GENEDU5030 - Final Data Analysis Project

Summary Report - Group 9

Jeanne Pulvermüller - A12221101
Teresa Wan (萬庭妤) - B08107052
Chris Milne (陳閔樂) - T12703104

1 Data Set Introduction

Primary Data Set

The following report details the analysis and analysis results of the primary data set (**EuropeanLoadData_Hourly_20192020.csv**) relating to 24 European countries' respective electricity consumption over an 18-month period. This data set includes hourly observations from 2019/01/01 00:00 (midnight) until 2020/06/30 23:00.

A sample of the data set is provided below in **Figure 1**.

	A	B	C	D	E	F	G	H
1	year	month	day	areaname	mapcode	timehourly	totalloadvalue_hourly	
2	2019	1	1	Austria	AT	01.01.2019 00:00	23410	
3	2019	1	1	Austria	AT	01.01.2019 01:00	22476,801	
4	2019	1	1	Austria	AT	01.01.2019 02:00	21295,6	
5	2019	1	1	Austria	AT	01.01.2019 03:00	21093,199	
6	2019	1	1	Austria	AT	01.01.2019 04:00	21756	
7	2019	1	1	Austria	AT	01.01.2019 05:00	22066,4	
8	2019	1	1	Austria	AT	01.01.2019 06:00	23791,6	
9	2019	1	1	Austria	AT	01.01.2019 07:00	24814,801	
10	2019	1	1	Austria	AT	01.01.2019 08:00	26176,801	
11	2019	1	1	Austria	AT	01.01.2019 09:00	27536,801	
12	2019	1	1	Austria	AT	01.01.2019 10:00	28255,199	
13	2019	1	1	Austria	AT	01.01.2019 11:00	28246,4	
14	2019	1	1	Austria	AT	01.01.2019 12:00	28162,801	
15	2019	1	1	Austria	AT	01.01.2019 13:00	27751,199	
16	2019	1	1	Austria	AT	01.01.2019 14:00	27904,801	
17	2019	1	1	Austria	AT	01.01.2019 15:00	29028,801	

Figure 1 - Sample data from primary data set, "EuropeanLoadData_Hourly_20192020.csv"

It should be noted that this data set therefore includes an approximate 3-month period (March-June 2020) where the impact of COVID-19 pandemic lockdowns and restrictions are potentially evident in the electricity consumption observations.

Secondary Data Sets

To supplement the primary data set, several additional data sets containing key additional characteristics and metrics pertaining to the 24 European countries for which electricity usage observations have been recorded have been sourced.

The additional characteristics and metrics contained within these secondary data sets have - wherever possible - similarly sourced to be temporally contemporary to the primary data set; that is, accurate as in 2019-20.

EconDataPop_edit.csv

Contained within this data set are a variety of characteristics and metrics pertaining to 174 countries, including the 2019 population, Human Development Index (HDI), Consumer Price Index (CPI), and geographic region.

It should be noted that information and values for the United Kingdom was manually appended to this data set, to ensure completeness in the subsequent data set merge operations.

weather_data.csv

Contained within this data set are hourly weather observations (temperature in degrees Celsius, and direct and diffuse radiation readings) for 28 European countries, spanning the period from 1980/01/01 00:00 to 2019/12/31 23:00.

API_NV.IND.MANF.ZS_DS2_en_csv_v2_6299867.csv

Contained within this data set are annual measures of the percentage of a country's GDP (or clustered region's GDP) contributed by the manufacturing industry, as recorded for 266 countries and regions.

UnitPrice.csv

Contained within this dataset are semi-annual measures of the electricity prices (in EUR per kWh) for countries in Europe, recorded for the years 2019 and 2020. The data is presented in two halves for each year, denoted as S1 (first half) and S2 (second half).

2 Analysis Goal

The central goal of this analysis is to identify key contributing factors (or proxies for underlying factors) that influence the consumption and usage of electricity at a country level. Although the results of the analysis will technically be valid only for the European countries for which the analysis is conducted, it is nonetheless expected that the results may be applicable and informative beyond the European context.

With electricity production accounting for approximately a quarter of global greenhouse emissions, understanding and responding to usage demand drivers is becoming increasingly critical to governments globally. Furthermore, with many electricity grids experiencing shortages and faults, often compounded by ageing infrastructure and underinvestment in transmission and distribution networks, intervention and targeted investment to rectify these issues has generated a critical need for specific and contemporary supporting data.

