Market Segmentation
of
ELECTRIC VEHICLES

BY
Ganesh Kumar

Fermi Estimation

- ★ Population of India: Approximately 1.3 billion people.
- ★ Number of households in India: Approximately 280 million households.
- ★ Percentage of households that own a car: Approximately 5% of households own a car, or 14 million cars in total.
- ★ Number of cars in India that are currently EVs: Approximately 1% of all cars, or 140,000 EVs.
- ★ Percentage of cars in India that are expected to be EVs by 2030: Assume a conservative estimate of 10%.
- ★ Number of EVs in India by 2030: 14 million cars x 10% = 1.4 million EVs.
- ★ Assuming an average lifespan of 10 years for a car, the annual demand for EVs in India would be approximately 140,000 cars per year (1.4 million cars / 10 years).
- ★ Considering the expected growth rate of the Indian economy and the increase in the middle-class population, we can assume that the demand for cars (both ICE and EVs) will increase in the coming years. Assume a conservative annual growth rate of 5%.
- ★ Therefore, the estimated demand for EVs in India by 2030 would be 140,000 cars x (1 + 5%)^9 = approximately 250,000 cars per year.

Data Sources

Electric & Alternative Fuel Vehicles US [2023]:
https://www.kaggle.com/datasets/nanduramesh/electric-vehicles
Cheapest Electric Cars:
https://www.kaggle.com/datasets/kkhandekar/cheapest-electric-cars
EVs - One Electric Vehicle Dataset - Smaller:
https://www.kaggle.com/datasets/geoffnel/evs-one-electric-vehicle-dataset?select=ElectricCarData_Norm.csv
Electric Vehicle Population:https:
//www.kaggle.com/datasets/ssarkar445/electric-vehicle-population

Electric Vehicle Population Data:
https://www.kaggle.com/datasets/ratikkakkar/electric-vehicle-population-data

EVPopulation:https:

//www.kaggle.com/datasets/vijayakishoredusi/evpopulation?select=Electric_Vehicle_Population_Data.csv

Electric Vehicle in India 2022:
https://www.kaggle.com/datasets/fathimaibrahimkunju/electric-vehicle-in-india-2022

Vehicle Registration India 17-22:
https://www.kaggle.com/datasets/prasenjitsharma/fuel-type-wise-vehicle-registration-india

Data Pre-processing

Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
```

Steps

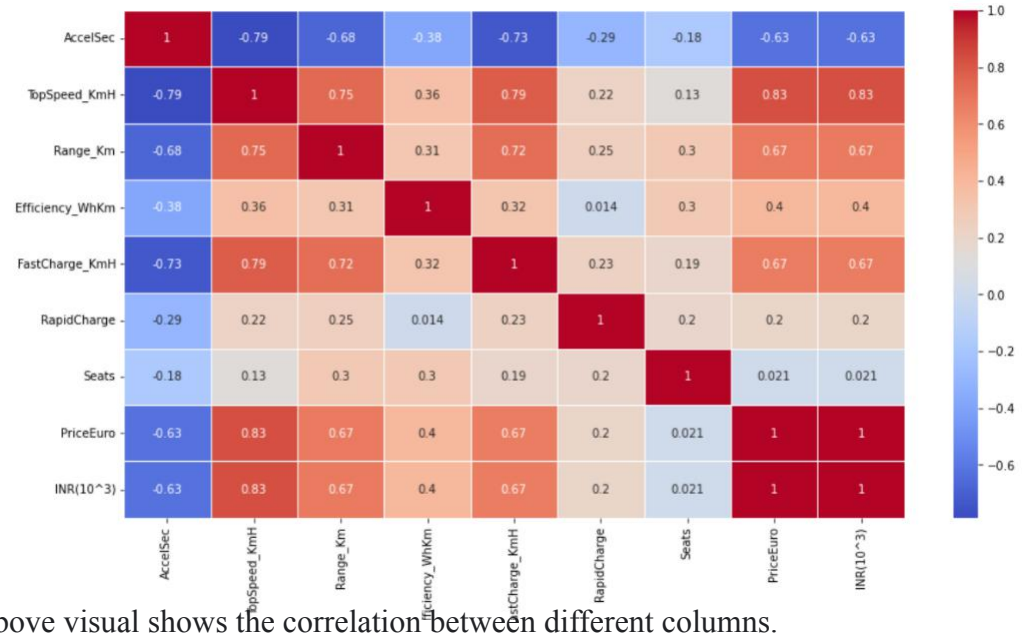
- Data collection: The first step is to identify the relevant data sources and collect the necessary data. This may involve scraping data from websites, using APIs, or working with databases.

- Data cleaning and preparation: The next step is to clean and prepare the data for analysis. This involves removing missing values, outliers, and any other data quality issues that could impact the analysis.

- Data exploration and visualization: Once the data is cleaned and prepared, the Data Analyst will explore and visualize the data to gain insights and identify patterns. This could involve creating histograms, scatter plots, heatmaps, and other types of visualizations.

