EV Market Segmentation

Report by :
Mohamed Bayas
Jyotsna Doonga,
Deepak Vishwakarma,
Nilesh Kumar,
Ramkhelavan Mandal,
Arnav Sharma,
Kushi Kiranraj

Abstract:

Electric vehicles (EV), as a promising way to reduce the greenhouse effect, have been researched extensively. In this study, the operating process of the various types of EVs will be explained. Battery technology and super capacitor technology will also be discussed as a possibility to increase the energy capacity of PHEV. The E vehicle is an emerging concept and the automobile industry is conducting extensive research to make the option feasible and commercially viable. There are already some first movers like Tesla who have successfully developed their model and moving forward. This paper analyses the strategy and leadership of Tesla which have enabled them to be take the first mover advantage. A descriptive research based on secondary data had been employed in this study.

Problem Statement:

To analyse the EV Market in India using Segmentation analysis and come up with a feasible strategy to enter the market, targeting the segments most likely to use Electric vehicles (EVs).

Context (Reason for growth of EV Market):

EVs are gaining attention across the globe as they help reduce emissions and depletion of natural resources. The Indian EV market is also evolving fast as close to 0.32 million vehicles were sold in 2021, up 168% YoY. Ongoing electric vehicle adoption in India is based on the Paris agreement to reduce carbon emissions, to improve the air quality in urban areas and reduce oil imports. Main reasons for the growth of EV Market are:

- Stricter Emission Regulations
- Consumer Acceptance
- Lower Battery Costs
- More Charging Stations
- Cost of petrol/diesel are higher
- Beneficial in the long run

Fermi Estimation (Breakdown of Problem Statement)

India’s Population = 130 Crore
Number of States and UTs in India = 28 + 8

The number of registered vehicles across India was 295 million in 2019.
Vehicle registrations compound annual growth rate = 10% (2007-19)

The electric vehicle market in India is expected to increase at a compounded annual growth rate of 49% between 2021 to 2030 while the EV segment's volume may cross annual sales of 17 million by 2030. Then we can assume an 8% Vehicle registrations compound annual growth rate after 2030. Also we will target the car segment in the Electric Vehicles Specifically.

Hence, it will be stable in the long run because it is renewable, environment friendly and will be economical in the long run than fossil fuels. We will see the state wise sale of vehicles and also see which company sells the most cars and generates more revenue including price range, safety and customer satisfaction. Then we will move to the KM range, durability, safety, charging time, customer satisfaction, battery etc.

Resources for datasets used:

- https://data.gov.in/
- https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset
- https://data.gov.in/resource/all-india-level-composition-vehicle-population-during-2015-2016

Preprocessing and EDA:

Firstly importing the necessary libraries and going through the data

```
[1] import numpy as np
    import pandas as pd
    import matplotlib.pyplot as plt
    import seaborn as sns
```

```
[2] df=pd.read_csv("ElectricCarData_Clean.csv")
```

```
[3] df
```

	Brand	Model	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro
0	Tesla	Model 3 Long Range Dual Motor	4.6	233	450	161	940	Yes	AWD	Type 2 CCS	Sedan	D	5	55480
1	Volkswagen	ID.3 Pure	10.0	160	270	167	250	Yes	RWD	Type 2 CCS	Hatchback	C	5	30000
2	Polestar	2	4.7	210	400	181	620	Yes	AWD	Type 2 CCS	Liftback	D	5	56440
3	BMW	iX3	6.8	180	360	206	560	Yes	RWD	Type 2 CCS	SUV	D	5	68040
4	Honda	e	9.5	145	170	168	190	Yes	RWD	Type 2 CCS	Hatchback	B	4	32997
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
98	Nissan	Ariya 63kWh	7.5	160	330	191	440	Yes	FWD	Type 2 CCS	Hatchback	C	5	45000
99	Audi	e-tron S Sportback 55 quattro	4.5	210	335	258	540	Yes	AWD	Type 2 CCS	SUV	E	5	96050
100	Nissan	Ariya e-4ORCE 63kWh	5.9	200	325	194	440	Yes	AWD	Type 2 CCS	Hatchback	C	5	50000
101	Nissan	Ariya e-4ORCE 87kWh Performance	5.1	200	375	232	450	Yes	AWD	Type 2 CCS	Hatchback	C	5	65000
102	Byton	M-Byte 95 kWh 2WD	7.5	190	400	238	480	Yes	AWD	Type 2 CCS	SUV	E	5	62000

