Exploring Homicide Trends in Toronto: A Seventeen-Year Statistical Overview (2004-2020)

Alexander Guarasci

January 24, 2024

This study investigates homicide trends in Toronto from 2004 to 2020, utilizing data from OpenDataToronto. The data shows a fluctuating increase in homicide rates over the city, with most homicides occurring in a relatively small subset of neighbourhoods. Additionally, the data shows the prevalence of gun violence, a surprising result in a country known for its strict gun control laws. This research is significant as it provides insights into urban violence in a major Canadian city. It contributes to the understanding of urban crime dynamics in the Greater Toronto Area.

Introduction

Urban violence and its impact on communities is a pressing concern in cities worldwide, and Toronto, Canada's most populous and diverse city, is no exception. The study of homicide trends within this urban setting is not only critical for ensuring public safety but also for understanding the underlying social and economic factors that drive such violence. This paper presents an analysis of the patterns and distribution of homicides in the Greater Toronto Area (GTA) from 2004 to 2020. The research is driven by a need to comprehensively understand the dynamics of these incidents, their geographic spread across the city, the predominant methods of homicide, and the broader implications for urban policy and community wellbeing.

In addressing this need, the study utilizes a robust dataset from OpenDataToronto (Gelfand 2022), specifically the "Police Annual Statistical Report - Homicides," published by the Toronto Police Services. This dataset provides a detailed account of all recorded homicides under the jurisdiction of the Toronto Police Services over a seventeen-year period. Our research offers an examination of both the spatial distribution of homicides across Toronto's neighbourhoods and the methods used in these violent acts. The analysis, conducted through the R programming language (R Core Team 2022) and supplemented with various analytical tools like Tidyverse

(Wickham et al. 2019), Janitor (Firke 2023), and Knitr (Xie 2014), delves deep into the trends, revealing patterns that could have implications for policy formulation and implementation.

The findings of this research are revealing. We observed a marginal yet noteworthy increase in overall homicide rates over the study period, a trend that calls for a critical evaluation of current approaches to managing urban violence. More strikingly, the analysis unveiled a disproportionate concentration of homicides in specific neighborhoods, suggesting a need for targeted intervention strategies. Furthermore, the predominance of shootings as the homicide method challenges existing perceptions about firearm accessibility and regulation in Canada, thus calling for a reevaluation of national gun control policies. These findings not only inform effective policy and law enforcement strategies but also contribute to the broader discourse on urban violence in multicultural settings.

The structure of the paper is as follows: Section 2 details the data processing techniques employed in the study, afterwards is a brief discussion of the data set, followed by insight into the variables of interest. Section 3 is devoted to presenting and discussing the key findings, including yearly homicide numbers and trends, the geographical distribution of these incidents, and the analysis of the prevalent methods of homicide. The concluding section synthesizes these insights, drawing out their implications for urban policy, community safety strategies, and the broader societal understanding of urban violence in Toronto. This study not only sheds light on the current state of urban homicides in Toronto but also sets the stage for future research and policy development in this critical area.

Code and data supporting this analysis is available.[1]

2 Data

The data used in this paper was gathered from OpenDataToronto (Gelfand 2022), a city-funded source with the explicit goal of making the city of Toronto more "transparent, accountable, participatory and accessible" ("Open Data" 1998). The dataset used in this analysis was called "Police Annual Statistical Report - Homicides", and was published by Toronto Police Services. In order to perform the analysis, the programming language for data analysis, visualization and statistical investigation R was used (R Core Team 2022), along with the R packages Tidyverse (Wickham et al. 2019), Janitor (Firke 2023) and Knitr (Xie 2014).

