Homocides in Toronto Neighbourhoods from 2004-2020*

Justin Teng

06 February 2022

Abstract

Understanding the datasets of homicides in The City of Toronto Neighbourhoods are crucial for determining the distribution of police resources among the neighborhoods. From the City of Toronto Open Portal, we obtain and analyze this data of homicide cases in Toronto neighbourhoods between the years of 2004-2020. We find that homicide occurrences have been slowly increasing each year from 2011 to 2020. Also, certain neighbourhoods in Toronto have relatively higher numbers of homicide occurrences. Moreover, there is a significant large percentage of homicides that are firearm-related. Our findings have implication for allocation of police surveillance, police patrols, and the management of weapons registration.

1 Introduction

As the largest city in Canada, Toronto has a population of about 6.5 million people. However, from 2002-2020, it is also the Metropolitan Area with the most homicide cases in Canada each year according to Statistics Canada(Canada 2021). Despite that overall police-reported crime in Canada has decreased in 2020, there are still 743 homicide cases being reported by the police. Toronto had the highest number of homicides of all Canada's Metropolitan Areas in 2020 with 105 victims, which was a decrease of 25 victims from 2019 (Armstrong and Jaffray 2021). Therefore, the analysis on this dataset of homicides in the neighbourhoods of Toronto from the year of 2004 to 2020 will be crucial in terms of understanding the occurrences of homicide over the years.

This dataset includes information on all reported homicides from 2004 to 2020 in the 140 Neighbourhoods of Toronto. However, the numbers by Division and Neighbourhood may not reflect the exact time or count of occurrences reported within these geographies, for the location of crime occurrences have been deliberately offset to the nearest road intersection for the privacy of parties involved in the occurrence (Gelfand 2020). From the data, we find that homicide occurrences have been slowly increasing each year from 2011 to 2020. Also, certain neighbourhoods in Toronto have relatively higher number of homicide occurrences. Furthermore, 52% of the homicide cases are firearm-related.

Our findings have important implications for the allocation of police surveillance and patrols among different neighbourhoods in Toronto. For instance, neighbourhoods with the highest number of homicide occurrences, such as Mosspark and Malvern, should have increase their police surveillance and patrol frequency. Also, due to the significant high percentage of firearms involvement in homicides, we know it is necessary to improve the management of weapons registration and firearm-related laws.

The remainder of this paper is: Section 2 explains the data.

*Code and data are available at: https://github.com/justinteng1999/starter_folder-main

Table 1: First Ten Rows of a Dataset that Shows Homicide Occurrences

Year of Occurrences,	Neighborhood of Occurences	Type of Homicide
2004	Yonge-St.Clair (97)	Other
2004	Woburn (137)	Shooting
2004	Malvern (132)	Shooting
2004	Dovercourt-Wallace Emerson-Junction (93)	Shooting
2004	Rouge (131)	Shooting
2004	Downsview-Roding-CFB (26)	Stabbing
2004	Downsview-Roding-CFB (26)	Shooting
2004	Mount Olive-Silverstone-Jamestown (2)	Other
2004	Malvern (132)	Shooting
2004	Clairlea-Birchmount (120)	Shooting

2 Data

In order to better understand and analyze occurrences of homicides over the years among different neighbourhoods in Toronto, I obtain the dataset from the City of Toronto open Data Portal, using the 'opendatatoronto' package(Gelfand 2020) and statistical programming language R (R Core Team 2020). The raw data includes all reported homicides in 140 neighbourhoods within the Greater Toronto Area from year 2004-2020, as well as the type of homicide in each case. However, there are also some variables that are not necessary for our analysis, such as Event Unique Id, Division ID, Object ID, Hood ID and Geometry. Therefore, I first read in the csv file of the dataset and cleaned out the unnecessary variables from it. Then, I extracted out the required variables for the exploratory data analysis. To accomplish these goals, I used R (R Core Team 2020), tidyverse(Wickham et al. 2019) and dplyr(Wickham et al. 2021).

After cleaning data and selecting the necessary variables, we can now show an extract of the dataset (Table 1).

