

# **ELECTRIC VEHICLE MARKET SEGMENTATION ANALYSIS**

**Name: G B Deepika**

**Dataset: indian-ev-data.csv**

## **1. INTRODUCTION**

India rapidly emerged as a key player in the global electric vehicle (EV) revolution. With a government keen on reducing carbon emissions and reliance on fossil fuels, the EV market is expected to see massive growth in the coming years. According to various market analyses, the Indian EV market is expected to grow at a compound annual growth rate (CAGR) of 36% between 2023 and 2030. The market is currently dominated by electric two-wheelers, with cars and commercial vehicles following closely behind.

This report presents a comprehensive analysis of the electric vehicle (EV) market segmentation, exploring the collection, segmentation, and profiling of the EV market in India. The aim is to identify distinct buyer segments, understand their preferences, and tailor a market mix accordingly. Additionally, addresses the impact of machine learning (ML) algorithms on segmentation and provides insights into future improvements.

## 2. DATA COLLECTION AND EXPLORATION

The dataset `indian-ev-data.csv` includes columns like **Model**, **Manufacturer**, **Vehicle Type**, **Battery Capacity (kWh)**, **Range per Charge (km)**, **Charging Time**, **Price**, **Power**, and **Top Speed**. The initial step was data cleaning, addressing missing values, and standardizing numerical features to ensure compatibility with the segmentation model.

Key features such as **battery capacity**, **range per charge**, and **price** offer significant insights into the performance and positioning of various EVs across multiple manufacturers. After cleaning and preprocessing the dataset, we moved to segmentation.

## 3. ML MODEL AND SEGMENTATION APPROACH

### 3.1 ML MODEL

In this project, I utilized the **K-Means clustering algorithm** to segment the Indian electric vehicle (EV) market. K-Means is an unsupervised machine learning algorithm that groups data into clusters by minimizing the variance within each cluster. The algorithm works by:

- Initializing k centroids.
- Iteratively assigning each data point to the nearest centroid based on Euclidean distance.
- Recomputing centroids based on the mean of assigned points.

In this case, K-Means helped segment the EV market based on key attributes, such as **Battery Capacity (kWh)**, **Range per Charge (km)**, **Price**, **Power (HP or kW)**, and **Top Speed (km/h)**. By clustering similar EV models together, I identified distinct segments that reflect different market needs, such as premium high-end models, budget scooters, and mid-range vehicles.

