**EV MARKET SEGMENTATION**

Report Prepared By : **Anjali**(Machine Learning Intern, Feynn Labs)

**PROBLEM STATEMENT**

**Question :**

Based on Markey Analysis, the entire segmentation problem boils down to Two main Questions :

1. What type of EV the company will produce?

2. Who are the target customer?

In this case I am specifically looking into the 2 wheeler EV market as it is much bigger and more dynamic market comparatively. Thus, I look for What type of EV 2 Wheelers a EV 2 Wheeler company should focus it’s resources on.

**Approach :**

Task is to analyse the Electric Vehicles Market in India using Segmentation analysis and come up with a feasible strategy to enter the market, targeting the segments most likely to use their product in terms of Geographic, Demographic, Psychographic, and Behavioural. In this report I have analysed the 2 Wheeler Electric Vehicles Market in India using segmentation analysis and tried to answer some of the crucial questions. Along with that, I have also performed Segmentation on customers and their Income Data. The Segmentation is Done using PCA and K-Means clustering and Hierarchical Clustering is also shown. At the end the possible segments to choose for such an EV company, are evaluated based of selected features.

**DATA COLLECTION:**

Both the Demographic dataset and EV bikes dataset are collected from <https://www.kaggle.com/>

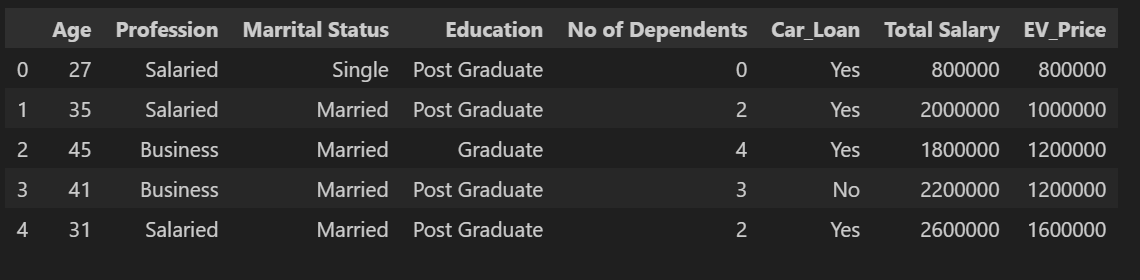
The bikes dataset needed to be modified to add some missing values like products from Popular brands. BEHAVIOURAL SEGMENTATION.

In this part I show the methods taken to do a demographic segmentation on the Dataset. But first some Exploratory Data Analysis is performed. An Exploratory Data Analysis or EDA is a thorough examination meant to uncover the underlying structure of a data set and is important for a company because it exposes trends, patterns, and relationships that are not readily apparent.

