ELECTRONIC VEHICLE MARKET IN INDIA

TEAM – HIMANSHI

MEMBERS – 1. HIMANSHI GARG (TEAM LEADER)

2. ARYAN KUMAR

3. P SADHASIVAM

4. KUSHAGRA BHATNAGAR

Problem Statement:

Electric vehicles (EVs) are automobiles powered by electric motors rather than internal combustion engines (ICEs) that rely on fossil fuels. They utilize rechargeable batteries or other energy storage systems to store electricity and provide power to the electric motor, which propels the vehicle. EVs offer several advantages over traditional ICE vehicles, making them an increasingly popular choice among consumers and an important component of sustainable transportation solutions.

Electric vehicle market penetration is growing globally, but it still faces significant challenges, including high prices, limited driving range, lack of charging infrastructure, and consumer awareness.

1. **High Prices**: Electric vehicles (EVs) tend to be more expensive than conventional vehicles, and this is a significant barrier for many potential buyers. The high cost is mainly due to the battery pack, which can account for up to half of the vehicle's cost. EVs may be cheaper to operate in the long run, but the upfront cost is still a challenge for many consumers.

2. **Limited Driving Range**: The driving range of most EVs is still limited compared to conventional vehicles. This means that EVs may not be suitable for long-distance travel or for people who do not have access to charging infrastructure at home or work. Range anxiety is a significant barrier for many consumers.

3. **Lack of Charging Infrastructure**: A lack of charging infrastructure is a significant challenge for the EV market. Many people do not have access to charging stations at home or work, and public charging stations are often not available or located in inconvenient locations. This makes it difficult for people to rely on EVs for their daily transportation needs.

4. **Consumer Awareness**: Many consumers are not aware of the benefits of EVs and may have misconceptions about them. Some people may be hesitant to switch to EVs because they are not familiar with the technology, or they may have concerns about the availability of charging infrastructure or the driving range. Lack of consumer awareness is a significant barrier to the growth of the EV market.

Addressing these challenges will be critical to the continued growth and success of the EV market. The industry will need to continue to invest in battery technology, charging infrastructure, and consumer education to overcome these barriers and accelerate the adoption of EVs.

Electric Vehicle Startup is seeking to enter the Indian market and capture a significant share of the growing demand for electric vehicles. To ensure a successful market entry, we need to conduct a thorough segmentation analysis to identify the most viable customer segments for our EVs. By understanding the unique characteristics and preferences of these segments, we can tailor our marketing and product strategies to effectively target and meet the needs of these specific groups.

Segmentation Analysis:

Geographic Segmentation:

Regional preferences: Study the demand for electric vehicles across different states and cities in India.

Infrastructure availability: Assess the presence of charging infrastructure and availability of government incentives for EVs in different geographic areas.

Climate and terrain: Analyze the impact of climate conditions and topography on the suitability and adoption of EVs.

Demographic Segmentation:

Age and income: Determine the age groups and income levels that are more likely to adopt electric vehicles.

Occupation and lifestyle: Explore the preferences of working professionals, entrepreneurs, and individuals with environmentally conscious lifestyles.

Family size: Examine the requirements of families and their transportation needs.

Psychographic Segmentation:

Environmental consciousness: Evaluate the interest and willingness of individuals who prioritize sustainability and eco-friendly practices.

Technological inclination: Identify early adopters who embrace new technologies and are open to electric vehicles.

Social status and image: Study the impact of electric vehicles on social status and the desire of individuals to showcase their eco-consciousness.

Behavioural Segmentation:

Commuting patterns: Analyse the travel habits and patterns of individuals, including daily commute distances, frequency, and charging needs.

Purchase decision factors: Understand the key considerations influencing the purchase decision for electric vehicles, such as cost savings, environmental benefits, and performance.

Brand loyalty: Assess the willingness of customers to switch from conventional vehicles to EVs and the factors influencing their brand loyalty.

By conducting an in-depth analysis of these market segments, we can develop a comprehensive understanding of the Indian electric vehicle market and identify the segments that offer the highest potential for adoption and success. This knowledge will enable us to create a targeted and effective market entry strategy, ensuring that our EVs meet the specific requirements and preferences of the identified segments.

Data Sources -

https://www.kaggle.com/datasets/karivedha/indian-consumers-cars-purchasing-behaviour

https://www.kaggle.com/datasets/kkhandekar/cheapest-electric-cars

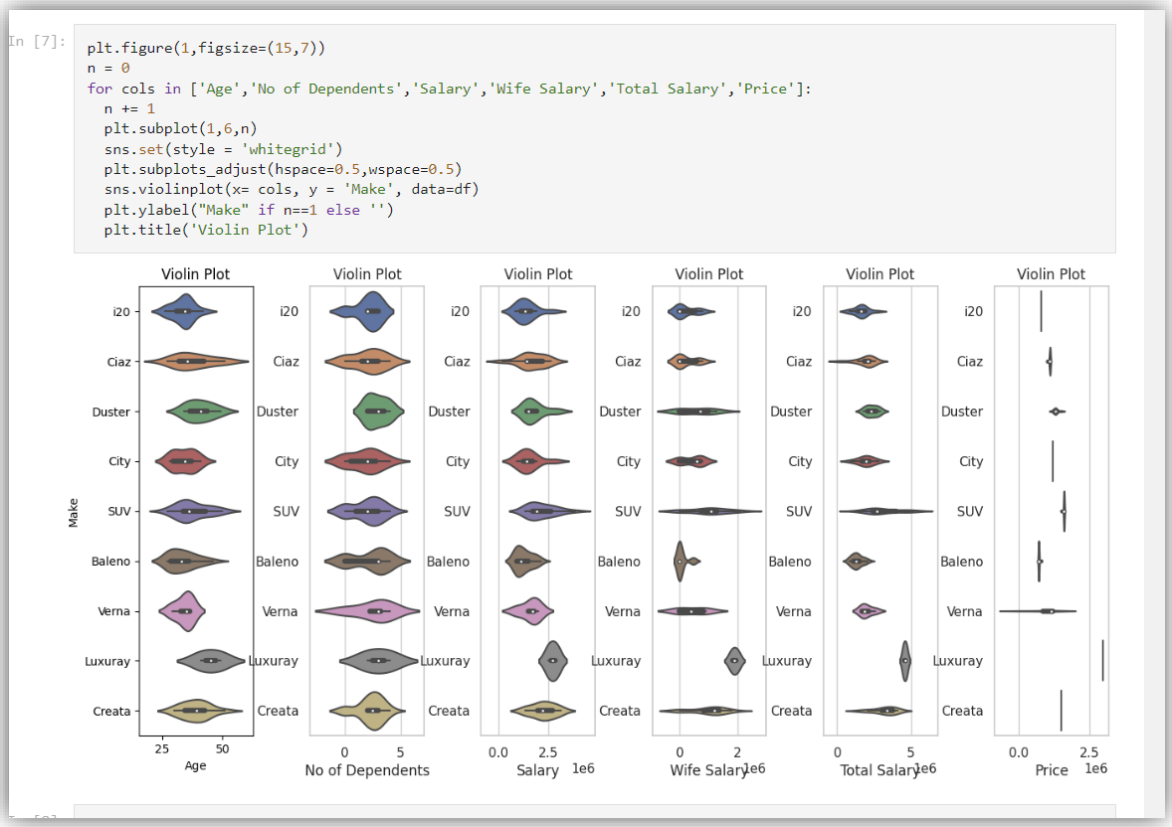
https://www.kaggle.com/datasets/deadprstkrish/ev-cars-user-reviews-india

https://dataspace.mobi/dataset/electric-vehicle-charging-station-list/resource/f39bb18a-bf5b-4e93-a22e-91f13b2ad9a7

