EV MARKET ANALYSIS

Report

(Task 2 Market Segmentation)

Date – 18-7-23

Ritik Mehrotra

(Machine Learning Intern @Fyenn labs)

Link → https://github.com/ritikdeveloper/Feynn-Labs-Internship

Overview of the Electric Vehicle (EV) Market Segment in India

As the world focuses on reducing greenhouse gas emissions and combating climate change, the adoption of electric vehicles (EVs) has become an essential component of sustainable transportation. In recent years, India has emerged as one of the key players in the global EV market, driven by its ambitious efforts to transition towards cleaner and greener mobility solutions. This overview will delve into the current state of the EV market segment in India, covering various aspects such as government initiatives, market trends, challenges, and future prospects.

The Indian government has been proactive in promoting EV adoption through a series of policies and incentives. One of the key initiatives is the Faster Adoption and Manufacturing of Electric Vehicles (FAME) scheme, launched in 2015 and subsequently revised in 2019. FAME II aims to accelerate the adoption of EVs in the country by providing incentives to manufacturers and buyers, encouraging the development of EV charging infrastructure, and supporting research and development in the sector.

Moreover, various states have also introduced their EV policies, offering additional incentives like subsidies, tax exemptions, and reduced registration fees to promote electric mobility at the regional level. These policy measures have significantly contributed to the growth of the EV market in India.

Market Trends and Growth:

The EV market in India has witnessed remarkable growth in recent years, driven by several factors. The rising awareness of environmental concerns, increasing fuel prices, and advancements in EV technology have all played a pivotal role in boosting consumer interest in electric vehicles. Additionally, various automakers, both domestic and international, have entered the Indian EV market, offering a diverse range of electric cars, two-wheelers, and commercial vehicles.

Electric two-wheelers have been particularly successful in India, owing to their affordability, convenience, and popularity in urban areas. Startups and established players alike have been vying to capture this segment, resulting in fierce competition and innovation.

Despite the progress, the Indian EV market faces several challenges that need to be addressed to ensure sustainable growth. One of the primary concerns is the lack of adequate charging infrastructure. While the government has been actively promoting the establishment of EV charging stations, the pace of development needs to be accelerated to allay range anxiety among potential buyers.

The cost of electric vehicles remains relatively higher compared to conventional internal combustion engine vehicles, primarily due to the cost of batteries. Although government subsidies and incentives help bridge the gap, continued efforts are required to make EVs more affordable for the masses.

Battery technology and indigenous battery manufacturing are also areas that need significant attention. India is heavily dependent on importing batteries, which not only increases costs but also raises concerns about supply chain disruptions. Encouraging domestic battery manufacturing and investing in research and development of advanced battery technologies can mitigate these challenges.

Despite the challenges, the future prospects of the EV market in India appear promising. The government's commitment to reducing carbon emissions and its emphasis on electric mobility as part of its sustainable development goals are strong indicators of continued support for the EV sector.

Moreover, several international automakers have shown interest in investing in India's EV market, paving the way for technological collaborations and knowledge transfer. This, coupled with the country's growing manufacturing capabilities, can lead to increased localization of EV production and a subsequent reduction in costs.

As battery technology improves and economies of scale kick in, the cost of electric vehicles is expected to decline further, making them more accessible to a wider audience. With the right blend of supportive policies, investment, and consumer awareness, India has the potential to become a global leader in the electric vehicle market.

Dataset Used

1) *Electric Vehicle Charging Stations in India*

The dataset available on Kaggle provides comprehensive information about the charging infrastructure for electric vehicles in India.

The dataset offers a detailed overview of electric vehicle charging stations across various regions in India.

It contains a rich collection of data on charging station locations, including latitude and longitude coordinates.

The dataset covers information about the charging station's operator, making it useful for analyzing the market players in the EV charging sector.

Each entry in the dataset includes the station's address, enabling users to assess the accessibility of charging infrastructure in different areas.

The dataset provides insights into the type of charging stations, such as fast chargers, slow chargers, or both, facilitating the understanding of charging capabilities.

Users can explore the available charging connectors at each station, helping EV owners find compatible charging options.

It contains data on the charging station's power capacity, which is crucial for understanding the charging speed and efficiency.

2) *State/UTs-wise Electric Vehicles (EVs) under Operation under Phase-II of Faster Adoption and Manufacturing of Electric Vehicles (FAME) Scheme*

The dataset available on data.gov.in provides a comprehensive overview of electric vehicle adoption across different states and union territories in India. Here are 30 lines highlighting key aspects of the dataset:

The dataset offers detailed information on the number of electric vehicles (EVs) currently in operation across various states and union territories in India.

It provides data specifically related to Phase-II of the Faster Adoption and Manufacturing of Electric Vehicles (FAME) Scheme, which focuses on accelerating EV adoption in the country.

Users can access statistics on the total count of electric vehicles categorized by vehicle types such as two-wheelers, three-wheelers, four-wheelers, and buses.

The dataset presents a temporal dimension, enabling users to observe the growth and trends of EV adoption over time during Phase-II of the FAME Scheme.

Researchers can analyze the state-wise distribution of EVs to identify regions with higher or lower adoption rates.

The dataset includes information on the type of technology used in the electric vehicles, such as battery electric vehicles (BEVs) and plug-in hybrid electric vehicles (PHEVs).

Users can assess the overall impact of the FAME Scheme on promoting EVs across different states and union territories.

It provides insights into the popularity of electric vehicles in urban and rural areas, facilitating regional analysis.

Researchers can compare the adoption rates of different types of EVs (two-wheelers, three-wheelers, etc.) to understand varying consumer preferences.

The dataset is useful for policymakers and stakeholders in the automotive industry to gauge the success of the FAME Scheme in achieving its objectives.

Implementation

Packages/Tools used:

[1]:
```
import pandas as pd
import numpy as np
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.preprocessing import LabelEncoder
from sklearn.decomposition import PCA
from sklearn import preprocessing
from bioinfokit.visuz import cluster
from sklearn.cluster import KMeans
from yellowbrick.cluster import KElbowVisualizer
from collections import Counter
from statsmodels.graphics.mosaicplot import mosaic
from itertools import product
```

Data-Preprocessing
Data Cleaning

Dataset 1:

In the given dataset, there are seven columns: 'name', 'state', 'city', 'address', 'latitude', 'longitude', and 'type'. However, it has been decided to remove the

columns 'city', 'address', 'latitude', and 'longitude' from the dataset as they are considered to have no impact on the overall analysis. The reasoning behind this decision may be that the location-specific information provided by these columns is not necessary for the analysis at hand.

After removing the aforementioned columns, the dataset now contains four columns: 'name', 'state', and 'type'. 'name' represents the name of each station, 'state' denotes the state in which the station is located, and 'type' is a numerical value representing the type of station.

The goal of the analysis is to count the total number of stations in each state. This step is essential to gain insights into the distribution of stations across different states. By calculating the frequency of stations in each state, it becomes easier to identify patterns, trends, or discrepancies in the data.

Dataset 2

After removing the columns "Ambulance/Hearses", "Construction Equipment Vehicle", "Special Category Vehicles", and "Public Service Vehicle" from the dataset due to their mostly null values, we are left with the following columns: "State Name", "Two Wheeler", "Three Wheeler", "Four Wheeler", "Goods Vehicles", "Other", and "Grand Total".

The data cleaning process involves filling in the null values using different techniques. These techniques are applied to handle missing or incomplete data, ensuring that the dataset is ready for analysis.

There are several common methods to handle missing data, and the choice of technique depends on the nature of the data and the specific analysis requirements. Here, I will describe a few common techniques that could be used to clean the data:

Mean/Median/Mode Imputation:

For numerical columns like "Two Wheeler", "Three Wheeler", "Four Wheeler", "Goods Vehicles", and "Other", we can use mean, median, or mode imputation to fill in the missing values. Mean imputation replaces missing values with the average of the non-missing values in the column, while median imputation uses the middle value, and mode imputation uses the most frequent value.

Forward Fill (or Backward Fill) Imputation:

If the data exhibits some sort of temporal or sequential ordering (e.g., time series data), we can use forward fill or backward fill to fill missing values. Forward fill fills missing values with the last known value, while backward fill uses the next known value.

Data Comparison

In each state, we have compared the number of different types of vehicles. The data provides insights into the transportation landscape of these regions.

The "Two Wheeler" category represents the number of two-wheeled vehicles, such as motorcycles and scooters, present in each state.

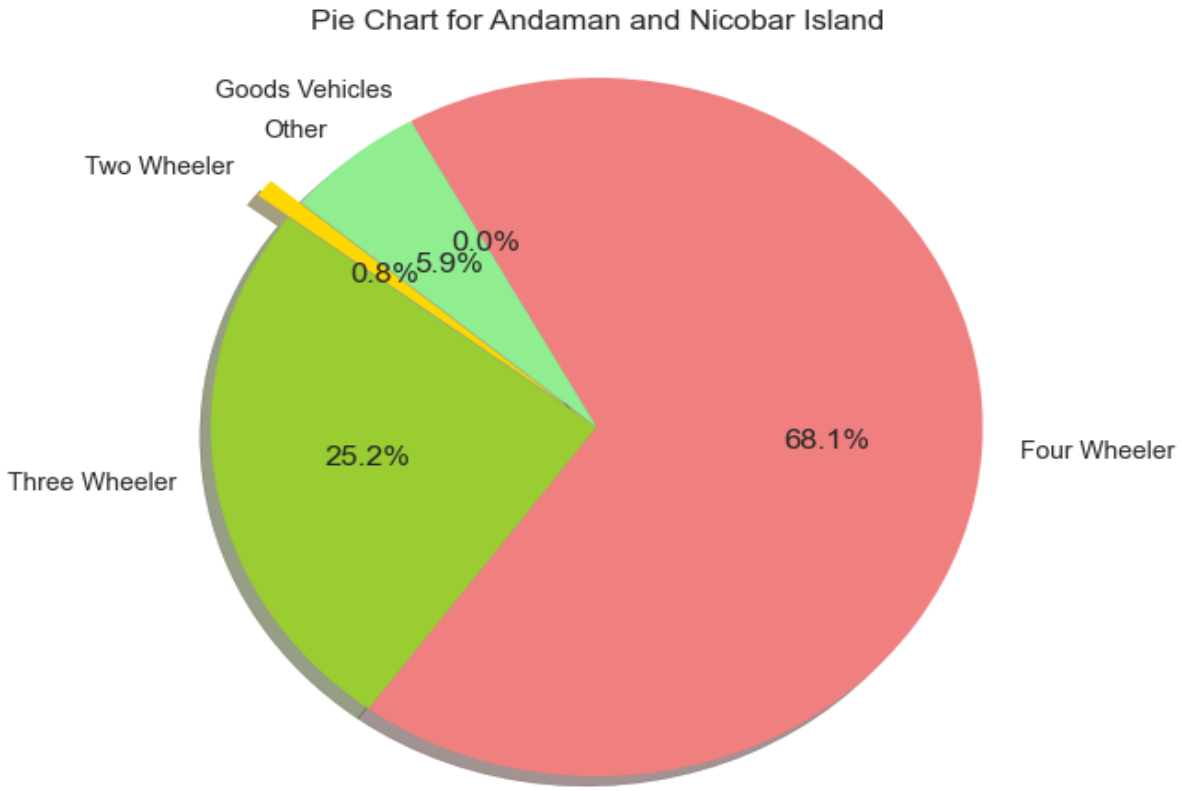
The "Three Wheeler" category includes three-wheeled vehicles, and this data point helps us understand the prevalence of auto-rickshaws or other similar vehicles in the respective states.

