Association Between Different Air-Quality Index with Suicide Deaths on Geospatial Level

By Wenhuan Tan, Xiange Wang

ABSTRACT

Previous studies have suggested a possible link between air quality and mental health. However, these studies are often conducted in small regions with a limited time span. As part of large-scale research in suicide risk management, we collected 20 years' worth of daily air quality data, such as PM 2.5, Ozone, SO2, and CO, as well as suicide data, to construct a county-level model. Four sub-periods: 2001--2005, 2006--2010, 2011--2015, and 2016--2019 are analyzed with linear regression method and results are very similar in different periods. Association between suicide rate and PM2.5, SO2, and CO were not found, but Ozone may be highly associated with suicide rate. However, since previous study have showed the correlation between Ozone and temperature, future research on the association between suicide rate and CO and temperature is needed. Our findings are incorporated into a community characterization model for policy-making assistance. All computation was carried out using Cori at NERSC.

DATA

The air-quality indexes data was sourced from the US Environmental Protection Agency (EPA) and the suicide deaths data was from the US Center for Disease Control and Prevention (CDC)
- All data was from the year 2001 to 2019
- From EPA, the pre-generated daily summary data of PM2.5, Ozone, SO2 and Co was used
- From CDC, suicide deaths were based on a 5-year range in order to collect more county level suicide deaths data
- When analyze the correlation between air-quality indexes and suicide rate, the null values are all ignored
- Not every county had all air-quality indexes data from 2001 to 2019. Different air-quality index in different period had different county-level amount of data. The detailed information is as follow:

County Amounts with Different Data Index in Different Periods

Data Index	Period 1	Period 2	Period 3	Period 4
PM2.5	855	833	823	805
Ozone	799	828	828	805
SO2	393	370	351	359
CO	268	249	210	182
Suicide Death	2036	2096	2318	2113

- Period 1: Year 2001 ~2005
- Period 2: Year 2006 ~ 2010
- Period 3: Year 2011 ~ 2015
- Period 4: Year 2016 ~ 2019

Note: The unit of Suicide Rate is 1/100000.

METHOD

Collect Data
- Collected PM2.5 Daily Summary Data from EPA
- Collected Ozone Daily Summary Data from EPA
- Collected SO2 Daily Summary Data from EPA
- Collected CO Daily Summary Data from EPA
- Collected 5-year Suicide Deaths Data from CDC

Preprocess Data
- Group data by county code, year, and period
- Merge and clean
- Remove incomplete data features
- Data aggregation to find national average suicide rate and national average suicide increase rate
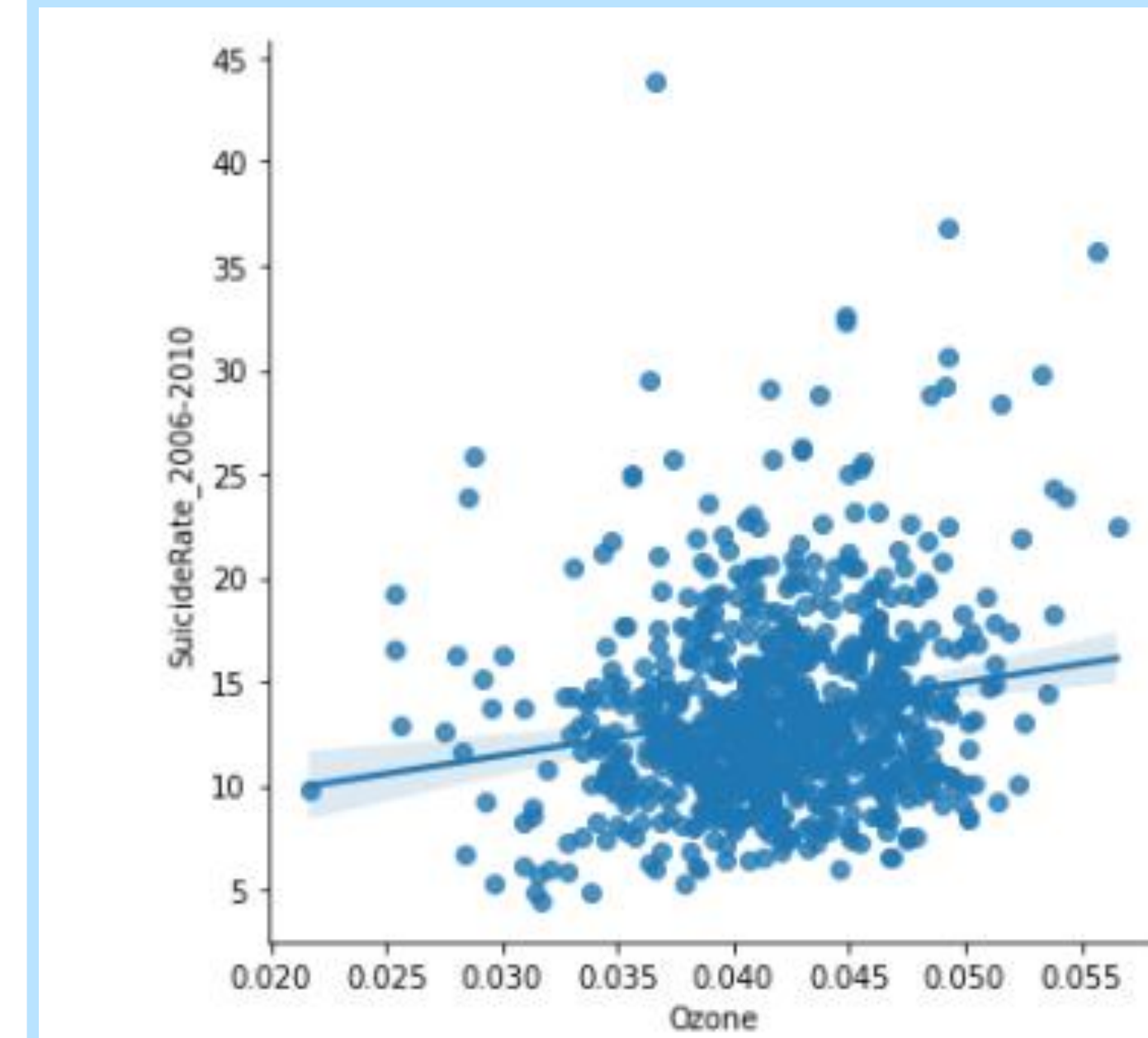
Linear Regression

