Sin and The City: Investigating the Number and Sex of Victims of Crimes Against the Person in Toronto Across 2014-2021.

Sofia Sellitto

February 3, 2023

Abstract

In the city of Toronto, citizens have become accustomed to fearing for their safety as crime appears to have increased in recent years. In this paper, I analyze the number of victims of crimes against the person in the City of Toronto from 2014 - 2021 to investigate if the number of Torontonians victimized by crime has increased over time and if sex differences are apparent across various categories of crime. By examining the data I found that in all categories of crime (assault, robbery and other) but sexual assault, the number of annual victims has decreased over time. Sexual violence appears to have increased in 2021 maintaining the largest sex difference, skewing drastically female. These results have implications for how Torontonians view their safety and consider how sex impacts their predisposition to certain crimes.

Table of contents

[1]

1 Introduction

Prior to the 2006 Canadian federal election, Stephen Harper, who would eventually be elected Prime Minister until 2016, led a conservative campaign based on crime prevention (Mallea 2011). The conservative party asserted that Canada was facing a period of increased crime and assured protection to Canadians by way of tougher sentencing/ parole, additional police funding, and investments in victim services, among a myriad of other things (Mallea 2011). In his 2011 campaign, Harper suggested that judicial opinions favoured the rights of criminals instead of Canadian citizens (Mallea 2011). Harper's claims of increased crime and strides for harsher punishment sparked debates of fear-mongering and inducing an era of incorrectly generated moral panic. According to statistics Canada, the total rate of crime had declined over 20 years since its peak in 1991 (Mallea 2011). This was in direct contradiction to Harper's claims.

Less than a decade following the Harper era, in the city of Toronto, violent crimes against the person have become customary as Torontonians regularly hear about and fear crime. Many of the crimes that have come

---

[1]Code and data supporting this analysis is available at: LINK

to shake Torontonians, specifically those which take place on the major transportation organization, the TTC (Toronto Transit Commission) appear to target random citizens and are often extremely violent forms of assault. Akin to Harper, it is challenging not to believe that crime has increased and feel at risk of being a victim. Nevertheless, as a young woman living in Toronto's Downtown core, it is difficult to avoid nihilistic thinking and continual fears concerning my safety as I exist in this city. While the analysis of cumulative crime rates is extremely valuable, I don't believe that this represents the root of Torontonian's fear - the possibility of being a victim themselves.

In this paper, I utilize Open Data Toronto (Gelfand 2022) to assess the number of victims of crimes against the person in Toronto across 2014 - 2021. This is to investigate how the number of victims of crime in Toronto has changed over time, if it is increasing, which categories of crimes are most prevalent and if sex differences are apparent across different forms of crime. I will commence with a discussion of my data source and its features. I will then conduct a breakdown of the average number of victimization reports for the sub-categories "assault", "robbery", "sexual violence" and "other", across 2014-2019. Next, I will move towards an analysis of how the number of victims has changed according to the category of crime, over time. Following, I will conduct an investigation of how the number of victimization reports for each category of crime, differs according to the sex of the victim. The paper will culminate with a discussion of my findings, limitations and future research possibilities.

2 Data

2.1 Data Source

The data source utilized for this paper was taken from the City of Toronto's open data portal (Gelfand 2022). The data was collected from the Toronto Police Service Annual Statistical Report (ASR) which conducts overviews of police-related statistics. For this paper specifically, the dataset is comprised of all identified victims of crimes against the person (victims of crimes) from 2014-2021.

2.2 Data Features

The data was collected across 8 years, spanning from 2014-2021. It includes all reports of identified victims of crimes against the person found within but not limited to the city of Toronto (some crimes may have occurred outside of city limits) and are not verified by location.

The data is categorized across 9 variables which include; the reported year of the crime, the category, category subtype and assault subtype of the crime, the sex of the victim(s), their age group and cohort, and the count (number) of identified victims. It was stated that duplicates concerning the count of identified victims are possible if the same individual was victimized for the same crime during the same period. This leads me to a discussion of potential biases included in the data source. To begin, it can often be difficult to fully trust that police statistics were ethically collected and reported. Oftentimes, citizens make statements that police fail to report or later on dismiss. Furthermore, the Toronto police may hold biases against communities of colour, thus potentially making them more likely to create victim reports against wrongfully accused individuals of colour.

This paper will follow an analysis of the following variables: the reported year of the crime, the crime category subtype, the number of identified victims and their respective sex. I have chosen to exclude the assault subtype of the crime as it included numerous missing values. I would have been forced to remove this data, thus negatively impacting my analysis. Because I am focused on victims of crime over time and potential gender differences, I have withdrawn from using data on age group or cohort.

