End Assignment | Data Analytics with Python | Gerben

Introduction

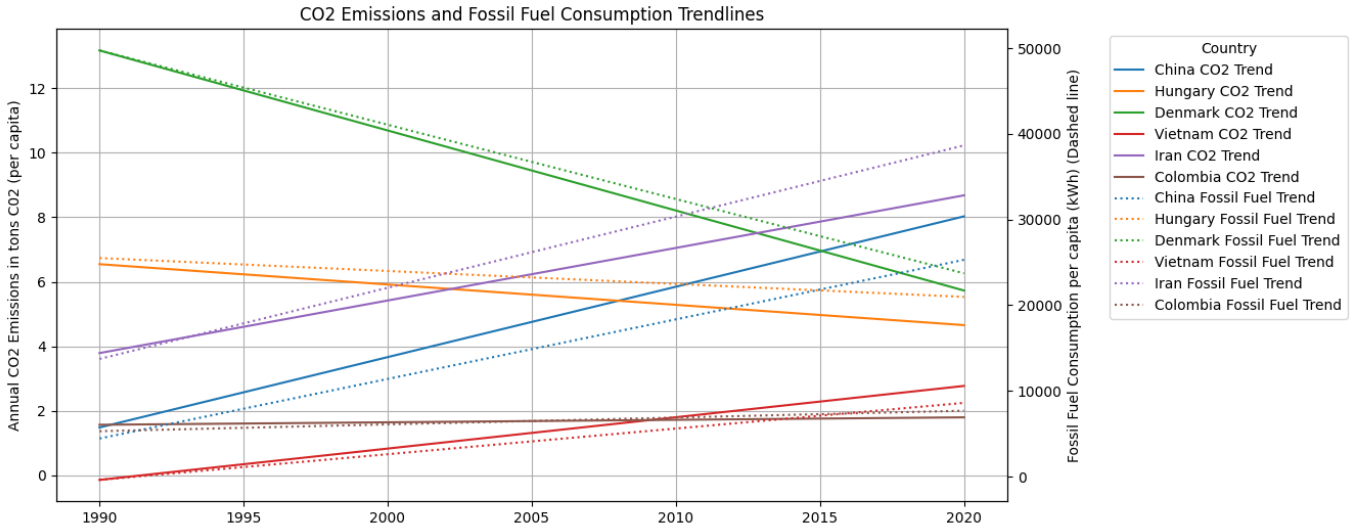
Climate change is an immense hot topic that brings forth a large amount of discussions. Such as:

What is the biggest predictor of a large CO2 output per capita of a country? Which countries are making the biggest strides in decreasing CO2 output? Or which non-fossil fuel energy technology will have the best price in the future?

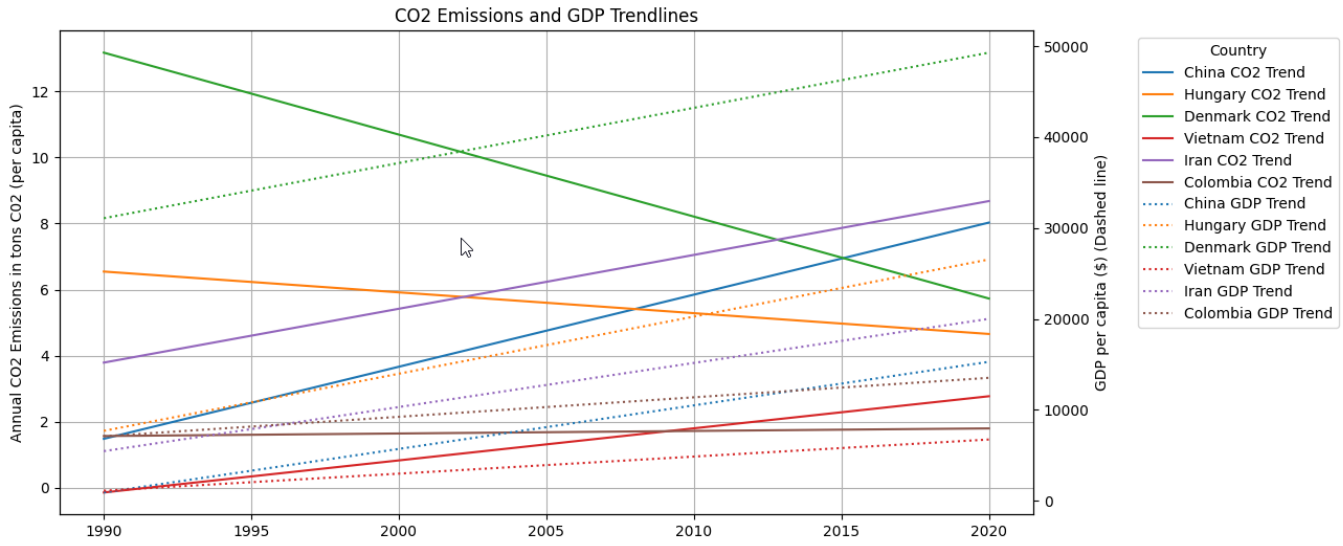
Method To analyse the data used in this case study, data was acquired mostly from the "Our World in data" webpage. Data was ingested from the downloaded csv files and then copied for use. After intial exploration data was cleaned, missed values dropped etc. For the in depth exploration see the Jupyter Notebook "Function Body" code block. It features column and dataset exploration functions. The datasets for the first 2 questions were then merged to form the final dataset. For the liniear regression question a seperate dataset was used from IRENA.

