

## 1.

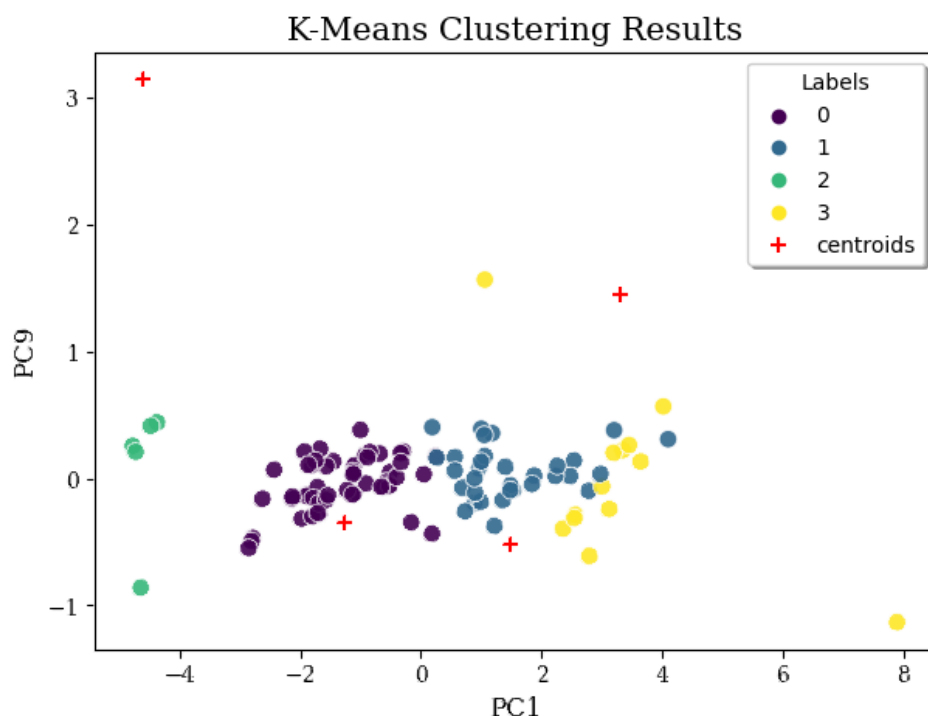
### Clustering

Clustering is a fundamental technique in unsupervised machine learning, where the objective is to group similar data points into clusters based on certain features or characteristics. Several algorithms can perform clustering, including k-means clustering, hierarchical clustering, and density-based clustering.

### K-Means Clustering

K-Means is a popular unsupervised learning algorithm used for clustering data points into K clusters. The algorithm works as follows:

1. Specify the number of clusters, K.
2. Initialize K random points in the dataset as centroids.
3. Assign each data point to the nearest centroid based on Euclidean distance.
4. Recalculate the centroids as the mean of the data points assigned to each cluster.
5. Repeat steps 3 and 4 until convergence, i.e., until centroids do not change significantly or a maximum number of iterations is reached.



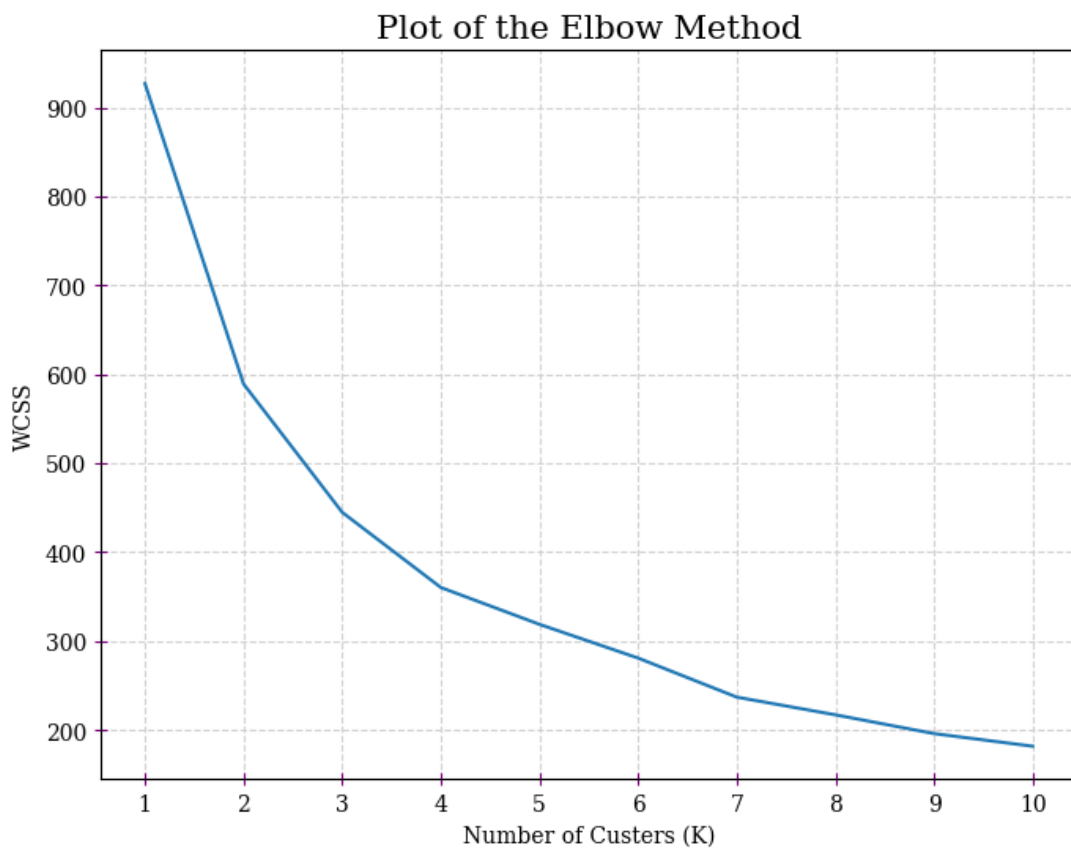
K-Means finds applications in various domains such as market segmentation, image segmentation, image compression, and document clustering.

## Principal Component Analysis (PCA)

PCA is a linear dimensionality reduction technique used to reduce the dimensionality of large datasets while preserving most of the information present in the original data. PCA achieves dimensionality reduction by projecting the data onto a lower-dimensional subspace defined by the principal components.

