

# MARKET SEGMENTATION ANALYSIS OF EV MARKET IN INDIA.

## Problem Statement:-

Task is to analyze 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 Behavioral.

## DATA SOURCES

The data used for this study are obtain from

‘Indian automobile buying behaviour study.csv’

<https://www.kaggle.com/datasets/karivedha/indian-consumers-cars-purchasing-behaviour>

## Data Pre-Processing

The Libraries that we have used for data pre-processing are as follows:-

- 1) Numpy
- 2) Pandas
- 3) Seaborn
- 4) Matplotlib

## Exploratory Data Analysis

It is important to understand the Car Buying Behaviour of the Indian Consumers. It will help the startup to target the correct segment and thereby start manufacturing of the correct type of vehicle in order to attract their target consumers. The first look of the dataset is as follows:-

```
In [2]: 1 df=pd.read_csv('Indian automobile buying behaviour study 1.0.csv')
        2 df.head()
```

Out[2]:

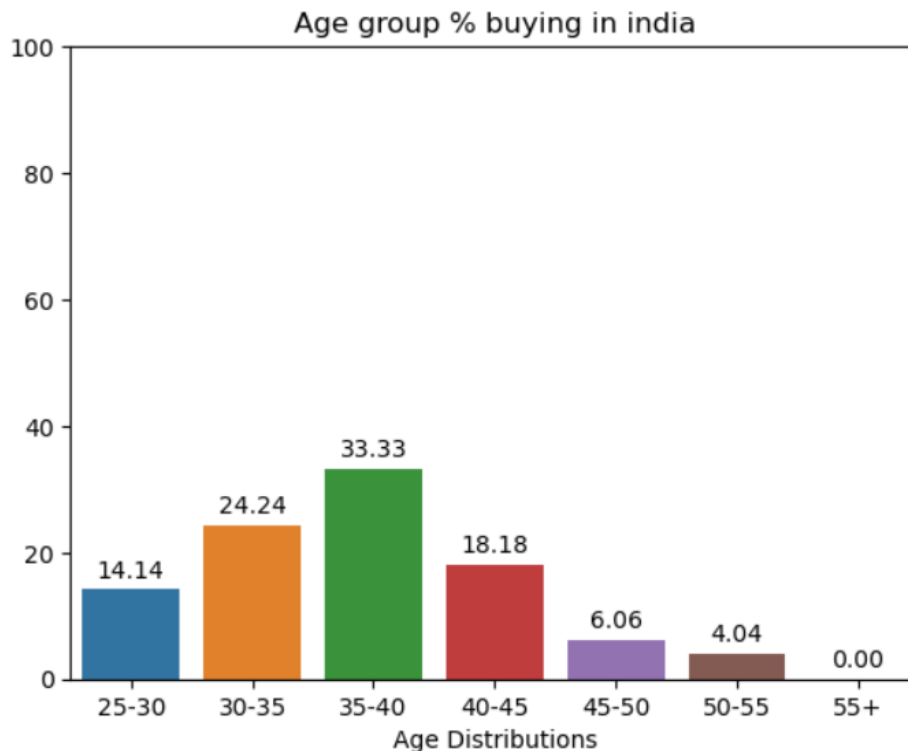
	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	House Loan	Wife Working	Salary	Wife Salary	Total Salary	Make	Price
0	27	Salaried	Single	Post Graduate	0	Yes	No	No	800000	0	800000	i20	800000
1	35	Salaried	Married	Post Graduate	2	Yes	Yes	Yes	1400000	600000	2000000	Ciaz	1000000
2	45	Business	Married	Graduate	4	Yes	Yes	No	1800000	0	1800000	Duster	1200000
3	41	Business	Married	Post Graduate	3	No	No	Yes	1600000	600000	2200000	City	1200000
4	31	Salaried	Married	Post Graduate	2	Yes	No	Yes	1800000	800000	2600000	SUV	1600000

