Electric Vehicle market in India

Date of Submission- 18-07-2022

Contributors: -

Ishu Khandelwal

Lakshya Pokharna

Sanchit Agarkar

Ajinkya Mahure

Harsh Makwana

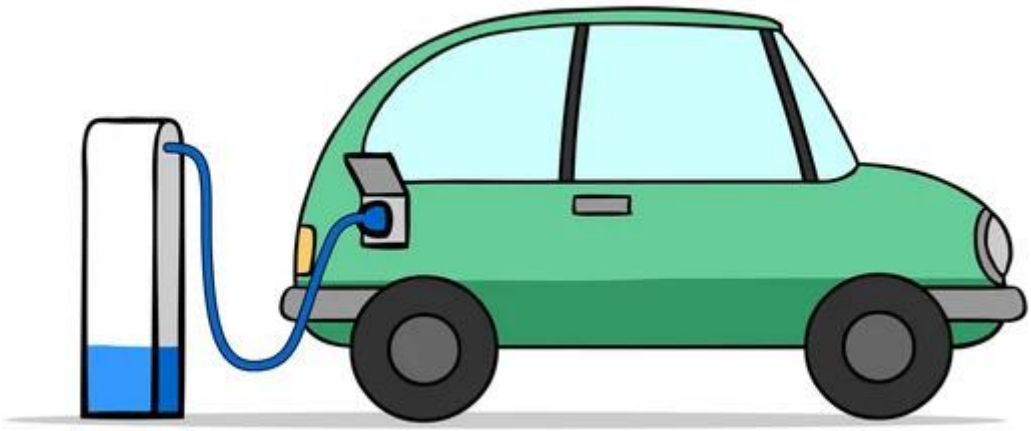


Table of Contents

1. Introduction

Market segmentation has become a vital strategy for emerging markets to investigate and execute for widespread adoption of emerging mobility technologies like electric vehicles (EVs). Due to their tremendous growth in popularity as low-emission and low-cost cars, EVs are projected to spark a great deal of interest in academic study in the near future. By utilizing an integrated research framework of "perceived benefits-attitude-intention," the primary goal of this study is to examine and identify several sets of possible buyer groups for EVs based on psychographic, behavioral, and socioeconomic characterization. In order to operationalize and validate segments from the data obtained from respondents via a cross-sectional online survey, the study used rigorous analytical techniques such as cluster analysis, multiple discriminant analysis, and the Chi-square test.

In this paper, our motto is to obtain information on the present electrical vehicle market in India using business analytics and market segmentation techniques. A detailed analysis of the various aspects of Indian EVs and their customer base is carried out through machine learning. The case study focuses on the clustering pipeline using limited amounts of data obtained from several trusted platforms, including government open source.

2. Problem Statement

The task is to analyze the electric vehicle market in India using segmentation analysis and come up with a feasible strategy to enter the market. In this report, our team has tried to analyze the electric vehicle market in India using segments such as price, top speed, range, safety, battery capacity, fuel types, fast charging, boot space, and much more.

3. Data Collection

Data was scraped from the website https://e-amrit.niti.gov.in/home. This website is the e-AMRIT (Accelerated e-Mobility Revolution for India's Transportation) portal for creating awareness about electric mobility in India. We gathered some specifications of electric vehicles from Kaggle. From the above links, we gathered datasets and merged them to create the final data. We have also taken the data from the survey we have done regarding electric vehicles. The data is partly used for visualization purposes and partly for clustering.

E-AMRIT -> https://e-amrit.niti.gov.in/home

Survey form link -> https://forms.gle/jJJeV1tTuAALscFP7

Library Used

1. Numpy: To calculate various calculations related to arrays.
2. Pandas: To read or load the datasets.
3. Seaborn: It is used for data visualization and exploratory data analysis.
4. Matplotlib: It is a data visualization and graphical plotting library for Python.

3.1 Data Pre-processing

[1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

import warnings
warnings.filterwarnings('ignore')
```

[2]:
```
data=pd.read_excel(r"C:\Users\Downloads\Final EV data.xlsx")
data.head()
```

[2]:

	Vehicle full name	Manufacturing	Model	Top speed (km/hr)	Price (INR)	Fuel Type	Wheelers type	Battery capacity [kWh]	Full charging time (HR)	Kerb weight (KG)	Range (km/hr)	Fast Charging	Drive Type	Number of Seats	boot space (L)	Number of Airbags	Type of brakes	Max Torque (N-M)	V
0	Revolt RV400	Revolt Motors	RV400	85.0	134000.0	Electric	Two wheeler	4.0	4.5	108.0	150.0	YES	Belt Drive	2	NaN	NaN	Disc	170.0	I
1	Revolt RV300	Revolt Motors	RV300	65.0	94999.0	Electric	Two wheeler	2.7	4.2	101.0	180.0	YES	Hub Drive	2	NaN	NaN	Disc	NaN	I
2	Tork Motors(Kratos)	Tork Motors	Kratos	100.0	192499.0	Electric	Two wheeler	4.0	5.0	NaN	180.0	NO	NaN	2	NaN	NaN	Disc	28.0	I
3	Tork Motors(Kratos R)	Tork Motors	Kratos R	105.0	207499.0	Electric	Two wheeler	4.0	5.0	NaN	180.0	YES	NaN	2	NaN	NaN	Disc	38.0	I
4	Oben Rorr	Kabira Mobility	Oben Rorr	100.0	102999.0	Electric	Two wheeler	4.4	2.0	110.0	200.0	YES	Belt Drive	2	NaN	NaN	Disc	NaN	I

[3]:

	States Name	Number of Electric Charging Station
0	Maharashtra	317
1	Andhra Pradesh	266
2	Tamil Nadu	256
3	Gujarat	228
4	Rajasthan	205

```
[4]: sales=pd.read_excel(r"C:\Users\Downloads\EV_sales.xlsx")
     sales.head()
```

[4]:

	Years	Two Wheeler	Three Wheeler	Four Wheeler
0	Year 2020	152000	140683	168300
1	Year 2021	143837	88378	134821
2	Year 2022	231338	384215	429217