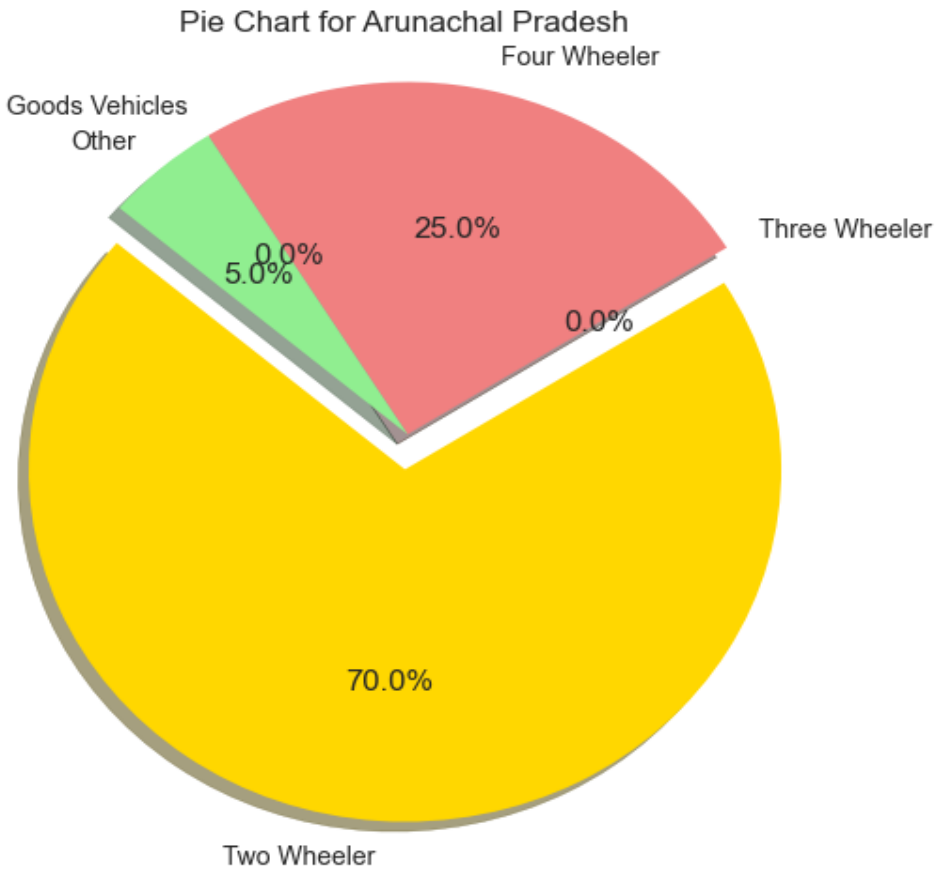
The "Four Wheeler" category encompasses all types of four-wheeled vehicles, including cars and small trucks, giving us an overview of private and commercial car ownership.

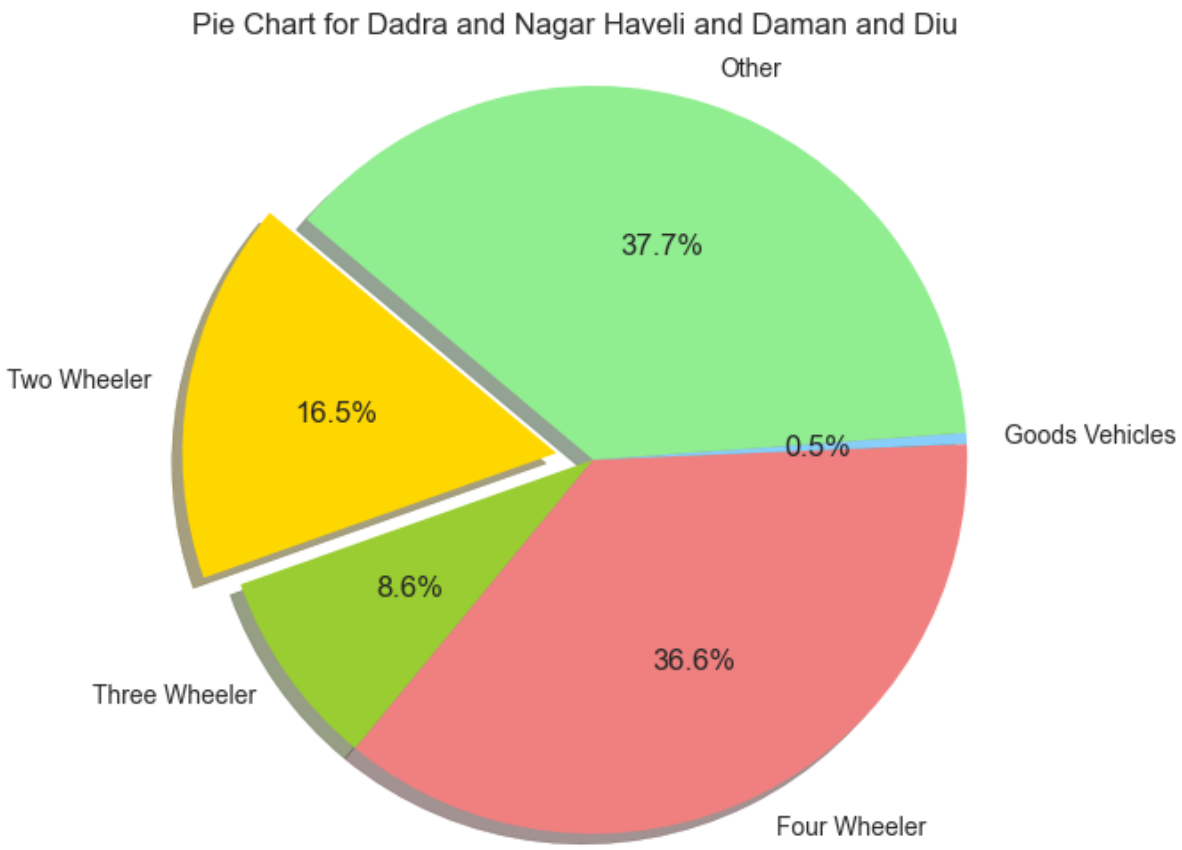
The "Goods Vehicles" category highlights the number of vehicles used for transporting goods and freight, offering valuable insights into the logistics and trade activities of each state.

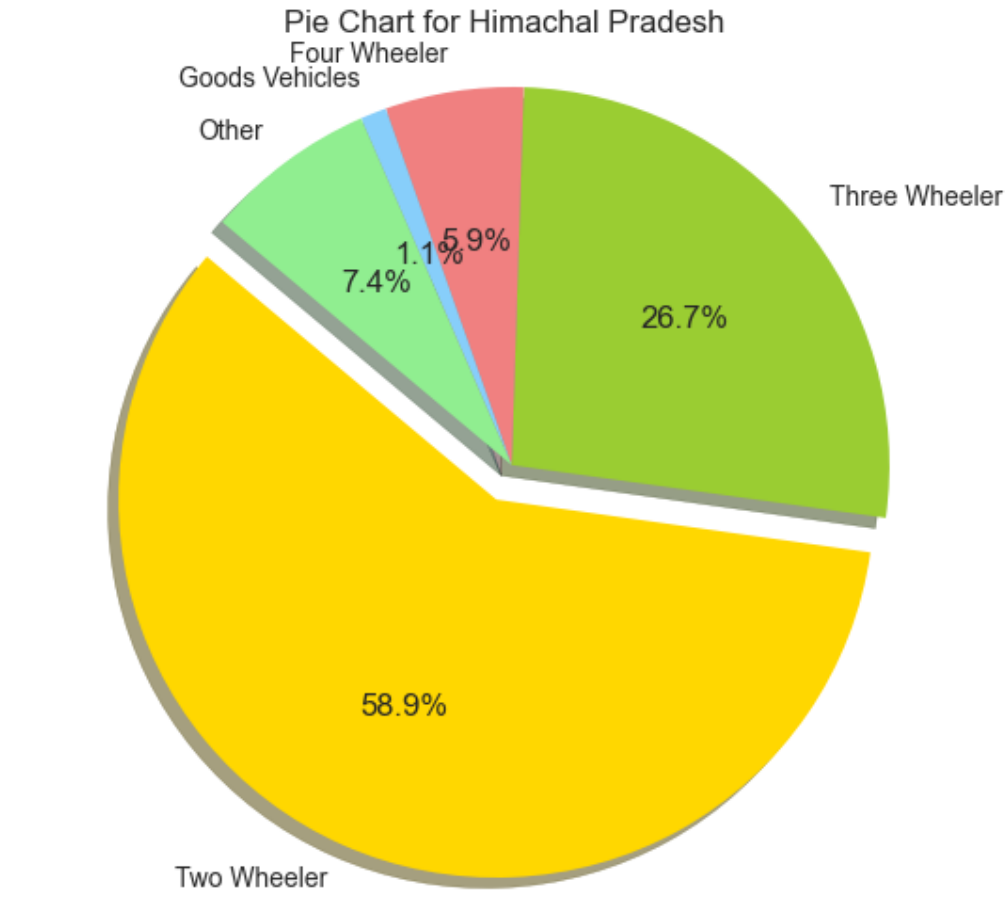
Lastly, the "Other" category accounts for any other types of vehicles not covered in the previous classifications, which could include specialized vehicles or unique transportation modes.

