

# The Electric Vehicle Market analysis in India

Monali Jadhav

## 1.0 Introduction

Electric vehicles (EVs) represent a revolutionary form of transportation powered solely by batteries. With their minimal noise and environmental footprint, EVs have garnered widespread popularity worldwide.

In today's world, the significance of electric vehicles cannot be overstated. As pollution levels rise and the impacts of global warming become more pronounced, transitioning away from traditional energy sources is imperative. Our planet has endured considerable harm from our relentless exploitation of its resources. Oil pollution, stemming from the combustion of fossil fuels like gasoline and diesel, poses a significant threat to the biodiversity of our environment. In contrast, electric vehicles produce no emissions during operation, making them a cleaner alternative to conventional vehicles.

The widespread adoption of Electric Vehicles (EVs) hinges significantly on two critical factors: public awareness regarding EVs and their benefits, and the availability of a reliable and extensive charging network. These elements are pivotal in assuaging concerns related to range and charge anxiety, which often deter potential buyers from embracing EV technology. The charging infrastructure landscape in India has experienced significant expansion in recent years.

## 2.0 Objective

The primary objective of this project is to determine what are the different geographic regions in India where the EV market shows the most potential based on factors like infrastructure development and consumer interest. Specifically, the project aims to:

- Segment customers based on demographic and psychographic factors to understand the target market for electric vehicles.
- Geographic distribution of active electric vehicles compared to total vehicles.
- Evaluate the availability and distribution of EV charging infrastructure at the state level in India.
- Examine the availability of EV charging infrastructure at the city level in India
- Investigate the availability of EV charging infrastructure along highways and expressways in India.

## 3.0 Dataset

Each dataset appears to serve a specific purpose in understanding different aspects of the electric vehicle ecosystem in India.

Here's a brief overview of each dataset based on the provided information:

1. **EV\_Data.csv:** This dataset is used for customer segmentation based on various aspects such as demographic and psychographic factors. It likely contains information about customers who own or are interested in electric vehicles. The source is Kaggle.
2. **Active\_EVs.csv:** This dataset provides insights into the distribution of overall vehicles versus electric vehicles actively used in different parts of India. It helps in understanding the penetration of electric vehicles in the market. The source is Kaggle.

3. **Charging\_station\_statewise.xlsx**: This dataset showcases the state-wise allocation of charging stations by the government for electric vehicles. It helps in understanding the infrastructure availability for EV charging across different states in India. The source is data.gov.in.
4. **Charging\_station\_citywise.xlsx**: This dataset provides information about city-wise allocation of charging stations by the government for electric vehicles. It helps in understanding the infrastructure availability for EV charging at a more granular level, focusing on cities. The source is data.gov.in.
5. **Highwise\_charging\_station.xlsx**: This dataset offers insights into different EV charging ports available in various highways and expressways across India. It helps in understanding the availability of charging infrastructure along major transportation routes. The source is not explicitly mentioned, but it seems to be related to government data or publicly available information.

By combining and analyzing data from these diverse sources, one can gain comprehensive insights into the electric vehicle market landscape, including customer preferences, infrastructure availability, and geographic distribution of charging stations. This information can be valuable to promote the adoption of electric vehicles in India.

## 4.0 Data Pre-Processing

Data preprocessing is a crucial step in data preparation, involving various processes to refine raw data for further analysis. This preparatory phase includes exploring the data to understand its characteristics and identifying the most appropriate algorithms for extracting valuable market segments. On a technical level, data exploration helps in:

- Determining the measurement levels of the variables.
- Examining the distributions of individual variables.
- Evaluating relationships and dependencies among variables.

## 5.0 Exploratory Data Analysis (EDA)

### EV\_Data.csv:

Performing exploratory data analysis (EDA) on the provided dataset involves several steps to understand the data, its distributions, relationships, and potential insights.

Here are the steps:

#### 1. Understanding the Dataset:

- Started by loading the dataset into your preferred data analysis tool (e.g., Python with pandas).
- There are 1000 rows and 15 columns.
- Inspected the first few rows of the dataset to understand the structure and types of data.

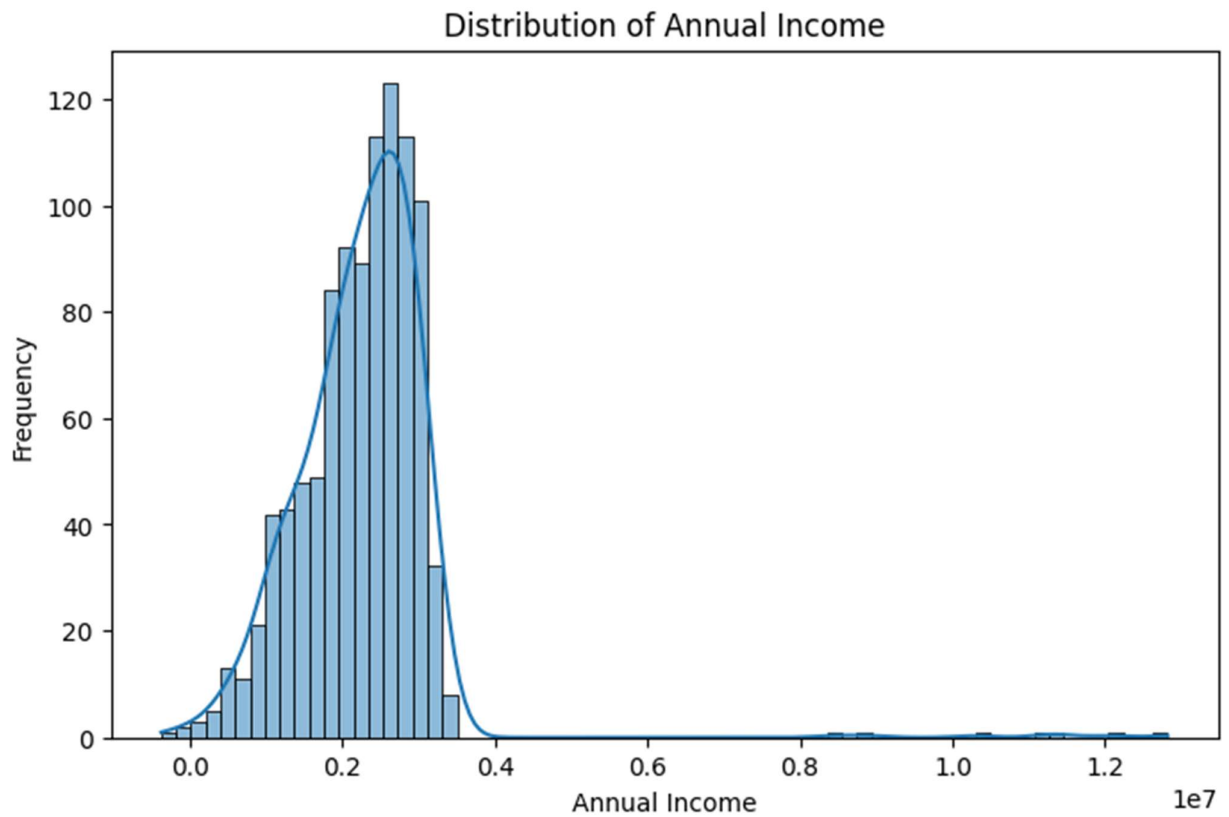
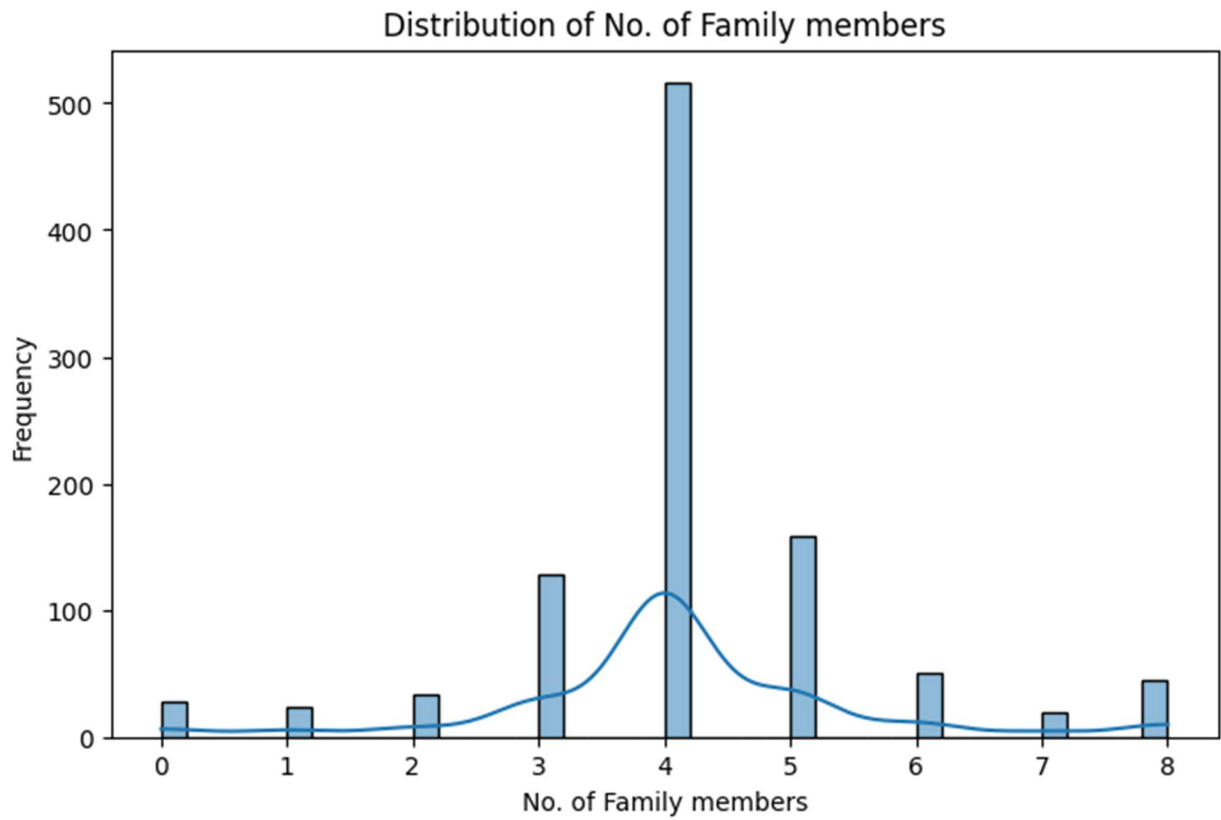
#### 2. Data Cleaning:

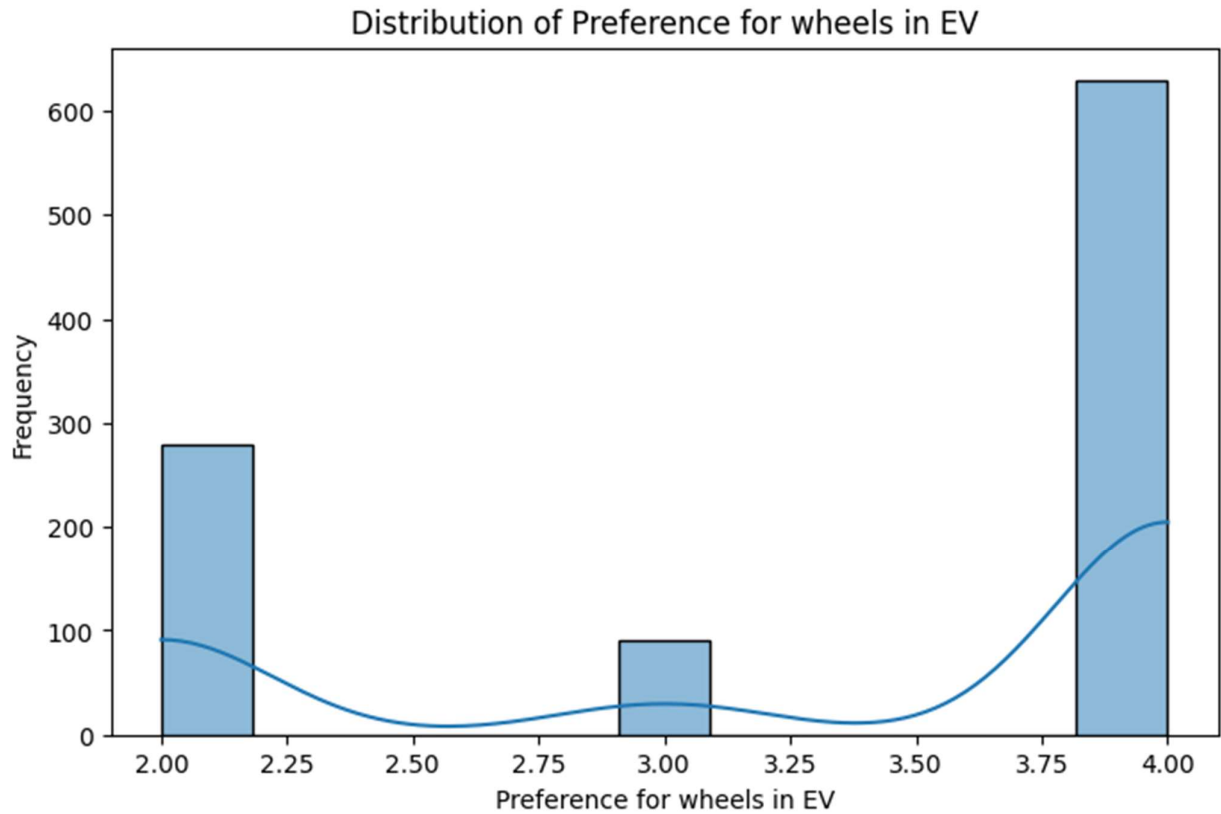
- Checked for missing values in each column, there are no null values.
- Dropped one column Unnamed as it is of no use.

#### 3. Univariate Analysis:

- For each variable (column), analyzed its distribution:
  - For numerical variables (e.g., Number of family members,

Annual Income), plotted histograms to visualize the distribution.

