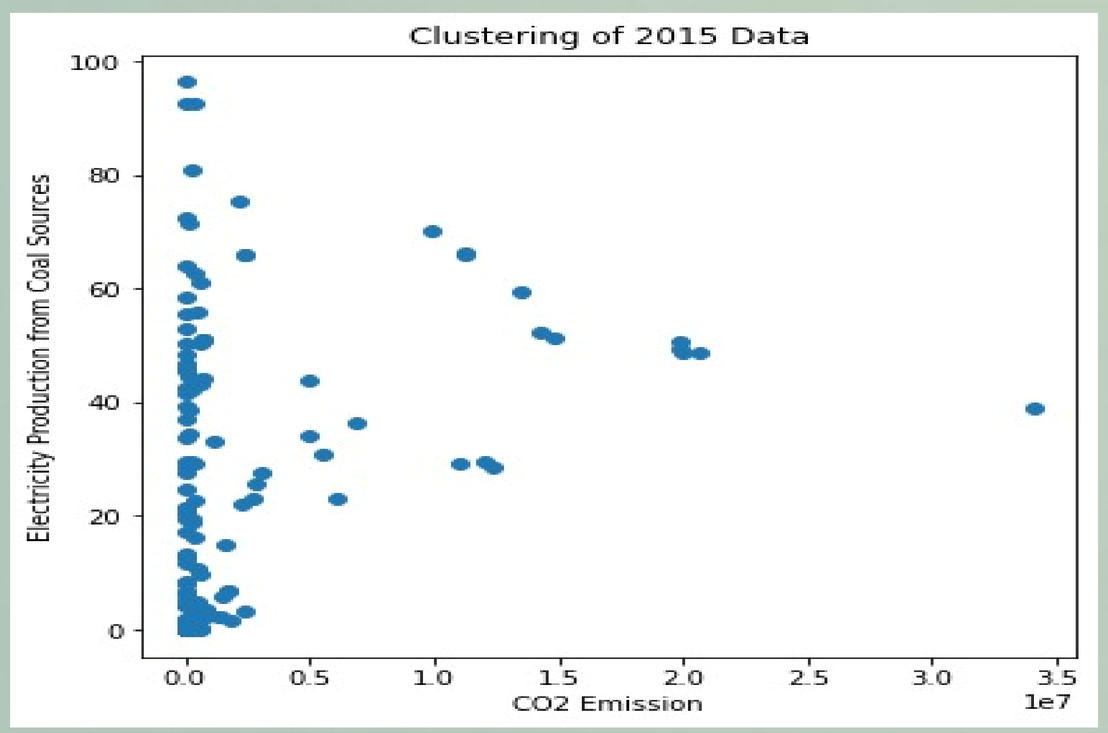
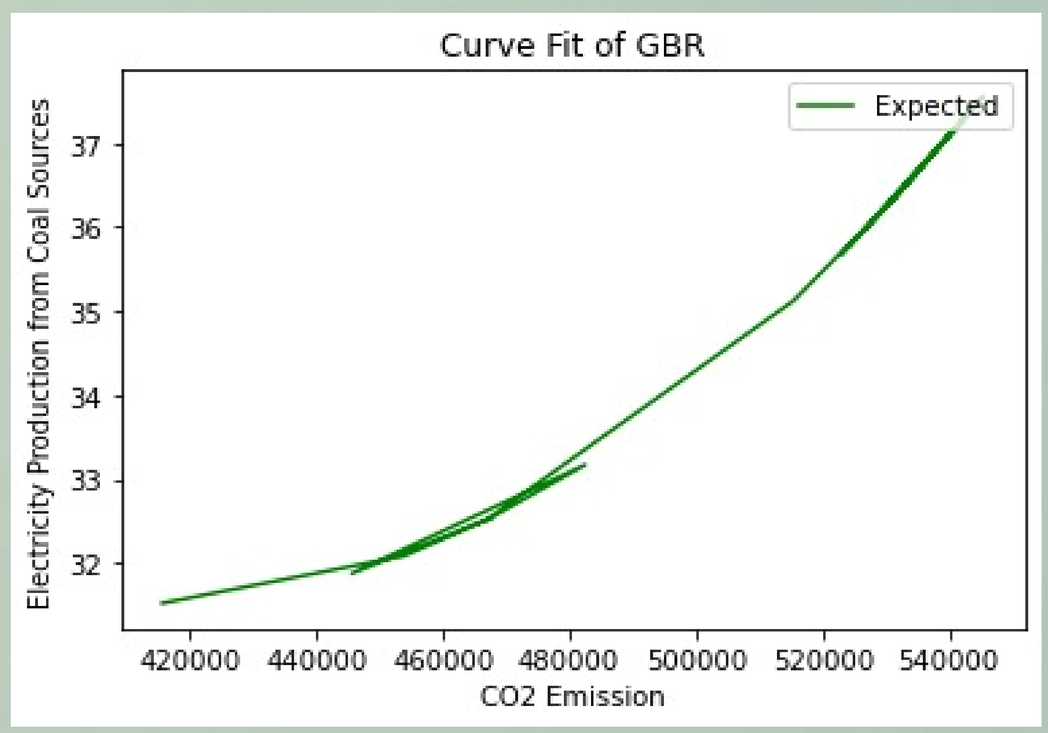
CO2 Emission VS Electricity Production Using Coal Sources

