EV MARKET SEGMENTATION

11th July, 2023

Sreeparvathy S Menon

INTRODUCTION

Electric vehicles have gained popularity in recent times of global warming and climate change with their environmentally friendly nature. With zero tailpipe emissions, customers are flocking into the new EV market looking to replace their outdated petrol diesel counterparts. EVs also promise lower running costs, low maintenance cost, convenience of charging at home and zero noise pollution. Government is also offering multiple tax and financial benefits.

National Electric Mobility Mission Plan (NEMMP) was created to encourage the adoption of EVs in India. Government has formulated a scheme FAME- Faster Adoption and Manufacturing of (Hybrid and) Electric Vehicles in India implementing Demand creation, Technology Platform, Pilot Project and Charging Infrastructure. Currently in its phase 2, it has been allocated a budget of Rs 10,000 crore. India is set to electrify automobiles by 2030.

Indian EV market is projected to grow to 113.99 billion in 2029 from 1.45 billion in 2021. Tata Motors dominates the space with 80 percent market share. With reasonable prices and easy availability of products Tata Motors attracts Indian consumers. Other major players include MG, Mahindra, Citroen and Hyundai.

PROBLEM STATEMENT

The aim of this project is to analyse the electric vehicle market in India and come up with feasible strategies to enter the market, targeting segments most likely to use electric vehicles.

First EV Vehicles available in the market is analysed and various data is explored thoroughly. To gain insights of Indian consumer behaviour, the details of customers of petrol diesel vehicles are analysed. This data is used to create market segments and determine which vehicles should be deployed to which segments. Lastly, the best location to enter the EV market is also determined.

DATA SETS USED

ElectricCarData_Clean_Me.csv

Sourced from Kaggle, it contains details about different electric cars available in the market. Details include 'Brand', 'Model', 'AccelSec', 'TopSpeed_KmH', 'Range_Km', 'Battery_Pack

Kwh', 'Efficiency_WhKm', 'FastCharge_KmH', 'RapidCharge', 'PowerTrain', 'PlugType', 'BodyStyle', 'Segment', 'Seats' and 'PriceEuro'.

Link to the dataset: https://www.kaggle.com/datasets/divyanshugupta95/cars-dataset-with-battery-pack-capacity

Indian automobile buying behaviour study 1.0.csv

Also sourced from Kaggle, it contains data about buying behaviour of customers of the automobile market of India. Note that this data is mostly for petrol and diesel vehicles as EV market is a relatively new market in India and not much data is available. Details include 'Age', 'Profession', 'Marrital Status', 'Education', 'No of Dependents', 'Personal loan', 'House Loan', 'Wife Working', 'Salary', 'Wife Salary', 'Total Salary', 'Make' and 'Price'.

Link to the dataset: https://www.kaggle.com/datasets/karivedha/indian-consumers-cars-purchasing-behaviour

Location_data.csv

Contains data like 'State Name', 'Total Electric Vehicle', 'Total Non-Electric Vehicle', 'Total' and 'No. of Operational PCS' from different states and union territories of India. Data is taken from different press releases by Ministry of Heavy Industries Government of India. Andhra Pradesh, Madhya Pradesh, Telangana, and Lakshadweep weren't available in Vahan4 and hence their data isn't included.

Link to data of 'Total Electric Vehicle', 'Total Non-Electric Vehicle' and 'Total': https://pib.gov.in/PressReleasePage.aspx?PRID=1842704

Link to data of Operational PCS (Public Charging Stations): https://pib.gov.in/PressReleasePage.aspx?PRID=1910392

MARKET SEGMENTATION

Market segmentation is a marketing strategy that involves dividing a broad target market into smaller, more defined segments based on commonalities, ranging from age, gender or location to priorities, values and behaviour. Market segmentation forces organisations to take look at where they stand, and where they want to be in future. It enables businesses to tailor their marketing efforts and offerings to different groups of customers, allowing for more effective and efficient marketing campaigns. Market segment are decided after taking multiple criteria into account. Common segmentation criteria are geographic, sociodemographic, psychographic and behavioural.

In geographic segmentation, each consumer can be assigned to a geographic unit and marketing campaigns can be targeted easily. People from the same geographical area may not share other characteristics relevant to the market which is a major shortcoming of geographic segmentation.

Socio-demographic segmentation criteria include age, gender, income and education. Members in a certain demography can be easily identified, but in many instances, the socio-demographic criterion doesn't provide sufficient market insight for proper segmentation.

Psychographic criteria refer to peoples' beliefs, interests, preferences, aspirations, or benefits sought. Psychographic criteria are more complex and require multiple segmentation variables to provide valuable insights into the market. The psychographic approach is more reflective of consumer behaviour.

In behavioural segmentation, the very behaviour of interest is used as the basis of segment extraction and if based on actual behaviour rather than reported one, they are found to be superior compared to other segmentation criteria.

Before market segmentation it is important to analyse the types of EVs available and their features

ELECTRIC CARS AVAILABLE IN THE MARKET

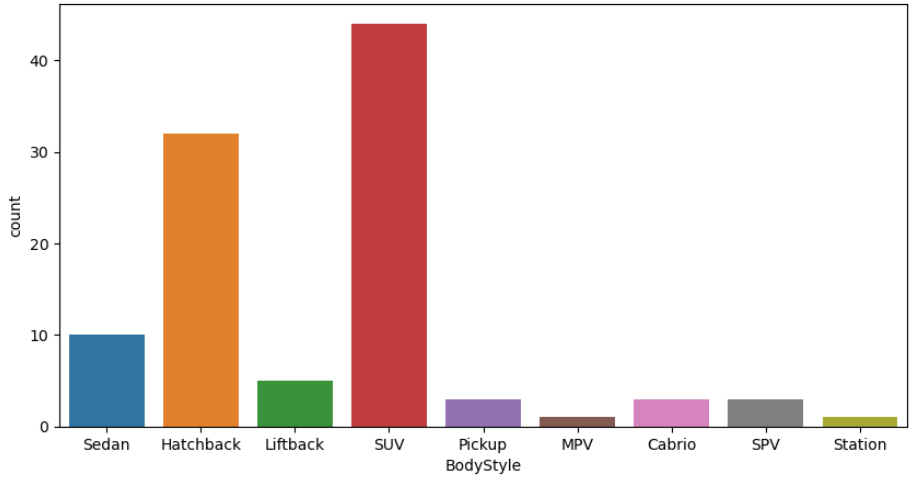
First 5 data points

	Brand	Model	AccelSec	TopSpeed_KmH	Range_Km	Battery_Pack Kwh	Efficiency_WhKm	FastCharge_KmH	RapidCharge	PowerTrain	PlugType	BodyStyle	Segment	Seats	PriceEuro
0	Tesla	Model 3 Long Range Dual Motor	4.6	233	460	70.0	161	940	Yes	AWD	Type 2 CCS	Sedan	D	5	55480
1	Volkswagen	ID.3 Pure	10.0	160	270	45.0	167	250	Yes	RWD	Type 2 CCS	Hatchback	C	5	30000
2	Polestar	2	4.7	210	400	75.0	181	620	Yes	AWD	Type 2 CCS	Liftback	D	5	56440
3	BMW	iX3	6.8	180	360	74.0	206	560	Yes	RWD	Type 2 CCS	SUV	D	5	68040
4	Honda	e	9.5	145	170	28.5	168	190	Yes	RWD	Type 2 CCS	Hatchback	B	4	32997

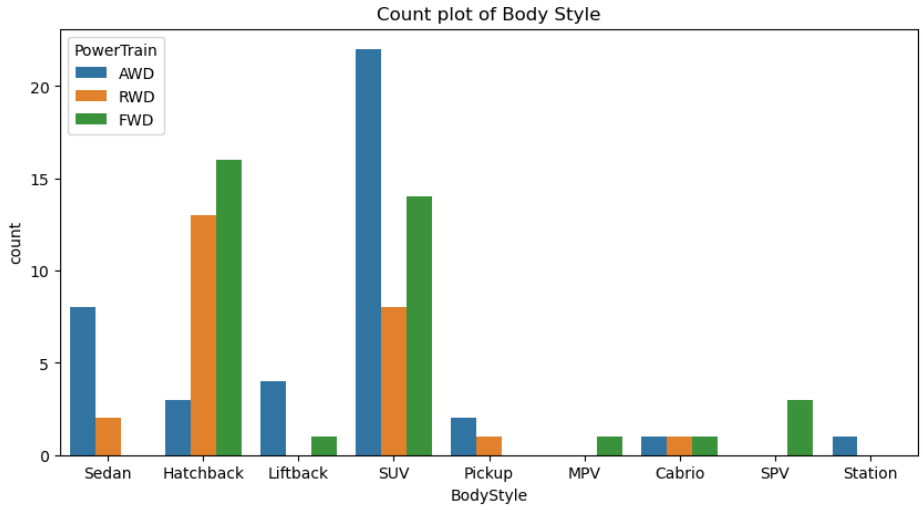
Describing the data.

