**A Study In Market Segmentation**

**What are the future prospects of Electric Vehicles in India ?**

The future prospects of the Electric Vehicle (EV) market in India look promising. With the increasing concern for environmental pollution and the need to reduce greenhouse gas emissions, the Indian government has been promoting the adoption of EVs by implementing various policies and incentives. Additionally, the growing awareness among consumers about the benefits of EVs, such as lower fuel costs and reduced maintenance expenses, has further increased the demand for electric vehicles in the country. This has resulted in a surge of investments in the EV industry in India, with both domestic and international players entering the market. The future looks bright for the EV industry in India, and it is expected to witness significant growth in the coming years.

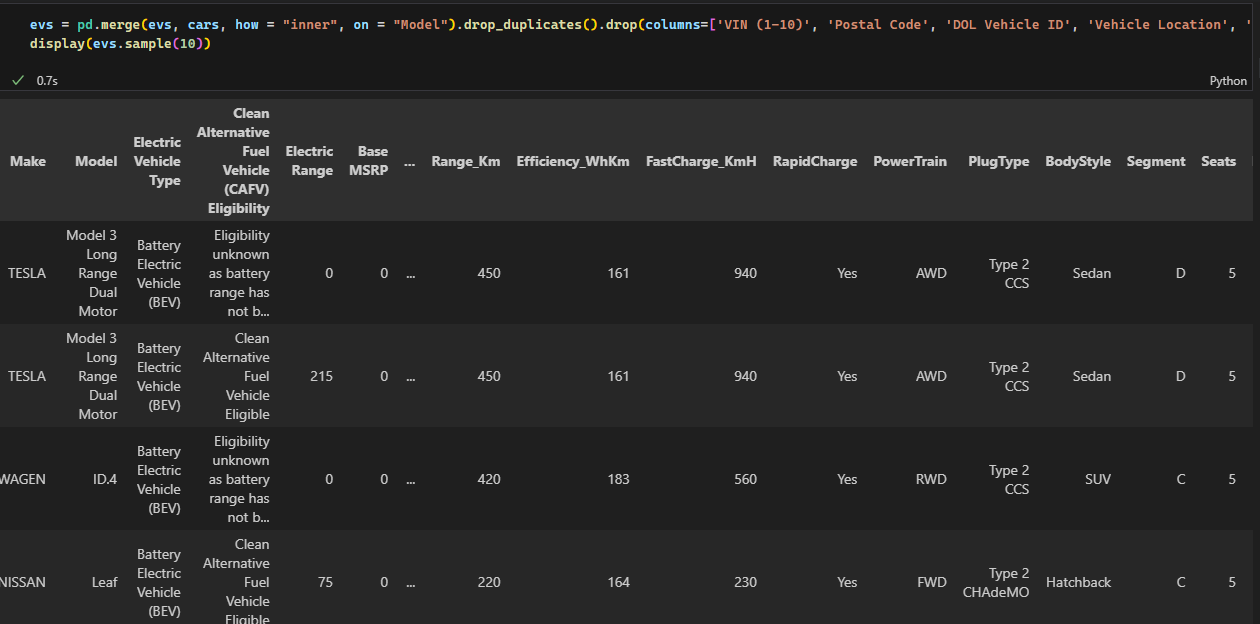
**Bsc Data Science**

**Pratyush Rao, Intern at Feynn Labs**

The dataset used for the analysis is a combination of two different datasets

* <https://www.kaggle.com/datasets/kkhandekar/cheapest-electric-cars>
* <https://www.kaggle.com/datasets/ratikkakkar/electric-vehicle-population-data>

The data was compiled an uploaded combined to provide with a dataset of 93k records which was then used for market segmentation and understand attributes of Electric vehicles on a broader scale and people’s biases towards it.



