Name:Vijay Pratap Yadav

Topic:Electric Vechicle Analysis

E-VEHICLE TRENDS IN INDIAN STATES

The present study helps to understand the current scenario of Electric Vehicles in Indian states and the reasons for their rising popularity among people. The study is based on secondary data taken from government reports that depict popularity as highly diverse among Indian states. States like Uttar Pradesh,Bihar, Karnataka,Rajasthan, Tamil Nadu, and Assam are at the top in adopting EVs that are away from traditional combustion engine vehicles.Although the government of India is running various schemes and programs for the growth of the electric vehicles industry, state policy is highly important in implementing the programs and schemes initiated by the central government. Bihar is keenly prioritizing the industrial policy in favor of EVs using the PPP model and is on the list of leading states in production and sales. These vehicles' lower maintenance and running costs make them popular among consumers. The need of the day is to work on the infrastructural front to give a further kick to the industry.

The Indian government has taken various steps to promote electric vehicle usage.

- FAME India scheme: The faster adoption and manufacturing of the Hybrid & Electric Vehicles scheme was launched in 2015 to reduce the dependency on fossil fuels and address the issue of vehicular emissions. Currently, phase II of the FAME India scheme is under implementation w.e.f. April 1, 2019.

- PLI Scheme: The production incentive scheme was launched in May 2021 to bring down the price of batteries in India, which may reduce the cost of electric vehicles.

- GST reduced from 12 % to 5%: The government of India has reduced the GST on electric vehicles to 5 percent. GST on charging stations was reduced from 18 percent to 5 percent.

- Exempt from permit requirements: Battery-operated electric vehicles are exempted from permit requirements and are given green license plates as per the announcement of the Ministry of Road Transport and Highways.

- Road Tax: The Ministry of Road Transport Highways has announced waiving road tax on Electric Vehicles, which will reduce their initial cost.

EV SALES TRENDS IN INDIAN STATES

Electric vehicles (EVs) are experiencing explosive growth, driven by a confluence of factors including technological advancements, government incentives, and a growing awareness of the environmental impact of traditional fossil fuel vehicles.

Sales of EVs have been doubling annually, with November 2023 alone witnessing a 27.15% increase compared to the previous

year same month. This upward trajectory shows no signs of slowing down, with experts predicting EVs to capture a significant portion of the global automotive market in the coming years. Several key factors fuel the surge in EV adoption.November 2023 witnessed a booming EV market, with sales soaring 9.54% from the previous month, reaching 152,514 units. This growth continues the upward trend, with a 27.15% increase from November 2022.

