Market Segmentation on the EV Market in India

Shraddha Saha

8th May,2024

The Electric Vehicle (EV) market in India has garnered significant attention in recent years, driven by the twin imperatives of environmental sustainability and energy security. This abstract provides an overview of the current landscape of the EV market in India, examining key trends, challenges, and opportunities.

Firstly, it outlines the rapid growth witnessed in the Indian EV market, propelled by government initiatives such as the Faster Adoption and Manufacturing of Electric Vehicles (FAME) scheme, which aims to incentivize the adoption of electric mobility. The abstract delves into the rising consumer interest in electric vehicles, spurred by increasing awareness of environmental concerns and the advantages of EVs in terms of reduced operating costs and lower emissions.

Secondly, it discusses the challenges confronting the Indian EV market, including infrastructural limitations such as inadequate charging infrastructure and battery technology constraints. Additionally, it highlights regulatory hurdles and policy ambiguities that hinder widespread EV adoption, alongside concerns regarding the affordability and accessibility of electric vehicles for the mass market.

Thirdly, the abstract explores the emerging opportunities within the Indian EV ecosystem. It examines the growing investments in research and development aimed at enhancing battery technology and charging infrastructure. Furthermore, it discusses the potential for collaboration between public and private stakeholders to address infrastructural gaps and promote EV adoption. Moreover, the abstract explores the burgeoning market for electric two-wheelers and three-wheelers, which present a significant opportunity for electrification in India's urban mobility landscape.

In conclusion, this abstract underscores the dynamic nature of the EV market in India, characterized by rapid growth, persistent challenges, and promising opportunities. It emphasizes the need for concerted efforts from various stakeholders, including policymakers, industry players, and consumers, to realize the full potential of electric mobility and drive sustainable transportation solutions in the country.

Problem Statement

Let us suppose, I work under an Electric Vehicle Start-up. The Start-up is still deciding in which vehicle/customer space it will be developing its EVs.

I have to analyze 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.

ELECTRIC VEHICLE: An electric vehicle (EV) is a vehicle that uses one or more electric motors for propulsion. The vehicle can be powered by a collector system, with electricity from extravehicular sources, or can be powered autonomously by a battery or by converting fuel to electricity using a generator or fuel cells. EVs include road and rail vehicles, electric boats and underwater vessels, electric aircraft and electric spacecraft.

How Do All-Electric Cars Work?

All-electric vehicles, also referred to as battery electric vehicles (BEVs), have an electric motor instead of an internal combustion engine. The vehicle uses a large traction battery pack to power the electric motor and must be plugged into a wall outlet or charging equipment, also called electric vehicle supply equipment (EVSE). Because it runs on electricity, the vehicle emits no exhaust from a tailpipe and does not contain the typical liquid fuel components, such as a fuel pump, fuel line, or fuel tank.

BENEFITS OF ELECTRIC VEHICLES

Transport is a fundamental requirement of modern life, but the traditional combustion engine is quickly becoming outdated. Petrol or diesel vehicles are highly polluting and are being quickly replaced by fully electric vehicles. Fully electric vehicles (EV) have zero tailpipe emissions and are much better for the environment.

Lower running costs

The running cost of an electric vehicle is much lower than an equivalent petrol or diesel vehicle. Electric vehicles use electricity to charge their batteries instead of using fossil fuels like petrol or diesel. Electric vehicles are more efficient, and that combined with the electricity cost means that charging an electric vehicle is cheaper than filling petrol or diesel for your travel requirements. Using renewable energy sources can make the use of electric

vehicles more eco-friendly. The electricity cost can be reduced further if charging is done with the help of renewable energy sources installed at home, such as solar panels.

Low maintenance cost

Electric vehicles have very low maintenance costs because they don't have as many moving parts as an internal combustion vehicle. The servicing requirements for electric vehicles are lesser than the conventional petrol or diesel vehicles. Therefore, the yearly cost of running an electric vehicle is significantly low.

Zero Tailpipe Emissions

Driving an electric vehicle can help us reduce our carbon footprint because there will be zero tailpipe emissions. We can reduce the environmental impact of charging our vehicles further by choosing renewable energy options for home electricity.

Tax and financial benefits

Registration fees and road tax on purchasing electric vehicles are lesser than petrol or diesel vehicles. There are multiple policies and incentives offered by the government depending on which state we are in.

Petrol and diesel use is destroying our planet

The availability of fossil fuels is limited, and their use is destroying our planet. Toxic emissions from petrol and diesel vehicles lead to long-term, adverse effects on public health. The emissions impact of electric vehicles is much lower than petrol or diesel vehicles. From an efficiency perspective, electric vehicles can convert around 60% of the electrical energy from the grid to power the wheels, but petrol or diesel cars can only convert 17%-21% of the energy stored in the fuel to the wheels. That is a waste of around 80%. Fully electric vehicles have zero tailpipe emissions, but even when electricity production is taken into account, petrol or diesel vehicles emit almost 3 times more carbon dioxide than the average EV. To reduce the impact of charging electric vehicles, India is ambitious to achieve about 40 percent cumulative electric power installed capacity from non-fossil fuel-based energy resources by the year 2030. Therefore, electric vehicles are the way forward for Indian transport, and we must switch to them now.

Electric Vehicles are easy to drive and quiet

Electric vehicles don’t have gears and are very convenient to drive. There are no complicated controls, just accelerate, brake, and steer. When you want to charge your vehicle, just plug it in to a home or public charger. Electric vehicles are also quiet, so they reduce noise pollution that traditional vehicles contribute to.

Convenience of charging at home

Imagine being at a busy fuel station during peak hours, and you are getting late to reach your workplace. These problems can easily be overcome with an electric vehicle. Simply plug your vehicle in at your home charger for 4-5 hours before you plan to go. If you are able to get a charger where you park at home, it is very convenient to plan your journeys in advance. What if you forget to plug in your machine someday? Then you can easily take the help of fast chargers or even battery swapping services if you are on a two-wheeler on the road.

No noise pollution

Electric vehicles have the silent functioning capability as there is no engine under the hood. No engine means no noise. The electric motor functions so silently that you need to peek into your instrument panel to check if it is ON. Electric vehicles are so silent that manufacturers have to add false sounds in order to make them safe for pedestrians.

Market segmentation is a marketing strategy that uses well-defined criteria to divide a brand's total addressable market share into smaller groups. Each group, or segment, shares common characteristics that enable the brand to create focused and targeted products, offers, and experiences.

In this market segmentation assignment, we'll use the concept of **Fermi estimation**. Fermi estimation empowers us to make informed decisions and verify claims in a time-efficient way, by breaking down these problems into manageable parts.

Segmentation Criteria

The term segmentation Criteria relates to the nature of the type of information used for market segmentation, unlike the segmentation variable which means the variable in empirical data in common sense segmentation for splitting the sample into market segments. In Segmentation, we usually find the identifiable characteristics of individuals in the data sample segment them into the same cluster and analyse the common interest needed to maximize the organization's profits. Segmentation Criteria are an important factor in market segmentation as well. The four main types of segmentation criteria are Geographic

Segmentation, socio-demographic segmentation, psychographic segmentation and behavioural segmentation.

Geographic Segmentation

In Geographic Segmentation the key criteria to form market segments is the geographic location or the residence of the customer. There are some specific advantages of doing geographic segmentation, they are, we can segment down all the customers in that particular area, do promotions which are meaningful in that area and even run ads in newspapers, television, etc. in that local area. The only key disadvantage is that it is not always the case that all the people residing in the same location will have the same opinions and preferences regarding the products.

Socio-Demographic Segmentation

Socio-demographic segmentation criteria include parameters like age, gender, education, income, etc. Forex, while buying cosmetics criteria associated is gender, while buying branded and luxury items criteria associated is income, while planning on vacation destination criteria associated is age (i.e., if people go in couple the vacation destination will be different if people going with children, then the vacation destination is different). The socio-demographic segmentation at times with better data can give us better market segments and give us clear clarity on who the customer is, this is achievable provided better data that provides sufficient insights about who the customer is and the market segments. However, in many cases, socio-demographic segmentation would not be the best fit for product preferences.

Psychographic Segmentation

For making market segments using the Psychographic segmentation the criteria is the Psychological criteria for grouping people. Parameters like interests, beliefs, aspirations, preferences, benefits, etc. can be used to define psychological criteria. Psychographic segmentation is more complex by nature compared to Geographic Segmentation and Socio-Demographic segmentation because, we cannot find a single fixed parameter for insights for better segmentation, there are a lot of factors affecting the psychographic criteria and the factors are different in each person. Therefore, we must use a lot of segmentation variables. And the main advantage that psychographic segmentation has is that clustering a common set of customers based on psychographic criteria for maximizing profits. For ex, people who want to go on a vacation and have a preference for attending historic pilgrims can be clustered and can be taken together which can reduce costs for the company and maximize the profit as well.

Behavioural Segmentation

In Behavioural segmentation, we can directly find similarities in the behaviours of customers. There can be many useful implementations possible for doing market segments.

Behavioural segmentation criteria depend on the way visitors interact with the website. Some data depends on their immediate online behaviour and giving positive feedback while other data depends on their past offline behaviour or negative feedback.

DATA COLLECTION: Data is collected manually from www.kaggle.com and https://www.data.gov.in

IMPLEMENTATION

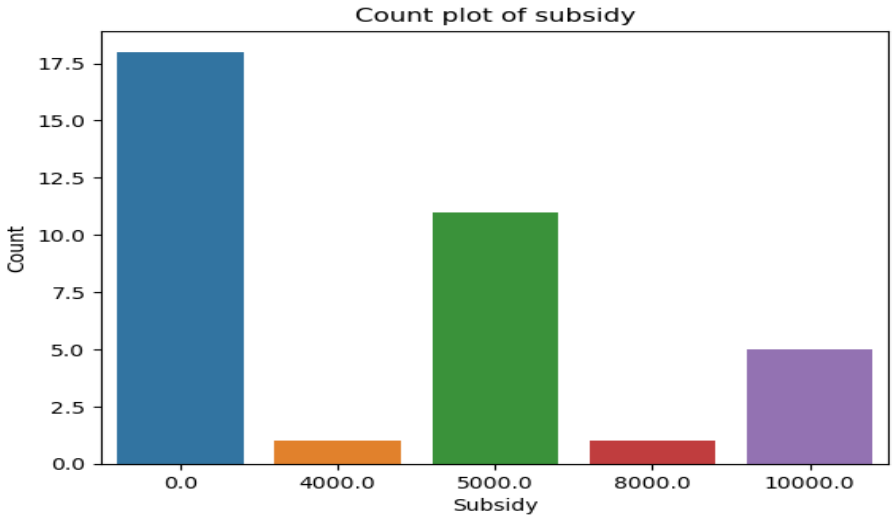
- Packages/tools used: In PYTHON, we used NumPy, Pandas, Sklearn, Matplotlib, Plotly, Statsmodels.