	AccelSec	TopSpeed_KmH	Range_Km	Battery_Pack Kwh	Efficiency_WhKm	FastCharge_KmH	Seats	PriceEuro
count	102.000000	102.000000	102.000000	102.000000	102.000000	102.000000	102.000000	102.000000
mean	7.391176	179.313725	338.627451	65.415686	189.303922	435.686275	4.882353	55997.588235
std	3.031913	43.771228	126.700623	29.955782	29.679072	220.447384	0.799680	34250.724403
min	2.100000	123.000000	95.000000	16.700000	104.000000	0.000000	2.000000	20129.000000
25%	5.100000	150.000000	250.000000	43.125000	168.000000	260.000000	5.000000	34414.750000
50%	7.300000	160.000000	340.000000	64.350000	180.500000	440.000000	5.000000	45000.000000
75%	9.000000	200.000000	400.000000	83.700000	204.500000	557.500000	5.000000	65000.000000
max	22.400000	410.000000	970.000000	200.000000	273.000000	940.000000	7.000000	215000.000000

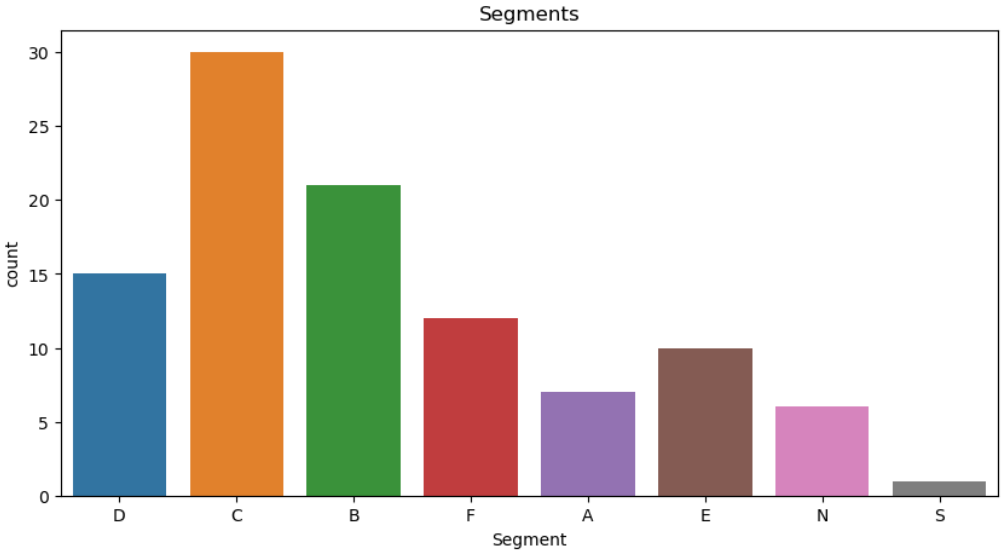
Body styles available



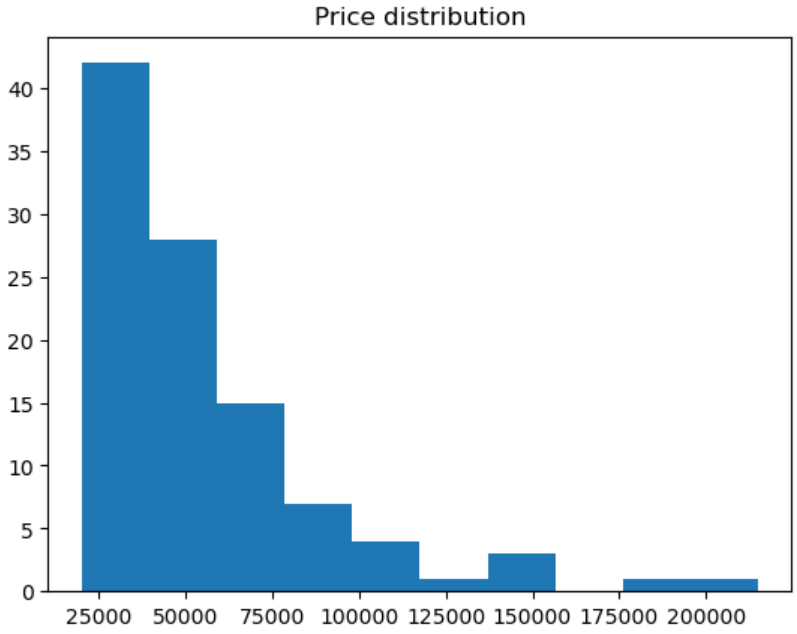
EVs mostly come as SUVs, Hatchbacks and Sedan



Most cars belong to segment C.

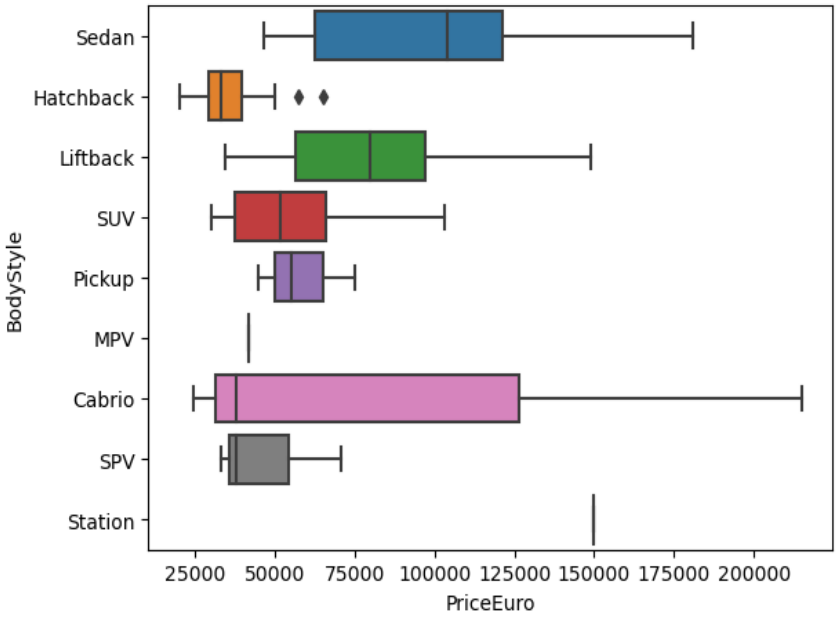


India is a very price and value sensitive market. In the data given Tesla Roadster that comes with 200kWh battery pack, 970 km range and 410km/h top speed is the most expensive one which cost 215k euros. Cheapest car is SEAT Mii Electric 130km/h top speed, 195km range and 32.3kW/h battery pack costing 20k euros.

