Rising Gun Violence in Toronto*

How Shootings Have Increased and Shifted Across Neighborhoods

Xingjie Yao

September 24, 2024

In this paper, we analyze shooting occurrences in Toronto between 2014 and 2019, focusing on trends and patterns across different geographical divisions. Through data visualization and analysis, we found that shooting incidents significantly increased over time, with specific divisions experiencing higher frequencies and greater variability. These findings highlight the growing prevalence of gun violence in Toronto and the unequal distribution of shootings across the city. Understanding these trends is crucial for informing law enforcement strategies and policy decisions aimed at reducing gun violence and improving community safety.

1 Introduction

In recent years, rising crime rates, particularly those involving shootings, have become a significant public concern. Shootings significantly impact community safety and public health, showing the importance of understanding their patterns and trends for effective policymaking and law enforcement. In this paper, we focus on analyzing shooting occurrences in Toronto, aiming to identify trends, patterns, and potential insights through the study of historical data. The dataset used in this analysis comes from the Toronto Open Data Portal, which provides a detailed account of shooting occurrences across various geographical divisions from 2014 to 2019. In this paper, we use data analysis tools in R Core Team (2023) such as Wickham et al. (2019), Wickham (2016), and Wickham et al. (2023) to explore both the temporal and geographical dimensions of shootings in the city. We use graphical visualizations (such as side-by-side boxplots and bar graphs) and numerical summaries to analyze key trends. The goal of this explanatory data analysis is to understand how the frequency and distribution of shooting occurrences have evolved across different divisions over the years, and our findings illustrate notable patterns in the distribution of shootings by year and geographical area in Toronto. These insights are valuable not only for the government but also for understanding

*Code and data are available at: https://github.com/Stella41603/Toronto-Shooting-Occurrences.git crime dynamics. By identifying high-risk areas and periods, resources can be better allocated to reduce gun violence.

The remainder of this paper is structured as follows: Section 2 provides a detailed description of the dataset and the exploratory data analysis conducted; Section 3 presents the conclusions drawn from the analysis, addressing key points discussed in Section 3.1 and Section 3.2; Section 3.3 discusses some weaknesses and potential future work. Through this structured approach, we aim to offer a comprehensive analysis of shooting occurrences in Toronto and contribute to crime prevention and community safety.

2 Data

The dataset used in this analysis is sourced from OpenDataToronto (Gelfand (2022)) and is titled "Shooting Occurrences". This dataset was selected as it specifically contains the detailed information required for our analysis, such as the number of shooting occurrences and the geographical divisions in which they took place. While there are two other similar datasets, "Shootings & Firearm Discharges" and "Neighbourhood Crime Rates", they were not suitable for our purposes. The "Shootings & Firearm Discharges" dataset lacks a variable indicating the total number of occurrences, and the "Neighbourhood Crime Rates" dataset does not contain geographical division information, both of which are crucial for our analysis. The chosen dataset consists of 96 observations and 6 variables:

- `_id`: Unique row identifier for Open Data database.
- **Index_**: Unique identifier.
- **OccurredYear**: Year shooting occurred.
- **GeoDivision**: Geographic division crime took place.
- **Category**: Shooting category.
- **Count_**: Total number of shooting occurrences.

For the purpose of data cleaning, we identified that both `_id` and Index_ served as unique identifiers, so we removed `_id` and kept Index_, renaming it to id for simplicity. Meanwhile, the Category variable was dropped since all observations had the same value ("Shooting Occurrence"). The three remaining variables (OccurredYear, GeoDivision, and Count_) are the primary focus of this analysis and are renamed as year, division, and count. After reviewing the data, we confirmed that there were no missing observations. Therefore, the cleaned dataset consists of 96 observations and 4 variables:

- **id**: Unique identifier for each observation.
- **year**: Discrete Numerical (ranging from 2014 to 2019); Year the shooting occurred.
- **division**: Nominal Categorical (codes such as "D11," "D12," etc); Geographic division where the shooting occurred.
- **count**: Discrete Numerical; Total number of shooting occurrences within the specified geographical division and year.

Based on the cleaned dataset, we then generated summary statistics and visualized the relationships between the variables.

Table 1: Summary Statistics for Division Shooting Occurrences by Year (2014-2019)

Year	minimum	Q1	median	Q3	maximum	mean	std deviation
2014	2	4.25	9.0	16.50	31	11.0625	8.33042
2015	3	10.5	14.5	21.75	50	18.0000	12.8686
2016	5	15.75	21.5	30.75	65	25.4375	16.0705
2017	6	13.25	23.5	30.25	68	24.5000	15.5435
2018	9	14.00	26.5	30.50	67	26.6875	15.1996
2019	9	18.25	25.0	40.25	64	30.7500	16.6032

The summary statistics in Table 1 show a notable upward trend in division shooting occurrences in Toronto from 2014 to 2019. The mean number of shooting occurrences increased from approximately 11.06 in 2014 to 30.75 in 2019, suggesting a growing problem of gun violence over the six-year period. The median value follows a similar pattern, rising from 9 in 2014 to 25 in 2019. This growth is also evidenced by the increasing spread of data that the standard deviation has more than doubled from 8.33 in 2014 to 16.6 in 2019, and the interquartile range and range of occurrences has widened over time. This indicates that the variability in shooting occurrences became more pronounced as time progressed. In terms of quartiles, the Q1, median, and Q3 values all increase steadily over the years, showing the fact that division shooting occurrences are generally rising.

