Investigating the Role of Green Spaces in Mitigating Urban Heat Island Effect

Yoonha Jeon
42791335
2023 WT2 CPSC 368 201

Introduction

The attention to the phenomenon of Urban Heat Island (UHI) effect has increased in recent years due to its significant implications for urban climate and sustainability. UHI is characterized by elevated temperatures in urban areas compared to their rural surroundings, driven by human activities such as industrialization, urbanization, and deforestation (Rutledge et al., 2023). As cities expand and populations grow denser, the exacerbation of UHI poses challenges for sustainability and climate resilience (Mohajerani et al., 2017). To resolve this concern of UHI, this project aims to investigate how the dynamics between urban green spaces, population density affects the temperature difference between urban and buffer areas. By examining datasets related to green spaces, population distribution, and urban heat data, the study seeks to explain the importance of green area such as parks, gardens, and other vegetated areas contributing to mitigating UHI.

Methods

To embark on our data analysis, aimed at understanding the impact of population and green spaces on urban heat levels, I collected three key datasets: the Global Urban Heat Dataset from 2013, annual population data spanning from 2010 to 2019, and park ranking data for parks across the USA. Ensuring the credibility of the selected datasets was a top priority. To achieve this, I opted for datasets gathered by reputable organizations: the UHI dataset was collected by NASA Earth Data Socioeconomic Data and Applications Center, while population data was sourced from the official website of the United States Census Bureau. Regarding green spaces, I compiled a summary dataset from various trustworthy sources, including the US Department of Energy for Green energy prevalence data, the Natural Resources Council of Maine for Open spaces and natural beauty, the U.S. Census and Environmental Protection Agency for Waste diversion and recycling data, Call2Recycle for recycling information, and the Environmental Protection Agency and Loyola University New Orleans for Social Justice and Access to Clean Outdoors and a social justice index respectively.

Before diving into the analysis, it was crucial to ensure the reliability and consistency of the datasets through several data cleaning steps. Initially, I narrowed down the population data to the year 2013 to match the collection year of the Urban Heat Island (UHI) data. Then, I converted the population data from its original text format, which included commas, to numerical values suitable for analysis. Next, to streamline the analysis process and prevent overfitting, I carefully selected 13 representative cities. These cities were chosen based on varying levels of green space per capita, as outlined in Angela Underwood's 2019 article titled "Cities with the Most Green Space per Capita". This selection allowed us to effectively demonstrate the relationship between green space, population density, and UHI. After identifying the cities, a crucial step was to harmonize the data collection levels. While UHI data was collected at the city level, population and green space data were collected at the state level. To address this inconsistency, I matched each city with its respective state, ensuring uniform granularity across all datasets for further analysis.

Results

The selected analytical approach aimed to provide robust insights into the intricate relationship between population density, green space, and UHI intensity. By employing a combination of descriptive and inferential techniques, the analysis sought to explore various dimensions of this relationship comprehensively.

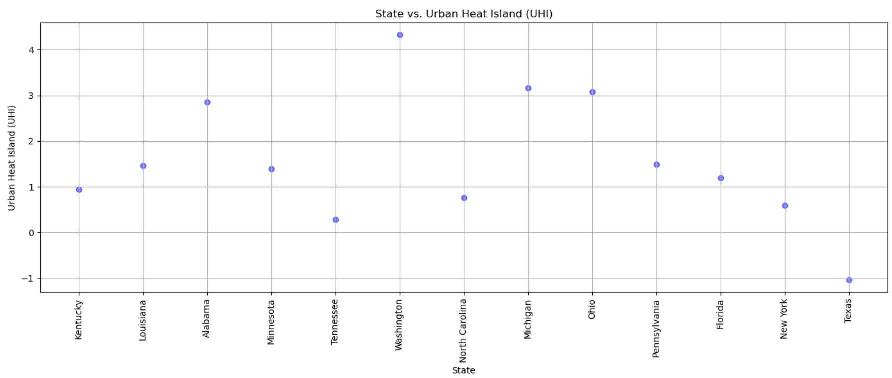

Figure 1 Scatter plot of UHI in 13 different cities in USA

Scatter plots were utilized to visualize the distribution of population density against UHI intensity, revealing no discernible pattern and suggesting a lack of significant correlation between the variables.

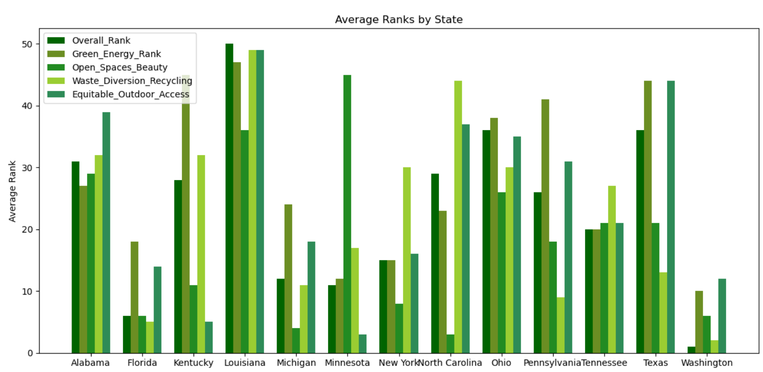

Figure 2 Side by side bar plot of green space rankings in 13 different cities in USA

Side-by-side bar plots were then generated to compare green space rankings across different states, highlighting variations in green space distribution and indicating a weak correlation in the ranking of green spaces.

