Electric Vehicle Market Segmentation

Name: Jatin Saini

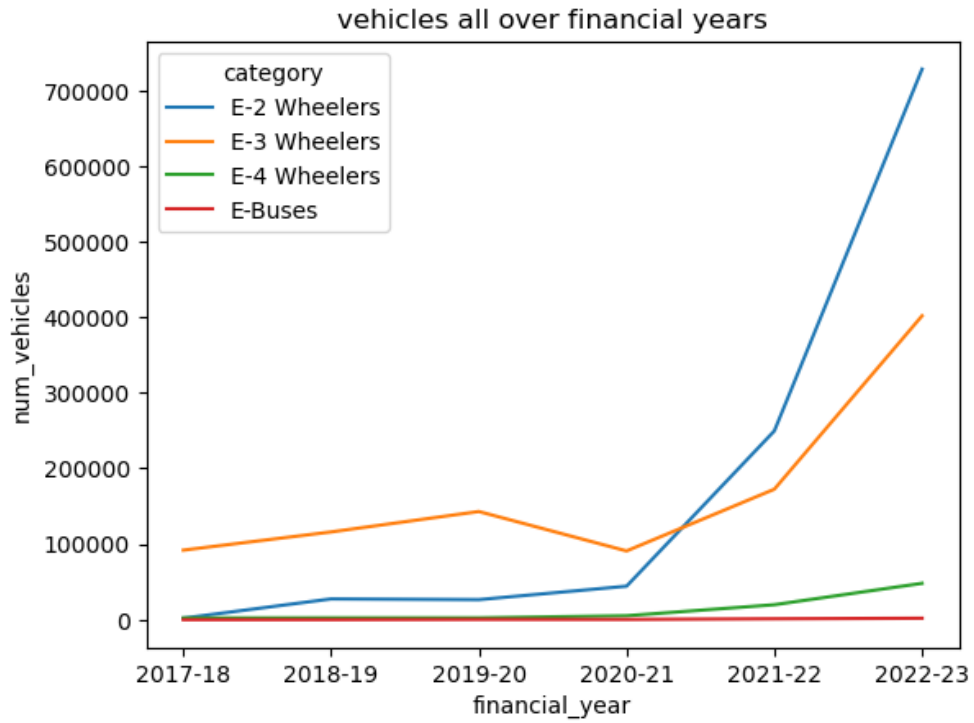
Code:

Electric-Vehicle-Market-Analysis/ev_market_analysis at main · jatin30121996/Electric-Vehicle-Market-Analysis · GitHub

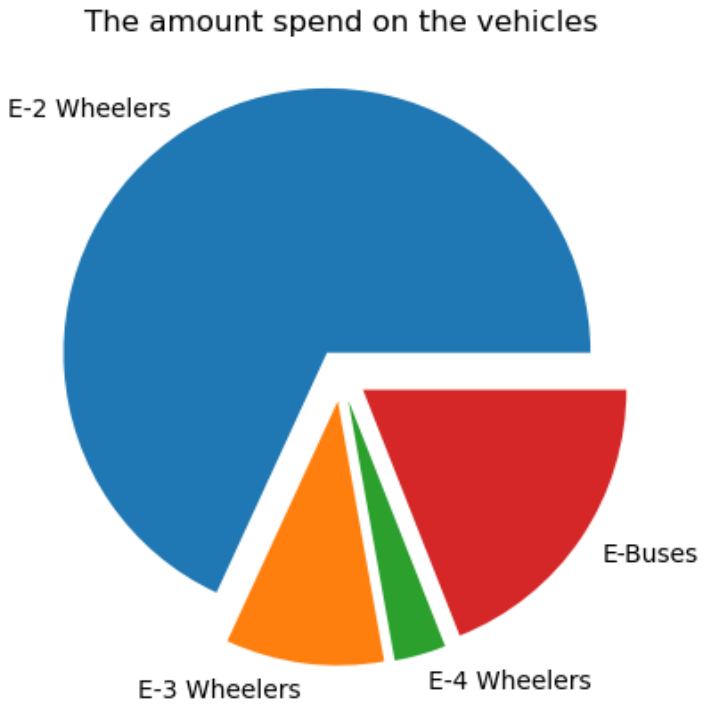
Overview

Electric Vehicle becomes the important part of the people life nowadays other then vehicle run of fuel. And the demand is increasing day by day, people start deviating towards the electric vehicle as they are user friendly as well as environment friendly and efficient in use. The restriction of using the electric vehicle is the not proper installing the charging ports across the country, however this lacking is overcoming by the initialization of installing the charging stations across the country. The demand of electric vehicle is increasing day by day.

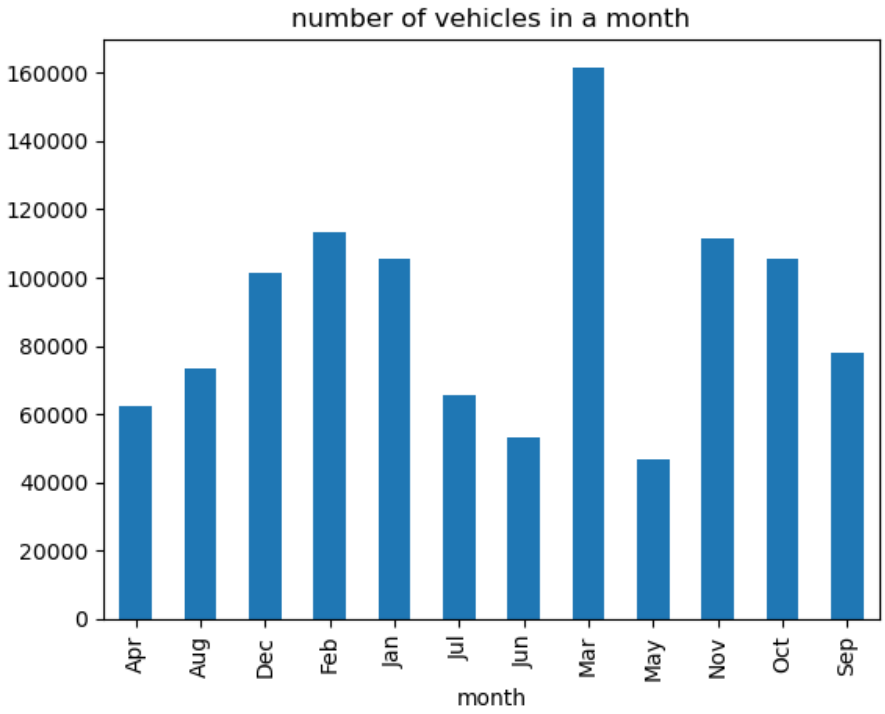
Market Overview



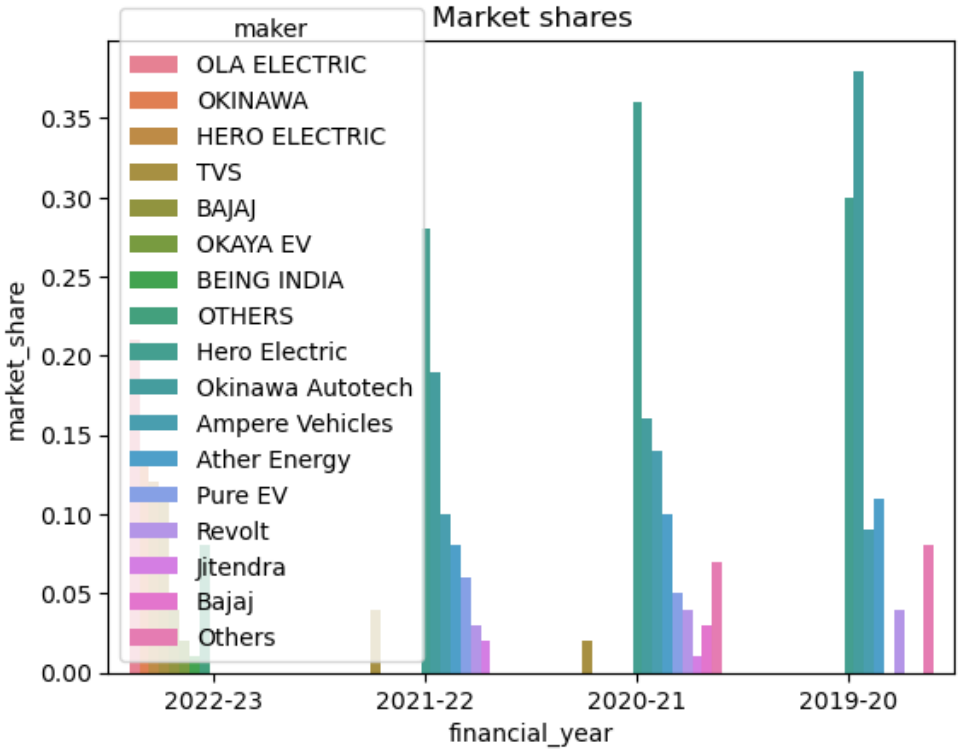
The two wheelers electric vehicle is the best choice shown by people and demand is increasing day by day. People deviating toward the electric vehicle more and more. Most used electric vehicle is 2 wheelers as it is the best choice shown by persons then other vehicles.



People spend more money on 2 wheelers electric vehicle then 3, 4 or E buses.



March is most selling month for E-Vehicle and May is least.



Paying the attention on market share in 2019-20 Hero electric was holding the most shares however till reaching the 2023 the Ola Electric replaces there place in holding the large amount of market share. Still there is large difference between the numerical value of shares as Hero electric shares were above 0.35 however Ola Electric shares were only 0.21.

Deciding not to segment

Electric Vehicle nowadays plays a crucial role in all over transportation. People start choosing the electric vehicle for movement from one place to another place. In that 2 Wheelers plays very crucial role, people start choosing the 2 wheelers over other Electric vehicle. Choosing whole market as a one is still effective because it will save the resources and cost. Companies not need to make any further plans for segmentation. But Heterogeneity among users can help to analyze the market in a better way and we can focus on the particular users of interest to give them better options, so they can choose to get the effective output in electric vehicle market.

Specifying the Ideal Target Segment

In 2 wheelers we have to choose the target segment very wisely so that we can analyze the market based on that and get the proper outcome from that. These have to give the whole market details to where and why the person use to and from where we can modify the criteria and help

the companies to get the proper result as well as they can improve the selves in the future. Selecting those ideal features for segmentations can truly help the companies to improve their market selling the product.