This dataset has no missing values as shown below

```
In [4]: 1 df.info()
```

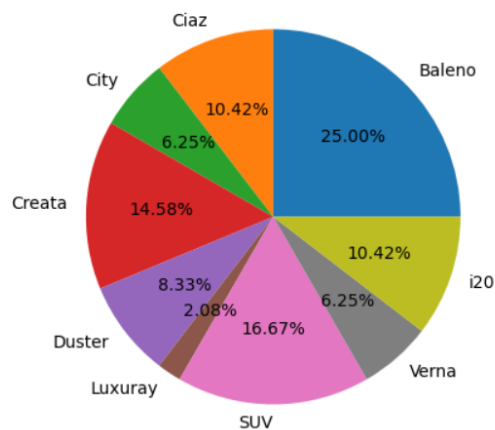
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 99 entries, 0 to 98
Data columns (total 13 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Age                   99 non-null    int64
1   Profession            99 non-null    object
2   Marrital Status      99 non-null    object
3   Education             99 non-null    object
4   No of Dependents     99 non-null    int64
5   Personal loan        99 non-null    object
6   House Loan           99 non-null    object
7   Wife Working         99 non-null    object
8   Salary               99 non-null    int64
9   Wife Salary          99 non-null    int64
10  Total Salary         99 non-null    int64
11  Make                 99 non-null    object
12  Price                99 non-null    int64
dtypes: int64(6), object(7)
memory usage: 10.2+ KB
```

The visualization of the dataset will help to provide a lot of useful information related to the dataset. It can allow us to analyse the Car Buying Behaviour of the Indian Consumers and hence make lots of inferences from it. Hence the visualization of the dataset containing Indian Consumer's Car Buying Behaviour is as follows –

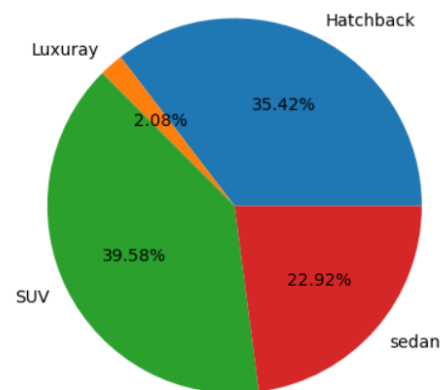


The above bar graph shows that the maximum people who are buying cars in India lie in the Age-Groups of 30 – 45 years. Hence targeting this age group seems to be a good option for the Company.

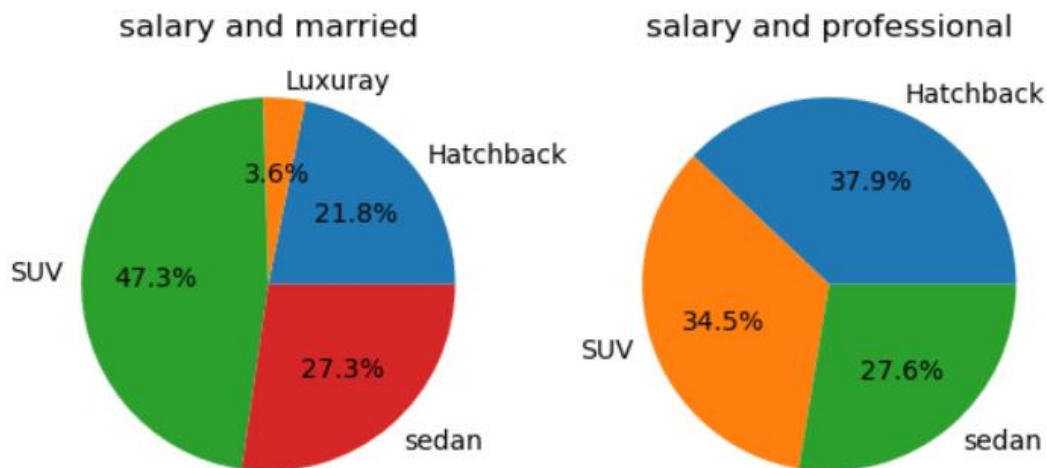
Married and No.of.Dependents>=3 'Car Make' chart



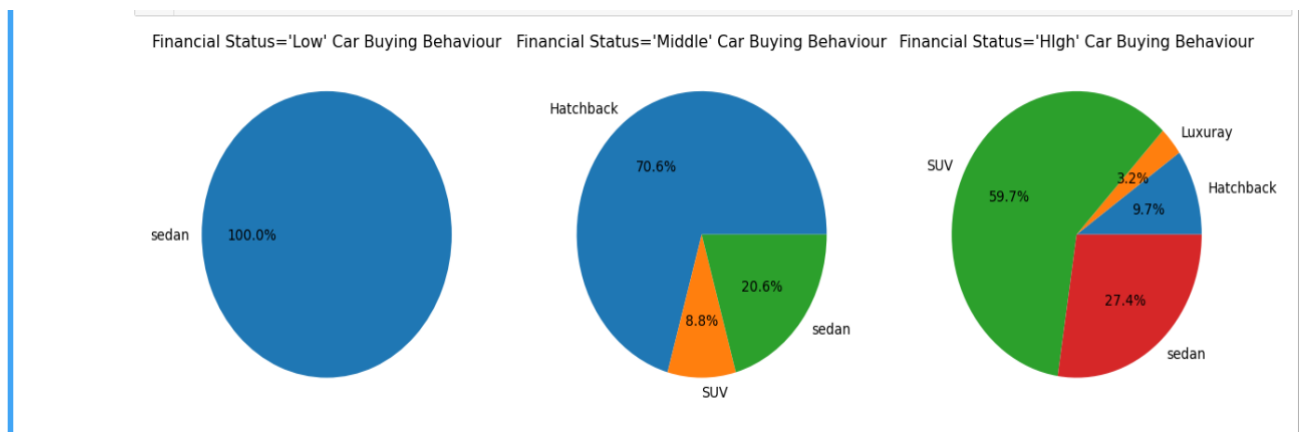
Married and No.of.Dependents>=3 'Car Type' chart



The above pie chart shows that for Married People who has number of dependents more than 3 people tend to buy SUV more than other Car types. This will give an idea to the Company that it will be a good idea to start manufacturing an Electric SUV in order to target these people.



The above pie chart shows that for Consumers who are salaried and marrieds about half of them tend to buy an SUV whereas about 35% of the consumers who are married and have their own Business also tend to buy SUV as well.