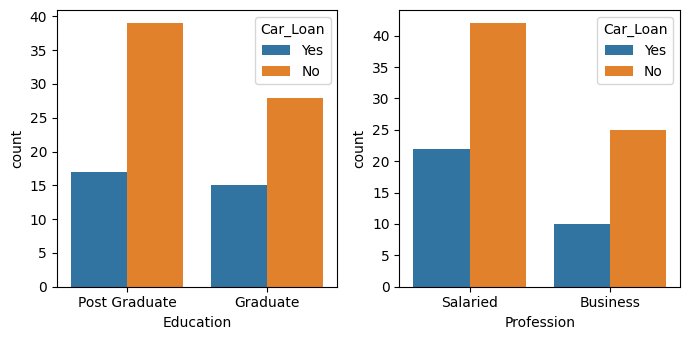
**BEHAVIOURAL SEGMENTATION :**

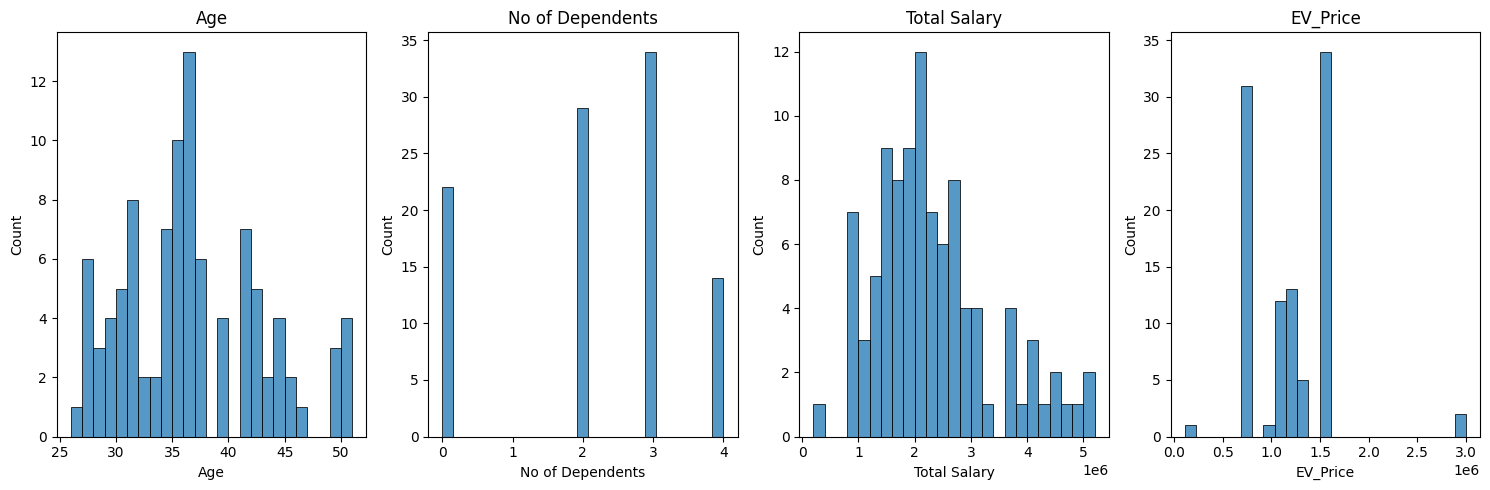
In this part I show the methods taken to do a demographic segmentation on the Dataset. But first some Exploratory Data Analysis is performed. An Exploratory Data Analysis or EDA is a thorough examination meant to uncover the underlying structure of a data set and is important for a company because it exposes trends, patterns, and relationships that are not readily apparent. Exploratory Data Analysis.

The Dataset –

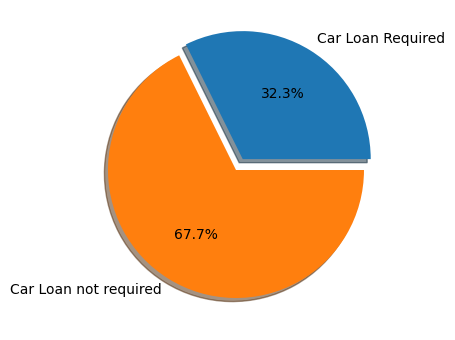


Count plots are used to show categorical features like Profession, Education, etc. and how many of them take Loan to buy EV.



Several Count plots are used to do Univariate Analysis on numerical data.

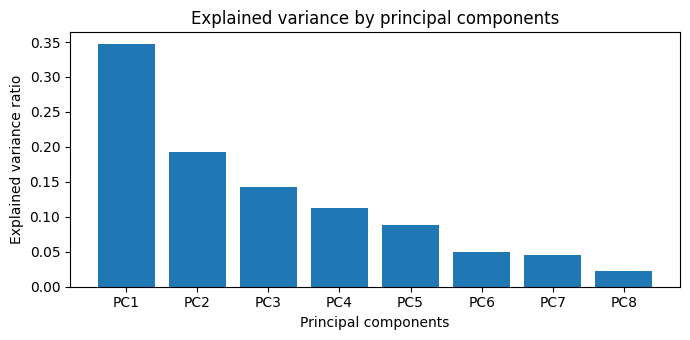
A data to show percentage of candidates taking loan, which might be useful for purchase options using EMI and other options given by the company.

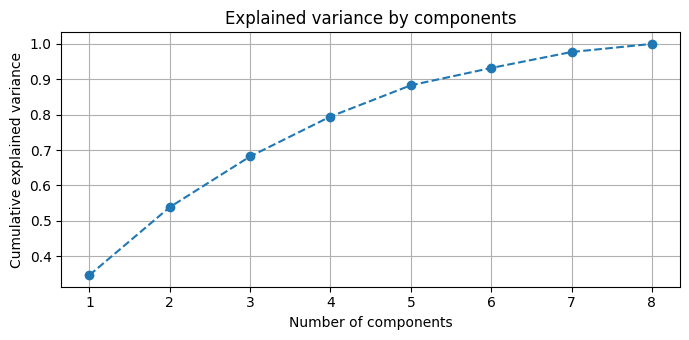


**Principal Component Analysis :**

Before Clustering the datapoints I have shown the correlation matrix where the dataset showing interdependency between features.

The data are preprocessed using Standard Scalar class in ScikitLearn and I proceed for PCA to extract the independent components and less than the number of features for which most of the information is intact i.e the explained variance.