```
[5]: data3=pd.read_csv(r"C:\Users\Downloads\Electric_Vehicle_Survey.csv")
     data3.head()
```

[5]:

	Name	Age	State / UT	Approx. Annual Income	Area of Living	Which vehicles do you prefer?	Are you willing to switch from Combustions Vehicles to EV?	What kind of Electric Vehicle do you want?	Budget
0	Sachin Bhandari	45	Gujarat	> 20 Lakh	Urban	Electric Vehicles	Yes	4 Wheeler	3000000
1	Lakshya Pokharna	21	Rajasthan	< 5 Lakh	Urban	Electric Vehicles	Yes	2 Wheeler	60000
2	Sunil Pokharna	53	Rajasthan	5 - 10 Lakh	Urban	Internal Combustion Vehicles	No	4 Wheeler	800000

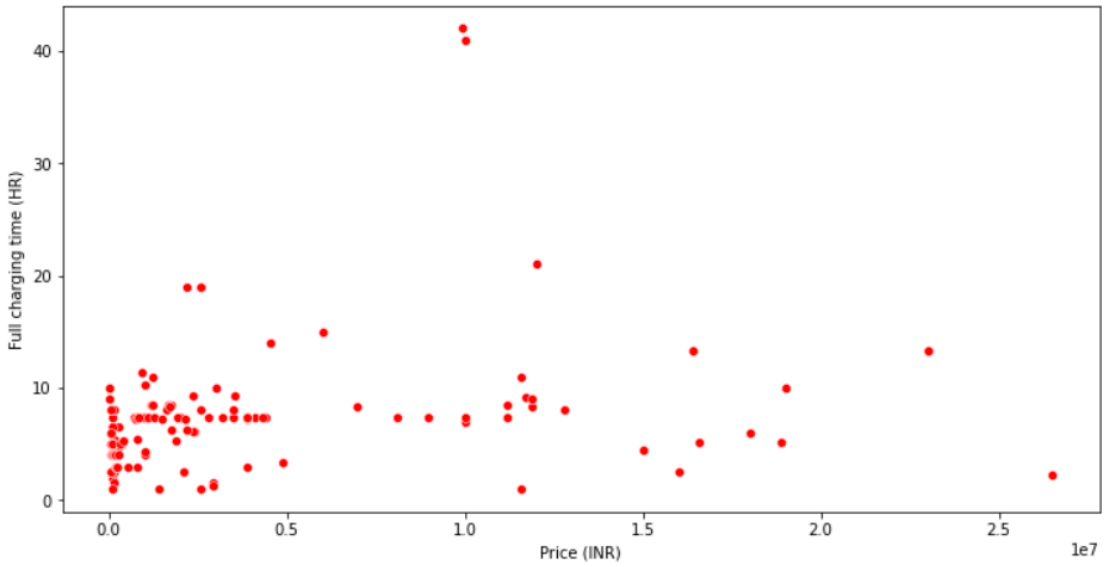
```
[21]: sales.head()
```

[21]:

	Years	Two Wheeler	Three Wheeler	Four Wheeler
0	Year 2020	152000	140683	168300
1	Year 2021	143837	88378	134821
2	Year 2022	231338	384215	429217

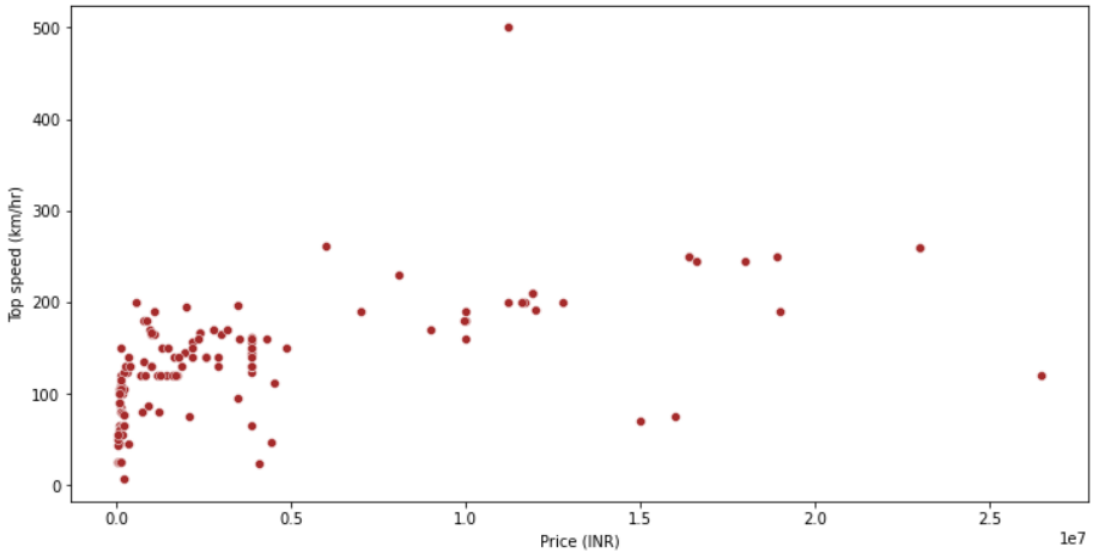
```
[22]: plt.figure(figsize=(12,6))
      sns.scatterplot(x='Price (INR)',y='Full charging time (HR)',data=data, color='red')
```

[22]: <AxesSubplot:xlabel='Price (INR)', ylabel='Full charging time (HR)'>



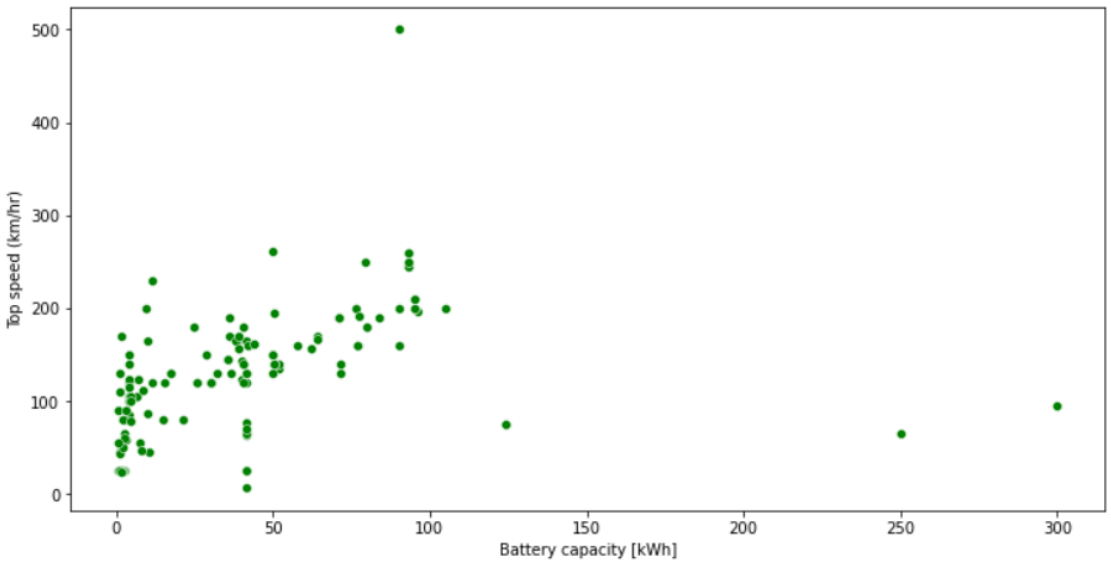
[23]:
```
plt.figure(figsize=(12,6))
sns.scatterplot(x='Price (INR)',y='Top speed (km/hr)',data=new_data, color='brown')
```

[23]: <AxesSubplot:xlabel='Price (INR)', ylabel='Top speed (km/hr)'>



[24]:
```
plt.figure(figsize=(12,6))
sns.scatterplot(x='Battery capacity [kWh]',y='Top speed (km/hr)',data=new_data, color='green')
```

[24]: <AxesSubplot:xlabel='Battery capacity [kWh]', ylabel='Top speed (km/hr)'>



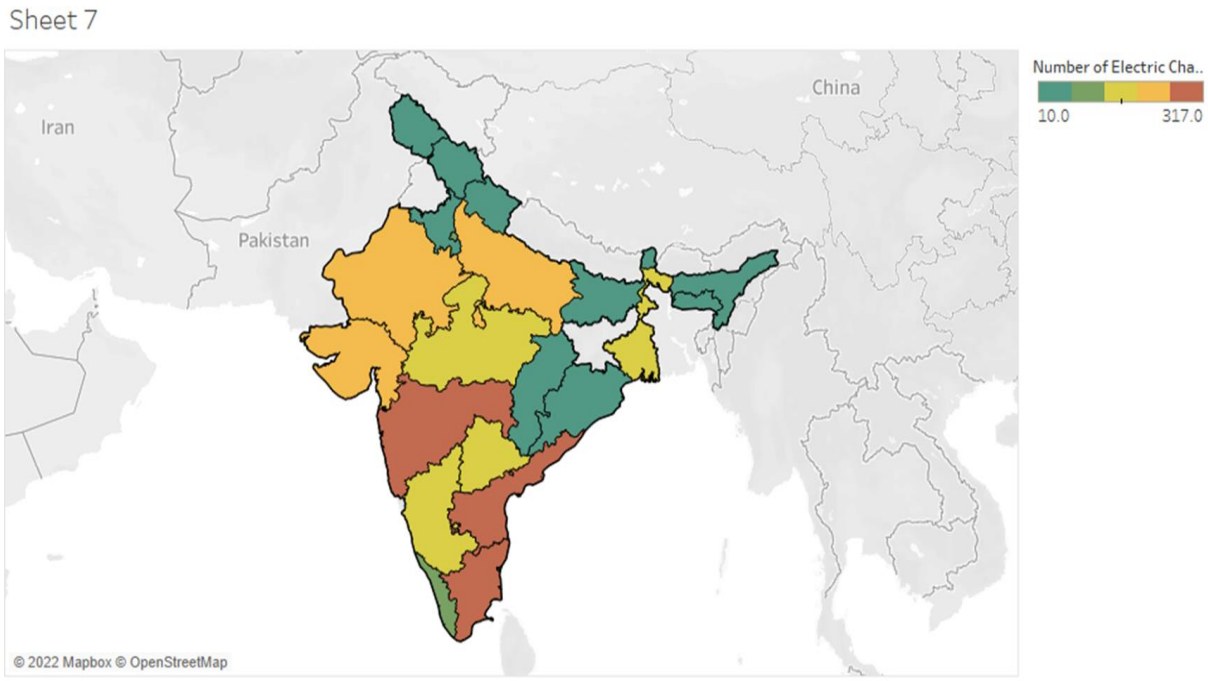
[25]:
```
from sklearn.preprocessing import LabelEncoder

features =['Wheelers type', ' Drive Type', 'Type of brakes','Fast Charging','Income','Fuel Type' ]