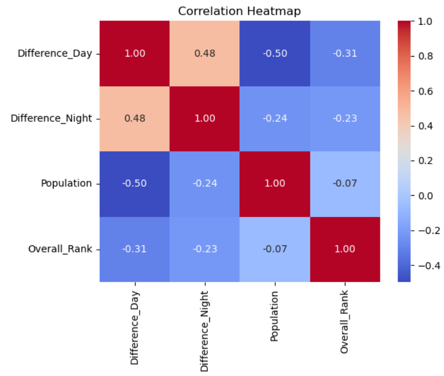

Figure 3 Correlation heat map

A correlation heat map was subsequently constructed to visualize the relationships between population density, green space, and UHI intensity. Despite weak correlations, the analysis supported the hypothesis that higher green space rankings and lower population densities corresponded to higher temperature differences between urban and buffer areas.

```
                            OLS Regression Results
==============================================================================
Dep. Variable:         Difference_Day   R-squared:                       0.362
Model:                            OLS   Adj. R-squared:                  0.235
Method:                 Least Squares   F-statistic:                     2.840
Date:                Wed, 27 Mar 2024   Prob (F-statistic):              0.106
Time:                        00:48:04   Log-Likelihood:                -19.666
No. Observations:                  13   AIC:                             45.33
Df Residuals:                      10   BIC:                             47.03
Df Model:                           2
Covariance Type:            nonrobust
================================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------
const            3.5641      0.926      3.847      0.003       1.500       5.628
Population   -1.068e-07    5.2e-08     -2.054      0.067   -2.23e-07    9.07e-09
Overall_Rank    -0.0352      0.026     -1.348      0.207      -0.093       0.023
==============================================================================
Omnibus:                        0.321   Durbin-Watson:                   2.460
Prob(Omnibus):                  0.852   Jarque-Bera (JB):                0.461
Skew:                           0.192   Prob(JB):                        0.794
Kurtosis:                       2.162   Cond. No.                     3.43e+07
==============================================================================
```

Table 1 Summary of MLR

A multiple linear regression model was employed to assess the predictive capability of the data. The model yielded a low adjusted R-squared value, indicating limited predictive power, and confidence intervals encompassing 0 suggested minimal impact of explanatory variables on UHI intensity.

In summarizing the analysis, the visualizations offer valuable insights into the intricate interplay among urban heat island (UHI), population, and green spaces. However, the results suggest a weak influence of green spaces and population on UHI, which may be attributed to limitations encountered during the analysis. Specifically, the discrepancy arises from the disparate collection levels of population and green space data at the state level, while UHI data were gathered solely from urban areas. This limitation underscores the need for future research where data is collected at a consistent level, potentially yielding statistically significant results. In forthcoming studies, I aim to revisit this research using datasets collected at uniform levels to further explore these dynamics.

Discussion

In conclusion, the data analysis conducted on the Urban Heat Island effect has yielded valuable insights, indicating that as population increases and green space decreases, the temperature gap between urban and buffer areas widens. However, an area for improvement lies in the database selection; while the project aimed for enhanced accessibility using Oracle, SQLite was utilized instead. Given more time, delving into Oracle database study and implementation would be a priority. Furthermore, future investigations into additional factors impacting the UHI effect beyond population and green space could further enhance the depth of analysis.

Long-term data storage in this project raises concerns about both data integrity and storage scalability. Given the involvement of multiple datasets from diverse sources, ensuring the integrity and consistency of the data over time is challenging, with potential changes or updates to datasets and formats impacting analysis compatibility. Additionally, the potential increase in data volume over time necessitates a storage solution capable of scaling to prevent data loss or performance issues. To preserve data provenance, comprehensive documentation, version control mechanisms, and metadata management strategies are essential, ensuring transparency, reproducibility, and trustworthiness for future researchers and collaborators.

For my project, a database was used instead of using multiple data files because database management system offers several benefits such as better data integration, minimized data inconsistency, and data abstraction. According to an article from Geeks for Geeks in 2013, database management systems organize and sync data neatly, making it much simpler to handle. They also provide a clear view of the data and different levels of abstraction, which makes it easier for users to interact with. Since my project involved using three datasets and analyzing them, it was important to have consistent data and an easy way to work with it. That was the reason for using a database was a better choice than using separate data file.

References

GeeksforGeeks. (2023, December 4). *Advantages of Database Management System*. https://www.geeksforgeeks.org/advantages-of-database-management-system/

Mohajerani, A., Bakaric, J., & Jeffrey-Bailey, T. (2017). The urban heat island effect, its causes, and mitigation, with reference to the thermal properties of Asphalt Concrete. *Journal of Environmental Management*, *197*, 522–538. https://doi.org/10.1016/j.jenvman.2017.03.095

MPHONLINE. (2021, April 8). *The Greenest States in America*. MPH Online. https://www.mphonline.org/green-states/

NASA Socioeconomic Data and Applications Center. (2013). *Satellite-derived environmental indicators*. SEDAC. https://sedac.ciesin.columbia.edu/data/set/sdei-global-uhi-2013

Rutledge, K., McDaniel, M., Teng, S., Hall, H., Ramroop, T., Sprout, E., Hunt, J., Boudreau, D., & Costa, H. (2023). *Urban heat island*. National Geographic. https://education.nationalgeographic.org/resource/urban-heat-island/

Underwood, A. (2019, July 26). *Cities with the most green space per capita*. Stacker. https://stacker.com/environment/cities-most-green-space-capita

US Census. (2021, December 16). *State Population Totals: 2010-2019*. Census.gov. https://www.census.gov/data/datasets/time-series/demo/popest/2010s-state-total.html