How do statewide measures impact statewide air pollution?

Climate-DEchangers
(Wes, Dylan, Dalton)

Motivation/Introduction

- Current forest fires have motivated us to find data about our air quality.
- Variables are complexed so we decided to focus on policy related measures.
- How do these policies affect the air pollution in their state?

Public Transportation

- Transportation score is based on:
- Public transit resources
- Safety and reliability
- Accessibility and convenience
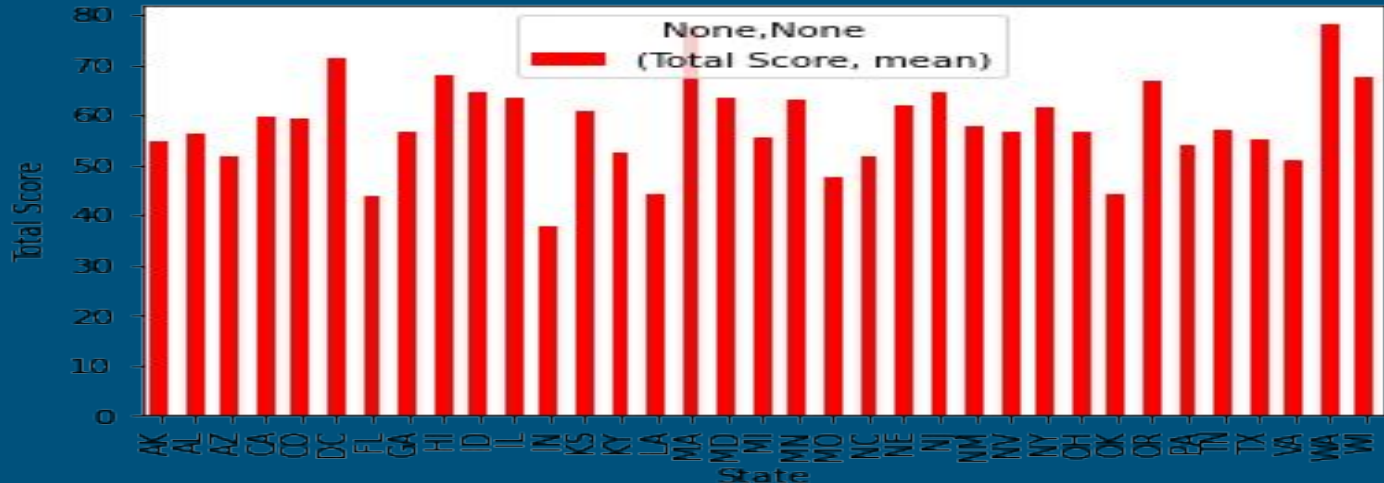
Public Transportation

```
[17]: file = 'Resources/public_transportation.csv'
      raw_df = pd.read_csv(file)
      next_raw_df = raw_df.sort_values(by=['Overall'])
      next_raw_df
```

[17]:

	Overall	City	Total Score	Accessibility	Safety	Resources
93	1	Seattle, WA	77.97	7	6	3
98	2	Boston, MA	77.84	2	34	10
99	3	San Francisco, CA	75.21	1	27	80
97	4	Washington, DC	71.23	3	48	52
54	5	Madison, WI	71.06	46	1	4
...	...	...	...	...	...	...
72	96	New Orleans, LA	42.43	28	92	100
47	97	Charlotte, NC	40.67	53	95	88
33	98	Tampa, FL	24.69	67	98	91
27	99	St. Petersburg, FL	24.03	73	98	91
13	100	Indianapolis, IN	21.13	87	100	39

100 rows × 6 columns

Subsidies

- Very large csv comprising of each state what subsidies were offered on each of the 64 Electric car types.

	state	statecode	year	time	makemodel	evtype	vehiclepurchaseincentive_dummy	vehiclepurchaseincentive_amount	homechargersubsidy_dummy
1920	Colorado	CO	2010	2010h1	Bmw Active E	bev	0	0.0	1
1921	Colorado	CO	2010	2010h1	Bmw I3	bev	0	0.0	1
1922	Colorado	CO	2010	2010h1	Bmw I3REx	phev	0	0.0	1
1923	Colorado	CO	2010	2010h1	Bmw I8	phev	0	0.0	1
1924	Colorado	CO	2010	2010h1	Bmw X5	phev	0	0.0	1
...	...	...	...	...	...	...	...	...	...

