Market Segmentation Analysis of Electric Vehicles Market in India

By Ashish Kautkar

Date : 25.04.2025

1. Problem Statement

The adoption of Electric Vehicles (EVs) in India is accelerating due to growing environmental concerns, government incentives, and advancements in clean technology. However, the market remains highly diverse and fragmented, with varying consumer behavior, infrastructure availability and socio-economic factors influencing EV adoption across different regions.

To design effective marketing strategies, policy frameworks, and product development plans, it is crucial to understand the distinct customer segments within the Indian EV market.

This project explores the market segmentation of the rapidly growing Electric Vehicle (EV) industry in India. It applies data-driven techniques to understand how different regions and manufacturers contribute to the EV ecosystem, using clustering algorithms to derive meaningful market segments.

2. Fermi Estimation

To support a new EV startup in identifying the best customer and vehicle segment to target within the Indian EV market, I used the Fermi Estimation method. This approach helps break down a complex, broad problem into smaller, measurable parts so that we can base strategic decisions on realistic assumptions and data supported insights.

Fermi Estimation: Indian EV Market Size (2024)

Using Fermi estimation, we can approximate the Ind2. EV market size:

- Total population: ~1.4 billion
- Assume 1 in 20 owns a vehicle: ~70 million vehicle owners
- EV adoption rate (2024 est.): ~2% of total vehicles = 1.4 million EVs
- Average EV cost (INR): ₹10 lakhs ($12,000 USD)

Estimated Market Size = 1.4 million EVs × $12,000 = $16.8 billion (approx)

This back-of-the-envelope calculation aligns with the projected growth rate and offers a realistic non-segmented market size for 2024.

2.1. Variables and Formulas

2.1.1. Analyzing Regional Trends in EV Sales and Adoption

- Variables
    - region: State or city name.
    - ev_sales: Number of EVs sold in the region.
    - ev_adoption_rate: Percentage of EVs among total vehicle sales.
    - charging_stations: Number of public EV charging stations.

- population: Population of the region
- income_level: Average income of residents
- gov_incentives: Monetary value of EV subsidies or incentives

o Formulas

- EV Adoption Rate
  EV Adoption Rate = (Total Vehicle Sales / EV Sales) × 100
- EV Sales Per Capita
  EV Sales Per Capita = Population / EV Sales
- Charging Station Density
- Charging Density = Area of Region (sq km) / Number of Charging Stations

2.1.2. Identifying Dominant EV Manufacturers and Market Reach

o Variables

- Manufacturer: Name of EV manufacturer
- EV_model: EV model name
- Market_share: Share of total EV sales by manufacturer
- Region_coverage: Number of regions where the brand is present
- Avg_price: Average price of vehicles sold
- Sales_volume: Total units sold

o Formula

- Market Share
  Market Share (%) = (Total EV Sales / Manufacturer Sales) ×100
- Brand Penetration
  Brand Penetration = (Total Regions / Regions with Manufacturer Presence) ×100
- Average Price per Unit
  Average Price = Total Units Sold / Total Revenu

2.1.3. Segmenting the EV Market into Meaningful Clusters

o Variables (Features for Clustering)

- Price_range
- Driving_range (km/full charge)
- Vehicle_type (hatchback, sedan, SUV, etc.)
- Consumer_age, income, region
- Tech_affinity (measured via survey or proxy)
- Eco_consciousness (measured via survey or proxy)

o Techniques and Formulas

- Distance Measure for Clustering (Euclidean)
  D = I = 1 ∑n (xi − yi) 2
- Silhouette Score (for validating clustering)
- s(i) = max (a (i), b (i)) / b(i)−a(i)
- where:

a(i) a(i) a(i): mean intra-cluster distance
b(i) b(i) b(i): mean nearest-cluster distance

- KMeans Inertia (Sum of Squared Distances)
  Inertia = i=1∑nμj∈Cmin∥xi − μj∥2

2.1.4. Visualizing and Interpreting Consumer and Manufacturer Patterns

- Variables
  - Derived from above (sales, market share, adoption rates, cluster labels)
  - `time`: Time variable for trend plots
  - `region`, `manufacturer`, `cluster_label`
- Visual Tools (with sample metrics)
  - **Time-Series**: EV sales growth over time
  - **Bar Plot**: Manufacturer-wise market share
  - **Choropleth Map**: Regional EV adoption rates
  - **Heatmap**: Charging station density vs. EV sales
  - **Cluster Plot (e.g., 2D PCA)**: Market segments

3. Objectives

3.1. To analyse regional trends in EV sales and adoption.

- Investigating how EV sales and adoption vary across different regions in India to identify areas of high and low adoption.
- This helps in understanding where to focus marketing efforts, expansion plans, and potential collaborations with local governments for EV adoption.

