EV-Market-Segmentation

Aanshu Maurya

May 2023

1 Introduction

The electric vehicle (EV) market in India has experienced substantial development and change recently. The Indian government has developed a number of regulations and measures to encourage the use of electric cars, with a heavy emphasis on lowering carbon emissions and dependency on fossil fuels. These initiatives, as well as rising environmental consciousness and improvements in EV technology, have helped the EV industry in India grow quickly.

A broad variety of vehicles, including electric automobiles, two-wheelers, threewheelers, and buses, are available in the Indian EV market. Consumers and manufacturers have been strongly encouraged to adopt electric mobility thanks to government incentives and subsidies, such as the Faster Adoption and Manufacturing of (Hybrid) Electric Vehicles (FAME) plan. In addition, the development of charging infrastructure networks throughout important cities and roads has allayed worries about range anxiety and accelerated the adoption of EVs.

Overall, due to government assistance, rising environmental consciousness, and technology developments, the EV market in India is primed for continued growth. In addition to providing a cleaner and more environmentally friendly transportation option, the move to electric mobility also presents the nation with major economic potential, including the development of jobs and a decreased reliance on imported fossil fuels.

2 Abstract

Market segmentation becomes a vital strategy for emerging markets to investigate and execute for widespread adoption of emerging mobility technologies like electric vehicles (EVs).

As a low emission and low operating cost vehicle, EV adoption is anticipated to increase drastically in the near future. As a result, it will stimulate a significant amount of future academic study interest.

By utilising an integrated research framework of "perceived benefits-attitudeintention," the primary goal of this study is to examine and identify several sets of possible customer groups for EVs based on psychographic, behavioural, and socio-economic characterisation. To operationalize and validate segments from the data gathered from 563 respondents via a cross-sectional online survey, the study used rigorous analytical techniques such cluster analysis, multiple discriminant analysis, and the Chi-square test.

According to the research, there are three separate youthful consumer groups known as "Conservatives," "Indifferents," and "Enthusiasts" who are thought to be forming relationships with EV customers. The consequences are suggested, which can provide some useful direction for academics and policymakers to promote the adoption of EVs against the backdrop of the developing sustainable transportation sector.

In this document, we will break down the issue, assess the data, and use Fermi Estimation to resolve it.

3   Data Collection

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

https://www.kaggle.com/datasets
https://data.gov.in/ https://www.data.gov/
https://data.worldbank.org/
https://datasetsearch.research.google.com/

4   Market Segmentation

4.1   Target Market

Geographic, SocioDemographic, Behavioural, and Psychographic Segmentation can be used to classify the target market for electric vehicle market segmentation.

4.2   Behavioral Segmentation

The process of behavioural segmentation looks specifically for patterns in observed or reported behaviour.

Examples include past use of the product, the cost of the purchase, etc.
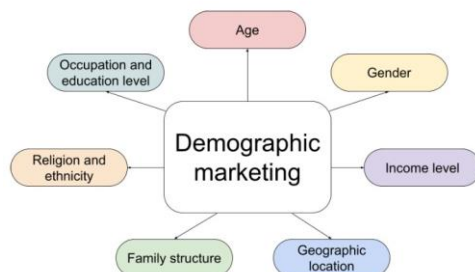
4.3   Psychographic Segmentation:

categorised according to the advantages customers seek while buying a product, such as beliefs, interests, preferences, or ambitions. Suitable for segmenting by lifestyle.
Involves a lot of segmentation factors.

4.4   Socio-Demographic Segmentation

It includes age, gender, income and education. Useful in industries.

5   Segmenting for Electric Vehicle Market

With the use of a comparable set of marketing methods, marketers may target actionable, controllable, homogeneous segments of specific clients using the market segmentation strategy.

There are two methods for segmenting the market in practise: a priori and posthoc.

In an a-priori method, segments are first determined based on factors like age, gender, income, and education, and then profiles are created based on a variety of assessed factors (behavioural, psychographic, or benefit).

On the other hand, with the post-hoc method to segmentation, the segments are determined based on the connection between the various measured variables. The'segmentation theme' is determined by the measured variables, which is a feature shared by both techniques.

