From the Suburbs to Downtown: An In-Depth Analysis of Major Crime Occurrences Throughout the City of Toronto's 158 Neighbourhoods (2019-2023)*

Shirley Chen

January 24, 2024

Criminal activity levels throughout the many neighbourhoods in the city of Toronto have fluctuated over the past few years. Using data from OpenData-Toronto, this paper will assess the various types of criminal offences that have occurred within Toronto's 158 neighbourhoods, comparing areas with both higher and lower crime rates. A deeper understanding of the patterns and implications for community safety and urban development are illustrated by conducting a thorough analysis on criminal activity data in the city of Toronto's 158 neighbourhoods.

Table of contents

1 Introduction **2**

2 Data **2**

2.1 Sources 2

2.2 Cleaning Procedure 3

2.3 Data Analysis 3

3 Conclusion **6**

References **6**

*Code and data in this report are available at: https://github.com/shirleychen003/toronto-crime-analysis.git

1 Introduction

Known as the largest city in Canada, the city of Toronto is made up of neighbourhoods that altogether encompass the diverse atmosphere of the city. As of March 2022, the city of Toronto identifies 158 distinct neighbourhoods, rather than the 140 that were established in the late 1990s (Toronto, n.d.). Among the historical, cultural, and socioeconomic diversity, includes the ranges of criminal activity that vary depending on the area. As the population of Toronto continues to expand, some criminal activity levels within the city are also bound to progressively increase. According to a survey of 1500 Torontonian adults conducted by the Toronto Police Association (TPA), along with issues such as housing accessibility and affordability, one of the top concerns that Toronto residents have relate to crime and public safety within the city (Sheppard and Coletto 2024). Residents have voiced how issues such as TTC safety, hate crimes, gang violence, and auto theft have worsened, conveying their uneasiness about how dangerous the city has become. Furthermore, everyday news articles continuously release the latest crime updates in the city, causing residents to question the security and safety in their area.

Understanding the complexities within the city is essential for numerous reasons. Those who may be unfamiliar with the city could be unaware of the more "safe" and more "dangerous" areas. Being informed about the criminal activity in different parts of Toronto assists with protecting one's well being and personal safety. On the other hand, knowing which areas are filled with higher crime rates and suspicious activity levels can also engage communities and promote developing initiatives to support the overall safety of residents. This can encourage the government to allocate resources to communities to address certain needs.

In the following paper, I investigate patterns and trends of criminal activity occurring during 2019-2023 within the City of Toronto's 158 neighbourhoods. The data section consists of the graphs produced using `ggplot2` (Wickham, Chang, et al. 2023) and tables produced using `knitr` (Xie 2023) and `kableExtra` (Zhu 2021). Any patterns or notable changes in data are discussed to present an in-depth analysis. Additionally, a description of the data cleaning processes applied to the dataset and the packages used are mentioned.

2 Data

2.1 Sources

The dataset used in this paper is titled "Major Crime Indicators (MCI)" and was obtained from the City of Toronto's open data portal, which was retrieved and saved in `R`(R Core Team 2023) using `opendatatoronto` (Gelfand 2022) and `readr`(Wickham, Hester, and Bryan 2024). The package `here`(Müller 2020) was used to manage file paths and both `gitcreds` (Csárdi 2022) and `usethis` (Wickham, Bryan, et al. 2023) were used for GitHub functions.

2.2 Cleaning Procedure

The data cleaning process included simplifying the dataset so that it could be understood and transformed into graphs. Packages used for this procedure included the `tidyverse`(Wickham 2023) package , `tidyr`(Wickham, Vaughan, and Girlich 2023),`tibble`(Müller and Wickham 2023), `janitor`(Firke 2023), and `dplyr`(Wickham, François, et al. 2023).