for i in features:
    new_data[i] =LabelEncoder().fit_transform(new_data[i])
new_data
```

[25]:

	Top speed (km/hr)	Price (INR)	Full charging time (HR)	Fuel Type	Battery capacity [kWh]	Range (km/hr)	Kerb weight (KG)	Fast Charging	Drive Type	Wheelers type	Number of Seats	Type of brakes	Max Torque (N-M)	Income
0	85.00000	1.340000e+05	4.500000	0	4.000000	150.000000	108.000000	1	11	3	2	1	170.00000	2
1	65.00000	9.499900e+04	4.200000	0	2.700000	180.000000	101.000000	1	15	3	2	1	346.74958	1
2	100.00000	1.924990e+05	5.000000	0	4.000000	180.000000	1506.382114	0	14	3	2	1	28.00000	2
3	105.00000	2.074990e+05	5.000000	0	4.000000	180.000000	1506.382114	1	14	3	2	1	38.00000	2
4	100.00000	1.029990e+05	2.000000	0	4.400000	200.000000	110.000000	1	11	3	2	1	346.74958	2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
138	65.00000	3.893761e+06	3.000000	0	250.000000	200.000000	1506.382114	1	14	0	31	2	346.74958	0
139	75.00000	1.600000e+07	2.500000	0	124.000000	150.000000	1506.382114	1	14	0	31	4	3000.00000	0
140	70.00000	1.500000e+07	4.500000	0	41.355385	300.000000	1506.382114	1	14	0	39	2	800.00000	0
141	129.76259	3.893761e+06	7.344911	0	41.355385	293.126929	1506.382114	1	14	0	43	2	346.74958	0
142	129.76259	3.893761e+06	7.344911	0	41.355385	293.126929	1506.382114	1	14	0	35	2	346.74958	0

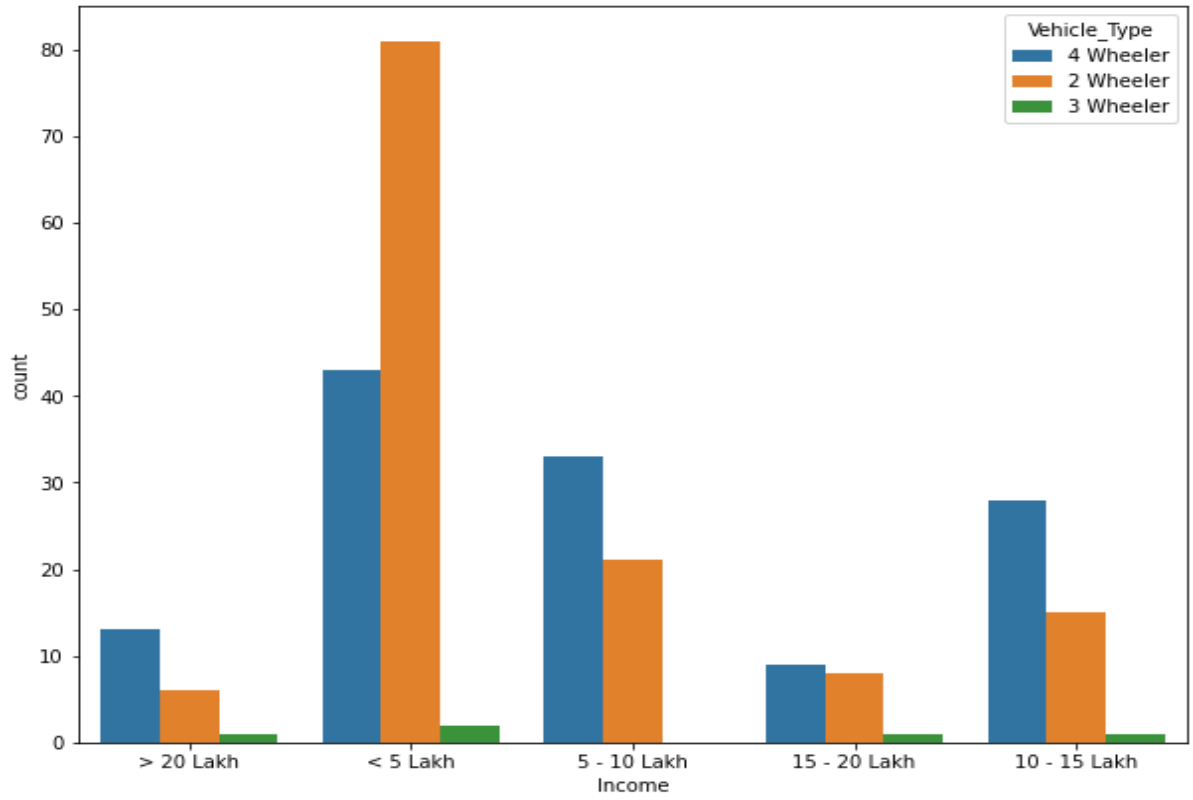
143 rows × 14 columns

Electric charging stations density by state:

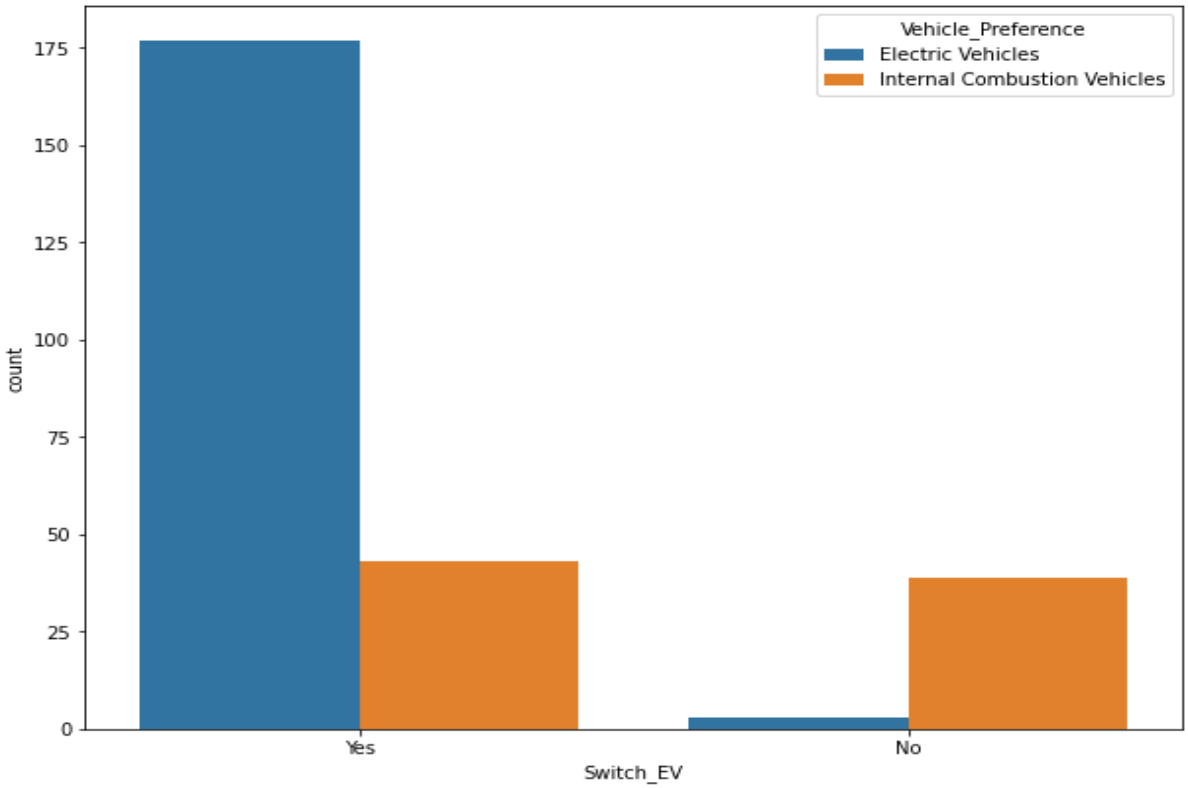


Map based on Longitude (generated) and Latitude (generated). Color shows sum of Number of Electric Charging Station. Details are shown for States Name.

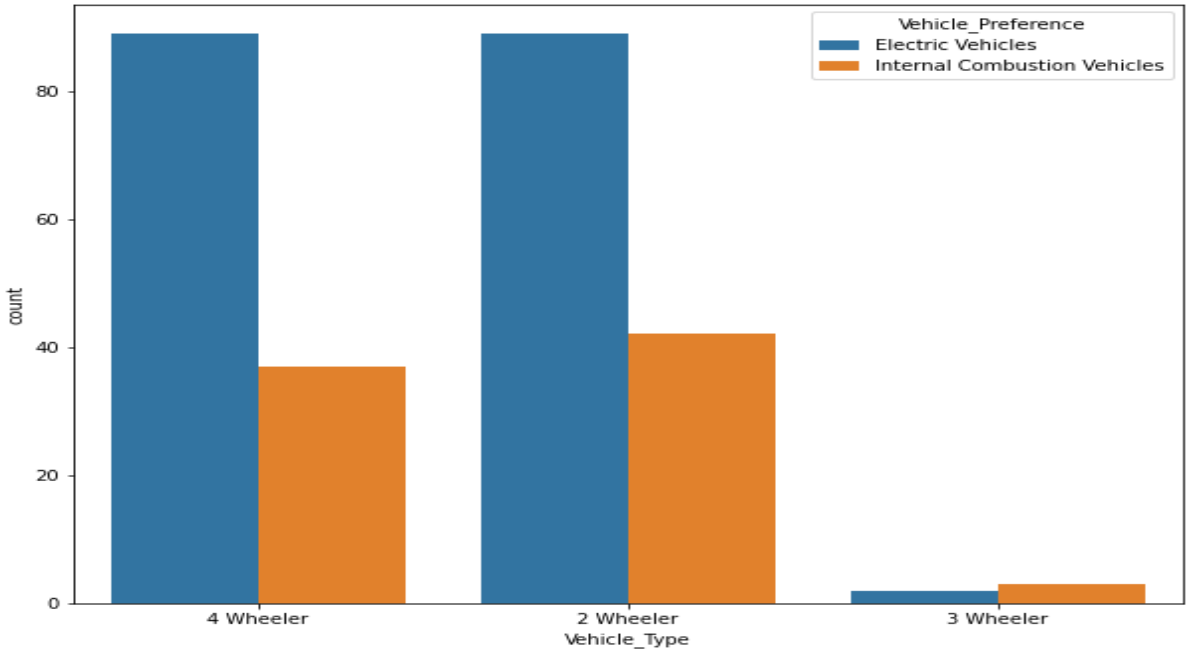
Vehicle Preferences:

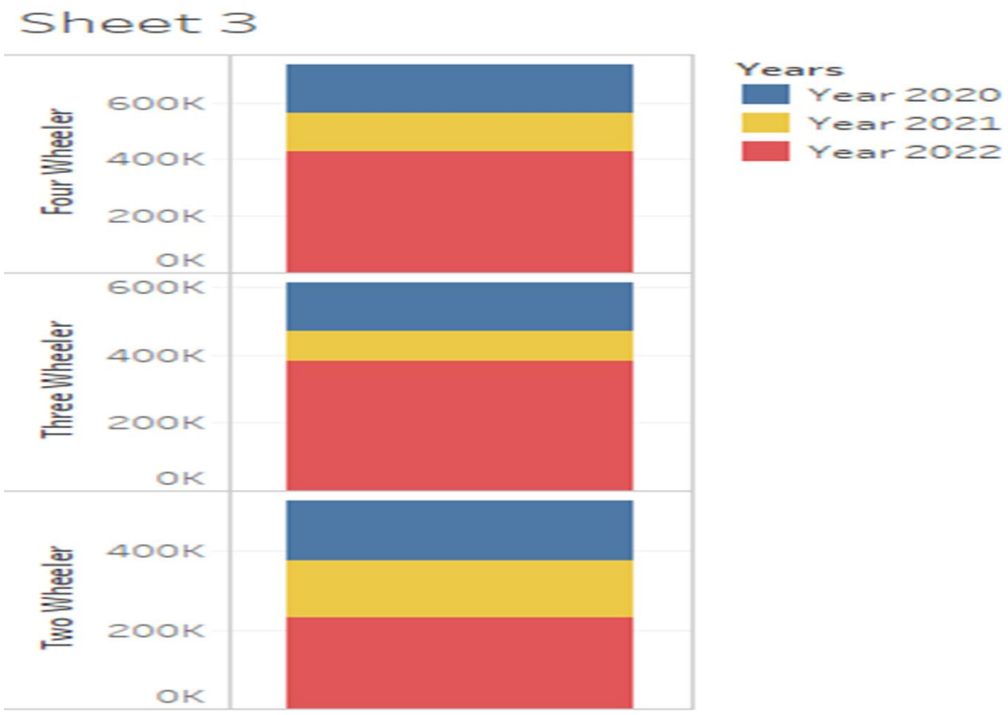


Here we observed that lower-income people preferred 2-wheelers and higher-income group preferred 4-wheelers.

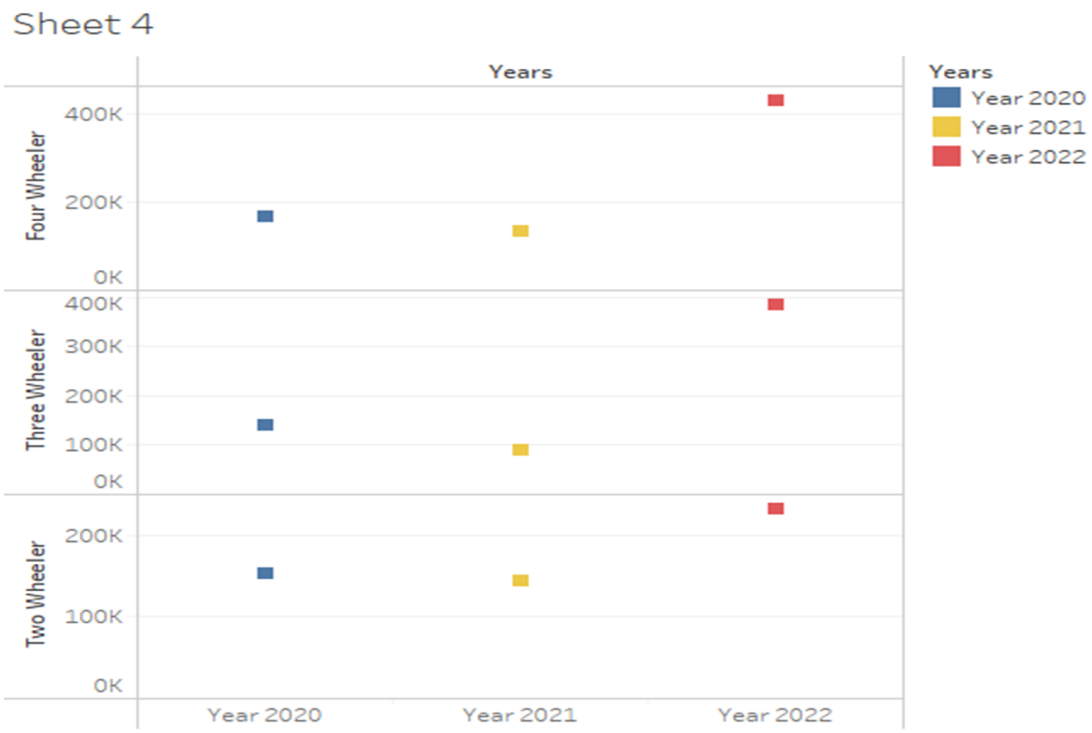


Here we see that most people prefer to switch to electric vehicles from internal combustion vehicles.



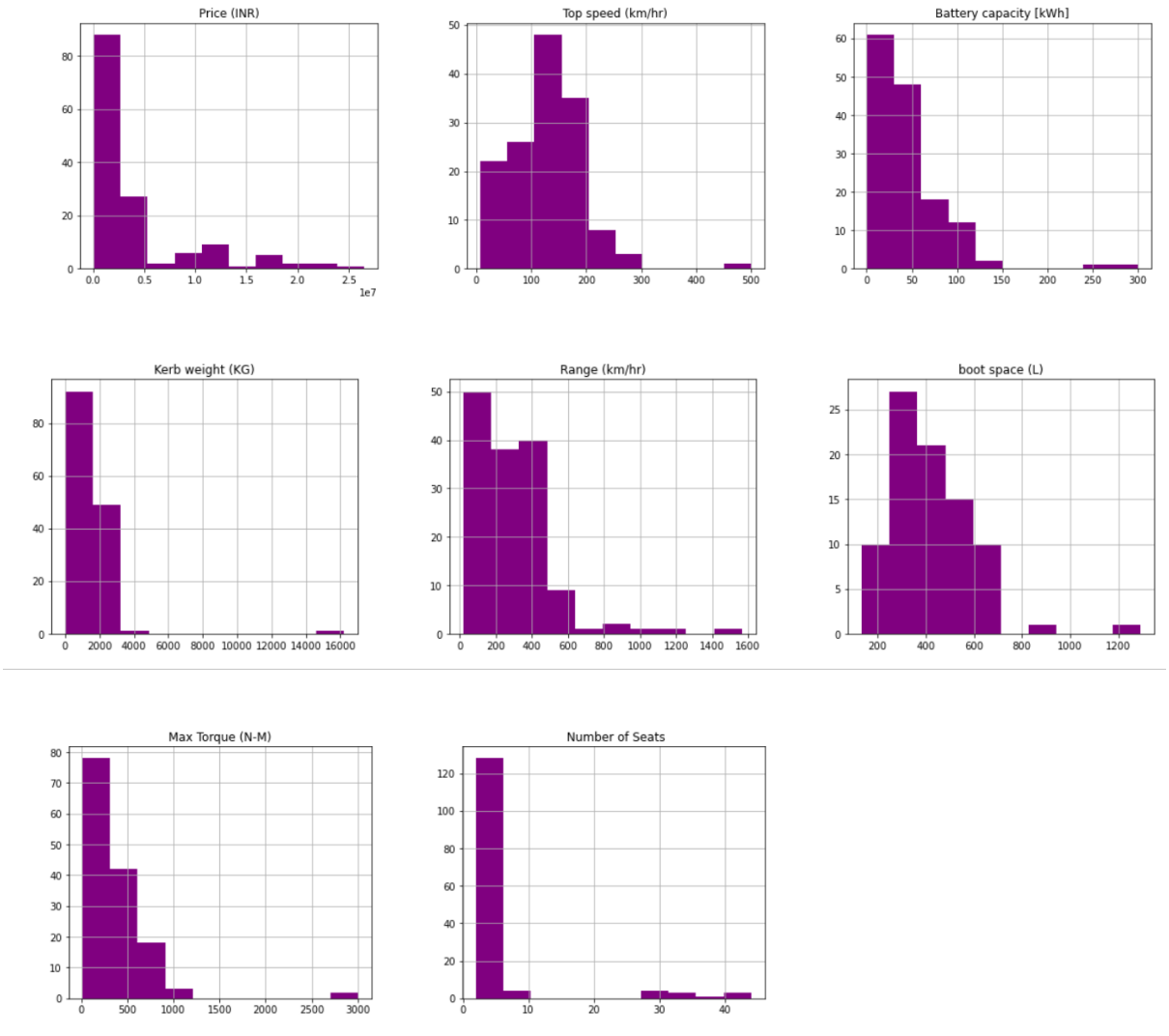


Sum of Four Wheeler, sum of Three Wheeler and sum of Two Wheeler. Color shows details about Years.



Sum of Four Wheeler, sum of Three Wheeler and sum of Two Wheeler for each Years. Color shows details about Years.

The figure above shows the proportion of two-wheeler, three-wheeler, and four-wheeler sales in the years 2020,2021, and 2022 respectively.



4. Segment Extraction

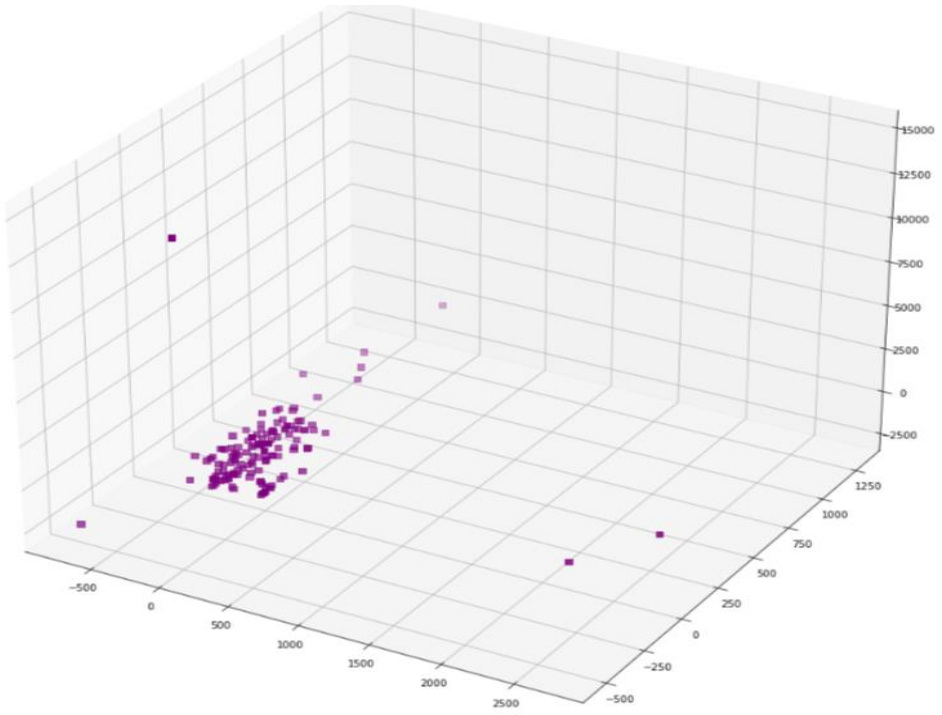
Clustering

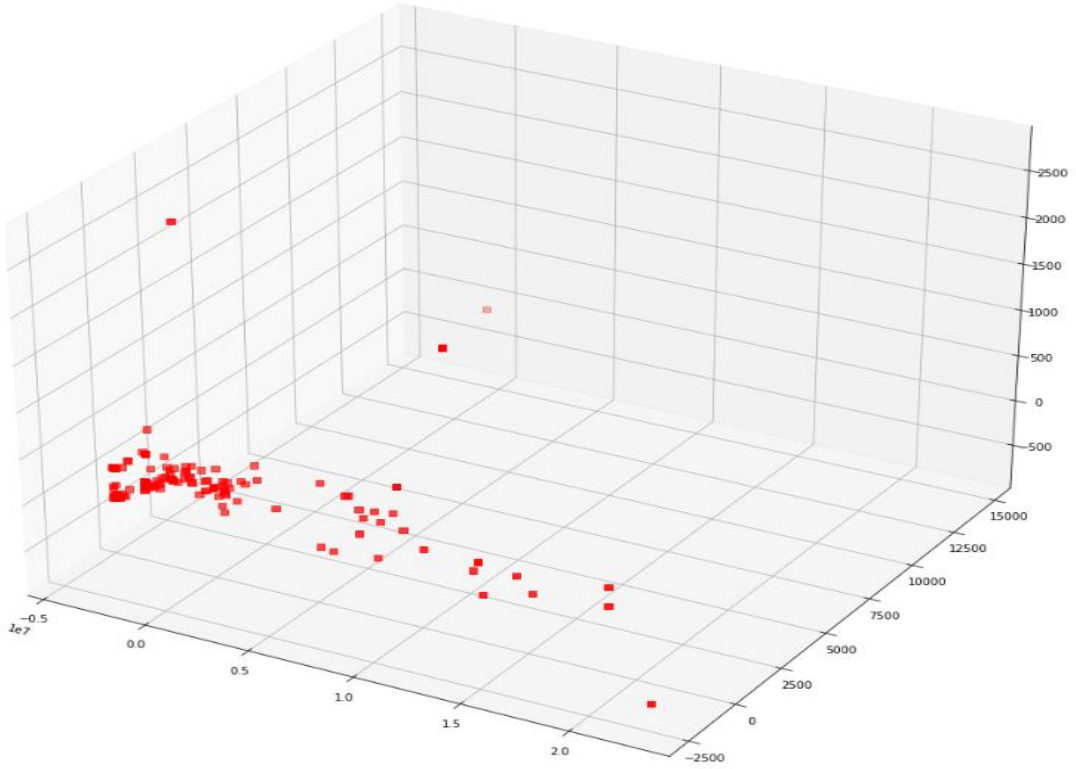