The present study utilizes an a-priori approach to segmentation so as to divide the potential EV customers into sub-groups According to the argument, using both socioeconomic and psychographic characteristics to segment the market allows for the creation of sub-market strategies that in turn cater to the distinct likes and preferences of the various customer groups.
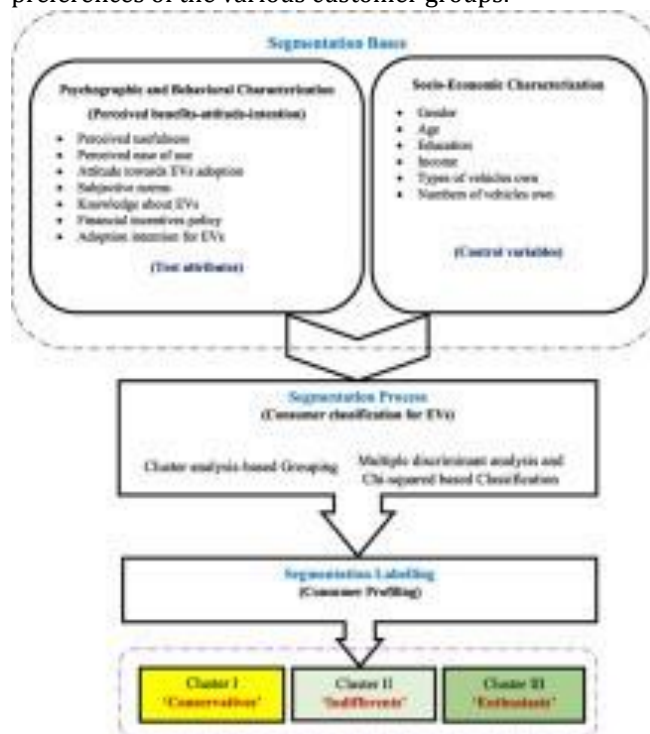
Based on customer evaluations of eco-friendly items, Straughan and Roberts compared the effectiveness of psychographic, demographic, and economic factors.

The study recommended using psychographic characteristics to profile the consumer segments in the market for eco-friendly products because they identified the perceived superiority of these traits over socio-demographic and economic ones in explaining the environmentally conscious consumer behaviour. For the purpose of segmenting the consumer market, the current study includes perceived-benefit characteristics that are directed by a combination of psychographic and socioeconomic factors.

6   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.

6.1   Data-Preprocessing

6.1.1   Data Cleaning

Compact data are gathered, and some of it is utilised for grouping and some of it for visualisation. The method makes use of Python modules like NumPy, Pandas, Scikit-Learn, and SciPy, and the outcomes are guaranteed to be repeatable.

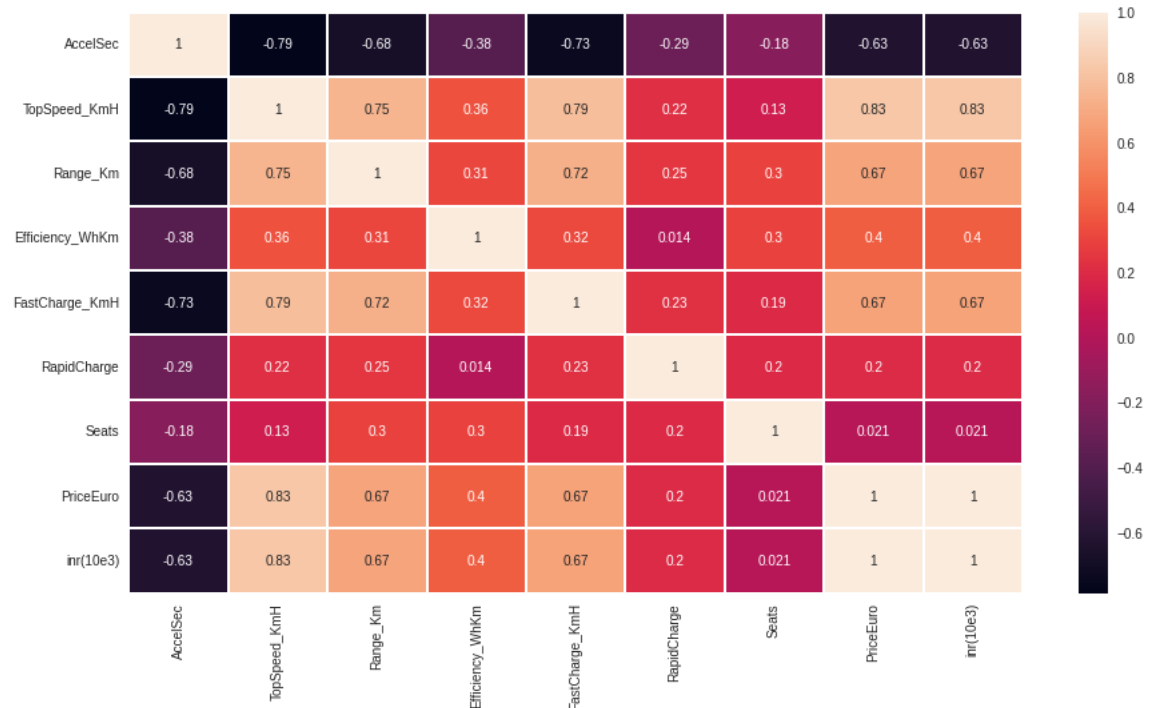It is shown in the Notebook.