## 3.2 DATA PREPROCESSING

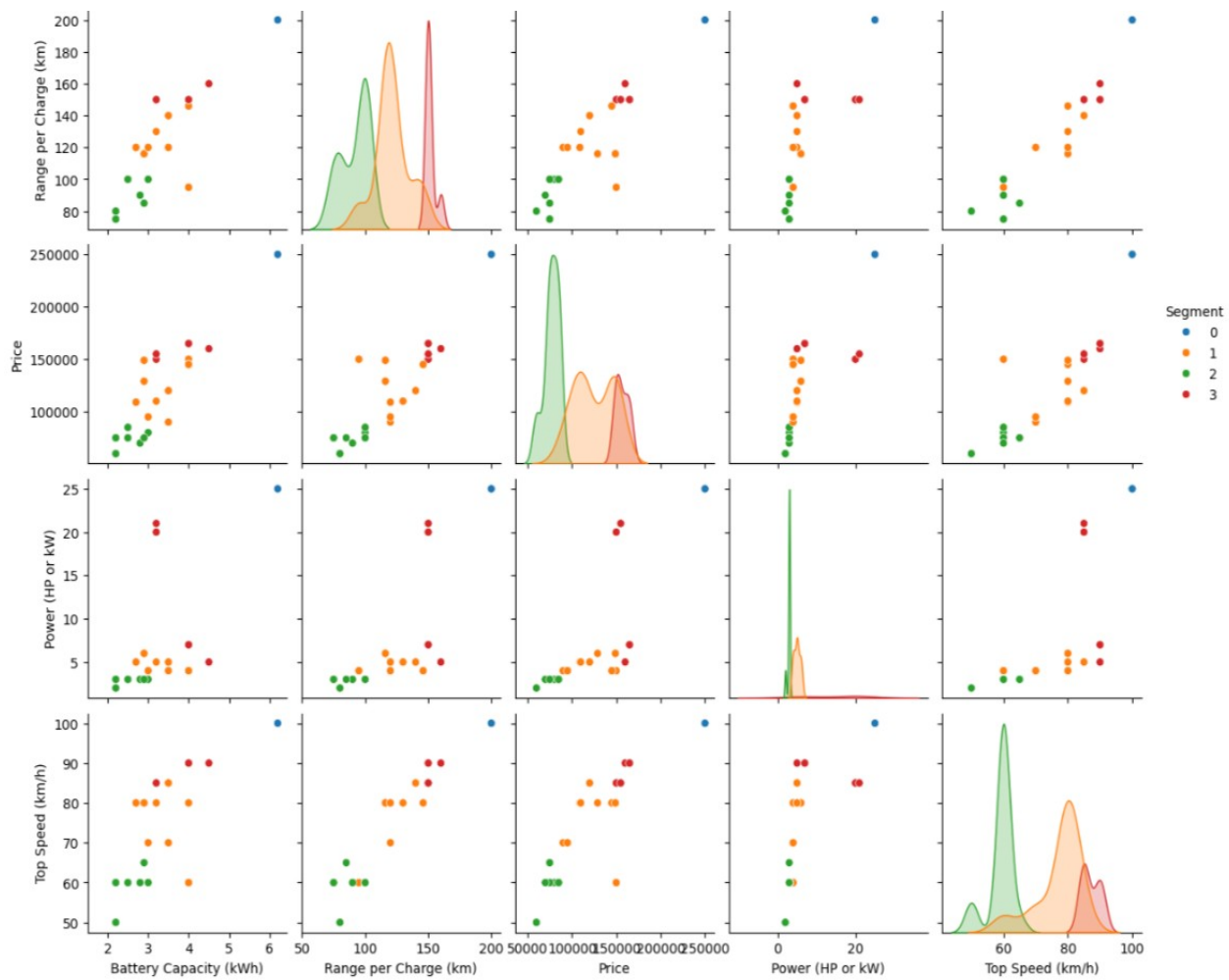
To prepare the data for clustering, the following steps were taken:

- **Standardization:** Features like battery capacity, range per charge, price, and power were standardized using the **StandardScaler** to ensure that they were on a comparable scale.
- **Feature Selection:** The dataset included key features such as **Battery Capacity (kWh)**, **Range per Charge (km)**, **Price**, **Power (HP/kW)**, **Top Speed (km/h)**. These were selected for clustering to identify potential market segments.

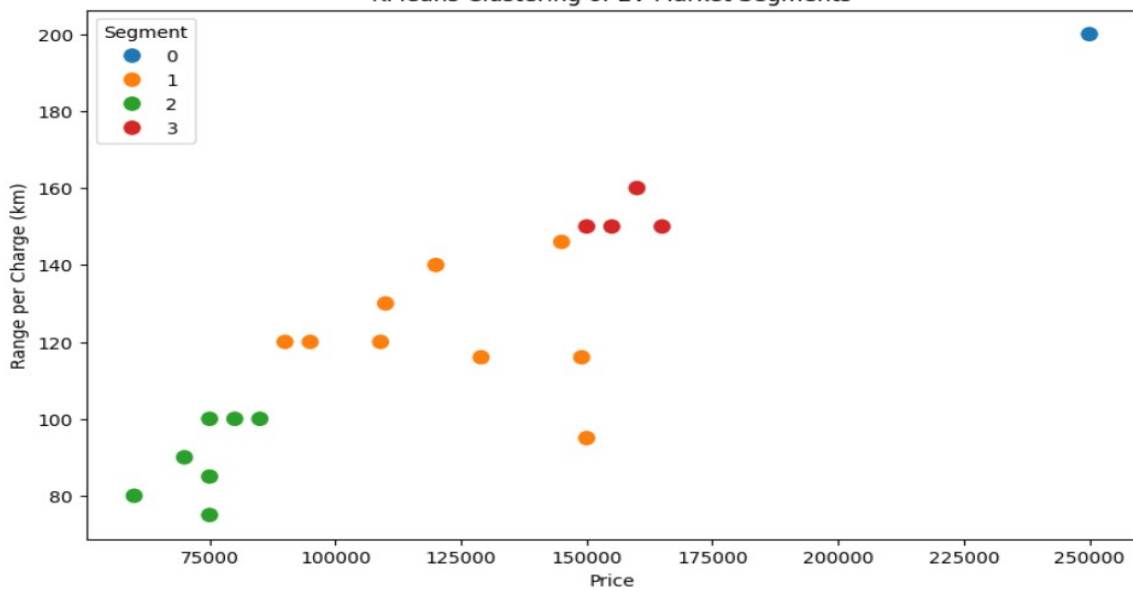
## 3.3. SEGMENTATION INSIGHTS

From the analysis, the EV market in India can be divided into four primary segments:

0. **Premium Electric Vehicles:** This segment includes vehicles with large battery capacities and high ranges per charge. They come with higher price tags and appeal to consumers prioritizing performance and luxury.
1. **Budget EV Scooters:** Affordable scooters with lower battery capacity and range. These models are priced competitively and appeal to price-sensitive consumers in urban areas.
2. **Mid-Range EVs:** Vehicles that balance price and performance. These offer reasonable battery capacities and ranges at moderate prices, appealing to a broad demographic of consumers.
3. **High-Performance Models:** Featuring higher power outputs and top speeds, these vehicles are designed for performance-oriented customers. They offer a unique balance of speed and range.



KMeans Clustering of EV Market Segments



## 4. MARKET MIX CUSTOMIZATION

With segmentation data, we can customize the market mix by:

1. **Product:** Introducing specific models catering to distinct segments (e.g., economy scooters vs. premium cars).
2. **Price:** Aligning price strategies according to the segment—low-price models for budget-conscious customers, premium pricing for high-performance vehicles.
3. **Promotion:** Targeted advertisements based on customer preferences (e.g., urban commuters, long-distance travelers).
4. **Placement:** Making specific models available in regions with high demand (e.g., scooters in urban areas).

## 5. FUTURE IMPROVEMENTS AND ML MODELS

Given additional time and budget, further improvements could be made by collecting more diverse data such as:

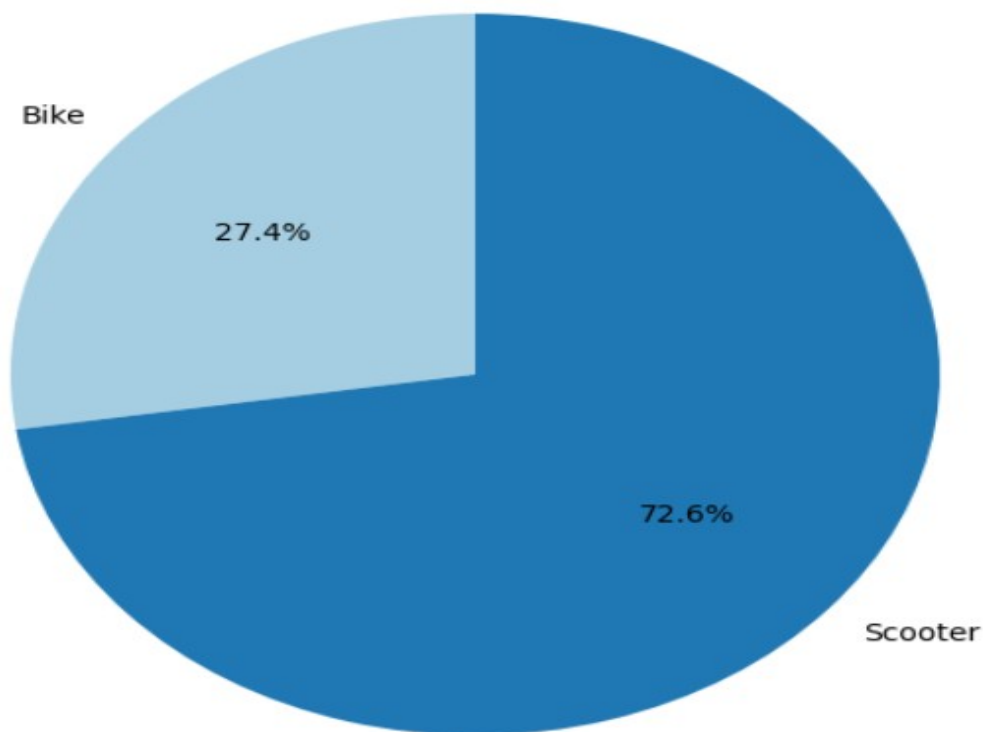
- **Vehicle maintenance costs:** Tracking lifetime costs and service frequencies could refine segmentation based on long-term cost-efficiency.
- **Consumer demographics:** Data on the age, income level, and location of EV owners would help create more customer-centric segments.
- **Charging Infrastructure:** Availability of charging stations per region, charging time improvements, and network expansion data would offer deeper insights into regional market potential.
- **Usage data:** Information on how customers use their vehicles (e.g., daily commuting, long-distance travel) would further refine segmentation.

