Clustering Based Analysis of EV Adoption in Washington State, USA

Presented By
Shailendra Kumar C S

ABSTRACT

The current work investigates electric vehicle (EV) trends within Washington State from a dataset containing more than 230,000 records. By pre-processing, applying statistical analysis and clustering (K-Means and hierarchical), the study identifies adoption patterns as well as drivers such as MSRP, Electric range and legislative districts. Outcomes enable policy and business planning for EV adoption.

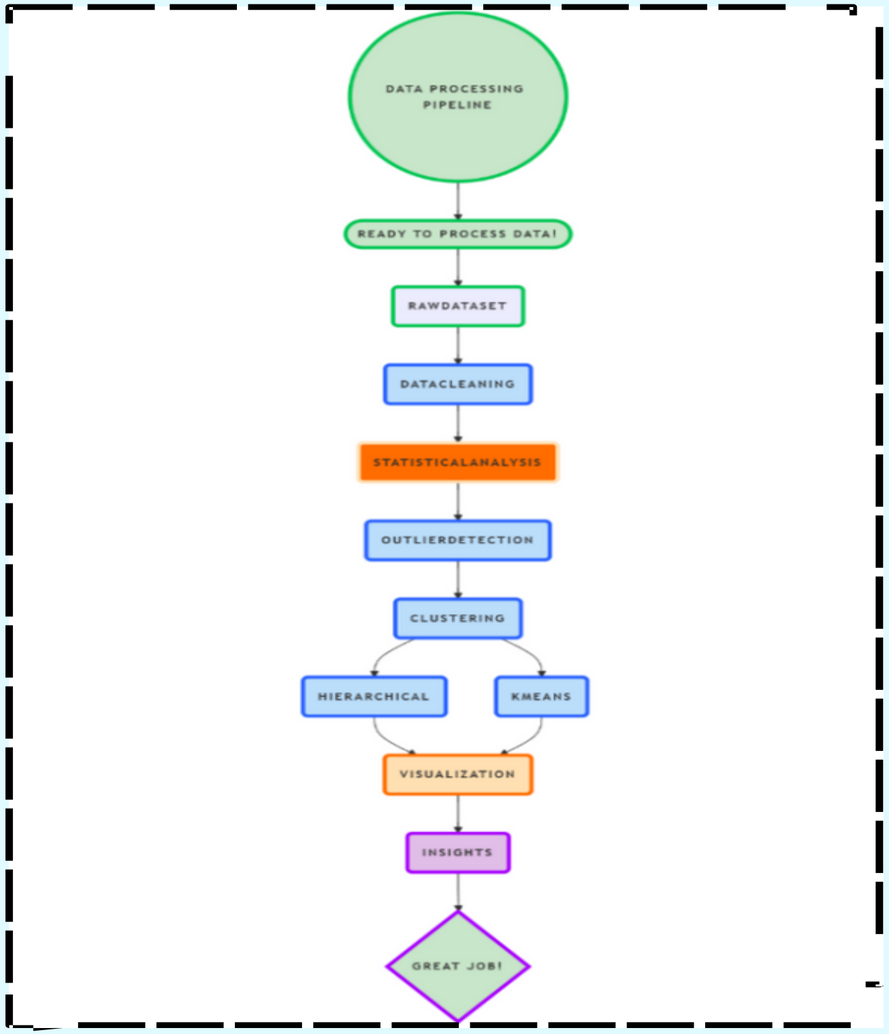
OBJECTIVES

-->Clean and prepare the EV dataset.
-->Analyze key features through EDA.
-->Handle outliers and missing values.
-->Apply clustering to segment EVs.
-->Visualize adoption patterns across time and regions.

PROPOSED METHODOLOGY

- Data Cleaning: Removed missing/zero values in crucial fields.
- Imputation: Applied mean/median filling for gaps.
- Outlier Detection: Used IQR and boxplots.
- Statistical Analysis: Performed skewness, kurtosis, correlation checks.
- Clustering: Applied hierarchical and K-Means clustering.
- Visualization: Created plots with heatmaps, histograms, and pie charts.

ARCHITECTURE DIAGRAM



MODULE DESCRIPTION

Hierarchical Clustering

- Used 1,000 samples with Ward’s method.
- Identified ~4 optimal clusters from dendrogram.

K-Means Clustering

- Based on MSRP and electric range.
- Elbow method confirmed 4 clusters.
- Scatter plot revealed distinct groupings.