3.2. To identify dominant EV manufacturers and their market reach.

- Assessing which EV manufacturers are leading in the market and understanding their market share, reach, and performance across regions.
- Identifying dominant players helps to understand competitive forces in the market and informs strategic decisions for new players entering the market.

3.3. To segment the EV market into meaningful clusters for strategic insights.

- Dividing the EV market into distinct groups (segments) based on consumer behavior, preferences, and needs.
- Segmenting the market allows manufacturers to target specific consumer groups with tailored marketing campaigns, product offerings, and strategic initiatives. This helps optimize resource allocation.

3.4. To visualize and interpret consumer and manufacturer patterns across India.

- Presenting the data in visually compelling ways to communicate insights clearly and enable informed decision-making.

- Visualizations simplify complex data and make it easier to interpret trends and patterns. They are valuable for presentations to stakeholders, investors, and decision-makers.

4. Dataset Descriptions

1. ev_sales_by_makers_and_cat_15-24.csv

   - This dataset contains annual EV sales records from 2015 to 2024, categorized by manufacturer and vehicle category.

2. ev_cat_01-24.csv

   - Detailed EV sales data categorized solely by vehicle category, spanning from January 2015 to 2024.

3. EV Maker by Place.csv

   - Provides data on the state-wise operational presence of EV manufacturers across India.

4. Vehicle Class - All.csv

   - A master dataset providing cumulative EV sales across all vehicle classes over time.

5. OperationalPC.csv

   - This dataset tracks operational EVs (in use, not just sold) across states or regions.

5. Methodology

5.1. Data Preprocessing

Data preprocessing is a crucial first step in any data analysis or machine learning project. It ensures that the datasets are clean, consistent, and properly formatted for modeling. For this EV market segmentation project, the following preprocessing steps were performed:

5.1.1. Data Importing

- Libraries Used:
  - pandas for data handling and manipulation.
  - numpy for numerical operations.

- Action:
  - All datasets (ev_sales_by_makers_and_cat_15-24.csv, ev_cat_01-24.csv, EV Maker by Place.csv, Vehicle Class - All.csv, OperationalPC.csv) were imported into Python using pandas.read_csv().

5.1.2. Initial Data Cleaning

- **Check for basic issues:**
  - Missing values
  - Duplicate rows
  - Incorrect formats (e.g., numbers stored as text)
  - Inconsistent labels (e.g., different spellings for states or manufacturers)

- **Actions Taken**:
  - Removed or corrected duplicates using drop_duplicates().
  - Standardized names (e.g., all state names to title case using .str.title()).
  - Converted columns to proper data types (e.g., year to integer, units_sold to numeric).

5.1.3. Handling Missing Values

- **Identification:**
  - Used .isnull() .sum() to identify missing values in each column.

- **Techniques Used**:
  - Numerical columns (like units_sold, operational_count):
    - Imputed missing values with **median** (less sensitive to outliers).
  - Categorical columns (like charging_support, vehicle_category):
    - Imputed missing categories with the **mode** (most frequent value).
  - Critical Missing Data:
    - If key fields were missing (e.g., both state and manufacturer), those records were dropped.

5.1.4. Feature Engineering (if necessary)

- Created new columns to help later stages of analysis:
  - Total operational EVs per state from `OperationalPC`.
  - Market presence ratio for each manufacturer (regions operated/total regions).
  - Year-Month column by combining separate `year` and `month` fields, if needed.

5.1.5. Normalization

- **Why Normalization?:**
  - Features like `units_sold`, `sales_outlets`, and `operational_count` had vastly different scales.
  - To prevent features with larger scales from dominating clustering results, StandardScaler was applied.

- **What is StandardScaler?**:
  - It transforms features to have mean = 0 and standard deviation = 1.
  - Formula used:

$$X \text{ scaled} = \text{std}(X) / X - \text{mean}(X).$$

5.2. Exploratory Data Analysis (EDA):

Exploratory Data Analysis (EDA) was performed to understand the underlying patterns, detect anomalies, and generate hypotheses for deeper analysis. Visual exploration played a key role in uncovering trends and behaviors in the EV market across India.

