# Project Proposal Document:

**Team:** Elie Hakim, Miguel Patxot

**Proposal Date:** 11/2/2019  
**Due Date:** 11/6/19 or 11/7/19

**Statement of work:**

We’re looking to observe the trends of fuel-ups on vehicles based on fuel ups from fuelly community online. With a given sample we’re going to dive into important questions:

What are Average Fill Ups looking like?

What Type of Vehicles are Consumers actively driving?

Are gas prices driving what type of cars we tend to be driving?

Do we believe people are more concerned with bigger cars? Smaller cars?

**Data sources**: (include details and why pick these data sources?)

Data from Fuelly website, and EPA Dataset on specifications of vehicles.

Further work: Gasfinder type of API that grabs prices of gas prices in specific area. Choose area for analysis with some of the top vehicles driven.

**Proposed ETL:**

**Extract –**

Webscrape Active Fuel Up Data from Fuelly website with regards to US drivers.

Utilize EPA Dataset to provide a baseline of metrics.

Develop schema for database tables.

**Transform-**

Normalize Data from Webscrape to reflect similar information to EPA dataset with Reflecting Make/Models/MPG

**Load-**

Loading of Datasets into PostgreSQL.

Reflecting Keys to associate makes with makes in EPA dataset.

Proposed Final Schema:

PostgreSQL

**Approved by: Satish Anthony on 11/02/2019.**

**Final Project Report (11/6/19 or 11/7/19)**

At about 8 PM, your team will submit a Final Report that describes the following:

* **E**xtract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).
* **T**ransform: what data cleaning or transformation was required.
* **L**oad: the final database, tables/collections, and why this was chosen.

Please upload the report to Github and submit a link to Bootcampspot.

Present 3-4 minutes on the project discussing some pain points and how did you resolve them. Only one student from each team should present. 1 min for any Q&A to the class.