

# The increasing trend of shooting occurrences in Toronto\*

Varun Vijay

07 February 2022

## Abstract

Gun violence is a very serious problem in the modern world so statistics about shooting occurrences are an important factor to consider. This paper analyzed yearly data about shooting occurrences that were reported in Toronto in different geographic divisions of the city from 2014 to 2019. The analysis showed that there was a mostly increasing trend in shooting occurrences in Toronto from 2014 to 2019 and a close correlation between current year and the number of shooting occurrences in the year. The findings in this paper have serious implications for the crime rate and the effectiveness of gun control in Toronto.

## Contents

<b>Introduction</b>	<b>1</b>
<b>Data</b>	<b>2</b>
Data source . . . . .	2
Source of data and Biases . . . . .	2
Characteristics of dataset . . . . .	2
Bar plots . . . . .	3
Summary statistics about the dataset . . . . .	3
Scatterplot about the dataset . . . . .	6
Packages . . . . .	6
<b>References</b>	<b>6</b>

## Introduction

The problem of gun violence is a very serious issue in the modern world. Guns are capable of dealing large losses to life in very short periods of time. In order to deal with this problem, proper gun control laws are require and must be implemented in an effective manner to reduce the number of shooting occurrences in Toronto. Statistics and data about shooting occurrences can be used to measure how effective the current gun control laws are in reducing the amount of shooting occurrences. The data could also be used as an indicator of police effectiveness and the crime rate.

---

\*Code and data are available at: [https://github.com/Varun1005473462/starter\\_folder-main](https://github.com/Varun1005473462/starter_folder-main).

In this paper, the dataset was analyzed using bar plots, scatterplots and various summary statistics about the dataset such as the correlation between year and the number of shooting occurrences in said year. This analysis was performed in R, an open-source language for statistical programming. A mostly increasing trend in shooting occurrences was observed from 2014 to 2019 with a slight decrease in 2017. However, the data shows that the number of reported shooting occurrences in Toronto have been steadily increasing from 2014 to 2019. Furthermore, it appears that in each year the largest number of shooting occurrences have occurred in the geographical division D31 closely followed by D23 which might suggest that these areas have a higher rate of gun violence than the other geographical divisions.

The results shown in the analysis dataset could imply that the effectiveness of gun control laws in Toronto has gone down since 2014 and that the crime rate has gone up, particularly in geographical divisions D31 and D23. It might mean that the Toronto government needs to seriously reexamine its laws in regards to guns and that policing must be more effective in the affected geographical divisions. The rest of this paper will explain the steps taken in analysing the dataset and the findings obtained in more detail.

## Data

### Data source

The dataset used for this paper was the Shooting Occurrences dataset provided by the Toronto's Police Service Annual Statistical Report (ASR), a detailed annual overview that covers police related statistics including reported crimes, victims of crime, missing persons, firearms, traffic collisions, personnel and other administrative information. The data was publicly available by using the City of Toronto open data portal, [opendatatoronto](https://opendatatoronto.com/en/datasets/shooting-occurrences/), allowing me to access it for this paper. I obtained the required dataset in csv format using the R package, [opendatatoronto](https://github.com/gelfand/opendatatoronto) (Gelfand 2020). The dataset was last refresed on November 18th, 2020.

### Source of data and Biases

The dataset includes all shooting related occurrences that were reported to the Toronto police service from 2014 to 2019 by occurred data divided by division. However, this dataset also includes shooting occurrences that were deemed to be false after investigation. These incidents could have been reported to the Toronto police service such as calls to emergency services such as 911. The ASR did not specify where the calls came from and and from whom. No personal information was included in the dataset due to the Municipal Freedom of Information and Protection of Privacy Act.

While this dataset contains all reported shooting occurrences in Toronto, it does not specify which of these reported occurrences turned out to be false after investigation. So the number of actual shooting occurrences is less than the number given in the dataset. This could mean that the observations made about the dataset are biased and exaggerated due to the actual number of shooting occurrences not being reported. Additionally, this dataset only specifies the number of shooting occurrences that were reported to the Toronto Police service, so the actual number of shooting occurrences that occurred from 2014 to 2019 could be underreported in this dataset. Both of these factors can lead to biases in the paper's findings.