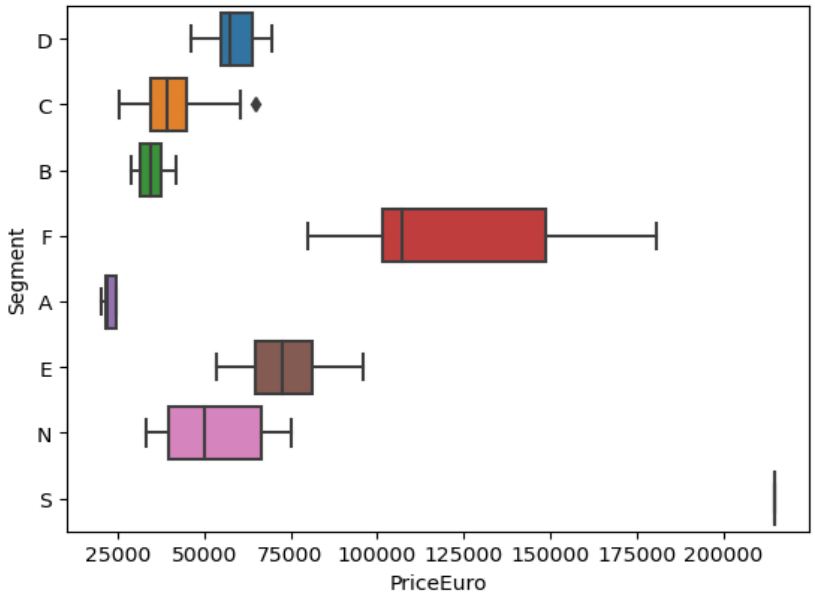


Body Style vs Price

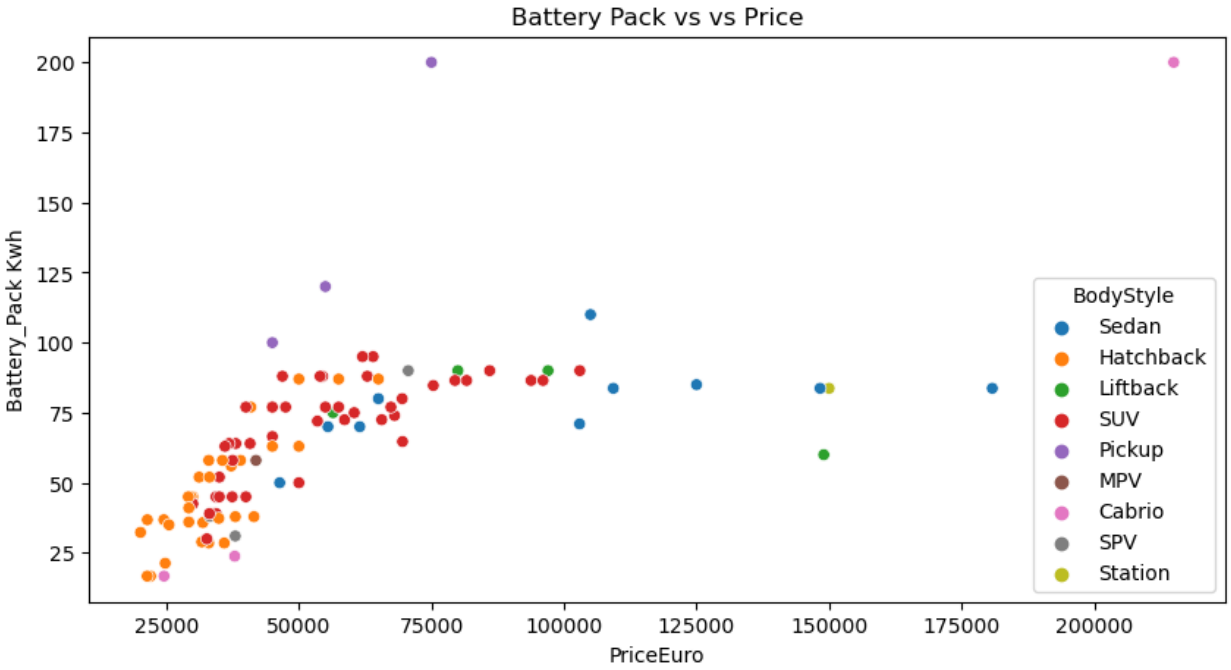
Sedan is on the expensive side, while SUVs are more affordable, and hatchbacks are the cheapest in the market.



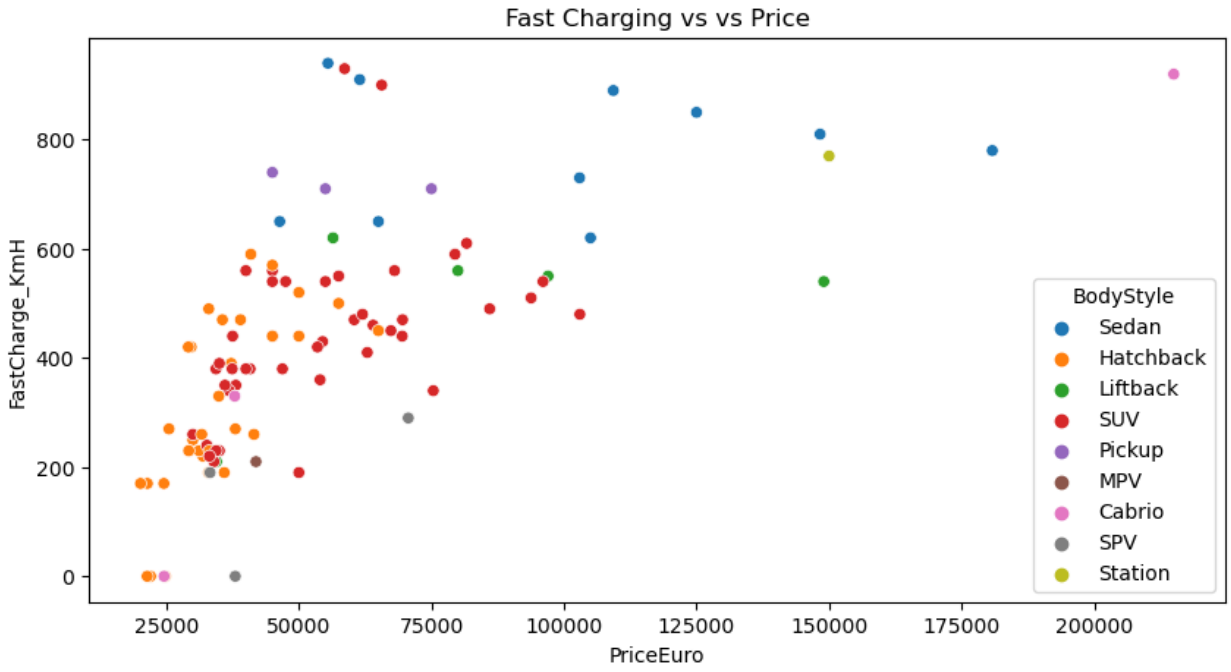
Segment vs Price



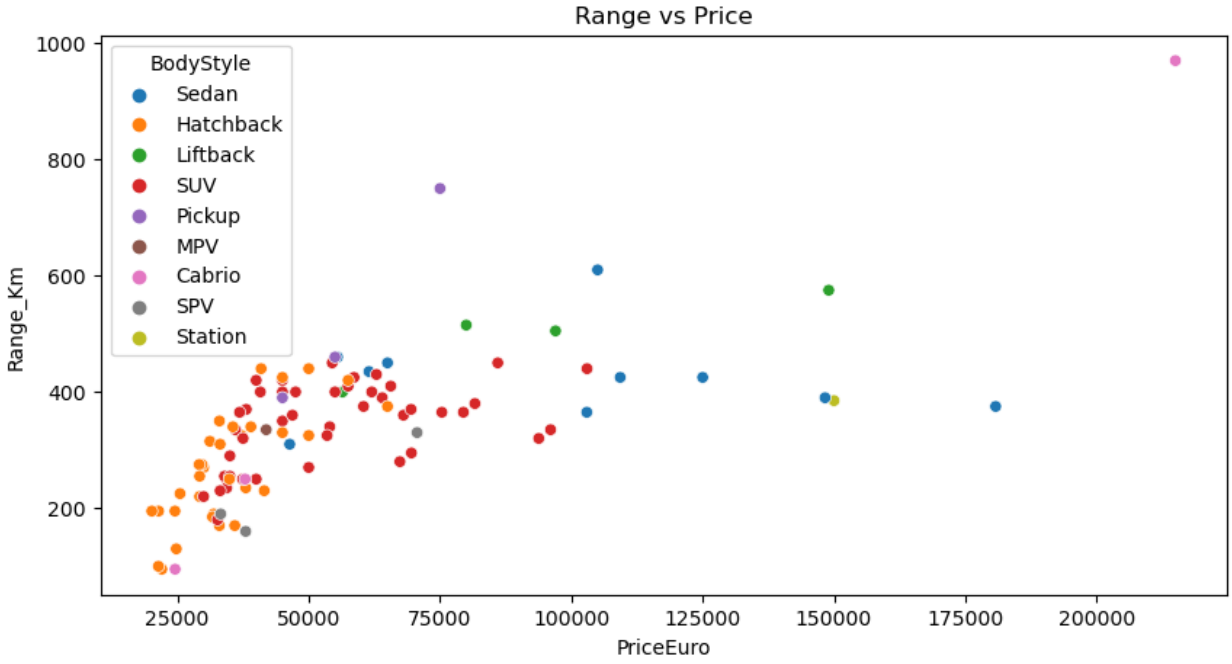
SUVs and pickups are the more efficient ones with reasonable prices.



Battery packs don’t go much higher than 100kWh



Most Sedans come with fast charging.



Both sedans and SUVs provide good Range.

MARKET SEGMENTATION USING DATA OF BEHAVIOUR OF CUSTOMERS OF AUTOMOBILE MARKET (INDIA)

	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	House Loan	Wife Working	Salary	Wife Salary	Total Salary	Make	Price
0	27	Salaried	Single	Post Graduate	0	Yes	No	No	800000	0	800000	i20	800000
1	35	Salaried	Married	Post Graduate	2	Yes	Yes	Yes	1400000	600000	2000000	Ciaz	1000000
2	45	Business	Married	Graduate	4	Yes	Yes	No	1800000	0	1800000	Duster	1200000
3	41	Business	Married	Post Graduate	3	No	No	Yes	1600000	600000	2200000	City	1200000
4	31	Salaried	Married	Post Graduate	2	Yes	No	Yes	1800000	800000	2600000	SUV	1600000

As make is only given, we are converting different make into different body styles available.

