**Household Electricity Prices in Europe**

**Abstract:**

**Keywords:**

**Project scope:** This project aims to analyse the evolution of end-consumer electricity prices in Europe over an extended period. The analysis involves the identification of trends for the whole region or country by country, the identification of anomalies in the price evolution that may be explained through past disruptive events such as the COVID-19 pandemic or the Ukraine war, and the use of multiple prediction algorithms in order to drive data-driven insights and recommendations for the future evolution of electricity prices in Europe. (Python was used + PANDA + tenserflow, etc.)

**1. Introduction:** At the base of this project stands the necessary data which needs to be analysed. A rather larger dataset with observations that span over a longer period of time for each or the majority of European countries is needed. Data crawling on the web was the designated method in order to gather an initial set of data. The folowing datasets were found and retreived form the web:

* Electricity prices by type of users [1]

Description: This indicator presents electricity prices charged to final consumers. Electricity prices for non-household consumers are defined as follows: *Average national price in Euro per kWh without taxes applicable for the first semester of each year for medium-sized industrial consumers* (Consumption Band Ic with annual consumption between 500 and 2000 MWh). Electricity prices for household consumers are defined as follows: ***Average national price in Euro per kWh including taxes and levies applicable for the first semester of each year for medium-sized household consumers*** (Consumption Band Dc with annual consumption between 2500 and 5000 kWh). [1]

* European wholesale electricity prices - monthly [2]

The mentioned dataset contains the prices of electricity in Europe, which are centralised in a monthly configuration. However, these are not the end consumer prices as they do not include taxes, levies, network charges, subsidies, and supplier profits. These are prices on what is called the spot market. [2] Unlike the previous dataset, which may be more relevant for this analysis, end-customer prices make a more relatable analysis.

* European Union Energy Market Data [3]

This dataset was collected from a Kaggle data source and it contains a large amount of data related to hourly updates on power princess across various systems. The purpose of the dataset is to be used for research and correlation insights as the European energy markets are highly dynamic due to factors such as renewable energy integration, supply-demand balance, and geopolitical influences [3]. What is interesting about this dataset is that it contains a categorization between fossil and renewable energy sources. It is not stated whether the prices contained in the dataset are end-consumer prices or prices at a different stage on the supply chain but for simplicity, we will consider them as end-consumer prices.

* Renewable energy share of total production - Europe [4]

This dataset is the result of a tool/methodology used by Eurostat to collect the mentioned data. The tool involves a standard for the calculation of the indicators related to the share of energy from renewable sources [4]. The dataset contains exactly the share of renewable energy for every European country starting from 2004 to 2023.

In the end, only the second dataset was used as the other four either contained a not big enough number of entries such as for the first and third or data out of the scope of the project, not related with the price dimension, such as the fourth dataset. Although the second dataset does not contain end-customer prices, the analysis focuses more on the evolution of this metric rather than its value. It could be argued that until the prices reaches the end customer the values of ratios for different countries may differ, however, in this project this was taken into account and the results were not present as fully conclusive.

**2. Data cleaning:** Data cleaning is required before any of the prediction algorithms are applied. Following this is a sample of the dataset. The datasets has 3504 observations, it contains monthly data for 31 countries starting with January 2015 and ending with December 2024. The sample contains the first 8 observations with the following features: ISO3 Code, Date and Price expresed in EUR/MWhe which is a unit that measures the power output of a power plant. For the scope of the project it can be refered to as MWh or megawatt per hour as is a more common term.

|  |  |  |  |
| --- | --- | --- | --- |
| Country | ISO3 Code | Date | Price (EUR/MWhe) |
| Austria | AUT | 1/1/2015 | 29.94 |
| Belgium | BEL | 1/1/2015 | 42.33 |
| Czechia | CZE | 1/1/2015 | 29.47 |
| Denmark | DNK | 1/1/2015 | 27.12 |
| Estonia | EST | 1/1/2015 | 33.84 |
| Finland | FIN | 1/1/2015 | 33.81 |
| France | FRA | 1/1/2015 | 40.94 |
| Germany | DEU | 1/1/2015 | 29.94 |

Astfel the dataset was introduced into the coding environment, to better evaluate the spread of data, Figure 1 was ploted. As it can be seen from this figure, disregarding the versatility of the plot, there is no observation with value 0 or missing and there are some observations with a flat evolution, i.e. Muntenegro and North Macedonia (Figure 2 & 3).