Clustering is one of the most common exploratory data analysis techniques used to get an intuition about the structure of the data. It can be defined as the task of identifying subgroups in the data such that data points in the same subgroup (cluster) are very similar while data points in different clusters are very different. In other words, we try to find homogeneous subgroups within the data such that data points in each cluster are as similar as possible according to a similarity measure such as euclidean based distance or correlation-based

distance. The decision on which similarity measure to use is application-specific. Clustering analysis can be done on the basis of features, where we try to find subgroups of samples based on features, or on the basis of samples, where we try to find subgroups of features based on samples.

PCA(Principal Component Analysis):

We have used Principal Component Analysis and have decreased the components and variability from 13 to 5 in the Final EV Data and then we have made 3D plotting and K-Means clusters.





K-Means Algorithm

K Means algorithm is an iterative algorithm that tries to partition the dataset into pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different (far) as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

The way k means algorithm works is as follows:

- Specify number of clusters K.

- Initialize centroids by first shuffling the dataset and then randomly selecting K data points for the centroids without replacement.
- Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.

The approach that k-means follows to solve the problem is **expectation maximization**.

The E-step is assigning the data points to the closest cluster. The M-step is computing the centroid of each cluster.

Below is a breakdown of how we can solve it mathematically,

The objective function is:

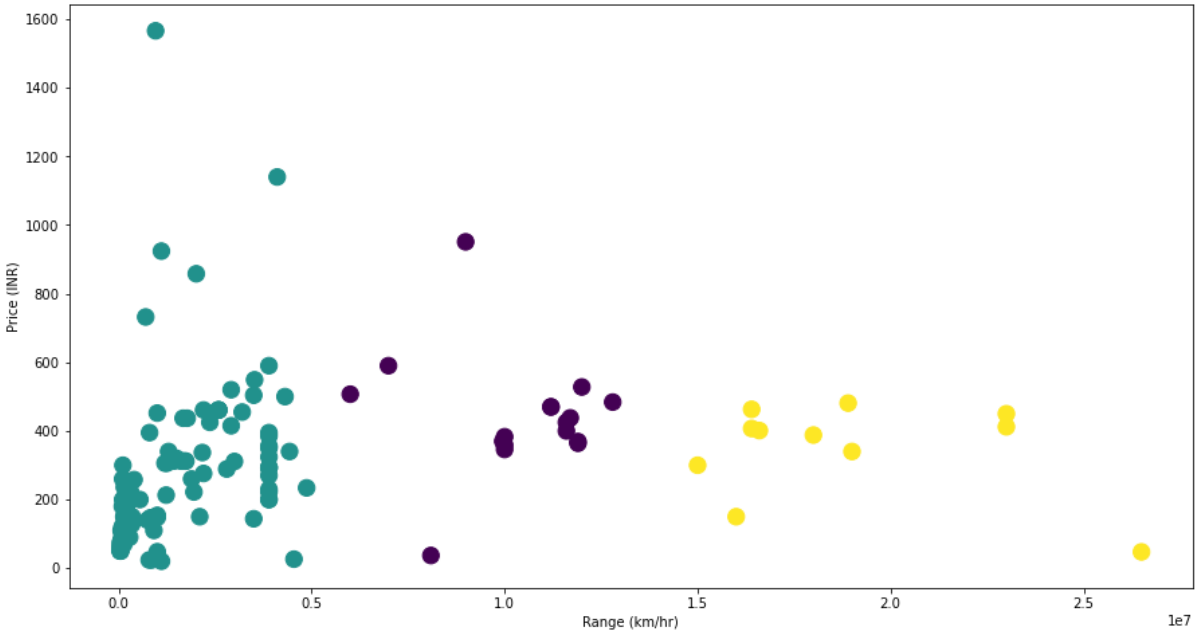
$$\boxed{J = \sum_{i=1}^{m} \sum_{k=1}^{K} w_{ik} ||x^i - \mu_k||} \quad (1)$$

And M-step is :

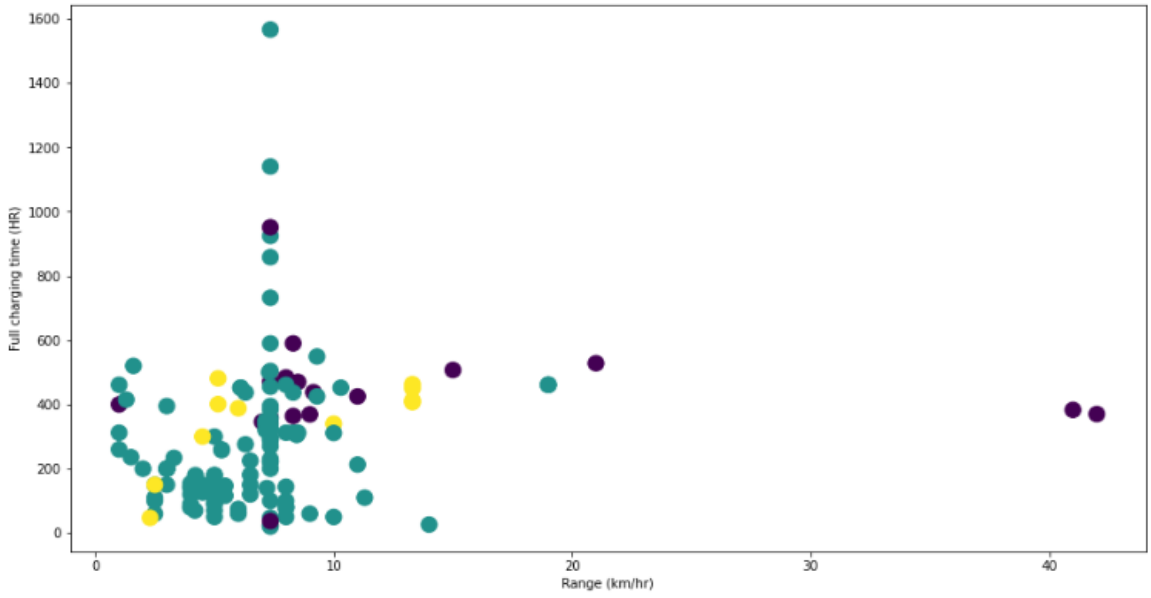
$$\frac{\partial J}{\partial \mu_k} = 2 \sum_{i=1}^{m} w_{ik}(x^i - \mu_k) = 0$$

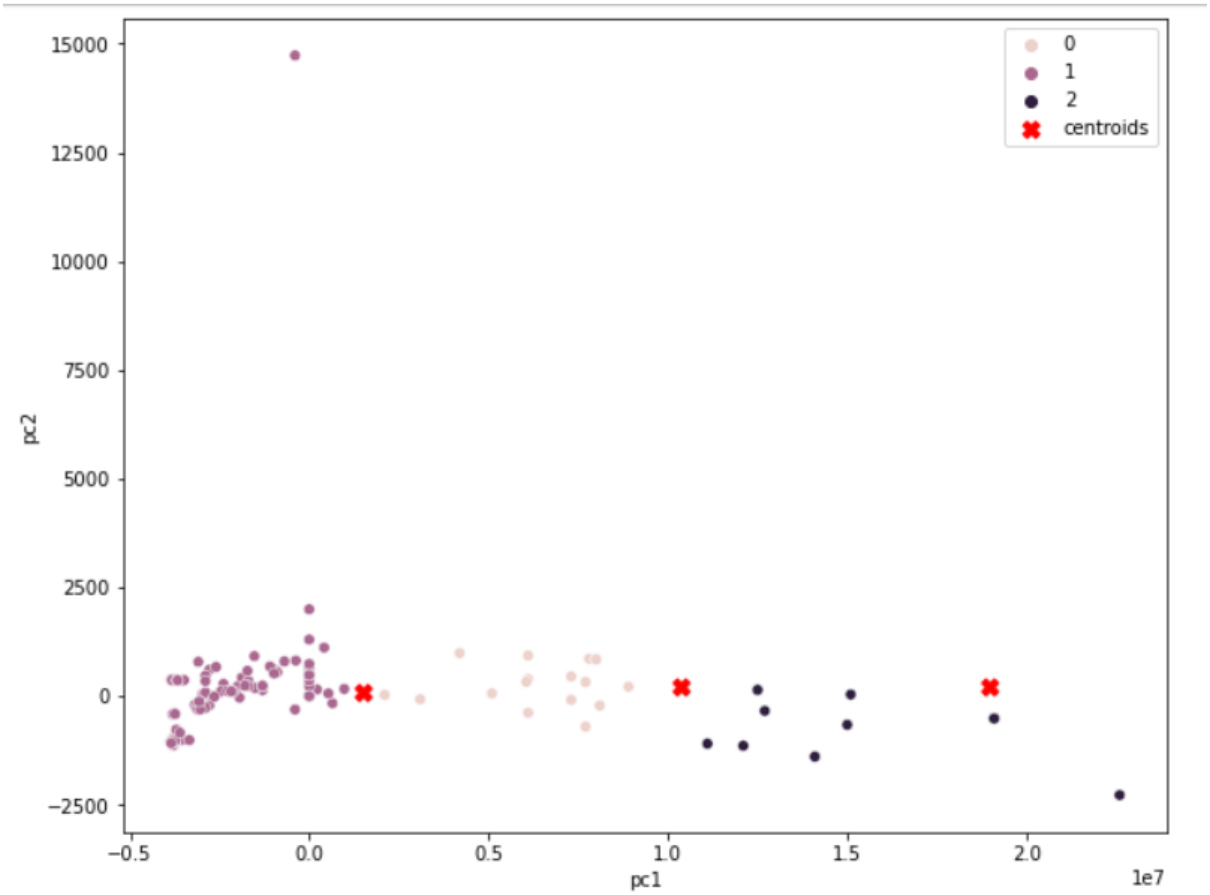
$$\Rightarrow \mu_k = \frac{\sum_{i=1}^{m} w_{ik} x^i}{\sum_{i=1}^{m} w_{ik}}$$

K- Means clustering / Centroid Method:



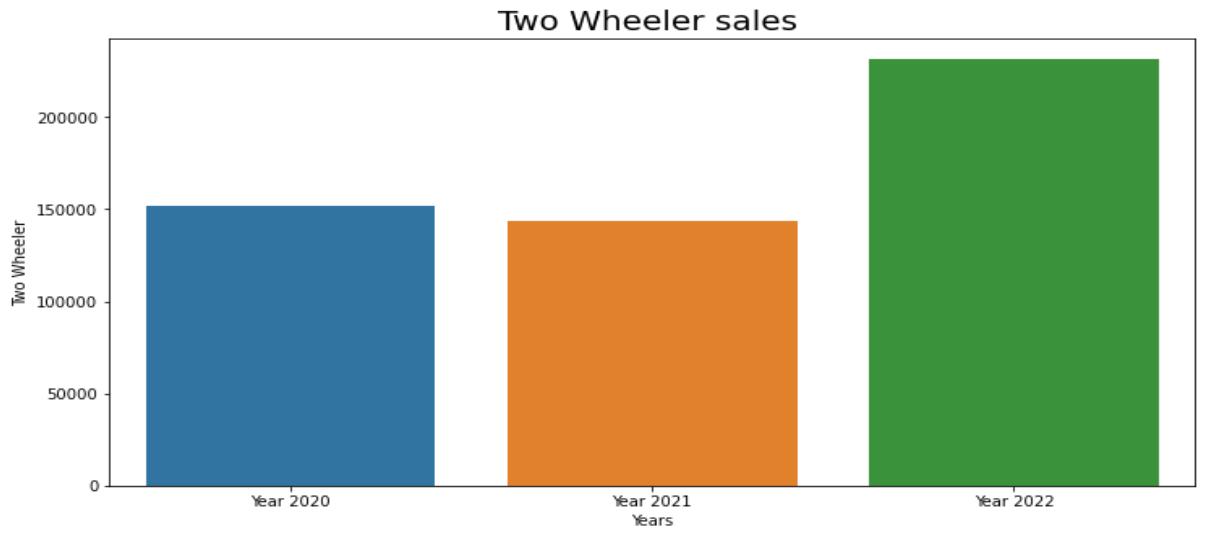
Here we used K means clustering while plotting Price(INR) vs Range(km/hr). Here we observed the three clusters are linearly separable.

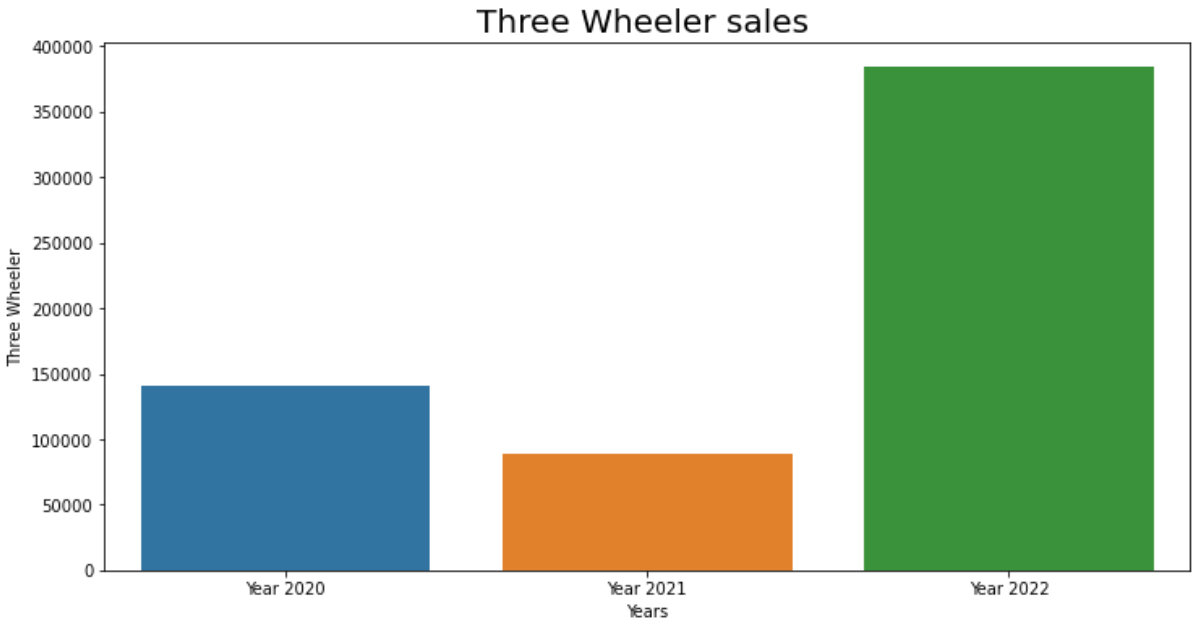


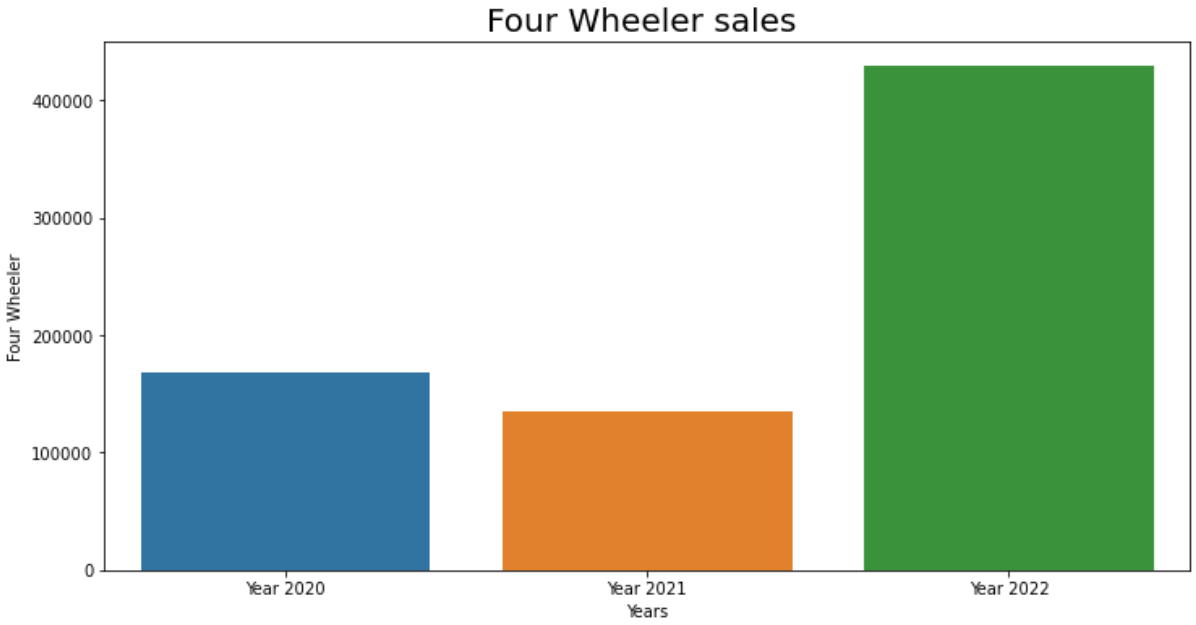


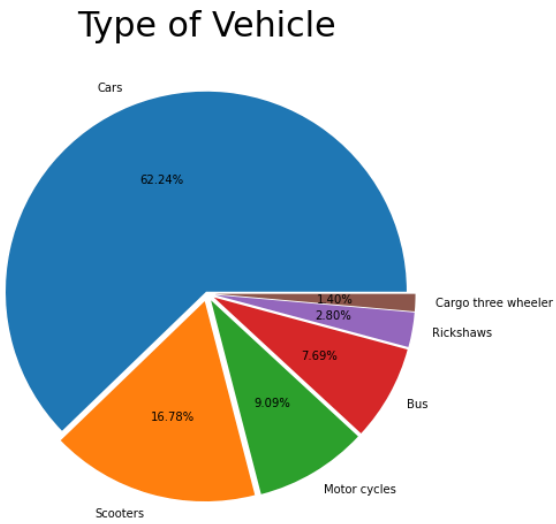
In the above Fig. we create 3 clusters by using K-Means Clustering and visualize for better understanding with Centroids.

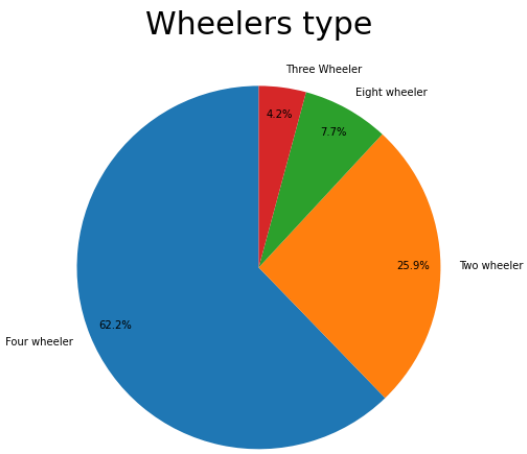
5. Selection of target Segment



