

EV MARKET SEGMENTATION PROJECT

- By S AJAY for FeyNN Labs

1. Market Variables

- Price
- Speed
- Distance range
- Popularity in Indian states

2. Datasets

Features/ column names include-

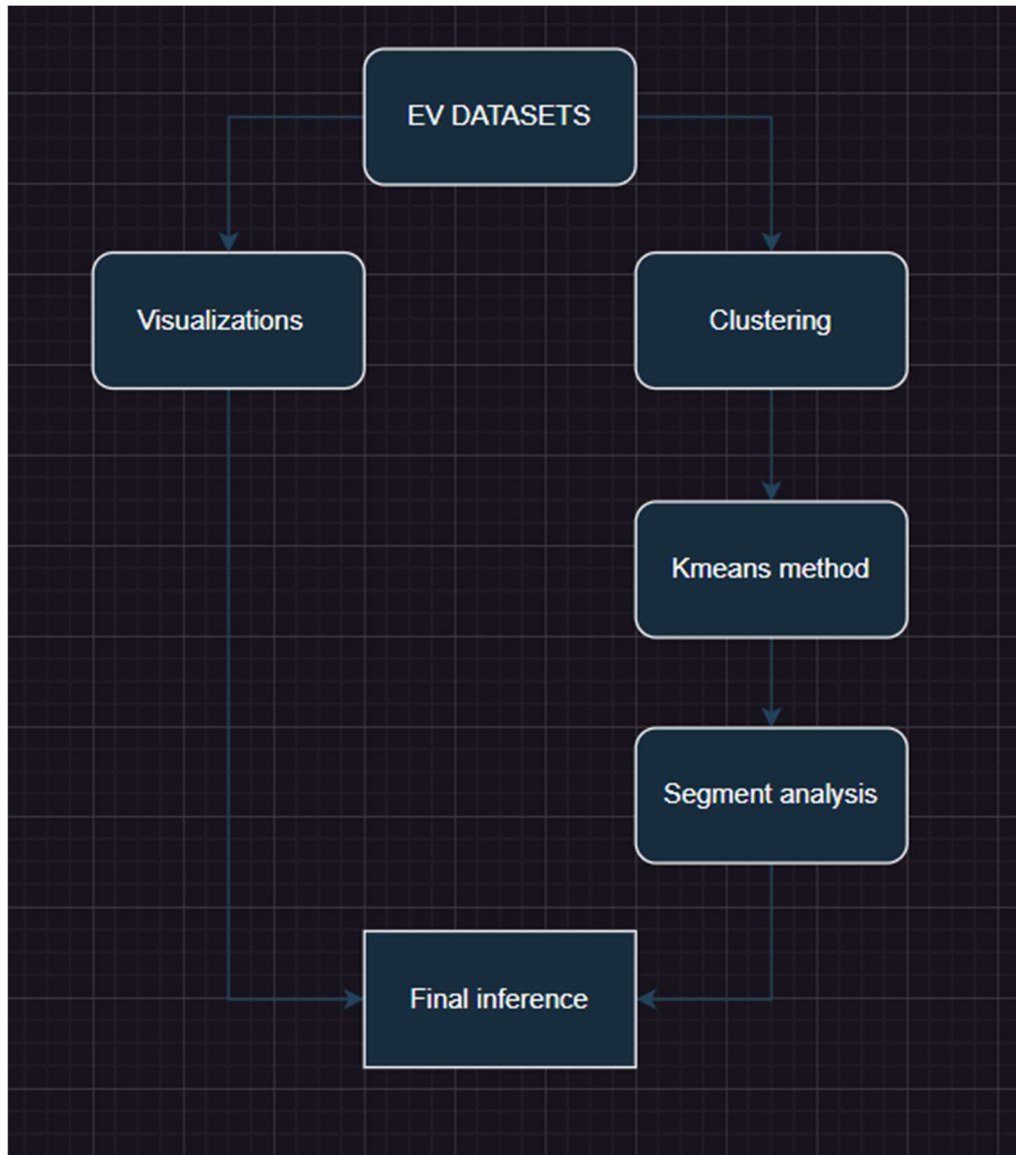
```
[17]: list(ev.columns.values)
```

```
[17]: ['Unnamed: 0',  
      'Brand',  
      'Model',  
      'AccelSec',  
      'TopSpeed_KmH',  
      'Range_Km',  
      'Efficiency_WhKm',  
      'FastCharge_KmH',  
      'RapidCharge',  
      'PowerTrain',  
      'PlugType',  
      'BodyStyle',  
      'Segment',  
      'Seats',  
      'PriceEuro']
```

```
: list(ev.columns.values)
```

```
: ['Sl. No',  
   'State',  
   'Two Wheelers (Category L1 & L2 as per Central Motor Vehicles Rules',  
   'Two Wheelers (Category L2 (CMVR))',  
   'Two Wheelers (Max power not exceeding 250 Watts)',  
   'Three Wheelers (Category L5 slow speed as per CMVR)',  
   'Three Wheelers (Category L5 as per CMVR)',  
   'Passenger Cars (Category M1 as per CMVR)',  
   'Buses',  
   'Total in state']
```

