EV MARKET SEGMENTATION

Vaishnavi Gunturi

Fermi Estimation:

The goal was to segment regions based on pollution levels, pollution density and also based on subsidies, road taxes and fuel prices. This analysis helps identify regions where introducing eco-friendly or electric vehicles would be most impactful. Also, by considering road taxes and fuel prices allows us to identify regions where electric vehicles could gain significant market traction.

Data Sources:

- Collected pollution data which included parameters like air quality index, population density and industrial activity.
- Another Dataset which includes economic factors such as subsidies, road taxes and fuel prices across different states.

Data Preprocessing:

- **Libraries used:** Pandas, NumPy, Scikit-learn, Matplotlib/Seaborn
- **Data cleaning:** Missing values in the AQI and population density fields were identified. Duplicate records were checked and removed to avoid bias in analysis.
- Numerical features like AQI and population density were standardized using z-score normalization.

- For skewed data, particularly in population density, a log transformation was applied to reduce skewness and stabilize variance.
- Regions were categorized based on AQI into bins to facilitate better clustering and segmentation.
- Economic features like subsidies, road taxes and fuel prices were normalized using min-max scaling. This Technique scales the data to a fixed range, typically between 0 and 1.
- A “Net Incentive” feature was created by subtracting road taxes from the subsidies, providing a clear indicator of the financial benefit to customers.

Segment Extraction (ML Techniques Used):

K means clustering is applied. This technique was used to segment regions based on pollution levels, population density, subsidies, road taxes and fuel prices. This helped in identifying high risk regions and states with favorable conditions for introducing fuel-efficient or electric vehicles and also where eco friendly interventions would have the most impact.

Profiling and Describing Potential Segments:

Pollution data segments:

1. **Segment 1: High pollution, High Population Density regions:**
    - Regions with severe air quality issues and dense populations.
    - High potential for eco friendly vehicles due to public health concerns and government intervention.

2. **Segment 2: Moderate Pollution, industrially active regions:**
    - Regions with moderate pollution levels but significant industrial activity.
    - Potential market for hybrid vehicles as a balance between cost and environmental impact
3. **Segment 3: Low Pollution, less populated regions:**
    - Regions with lower pollution and less dense populations.
    - Lower immediate demand for eco friendly but could be future growth markets.

States car data segments:

1. **Segment 1: High subsidy, Low Road tax states:**
    - States offering substantial financial incentives for vehicle purchases, particularly for electric vehicles.
    - High potential market for electric and hybrid vehicles.
2. **Segment 2: High fuel prices, Moderate Subsidy states:**
    - States where fuel prices are high, creating a demand for more fuel-efficient vehicles.
    - Potential for both hybrid and electric vehicles.
3. **Segment 3: Low Subsidy, High Road Tax states:**
    - States with minimal financial incentives and higher taxes, potentially challenging markets.
    - Requires strategic pricing and aggressive marketing**.**