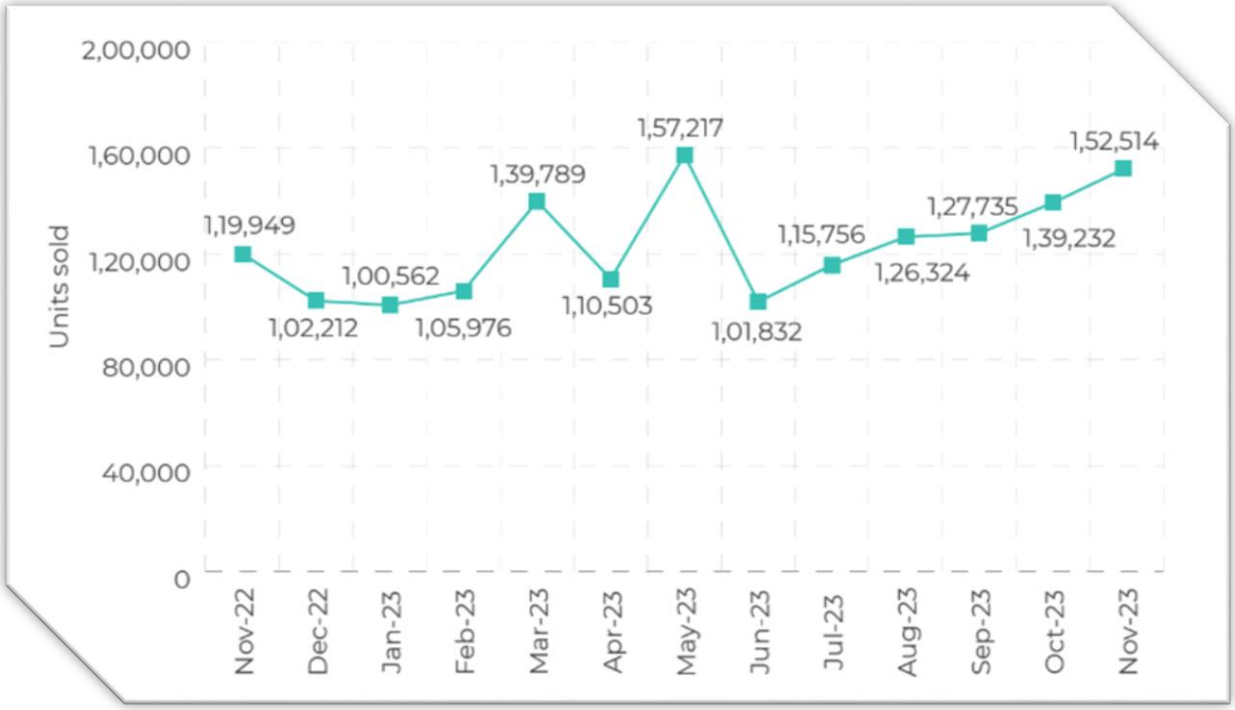
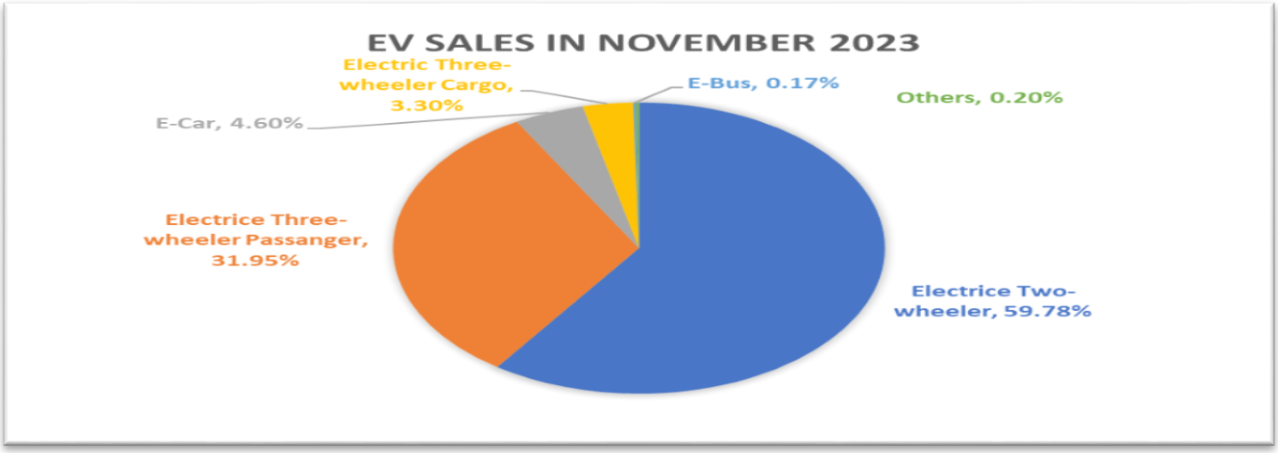


Fig. Trends of Sales of EV vehicles

Electric two-wheelers and three-wheelers saw a significant increase in registrations in November 2023, contributing to 91.73% of the total and solidifying their position as the leading segment of the Indian EV market.



Data Collection

The data has been collected manually, and the sources used for this process are listed below:

- **https://www.kaggle.com/datasets**
- **https://data.gov.in/**
- **https://www.data.gov/**

Implementation

Packages/Tools used:

1. Numpy: To calculate various calculations related to arrays.
2. Pandas: To read or load the datasets.
3. SKLearn: Used for K-means clustring.

Data-Preprocessing

Data Cleaning

Unnamed: 0	State Name	Two Wheeler	Three Wheeler	Four Wheeler	Goods Vehicles	Public Service Vehicle	Special Category Vehicles	Ambulance/Hearses	Construction Equipment Vehicle	Other	Grand Total	total-charging-stations	
0	0	Andaman and Nicobar Island	1	30	81	0	40	0	0	0	7	159	3
1	1	Arunachal Pradesh	14	0	5	0	0	0	0	0	1	20	9
2	2	Assam	721	47041	161	7	15	0	0	0	2	47947	10
3	3	Bihar	5003	59079	114	11	26	0	0	0	8	64241	9
4	4	Chandigarh	298	1410	182	0	40	0	0	0	1	1931	2
5	5	Chhattisgarh	6424	5341	117	1077	1	0	0	368	100	13428	18