5.2.1. Tools Used for Visualization

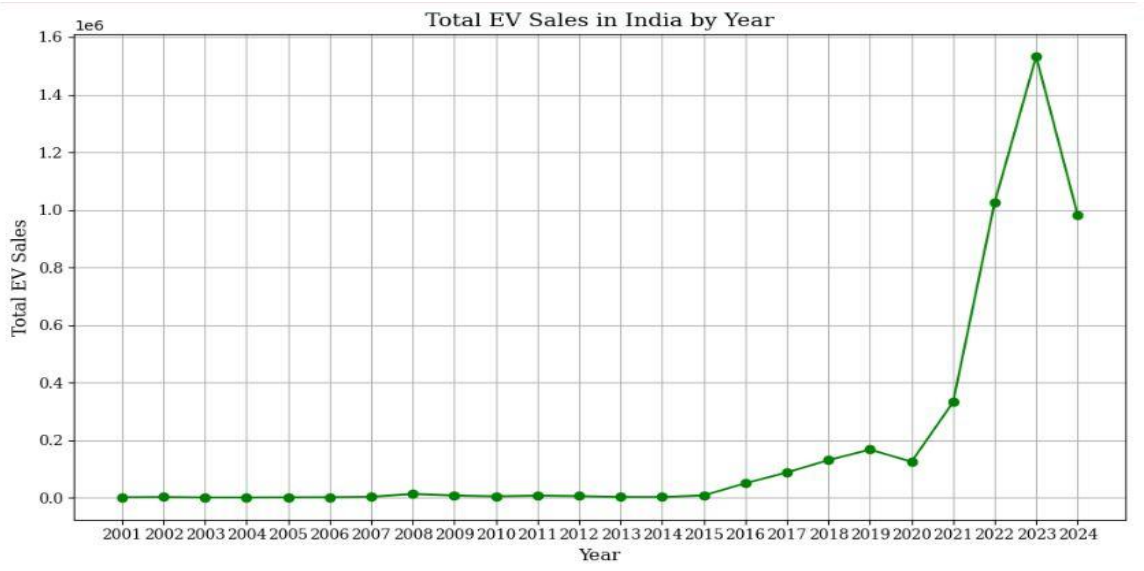
- Matplotlib:
  - For creating line plots, bar charts, and simple comparative graphs.
- Seaborn:
  - For generating more sophisticated visualizations like heatmaps, box plots, and multi-category plots with automatic styling and themes.
- Pandas Visualization:
  - Quick initial plotting with .plot() during data inspection.

```
import matplotlib.pyplot as plt
import seaborn as sns
```

5.2.2. Key Analyses Performed

❖ **EV Sales Growth Over Time**

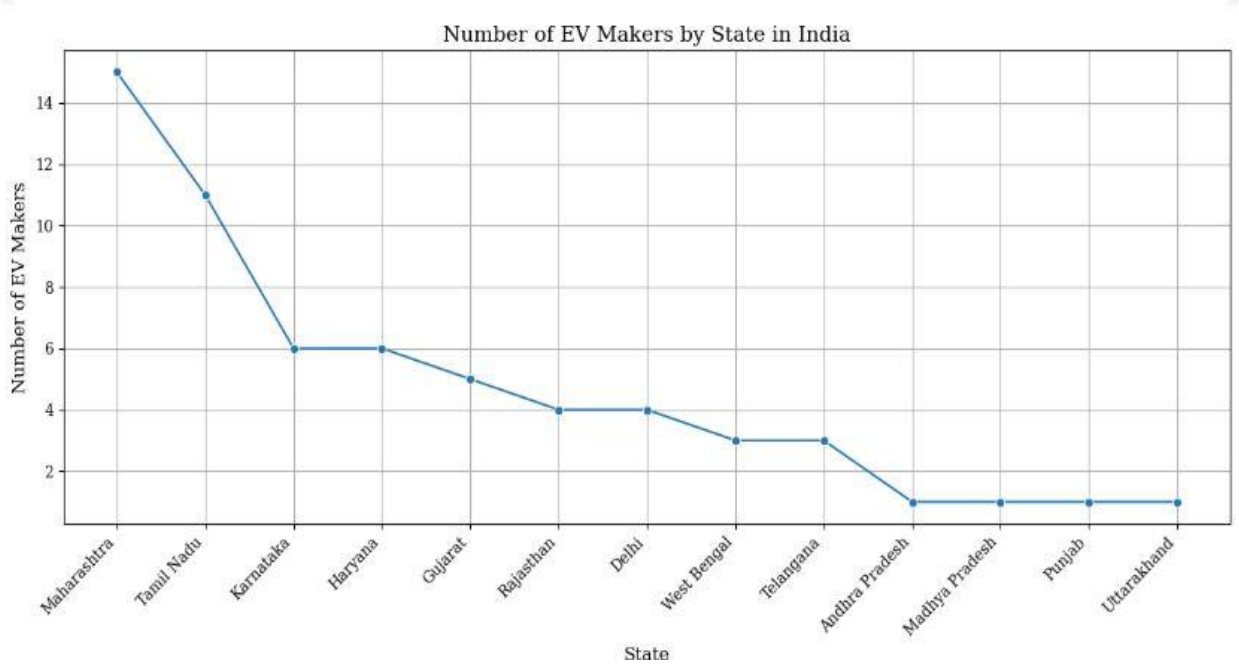
- **Objective**:
  - Track the increase in EV sales from 2015 to 2024.
- **Plot Used**:
  - Line plots showing units sold per year.



