

Market Segmentation

Segmentation Analysis of the Indian Market for an Electric Vehicle Start-up

by

Sumit kushawah

Prajakta Khairnar



Problem Statement

The task is to analyze the Electric Vehicle (EV) market in India by leveraging multiple real-world datasets related to EV manufacturers, EV sales, public charging stations, and vehicle types across different states and cities.

This report aims to perform detailed market segmentation based on geography (state, region, city) and explore EV adoption trends through the lens of sales volume, vehicle type (2W, 3W, 4W), manufacturing concentration, and charging infrastructure availability.

The goal is to identify high-potential regions and vehicle categories to help stakeholders make informed strategic decisions for market entry, expansion, and investment.

The analysis includes descriptive insights, visual exploration, clustering for segmentation, and predictive modeling of EV sales, all focused around geographic factors.

Fermi Estimation

Objective:

Estimate the potential Electric Vehicle (EV) sales in India over the next year based on current adoption trends, geographic spread, and infrastructure availability.

Assumptions for Estimation:

India's current EV adoption is concentrated in top 10 states, contributing around 70% of total EV sales (based on sales and PCS data). Suppose India's current annual EV sales is approximately 1 million units (industry estimate as of 2024). With increasing charging infrastructure and government incentives, assume a growth rate of 30% year-on-year for EV adoption.

2-wheelers dominate the EV market, contributing about 60% of total EV sales, followed by 3-wheelers (25%) and 4-wheelers (15%). Charging station growth is expected to increase by about 20% annually, facilitating faster adoption in urban and semi-urban areas.

Estimation Calculation:

Current sales: 1 million EVs per year.

Next year sales prediction = $1 \text{ million} \times (1 + 0.30) = 1.3 \text{ million EVs}$.

Breakdown by vehicle type:

2-Wheelers: 60% of 1.3 million $\approx 780,000$ units

3-Wheelers: 25% of 1.3 million $\approx 325,000$ units

4-Wheelers: 15% of 1.3 million $\approx 195,000$ units

Data Collection

Data was extracted from the various websites mentioned below for EV market segmentation.

Link for data extraction:

- [The Electric Vehicle Ecosystem in India: A Look at the Progress So Far](#)
- [Ashwini3535/EV-MARKET-IN-INDIA: To analyze the Electric Vehicle market in India using Segmentation Analysis.](#)
- [Market-Segmentation-for-Electric-Vehicles-in-India/Market_Segmentation.ipynb at main · Marisha18/Market-Segmentation-for-Electric-Vehicles-in-India](#)

Data from those links are extracted by Google play scraper available on libraries package. There are multiple datasets get extracted from those websites in CSV and Excel formats. There are some pdfs also which contains valuable information regarding the EV market. We have extracted data from those pdfs as well.

Raw data generated:

- [Feynn-Labs-Internship-2022/Project 2.1: Market Segmentation of Electric Vehicles in India/Datasets/1_ev_charger_dataset.csv at main · Rohit-Rannavre/Feynn-Labs-Internship-2022](#)
- [Feynn-Labs-Internship-2022/Project 2.1: Market Segmentation of Electric Vehicles in India/Datasets/2_ev_charging_station_dataset.xlsx at main · Rohit-Rannavre/Feynn-Labs-Internship-2022](#)

Data Preprocessing

For this project, data preprocessing was performed on the Electric Vehicle (EV) sales dataset to prepare it for predictive modeling and segmentation analysis. The following steps were taken:

Handling Dates and Time Features:

The dataset contained a Month_Name column representing the month of sale.

A mapping was created to convert month names into corresponding month numbers (e.g., January → 1, February → 2, etc.) to make it numerically understandable by machine learning models.

```
# Convert Month_Name to Month Number
month_map = {'January':1, 'February':2, 'March':3, 'April':4, 'May':5, 'June':6,
             'July':7, 'August':8, 'September':9, 'October':10, 'November':11, 'December':12}
df3['Month'] = df3['Month_Name'].map(month_map)
```

Dropping Irrelevant Columns:

The Date and Month_Name columns were dropped after extracting necessary information, to reduce redundancy and simplify the dataset.

```
# Drop unnecessary columns
df = df3.drop(['Month_Name', 'Date'], axis=1)
```

Encoding Categorical Variables:

Several important categorical columns, such as State, Vehicle_Class, Vehicle_Category, and Vehicle_Type, were encoded using Label Encoding.

This transformation was necessary to convert string labels into numeric format suitable for machine learning algorithms like Random Forest.

```
# Encode categorical columns
le = LabelEncoder()
for col in ['State', 'Vehicle_Class', 'Vehicle_Category', 'Vehicle_Type']:
    df[col] = le.fit_transform(df[col])
```

Feature-Target Separation:

The feature matrix X was created by excluding the target variable (EV_Sales_Quantity).

The target vector y was defined as the EV_Sales_Quantity, which we aimed to predict.

```
# Features and Target
X = df.drop('EV_Sales_Quantity', axis=1)
y = df['EV_Sales_Quantity']
```

Train-Test Split:

The dataset was split into training (80%) and testing (20%) subsets using random state 42 to ensure reproducibility.

```
# ===== 2. Train-Test Split =====
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Standardization (Optional for Clustering):

Standard scaling was also considered while performing clustering tasks like K-Means or PCA, to bring all numeric features to the same scale.

