Exploring the Impact of Population Density on Crime Distribution in Toronto Neighborhoods (2014–2023)*

Yongqi Liu

September 26, 2024

This study examines the relationship between population density and crime distribution in Toronto neighborhoods from 2014 to 2023. Using data from Toronto Police Services, the paper analyzes how crime patterns vary across the city and identifies key neighborhoods where crime occurance is concentrated.By focusing on the types of crime most influenced by urban density, this study offers data-driven recommendations for targeted crime prevention strategies and policy interventions. The findings highlight the need for more nuanced urban planning and law enforcement efforts to address the underlying causes of crime in high-density areas.

Table of contents

*Code and data are available at:https://github.com/Cassieliu77/Crime-Frequency-and-Population-Density.git.

1 Introduction

Urban crime has long been a focus for researchers and policymakers due to its profound impact on public safety and quality of life. In a rapidly growing metropolis like Toronto, population density plays a significant role in shaping the spatial distribution of crime. This study explores how population density correlates with crime frequency across Toronto neighborhoods between 2014 and 2023. Additionally, the research examines the influence of socioeconomic factors in high-density areas, aiming to identify neighborhoods that require targeted crime prevention strategies. Through this analysis, we hope to shed light on the dynamics of urban crime in Toronto and contribute to more effective policies for mitigating crime in high-density areas. All of the analyses in the paper are conducted in R Core Team (2021).

2 Data

2.1 Data Source

The Neighborhood Crime Rates dataset, obtained from Gelfand (2022), spans from 2014 to 2023 and covers 158 Toronto neighborhoods.It is designed to offer communities insights into public safety and awareness. The data provided to the Toronto Police Service by reporting parties is preliminary and may not have undergone full verification at the time of publication. The dataset includes various crime types such as Assault, Auto Theft, Robbery, Break and Enter, Bike Theft, Shooting, and Homicide. According to the standard definition by Statistics Canada, the crime rate is calculated as the number of crimes per 100,000 people each year. Besides, crime counts per crime type per year are also shown in the data. In Gelfand (2022), another variable of interest *POPULATION_2023* refers to the population in the 2023 year in each neighborhood, enabling the analysis of the relationship between population density and crime frequency. Geometry column is transformed into spatial data to build up the Toronto city map to observe population density and crime distribution.

As for the measurement, the crime rate is measured using population estimates from Environics Analytics, which is in line with the standard definition by Statistics Canada. The dataset records crimes that got Toronto Police intervention. This metric allows for fairer comparisons between neighborhoods with different population sizes. Also note that this dataset incorporates data for the new structure of the 158 City of Toronto neighborhoods, enabling a more comprehensive geographic analysis, and it does not contain occurrences of crime that were deemed "unfounded". Fields have been updated to reflect the new structure of the 158 City of Toronto neighborhoods.Population figures only account for the resident population in each region and do not include temporary populations like commuters or business patrons. To protect the privacy of those involved, the locations of crime incidents have been intentionally shifted to the nearest road intersection node. As such, all location data should be regarded as approximate, and users are cautioned not to interpret these locations as linked to any specific address or individual.

Crime rates offer a more balanced comparison over time as they account for population changes in the area and give us a fairer way to see the crime trends in the past decade, which explains why a similar dataset called "Police Annual Statistical Report - Reported Crimes" is not be chosen as my dataset, which only contains the reported crime count for a certain year and a certain division, which lack a standard way to show the crime trends and spatial attribute to see the geography distribution.

Figure 1 gives us a glimpse of how the dataset looks. Each neighborhood is identified by its *AREA_NAME* and a unique *HOOD_ID*.

HOOD_ID	POPULATION_2023	ASSAULT_2023	ASSAULT_RATE_2023
174	21987	101	459.3624
173	15077	105	696.4250
172	13837	104	751.6080
171	26240	229	872.7134
170	14731	543	3686.1042
169	19055	142	745.2112

Figure 1: A Lookup for the Raw Data

2.2 New Variable Constructure

The average crime rates represent the mean crime rate across 158 neighborhoods from 2014 to 2023, and were calculated to analyze long-term trends and highlight crime types that have seen significant increases. The variables *Crime_Type* and *Crime_Count* are based on 2023 data (Figure 2).

Each neighborhood's average rank was determined by ranking all 158 neighborhoods based on their total annual crime count. A rank is assigned for each year, and the average rank over the past decade (2014–2023) is calculated by averaging these yearly ranks. A lower rank indicates a higher crime count. Summaries of the top 10 and bottom 10 ranked neighborhoods are provided in Figure 3 and Figure 4.

