**MARKETING SEGMENTATION FOR AN ELECTRIC VEHICLE STARTUP IN UPCOMING MARKET(INDIA)**

By Sandeep Singh Sajwan

*Dated: Oct 2, 2024*

**Objective:**  
This report provides a comprehensive analysis of the Electric Vehicle (EV) market based on key performance indicators such as battery capacity, range per charge, charging time, price, power, and top speed. The goal is to segment the EV market into distinct clusters using K-Means clustering, conduct a Principal Component Analysis (PCA) to understand key drivers, and identify the optimal cluster representing the perfect choice of EV for the upcoming market.

1. **Problem Statement:**

An Electric Vehicle (EV) startup in India faces a crucial decision: who to target. The diverse Indian market offers numerous possibilities, but choosing the right segment is essential for success. This project tackles this challenge through market segmentation analysis. By understanding different customer groups (geographic, demographic, etc.), we will assess their needs and the competition. This will help us identify the ‘sweet pot’: the segment most likely to embrace the startup’s EVs. Considering data limitations and market dynamics, we will develop a targeted entry strategy, positioning the startup for long-term growth in the electrifying Indian market.

1. **Data Collection:**

To initialize the implementation of the market segmentation analysis for our EV startup’s Indian launch, I started with the data acquisition efforts. Through meticulous research, I delved into various data sources available on the internet to gather appropriate and relevant data for the project. This comprehensive data collection exercise lays the groundwork for the next crucial step: identifying the most promising segment for our startup’s successful entry into the electrifying Indian EV market.

***Websites used for collecting the data:***

• <https://www.kaggle.com/>

• <https://www.data.gov.in/>

• <https://datasetsearch.research.google.com/>

• <https://trends.google.com/trends/explore>

**Dataset that I used for the project:**

• The Fist dataset specifies the total no. of vehicles registered in different segments in different states of India till 2023. This data has been categorized into states of the different segments of vehicle registered.

• Analyzing this data would give an idea about the purchasing trends in different states of the Indian Market and upcoming scope of market.

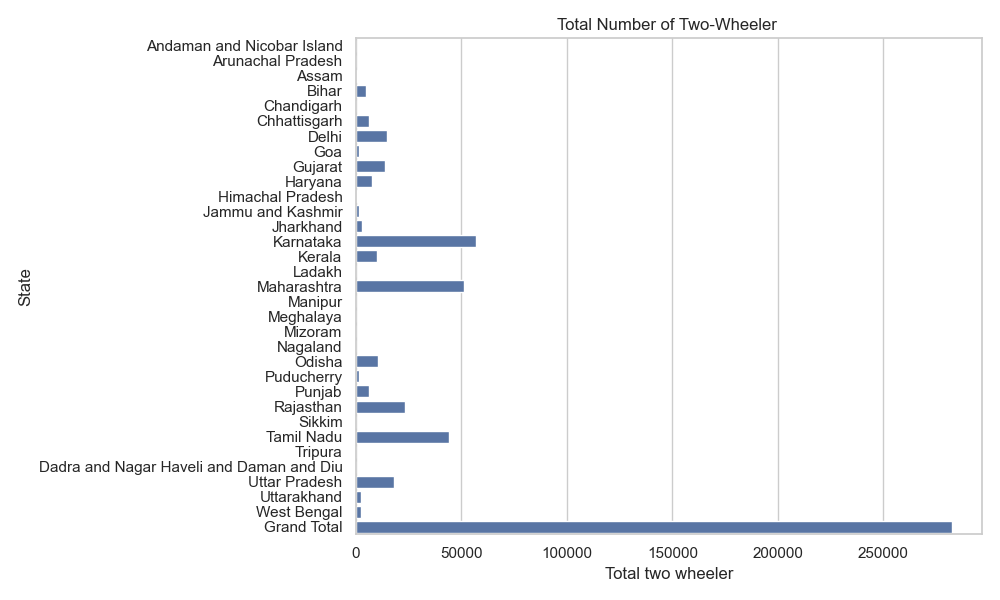
• The Second dataset specifies the list of market holders in 2 Wheeler segment i.e. Scooters and bikes. This list encapsulates all the specific details targeting the customer. With the help of this data, we can further see the future venture in the EV segment.

• The Third dataset shows the sales figures of EVs from Apr-2017 till May-2023. This further supports our vision for the growth of 2W in Upcoming years.

1. **Exploratory Data Analysis:**

An Exploratory Data Analysis or EDA is a thorough examination meant to uncover the underlying structure of a data set and is important for a company because it exposes trends, patterns, and relationships that are not readily apparent.

