

DS8007 Project Documentation

World Energy Consumption

April 21st, 2023

Tara Shingadia

Table of Contents

1. Introduction 3

2. Data Description 4

3. EDA 5

4. Business problem framing 3

5. Dataset 3

6. References 3

# Introduction

According to the United Nations (UN), “the energy sector accounts for about two-thirds of global greenhouse gas emissions attributed to human activity.” The dataset being analyzed is the World Energy Consumption dataset. The problem I will be answering with this dataset is understanding if there is a trend amongst different countries in different parts of the world regarding consumption levels and the types of energy used. We can see which countries use cleaner energy than dirty fossil fuels by considering the energy types and ultimately see the transition over time.

Some of the questions that may provide insights include what is the distribution of energy usage by continent and country? This will tell us which countries/continents are the highest and lowest energy consumers. How has energy usage changed over time? Is there a correlation between energy consumption and economic growth? Has the energy usage per person increased?

The motivation behind this research is to understand better how energy usage is changing. We know there are many types of energy producers, and seeing how countries with the highest GDP progress could help other countries with energy production decision-making. Clean energy is also very important to me and something of interest, so it would be interesting to see how clean energy compares to other forms of energy.

I will utilize Python, and several visualization tools learnt throughout the DS8007 course to help solve the problem previously defined for this dataset. Python will be used to process and analyze our data, including univariate and bivariate analyses such as correlation plots. Several visualizations will also be made using Python, such as bar, pie, and line charts to visualize our data. As well as geo-visualization using Python or R to show data distribution as a heat map.

# Data Description

The link to the open-source dataset is [World Energy Consumption | Kaggle](https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption?resource=download). The dataset was created by ‘Our World Data’ and contained 122 columns and 17432 records. The dataset shows continents, countries, country iso-codes, continents and countries’ energy usage, year, type of energy, population, GDP, and consumption per capita, and change as a percentage measured in terawatt-hours and electricity generation. The different types of energy include coal, gas, hydro, nuclear, oil, wind, and carbon. There is also data on different types of electricity such as coal, fossil, gas, hydro, nuclear, oil, and other renewable electricity such as solar and wind.

From the dataset, 26 features were selected, and the data were filtered from 1990 to 2020 for a portion of the analysis.

