

Distr.
LIMITED
LC/CEA.11/5
10 November 2021
ENGLISH
ORIGINAL: SPANISH
21-00659

Eleventh meeting of the Statistical Conference of the Americas
of the Economic Commission for Latin America and the Caribbean

Virtual meeting, 23–25 November 2021

**METHODOLOGICAL DOCUMENT ON STATISTICAL USE
OF ECONOMIC ADMINISTRATIVE RECORDS**

SUMMARY



CONTENTS

	<i>Page</i>
INTRODUCTION	3
A. CHAPTER I. IMPORTANCE OF ADMINISTRATIVE RECORDS	6
B. CHAPTER II. VISION OF A SYSTEM OF ECONOMIC ADMINISTRATIVE RECORDS	6
C. CHAPTER III. SPECIFICATION OF NEEDS	9
D. CHAPTER IV. DESIGN.....	10
E. CHAPTER V. CONSTRUCTION	10
F. CHAPTER VI. DATA COLLECTION	11
G. CHAPTER VII. PROCESSING.....	12
H. CHAPTER VIII. ANALYSIS	12
I. CHAPTER IX. DISSEMINATION	13
J. CHAPTER X. EVALUATION.....	13
K. CHAPTER XI. CONCLUSIONS AND RECOMMENDATIONS.....	13
L. ANNEXES	14
M. GLOSSARY	14
N. BIBLIOGRAPHY	14

INTRODUCTION

This document summarizes the main elements of the “Methodological document on statistical use of economic administrative records”, prepared by the working group established by resolution 11(X) of the Statistical Conference of the Americas of the Economic Commission for Latin America and the Caribbean (ECLAC). The working group was coordinated by Colombia (National Administrative Department of Statistics (DANE)) and Mexico (National Institute of Statistics and Geography (INEGI)).¹

The objective was to prepare a methodological document for the statistical use of economic administrative records, to generate new statistics or improve existing ones in the countries of Latin America and the Caribbean, taking into account the conditions and level of statistical development of the countries in question.

In recent years, several countries around the world have focused on using administrative records as a source of information for the production of statistics. In the case of developed countries, there is a statistical infrastructure that includes, at least, data on the identification and location of observation units (persons, dwellings, establishments, businesses, properties, or others), which are completely updated through administrative records.

In Latin American and Caribbean countries, the situation is different. There are still several problems that prevent the creation of a similar infrastructure to generate accurate and reliable economic statistics.

The problems faced by the countries of the region include the following:

- The presence of rigid legal frameworks prevents government agencies from sharing administrative records with the national statistical offices.
- Administrative records have specific objectives that do not necessarily have a statistical purpose. They are also based on conceptual frameworks specific to the objective they pursue, which are not always in line with those developed by the national statistical offices.
- Government agencies and private entities that have administrative records use classifiers that meet their own needs, but differ from those used by the national statistical offices.
- National statistical offices construct independent sampling frameworks for the statistical design of their surveys.
- Although there are various international recommendations on using administrative records for statistical purposes, there is no methodology nor statistical and information technology tools that would enable the countries of the region to build a statistical infrastructure based on administrative records to be used to statistical purposes.

The document identifies the need for conceptual definitions and guidelines for the statistical use of specific administrative records in the region’s countries, considering their current situation, their level of statistical development and the availability of administrative records, among other aspects. A special section of the document provides definitions for the terms used in the statistical use of administrative records.

¹ The authors are grateful to the agencies of various countries in the region and to United Nations entities for valuable assistance provided in the preparation of the document.

Accordingly, the methodological document is intended to provide the countries of the region with a set of guidelines and recommendations for producing innovative and timely economic statistics, and to provide greater benefits by making statistical use of administrative records.