- Found the average amount of subsidy given total by state by year

Out[4]:

|	subsidies excluding hov	year	subsidies	state
0	2062.5	2010	5359.634	California
0	2062.5	2011	5310.283	California
0	2062.5	2012	5304.900	California
0	2062.5	2013	5303.941	California
0	2062.5	2014	5287.633	California
0	2062.5	2015	5344.072	California

Subsidy graphs
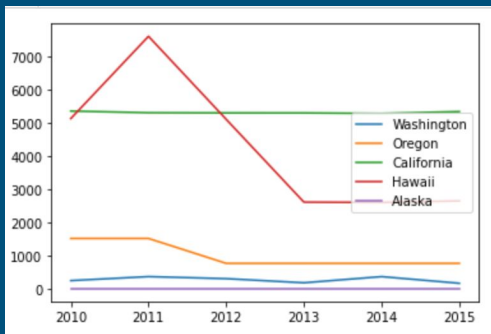
```python
def state_graph(state_list):
    for state in state_list:
        state_df = useful_df.loc[useful_df['state'] == state]
        plt.plot(state_df['year'],state_df['subsidies'])
    plt.legend(state_list,loc='right')
```
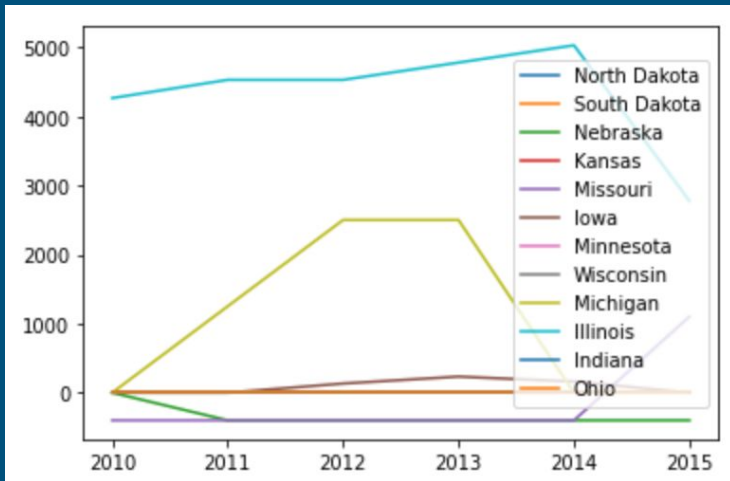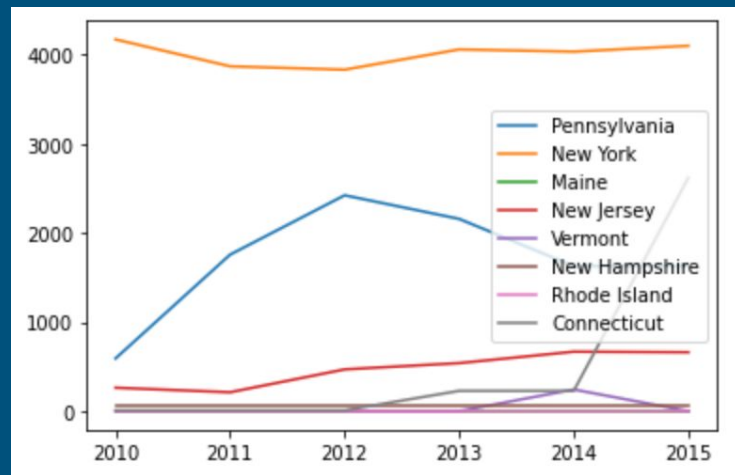

