

FINAL DATA ANALYTICS PROJECT REPORT

A Data-Driven Analysis of Electric Vehicle Population in Washington State

Stage 1 – Problem Definition and Dataset Selection

Problem Statement

The adoption of electric vehicles (EVs) is increasing rapidly as part of global efforts to reduce carbon emissions and promote sustainable transportation. Understanding adoption trends, vehicle characteristics, and geographic distribution of electric vehicles is essential for policymakers, utility providers, and urban planners. This project aims to analyze electric vehicle population data from Washington State to identify adoption patterns, technological trends, and regional differences.

Expected Outcome

The expected outcome of this analysis is to generate actionable insights related to electric vehicle growth, consumer preferences, infrastructure needs, and policy impact using data-driven methods.

Dataset Description

The dataset used for this project is the *Electric Vehicle Population Data* sourced from the Washington State Department of Licensing. It contains 271,113 records and 16 features covering vehicle specifications, electric range, CAFV eligibility, manufacturer details, and geographic attributes. The dataset spans model years from 1999 to 2026 and focuses exclusively on Washington State.

Initial Data Exploration

Initial exploratory steps included reviewing the dataset structure, inspecting column data types, checking for missing values and duplicates, and generating summary statistics to understand overall data quality and distribution.

Stage 2 – Data Cleaning and Pre-processing

Missing Value Handling

Minor missing values were observed in geographic and numeric fields. Categorical fields were handled appropriately, and numerical fields were imputed using median values where necessary to preserve data integrity.

Duplicate Handling

Duplicate records were checked, and no significant duplication was found that required removal.

Column Standardization

Column names were standardized by converting them to lowercase and replacing spaces with underscores to ensure consistency and ease of analysis.

Outlier Analysis

Outliers were identified in the electric range variable using boxplots. These outliers represent high-performance electric vehicles and were retained, as they are valid real-world observations.

Skewness Treatment

The electric range variable showed high positive skewness. A logarithmic transformation was applied to reduce skewness while retaining all observations.

Feature Engineering

New features were created to enhance analysis, including vehicle age (derived from model year), electric range categories, and log-transformed electric range. These transformations improved interpretability and supported deeper analysis.

Stage 3 – Exploratory Data Analysis (EDA) and Visualizations

Univariate Analysis

Univariate analysis examined individual variables such as electric vehicle type, electric range distribution, and CAFV eligibility. Results showed that Battery Electric Vehicles dominate the market and most vehicles fall within low to medium electric range categories.

Bivariate Analysis

Bivariate analysis explored relationships between variables such as model year and electric vehicle count, vehicle type and electric range, and CAFV eligibility and adoption. A strong upward trend in EV adoption over time was observed, with newer vehicles offering better electric range.

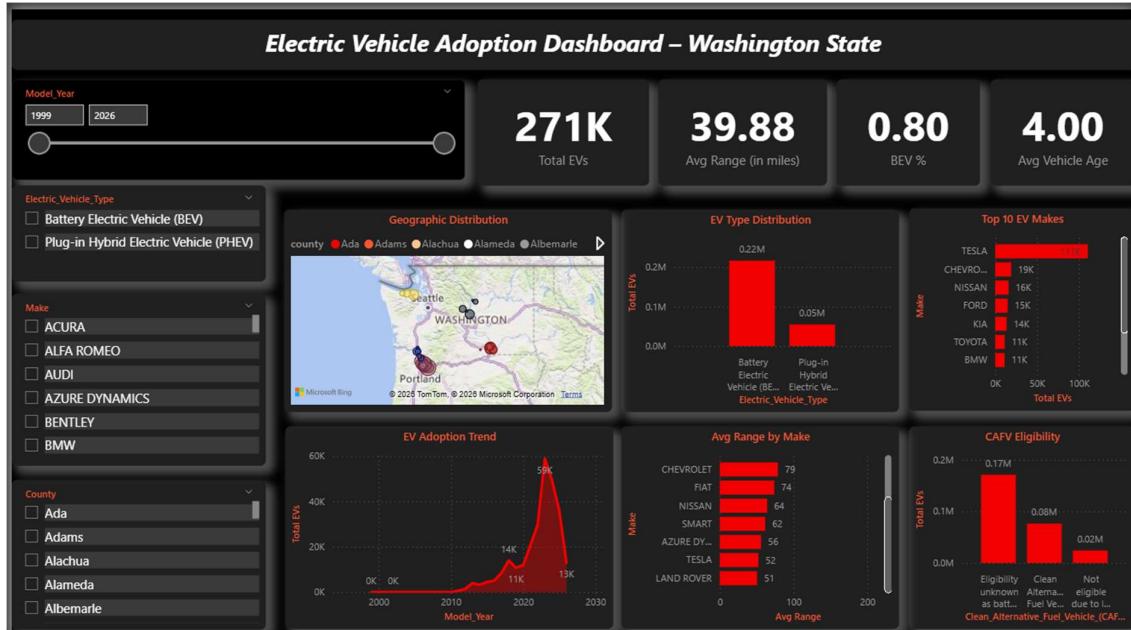
Multivariate Analysis

Multivariate analysis included correlation heatmaps, grouped comparisons, and trend analysis across multiple dimensions. The analysis confirmed that newer vehicle models are associated with higher electric range and lower vehicle age, reflecting technological advancements.

Each visualization was accompanied by interpretation focusing on business meaning rather than just numerical patterns.

Stage 4 – Documentation, Insights, and Presentation

Dashboard



Dashboard Integration (Python + Power BI)

The cleaned dataset prepared in Python was imported into Power BI to create an interactive dashboard. The dashboard integrates KPIs, trend analysis, vehicle type distribution, manufacturer comparison, geographic mapping, and CAFV eligibility analysis. Interactive slicers allow users to explore insights dynamically by model year, county, vehicle type, and manufacturer.

Summary of Findings

Electric vehicle adoption in Washington State has grown rapidly, particularly after 2018. Battery Electric Vehicles dominate the market, indicating strong consumer confidence in fully electric technology. Adoption is concentrated in urban counties where charging infrastructure is more developed. Newer vehicles consistently offer higher electric range due to improvements in battery technology. Policy incentives such as CAFV eligibility significantly contribute to adoption growth.

Key Insights

1. Electric vehicle adoption has increased sharply in recent years.
2. Battery Electric Vehicles account for approximately 80% of total EVs.

3. Urban regions show significantly higher EV adoption than rural areas.
4. Newer vehicle models provide improved electric range.
5. Policy incentives positively influence electric vehicle adoption.

Business and Policy Recommendations

- Expand charging infrastructure in rural and semi-urban regions.
- Strengthen incentive programs for Battery Electric Vehicles.
- Promote adoption of newer high-range electric vehicle models.
- Use county-level data to guide infrastructure investment.
- Collaborate with utility providers to ensure grid readiness for increased EV demand.

Final Story Using the Dashboard

The Power BI dashboard presents a complete narrative of electric vehicle adoption in Washington State. It begins with high-level KPIs summarizing adoption scale, followed by trend analysis highlighting rapid growth. Distribution charts show strong preference for Battery Electric Vehicles, while geographic visuals reveal regional adoption patterns. Manufacturer-level analysis demonstrates technological competition, and CAFV eligibility confirms the role of policy incentives. Together, the dashboard explains how technology, infrastructure, and policy drive the transition toward sustainable transportation.

Conclusion

This project demonstrates end-to-end data analytics capability, from data cleaning and exploratory analysis using Python to interactive visualization and business storytelling using Power BI. The insights derived from this analysis support informed decision-making for sustainable transportation planning and electric vehicle infrastructure development.

END OF REPORT
