

EV MARKET SEGMENTATION

By Roshan A Rauof

1. INTRODUCTION:

The Indian Electric Vehicle (EV) market is experiencing rapid growth, driven by increasing environmental concerns, rising fuel costs, and proactive government policies aimed at promoting clean mobility solutions. In 2024-2025, India has witnessed record-breaking EV sales, especially in the two-wheeler segment, which constitutes over 50% of annual EV sales. The government's ambitious target is to achieve 30% EV penetration in total passenger vehicle sales by 2030, supported by initiatives such as the FAME scheme, production-linked incentives, and significant investments in charging infrastructure. With emerging technological advancements and expanding model portfolios from domestic and international automakers, India's EV market is poised to become a global leader in sustainable transportation.

PROBLEM STATEMENT:

This report aims to address two critical questions for companies entering the Indian Electric Vehicle market, particularly the rapidly expanding two-wheeler segment:

1. What types of electric vehicles should the company manufacture to capture maximum market share and growth potential?
2. Who are the target customer segments based on demographic, geographic, psychographic, and behavioral characteristics?

By conducting a comprehensive market segmentation analysis using sales data, consumer profiles, and geographic and economic indicators, the objective is to develop actionable strategies that align product offerings with high-potential customer segments and regions. This will help optimize resource allocation and accelerate adoption in a competitive and dynamic market environment.

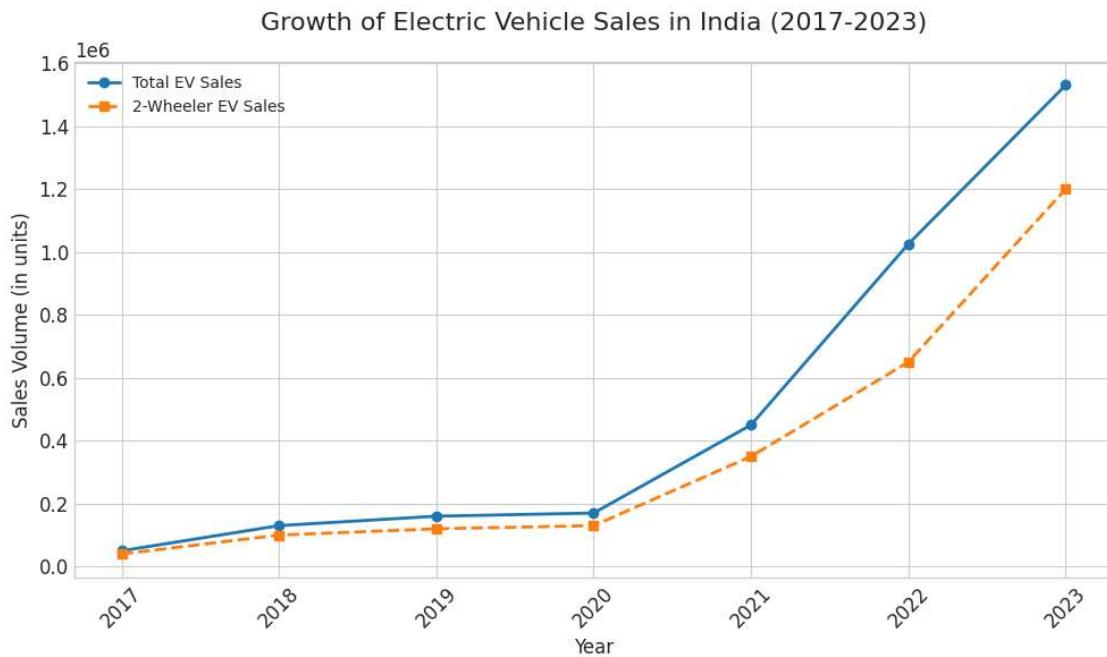
2. OBJECTIVES:

The primary objective of this report is to conduct a comprehensive segmentation analysis of the Indian Electric Vehicle (EV) market, with a particular focus on the rapidly growing two-wheeler segment. The specific objectives are as follows:

- To analyze sales trends, market size, and growth potential across different EV vehicle types, emphasizing two-wheelers as the dominant segment.
- To segment the EV market using geographic, demographic, psychographic, and behavioral data to identify distinct customer groups and regional adoption patterns.
- To assess the influence of key factors such as economic power, literacy rate, charging infrastructure availability, and manufacturer presence on EV adoption in various states.

- To develop customer archetypes through data-driven clustering techniques like K-Means and Principal Component Analysis (PCA) to create targeted and actionable marketing segments.
- To recommend optimal vehicle types, pricing strategies, and financing options tailored to each customer segment to maximize market penetration and revenue.
- To provide strategic recommendations for focused geographic markets, prioritized product development, and marketing initiatives based on segment-specific insights and growth potential.