Pacific

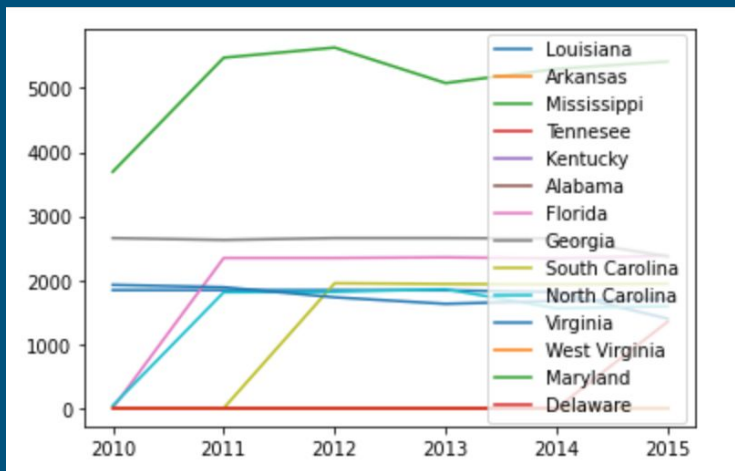Rocky Mountains
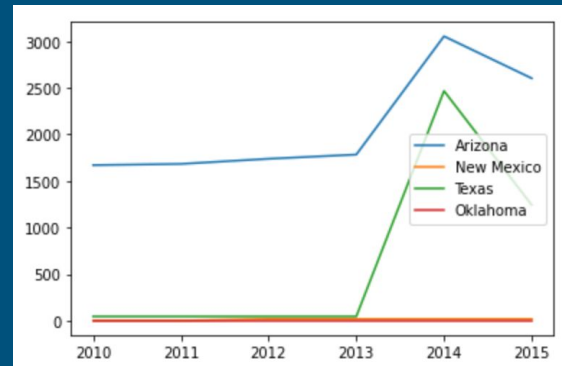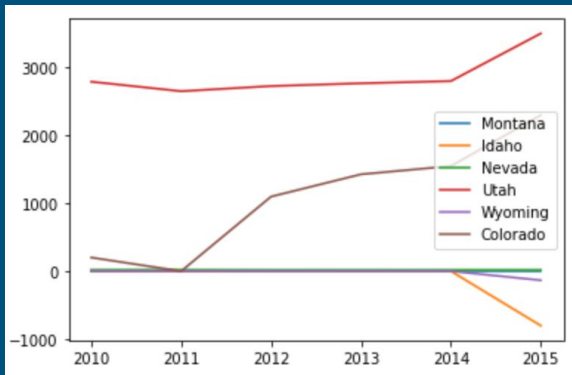
Southwest

More graphs!

Southeast

Midwest

Northeast

Air quality

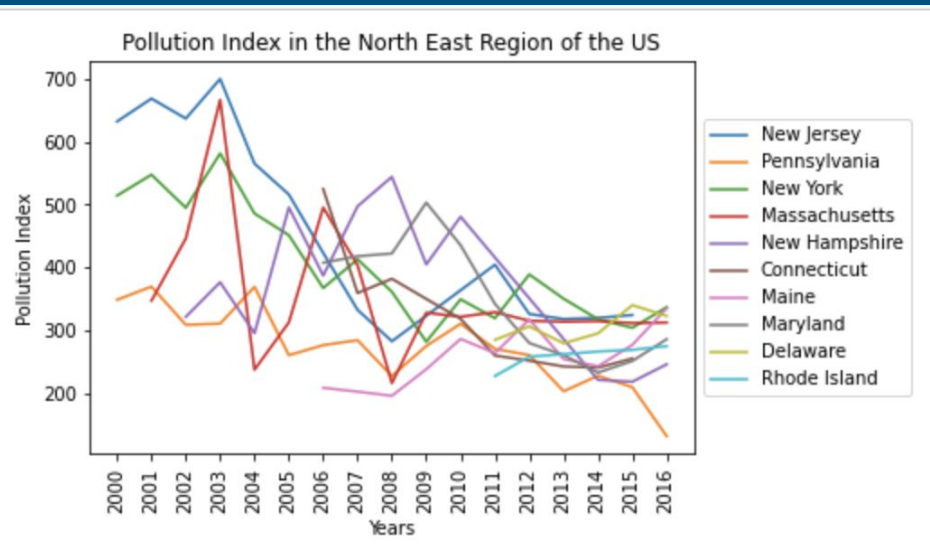|	State	Year	NO2 Mean	O3 Mean	SO2 Mean	CO Mean	Pollution Index
2	Alabama	2013	12.065635	12.098839	0.750311	211.006887	235.921672
3	Alabama	2014	9.411746	24.660149	1.186517	205.921470	241.179882
4	Alabama	2015	9.078580	23.999967	1.009429	221.026680	255.114656
5	Alabama	2016	9.524873	28.347745	0.835118	209.514138	248.221873
6	Alaska	2014	11.658510	10.306575	5.951531	455.443869	483.360485

