# ELECTRIC VEHICLE MARKET SEGMENTATION FOR A STARTUP COMPANY IN INDIA

Gaurav Prakash

21 May 2024

GitHub

## Fermi Estimation

This Project is made working under an Electric Vehicle Startup. The objective is to help the Startup in deciding that which Vehicle/Customer it is going to develop its EVs. And to come up with a feasible strategy to enter the market, targeting the segments most likely to use Electric Vehicles. There is basically 2 possible client requirements-

**1. Product based:** This focuses on the Vehicle type the company is going to produce. Available criteria are-

- 2 Wheelers including both scooters and motorbikes,
- 4 Wheelers including Hatchback, Sedan, SUV etc.,
- 3 Wheelers including e-rikshaws, and
- Electric Buses.

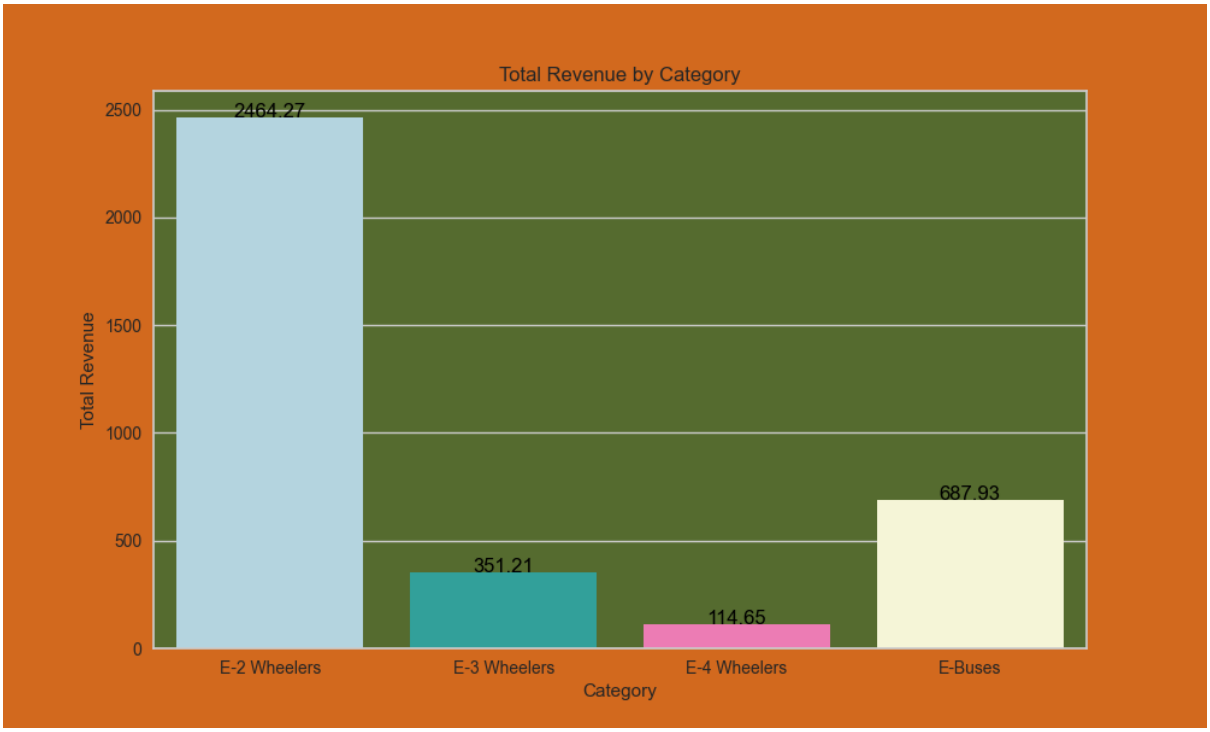
**2. Customer based:** This focuses on the question "Who our customers are?', The possible Segmentation/Groups are-

- **Psychographic Segmentation:** Consumers who prioritize sustainability and innovation in EVs. These segments often overlap with those interested in smart home technologies and renewable energy solutions.
- **Behavioural Segmentation:** Early adopters and frequent commuters are significant segments. Benefits sought include reduced operating costs, environmental benefits, and access to new technologies.
- **Demographic Segmentation:** The primary demographic for EVs includes tech-savvy young professionals, middle to upper-income groups, and environmentally conscious individuals.
- **Geographic Segmentation:** Urban areas, especially Tier 1 and Tier 2 cities, show higher EV adoption rates. Regions like Maharashtra, Delhi-NCR, and Karnataka are leading in EV adoption.

## 1. Data Sources

There are total of 3 datasets used in this project. Two of them are obtained from bikewale.com; It is one of the India's leading sources of bike and scooter related information. Out of the two datasets, one contains Price and Technical specifications of various E-2 Wheeler brand models. Another one of the datasets contains reviews of various customers of 2 Wheelers EVs, which aids in psychographic and behavioural segmentation.

