EV MARKET SEGMENTATION

Report Prepared by: *Vishal Pal* (Machine Learning Intern, Feynn Labs)

20 July 2024

Question.

Based on market analysis, the segmentation challenge can be summarized into two key questions:

1. What kind of electric vehicle (EV) will the company manufacture?
2. Who is the target customer?

In this context, the focus is specifically on the 4-wheeler EV market, which is considerably larger and more dynamic compared to others. Therefore, the objective is to determine which types of 4-wheeler EVs a company should allocate its resources towards.

Approach.

The objective is to analyze the electric vehicles (EV) market in India using segmentation analysis and develop a viable strategy for entering the market, focusing on the segments most likely to adopt their product. The analysis will consider geographic, demographic, psychographic, and behavioral factors.

In this report, I have examined the four-wheeler electric vehicles market in India using segmentation analysis to address some of the critical questions. Additionally, I have performed segmentation on customers and their income data.

The segmentation was conducted using Principal Component Analysis (PCA) and K-Means clustering, with hierarchical clustering also demonstrated. Finally, the potential segments for an EV company to target are evaluated based on selected features.

The GitHub Link to the Analysis:
https://github.com/vissu1299/FeynnlabProject_T2_EV_MarketSegmentation.git

Demographic dataset and EV bikes dataset are collected from https://www.kaggle.com/

The bikes dataset needed to be modified to add some missing values like products from Popular brands.

PSYCHOLOGICAL AND BEHAVIOURL SEGMENTATION:

In this section, I outline the methods used to perform demographic segmentation on the dataset. Initially, an Exploratory Data Analysis (EDA) is conducted. EDA is a comprehensive examination that reveals the underlying structure of a dataset, making it crucial for a company. It helps uncover trends, patterns, and relationships that are not immediately obvious.

Exploratory Data Analysis (EDA):

The datasets:

	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

The analysis of the dataset, with columns such as Exterior, Comfort, Performance, Fuel Economy, Value for Money, and Rating, indicates that consumers are generally satisfied with all attributes. However, the reviews suggest that consumers are particularly more satisfied with the appearance of the EVs compared to the value for money they provide. This highlights the importance of aesthetic appeal in influencing consumer satisfaction within the electric vehicle market.
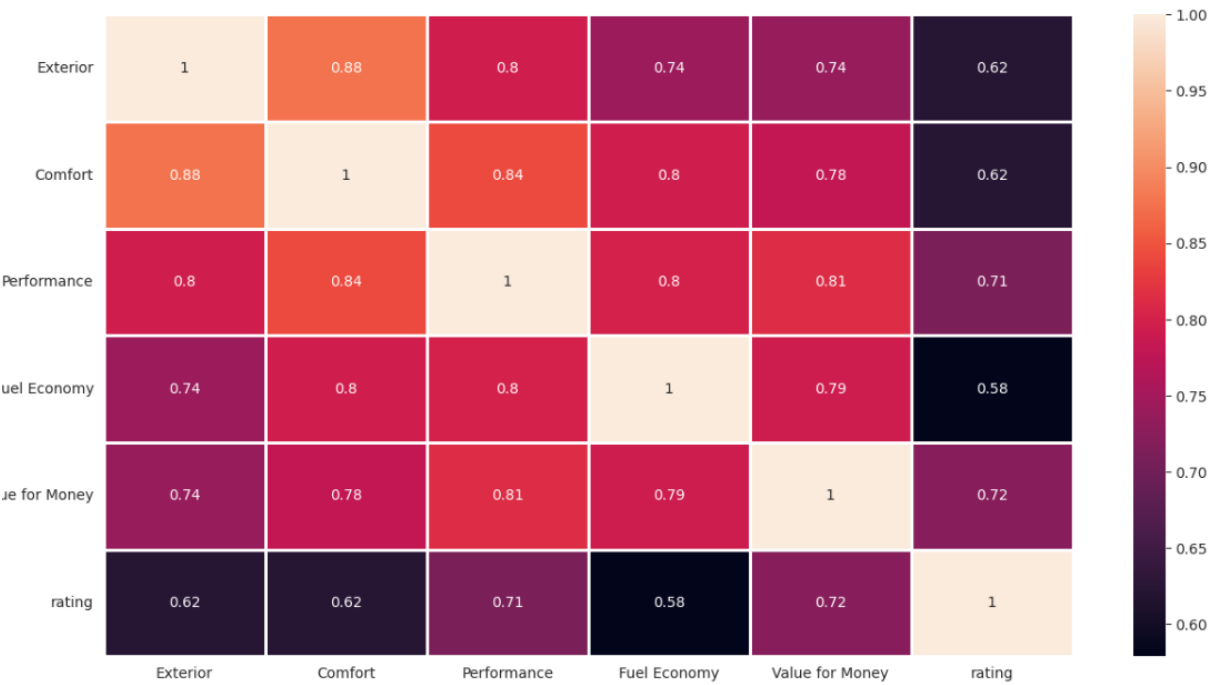
	Exterior	Comfort	Performance	Fuel Economy	Value for Money	rating
count	129.000000	129.000000	129.000000	129.000000	129.000000	129.000000
mean	4.472868	4.418605	4.418605	4.418605	4.162791	4.341085
std	0.968871	1.036051	1.150392	1.122899	1.345076	1.389110
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000
50%	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000
75%	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000
max	5.000000	5.000000	5.000000	5.000000	5.000000	5.000000

|	Exterior	Comfort	Performance	Fuel Economy	Value for Money	rating
Exterior	1.000000	0.875308	0.795318	0.742981	0.737783	0.622239
Comfort	0.875308	1.000000	0.841613	0.795066	0.780424	0.621994
Performance	0.795318	0.841613	1.000000	0.800712	0.813931	0.711727
Fuel Economy	0.742981	0.795066	0.800712	1.000000	0.792479	0.578896
Value for Money	0.737783	0.780424	0.813931	0.792479	1.000000	0.722676
rating	0.622239	0.621994	0.711727	0.578896	0.722676	1.000000

The findings indicate a high correlation between the appearance and comfort of the electric vehicles. Interestingly, the ratings given by consumers do not show a correlation with the vehicle's fuel economy. Instead, higher ratings are awarded

when the vehicle offers good value for money. Additionally, both performance and value for money play crucial roles in driving consumer satisfaction.

	Exterior	Comfort	Performance	Fuel Economy	Value for Money	rating
Exterior	1.000000	0.875308	0.795318	0.742981	0.737783	0.622239
Comfort	0.875308	1.000000	0.841613	0.795066	0.780424	0.621994
Performance	0.795318	0.841613	1.000000	0.800712	0.813931	0.711727
Fuel Economy	0.742981	0.795066	0.800712	1.000000	0.792479	0.578896
Value for Money	0.737783	0.780424	0.813931	0.792479	1.000000	0.722676
rating	0.622239	0.621994	0.711727	0.578896	0.722676	1.000000