Data Pre-processing -

Libraries like Pandas and NumPy has been used. Pre-processing that has been done includes finding null values, checking if there is any missing data, new criteria formed for price ranges, checking if any column requires a change in its label.

Demographic Analysis:



Conclusion from above violin plots:

Customer less than age 25 purchase less expensive cars. If numbers of dependents are more consumers are likely to buy cars having a greater number of seats like SUV's. If you overlap the normalised salary plots with price plot, you would observe the median of salary violin plot matches that of the price of the vehicle indicating a very direct relationship.

Other analysis is also done to establish the relationship between the price of vehicles corresponding to different factors such as age, salary etc.

Model Deployment:

There are 2 models shown in the notebook which are liner regression and k-means algorithm along with their appropriate accuracy measure

Most Optimal Market Segment:

Determining the most optimal market segment for electric vehicles in India depends on several factors, such as consumer behaviour, government policies, infrastructure, and economic conditions. Here are some potential market segments for electric vehicles in India that could be considered optimal:

Urban Commuters: With rising traffic congestion and air pollution in Indian cities, there is a growing demand for efficient and eco-friendly transportation options. Electric scooters and bicycles could be popular among urban commuters who have shorter travel distances and need a reliable and affordable mode of transportation.

Middle-Class Families: Indian families often prioritize affordability and practicality when buying a car. The lower operating costs of electric cars and government subsidies and incentives could make electric vehicles more appealing to middle-class families who are looking for a more environmentally friendly option.

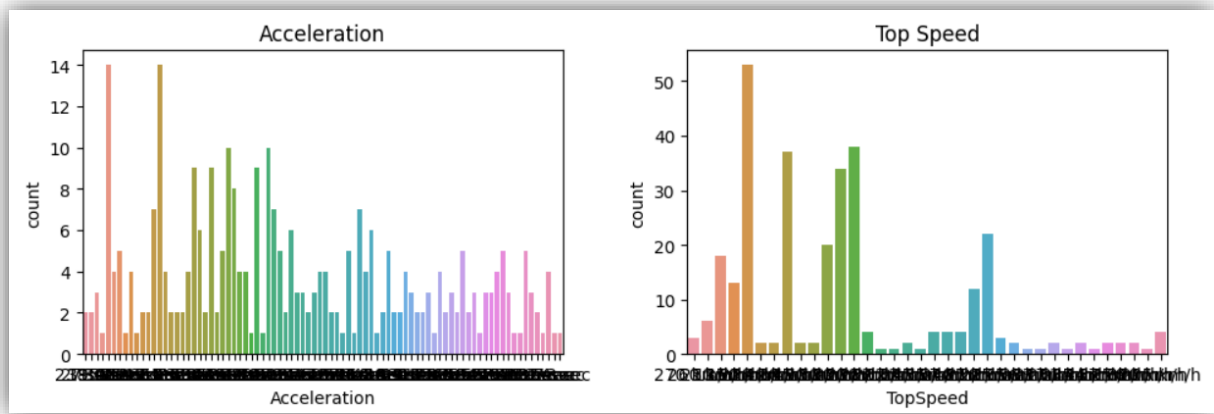
Commercial Fleets: The logistics industry in India is rapidly growing, and companies are looking for ways to reduce fuel costs and emissions. Electric trucks and vans could be an attractive option for commercial fleets that travel short distances and have a predictable daily usage pattern.

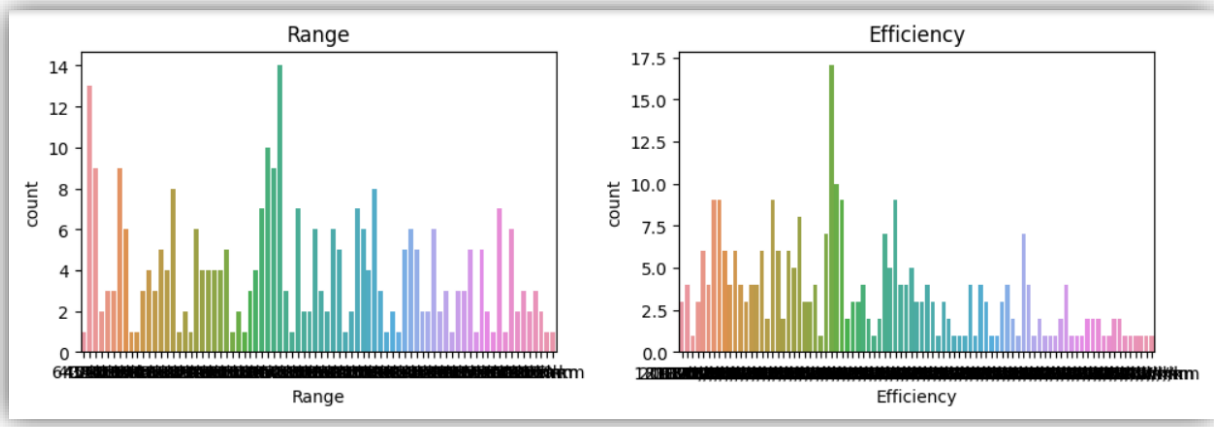
Luxury Car Buyers: Indian consumers are increasingly interested in luxury cars and are willing to pay a premium for high-end features and performance. Electric luxury cars from brands like Tesla and Mercedes-Benz could appeal to this segment of consumers who are interested in environmentally friendly options but do not want to compromise on luxury.