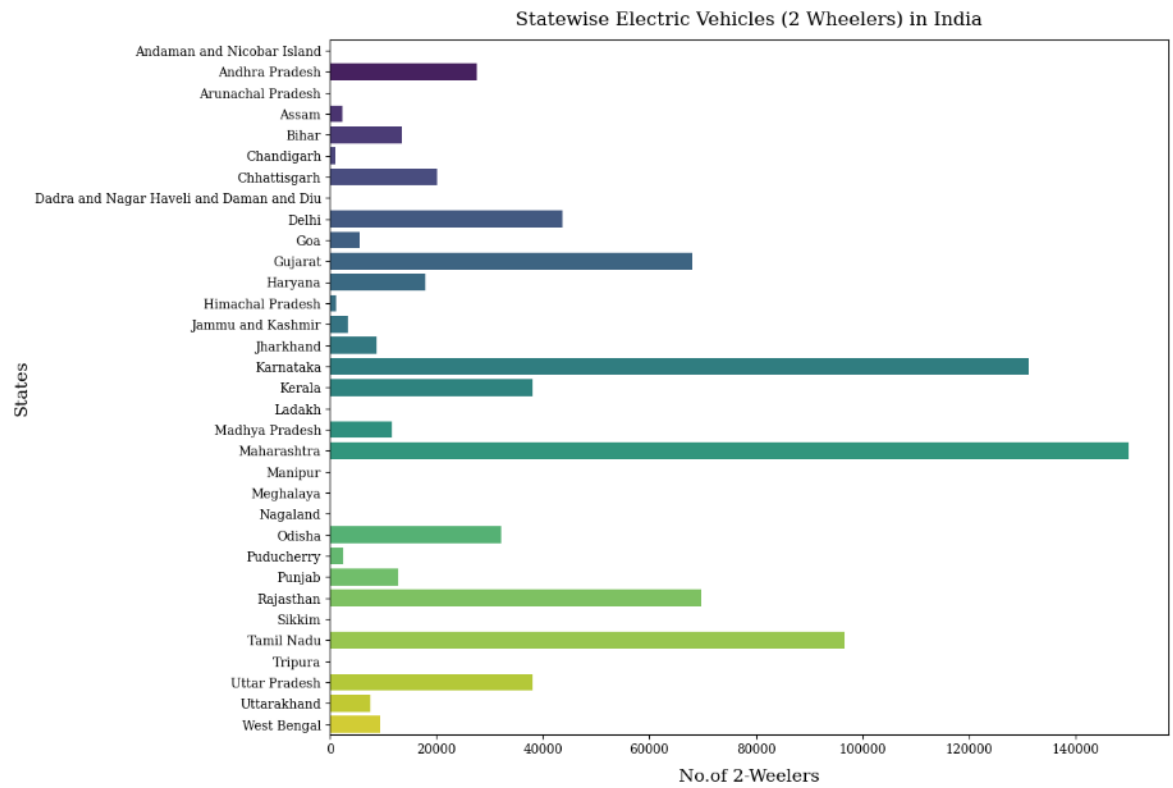
## Elbow Method

The Elbow method is a heuristic used to determine the optimal number of clusters (K) in K-Means clustering. It involves calculating the Within Cluster Sum of Squared Errors (WCSS) for different values of K and selecting the value of K at which the rate of decrease of WCSS starts to diminish significantly. The point at which the rate of decrease sharply decreases resembles an elbow in the plot, hence the name.

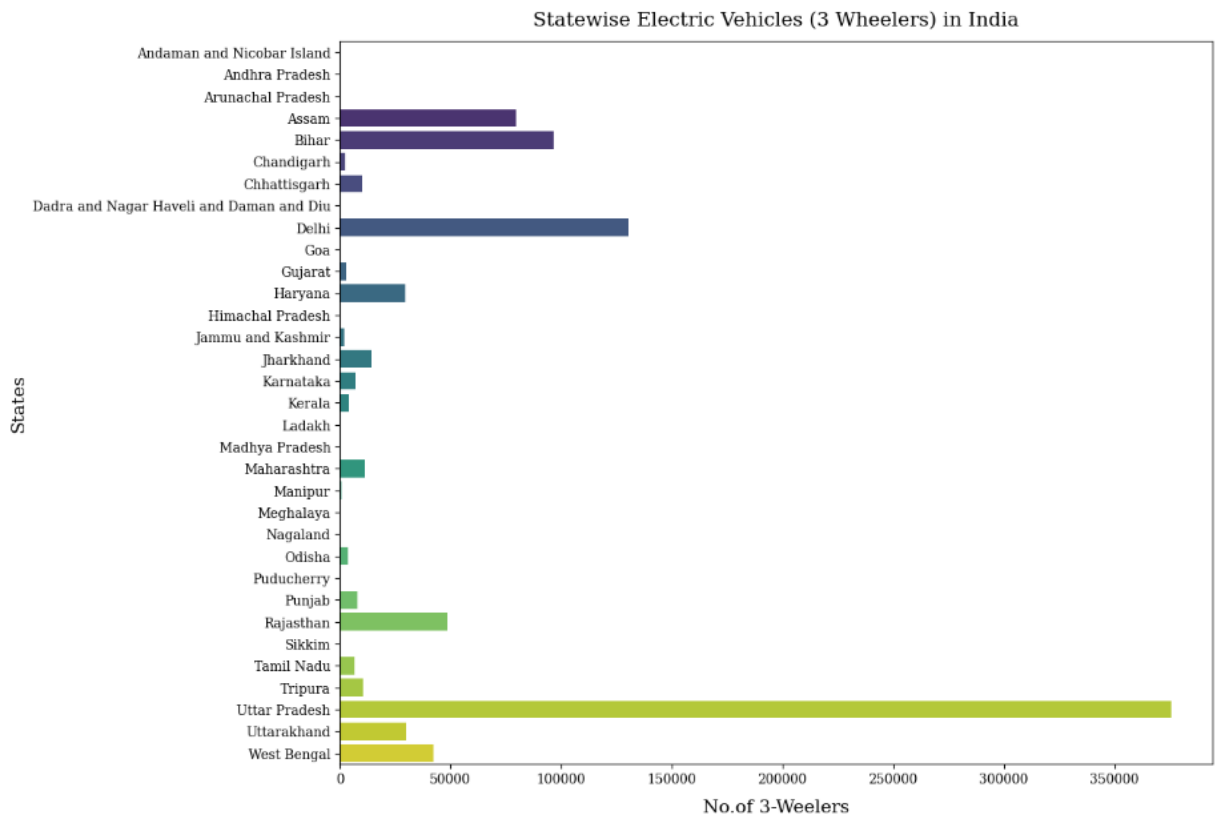


1.