Originally, the dataset contained information dating back to 2009. However, I chose to focus on the data occurring in the past 5 years, thus I removed all the rows which were not involved in the specific time range. Additionally, I wanted to focus on the date that the offences had occurred, not reported. Therefore, I removed any data that included the reporting date. Otherwise, I changed several column names for better understanding and removed any other unnecessary columns that were not relevant to my analysis.

2.3 Data Analysis

Below shows a table of the total number of criminal offences that occurred per year from 2019 to 2023.

Table 1: Number of Incidents by Year

Years	Number of Incidents
2019	40098
2020	35196
2021	34777
2022	41299
2023	47833

As we can see, there is a significant decline from 2019 to 2020. This drastic change may be attributed to the COVID-19 pandemic, as less of the Toronto population were spending their time in public or leaving their residence due to worries of contracting the virus. Even until 2021, we see that there are less incidents of criminal activity compared to the other years, which is around the time that stay-at-home orders were coming into effect and case numbers were rapidly rising.

Once life became more "normal" after the pandemic, citizens were more open to being in public once again. However, there was also more criminal activity happening. In 2022, the incidents began to increase again and in 2023, there is a significant increase of 6,534 incidents. The incident numbers in 2019 are quite similar to those in 2022, but we see a considerable increase when comparing 2019 and 2023. This may be due to post-pandemic issues, such as effects on mental health and fluctuations in the housing market.

Figure 1 illustrates that the most occurred offence type is assault and the least occurred offence type is theft over. As shown below in the graph, the numbers are drastically different.

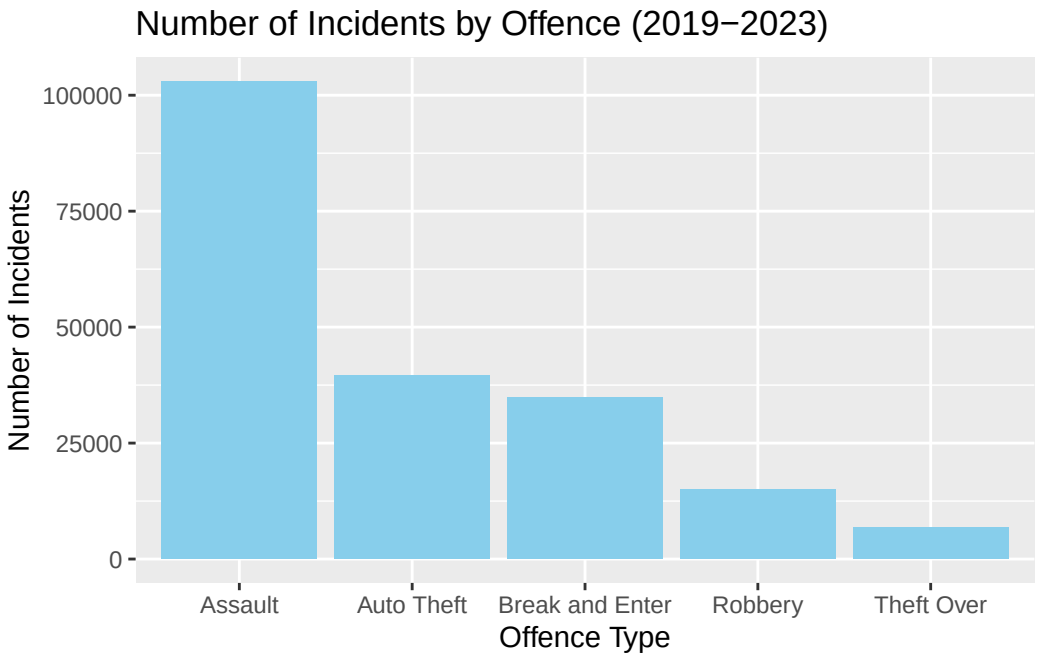

Figure 1: Number of Incidents by Offence Type

Finally, location is an unavoidable and crucial component to consider when analyzing criminal activity in the city. Depending on the individual, this knowledge can be used for personal safety reasons or for social impact motivations. Below, Figure 2 illustrates the areas with the most incidents of crime.

