

Market Segmentation

Electric Vehicle Market Analysis

Kiran Thomas V

(Team Sandeep)

Fermi Estimation

Problem statement:

You are a team working under an Electric Vehicle Startup. The Startup is still deciding in which vehicle/customer space it will develop its EVs.

You have to analyse the Electric Vehicle market in India using Segmentation analysis and come up with a feasible strategy to enter the market, targeting the segments most likely to use Electric vehicles.

(CUSTOMER/VEHICLE/B2B) SEGMENTS: Apart from Geographic, Demographic, Psychographic, Behavioural segments, teams can consider different CATEGORY of Segments for the Segmentation Tasks, based on AVAILABILITY OF DATA. Market Segmentation comes with wide scope of possibility and Segments created can change based on different datasets collected.

- Population of India: As of 2024, India's population is over 1.4 billion people.
- Percentage of population with access to vehicles: Around 36% of the population in India has access to some form of personal transportation, whether it's a car, motorcycle, or bicycle
- Electric Vehicle (EV) adoption rate: The adoption rate of EVs in India is relatively low compared to traditional vehicles. Let's estimate it to be around 1% of the vehicles on the road.
- Types of vehicles: Considering the market, we have various types of vehicles - cars, motorcycles, scooters, buses, and trucks.
- Geographic distribution: EV adoption might be higher in urban areas due to better infrastructure and awareness. Let's estimate around 70% of the EVs being used in urban areas.
- Demographic segments: Young professionals in urban areas might be more inclined towards EVs due to environmental concerns and technological interest. Let's estimate this segment to be around 30% of the urban population.
- Psychographic segments: Environmentally conscious individuals or those concerned about air pollution might be more inclined towards EVs. Let's estimate this segment to be around 20% of the urban population.
- Behavioral segments: People who have previously owned hybrid vehicles or have shown interest in green initiatives might be more likely to adopt EVs. Let's estimate this segment to be around 15% of the urban population.
- B2B segments: Companies and organizations might be interested in fleet electrification for cost savings and sustainability. Let's estimate this segment to be around 10% of the overall EV market.

With these estimations, we can now focus our strategy on targeting urban areas, particularly young professionals and environmentally conscious individuals, while also considering B2B opportunities for fleet electrification. This would involve developing EV models that cater to

urban commuting needs, with features that appeal to these target segments such as affordability, range, and charging infrastructure support. Additionally, partnerships with businesses for fleet electrification could be explored to tap into the B2B segment.

Data Sources

mega_cars.csv: This dataset is used for the cluster analysis of electric cars in India. This dataset was obtained from a git hub repository. The link to the repo is:

<https://github.com/shiv2110/EV-Market-Segmentation-India.git>

EV_india.csv: This dataset is used for the geographic market analysis of electric vehicles. This dataset was obtained from a Kaggle project. It contains the number of EV vehicles that is currently being used in each state, also the number of non-EV vehicles that are in use in the same states. This is used to find out the percentages in each state to understand which state can be a good market for new EV vehicles.

electric_bikes_f.csv: This dataset is obtained from a website called zigwheels.com (<https://www.zigwheels.com/newbikes/best-electric-electric-two-wheelers-bikes>). This dataset was web scrapped from the above given website and then cleaned to get another csv filed called EV_bike_data.csv. This file contains the EV two wheelers that are available in market and a few of their features. This was used for the cluster analysis of electric two wheelers in India to understand the market of EV two wheelers.

Data Pre-processing

In the code for analysing the market potential in each state the data goes through some basic pre-processing steps.

- First the null vales are dropped/ removed from the dataset.
- Then we find the percentage of EV vehicles and non-EV vehicles. This is done by using the formulae:
 - $\text{No. of EV vehicles} / \text{total no. of vehicles}$
 - $\text{No. of non-EV vehicles} / \text{total no. of vehicles}$

In the code for segmentation of EV cars the data goes through similar basic pre-processing steps.

- The unit (L) in the bootspace column was removed by subscripting each data in the same column. The data type of the same was changed from string to int.
- The Drive type column contained string data. We need numerical data so we performed one hot encoding on that column and split it into three different fields.

- Then a confusion matrix between the different features was plotted which led to the discovery that the features were highly correlated, which led to the conclusion we can use less features for the ml model. Therefore, dimensionality reduction (PCA) was performed on the data.

In the code for segmentation of EV two wheelers the data goes through a lot of pre-processing steps. As the data was scrapped from the internet there were many inconsistencies with the data.

