EV Market Segmentation Analysis Report

by VISHNUSAI RACHAKATLA

Domain: Electric Vehicle (EV) Sales by Manufacturers and Categories (2015–2024)

---

1. ML Model Application: K-means Clustering

Objective

To segment EV manufacturers based on their sales performance across vehicle categories (2W, 3W, LMV) and identify dominance patterns.

Methodology

- **Dataset**: Sales data for 1,000+ manufacturers across 10 years (2015–2024).
- **Preprocessing**:
  - Handled missing values (e.g., zero sales entries).
  - Normalized sales data to account for scale differences.
- **Algorithm**: **K-means Clustering**
  - **Features**: Yearly sales (2015–2024) for each vehicle category (2W, 3W, LMV).
  - **Clusters**: Grouped manufacturers into clusters based on sales trends.
  - **Optimal K**: Determined using the **Elbow Method** (K=4–6).

Challenges

- **High Sparsity**: Many manufacturers had sporadic sales (e.g., 0 sales in early years).
- **Overlap**: Companies like **Mahindra & Mahindra** and **Piaggio** operated across multiple categories, leading to ambiguous clusters.
- **Noise**: Small manufacturers with inconsistent sales skewed clusters.

Outcome

- Identified **4 clusters**:
  1. **Dominant 2W Players**: High 2W sales (e.g., Bajaj Auto, TVS Motor, Ola Electric).
  2. **3W Specialists**: Focused on 3W (e.g., Piaggio, Mahindra).
  3. **LMV Niche**: LMV-focused (e.g., Hyundai, Tata Motors).
  4. **Multi-Category Players**: Balanced sales across categories (e.g., Greaves Cotton).

---

2. Key Insights & Conclusions

Category Dominance

Category	Top Players	2024 Sales	Trend
2W	Ola Electric, Bajaj Auto, TVS	~550,000 units	Rapid growth post-2020 (EV policy push).
3W	Piaggio, Mahindra	~320,000 units	Steady growth, driven by last-mile logistics.
LMV	Tata Motors, Hyundai	~15,000 units	Emerging premium segment (post-2022).

Emerging Trends

- **Startup Surge**: New entrants like **Ather Energy** (58,547 units in 2024) and **Altigreen** (939 units in 2024) captured niche markets.
- **Decline of Legacy Players**: **Hero Electric** shifted focus from 2W to 3W after 2020.
- **Fragmentation**: 45% of manufacturers fell into the **"Others"** category, indicating a highly fragmented market.

Segmentation Ambiguity

- **Overlap**: Companies like **Greaves Cotton** (2W, 3W, LMV) blurred cluster boundaries.
- **Data Sparsity**: 60% of entries had ≤5 years of sales data, reducing clustering accuracy.

---

3. Proposed Improvements

Enhanced Data Collection

Data Column	Purpose
Geographic Sales Distribution	Identify regional dominance (e.g., 3W in Uttar Pradesh).
Battery Type (Li-ion/Lead-Acid)	Segment by technology adoption.
Pricing Tier (Budget/Premium)	Link pricing to market penetration.
Charging Infrastructure	Correlate infrastructure with EV adoption.

Advanced Models

- **DBSCAN**: To handle noise and outliers (e.g., sporadic sales).

- **Hierarchical Clustering**: For multi-level segmentation (e.g., regional → category).
- **LSTM Networks**: Predict future sales trends using temporal patterns.

Budget Allocation

- **Purchase External Data**:
  - **Government Subsidy Records**: Track policy impact.
  - **Consumer Surveys**: Demographics (age, income) linked to EV preferences.

---

4. Estimated Market Size (2024)

Category	Units Sold (2024)	Key Players
2W	550,000	Ola Electric, Bajaj Auto, TVS, Okinawa
3W	320,000	Piaggio, Mahindra, Mini Metro EV LLP
LMV	15,000	Tata Motors, Hyundai, MG Motor
Total	**885,000**	

---

5. GitHub Repository

🔗 **Code & Analysis**: GitHub Link

- Includes:
  - Jupyter Notebook (K-means implementation).
  - Processed dataset (CSV).
  - Visualization scripts (Matplotlib/Seaborn).

Prepared by: VISHNUSAI RACHAKATLA
Date: 28 APRIL 2025