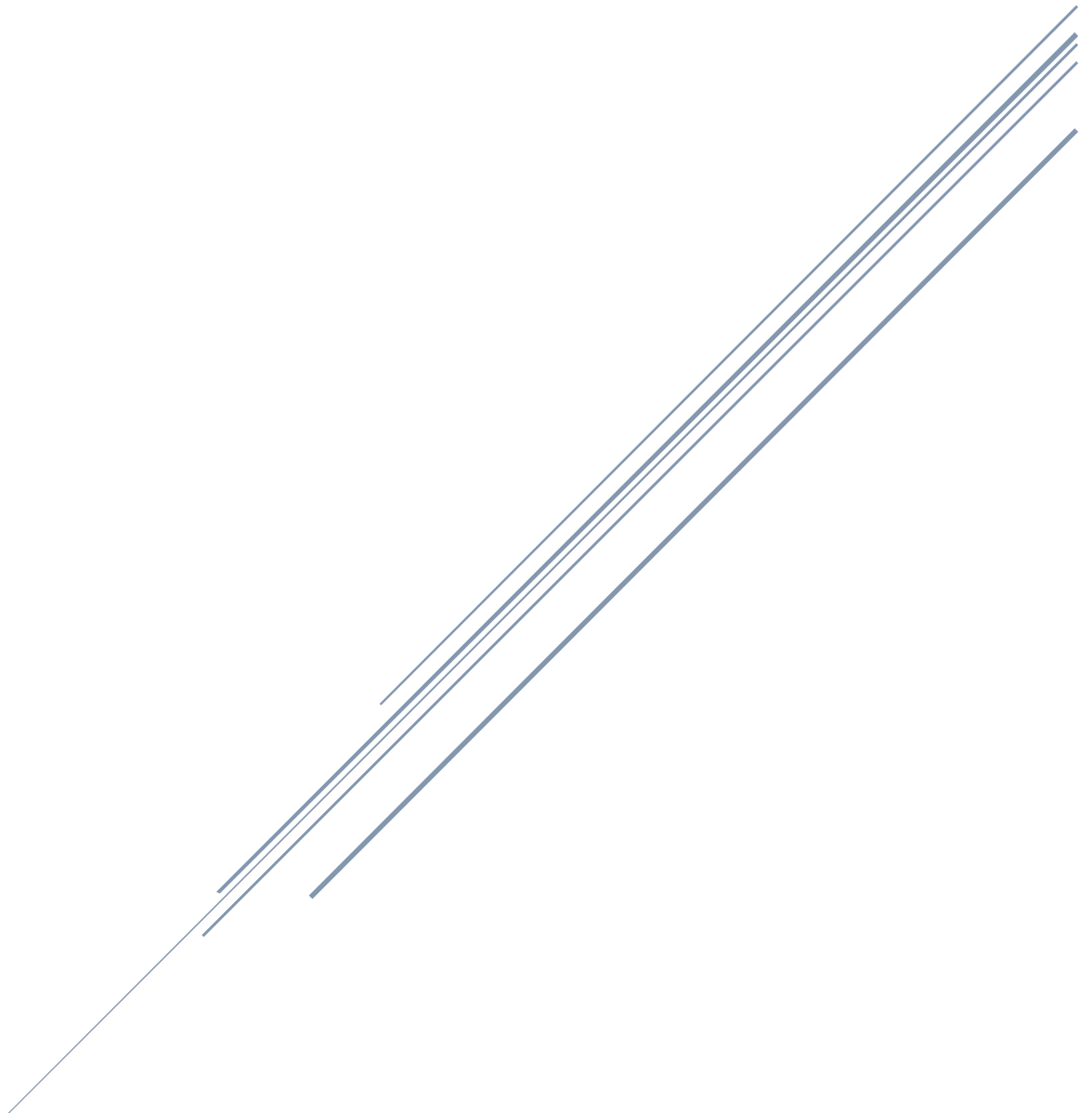# Fuel Usage, Electricity Generation by Fuel Types, Consumption of Electricity, and Prices of Electricity

Ciarán Crowley

20080223

19/Dec/2020
Brenda Mullally

Contents

What is contained within each sheet?