3. Methodology

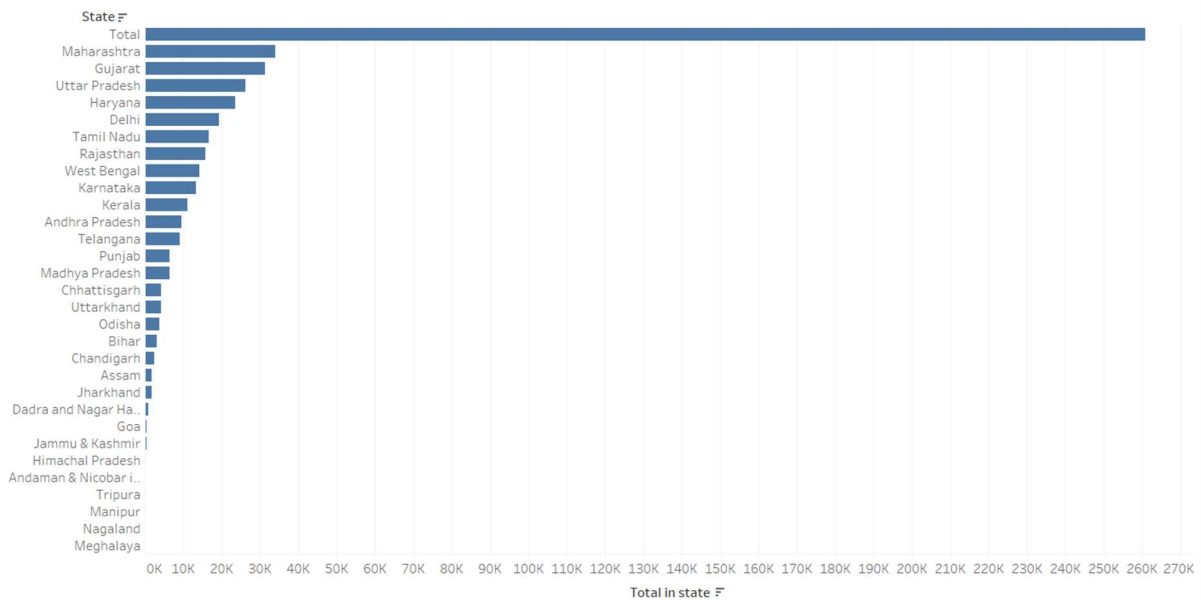


K Means clustering algorithm- It is a ML technique that uses data analysis to discover interesting structures within data. K-means clustering helps in market segmentation by identifying groups with similar characteristics in the EV market such as price, size, speed in our datasets.