103 rows × 14 columns

Changing the price to INR and few other changes

```
df['INR'] = df['PriceEuro']*79.81
df['RapidCharge'].replace(to_replace=['No','Yes'],value=[0, 1],inplace=True)
df.head()
```

	Brand	Model	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro	INR
0	Tesla	Model 3 Long Range Dual Motor	4.6	233	450	161	940	1	AWD	Type 2 CCS	Sedan	D	5	55480	4427858.80
1	Volkswagen	ID.3 Pure	10.0	160	270	167	250	1	RWD	Type 2 CCS	Hatchback	C	5	30000	2394300.00
2	Polestar	2	4.7	210	400	181	620	1	AWD	Type 2 CCS	Liftback	D	5	56440	4504476.40
3	BMW	iX3	6.8	180	360	206	560	1	RWD	Type 2 CCS	SUV	D	5	68040	5430272.40
4	Honda	e	9.5	145	170	168	190	1	RWD	Type 2 CCS	Hatchback	B	4	32997	2633490.57

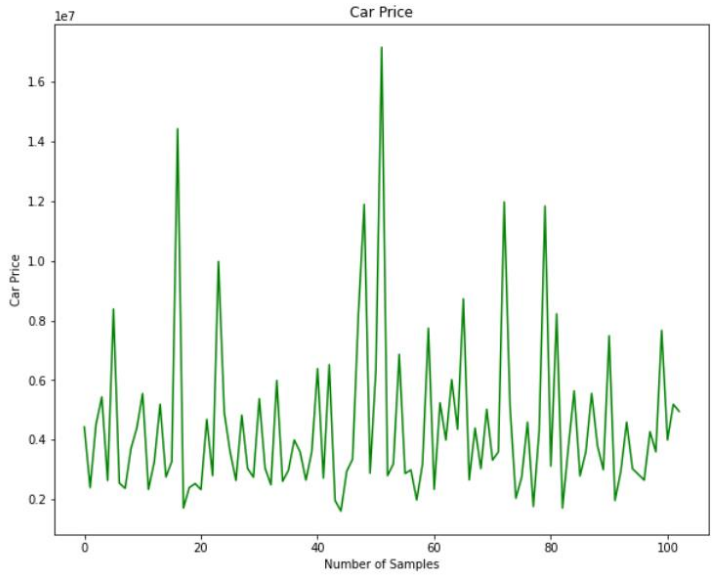
Again looking at the data and looking for any null values

```
df.describe()
```

	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	RapidCharge	Seats	PriceEuro	INR
count	103.000000	103.000000	103.000000	103.000000	103.000000	103.000000	103.000000	1.030000e+02
mean	7.396117	179.194175	338.786408	189.165049	0.951456	4.883495	55811.563107	4.454321e+06
std	3.017430	43.573030	126.014444	29.566839	0.215963	0.795834	34134.665280	2.724288e+06
min	2.100000	123.000000	95.000000	104.000000	0.000000	2.000000	20129.000000	1.606495e+06
25%	5.100000	150.000000	250.000000	168.000000	1.000000	5.000000	34429.500000	2.747818e+06
50%	7.300000	160.000000	340.000000	180.000000	1.000000	5.000000	45000.000000	3.591450e+06
75%	9.000000	200.000000	400.000000	203.000000	1.000000	5.000000	65000.000000	5.187650e+06
max	22.400000	410.000000	970.000000	273.000000	1.000000	7.000000	215000.000000	1.715915e+07

```
df['INR'].plot(figsize = (10,8),title='Car Price',xlabel = 'Number of Samples',ylabel = 'Car Price',color = 'green')
```

```
<matplotlib.axes._subplots.AxesSubplot at 0x7f16c5d62f90>
```