	Make
SUV	19
Baleno	19
Creata	14
i20	12
Ciaz	12
City	10
Duster	7
Verna	4
Luxury	2

Before conversion

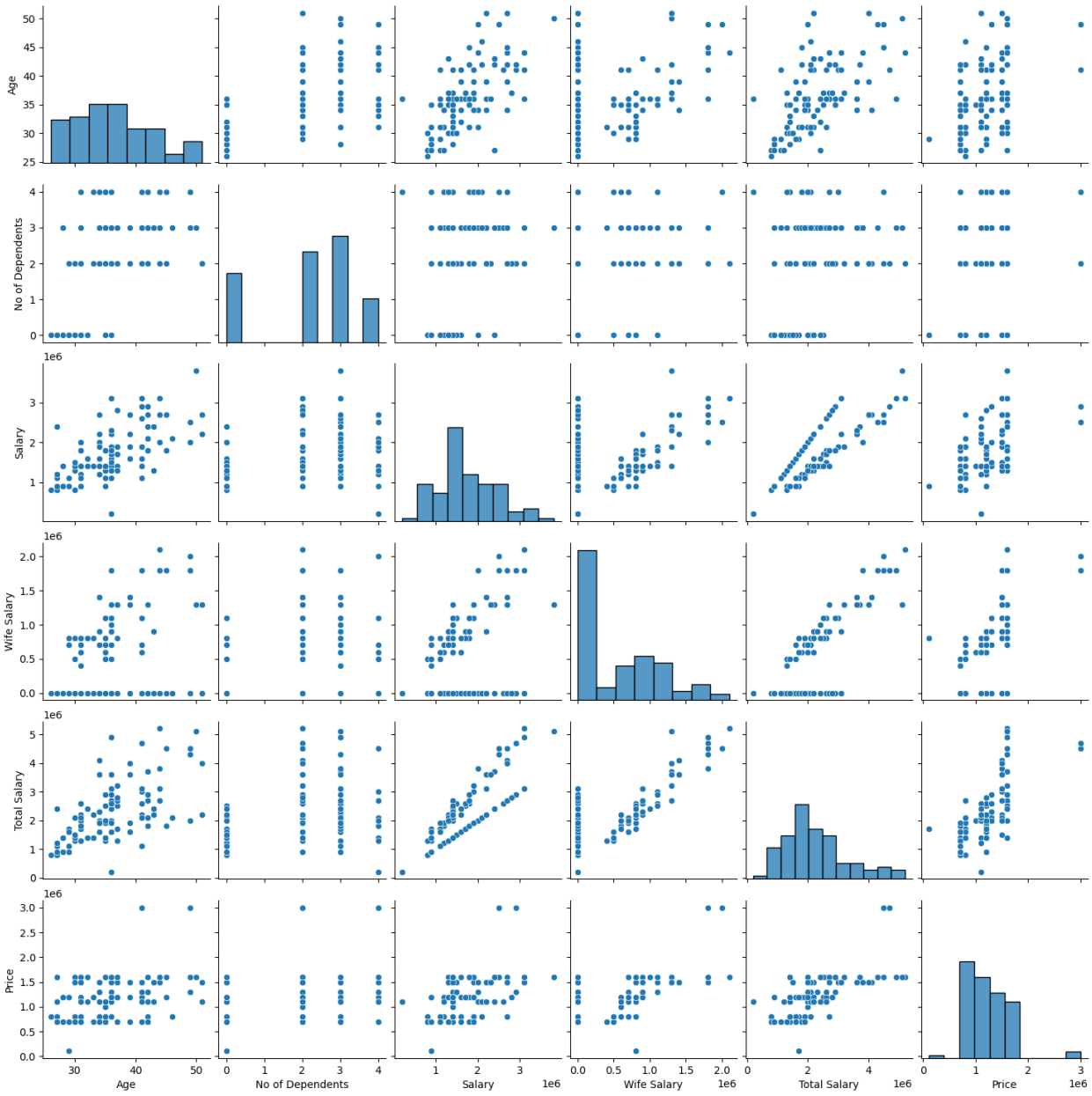
	Make
SUV	40
Hatchback	31
Sedan	26
Luxury	2

After Conversion

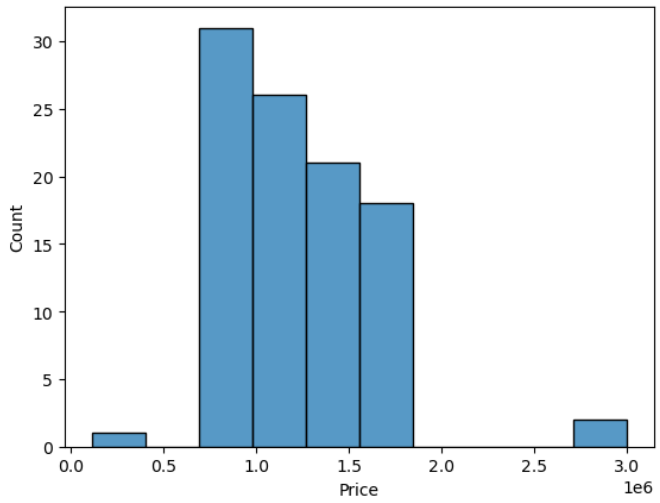
Describing Data

	Age	No of Dependents	Salary	Wife Salary	Total Salary	Price
count	99.000000	99.000000	9.900000e+01	9.900000e+01	9.900000e+01	9.900000e+01
mean	36.313131	2.181818	1.736364e+06	5.343434e+05	2.270707e+06	1.194040e+06
std	6.246054	1.335265	6.736217e+05	6.054450e+05	1.050777e+06	4.376955e+05
min	26.000000	0.000000	2.000000e+05	0.000000e+00	2.000000e+05	1.100000e+05
25%	31.000000	2.000000	1.300000e+06	0.000000e+00	1.550000e+06	8.000000e+05
50%	36.000000	2.000000	1.600000e+06	5.000000e+05	2.100000e+06	1.200000e+06
75%	41.000000	3.000000	2.200000e+06	9.000000e+05	2.700000e+06	1.500000e+06
max	51.000000	4.000000	3.800000e+06	2.100000e+06	5.200000e+06	3.000000e+06