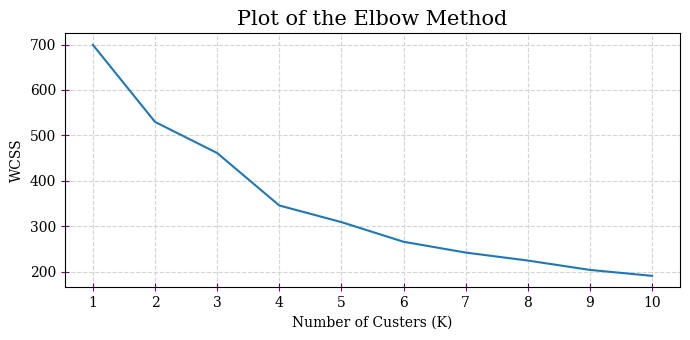




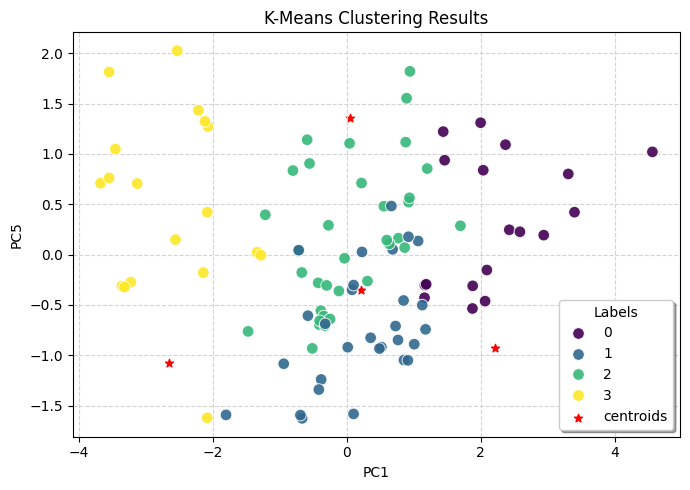
It is obvious from the PCA analysis by taking all the features that only 5 components explain more than 90% of the variance.

**K-Means Clustering :**

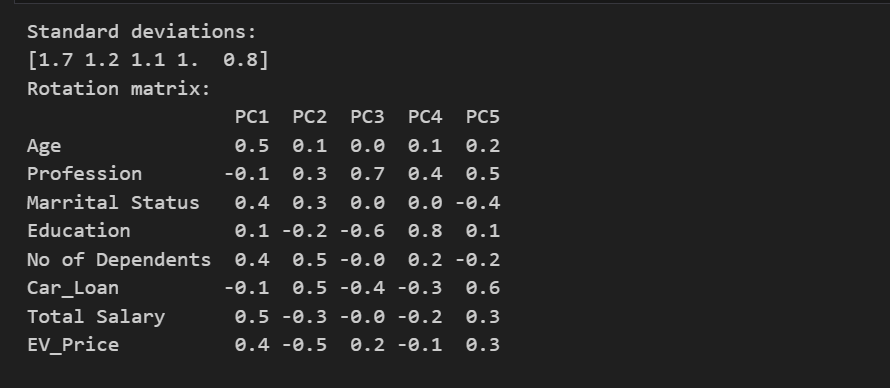
Now I performed K-means clustering for different number of clusters and plot the Elbow curve to determine the number of clusters, as the algorithm needs the number of clusters to be given as an input. One point to note that I have used “k-means++”. Only difference from normal one is that it initialises the clusters smartly rather than randomly in normal k-means.



According to the Elbow curve I have chosen 4 clusters to preform K-Means. The clusters are shown in the plot(First vs Last principal components).

