Electric Vehicle Market in India

Market Segmentation

Team B

Aditya Saxena

Abstract

EVs adoption is expected to grow phenomenally in near future as low emission and low operating cost vehicle, and thus, it drives a considerable amount of forthcoming academic research curiosity. The findings posit that the three distinct sets of young consumer groups have been identified and labelled as 'Conservatives', 'Indifferent', and 'Enthusiasts' which are deemed to be buddying EV buyers The implications are recommended, which may offer some pertinent guidance for scholars and policy- makers to encourage EVs adoption in the backdrop of emerging sustainable transport market.

Data Collection

The data has been collected manually, and the sources used for this process are listed below:

- https://www.kaggle.com/datasets

Market Segmentation for EV Market

In practice, there are two ways of segmenting the market: a-priori and post-hoc. In the post-hoc approach to segmentation on other hand, the segments are identified based on the relationship among the multiple measured variables. The commonality between both approaches lies in the fact that the measured variables determine the 'segmentation theme.' The present study utilizes an a-priori approach to segmentation to divide the potential EV customers into sub-groups.
It is argued that the blended approach of psychographic and socioeconomic attributes for market segmentation enables the formulation of sub-market strategies which in turn satisfy the specific tastes and preferences of the consumer groups.

They pinpointed the perceived superiority of the psychographic characteristics over the socio-demographic and economic ones in explaining the environmentally-conscious consumer behavior and thus, the study recommended the use of psychographic characteristics in profiling the consumer segments in the market for eco-friendly products. The present study adds perceived-benefit characteristics guided by blended psychographic and socio-economic aspects for segmenting the consumer market.

Implementation

Packages/Tools used:

1. **Numpy:** To calculate various calculations related to arrays.
2. **Pandas:** To read or load the datasets.
3. **SKLearn:** We have used LabelEncoder() to encode our values.