Collecting data

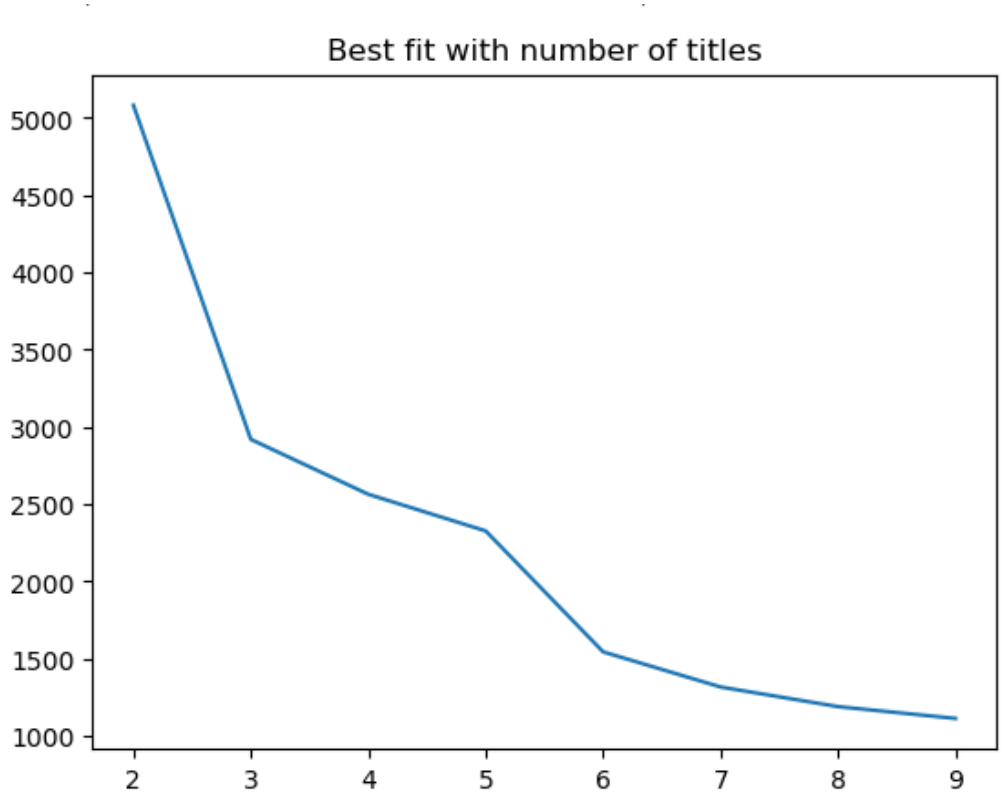
Data is collected from different sources mostly from the kaggle.

Exploring Data

For data segmentation we choose the few features which seem to be unique and effective so that we can analyze the data in a better way without any complications. These features contain the numeric values and some Nan values which can be replaced by the some values like 0. We choose the rating, visual appeal, reliability, performance, service experience, extra features, comfort, Maintenance cost, value of money for segmentation. These features look similar to each other and do not contain the regular data that used for regression. These features contain the classes' type numeric data which can be easily divided into clusters and that can be classifying the data.

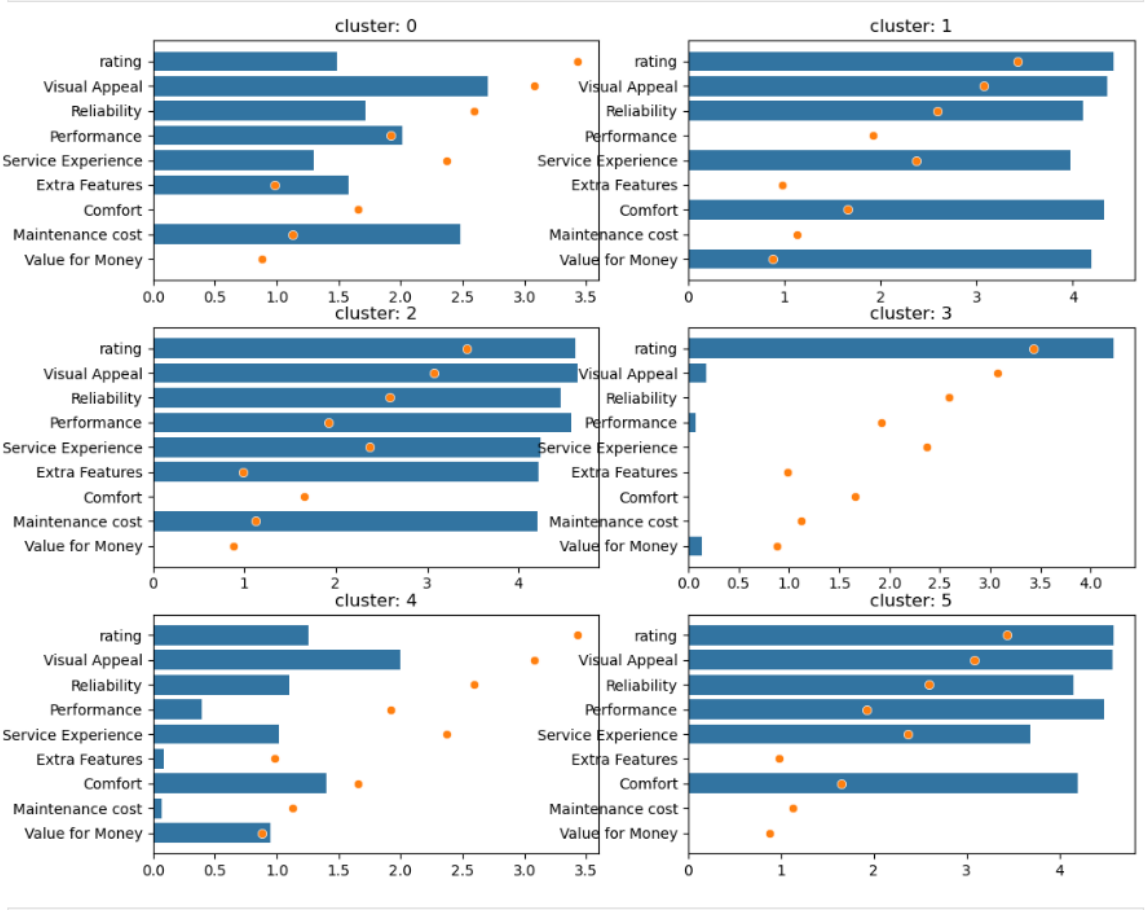
Extracting Segments

For extracting the segments we use the KMeans clustering algorithm for extracting the segments.



From the above elbow method we choose 6 segments as data line was strictly falling from above and after that line does not show that much variation.

Profiling Segments

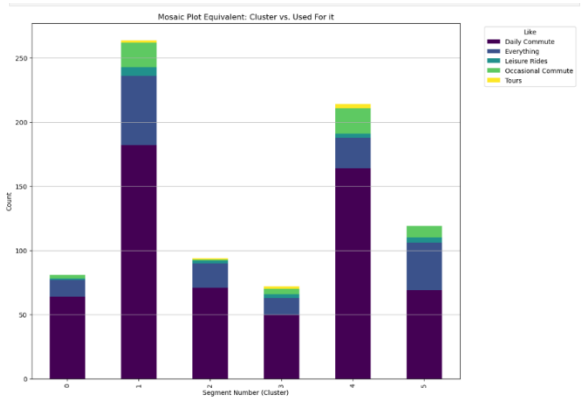


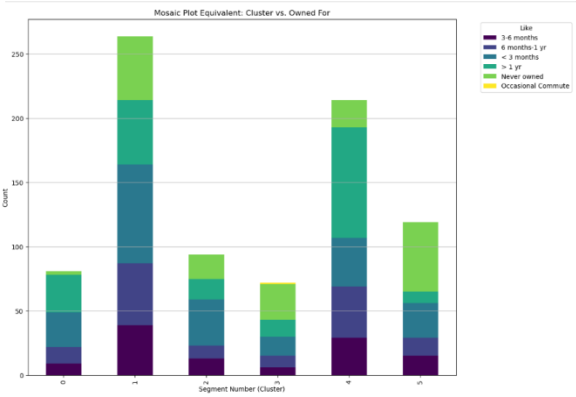
The second cluster contains the most of the consumers and cluster 3 contain the least consumers, however cluster 5 and cluster 1 contain the almost same consumers with 2 cluster. So these 1 and 2 and 5 covers more consumers and remaining the least.

More consumers is on the 1, 2, 5 segments. It almost contain 75% of total consumers is covered by these segments. However 0, 3 and 4 contain the least consumers.

So, it easy to divide the consumers based on these criteria and we can provide specific features to the users so the steps will be taken to highlight the particular segment and enhance the productivity as well as which segments are lacking we can make them to increase the productivity by making some crucial steps.

Describing Segments





People used the Electric vehicle for daily commute the most and tours for least. Cluster 1 and cluster 4 contain the more than half of the consumers and almost 80 percent people used it for daily commute. Some people used it for their all purpose. A very small fraction used it for everything. There are equal fraction can be seen in owned for different purposes. People rent it for various purposes like daily commute for most for different interval like for greater than 1 year or 6 months and et-cetera. The large number of fraction ride it for more than 5000 Km, almost 75 percent people used it for it. Some people ride it for 15000 km but very small fraction of people lies there for it.

Solution

Ola is the number 1 electric vehicle company at the moment. The reason is it provide large features to their users like they can used it for according to there will if they want to use it for ride more than 15000 km they simply use it for this purpose. So, people do not have restrictions or limitations for it. They can use vehicle according to their will. Moreover they provide extra thing like online application to own the vehicle for ride it. More over for online delivery of food or like rent it as a taxi, the Ola provide these features. People are fond of it. Even people replace the name of taxi as an Ola. If someone wants to call up the taxi they do not say I call the taxi instead they will say I call the Ola for travelling. These great features make the people to fond the Ola and large fraction of people prefer Ola and that's the reason their market shares are at peak. The solution for other companies is simple which are, they do not have to give them limited options for rent the bike. Moreover companies do not limit themselves to the offline business. Online becomes the major part now for business, they have to step in to the online business as well. They not only have to sell the bike but also they have to give the option to rent it as long as the user want and for all purposes. Even for taxi as a bike or food or everything. They have to start from the larger cities if they are successful then they have to go the less dense cities. If there are a more options in front of people they have the more chance to select the preferred company vehicle.

Name: Ashutosh Sunil Bhirud

Code:

https://github.com/asbhirud/Data-Science-Projects/tree/c91b84012bff3a0f3a36d47febc483b97ca297ae/EV%20Market%20segmentation

Conclusion

At the conclusion of our research, we successfully identified our target audience: individuals with moderate incomes, postgraduate education, and who are married, typically residing in areas with a high availability of charging points. While demand fluctuated between two-wheelers and four-wheelers, two-wheelers remained more popular due to their mobility and agility.

This report has been instrumental in helping us identify potential buyers for our upcoming electric vehicles. Most of the data was sourced from reputable websites, ensuring the reliability of the findings.

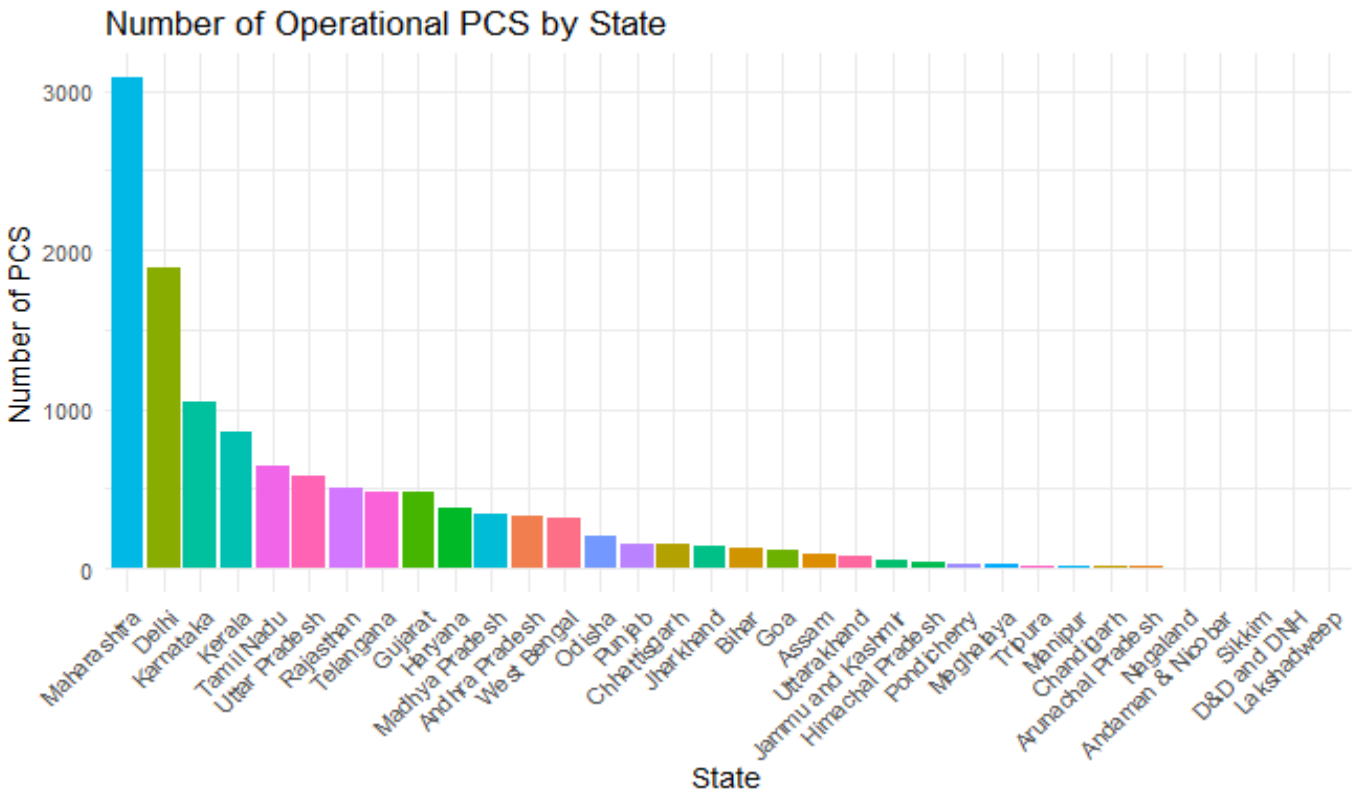
The best market segment for electric vehicles (EVs) in India is a complex choice that depends on factors like market conditions, consumer preferences, and the startup's goals. India's fast-changing market presents many opportunities for EV adoption across different segments.