```
[36] # Step 4: Feature Scaling
    scaler = StandardScaler()
    X_scaled = scaler.fit_transform(X)

# Step 5: PCA
pca = PCA(n_components=X.shape[1]) # Keep all components first
X_pca = pca.fit_transform(X_scaled)
df_pca = pd.DataFrame(X_pca, columns=[f'PC{i+1}' for i in range(X_pca.shape[1])])
```

Exploratory Data Analysis (EDA)

The goal of EDA was to understand patterns, trends, and anomalies in the Indian EV market across multiple dimensions like geography, vehicle types, and infrastructure availability. Here's a summary of the insights obtained:

1. Dataset Overview:

Multiple datasets were combined to cover different aspects:

EV Makers by Place (Manufacturers and their locations)

State-wise EV Charger Availability

EV Sales Data (State, Vehicle type, Year, Month)

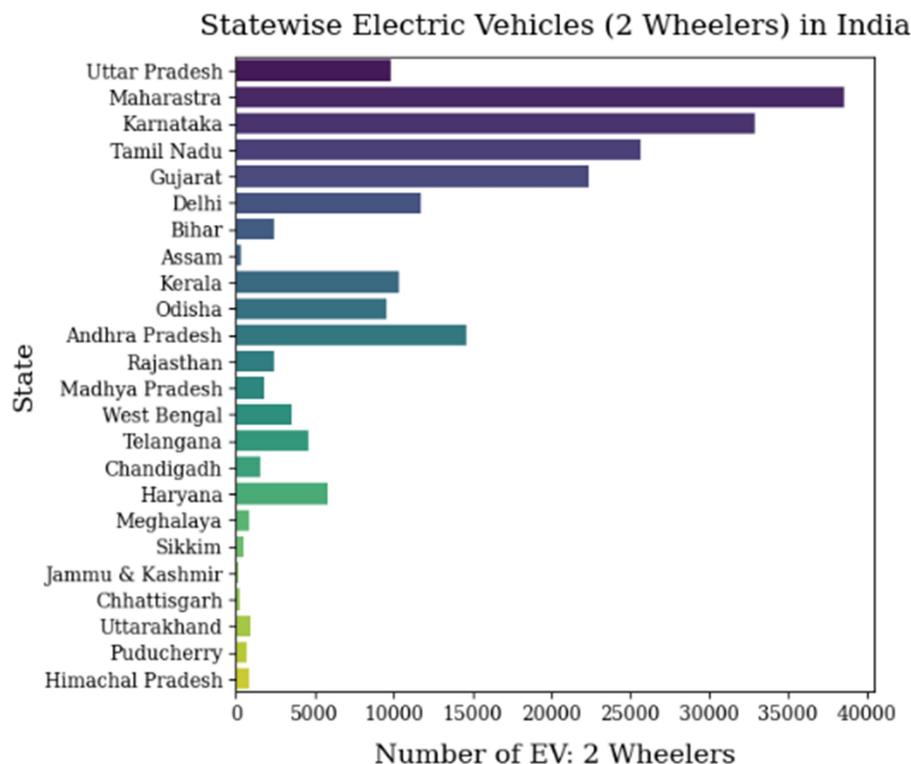
Public Charging Stations by State

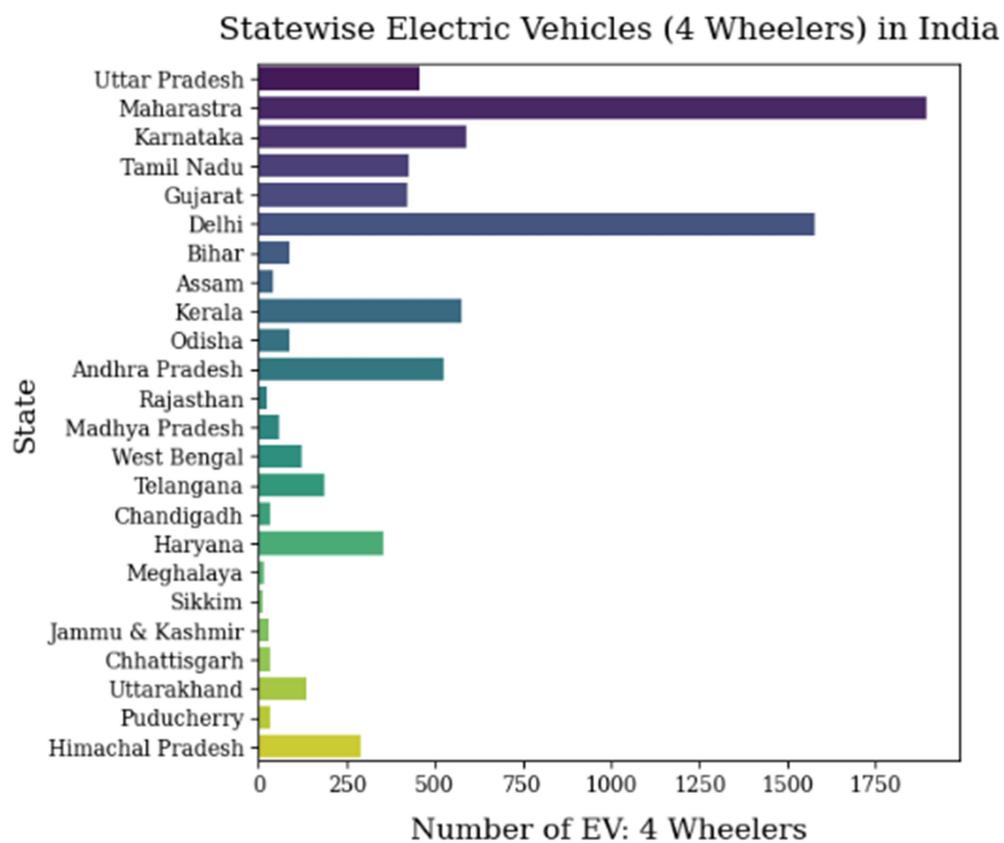
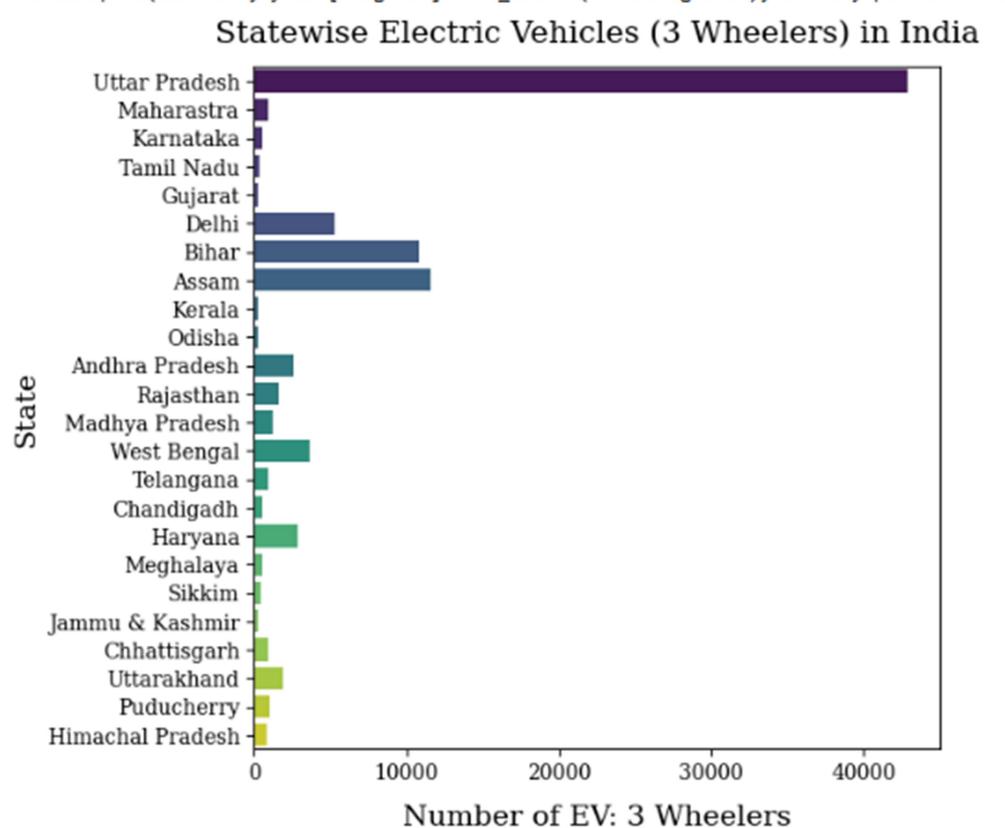
Each dataset was inspected for missing values, data types, and basic statistics to understand its structure.

2. State-wise EV Distribution:

2 Wheelers, 3 Wheelers, and 4 Wheelers were visualized across states.

States like Uttar Pradesh, Maharashtra, and Karnataka showed the highest adoption in various categories.





Insights:

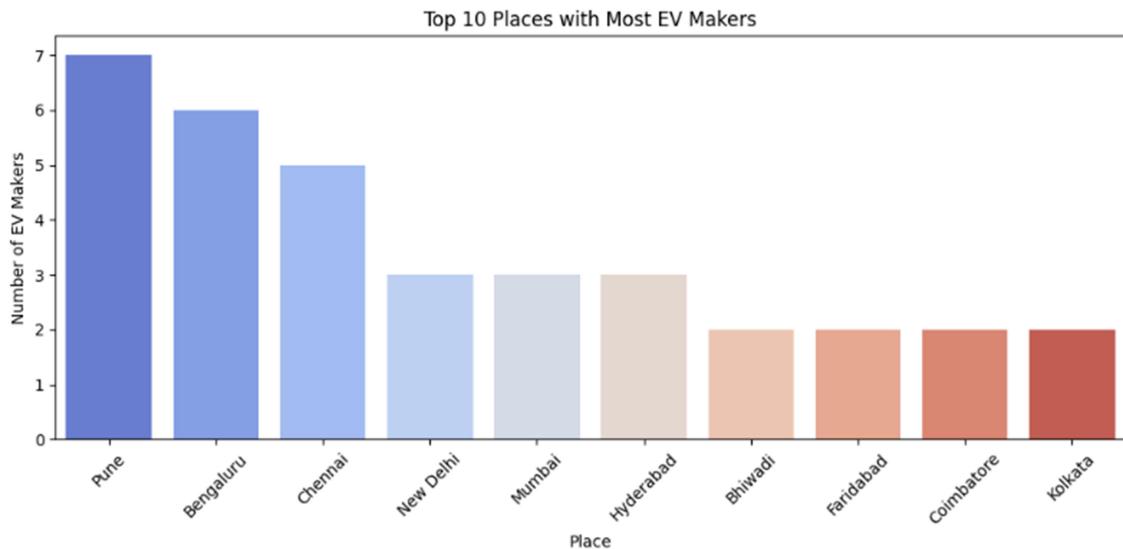
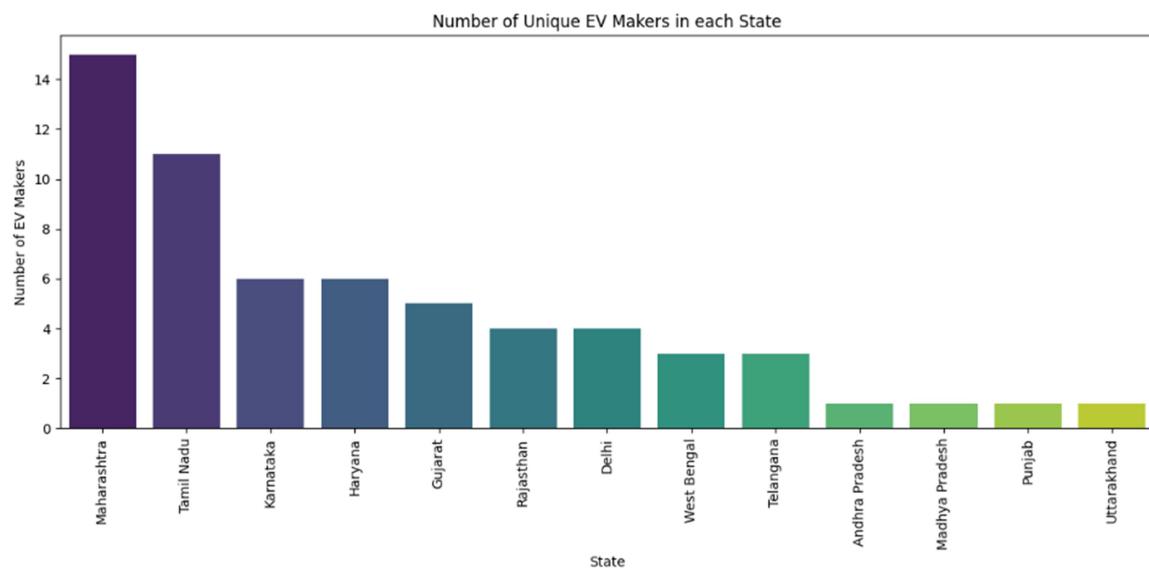
2-wheelers dominate the EV segment in India, especially in populous states.

3-wheelers (electric autos) also have significant penetration, hinting at strong commercial use in urban and semi-urban areas.

3. EV Makers Distribution:

Analysis of the number of unique EV manufacturers per state revealed that Maharashtra, Tamil Nadu, and Karnataka have the highest number of EV manufacturers.

Top cities like Pune, Chennai, and Bengaluru host the maximum number of EV makers.



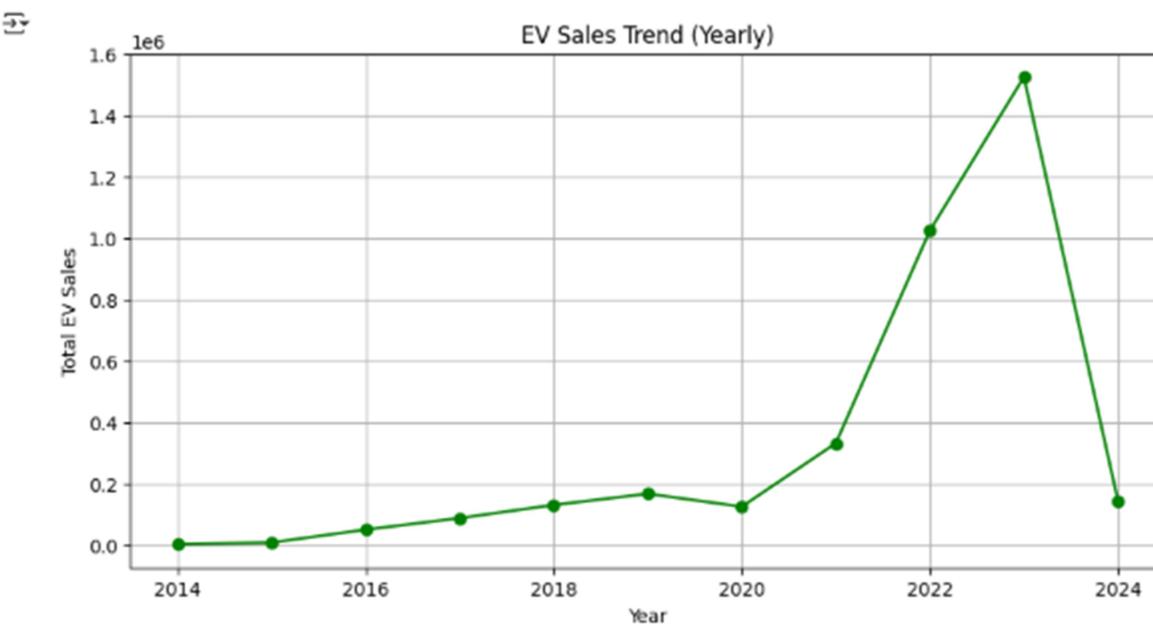
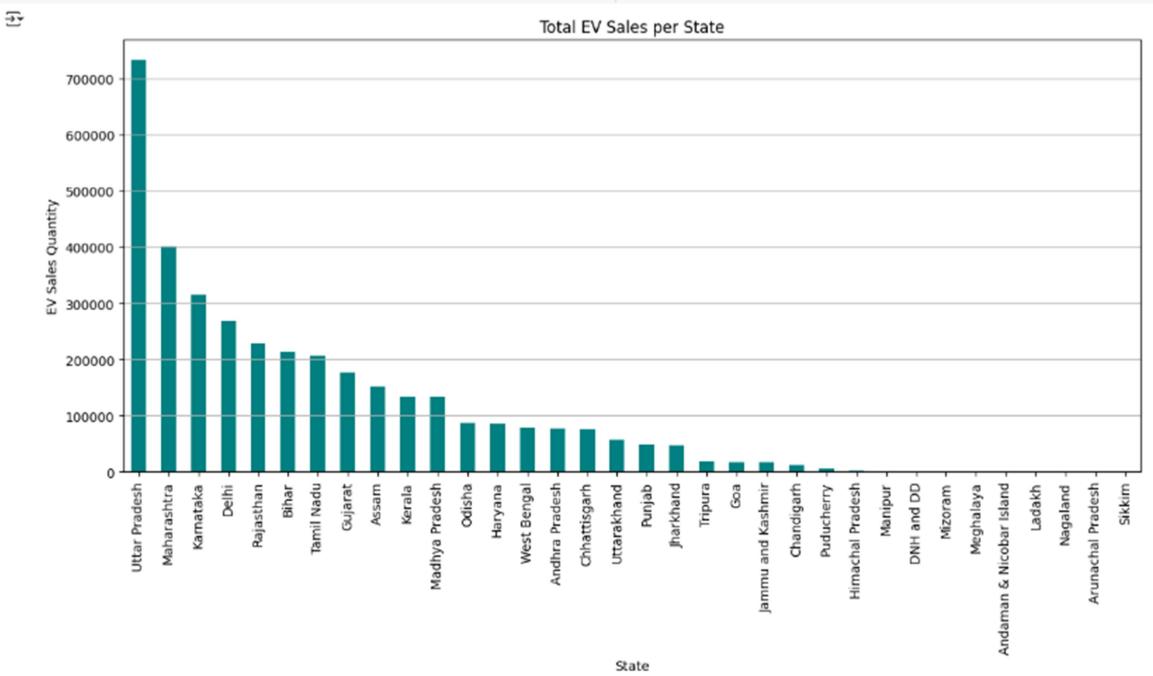
Insights:

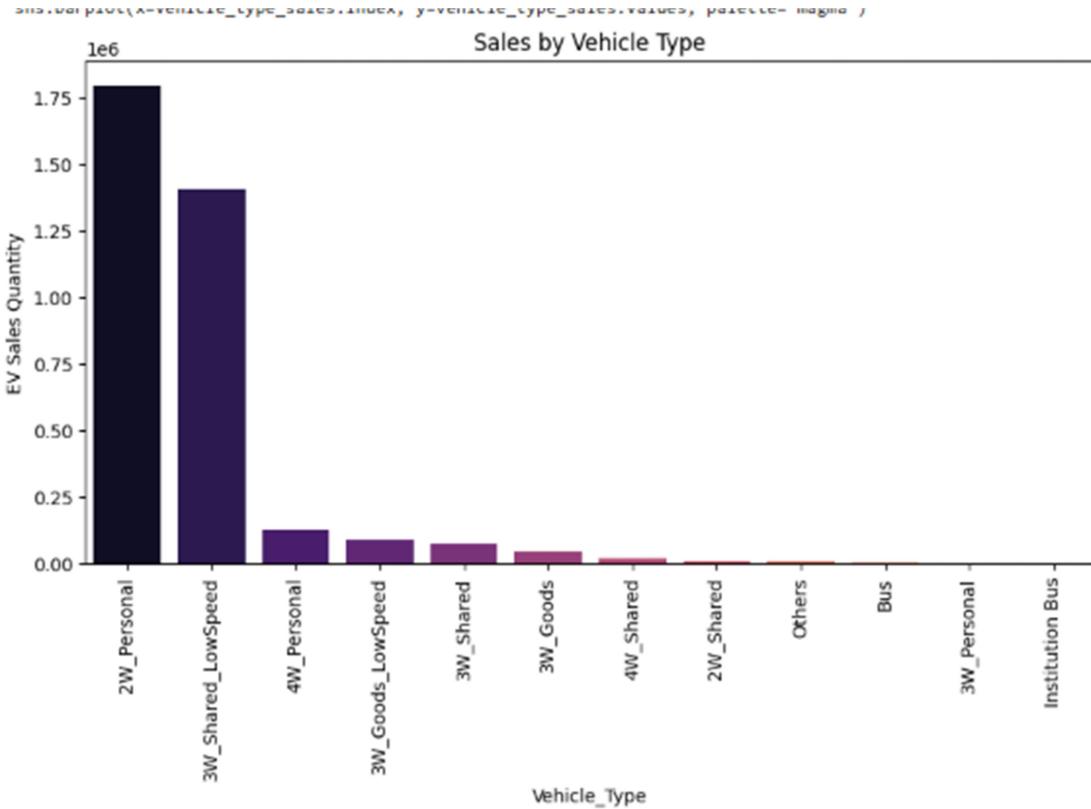