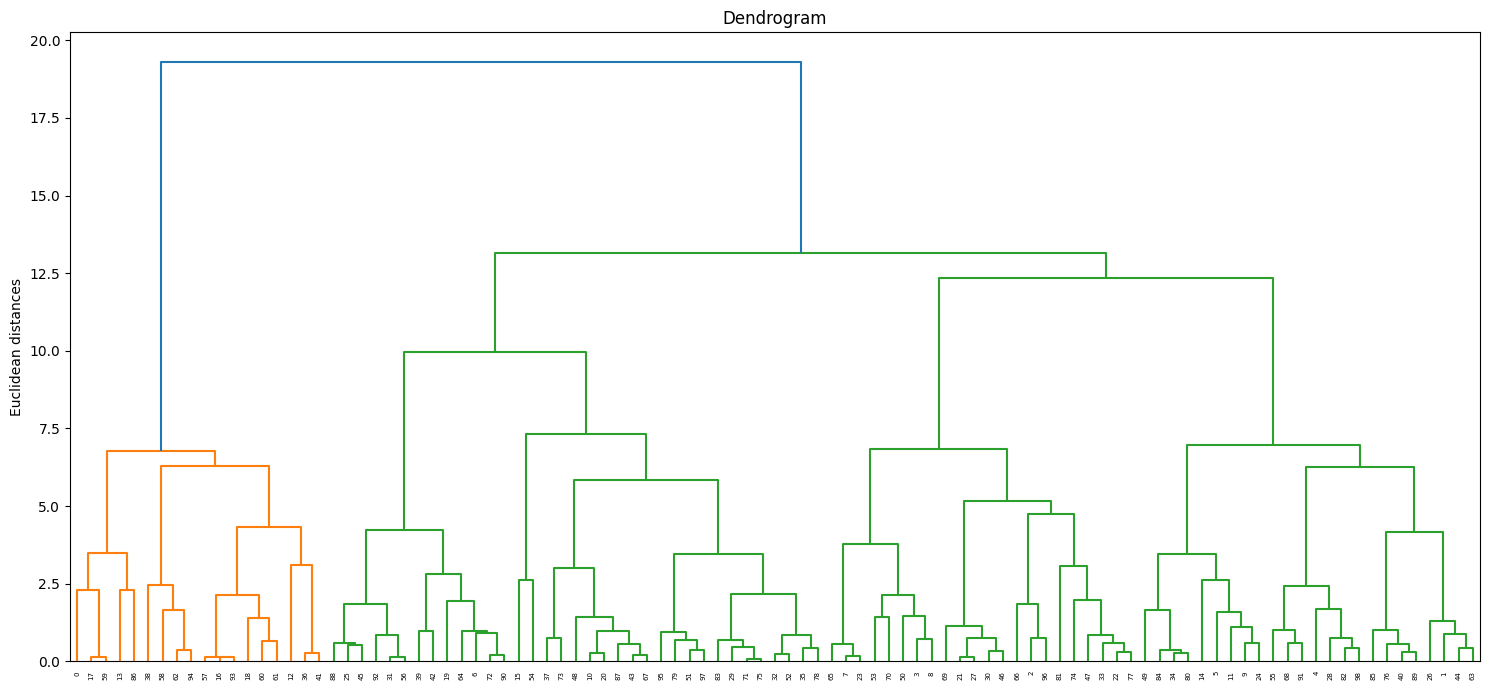


The standard deviations and rotations for each principal components are shown below.



**Hierarchical Clustering :**

A similar hierarchical clustering is also performed which shows 2 different components, shown in the Dendrogram.



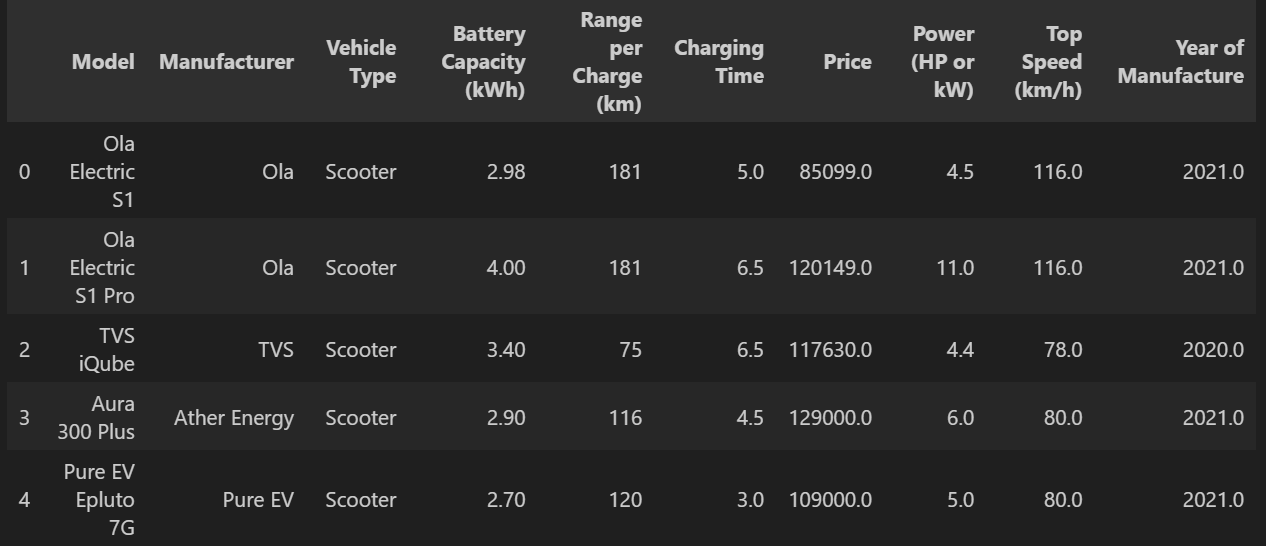
**Making Predictions :**

In this subsection I try to answer the questions that posed initially, by selecting target clusters which to the 2 wheeler EV company wants to cater to. The top 5 variables on which the segment can be targeted are- “Age”, “Total Salary”, “Profession”, “Education” and “Car\_Loan”. The target variables and chosen cluster can depend on the company policies and their goal. Some possible target clusters are chosen here. The feature values for different clusters are shown here.

