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 27, 2024

This paper looks at the differences in total major crime incidents per neighbourhood, specifically isolating and analyzing the data from neighbourhoods with high numbers of total crime. The paper also compares these numbers to the number of specific crimes per neighbourhood and assesses any potential correlations between specific crimes and total crimes. The findings reveal that the 'Waterfront Community' neighbourhood had the highest number of total crimes among other neighbourhoods in the Southern portion of Toronto. The findings also reveals that the number of assault had a more obvious positive correlation to the total number of crimes compared to other crime types. This leads to indications of safety and chance of specific crimes in neighbourhoods in the Greater Toronto Area (GTA) and can be used as a guideline to select neighbourhoods to live in. Inferences can also be made regarding the relationship and tendency of specific crimes to occur in neighbourhoods with high overall crime numbers.

*Code and data are available at: https://github.com/brucejczhang/Toronto_Crimes_By_Neighborhood.

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. Overall, in cities across the world, it is also a trend that the distribution of crimes and levels of safety vary between different regions (Wang, Lee, and Williams (2019)). However, it is hard to determine 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 based on the data (SocialDevelopment, Finance, and Administration (2014)). Having this information with statistical support can guide citizens of the city in residence selection.

Another potential lack of knowledge is the pattern of specific types of crimes relative to the safety of the neighbourhood. 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 than security due to the more violent nature of the crime. In other words, there are patterns in crimes. Occurence of crimes often have specific patterns. Some of the patterns may be geographical characteristics or socially related (DeLisi and Vaughn (2015)). This information may be more useful than an overall number of crimes when considering the general social background of specific neighbourhoods (Fitzgerald, Wisener, and Savoie (2004)). The paper will 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. We will also attempt to look into the potential patterns based on the data present for crime abundances in different geographic locations (neighbourhoods) in the GTA.

In the results section, a table was constructed to offer an overall representation of the data, a bar graph of the top 50 neighbourhoods in total major crime incidents was made, and a series of scatter plots were created to display the relationship between the total number of crimes and specific types of crime. The bar graph demonstrated that the neighbourhood named "Waterfront Communities - The Island" had the highest number of total crimes. Other neighbourhoods that were high in total number of crimes include many that are located in the Downtown Toronto region. This information can be important because it acts as an indicator of the overall safety level of neighbourhoods across the GTA and can assist people in selecting areas to reside.

The scatter plot showed that specific crime types, including assault, theft, break and enter, sexual assault, vehicle theft, drug arrest, and hazardous incidents, all had a positive correlation with the total number of crimes. All correlations were relatively weak, but assault had a comparably stronger positive correlation to total major crimes. This demonstrates a key relationship between assaults and total crimes. The information can indicate social characteristics of high-crime neighbourhoods and better aid residence selection of the population.

This introduction went over the background of the paper and summarizes the overall procedures performed. It also highlighted 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. The discussion of what the high crime numbers (Section 3.1), the relationships between total crimes and specific crimes (Section 3.2), and further exploration of the relationship between assault and total crimes (Section 3.3) will be broken down. Potential ways of interpreting these pieces of information will be put into a real-world setting. The weaknesses and potential next steps of the data analysis will also be discussed (Section 3.4).

This paper was created using statistical programming language python (Python Core Team (2019)). Packages numpy (Harris et al. (2020)), matplotlib (Hunter (2007)) and pandas (McKinney (2010)) were used as part of the code.

2 Data

2.1 Overview

The data set used for this paper (SocialDevelopment, Finance, and Administration (2014)) was obtained from Open Data Toronto. The measurement method of the data was by obtaining police and government related information. The publisher of the data acknowledged potential measurement errors due to method of measurement and ways that the neighbourhoods were separated. The direction of effect would be that the number of crimes may have been under-counted.

For the data section, this paper will focus on identifying the 50 most dangerous neighbourhoods in the GTA based on the number of total crime incidents. Then, the paper will look at how different crime types correlate with the number of total crimes within these neighbourhoods. This will help identify trends of crime relationships in the most crime-abundant areas of the GTA.