DATASET

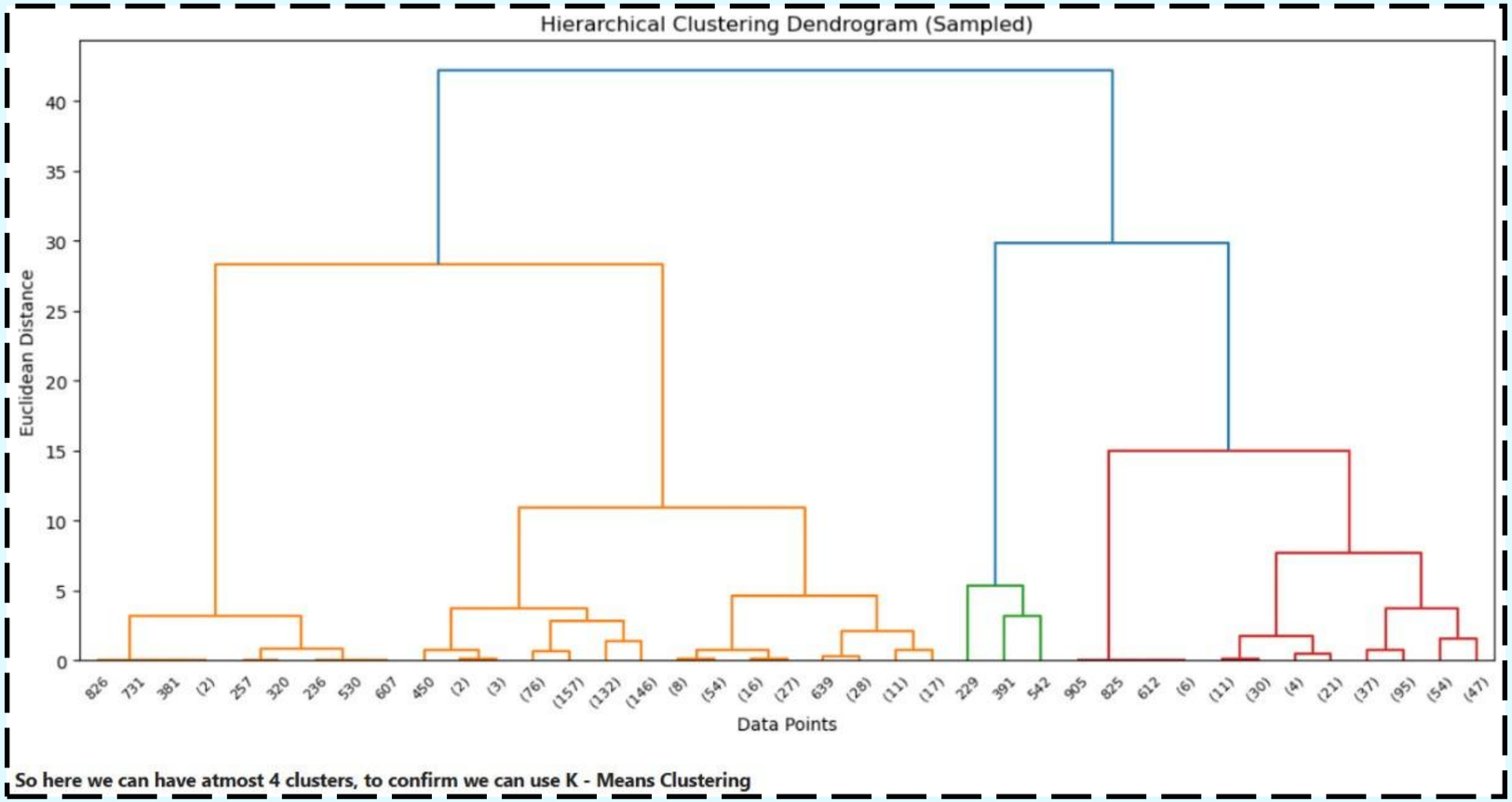
- Dataset Link: https://catalog.data.gov/dataset/electric-vehicle-population-data

- Source: Public data from Washington State
- Size: 235,000 rows × 17 columns
- Important Features: Model year, Make, EV type, Electric range, Base MSRP, County, City, Legislative district

RESULT

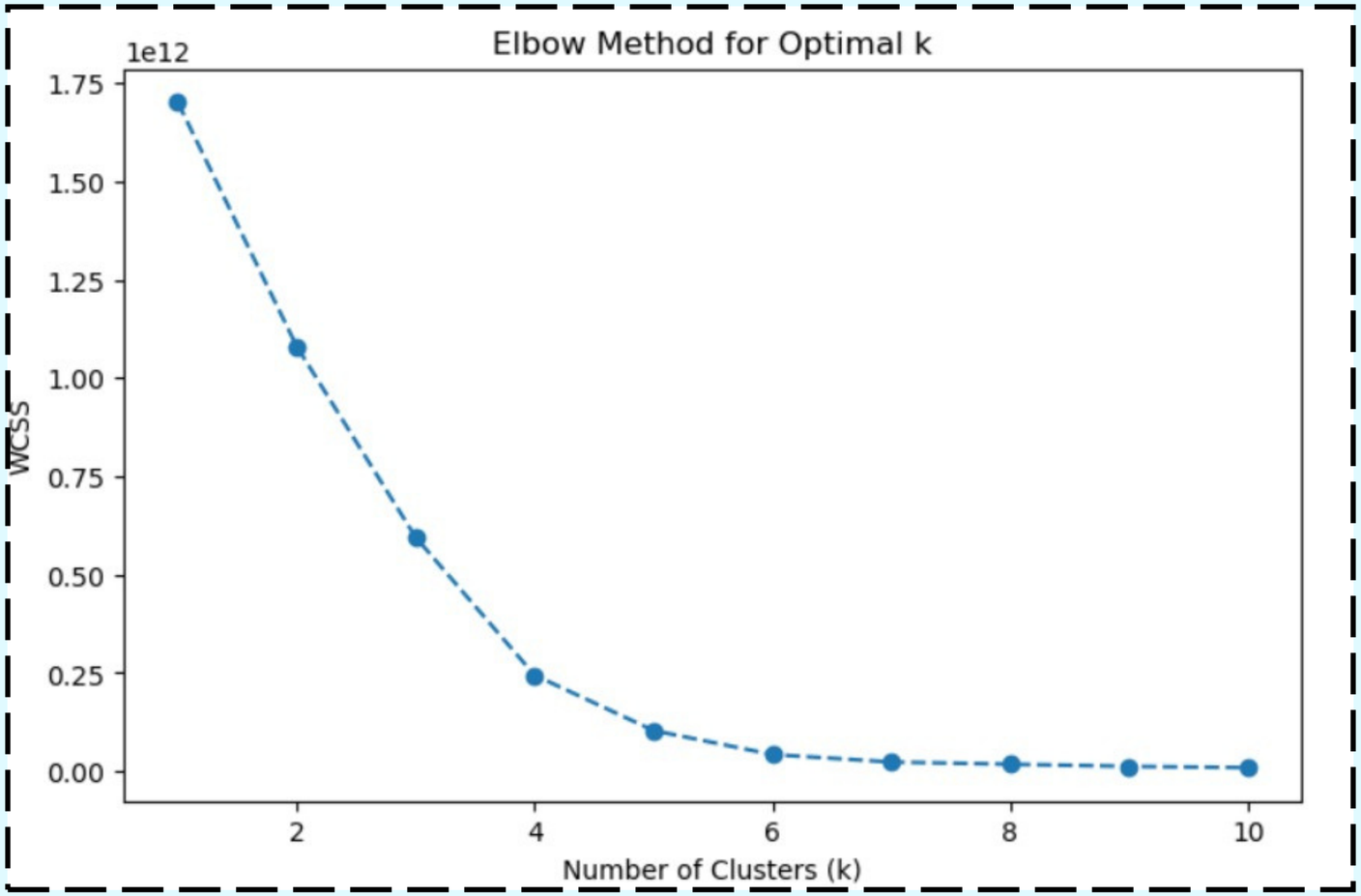
- **Top Counties:** King County leads in EV count.
- **Popular Models:** Tesla Model Y and 3 top the list.
- **Growth Trend:** Surge in EV adoption post-2015, peaking in 2023.
- **EV Type:** BEV and PHEV are evenly distributed.
- **Mileage Leaders:** Jaguar and Tesla offer the longest range.
- **Cluster Summary:**
  - **Cluster 1: Low cost, low range**
  - **Cluster 2: Mid-range**
  - **Cluster 3: High range & cost**
  - **Cluster 4: Luxury outliers**