HOOD_ID	AREA_NAME	Crime_Type	Crime_Count
174	South Eglinton-Davisville	Assault	101
174	South Eglinton-Davisville	Robbery	3
174	South Eglinton-Davisville	Break and Enter	43
174	South Eglinton-Davisville	Theft Over	8
174	South Eglinton-Davisville	Auto Theft	21
174	South Eglinton-Davisville	Bike Theft	37
173	North Toronto	Assault	105
173	North Toronto	Robbery	14
173	North Toronto	Break and Enter	31
173	North Toronto	Theft Over	7

Figure 2: Summary Table for the 2023 Crime Count

Neighborhood_ID_Name	Average_Rank
1 - West Humber-Clairville	1.60
27 - York University Heights	6.10
164 - Wellington Place	6.55
95 - Annex	7.50
73 - Moss Park	7.85
78 - Kensington-Chinatown	9.50
70 - South Riverdale	10.10
166 - St Lawrence-East Bayfront-The Islands	10.55
168 - Downtown Yonge East	10.90
136 - West Hill	11.00

Figure 3: Summary Table for Top10 Ranked Neighborhoods

Neighborhood_ID_Name	Average_Rank
15 - Kingsway South	144.50
12 - Markland Wood	144.60
49 - Bayview Woods-Steeles	146.20
97 - Yonge-St.Clair	146.25
112 - Beechborough-Greenbrook	146.65
58 - Old East York	147.30
140 - Guildwood	149.70
60 - Woodbine-Lumsden	149.80
29 - Maple Leaf	153.45
114 - Lambton Baby Point	156.20

Figure 4: Summary Table for Bottom 10 Ranked Neighborhoods

3 Results

3.1 Crime Rates Over the Past Decade

Figure 6 uses the variable *Average_Crime_Rate* (Figure 5) to show the crime rate trend between 2014 and 2023 and demonstrate the evolution of different crime types in the city of Toronto. It can be seen that the visualization reveals certain patterns in some specific crime types.

AREA_NAME	Crime_Type	Year	Average_Crime_Rate
Toronto Average	Assault	2014	610.4018
Toronto Average	Assault	2015	656.9970
Toronto Average	Assault	2016	670.1699
Toronto Average	Assault	2017	673.5303
Toronto Average	Assault	2018	680.1286
Toronto Average	Assault	2019	706.7088
Toronto Average	Assault	2020	605.1265
Toronto Average	Assault	2021	637.6509
Toronto Average	Assault	2022	703.0305
Toronto Average	Assault	2023	806.3105

Figure 5: Average Crime Rate in Toronto

Figure 6 reveals a significant upward trend in **Auto Theft** and **Assault**, particularly after 2020. The crime rate for **Assault** consistently remained the highest across all crime types, starting at around 600 per 100,000 population in 2014 and reaching a peak of approximately 800 per 100,000 population in 2023. This consistent rise suggests that assault has been a persistent issue in Toronto, with an even sharper increase following the onset of the COVID-19 pandemic.**Break and Enter**,**Robbery** and **Theft Over** all experienced increases after 2021. This rise suggests a sharp increase in private property related crimes, possibly influenced by socioeconomic factors or changes in law enforcement or technology. In contrast, **Homicide** and **Shooting** had relatively lower crime rates and displayed more stability over time, with little significant change compared to **Assault** and **Auto Theft**.

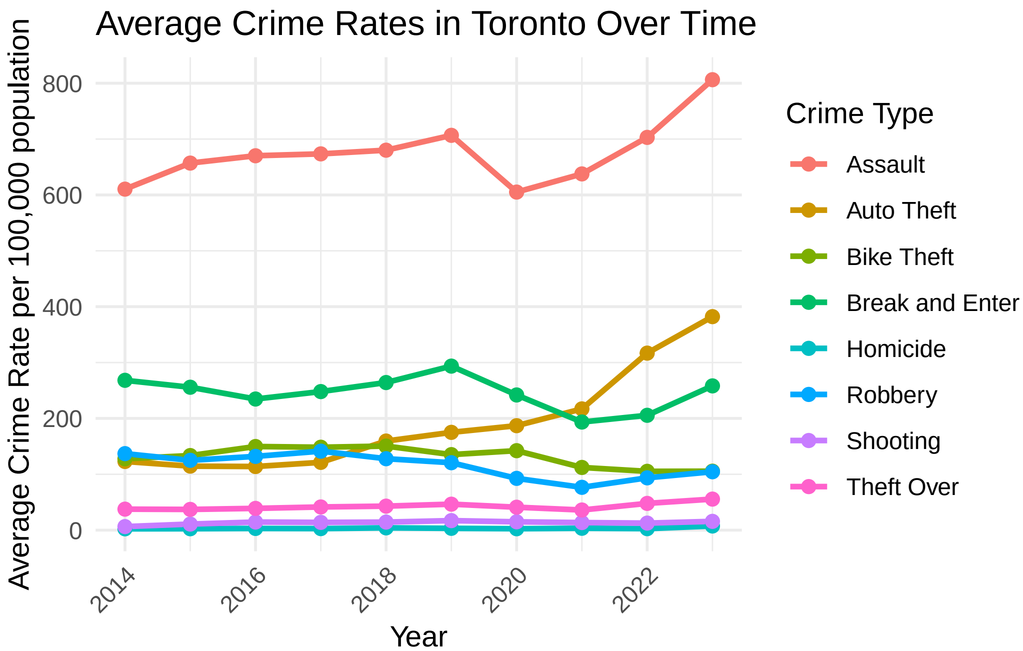

Figure 6: Average Crime Rates in Toronto Over Time

3.2 Toronto Population Density

Figure 7 shows the population density in Toronto, constructed by using variable *POPULATION_2023* and geometry data. Higher saturation represents higher population density in 2023. Neighborhoods with lighter shading have smaller populations, such as neighborhood 12 and 13. Notably, neighborhoods 1 and 95 exhibit the darkest saturation, indicating the highest population density. In the following sections, we will explore whether these high-density neighborhoods correspond to areas that are more prone to crime.

