Analysis of Annual Natural Gas Consumption Trends in the United States: 2014-2024

GitHub: https://github.com/azharasif/clusteringAssignment

Assignment : CLUSTERING AND FITTING Student Name :Muhammad Azhar   Student ID: 23017233

Introduction

This report explores the statistical analysis of US natural gas usage from 2014 to present. By studying consumption patterns, we aim to uncover insights on energy demand, economic trends, and environmental impacts. Leveraging Python programming and advanced visualization techniques, we provide a detailed examination of annual consumption trends

Data Collection and Preprocessing

We read the file that contains the data from 2014-2024.Each row represents the natural gas consumption value for a specific state, sector, and process. The "value" column provides the consumption amount in millions of cubic feet (MMcf).

The key columns are:

- duoarea: State abbreviation
- area-name: State name
- product: Energy product (all rows have "EPG0" for Natural Gas)
- sector: Consumption
- process: Specific consumption process within the sector
- value: Monthly consumption in millions of cubic feet (MMcf)

Visualization and Analysis
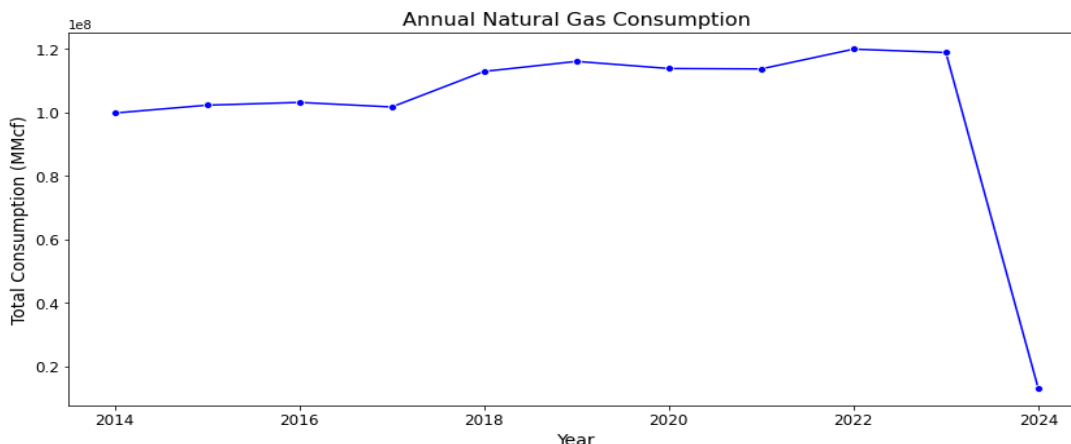
Our analysis begins with plotting annual trends in natural gas consumption, revealing overall consumption patterns over the years. The **plotted graph** illustrates the total consumption of natural gas (in mmcf) from 2014 to 2024. This visualization allows us to observe any significant trends or fluctuations in consumption over the years.

Change in Natural Gas Consumption from 2014 to 2024
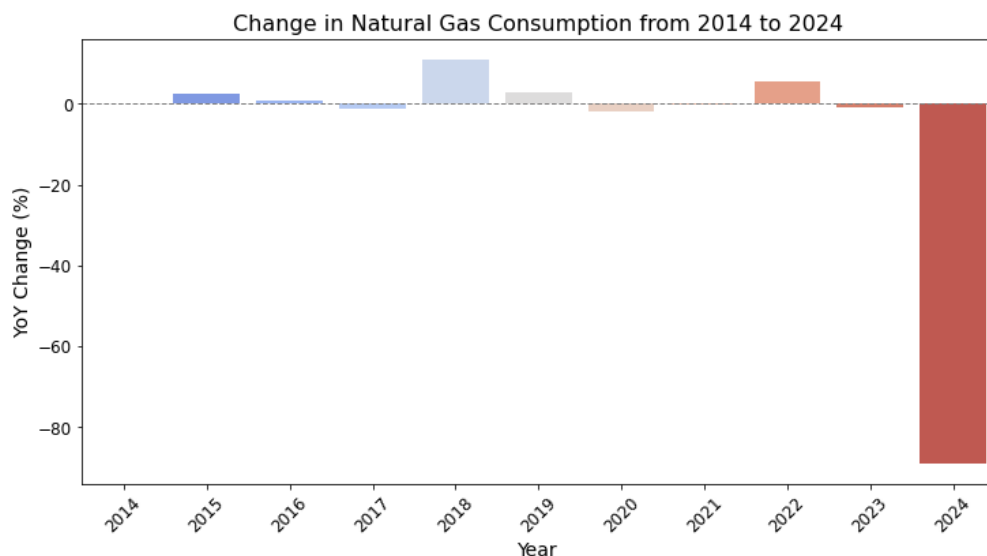
Next we need assessing the year-over-year change in natural gas consumption. By calculating the percentage we can see the insights.The **bar plot** displays the year-over-year change in consumption percentages, enabling us to identify periods of significant change and understand the underlying factors driving these fluctuations.