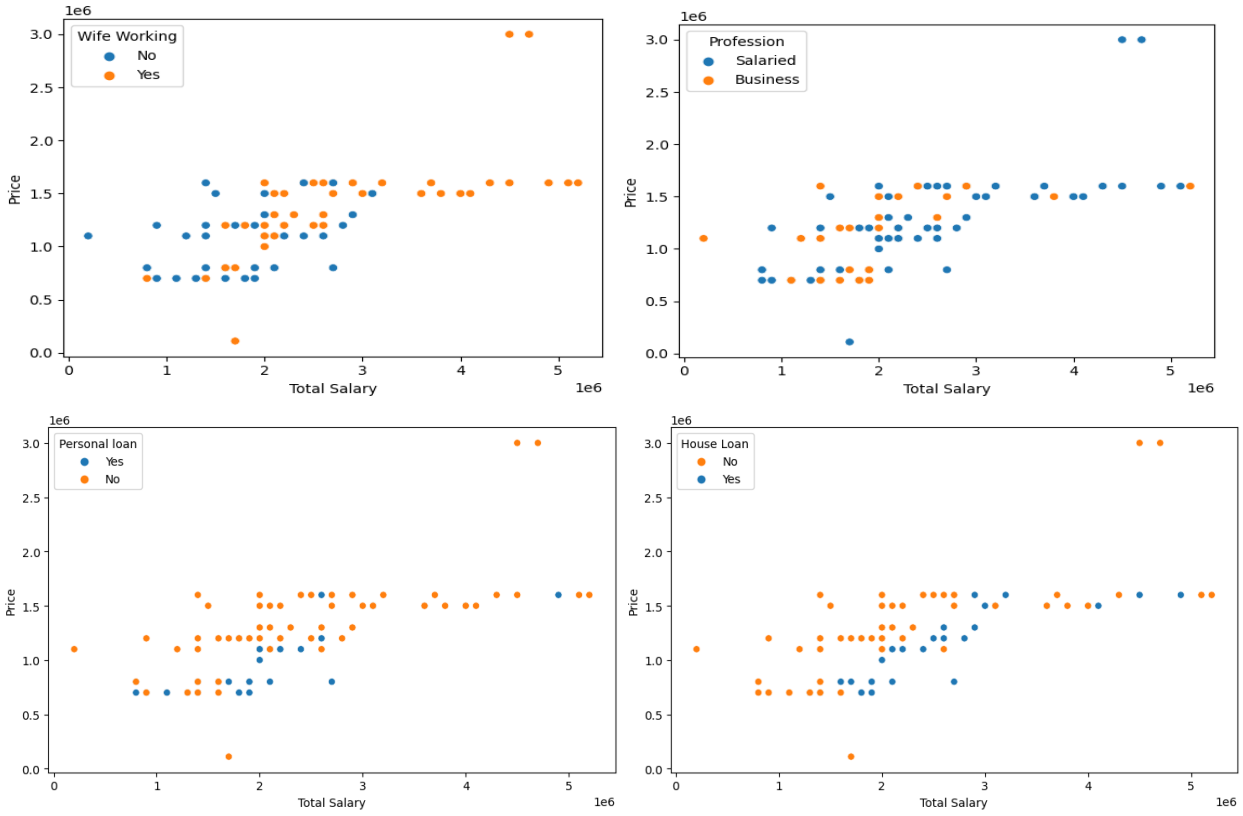
Pair Plot



Price Distribution



Price vs Salary



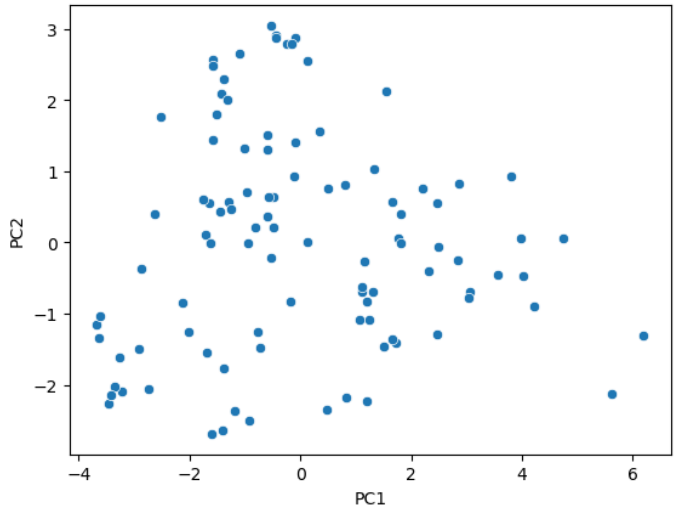
With working wife, an increase in total salary is seen. Professional status doesn't matter as much as the total salary. Price is proportional to Total Salary, but after 30 lakh of salary, price increase isn't seen.

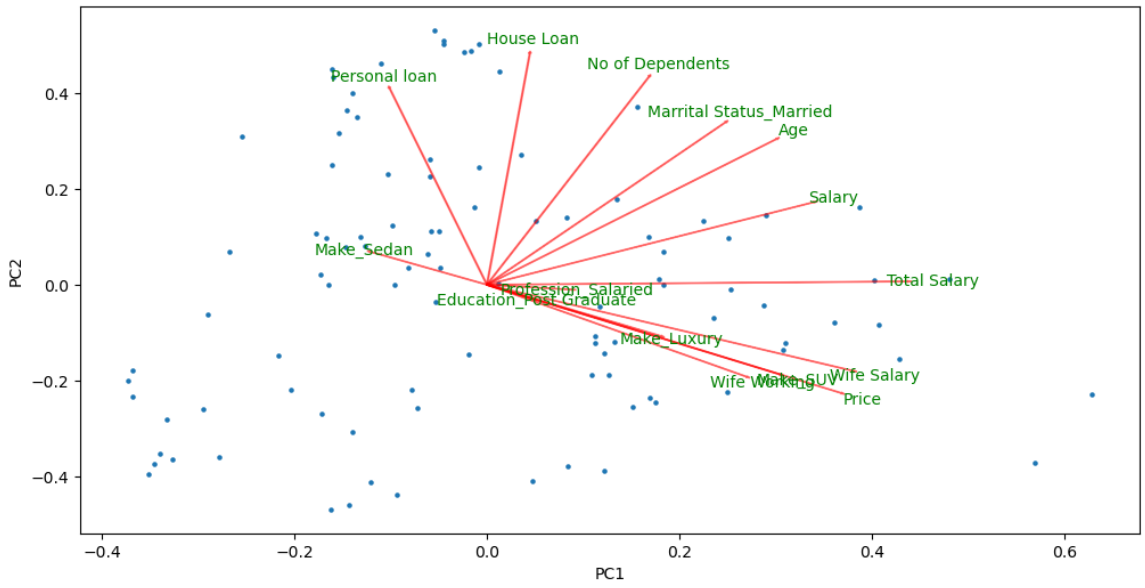
Customers buying the vehicles with higher retail prices do not have many loans.

PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA is a dimensionality reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set. Principal component analysis can be used for feature extraction also. First categorical data is converted to discrete and data is scaled. There are 15 features and principal component analysis is done. The cumulative variance of each of the components are 0.313, 0.467, 0.56, 0.649, 0.732, 0.804, 0.862, 0.908, 0.934, 0.96, 0.977, 0.993, 0.998, 1., 1.

First two components explain 46.7% of the data.





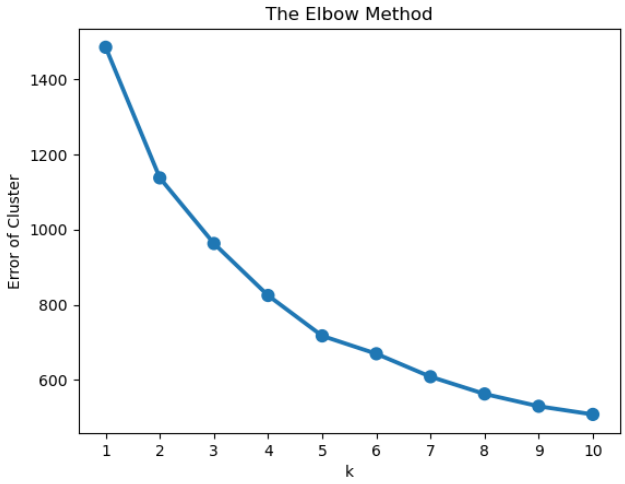
CLUSTERING

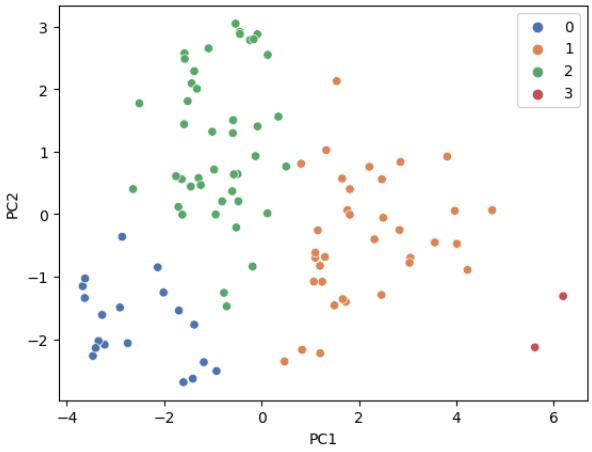
In this segmentation, market segments are being decided by clustering. Clustering is an unsupervised machine learning technique that divides the population into several clusters. The data points in a cluster are similar to each other, and data points in different clusters are dissimilar.

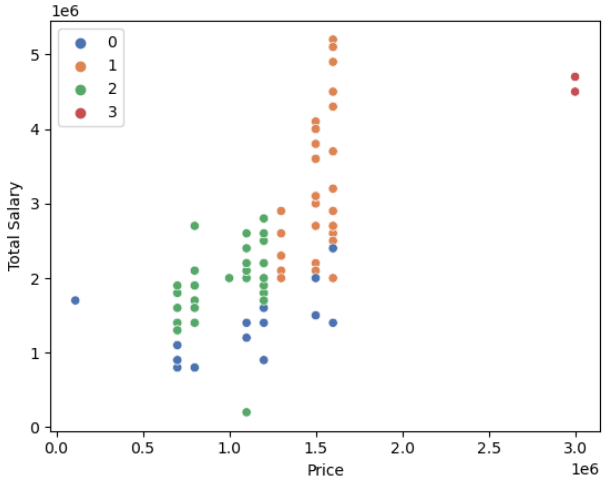
K-MEANS CLUSTERING

K-Means is a popular clustering algorithm. K-Means clustering algorithm aims to partition n observations into k clusters in which each observation belongs to the cluster with the nearest mean. K-Means first find k-centroid points and every point in the dataset will be assigned to either of the k-sets having minimum Euclidean distance.

K-means is limited to numerical data and there is categorical data in the dataset. First, we attempt to segment the market using k-means clustering after converting categorical data to discrete data and scaling it. The number of clusters is found using the elbow method. The number of clusters is chosen to be 4.



