Exploring CO2 Emissions, Greenhouse Gas Emissions, Fossil fuel energy consumption, and Oil-based Electricity Production: Statistical Analysis & Correlations

Ghazanfar Ali

Abstract- This report presents a statistical analysis and correlation insights on multiple environmental indicators, including CO2 emissions, greenhouse gas emissions, Fossil fuel energy consumption (% of total), and electricity production from oil sources (% of total). The report aims to analyze and understand the relationship between these key environmental factors and provide insights into their visualizations and correlations.

https://github.com/Ghazanfar999/ADS-1-assignment-2-rework/tree/cdf4f3e237c4955713bf9e5864bb23906a620c7f

I. INTRODUCTION

We evaluate the trends, correlations, and statistical properties of these factors using Python programming and data analysis techniques. The objective is to acquire a deeper understanding of their interrelationships and their implications for environmental sustainability and energy planning.

We reveal the interdependencies between CO2 emissions, greenhouse gas emissions, Fossil fuel energy consumption (% of total), and oil-based electricity production using visualizations and correlation analysis. These findings can assist policymakers and stakeholders in making informed decisions and devising sustainable future strategies.

It is essential to recognize the limitations inherent in the dataset and the World Bank's credibility as a source of climate change data.

Using data from the World Bank, this report provides a concise analysis of important environmental factors in order to comprehend their interrelationships and contribute to environmental sustainability and energy planning.

II. EXPLANATION

This report is mostly about analyzing statistics from the World Bank about the environment. The data set includes information about important things like CO2 emissions, greenhouse gas emissions, Fossil fuel energy consumption (% of total), and power production from oil sources as a percentage of the total energy mix.

Downloading the CSV files from the World Bank was the first step in the research. Then, these files were cleaned to get rid of any missing values or columns that weren't needed. This made sure that the data could be used for further research.

Then, different statistical methods were used to learn more about the information. Measures like mean, standard deviation, and summary statistics were used to figure out how big the surrounding factors were and how much they changed over time. This gave a full picture of the statistical properties of the sample.

By figuring out the correlation values, it was possible to find out how strong and in what direction these relationships were. This study was helpful because it showed how changes in one variable might affect the others and how they might be linked.

Line graphs, Bar graph pie charts were used throughout the report to show the data in a way that was clear and easy to understand. These pictures made it easier to find trends, patterns, and comparisons between the variables, which made the dataset easier to understand and analyze.

For code, importing the necessary libraries was the initial step: **pandas, matplotlib.pyplot, numpy, and scipy**. These libraries would provide us with the tools and functions necessary to effectively manipulate the data.

Next, we defined the "**read_data**" function. Reading and manipulating CSV data required this function. It processed dataframes from filenames. The code read and manipulated the CSV file using pandas' read_csv method.

Then we defined **“explore_statistics”** and "**explore_correlations**" while searching for interesting concepts. This function let us compare two indication dataframes. Returning a correlation matrix from two dataframes It allowed for extra correlation analyses, allowing for data exploration.

After understanding the functions,we have to load these datasets from local files into dataframes. The file paths were variables: "**co2_emissions_filename**" and "**ghg_emissions_filename**."

We called the **“explore_statistics”** and "**explore_correlations**" function for both CO2 emissions, greenhouse gas emissions, Fossil fuel energy consumption (% of total), and oil-based electricity production dataframe. The function provided summary statistics for each indicator, revealing distribution, central tendency, and variability. These statistics would help us comprehend the data and find patterns.

Finally, the report employs the use of visual representations such as line graphs, bar graphs, and pie charts to make the facts presented clearly and easily.

Overall, The study talks about how important the World Bank is as a reliable source of information about climate change.

III. INTERPRETATION OF THE RESULTS

In order to present our results and give context to the data, we prepared a table of the CO2 emissions from the output generated by the code. Although other tables are also generated from the code(see code file **22014947.py**).

The CO2 emission data provided in Table I above was generated by Python code. The data in the table were derived through statistical analyses and data manipulation techniques implemented in this code. When the code is run, the unstructured emissions data is converted into a more useful format. The code output appears to be a table displaying statistics related to greenhouse gas emissions for different countries or regions (see Table I).

The table (Table II) displays the country names in the first column and the correlation values in the second column.

The correlation value (see Table II). represents the strength and direction of the linear relationship between the "Country Name" and the corresponding variable.

Table II. Co_2Emissions Statistics Code output

Country Name	Correlation
Africa Eastern and Southern	0.981546
Afghanistan	-0.685017
Africa Western and Central	0.965438
Angola	0.983092
Arab World	0.801914

Table I. Co_2Emissions Statistics Code output

Country Name	count	mean	std	min	25%	50%	75%	max
Africa Eastern and Southern	63	10.59904	2.516189	6.444405	8.473174	10.50617	12.68458	14.87672
Afghanistan	63	9.538279	3.563721	3.309401	5.716158	11.10235	11.89748	15.06442
Africa Western and Central	63	12.33485	3.861425	4.885585	9.210733	13.58698	15.66127	17.66161
Angola	63	14.35061	6.956907	4.095931	8.587819	12.93611	20.72271	25.15524
Arab World	63	20.26765	2.377067	14.19206	19.32466	21.22899	22.09476	23.15892
United Arab Emirates	63	55.62204	5.616824	42.71431	54.30808	56.10058	59.03505	66.62229
Argentina	63	43.38793	0.775523	41.11526	42.87691	43.39967	44.07455	44.57091
Armenia	63	33.85761	2.87924	28.24145	32.23808	34.25105	35.83722	39.27496
Australia	63	60.17349	2.232119	54.19995	59.0251	60.95721	61.50955	63.45899
Austria	63	20.76366	1.016846	19.30892	20.04164	20.57822	21.53373	23.08216
...	...	...	...	...	...	...	...	...
South Africa	63	26.84176	4.677401	22.20657	23.18736	24.26585	30.66385	36.94971
Zambia	63	9.737072	3.142644	2.91534	7.939628	9.952483	11.96466	

