Shooting Occurrences in Toronto have increased substantially all throughout the city: A Data Analysis*

Thomas D'Onofrio

04 February 2022

Abstract

Toronto, which was once seen as a safe place to live, has seemingly had a lot more gun violence occur within the city over the years. Data on shooting occurrences in Toronto, including the years and police divisions in which they occurred ,has been obtained, and will be analyzed in this paper with the use of several figures. It was found that shooting occurrences have greatly increased all throughtout the city of Toronto, and overall have more than doubled since 2014. Our discovery unfortunately has some implications, as police have racial biases against people of colour which have caused the dataset worked with to be biased.

1 Introduction

Canada was once thought of as an extremely safe and peaceful place to live, with very little violence. Unfortunately, Toronto, Canada's most populated city, is starting to change that narrative. Although overall crime rate has decreased about 30% in Canada over the past 20 years (Statista Research Department 2021), there has been a recent increase in what may be the most worrisome crime of them all, gun violence. As a resident of Toronto, it seemed as if every other day you were hearing about a new shooting in the news, whereas in the past it was rarely mentioned. This raised a lot of questions about gun violence in Toronto and how it has changed over the past several years, with many worried about this recent upwards trend. The COVID-19 pandemic has caused the focus on gun violence in Toronto to die down a little, as there has been a 35% decrease in shooting occurrences since it started (Vilkov 2021), but Toronto should not let this distract them from the severe issue that was becoming worse and worse every year before 2020.

In this paper, we use R (R Core Team 2020) to analyze a dataset from opendatatoronto (Gelfand 2020) on shooting occurrences in Toronto in order to get a better idea on how exactly gun violence has differed in Toronto recently and tell a story about these changes. The dataset being analyzed contains observations on how many shootings occurred each year in each Toronto police division from 2014-2019. Using tables and useful visualizations like scatter plots and bar graphs, the trends of shooting occurrences in Toronto were able to be seen. It was discovered that shooting occurrences in Toronto have been increasing steadily over the years, and has more than doubled from 2014 and 2019. It was also found that this increase in shootings remained true for all areas and police divisions in Toronto, especially the most populated and largest, meaning this is a citywide issue. Overall, the trends in the data confirm that gun violence and shooting occurrences was becoming an increasingly bigger issue in Toronto overtime.

These findings do unfortunately have some implications to them. There is a lot of racial bias in the police force which tends to result in Black people and other people of colour to be stopped and charged in relation to gun related incidents much more than White people (Gelman, Kiss, and Fagan 2005). This racial bias that police have against minorities means that the dataset being worked with is most likely biased, and has been effected by the racial injustice shown towards people of colour (Bronner 2020). This means we must view our findings with the understanding that they may not be perfectly correct as a result of this bias, but overall there is still a lot to be learned from the data and results.

---

*Code and data are available at: https://github.com/TDonofrio62/Paper-on-Shooting-Occurrences-in-Toronto.

The rest of this paper will be as follows: Section 2 goes over the dataset being used and analyzes it with the use of tables and figures to tell a story about shooting occurrences in Toronto.

2 Data

The dataset being used in this paper was taken from the City of Toronto's open data portal using the opendata toronto package (Gelfand 2020). It is a subset of a larger collection of data called The Toronto's Police Services Annual Statistical Report. This report contains data based on police related topics including a variety of different types of crimes and other incidents, and administrative information. The dataset that will be analyzed draws from the "Shooting Occurrences" portion of the annual statistical report. It contains all Toronto shooting incidents between 2014 and 2019. A shooting occurrence is described as an altercation where a projectile was shot from a firearm and injured a person (Toronto Police Services 2022a). Toronto Police Services do not count suicidal attempts or shots taken by Police in the data.

The data is collected by Policemen through the Toronto Police system. Any time a shooting is reported by an officer, it is put into the system, and is eventually added as a observation in the overall dataset. This data technically is the population of all police reported shooting incidents by their definition, but it could also be considered a sample of all shooting incidents since many go unreported. As the dataset is made up of observations that follow a specific definition, it helps to avoid many of the biases we typically see in datasets as there is less subjectivity. All in all, the system Toronto Police uses allows for this dataset to be very reliable and trustworthy. Toronto Police takes their data very seriously, which is evident by their Public Data Safety Portal (Toronto Police Services 2022b), and thus the dataset this report is working with is quite strong.