As such, it is expected that the results of this analysis may be valuable to policy makers and infrastructure investment decision makers. These parties have the capacity to significantly influence the electricity demand and supply characteristics of populations, both at a national and local level, which is increasingly critical as the world pushes to decarbonise and transition away from fossil fuels.

3 Related Work

Numerous studies have explored the interplay between a population's energy consumption and various influencing factors. However, many of these studies have been limited in scope, often focusing on just one or two variables. In contrast, our research has taken a broader approach, examining a broader, diverse range of factors.

One noteworthy study [1] highlights temperature as a significant variable affecting energy consumption in developing countries, though its impact appears to be less pronounced in developed nations. This finding aligns with our observation that temperature does not play a consistently major role in the energy consumption patterns of European countries. Moreover, this study underscores the importance of Gross Domestic Product (GDP) in this context, revealing a bidirectional relationship between GDP and energy consumption. This suggests that as a country's economy expands, its energy demands increase, which in turn can stimulate further economic growth.

Similarly, another investigation [2], focusing on Taiwan's residential sector, corroborates the substantial influence of GDP on electricity consumption. However, it also posits that the impact of electricity prices on the overall electricity demand is relatively marginal.

In our study, we examined two geographical factors influencing electricity consumption: weather and historical urbanisation. Our findings indicate that while weather has a minor effect, the impact of urbanisation, shaped by a country's historical development, is notably more significant. This highlights urbanisation, rather than weather, as a key determinant in a nation's energy usage patterns. Several studies support this perspective. For instance, research [3] examining the Bangkok Metropolitan Area identified urbanisation, particularly in the industrial sector, as a critical determinant of energy usage. Additionally, another study [4] investigating MENA countries established a positive, bidirectional relationship between urbanisation and energy consumption.

In conclusion, our research contributes to the existing body of knowledge by offering a more comprehensive examination of the factors influencing energy consumption, particularly in European countries. This holistic approach not only corroborates findings from previous studies but also provides new insights into the complex interdependencies between economic, geographical, and social factors and their collective impact on energy consumption patterns.

4 Methods Applied

The data preparation (cleaning, merging, aggregation, etc.) and subsequent analyses were conducted primarily using Excel, PowerBI and PowerQuery.

Where data tables were merged in Excel, VLOOKUP() was heavily utilised, with country name or acronym generally used as the lookup variable.

In PowerBI (and PowerQuery), data merging processes at times merged on multiple variables (for example, the country acronym and the date-time value when merging the unpivoted weather data with the European electricity consumption data).

A key challenge faced was the size of the datasets being merged and operated upon, which saw significant reduction in computing performance.

Additionally, PowerBI was not readily able to superimpose multiple time periods on a common axis, instead only extending the time period as the variable over which the data was plotted. This limited the analyses between seasons (or even on the COVID-19 effect) due to poor visualisation options.

Visualisation limitations were a common challenge encountered during the analysis process, especially when looking to qualitatively assess the effect of various factors on the electricity consumption of a country, as the relative variable scales frequently obscured any potential relationship. To mitigate this issue, percentage scaling (relative to the other countries in the data sets) of variables was utilised, to allow the comparison of variables of significantly different scales.

5 Results

5.1 Broad Preliminary Analyses

Numerous preliminary exploratory analyses were conducted to determine the broad shape and coverage of the data set.

These analyses included the summation of the total electricity usage across Europe, aggregated quarterly, as shown below in **Figure 2**.

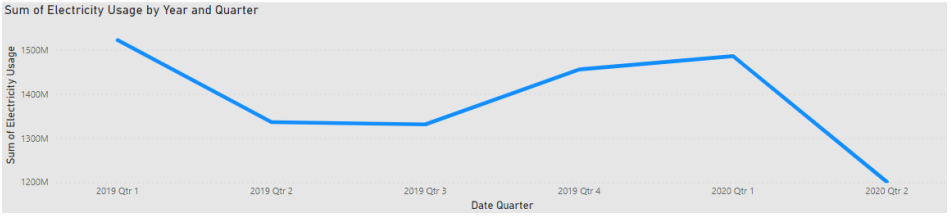

Figure 2 - Sum of electricity usage across Europe by quarter

From **Figure 2** (above), it can be noted that the winter quarters generally see elevated electricity usage across Europe. In addition, there effect of COVID-19 can be detected in the notably decreased usage during 2020 Q2, as compared to 2019 Q2.