1 - What is the biggest predictor of a large CO2 output per capita of a country?

The biggest predictor of CO2 output shares the same trendline. The fossil energy consumption rate is the biggest predictor of CO2 output. This indicates that their closely correlated is some way or another. It is a good predictor because it takes in account the literal (fossil) energy usage per capita. This way energy share, industrial activity (consumption patterns) and population are taken into account. To get a general idea a sample of 5 countries was used. This way we can look at multiple countries without losing detail. It does although create bias for not finding outlier countries.



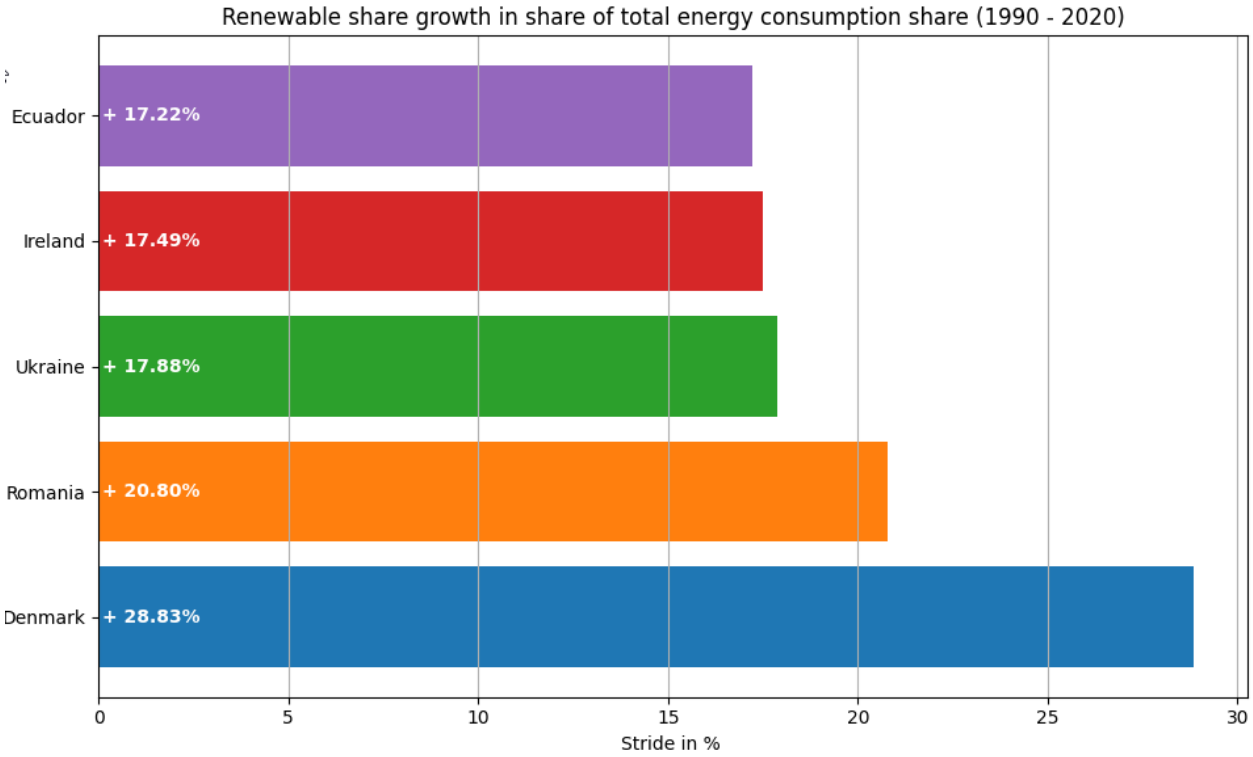
We should note that economical development (GDP) isn't taken into account. To show that economical development isn't as directly correlated see this figure below. Note the difference that countries with a higher GDP such as Hungary or Denmark are inclined to be even contrary. We can argue that GDP does however correlate indirectly by how countries are able to move away from fossil fuel consumption. This could be a topic of further research.



