An Analysis of the Relationship between Presence of Green Space in 330 Cities in the United States and the Health Status of Their Citizens

Naomi Mauss, Department of Data Science

Defining Health Status:

- Kidney disease
- Asthma
- Cancer
- High Blood Pressure
- Arthritis
- Diabetes
- Chronic Obstructive Pulmonary Disease (COPD)
- Binge drinking
- Poor mental health

Datasets

US Parks

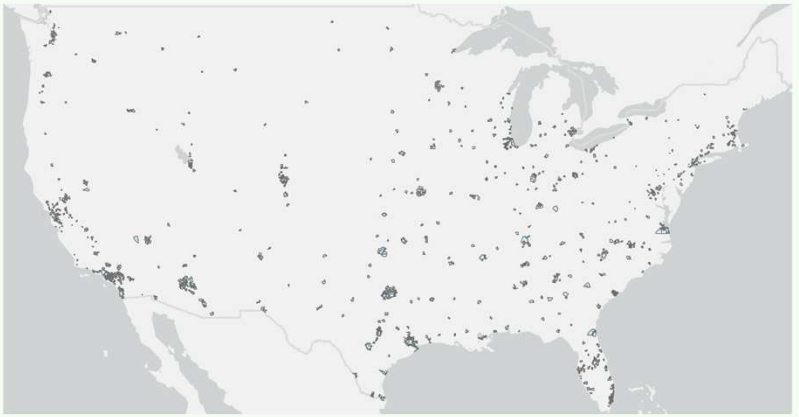
Data cleaning:

- Remove duplicates
- Calculate geometry
- Pairwise Clip
- Explode
- Comparative statistics

Data merging:

- Select by Attributes
- Spatial Join
- Pairwise Intersect

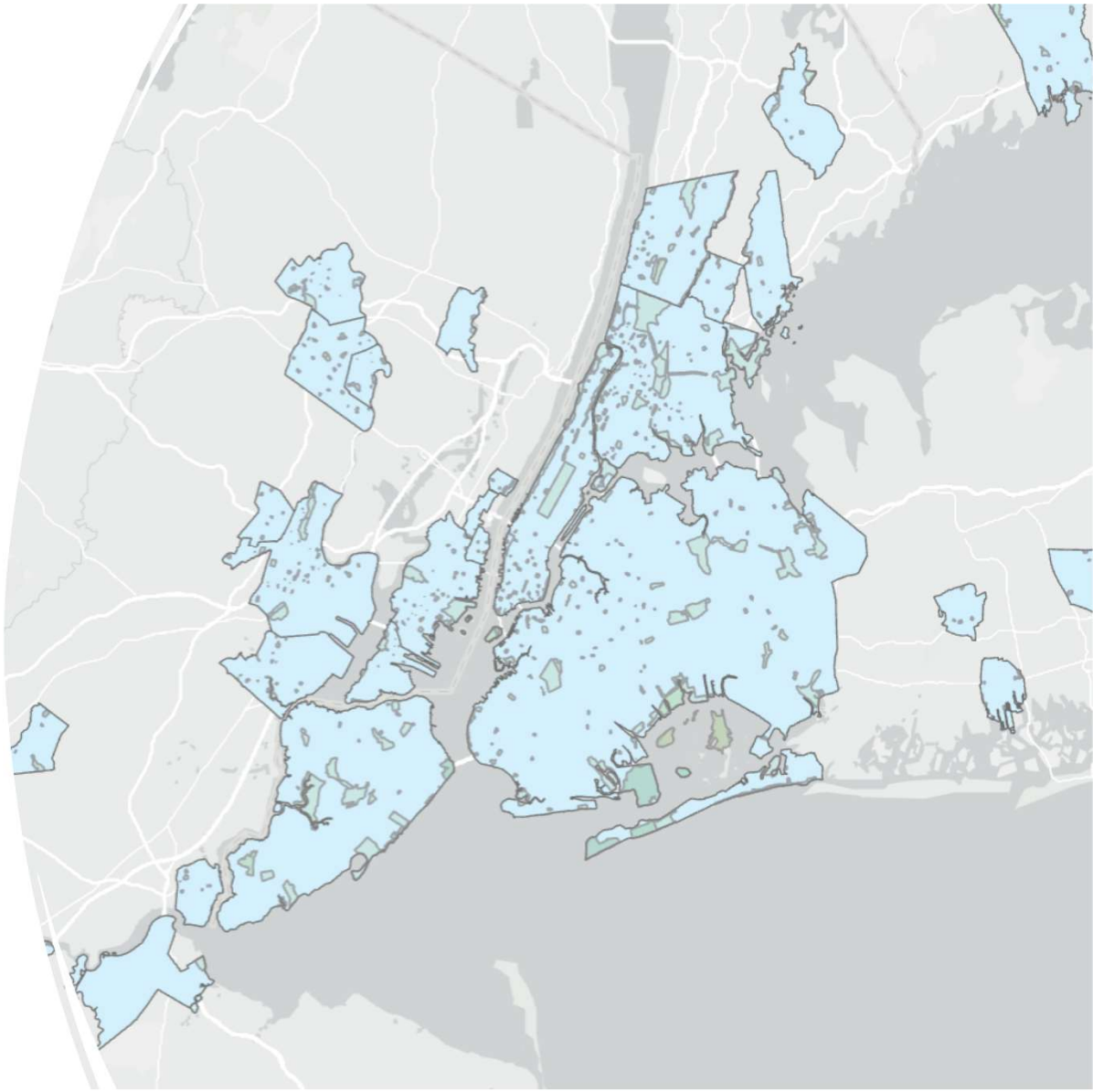
US Cities



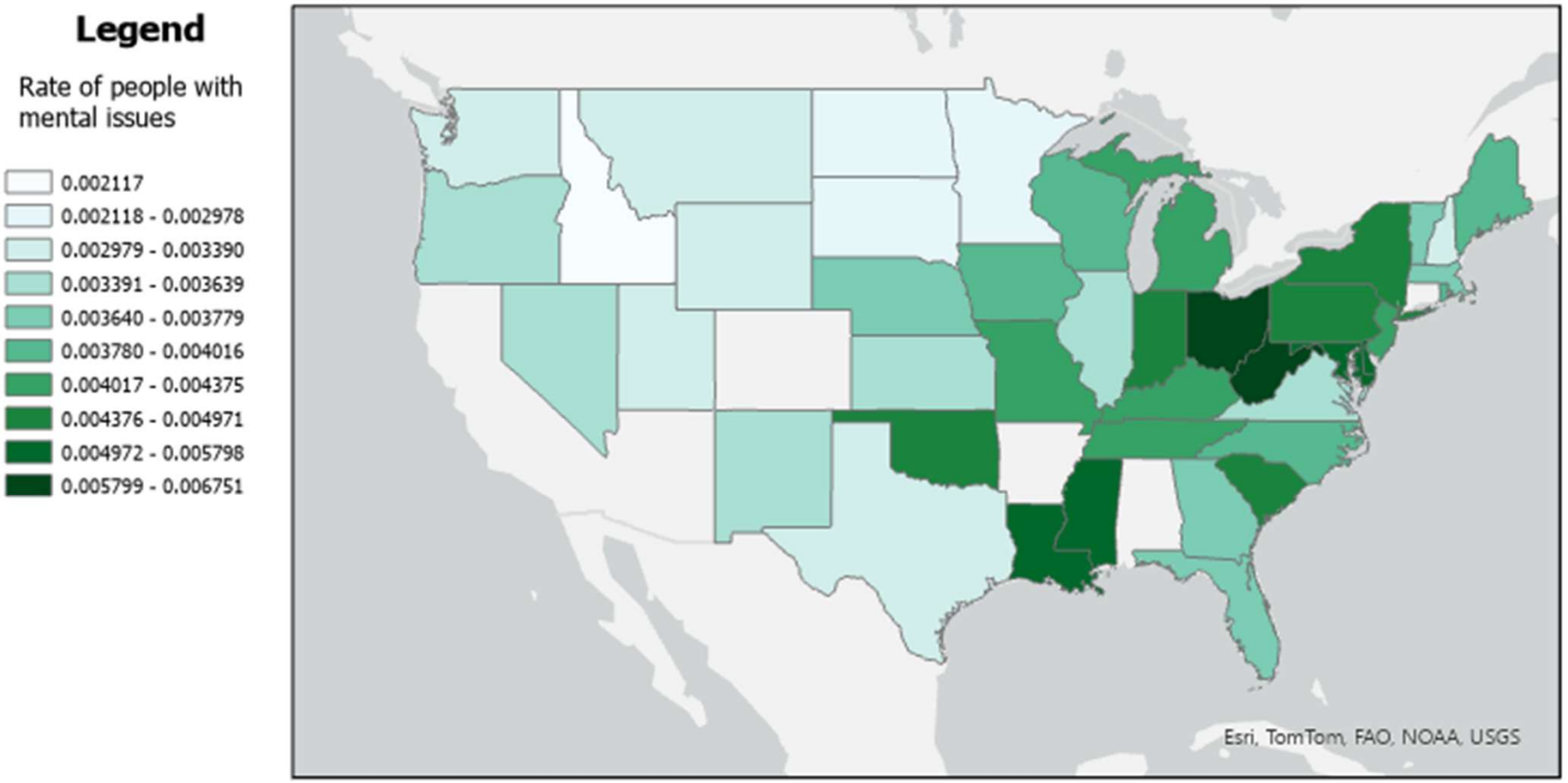
CDC's PLACES Dataset

Output:

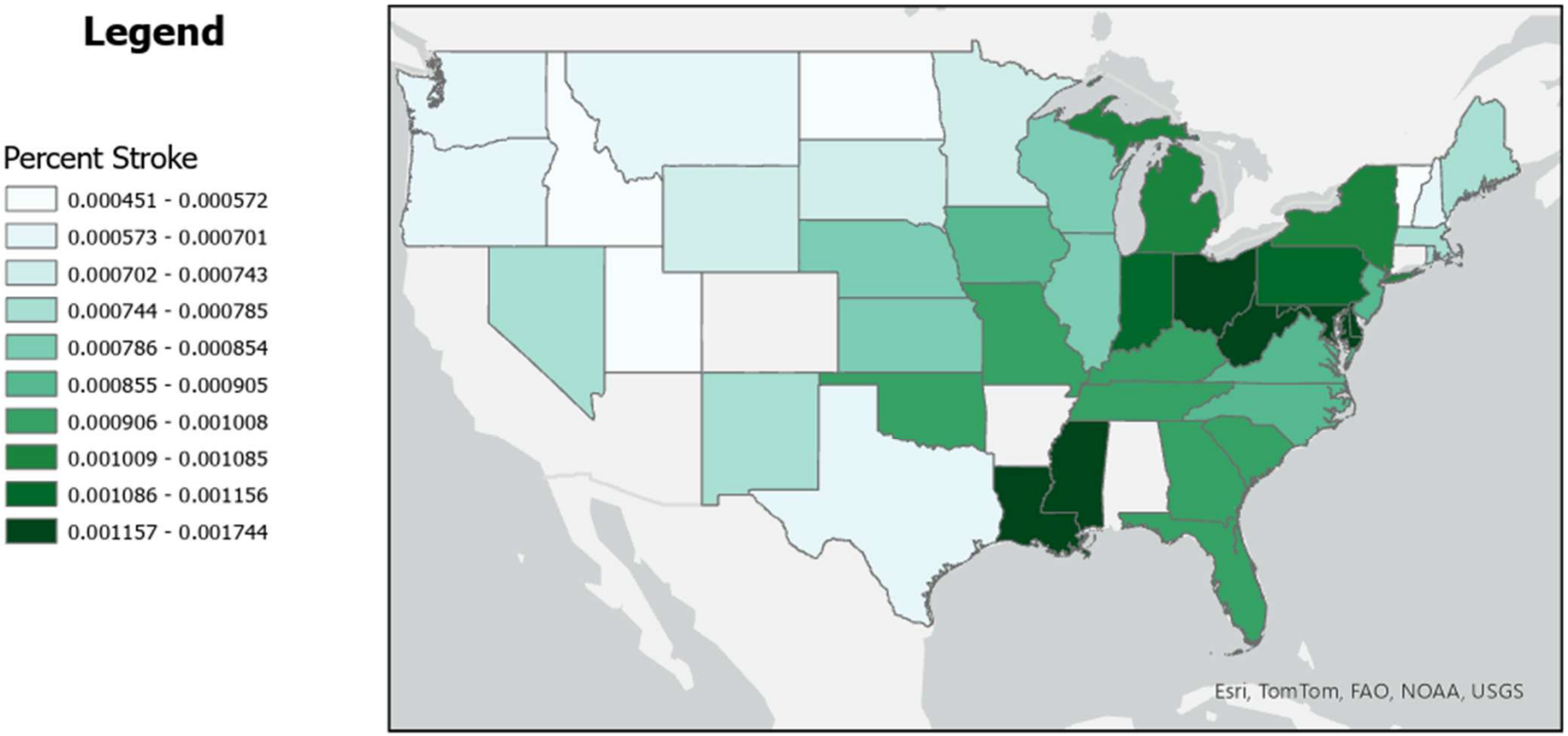
- Parks have a calculated area inside each city
- Health data is joined to each city by FIPS code