The third dataset which is used in the beginning of the Project contains various information regarding Sales of various types of EVs over the period of 2017 to 2023. This dataset tells us exactly the market trends of various vehicle types including 2 Wheelers, 3 Wheelers, 4 Wheelers, and E-Buses. The dataset is made available by the Society of Manufacturers of Electric Vehicles India.




The above graph shows the total revenue made by each of the 4 categories of EV, and we can definitely observe in an instance that 2 Wheelers clearly stands out with an outstanding revenue of 2464.27 Cr. made so far through the year 2017-2023.

## 2. Data Pre-Processing (Steps and Libraries Used)

A brief overview of the preprocessing steps undertaken-

- **Data Loading:** Multiple datasets related to EV sales, market, specifications, and reviews were loaded from Excel and CSV files using pandas.

- **Data Merging:** The datasets were merged on common keys, such as model names, to create a comprehensive dataset that includes both sales figures and customer reviews.

- **Handling Missing Values:** Missing values in the dataset were identified and handled appropriately. For instance, missing review sentiments were filled with 'neutral', while other missing numeric values were filled with zeroes or mean values as necessary.

- **Standardizing Categorical Values:** Brand names and other categorical values were standardized to ensure consistency. For example, different variations of the same brand name were unified.

- **Sentiment Analysis:** Sentiment analysis was performed on customer reviews. Reviews were classified as positive, negative, or neutral based on sentiment scores.

- **Feature Selection:** Relevant features for clustering were selected, including review aspects like performance, reliability, and service experience.
- **Scaling:** The selected features were scaled using Standard Scaler from scikit-learn to standardize the range of values, which is important for distance-based clustering algorithms like K-Means.

- **Dimensionality Reduction:** Principal Component Analysis (PCA) was applied to reduce the dimensionality of the data, making it easier to visualize and analyse while retaining the most important information.

- **Clustering:** The processed data was then used to perform clustering using the K-Means algorithm. The optimal number of clusters was determined using the Elbow method and silhouette scores.

These preprocessing steps ensure that the data is clean, consistent, and ready for detailed analysis, leading to accurate and actionable insights for the EV startup.

The major libraries used in this project are –

- **Pandas:** For data exploration and analysis and dealing with available data in tabular format i.e. Data Frame.
- **NumPy:** For numerical python calculations.
- **Matplotlib, Seaborn:** For Data Visualization and insights through charts and trends.
- **Scikit-Learn:** For implementing Machine Learning Algorithms, like K-Means, metrices, PCA, etc.
- **NLTK:** For Sentiment Analysis of review text.
- **Yellowbrick:** For model selection process by plotting the optimum number of clusters.

# 3. Segment Extraction

Segment extraction is the process of identifying distinct groups within the data that share similar characteristics. Here's a brief overview of the segment extraction process undertaken in this project:

### Clustering Algorithm:

- **K-Means Clustering:** The K-Means algorithm was used to partition the pre-processed data into clusters. The choice of K-Means is due to its efficiency and effectiveness in handling large datasets.

### Determining Optimal Number of Clusters:

- **Elbow Method:** This method was used to plot the within-cluster sum of squares (WCSS) against the number of clusters to find the 'elbow point' where the WCSS begins to diminish significantly. This helps in identifying an appropriate number of clusters.

- **Silhouette Score:** This metric was used to measure the quality of clustering. Higher silhouette scores indicate better-defined clusters. The optimal number of clusters was chosen based on the combination of these methods.

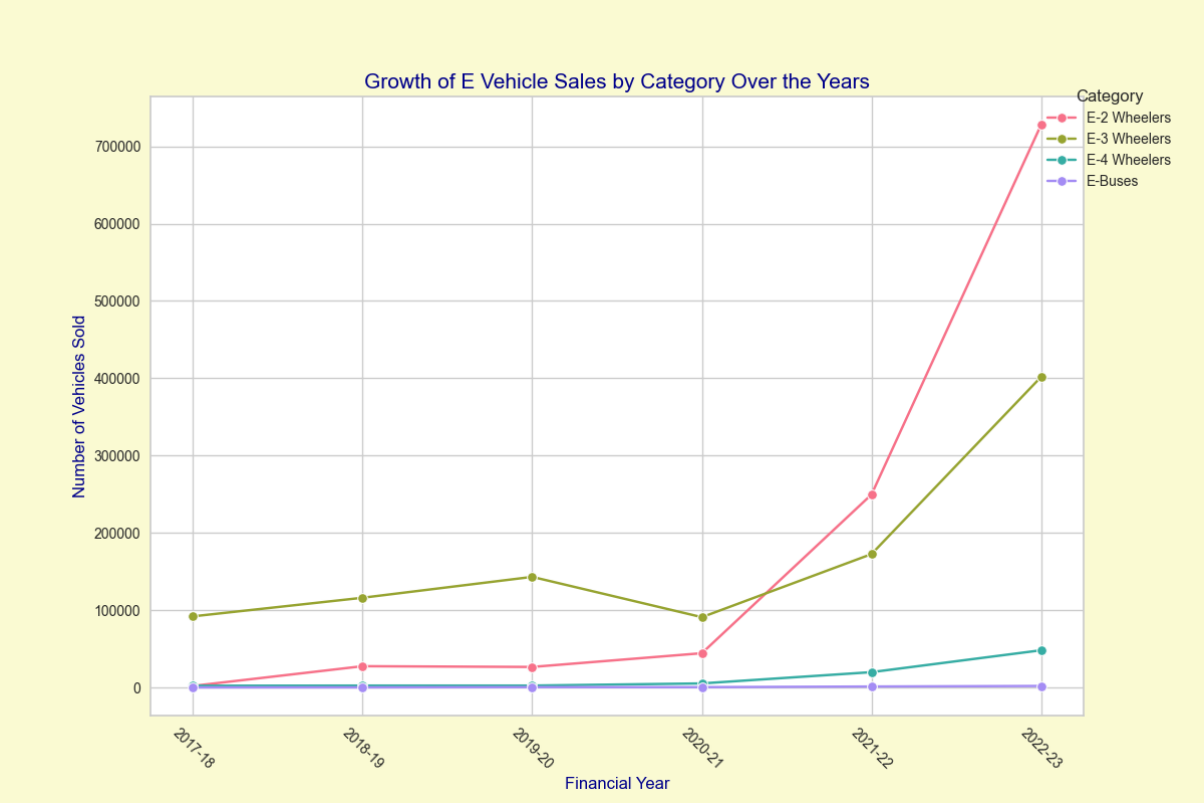
### Feature Importance:

- **Principal Component Analysis (PCA):** Before clustering, PCA was used to reduce the dimensionality of the dataset while retaining the most significant features. This helps in visualizing the clusters and understanding the feature importance.

### Cluster Analysis:

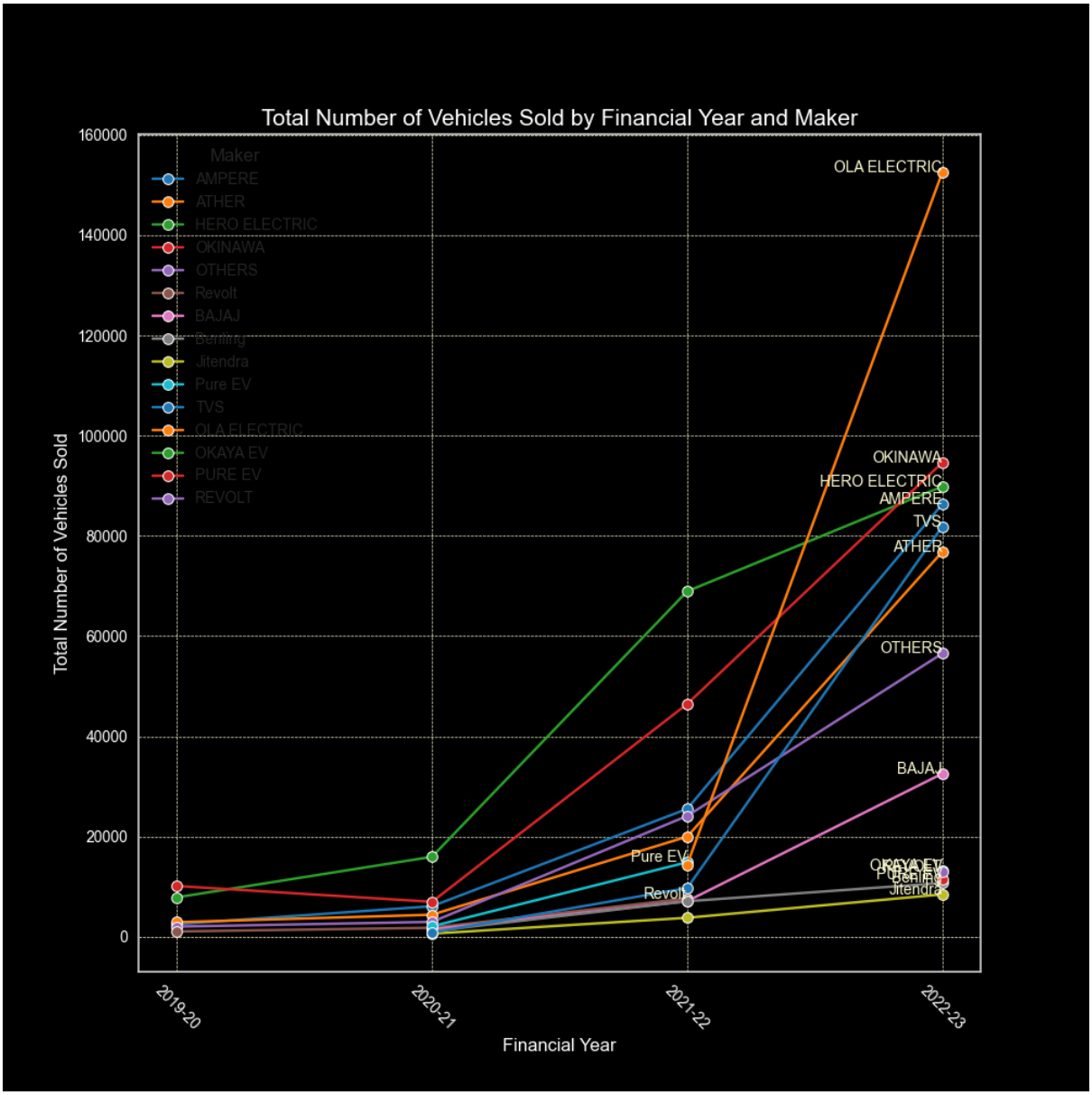
- **Cluster Profiles:** Each cluster was analysed to identify its unique characteristics. This involved looking at the central tendencies of features such as brand preferences, sentiment scores, and key performance metrics.
- **Visualization:** Clusters were visualized using scatter plots and PCA component plots to understand the distribution and separation of different segments.