Included within the crime category subtype are (1) assault, (2) robbery, (3) sexual violation and (4) other. To further discuss these categorizations, in Canada, assault is classified as intentionally applying force to an individual indirectly or directly without consent. This can then include assault to cause bodily harm, with a deadly weapon, aggravated assault or attempted murder. In Canada, robbery is defined as stealing for

extortion of the article stolen and can include the use of violent threats, personal violence or with a weapon. In Ontario, sexual violence describes any physical or psychological violence conducted or targeted in a sexual means. Unfortunately, no information was provided by the City of Toronto nor the Toronto Police Service Annual Statistical Report (ASR) on the nature of the "other" category. We are forced to assume that this may include break and enter, homicide, domestic violence or various forms of crimes against the person.

2.3 Data Analysis

Using the R coding language (R Core Team 2022), I processed, cleaned and analyzed the data set . I then loaded the packages tidyverse (Wickham et al. 2019), janitor (Firke 2021) and dplyr (Wickham et al. 2022). These were used throughout my analysis to call on, sort, categorize and clean my data. To display my data, the packages ggplot2 (Wickham 2016) and knitr (Xie 2023) were loaded and used respectively for creating graphs and tables.

I initiated my analysis by cleaning and sorting my data. I first selected the variables of interest, checked for missing values, cross-tabulated to understand the pattern of unidentified data and dropped unknown variables. Included in the sex variable was "U" for unknown, which I was unfortunatelyforced to remove it from my dataset. This is a limitation of my analysis that will be further examined in my discussion.

My investigation began with an overview of the average number of people victimized by crime in Toronto across 2014-2021 based on the sub-category of crime. I visualized this data to create categorical organization from the highest to lowest average victims of crime, see table 1.

Table 1: Average number of victims of crimes in Toronto across 2014 -2021, grouped by category of crime

Category of Crime	Average Number of Victims
Assault	16858.625
Other	4521.500
Robbery	3059.250
Sexual Violation	2379.125

From this table, it is apparent that assault has the highest average number of victims total spanning 2014 - 2021.

From there, I began preparing the data frame I would need to visualize how the number of victims in each sub-category of crime has changed over time. I included the 3 variables; sub-categories, years and the number of victims. Because my data frame still remained relatively large, below, I have visualized the data in a table (see Table 2), which presents only the **first 6 rows** of the data frame I created to be called on later in my analysis.

Table 2: The first 6 rows of the data frame used to represent the number of victims according to category of crime in Toronto Across 2014 - 2021

Sub-Category of Crime	Year	Number of Victims
Assault	2014	15141
Assault	2015	15897
Assault	2016	16550
Assault	2017	17089
Assault	2018	17799
Assault	2019	18819

Similarly, to create the data frame that I would use to visualize how sex impacts the victims of crimes in Toronto, I created a data frame, akin to the one above, but instead included the variable 'sex'. Again, the

data frame remained rather large, so below I have visualized the data in a table (see Table 3), which presents the **first 6 rows** of the data frame I created to call on in my analysis.

Table 3: The first 6 rows of the data frame used to represent the sex of the number of victims according to category of crime in Toronto Across 2014 - 2021

Sub-Category of Crime	Sex	Year	Number of Victims
Assault	F	2014	7290
Assault	F	2015	7502
Assault	F	2016	7714
Assault	F	2017	7880
Assault	F	2018	8221
Assault	F	2019	8679

3 Results

As per my primary objective for this paper, I wish to understand how the number of victims of crime in Toronto has changed over time and to analyze if victimization numbers are increasing over time. Because of this, I created a line graph that showcases the number of victims according to each subcategory of crime over 2014 - 2019 (See Figure 2). To do this, I called on the data frame visualized above (See Table 2). This visualization allows for comparison across the different subcategories of crimes while simultaneously presenting a rich depiction of how the number of victims has changed over time, the year they peaked and at their lowest. Using the data previewed in Table 2, I was able to discriminate the number of total victims for the peak years of each sub-category of crime.
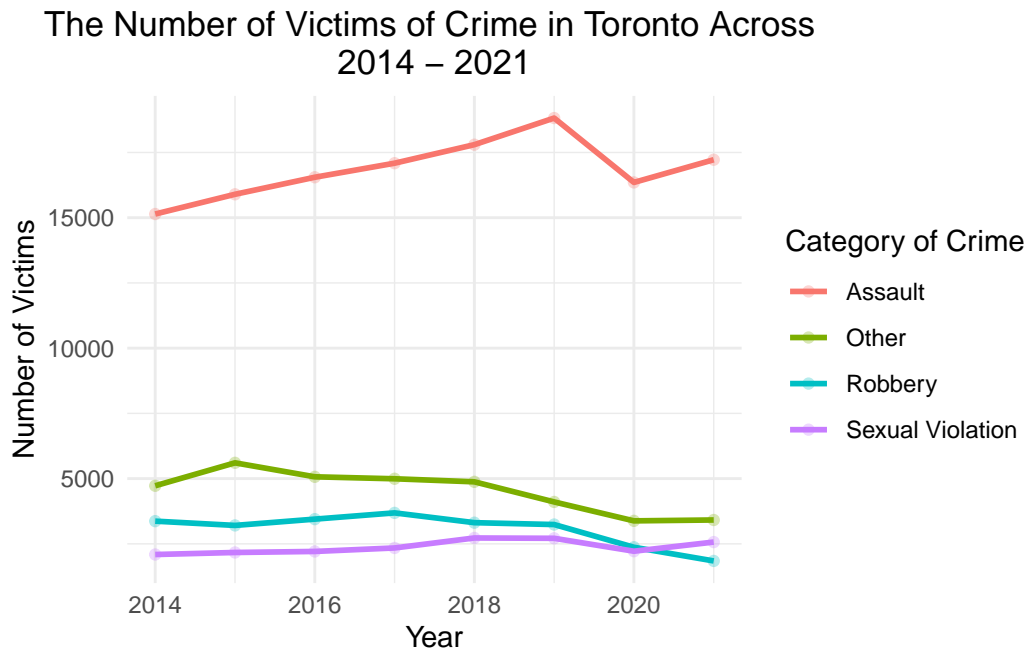
See **Figure 1**