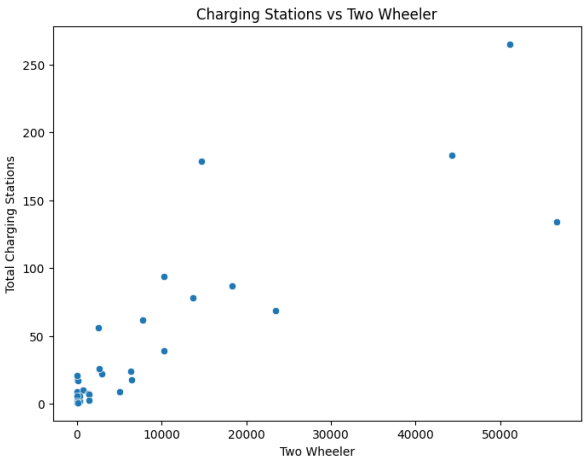
The data collected is compact and is partly used for visualization purposes and partly for clustering. Python libraries such as NumPy, Pandas, Scikit-Learn, and SciPy are used to perform several steps.

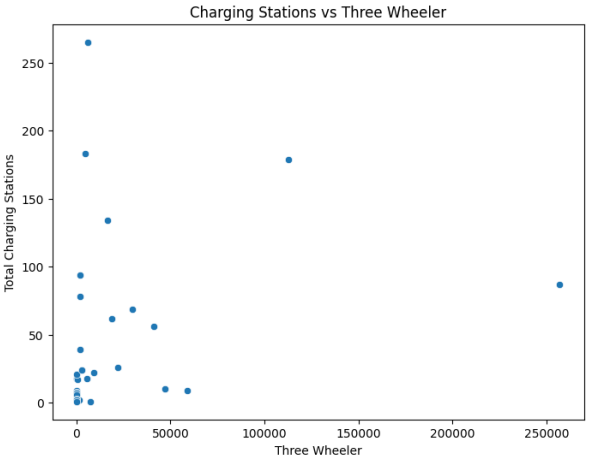
The dataset looks like given below:

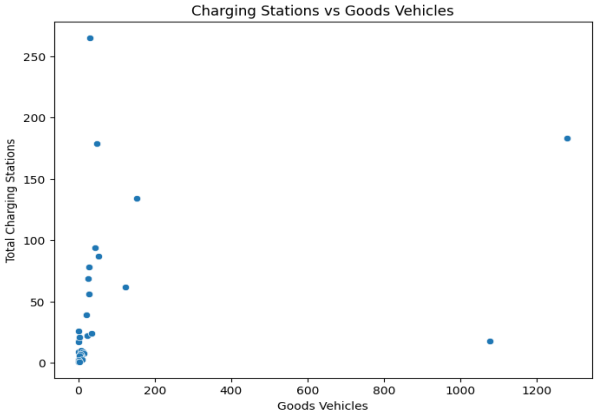
EDA

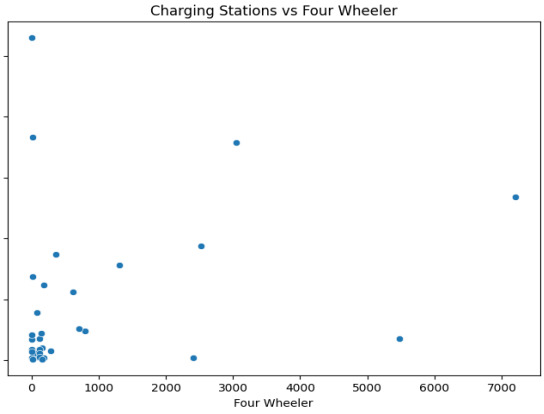
Tried to extract important insights form dataset by doing Exploratory Data Analysis with the help of pandas and matplotlib.