Data-Preprocessing

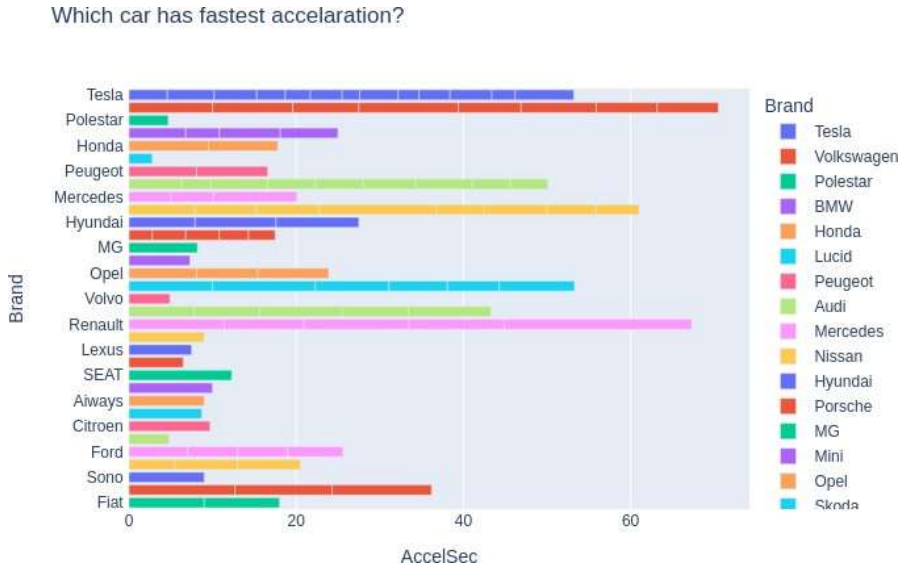
Data Cleaning

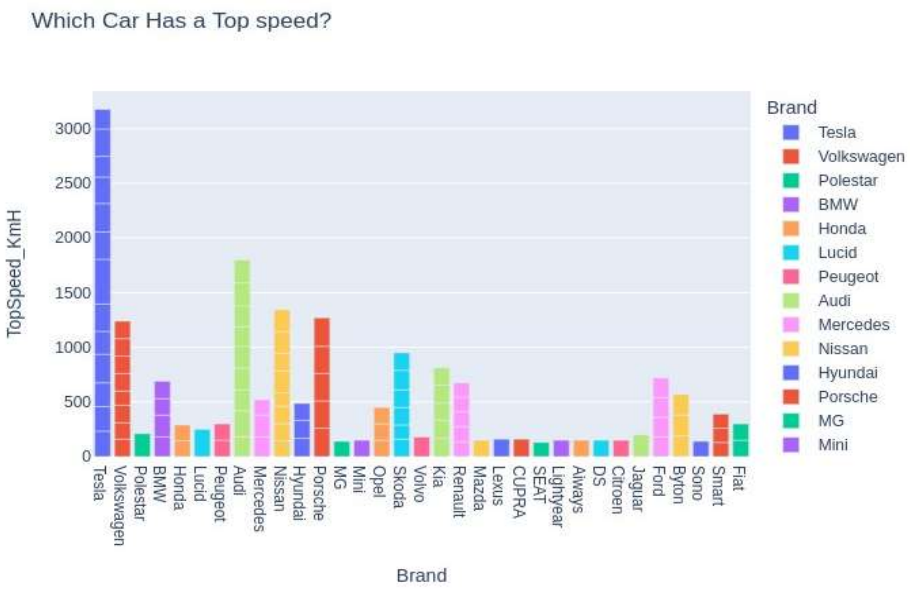
Python libraries such as NumPy, Pandas, Scikit-Learn, and SciPy are used for the workflow, and the results obtained are ensured to be reproducible.

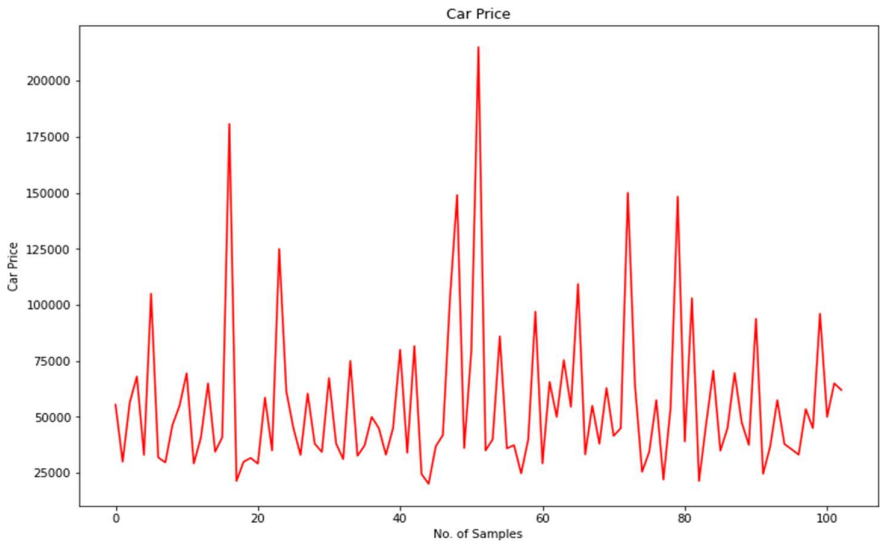
EDA

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 statis- tical process that converts the observations of correlated features into a set of linearly uncorrelated features with the help of orthogonal transformation. These new trans- formed features are called the Principal Components.