Figure S1 – Linear Regression Example

In my **Linear Regression Model**, X-axis represent a single air-quality index and Y-axis represents Suicide Rate. If the value in X-axis and Y-axis increase at the same time, like the above figure S1, it indicates this air-quality index may be highly associated with the Suicide Rate.
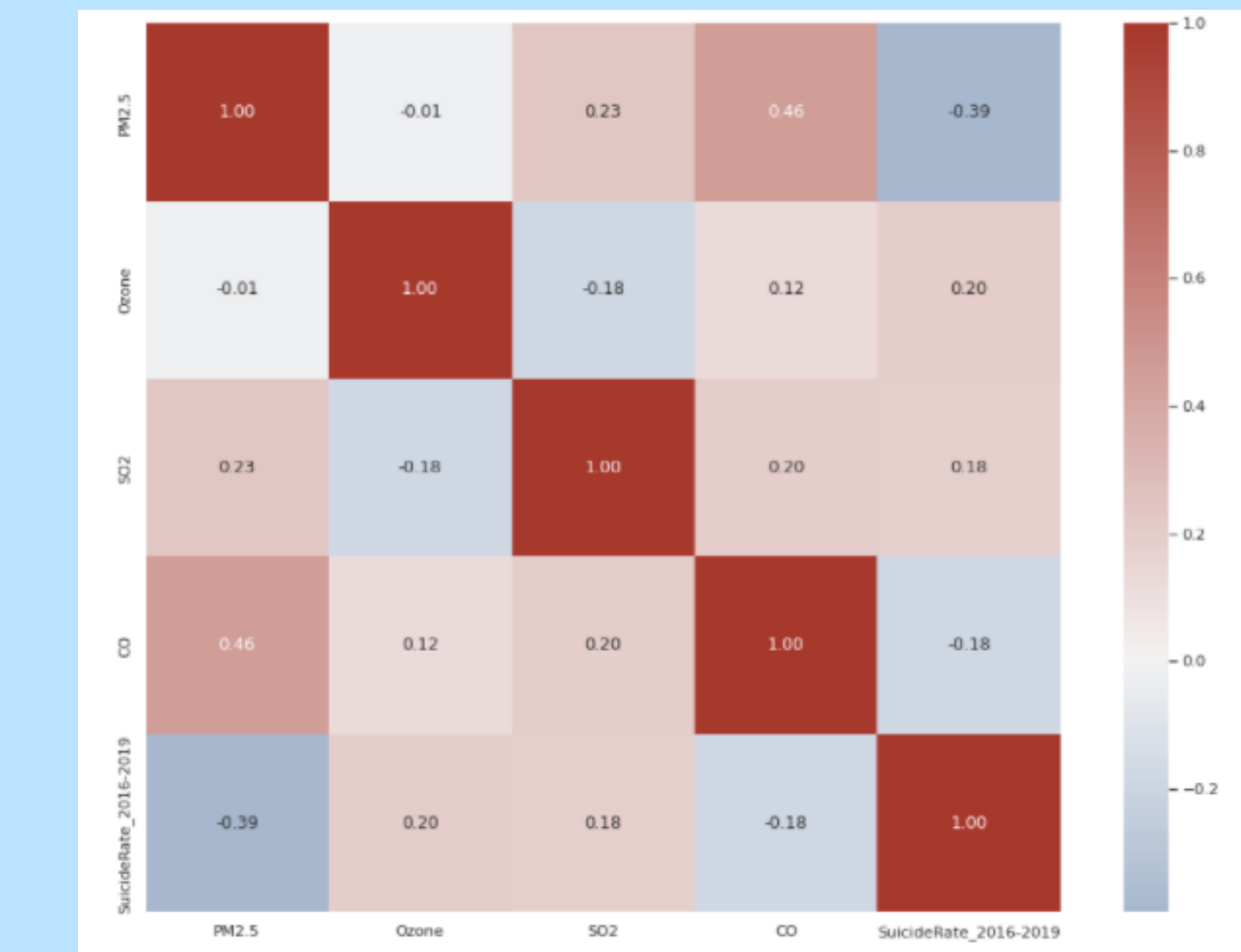
Pairwise Correlation