2.1 The Dataset

The dataset utilized in this paper is the aforementioned "Police Annual Statistical Report - Homicides", which includes data from every homicide under the Toronto Police Services jurisdiction between 2004 and 2020. The data set itself contains 1166 samples each with

---

[1]https://github.com/AlexanderG123/toronto_homicide

10 variables. Because Toronto Police Services is the only institution qualified to deal with homicides, no other datasets were considered. One potentially significant concern with this data, is the possibility that it is incomplete due to undetected homicides. However, I do not believe that there are sufficient undetected homicides to greatly impact the ensuing analysis. Another problem with the data lies in how it handles the location. The Toronto Police Services have deliberately made the location of each occurrence at the nearest intersection, and therefore the data is not flawless by the admission of the data procurers. Finally, the relatively small number of homicides per year in Toronto renders the data somewhat difficult to analyze. When trying to examine trends within certain communities and regions, the deviation was too high to extrapolate anything of interest.

Making the data manageable was simply a matter of removing the useless variables (I am sure the police department would argue that homicide_id is relevant for their work but it is useless for the purposes of our analysis) and summarizing certain elements (ie. explicitly stating the year each homicide happened in order to perform analysis on a year by year basis).

2.2 Variables of Interest

The data includes the date of the event, the type of homicide, as well as the neighbourhood in which the homicide took place. The date is unambiguous and needs no further explanation. The type of homicide is broken into three clear sub-categories which are "shooting", "stabbing" and "other". For neighbourhood, Toronto is broken up into 158 "social planning neighbourhoods" ("About Toronto Neighbourhoods" 2022). These are each numbered and reported independently for the purposes of this paper, (this is also how Statistics Canada breaks Toronto into neighbourhoods for all of their relevant data sets).

With regard to measurement, the precise location is inaccurate, but the important aspects of the data set are unambiguous in terms of their ability to be measured. A homicide is a countable event that does not need to be measured, the categorization of homicide is the same. Only the date may have measurement errors but these would be unimpactful for the following analysis.

3 Analysis and Data Visualization

Our analysis delves into the details of homicide trends in the GTA from 2004 to 2020, revealing insights about the nature and distribution of these violent incidents.
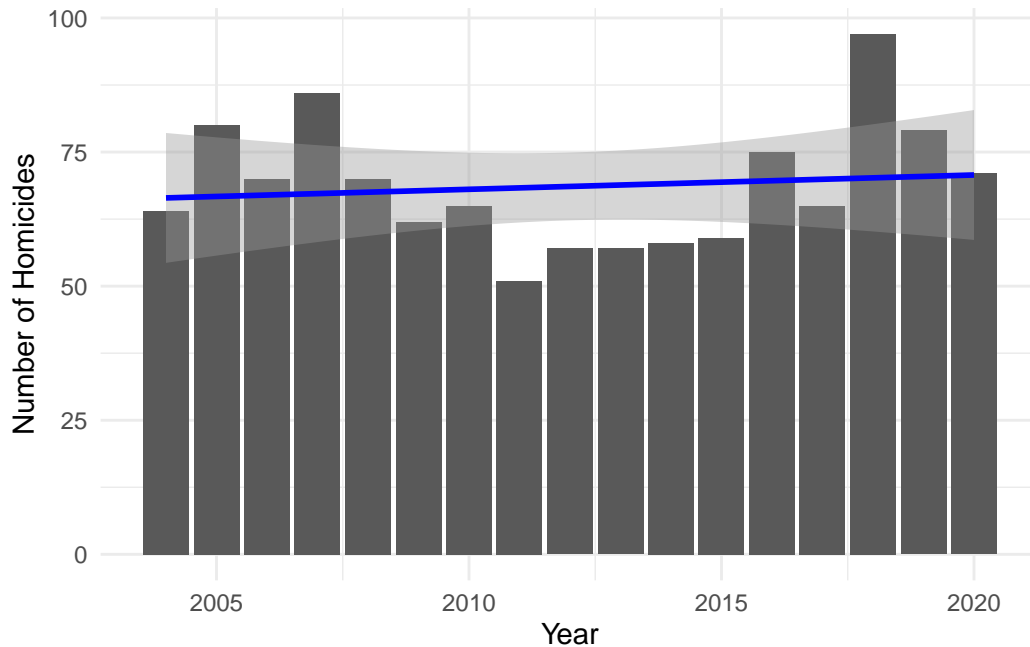
3.1 Yearly Homicide Trends

