Heat island index is the measurement of temperatures over time in a census tract, generally categorized into three communities, urban, suburban, and rural. However, urban environments are one of the areas with the greatest trend of heat island index due to air pollution, greater population density in smaller areas, and much more. With a higher heat island index, this variable discovers how it can impact the population in these urban areas, such as increasing health risks like heat-related illness. Another usage of this variable creates a relationship, whether negative or positive, with factors such as air pollution or plant population.

Plants, especially trees, can decrease surface and air temperature, “by providing shade and cooling through evaporation and transpiration, also called evapotranspiration” (US EPA). Evapotranspiration is the process where plants take in water for nourishment and thus cool the surrounding areas when they throw away the byproduct, water vapor. In this way, many can assume that a higher plant population can help mitigate a higher heat island index. There have even been studies on this subject by the Environmental Protection Agency that stated, “Researchers have found that planting deciduous species to the west is typically most effective for cooling a building... (US EPA,3).

In this project, the variables heat island index and tree crown density are used to find if there is a relationship between them in Chicago in the timespan of 2018. The hypothesis is that an increase in plant density population will decrease the heat island index temperatures, namely a strong negative correlation. However, factors such as higher human density may decrease this correlation. The data used in this project is from the Chives project, using the file GEOJSON from “https://chichives.com/data”. This project compiled data from a timespan of 2014 to 2018. The variables used were the heat island index, tree crown density, and acs\_population. The timespans for these variables were from the year 2018. The current data release was from 2023. The variables used in this project are from the census tract and tree crown polygons.

The main methods used in this project are initial summary statistics, LISA spatial autocorrelation, and correlation/hypothesis testing using the Pearson method. The start of the project was to download all the variables for the project from the ChiVes dataset. The three variables were heat island index, number of trees, and acs population. Then since percentages are easier to evaluate than whole numbers, there was a manipulation of the number of trees and ACS population variable by summing the variables and then dividing each census tract with the total population to find the percent in each census tract. With these three variables downloaded into the project, each variable was visualized using a choropleth map with 20 divisions in the percentage numbers to better visualize.

The second part of the project was to use LISA spatial autocorrelation to better understand how each variable correlated with its surroundings. The reason for using LISA in this project is to point out if there are any anomalies in the variables that cannot be seen by the maps and the summary statistics. It was also used to see how each variable worked when clustered with similar values around them. One example of this would be in the heat island index which has a negative spatial autocorrelation around the areas connected to Lake Michigan. This pattern shows that the lake could cool down areas around it.

After looking at the variables independently and understanding their relationships with their surroundings instead of with each other, a hypothesis test was used to understand the relationship between tree population and heat island index.  Looking at this relationship, as per the example of Lake Michigan and the cooling in the areas around it. Human density was a factor investigated to see if it could affect the relationship between heat island indexes. The hypothesis for this part was that human density would cause the heat island index to increase and thus make a higher plant population ineffective in lowering temperatures. These three variables were modeled by a bubble map and then overlapped to show the relationship between each other.

Chart of the methods and plan for the project

A diagram of a flowchart

Description automatically generated

The results were shown based on maps and correlation tests. Starting with the choropleth map:

Human density per census tract in Chicago 2018

IA map of a city with pink squares

Description automatically generated

Heat island index per census tract in Chicago 2018

A map of the chicago area

Description automatically generated

Tree percentage density per count in Chicago 2018

A map of a city with green and white squares

Description automatically generated

The three maps are choropleth maps of the variable’s human density, heat island index, and tree percentage.  To compare the maps, there is a big area at the bottom that shows the heat island index to be lower while there is a higher tree percentage density. This could show a negative correlation between the variables. However, when looking at the human percentage map, the big area shown also has a lower population which may have created a lower heat island index. However, another observation is on most of the left side of the heat island index map that shows a high heat island index and relatively normal human population, but it also shows that the plant population there is in the range of 0% to about 0.004 % which indicates a small plant population. This mapping could show a correlation between heat island index and plant population. Overall, the shows a trend of when plant population increases, the heat index temperature decreases and vice versa.

