Understanding the Relationship Between Infrastructure and Women's Public Safety: A Case Study of New York City

Haobing Liu

Stuart Weitzman School of Design, University of Pennsylvani

Instructor: Elizabeth Delmelle

CPLN 6800: MUSA Capstone

Abstract

In New York City, 203,516 incidents of violence against women occurred during 730 days (2020 - 2022), according to incomplete data from the New York City Police Department (NYPD), and 25.88 percent of incidents happened in public places. This alarming statistic emphasizes the urgency of addressing women's safety in urban spaces. This study explores the association between urban infrastructure and the number of crimes affecting women in public and transportation places in New York City. Utilizing the NYPD Complaint dataset, the study scrutinizes the influence of damaged streetlights, pavement conditions, vacant and unsecured buildings, and the distance to women's resources, facilities, and subway stations on women's safety. Implementing linear regression and random forest models and spatial analysis of various datasets, this paper emphasizes the significance of gender equality considerations in urban infrastructure development, particularly in public space settings. The findings underscore the necessity for enhanced urban planning policies and safer urban environments for women while offering valuable insights into the efficacy of existing infrastructure planning.

Keywords: urban infrastructure inequality, women's public safety, spatial analysis, public crimes, New York City

1. Introduction

A research survey [1] in 2022 stated that 32 percent of women respondents feel unsafe in public spaces at night. Another study [2] of 28 global cities that found women were 10 percent more likely than men to feel unsafe in metros and 6 percent more likely to feel unsafe on buses. Space is fundamental in any form of communal life and exercise of power [3]. According to James Scott's theory in 1988 and Michael Mann's concept of infrastructure power in 1984, infrastructure is one of the primary instruments the state uses to organize society [4].

2. Literature Review

There are many different forms of infrastructural violence and gender inequality in the way that our cities are planned. Gender power relations are constantly being rearranged, reshaped, embodied, and embedded in even the most common urban infrastructures.[5] The fact that public infrastructures in urban spaces are touched and experienced physically reinforces the idea of infrastructural violence, forcing women to contend with the limitations imposed by time and space. Fear of violence can undermine women's confidence and limit their activity accessibility in public spaces [6]. It is unacceptable that women should be forced to give up their right to access public spaces out of fear. For example, female cyclists are more concerned about overall safety than male cyclists [7]. Additionally, it has an impact on women's activity duration, which means women are able to remain outside until late in the evening without feeling anxious or depressed because of the darkness. Therefore, the requirements for women's safety must be taken into consideration when developing infrastructure. Characteristics providing prospect, escape, and sufficient lighting should already be considered at the early stage of infrastructure planning[8]. A good public space atmosphere in a neighborhood has, from the point of view of giving women a sense of being safe and comfortable, the following characteristics:

Sufficient illumination: Street lighting affects people's perception of safety, and people prefer the light in their own surroundings [9]. Proper street lighting

is a major factor in enhancing street safety[10]. Research showed that fears in relation to these features might be reflected on campuses and public spaces ('Hot spots' of fear and crime), where policy calls for the lighting of dark pathways, and lighting and cameras in parking lots[11].

Openness: Unobstructed views and ample space with good all-around visibility ensure safety ahead[12].

Utilizing public facilities freely: It refers to the freedom to use public amenities like public restrooms, take public transit, walk alone, ride a bicycle, and so on.

Social Interaction: Activities are accessible for women within walking distance. It takes a lot of energy for women to address safety issues before they can start enjoying a safe, active lifestyle[13].

This article explores the association between physical infrastructure (damaged streetlights, pavement condition, vacant and unsecured buildings, the distance to women's resources, facilities, and subway stations) and the public violence against women in New York City. It investigates a spatial issue in urban planning, which is the overlooked requirement for women's safety in infrastructure construction.

3. Data

In this study, the primary dataset used is the NYPD complaint dataset, which offers extensive information on reported incidents1 of violence against women in New York City, revealing 203,516 incidents between 2020 and 2022.