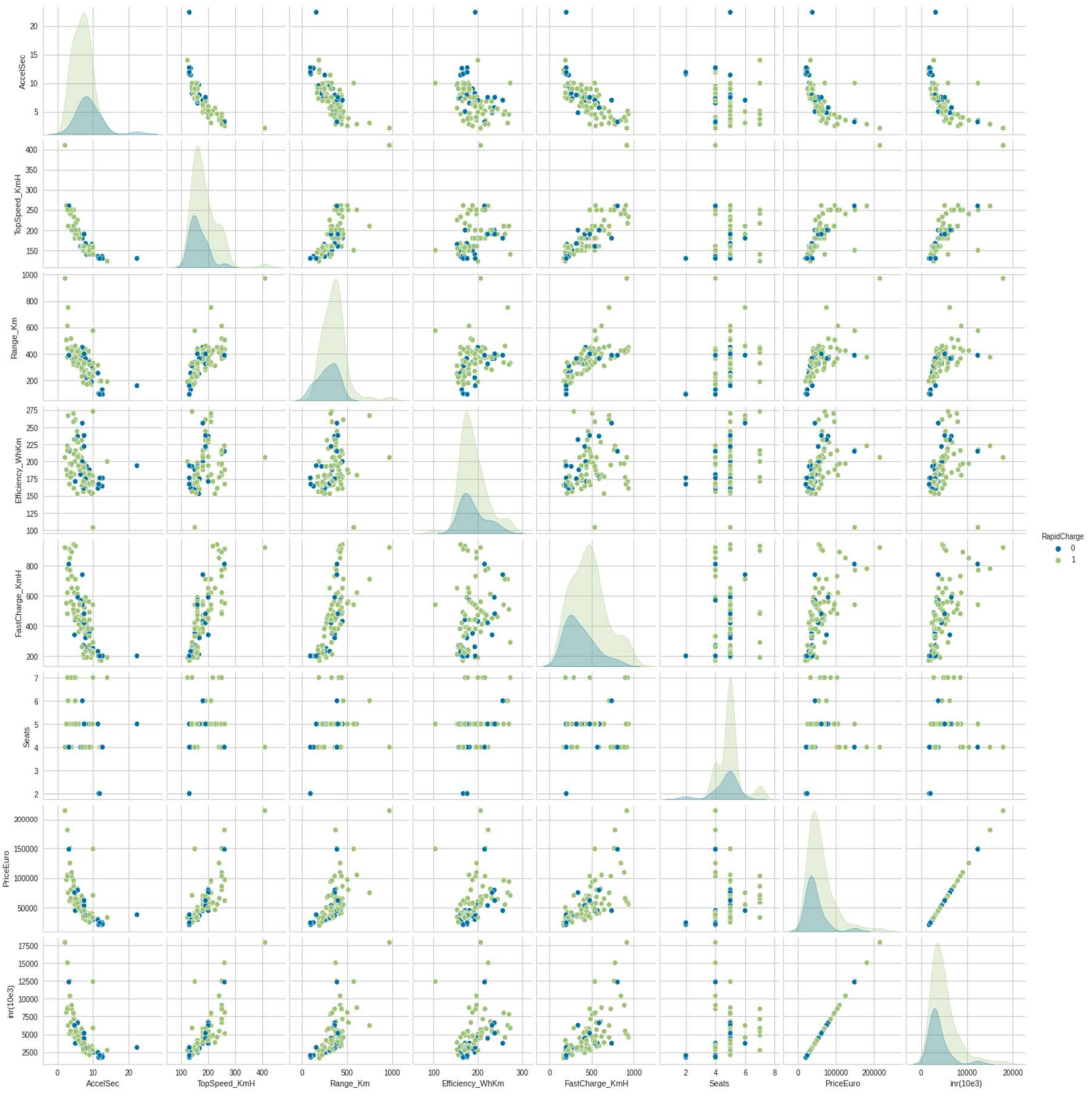
Comparision of cars in our data







For Electric Vehicle Market one of the most important keys is Charging:



Correlation Matrix: A correlation matrix is simply a table that displays the cor- relation. It is best used in variables that demonstrate a linear relationship between each other. Coefficients for different variables. The matrix depicts the correlation be- tween all the possible pairs of values through the heatmap in the below figure. The relationship between two variables is usually considered strong when their correlationcoefficient value is larger than 0.7.