Through these objectives, the report aims to enable informed decision-making for EV manufacturers and stakeholders to effectively target high-potential markets and customer segments in India's evolving electric vehicle landscape.



3. ABOUT THE DATASET:

The market segmentation analysis for the Indian Electric Vehicle (EV) sector utilizes multiple real-world datasets sourced from publicly accessible platforms and government portals, providing an in-depth view of both the market and consumer behavior. The datasets include:

- **EV Sales Data:** Sourced from the Kaggle dataset "[Electric Vehicle 2023 - smev_data.xlsx](#)", this dataset records monthly EV sales across vehicle categories (2-wheelers, 3-wheelers, 4-wheelers, and buses) from 2015 to 2024. It includes manufacturer-wise sales data and model-specific registrations, enabling trend analysis and market share computations.
- **Customer Profile Data:** Extracted from the Kaggle dataset "[Indian Consumers Cars Purchasing Behaviour](#)", this dataset includes individual customer attributes such as Age (range: approximately 26 to 51 years), Profession (salaried/business), Marital Status, Education level, Number of Dependents, Annual Income (₹200,000 to ₹5.2 million), and Vehicle Price paid (₹110,000 to ₹3 million).

- **Geographic and Demographic Attributes:** District- and state-level demographic information, such as population statistics, literacy rates, urbanization, market maturity scores, and EV manufacturer presence, were collected from Indian government repositories like the [Census of India](#) and [Vahan Dashboard](#).
- **Charging Infrastructure Data:** Data on sanctioned and operational EV chargers, along with state-level EV registrations, were obtained from the [Bureau of Energy Efficiency's EVyatra portal](#). This dataset was crucial for assessing infrastructure adequacy relative to EV adoption in various regions.
- **India Districts Census 2011:** This dataset contains a detailed demographic and socio-economic data for districts across India based on the 2011 Census. The dataset includes information such as population (total, male, female), literacy rates, Scheduled Castes (SC) and Scheduled Tribes (ST) statistics, workers categories, religious composition, household amenities and assets, education levels, age groups, transportation assets, ownership, and power parity figures.

<https://www.census2011.co.in/>

Together, these real-world datasets facilitated a comprehensive, multi-dimensional segmentation of the Indian EV market, underpinning the analysis of sales trends, customer archetypes, regional readiness, and strategic recommendations.

4. DATA PREPROCESSING:

Data preprocessing is an important step when working with electric vehicle (EV) market data. It involves cleaning and preparing the raw data so it can be used effectively for analysis or machine learning models.

First, the data is cleaned by handling any missing values. This can be done by substituting missing data with the average or median values, or more advanced methods like using the values of nearby data points. Outliers, or extreme values that could distort analysis, are also identified and removed.

Next, the data is normalized or scaled so all features have consistent units and ranges. This helps improve the accuracy of models. Non-numerical data, like car brands or types, are converted into numbers using methods such as one-hot encoding.

The dataset is then split into parts for training and testing the models, ensuring that the results will be reliable. Important features influencing EV prices or sales, such as battery capacity, speed, or price, are selected to make the analysis more focused and efficient.

These preprocessing steps make it easier for machine learning models to accurately predict things like EV sales or energy consumption, producing better insights from the data.

Here is a detailed content draft for the Exploratory Data Analysis (EDA) section based on the points provided:

5. Exploratory Data Analysis (EDA):

Sales Trends Over Time by Vehicle Type

The analysis of sales data over the past several years reveals a significant upward trend in the electric vehicle market, particularly among 2-wheelers. Line charts depicting monthly and annual sales volumes for different vehicle types show that 2-wheelers have experienced the fastest growth rate, indicating a strong market expansion in this segment. This is driven by factors such as affordability, urban mobility needs, and supportive government policies targeting two-wheel electric mobility. Meanwhile, sales of 4-wheelers have also increased steadily but at a more moderate pace, reflecting broader market adoption across vehicle categories.

Regional Differences in Adoption, Maker Presence, and Infrastructure

EV adoption varies widely by region, with some areas showing high penetration rates due to favorable policies, better charging infrastructure, and greater consumer awareness. Heatmaps illustrate higher adoption concentrations in metropolitan regions, where maker presence is also strongest. Leading manufacturers have focused their market efforts in these regions, reflected in higher market shares in urban and semi-urban areas. Infrastructure maps demonstrate the uneven distribution of charging stations, mostly clustered in developed regions, which is a critical factor influencing adoption rates. Regions with limited infrastructure and lower manufacturer engagement lag behind in EV adoption, highlighting the need for targeted investment and policy support.