```
sns.heatmap(df.corr(),linewidths=1,linecolor='white',annot=True, cmap ='coolwarm')
```

```
Out[27]: <AxesSubplot:>
```



- The above visual shows the correlation between different columns.

- Feature engineering: The next step is to identify and create relevant features that will be used in the machine learning model. This could involve transforming or combining existing features, or creating new ones from scratch.

- Model selection: After feature engineering, the Machine Learning Engineer will select an appropriate model for the problem at hand. This could involve selecting a linear regression model, decision tree, random forest, or other type of model.

```
model.summary()
```

OLS Regression Results

Dep. Variable:	INR(10^3)	**R-squared:**	0.721
Model:	OLS	**Adj. R-squared:**	0.704
Method:	Least Squares	**F-statistic:**	41.36
Date:	Sat, 18 Mar 2023	**Prob (F-statistic):**	1.57e-24
Time:	11:59:34	**Log-Likelihood:**	-904.10
No. Observations:	103	**AIC:**	1822.
Df Residuals:	96	**BIC:**	1841.
Df Model:	6		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	-9021.3591	2086.474	-4.324	0.000	-1.32e+04	-4879.741
AccelSec	153.3875	91.766	1.672	0.098	-28.767	335.542
Range_Km	3.1762	1.980	1.604	0.112	-0.754	7.107
TopSpeed_KmH	50.9030	7.014	7.257	0.000	36.981	64.825
Efficiency_WhKm	10.2960	6.152	1.674	0.097	-1.915	22.507
RapidCharge	128.2373	393.484	0.326	0.745	-652.822	909.296
PowerTrain	458.1352	258.671	1.771	0.080	-55.322	971.592

Omnibus:	84.867	**Durbin-Watson:**	2.060
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	741.645
Skew:	2.644	**Prob(JB):**	8.99e-162
Kurtosis:	15.036	**Cond. No.**	5.79e+03

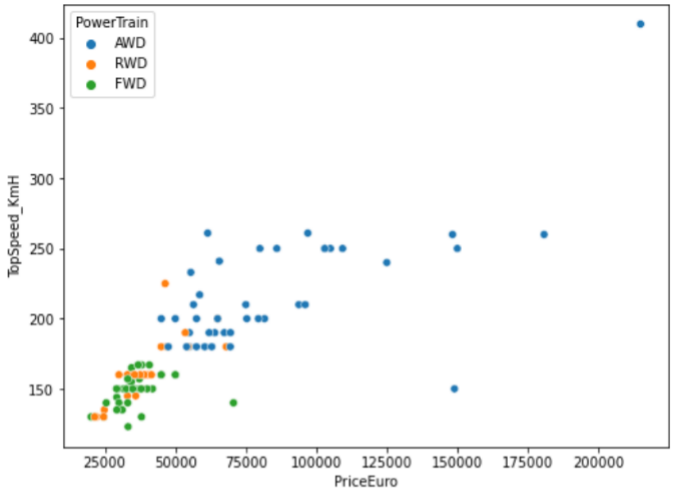
- Model training and evaluation: Once the model is selected, the Machine Learning Engineer will train the model on the prepared data and evaluate its performance. This could involve using techniques such as cross-validation and regularisation to optimise the model's performance.

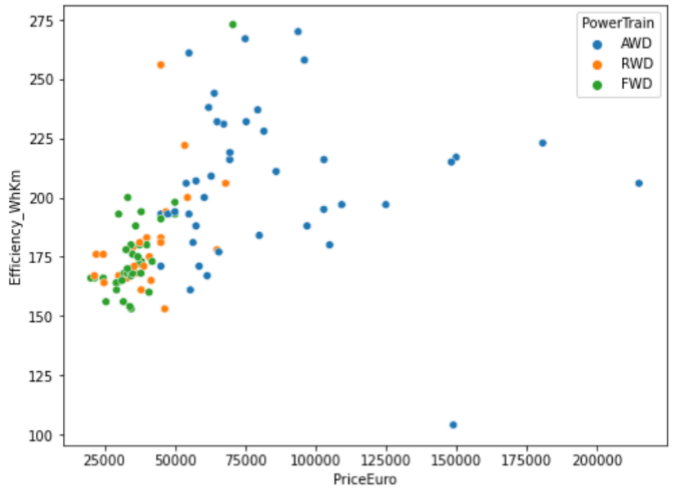
- Model deployment: After the model has been trained and evaluated, it can be deployed to make predictions on new data. This could involve deploying the model to a web service, or integrating it into an existing application.

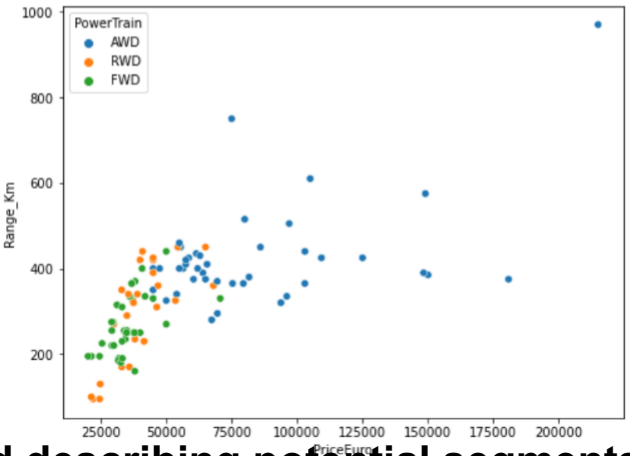
3. Segment Extraction