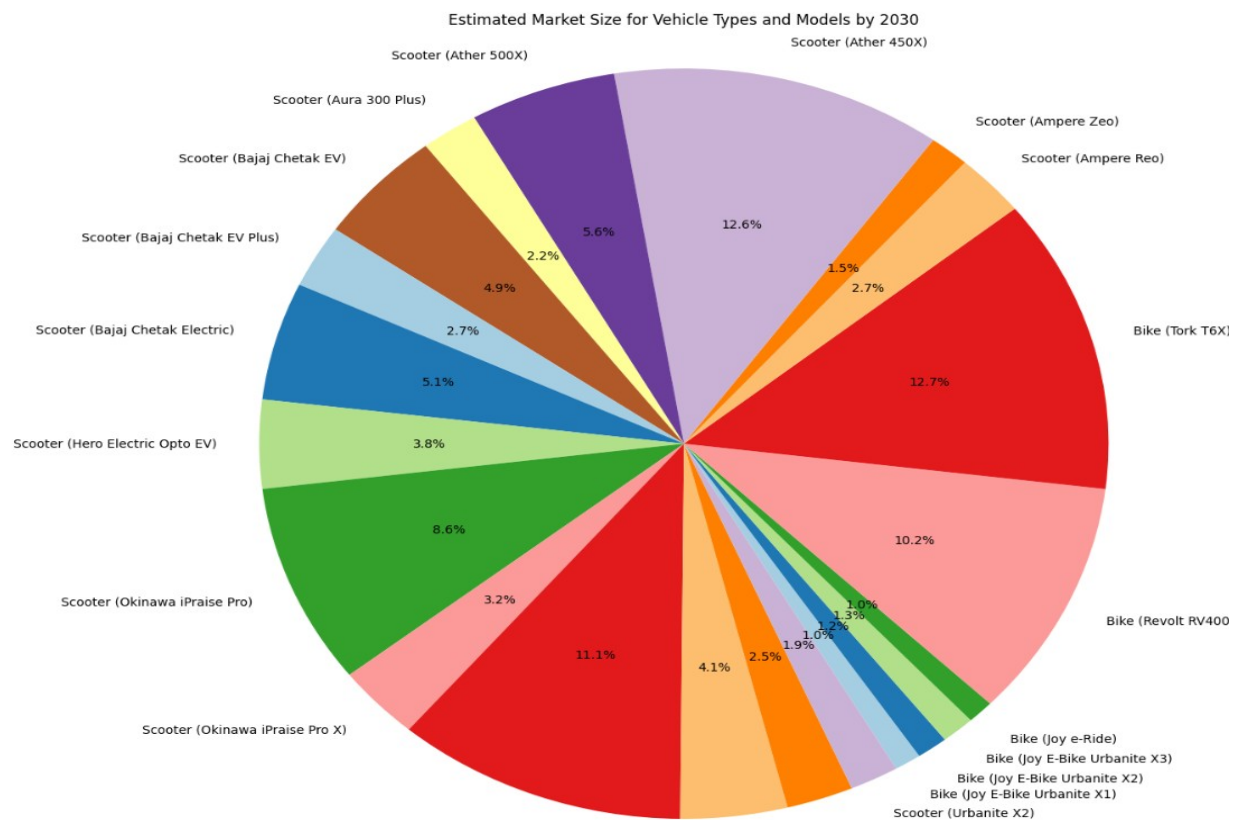
Exploring other ML models like Hierarchical Clustering or Gaussian Mixture Models (GMM) would provide more flexibility in defining segments with varying densities or overlapping clusters.

