**Data Analysist Report on CO2 emission**

*Title Page:* CO2 *predictors, decrease and renewable energy*

Prepared by: *Joëlle Kappert*

Organization: *Winc Academy*

Date*: February 28, 2025*

Table of Contents

[*1. Executive Summary* 1](#_Toc191645840)

[*2.* *Introduction* 1](#_Toc191645841)

[*3. Methodology* *and* *definition* *of* *choices* 2](#_Toc191645842)

[*4. Data* *Analysis* 3](#_Toc191645843)

[*5. Result* 3](#_Toc191645844)

[5.1 What is the biggest predictor of a large CO2 output per capita of a country? 3](#_Toc191645845)

[5.2 Which countries are making the biggest strides in decreasing CO2 output? 4](#_Toc191645846)

[5.3 Which non-fossil fuel energy technology will have the best price in the future? 5](#_Toc191645847)

[*6.* *Conclusion* 6](#_Toc191645848)

[*7. Appendices* 6](#_Toc191645849)

[*8.* *References* 7](#_Toc191645850)

## 1. Executive Summary

This report provides an in-depth analysis of CO2 emissions from data analyses and research organisation Our World of Data. The objective is to discover which factors have an impact on CO2 emissions output, such as GDP per capita, transport, fossil fuel consumption and population among others. And what actions have been taken to reduce the emission output.

Investments are increasing in renewable energy technologies all over the world, due to the Paris Accord.

## 2. Introduction

CO2 emission is based on greenhouse gas. This report examines data from the past 5 years (2019-2023) to see the movement of the emission output, other influential impact, and the correlation between these factors. The final goal is to visualize the outcomes of the data analyses in order to answer the questions about CO2 emission output, steps taken to reduce emission output and the prices of future of renewable energy technologies.

## 3. Methodology and definition of choices

Analytical tools such as Python and Pandas were used to analyse and interpret the different, large datasets in Google Colab Notebook. MatplotLib was used to visualize the outcomes into graphs (plots).

Multiple large datasets were selected to analyse and find the results required. The number of cars per capita was outdated (till 2017) and therefore not analysed. This was replaced by the sales of electric cars which was used for the reduction of emission output analysis. This dataset was more suitable for the analysis of having an active plan to reduce emission output.

The first step in analysing datasets is to clean/organize them. This consists of dropping unnecessary columns, for example country codes and countries that are not actually countries, like Africa, Asia, etc. Renaming columns and rounding numbers to make the datasets easier to read and selecting the past 5 years (2019-2023).

After cleaning, checking for missing data (NaN, null or 0.0) was executed. If the missing data was less than 6%, the decision was made to do nothing with it. In some cases it was necessary to drop the missing data in order to combine (concatenate) the results for the graphs. This can be seen from the coding used for concatenating (*pd.concat()*). To get the results, three methods were used to make sure the outcomes were similar. These methods are: groupby in combination with aggregation, idxmax or idxmin and nlargest or nsmallest.

In some cases, additional analyses/calculations had to be made. For example, the year 2023 was missing in the dataset of “GDP per capita”. For the dataset “CO2 emission Transport” the years 2022 and 2023 were missing. These were added by calculating the mean (average) and finally used for both graphs: ‘*Qatar*’ and *‘Predictors of CO2 output 2023’*. The latter is shown further in this report. Both graphs were made in order to discover the correlation and/or causation of CO2 emission output. With the dataset on “Dietary composition” the column names were exceptionally long and therefore a different code is used (see Notebook CO2 *Output).*

For the datasets *“Fossil fuels per type”, “Dietary composition” and “Renewable energy production”* a new column ‘Total’ was added. This made it easier to analyse the top or last 5. The same was done for ‘*Electric car sales*’, only the added column calculated the relative of total sales (%).

Concatenating different datasets before beginning the analysis, can cause more missing data. Therefore, I used pre-combined dataset between ‘CO2 emission’, ‘GDP’ and ‘Population’ and dropped the ‘GDP’ column since it was not necessary for the analysis of actively emission reduction. In order to gather the data for the chart, I also dropped the missing data. Struggling with creating the chart because of the use of relative numbers instead of regular numbers, has resulted in displaying only the CO2 decrease without the population (see graph: *Largest decrease CO2 emission in 2023*).

A map of countries with active CO2 emission reduction plans is included in the Notebook as a reference which I downloaded from Our World in Data/United Nations (see Notebook *Active decrease* CO2*).*

For the last analysis on renewable energy, two different datasets were used. The investments in these technologies and the cost in $ per kWh. In order to make a future prediction of renewable energy prices, the world was selected for the linear regression plots (graphs). For each energy technology a separate graph was made. These graphs are added to this report as an appendix and are included in the Notebook: *Renew. Tech*.