According to the analysis above, segment extraction is based on clustering methods, notably k-means clustering. K-means clustering is a popular unsupervised machine learning method for creating clusters of related data points based on their attributes. Range, efficiency, charging speed, and acceleration are among the variables related to electric vehicles that were most likely clustered in the aforementioned analysis.. Once the clusters were formed, the segment labels were assigned based on the characteristics of the vehicles within each cluster,

such as their size or intended use. This type of segmentation can be useful for understanding market trends and developing targeted marketing strategies for different customer segments.







4. Profling and describing potential segments

EV (Electric Vehicle) markets in India are set to explode due to a number of factors such as improved charging infrastructure, increased awareness about the potential for cost savings over the long run and availability of financing options. To ensure success in launching an EV product or service, it is important to identify the key market segments that could benefit from your offering.

The key segments identified as:

1) **Professionals** - High income earners seeking convenience and luxury; often preferring established brands with strong brand presence.

2) **Eco-Conscious Consumers** - Millennials who prioritise environmental sustainability when making a purchase decision; they prefer local eco-friendly companies which can customize EVs according to their needs.

3) **City Residents** - Value conscious individuals looking for low priced modes of transport within city limits while avoiding pollution caused by petrol/diesel running vehicles; usually willing to trade off comfort levels for price performance ratio .

4) **Rural Areas Residents** – Require economical means of transport covering rural distances at reasonable speeds without paying extra money on fuel refills & maintenance costs incurred during frequent repairsIn India, the potential markets for Electric Vehicle (EV) adoption are increasingly diverse. As with any target demographic, it is critical to first understand who those consumers might be in order to formulate initiatives that will effectively reach and engage them. Here we have provided a profile of three specific segments within India's EV market: cost-sensitive city dwellers, environmentally conscious commuters, and luxury buyers.