- The Fuel input sheet contains information the fuel input for electricity generation between the year 1920 and 2019. It records the total of all the fuels used in MTOE (Million Tonnes of Oil Equivalent)
- The Supply, Availability & Consumption sheet details how much electricity was supplied, purchased, available, lost, imported.
- The generated and supplied chart details how much electricity was generated via each method of electrical generation.
- The Capacity sheet details how much electricity is stored

What would I like to discover from the data set?

I would like to discover an increase in the amount of fuel input each year and the amount of electricity generated. I would like to see if small increases in fuel input led to much higher amounts generated.

3 questions I would like to answer

- Does the efficiency of fuel generation increase?
- Does the consumption of electricity increase or decrease with greater production?
- Does electricity get cheaper as more is produced?

Data Mining Methods

Data Cleaning and Preparation

Data cleaning and preparation is an important part of data mining. Raw data must go through the cleaning process and be prepared to suit various analysis methods. The cleaning and preparation process includes different components of data modeling, data integration, aggregation, transformation, ETL, and ELT. Data cleaning and preparation is required in order to gain an understanding of the basic attributes of the data in question. It is also the most basic step in data mining. This is the method I will be using to gather the relevant data for my questions

Classification

Classification involves analyzing the attributes related to different data types. Once the primary qualities of these data types are identified, data analysts can then sort the data by categories or classify the data.

Clustering

Clustering is a data analysis method that requires the analyst to have a visual on the data in order to gain an understanding of the data. Clustering makes use of graphs to display where the data is situated in relation to various metrics set out by the analyst. Making use of charts and graphs is often the best way for using cluster analytics. Analysts can see how data is spread out to discover trends that are relevant to their objectives.

Neural Networks

Neural Networks are a type of machine learning algorithm that sees heavy use in AI and deep learning. Neural Networks are made up of layers which simulate the way the neurons in our brains work. They are one of the most accurate methods of machine learning in use today. Although they are a powerful method of data mining, they are also incredibly complex.

Sequential Pattern Mining

Sequential Pattern Mining is a type of data mining where patterns between data sets are discovered and displayed in a sequence. It is most useful for data made up of transactions.

Decision Tree

Decision Tree mining is a data mining method used to build classification models. It builds these models in a tree like structure. It is best used in machine learning.

What Sheets and columns will I be using?

I will be using the Fuel Input sheet, the Generated and Supplied sheet, and the Electricity Prices Sheet. I have selected these sheets as I feel they contain the data that holds the best possibilities of answering the questions I want answered, which can be seen above. The columns I have selected are directly related to the questions I want to answer, and they can also be used to answer any extra questions I may have as I work through this assignment.

Fuel Input Sheet

In this sheet, I will most likely make some use of each column available in the sheet with the exception of the Coke and Breeze column. Each one provides valuable information. I will use this data to compare fuel inputs and electricity outputs in order to answer my first question. I will likely combine all the fuel inputs that can be classified as fossil fuels and compare them to alternative forms of energy production.

Generated and Supplied Sheet

In this sheet, I will make use of columns B, C, D, G, J, K, L, M. These columns provide the data which is most relevant to my second question. These columns tell me how much energy was produced, how much was used in the production of energy, how much was supplied to the grid, and how much customers used.

Electricity Prices Sheet

I will only be using column C in this sheet. I feel that this is the only column that is relevant to my third question. This column shows how much electricity cost per kilo watt.

Cleaning the data

There wasn't much data cleaning to do. In the Fuel Input Sheet, there was two rows dedicated to 1987. Both rows contained two different values. I didn't want to have two data sets for 1987 in Fuel Input, so I averaged the values for 1987.

Summary of Data

Above I have talked about what data I have chosen to use for this assignment. Here I will describe it.