4. Visualizations

4.1. Total sales in each state

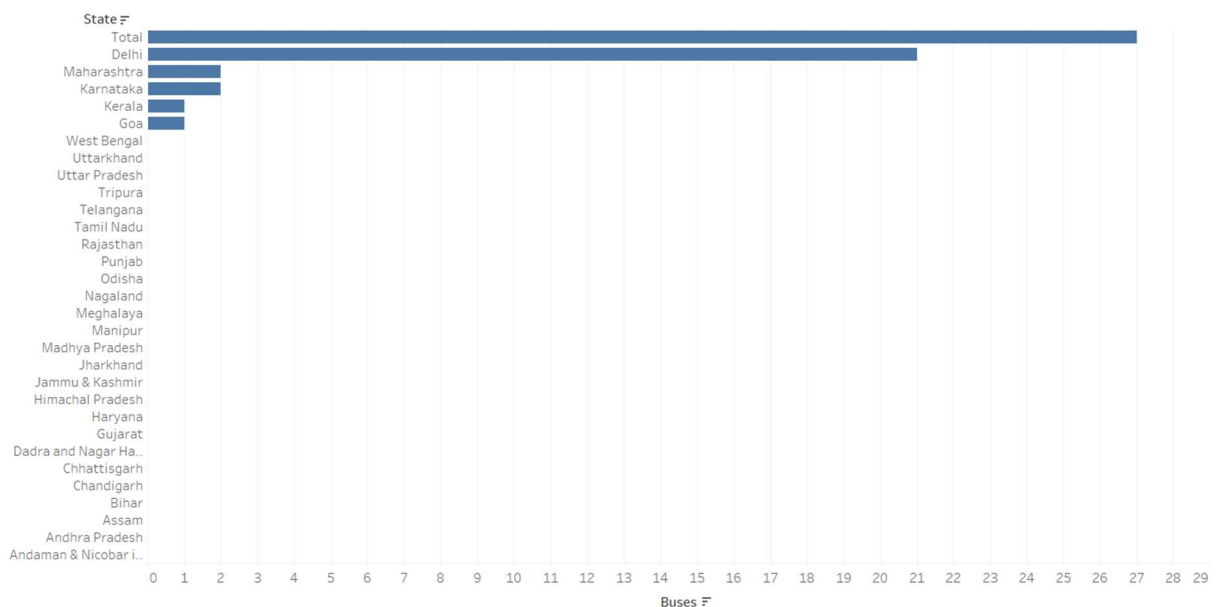
EV sales in Indian States



As we can see, Maharashtra, Gujarat, Uttar Pradesh account for top 3 EV sales.

4.2. Electric buses sales in Indian states

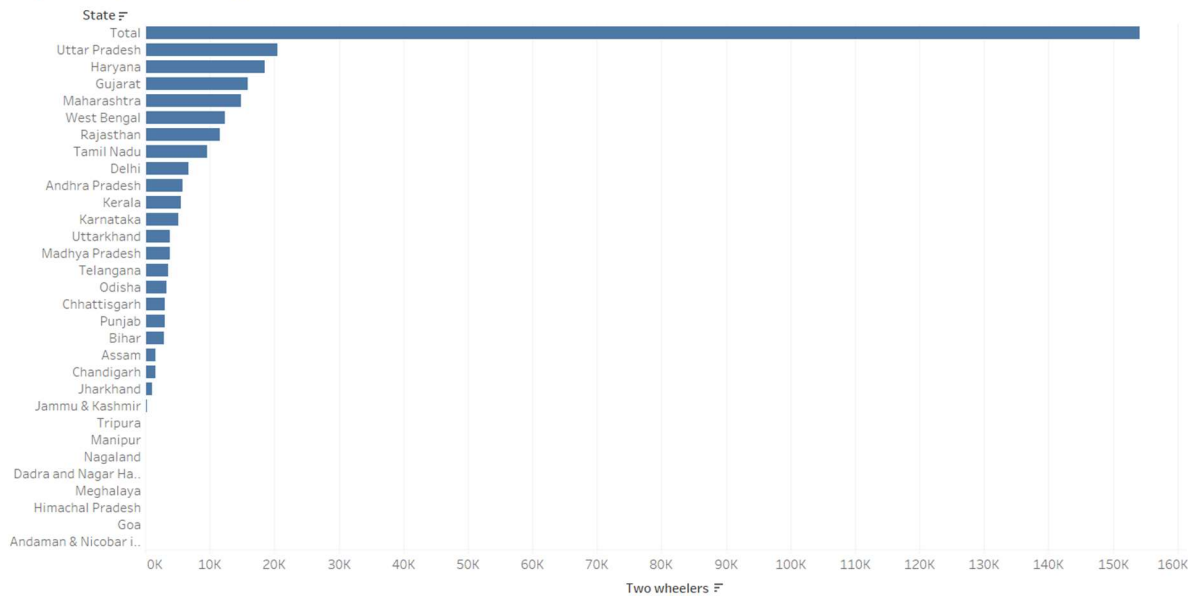
Electric buses sales in Indian States



As we can see, Delhi accounts for higher public transport in electric vehicles.

