Homicide in Toronto Neighbourhoods - Investigation on Occurrence Trend Over Years, Months, and Types of Homicide

Yuzhi Pi

02/06/2022

Abstract

Data on Toronto Homicide cases from 2005 to 2020 was pulled from Open Data Toronto. The dataset is cleaned and used to analyze the trend in the occurrence of Homicide in Toronto Neighbourhoods. Through the analysis process, there were three major findings: 1) there is a downward trend in the number of homicide cases from 2018 to 2020; 2) homicide is more likely to occur in the summer and early fall, while Division 31 in Toronto is the division with the most homicide cases, and 3) shooting is the most common type of homicide. These findings allow the user of the dataset to have a better understanding of the trends in homicide in Toronto, implementing better policies to more effectively allocate police resources in order to prevent and react to crime.

Introduction

Home to approximately 2.8 million people, the city of Toronto is the largest city in Canada and the fourth-largest city across North America. With its great reputation in education, finance, and cultural diversity, Toronto attracts many international students and immigrants, including myself, every year to Canada - in 2016, 47% of Toronto's population is composed of immigrants (Statistics Canada 2016). Immigration is a key factor for achieving economic prosperity in Canada. According to the Government of Canada, immigrants increase the size of the labour force, tax revenue, and general goods consumption, which would ultimately increase Canada's GDP (Government of Canada 2022).

With this, the safety of Toronto is often seen as a key evaluative criterion for one to consider when moving to Toronto or to Canada. Among all crimes, homicide is the most dangerous and immoral crime type. Although Toronto is not notorious for being a violent crime city, Toronto's beginning for this new year has been brutal - there have already been 5 homicide cases taken place during the first three days of 2022 (Ngabo 2022) . There has been criticism over whether this is a signal that the city has become more violent, which would endanger the safety of the Torontonians and decrease Toronto residents' living quality. However, more analysis needs to be conducted to portray the true picture of Toronto's homicide case trends, and this is one of the major motivation of this report.

More specifically, the analysis in this report will primarily focus on three aspects of Toronto's homicide trends - its occurrence (measured in both years and months), the location (segregated based on Toronto police districts), and types of homicide. For this report, we are specifically interested in investigating the trend of homicide, to testify whether there has been a rising trend. Other than this, the statistical analysis will provide an insight into the occurrence of homicide - whether homicides are especially likely to occur in a certain month or location. Understanding the occurrence pattern of homicide will significantly improve Toronto police's ability to react, or even prevent the crime. More budget and resources could be allocated to the locations that are more likely for crime to take place, while the police could also increase the frequency of patrol duty during more dangerous months. Lastly, this report will provide a better understanding of the types of homicide, which is especially important for politicians. This could help them to understand whether the "gun ban" has been effective at cutting the frequency of shootings, and what else should the government

consider other than banning certain types of guns from licensed owners in order to better protect Canadians (Lilley 2022).

All the analysis in this report will be based on data obtained from the Police Annual Statistical Report, published by Toronto Police Service on Open Data Toronto (Toronto Police Services 2021), and analyzed using R (R Core Team 2021). Resources regarding this report are available on GitHub.[1]

Data Section

Methodology and Approach

The homicide dataset used in this report includes all reported homicide cases in the city of Toronto from 2004 to 2020. For a case to be included in the dataset, it has to be classified as homicide by the Toronto Police. From the open data documentation published by Toronto Police Service, a homicide occurs when "a person directly or indirectly, by any means, causes the death of another human being. Deaths caused by criminal negligence, suicide, or accidental or justifiable homicide (i.e self-defence) are not included" (Toronto Police Service 2022). The dataset is compiled based on the Homicide Squad Case List Log, which is maintained by the Toronto Police Service. The log itself contains the information on the victim, types of homicide, location, etc. Then the information will be updated to the annual statistical report and the dataset used in this report.

However, the fact that this dataset only includes reported homicide could lead to the bias where the number of homicide does not reflect the actual homicide cases that have taken place. A portion of the homicide cases are sometimes unreported, therefore not recorded by the police or being entered into the dataset. For example, homicide cases that relate to domestic abuse often go unreported. Most domestic abuse victims, mostly female, suffer from long-term physical and mental abuse by their partners. It is not rare that the abuse evolves to homicide, but such homicide is rarely reported or investigated by the police, given the fact that most victims were already in an unstable state that they typically decease without anyone's notice, plus such cases are frequently treated as "family businesses" (Gracia 2004).