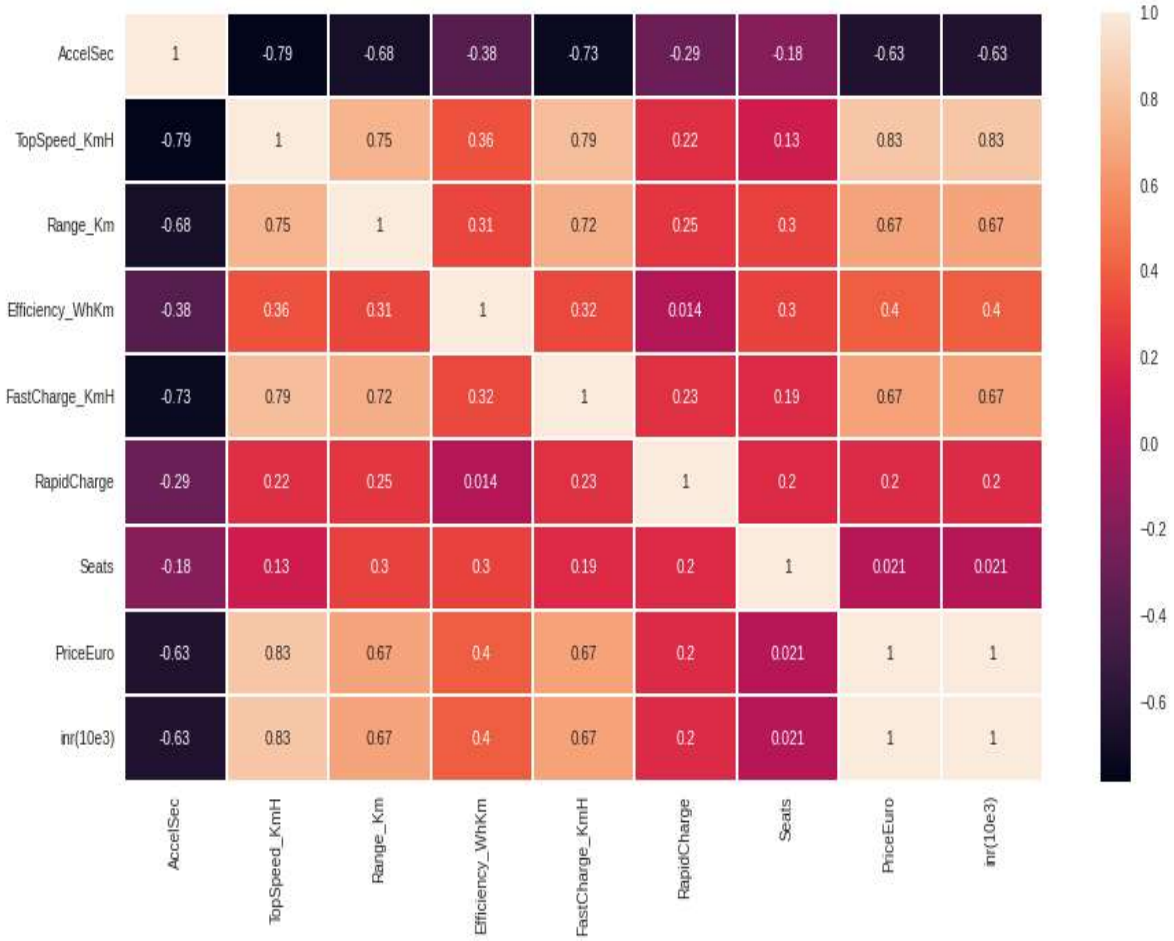
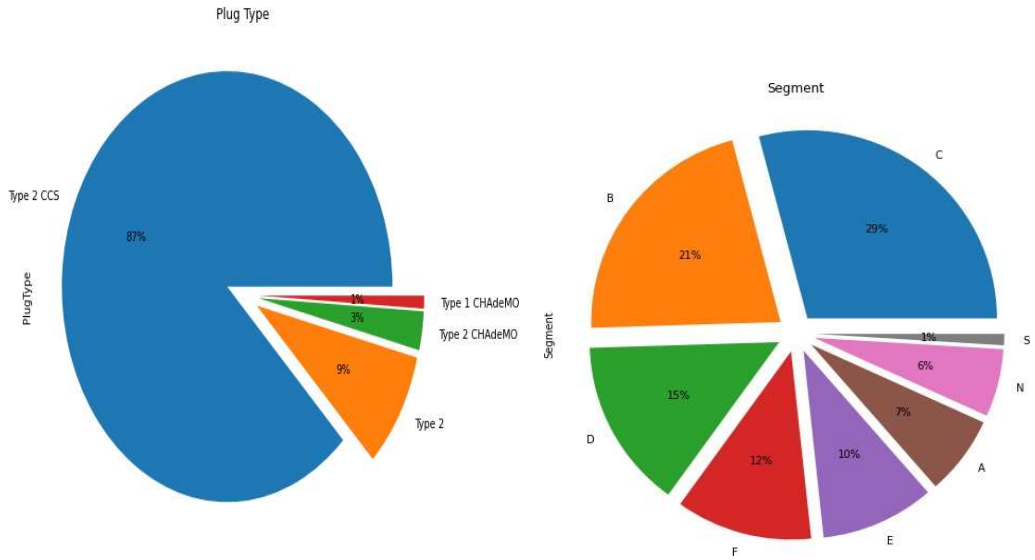