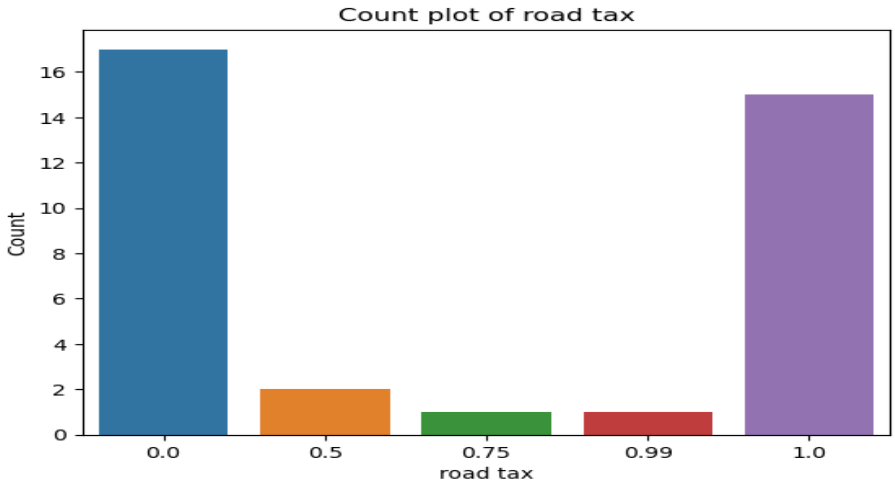
- Pre-processing data before performing Segmentation

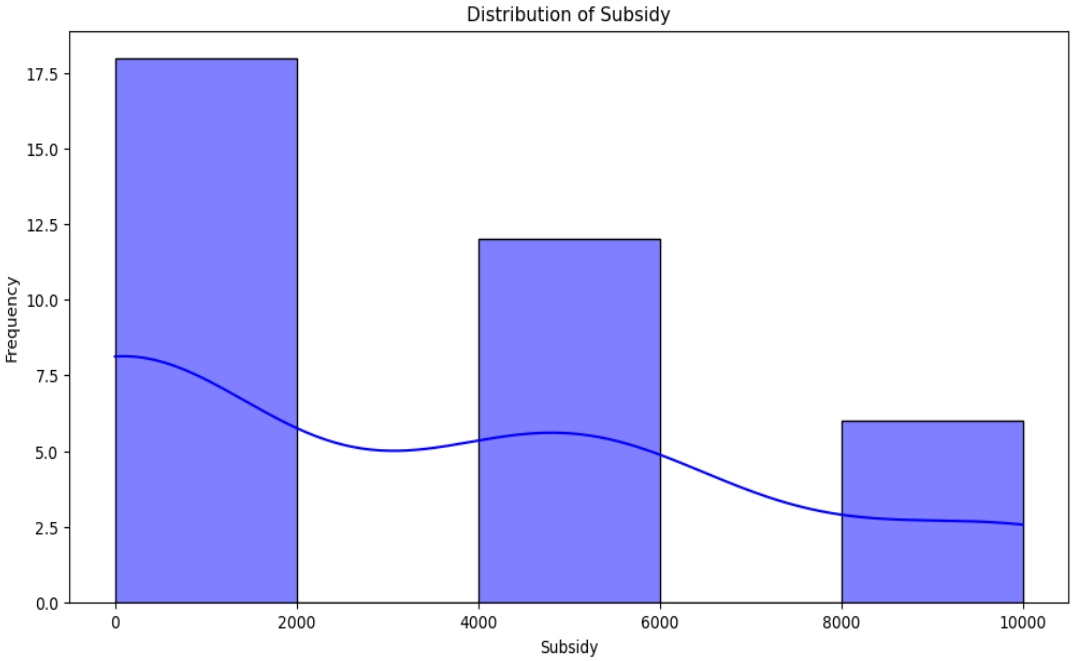
    - Data cleaning: Data cleaning is the process of fixing or removing incorrect, corrupted, incorrectly formatted, duplicate, or incomplete data within a dataset. When combining multiple data sources, there are many opportunities for data to be duplicated or mislabeled.

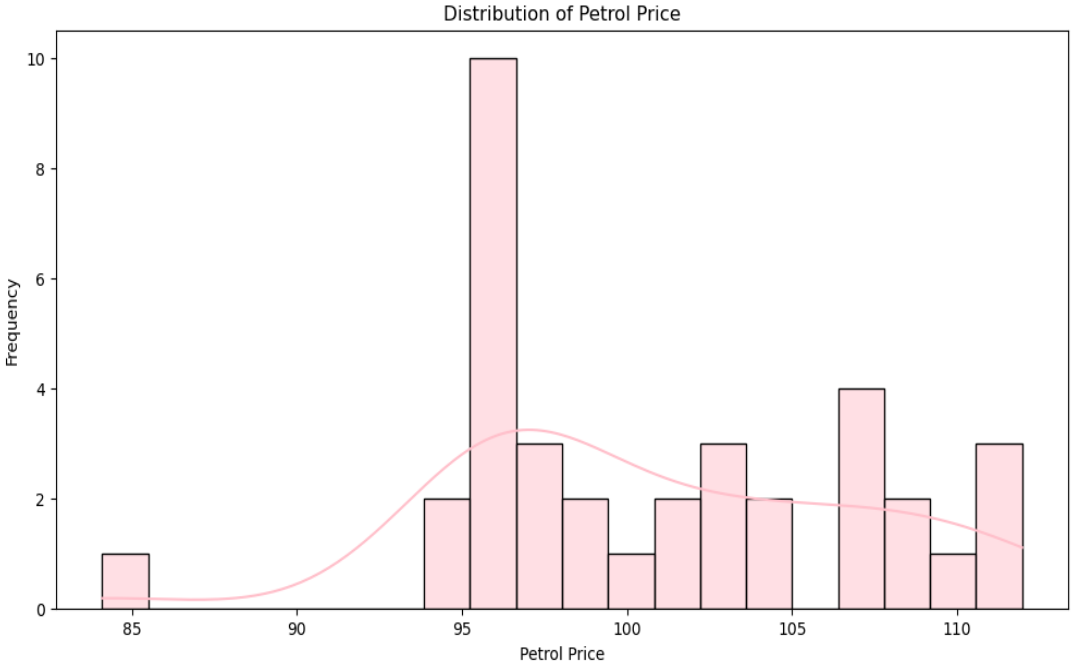
    - EDA:
      We start the Exploratory Data Analysis with some data Analysis drawn from the data without Principal Component Analysis and with some Principal Component Analysis in the dataset obtained from the combination of all the data we have. PCA is a statistical process that converts the observations of correlated features into a set of linearly uncorrelated features with the help of orthogonal transformation. These new transformed features are called the Principal Components. The process helps in reducing dimensions of the data to make the process of classification/regression or any form of machine learning, cost-effective.

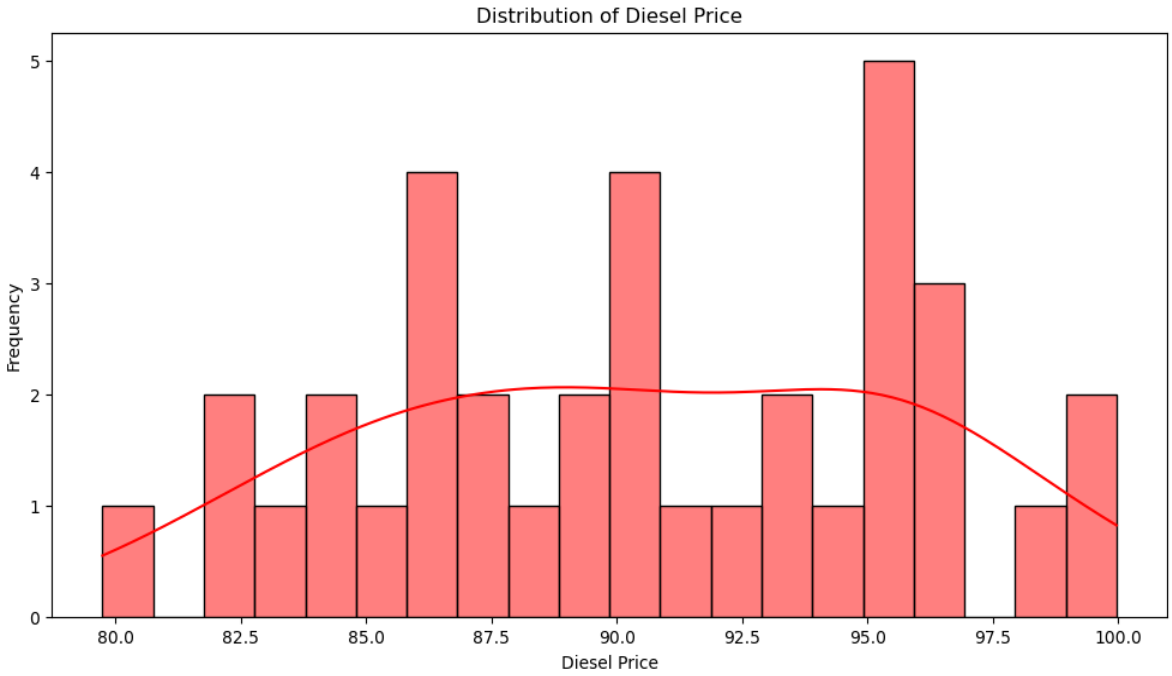
Following are the findings of the Exploratory data analysis from our first dataset “statesdataEVproject.csv” :

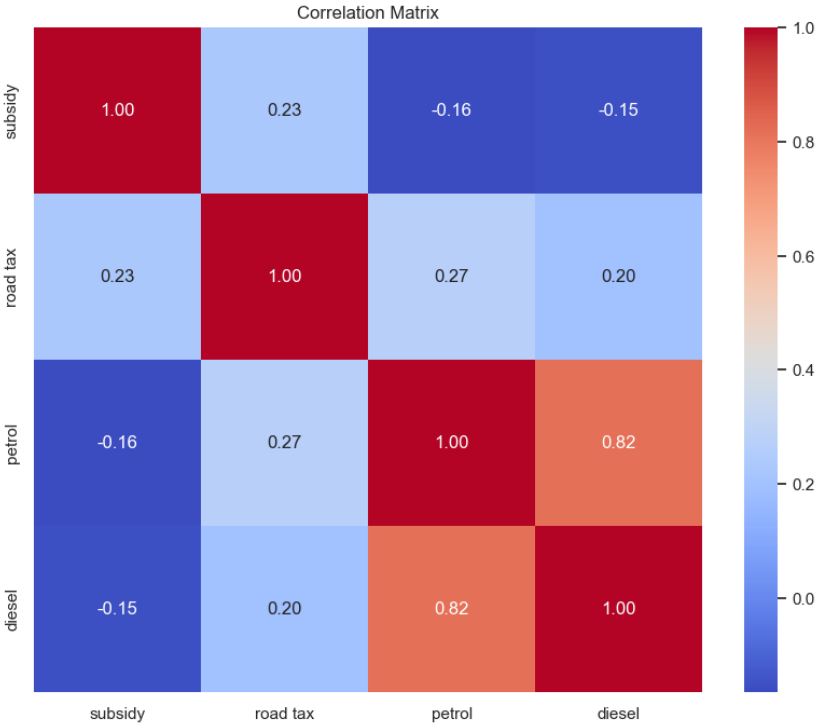


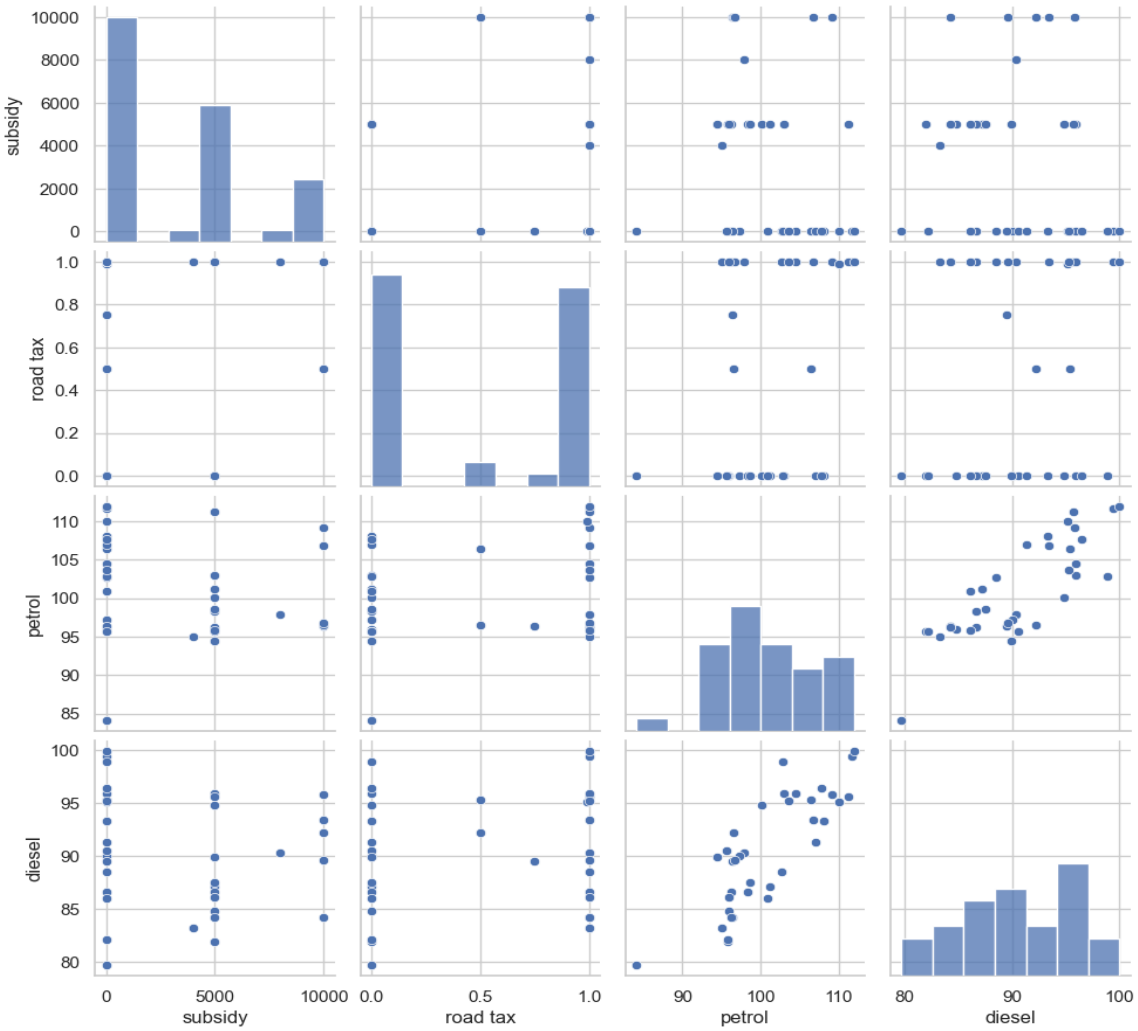



Distribution of Subsidy
Frequency
Subsidy
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
0
2000
4000
6000
8000
10000

