Center for Urban Science + Progress

NYU | TANDON

Understanding the Dynamics of Social Encounters in NYC Public Spaces Using Human Mobility Data

Ruixin Gan rg4743@nyu.edu
Zheyan Chen zc3053@nyu.edu
Yin Wang yw7422@nyu.edu

Capstone Instructor: Takahiro Yabe, Graham Dove
Mentor: Vaidehi Raipat

Background

Each summer, around 350 NYC residents die from heat stress, often at home. This study analyzes park population patterns under heat to improve design.

Research Question 1: How does park visitation (visitors count) differ between extremely hot and normal temperature days or hours during the summer?

Hypothesis 1: Visitation decreases during hot hours but increases during cooler hours.

Research Question 2: How is visitor distributed across the park under different weather conditions (e.g., near shaded areas or water features)?

Hypothesis 2: Visitors concentrate near shaded areas or water features during heat.

Methodology

Data Preprocessing

Hourly visitation estimates were derived by redistributing daily totals using day-of-week proportional weights to preserve temporal patterns. The Heat Index (HI) was calculated using daily temperature and relative humidity data, Heat Index values were then classified into four categories (Extreme Hot, Hot, Average, and Cool) based on thresholds established by the New York City Emergency Management. Park-specific features were explicitly noted to contextualize results. Points of Interest (POIs) were categorized by shade coverage, water proximity, and functional use using NYC Parks Zone Map data.

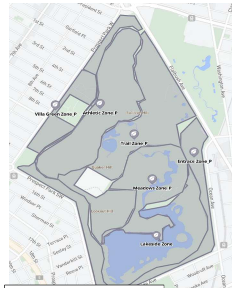
Model: Multi Linear Regression

This study employs multivariable Ordinary Least Squares (OLS) regression to analyze park visitation patterns, focusing on environmental, temporal, and social predictors.

Categorical variables were converted to binary (0/1) variables to isolate their effects. Multicollinearity was addressed via Variance Inflation Factor (VIF) analysis (threshold: VIF < 5), with moderate inflation observed only in Cool and Water_Area (VIF = 4.45, 4.60)

Stepwise feature selection prioritized variables with $p < 0.05$, though non-significant predictors (Average Temperature, Small Events) were retained for theoretical completeness.
Park-specific adjustments included merging overlapping zones (e.g., Water Area and Shaded Area in Kissena Park) and adding contextual variables (e.g., Recreation Zone in Flushing Meadows).

Flushing Meadows Corona park POI



Prospect Park POI

Kissena Park POI

Results

We found strong variations in park visitation patterns in response to weather, time, and event conditions across the three parks. Visitation consistently decreased during extreme heat and early mornings, supporting Hypothesis 1. For example, morning visitation dropped by 48.7% in Flushing, 52.8% in Prospect, and 22.2% in Kissena Park. Conversely, weekends boosted visitation across all sites by over 50%, while large events led to dramatic spikes—up to +983% in Flushing Park.