4.3. Electric Two wheelers sales in Indian States

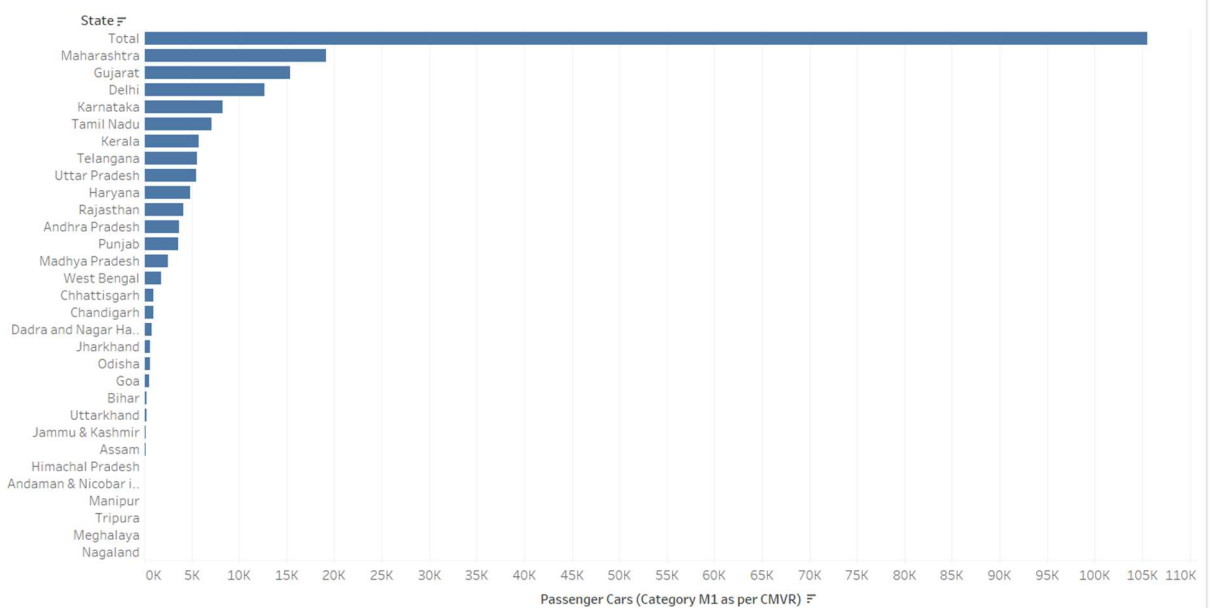
2 wheelers sales in Indian States



As we can see, UP, Haryana and Gujarat are the top 3 states with highest 2 wheeler sales