The Kernel Density Estimate (KDE) plot illustrates that there is a high correlation between the appearance and comfort of electric vehicles. It also shows that consumer ratings are not correlated with the vehicle's fuel economy. The plot highlights that vehicles perceived as offering better value for money tend to receive higher ratings. Moreover, it underscores the significant roles that performance and value for money play in consumer satisfaction.
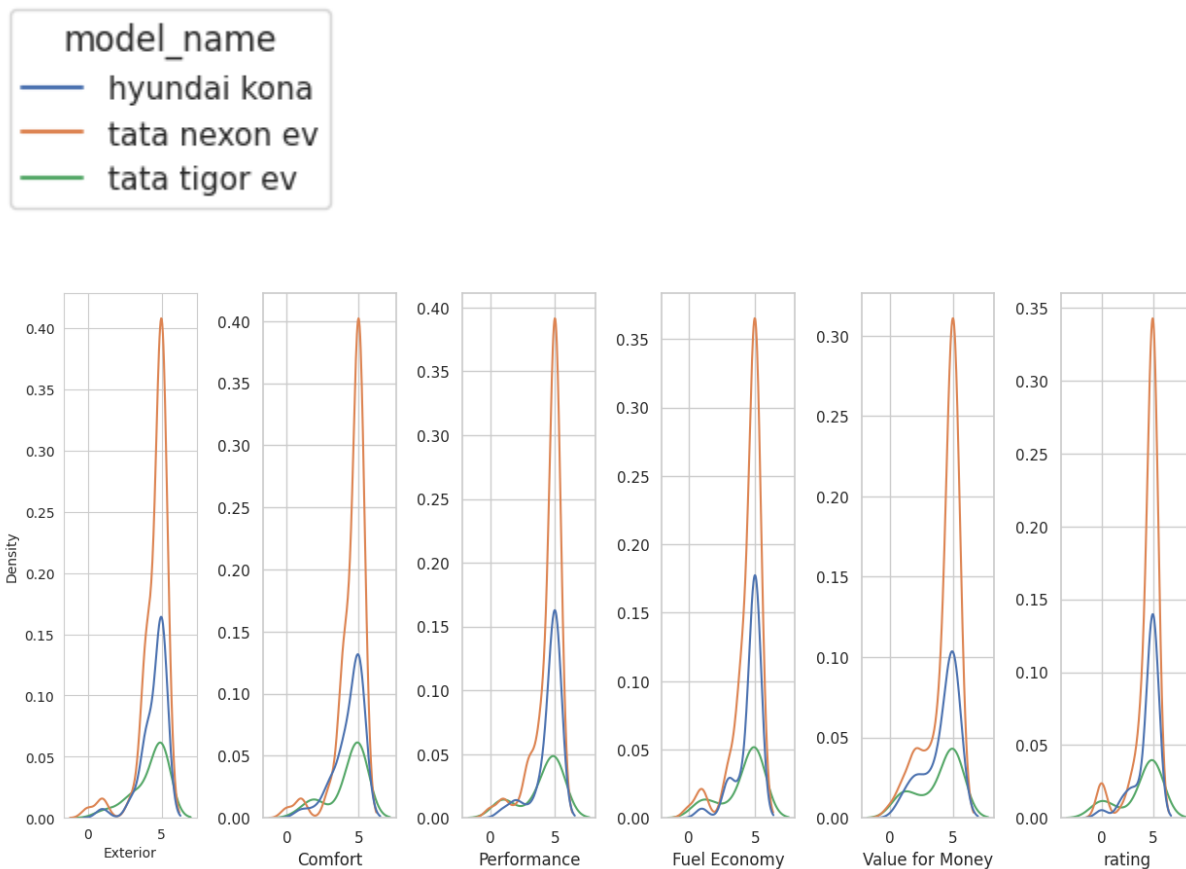

The boxplot reveals a high correlation between the appearance and comfort of electric vehicles, while consumer ratings show no correlation with fuel economy. It also indicates that better value for money and strong performance are key factors in achieving higher consumer satisfaction.
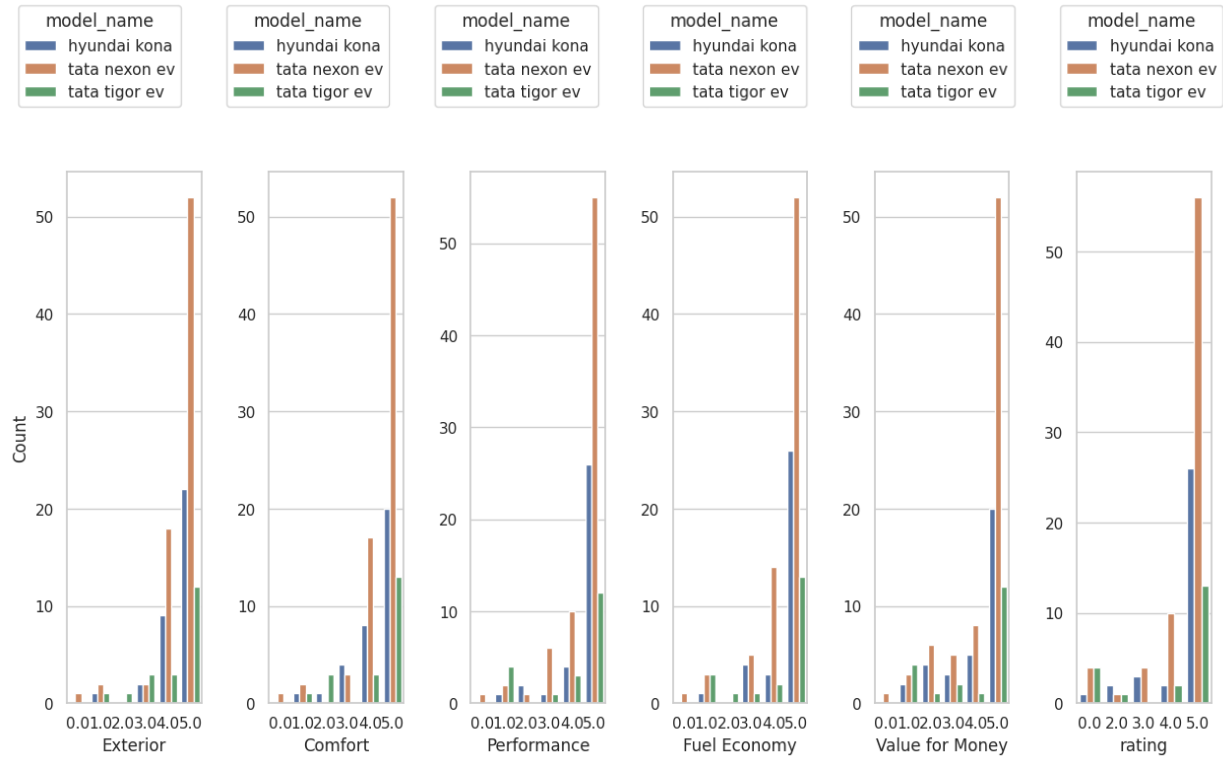
A data to show percentage of candidates having different electric vehicle options given by the company.
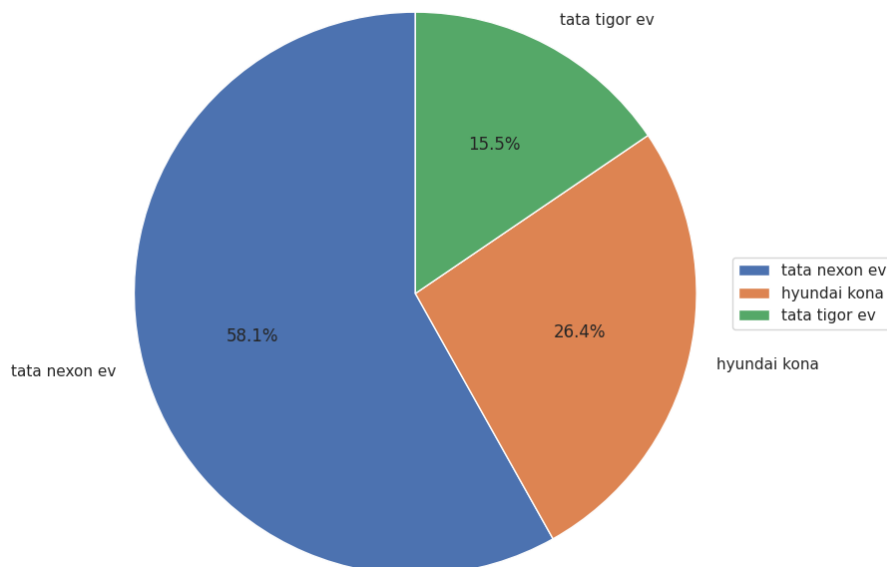
The findings indicate that electric vehicles are primarily used for short to mid-distance travel by 28.7% of users. A significant portion, 26.4%, use them for long-distance travel. Additionally, a substantial 35.7% of the population, who have not yet purchased an EV but have driven one and are satisfied with the experience, appear willing to adopt EVs based on the ratings given.
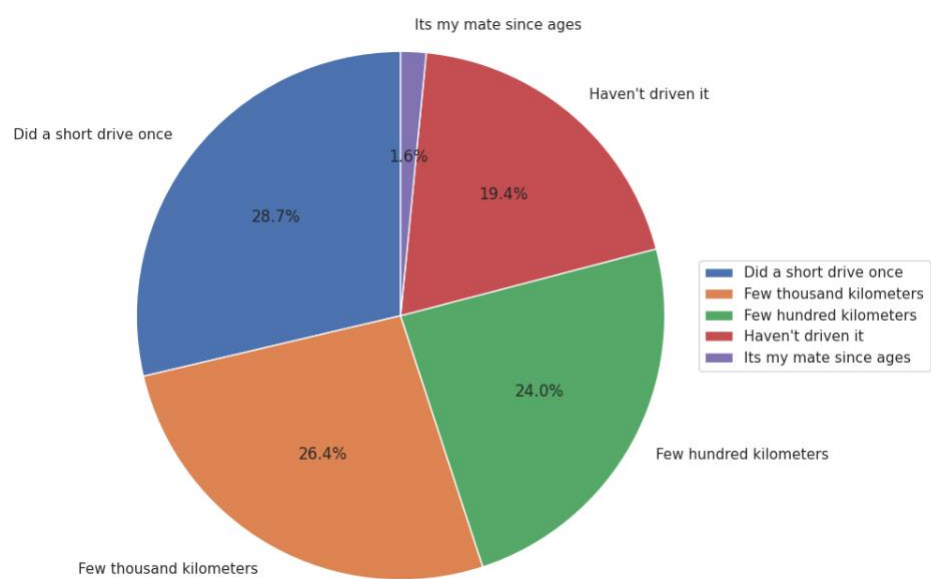