The original dataset is categorized and reclassified into five groups: Outdoor, Residential, Commercial, Transport, and Other 2. Specifically, the Outdoor category encompasses streets, parking lots or garages (private), parks or playgrounds, parking lots or garages (public), highways or parkways, abandoned buildings, bridges, tunnels, and open areas (open lots, construction sites). The Residential category includes various types of residences, homeless shelters, pri-

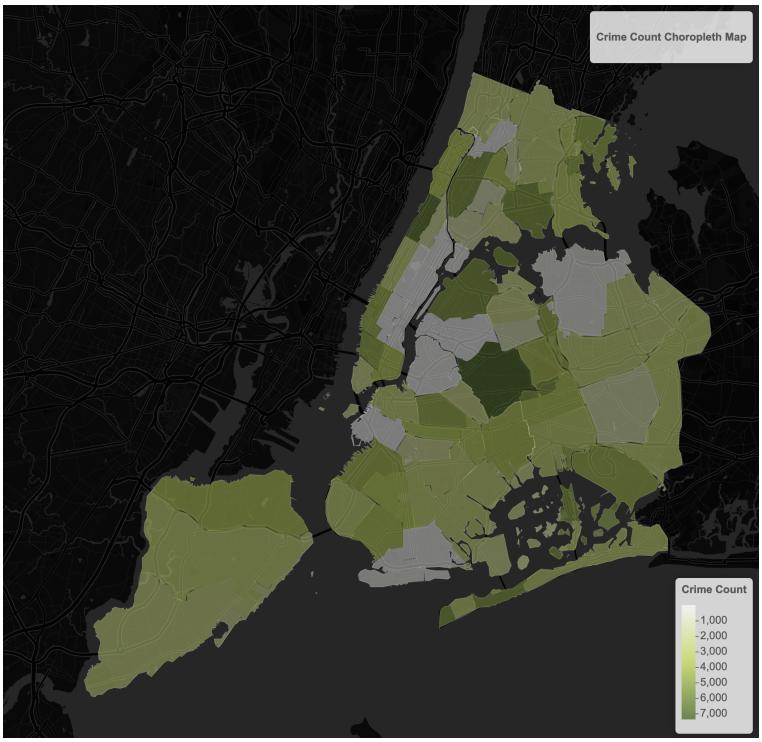

Figure 1: The correlation between the number of crimes and infrastructure features

vate or parochial schools, hotels or motels, and real estate establishments. The Commercial category covers a wide range of businesses, including restaurants, retail stores, banks, and service providers. The Transportation category includes bus stops, transit facilities, buses, taxis, and tramways. In this report, crimes occurring in outdoor and transportation spaces are defined as "public crimes"3.

Furthermore, several datasets between 2020 and 2022 are integrated to provide a comprehensive understanding of urban infrastructure factors potentially associated with crime rates impacting women. These datasets include streetlight condition records from 311 requests, valid facilities data collected by the Department of City Planning, pavement rating data from the Department of Transportation, and information on vacant and unsecured buildings.