6.1.2   EDA

Beginning with some data analysis without principal component analysis and some principal component analysis in the dataset created by combining all of the data we have, we begin the exploratory data analysis.
With the use of orthogonal transformation, PCA is a statistical procedure that transforms the observations of correlated features into a collection of linearly uncorrelated features. The Principal Components are these newly altered features. The method aids in the cost-effectiveness of classification, regression, or any other type of machine learning by lowering the dimensions of the data.

6.1.3   Correlation Matrix:

Simply said, a correlation matrix is a table that shows the correlation. It works well with variables that have a linear connection to one another. distinct variables' coefficients.

The heatmap in the next graphic shows how the matrix represents the correlation between all possible pairings of values. When the correlation coefficient between two variables is more than 0.7, that link is typically regarded as strong.

|	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	Seats	PriceEuro	inr(10e3)
AccelSec	1	-0.79	-0.68	-0.38	-0.73	-0.29	-0.18	-0.63	-0.63
TopSpeed_KmH	-0.79	1	0.75	0.36	0.79	0.22	0.13	0.83	0.83
Range_Km	-0.68	0.75	1	0.31	0.72	0.25	0.3	0.67	0.67
Efficiency_WhKm	-0.38	0.36	0.31	1	0.32	0.014	0.3	0.4	0.4
FastCharge_KmH	-0.73	0.79	0.72	0.32	1	0.23	0.19	0.67	0.67
RapidCharge	-0.29	0.22	0.25	0.014	0.23	1	0.2	0.2	0.2
Seats	-0.18	0.13	0.3	0.3	0.19	0.2	1	0.021	0.021
PriceEuro	-0.63	0.83	0.67	0.4	0.67	0.2	0.021	1	1
inr(10e3)	-0.63	0.83	0.67	0.4	0.67	0.2	0.021	1	1

6.1.4   Scree Plot

Is a standard technique for visualising the number of PCs that should be preserved. The eigenvalues for each individual PC are displayed using a straightforward line segment plot.

It displays the number of components on the x-axis and the eigenvalues on the y-axis.

It consistently shows a declining slope. The majority of scree plots have a form that is essentially the same: they begin high to the left, drop swiftly, and then eventually flatten out.