Southern and Western India are leading in manufacturing hubs, likely due to better industrial policies and infrastructure.

4. Sales Analysis:

State-wise total EV sales showed Uttar Pradesh and Maharashtra as the top-performing states.

Sales trend over years indicated a steady and sharp rise post-2019, coinciding with government incentives and greater public awareness.





Insights:

EV adoption is accelerating, with newer states also showing gradual pickup.

5. Vehicle Type Popularity:

2-wheelers are the most sold EVs, followed by 3-wheelers and 4-wheelers.

Commercial vehicles (autos, loaders) constitute a major chunk of 3-wheeler EVs, reflecting cost-driven usage.

Insights:

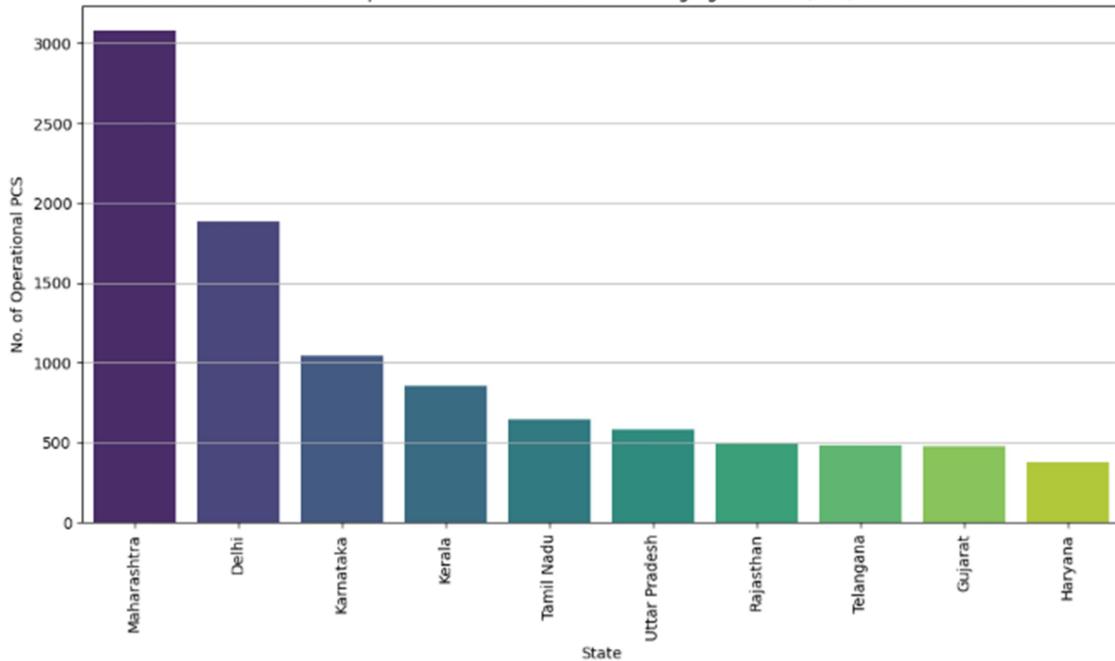
Entry-level and mid-level two-wheelers are the primary choice for private users.

3-wheeler sales suggest a strong commercial demand in urban transport and last-mile delivery.

6. Charging Infrastructure Analysis:

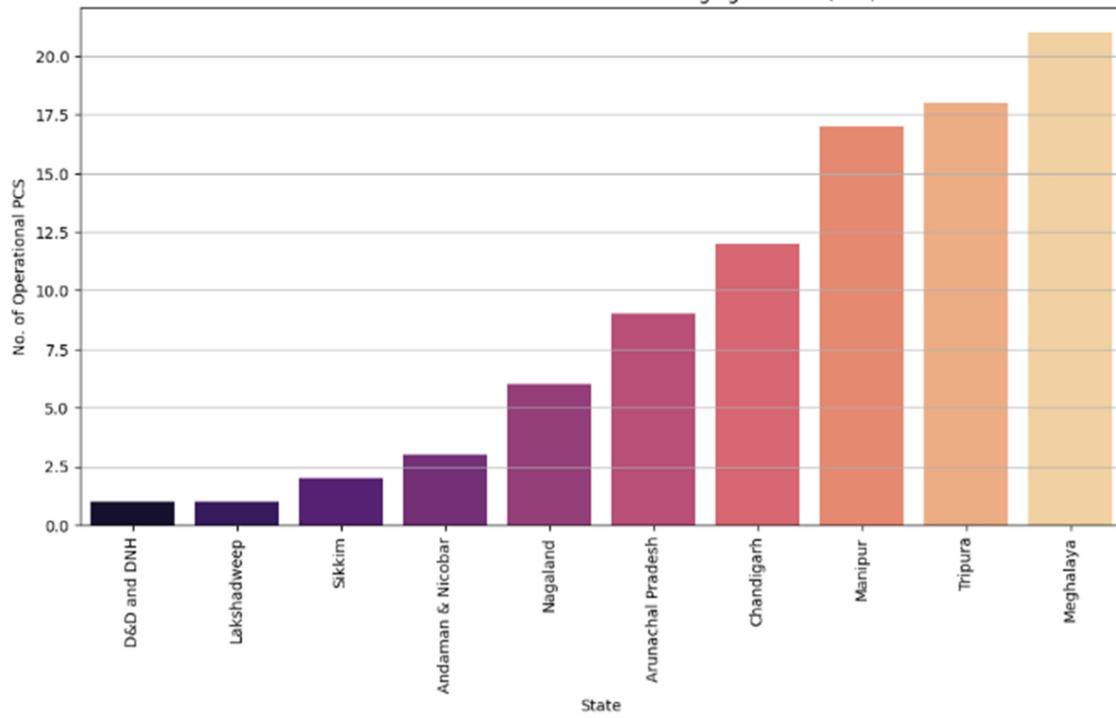
Top 10 states with most public charging stations (PCS) were identified — Maharashtra, Delhi, and Karnataka led the list.

Top 10 States with Most Public Charging Stations (PCS)



Bottom 10 states highlighted the gap in infrastructure for states like Bihar, Jharkhand, and Chhattisgarh.

Bottom 10 States with Least Public Charging Stations (PCS)



Insights:

While manufacturing and sales are growing, charging station penetration remains uneven, creating an opportunity for infrastructure expansion.

A donut chart showed the contribution share of top states in the public charging station network.

7. Segmentation Analysis (Basis for Market Strategy):

Geographic Segmentation: States with maximum sales, manufacturing hubs, and charging stations.

Behavioral Segmentation: Preference for 2-wheelers and 3-wheelers.

Infrastructure Segmentation: Distribution of public charging stations and urban dominance.

Segment Extraction

K-Means Clustering is one of the most popular Unsupervised Machine Learning Algorithms Used for Solving Classification Problems. K Means segregates the unlabeled data into various groups, called clusters, based on having similar features, common patterns.