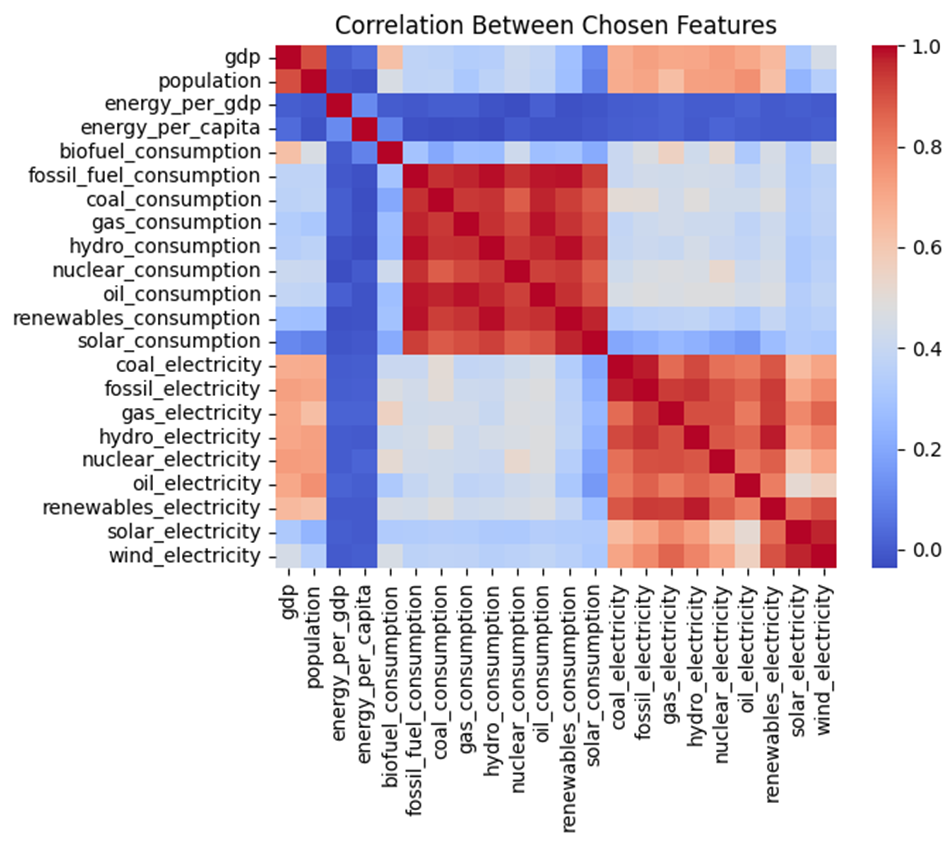
**Table

Description automatically generated**

**Fig**: **list of features and descriptions used in our analysis**

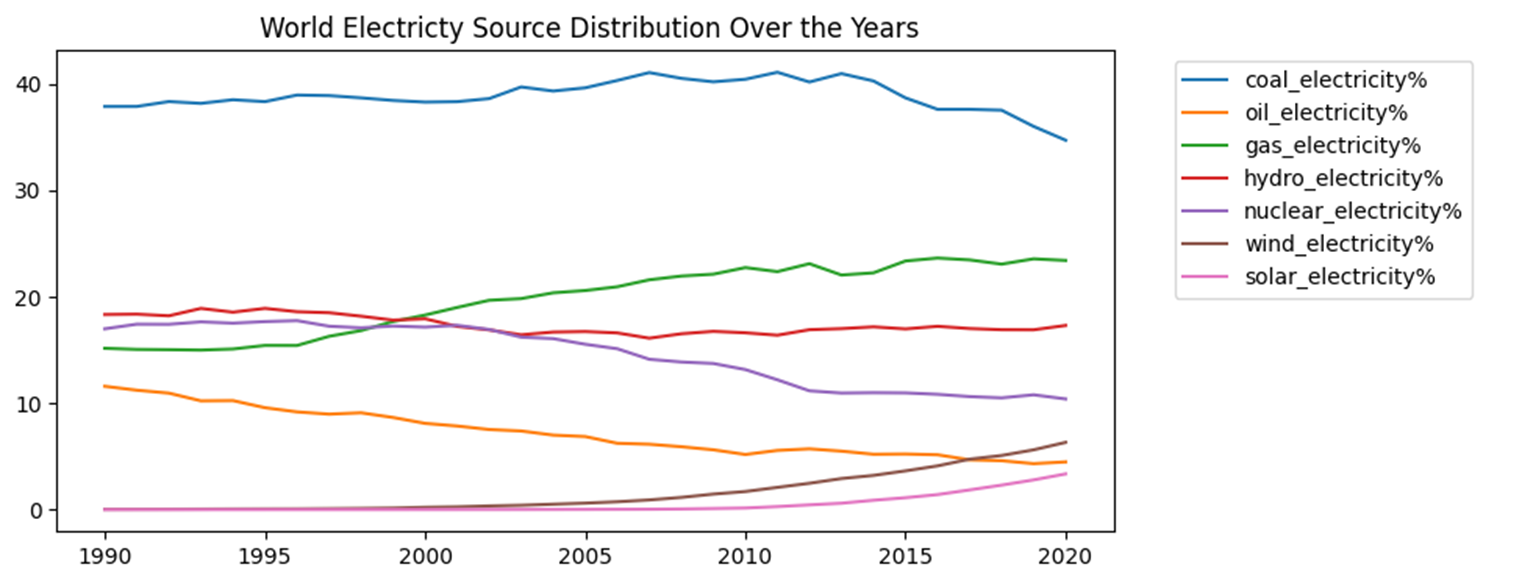
# Exploratory Data Analysis

The correlation between variables was visualized to help us understand whether there were any strong correlations between the features. From this figure, we are able to see a strong positive correlation between GDP and population. There was also a positive correlation between GDP, Population compared with the production of different types of energy and a lower correlation with the consumption of these energies. It was also interesting to note the moderate correlation between production and consumption of the different energy types.

****

**Fig: correlation between chosen features**

Using a line graph, it was easy to show the distribution of energy production over the years. The line graph colour codes the lines by energy type and looks visually appealing and easy to understand. We can see that coal production started to dip after 2015 but is still the most predominant type of energy produced. Hydroelectricity production has remained fairly stable over the past 30 years. There has been an increase in electricity produced by gas and a steady drop in the production of nuclear and oil electricity. It’s also interesting to see wind and solar energy on the rise since 2010.

****

**Fig: change in electricity type production from 1990 to 2020**

Bar charts were chosen to show the electricity distribution by type, comparing 1990 and 2018 because they make it easy to compare percentages. We can see that solar and wind were not used back then, but in 2018 they made up a small percentage of the world’s electricity distribution. It’s also important to note that coal has remained the most used source of energy, and hydro has remained stable. Oil and nuclear have decreased quite a bit and gas electricity has increased.

**Chart, pie chart

Description automatically generated**

**Fig: world electricity distribution in 1990 (left) and 2018 (right)**

The total energy consumption vs production was plotted using a scatter plot below to see if there was a trend, as scatterplots make it easy to see if there is a linear or another type of trend. This graph shows a linearly positive trend as there is a steady upwards line.

