October 27, 2025

1 Final Project Proposal: Examining the Relationship between Green Space Access, Income, and Health

4 November 2025

Authors

Name	Student #	Email
Patrick A. Mikkelsen	1010572514	patrick.mikkelsen@utoronto.ca
Hugo Cheng	1009827323	hugo.cheng@mail.utoronto.ca
Mike McCracken	[Mike please add yours]	mike.mccracken@utoronto.ca

Course Information **Instructor:** Professor Ignacio Tiznado-Aitken

Course: GGR375H1 F: Introduction to Programming in GIS

TA: Evan Powers

1.1 1. Introduction and Research Question

Urban green spaces—including parks, trees, and natural vegetation—provide substantial benefits for public health and environmental resilience. Research demonstrates that residential greenness is associated with reduced mortality risk, improved cardiovascular health, better mental wellbeing, and protection against adverse environmental conditions such as extreme heat and flooding (Pinault et al. 2021). The COVID-19 pandemic further underscored the critical importance of accessible green spaces for physical and mental health, as urban residents increasingly relied on nearby parks and natural areas for recreation, stress relief, and social connection (Vabi 2022).

However, access to these vital environmental resources is not equitably distributed across urban populations. Studies across North America have documented significant disparities in residential greenness according to socioeconomic status and ethnocultural identity, with lower-income households and racialized communities consistently experiencing reduced access to quality green space (Pinault et al. 2021). In Toronto specifically, research has shown that low-income and racialized

communities have significantly fewer trees and less access to parks compared to wealthier, predominantly white neighborhoods (Ward 2020).

These environmental inequities represent an urgent matter of environmental justice, as they may contribute to producing or exacerbating existing health inequalities in urban populations (Vabi 2022; Pinault et al. 2021). Understanding the spatial patterns and demographic correlates of green space access is essential for informing urban planning policies that promote environmental equity and public health. This study aims to examine the relationship between population demographics, income levels, and access to green space across Toronto's census tracts. Building on previous literature documenting green space inequities (Ward 2020; Pinault et al. 2021; Vabi 2022), we employ geospatial analysis techniques to investigate how socioeconomic and demographic factors correlate with the distribution of parks, trees, and vegetated areas throughout the city.

1.1.1 Research Question

How does access to green space vary across Toronto census tracts according to household income and population demographics, and what spatial patterns of environmental inequity can be identified through geospatial analysis?

1.2 2. Background and Motivation

[Explain why this topic is important and what gap in knowledge you're addressing]

1.3 3. Data Sources

[List the datasets you plan to use, including: - Dataset names - Sources/URLs - Spatial resolution - Temporal coverage - Key variables]

1.3.1 Required Libraries

This analysis utilizes Python libraries for geospatial analysis (GeoPandas), statistical computing (SciPy, NumPy), data manipulation (Pandas), and visualization (Matplotlib, Seaborn, Folium).

[1]:
```
import geopandas as gpd
import pandas as pd
import shapely as shp
import folium
import numpy as np
from scipy import stats
import matplotlib.pyplot as plt
import seaborn as sns

# Set visualization style
sns.set_style("whitegrid")
plt.rcParams['figure.figsize'] = (12, 6)
plt.rcParams['font.size'] = 10
```

[ ]:
```
```

```
census_boundaries_path = "Green Space & Income/data/ldb_000a21a_e/ldb_000a21a_e.
 ↪shp"
census_data_path = "Green Space & Income/data/98-401-X2021007_eng_CSV/
 ↪98-401-X2021007_English_CSV_data.csv"
census_sd_path = "Green Space & Income/data/lcsd000b21a_e/lcsd000b21a_e.shp"
```

1.3.2 Census Tract Boundaries and Spatial Delimitation to Toronto

Census tract boundaries from Statistics Canada's 2021 Census are loaded and transformed to the WGS84 coordinate reference system (EPSG:4326). Census tracts represent small, relatively stable geographic units designed to be homogeneous with respect to population characteristics, economic status, and living conditions.

To focus the analysis on Toronto proper (Census Subdivision Code 3520005), census tract boundaries are immediately clipped to the municipal boundary, excluding surrounding municipalities within the Greater Toronto Area.