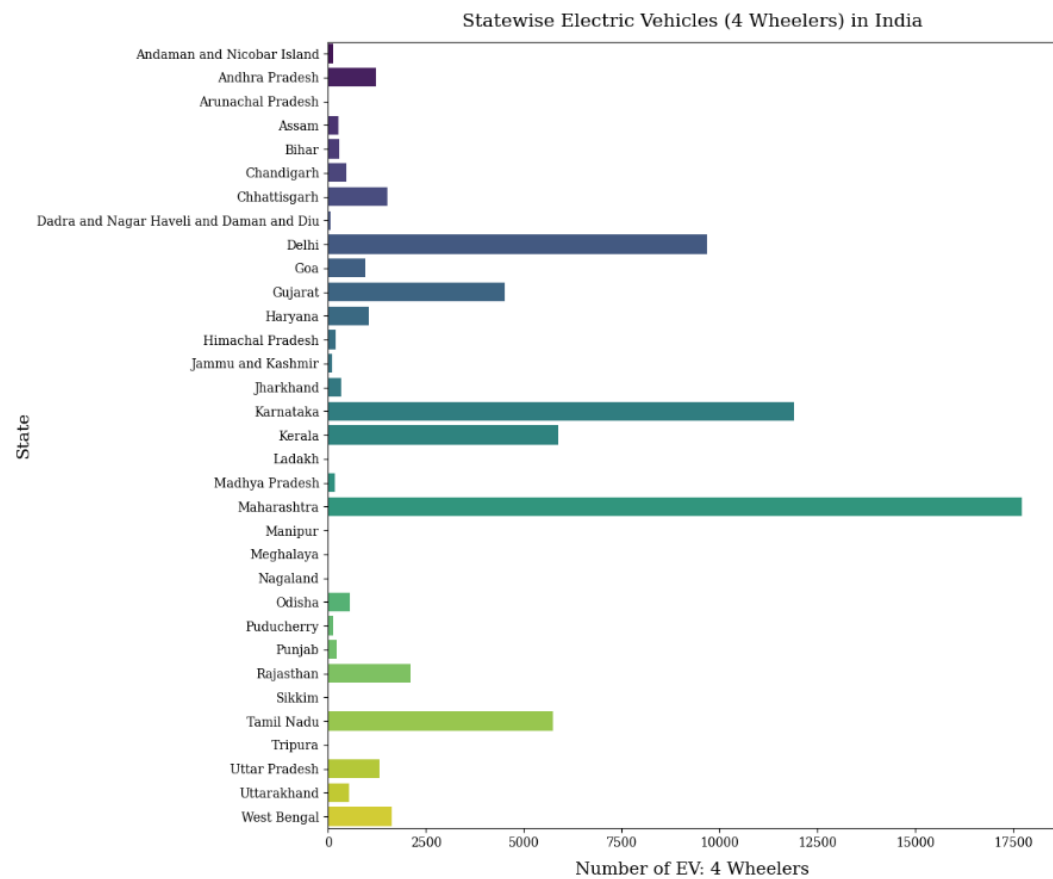
## 2-wheeler EV's



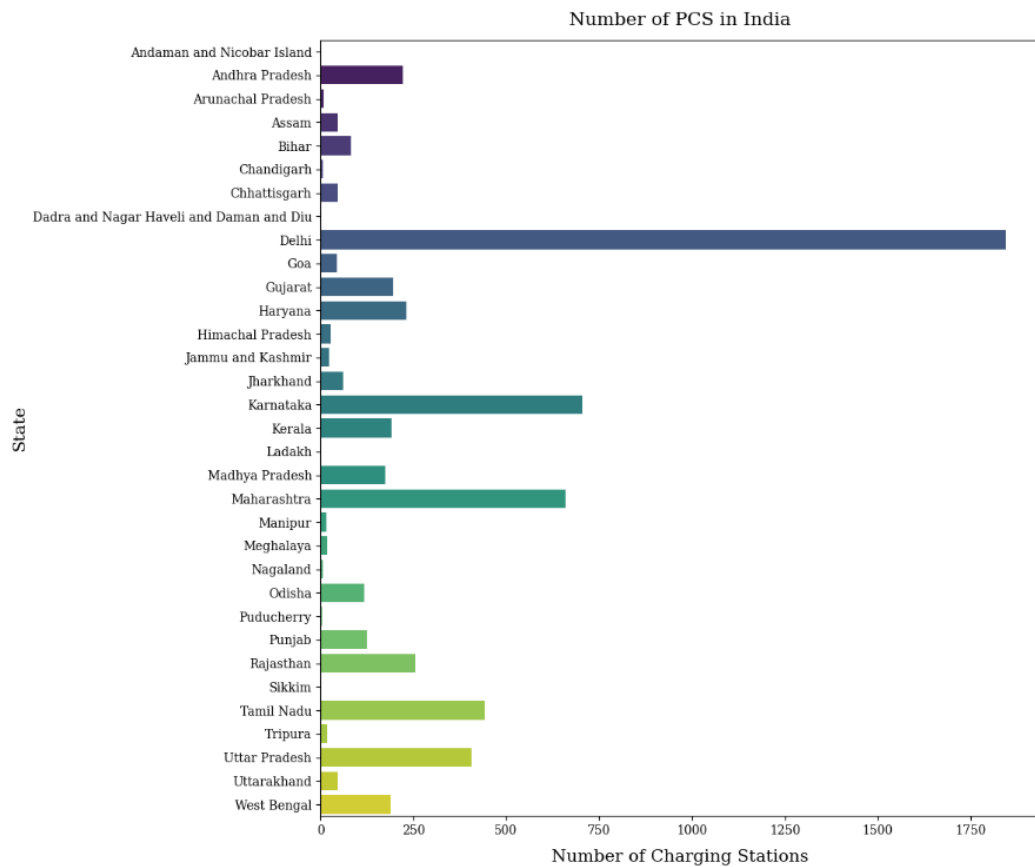
## 3-wheeler EV's



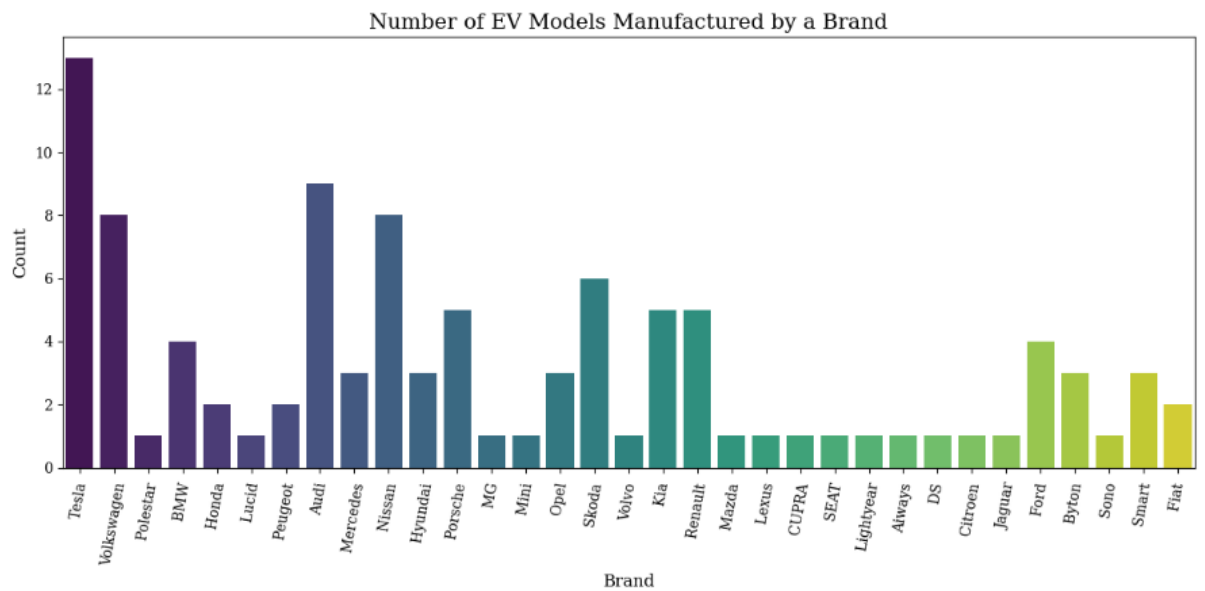
## 4-wheeler EV's



## Public Charging Stations (PCS) in India

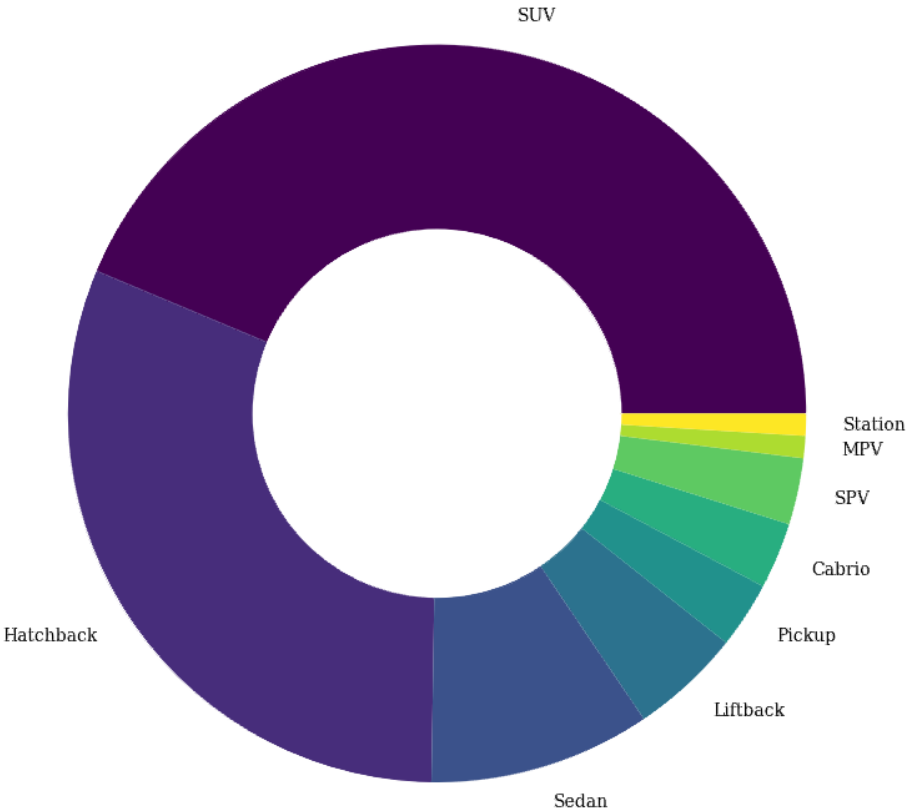


