

# The Same Societal Conditions Can Impact Toronto’s Types of Major Crime Very Differently\*

Jayden Jung

February 3, 2023

## Abstract

Measurements of crime is often used to inform vital policies, and even to indicate the quality of a society itself. This paper draws upon Toronto’s recorded Major Crime Indicators in recent years and cross references the observations with secondary research on varieties of societal circumstances and events. We find that one instance of a societal condition can affect different types of crime in nuanced ways. A certain event may increase the rate of one crime type while simultaneously decreasing another, for example. Policies intended to decrease crime should be conscious of this, ensure careful approaches that do not yield unintentional adverse effects, and try to better understand the root of different crimes.

## Introduction

“Crime is a symptom of a diseased society.”

— Alexandria Ocasio-Cortez, New York City Congresswoman, 2020.

As the quote above (The Hill 2020) claims, crime rate is often considered as a measurable indicator of the quality of a societal organization, as there is a meaningful relationship between the two. Some claim that crime is a problem on its own that negatively affects society, and that policies that aim at directly lowering crime will improve society, but this is a mistaken allocation of responsibility, as it wrongly suggests that crime exist without the definite influences of the society it occurs in. Crime is much more symptom-like in nature than this, and seeing data on it can be useful to infer the state of society in given scenarios.

Data on crime can and should be analyzed carefully to inform how a society can be organized better. This paper proposes many ways of visually representing this information with varying axes, specifically with data on Toronto’s major crimes in recent years (2018-2021). Insights from these illustrations of data are looked at alongside secondary resources regarding what was happening at the time in Toronto, Canada, and the world, and connections are drawn where they can be seen. It is found that crime very much fluctuates according to ongoing societal circumstances, and specific crime types are affected very differently by the same events. These conclusions will hopefully mobilize more conscious and intentional consideration of using crime rates to inform policy, facilitating the movement towards more thorough investigation on the sources of crime and how they can be mitigated.

This paper first considers how crime varies across the months of the year, days of the week, and times of day, showing that, while there are general trends that apply to all crime types, there is more nuance in specific crime types’ trends. Similar points are then made about how, though one crime might be more prevalent overall, looking at specific types of premises for their occurrence may reveal more about its severity.

After, the paper delves into how general monthly crime rates have fluctuated over the years, mainly relative to the COVID-19 pandemic persisting through 2020 and 2021. Finally, it is shown that specific crime types have been affected very differently over the same course of time, some decreasing while others increased as the pandemic went on.

---

\*Code and data are available at: <https://github.com/jj-andj/analysis-toronto-mci>

The analysis in this paper is done by using R (R Core Team 2020) and packages like `dplyr` (Wickham et al. 2022), `knitr` (Xie 2014), `tidyverse` (Wickham et al. 2019), `bookdown` (Xie 2016), `janitor` (Firke 2021), and `lubridate` (Grolemund and Wickham 2011). Figures are generated with `ggplot` (Wickham 2016) and `kableExtra` (Zhu 2021), for graphs and tables respectively.

## Data

### Source, Data Collection, & Cleaning

This paper uses data from *Major Crime Indicators* (2022) as read in from the City of Toronto’s Open Data Portal, which was accessed with R and the package `opendatatoronto` by Gelfand (2022). This data set is comprised of Major Crime Indicator (MCI) occurrences in Toronto. This was chosen over other available options on `opendatatoronto` because it is comparably very well organized and complete and encompasses multiple categories of major crime. Other data sets are often too specific, focusing on one type of crime. Using this one allowed this paper to make comparisons across different crime types will operating on a single data set.

The MCI categories recorded are: Assault, Break and Enter (B&E), Auto Theft, Robbery, and Theft Over \$5000 (referred to as Theft Over). Sexual assaults are not part of this data set.

This data represents MCI occurrences that were reported to the Toronto Police and been considered founded, meaning that it only includes instances that were investigated and deemed that the reported offence occurred (*Major Crime Indicators* 2022). It should be noted that there may be discrepancy in reported and founded occurrences and unreported or wrongly unfounded ones. There is a notable history of police reports not being entirely reflective of the reality of crime present in a society. For example, people of marginalized communities justifiably are weary of reporting to the police out of fear that they may be harmed. This paper does not aim to ignore these aspects, but does use the large data set that is available to try to derive general insights about crime trends.

Each entry includes information on both the occurrence itself (the MCI category, date, year, month, day, day of week, time, type of premise) and the reporting of the occurrence (date, year, month, day, day of week, and time), and there is additional information regarding a unique event ID, police division where the offence occurred, identifier and name of the neighborhood it occurred in, and longitude and latitude coordinates if available (Toronto Police Services Analytics and Innovation 2023). As this paper focuses solely on the actual occurrences, the reporting data was not used. Very specific location data was not used, either, as there were many empty entries and it would not have yielded complete analysis.