Figure S2 – Pairwise Correlation Example

Pairwise correlation was used to analyze the correlation between each air-quality index. From the above Figure S2, we can see that the all the coefficients were close to zero, which indicates no correlation was found between any two air-quality indexes.

RESULTS

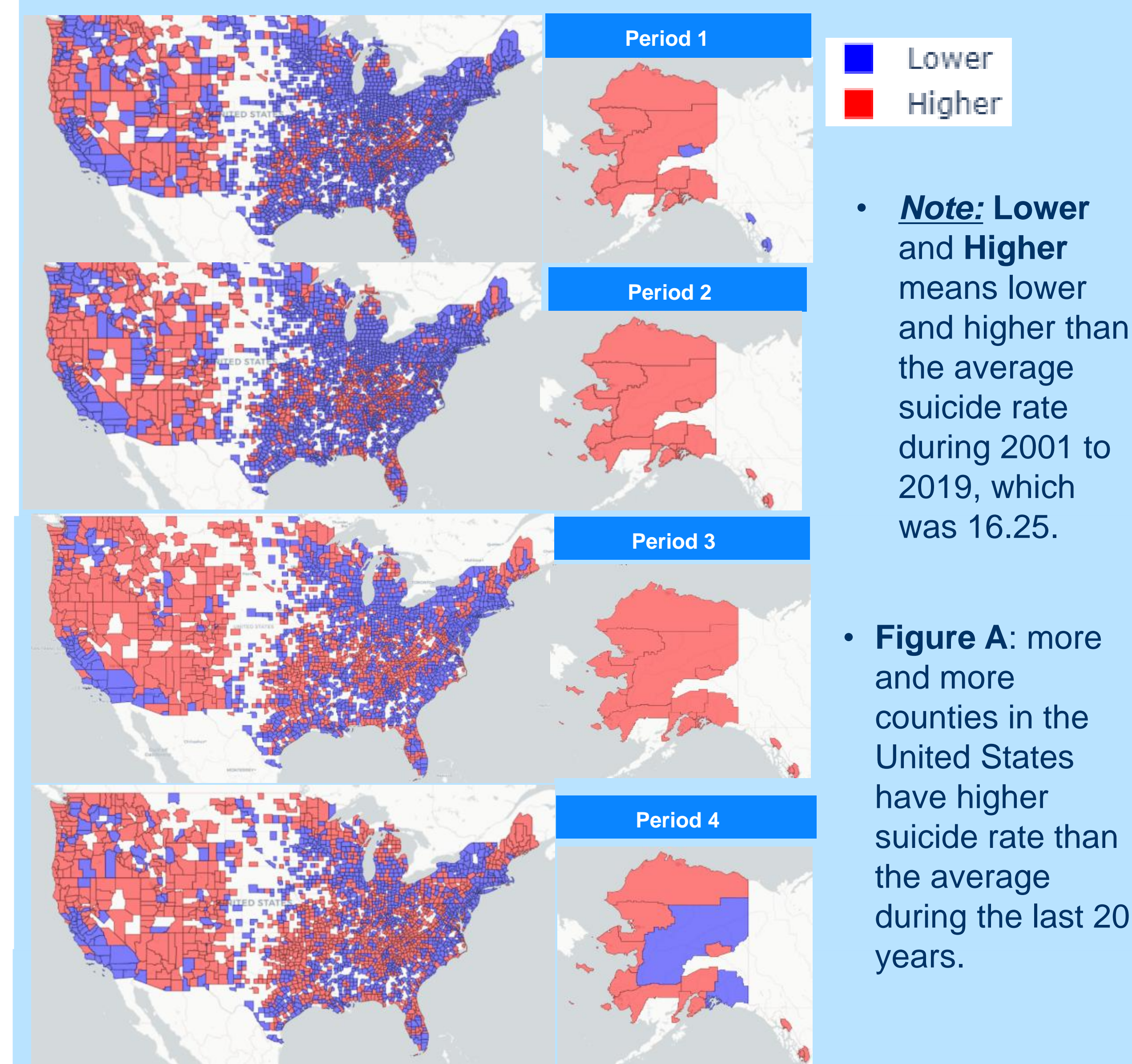### Suicide Rate Changes on Geospatial Level over 4 Periods