This comprehensive comparison across all states helps us gain a better understanding of the diverse transportation preferences and needs in different regions


Pie Chart for Andaman and Nicobar Island
Goods Vehicles
Other
Two Wheeler
0.0%
0.8%
5.9%
25.2%
68.1%
Three Wheeler
Four Wheeler

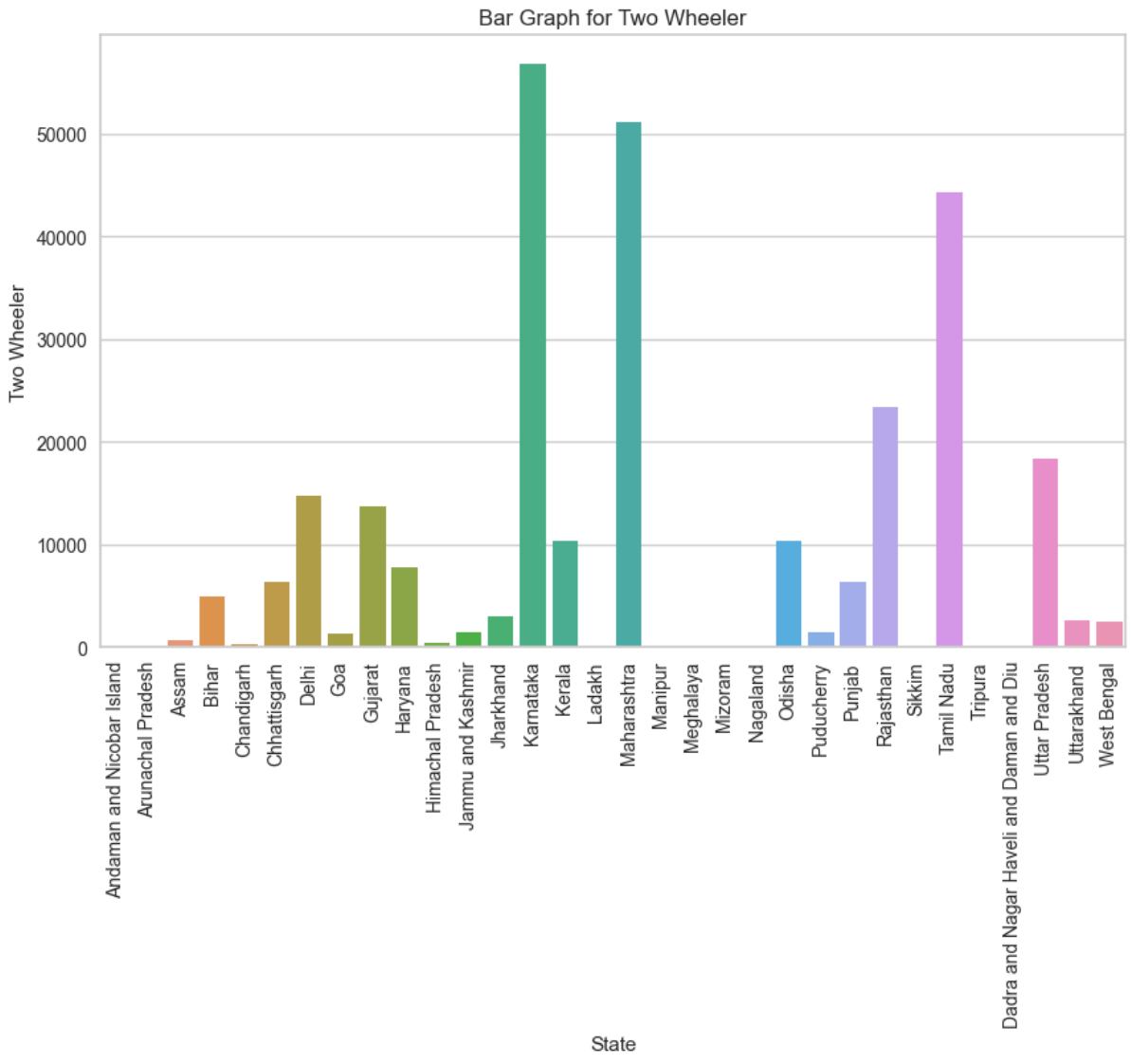

Pie Chart for Arunachal Pradesh
Four Wheeler
Goods Vehicles
Other
25.0%
Three Wheeler
0.0%
5.0%
0.0%
70.0%
Two Wheeler


Pie Chart for Dadra and Nagar Haveli and Daman and Diu
Other
37.7%
Two Wheeler
16.5%
Goods Vehicles
0.5%
8.6%
Three Wheeler
36.6%
Four Wheeler

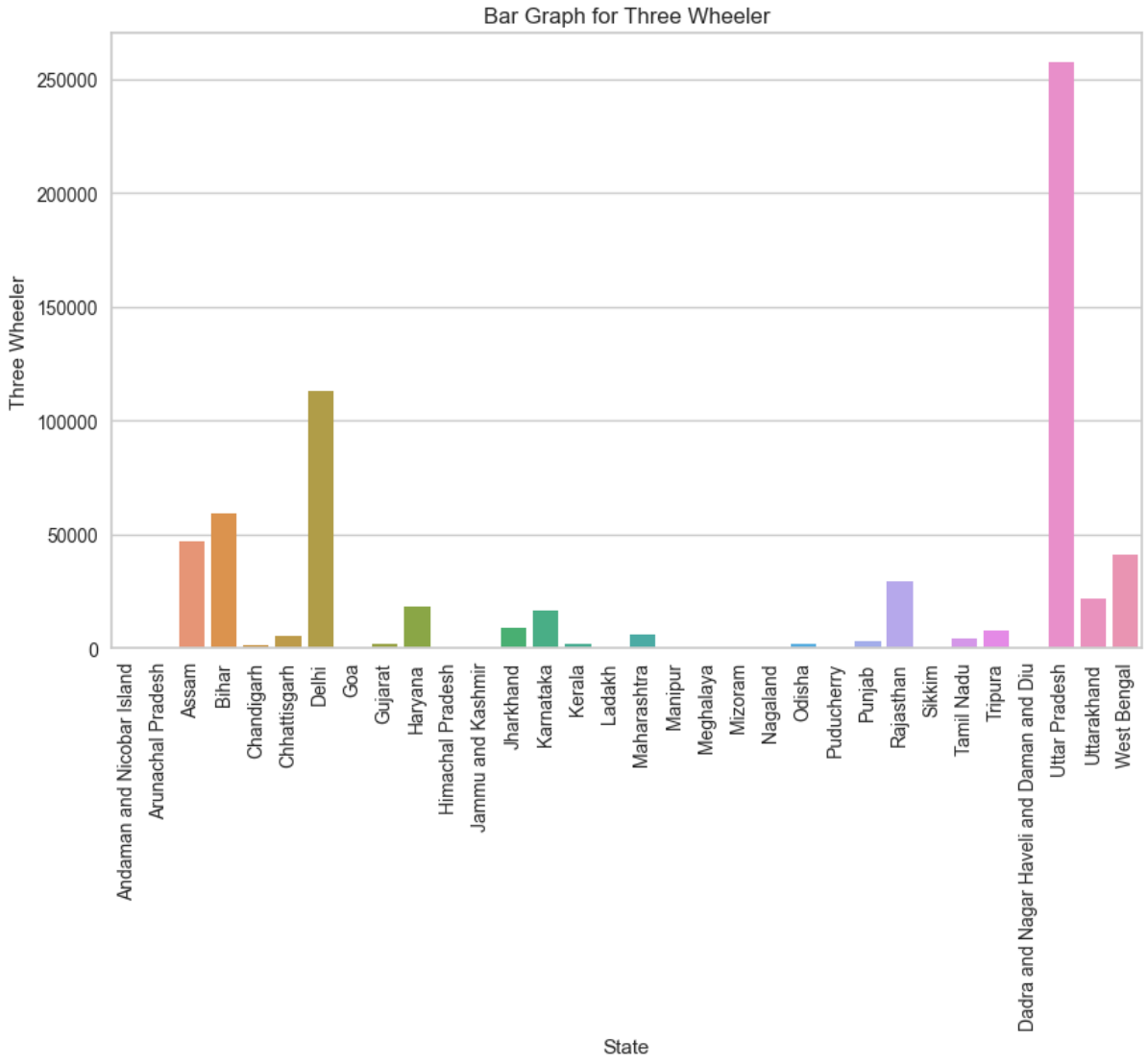

Pie Chart for Himachal Pradesh
Four Wheeler
Goods Vehicles
Other
Three Wheeler
5.9%
1.1%
7.4%
26.7%
58.9%
Two Wheeler

The dataset presents a comprehensive comparison of the number of vehicles of various types across all states. This extensive analysis offers valuable insights into the transportation patterns and preferences prevalent throughout the entire region.