The layout, design, and quality of streets and sidewalks considerably influ-

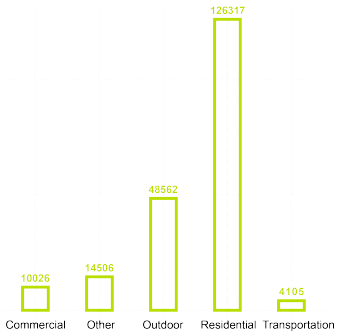

Figure 2: The number of crimes by location

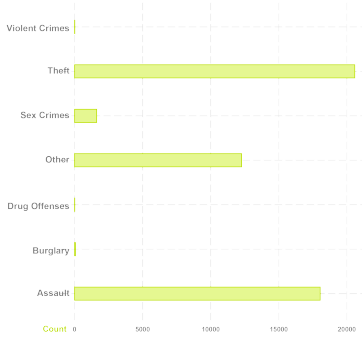

Figure 3: The number of public crimes by type

ence the walking experience in urban environments. Pavement conditions can be rated on a scale of 1 to 10, with 8 to 10 representing good quality, 4 to 7 indicating fair conditions, and 1 to 3 denoting poor conditions. In New York City, "dark spots" present concerns for pedestrian safety. The city has approximately 315,000 streetlights, and the analysis of 311 requests reveals complaints about damaged streetlights leading to inadequate lighting conditions. Additionally, the study aims to examine the impact of social involvement factors, such as women's resource networks, facilities, and subway stations. Four types of facilities are considered: core infrastructure and transportation, parks, gardens, and historical sites, public safety, emergency services and administration of justice, and library and cultural programs. Accessibility to public transportation is another vital aspect of urban infrastructure. With 472 subway stations in New York City, the subway system operates 24 hours a day, but not all lines run at all times. The temporal heatmap analysis indicates that public incidents occur more frequently during evening and nighttime hours4 and are distributed irregularly throughout the year5.

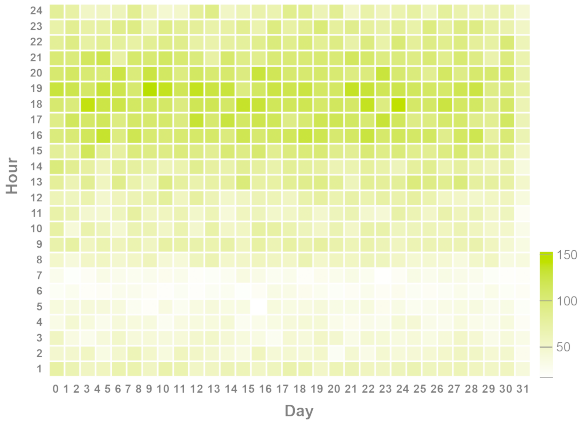

Figure 4: The public crime frequency by hours

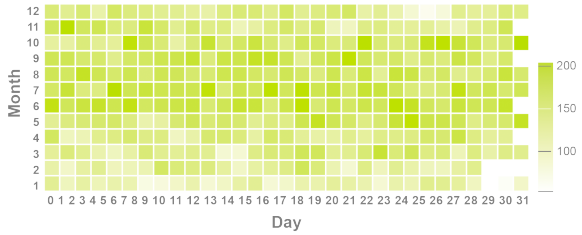

Figure 5: The public crime frequency by days

4. Methodology

In this study, several spatial analytics were used. Kernel Density Estimation (KDE) is employed to analyze the spatial distribution of vacant and insecure buildings, damaged streetlight conditions, and poorly maintained pavement in New York City.

The KDE results reveal distinct patterns6 and concentrations for each aspect of urban infrastructure:

1. Vacant and Insecure Buildings: The KDE analysis indicates a higher concentration of vacant and insecure buildings in Brooklyn, suggesting a potential area of concern for urban planners and policymakers when address-

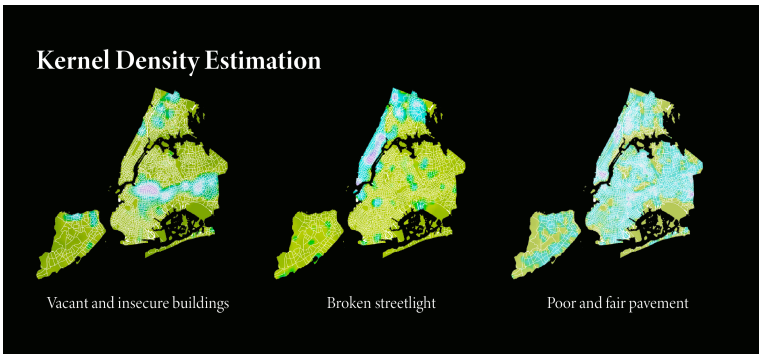

Figure 6: Kernel Density Estimation

ing women's safety and overall neighborhood quality.

2. Streetlight Conditions: The KDE results for streetlights show a notable concentration of broken streetlights in Manhattan. This finding highlights the importance of ensuring proper maintenance and functioning of streetlights in this borough to enhance safety and comfort for pedestrians, especially women, during nighttime hours.

3. Pavement Quality: The analysis of pavement conditions in New York City reveals that the overall quality is not exceptional but rather moderate. This outcome underscores the need for improvements in pavement quality across the city to facilitate safer and more enjoyable walking experiences for all residents.

