Executive Summary

MQE Capstone for the Pittsburgh Parks Conservancy: Measuring and Mapping Community Investment Need
Authors: Lap Pham, Patrick Swain, José de los Ríos, Maxwell Snodgrass
University of Pittsburgh Department of Economics, April 2025

Overview: This capstone project for the Pittsburgh Parks Conservancy presents a data-driven framework to support equitable investment decisions in city parks. Developed by Master of Quantitative Economics candidates at the University of Pittsburgh, the project analyzes and visualizes standardized data on demographics, environment, public health, and crime across Pittsburgh's parks and their associated walksheds, defined as the area within a five-minute walk of a park.

Methodology: Geographic data at the Census tract level was aligned to custom park and walkshed geometries using areal interpolation, enabling the transfer of population-weighted estimates from larger units to finer, park-specific zones. Quantitative variables were cleaned, aggregated, and standardized using z-scores, with additional metrics generated such as counts, proportions, index values, and ranks. Visualization tools included R (ggplot2, sf, tmap), Python (seaborn, matplotlib), and QGIS. The deliverables include a master dataset, ranked index values, comparative graphics, and interactive choropleth maps.

Key Findings:

- Census data was used to evaluate race, age, poverty, and housing vacancy across walksheds. The most consistent pattern was a strong negative association between Black and White population proportions, highlighting the persistence of racial residential segregation in Pittsburgh. Walksheds with the highest concentrations of Black residents — including Homewood, the Hill District, Knoxville, and the North Side — also exhibited higher poverty rates and housing vacancy. Walksheds with the highest median age were concentrated near senior centers (e.g., Glen Hazel, Overbrook), while the youngest walksheds were in student-heavy neighborhoods like Oakland and the Bluff.
- Environmental analysis showed that tree canopy coverage varied widely by park and neighborhood. The average park tree canopy was 26.4%, but 25 parks had at least 50% lower canopy than their surrounding walkshed. Parks like Hays Woods and Riverview Park had the highest coverage, while urban parks such as Southside Market House and Denny Park had the least. Air quality was relatively uniform, but worse in North Side parks. The only significant outlier was Warrington Park, which had notably low pollution levels. Sewershed priority, measured using PWSA classifications, revealed that 24 parks (14.3%) fall entirely or predominantly within high-priority sewershed areas.
- Crime data from 2016-2023 showed that nonviolent crime occurred more frequently than violent crime. Crime per acre was calculated for each park and walkshed. Parks like Dallas Park and Southside Market House had exceptionally high crime densities in single years, while walksheds around Market Square and Mellon Square consistently ranked highest over time.
- Public health indicators — specifically asthma, obesity, depression, and diabetes — were evaluated using CDC estimates at the tract level. These were spatially joined to walkshed geometries and population-weighted. Manchester Park and other North Side walksheds appeared in the top ranks for asthma, depression, and obesity prevalence, both in relative and absolute terms. Chadwick Park, Baxter Park, and the Homewood Senior Center walksheds had the highest relative rates of diabetes.
- In correlation analyses, several strong relationships emerged: asthma and obesity and vacancy and violent crime had coefficients above 0.5. Black population share and poverty, vacancy, diabetes, and obesity all had positive correlations. Negative correlations were also found between White population share and diabetes rates, and between tree canopy and total and nonviolent crime.

Conclusion: Rather than calculate a single investment index, the project preserves flexibility for the Conservancy to apply its own weights – ultimately value-based policy decisions. The provided data and visualizations enable transparent, replicable comparisons between parks, empowering decision-makers to align investments with community-specific needs across environmental, social, and health dimensions.

Deliverable: The shared Google Drive folder contains a document detailing the data, methodology, and analysis of the project, along with subfolders of datasets, visualizations, and code. Parks_Master.csv contains population- and size-scaled information for relevant variables across all parks and walksheds.