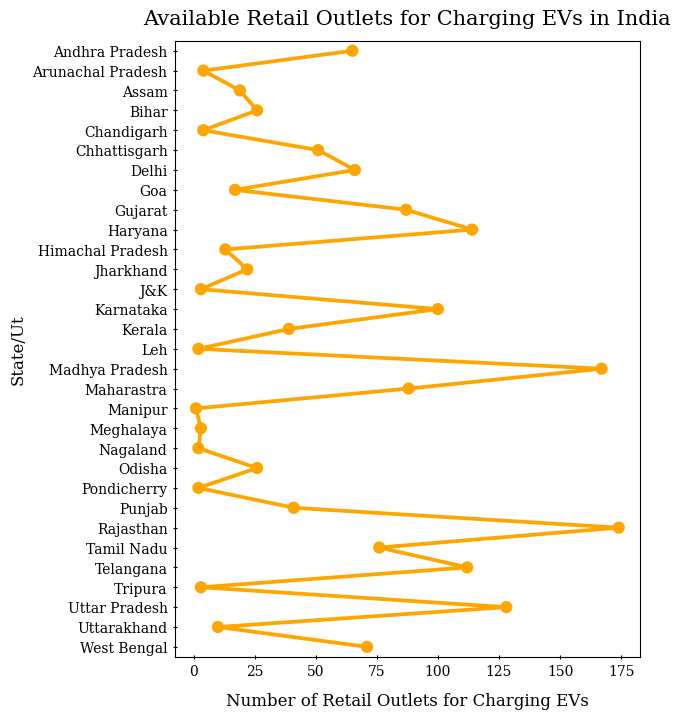
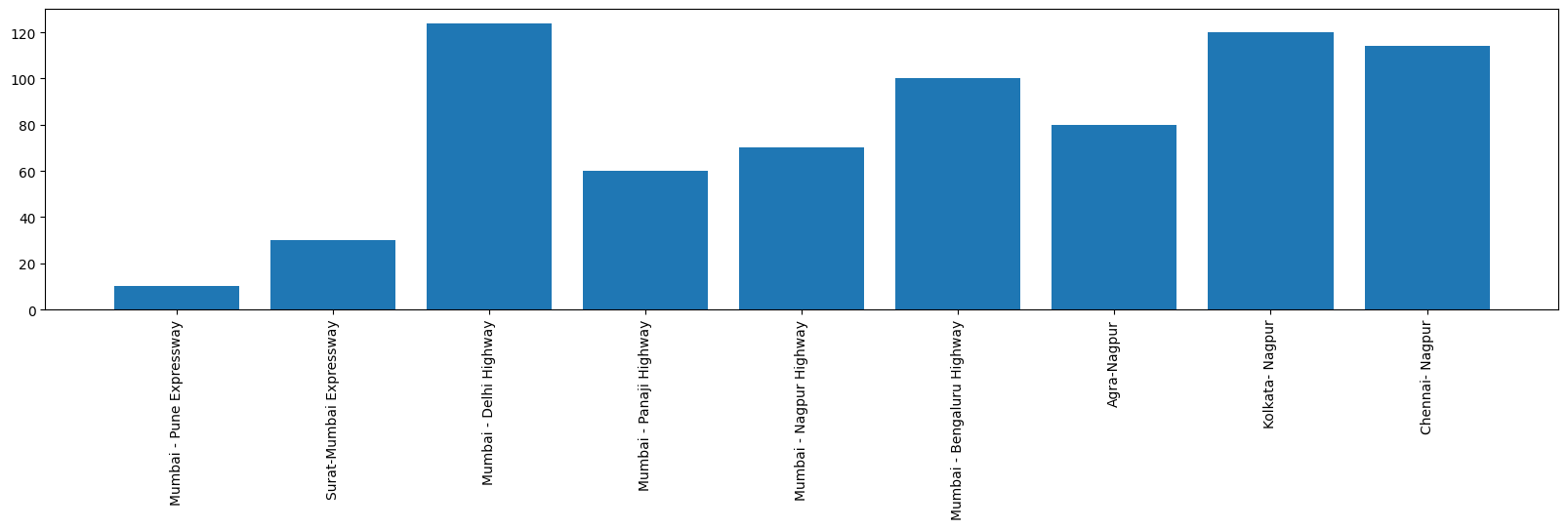