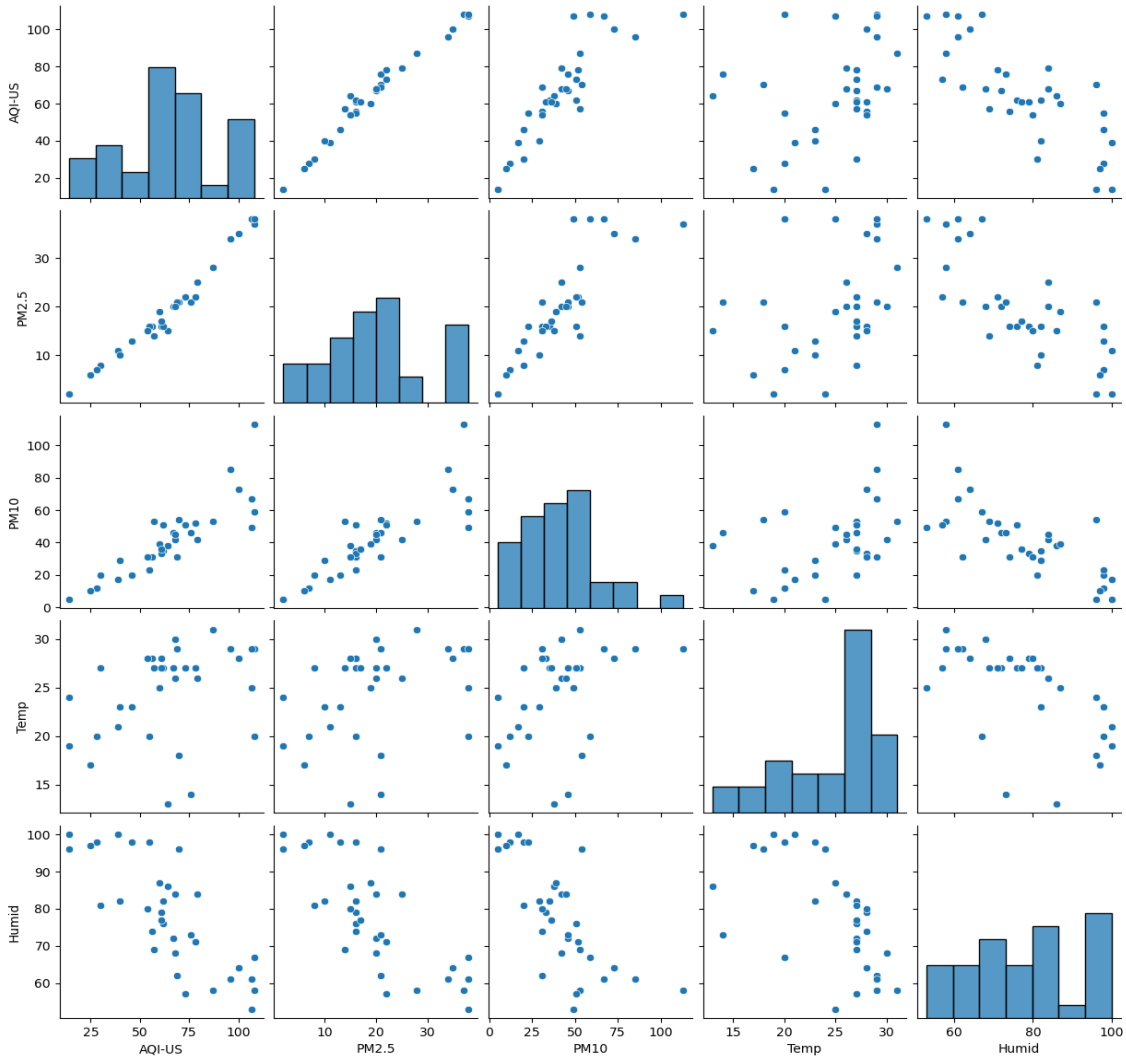

Distribution of Petrol Price
Frequency
Petrol Price
0
2
4
6
8
10
85
90
95
100
105
110

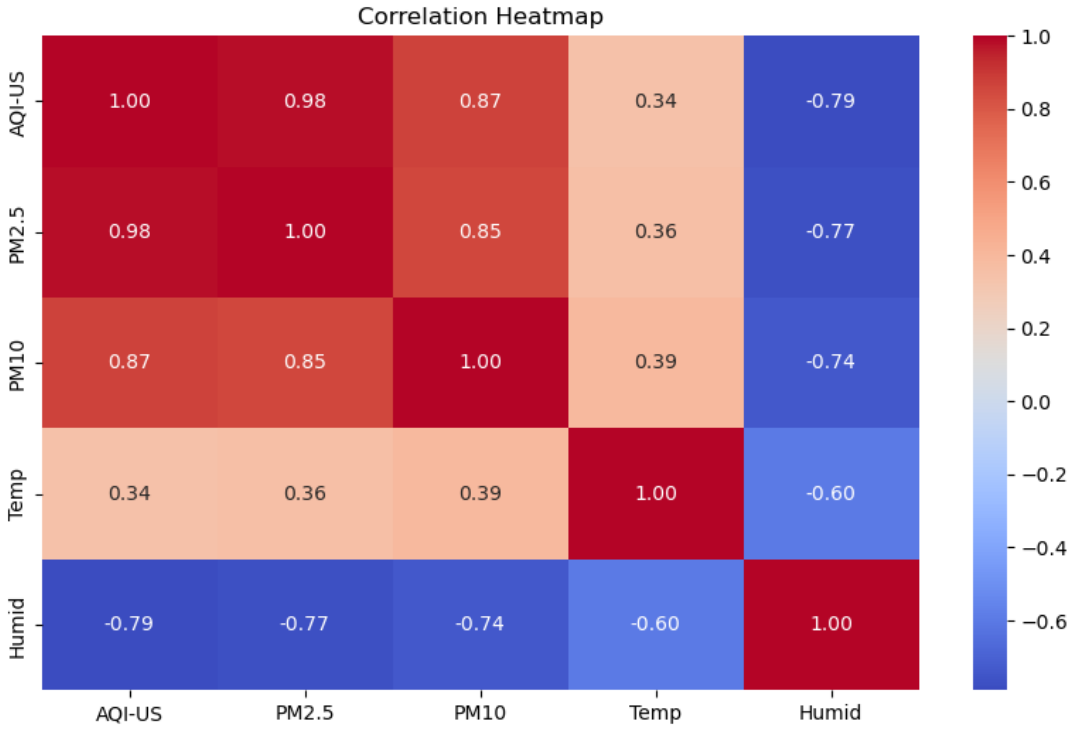

Distribution of Diesel Price
Frequency
Diesel Price
0
1
2
3
4
5
80.0
82.5
85.0
87.5
90.0
92.5
95.0
97.5
100.0

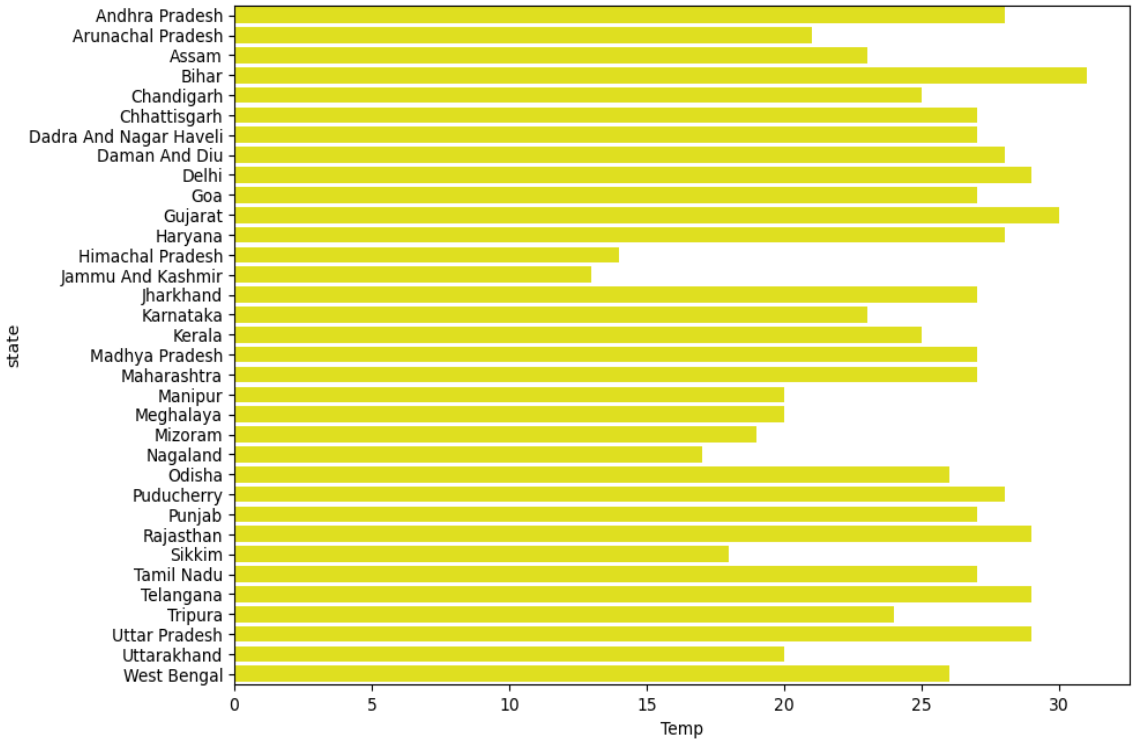

Correlation Matrix
subsidy
road tax
petrol
diesel
1.00
0.23
-0.16
-0.15
0.23
1.00
0.27
0.20
-0.16
0.27
1.00
0.82
-0.15
0.20
0.82
1.00
1.0
0.8
0.6
0.4
0.2
0.0