RESULT

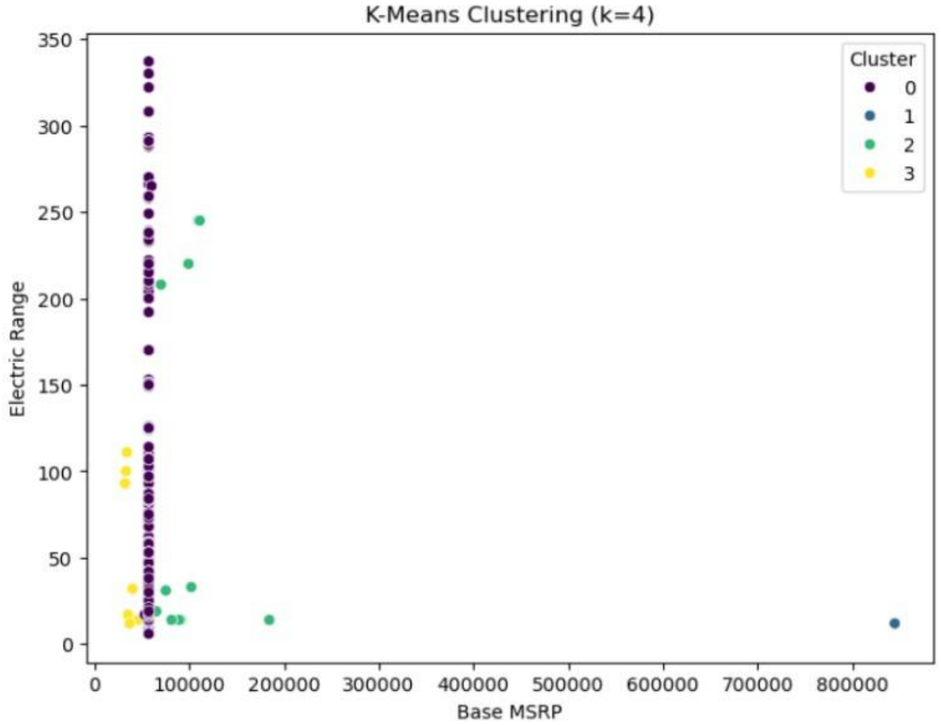


So here we can have atmost 4 clusters, to confirm we can use K - Means Clustering

RESULT


Elbow Method for Optimal k
1e12
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
WCSS
2
4
6
8
10
Number of Clusters (k)

