Final Assignment CO2 Emission

Our report aims to address three key questions, with data sourced from ourworldindata.org. The questions seek to answer are:

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

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

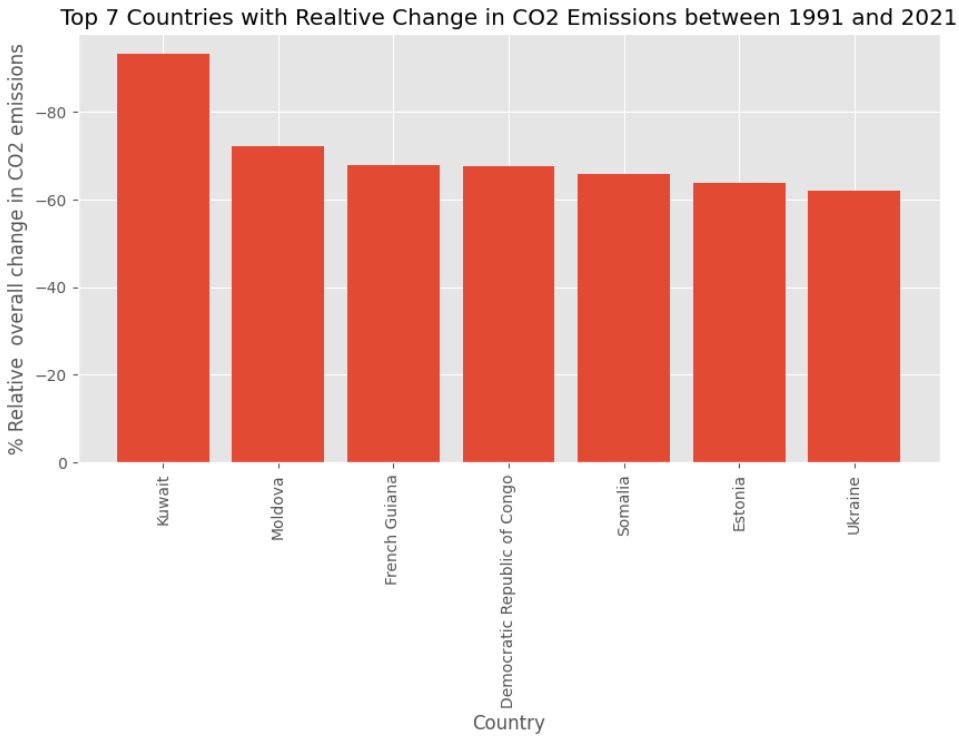
I curated and analysed datasets from ourworldindata.org to answer these questions and present our findings with visual aids. My analysis revealed that the largest predictor of high CO2 output per capita is fuel. As individuals gain more prosperity, they tend to purchase more appliances, electronics, and cars, ultimately leading to greater environmental impact. Our calculations using the Pearson correlation coefficient shod a strong correlation of 0.74 between GDP per capita and CO2 emissions per capita.

Additionally, our analysis shod that the rise in CO2 emissions starts to stabilize and even decrease at a GDP per capita of approximately $5400. This trend is not because use less energy, but because can afford less environmentally damaging and more energy-efficient goods. Finally, while total energy consumption has continued to rise in conjunction with GDP, CO2 emissions have stagnated.

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

To calculate this, we used a data frame with the population and CO2 emissions per country and per year. We focused on data from 1990 to 2020 as data before that period was missing for many countries. This left us with a dataset of 217 countries. Since population size has a significant influence on emissions, we calculated the relative change in population between 1990 and 2020. We measured four points: 1990, 2000, 2010, and 2020, with in-between points to ensure that extreme changes in population were correct and not caused by incorrect data. We did the same for CO2 emissions.

On average, population growth per country was 163%. Therefore, we removed the countries with more than 300% relative growth and less than 70%, excluding seven countries. We also filtered out small countries with less than 50k inhabitants (15 countries). For the 190 remaining countries, we calculated the relative change and identified the top 10.