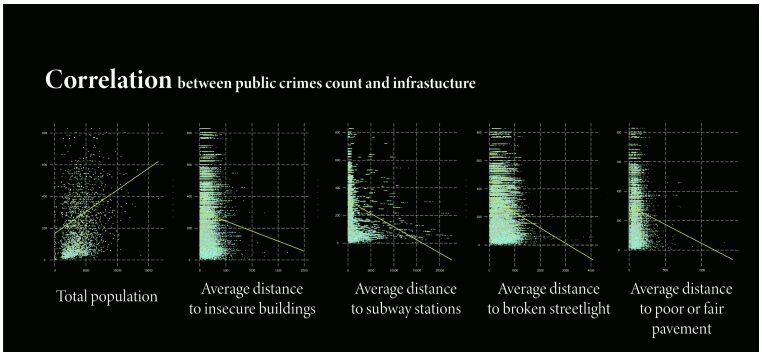

Figure 7: The correlation between the number of crimes and infrastructure features

Correlations have been identified between these urban infrastructure factors and the number of crimes. Although correlation does not imply causation, the findings suggest that the closer women are to areas with damaged streetlights, subway stations, facilities, and poor pavement, the higher the likelihood of an incident occurring7. However, different algorithms may assign varying levels of importance to these factors. Moreover, other socioeconomic factors have also been found to be significantly associated with the number of crimes.

The k-NN algorithm is a non-parametric method used for classification and regression and provides a quantitative understanding of spatial relationships in the study area. In this particular case, it is utilized to analyze distances between various geographical features and the location of public crimes. Following the function definition, the mean average distances from each geographical feature—vacant buildings, damaged streetlights, poorly maintained pavements, women's resources, subway stations, and facilities—are calculated. Then the correlation scatterplots7 give further exploration of the spatial relationships between these features and the number of crimes.

5. Models

Statistical analysis plays a crucial role in identifying key factors associated with violent incidents. However, different methods can yield slightly different perspectives on the relationships between these factors and crime rates.

The linear regression model presented here aims to predict the number of crimes against women in outdoor and public transportation spaces based on various urban infrastructure and socioeconomic factors. The model uses the following predictors: total population, mean distance to damaged streetlights, mean distance to vacant buildings, mean distance to facilities, mean distance to areas with poor pavement, mean distance to women's resources, mean distance to subway stations, median contract rent, median house value, percentage of people below poverty, and types of offenses.

The result from a linear regression model8 shows that the distance to dam-

	Dependent variable:
	num_crimes
total_pop	0.007***
	(0.0005)
mean_streetlight_distance	-0.056***
	(0.003)
mean_building_distance	-0.003
	(0.007)
mean_facility_distance	-0.442***
	(0.011)
mean_poorpavement_distance	-0.087***
	(0.012)
mean_womenres_distance	-0.021***
	(0.001)
mean_subway_distance	0.003***
	(0.001)
median_contract_rent	0.029***
	(0.002)
median_house_value	-0.00002***
	(0.00000)
people_below_poverty	0.074***
	(0.001)
offenseBurglary	-31.119
	(20.603)
offenseDrug Offenses	10.092
	(33.520)
offenseOther	-5.377***
	(2.004)
offenseSex Crimes	145.263***
	(4.445)
offenseTheft	-11.245***
	(1.752)
offenseViolent Crimes	-20.427
	(31.744)
Constant	240.559***
	(3.603)
Observations	52,660
R^2	0.246
Adjusted R^2	0.246
Residual Std. Error	170.784 (df = 52643)
F Statistic	1,075.249*** (df = 16; 52643)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 8: Linear Regression Model

aged streetlights, subway stations, facilities, poor pavement, and different offense types significantly affects the probability of incidents, indicating that increasing distance from damaged streetlights, facilities, the poor pavement, subway stations is associated with lower crime rates. Among the types of offenses, sex crimes have a strong positive association with the number of crimes.