Figure 1: Number of victims according to sub-category of crime in Toronto Across 2014 - 2021

By drawing your attention to Figure 1, it is apparent from 2014 - 2021, the most prevalent crime that Torontonians were victims of was "Assault". From 2014 - 2018, the number of annual victims gradually

increased until it peaked in 2019, with 18819 total victims. The number of victims of assault then began to decrease following 2019, where it spiked in 2021 but remained below 17500 victims.

The "Other" category held the second-highest number of victims from 2014 - 2021. Unfortunately, due to the nature of this variable, valuable insights cannot be drawn. However, it peaked in 2015, seeing 5605 victims, where it steadily decreased until 2021.

"Robbery" had the third most victims of crime from 2014 - 2021. The peak number of victims for this variable is not very distinct, but saw its highest in 2017 with 3686 victims that year. From 2017 onward the number of victims of robberies steadily decreased.

"Sexual Violence" had the least number of victims of crime from 2014 - 2021. However, contrary to the others, sexual violence saw its peak in 2021 with 2568 victims. From 2014 onward, the number of victims gradually increased until it peaked in 2021.

In representing how sex impacts the victims of crimes, I chose to display this data using 4 bar charts based on the categories "Assault", "Other", "Robbery" and "Sexual Violence". Each plot assesses the number of victims across 2014 - 2021 and the number of victims that were male (M) and female (F), respectively. Using the data set from Table 3, which was previewed above, I determined the number of victims that were male and female for the peak years of each sub-category of crime and calculated percentage differences.
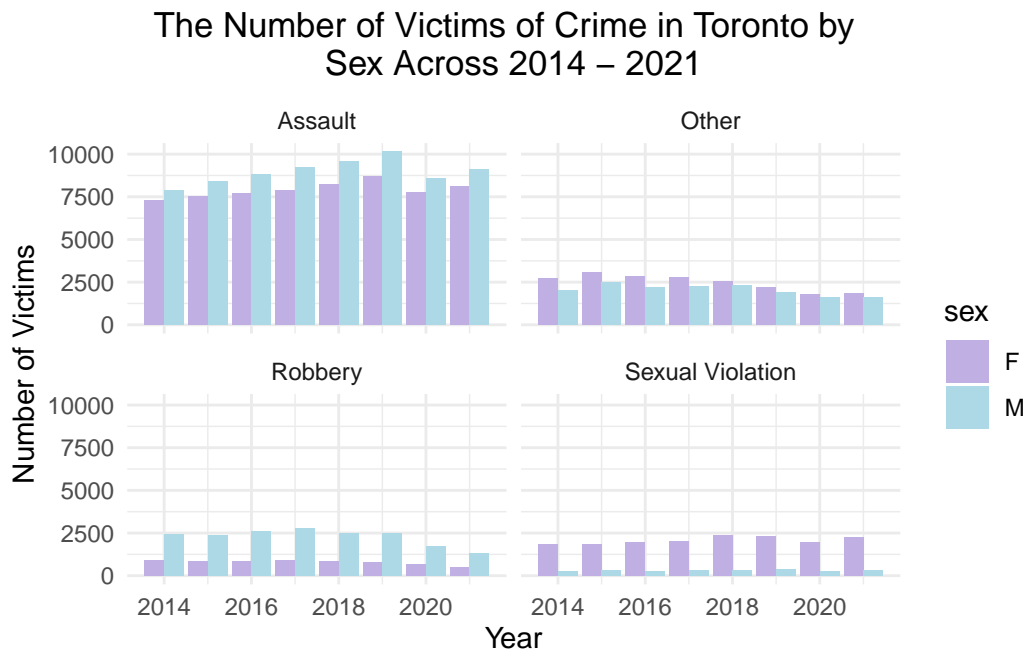
See **Figure 2**

Figure 2: How the victims of crimes varries across sex in Toronto from 2014-2021

In drawing your attention to Figure 2 contained in the "Assault", on the top left, it is apparent that from 2014 - 2021, more male Torontonians were victims of assault than women. In its peak year, 2017, of the 18819 victims, 8679 were women and 10140 were men. This would mean that in 2017, there was 1461 more male than female victims, with a difference of 15.53%.

