**PROJECT REPORT**

**Project Proposal**

We would like to understand the relationship between fuel price and the proportion of each US State’s privately owned electric cars. Additionally, we would like to find the correlation of the registration of electric vehicles by State, population, and income. Our hypothesis is that gas prices play a large part in influencing vehicle fuel type choice, and because of this, states with higher average fuel prices will also have higher proportions of privately owned electric vehicles making up their privately owned car population.

**Data Sources**

<https://afdc.energy.gov> – From this website, data for electric vehicles registered per State for the year of December … was extracted.

<https://gasprices.aaa.com/state-gas-price-averages/> - from this website, the average gas prices per State was extracted.

[US States - Ranked by Population 2020 (worldpopulationreview.com)](https://worldpopulationreview.com/states) – from this website, the population per State for the year 2020 was extracted.

[Median Household Income by State 2020 (worldpopulationreview.com)](https://worldpopulationreview.com/state-rankings/median-household-income-by-state) – the average population income for the year 2020 was extracted from this website.

**How data was formatted**

Data (the gas prices and the electric vehicle registration) was downloaded and a clean copy of data was copied in another excel document and saved as csv.

State population and income were extracted as csv files.

In regard to having the predefine structure of the documents in our collections by using SQL, we believe that this saved us a substantial amount of time because we were able to initialize the fields within the documents we were creating from our data as we wrote the code in our Jupyter notebook and what the schemas of each table would be after we used SQL Postgres

**Extract transform load (ETL)**

We built four datasets from four sources of data so we could better explore our question. We used one data source to build our EV analysis dataset. Data extract, transform and load were executed in EV\_Registration\_By\_State\_Anlysis Jupyter notebook. Work was done using Python its Pandas library and PostgreSQL.

We download the Electric Vehicle Registration Counts by State dataset from the U.S. Department of Energy website (ttps://afdc.energy.gov).