The above pie charts compare the Financial Status of the Indian Consumers. It shows that all those people whose financial status is low end up buying Sedan car type as it is affordable for them. Whereas the consumers whose Financial status is Middle Class tend to buy Hatchback car type more as compared to an SUV or a Sedan. On the other hand the consumers whose Financial Status is high tend to buy SUV more often than Sedan or Hatchback. Hence the Financial Status of a consumer greatly affects their choice of Car. The company can use this information to target consumers of a particular target segment and then can make Electric Car of the particular type which can woo its target consumers.

After doing all these data visualizations it becomes very important to convert the categorical data into numerical data so that it can be further used for the Segmentation purposes. Hence to convert the categorical data into numeric data dictionaries were used in which the keys

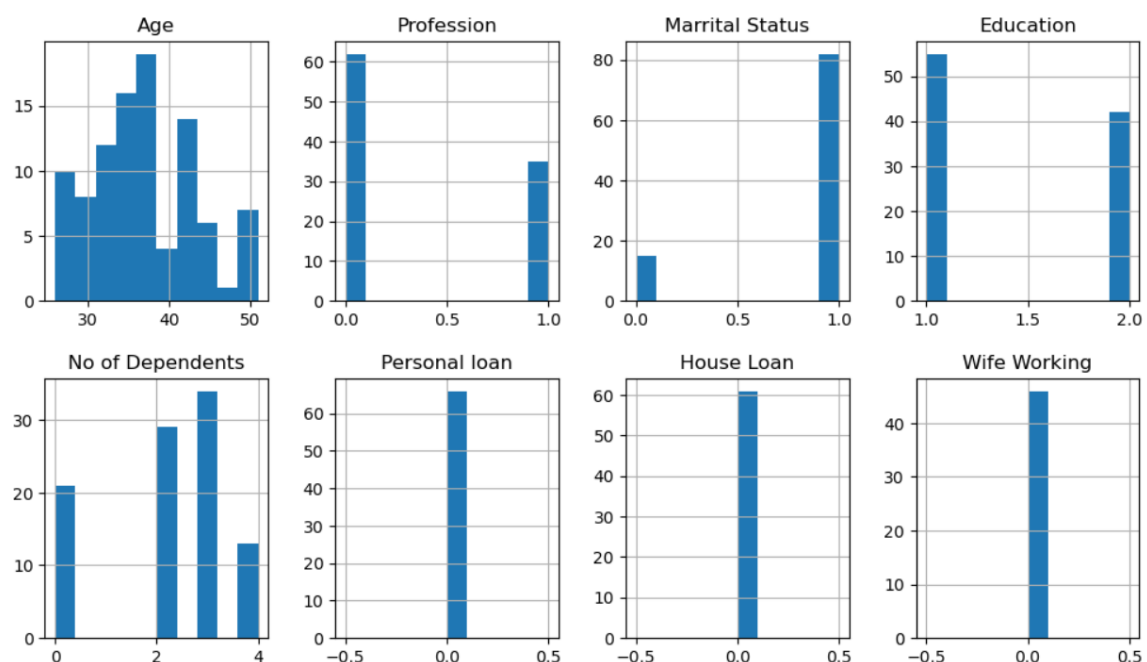
were the categorical values and the values were their corresponding numerical values. Once such dictionaries were created they were mapped into the dataset so that the dataset can be converted from categorical values to numerical values.

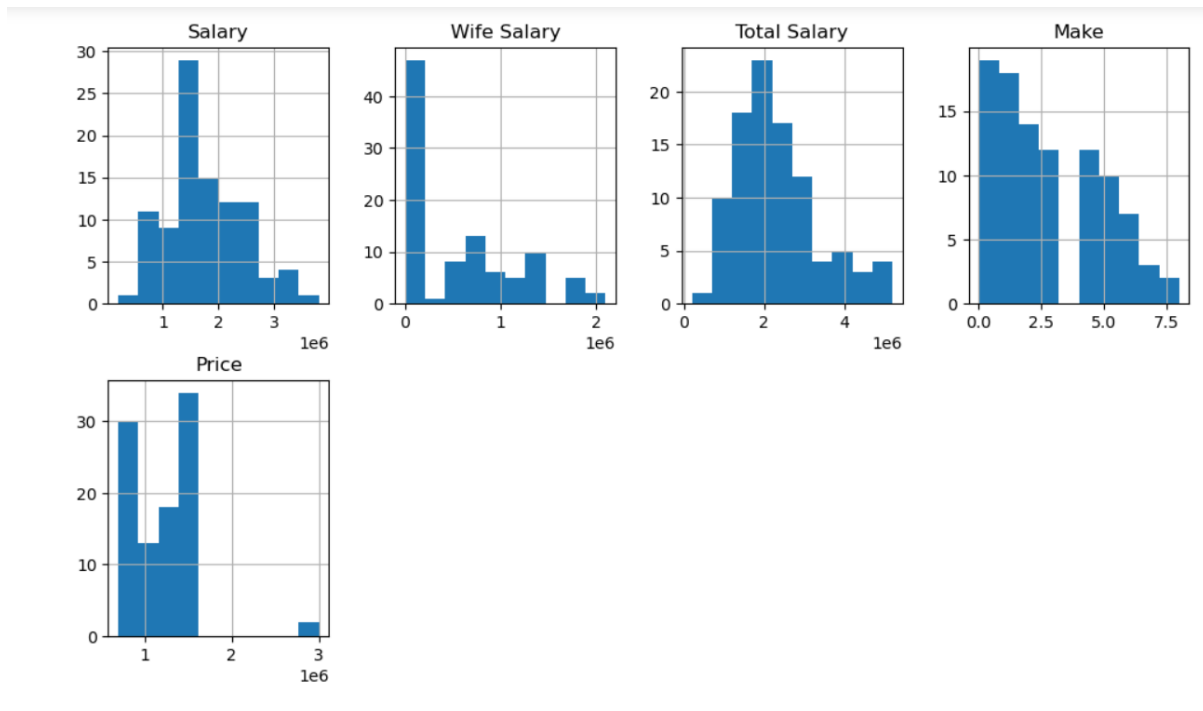
Columns having categorical values are being converted into integers by using the below mentioned transformations.

- Column 'Profession' has two values, 'Salaried' and 'Business', change it to '0' and '1'.
- Column 'Marrital Status' has two values, 'Married' and 'Single', change it to '1' and '0'.
- Column 'Education' has two values, 'Post Graduate' and 'Graduate', change it to '1' and '0'.
- Column 'Personal loan' has two values, 'Yes' and 'No', change it to '1' and '0'.
- Column 'House Loan' has two values, 'Yes' and 'No', change it to '1' and '0'.
- Column 'Wife Working' has three values, 'Yes', 'No' and 'm'. There is only one row which has 'm' as 'Wife Working', we are not sure if that refers to 'Yes' or 'No', so better to delete that row.
- Column 'Make' has these nine values, 'SUV', 'Baleno', 'Creat', 'i20', 'Ciaz', 'City', 'Duster', 'Verna' and 'Luxury', change them to '0', '1', '2', '3', '4', '5', '6', '7', '8' and '9' respectively.

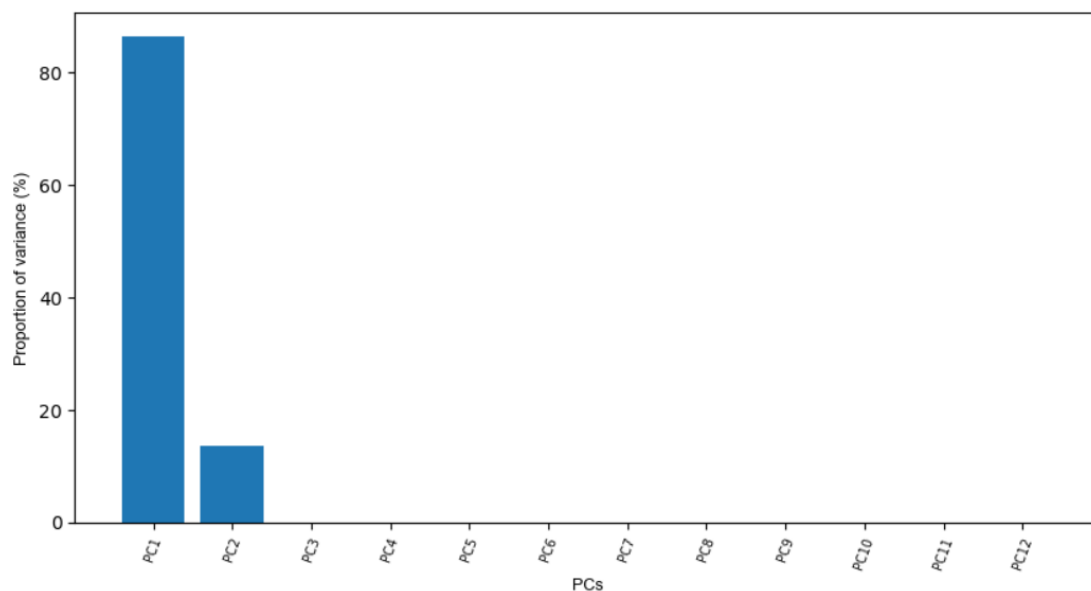
```
: 1 dic_6={'SUV':0,'Baleno':1,'Creat':2,'i20':3,'Ciaz':4,'City':5,'Duster':6,'Verna':7,'Luxuray':8}
  2 df['Make']=df['Make'].map(dic_6)
```

Then the histograms of the data were drawn as below –