Figure 3 (below) shows the country-by-country usage patterns. Given the magnitude of Germany's usage, the scale of the plot is dominated by Germany's usage.

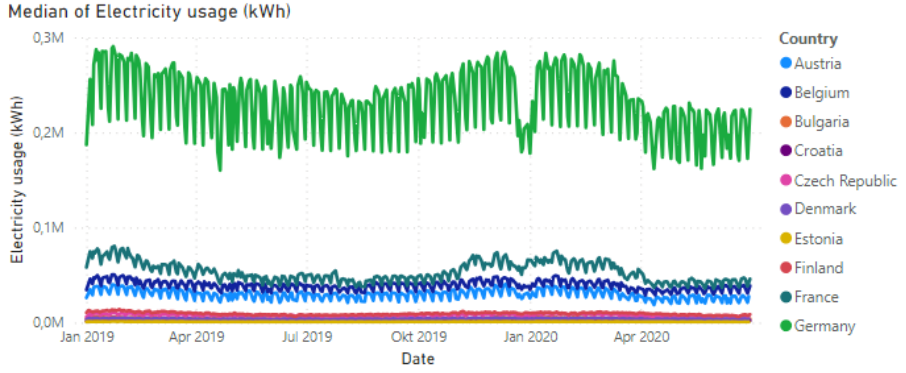

Figure 3 - Electricity usage patterns, by country

Furthermore, the evident weekly oscillations between the workdays and weekends indicates that the peak electricity usage tends to occur during the working week. Thus, the comparative electricity usage trough observed during the Christmas-New Year holiday (despite being during the otherwise elevated electricity usage period during the European winter) is explained by industry and businesses closing for the holiday season.

Some COVID-19 effects can also be seen at a country level, in the plot above. As such, consideration for the effect of this abnormal event on the integrity of the later data analyses should be given.

Adjusting a country's electricity usage for its population (as shown in **Figure 4** and **Figure 5**, below) makes the time series plotted in **Figure 3** more meaningful.