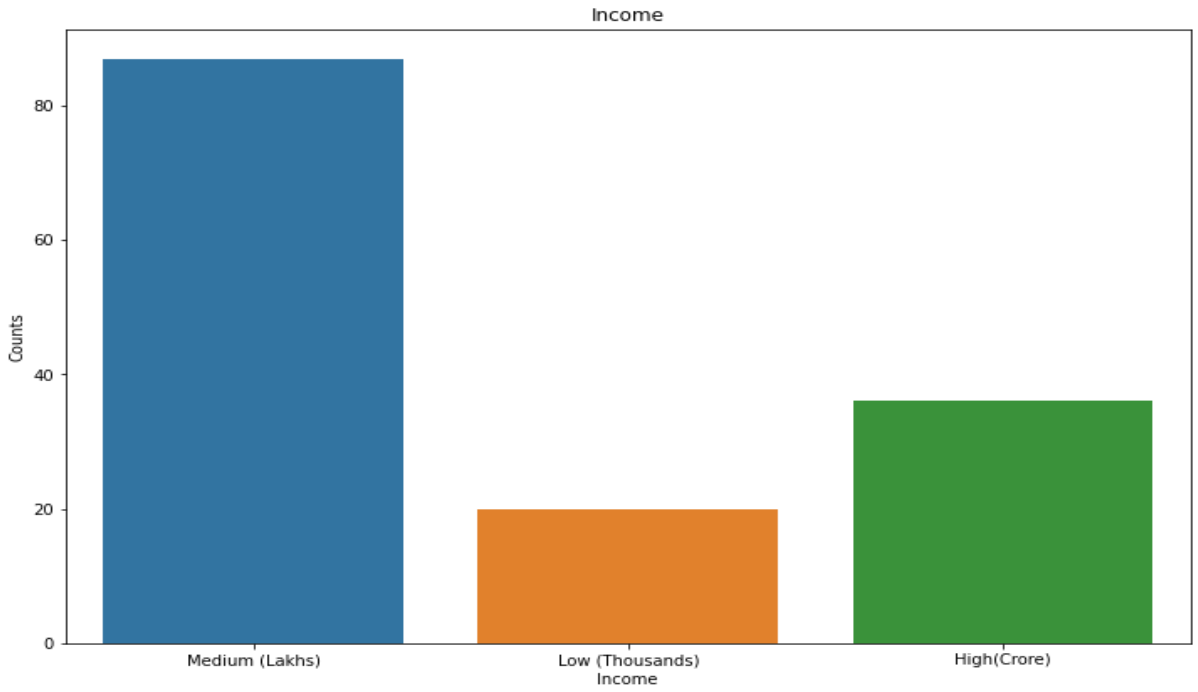

Three Wheeler sales
400000
350000
300000
250000
200000
150000
100000
50000
0
Three Wheeler
Year 2020
Year 2021
Year 2022
Years


Four Wheeler sales
400000
300000
200000
100000
0
Four Wheeler
Year 2020
Year 2021
Year 2022
Years

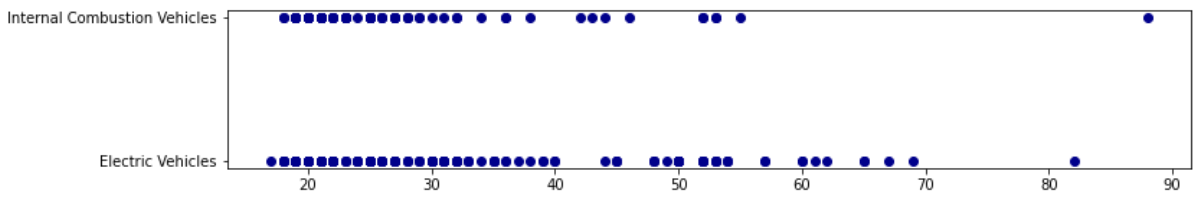




From the above graphs, we can observe that sales of four-wheelers are very high. So our main focus should be on this wheeler only. They may be the main thing in the market of Electric Vehicle.



Age Segmentation:



Here age segmentation shows a higher preference level for electric vehicles among the young population.

6. Customizing the Market Mix

The marketing mix refers to the set of actions, or tactics, that a company uses to promote its brand or product in the market.

The 4Ps make up a typical marketing mix - Price, Product, Promotion, and Place

- Price: Refers to the value that is put on a product. It depends on the cost of production, the segment targeted, the ability of the market to pay supply-demand, and a host of other direct and indirect factors. There can be several types of pricing strategies, each tied in with an overall business plan.
    - Product: Refers to the item actually being sold. The product must deliver a minimum level of performance; otherwise, even the best work on the other elements of the marketing mix won't do any good.
    - Place: Refers to the point of sale. In every industry, catching the eye of the consumer and making it easy for her to buy it is the main aim of a good distribution or 'place' strategy. Retailers pay a premium for the right location. In fact, the mantra of a successful retail business is "location, location, location."
    - Promotion: This refers to all the activities undertaken to make the product or service known to the user and trade. This can include advertising, word of mouth, press reports,

incentives, commissions, and awards to the trade. It can also include consumer schemes, direct marketing, contests, and prizes.

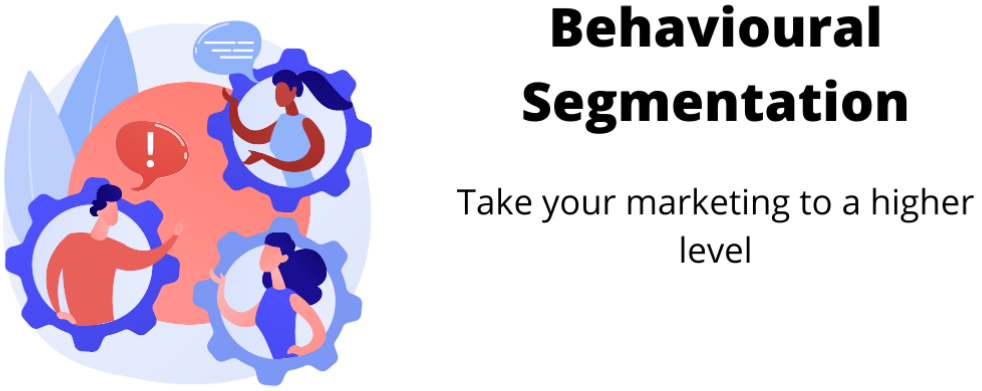
All the elements of the marketing mix influence each other. They make up the business plan for a company and handle it right, and can give it great success. The marketing mix needs a lot of understanding, market research and consultation with several people, from users to trade to manufacturing and several others.

7. Market Segmentation

Target Market: The target market of electric vehicle market segmentation can be categorized into geographic, sociodemographic, behavioral, and psychological segmentations.

Behavioral Segmentation