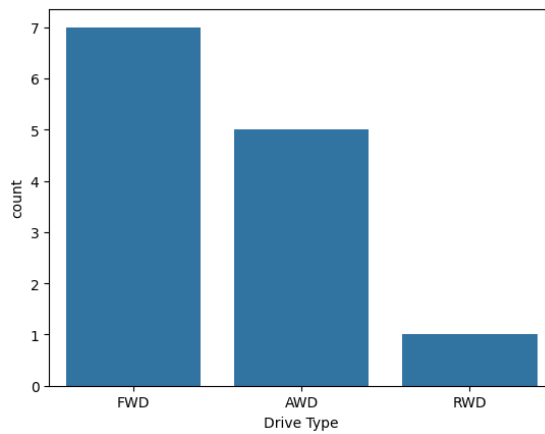
- The Kwh, Range and Max speed was extracted as one text. This one field was split into three using pandas.
- The price had different types of values in the field. Some were in lakhs; some were in thousands and some were even in ranges. Therefore, normalization was required.
First the word (Lakh) was removed from each data field. Then in fields where the price was in a range the maximum value was taken. Then the fields where the price was in lakhs was multiplied by 1000000 so every price was in thousands.
- The kwh data field had kwh as a unit that was removed. And any data that wasn't kwh was replaced by null values.
- The range field had other values in some of the rows, these were replaced with null values and the km/range unit was removed.
- All this data was converted to float.
- All the null values were dropped.

Segment Extraction

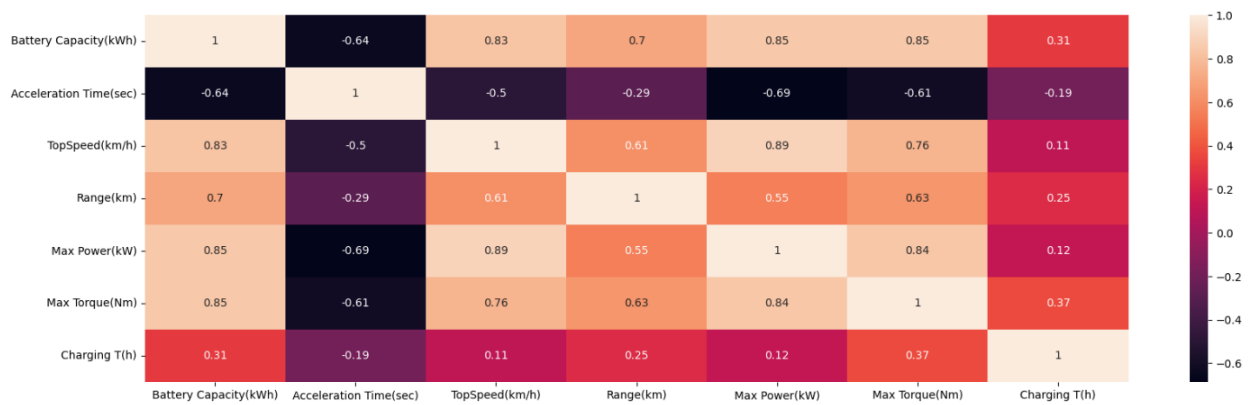
Segmentation of four wheelers:

The data was collected. The dataset used is called mega_cars.csv which contained information about the EV cars that are available in India. Some EDA was done in which we found the following:

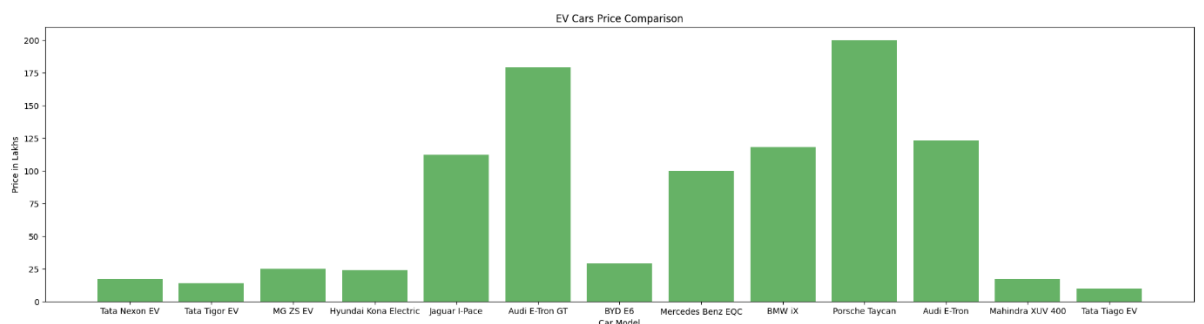
Most of the cars available are Four Wheel Drives.