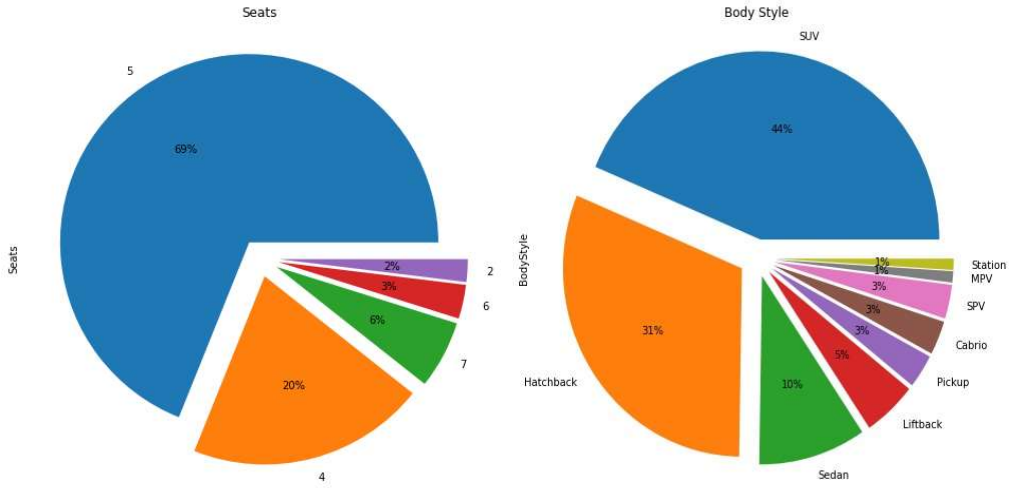