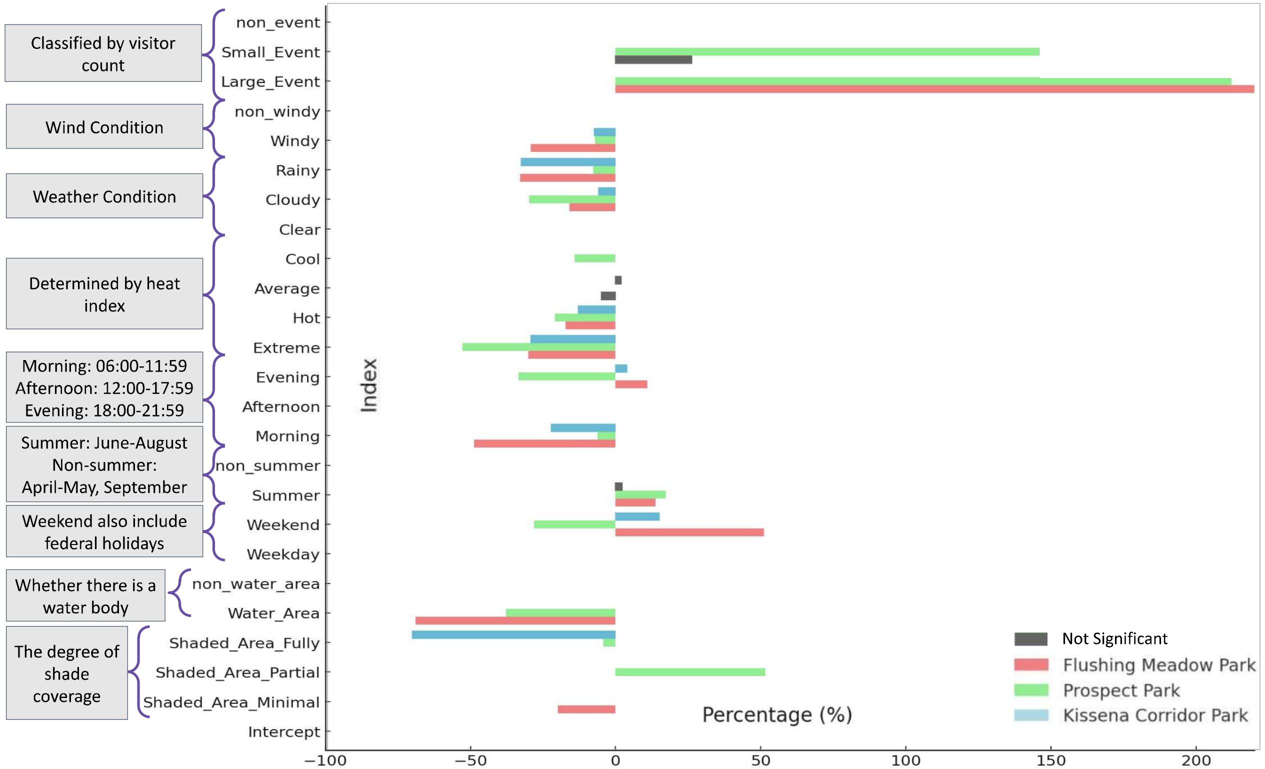
Qualitative patterns also revealed behavioral shifts. Visitors preferred cooler times of day and tended to avoid fully shaded or waterfront areas, likely due to activity constraints in those spaces. Kissena Park showed fewer positive responses to weather and event changes, suggesting different usage dynamics.

Our regression models had moderate explanatory power (R^2 ~0.31–0.45), validating the relevance of selected predictors while acknowledging unexplained variance. Despite some statistically insignificant variables, all hypothesized factors were retained to preserve interpretability.

Discussion

Our findings support Hypothesis 1: visitation decreases during hot hours and increases during cooler times. However, Hypothesis 2 was not supported. Fully shaded and waterfront areas, though thermally comfortable, saw lower visitation—likely due to limited access or activity options. This suggests that thermal comfort alone doesn't drive behaviour; functionality and design matter. While models explained moderate variance (R^2 = 0.31–0.45), unobserved factors likely play a role. Future designs should combine shade and cooling features with inviting, usable park spaces.

Percentage Bar Plot



Datasets & References

National Oceanic and Atmospheric Administration National Weather Service. (2024). NOAA Online Weather Data. United States Department of Commerce. https://www.weather.gov/wrh/Climate?wfo=okx

The Weather Company. (2024). New York City, NY Weather History. https://www.wunderground.com/history/monthly/us/ny/new-york-city/KLGA/date/2024-6

Placer Labs Inc. (2024). Visitation Insights for Complexes. https://analytics.placer.ai/

NYC Department of Parks & Recreation. (2024). Free and Low-Cost Events: NYC Parks. https://www.nycgovparks.org/events/keyword%2520Flushing%2520Meadows%2520Corona%2520Park/f2024-06-01/t2024-08-31

Acknowledgements

We would like to express our gratitude to Professor Takahiro Yabe, Professor Graham Dove and PhD Student Vaidehi Raipat for for their help and advice that made this research possible.