Despite this, the data does not come without any implications. Considering we are working with crime and police data, unfortunately there are historically many social and cultural issues that impact our data and make it slightly biased, and thus must be considered from an ethical perspective (Crawford 2021). As most are aware, there has been a lot clear evidence compiled over the years that show racial injustice against non-white people by the police force. Studies have shown that in the USA, Black and Hispanic people are stopped at a much higher rate than whites, and are twice as likely to be stopped and reported for violent crimes and weapon related issues (Gelman, Kiss, and Fagan 2005). This racial bias occurs in Toronto as well, as Black people account for 32% of all charges despite only making up 8.8% of the total population (Wortley, Laniyonu, and Laming 2020). That means that this bias towards stopping non-white people, especially in regards to gun related incidents, most like carries over into our shooting occurrences dataset. For example, there may have been more shootings reported in police divisions with a larger minority populations compared to others. This means that our data comes from a biased sample as opposed to a full one, as the police and their racist tendencies have effected the observations in our dataset (Bronner 2020). One may argue that this bias is less likely to occur since the data was taken from Toronto

Although our dataset may be biased, there is still a lot that can be learned from it. The data can be analyzed in R (R Core Team 2020), using packages such as Tidyverse (Wickham et al. 2019), Dplyr (Wickham et al. 2021), and Knitr (Xie 2014) in order to produce tables and figures that tell us important information about the dataset and can be use to tell a story about shooting occurrences in Toronto.

Table 1 shows an extract of 12 data points from the overall dataset which contained 96 observations. Two observations from each year are shown in the extract. The original data had 6 variables, including 3 redundant identifier variables, so it was narrowed down to the three main variables seen in the extract, plus one identifier variable was kept for organizational purposes. These three main important variables that will be worked with in the paper are Year, Geographic Division, and Number of shootings, which are occurred_year, geo_division, and num_shootings respectively in the dataset.

Occurred year is fairly self explanatory, and is the year in which all the shootings in an observation took place. It is a numerical variable, but was also converted to a categorical at points for use in visulaizations seen later. Geo division is a categorical variable representing the geographically determined area boundaries for each police division. Num of shootings is a numerical variable that counts the number of shootings corresponding to each observation.

Table 1: 12 Observations from dataset of shooting occurences in Toronto

Year	Geographic Division	Number of Shootings
2014	D11	2
2014	D12	20
2015	D12	36
2015	D13	4
2016	D14	23
2016	D22	15
2017	D23	44
2017	D31	68
2018	D32	23
2018	D33	12
2019	D41	23
2019	D42	35

Table 2: Number of Shootings in Toronto each year from 2014-2019

Year	Number of Shootings
2014	177
2015	288
2016	407
2017	392
2018	427
2019	492

As shown in Table 1, each observation tells us the number of shootings that occur in a specific police division, based on the geographical boundaries described above, in a certain year. For example, the last observation seen in Table 1 tells us that in 2019, 35 shootings occurred within police division 42. These data points can be analyzed together in order to tell us a story about shooting incidents in Toronto.

By grouping the data by year, Table 2 shows how many shootings have occurred each year in Toronto. It can be seen that shootings have increased dramatically from 2014 to 2019, more than doubling, going from 177 to 492 shootings. For the most part, shootings have also increased each year apart from 2017. This increase in shootings is something worth analyzing, and can be better seen and understood through the use of a visual.

Figure 1 shows this relationship between the number of shootings in Toronto and the year in which they occurred, while also showing which police divisions each observation occurred in.

Plotting every single observation in the dataset gives us a better idea about the changes for Toronto's shooting occurrences, and also allows us to link them to each division. The plot points for each year seem to shift up as a collective year after year, suggesting that shootings are increasing all around Toronto. No police division's plot point is lower in 2019 than it was in 2014, which confirms this thought. Police Divisions that have the most shootings see much bigger yearly increases compared to the regions with less shootings when judging based on magnitude of the increases. Relative to how many shootings occur in each division, all divisions seem to increase by similar amounts.

A bar plot will make it easier to see the increasing relationship between the year and the number of shootings. Figure 2 shows this relationship.

