

Investigating the Safety of Different Toronto Neighbourhoods in 2011 Based on The Number of Crimes Reported*

Determining Which Neighbourhoods have the Highest Number of Crimes and How Different Crimes are Correlated.

Bruce Zhang

September 23, 2024

The paper looks at how the crime rates in different Toronto neighbourhoods are different. The data will look into the differences in total major crime incidents per neighbourhood and compare these numbers to the number of specific crimes per neighbourhood. The paper will then assess any potential correlations between specific crimes and total crimes. This will lead to indications of safety and chance of specific crimes in neighbourhoods in the Greater Toronto Area (GTA).

*Code and data are available at: [LINK](#).

1 Introduction

This paper looks at how Toronto neighbourhoods differ in the number of crimes committed. As those living in Toronto would know, certain areas of the city feel less safe than others. However, it is barely ever known whether these feelings of safety are proven by data or if it's just a non-evidence-based thought that people get. This paper aims to answer that question by looking at how the number of crimes differ between neighbourhoods.

People also often think of different crimes differently when considering safety. For example, people may be more concerned about assault than theft and associate it more with safety due to the more dangerous nature of the crime. The paper will also look at how these specific types of crimes relate to the total number of major crimes and how that may indicate certain characteristics about the safety of the neighbourhood.

The introduction section goes over the background of the paper and summarizes the overall procedures performed. It also highlights the potential real-world implications of the analysis of the data. The remainder of this paper is structured as follows. The data section (Section 2) will highlight the characteristics of the dataset used for this paper. It will look at the types of parameters that the initial data looks at and how these can be looked at. The section will dissect how the data can be interpreted via graphs and tables. The discussion section (Section 3) will go over the real-world implications in more detail based on the findings of the data. Potential ways of interpreting the total major crime numbers and the correlation between types of crimes will be further broken down and put into a real-world setting. The weaknesses and potential next steps of the data analysis will also be discussed.

2 Data

Based on the raw data, different neighbourhoods have a different amount of crimes by category. The table (Figure 1) shows an example of this data for 10 neighbourhoods.

The bar graph (Figure 2) shows the top 50 neighbourhoods ranked by the total major crime incidents. This helps narrow down the high-crime communities where correlational relationships between total major crime incidents and specific crimes can be better analyzed. As the graph suggests, the neighbourhood named “Waterfront Communities - The Island” has a substantially higher number of total major crime incidents compared to the others. Other neighbourhoods such as the “West Humber-Clairville” neighbourhood, “Bay Street Corridor”, “Church-Yonge Corridor”, and a few others also have a noticeably higher numbers of total major crime incidents.

Scatter plots (Figure 3) were constructed to examine the potential correlations between major crimes and specific crimes.

Neighbourhood	Assaults	Break & Enters	Drug Ar- rests	Hazardous Incidents	Sexual As- saults	Thefts	Total Major Crime Incidents	Vehicle Thefts
Waterfront	892	139	148	272	41	17	1393	61
Communities- The Island								
West Humber- Clairville	390	175	62	210	68	54	1119	288
Bay Street Corridor	554	129	108	179	47	29	1017	26
Church-Yonge Corridor	535	143	99	257	60	13	1006	31
Downsview- Roding-CFB	402	187	126	233	50	8	981	128
Moss Park	371	116	302	215	36	5	921	27
Kensington- Chinatown	369	166	197	146	25	14	869	29
Islington-City Centre West	268	207	71	249	37	23	853	188
Woburn	412	128	77	223	29	7	808	45
York University Heights	282	150	56	186	49	16	776	122