In order to pick all components, the scree plot criterion locates the "elbow" in the curve and does so immediately before the line flattens out. The plot of
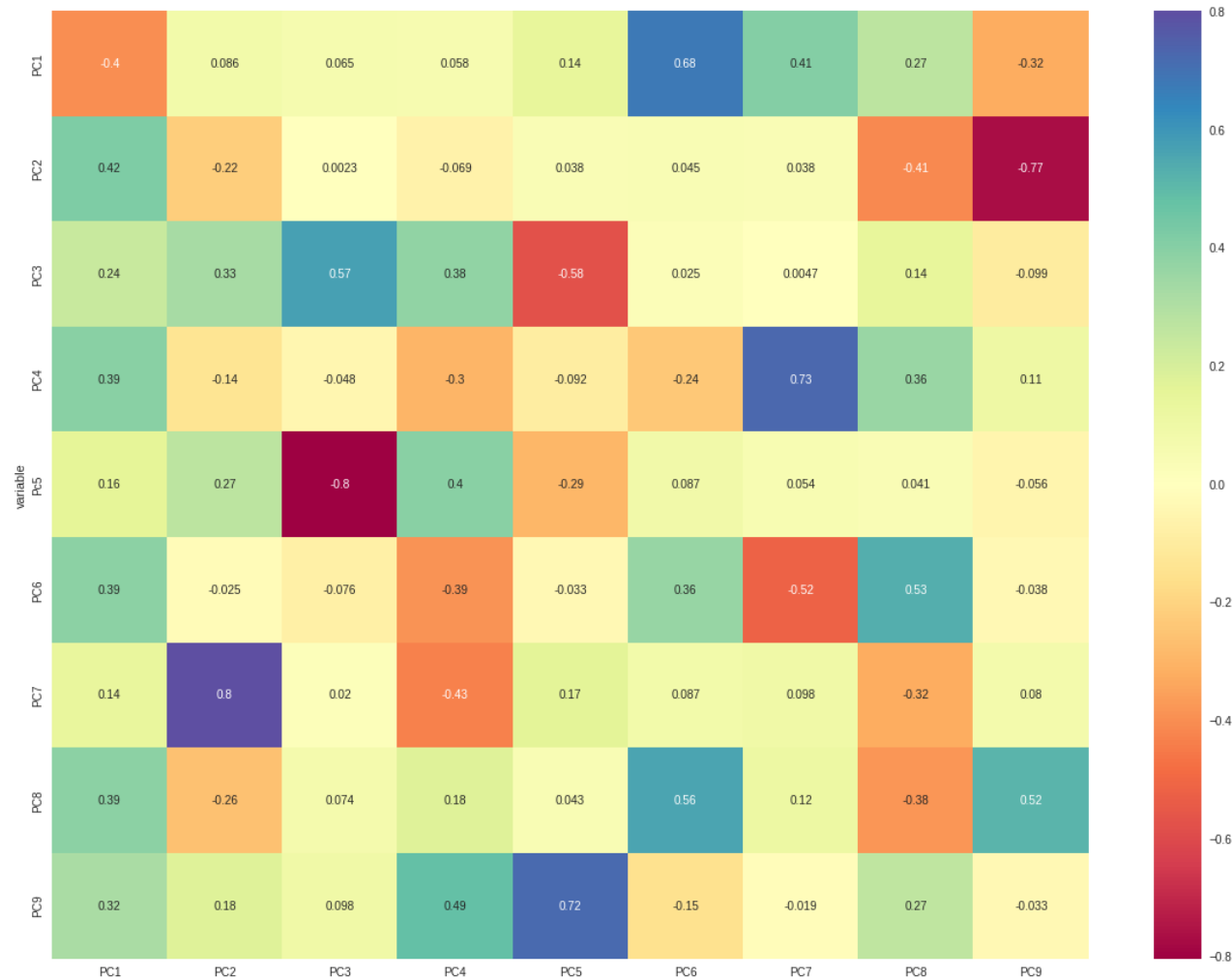
proportional variance: At least 80% of the variation should be able to be explained by the chosen PCs.