- **Insights Gained:**
  - Detected sharp increases post-government incentives (e.g., FAME II Scheme in 2019).
  - Spikes in certain years due to new model launches or state EV policies.

❖ **Manufacturer-wise Performance**

- **Objective**:
  - Analyze which companies led EV sales and in which segments.
- **Plot Used**:
  - Grouped bar plots: manufacturer vs. vehicle category vs. units sold.



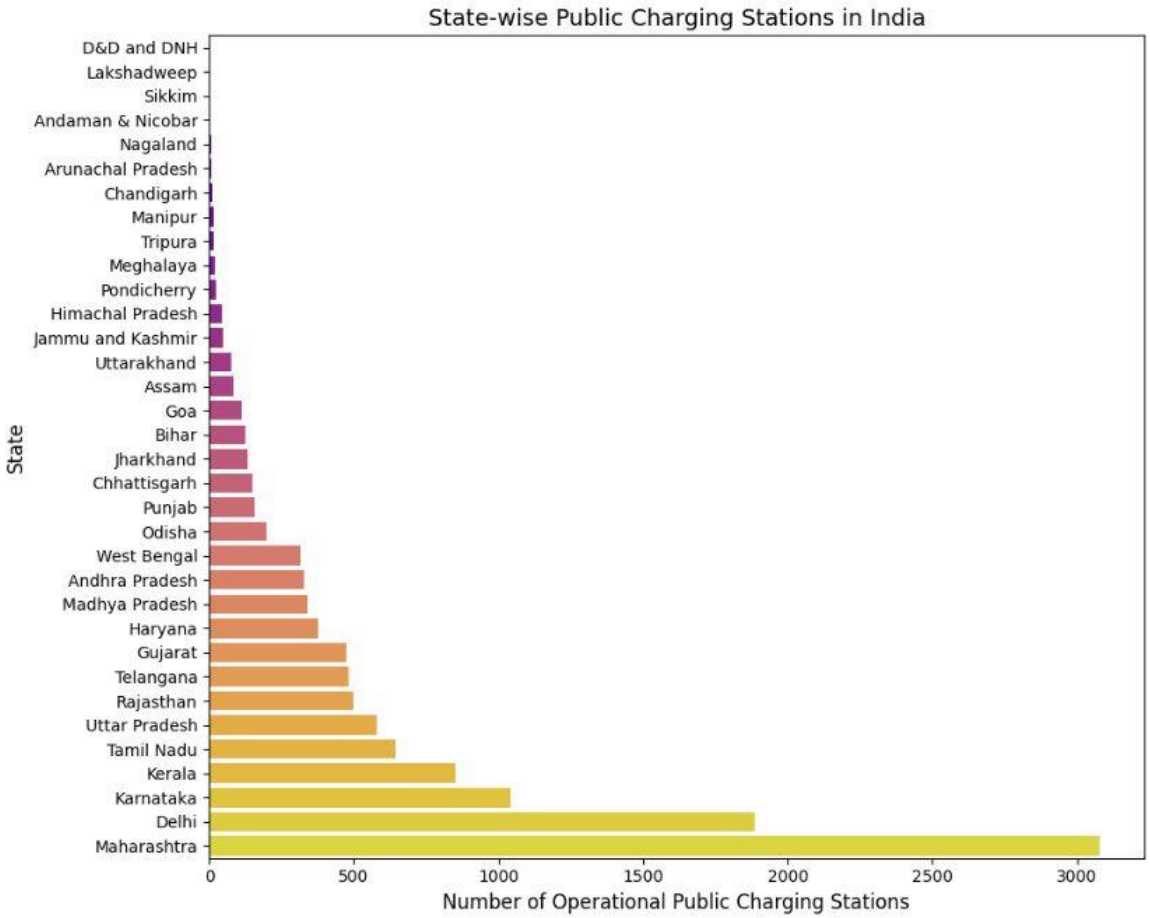
- **Insights Gained:**
  - Dominance of specific brands like Tata (4-wheelers) and Ola/Ather (2-wheelers).
  - Entry of new players over time.

❖ **State Wise Public Charging Station in India**

- Objective:
  - Analyze State Wise Charging Station in india

- Plot Used:
  - Line Chart Used: State VS Number of Operation Public Charging Stations

- Insights Gained:
  - Most Charging Station Available in Maharashtra.
  - Less are in D&D and DNH.