The four fundamental inputs for the development of the methodology are:

- (i) The book *Register-based Administrative Data for Statistical Purposes* by Anders Wallgren and Brit Wallgren. This volume contains an exhaustive presentation of the administrative records that feed the four basic registers —population register, business register, real estate register and activity register— which in turn support the national statistical system (NSS).
- (ii) The Economic Commission for Europe (UNECE) *Guidelines on Statistical Business Registers*. The methodology proposed in this document is aligned with these international recommendations, viewing the statistical business register as promoting and optimizing the production of economic statistics.
- (iii) The UNECE Generic Statistical Business Process Model (GSBPM). The generic model of the statistical process is the framework within which the guidelines and recommendations are developed.
- (iv) The experience of Latin American and Caribbean countries. This was one of the most important inputs used to develop the methodology. The methodology incorporates good practices that these countries have implemented and shows how they have solved problems that are common across the region.

The methodological document consists of 11 chapters. The first chapter highlights the importance of administrative records. The second chapter presents the vision of a system of economic administrative records. It also briefly describes how administrative records can be used to update the statistical business register.

The third chapter develops the guidelines for the statistical use of administrative records based on GSBPM. A specific chapter is devoted to the guidelines and recommendations for each of the eight phases of the model —specify needs, design, build, collect, process, analyse, disseminate and evaluate— based on the best practices of the countries of the region.

Table 1 displays the eight phases and 44 subprocesses of BSBPM, which form the reference framework for developing the methodology. The crosscutting processes for quality management and metadata preparation are also indicated.

The methodological document ends with a chapter of conclusions and recommendations. Lastly, the annexes include examples of good practices in this area that have been carried out in Latin American and Caribbean countries.

Table 1
Phases and subprocesses of the Generic Statistical Business Process Model

Management of quality and metadata							
Specification of needs	Design	Construction	Collection	Processing	Analysis	Dissemination	Evaluation
1.1 Identification of needs	2.1 Conceptual design	3.1 Construction of collection instruments	4.1 Formation of the sampling framework and selection of the sample	5.1 Integration of data	6.1 Preparation of draft results	7.1 Updating of output systems	8.1 Concentration of inputs for the evaluation
1.2 Consultation and confirmation of needs	2.2 Design of variable descriptions	3.2 Construction or improvement of components	4.2 Preparation of the collection	5.2 Classification and codification	6.2 Validation of results	7.2 Generation of dissemination products	8.2 Evaluation
1.3 Definition of objectives	2.3 Design of collection	3.3 Construction or improvement of dissemination components	4.3 Execution of the collection	5.3 Revision and validation	6.3 Interpretation and explanation of results	7.3 Management of the publication of dissemination products	8.3 Determination of a plan of action
1.4 Identification of concepts	2.4 Determination of sampling framework and type	3.4 Organization of workflows	4.4 Closure of the collection	5.4 Editing and imputation	6.4 Application of control of dissemination	7.4 Promotion of dissemination products	
1.5 Verification of data availability	2.5 Design of processing and analysis	3.5 Tests of the production system		5.5 Derivation of new variables and units	6.5 Finalization of results	7.5 Management of user support	
1.6 Preparation of the general plan	2.6 Design of production systems and workflows	3.6 Pilot test of the statistical process		5.6 Calculation of weights			
		3.7 Finalization of the production system		5.7 Calculation of aggregations			
				5.8 Finalization of data files			

Source: Economic Commission for Europe (UNECE), Generic Statistical Business Process Model (GSBPM) version 5.0 (English), Geneva, 2013.

A. CHAPTER I. IMPORTANCE OF ADMINISTRATIVE RECORDS

This chapter describes the conceptual reference framework used in the document. It also highlights the importance of administrative records, their proper treatment and processing, their comparability and their integration, which is highly relevant for constructing robust indicators that support government decision-making.

The interest in making statistical use of administrative records stems mainly from their qualities, as follows:

- coverage and continuity of data collection on events that may be of general interest;
- coverage and geographic disaggregation of the records, particularly when they are national in scope;
- feasible characterization of the phenomenon given the variables captured, in accordance with the thematic coverage that is confined to the legal jurisdiction of the organizations or agencies that compile the record;
- possibilities of broadening the conceptual coverage of data collection; and
- marginal costs relative to current statistical use.