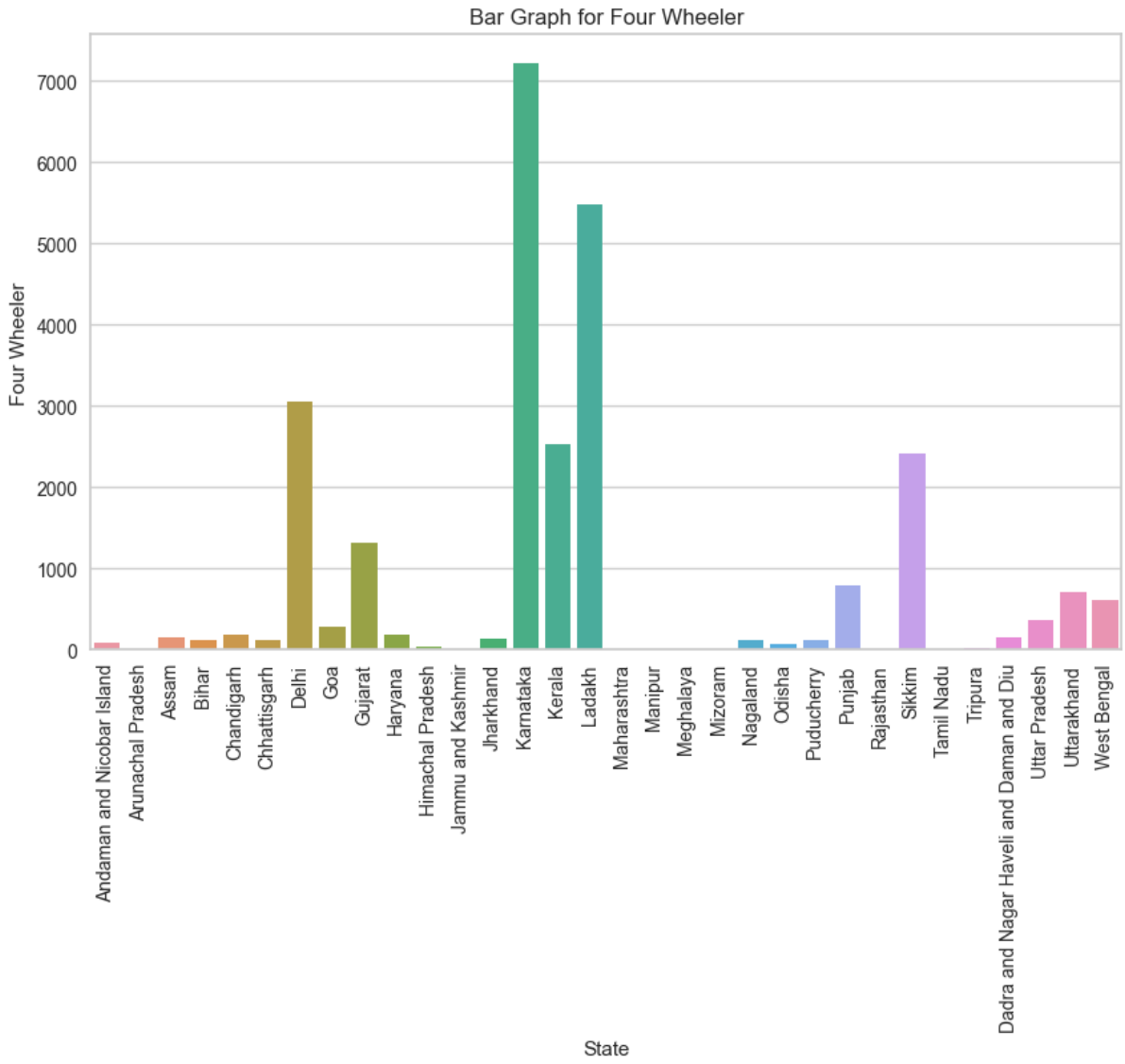
Starting with the "Two Wheeler" category, which includes motorcycles and scooters, we observe the distribution of these popular two-wheeled vehicles in each state. This data sheds light on the prominence of these nimble and efficient modes of personal transportation across the entire country.



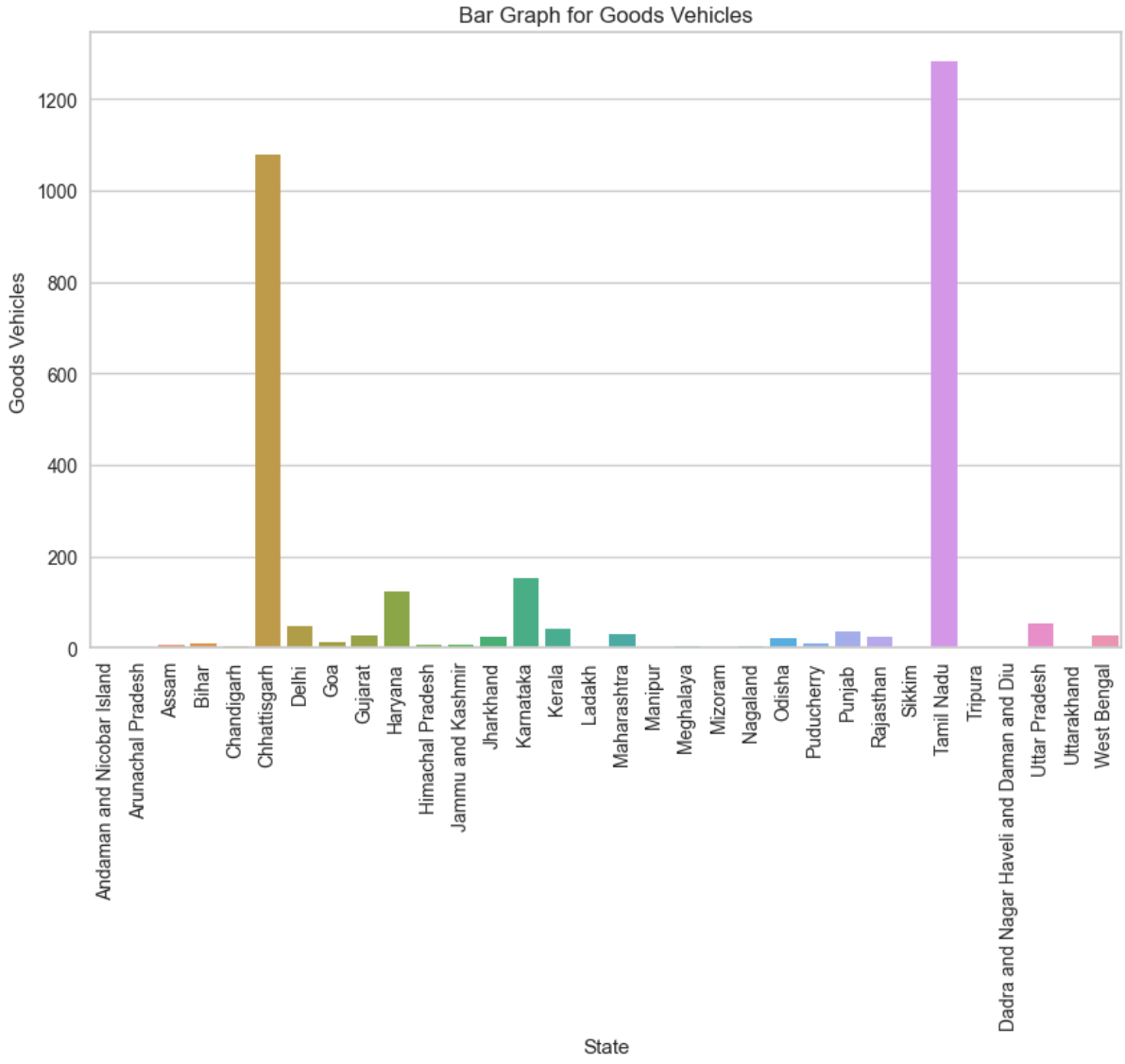
Moving on to the "Three Wheeler" category, encompassing vehicles like auto-rickshaws or similar tri-wheelers, we gain an understanding of their prevalence and usage as a vital mode of public and intermediate transportation across different states.



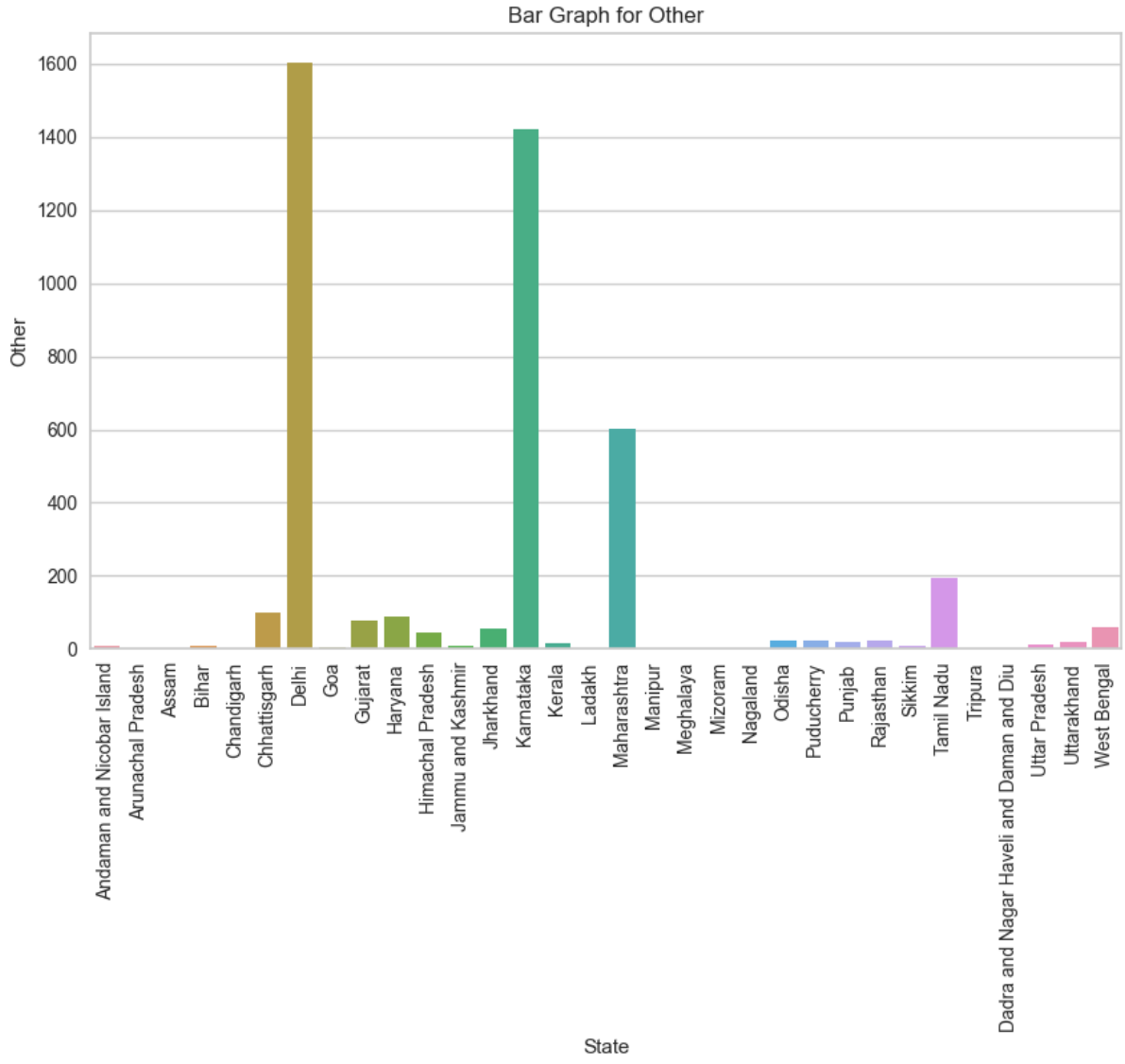
In the "Four Wheeler" category, which comprises cars, small trucks, and other four-wheeled vehicles, we explore the landscape of private and commercial transportation. This data allows us to analyze the patterns of car ownership and commercial vehicle usage, which play crucial roles in the country's economy and daily life.