Figure 1: Number of Different Types of Crimes By Neighbourhood

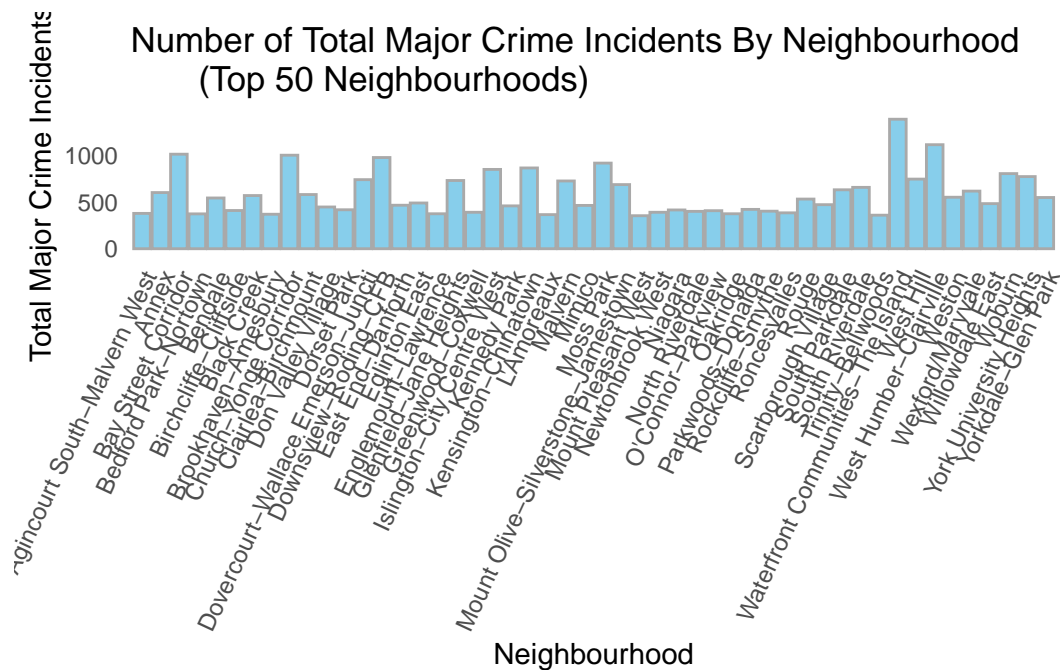


Figure 2: Top 50 Neighbourhoods in Total Major Crime Incidents

Based on the scatterplots in Figure 3, it can be seen that the specific types of crimes examined, including assaults, break and enters, drug arrests, hazardous incidents, sexual assaults, thefts, and vehicle thefts, all show positive correlation to the total number of major crime incidents. However, it is also clear that most specific crimes are very weakly correlated the the total number of crimes. The only crime with a relatively strong correlation to the total number of major crime incidents, based on visual interpretations, is assault.