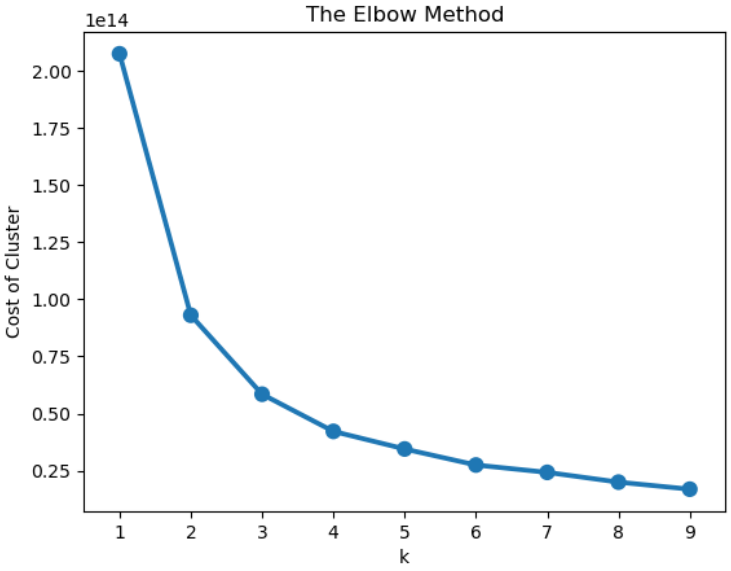


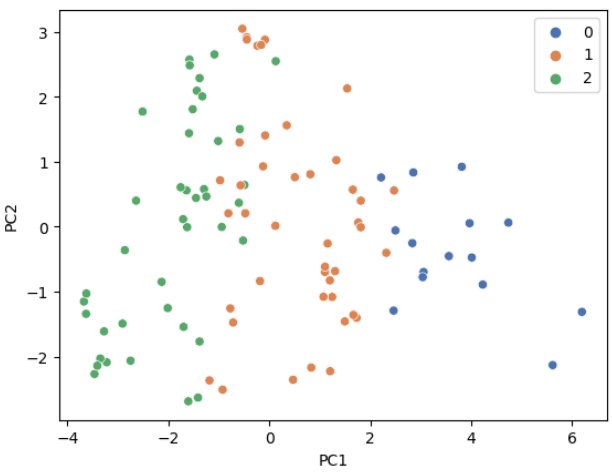


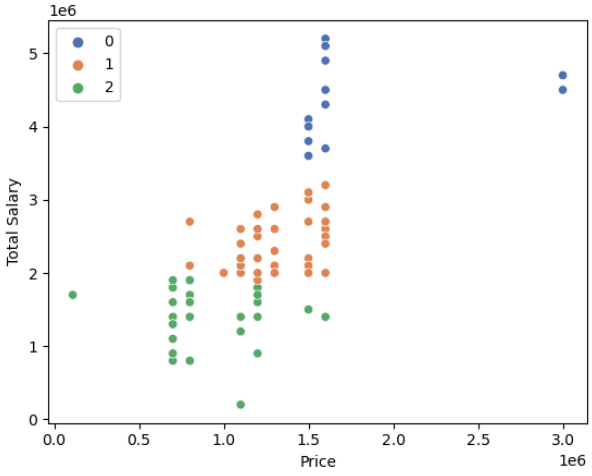
K-PROTOTYPE CLUSTERING

The k-Means algorithm is not applicable to categorical data, as categorical variables are discrete and do not have any natural origin. So, computing Euclidean distance for such a space is not meaningful. The K-Modes algorithm modifies the standard K-Means process for clustering categorical data by replacing the notion of distances with dissimilarities. That means instead of measuring distances, it counts the total number of mismatches between two observations the smaller the count the more similar two observations are. Also, instead of using the mean to find the centroid of the cluster it uses the mode. These changes allow the algorithm to converge to a local minimal for categorical datasets.

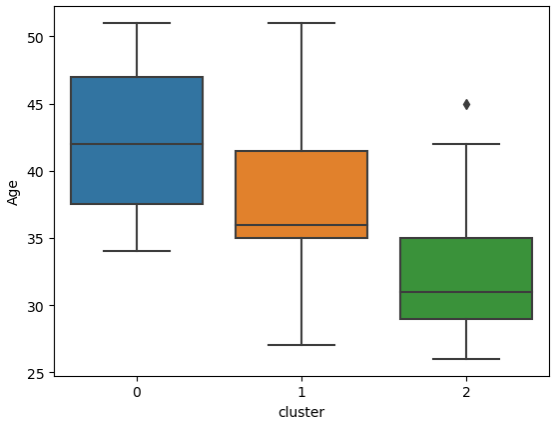
The k-Prototype algorithm is an extension to the k-Modes algorithm that combines the k-modes and k-means algorithms and is able to cluster mixed numerical and categorical variables. Number of clusters are determined to be 3 using the elbow method.