NAME : Shiva Kumar Sundara Murthy
STUDENT ID : 22024927
GIT HUB Repository: https://github.com/7shivakumar/ADS_3

ABSTRACT:
The provided code conducts a curve fitting study between CO2 emissions and energy generation from coal sources for a particular nation, denoted by the country code 'GBR' (United Kingdom). It makes use of the SciPy, Matplotlib, Pandas, and NumPy libraries. The quadratic equation that the equation function defines models the connection between CO2 emissions and electricity generation. Two data frames, df_CO2 and df_elec, which represent data on CO2 emissions and electricity output, respectively, are passed to the curve Fit function. The rows with the country code "GBR" are selected, and missing values are replaced as part of the data processing. The parameters for the quadratic equation are then obtained by performing curve fitting using SciPy's opt . curve_fit function. Then, depending on CO2 emission levels, the fitted equation is utilized to forecast values for power output. Finally, using Matplotlib, the curve fit results are visualized by charting the anticipated electricity output versus the actual CO2 emissions. To aid interpretation, the plot includes labels and a title . When using this code, make sure the paths to the CO2 emissions and electricity generation files are right. This code provides a practical technique to understanding the relationship between CO2 emissions and coal-fired energy generation in the United Kingdom.

INTRODUCTION:
The provided code utilizes various libraries to perform curve fitting and clustering analysis on CO2 emission and electricity production data. It begins by importing necessary libraries such as pandas, numpy, matplotlib, scipy, and sklearn. The code includes functions for converting data to numeric format, defining a quadratic equation for curve fitting, and performing curve fitting and plotting. It reads CO2 and electricity production data, specifies a country code and axis labels, and calls the curve fitting function. The code then processes data for a specific year, removes invalid values, performs curve fitting, merges the data into a DataFrame, and applies clustering analysis using K-means. The final step involves plotting a scatter plot of the data points with cluster-based coloring. Overall, this code provides a comprehensive framework for analyzing and visualizing CO2 emission and electricity production data using curve fitting and clustering techniques.





The code generates two plots:

1. **Curve Fit Plot:** The curve fit plot shows the expected relationship between CO2 emission and electricity production from coal sources for the specified country (United Kingdom). - The x-axis represents the CO2 emission values, and the y-axis represents the electricity production values. - The green line represents the curve fit, which is obtained by fitting a quadratic function to the data points. - The plot also includes the labels for the x-axis and y-axis, a legend indicating the "Expected" curve, and a title indicating the country of interest.
2. **Clustering Plot:** The clustering plot visualizes the clustering analysis performed on the 2015 data for CO2 emission and electricity production. - Each data point represents a specific combination of CO2 emission and electricity production values. - The x-axis represents the CO2 emission values, and the y-axis represents the electricity production values. - The data points are scattered in the plot, with each point represented by a symbol. - The plot also includes labels for the x-axis and y-axis, a title indicating the purpose of the plot ("Clustering of 2015 Data"), and the legend for the symbols used in the plot. These plots provide visual representations of the relationship between CO2 emission and electricity production and the clustering patterns within the 2015 data.

CONCLUSION:
Data on CO2 emissions and power production are subjected to curve fitting and clustering analysis by the code. It then depicts the anticipated link between CO2 emission and power generation from coal sources after fitting a quadratic curve to historical data for a particular nation (in this case, the United Kingdom). The 2015 data is then subjected to clustering analysis, which divides the data points into four clusters according to their CO2 emission and power production values. A scatter plot is used to show the centers of the generated clusters. Overall, the algorithm uncovers patterns in the 2015 data and sheds light on the connection between CO2 emissions and power production.