The findings reveal that the Tata Nexon EV has the largest market presence and is the best-performing vehicle, achieving the highest average ratings overall. In contrast, the Tata Tigor EV has the lowest market presence and the poorest consumer satisfaction. The Hyundai Kona, while competitive, has a lower market presence.

| model_name	Exterior	Comfort	Performance	Fuel Economy	Value for Money
hyundai kona	4.500000	4.323529	4.529412	4.558824	4.088235
tata nexon ev	4.533333	4.520000	4.493333	4.453333	4.293333
tata tigor ev	4.200000	4.200000	3.950000	4.050000	3.800000

K-Means Clustering:

I performed K-means clustering for various cluster counts and plotted the Elbow curve to determine the optimal number of clusters, as the algorithm requires this input. It's worth noting that I used the "k-means++" initialization method, which smartly initializes clusters rather than doing so randomly as in the standard K-means.

KMeans Clustering Visualization

Elbow Method Plot
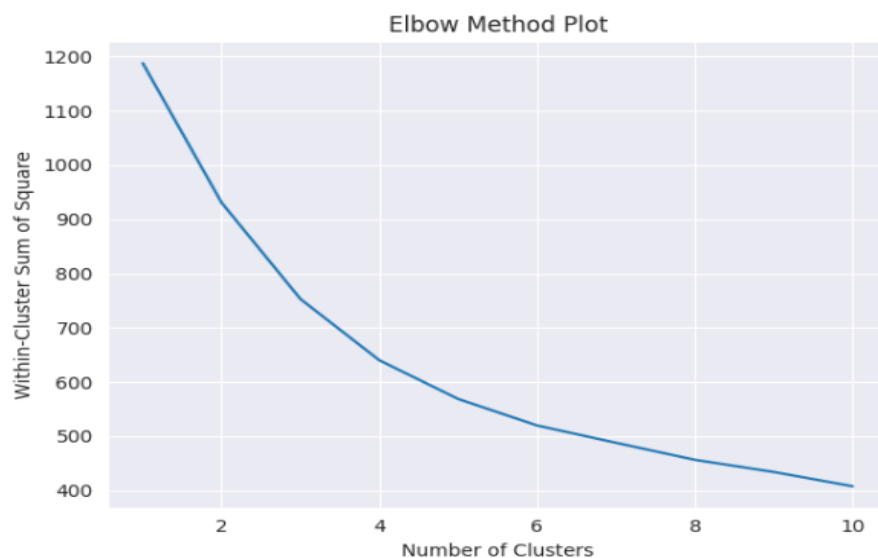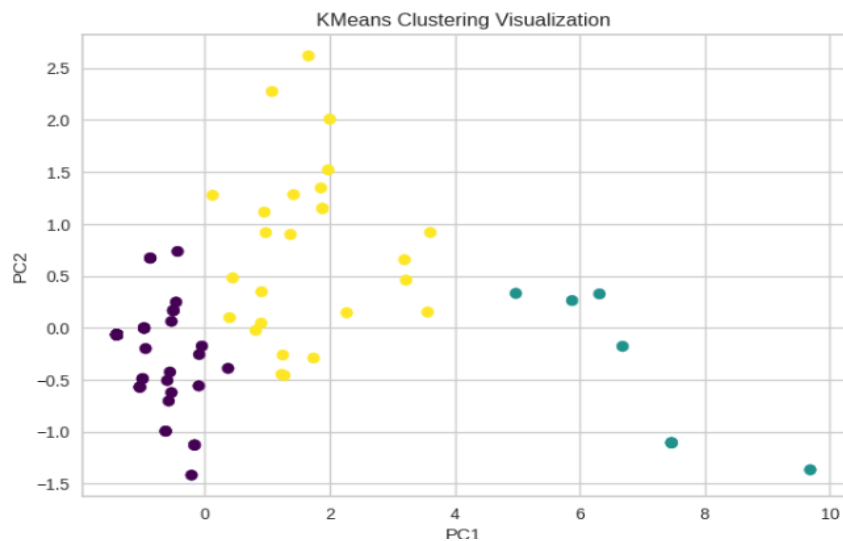
Hierarchical Clustering:

A similar hierarchical clustering is also performed which shows 2 different components, shown in the Dendrogram.

Hierarchical Clustering Dendrogram
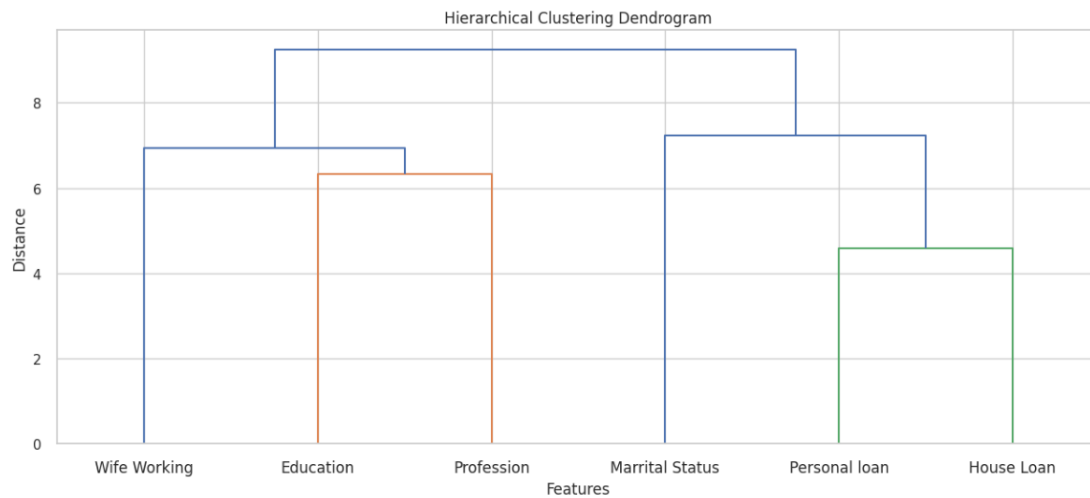
EXTRACTING SEGMENT:

Segment 0 : High satisfaction on every aspect ,indicates higher willing to adoption of EV, forms most of the population - 72.9%; Most desired EV is Tata Nexon; These are Early majority

Segment 1 : Unsatisfied on all aspects, didn't find EV to be value for money ; forms 7% of population ; Found Tata tigor EV to be most likeable; can be considered Late adapters.