The below figure shows the trend comparison of Sales over the years for different EV types.




We can clearly see how from year 2021-22 to year 2022-2023 the E-2 Wheelers skyrocketed in terms of growth aspect.

After this I checked and analysed how and which brands in 2 Wheelers EVs performed, and the below figure states exactly which of the brands performed the best.




As we earlier find out that among 4 e-vehicle types, the sales and revenue of 2 wheelers are highest across many years, so on that basis we plotted 2 wheelers to find out which of the brands perform optimum in the available segments. we found out that Ola Electric 2 wheelers are highest especially the uptrend from fy21-22 to fy22-23 just skyrocketed contributing 152.567k to the market. following the ola electric, there's Okinawa and Hero Electric also holding 2nd and 3rd place respectively. Amber, TVS, and Ather also follows within the same range.

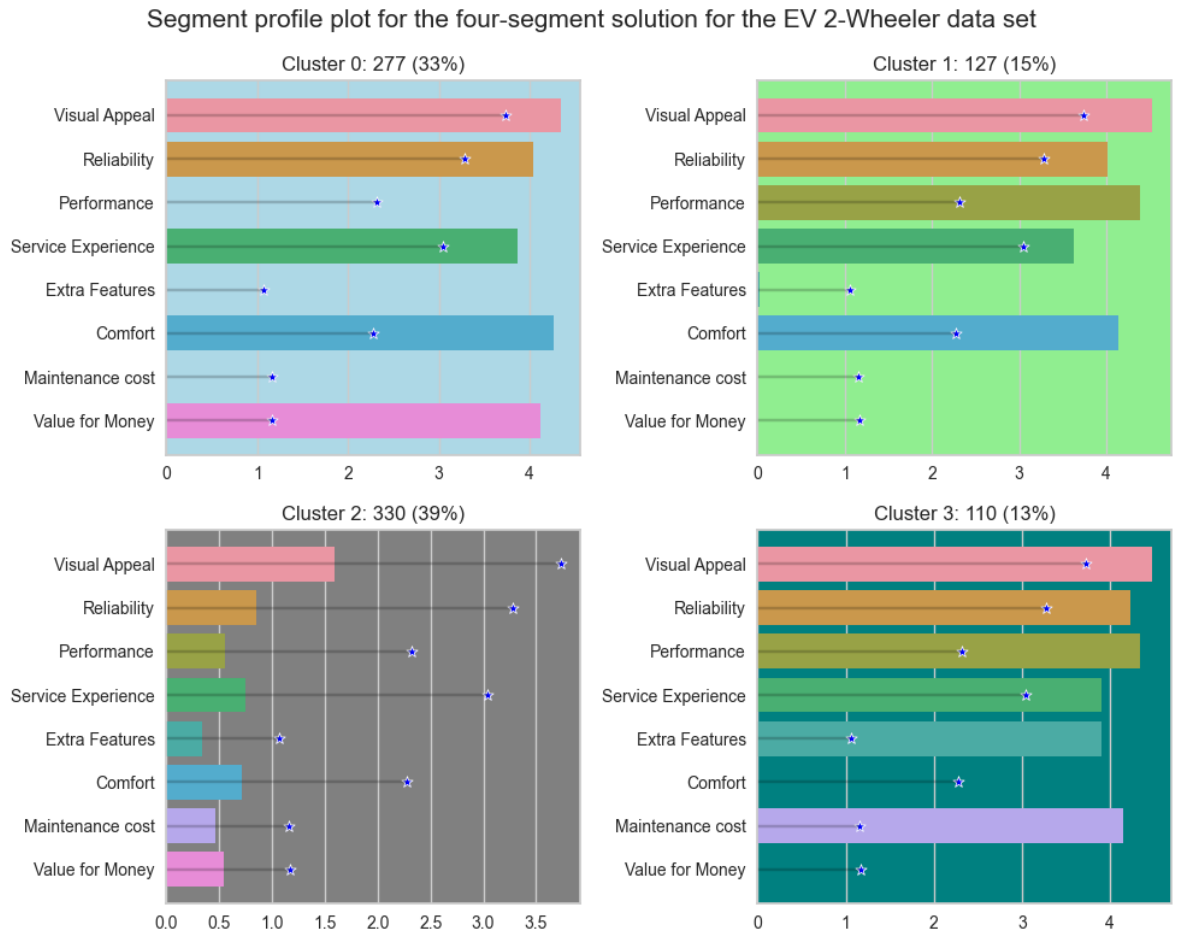
# 4. Profiling and Describing Potential Segments

Profiling and describing potential segments are crucial steps in understanding distinct customer groups within the EV market. Here's an overview of how this was done in the project.