Air quality

```
#Northeast
north_east = new_df.loc[((new_df["State"] == "Maine")) | ((new_df["State"]=="Massachusetts")) |
                        ((new_df["State"] == "Rhode Island"))| ((new_df["State"] == "Connecticut"))|
                        ((new_df["State"] == "New Hampshire")) | ((new_df["State"] == "Vermont"))|
                        ((new_df["State"] == "New York"))|((new_df["State"] == "Pennsylvania"))|
                        ((new_df["State"] == "New Jersey"))|((new_df["State"] == "Delaware"))|
                        ((new_df["State"] == "Maryland"))]

north_east = north_east.sort_values("Year").reset_index(drop=True)
ne = sns.lineplot(data=north_east, x="Year", y="Pollution Index", hue="State")
plt.xlabel("Years")
plt.ylabel("Pollution Index")
plt.title(" Air Pollution in the Northeast Region of the US")
ax.yaxis.set_major_locator(ticker.MultipleLocator(5))
ax.yaxis.set_major_formatter(ticker.ScalarFormatter())
plt.xticks(rotation = 'vertical')
plt.legend(loc='center left', bbox_to_anchor=(1.0, 0.5))
plt.show()
```
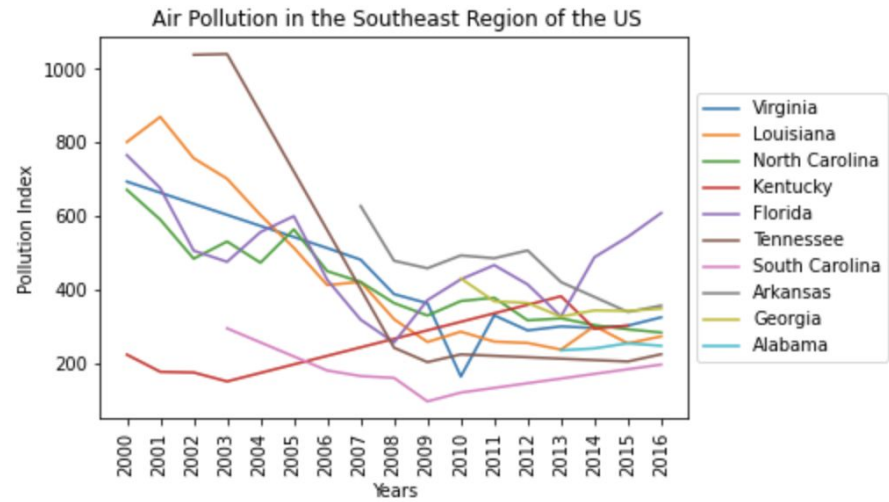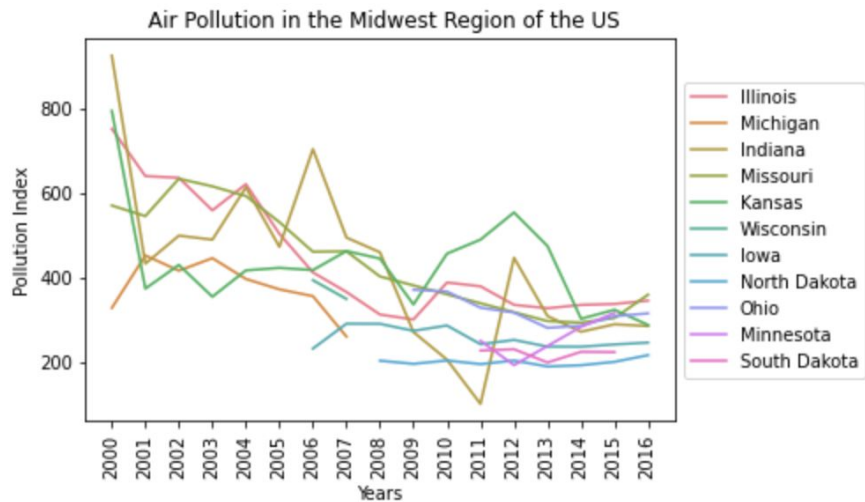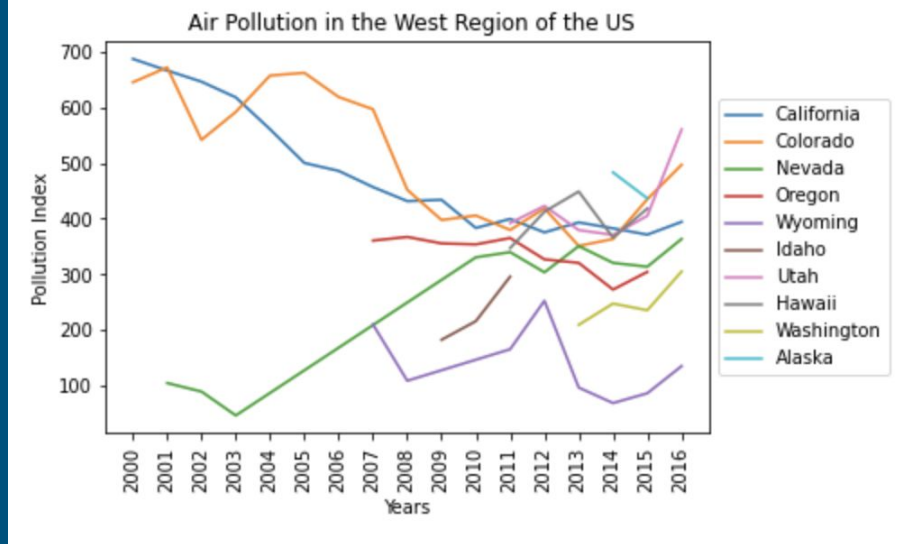
Air quality

