Data Analytics | Data Explanation

Full Forms:

* GW = gigawatts
* TWh = terawatt-hours
* MW = megawatts

IMP Pointers:

* Energy and Electricity are two separate things.
  + Primary energy
    - Primary energy is the energy available as resources – such as the fuels burnt in power plants – before it has been transformed. This relates to the coal before it has been burned, the uranium, or the barrels of oil.
    - Primary energy includes energy that the end user needs, in the form of electricity, transport and heating, plus inefficiencies and energy that is lost when raw resources are transformed into a usable form.
* "Installed capacity"
  + refers to the maximum amount of electricity that can be generated by a power plant or a group of power plants under specific conditions. It is typically measured in megawatts (MW) or gigawatts (GW). This capacity represents the total potential output of the renewable energy source (like wind, solar, or hydro) if it were to operate at full capacity continuously.
  + In simpler terms, installed capacity is like the horsepower of a car's engine—it tells you the maximum power the engine can produce, but not necessarily how much power it actually produces all the time. The actual output can vary depending on factors like weather conditions, maintenance, and efficiency of the technology.

Data/1-country-renewable-production/**modern-renewable-energy-(consumption)**.csv

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Entity | Code | Year | Other renewables (including geothermal and biomass) electricity generation - TWh | Solar generation - TWh | Wind generation - TWh | Hydro generation – TWh |
|  |  |  |  |  |  |  |

Data/1-country-renewable-production/**modern-renewable-production**.csv

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Entity | Code | Year | Electricity from wind - TWh | Electricity from hydro - TWh | Electricity from solar - TWh | Other renewables including bioenergy - TWh |
|  |  |  |  |  |  |  |

Data/1-country-renewable-production/**renewable-share-energy**.csv

Title: Share of primary energy consumption from renewable sources, 2023

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Renewables (% equivalent primary energy) |
|  |  |  |  |

Data/1-country-renewable-production/**share-electricity-renewables**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Renewables - % electricity |
|  |  |  |  |

Data/1-country-renewable-production/**hydropower-consumption**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Electricity from hydro - TWh |
|  |  |  |  |

Data/1-country-renewable-production/**hydro-share-energy**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Hydro (% equivalent primary energy) |
|  |  |  |  |

Data/1-country-renewable-production/**share-electricity-hydro**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Hydro - % electricity |
|  |  |  |  |

Data/1-country-renewable-production/**installed-solar-pv-capacity**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Solar energy capacity - GW |
|  |  |  |  |

Data/1-country-renewable-production/**share-electricity-solar**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Solar - % electricity |
|  |  |  |  |

Data/1-country-renewable-production/**solar-energy-consumption**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Electricity from solar - TWh |
|  |  |  |  |

Data/1-country-renewable-production/**solar-share-energy**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Solar (% equivalent primary energy) |
|  |  |  |  |

Data/1-country-renewable-production/**share-electricity-wind**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Wind - % electricity |
|  |  |  |  |

Data/1-country-renewable-production/**wind-generation**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Electricity from wind - TWh |
|  |  |  |  |

Data/1-country-renewable-production/**wind-share-energy**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Wind (% equivalent primary energy) |
|  |  |  |  |

Data/1-country-renewable-production/**cumulative-wind-gigawatts**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Wind energy capacity - GW |
|  |  |  |  |

Data/1-country-renewable-production/**installed-geothermal-capacity**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Geothermal energy capacity - MW |
|  |  |  |  |

Data/1-country-renewable-production/**biofuel-production**.csv

|  |  |  |  |
| --- | --- | --- | --- |
| Entity | Code | Year | Biofuels Production - TWh |
|  |  |  |  |

Data/2-energy-production-and-consumption/electricity-generation.csv

The data is in time series format

* Entity: shows the countries and continents
* Code: code for countries not included for continents
* Year: the year the data was recorded
* Annual change in primary energy consumption: percentage change in primary energy.

Data/2-energy-production-and-consumption/electricity-generation.csv

The data is in time series format

* Entity: includes countries, regions and even associations like ASEAN(Ember) which are 11 south east asian countries.
* Code: code for countries not included for continents
* Year: the year the data was recorded
* Electricity generation - TWh: tells electricity generation in one trillion watt hours

Data/2-energy-production-and-consumption/global-energy-substitution.csv

* Entity: only contains the world (as 1 place)
* Code: code for the world which is OWID\_WRL

"Substituted energy" refers to an adjustment applied to energy sources like renewables and nuclear to estimate what their consumption would look like if they were as inefficient as fossil fuels.