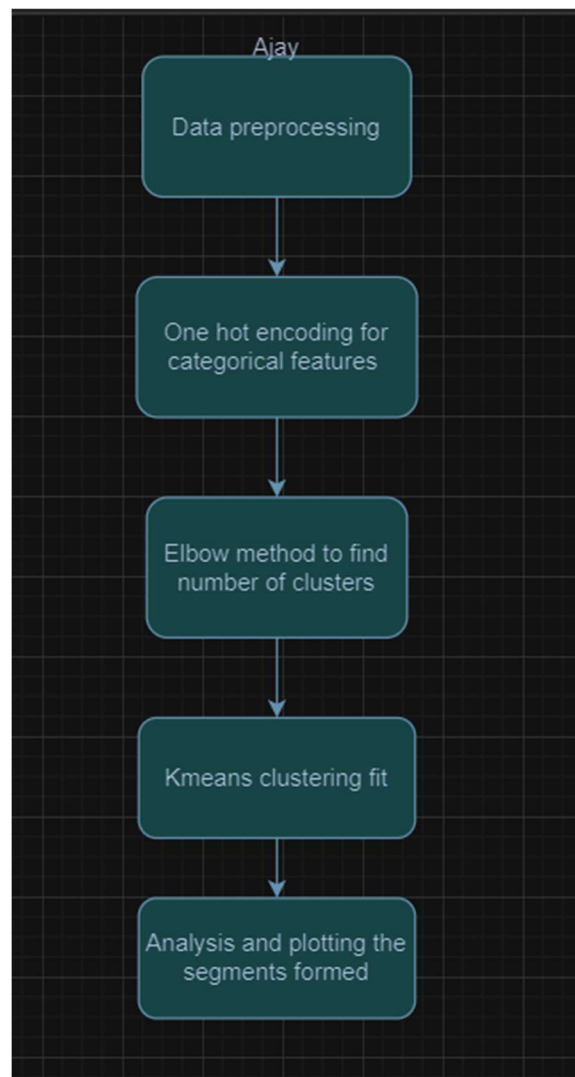
4.4. Electric Cars

2 wheelers sales in Indian States



As we can see, Maharashtra and Gujarat have highest EV 2 wheeler sales.

5. Segmentation using K Means Clustering



5.1. Data Pre processing

```
ev2 = ev.drop(ev.columns[[0, 1, 2, 9, 10, 13, 12]], axis = 1)
ev2.head()
```

```
ev2
```

	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	RapidCharge	BodyStyle	PriceEuro
0	4.6	233	450	161	940	Yes	Sedan	55480
1	10.0	160	270	167	250	No	Hatchback	30000
2	4.7	210	400	181	620	Yes	Liftback	56440
3	6.8	180	360	206	560	Yes	SUV	68040
4	9.5	145	170	168	190	Yes	Hatchback	32997

Dropping excess columns

5.2. One hot encoding

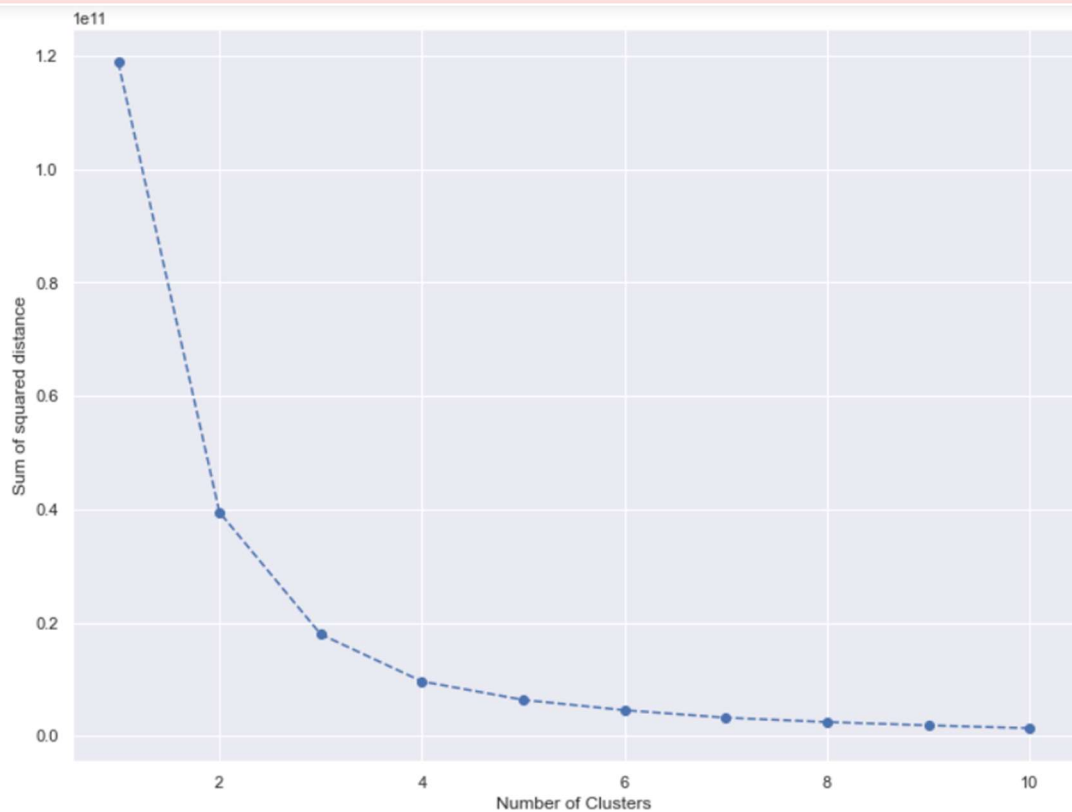
```
ev3 = pd.get_dummies(ev2, columns = ['BodyStyle', 'RapidCharge'])
ev3.head()
```