Success will come from thorough market research to find the segments with the most growth potential. Urban commuters, delivery services, fleet operators, eco-conscious consumers, and government fleets are all promising groups, but the right choice should match the startup's strengths, resources, and long-term plans.

Additionally, addressing India's unique challenges—such as the need for better charging infrastructure, government incentives, and making EVs affordable—will be key to making them more attractive and accessible to consumers.

As the EV industry grows, startups should stay flexible and responsive to changes in the market. By keeping up with consumer trends, using new technology, and working with important partners, startups can help drive EV adoption in India and support a more sustainable future in transportation.

Name: Ayan Bashir Sheikh



Maharashtra has the highest number of PCS, significantly outpacing **Delhi**, the second-highest state. - **Karnataka** and **Kerala** also show substantial operational PCS numbers.

Moderate States: States like **Tamil Nadu**, **Uttar Pradesh**, and **Telangana** have moderate PCS counts, indicating development potential.

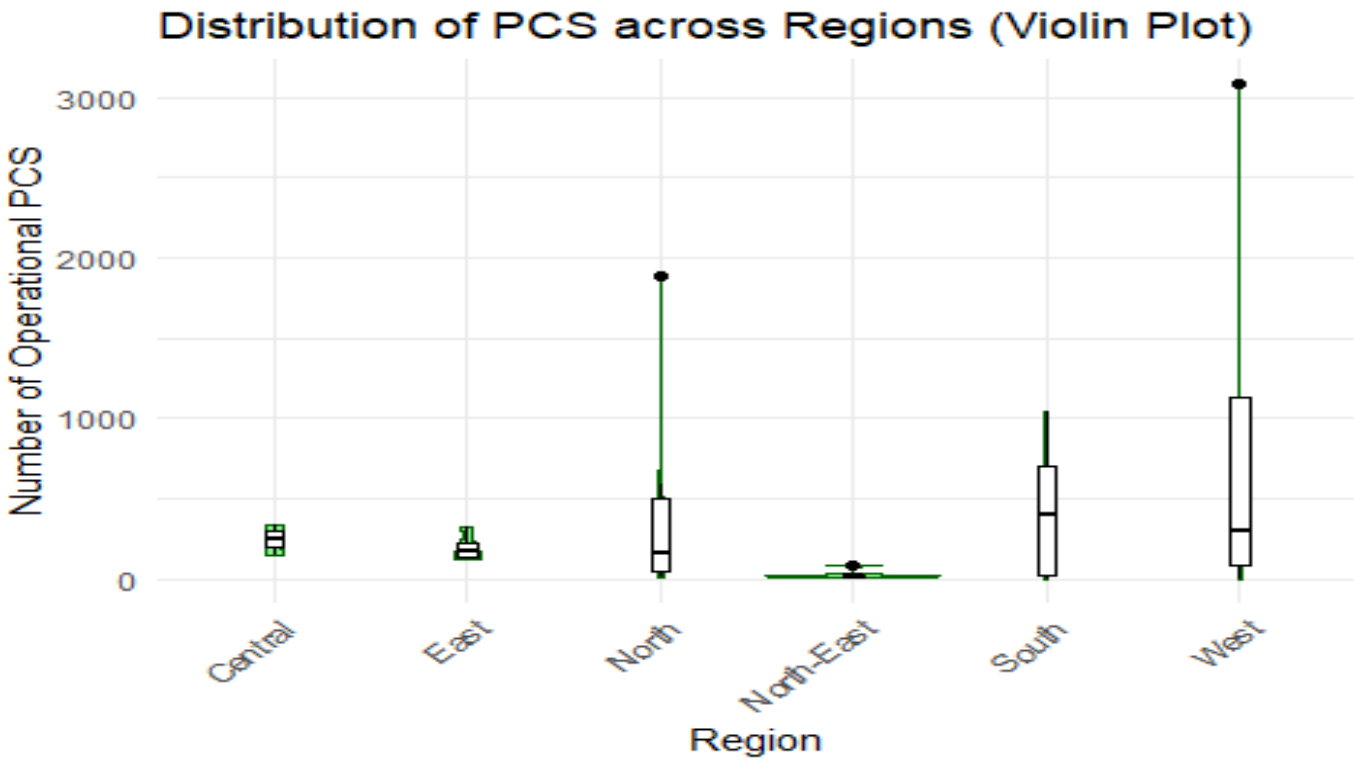
Low Representation: Several states, including **Arunachal Pradesh**, **Nagaland**, and **Lakshadweep**, exhibit very low operational PCS, suggesting infrastructural challenges.

Hypotheses Testing

To test if the average number of Power Conversion Systems (PCS) is equal across different regions in India, we formulate the following hypotheses:

- **Null Hypothesis (H_0)**:

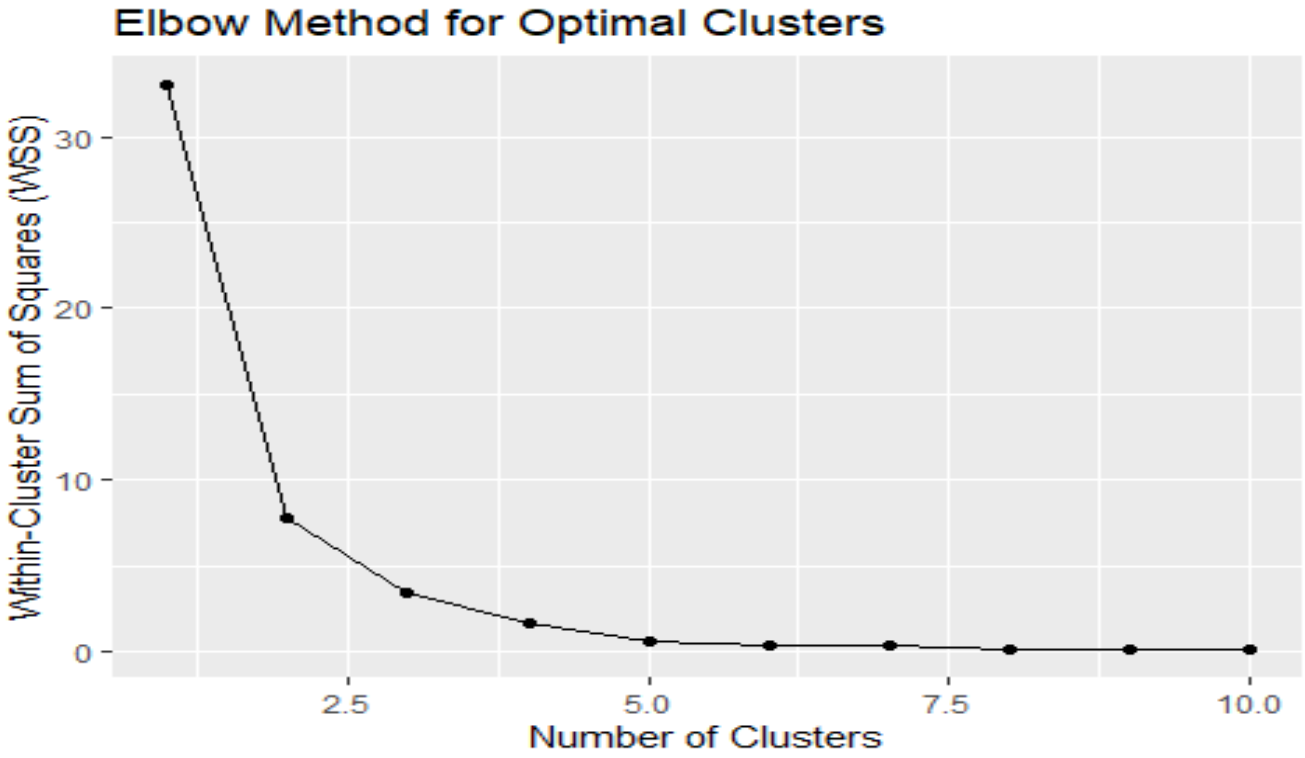
    - The average number of PCS is equal across all regions.
    - *There is no significant difference in the average number of PCS between the regions.*
- **Alternative Hypothesis (H_1)**:
    - The average number of PCS is **not** equal across all regions.
    - *At least one region has a significantly different average PCS count.*



1. **North Region**:
    - Exhibits the **highest number of PCS**, indicating robust infrastructure and investment in electric vehicle technology.
    - The wide distribution suggests significant variability, with some states having exceptionally high numbers of PCS.
2. **South and West Regions**:
    - Show moderate PCS counts, suggesting decent infrastructure but less than the North. The presence of outliers indicates some states may have invested significantly more.
3. **East and Central Regions**:

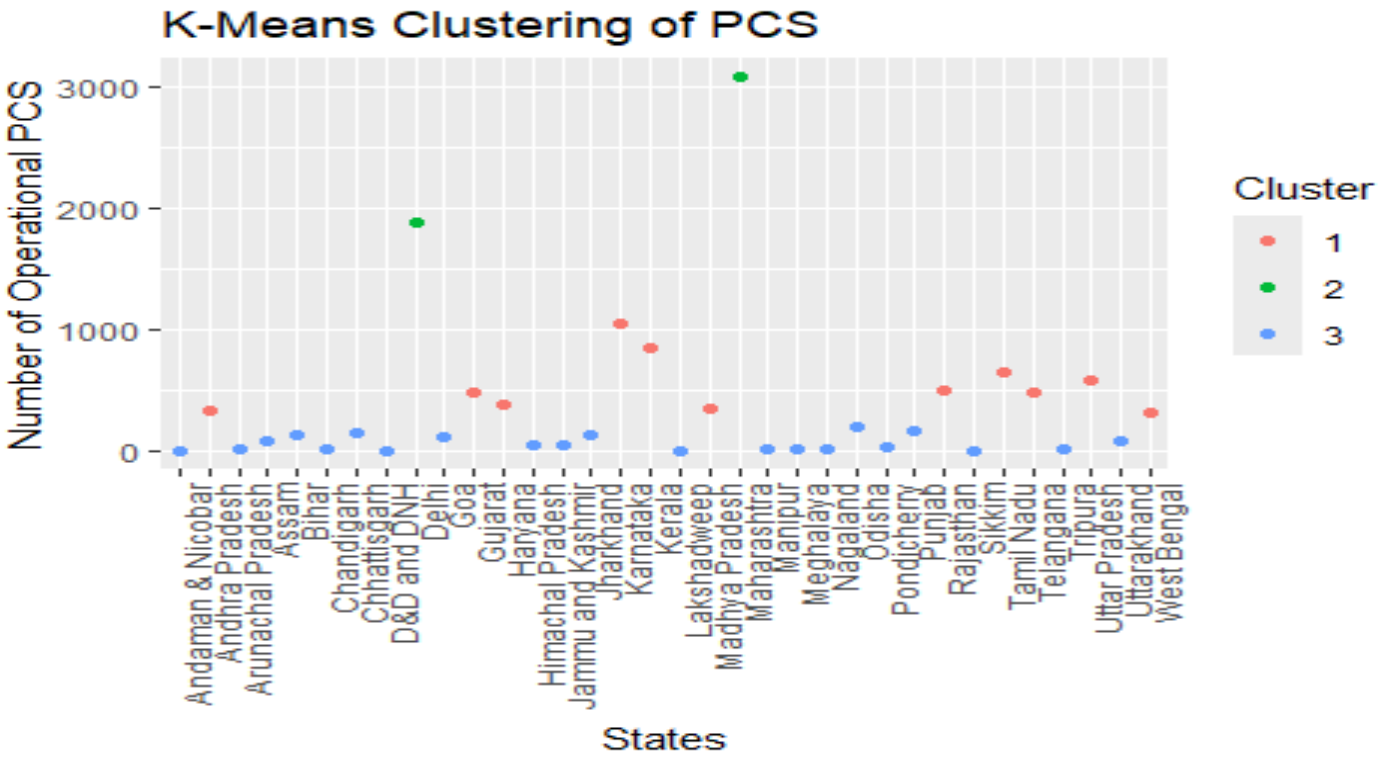
- Display narrow distributions and low PCS numbers, signaling potential gaps in electric vehicle infrastructure and opportunities for growth.

4. **North-East Region**:
    - Has the **lowest number of PCS**, highlighting a critical area for development and support to enhance electric vehicle adoption.



- The plot exhibits a noticeable “elbow” at **3 clusters**, where the WSS sharply decreases and then levels off.
- This suggests that **3 clusters** is the optimal number for grouping the data, as adding more clusters yields diminishing returns in improving the cluster compactness.

K-Means clustering of operational PCS (Public Distribution System Centers) across various Indian states.



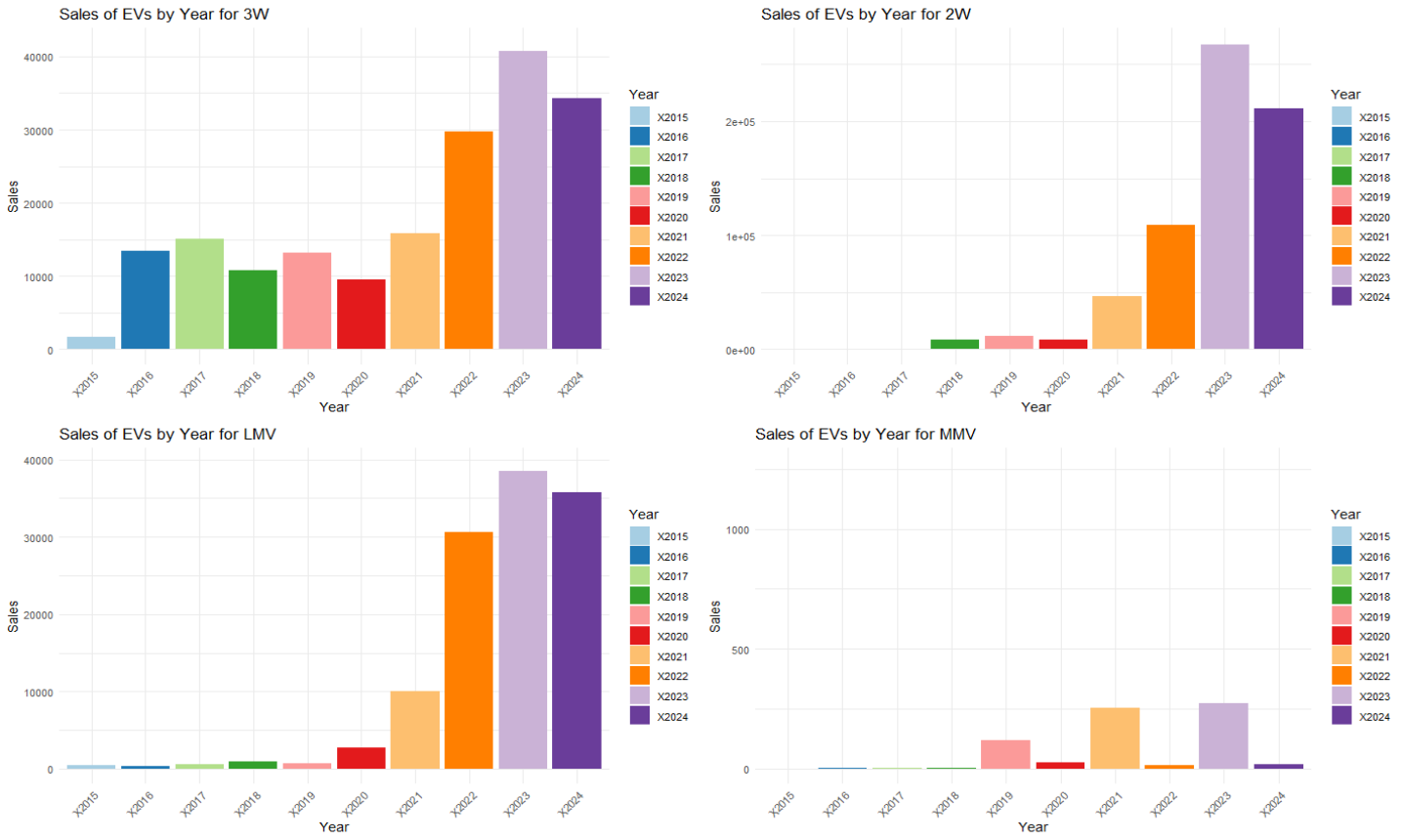
- **X-axis**: Indian states.
- **Y-axis**: Number of operational PCS.

Cluster Information:

1. **Cluster 1 (Red)**: Represents most states with a moderate number of PCS.
2. **Cluster 2 (Green)**: Includes Uttar Pradesh and Gujarat, which have the highest number of PCS, with Uttar Pradesh having the most (~3200 PCS).
3. **Cluster 3 (Blue)**: Contains states with the fewest PCS, such as Lakshadweep, Sikkim, and Arunachal Pradesh.
4. **Uttar Pradesh is a clear outlier, having the most operational PCS.**

5. The clustering effectively groups states into high, moderate, and low PCS categories.

Sales



1. 3W (Three-Wheelers):

- Sales remained low from 2015 to 2020.
- There was a significant increase starting from 2021, peaking in 2023, with a slight decrease in 2024.

2. 2W (Two-Wheelers):

- Sales were low until 2020, after which a large increase occurred, especially in 2022 and 2023, with the highest numbers in 2023.
- A slight decline is seen in 2024.

3. LMV (Light Motor Vehicles):

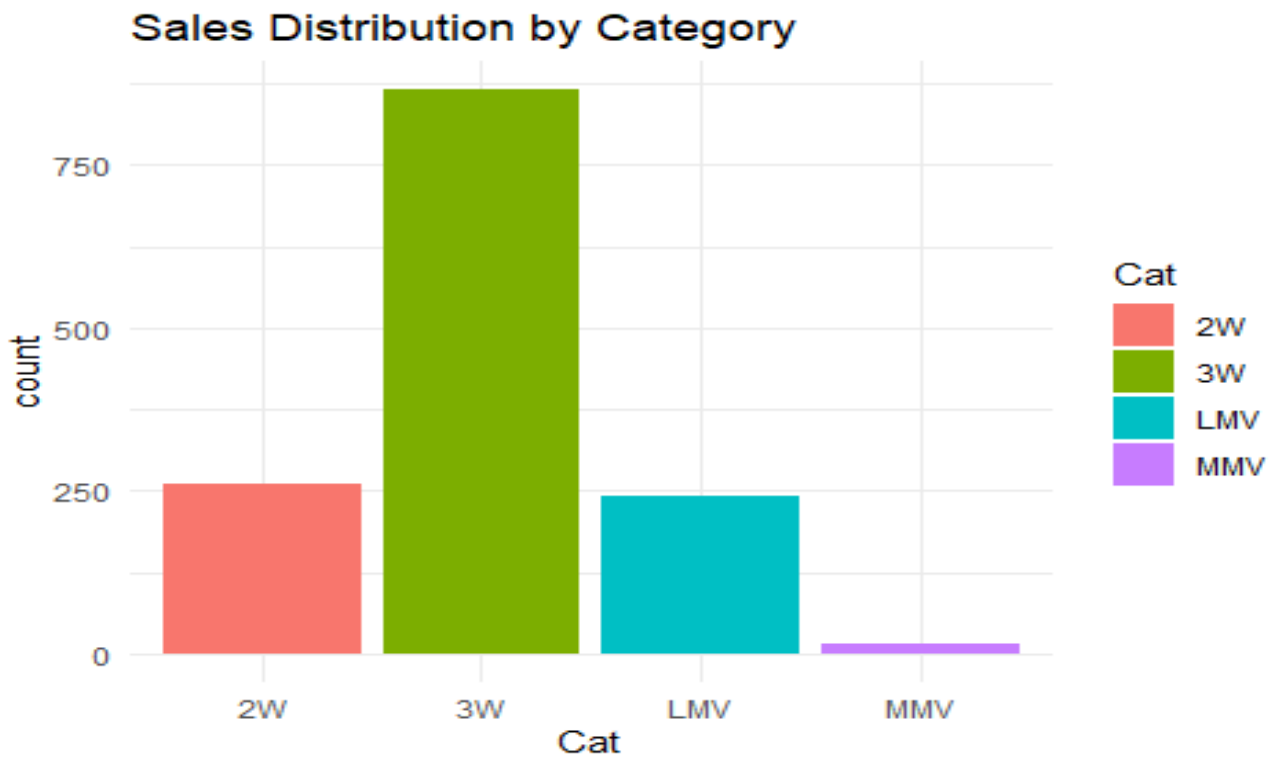
- Similar to 3W, sales were low until 2021, followed by a sharp rise in 2022 and 2023.
- Sales remained high in 2024, with minimal decline from the 2023 peak.

