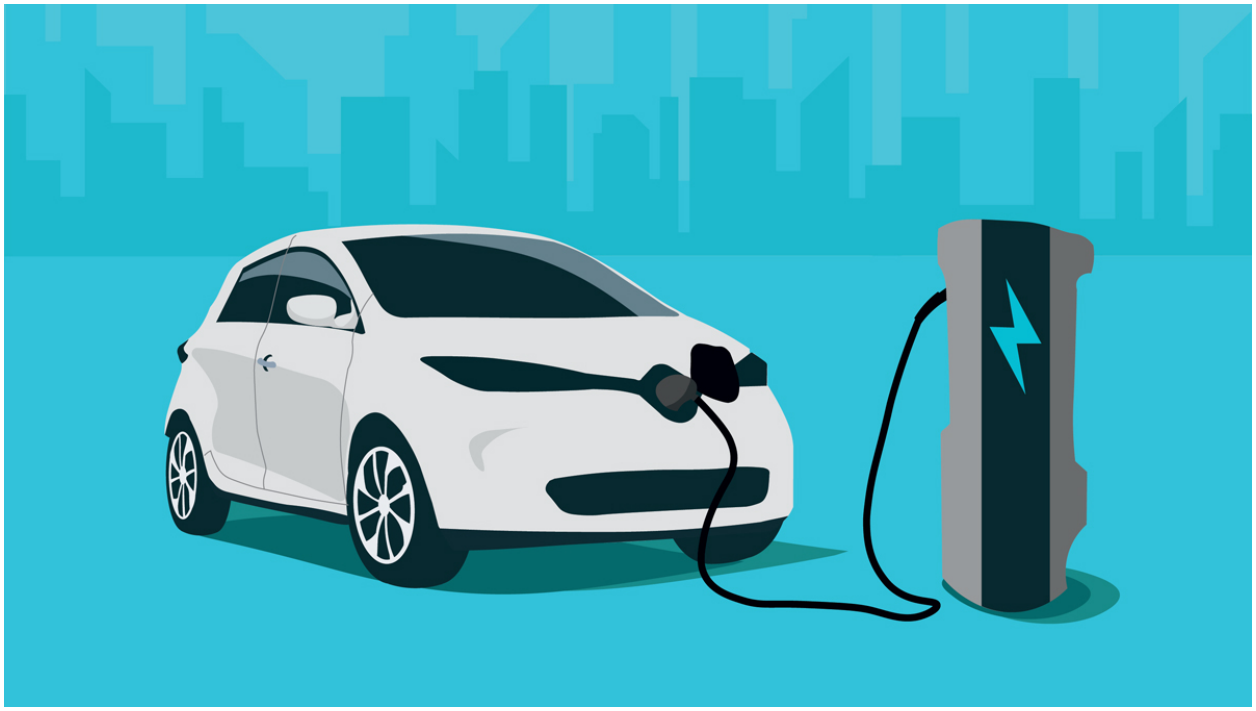
ANALYZING THE ELECTRIC VEHICLE MARKET USING SEGMENTATION

-PADMASHRI B V



The electric vehicle (EV) market in India has undergone a significant transformation in recent years, reflecting a global shift towards sustainable and environmentally friendly transportation solutions. As the world grapples with the challenges posed by climate change and seeks cleaner alternatives to traditional fossil fuel-powered vehicles, India has emerged as a key player in the growing electric vehicle market.

India's electric vehicle market has witnessed substantial growth in recent years, driven by a combination of government initiatives, technological advancements, and an increasing awareness of environmental sustainability. The government of India has introduced various policy measures and incentives to promote the adoption of electric vehicles, including the Faster Adoption and Manufacturing of Hybrid and Electric Vehicles (FAME) scheme. Under FAME, subsidies and incentives are provided to manufacturers and consumers to encourage the production and purchase of electric vehicles.

Major players in the automotive industry, both domestic and international, have responded to the changing landscape by investing in the development and production of electric vehicles. Established automakers and new entrants alike are introducing a diverse range of electric vehicles, from electric cars to two-wheelers and commercial vehicles. This diversification is contributing to a broader acceptance of electric mobility across various segments of the population.

The future of the electric vehicle market in India holds immense potential, with several factors contributing to its optimistic outlook. Collaborations between government bodies, private enterprises, and international organizations can accelerate the development of charging infrastructure and technological innovation. Partnerships with renewable energy providers can enhance the sustainability of electric vehicles by promoting clean energy sources for charging.

The electric vehicle market in India is at a pivotal juncture, poised for substantial growth and transformative change. While challenges persist, the commitment of the government, the dynamism of the automotive industry, and advancements in technology create a promising environment for the widespread adoption of electric vehicles.

The ongoing collaborative efforts and a focus on addressing key challenges will determine the trajectory of India's electric vehicle market, shaping a sustainable and environmentally conscious future for transportation in the country.

DATASET USED FOR ANALYSIS

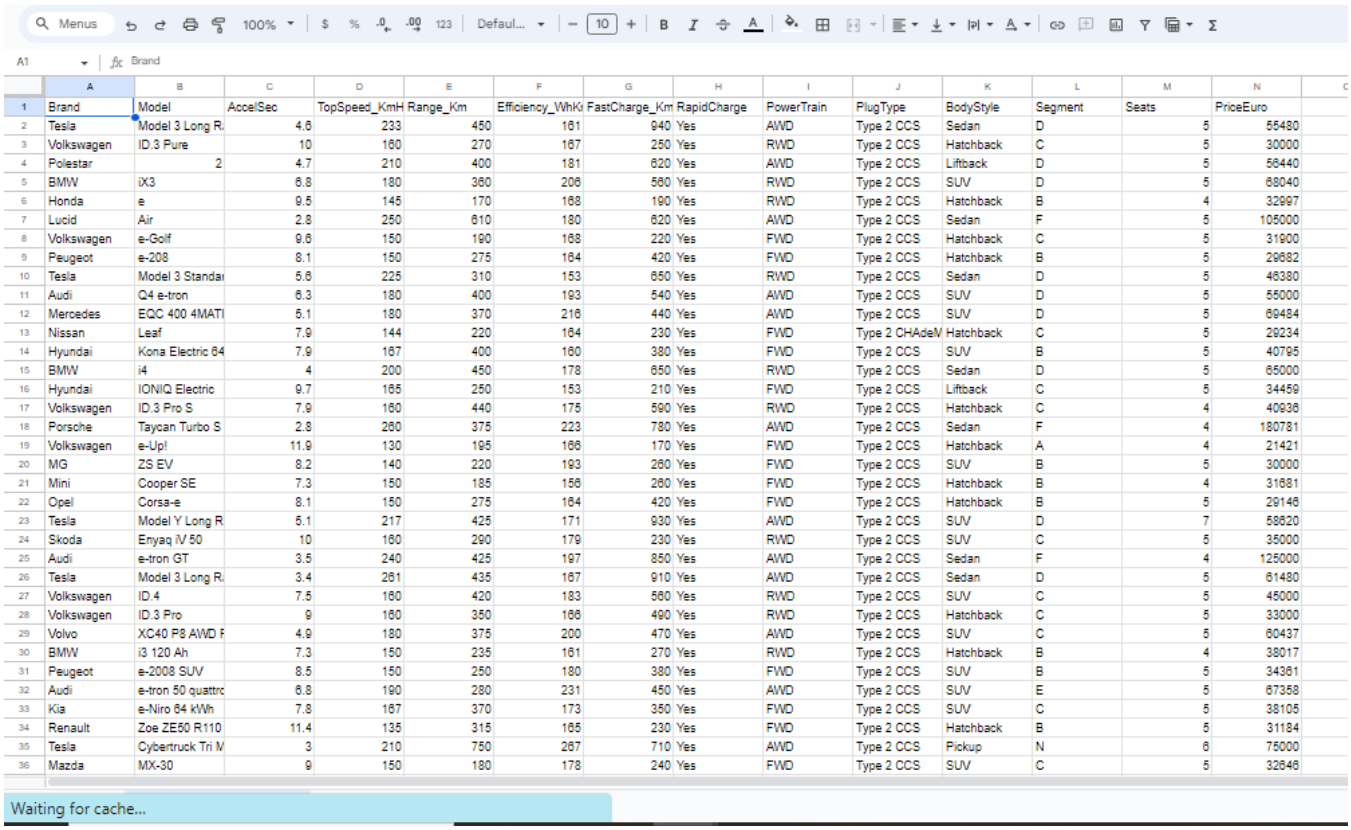
DATASET1-https://drive.google.com/file/d/1oxeZejiNFcGVvcAT38B-0e7AFkEIkbsV/view?usp=drive_link

DATASET2-https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset/data?select=ElectricCarData_Clean.csv

DATASET 1 is ev stats in india which consists of electric vehicle data of two wheelers ,three wheelers,passenger cars and buses and total number electric vehicle used in all states

Sl. No	State	Two Wheelers ((	Two Wheelers ((	Two Wheelers (I	Three Wheelers	Three Wheelers	Passenger Cars	Buses	Total in state
1	Andhra Pradesh	431	692	4689	0	0	3680	0	9492
2	Assam	463	138	1006	0	117	151	0	1875
3	Bihar	252	430	2148	6	64	271	0	3171
4	Chhattisgarh	613	382	2078	58	106	997	0	4234
5	Delhi	1395	251	5018	0	1	12695	21	19381
6	Goa	0	0	0	0	0	513	1	514
7	Gujarat	7182	217	8476	0	4	15388	0	31267
8	Haryana	3162	1504	13908	113	24	4878	0	23589
9	Himachal Prade	0	0	0	0	0	98	0	98
10	Jammu & Kashn	2	76	152	0	0	208	0	438
11	Jharkhand	75	228	736	9	7	655	0	1710
12	Karnataka	784	1104	3252	2	0	8242	2	13386
13	Kerala	432	78	4961	1	0	5729	1	11202
14	Madhya Pradesl	503	378	2904	8	106	2562	0	6461
15	Maharashtra	2630	2097	10146	6	3	19129	2	34013
16	Manipur	16	8	11	0	5	12	0	52
17	Meghalaya	0	0	0	0	0	6	0	6
18	Nagaland	0	20	3	0	0	1	0	24
19	Odisha	377	824	2031	0	37	594	0	3863
20	Punjab	698	300	1968	0	5	3567	0	6538
21	Rajasthan	2036	1153	8375	19	64	4116	0	15763
22	Tamil Nadu	491	863	8260	0	0	7132	0	16746
23	Telangana	535	711	2256	2	0	5530	0	9034
24	Tripura	28	9	36	0	0	8	0	81
25	Uttar Pradesh	2954	2355	15199	117	139	5445	0	26209
26	Uttarkhand	423	168	3239	45	38	265	0	4178
27	West Bengal	1451	65	10781	3	0	1840	0	14140
28	Andaman & Nic	0	0	0	0	0	82	0	82
29	Chandigarh	612	18	896	0	0	974	0	2500
30	Dadra and Naga	4	0	9	0	0	803	0	816
31	Total	27549	14069	112538	389	720	105571	27	260863