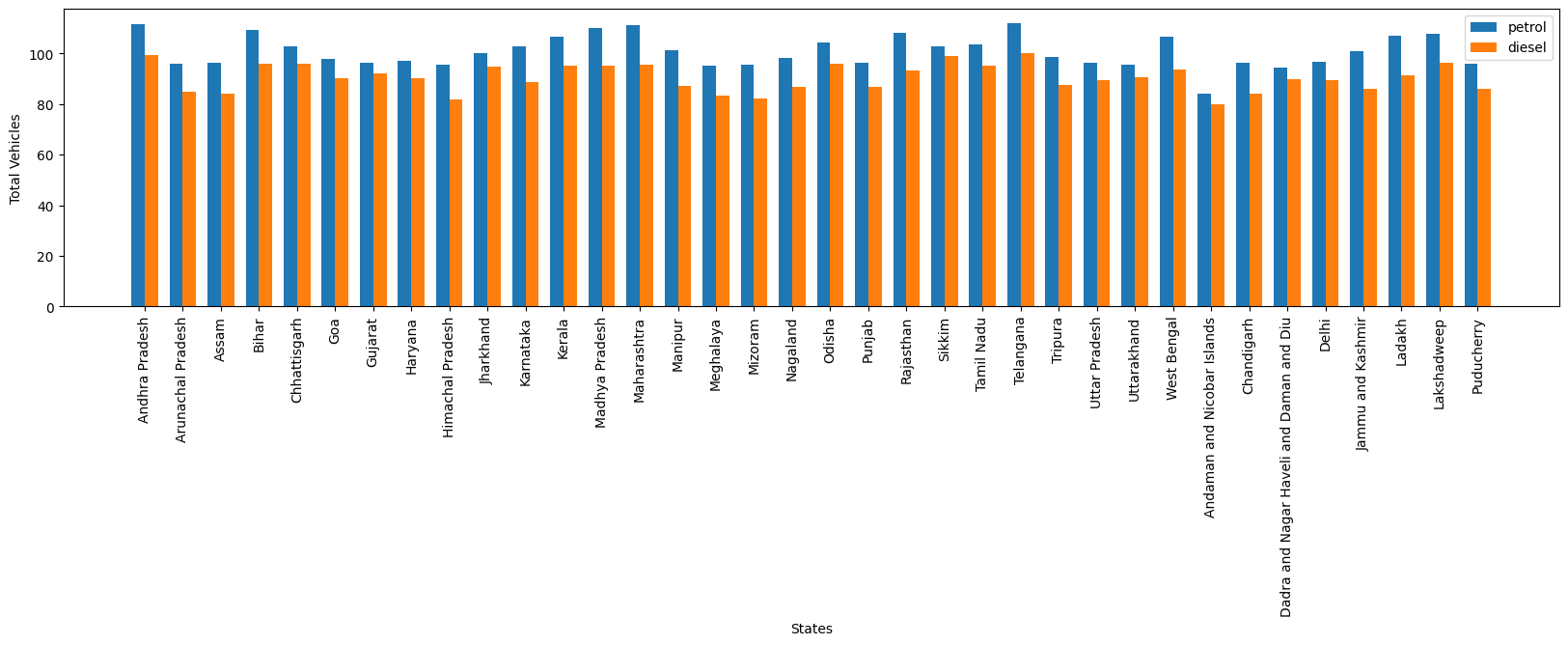
**Line Graph**