Figure 5: *Correlation Matrix for the dataset*





Now we can see that the requirements of what type of cars are most needed for customers and from the past 10 years there is a rapid growth of Electric vehicles usage in India.

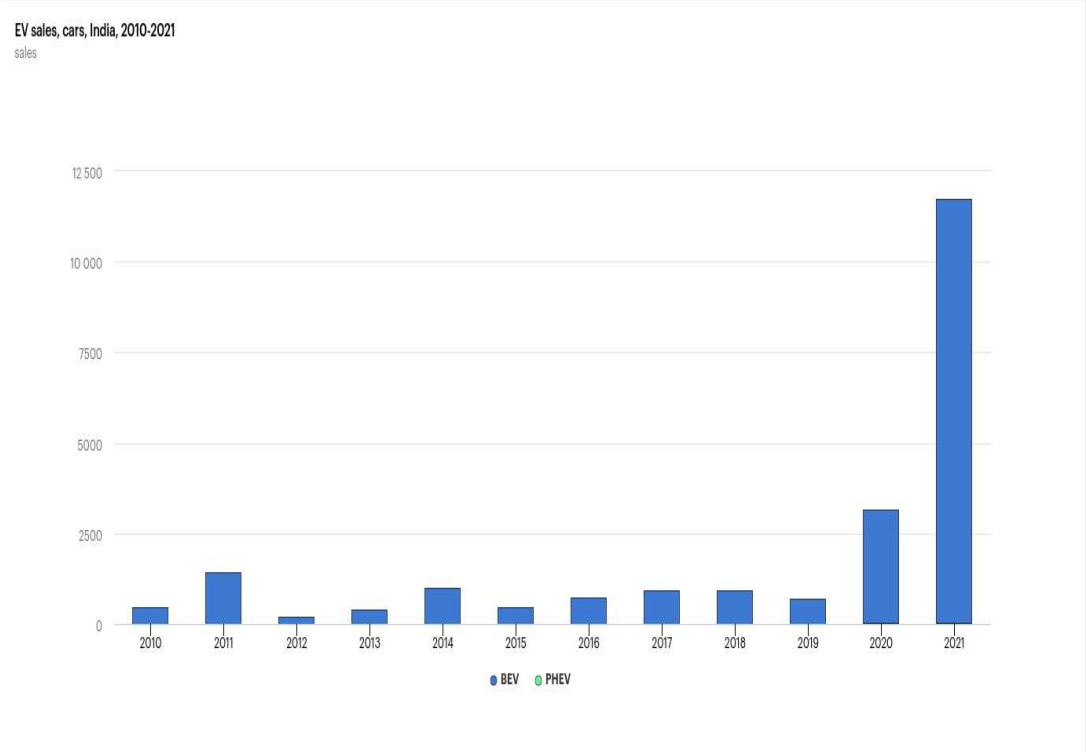


Figure 6: *Electric Cars sales in India*

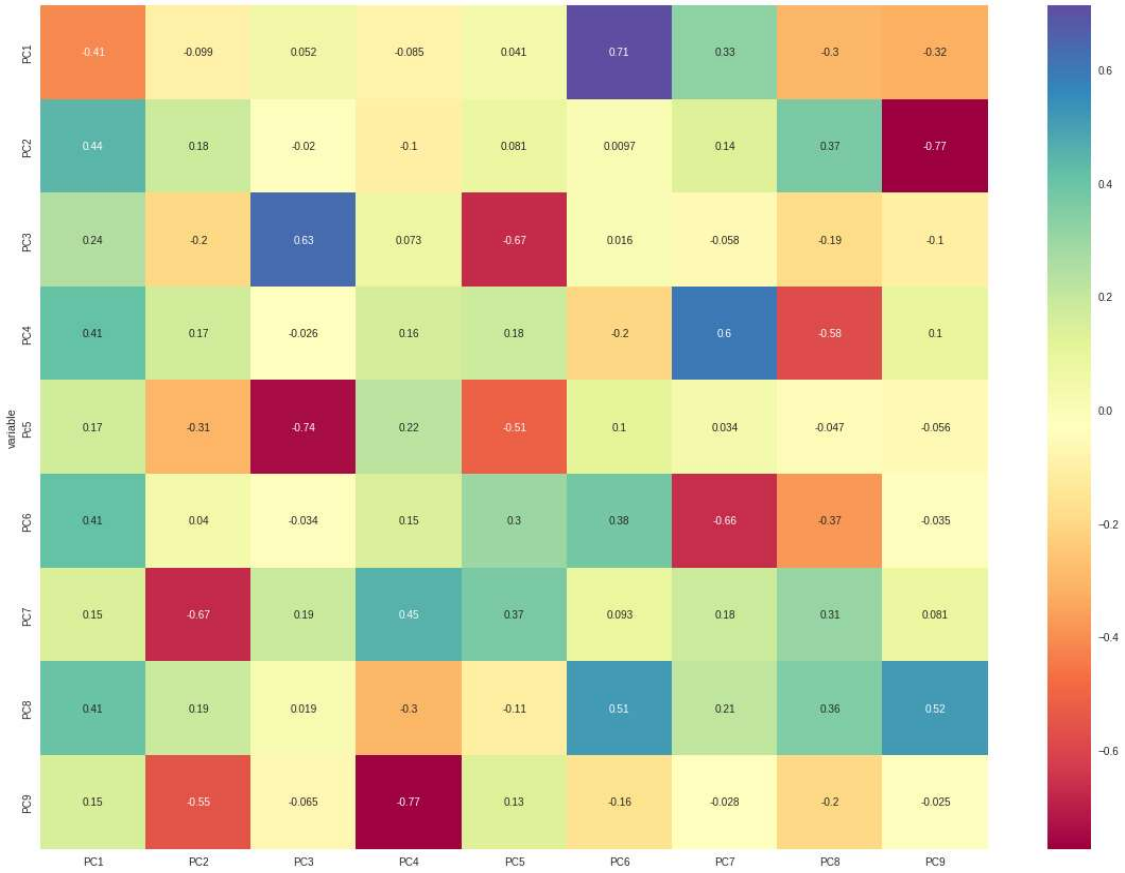


Figure 7: *Correlation matrix plot for loadings*