The "Goods Vehicles" category offers critical insights into the logistics and transportation of goods and freight across all states. By examining this data, we can gauge the efficiency and magnitude of trade and distribution activities in each region.



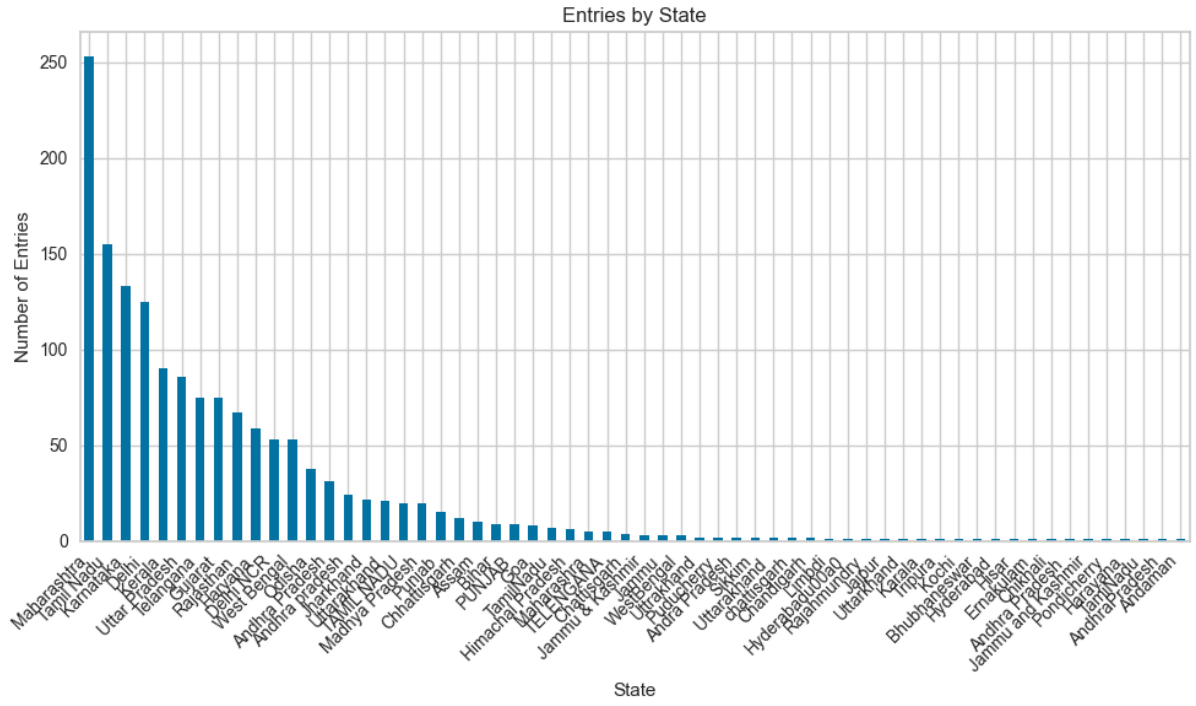
Lastly, the "Other" category accounts for any unique or specialized vehicles not covered in the previous classifications. This category's inclusivity allows us to discover and study unconventional modes of transportation that may be specific to certain regions or have niche applications.



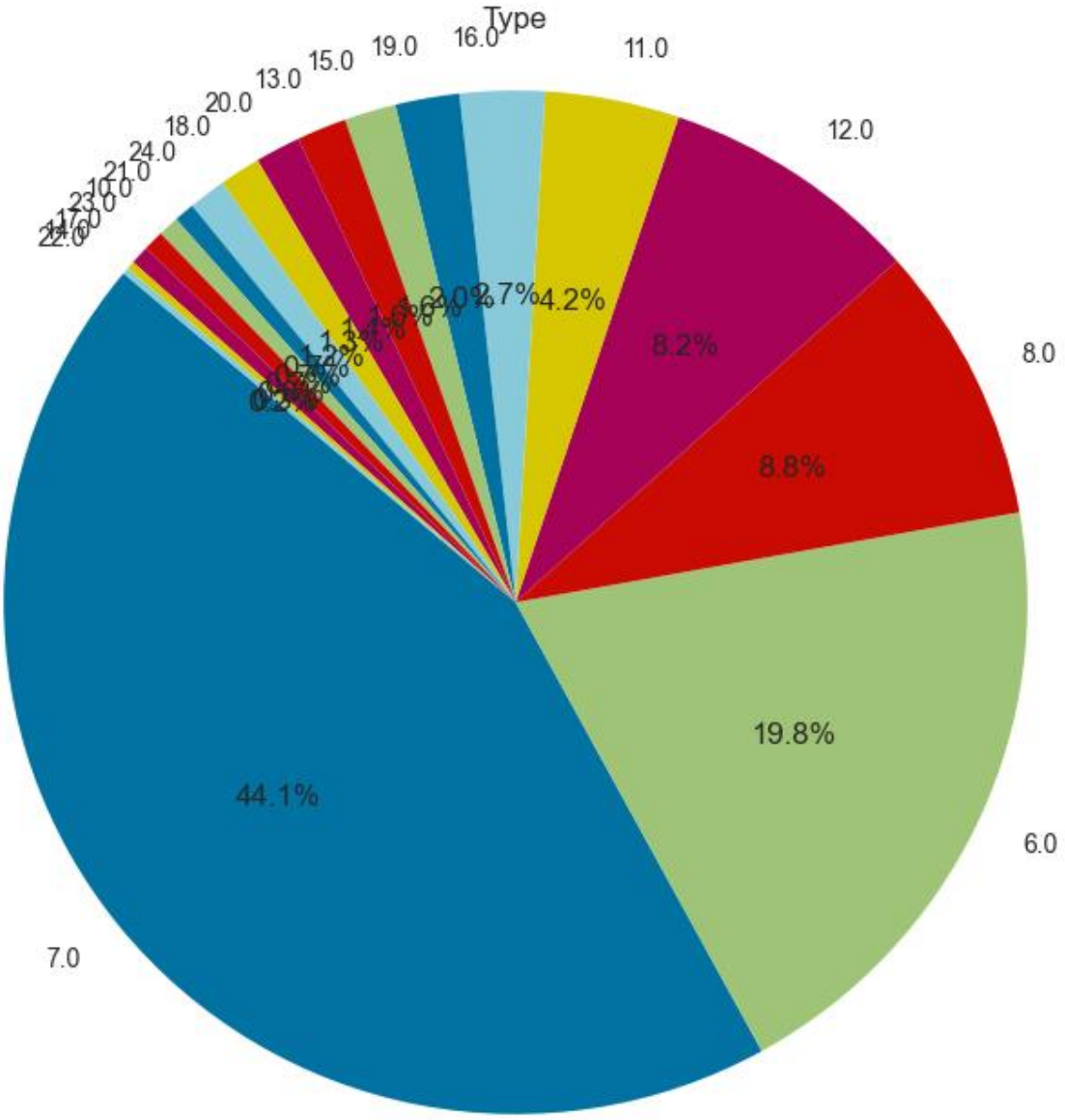
Dataset 1 →

The dataset also provides valuable information regarding the total number of charging stations in each state. Charging stations play a pivotal role in supporting the growth of electric vehicles (EVs) and are instrumental in promoting sustainable and eco-friendly transportation solutions.

