Hannah Yu*

Hannah Yu

January 23, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

Table of contents

1 Introduction

The homicide rate plays a pivotal role in assessing the state of violence in society. In 2022, Canadian police services documented 874 cases of homicide, with 2.25 homicides per 100,000 people. With an 8% increase in the national homicide rate, this increase signifies the fourth consecutive annual rise in the national homicide rate (Canada 2024). As one of Canada's most vibrant and diverse cities, Toronto is the home of 2.93 million people. Toronto's safety and stability are extremely important for all the residents of the city. Understanding the patterns, factors, and implications is important in ensuring Toronto's public safety and establishing targeted interventions.

This research seeks to analyze the spatial and temporal distribution of Toronto's homicide cases. By investigating whether specific times or certain areas are prone to be associated with

*Code and data are available at:

homicide incidents, we want to shed light on the dynamics of and think about addressing and preventing future prevention for homicide cases.

The Police Annual Statistical Report - Homicides (Services 2024) from (Gelfand 2022) provides a record of all homicide cases in Toronto from 2004-2020 including the time of occurrence, location, and type of homicide. we could draw information about homicide cases and their related characteristics from this extensive dataset, and investigate Toronto's homicide patterns across time. In Section 2, I introduce the dataset and some general information. In (**sec1-subdata?**), I navigate the change in homicide case numbers across time with various characteristics and perform some analysis on the relationship between homicide cases and time variables. And in (**ref?**)(subsec2-data), I redo most of the analysis on the top five neighborhood with the most amount of homicide cases. Lastly, I discuss the implications and further discussions on these results in Section 3.

2 Data

To investigate the homicide cases in Toronto, I obtained the dataset " Police Annual Statistical Report - Homicides (Services 2024) from the Toronto Open Data Portal (Gelfand 2022). Data was cleaned and analyzed using the open source statistical programming language R (R Core Team 2023), with additional support from 'tidyverse' (Wickham et al. 2019), 'dplyr' (Wickham et al. 2023), 'janitor'(Firke 2023), 'readr' (Wickham, Hester, and Bryan 2024) , 'stringr' (Wickham 2023), 'lubridate' (Grolemund and Wickham 2011), 'ggplot2' (Wickham 2016), 'kableExtra' (Zhu 2021), ' ' Further details about data extraction and analyzation processes will be discussed in the following sections below.

This dataset includes a thorough list of alongside every homicide case in Toronto from the year 2004-2020. We cleaned the dataset and selected variables that are relevant to our analysis of homicide cases. Our variable of interest includes the year the offence occurred, type of homicide (shooting, stabbing, other), date of offence, and identifier of neighborhood using Toronto's 140 neighborhood structure. In addition to these variables, we also constructed three more variables: day of the week when the offence occurred (Monday-Sunday), season when the offence occurred (fall, winter, spring, summer), and count of homicide cases by year to further aid our data analysis Table 1. This dataset was last refreshed on January 11, 2024.

2.1 General Homicide Analysis

We listed out the number of homicide cases each year in Table 2 and created the summary statistics table of the homicide count per year. We found the year 2011 has the least cases of homicide with only 51 cases while the year 2018 with the total of 97 cases has the most number of homicide cases.

Table 1: Sample of Cleaned Homicide Data

Year	Homicide Type	Date	Hood ID	Day of the Week	Season	Homicide Count
2004	Other	2004-01-03	97	Saturday	Winter	64
2004	Shooting	2004-01-08	137	Thursday	Winter	64
2004	Shooting	2004-01-08	132	Thursday	Winter	64
2004	Shooting	2004-01-25	93	Sunday	Winter	64
2004	Shooting	2004-01-25	131	Sunday	Winter	64

Table 2: Total Homicide Cases Each Year

Total Homicide Cases Each Year

occurrence_year	total_cases
2004	64
2005	80
2006	70
2007	86
2008	70
2009	62
2010	65
2011	51
2012	57
2013	57
2014	58
2015	59
2016	75
2017	65
2018	97
2019	79
2020	71

Of all the homicide cases, the cases caused by shooting is significantly higher than other type of homicides. In the years of our examination, shooting takes up more than 600 cases of homicides while stabbing takes over more than 250, and all other types together takes up around 300 (Figure 1).

We then shift our attention to the homicide trend over time. Figure 2 portrays the trend of homicide cases throughout the span of almost 2 decades. With no obvious trend, We do not observe a clear increase or decrease of cases across time. However, the homicide cases are decreasing after it peaked in 2018.

