Market Segmentation

Analysing the Electric Vehicle market in India using Segmentation analysis for an Electric Vehicles Startup and coming up with a feasible strategy to enter the market, targeting the segments most likely to use Electric vehicles.

GROUP 2

Marisha Bhatti
Hari G
S. Siranjeevi

Github Link:
https://github.com/Marisha18/Market-Segmentation-for-Electric-Vehicles-in-India

Overview

More than 90% of vehicles all over the world run on oil, there is a noticeable trend of desire to power vehicles with alternative energy sources. As a result, the subject of electric vehicles (EVs) is gaining popularity. An electric vehicle is one that operates on an electric motor instead of an internal combustion engine, which generates power by burning a mix of fuel and gases. Therefore, an electric vehicle is seen as a possible replacement for the current-generation automobile in the near future. As the problem of rising levels of global air pollution is serious, the use of electric cars can be a response to the achievement of sustainable development goals. With a pressing need for smarter infrastructure and friendlier government policy, electric vehicles have an important role to play in India's energy and mobility markets. In India the current market share of EV/HEV/PHEV is around 0.1%. At present almost all vehicles rely on fossil fuel-based transportation. This pollutes the atmosphere by the emission of greenhouse gases & causes global warming. The Indian transportation sector is growing very fast. The gap between domestic crude oil production and consumption is widening. India is a country which imports around 70% of oil required per year. Hence, there is an urgent need to investigate factors and challenges for the development of sustainable and clean alternatives for transportation systems. Electrified vehicles are one of the promising, clean and sustainable forms of transportation.

The recent scenario of the road transportation sector can be highlighted as:

- Energy consumption: 524 million tons of oil equivalent
- Vehicle to people ratio: 1:56.3
- Per capita energy: 442 kg of oil equivalent
- GHG emissions: 1730 million tons of CO_2 equivalent
- Electric Vehicles sold (2016): 25000 (all) and 2000 (cars)

Unlike other countries the vehicle to people ratio is very high, however, the population is more and emission is high. India stands third with the CO_2 emission of 1.726 billion Metric ton. Hence, there is an urgent need to focus towards EV technology which has the capability towards zero emission for sustainable transportation.

In addition, due to urbanization and decentralization of city areas, a rapid increase in personal vehicles has been observed.

EV (Electric Vehicle) /HEV (Hybrid Electric Vehicle) /PHEV (Plug-in Hybrid Electric Vehicle) can be more beneficial for Indian roads due to the following reasons:

1. Hybrid or electric powertrains operate at much higher efficiency at low Indian driving speeds than an Internal Combustion Engine.
2. A higher share of energy per Indian trip is lost in braking, which is almost recovered in a hybrid-electric vehicle (HEV) and EV (Regenerative braking).
3. HEVs and EVs use no fuel during idling and the share of idling time in traffic is much higher in India (than the U.S. & Europe).
4. The average range travelled in India is much smaller than in the U.S. & Europe, making EVs much more feasible and with no range problem with a single charge.
5. Vehicle use and vehicle distance – Urban driving cycle patterns have a frequent start and stop, high traffic benefits to provide high efficiency electric vehicles.

Market Overview

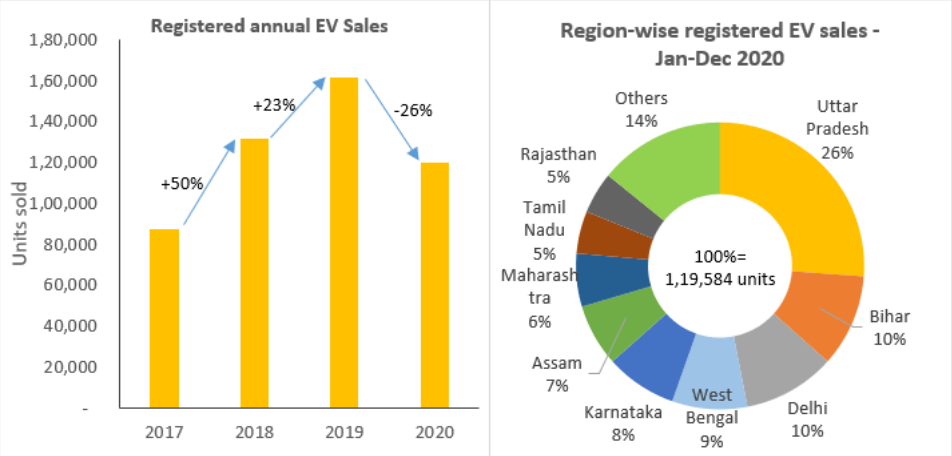
The Indian Electric Vehicle Market is segmented by Vehicle Type and Power Source.

- By Vehicle Type, the market is segmented into Passenger Cars, Commercial Vehicles, and Two- and Three-wheelers.
- By Power Source Type, the market is segmented into Battery Electric Vehicle, Plug-in Electric Vehicle, and Hybrid Electric Vehicle.

Our report mainly focuses on the Indian Electric Vehicle Market segmented by Vehicle Type. However, accessibility to Power Sources for Electric Vehicles affects the market and would be slightly discussed in the report.

The Indian Electric Vehicle Market was valued at USD 5 billion in 2020, and it is expected to reach USD 47 billion by 2026, registering a compound annual growth rate (CAGR) of above 44% during the forecast period (2021-2026).

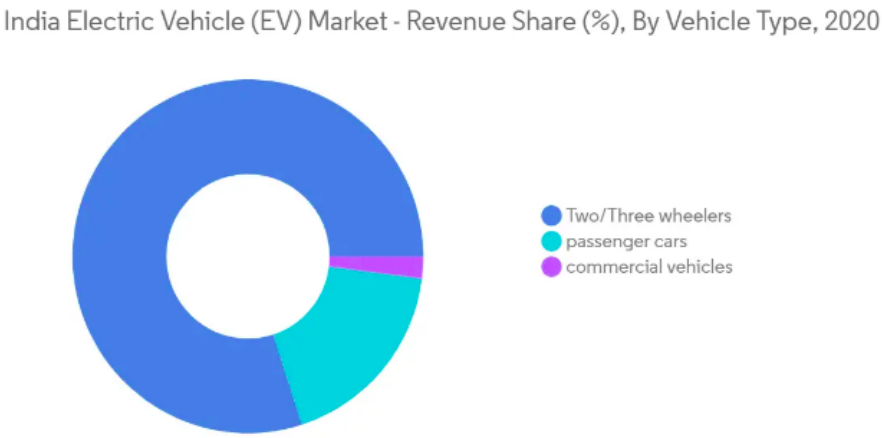
The Indian Electric Vehicle Market has been impacted by the outbreak of the COVID-19 pandemic due to supply chain disruptions and halt of manufacturing units due to continuous lockdowns and travel restrictions across the country. However, the electric vehicle (EV) market is still in its nascent stage in India. It is expected to grow at a much faster rate during the forecast period due to various government initiatives and policies.



E-commerce companies (Amazon, for example) are launching initiatives to use e-Mobility for last-mile deliveries to reduce carbon footprint. India is experimenting with e-Mobility for public transport, and the country has deployed electric intercity buses across some major cities. In addition, state governments are also playing an active role in the deployment of policies encouraging the usage of EVs. For instance,

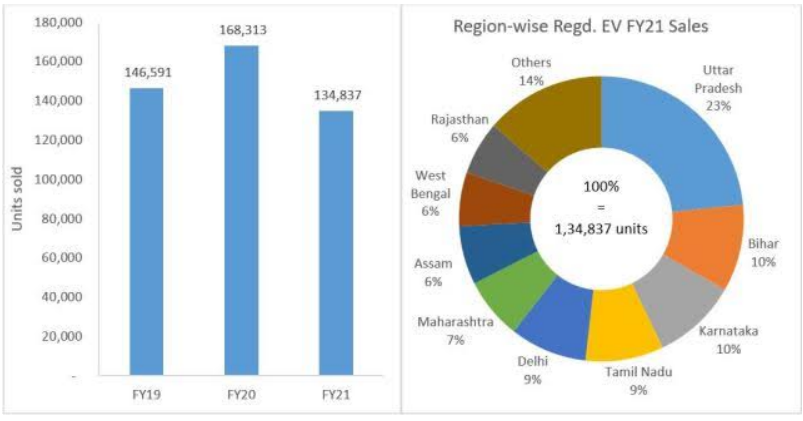
- Kerala aims to put one million EV units on the road by 2022 and 6,000 e-buses in public transport by 2025.
- Telangana aims to have EV sales targets for 2025 to achieve 80% 2- and 3-wheelers (motorcycles, scooters, auto-rickshaws), 70% commercial cars (ride-hailing companies, such as Ola and Uber), 40% buses, 30% private cars, and 15% electrification of all vehicles.