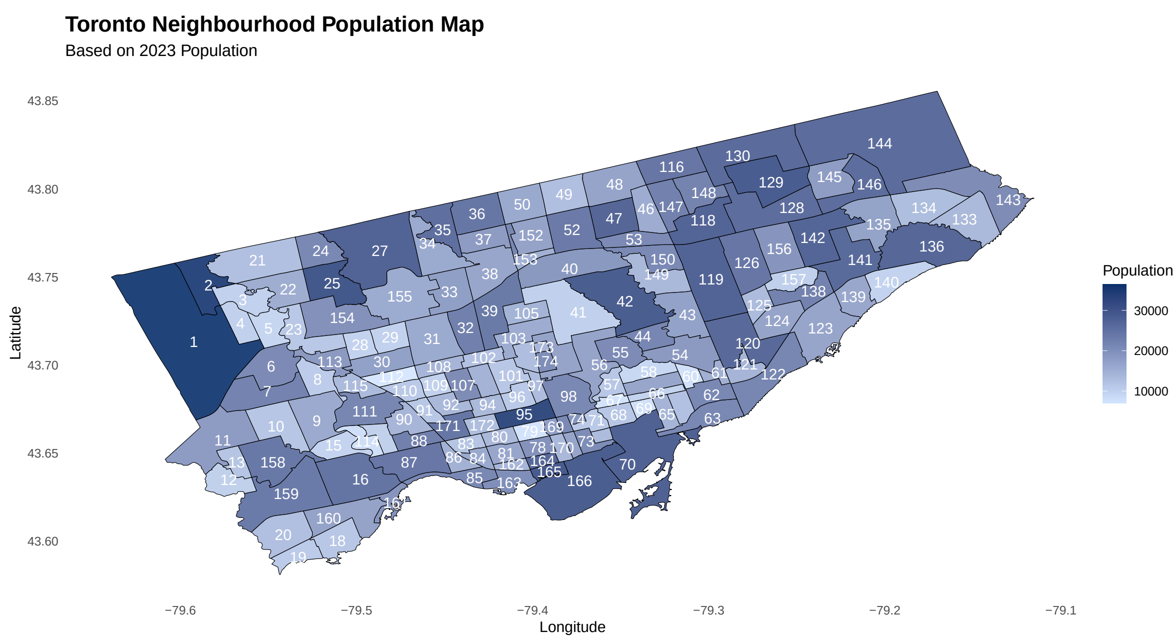

Figure 7: Population Density in Toronto

3.3 Relationship between Crime Count and Population

The scatterplots in Figure 8 illustrate the relationship between crime count and population for various crime types across Toronto neighborhoods in 2023. Each crime type is represented by a separate plot, and a linear trendline has been added to demonstrate the relationship between crime count and population size.