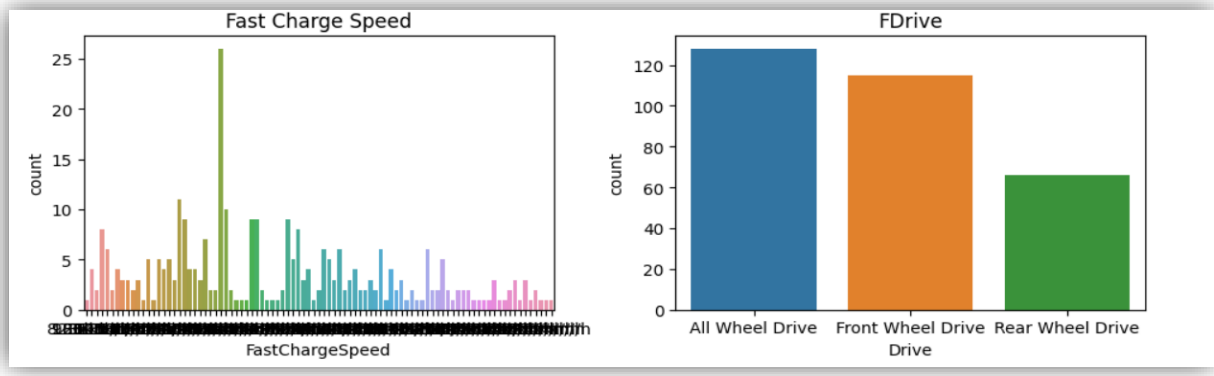
Government and Public Institutions: Indian government policies are promoting the adoption of electric vehicles through subsidies, incentives, and tax breaks. Government agencies, public institutions, and other organizations could be an important market segment for electric vehicles as they seek to reduce their carbon footprint and meet sustainability goals.

Ultimately, the most optimal market segment for electric vehicles in India will depend on a variety of factors, including consumer preferences, government policies, and infrastructure development. Companies that can successfully identify and cater to these market segments will be well positioned to capture a share of the growing Indian electric vehicle market.

Behavioural Analysis:







Dataset Overview: The dataset contains information about cars, including features such as names of cars, subtitles, acceleration, top speed, range, efficiency, fast charge speed, drive, and number of seats. It includes multiple categorical variables and some numerical variables.

Exploratory Data Analysis (EDA): Exploratory data analysis techniques were applied to gain insights into the dataset. This involved analyzing the distribution of categorical variables using bar plots and count plots. Histograms and box plots were used to explore the distribution and outliers of numerical variables.

Correlation Analysis: A correlation matrix was generated to understand the relationships between numerical variables. The correlation matrix can be visualized using a heatmap, which provides an overview of the correlation coefficients between variables.

Graphical Analysis: Various graphical analyses were performed to understand the relationships between variables. This included bar plots, histograms, box plots, scatter plots, and pair plots. These plots helped visualize the distributions, relationships, and potential patterns across the variables.

Geographic Analysis:

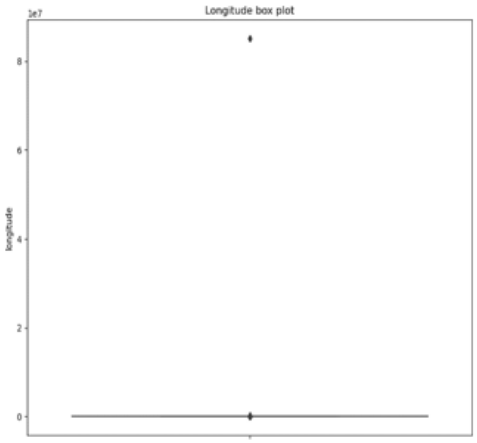
Geographical Segmentation Introduction: The aim of this segmentation analysis is to identify and understand the EV market in India using a dataset of EV charging stations. The dataset includes important attributes such as the name, state, city, address, latitude, longitude, and type of each charging station. By analysing this dataset, we can gain valuable insights into the geographic distribution of charging stations, which can aid in segmenting the EV market based on location.

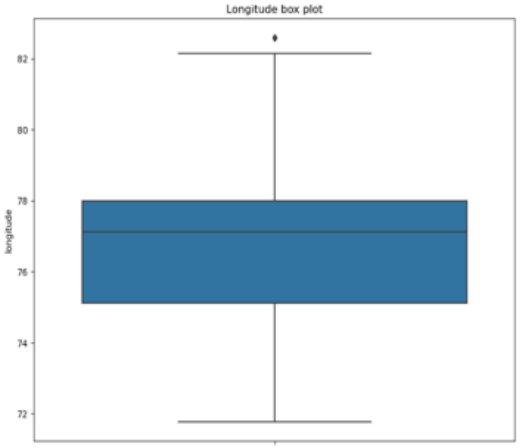
	name	state	city	address	lattitude	longitude	type
0	Tata Power	Andaman	Port Blair	Gennext Motors , Pahargaon, Ground Floor New, ...	11.630305	92.724835	7.0
1	E Planet, Thalassery	Andhra pradesh	kannur	AVK Nair Road, Pilakool, Thalassery, Kerala	11.74554436	75.495020	6.0
2	M R A Restaurant	Andhra pradesh	kannur	Kasaragod - Kannur Rd, Trichambaram, Taliparam...	12.02562662	75.367177	6.0
3	Matsya Amazon Kitchens, Kurnool Rd	Andhra pradesh	Kurnool	NH 44; Bangalore - Hyderabad highway, Kurnool...	15.576915	77.937040	7.0
4	Blue Moon Highway Restaurant, Bangalore-Hydera...	Andhra pradesh	Anantapuram	NH 44; Bangalore - Hyderabad highway, 116/1B,...	15.02395	77.620628	7.0
5	Aavas Hotels & Resorts, NH 16 Velampalli	Andhra pradesh	Ongole	NH 16, Vijayawada - Nellore Hwy, Vellampalli (...	15.6382976	80.016753	7.0
6	IOCL - NAVABHARAGAVA FILLING STATION, Vijayawa...	Andhra pradesh	Vijayawada	Vijayawada - Hyderabad Hwy, Ibrahimpatnam, And...	16.5812639	80.526349	7.0
7	Tristar Auto Agencies, RK Beach Junction	Andhra pradesh	Vizag	Beside Kali Matha Mandhir, RK Beach Junction, ...	17.71241304	83.318481	7.0
8	Highway Grand World, Chennai-Tirupati Highway ...	Andhra pradesh	Tirupati	Chennai - Thiruttani - Renigunta Hwy, NH716 - ...	13.29360934	79.581709	7.0
9	Hotel Sapthagiri Holiday Resorts,Tirupati u00e...	Andhra pradesh	Chittoor	NH140, Tirupati u00e2u0080u0093 Bangalore Hig...	13.37412051	79.092103	7.0

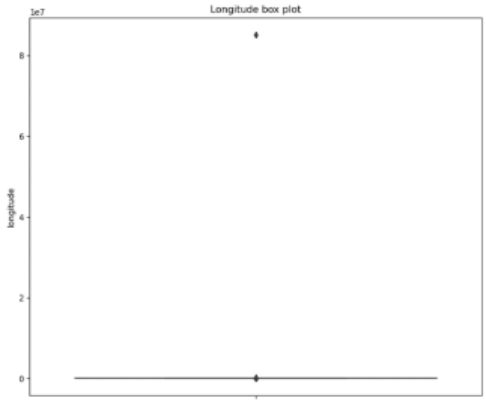
In the data preprocessing phase, we addressed two key aspects: handling null values in the address column and removing outliers from the latitude and longitude columns. Null values in the address column were replaced using reverse geolocation techniques. Since latitude and longitude coordinates were available for each charging station, we utilized a reverse geolocator to obtain the corresponding addresses. To maintain the integrity of the analysis, we also addressed outliers present in the latitude and longitude columns. Outliers can potentially skew the results and impact the accuracy of subsequent analyses. To mitigate this, we employed the Interquartile Range (IQR) method.