Following the grand picture across time, Figure 3 breaks down the homicide cases over time into further details based on the type of homicide. The trend of the three types of homicide cases generally follow the overall trend in Figure 2. Shooting always prevail over other types of homicide cases even in its lowest year. We conclude that there is no clear increase or decrease of homicide across time.

We want to determine whether if there are a specific time in the year when homicides are more likely to occurr. To investigate our interest, we created Figure 4 that displays the homicide counts by every day of the week. Sunday with a total of 200 cases is the day with the most cases following by Saturday and Friday with 197 and 175 respectively. Thursday is the day

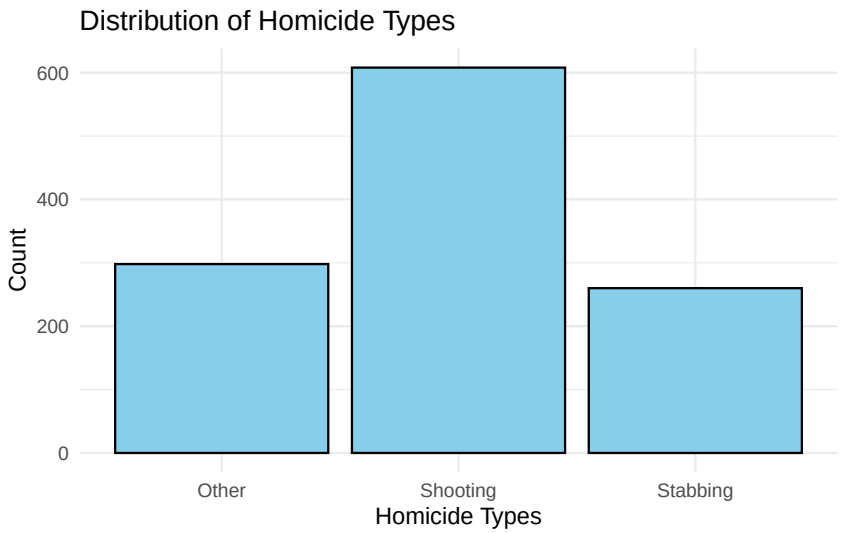

Figure 1: Distribution of Homicide Types

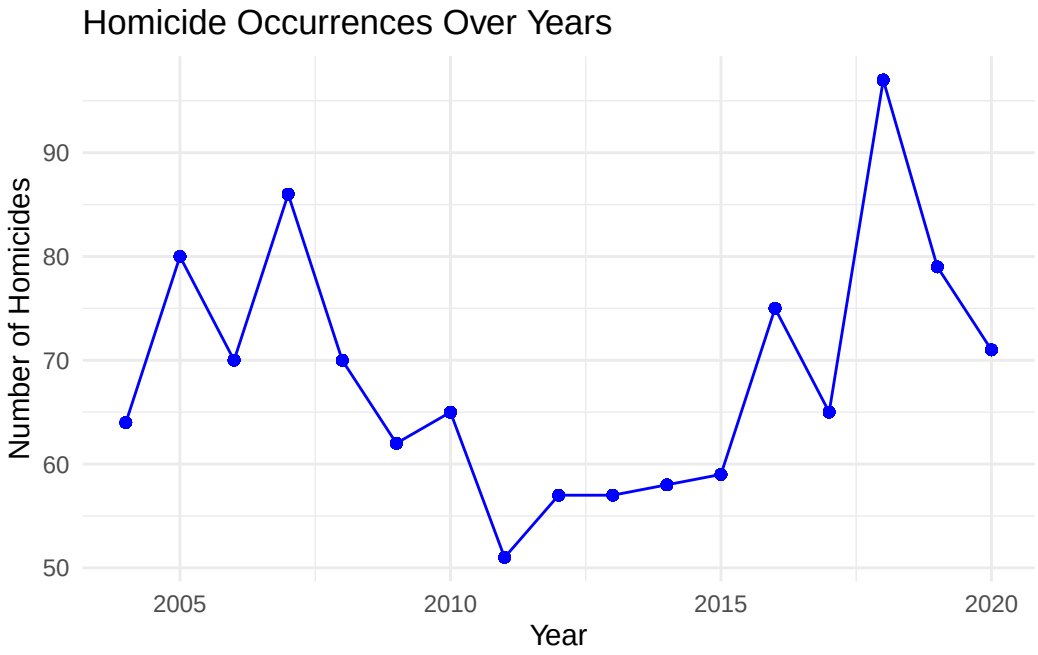

Figure 2: Homicide Occurrences Over Years

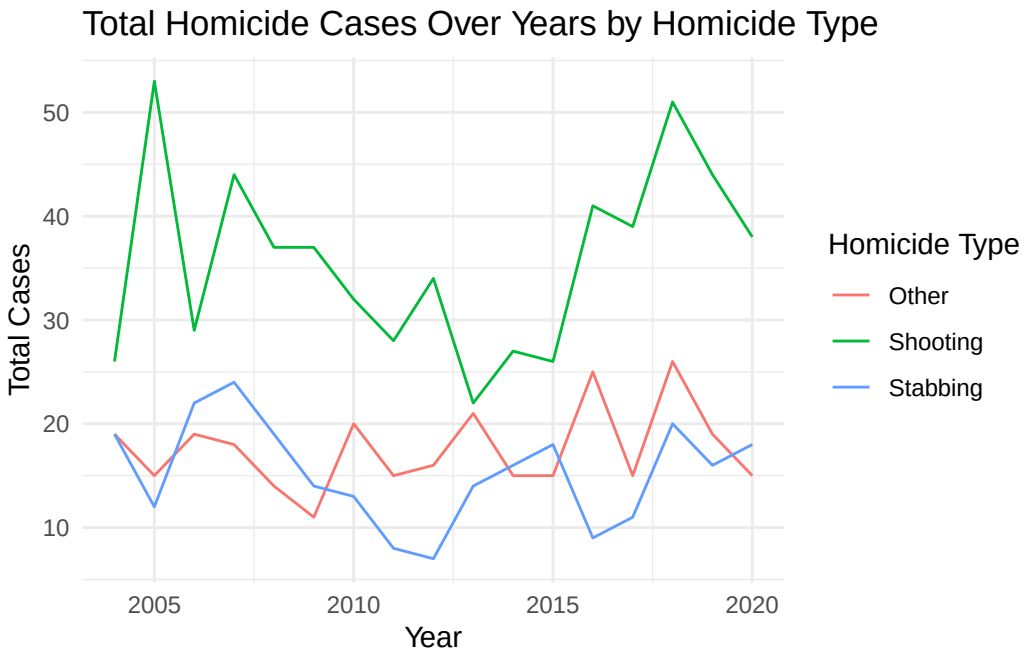

Figure 3: Total Homicide Cases Over Years by Homicide Type

with the least cases with 124 cases. We conclude that homicides are slightly more likely to occur in weekends.

We perform similar analysis with the count of homicide cases by season and found that the seasonal data are distributed somewhat evenly with fall having slightly more cases than other weather but the difference is to small and not worth noting Figure 5.

To investigate the relationship between geographical location and homicide cases, we constructed Figure 6. This bar plot illustrates the distribution of homicide cases in Toronto's 140 neighborhoods. This visualization enables the identification of patterns or areas of concerns in the city. While most neighborhoods have fewer than 10 cases, we notice that several neighborhoods stand out with significantly higher homicide incidences compared to the rest.

2.2 Homicide Analysis in Top Five Neighborhood

Having observed the obvious significantly higher cases of homicide in several neighborhoods, I selected the five neighborhoods with the highest amount of homicide cases and performed the following analysis that focus on these five neighborhoods. The ids of the five neighborhoods are 73, 25, 132, 26, 2 respectively with neighborhood 73 having the highest number of cases of 31 and neighborhood 2 with 28 cases. Through these analysis, we could understand if there is specific patterns or trends in the occurrence of homicide incidences in these top five neighborhoods and genearlize the conclusion in larger areas.