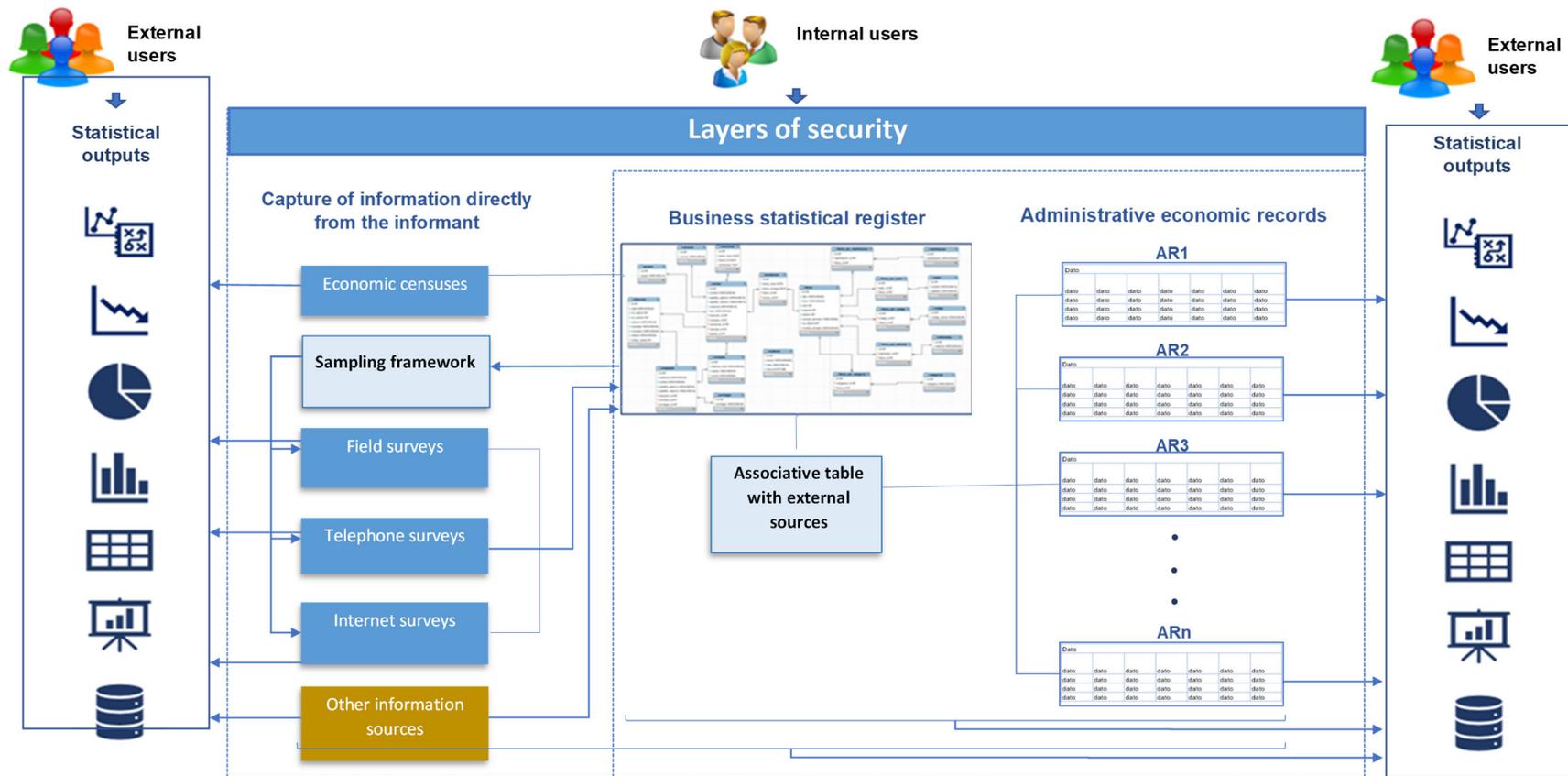
B. CHAPTER II. VISION OF A SYSTEM OF ECONOMIC ADMINISTRATIVE RECORDS

The vision of the role played by economic administrative records in relation to the statistical business register is proposed, mainly with regard to the following aspects:

- Updating of the statistical business register using economic administrative records (taxes, electric power and water services, among others);
- Production of statistics through the statistical use of administrative records:
 - traditional exploitation of the administrative record: statistical use of the statistical record only, using descriptive statistical techniques;
 - production of statistics by linking administrative records with the statistical business register: statistical use of various administrative records linked through the statistical business register, combining information from various sources using descriptive and inferential statistical techniques; and
 - integration of a system of economic administrative records consisting of the business statistical register and economic administrative records, for the purpose of establishing interrelationships to meet the requirements arising from implementation of the 2030 Agenda for Sustainable Development.

Diagram 1 illustrates the elements that make up a system of economic administrative records. It encompasses the use of various administrative records and additional sources of information, how they are used as a support for sampling frameworks, and the production of statistical outputs to meet the requirements of information users.

Diagram 1
View of an economic administrative records system



Source: Prepared by the authors, on the basis of National Institute of Statistics and Geography (INEGI).

This chapter also addresses international trends in administrative records and in the innovation processes of both public and private sector institutions.

C. CHAPTER III. SPECIFICATION OF NEEDS

Administrative records are a secondary statistical source; that is, they were designed with nonstatistical objectives by administrative areas of the different public (national and subnational) and private administrations. This means that there is an “owner” of the administrative record that is not the national statistical office.

This “right of ownership” is vested in each secretariat, ministry, directorate, department or other agency in each country —entities that have control over the administrative records they hold.

Optimizing the statistical use of administrative records requires good will on the part of multiple political and bureaucratic entities, which means the national statistical offices have to negotiate with the direct and political “owners” of the records in question. Thus, the task of developing a system of administrative economic records based on administrative records involves the confluence of various actors: political, technical and administrative.

In general, the countries of Latin America and the Caribbean have, at both the national and subnational levels, a large volume of administrative data that has not been exploited statistically. However, it is essential to select those that meet the country’s needs.

These elements are addressed in this chapter, along with the desirable characteristics that administrative records should possess if they are to be used for statistical purposes. The following elements, among others, are described:

- desirable characteristics of the administrative records (coverage and completeness, among others);
- observation units;
- identification of administrative records with statistical potential;
- metadata of the administrative records;
- protocols for the exchange of information (interagency agreements, memoranda of understanding);
- inclusion of the issue of “administrative data as a statistical source” on the political agenda;
- management of administrative records;
- evaluation of the quality of administrative records;
- standardization of administrative records with the business statistical register;
- linkage of administrative records with the business statistical register; and
- preservation of the information.

D. CHAPTER IV. DESIGN

The design of a system of economic administrative records focuses on the process to be followed in turning an administrative record into a statistical record and for its results to be comparable and consistent with those obtained from other information sources: censuses or surveys.

In the conceptual design of projects or programs based on administrative records, the following should be considered:

- Specification of the objectives of the economic statistical register;
- Definition of the set of objects of the register (observation unit);
- The unique identification number for the observation units, for public use;
- Preparation of an inventory of sources (the objectives, scope, attributes and coverage with which the administrative register from which the data will be obtained operates);
- the criteria for the selection of sources;
- the statistical outputs to be generated by the project;
- Identification or selection of variables, following statistical and conceptual criteria (the data to be processed to generate the variables of interest to the project or program);
- Identification, georeferencing and control data (identification keys) that make it possible to link the data contained in two or more administrative records, and
- Data that help in the creation of base registers for units or objects of observation of strategic interest, among others.