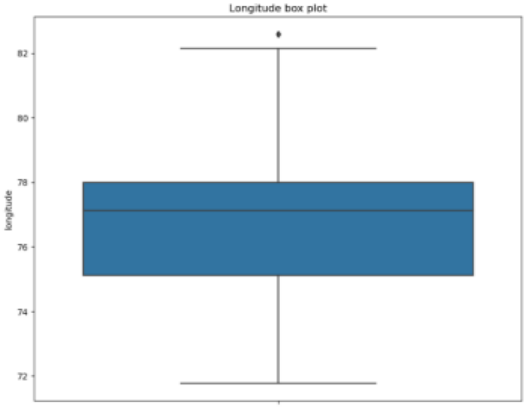
	name	state	city	address
195	AARGO CSC MINTO ROAD	Delhi	NEW DELHI	NaN
392	Savoy Suites Manesar	Haryana	Gurugram	NaN
393	Mehak Garden Charging Station	Haryana	Sirsa	NaN
394	Aman Hotel Kurukshetra DC Charging Station	Haryana	Shabhad	NaN
395	Statiq RnD Centre DC Charging Station	Haryana	Gurgaon	NaN
396	Park Plaza Gurgaon DC Charging Station	Haryana	Gurugram	NaN
397	Monsoon Breeze AdPod	Haryana	Gurugram	NaN
411	AARGO GOLDEN GALAXY HOTEL	Haryana	FARIDABAD	NaN
418	AARGO GOLDEN GALAXY HOTEL	Haryana	FARIDABAD	NaN
440	Hotel Prabhat Palace DC Charging Station	Jammu & Kashmir	Katra	NaN
472	MESCOM Shivamogga Charging Station 1	Karnataka	Shivamogga	NaN
473	MESCOM Mudigere Charging Station	Karnataka	Karadagodu	NaN
474	Karkala Charging Station	Karnataka	Karkala	NaN

	name	state	city	address
397	Monsoon Breeze AdPod	Haryana	Gurugram	Monsoon Breeze, Manesar, Gurugram, Haryana, 12...
411	AARGO GOLDEN GALAXY HOTEL	Haryana	FARIDABAD	Golden Galaxy, NH44, Seekri, Ballabgarh, Farid...
418	AARGO GOLDEN GALAXY HOTEL	Haryana	FARIDABAD	Golden Galaxy, NH44, Seekri, Ballabgarh, Farid...
440	Hotel Prabhat Palace DC Charging Station	Jammu & Kashmir	Katra	Continent Panchvati, Katra Bypass, Katra, Reas...
472	MESCOM Shivamogga Charging Station 1	Karnataka	Shivamogga	Basavana Gudi, Shivamogga, Shivamogga taluk, S...
473	MESCOM Mudigere Charging Station	Karnataka	Karadagodu	Indiranagara, Hesagal, Mudigere taluk, Chikkam...
474	Karkala Charging Station	Karnataka	Karkala	Nakre, Karkala, Karkala taluk, Udupi district,...
475	MESCOM Surathkal Charging Station	Karnataka	Surathkal	Hotel Lalit International, SH101, Iddya, Surat...
476	Kaup Charging Station	Karnataka	Kaup	Polipu, Kapu, Kapu taluku, Udupi district, Kar...
477	Bramavara Charging Station	Karnataka	Bramavara	Bramhavara Akashavani, Brahmavara, Brahmavara ...
478	Kundapura Charging Station	Karnataka	Kundapura	Vaderahobali, Kundapura, Udupi district, Karna...