- Scatter plot for each vehicle type vs. charging stations

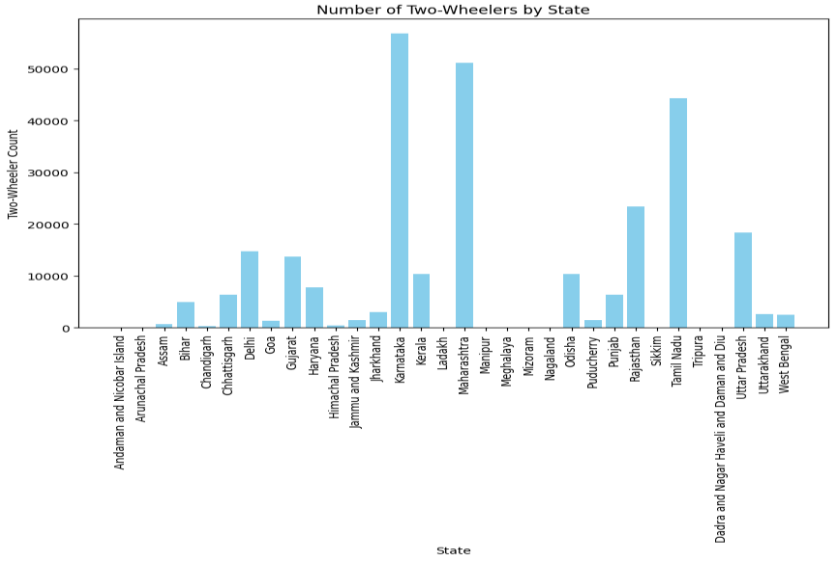




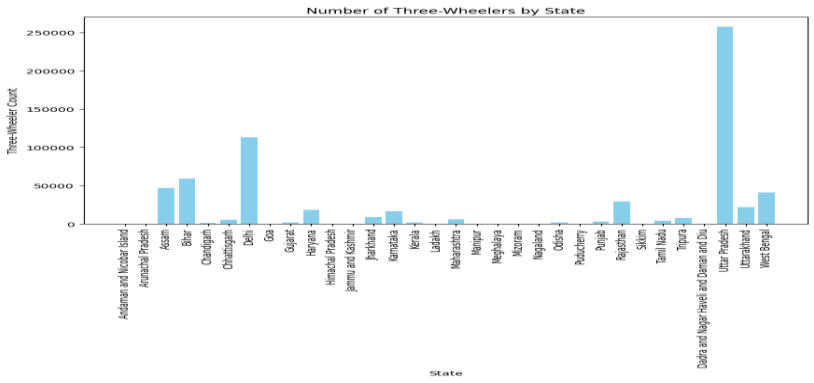