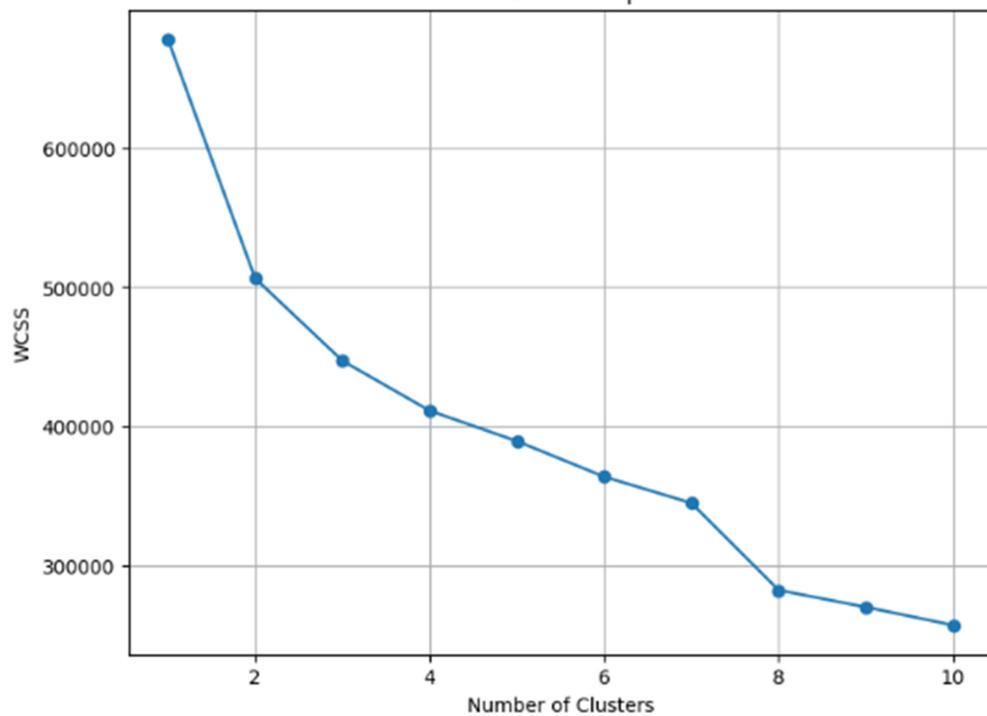
Suppose we have N number of Unlabeled Multivariate Datasets of various features like wateravailability, price, city etc. from our dataset. The technique to segregate Datasets into various groups, on the basis of having similar features and characteristics, is called Clustering. The groups being Formed are known as Clusters. Clustering is being used in Unsupervised Learning Algorithms in Machine Learning as it can segregate multivariate data into various groups, without any supervisor, on the basis of a common pattern hidden inside the datasets.

In the Elbow method, we are actually varying the number of clusters (K) from 1 – 10. For each value of K, we are calculating WCSS (Within-Cluster Sum of Square). WCSS is the sum of squared distance between each point and the centroid in a cluster. Whenwe plot the WCSS with the K value, the plot looks like an Elbow.

As the number of clusters increases, the WCSS value will start to decrease. WCSS value is largest when K = 1. When we analyze the graph, we can see that the graph will rapidly change at a point and thus creating an elbow shape. From this point, the graph starts to move almost parallel to the X-axis. The K value corresponding to this point is the optimal K value or an optimal number of clusters.

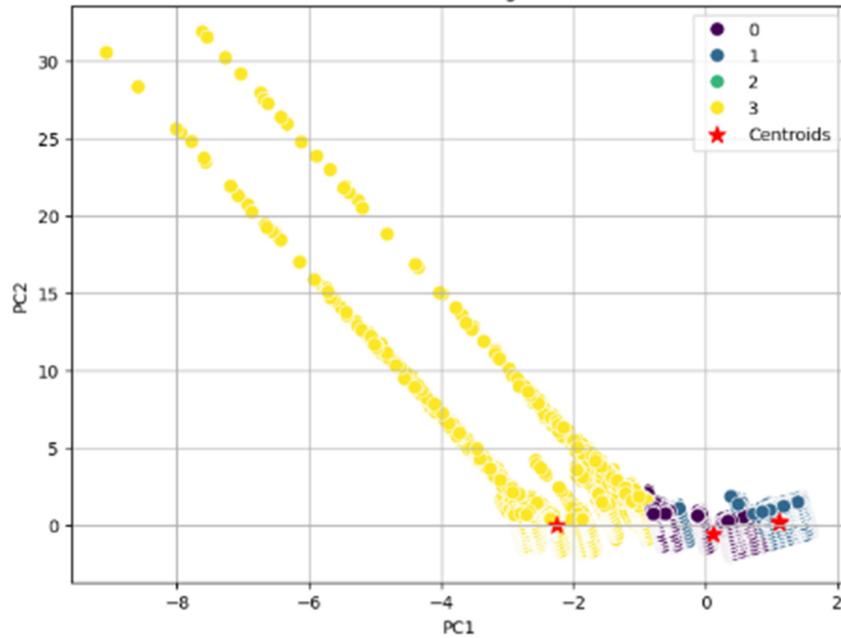
→

Elbow Method for Optimal K



→

K-Means Clustering Results



Profiling and Describing Potential Segments

After applying K-Means clustering to the dataset (or image), the resulting segments represent different groups that share similar characteristics.

Each cluster or segment can be profiled based on the features that are most prominent within it.

For example:

In an image, clusters might group similar colors or patterns.

In a customer dataset, clusters might represent different types of customers based on demographics, purchasing behavior, etc.