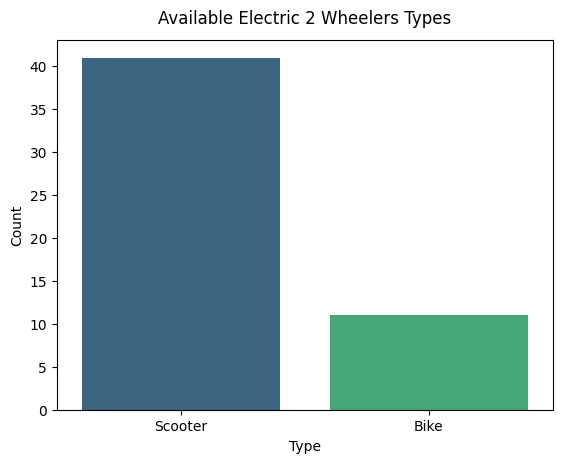
**VEHICLE SEGMENTATION** **:**

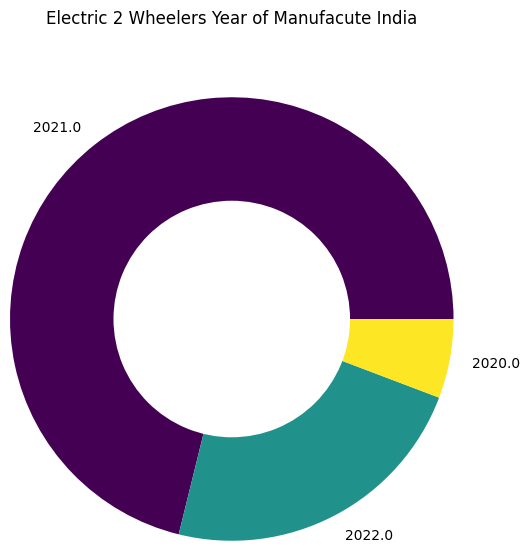
Here I perform a segmentation analysis on the possible segment to target depending on the product(2 wheelers) features during product development. Let’s explore the data through Exploratory Data Analysis.

**Exploratory Data Analysis :**



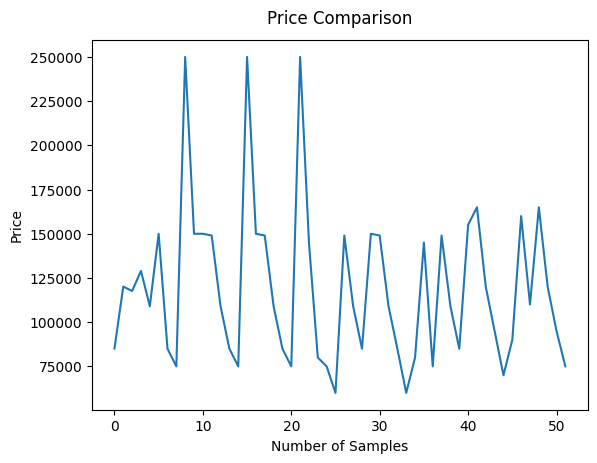
There exists two Types of 2 wheelers: Bikes and Scooters, their counts along with the year of manufacture of the 2 wheelers in the dataset :





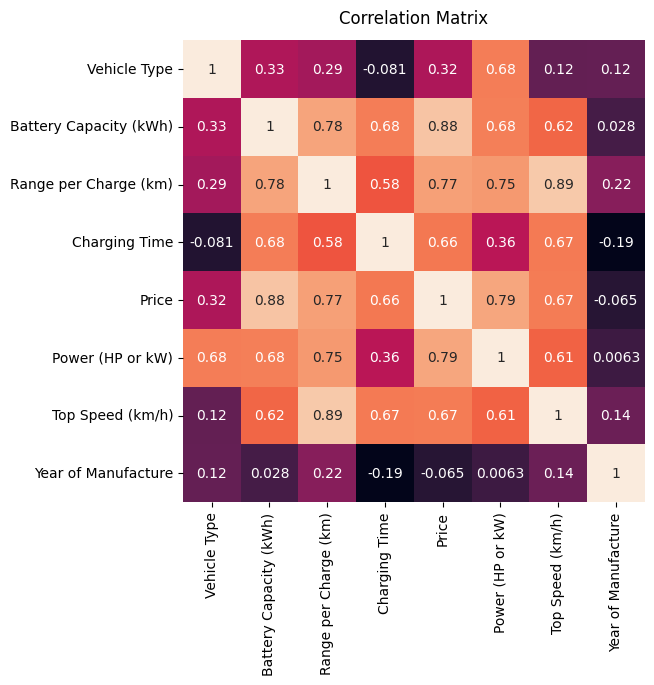
The most important features that the buyer might look for in an EV are Range, Battery Size and probably the Top Speed. This information help the buyer to choose a product. These features are plotted in bar plot.