Segment 2: They liked Exterior and Comfort of EVs but didn't find it Value for money much; form 20.2% population ;Found each EV equally likeable; can be considered Early adaptors
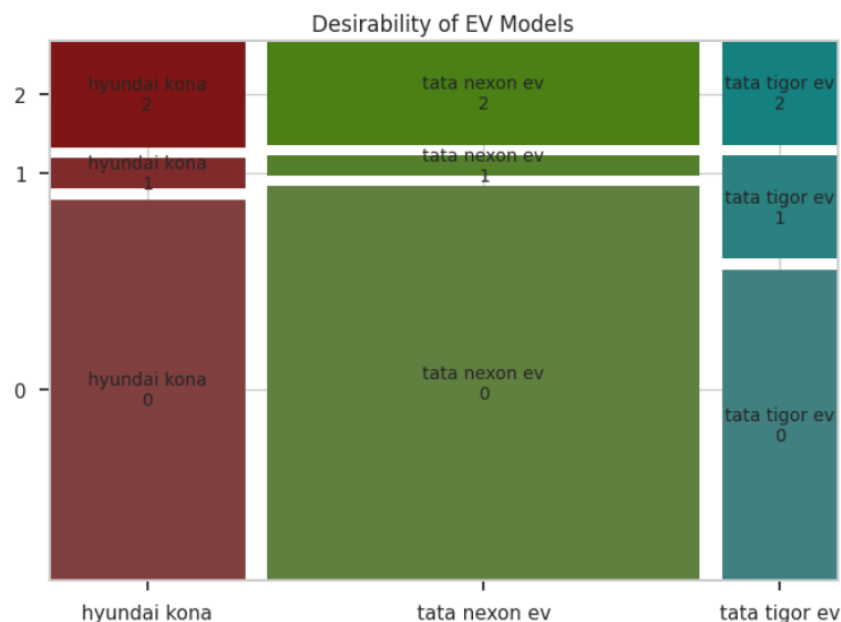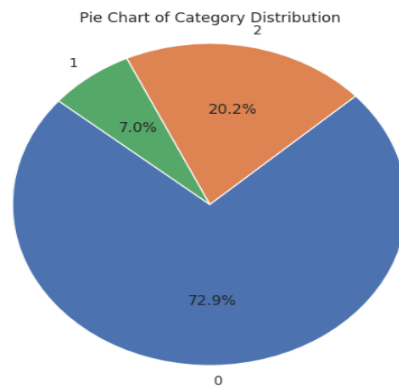
Data Exploration on Buying behaviour:

- On average Consumer is 36 years old Salaried professional, Married, having higher education with 2 dependents, No ongoing loans, having a working wife, earning household income of 22 Lacs , Opting for SUV pricing up to 12 lacs
- Findings:
    1. Price of car purchased is highly correlated with Total salary

    2. No. of dependents play no role in deciding price of car to purchase or even make.
    3. Age is somewhat related to Salary earned but not so much to the caprice.

- Findings:
    1. Younger people Mostly opt for Hatchbacks like Baleno or i20 and some go for Sedans ,there is small population of youngsters who also go for Suv but none go for Luxury
    2. Higher Age people mostly opt for Luxury of Premium Sedans like Ciaz.
    3. People who Earn Higher salary go for SUVs or Luxury.

    4. i20 and Baleno (Hatchbacks) are most popular vehicle of Lower income group.

    5. When Wife's salary is also high and resultantly Total salary also hikes up ,choice of vehicles becomes Luxary and SUVs

    6. Luxary segment has highest market price and Baleno has lowest.


Multiple Subplots with Titles

Salary

Wife Salary

Total Salary

Price
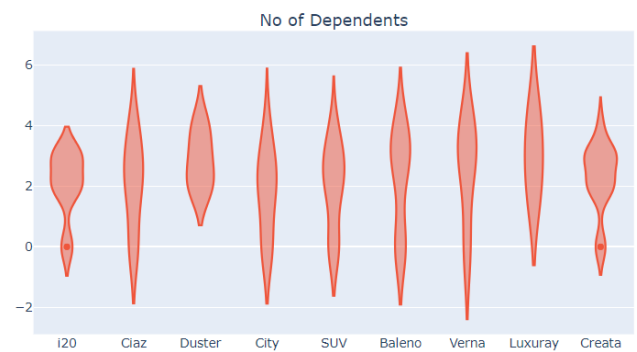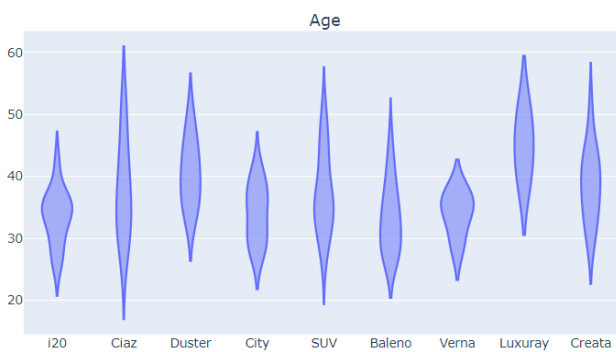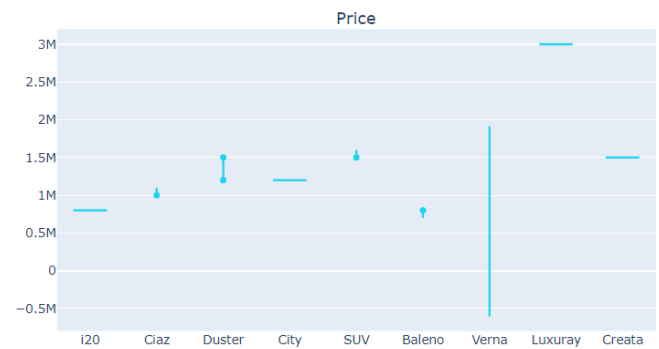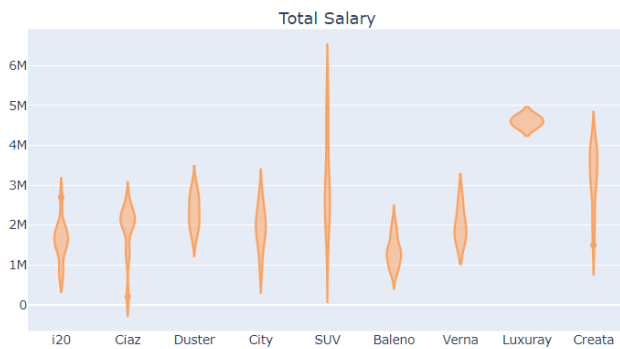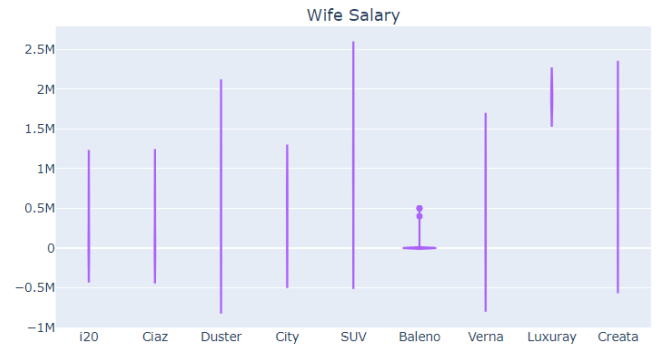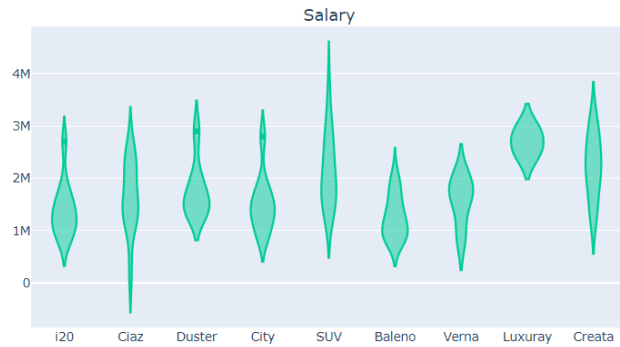
- **Notable Findings:**