Moreover, homicide towards children is another homicide category that is often overlooked and therefore underreported. Homicide against children are hard to detect by the police as it is extremely challenging for the victim, in this situation, the children, to report the crime or the abusive activities that eventually lead to homicide. It is also a lot easier for the offenders, most of the cases are childrens' parents, family members to report such crime with a false cause, therefore avoiding the investigation and will not be recorded by the police (UNODC 2019). Hence, the major bias of this dataset is that it could be missing records on intrafamilial homicide, including the murder of intimate partner and children.

On the other hand, the strength of the source is that it is maintained by the Toronto Police Service, which is the organization that should be able to obtain homicide data that is the closest to the actual homicide data. Although, as discussed above, a portion of the homicide case will be underreported and therefore excluded from the list, the dataset will contain 100% of the reported homicide. Consequently, the data used in this report will be the closest representation of the actual homicide cases, if considering the unreported cases.

Data Characteristic

The original Homicide dataset from Open Data Toronto includes 1166 entries of data and 10 columns in total. Below is an in-depth explanation on the meaning of each column, while the title of the column as shown per dataset is listed in brackets. Each row in the dataset represents one victim.
1) an ID column that acts as a row identifier for the dataset (_id)
2) a column that lists the occurrence number, which is unique for each case (Event_Unique_Id). For a case with two victims, it will take up two rows in the dataset but the event ID will be the same for the two rows
3) a column that lists the occurrence year of the homicide (Occurrence_year)

---

[1] https://github.com/YuzhiPi/STA304

4) a column that indicates the specific police division where the homicide took place (Division)

5) a column that specifies the types of homicide, classified into three categories - shooting, stabbing, and other (Homicide_Type)

6) a column that specifies the number identifier of the neighbourhood where the homicide took place (Hood_ID)

7) a column that specifies the date of homicide in the format of yyyy-mm-dd-time, but time has been set to default at 00:00 (Homicide_Date)

8) a column that identifies the actual name of the neighbourhood where the homicide took place (Neighbourhood)

9) a column that outlines the autogenerated unique record identifier (ObjectId)

10) a column that includes the longitude and latitude coordination (offset to the nearest intersection) of the location where the homicide took place (geometry)

For the purpose of this report, the Homicide dataset was modified and cleaned using R with package dplyr (Wickham et al. 2021) and tidyverse (Wickham et al. 2019). Column 1, 2, and 10 were removed as they are identifiers. Column 10 is also removed, as it is easier to generalize the finding on crime trends using police divisions and neighbourhoods. It would be very challenging to summarize trends by looking at specific geographical points. Column 7 was decomposed by separating the date based on "-" using function "separate". In this way, three new columns will be generated and they were named "year", "month", and others. Only the month column is kept and the other two new columns were deleted.

Analysis on Key Variables

All graphs presented below were constructed using ggplot2 (Wickham 2016).

General Trend of Homicidee Occurance

As Toronto's start of 2022 has been significantly more violent than previous years, some have criticized that this indicates an overall uprising trend in Toronto's homicide case. However, the dataset suggests otherwise. Figure 1 displays the historical homicide crime case per year, from 2004 to 2020. From the graph, there is an obvious downward trend in the number of homicide cases from 2018 to 2020 - the number of cases has dropped by 26.8% from 2018 to 2020. This means that at least given the currently available data, it should not be concluded that there is an uprising trend in homicide cases overall. In reality, based on past data, Toronto's average homicide cases range between five to eight for the month of January (Ngabo 2022). Therefore, there is a lack of evidence to conclude the overall trend based on the number of killings during the first three days of 2022.