This information is provided by the Toronto Police Services and is updated semi-annually. Though there is data ranging from occurrences reported from 2014 until mid-2022, this paper narrows in only on data on occurrences that occurred between 2018 and 2021 both to focus on more recent data and to ensure completeness of the data. First, some reports made in recent years actually refer to occurrences that occurred many years ago, earliest in the entry being 1966. There are many older years that have very incomplete data, so we set the lower limit to 2018. Though there is 2022 data, at the time of this paper, there is only information for up until June, making it incomplete. Such, we set our upper limit to 2021. Other than the date-related information for the occurrences, we only included the MCI category and premise type (one of Apartment, House, Outside, Commercial, Educational, Transit, and Other) for the purposes of this paper.

### Observations & Analyses

The data set for analysis had the columns for Occurrence Date, Year, Month, Day, Day of the Week, Hour of the day, the MCI Category, and Premise type. After cleaning for data between 2018-2021 and omitting rows with N/A, we were working with 136,680 entries. Much of the analysis uses the date-related information and MCI category to identify trends over time.

Figure 1 shows all MCI occurrences in the four year range by day of the week that the incident occurred, and Figure 2 does the same by time of day.

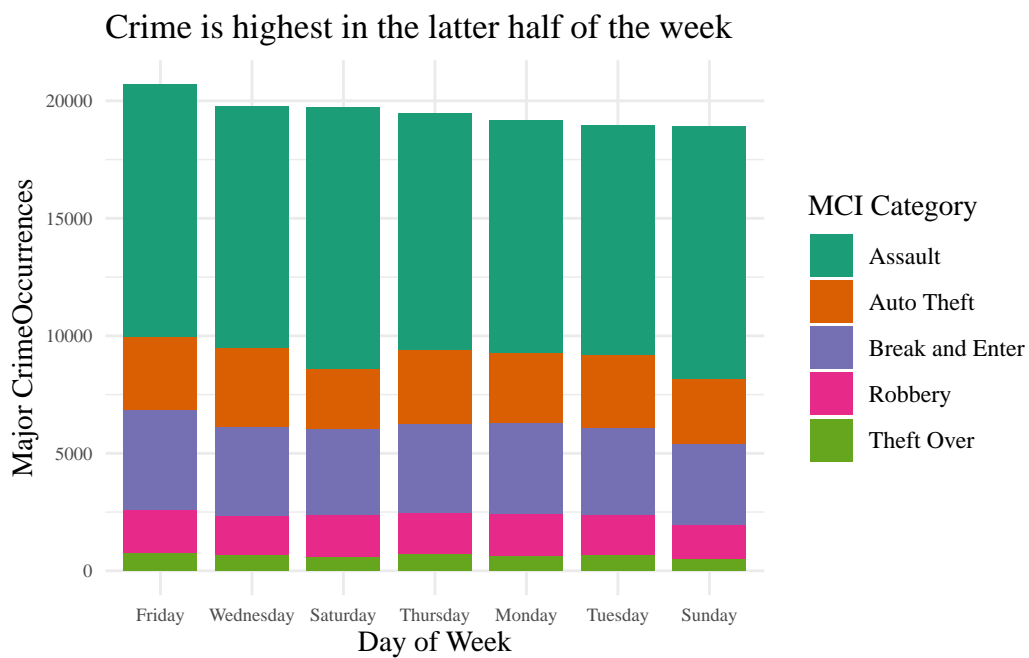


Figure 1: Crime occurrences by day of week.