Finally compiled dataset image

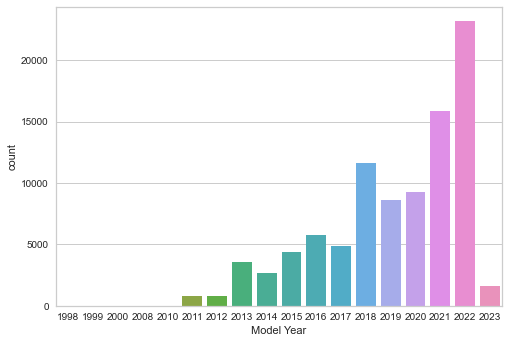
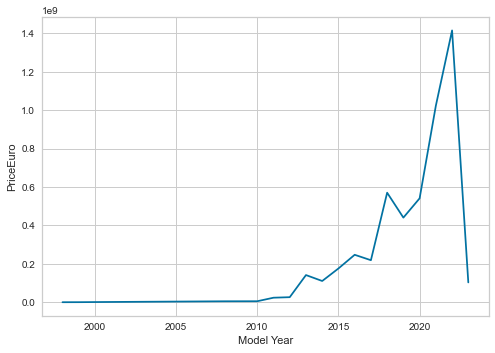
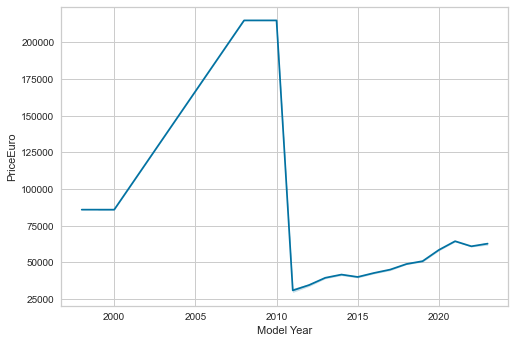


Image of sales increasing over the years

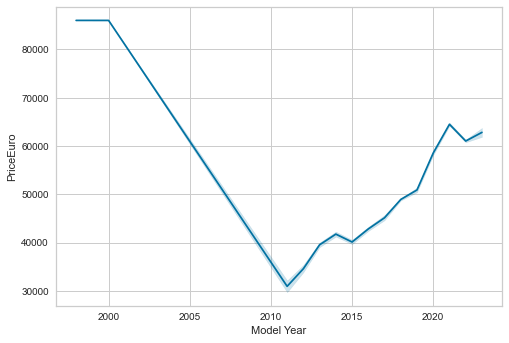
It was observed the total sales of EV’s in the past few years had increased exponentially, and although the dataset does not originate in India we can assume that we will have similar prospects as the prices have only gone cheaper as you can see in the image below.



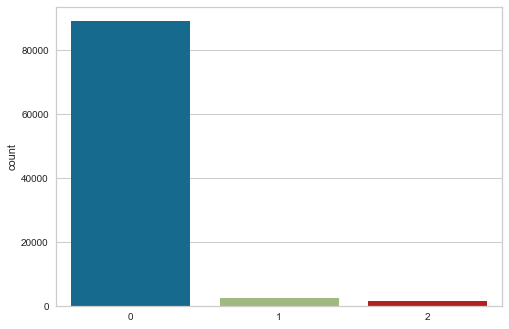
\* Only one model called the TESLA Roadster was sold for 215000 Euros in the years from 2008-2010

\* If the roadster model with the 47 models is ignored we should notice a downward trend from 2000-2010 and spiking trend from the years 2010 to 2020 because of the increase in sales of electric cars.

\* The trend shows the prices have dropped drastically

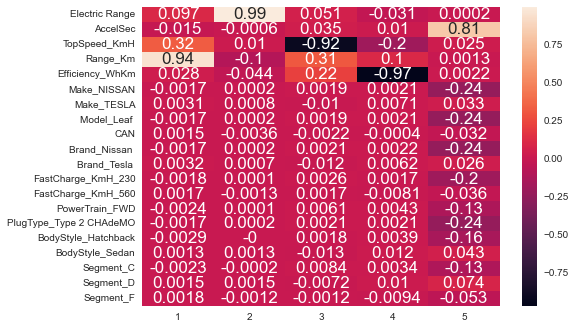


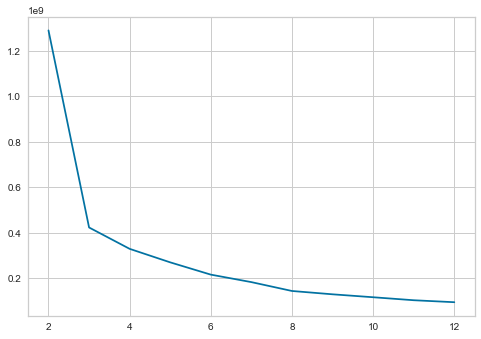
Using feature selection methods we were able to select 20 best features that provide optimal clustering. However we noticed that the dataset contained a lot of outliers because of which the clustering algorithm was isolating them resulting in poor quality clusters.