2 wheelers data visualization from dataset 1

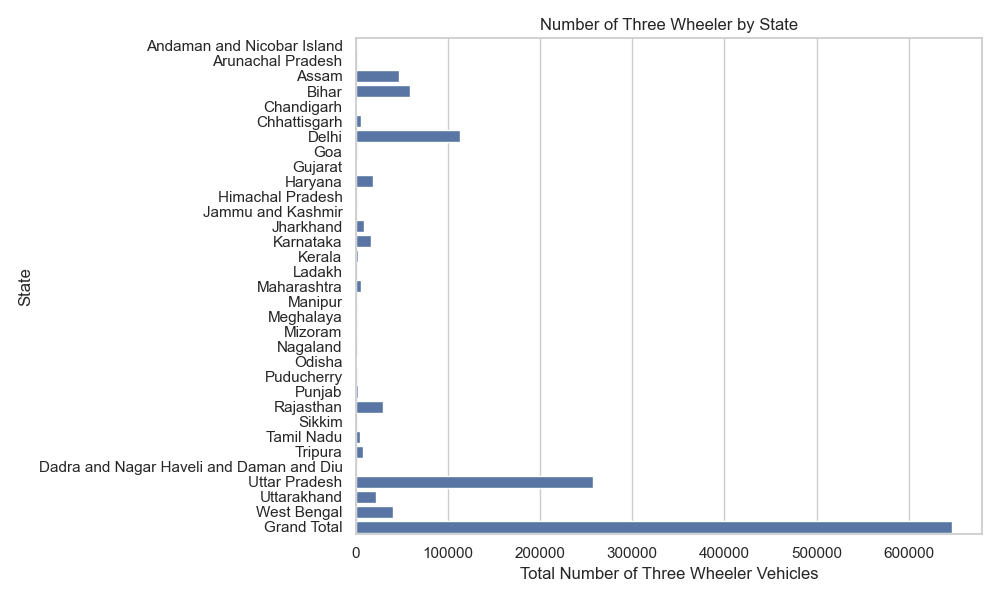


• Out of all the states in India, the top 3 states with the highest number of 2-W Electric Vehicles comes to be Karnataka, Maharashtra and Tamil Nadu and seems to suggest a growing market in the near future.

• The Lowest number of 2-W Electric Vehicles is in Andaman and Nicobar Islands, Sikkim, Ladakh and Arunachal Pradesh.

• The Total two Wheeler segment also seems promising for an upcoming market as we will see in later results.

3 wheelers data visualization from dataset 1

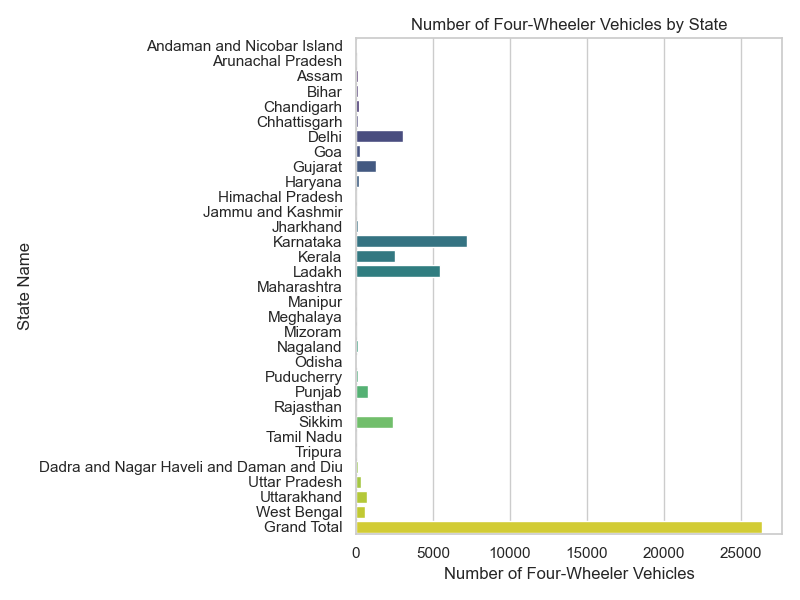


• Out of all the states in India, the top 2 states with the highest number of 3-W Electric Vehicles comes to be Uttar Pradesh and Delhi and seems to suggest a growing market in the near future.

• The Lowest number of 3-W Electric Vehicles is in Nagaland, Mizoram, Sikkim, Ladakh and Arunachal Pradesh.