Figure A - Suicide Rate Changes on the map in 4 periods

- **Lower** / **Higher**

- *Note:* Lower and **Higher** means lower and higher than the average suicide rate during 2001 to 2019, which was 16.25.

- **Figure A**: more and more counties in the United States have higher suicide rate than the average during the last 20 years.
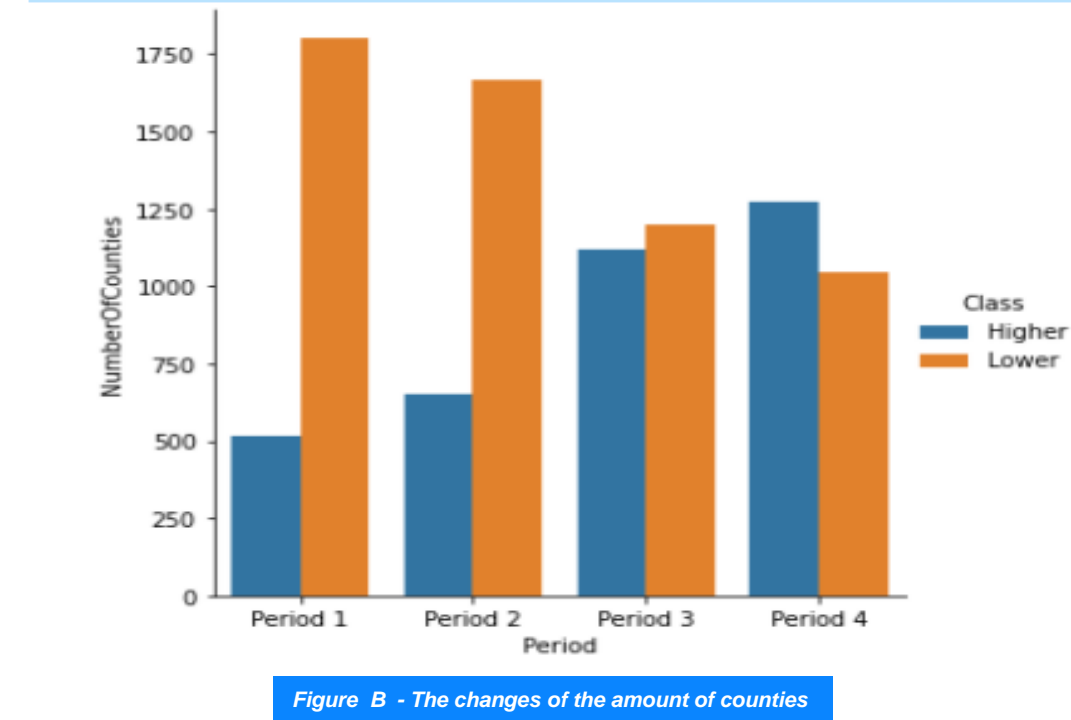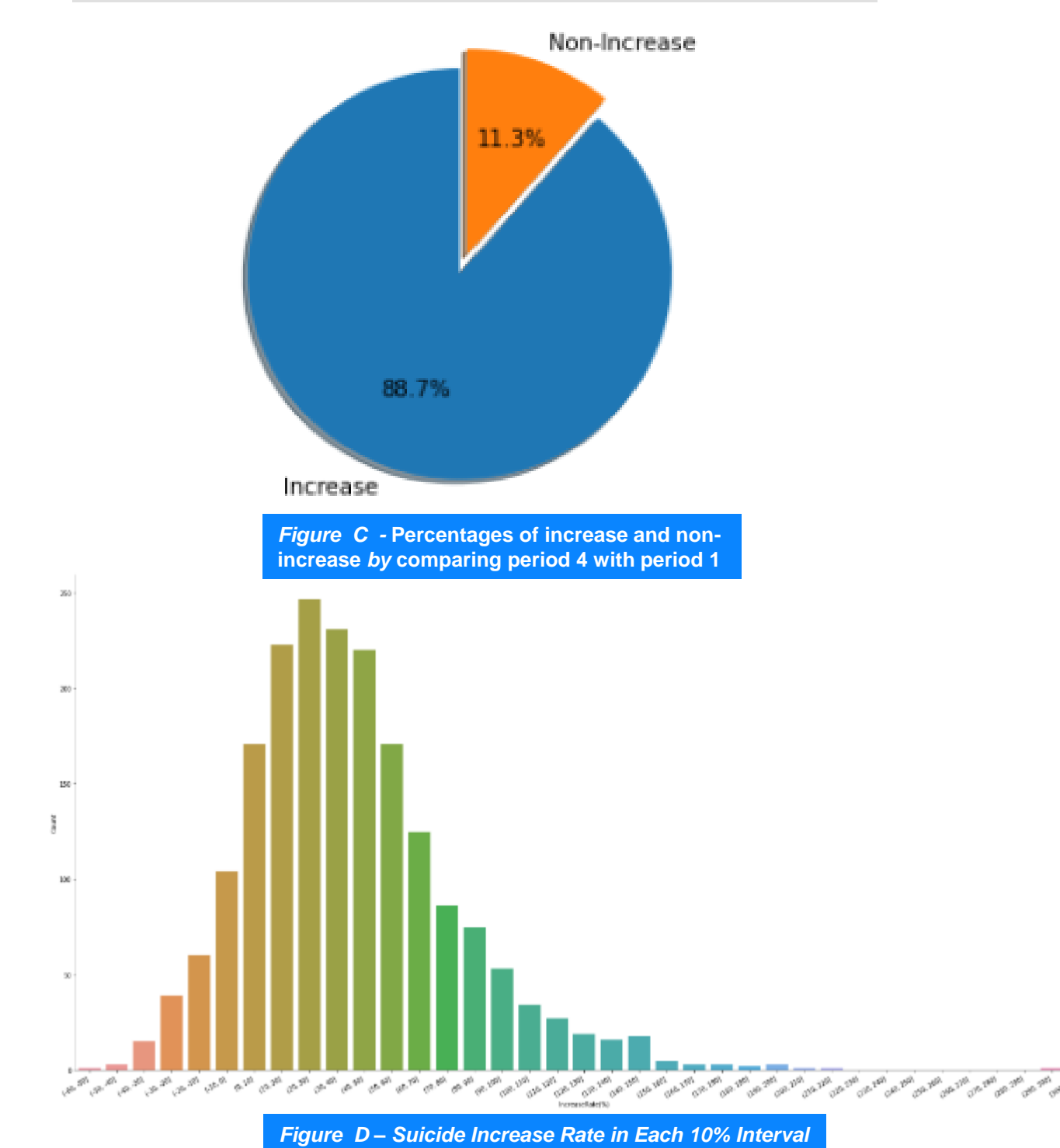