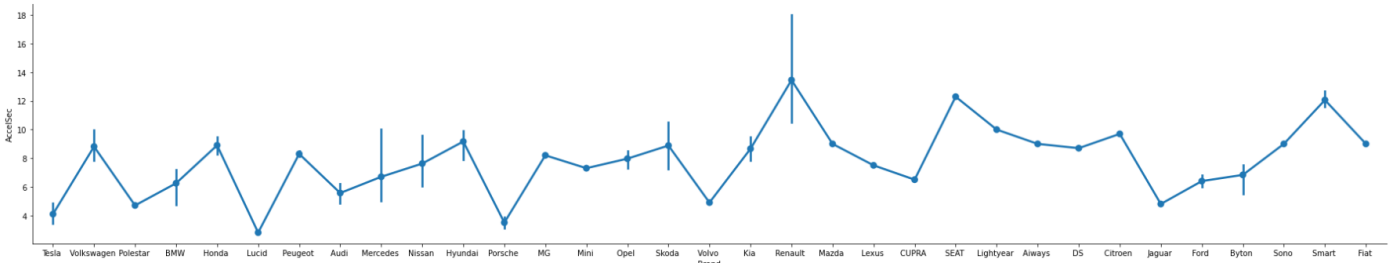


Overview of acceleration and speed of cars

Acceleration

```
sns.factorplot(x="Brand",y="AccelSec",data=df,aspect=40/8)
```

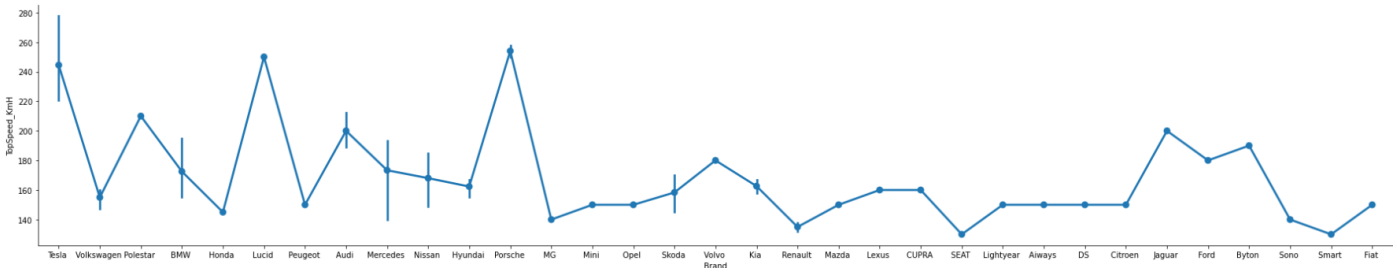
```
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:3717: UserWarning: The `factorplot` function has been renamed to `catplot`. The original name will be removed in a future release.
  warnings.warn(msg)
<seaborn.axisgrid.FacetGrid at 0x7f16c57d4f90>
```

