistics (LGAE0) for the Portuguese National Statistical System, for the period 2013-2017, namely in what concerns:
 - Objective 1: “To reinforce the quality of official statistics, ensuring the optimization, improvement, flexibility, modernization and efficiency of the statistical production process through their methodological, scientific and technological development”, and in four Action Lines (LA):

LA 2 – To ensure intervention by Statistical Authorities from the inception of process to design/revise administrative acts, so as to guarantee their use for statistical purposes.

LA 3 – To call the attention of entities under the State’s direct and indirect administration holding administrative data to the importance of supplying these data for the production of official statistics and promote in those entities mechanisms to facilitate and develop the data utilization process.

¹ Article 4 (Statistical authority) / item 2 — (...) “for all purposes, the use of administrative records for official statistical purposes shall be considered one of the key objectives of their collection.”

Article 8 (statistical accessibility) / item 2 — “Official statistics are considered a public good. They shall be proportionate to user needs and have a non-excessive burden on respondents, namely through a wider use of administrative data.”

LA 4 – To trace and reinforce the use of administrative sources in the production of official statistics, so as to streamline allocated resources and reduce the burden on respondents.

LA 5 – To continue developing a new population and housing census model essentially focused on the use of administrative files.

- The European Statistics Code of Practice – Principles 8 (Appropriate Statistical Procedures) and 10 (Cost Effectiveness), highly adopted by SP and Other National Authorities (ONA) from its early beginning, that include a set of indicators that should reflect the efforts and actions taken by Statistical Authorities to enhance the increase use of Administrative data (namely Indicators 8.1; 8.7; 8.8; 8.9 and 10.3).
- The amended Regulation n.º 223/2009 (Nº 759/2015, de 29th April), which reinforces, in particular, the coordinating role of Statistical Authorities regarding the production and dissemination of European statistics at the national level and improve the right of access to administrative data sources, promoting the involvement of Statistical Authorities in the design, development and elimination of administrative records used for statistical purposes.
- Lastly, a reference to the recent Peer Review exercise should be made, in what concerns the implementation of the European Statistics Code of Practice, run in Portugal (both to SP and ONAs), in January 2015. In fact, the Peer Review Report includes the analysis of processes and situation concerning the use of administrative data in Portugal, and two direct recommendations were made:
 - Recommendation 3 - The responsible Portuguese authorities should amend the Law on the National Statistical System to strengthen the right of Statistics Portugal to access administrative data and to make explicit the obligation of data owners to provide these data on demand for statistical purposes. (European statistics Code of Practice, indicator 2.2.)

- Recommendation 4 - The responsible Portuguese authorities should amend the Law on the National Statistical System to make explicit the right of Statistics Portugal to be involved in the design and revision of any administrative data system in the public sector in order to make administrative data more suitable for statistical purposes and to monitor their quality (European statistics Code of Practice, indicators 8.7., 8.8. and 8.9.).

Despite this framework and SP efforts in this area in recent years (and even though administrative data is appropriated for statistical purposes in several domains, and good examples exist at national level), constraints on its access persist in some areas, remaining a big challenge for SP:

- Certain areas/domains have legal specific rules for access to data that are not consistent with legislation that allows access to administrative sources by statistical authorities;
- Lack of understanding of the utility to society and cooperation of entities owners of such data with the national and European strategy in this area;
- Poor quality of administrative data produced, stored and updated by some entities;
- Lack of an institutional procedure to monitor the progress of legal acts involving the process of the creation of administrative data, verifying often that any attempt to use of administrative data sources occurs outside the legal act and are based on knowledge of *ad hoc* existing sources;
- Lack of an effective mechanism for ensuring the respect for the rights of the statistical authorities in the field of administrative data usage.

2. Administrative data already used by Statistics Portugal, inventory processes

Despite the mentioned constraints, the challenges persist and several activities already benefit from the use of administrative data, specifically with the following objectives:

- Total or partial replacement of primary statistical information needs, that is, which is collected via traditional surveys;

- Validation/calibration of the primary statistical information;
- Update of statistical units registers, essential to provide population and sampling frames with quality.

A list of what is already in use in constantly updated, ready to use for several purposes and domains, and at internal level.

Besides the information that is already being used it is important to describe how SP organizes itself to deal with the constraints and develops actions to overcome them.

One of the actions is to undertake, on a yearly basis and per domains level, feasibility studies, contacting the entities that have potential administrative data for statistical purposes. The studies are done on a bilateral basis and the “stability” of the contact person on the side of the partner entity is crucial. These *ad hoc* activities are in line with the National Strategy and are part of the Annual Activity Plan. They also make part of indicators of our Responsibility and Assessment Framework (QUAR).