E. CHAPTER V. CONSTRUCTION

This chapter describes the modules that a system of economic administrative records based on administrative records should contain, in order to guarantee conceptual uniformity and full comparability of the statistical and geographic information produced at both the national and the international levels.

These modules should correspond to the needs of the countries, so that the information contained in the administrative records can be used in a cross-cutting manner by all administrative areas of the national statistical offices, while maintaining the confidentiality of the information.

To this end, technological tools must be developed to support the activities of transmission and reception of the variables in the administrative records, derivatives, files or statistical or non-statistical databases that will be the object of study. This means developing the technological means of transmission and interoperability of data obtained from administrative records, files or statistical or non-statistical databases; developing the techniques and tools to verify that the information sent or received satisfies the quality conditions established for data exchange; and preparing the procurement plan, as well as the bids, contracts, agreements or arrangements necessary to perform the statistical operation.

The importance of developing technological means of transmission and interoperability must be considered. This is becoming a challenge for data exchange, owing to the increasing volume of information available for use in statistical operations. Thus, national statistical offices, in their different realities, need to be prepared to address these challenges.

They must also provide for the capture of diverse types of data, which can be classified according to their origin, their range or their structure, among other factors. According to their structure, three types of data can be defined: structured, semi-structured and unstructured.

Historically, national statistical offices have only focused on extracting and analysing information from structured data. However, with the current increase in information in semi-structured and unstructured form, national statistical offices must seek solutions that can help them analyse all three types of data, as this represents an analytical advantage and, in fact, the ability to extract value from them is one of the main drivers of the rapid growth of big data.

F. CHAPTER VI. DATA COLLECTION

This chapter describes mechanisms that can be used for information exchange between administrative record-keeping entities and the national statistical offices, in such a way that they are efficient and guarantee the confidentiality of the information.

In the collection phase, data can be obtained from various sources, such as traditional ones from statistical operations; and data can be compiled from those generated by the information-producing units based on administrative and geographic records and images, and so forth, which have been derived from databases or microdata.

Given the reality of each country, and of each unit that keeps administrative records, there are several scenarios for requesting data and their statistical use; however, it is important to distinguish two situations or levels that occur in this process, the collection of data by the units holding administrative records, and access by the national statistical offices to the data or information for statistical use.

Considering the issues discussed in the previous phases, the information collection process should be underpinned by a methodology established for the purpose, including the processes, the availability of appropriate technology, materials and other relevant elements for the preparation and execution of the collection.

The interagency agreements or memoranda of understanding that are established should consider mutual interest; in other words, all participants should obtain benefits. In this sense, this is an important opportunity for national statistical offices to fulfil their role as lead agencies and contribute their expertise.

The chapter also discusses the advantages of drawing up interagency agreements or memoranda of understanding, covering the following issues, among others: exchange mechanisms, responsible entities, periodicity, delivery dates, information validation criteria, standardization of criteria, use of classifiers and the preparation of diagnostic assessments.

G. CHAPTER VII. PROCESSING

This chapter addresses the need to define criteria for consistency analysis and information integration procedures, as well as the importance of having data models to store and retrieve information efficiently. Consideration should be given to the following activities, among others:

(i) *Processing of information from administrative records.*

This stage includes the process of converting administrative records into statistical records, which requires data purging, considering an analytical-conceptual and quality assurance framework; subsequently, the information is organized for processing. During this stage, the data is analysed to detect errors in order to impute or eliminate them,—redundant, incomplete, incorrect or other types of data— as the case may be, in order to start creating quality data for the production of statistics. Once the data has been purged, it is entered into the economic administrative records system using integration methods and procedures.

To make the information transfer simple, the records should have a standard structure, for which purpose the national statistical offices should undertake conceptual and information technology standardization work in conjunction with the entities that hold administrative records.

(ii) *Information processing*

This phase consists of a series of steps, most of which are computerized and make it possible to prepare the initial database for analysis and exploitation, and for data storage. This database must include a variable that allows a record to be unequivocally identified, in addition to generating a data dictionary.