1. **SUVs holds most of the market capital** with their primary Customers being **Post Graduate ,High Income Married , Salaried Professionals with Working wife**; **Singles** of this group prefer **City or Creta** (Sedan/Compact SUV)
2. Second Fav Vehicle in this segment is **Creta**
3. **Highly Educated Married Businessmen** Prefer **Hatchback &sedan** whereas Single of this catagory also prefer SUV
4. Overall Post graduates Salaried Professionals prefer SUV and Businessmen prefer Baleno So we can say These two are most popular catagories.
5. Overall Graduates Salaried Professionals prefer Ciaz and Businessmen prefer SUV So we can say These two are most popular catagories.
6. Overall Most Popular vehicle segments in decreasing order are
   : **Suv > Baleno(Hatchback) > CIAZ(Premium sedan)**
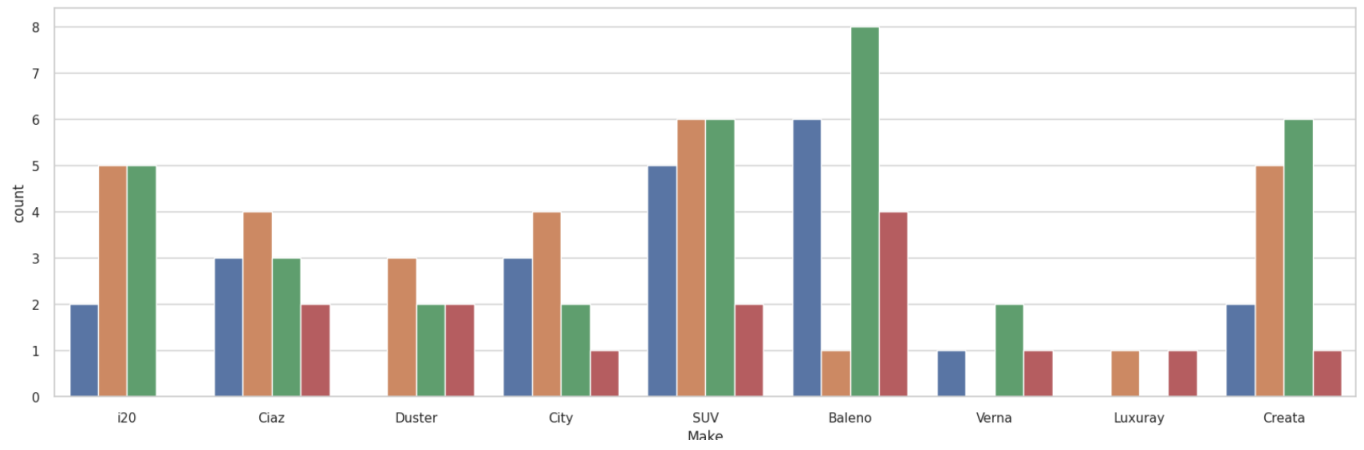
DEMOGRAPHIC SEGMENTATION:

Describing Datasets:

|	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

```
Age : [27 35 45 41 31 28 33 34 29 30 49 26 37 36 43 42 32 44 39 46 50 51]
Profession : ['Salaried' 'Business']
Marrital Status : ['Single' 'Married']
Education : ['Post Graduate' 'Graduate']
No of Dependents : [0 2 4 3]
Personal loan : ['Yes' 'No']
House Loan : ['No' 'Yes']
Wife Working : ['No' 'Yes' 'm']
Salary : [ 800000 1400000 1800000 1600000  900000 1200000 2000000 1300000 2500000
 1700000 1100000 1900000 2100000 2400000 2200000  200000 1500000 2700000
 2900000 3100000 2600000 2300000 2800000 3800000]
Wife Salary : [      0  600000  800000  700000  400000 2000000  500000 1000000 1100000
  900000 1300000 1400000 1800000 2100000]
Total Salary : [ 800000 2000000 1800000 2200000 2600000  900000 1400000 1900000 1700000
 1300000 4500000 2500000 2400000 2900000 1600000 2700000 1100000 2100000
 3000000 3700000 2300000 3600000  200000 3100000 4300000 3800000 4700000
 1200000 1500000 4000000 3200000 5200000 4100000 4900000 2800000 5100000]
Make : ['i20' 'Ciaz' 'Duster' 'City' 'SUV' 'Baleno' 'Verna' 'Luxuray' 'Creata']
Price : [ 800000 1000000 1200000 1600000  700000 1100000  110000 3000000 1300000
 1500000]
```
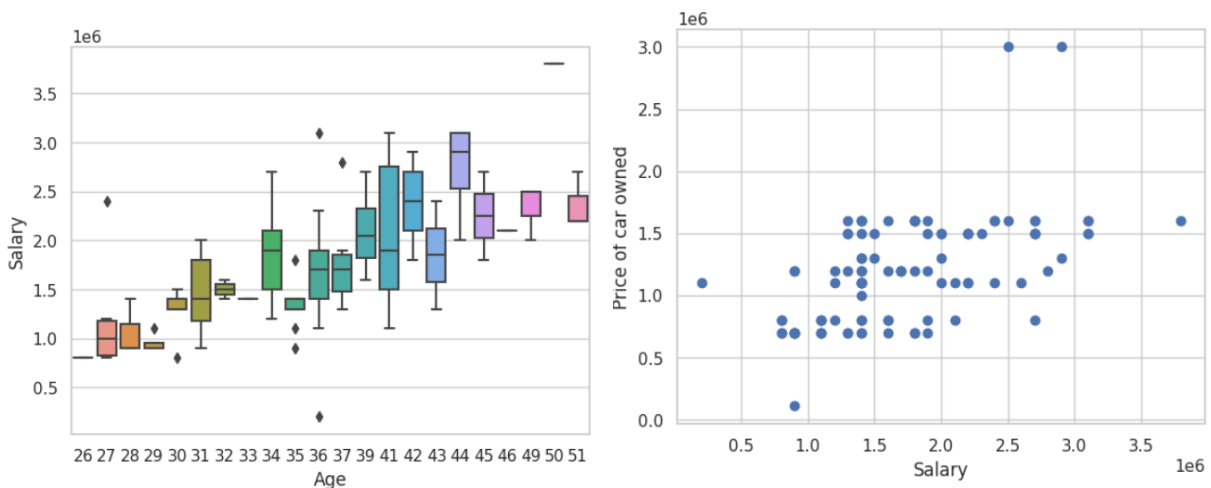
Exploratory Data Analysis (EDA):
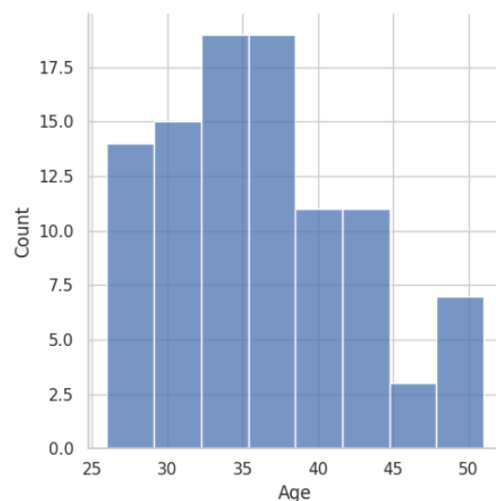
Loan:

Although Personal and home loans do not seem to have any significant impact on EV purchasing patterns.
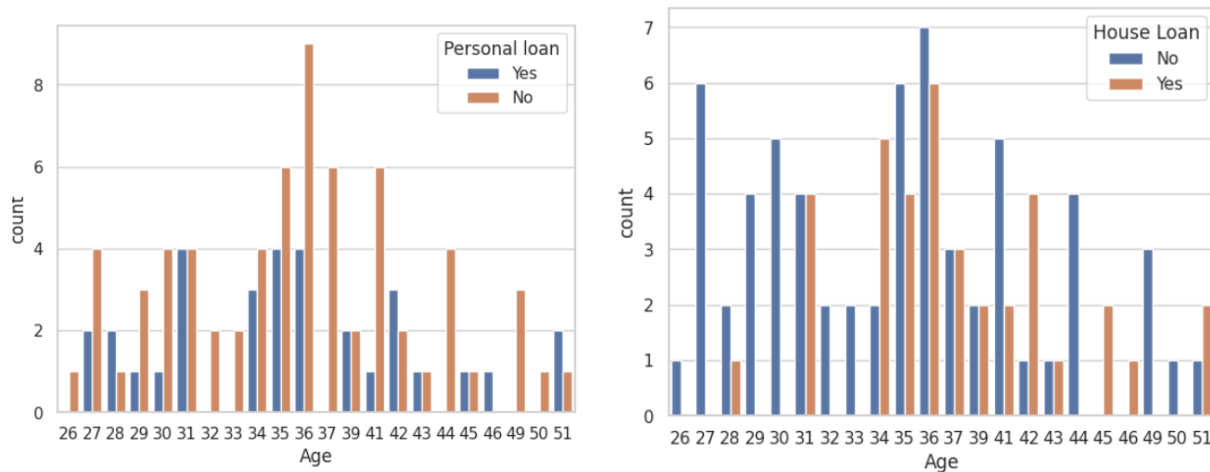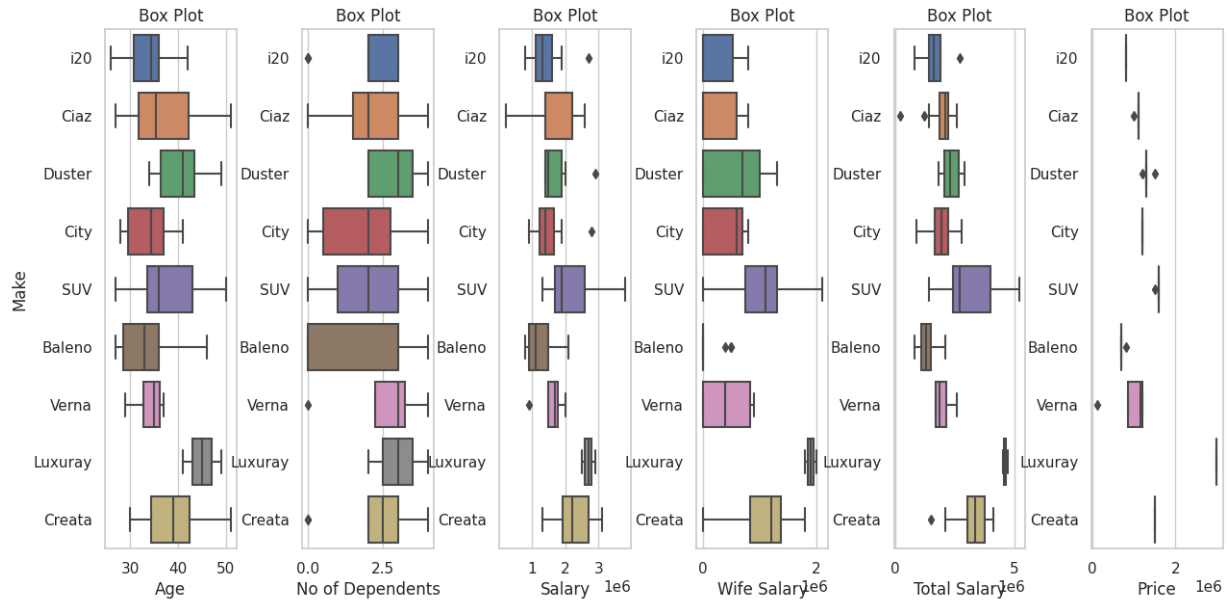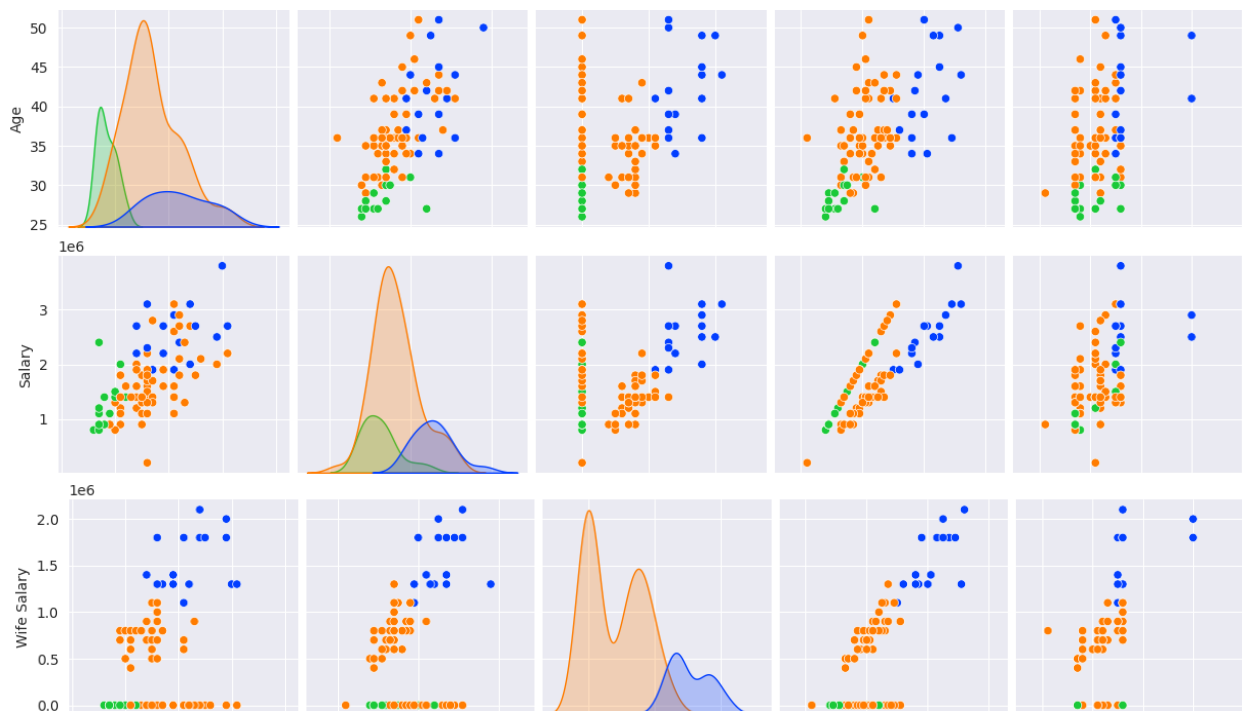