Similar observations can be made from the side by side boxplots Figure 1 which shows the distributions of division shooting occurrences by year. The boxplots moved up across the years and the middle box has become wider and wider which shows that the frequency of shootings increases over time.

Table 2: Summary Statistics for Annual Shooting Occurrences by GeoDivision

Division	minimum	Q1	median	Q3	maximum	mean	std deviation
D11	2	8.25	13.5	15.75	18	11.6667	6.08824
D12	20	27.75	36.0	36.00	47	33.3333	9.54289
D13	2	6.50	15.0	16.75	27	13.3333	9.20145
D14	7	15.25	19.5	23.00	27	18.5000	7.25948
D22	8	14.25	15.0	16.50	23	15.3333	4.84424
D23	18	38.00	46.5	52.75	61	43.6667	15.1877
D31	31	53.50	64.5	66.50	68	57.5000	14.5430
D32	11	11.50	18.0	26.00	37	20.3333	10.5578
D33	5	11.00	11.0	11.75	22	12.0000	5.51362
D41	15	21.50	23.5	26.25	30	23.3333	5.16398

Division	minimum	Q1	median	Q3	maximum	mean	std deviation
D42	10	23.00	30.5	34.25	35	27.0000	9.81835
D43	7	17.50	26.0	30.00	40	24.1667	11.7033
D51	14	16.75	19.5	23.00	30	20.5000	5.78792
D52	2	6.00	9.0	9.75	14	8.16667	4.16733
D53	2	3.75	7.5	9.00	13	7.00000	4.14729
D54/D55	14	20.75	28.5	35.50	41	28.0000	10.3923

Table 2 shows the distribution of annual shooting occurrences by geographic division, providing insight into how shootings are concentrated in different areas of Toronto. The summary statistics demonstrate significant variation across divisions. Division D31, with a mean of 57.5 occurrences and a maximum of 68, stands out as having the highest average number of annual shootings, while Division D52, with a mean of 8.17, experiences far fewer annual shootings on average. The standard deviation for Division D31 is relatively high at 14.54, indicating significant variability in shooting occurrences across years. Similarly, Division D23 (mean 43.67, standard deviation 15.19) and Division D54/D55 (mean 28, standard deviation 10.39) also show notable fluctuations in shooting occurrences, suggesting these areas may experience more accidental gun violence compared to other divisions. On the other hand, some divisions (like D52 and D53) show relatively low means and standard deviations, implying fewer and more consistent shooting occurrences over the years. For example, Division D53 has a mean of 7 and a standard deviation of just 4.15, indicating less volatility in gun violence in this area. The interquartile range for most divisions is moderate, but divisions D23 and D31 have larger spreads, showing the observation that shooting occurrences in these areas are more frequent and variable.

The bar graph (Figure 2) illustrates the annual shooting occurrences across various divisions in Toronto from 2014 to 2019, with each year represented by a different color as shown in the legend. Each subplot corresponds to a specific division, displaying the changes in shooting occurrences over time. Divisions D31 and D23 show significantly higher shooting occurrences compared to others, with tall bars, particularly in the later years (2017–2019) which suggests these areas experienced more frequent gun violence. Divisions D52 and D53 show consistently low levels of shooting occurrences, as indicated by the short bars across all years which suggests these divisions face fewer issues with gun violence. Some divisions display a clear upward trend, with occurrences increasing gradually each year; while other divisions show no clear pattern or trend, showing variability from year to year without consistent growth or decline. The colors representing 2018 and 2019 (blue and pink) are outstanding in many divisions, indicating that the later years saw a rise in shooting occurrences across multiple areas. On the other hand, 2014 (red) shows the smallest bars in almost every division, suggesting a lower number for shooting incidents at the beginning of the period.