The EV market in India has gained significant momentum after the implementation of the (Faster Adoption and Manufacturing of (Hybrid &) Electric Vehicles in India) FAME India scheme with its aim of shifting toward e-mobility in the wake of growing international policy commitments and environmental challenges. Moreover, India offers the world's largest untapped market, especially in the electric two-wheeler segment. As 100% foreign direct investment is allowed in this sector, the automatic route market is expected to gain momentum during the forecast period.



Market Dynamics

In the financial year 2020-21, the leading type of electric vehicles sold in India was two-wheelers, reaching around 144 thousand units. This was a five percent decrease from the previous year's 152 thousand units. The only section that saw growth was four-wheelers.



The Indian EV market is consolidated with the presence of major players in the market, owing to cheap and readily available manpower. However, established players in the market are introducing new models to gain a competitive edge over other players. The startups are expanding their presence by raising funds from investors and tapping into new and unexplored cities. Companies are investing a tremendous amount in R&D and launching new models to mark their presence in the market.

GOVERNMENT INITIATIVES AND POLICIES SUPPORTING THE EV INDUSTRY

The Indian government has also taken initiatives, like FAME, which will contribute towards the boom in the EV market. The Indian government announced its National Electric Mobility Mission Plan (NEMMP), to support the manufacturing capabilities of local automakers. With this roadmap, the government plans to make electric vehicles economically viable and self-sustaining, by 2020. The government also announced investments of over INR 13,000 crore for demand incentives, INR 1,800 crore for R&D investments, INR 5,000 crore for power infrastructure, and INR 1,200 crore for charging infrastructure. This plan aims at encouraging reliable, affordable, and efficient EVs that can meet the consumers' performance and price expectations. Additionally, it involves a collaboration between the government and the industry for the promotion and development of indigenous manufacturing capabilities, consumer awareness, technology, and required infrastructure, thereby, helping the country emerge as a global leader in the global two-wheeler and four-wheeler electric vehicles market, by 2022.

MARKET CHALLENGES

The push for electric vehicles (EVs) in India seems to be coming at a rapid pace, but the hype does not seem to match the sales of electric vehicles in the country. The slow progress of EV sales is due to various factors, such as limited options in the passenger car segment, driving range of vehicles, lack of affordability, and lack of charging infrastructure.

Affordability is playing a significant role in hindering the growth of the market studied. India is a price-sensitive country, where the majority of people

consider the price of the vehicle first rather than any other factor or aspect. At present, EVs are not affordable for a large section of people who cover a significant sales share of vehicles in the country.

As the electric vehicles market (EVs) in India is at its very nascent stage, the charging infrastructure is also at its minimum, whereas developed countries have well-established charging stations that are more accessible to people for charging their vehicles. Considering the expected increase in the sales of EVs, the development of charging infrastructure becomes very important for the development of a suitable ecosystem. Further, in terms of driving range, very few variants available in the market go beyond 150 km/charge.

COMPETITIVE LANDSCAPE

The Indian EV market is consolidated with the presence of major players in the market, owing to cheap and readily available manpower. However, established players in the market are introducing new models to gain a competitive edge over other players. For instance,

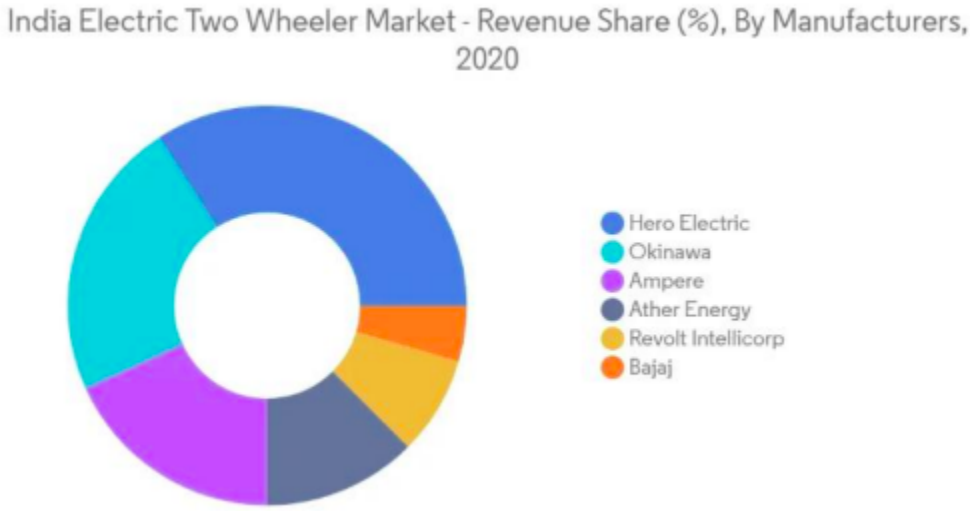
- In January 2020, Morris Garages Motor India launched its first electric internet SUV, and the car has a driving range of 340 km on a full charge.
- In 2019, Tata Motors announced its EV technology ZIPTRON, which will power all future Tata electric cars. It consists of a highly efficient permanent magnet AC motor, providing excellent performance. It will also offer a dust and waterproof battery system.

The startups are expanding their presence by raising funds from investors and tapping into new and unexplored cities. Companies are investing a tremendous amount in R&D and launching new models to mark their presence in the market.

OPTIMISTIC GROWTH FOR ELECTRIC BUSES AND TWO-WHEELER VEHICLES

India is also pushing hard for the electrification of buses. Many state governments have already started procuring electric buses from Chinese and local electric bus manufacturers. Many local bus manufacturers who are in collaboration with some Chinese manufacturers are trying to cater to the

rising demand for electric buses in India. With transportation still being a challenge in India, a lot of people in these segments look forward to the two-wheeler industry in India. As a result of the surging pollution, the national government has launched stringent policies to curb vehicular emissions. Furthermore, the availability of a considerable number of electric two-wheeler models, their low cost, as well as their availability as a substitute for conventional fuel-based vehicles. These aforementioned factors are fueling the demand in the Indian electric vehicle market.



Market Segmentation

As established in the beginning of this report, the Electric Vehicle market in India has just started to gain momentum, there are not a lot of statistics to provide an insight on Electric Vehicles consumers. So we changed our approach and we have collected consumer data on existing fuel-based vehicles and we would perform simple behavioral and demographic analysis on this data and try to understand the market. Next, for geographic analysis we have used some state-wise statistics to understand which region is most likely to be a good market for which type of Electric Vehicle.

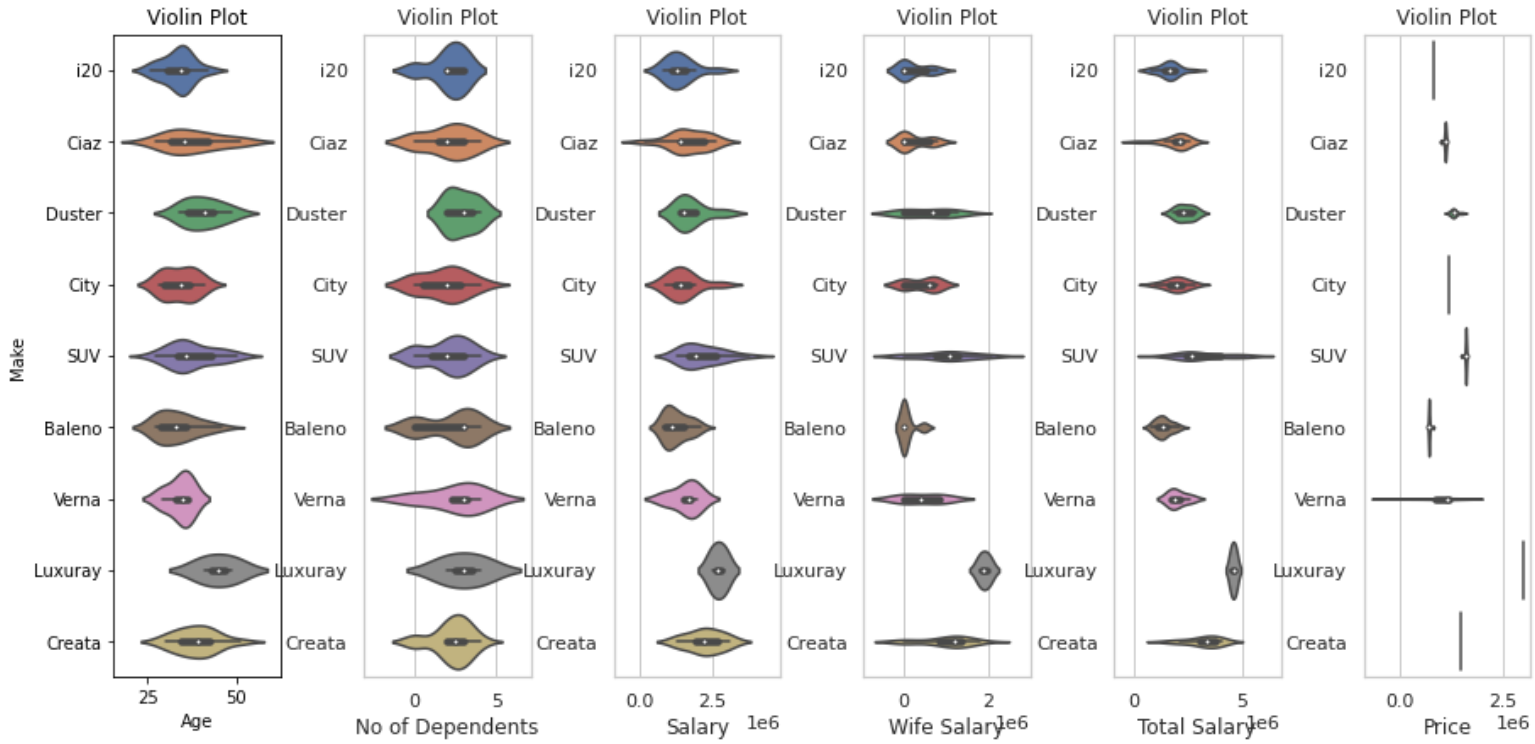
Following this analysis, we can understand important attributes of the segment we aim to target and use them for market segmentation using model-based algorithms.