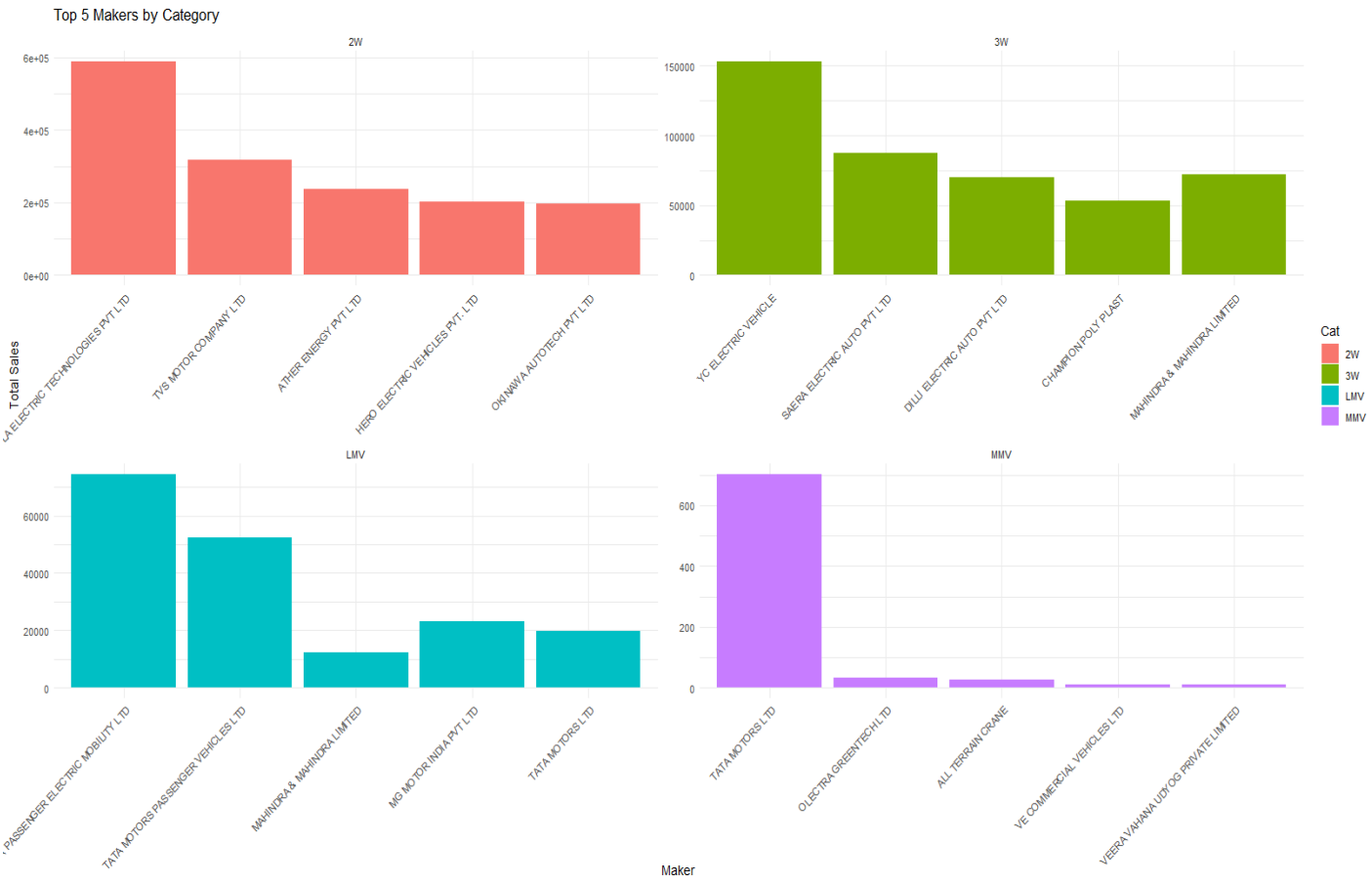
4. MMV (Medium Motor Vehicles):

- Sales remained consistently low across all years, with a minor increase in 2022 and 2023.
- Overall, sales for this category are negligible compared to the others.

Overall Trend:

- **2W and 3W categories** have seen a dramatic increase in sales from 2021 onwards, especially in 2022 and 2023.
- **LMV** sales also grew significantly after 2021, while **MMV** sales have remained very low throughout.





2W (Two-Wheeler)

- **Leading Manufacturer**: YE TECHNOLOGIES PVT LTD with approximately 600,000 units sold.
- **Statistical Outlier**: This manufacturer significantly outperforms the others, with a sales figure that is 50% higher than the second-leading manufacturer.
- **Range**: The range of sales spans from ~200,000 to ~600,000 units, indicating a wide disparity among the top manufacturers.

3W (Three-Wheeler)

- **Leading Manufacturer**: YC ELECTRIC VEHICLE with around 150,000 units sold.
- **Top Performers**: The top three manufacturers (YC ELECTRIC VEHICLE, SAERA ELECTRIC AUTO PVT LTD, and DILLI ELECTRIC AUTO PVT LTD) dominate, covering a majority of the sales.
- **Range**: Sales figures range from ~50,000 to ~150,000 units, with the leading manufacturer selling three times more than the lowest in the top five.

LMV (Light Motor Vehicle)

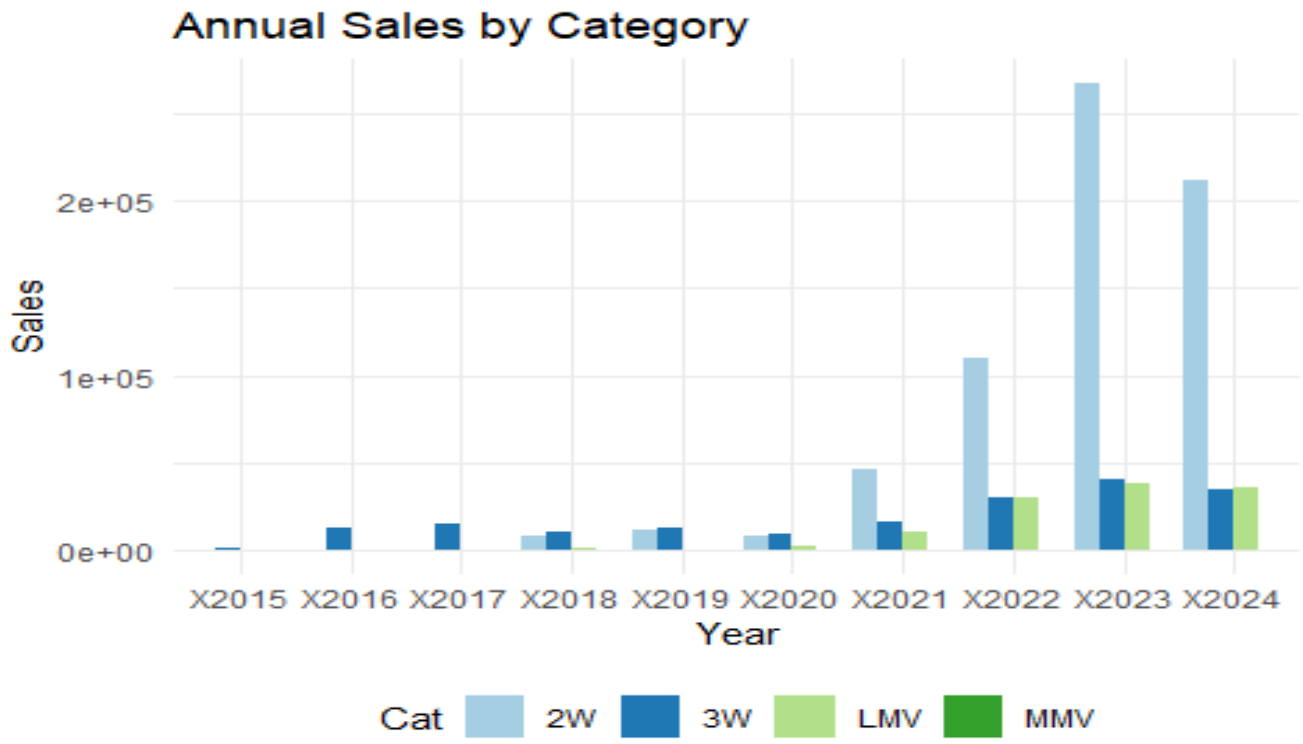
- **Dominant Player**: TATA MOTORS LTD with ~60,000 units sold.
- **Market Concentration**: TATA MOTORS LTD and TATA MOTORS PASSENGER VEHICLES LTD together dominate more than 70% of this category.
- **Range**: Sales figures range from ~10,000 to ~60,000 units, showing a moderate spread with two primary players.

MMV (Medium Motor Vehicle)

- **Top Manufacturer**: TATA MOTORS LTD with ~600 units sold.
- **Sparse Market**: This category has significantly lower numbers compared to others, indicating a niche market.
- **Range**: Sales range from ~10 to ~600 units, with TATA MOTORS LTD dominating over 90% of this segment.

Across all categories, certain manufacturers like TATA MOTORS LTD consistently dominate, particularly in the LMV and MMV segments. The two-wheeler segment shows a competitive market with a significant outlier, while the three-wheeler market is moderately competitive with the top few manufacturers leading significantly.

This statistical analysis reveals not only the leaders in each category but also the extent of their dominance and the competitive landscape of each segment.



.

- **2015 to 2020**: Sales are relatively low across all categories, with no significant spikes.
- **2021**: There is a noticeable increase in sales for the Cat category (light blue), reaching approximately 50,000. Other categories remain relatively low.
- **2022**: Sales for the Cat category (light blue) see a significant spike, reaching approximately 200,000. The 2W (blue) and LMV (light green) categories also show an increase, reaching around 50,000 each.
- **2023**: The Cat category (light blue) maintains high sales, slightly below 200,000. The 2W (blue) category also shows high sales, around 100,000. The LMV (light green) category remains steady at around 50,000.
- **2024**: Sales for the Cat category (light blue) decrease but remain significant, around 100,000. The 2W (blue) and LMV (light green) categories show consistent sales, around 50,000 each.

Statistical Summary

- **Cat Category**:
  - Peak Sales: ~200,000 units in 2022
  - Notable Increase: Starting from 2021, peaking in 2022
- **2W Category**:

    - Peak Sales: ~100,000 units in 2023
    - Steady Increase: Notable from 2022 onwards
- **3W Category**:
    - Consistently Low: No significant spikes observed
- **LMV Category**:
    - Peak Sales: ~50,000 units in 2022 and 2023
    - Steady Performance: Consistent sales around 50,000 units from 2022 onwards
- **MMV Category**:
    - Consistently Low: No significant spikes observed

The chart indicates a significant increase in sales for the Cat category starting from 2021, peaking in 2022 and 2023, and then slightly decreasing in 2024. The 2W and LMV categories also show notable increases in sales during the same period. This trend suggests a growing market or increased demand for these categories during these years.

Hypotheses for ANOVA Test

- **Null Hypothesis (H_0)**: There is no significant difference in mean sales among the different categories.
- **Alternative Hypothesis (H_a)**: At least one category has a significantly different mean sales compared to the others.

```
ANOVA summary
##                Df    Sum Sq   Mean Sq F value   Pr(>F)
## Cat             3 1.030e+09 343421962   18.85 3.41e-12 ***
## Residuals   13856 2.525e+11  18222063
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

## Reject the null hypothesis. There is a significant difference in mean
sales among the different categories.
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Sales ~ Cat, data = ev_data_long)
##
## $Cat
##               diff        lwr          upr     p adj
## 3W-2W   -649.99572  -894.9041 -405.0873069 0.0000000
## LMV-2W  -788.06743 -1097.5791 -478.5557762 0.0000000
## MMV-2W  -868.62231 -1736.7800   -0.4645979 0.0498178
## LMV-3W  -138.07171  -390.2596  114.1161645 0.4951044
```

```
## MMV-3W  -218.62659 -1068.0376  630.7844255 0.9115767
## MMV-LMV  -80.55489  -950.7942  789.6843970 0.9952689
```

