A Study on Assault Crime Rates in Toronto Neighborhoods (2014-2023)*

Top 10 Assault Crime Hotspots and Population Impact: An Analysis of Toronto (2014-2023)

Jack Zhou

September 27, 2024

This paper analyzes trends in assault crimes across various neighborhoods in Toronto from 2014 to 2023. By examining the total assaults over the past decade, this study highlights key patterns in the growth or decline of assault-related incidents. Furthermore, the relationship between population size and assault rates is explored to determine if higher populations correlate with increased assault numbers. Finally, the analysis identifies the top 10 neighborhoods with the highest total assaults, pinpointing crime hotspots within the city. These findings provide insights that could aid urban planners and policymakers in improving public safety and resource allocation.

Table of contents

*A GitHub Repository containing all data, R code, and other files used in this investigation is located here: https://github.com/Shuhengzhou03/Initial-submission-paper-1.git

1 Introduction

In 2023, as Toronto's urban population continues to grow, assault crimes have become a focal point in public safety discussions. Assault crimes not only pose a direct threat to the daily lives of residents but also reveal deeper issues related to the unequal distribution of urban resources (Luca et al. 2023). Assault incidents in various neighborhoods across Toronto exhibit distinct trends and characteristics, reflecting differences in economic conditions, population density, and urban planning across the city (Wang, Lee, and Williams 2019). These trends suggest the importance of analyzing the broader social and demographic factors that influence crime in different communities.

This paper seeks to address this gap by analyzing assault crime data across Toronto neighborhoods from 2014 to 2023. The focus of the analysis is twofold: first, it examines the overall trends in assault crime rates over the past decade, identifying key patterns in the increase or decrease of assault incidents. Second, it explores the relationship between population size and assault rates, determining whether higher populations correlate with an increase in assault crimes. Finally, the analysis highlights the top 10 neighborhoods with the highest number of total assaults, revealing hotspots where public safety concerns are particularly acute. These insights are critical for urban planners and policymakers, who need a clear understanding of how population and resource disparities impact crime.

The findings of this study reveal that while assault crimes have generally increased in Toronto between 2014 and 2023, the relationship between population size and assault rates is not uniform. Some densely populated neighborhoods report higher crime rates, but several smaller neighborhoods also exhibit disproportionately high levels of assaults. This suggests that other factors, such as socioeconomic status or access to public services, may have a more significant impact on crime rates than population size alone. Understanding these dynamics is vital for developing targeted crime prevention and resource allocation strategies that can improve public safety in Toronto (Onyeneke and Karam 2022).

This paper is structured as follows: Section 2 outlines the data and methods used, including a discussion of measurement and statistical analysis. Section 2.2 presents the key results, including trends in assault crimes and the population-crime relationship. Section 3 discusses the implications of these findings for urban policy and planning, with recommendations for future research.

2 Data

The raw data used in this study was sourced from the `opendatatoronto` (Gelfand 2022),which provides access to detailed datasets on crime rates and population statistics. Specifically, the dataset titled "Neighbourhood Crime Rates (2014-2023)" was downloaded in CSV format. The data was then processed and analyzed using the R programming language (R Core Team 2022).

During the data processing, tools from the tidyverse suite (Wickham et al. 2019)were used to clean and structure the data. After importing the data through CSV files, ggplot2 (Wickham 2016) was employed to create bar plots and scatter plots to illustrate assault crime patterns across Toronto's neighborhoods. The entire analysis, including data cleaning and visualization, was conducted in the R programming environment to ensure reproducibility and efficiency. In addition, I also used styler (Müller and Walthert 2024) and lintr (Hester et al. 2024) to format my R code to follow a consistent style, improving readability and making collaboration easier

2.1 Measurement

This study utilizes the Neighborhood Crime Rates dataset provided by the City of Toronto's Open Data Portal. The dataset includes statistical data on various types of crimes such as Assault, Auto Theft, Break and Enter, Robbery, Theft Over, Homicide, and Shooting & Firearm Discharges. Additionally, the dataset provides crime rates per 100,000 people, calculated based on population estimates from Environics Analytics. Each entry represents a summary of crime victims by various demographic and crime categories, as reported by the Toronto Police Department. The data originates from actual incidents and victim reports, which are then structured into records that quantify crime events over time.

One challenge in using this dataset is the presence of missing data (NA values) for certain neighborhoods and years, which may reflect unreported incidents or gaps in data collection. To address this issue, imputation strategies such as filling missing data with zeros were employed. However, while this approach maintains the continuity of the analysis, it may obscure actual changes in crime patterns, potentially introducing bias and uncertainty. Additionally, this study focuses on victim data, providing crucial insights into the demographics and experiences of crime victims, which is essential for understanding the impact of crime within Toronto's communities. This approach differs from other studies that often use datasets focusing more on offenders and general crime reports.