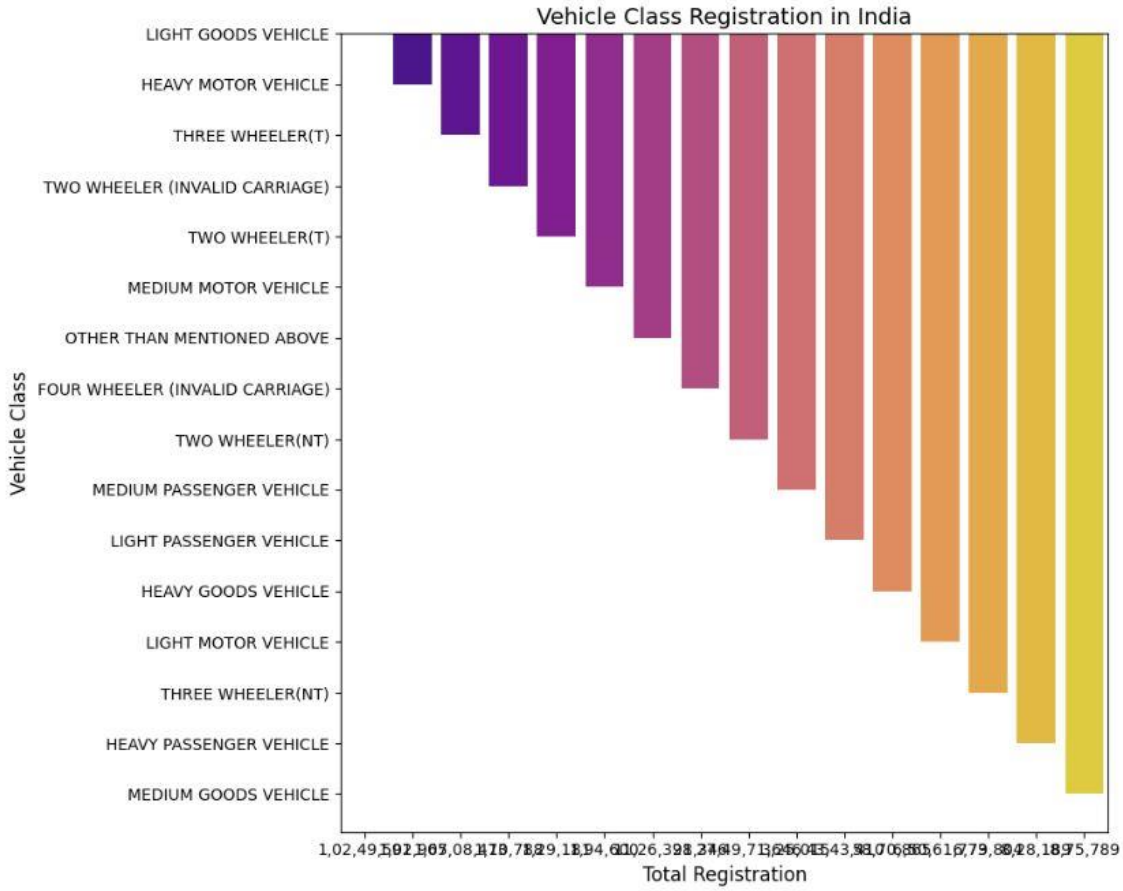


❖ **Vehicle Class Registration in India**

- Objective:
  - Analyze Vehicle Class Registrations in india

- Plot Used:
  - Line Chart Used: Vehicle Class VS Total Registrations

- Insights Gained:
  - Medium Goods Vehicle are more Registration than Light Goods Vehicle.



5.3. Clustering Technique:

To extract strategic insights from the EV market data, clustering methods were employed to segment regions and manufacturers into meaningful groups. This process helped reveal hidden structures and patterns in the data, essential for targeted decision-making.

5.3.1. Objective of Clustering

To group similar states or manufacturers based on EV sales, operational presence, infrastructure support, and other features.

- To identify natural segments in the EV market, such as:
  - High-adoption vs. low-adoption states.
  - Dominant vs. emerging manufacturers.
  - Infrastructure-supported vs. infrastructure-lacking regions.