## Profiling Segments:

**Cluster Identification:**

- Method: The project utilized K-Means clustering to segment customers based on their travel behaviours and characteristics from the dataset.
- Clusters Found: Four primary clusters were identified: Cluster 0, Cluster 1, Cluster 2, and Cluster 3.



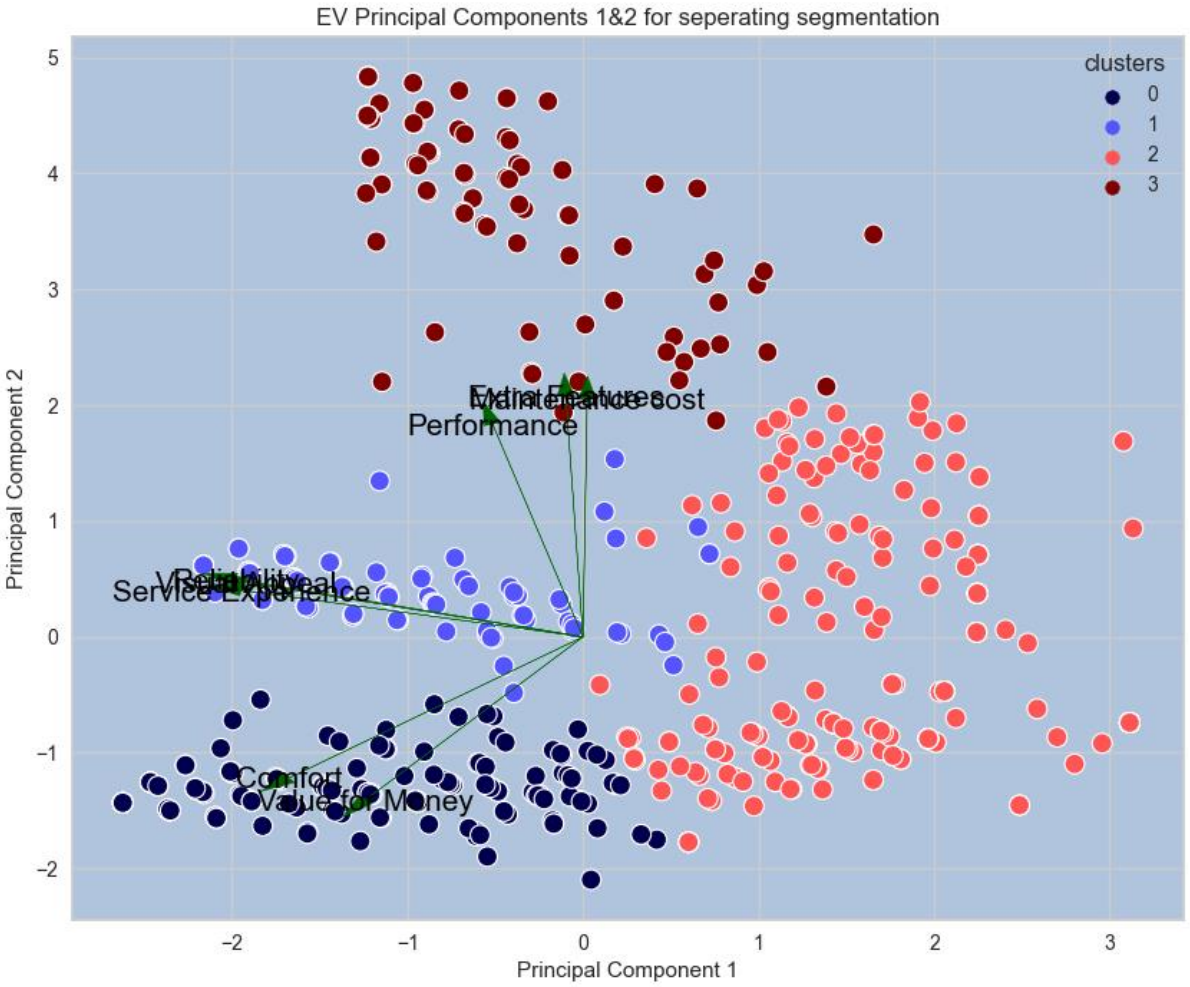

The above graph shows the level of customer's interest from each segment in 8 different features of EVs. It gives us the insight as follows-

- Segment 0 customers which comprises of 33% of the Total customers are highly satisfied with 'Visual appeal', 'Comfort', and 'Value for Monay' features of their 2 Wheelers EVs.