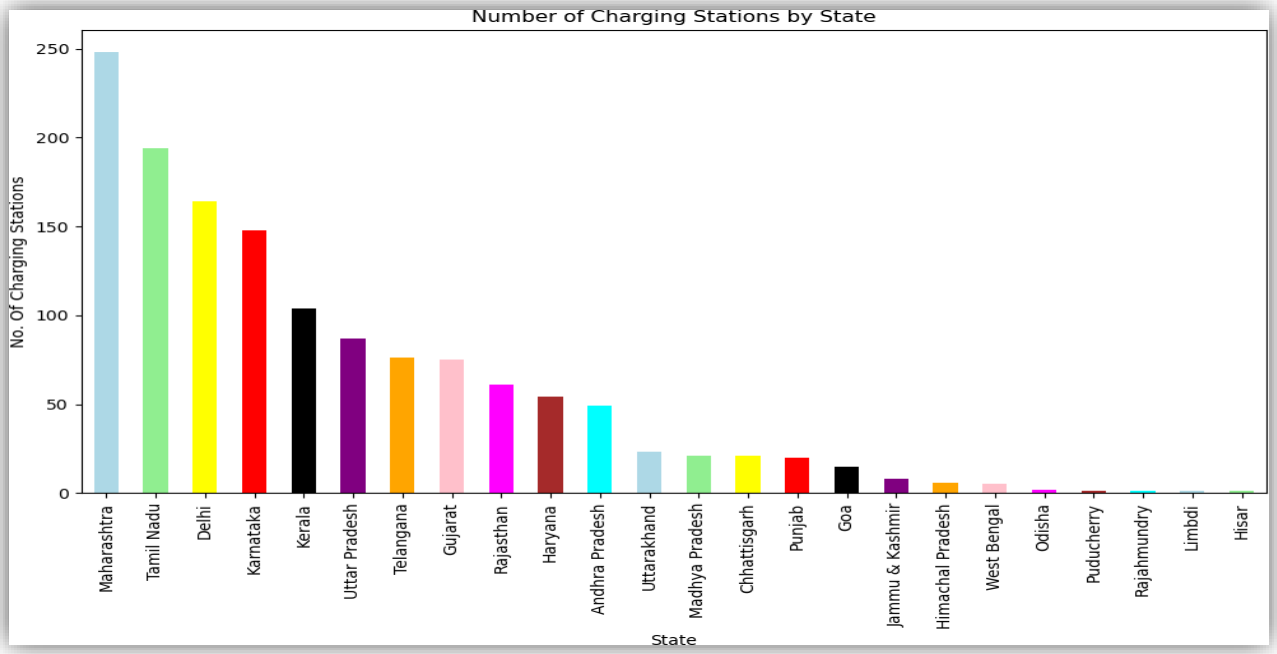


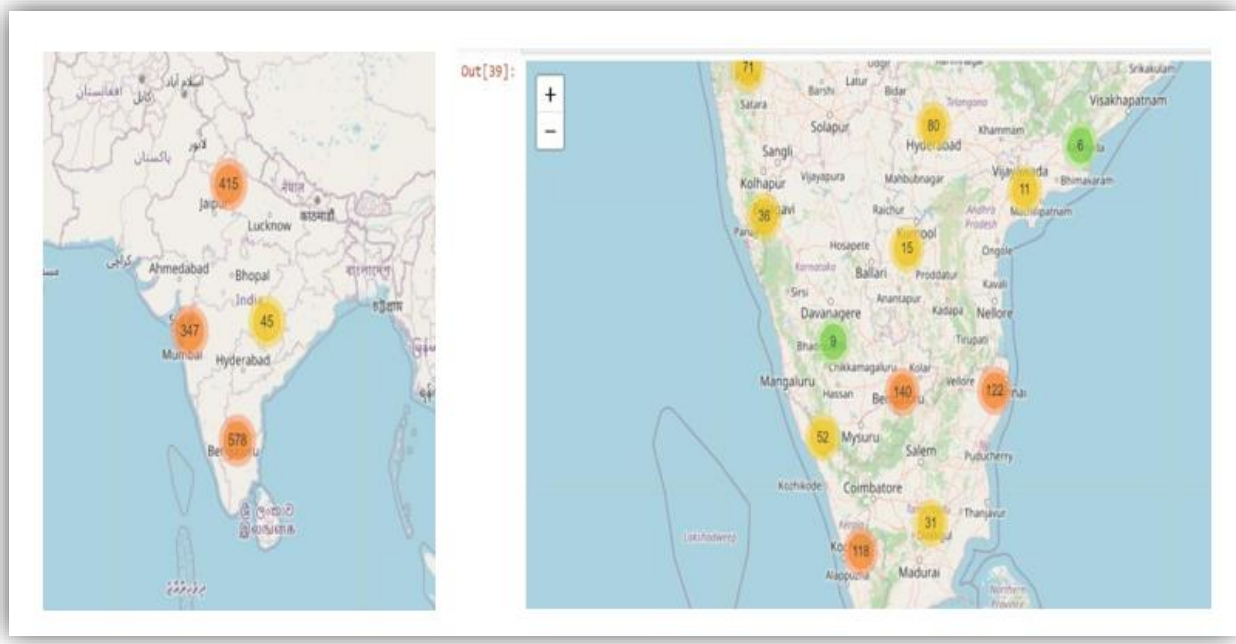


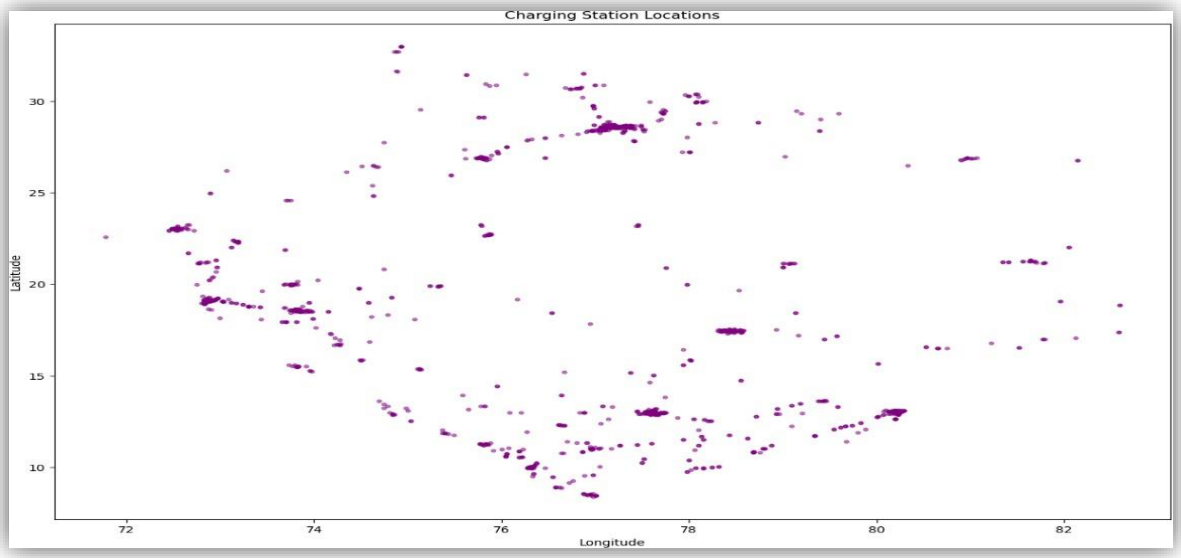


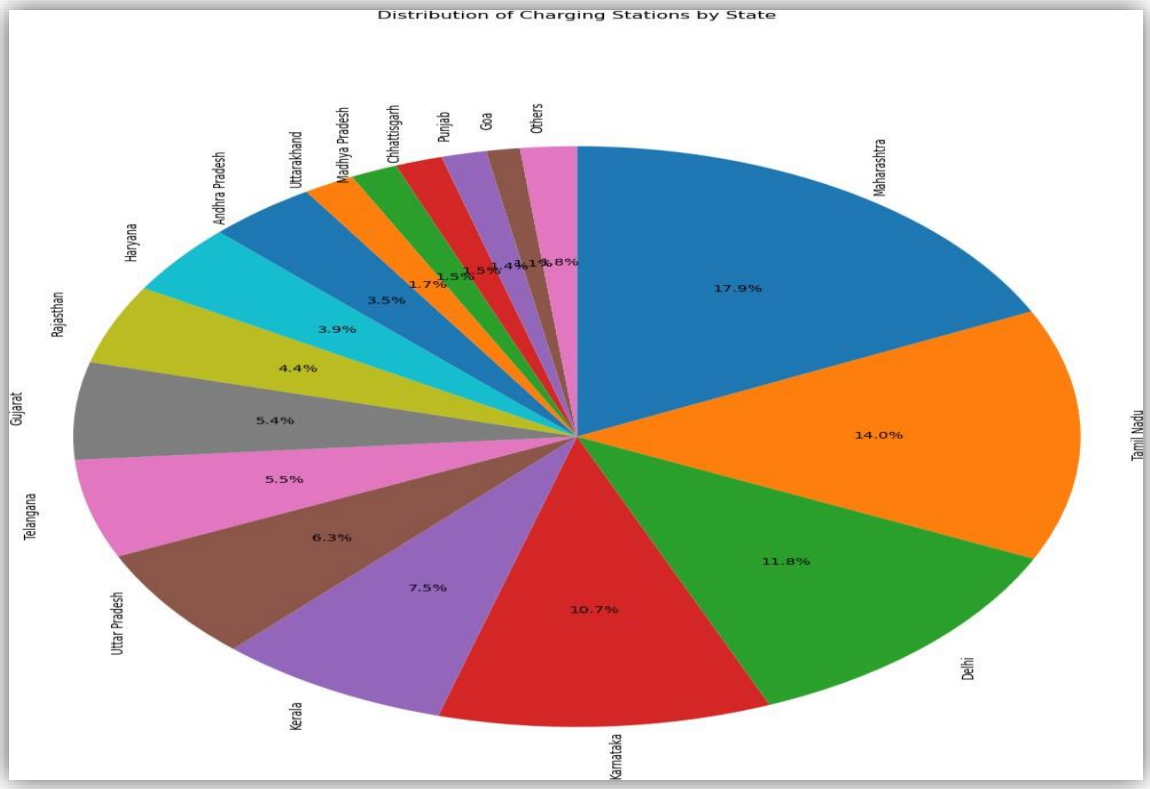
We began by creating a bar plot to visualize the state-wise distribution of charging stations. This plot allowed us to identify the top states with the highest number of stations. Maharashtra emerged as the state with the highest concentration of charging stations, followed by Tamil Nadu and Delhi. This information provides valuable insights into the regions where EV infrastructure is well-established and where the market potential for EV vehicles may be high.