The buying decisions also depends on the price of the product. The price values from the dataset:



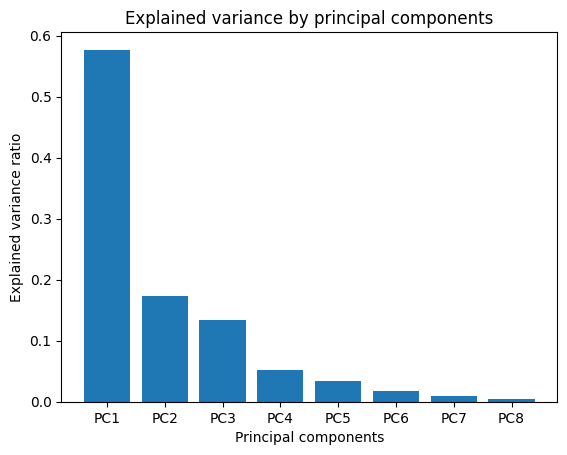
**Principal Component Analysis :**

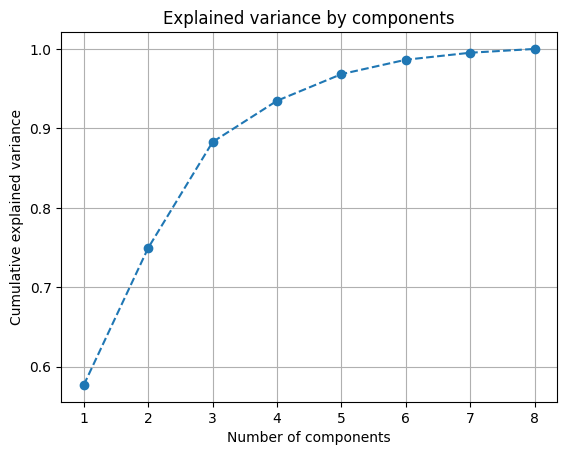
I have performed correlation analysis among the features and they had significant correlation.



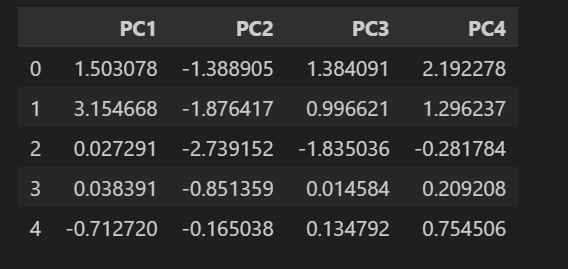
Therefore, it is necessary to preform PCA to extract independent principal components and only those which have significant information of the entire dataset.

We first scale the data to perform PCA. Taking all the features to do PCA we get the explained variance as :



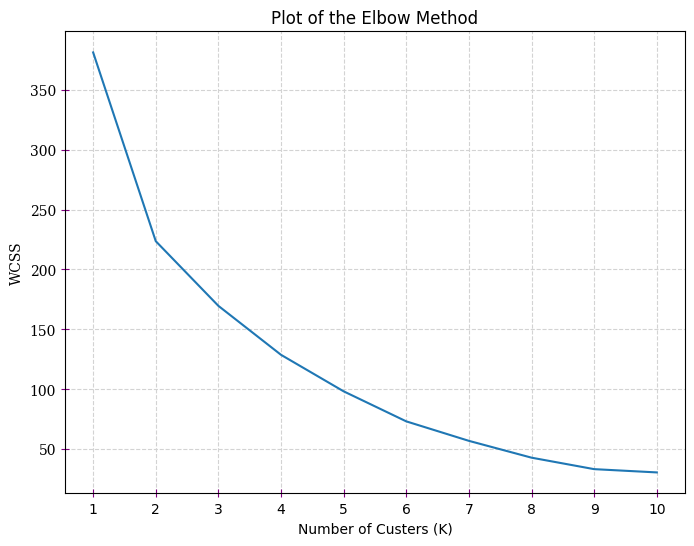


I have taken 4 Principal Components as it is enough for more than 90% of explained variance. The principal components along with the weightage of each feature with the components are :