From the Fuel Input Sheet, I have chosen to make use of all columns but one, the Coke and Breeze column. Initially, I didn't know what this was, and after some research, I decided that it was not relevant to my questions. The remaining columns record how much fuel was used to generate different forms of electricity. Each piece of data is measure in MTOE, Million Tonnes of Oil Equivalent.

From the Generated and Supplied sheet records how much electricity was generated each year. It records how much was generated by each method of electrical generation as well as how much was used in the generation of electricity. All data recorded in this sheet is measured in GWh, Giga Watt hours.

From the Price sheet, I have only chosen the p/KWh, price per Kilo Watt Hour column. This is the only column that makes sense to me from this sheet and it gives valuable context to my questions.

Answering my questions

To answer this question, I needed to measure the amount of fuel input and compare it to the amount of electricity generated. The 1980's and 2010's are good time frames to take these measurements. I chose the 1980's because there are no years where there is a sudden increase or decrease in consumption or the amount generated. I chose the 2010's because it is a time where electrical generation methods and technologies are at their most modern and most efficient.

Unfortunately, I cannot measure the output of coal, oil, and gas individually. It seems that these three are pooled together under the Conventional & thermal & others column. To compensate for this, I added the fuel input values for these fossil fuels together to create the tables which would compare these generation methods.

I made four graphs, two detailing the inputs of fossil and nuclear fuels from 1980 to 1989 and 2010 to 2019, and two detailing the electricity generated by both fuel types in the same time spans. All of these graphs allowed me to very clearly compare the fuel in vs the electricity out for both fuel types. Looking at the fuel in vs the electricity out between 1980 and 1989, we can see that the fuel input and the electricity generated for fossil fuels are directly correlated, with the graphs showing very similar trend lines. If we look at the inputs of nuclear fuel versus the electricity generated, we can see that the small increases of fuel going in results in much larger increases in electricity being generated. This tells us that, for fossil fuel, there were no significant improvements in how electricity is generated, whereas for nuclear, the increases in electricity out for small increases in fuel inputs tells us that the technology required to generate power in a nuclear reactor improved and resulted in increased efficiency of electrical generation. If we move onto the fuel inputs and electricity generated between 2010 and 2019, we can see the decline in use of fossil, which results in the obvious decline in electricity being generated by fossil fuels. The use of nuclear fuel is very stable in this time span, with no major increases or decreases in fuel used, but the power output is far higher, again demonstrating the progressing technologies used to generate nuclear power.

fuel is generating far less electricity when compared to coal, oil, and gas in the 80's. This is obviously because they are pooled together and the difference in fuel inputs is phenomenal. In 1980, there was 59.10 million tonnes of coal, oil, and gas input, which resulted in 215,418 GWh of electricity being produced. Coal accounted for 51.01 million tonnes of fuel, oil accounted for 7.67 million tonnes, and gas accounted for 0.42 million tonnes. In 1983, the least amount of these fuels was used, 52.51 million tonnes, which resulted in 198,822 GWh of electricity. In the 2010's, there is a drastic decrease in the use of fossil fuels. This is likely caused by the need to move to more efficient and renewable energies was well as the need to cut down on carbon emissions. From the fuel input graph depicting fuel inputs for the 2010's, the decrease in inputs is surprising stable, with a large increase in 2012, likely caused by the London Olympics. Moving to the generation graph, we can see aa huge decline in the amount of

energy generated, with an uptick in production in 2012 correlating to the increase in fuel input. I feel that this best shows the how efficient the generation of fossil fuels is.

The efficiency of fossil fuels is negligible during the 1980's. The small increases and decreases in fuel input directly result in small increases in the electricity generated. This method of generation seems to have been most efficient in 9187, where 57.25 million tonnes was used to make 215,790 GWh. A decrease of 2 million tonnes led to a decrease of just 200 GWh when compared to 1980. During the 2010's, we can really see the fall in efficiency. There is a steady fall in the amount of fuel input, but there is a huge drop in the electricity produced. In 2010, there was 59.17 million tonnes of coal, oil and gas used to generate electricity, which resulted in 157,818 GWh. Compare this to the 59.10 million tonnes used in 1980 and the 215,418 GWh produced. The difference in fuel is minimal, but there is a drop of 57,600 GWh in electricity produced. This all the more telling when we look at the fuel input in 2019 vs the electricity generated. A drop of 33.57 million tonnes resulted in 25,403 GWh. It seems that, as the need for alternate forms of energy became more apparent, the efficiency of generating electricity using fossil fuels fell greatly.

