

# Analysis Report

## 1. Introduction

The report delves into an analysis of a dataset containing information on electric vehicle purchases in Washington State. The primary objective is to understand the dataset's characteristics, perform exploratory data analysis (EDA), and build a classification model for predicting Clean Alternative Fuel Vehicle (CAFV) eligibility.

## 2. Methodology

### Exploratory Data Analysis (EDA)

Descriptive Statistics:

- Summary Statistics:
- The dataset comprises 159,467 entries, showcasing various attributes related to electric vehicle purchases.
- Key statistical insights across columns such as Model Year, Electric Range, Base MSRP, Legislative District, DOL Vehicle ID, and 2020 Census Tract.

Data Distribution and Relationships:

- Visualizations depicting the distribution of Electric Vehicle Types, relationships between Model Year and Electric Range, among others.

### Classification Modeling

Data Preparation:

- Selection of Features:
- Identified relevant features including Model Year, Electric Range, and others for predicting CAFV eligibility.

Model Development:

- Utilized a HistGradientBoostingClassifier model from Scikit-learn for classification.

Model Evaluation:

- Achieved 100% accuracy on the test dataset for CAFV eligibility prediction.
- Presented precision, recall, and F1-score metrics for each class.

## 3. Findings

Model Year Distribution:

- Majority of purchases occurred in recent years, indicating an increasing trend in electric vehicle adoption.

Electric Range and Base MSRP:

- Wide variability observed, reflecting diverse consumer preferences and pricing structures.

Geographical Representation:

- Legislative Districts and Census Tracts demonstrate diverse representation across regions.

Model Performance

- The classification model displayed exceptional accuracy (100%) in predicting CAFV eligibility across all categories.

## **4. Key Business Insights**

### **Trends in Electric Vehicle Adoption:**

- Notable surge in electric vehicle purchases in recent years, signifying an escalating trend in adoption within Washington State.

### **Consumer Preferences and Pricing:**

- Wide variability in Electric Range and Base MSRP suggests diverse consumer preferences and pricing structures. This variability underscores the importance of catering to varied consumer needs in the market.

### **Predictive Model Performance:**

- The high accuracy of the classification model in predicting CAFV eligibility categories indicates its potential applicability for stakeholders in evaluating vehicle eligibility accurately.

## **5. Conclusion**

The analysis provided valuable insights into electric vehicle purchases in Washington State. The EDA highlighted trends in vehicle attributes and geographical representation, while the classification model exhibited excellent accuracy in predicting CAFV eligibility. Further investigation into influential factors and model validation could offer deeper insights beneficial for stakeholders in the electric vehicle domain.