### Characteristics of dataset

The shooting occurrences dataset contains all shooting occurrences reported to the Toronto Police Service from 2014 to 2019 aggregated by division. There are 96 observations and 6 variables. These variables are `_id`, `index`, `OccurredYear`, `GeoDivision` (*Geographic division crime took place*), `Category` (*Shooting category*) and `Count` (Total number of shooting occurrences reported). The first two `_id` and `_index` are the same and were removed prior analysis. Category was also removed because all of the occurrences were the same category, Shooting occurrence, and thus it had no impact on the analysis of the dataset. Shooting occurrences in each

year were aggregated by the geographical division they took place in. A sample view of the dataset can be viewed [here](#).

Extract of dataset is shown in table 1 (Table @ref(tab:dataextract))

Table 1: Table 1: First ten rows of a dataset that shows shooting occurrences in Toronto

Occured Year	Geo Division	Count
2014	D11	2
2014	D12	20
2014	D13	2
2014	D14	7
2014	D22	8
2014	D23	18
2014	D31	31
2014	D32	13
2014	D33	5
2014	D41	15

As you can see in Table 1 (Table @ref(tab:dataextract)) above, the dataset has three variables, the year the shooting was reported, the geographical division of the city the shooting took place in and the number of times it happened.

## Bar plots

The relationship between the current year and the number of shooting occurrences that took place in said year is a useful way to show the data as it can be interpreted as a measure of how effective gun control policies are in Toronto from year to year.

Figure 1 (Figure @ref(fig:bar plot 1)) shows the relationship between year and shooting occurrences.

From figure 1 (Figure @ref(fig:bar plot 1)), there appears to be a mostly increasing trend in shooting occurrences over the years. It decreases from 2016 to 2017 but then increases again from 2017 to 2018 and maintains the trend till 2019. This increasing trend could be interpreted as a rise in gun violence over the course of 5 years in Toronto and the relaxation or decreased effectiveness of gun control laws.

Showing how shooting occurrences in each year were aggregated by the geographical division it took place in is also a useful thing to show as it shows us which division had the highest amount of reported shooting occurrences and could be interpreted as that division having the highest crime rate.

Figure 2 (Figure @ref(fig:bar plot 2)) shows shooting occurrences for each year aggregated by geographical division.

From figure 2 (Figure @ref(fig:bar plot 2)), it can be observed that division D31 had the highest reported amount of shooting occurrences in each year from 2014 to 2019 closely followed by division D23 in each year. So, it could be interpreted as these divisions having the highest crime rate and gun violence in Toronto out of all of the divisions.

## Summary statistics about the dataset

Table 2 (Table @ref(tab:datayear)) shows the maximum number of shooting occurrences that occurred for each year.