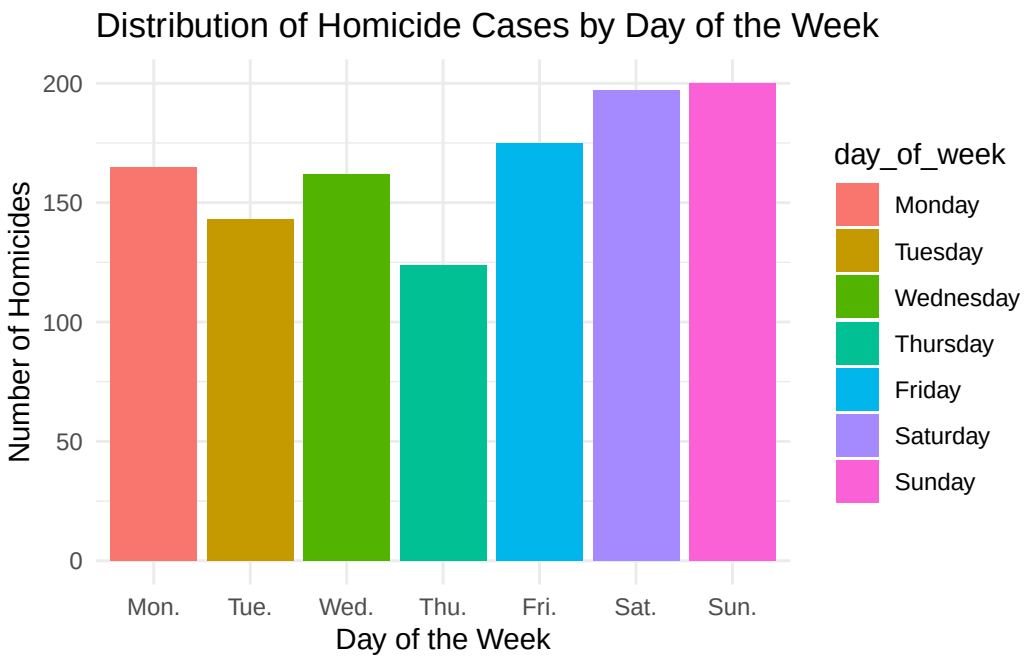

Figure 4: Distribution of Homicide Cases by Day of the Week

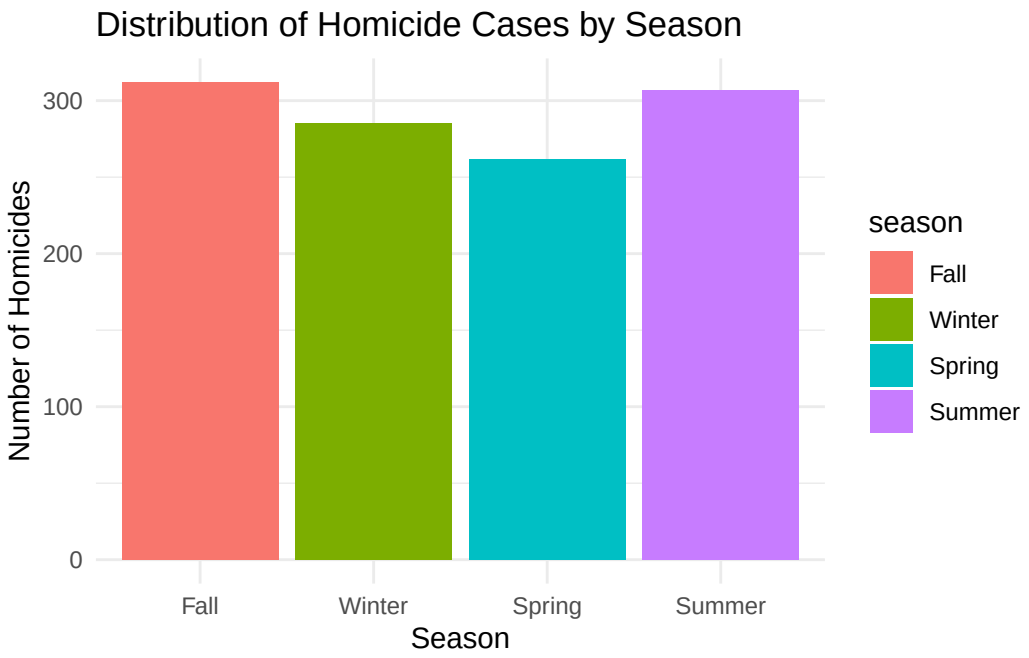

Figure 5: Distribution of Homicide Cases by Season

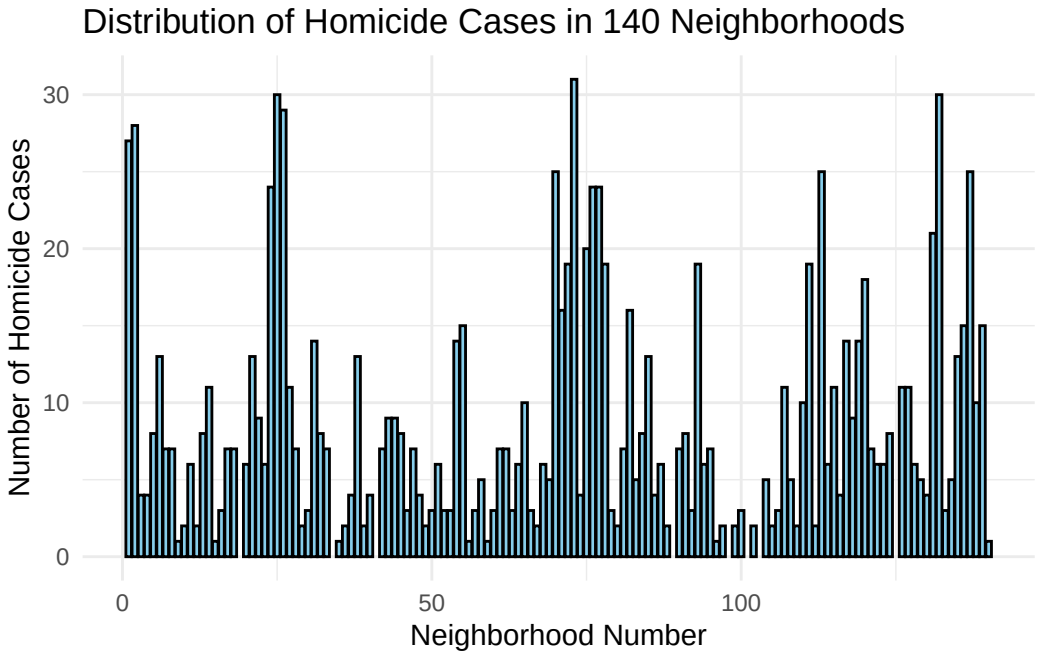

Figure 6: Distribution of Homicide in Neighborhoods

In order to observe the relationship between relating factors and homicide cases, I investigate the correlation of the occurrence of each case with each day of the week, each season, and type of homicide (Table 4, Table 3, Table 5). We do not observe a specific trend in the relation between homicide cases and day of the week. Just like our previous analysis on Figure 5, the correlation between the number of homicide cases and with a specific day of the week or the weekends are also not noticeable. We could also attribute this result to the fact that our sample size is not big enough to observe a difference as there are only limited amount of homicide cases in the city. However, we do notice that just like the previous, the amount of homicide caused by shooting is higher than any other methods in each of the top five neighborhoods. This is a confirmation that shooting is the main type of homicide in Toronto.

Table 3: Homicide Cases in Most Frequent Neighborhoods by Season

Homicide Cases in Most Frequent Neighborhoods by Season

	Top 5 Neighborhoods				
season	73	25	132	26	2
Fall	4	9	5	9	10
Winter	7	5	15	9	9
Spring	5	6	4	4	5
Summer	15	10	6	7	4

Table 4: homicide cases in most frequent neighborhoods by day of the week

Homicide Cases in Most Frequent Neighborhoods by Day of the Week

	Top 5 Neighborhoods				
day_of_week	73	25	132	26	2
Monday	8	7	3	2	4
Tuesday	3	7	6	2	5

	Top 5 Neighborhoods				
homicide_type	73	25	132	26	2
Other	10	2	10	3	3
Shooting	13	21	20	17	21
Stabbing	8	7	0	9	4

‘

3 Discussion

The result of our general analysis of homicide cases by season and by each day of the week throughout Toronto generally align with our specific focus into the top five highest homicide cases neighborhoods. We cannot obverse any significant correlation between homicidece occurrences and particular timings. Homicides may have a slightly higher chance in occurring on a weekend than weekdays and are not more likely to occur in a particular season.