Key actions while profiling:

Analyze the centroids: Look at the average values (means) of the features in each cluster.

Identify common traits: For instance, one segment might be characterized by high-value purchases and another by frequent but low-value transactions.

Give each segment a descriptive name: For marketing, you could name segments like "Premium Customers," "Price Sensitive Buyers," or "Frequent Shoppers."

Example (if applied to customer data):

Segment	Description
Cluster 1	Young, tech-savvy, high online spending
Cluster 2	Middle-aged, value-conscious, moderate spending
Cluster 3	Senior customers, prefer offline shopping, low spending

Selection of Target Segment

Once all segments are profiled, it's important to decide which segment(s) to target.

This decision depends on several factors:

Market Size: How large is the segment?

Profitability: What is the potential revenue/profit from the segment?

Accessibility: Can you easily reach and communicate with this group?

Strategic Fit: Does the segment align with the company's strengths and mission?

Example:

If Cluster 1 ("Young, tech-savvy, high spenders") is highly profitable and matches the brand image, it would be the best target.

Sometimes multiple segments can be targeted with different strategies (multi-segment marketing).

Customizing the Marketing Mix (4Ps) for the Target Segment

After selecting the target segment, customize the 4Ps (Product, Price, Place, Promotion) to meet their needs:

4P	Strategy Example
Product	Offer products/services that match the segment's preferences (e.g., high-tech gadgets for tech-savvy youth)
Price	Set a pricing strategy suited for the segment (e.g., premium pricing for quality seekers, discounts for price-sensitive shoppers)
Place	Choose the right distribution channels (e.g., online platforms for younger, tech-savvy customers)
Promotion	Tailor marketing messages and campaigns (e.g., social media campaigns for young customers, TV ads for older audiences)