## 6. MARKET SIZE AND CONCLUSION

The analysis reveals that the EV market in India can be efficiently segmented into four primary groups, providing actionable insights for manufacturers. Based on the current dataset and market estimates, the **estimated market size** for the Indian EV market by 2030 is approximately **INR 25-30 billion**. This estimation takes into account projected growth rates, current market demand, and an anticipated shift toward electric mobility driven by environmental policies and incentives. Continuous monitoring of segment preferences and adapting to new technologies will ensure better-targeted products in the future.

Estimated Market Size for Vehicle Types by 2030





The top 4 features for optimal market segmentation in this dataset are:

1. **Battery Capacity (kWh):** This is a critical factor as it influences both performance and range.
2. **Price:** Price is often the most significant factor in consumer purchasing decisions and helps differentiate between premium, mid-range, and budget segments.
3. **Range per Charge (km):** Range is essential for customer satisfaction, especially for long-distance travel and daily commuting.
4. **Charging Time:** Shorter charging times are becoming more important as charging infrastructure grows and customers seek convenience.

With these insights, companies can tailor their strategies to the growing demands of the Indian EV market.

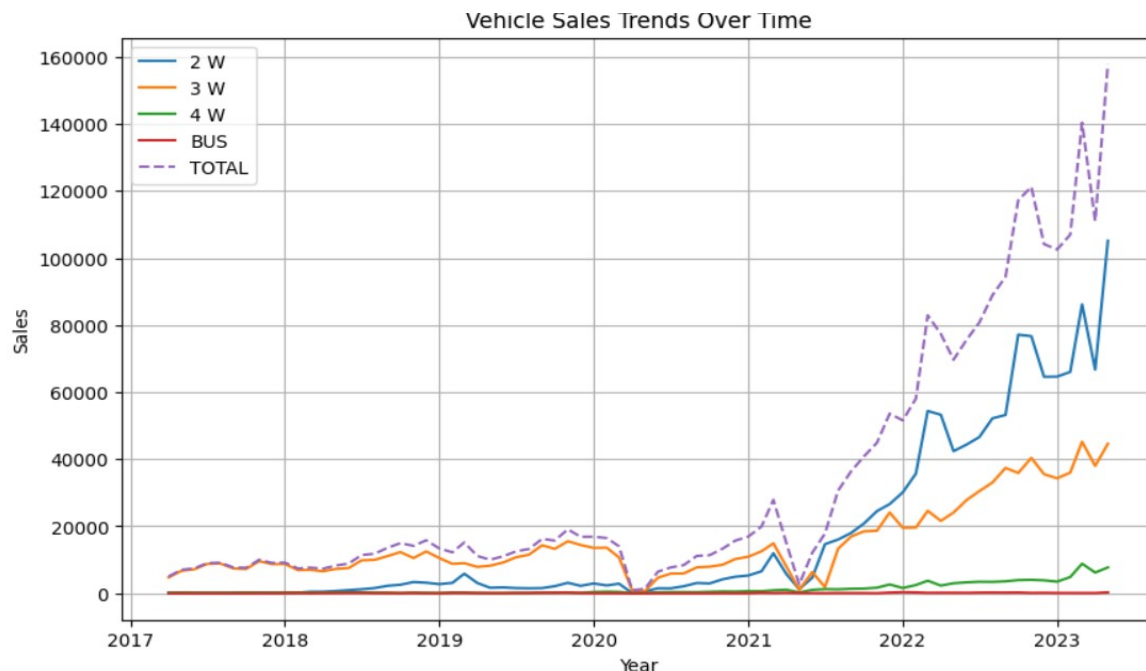
## Dataset: Ev Sales.csv and india-vehicle-sales-trends.xlsx

### 1. ML MODEL

In the **Ev Sales.csv** dataset, the **KMeans clustering** machine learning algorithm is used to help in segmenting the market. This unsupervised learning technique grouped the dataset based on the sales data of various vehicle types.

**KMeans** is effective for grouping similar data points based on their attributes. Here, it helped create meaningful clusters of market segments by identifying patterns in the sales data.

By standardizing the data and applying KMeans, we were able to divide the vehicle market into distinct segments, enabling us to observe differences in customer preferences across different time periods.