The National Public Administration approach in assessing performance in public organizations - the Responsibility and Assessment Framework (QUAR) - is an integrated approach in defining and assessing the annual main objectives and the indicators that are related to them at organizational level, but also at unit and employee level. This framework has helped organizations to find a global approach for assessing annual performance, dully aligned with the Strategic Objectives, whose results are used for measuring organization Excellency amongst other organizations, based on similar criteria of the European Foundation for Quality Management (EFQM) Excellency Model.

For example, the following feasibility studies per domains have been undertaken and are part of the QUAR 2013/2015:

- Study on the administrative data use from the Municipalities for the Indicators System of Urban Operations;
- Studies on increasing of use of administrative data from entities of the Ministries of Agriculture and Economy;

- Feasibility study on the use of administrative data from entities related to public transport system;
- Study on the use of administrative data from entities related to public transport system;
- Study on the administrative data use sources to enhance statistics on internal migrations within the resident population estimates;
- Studies of use of administrative data sources in the new census model for 2021.

The results achieved are very important in our way to a more intensive use of administrative data.

Another good practice that SP undertook was the development of an IT platform *Inventory of Administrative Sources* (IFA), which allowed the Ministry of Agriculture and Sea to make an inventory of potential administrative sources for statistical purposes. This platform was developed by SP and also brought benefits of systematizing the information to the Ministry itself. This flexible tool can be used by other entities, with eventual adaptations, allowing a detailed and current view of the existing reality of the availability of administrative data.

3. Potentialities of administrative data for statistical purposes – Mapping exercise

In spite of all the developments that were achieved in recent years, in which some good practices were established, SP decided to analyze in detail the potentialities of administrative data in public administration making its own inventory.

In fact, there are unexplored administrative data sources that are available for SP to analyze and to perceive their potential. And it is crucial to evaluate whether this potential administrative data, which are available in public administration, can be used in a short and medium term or only in a long term horizon, implying, for instance, the change of legislative acts. And this is a very important and difficult dimension/attribute of administrative data that SP has to cope with.

In this context, the decision to appropriate administrative data should be taken based on the analysis of various dimensions of quality. We highlight the following five critical factors:

- The adequacy of concepts, classifications, nomenclatures and variables of administrative data in what concerns statistical compilation needs;
- The quality of the administrative data itself;
- The access conditions and terms of transmission of administrative data;
- The stability of the administrative source and its contents;
- The operational capacity of the statistical authority in the management and storage of a specific administrative data, namely in huge data bases.

Indeed, there are several dimensions/quality indicators that should be taken into account in order to identify administrative sources that actually can (and should) be used in statistical production. The mapping exercise done within SP was focused at the stage of the analysis of the relevance of the administrative source.

The relevance of the source was made taking into account the SP Annual Activities Plan of official statistics, at the level of the statistical activity (taking into account SP Classification Activity System) and for every domain. In this first phase, the focus was on the administrative data of a business nature. The mapping exercise was undertaken within SP, involving several internal units, and also at an institutional level: the results were presented to a working group in the context of public administration concerning administrative statistical simplification, in which different Public Administration entities' representatives participate.

This exercise had two major objectives: i) The mapping exercise itself, allowing SP to better systematize the potentialities of administrative data at the level of the statistical activities, independently on its type of usage; and ii) showing to Public Administration Entities, the potentialities of the use of the information they own for statistical purposes.

It should be noted that this mapping does not imply an effective appropriation, but a starting point for the study of its relevance and type of use.

In this first systematized mapping exercise SP identified:

- 44 entities with potential administrative sources;
- 127 potential administrative sources, relevant to 52 statistical activities. For this 127 administrative sources:
 - 13 involves the need to amend the legislative framework;
 - For 106, its possibility of usage can be a reality in the short-medium term.
 - And 8 potential administrative sources are still being analyzed whether their possibility of usage is within the short/ medium term or long term.

With the mapping exercise a comprehensive description was done, taking into account the following information (filled out by each department/unit), being possible to correlate:

Identification of the administrative source

Designation of the administrative source
Entity owning the administrative source
Department or unit of the entity which owns the administrative source
Ministry of the entity

General characteristics of the administrative “act”

Legislation that establishes the administrative “act”
Closing dates of the administrative act for the purposes of data collection
Geographical scope of the administrative act
Frequency of the administrative act

Administrative data with potential use or already in use, related to each administrative source

Administrative source with potential use or already in use for statistical purposes
For potential sources identification if its usage for statistical purposes can be accomplished in the medium or long term. In the long run it means there is a need to amend the law of the administrative act.
Type of use of administrative data (already in use): (i) Total or partial replacement of primary statistical information; (ii) Validation/calibration; (iii) Update of statistical units registers.