The increase in shootings per year is even more evident now. The bar for 2019 towers over the much smaller bar representing 2014 shootings, once again showing just how much worse shooting incidents have gotten over these 5 years. On a year to year basis, we see the biggest growth in shootings from 2014 to 2016, suggesting
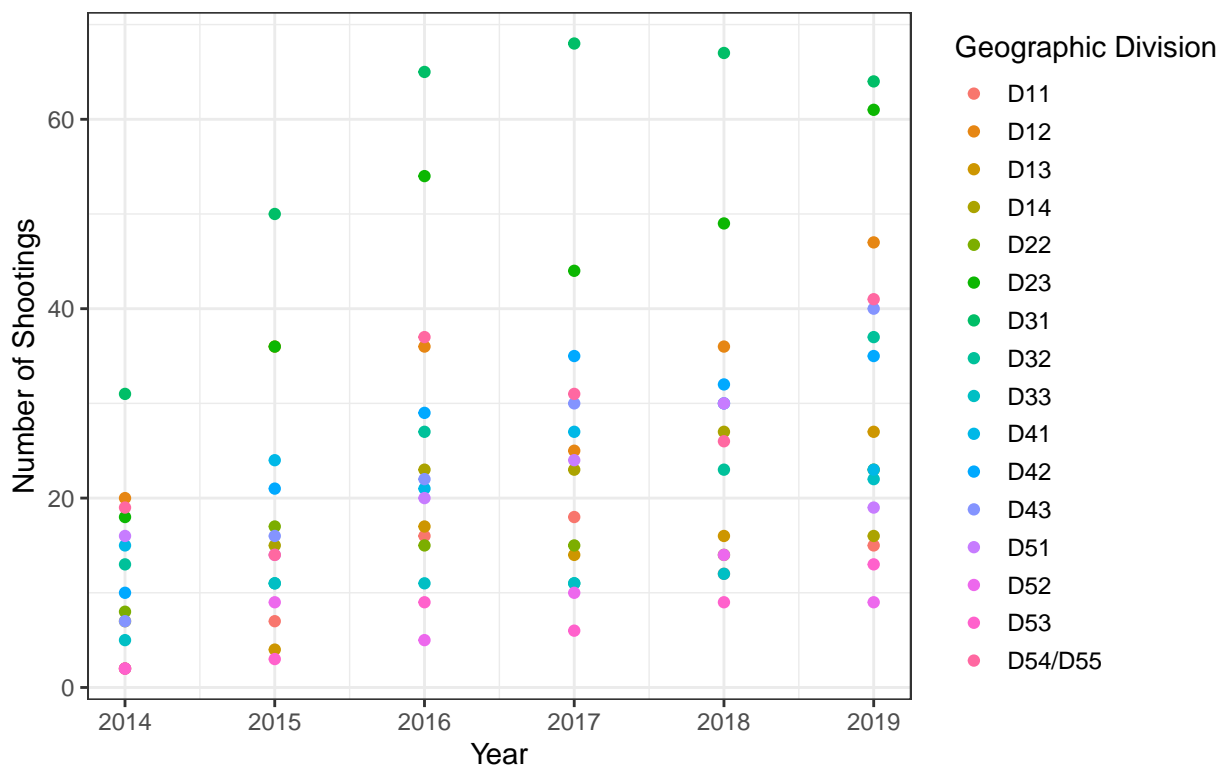
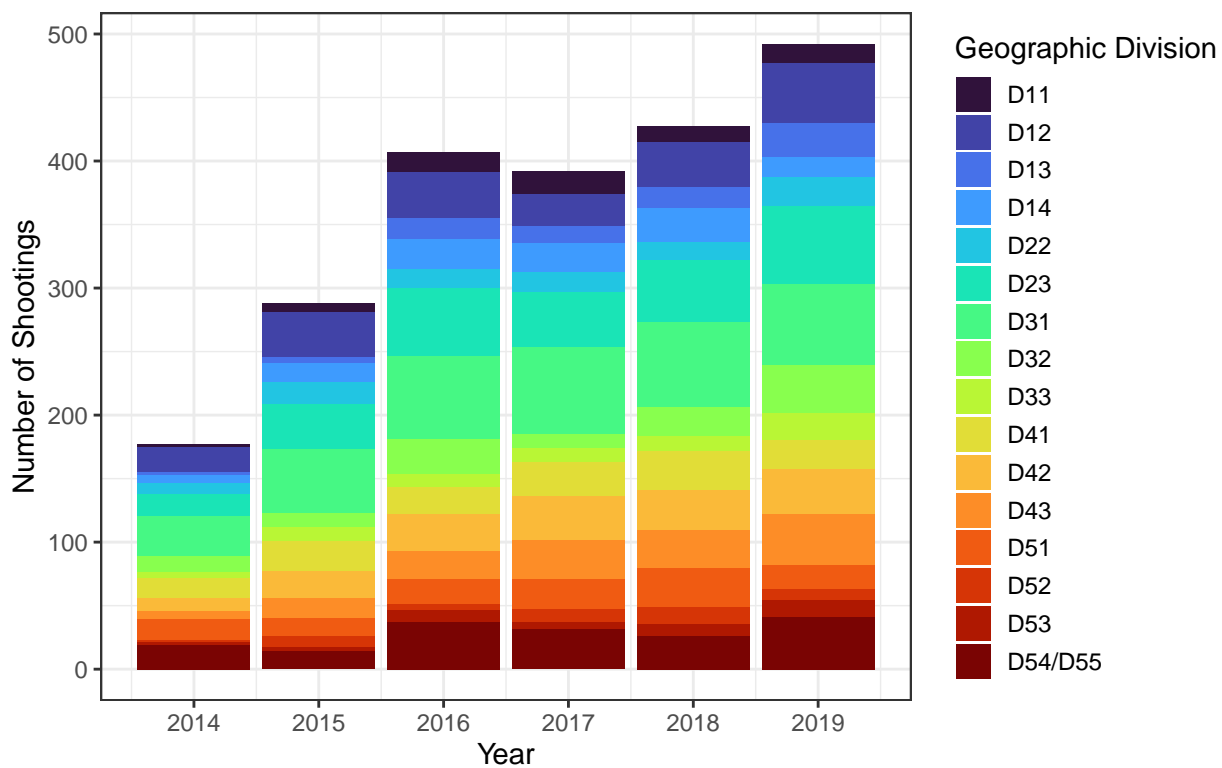Figure 1: Relationship between number of shootings in each police division and year