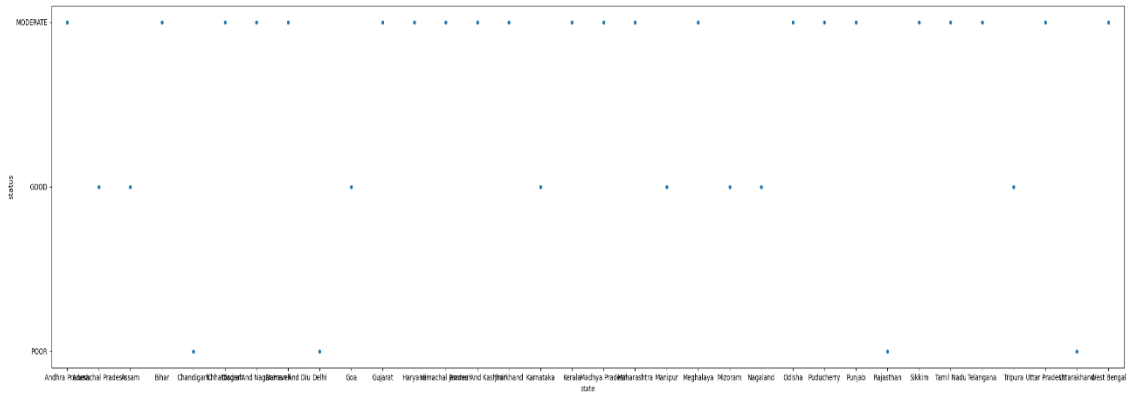

subsidy
road tax
petrol
diesel
0
5000
10000
0.0
0.5
1.0
90
100
110
80
90
100

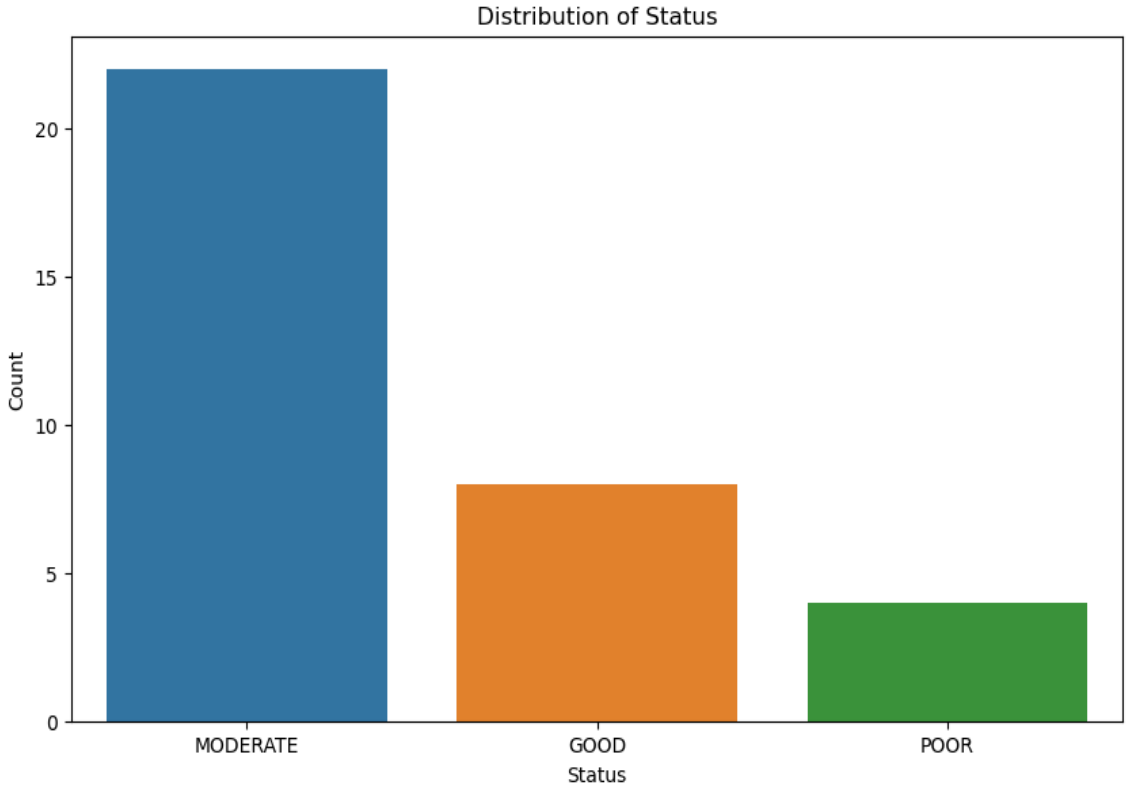
Following are the findings of the Exploratory data analysis from our second dataset “pollutiondataEVproject.csv” :



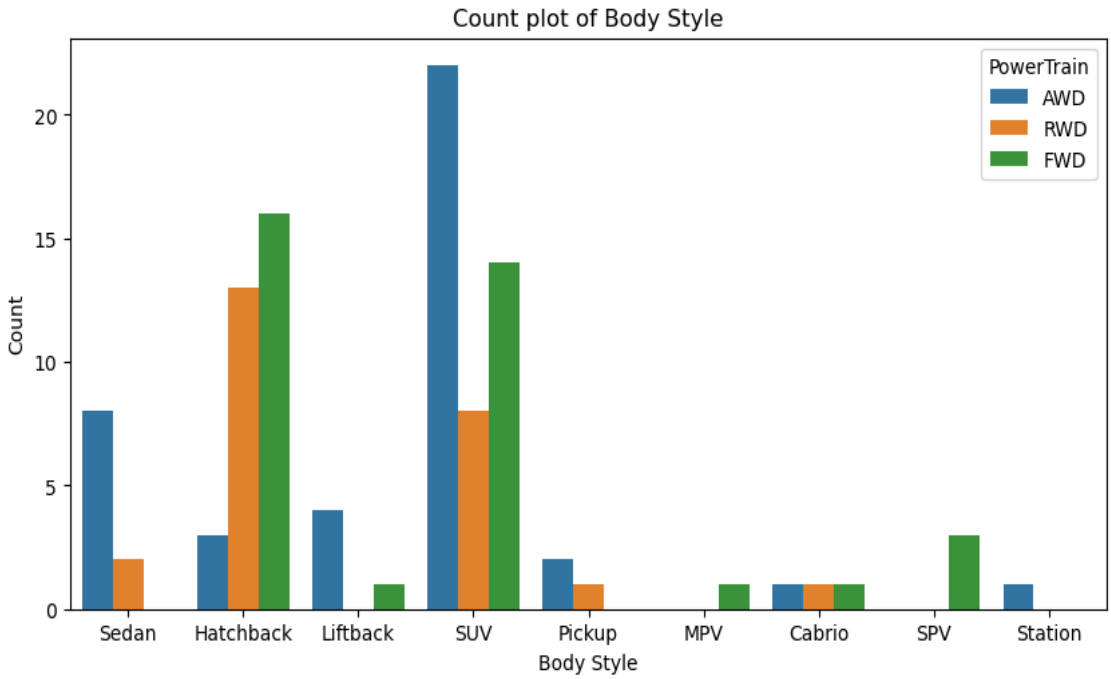

Correlation Heatmap
AQI-US
PM2.5
PM10
Temp
Humid
1.00 0.98 0.87 0.34 -0.79
0.98 1.00 0.85 0.36 -0.77
0.87 0.85 1.00 0.39 -0.74
0.34 0.36 0.39 1.00 -0.60
-0.79 -0.77 -0.74 -0.60 1.00
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

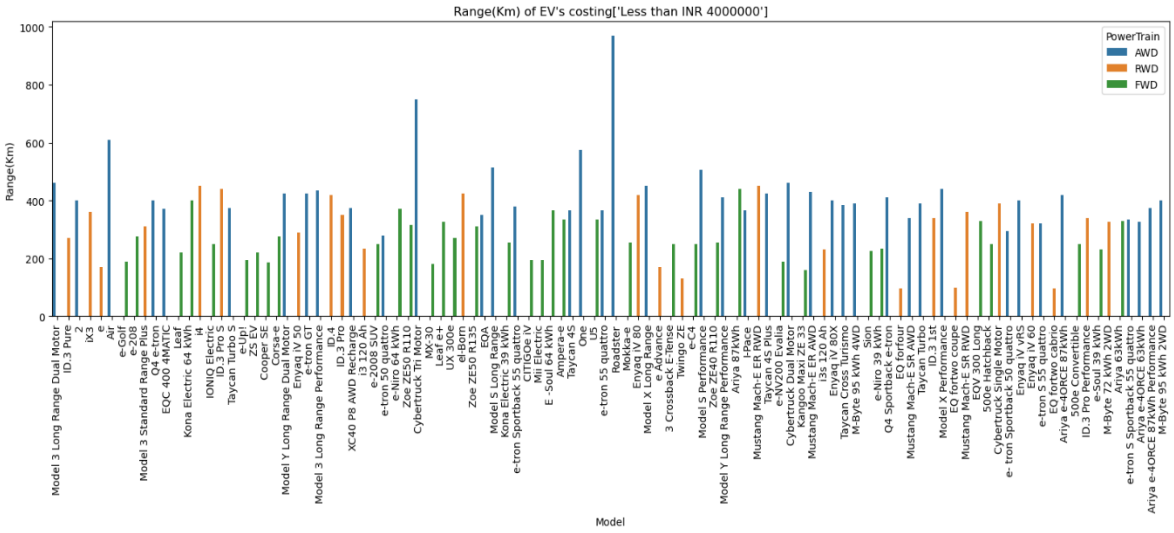

state
Andhra Pradesh
Arunachal Pradesh
Assam
Bihar
Chandigarh
Chhattisgarh
Dadra And Nagar Haveli
Daman And Diu
Delhi
Goa
Gujarat
Haryana
Himachal Pradesh
Jammu And Kashmir
Jharkhand
Karnataka
Kerala
Madhya Pradesh
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Odisha
Puducherry
Punjab
Rajasthan
Sikkim
Tamil Nadu
Telangana
Tripura
Uttar Pradesh
Uttarakhand
West Bengal
0
5
10
15
20
25
30
Temp

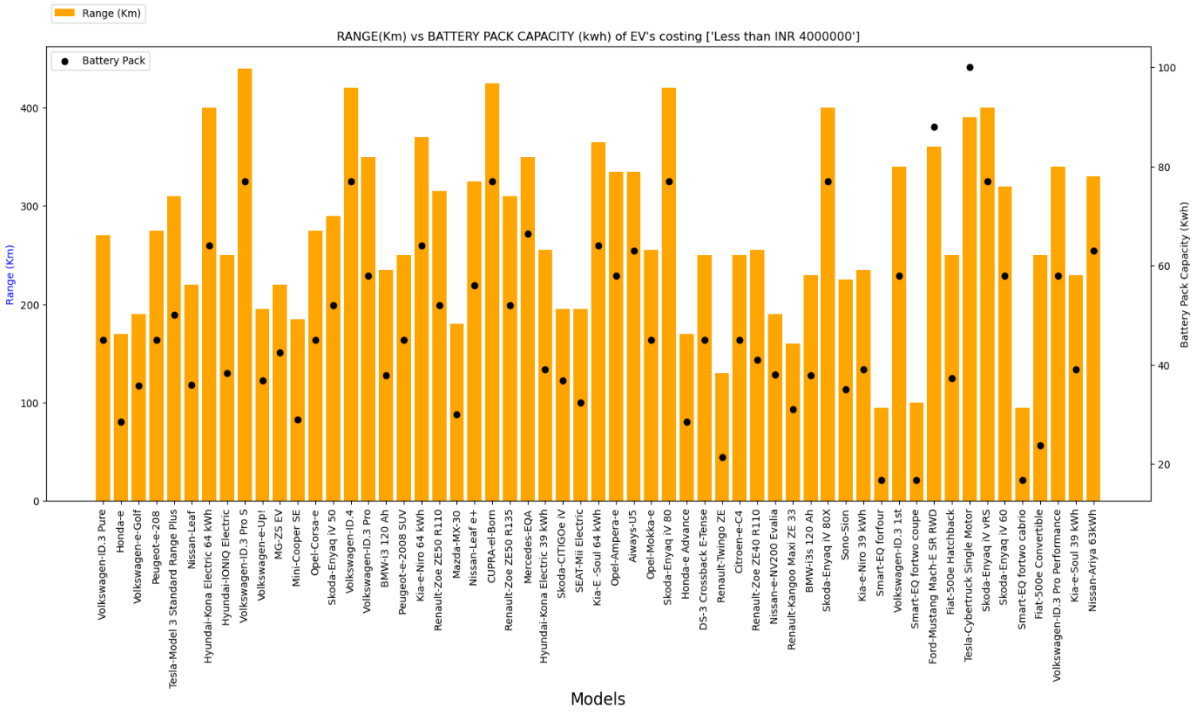

MODERATE
GOOD
POOR
status
Andhra Pradesh
Assam
Bihar
Chandigarh
Delhi
Goa
Gujarat
Haryana
Karnataka
Kerala
Maharashtra
Manipur
Meghalaya
Mizoram
Nagaland
Odisha
Puducherry
Punjab
Rajasthan
Sikkim
Tamil Nadu
Telangana
Tripura
Uttar Pradesh
Uttarakhand
West Bengal
state