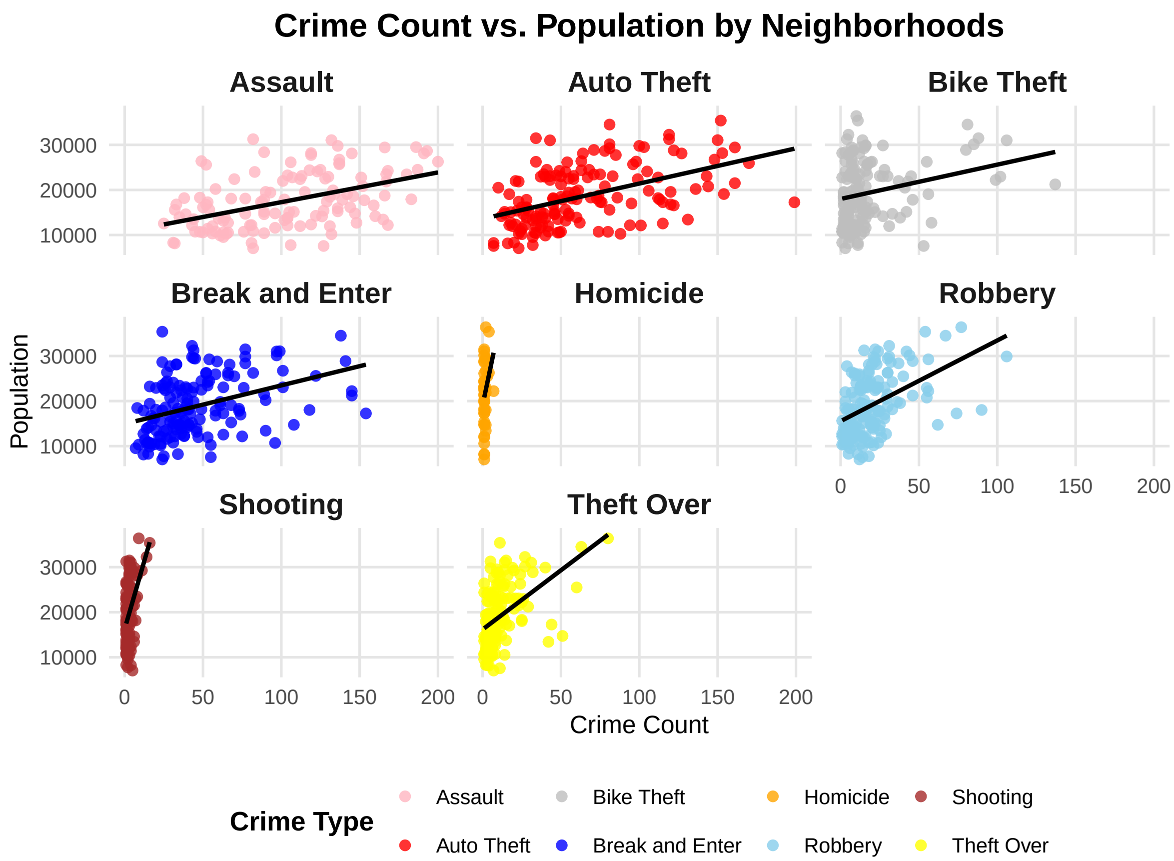

Figure 8: Crime Count with Population in 2023

Assault and **Auto Theft** exhibit strong positive correlations with population, as indicated by the upward slope of the trendlines. This suggests that as the population of a neighborhood increases, the number of assaults and auto theft incidents tends to rise as well. **Break and Enter** shows a similar trend, with higher crime counts in neighborhoods with larger populations.

In contrast, crimes such as **Homicide** and **Shooting** display much smaller crime counts overall, and their trendlines suggest weaker correlations with population size. These crimes are less frequent, and the relationship between population and crime count is less pronounced compared to property crimes.

Robbery and **Theft Over** also show a positive relationship with population, although the correlation is not as strong as it is for Assault or Auto Theft. **Bike Theft**, while showing an upward trend, appears more dispersed, indicating a weaker correlation between population size and the number of incidents.

Overall, the graph highlights that property crimes, such as **Assault**, **Auto Theft**, and **Break and Enter**, are more likely to increase with population density, whereas violent crimes like **Homicide** and **Shooting** remain relatively low regardless of neighborhood size. From Figure 8, these scatter plots represent upward trend between Crime Count and Population is shown. It seems that the lack of cases in **Shooting** and **Homicide** makes most dots overlap together. So, the map for each crime is created respectively to help us see the distribution of these 2 particular crime types.

As for the Homicide and Shooting,………….

