

Chapter 6 Statistical Units and Data Items

Section A: Statistical Units

*Below is a provisional text for the section A of Chapter 6 - Statistical Units. The text is based on the similar text adopted by the UN Statistical Commission for other areas of economic statistics. This section aims at describing the concepts of statistical units and provides recommendations on which statistical units to use for energy statistics. Please review and suggest amendments to better reflect needs and priorities of energy statistics. The proposed recommendations and encouragements are in **bold**.*

I. Statistical units

6.1. One of the preconditions for effective compilation and interpretation of energy statistics as well as for ensuring comparability of energy data with other statistics both nationally and internationally is a clear definition of the statistical units used in the statistical process. The universe of entities engaged in production, transformation and consumption of energy is very vast. It varies from the small entities engaged in one or very few activities undertaken at or from one geographical location to large and complex entities engaged in many different activities that may be carried out at or from many geographical locations. These entities vary in their legal, accounting, organizational and operating structures and have different abilities to report data.

6.2. *Collection of data from economic entities.* Economic statistics in general and energy statistics in particular have to take into account those structures in order to compile data which are most useful for economic analysis. However, legal and operational structures of economic entities as well as their record keeping practices are not developed in most countries to suit statistical purposes. Therefore, it is desirable to have guidelines on definition of statistical, reporting and collection units for use in data collection so that comparable national and international statistics can be produced.

6.3. A *statistical unit* is an entity about which information is sought and for which statistics are ultimately compiled. It is the unit at the basis of statistical aggregates and to which tabulated data refer. These units can be divided into two categories:

(a) *observation units* – identifiable legal/organizational or physical entities which are able, actually or potentially, to report data about their activities;

(b) *analytical units* – entities created by statisticians (also referred to as statistical constructs), often by splitting or combining observation units in order to compile more detailed and more homogeneous statistics than it is possible by using data on observation units. Analytical units are not able to report data themselves about their activities, but there exist indirect methods of statistical estimation including imputation of such data.

6.4. *Collection and reporting units.* A collection unit is the unit *from which* data are obtained and by which statistical forms are completed. In fact, it is more a contact address or contact person than a unit. Any entity which possesses statistically relevant information about statistical

units can, potentially, serve as a collection unit. For, example, if entities leave the form filling to a bookkeeping office, the latter is the collection unit. *Reporting units*. A reporting unit is the unit *about which* data are reported by the collection unit.

6.5. It should be noted that in practice the statistical, reporting and collection units often coincide with each other. For example, the information office of a power plant reports data about its production of electricity. That power plant is the unit which reports data (collection unit) as well as the unit about which data are reported (reporting unit) and about which data are sought by the energy statistics compiler (statistical unit). However, this is not the case if the company is a complex enterprise with production units located in different parts of the country (see paras 6.39-6.41 below for details).

6.6. Energy statistics compiler should distinguish between these units because they apply to different stages of data collection/compilation process. In the context of these recommendations statistical units are of primary interest as they are the basis of statistical aggregates and to which all data items refer. Collection and reporting units are especially relevant in the sampling and data collection stage. If an economic entity is engaged in several kinds of activity and if a separate statistical unit is associated with each of them, this entity is statistically complex. Provided that this unit (or units) is (are) able to supply all the required data, no additional data collection is necessary with respect to the entity as a whole. However, if the complete set of data cannot be obtained in this way, the data collection should target the entity as a whole as well and, if successful, apportion additional information to the statistical units. While doing this it is important to ensure that the reported data does not contain double counting.

II. Definition of selected statistical units

6.7. The full discussion of the statistical units relevant for various areas of economic statistics is available in *International Recommendations for Industrial Statistics* (IRIS) and in introduction to *International Standard Industrial Classification of All Economic Activities* (ISIC) and is not reproduced here. **It is recommended that** those documents are used for reference. Below are description of several statistical units which are important in the context of energy statistics.

6.8. *Enterprise*. An economic entity in its capacity as a producer of goods and services is considered to be an enterprise if it is capable, in its own right, of owning assets, incurring liabilities and engaging in economic activities and in transactions with other economic entities. An enterprise is an economic transactor with autonomy in respect of financial and investment decision-making, as well as authority and responsibility for allocating resources for the production of goods and services. It may be engaged in one or more productive activities at one or more locations.