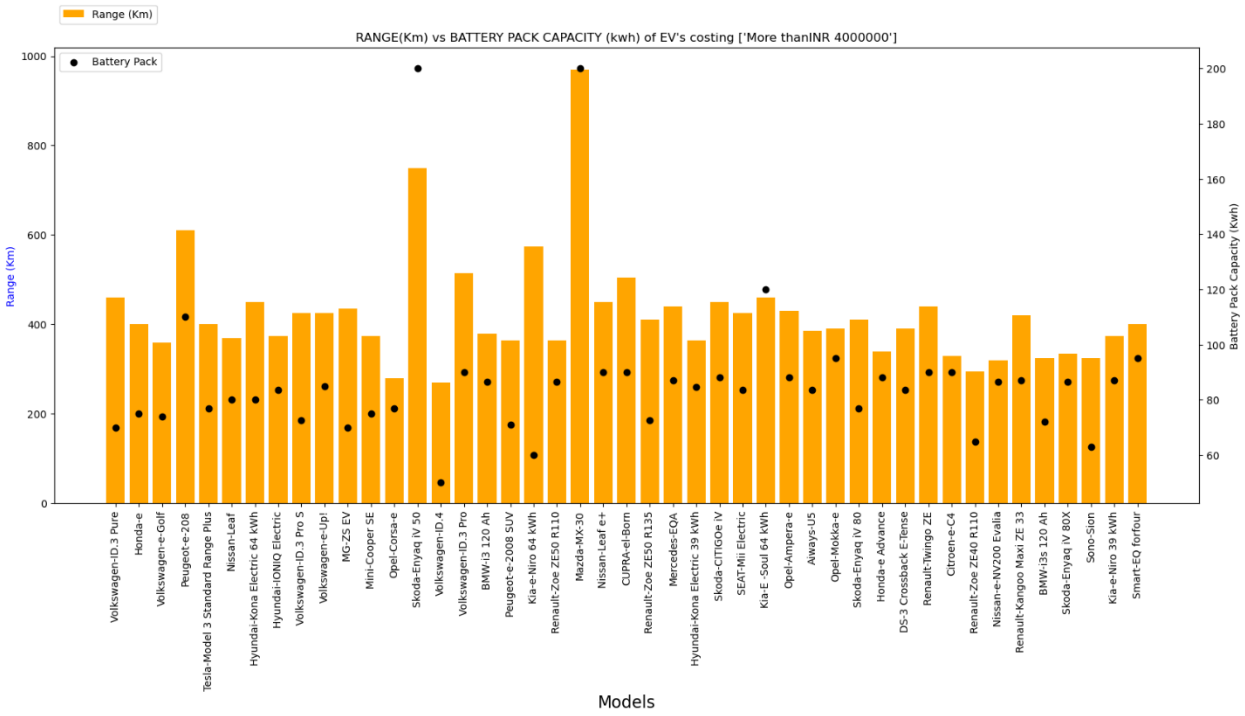

Distribution of Status
Count
0
5
10
15
20
MODERATE
GOOD
POOR
Status

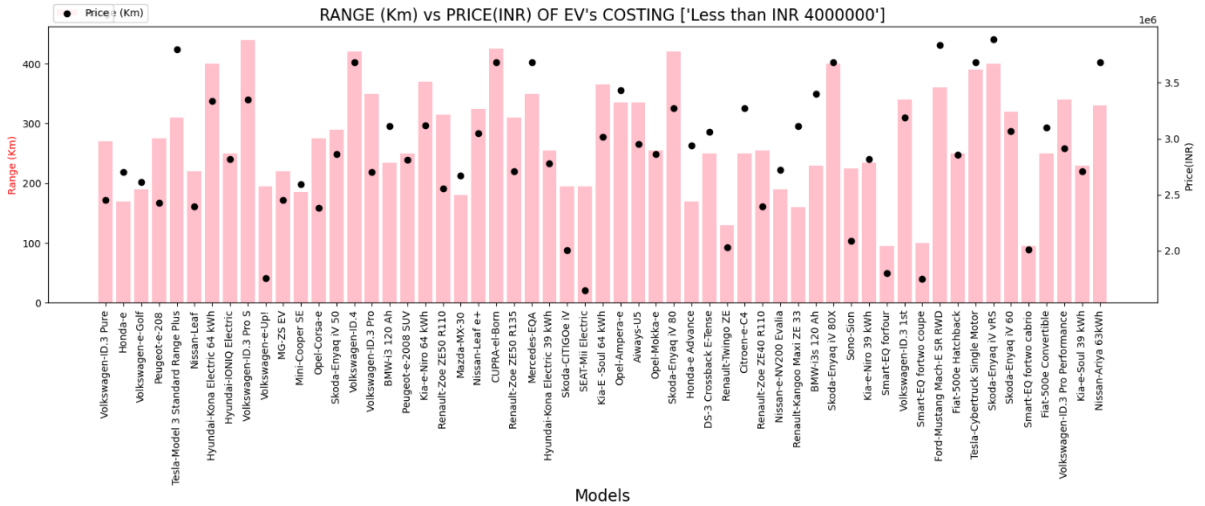
Following are the findings of the Exploratory data analysis from our third dataset “electriccardataEVproject.csv” :