## Brand-wise EV's



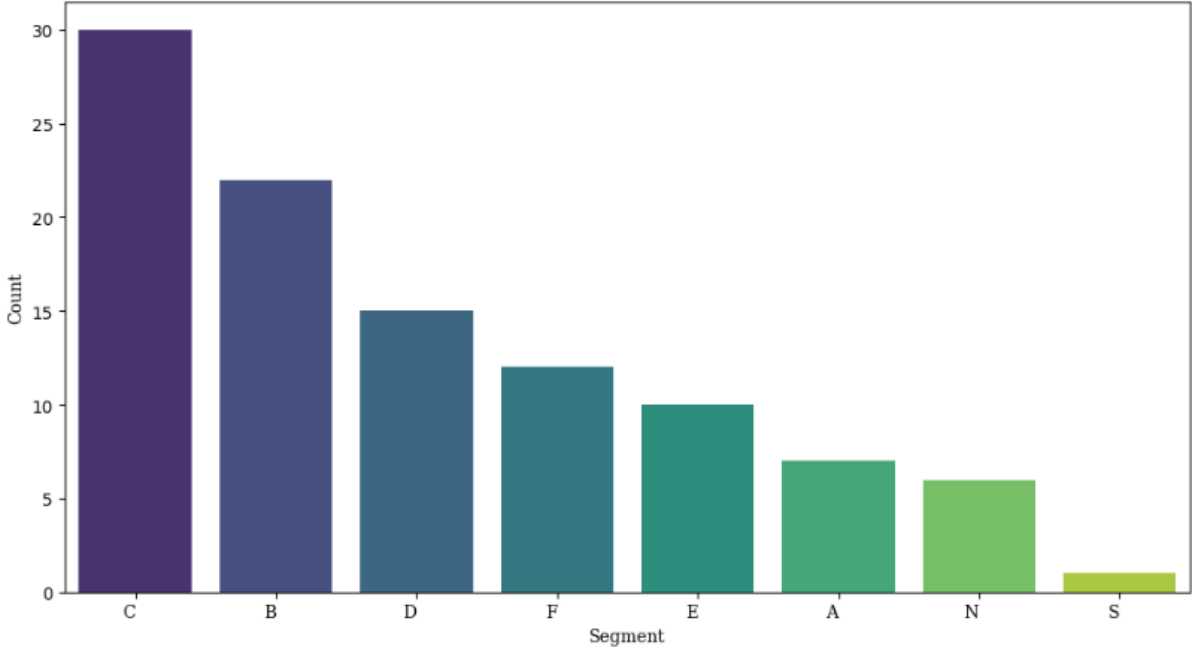
## Body Styles of EV's

Electric Vehicles of Different Body Types in India



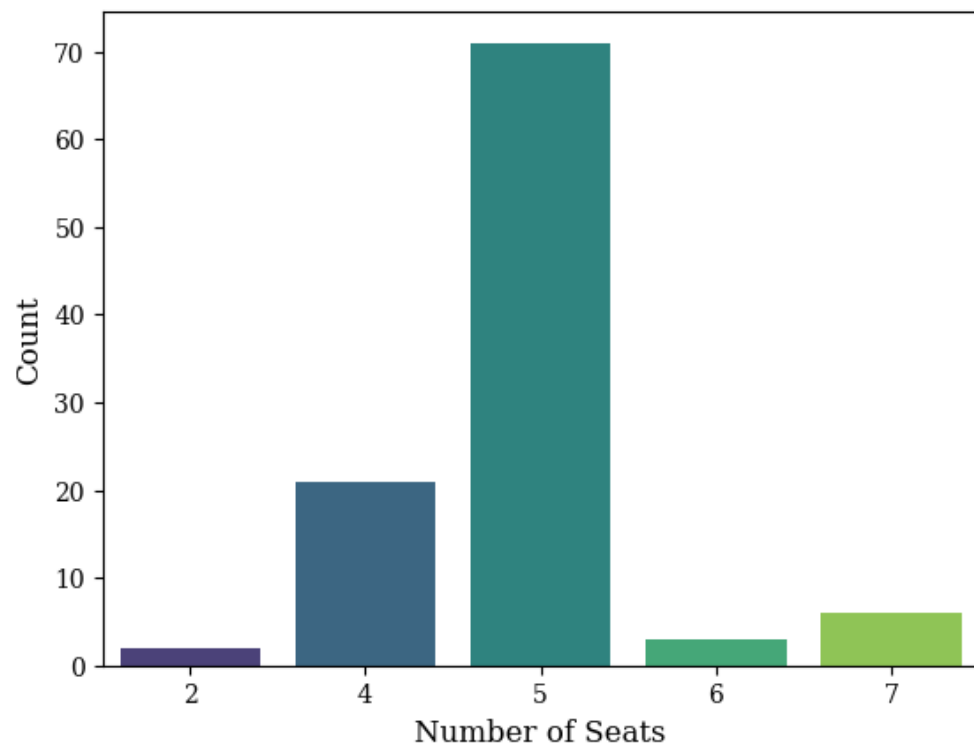
Segments of EV