6.9. *An enterprise group* is an association of enterprises bound together by various types of links such as ownership, controlling interest and management. An enterprise group can have more than one decision-making centre, especially for the policy on production, sales and profits and may centralize certain aspects of financial management and taxation. The enterprise group unit often corresponds to a conglomerate bound together by a network of complex relationships and frequently covers a very wide range of activities. Often, different subgroups can be identified within an enterprise group. There are some difficulties in the use of such a

unit for statistical purposes, in particular because of the problems of identifying and keeping track of sometimes unstable links between different enterprises.

6.10. *Establishment.* An establishment is defined as an enterprise or part of an enterprise that is situated in a single location and in which only a single productive activity is carried out or in which the principal productive activity accounts for most of the value added. Although the definition of an establishment allows for the possibility that there may be one or more secondary activities carried out, their magnitude should be small compared with that of the principal activity. If a secondary activity is as important, or nearly as important, as the principal activity, then the unit is more like a local unit, that is, an enterprise which engages in a productive activity at or from one location.

6.11. In the case of most small and medium-sized businesses, the enterprise and the establishment will be identical. In general, **it is recommended that** large enterprises engaged in many economic activities that belong to different industries be broken into one or more establishments, provided that smaller and more homogeneous units can be identified for which data on energy production or other activities attributed to energy sector may be meaningfully compiled.

6.12. *Kind-of-activity unit.* Based on the way the enterprise unit is constructed and defined, it may already have a certain degree of homogeneity with respect to its economic activities; however, energy statistics may require an even higher degree of homogeneity. It is for this purpose that the kind-of-activity unit has been defined and is used. A kind-of-activity unit is an analytical statistical unit and is defined as an enterprise or part of an enterprise that engages in only one kind of productive activity or in which the principal productive activity accounts for most of the value added. There is no restriction placed on the geographical area in which the activity is carried out. In order to create such homogeneous units, the enterprise must be partitioned into narrower, more homogeneous parts.

6.13. *Unit of homogeneous production.* It is defined as a production unit in which only a single (non-ancillary) productive activity is carried out. It is characterized by a single activity which is identified by its homogeneous inputs, production process and outputs. Most units are engaged in a combination of activities at the same time. If a unit carries out a principal activity as well as one or more secondary activities, it will be subdivided into the corresponding number of units of homogeneous production, and the secondary activities will be reclassified under different headings from the principal activity. Units of homogeneous production are normally not observable and are more an abstract or conceptual unit underlying the symmetric input-output tables as it may not be possible to collect directly from the enterprise or establishment data corresponding to units of homogeneous production. In practice, data on units of homogeneous production are calculated/estimated by transforming the data supplied by establishments or enterprises on the basis of various assumptions or hypotheses.

6.14. *Households.* Households are important energy consuming units. In some cases, household may also produce energy products for sale or for own use. A household is defined as a group of persons who share the same living accommodation, who pool some, or all, of their income and wealth and who consume certain types of goods and services collectively, mainly housing and food. In general, each member of a household should have some claim upon the

collective resources of the household. At least some decisions affecting consumption or other economic activities must be taken for the household as a whole¹.

III. Classification of statistical units by kinds of economic activity

6.15. To compile statistics by economic activity classes countries have developed national activity classifications. The UN Statistical Commission recommended that ISIC be used as the basis for such classifications. To facilitate international comparisons, **it is recommended that** countries following another activity classification develop full correspondence of their national classification with ISIC at 2-digit level. **This recommendation should be fully implemented in energy statistics** in view of a paramount importance of ensuring availability of internationally comparable energy data.

6.16. Each statistical unit can engage in a number of activities. The activities are classified into principal, secondary and ancillary activities.

