Market Segmentation Analysis on EV Charging Patterns using Machine Learning

By Riya Maurya

1. Introduction

With the rapid growth of electric vehicle (EV) adoption, understanding user charging behavior is crucial for optimizing infrastructure, energy distribution, and policy planning. The goal of this project is to analyze EV charging patterns and identify distinct user segments using unsupervised machine learning techniques, particularly KMeans clustering.

2. Dataset Description

The dataset used in this project contains detailed information on EV charging sessions. The key features include:

- Battery Capacity (kWh)
- Energy Consumed (kWh)
- Charging Duration (hours)
- Charging Rate (kW)
- Charging Cost (USD)
- State of Charge (Start and End %)
- Distance Driven since last charge (km)
- Temperature (°C)
- Vehicle Age (years)

The dataset has been cleaned and preprocessed for analysis.

3. Methodology

3.1 Data Preprocessing

- Dropped irrelevant or identifier columns
- Checked and confirmed no null values
- Scaled numerical features using `StandardScaler`
- Encoded categorical variables (if present) using one-hot encoding

3.2 Dimensionality Reduction

- Principal Component Analysis (PCA) was applied for visualization purposes
- However, since the first two components explained limited variance, PCA was not used for clustering

3.3 Clustering

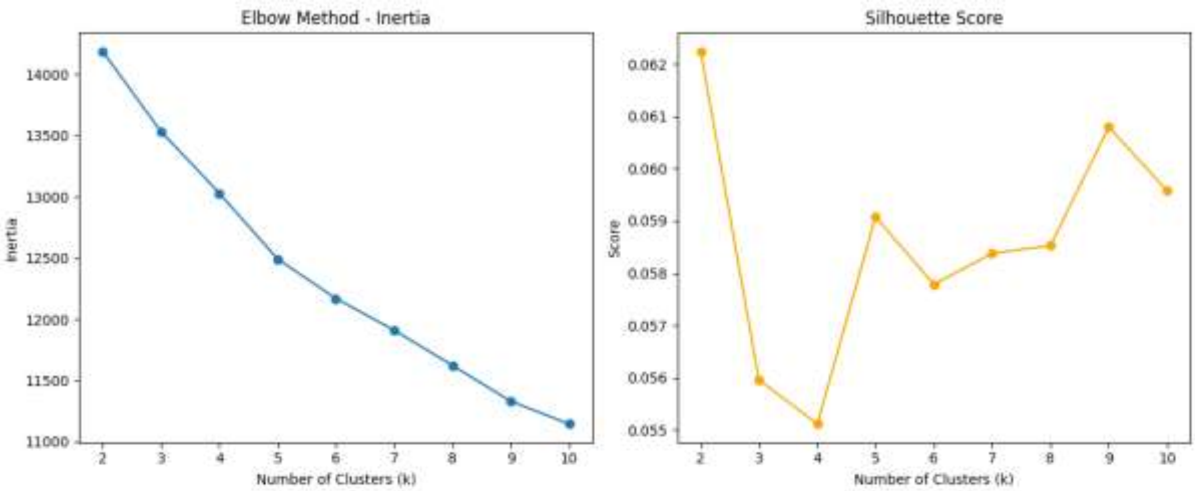
- KMeans clustering was used to segment the dataset

- Elbow Method and Silhouette Score were used to determine the optimal number of clusters
- 5 clusters were selected based on results