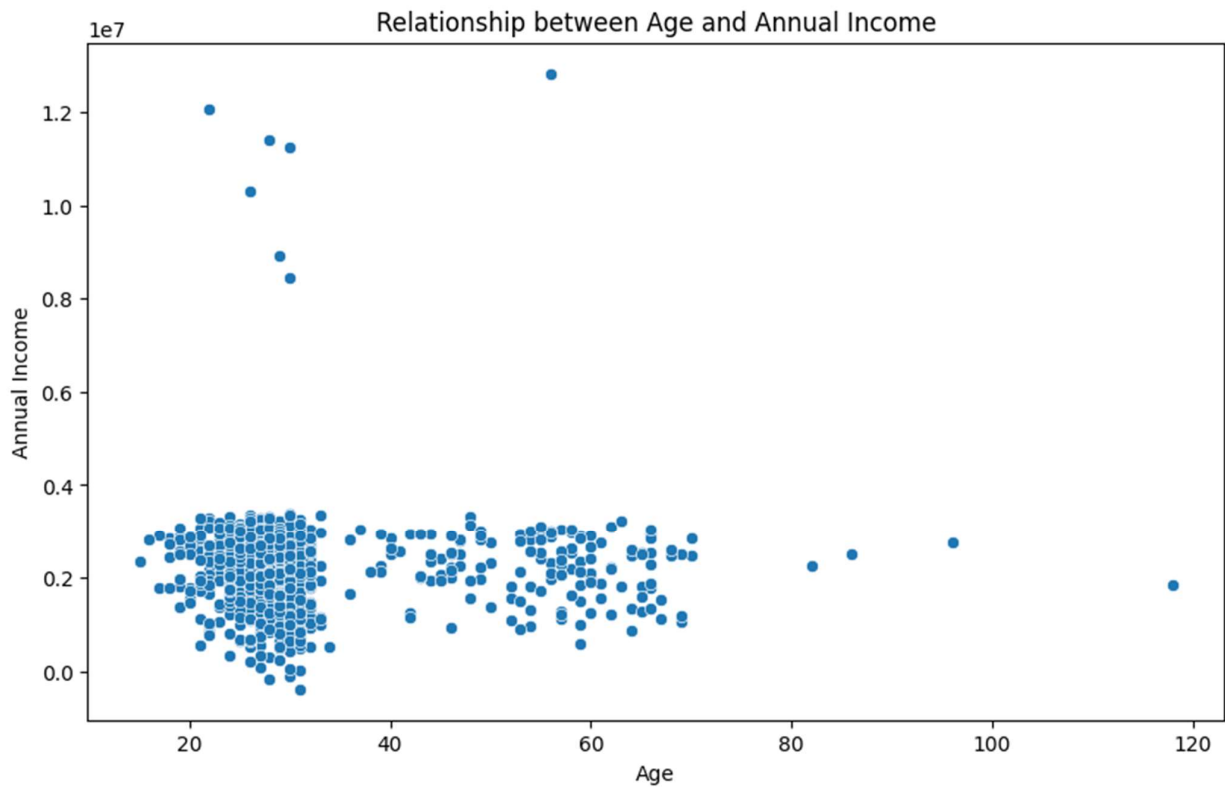




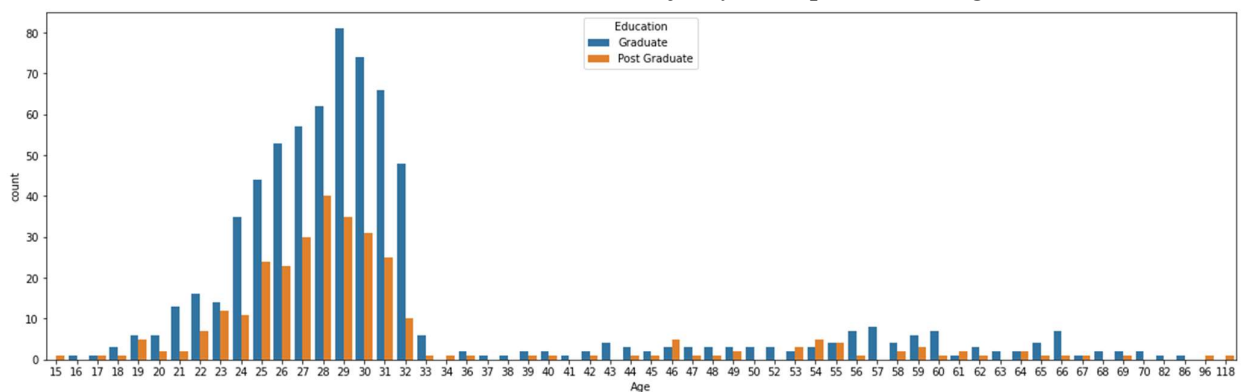
- For categorical variables (e.g., City, Marital Status), plotted bar charts to see the frequency of each category.
- Computed summary statistics (mean, median, mode, standard deviation, etc.) for numerical variables.

#### 4. Bivariate Analysis:

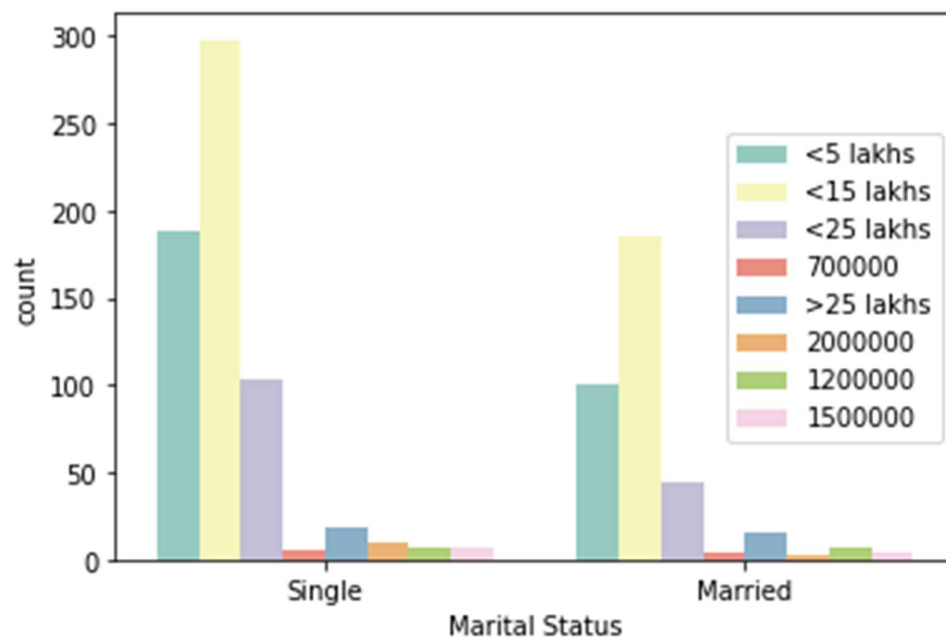
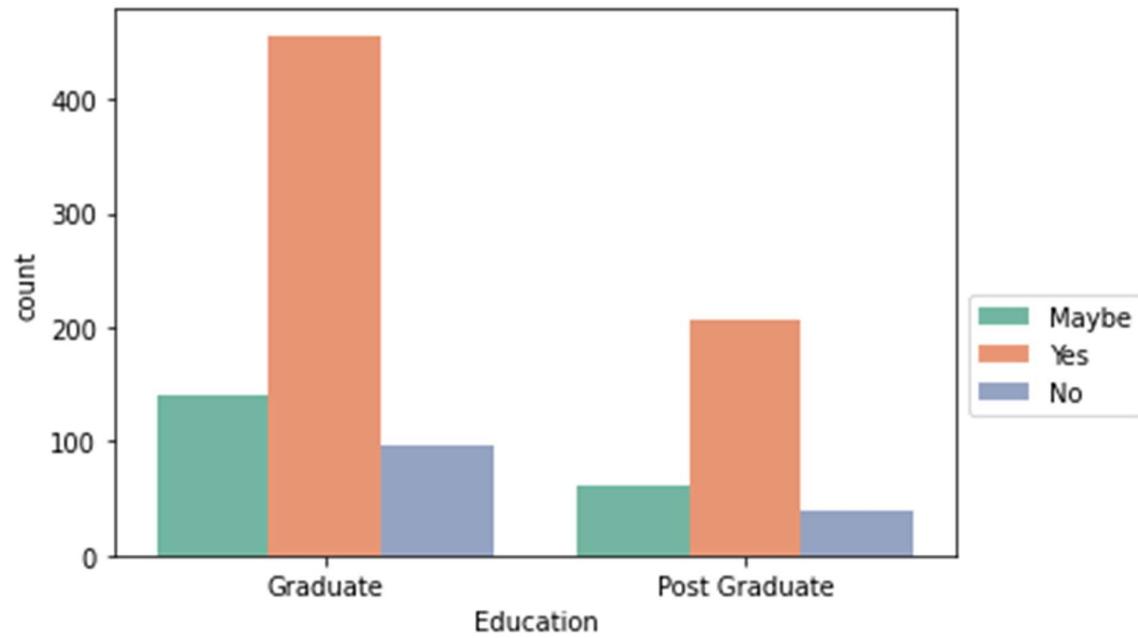
- Explored relationships between pairs of variables:
  - For numerical vs. numerical variables: used scatter plots or correlation analysis
    - Annual Income: The average annual income is around 2.3 million INR, with variations across respondents.
    - Age: The average age of respondents is approximately 29 years, with a minimum age of 15 and a maximum age of 118.