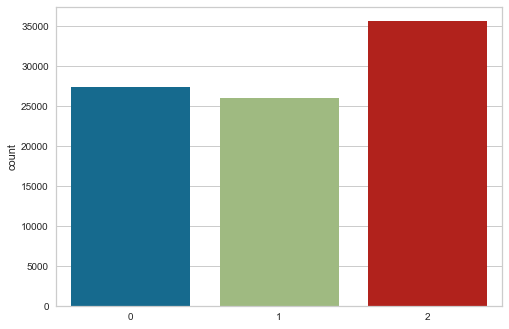
Poor quality clusters

To rectify this we dropped the clusters and then restarted the whole process to obtain optimum clusters on K selected best features. Following is the plot of contribution of each column in the dataset to the derived compressed components of PCA.

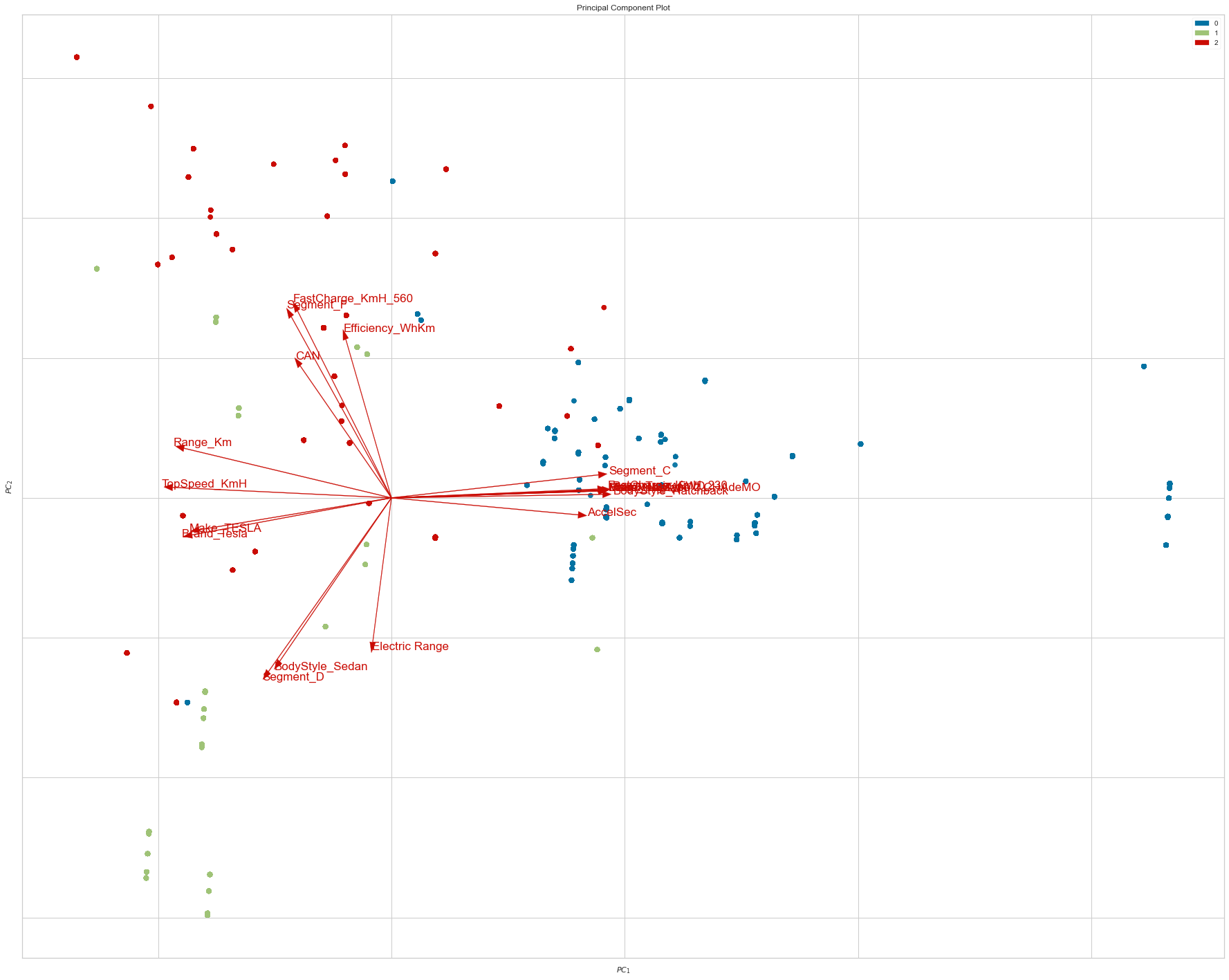




Using the elbow method we identified the optimum number of clusters was 2.