2.2 Data Table

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 data shown has been clean to be ordered in a way where the neighbourhoods with the highest total number of crime incidents would be listed first.

The data from the table (Figure 1) show that the Waterfront Communities have the most total number of crime incidents in 2011. The remaining neighbourhoods in the top 10 include those on the left-most column labelled “Neighbourhood” in (Figure 1). The types of specific

Neighbourhood	Assaults	Break & Enters	Drug Arrests	Hazardous Incidents	Sexual Assaults	Thefts	Total Major Crimes	Vehicle Thefts
erfront Communities-The	892	139	148	272	41	17	1393	61
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 We	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

crimes that are viewed are in the column headings in row 1 of (Figure 1). These are the most common crimes that have occurred in the GTA in 2011, as represented by the data set.

2.3 Graphical Representations

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.

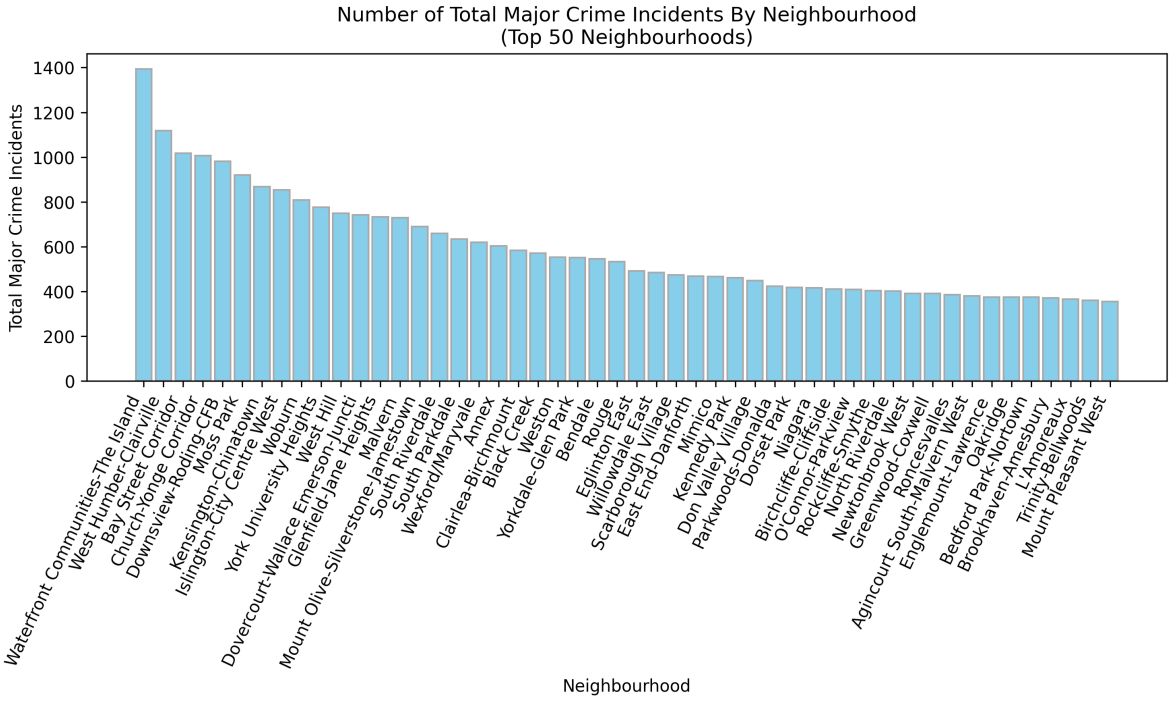
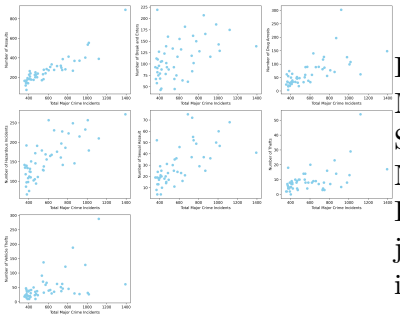