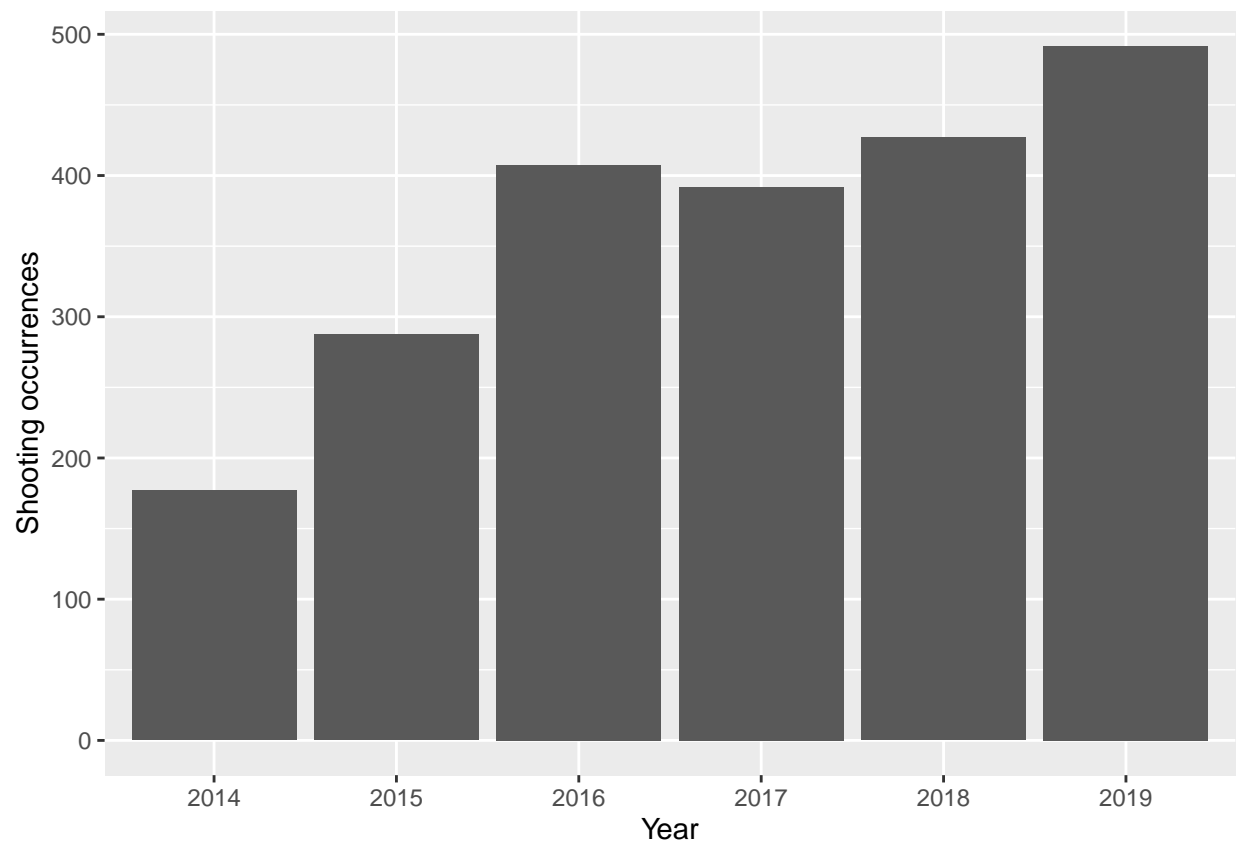


Figure 1: Figure 1: Number of shooting occurrences from 2014 to 2019

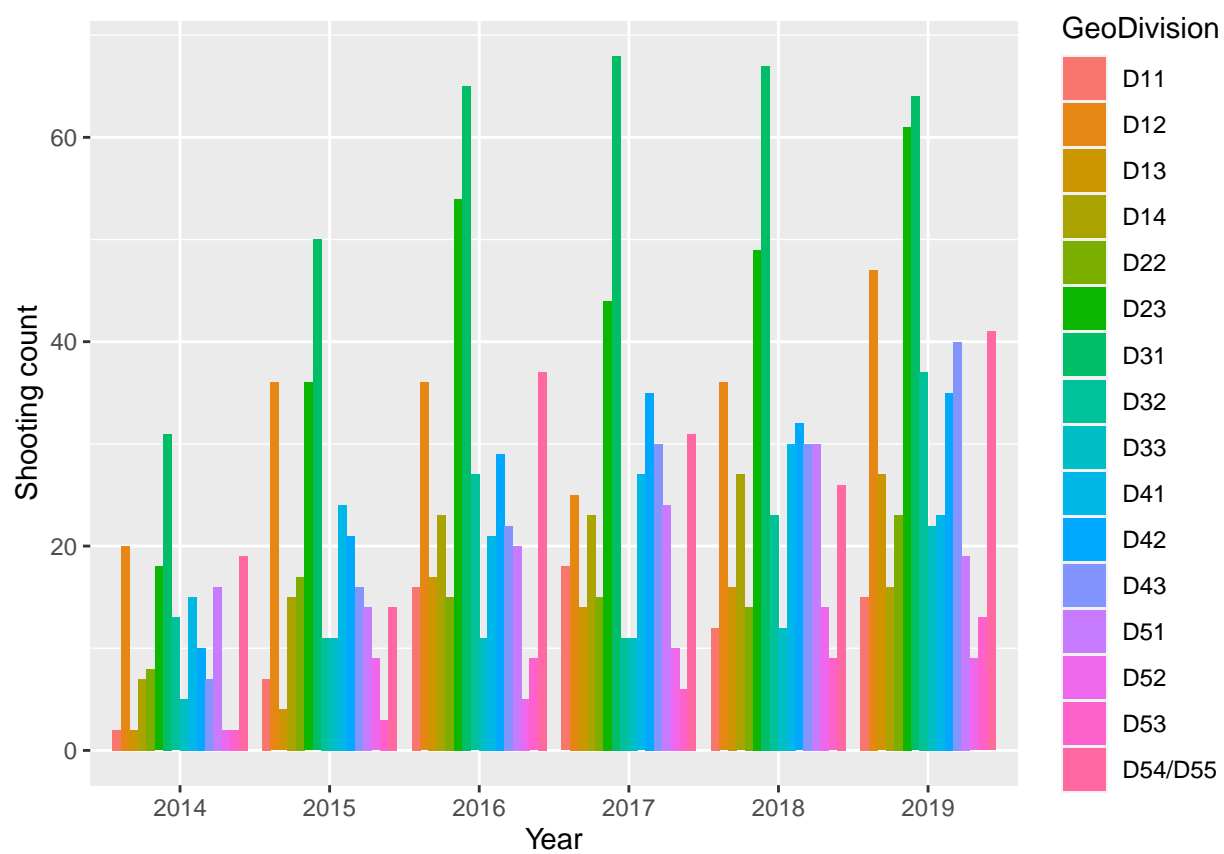


Figure 2: Figure 2: Number of shooting occurrences from 2014 to 2019 aggregated by geographical division

Table 3: Table 3: correlation between year and shooting occurrences reported in that year

	Year	Shooting_count
Year	1	.
Shooting_count	0.94	1

Table 2: Table 2: Dataset showing shooting count for each year from 2014 to 2019

Year	Shooting_count
2014	177
2015	288
2016	407
2017	392
2018	427
2019	492

From the table 2 (Table @ref(tab:datayear)), it can be observed that shooting occurrences have been increasing from 2014 to 2019 with a slight decrease in 2017. From 2017, there have only been slight increases in reported shooting occurrences but there has been no decrease in shooting occurrences over the years shown in the dataset. This table reinforces the earlier theory that gun laws have become less effective since 2014 in Toronto and that more serious measures should have been implemented in these years.

Table 3 (Table @ref(tab:datacorrelation) )shows the correlation between year and shooting occurrences reported in that year.

Table 3 (Table @ref(tab:datacorrelation)) shows that there is a very close correlation between year and reported shooting\_occurrences in that year.

## Scatterplot about the dataset

Figure 3 (Figure @ref(fig:scatterplot)) shows us the relationship between year and shooting occurrences in a scatterplot.

Figure 3 (Figure @ref(fig:scatterplot)) shows that there is something closely resembling a linear relationship between year and shooting occurrences in that year.

## Packages

The packages used for this paper were tidyverse, modelsummary, knitr, opendatatoronto, dplyr, and readr. Some of these packages such as opendatatoronto and dplyr were used to upload the dataset used for this project. Others were used for the various plots and tables included in this paper.

## References

- R Core Team (2020)
- (Gelfand 2020)
- (Wickham et al. 2019)