As for Nuclear energy, there was 9.91 million tonnes of nuclear fuel used, which resulted in 32,291 GWh of electricity. In 1989, 17.74 million tonnes of nuclear fuel generated 63,602 GWh of power. In these nine years, an increase of 8 million tonnes of fuel, almost doubled the electricity produced. In 2010, only 13 tonnes of nuclear fuel generated 56,442 GWh. The fuel input versus the electricity generated using nuclear fuel became more efficient as time went on. This is most apparent in the 2010's where the average amount of fuel used is not much less than that of the 1980's and the electricity generated is often 10,000 GWh higher, for example, in 1984 and 2015, 14.50 and 13.85 million tonnes of nuclear fuel was used respectively. In 1984, this fuel resulted in 47,256 GWh, whereas in 2014, this fuel resulted in 57,903 GWh of electricity. Now why is nuclear power not more widespread? This can be answered very easily, cost. Nuclear power plants are extremely expensive and time consuming to build, often taking upwards of 6 years to see the end of construction, and another 15 years to begin making a profit. This is a very interesting topic and I highly recommend watching this video by the Irish educational YouTuber, Brian McManus, who runs the Real Engineering YouTube channel. His video on The Economics of Nuclear Energy is very detailed in the costs of building and running a nuclear power plant. Unfortunately, this topic is not related to my question, so I will not be discussing it any further.

After comparing these different methods of generating electricity, we can see that the efficiency of generating electricity using fossil fuels seems to drop. This is apparent when we look at the fuel input and electricity generated in 1980 versus 2010. When we look at nuclear power, we can see that the efficiency improves. In 1980, 9 million tonnes of fuel resulted in 32 thousand GWh of electricity. In 2010, 13 million tonnes of fuel resulted in 56 thousand GWh of fuel.

Between 1965 and 1986, the production if electricity is greater than what is consumed. This is most likely due to the lack of consumer electronic devices. It is only from 1987 onwards is the consumption of electricity greater than that of which is supplied. Electricity supplied and consumed is at its highest in 2005, likely due to the financial boom of the early 2000's. This is when everyone has a cell phone and computers are everywhere. In 2009 there is a large decrease in what is supplied and consumed, and we can mark this down to the financial crash of 2008. Beyond this point, the gap between electricity supplied and consumed continues to grow. The increased consumption is caused by the increased number of electronic devices present in every household across Britain, as well as the increasing reliance on cloud data storage and data farms, which use a huge amount of electricity on cooling alone.

The decrease in electricity supplied is caused by infrastructure. Britain may not have the infrastructure in place to meet demand, and due to the countries location, it likely purchases large amounts of electricity from France and Germany. The gap present in my graphs is probably caused by this. Purchases from other producers in the data set provided, being as vague as it is, can be interpreted as purchases from other nations, and ceases to provide any information past 1986, and as such, we see the fall in supply. This explains why the consumption of electricity quickly outpaces the electricity supplied.

Consumption has increased at a very similar rate to what is being generated and supplied, and it likely remains this way post 1986, otherwise, consumption rates wouldn't be so high. The progression of technology and the increase in the number of electronic consumer products, such as phones, games consoles, and computers, has led the steady increase of consumption up to 2005. As I mentioned earlier, this is soon followed by the financial crisis, leading to a decreased consumption rate. After this point, the consumption of electricity begins to steadily decrease. The most probable cause of this decrease is the increasing efficiency of electronic devices.

The consumption of electricity has greatly increased with higher rates of generation, but it has also decreased due to the efficiency in power usage of the devices that we use daily.