Natural Gas Consumption within Different sectors
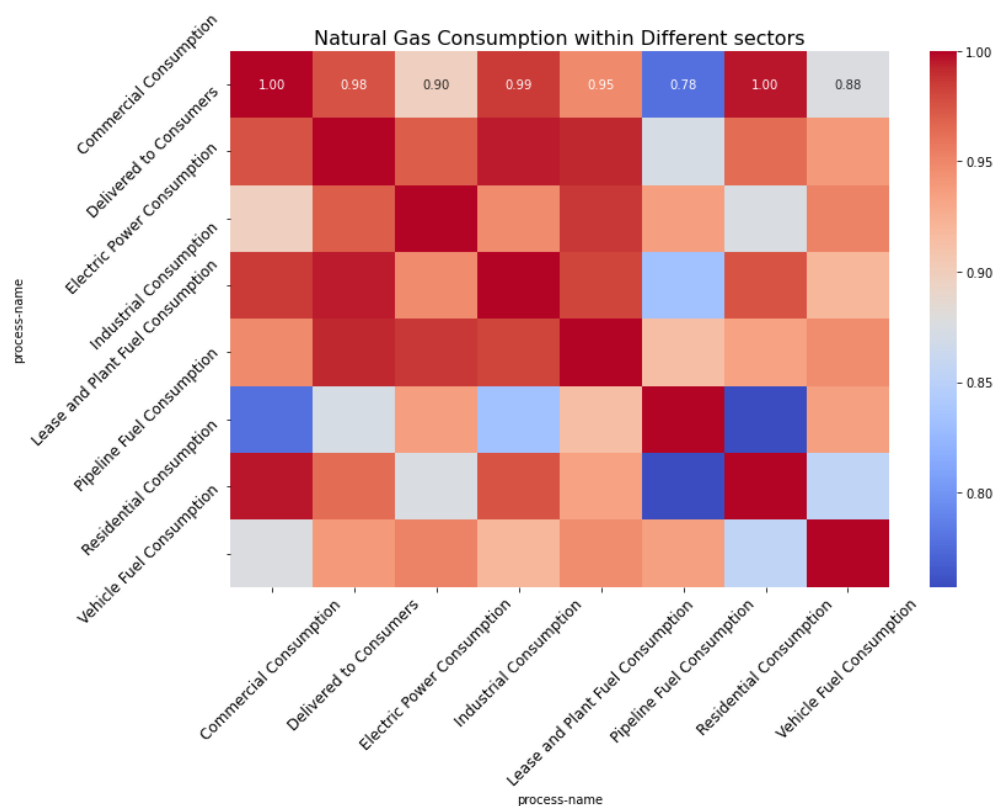
To explore the inter-sector correlation we use **heatmap visualization technique**. This heatmap illustrates the correlation matrix between different consumption sectors across various years..