Scree Plot
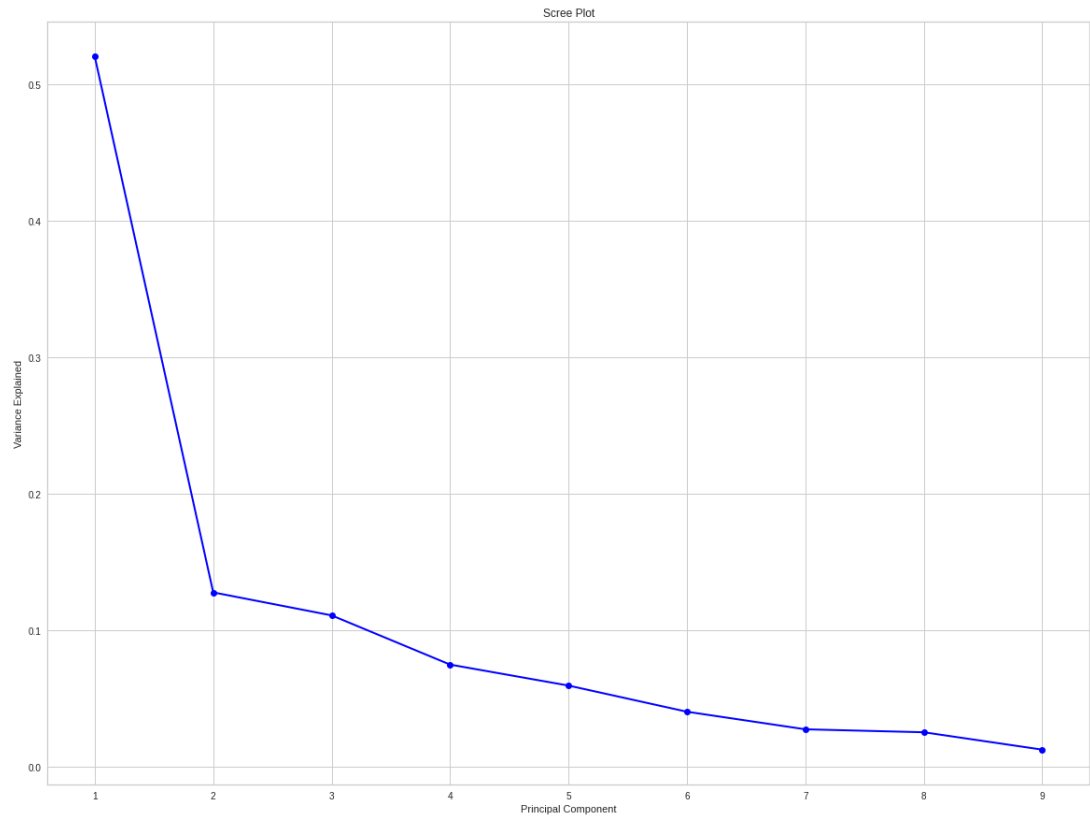
7   Extracting Segments

7.1   Dendrogram

This strategy is unique to the agglomerative hierarchical clustering approach. Each point is initially taken into account as a distinct cluster in the agglomerative hierarchical technique of clustering, and points are then connected to clusters in a hierarchical manner depending on their distances.

In order to determine the ideal number of clusters for hierarchical clustering, we employ a dendrogram, a tree-like graph that displays the sequences of cluster mergers or splits.

The dendrogram will join two clusters together if they are combined, and the height of the join will be determined by the separation between the two clusters. Figure illustrates how we may select the ideal number of clusters depending on the dendrogram's hierarchical structure.

