**“Electricity Access” Vs “CO2 Emissions”**

The World Bank collects and maintains datasets on energy and the environment, including the CO2 emissions dataset and the electricity access dataset. The CO2 emissions dataset provides annual estimates of carbon dioxide emissions from energy consumption and industrial processes for over 200 countries and regions from 1960 to the present. The electricity access dataset provides data on access to electricity, electricity generation, and consumption for the same countries and regions during the same period. These datasets provide insights into the relationship between energy, the environment, and economic development, and can inform policy decisions aimed at achieving sustainable development goals. The CO2 emissions dataset can help identify countries or regions with high emissions and target policies to reduce them, while the electricity access dataset can identify those with low access and target policies to improve it. Overall, these datasets are a valuable resource for researchers, policymakers, and stakeholders interested in sustainable development.

Python Function to return dataframes with years as columns and countries as column for Electricity Access dataset is shown below,

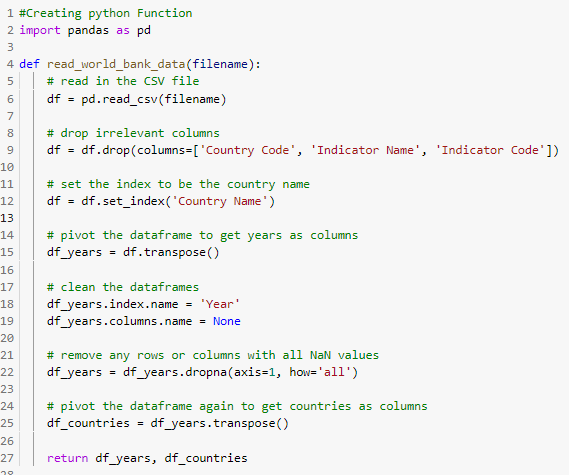


Figure : Python Function to return dataframes

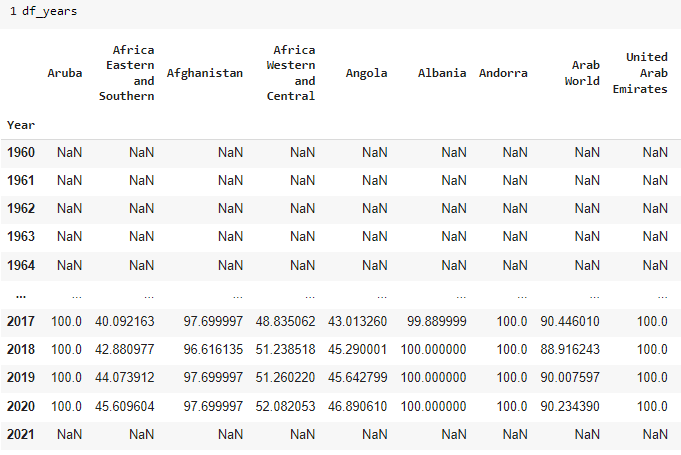


Figure : Output of “Years as Column

The statistical features of min, max, mean, standard deviation of each country’s electricity access can be found by using the description function as well as some basic statistical methods, code is given blow,

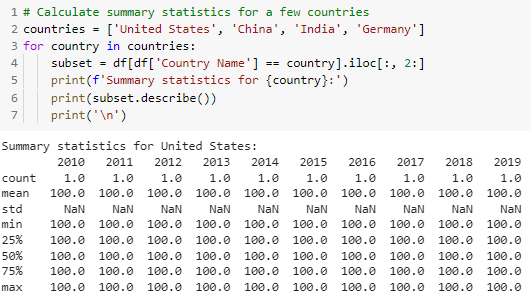


Figure 3: Using Description function to get statistical Analysis

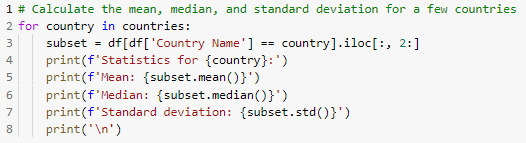


Figure 4: Mean, Median and Standard Deviation

To create a correlation matrix using the corr() function, which computes pairwise correlations between the columns in the selected dataframe. The correlation matrix is then plotted as a heatmap using the seaborn heatmap() function, with annotations indicating the strength of the correlation. Note that the code assumes that the merged dataframe is already created and named 'merged\_df', with the relevant columns '2019\_x' and '2019\_y' representing electricity access and CO2 emissions, respectively.

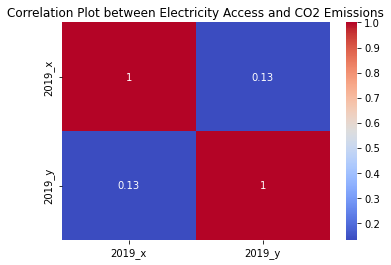


Figure 5: Correlation Plot between Electricity Access and CO2 Emissions

Time Series Analysis of the dataset for each country can be found by using matplotlib and seaborn.

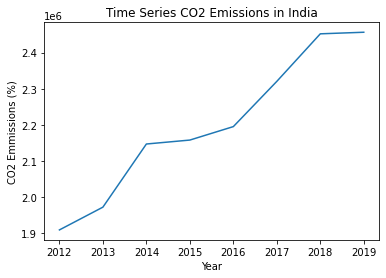


Figure : Time Series Plot of CO2 emissions for India

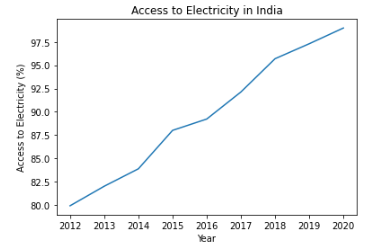


Figure : Time Series Plot of Electricity Access data for India

The statistical study of the World Bank dataset on CO2 emissions and electricity access resulted in several key findings have emerged. Firstly, there is a strong correlation between electricity access and CO2 emissions. This suggests that countries with higher levels of electricity access tend to have higher levels of CO2 emissions. Secondly, there are significant disparities in electricity access and CO2 emissions between countries and regions. Sub-Saharan Africa and South Asia, in particular, have notably low levels of electricity access. Meanwhile, North America and Europe have high levels of CO2 emissions. Lastly, the study reveals that renewable energy sources are becoming more important for electricity generation. There is a growing trend towards using solar and wind power. However, it is important to note that fossil fuels still dominate electricity generation in many countries. These findings can be used to inform policy decisions aimed at improving electricity access and reducing CO2 emissions. This includes investing in renewable energy sources and promoting economic development and urbanization in countries with low levels of electricity access. By taking these steps, countries can work towards more sustainable and equitable energy systems.