# Final Project

Paige Wiaduck

# What environmental challenge is the startup solving?

The environmental startup will solve the issue of unsustainable transportation by focusing on the distribution of alternative fueling stations across the United States compared to the number of driving trips taken by citizens. Exploring this will provide guidance for policymakers and government officials to determine the ideal locations for alternative fueling centers.

My hypothesis is that locations with higher number of driving trips will have more alternative fueling stations.

# Why do we need data to solve this problem?

Data is essential to solving the problem because it offers information about the number and locations of different alternative fueling stations. Datasets containing the coordinates of the stations and the numerical count of driving trips taken by citizens can help identify patterns. This will help combat the reliance on non-renewable resources for transportation by identifying what population is best to switch to electric vehicles.

# What data will you use to solve this problem?

The potential data the environmental startup will use is:

Trips by Distance per state

Alternative Fueling Stations in the US

EV Data 1997-2024

# Which two datasets did you select? Or will you use all three?

From the potential datasets, I chose to focus on two:

Trips by Distance per state

Alternative Fueling Stations in the US

These datasets will provide information to analyze if areas with more trips have more alternative fueling stations for electric vehicles.

# Which type of analysis will you do?

In this project I will complete a exploratory analysis regarding the state distribution of alternative fueling stations and number of driving trips citizens take.

# What type of data have you collected?

The data I have collected is text, numerical, and geospatial data. Some of the textual data represents dates, so I will convert them into datetime objects for analysis.

Most of the numerical data refers to the number of individual driving trips and the coordinates of the alternative fueling stations. The textual data mostly represents characteristics such as state and city.

# How big are the datasets?

The dataset containing trips by distance per state is about 483 MB with 5,689,234 rows and 22 columns.

The dataset with the alternative fueling stations in the UD is about 29.5 MB with 78,865 rows and 79 rows.

Since these datasets are so large, I will filter that data to only contain information about the current year, 2023.

# What is the conclusion of your analysis (why would anyone care)?

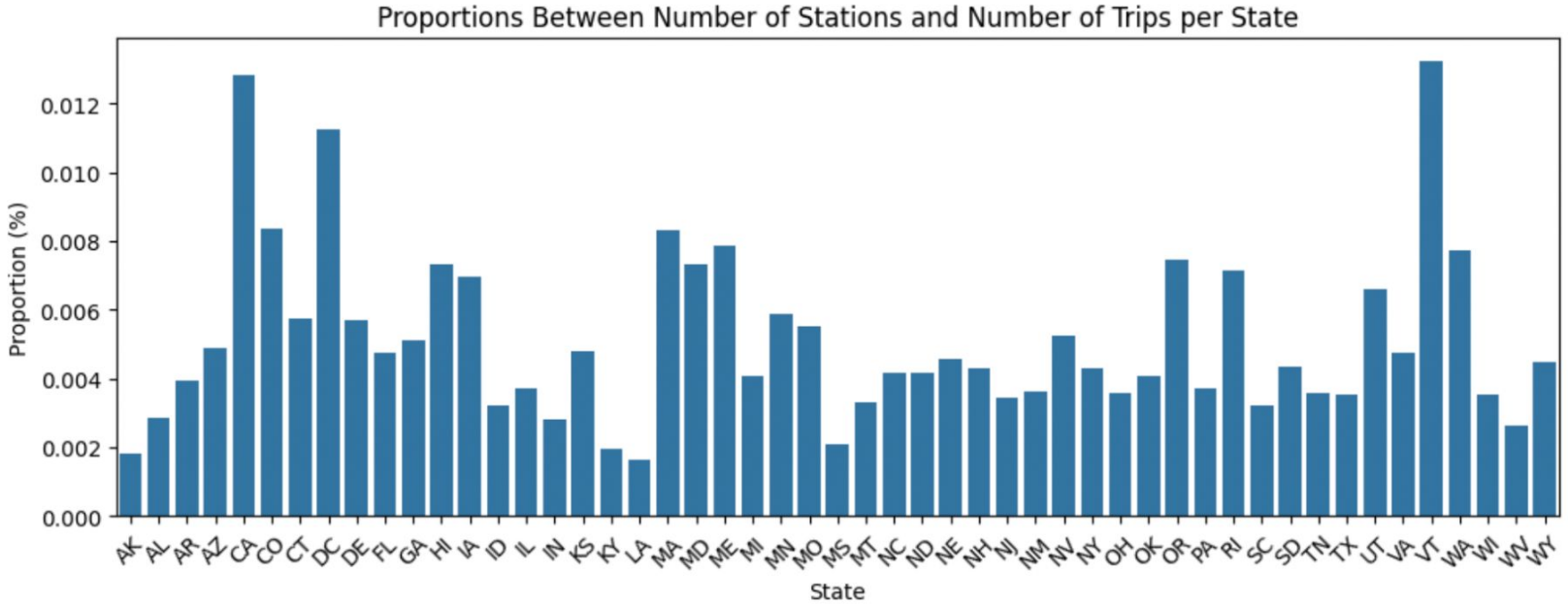
The conclusion of my analysis is that alternative fueling stations are more popular in areas with higher number of driving trips. This was proven by California being the state with the greatest count of driving trips and alternative fueling stations as of 2023. The data also shows that the major cities of California are the ones with the most stations too.

