

2022

Supply, transformation and consumption of electricity in European countries

(Open-source data analysis)

Sonali

Career Foundry

6/10/2023

Executive summary

During last decade, climate change-oriented policies have caused a significant transformation in the energy sector. Moreover, the recent war in Ukraine has led to a sharp increase in energy prices and significant unpredictability in energy markets. Due to the heavy reliance on Russian supplies before the invasion, euro area energy markets have been especially impacted by recent political events. Generating electricity from renewable sources today is the cheapest possible substitute and brings variable energy sources (wind and solar) into the basket of the EU electricity market. The low-carbon transition, independency from fossil fuels and providing a more stable network of energy supply, will require more focus on sustainable electricity networks. To have a better insight about the current situation and make some prediction about the future trends, this analysis provides an overview of the generation, transformation and consumption of the electricity across European countries using the open source data sets available in [European Union Open Data Portal](#). In the following the data sets selected for this analysis, a brief data profile and finally the main questions which are going to be answered in will be introduced.

Data source

The initial raw data sets which are selected for this project are listed below. It should be noticed that the number of rows and attributes are from the original data set and after data cleaning and extraction of required sections, these numbers will be changed. For example, in annual energy consumptions, all energy sources are involved but we just intend to extract the electricity consumption in all sections. Moreover, there are some attributes in these tables which are related to datacenter information (like dataflow, last update or freq.). These columns will be removed in data cleaning process.

Data set Name	Description	source	# Attributes	# Rows
nrg_cb_em	Supply, transformation and consumption of electricity - monthly data	https://ec.europa.eu/eurostat/databrowser/view/NRG_CB_EM/default/table	10	28423
nrg_cb_pem	Net electricity generation by type of fuel - monthly data	https://ec.europa.eu/eurostat/databrowser/product/page/NRG_CB_PEM	9	55073
ten00125	Annual energy consumption in households by type of fuel	https://ec.europa.eu/eurostat/databrowser/product/page/TEN00125	10	5423
ten00126	Annual energy consumption in transport by type of fuel	https://ec.europa.eu/eurostat/databrowser/product/page/TEN00126	10	6902
ten00127	Annual energy consumption in road transport by type of fuel	https://ec.europa.eu/eurostat/databrowser/product/page/TEN00127	10	5916
ten00128	Annual energy consumption in services by type of fuel	https://ec.europa.eu/eurostat/databrowser/product/page/TEN00128	10	5892
ten00129	Annual energy consumption in industry by type of fuel	https://ec.europa.eu/eurostat/databrowser/product/page/TEN00129	10	5892

Based on Reference Metadata of Energy statistics in Eurostat, the Statistical Office of the European Union, the following information regarding ethics, limitations and data collection and quality have been extracted. For more detailed information refer to the [online documentations](#).

Data collection

All data transfers to Eurostat come via Single Entry Point of data, which is currently implemented by EDAMIS. For national practices:

National competent authorities collect the data by using census, surveys, administrative data, association's data, customs data and by modelling including interpolation and extrapolation. Even estimates are occasionally submitted if real data is not available. The mixture of sources differs between countries.

Data limitation and Ethics

Net electricity generation is available monthly from 1/2016 till 06/2022 and supply and transformation data are monthly from 01/2008 till 06/2022. Energy consumption in each section is reported annually from 1990 and we don't have monthly data. The inflowing data are delivered by the reporting countries to Eurostat via separate dedicated questionnaires. For energy statistics, definitions, concepts and methodology are harmonized at international level. Eurostat copyright rules applies: *free re-use, both for non-commercial and commercial purposes*.

Quality assurance

Eurostat carries out quality tests, mainly on the coherency of the provided information. In addition, the questionnaires used for data transmission also have built-in coherency tests. Energy data are subject to several validation checks, among which the time series checks, the consistency & completeness checks, plausibility checks, energy transformation efficiency checks and other checks. If there are any doubts as regards data quality, Eurostat contacts reporting countries to provide necessary justifications or corrections.

Accuracy

Data on energy is submitted based on energy questionnaires employing a standardized methodology. Eurostat receives disaggregated data which are used to countercheck the results and to ensure consistency with the total amount of energy consumption. The accuracy of the basic data depends on the quality of the national statistical systems and may vary from country to country. Eurostat checks each data collection against analog elements in other data collection.

Timeliness and punctuality

1. Annual energy data are transmitted from countries to Eurostat 11 months after the end of the reference year.
2. For monthly energy data: 3 months after the reference period
3. For gas and for electricity (SEG-GAS and SEG-ELE): within one month following the reported month.