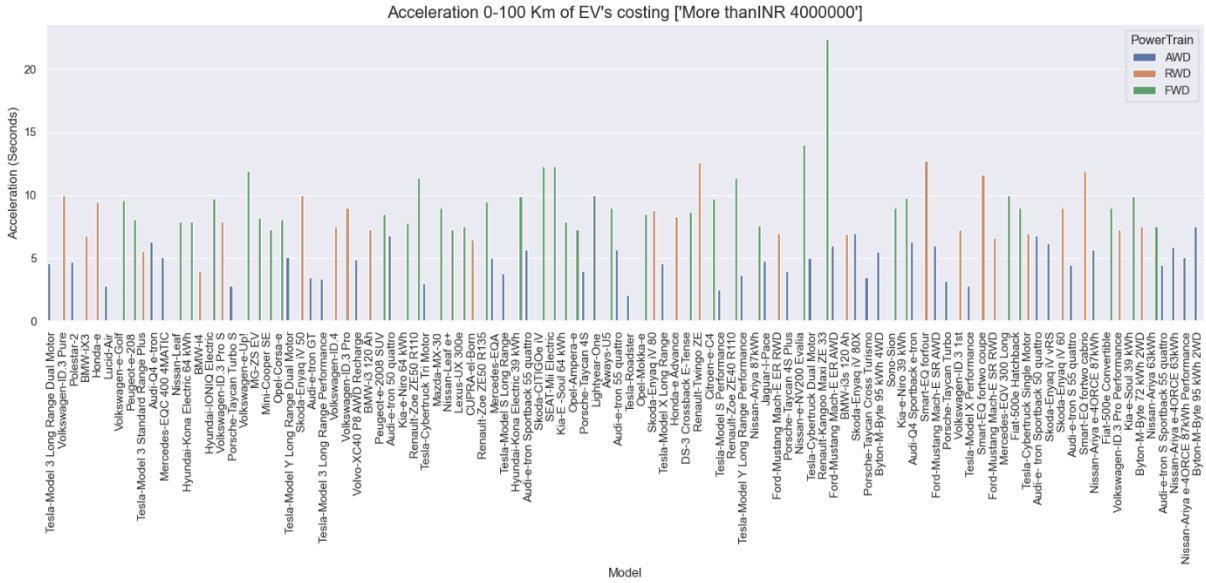


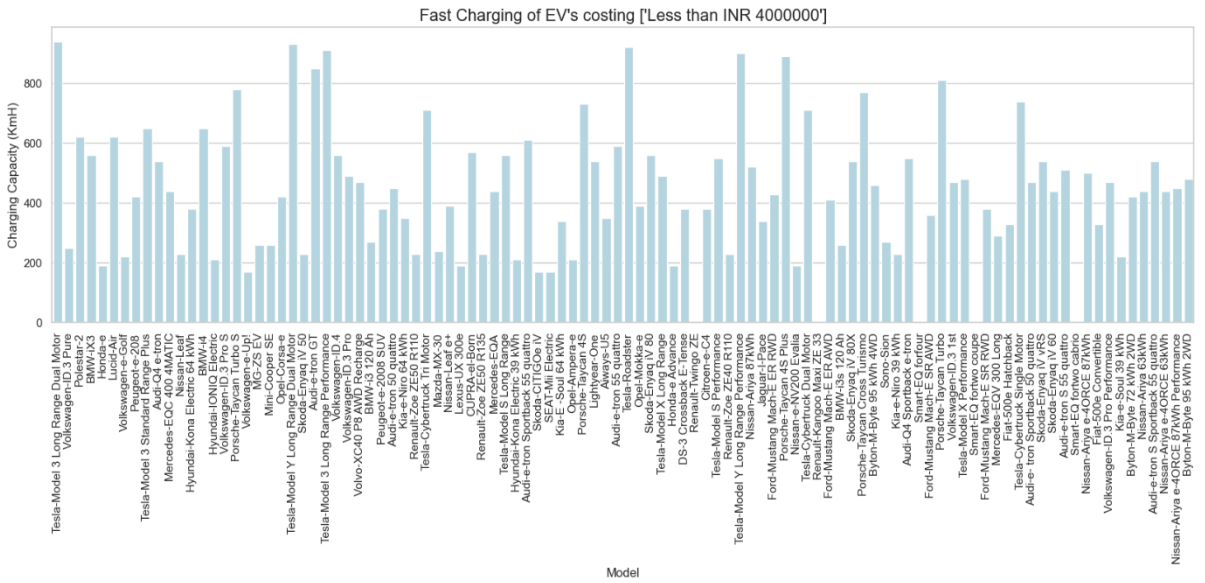


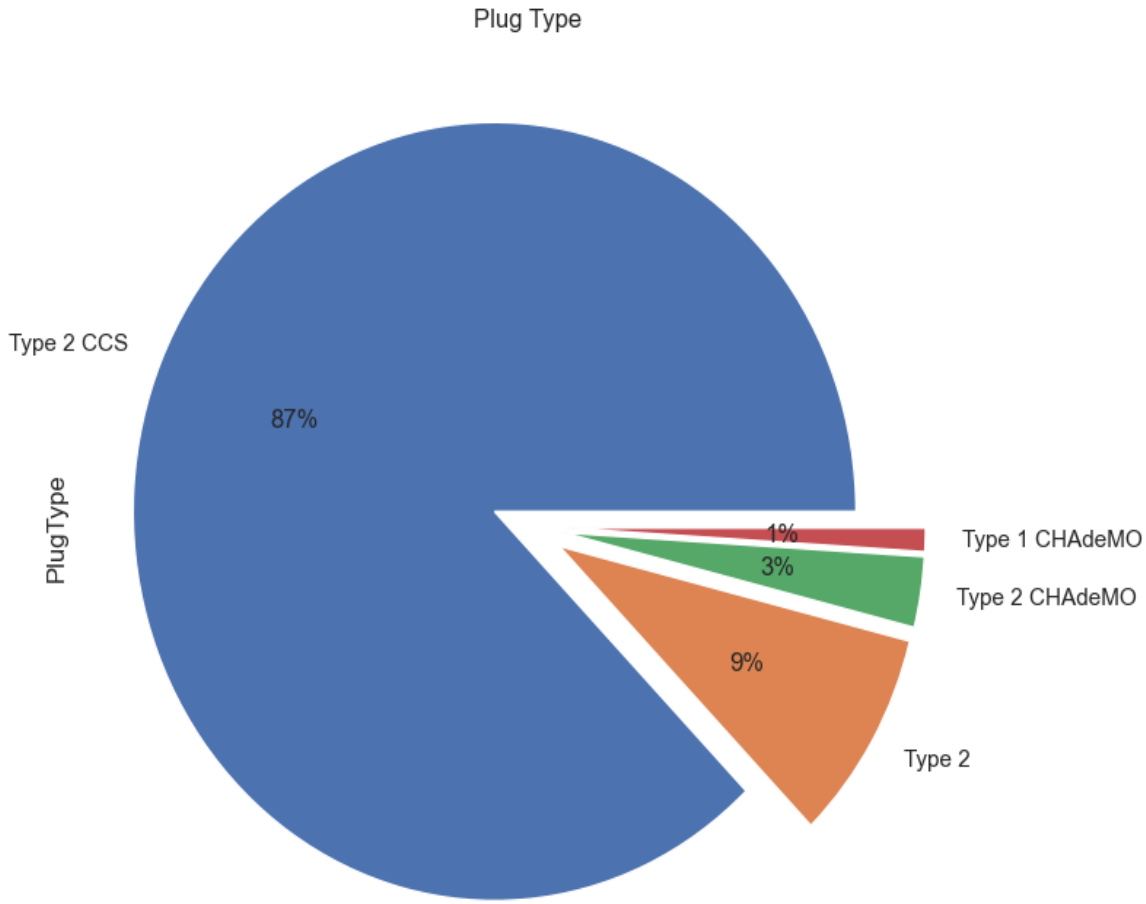

Range (Km)
Battery Pack
RANGE(Km) vs BATTERY PACK CAPACITY (kwh) of EV's costing ['Less than INR 4000000']
Range (Km)
Battery Pack Capacity (Kwh)
Volkswagen-ID.3 Pure
Honda-e
Volkswagen-e-Golf
Peugeot-e-208
Tesla-Model 3 Standard Range Plus
Nissan-Leaf
Hyundai-Kona Electric 64 kWh
Hyundai-IONIQ Electric
Volkswagen-ID.3 Pro S
Volkswagen-e-Up!
MG-ZS EV
Mini-Cooper SE
Opel-Corsa-e
Skoda-Enyaq iV 50
Volkswagen-ID.4
Volkswagen-ID.3 Pro
BMW-i3 120 Ah
Peugeot-e-2008 SUV
Kia-e-Niro 64 kWh
Renault-Zoe ZE50 R110
Mazda-MX-30
Nissan-Leaf e+
CUPRA-el-Born
Renault-Zoe ZE50 R135
Mercedes-EQA
Hyundai-Kona Electric 39 kWh
Skoda-CITIGOe iV
SEAT-Mii Electric
Kia-E -Soul 64 kWh
Opel-Ampera-e
Aiways-U5
Opel-Mokka-e
Skoda-Enyaq iV 80
Honda-e Advance
DS-3 Crossback E-Tense
Renault-Twingo ZE
Citroen-e-C4
Renault-Zoe ZE40 R110
Nissan-e-NV200 Evalia
Renault-Kangoo Maxi ZE 33
BMW-i3s 120 Ah
Skoda-Enyaq iV 80X
Sono-Sion
Kia-e-Niro 39 kWh
Smart-EQ forfour
Volkswagen-ID.3 1st
Smart-EQ fortwo coupe
Ford-Mustang Mach-E SR RWD
Fiat-500e Hatchback
Tesla-Cybertruck Single Motor
Skoda-Enyaq iV vRS
Skoda-Enyaq iV 60
Smart-EQ fortwo cabrio
Fiat-500e Convertible
Volkswagen-ID.3 Pro Performance
Kia-e-Soul 39 kWh
Nissan-Ariya 63kWh
Models

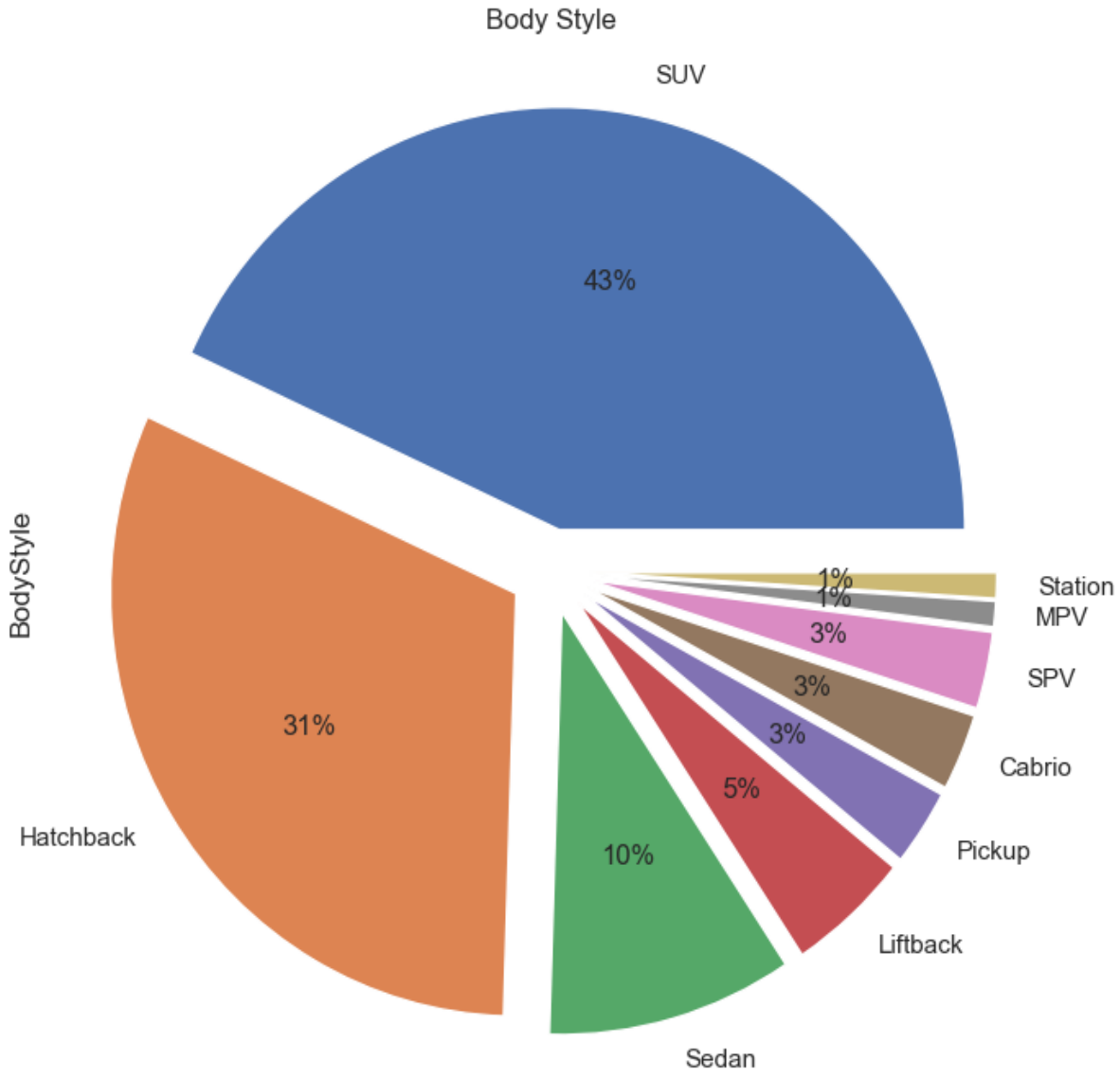

Range (Km)
Battery Pack
RANGE(Km) vs BATTERY PACK CAPACITY (kwh) of EV's costing ['More thanINR 4000000']
Range (Km)
Battery Pack Capacity (Kwh)
Volkswagen-ID.3 Pure
Honda-e
Volkswagen-e-Golf
Peugeot-e-208
Tesla-Model 3 Standard Range Plus
Nissan-Leaf
Hyundai-Kona Electric 64 kWh
Hyundai-IONIQ Electric
Volkswagen-ID.3 Pro S
Volkswagen-e-Up!
MG-ZS EV
Mini-Cooper SE
Opel-Corsa-e
Skoda-Enyaq iV 50
Volkswagen-ID.4
Volkswagen-ID.3 Pro
BMW-i3 120 Ah
Peugeot-e-2008 SUV
Kia-e-Niro 64 kWh
Renault-Zoe ZE50 R110
Mazda-MX-30
Nissan-Leaf e+
CUPRA-el-Born
Renault-Zoe ZE50 R135
Mercedes-EQA
Hyundai-Kona Electric 39 kWh
Skoda-CITIGOe iV
SEAT-Mii Electric
Kia-E -Soul 64 kWh
Opel-Ampera-e
Aiways-U5
Opel-Mokka-e
Skoda-Enyaq iV 80
Honda-e Advance
DS-3 Crossback E-Tense
Renault-Twingo ZE
Citroen-e-C4
Renault-Zoe ZE40 R110
Nissan-e-NV200 Evalia
Renault-Kangoo Maxi ZE 33
BMW-i3s 120 Ah
Skoda-Enyaq iV 80X
Sono-Sion
Kia-e-Niro 39 kWh
Smart-EQ forfour
Models

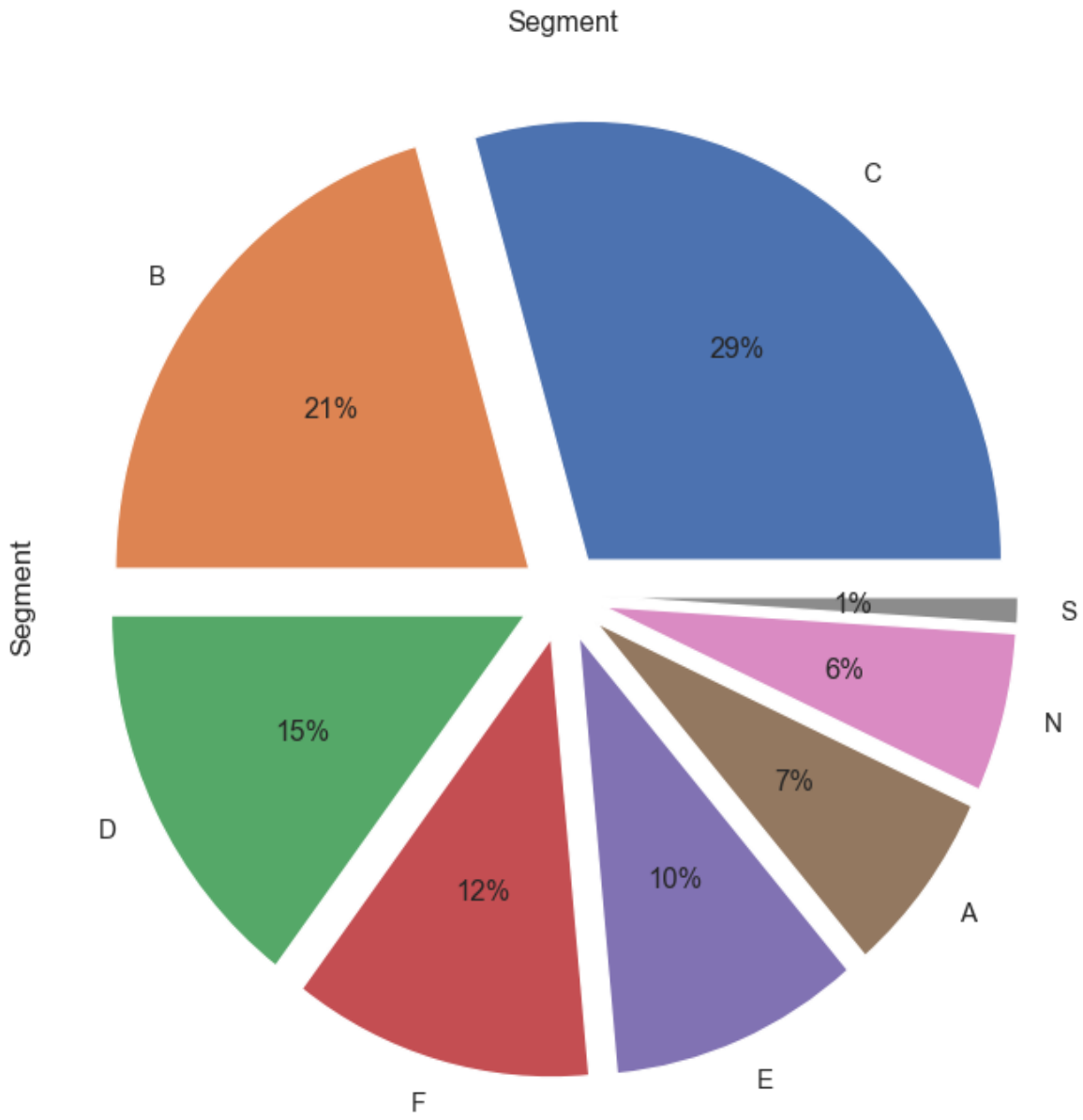

RANGE (Km) vs PRICE(INR) OF EV's COSTING ['Less than INR 4000000']
Price (Km)
Range (Km)
Price(INR)
1e6
Models