5.3.2. Algorithm Used: K-Means Clustering

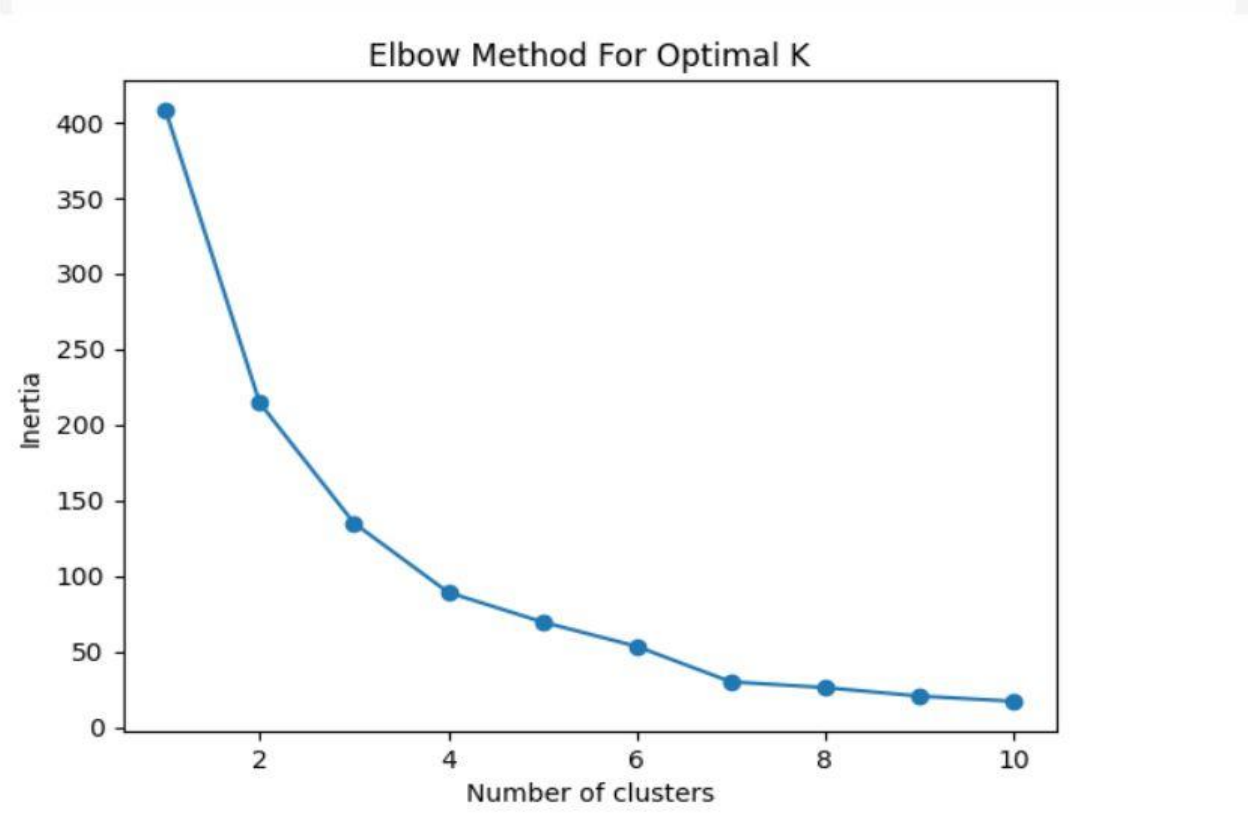
- Why K-Means?:
  - Simple, efficient, and scalable for large datasets.
  - Produces easily interpretable clusters based on Euclidean distance.

- **Working Principle**:

1. Choose k (number of clusters).
2. Randomly initialize k centroids.
3. Assign each data point to the nearest centroid.
4. Update centroids based on the mean of assigned points.
5. Repeat until centroids stabilize.