- Segment 1 customers comprises of only 15% customers out of total 2 Wheelers EV users in the market. they are satisfied mainly with 'visual appeal', 'performance', and 'comfort', and not so impressed with 'value for money', 'maintenance cost', and 'extra features'.

- Segment 2 customers are the highest in numbers among all clusters with 39% of the total consumers. They show less interest in all of 8 EV feature segments, seems to be highly dissatisfied with. A possible assumption for this might be that they are 4 Wheelers users for most of the occasions.

- Segment 3 customers comprising 13% the lowest among all segments show highest level of satisfaction in most of the features except for ‘comfort’, and ‘value for money’.




The above figure further differentiates between segments by using Principal Components Analysis. PCA based segment opinions shows Segment 3 is the most visible in opinions, contrary to segment2 (Highest number of Customers) which is showing lack of opinions.

## Describing Segments:

In this section the goal is to describe the customer segments based on their usage, ownership duration, riding distance, sentiments, and technical specifications of 2-wheeler electric vehicles (EVs). The four key categorical columns analysed are namely- "Used it for", "Owned for", "Ridden for", and "Sentiment".

**Used it for:**

- Most customers use their 2-wheeler EVs for daily commutes, with very few using them for tours.

**Owned for:**

- Segment 2 primarily consists of customers who have owned their vehicles for more than a year.
- Segment 1 mostly includes customers who have never owned the vehicle.
- Segment 0 shows a balanced distribution of ownership durations.

**Ridden for:**

- Across all segments, the majority of customers have ridden their EVs for distances within 5000 km.

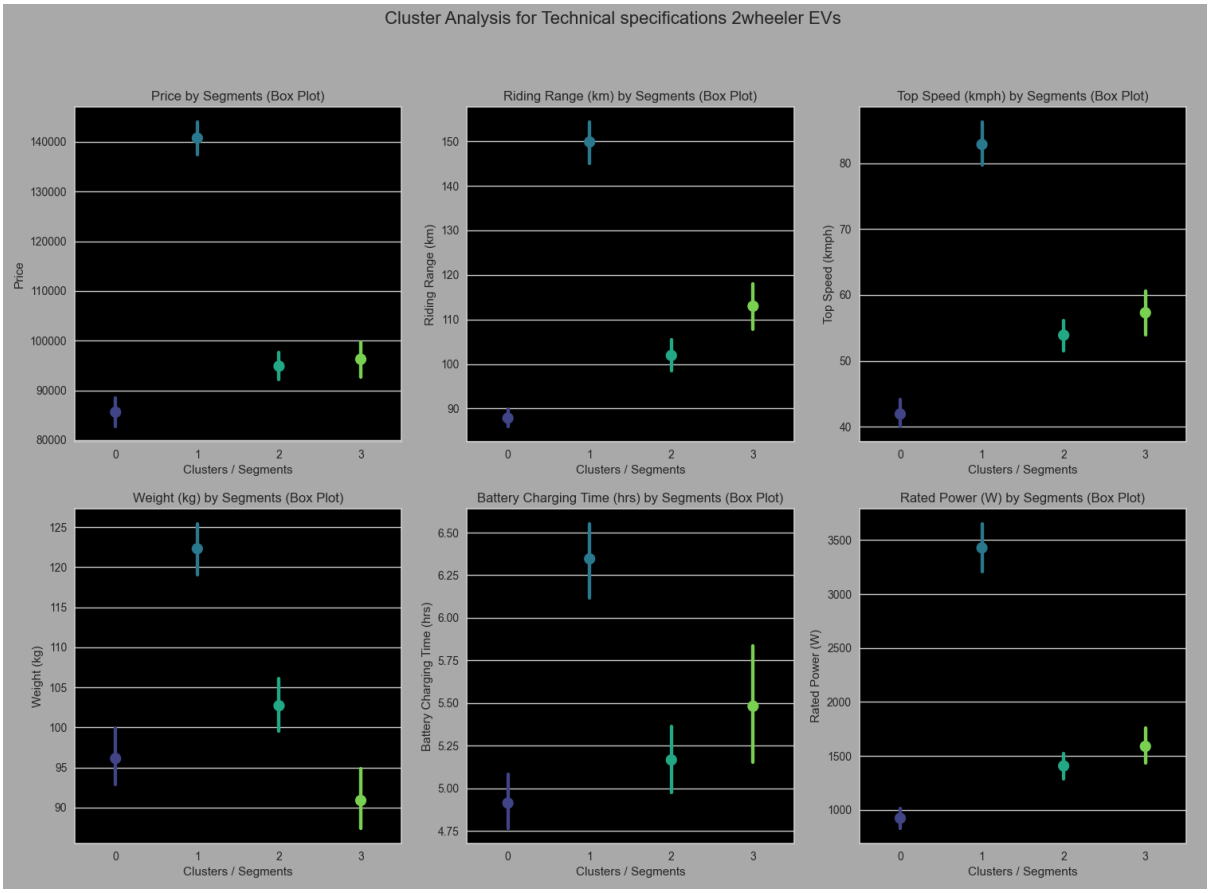
**Sentiments:**

- Segments 0, 1, and 3 exhibit overall positive sentiments towards their 2-wheeler EVs.
- Segment 2 shows a significant portion (62%) of negative sentiments, indicating dissatisfaction.