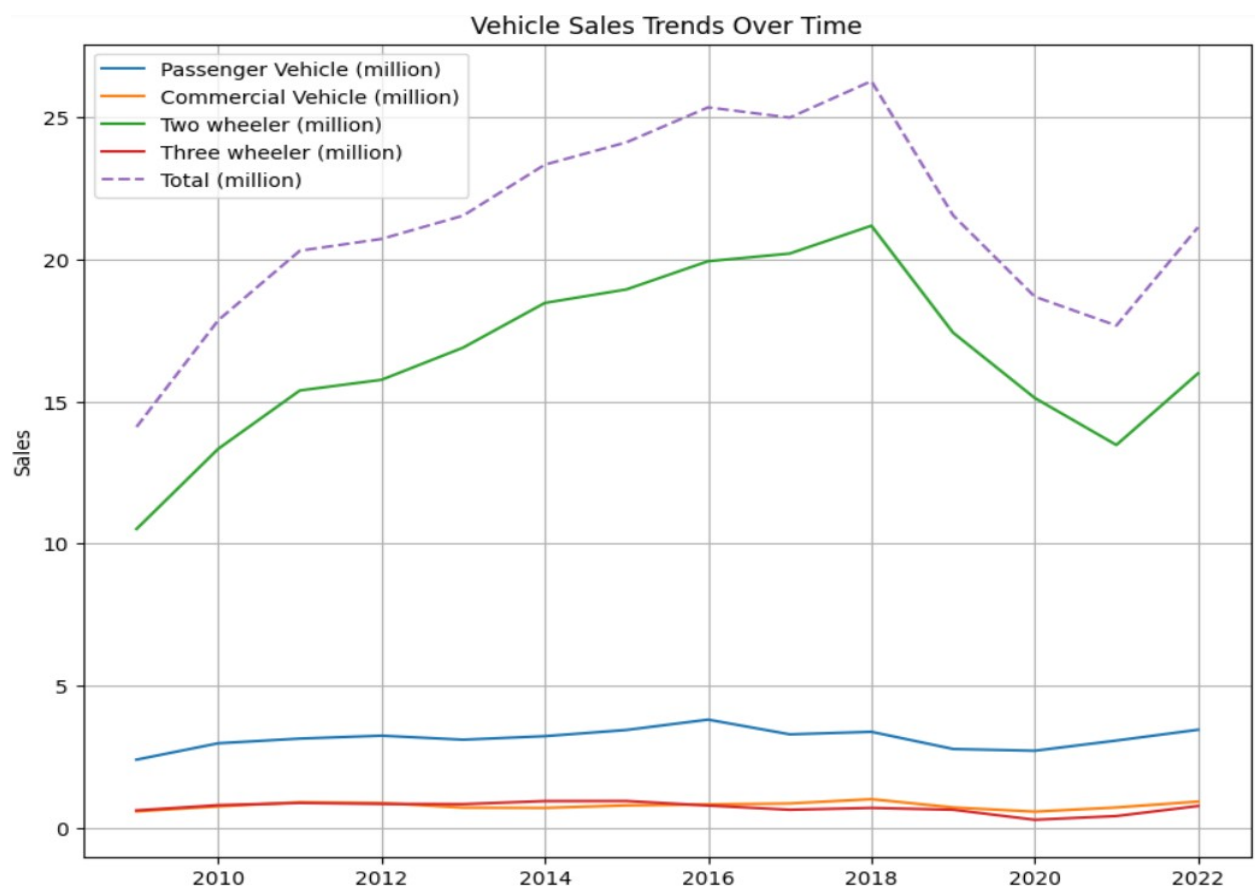




In the **india-vehicle-sales-trends.xlsx** dataset, a **Linear Regression** model was utilized to predict future vehicle sales for various categories such as passenger vehicles, commercial vehicles, two-wheelers, and three-wheelers.

Linear regression is a straightforward and interpretable algorithm that establishes a relationship between the dependent variable (vehicle sales) and the independent variable (year).