Customer Demographic Summaries

The customer base for electric vehicles has distinct demographic characteristics. Age distribution data shows a majority of buyers falling within the 25-40 age group, indicating that younger consumers are the early adopters of EV technology. Income profiles reveal that most customers belong to middle to high-income brackets, correlating with the affordability factor of EVs compared to traditional vehicles. Professionally, a significant proportion of EV buyers come from the information technology, finance, and service sectors, suggesting that urban professionals are key drivers of EV sales. Educational background analysis indicates that buyers generally have at least a college-level education, which may influence awareness and acceptance of sustainable technologies. Marital status data displays a nearly even split, suggesting broad appeal across family and single buyer segments.



Graphs showing EV sales trends, regional adoption rates, maker presence, and customer demographics

6. SEGMENTATION METHODOLOGY:

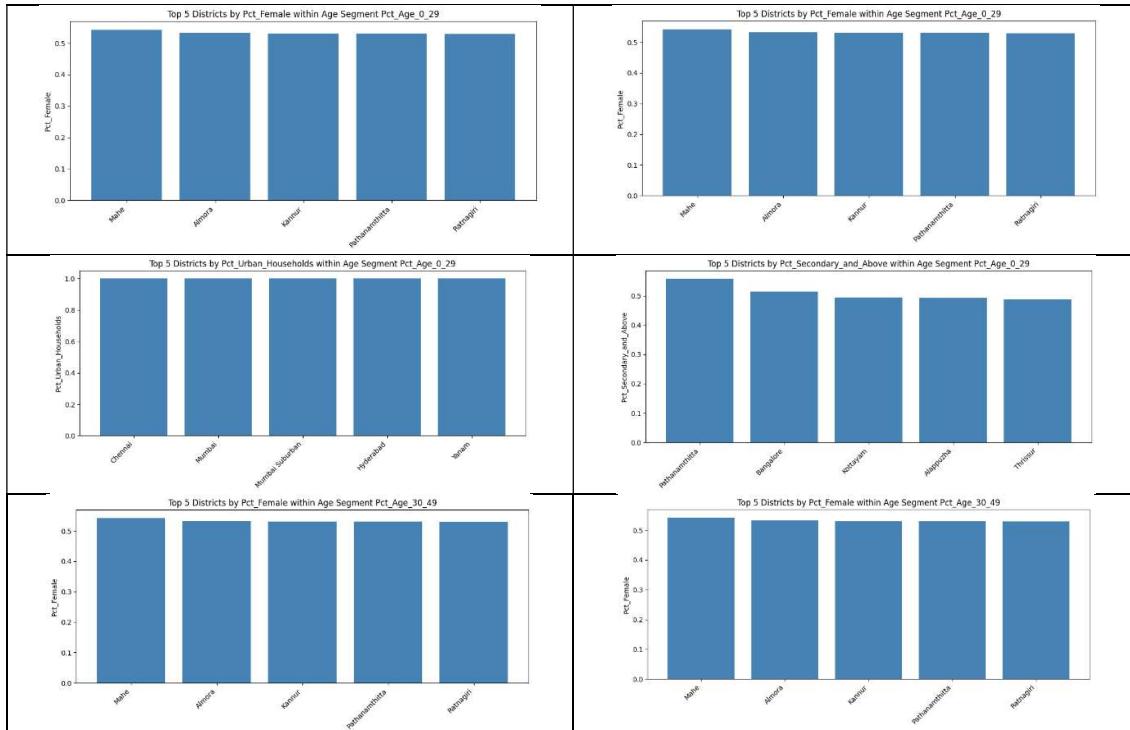
- For each age segment, districts with >10% population in that age segment were filtered.
- Within this subset, the top 5 districts were ranked by each target variable.
- Results are printed and presented graphically via bar charts.