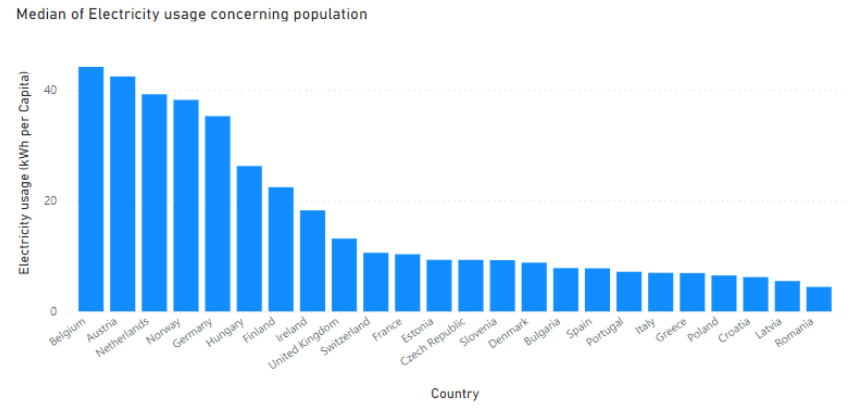

Figure 4 - Electricity usage per Capita, by country

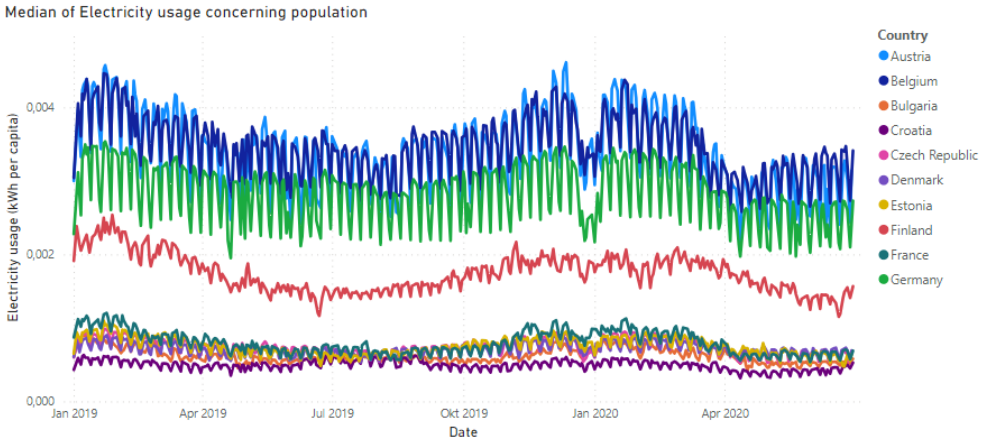

Figure 5 - Electricity usage per Capita patterns, by country

5.2 Factor Analyses

Investigation of potential contributing or explanatory factors influencing a country's electricity usage was subsequently conducted.

Firstly, the mean per Capita electricity usage of the sample countries was plotted geographically, in order to determine if there were evident clusterings - this output map graphic is shown below in **Figure 6**.

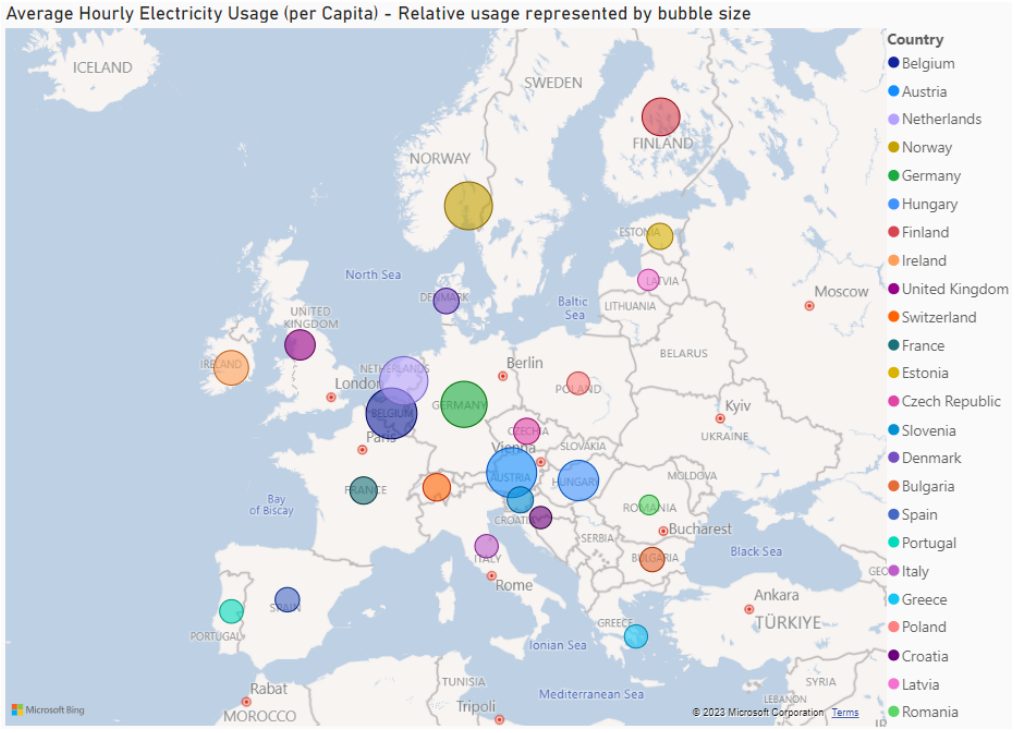

Figure 6 - Mean hourly electricity usage per Capita

From **Figure 6**, it can be observed that there is a broad trend for more northern, western European countries (with colder winter temperatures) tended to use electricity at an increased level compared to their southern and eastern European counterparts.

Additionally, the western European countries also have a higher level of wealth than many of their eastern European neighbours, which is partly due to the respective geopolitical history of the continent and the consequent ramifications on the development of independent and prosperous industries and economies.

Analysis of the relative Consumer Price Index (CPI) of the sample European countries saw the sample divided in half; the 'high CPI' and 'low CPI' clusterings. The per Capita electricity usage of these grouped clusters is shown below in **Figure 7**, which suggests that countries with a higher CPI generally consume more electricity per Capita.