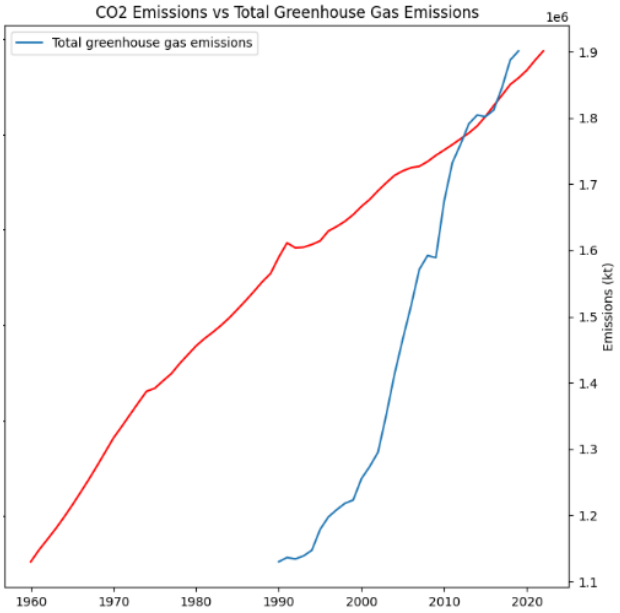

Fig.1. Co_2Emissions vs Total Greenhouse gas Emission

Figure 1 compares CO2 and greenhouse gas emissions over time. The graph's two y-axes indicate CO2 emissions (red) and total greenhouse gas emissions (blue). The x-axis shows the time and kilotons (kt) of emissions.

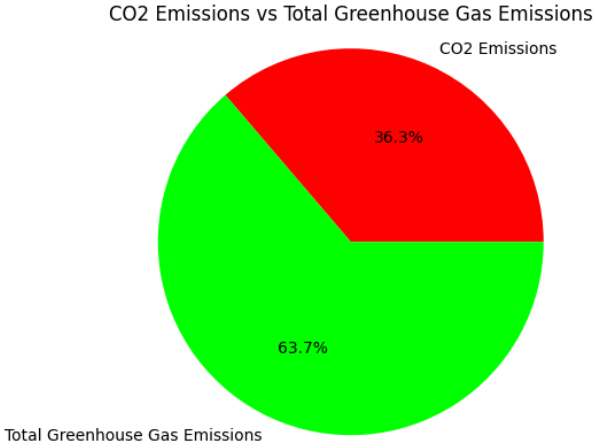

Fig.2. Co_2Emissions vs Total Greenhouse gas Emission(pie graph)

The second diagram is a pie chart, and it shows how much of the total greenhouse gas emissions come from CO2 emissions. Each "slice" of the pie chart shows the percentage of data for one of the categories. Different sized slices represent different amounts of emissions. (See Figure 2).

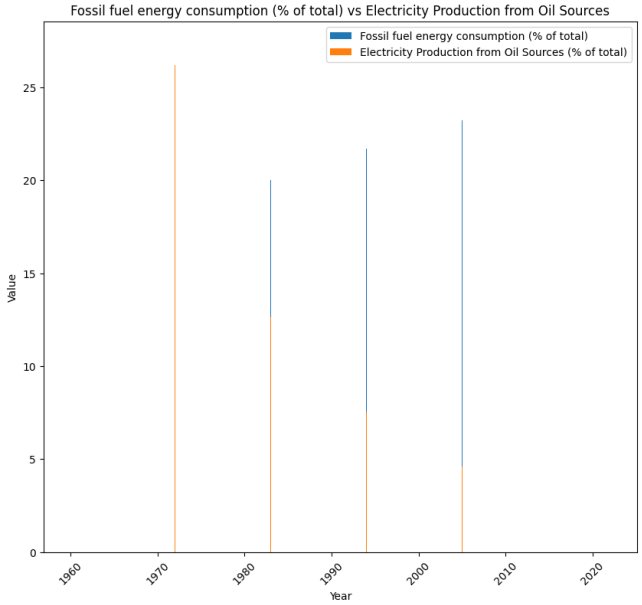

Fig.3. Fossil fuel energy consumption vs Electricity Production from Oil Sources (Bar graph)

Lastly, we generate a bar graph (see Figure 3) comparing "Fossil fuel energy consumption (% of total)" and "Electricity production from oil sources (% of total)." The years are shown by the x-axis, while the values are represented by the y-axis.

IV. CONCLUSION

The World Bank's analysis of environmental data reveals significant correlations between CO2 emissions, greenhouse gas emissions, Fossil fuel energy consumption, and oil-based electricity production. These findings highlight the significance of environmental impact consideration in energy planning. The reliability of the World Bank's data and the use of Python programming enhance the validity of the analysis.