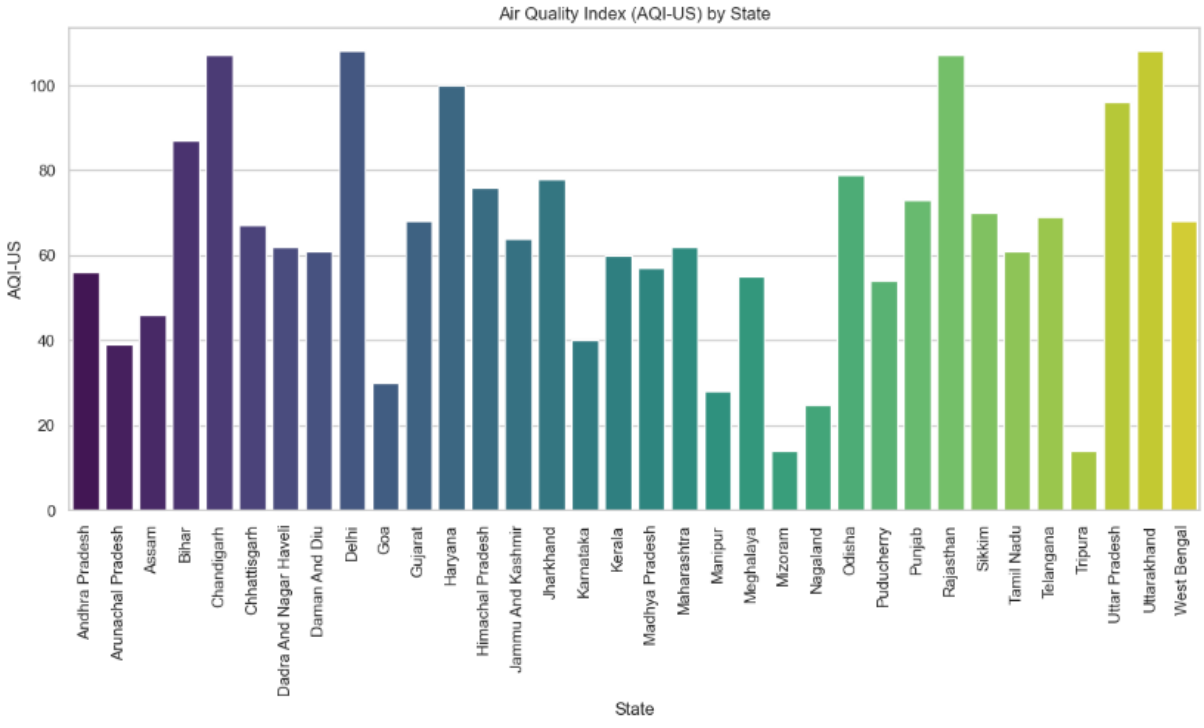
Selection of Target Segments:

1. **High pollution, High population Density regions:** These regions face significant environmental challenges and would benefit most from the introduction of eco friendly vehicles. This segment is highly attractive because of the public health concerns and potential govt policies
2. **High subsidy, Low Road tax states:** Financial incentives and low operational costs make these states ideal for launching electric and hybrid vehicle. Consumers in these states are likely to adopt eco-friendly vehicles more rapidly due to the favourable economic environment.

Visualizations and Insights:

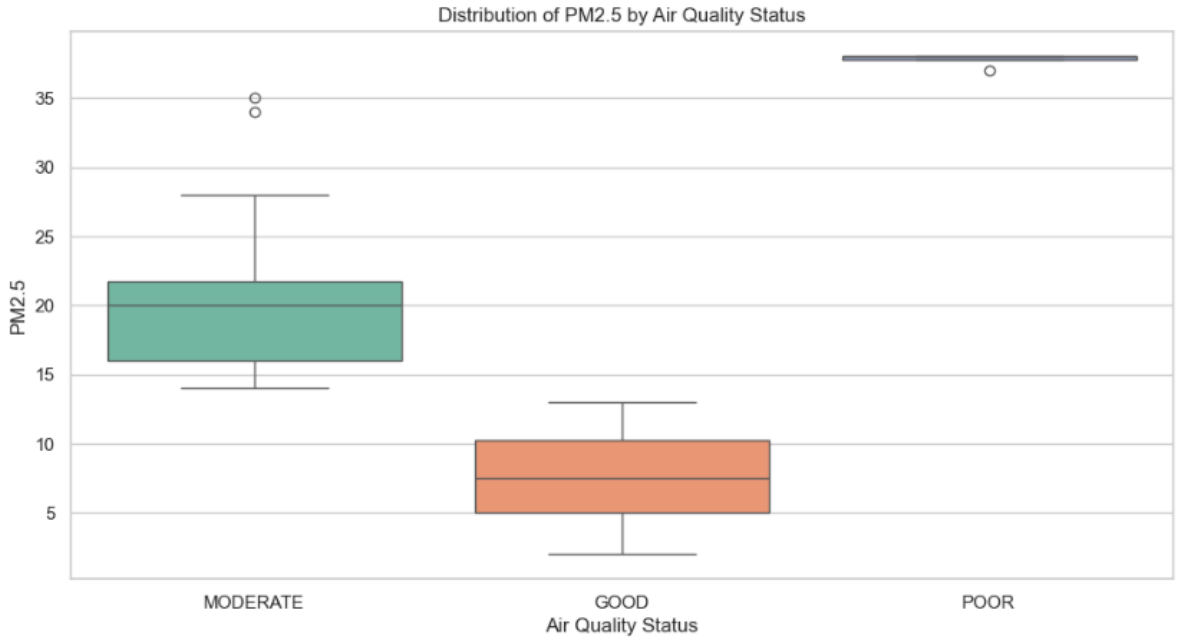
1. **Bar plot of pollution levels across regions:**