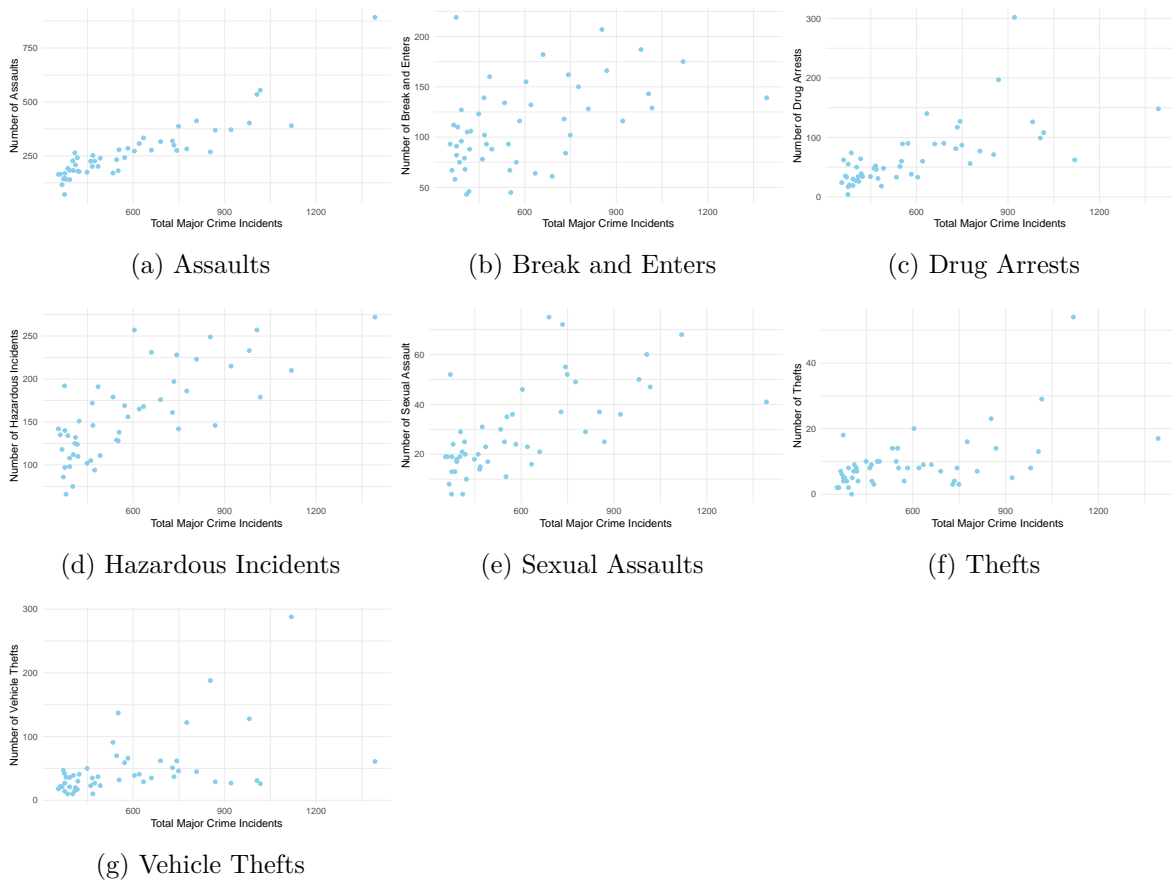


Figure 3: Relationship between Total Major Crime Incidents and Specific Crimes for The 50 Neighbourhoods with the Highest Number of Total Major Crime Incidents in Toronto in 2011

3 Discussion

3.1 Implication of High Number of Total Major Crime Incidents

Based on what was seen in the data, the Toronto neighbourhoods with the highest total number of major crimes can be easily determined. According to what was shown in the bar graph (Figure 2), neighbourhoods located near the South part of the Greater Toronto Area (GTA), especially Downtown regions of the city, were highly represented in the top 50 for total crimes.

One interpretation is that this indicates that the safety of the downtown region is considerably lower than other regions of the city. The city departments that play a role in citizen regulation and safety control may be less competent than those departments in regions with lower numbers of total major crime incidents.

However, another interpretation could be related to the population and abundance of people. The regions with the highest total major crime incidents also tend to be regions with a higher population abundance. The Downtown Toronto area is significantly more populated than Northern GTA. Since the data presented was not standardized for population size, the numbers reflected can largely be a result of that difference. Having more people in a given neighbourhood would likely impact the number of crimes committed.

3.2 Relationship Between Specific Crimes and Total Crimes of Neighbourhoods

Based on what was seen in the results displayed in Figure 3, all specific crimes show a positive but relatively weak correlation to the total number of major crime incidents. This indicates that, although a high total number of crimes may suggest a higher likelihood of each specific crime, the relationship is not strong. A neighbourhood with the highest number of total number of major crimes may not consequently have the most number of vehicle theft, for example. This can be further examined by calculating crime rates and looking at the correlations again.

The crime that had notably stronger correlation to the total major crime incidents was assault, as seen in sub figure (a) of Figure 3. This suggests that there may be more of a relation between the total number of crimes in the area and assault crimes in the area. In other words, the general safety or likelihood of crime in a neighbourhood in Toronto can indicate the likelihood of assault happening in that area more so than it can indicate the likelihood of other crimes. The cause of this stronger relationship requires further research and investigation to determine if the stronger relationship is significant and a reproducible result.

3.3 Weaknesses and next steps

One weakness, which was previously mentioned, is how the data is not standardized for different population sizes. This results in the data across different communities being incomparable. A next step to solve this issue would be to obtain population data and calculate the crime rate instead of the number of crime incidents. This would allow the data from each neighbourhood to be standardized. They can then be compared to determine the levels of safety of different communities relative to each other.

Another weakness is the lack of generalizability of this data and correlational analysis. Since the data is solely from Toronto neighbourhoods and is obtained in 2011, the correlations between crimes may not hold relevant in any other city in the world. It may also no longer be relevant at the current time because of how much the city has changed and developed over the past 13 years. To address this, a more thorough analysis on different cities across the world would need to be done. It would also be important to use data that is more recent in order to have implications on the future.

4 References

R Core Team (2023) Wickham et al. (2019)

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

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