BEHAVIORAL AND PSYCHOGRAPHIC ANALYSIS

Behavioral Segmentation is a form of customer segmentation that is based on patterns of behavior displayed by customers as they interact with a company/brand or make a purchasing decision. It allows businesses to divide customers into groups according to their knowledge of, attitude towards, use of, or response to a product, service or brand. Psychographic segmentation approach involves an understanding of a consumer's lifestyle, interests, and opinions. We have combined the two types of analysis because a consumer's lifestyle, interests and opinions are mirrored in their purchasing behavior.

The dataset we have used is a survey of people who own particular brands of fuel-based vehicles and it contains some basic information such as their age, salary, loan status, marital status, number of dependents, education, occupation and the make of their car and its price.

The violin plot below gives us some insight on the relation between the segmentation and descriptive variables in our data.

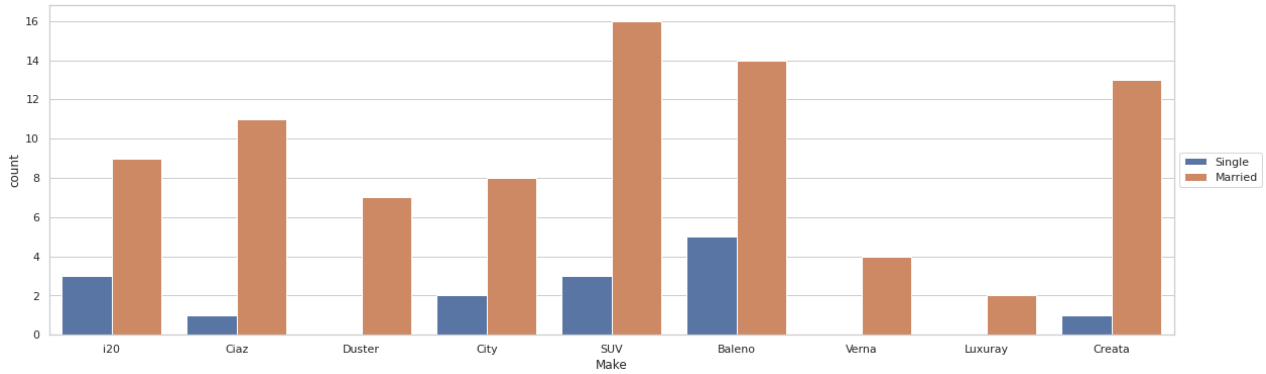


Observations:

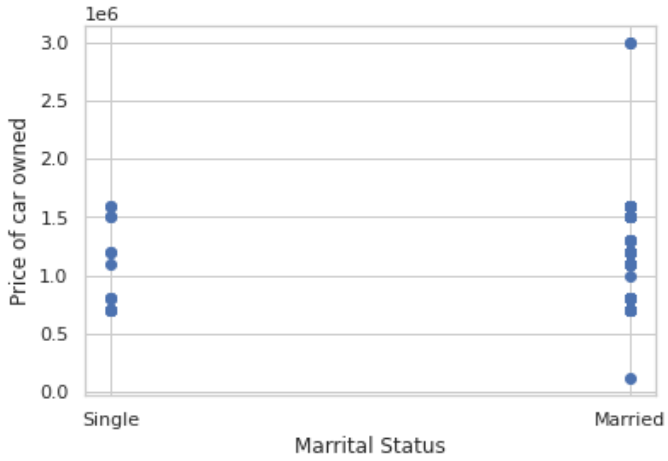
- Age: Younger consumers purchase less expensive vehicles. This can be explained simply as they have lesser dependents, lesser income and are single, and so they don't have both the option and the need to buy more expensive vehicles.
- Number of Dependents: Greater number of dependents makes the consumer buy a vehicle with more seats and so they tend to prefer SUVs.
- Salary: If you overlap the normalised salary plots with price plot, you would observe the median of salary violin plot matches that of the price of the vehicle indicating a very direct relationship, which makes sense as most people would buy vehicles they can afford.

Dependency of make and price of vehicles on other descriptor variables:

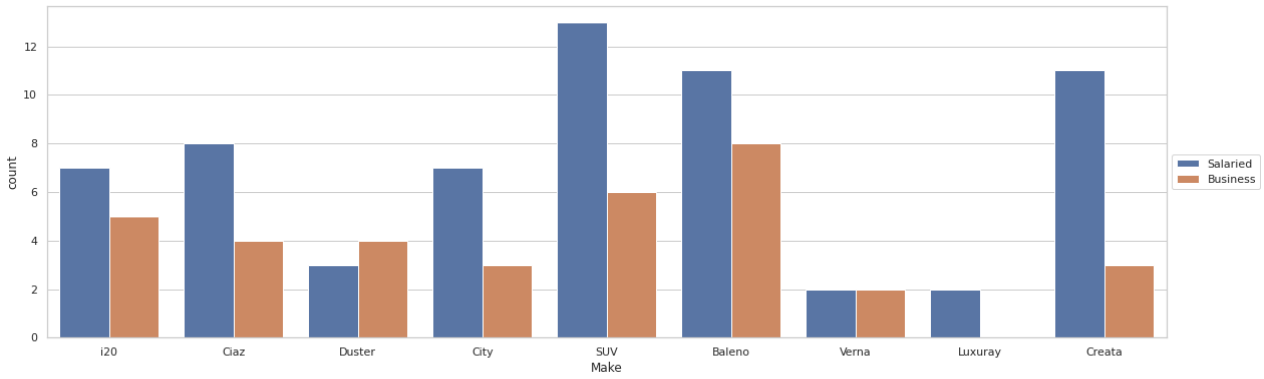
- Marital Status:
    - Make of vehicles they tend to purchase:



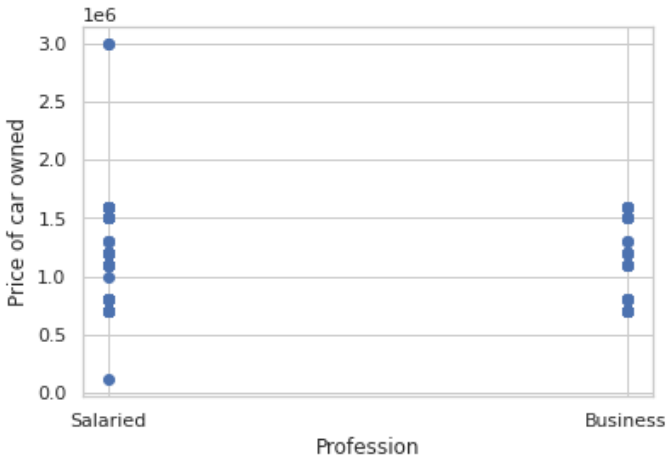
    - Price of vehicle owned:



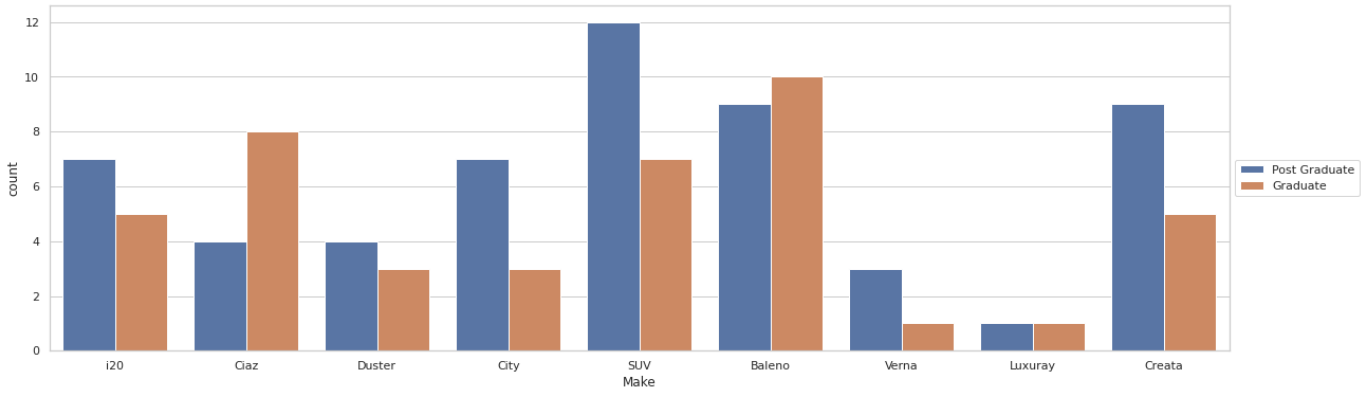
- Profession:
    - Make of vehicles they tend to purchase:



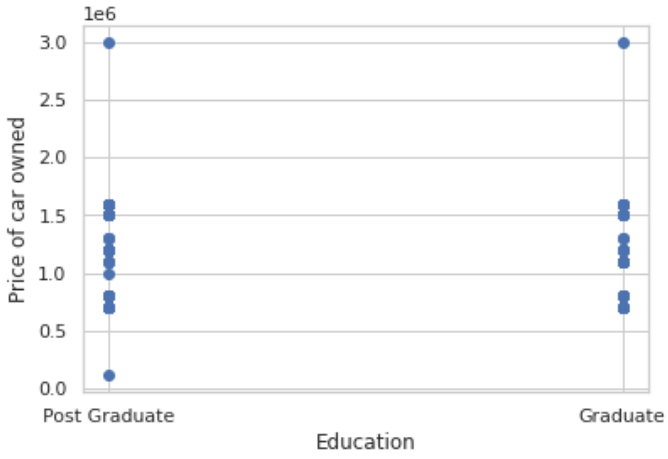
    - Price of vehicle owned:



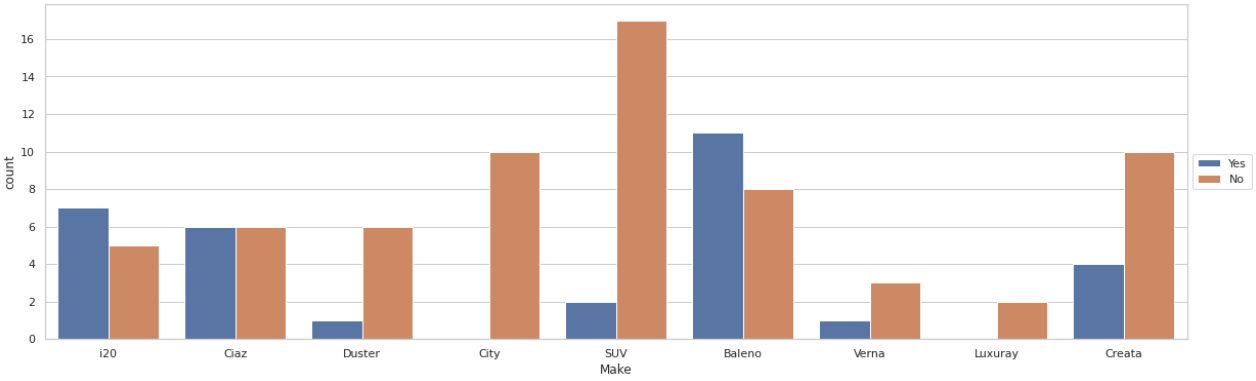
- Education:
    - Make of vehicles they tend to purchase:



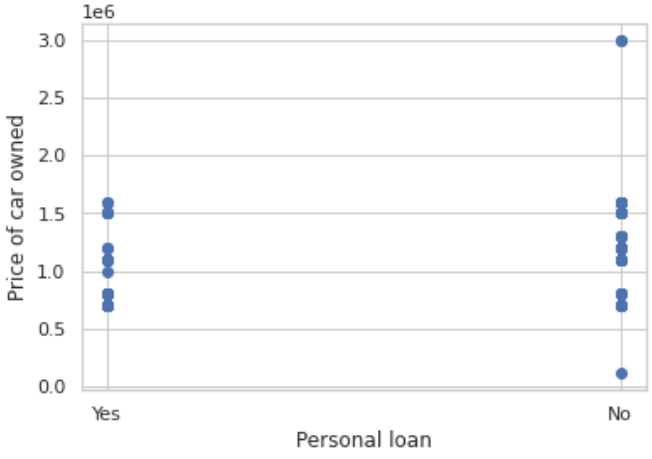
- Price of vehicle owned:



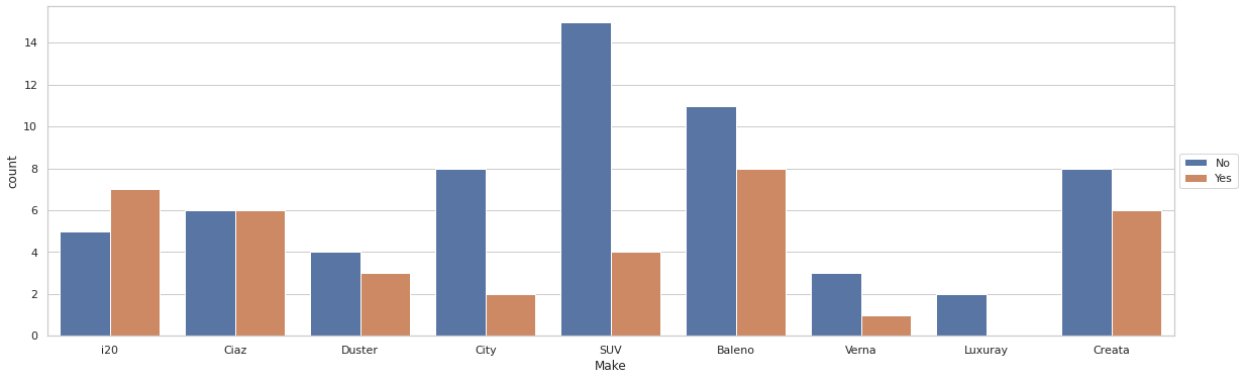
- <u>Personal Loan</u>:
  - Make of vehicles they tend to purchase:



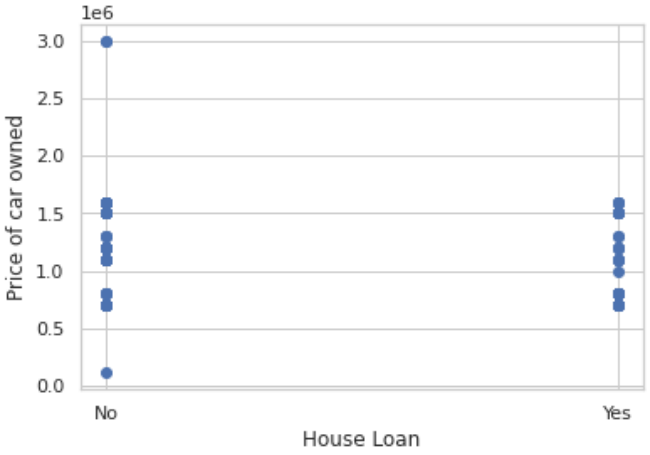
  - Price of vehicle owned:



- House Loan:
    - Make of vehicles they tend to purchase:



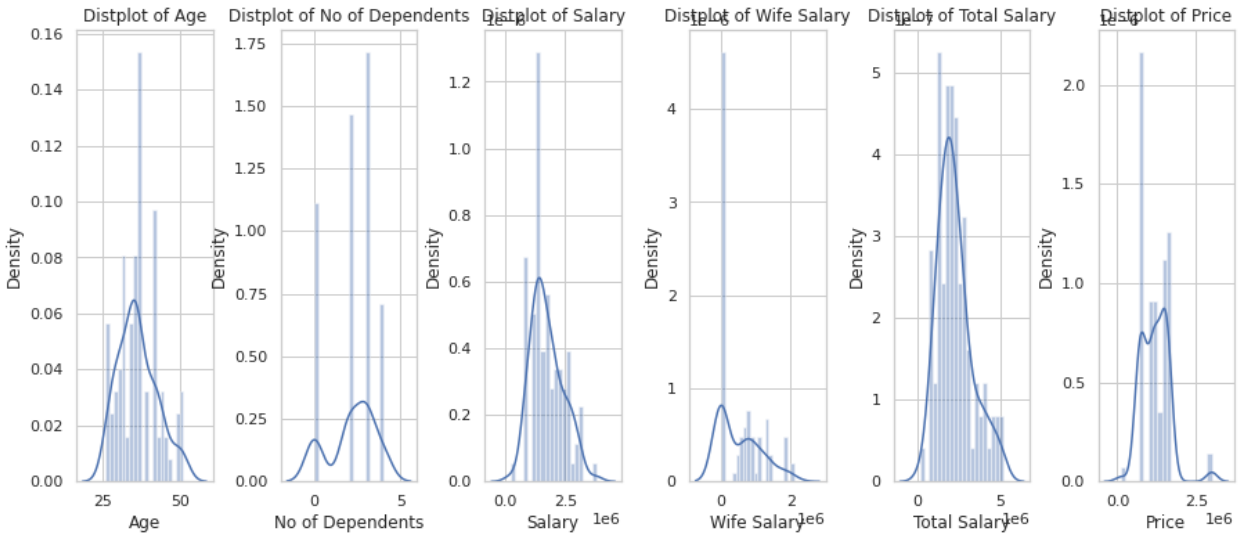
    - Price of vehicle owned:



DEMOGRAPHIC ANALYSIS

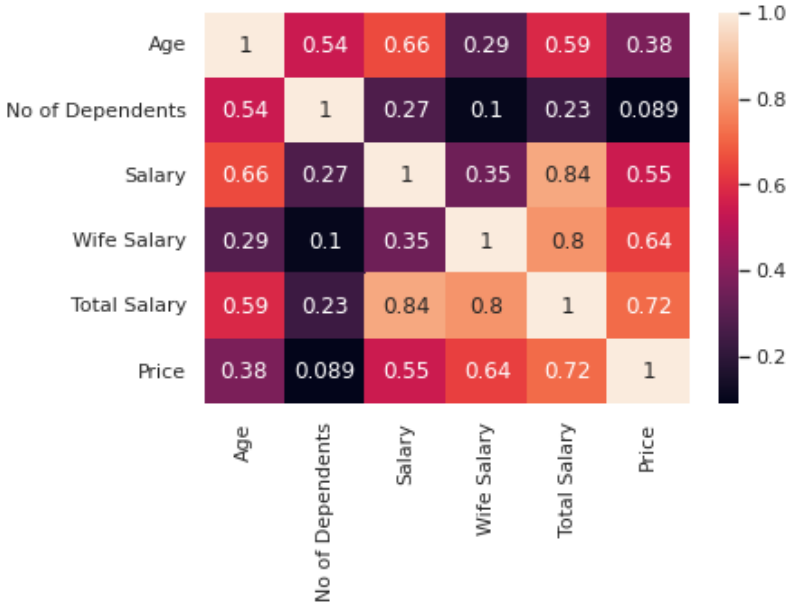
Demographic segmentation groups customers and potential customers together by focusing on certain traits such as age, gender, income, education, occupation and family status. Demographic segmentation is based on the assumption that consumers in the same demographic group will have similar needs. Demographic customer segmentation helps organizations to develop market outreach for better marketing strategies. When an organization looks at the demographic segmentation, it focuses on the people who are most likely to buy a product. This helps in identifying the target market.

We have used the same dataset we used for behavioral and psychographic analysis and the following plots help us understand the socio-demographic structure of the market:

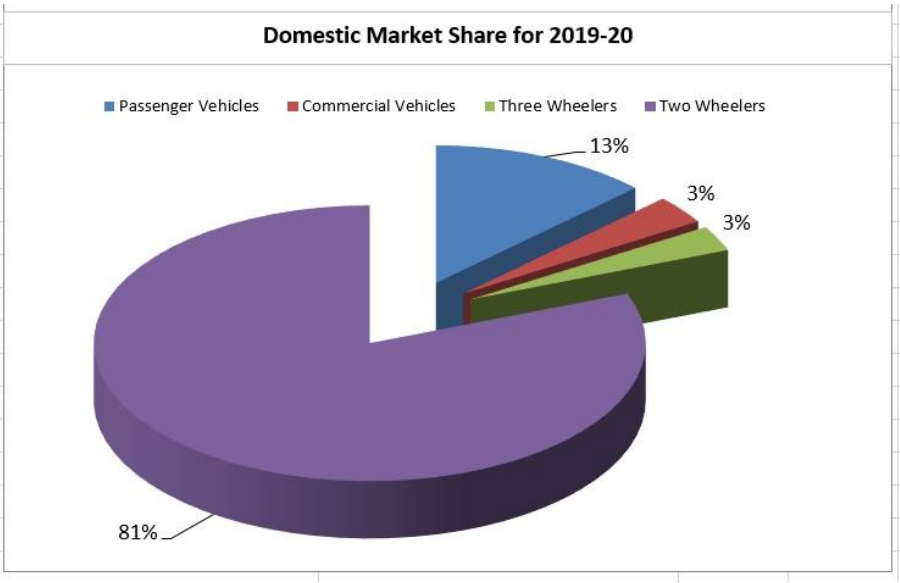


Observations from the distribution plot:

- People between the age group of 25 to 50 constitute most of the consumer market.
- Most people having an average total salary of around 30 lakh INR tend to purchase vehicles more.
- Most people spent around 10 to 20 lakh INR for vehicles.



Observations from the heatmap: There isn't any striking new relation found, but it confirms our previous observations.