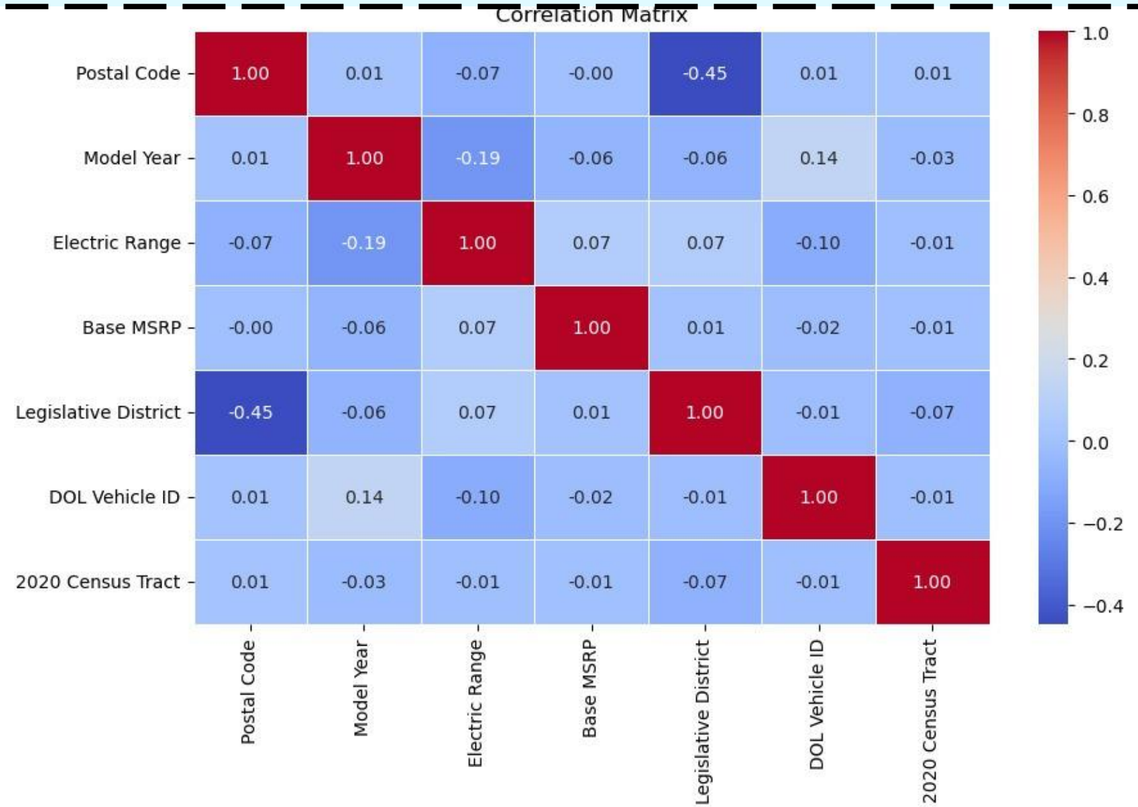
RESULT



WCSS (Within-Cluster Sum of Squares) measures the compactness of the clusters in K-Means clustering. It calculates the sum of squared distances between each data point and the centroid of its assigned cluster.

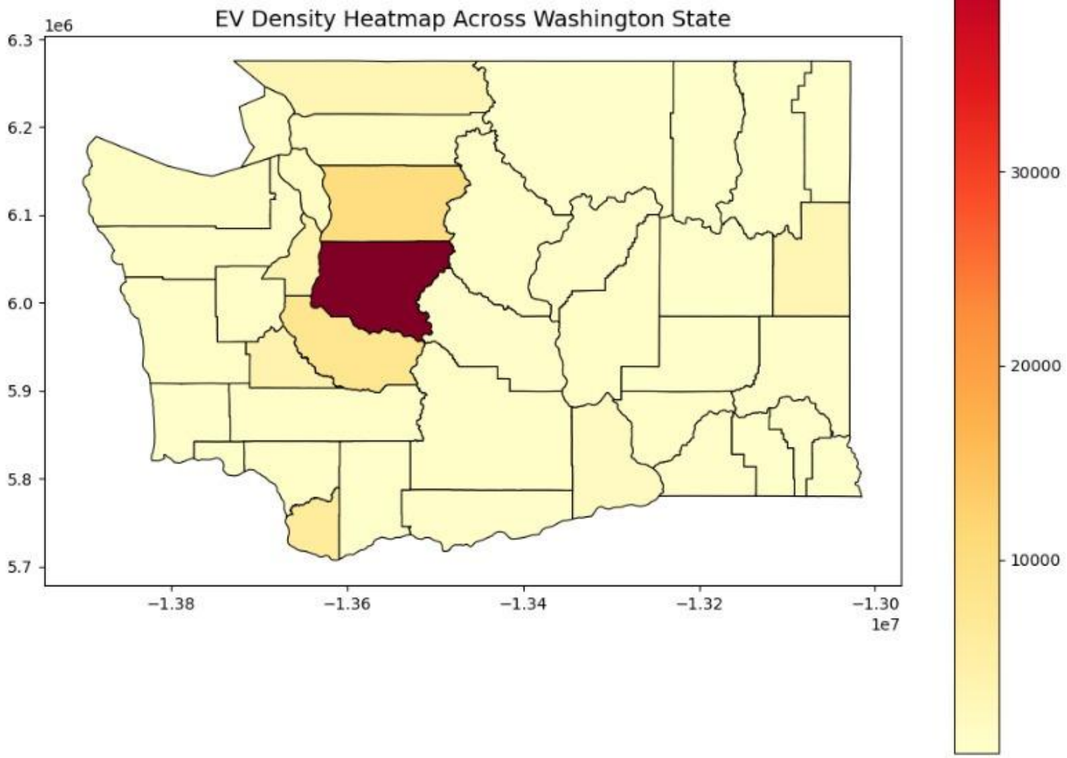
From the K - Means Clustering using the elbow method, we can say that only 4 clusters are possible.