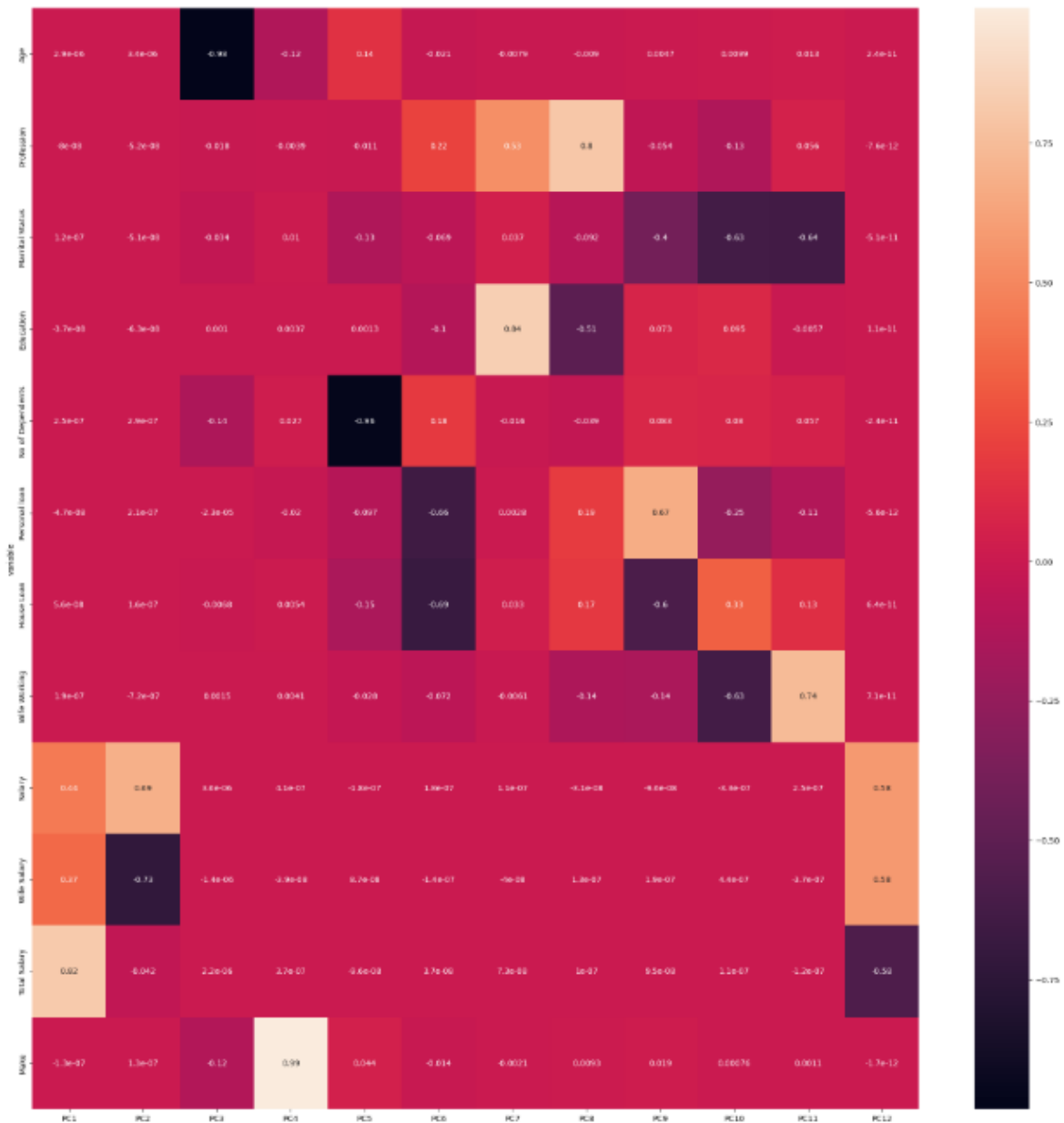


Once the histogram was plotted then the Principal Component Analysis(PCA) of the dataset was done. The PCA is a statistical process which converts a set of Correlated features into a set of linearly uncorrelated features with the help of Orthogonal Transformation. These new Transformed Features are called Principal Components. This process reduces the dimensions of the dataset and makes it cost-effective to carry out clustering operations. The Principal Components of the Indian Consumer Car Buying Behaviour dataset is as Follow –



From the above bar chart it can be clearly seen that the first 2 Principal Components themselves account for 100% of the variance in the data.

Also the heatmap between the original variables and the principal components can be shown as follows –



From the above heatmap it can be inferred that for the first principal component Salary, Wife Salary and Total Salary are the original variables which are highly correlated with it and similarly for other principal components as well.

## Segment Extraction(K-Means Methods)

Now for the extraction of the segments the K-Means algorithm has been used. K-Means algorithm is an unsupervised learning algorithm which is used to solve clustering problems in ML. It allows us to cluster the dataset into K number of group in a convenient way. The number K either must be known in advance or it must be found out using the elbow method. This algorithm allows us to identify segments of groups in an unlabeled dataset without the need for any training. The main aim of the algorithm is to divide the dataset into k number of segments in such a way that there is minimum sum of distance between a data point and its corresponding segment. First in order to extract the segments it is important to know the correct number of segments in the data. The correct number of segments can be known by

using a method called elbow methods. The elbow method help us to choose the optimum value of k by fitting the model with a range of k values. The sum of squares of distances also called the epsilon or the cost function of the K-means algorithm is used to find the elbow. The epsilon is expected to decrease with increasing value of k. The value of k after which the value of epsilon decreases very less is considered to be the optimum value of k. In our dataset we performed the elbow method 10 times to get the average value of k.