The data is highly correlated, which means we can use less features for the ml model.

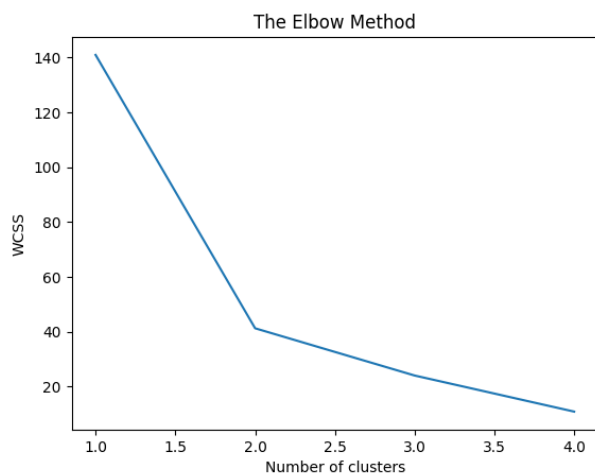


A graph for price comparison was also plotted



As we found earlier that the data was highly correlated. The data underwent dimensionality reduction. PCA (principal component analysis) done on the dataset. PCA component was decided to be 3.

Then the k value for k-means clustering was decided by the Elbow method (WCSS)



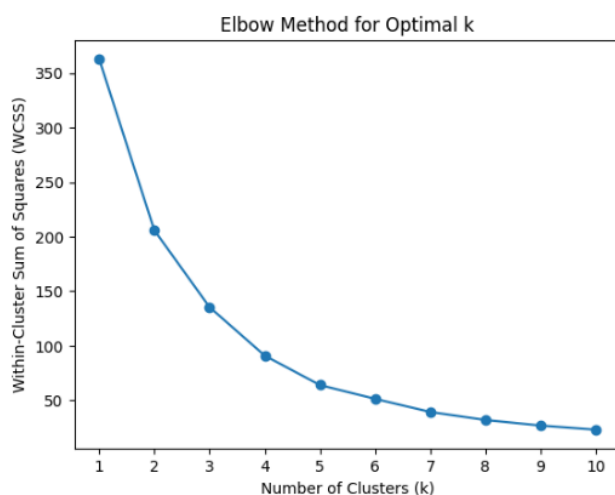
The elbow method gave us a k of 2. Therefore, k-means clustering was done by setting the hyper parameter ($k=2$). Then we got these segments.

Segmentation of two wheelers:

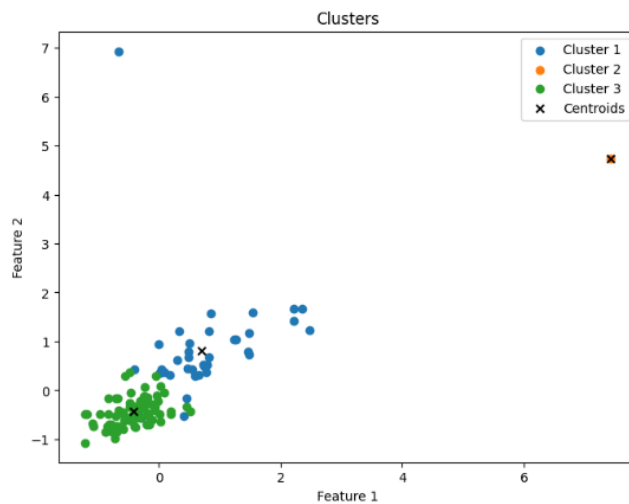
The data was collected by scrapping it from a website called zigwheels. The dataset used is called electric_bikes_f.csv which contained information about the EV two wheelers that are available in India.

As there were only two features in the dataset it was decided that there is no need for any dimensionality reduction.