BodyStyle_Cabrio	BodyStyle_Hatchback	BodyStyle_Liftback	BodyStyle_MPV	BodyStyle_Pickup	BodyStyle_SPV	BodyStyle_SUV	BodyStyle_Sedan	BodyStyle_Static
0	0	0	0	0	0	0	0	1
0	1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0
0	1	0	0	0	0	0	0	0

5.3. Elbow method

```
results = {}

for i in range(1, 11):
    kmeans = KMeans(n_clusters=i, init='k-means++', random_state=42)
    kmeans.fit(ev3)
    results[i] = kmeans.inertia_
plt.figure(figsize=(12, 9))
plt.plot(results.keys(), results.values(), marker='o', linestyle='--')
plt.xlabel('Number of Clusters')
plt.ylabel('Sum of squared distance')
plt.show()
```



We will take 4 clusters as per the elbow obtained from the diagram.

5.4. KMeans label assignment to segments

```
[25]: kmeans = KMeans(n_clusters=4, init='k-means++', random_state=42)
      kmeans.fit(ev3)
```

```
[25]: KMeans(n_clusters=4, random_state=42)
```

```
[45]: ev_kmeans = ev3.copy()
      ev_kmeans['Segment'] = kmeans.labels_
      ev_kmeans
```

5.5. Segment analysis

	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	PriceEuro	BodyStyle_Cabrio	Body
Segment								
0	5.735	195.548	399.677	210.097	545.161	62917.452	0.000	
1	4.320	266.000	539.000	193.000	764.000	168616.400	0.200	
2	9.124	153.345	274.741	174.034	334.138	35120.707	0.034	
3	3.689	241.222	430.556	212.444	628.889	102007.444	0.000	

Upon renaming the segments based on the speed, price, size

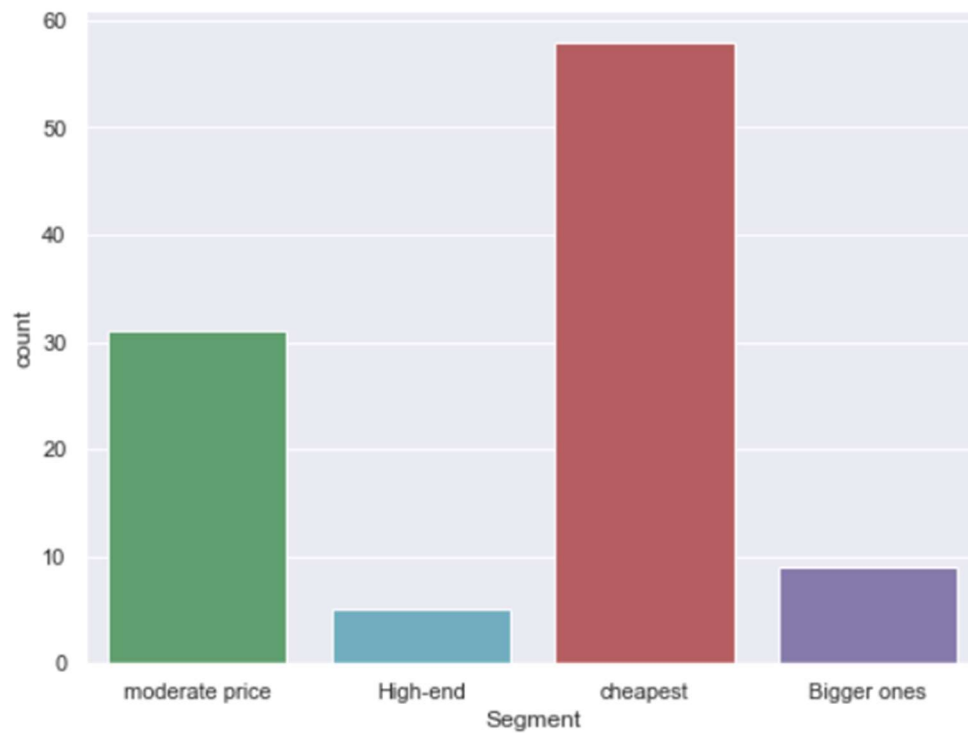
	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhKm	FastCharge_KmH	PriceEuro	Body
Segment							
moderate price	5.735	195.548	399.677	210.097	545.161	62917.452	
High-end	4.320	266.000	539.000	193.000	764.000	168616.400	
cheapest	9.124	153.345	274.741	174.034	334.138	35120.707	
Bigger ones	3.689	241.222	430.556	212.444	628.889	102007.444	