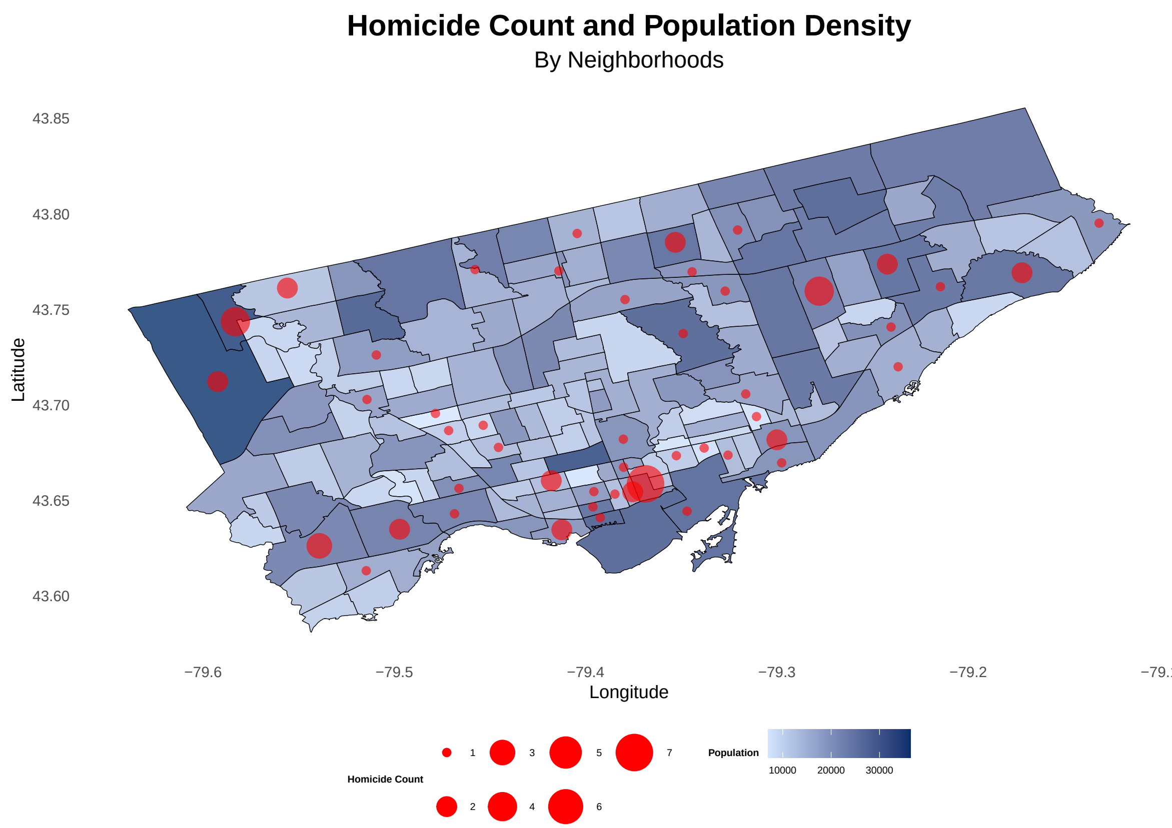

Figure 9: 2023 Homicide Distribution

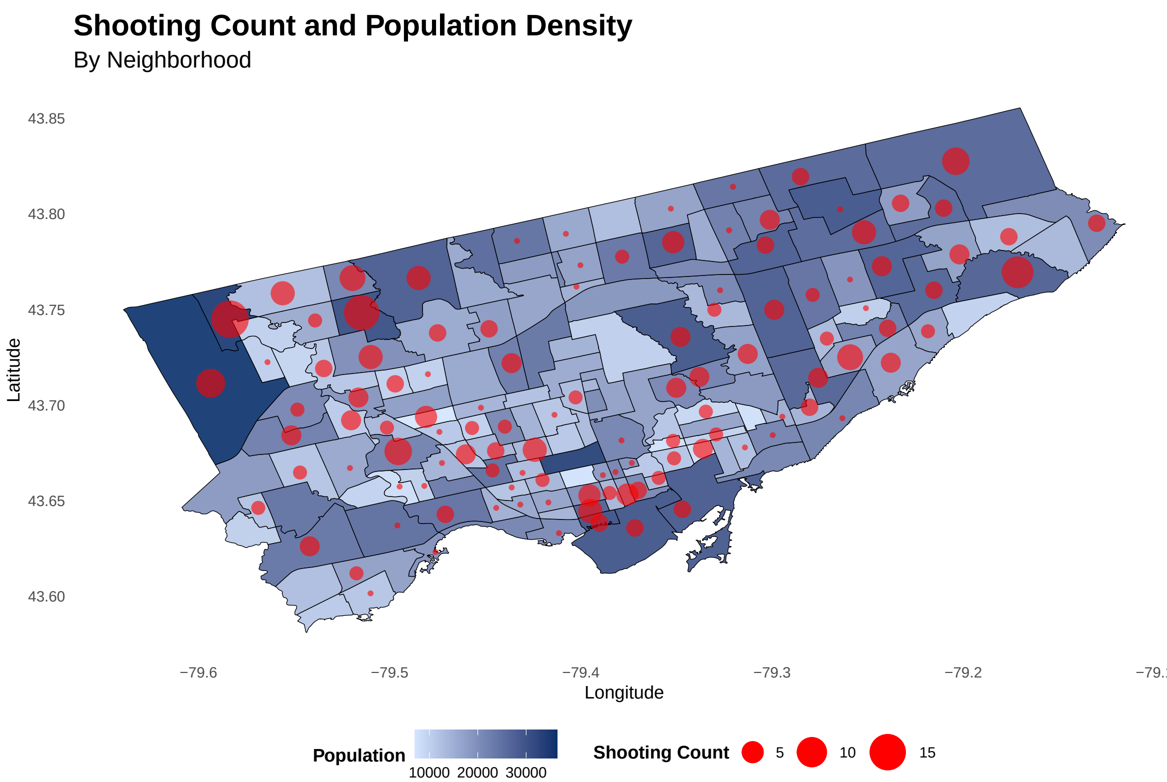

Figure 10: Shooting Count Distribution in 2023

3.4 Neighborhood Analysis

After looking at the trend of crimes in the whole Toronto, it is more important to take a look into those small neighborhoods. Top10 safest and Top10 Crime Prone neighborhoods are shown in the following content. It give police and security department to take different measures to differnt neighborhoods.Top 1 crime prone neighborhoods is West Humber-Clairville. Based on the map, it lcoates