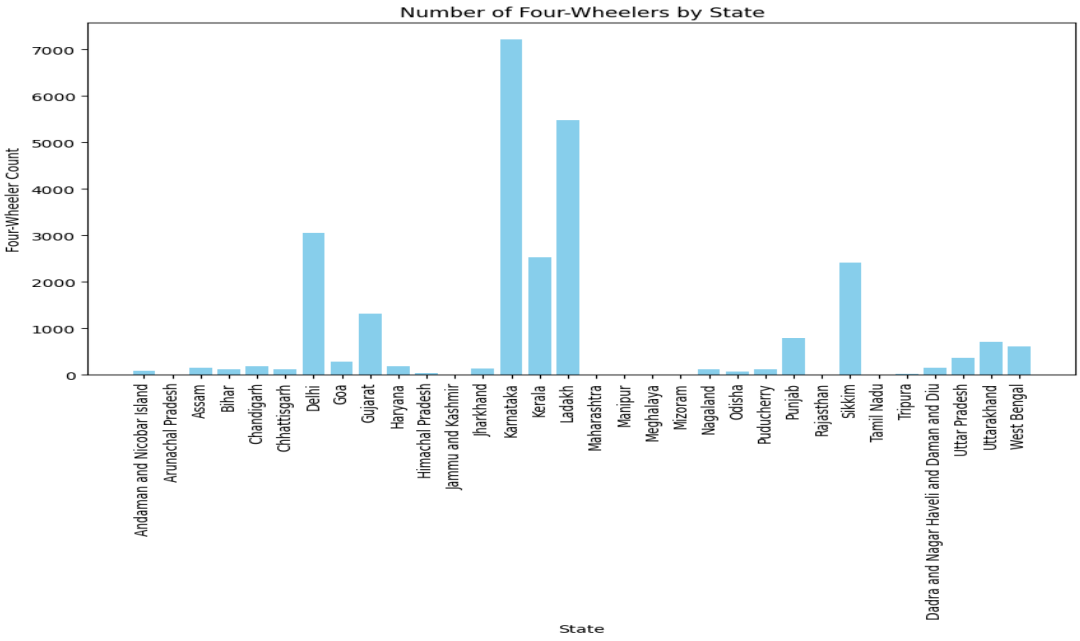
- Number of two wheeler in each states



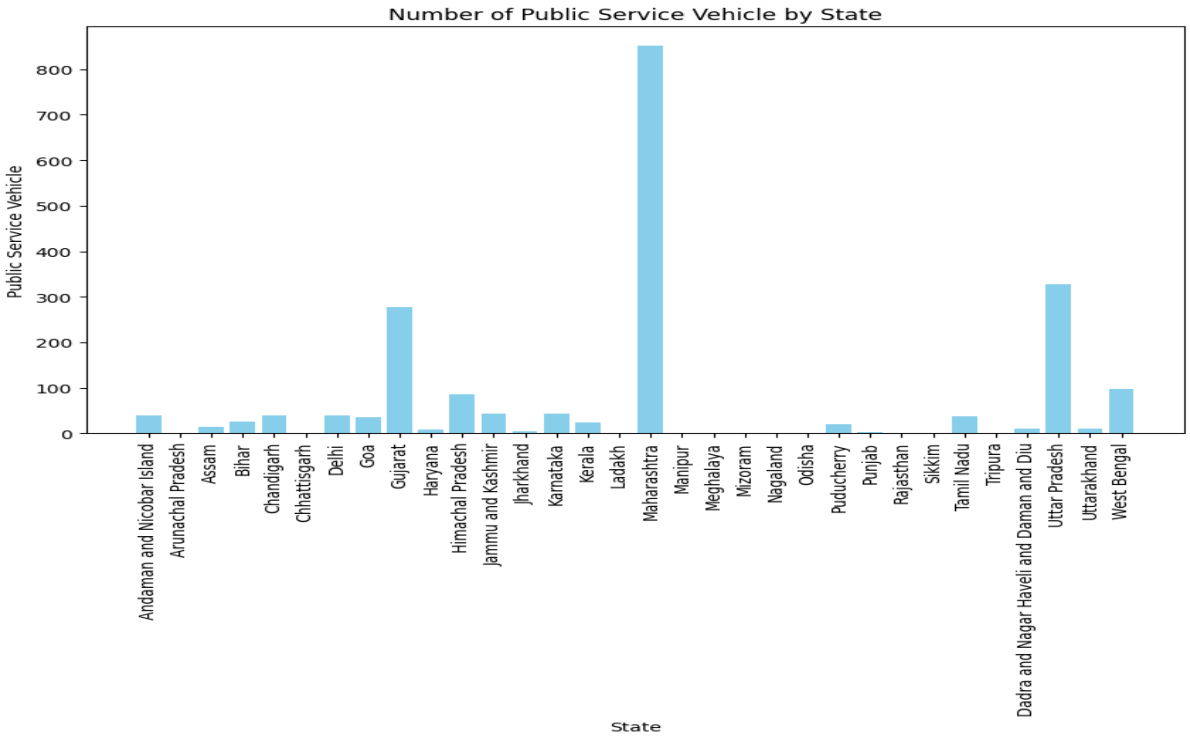
- Number of three wheeler in each states