We see that the 5 neighbourhoods with the most criminal incidents are (decreasing in order) West Humber-Clairville, Moss Park, Downtown Yonge East, York University Heights, and Yonge-Bay Corridor. With West Humber-Clairville as an exception, the other 4 of these neighbourhoods are located closer to the downtown area, where the population is more dense. Since there are more people, there is more chance for crime offence numbers to be higher.

In contrast, Figure 3 displays the neighbourhoods with the least incidents of crime. According to this graph, the neighbourhoods that may be considered the "safest" are Woodbine-Lumsden, Guildwood, Lambton Baby Point, Humber Heights-Westmount, and Old East York. Most of these neighbourhoods are in out of the city in the suburbs, in less densely populated areas. Individuals who are interested in buying property are likely to take these statistics in to consideration.

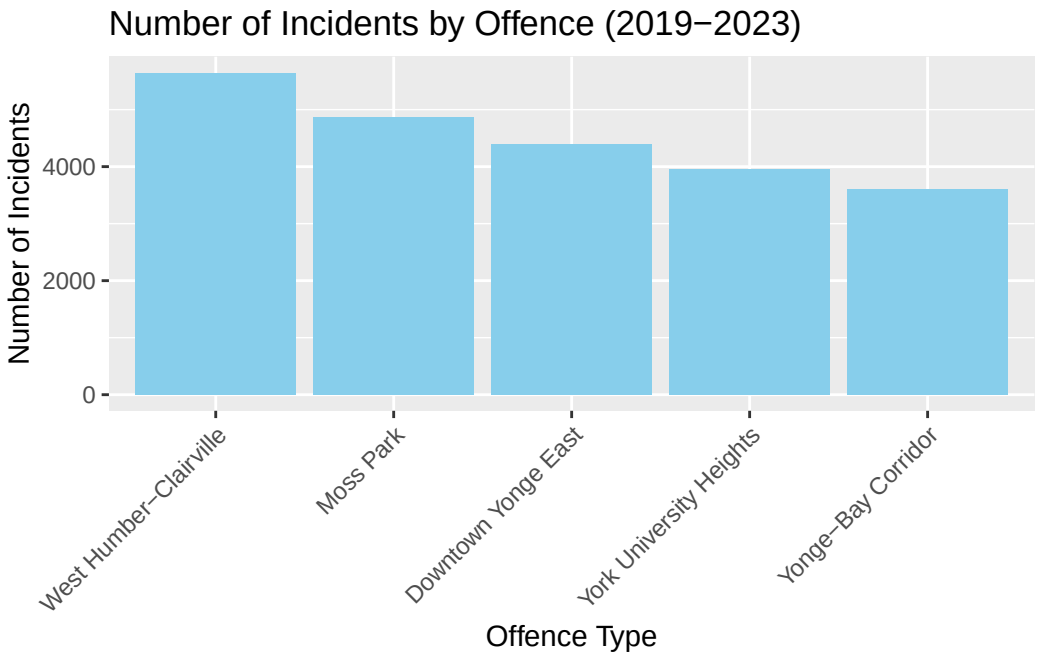

Figure 2: Areas with the most incidents of crime

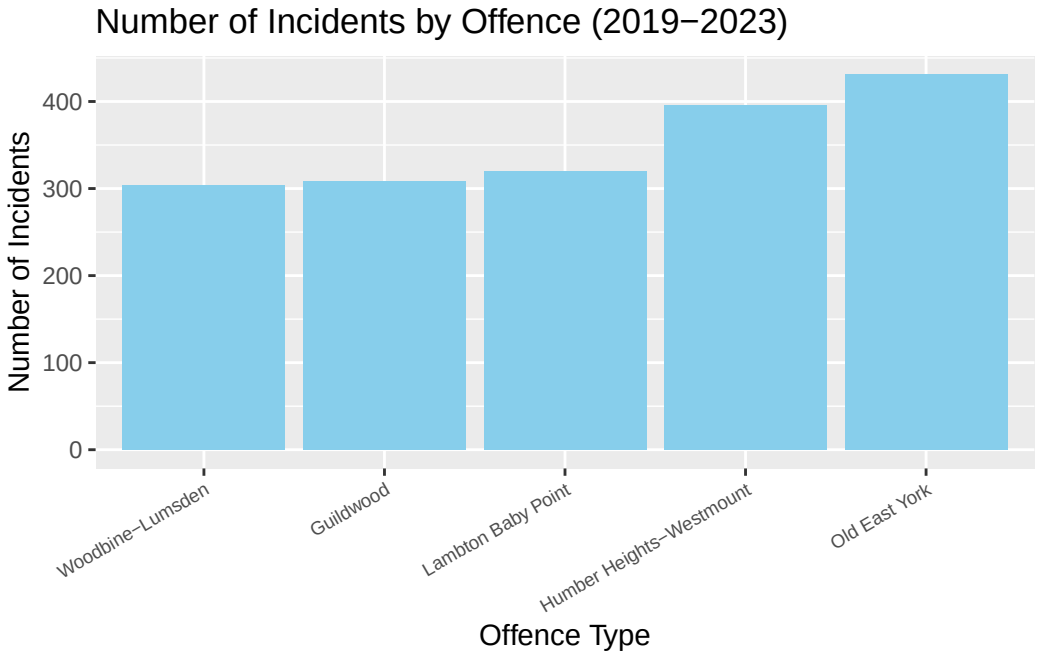

Figure 3: Areas with the least incidents of crime

3 Conclusion

This paper investigates the growing trends of criminal activity within Toronto's 158 neighbourhoods. Over the course of 5 years between 2019 and 2023, data collected showed that 2023 had the highest rates of criminal activity, whereas the lowest rates were in 2020 and 2021. The 2020 and 2021 numbers can be attributed to the pandemic. Further, assaults were recorded to be the most occurred offence type. Finally, it was demonstrated that West Humber-Clairville, Moss Park, Downtown Yonge East, York University Heights, and Yonge-Bay Corridor contained the most amount of criminal activity, while Woodbine-Lumsden, Guildwood, Lambton Baby Point, Humber Heights-Westmount, and Old East York contain the least amount of criminal activity. Overall, with crime rates fluctuating in the city of Toronto, the analysis of Toronto major crimes assists with informing individuals of the diverse regions and activity levels within the city.