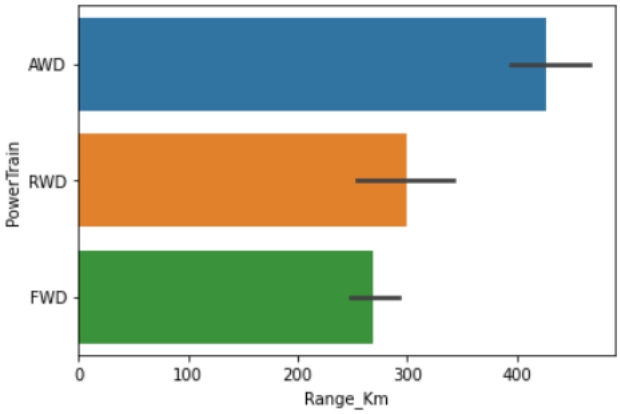


Speed

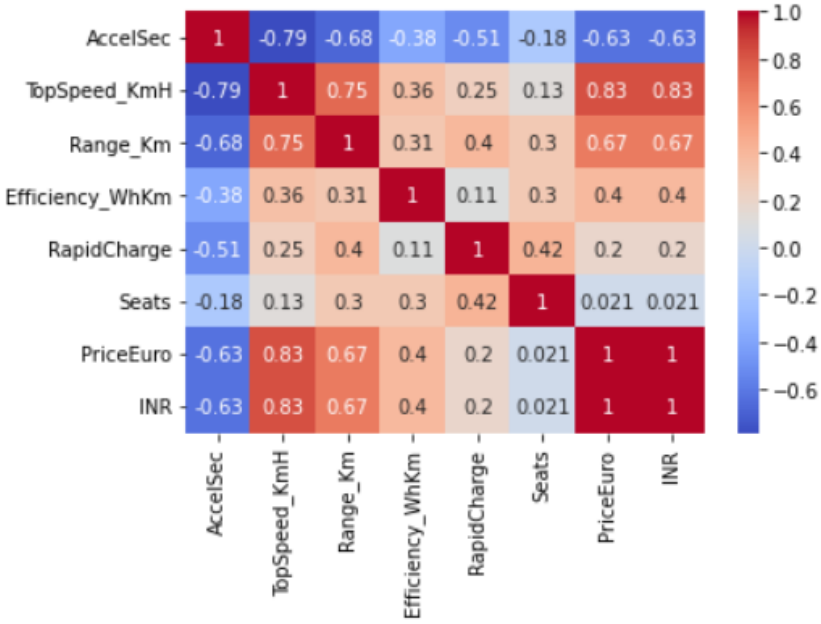
```
sns.factorplot(x="Brand",y="TopSpeed_KmH",data=df,aspect=40/8)
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:3717: UserWarning: The `factorplot` function has been renamed to `catplot`. The original name will be removed in a future release.
  warnings.warn(msg)
<seaborn.axisgrid.FacetGrid at 0x7f16c57dd9d0>
```





Now looking at the heat map of the data



After this we checked various factors that are interconnected like Range vs Brand etc, like
Type of Plug used for charging and found that Most companies use Type 2 CCS and the Type 1 CHAdeMo are used the least.
Cars and their body style most cars are eiher SUV or Hatchback.

Extracting Segments :

Selecting Target Segments :

After segment extraction, the next step is to select the target segment.
This is can be done by clustering. This is the sub-step of the previous one.
After selecting different segments now it's time to select a target segment.
The Target Segment is one that really solves our problem statement.
It tells us which segment would be preferable to increase the revenue or profit of the product. This is done based on the concept of Clustering. The customers of the same characteristics are clustered together. Here we use **k-Means** clustering as the machine learning algorithm.

- Using Visualisations to understand cluster formation

The following figure shows the clustering of two descriptor variable that are *Efficiency_WhKm* and *PriceEuro*.

Fig 1:

There are 3 colors depicting 3 clusters:

- **Cluster 1**
- Cluster 2
- Cluster 3

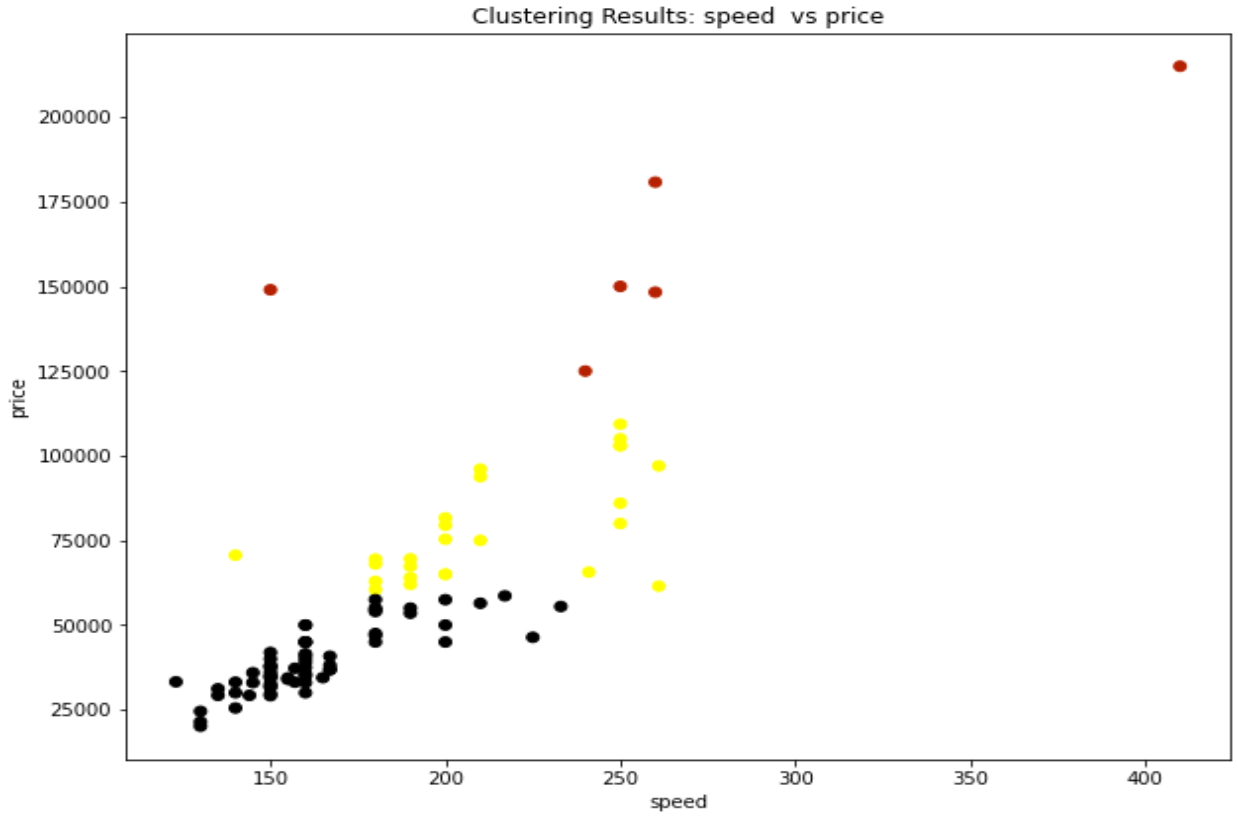
Observations:

- ★ **Cluster 1** shows medium efficiency of the car and low price of the car.
- ★ Cluster 2 shows higher efficiency of the car and medium price of the car.
- ★ Cluster 3 shows medium efficiency of the car and higher price of the car.

Let's check clustering in another descriptor variable with *PriceEuro*. Because we should check different segments and only conclude the better one.

The following figure shows the clustering of two descriptor variable that are *ToSpeedkmH* and *PriceEuro*

Fig 2:



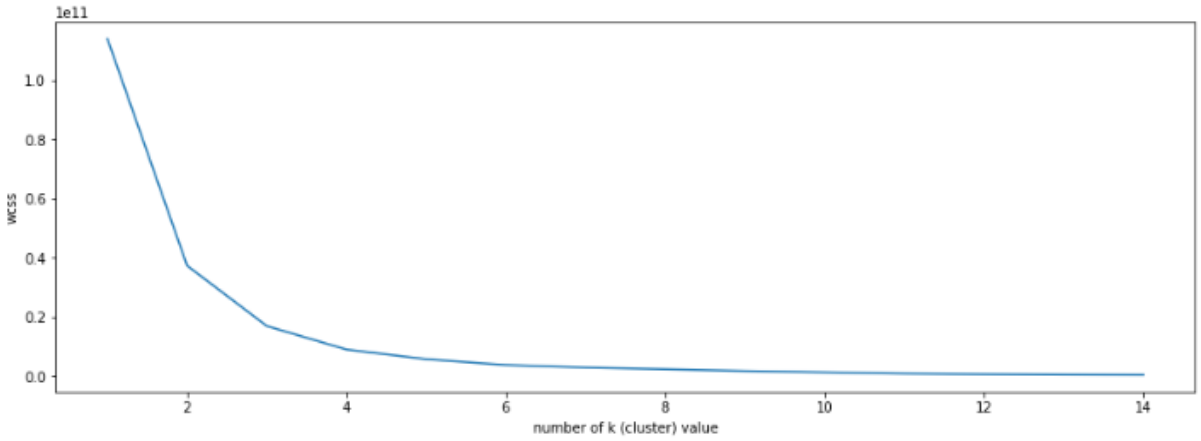
Observations:

- ★ **Cluster 1** shows the lower speed of the car and lower price of the car.
- ★ Cluster 2 shows medium speed of the car and the medium price of the car.
- ★ Cluster 3 shows medium speed of the car and higher price of the car.
- ● The electric vehicle car more depends on efficiency hence we will consider the fig1.

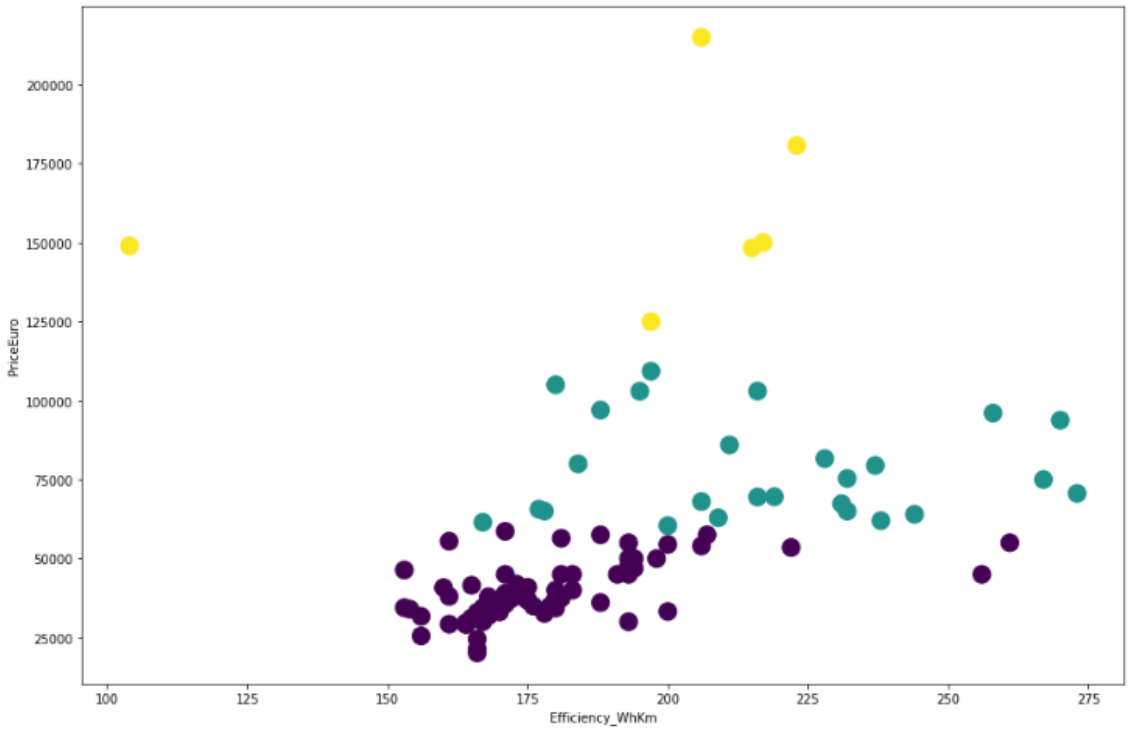
Conclusion : We conclude that our target segment would be **cluster 2** because it is showing cars of higher efficiency and medium price. That sounds great and will attract more customers because customers want cars with higher efficiency and which costs low.