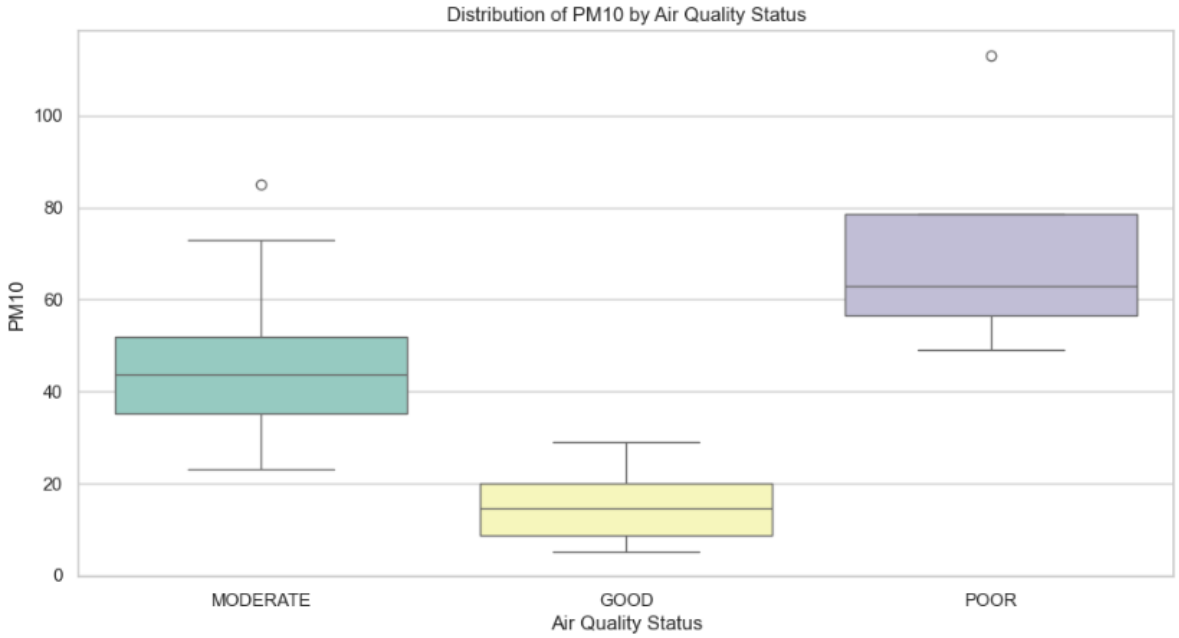
   Visualized the distribution of pollution levels, highlighting hotspots where interventions are most needed.



2. **Box Plot of population Density vs AQI**:

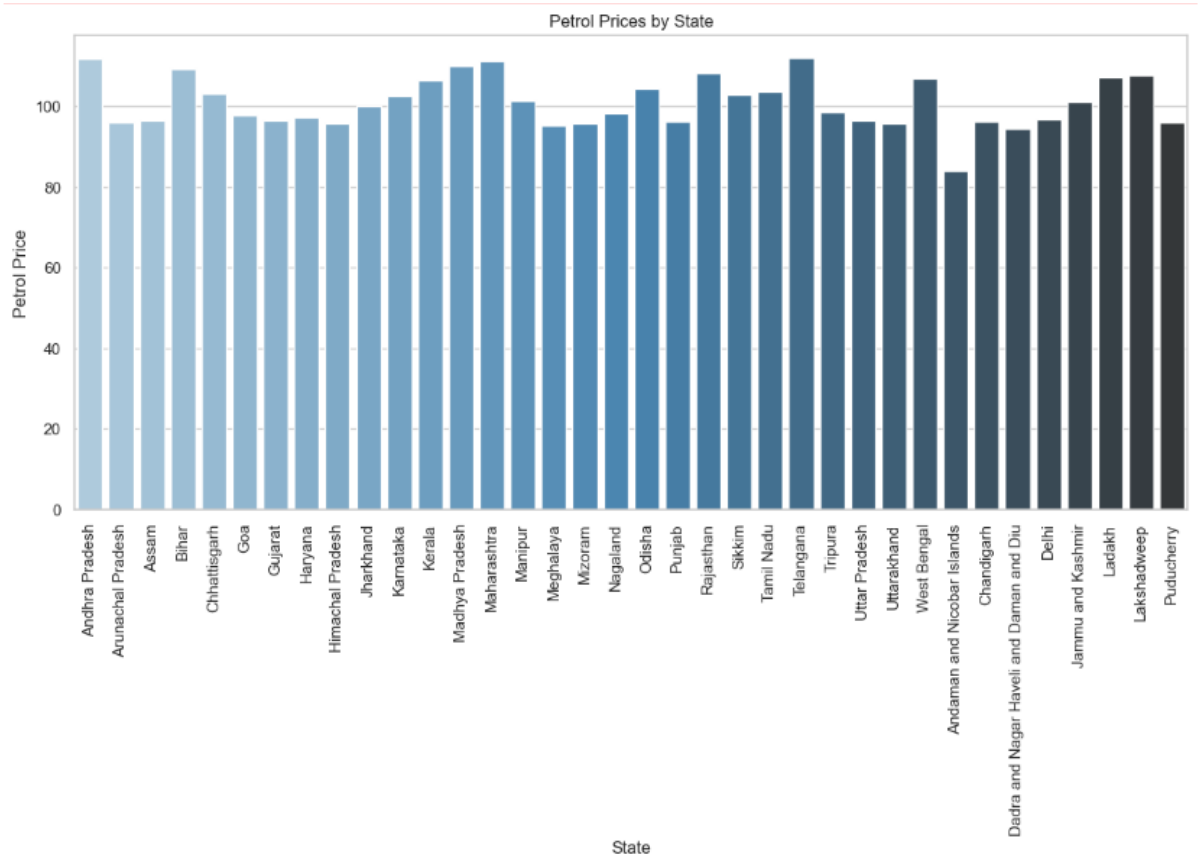
   A Box plot was created with population density on the x axis and AQI on the y axis.