Figure 1: Toronto homicides per year 2004 - 2020

Looking at the yearly trends in homicides in Toronto, Figure 1, which combines a bar chart with a linear regression line[2], shows a discernible increase in homicides over the 17 years. Despite some volatility in yearly figures, a worrying upward trend is apparent in the positively-sloped linear regression line. To quantify this rise, the average number of homicides in the first half of the period was 67 per year, which increased to 69 in the second half. This nearly 3% rise in homicides, although seemingly small, is significant given its societal implications. However, it's important to note that these figures do not account for other factors like population growth.

---

[2]A linear regression is a statistical model that attempts to show a linear relationship between two variables, and thusly be able to predict how a change in one of the variables will impact a change in the other. In this case it is showing that over the course of the study, homicides are increasing

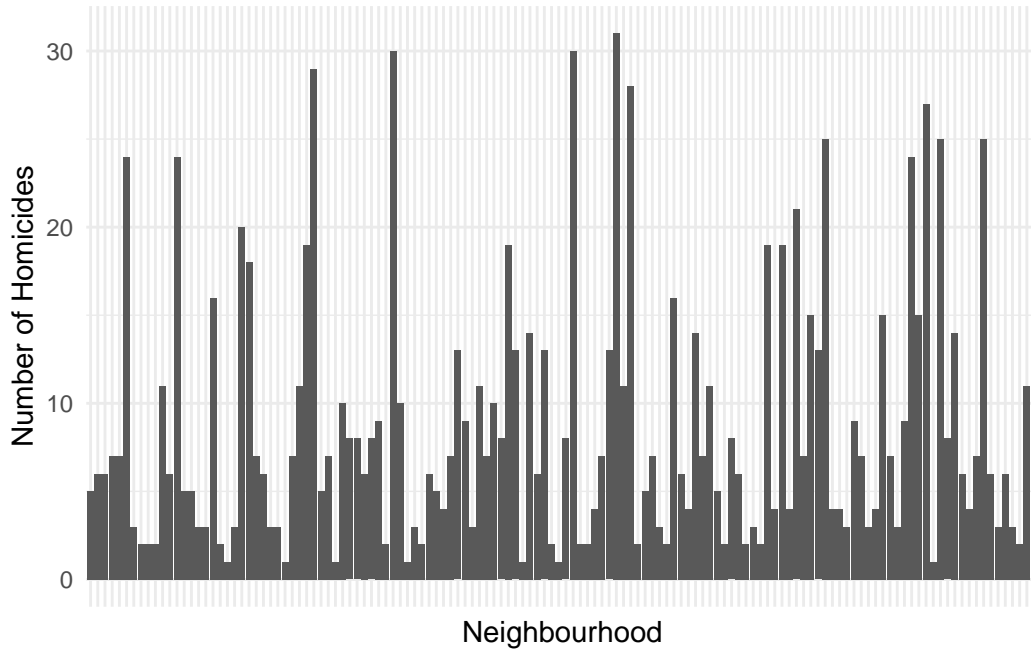## 3.2 Geographical Distribution of Homicides