It searches directly for similarities in behavior or reported behavior. Example: prior experience with the product, amount spent on the purchase, etc.



Basic segmentation examines the characteristics and general activity of your customers (such as their geography, company size, website visits, form-filling activity, etc.). However,

behavioral segmentation provides a more in-depth examination of consumer behaviors, including browsing preferences, current engagement and frequency, buying trends, and other online and offline behaviors.

Customer behavior gives you the chance to influence them early and move them to the next step in the buying process since it is a reliable indicator of the customer's purchase intent or buying journey stage.

The Lead Intuition product, ACTIVE DEMAND, offers dynamic, real-time behavioral segmentation as the prospect interacts with you.

Psychographic segmentation

Psychographic segmentation consists of grouping the target audience based on their behavior, lifestyle, attitudes, and interests. For understanding the target audience, market research methods such as focus groups, surveys, interviews, and case studies can be successful in compiling this type of conclusion.

1. Lifestyle: A consumer whose profession is more time-consuming than other average consumers, that consumer may select a vehicle that takes less time to charge a vehicle. This group of consumers only focuses on the time required to charge an EV.

2. Interests: Some consumers may have an interest in particular manufacturing companies. Some consumers may prefer only vehicles made by the Tata company.
3. Behavior: The behavior of consumers is the most important factor in the market segment. Does it show what exactly consumers want from us? Some consumers may want an EV that will cover a long distance per charge.

Demographic segmentation

Demographic segmentation consists of dividing the market through different variables such as age, gender, nationality, education level, family size, occupation, income, etc. This is one of the most widely used forms of market segmentation since it is based on knowing how customers use your products and services and how much they are willing to pay for them.

1. Income: Income levels have a significant effect on consumer purchasing decisions. Those with higher-income levels may prefer luxury vehicles. Conversely, individuals with lower income levels may prefer to get the best deal on vehicles and are likely to choose inexpensive products/services.
2. Family size: family size also determines consumers' purchase decisions. Those who have large family members may choose four-wheelers and those who have smaller family members will choose two-wheelers.