Cost Sensitive City Dwellers – These individuals primarily live in urban areas where public transportation options are often plentiful but unaffordable or inconvenient at times. They want an affordable electric car option that can provide convenience when needed without breaking their budget.

Environmentally Conscious Commuters – This segment wants an efficient mode of transportation but has strong concerns about environmental impact related to conventional gas cars as well as the sustainability of energy sources used by EVs.

Luxury Buyers – Generally affluent yet equally concerned about the environment compared to eco-conscious commuters, these vehicle adopters prioritise fashion over function which helps explain their higher rate of interest in luxurious EV models such as high performance sports cars or large SUV sedans geared towards family travel adventures. The electric vehicle market in India is a diverse and rapidly growing one. As such, there are many potential customer segments for EV manufacturers to target. The key to successful profiling of these groups is understanding the motivations that lead people towards electric vehicles. Three

main factors influence consumer decisions on whether or not adopt an EV: affordability, environmental benefits, and convenience/practicality.

Therefore, when profiling potential segments in Indian markets it's important to consider economic constraints as well as the types of features they are looking for from their EVs (e.g., range anxiety). Other important criteria include lifestyle interests such as eco-friendly commuting options, while also targeting age demographics with different financial means (18-35yo) who may care more about sustainability considerations than older consumers might be considered too. With this information nearby dealerships can better create campaigns tailored towards specific audiences and make sure their products reach niche buyers without breaking budgets by overspending on ineffective advertising efforts across broad swaths of customers likely unreceptive to expensive products like high end EVs

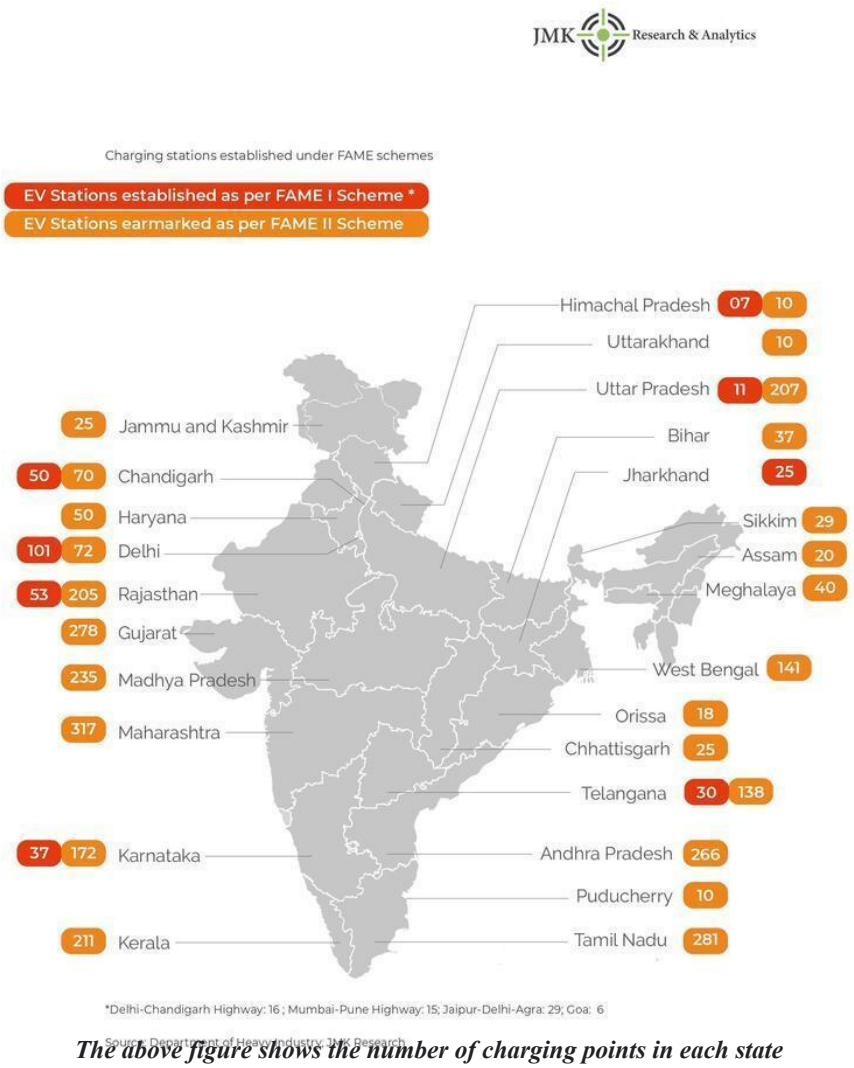
Behavioural Segmentation: Behavioural segmentation is a form of marketing segmentation that divides people into different groups who have a specific behavioural pattern in common. Users may share the same lifecycle stage, previously purchase particular products, or have similar reactions to your messages.