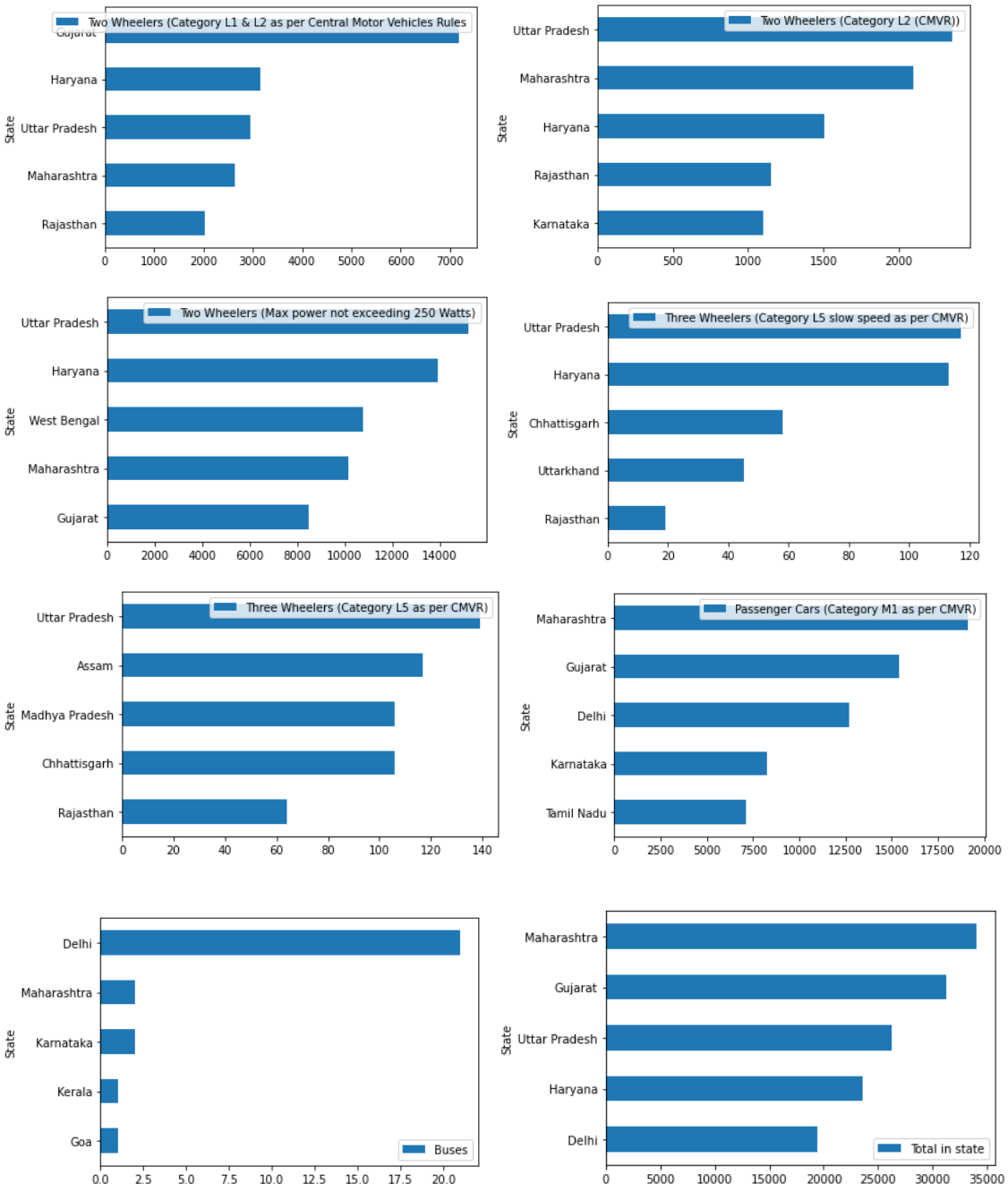


The pie chart above shows the domestic market share of automobiles in India, taken from the Society of Indian Automobile Manufacturers (SIAM) (see references). We can notice that the market share for two-wheelers in India is significantly huge compared to any other vehicle type. Therefore, for any electric vehicle startup it would be profitable if they focus on two-wheelers electric vehicles.

GEOGRAPHIC ANALYSIS

It is a component that competently complements a marketing strategy to target products or services on the basis of where their consumers reside. Division in terms of countries, states, regions, cities, colleges or areas is done to understand the audience and market a product/service accordingly. Here we have made divisions in terms of states and union territories in India.

For geographic analysis we used state-wise sales of different types of Electric Vehicles dataset which would help us understand our target region. Based on the type of electric vehicle, states with higher numbers of electric vehicles can be targeted as people in these states are more likely to purchase them. Given below are bar charts showing the top 5 states in sales of a particular EV type:



Depending on the type of Electric Vehicle the startup comes with, it can target that particular state. What is important to consider is that for most of these electric vehicles that market would be a fairly developed city in that state, because consumers should be willing to purchase the electric vehicle and factors like cost versus average consumer income and the resources to charge the EV (e.g. Charging Stations) and being able to maintain it are important.

Approaches Used For Segmentation

To perform market segmentation, we are using the population behavioral study where 100 people out of the entire population are selected and data relevant to our goal which is to know about the automobile purchase capability is noted. Since we are trying to find the ideal target segment for market penetration, we will classify the market into various segments. There are 2 general ways for classification: common sense classification and data-driven classification. Here we will be implementing one of the data-driven classification (i.e., K-Means Clustering).

ALGORITHM : K-MEANS CLUSTERING

K-means clustering is a type of unsupervised learning, which is used when you have unlabeled data (i.e., data without defined categories or groups). The goal of this algorithm is to find groups in the data, with the number of groups represented by the variable *K*. The algorithm works iteratively to assign each data point to one of *K* groups based on the features that are provided. Data points are clustered based on feature similarity. The results of the *K*-means clustering algorithm are:

1. The centroids of the *K* clusters, which can be used to label new data.
2. Labels for the training data (each data point is assigned to a single cluster)

The *'means'* in the K-means refers to averaging of the data; that is, finding the centroid.

The steps followed by the K-Means Clustering are :

1. Specify the desired number of segments k.
2. Randomly select k observations (consumers) from data set X and use them as the initial set of cluster centroids C = {c1, . . . , ck}.
3. Assign each observation xi to the closest cluster centroid to form a partition of the data, that is, k market segments S1, . . . ,Sk where

$$\mathcal{S}_j = \{\mathbf{x} \in \mathcal{X} | d(\mathbf{x}, \mathbf{c}_j) \leq d(\mathbf{x}, \mathbf{c}_h), \ 1 \leq h \leq k\}$$

This means that each consumer in the data set is assigned to one of the initial segment representatives. This is achieved by calculating the distance between each consumer and each segment representative, and then assigning the consumer to the market segment with the most similar representative.

4. Recompute the cluster centroids (segment representatives) by holding cluster membership fixed, and minimising the distance from each consumer to the corresponding cluster centroid.

$$\mathbf{c}_j = \arg\min_{\mathbf{c}} \sum_{\mathbf{x} \in S_j} d(\mathbf{x}, \mathbf{c}).$$

5. Repeat from step 3 until convergence or a pre-specified maximum number of iterations is reached. This is when the stepwise process of the partitioning algorithm stops and the segmentation solution is declared to be the final one.

LIBRARIES USED :

- NumPy : Scientific Computing Library
- Pandas : Data Analysis Library (mainly used to manage dataframes)
- Matplotlib : Data Visualization Library
- seaborn : Data Visualization Library
- scikit-learn : Machine Learning Library

IMPLEMENTATION

Data Sources:

We have taken the automobile buying behavior of Indian people as our dataset. With this dataset, we will find the various attributes related to the buying behaviors.

```
# Importing consumer buying behavior study dataset
df = pd.read_csv('https://raw.githubusercontent.com/Marisha18/Market-Segmentation-for-Electric-Vehicles-in-India/main/Indian%20automoble%20buying%20behavour%20study%20
df.head()
```

	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

We find that there are 14 attributes in the dataset (which includes categorical and numerical variables).

```
[ ]  df.columns
```

```
Index(['Age', 'Profession', 'Marrital Status', 'Education', 'No of Dependents',
       'Personal loan', 'House Loan', 'Wife Working', 'Salary', 'Wife Salary',
       'Total Salary', 'Make', 'Price'],
      dtype='object')
```

There are 2 types of categorical variables present in the dataset (i.e., binary variables and ordinal variables).

```
[ ]  # Observing unique value for object dtype columns
     for col in ['Profession','Marrital Status','Education','Personal loan','House Loan','Wife Working','Make']:
       print(col,':',df[col].unique())
```

```
Profession : ['Salaried' 'Business']
Marrital Status : ['Single' 'Married']
Education : ['Post Graduate' 'Graduate']
Personal loan : ['Yes' 'No']
House Loan : ['No' 'Yes']
Wife Working : ['No' 'Yes' 'm']
Make : ['i20' 'Ciaz' 'Duster' 'City' 'SUV' 'Baleno' 'Verna' 'Luxuray' 'Creata']
```

After finding the various categories, we find that the category ‘m’ in Wife Working is a human error. (Since it's a binary response attribute - either a wife can be working or not working).

Data Preprocessing :

After looking into the category, we realize that it can be classified under another category in the same attribute (i.e., No).

```
[ ]  df.loc[df['Wife Working'] == 'm']
```

	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	House Loan	Wife Working	Salary	Wife Salary	Total Salary	Make	Price
11	35	Salaried	Married	Graduate	4	Yes	Yes	m	1400000	0	1400000	Baleno	700000

After classifying it properly, we now look for null entries that might be present in the dataset. In our case, there are no such entries.

```
[ ]  ## Double checking the percentage of empty entries column wise
     df.isnull().sum() / df.shape[0] * 100.00
```

```
Age                0.0
Profession         0.0
Marrital Status    0.0
Education          0.0
No of Dependents   0.0
Personal loan      0.0
House Loan         0.0
Wife Working       0.0
Salary             0.0
Wife Salary        0.0
Total Salary       0.0
Make               0.0
Price              0.0
dtype: float64
```

Since we can't use categorical variables for K-Means Clustering, we will be encoding various attributes to a copy of the original dataset and use that for training the model.

```
[ ]  encoding = {"Profession":{"Salaried": 0, "Business": 1},
                 "Marrital Status":{"Single": 0, "Married": 1},
                 "Education":{"Graduate": 0, "Post Graduate": 1},
                 "Personal loan":{"No": 0, "Yes": 1},
                 "House Loan":{"No": 0, "Yes": 1},
                 "Wife Working":{"No": 0, "Yes": 1}
                 }
```

```
[ ]  obj_df = X.replace(encoding)
     obj_df.head()
```

	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	House Loan	Wife Working	Salary	Wife Salary	Total Salary	Price
0	27	0	0	1	0	1	0	0	800000	0	800000	800000
1	35	0	1	1	2	1	1	1	1400000	600000	2000000	1000000
2	45	1	1	0	4	1	1	0	1800000	0	1800000	1200000
3	41	1	1	1	3	0	0	1	1600000	600000	2200000	1200000
4	31	0	1	1	2	1	0	1	1800000	800000	2600000	1600000