As highlighted by other cluster validation metrics, four to five clusters can be considered for the agglomerative hierarchical as well.
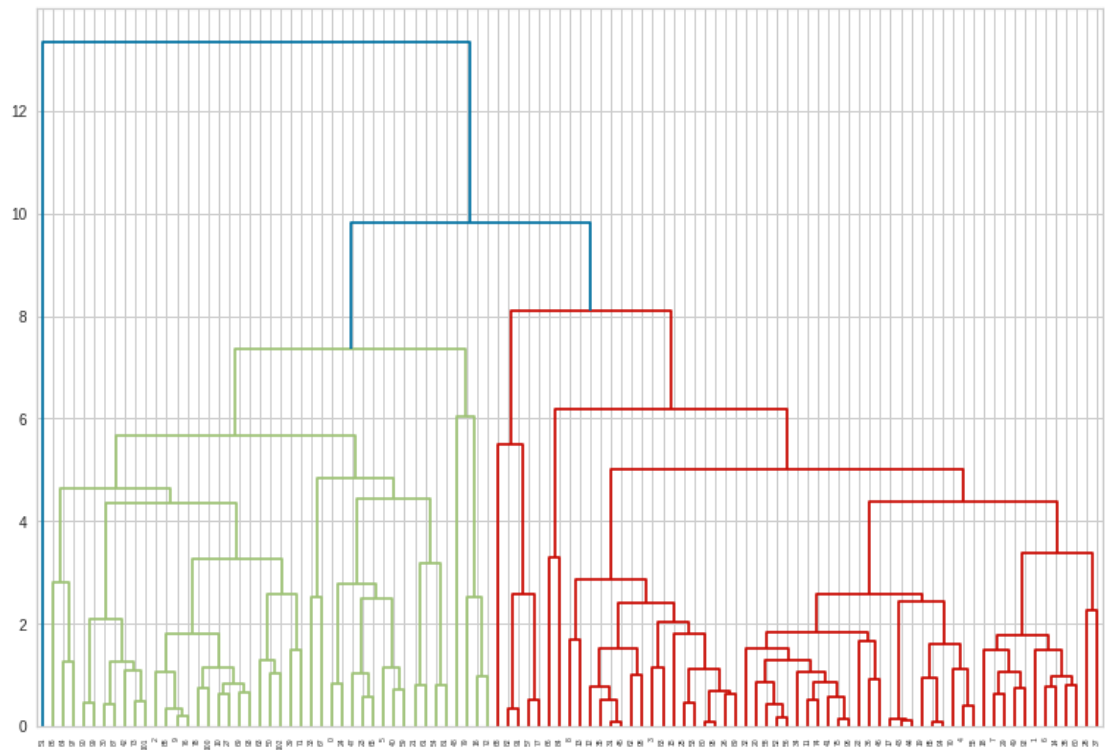
7.2   Elbow Method

The Elbow approach is a well-liked technique for figuring out how many clusters are ideal.

The approach is based on computing the Within-Cluster-Sum of Squared Errors (WSS) for various numbers of clusters (k), then choosing the k for which change in WSS first begins to decline.

The elbow approach is based on the premise that the explained variation varies quickly for a few clusters, then it slows down, forming an elbow in the curve. The amount of clusters we may employ for our clustering process is represented by the elbow point.

For cluster numbers ranging from 2 to 8, the Elbow Visualizer function matches the KMeans model. The elbow point is reached, as seen in Figure, which is underlined by the function itself.

The green line in the function shows us how long it took to draw models for different numbers of clusters.

Calinski Harabasz Score Elbow for KMeans Clustering
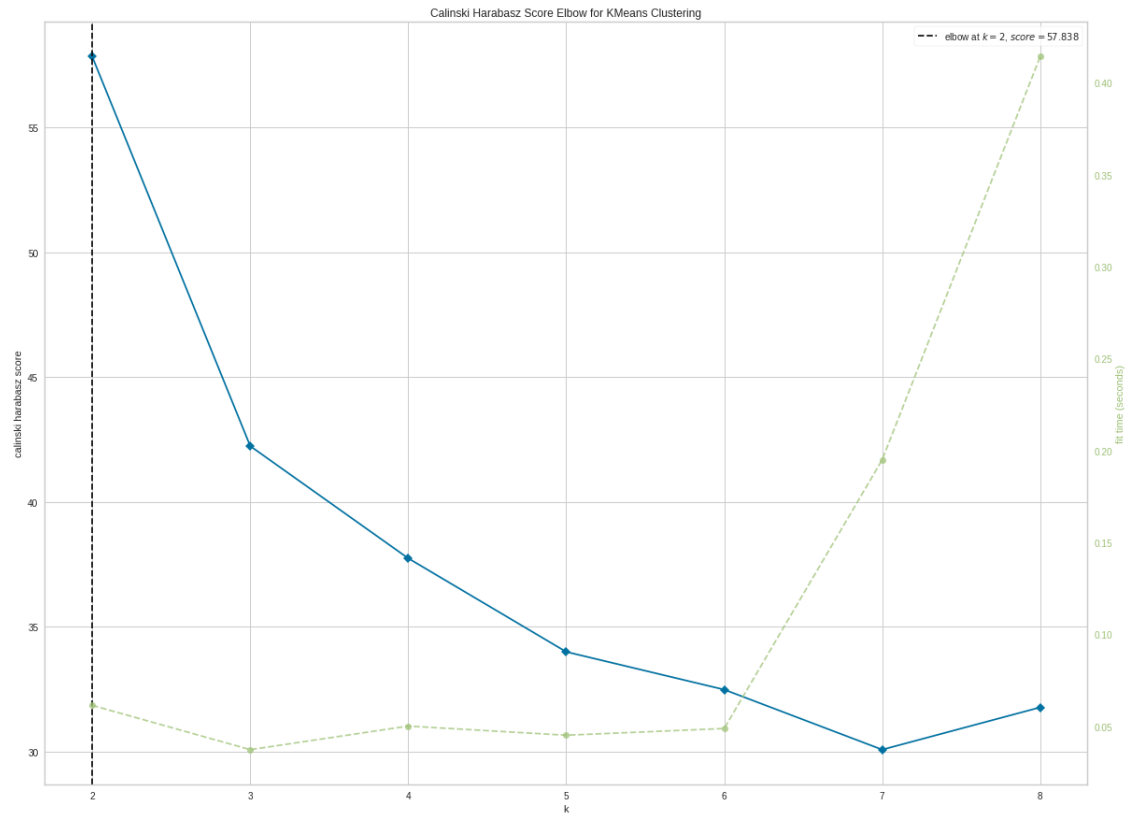
8 Analysis and Approaches used for Segmentation