Next, we explored the relationship between latitude and longitude coordinates using a scatter plot. This visualization technique allowed us to observe the spatial distribution of charging stations across India.

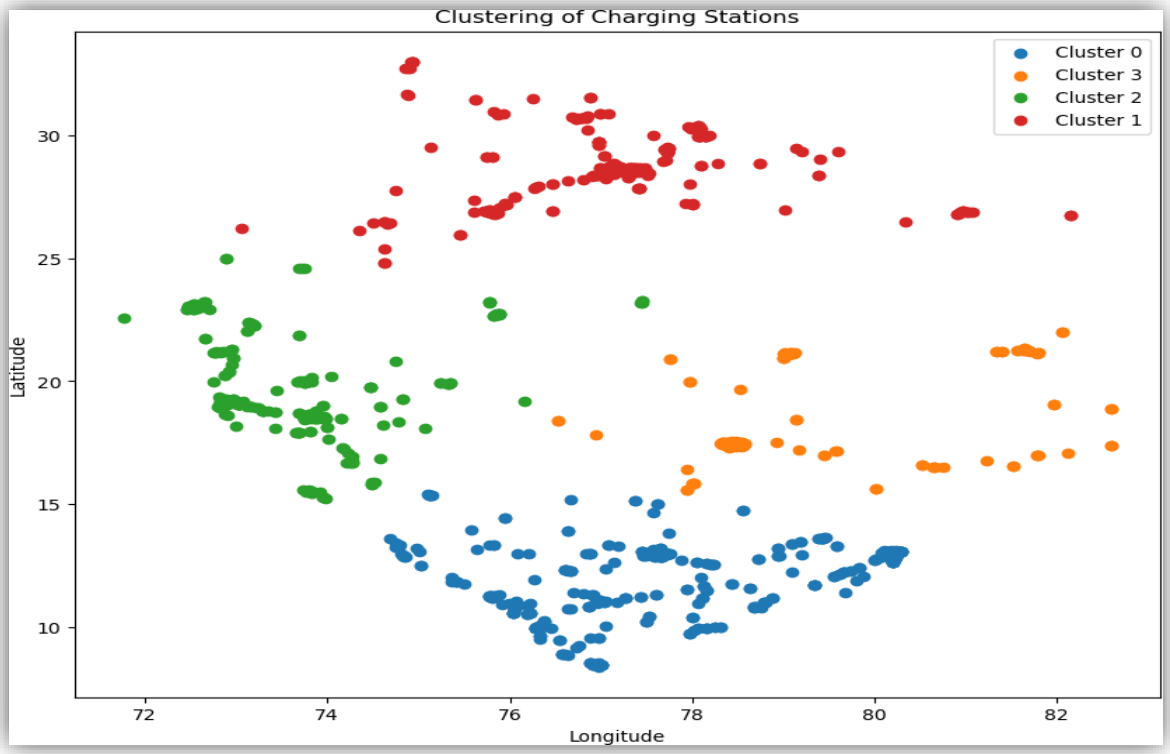




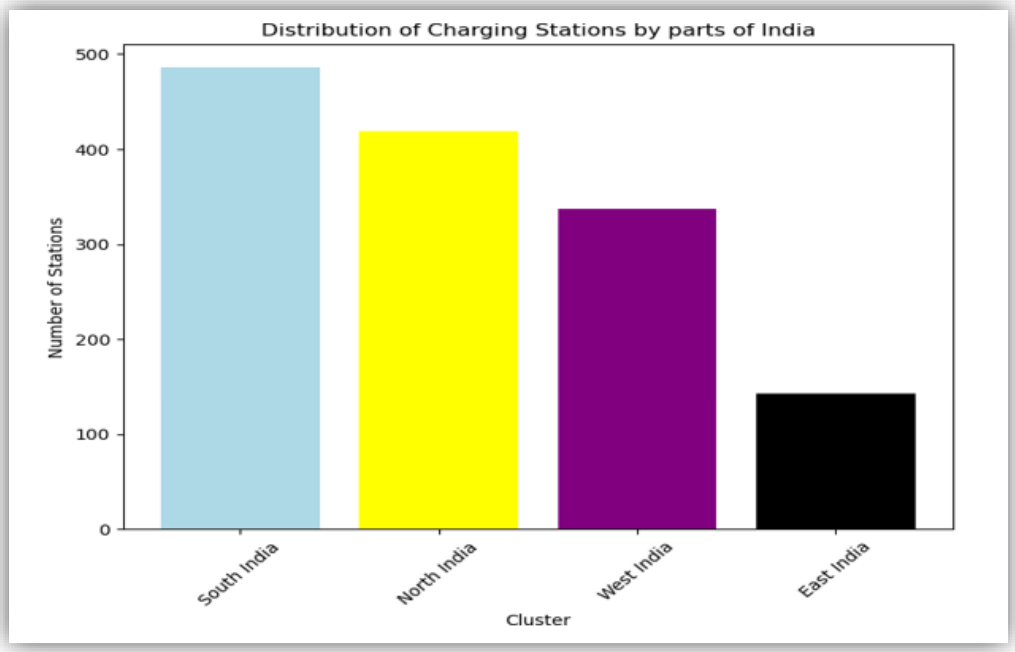




In the model building phase, we aimed to cluster the EV charging stations based on their geographic coordinates using the K-means algorithm. We chose the K-means algorithm for several reasons. Firstly, it is widely used and well-established in the field of machine learning and clustering. Secondly, it is particularly effective for geographic segmentation, as it can group data points based on their proximity in latitude and longitude space. This aligns with our objective of segmenting the EV market based on geographic factors.



