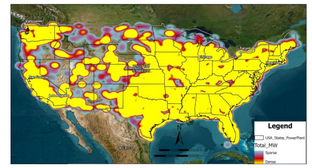
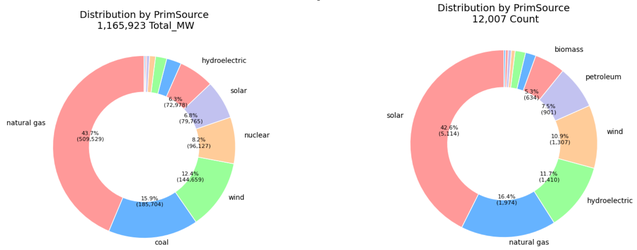
Analysis of Power Plant Energy Generation in the United States using Machine Learning and Geographic Information System (GIS)



Cristian C Noriega M, Blessing Austin-Gabriel, Ernest Chianumba, Rolih Ferdinand, Dr. Aparna Varde
(School of Computing, and CESAC)

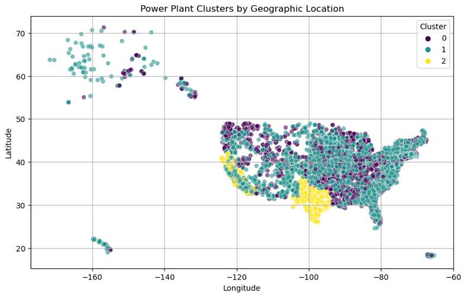
Goal

While Natural Gas is leading energy source in the United States, renewable energy sources are poised for future dominance. We aim to analyze data on them.

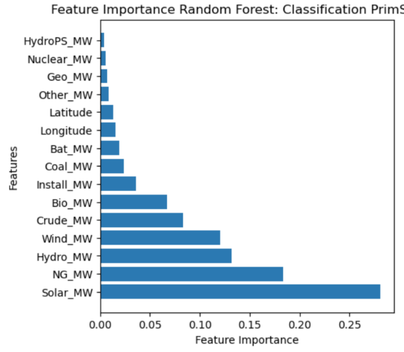


The main goal is to study the distribution and classification of energy generation based on demographics and power source.

Methods



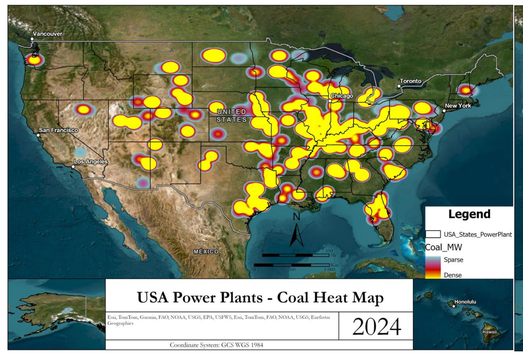
Unsupervised machine learning using K-Means clustering to fathom how power generation areas are grouped.

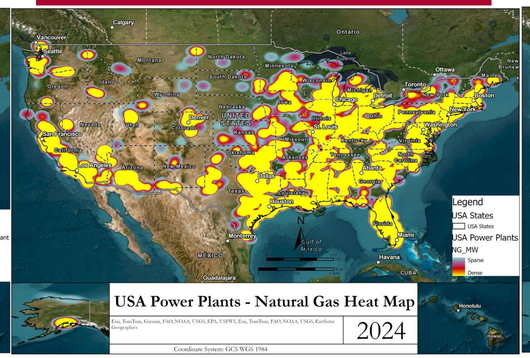


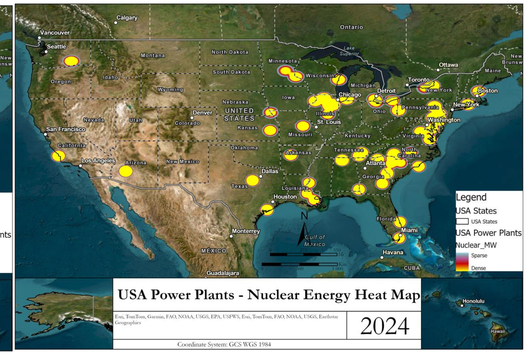
Supervised machine learning using decision trees and random forest to classify the power sources.

Decision Trees & Random Forest: What exactly is causing these & why?

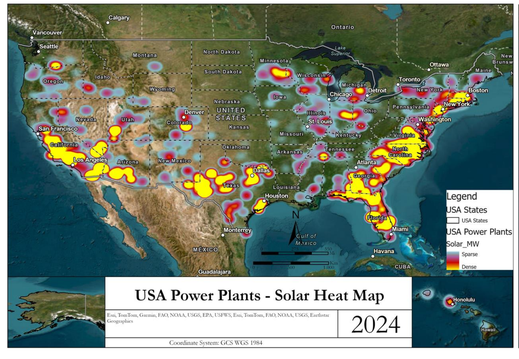
Non-renewable energy

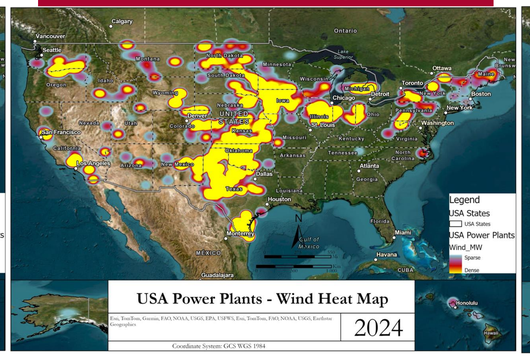


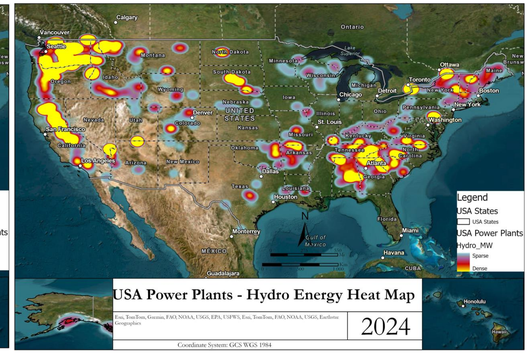




Renewable energy







Conclusions

- The greatest energy producing source in the US is natural gas.
- The greatest number of power plants are solar.

Data source

The U.S Energy Information Administration, May 2023

Future work

- How can the energy distribution and findings be modeled using GIS in a user-friendly and interactive manner?
- Address CO2 emissions with respect to transition towards cleaner energy, and the demographics.