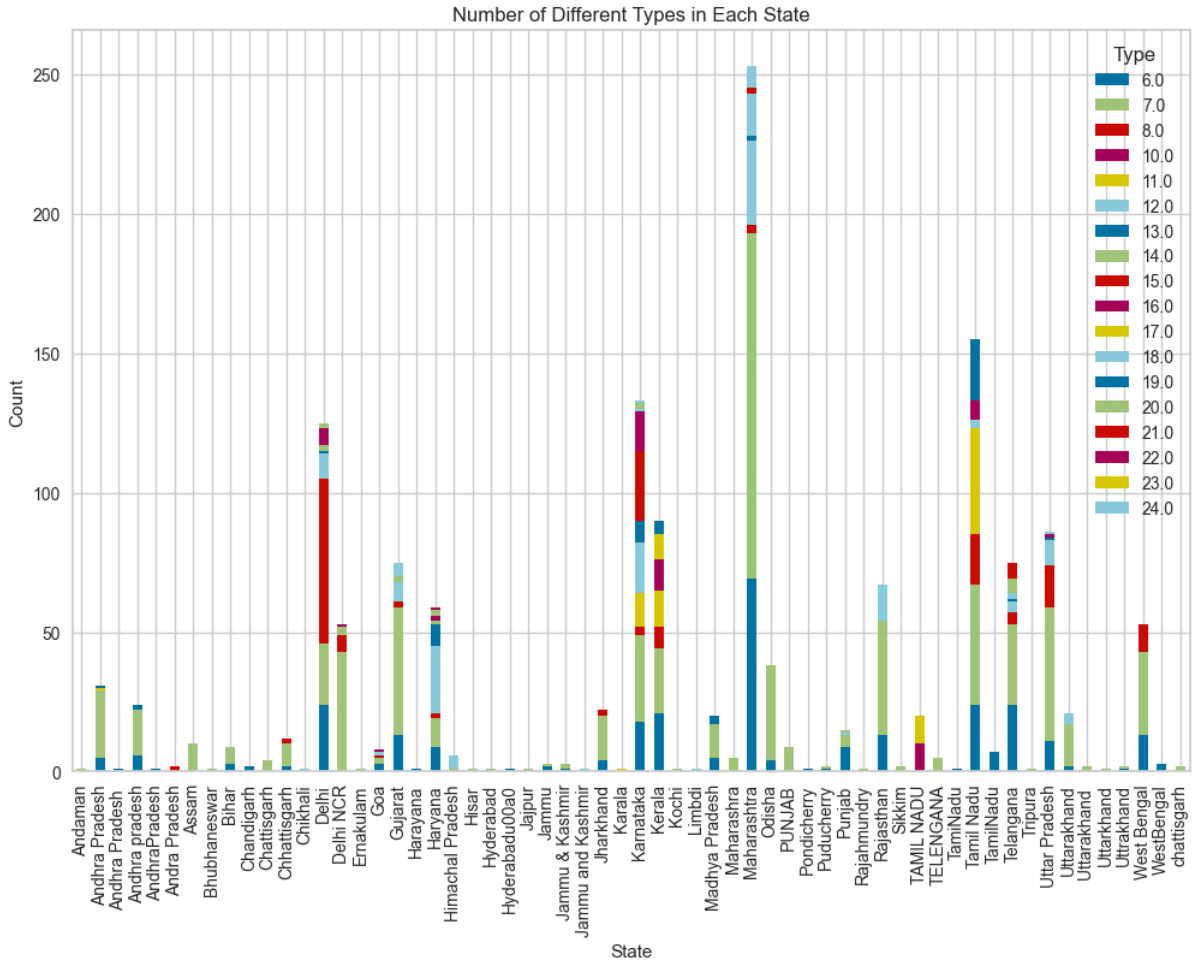
By examining the total number of charging stations across different states, we can gain insights into the level of infrastructure development and the readiness for electric mobility adoption in each region.



Different Types of Charging Station in country



Overall Types of charging Station in Each State



Extracting Segments

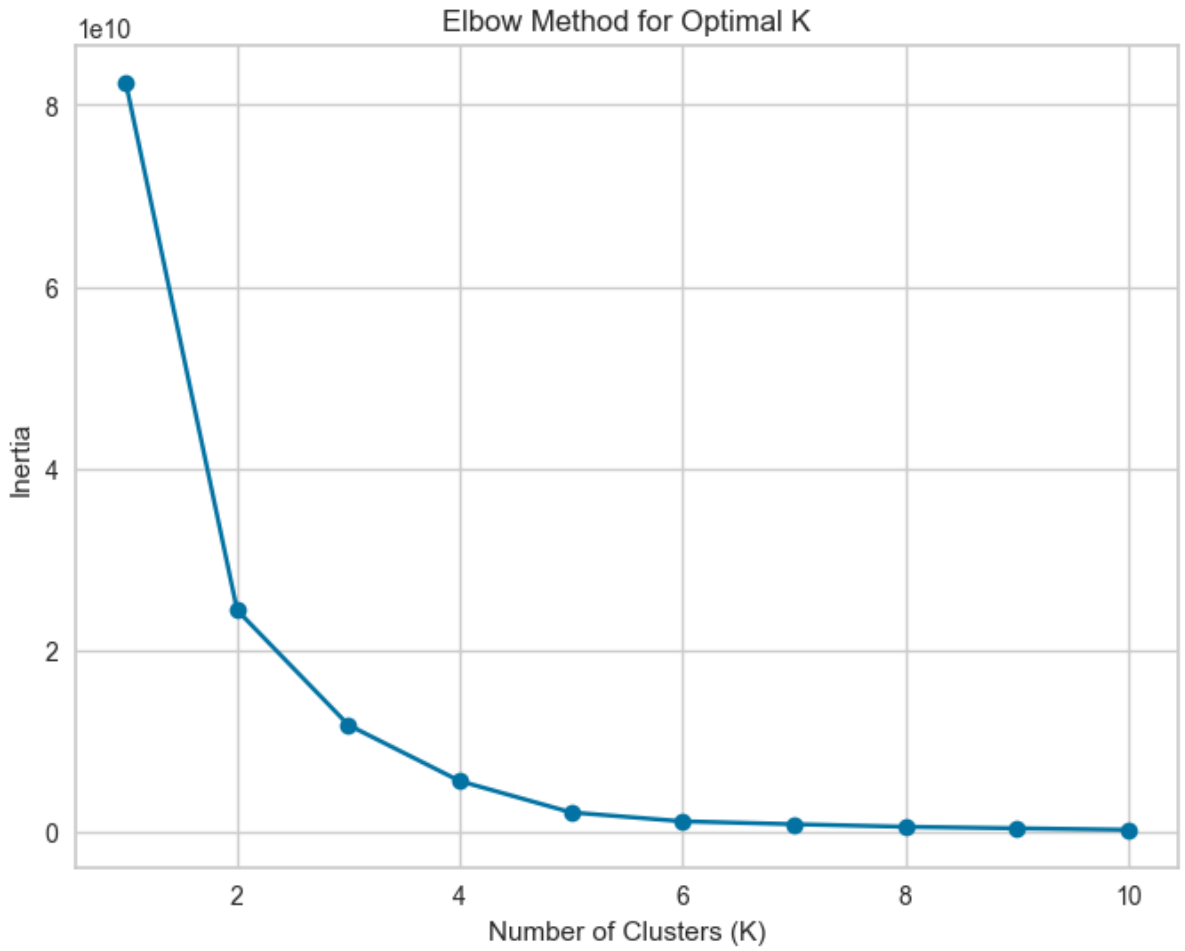
The *Elbow Method* is a popular technique used in data analysis, particularly in the context of clustering algorithms like K-means. Its primary purpose is to help determine the optimal number of clusters to use in a dataset.

When performing clustering, it's essential to find the right number of clusters that can effectively group similar data points together while maintaining a meaningful level of distinction between the clusters. The Elbow Method aids in this decision-making process by plotting the cost (or inertia) of the clustering algorithm against different values of K (the number of clusters).

The cost represents how far the data points within each cluster are from their respective cluster centers. As the number of clusters increases, the cost

typically decreases because the clusters become more specific and data points are closer to their cluster centers. However, at a certain point, adding more clusters may not significantly reduce the cost further.

The Elbow Method helps identify this point by plotting the cost against K and looking for the "elbow" or the point at which the cost curve starts to flatten out. This elbow point indicates the number of clusters that strike a balance between capturing enough variance in the data while avoiding excessive fragmentation. The optimal value of K is usually chosen at or just before the elbow point on the graph.



Analysis and Approaches used for Segmentation

Clustering is a popular unsupervised machine learning technique used to group similar data points together based on their similarities or distances in a dataset. The primary objective of clustering is to partition the data into distinct groups, known as clusters, where data points within the same cluster are more

similar to each other than to those in other clusters. This helps to identify underlying patterns, structures, or natural groupings in the data, even when the class labels are not known or unavailable.

Clustering is widely used in various domains, including customer segmentation, image segmentation, document clustering, anomaly detection, and recommendation systems, among others. It is an essential exploratory data analysis technique that allows researchers and data scientists to gain insights into the inherent structure of data.

One of the most popular clustering algorithms is the K-means clustering algorithm.

K-means Clustering:

K-means is an iterative algorithm that partitions a dataset into K clusters, where K is a user-defined hyperparameter representing the number of clusters desired. The algorithm aims to minimize the sum of squared distances (often called inertia) between data points and their corresponding cluster centers.

Here's a step-by-step overview of the K-means clustering algorithm:

Initialization: Randomly select K data points from the dataset as initial cluster centers (centroids).

Assignment: Assign each data point to the nearest cluster center based on the Euclidean distance or other distance metrics.

Update Centroids: Recalculate the cluster centers (centroids) by taking the mean of all data points assigned to each cluster.

Re-assignment: Reassign each data point to the nearest cluster center based on the updated centroids.