Figure 2: Toronto distribution of homicides per neighbourhood 2004 - 2020

Our analysis further explores the uneven distribution of homicides across Toronto's neighbourhoods. As illustrated in Figure 2, there were 1166 homicides recorded between 2004 and 2020, with a pronounced concentration in certain areas. For example, just 24 out of 158 neighbourhoods, which is merely 15.2% of the total, witnessed 15 or more homicides, accounting for 45.6% of the overall figure. Additionally, 14 neighbourhoods experienced over 20 homicides, making up 31.1% of the aggregate, while a mere 9 neighbourhoods, each with more than 25 homicides, constituted 21.4% of the total. This lopsided distribution highlights a pattern where most neighborhoods encounter few or no homicides, while a select few bear a disproportionate share. Such insights are crucial for devising policing strategies and allocating resources. They underscore the potential effectiveness of a targeted approach but also bring to light the challenges of potential over-policing and its effects on communities already burdened by high crime rates. Consideration of alternatives, such as increased educational and social support spending, may offer pathways to reduce homicide rates in these neighbourhoods. A deeper analysis with a more comprehensive data set is essential to inform any policy decisions.

3.3 Predominance of Shooting as a Homicide Method

Perhaps the most striking finding of our study is the predominance of shooting homicides.
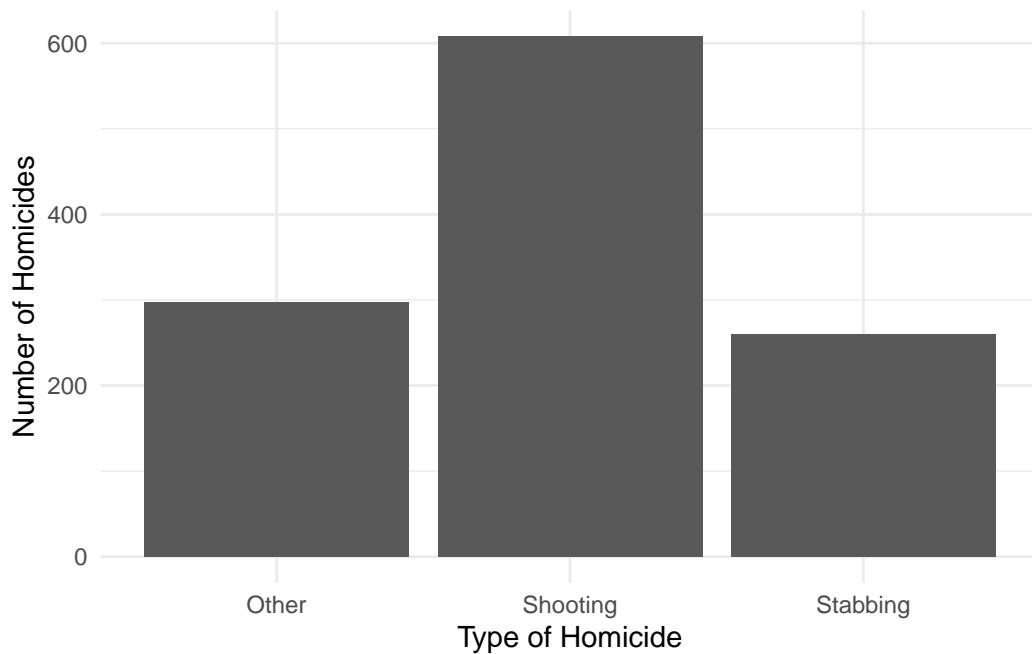