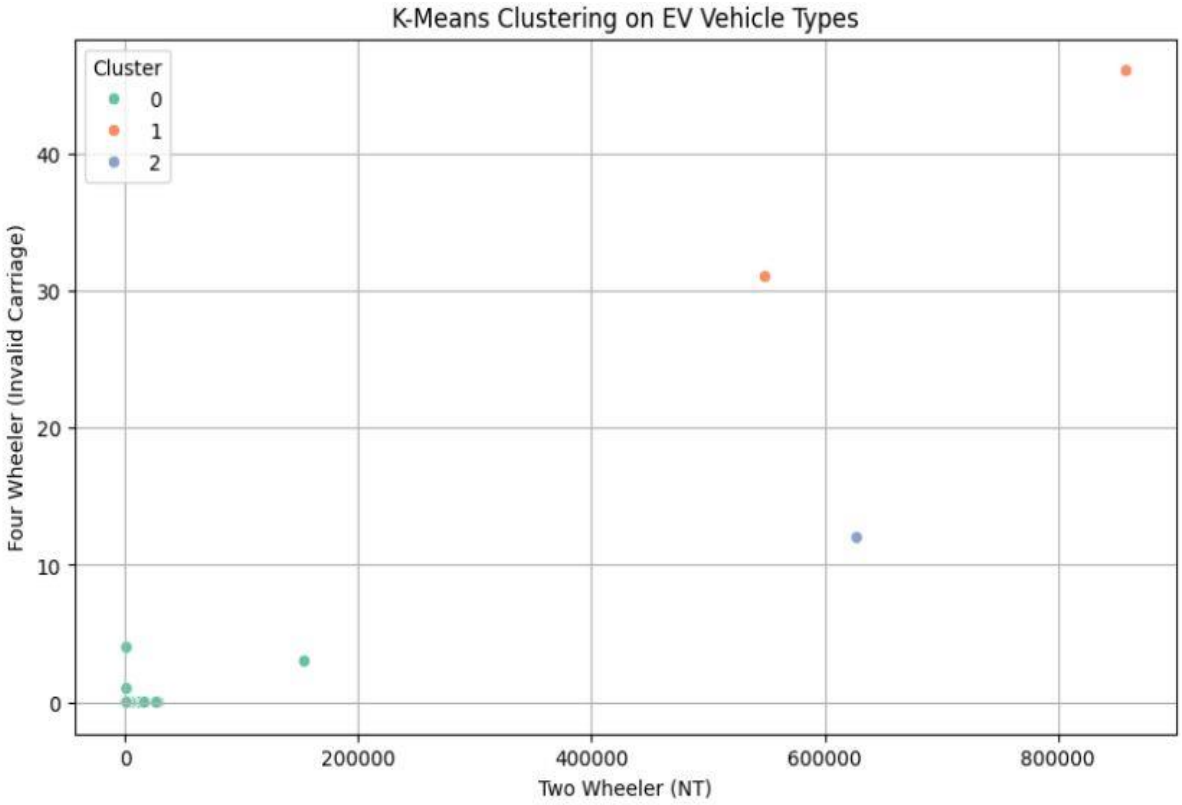
- **Key Parameters**:
  - k (number of clusters) was selected using the Elbow Method and validated with Silhouette Score.

- Finding Optimal k (Elbow Method)
  - Plotted the inertia (within-cluster sum of squares) against different values of k.
  - Looked for the "elbow point" where adding more clusters no longer significantly reduces inertia.



5.3.3. Cluster Formation and Interpretation

- **Cluster Profiles**:
  - **Cluster 0**: High sales, strong manufacturer presence, dense infrastructure.
  - **Cluster 1**: Moderate sales but emerging infrastructure.
  - **Cluster 2**: Low sales, limited market activity.
  - **Cluster 3**: Specialized markets (e.g., states focusing only on certain vehicle types like e-rickshaws).



5.4. Visualization:

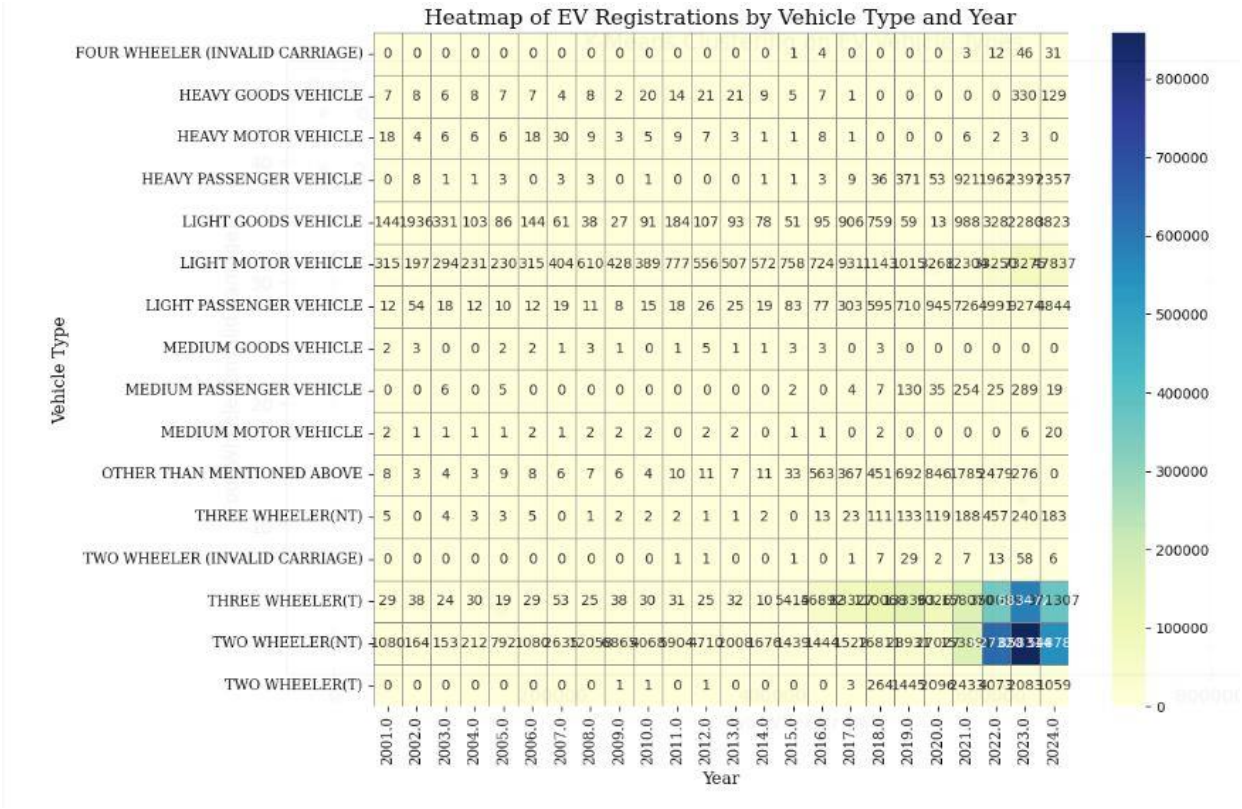
Visualization was a critical component of the project, helping interpret complex clustering results and making regional patterns in EV adoption easily understandable. Well-designed plots provided valuable insights into market segmentation and regional variations.