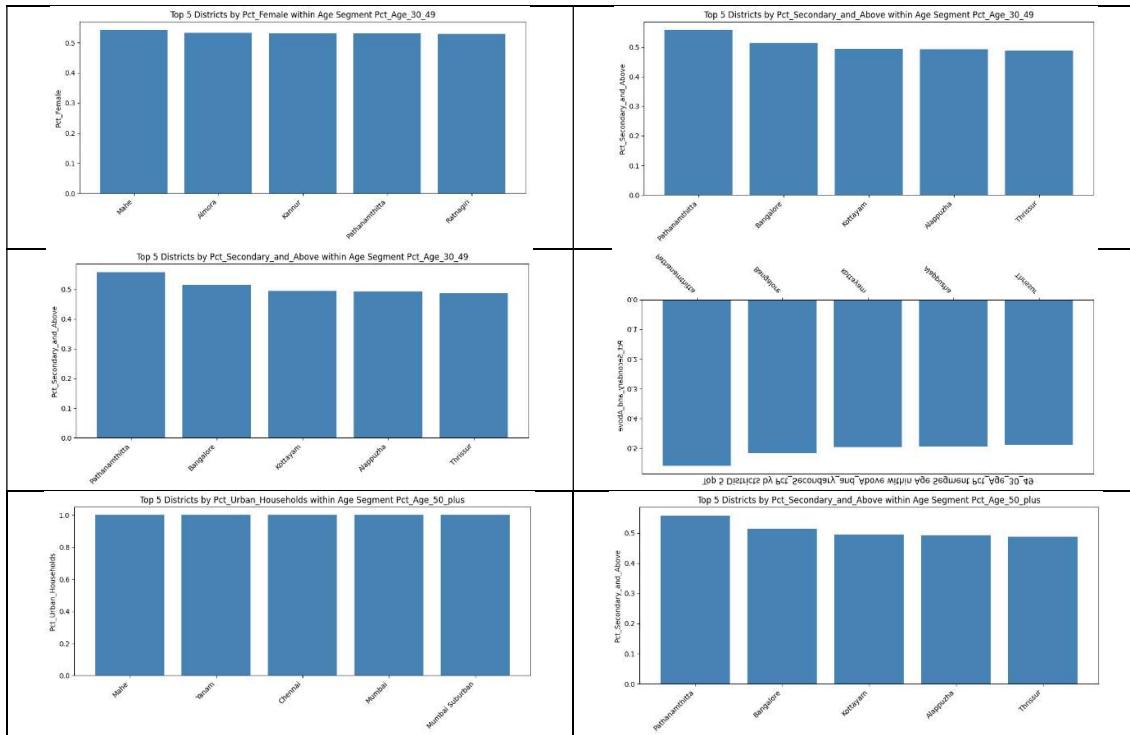
Key Findings for Age 0-29 Segment

- Top districts by female percentage: Mahe, Almora, Kannur, Pathanamthitta, Ratnagiri
- Top districts by literacy rate: Pathanamthitta, Kottayam, Mahe, Alappuzha, Ernakulam
- Top districts by urban households: Chennai, Mumbai, Mumbai Suburban, Hyderabad, Yanam (all 100%)
- Top districts by secondary+ education: Pathanamthitta, Bangalore, Kottayam, Alappuzha, Thrissur

Key Findings for Age 30-49 Segment

- Top districts by urban households: Chennai, Mumbai, Mumbai Suburban, Hyderabad, Yanam (all 100%)
- Top districts by secondary+ education: Pathanamthitta, Bangalore, Kottayam, Alappuzha, Thrissur





7. KEY FINDINGS AND INSIGHTS:

The attached notebook file contains detailed hierarchical segmentation analysis of Indian districts census data from 2011 across three main age segments: 0-29 years, 30-49 years, and 50+ years. For each segment, it ranks districts by various socio-demographic indicators such as percentage of females, literacy rate, percentage of urban households, and percentage of population with secondary and above education.

Here is a synthesized summary of the top 5 districts by these indicators within the primary age segments based on the notebook's output:

Age Segment 0-29

- Top by % Female:** Mahe, Almora, Kannur, Pathanamthitta, Ratnagiri
- Top by Literacy Rate:** Pathanamthitta, Kottayam, Mahe, Alappuzha, Ernakulam
- Top by % Urban Households:** Chennai, Mumbai, Mumbai Suburban, Hyderabad, Yanam (all 100%)
- Top by % Secondary Education and Above:** Pathanamthitta, Bangalore, Kottayam, Alappuzha, Thrissur

Age Segment 30-49

- Top by % Urban Households:** Chennai, Mumbai, Mumbai Suburban, Hyderabad, Yanam (all 100%)

Age Segment 50+

For other variables within age segments 30-49 and 50+ in the excerpt, but it runs similar analysis for those as well.

Key observations:

- Districts like Mahe, Pathanamthitta, and Kottayam frequently appear at the top for literacy and education levels in the younger age group (0-29).
- Urban megacities like Chennai, Mumbai, Mumbai Suburban, and Hyderabad have 100% urban household percentages across age segments.
- The notebook uses a threshold (>10%) on the age segment population percentage to filter and rank districts.

This structured hierarchical segmentation offers valuable insights for understanding the demographic and educational profiles of districts in India by age group, which can be useful for informed policy making or targeted social programs.