The segments obtained can usually be classified as

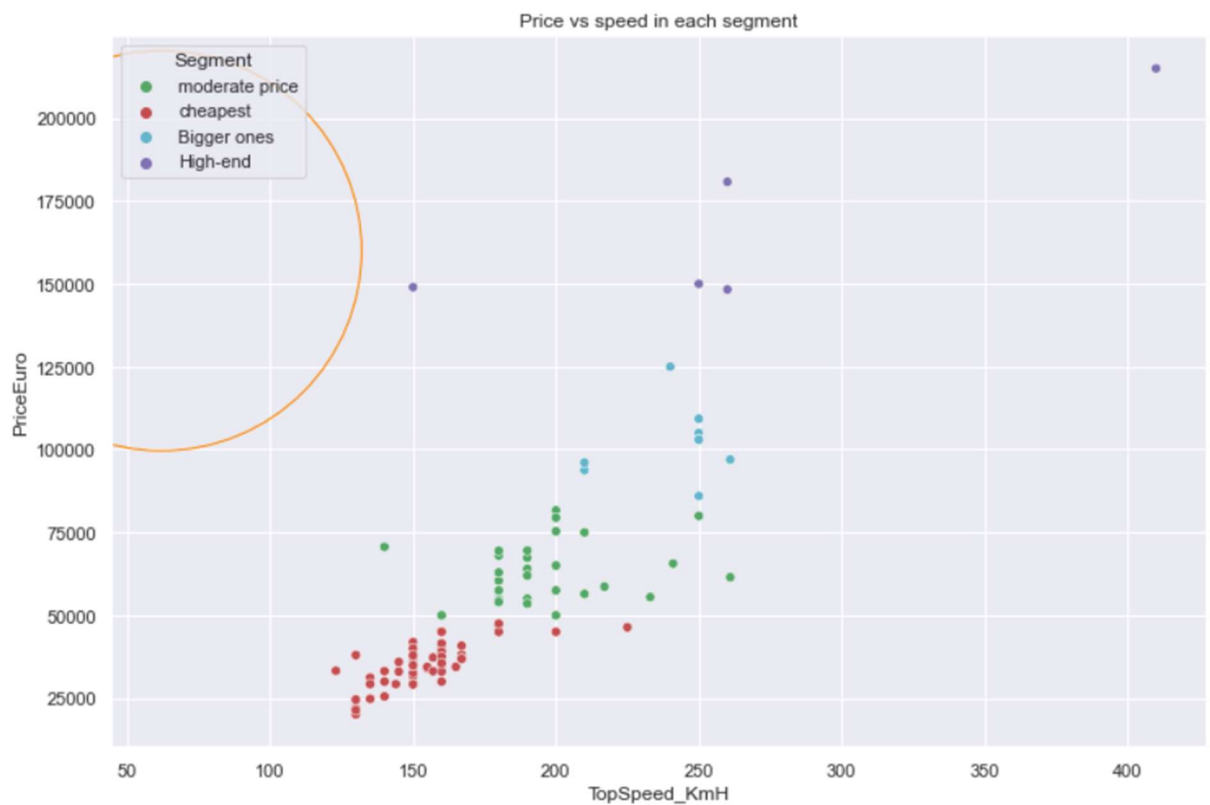
- Moderate price- slower- SUVs
- High end – faster- sedans
- Cheap- slower- Hatchback
- Bigger SUVs-Sedan