[ ]:
```
# Load census tract boundaries
census_boundaries = gpd.read_file(census_boundaries_path)
census_boundaries = census_boundaries.to_crs(epsg=4326)

# Load census subdivision boundaries and subset to Toronto
csd_boundaries = gpd.read_file(census_sd_path)
toronto_csd = csd_boundaries[csd_boundaries['CSDUID'] == '3520005'].copy()

# Ensure CRS match
if toronto_csd.crs != census_boundaries.crs:
    toronto_csd = toronto_csd.to_crs(census_boundaries.crs)

# Clip census boundaries to Toronto
toronto_boundaries = gpd.clip(census_boundaries, toronto_csd)
toronto_boundaries = toronto_boundaries.to_crs(epsg=4326)

# Filter to only polygon/multipolygon geometries (exclude any line features)
toronto_boundaries = toronto_boundaries[toronto_boundaries.geometry.type.
 ↪isin(['Polygon', 'MultiPolygon'])].copy()

# Save the Toronto census tracts
toronto_boundaries.to_file("my_data/toronto_census_tracts.shp")

print(f"Loaded {len(census_boundaries)} census tracts total")
print(f"Subset to {len(toronto_boundaries)} census tracts in Toronto")to
```

```
Loaded 6247 census tracts total
Subset to 585 census tracts in Toronto
```

You're gonna wanna restart your kernel now...

[2]:
```
toronto_boundaries = gpd.read_file("my_data/toronto_census_tracts.shp")
toronto_boundaries.explore()
```

[2]:
```
<folium.folium.Map at 0x75f530827f50>
```

1.3.3 Census Demographic and Economic Data

The 2021 Census Profile dataset contains demographic and socioeconomic characteristics for all census geographic areas in Canada. Census tract-level records are extracted for subsequent analysis.

[7]:
```
census_data = pd.read_csv(census_data_path,
                          encoding="latin1", low_memory=False)
ct_data = census_data[census_data['GEO_LEVEL'] == 'Census tract'].copy()
```

```
The Kernel crashed while executing code in the current cell or a previous cell.


Please review the code in the cell(s) to identify a possible cause of the␣
 ↪failure.

Click <a href='https://aka.ms/vscodeJupyterKernelCrash'>here</a> for more info.


View Jupyter <a href='command:jupyter.viewOutput'>log</a> for further details.
```

1.3.4 Population and Income Variables

Population counts and median household income data are extracted from the census profile. These variables serve as the dependent and independent measures for analyzing green space equity.

[ ]:
```
population_data = toronto_ct_data[
    (toronto_ct_data['CHARACTERISTIC_NAME'] == 'Population, 2021')
][['DGUID', 'C1_COUNT_TOTAL']].copy()
population_data.rename(columns={'C1_COUNT_TOTAL': 'POPULATION'}, inplace=True)
population_data['POPULATION'] = pd.to_numeric(population_data['POPULATION'],␣
 ↪errors='coerce')

income_data = toronto_ct_data[
    toronto_ct_data['CHARACTERISTIC_NAME'].str.contains('Median total income of␣
 ↪household', case=False, na=False)
][['DGUID', 'C1_COUNT_TOTAL']].copy()
income_data.rename(columns={'C1_COUNT_TOTAL': 'MEDIAN_INCOME'}, inplace=True)
income_data['MEDIAN_INCOME'] = pd.to_numeric(income_data['MEDIAN_INCOME'],␣
 ↪errors='coerce')
```

1.3.5 Green Space Data

Municipal green space data includes parks, ravines, golf courses, and other vegetated areas within Toronto's boundaries. Spatial data are standardized to the WGS84 coordinate system for compatibility with census boundaries.

[ ]:
```
green_spaces = gpd.read_file("Green Spaces - 4326/Green Spaces - 4326.shp")
```

1.4 4. Methodology

[Outline your analytical approach: - Data preprocessing steps - Spatial analysis techniques - Statistical methods - Python packages you'll use (e.g., geopandas, rasterio, scikit-learn)]

1.5 5. Expected Outputs

[Describe what you plan to produce: - Maps/visualizations - Statistical results - Python scripts/modules]

1.6 6. Timeline

[Break down the project into tasks with tentative deadlines]

1.7 7. Challenges and Limitations

[Identify potential obstacles and how you plan to address them]

1.8 8. References

Pinault, Lauren, Tanya Christidis, Olaniyan Toyib, and Dan L. Crouse. 2021. "Ethnocultural and Socioeconomic Disparities in Exposure to Residential Greenness within Urban Canada." Health Reports (Ottawa, Canada) 32 (5): 3–14. https://doi.org/10.25318/82-003-x202100500001-eng.

Vabi, Vilbert. 2022. "Parks and Forests Are Missing in Marginalized Neighbourhoods." Nature Canada, March 18. https://naturecanada.ca/news/blog/parks-and-forests-are-missing-in-marginalized-neighbourhoods/.

Ward, Christine. 2020. "Toronto's Low-Income and Racialized Communities Have Fewer Trees: U of T Researchers | University of Toronto." News. U of T News, October 26. https://www.utoronto.ca/news/toronto-s-low-income-and-racialized-communities-have-fewer-trees-u-t-researchers.