The Future of Electric Vehicle: Exploration of US EV Adoption

Team 14: Charlotte Chatfield, Christina Ho, Cindy Jin, Laurné Jones, Princess Adentan, Rime Jebai, Sophia Zhang

Highlights

- **Rich Data:** By collecting and merging datasets from different sources, we build a unique and robust data set
- **Careful Cleaning:** By feature engineering and data balancing, we get high quality data to train the model
- **Best Model:** By building and comparing different classification models, we choose the best performing model
- **Meaningful Insights:** Based on data analysis, we identify a target customer profile most likely to adopt EVs
- **Recommendations:** Based on the insights, we offer recommendations to public and private stakeholders

Background

Interest in electric vehicles (EVs) is growing given the potential environmental impact from Greenhouse Gas (GHG) emission reduction; however, US consumers have been slow to adopt EVs. This project predicts the consumers who are most likely to purchase an electric vehicle, with the goal of (1) helping companies identify their target customer base and make strategic business decisions and (2) providing recommendations to policy makers on necessary EV infrastructure development.

Data Sources and Wrangling

What's most important in predicting who will adopt an EV? Are Demographic attributes good predictors or are other variables (Geographic, Economic) more important?

2017 National Household Travel Survey (NHTS) ⟷ Merge data on **STATE** variable ⟷ Alternative Fuels Data Center (AFDC)

Fill missing values

Feature Engineering
- Drop feature with high multi-colinearity
- Drop features similar with each other
- Drop features with mainly invalid responses

One hot encoding

Data Imbalance Treatment with SMOTE

Target Label
EV or not

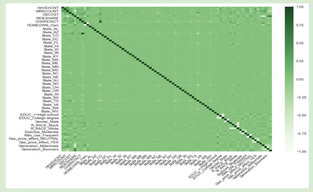
Demographic Features
- Gender
- Generation
- Education
- Home Ownership
- ...

Geographic Features
- State
- Charging Station Count
- City Population
- ...

Economic Features
- Income
- Gas Price
- ...

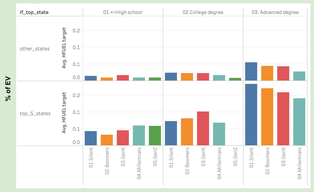
Exploratory Data Analysis

For the exploratory analysis, we segmented the data into 5 states with the highest rates of EV adoption and other states.



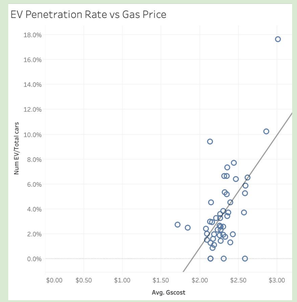
Correlation Coefficient Heatmap

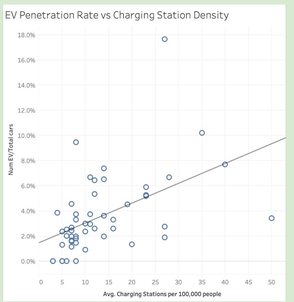
Selected features do not indicate a collinear relationship with any of the other features.



Education levels and generation

Among advanced degree holders, the older generations are more likely to adopt an EV.





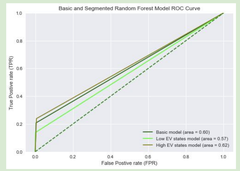
Regression Analysis

Graph 1:
EV adoption rate increases as the average charging station density increases.

Graph 2:
EV adoption rate increases as the average price of gas increases.

Models and Key Insights

We built several machine learning models including Logistic Regression, Naïve Bayes, SVM, Decision Tree, and Random Forest.



Model Performance: Random Forest performs best. AUC value 0.6. Accuracy rate 97%.

All States	Top 5 States	Other States
Gas cost	**Gas cost**	Income
Charging stations	City population	**Gas cost**
City population	Income	#Household vehicles

Feature Importance: Geographic & Economic features are far more predictive than Demographic features.

Key Insights: Individuals with the following features are more likely to adopt an EV:

- High gas price
- High population cities
- High household income
- High charging station density

Recommendations

Target customer profile is high income families in the most highly populated cities in states with high gas prices + moderate EV penetration.

10 Priority Cities: *Baltimore, Columbia, Phoenix, Tucson, Burlington, Essex, Newark, Jersey City, New York City and Hempstead*

- **Supply-side | Manufacturers + Dealers**: Concentrate EV distribution in 10 priority cities to increase supply to target customer
- **Demand-side | Manufacturers + Dealers + Financial Institutions**: Offer financial incentives to encourage high income families in 10 priority cities to adopt EVs
- **Ecosystem | Government**: Develop strategies/incentives to accelerate charging station network density in 10 priority cities