Goal: Deliver a personalized experience so that the target customers feel the brand understands and meets their needs better than competitors.

Product Mix Customization for Target Segment

The Product Mix refers to the complete range of products or services that a company offers to its customers.

When customizing the product mix for a specific target segment identified through market segmentation, the goal is to ensure that the products meet the specific needs, preferences, and expectations of that segment.

- **Product Variety:** Offer a range of products that align closely with the lifestyle and demands of the segment.

Example: For a "Young, Tech-Savvy" segment, offer a variety of latest gadgets, accessories, and innovative tech solutions.

- **Product Features:** Customize product features to fit the specific desires of the target customers.

Example: Emphasize lightweight design, latest technology, eco-friendly materials for environmentally conscious young consumers.

- **Product Quality :** Maintain the quality level expected by the segment.

Example: For premium buyers, ensure high-end finishing, durability, and top-notch functionality.

- **Branding :** Position the brand to resonate with the segment's values and self-image.

Example: Use youthful, innovative branding for young customers, and trustworthy, heritage branding for older audiences.

Prices and Pricing Strategies

Pricing plays a critical role in positioning the product for the target segment. After segmentation, prices must reflect the perceived value, affordability, and expectations of the selected group.

Key Pricing Strategies:

- **Value-Based Pricing:**

Set the price based on the value the product provides to the customer, not just the cost.

Example: Premium gadgets for tech-savvy users priced higher due to innovative features.

- **Penetration Pricing:**

Introduce products at a low price to attract customers quickly and gain market share.

Example: A new smart gadget at an affordable price for mass adoption among young users.