Figure B - The changes of the amount of counties

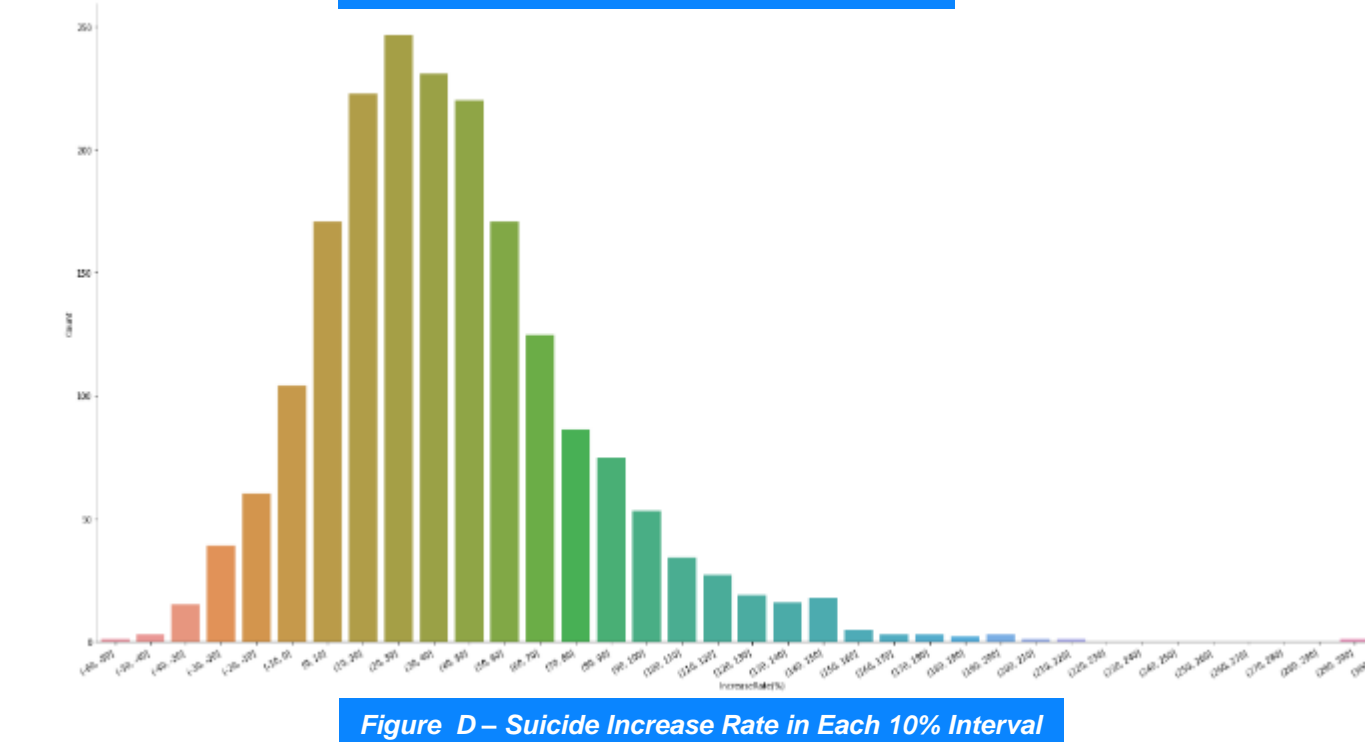Figure C - Percentages of increase and non-increase by comparing period 4 with period 1