According to "Other", on the top right, it is apparent that from 2014 - 2021, more female Torontonians were victims than men. Again, because of the nature of this variable, insightful analysis cannot be made, however, the difference between male and female victims is minimal.

As per "Robbery", on the bottom left, it is apparent that from 2014 - 2021, more male Torontonians were victims of robbery than women. Of the 3686 victims in its peak year, 2017, 909 victims were female, and 2777 were men. This would mean that in 2017, there were 1868 more male victims than females, with a difference of 101.35%.

Lastly, "Sexual Violence", on the bottom right, shows the most apparent sex difference from 2014-2021, where more female Torontonians were victims of sexual violence than men. This sex difference was the most defined across all the sub-categories, with 2244 of the total 2568 victims in its peak year, 2021, being female, and 334 being male. This means a difference of 147.9%.

4 Discussion

4.1 Findings

In this paper, I have utilized The City of Torontos open data portal (Gelfand 2022), supplied by the Toronto Police Service Annual Statistical Report (ASR) to assess the number of victims of crimes against the person in Toronto across 2014 - 2021. The primary objective of this paper was to assess if the number of Torontonians victimized by crime has changed, if it has increased over time, which categories of crimes against the person are most prevalent and if sex differences are apparent. By utilizing the Coding language R (R Core Team 2022) I was able to import, clean and asses the raw data made available by the City. In doing so I discovered that for the sub-category of crimes, assault, robbery, and other, the number of Torontonian victims peaked in the years 2019, 2017, and 2015 respectively and decreased until the end of the data collection date (2021). The most prevalent crime that Torontonians were victims of was assault, maintaining the highest cumulative average across 2014-2021. Contrarily, sexual violence saw its peak in 2021, meaning it has not decreased over time. In assessing how sex affects the number of victims for each sub-category of crime, it was apparent that sexual violence had the largest difference between male and female victims, skewing drastically female, maintaining a difference of 147.97%. Robbery maintains the second largest difference between male and female victims, skewing towards men with a difference of 101.35%. "Other", had the third largest difference, with more female victims than male. Assault had the least large sex difference, with slightly more male victims than feel with a difference of 15.53%.

4.2 Limitations

Because of the nature of the data set, my analysis is limited in many regards. To begin with the source, Torontos Police Service Annual Statistical Report (ASR), it is possible that data was misrepresented, incorrectly input, missing, duplicated or potentially falsified. While it is important to avoid speculation about these statistics' validity, I believe it is necessary to highlight possibilities for error. Another vital point to acknowledge is the number of individuals who may have fallen victim to a crime against the person and failed to report it or those who intended to file a report but were dismissed. The latter is an extremely apparent phenomenon for women who attempt to report sexual violence. While this data represents the number of individuals who have fallen victim to crimes against the person across 2014 - 2021, it cannot account for unfilled and dismissed reports. Meaning that the number of victims could potentially be far higher than portrayed in the analysis. Regarding the data set itself, removing missing and unknown data will impact the analysis in various regards. More specifically, concerning the variable sex including the categories "male", "female" and "unknown", I was forced to remove "unknown" as further insights were not provided on this denomination. Removing this data could have prevented insightful analysis of non-binary and Agender individuals.

4.3 Future Research

In the future, I believe that efforts should be made to analyze and investigate data on crime and victimization rates for the most updated reports from 2022-2023. As mentioned earlier in my paper, throughout 2022 and 2023, assaults occurred in Toronto's major transportation commission, the TTC. In the media, it appeared as though these crimes were novel and occurring far more frequently than ever before. While my results showed that the number of assault victims peaked in 2019, and began to decline until 2021, it is possible that this

data looks different when accounting for the current day. Furthermore, I believe that future research on the rate of crime and victims on the TTC would allow for insightful and meaningful findings for Torontonians.

References

Firke, Sam. 2021. *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.

Mallea, Paula. 2011. *Fearmonger: Stephen Harper's Tough-on-Crime Agenda.* James Lorimer & Company.

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. 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. 2022. *Dplyr: A Grammar of Data Manipulation.* https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r.* https://yihui.org/knitr/.