**Chart, scatter chart

Description automatically generated**

**Fig: world energy consumption vs production trend**

Using a line plot, we can see the trend, if it’s steadily increasing or if there are any dips and what years those dips occur. Energy consumption is steadily increasing over the years despite a small drop around 2010.

**Chart, line chart

Description automatically generated**

**Fig: world energy consumption from 1990 to 2020**

Using another scatterplot, we can see the average energy usage per capita for each continent and how that trend has changed over the years. The highest average energy users have been in North America for the last 30 years (1990-2020). The lowest energy usage per capita has been in Africa.

**Chart

Description automatically generated**

**Fig: world energy usage per capita from 1990 to 2020 by continent**

Using a heat map, renewable energy consumption was plotted so that we could visualize which countries were the largest consumers and which ones were the least. From the visualization below, we can see that Zambia, Yemen and Zimbabwe have had the highest renewable energy consumption in the world over the past 30 years.

**Map

Description automatically generated**

**Fig: world renewable energy usage from 1990 to 2020 by country**

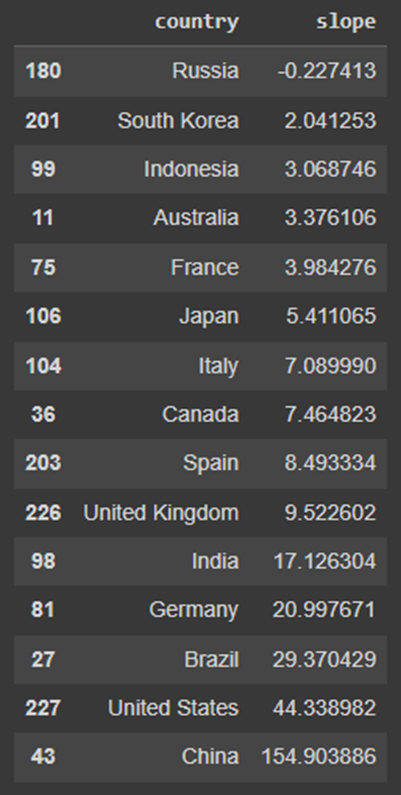
Another heat map was used to visualize the energy consumption by GDP because world heatmaps make it easy to visualize and see if there are trends by continent as well. This heatmap highlights Iceland, Libya, and Paraguay as the highest energy users per GDP. This graph was created to answer the question of whether higher energy consumption is necessary for a higher GDP. These results may be the case because the countries mentioned don’t use energy efficiently.

Map

Description automatically generated

**Fig: world energy usage per GDP by country from 1990 to 2020**

Next, we explore the top 10 countries in the world with the highest GDP’s renewable energy consumption. The slope is calculated so we can see how much of an increase was made over the past 30 years. From this, we can see that China, The United States and Brazil were the highest users of renewable energy sources. The bar graph was chosen with the slopes because it allows for an easy side-by-side comparison which also takes into consideration change over the years.

Chart, histogram

Description automatically generated

Similarly, to the graph above, we look at fossil fuels consumption this time for the Top 10 countries with the highest GDP. China, India, and South Korea are the highest consumers of fossil fuel energy. Germany, Russia, and the United Kingdom have had a negative slope of fossil fuel energy consumption.

A screenshot of a computer screen

Description automatically generated with medium confidenceChart, histogram, waterfall chart

Description automatically generated

# Findings, Results and Conclusions

From all the exploratory work that was done, we can see that world energy consumption is increasing but is quite stable per capita. Of the countries with the highest GDP, China is the largest consumer of both renewable and fossil fuel energy. Germany, Russia, and The United Kingdom have been decreasing their fossil fuel consumption whilst maintaining their high GDP. This shows that it is possible for other countries to follow in their footsteps. Wind and Solar energy are on the rise, these being renewable energy sources means this is a great thing for the environment as resources are not being depleted. These methods can also be implemented in countries with lower infrastructure. Nuclear and Oil energy usage is on the decline, which is another positive takeaway as these are not as safe for the environment as alternatives. The next steps include looking at things on a more granular level and seeing how the distribution of energy usage is by specific countries.

# References

Ritchie, H., Roser, M., & Mathieu, E. (n.d.). *World Energy Consumption*. Kaggle. Retrieved March 14, 2023, from https://www.kaggle.com/datasets/pralabhpoudel/world-energy-consumption?resource=download

Environment, U. N. (n.d.). *Why does energy matter?* UNEP. Retrieved March 14, 2023, from https://www.unep.org/explore-topics/energy/why-does-energy-matter#:~:text=Energy%20and%20technology&text=The%20energy%20sector%20accounts%20for,waste%20for%20cooking%20and%20heating.