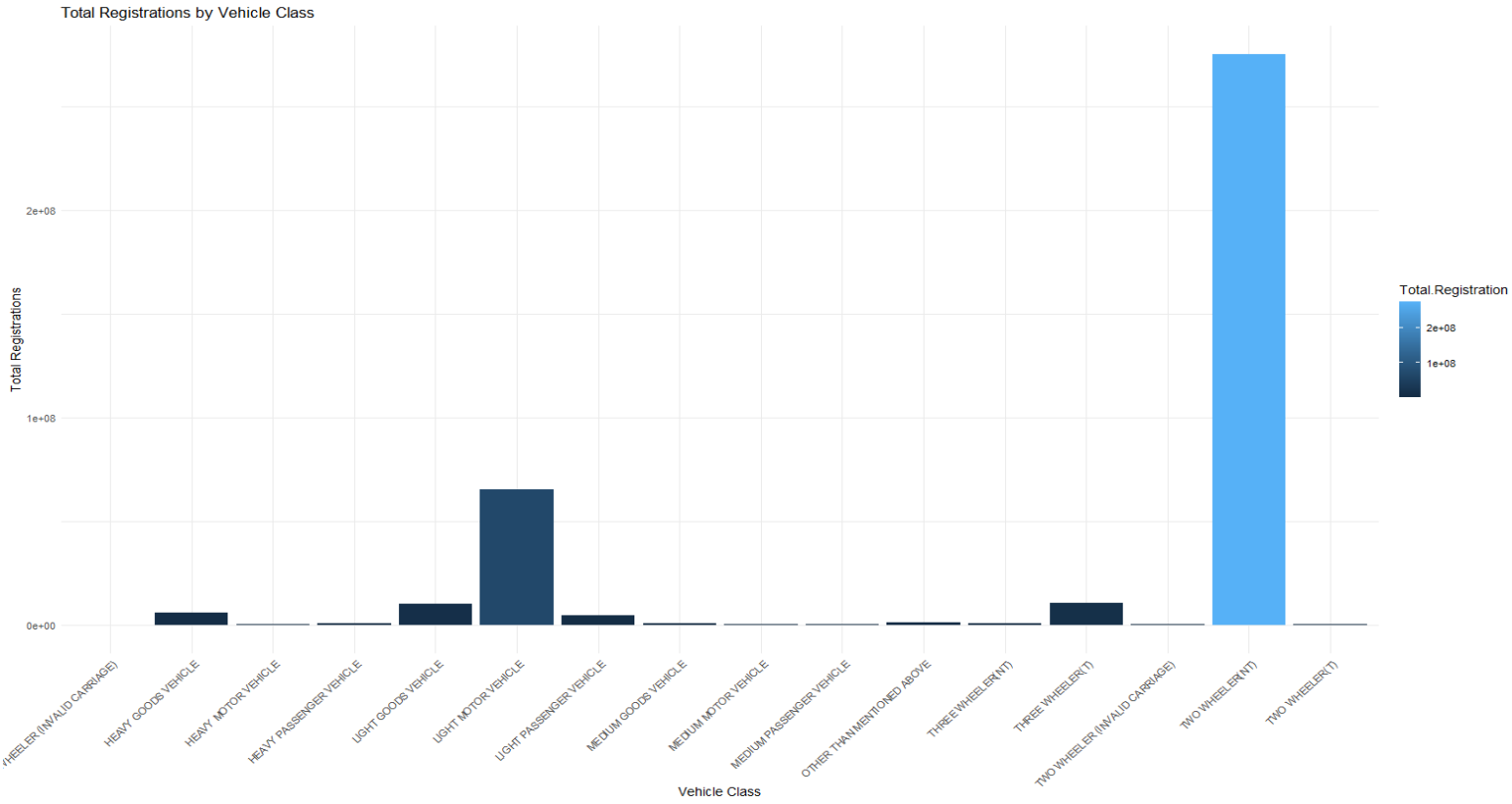
Post Hoc Test Results Summary

- **Significant Differences**:
  - **3W vs. 2W**: Average sales for 3W are significantly lower than for 2W ($p < 0.001$).
  - **LMV vs. 2W**: Average sales for LMV are significantly lower than for 2W ($p < 0.001$).
  - **MMV vs. 2W**: Average sales for MMV are significantly lower than for 2W ($p = 0.0498$).
- **No Significant Differences**:
  - **LMV vs. 3W**: No significant difference ($p = 0.4951$).
  - **MMV vs. 3W**: No significant difference ($p = 0.9116$).
  - **MMV vs. LMV**: No significant difference ($p = 0.9953$).

Conclusion

The 2W category shows significantly higher average sales compared to 3W, LMV, and MMV categories, indicating that category type affects sales performance.

```
## # A tibble: 13,860 × 4
## # Groups:   Year [10]
##    Cat   Maker                                Year  Sales
##    <chr><chr><dbl><int>
##  1 2W    OLA ELECTRIC TECHNOLOGIES PVT LTD  2023 267355
##  2 2W    OLA ELECTRIC TECHNOLOGIES PVT LTD  2024 211273
##  3 2W    TVS MOTOR COMPANY LTD              2023 166580
##  4 2W    OLA ELECTRIC TECHNOLOGIES PVT LTD  2022 109398
##  5 2W    ATHER ENERGY PVT LTD               2023 104735
##  6 2W    OKINAWA AUTOTECH PVT LTD           2022 103620
##  7 2W    TVS MOTOR COMPANY LTD              2024  98516
##  8 2W    HERO ELECTRIC VEHICLES PVT. LTD    2022  97828
##  9 2W    AMPERE VEHICLES PRIVATE LIMITED    2022  79857
## 10 2W    BAJAJ AUTO LTD                     2024  73698
```




```
## # ℹ 13,850 more rows
```

“TWO WHEELER (NT)” leads vehicle registrations with over 200 million units, followed by “MEDIUM GOODS VEHICLE” with around 100 million units. Other classes, like “HEAVY GOODS VEHICLE” and “LIGHT GOODS VEHICLE,” show moderate figures, while categories such as “INVALID CARRIAGE” and “THREE WHEELER (INVALID CARRIAGE)” have negligible counts. This

distribution highlights the dominance of two-wheelers and medium goods vehicles in the market.

Conclusion

1. **Public Charging Stations (PCS)**:
    - **Maharashtra** leads in PCS, followed by **Delhi**, **Karnataka**, and **Kerala**.
    - **Tamil Nadu**, **Uttar Pradesh**, and **Telangana** have moderate counts, while **Arunachal Pradesh**, **Nagaland**, and **Lakshadweep** show very low numbers.
2. **Regional Insights**:
    - The **North Region** has robust infrastructure; the **South and West** are moderate; the **East and Central** have gaps; and the **North-East** requires urgent development.
3. **Electric Vehicle Sales Trends**:
    - Significant increases in sales for **Two-Wheelers (2W)** and **Three-Wheelers (3W)** began in 2021, peaking in 2023. **Light Motor Vehicles (LMV)** also rose sharply, while **Medium Motor Vehicles (MMV)** remain low.
4. **Manufacturer Performance**:
    - **2W**: YE TECHNOLOGIES PVT LTD leads with ~600,000 units.
    - **3W**: YC ELECTRIC VEHICLE tops with ~150,000 units.
    - **LMV**: TATA MOTORS LTD dominates with ~60,000 units.
    - **MMV**: TATA MOTORS LTD leads with ~600 units.
5. **Vehicle Registrations**:
    - **TWO WHEELER (NT)** has over 200 million registrations, indicating strong market preference.

The analysis reveals regional disparities in electric vehicle infrastructure, with Maharashtra at the forefront in PCS. A growing market for 2W and LMV is evident, emphasizing the need for targeted development in underrepresented areas to enhance electric vehicle adoption across India.

Name: Lakshmithra K

Code:

https://github.com/Mithra09/EV-Bunks/blob/main/EV_Charging_bunks.ipynb

OVERVIEW OF CHARGING INFRASTRUCTURE

Charging infrastructure by region plays a crucial role in determining the adoption rates of electric vehicles (EVs). Charging stations are to EVs what gas stations are to traditional vehicles, so the availability and accessibility of charging infrastructure significantly impact the decision-making process of potential EV buyers. Here's how we analysed and what concluded from charging infrastructure density by region:

Key Conclusions from Charging Infrastructure Density by Region

1. Higher Charging Infrastructure Density Correlates with Higher EV Adoption

Regions with dense charging networks, such as Delhi, Maharashtra, and Karnataka (Bangalore), tend to show higher EV sales and adoption rates. In areas with more charging stations, potential EV buyers feel confident that they won't face range anxiety (fear of running out of battery without access to a charger). This boosts consumer confidence in switching to EVs. Focus marketing campaigns and sales efforts in these high-density regions to capitalize on an already favorable infrastructure.

2. Emerging Markets Need Improved Infrastructure for EV Growth

Regions like Kerala, Assam, and Odisha are showing potential for EV growth, but the lower density of charging stations may be a barrier to wider adoption.Although these states have government incentives or rising consumer interest in EVs, the limited availability of charging stations restricts their ability to scale adoption rates. Many consumers may delay or avoid purchasing EVs until they are confident that charging will be convenient.Collaborate with local governments and private sector companies to expand charging networks in these emerging regions to stimulate EV sales.

3. Urban Areas Dominate EV Charging Infrastructure

Urban centers like Mumbai, Delhi, and Bangalorehave the highest density of charging stations compared to rural areas. This concentration supports high EV adoption in cities, but rural EV adoption remains low.The infrastructure gap between urban and rural areas is significant. In rural regions, the lack of fast and convenient charging stations discourages EV purchases, as long-distance travel requires reliable access to chargers.To expand EV adoption beyond urban markets, there is a need to focus onrural electrification efforts, specifically on building highway

charging stations and decentralized rural networks. This will open up new customer segments for EV manufacturers.

4. Government Policy and Incentives Correlate with Infrastructure Development

States with stronger government backing, such as Delhi (EV policy)and Karnataka, tend to have more robust charging infrastructure. These regions have seen faster deployment of charging stations due to both public investment and favorable policies that encourage private investment.Government policies that provide financial incentives, subsidies, and simplified procedures for setting up charging infrastructure lead to faster expansion. States without such support lag behind in terms of both infrastructure and EV adoption.Advocate for stronger state-level policies in regions where infrastructure is lacking, and partner with the government to promote private sector involvement in building charging stations.

Conclusion

The density of EV charging infrastructure is a pivotal factor in influencing EV adoption across different regions in India. Regions with high infrastructure density (urban areas, regions with government support) see significantly higher EV sales. Meanwhile, improving infrastructure in emerging and rural areas presents an opportunity for market expansion.

The strategic development of charging stations—particularly fast chargers—across urban and rural markets can serve as a catalyst for the next wave of EV adoption in India. Focusing on infrastructure expansion will not only remove the "range anxiety" barrier but also unlock new consumer segments.

Data Sources

We have gathered dataset on the electric vehicle charging station from Kaggle and other sites:

Packages/ Tools used:

1. Numpy: To calculate various calculations related to arrays.

2. Pandas: To read or load the datasets.

Data Pre-processing (Steps and Libraries used)

Preprocessing Steps

1. Handling Inconsistent Data Entries

In the dataset we are using steps where the state column is made consistent by converting values to lowercase and stripping any leading or trailing spaces.A state-mapping dictionary is created to

standardize variations in state names (e.g., handling cases like "AndhraPradesh" or spelling errors).

- Text Normalization:

  Lowercasing: Converting all text to lowercase helps eliminate case sensitivity, making comparisons more straightforward.

  Stripping Whitespace: Removing leading or trailing spaces ensures that no extra characters interfere with comparisons.

- State Mapping Dictionary:

  This dictionary maps variations of state names to a standard format. This is useful for correcting common misspellings and variations (e.g., "AndhraPradesh" vs. "Andhra Pradesh").

  Example:

  ```
  state_mapping = {
  "Andhra Pradesh": "Andhra Pradesh",
  "AndhraPradesh": "Andhra Pradesh",
  "A.P.": "Andhra Pradesh",
  # Add more mappings as needed
  }
  ```

2. Data Exploration and Cleaning

   In this phase involves inspecting the dataset for issues such as missing values, incorrect entries, or anomalies.

   - Identifying Missing Data:

     The dataset is checked for missing values using methods like .isnull().sum() to get a count of missing entries for each column.