The code for elbow method is as follows –

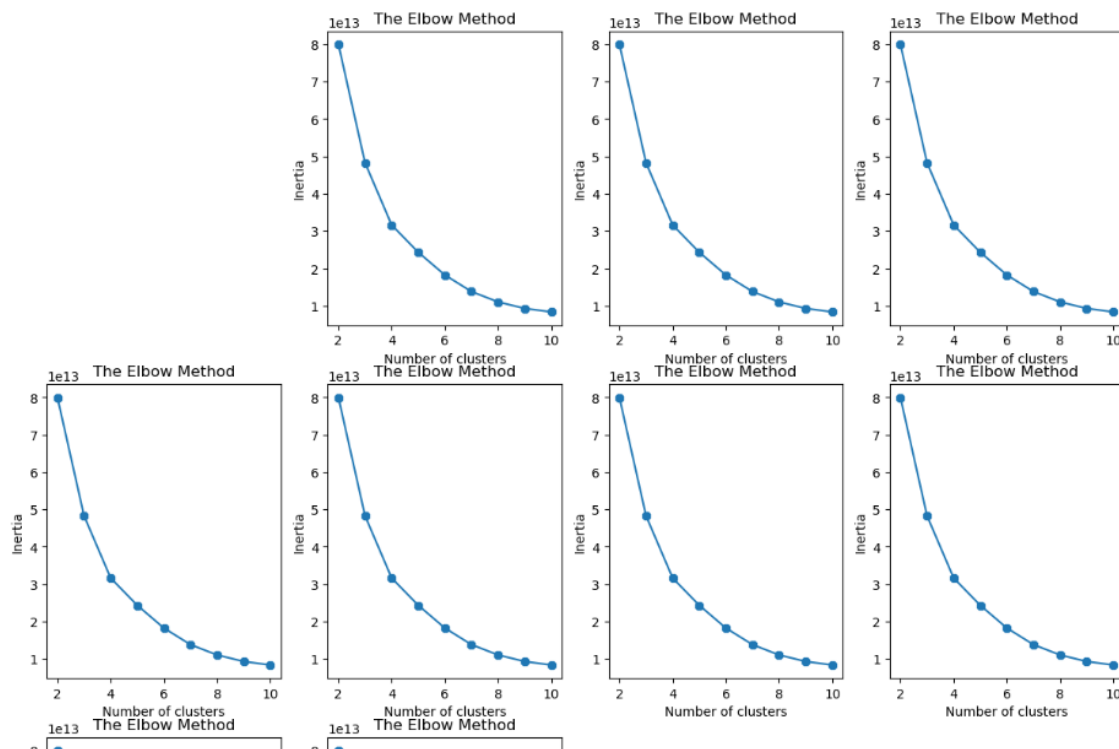
```
In [37]: 1 plt.figure(figsize=(15,14))
2 from sklearn.cluster import KMeans
3 # Get the inertia values for number of clusters varying from 2 to 10
4 for j in range(2,11):
5     inertia = []
6     for i in range(2, 11):
7         kmeans = KMeans(n_clusters = i, init = 'k-means++', max_iter = 300, n_init = 10, random_state = 0)
8         kmeans.fit(df1)
9         inertia.append(kmeans.inertia_)
10
11 # Plot the inertia values for these clusters to choose the appropriate number of clusters
12 plt.subplot(3,4,j)
13 plt.plot(range(2, 11), inertia,marker = "8")
14 plt.title('The Elbow Method')
15 plt.xlabel('Number of clusters')
16 plt.ylabel('Inertia')
17 plt.show()
```

--- The Elbow Method

--- The Elbow Method

--- The Elbow Method

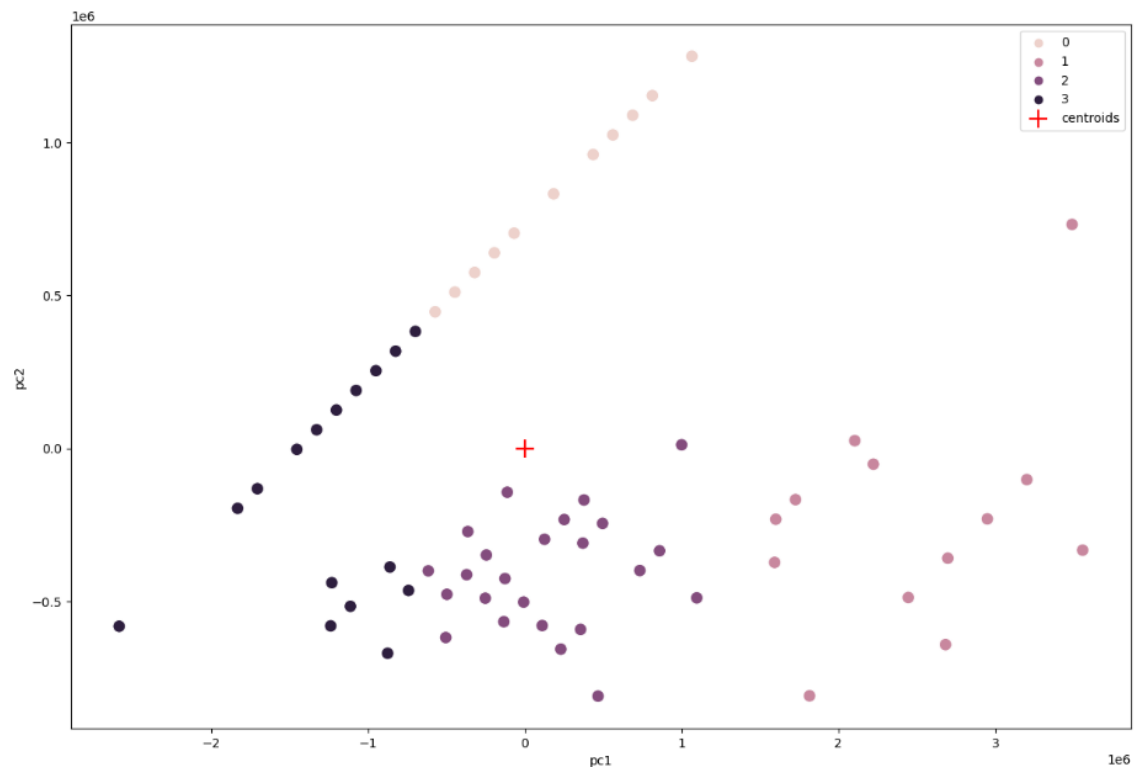
The graphs showing the decrease in the value of epsilon are as follows –



So from observing all the 10 graphs it can be concluded that the optimum number of clusters will be 4. Hence the value of K for K-Means algorithm will be taken as 4.

The clusters obtained after applying K-Means algorithm will be as follows –

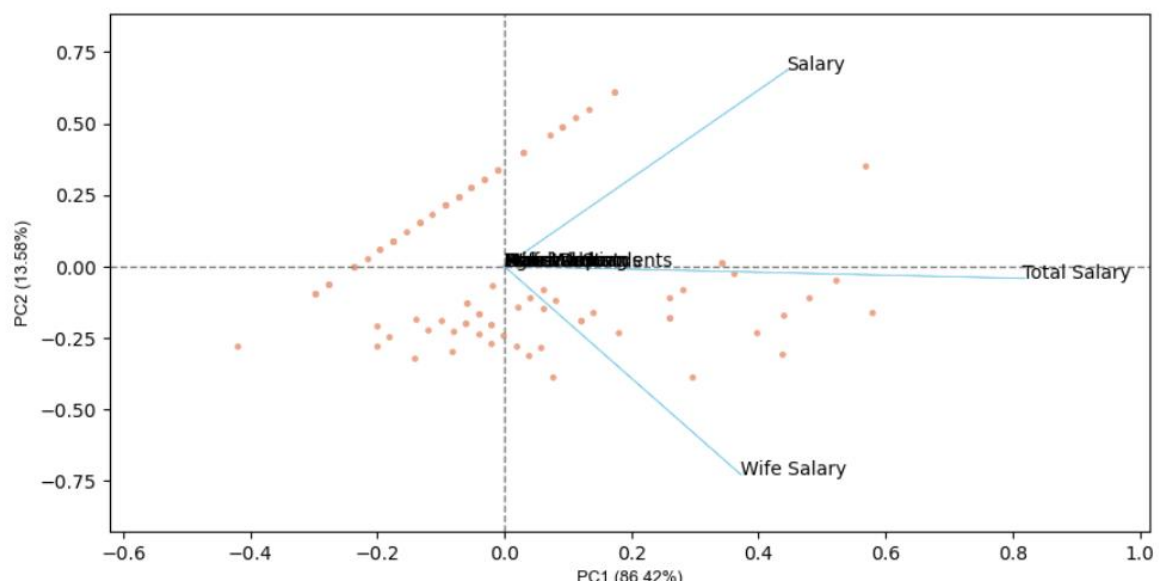




## Profiling and Describing Potential Segments

Now once we have extracted the segments from the dataset the next thing to do is to profile the segments. The Segments can be profiled by drawing a perceptual map of the dataset. This perceptual map will indicate the columns which are dominant in each segment. This perceptual map is used for segment separation using the Principal Components 1&2 for the Indian Consumers car Buying Behaviour Dataset.

The Perceptual Map will be as follows –

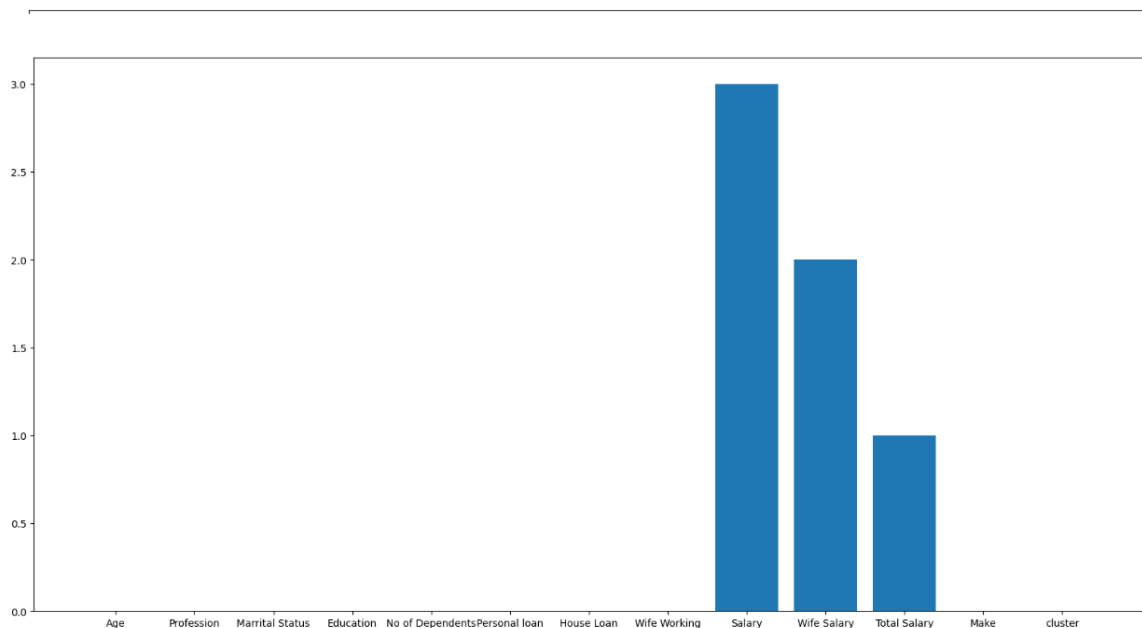


The perceptual map above shows that in the four segments the Segmentation Variables Salary, Total Salary, Wife Salary will be dominant in describing 3 segments. This can further be confirmed by plotting a segment profile plot. It can be plotted by first doing hierarchical clustering of the segmentation variables and then plotting them into a bar graph. The code for plotting the segment profile plot is as follows –