The K-means algorithm assigned each charging station to one of these four clusters based on its latitude and longitude coordinates. This segmentation allowed us to identify the charging stations located in the eastern, western, northern, and southern parts of India. By doing so, we gained insights into the geographic distribution and concentration of EV charging stations across different regions of the country.



By incorporating the K-means algorithm into our analysis, we gained a deeper understanding of the geographic segmentation of the EV market in India. The clusters identified by the algorithm help delineate distinct regions or zones with similar characteristics, enabling targeted marketing strategies and infrastructure planning based on the specific needs and preferences of each segment. The application of the K-means algorithm served as a crucial step in our analysis, allowing us to identify geographic segments within the EV market and facilitating more targeted and effective decision-making processes for market expansion and development.

Summary of Segmentation Analysis:

The segmentation analysis focused on understanding the EV market in India through geographic segmentation. By utilizing a dataset of EV charging stations, we explored the distribution and concentration of charging infrastructure across different regions of the country. The analysis encompassed data preprocessing, visualization, model building using K-means clustering, and assigning meaningful names to the clusters.

Through data visualization techniques such as bar plots, scatter plots, and pie charts, we gained insights into the state-wise distribution and geographic patterns of EV charging stations. We observed that Maharashtra, Tamil Nadu, and Delhi emerged as the states with the highest concentration of charging stations. The scatter plot visualized the spatial distribution of stations, revealing a higher concentration in the southern region of India.

To further analyze and segment the market, we employed the K-means clustering algorithm. The algorithm grouped the charging stations into four clusters representing East India, West India, North India, and South India. This segmentation highlighted the geographic proximity of charging stations and provided a meaningful framework for understanding the regional differences in the EV market.

Based on the clustering results, we inferred that the southern part of India exhibits a significant concentration of charging stations. This finding suggests a higher level of EV infrastructure development and market potential in the southern states. This insight has important implications for infrastructure planning, market expansion, and targeted marketing efforts in the EV sector.

Psychographic Analysis:

Electric Vehicle is an emerging market in India, new businesses open to investing in this market require to study the trends and needs of the market in India, thus market segmentation is an important strategy that needs to be implemented by the businesses.

Among the many type of segmentations, psychographic segmentation is needed to study the requirements, needs, and behaviour of the customers.

To perform psychographic segmentation, a dataset containing the customer reviews, their needs and expectations can be used. The dataset can be divided into segments based on the features that appeal to them the most.

The original structure of the dataset is as follow:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 129 entries, 0 to 128
Data columns (total 10 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   review           129 non-null    object
 1   Exterior         129 non-null    float64
 2   Comfort          129 non-null    float64
 3   Performance      129 non-null    float64
 4   Fuel Economy     129 non-null    float64
 5   Value for Money  129 non-null    float64
 6   Condition        129 non-null    object
 7   driven           129 non-null    object
 8   rating           129 non-null    float64
 9   model_name       129 non-null    object
dtypes: float64(6), object(4)
memory usage: 10.2+ KB
```

There are 3 car models in this dataset:

- Tata Nexon EV
- Hyundai Kona
- Tata Tigor EV

The features ‘exterior’, ‘comfort’, ‘performance’, ‘fuel economy’, ‘value for money’ and ‘rating’ contains rating from 0.0 to 5.0 with 5.0 being the best.

The ‘review’ column contains detailed reviews by the users in natural language.

The ‘condition’ column contains the condition of the vehicle as new, old and ‘not purchased’

The ‘driven’ column contains the amount that the vehicle was driven for. The values in this column are:

- Did a short drive once
- Few thousand kilometres
- Few hundred kilometres
- Haven't driven it
- It’s my mate since ages

Text Analysis was performed on the reviews column to get the sentiment of the language used to describe the experience of different users.

The reviews were first cleaned using NLTK’s PorterStemmer and stopwords and regex libraries.

The cleaned reviews are then passed to the sentiment.polarity method from the TextBlob library to measure the polarity of the text in the reviews. If the polarity is between 0 and 1,

the review is positive, if it is between -1 and 0, the review is negative, and if it is 0, the review is neutral.

Two more columns are added to the dataset named ‘polarity’ and ‘analyses.’ The ‘polarity’ column contains values from -1 to +1 depending on the sentiment of the reviews. The ‘analysis’ column contains distinct values of 1, 0 or -1 as per the polarity of the review.

The new structure of the dataset is as follow:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 129 entries, 0 to 128
Data columns (total 13 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   review           129 non-null    object
 1   Exterior         129 non-null    float64
 2   Comfort          129 non-null    float64
 3   Performance      129 non-null    float64
 4   Fuel Economy     129 non-null    float64
 5   Value for Money  129 non-null    float64
 6   Condition        129 non-null    object
 7   driven           129 non-null    object
 8   rating           129 non-null    float64
 9   model_name       129 non-null    object
 10  clean_reviews    129 non-null    object
 11  Polarity         129 non-null    float64
 12  Analysis         129 non-null    int64
dtypes: float64(7), int64(1), object(5)
memory usage: 13.2+ KB
```

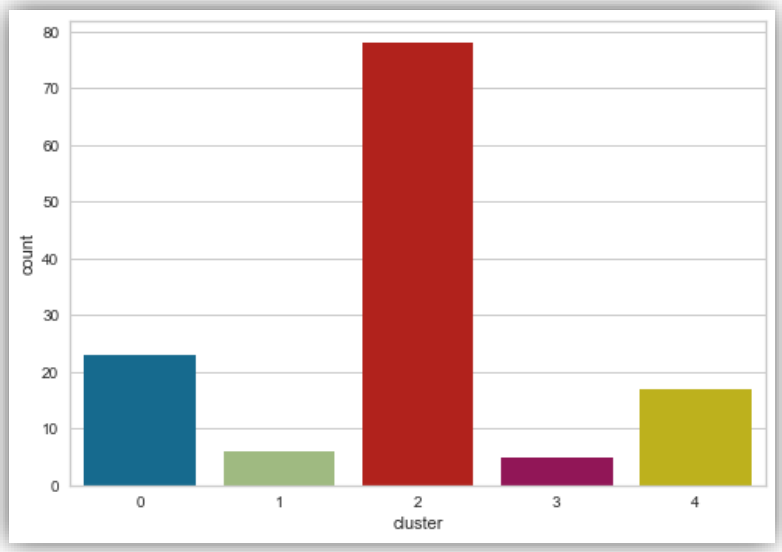
PCA is performed on the features containing numerical values in the dataset, the results of PCA are stored.

