



Eurostat / IEA DSD for Energy statistics

InterEnerStat, Paris 14/12/2017

Bart De Norre, Eurostat

DSD Energy statistics

- Supply-Consumption chain:
 - **Joint annual data collections (JAQ) of IEA-Eurostat-UNECE; monthly data collections (joint, IEA, Eurostat)**
 - **IRES and SIEC**
 - **International / Standards / Classification(s)**
- SDMX guidelines:
 - **THE DESIGN OF DATA STRUCTURE DEFINITIONS**
https://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_DSDs_1.0.pdf
 - **THE CREATION AND MANAGEMENT OF SDMX CODE LISTS**
http://sdmx.org/wp-content/uploads/SDMX_Guidelines_for_CDCL.doc

DSD Energy statistics

Joint development Eurostat - IEA

- IEA had done work in 2009-2012
- Eurostat requested IEA in May 2015 to take up together DSD development
- Series of video conferences / mail exchanges
October 2015 – July 2016
- Explanatory document + Inventory code lists
- Testing DSD mapping with some questionnaires on-going
- Gradual verification/update
(revised monthly coal, new joint IEA monthly electricity, Revised JAQ 2017)

DSD Energy statistics

Concepts in the energy domain (of the scope defined)

- **Definition** (concept code name and description)
- **Role** (dimension, primary measure, attribute)
- **Level** (attribute relevant at observation, series or dataset level)
- **usage status** (mandatory or conditional/optional attributes)
- **code list or format**

DSD Energy statistics

Dimensions / structural principles (SDMX guidelines)

- 1. Parsimony:**
no redundant dimensions for identifying
- 2. Simplicity:**
keep identifiers short / keep number of dimensions low
- 3. Purity:**
dimensions relate to one pure concept, not to a combination
- 4. Density and sparseness**
("not available" values in the dimension combinations)
- 5. Unambiguousness**
(avoid one observation to be expressed by multiple combinations of dimension values (keys))
- 6. Exhaustiveness**
(includes every piece of information that is required to unambiguously represent a data point and to correctly interpret it outside its usual context)

DSD Energy statistics

7. Orthogonality:

independence of the meaning of a value of one dimension from the values of any other dimensions

8. User friendliness

(While a simple DSD consisting of a few dimensions only may be easier to understand by a human data consumer, a more complex, but purer DSD is typically more flexible in terms of further usage in automated processes.)

9. Fitness for use throughout the entire statistical business process

- Re-use concepts / code lists
(frequency, observation status)
- Extensible for potential future needs

DSD Energy statistics

- Further design principles:
 - **designed independently of the layout or technical features of existing EXCEL questionnaires and database structures in place at IEA and Eurostat**
 - **one DSD which is based on a clear logical model and flexible enough to cover all data and metadata from all concerned questionnaires**
- "Remarks" sheets
(Attribute "COMMENT_OBS"; explanatory "free text" to future MSD)
- Link to for future dissemination – energy balances
- SDMX code list guidelines:

DSD Energy statistics

14 DIMENSIONS (identifying concepts)

1. QUEST_SOURCE
2. REF_AREA
3. TIME_PERIOD
4. FREQ
5. ENERGY_PRODUCT
6. MAIN_FLOW
7. FLOW_BREAKDOWN
8. PLANT_TECH
9. PLANT_TYPE
10. STOCKS
11. INFRASTRUCTURE_IND
12. VIS_A_VIS_AREA
13. MEASURE_VALUE_TYPE
14. FACILITY_ID

DSD Energy statistics

REF_AREA

- **SDMX promotes one code list across SDMX domains**
- **Based on the one of National Accounts**

ENERGY_PRODUCT

- **all primary and secondary energy products or commodities and their aggregates as used in the energy questionnaires (and energy balances)**
- **Align to SIEC**
- **Complete SIEC based code list**
- **Align to SDMX code list guidelines**
- **Codes based on SIEC hierarchical numbering**

DSD Energy statistics

MAIN_FLOW and FLOW_BREAKDOWN

- **two-level hierarchical approach according IRES**

- **18 codes for MAIN_FLOW**

1. Production
2. Imports
3. Exports
4. International marine bunkers
5. International aviation bunkers
6. Stock changes
7. Transfers
8. Supply

9. Transformation

10. Net production (of electricity or heat)
11. Gross production (of electricity or heat)
12. Consumption
13. Energy use
14. Losses
15. Non-energy use
16. Final consumption
17. Statistical differences

- **Around 110 codes for FLOW_BREAKDOWN**

DSD Energy statistics

MAIN_FLOW and FLOW_BREAKDOWN

- Many main flows are split in more detailed flows
 - There are multiple electricity and heat production flows because of all possible energy input, plant technologies and plant types.
 - Stock changes apply to a broad diversity of types of stocks
 - Import and export flows are detailed by the country/region from where is imported resp. the country/region to where is exported
- **For some detailed flows additional dimensions are needed:**
- **PLANT_TECH**
 - **PLANT_TYPE**
 - **STOCKS**
 - **VIS_A_VIS_AREA.**

DSD Energy statistics

PLANT_TECH:

- technologies used in plants for production of electricity and / or heat
- This code list is not a straight hierarchical classification:
different perspectives are used in classifying/grouping power and heat plants
(product based, single/multi-fired, and technical type of generation)

COMBFUEL - Combustible Fuels; HYDRO - Hydro (all, unspecified)

PLANT_TYPE:

- main classification of electricity and heat plants

MAINELEC - Main Activity Producer Electricity Plants

INFRASTRUCTURE_IND:

- A number of data in the questionnaires describe infrastructure characteristics

GROSSCAP - Gross capacity (of electricity and/or heat);

SOLARSUR - Solar collectors surface

DSD Energy statistics

MEASURE_VALUE_TYPE

several measurement concepts used in reporting of energy data values

ENERGY Measure of Heat or Electricity; NCV Net Calorific Value

VIS_A_VIS_AREA (COUNTERPART_AREA)

FACILITY_ID

an identifier key for the storage locations for gas and refineries (from the JAQ 2017 onwards)

DSD Energy statistics

OBS_VALUE

6 ATTRIBUTES:

UNIT_MEASURE

KT – Kilotonne; TJ_NCV - TeraJoule (NCV)

OBS_STATUS (SDMX Standard)

normal, missing, estimated

CONF_STATUS (SDMX Standard)

SUBMISSION (date of the submission of the questionnaire)

COMMENT_OBS (short free text related to one or more observations)

FACILITY_TYPE (types of gas storage facilities and refineries)

FACILITY_NAME

Annex: SIEC based code list ENERGY_PRODUCT

Extract

SIEC based for DSD	aggregation formula	SIEC		
		Section Division Group	class	SIEC label
C0000	C0000 = C0100 + C0200 + C0300 + C0370 + C0390	0		Coal
C0100	C0100 = C0110 + C0120	01		Hard coal
C0110		011	0110	Anthracite
C0120	C0120 = C0121 + C0129	012		Bituminous coal
C0121			0121	Coking coal
C0129			0129	Other bituminous coal
C0200	C0200 = C0210 + C0220	02		Brown coal
C0210		021	0210	Sub-bituminous coal
C0220		022	0220	Lignite
C0300	C0300 = C0310 + C0320 + C0330 + C0340 + C0350 + C0360	03		Coal products
C0310	C0310 = C0311 + C0312 + C0313 + C0314	031		Coal coke
C0311			0311	Coke oven coke
C0312			0312	Gas coke