Clustering Code:

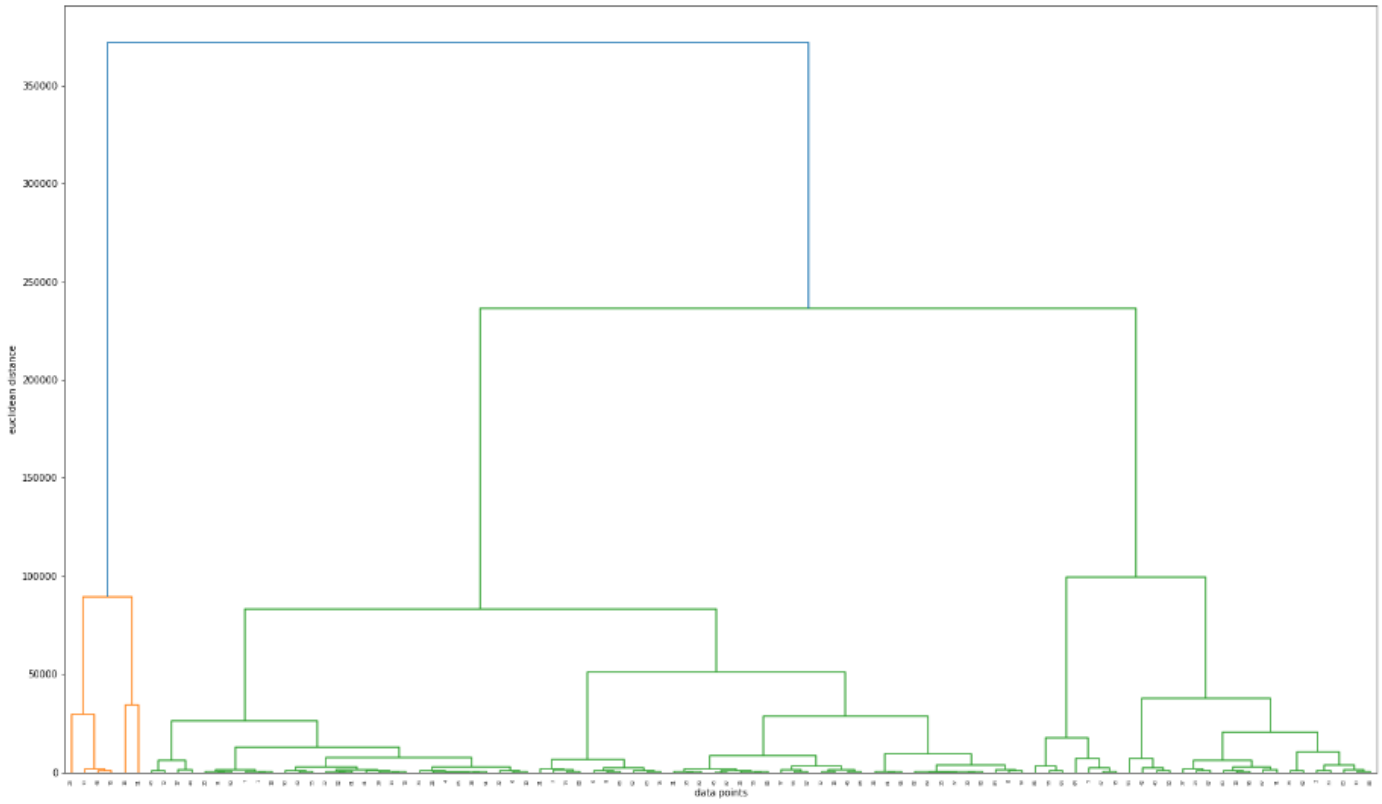
```
wcss = []
data_model = df.drop(['Brand','Model','RapidCharge'],axis=1)
for k in range(1,15):
    kmeans = KMeans(n_clusters=k)
    kmeans.fit(data_model)
    wcss.append(kmeans.inertia_)

# the best value is elbow value. It's 3.
plt.figure(figsize=(15,5))
plt.plot(range(1,15),wcss)
plt.xlabel("number of k (cluster) value")
plt.ylabel("wcss")
plt.show()
```