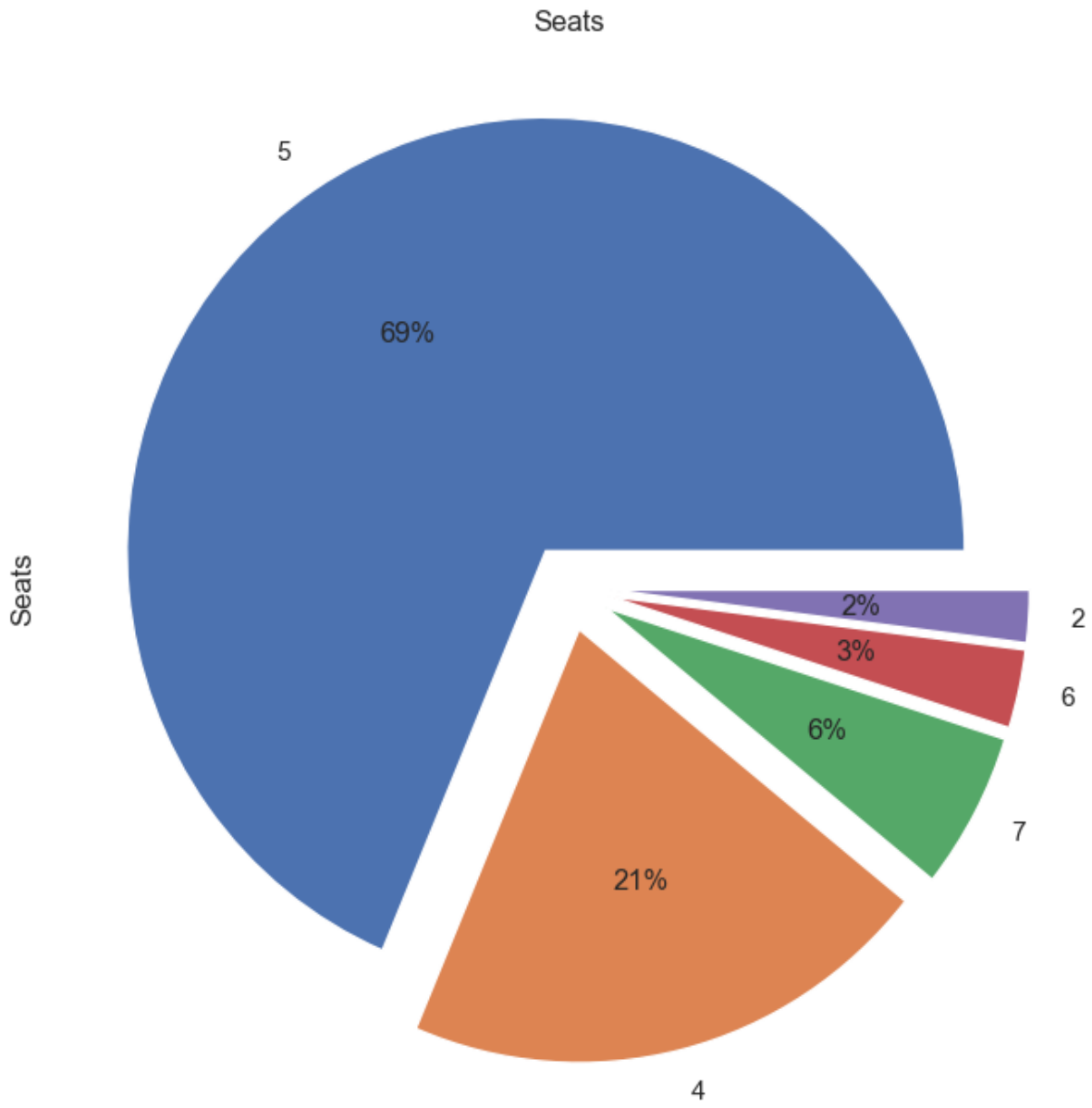

Acceleration 0-100 Km of EV's costing ['More thanINR 4000000']
PowerTrain
AWD
RWD
FWD
Acceleration (Seconds)
Model

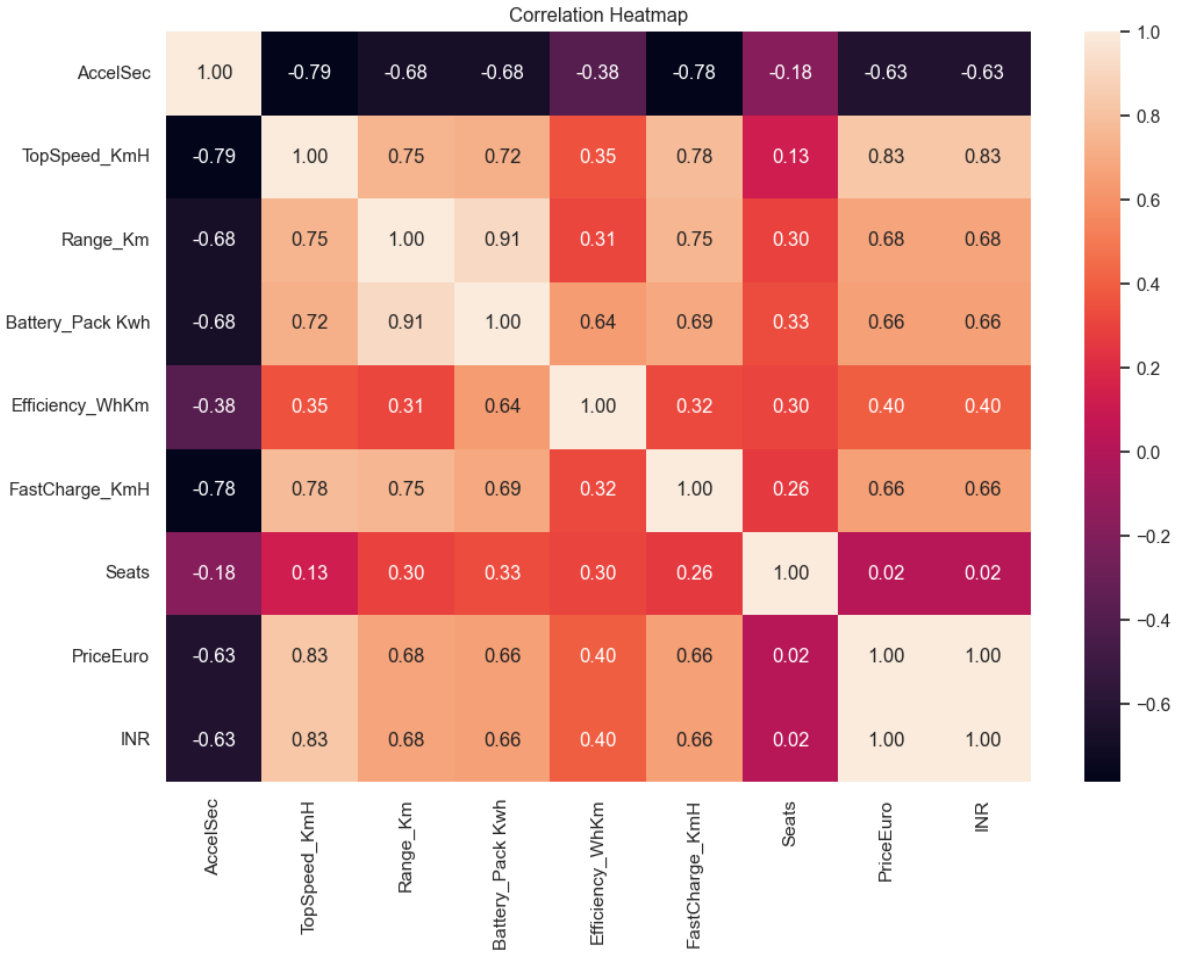

Fast Charging of EV's costing ['Less than INR 4000000']
Charging Capacity (KmH)
Model


Plug Type
PlugType
Type 2 CCS
87%
1%
Type 1 CHAdeMO
3%
Type 2 CHAdeMO
9%
Type 2


Body Style
SUV
43%
BodyStyle
31%
Hatchback
10%
Sedan
5%
Liftback
3%
Pickup
3%
Cabrio
3%
SPV
1%
MPV
1%
Station

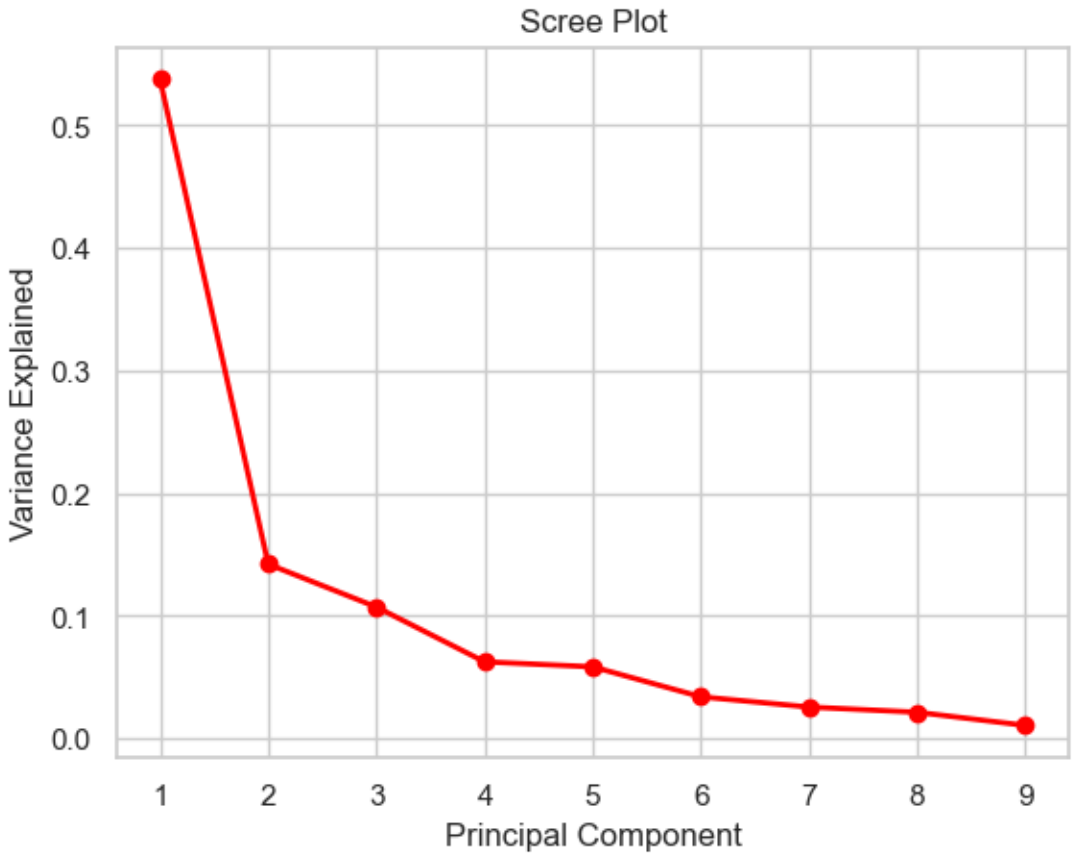

Segment
Segment
C
29%
B
21%
S
1%
N
6%
A
7%
E
10%
F
12%
D
15%


Seats
Seats
5
69%
2%
2
3%
6
6%
7
21%
4


Correlation Heatmap
AccelSec 1.00 -0.79 -0.68 -0.68 -0.38 -0.78 -0.18 -0.63 -0.63
TopSpeed_KmH -0.79 1.00 0.75 0.72 0.35 0.78 0.13 0.83 0.83
Range_Km -0.68 0.75 1.00 0.91 0.31 0.75 0.30 0.68 0.68
Battery_Pack Kwh -0.68 0.72 0.91 1.00 0.64 0.69 0.33 0.66 0.66
Efficiency_WhKm -0.38 0.35 0.31 0.64 1.00 0.32 0.30 0.40 0.40
FastCharge_KmH -0.78 0.78 0.75 0.69 0.32 1.00 0.26 0.66 0.66
Seats -0.18 0.13 0.30 0.33 0.30 0.26 1.00 0.02 0.02
PriceEuro -0.63 0.83 0.68 0.66 0.40 0.66 0.02 1.00 1.00
INR -0.63 0.83 0.68 0.66 0.40 0.66 0.02 1.00 1.00
AccelSec
TopSpeed_KmH
Range_Km
Battery_Pack Kwh
Efficiency_WhKm
FastCharge_KmH
Seats
PriceEuro
INR
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6