We are interested in the number of homicide occurrences each year, and the neighbourhoods with highest number of homicide cases, as well as the type of homicide in each case. This is the measure of how different types of homicide cases are distributed throughout the years. Most importantly, it has implications for allocation of police resources. In order to obtain the total number of homicides with each variable, I sum up the occurrences of homicide for every item in each variable, and then I used tibble (Muller and Wickham 2021) to store the values within each variable.

(Figure 1) shows the number of homicide occurrence each year from 2004-2020. The barplot was generated using ggplot(Wickham 2016). From the graph we obtained, we can see that the number of homicides were relatively high between 2004-2010. However, number of homicides were significantly low between 2011-2015, with 2011 as the minimum number of 51 homicide cases. Nonetheless, the number of homicides were slowly increasing each year until reaching its peak with the maximum of 97 cases on 2018. In the year of 2020, it is believed that the number of homicides had decreased in relationship to the rise of Covid-19 Pandemic(Armstrong and Jaffray 2021).

(Figure 2) shows the top 10 neighbourhoods with most number of homicide occurrences, and they are ordered accordingly. The barplot was generated using ggplot(Wickham 2016).

(Figure 3) shows number of homicide occurrences for each homicide type. The barplot was generated using ggplot(Wickham 2016). From the graph we obtained, we can see there are 608 homicide cases of that are caused by shooting, which is more than half of the total homicide occurrences. Having the high percentage of 52% of all homicide cases, firearm-related homicides was the main type of homicides throughout 2004-2020.

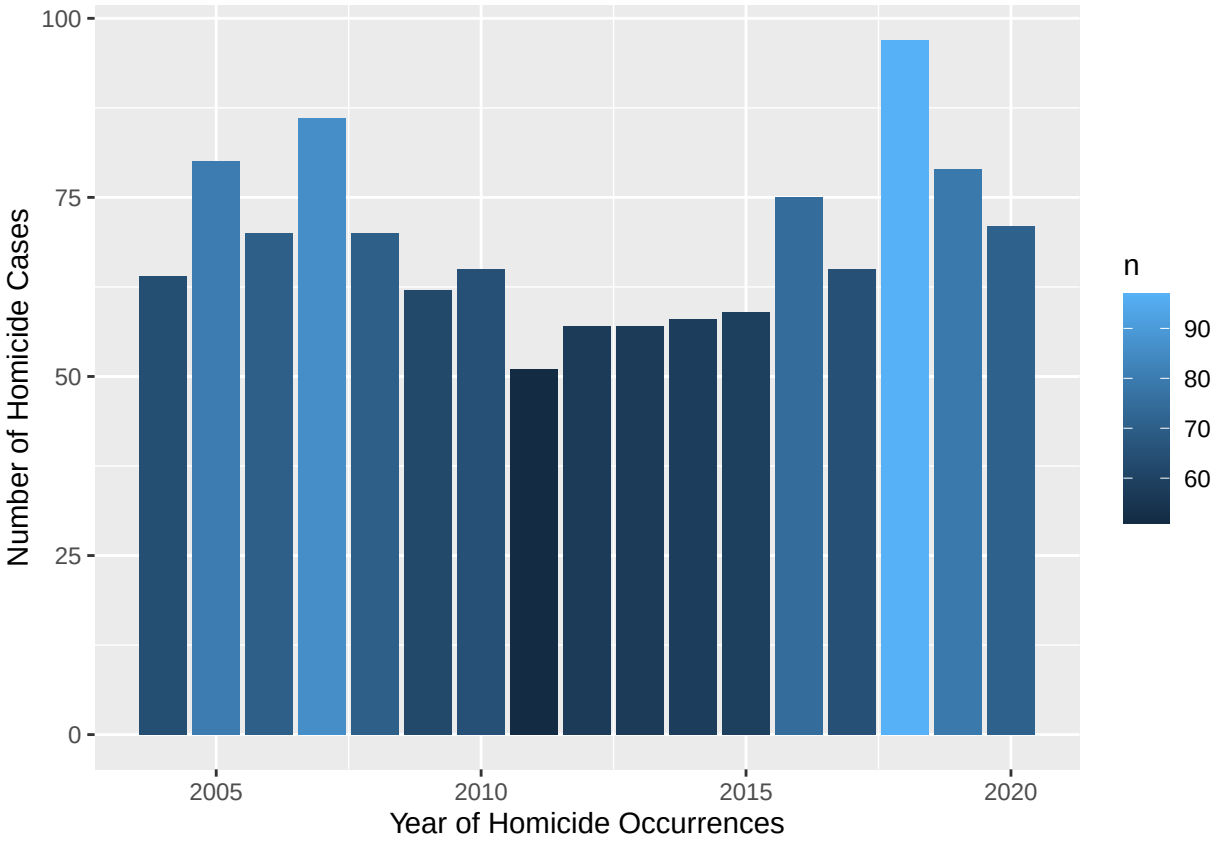

Figure 1: Homicide Occurrences each Year (2004-2020)

