

Feyn Labs: Project 2

Link: <https://github.com/saxenasajal03/Fyenn-Labs/tree/main/Project%20>

Market Segmentation Analysis of Electric Vehicles in India

Buying Behavioral Analysis

By: Sajal Saxena



Market Segmentation Report: Indian Automobile Buying Behavior

Project Overview:

This project analyzes consumer buying behavior in the Indian automobile market. We aim to identify distinct customer segments using various demographic, financial, and behavioral features. The segmentation will help in tailoring marketing strategies to meet the specific needs of different customer groups.

Machine Learning Model Used:

For the second project, which is market segmentation, we used the K-Means clustering algorithm. This unsupervised learning algorithm helps group similar data points into clusters based on their features, allowing us to identify distinct customer segments.

Why k-means?

- **Simplicity & Efficiency:** K-Means is easy to implement and computationally efficient for large datasets.
- **Interpretability:** The clusters created by K-Means are easy to interpret and help in grouping customers based on similar behavior or characteristics.

In this project, we used features such as **Age, Salary, Number of Dependents, Total Income, Car Price, Profession, and Marital Status** to create customer segments. K-Means helped uncover distinct patterns in these features, which were used to form clusters representing different customer groups.

Final Conclusion & Insights Gained from the Research:

After applying K-Means clustering with 4 clusters, we gained the following key insights into the Indian automobile buying behavior:

Key Customer Segments:

1. **Cluster 1:**
 - **Age:** Mid-career professionals (Average age ~40)
 - **Salary:** High-income earners (~₹2 million annual salary)
 - **Car Preference:** Mid-range cars (~₹1.34 million price)
 - **Dependents:** 2-3 dependents
2. **Cluster 2:**
 - **Age:** Younger professionals (Average age ~29)
 - **Salary:** Lower-income bracket (~₹1.2 million salary)
 - **Car Preference:** Budget cars (~₹1.08 million price)
 - **Dependents:** Mostly single or no dependents
3. **Cluster 3:**

- **Age:** High-income older individuals (Average age ~43)
- **Salary:** Premium earners (~₹2.9 million salary)
- **Car Preference:** Premium cars (~₹1.77 million price)
- **Dependents:** Larger families (2-3 dependents)

4. Cluster 4:

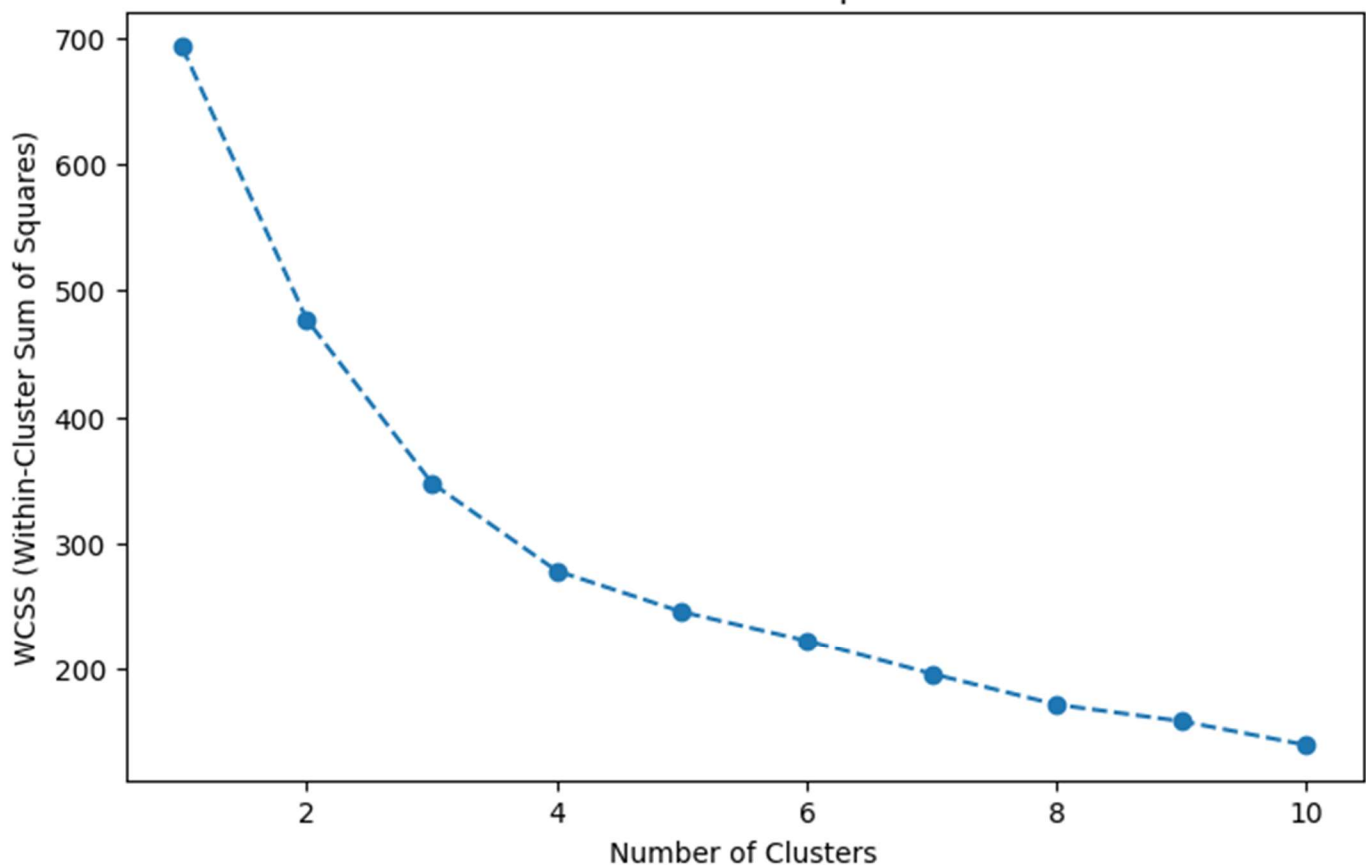
- **Age:** Mid-career professionals (Average age ~34)
- **Salary:** Moderate income (~₹1.34 million salary)
- **Car Preference:** Budget cars (~₹0.87 million price)
- **Dependents:** Moderate family size