8.1 Clustering

One of the most popular methods for gaining a general understanding of the data's structure is clustering.

It may be described as the process of finding data subgroups where data points in the same subgroup (cluster) are extremely similar and other data points in other clusters are quite dissimilar.

To put it another way, we look for homogeneous subgroups within the data so that the data points in each cluster are as comparable as feasible in terms of a similarity metric like euclidean-based distance or correlation-based distance. Choosing the similarity metric to employ depends on the application.

Clustering analysis may be carried out either on the basis of samples or on the basis of characteristics, where we attempt to identify subgroups of samples based on features.

8.2   K-means Clustering

The iterative K-means clustering technique seeks to divide the dataset into clusters (pre-defined discrete, non-overlapping subgroups) where each data point only belongs to one cluster.

While maintaining the clusters as distinct (far) apart as feasible, it aims to make the intra-cluster data points as comparable as possible.

It places data points into clusters so that the total squared distance between each data point and the cluster's centroid—the average value of all the data points in the cluster—is at its smallest.

The homogeneity (similarity) of the data points inside a cluster increases as the amount of variance within the cluster decreases.

9   Prediction of Prices most used cars

A machine learning algorithm built on supervised learning is linear regression. It carries out a regression operation. Based on independent variables, regression models aim to predict a value.

It is mostly used to determine how variables and forecasting relate to one another. Here, we employ a linear regression model to forecast the costs of various Electric vehicles in various businesses.

The dependent Prices are represented by y, whereas the independent variables are included in X.

We split the data into a 4:6 ratio to train our model, which means that 40% of the data is utilised.

LinearRegression().The data collection is fitted into the model using the fit(Xtrain,ytrain) function.

The graphic gives information about the intercept, coefficient, and cumulative distribution function (CDF) values.

We test the remaining 60% of the data on the model when the model-training procedure is complete.