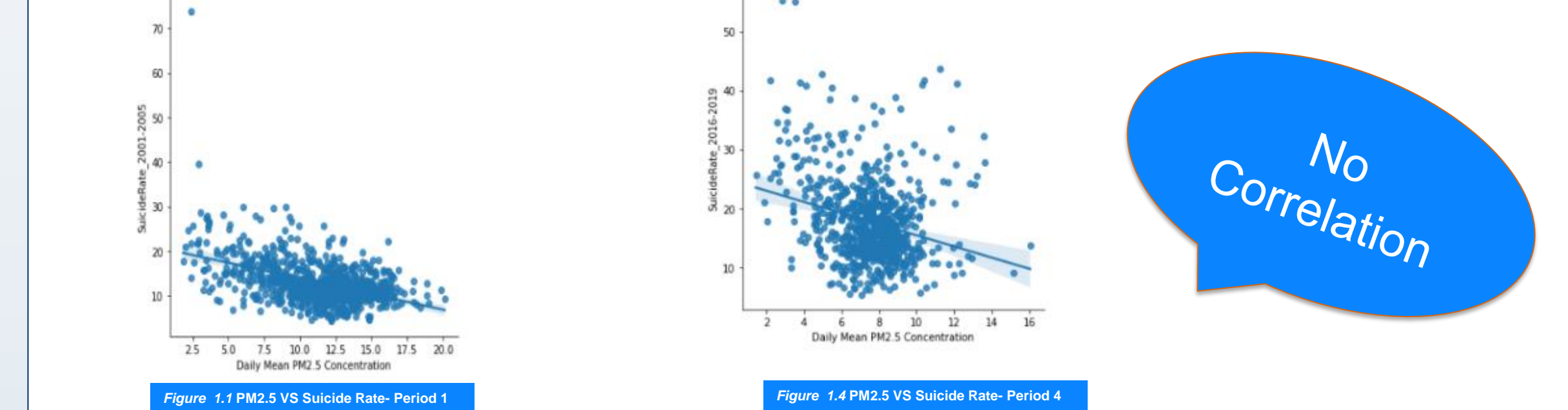Figure D – Suicide Increase Rate in Each 10% Interval