From 1920 and through most of the 1930's electricity was very cheap. The amount being produced, and the time period meant that there was no major uses for it. Prices of electricity continued to fall with more being produced. The amount of electricity continued to increase each year, and the price of electricity, although there is data missing between 1961 and 1954, and again in 1956, 57, 58, 59, and 61, the price of electricity remained stable. It is not until the 1970's that we begin to see the price per kilo watt increase. The 70's saw low unemployment rates and a surprisingly high inflation rate of 16.04%. This high inflation rate can likely be blamed for the sudden explosion in electricity prices. This trend of increasing prices keeps going, with a small slowdown in the early and mid 80's. Electricity has remained relatively cheap, and as more is produced, prices have increased. But this increase in price cannot account for the increasing generation rates.

The Dashboard

The purpose of the dashboard is to the display the data I have chosen to answer my questions in an easy-to-understand manner. For question 1, I began by creating a pivot table containing data on the fuel inputs and electricity output by fossil fuels and nuclear fuel. Then I added a group box with four option buttons. The option buttons are linked to a cell which displays which button is selected. The number displayed in this cell correlates to a column in the pivot table, depending on which option button is selected. I then created a second table with a year column and a column that would display data based on what options button was selected. This column dynamically changes through the use of the 'INDEX' function. The formula is as follows: 'INDEX(B20:E20,0,A19)'. The array highlighted in yellow represents the heading of the fuel inputs and electricity outputs of the pivot table, and the green highlight represents the cell that each option button is linked to. This allows the column to changed and display whatever data is linked to each options button. I then made a graph using the 'INDEX' function, this time the yellow highlight included all of the fuels inputs and electricity outputs for each year in the pivot table. Next I used the 'MATCH' function to match the years, and lastly, I used the green highlight, cell A19, to make the graph display the data corresponding to whatever options button was displayed.

I did this for all three of my questions, starting with the pivot table, adding the options buttons, the dynamic table, and then the dynamic graphs. I also added static charts to display more detailed information comparing fuel types and electricity outputs in specific decades, consumption versus supply, and price versus supply.

I also added a green background to the dashboard with a yellow divider between tables and each question. I feel that this provides contrast between text, charts, and clearly shows where each question begins and ends.
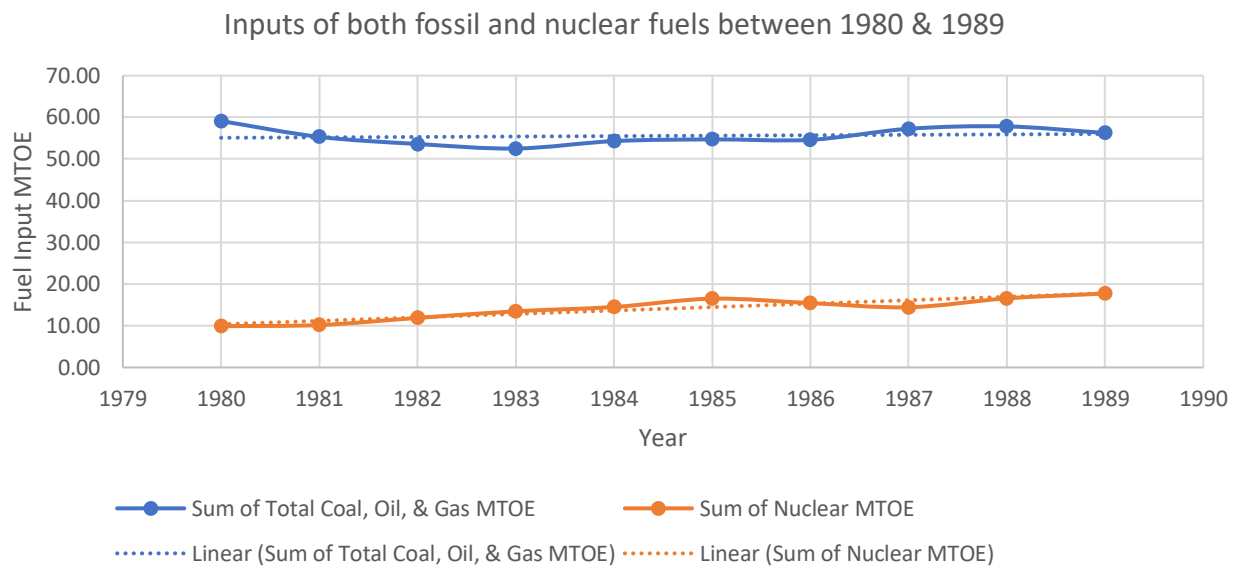
Graphs Referenced in this Report