A scatter plot between the predicted values and the original test data set for the dependent variable, acquired similarly to a straight line as shown in the image, is used to compare the findings to ensure they are accurate. The density function is also normally distributed.
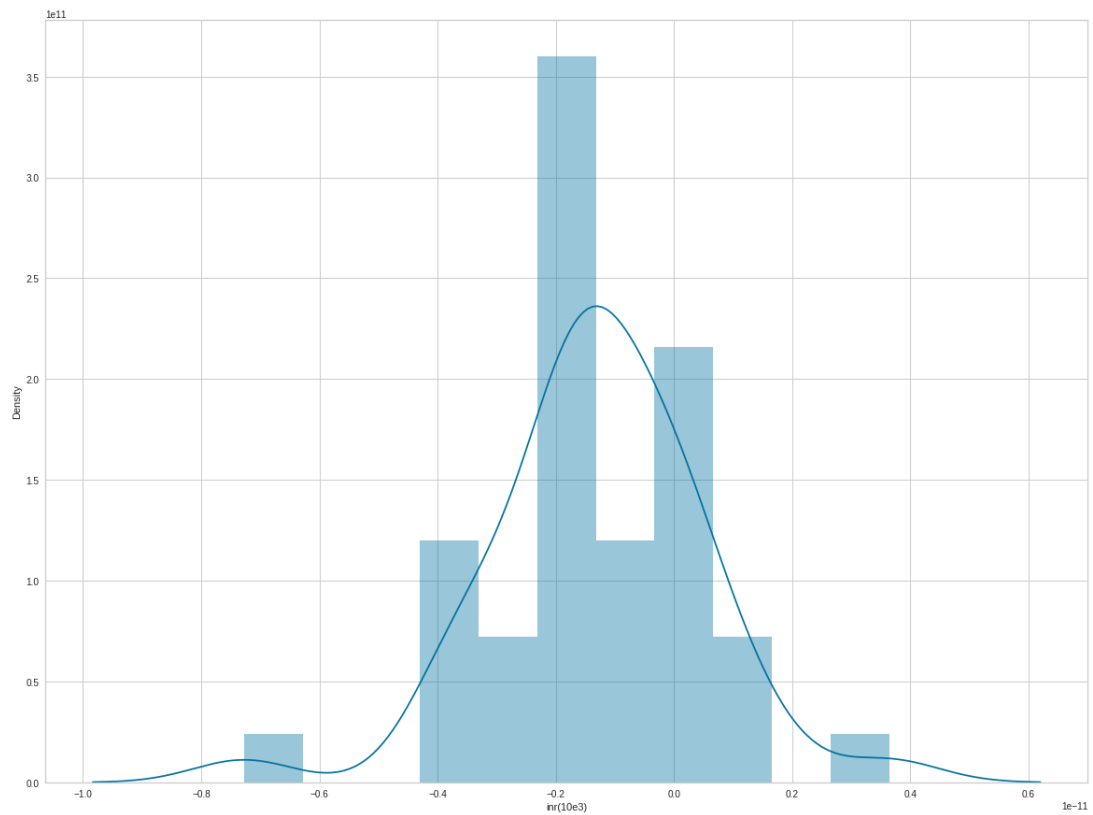
10 Target Segments:

Therefore, based on the data, it is clear that the following groups should comprise the ideal targeted market.

10.1 Behavioral:

Mostly from our analysis there are cars with 5 seats.

10.2 Demographic:

Top Speed Range: Due to the size of the market, the cost of automobiles is influenced by their top speeds and maximum range.

• Effectiveness: The portions tend to be the most effective.

10.3   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.

11   Customizing the Marketing Mix

The marketing mix is the collection of activities, or methods, that a business does to sell its brand or merchandise. A typical marketing mix is made up of the four Ps: price, product, promotion, and place.

•   Price: This term describes the value assigned to a product. It is dependent on the target market segment, the capacity of the businesses to pay, the capacity of the consumers to pay, supply and demand, and a variety of other direct and indirect considerations.

•   Product: The actual thing being sold, which in this case is the service. Even the finest effort on the other components of the marketing mix will be useless if the product doesn't perform to a certain level.

•   Place: Point of sale is referred to as "place." In any sector, the basic goal of an effective distribution or "place" strategy is to draw the consumer's attention and make it simple for her to purchase it.
For the ideal location, retailers fork up a premium.
'Location, location, location' is actually the motto of a prosperous retail firm. •

Promotion: This term refers to any actions made to spread awareness of the item

or service among consumers and businesspeople.

This can involve marketing, word-of-mouth recommendations, news coverage, rewards for the trade, and incentives. It can also include direct marketing, competitions, and giveaways.

The marketing mix's components all interact with one another. They create a company's business strategy, manage it properly, and may ensure its success. The marketing mix requires a great deal of comprehension, market research, and consultation with many parties, including users, traders, manufacturers, and many others.

GitHub Respositiory:

This is my github respositiory link in which I used to provide my notebook and the relevant things related to this project.

https://github.com/Ansh420/EV_Market_Segmentation