A graph of a graph

Description automatically generated

Figure 1. Dataset plotting for visualization

A graph with a line graph

Description automatically generated with medium confidence

Figure 2. Price evolution of Muntenegro

A graph with a line

Description automatically generated

Figure 3. Price evolution of North Macedonia

The decision for this observations was to remove them as they are also not relevant as main drivers of the price such as bigger European economies. Once those were remover the data was sorted by country name and then by the date in order to have them in a more categorized order. Moreover, it was noticed that 4 countries (Bulgaria, Coatia, Serbia and United Kingdom) were missing the price values for the first part of 2015. To use an uniform dataset, those values were filled using the average electricity price of that month for the whole countries, that way avoinding any biar or skewing of the data.

**3. Statistics:** As a prerequisite of the prediction algortihms it would be necessary to birefly understand the price evolution of electricity in Europe over the analysed period. Figure 4 displays the evolution of the computed average price over the analysed period.

A graph with a line

Description automatically generated

Figure 5. Average electricity price for all European countries.

It is evidently that the price mantained itself at a lower level belowe 60 euros while in the half of 2021 „in the wake of Covid-19 pandemic” the price started to increase considerably reaching values of around 175 euros at the start of 2022, when the Russia invasion of Ukrained lead to surging prices of 400 euros per MWh [#].

In [#] we learn that the sharp increase in energy import prices at the end of 2021, which more than doubled between December 2020 and December 2021, was driven by rising global demand and limited supply as economies recovered from the pandemic. This surge was unprecedented, as energy prices, despite their usual volatility, typically do not change by more than 30% in a year. The situation worsened in 2022 with Russia’s war on Ukraine and its suspension of gas supplies to some EU member states, which pushed gas and electricity prices to record highs. Summer heatwaves further strained energy markets, increasing demand for cooling while droughts reduced hydropower supply.

On the other hand, it can also be observed that prices have started to decrease around 2023 and got to a level more relatable to the period before this two disruptive events described.

Looking back at figure 1, it could be observed that when the price disruptions appeared some countries performed better than others. In order to visualize which ones were those the following box plot was drafted (Figure 6).