References

Csárdi, Gábor. 2022. *Gitcreds: Query Git Credentials from r*. https://gitcreds.r-lib.org/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://github.com/sfirke/janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://here.r-lib.org/.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://tibble.tidyverse.org/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Sheppard, Eddie, and David Coletto. 2024. "Toronto's Safety Crossroads: Unveiling Concerns, Policing Priorities, and the Push for Adequate Funding." Abacus Data. 2024. https://abacusdata.ca/tpa-torontos-safety-crossroads/.

Toronto, City of. n.d. "About Toronto Neighbourhoods." https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/neighbourhood-profiles/about-toronto-neighbourhoods/.

Wickham, Hadley. 2023. *Tidyverse: Easily Install and Load the Tidyverse*. https://tidyverse.tidyverse.org.

Wickham, Hadley, Jennifer Bryan, Malcolm Barrett, and Andy Teucher. 2023. *Usethis: Automate Package and Project Setup*. https://usethis.r-lib.org.

Wickham, Hadley, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2023. *Ggplot2: Create*

Elegant Data Visualisations Using the Grammar of Graphics. https://ggplot2.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2024. *Readr: Read Rectangular Text Data*. https://readr.tidyverse.org.

Wickham, Hadley, Davis Vaughan, and Maximilian Girlich. 2023. *Tidyr: Tidy Messy Data*. https://tidyr.tidyverse.org.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with Kable and Pipe Syntax*. http://haozhu233.github.io/kableExtra/.