- For numerical vs. categorical variables: use bar plots:
  - Age: The average age of respondents is approximately 29 years, with a minimum age of 15 and a maximum age of 118
  - Education: The majority of respondents are graduates

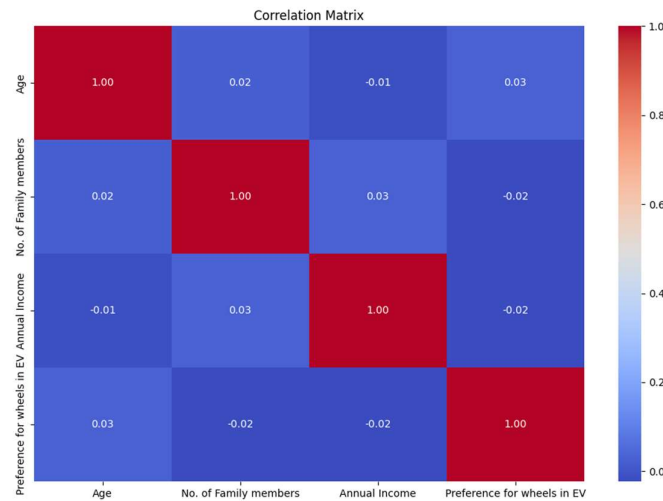


- For categorical vs. categorical variables: stacked bar charts:
  - Marital Status: A significant portion of respondents are single.
  - Education: The majority of respondents are graduates.



##### 5. Multivariate Analysis:

- Explore relationships between multiple variables simultaneously:
  - Used techniques like heatmaps for correlation analysis among numerical variables.



- Explored interactions between variables through advanced visualization techniques.

### Active\_EV.csv:

Performing exploratory data analysis (EDA) on the provided dataset involves several steps to understand the data, its distributions, relationships, and potential insights. Here are the steps:

#### 1. Understanding the Dataset:

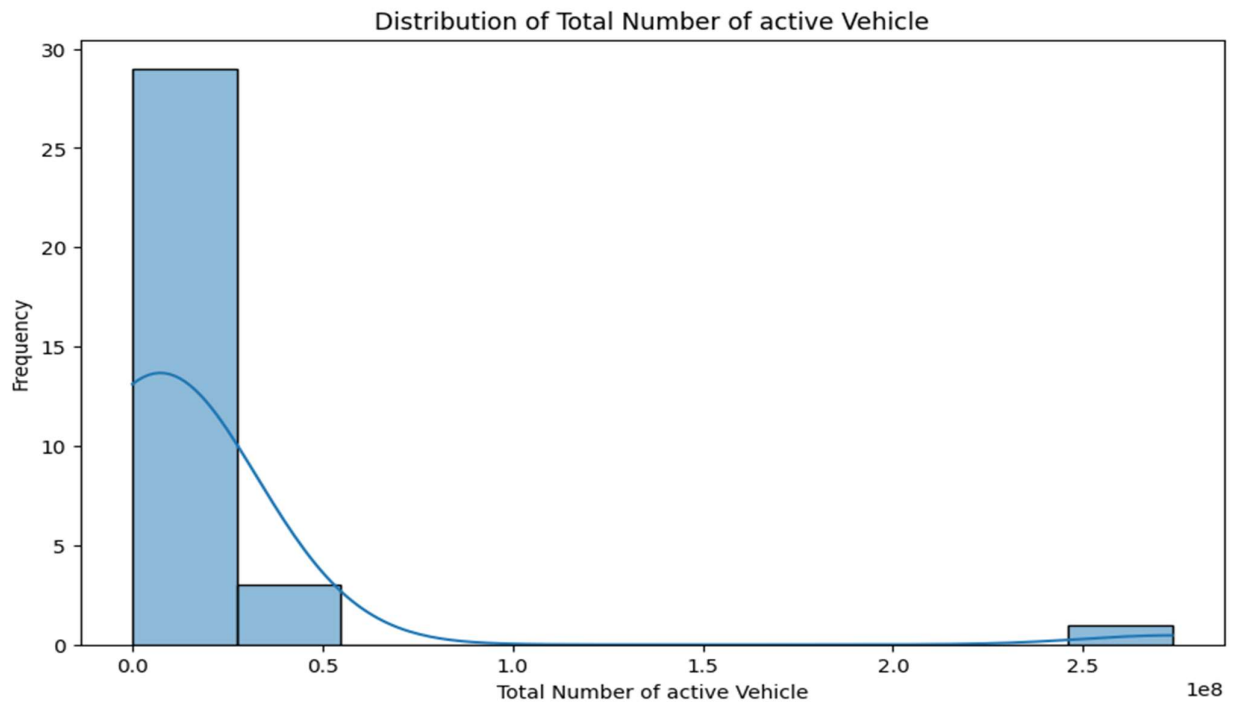
- Started by loading the dataset into your preferred data analysis tool (e.g., Python with pandas).
- Checked the dimensions of the dataset.
- Inspected the first few rows of the dataset to understand the structure and types of data.

#### 2. Data Cleaning:

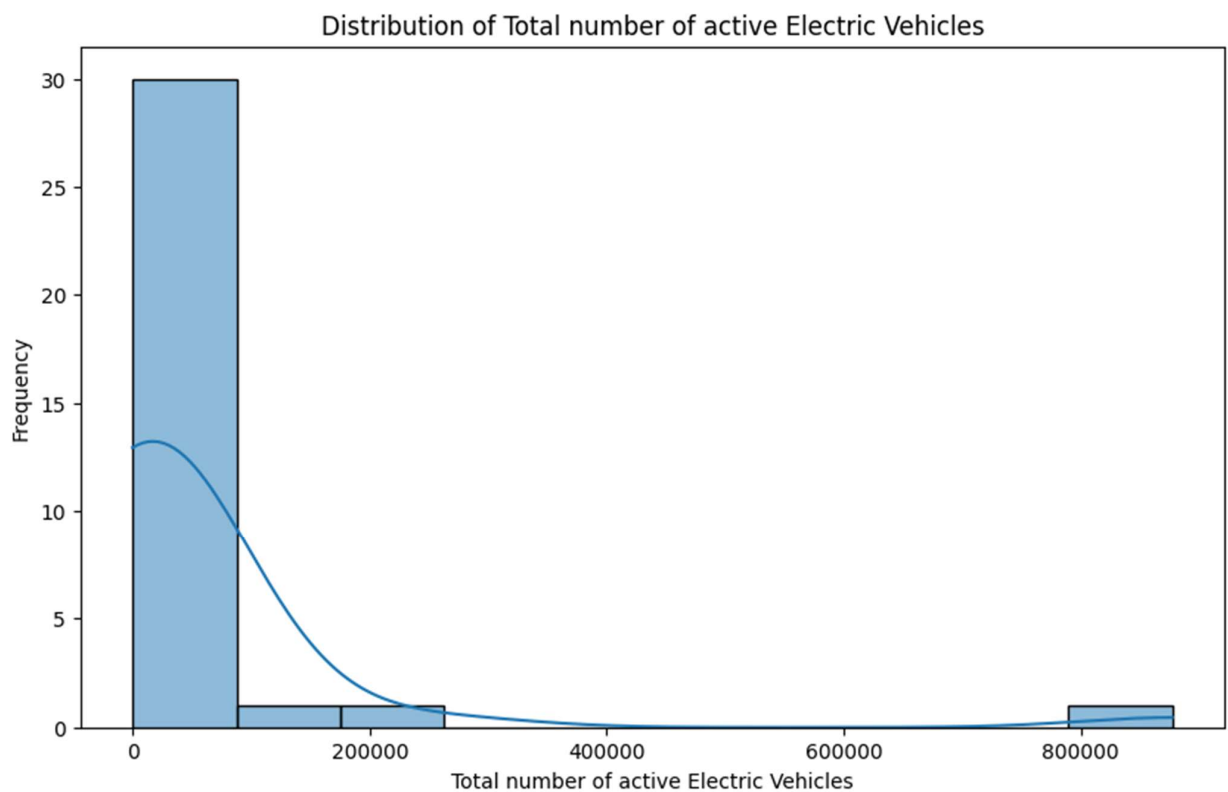
- Checked for missing values in each column and decide how to handle them (e.g., imputation or removal). There are no missing values.