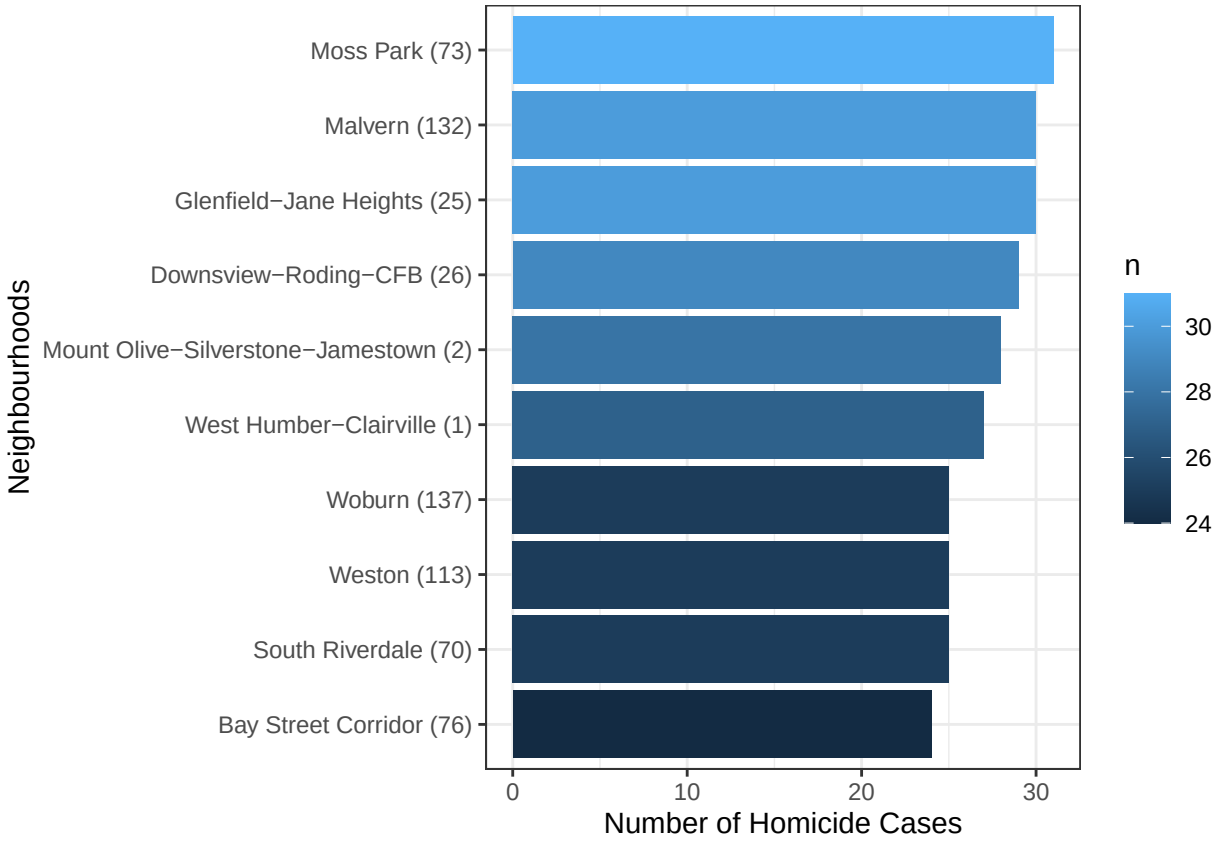

Figure 2: Top 10 Neighbourhoods with Most Homicide Occurrences (2004-2020)