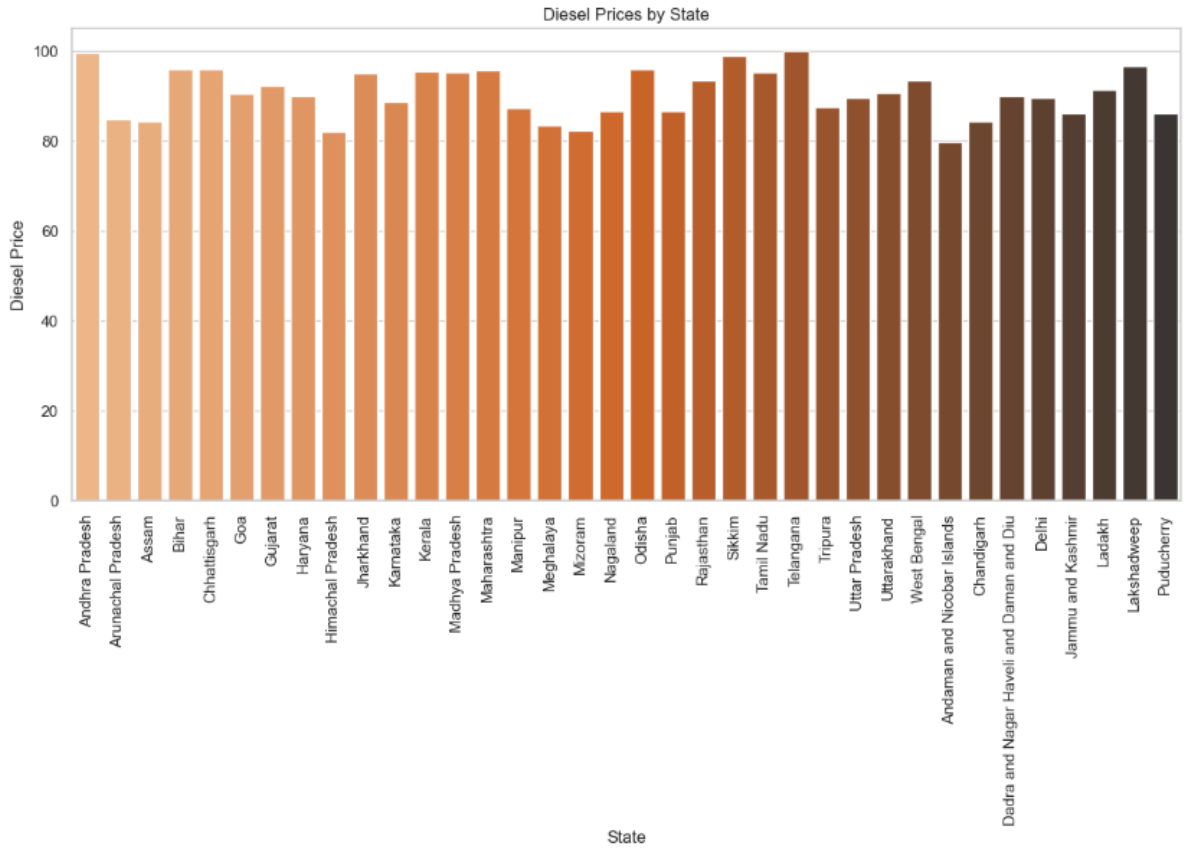




3. **Bar Plot of Petrol and Diesel Prices by State:**

   Bar plots were used to compare petrol and diesel prices across different states. The states were grouped by the clusters identified during segmentation**.**