#### 3. Univariate Analysis:

- For each variable (column), analyze its distribution:
  - For numerical variables (e.g., Total Number of active Vehicle), plotted histograms to visualize the distribution



- Plotted bar charts to see the frequency of Total number of active electric vehicles.



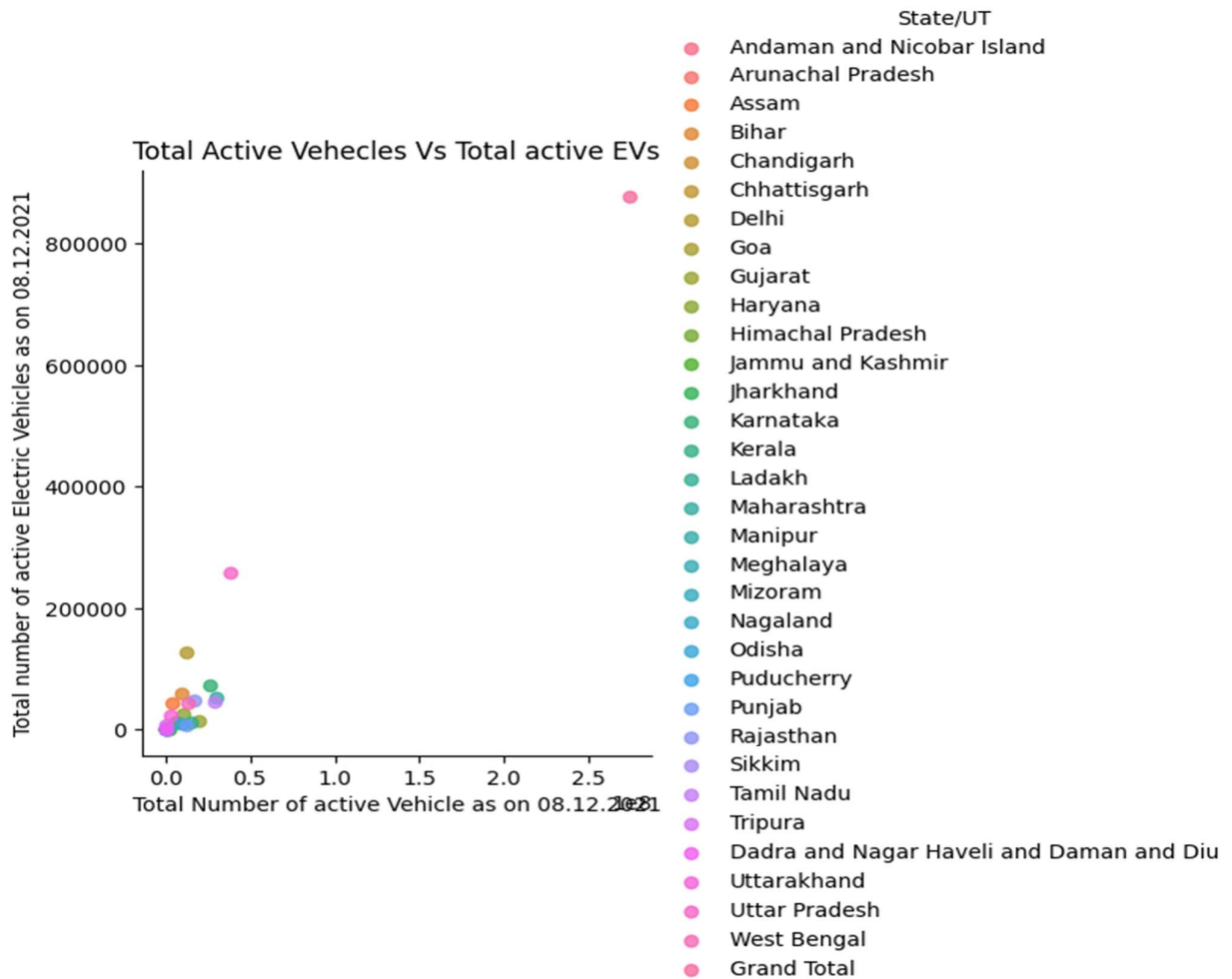
- Compute summary statistics (mean, median, mode, standard deviation, etc.) for numerical variables.

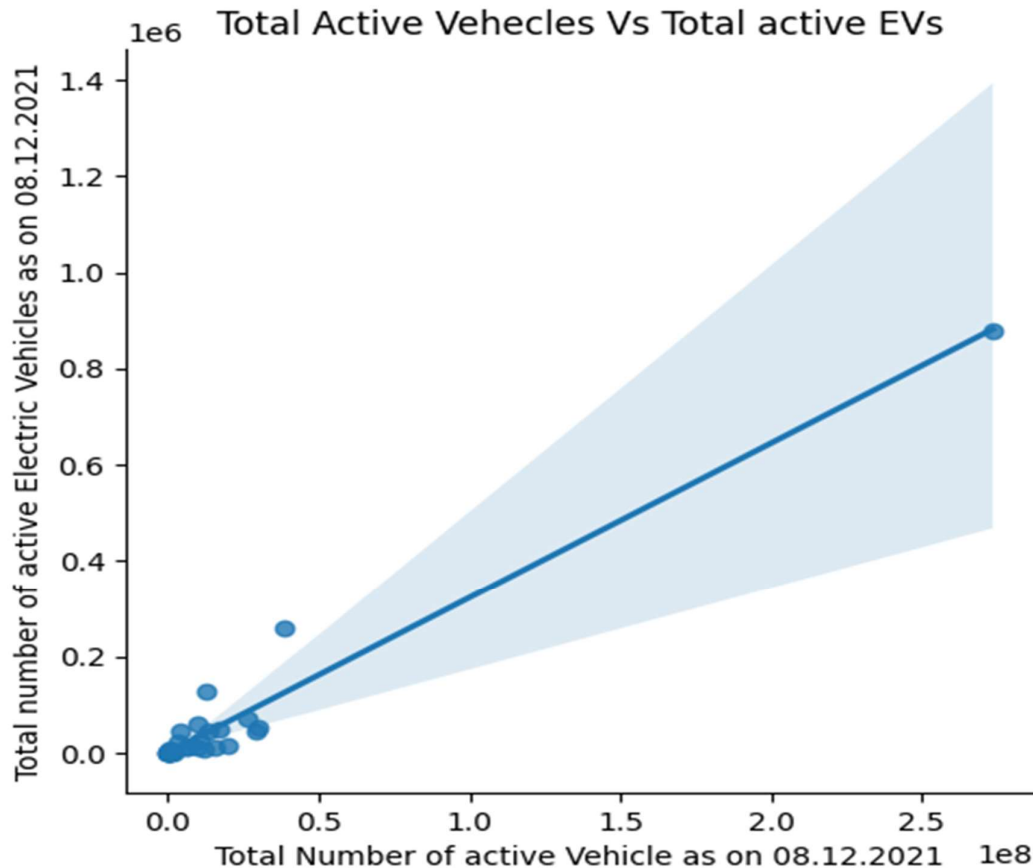
#### 4. Bivariate Analysis:

- Explore relationships between pairs of variables:
  - For numerical vs. numerical variables: use scatter plots or correlation analysis.