Electric Vehicles of Different Body Types in India

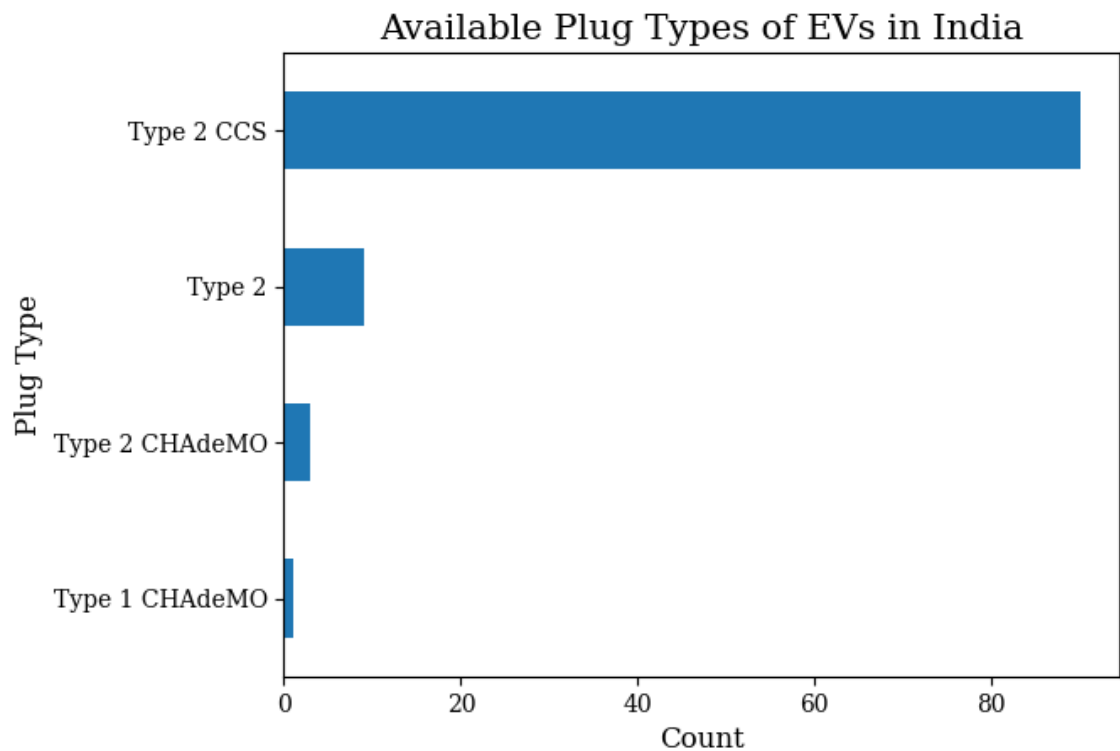


EV's with different No. of seats

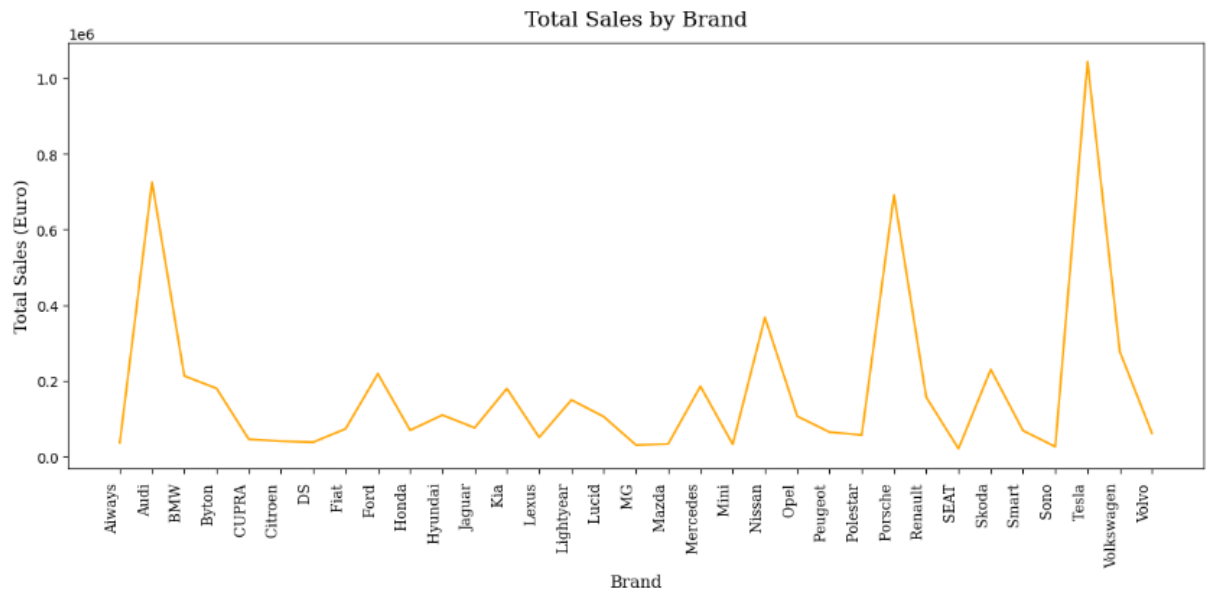
## Available Electric Vehicles of Different Number of Seats in India



