MARKET SEGMENTATION ANALYSIS OF ELECTRIC VEHICLES IN INDIA

Aparajita Goswami

25/10/2023

Problem Statement:

For this task, we have to do market segmentation on the Electric Vehicles Market in India 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 traits.

Introduction

Electric vehicles (EVs) are vehicles powered by electric motors. They can use external electricity sources or batteries for propulsion. EVs, including cars, scooters, and three-wheelers, are gaining popularity due to environmental concerns. In India, the EV industry is in its early stages, accounting for less than 1% of vehicle sales. However, it has the potential to grow to more than 5% with government incentives. Most Indian EVs are low-speed scooters with lead batteries, facing battery life issues. Charging infrastructure is being developed.

Datasets used:

1. https://drive.google.com/file/d/1J1eoyfV1KHxdWqgSaTY5MfNFeH4IqvFB/view?usp=drive_link
2. https://drive.google.com/file/d/1SZVWhRMN8MjLp3M4u471AzDwYl2TZhWX/view?usp=drive_link

DATASET 1:

Machine Learning Model(Algorithm) Used:

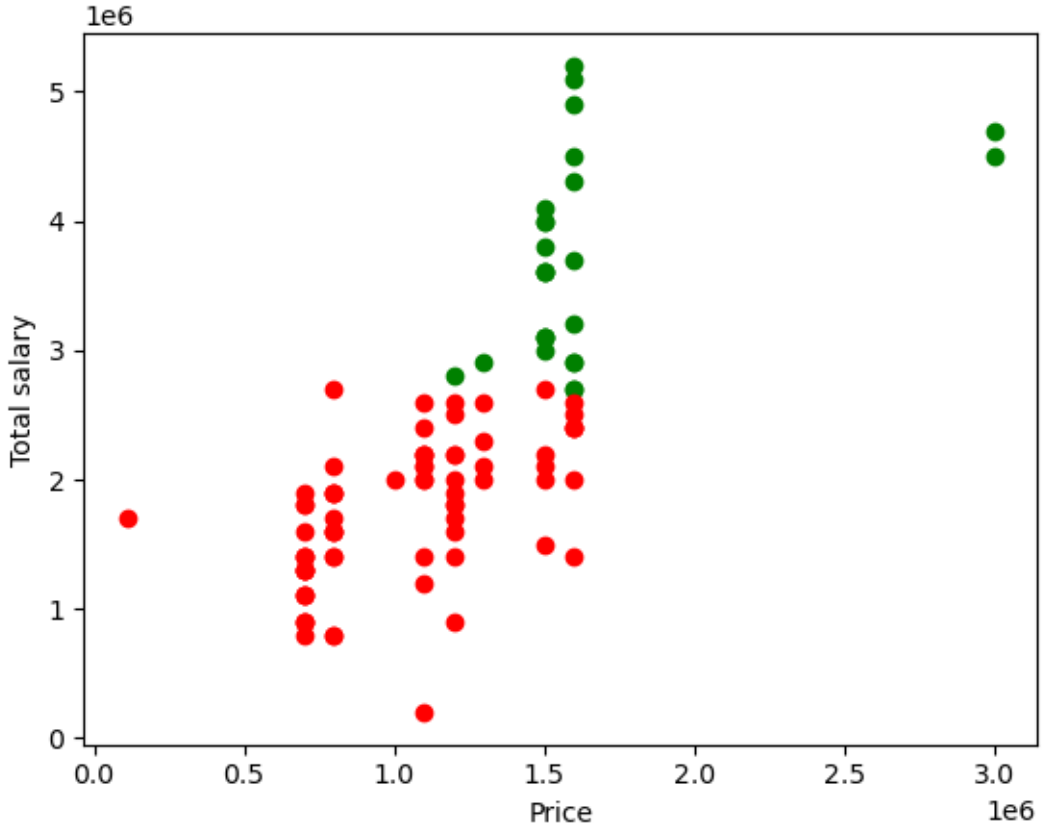
1. K-Means Clustering to form clusters based on the price of vehicles and user salary

Observations/Conclusions:

1. We found that 2 clusters groups can be formed from the data given(based on price of the vehicle) using the K-Means algorithm.

CODE:

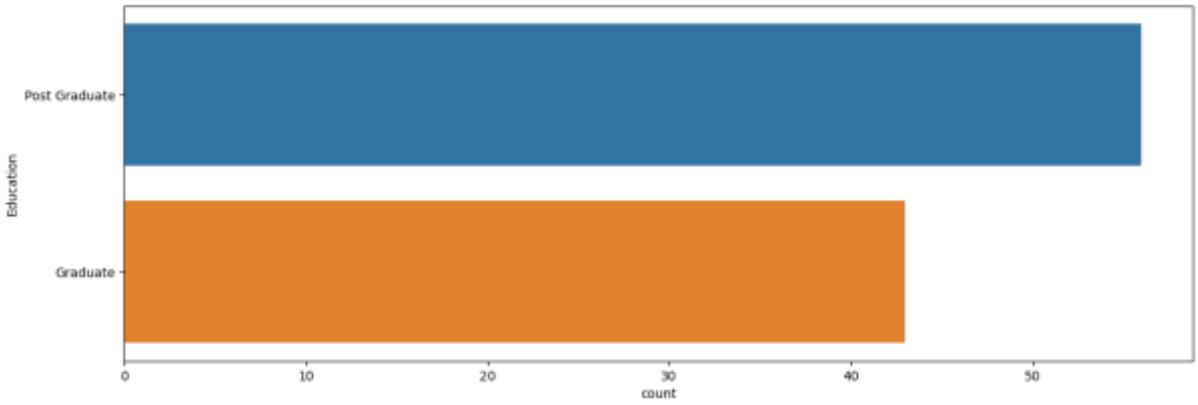
```
d1 = d[d.clusters==1]
d0 = d[d.clusters==0]
plt.scatter(d1.Price, d1['Total Salary'],color='green')
plt.scatter(d0.Price, d0['Total Salary'],color='red')
plt.xlabel('Price')
plt.ylabel('Total salary')
```



2. While looking at the patterns, we find that as the level of education increases the chances of buying an electric vehicle also rises.