Figure 1: Summmary of Homicide by Occurrence Year
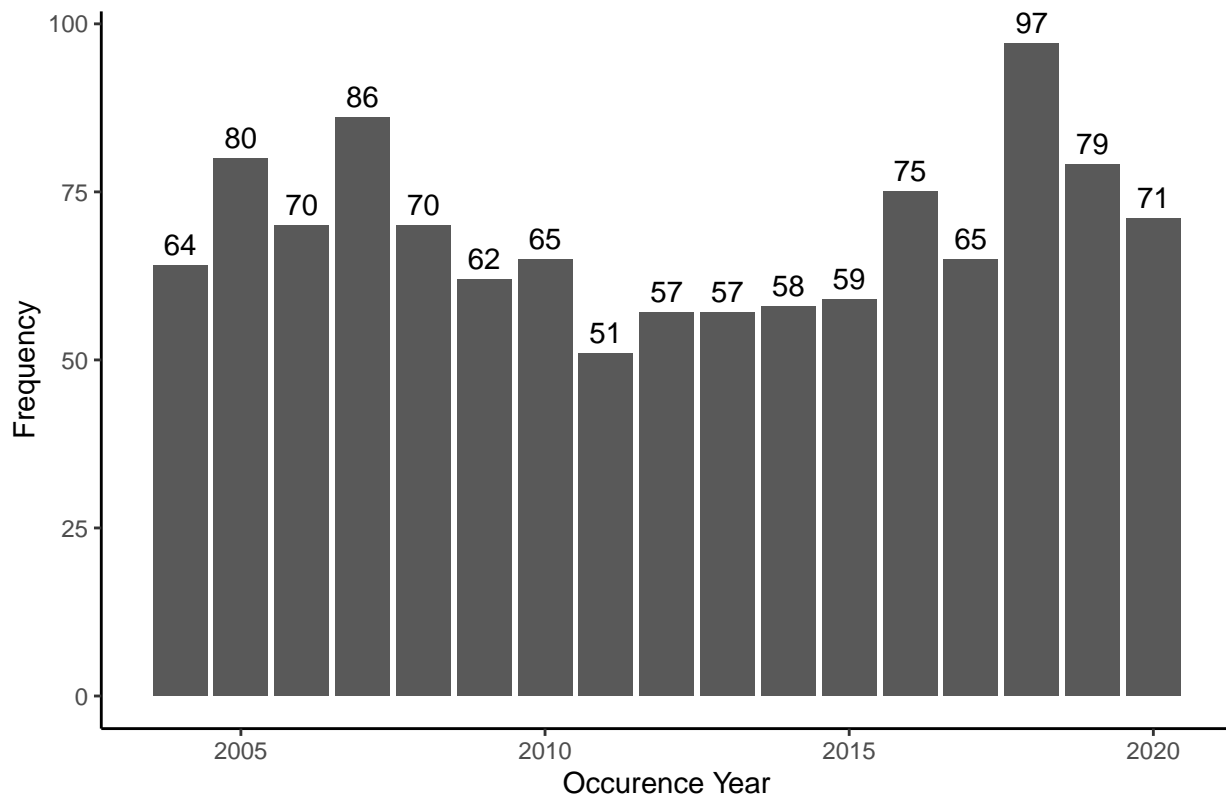
On the other hand, it is noticeable that the case number increased dramatically in 2018 from 65 cases in 2017 to 97 cases in 2018. There is an obvious trend that the number of cases usually ranges around 60 from 2009 to 2017 (62, 65, 51, 57, 57, 58, 59, 75, 65 cases, respectively), with a sudden increase to 97 in 2017. This is potentially due to the uprising amount of gun violence in the city. Toronto's police chief Mark Saunders also commented on the spike of homicide cases in 2018, making the connection between the surge in numbers of handguns recovered on the street (from 292 in 2017 to 514 in 2018) (McQuigge 2019).

Seasonality Analysis of Homicidee and Trends in Occurrence Location

In terms of the potential trends in the occurrence of homicide, this report will look at two factors: the occurrence of homicide based on months and geometric locations (Toronto Police Districts).