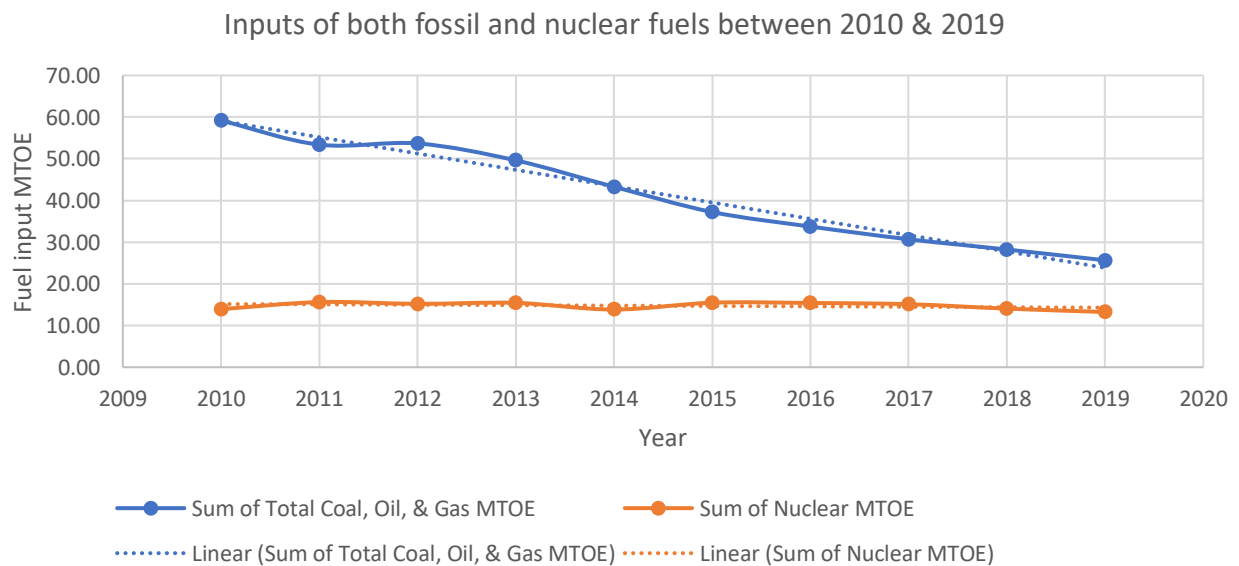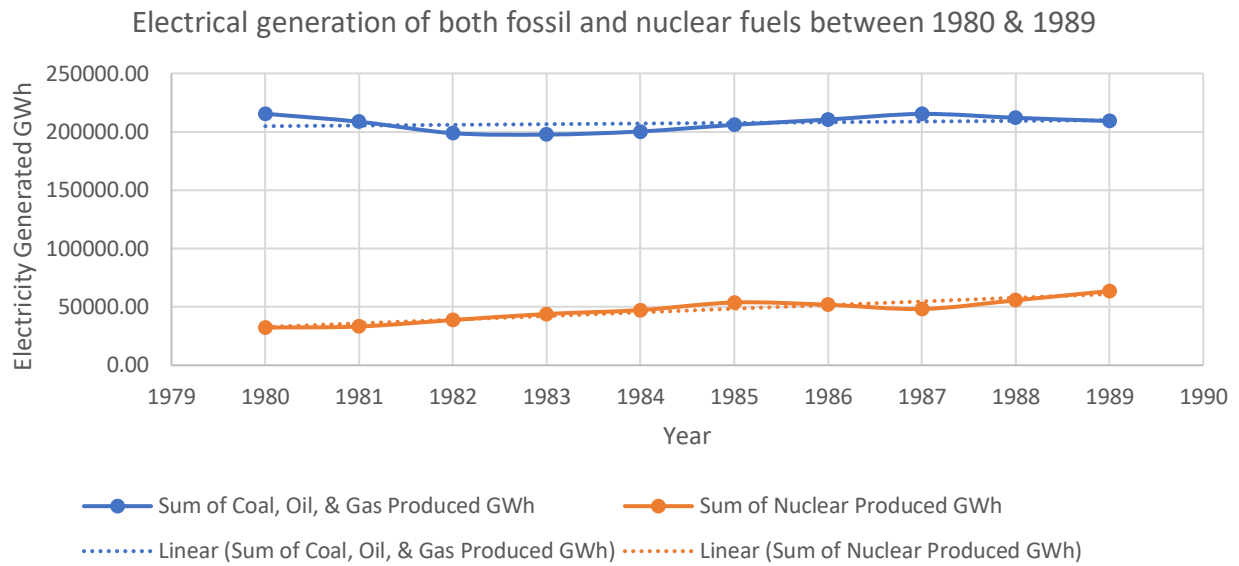## Inputs of both fossil and nuclear fuels between 1980 & 1989

