



International Recommendations for Energy Statistics (IRES)

Chapter 7 – Data sources and collection strategies

Secretaría de Energía (SENER)
Instituto Nacional de Estadística y Geografía (INEGI)
MEXICO

contents

- A Introduction**
- B Data Sources**
- C Data compilation methods**
- D Data collection strategies**
- E Scope and coverage of data collection**
- F Institutional framework**
- Conclusions**

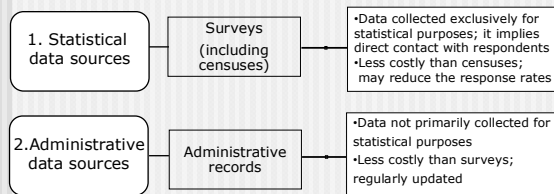
A. Introduction

- The presentation has been worded according to the version of Chapter 7 raised from the VM2 and also the complete version of IRES
- In the draft still remain certain confusion regarding the terms 'data collection' and 'data compilation'

B. Data sources

The energy data sources refer to origins, characteristics and components of the raw data used for Collecting energy statistics from both the **supply** and **demand** side for each energy commodity.

They require a set of sources and their proper strategies and methods, and can be categorized into two basic categories



B. Data sources

1. Statistical surveys

(a) Household surveys and mixed household-enterprise surveys:

- Not designed specifically for energy data collection
- Their metadata may be useful for estimates and other measurements
- High variation due to changes in prices, technologies and fuel availability
- Representativeness not only nationally but also in rural and urban areas and regions

(b) Enterprises surveys and Business register:

Enterprise surveys: the availability of a sampling frame is assumed

Business registers: it represent advantages for the energy data compilation

B. Data sources

(c) Ad-hoc energy statistics surveys:

- Allow to cover the lack of information and gaps from earlier surveys
- High designing and implementation costs
- Participation of the NSO or an academic national or international institute and stakeholders
- Reliability: being capable of being repeated periodically
- Data must cover the weights or volumes of different fuels

B. Data sources

2- Administrative data sources

Administrative data sources collect data in response to a legislation and/or regulation. For example to:

- (i) monitor of diverse forms of intervention or activities carried out by diverse agents related to products and/or services;
- (ii) enable regulatory activities and audit actions; and
- (iii) target outcomes of management policies, programmes and instruments

The register and data are referred to by collectively the statistical office as administrative data

B. Data sources

Advantages in the use of administrative data:

- ✓ reduction of the overall cost of data collection;
- ✓ development of statistical systems within agencies; and
- ✓ possible use as a framework for statistical surveys

Other:

- ✓ reduction of the response burden;
- ✓ there is a smaller errors than those arising from sample survey;
- ✓ minimal additional cost and long-term accessibility;
- ✓ regularly updated

B. Data sources

When using administrative data, special attention is given to their limitations and a description of these is given in the metadata.

Possible limitations include:

- inconsistencies in the concepts and definitions;
- discrepancies in the use of statistical units;
- possible breaks in the time series because of changes in regulations/legislations; and
- legal constraints with respect to access and confidentiality

B. Data sources

Legal constraints with respect to confidentiality when using administrative data (see also Chapter 10)

Each country must have proper mechanisms and procedures to ensure confidentiality

Lack of enforcement to guarantee confidentiality leads to incomplete and imprecise data reports from the agencies

It is important that compilers of energy statistics identify and review the available administrative data source in their countries

C. Data compilation methods

Data compilation: operations performed on data to derive new information according to a given set of rules, that is the statistical procedures used for producing intermediate data and final statistical outputs.

There are several data compilation methods:

- Identify the particular needs of the sector;
- Identify data sources in order to exploit the available data from existing surveys, censuses and administrative registers;
- Define whether there are intermediates or not;
- Use an appropriate classification and precise definition;
- Define the validation criteria

C. Data compilation methods

The data compilation process covers:

- the use of weighting schemes,
- methods for imputing missing values or source data,
- statistical adjustment,
- balancing/cross-checking techniques and relevant characteristics of the specific methods applied

It is necessary to analyze:

- the international recommendations,
- the conceptual and methodological infrastructure,
- the availability and quality of information, and
- to study the laws and regulations applicable to the matter being addressed

D. Data collection strategies

Countries to be able to collect **production**, **supply**, **transformation** and **consumption** statistics, for each fuel and for energy total.

When data can not be collected from existing sources, it is necessary to design strategies to collect this information, according to the following steps:

- Identify needs and establish the specific goals;
- Select the categories and variables;
- Select the target population or sample;
- Design a clear questionnaire or format to be filled
- Test the questionnaire or format under a similar context;
- Define the validation criteria;
- Other...

D. Data collection strategies

Most common and most critical sources of error:

- incomplete coverage,
- non-response,
- out of range responses,
- multiple responses,
- inconsistencies or contradictions, and
- responses to questions that do not apply

These errors may be caused by:

- deficiencies in the questionnaires design,
- lack of proper training to the interviewers,
- difficulty from the interviewee to provide the required data,
- issues related to the processing of the data

D. Data collection strategies

Three groups of units must be distinguished:

- Energy supply industries,
- Other industries/organizations producing energy, and
- Energy consumers

This distinction is important as data collection methods may be different for each group of units.