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. 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.
Clustering analysis can be done based on features where we try to find subgroups of samples.

K-Means Algorithm

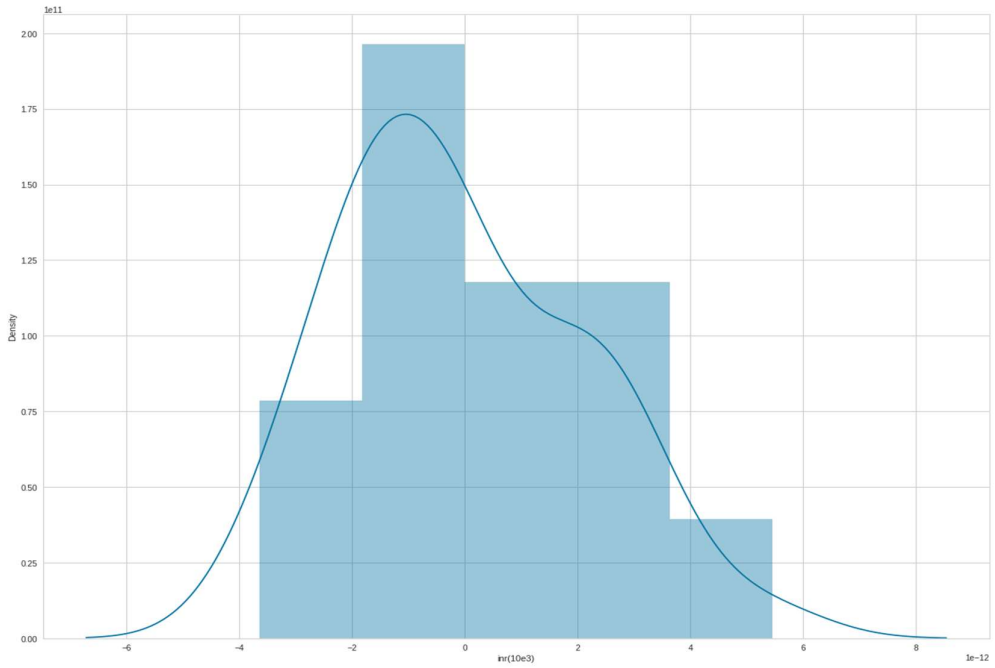
It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid is at the minimum. The less variation we have within clusters, the more homogeneous the data points are within the same cluster.