CODE:

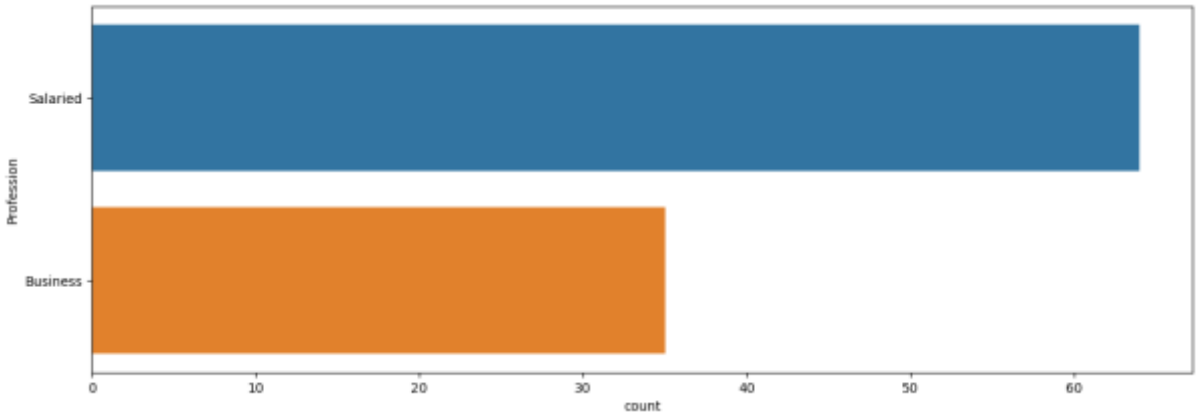
```
plt.figure(figsize=(15,5))
sns.countplot(y='Education',data=data)
plt.show()
```



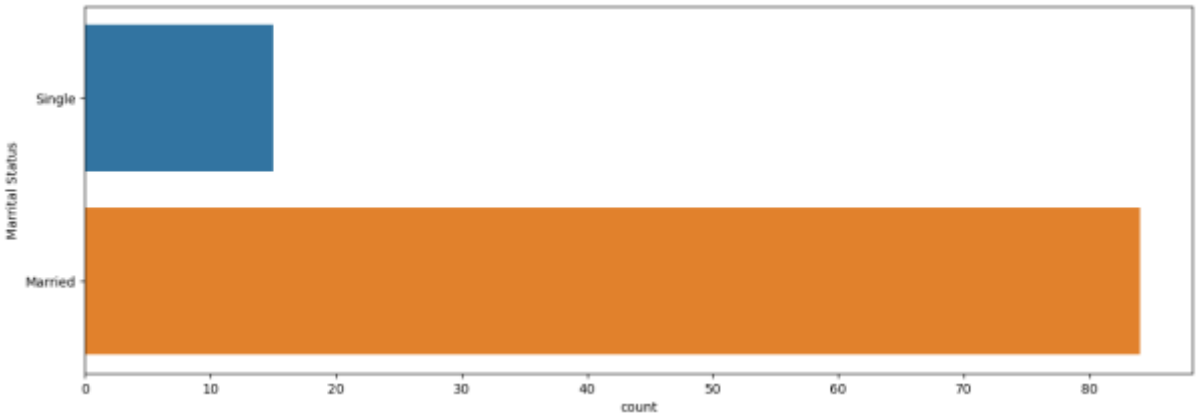
3. Also, amount spent on the car goes up for salaried people.

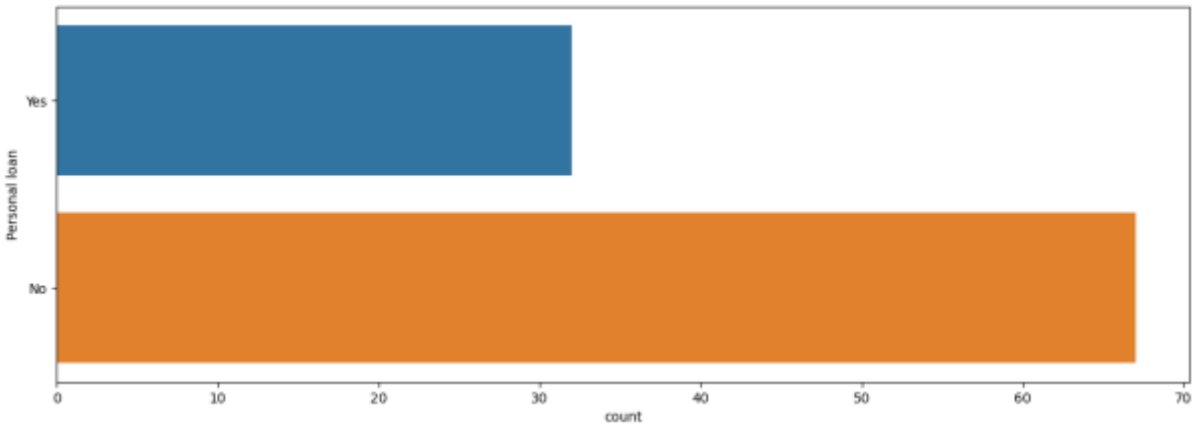
   CODE:

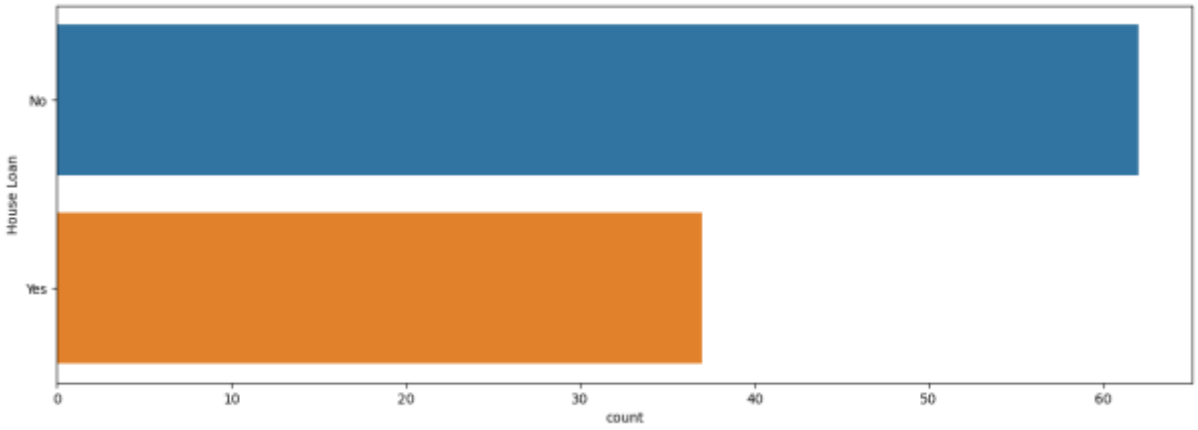
```
plt.figure(figsize=(15,5))
sns.countplot(y='Profession',data=data)
plt.show()
```