```
1 from sklearn.cluster import AgglomerativeClustering
2 from scipy.cluster import hierarchy as shc
3 from sklearn.metrics import pairwise_distances
4 clust2 = AgglomerativeClustering(n_clusters = 4).fit_predict(pairwise_distances(df1.transpose()))
5 plt.figure(figsize = (20,10))
6 plt.bar(df1.columns,clust2)
```

C:\Users\HP\anaconda3\lib\site-packages\scipy\cluster\hierarchy.py:834: ClusterWarning: scipy.cluster: The symmetric non-negative hollow observation matrix looks suspiciously like an uncondensed distance matrix  
return linkage(y, method='ward', metric='euclidean')

<BarContainer object of 13 artists>



From the above bar plot it can be confirmed that the Segmentation variables in the 4 segments are as follows –

Segment 0 – Age, Profession, Marital Status, Education, No of Dependents, Personal Loan, House Loan, Wife Working, Make.

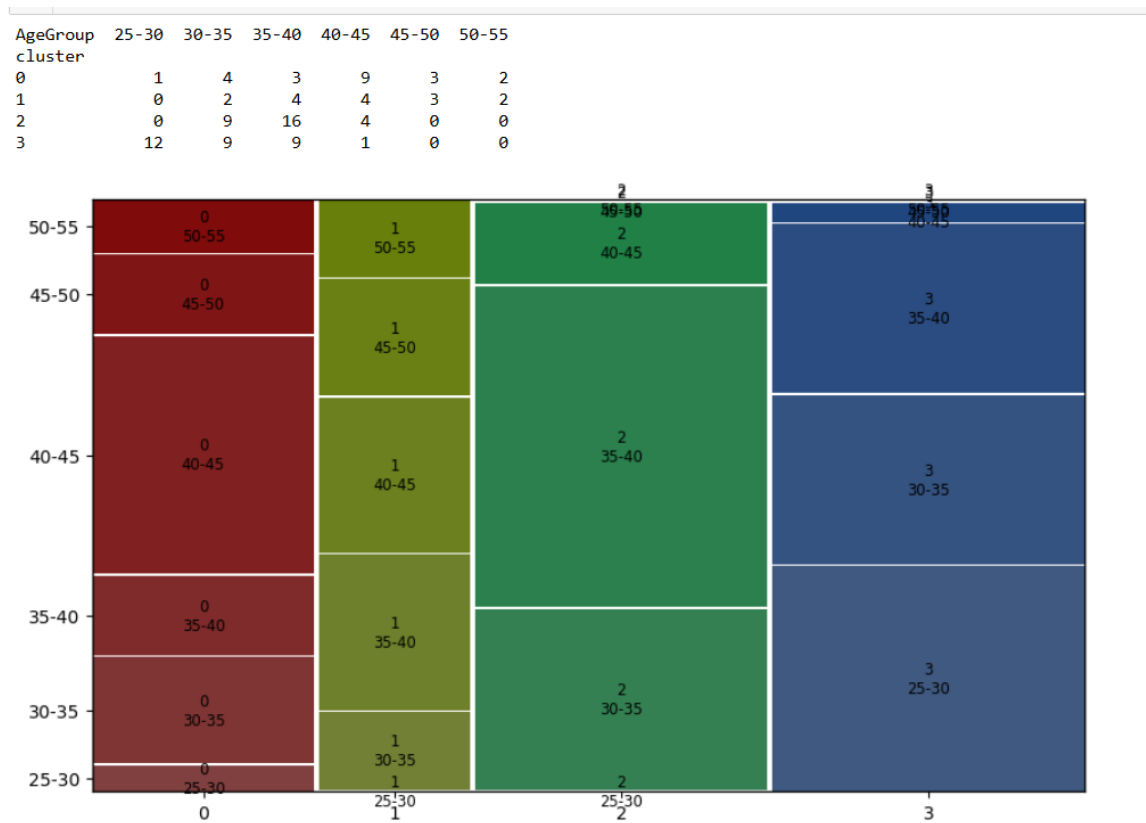
Segment 1 – Total Salary.

Segment 2 – Wife Salary.

Segment 3 – Salary.

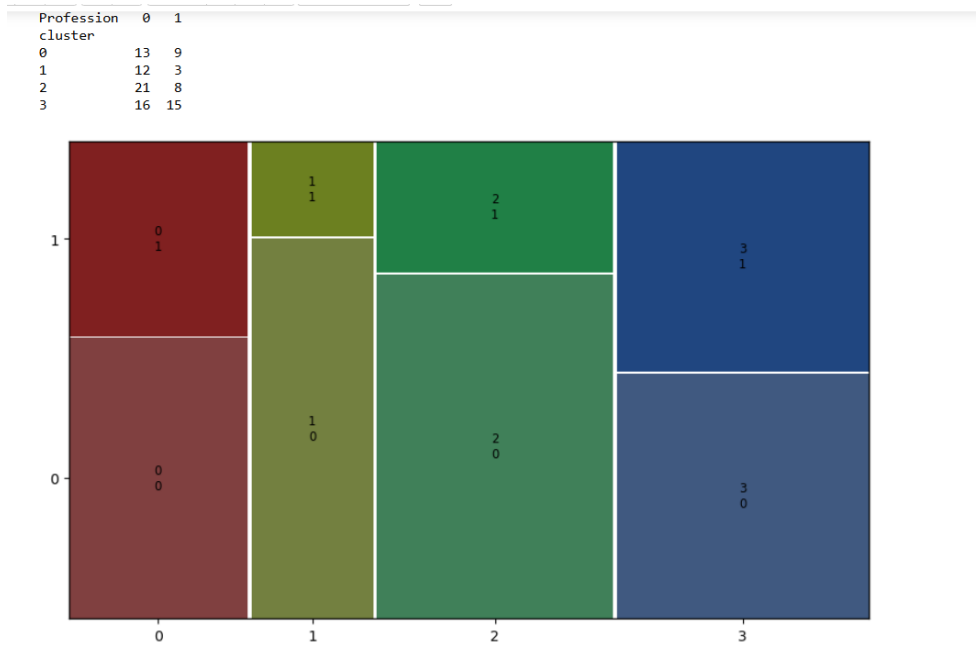
Now these segments can be described further using various visualisations. In order to describe these segments we will plot the mosaic plot of these segments which will describe these segments to us on the basis of the variables. Some of these mosaic plots are as follows –

## Mosaic Plot for AgeGroup



The above mosaic plot tells us that the Segment 0 has more people of the AgeGroup 40-45, the Segment 1 has more people in the AgeGroup 35-45, the Segment 2 has more people in the AgeGroup 35-40 and the Segment 3 has more people in the AgeGroup 25-30.

## Mosaic Plot for Profession



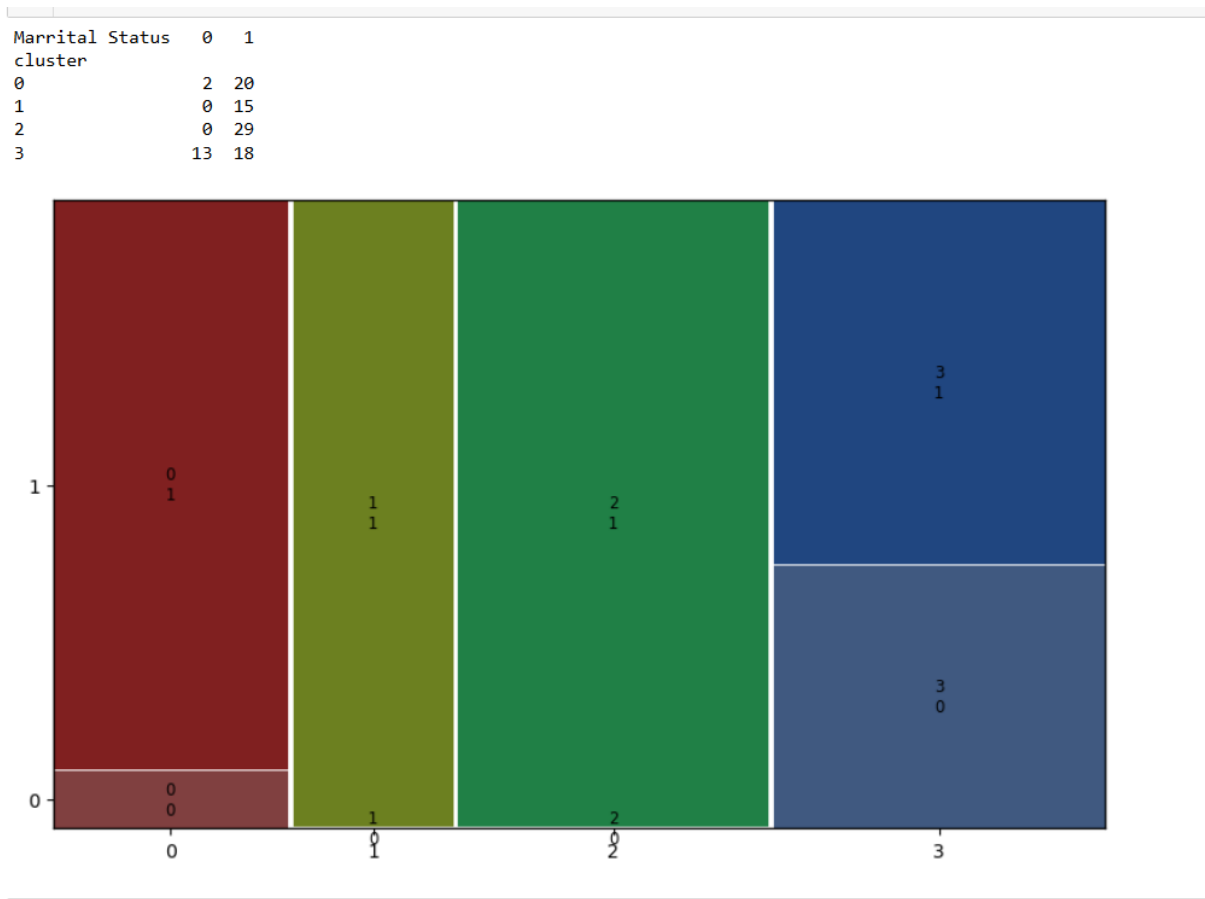
The above mosaic plot shows the relationship between Segments and profession as follows –  
Segment 0 = Salaried People -13, Business People – 9.

Segment 1 = Salaried People – 12, Business People – 3.

Segment 2 = Salaried People – 21, Business People – 8.

Segment 3 = Salaried People – 16, Business People – 15.

### **Mosaic Plot for Marital Status**



The above plot can be described as follows –

Segment 0 = Unmarried – 2, Married – 20.

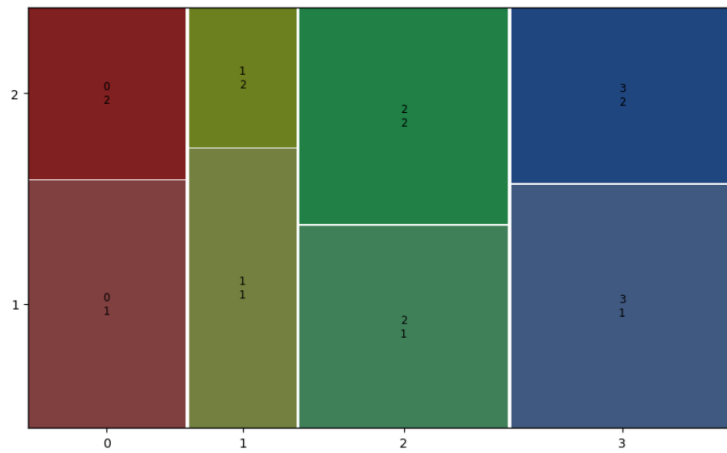
Segment 1 = Unmarried – 0, Married – 15.

Segment 2 = Unmarried – 0, Married – 29.

Segment 3 = Unmarried – 13, Married – 18.

### **Mosaic Plot for Education**

Education cluster	1	2
0	13	9
1	10	5
2	14	15
3	18	13



The above mosaic plot can be described as follows –

Segment 0 = Graduate – 9, PostGraduate – 13.

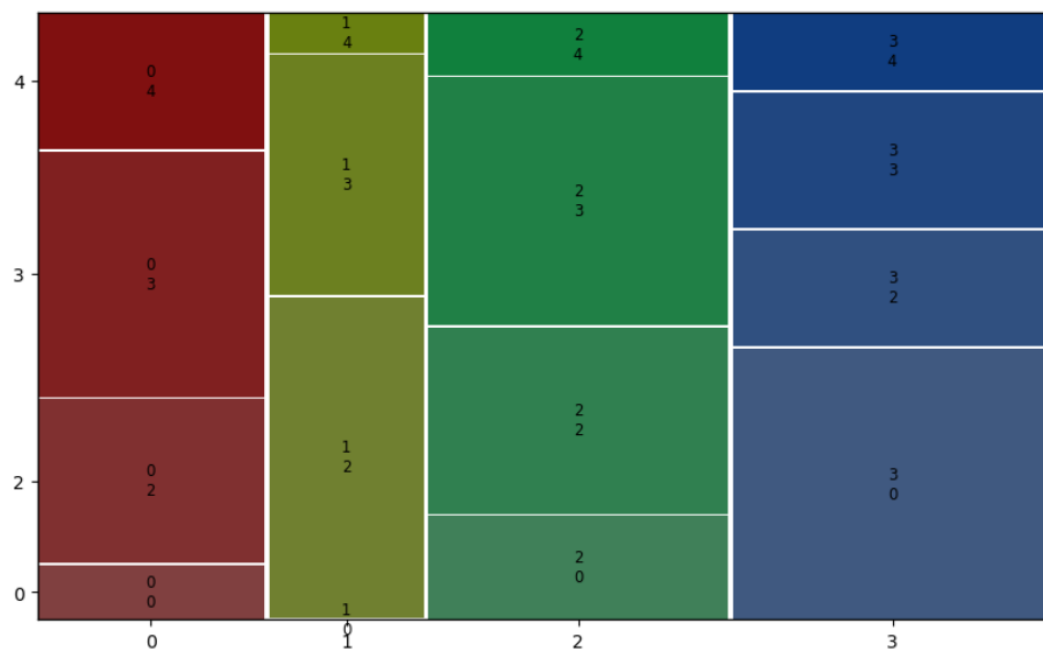
Segment 1 = Graduate – 5, PostGraduate – 10.

Segment 2 = Graduate – 15, PostGraduate – 14.

Segment 3 = Graduate – 13, PostGraduate – 18.

### Mosaic Plot for Number of dependents

No of Dependents cluster	0	2	3	4
0	2	6	9	5
1	0	8	6	1
2	5	9	12	3
3	14	6	7	4

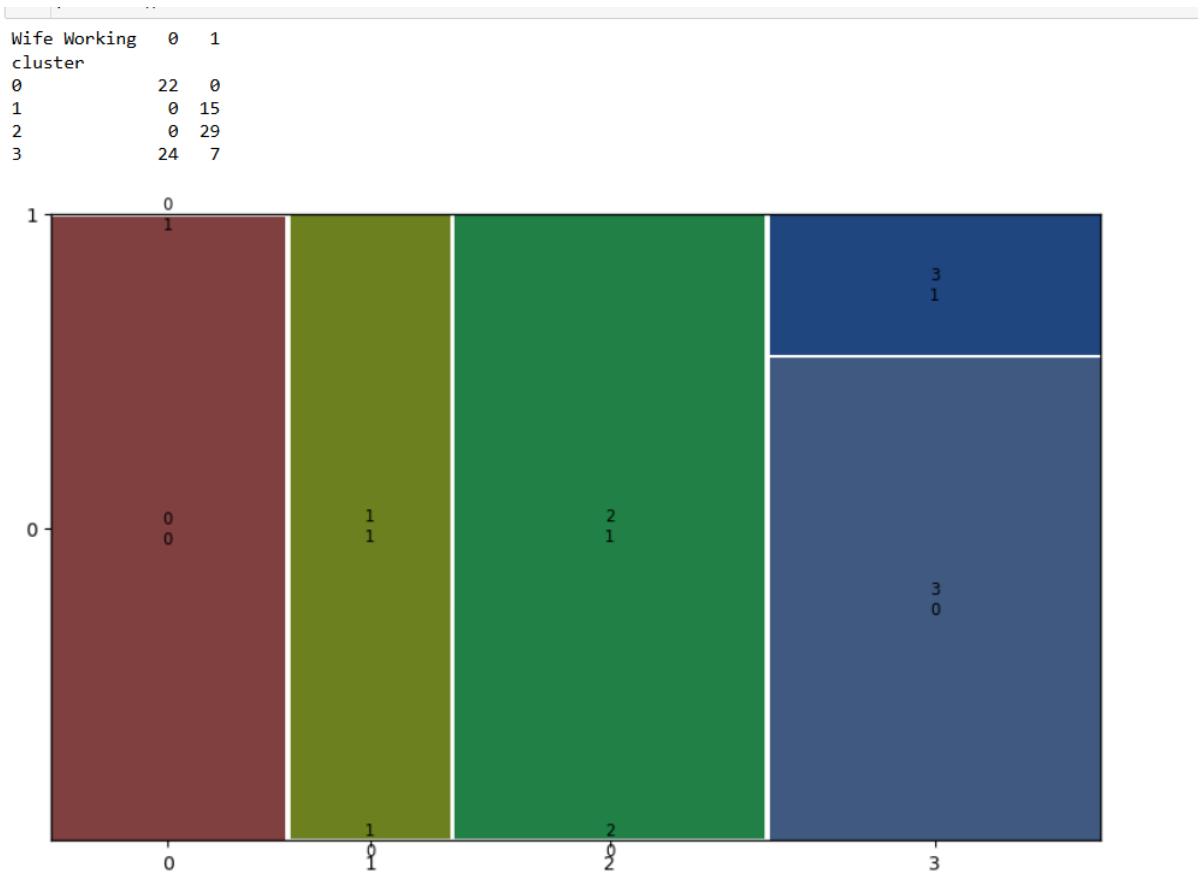


In the above mosaic plot it can be seen that –

Segment 0 and Segment 2 has most no of consumers having no of dependents 2 or 3.  
 Segment 1 has most no. of consumers having 2 number of Dependents.

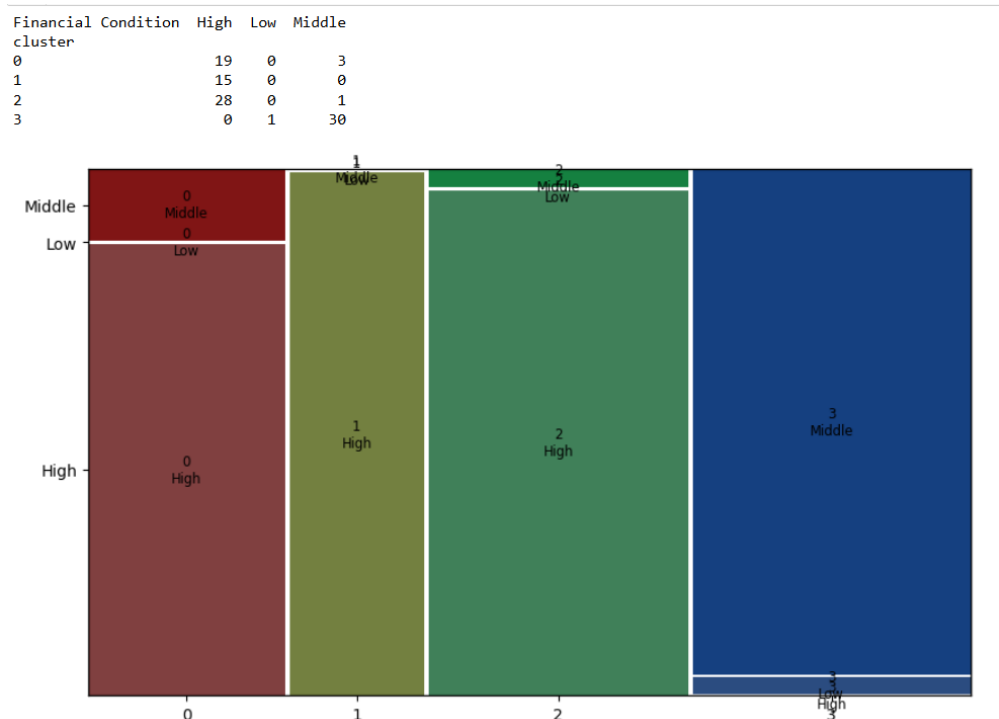
Segment 4 has most number of consumers having 0 Dependents.

### Mosaic Plot for Wife Working



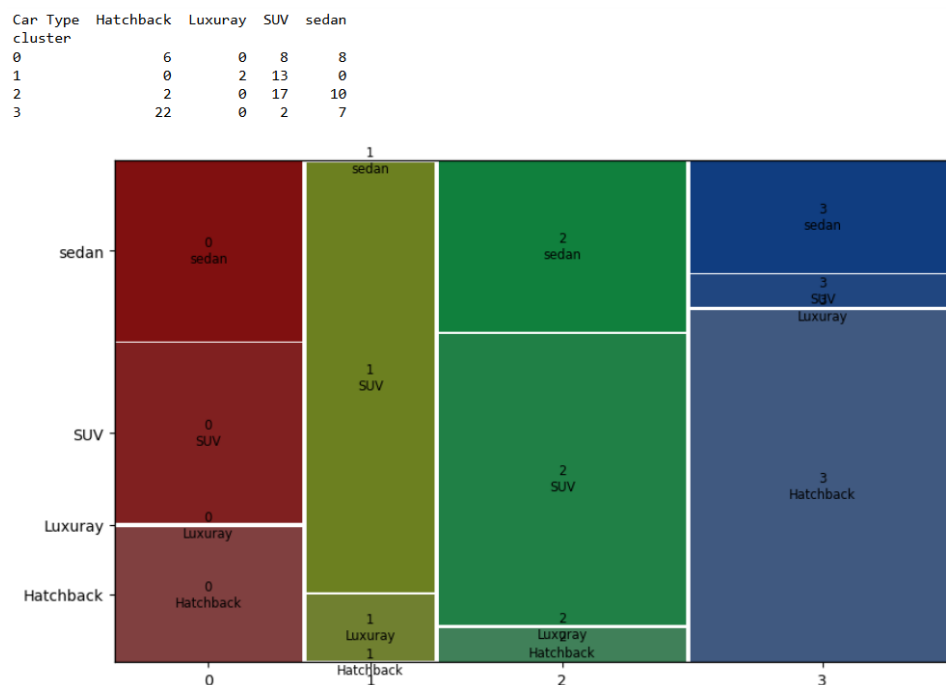
In the above mosaic plot it can be seen that Segment 0 has only Non-Working Wife whereas Segment1 and Segment 2 have only Working Wife. On the other hand Segment 3 has Majority of Non-Working Wife and minority of Working Wife.

### Mosaic plot for Financial Condition



In the above Mosaic plot it can be seen that in Segment 0, Segment 1, Segment 2 the number of people having High Financial Conditions are more whereas in Segment 3 the number of people having Middle Financial Condition are more.

### Mosaic Plot for Car Type



In the above mosaic plot it can be seen that people from Segment 0, Segment 1, Segment 2 are more inclined towards buying an SUV whereas the people in Segment 3 are more inclined towards buying the Hatchback Car Type.



## Selecting Target Segment

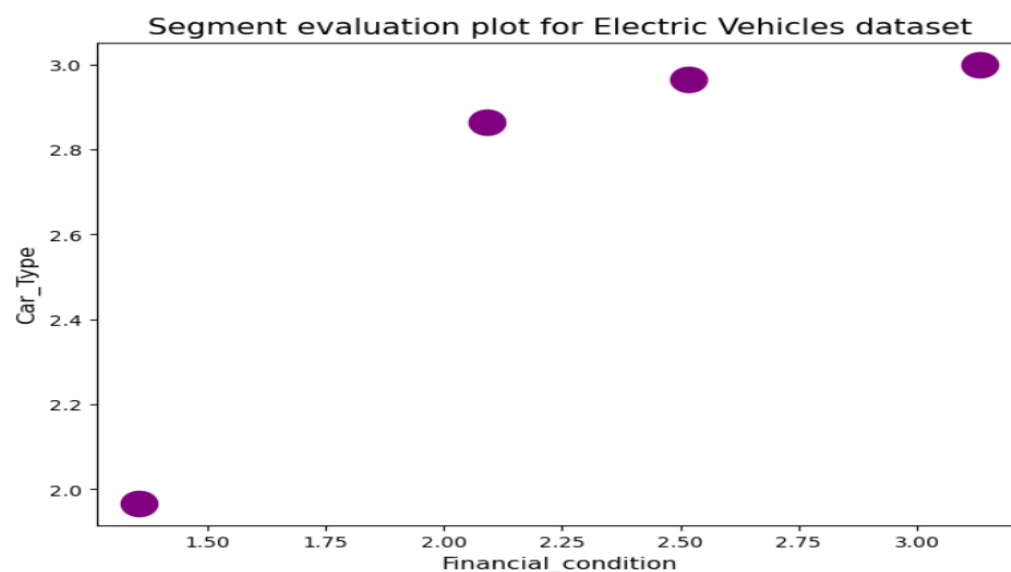
