Megha Jain

Thomas Sheridan

Brandon Bennett

EVs by state

California (39m) – 2.5m EVs

Texas (29m) - 1.02m EVs

Michigan (10m) 42k

Florida (21m) 291k EVs

NY (19m) 1.03m

NJ (9.5m) 304k

\*\*(Total Pop of 6 States 127.5 of 332m) (38% of US)

EV Charging Stations by State (Public/Electric)

California – 14,983

Texas – 2,573

Florida – 2,878

Michigan – 1,108

NY – 3,458

NJ - 915

Make/Model (1,214,000 EVs between those 6 States)

/////

Slide Notes

**Intro**:

Good Evening Everyone,

For the Team Electric project that Megha, Brandon, and I are worked on, we chose to investigate whether the average per capita income in each county connects with the sales of electric vehicles as well as whether availability to charging stations in a given location.

To do so we pulled three primary sources of data and manipulated them to fit our needs and applied them within our machine learning model.

Data Sources

State EV Registration Data : Per State EV Data

*This data was pulled from Atlas EV Hub, 6 states (NY, NJ, Cali, Texas, Michigan& FL,* *\*\*(Total Pop of 6 States 127.5 of 332m) (38% of US ) (& 1.59% of the world’s population.)*

Country Wide Fueling Data : Alternative Fueling Station Data

*This data set was pulled from the US Department of Transportation*

Personal/Household Income by county : County Income

*This data was pulled from the Bureau of Economic Commerce, US Dept of Commerce*

Data Exploration

We used Jupiter Notebook to explore our data,

**Understanding the relationship between county income, the number of EVs, and the number of charging stations in that area is the goal of this data exploration.**

**We'll start by utilizing a scatter plot to visualize the data, which can help us see any patterns or trends.**

**In order to analyze the relationship between the two variables and estimate the number of EVs in a particular county based on the income level, we also employed multiple regression analysis.**

**Using the multiple regression model with K means, we also looked at the association between county income, the number of EVs, and the number of charging stations in a county.**

SQL Database connections

Our project was built around data relationships, and we conceptually represented our data using the QuickBase relationship Diagram. State and County served as our keys, connecting them across various data sources.

This data is hosted Postgres which we hosted on AWS’ cloud.

ETL Slide

Jupyter Notebooks were used to extract, transform and load the data from CSV files downloaded from various sites to the AWS postgres database

Data was read in Pandas dataframes, functions dtypes and value\_counts() were used for initial exploration

Certain states data was merged from two CSV files into one dataframe for EV\_registration DB table

Filtered rows based on certain values

Unwanted columns were dropped

Rows with NaN were dropped

Column with state code was added

Dataframe columns names were renamed

Datatype of string to datetime was done for certain columns

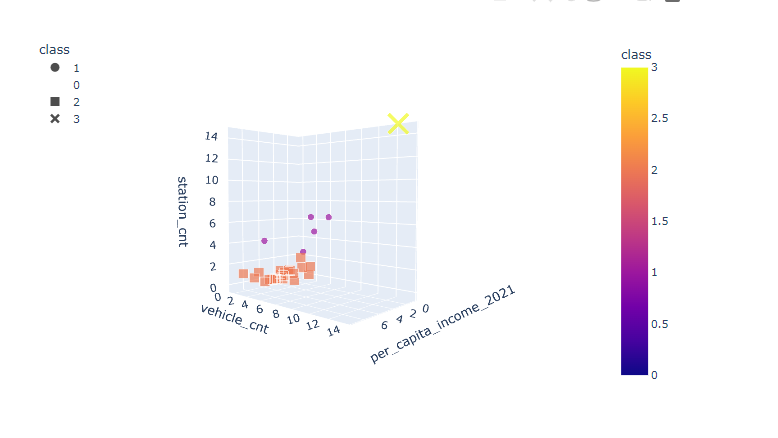
Two columns were merged into a single column for certain datasets

Library Psycopgy2 and sqlalchemy’s create\_engine function was used to insert DF data into DB tables

SQL queries were used to clean and manipulate data further in the DB tables such as removal of null or unknown values and so on

County table data was loaded using PG admin’s load utility and then cleaned using SQL queries

Machine Learning



**Ending**

Due to a number of factors, counties with greater incomes may be more likely to experience higher EV sales.

People with greater incomes may be better able to afford the higher upfront cost of EVs.

People may find it simpler to acquire and operate EVs in higher income regions because there may be more charging facilities in those areas.

Due to variables like more access to information and higher levels of education, higher income counties may have a stronger awareness of and acceptance for electric vehicles.

**Market modeling** entails creating a simulation of the EV market based on variables including consumer demand, rivalry, and prevailing conditions.

**Survey-based strategies**: These entail gathering customer information through surveys or focus groups and using that data to forecast future EV sales.

**Expert judgment**: This entails asking professionals in the subject, such as researchers or industry analysts, their expectations about EV sales.

The cost and accessibility of charging infrastructure, the availability and price of gasoline, and government subsidies for EVs are just a few of the numerous other variables that might have an impact on EV sales.

In order to determine strategies for boosting EV adoption across all income levels and to fully grasp the precise relationship between county income and EV sales, more research is required.