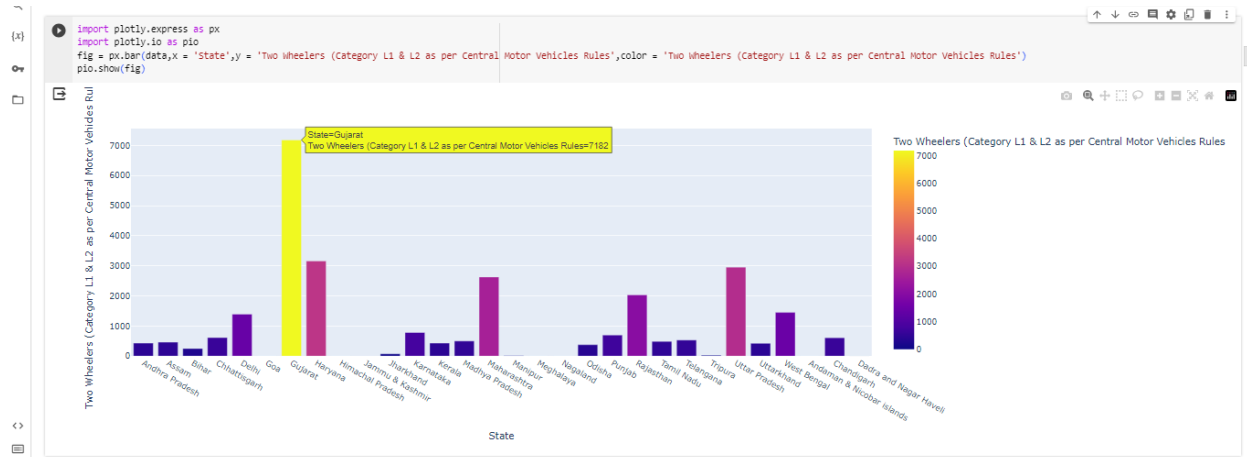
DATASET 2 consists electric car data which consists of brand, model, Accelsec, Topspeed_kmh, Range_km, efficency_kmh, fastcharge_kmh, Rapidcharge , powertype, Seat, Plug_type, segment, body style, PriceEuro



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Brand	Model	AccelSec	TopSpeed_KmH	Range_Km	Efficiency_WhK	FastCharge_Km	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro
2	Tesla	Model 3 Long R	4.6	233	450	161	940	Yes	AWD	Type 2 CCS	Sedan	D	5	55480
3	Volkswagen	ID.3 Pure	10	160	270	167	250	Yes	RWD	Type 2 CCS	Hatchback	C	5	30000
4	Polestar	2	4.7	210	400	181	620	Yes	AWD	Type 2 CCS	Liftback	D	5	56440
5	BMW	iX3	6.8	180	360	206	560	Yes	RWD	Type 2 CCS	SUV	D	5	68040
6	Honda	e	9.5	145	170	168	190	Yes	RWD	Type 2 CCS	Hatchback	B	4	32997
7	Lucid	Air	2.8	250	610	180	620	Yes	AWD	Type 2 CCS	Sedan	F	5	105000
8	Volkswagen	e-Golf	9.6	150	190	168	220	Yes	FWD	Type 2 CCS	Hatchback	C	5	31900
9	Peugeot	e-208	8.1	150	275	164	420	Yes	FWD	Type 2 CCS	Hatchback	B	5	29682
10	Tesla	Model 3 Standa	5.6	225	310	153	650	Yes	RWD	Type 2 CCS	Sedan	D	5	46380
11	Audi	Q4 e-tron	6.3	180	400	193	540	Yes	AWD	Type 2 CCS	SUV	D	5	55000
12	Mercedes	EQC 400 4MATI	5.1	180	370	216	440	Yes	AWD	Type 2 CCS	SUV	D	5	69484
13	Nissan	Leaf	7.9	144	220	164	230	Yes	FWD	Type 2 CHAdeM	Hatchback	C	5	29234
14	Hyundai	Kona Electric 64	7.9	167	400	160	380	Yes	FWD	Type 2 CCS	SUV	B	5	40795
15	BMW	i4	4	200	450	178	650	Yes	RWD	Type 2 CCS	Sedan	D	5	65000
16	Hyundai	IONIQ Electric	9.7	165	250	153	210	Yes	FWD	Type 2 CCS	Liftback	C	5	34459
17	Volkswagen	ID.3 Pro S	7.9	160	440	175	590	Yes	RWD	Type 2 CCS	Hatchback	C	4	40936
18	Porsche	Taycan Turbo S	2.8	260	375	223	780	Yes	AWD	Type 2 CCS	Sedan	F	4	180781
19	Volkswagen	e-Up!	11.9	130	195	166	170	Yes	FWD	Type 2 CCS	Hatchback	A	4	21421
20	MG	ZS EV	8.2	140	220	193	260	Yes	FWD	Type 2 CCS	SUV	B	5	30000
21	Mini	Cooper SE	7.3	150	185	156	260	Yes	FWD	Type 2 CCS	Hatchback	B	4	31681
22	Opel	Corsa-e	8.1	150	275	164	420	Yes	FWD	Type 2 CCS	Hatchback	B	5	29146
23	Tesla	Model Y Long R	5.1	217	425	171	930	Yes	AWD	Type 2 CCS	SUV	D	7	58620
24	Skoda	Enyaq iV 50	10	160	290	179	230	Yes	RWD	Type 2 CCS	SUV	C	5	35000
25	Audi	e-tron GT	3.5	240	425	197	850	Yes	AWD	Type 2 CCS	Sedan	F	4	125000
26	Tesla	Model 3 Long R	3.4	261	435	167	910	Yes	AWD	Type 2 CCS	Sedan	D	5	61480
27	Volkswagen	ID.4	7.5	160	420	183	560	Yes	RWD	Type 2 CCS	SUV	C	5	45000
28	Volkswagen	ID.3 Pro	9	160	350	166	490	Yes	RWD	Type 2 CCS	Hatchback	C	5	33000
29	Volvo	XC40 P8 AWD E	4.9	180	375	200	470	Yes	AWD	Type 2 CCS	SUV	C	5	60437
30	BMW	i3 120 Ah	7.3	150	235	161	270	Yes	RWD	Type 2 CCS	Hatchback	B	4	38017
31	Peugeot	e-2008 SUV	8.5	150	250	180	380	Yes	FWD	Type 2 CCS	SUV	B	5	34361
32	Audi	e-tron 50 quattro	6.8	190	280	231	450	Yes	AWD	Type 2 CCS	SUV	E	5	67358
33	Kia	e-Niro 64 kWh	7.8	167	370	173	350	Yes	FWD	Type 2 CCS	SUV	C	5	38105
34	Renault	Zoe ZE50 R110	11.4	135	315	165	230	Yes	FWD	Type 2 CCS	Hatchback	B	5	31184
35	Tesla	Cybertruck Tri M	3	210	750	267	710	Yes	AWD	Type 2 CCS	Pickup	N	6	75000
36	Mazda	MX-30	9	150	180	178	240	Yes	FWD	Type 2 CCS	SUV	C	5	32846

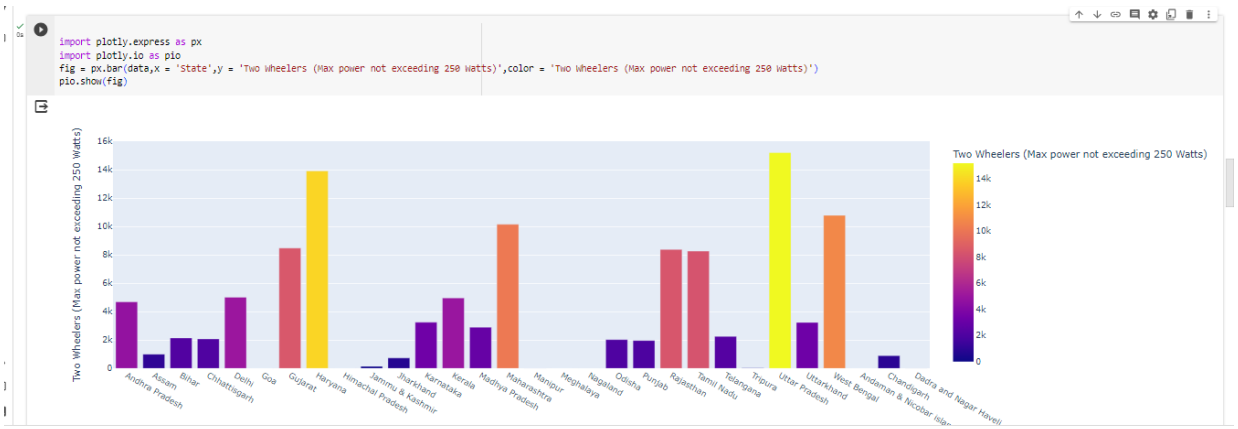
Waiting for cache...

EXPLORATORY DATA ANALYSIS OF EV STATS OF INDIA AND THEIR INTERPRETATION

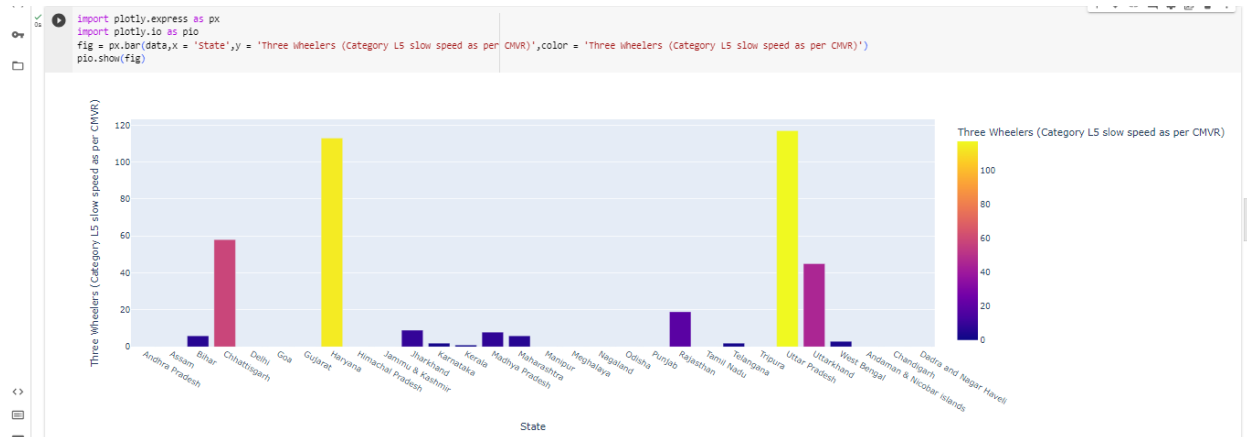


- **From the above Figure it is identified that Gujarat has highest number of Two Wheelers (Category L1 & L2 as per**

Central Motor Vehicles Rules), but the states like Goa, Himachal pradesh, Jammu & Kashmir, Meghalaya, Nagaland, Andaman and nicobar islands, Dadra and Nagar Haveli the Two Wheelers (Category L1 & L2 as per Central Motor Vehicles Rules) does not exist