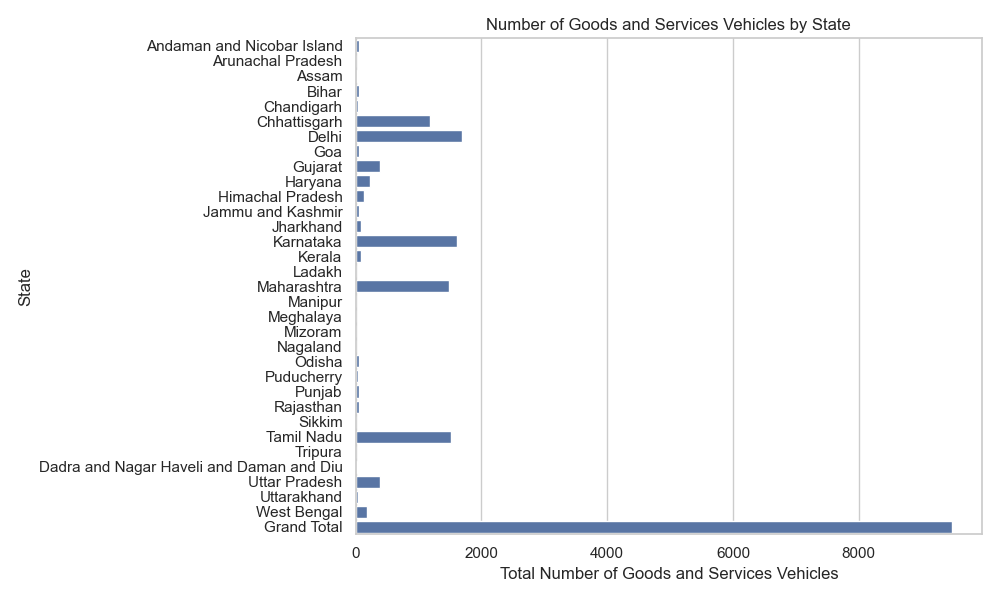
• The Total three Wheeler segment also seems promising for an upcoming market as we will see in later results.

4 wheelers data visualization from dataset 1

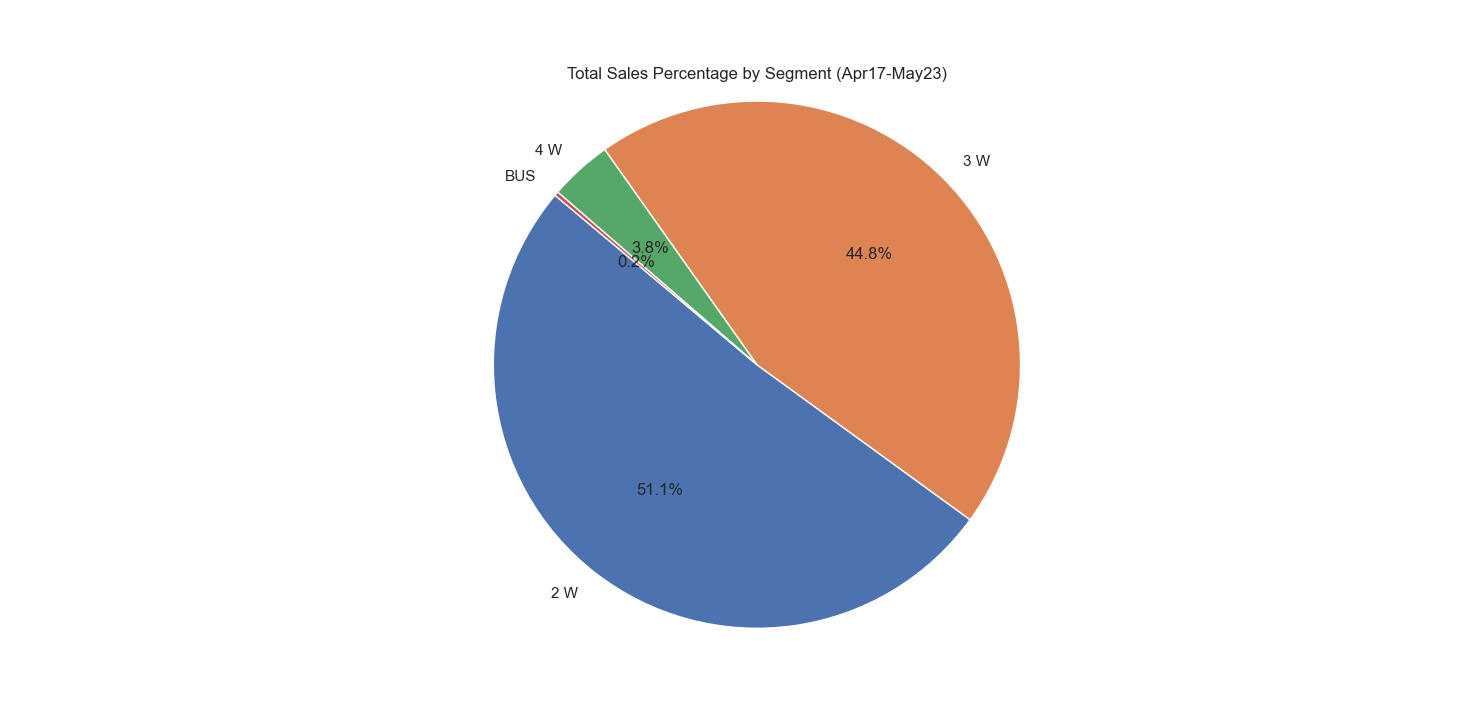


• From this graph, we can Karnataka and Ladakh among the states with the highest no. of 4W Electric Vehicles.