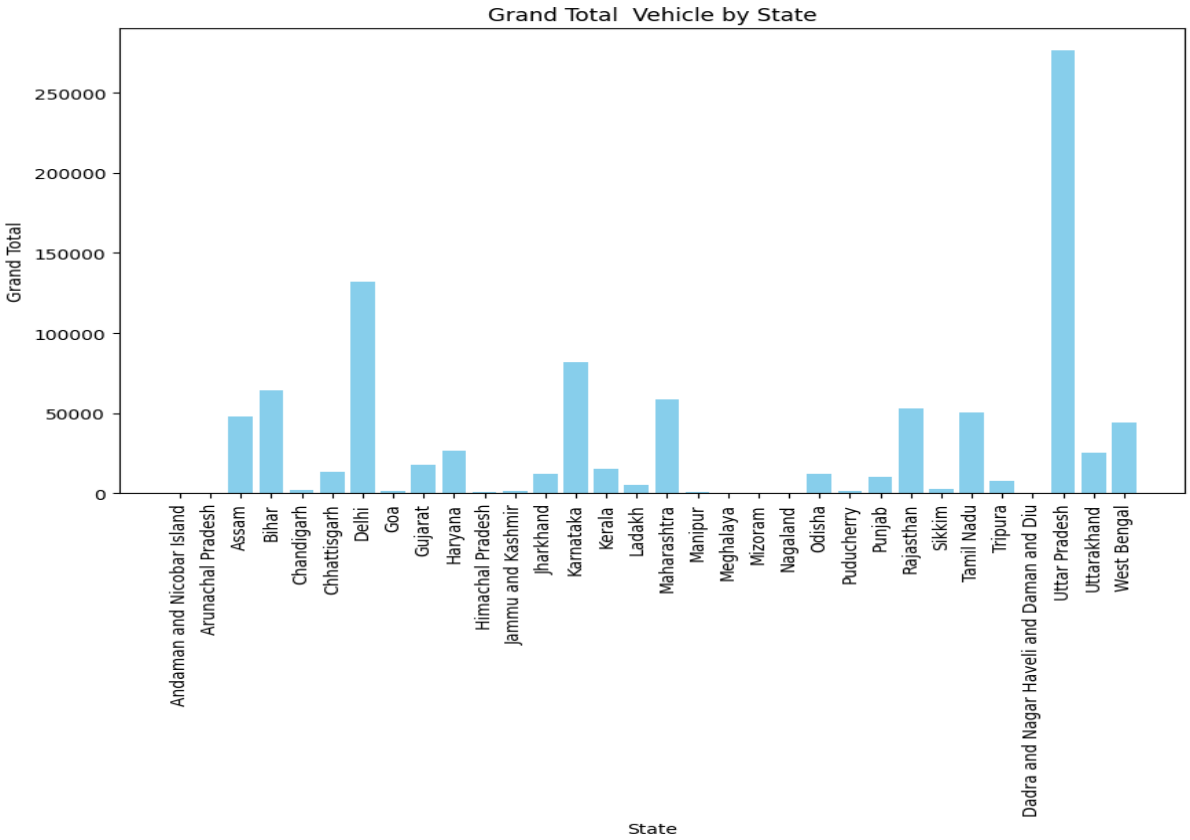
- Number of four wheeler in each states



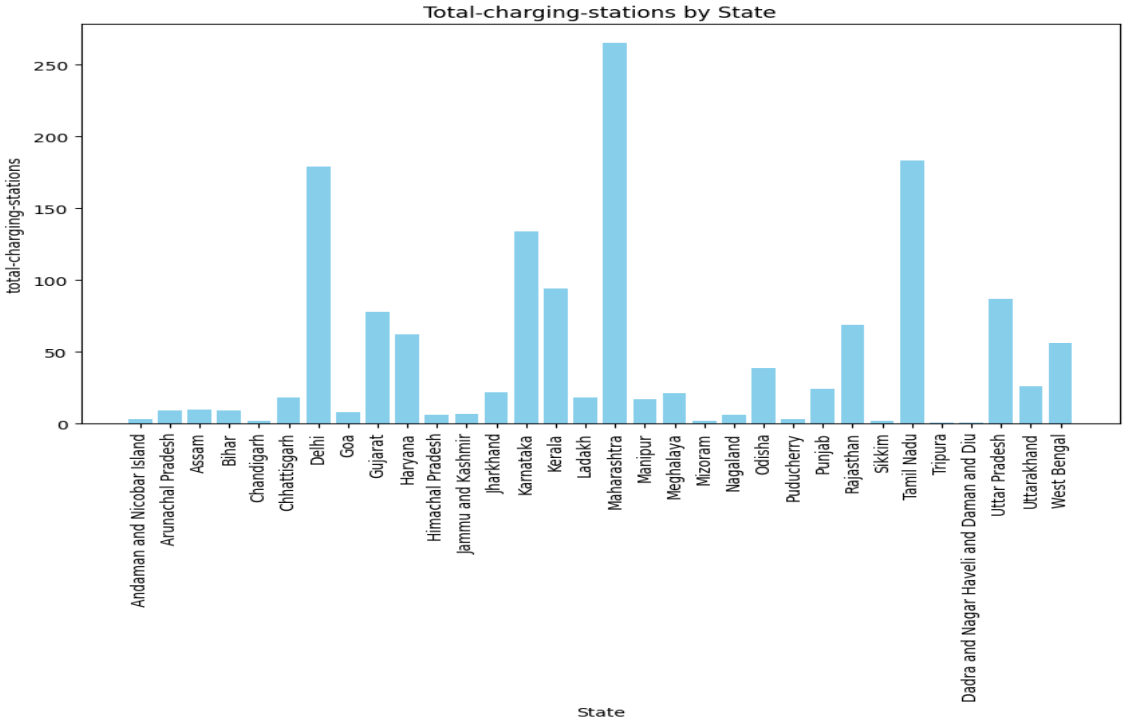
- Number of public service vechiles in each states