Rates of mental illness by city generalized to state



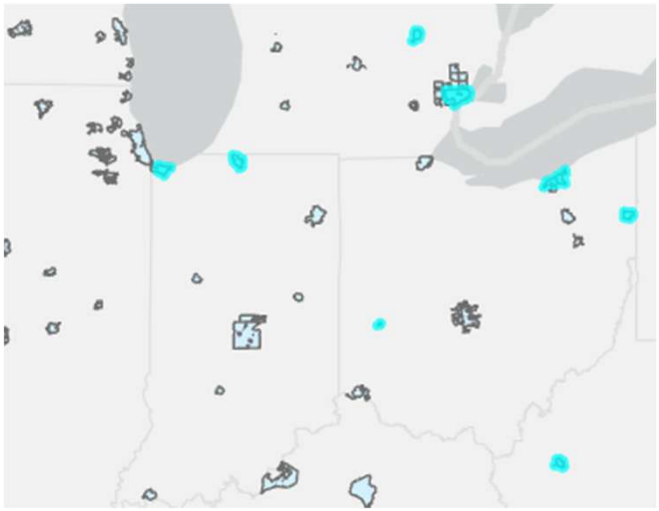
Stroke rates by city generalized to state

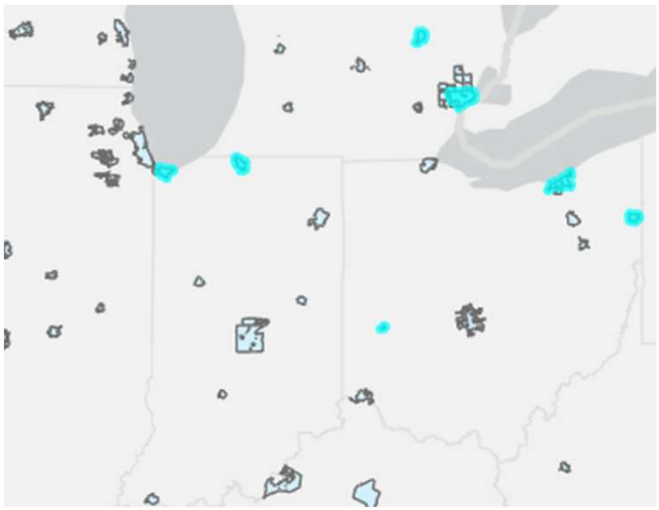


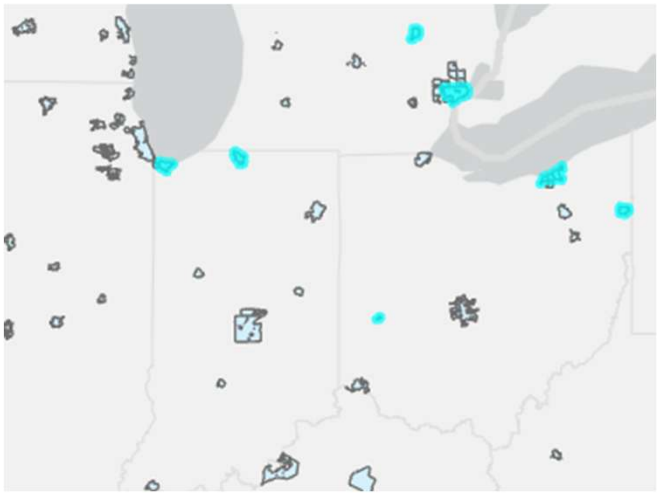
In Conclusion: No correlation

Health Condition	Correlation coefficient
Binge Drinking	0.000
Asthma	0.003
Arthritis	0.008
Cancer	0.004
Kidney Disease	0.005
High Blood Pressure	0.005
Diabetes	0.003
Mental Issues	0.001
COPD	0.010

Is there a correlation between geographical attributes of an area and health level?







Is there a significant difference in the health levels of people living within 30 miles of a lake compared to the rest of the country?

How to determine if a correlation exists?

- Using the Buffer and Pairwise Intersect tools, separate the cities
- Use the Explore Statistics tool to find the mean and standard deviation of the data for each group