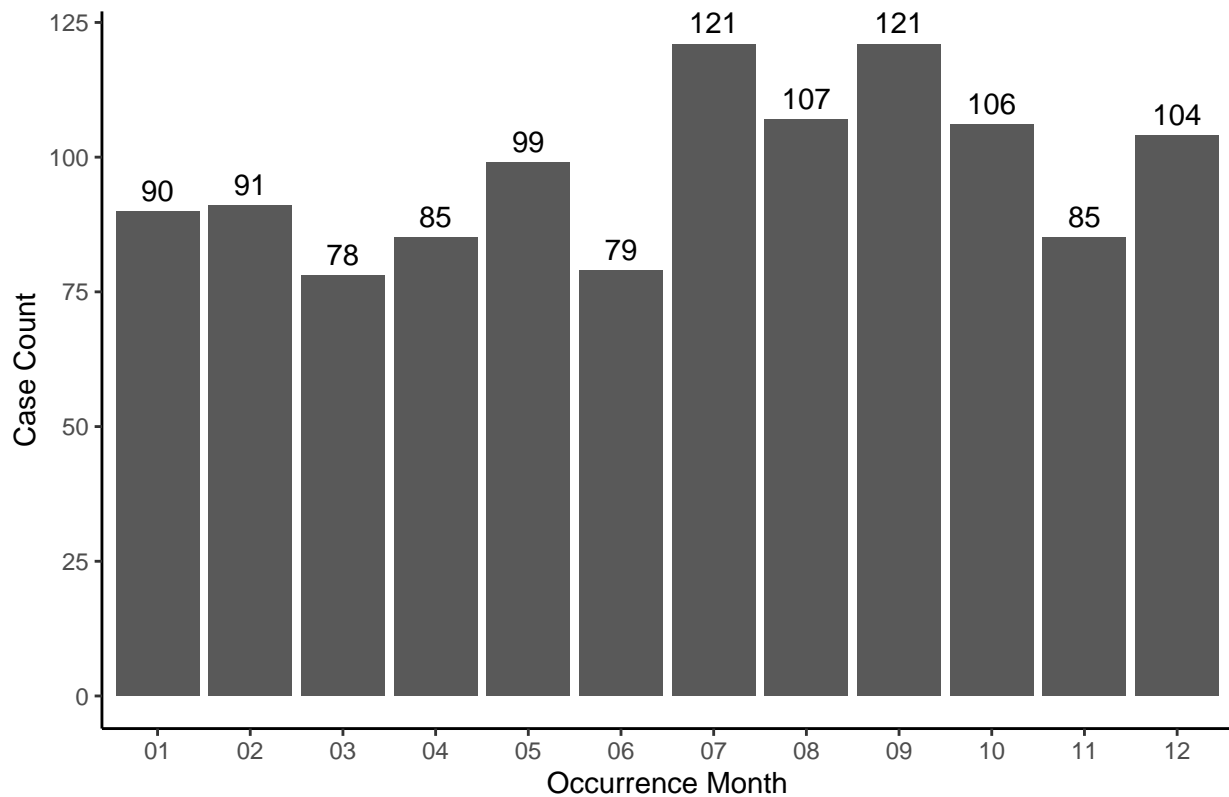## Figure 2: Summmary of Homicide Occurrence Month

Figure 2 summarizes the occurrence of homicide based on twelve months, and the graph is created using the month data obtained earlier from the data cleaning process. Over the course of 17 years, it is clear that more crime takes place during summer and early fall. The average number of cases of homicide is 113.75 for the four months of July, August, September and October, while the average number of cases for the rest of the year is 88.88. This phenomenon is often described as the seasonality of crime, and the figure above seems to support this theory.

Violent crimes, including homicide, are more likely to peak during summer and also in December. This is closely connected to the temperature and climate, as they largely shape people's activity routine. During warmer times, people are more likely to go out to enjoy the weather. Summer and Christmas holiday is another factor why violent crimes are more common during those months, as the climate and the holidays both increases outdoor activities, which ultimately increases the chances of being exposed to violent crime (McDowall, Loftin, and Pate 2011).