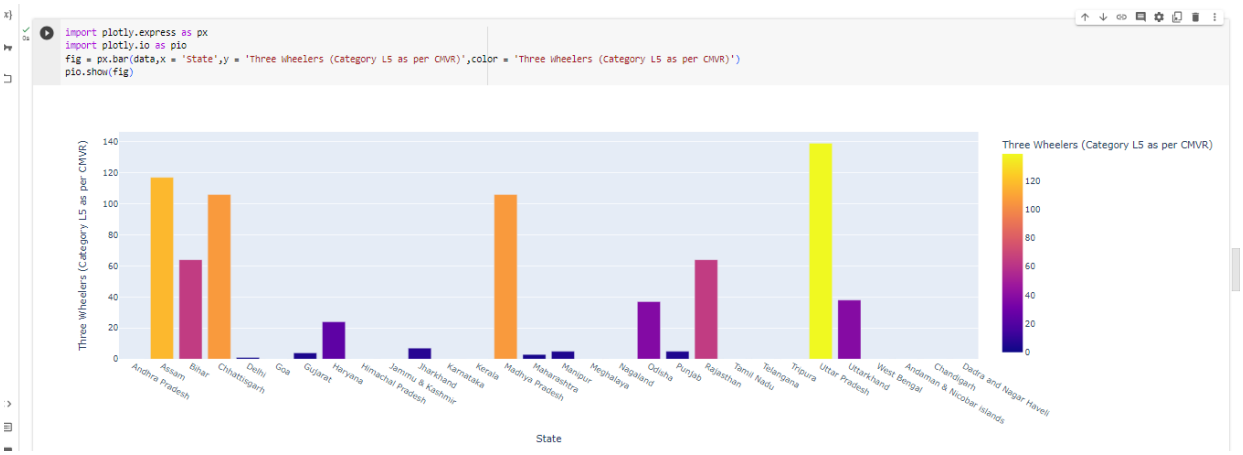


- **From the above figure it is identified that Uttar Pradesh has the highest number of Two Wheelers (Category L2 (CMVR)) ,but the states like Goa, Himachal pradesh, Meghalaya, Andaman and nicobar islands, Dadra and Nagar Haveli the Two Wheelers (Category L2 (CMVR)) does not exist**



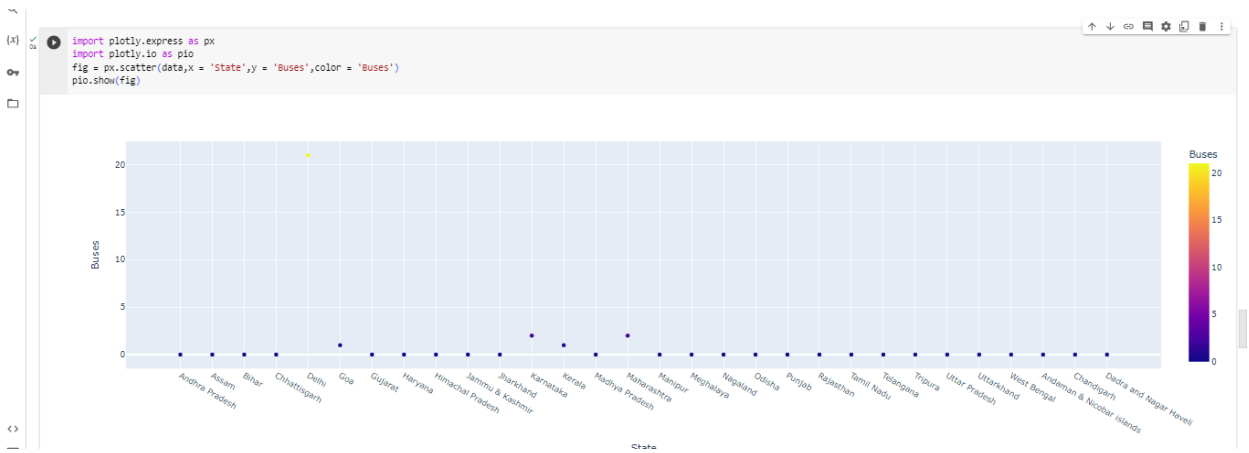
- **From the above Figure it is identified that Uttar Pradesh has the highest number of Two Wheelers (Max power not exceeding 250 Watts) ,but the states like Goa, Himachal pradesh, Jammu & Kashmir, Meghalaya, Nagaland, Manipur,Andaman and nicobar islands, Dadra and**

Nagar Haveli the Two Wheelers (Max power not exceeding 250 Watts) , does not exist

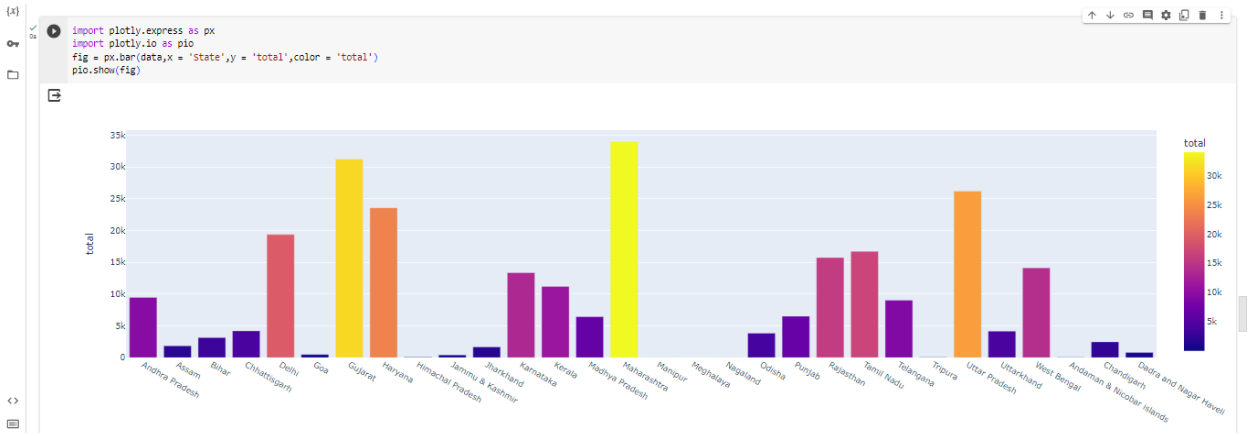


- **From the above figure it is identified that only Uttar Pradesh and Haryana has the highest number of Three Wheelers (Category L5 slow speed as per CMVR) and many states less count and some states the Three Wheelers (Category L5 slow speed as per CMVR) does not exist**

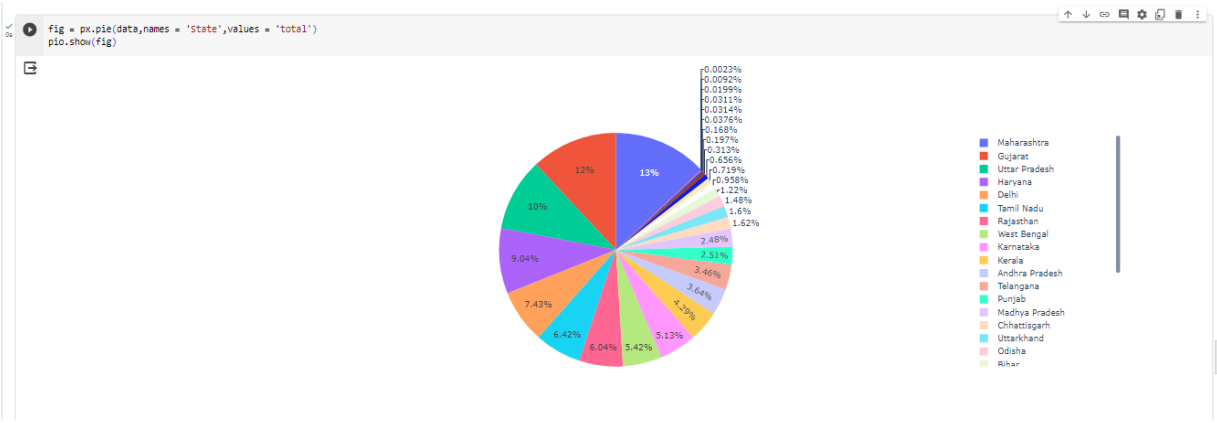
- **From the above figure it is identified that Uttar pradesh has highest number of Three Wheelers (Category L5 as per CMVR) , but in south india and many states it does not exist**



- **From the above figure it is identified that Delhi has highest number of Buses and other states has very low and in many states no Buses exist**

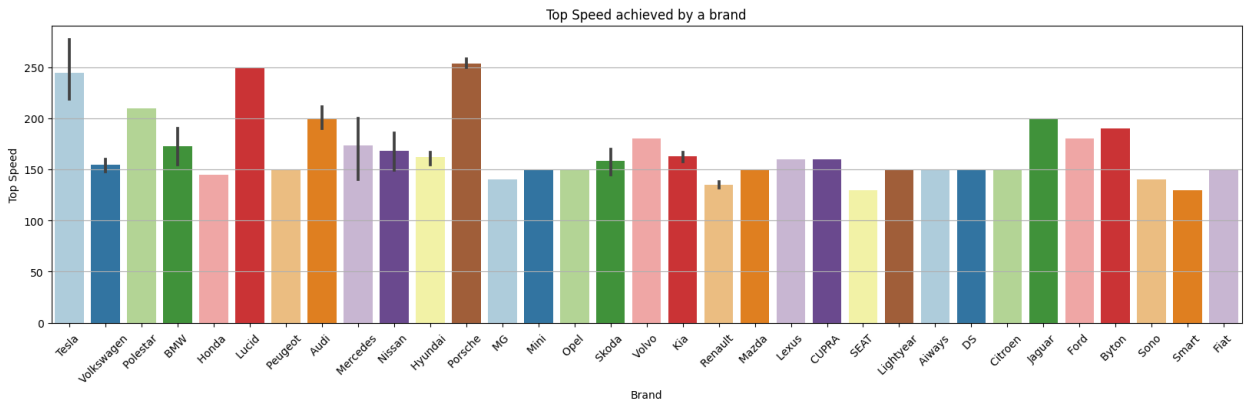


- **From the above figure it is identified that Maharashtra has highest number of Passenger Cars (Category M1 as per CMVR) but states like manipur ,meghalaya, nagaland, Tripura the Passenger Cars (Category M1 as per CMVR) does not exist**