Figure 2: Bar plot showing number of shootings in each police division for every year

something important changes around this time period. There is then a small dip in 2017, before a steady increase in shootings starts to occur again from this point on wards. A notable increase can be seen from 2018 to 2019, as 2019 saw the biggest increase in shootings since 2016. If this is a sign of another period of large growth like there was from 2014 to 2016, Toronto should be very worried. Regardless, there is certainly no signs of the trend slowing down anytime soon, and the number of shootings should continue to increase in future years, especially once the world recovers from the COVID-19 pandemic.

Figure 2 also makes the changes in number of shootings in each police division more evident. Every single coloured section representing the divisions are larger in 2019 than 2014, and the number of shootings for most divisons increased every single year. The larger police divisions specifically saw massive growths in the size of their colour in the bars of the Figure Figure 2, which was also, albeit slightly less, clear in Figure 1. This can be seen in D12's, D23's, and D31's coloured parts of the bars, as their filled in areas in the bars are very large in 2019 compared to what they were in 2014. A final figure, Figure 3, can be used to show how the number of shootings has changed in each division even further.
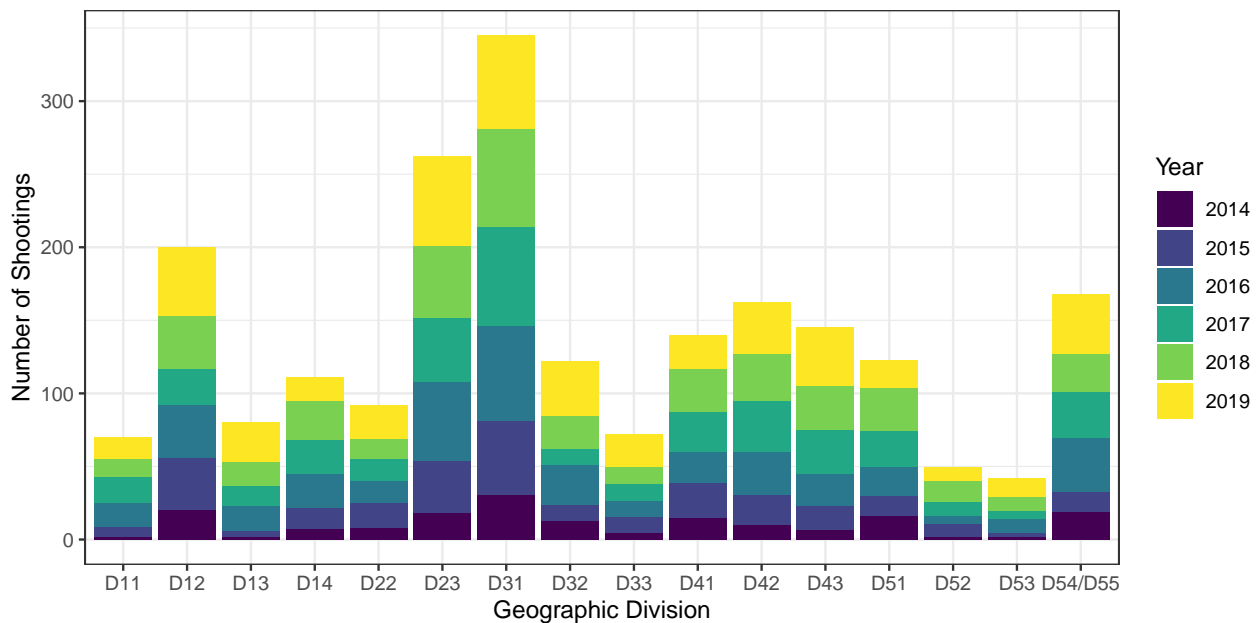