- **Psychological Pricing:**

Pricing products slightly below a round number (like ₹9,999 instead of ₹10,000) to make them seem more affordable.

- **Segmented Pricing:**

Different prices for different customer segments, even if the product is the same.

Example: Special discounts for students or young professionals.

Promotional Mix

The promotional mix refers to the blend of different communication tools and strategies used to promote the product to the target segment. After segmentation, the promotional efforts are tailored to match the preferences and behaviors of the selected audience.

Advertising:

Use targeted ads through channels like social media, search engines, TV, or print, depending on where the segment is most active.

Sales Promotion:

Short-term incentives like discounts, free trials, or cashback offers to quickly boost sales and attract attention.

Public Relations (PR):

Building a strong brand image through media coverage, press releases, influencer partnerships, and event sponsorships.

Personal Selling:

One-on-one interaction through sales representatives to build relationships and provide customized solutions, especially useful for premium segments.

Direct Marketing:

Personalized communication like emails, SMS, or app notifications to directly engage customers with offers and information.

Place/Distribution

Place (or distribution) refers to how the product or service is delivered to the target customers. It ensures that the product is available at the right location, at the right time, and in the right quantity to meet customer demand.

Distribution Channels:

Choosing the best channels such as online platforms (websites, apps), retail stores, wholesalers, or direct-to-consumer models based on the segment's buying behavior.

Logistics and Supply Chain:

Managing inventory, warehousing, and transportation efficiently to ensure product availability without delays.

Coverage Strategy:

Deciding whether to go for intensive (everywhere), selective (specific outlets), or exclusive (limited and premium locations) distribution depending on the segment's needs.

Technology Integration:

Using e-commerce, online delivery partners, and digital payment systems to make the buying process seamless for tech-savvy segments.

Most Optimal Market Segment

There are many EV manufacturing companies in the country like Hero Electric, Tata Motors, Ather Energy, Ashok Leyland, Hyundai Kona Electric, etc. Tesla has also arrived; the demand will get higher & higher since it is automotive so the investments and policies and all that would be bigger but it will take some time to perfectly settle in India. The following are the key insights of the

project:

- The electric vehicle industry has not done that much good due to the devastating hit of the Covid outbreak but it will take a huge jump in upcoming years.
- The use of EVs will be game-changing in terms of environment, air, noise pollution-free, postelectric, and much more.
- The company should plan to establish local operations in India either by partnering with a local company or by setting up its own manufacturing/ development unit, potentially combined with imports of specific components.
- The company would expect to further grow in India, underpinned by a growing commercial fleet market for two-wheelers and three-wheelers especially for last km delivery/urban freight services. The company must see opportunities across the supply chain in the battery, EV component and charging infrastructure segments including the machinery and equipment needed for establishing manufacturing plants, training and provision of skilled workforce etc.
- The company should start their business from Metro Cities in India and then after considerable business expand to other cities of the same state of the Metro Cities. This will help the company to expand easily as they will be having a prior knowledge of business from Metro Cities and Network of Supply chain will be easy for the company as the time goes in business.