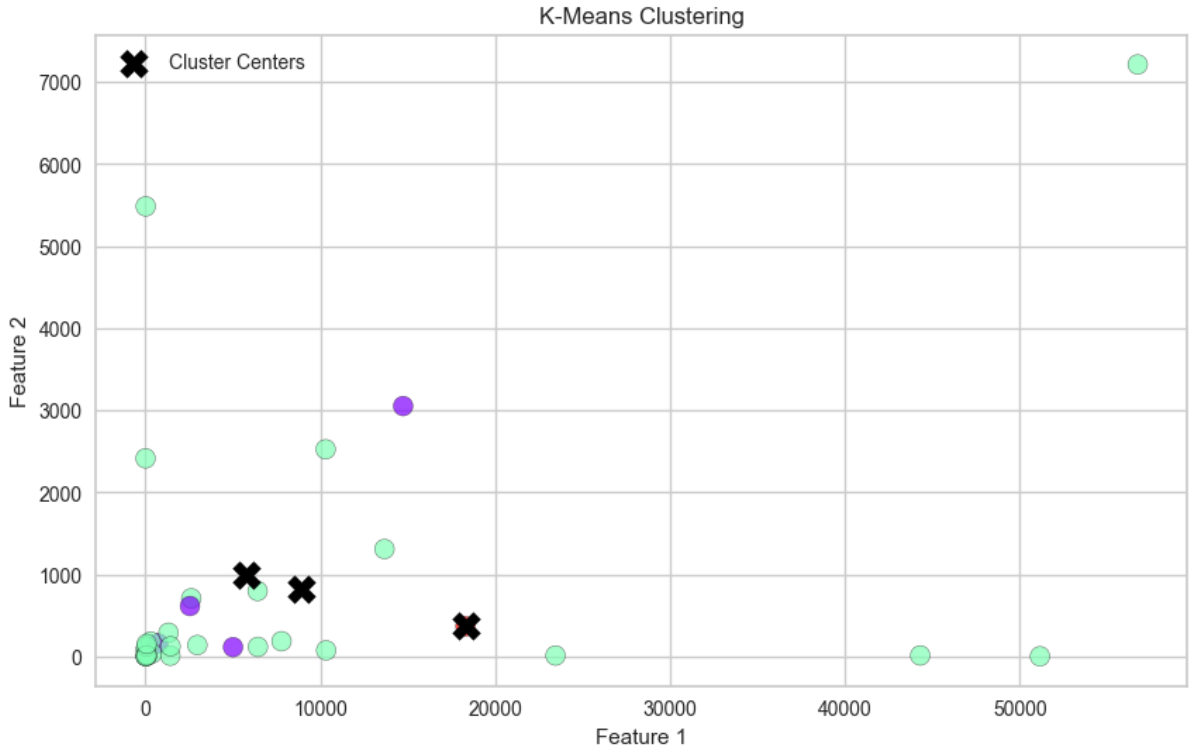
Iteration: Repeat steps 3 and 4 until convergence. Convergence occurs when the cluster assignments do not change significantly, or a maximum number of iterations is reached.

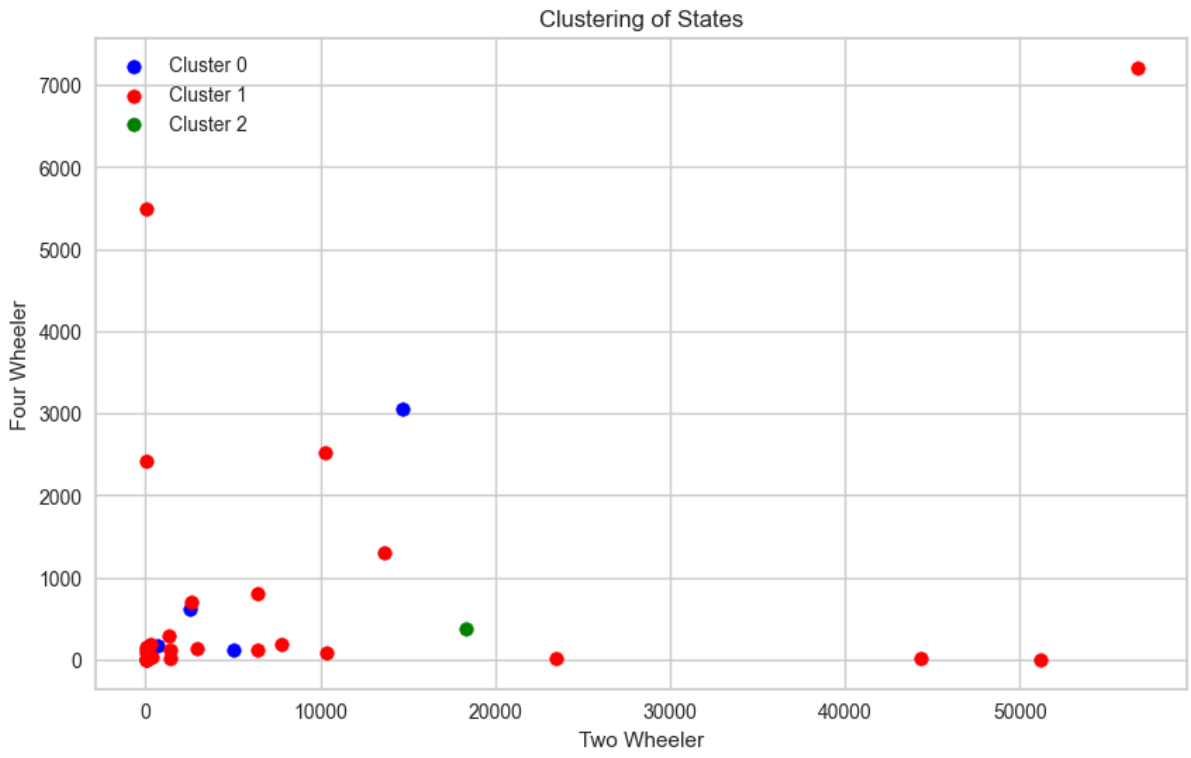
Final Result: The algorithm converges to a set of K cluster centers, and each data point belongs to one of the K clusters.

It's important to note that K-means is sensitive to the initial random selection of centroids. To mitigate this, multiple runs of the algorithm with different

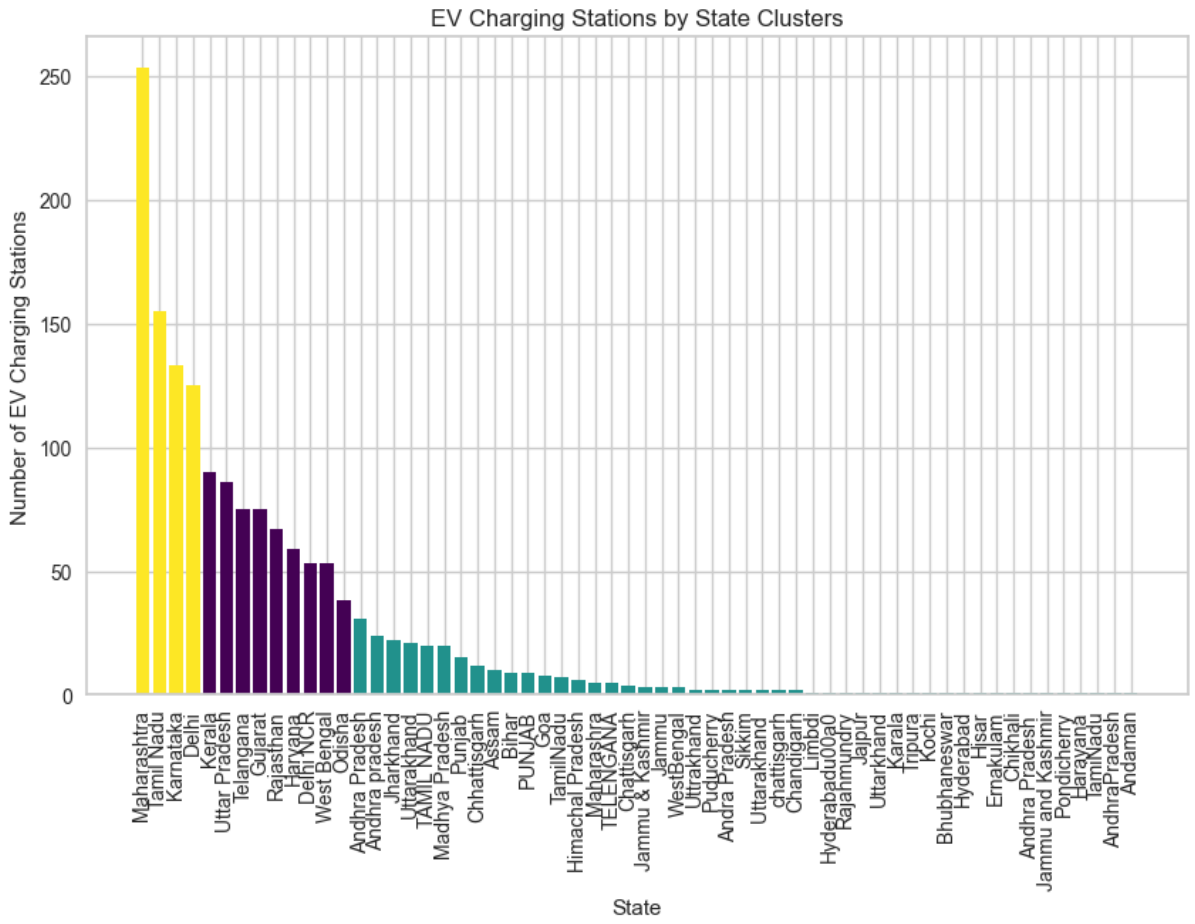
initializations are often performed, and the best clustering result (with the lowest inertia) is chosen based on evaluation metrics or domain knowledge

For Dataset (no of evs)





For Dataset(no of charging station)



Target Segments

Based on the analysis performed, it is evident that the most promising target segments for our product or service are the "Innovators" and "Early Adopters." These segments play a crucial role in the adoption and diffusion of new technologies and innovations. Let's explore these target segments in more detail:

Innovators:

Innovators are the trailblazers and visionaries among consumers. They are known for their willingness to take risks and embrace new technologies or ideas at the earliest stage. This segment represents a small percentage of the total market but holds significant influence. Innovators are often eager to try

out novel products or services and provide valuable feedback, acting as evangelists for the brand.

Characteristics:

High openness to new ideas and experiences.

Willing to invest time and resources in exploring innovative solutions.

Adept at understanding and adapting to emerging technologies.

Thrive on being early adopters and influencers within their social circles.

Tend to have a strong interest in research and development.

Marketing Strategy:

To attract Innovators, our marketing strategy should emphasize the uniqueness and cutting-edge features of the product or service. Highlighting its novelty, functionality, and potential to revolutionize the industry will appeal to this segment's desire for experimentation and staying ahead of the curve. Leveraging social media, tech forums, and industry events can help create a buzz and attract Innovators to be the first to experience our offering.