**Technical Specifications:**

- Segment 0: Prefers budget friendly EVs with lower specs.
- Segment 1: Prefers high-end EVs with superior performance and luxury, and they ride the longest distances.
- Segment 2: Satisfied with average specs in their EVs.
- Segment 3: Prioritizes battery charging time and riding range, similar to the Segment 2 in terms of long-distance travel needs.

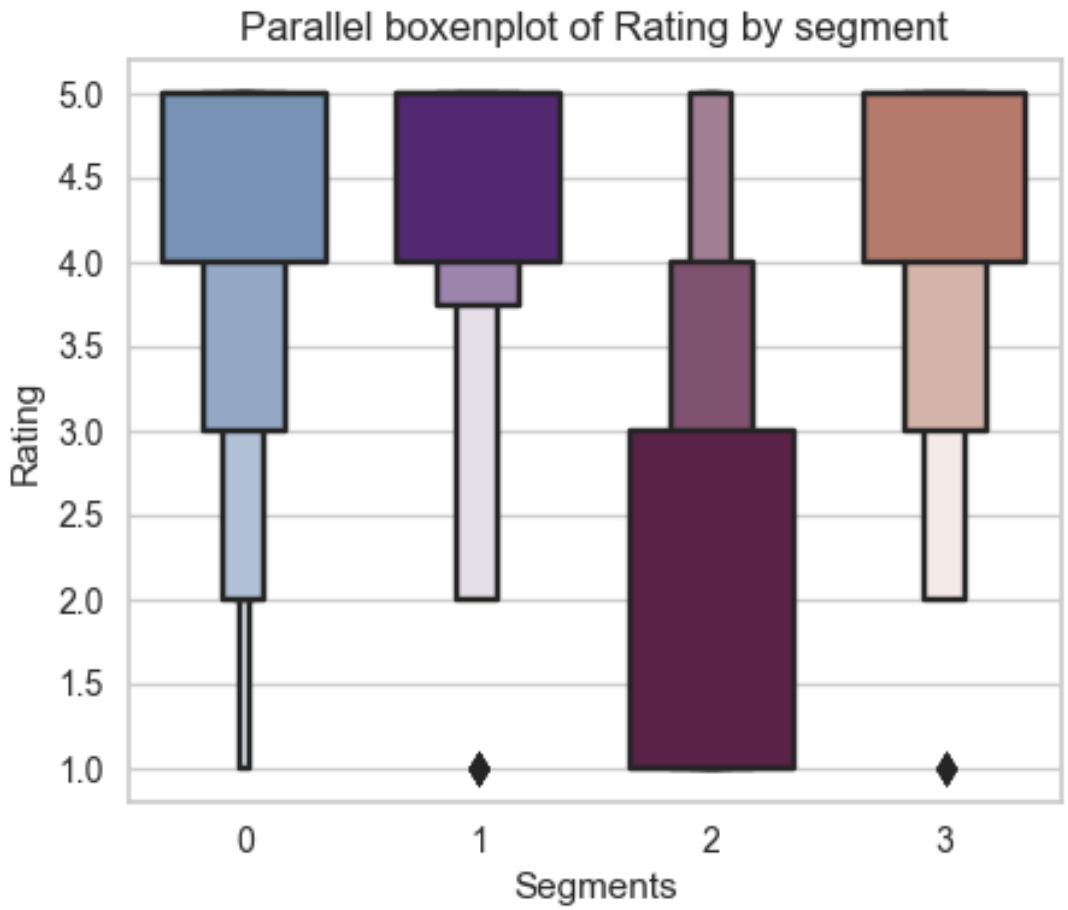
The below chart shows technical specs with respect to each segment of customers.



**Rating Analysis:**

- Segments 0, 1, and 3 generally rate their EVs between 4 and 5, indicating high satisfaction.
- Segment 2 mostly rates between 1 and 3, reflecting their dissatisfaction.

The below boxen plot chart shows the above analysis.




# 5. Selection of Target Segment

In all these analysis and exploration of various Segments I concluded that 2 segments, Segment 1 and Segment 2 are worth focusing on as our target segmentation. If possible, by keeping in mind the limited resources of the startup both of these can be our target segments as segment 1 consists of those customers who are extremely satisfied and happy with the product specs and features and they even need more as they seems to be willing to pay higher price and have taste in luxury, they consist of 15% of the total market customers.

Then 2nd cluster that is segment 2 consist of customers who need to be introduced with the features and specs that can meet their satisfaction level as they are most dissatisfied by majority of variables(specs & features) and keeping in mind they are the largest consumer group in total EV market consisting of almost 40 % of the total market customer so we can't simply overlook this while deciding our target segments.