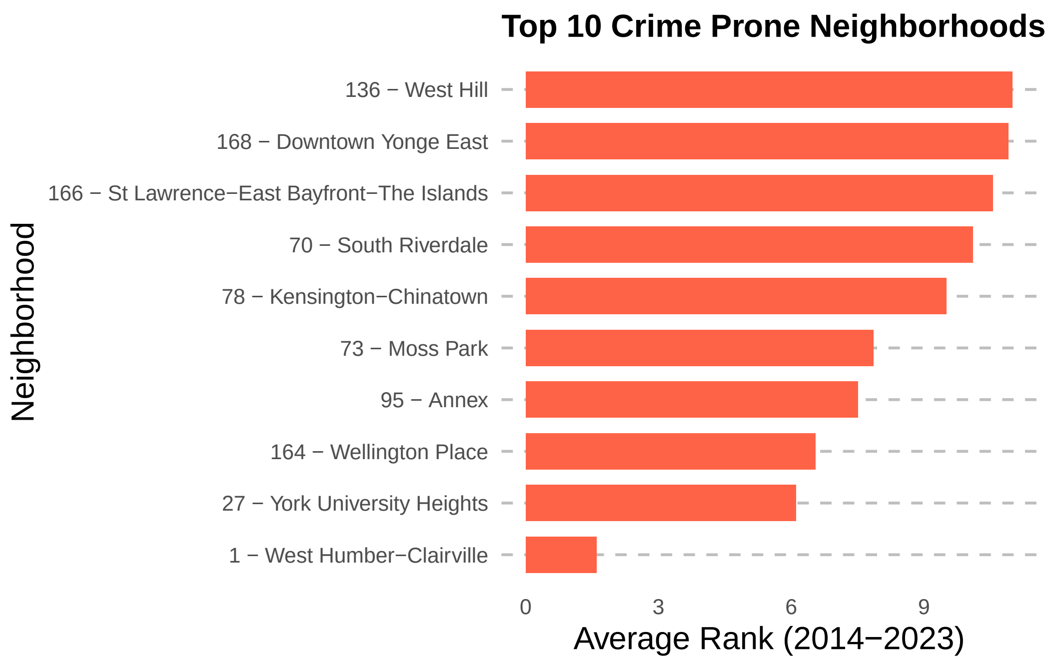

Figure 11: Top10 Ranked Crime Prone Neigborhoods Based on Crime Count

In Figure 11, neighborhoods such as West Humber-Clairville, York University Heights, and Wellington Place have consistently ranked among the top 10 crime-prone areas over the past decade. These neighborhoods are characterized by higher population densities and socioeconomic challenges, making them more vulnerable to crime. The increasing trend in auto theft and assault in these neighborhoods underscores the need for targeted interventions and enhanced security measures. In contrast,Figure 12 neighborhoods like Maple Leaf and Lambton Baby Point have consistently ranked as some of the safest areas in Toronto. These neighborhoods tend to have lower population densities and are more affluent, with better access to resources and community support. The stability of crime rates in these areas suggests that socioeconomic factors play a critical role in maintaining low crime levels.

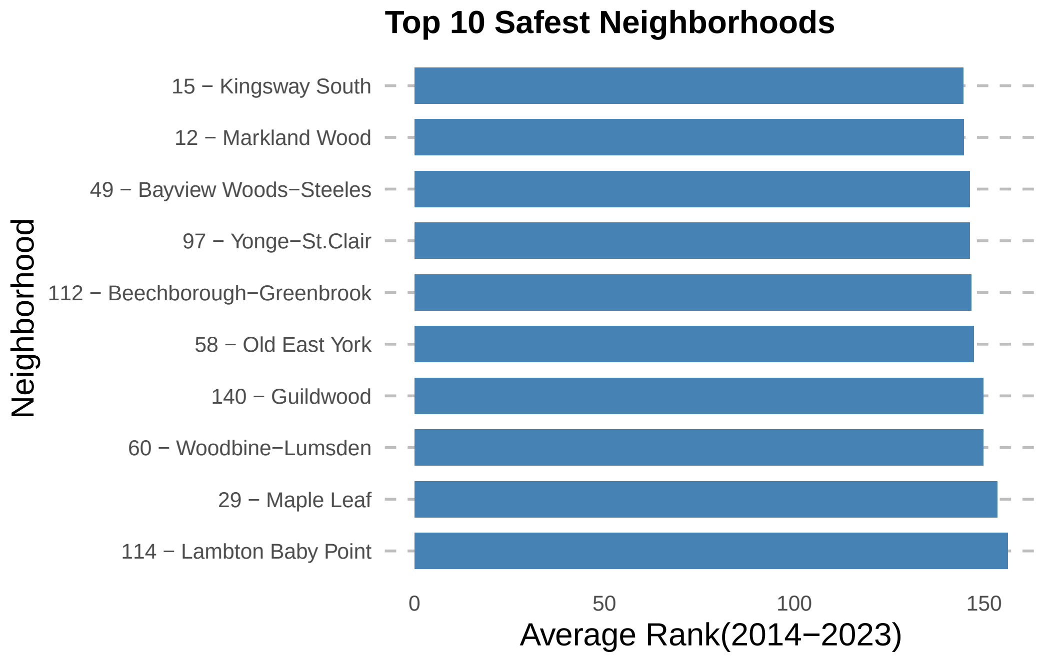

Figure 12: Top10 Ranked Safest Neigborhoods Based on Crime Count

4 Discussion

4.1 Summary of Findings

The analysis revealed significant variations in crime rates over time, particularly in property-related offenses such as Auto Theft and Break and Enter. Assault remains the most common crime, while Homicide and Robbery rates have remained stable. The data shows an upward trend in crime rates after 2020, particularly in Assault and Auto Theft, suggesting a link between population density and increased criminal activity, especially post-pandemic.