RESULT



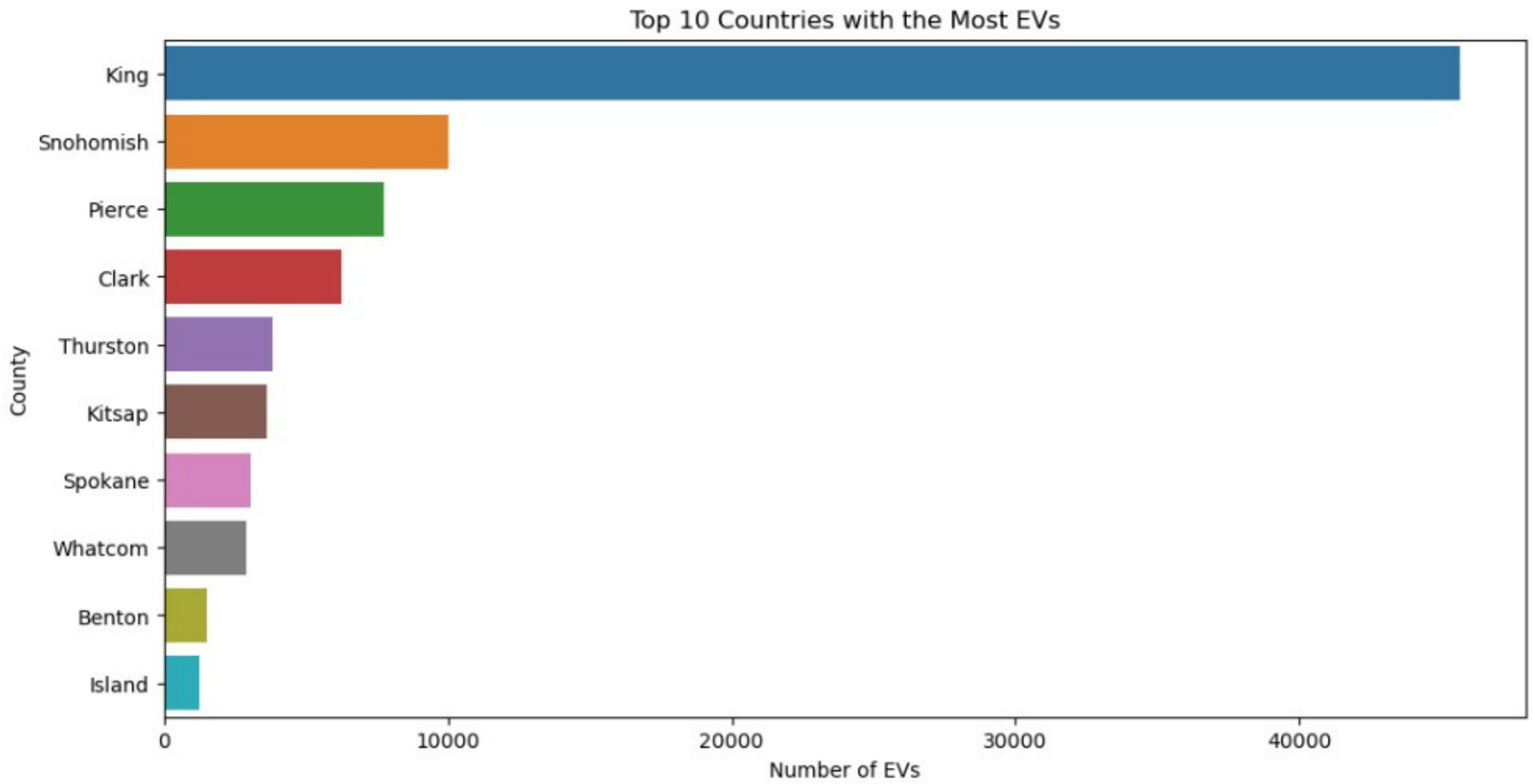
Here in the given Correlation Matrix, None of the Variables are Correlated to each other. But Legislative District and Postol code has somewhat weak negative correlation of -0.45

RESULT



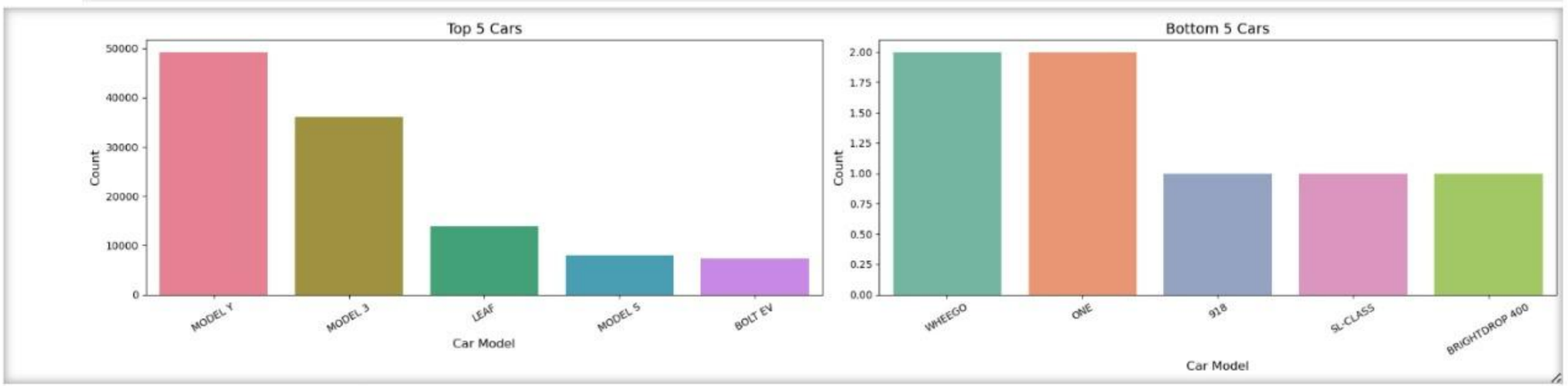
Here we can see that the King State has highest number of Electric Vehicles.