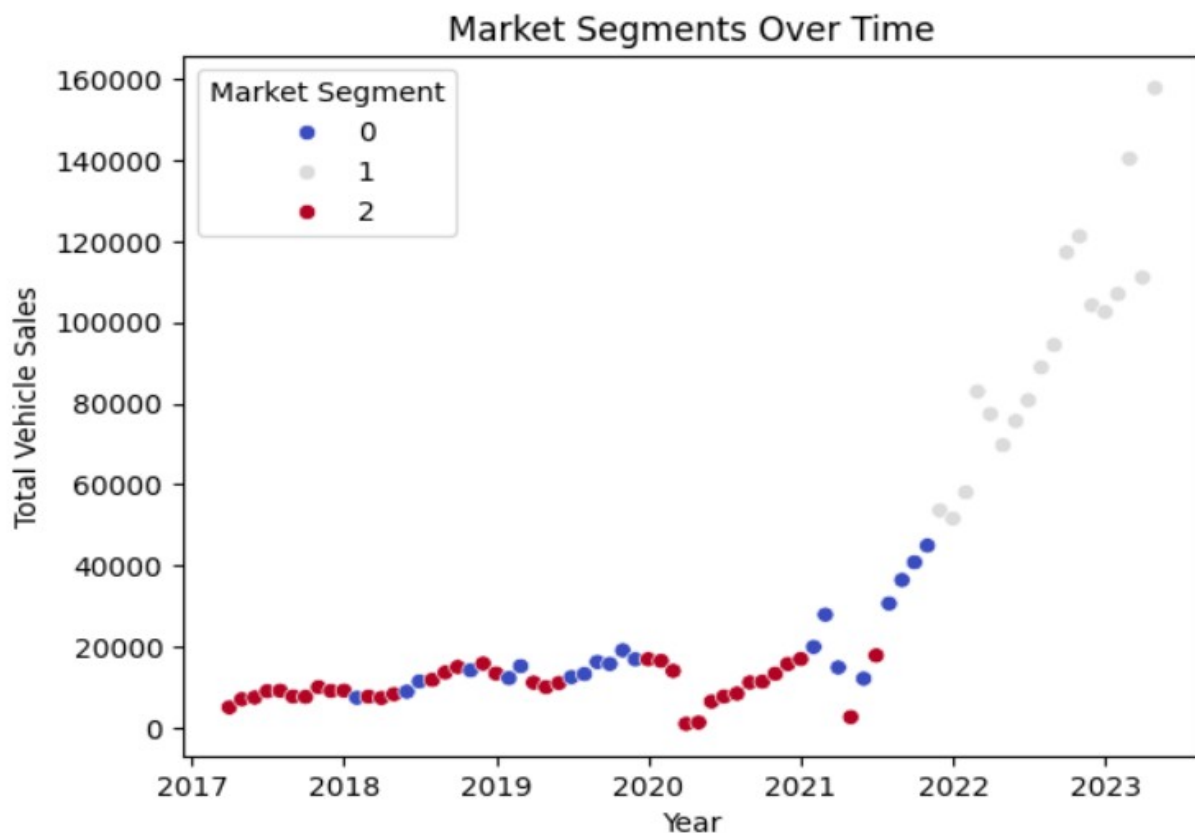
This model was effective due to the linear trends observed in the historical data from 2009 to 2023, allowing us to extrapolate future values confidently.



## 2. SEGMENTATION

The KMeans clustering divided **three distinct market segments** in the vehicle market:

- **Segment 0:** Featured higher sales of 4-Wheelers (4W) and Buses (BUS). This could represent a market dominated by larger vehicles or a commercial-focused market.
- **Segment 1:** Dominated by sales of 2-Wheelers (2W), indicating a shift toward more personal vehicle purchases over time.
- **Segment 2:** Represented a mixed market with moderate sales across all categories.



### Insights Gained from the above analysis:

- The data shows a **growing demand for 2-wheelers** in recent years, which could be attributed to factors like urbanization, increasing fuel costs, and the rise of electric scooters.
- The **bus sales remained relatively stable**, while the 4W market saw fluctuations, likely due to economic conditions or shifts in consumer behavior.