Scree Plot: This is a common method for determining the number of PCs to be retained via graphical representation. It is a simple line segment plot that shows the eigenvalues for each individual PC. It shows the eigenvalues on the y-axis and the number of factors on the x-axis. It always displays a downward curve. Most scree plots look broadly similar in shape, starting high on the left, falling rather quickly, and then flattening out at some point. This is because the first component usually explains much of the variability, the next few components explain a moderate amount, and the latter components only explain a small fraction of the overall variability. The scree plot criterion looks for the "elbow" in the curve and selects all components just before the line flattens out. The proportion of variance plot: The selected PCs should be able to describe at least 80% of the variance.

```
PC_values = np.arange(pca.n_components_) + 1
plt.plot(PC_values, pca.explained_variance_ratio_, 'o-', linewidth=2, color='red')
plt.title('Scree Plot')
plt.xlabel('Principal Component')
plt.ylabel('Variance Explained')
plt.show()
```



EXTRACTING SEGMENTS:

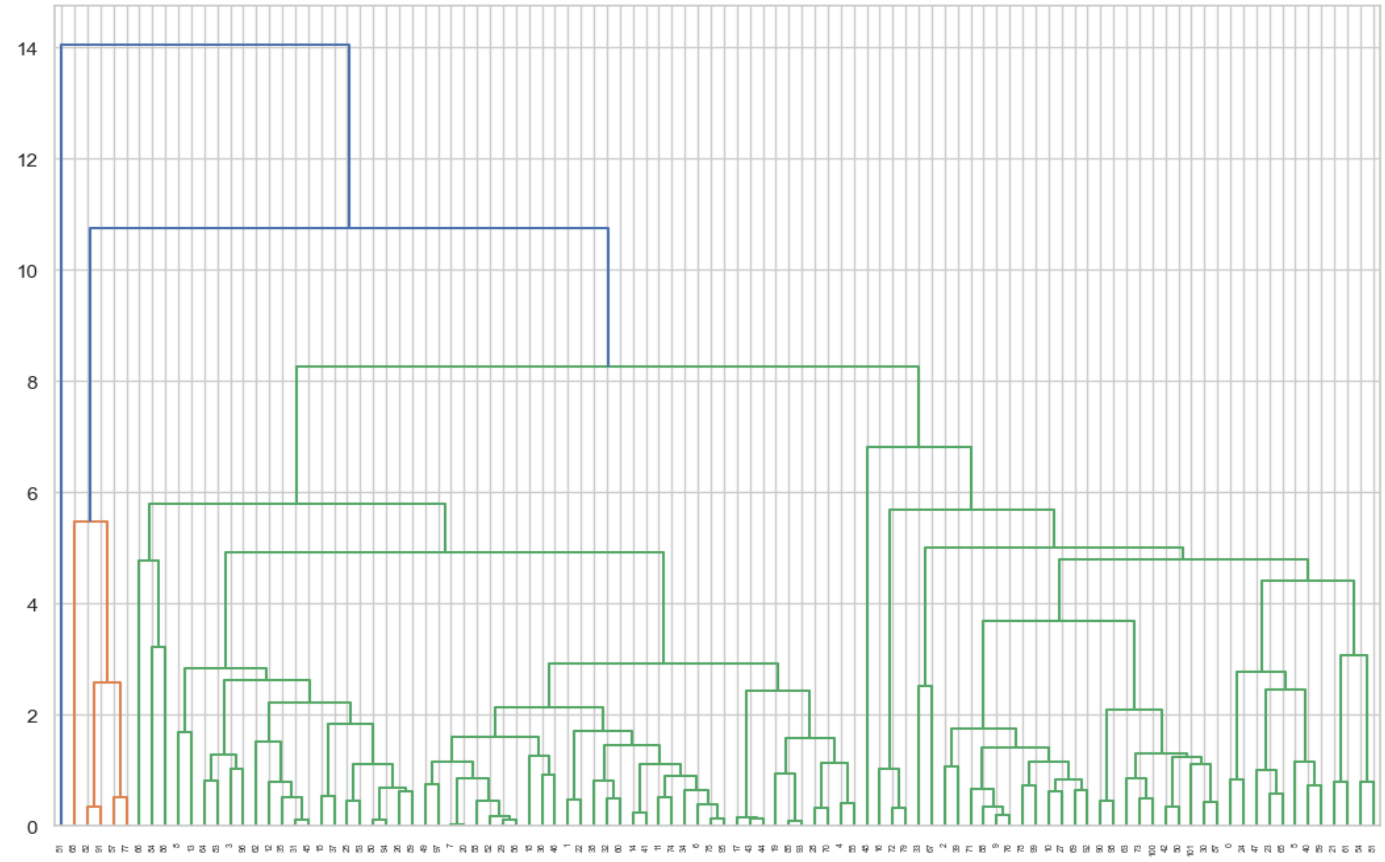
DENDROGRAM: A dendrogram is a diagram that shows the hierarchical relationship between objects. It is most commonly created as an output from hierarchical clustering. The main use of a dendrogram is to work out the best way to allocate objects to clusters. The dendrogram below shows the hierarchical clustering of six observations shown on the scatterplot to the left.

```
linked = linkage(data2, 'complete')

plt.figure(figsize=(13, 9))

dendrogram(linked, orientation='top')

plt.show()
```



Analysis and Approaches Used for Segmentation:

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

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.

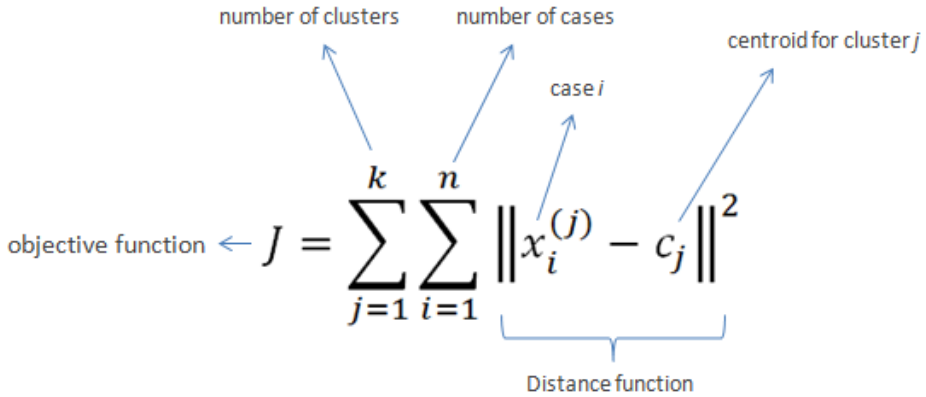
The way k means algorithm works is as follows:

• Specify the 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 k-means follows to solve the problem is the 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,



Applications

K means algorithm is very popular and used in a variety of applications such as market segmentation, document clustering, image segmentation and image compression, etc.

The goal usually when we undergo a cluster analysis is either:

1. Get a meaningful intuition of the structure of the data we're dealing with.

2. Cluster-then-predict where different models will be built for different subgroups if we believe there is a wide variation in the behaviours of different subgroups.

The k-means clustering algorithm performs the following tasks:

• Specify the number of clusters K

• Initialize centroids by first shuffling the dataset and then randomly selecting K data points for the centroids without replacement.

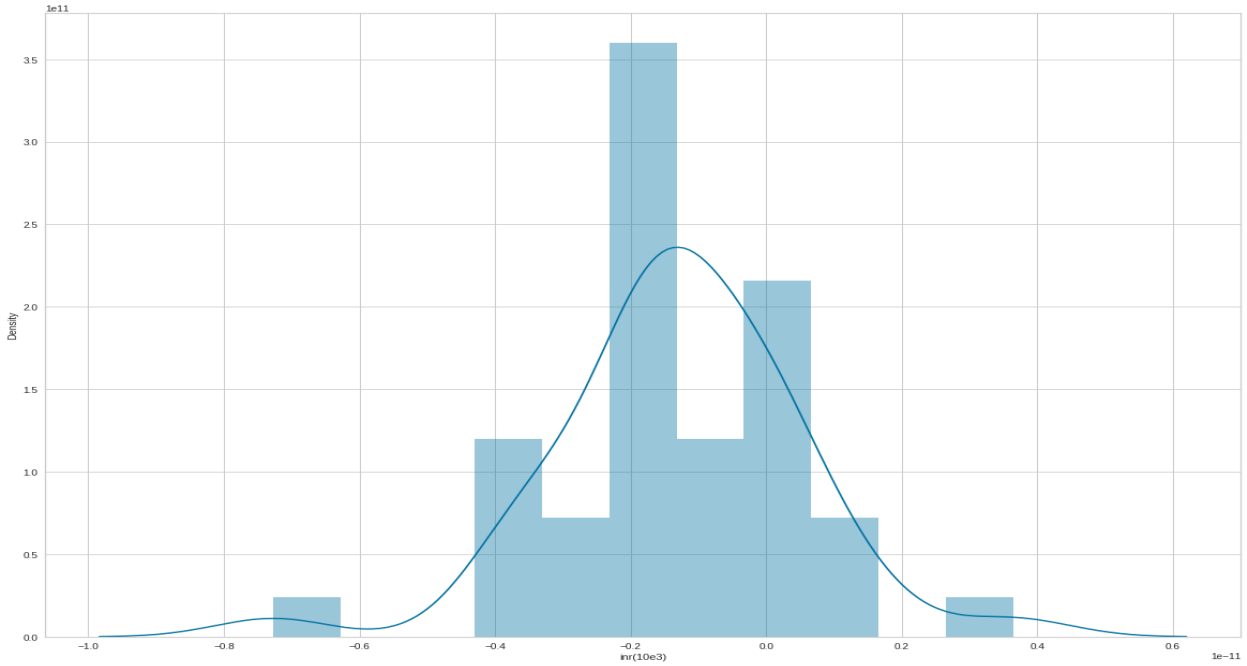
• Compute the sum of the squared distance between data points and all centroids.

• Assign each data point to the closest cluster (centroid).

• Compute the centroids for the clusters by taking the average of all data points that belong to each cluster.

• Keep iterating until there is no change to the centroids. i.e. assignment of data points to clusters isn't changing.

After completion of training the model process, we test the remaining 60% of data on the model. The obtained results are checked using a scatter plot between the predicted values and the original test data set for the dependent variable and acquired similar to a straight line as shown in the figure and the density function is also normally distributed.



Target Segments:

So from the analysis, we can see that the optimum targeted segment should be belonging to the following categories:

Behavioral: Mostly from our analysis there are cars with 5 seats.

Demographic:

• *Top Speed & Range* : With a large area of market the cost is dependent on Top speeds and Maximum range of cars.

• *Efficiency* : Mostly the segments have with most efficiency.

Psychographic:

• Price : From the above analysis, the price range is from16 lkhs to 180 lkhs.

Github link : https://github.com/ShraddhaSaha/EV-market-Segmentation-by-Shraddha-Saha