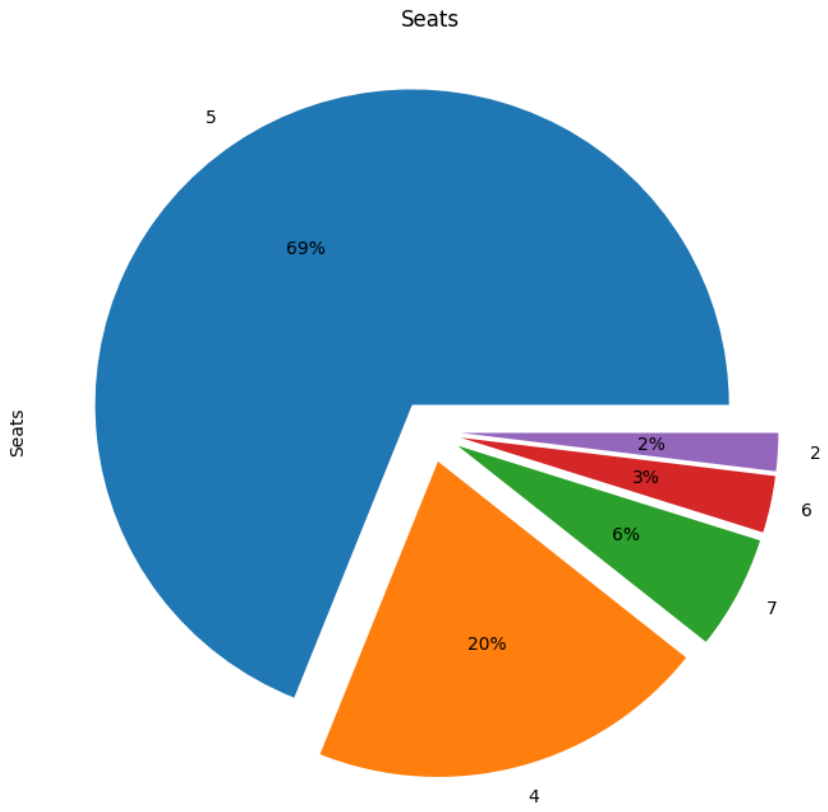


- **From the above figure it is identified that overall counting of the every vehicle numbers of each Maharashtra has highest number of Electric vehicles and Manipur, Nagaland ,Meghalaya has no Electric Vehicles in their State**

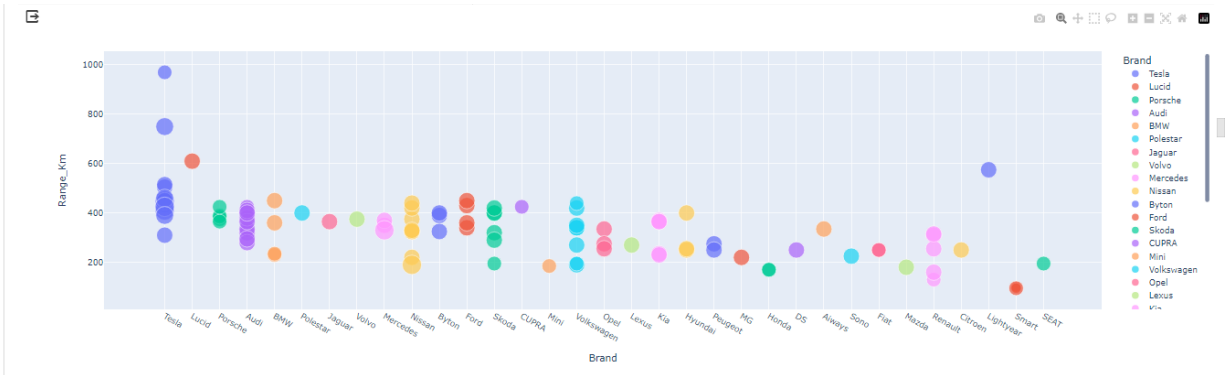
EXPLORATORY DATA ANALYSIS OF ELECTRIC CAR DATA AND THEIR INTERPRETATION



Porsche achieves top speed among all other brands and SEAT has low top speed among all brand

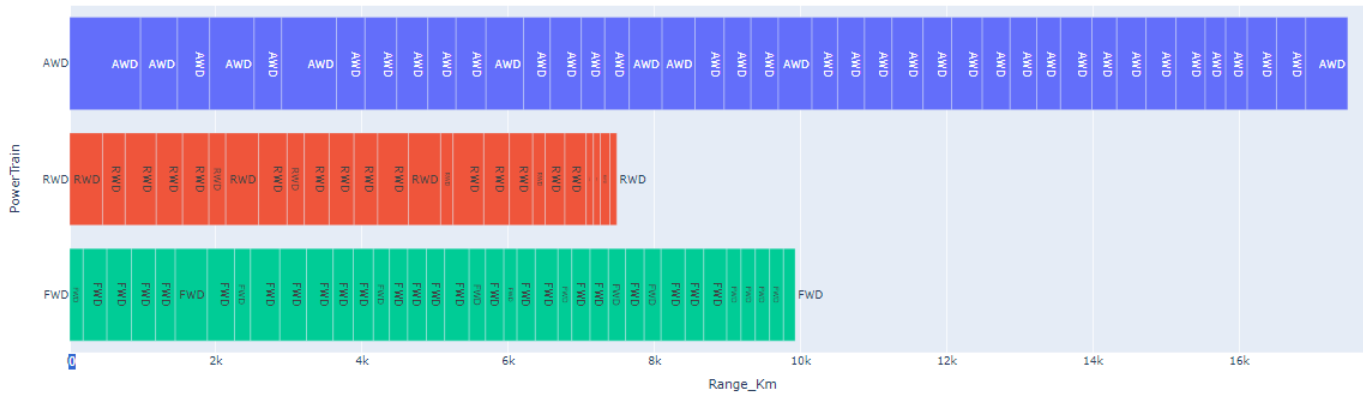


Many cars have seats of 5 and seats of 4 compared to others

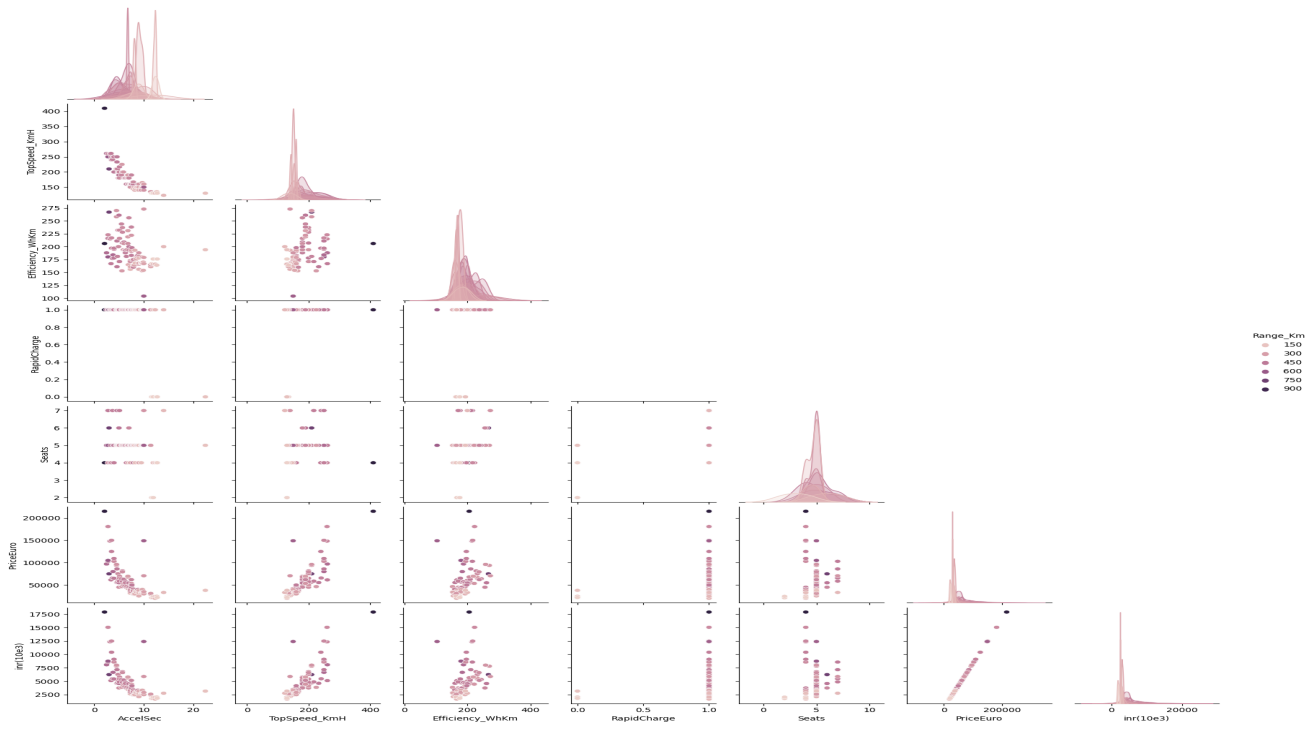


Tesla has highest electric vehicle Range per km among all the brands

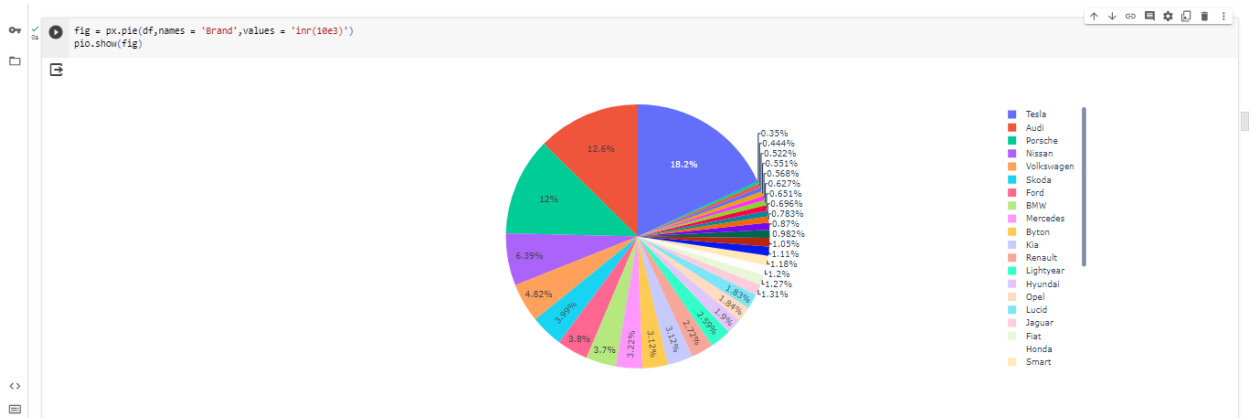
```
import plotly.express as px
import plotly.io as pio
fig = px.bar(df,x = 'Range_Km',y = 'PowerTrain',color = 'PowerTrain',text='PowerTrain')
pio.show(fig)
```



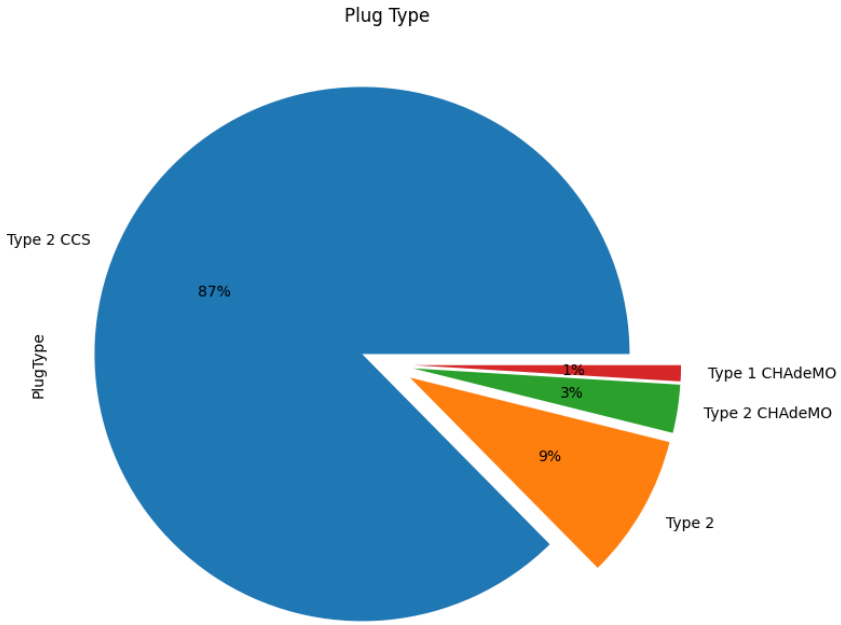
Distribution of Range per km with Power Train , the AWD has high Range per km