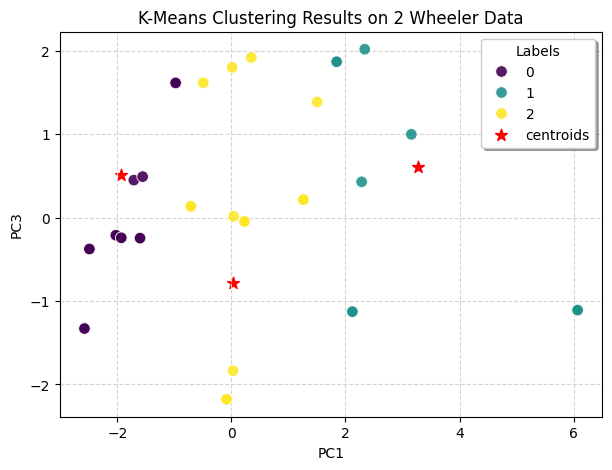


**K-Means Clustering :**

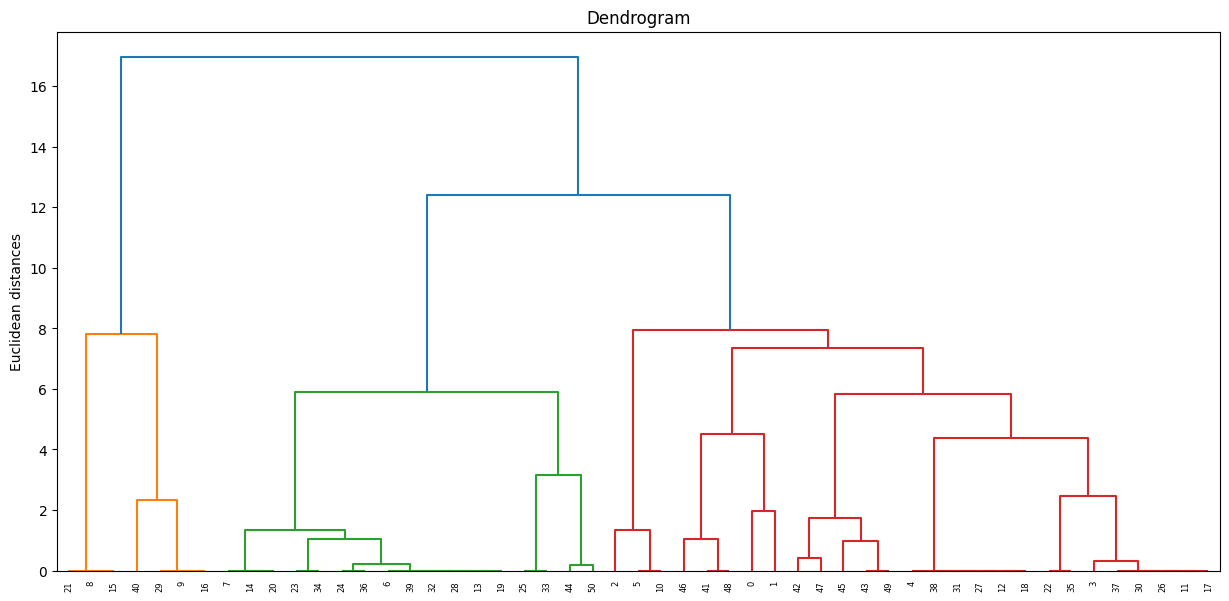
K-means takes number of clusters as an input as it itself can’t determine number of clusters. To do that I rum k-means for various number of clusters and choose the optimal one. One point to note that I have used “k-means++”. Only difference from normal one is that it initialises the clusters smartly rather than randomly in normal k-means. The elbow plot :