The k value for k-means clustering was decided by the Elbow method (WCSS)



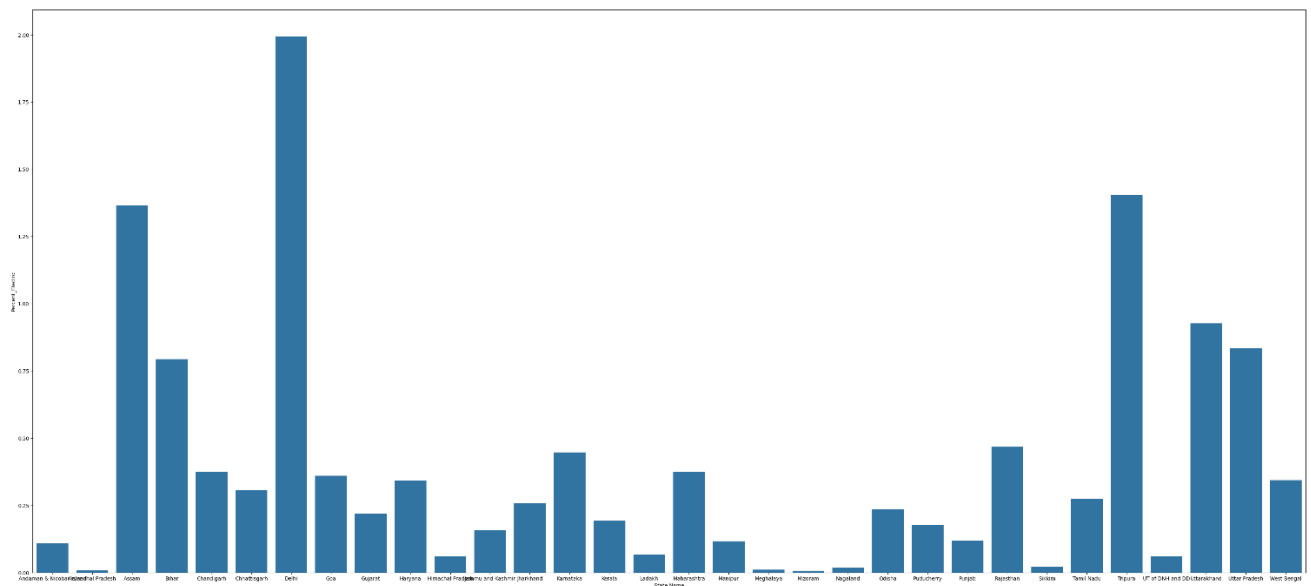
The elbow method gave us a k of 3. Therefore, k-means clustering was done by setting the hyper parameter ($k=3$). Then we got these segments.

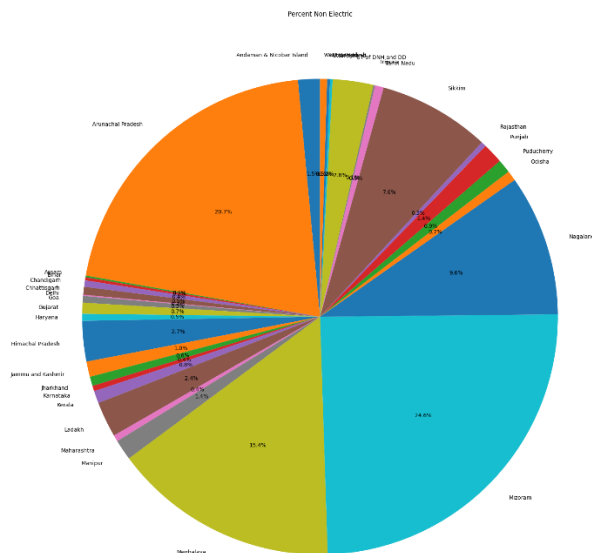


As we can see there are a few outliers and one of them don't belong to both the segments therefore it is its own cluster.

Market Analysis (by state):

The data was collected. Few pre processing steps were done. The percentages were calculated for each state. Then the graph was plotted.





As we can see from the graphs there are states where there are a lot of EVs already in use. These states can be potential markets as they have received EV vehicles well. The one drawback of these segments is that there is a lot of competition in these segments already.

There are segments where the no. of EVs in use are very less. These states can be untapped markets where the company could deploy EV models suitable to the demographic in the given area. The benefit is that there is less to no competition in these states. The drawback is that we don't know how well EVs will be received in these states.

Then the states where there is a decent amount of EVs in use. These are middle grounds in the market where there is decent competition and where we know there are some EVs in use. This can be used to our advantage and develop our product by understanding what is lacking in the EVs already in use in these states and implementing that into our design. This is a safe market segment and can be used to our advantage if more in-depth analysis is done of that segment alone.