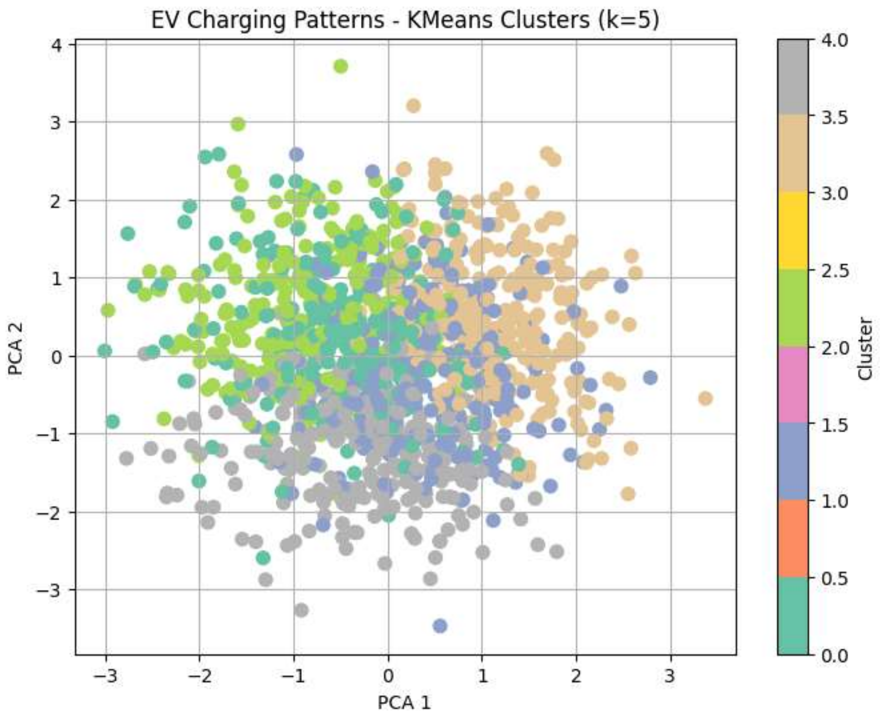
4. Results & Visualizations

Clustering visualizations include:

- Elbow plot to find optimal k and Silhouette score comparison



- PCA scatter plot (for visualization only)



- Cluster-wise average values for all features

Cluster	Battery Capacity (kWh)	Energy Consumed (kWh)	Charging Duration (hours)	Charging Rate (kW)	Charging Cost (USD)	State of Charge (Start %)	State of Charge (End %)	Distance Driven (since last charge) (km)	Temperature (°C)	Vehicle Age (years)
0	89.080660	43.082492	2.068051	32.712746	19.571353	31.109614	75.070703	89.718297	13.676615	3.161247
1	70.240492	38.898326	1.808183	25.226734	17.324348	72.524147	69.896470	143.503221	9.125011	3.004304
2	71.254030	52.396039	3.115557	19.077695	18.481908	42.941744	83.729647	176.133459	7.191841	4.861444
3	71.608572	43.123646	2.082964	23.201145	29.576523	57.326416	75.540930	118.183944	28.318257	4.974694
4	71.226795	36.875683	2.366933	29.369577	26.700722	39.488720	72.448059	240.465648	16.199018	2.107896

5. Insights from Clustering Analysis

1. **Cluster 0 – Fast Chargers, Short Rides**
    - Vehicles in this cluster use **fast chargers** with **high power (avg ~33 kW)** and **short charging times** (~2 hrs).
    - **Lower starting SOC (~31%)** and **high battery capacity (~89 kWh)** suggest quick, frequent charging.
    - Likely used in **urban areas for shorter trips** with newer EVs (~3 years old).

2. **Cluster 1 – Efficient Chargers**
    - These users **start and end charging at similar SOC levels**, indicating more **planned charging patterns**.
    - **Low cost and duration** make it suitable for **personal or fleet vehicles**.
    - Vehicles drive moderate distances (~143 km) and are used in **colder weather**.

3. **Cluster 2 – Long-Distance Users**
    - Users in this cluster travel the **farthest (~176 km)** and charge **slowly** (~19 kW) over **longer durations (~3.1 hrs)**.
    - SOC increases significantly during each session.
    - Indicates **intercity or highway travel** with older vehicles.

4. **Cluster 3 – High Cost in Hot Regions**
    - These EVs are charged in **hot environments (~28°C)** and have the **highest charging costs (~$29.6)**.
    - Moderate SOC changes and distance imply **heavy usage in warm climates**, possibly **commercial or delivery vehicles**.

5. **Cluster 4 – Maximum Distance Drivers**
    - **Drive the longest distances (~240 km)** between charges.
    - Use relatively **fast chargers** and charge for ~2.4 hrs.
    - Likely **ride-sharing** or **long-route service** vehicles with consistent energy needs.

6. Conclusion

The clustering analysis revealed **five distinct types of EV charging behaviors**, each representing different user needs and patterns. These include:

- **Urban fast chargers** who charge quickly and frequently,
- **Planned chargers** with efficient usage,
- **Long-distance travelers** needing longer, slower charges,
- **Cost-intensive users** operating in hot climates, and
- **High-mileage drivers** with consistent, long-range usage.

These insights help us understand that **EV users are not a single group** but have **diverse habits and needs**. By recognizing these patterns:

- Charging infrastructure can be **strategically planned** for different areas,
- **Customized energy plans** can be designed for each user type,
- And EV adoption can be made **more efficient and user-friendly**.

In short, this segmentation provides valuable knowledge for businesses, energy providers, and policymakers to support smarter decisions in the evolving EV ecosystem.

7. Project Repository

Github Link: https://github.com/riya-maurya/EV_Market_Segmentation.git

8. References

- Python Libraries: pandas, matplotlib, seaborn, scikit-learn
- Dataset- kaggle