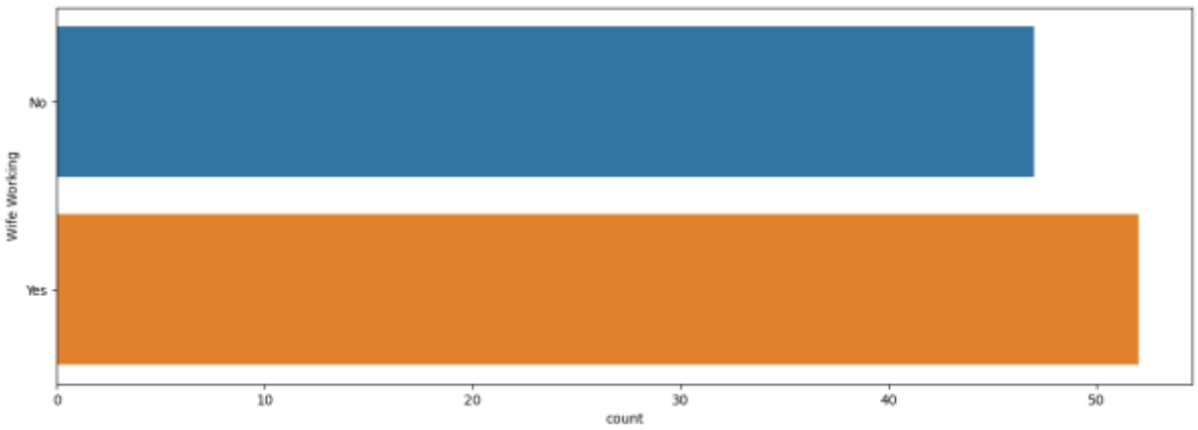
4. Similarly, we can see that a person is more likely to buy an electric car, if he/she is married, with no personal or house loans.



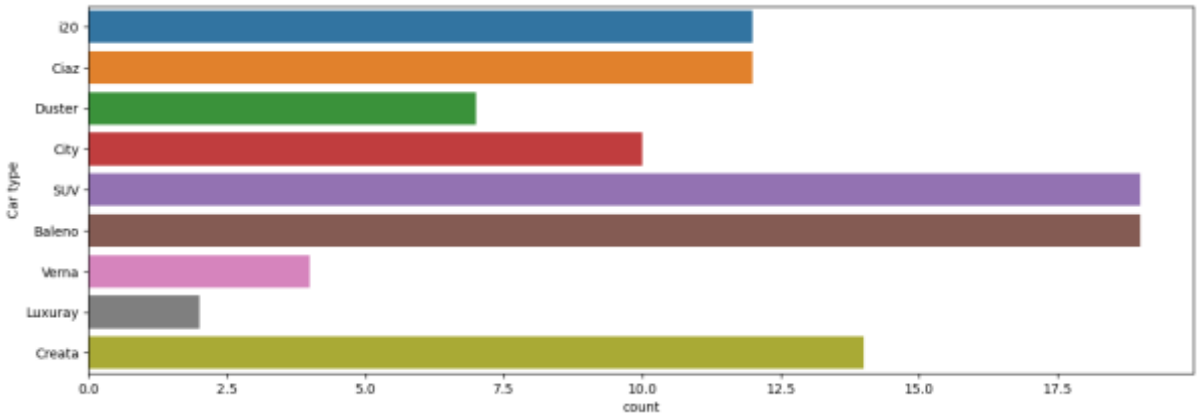




5. The visualizations suggest that couples where both people are working, are slightly more likely to buy an electric vehicle.



6. Also, the following plot gives an idea of the electric vehicle makes that are mostly bought, i.e., the ones that are the most popular among the customers.



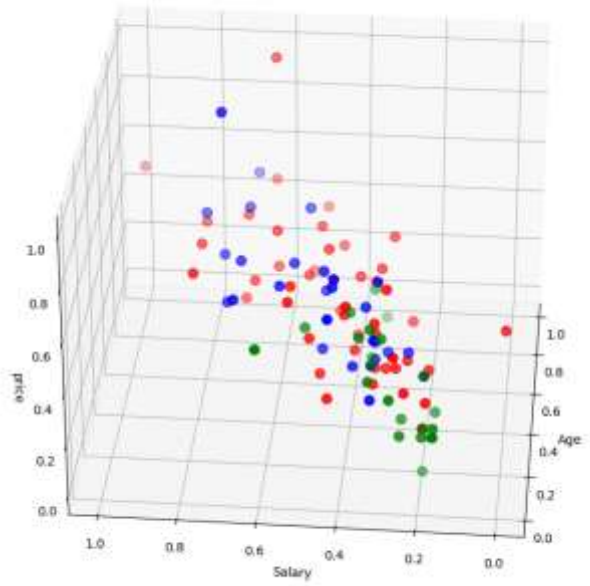
7. 3-Dimensional visualization - Age, Salary, Price:

CODE:

```
fig = plt.figure(figsize=(20,10))

ax = fig.add_subplot(111, projection='3d')

ax.scatter(d.Age[labels == 0], d["Salary"][labels == 0],
d["Price"][labels == 0],  c='blue', s=60)
ax.scatter(d.Age[labels == 1], d["Salary"][labels == 1],
d["Price"][labels == 1],  c='red', s=60)
ax.scatter(d.Age[labels == 2], d["Salary"][labels == 2],
d["Price"][labels == 2],  c='green', s=60)
ax.view_init(30, 185)
plt.xlabel("Age")
plt.ylabel("Salary")
ax.set_zlabel('price')
plt.show()
```



Scope for improvements in this project

The project/research can be further improved, if some more variables or factors can be added to the dataset. For e.g., if the dataset had a column for year of purchase, we could have analysed that to find out which year had the most sales, also we could analyse, which make was popular or was selling best each year. Also, if the data had an added column which detailed the locations, based on the purchasing showrooms, where electric cars were more in use, it would be useful, as that would enable us to analyse those locations to understand what factors facilitated those cars to be in more use in those locations.