Pollution Index in the North East Region of the US

AIR POLLUTION IN THE UNITED STATES - NORTHEAST

Air quality

Air Pollution in the Midwest Region of the US

Air Pollution in the Southeast Region of the US

Air quality

Air Pollution in the Southwest Region of the US

Air Pollution in the West Region of the US
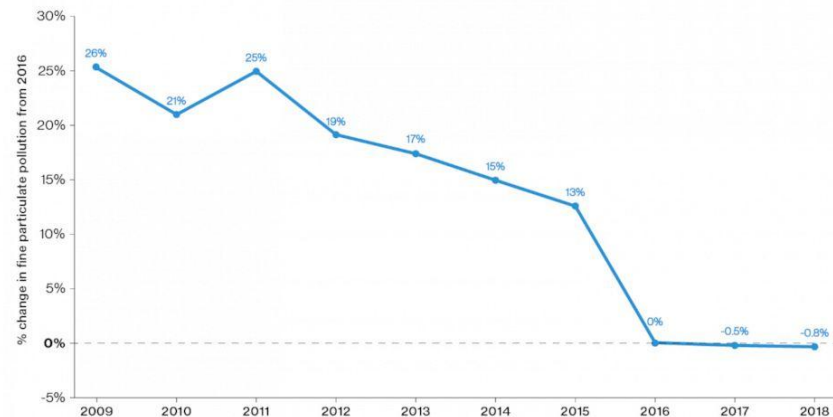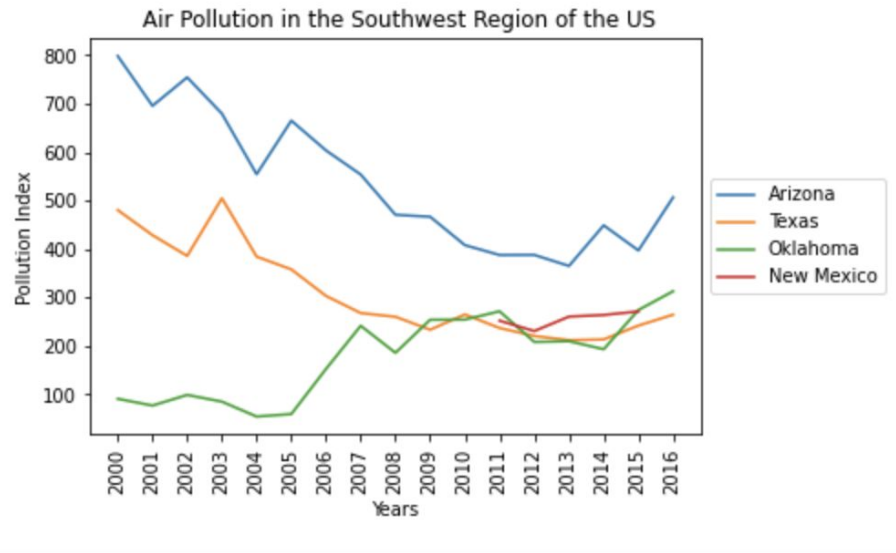
Putting them together

Subsidies!

R-Value: -0.02428867381001295

Subsidy amount vs. Sulfur Dioxide levels

y = -0.0x + 0.566
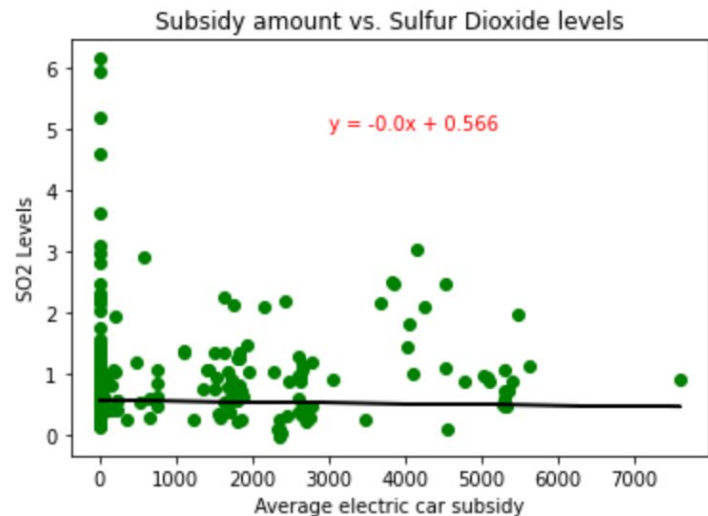
R-Value: -0.04041105300347709
Slope: -1.1323061372474311e-07