5.6. Cluster visualizations

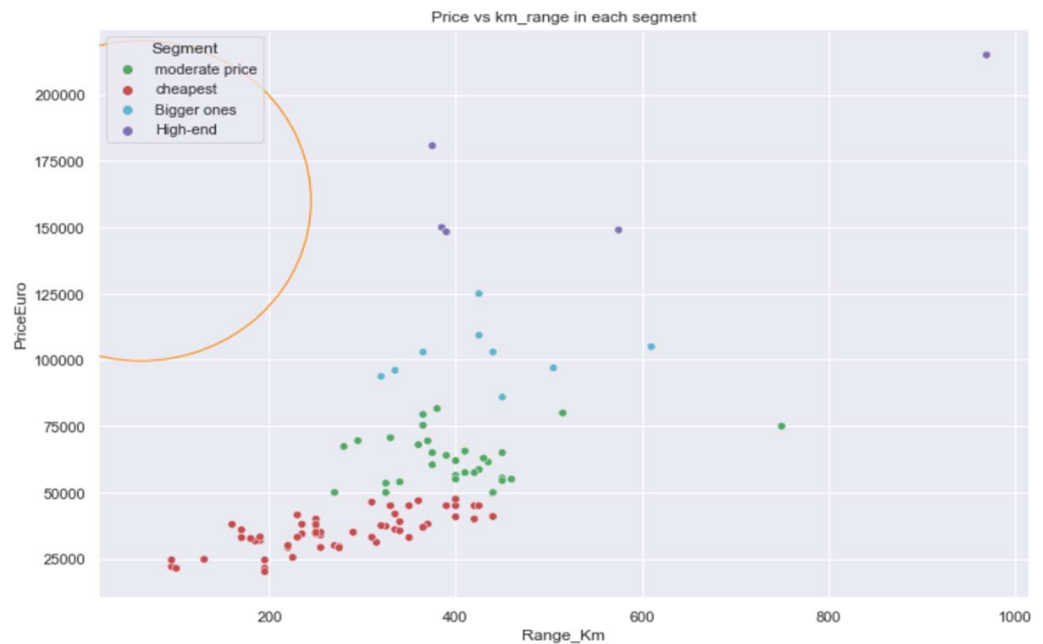
5.6.1. Number of data in each segment



5.6.2. Price vs Speed in segments



5.6.3. Price vs Distance range in segments



6. FINAL INFERENCE

- 4 segments were identified using KMeans clustering algorithm, the 4 namely
 - Moderate price- slower- SUVs: This segment of EV Market pertains to Moderately priced but slower cars. The size and type of cars available are mostly SUVs and few sedans.
 - High end – faster- sedans: The costliest cars with much higher speed and distance range. They have the best fast charging capacity. They are mostly available in Sedan type cars. Very few of them are available in the market.
 - Cheap- slower- Hatchback: Cheapest ones often lesser in range and has the lowest speed. But they form the largest portion of the market.
 - Bigger SUVs-Sedan: They typically consists of larger vehicles with moderate speed.
- All segments have rapid charge facility and are decently energy efficient.
- Indian market size is about 3 lakh electric vehicles per year. Uttar Pradesh, Gujarat, Delhi forms the largest of buyers. The EV Car is growing, specifically cheaper ones are preferred.
- Production of electric 2 wheelers can be focused.
- As far as passenger cars concerned, cheaper ones can be made and advertisement focused on states like Maharashtra, Gujarat, Delhi.

7. FUTURE SCOPE/ IMPROVEMENTS

- For better segmentation, we can focus more on collecting more demographic variables. Features such as age, income of buyers, and other characteristics can help define the target segments efficiently.
- Including hierarchical clustering algorithms and PCA for visualization can provide us with enhanced visualization.
- Ideally, I would like Income level of buyers, environmental consciousness measure of people, Electric Vehicle cost, geography as my top market variables.

8. GIT HUB LINK

https://github.com/Ajay2135/project2_marketsegmentation