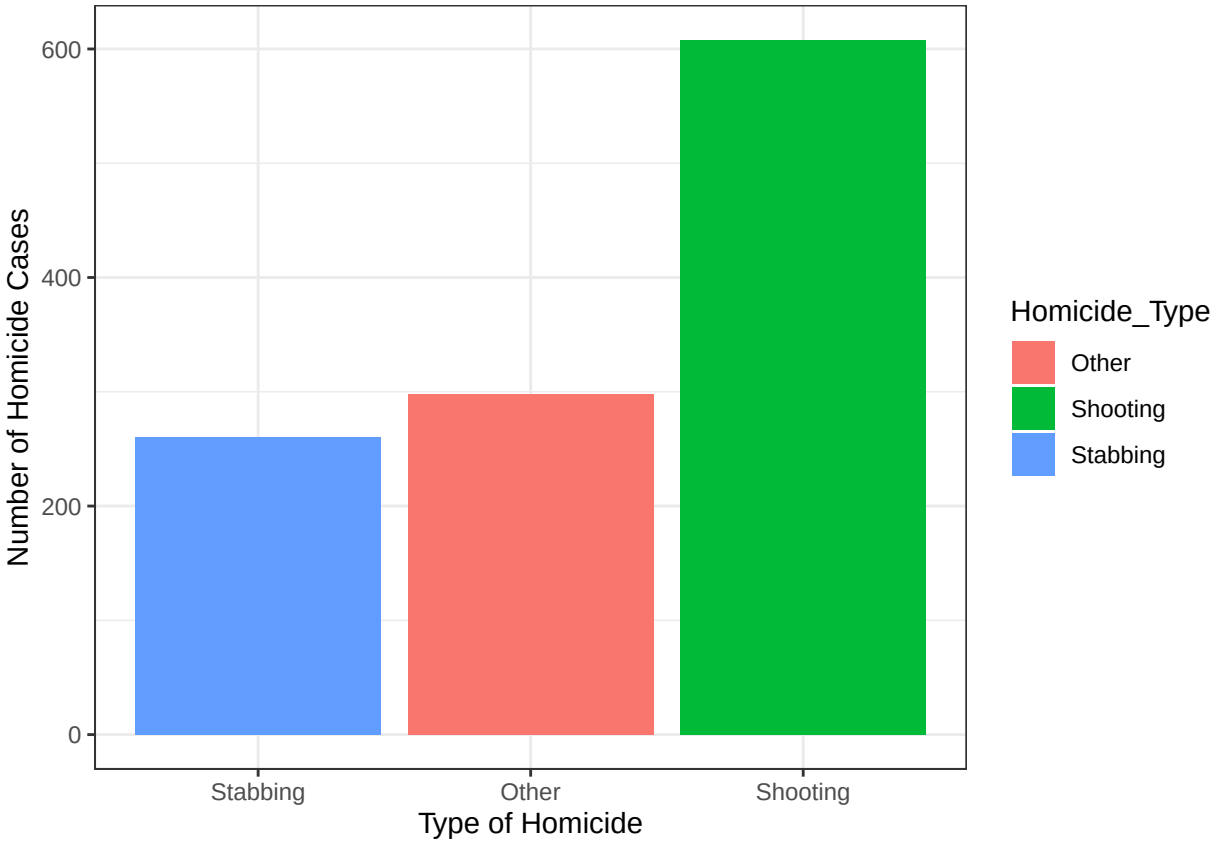

Figure 3: Type of Homicide Cases (2004-2020

References

Armstrong, Amelia, and Brianna Jaffray. 2021. *Homicide in Canada, 2020*. https://www150.statcan.gc.ca/n1/en/pub/85-002-x/2021001/article/00017-eng.pdf?st=cgxTVe4G.

Canada, Statistics. 2021. *Number and Rate of Homicide Victims, by Census Metropolitan Areas*. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=3510007101&pickMembers%5B0%5D=2.1&cubeTimeFrame.startYear=2002&cubeTimeFrame.endYear=2020&referencePeriods=20020101%2C20200101.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

Muller, Kirill, and Hadley Wickham. 2021. *Tibble: Simple Data Frames*. https://CRAN.R-project.org/package=tibble.

R Core Team. 2020. *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.

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 Francois, Lionel Henry, and Kirill Muller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.