8. Fermi Estimation:

Total Population of India- 1.4 billion.

Population in the range of 18-60 is the working population which approximated as 60% = 1.4*(0.6) = 840 million.

The population can be further divided according to their financial conditions and their requirements/Family sizes. Considering the reports and survey employment ratio in India is 70%= 588 million.

The young Population who are unmarried is between 18-26 which further constitutes about 30% of the working population = 588*(0.3) = 176.4 million.

We will take the conversion rate of people from Combustion Vehicles to E-Vehicles by 60%. Since Petrol, Diesel is getting exhausted = 176.4*(0.6) = 105.84 million.

Now consider that 30% of people between the age of 18-26 are interested in E-Cars

= 105.84*(0.3) = 31.752 million

60% of people are more interested in E-bikes

= 105.84*(0.6) = 63.504 million

Rest, 10% of people are more interested in 3 wheelers

= 105.84*(0.1) = 10.584 million.

Now People above 26 are married and they might be having a family of 4 or more if they are having kids.They constitute about 70% of the total working population who are employed = 588*(0.7) = 411.6 million.

We will take the conversion rate of people from Combustion Vehicles to E-Vehicles by 60%.

Since Petrol, Diesel is getting exhausted people will most likely switch = 411.6*(0.6) = 246.96 million.

70% of people in this group will buy E-car = 246.96*(0.7) = 172.82 million.

These vehicles are generally used for a family of 4 or more we will take 4 as the ideal family size and normal car capacity = 172.82/4 = 43.205 million.

Salary being a barrier 20% of people will buy only E-bikes = 246.96*(0.2) = 49.392 million.

Since a bike is a 2-seater, we will divide it by 2 = 49.392/2 = 24.696.

10% of people in this group will buy 3-wheeler because of either inability in driving normal vehicles or they constitute the Auto Drivers = 246.96*(0.1) = 24.696 million.

The demand for E-cars = 74.957 million.

The demand for E-bikes = 88.2 million.

The demand for E-auto or E-3 wheeler bikes = 35.28 million.

Estimated Car Price = 700,000 rupees = 700,000*74957000 = 52.4 trillion rupees.

Estimated Bike Price = 70,000 rupees = 70,000*88200000 = 6.174 trillion rupees.

Estimated 3-wheeler Price = 90,000 rupees = 90,000*35280000 = 3.1752 trillion rupees.

This money will further divide among the competing companies Considering 10 companies preparing the E-Vehicles.

Considering the market share of our new company be 0.08.

Estimated car market = 4.192 trillion rupees = 4192 billion rupees.