Top 4 Variables/Features Which Can Be Used To Create Most Optimal Market Segments For This Market Domain

In this dataset, the top 4 variables/features that can be used to create the most optimal market segments are:

1. Salary type and salary amount of individual.
2. Age and marital status of individual.
3. If the individual has taken any personal/home loans.
4. The price and make of the electric car.

DATASET 2:

Machine Learning Model(Algorithm) Used:

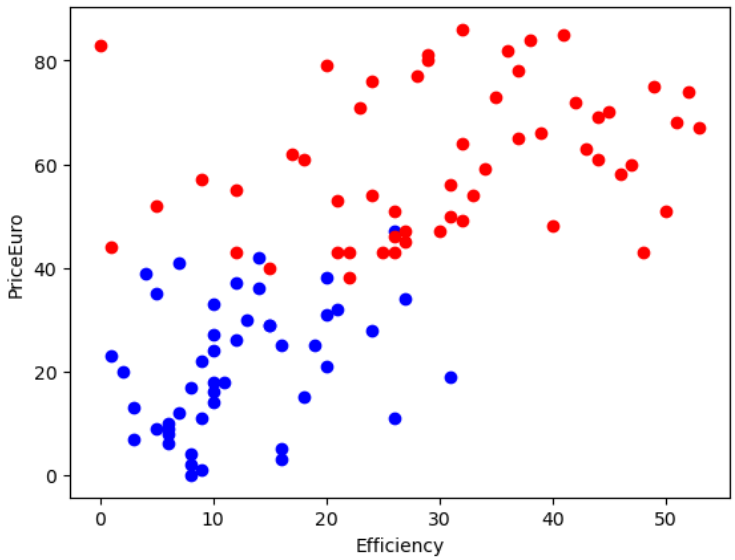
1. K-Means Clustering

Observations/Conclusions:

1. We found that 2 clusters groups can be formed from the data given(based on price of the vehicle and the efficiency) using the K-Means algorithm.

   CODE:

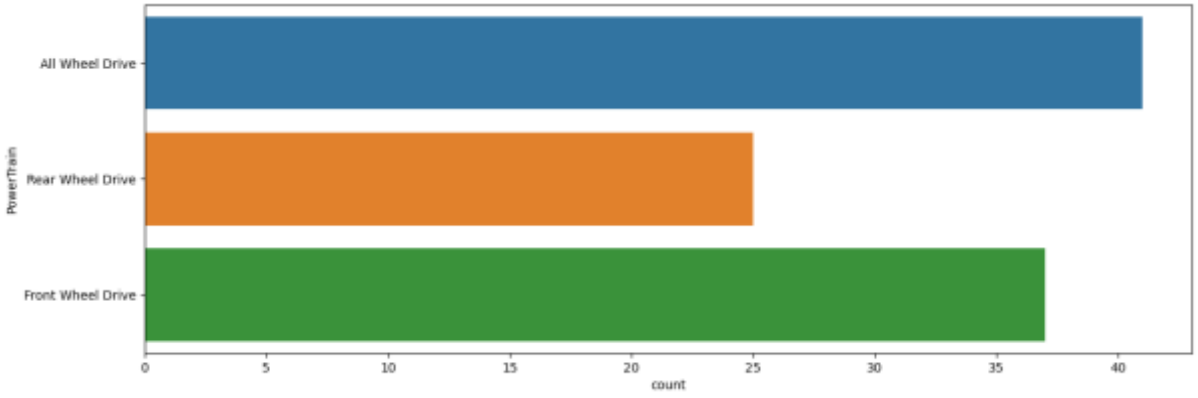
```
d1 = d[d.clusters==1]
d0 = d[d.clusters==0]
plt.scatter(d1.Efficiency, d1['PriceEuro'],color='blue')
plt.scatter(d0.Efficiency, d0['PriceEuro'],color='red')
plt.xlabel('Efficiency')
plt.ylabel('PriceEuro')
```



2. Analysing the given car data, we can also see the most used power train in cars, presently. All-Wheel Drive is the most in use currently, followed by Front-Wheel Drive and Rear-Wheel Drive.

CODE:

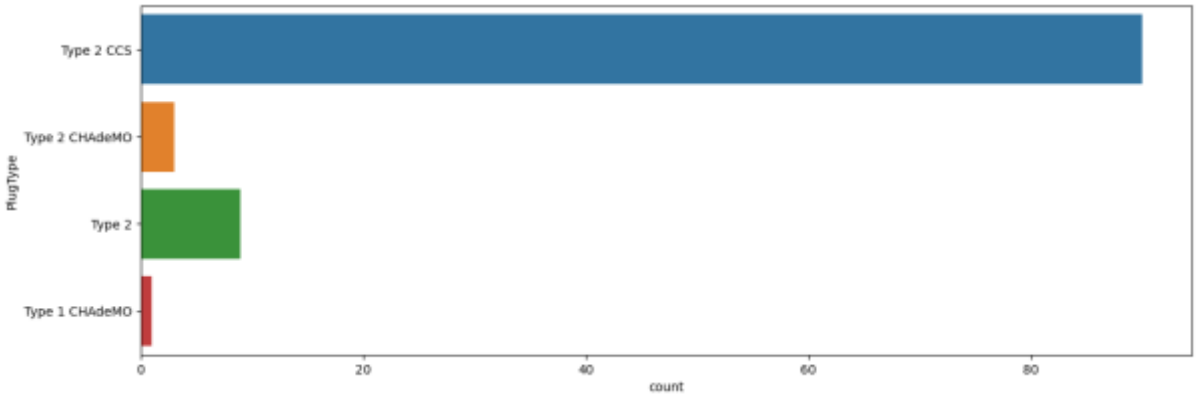
```
plt.figure(figsize=(15,5))
sns.countplot(y='PowerTrain',data=data)
plt.show()
```



3. Also, we can see that at present, the count of electric cars is significantly related to the Plug type. Cars with Type 2 CCS plug are the most in use, while cars with Type 1CHAdemo plugs are the least in use.