Benefits of Behavioural Segmentation:
• Improves targeting accuracy
• Helps provide better-personalised experience
• Sifts engaged users from uninterested
• Saves money
• Makes it easier to track success
• Helps build loyalty to your brand

Psychographic Segmentation: Psychographic segmentation's emphasis on characteristics like personality and values differs from demographic segmentation, which uses a specific trait (like gender, age, income, etc.) to categorise potential audiences. Market researchers use psychographic characteristics to help develop and position their products and marketing messages for different target groups Marketers use both demographics and psychographics in their market research to create their marketing strategy. So we will combine these both categories as well.

Demographic Segmentation: Demographic segmentation is a market segmentation technique where an organisation's target market is segmented based on demographic variables such as age, gender, education, income, etc. It helps organisations understand who their customers are so that their needs can be addressed more effectively. When an organisation 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 behavioural and psychographic analysis and the following plots help us understand the socio-demographic structure of the market:

Geographic Segmentation: Geographic segmentation is the process of dividing people into groups based on location, such as city, country, state, region, and even continent. It can help you tailor your approach during seasons customers may need your product. For example, a fisherman in Alaska may only buy more equipment leading up to the salmon season. Whereas a fisherman in Orange Beach, Alabama, might purchase equipment all year round. In contrast to other types of segmentation — demographic, psychographic, and behavioural — location-based segmentation analysis is easier to see results from. It doesn't take a lot of research to identify someone's location and the characteristics of a certain area, versus figuring out potential customers' purchasing behaviours and psychographics. 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.



The above figure shows the number of charging points in each state

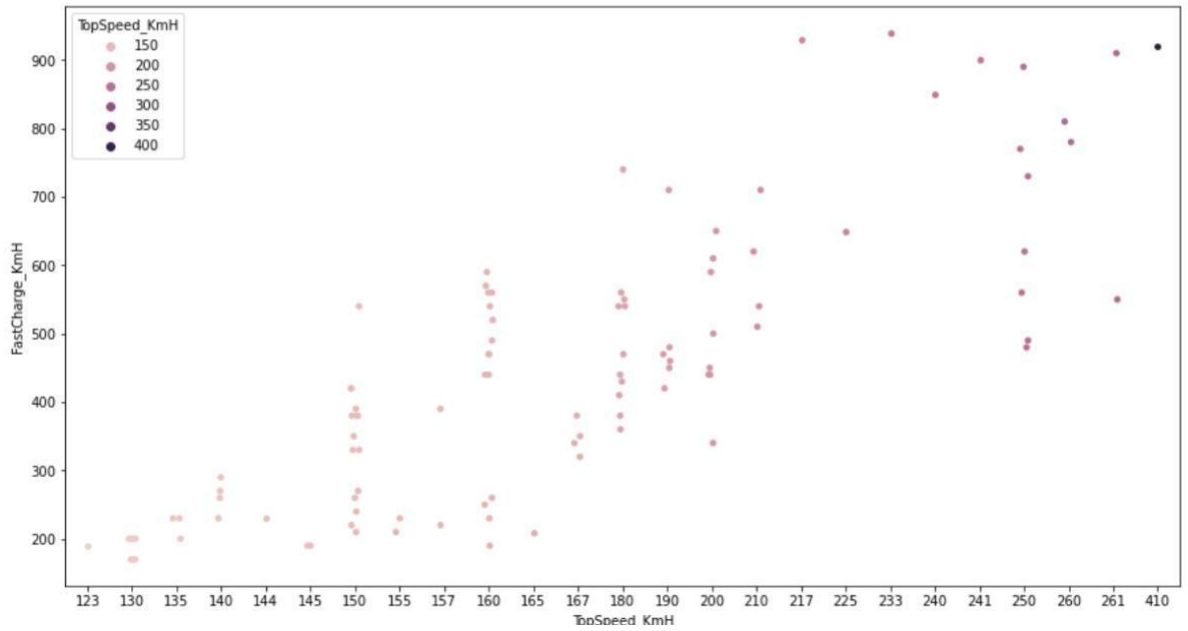
Model fitting:

K-Means Clustering Algorithm

Clustering