Figure 3: Summmary of Homicide Occurrence Location
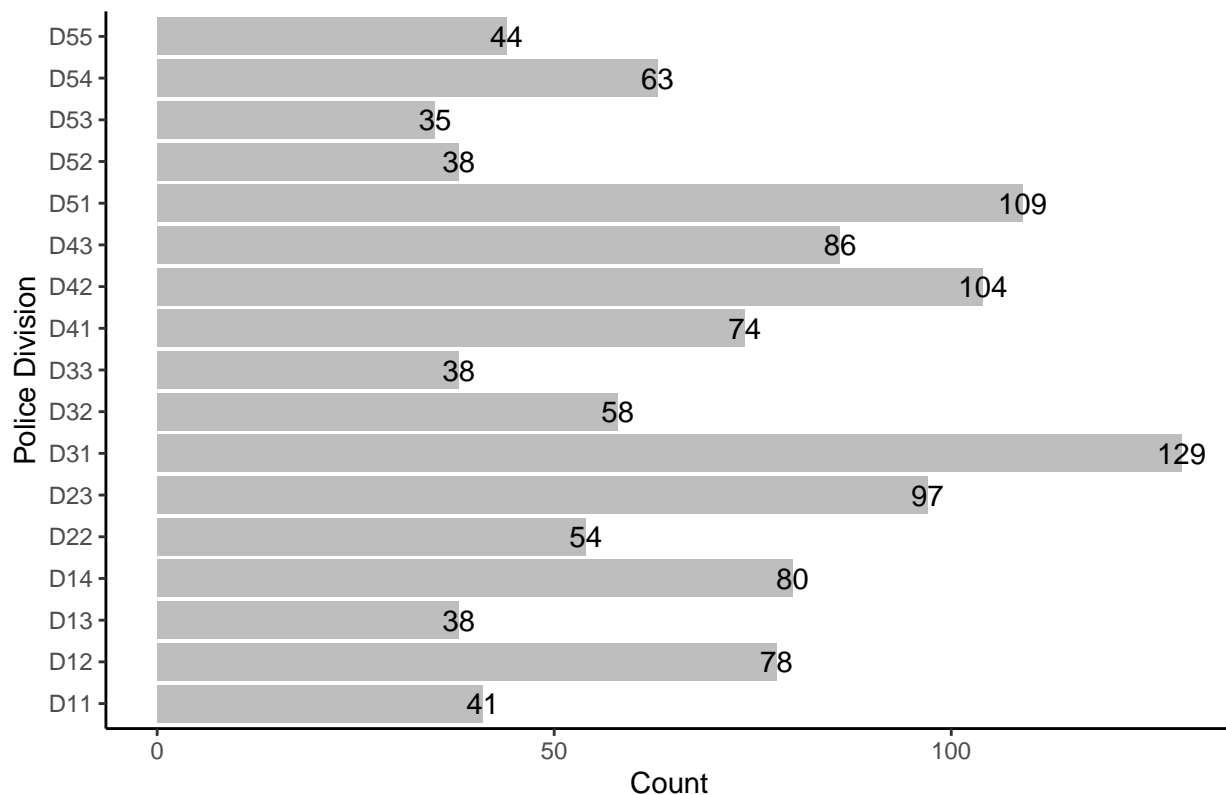
Other than the seasonality of homicide, the dataset also allows us to further investigate the hotspot of homicide. Figure 3 summarizes the total homicide cases based on the Police District the crime took place. The top three police districts with the most homicide cases are D31, D51, and D42. D31 has the largest total homicide cases, with 129 victims in total over the 17 years. In comparison, other police divisions such as D53, D52, D33, D13 and D11 see much less killing in the past, with average victims of 38 from 2007 to 2020.

Figure 3 outlines the tremendous difference in the number of homicide cases between different police districts and neighbourhoods. It is clear that some divisions are significantly more dangerous and deadly than others. There could be multiple factors influencing the number of homicides in a division, but the disparity suggests that there is a great need for Toronto Police to allocate more resources to these divisions.

Trends in Different Types of Homicide

The dataset also includes information on the types of homicide, and all the cases are classfied in three categories: shooting, stabbing, and other. Table one outlines the total number of cases of each type of homicide per year, and the proportion of each type for the year. I constructed the table using the package knitr (Xie 2021) and kableExtra (Zhu 2021). By grouping the data first by its occurrence year using function "group_by" and count each type of homicide within each year using function "count", I was able to get the number of each type of homicide case per year and then calculate the percentage accordingly. The functions are from package dplyr (Wickham et al. 2021).

The result of the calculation is presented in table one.

Table 1: Summary Table of Homicide Type and Percentage

Occurance Year	Homicide Type	Number of Cases Per Year	% of Total Case For the Year
2004	Other	19	29.7 %
2004	Shooting	26	40.6 %