The rest of the data calculates different types of energy produced

* Biofuels (TWh, substituted energy)
* Solar (TWh, substituted energy)
* Wind (TWh, substituted energy)
* Hydropower (TWh, substituted energy)
* Nuclear (TWh, substituted energy
* Gas (TWh, substituted energy)
* Oil (TWh, substituted energy)
* Coal (TWh, substituted energy)
* Traditional biomass (TWh, substituted energy)
* Other renewables (TWh, substituted energy)

Data/2-energy-production-and-consumption/per-capita-electricity-generation.csv

The data is in time series format

* Entity: shows the countries and continents
* Code: code for countries not included for continents
* Year: the year the data was recorded
* Per capita electricity - kWh: energy consumption per person

Data/2-energy-production-and-consumption/per-capita-energy-use.csv

The data is in time series format

* Entity: shows the countries and continents
* Code: code for countries not included for continents
* Year: the year the data was recorded
* Primary energy consumption per capita (kWh/person): calculates the average total energy consumed per person, including all forms of energy (fossil fuels, nuclear, renewables, etc.) before any conversion losses are considered.

Data/2-energy-production-and-consumption/primary-energy-cons.csv

The data is in time series format

* Entity: shows the countries and continents
* Code: code for countries not included for continents
* Year: the year the data was recorded
* Primary energy consumption (TWh): calculates the average total energy consumed per entity, including all forms of energy (fossil fuels, nuclear, renewables, etc.) before any conversion losses are considered.

Data/3-electricity-production-from-renewables/share-electricity-renewables.csv

The data is in time series format

* Entity: includes countries, regions and even associations like ASEAN(Ember) which are 11 south east asian countries.
* Code: code for countries not included for continents
* Year: the year the data was recorded
* Renewables - % electricity: renewable energy as a percentage of total electricity.

Data/4-gdp-data-world-bank/API\_NY.GDP.MKTP.CD\_DS2\_en\_csv\_v2\_9865.csv

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Country Name | Country Code | Indicator Name | Indicator Code | Years (1960 to 2023) |

The dataset columns and their explanations are as follows:

1. **Country Name**: The name of the country or region being analyzed.
2. **Country Code**: A three-letter code representing the country (ISO 3166-1 alpha-3 format).
3. **Indicator Name**: The name of the metric being measured, which is "GDP (current US$)" for all entries in this dataset.
4. **Indicator Code**: A unique identifier for the indicator, in this case, "NY.GDP.MKTP.CD," signifying GDP measured in current U.S. dollars.
5. **Years (1960 to 2023)**: Each of these columns represents GDP data for the respective year, showing the total market value of all goods and services produced within the country for that year, expressed in current U.S. dollars.

**Domain-specific keywords**:

* **GDP (current US$)**: Gross Domestic Product expressed in current U.S. dollars, reflecting the market value of all goods and services produced in a year without adjusting for inflation. Current GDP is useful for analysing economic growth over time within the context of nominal values.
* **Current US$**: A currency metric that reflects the value without adjustments for inflation or deflation, making it suitable for contemporary comparisons but less so for analysing trends over time without considering inflation.

This dataset is from the World Bank and is used to track and analyse the **Gross Domestic Product (GDP)** of various countries and regions over time, specifically measured in **current U.S. dollars**. It provides a historical overview of the economic size and growth of countries from 1960 to 2023.

The GDP metric reflects the total monetary value of all goods and services produced within a country's borders in a given year. By using current U.S. dollars, this dataset allows for a straightforward comparison of economic performance among different countries for specific years, although it does not adjust for inflation or purchasing power parity.

Data/4-gdp-data-world-bank/Metadata\_Country\_API.csv

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Country Code | Region | IncomeGroup | SpecialNotes | TableName |

The dataset columns and their explanations are as follows:

1. **Country Code**: A three-letter code representing each country (ISO 3166-1 alpha-3 format).
2. **Region**: The geographical region to which the country belongs (e.g., Latin America & Caribbean, Sub-Saharan Africa).
3. **Income Group**: The World Bank classification of a country's income level (e.g., High income, Low income, Lower middle income).
4. **Special Notes**: Additional information or context related to the country or region, such as economic assessments or unique characteristics.
5. **Table Name**: The formal or standard name of the country as it should appear in tables and reports.

This dataset if required for providing contextual information about various countries. It complements numerical data files by giving additional qualitative details such as the **geographical region**, **income classification**, and **special notes** relevant to each country. This information helps analysts and researchers better understand and categorize countries when working with related data sets (e.g., GDP, energy consumption, etc.).