**Bar plot**



**Stacked Bar Graph**



**Observation**

The first visualisation tells us about the no. of EV charging stations on highways and the line graphs gives insights about the no. of EVs charging outlets in various Indian states. from both the visualisation we can observe that overall Maharastra and Karnataka regions leads the tally

**Target Place**

We can launch our product in the Mumbai region of Maharastra and Bengalore region of Karnataka as

* **Excellent EV Ecosystem**: Bangalore and Mumbai already have an established EV ecosystem with a growing number of electric vehicles, charging infrastructure, and support services.
* **Market Potential:** Bangalore and Mumbai are major urban centers with a high population density and a growing market for electric vehicles.
* **High Petrol & Diesel Prices-**The high petrol and diesel prices in Maharastra and Karnataka is also a important factor in EV growth.
* **Infrastructure:** Both Bangalore and Mumbai have well-developed

infrastructure and transportation networks, making them suitable for the launch of an EV product.

**K-MEANS CLUSTERING**

K-means clustering is a popular unsupervised machine learning algorithm used for partitioning a dataset into distinct groups or clusters. It aims to assign each data point to the nearest cluster centroid based on the Euclidean distance. The algorithm starts by randomly initializing K centroids and iteratively updates them until convergence. During each iteration, data points are reassigned to the cluster with the closest centroid, and the centroids are recalculated based on the new cluster assignments. This process continues until the centroids no longer change significantly. K-means clustering is widely used in various fields, such as image segmentation, customer segmentation, and anomaly detection, to uncover meaningful patterns and structures within data

**Packages/ Tools used:**

1. NumPy: To calculate various calculations related to arrays.

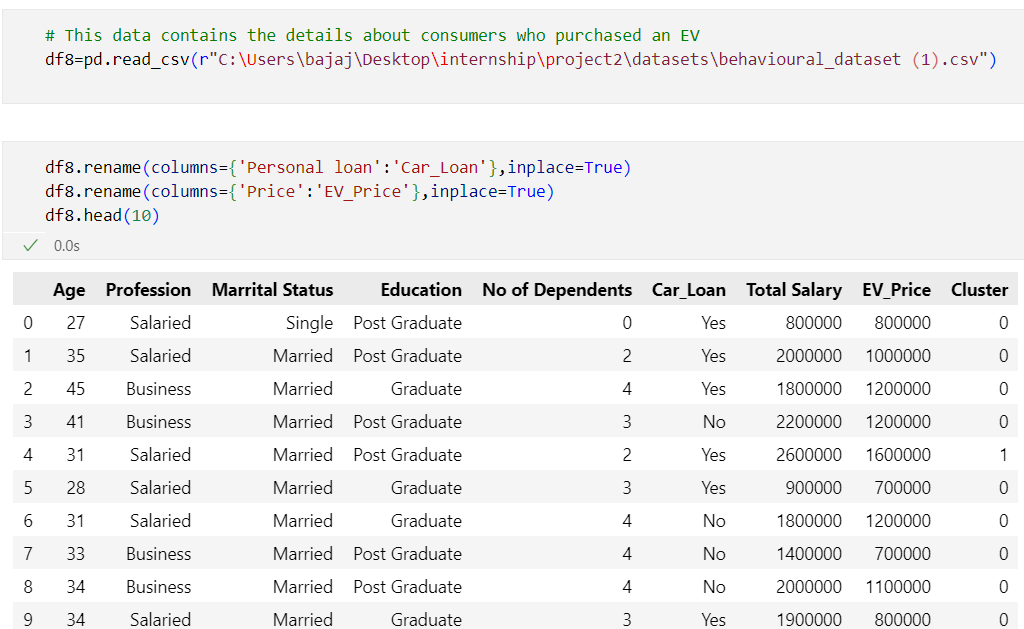
2. Pandas: To read or load the datasets.