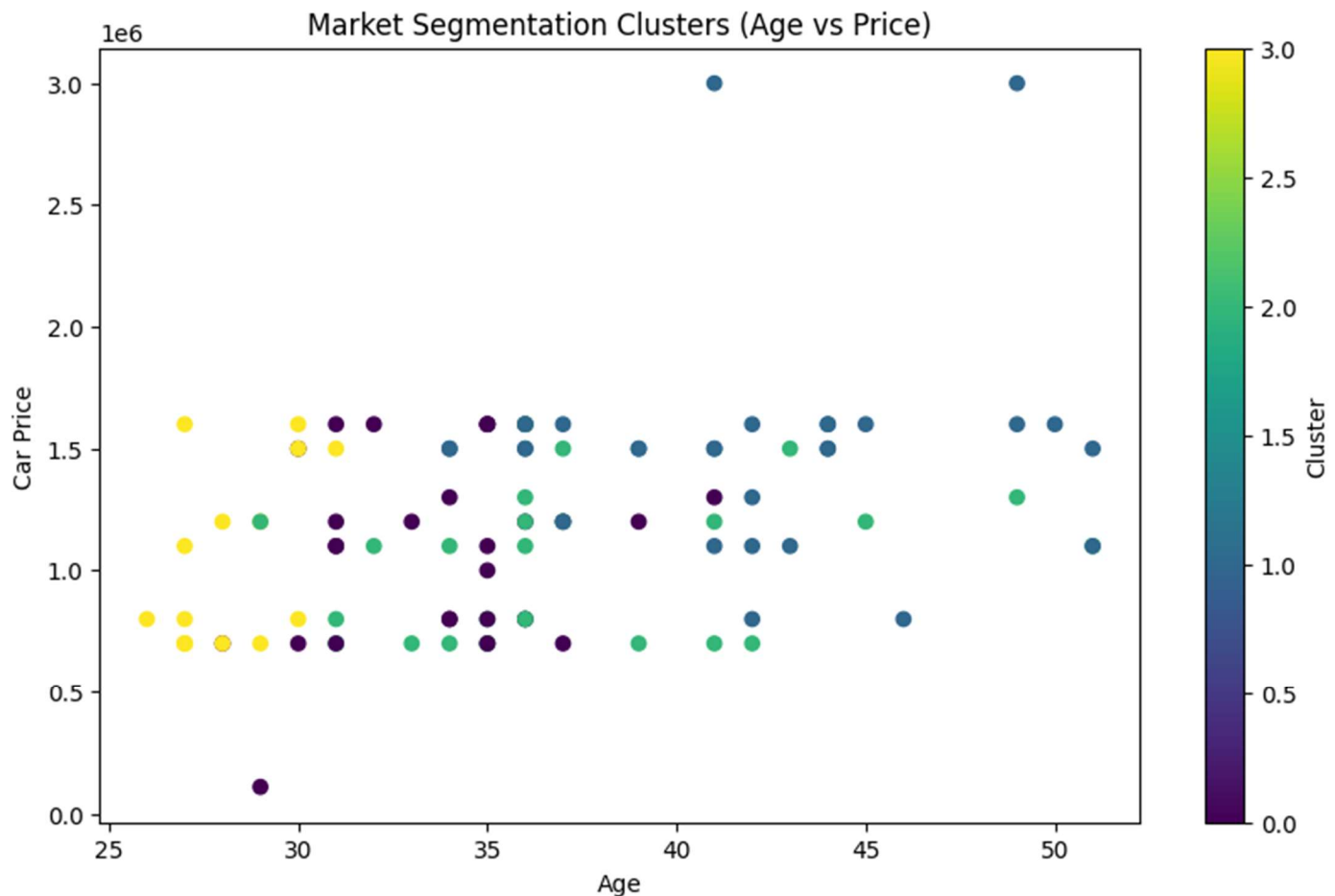
Conclusions:

- Younger professionals with fewer dependents (Cluster 2) tend to prefer budget cars.
- High-income individuals (Cluster 3) prefer premium cars, with larger family sizes and higher financial capacity.
- Mid-career professionals with moderate incomes (Clusters 1 and 4) show varying car preferences depending on family size and total income.

These insights can help marketers develop targeted campaigns for each segment, enhancing customer engagement and optimizing product offerings.

Elbow Method for Optimal K





Improvements to the Market Segmentation Project:

Given more time and budget, the following improvements could be made:

Data Collection:

Additional datasets could enhance the segmentation:

- **Credit Score:** To measure the financial health of customers and their ability to afford cars.
- **Car Usage:** Data on whether cars are used for personal, business, or leisure purposes.
- **Geographical Location:** Urban vs rural, which can influence car preferences.
- **Driving Frequency:** Information on how often customers drive.
- **Purchase Mode:** Whether cars are purchased outright or financed through loans.

Additional Features:

Additional columns could be included for a more detailed analysis:

- **Credit Score**
- **Car Usage Frequency**
- **Customer Location**
- **Loan Type and Duration**

Advanced ML Models:

- **Gaussian Mixture Models (GMM):** Can be used for overlapping clusters, where customers might belong to multiple segments.

- **DBSCAN:** Detects outliers and noise, which could identify niche customer segments or unique behaviors.

These improvements would yield more granular and actionable insights.

Estimated Market Size:

The estimated size of the Indian automobile market is approximately **\$100 billion** as of 2022. This includes both premium and budget car segments. The market is growing rapidly, especially in urban areas where car ownership is rising.

To further refine the market size, detailed sales data by segment would be necessary, as well as macroeconomic data such as disposable income and urbanization rates.

Top Variables for Optimal Market Segmentation:

Based on the analysis, the top 4 variables for creating the most optimal market segments are:

1. **Age:** Segments customers based on life stages (e.g., young professionals vs. older family-oriented buyers).
2. **Salary:** Helps define segments based on financial capacity and affordability.
3. **Car Price:** Indicates the type of car preferred (budget vs. premium).
4. **Number of Dependents:** Captures family size, which affects the need for larger, more expensive vehicles.

These variables, combined with other features such as profession and marital status, help identify key patterns in consumer behavior, facilitating better market segmentation.

Conclusion:

The market segmentation analysis successfully identified four distinct customer segments in the Indian automobile market. By using K-Means clustering, we were able to categorize customers based on key features such as age, income, car price, and family size. These insights are invaluable for tailoring marketing strategies and optimizing product offerings. Future work could involve adding more features and experimenting with advanced clustering techniques for more detailed segmentation.