The dataset is divided into multiple set of clusters using K-means algorithm, and the within cluster distance is measured. The set of cluster with minimum within cluster distance is taken for segmentation of the dataset.

In this project, the minimum within cluster distance was a result of a set of cluster containing 5 clusters, therefore the dataset was divided into 5 segments.

The dataset was divided into 5 segments, with the number of customers per cluster as follow:

- Cluster 2 78
- Cluster 0 23
- Cluster 4 17
- Cluster 1 6
- Cluster 3 5



The ‘driven’, ‘condition’ and ‘model_name’ columns were converted to numerical type in order to fit the decision tree algorithm that was later used to determine the importance of various features in the segmentation process of the dataset.

After the segmentation, and changing the type of the 3 features to int, the structure of the dataset is:

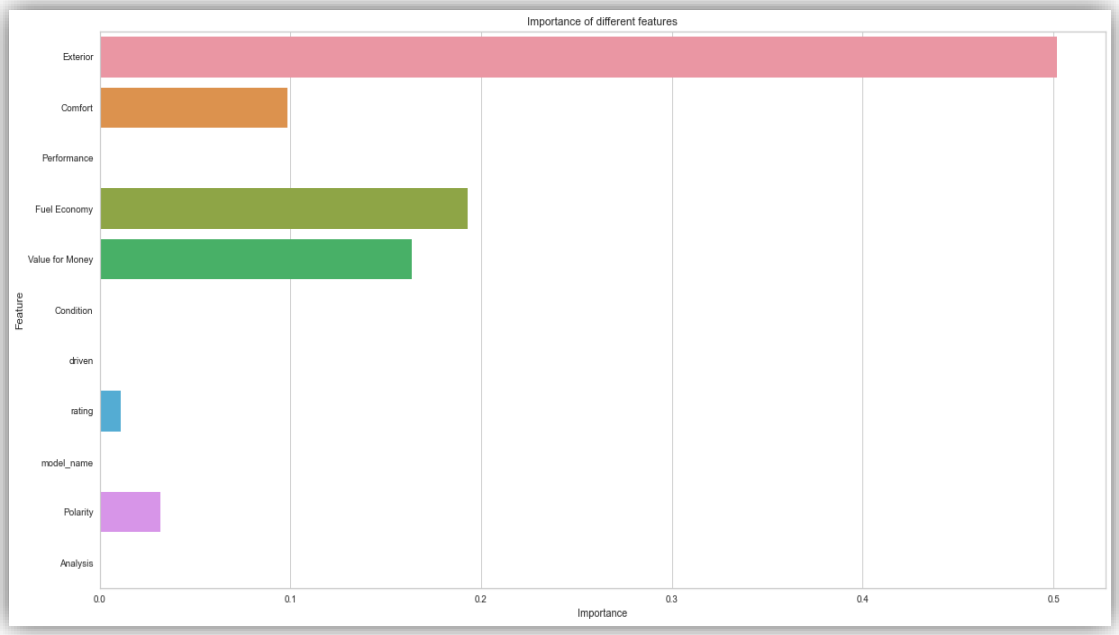
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 129 entries, 0 to 128
Data columns (total 14 columns):
 #   Column           Non-Null Count  Dtype
---  ------           --------------  -----
 0   review           129 non-null    object
 1   Exterior         129 non-null    float64
 2   Comfort          129 non-null    float64
 3   Performance      129 non-null    float64
 4   Fuel Economy     129 non-null    float64
 5   Value for Money  129 non-null    float64
 6   Condition        129 non-null    int32
 7   driven           129 non-null    int32
 8   rating           129 non-null    float64
 9   model_name       129 non-null    int32
 10  clean_reviews    129 non-null    object
 11  Polarity         129 non-null    float64
 12  Analysis         129 non-null    int64
 13  cluster_num      129 non-null    int32
dtypes: float64(7), int32(4), int64(1), object(2)
memory usage: 12.2+ KB
```

The dataset was split into 70:30 ratio for training and testing the decision tree algorithm.

The decision tree could predict the segment of the user based on their review with 92.3076% accuracy.

The most feature playing the most important role in the segmentation was the Exterior of the car, followed by value for money and comfort. The fuel economy also played an important role in the segmentation.

The importance of various features is shown in the bar graph below:



Selection of target segment -

Based on the analysis, the target segment can be narrowed down to EVs having:

- Psychographic factors such as Comfort and Value for Money
- Behavioural factors such as good Acceleration and viable Price range
- Geographic factors such as States which are more market friendly

In conclusion, the target segment should comprise of EVs having Acceleration of 7.5-10 sec, High in Comfort and Value for Money ratings, have a Price range of 20-30 Lakhs, and be focused mainly on States such as Maharashtra, Karnataka, Tamil Nadu and Rajasthan

Github Links -

1. Himanshi - https://github.com/Ghimanshigit03/FeynnLabs-Internship2023/tree/main/Project2.1-Electric%20Vehicle%20Market%20In%20India
2. Aryan - https://github.com/aryan311/Fynn_labs/tree/main/Electric_vehicles
3. Sadhasivam - https://github.com/Sadhasivam9/EV-Market-Segmentation-.git
4. Kushagra - https://github.com/Bhatnagar621/MS_EV.git

Summary -

There are many EV manufacturing companies in the country like Hero Electric, Tata Motors, Ather Energy, Ashok Leyland, Hyundai Kona Electric, etc. Tesla has also arrived; the demand will get higher & higher since it is automotive so the investments and policies and all that would be bigger but it will take some time to perfectly settle in India. The following are the key insights of the project:

¬ The electric vehicle industry has not done that much good due to the devastating hit of the Covid outbreak but it will take a huge jump in upcoming years

¬ The use of EVs will be game-changing in terms of environment, air, noise pollution-free, postelectric, and much more

¬ The company should plan to establish local operations in India either by partnering with a local company or by setting up its own manufacturing/ development unit, potentially combined with imports of specific components

¬ The company would expect to further grow in India, underpinned by a growing commercial fleet market for two-wheelers and three-wheelers especially for last km delivery/urban freight services. The company must see opportunities across the supply chain in the battery, EV component and charging infrastructure segments including the machinery and equipment needed for establishing manufacturing plants, training and provision of skilled workforce etc.

¬ The company should start their business from Metro Cities in India and then after considerable business expand to other cities of the same state of the Metro Cities. This will help the company to expand easily as they will be having a prior knowledge of business from Metro Cities and Network of Supply chain will be easy for the company as the time goes in business

IN THE CONCLUSION, ELECTRIC VEHICLES ARE THE FUTURE!!