6.17. *Principal and secondary activities.* In general, the *principal activity* of a unit can be determined from the goods that the unit produces or the services that it renders to other units or consumers. It is the activity that contributes most to the value added of that unit, or the activity the value added of which exceeds that of any other activity of the unit. It is not necessary that the principal activity account for 50 per cent or more of the total value added of a unit. A *secondary activity* is an activity carried out by a unit in addition to the principal activity and whose output, like that of the principal activity, must be suitable for delivery outside the unit. The value added of a secondary activity must be less than that of the principal activity. In general, most units have at least some secondary activities.

6.18. An example of secondary activity is the generation of electricity which can be sold or not in establishments engaged in the production of sugar cane products (sugar, alcohol, molasses, etc.). The fibrous by-product of the sugar cane after the cane juice is extracted, called bagasse, is usually incinerated in small thermal electricity plants. The electricity generation is considered as secondary activity provided that the value added of the electricity produced is less than that of the principal activity.

6.19. *Ancillary activities.* A productive activity undertaken with the sole purpose of producing one or more common type of services for own use within the same enterprise is defined as an ancillary activity. Such units are normally classified in accordance with the principal activity of the enterprise to which it belongs. However, this may not be desirable from the point of view of energy statistics. For example, a separate establishment may be generating electricity as ancillary activity of the enterprise to which it belongs. If such an establishment is (1) statistically observable, in the sense that separate accounts for the production it undertakes are readily available, or (2) if it is in a geographically different location from the establishments it serves, it may be desirable and useful to consider it as a separate (statistical) unit and allocate it to the corresponding energy producing activity.

6.20. To determine the kind-of-activity, including whether or not the unit belongs to the energy sector, of the statistical units in terms of ISIC Rev. 4, the following general principles should be followed:

¹ 2008 SNA para 1.149

(i) The kind-of-activity of a statistical unit is determined by the kind of its principal activity. Secondary and ancillary activities (except when units undertaking ancillary activities are identified as separate establishments) are to be disregarded when classifying a unit;

(ii) If the unit is engaged in several types of independent activities, but the unit itself cannot be segregated into separate statistical units, its kind-of-activity should be determined according to the ISIC class with the largest share of value added by using the “top-down” method. The “top-down” method means that first the appropriate highest classification level (1-digit) should be determined, then the lower (2- and 3-digit) levels and finally the class (4-digit level).

(iii) In cases where the value added cannot be determined for the activities involved, the principal activity should be determined using other criteria provided these are applied consistently overtime to different activities involved. The following alternative criteria **are recommended**:

(a) based on output: output of the unit that is attributable to the goods or services associated with each activity; and value of sales, shipment or transfers to other establishments of those groups of products falling within each activity;

(b) based on input: wages and salaries attributable to the different activities; or employment in the activities according to the proportion of people engaged in the different activities of the unit.

6.21. Instances may arise where considerable proportions of the activities of a unit are included in more than one class of ISIC. These cases may result from horizontal or vertical integration of activities. Such cases will require special attention.

6.22. The principal activity of producer units may change from one statistical period to the next, either because of seasonal factors or the management decision to vary the pattern of output. This change in activity would require the correspondent change in the classification of the unit. Frequent changes however, need to be avoided as it may render statistics on energy sector of an economy difficult for interpretation. In this connection countries **are encouraged** to develop a *stability rule*. Without such a rule there would be apparent changes in the economic demography of the business population which would be no more than statistical artefacts. The **recommended working rule** is that the secondary activity should exceed the activity to which the unit is classified for two years before the activity classification is changed. Similarly, if a unit engages in a mix of activities that are almost balanced, raising the risk of changes for the principal activity, the ratio of activities over the past two to three years should be taken into account for determining the principal activity.

6.23. **It is recommended that** countries change the activity classification of units for the purpose of statistical inquiries no more than once a year, either at a fixed date or as the information becomes available. More frequent changes would result in inconsistency between infra-annual and annual statistics.

IV. Characteristics of statistical units

6.24. Characteristics of statistical units are data items used for their unique identification, their classification within particular activity area of energy and for description of various aspects of their structure, operation and relationship with other units. Availability of

information on characteristics of the statistical units is a precondition for an effective organization of the statistical sample surveys as well as for comparisons and links to be made between data from different data sources thus, significantly reducing the duplication in data collection and response burden.