Subsidy amount vs. O3 levels

y = -0.0x + 0.027

Transportation scores

Sulfur Dioxide levels vs. Public transportation score

$y = -0.017x + 2.916$
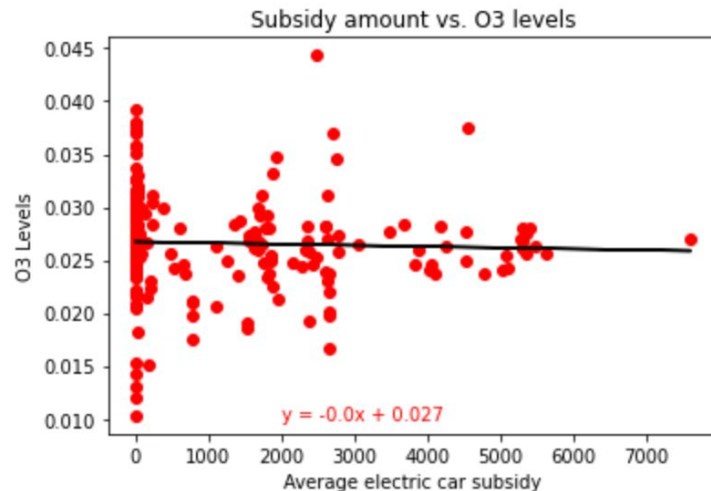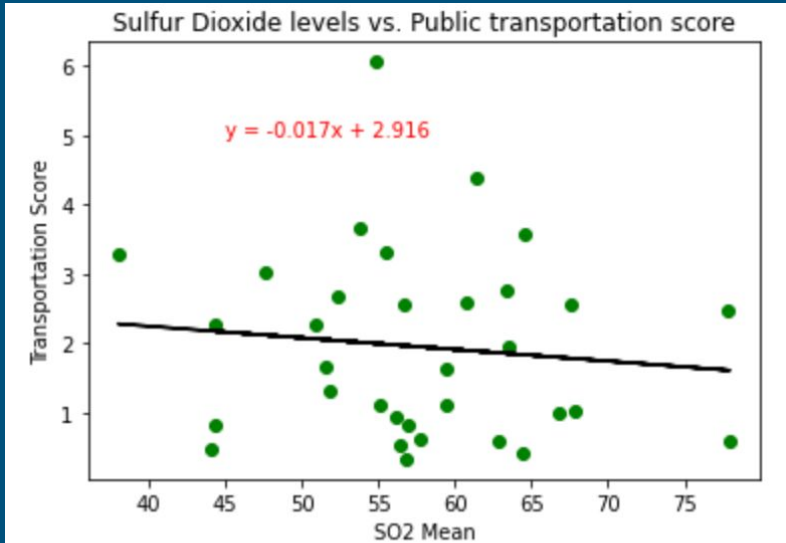
R-Value: -0.11092648809339511

Ozone levels vs. Public transportation score

$y = -0.00021x + 0.037$
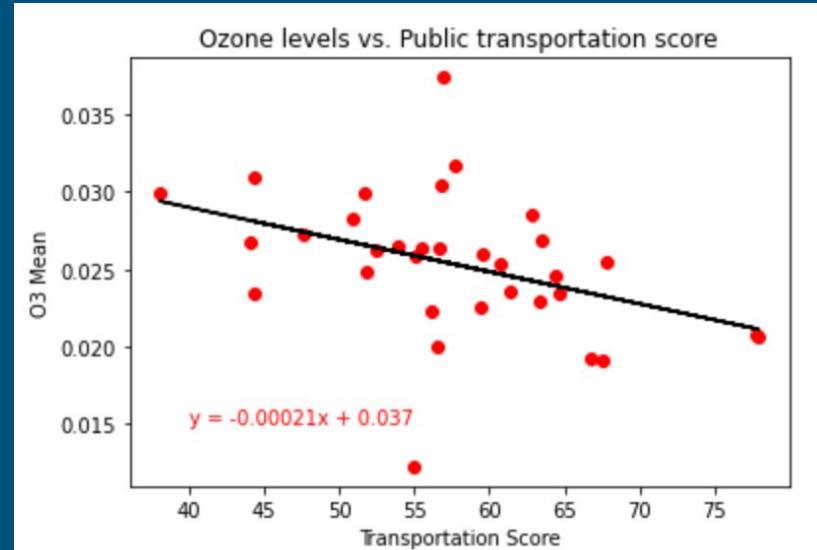
R-Value: -0.4031468048697786

Conclusion

- Could not find a clear relationship between our selected measures and air pollution
- We need to look at more variables such as wildfires, factories, etc.
- A lot of our data varied which could lead to inconsistent results
- We still need to run a hypothesis test for the last 2 graphs in order to understand if subsidies and public transportation has a significant effect on air pollution

Data Sources

Subsidies:

https://www.sciencedirect.com/science/article/pii/S2352340919300071?via%3Dihub

Air Quality:

https://www.kaggle.com/sogun3/uspollution

Public Transportation:

https://wallethub.com/edu/cities-with-the-best-worst-public-transportation/65028