CODE:

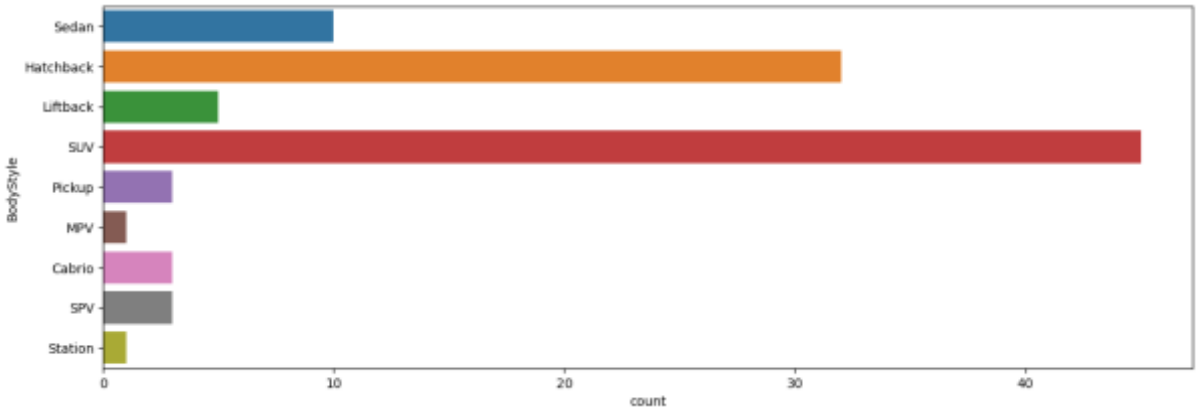
```
plt.figure(figsize=(15,5))
sns.countplot(y='PlugType',data=data)
plt.show()
```



4. The visual representation of the count of electric cars with respect to the body styles is given below:

CODE:

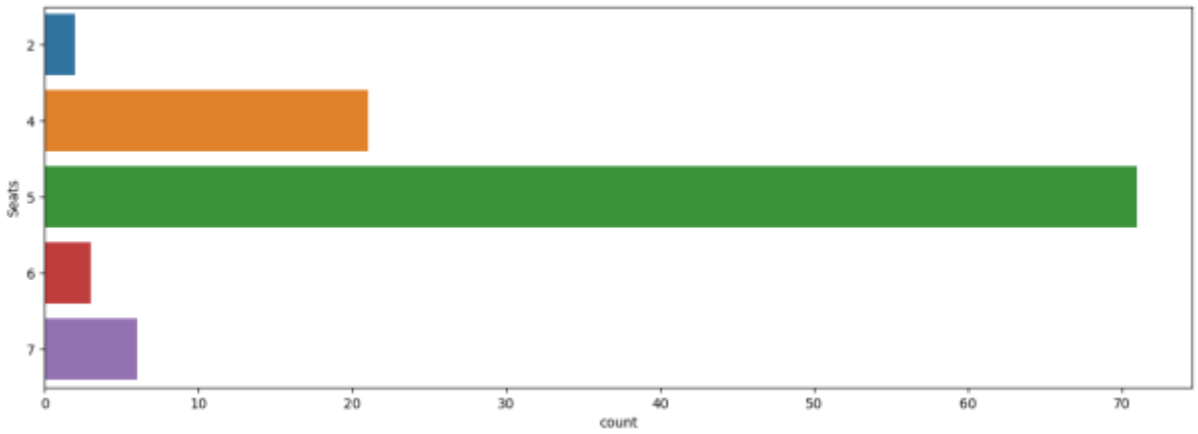
```
plt.figure(figsize=(15,5))
sns.countplot(y='BodyStyle',data=data)
plt.show()
```



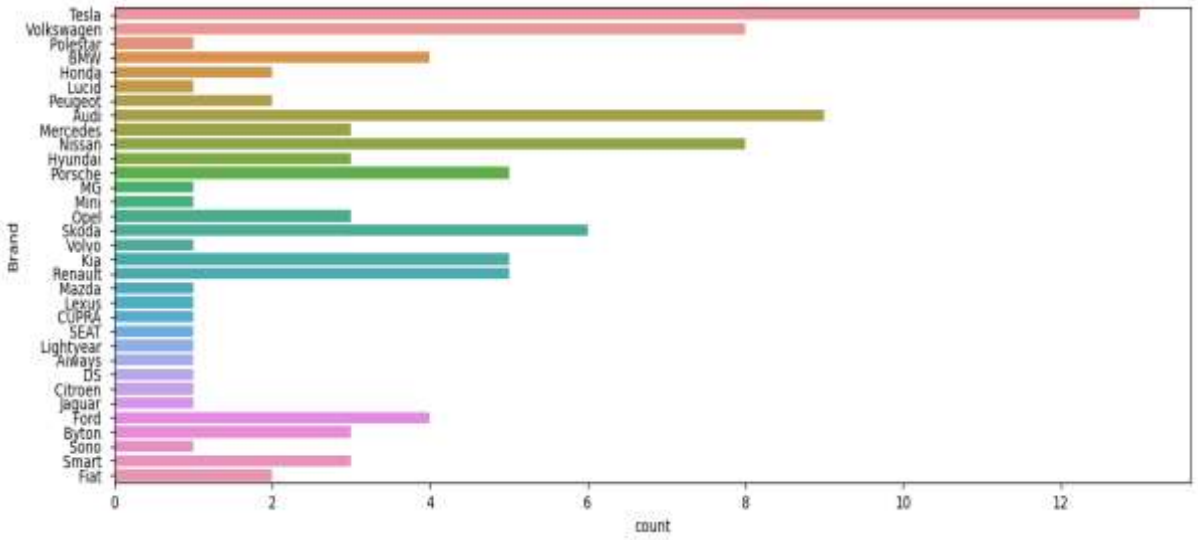
5. Another observation is that, of the electric cars sold, 5-seater cars were the maximum in count, while 2 seaters were the least in count.