Selecting the target segment of many marketing segments is the most important decision. The selection of one or more target segments is a long-term decision significantly affecting the future performance of an organization. Below we can see a table which compares the values of all of the variables from each market segment.

	cluster	Financial Condition	Wife Salary	Salary	Total Salary	Age	Marrital Status	Profession	Education	No of Dependents	Car Type	Pri
0	0	2.863636	0.000000e+00	2.290909e+06	2.290909e+06	40.545455	0.909091	0.409091	1.409091	2.681818	2.090909	1.136364e+1
1	1	3.000000	1.586667e+06	2.653333e+06	4.240000e+06	42.200000	1.000000	0.200000	1.333333	2.533333	3.133333	1.740000e+1
2	2	2.965517	8.655172e+05	1.517241e+06	2.382759e+06	35.379310	1.000000	0.275862	1.517241	2.275862	2.517241	1.341379e+1
3	3	1.967742	1.032258e+05	1.141935e+06	1.245161e+06	31.612903	0.580645	0.483871	1.419355	1.580645	1.354839	8.838710e+1

From the above table we can infer that our target segment should be Segment 2. The reason for this selection is as follows –

- 1) Consumers in Segment 2 have high Financial Status and hence they can easily afford to buy Electric Vehicles.
- 2) The average age in Segment 2 is close to 35 and hence this segment is more likely to buy EV as compared to others according to various researches which states that men aged between 30-45 years are most likely to buy EVs.
- 3) Also most of the consumers in this segment are salaried people and hence they have a stable income and comparatively more job secure and hence their chances of buying EV will be high.
- 4) Also the preferred Car Type of the Segment 2 is SUV and the average price range of an EV SUV is 14 lakh – 17 lakh which is closer to the price paid by this segment for their conventional SUVs. Hence EV SUVs will be able to easily attract this market segment.

Hence after going through all the analysis conducted above the target segment for an EV startup company should be segment 2. Now the other dilemma is to establish the type of Car to be manufactured to attract this market segment. Hence to analyse that we will plot a Segment Evaluation plot.



In the above segment evaluation plot it is clear that Segment 2 which is having good Financial Condition are preferring to Buy SUVs over other Car Types. Hence it can be concluded that the new EV Startup Company should initially focus on manufacturing EV SUVs in order to attract this particular market segment which will help them to create a strong potential customer base in the early market.

## **Marketing Mix**

For marketing, geometric segmentation plays an important and crucial role.

Geographic segmentation categorizes the target market into segments where marketers can better serve customers in a specific region. This kind of Market segmentation is based on the geographical entities themselves (countries, states, city, etc.), but also depends on various geographical factors such as climate, culture, etc. settings, population, etc. Geographic segmentation includes segmentation Target users based on where they live or work. You can do this any number of times Possibilities: Group customers by country of residence or smaller geographical countries Department from region to city to postal code.

Geographic segmentation is probably the simplest market segmentation Please look back. However, there are still many opportunities that companies will never take advantage of think about. The size of the target area should be changed as needed business. In general, the bigger the company, the bigger the area it belongs to alignment. Ultimately a wider potential audience targeting each zip code individually is not cost effective.

After our analysis and research we came to the conclusion that in India there are 4-5 states such as Uttar Pradesh, Delhi, Maharashtra, Karnataka, Tamil Nadu etc which provide suitable locations for EV startups to start their operations. These states have announced several policies for the growth of EV Vehicles and hence are ideal destinations for any EV Start-up company to start their operations.