However, one characteristics that stand out throughout the entirety of our analysis is that how shooting is always the highest type of homicide. It started as the highest cause of homicide in 2004 and remained until 2020. We could conclude that is the highest cause of homicide across time.

We do not observe a general increase or decrease of homicide cases throughout time. However, afte the number of homicide cases peaked in 2018, the cases are experiencing a decline. Cases caused by shooting are also in decline during the same period. This decline may be associated with how Toronto gun violence has dropped by 30% over the past five years (Balintec (2023)).

Our analysis would be more comprehensive if there are more detailed data about more relevant characteristics about homicides and more cases of homicide. But due to the scope of our dataset and the nature of homicide occurrences being generally rare compared to other research topics, it is very hard to give out a firm statement such as a certain factor causes more homicide incidences or homicide incidences are more likely to happen on a certain dates.

References

Balintec, Vanessa. 2023. *Toronto Gun Violence Dropped 30*. https://www.cbc.ca/news/canada/toronto/gun-violence-trend-2023-toronto-1.7064901.

Canada, Statistics. 2024. *Homicide Trends in Canada, 2022*. https://www150.statcan.gc.ca/n1/daily-quotidien/231129/dq231129b-eng.htm.

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/.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. https://www.jstatsoft.org/v40/i03/.

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

Services, Toronto Police. 2024. *Police Annual Statistical Report - Homicides*. https://open.toronto.ca/dataset/police-annual-statistical-report-homicide/.

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

———. 2023. *Stringr: Simple, Consistent Wrappers for Common String Operations*. https://stringr.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 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.

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