RESULT



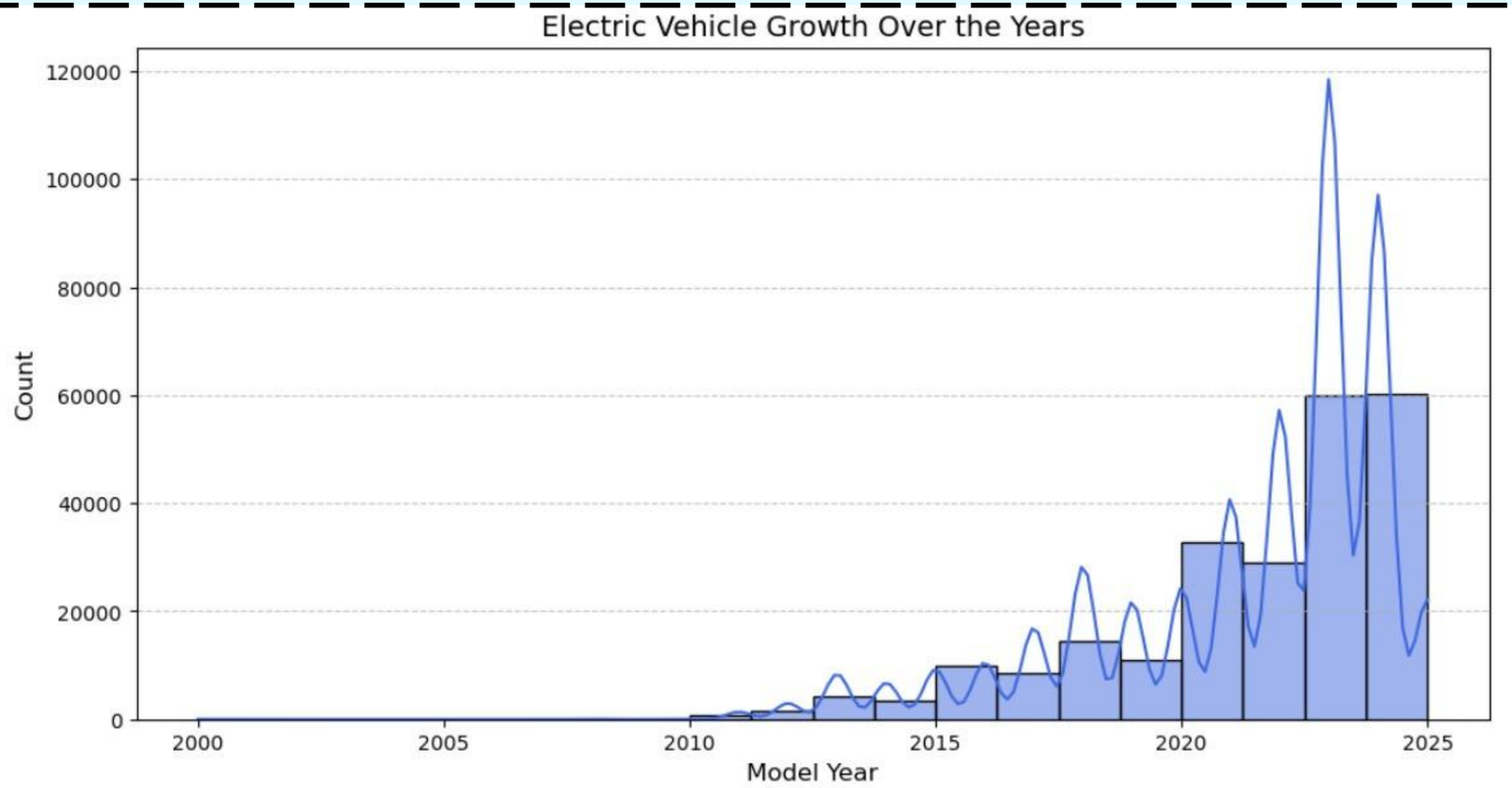
From the given Interpretation we can say that 'King' Country has Highest Number of Electric Vehicles