## 4. Data Analysis

Factors analysed from 2019-2023:

**CO2** **emission per capita**: Qatar has the largest carbon dioxide output of an average 20%. growth.

**GDP per capita:** The GDP has reduced over the years, Qatar still on top.

**Fossil fuels:** China has the largest fossil fuel consumption per fuel type (overall total), Qatar has the largest per capita.

**Population:** The largest world population is between China and India, follow by USA.

**Dietary:** The outcome per category differs. For example China has the largest vegetables composition, where the USA has the largest vegetable oils and sugar & sweeteners, Bangladesh the largest rice, Luxembourg the largest alcohol, etc. Summing all of the categories, Ireland has the largest dietary composition.

**Transport:** China and the USA share the largest CO2 emission output caused by transport.

**Electric car sales:** From the total cars sold, Norway sold 93% electric cars, followed by China with 38%.

**Renewable energy technology:** China and the United States have the largest total production of renewable energy. Iceland has the largest renewable energy per capita. Investments are on-going, prices drop. Onshore wind energy has the lowest price and will remain so in the future.

## 5. Result

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

Analyses on all the datasets that answer this question have been analysed on the past five years (2019-2023). In order to keep the results short and simple, only the results of 2023 are shown.

In order to answer this question analyses were done on multiple factors that can have in impact on emission output. Population, GDP per capita, fossil fuels, are a few examples. The results of each analysis on all of these factors are described below.

Starting with the population. The largest populations in 2023 are that of India with 1.43 billion, China with 1.42 billion and USA with 343 million inhabitants. In 2019 China was number one on the list.

CO2 emission, GDP, and fossil fuel per capita are all the highest for Qatar. To start with the most important, the annual CO2 emission per capita in 2023 was 38.84 for Qatar, followed by Brunei with 25.64 and Bahrain with 23.27. The GDP per capita for Qatar was 149.171 followed by Norway with 88.366 and Singapore with 80.320. The fossil fuels emission per capita for Qatar in 2023 show 225.427 (kWh), followed by Singapore with 159.430 (kWh) and the United Arab Emirates with 137.598 (kWh).

Following up with the fossil fuels per type (oil, coal, and gas) China, USA and India are at the top. Per type, USA has the largest gas emission with 8864.67 (TWh) and oil emission with 9960.66 (TWh), China has largest coal emission with 25538.52 (TWh). Overall China takes the lead with a total of 38.677 (TWh), USA with 21.102 (TWh), and India with 9.669 (TWh) emission output.

The CO2 emission coming from transport which does include industrial and internation transport from shipping and aviation (excl. tourism) as well, reveal that the USA is on top of the list. Unfortunately, this dataset only has data till 2021. Therefore analysing the difference between 2019 and 2021. A calculation for 2023 was made for the USA in order to show it in the graph below. The top 3 are yet again USA, China and India. The difference between the two years here is that Japan was replaced by Brazil in 2021 at the bottom of the list. Russia takes the fourth position. In 2021 USA had the largest transport emission with 1.757.339.900 (1.76 billion). Unknown is how these numbers are defined, assuming it is in kilograms (kg).

The dataset on dietary composition is more complex than the others. Analysis show that Ireland and Germany are the top two countries that are most capable of providing their population with all of the different food products combined. Meaning that Ireland can provide 3641 kilocalories (kcal) per day per capita. On the other hand this does not mean that Ireland can produce all of the food products itself. Fruits, nuts and other exotic foods are imported. This has in impact on the overall transport emission, especially industrial and international shipping/aviation. Looking at the food product separately, the outcomes differs. For example China has the highest composition on vegetables with 300 kcal, USA has sugar & sweetener with 603 kcal, Iceland has seafood with 169 kcal, etc., etc.

The graph below shows the countries of the different factors analysed.

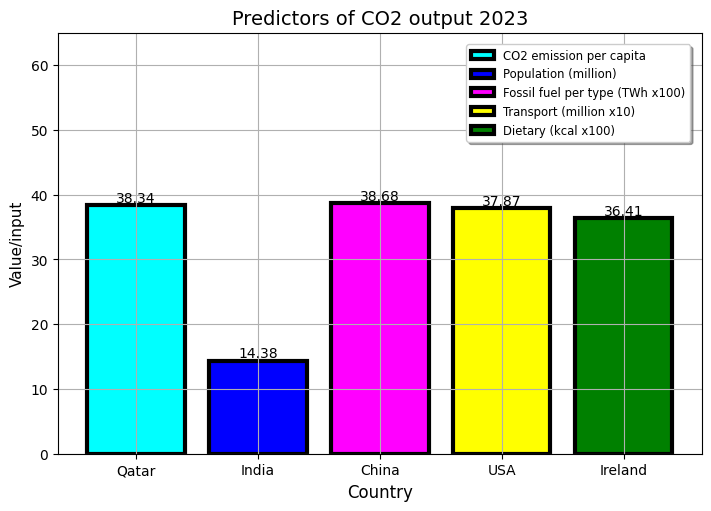


Chart: *Predictors of CO2 output*

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

The Paris Accord was signed by 195 countries/nations in 2016, meaning all of these countries are actively reducing their carbon dioxide emissions and there is no more difference between developed and developing countries. (source: *United Nations* (UN))

For this topic, decrease in CO2 output, analyses are made for 2023 only. Starting with the relative CO2 emission versus population, the top 5 countries are from Europe (see chart below). Bulgaria has the highest relative reduction of -23% and a population reduction of -3%.