6

Occurance Year	Homicide Type	Number of Cases Per Year	% of Total Case For the Year
2004	Stabbing	19	29.7 %
2005	Other	15	18.8 %
2005	Shooting	53	66.2 %
2005	Stabbing	12	15 %
2006	Other	19	27.1 %
2006	Shooting	29	41.4 %
2006	Stabbing	22	31.4 %
2007	Other	18	20.9 %
2007	Shooting	44	51.2 %
2007	Stabbing	24	27.9 %
2008	Other	14	20 %
2008	Shooting	37	52.9 %
2008	Stabbing	19	27.1 %
2009	Other	11	17.7 %
2009	Shooting	37	59.7 %
2009	Stabbing	14	22.6 %
2010	Other	20	30.8 %
2010	Shooting	32	49.2 %
2010	Stabbing	13	20 %
2011	Other	15	29.4 %
2011	Shooting	28	54.9 %
2011	Stabbing	8	15.7 %
2012	Other	16	28.1 %
2012	Shooting	34	59.6 %
2012	Stabbing	7	12.3 %
2013	Other	21	36.8 %
2013	Shooting	22	38.6 %
2013	Stabbing	14	24.6 %
2014	Other	15	25.9 %
2014	Shooting	27	46.6 %
2014	Stabbing	16	27.6 %
2015	Other	15	25.4 %
2015	Shooting	26	44.1 %
2015	Stabbing	18	30.5 %
2016	Other	25	33.3 %
2016	Shooting	41	54.7 %
2016	Stabbing	9	12 %
2017	Other	15	23.1 %
2017	Shooting	39	60 %
2017	Stabbing	11	16.9 %
2018	Other	26	26.8 %
2018	Shooting	51	52.6 %
2018	Stabbing	20	20.6 %
2019	Other	19	24.1 %
2019	Shooting	44	55.7 %
2019	Stabbing	16	20.3 %
2020	Other	15	21.1 %
2020	Shooting	38	53.5 %
2020	Stabbing	18	25.4 %

Although the table is able to present all data with accurate numbers, to further analyze the trend, stacked line graph will be more appropriate at visulizing the trend. Figure 4 illustrates the actual number of cases for

different types of homicide from 2004 to 2020. It is clear that the numbers of homicide caused by shooting have alway surpass the other two categories. It is no surprise that most of the homicide victim is killed due to gun violence, as weapon itself is extremly leathal. Gun violence saw a decreasing trend from 2005 to 2015, where the number of homicide caused by shooting dcreased by 50.9% in total over the 10 years. However the number began to increase since 2016 and reached a second peak in 2018, totalling 51 cases of shooting.

On the other hand, numbers of other types of homicide have stayed relatively steady in the past years. Stabbing used to be the second common ways for people to get killed, as the number of stabbing cases were larger than all other types from 2006 to 2009. However, since 2010, stabbing has become less dominate. The numbers of homicide caused by stabbing were below other means of homicide between both 2010-2014 and 2016-2019.