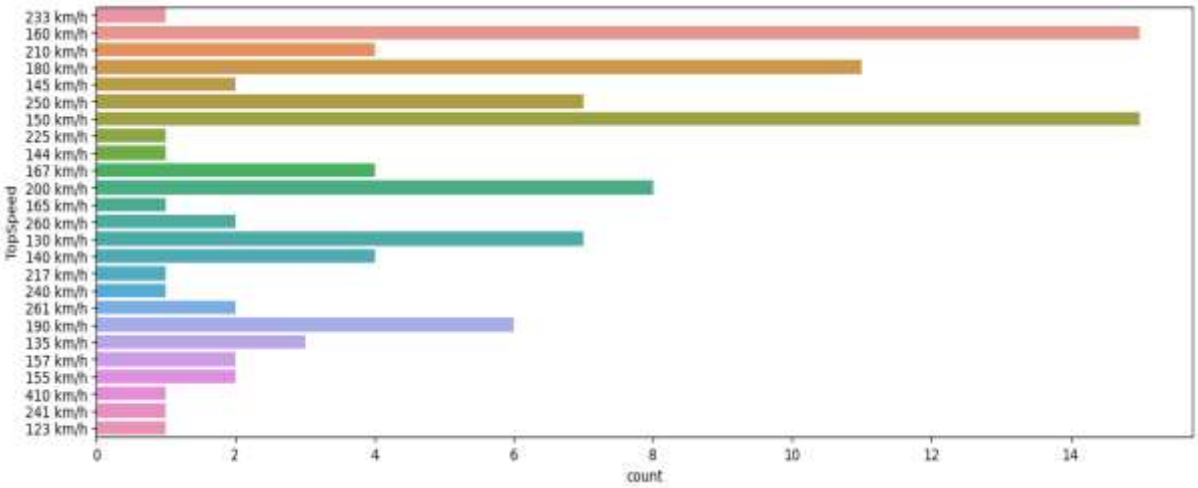
CODE:

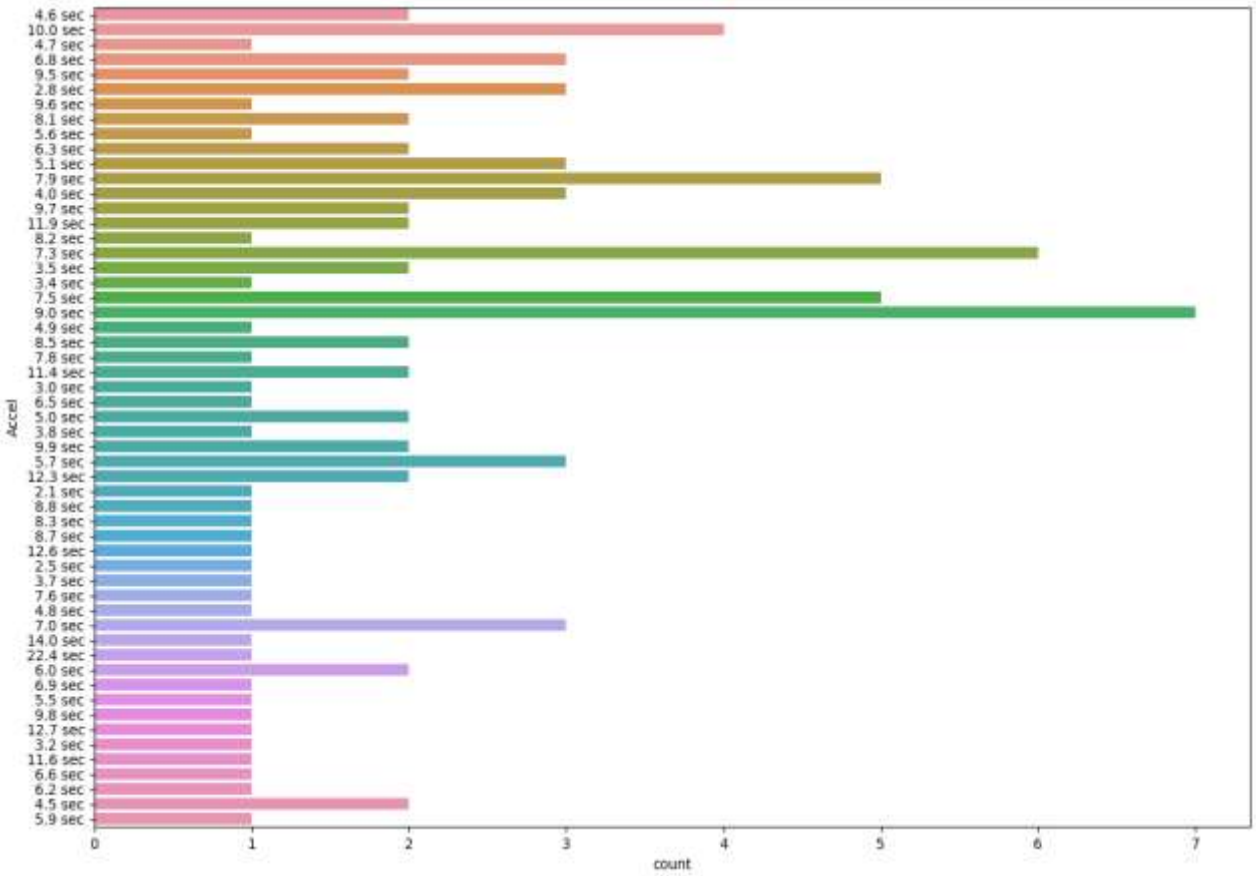
```
plt.figure(figsize=(15,5))
sns.countplot(y='Seats',data=data)
plt.show()
```