To ensure the robustness of the analysis, systematic data cleaning processes were applied, with a focus on aggregating total assault cases across neighborhoods.

2.2 Results

After loading the dataset using the R programming language (R Core Team 2022), the `tidyverse` (Wickham et al. 2019) package was used to generate graphs.

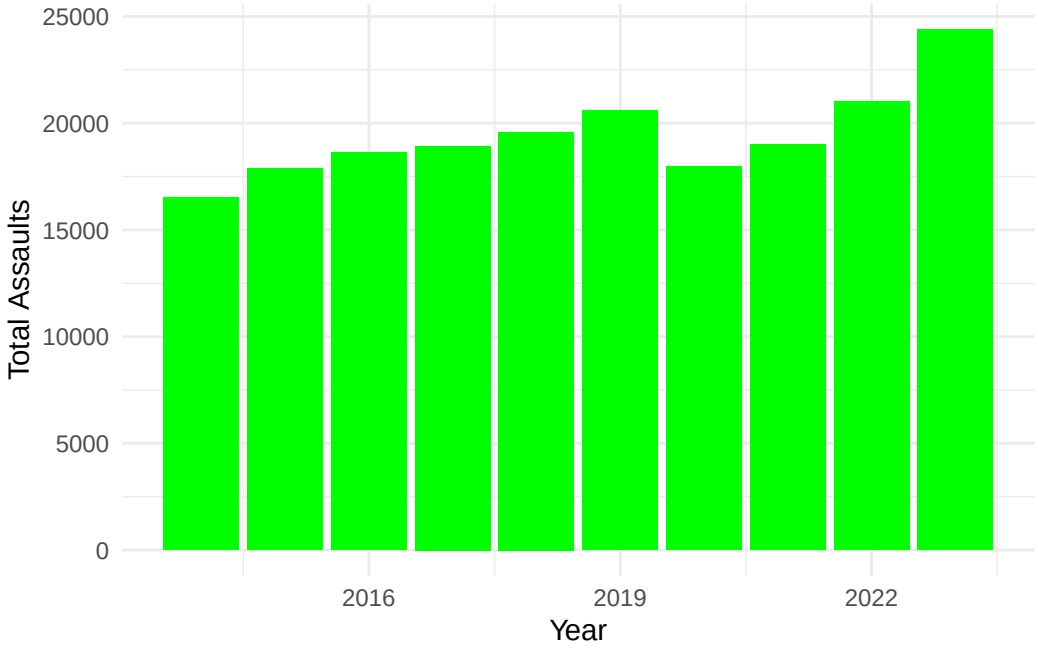

Figure 1: Total Number of Assault Cases in Toronto Neighborhoods (2014-2023)

Figure 1 illustrates the overall trend of assault cases across various neighborhoods in Toronto from 2014 to 2023. As seen in the graph, the total number of assaults fluctuated during this period. In 2014, the number of reported assault cases was around 15,000, and it steadily increased over the next few years, reaching close to 20,000 by 2018. Afterward, there was a slight decline, but from 2021 onward, the number of assaults surged again, peaking at approximately 25,000 cases in 2023—the highest point in the time frame.