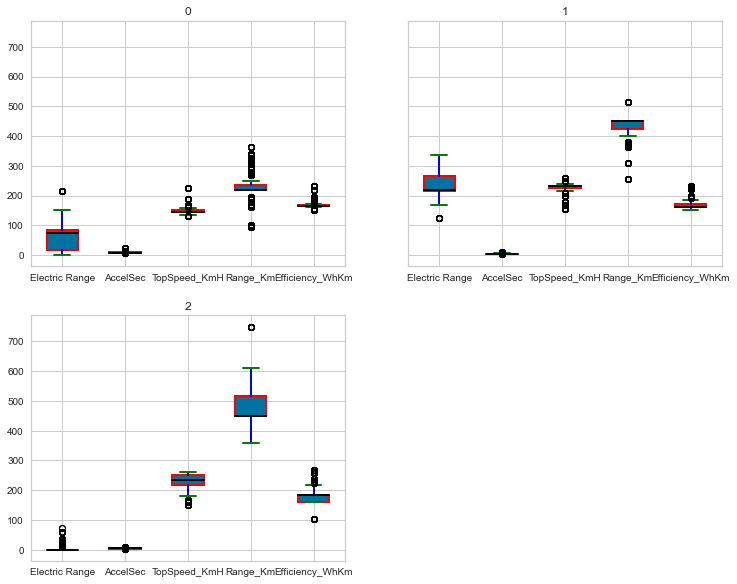
And since outliers are eliminated the clusters were very distinct as well. The next step of course was describing the segments.



\* You can notice that certain parameters point in certain cluster directions.

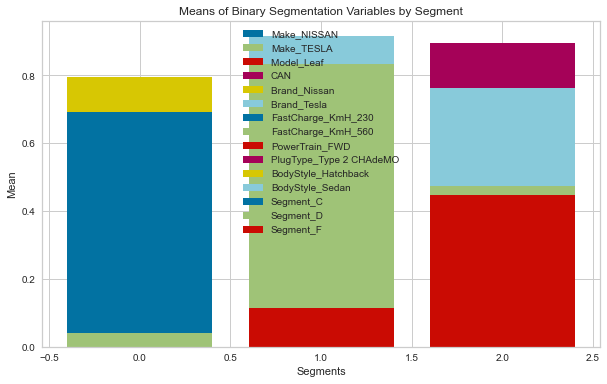
\* This information is useful because this means that different columns have different weightage on determining the clusters.

\* Now its time we analyze our clusters and identify the defining factors.



\* The distribution of the numerical values seems almost equivalent in all the columns except for Electric ranges.

\* Electric range from segment two has the lowest values, segment the second most lowest and from segment 1 the highest.



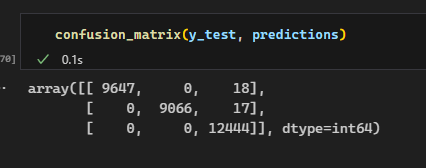
\* Segment 0 consists of mostly nissan users of Segment\_D column

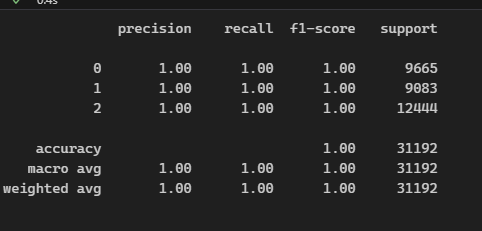
\* Segment 1 consists of Fast\_Charge\_KmH\_560 showing that people prefer fast cars in segment 1. It also consists of Segment\_F column.

\* Segment 2 can be mostly classified as customers preferring Segment\_F, and CAN, prefer sedan bodystyled tesla.

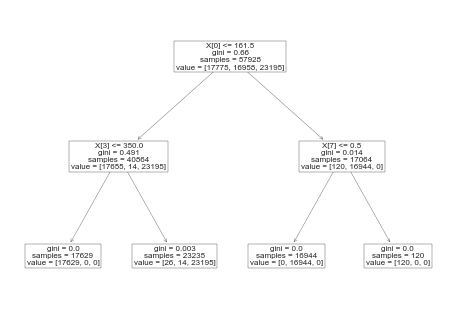
\* You can also always train a descision tree to identify segments.

The results of the decision tree were absolutely remarkable with a confusion matrix and classification report looking like this.





This is an indication that the clustering was well performed. Finally the tree that will help us classify looks like this.



You can mine this tree to identify segments and get desirable results.