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

According to the U.S. Energy Information Administration (EIA) survey from May 2023, 42.6% of power plants in the United States are solar, followed by natural gas at 16.4%. However, natural gas contributes 43.7% of the total energy generation in the US, while solar only contributes 6.8%. Additionally, while Texas is the major contributor to energy generation in the US, California boasts double the number of power plants. This study employs unsupervised and supervised Machine Learning (ML) techniques, such as K-Means and K-Nearest Neighbors (KNN), alongside Geographic Information System (GIS), to analyze the energy generation pattern in the US. Key attributes for clustering analysis include the source of energy, total installed capacity, total energy generated, and geographical coordinates.

Notably, one cluster comprises natural gas, nuclear, and coal, with Texas and Florida as top contributors, while California does not rank within the top 10. Remarkably, this cluster contains no renewable energy sources and exhibits the characteristic of the greatest disparity between total installed capacity and total energy generation. Likewise, another cluster comprises wind and solar with Texas and California as top contributors, showing the lowest disparity between total installed capacity and total energy generation.

A graph of different colored bars

Description automatically generated

CLUSTER 1: No renewable energy (Texas & Florida as the top contributors)

A graph of different states

Description automatically generated

CLUSTER 2: Renewable energy (Texas & California as the top contributors)