Estimated bike market = 0.493 trillion rupees = 493 billion rupees.

Estimated 3-wheeler market = 0.254 trillion rupees = 254 billion rupees.

9. Insight and Recommendation

Since EVs are more effective, their adoption has increased dramatically in India during the past five years. Additionally, rising fuel costs are contributing to a significant increase in product adoption, mostly because of the increased efficiency and range of these products. Manufacturers have been forced to provide electric vehicles all around the world due to factors including rising demand for low-emission commuting and governments encouraging long-range, zero-emission vehicles through subsidies & tax refunds. We develop several segment kinds based on this information to influence consumers' purchase decisions. Geographic segmentation refers to the locations, cities, and states that influence market sales based on customer residence. For example, there may be fewer charging stations available to consumers who reside in rural locations than in metropolitan ones. Only approx.. 1742 public charging stations are currently accessible as of 2022. Because of this, consumers from states with more charging stations have a higher likelihood of making purchases than consumers from states with less charging stations. Since it is focused on understanding how customers use your products and services and how much they pay for them, demographic segmentation places a focus on factors like education level, family size, employment, income, etc. That depends on the level of education, financial situation, and reason for the purchase of EVs. Customers who want to purchase an EV for the purpose of carrying products between cities or states will pay particular attention to a vehicle's cargo area and maximum range. On a psychological level, some customers could choose a product that makes them happy, while others might choose one that is less expensive and their other elements are average.

GitHub link-

Sanchit Agarkar

Lakshya Pokharna

Ajinkya Mahure

Ishu Khandelwal

Harsh Makwana

References

E-amrit.niti.gov.in. 2022, *e-AMRIT*, https://e-amrit.niti.gov.in/home

Jaiswal, D., Deshmukh, A. and Thaichon, P., 2022. Who will adopt electric vehicles? Segmenting and exemplifying potential buyer heterogeneity and forthcoming research. *Journal of Retailing and Consumer Services*, 67, p.102969.

Kaggle.com. 2022. *Kaggle: Your Machine Learning and Data Science Community*, https://www.kaggle.com/

Kizielewicz, B. and Dobryakova, L., 2020. How to choose the optimal single-track vehicle to move in the city? Electric scooters study case. *Procedia Computer Science*, 176, pp.2243-2253.