This research helps identify areas that lack fueling stations for electric vehicles, thus still relying on non-renewable transportation. This information is essential for policymakers and government officials to determine locations that need improvement with eco-friendly transportation.

# Applied Skills

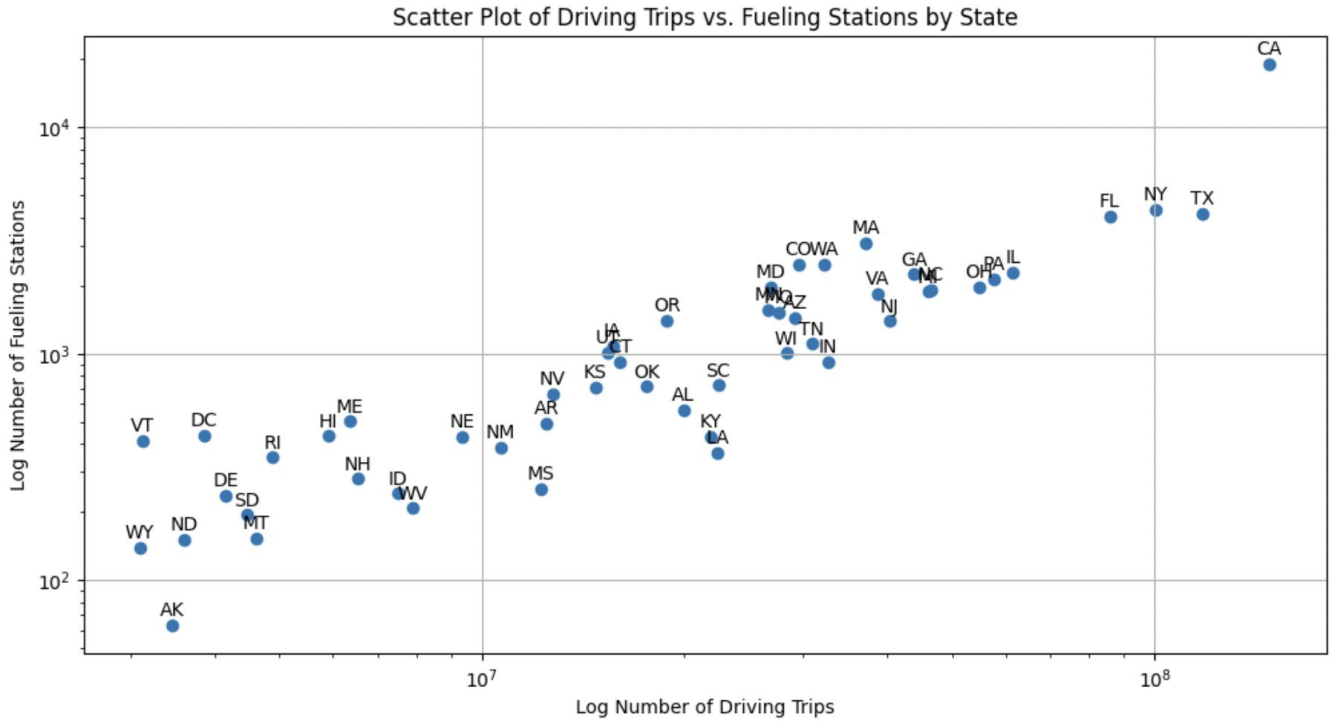
1. Groupby function
2. Merge
3. Apply
4. Time series
5. Plotting with Seaborn and Matplotlib
6. Plotting geospatial data