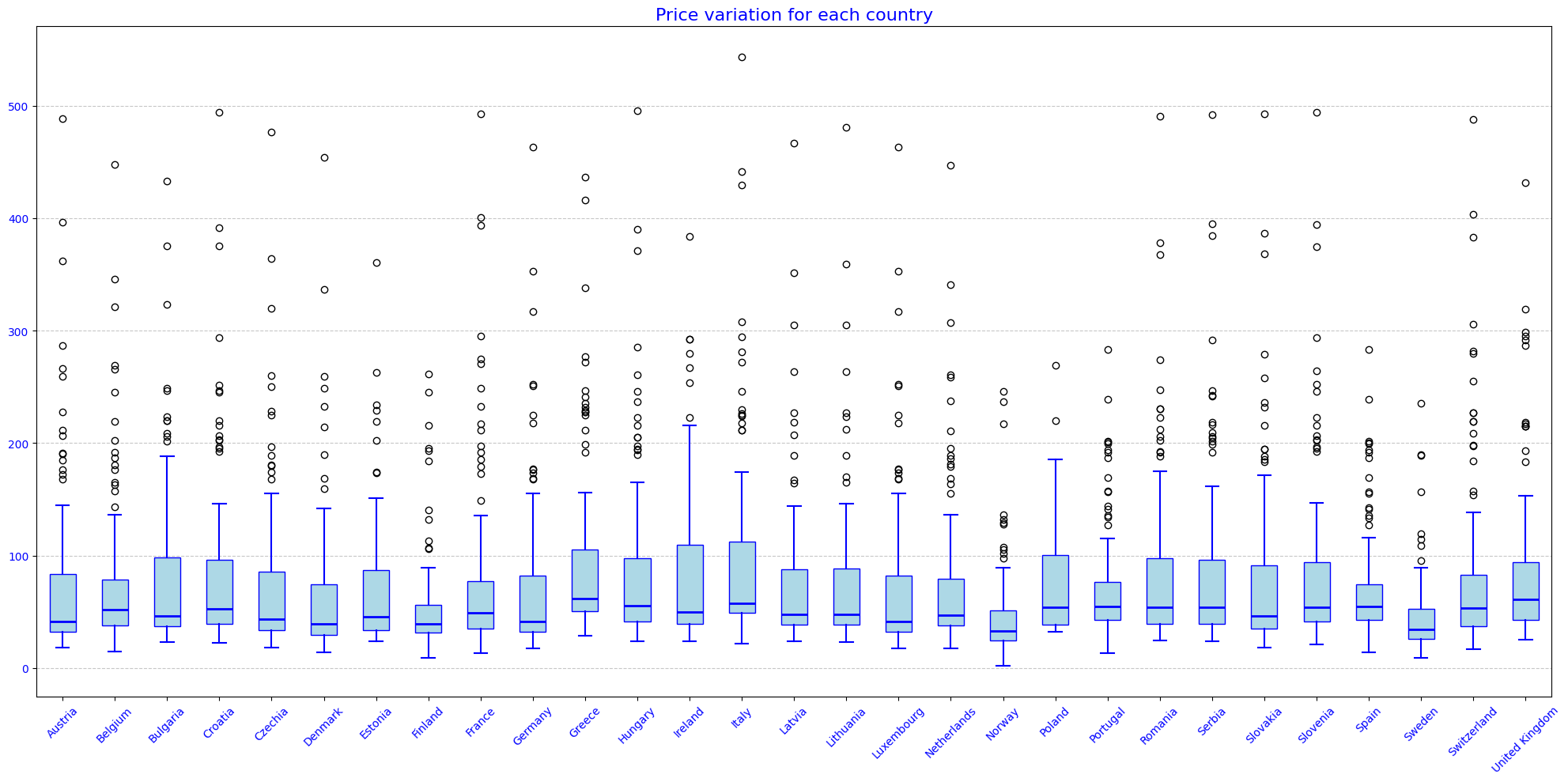


Figure 6. Price variation of each country

As it can be observed the mean values is rather similar for all countries. However, when looking at the outlier prices it can be spoted that some countries have much smaller values for upper outliers than other countries. This could be inderpreted as having more stability in times of crises such as the two disruptive events. For this reason, the variance value of the price metric for each country was calculated and ploted in Figure 7 in order to identify which are the countries that had the best price stability over the time period analysed.

A graph of blue bars with white text

Description automatically generated

Figure 7. Price Variance of each country

The top 5 most stable countries were Sweden, Norway, Finland, Poland and Spain. One potential reason for this stability may be that according to [#] these nordic countries as well as Spain, Estonia or Portugal has high and even the highest share of renewable energy as an energy source. Such an asset may be used to cope for the dependency of external suppliers of energy and keep the price at a more stable level than the markets.

To support again the point already made, Figure 8 and 9 show the difference in price for the most instable and stable countries in terms of price evolution.

A graph with blue and green squares

Description automatically generatedA graph with blue and white squares

Description automatically generated

Figure 8. Price statistics yearly for Italy

Figure 9. Price statistics yearly for Sweden

The difference can be best seen in the years 2021 and 2022 where the difference in prices are as much as 200 to 250 euros.

Following there is a final visualisation of the price variance for each of the analysed countries. The visualisation was created using SAS Viya Data Analytics platform and it shows again the countries with a better stability in the electricity price over the others.

A map of the world with different colored circles

Description automatically generated

Figure 10. Price Variance of each country

**4. Data prediction:**

**Methodology:** As mentioned, this research aims to analyse the evolution of electricity prices in Europe over an extended period. This analysis is primarily quantitative.

The data used for the analysis was collected from multiple sources and centralized in a main dataset which contains the necessary parameters. For this step data cleaning was required. The environment and programming language used was python and jupyter notebook. Furthermore the following libraries were used….

Describe prediction algorithm 1

Describe prediction algorithm 2

**Conclusions and Recomandations:**

**References:**

[1]. Eurostat Database. Electricity prices by type of user. Source:

<https://ec.europa.eu/eurostat/databrowser/view/ten00117/default/table?lang=en&category=t_nrg.t_nrg_indic>

[2]. European Wholesale Electricity Price Data. Source:

<https://ember-energy.org/data/european-wholesale-electricity-price-data/>

[3]. Kaggle. European Union Energy Market Data. Source:

<https://www.kaggle.com/datasets/pythonafroz/european-union-energy-market-data?resource=download>

[4]. Eurostat Database. Short Assessment of Renewable Energy Sources (SHARES). Source:

<https://ec.europa.eu/eurostat/web/energy/database/additional-data#Short%20assessment%20of%20renewable%20energy%20sources%20(SHARES)>

<https://www.eea.europa.eu/en/analysis/indicators/share-of-energy-consumption-from>

<https://www.consilium.europa.eu/en/infographics/energy-prices-2021/>