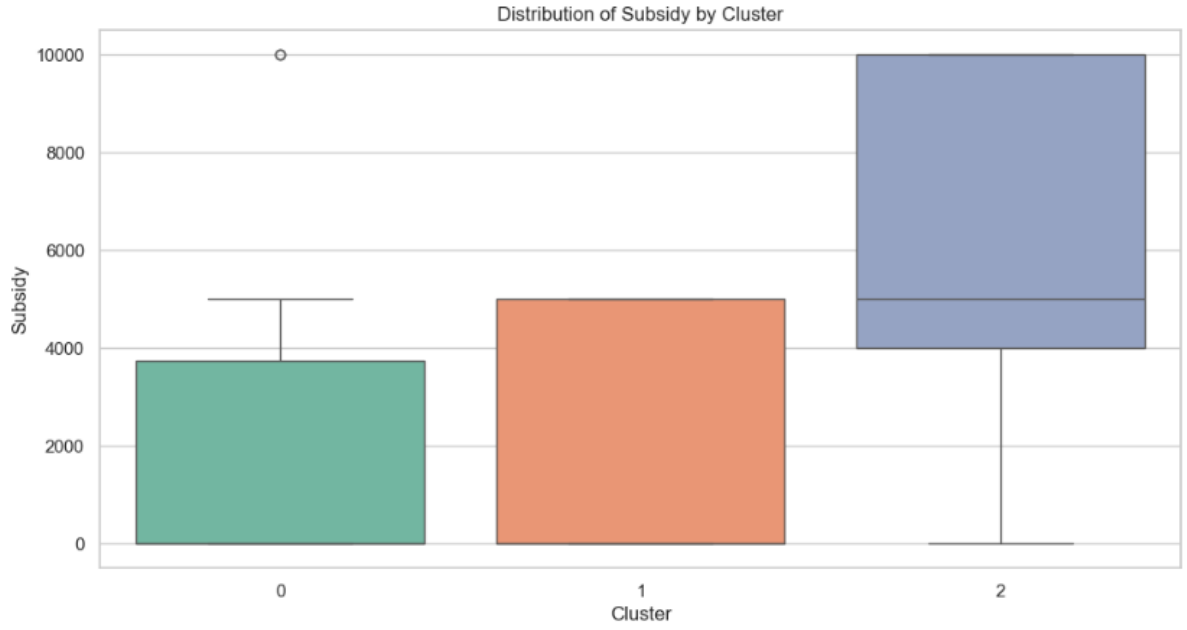


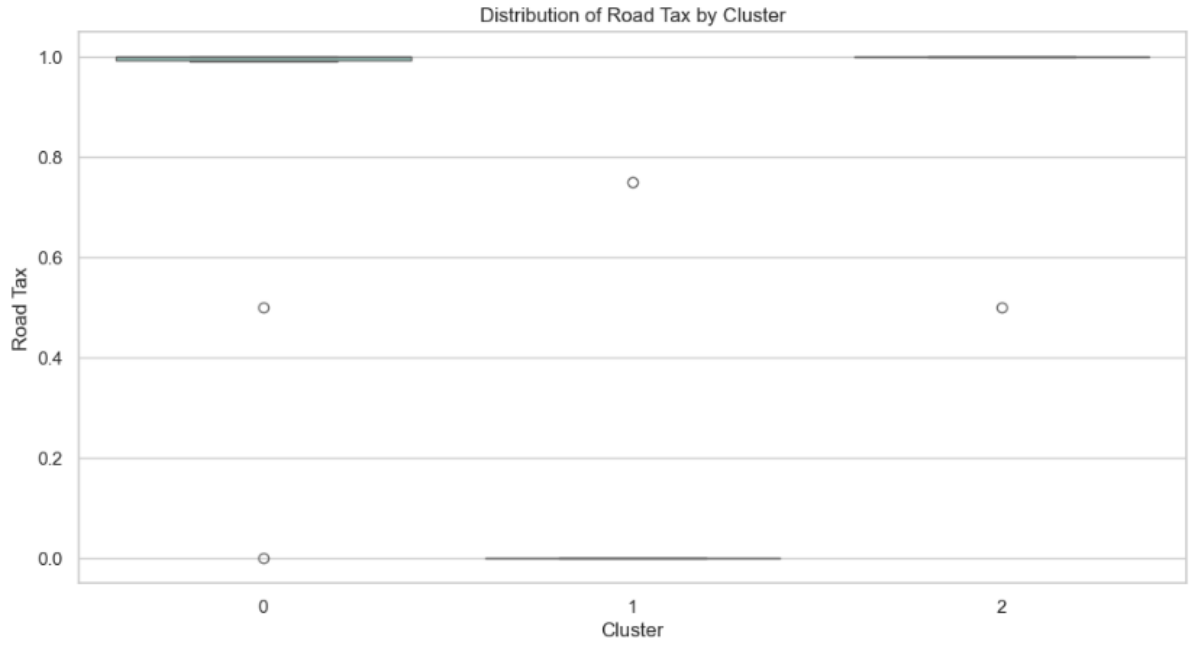


The bar plots highlighted significant disparities in fuel prices across states. States in Cluster 2 (high fuel prices) were clearly distinguished from those in Cluster 1 (lower fuel prices with high subsidies). The high fuel prices in certain states indicated a strong potential market for fuel-efficient and electric vehicles, as consumers in these regions are more likely to seek cost-saving alternatives.

4. **Box Plot of Subsidy and Road Tax by Cluster:**

   A box plot was created to visualize the distribution of subsidies and road taxes across the different clusters. Each cluster's subsidy and road tax data were plotted, showing the range, median, and outliers.