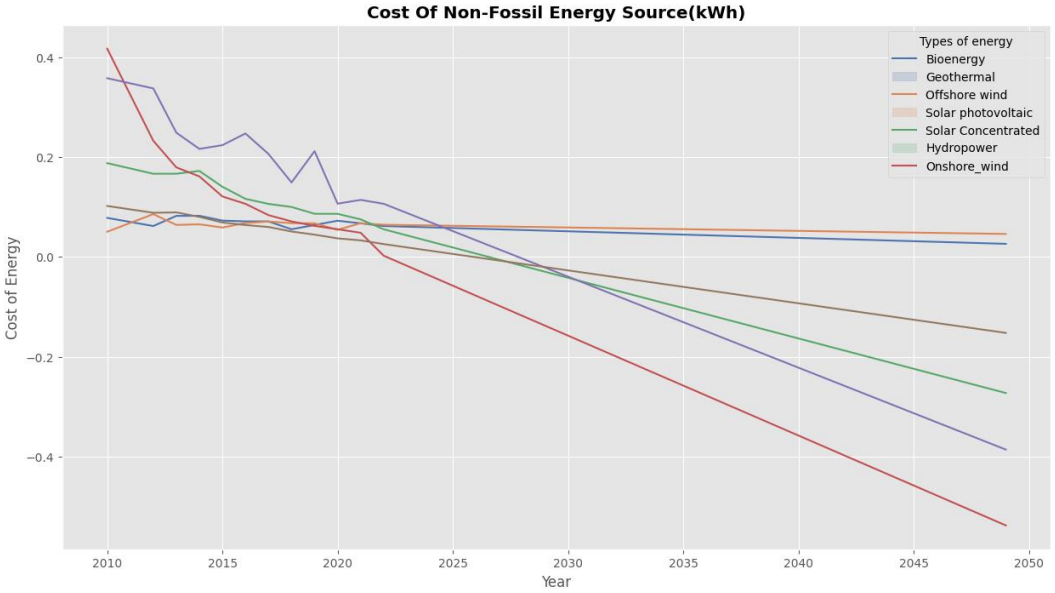
As seen below, Moldova made the biggest stride in CO2 reduction. However, despite a shrinking population, emissions are climbing again, raising questions about how long this will last since they have a relative population change of 92. As Moldova separated from the USSR in 1991, we cannot determine how the emissions for 1990 were calculated. We did not detect any outliers in the dataset for Moldova. Slovakia, Romania, North Korea, and Denmark are the countries in the top 10 that are on a steady decline.

It's worth noting that most of these countries have a relative growth below 100 since 1990, indicating a shrinking population. However, Denmark, North Korea, and Slovakia are exceptions to this trend.

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

To answer this question, we utilized a dataset containing the prices of various sustainable energy sources. These sources include bioenergy, geothermal energy, energy from offshore wind turbines, energy from onshore wind turbines, energy from solar panels, solar thermal energy, and energy generated from flowing water (hydropower). Only the world-level data are selected for analysis. For each category, a separate dataset is created consisting of the years and the price per kWh. Rows with missing values are removed. When the data is examined, it is clear that in the case of energy from solar panels, there is an exponential decline. Therefore, we chose to convert these measurements to logarithmic values to perform a more reliable linear regression. After the linear regression, the values are converted back to normal values. After preparing the data, linear regression is performed using the sklearn.linear_model library. Here, the data is divided into test data and train data using the train_test_split function of the sklearn.model_selection library. We chose to use 80% of the data to train the model and 20% of the data to test the model. For the random_state parameter, the value '42' is used to ensure that the same result is generated each time. The model is evaluated using the mean_absolute_error and mean_squared_error functions of the sklearn.metric library. Then, a linear formula is created for each energy source using the results of the linear regression. For 2022 through 2050, the price per kWh is calculated using these formulas. The data from all sustainable energy sources is combined into one dataset and plotted together in a line chart using the seaborn library.

The conclusion of a study on sustainable energy sources. The dataset used for the analysis had limited global measurements available, with only a small number of measurements for each type of energy source. The prices per kWh in US dollars were plotted against the years for each energy source, and linear regression models were developed based on these data.



One interesting observation was that the calculated price per kWh for geothermal energy, onshore wind energy, and offshore wind energy ultimately became negative, which is unlikely. Therefore, the reliability of the future prices calculated by the models is somewhat limited.

To answer the question of which sustainable energy source will have the best price in the future, the authors looked at the energy sources with a realistic price trend. Based on their analysis, energy generated from solar panels is expected to have the best price in the future.