The lowest consumption of fossil fuels (kWh) per capita is by Bangladesh, whereas the fossil fuels by fuel type (total) is by Latvia.

Analysing the car sales for electric/hybrid vehicles, Norway stands out with 93% of the total cars sold that were electric or hybrid, in total 110.000 cars. The highest total of electric cars sold is 8.100.000 (8.1 million) by China which is 38% of the total amount of cars sold in 2023.

Analysis on renewable energy technologies reveals that China has the largest production of all renewable energy sources. The USA follows on wind- and solar energy. Iceland is the largest with renewable energy per capita, followed by Norway.

The overall outcome per major continent is that China , USA and Scandinavian countries exceed in actively reducing their CO2 emissions output by investing, producing and distributing renewable energy.

The graph below shows the top 5 countries with the largest emission reduction in 2023.

A graph of the co2 emissions

AI-generated content may be incorrect.

Chart: *Largest decrease CO2 emission in 2023*

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

After making a linear regression for each renewable energy technology, it is visible that most of these technologies are reducing in price and will in the future as well. Hydropower technology is the only one that is increasing in price per kWh. Predictions show that the price will rise to $0.06 per kWh in the future.

The best price now and in the future is for Onshore Wind energy. In 2023 the price was still above $0.03 per kWh, future predictions show that the price will drop below $0.03 per kWh. (see graph below). Solar photovoltaic energy technology is also getting cheaper in the future. In 2023 it was still above $0.04 per kWh, and is expected to drop to $0.03 kWh.

Bio-energy and Geothermal energy technologies are slowly reducing in their prices compared to all wind and solar energy technologies. This is visible by looking at both the blue dots and red line. In the graph below the red line is very steep, this means that the prices are reducing faster than Bio-energy and Geothermal technology.

The most expensive renewable energy is for Concentrated Solar energy. Even though the prices of the Hydropower energy is rising, in 2023 the price for Concentrated Solar was $0.12 per kWh and is expected to drop to $0.10 per kWh. Future predictions for Hydropower energy will be between $0.05 and $0.06 per kWh, making Concentrated Solar energy the most expensive of all the renewable energy technologies. For all the graph see appendix: *Graphs Renewable Energy.*

The graph below shows the prediction of future prices for Onshore Wind energy technology.

*A graph with a red line and blue dots

AI-generated content may be incorrect.*

Chart: Linear regression/ prediction Onshore wind energy price

## 6. Conclusion

The population has the largest impact on the overall CO2 emission output and per capita. The larger the population, the lower the numbers per capita become. For example, Qatar has the largest fossil fuel per capita but China has the overall largest fossil fuel consumption. This is because the population in China is much larger than in Qatar. Therefore China, India and USA aren’t in the top 5 with the fossil fuel consumption per capita or Qatar not in the fossil fuel per type overall. Only the GDP per capita reduces when the population grows.

Other factors like transport and dietary composition are also influenced by the population. More people mean more food is needed, therefore food production (greenhouse gas) and distribution (transport) is needed as well. Other industries such as car sales and (inter)national shipping are influenced by the population and the consumption of fossil fuels.

Due to the Paris Accord, a lot of countries are making efforts to reduce the overall CO2 emission output. China, USA and Europa are leading the way. The largest electric car sales in 2023 is by Norway, Iceland has the largest renewable energy consumption per capita, and both China and USA are investing and producing in renewable energy technologies. Future predictions show that the prices for most of these technologies will reduce further.

## 7. Appendices

Github – <https://github.com/JoKa1986/Final-Project-CO2-emission---Winc-Academy>

Analysis Report: placed in Guthub map

Graphs of Renewable Energy: placed in Guthub map

Analysed datasets: placed in Github map

Colab Notebooks: placed in Github map

Colab Notebooks links:

* *CO2 output*:

<https://colab.research.google.com/drive/1Nf8757FkmgZlecEHYOoq-LbpqDr4zOD0>

* *Active Decrease CO2:*

<https://colab.research.google.com/drive/1nSxee6nOVlpVbXISOTUvPg_FnTO9oe_h>

* *Renew. Tech:*

<https://colab.research.google.com/drive/1kItj19nOdvd9AX48D2ZxFWG4a7R99ncx#scrollTo=tgKlQuEcLR3Q>

## 8. References

United Nations (UN) - Paris Accord (2015) – list of countries

Our World in Data – for datasets and reference map of countries with active plan

Matplotlib documentations – for charts

Pandas documentations – for analysis tools

Google Search – example reports and charts