K-means Clustering of Annual Natural Gas Consumption

Elbow Plot for K-means Clustering
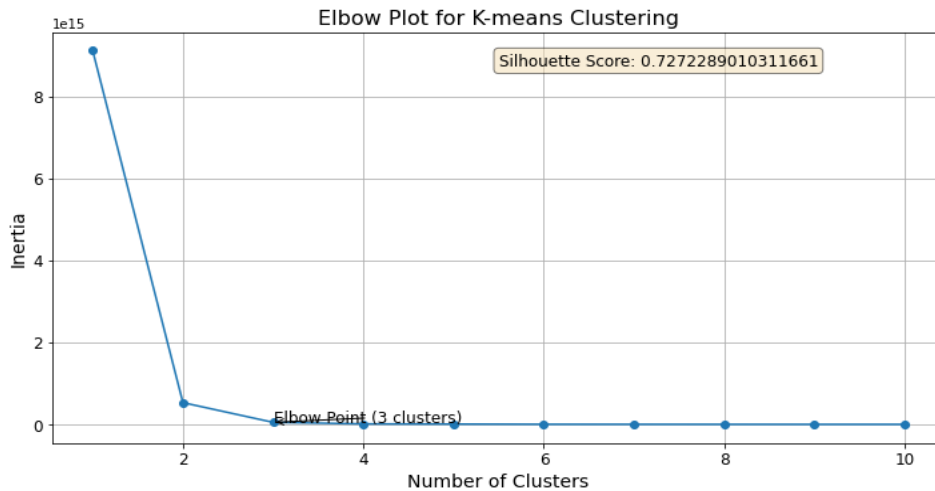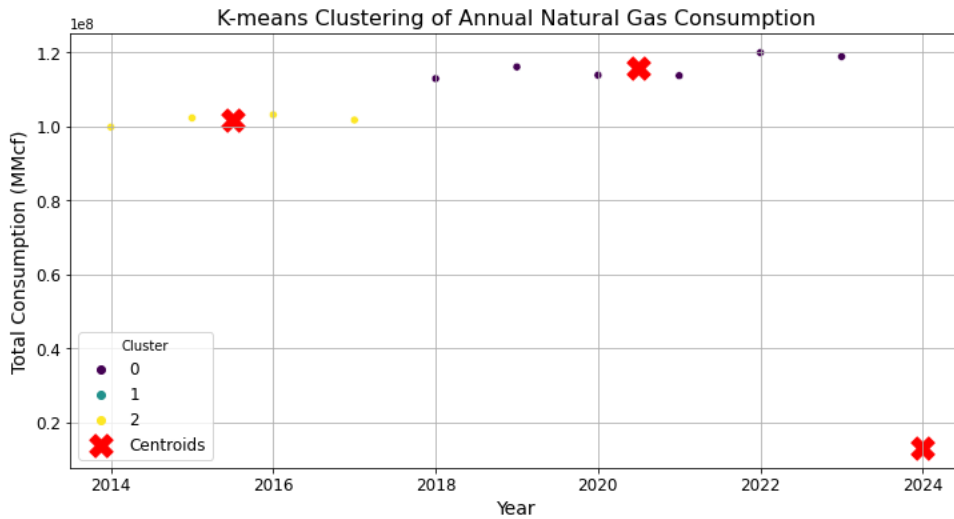
Furthermore, We employ the elbow method to ascertain the ideal number of clusters for K-means clustering analysis, opting for k=3 for the appropriate number of clusters. We calculated the **Silhouette Score** for this clustering to be **0.72** which means they are well separated overall.

Fitting Prediction of Annual Natural Gas Consumption
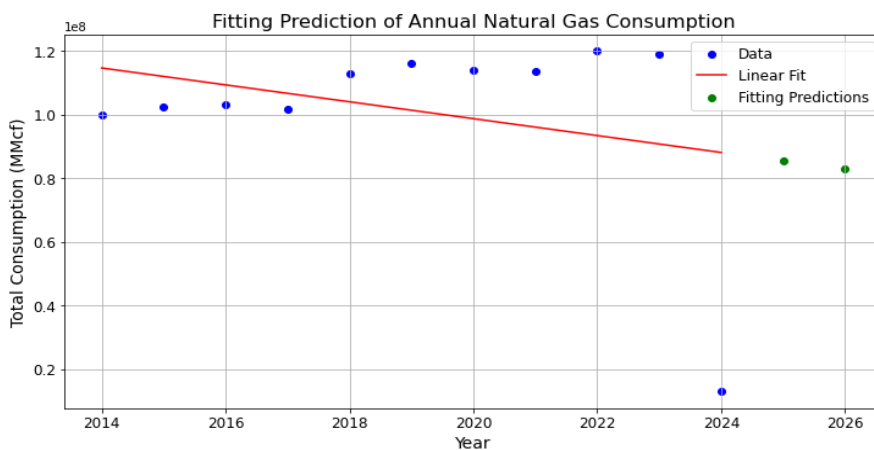
The linear regression Fitting predictions for the years 2025 and 2026 suggest consumption levels of around that will continue in the future.

References

US Energy Information Administration. Natural Gas Data. Retrieved from https://www.eia.gov/naturalgas/data.php.