Figure 3: Toronto types of homicides breakdown 2004 - 2020

Figure 3 offers a revealing glimpse into the distribution of homicides over the span of 17 years, clearly demonstrating that shootings were substantially more prevalent than stabbings. However, the data becomes even more informative when dissected on an annual basis.
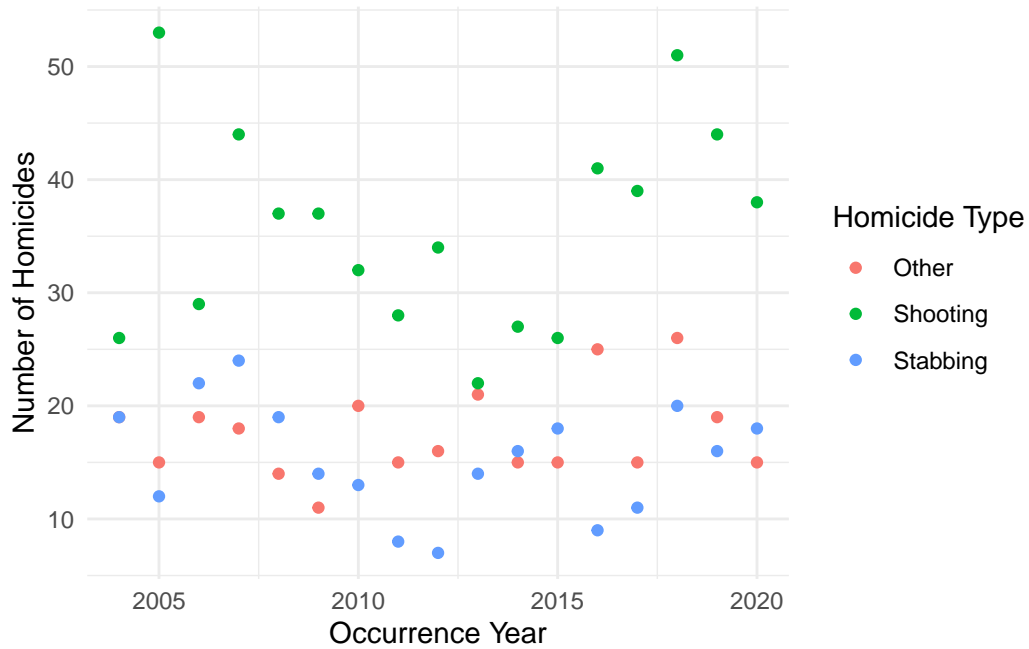
Figure 4: Types of homicides per year

Examining Figure 4, which depicts homicide trends over the 17-year study period reveals a consistent pattern: shootings were the most common type of homicide every year. There was not a single year when shootings were outnumbered by stabbings or other forms of homicide. This observation raises critical questions about Canada's gun control laws. It challenges the widely held perception of limited access to guns among Canadians, prompting a deeper exploration of whether this is a true representation of the situation or if it is more a reflection of population size or other variables. The data suggests a reevaluation of current policies and societal attitudes toward gun ownership and accessibility in Canada.
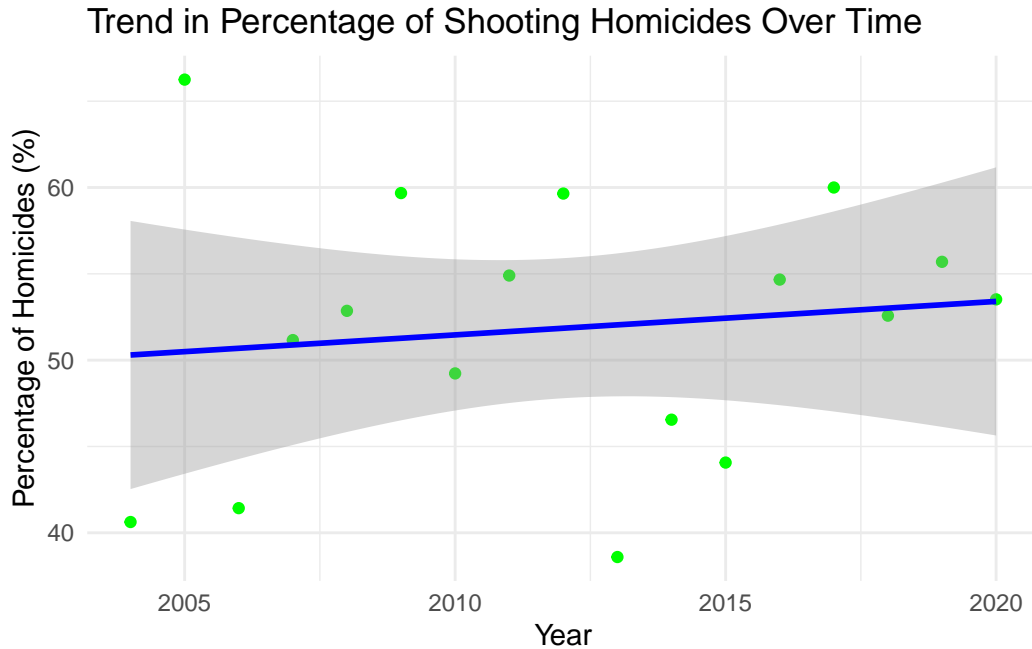
Figure 5: Toronto shooting percentage of total homicides by year 2004 - 2020