- **Figure B**: In period 1, only 516 counties had higher suicide rate than the average suicide rate during the last 20 years, however, 1270 counties in period 4 had higher suicide rate than the average, which was more than doubled comparing to period 1.
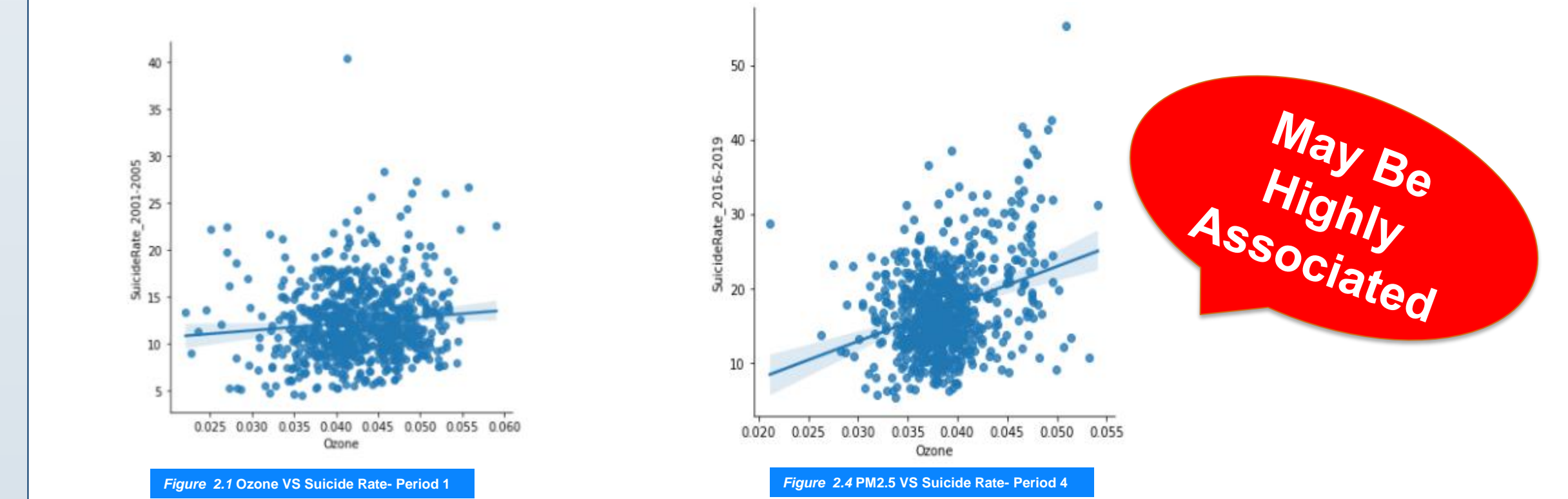
- **Figure C**: Comparing with period 1, the suicide rate in 88.7% of the counties got increased and only 11.3% of them got decreased or kept the same in period 4.