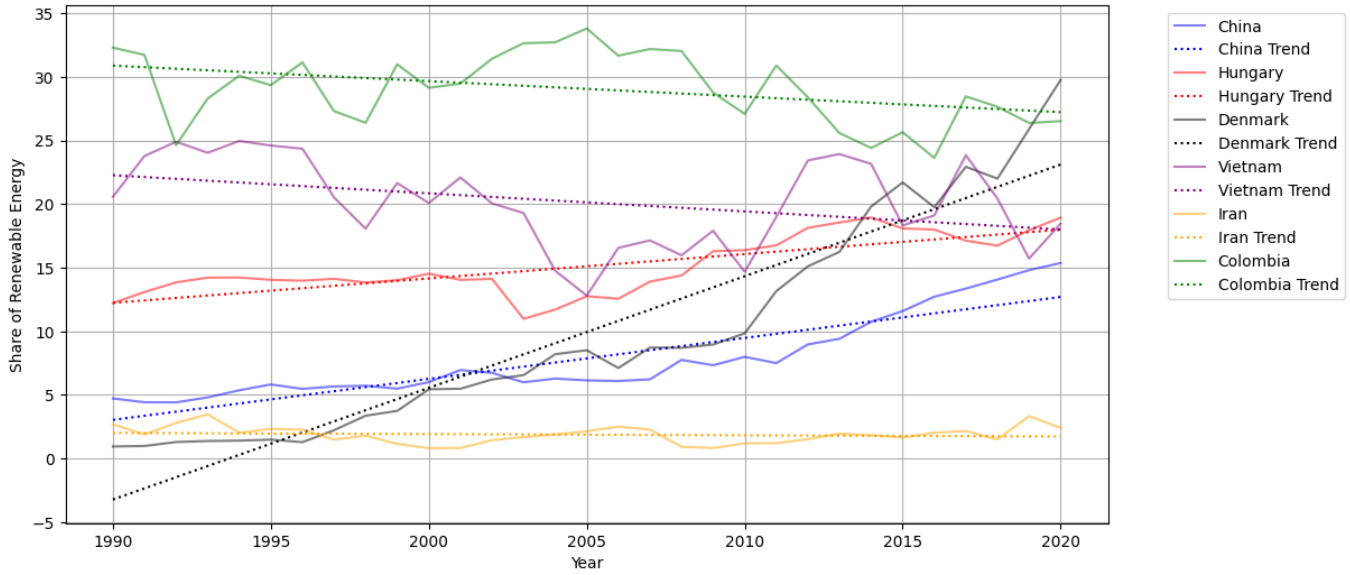
See the Jupyter Notebook codeblock 'Question 1' for more figures. These are used as stepping stones to create the figures shown above. See the Jupter Notebook codeblock 'Main Body' on how the dataset was constructed.

2 - Which countries are making the biggest strides in decreasing CO2 output?

The best reducers of CO2 are those countries whose renewable energyshare in the energy consumption share has grown the most. To get a good overview we take a look at the period 1990-2020. This period was chosen because of the realiability of the dataset used. From 1990 on nearly all countries started tracking their CO2 output. See the figure below for the 5 countries who have taken the biggest stride. Their share was calculated by aggregating the energy consumption of each technology and comparing it to the total energy consumption.



The horizontal bar chart misses some detail though. Because we calculate the difference between 1990-2020 we dont see what has happened in between. We miss out on factors like war, geopolitical struggles or economic influences. A good further topic would be to investigate if the trend line of the progress would be a better way to define the stride. The average gives a better absolute answer while the trend line would give more of a relative stride. See the figure below to visualize the 'noise' that the bar chart wont show:

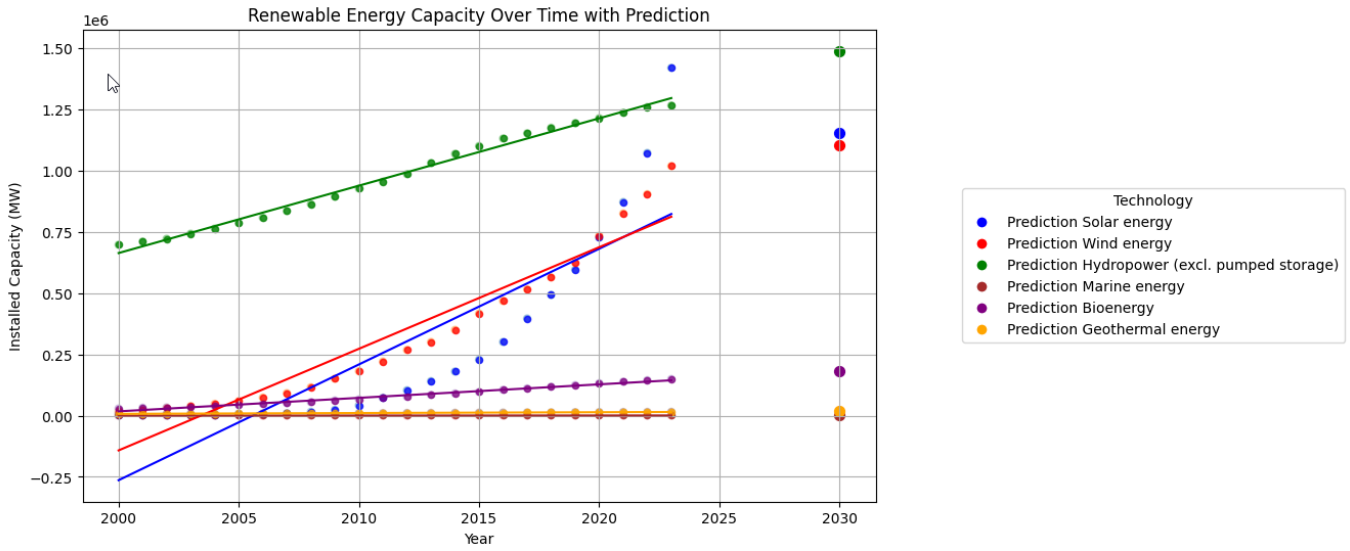


See the Jupter Notebook codeblock 'Question 2' for more figures. These are used as stepping stones to create the figures shown above. See the Jupter Notebook codeblock 'Main Body' on how the dataset was constructed.

3 - Which non-fossil fuel energy technology will have the best price in the future?

Electricity from renewable energy follows learning curves (See Our World in Data article on Renewable Energy). With each doubling of the cumulative installed capacity of a technology the price declines the same fraction. Fossil energy technologies do not follow this learning curve. So to predict which type of renewable energy will have the best price we need to predict which learning curve will be the strongest in the future. To predict which learning curve is the strongest we can argue that if we look at the prediction of installed capacity this could indicate a fair share.

If we plot the capacity for each technology and predict the values using a liniear regression model we see that Hydropower will have the biggest installed capacity. Closely followed by solar and wind energy.



Note that Hydropower has its limitations. It is much more limited to geographic conditions in comparison to solar and wind energy for instance. Solar would be the best non-fossil fuel energy in terms of the prediction and allround availability.

The dataset used also is not that large which skews the answer. For a machine learning model needs more data to train itself well. With more datapoints it could also be compared to a ridge regression model for instance. Also, reliable data on energy prices for each technology is hard to find and differs greatly which makes a good topic for further investigation.

See the Jupter Notebook codeblock 'Question 3' for more figures. These are used as stepping stones to create the figures shown above. See the Jupter Notebook codeblock 'Main Body' on how the dataset was constructed.

4 - Appendix

The notebook used: End Assignment Notebook

Sources

df1, Global Carbon Budget (2023); Population based on various sources (2023) – with major processing by Our World in Data https://ourworldindata.org/grapher/co-emissions-per-capita

df2, GDP per Capita - Bolt and van Zanden - Maddison Project Database 2023 – with minor processing by Our World in Data https://ourworldindata.org/grapher/gdp-per-capita-maddison

df3, Per capita primary energy consumption by source, 2023 by Our World in Data https://ourworldindata.org/grapher/per-capita-energy-stacked

df4, Electricity statistics (MW/GWh) by Region, Technology, Data Type and Year IRENASTAT Online Data Query Tool by International Renewable Energy Agency (IRENA) https://pxweb.irena.org/pxweb/en/IRENASTAT/IRENASTAT__Power%20Capacity%20and%20Generation/

Article on Renewable energy | Our World in Data: https://ourworldindata.org/renewable-energy