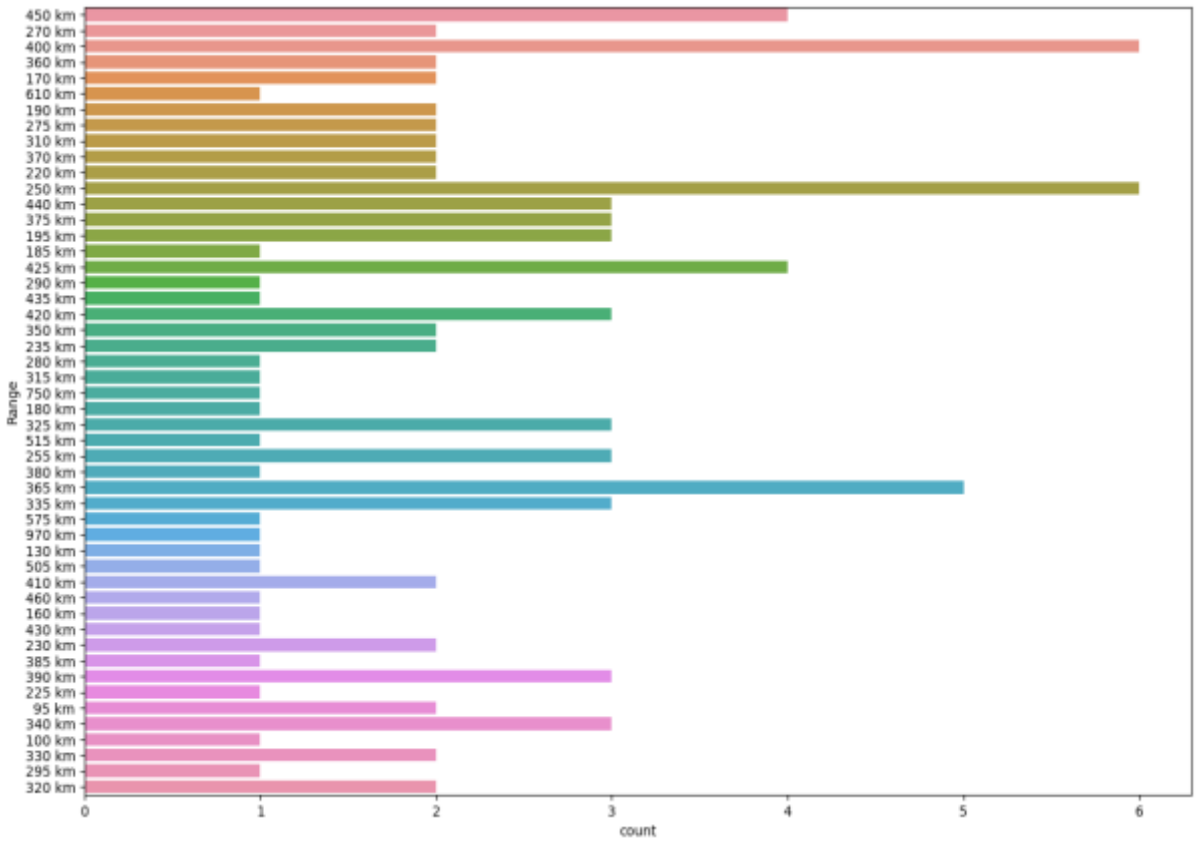


6. Some other graphs related to the count of electric cars, given in the dataset are as follows:

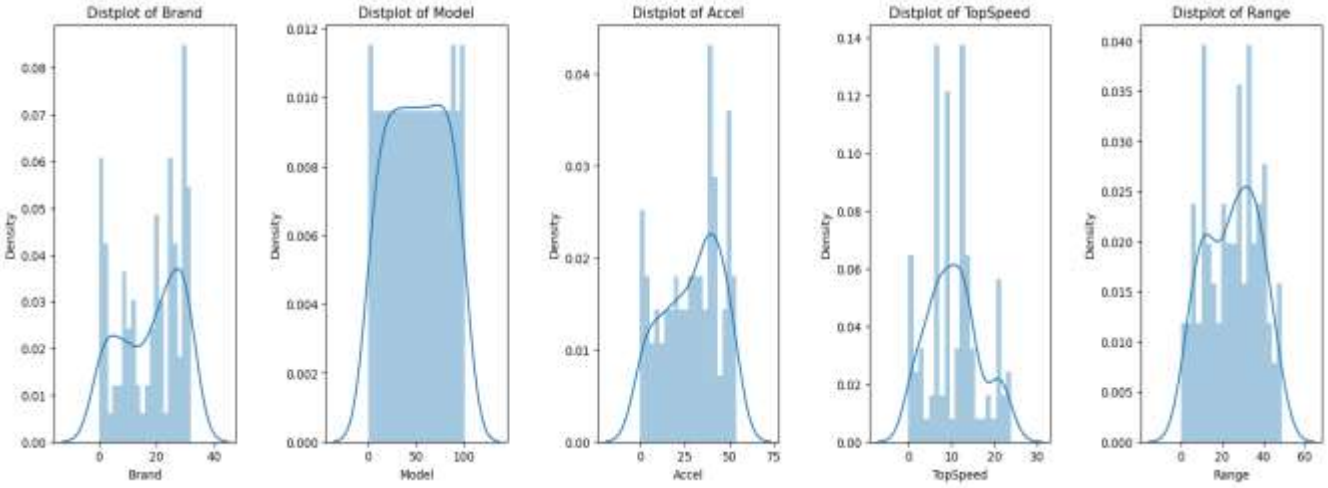


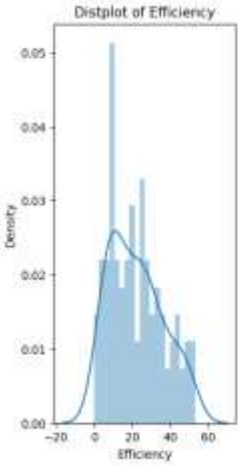


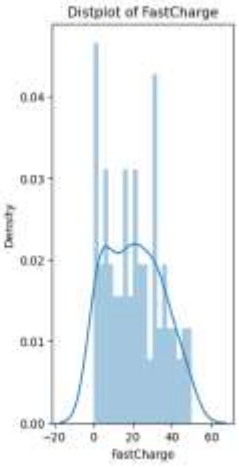

Accel
4.6 sec
10.0 sec
4.7 sec
6.8 sec
9.5 sec
2.8 sec
9.6 sec
8.1 sec
5.6 sec
6.3 sec
5.1 sec
7.9 sec
4.0 sec
9.7 sec
11.9 sec
8.2 sec
7.3 sec
3.5 sec
3.4 sec
7.5 sec
9.0 sec
4.9 sec
8.5 sec
7.8 sec
11.4 sec
3.0 sec
6.5 sec
5.0 sec
3.8 sec
9.9 sec
5.7 sec
12.3 sec
2.1 sec
8.8 sec
8.3 sec
8.7 sec
12.6 sec
2.5 sec
3.7 sec
7.6 sec
4.8 sec
7.0 sec
14.0 sec
22.4 sec
6.0 sec
6.9 sec
5.5 sec
9.8 sec
12.7 sec
3.2 sec
11.6 sec
6.6 sec
6.2 sec
4.5 sec
5.9 sec
0
1
2
3
4
5
6
7
count

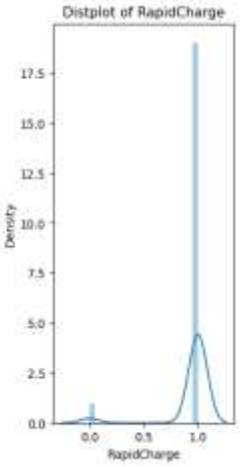

Range
450 km
270 km
400 km
360 km
170 km
610 km
190 km
275 km
310 km
370 km
220 km
250 km
440 km
375 km
195 km
185 km
425 km
290 km
435 km
420 km
350 km
235 km
280 km
315 km
750 km
180 km
325 km
515 km
255 km
380 km
365 km
335 km
575 km
970 km
130 km
505 km
410 km
460 km
160 km
430 km
230 km
385 km
390 km
225 km
95 km
340 km
100 km
330 km
295 km
320 km
0
1
2
3
4
5
6
count

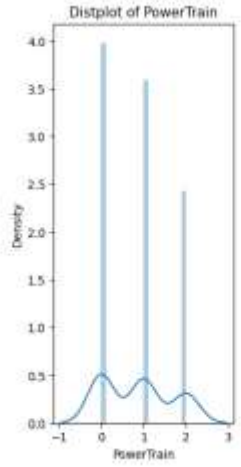
7. Displots of the various attributes of the electric cars given in dataset:

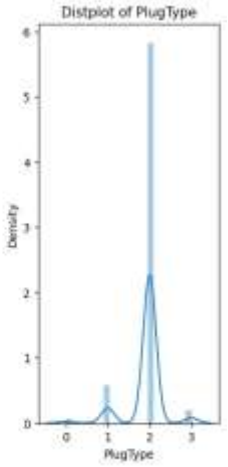


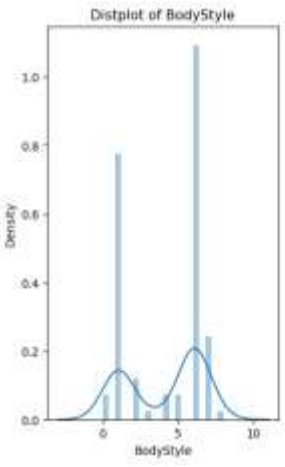


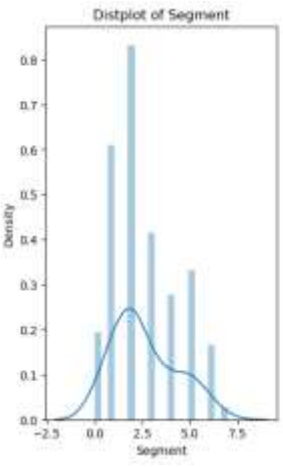


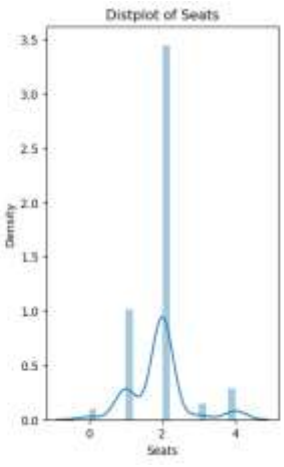


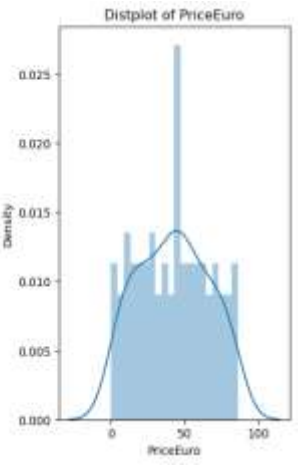












Scope for improvements in this project

The project/research can be further improved, if some more variables or factors can be added to the dataset. This dataset primarily focuses on the attributes or features of the electric cars. There is no customer purchase details, that might help in understanding the consumer market better. For e.g., if the dataset had a column for the salary of the purchasers, we could analyse that with respect to the selling price of the cars, to group the customers. Also, if the data had an added column which detailed the locations, based on the purchasing showrooms, where each of the cars were more in use, it would be useful, as that would enable us to analyse those locations to understand what factors facilitated those cars to be more in use in those locations.

Top 4 Variables/Features Which Can Be Used To Create Most Optimal Market Segments For This Market Domain

In this dataset, the top 4 variables/features that can be used to create the most optimal market segments are:

1. Efficiency of the electric vehicle
2. Price of the electric vehicle
3. Range and Top Speed of the electric vehicle
4. Fast Charge of the electric vehicle

Conclusion

In this market segmentation analysis for the Electric Vehicles (EVs) market in India, we utilized two datasets and applied K-Means clustering to identify distinct market segments based on various features. The goal was to understand customer behaviour and preferences to devise a feasible market entry strategy targeting the most promising segments.