   - Correcting Incorrect Entries:

     Any misclassified entries (e.g., city names appearing in the state column) are addressed by correcting them either manually or through mapping techniques.

   - Handling KeyErrors:

     A KeyError might occur if the code references a column that doesn't exist in the DataFrame, such as "population." This suggests that either the column was never

there or it was dropped/renamed during earlier processing steps. It highlights the need for careful tracking of column names throughout preprocessing.

Libraries Used

The libraries used

1. Pandas:

   Role: The primary library for data manipulation and analysis in Python. Itprovides DataFrame objects for storing and manipulating structured data.

   Common Methods: pd.read_csv(), .isnull(), .fillna(), .dropna(), .replace(), .map(), and .apply().

```
import pandas as pd

data = pd.read_csv('ev_charging_stations.csv')
```

2. NumPy:

   Role: A library for numerical computing in Python. It's often used in conjunction with Pandas for efficient calculations.

   Common Methods: Array manipulations, mathematical functions, and logical operations.

```
import numpy as np
```

3. Regular Expressions (re):

   Role: A module for working with regular expressions, which is useful for cleaning and normalizing text data.

   Common Functions: re.sub(), re.findall() to identify patterns in text.

```
import re
```

4. Matplotlib / Seaborn:

   Role: Libraries for data visualization. They are often used for plotting data distributions, missing value heatmaps, or exploratory data analysis.

   Common Functions: plt.plot(), sns.heatmap().

```
import matplotlib.pyplot as plt

import seaborn as sns
```

5. Scikit-Learn:

   Role: A library for machine learning that includes preprocessing functions, such as scaling and encoding categorical variables.

   Common Methods: StandardScaler, OneHotEncoder.

from sklearn.preprocessing import StandardScaler, OneHotEncoder

Conclusion

Data preprocessing is a crucial step in preparing your EV charging stations dataset for analysis and modelling. By addressing inconsistencies, handling missing values, and ensuring the correct data structure, you can enhance the quality of your dataset, leading to more accurate insights and model performance. The use of libraries like Pandas and NumPy streamlines this process and provides powerful tools for data manipulation.

Analysis and Approaches used for Segmentation

K-MEANS CLUSTERING

K-Means clustering is a popular unsupervised machine learning algorithm used to partition a dataset into distinct groups (clusters) based on feature similarity. Explanation of how K-Means works, its key components, and applications:

1. Unsupervised Learning: K-Means is an unsupervised learning algorithm, meaning it does not require labeled data. Instead, it identifies patterns or groupings in the data based on the inherent structure.

2. Clusters: A cluster is a collection of data points that are more similar to each other than to those in other clusters. The algorithm seeks to minimize the distance between data points within the same cluster.

3. Centroids:Each cluster is represented by a centroid, which is the mean position of all the points in the cluster. The centroid acts as the "center" of the cluster.

K-Means clustering can significantly enhance your analysis of the dataset containing information about electric vehicle (EV) charging stations across cities. Here’s how K-Means can be specifically useful for your dataset:

How we are using it for our dataset:

1. Identification of Patterns and Trends

Segmentation of Cities: K-Means allows you to group cities based on the number of EV charging stations. By clustering cities, you can identify patterns, such as which cities are underserved in terms of charging infrastructure or which have a high density of stations.

Understanding Distribution: You can visualize how charging stations are distributed across different clusters. This helps identify geographical areas that might require more attention or investment.

2. Market Opportunity Analysis

Targeting Under-Served Areas: By clustering cities, you can easily spot those with fewer charging stations. These clusters represent potential markets for expanding charging infrastructure, helping you prioritize new installations where they are most needed.

Resource Allocation: Understanding which clusters have high demand for charging stations allows businesses to allocate resources more effectively. For instance, more marketing and installation efforts can be directed toward cities that show high potential for EV adoption but currently lack sufficient charging options.

3. Strategic Planning and Decision-Making

Informed Decision-Making: The insights gained from clustering provide a data-driven basis for strategic planning. Instead of relying solely on intuition, you can make informed decisions about where to invest in new charging stations or marketing campaigns.

Partnership Opportunities: Identifying clusters of cities with similar needs can also reveal potential partnership opportunities with local businesses or municipalities interested in promoting EV adoption.

4. Enhanced Visualization and Communication

Clear Visualization: Clustering allows for effective visualization of the data. By plotting the clusters, you can create visual representations that make it easier to communicate findings to stakeholders, investors, or partners, demonstrating areas of opportunity or need.

Stakeholder Engagement: Clear and compelling visualizations can facilitate discussions with stakeholders about market potential and the rationale for investment in certain areas.

Conclusion

K-Means clustering helps you make sense of your dataset by revealing underlying patterns and opportunities in the distribution of EV charging stations across cities. This analysis not only aids in identifying market opportunities but also supports strategic planning and effective resource allocation. Overall, K-Means clustering provides a structured approach to analyzing your dataset, facilitating informed decisions that can drive business growth in the EV infrastructure sector.

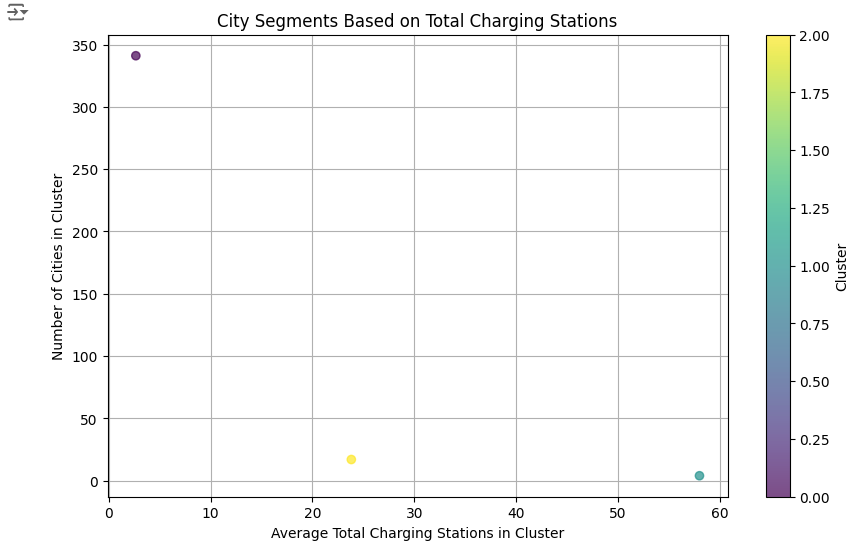
Clustering

Clustering is one of the most common exploratory data analysis techniques used to get 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 each cluster are as similar as possible according to a similarity measure such as Euclidean-based distance or correlation-based distance. The decision of which similarity measure to use is application-specific.

```
# Count the number of cities in each cluster
cluster_city_counts = city_segments.groupby('cluster').agg(
    total_stations_mean=('total_stations', 'mean'),
    city_count=('city', 'count')
).reset_index()

# Plotting the total charging stations mean vs. number of cities in each cluster
plt.figure(figsize=(10, 6))
plt.scatter(cluster_city_counts['total_stations_mean'], cluster_city_counts['city_count'],
            c=cluster_city_counts['cluster'], cmap='viridis', alpha=0.7)

plt.title('City Segments Based on Total Charging Stations')
plt.xlabel('Average Total Charging Stations in Cluster')
plt.ylabel('Number of Cities in Cluster')
plt.colorbar(label='Cluster')
plt.grid(True)
plt.show()
```



Selection of target segment

1. Data Exploration

The first step in selecting a target segment for the electric vehicle (EV) business involves a comprehensive exploration of the electric charging stations dataset. This entails examining key attributes such as the locationof charging stations, their types (e.g., Level 1, Level 2, DC fast chargers), and usage statistics (if available). Understanding the distribution and characteristics of charging stations can provide insights into areas with a strong demand for EV infrastructure and highlight opportunities for expansion.

2. Market Demand Analysis

Analyzing market demand is crucial to identify high-potential areas for EV adoption. By examining the dataset, one can identify regions with a high density of vehicle ownership but limited access to charging stations. This analysis can reveal urban centers where the adoption of EVs is rising, often driven by increasing environmental awareness and supportive government policies. Additionally, comparing charging station availability in suburban versus urban areas can help pinpoint regions where there is significant potential for growth, particularly in suburbs that may lack sufficient charging infrastructure.

3. User Demographics and Preferences

If the dataset provides insights into user demographics, it becomes imperative to analyze these aspects to understand the preferences and characteristics of potential EV customers. Demographic factors such as age, income level, and lifestyle choices can indicate which segments are more inclined to adopt electric vehicles. For instance, targeting younger, environmentally conscious consumers in urban settings may yield higher adoption rates compared to older demographics.

4. Government Incentives and Support

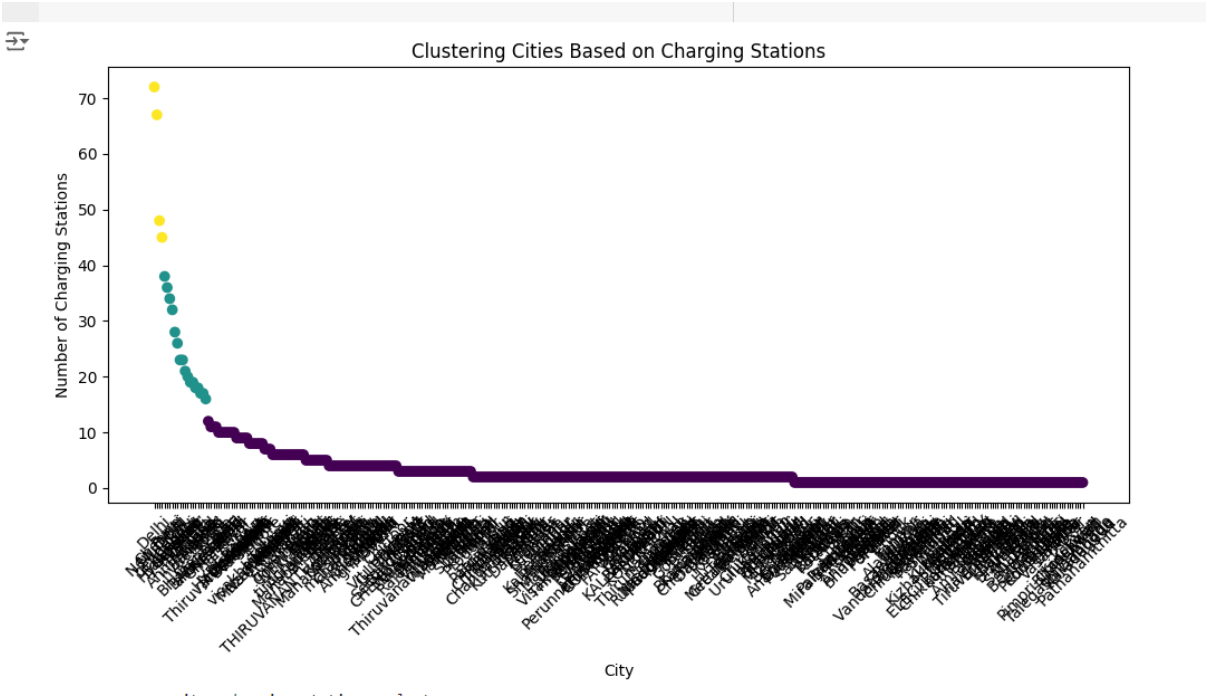
A critical factor influencing EV adoption is the presence of government incentives and supportive policies. Investigating local government initiatives, such as tax breaks, rebates for EV purchases, or funding for charging infrastructure, can guide segment selection. Areas with strong governmental support for EV adoption are likely to experience higher growth rates in the EV market, making them prime candidates for targeting.