# Visualizations

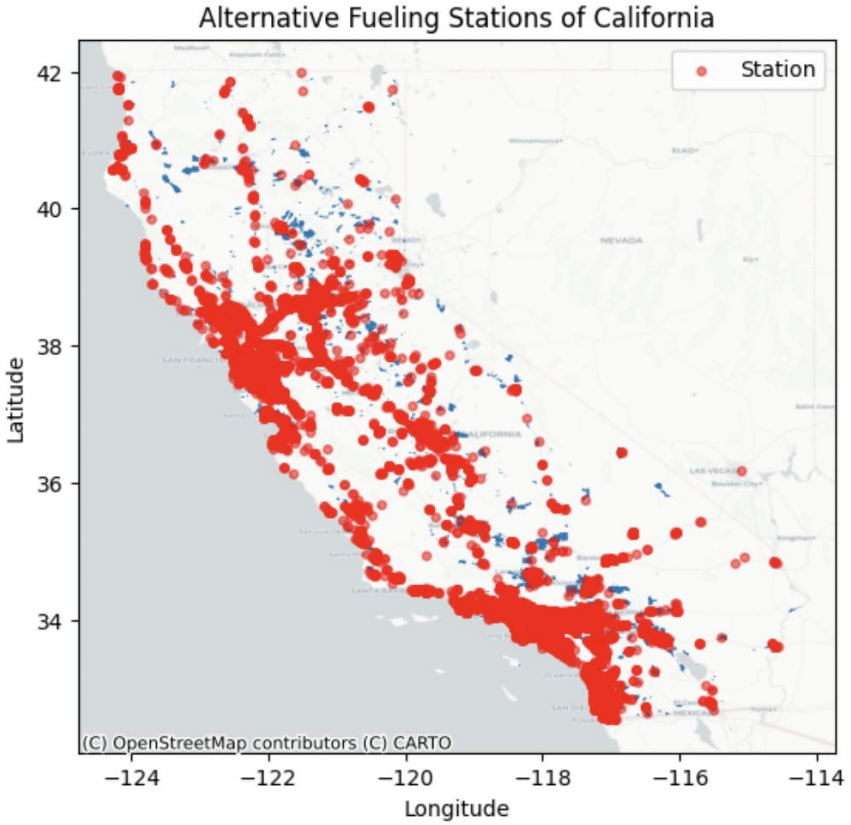


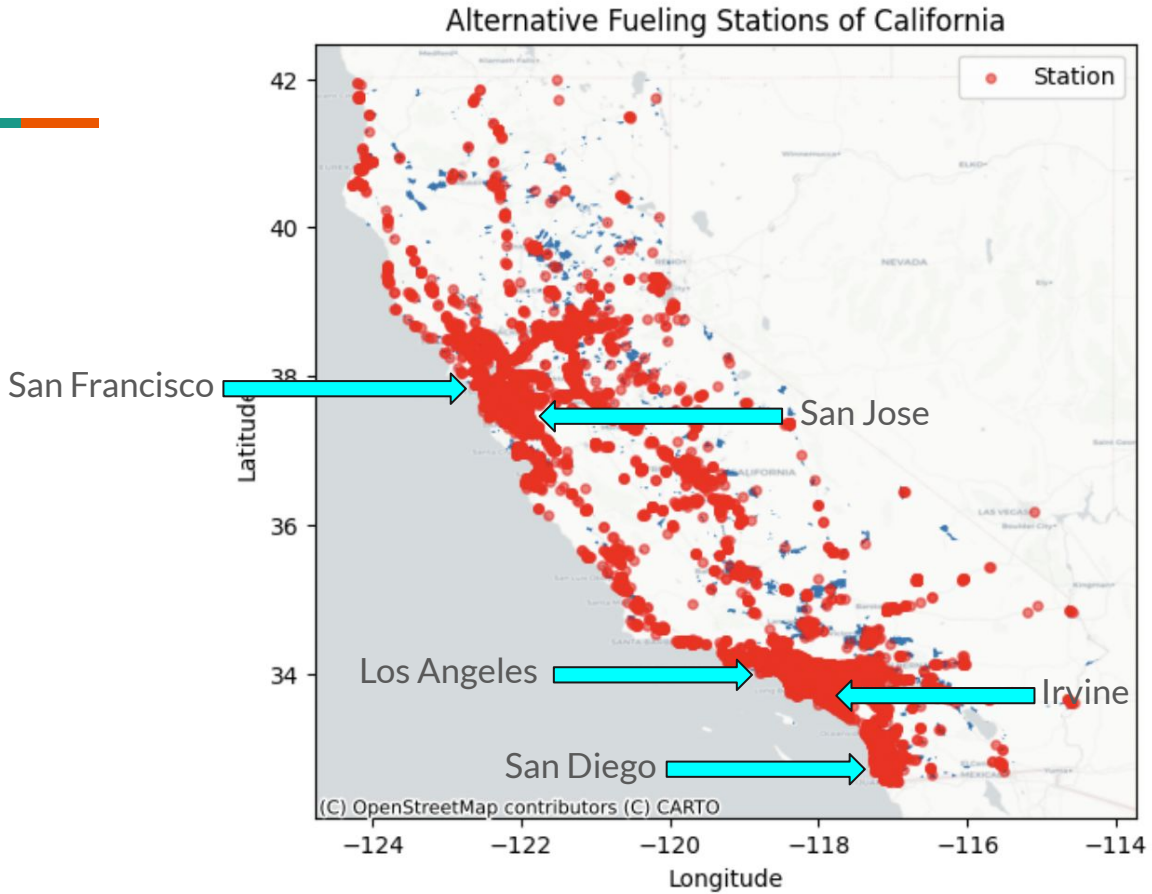
[30]:

| | Number of Trips | Population Not Staying at Home | Population Staying at Home | Number of Stations | Proportion (%) |
|---|---|---|---|---|---|
| **VT** | 3.130105e+06 | 4.928074e+05 | 1.311816e+05 | 414 | 0.013226 |
| **CA** | 1.483283e+08 | 2.997671e+07 | 9.535510e+06 | 19037 | 0.012834 |