Linkage between the administrative source and the statistical activities (linked to the SP Activities Plan)

Identifies the entity responsible for the statistical activity: <ul style="list-style-type: none"> • SP • ONA
SP Department/unit responsible for the statistical activity
Statistical area (code and designation)
SP Classification Activity System (CGA) (code and designation)
Basis and legislative and regulatory framework of SP/ONA activity - description

All this information can be correlated and organized in different ways.

The list does not exhaust the possibilities of appropriation of administrative data. In addition to the unexplored sources, there is also potential for a more intensive use of sources that SP already has been working with.

Taking into account the quality dimensions of quality of administrative data, experience has shown some weaknesses in the administrative data appropriation process:

- Coverage and timeliness of administrative data not compatible with the European Regulation requirements of statistical data production, Confidentiality of data is very

often an excuse/reason to not give administrative data to SP; in some case information is given but in an anonymised way, which makes impossible to cross it with additional information sources;

- Simplification, modification and/or elimination of administrative acts done unilaterally, without taking into account statistical needs.

4. Final remarks and next steps

For the short term much of this potential administrative data could be effective if from the part of some entities, there was better understanding of the legal regime of official statistical production, particularly on two principles: i) the right of SP to access to administrative information and the participation of its involvement in the improvement of administrative data; ii) the individual data protection guarantee as a result of the statistical confidentiality principle, which under the NSS Law will override the other duties of confidentiality of entities producing administrative data.

The correct application of these principles should allow widespread access by SP to administrative data not anonymised, and enabling the crossing of information from various sources. Moreover, this situation already exists in some particular areas but very relevant to the official statistical production, including national accounts, such as the Simplified Business Information. In a program set up by the Portuguese Government with the purpose of simplifying and modernizing the administration in order to make life easier for enterprises and citizen, four national public entities (Ministry of Finance, Ministry of Justice, Statistics Portugal and Portuguese Central Bank) took part in the creation of a system that made it possible to acquire administrative data while simultaneously complying with four legal obligations. Since 2006, SP works with micro data from financial statements, in an electronic way, not anonymised, for all companies that operate in Portugal. The whole business production system was reengineered, and the advantages were huge, namely in what concerns burden reduction, availability of information, and more efficiency in sampling designs.

In order to create more of the cases of success, there are a set of guidelines to be pursued, such as:

- Identify in detail the legal changes to be made so that it becomes feasible to use the result of the mapping exercise and the inventory of administrative sources/data;
- Identify the requirements for the use of data from administrative sources whose use was considered short and medium term possible;
- Continue with the inventory studies of administrative sources, namely in other statistical domains, such as social statistics;
- Continue the development of protocols with entities possessing the administrative sources for the transfer of data for statistical purposes;
- Participate in the design process / review of administrative acts in order to ensure their use for statistical purposes;
- As last but not least, continue to build a catalog of administrative sources/data for statistical purposes from public administration.

References

Portugal, National Statistical System (2008), Law No. 22/2008, of 13 May, approves the general basis of the National Statistical System,

http://cse.ine.pt/xportal/xmain?xpid=CSE&xpgid=cse_main&cont_cse=139660987

Official Journal of the European Union (2009), Regulation (EC) No 223/2009 of the European Parliament and of the Council, [http://eur-](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:087:0164:0173:EN:PDF)

[lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:087:0164:0173:EN:PDF](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:087:0164:0173:EN:PDF)

Statistics Netherlands (2009), Checklist for Quality evaluation of Administrative Data Sources, Statistics Netherlands,

http://ec.europa.eu/eurostat/documents/64157/4374310/43_CHECKLIST-QUALITY-STATISTICAL-OUTPUT-CBS-2009.pdf/2000653a-e8bc-4979-8e13-bb455b5ddc5

Portugal, Statistical Council (2012), General Guidelines of Official Statistical Activity 2013-2017, http://cse.ine.pt/xportal/xmain?xpgid=cse_home&xpid=CSE&xlang=en

ESS (2011), The European Statistics Code of Practice,

https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_cont_inst&INST=53871&xlang=en

Rockmann, U., Vaicenavicius, R.J., Villar J.G. (2015), Peer Review Report on compliance with the Code of Practice and the Coordination role of the National Statistical Institute,

<http://ec.europa.eu/eurostat/web/quality/peer-reviews>

Statistics Portugal, Economic and Social Council, Economic Commission for Europe,

Conference of European Statisticians (2009), Simplified Business Information – Producing

Quality Statistics based on Administrative data: a true possibility? ,

<http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2009/14.e.pdf>