The chapter also stresses the importance of the aspects of the computing systems that are needed to automate procedures (access, security and linkage, among others) to generate results in a timely manner.

H. CHAPTER VIII. ANALYSIS

This chapter stresses that it is essential to verify the statistical results obtained from administrative records and check that the validation mechanisms ensure the congruence and consistency envisaged in the processing design.

Emphasis is placed on the need to analyse the results by comparing them with those of other data sources (censuses or sample surveys).

It also notes the importance of ensuring that the statistical outputs and services resulting from the use of administrative records meet the following criteria:

- that the statistical outputs respond to the conceptual scheme envisaged in the design phase;
- that the output formats, dissemination media and support elements (explanatory notes and expressions “include” or “exclude”, among others) facilitate consultation and correct interpretation of the statistical information; and
- that the data and metadata published satisfy the principles of confidentiality.

The reality of the region's countries is considered in terms of the quality of administrative records, establishing indicators of under-coverage and over-coverage, for example.

I. CHAPTER IX. DISSEMINATION

This chapter considers the relevance of the statistical information disseminated to the population at large, which should have the corresponding metadata.

In accordance with what was established in the design phase, it is essential to define and design statistical outputs that meet users' needs, including anonymized microdata for public use.

J. CHAPTER X. EVALUATION

This chapter discusses the importance of collecting the necessary information and preparing an action plan and an evaluation report that integrates the results, diagnostic studies and recommendations obtained from the assessment of a set of issues associated with the execution of the previous phases of this process. In particular, it recommends evaluating, among other issues, whether the statistical records generated from administrative records:

- meets the existing statistical or geographic information needs;
- is comparable and consistent with the results obtained from other sources of information or even by other methods;
- fulfils the principles of confidentiality and information security;
- reduces costs or informants' workloads; and
- complies with the international standards and subprocesses defined for source selection and quality assessment, data standardization, record linkage and transformation of administrative records into statistical records.

It also recommends evaluating whether:

- the IT solution implemented is aligned with the established standards, and
- the actions established to promote quality in the entities holding administrative records have been carried out.

K. CHAPTER XI. CONCLUSIONS AND RECOMMENDATIONS

This chapter presents the document's recommendations for implementation by the different countries of Latin America and the Caribbean.

L. ANNEXES

The countries' experiences are presented in eight annexes, which cover the separate phases of the Generic Model of Statistical Institutional Processes:

- Annex 1: For the needs specification phase, several examples of case studies from Costa Rica are presented, referring to each subprocess of this phase.
- Annex 2: For the design phase, the practical experience of Cuba in integrating the statistical business registry is presented.
- Annex 3: As an example of the processing phase, Cuba's experience in integrating an inventory of administrative records for statistical purposes is described.
- Annex 4: For the construction phase, two experiences of the National Institute of Statistics (INE) of Chile are presented, associated with two subprocesses: construction of data collection instruments and construction or improvement of components.
- Annex 5: For the processing phase, Peru's experience of processing the information collected with a unique code is described. Before the information is considered validated, it is subjected to validation tests for variables such as economic activity, company address or geographic location.
- Annex 6: For the analysis phase, Brazil's experience in coordinating services and commerce with structural surveys by firm in the adoption of GSBPM is presented.
- Annex 7: For the dissemination phase, Brazil's experience with subnational initiatives for using fiscal and administrative records in statistical production is presented.
- Annex 8: For the evaluation phase, Argentina's experience with questionnaires that make it possible to monitor and, at the same time, compile information from each phase for the comprehensive evaluation of the model is presented.

M. GLOSSARY

A glossary has been added to facilitate the standardization of concepts among the countries of Latin America and the Caribbean and to facilitate the exchange of economic statistical information.

N. BIBLIOGRAPHY

In addition, a bibliography is included to support the application of the methodology proposed for the use of administrative economic records.