Coherence and comparability

Energy domain provides highly comparable annual energy statistics. Eurostat verifies, to the extent possible, if the reported data respect the prescribed methodology. The underlying data collection methods are however the responsibility of the reporting countries. The methodology for reporting energy statistics is harmonized for all European countries and is adopted by the United Nations Statistical Commission in 2011.

Data are checked in detail for internal consistency/coherence in full cooperation with the countries concerned. Details checks involve the consistency of time series, detection of outliers, efficiency of energy transformation processes and checks on utilization factors.

As a result, we can trust the validity and accuracy of available data, we just extract the required tables from original datasets and clean them according to our project purposes.

Data preparation

For 'nrg_cb_pem' data set the following data cleaning were done:

- Unnecessary columns were removed
- Column names were replaced by more understandable names
- Abbreviations and codes for country names and fuel types were substituted by complete names
- Data types were checked, and there were no mixed data type
- Accuracy and completeness showed there are one negative value and 289 null values in '*value(GWH)*'. All replaced with 0.
- Data rows for 'EU27_2020' & 'EA19' were removed from data set as they are summations for countries in EU and they are not helpful in our analysis.
- There were no duplicate rows
- The final data profile is as bellow:

Column name	Description	Data type	Structured?	Time variant?
Country	Name of European countries	Qualitative, nominal	yes	no
Time-period	The month & year of the measurement	Qualitative, ordinal	yes	yes
Fuel-type	The energy source for electricity generation	Qualitative, nominal	yes	No
Value(GWH)	Generated electricity in gigawatt-hour	Quantitative, continuous	yes	Yes
Initial dataset dimensions: 9×55073 , Initial dataset name: nrg_cb_pem				
Final dataset dimensions: 4×52347 , Final dataset name: elect_production				

For 'nrg_cb_em' data set the following data cleaning were done:

- Unnecessary columns were removed
- Column names were replaced by more understandable names
- Abbreviation and codes for country names and energy balance were substituted by complete names
- Data types were checked, there were no mixed data type
- Accuracy and completeness showed there are one negative value and 69 null values in '*value(GWH)*'. All replaced with 0.
- Data rows for 'EU27_2020' & 'EA19' were removed from data set as they are summations for countries in EU and they are not helpful in our analysis.
- There were no duplicate rows
- The final data profile is as following:

Column name	Description	Data type	Structured?	Time variant?
Country	Name of European countries	Qualitative, nominal	yes	no
Time-period	The month & year of the measurement	Qualitative, ordinal	yes	yes
Energy-balance	The type of flow in electricity energy balance	Qualitative, nominal	yes	no
Value(GWH)	Generated electricity in gigawatt-hour	Quantitative, continuous	yes	yes
Initial dataset dimensions: 10 × 28423 , Initial dataset name: nrg_cb_em				
Final dataset dimensions: 4 × 27009 , Final dataset name: elect_balance				

For ‘ten00125’ to ‘ten00129’ data sets the following data cleaning were done:

- Unnecessary columns were removed
- Column names were replaced by more understandable names
- Abbreviation and codes for country names, energy-type types and sectors were substituted by complete names
- Data types were checked, there were no mixed data type
- Accuracy and completeness showed there are some null values in ‘*value(KTOE)*’. All replaced with 0.
- Data rows for 'EU27_2020', 'EA19' & 'EU28' were removed from data set as they are summations for countries in EU and they are not helpful in our analysis.
- There were no duplicate rows
- All 5 data sets were merged to produce a unique table of energy consumptions in different sectors
- The final data profile is as bellow:

Column name	Description	Data type	Structured?	Time variant?
Country	Name of European countries	Qualitative, nominal	yes	No
Time-period	The month & year of the measurement	Qualitative, ordinal	yes	Yes
Sector	The section, where using energy (household)	Qualitative, nominal	yes	No
Energy-type	The type of energy consumed in each sector	Qualitative, nominal	yes	No
Value(KTOE)	Generated electricity in Kilo-Ton Oil Equivalent	Quantitative, continuous	yes	Yes
Final dataset dimensions: 5 × 28609 , Final dataset name: energy_consumption				

Questions to explore

- 1- What is the share of each energy resource in electricity production during recent years in different regions of Europe? What types of energy resources are more vital for European countries?
- 2- How is the energy balance in case of electricity energy for European countries? Which are main electricity exporters, and which are the main electricity importers?
- 3- How different types of energy (in this analysis with more emphasize in electricity energy) contribute in different parts of European lifestyle (household, services, industry and transportation)?
- 4- In case of low gas supplies which sources of energy are good candidates to replace it? Which aspects of consumptions will be better to be revised to save more energy?