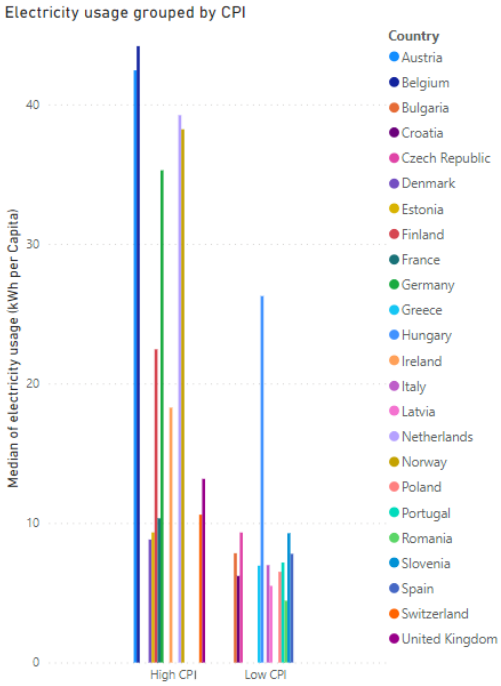

Figure 7 - Per Capita electricity usage clustered by high vs. low CPI

The CPI variable, as a measure of the cost of goods, and implicitly services too, suggests that the more expensive a country's living costs, the higher their energy consumption. This may be simply due to the greater economic output of a country (i.e., their GDP) being highly correlated with greater electricity consumption, as CPI is often regarded as a proxy for economic activity given its intrinsic connection to economic output. However, the CPI also captures the inherent impact on the cost of electricity (both as a unit price and as a quantity of consumption valuation), which is further analysed below in the *Unit Price* discussion.

Correlation between increased electricity usage and a greater amount of manufacturing output has been reported in both literature and recent media reports analysing the uneven impact of industry types on electricity usage. Consequently, information on the relative manufacturing contribution to a country's GDP was utilised to analyse the implications on the respective electricity usage of the country. The results are shown in **Figure 8**. Note, the y-axis has been scaled so that the relative energy consumption per capita amongst the sample countries is matched to the relative percentage of a country's manufacturing contribution to their GDP amongst the sample countries; that is, the relative scale of the bars provide qualitative information, but exact quantitative information cannot be determined due to significant variable scaling discrepancy.

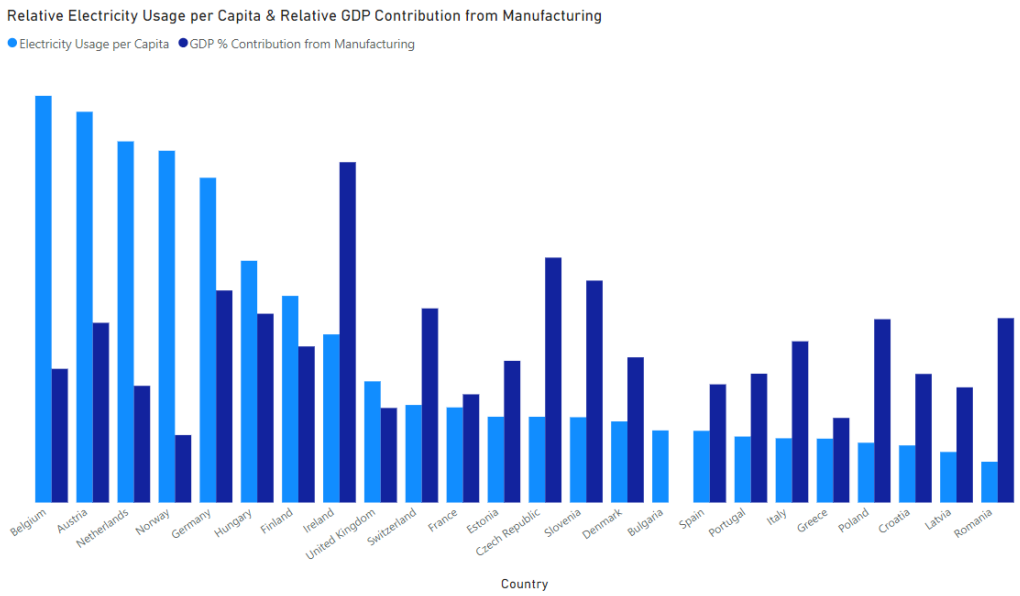

Figure 8 - Relative electricity usage per Capita and relative GDP contribution from manufacturing sector

As can be observed from **Figure 8** above, there is no definitive relationship between the relative size of a country's manufacturing sector and the respective per Capita electricity usage. This therefore implies that although the electricity usage of a country is heavily impacted by the 'work day' (as seen in **Figure 3** and **Figure 5**), it is not necessarily manufacturing and heavy industry that contributes predominantly to this usage in the sample European countries. Instead, electricity usage is likely by the climate control requirements for more general (and widespread) office buildings and company offices.