Finally before implementing the K-Means Clustering algorithm, we will be scaling the entire dataset using the StandardScaler() function.

```
[ ]  X_scaled = StandardScaler().fit_transform(obj_df)
     X_scaled = pd.DataFrame(X_scaled,columns=['Age', 'Profession', 'Marrital Status', 'Education', 'No of Dependents',
                                               'Personal loan', 'House Loan', 'Wife Working', 'Salary', 'Wife Salary',
                                               'Total Salary','Price'])

     x = X_scaled.to_numpy()
     X_scaled
```

	Age	Profession	Marrital Status	Education	No of Dependents	Personal loan	House Loan	Wife Working	Salary	Wife Salary	Total Salary	Price
0	-1.498630	-0.739510	-2.366432	0.876275	-1.642313	1.446980	-0.772512	-1.051847	-1.397118	-0.887055	-1.406760	-0.904843
1	-0.211304	-0.739510	0.422577	0.876275	-0.136859	1.446980	1.294479	0.950708	-0.501877	0.108995	-0.258937	-0.445579
2	1.397855	1.352247	0.422577	-1.141195	1.368594	1.446980	1.294479	-1.051847	0.094950	-0.887055	-0.450240	0.013685
3	0.754191	1.352247	0.422577	0.876275	0.615867	-0.691095	-0.772512	0.950708	-0.203464	0.108995	-0.067633	0.013685
4	-0.854967	-0.739510	0.422577	0.876275	-0.136859	1.446980	-0.772512	0.950708	0.094950	0.441012	0.314975	0.932213
...	...	...	...	...	...	...	...	...	...	...	...	...
94	-1.498630	1.352247	-2.366432	-1.141195	-1.642313	-0.691095	-0.772512	-1.051847	0.990190	-0.887055	0.123671	0.932213
95	2.202434	-0.739510	0.422577	0.876275	0.615867	-0.691095	-0.772512	0.950708	3.079085	1.271054	2.706274	0.932213
96	2.363350	1.352247	0.422577	-1.141195	-0.136859	1.446980	1.294479	-1.051847	0.691777	-0.887055	-0.067633	-0.215947
97	2.363350	-0.739510	0.422577	0.876275	-0.136859	-0.691095	-0.772512	0.950708	1.437811	1.271054	1.654102	0.702581
98	2.363350	-0.739510	0.422577	0.876275	-0.136859	1.446980	1.294479	-1.051847	0.691777	-0.887055	-0.067633	-0.215947

99 rows × 12 columns

Model Deployment :

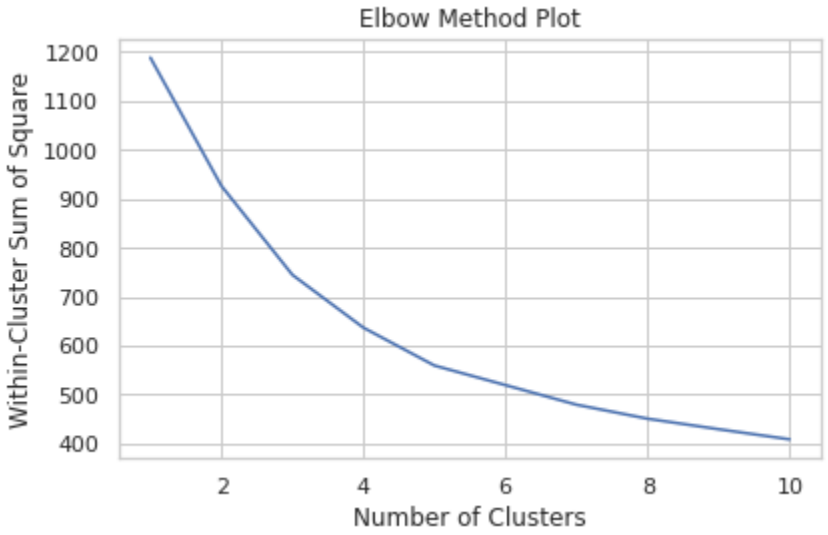
Initially we are trying to find the optimal K value using the Elbow Method wherein we will be finding the Within Cluster Sum of Square (WCSS) and try to find the point where it rapidly decreases which makes the graph look like an “elbow” there. The K value corresponding to that point is the optimal K value.

```
[44] wcss = []

     for i in range(1, 11):
         kmeans = KMeans(n_clusters = i, init = 'k-means++',
                         max_iter = 300, n_init = 10,random_state = 0)
         kmeans.fit(X_scaled)
         wcss.append(kmeans.inertia_)
```

```
plt.plot(range(1, 11), wcss)
plt.title('Elbow Method Plot')
plt.xlabel('Number of Clusters')
plt.ylabel('Within-Cluster Sum of Square') # Within cluster sum of squares
plt.tight_layout()
plt.show()
```

After looking into the plot, we can find that there are 2 points at which elbows are formed (which can be seen with the slight bent at K=3 and K=5). After finding the possible optimal K value, we will try to find the K value which provides us with the right clustering.



Therefore we will try to train K-Means Clustering by taking K = 3 and K = 5.

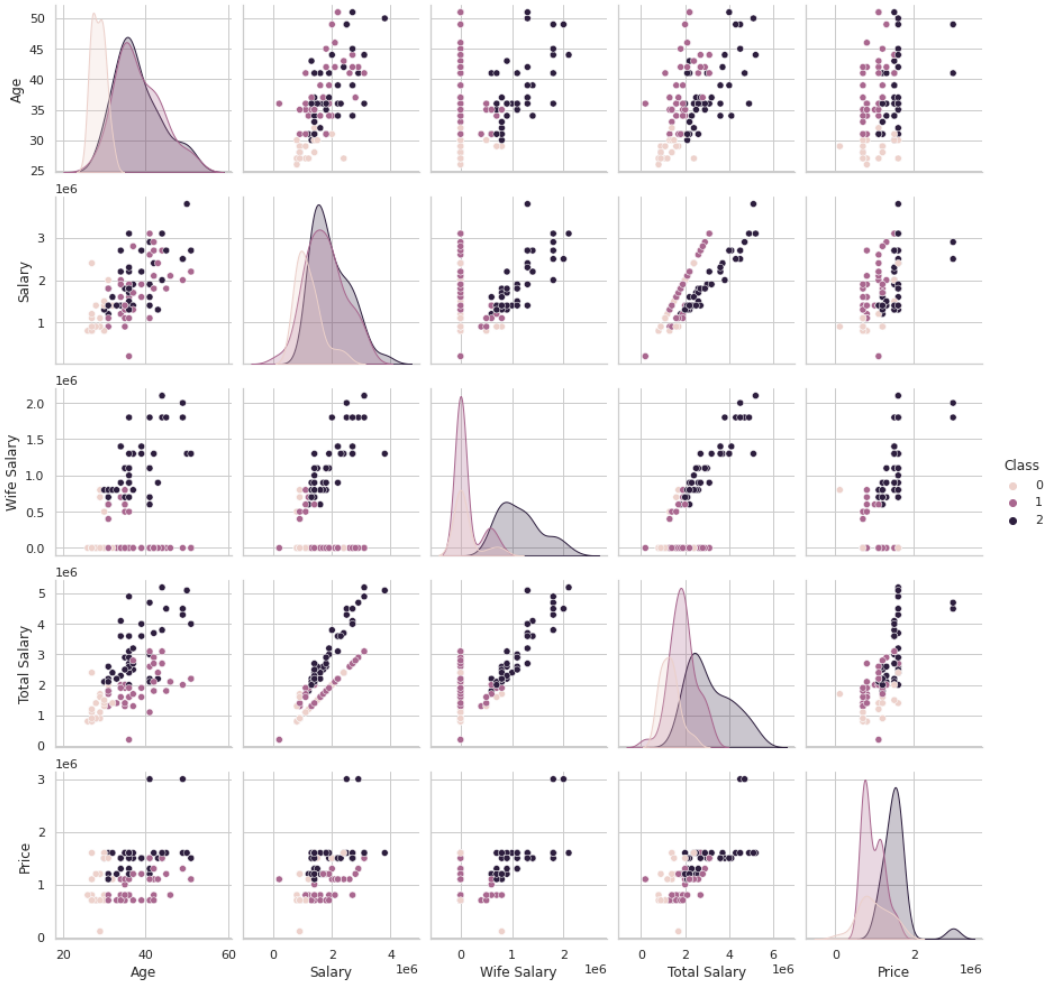
```
[49] kmeans = KMeans(n_clusters = 3, init = 'k-means++',
                     max_iter = 300, n_init = 10,random_state = 42)
     kmeans.fit(X_scaled)
```

