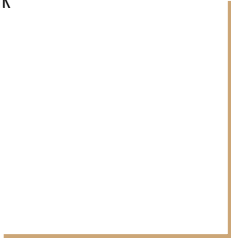# Does days with ozone pollution predict the median AQI for counties in the US in 2021 controlling for days where AQI measurements were taken?

By Krystin Sinclair and John Wawrzyniak
EPBI 5002
Dr. Wu
April 29, 2022

Question and Hypotheses

Research Question: Does days with ozone pollution predict the median AQI for counties in the U.S. in 2021 controlling for days where AQI measurements were taken?

Null Hypothesis: Days with ozone pollution does not significantly predict the median AQI for counties in the US in 2021 with control for days where AQI measurements were taken.

Alternative Hypothesis: Days with ozone pollution does significantly predict the median AQI for counties in the U.S. in 2021 with control for days where AQI measurements were taken.

Methods

Sample: AQI measurement data taken from 1,000 U.S. counties in 2021.

Explanatory Variable 1: Days with ozone pollution for counties in the U.S. in 2021.

Explanatory Variable 2: Days where AQI measurements were taken in U.S. counties in 2021. 1=More than half of the year has measurements, 0=Less than half of the year has measurements.

Response Variable: Median AQI for counties in the U.S. in 2021.

Statistical Test: Multiple regression, two-tailed, alpha level 0.05.

Results

Table 1: Characteristics of the variables in the study

Variable	Summary Statistic
Days with ozone pollution	Mean: 126.7. Standard Deviation: 88.7.
Days measurements were taken	
1=More than half of the year	615 or 61.6%.
0=Less than half of the year	384 or 38.4%.
Median AQI	Mean: 36.1. Standard Deviation: 11.5.

The correlation between days with ozone pollution and median AQI was significant. r=0.533 and the p-value is less than 0.0001.

The mean median AQI was 38.48 (standard deviation 10.38) where the variable days measurements were taken=1 (measurements were taken for more than half of the days in the year). The mean median AQI was 32.27 (standard deviation 12.22) where the variable days measurements were taken=0 (measurements were taken for less than half of the days in the year.).

Using a Two Sample t-Test, we found that the difference between groups was significant t(df=715.05)=-8.27, p-value is less than 0.0001.

Results

The test statistic for the regression model: $F_{(2, 996)}=198.19$, p-value is less than 0.0001.

Days with ozone pollution significantly predicts median AQI (b=0.07, $t(996)$=17.34, p-value is less than 0.0001) with control for days where measurements were taken.

We reject the null hypothesis.

Table 2. Results of a Multiple Regression Analysis using days with ozone pollution and days with measurement taken to predict median AQI.

R-squared= 0.2847

Variable	b	t	p-value
Days with Ozone Pollution	.07	17.34	<.0001
Measurement Taken (0 vs 1)	.29	.39	.69

Conclusion

We found evidence that days with ozone pollution significantly predicts median AQI with control for days with measurement taken, such that increased days with ozone pollution is associated with increased median AQI.