6.25. The main characteristics of the statistical unit are: its identification code, location, kind of activity, type of operation, type of economic organization, type of legal organization, type of ownership, size, and demographic characteristics. The list is not exhaustive but those characteristics represent the most important ones from the viewpoint of international comparability, as well as those considered to be of significant national interest.

1. Identification code

6.26. The identification code is a unique number assigned to a statistical unit which may comprise digits identifying its geographic location, kind-of-activity, whether a unit is a principal producing unit or an ancillary unit, link to its subsidiaries/principal if any etc.. The unique identification of statistical units is necessary in order to: (i) allow their registration in statistical business register or their inclusion in the sampling frame; (ii) permit the collection of information about them via administrative sources; (iii) provide a sampling base for statistical surveys; and (iv) permit demographic analysis of the population of units. The identification code must not change throughout the life of the unit, although some of the other unit's characteristics may change. Common identification codes, shared with administrative authorities and other government departments greatly facilitate the statistical work, including the connection of the statistical business register, if such is established, with other registers.

2. Location

6.27. The location is defined as the place at which the unit is physically performing its activities, not as where its mailing address is. This characteristic serves two important purposes. First, to identify the units and to classify them by geographical regions, at the most detailed level as demanded by the statistical programme. Second, if a unit operates in more than one location, to allocate its economic activity to the location in which it actually takes place. The latter is important for regional analyses. Since the classification of units by location is of particular national interest, any geographical classification should distinguish the major economic regions or administrative divisions of the country ranging from large areas (states or provinces) to intermediate areas to local areas (towns). **[Is this relevant also for energy statistics?]**

6.28. The details about mailing address, telephone and fax numbers, e-mail address and contact person are also important identification variables since these details are used for mailing the statistical questionnaires, written communication with the unit or making ad-hoc queries about its activity. Up-to-date information about any changes in those variables is crucial for the efficient work of statistical authorities.

6.29. *Location in case of multi-establishment enterprises.* Where an enterprise has more than one establishment; it may or may not have one location and address. Often, the enterprise address is used for administrative purposes and the establishment address for statistical purposes. There is a need, however, for care when dealing with large complex enterprises. **It is recommended that** the multi-establishment enterprise be requested to provide location details about each establishment it has, or the establishment may be asked about the name and location

of the enterprise that owns it so that a data set in the register on the enterprise and its own component establishments can be established. In some cases, it may be necessary to correspond with both the establishment and the enterprise because in general, the unit supplying for example employment details is different from one providing financial details.

3. Kind-of-activity

6.30. The kind-of-activity is a variable defined as the type of production in which a unit is engaged and should be determined in terms of ISIC, Rev. 4. The breakdown of economic activities for statistical units in the energy sector should be more detailed as described in Chapter 5.

4. Period of operation

6.31. This indicates the period during which the establishment has been in operation during the reference period. It would be useful to seek information under the following alternative items: (a) in operation since (date), important for instance for determining the electricity installed capacity as of a determined date; (b) temporarily or seasonally inactive, useful for example to track refinery shut-downs which might explain decrease in annual refinery throughput/output; (c) ceased operation (date), also important for determining installed capacity; and (d) sold or leased to another operator (name of new operator), which might explain changes in electricity capacity/production between main and self-producers. Besides the information that this characteristic provides about the activity status of the unit (active or temporarily inactive), it also helps in interpreting the returns made by statistical units that are affected by seasonal factors and those made by statistical units that began or ceased operations during the reference period. Most of such information lies on the level of metadata and is useful for data quality checks.

5. Type of economic organisation

6.32. The enterprise and the establishment are the main units used by countries for conducting industrial surveys. The characteristic “type of economic organisation” is intended to indicate whether the establishment is the *sole establishment* of the enterprise of immediate ownership or is a part of a *multi-establishment enterprise*. If further details are required on this aspect of the industrial structure, the multi-establishment enterprises might be divided into classes according to the number of their constituent establishments or by the criteria used for classifying establishments (employment, value added) that are most appropriate for each country.