Additionally, an analysis of Figure 5 reveals a concerning upward trend in the total percentage of homicides by gunshot. This escalation not only accentuates the growing problem of gun violence in Toronto but also signals a potential shift in the nature of criminal activity and societal issues. The increasing reliance on firearms as a means for committing homicide points towards a deeper issue, possibly linked to the accessibility of firearms, changes in criminal behaviour, or broader socio-economic factors. The rising trend underscores the urgency for a thorough examination of gun control measures, policing strategies, and community interventions aimed at addressing the root causes of this surge in gun-related violence. It also invites a broader discussion about the societal factors contributing to this alarming trend, calling for a multi-faceted approach to tackling the complexities of gun violence in modern urban settings.

4 Conclusion

Our in-depth analysis of Toronto's homicide trends from 2004 to 2020 reveals critical insights into the nature and distribution of urban violence in Canada's largest city. The data unearthed in this study paints a complex picture, indicating a marginal but noticeable increase in overall homicide rates over the years. This uptick, while seemingly slight, necessitates a reevaluation

of existing strategies to combat urban violence. The study highlights the importance of contextual factors, such as population growth and urbanization, in understanding these trends. Moreover, the disproportionately high concentration of homicides in a small number of neighbourhoods demands a targeted approach. This pattern underscores the need for strategic resource allocation and policy planning, aiming to address the unique challenges in these high-risk areas. Balancing enhanced policing with increased investment in social and educational programs could mitigate the root causes of violence, avoiding the pitfalls of over-policing and its potential negative impact on community relations.

Equally significant is the predominance of shooting as the method of homicide, consistently observed throughout the study period. This trend challenges the existing perceptions about gun accessibility and usage in Canada, calling for a critical review of the nation's gun control policies. The increasing reliance on firearms in homicides suggests a shift in criminal behaviour and possibly points to broader socio-economic issues. In conclusion, the complexities uncovered in our study advocate for a multifaceted approach to addressing urban violence. Effective law enforcement combined with proactive social policies is crucial, not only for immediate crime reduction but also for addressing the underlying socio-economic factors that contribute to such environments. As Toronto continues to evolve as a multicultural urban setting, it is imperative that future policies and interventions are informed by comprehensive data analysis and a deep understanding of the city's unique social dynamics.

Appendix

Data Cleaning

As mentioned briefly in 2.1, the cleaning of the data set was straight forward. I started by using the wday function (Wickham et al. 2019) to create a column in the data set that has the day of the week each event occurred, this was not ultimately helpful in my final analysis but it was interesting see the breakdown of homicide by weekday in my data exploration. After this, I used the select function (Wickham et al. 2019) to remove all unnecessary columns so that only the interesting variables would be in my cleaned data set. I proceeded to use clean_names function (Firke 2023) to make the column names more appropriate, and finally I used as.Date (R Core Team 2022) to get rid of the time stamps shown in the date column.

References

"About Toronto Neighbourhoods." 2022. https://www.toronto.ca/city-government/data-research-maps/neighbourhoods-communities/neighbourhood-profiles/about-toronto-neighbourhoods/.

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal.* https://CRAN.R-project.org/package=opendatatoronto.

"Open Data." 1998. https://www.toronto.ca/city-government/data-research-maps/open-data/.

R Core Team. 2022. *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 Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.