Figure 1

Inputs of both fossil and nuclear fuels between 2010 & 2019

Figure 2

Figure 3

Figure 4

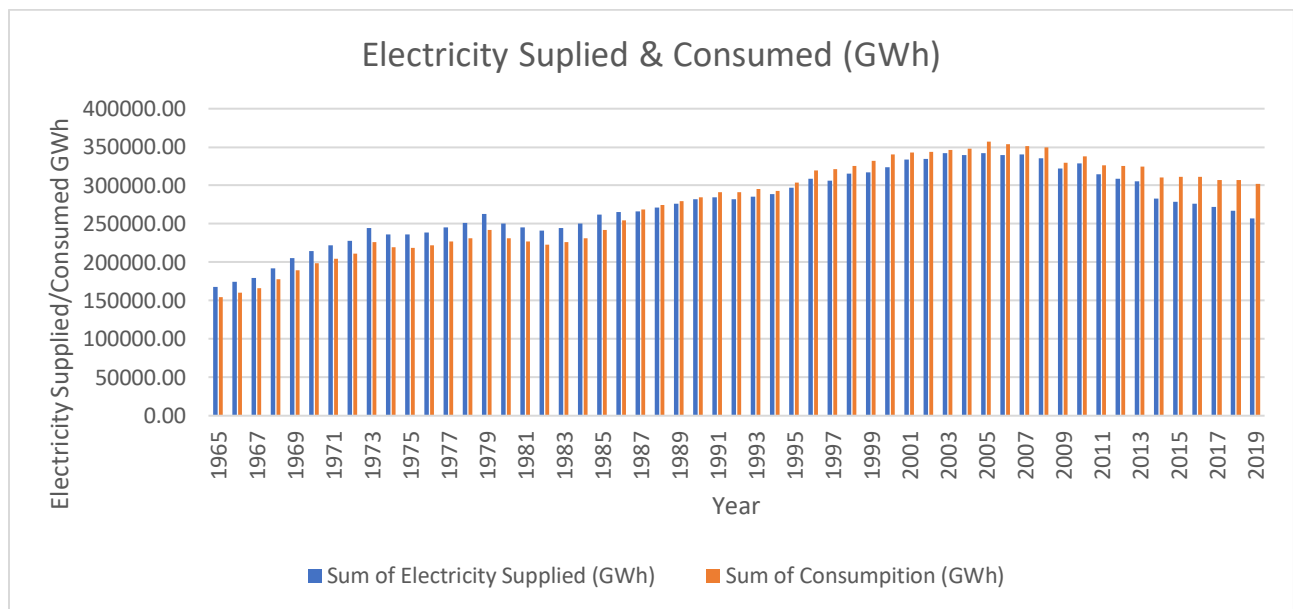Overlay of Price Per Kilo Watt & Total Electricity Produced between 1920 and 1990
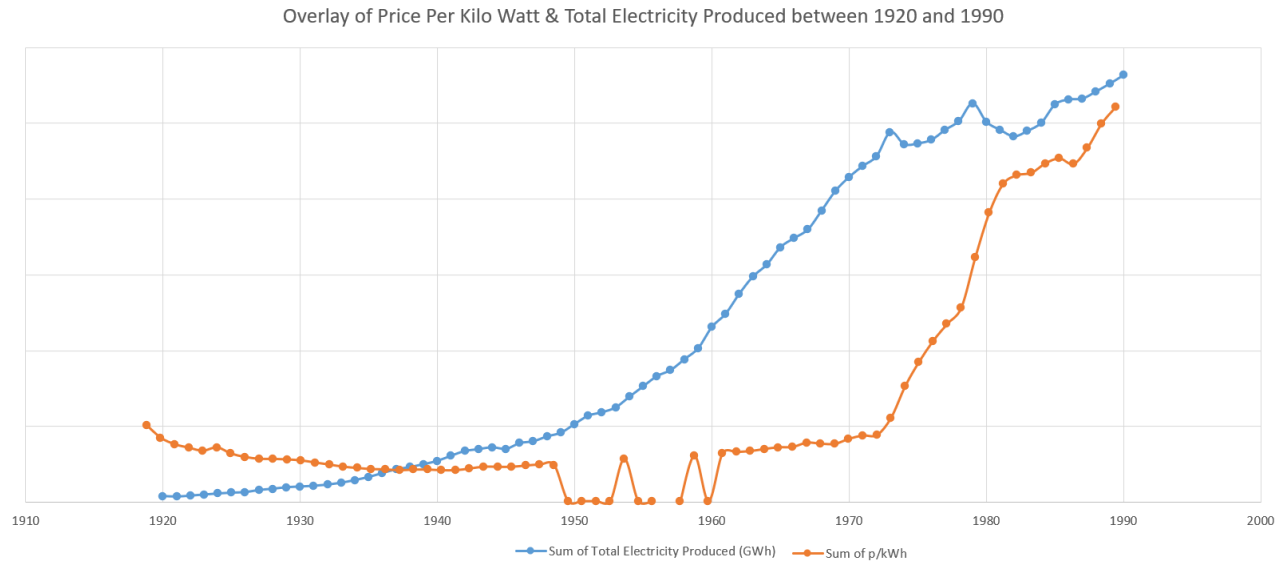
Figure 5

References

Talend - A Cloud Data Integration Leader (modern ETL). 2020. 16 Data Mining Techniques: The Complete List - Talend. [ONLINE] Available at: https://www.talend.com/resources/data-mining-techniques/. [Accessed 19 December 2020].

Decision Tree Algorithm Examples in Data Mining. 2020. Decision Tree Algorithm Examples in Data Mining. [ONLINE] Available at: https://www.softwaretestinghelp.com/decision-tree-algorithm-examples-data-mining/. [Accessed 19 December 2020].

Wikipedia. 2020. Sequential pattern mining - Wikipedia. [ONLINE] Available at: https://en.wikipedia.org/wiki/Sequential_pattern_mining. [Accessed 19 December 2020].

. 2020. [PDF] A taxonomy of sequential pattern mining algorithms | Semantic Scholar. [ONLINE] Available at: https://www.semanticscholar.org/paper/A-taxonomy-of-sequential-pattern-mining-algorithms-Mabroukeh-Ezeife/d67cd96b0a30c5d2b2703491e02ec0f7d1954bb8. [Accessed 19 December 2020].

Real Engineering. (2020). *The Economics of Nuclear Energy*. [Online Video]. 6 June 2020. Available from: https://www.youtube.com/watch?v=UC_BCz0pzMw&t=1s. [Accessed: 15 November 2020].