6.33. For the purpose of accurate measurement of energy production and other energy flows and for compilation of various energy indicators, it is desirable to have the links between individual establishments and their parent enterprise clearly defined. More importantly, these links are fundamental for the efficient sampling design because one survey might gather information on value added, employment and production statistics usually available at establishment level, while another may collect data from consolidated financial statements compiled mainly at the enterprise level.

6. Size [is it important for energy statistics?]

6.34. A size measure of a statistical unit is an important stratification characteristic, essential for sample design and grossing up techniques. In general, the size classes of statistical units can

be defined in terms of physical units (e.g. employment) or in monetary units (e.g. turnover, amount of net assets). Monetary criteria can be used separately or in conjunction with employment criterion.

6.35. In energy statistics there may be the need to define two size measures depending on the main objective of the analysis. To study the production/generation of energy it may be more appropriate to define the size of an establishment in terms of the maximum capacity of the establishment to generate energy products. This, however, may not be applicable to all energy products.

6.36. To study the consumption of energy products, it may be more appropriate to measure the size of a unit by employment for establishment and number of persons for households.

[to be further discussed.]

The choice of the appropriate size measure of a unit for energy statistics should be discussed. This is linked also to data collection strategies and may need to be decided together with Chapter 7]

7. Type of operation

6.37. Type of operation refers to the type of conversion process the energy products undergo within the unit (e.g. coke oven, blast furnace, etc.). **[more input is needed for further development of this topic]**

V. Recommendations on statistical, reporting and collection units

6.38. Countries **are encouraged** to use the establishment as a statistical unit for energy statistics in view: (1) of the clear need for the most accurate data on production, transformation and distribution of energy; (2) in recognition of the recommendation to use the establishment as the most appropriate statistical unit for production and employment data and to ensure the compilation of homogeneous and geographically distributed data. The use of this unit is especially important for the provision of data needed for various kinds of macroeconomic analysis including compilation of input-output tables.

6.39. In the majority of cases, the establishment and the enterprise are the same, hence all types of data can be obtained from the same source. In such cases, an establishment/enterprise can be not only a statistical unit but also a reporting and collection unit. However, if an establishment is a part of a multi-establishment enterprise, it may not have access to all the necessary (for example, financial) information. Under these circumstances, the enterprise to which a given establishment belongs may serve as a collection unit which provides data about activities of that establishment to the statistical authorities.

6.40. Countries **are encouraged** to collect data for all relevant establishments belonging to a multi-establishment enterprise within a country. In the case of enterprises operating branches in economic territories of several countries, special care should be taken to ensure that data reflecting activities of establishments that are resident units in other economies not be included in energy statistics of the compiling country. If country does not have capacity to do so it should be clearly explained in the energy statistics metadata.

6.41. If a sufficient degree of homogeneity and desired geographical distribution can be obtained by other means, or data at the establishment level are not available, the enterprise can

be used as the statistical unit and countries may limit their data-collection activities to the enterprises whose main activity falls within boundary of energy sector. Some countries aiming to achieve maximum possible homogeneity and detailed geographical distribution may find it practical to use the establishment as the statistical unit for the collection of non-financial data items while using the enterprise for the collection of financial data.

6.42. The kind-of-activity unit differs from the establishment in that there is no restriction in respect of the geographical area in which a given kind of activity is carried out by a single legal entity. In certain instances, the availability of data on a kind-of-activity-unit basis may warrant the utilization of this unit rather than the establishment in energy surveys. For example, in some cases, data on fixed capital formation, inventories and sales may be more easily available in respect of kind-of-activity units but not in respect of establishments; at the same time, interest in the classification of the data according to area or size of establishment may be minimal. More generally, the kind-of-activity unit may be considered, for many purposes, a suitable alternative to the establishment in those countries where the larger multi-establishment enterprises organize their records on this basis. However, if the kind-of-activity unit is utilized in such cases, it would be useful to indicate the relationship between these units and the units used in other surveys.

VI. Statistical units of the informal sector

6.43. *Informal sector.* In many countries economic units of informal sector are not only final consumers but also producers of energy. In this connection it is important that countries define statistical units relevant to energy statistics of this sector.

[To be developed, please advise]