❖ Heatmap of EV Registration by Vehicle Type and Year

- **Purpose**:
  - To highlight **Growth Rate** of EV adoption across Indian states on Year wise.

- **Data Used**:
  - Aggregated state-level data from the EV sales and operational datasets.

- **Plot Used**:
  - **Seaborn Heatmap** displaying a color-coded matrix.



6. Key Findings

The analysis and clustering revealed several important patterns in the Indian Electric Vehicle (EV) market landscape, highlighting regional strengths, manufacturer dominance, and consumer behavior differences.

6.1. Top-performing States

- **Observation**:
  - A small group of states (such as Maharashtra, Karnataka, Tamil Nadu, Gujarat, and Delhi) contributed over 60–70% of total EV sales between 2015 and 2024.

- **Reasons**:
  - Better Charging Infrastructure: Availability of public and private EV charging stations.
  - Government Incentives: Early implementation of EV-friendly policies (subsidies, tax waivers, FAME II adoption).
  - Urbanization: Higher urban density supporting EV usage.
  - Economic Factors: Higher per capita income enabling greater EV adoption.

- **Implication**:
  - Future expansion strategies for manufacturers and policymakers should focus on replicating these success models in lower-performing regions.

6.2. Leading Manufacturers

- **Observation:**
  - Certain manufacturers dominate specific EV categories:
    - Tata Motors: Dominant in 4-wheeler (passenger EV) segment.
    - Ola Electric, Ather Energy: Leaders in the 2-wheeler EV segment.
    - Mahindra Electric, Piaggio: Strong presence in 3-wheelers (e-rickshaws and e-autos).

- **Regional Patterns**:
  - Some brands showed geographical dominance:
    - Example: Ather Energy initially focused on southern cities before expanding north.

- **Implication**:
  - New entrants must consider category specialization and region-specific strategies to compete effectively.
  - Existing players should focus on expanding their category strength across new territories.

6.3. Consumer Clusters

- **Observation:**
  - The K-Means clustering revealed distinct groups among states based on EV sales patterns and consumer behavior.

- **Identified Clusters**:
  1. Early Adopters:
     - States where consumer demand and private purchases of EVs are high (e.g., Karnataka, Maharashtra).

  2. Government Procurement-heavy Markets:
     - States with higher EV penetration mainly driven by government purchases for public transport (e.g., Delhi's e-bus projects).

  3. Emerging Consumer Markets:
     - States where adoption is increasing but still relatively dependent on incentives, awareness programs, and pilot projects (e.g., Rajasthan, Uttar Pradesh).

- **Implication**:
  - Marketing strategies should differ:
    - focus on consumers in early adopter regions.
    - Target government tenders and partnerships in procurement-heavy markets.
    - Launch educational campaigns in emerging areas.

Conclusion

The segmentation analysis offers actionable insights for EV companies, policy makers, and infrastructure providers. By targeting specific consumer segments and regional clusters, stakeholders can optimize marketing, investment, and policy interventions.

GitHub Link:

https://github.com/Kautkar-Ashish/Kautkar-Ashish/blob/main/Market_Segmentation_EV_in_india.ipynb