- States like Delhi, Maharashtra, Karnataka, Tamil Nadu, and Uttar Pradesh have relatively higher numbers of active electric vehicles compared to others.
- On the other hand, states like Arunachal Pradesh, Ladakh, Meghalaya, Mizoram, and Sikkim have significantly lower numbers of electric vehicles, indicating lower adoption rates and potential challenges in EV infrastructure development.





##### 5. Correlation between Total Vehicle Population and EV Adoption:

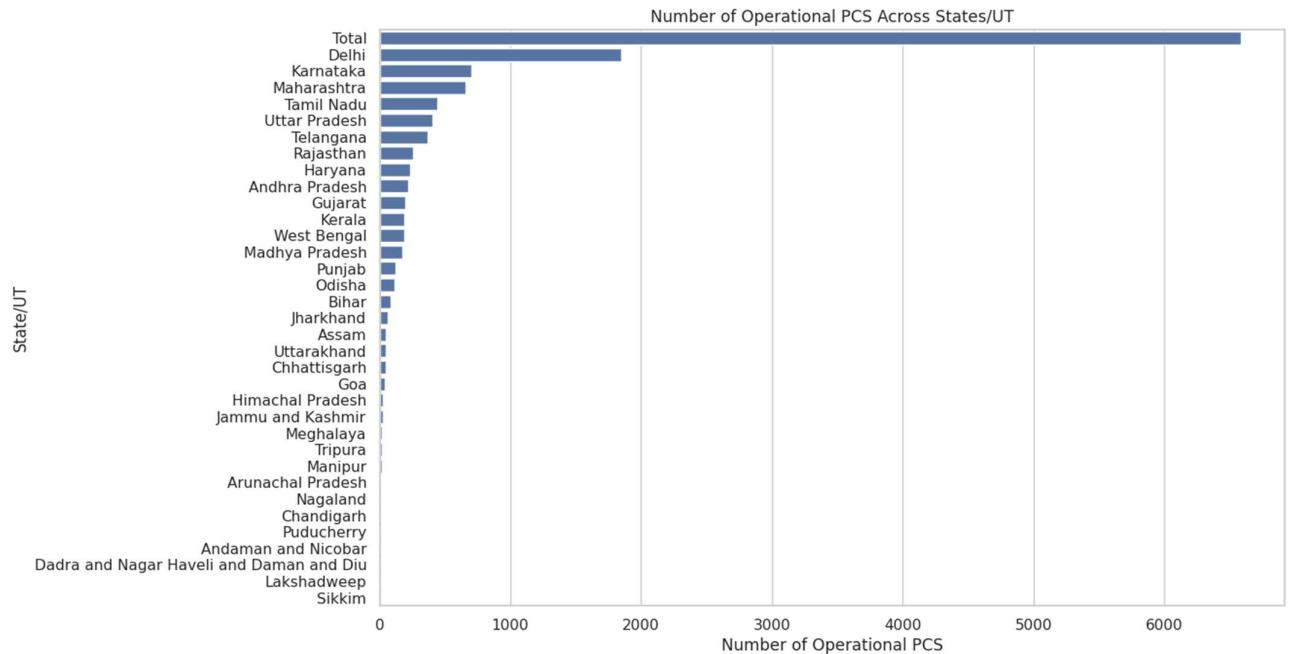
- There appears to be a positive correlation between the total number of active vehicles and the number of active electric vehicles in each state/UT.
- States with larger overall vehicle populations tend to have more electric vehicles, possibly due to increased awareness, infrastructure development, and economic factors driving EV adoption.

##### Charging\_station\_statewise.xlsx:

1. **Data Loading:** Loaded the dataset containing the number of operational Power Conditioning Systems (PCS) for each State/UT.
2. **Data Cleaning and Preprocessing:**
  - Checked for missing values, duplicate rows, and inconsistencies in the data.
  - Ensured uniformity in the formatting of state/UT names.
3. **Descriptive Statistics:**
  - Calculated summary statistics such as mean, median, mode, minimum, maximum, and standard deviation for the number of operational PCS.
  - Visualized the distribution of the number of operational PCS using histograms or box plots.
4. **State-wise Analysis:**
  - Explored the variation in the number of operational PCS across different states and UTs.
  - Identified states with the highest and lowest numbers of operational PCS: States like Delhi, Maharashtra, Karnataka, Tamil Nadu, and Uttar Pradesh have relatively higher numbers of operational PCS, indicating greater investment in

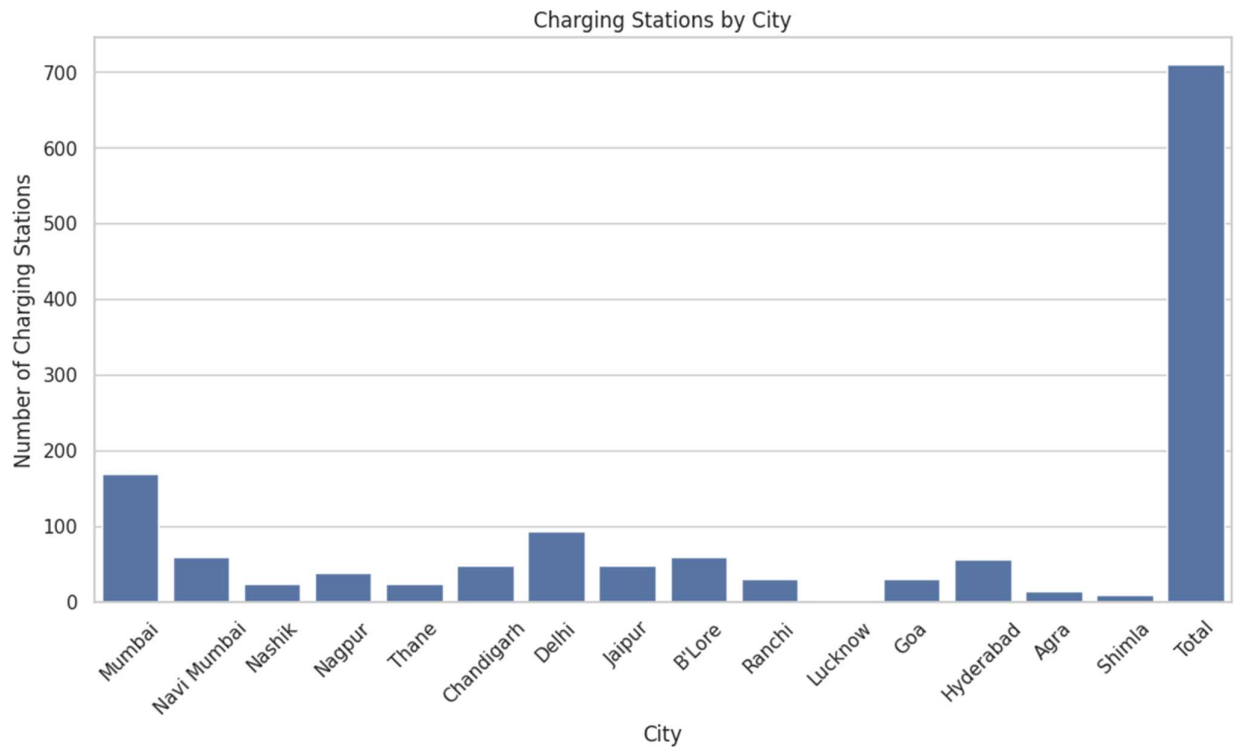
renewable energy infrastructure.

