Urban Heat Island Effect Project Report

CHIA YUE ZHU AMELINE, A0211277L

GOH WEI JIE DARREN , A0265959E

RAKHSHANA PARVEEN , A0227119H

WILLIAM JOSE , A0245394X

2024-11-17

# INTRODUCTION

For this report, the **Problem Statement is: Which factors (population density, energy consumption, and green area size) most significantly impact the intensity of urban temperature?** With urban areas worldwide experiencing rising temperatures due to industrialization and rapid urbanization, it is crucial to understand and mitigate the factors driving urban heat. Therefore, this project seeks to analyze and predict urban heat island (UHI) intensity by examining how population density, energy consumption, and green space availability contribute to temperature variations across diverse geographic regions.

Using datasets download from Kaggle on temperature, population, energy consumption, and land use, we will explore the correlation between population density and urban temperatures, assess the influence of various environmental and socio-economic variables on UHI, and examine how energy consumption exacerbates UHI effects. By integrating these datasets and conducting exploratory data analysis, we aim to build predictive models that accurately estimate UHI intensity, enabling policymakers and urban planners to prioritize effective strategies to mitigate UHI impacts.

Ultimately, the findings from this project will provide public policymakers and urban planners with insights into how urban environmental factors—such as population density, green area size, and energy consumption—affect temperature variations in cities. These insights can guide the development of strategies for urban heat mitigation, support the creation of climate-adaptive policies, and inform sustainable urban planning practices focused on managing and reducing urban heat intensity.

# BACKGROUND

Urban areas worldwide are experiencing rising temperatures, often due to industrialization and rapid urbanization. For this reason, many cities are significantly warmer than their surrounding rural areas. According to a study conducted by the **Queensland University of Technology,** *titled -* [*Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures*]((https://eprints.qut.edu.au/115400/1/Manuscript%20Accepted%20Version.pdf)), several factors contribute to UHI, including population density, energy consumption, and the lack of green spaces. This study systematically reviews and highlights how spatial and temporal factors influence temperature intensity and emphasizing the importance to analyze these variables effectively.

As industrialization and urbanization continue to grow, it becomes increasingly important to examine how various factors contribute to UHI. Therefore, this report hypothesizes that in high-density urban areas, the combination of built infrastructure and human activities results in higher temperatures due to reduced green space and increased heat emissions. Accordingly, our study will aim to investigate this hypothesis by analyzing the underlying relationships between these variables and UHI intensity.

# DATA

## STEP 01 - DATA COLLECTION

The data used in this project comes from two primary sources: R’s Natural Earth data package and Kaggle.

During the data sourcing process, we encountered challenges finding comprehensive spatial polygon dataframes that aligned with our project requirements. Many available datasets we came across on Kaggle and NASA earth data either lacked polygon scales or, included them but did not contain the spatial attributes needed to address our problem statement. As a result, using R libraries, we downloaded world and city data from the Natural Earth data package and integrated the spatial data libraries with datasets from Kaggle to form spatial polygon dataframes for our analysis.

On the Kaggle platform, we downloaded five additional datasets. Two of these represent average temperatures across various countries and cities over time, while the remaining three datasets provide supplementary variables that could influence temperatures in these areas. Specifically:

* **The population dataset** includes information such as country area size, population density, total population, and grow \*\*th rate
* **The energy dataset** provides details on energy production and consumption within each country across multiple decades
* **The green area dataset** contains information on the extent of green spaces within countries over the years

## STEP 02 - DATA CLEANING AND TRANSFORMATION

### DATA CLEANING

Throughout the data cleaning process, we check and remove records with values such as:

* “NA” or NA values
* Duplicated records
* Inconsistent data records:

—– Explanation ——

## DATA TRANSFORMATION

—– Explanation ——

# EXPLORATORY DATA ANALYSIS

—– Explanation ——

# METHODS AND MODELINGS

—– Explanation ——

# ANALYSIS

—– Explanation ——

# CONCLUSION

—– Explanation ——

# EXTRA MILE

—– Explanation ——