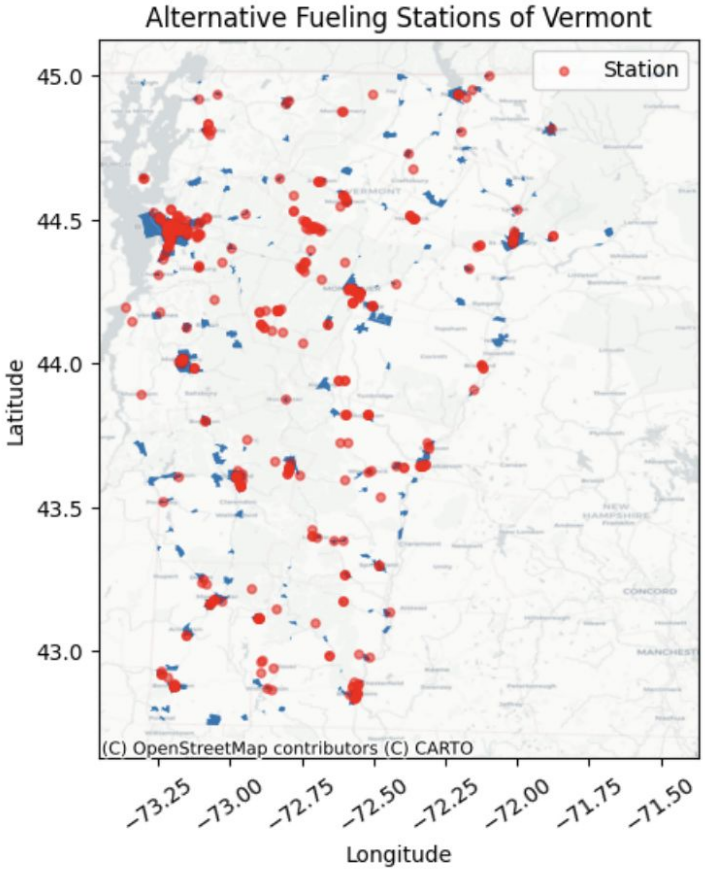

Scatter Plot of Driving Trips vs. Fueling Stations by State
Log Number of Fueling Stations
Log Number of Driving Trips
$10^4$
$10^3$
$10^2$
$10^7$
$10^8$
CA
FL
NY
TX
MA
COWA
MD
GA
VA
NJ
OR
TN
WI
IN
NV
KS
OK
SC
AL
KY
LA
AR
NE
NM
MS
VT
DC
RI
HI
ME
NH
ID
WV
DE
SD
MT
WY
ND
AK

To create this map, I found a cartographic boundary shapefile of California from the United States Census Bureau.

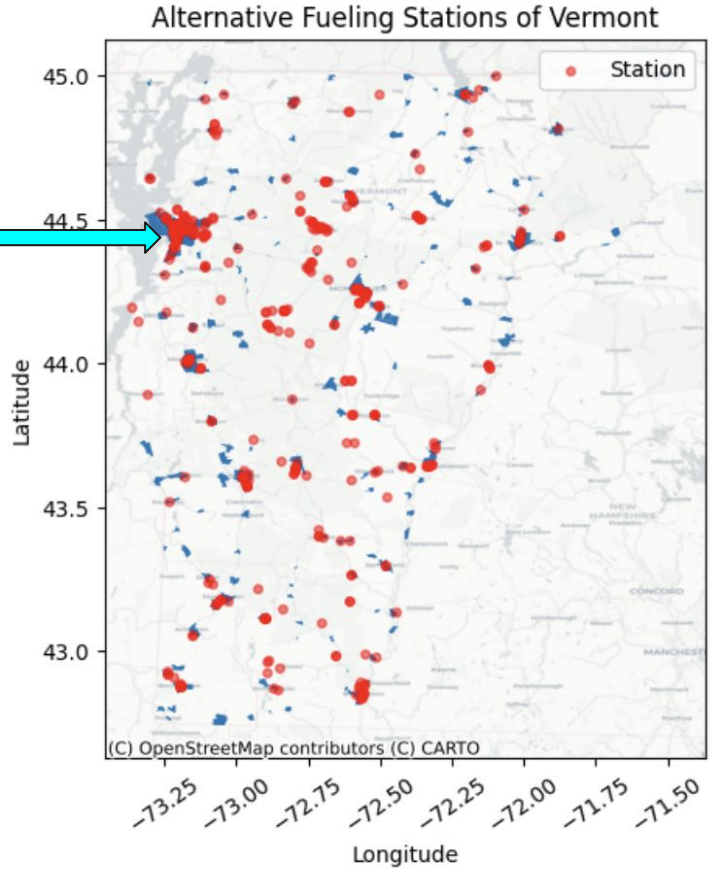



Alternative Fueling Stations of California
Station
San Francisco
San Jose
Los Angeles
Irvine
San Diego
(C) OpenStreetMap contributors (C) CARTO
Latitude
Longitude
42
40
38
36
34
−124
−122
−120
−118
−116
−114

To create this map, I found a cartographic boundary shapefile of Vermont from the United States Census Bureau.



Burlington


Alternative Fueling Stations of Vermont
Station
45.0
44.5
44.0
43.5
43.0
Latitude
(C) OpenStreetMap contributors (C) CARTO
−73.25
−73.00
−72.75
−72.50
−72.25
−72.00
−71.75
−71.50
Longitude

# What kind of data analysis would you do on this project in the future?

In the future, I would include average income per city/state to see if areas with higher incomes are likely to have alternative fueling stations for electric vehicles. I also think it would be interesting to compare the differences of private vs. public charging stations and research the patterns of fares for charging. These data analyses would provide further insight into the distribution of alternative charging stations and sustainable transportation.