Considering instead a country's relative health, as measured by its Human Development Index (HDI), it can be determined that while a weak correlation between the health of the population and its per Capita electricity appears to exist (see **Figure 9**), this is not a pronounced relationship, and no inference around causality can be made.

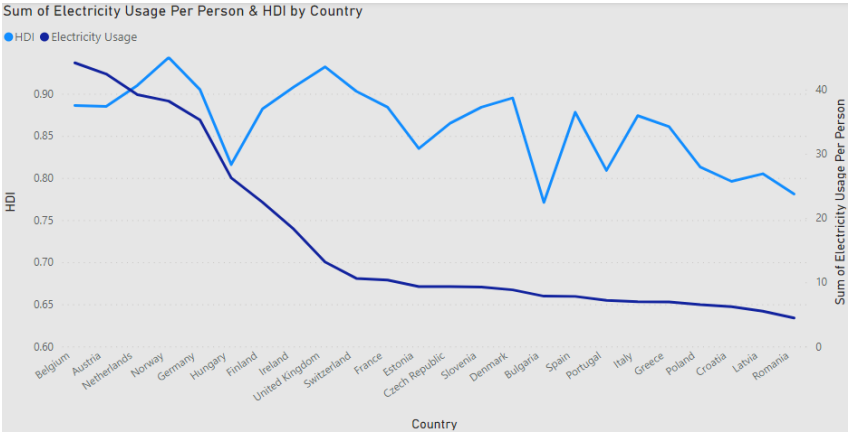

Figure 9 - Electricity usage per Capita and HDI

However, it may be hypothesised that the overall development of a country (in both public health infrastructure, but also in its education and private facility technological advancement) is a contributor to electricity demand. HDI may be regarded as a proxy metric of a country's development and investment in these areas. This effect becomes more pronounced when the European countries are clustered by geographic region (see **Figure 10**, below) - again, the historical geopolitical context that has affected countries' relative development is evident, at least in the most developed countries having higher electricity consumption patterns.

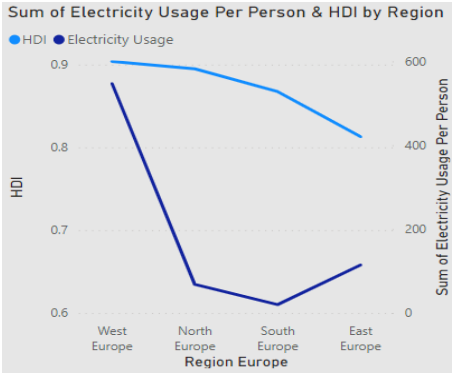

Figure 10 - Electricity usage per Capita and HDI, by geographic region

Where the CPI's value as a predictor for the electricity usage of a population (**Figure 7**), the nature of whether the implicit price of the electricity used in the production of goods which the price of which are then reflected in the CPI of a country was uncertain. Analysis of the unit price of electricity on the usage of electricity (**Figure 11**, below) suggests that the price of electricity has minimal discernible impact on the quantity of electricity consumed by a population.

