

Final Report - EV Clustering Project

Dataset Used: Electric Vehicle Population (Kaggle)

Model Applied: K-Means Clustering

Preprocessing Steps:

- Removed duplicates and filtered years (1990-2025)
- Selected important columns
- Normalized numbers: Electric Range, Base MSRP, Model Year
- Turned Electric Range into categories: Low, Medium, High

Why K-Means?

This model is simple and good for grouping numbers into clusters. Our data is already clean and scaled, so K-Means works well here.

GitHub Link: <https://github.com/dwar32/ev-clustering>

Findings:

We used Elbow Method and found 3 clusters.

- Cluster 0: Newer cars, low price, good range
- Cluster 1: Old or cheap cars with low range
- Cluster 2: Expensive cars with small range (maybe hybrids)

This helps to understand what kind of EVs exist and how they differ by price and performance.