- **Figure D**: Comparing the period 4 and period 1, the average suicide increase rate was 40% and the graph in the left showed the amounts of counties in each 10% increase internal. **The highest suicide increase rate reached to more than 300%.**
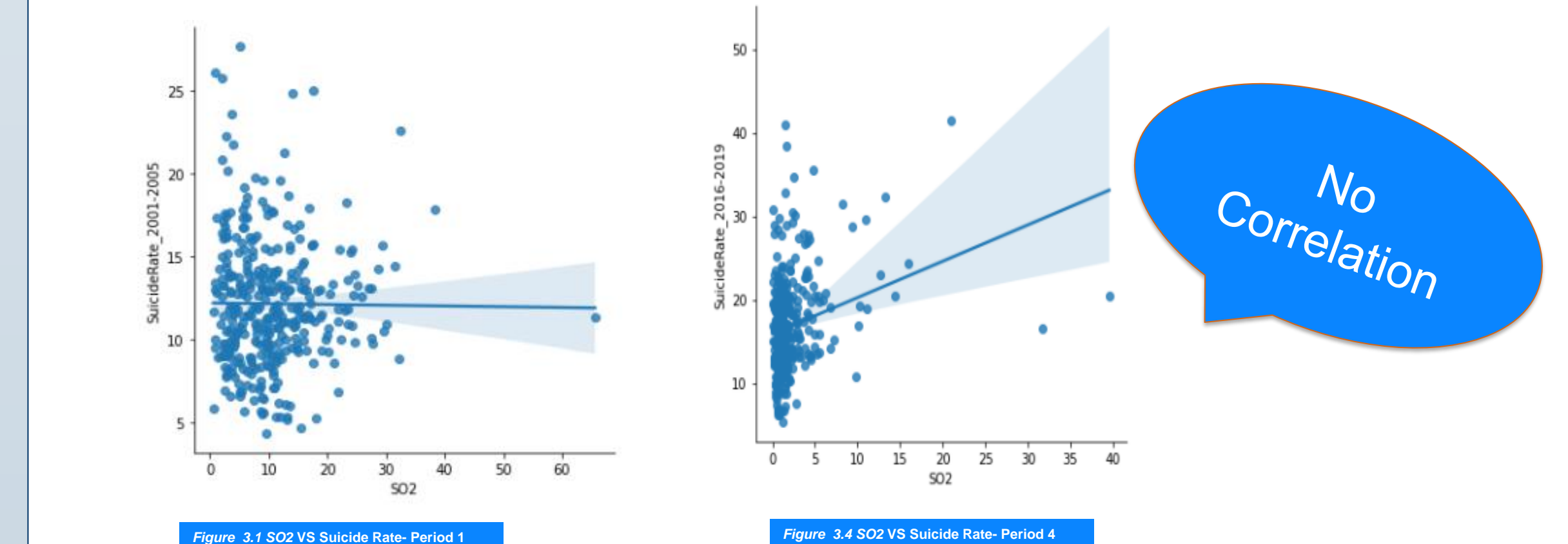
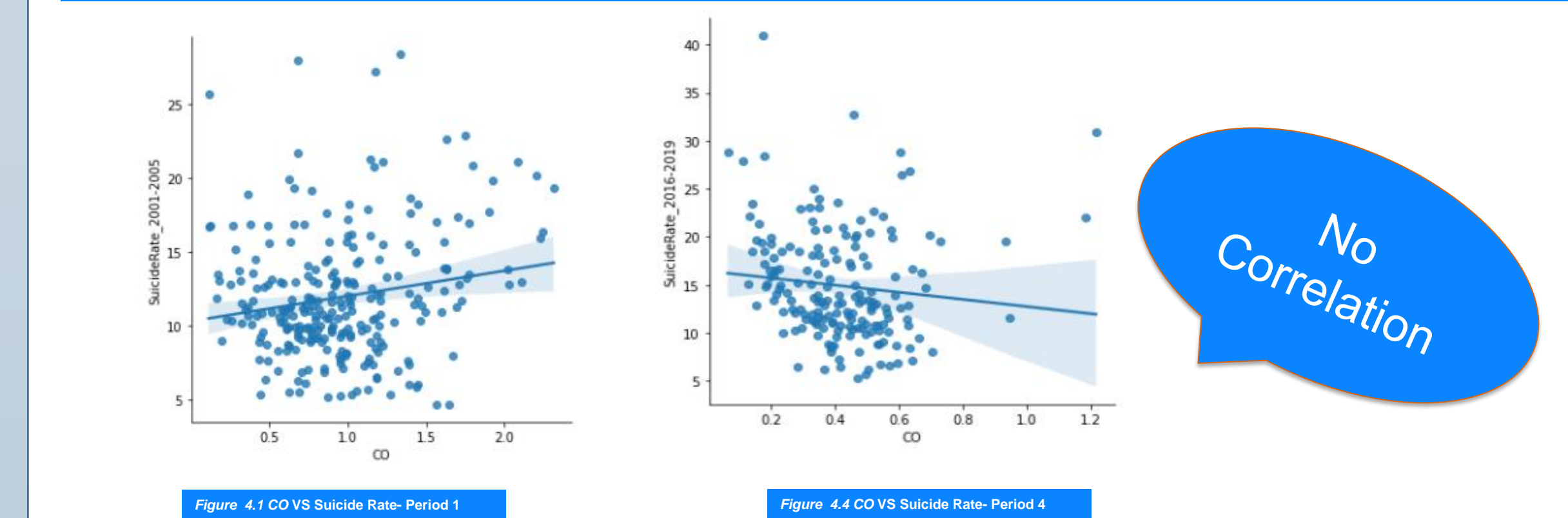### Relation Between Suicide Rate and PM2.5 in Period 1 and Period 4

No Correlation

Figure 1.1 PM2.5 VS Suicide Rate- Period 1 / Figure 1.4 PM2.5 VS Suicide Rate- Period 4

Relation Between Suicide Rate and Ozone in Period 1 and Period 4

May Be Highly Associated

Figure 2.1 Ozone VS Suicide Rate- Period 1 / Figure 2.4 PM2.5 VS Suicide Rate- Period 4

Relation Between Suicide Rate and SO2 in Period 1 and Period 4

No Correlation

Figure 3.1 SO2 VS Suicide Rate- Period 1 / Figure 3.4 SO2 VS Suicide Rate- Period 4

Relation Between Suicide Rate and CO in Period 1 and Period 4

No Correlation

Figure 4.1 CO VS Suicide Rate- Period 1 / Figure 4.4 CO VS Suicide Rate- Period 4

Coefficients Between Suicide Rate and Each Air-Quality Index

	Features	Period 1	Period 2	Period 3	Period 4
0	PM2.5	-0.46	-0.60	-1.09	-1.36
1	Ozone	149.36	167.93	241.35	349.70
2	SO2	0.07	0.06	0.41	1.21
3	CO	0.14	0.35	-2.25	-2.83

From all the coefficients in the left chart, we can see only the coefficients of Ozone are big positive values, which match with the Linear Regression graphs and indicate that Ozone may be highly associated to Suicide Rate.

FUTURE WORK

- Temperature data needs to be collected
- Correlation between Ozone and Temperature needs to be analyzed
- Correlation between Temperature and Suicide Rate needs to be analyzed

ACKNOWLEDGEMENTS

CONTACT

- wenhuan.tan@bellevuecollege.edu
- xiangewang@lbl.gov
- liu@hood.edu