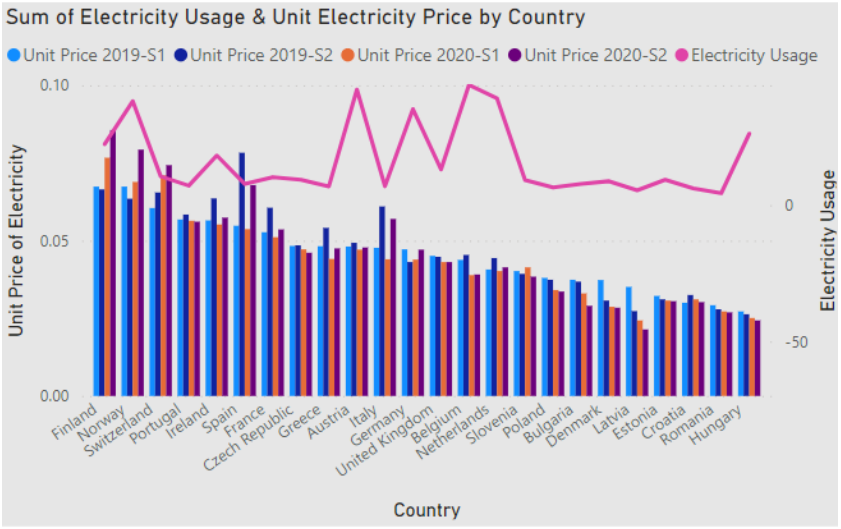

Figure 11 - Electricity usage per Capita compared with electricity unit price (2019-20)

Analysis of the effect of electricity unit prices on consumption when aggregated on a regional basis (**Figure 12**) again suggests little discernible impact. Perhaps the most critical insight available is that

the price of electricity is largely affected by the readiness and cost efficiency of its production, with Eastern European countries having notably cheaper access to electricity than the rest of the continent; however, this did not appear to either increase or decrease their electricity usage.

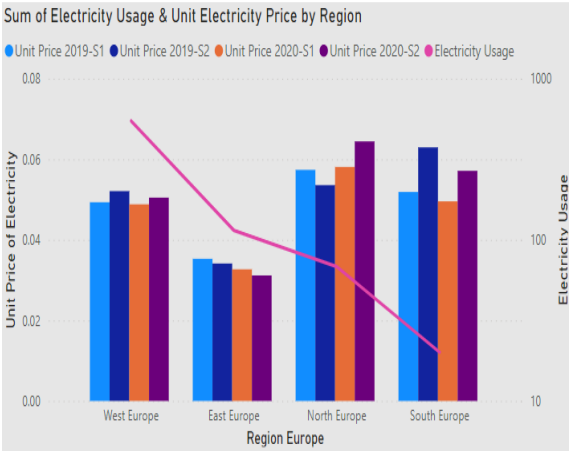

Figure 12 - Regional electricity unit prices compared with electricity consumption patterns

6 Conclusion

In summary, while the energy consumption of the European countries analysed are evidently influenced by a number of (often overlapping) factors, there is no one individual dominating factor than can be used to predict or explain the electricity usage of a country's population.

The strongest indicators of a country's electricity consumption patterns include the wealth and health of a country, with usage driven by business activity more than private consumption. Furthermore, the geographic context of a country - weather, urbanisation and historical geopolitical context - also significantly affect a country's electricity usage, however the effect is generally idiosyncratic, with individual situations and contexts overcoming any broader trend.

As a result, where governments are looking to amend energy policy or invest in new electricity infrastructure (or otherwise invest in maintaining or future-proofing existing infrastructure), location-specific analyses should be conducted.

Further research into the statistical causal relationships that may exist between electricity usage and the variables investigated graphically above should be pursued, even if the results of such analyses only provide a clearer ranking of variable import on consumption patterns, rather than a firm regression or predictive model.

References

[1] I. Dokas, M. Panagiotidis, S. Papadamou and E. Spyromitros, "The determinants of energy and electricity consumption in developed and developing countries: International evidence," *Energies*, vol. 15, no. 7, pp. 2558, Mar. 2022.

[2] Y. T. Chen, “The factors affecting electricity consumption and the consumption characteristics in the residential sector - A case example of Taiwan,” *Sustainability*, vol. 9, no. 8, pp. 1484, Aug. 2017.

[3] C. T. Nguyen, D. T. H. Nguyen and D. K. Phan, “Factors affecting urban electricity consumption: a case study in the Bangkok Metropolitan Area using an integrated approach of earth observation data and data analysis,” *Environ. Sci. Pollut. Res.*, vol. 28, pp. 12056-12066, Mar. 2021.

[4] U. Al-mulali, H. G. Fereidouni, J. Y. M. Lee, C. N. B. Che Sab, “Exploring the relationship between urbanisation, energy consumption, and CO2 emission in MENA countries,” *Ren. and Sust. Ener. Rev.*, vol. 23, pp. 107-112, Jul. 2013.