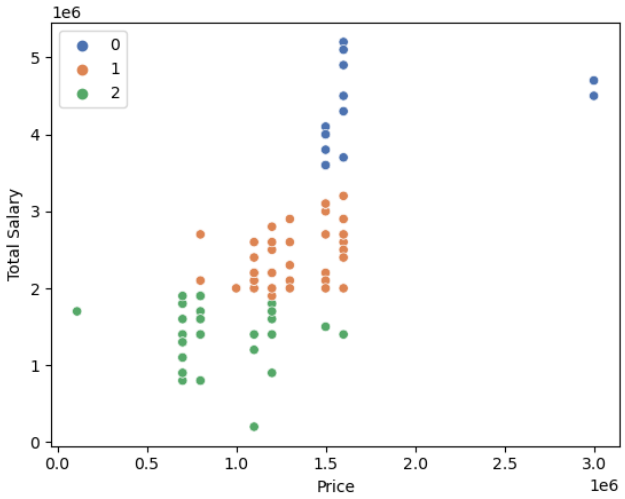




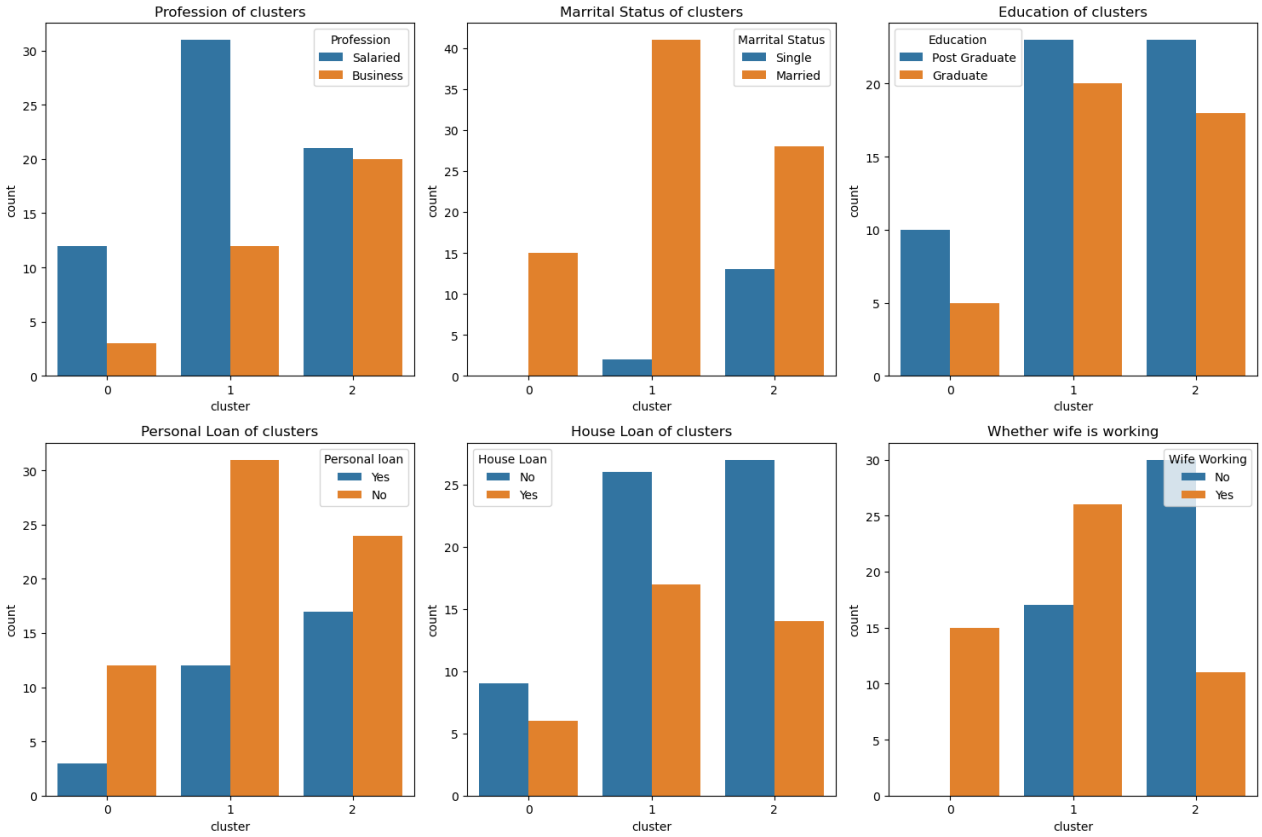


Analysing clusters from K-Prototype clustering

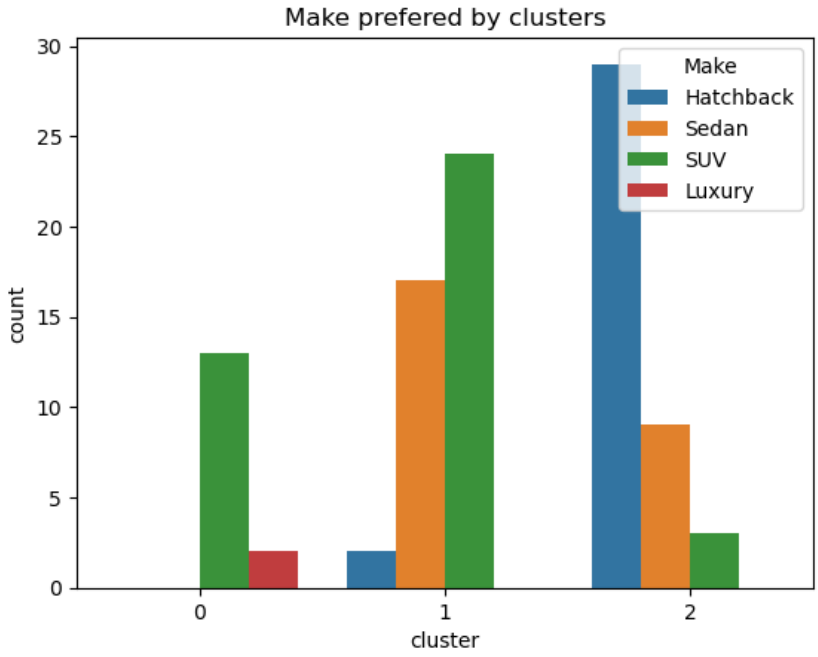




Cluster 0 has total salary above 30 lakh, cluster 1 has salary between 20 to 30 lakh and cluster 2 has salary below 20 lakhs. Cluster 0 is the oldest and cluster 2 is youngest. It also shows that only people in late 30s are able to afford expensive cars.



People doing business is mostly in the younger generation, cluster 2. Youngest generation (cluster 2) aren't married and have a lot of personal loans.



Cluster 2 prefers hatchback while cluster 0 and 1 prefer SUVs. Sedan is bought by cluster 1, luxury cars are bought by cluster 0.

GEOGRAPHICAL SEGMENTATION

India has an underdeveloped charging infrastructure. The availability of this charging infrastructure is different at different states of India. Also, few states are adopting EVs faster while few are resisting changes. EVs also need separate maintenance centres and location of these need to be determined. Hence geographical segmentation is important.

Andhra Pradesh, Madhya Pradesh, Telangana, and Lakshadweep weren't included in Vahan4 and hence their data isn't available.

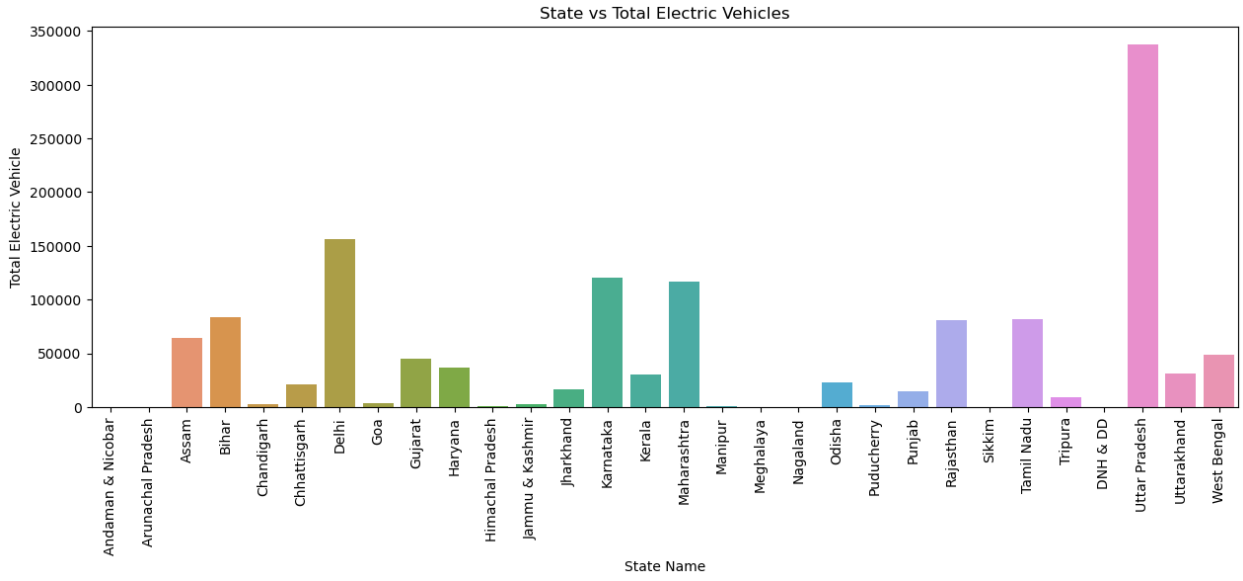
	State Name	Total Electric Vehicle	Total Non-Electric Vehicle	Total	No. of Operational PCS
0	Andaman & Nicobar	162	146945	147107	3
1	Arunachal Pradesh	20	252965	252985	9
2	Assam	64766	4677053	4741819	48
3	Bihar	83335	10407078	10490413	83
4	Chandigarh	2812	746881	749693	6

Describing data

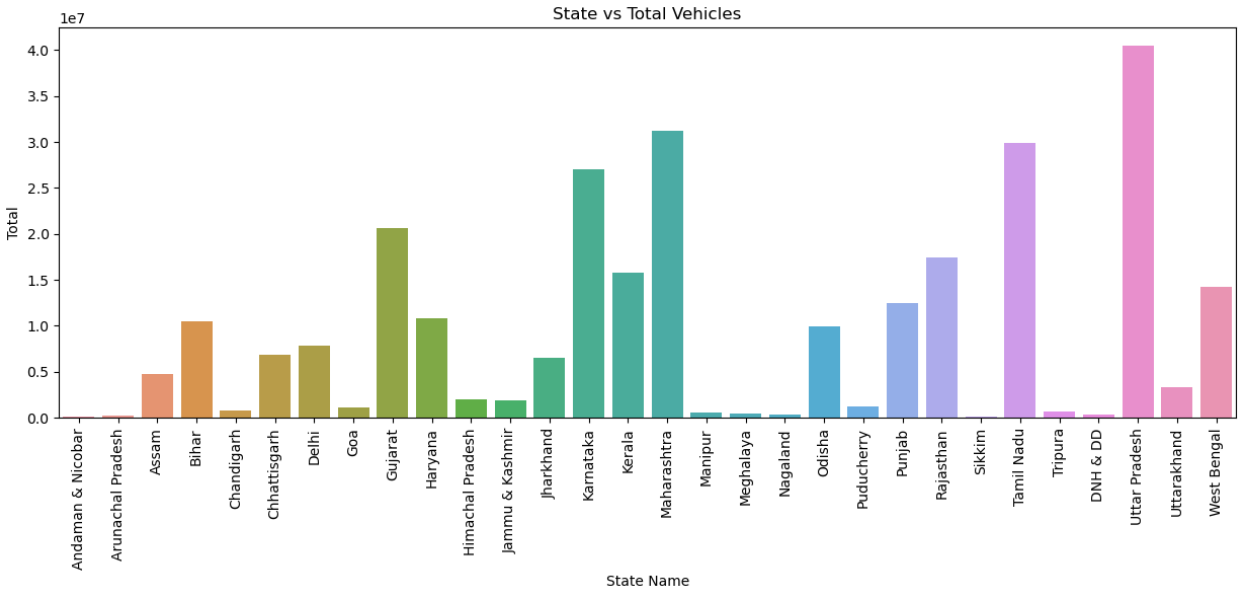
	Total Electric Vehicle	Total Non-Electric Vehicle	Total	No. of Operational PCS
count	30.000000	3.000000e+01	3.000000e+01	30.000000
mean	44477.933333	9.260523e+06	9.305001e+06	194.133333
std	69664.616868	1.091429e+07	1.096983e+07	364.077269
min	20.000000	9.718900e+04	9.721000e+04	1.000000
25%	1418.500000	6.742398e+05	6.818892e+05	16.500000
50%	18888.500000	5.581995e+06	5.622784e+06	48.000000
75%	60766.250000	1.371638e+07	1.375666e+07	194.250000
max	337180.000000	4.009249e+07	4.042967e+07	1845.000000