- Plotted a bar chart to visualized the distribution of operational PCS across states/UTs.



### Charging\_station\_citywise.xlsx:

1. **Data Loading:** Loaded the dataset containing the number of charging stations in different cities.
2. **Data Cleaning and Preprocessing:**
  - Checked for missing values, duplicate rows, and inconsistencies in the data.
  - Ensured uniformity in the formatting of city names.
3. **Descriptive Statistics:**
  - Calculated summary statistics such as mean, median, mode, minimum, maximum, and standard deviation for the number of charging stations.
  - Visualized the distribution of the number of charging stations using histograms or box plots.
4. **City-wise Analysis:**
  - Explored the distribution of charging stations across different cities.
  - Identified cities with the highest and lowest numbers of charging stations: Cities like Mumbai, Navi Mumbai, Delhi, Hyderabad, and Bengaluru have relatively higher numbers of charging stations, reflecting their status as major urban centres with high electric vehicle adoption rates.
  - Plotted a bar chart or a pie chart to visualize the distribution of charging stations across cities.

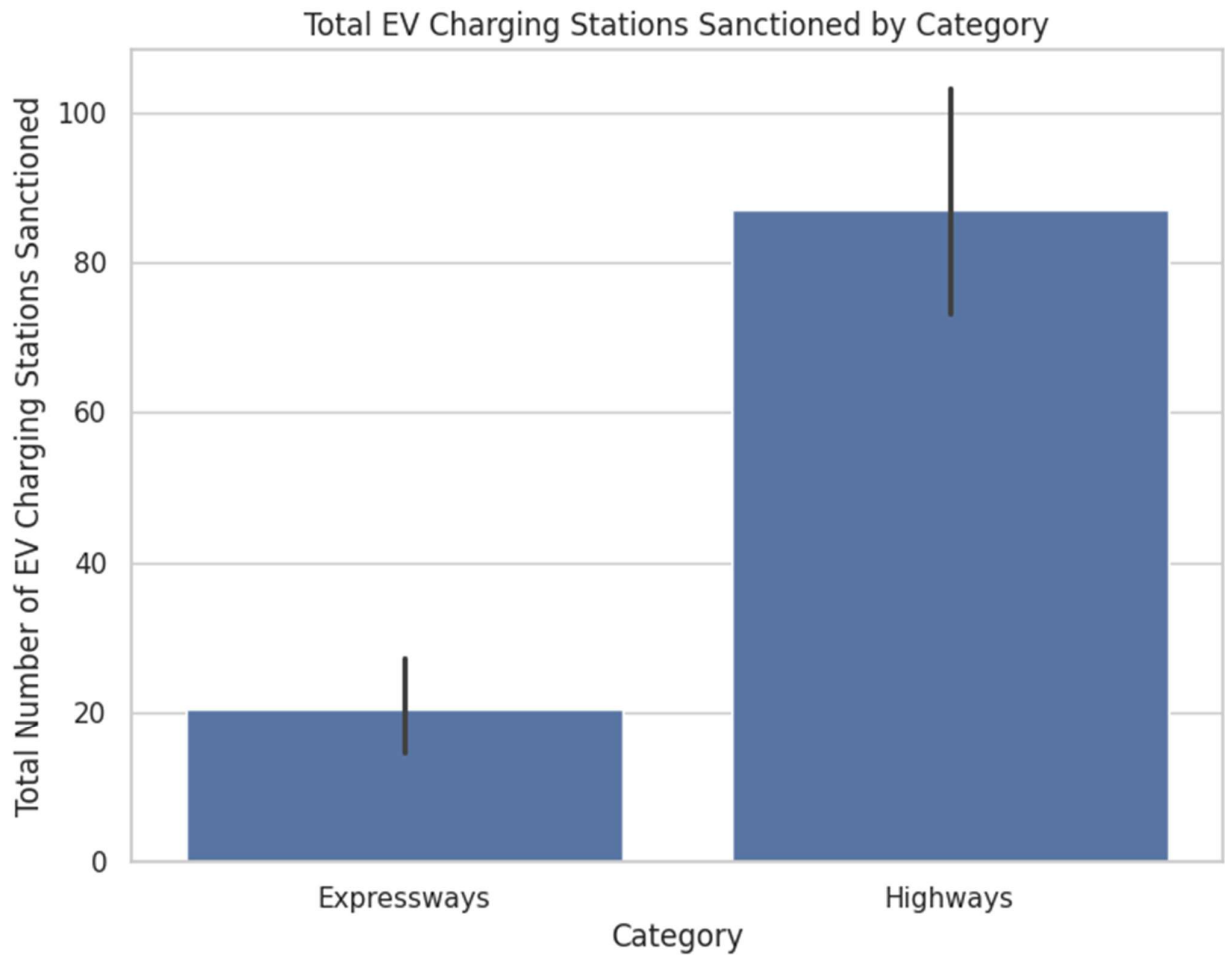


#### Highwise\_charging\_station.xlsx:

1. **Data Loading:** Loaded the dataset containing information about expressways, highways, and the number of EV charging stations sanctioned along each route.
2. **Data Cleaning and Preprocessing:**
  - Checked for missing values, duplicate rows, and inconsistencies in the data.
  - Ensured uniformity in the formatting of category names, expressway/highway names, and other relevant columns.

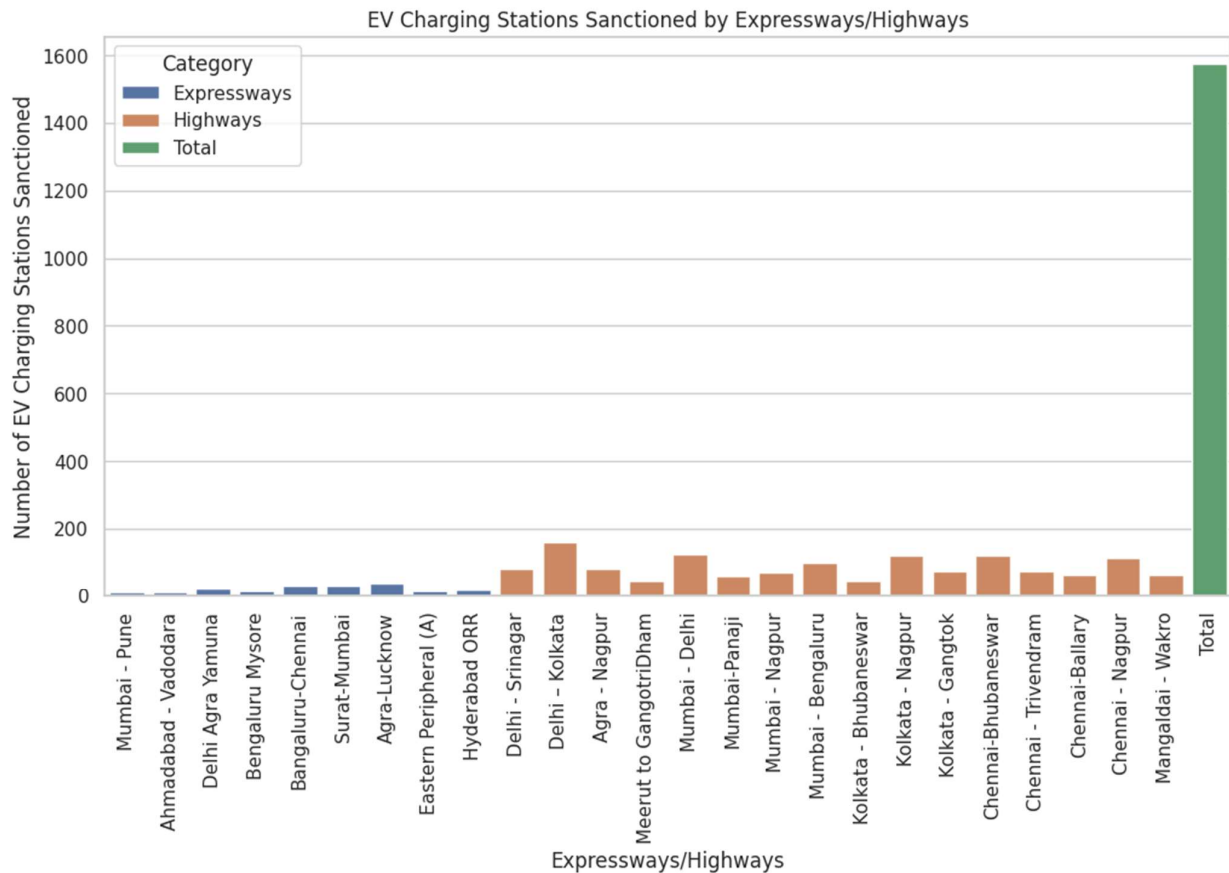
#### Descriptive Statistics:

- Calculated summary statistics such as mean, median, mode, minimum, maximum, and standard deviation for the number of EV charging stations sanctioned.
  - Visualized the distribution of sanctioned EV charging stations across expressways and highways using histograms or box plots.
3. **Category-wise Analysis:**
    - Explored the distribution of sanctioned EV charging stations across expressways and highways.
    - Identified the categories (expressways or highways) with the highest and lowest numbers of sanctioned EV charging stations.
    - Plotted a box plot to visualize the distribution of sanctioned EV charging stations by category.



**4. Route-wise Analysis:**

- Analyzed the number of EV charging stations sanctioned along each expressway and highway route.
- Identified routes with the highest and lowest numbers of sanctioned EV charging stations: Routes with high traffic volumes, such as Mumbai-Delhi and Delhi-Kolkata highways, have a larger number of sanctioned EV charging stations, reflecting the demand for electric vehicle infrastructure along these busy corridors. .
- Plotted a box plot to visualize the number of sanctioned EV charging stations along different routes.



## 6.0 Segmentation

Segment extraction in marketing involves the process of analyzing customer analytics and segmenting them based on various characteristics, allowing for targeted presentation of products or services tailored to their interests. Machine learning techniques are often employed for this purpose, and one commonly used method is K-Means Clustering.

K-Means Clustering is an Unsupervised Learning algorithm that partitions an unlabeled dataset into clusters. The parameter K represents the number of clusters to be created during the process. By iteratively optimizing cluster centroids, the algorithm assigns data points to the nearest centroid, forming distinct clusters based on proximity.

This centroid-based approach minimizes the total distance between data points and their respective centroids, effectively grouping similar data points together. Through this process, K-Means Clustering enables the discovery of meaningful segments within the dataset without the need for explicit training data.

Overall, K-Means Clustering serves as a valuable tool for segment extraction in marketing analytics, facilitating the identification of distinct customer groups for targeted marketing strategies.

The clustering algorithm identified several distinct customer segments for dataset based on age, location, profession, income, and preferences regarding electronic vehicles (EVs).

1. **Age of Customer:** Upon conducting univariate analysis of the EV\_data.csv, it becomes

evident that the predominant age group among potential customers falls within the range of 28 to 31 years.

2. **Customer Location:** The geographical location of customers stands as a crucial factor shaping market dynamics and customer outreach strategies. Further exploration into the city variable unveils promising markets for the EV industry, with Pune, Mumbai, and New Delhi emerging as focal points.
3. **Family Size of Customer:** In contemporary Indian society, vehicle ownership has become intrinsic to household transportation needs. Evaluating family size aids in delineating vehicle preferences and sizing requirements, thereby informing product development and marketing strategies. Family with **3-5** members are more interested in EV.
4. **Price of EV:** The selling price of electric vehicles offers a rich opportunity to delineate multiple market segments, allowing for tailored products that cater to the preferences and needs of affluent consumers.
5. **Wheel Type:** The configuration of wheels on electric vehicles presents another avenue for segmenting the market, accommodating diverse customer preferences. Leveraging this information can yield targeted market segments aligned with specific wheel type preference

### Selection of target segment

The process of market selection, known as Segmentation-Targeting-Positioning (STP), is crucial in crafting effective marketing strategies. Before positioning a product, it's imperative to segment the market and identify the target audience(s) for the marketing campaign.

After analyzing the data on electric vehicles in India, the following target segments have been inferred:

- **Target Segments based on Demographic Segmentation:** It has been observed that the most favorable target demographic for this market comprises individuals aged between 28 to 31, preferably holding graduate degrees. In terms of income, the focus should be on individuals earning between Rs. 20,64,995 to Rs. 28,12,149 annually. These segments represent key opportunities for targeted marketing strategies.
- **Geographic Segmentation Targets:** Based on segmentation analysis and the presence of government-sanctioned infrastructure such as EV charging stations, the primary locations for targeting should include Pune, Mumbai, New Delhi, and Bengaluru.
- **Psychographic (Lifestyle) Segmentation Targets:** Through data analysis and segmentation, it has been discerned that the focal point of the marketing strategy should revolve around unmarried individuals, followed by families with smaller household sizes ranging from 3 to 5 members.

## 8.0 Conclusions

- The analysis reveals diverse demographic representation, highlighting the broad spectrum of views on EVs across different age groups, professions, and income levels.
- States like Delhi, Maharashtra, Karnataka, Tamil Nadu, and Uttar Pradesh have relatively higher numbers of operational PCS, indicating greater investment in renewable energy infrastructure.

- Mumbai, Delhi, Bengaluru, Hyderabad Considering based on Segmentation Analysis and the Infrastructures sanctioned by Govt..
- Highways such as Delhi-Kolkata and Mumbai-Delhi stand out with a higher number of sanctioned charging stations, likely due to their status as major transportation arteries. Investing in robust charging infrastructure along these high-traffic corridors is essential to support seamless EV travel and alleviate range anxiety among users.
- Key insights include the influence of income on EV investment propensity, brand loyalty towards manufacturers like Tata and Hyundai, and affordability concerns among respondents. Moreover, the perception of EVs being economical among a significant portion of respondents underscores a positive attitude towards EV adoption, signaling opportunities for market growth.

#### Recommendations:

1. **Market Segmentation:** Further analyze demographic and preference data to identify specific customer segments and tailor marketing strategies accordingly, ensuring targeted outreach to diverse consumer groups.
2. **Geographic Expansion:** Expand EV infrastructure and awareness campaigns beyond Pune, Mumbai to capitalize on potential markets in other cities and regions, leveraging insights from geographical trends.

## 7.0 GITHUB Link

[https://github.com/Monalij28/EV\\_Market\\_segmentation/](https://github.com/Monalij28/EV_Market_segmentation/)

## 8.0 References

1. <https://www.kaggle.com/>
2. <https://data.gov.in/>
3. <https://www.google.co.in/>
4. <https://www.linkedin.com/pulse/mastering-marketing-mix-4ps-beyond-effective-business-mubashar/>