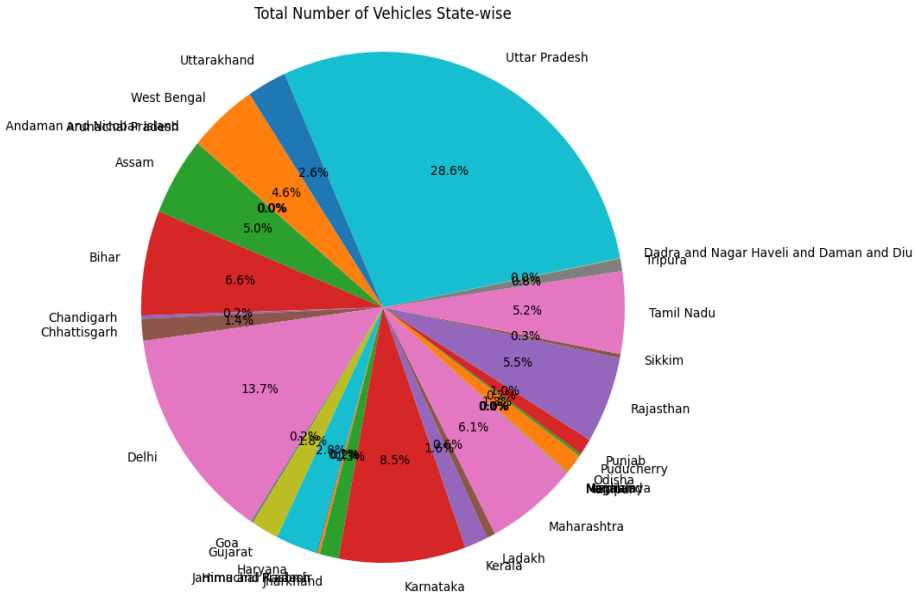
- Total Numbers of vechiles in each states



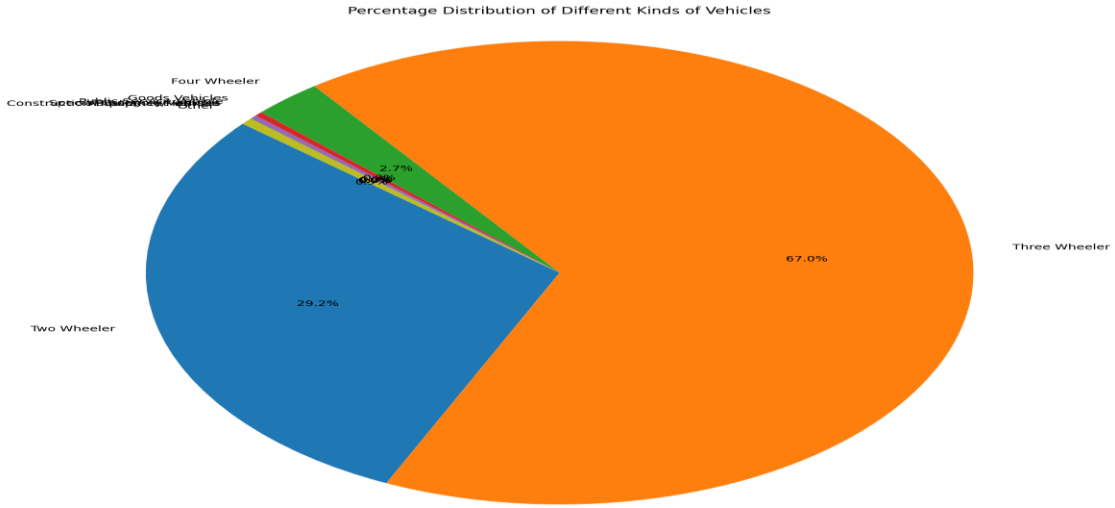
- Number of power station in each states