Figure 4: Count of Different Homicide Type
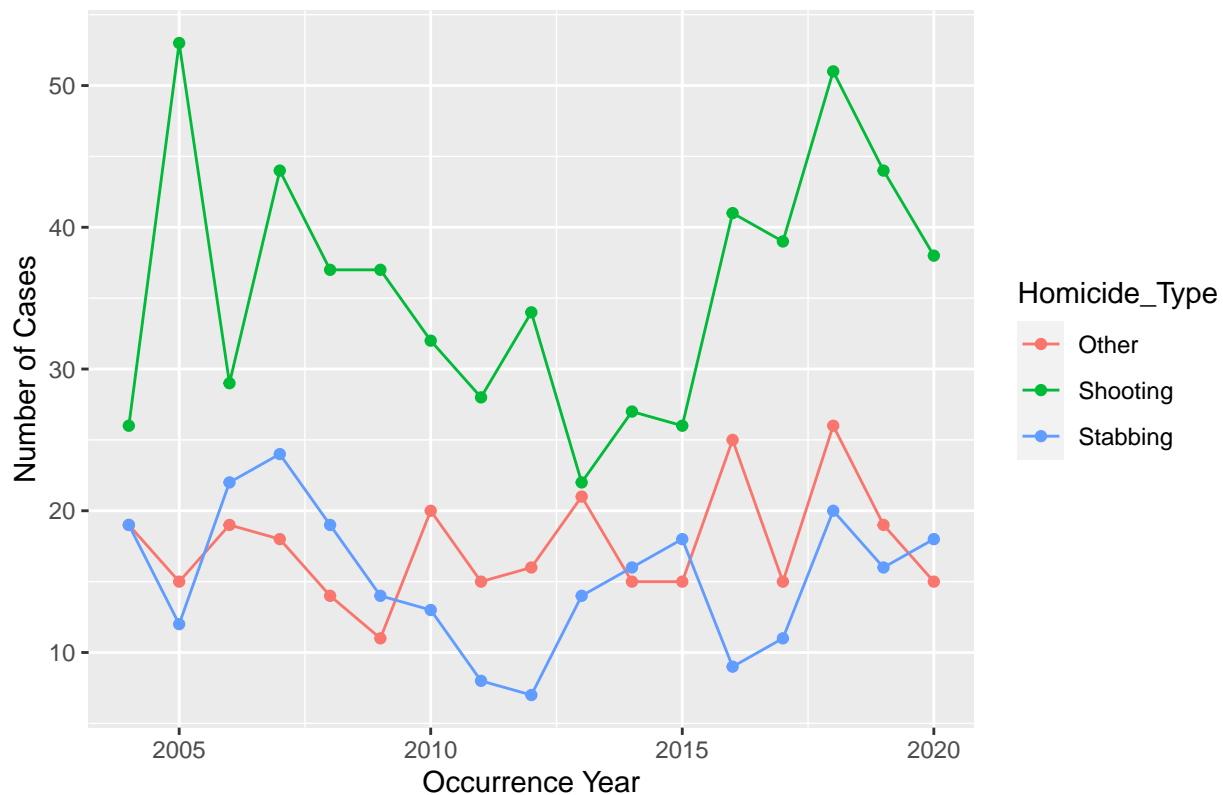
To better illustrate the dominance of shooting, Figure 5 illustrates the proportion of each homicide type by years. Shooting on average makes up 51.85% of all homicide cases over the 17 years. It seems like although firearms are fedreally controlled in Canada, it is still the most common reason for people to get murdered. In response to this, the government of Canada has implemented a series of policies to tighten the current gun control laws - in 2018, the federal government reinforced Bill C-71 so that gun dealers are required to keep records and ensure their availability to the police, and in 2020, the federal government announced a ban against around 1,500 models of semi-automateed rifles under the Criminal Code. These could contribute to the recent decline in homicide cases caused by shooting from 2018 to 2020 as shown per Figure 4. However, shooting is still the primary type of homicide and the proportion did not change much from 2018 to 2020 as shown per Figure 5.

Figure 5: Percentage of Different Homicide Type
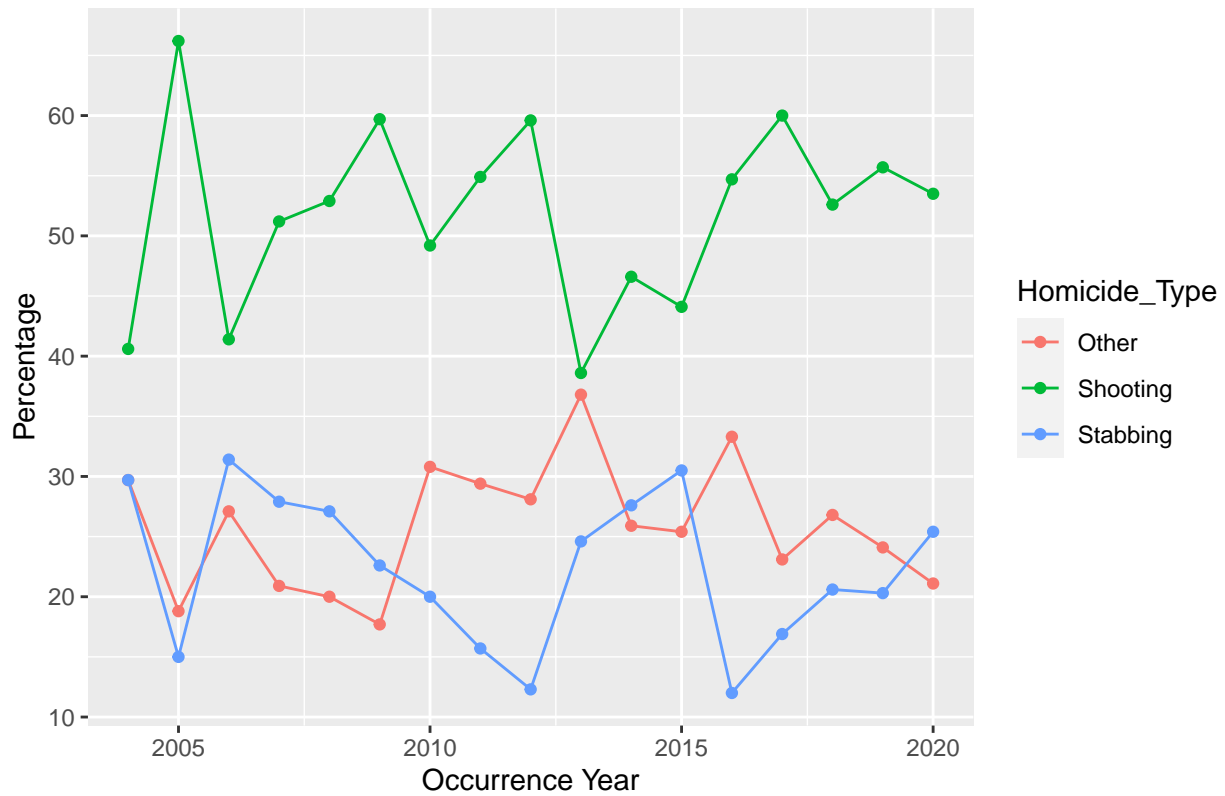
Overall, the dataset illustrates the current trend in homicide cases in Toronto. The data does not confirm the spectualation of the uprising trend in homicide, while it provided us with the insight on the occurrence of the crime. Toronto sees more homicide cases during summer and early fall, and this aligns with others' previous research on the seasonality of crime. Other than season, some police division sees significantly more crime than other divisions as well, with D31 tops the chart with 129 homicide victims over the 17 years. With this, the government should definitely consider to allocate more police resource during summer times and in those divisions to better prevent and react to the crime. The effectiveness of the gun controlling policy is also reflected in the numbers of people killed by shooting - shooting count has decreased since 2018. Nevertheless, shooting remains the most common type of homicide, and more effort could be made to reinforce the control on gun smuggling to cut off the sources of gun trades.