Moreover, a Random Forest model is employed to predict the number of public crimes against women using various urban infrastructure and socioeconomic factors as predictor variables. The importance of each predictor variable in the Random Forest model is assessed by calculating the percent of IncMSE and IncNodePurity values. The percent of IncMSE value represents the increase

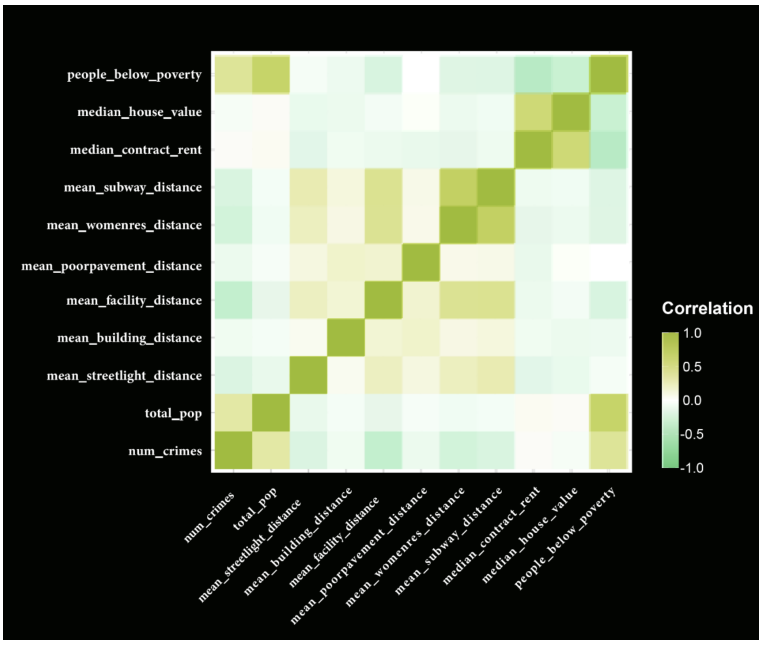

Figure 9: Correlation Matrix

in the mean squared error of predictions when a specific variable is randomly permuted, while the IncNodePurity value measures the total decrease in node impurities (i.e., residual sums of squares) resulting from splitting on a given variable, averaged across all trees in the forest.

Yet, when employing random forest10, the relative importance of these factors might differ. In addition to urban infrastructure factors, other socioeconomic factors have been found to be significantly related to the number of crimes. These importance values can be interpreted as a measure of the relative contribution of each predictor variable to the model's predictive performance. In this case, the variables with the highest importance values are the percentage of people below poverty, total population, and median contract rent, indicating that these variables are particularly influential in predicting the number of crimes against women in outdoor and public transportation spaces.

By employing multiple analytical approaches, we can gain a more comprehensive understanding of the complex interplay between urban infrastructure,

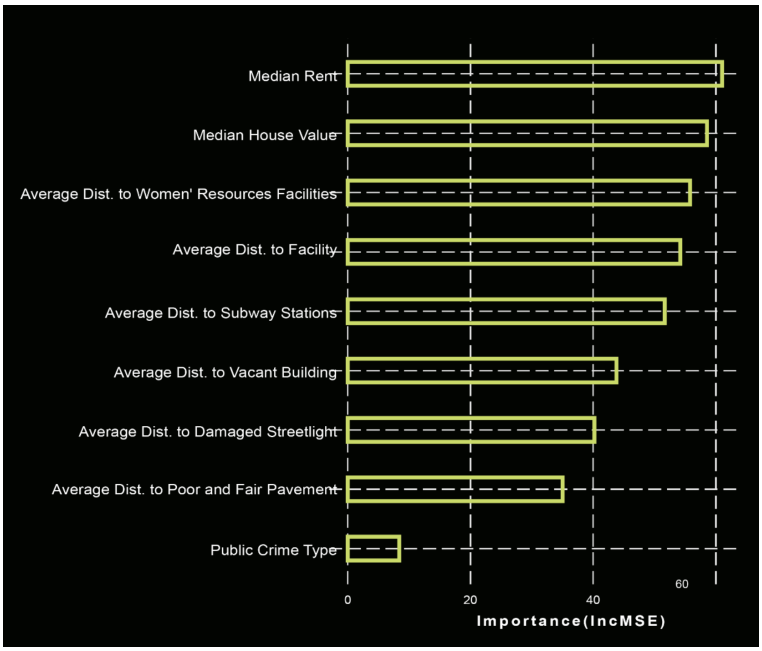

Figure 10: Feature Importance the increase in Mean Squared Error in Random Forest

socioeconomic factors, and women's safety in outdoor and public transportation spaces. This understanding can ultimately inform the development of evidence-based urban planning policies in New York City.