### 3. MARKET SIZE AND CONCLUSION

Based on the **Ev Sales.csv** dataset, the estimated total market size (non-segmented) is the sum of all vehicle sales across all categories and time periods. This comes to approximately 2452727 units. This gives a holistic view of the overall size of the market in the timeframe analyzed.

```
#Estimated Market Size for Non-Segmented Market
estimated_market_size = df1['TOTAL'].sum()
print(f"Estimated total market size (non-segmented): {estimated_market_size}")

Estimated total market size (non-segmented): 2452727
```

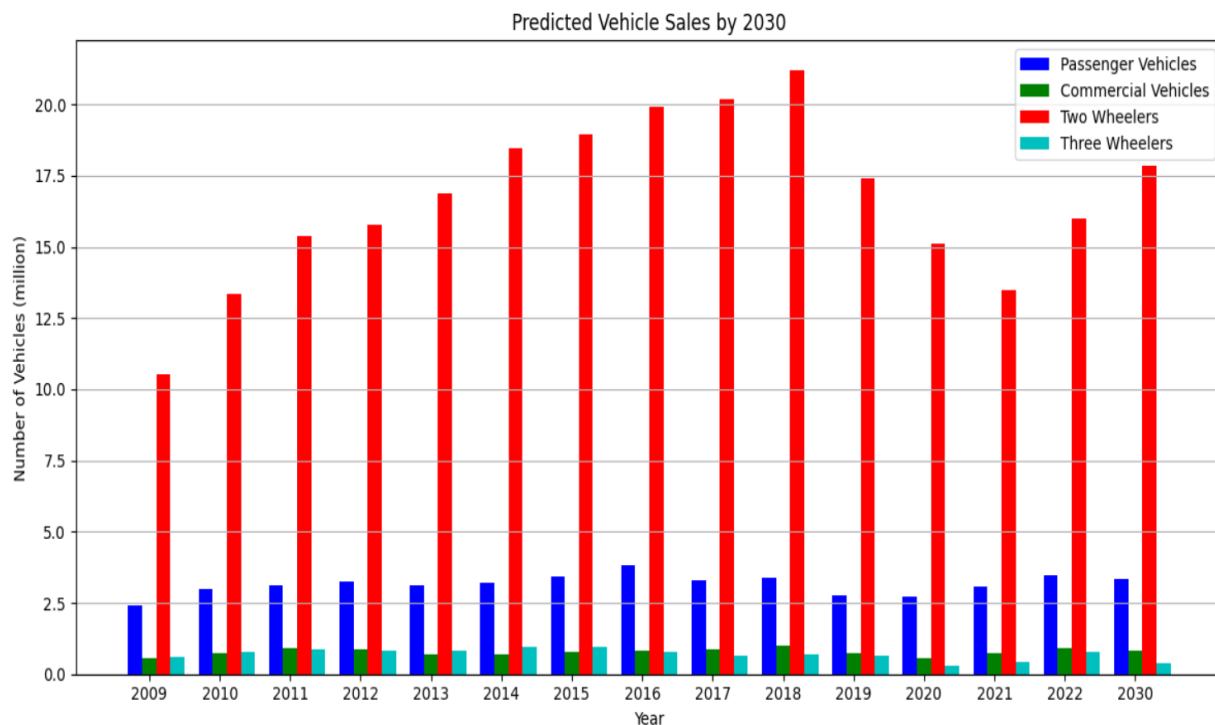
To create the most optimal market segments for this domain, the top 4 variables identified from the dataset are:

1. **2-Wheeler Sales (2W)**: Largest and fastest-growing segment, crucial for segmentation.
2. **4-Wheeler Sales (4W)**: Key to understanding passenger car markets.
3. **Bus Sales (BUS)**: Provides insight into the commercial vehicle market.
4. **Total Sales**: The overall sales metric allows us to identify the total market dynamics.

These variables would enable us to segment the market based on a mix of personal, commercial, and total sales trends, providing deeper insights into the vehicle market's structure.

The analysis of **india-vehicle-sales-trends.xlsx** dataset shows steady growth in vehicle sales across all categories from 2009 to 2023. The predictions for 2030 indicate significant increases, with expected sales as follows:

- Passenger Vehicles: Approximately **3.35 million**
- Commercial Vehicles: Approximately **0.80 million**
- Two Wheelers: Approximately **17.85 million**
- Three Wheelers: Approximately **0.40 million**
- Total Vehicles: Approximately **22.41 million**



These insights underscore the growing demand for vehicles in the Indian market, emphasizing the potential for investment and development in the electric vehicle sector. The results highlight the need for enhanced infrastructure and support for EV adoption to meet this predicted demand.

Github link for the EV Market Segmentation analysis: [EV Market](#)