## **Potential Customer Base for Business Markets**

There are 87,554 electric vehicles sold during 2017-2018, it increased to 131,554 units in 2018-2019 and again raised in 2019-2020 to 161,308 further due to Covid-19 it became 119,650 units. This number included both two-wheelers, three-wheelers, four wheelers and buses but two wheeler sales have been increasing. A significant portion of the electric vehicles registered in the country are two- and three-wheelers.

- The Electric 2-Wheeler sales have increased by 463.61% from FY 2021 to FY2022. It is increasing rapidly; the Electric 2-Wheeler sale will reach 400,000 units to 450,000 units in FY 2023.
- The Electric 3-Wheeler sale has increased by 212.45% from FY 2021 to FY2022. It is increasing rapidly the Electric 3-Wheeler sale will reach 75,000 units to 80,000 units in FY 2023.
- The Electric 4-Wheeler sales have increased by 344.65% from FY 2021 to FY2022. It is increasing rapidly the Electric 4-Wheeler sale will reach 50,000 units to 55,000 units in FY 2023.

- Karnataka, Tamil Nadu, Maharashtra, Telangana and Rajasthan have combined 67% sales share in high speed 2-wheeler electric vehicles.
- Uttar Pradesh, Bihar, Delhi and Assam have highest percentage of L3 category 3- wheelers electric vehicles sales.
- Telangana, Karnataka and Delhi have highest number of high speed L5 category electric 3-wheeler vehicles sales.
- Maharashtra has highest 4-wheeler EV sales in 2021 (26%) followed by Delhi and Telangana(13% each).
- Maharashtra and Gujarat have highest e-Bus sales accounting for 47% and 18% units sold respectively.

### **Potential Customer Base \* Your Target Price Range = Potential Profit.**

- For Electric 2-Wheeler, The per unit average price will be 1,00,000 and the number of units sold will be around 4,25,000.

Potential Profit in India =  $425000 * 100000 = \text{Rs } 42.5 \text{ Billion}$

Potential Profit in Karnataka =  $35,000 * 1,00,000 = \text{Rs } 3.5 \text{ Billion}$  .

Potential profit in Maharashtra and Tamil Nadu =  $30,000 * 1,00,000 = \text{Rs } 3 \text{ Billion}$  each in both the states.

- For Electric 3-Wheeler, The per unit Average price will be 2,00,000 and the number of units sold will be around 150000.

Potential Profit in India =  $150000 * 200000 = \text{Rs } 30 \text{ billion}$

Potential profit in Uttar Pradesh =  $65000 * 200000 = \text{Rs } 13 \text{ Billion}$ .

- For Electric 4-Wheeler, The per unit Average price will be 15,00,000 and the number of units sold will be around 50000.

Potential Profit =  $50000 * 15,00,000 = \text{Rs } 75 \text{ billion}$ .

Potential Profit for Electric 4-wheelers in Maharashtra =  $5000 * 15,00,000 = \text{Rs } 7.5 \text{ Billion}$

Potential Profit for Electric 4-wheelers in Delhi and Telangana =  $2500 * 15,00,000 = \text{Rs } 3.75 \text{ Billion}$  each in both these states.

### **MOST OPTIMAL MARKET SEGMENTS**

From the above report, we conclude that to create an Electric Vehicle startup in India, the most optimal market segment for us will be based on Geographic and Demographic segments which would be the most amount of EVS sold in particular states and the type of electric vehicle respectively.

After analysing the EV market using Market Segmentation Analysis, the feasible strategy that we have come up with is that we will be focussing on the states that have more demand for EVs like Uttar Pradesh, Delhi and Karnataka, Maharashtra. Also, one more reason to set up the startup in these 3 states is that the infrastructure required for the EVs including the

charging station is available which would ease the burden in setting up of the startup process and also the governments are providing a lot of subsidies to the people as well as for the company in order to encourage electric vehicles.

Also the EV start-up company should focussing on developing EV 2-Wheelers and SUVs in EV 4-wheelers which will help them get a good potential customer base according to our research and analysis and will help the company to grow in the early market.