# 6. Customizing the Marketing Mix

In this project, I've identified distinct customer segments for 2-wheeler electric vehicles (EVs) using cluster analysis. Each segment has unique preferences and behaviours. To effectively target these segments, we need to customize the marketing mix—Product, Price, Place, and Promotion—accordingly. Here's a brief on how to do this:

## Product

**Segment 0:** Budget-conscious customers who prefer lower-spec EVs.

- **Customization:** Focus on affordability and essential features. Highlight reliability and cost-effectiveness. Offer basic models with good value for money.

**Segment 1:** High-end customers seeking luxury and performance.

- **Customization:** Emphasize premium features, superior performance, and advanced technology. Offer top-tier models with high-end specs and consider exclusive customizations.

**Segment 2:** Customers satisfied with average specs.

- **Customization**: Provide balanced models with a good mix of features and affordability. Ensure these EVs are versatile for general use without focusing too much on either end of the spec spectrum.

**Segment 3:** Customers prioritizing battery life and riding range.

- **Customization:** Highlight long battery life and extended range. Offer models with efficient battery management and fast charging capabilities.

## Price

**Segment 0:** Price-sensitive.

- **Strategy:** Use competitive pricing, discounts, and financing options. Highlight the long-term savings from using EVs.

**Segment 1:** Less price-sensitive, value high performance and luxury.

- **Strategy:** Premium pricing strategy. Offer financing plans and warranty packages to emphasize quality and exclusivity.

**Segment 2:** Looking for value.

- **Strategy:** Mid-range pricing. Bundle features to provide perceived added value, and consider loyalty programs or trade-in options.

**Segment 3:** Willing to pay for efficiency.

- **Strategy:** Pricing that reflects the advanced battery technology and longer range. Emphasize cost savings from reduced charging frequency and higher efficiency.

## Place

In case of Place customization we will focus on 2 aspects, namely-

**Convenience oriented:** Focus on online sales and urban retail locations. Ensure easy access to service centres and parts.

**Premium shopping experiences oriented:** Sell through exclusive dealerships and flagship stores. Provide personalized shopping experiences and at-home test drives.

### Promotion

**Segment 0:** Cost-conscious messaging.

- **Strategy:** Highlight cost savings, reliability, and efficiency. Use social media, online reviews, and word-of-mouth campaigns.

**Segment 1:** Luxury and performance-focused messaging.

- **Strategy:** Use high-end advertising, influencer partnerships, and experiential marketing events. Emphasize performance, innovation, and exclusivity.

**Segment 2:** Balanced approach.

- **Strategy:** Use testimonials, comparative ads showing value for money, and emphasize versatility. Utilize a mix of digital and traditional media.

**Segment 3:** Efficiency and convenience-focused messaging.

- **Strategy:** Highlight the benefits of long battery life and extended range. Use educational content about the benefits of EVs for long-distance travel, and target areas with high travel needs.

By tailoring the marketing mix to each segment's unique preferences and needs, you can enhance customer satisfaction, increase market share, and foster brand loyalty among diverse customer groups.

## 7. Potential Early Market Customer Base

In our analysis, we identified two major segments, Segment 1 and Segment 2. Segment 1 comprises of 127 customers. Segment 2 comprises of 330 customers.

**1. Segment 1:** High-End Enthusiasts

- **Assumed Segment Size (10 times of the segment size):** 1270 customers.
- **Adjusted Average Target Price:** ₹112,500
- **Potential Profit:** 1270 * ₹112,500 = ₹142,875,000

**2. Segment 2:** Average Spec Seekers

- **Assumed Segment Size (10 times of the segment size):** 3300 customers.
- **Adjusted Average Target Price:** ₹67,500
- **Potential Profit:** 3300 * ₹67,500 = ₹222,750,000

**Total Potential Profit:**

- **Total Potential Customer Base:** 127 (Segment 1) + 330 (Segment 2) = 457 customers
- **Total Potential Profit:** ₹142,875,000 (Segment 1) + ₹222,750,000 (Segment 2) = ₹365,625,000

# 8. Most Optimal Market Segments

After thorough market research and segmentation analysis, **Segment 2** emerges as the most optimal market segment to target for opening in the market. Comprising 39% of the total customer base, segment 2 represents a substantial portion of the market with significant growth potential. Its characteristics align closely with the product or service offerings, indicating a strong fit for the business strategy. By focusing efforts on catering to the specific needs and preferences of this segment, the business can capitalize on its sizable market share and drive sustainable growth and profitability.

These are the features that our target Segment 2 shows high interest into.

Price Range ---------- 71,000 – 1,30,000

Battery Charging Time -------- 3.5 – 5 hrs.

Top Speed ---------- 60 – 120 Kmph

Rated Power ---------- 1200 – 5500 W

Riding Range ---------- 90 – 185 Km

Weight ---------- 72 – 125 Kg

# Conclusion

In conclusion, targeting segment 2, which represents 39% of the total customer base, offers a strategic advantage for market entry. Its significant size and alignment with our offerings present an excellent opportunity for growth and success. By focusing on this optimal segment, we can effectively meet customer needs and establish a strong market presence, ensuring long-term profitability and business sustainability.