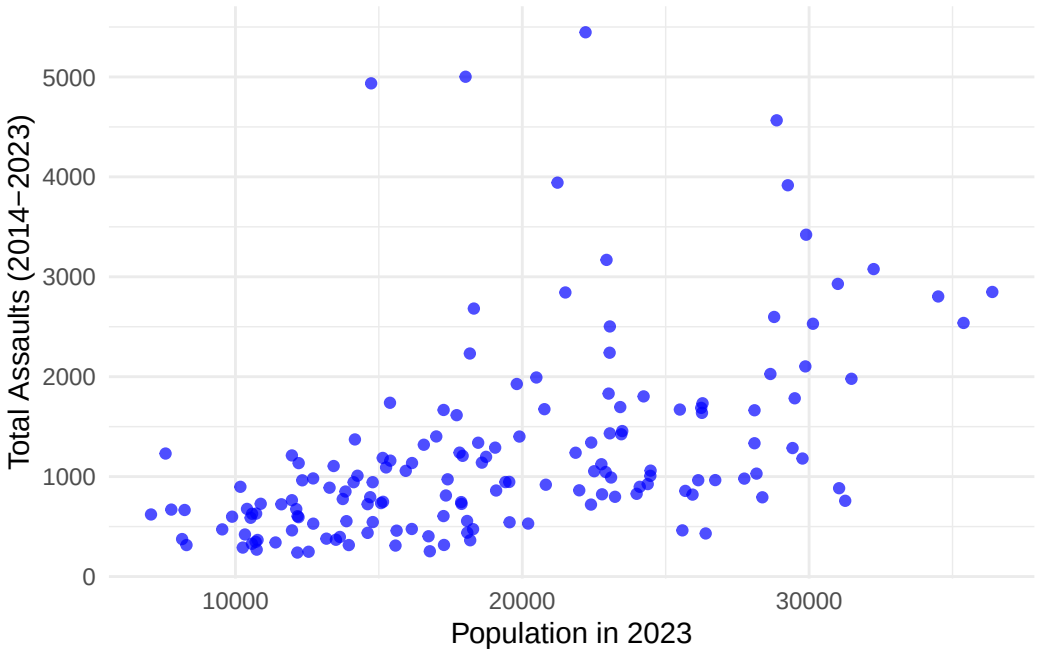

Figure 2: Relationship Between Population Size in 2023 and Total Assaults (2014-2023) in Toronto Neighborhoods

Figure 2 illustrates the relationship between the population of various neighborhoods in 2023 and the total number of assaults from 2014 to 2023. As shown in the plot, the population of most neighborhoods in 2023 is concentrated between 10,000 and 25,000, and the total number of assaults in these neighborhoods generally ranges from 0 to 300. However, there are some notable outliers, indicating that in certain smaller neighborhoods (e.g., those with populations below 15,000), the total number of assaults can exceed 500.

As the population increases, the total number of assaults shows an overall upward trend, particularly in neighborhoods with populations over 25,000 in 2023, where the total number of assaults often exceeds 300. Although the relationship between population and the total number of assaults is not strictly linear, the overall trend suggests that neighborhoods with larger populations tend to report higher numbers of assaults. This indicates that, aside from population size, other factors such as socioeconomic conditions or the availability of public services may influence the crime rate in these neighborhoods.

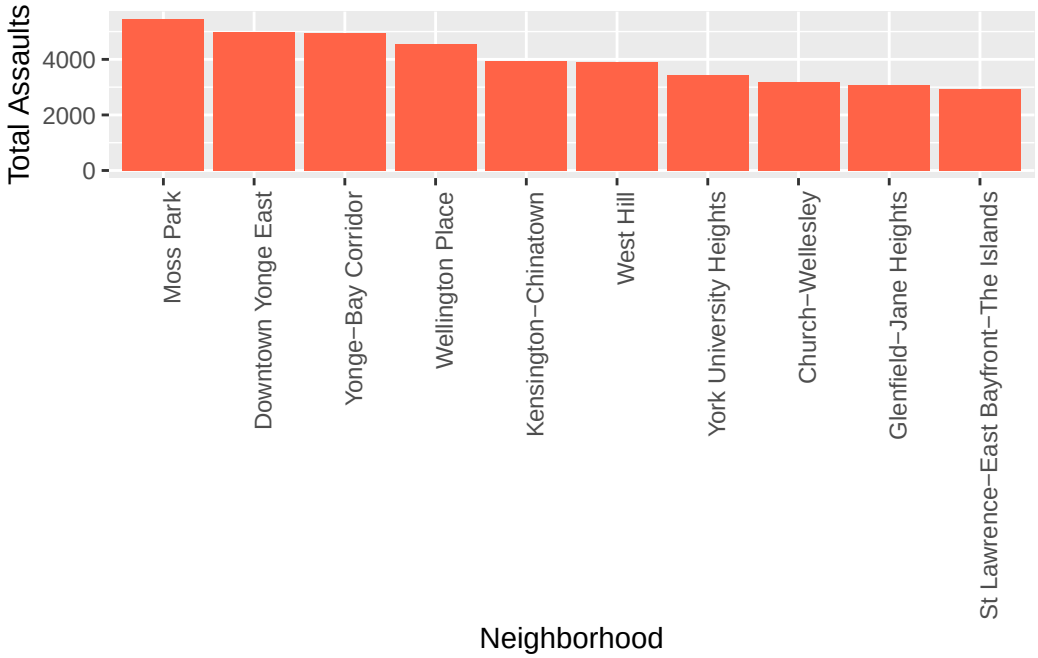

Figure 3: Top 10 Neighborhoods by Total Assaults in Toronto (2014-2023)

Figure 3 displays the top ten neighborhoods in Toronto with the highest total number of assaults from 2014 to 2023. The neighborhood names are shown on the x-axis, while the y-axis represents the total number of assaults, ranging from 3,000 to over 5,000. "Moss Park" has the highest number of assaults, with over 5,000 cases, followed by "Downtown Yonge East" with approximately 4,500 cases, and "Yonge-Bay Corridor" with around 4,000 cases. Other neighborhoods, such as "St. James Town" with approximately 3,800 cases, "Bay Street Corridor" with about 3,600 cases, and "Waterfront Communities-The Island" with approximately 3,400 cases, also rank high.