Figure 2: Top 50 Neighbourhoods in Total Major Crime Incidents

The graph visually demonstrates what was quantitatively seen in Figure 1. 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", "Bay Street Corridor", "Church-Yonge Corridor", and a few others also have a noticeably higher numbers of total major crime incidents. The other neighbourhoods from the top 50 in total crime incidents that are represented all have an estimate of 350 total major crimes in 2011. This is significantly less than than "Waterfront Communities - The Island", which had 1393 total crime incidents in 2011, as indicated by Figure 1.

Scatter plots (Figure 3) were constructed to examine the potential correlations between major crimes and specific crimes. The purpose of these was to identify the patterns of crimes relative to the number of total crimes, or general safety level, of the neighbourhood.



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

Figure 3: Assaults

Based on the scatter plots 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. This can be seen by the overall increasing trend of the number of each specific crime as the number of total crimes increase. 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 (a)). The stronger correlation is determined by the more clustered and data points and the clearer trend of increase in number of assaults as the total number of crime incidents increase.

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 that were highest in total number of crimes tended to be located towards the South part of the Greater Toronto Area (GTA), especially Downtown regions of the city. The higher crime numbers can be interpreted as an overall well-being of the community and the safety of the neighbourhood, as indicated in the original data set (SocialDevelopment, Finance, and Administration (2014)).

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. On a similar note, the social status and financial ability of people in different neighbourhoods vary. Those neighbourhoods with higher numbers of crime could be the ones where preferred victims of criminals live or where criminals themselves tend to reside.

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 positve 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 Further Interpretation of The Relationship Between Assault and Total Crimes

The potential reasons behind assault numbers being correlated to the total number of major crimes can vary largely. This can interpret One way of interpreting this relationship is that, people who commit assault crimes tend to do so in neighbourhoods that have more crime

overall. This tendency is stronger than people committing other types of crimes. Many other potential explanations may be valid. Further research would be required for a specific conclusion on the reason behind this correlation.

3.4 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.

A final weakness is the timeliness of this data. The data set is from 2011, which is out of date to be used for analysis of more recent trends. The neighbourhood crime numbers could have shifted largely since 2011. There may have also been expansions of the city, where many neighbourhoods present today may not have been present in 2011. This means that the data that is looked at in this paper may not be the best source to use to make residence suggestions to people living in the city today. To improve this, a more recent set of data must be used. Similar ways of analyzing the data can be done to achieve the effects of suggesting residency for the current population.

References

DeLisi, Matt, and Michael G. Vaughn. 2015. "Correlates of Crime." *The Handbook of Criminological Theory* Ch.2: 18. https://doi.org/10.1002/9781118512449.ch2.

Fitzgerald, Robin, Michael Wisener, and Josée Savoie. 2004. *Neighbourhood Characteristics and the Distribution of Crime in Winnipeg*. Canadian Centre for Justice Statistics. https://www.publicsafety.gc.ca/lbrr/archives/cnmcs-plcng/statcan-cjrps-no4-eng.pdf.

Harris, Charles R., K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, et al. 2020. "Array Programming with NumPy." *Nature* 585 (7825): 357–62. https://doi.org/10.1038/s41586-020-2649-2.

Hunter, John D. 2007. "Matplotlib: A 2D Graphics Environment." *Computing in Science & Engineering* 9 (3): 90–95.

McKinney, Wes. 2010. "Data Structures for Statistical Computing in Python." Edited by Stéfan van der Walt and Jarrod Millman, 51–56.

Python Core Team. 2019. *Python: A dynamic, open source programming language*. Python Software Foundation. https://www.python.org/.

SocialDevelopment, Finance, and Administration. 2014. "About Wellbeing Toronto - Safety." Open Data Toronto. https://open.toronto.ca/dataset/wellbeing-toronto-safety/.

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." *International Journal of Geo-Information* 8 (1): 51. https://doi.org/10.3390/ijgi8010051.