Goods and Services Vehicles data visualization from dataset 1

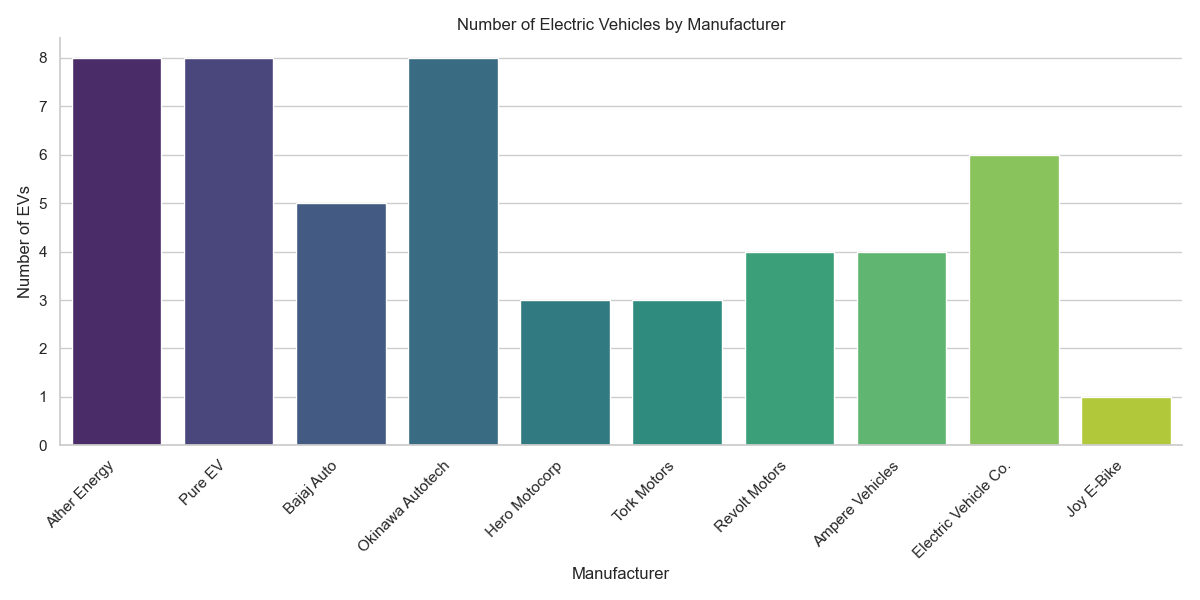
• This representation shows that Delhi, Karnataka, Maharashtra, Tamil Nadu and Chhattisgarh emerging as states with highest no. of Goods and Services Electric Vehicles.

1. **Analysis of distribution of Sales in percentage by segments:**



This Pie-Chart shows the sales of different segment of EVs in indian market from Apr-2017 to May-2023. Through this analysis, we can clearly state that the entire market is held by two segments of EV, namely, 2W(51.1%) and 3W(44.8%).

1. **Brand-wise 2-Wheeler EV segmentation**



Ather Energy, Pure EV and Okinawa Autotech tops the industry of making highest no. of models in 2-W segment(Scooter and bike).

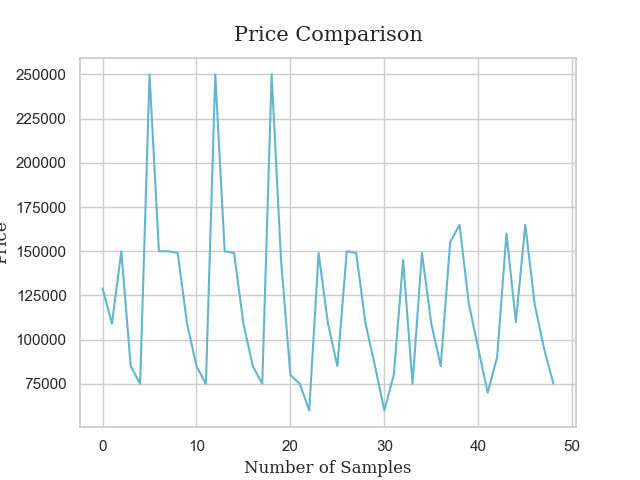
1. **Analysis of Prices for Customer Segmentation**