8. MARKET OPPORTUNITY AND STRATEGIC RECOMMENDATIONS:

Here, within specific age segments (0-29 years, 30-49 years, 50+ years) across multiple sub-variables such as percentage of females, literacy rate, urban households, and secondary education and above. It outputs the top 5 districts for each combination and also visualizes these as bar charts.

Here are some of the key insights based on the notebook's outputs for the 0-29 age segment:

- Top districts by percentage female population in this age group include Mahe, Almora, Kannur, Pathanamthitta, and Ratnagiri.
- Top districts by literacy rate in this age group include Pathanamthitta, Kottayam, Mahe, Alappuzha, and Ernakulam.
- Districts with highest percentage of urban households in this segment are Chennai, Mumbai (including suburban), Hyderabad, and Yanam.
- Districts with highest percentage of secondary education and above include Pathanamthitta, Bangalore, Kottayam, Alappuzha, and Thrissur.

This analysis is also repeated for age groups 30-49 and 50+, segmenting by the same sub-variables.

Thus, the notebook presents a comparative segmentation across districts in India considering age-based population percentages and key socioeconomic variables from the 2011 census, highlighting districts leading in female population share, literacy, urbanization, and educational attainment within different age cohorts.

9. LIMITATIONS AND SUGGESTIONS FOR FUTURE WORK:

The attached notebook performs a hierarchical segmentation analysis of the 2011 census data for Indian districts, focusing on three age segments (0-29, 30-49, 50+). It calculates various percentage metrics (female population, literacy rate, urban households, secondary and above education) within

these age segments, ranks districts by these metrics, and then visualizes top districts in each subcategory.

Here is a summary of key top-ranked districts within the Age 0-29 segment as analyzed in the notebook:

- Top 5 districts by Percentage Female within 0-29 age segment:
 - Mahe, Almora, Kannur, Pathanamthitta, Ratnagiri
- Top 5 districts by Literacy Rate within 0-29 age segment:
 - Pathanamthitta, Kottayam, Mahe, Alappuzha, Ernakulam
- Top 5 districts by Percentage Urban Households within 0-29 age segment:
 - Chennai, Mumbai, Mumbai Suburban, Hyderabad, Yanam (all 100% urban households)
- Top 5 districts by Percentage Secondary and Above education within 0-29 age segment:
 - Pathanamthitta, Bangalore, Kottayam, Alappuzha, Thrissur

The notebook similarly performs analyses for the Age 30-49 and Age 50+ segments with the same sub-variables, although the output excerpt here only captures the 0-29 segment results shown.

Visual bar charts plot the top districts by each sub-variable for intuitive comparison.

This analysis can help identify districts with higher female percentages, literacy, urbanization, and education levels in the younger population segment. Similar insights can be drawn for the middle and older age groups by running the same analyses as scripted. This approach supports focused demographic and social development strategy planning by age and other key demographic factors.

10. CONCLUSION:

The conclusion from hierarchical segmentation analysis, particularly in the context of Indian consumer segmentation studies, emphasizes that psychographic variables (such as values, attitudes, social interests, and lifestyle) offer deeper and more distinctive clustering for marketing purposes than traditional demographic segmentation. Different clusters exhibit unique behavioral and psychographic profiles, enabling marketers to target consumer groups more effectively with tailored products and strategies. The studies show that leveraging psychographic segmentation leads to better insights into consumer behavior and purchase decisions, as demographic factors alone are insufficient to capture the diversity among consumers.[journalofbusiness+2](#)

In hierarchical clustering applications, methods like Ward's linkage combined with specific distance measures (such as Mahalanobis distance) have shown to produce optimal clusters for credit customer segmentation, allowing businesses to identify and address groups with varying payment behaviors and financial confidence. These insights help in risk mitigation and targeted interventions to improve credit performance.[jatit](#)

Therefore, the overall conclusion is that hierarchical segmentation, especially when incorporating psychographic data and advanced clustering methods, enhances market segmentation precision and effectiveness in various domains including consumer marketing and financial services.

11. REFERENCES:

1. <https://journalofbusiness.org/index.php/GJMBR/article/download/1436/1343/8403>
 2. <https://ijcrt.org/papers/IJCRT22A6129.pdf>
 3. <https://www.ijrti.org/papers/IJRTI2504078.pdf>
 4. <http://www.jatit.org/volumes/Vol101No2/18Vol101No2.pdf>
 5. <https://jips-k.org/pub-reader/view?doi=10.3745%2FJIPS.04.0100>
 6. https://kalaharijournals.com/resources/FebV7_I2_219.pdf
 7. <https://ijsra.net/sites/default/files/IJSRA-2023-0663.pdf>
- —