3)Matplotlib: to plot visuals on the data

**Data loading and preprocessing**

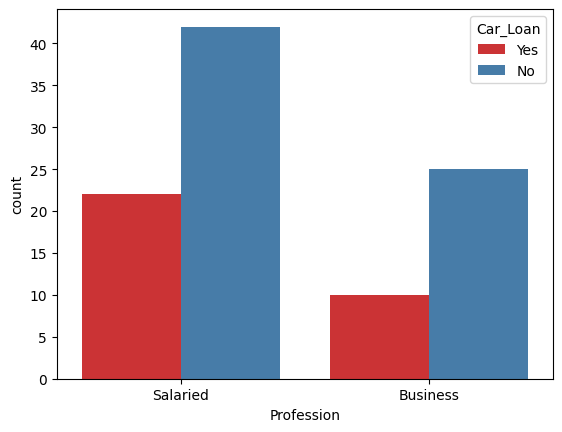
The data contains various variables among which few important variables are:

1)Age,2)Education,3)Car loan,4)Total salary,5)EV\_price

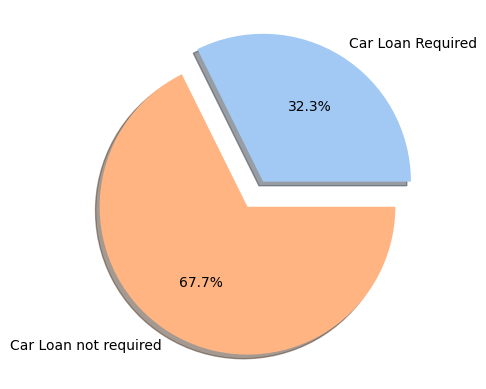
****

**Data Visualisation**

**Bar chart**

****

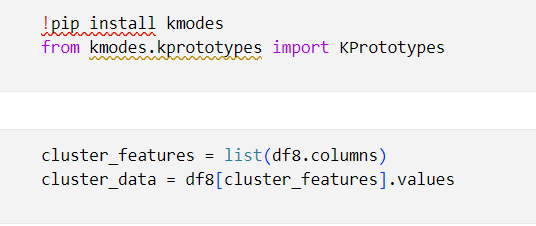
**Pie chart**

****

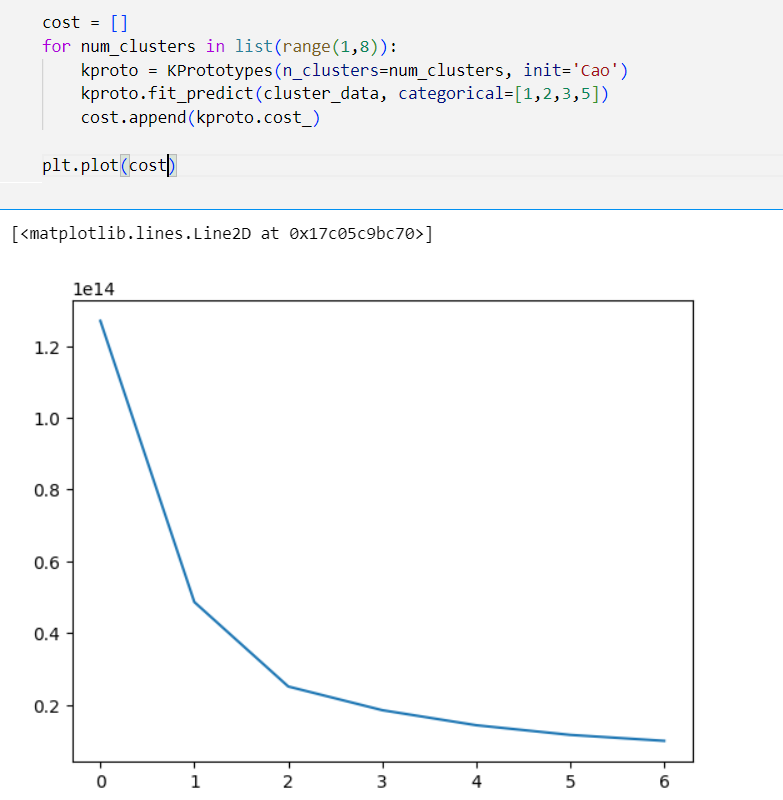
The 1st graph gives us insight about the count of no. of married and single people taking car loan

The 2nd graph gives us insight about the percentage of total no. people taking car loan & % of people not taking car loan.