5. Market Readiness Assessment

Assessing the readiness of various segmentsfor EV adoption is essential. This can be achieved by analyzing existing EV ownership rates, consumer sentiment, and overall interest in electric mobility. Identifying regions with a high concentration of EV owners or those demonstrating increasing registrations for electric vehicles can help determine where to focus marketing efforts effectively.

7. Segment Identification and Selection

Based on the comprehensive analysis of the dataset and the factors outlined above, businesses can identify several promising target segments for the EV market:Uban Areas with High Charging Station Demand: These regions often have significant populations and existing infrastructure to support EV growth.Suburban Areas with Limited Charging Options: Suburban communities showing interest in EVs but lacking adequate charging facilities present a unique opportunity for market penetration.College Towns: Engaging with educational institutions can tap into a demographic of environmentally conscious students and faculty members who are more likely to adopt EVs.Government-Supported Regions: Areas where local governments actively promote EV adoption through incentives and policies are likely to see increased consumer interest and market growth.Corporate Fleets: Targeting businesses seeking to transition their vehicle fleets to electric options can provide stable revenue streams while supporting corporate sustainability initiatives.



City

	city	charging_stations	cluster
0	Delhi	72	2
1	New Delhi	67	2
2	Chennai	48	2
3	Mumbai	45	2
4	bengaluru	38	1
..	...	...	...
357	Kosi Kalan	1	0
358	Kotwan	1	0
359	Shivamogga	1	0
360	Karadagodu	1	0
361	Pathanamthitta	1	0

[362 rows x 3 columns]

Conclusion

Selecting a target segment for an EV business using the electric charging stations dataset requires a thorough and methodical approach. By exploring the dataset, analyzing market demand, understanding user demographics, assessing government support, evaluating market readiness, and conducting competitor analysis, businesses can strategically target segments with the highest potential for growth. This data-driven methodology ensures that the EV business is well-positioned to meet the needs of consumers seeking sustainable transportation options, ultimately contributing to the broader transition toward electric mobility.

Profiling and describing potential segments

Profiling and describing potential segments in your EV charging dataset can provide valuable insights for targeted strategies and decision-making.

Potential Segments

1. High Availability Segment: Cities in this segment have a high number of EV charging stations relative to their population or area.

Characteristics: Established EV infrastructure. Likely to have a higher adoption rate of electric vehicles. Potential for expanding services related to EVs (e.g., maintenance, accessories).

Examples: Major metropolitan areas, tech hubs, or cities with strong government incentives for EV adoption.

Marketing Strategies :Promote additional services (e.g., subscription models, loyalty programs).Collaborate with local businesses to enhance EV-related offerings.Leverage the existing infrastructure to attract EV users from nearby regions.

2. Medium Availability Segment: Cities with a moderate number of charging stations. These cities have some infrastructure but may not fully meet the growing demand for EV charging.

Characteristics:Emerging markets for electric vehicles.May have some government initiatives promoting EV adoption. Potential for growth in charging station deployment.

Examples: Suburban areas or smaller cities that are beginning to invest in EV infrastructure.

Marketing Strategies: Educate residents on the benefits of EVs and available charging options. Work with local governments to promote incentives for EV purchases.Plan new installations in areas identified as underserved within the segment.

3. Low Availability Segment: Cities in this segment have a limited number of charging stations, indicating a significant opportunity for growth.

Characteristics: Limited EV infrastructure and low adoption rates of electric vehicles. Potential barriers include lack of awareness, incentives, or charging station availability.

Examples: Rural areas or regions with less focus on environmental initiatives.

Marketing Strategies: Conduct awareness campaigns to educate residents about EV benefits and charging options.Partner with local governments to introduce incentives for charging station installation.Explore opportunities for community-based charging solutions (e.g., charging at workplaces, residential complexes).

4. Emerging Markets Segment:Cities that are beginning to see an increase in EV interest but currently have low infrastructure.

Characteristics: Growing awareness and interest in EVs.Potential for rapid growth as charging stations become available.Young, tech-savvy populations that may adopt EVs quickly. Examples: College towns or regions with initiatives promoting clean energy.

Marketing Strategies:Leverage social media and local influencers to promote EV benefits.Organize events showcasing EV technology and local charging options. Engage with local businesses to develop community charging solutions.

Conclusion

By profiling these segments, you can develop targeted marketing strategies and investment plans tailored to the specific needs of each group. This approach enables you to effectively leverage your dataset for strategic planning, resource allocation, and enhancing the adoption of electric vehicles in various regions.

MARKET SEGMENTS

The most optimal maret segmentation to start the electric vehicle business is as follows:

Urban Areas with High Population Densityare among the most promising segments for the EV business. These areas typically feature a large number of residents and vehicles, resulting in heightened potential for EV adoption. With a strong public interest in sustainability and environmental initiatives, urban populations are increasingly seeking eco-friendly transportation options. Furthermore, the existing infrastructure in cities often supports the convenience of owning an EV, such as access to charging stations and proximity to amenities and public transportation. This environment presents opportunities for partnerships with local governments and businesses to promote EV initiatives, enhancing brand visibility and consumer engagement.

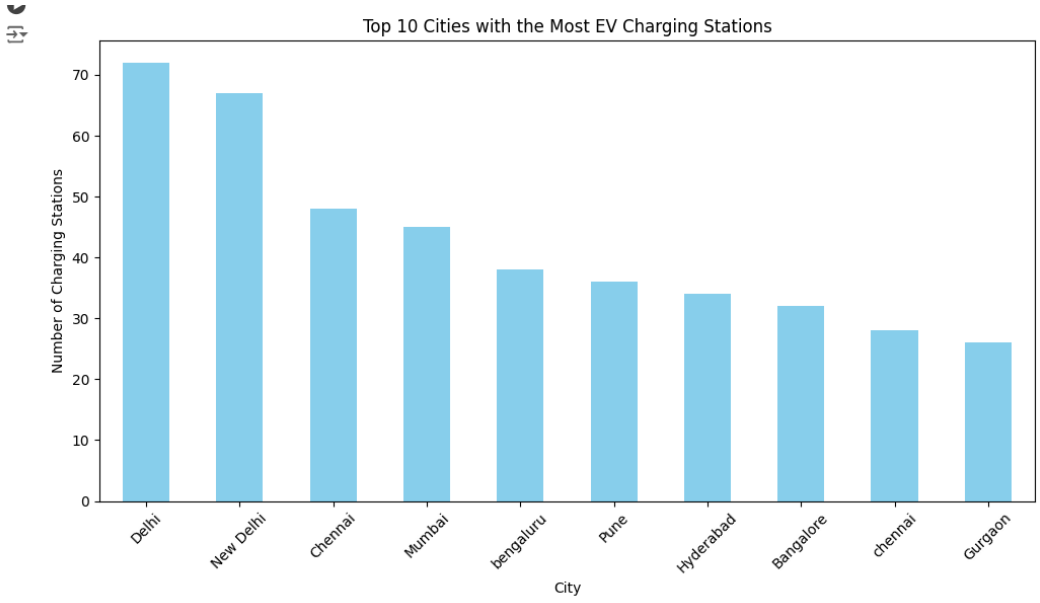
Suburban Areas with Growing Interest in EVs represent another optimal market segment. Many suburban communities are family-oriented, where long commutes to urban centers are common. As awareness of electric vehicles increases among residents, there is a significant opportunity for growth in EV adoption in these regions. Suburban homes often have garages or designated parking spaces, facilitating the practicality of home charging. With limited existing EV

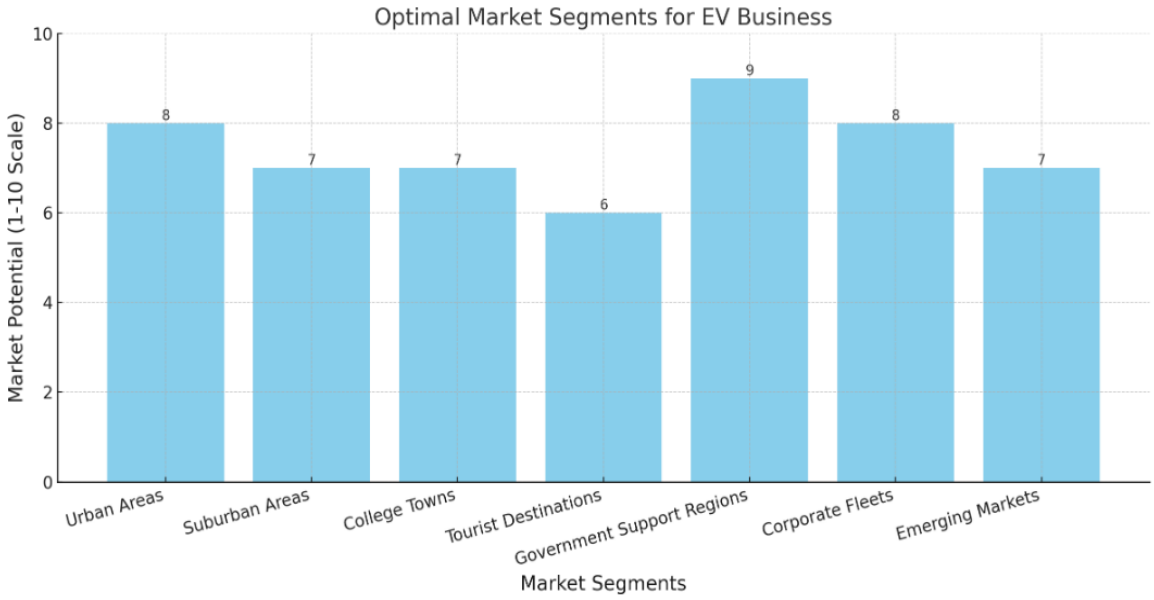
infrastructure, businesses have the chance to establish charging stations that meet the growing demand for electric vehicles, making this segment particularly appealing for market entry

College Towns and University Campuses present a unique opportunity for the EV business. These areas typically have large populations of environmentally conscious students and faculty members who are interested in sustainable practices. Many universities are adopting sustainability initiatives, further encouraging the adoption of electric vehicles. The high foot traffic and community engagement within college towns create an ideal environment for promoting EVs and charging infrastructure. Students, often looking for cost-effective and eco-friendly transportation solutions, may be inclined to choose electric vehicles, leading to potential brand loyalty and long-term customer relationships.

Tourist Destinations and Recreational Areas also represent an attractive market segment for the EV business. Popular attractions draw significant visitor traffic, many of whom may already own electric vehicles or are considering their options. The demand for charging stations in these areas is often high, especially during peak tourist seasons. Establishing charging stations in high-traffic locations allows for strategic partnerships with local businesses and attractions, enhancing brand visibility and sales opportunities. By catering to the needs of EV driversduring their travels, businesses can create a compelling value proposition that aligns with consumers' desires for convenience and accessibility.

Regions with Strong Government Support for EV Adoptionform another optimal segment. Areas that actively promote electric vehicles through incentives, rebates, and supportive policies are likely to see increased adoption rates. Government initiatives aimed at reducing emissions and promoting sustainable transportation can significantly influence consumer behavior. In such markets, EV businesses can benefit from lower total ownership costs due to incentives, creating a more attractive proposition for potential buyers. Additionally, the increasing investment in charging infrastructure driven by public policy further enhances market potential, making these regions ripe for EV business expansion.





This overview provides a comprehensive understanding of the optimal market segments for an EV business, highlighting the unique opportunities and characteristics of each segment.

CONCLUSION

From the above graphs and the calculations on the dataset we can conclude that these cities are the best choices for our business. We can also find the cities that are adapting to the electric vehicles and are developing in the area of electric vehicles. So based on these findings we can choose a optimal location for starting the business.