Two-value Hypothesis testing

Health Issue	Mean Value Lake Cities	Stdev Lake Cities	Mean Value Non-Lake	Stdev Non-Lake	Difference in means
kidney	0.000895	0.000365	0.000821	0.000291	0.000074
asthma	0.002931	0.000918	0.002628	0.000788	0.000303
cancer	0.001735	0.000918	0.001601	0.000498	0.000134
high_bp	0.009048	0.003408	0.008549	0.002938	0.000499
arthritis	0.006920	0.002368	0.006139	0.002191	0.000781
diabetes	0.003069	0.001577	0.002868	0.001162	0.000201
COPD	0.002025	0.000931	0.001853	0.000825	0.000172
binge_drinking	0.005372	0.001400	0.004728	0.001226	0.000644
mental_health	0.004093	0.001439	0.003820	0.001245	0.000273

$$Z = \frac{(p_1 - p_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Health Condition	Z-Score	P-Value	Significant at α=0.05?
Kidney	1.437	0.151	No
Asthma	2.315	0.021	Yes
Cancer	1.074	0.283	No
High Blood Pressure	1.026	0.305	No
Arthritis	2.286	0.022	Yes
Diabetes	0.912	0.362	No
COPD	1.289	0.197	No
Binge Drinking	3.217	0.001	Yes
Mental Health	1.329	0.184	No

IOWA STATE UNIVERSITY
Department of Computer Science

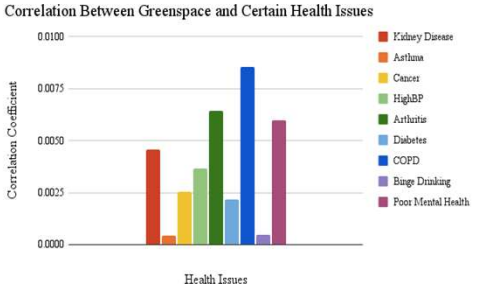
CRP 551 – Semester Project Poster Presentation
12/11/2024

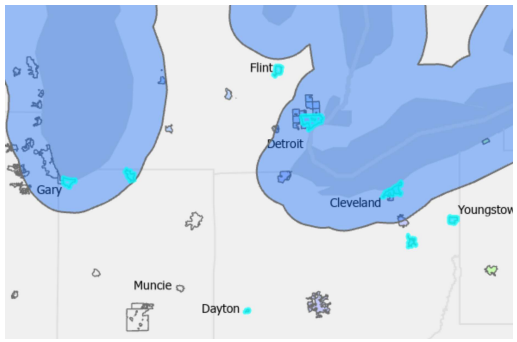
Principal investigators

An Analysis of the Relationship between Presence of Green Space in 330 Cities in the United States and the Health Status of their Citizens

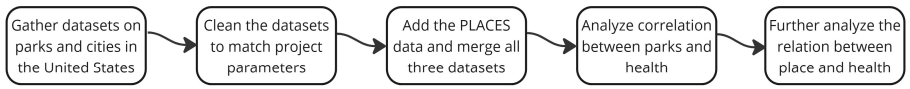
Abstract

Increasingly, scientific literature suggests that there is a positive correlation between the amount of green space in an area–the amount of plant matter growing in an area–and the health of the population living in that area. A meta-analysis published in *The Lancet Planetary Health* in 2019 found that there is a link between a higher percentage of green space in an area and the mortality rate of people living in that area. A different meta-analysis published in *Science of The Total Environment* in 2019 found a positive correlation between an area's Normalized Difference Vegetation Index (NDVI) and a decrease in the prevalence of hypertension in the population.





Methodology



Conclusion

The results of this study indicate there is less than a 0.0085 coefficient of determination between the percent of a city dedicated to parks and any of the health metrics for that city. However, regional differences in health may be acting as a confounding variable, causing the apparent lack of correlation. Specifically, the cities within 30 miles of lakes have higher levels of asthma, arthritis, and binge drinking than cities outside of that lake buffer. Further studies would be needed in to determine whether a smaller geographic scope would result in a significant correlation between the amount of greenspace and the rate of health issues.

Goal

The goal of this research was to examine the correlation between different health conditions and the prevalence of green space for the 330 largest cities in the United States. The study considered kidney disease, asthma, cancer, high blood pressure, arthritis, diabetes, Chronic Obstructive Pulmonary Disease (COPD), binge drinking, and poor mental health by city, as obtained from the CDC's PLACES survey. This study measured greenspace as the percentage of the city designated as parks. As used here, a park includes city, state, and national parks, as well as city and state forests.

Health Condition	Z-Score	P-Value	Significant at α=0.05?
Kidney	1.437	0.151	No
Asthma	2.315	0.021	Yes
Cancer	1.074	0.283	No
High Blood Pressure	1.026	0.305	No
Arthritis	2.286	0.022	Yes
Diabetes	0.912	0.362	No
COPD	1.289	0.197	No
Binge Drinking	3.217	0.001	Yes
Mental Health	1.329	0.184	No

Acknowledgements: Green Spaces and Mortality: A Systematic Review and Meta-Analysis of Cohort Studies. ESRI, USGS
Association between Greenspace and Blood Pressure: A Systematic Review and Meta-Analysis