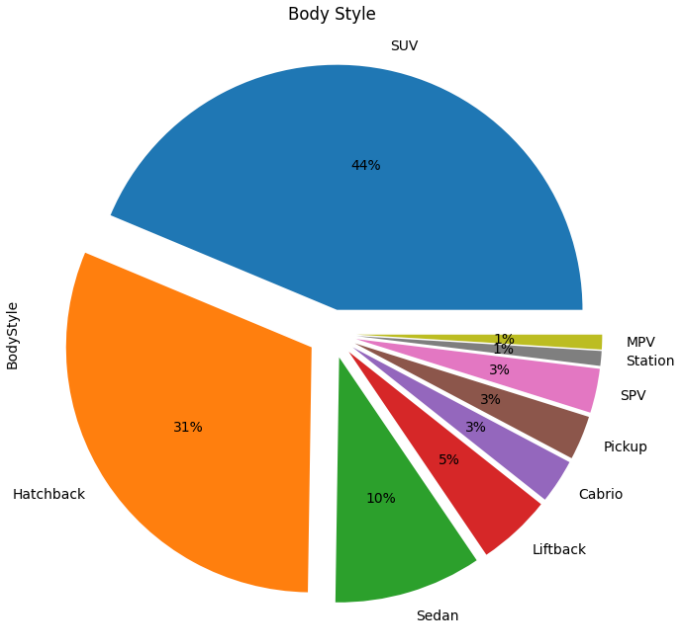
Pairplot with hue of Range per km with numeric features



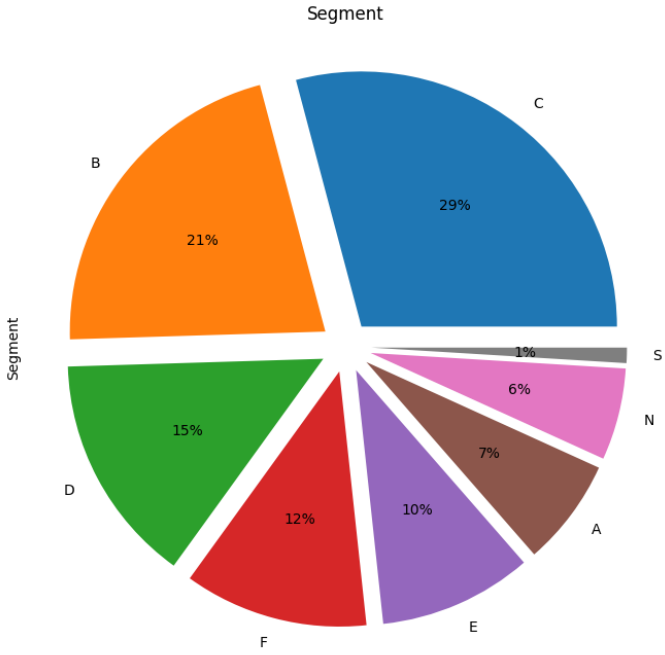
Tesla has more Brand value of inr compare to others



Type 2 CCS is the most used plug type



SUV is most used body type in electric vehicle



Most vehicles belong to Segment C

MODEL BUILDING OF BOTH DATASET USING K-MEANS ALGORITHM AND PRINCIPAL COMPONENT ANALYSIS

The K-means algorithm is widely employed in market segmentation to group similar customers or entities based on their characteristics, behaviors, or preferences. In this context, the algorithm helps businesses analyze large datasets containing diverse customer attributes, such as purchasing patterns, demographics, and psychographics. By iteratively partitioning the dataset into K clusters, where K represents the predetermined number of segments, K-means enables companies to identify homogeneous groups within their target market. This segmentation facilitates more effective and personalized marketing strategies, as businesses can tailor their approaches to the distinct needs and preferences of each cluster. Through the iterative process of assigning data points to clusters and optimizing centroids, K-means optimally divides the market, allowing organizations to enhance their understanding of customer segments and consequently refine product offerings, marketing messages, and customer experiences for improved overall business performance.

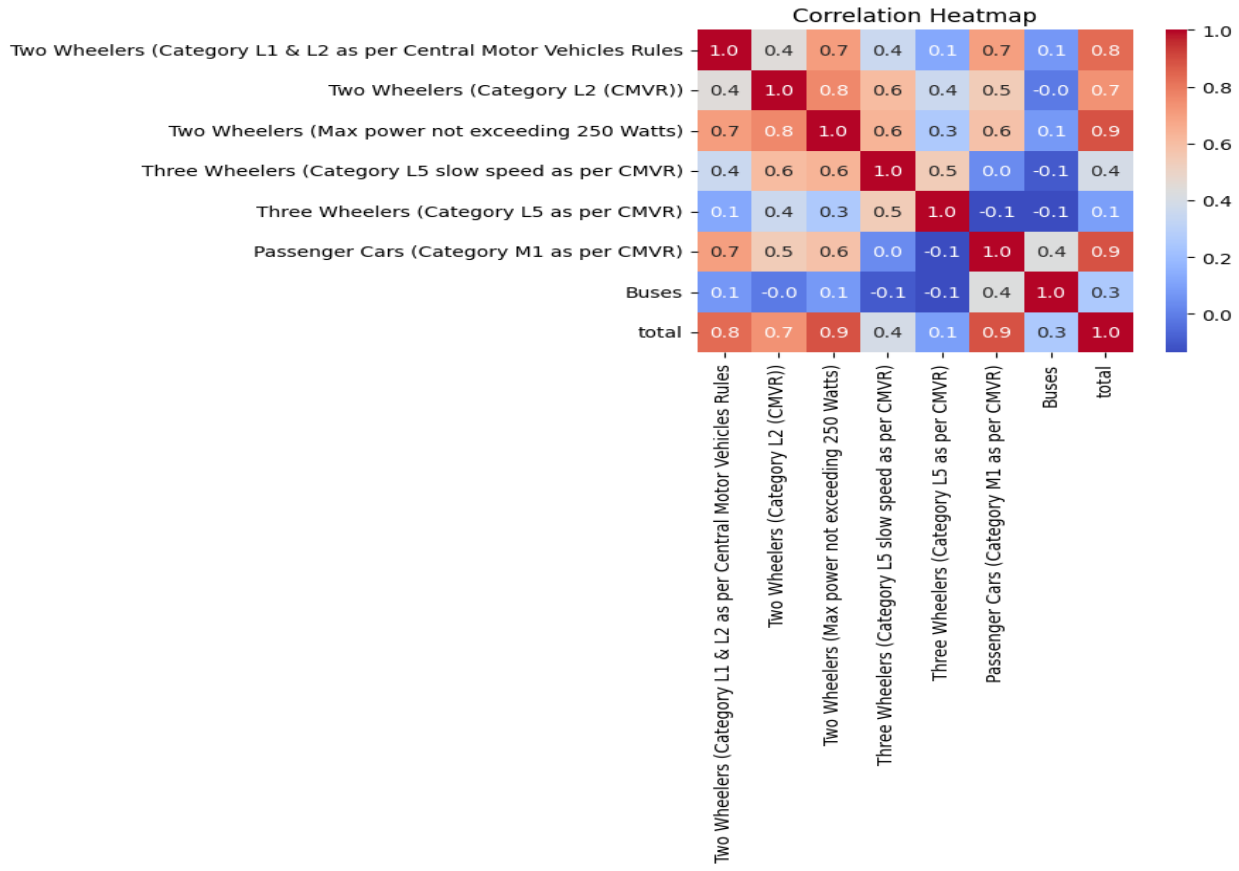
Principal Component Analysis (PCA) is a valuable technique in market segmentation, particularly when dealing with high-dimensional datasets. In the context of market research, PCA can be applied to reduce the dimensionality of the data by identifying the most significant features or variables that capture the underlying patterns of customer behavior or preferences. By transforming the original variables into a new set of uncorrelated variables, known as principal components, PCA allows businesses to focus on the key aspects that drive market segmentation. This reduction in dimensionality not only simplifies the analysis but also helps in visualizing and interpreting the data more effectively. Through PCA, businesses can uncover latent structures within their customer

data, discerning the major factors influencing market segments. This information proves instrumental in crafting targeted marketing strategies and tailoring products or services to meet the specific needs of distinct customer segments, ultimately enhancing the precision and effectiveness of market segmentation efforts.

In our dataset the first one comes under the segmentation category of GEOGRAPHIC SEGMENTATION because the values statistics of electric vehicles were divided state wise. The second dataset of electric car data and most of the feature come under BEHAVIORAL SEGMENTATION.

EV STATS MODEL BUILDING

This dataset has very vital information in less data.