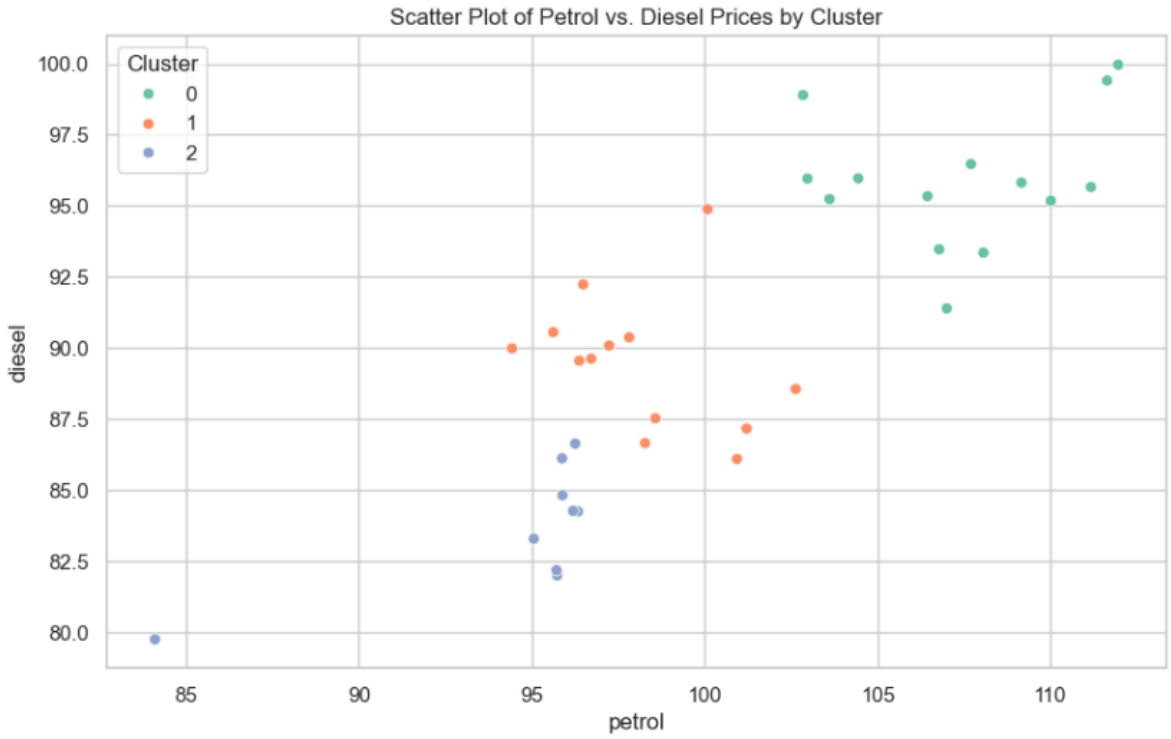




The box plot revealed the financial environment of each cluster. For example, Cluster 1 had higher median subsidies and lower road taxes, making it an attractive market for vehicle manufacturers. The plot also helped identify outliers in the data, such as states with exceptionally high road taxes or unusually low subsidies, which could be outliers or represent unique market conditions.

5. **Scatter Plot of Petrol vs. Diesel Prices by Cluster:**

   A scatter plot was used to compare petrol prices(x-axis) and diesel prices(y-axis) across states, with points color-coded by cluster



   The scatter plot showed a strong correlation between petrol and diesel prices within each cluster, indicating that states generally have similar pricing strategies for both types of fuel. The scatter plot visually differentiated the clusters, showing that Cluster 2 states had consistently higher fuel prices, which aligns with the need for fuel-efficient vehicles in these regions.

Overall Insights:

- **High Risk Areas**: The visualizations helped identify regions with both high pollution and high population density. These areas are prime

candidates for eco-friendly vehicle interventions, where public health concerns and potential government policies could drive demand.

- **Economic Incentives**: Visualizations highlighted the economic landscape across states, showing where subsidies and taxes create favourable or challenging markets for vehicle manufacturers.
- **Strategic Market Entry:** The identification of high-subsidy, low-tax states as the most attractive markets provides a clear strategy for launching new vehicle models, particularly electric and hybrid vehicles**.**

<u>Customizing the Marketing Mix</u>:

Pollution data – Segment 1:

- Product: Focus on electric and hybrid vehicles with low emissions.
- Price: Competitive pricing, leveraging government subsidies.
- Place: Target urban areas in high-pollution regions.
- Promotion: Emphasize health benefits and environmental impact**.**

States Car Data – Segment 1:

- Product: Electric vehicles, supported by strong government incentives.
- Price: Use subsidies to reduce the upfront cost for consumers.
- Place: Focus on states within the high-subsidy, low-tax segment.
- Promotion: Highlight long-term savings and environmental benefits**.**

The most optimal market segments:

- High Pollution, High Population Density Regions: High demand for eco-friendly vehicles.
- High Subsidy, Low Road Tax States: Strong potential for electric and hybrid vehicle adoption.

Potential Customer Base in the Early Market:

- From the previous analysis, let's assume the startup is targeting high-priority regions identified in the segmentation.
- Early adopters typically represent a small percentage of the total market. Let's estimate that 5% of the identified target market is composed of early adopters.
- If the target market size is 1 million potential customers, the early adopter segment would be 50,000 customers.
- Assume the average price of the eco-friendly vehicle is ₹5,00,000.
  Potential Sales = 50,000 customers × ₹5,00,000 = ₹25,000,000,000 (₹25 billion).
- Assume that the combined cost of manufacturing, distribution, and operations is 70% of revenue.
  Total Costs = 70% of ₹25 billion = ₹17.5 billion.
- Potential Profit = Total Revenue - Total Costs
  Potential Profit = ₹25 billion - ₹17.5 billion = ₹7.5 billion.

Summary

- Potential Customer Base in Early Market: 50,000 customers.
- Potential Sales: ₹25 billion.

- Potential Profit: ₹7.5 billion.

GitHub link: https://github.com/vaishnavig06/EV-market-segmentation