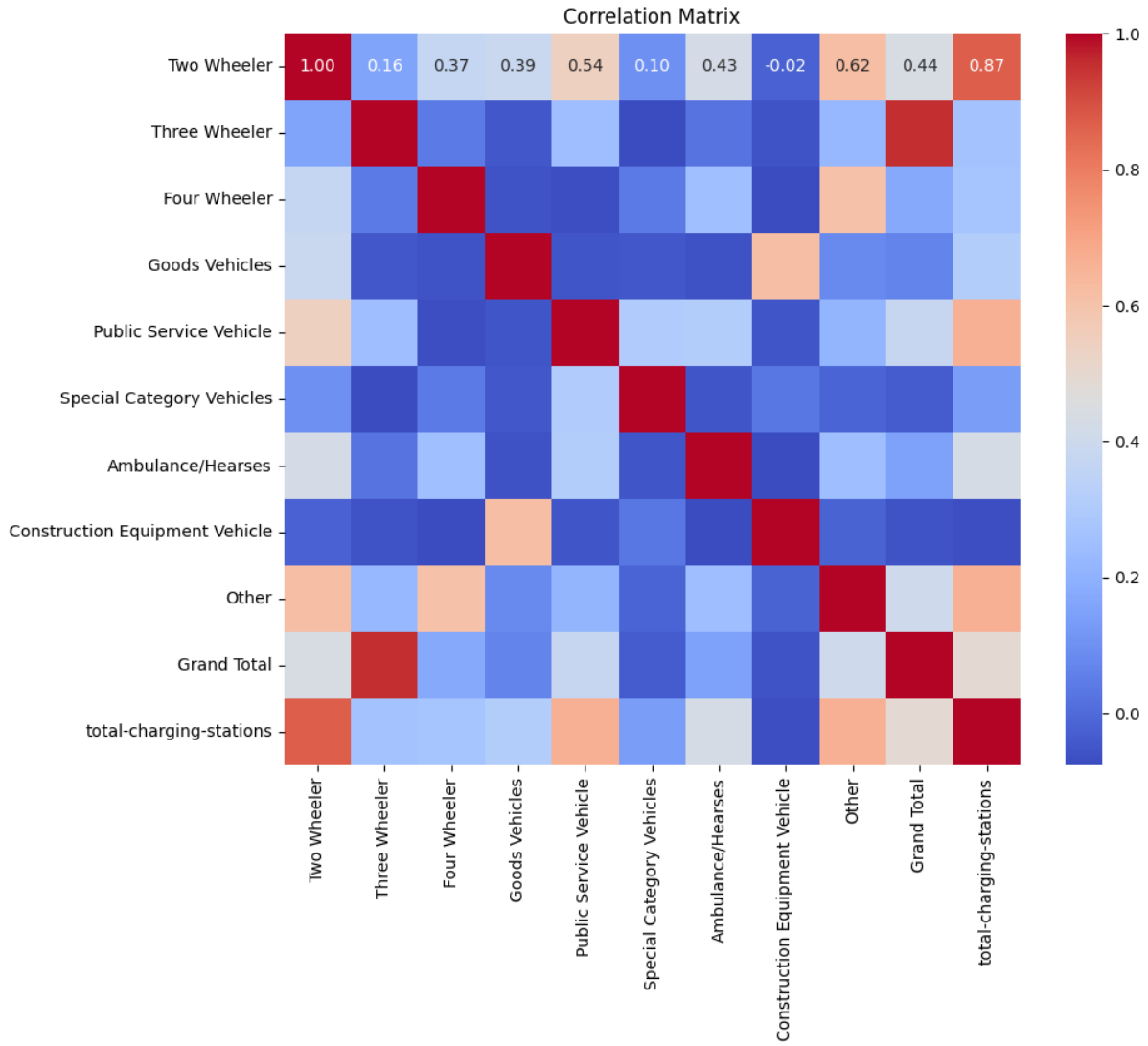
- Pie chart of total number vechiles statewise



- Percentage Distribution of Different Kinds of Vehicles



Correlation Matrix: A correlation matrix is simply a table that displays the correlation. It is best used in variables that demonstrate a linear relationship between each other. Coefficients for different variables. The matrix depicts the correlation between all the possible pairs of values through the heatmap in the below figure. The relationship between two variables is usually considered strong when their correlation coefficient value is larger than 0.7.



Clustering

Clustering is one of the most common exploratory data analysis techniques used toget an intuition about the structure of the data. It can be defined as the task of identifying subgroups in the data such that data points in the same subgroup (cluster) are very similar while data points in different clusters are very different. In other words,we try to find homogeneous subgroups within the data such that data points in eachcluster are as similar as possible according to a similarity measure such as euclideanbased distance or correlation-based distance.

The decision of which similarity measure to use is application-specific. Clustering analysis can be done on the basis of features where we try to find subgroups of samples based on features or on the basis of samples where we try to find subgroups of features based on samples.

K-Means Algorithm

K Means algorithm is an iterative algorithm that tries to partition the dataset into pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to only one group. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different (far) as possible. It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster’s centroid (arithmetic mean of all the data points that belong to that

cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster.

The way k means algorithm works is as follows:

• Specify number of clusters K.

• Initialize centroids by first shuffling the dataset and then randomly selecting K

data points for the centroids without replacement.

• Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.

The approach k-means follows to solve the problem is expectation maximization

The E-step is assigning the data points to the closest cluster. The M-step is computing the centroid of each cluster.

Applications

K means algorithm is very popular and used in a variety of applications such as market segmentation, document clustering, image segmentation and image compression, etc.

The goal usually when we undergo a cluster analysis is either:

1. Get a meaningful intuition of the structure of the data we're dealing with.

2. Cluster-then-predict where different models will be built for different subgroups if we believe there is a wide variation in the behaviors of different subgroups.

The k-means clustering algorithm performs the following tasks:

• Specify number of clusters K

• Initialize centroids by first shuffling the dataset and then randomly selecting K

data points for the centroids without replacement.

• Compute the sum of the squared distance between data points and all centroids.

• Assign each data point to the closest cluster (centroid).

• Compute the centroids for the clusters by taking the average of the all data points that belong to each cluster.

• Keep iterating until there is no change to the centroids. i.e assignment of data points to clusters isn't changing.