Early Adopters:

Early Adopters follow closely behind the Innovators and are quick to embrace new technologies once they see a compelling value proposition. They represent a slightly larger portion of the market compared to Innovators and often act as opinion leaders within their social and professional networks. Early Adopters' willingness to try new things makes them vital for driving the initial growth of an innovative product or service.

Characteristics:

High social influence; their opinions are respected and followed by others.

Keen interest in technological advancements and innovations.

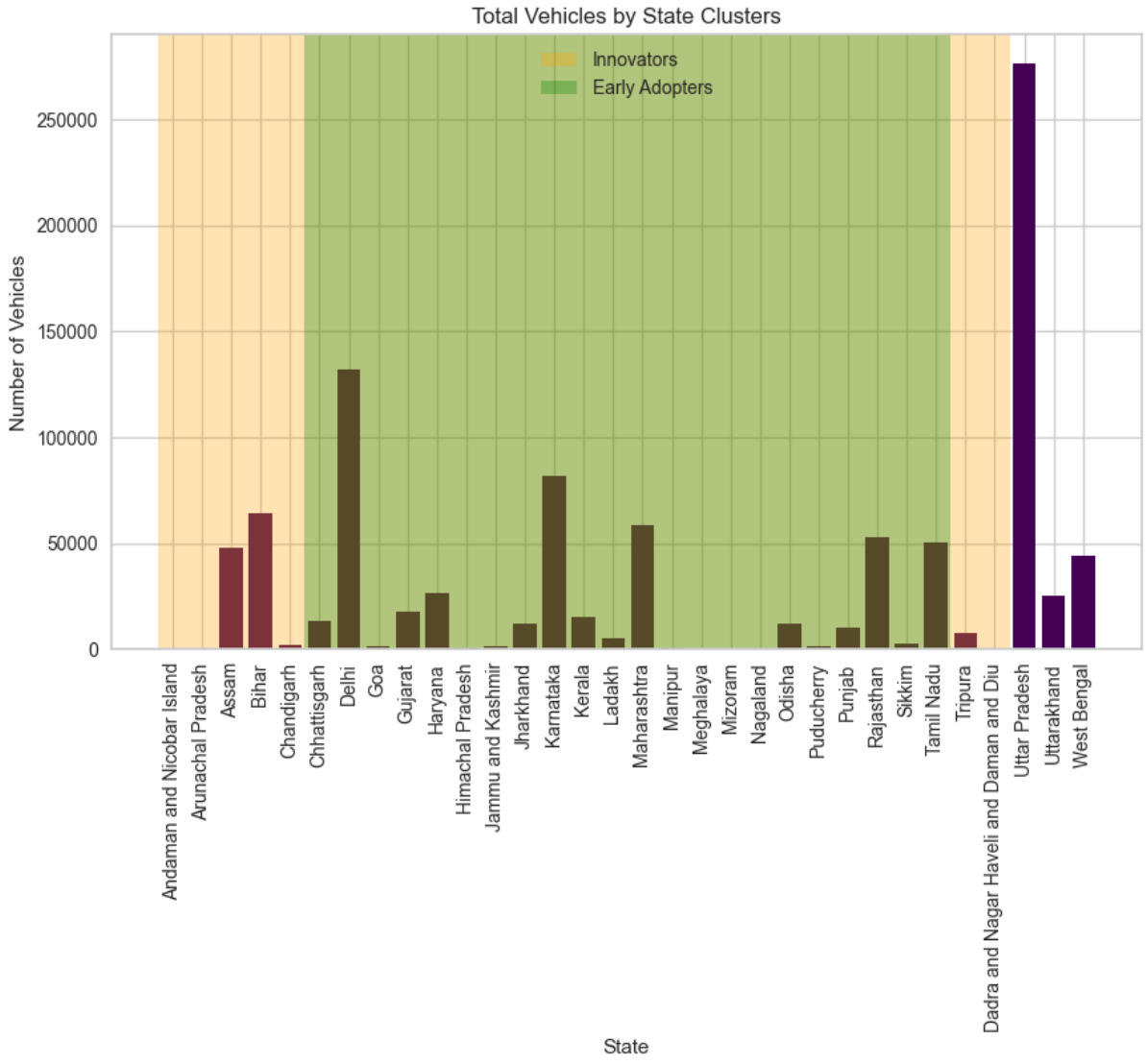
Willing to invest in new products to gain a competitive advantage.

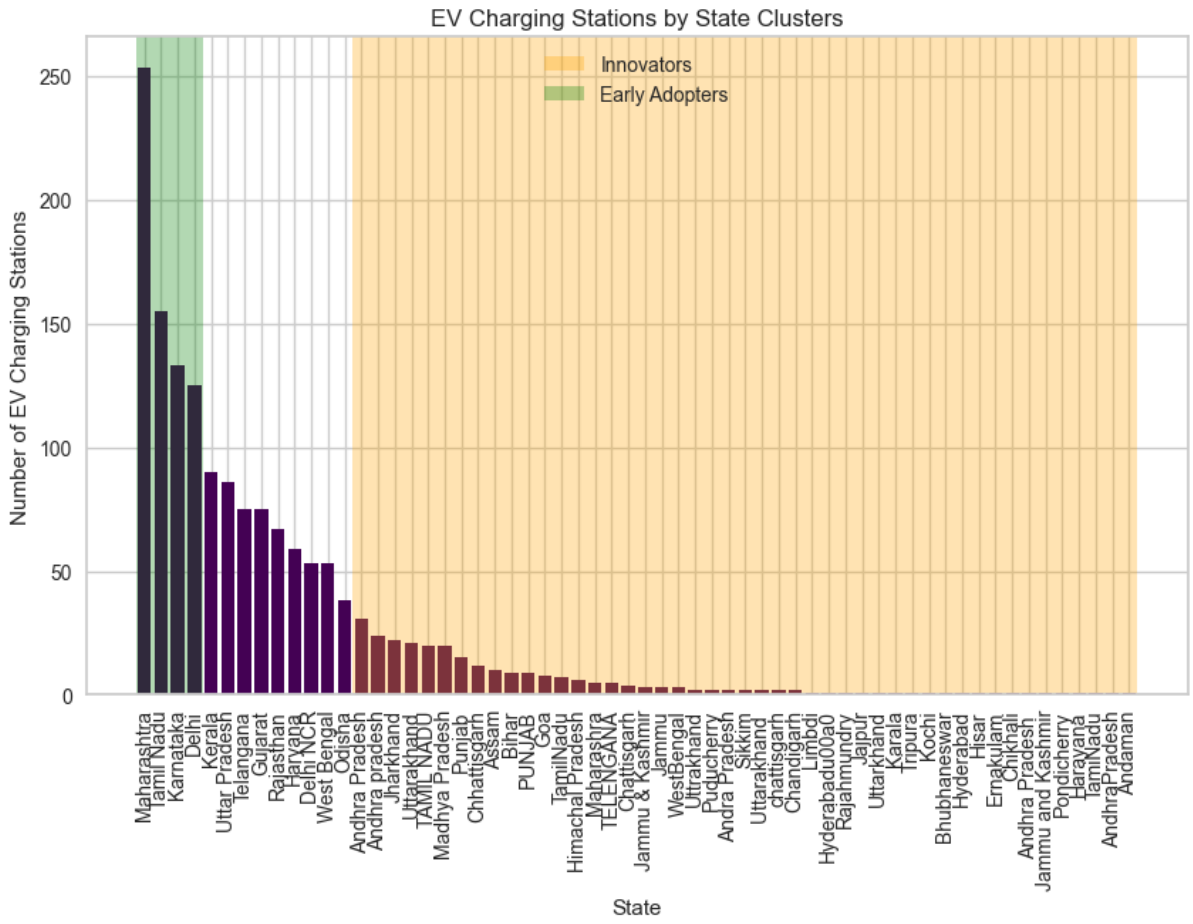
Prefer being among the first to adopt novel solutions.

Provide valuable feedback and recommendations to refine the offering.

Marketing Strategy:

Engaging Early Adopters requires building credibility and trust in the product or service. Leveraging testimonials, case studies, and endorsements from Innovators who have already embraced the solution can help build this trust. Offering early access, special incentives, and personalized communication can also attract Early Adopters to the offering. Additionally, fostering a sense of community and encouraging feedback will create a positive experience for them.





Conclusion

In conclusion, the analysis has provided valuable insights into the dataset, offering a comprehensive understanding of the transportation landscape in various states. We have explored the distribution and comparison of different types of vehicles, including Two Wheelers, Three Wheelers, Four Wheelers, Goods Vehicles, and Other vehicles.

The dataset's comparison across all states has shed light on the prevalence of each vehicle category, enabling us to identify regional preferences and transportation patterns. Additionally, the inclusion of null values and their subsequent filling based on specific ratios has enhanced the dataset's completeness and accuracy, providing a more robust foundation for further analysis.

Furthermore, the examination of the number of charging stations in the entire country has highlighted the advancements in electric vehicle infrastructure.

The different types of charging stations, such as Level 1, Level 2, DC Fast Charging, and others, underscore the growing support for electric mobility and sustainable transportation options nationwide.

With this wealth of information, stakeholders, policymakers, and businesses can make informed decisions to address transportation challenges, promote green mobility solutions, and ensure efficient and sustainable transportation networks. The insights gained from the clustering analysis can aid in targeted marketing strategies, resource allocation, and infrastructure development, focusing on the Innovators and Early Adopters as key segments for driving initial adoption and market growth.

Overall, the dataset's thorough analysis and subsequent clustering have provided a solid foundation for strategic planning and decision-making, fostering a better understanding of the transportation landscape and contributing to the development of a more sustainable and connected future. As we move forward, these insights will serve as a valuable guide in shaping transportation policies, advancing electric mobility, and meeting the evolving needs of the transportation industry.

References :

https://data.gov.in/resource/stateuts-wise-electric-vehicles-evs-under-operation-under-phase-ii-faster-adoption-and

https://www.kaggle.com/datasets/saketpradhan/electric-vehicle-charging-stations-in-india