## Types of connectors for charging

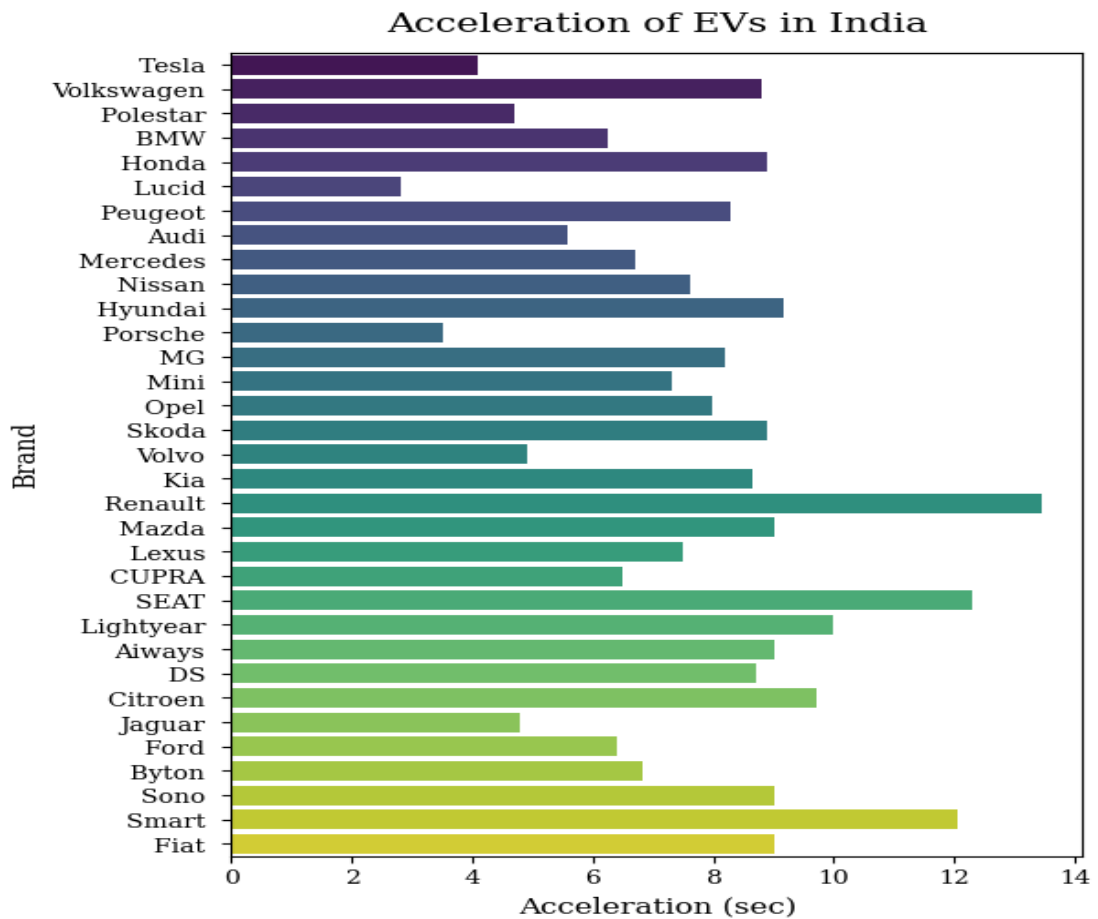


## Brand Sales

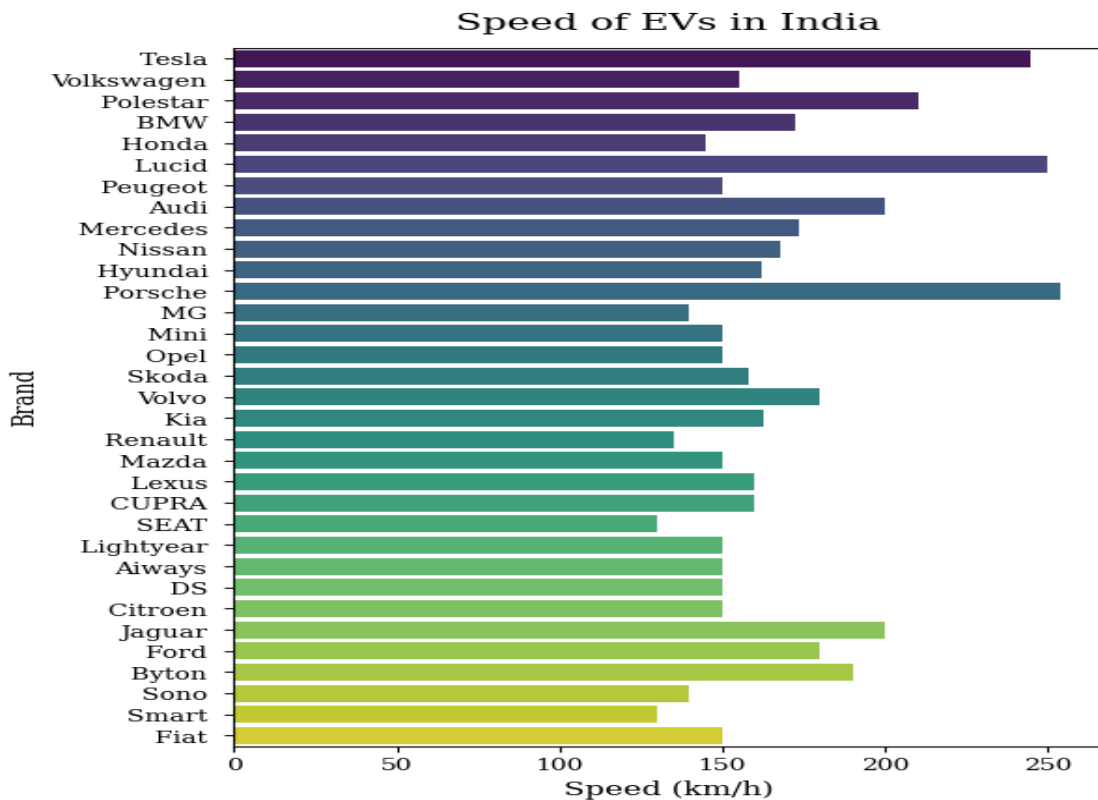


## Brand-wise acceleration

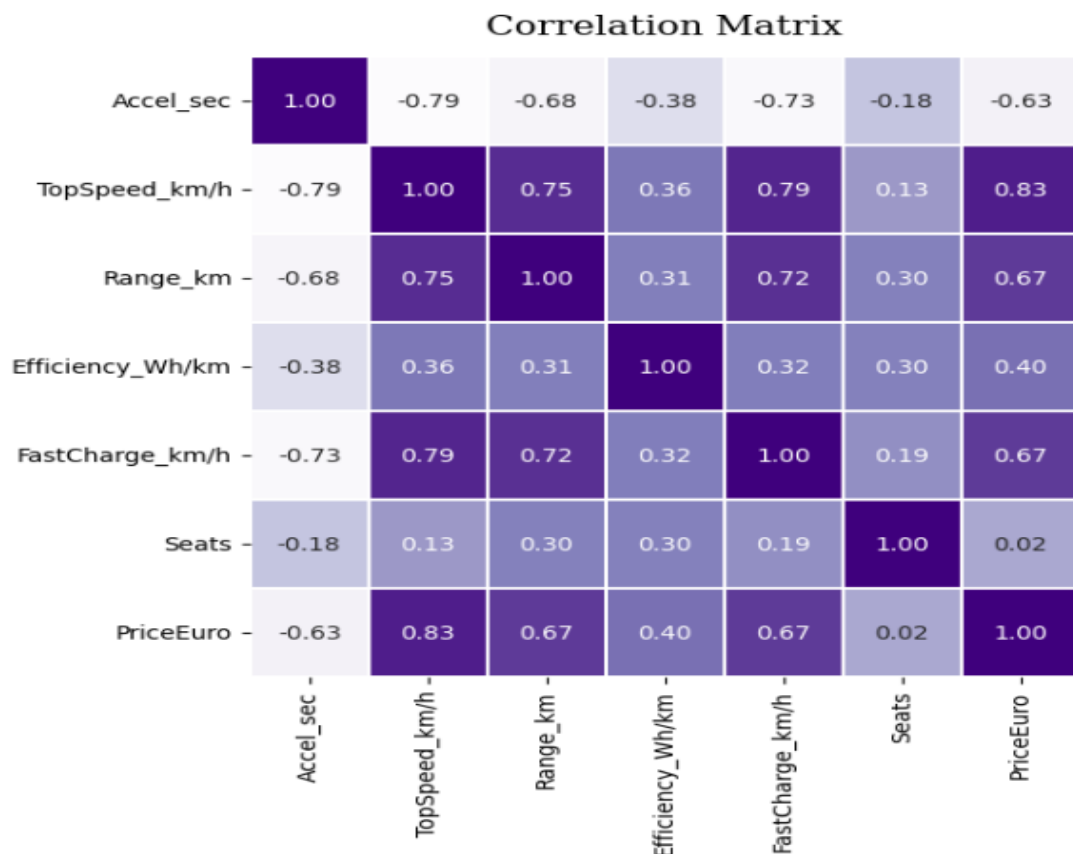




## Brand-wise speed



## Correlation Matrix



**a) Regional Trends:**

- Maharashtra, Karnataka, Tamil Nadu, Gujarat, Rajasthan, and Delhi emerge as key states in the adoption of 2-wheelers EVs.
- Uttar Pradesh, Delhi, Bihar, and Assam lead in the usage of 3-wheeler EVs.
- Maharashtra, Karnataka, Delhi, Kerala, and Tamil Nadu stand out for 4-wheeler EV usage.
- Delhi, Karnataka, Maharashtra, Tamil Nadu, and Uttar Pradesh are prominent in terms of Public Charging Stations (PCS) for EVs.

**b) Brand Analysis:**

- Tesla, Audi, Volkswagen, Nissan, and Skoda are the top brands with the highest number of EV models in the Indian market.
- Tesla, Audi, and Porsche are the leading brands with the highest sales in the EV market.
- Renault, SEAT, and Smart are notable for their high acceleration in EV products.
- Tesla, Lucid, and Porsche are recognized for their speed in the EV segment.