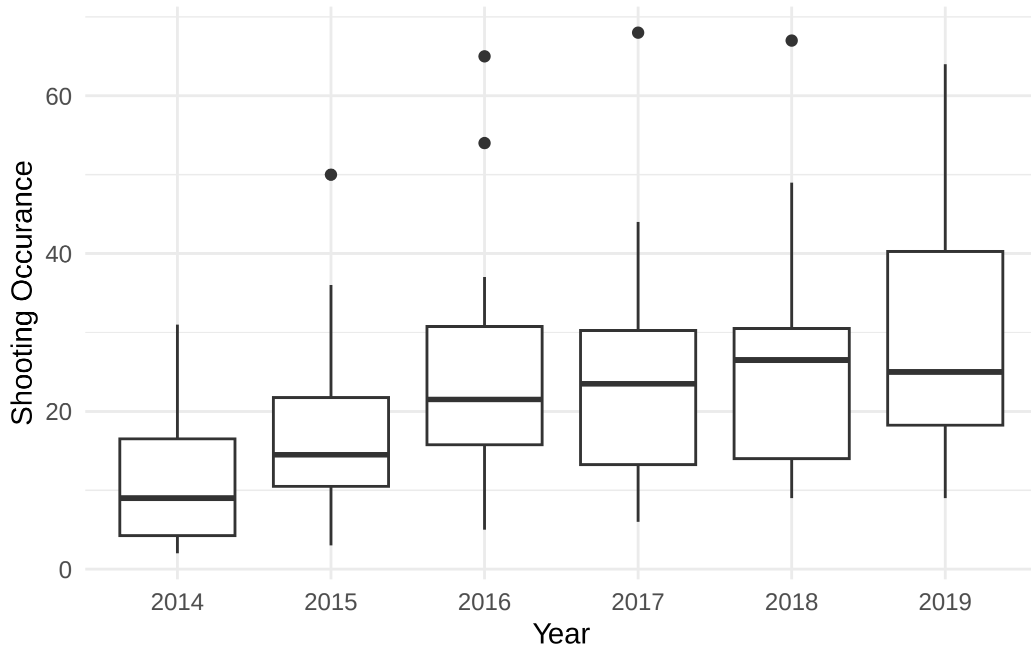

Figure 1: Distribution of GeoDivision Shooting Occurrences by Year (2014-2019)

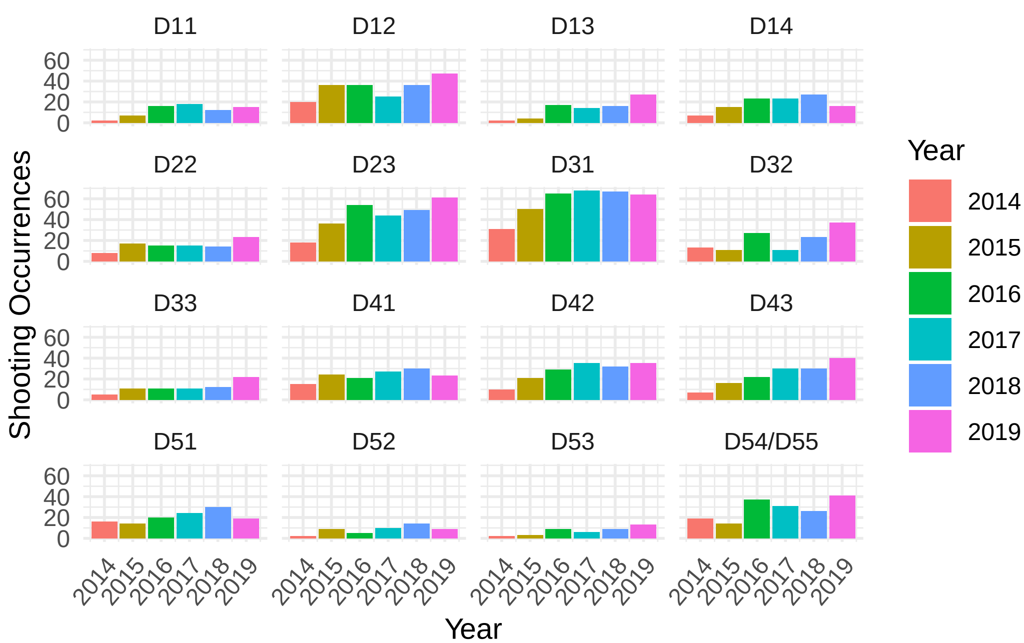

Figure 2: Shooting Occurrences by Year (2014-2019) Across Toronto GeoDivisions

3 Discussion

3.1 First Discussion Point - Temporal Trends

The overall upward trend in shootings from 2014 to 2019 suggests a worsening issue with gun violence in Toronto. The increasing variability across years (shown by growing standard deviations and wider ranges) indicates a disparity in the frequency of shootings with more extreme values occurring in recent years. This highlights the need for a time-based analysis of the factors contributing to the rise in gun violence.

3.2 Second Discussion Point - Geographic Hotspots

The varying trends across divisions suggest that the problem of gun violence is not uniform throughout the city. Some divisions, like D31 and D23, have seen a steady rise in violence, whereas others, such as D52 and D53, show more stable or fluctuating patterns with lower shooting occurrences. This spatial variation implies that localized efforts targeting specific divisions might be more effective in reducing gun violence. Meanwhile, tailored interventions and resource allocation should focus on high-risk areas to curb the rising trend in certain divisions.

3.3 Weaknesses and Next Steps

One key weakness of the analysis is the lack of data beyond 2019, limiting our understanding of recent trends in shootings. Future work could extend the time range to include more recent years, as this could provide insights into whether the upward trend has continued, stabilized, or reversed. Expanding the dataset would improve the accuracy of forecasts and inform more effective strategies for violence prevention.

References

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

R Core Team. 2023. *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 François, Lionel Henry, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.