RESULT



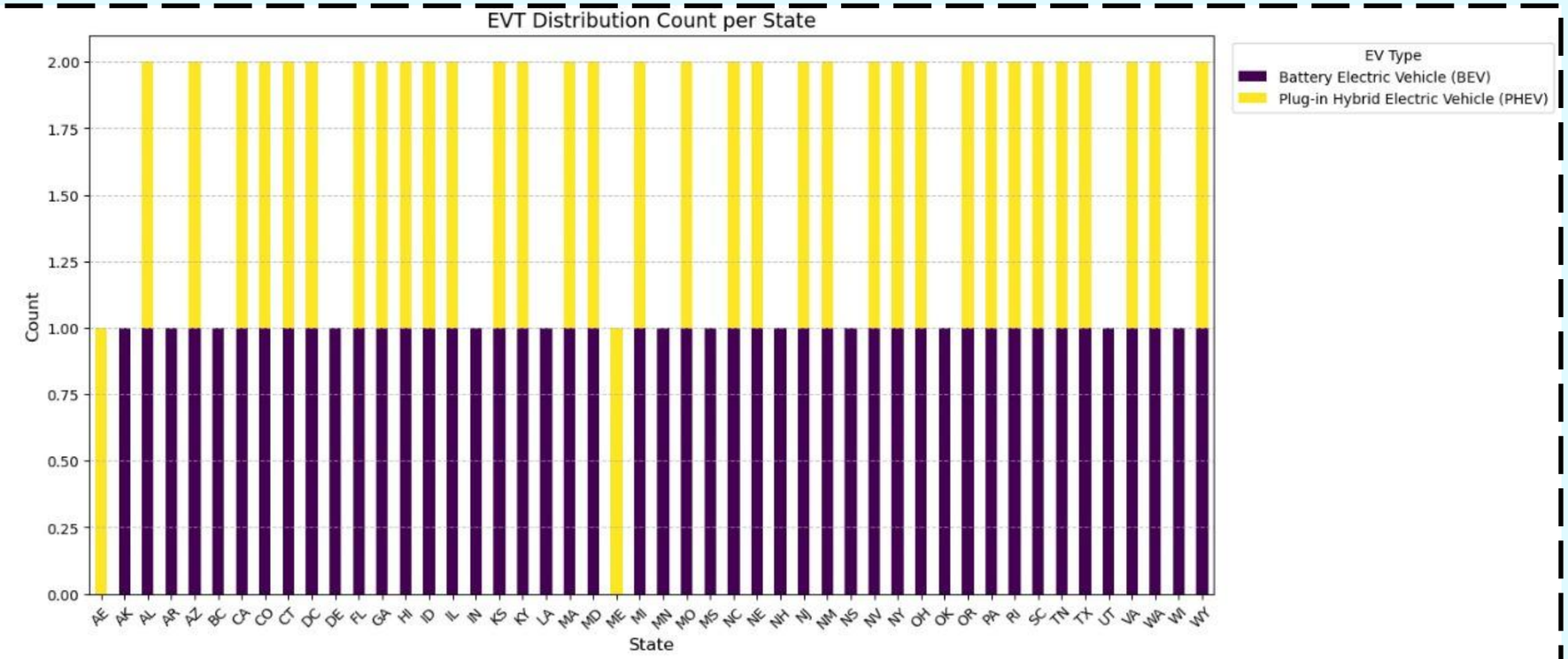
The most popular EVs in Washington are Tesla Model Y and Model 3, with registrations nearing 50,000 and 36,000 respectively. In contrast, rare models like WHEEGO and BRIGHTDROP 400 have minimal presence, with only 1–2 registrations each.

RESULT



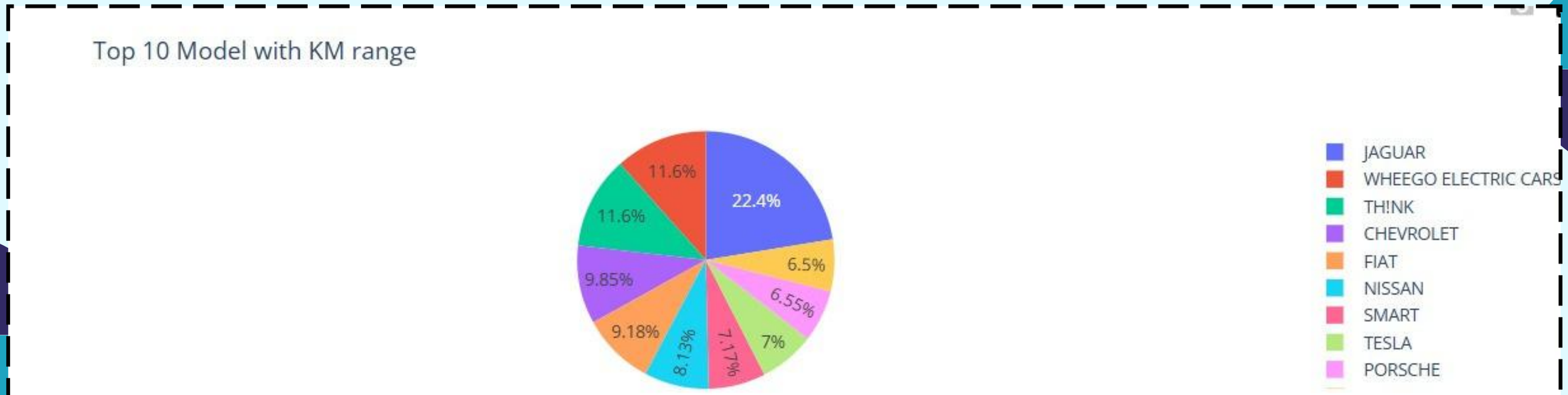
The chart shows a significant rise in electric vehicle counts starting around 2015, with sharp growth peaking between 2022 and 2024. This indicates rapid adoption of EVs in recent years.

RESULT



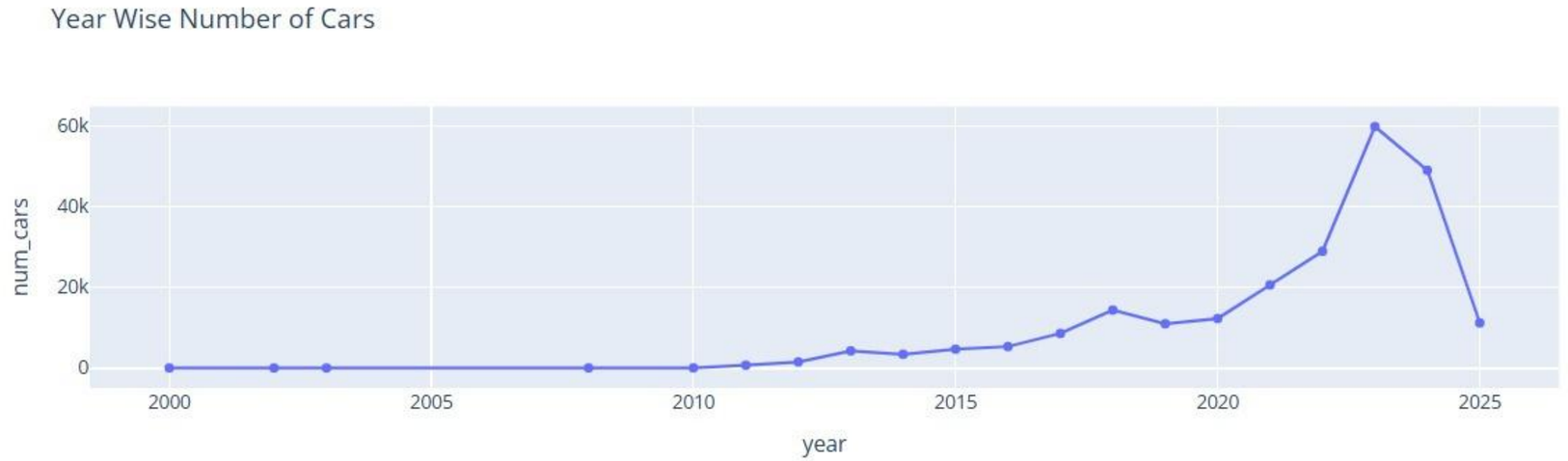
The bar chart shows that every state has 1 count each of Battery Electric Vehicles (BEVs) and Plug-in Hybrid Electric Vehicles (PHEVs), indicating an even distribution. The total count per state is consistently 2, split between the two EV types.