Total sum

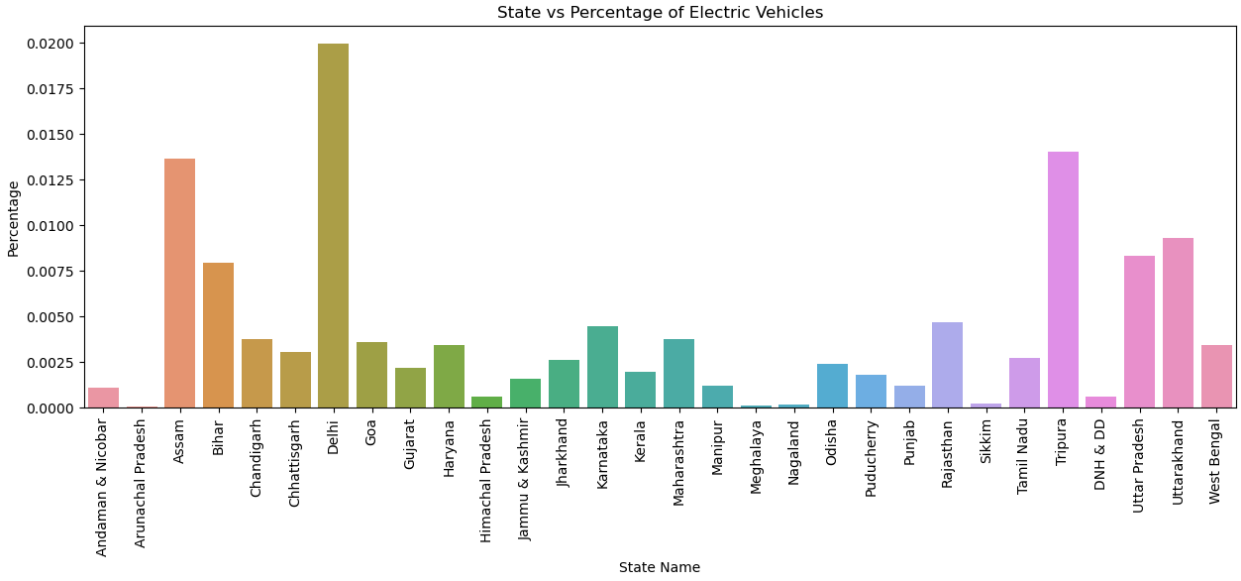
Total Electric Vehicle	1334338
Total Non-Electric Vehicle	277815703
Total	279150041
No. of Operational PCS	5824



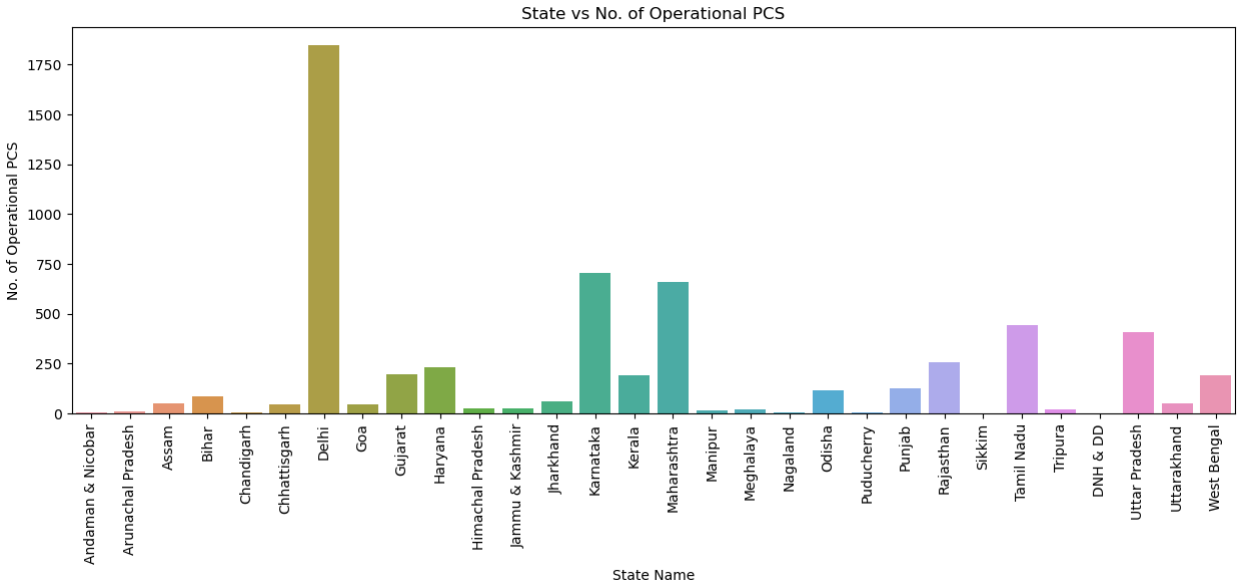
Uttar Pradesh, Delhi and Karnataka have most no of electric vehicles



Uttar Pradesh, Maharashtra and Tamil Nadu have the greatest number of vehicles.



Delhi, Tripura, Assam are adopting EVs faster than other states.



Delhi, Karnataka and Maharashtra have good infrastructure.

CONCLUSIONS FROM MARKET SEGMENTATION

- India is predicted to have rapid growth. With growing environment concerns of conventional petrol and diesel vehicles, transition to electric vehicles is unavoidable.
- India is a very price sensitive market. Most EVs come below 50,000 euros or 45 lakh rupees.
- Most EVs come as SUVs, then hatchback and then sedan. EVs are belonging predominantly in C segment.
- Sedan is on the expensive side, while SUVs are more affordable, and hatchbacks are the cheapest in the market.
- EV prices are marked at 20k euros or 18 lakh rupees (SEAT Mii Electric) to 215k euros or 1.94 crore rupees (Tesla Roadster).
- SUVs have good efficiency; sedans have fast charging while hatchbacks are cheaper in comparison. Both sedans and SUVs provide good range.
- Three clusters were determined through K-Prototype Clustering.
- Cluster 0 contains older people having high bracket of salaries. Luxury cars, sedans and SUVs need to be marketed towards this cluster.
- Cluster 1, younger demographic than cluster 0, has a good amount of salary. This segment prefers SUVs and sedans
- Cluster 2 is the youngest demographic having lower salaries compared to other clusters. Hatchbacks need to be marketed towards this segment.

- Peoples' interest towards EVs and the charging infrastructure is distributed disproportionately in different parts of India.
- Top three states in different parameters used to determine best location are given below

PARAMETER	FIRST	SECOND	THIRD
Total Electric Vehicle	Uttar Pradesh	Delhi	Karnataka
Total Vehicles	Uttar Pradesh	Maharashtra	Tamil Nadu
Percentage of EVs	Delhi	Tripura	Assam
No. of PCS	Delhi	Karnataka	Maharashtra

- **Delhi is the best location to enter the EV Market.** Delhi boasts of good infrastructure and high density of EV consumers. There is also a high converting rate from petrol and diesel to EV vehicles. Due to high pollution in Delhi, government is also very keen to completely electrify the automobile sector.
- **Karnataka** is also a good location with good infrastructure and large number of EV customers.
- **Uttar Pradesh** has high consumers, but it is a very large state with high population (and hence many consumers) with very less infrastructure.
- There is untapped potential in **Maharashtra** with a lot of automobile consumers situated there. It already has good infrastructure.
- **Tamil Nadu** has less EV customer even though it is the state with 2nd largest no of vehicles. If infrastructure of Tamil Nadu becomes better, this would be also a good market place.

REFERENCES

https://e-amrit.niti.gov.in/benefits-of-electric-vehicles

https://www.fortunebusinessinsights.com/india-electric-vehicle-market-106623

https://towardsdatascience.com/clustering-algorithm-for-data-with-mixed-categorical-and-numerical-features-d4e3a48066a0

https://towardsdatascience.com/the-k-prototype-as-clustering-algorithm-for-mixed-data-type-categorical-and-numerical-fe7c50538ebb

https://www.rushlane.com/electric-cars-sales-may-2023-tata-mahindra-mg-byd-12471564.html

https://bolt.earth/blog/future-electric-vehicles-india

CODE

GitHub link: https://github.com/Sreeparvathy-S-Menon/EV_MARKET_INDIA