Indeed, criminology is a complex field that is deeply influenced by various human factors. While spatial analysis offers valuable insights by uncovering hidden patterns and relationships, its scope remains limited, and the findings must be considered with caution. This study serves as a reference for decision-makers, urban planners, and advocates of urban development who aim to create cities that cater to the needs and safety of women.

As noted by Leslie Kern, cities have predominantly been built, designed, and planned by men, leading to a male-centric perspective on urban functioning and design. This reinforces the importance of incorporating diverse perspectives, including those of women, in the process of urban planning and development. By doing so, we can create more inclusive, safe, and equitable urban spaces that accommodate the needs and preferences of all inhabitants, fostering a more just

and harmonious society.

6. Acknowledgment

I would like to express my deepest appreciation to all those who gave me supportive advice both in my study and in life.

Special gratitude goes to my academic advisor, Elizabeth for wonderful instruction, variable guidance, and insightful feedback throughout the course of this project. Her expertise and support have been instrumental in shaping my project and bringing it to fruition.

I am also grateful to my friends Ling, Ella, Tina, and Angel at UPenn for their joyful time and care. I would like to extend my thanks to my bestie Yalin who is far away in China Mainland for providing me with emotional support to enable me to keep going on my study. Furthermore, I would like to express my heartfelt gratitude to Jeremy for his love and patience throughout this academic journey, and to the cute puppy Sammy for her lovely companion. Lastly, I would like to thank my parents for their long-distance care and support to make all the thing possible.

References

[1] M. Elkin, www.ons.gov.uk (2022). URL: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/bulletins/perceptionsofpersonalsafetyandexperiencesofharassmentgreatbritain/16februaryto13march2022.

[2] L. A. B. Ouali, D. Y. Graham, M. Trompet, A. R. Barron, Journal of the Royal Statistical Society 183 (2020) 737–769. doi:10.1111/rssa.12558.

[3] E. Heroux, M. Foucault, J. D. Faubion, R. W. Hurley, Sub-stance 30 (2001) 143. doi:10.2307/3685769.

[4] D. Rodgers, B. C. O'Neill, Ethnography 13 (2012) 401–412. doi:10.1177/1466138111435738.

[5] S. L. Star, American Behavioral Scientist 43 (1999) 377–391. doi:10.1177/00027649921955326.

[6] H. Koskela, Gender Place and Culture 4 (1997) 301–320. doi:10.1080/09663699725369.

[7] M. Graystone, R. Mitra, P. L. Hess, Transportation Research Part D-transport and Environment 105 (2022) 103237. doi:10.1016/j.trd.2022.103237.

[8] A. Blöbaum, M. Hunecke, Environment and Behavior 37 (2005) 465–486. doi:10.1177/0013916504269643.

[9] A. A. Haans, Y. Y. De Kort, Journal of Environmental Psychology 32 (2012) 342–352. doi:10.1016/j.jenvp.2012.05.006.

[10] Y. Park, M. Garcia, International Journal of Sustainable Transportation 14 (2020) 860–871. doi:10.1080/15568318.2019.1641577.

[11] J. L. Nasar, B. S. Fisher, Journal of Environmental Psychology 13 (1993) 187–206. doi:10.1016/s0272-4944(05)80173-2.

[12] I. Gargiulo, X. Garcia, M. Benages-Albert, J. Martinez, K. Pfeffer, P. Vall-Casas, Landscape and Urban Planning 198 (2020) 103779. doi:10.1016/j.landurbplan.2020.103779.

[13] K. A. Henderson, M. D. Bialeschki, Journal of physical education, recreation dance 64 (1993) 44–47. doi:10.1080/07303084.1993.10606673.