**K - Means Clustering :**

****

**Hierarchical Clustering :**

****

The EV Manufacturer Company might choose the 1st Segment having feature values-

• Manufacture EV Scooter

• Price Point of 85k-95k

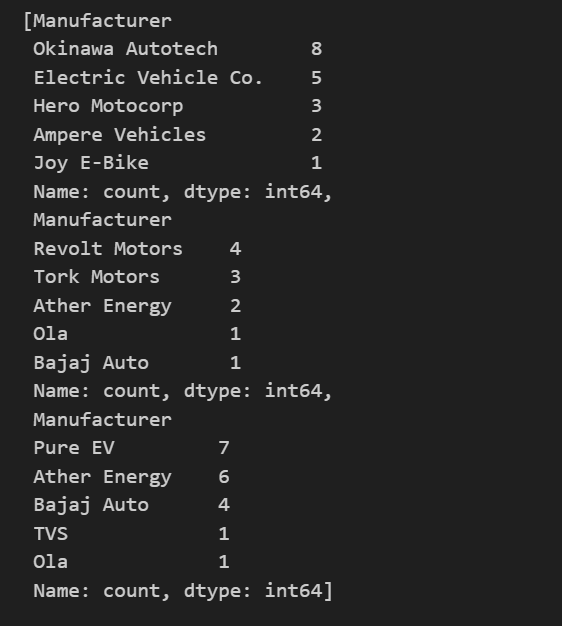
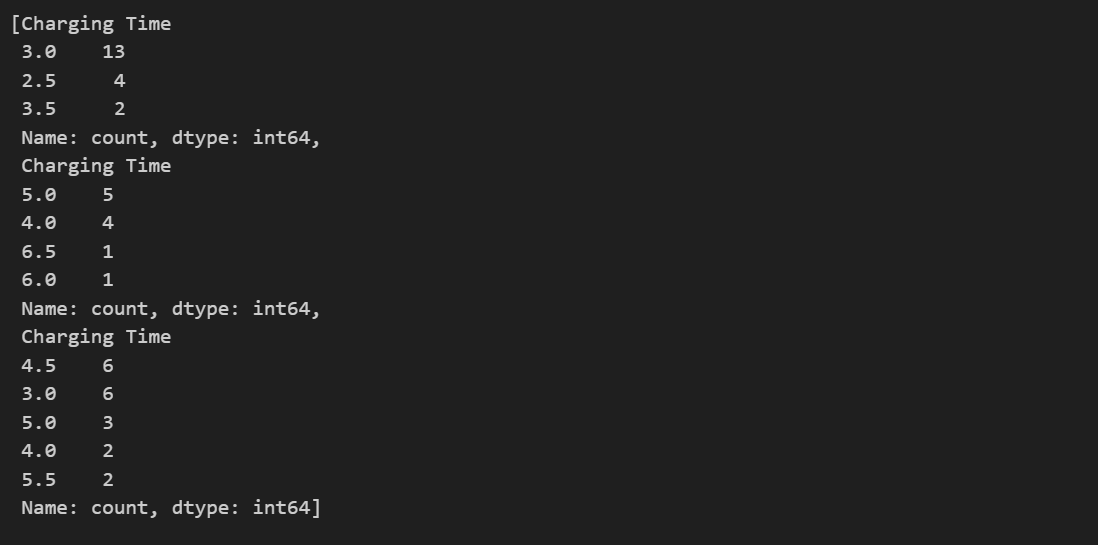
• Top Speed having 50-70 km/h

• Range Having between 80-120 km,

• Battery Capacity of 2.2-3 kWh

• Charging Time of 2.5-2.5 Hrs

• Already Present Manufacturers in the Segment are Hero, Okinawa, Ampere, etc.

**Another Possible Segment :**

to target is 3rd Segment with features as-

• Manufacture EV Scooter

• Price Point of 85k-149k

• Top Speed having 80-116 km/h

• Range Having between 95-181 km,

• Battery Capacity of 2.7-4 kWh

• Charging Time of 3-5.5 Hrs

• Already Present Manufacturers in the Segment are Ola, Aether, Bajaj , TVS, etc.

**MARKET MIX :**

The four main points of a marketing mix called 4P’s- Product, Place, Promotion, and Price. They specify the approaches and strategies that address the target market, based on the details of the marketing plan. Product Each product line represents a group of outputs or products. The set of all the product lines is called the product mix. The company can have different Kind of Products, Models of EV. They might offer-

• Automobiles

• Automobile parts

• Commercial vehicles

• Financial services

**Pricing Strategies :**

The Company should have a strategy for the pricing of it’s products depending on the segments the company targeting. Both Premium and Affordable strategies are necessary e.g-

• Market-oriented pricing strategy

• Premium pricing strategy

**Promotion :**

One of the most important task for the Company is to tell the customer about the product they have produced in a convincing way. The activities are-

• Advertising (Primary Method)

• Direct marketing

• Personal selling

• Sales promotion

• Public relations

**Place :**

In this aspect, the virtual or physical locations of transactions are considered. Such locations are significant because they enable the company to reach target customers in specific markets.

The places can be-

• Official websites

• Dealerships

• Automotive shows and exhibits

**CONCLUSION :**

There are many 2 wheeler EV manufacturing companies in the country like Ola Electric, Hero Electric, Aether Energy, TVS, Bajaj, Okinawa, etc. 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 even though we have already started to see them spreading throughout the country. Such a segmentation analysis on EV Product, Demographic Segmentation, etc. proves to be invaluable resource to New EV Brands as well as already established brands to better understand the market and predict the dynamics of such and early market especially where investments are happening.