```
ax=plt.subplots(figsize=(15,8))
sns.stripplot(x='TopSpeed_KmH', y='FastCharge_KmH', data=df, jitter=True,hue = 'TopSpeed_KmH')
```

```
<AxesSubplot:xlabel='TopSpeed_KmH', ylabel='FastCharge_KmH'>
```



K-Means Clustering is an unsupervised learning algorithm that is used to solve the clustering problems in machine learning or data science. In this topic, we will learn what is K-means clustering algorithm, how the algorithm works, along with the Python implementation of k-means clustering.

It is an iterative algorithm that divides the unlabeled dataset into k different clusters in such a way that each dataset belongs only one group that has similar properties. It allows us to cluster the data into different groups and a convenient way to discover the categories of groups in the unlabeled dataset on its own without the need for any training. It is a centroid-based algorithm, where each cluster is associated with a centroid. The main aim oif this algorithm is to minimise the sum of distances between the data point and their corresponding clusters. The algorithm takes the unlabeled dataset as input, divides the dataset into k-number of clusters, and repeats the process until it does not find the best clusters. The value of k should be predetermined in this algorithm.

The k-means clustering algorithm mainly performs two tasks:

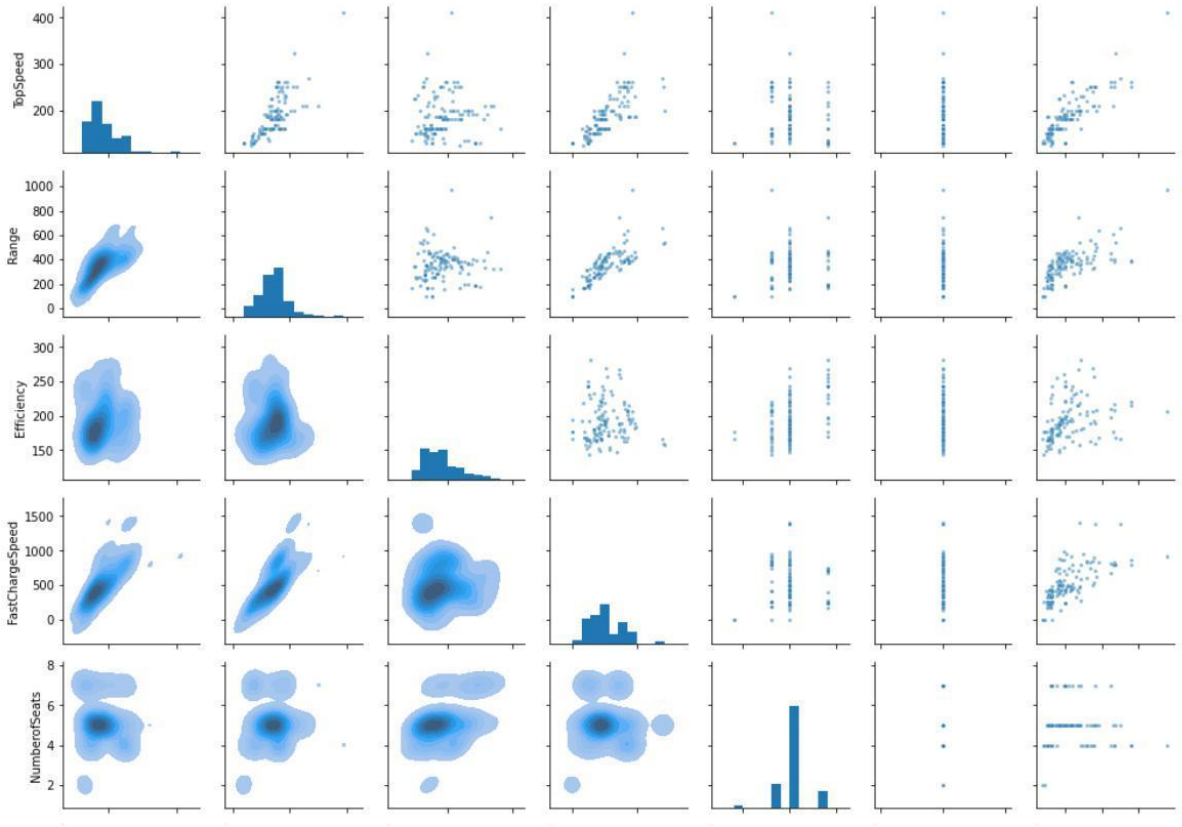
- o Determines the best value for K centre points or centroids by an iterative process.
- o Assigns each data point to its closest k-centre. Those data points which are near to the particular k-centre, create a cluster.

Data pre-processing:

The libraries used are:
- Numpy – For Computations
- Pandas – Manipulating the datasets
- Scikit-learn – For ML-based applications

With the help of this dataset, we can implement Behavioural, Psychographic and Demographic Segmentation of Indian Automobile Market. This helps us with understanding the various attributes leading to the consumer buying behaviour. After uploading the 'Indian automobile buying behaviour study 1.0' dataset, we inspect the dataset. It has 99 rows and 13 columns.