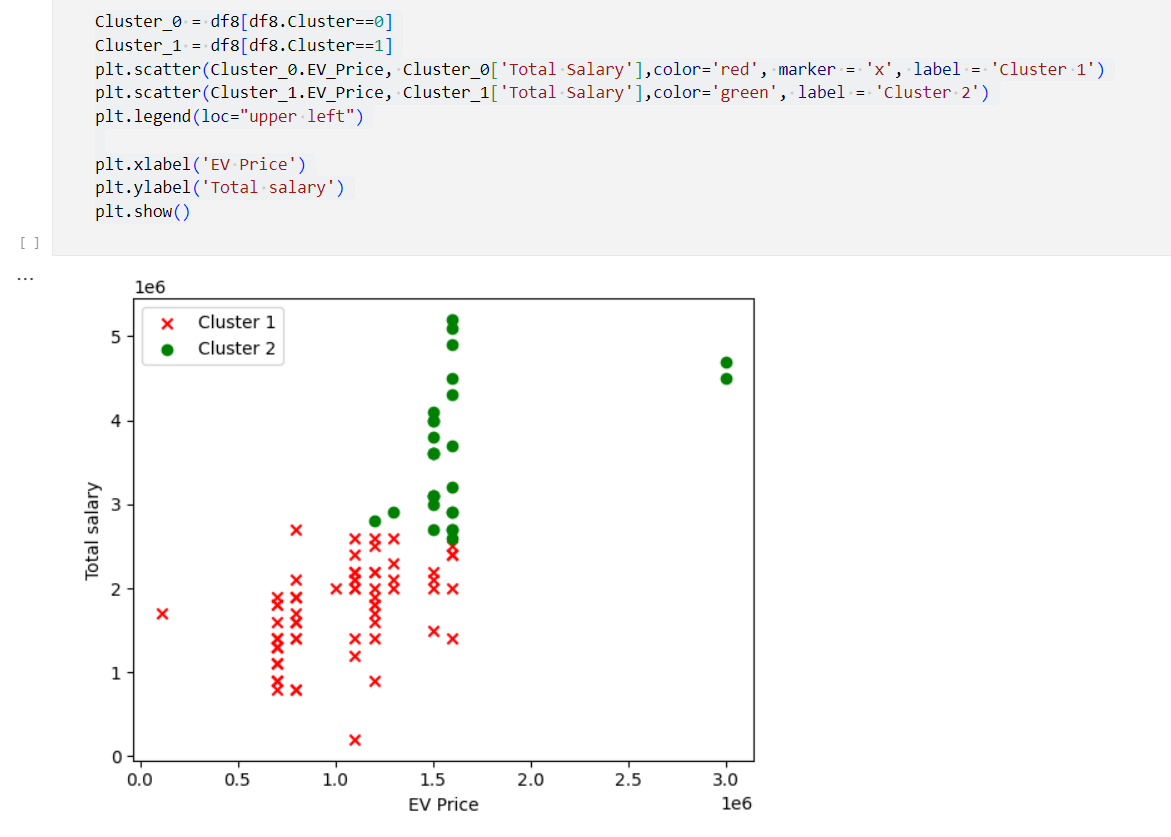
**MODEL TRAINING & FITTING**

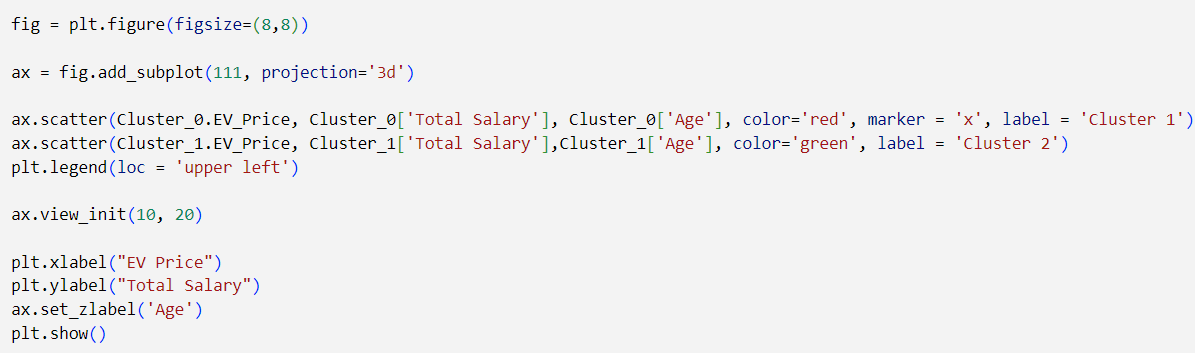
****

Here we have used the **PCA** technique to reduce the dimensionality of the data.We have used the elbow curve to find the optimal number of clusters.

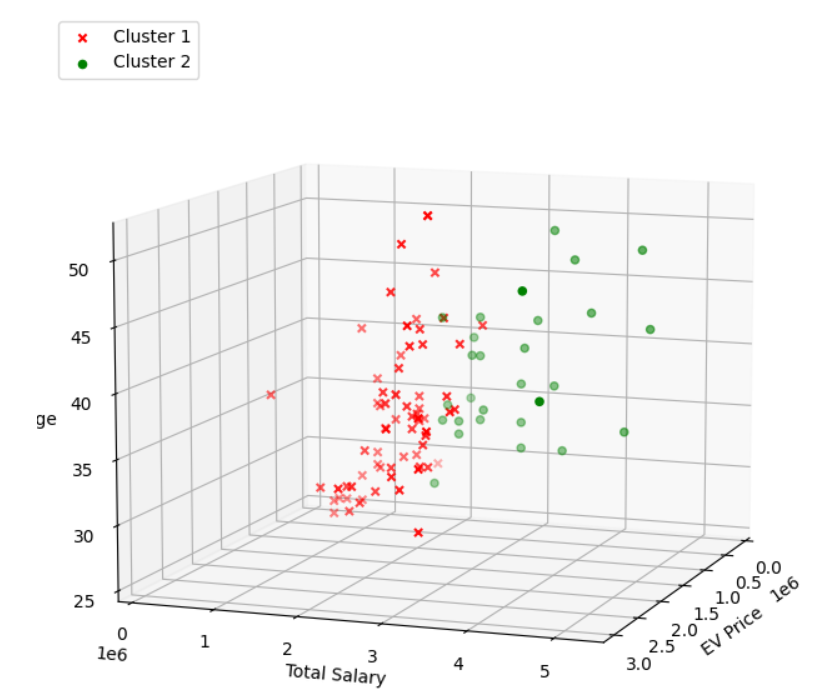
****

**ANALYSING CLUSTERS**





**3D VISUALISATION**



**Observations & Insights**

1)The optimal number of clusters for the given data comes out to be 2.

2)As the total salary increases the amount of money spent on the car increase.

3)The amount of money spent is also proportional to age as age increases, money spent also increase.

**Target Segment**

The above data visualisation and analysis completely tells about the trend that high salaried,old people prefers the Electric vehicles but keeping in mind the rising prices of petrol and diesel and increasing awareness about vehicle pollution in the younger people influence their decision to buy EVs.So the prefer market segment must be the **mid tier** which caters to both the need of young as well as old people.