**c) Vehicle Attributes:**

- SUVs and hatchbacks are the most popular body styles among EVs.
- C and B body segments dominate the EV market, with S and N segments being less common.
- 5-seater EVs are the most popular choice among consumers.
- Type 2 CCS connectors are the most widely used for charging EVs.

**d) Correlation Insights:**

- The correlation matrix might reveal relationships between different variables such as brand sales, acceleration, speed, and other factors. For instance, brands with high-speed correlate with higher sales.

**e) Future Implications:**

- Insights from this analysis can inform policymakers, manufacturers, and investors about the dynamics of the EV market in India.
- Understanding regional preferences, brand popularity, and vehicle attributes can help stakeholders tailor their strategies to target specific market segments effectively.
- As the EV market evolves, continuous monitoring of trends and consumer preferences will be crucial for adapting to changing dynamics and fostering sustainable growth.

## 2.

➤ **Dataset Collection:**

**Demographic Data:**

- Age: Understanding age demographics can help in tailoring marketing strategies and product offerings.
- Income Levels: Income data can provide insights into purchasing power and preferences.
- Education Level: Education can influence consumer behaviour and preferences.
- Occupation: Different occupations may have distinct preferences and needs.
- Location: Detailed geographical data can provide insights into regional preferences and market trends.

**Consumer Behaviour Data:**

- Purchase History: Historical data on purchases can help in understanding buying patterns and preferences.
- Online Behaviour: Data on online interactions, such as website visits, clicks, and duration of visits, can provide insights into consumer preferences and interests.
- Social Media Activity: Analyzing social media engagement can help in understanding consumer sentiment and preferences.

**Environmental Factors:**

- **Climate Data:** Climate information can influence consumer preferences for certain types of vehicles or products.
- **Government Policies:** Data on government policies and incentives related to electric vehicles can impact market dynamics.

➤ **Additional ML Models to Try:**

**Clustering Algorithms:**

- **K-Means Clustering:** To segment customers based on similarities in demographic and behavioural attributes.
- **Hierarchical Clustering:** To identify hierarchical relationships among customer segments.

**Ensemble Learning Methods:**

- **Random Forest:** To improve predictive accuracy by aggregating predictions from multiple decision trees.
- **Gradient Boosting:** To sequentially train models and correct errors of the previous models, potentially enhancing predictive performance.

**Neural Network Architectures:**

- **Multilayer Perceptron (MLP):** To capture complex relationships between features and target variables.
- **Convolutional Neural Networks (CNNs):** If image data is available (e.g., for vehicle types or body styles), CNNs can extract relevant features for segmentation.

**Natural Language Processing (NLP) Techniques:**

- **Sentiment Analysis:** To analyze customer reviews and feedback, providing insights into consumer sentiment towards different brands or products.
- **Topic Modelling:** To identify latent topics in textual data, such as online reviews or social media comments, helping in understanding consumer preferences and concerns.

**Model Evaluation and Validation:**

- Implement rigorous cross-validation techniques to assess the generalization performance of the models.
- Use metrics such as silhouette score, Davies–Bouldin index, or adjusted Rand index for evaluating clustering algorithms.
- Conduct A/B testing to validate the effectiveness of marketing strategies targeted at different customer segments.

**Collaborations and Data Partnerships:**

- Collaborate with industry partners and data providers to access proprietary datasets and gain deeper insights into consumer behaviour and market trends.
- Explore opportunities for data sharing and collaboration with government agencies, research institutions, and other stakeholders involved in electric vehicle adoption and sustainability initiatives.

### 3.

To estimate the market size, we need to consider various factors including:

- **Total Number of EVs Sold:** This includes 2-wheelers, 3-wheelers, 4-wheelers, and other categories of electric vehicles.
- **Average Selling Price (ASP) of EVs:** Understanding the average price of EVs across different segments is crucial for estimating revenue.
- **Historical Sales Data:** Analyzing past sales data can provide insights into market growth trends and seasonality.
- **Projected Growth Rate:** Considering factors such as government incentives, infrastructure development, and consumer awareness, we can estimate the future growth rate of the EV market.
- **Market Share of Different Brands:** Understanding the market share of various brands can help in estimating the distribution of sales and revenue.
- **Geographical Distribution:** Considering regional variations in EV adoption rates and preferences can provide a more accurate estimation of the market size.
- **Market Penetration:** Assessing the percentage of households or businesses adopting EVs can help in estimating the total addressable market.

To provide a rough estimate, we would need access to extensive market research data, industry reports, and sales figures from relevant sources such as automotive associations, government agencies, and market research firms. Conducting surveys and interviews with industry experts and stakeholders can also provide valuable insights into market dynamics and trends.

### 4.

#### **Geographical Location:**

- Regional differences play a significant role in consumer preferences, adoption rates, and infrastructure availability.
- Different states or cities may have varying levels of government incentives, charging infrastructure, and environmental awareness, influencing EV adoption.
- Segmenting the market based on geographical location allows for targeted marketing strategies and product offerings tailored to specific regions.

#### **Vehicle Type and Body Style:**

- Segmenting based on vehicle type (e.g., 2-wheelers, 3-wheelers, 4-wheelers) and body style (e.g., SUVs, hatchbacks, sedans) helps in understanding diverse consumer preferences.
- Different vehicle types cater to distinct market segments with varying needs, usage patterns, and price sensitivities.
- Analyzing the popularity of different vehicle types and body styles within specific segments can guide product development and marketing efforts.

#### **Price Sensitivity and Affordability:**

- Understanding consumer affordability and price sensitivity is crucial for market segmentation.

- Segmenting consumers based on income levels, purchasing power, and willingness to pay helps in targeting the right market segments with appropriate pricing strategies.
- Different segments may prioritize factors such as upfront costs, total cost of ownership, and return on investment differently, influencing purchasing decisions.

**Charging Infrastructure and Range Requirements:**

- Availability and accessibility of charging infrastructure significantly impact EV adoption rates and usage patterns.
- Segmenting based on consumers' access to charging facilities (e.g., home charging, public charging stations) and range requirements helps in addressing range anxiety and infrastructure challenges.
- Understanding the importance of factors such as charging speed, convenience, and interoperability can help in designing tailored solutions and services for different market segments.