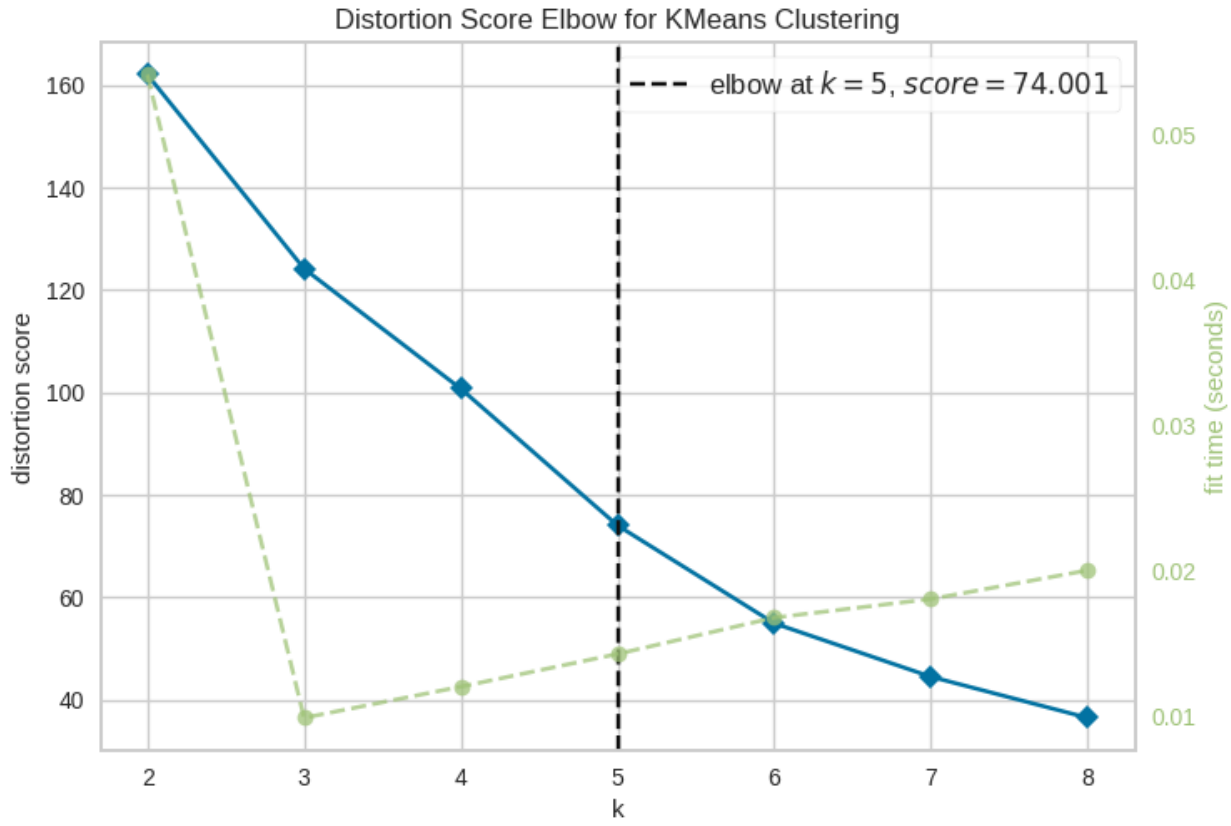


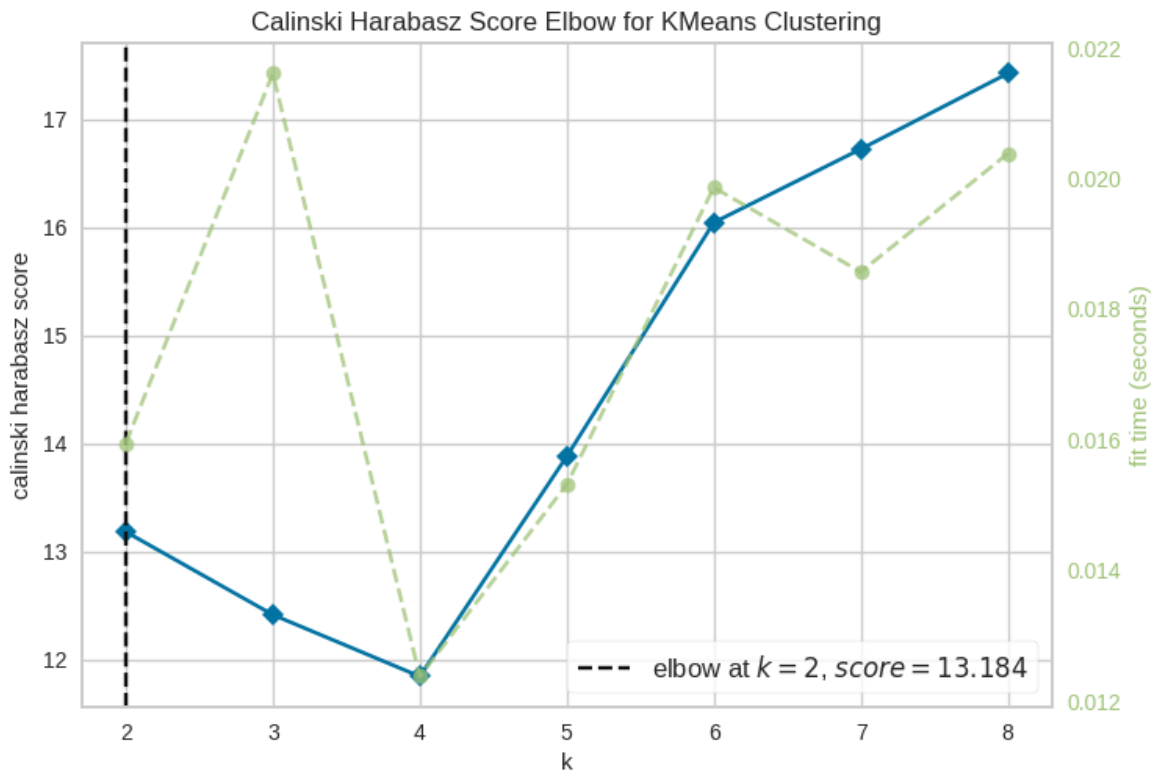
```
[ ] from sklearn.model_selection import train_test_split
    from sklearn.preprocessing import StandardScaler
    from sklearn.decomposition import PCA
    from sklearn.cluster import KMeans
```

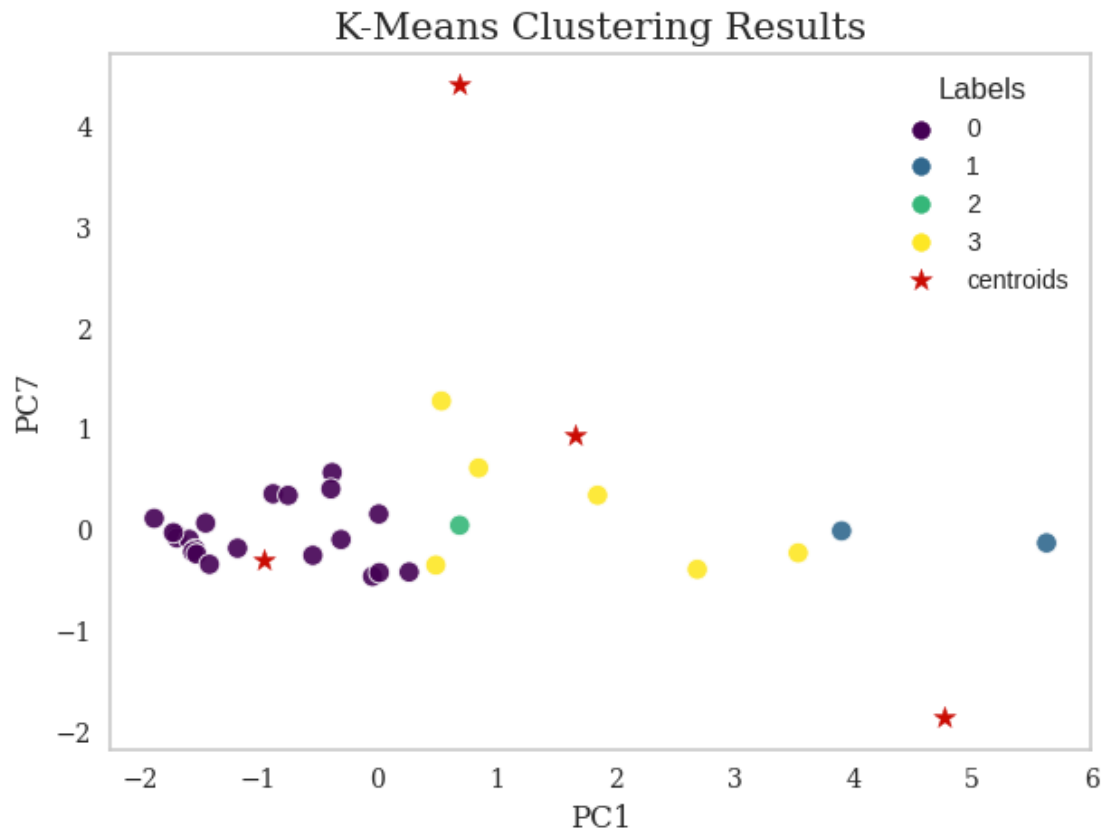
```
[ ]
    # feature scaling
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X)
```

```
    # applying Principle Component Analysis (PCA)
    pca = PCA(n_components=7)
    X_pca = pca.fit_transform(X_scaled)
    df_pca = pd.DataFrame(X_pca, columns=['PC1', 'PC2', 'PC3', 'PC4', 'PC5', 'PC6', 'PC7'])
    df_pca.head()
```

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
0	-0.379827	0.506616	0.758411	-0.789536	-1.131726	-0.317878	0.563965
1	-0.877028	-1.194175	2.121372	-0.267373	1.012621	-0.947004	0.353495
2	-0.750722	-0.729245	1.426220	-0.308906	-0.051495	-0.532041	0.338402
3	0.266808	-1.788843	2.061258	0.060068	0.291974	0.077175	-0.421508
4	0.691665	4.406813	2.026130	3.004043	0.416916	0.669018	0.041862

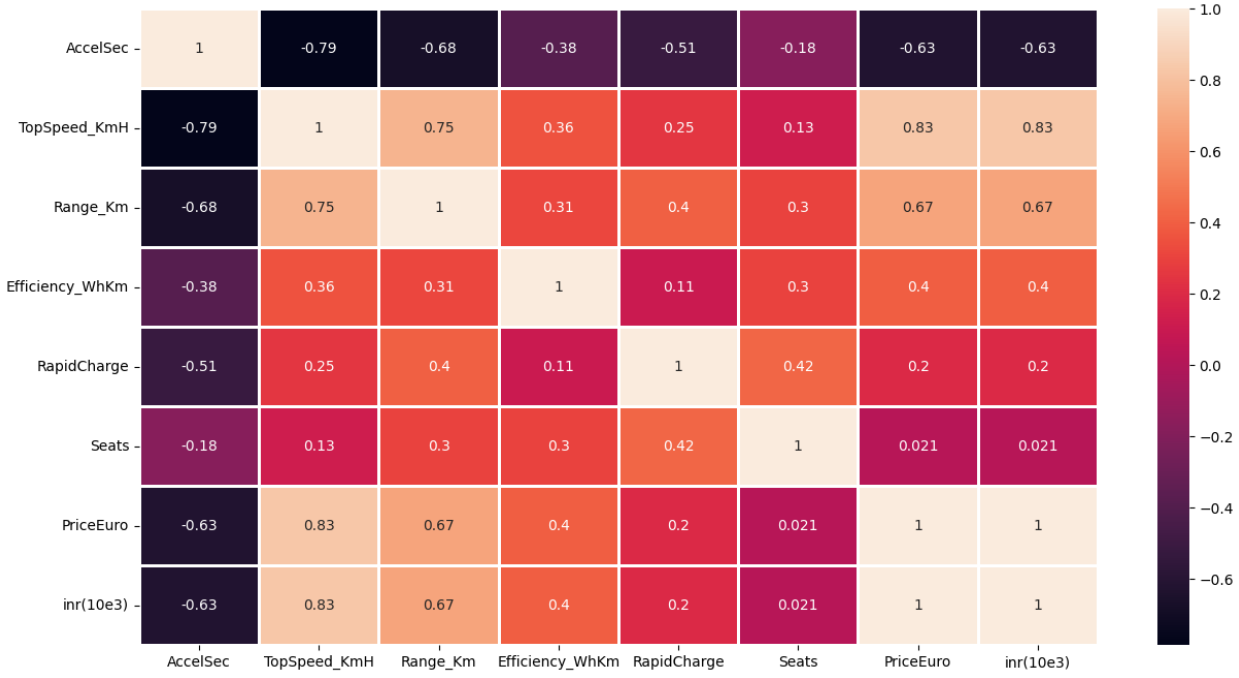



Calinski Harabasz Score Elbow for KMeans Clustering
calinski harabasz score
fit time (seconds)
k
elbow at $k = 2$, $score = 13.184$