Prediction of Prices most used cars

Linear regression is a machine learning algorithm based on supervised learning. It per- forms a regression task. Regression models targets prediction value based on independent variables. It is mostly used for finding out the relationship between variables and forecasting. Here we use a linear regression model to predict the prices of different Electric cars in different companies. X contains the independent variables and y is the dependent Prices that is to be predicted. We train our model with a splitting of data into a 4:6 ratio, i.e., 40% of the data is used to train the model.

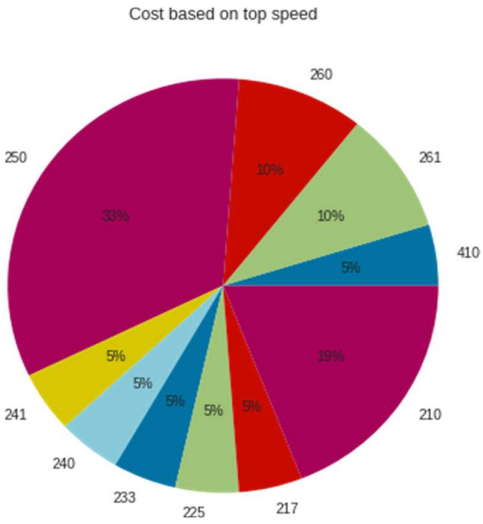
LinearRegression().fit($X_{train,ytrain}$) command is used to fit the data set into model. The values of intercept, coefficient, and cumulative distribution function (CDF) are described in the figure.

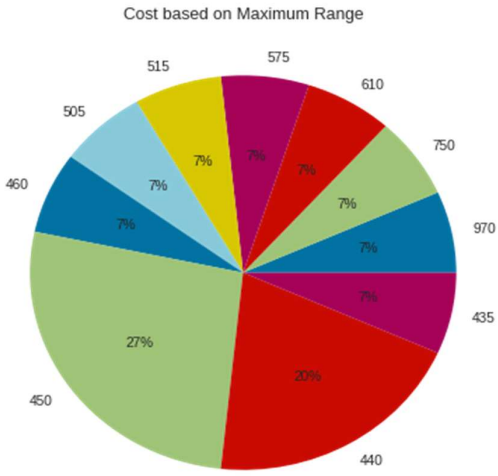
After completion of training the model process, we test the remaining 60% of data on the model. The obtained results are checked using a scatter plot between predicted values and the original test data set for the dependent variable and acquired like a straight line as shown in the figure and the density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

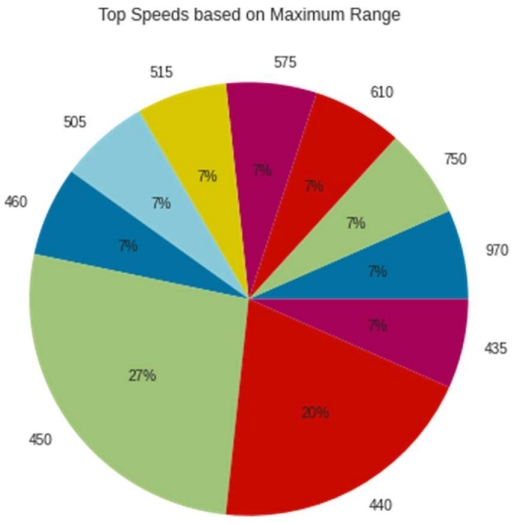


Profiling and Describing the Segments

Sorting the Top Speeds and Maximum Range in accordance to the Price with head ()we can view the Pie Chart.







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 are with most efficiency.

Psychographic:

- *Price:* From the above analysis, the price range is between 16,00,000 to 1,80,00,000.

Finally, our target segment should contain cars with most **Efficiency**, contains **Top Speed** and price between **16 to 180 lakhs** with mostly with **5 seats**.

References

[1] Dolnicar, S., Grun Bettina, amp; Leisch, F. (2019). *Market segmentation analysisunderstanding it, doing it and making it useful*. Springer Nature.

[2] McDonald, M., amp; Dunbar, I. (2003). /emphMarket segmentation. Butterworth-Heinemann.

GitHub:

https://github.com/rohansaxena2002/EVMarketSegmentation