The various features include Age, Profession, Education, Number of dependents of the person buying the Vehicle and if the buyer is married or not, spouse is earning or not to ascertain the net income of the family. It further includes the Maker and the Price of the Vehicle of the buyer. We then use describe() function to gain in-depth insight about the mean, median and other statistical figures about the attributes of the dataset. This gives us an idea about the buyers' essential ranges. There aren't any missing values or irrelevant attributes and thus needs no data handling. In most of the algorithms, categorical values cannot be handled. Thus, there is a need to convert these categorical values to numerical levels.

ML algorithms work better when feature values are on relatively on a similar scale and close to normalised distribution. Using StandardScaler(), we scale the entire numerical portion of the dataset for an enhanced productive result.



Conclusion:

Top 3 states where we can sell our EV model are:
Maharashtra: It is the most richest state of India with gdp contribution of about 13.90% of all. People can easily buy the EV models here.
Tamil Nadu: It takes second spot in our table. People here are more eco-friendly. They would prefer these EV models due to its eco-friendliness.
Uttar Pradesh: It takes third spot in our table with gdp contribution of 8.35% of all. It is the most populous state of India. People of UP prefer cars much more than any other state of India. So EV model sales will be also very high here.
We can also reach out to Gujarat and Karnataka for the sales of EV models as these takes the 4th and 5th spot in gdp contribution of India.

It is better to kick start the marketing in these states to increase the sales.

In metropolitan areas, vehicular emissions are a significant source of air pollution.

At the breathing level of air quality, vehicular emissions typically contribute 20-30% of Particulate Matter (PM) 2.5.

PM2.5 refers to particles that are longer-lasting in suspension and have a diameter of fewer than 2.5 micrometres (more than 100 times smaller than a human hair).

Studies show that each year, cars generate roughly 290 gigagrams (Gg) of PM2.5.

In India, the transport industry is responsible for about 8% of all greenhouse gas (GHG) emissions, and in Delhi, that number approaches 30%.

Global Vehicle Emissions:

A quarter of all emissions come from the transportation sector, of which road transport is responsible for 75% of emissions (and 15% of all global CO2 emissions).

passenger cars