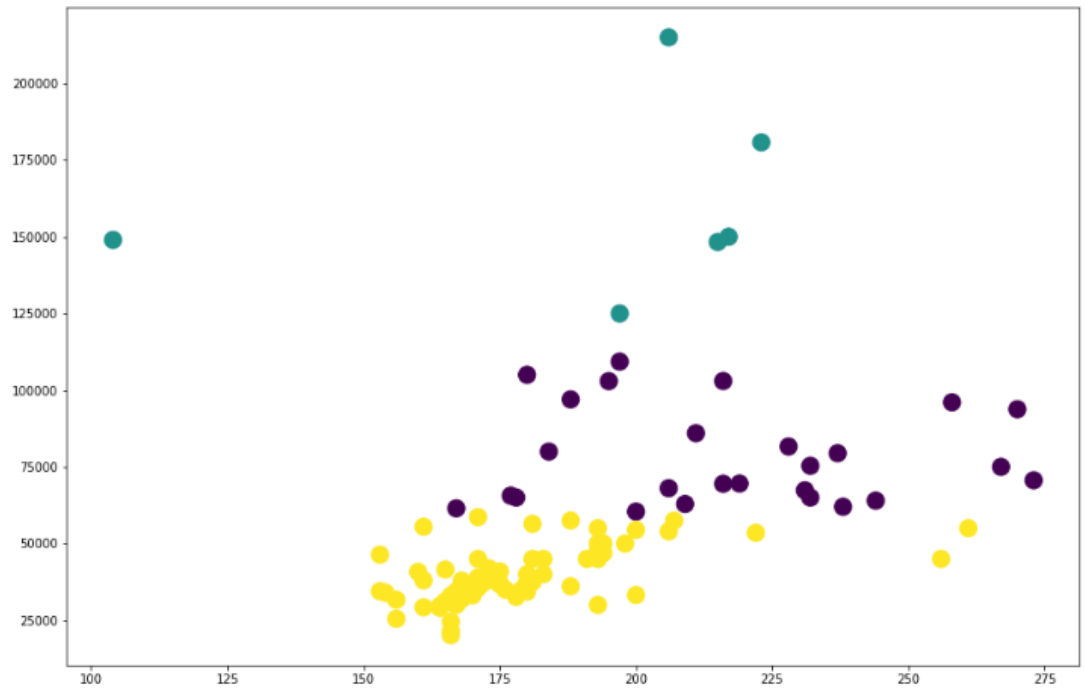

```
kmeans = KMeans(n_clusters=3)
data_predict = kmeans.fit_predict(data_model)

plt.figure(figsize=(15,10))
plt.scatter( x = 'Efficiency_WhKm' ,y = 'PriceEuro' , data = data_model , c = data_predict , s = 200 )
plt.xlabel("Efficiency_WhKm")
plt.ylabel("PriceEuro")
plt.show()
```



```
from scipy.cluster.hierarchy import linkage, dendrogram
#create demogram and find the best clustering value
merg = linkage(data_model,method="ward")
plt.figure(figsize=(25,15))
dendrogram(merg,leaf_rotation = 90)
plt.xlabel("data points")
plt.ylabel("euclidean distance")
plt.show()
```




```
from sklearn.cluster import AgglomerativeClustering
#create model
hiyerartical_cluster = AgglomerativeClustering(n_clusters = 3,affinity= "euclidean",linkage = "ward")
data_predict = hiyerartical_cluster.fit_predict(data_model)
plt.figure(figsize=(15,10))
plt.scatter( x = 'Efficiency_WhKm' ,y = 'PriceEuro' , data = data_model , c = data_predict , s = 200 )
plt.show()
```



Link to code:
https://colab.research.google.com/drive/1OsnJOrjebTHu0610rK5q6Xkg6JDnsqmQ

Conclusion:
Electric vehicles are the wave of the future! Manufacturing businesses are putting more effort into discover continuous electric vehicles rather if a start-up concentrate on their (Efficiency, Price) it may be base for their growth of their EV. There are several advantages to owning an electric car with the appropriate level of functionality and infrastructure. According to our market segmentation of EV the market is going to flourish in the future.