Reference

Government of Canada. 2022. "#ImmigrationMatters: Canada's Immigration Track Record." https://www.canada.ca/en/immigration-refugees-citizenship/campaigns/immigration-matters/track-record.html.

Gracia, Enrique. 2004. "Unreported Cases of Domestic Violence Against Women: Towards an Epidemiology of Social Silence, Tolerance, and Inhibition." *J Epidemiol Community Health*, 536. https://doi.org/10.1136/jech.2003.019604.

Lilley, Brian. 2022. "Apparently Criminals Pay No Heed to Trudeau's Gun Ban Announcements." https://torontosun.com/opinion/lilley-apparently-criminals-pay-no-heed-to-trudeaus-gun-ban-announcements.

McDowall, David, Colin Loftin, and Matthew Pate. 2011. "Seasonal Cycles in Crime, and Their Variability." *J Quant Criminol*, 390. https://doi.org/10.1007/s10940-011-9145-7.

McQuigge, Michelle. 2019. "Police Chief Says High Homicide Rate in 2018 Toronto Was a 'Unique' Year." https://globalnews.ca/news/4815150/police-chief-says-high-homicide-rate-in-2018-toronto-was-a-unique-year/.

Ngabo, Gilbert. 2022. "Toronto Has Had a Violent Start to 2022. Why It's Too Early to Call It a Trend." https://www.thestar.com/news/gta/2022/01/05/toronto-has-had-a-violent-start-to-2022-why-its-too-early-to-call-it-a-trend.html.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Statistics Canada. 2016. "Focus on Geography Series, 2016 Census." https://www12.statcan.gc.ca/census-recensement/2016/as-sa/fogs-spg/Facts-cma-eng.cfm?LANG=Eng&GK=CMA&GC=535&TOPIC=7.

Toronto Police Service. 2022. "PUBLIC Safety Data Portal: OPEN Data Documentation." https://torontops.maps.arcgis.com/sharing/rest/content/items/c0b17f1888544078bf650f3b8b04d35d/data.

Toronto Police Services. 2021. *Police Annual Statistical Report - Homicides*. https://open.toronto.ca/dataset/police-annual-statistical-report-homicide/.

UNODC. 2019. "Global Study on Homicide." https://www.unodc.org/documents/data-and-analysis/gsh/Booklet_6new.pdf.

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

Xie, Yihui. 2021. *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*. https://CRAN.R-project.org/package=kableExtra.