

USA Power Plant CO2 Emissions Analysis

USA CO2 Emissions 2016-2019

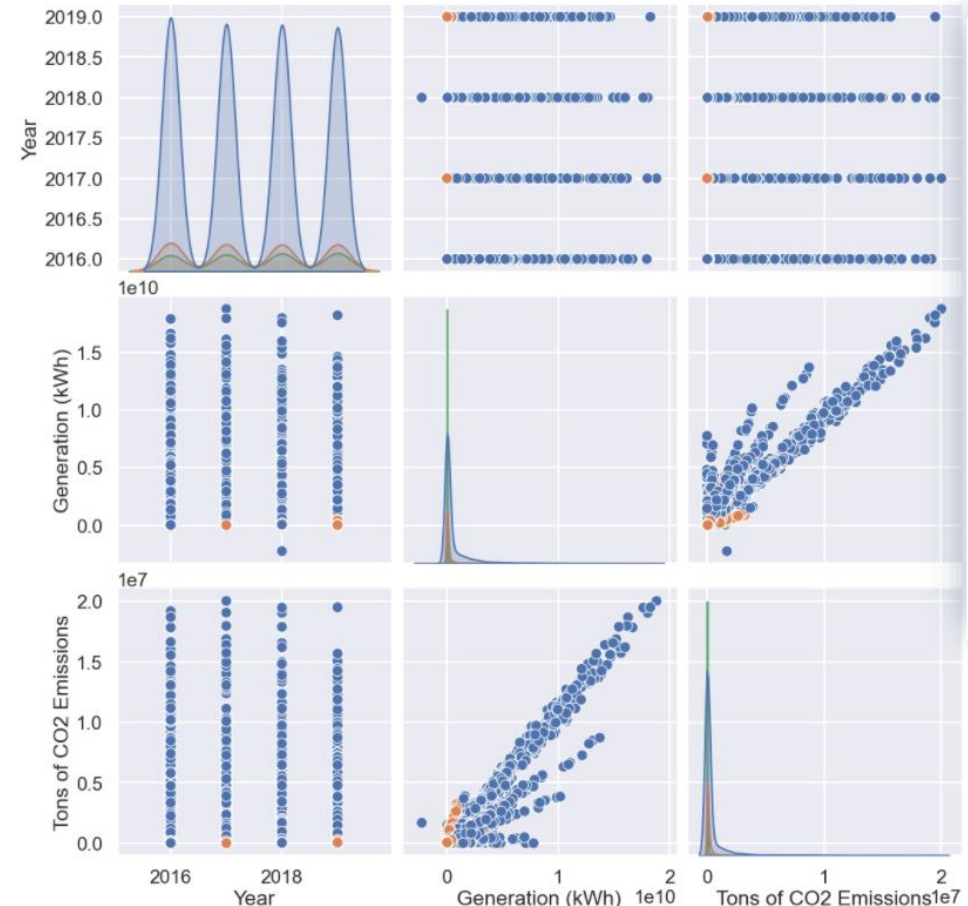


Objective and Key Questions

This Power Plant CO2 Emissions analysis focused on finding demonstrable statistical evidence showing that electric power generation has a direct impact on CO2 emissions, but the following questions were also posed:

- How much CO2 emissions are associated with electricity generation?
- How much electricity is produced in each sector group, and how much fuel is required to produce this?
- Which state, sector group, and NERC regions produce the highest total emissions?
- Which electric plants emit the most CO2 emissions, and in which state, sector group, and NERC region are they located in?

Finding Relationships in the Data



Electric Power
49%

Commercial
27%

Industrial
14%

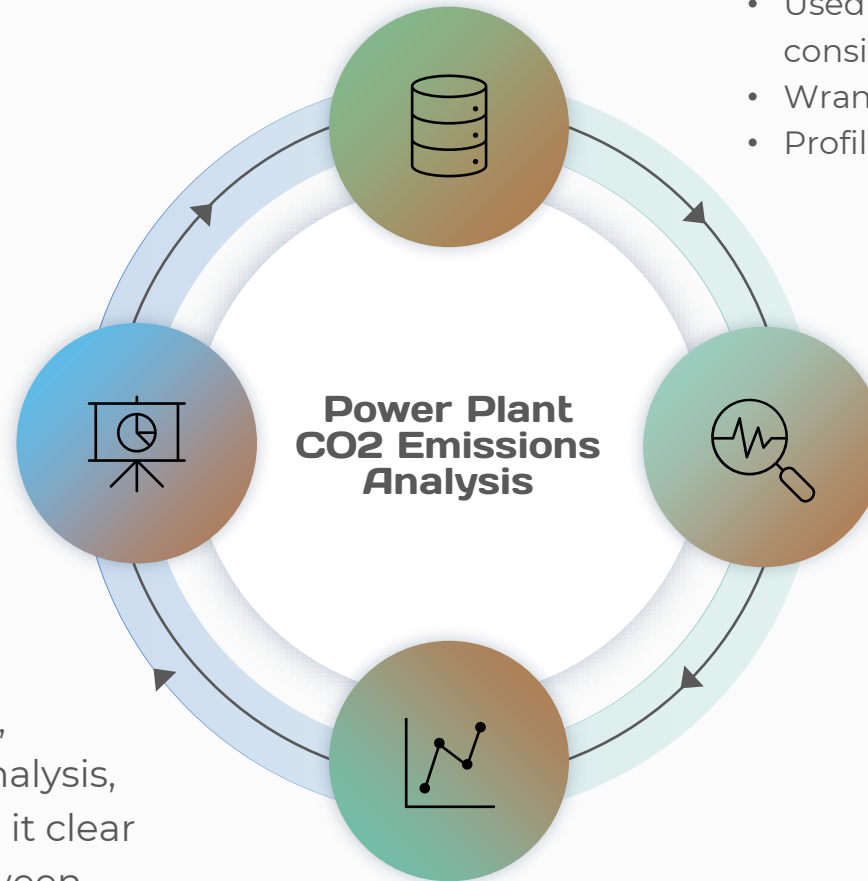
Project Life Cycle

Notes

- It is likely that the stationarized data is sufficient for fore-casting since the EIA (Energy Information Administration) has only annual data, and any monthly or daily imputations would insert possible bias into the data.

Statistical Analysis

- The results of exploring relationships, performing regression and cluster analysis, and use of geospatial analysis makes it clear that there is a close relationship between power generation and higher CO2 emissions.
- This allows us to use the cluster model on other variables to gain greater insights



Data Preparation

- Sourced public data from a reliable source
- Used Python to assess data accuracy, quality, consistency, completeness, uniqueness, and timeliness
- Wrangled, cleaned, transformed and merged the data
- Profiled the data using basic statistical methods

Data Exploration

- Explored data relationships using correlation heatmaps, scatterplots, pair plots, and categorical plots
- Performed linear regression analysis and cluster analysis using the elbow technique and k-means clustering
- Used shapefiles in a geospatial analysis
- Explored geographic variables using choropleth maps to draw early insights
- Performed time series analysis, including decomposition, tests for stationarity, and the stationarization of the data

Conclusions and Recommendations

The results of the analysis has demonstrably and statistically answered the most important question about power plants and CO2 emissions, and we are now able to ask and answer the other questions by exploring the charts and map on the [interactive dashboard here](#).

My recommendation is to glean as much insights from the dashboard charts to narrow the focus and take the next step in learning more about power plant CO2 and other harmful emissions and the reality of our reliance on the electricity they produce.

Further details of this **CO2 Emissions Analysis** can be found in my [GitHub repository here](#).