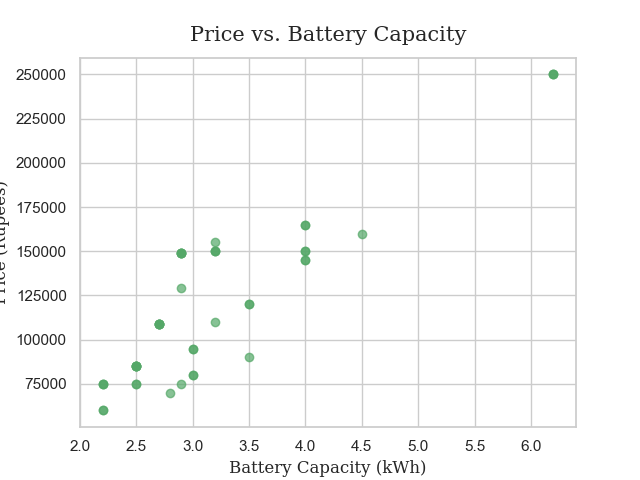
This shows that majority of purchased vehicles is in low price segment(60k-100k), followed by low to medium group(100k-150k).

1. **Analysis of Price Comparison**

Another Graph of Price Comparison (below) shows similar results as in the above graph. With the increasing no of samples, the price shows downtrend.



1. **Analysis of Prices for Different Battery Capacity**

****

The above plot shows with the increase in battery capacity (kWh), there is relative increase in prices too.

1. **Average price range of 2W Electric Vehicles:**

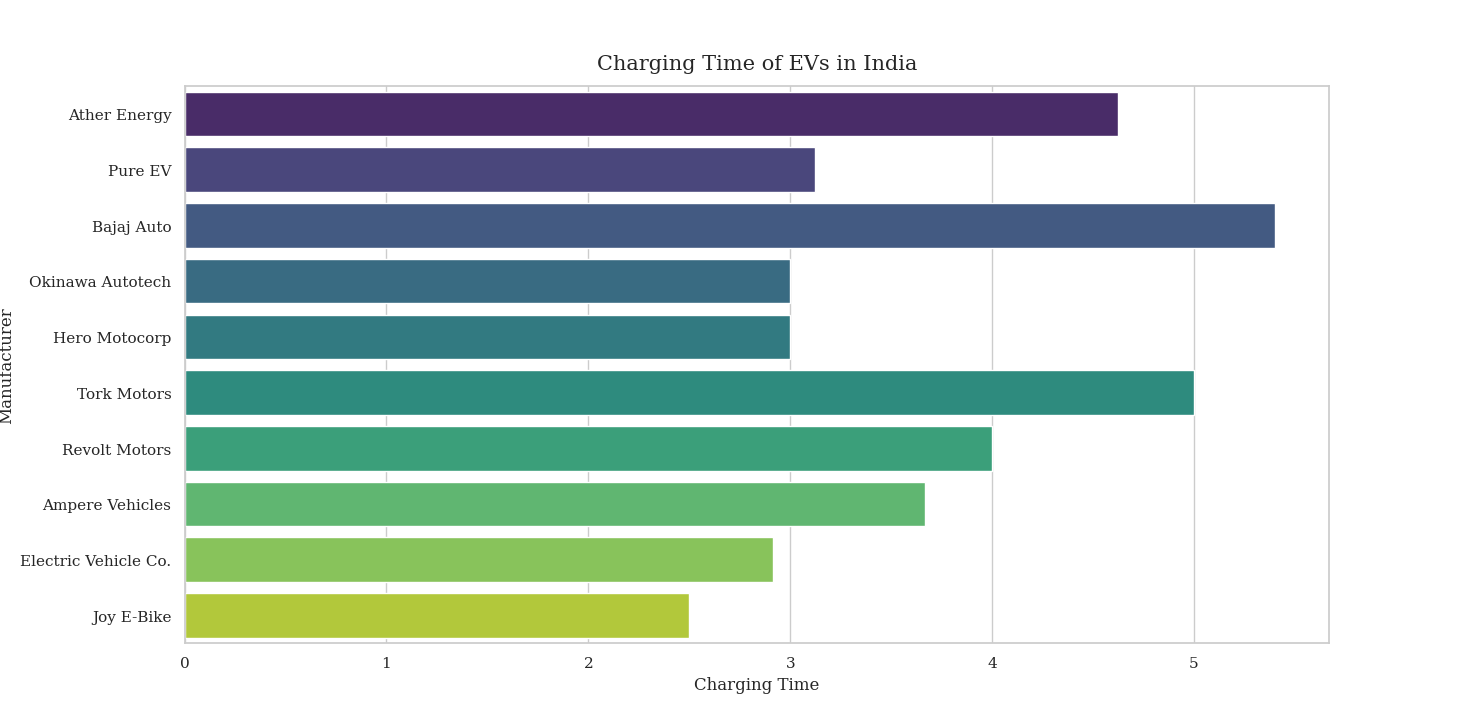


Average Price Ranges comes in the low to medium price group for customer segmentation.

