

# **Power Generation in the US**

Kiran Kour

MIS-64060: Fundamentals of Machine Learning

Professor Dr.Murali Shanker

Kent State University

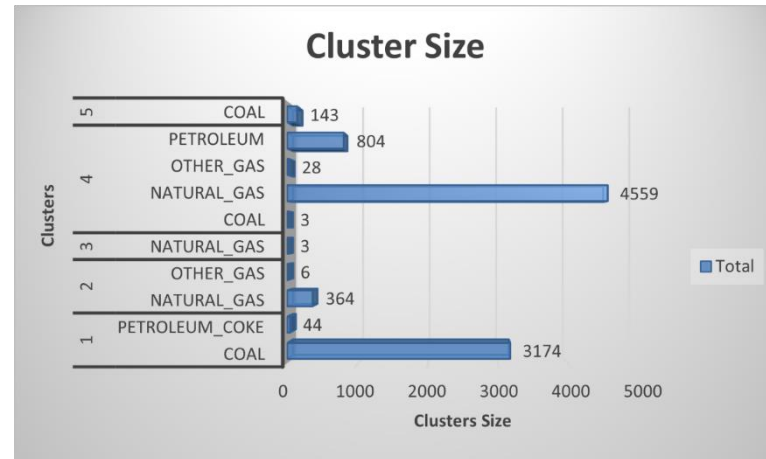
## **Executive Summary**

This report summarizes the use of machine learning algorithm to understand power generation in the U.S. By using k-means clustering, the clusters formed were analyzed and interpreted to understand the power generation and how the best segmentation will help make a clean energy future in the U.S considering the type of fossil fuel that will help to achieve this goal. The best Cluster was selected considering the variables used in the analysis. Cluster **The Clean Power Future** was the economical cluster with less heat production, the average cost per MMBtu of heat content, and less emission of harmful pollutants. The EIA sources project that the share of natural gas-fired generation in the United States will remain relatively constant through 2050. The contribution from coal and other fuel types will drop by half. With the maximum number of Natural gas units and a small amount of coal, petroleum, and other gases, Cluster **The Clean Power Future** seems to be an ideal cluster to achieve this goal and explain the future generation of clean power in the U.S. In my opinion, use of Natural gas in the future should be increased in the power generation keeping in mind the demand and contributions it will make towards the clean future.

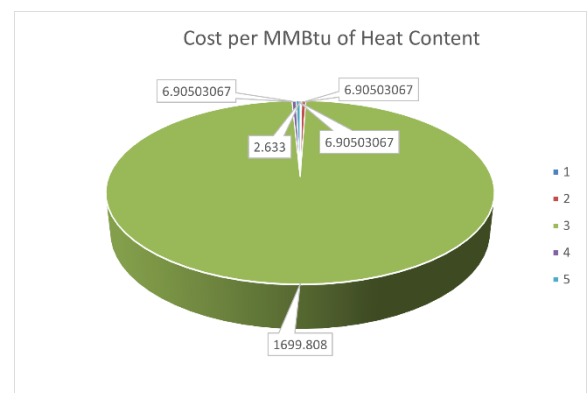
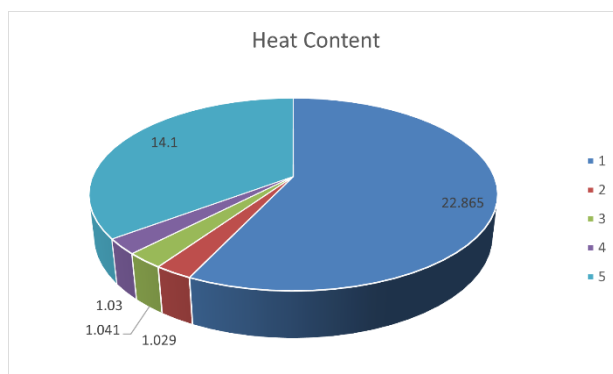
## **Problem Statement**

- Which best segmentation will help to understand the power generation in the U.S.?
- To predict which segmentation will help make a clean energy future in the U.S.
- To identify which fossil fuel will contribute to making a clean and healthy environment in the future.

- To compare the cost and units of fuels being used between the best segmentation helping to understand power generation and the segmentation that will help to achieve the goal of reducing harmful pollutants.



The graph above shows the distribution of fossil fuels among the clusters and the size of each cluster.