Salary:

It is to be noted that regardless of even high salaries, average to low price-ranged cars are being preferred.
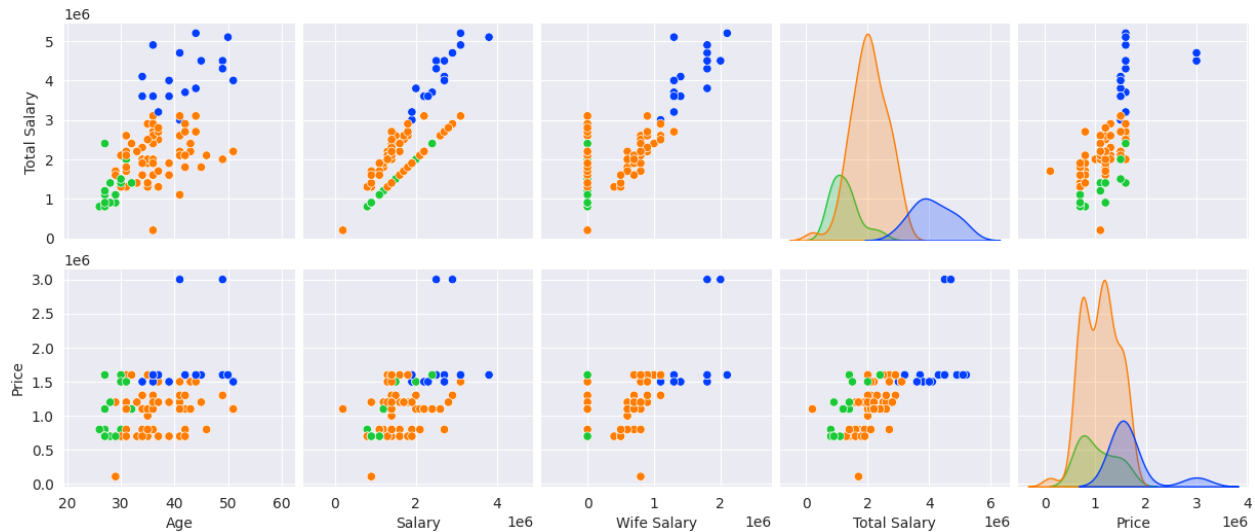
We see the people buying over a range of prices when it comes to Verna, whereas in case of other brands, probably some specefic models are being preferred.

Notable findings:

we see that we can get a more meaningful analysis for k = 3 Let's see the characteristics of every cluster:

Cluster 1 (Green): Prefer low to moderately priced Electric vehicles

- Age group 20-30
- Moderate to average salary
- Wife's salary - Low

Cluster 2 (Orange) : Prefer low to moderately priced Electric vehicles

- Age group 30-45
- Average to high salary
- Wife's salary - low to

Cluster 3 (Blue) : Prefer moderate priced Electric vehicles and experiment with the high priced segment.

- Age group 30-45
- Average to very high
- Wife's salary - high

1. Moderately priced EVs are the most preferred across all the classes/clusters.
2. Most purchasing acitivity is seen in the age range of 26-38.
3. Regardless of even high salaries, average to low price-ranged cars are being preferred. Personal and home loans do not seem to have any significant impact on EV purchasing patterns.
4. We see the people buying over a range of prices when it comes to Verna, whereas in case of other brands, probably some specefic models are being preferred.
5. We observe high correlation between total salary (salary + wife salary) and price.

Analysis based on efficiency achieved by a brand:

Analysis based on efficiency achieved by a brand:
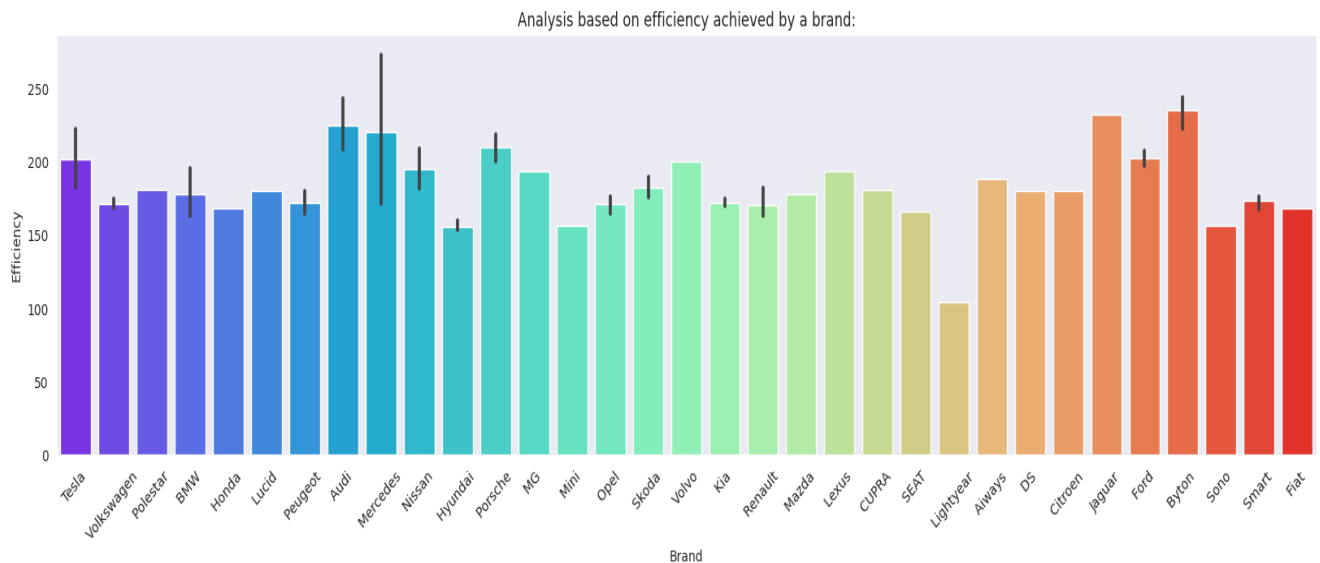
Analysis based on brand price:

This is the price distribution of Electric Vehicles visualised over different brands.
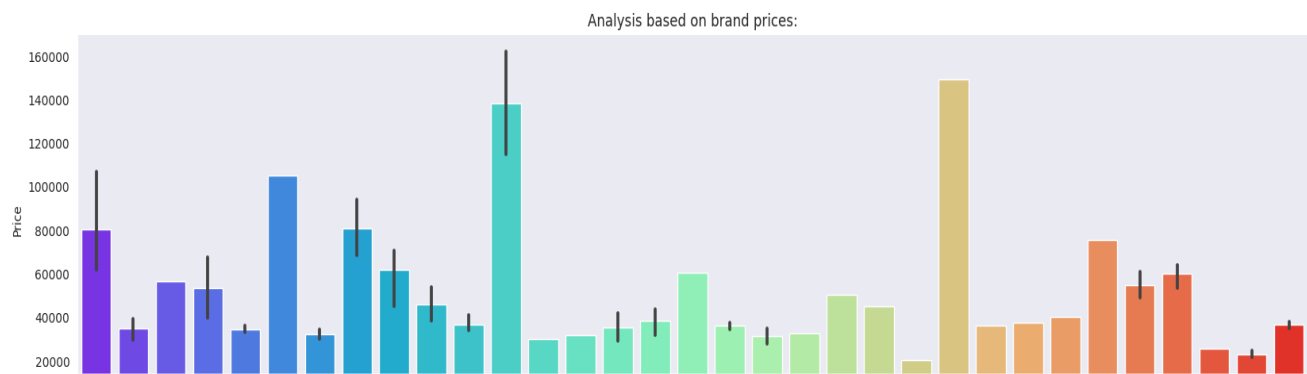

Optimal Target Profile:

Next, I aim to address the posed questions by identifying potential target segments. The selection of target variables and clusters can vary based on the company's policies and objectives. Here, I have identified possible target clusters.