4.2 Population Density as a Key Driver of Crime

The increasing trend in crimes like auto theft and assault calls for immediate attention from policymakers. As shown in the data, crime rates do not arise uniformly across all neighborhoods but are concentrated in specific areas. This concentration suggests that socioeconomic factors, along with population density, play a significant role in the prevalence of crime. The sharp increase in auto theft post-2020 highlights the need for enhanced security measures, while the stable yet high assault rates point to the need for sustained law enforcement efforts in vulnerable neighborhoods. The data clearly shows a significant correlation between population density and crime frequency in Toronto neighborhoods. Densely populated areas such as Moss Park, St. Lawrence-East Bayfront, and Kensington-Chinatown consistently rank among the highest in terms of crime rates. The findings support criminological theories suggesting that increased population density creates more opportunities for crime due to social fragmentation and anonymity. Urban planners and policymakers must consider interventions such as improved public surveillance and community engagement initiatives to reduce crime in high-density areas. The relationship between population density and crime suggests that urban planners and local governments should consider designing interventions that specifically address high-density environments. Strategies such as better lighting, increased public surveillance (e.g., CCTV), and more frequent police patrols can act as deterrents in areas where crime tends to cluster. Additionally, promoting community cohesion through events, neighborhood associations, and resident engagement initiatives can reduce the social disorganization that often accompanies densely populated urban settings.

4.3 Socioeconomic Disparities and Crime Hotspots

Socioeconomic factors, alongside population density, play a significant role in shaping crime patterns. Neighborhoods with higher levels of economic disadvantage, such as Moss Park and York University Heights, tend to have elevated crime rates, while affluent areas like Kingsway South report significantly lower crime levels. This suggests that addressing socioeconomic disparities is essential for reducing crime. Targeted social interventions, such as improving access to education, healthcare, and employment opportunities, could help mitigate the underlying causes of crime. Figure 6 shows several crime types, notably Assault and Auto Theft, experienced noticeable increases after 2020, possibly indicating societal shifts, such as the impacts of the pandemic on public safety and economic conditions. Crime rates also vary between urban and rural areas. Densely populated urban areas typically experience higher crime rates due to greater social anonymity, economic disparities, and more opportunities for criminal activity. Conversely, remote rural areas often face challenges related to limited police resources, complicating crime prevention and law enforcement efforts. Socioeconomic inequality drives crime dynamics by creating social strain and increasing the likelihood of economic crimes such as theft, robbery, and property damage. Policymakers should prioritize reducing economic inequality through programs that address poverty, unemployment, and lack of access to education. Neighborhood revitalization initiatives that improve housing, create job opportunities, and enhance access to essential services like education and healthcare could help address the root causes of crime and reduce its prevalence in disadvantaged areas.

4.4 Targeted Crime Prevention Strategies

Lu, Lee, and Ian (2019)'s spatial analyses support previous scholarly findings that crime is not randomly dispersed across areas, but is instead concentrated in neighborhoods that exhibit specific characteristics. The geographic concentration of crime in specific neighborhoods has significant implications for law enforcement and public policy. Policymakers and law enforcement agencies must recognize that crime prevention cannot be approached with a one-size-fits-all strategy. Instead, targeted interventions are necessary to address the unique challenges faced by high-risk neighborhoods. he data reinforces the idea that crime is not uniformly distributed across Toronto neighborhoods but is concentrated in specific high-risk areas. This suggests that a "one-size-fits-all" approach to crime prevention would be ineffective. Instead, law enforcement agencies should focus their efforts on areas where crime is most prevalent, such as Moss Park and Kensington-Chinatown, by deploying more officers and allocating more resources to these neighborhoods.

At the same time, the study highlights the importance of community-based crime prevention strategies. Engaging local residents in neighborhood watch programs, providing education on crime prevention techniques, and fostering collaboration between police and community organizations can empower communities to take ownership of their safety. This approach not only reduces crime rates but also builds trust between law enforcement and the communities they serve, which is essential for long-term success in crime reduction.

For neighborhoods like Moss Park and St. Lawrence-East Bayfront, which consistently rank among the highest neighborhodos in crimes, a multifaceted approach is required. This might include increased policing and community-based programs aimed at addressing the root causes of crime, such as poverty, unemployment, and lack of access to education and healthcare. Research suggests that addressing these underlying issues can have a significant impact on reducing crime rates, particularly in economically disadvantaged areas.