Clusters	fuel_received_units	fuel_mmbtu_per_unit	fuel_cost_per_mmbtu	gases_emission
1	23170.5	22.865	2.633	10.14
2	2586365	1.029	6.90503067	0.008215271
3	10	1.041	1699.808	0.008215271
4	16719	1.03	6.90503067	0.008215271
5	2950	14.1	6.90503067	38.71821527

According to the above two Pie chart figures and the aggregate function table:

The best segmentation that explains the power generation in the U.S. is **Cluster 1- The Thermogenic plants.** As the name suggests, it produces the maximum heat essential for producing power. This cluster, consisting of Coal as its prominent fossil fuel, has a maximum heat content of 22.86 MMBtu compared to the other clusters. The higher the heat content, the higher the power-generating capacity. Therefore, this cluster would be the best segmentation to understand power generation.

### **Facts:**

According to EIA, in 2021, the electric power sector accounted for about 37% of U.S. natural gas consumption. Natural gas was the source of about 32% of the U.S. electric power sector's primary energy consumption. As stated by EIA, **natural gas accounted for about 38% of all sectors' total utility-scale U.S. electricity generation in 2021.**

#### **From the environmental point of view:**

- ❖ Burning natural gas for energy results in fewer emissions of nearly all air pollutants than burning coal or petroleum products to produce an equal amount of energy.
- ❖ Within the power sector, coal accounts for 98% of sulfur emissions and 94% mercury emissions.
- ❖ The ash emitted by a power plant- a by-product from burning coal for electricity- carries into the surrounding environment 100 times more radiation than a nuclear power plant producing the same amount of energy.
- ❖ On the other hand, Natural gas is not a significant source of mercury, sulfur, and ash.

Considering the facts presented and identifying the best segmentation that will help make a clean energy future in the U.S. is **Cluster 4- The Clean Power Future.** The reason is that **this cluster consists** of four fuel categories, among which **Natural gas contains the maximum percentage**, followed by petroleum, other gases, and **lastly, coal**. The **heat content** production, the **average cost** of fuel per MMBtu of heat content, and the **air pollutant**

**emission** from this cluster are also **very minimal**. With **maximum units of Natural gas** and a **minimal** amount of heat production, gases emissions in this cluster will help in less emission of harmful pollutants, thereby reducing the overall harmful impact of coal, petroleum, and other gases.

### **Findings:**

- The two prominent fossil fuels used in power generation are Coal and Natural Gas.
- In terms of finding the best segmentation to understand the power generation in the U.S. is **Cluster 1- The Thermogenic plants** due to the maximum heat production and nominal cost of fuel
- In contrast, if we look at the segmentation, which will help reduce the impurities and pollutants from the environment, **Cluster 4- The Clean Power Future** would be the best choice as it consists of Natural gas as its leading fossil fuel and as we know Natural gas is the Clean-burning fossil fuel.
- According to the cluster analysis and interpretation, the agencies grouped in **Cluster 4- The Clean Power Future** seem to be major power-generating agencies that will be using Natural Gas as its foremost power-generating fuel in the U.S., keeping in mind the emerging importance of Natural gas in the power sector.
- Comparing the cost and units received by these two clusters, even though the cost of **Cluster 1 – The Thermogenic Plants** is comparatively less than **Cluster 4- Clean Power Future**, the units received are much higher than Clean Power Future. And as mentioned, Cluster The Thermogenic Plants is good for understanding power generation but not from the environmental point of view, because the gas emissions from this cluster are significant compared to Clean Power Future.

### **Conclusion**

To conclude, Cluster **The Clean Power Future** will be the best segmentation to explain future power generation in the U.S. to achieve a Clean energy future. As it is seen that previous records show that power-generating sectors are already increasing the use of Natural gas, as its one of the prominent fuels, to achieve a clean energy future by decreasing the use of fuels like coal, petroleum, and other gases to decrease the air pollutants and land

pollutants. This Cluster will not only achieve the goal but also comes out as one of the economic clusters from among all the other clusters. Even the analysis and interpretations drawn from categorical variables prove Cluster **The Clean Power Future** to be the best segmentation. The demand for Natural gas was seen when it was the only maximum purchased fuel on the spot and through a contract with a maximum number of suppliers.

But when the objective is to increase the power generation, then **Cluster The Thermogenic Plants** plays the major role, with Coal being its primary fossil fuel, containing the maximum power units would produce power on a large scale at the cost of environment and human health.

### **Appendix:**

- <https://www.eia.gov/energyexplained/natural-gas/use-of-natural>
- <https://www.epa.gov/energy/about-us-electricity-system-and-its-impact-environment>
- <https://www.fossilconsulting.com/2020/09/17/challenges-facing-the-power-industry/>