Data/4-gdp-data-world-bank/Metadata\_Indicator\_API.csv

|  |  |  |  |
| --- | --- | --- | --- |
| INDICATOR\_CODE | INDICATOR\_NAME | SOURCE\_NOTE | SOURCE\_ORGANIZATION |

The dataset columns and their explanations are as follows:

1. **INDICATOR\_CODE**:
   * The unique code assigned to each indicator (e.g., "NY.GDP.MKTP.CD" for GDP in current US dollars).
2. **INDICATOR\_NAME**:
   * The descriptive name of the indicator (e.g., "GDP (current US$)").
3. **SOURCE\_NOTE**:
   * A detailed explanation of what the indicator measures and any important notes regarding its calculation or use.
4. **SOURCE\_ORGANIZATION**:
   * The organization or data source responsible for providing or compiling the data (e.g., World Bank national accounts data, OECD National Accounts).

Data/5-gdp-data-oecd/oecd-gdp-data.csv

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| STRUCTURE | STRUCTURE\_ID | STRUCTURE\_NAME | ACTION | FREQ | Frequency of observation | REF\_AREA | Reference area | SECTOR | Institutional sector |
| COUNTERPART\_SECTOR | Counterpart institutional sector | TRANSACTION | Transaction | INSTR\_ASSET | Financial instruments and non-financial assets | ACTIVITY | Economic activity | EXPENDITURE | Expenditure |
| UNIT\_MEASURE | Unit of measure | PRICE\_BASE | Price base | TRANSFORMATION | Transformation | TABLE\_IDENTIFIER | Table identifier | TIME\_PERIOD | Time period |
| OBS\_VALUE | Observation value | REF\_YEAR\_PRICE | Price reference year | CONF\_STATUS | Confidentiality status | DECIMALS | Decimals | OBS\_STATUS | Observation status |
| UNIT\_MULT | Unit multiplier | CURRENCY | Currency | UNIT\_MULT |

The dataset columns and their explanations are as follows:

1. **STRUCTURE**: Description of the dataset structure or data category (e.g., GDP and consumption data).
2. **STRUCTURE\_ID**: Unique identifier for the data structure (e.g., "OECD.SDD.NAD").
3. **STRUCTURE\_NAME**: Name of the dataset structure (e.g., "Annual GDP and consumption per capita").
4. **ACTION**: Indicates the type of action performed on the data (e.g., insertion, update).
5. **FREQ / Frequency of observation**: The frequency at which the data is collected or reported (e.g., annual).
6. **REF\_AREA / Reference area**: Code representing the country or region (e.g., "LVA" for Latvia).
7. **Reference area**: The name of the country or region.
8. **SECTOR / Institutional sector**: Specifies the economic sector (e.g., "Total economy").
9. **COUNTERPART\_SECTOR / Counterpart institutional sector**: Counterpart sector related to the transaction.
10. **TRANSACTION / Transaction**: Describes the type of financial or economic transaction recorded.
11. **INSTR\_ASSET / Financial instruments and non-financial assets**: Specifies whether the data involves financial or non-financial assets.
12. **ACTIVITY / Economic activity**: Indicates the type of economic activity covered.
13. **EXPENDITURE**: Specifies the expenditure category.
14. **UNIT\_MEASURE / Unit of measure**: Unit in which the data is measured (e.g., USD, constant prices).
15. **PRICE\_BASE**: The price base used for the data (e.g., constant prices).
16. **TRANSFORMATION**: Indicates any transformation applied to the data (e.g., seasonal adjustment).
17. **TABLE\_IDENTIFIER**: Identifier for the table or dataset from which the data originates.
18. **TIME\_PERIOD**: The year or time period when the data was recorded.
19. **OBS\_VALUE**: The actual value of the observation for the corresponding year.
20. **REF\_YEAR\_PRICE / Price reference year**: The reference year for constant price calculations.
21. **CONF\_STATUS / Confidentiality status**: Indicates if the data is confidential or free for publication.
22. **DECIMALS / Decimals**: Number of decimal places used in the data.
23. **OBS\_STATUS / Observation status**: The status of the observation (e.g., normal value, estimated).
24. **UNIT\_MULT / Unit multiplier**: Multiplier for the unit of measure (e.g., units, thousands).
25. **CURRENCY**: The currency used (e.g., USD).
26. **Currency**: Clarifies the currency applied or if it is not applicable.

This dataset contains detailed annual GDP and related economic data for different countries, structured according to the OECD's reporting standards.