D. Data collection strategies

(a) Energy supply industries:

Producers, distributors and international traders directly related to energy

They usually "concentrate in one particular fuel, and often in only one part of the overall supply" chain.

These energy supply companies can be grouped according to their public-private status:

- a) Private industries belonging to a particular person or group;
- b) Public industries managed by the government; and
- c) Public-private industries

D. Data collection strategies

(b) Other industries and organizations producing energy (supplying energy is a secondary activity):

Industries producing energy for **self-consumption** and supply it to **other consumers**, but they must not be omitted from national energy statistics

These companies could have not the same amount of detailed information readily or be obliged to provide data

There is a trend towards an increase in their participation in the energy sector, and it is directly related to both **energy savings** and the **efficient use of energy**

Usually, the **auto-production** requires a government's permission, which facilitates the monitoring of these companies and the gathering of the required data

D. Data collection strategies

(c) Energy consumers:

Final energy consumers can be grouped according to the economic activity: (industry, transport, agriculture, residential, commerce and public services, forestry and fishing)

The sectors can be divided into more specific categories

Given the diversity, mobility and multipurpose forms of this consumers, it is necessary to classify them and to design specific methodologies and collection strategies

However, to fill the remaining gaps and have more detailed information, direct consumer surveys and censuses could be done

D. Data collection strategies

(d) A programme for data collection:

The development of an overall data collection strategy has to adapt to the policy needs and the resources availability of a country

An approach is to exploit the available data from surveys and administrative registers from which useful data on energy final use can be derived

Steps for data collection:

- (a) to sketch the production, supply and consumption flow charts for each fuel
- (b) to outline the data sources for each stage of the flow
- (c) to close collaboration between owners or compilers of the statistical surveys and energy statisticians

D. Data collection strategies

Collection of energy data as a part of regular Programmes of Surveys:

This Programme is directed to:

- obtain a wide coverage of accurate, detailed and timely energy statistics, and
- cover those industrial and other organizations that are engaged in supplying energy as a secondary activity

3 main methodological sources (from the demand side):

- (a) Censuses: Economic, Population, Agriculture,
- (b) Surveys: Establishments, Households, mixed household-enterprise surveys; and,
- (c) Administrative Records: Energy sector and others.

E. Scope and coverage of data collection

The scope and coverage of the data collection are defined according to:

Conceptual design: It must take into account the type of statistics to collect, the units of measurement, international standards, the target population

Geographical coverage: For policy purposes it is fundamental to collect statistics at the national level. However, to compile energy statistics at a sub national level which implies a better coverage of the information

In the consumption side, there is necessity for regional disaggregation since the energy use could vary according to climate, local behaviour and customs, economic activities, etc.

E. Scope and coverage of data collection

Frequency of data collection: it represents a balance between the availability of data and resources

Annual data should be the first target, but it is preferable to have 3 types of frequency: annual, infra annual or infrequent

Annual data collection: Should comprise the energy statistics surveys that collect the basic and most appropriate data for information needs

Infra annual data collections (quarterly, monthly etc.) are necessarily more restricted in scope than those carried out annually

Infrequent data collections: are generally conducted for specialized topics or to fill in remaining data gaps or to provide baseline information relating to biofuels

E. Scope and coverage of data collection

Point in time submission: the statistics collected could have a cyclical or seasonal behaviour, making a big difference in the analysis

General recommendations:

- The quality principles for the production of official statistics are followed for energy statistics;
- Encouraged transparency in principles and, in a decentralized statistical system, legal foundation for coordinating official energy statistics;
- The quality dimensions should be clearly articulated and promoted

F. Institutional framework

Legal principles guiding an integrated approach to the development of data collection strategies in energy statistics

(1) each participant is enforced to provide good quality energy data, through well designed channels and instruments that ensure the consistency and harmonization of energy statistics,

(2) responsibilities for collection, compilation and maintenance of the different data components among different government bodies are fixed

F. Institutional framework

National Systems of Energy Statistics: a comparison and recommendations for improvements

- A NSES is crucial for a sectorial analysis and for the design and implementation of public policies on the matter
- It integrates, in one central database, information scattered among the different entities or agencies involved in the energy sector
- Their concepts and methods ought to be harmonized with international standards
- The NSES could be managed by the ministry of energy, the national statistics office or another specialized agency

F. Institutional framework

Principles guiding effective institutional arrangements on compilation of energy statistics

- ❖ great potential for coordinating data collection, reduce response burden and improve timeliness of the data;
- ❖ to ensure the integration of data as concepts, definitions and methods

An institutional arrangement:

- (a) It defines roles and responsibilities for the parties involved, and the cooperation and mutual agreement
- (b) The data compilation strategies, concepts and methods are regulated
- (c) The collaboration between the energy statisticians and the energy policy community is also an important
- (d) Close collaboration between owners or compilers and energy statisticians is also essential to enhance the use of information

Conclusions

- ❖ Collection: is related to production of information, and Compilation is related to use of information (as it can be inferred from the *UN Energy Statistics: A Manual for developing countries*, Series F, No. 56, 1991)
- ❖ A definition of Data Collection must be included into the Glossary
- ❖ Suggestion: the title of the chapter 7 should be as *Data sources, collection and compilation strategies*