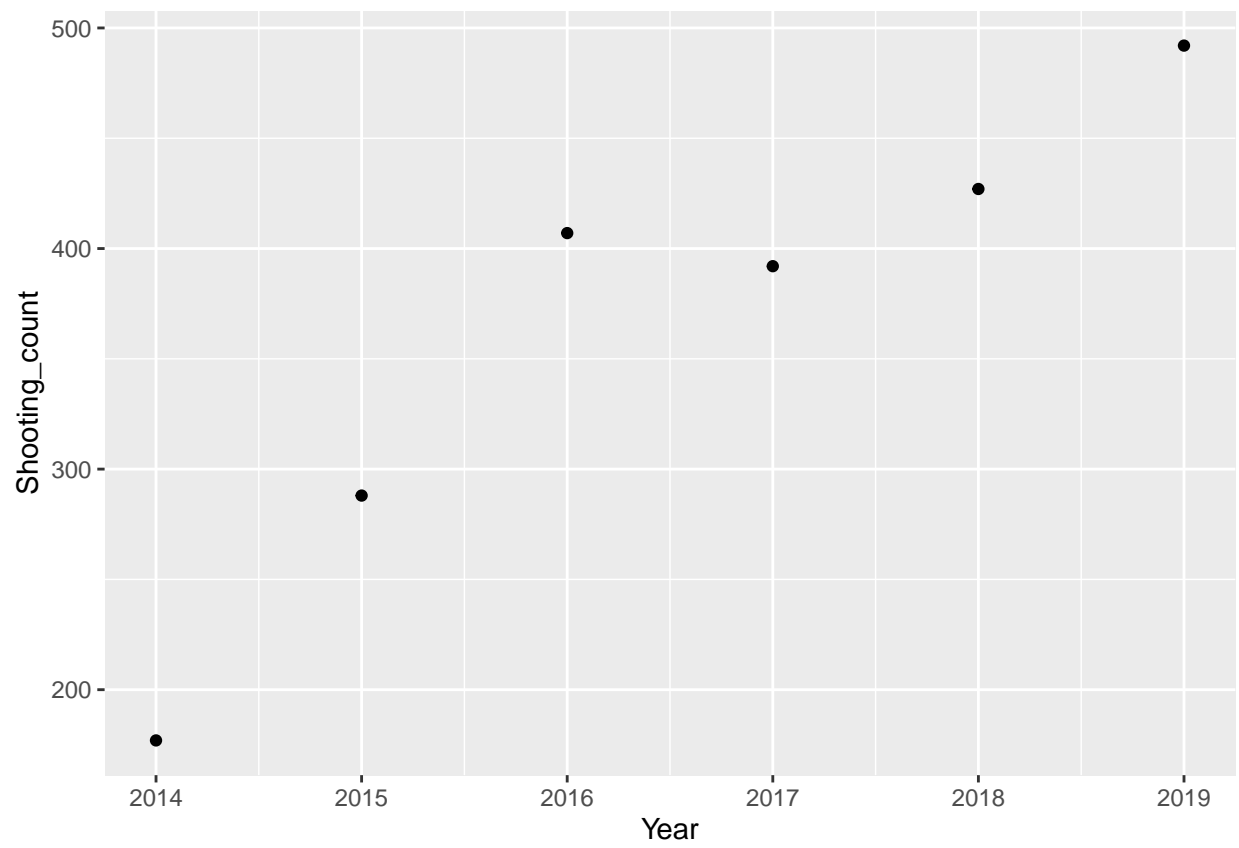


Figure 3: Figure 3: Relationship between year and shooting occurrences

- (Xie 2020)
- (Zhu 2020)
- (Müller 2020)
- (Xie 2021b)
- (Xie 2021a)
- (Wickham 2016)
- (Wickham et al. 2021)
- (Wickham et al. 2022)
- (Arel-Bundock et al. 2022)
- (C. of Toronto, n.d.)
- (O. D. Toronto 2020)
- Arel-Bundock, Vincent, Joachim Gassen, Nathan Eastwood, Nick Huntington-Klein, Moritz Schwarz, Benjamin Elbers, and Grant McDermott. 2022. *Modelsummary: Summary Tables and Plots for Statistical Models and Data: Beautiful, Customizable, and Publication-Ready*. <https://cran.r-project.org/package=modelsummary>.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. <https://CRAN.R-project.org/package=here>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Toronto, City of. n.d. “City of Toronto Open Data Portal.” City of Toronto Open Data Portal. <https://open.toronto.ca/>.
- Toronto, Open Data. 2020. “About Police Annual Statistical Report - Shooting Occurrences.” City of Toronto Open Data Portal. <https://open.toronto.ca/dataset/police-annual-statistical-report-shooting-occurrences/>.
- 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 Golemund, 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>.
- Wickham, Hadley, Jim Hester, Romain Francois, Jennifer Bryan, Shelby Bearrows, Rstudio, Jukka Jylänki, and Mikkel Jørgensen. 2022. *Readr: Read Rectangular Text Data*. <https://cran.r-project.org/package=readr>.
- Xie, Yihui. 2020. *Bookdown: Authoring Books and Technical Documents with r Markdown*. <https://github.com/rstudio/bookdown>.
- . 2021a. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://yihui.org/knitr/>.
- . 2021b. *Tinytex: Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents*. <https://github.com/yihui/tinytex>.
- Zhu, Hao. 2020. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.