1. **Charging Time Visualization**

The charging time is lowest for Joy E-bike, whereas it is highest for manufacturer like Bajaj Auto.

The median seems to be around 3hr which is achieved by Electric Vehicles Co., Hero Motocorp, Okinawa Autotech and Pue EV. These same companies are also leading manufacturers in this segment as already shown by previous analysis.



1. **Correlation Matrix**

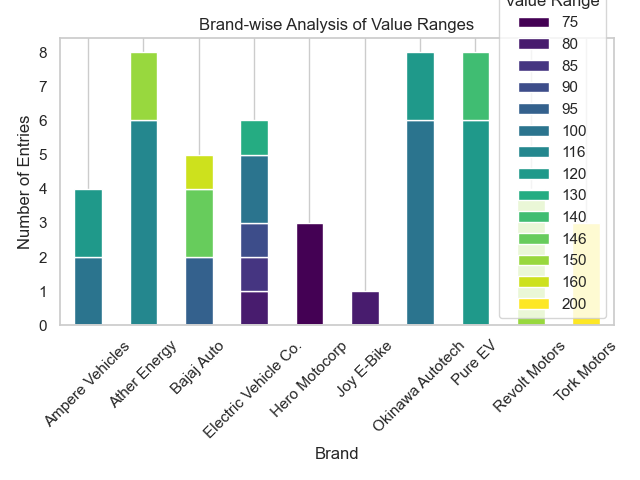


• Strong Positive Correlation exists between 'Top Speed (km/h)' and 'Range per Charge (km)' (0.91).

• Strong Negative Correlation exists between ‘Price’ and ‘Manufacturer’ (-0.08).

• Moderate Positive Correlation exists between ‘Price’ and ‘Charging Time’ (0.80).

1. **Brand-wise Analysis**



### Step 1: Data Preprocessing

Before applying any clustering algorithms, the dataset was cleaned and preprocessed:

* **Missing values** were handled.
* **Feature scaling** was applied using StandardScaler to normalize the range of all features.
* Unnecessary rows and columns were removed.

### Step 2: K-Means Clustering

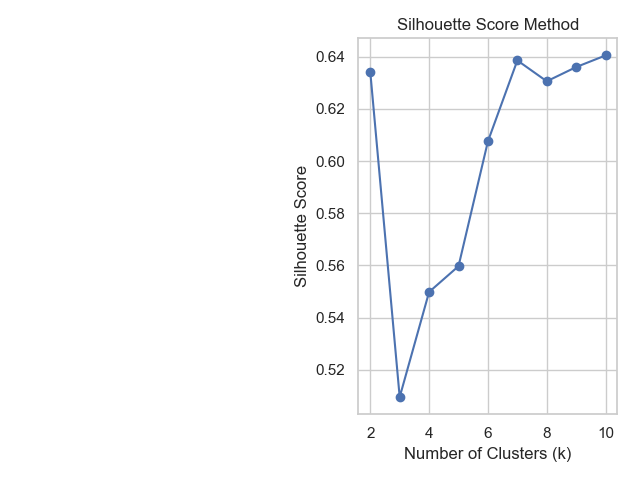
We applied **K-Means Clustering** to segment the dataset into distinct groups of EVs. The Elbow method was used to determine the optimal number of clusters.

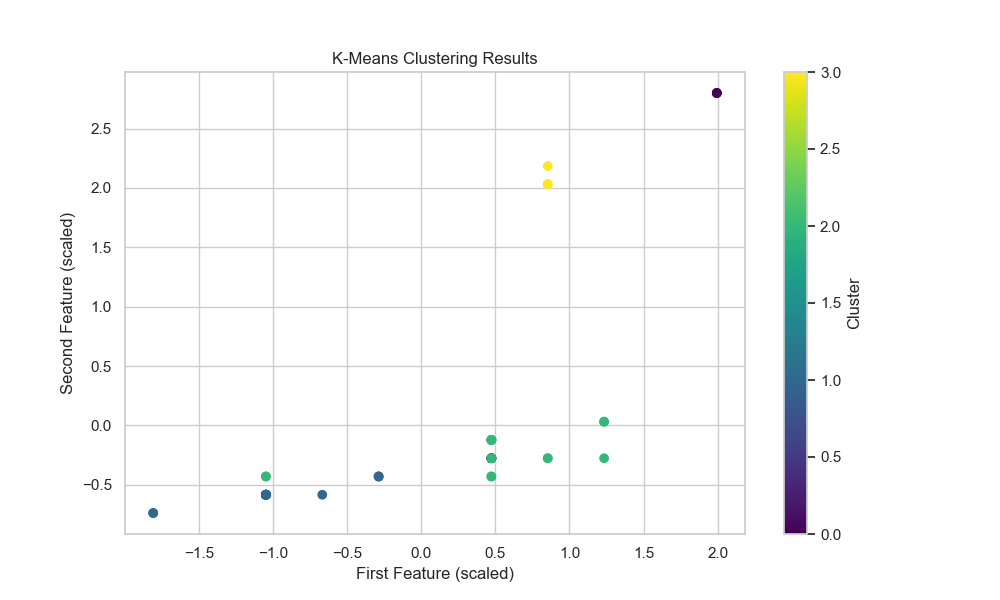
#### Elbow Method Results:

* We plotted the sum of squared distances (inertia) for different numbers of clusters (k=2 to k=10).
* The **optimal number of clusters** was found to be **4**, as indicated by the "elbow" in the plot.

#### K-Means Clustering Summary:

The K-Means algorithm was trained with **k=4**, creating 4 distinct clusters of EVs. Each cluster groups vehicles with similar characteristics in terms of battery capacity, range, power, and price.





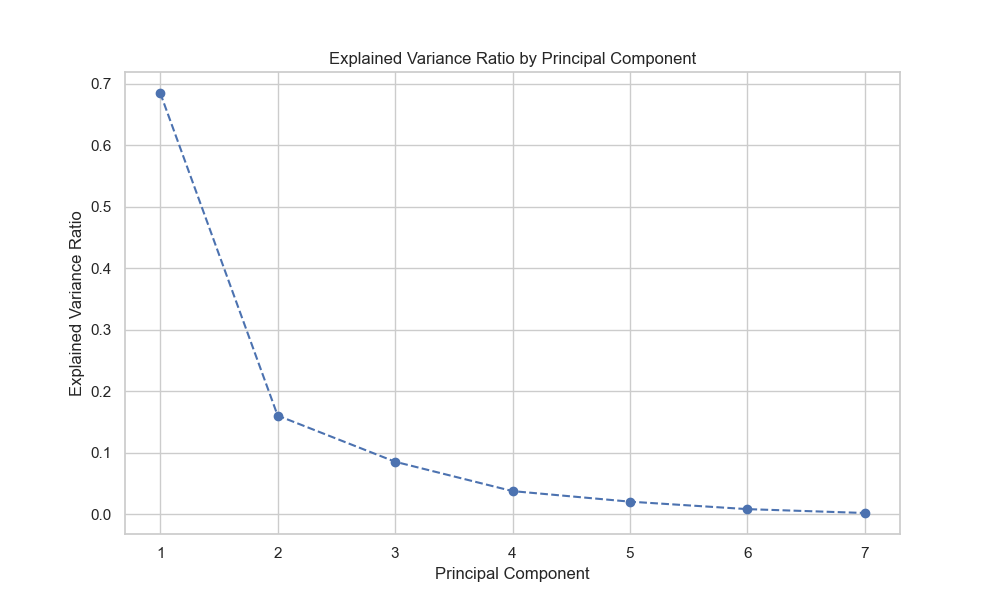
### Step 3: Principal Component Analysis (PCA)

To reduce the dimensionality and understand the primary factors driving EV differentiation, we applied **Principal Component Analysis (PCA)**.

#### PCA Results:

* **7 components** were extracted, accounting for over **75% of the total variance** in the dataset.
* The first 3 principal components captured the majority of the explained variance:
  + **PC1**: Primarily represented price and battery capacity.
  + **PC2**: Strongly influenced by range and power.
  + **PC3**: Represented top speed and charging time.

This analysis revealed that **price, battery capacity, and range per charge** are the most critical factors distinguishing different types of EVs.