Furthermore, the rise in auto theft and break-and-enter incidents post-2020 suggests that law enforcement must adapt to changing crime patterns. Technological advancements, such as the increased use of vehicle tracking systems and smart security devices, could help curb the rising rates of property crime. In addition, public awareness campaigns aimed at promoting better security practices, particularly in high-risk neighborhoods, may also prove effective in preventing crime.

4.5 Limitations and Future Research

While this study represents the relationship between population density and crime frequency in Toronto, several limitations must be acknowledged. Firstly, the analysis relies on reported crime data, which may not fully capture the extent of criminal activity, particularly for under-reported crimes such as domestic violence and cybercrime. Oliveira (2021)'s findings indicate that the assumption of linear crime growth is inaccurate. In over half of the analyzed data sets, we observed evidence of nonlinear crime growth, meaning that crime tends to increase with population size at a rate different from per capita expectations. This nonlinearity introduces a population effect that impacts crime rates and alters city rankings. We showed that ranking cities based on crime rates alone yields significantly different results compared to rankings that account for population size. So, the further

Secondly, the study focuses primarily on the relationship between population density and crime rates, without accounting for other factors that may also influence crime, such as the presence of law enforcement, urban infrastructure, and even social events. Future research could incorporate these variables to provide a more nuanced understanding of crime dynamics in urban areas.

Finally, the analysis spans a relatively short period (2014–2023), which may not capture long-term trends or shifts in crime patterns. Expanding the dataset to include earlier years or incorporating data from other cities could offer a broader perspective on crime dynamics in Toronto and beyond.

Besides, Lu, Lee, and Ian (2019) mentions the digital era has ushered in a new type of criminal activity: cybercrime and telecom crime. The Canadian Centre for Cyber Security (CCCS) reports a consistent rise in cybercrime incidents, which amounted to 74,073 police-reported cases as of recent data from Statista Lu, Lee, and Ian (2019). This kind of crime count is missing in our dataset.While obtaining a definitive national cybercrime rate is challenging due to underreporting, the CCCS highlights phishing scams, malware attacks, and ransomware as significant threats. Besides, as for differnt crime types, there may be differnt potential factors influecning the crime occurance.

Future research should expand on these findings by incorporating additional data sources to capture a broader range of criminal activity. It would also be useful to explore how different types of urban infrastructure, such as public transportation or lighting, influence crime rates. Finally, comparative studies with other major cities could provide a more comprehensive understanding of urban crime dynamics, enabling better-targeted interventions for crime prevention.

5 Acknowledgements

All data is analyzed in R (R Core Team (2021)), using packages Tidyverse (Wickham et al. (2019)), Dplyr(Wickham et al. (2023)), Geojsonsf(Cooley (2022)), Httr (Wickham (2023)), Knitr(Xie (2014)) and ggplot2 (Wickham (2016)), sf((**sf?**)), gt(Iannone et al. (2024)), lintr(Hester et al. (2024)) in order to visualize the data and communicate information intuitively.

We would like to express our appreciation to the City of Toronto and the Open Data Portal for granting access to the dataset via the Open Data Toronto package (Gelfand 2022). Besides, R codes from (**citerohan?**) are used to provide tables in this paper.The provision of high-quality, open-source data is essential for enabling precise research. Thanks to OpenAI (2023), which is used in this research paper.Code written in the scripts was checked and styled with lintr (**citelintr?**) and styler (**citestyler?**).

References

Cooley, David. 2022. *Geojsonsf: GeoJSON to Simple Feature Converter*. https://CRAN.R-project.org/package=geojsonsf.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Hester, Jim, Florent Angly, Russ Hyde, Michael Chirico, Kun Ren, Alexander Rosenstock, and Indrajeet Patil. 2024. *Lintr: A 'Linter' for r Code*. https://CRAN.R-project.org/package=lintr.

Iannone, Richard, Joe Cheng, Barret Schloerke, Ellis Hughes, Alexandra Lauer, JooYoung Seo, Ken Brevoort, and Olivier Roy. 2024. *Gt: Easily Create Presentation-Ready Display Tables*. https://CRAN.R-project.org/package=gt.

Lu, Wang, Gbby Lee, and Williams Ian. 2019. "The Spatial and Social Patterning of Property and Violent Crime in Toronto Neighbourhoods: A Spatial-Quantitative Approach." *ISPRS International Journal of Geo-Information* 8 (1). https://doi.org/10.3390/ijgi8010051.

Oliveira, M. 2021. "More Crime in Cities? On the Scaling Laws of Crime and the Inadequacy of Per Capita Rankings—a Cross-Country Study." *Crime Science* 10 (27). https://doi.org/10.1186/s40163-021-00155-8.

OpenAI. 2023. "ChatGPT: GPT-4 Model." https://openai.com/chatgpt.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2023. *Httr: Tools for Working with URLs and HTTP*. https://CRAN.R-project.org/package=httr.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in r." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.