Figure 3: Bar plot showing number of shootings every year for each police divison

It can further be seen in Figure 3 that every police division had more shootings occur within it in 2019, the yellow part of every bar, compared to in 2014 which is shown by the dark purple parts of the bars. It is also clear how almost every police division has seen a steady increase in shootings since the size each colour representing a year are almost always increasing in size.

The divisions with much more shooting activity than the rest are very evident in Figure 3 as their overall bar size are much taller than the rest of the divisions, which is important to note as Toronto Police may opt to add more police presence to these divisions in the future. The fact that the yellow portions of the bars representing 2019 are so much larger relative to the dark purple bars representing 2014 really shows just how bad the number of shootings in these areas specifically has gotten.

Overall, our data has painted a clear story about the shooting occurences in Toronto. It is extremely evident based on Table 2, Figure 1, and Figure 2 that the number of shooting occurrences in Toronto have increased greatly from 2014 to 2019, and that this increase has been occurring steadily over the years since 2014. While the other figures helped, Figure 3 made it clear that shooting occurrences have been increasing all throughout Toronto in every single police division, although some of the larger divisions in particular have seen even bigger increases than the rest. It is still important to note that the racial bias towards people of colour by the Toronto Police has caused our dataset to be biased, which may have caused implications for our findings.

References

Bronner, Laura. 2020. *Why Statistics Don't Capture the Full Extent of the Systemic Bias in Policing.* FiveThirtyEight. https://fivethirtyeight.com/features/why-statistics-dont-capture-the-full-extent-of-the-systemic-bias-in-policing/.

Crawford, Kate. 2021. *Atlas of AI.* Yale University Press.

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal.*

Gelman, Andrew, Alex Kiss, and Jeffrey Fagan. 2005. *An Analysis of the NYPD's Stop-and-Frisk Policy in the Context of Claims of Racial Bias, Columbia Public Law Research Paper.* Columbia Public Law Research Paper. https://scholarship.law.columbia.edu/faculty_scholarship/1390.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Statista Research Department. 2021. *Crime in Canada - Statistics & Facts.* Statista. https://www.statista.com/topics/2814/crime-in-canada/#dossierKeyfigures.

Toronto Police Services. 2022a. *TPS Crime Statistics - Shootings & Firearm Discharges.* https://data.torontopolice.on.ca/pages/shootings.

———. 2022b. *TPS Public Safety Data Portal.* https://data.torontopolice.on.ca/.

Vilkov, igor. 2021. *Toronto Crime Rates in 2021.* Vilkov Law. https://vilkhovlaw.ca/toronto-crime-rates/.

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.

Wortley, Scot, Ayobami Laniyonu, and Erick Laming. 2020. *Use of Force by the Toronto Police Service.* Ontario Human Rights Commission. https://www.ohrc.on.ca/sites/default/files/Use%20of%20force%20by%20the%20Toronto%20Police%20Service%20Final%20report.pdf.

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. http://www.crcpress.com/product/isbn/9781466561595.