K-Means Clustering Results
Labels
0
1
2
3
centroids
PC7
PC1

ELECTRIC CAR DATA

The dataset consists of electric car data with 107 data

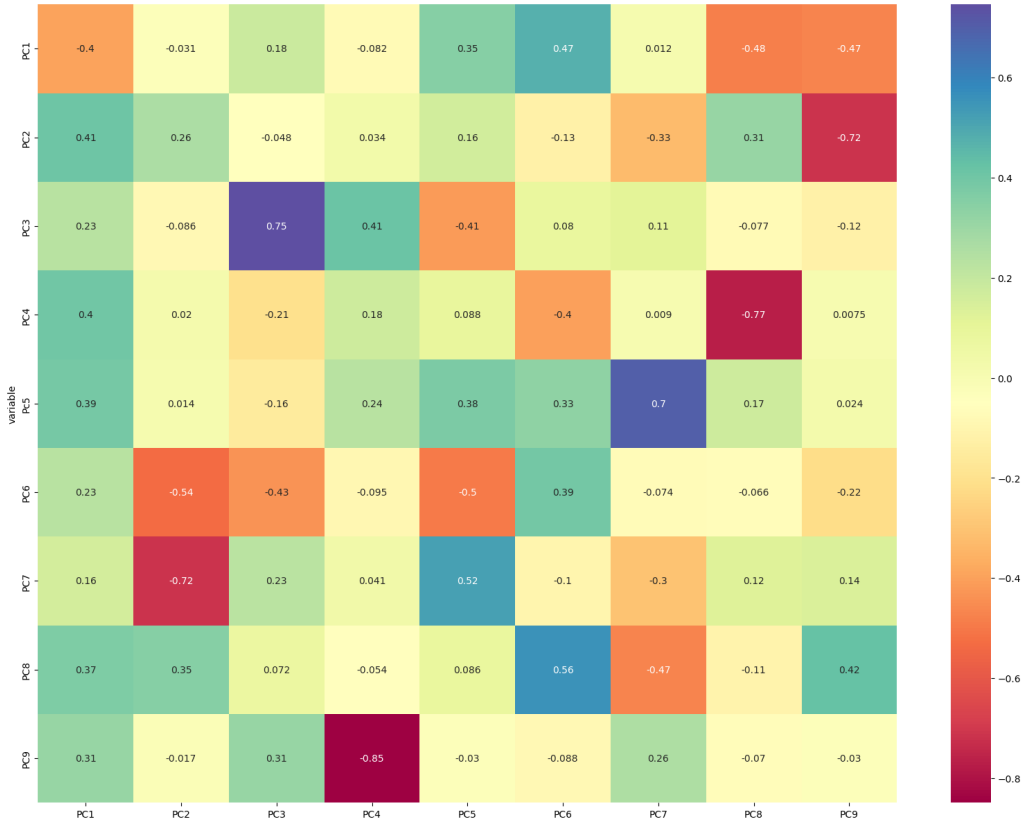



```
[44] df['FastCharge_KmH'] = df['FastCharge_KmH'].replace('-', 0)

df['PowerTrain'].replace(to_replace=['RWD','AWD','FWD'],value=[0, 2,1],inplace=True)
#X=df[['AccelSec','Range_Km','TopSpeed_KmH','Efficiency_WhKm', 'RapidCharge','PowerTrain']]
y=df['PriceEuro']
# selecting features for building a model
X = df[['AccelSec','TopSpeed_KmH','Efficiency_WhKm','FastCharge_KmH', 'Range_Km', 'RapidCharge', 'Seats', 'PriceEuro','PowerTrain']]

# feature scaling
from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)

[47] from sklearn.decomposition import PCA
pca = PCA(n_components=9)
t = pca.fit_transform(X_scaled)
data2 = pd.DataFrame(t, columns=['PC1', 'PC2','PC3','PC4','Pc5','PC6', 'PC7','PC8','PC9'])
data2
```

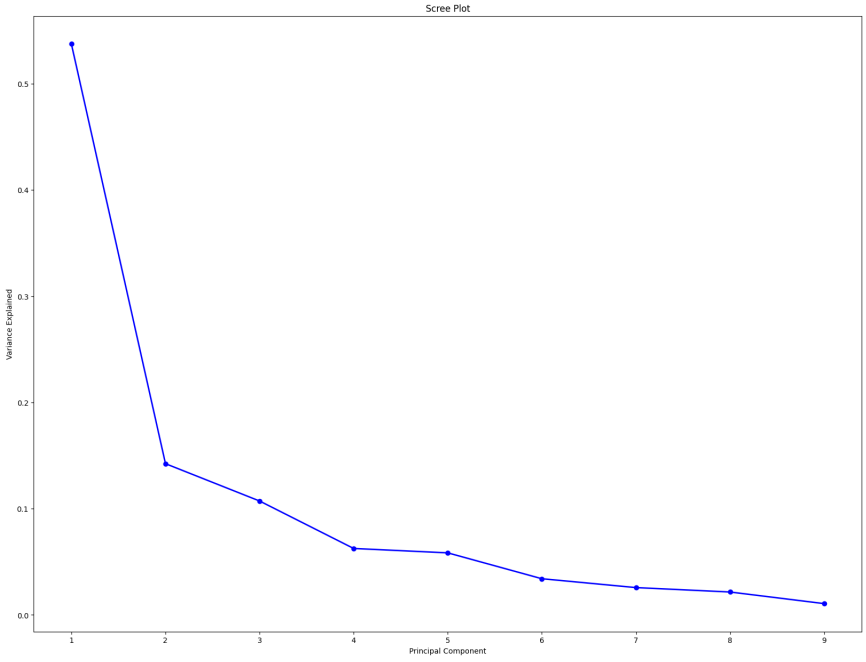
	PC1	PC2	PC3	PC4	Pc5	PC6	PC7	PC8	PC9
0	7.774740	3.820347	-1.066919	0.902984	1.782470	2.026297	0.196807	0.877754	-1.168988
1	2.990394	0.759728	-0.369588	-0.408030	0.334161	-0.134119	-0.074877	0.994185	-0.090111
2	3.204563	-1.152740	1.104266	-0.114819	1.039833	-0.187842	-1.097982	1.242346	0.344728
3	3.322837	0.813423	-0.722346	-0.296509	0.805604	0.202476	0.456631	0.760691	0.200850
4	3.955425	2.417234	0.371327	-0.168104	-0.872066	0.748932	-1.435167	-0.675402	0.582604
...	...	...	...	...	...	...	...	...	...
98	-2.744374	0.038335	-0.289711	-0.831001	-0.593404	0.671974	0.201753	-0.382816	-0.487717
99	-4.742341	2.564124	1.526276	0.400489	1.588351	-0.878552	-0.203239	0.513826	0.432071
100	-4.845600	2.465435	1.879882	0.501194	1.300045	-0.953155	-0.275843	0.390910	0.407617
101	-2.053409	-2.744705	1.548710	-0.255738	1.087578	0.748166	-0.906018	-0.434977	-0.136518
102	-5.013260	1.558293	3.555516	-0.444261	3.032839	0.811277	-0.083230	-1.112883	-0.836732