The EV Manufacturer Company might choose the 1st Segment (Demographical profile) having feature values-

- *Cluster 0: Young Age , No Business, Graduate, Single*
- *Cluster 1: Young age , working Professional, Graduate, Married.*
- *Cluster 2: Middle aged(30+) , Businessmen, Post Graduate*
- *Cluster 3: Middle aged, Salaried professional, Married*
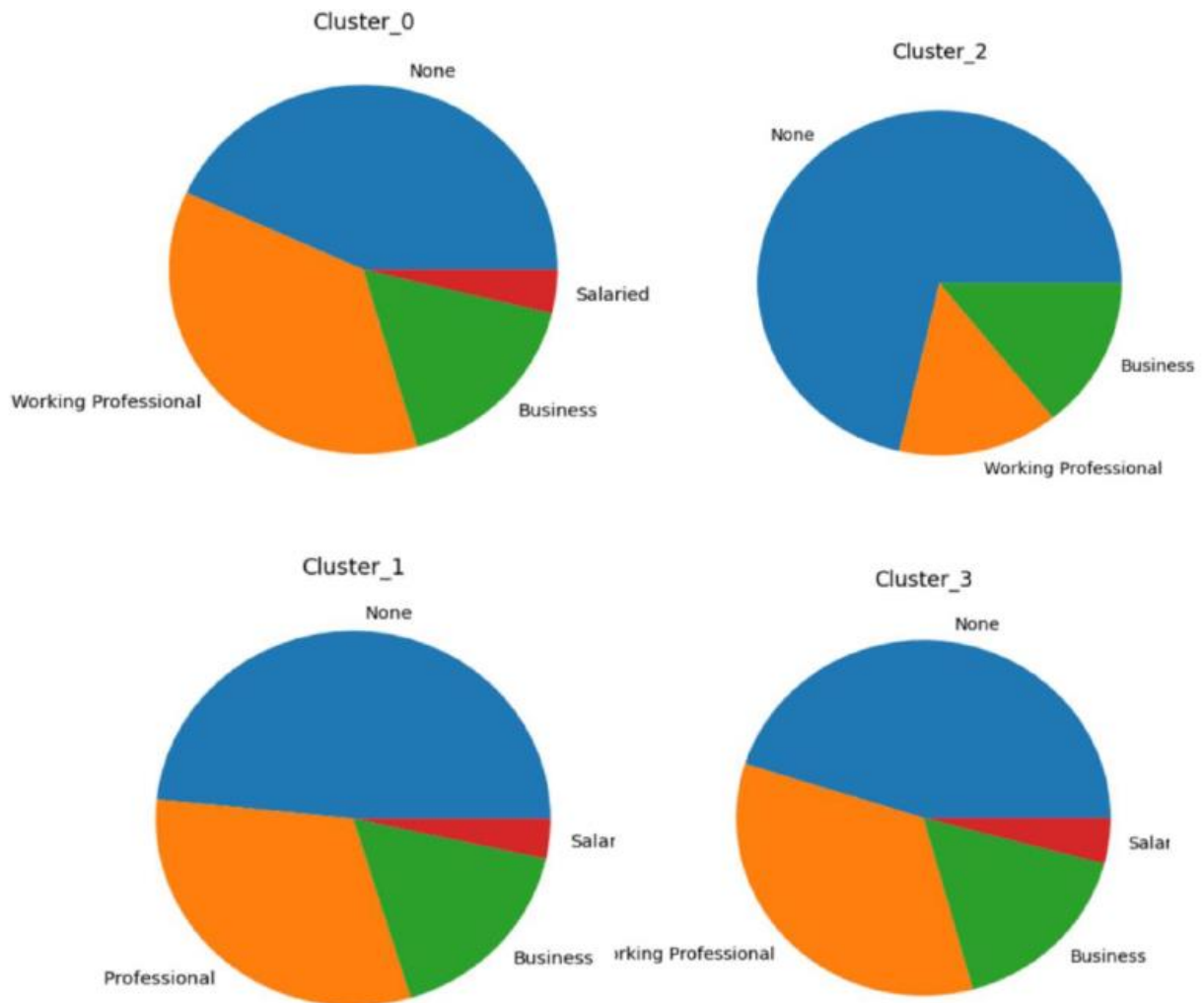
The EV Manufacturer Company might choose the 2st Segment (Geographical profile) having feature values-

- *High EV Adoption, Moderate Charging Infrastructure, Moderate EV Charging Sanctions.*
- *States - Tamil Nadu and Chhattisgarh*

- *Low to Moderate EV Adoption, Low Charging Infrastructure, Low EV Charging Sanctions.*
- *Odisha, Punjab, Bihar, Assam, Haryana, Ladhakh, Sikkim, Jharkhand, Puducherry, Goa, Jammu Kashmir*
- *High EV Adoption, High Charging Infrastructure, High EV Charging Sanctions.*
- *States: Karnataka And Delhi*
- *Moderate EV Adoption, Moderate Charging Infrastructure, High EV Charging Sanctions.*
- *Maharashtra, Rajasthan, West Bengal, Gujarat, Kerala*
- *High EV Adoption (Mainly Three Wheelers and Four Wheelers), Low Charging Infrastructure, Low EV Charging Sanctions.*
- *States: Uttar Pradesh*

POSSIBLE IMPROVEMENTS:

Although the analysis has been conducted and results have been obtained, there is always room for improvement. Alternative algorithms such as DBSCAN or Gaussian models could be utilized for additional insights.

Additionally, having clearer data and more features would enhance the precision of the segmentation. Such data could be collected through methods like web scraping, manual data collection, and surveys.

MARKET MIX:

To effectively target the market, it's important to consider these four key factors: Innovation, Infrastructure, Customer Experience, and Sustainability.

Innovation:

Constant innovation is crucial in the electric vehicle market to stay competitive and meet evolving consumer demands. This includes:

- Advanced battery technologies
- Autonomous driving features
- Smart connectivity and integration with other technologies
- Continuous improvement in vehicle design and performance

Infrastructure:

Building and maintaining a robust infrastructure is vital for the adoption of electric vehicles. Key aspects include:

- Availability and accessibility of charging stations
- Efficient service and maintenance facilities
- Integration with smart grid systems
- Partnerships with local governments and private sectors for infrastructure development

Customer Experience:

Providing an exceptional customer experience is essential to attract and retain customers. This can be achieved through:

- Comprehensive after-sales services
- User-friendly mobile apps for vehicle management
- Personalized customer support
- Offering test drives and educational workshops on EV benefits and usage

Sustainability:

Emphasizing sustainability can enhance brand reputation and attract environmentally conscious consumers. This involves:

- Using eco-friendly materials in manufacturing
- Implementing recycling programs for batteries and parts
- Reducing the overall carbon footprint of the production process

CONCLUSION:

This report explores the introduction of eco-friendly car technology, focusing on four-wheeler electric vehicles (EVs). The adoption of EVs depends on perceived product attributes, personal traits, and background variables. Favorable attitudes towards EVs, based on perceived usefulness and technical performance, drive adoption.

Machine learning in market segmentation has provided valuable insights into distinct customer segments and preferences within the EV market. This analysis reveals key patterns in consumer behavior, guiding product development, infrastructure, and marketing strategies.

Segmentation has uncovered significant profitability opportunities, driven by rising fuel costs, making EVs more attractive for their cost-efficiency and energy conservation. Overall, machine learning-driven segmentation has enabled data-driven decisions and refined marketing approaches in the promising electric vehicle market.