EV Charging Station Segmentation and Service Analysis

Ömer Kadıoğlu 28812

1. Goal

Clustering: Segment charging stations in Turkey based on the diversity of amenities/services offered.

Regression: Analyze the relationship between charging price and the number of amenities.

Geographical analysis: Explore spatial patterns and highlight disparities in infrastructure.

2. Data Preparation

2.1. Source Files

- ev_charging_stations.json: Structured data from a real CPO, including details about each station, location, amenities, pricing, and more.

2.2. Data Flattening and Selection

- Flattened nested structures (services, stations, sockets) using pandas.
- Extracted amenity fields (parking, restroom, café, market, wifi, playground) to dedicated columns.
- Filtered to stations with at least one amenity for clustering.

2.3. Cleaning and Feature Engineering

- Standardized city names, handled duplicates (e.g., “Istanbul”/“İstanbul”/“istanbul”).
- Marked missing amenity data and flagged incomplete records.
- Created a “num_services” feature (number of distinct amenities at each station).

2.4. Exploratory Highlights

- **Major amenities:** WC and parking present in most stations; WiFi and playground are rare.
- **City distribution:** Stations concentrated in major cities, especially Istanbul, with service diversity highest in metropolitan areas.
- **Missing data:** Roughly 50% of stations lacked detailed amenity info.

2.5. Final Dataset

- 327 rows × 7 features (amenities) + 3 targets (cluster, price, region); ready for clustering and regression.

3. Modeling

3.1. Experimental Setup

- Clustered stations using K-means (k=3) on binary amenity features.
- Used StandardScaler for normalization.
- Analyzed price ("kW_ucret") and amenities relationship with linear regression.

3.2. Cluster Interpretation

- Cluster 0 ("Full-featured"): All amenities present; highest average price.
- Cluster 1 ("Standard"): Most basic amenities, but fewer premium features; moderate price.
- Cluster 2 ("Minimal"): Only restrooms or parking; lowest average price.

3.3. Regression Analysis

- Weak positive correlation between number of amenities and price.
- Linear regression shows price tends to rise with more amenities, but effect size is small.

3.4. Key Insights

- Most Turkish charging stations provide only basic services.
- Service diversity is highest in large urban centers.
- Operators could leverage amenity-rich stations for better customer experience and value proposition

4. Conclusion

This study mapped and segmented the Turkish EV charging network using real-world data, revealing significant variation in both infrastructure and service diversity. Although most stations offer only basic amenities, a minority provide extensive customer-focused features. There is a modest association between price and service richness. Future work could combine usage statistics, external regional data, and user feedback to deepen insights and guide infrastructure improvements.

5. Limitations and Future Work

- **Data completeness:** Many stations lacked detailed amenity information, which may bias clustering.
- **Broader context:** Analysis was limited to one operator and one snapshot in time.
- **Future:** Integrate time-series usage data, compare with other CPOs, and enrich with demographic and traffic data.