RESULT



Jaguar leads with the largest share at 22.4%, followed by Wheego Electric Cars and TH!NK, each with 11.6%. Other notable brands include Chevrolet, Fiat, and Tesla, each contributing between 6.5% to 10% of the total.

RESULT



The number of cars remained nearly flat from 2000 to around 2010, indicating minimal growth.

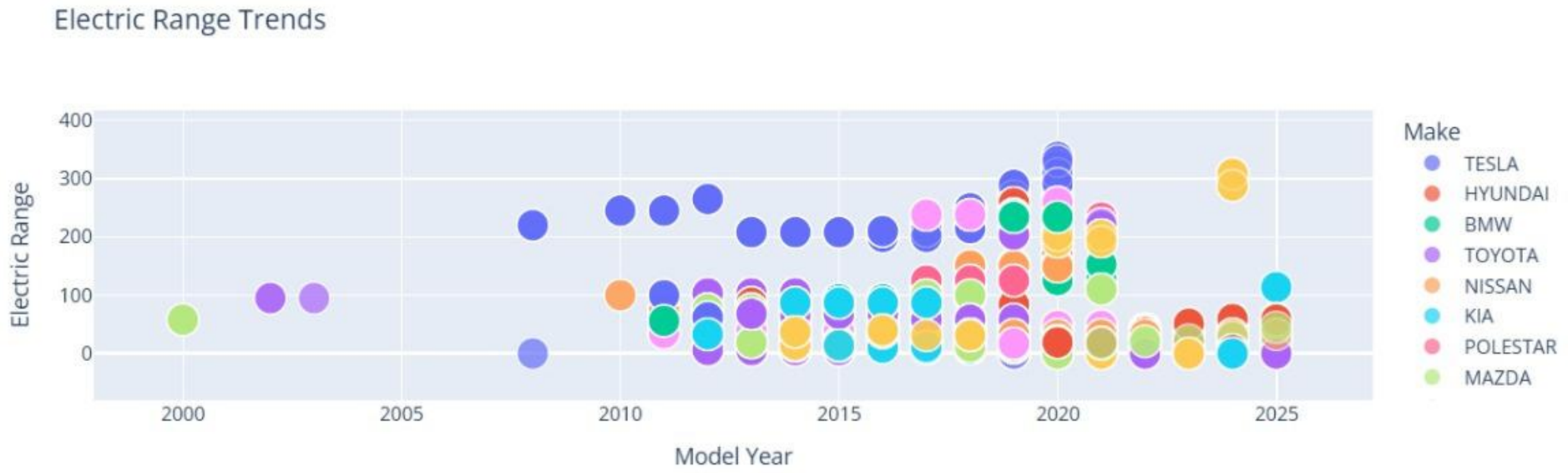
A gradual rise began post-2010, with noticeable acceleration around 2016.

A sharp surge occurred between 2021 and 2023, peaking in 2023 with about 60,000 cars.

There was a drop in 2024, with car numbers falling to around 50,000.

A steep decline followed in 2025, dropping to under 15,000 cars.

RESULT



Electric vehicle range has steadily increased over the years, especially post-2015. Tesla consistently leads with the highest electric range across model years.

CONCLUSION

Clustering revealed distinct EV market segments, supporting targeted strategies for stakeholders. The rise in EVs reflects growing green transport interest. This project demonstrates the impact of EDA and machine learning in policy and market insights.

GITHUB LINK

- https://github.com/RadheKrishna-RadheRadhe/EV_Data-Analysis/tree/main

The Following **GitHub Link Contains** the full **Data Analysis Report** in format of **Python Notebook file (.ipynb)** and **pdf File (.pdf)**, the following link also contains the **shape file (.shp)** of **Washington State.**

REFERENCES

- https://data.wa.gov/Transportation/Electric-Vehicle-Population-Data/f6w7-q2d2/about_data
- https://catalog.data.gov/dataset/electric-vehicle-population-data
- https://sdgresources.relx.com/electric-vehicles-evs#:~:text=EVs%20also%20play%20a%20role,which%20has%20significant%20health%20impacts
- https://sdgresources.relx.com/goal-7-affordable-and-clean-energy
- https://sdgresources.relx.com/goal-11-sustainable-cities-and-communities
- https://sdgresources.relx.com/goal-13-climate-action

THANK YOU