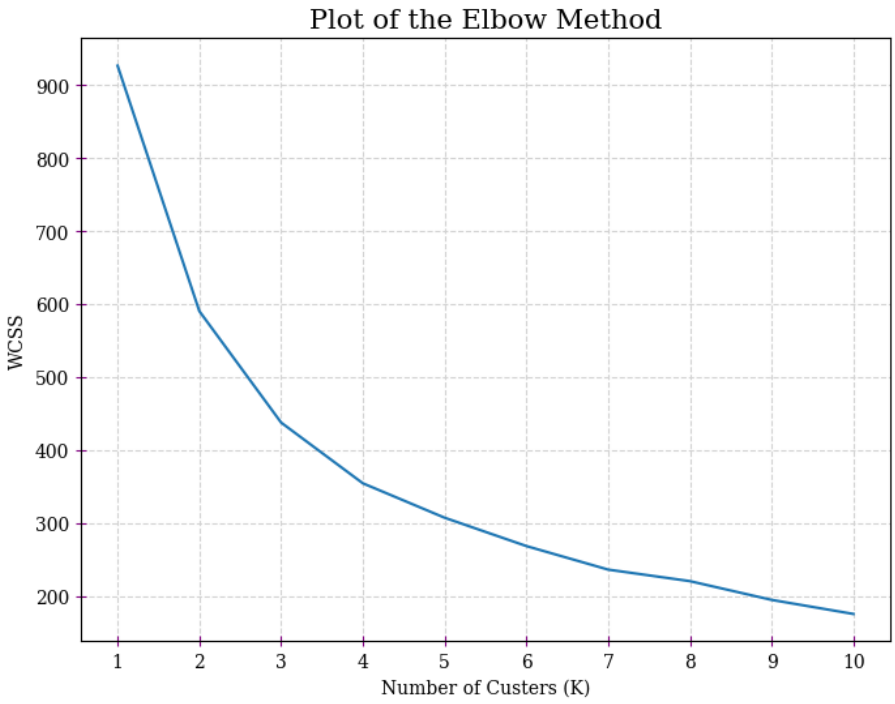
103 rows × 9 columns

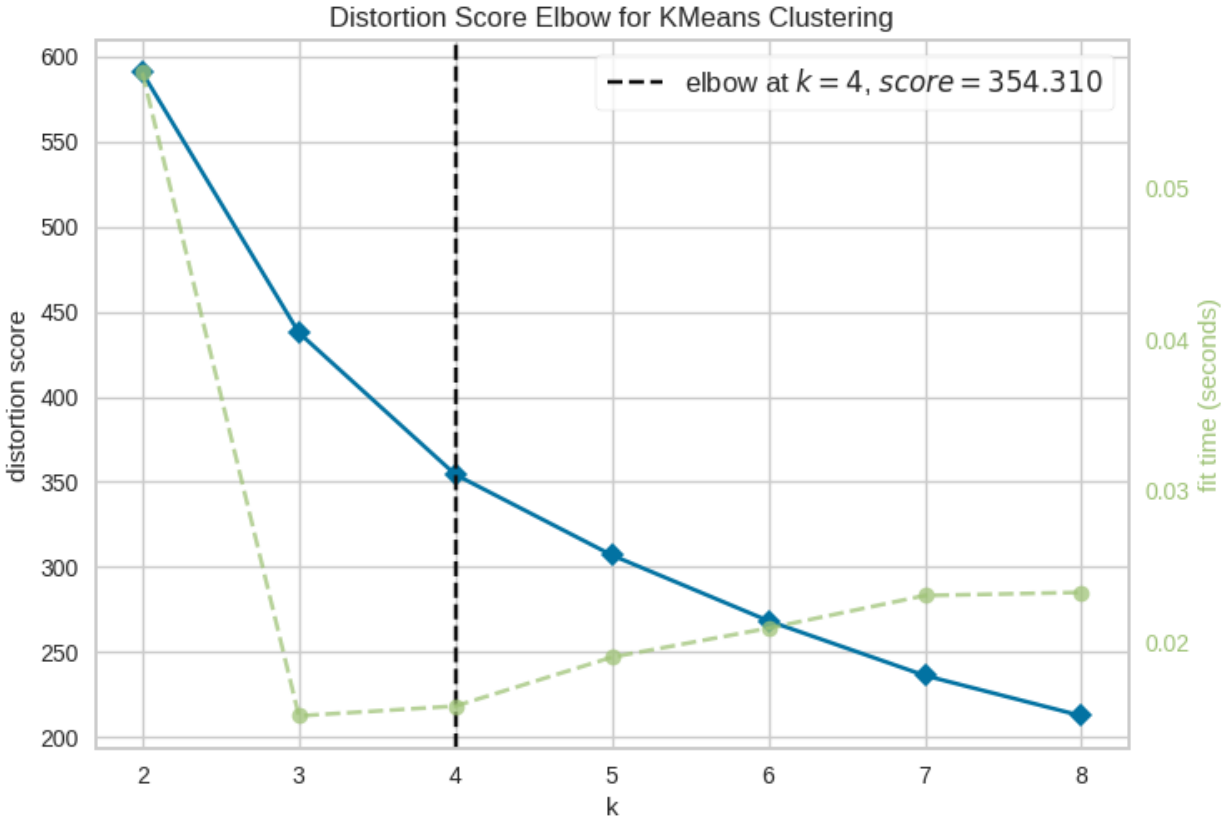
✓ 0s completed at 7:10 PM

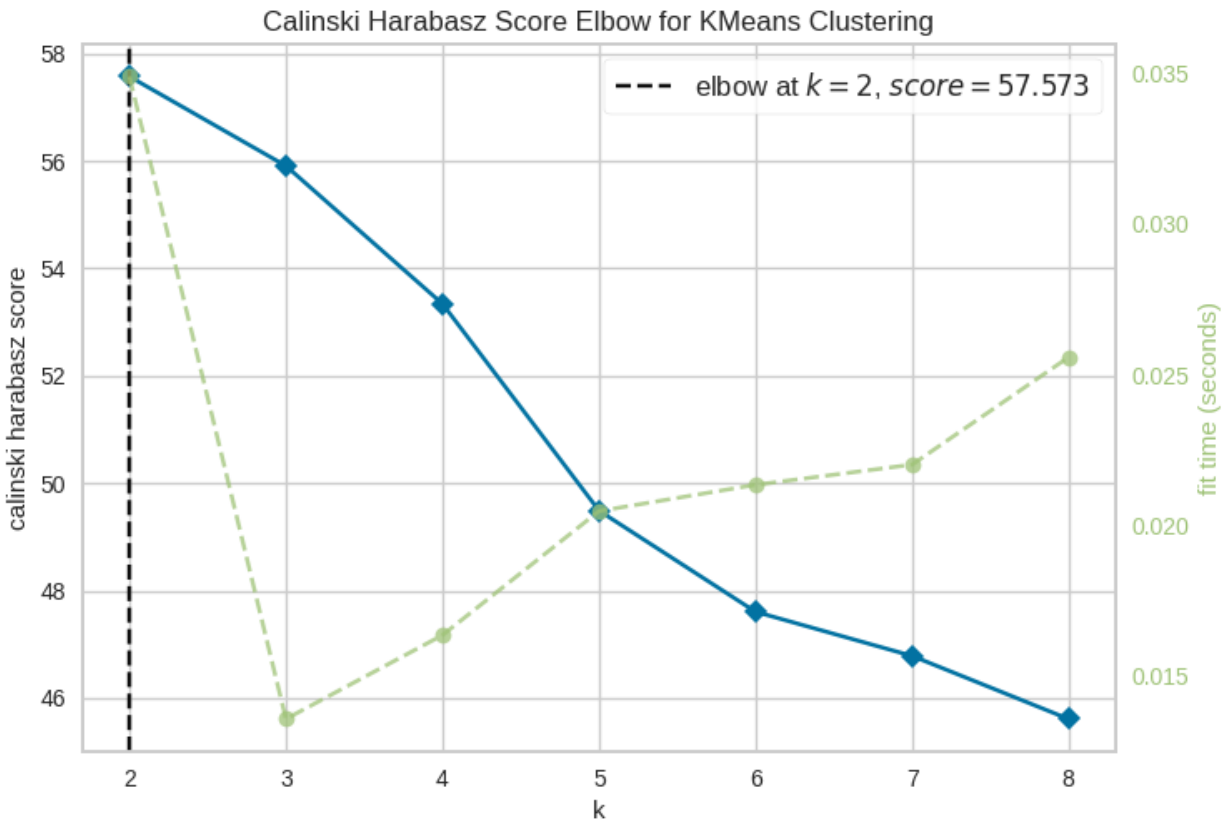


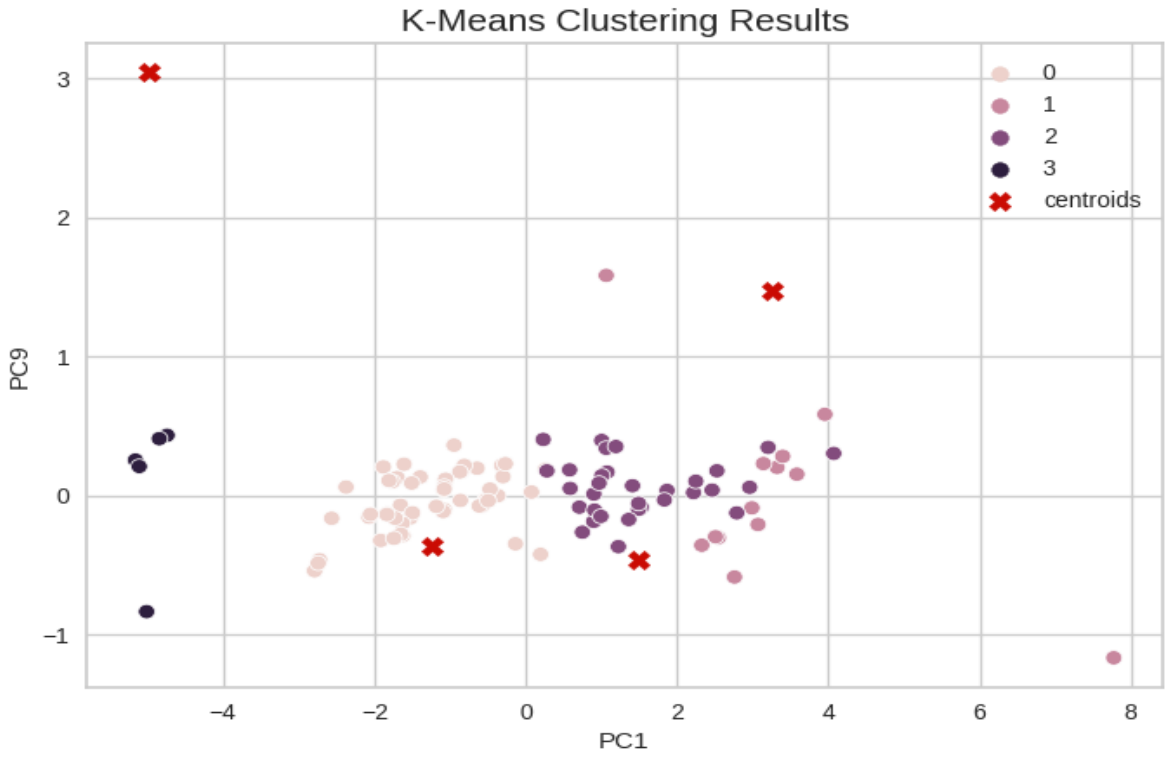

```
#Correlation matrix plot for loadings
```




Plot of the Elbow Method
WCSS
Number of Custers (K)
900
800
700
600
500
400
300
200
1
2
3
4
5
6
7
8
9
10


Distortion Score Elbow for KMeans Clustering
elbow at $k = 4$, $score = 354.310$
distortion score
fit time (seconds)
k
600
550
500
450
400
350
300
250
200
0.05
0.04
0.03
0.02
2
3
4
5
6
7
8


Calinski Harabasz Score Elbow for KMeans Clustering
elbow at $k = 2$, $score = 57.573$
calinski harabasz score
fit time (seconds)
k
58
56
54
52
50
48
46
0.035
0.030
0.025
0.020
0.015
2
3
4
5
6
7
8


K-Means Clustering Results
0
1
2
3
centroids
PC9
PC1
3
2
1
0
−1
−4
−2
0
2
4
6
8

FINAL ANALYSIS OF WORK

In first dataset , based on Exploratory Data Analysis we can see over all north indian states like maharashtra , Uttar pradesh,Gujarat are those use more electric vehicles but in south indian states are very less popular compared to north india and not does not states like manipur , Meghalaya, Nagaland. In terms of business model we have to expand and make our EV market popular in south india and North-eastern states of india to achieve more profit in this field. In model building we used k-means and PCA ,we used this algorithm because of performance popularity in market segmentation. The visualization of k-means are easier to interpret , in future with more data added to this we can get better results as the model was built in very less data

In second dataset, we see more behavioral features of car being compared in Exploratory Data Analysis like top speed achieved by brand, plug type,body type, categorizing the powertrain and Highest range achieved by brand and many more behavioral feature visualization which helps us to go through easily about dataset .In model building we used K-means and PCA , this algorithm can work very well in market segmentation and this dataset has little more features than the previous but both come under types of segmentation criteria . The model building is done with quite basic approaches inspired from various examples with some more added data. In the future we can build a better model and easily segment it .

In our both datasets we have very vital information about electric vehicles but we have less data, with additional time and budget, the market segmentation project can be substantially enhanced by investing in more comprehensive and granular data sets. This could involve acquiring data from diverse sources such as consumer behavior analytics, social media insights, and third-party databases to gain a holistic understanding of customer preferences and demographics.

Additionally, conducting in-depth surveys or interviews with a representative sample of the target audience can provide qualitative insights, complementing quantitative data. Advanced analytics tools and machine learning algorithms can be employed to uncover hidden patterns and correlations within the data, facilitating a more nuanced segmentation. Furthermore, the extended time frame allows for iterative testing and refinement of the segmentation model, ensuring its accuracy and relevance.

The EV Market in India is estimated at USD 5.61 billion in 2023, and is expected to reach USD 37.70 billion by 2028, growing at a CAGR of 46.38% during the forecast period (2023-2028). The top 4 features that can be used to create the most optimal market segments for electric vehicles are Driving Patterns and Range Requirements ,Charging Infrastructure Accessibility ,Pricing Sensitivity and Incentive Eligibility ,Technological Preferences and Features.

CODING LINK:

https://colab.research.google.com/drive/1dvE6rXMfEFr2AyETgkT50VQnFhIwaVgv?usp=sharing

https://colab.research.google.com/drive/1fszwmrdtbtVbFXQj1E6FwbSy7CNjWEcD?usp=sharing