3 Discussion

In this analysis, we found that the number of assault cases in certain neighborhoods in Toronto is significantly higher than in others. The concentrated distribution of assault cases reflects the population size, socioeconomic status, and uneven distribution of resources in specific communities. This phenomenon suggests that areas with higher population density and poorer economic conditions are more likely to become crime hotspots, especially in places where resources are scarce and public services are limited, making this trend even more pronounced.

Notably, this study revealed a positive correlation between population size and the total number of assault crimes, particularly in neighborhoods with populations exceeding 25,000, where the number of assault cases increased significantly. This may be related to factors such as

the mobility of residents, social class differences, and higher stress levels in these communities. Additionally, the concentration of assault cases in specific neighborhoods may be linked to the lack of law enforcement resources and social service distribution (Bones and Hope 2015). In high-crime areas, the appropriate allocation of resources and more effective crime prevention measures could have a positive impact on reducing crime rates.

However, this study has certain limitations. First, the study only considered reported assault cases and did not include unreported crimes, which may lead to an underestimation of the total number of assault cases. Secondly, while this study analyzed the relationship between assault cases and population size, it did not fully examine other factors influencing crime rates, such as the economic conditions of the community, education levels, and social support networks (Gokmenoglu, Yıldız, and Kaakeh 2020). Future research should further explore the impact of these variables on assault crime rates to gain a more comprehensive understanding of crime patterns in Toronto neighborhoods.

Furthermore, this study emphasizes that policymakers should prioritize resource allocation in high-crime neighborhoods. Strengthening law enforcement efforts, improving community infrastructure, and providing more social services in these areas could effectively reduce the occurrence of assault cases (Lanni 2022). Future policies should focus more on densely populated and economically disadvantaged neighborhoods, developing targeted crime prevention strategies to enhance public safety levels.

In conclusion, this study provides important insights into the understanding of assault crimes in Toronto neighborhoods and offers a basis for urban planners and policymakers to develop more effective resource allocation strategies and crime prevention measures.

A Appendix

A.1 Dataset and Graph Sketches

Sketches depicting both the desired dataset and the graphs generated in this analysis are available in the GitHub Repository.

A.2 Data Cleaning

The data cleaning process involved filtering out some of the columns from the raw dataset and renaming some of the data entries for clarity and simplicity.

A.3 Attribution Statement

"Contains information licensed under the Open Government Licence – Toronto" (City of Toronto, n.d.).

References

Bones, Paul DC, and Trina L Hope. 2015. "Broken Neighborhoods: A Hierarchical Spatial Analysis of Assault and Disability Concentration in Washington, DC." *Journal of Quantitative Criminology* 31: 311–29.

City of Toronto. n.d. "Open Data License." https://open.toronto.ca/open-data-license/.

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

Gokmenoglu, Korhan K, Bünyamin Fuat Yıldız, and Mohamad Kaakeh. 2020. "Examining the Impact of Socioeconomic Factors on Crime Rates: A Panel Study." In *Annual Conference on Finance and Accounting*, 409–20. Springer.

Hester, Jim, Florent Angly, Russ Hyde, Michael Chirico, Kun Ren, Alexander Rosenstock, and Indrajeet Patil. 2024. *Lintr: A 'Linter' for r Code*. https://github.com/r-lib/lintr.

Lanni, Adriaan. 2022. "Community-Based and Restorative-Justice Interventions to Reduce over-Policing." *American Journal of Law and Equality* 2: 69–84.

Luca, Massimiliano, Gian Maria Campedelli, Simone Centellegher, Michele Tizzoni, and Bruno Lepri. 2023. "Crime, Inequality and Public Health: A Survey of Emerging Trends in Urban Data Science." *Frontiers in Big Data* 6: 1124526.

Müller, Kirill, and Lorenz Walthert. 2024. *Styler: Non-Invasive Pretty Printing of r Code*. https://CRAN.R-project.org/package=styler.

Onyeneke, Christopher Chimaobi, and Aly H Karam. 2022. "An Exploratory Study of Crime: Examining Lived Experiences of Crime Through Socioeconomic, Demographic, and Physical Characteristics." *Urban Science* 6 (3): 43.

R Core Team. 2022. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wang, Lu, Gabby Lee, and Ian Williams. 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): 51.

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

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.