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
       n_clusters=3, n_init=10, n_jobs=None, precompute_distances='auto',
       random_state=42, tol=0.0001, verbose=0)
```

```
[54] kmeans1 = KMeans(n_clusters = 5, init = 'k-means++',
                      max_iter = 300, n_init = 10,random_state = 42)
     kmeans1.fit(X_scaled)
```

```
KMeans(algorithm='auto', copy_x=True, init='k-means++', max_iter=300,
       n_clusters=5, n_init=10, n_jobs=None, precompute_distances='auto',
       random_state=42, tol=0.0001, verbose=0)
```

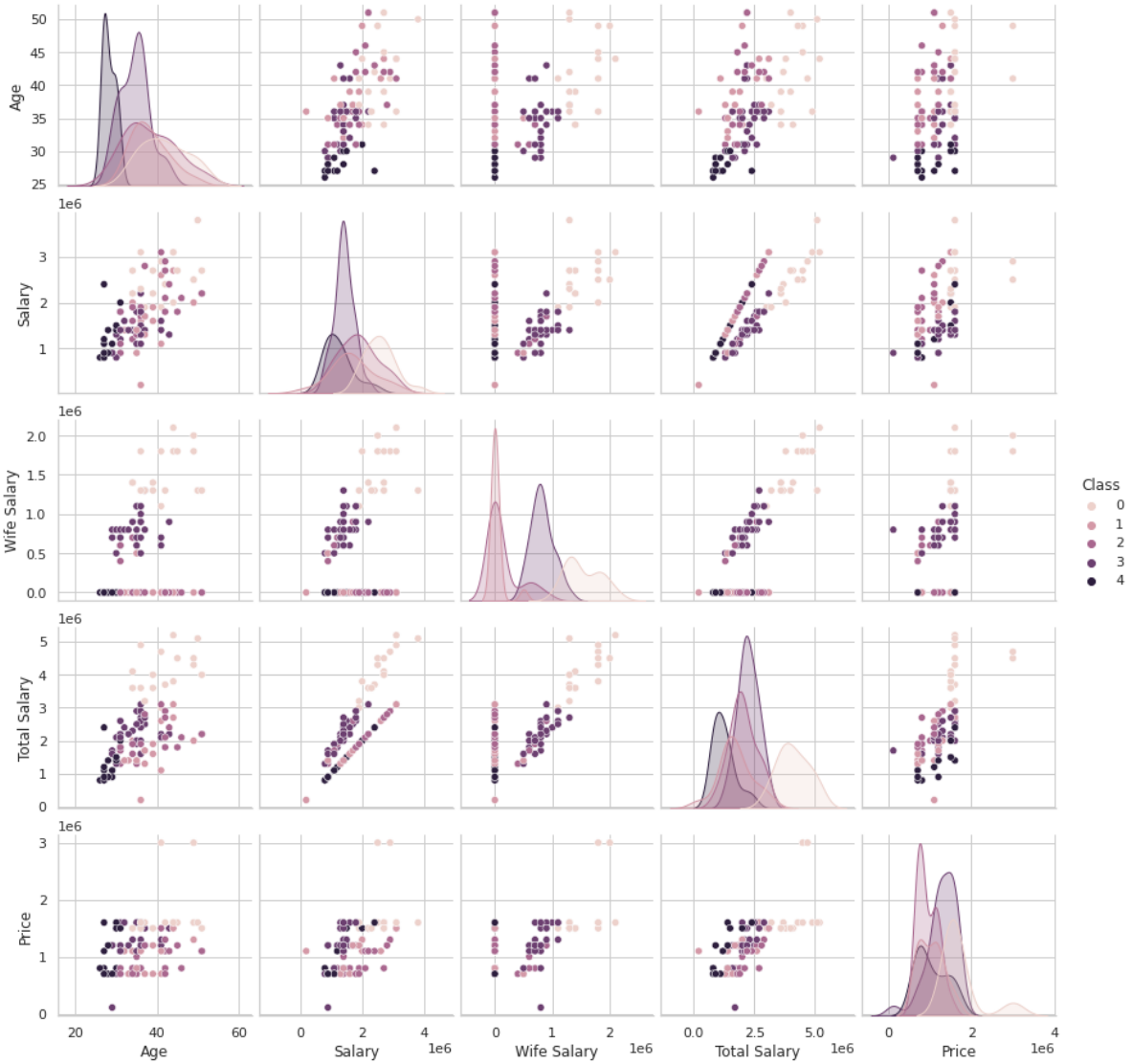
From the behavioral, psychographic, geographic and demographic analysis, we were able to see some attributes having an impact in the way of grouping consumers. However while looking into the clustered dataset, we were able to find 5 attributes that contributed most to the clustering (i.e., Age, Salary, Wife Salary, Total Salary, and Price). This can be seen in the pairplot done below for both K=3 and K=5 conditions.



K = 3

In this case, we are able to see the dataset being clustered naturally. From here we are also able to see that the model is trying to cluster people on the basis of their total income wherein :

- Class 1 are the group of people who have Total Salary close to Salary (husband's salary)
- Class 2 is the group of people who have Total Salary higher than Salary (husband's salary)
- Class 0 are the group of people who have Total Salary close to Salary (husband's salary) but their total salary is relatively less compared to other people.



K = 5

In this case, we are able to see the dataset is being clustered into very small groups of people that the model recognizes as a trend which is however not the case. We don't want to lose the homogeneity between our segments, therefore, going with K=3 will give the best results in the clustering analysis that is being done.

Target Segment

The younger population is more likely to purchase products with new technology, especially Electric Vehicles as they are aware of the environmental benefits and would like to bring that change, but our report showed that younger population buys less expensive vehicles and so Electric Vehicles not being affordable can be a downside. It is then suggested to target a segment which is still eager to try new technologies but financially well enough to be able to afford Electric Vehicles. These people are likely to be in an age-group of 30 to 40 years.

People from urban cities with available infrastructure and education about technology and its benefits will tend to purchase electric vehicles more.

People who are married and who have dependents are more likely to go ahead and purchase a vehicle and so they could be targeted.

Average salary of people who buy vehicles is around 30 lakh and the most purchases for automobiles lies in the range 10-20 lakh and lesser for two-wheelers. These aspects need to be kept in mind too.

Marketing Mix

Setting prices for our products is both an art and a science. Most importantly, you must know and understand your cost of production. From there you can adjust based on product characteristics, a specific pricing strategy, customer price sensitivity, customer values, and other factors. Marketing Mix helps understand what our product or service can offer to our customers and helps plan a successful product offering. Helps with planning, developing and executing effective marketing strategies. Help determine whether your product or service is suitable for your customers.

PRODUCT

The type of product would obviously depend on the EV Startup, but throughout our analysis we figured that for India it is best to enter the market with two-wheelers because the most automobile market-share is of two-wheelers. Most people would purchase a two-wheeler because it is cost effective, and the current infrastructure would support that.

Another type of product EV Startup can look into is public transport vehicles, because the current government policies are supportive for revamping public transport to electric-based engines.

PRICE

Affordability is a major issue with the growth of Electric Vehicles. It is important to keep in mind that in order to appeal to the consumers, the company's product has to be cost effective to both purchase and maintain. The product's price should ideally range between 10 to 20 lakh, as most people would make a purchase in this range.

PLACE

Infrastructure is another important aspect that has to be kept in mind while creating any product and launching it. Major urban cities of the country should be targeted as these are the places where infrastructure would support. Another reason for targeting urban cities is that here it is more likely to have an educated population willing to buy Electric Vehicles because they are aware of the environmental benefits.

For different types of vehicles, the list of top states which will promise a good market have been given in our geographical analysis.

PROMOTION

Promotion is product dependent. The best possible promotion is to educate people of the benefits of EV/HEV/PHEV over fuel-based vehicles. If the Startup comes up with an affordable product that should definitely be promoted.

References

- https://www.researchgate.net/publication/325801124
- https://www.mordorintelligence.com/industry-reports/india-electric-vehicle-market
- https://samples.mordorintelligence.com/69655/Sample%20-%20India%20Electric%20Vehicles%20Market%20(2020%20-%202025)%20-%20Mordor%20Intelligence.pdf
- https://jmkresearch.com/registered-ev-sales-in-india-in-2020-dropped-by-26-on-yoy-basis/
- https://jmkresearch.com/registered-ev-sales-drop-20-y-o-y-in-fy2021/
- https://www.siam.in/statistics.aspx?mpgid=8&pgidtrail=12

FOR DATASETS

- https://electricvehicles.in/electric-vehicles-sales-report-in-india-2018/
- https://www.kaggle.com/karivedha/indian-consumers-cars-purchasing-behaviour