Next was the LISA spatial autocorrelation maps which depict positive, negative, and outlier correlations.

Human density LISA spatial autocorrelation

A map of a city with red and blue squares

Description automatically generated

Heat island index LISA spatial autocorrelation

A map of different colored areas

Description automatically generated

Plant percentage index LISA spatial autocorrelation

A map of a plant population

Description automatically generated

Each LISA spatial autocorrelation shows the correlation between the census tract and its clusters of similar values. Some observations are the bottom area touched before in the choropleth maps. In the plant population map, there is a significant high correlation at the bottom side of Chicago which when compared with the heat island index could show a negative correlation since the heat island index in that area has a negative correlation. This means a decreased temperature in that cluster. When human density is factored in, there is no significant correlation in that area implying that plant population mostly affects the heat island index in that region and not human population as what it looked like in the choropleth maps.

The last step was the look at the correlation first between the plant percentage and heat island index. This was done by calculating:

*#correlation test with plant and heat island index*

correlation = cor(hp$heatisl, hp$plant\_percentage, method = "pearson")

head(correlation)

## [1] -0.08542622

Using Pearson, which depicts a more linear relationship. The result that came out was -0.08 which is a very low negative correlation, and since it is at almost zero, there could be said to be zero correlation between heat island index and plant percentage. Thus, meaning that there is not a relationship between plant percentage and heat island index. The hypothesis for this project which states: “There is a negative correlation between heat island index and plant percentage population in Chicago around the time span of 2018.”.

Bubble map of plant percentage, heat island index, and human percentage

A map of a plant

Description automatically generated

This map shows all three variables in a bubble map to visualize the correlation between all of them. What can be shown is that there is not much of a presence of plant population while heat island index and human percentage dominate the map. The relationship between these two variables shows that the higher the human density in that area, the higher the heat island index is.

To conclude, from the final correlation test and bubble map, there does not seem to be much of a relationship between plant population and heat island index. It seems that the human population affects the heat island index more, contrary to the spatial autocorrelation and choropleth map. This could mean a few things, such as too small of a sample size, unique ecological and urban background, or a need to compare with other settings such as other urban areas or suburban/rural areas. The reason for this thinking is because places like Chicago do not have a big plant population as seen by the compact buildings that create very little space for greenery but a lot of space for humans. If there was a need to see if plant population affects the heat island index, it is best to expand the dataset outside of just Chicago and create a more diverse set with other community types across the country. Still, there may be some other limitations to this project like other environmental settings. The city of Chicago stands right next to Lake Michigan, a huge lake that can create a phenomenon called lake breeze. Lake breeze is the movement of the wind above the lake that cools the temperatures as it goes towards land. Such a factor could create a more lasting impression of heat island index in Chicago than the plant population, however, other urban areas that are not close to a body of water may have a different constitution when it comes to the relationship between heat island index and plant population. The environmental setting of any place is very complex. It would take greater and more recent data sets to create better results, it is best to find as many factors as possible that can affect one variable and see how they are interconnected with each other.

Works cited:

California, State of. “Understanding the Urban Heat Island Index.” *CalEPA*, calepa.ca.gov/climate/urban-heat-island-index-for-california/understanding-the-urban-heat-island-index/. Accessed 8 May 2024.

*EPA*, Environmental Protection Agency, www.epa.gov/green-infrastructure/reduce-urban-heat-island-effect. Accessed 8 May 2024.

*EPA*, Environmental Protection Agency, www.epa.gov/heatislands/using-trees-and-vegetation-reduce-heat-islands#:~:text=Trees%20and%20vegetation%20(e.g.%2C%20bushes,and%20transpiration%2C%20also%20called%20evapotranspiration. Accessed 8 May 2024.

Kolak, Marynia, et al. “Data.” *Data*, 21 Apr. 2023. *Chives-data-minusNU.geojson*

chichives.com/data. Accessed 8 May 2024