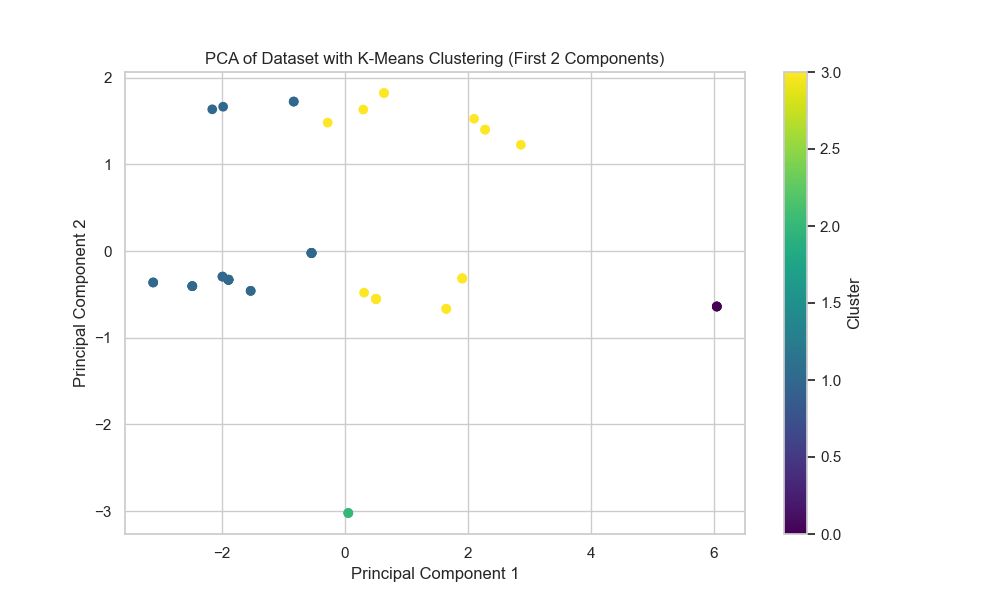


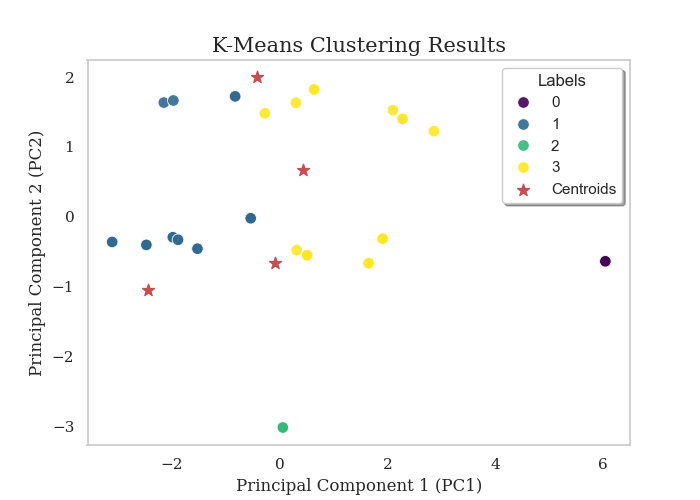
### Step 4: Cluster Profiles

To understand the characteristics of each cluster, we calculated the mean values of the key features for each group:

#### Cluster Profiles:

1. **Cluster 0** (Affordable & Efficient EVs):
   * **Average Price**: Low to medium price range.
   * **Average Range**: Moderate range per charge.
   * **Power & Speed**: Low to medium power and speed.
   * **Ideal for**: Cost-conscious consumers looking for efficient daily-use EVs.
2. **Cluster 1** (High-Performance EVs):
   * **Average Price**: High-end pricing.
   * **Average Range**: Long range per charge.
   * **Power & Speed**: High power output and top speed.
   * **Ideal for**: Luxury buyers who prioritize performance and long-range travel.
3. **Cluster 2** (Budget EVs):
   * **Average Price**: Low price range.
   * **Average Range**: Shorter range per charge.
   * **Power & Speed**: Lower power and speed.
   * **Ideal for**: Entry-level EV buyers or city dwellers who prioritize affordability over performance.
4. **Cluster 3** (Balanced Mid-Range EVs):
   * **Average Price**: Mid-range pricing.
   * **Average Range**: Sufficient range for most users.
   * **Power & Speed**: Balanced power and speed.
   * **Ideal for**: Consumers seeking value for money with balanced features.





**Project GitHub Link:**

***GitHub Link:*** [**https://github.com/sandeep4seyeon/EV-Segmentation-Startup**](https://github.com/sandeep4seyeon/EV-Segmentation-Startup)

**Conclusion: The Perfect Choice for the Upcoming Market**

The analysis suggests that **mid-range EVs** with a balance between price, range, and power are best suited for the upcoming market. These vehicles, which offer a mix of affordability and performance, align with consumer preferences for long-range capabilities without breaking the bank. EV manufacturers should focus on optimizing battery technology and reducing costs to provide the best value for this segment.

By targeting Cluster 3, manufacturers can capture the majority of consumers looking for reliable, moderately priced EVs with good range and performance.

**Visualizations**

**1. Elbow Method (Optimal K Selection)**

* Displays the optimal number of clusters based on the elbow point.

**2. PCA Explained Variance Plot**

* Shows how much variance each principal component explains.

**3. K-Means Clustering Visualization**

* Scatter plot of EVs segmented into clusters, with centroids marked for each group.

**Recommendations:**

1. **Focus on Mid-Range 2W EVs**: Prioritize vehicles in the mid-range segment with balanced performance features.
2. **Battery and Range Improvements**: Continue optimizing battery technology to improve the range and reduce charging times.
3. **Affordable Pricing**: Ensure pricing remains competitive, especially for mid-range EVs.
4. **User Experience**: Enhance user experience by focusing on practical features like fast charging, intuitive technology, and comfort.

This analysis will guide strategic decisions for EV manufacturers and marketers targeting the most attractive segment in the evolving electric vehicle market.