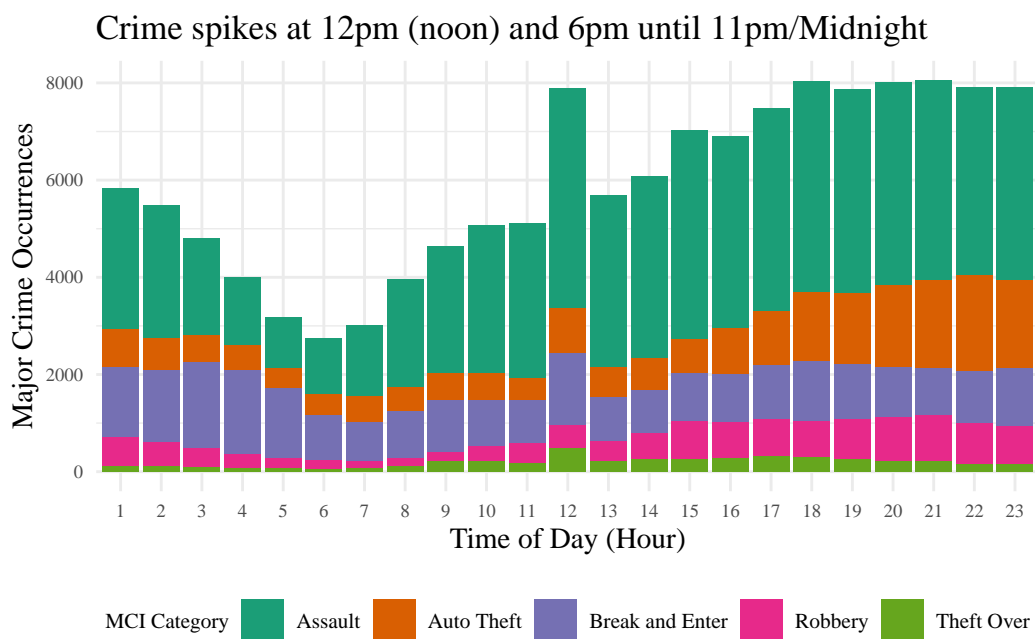


Figure 2: Crime occurrences by time of day.

These figures illustrate that crime is generally higher on the latter half of the week (Wednesday, Thursday, Friday, Saturday) and in the late evening (6pm - 12am). Sundays have the least crime in a week, and 3 am - 9 am sees less crime. Crime doesn't fluctuate a greatly noticeable amount between different days of the week, but it does for hours in a day, the peaks being almost triple the height of the lowest point (6 am). These findings seemingly follow common assumptions that the weekend and the night is more dangerous, though we see that this is not so true the later it gets in the day (1 am onward).

Figure 1 and Figure 2 breaks down each count by the MCI categories. Assaults definitely take up a majority, and there is usually a pretty even proportion throughout each category at different times. One large outlier, though, is that crime greatly spikes in the middle of the day at 12pm. This is the point in the day that has the highest rate of Theft Over and relatively high counts of B&E. Though this time is not so distinct from the other times with high numbers of overall crime (i.e. 6pm-12am), it stands out in over representing these certain MCI types.

On a similar note Figure 3 illustrates the total number of crime by MCI category type with distinction on what the premise types the occurrences happened at, and here also we see that, looking beyond the general trends, we can analyze over representation of certain MCI-types for certain premise types.

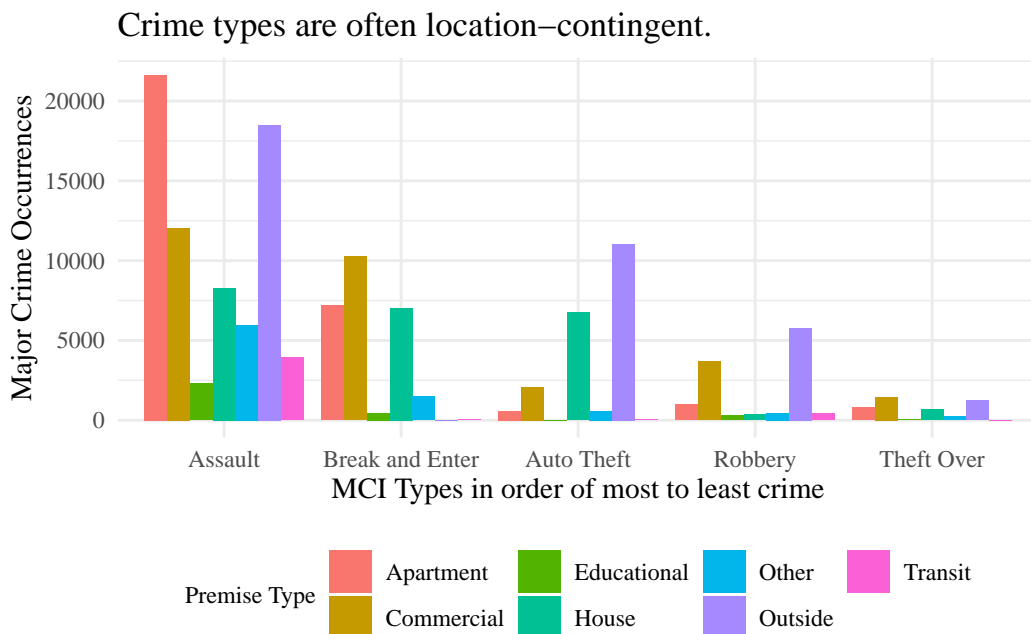


Figure 3: Crime occurrences at different premise types by MCI type

For example, Assault is the highest occurring crime overall, and it is a crime that can occur at every premise type. B&E cannot happen or is very unlikely to happen Outside or on Transit, as the nature of the crime is breaking into restricted or private areas. Considering this, even though B&E occurs much less than assault overall, it occurs on Commercial and House premises as much as if not more than Assault does. This is a more accurate understanding of interpreting these crime rates.

This analysis can also be applied to Auto Theft, a crime that can only realistically occur Outside or in front of a House, and possibly in a Commercial parking garage. With this in mind, we again see how common auto theft is, being comparable to the amount of Assault and B&E that occurs at houses, and noticeably taking up a lot of the crime that occurs outdoors.

These points exemplify that we should look much closer at crime rate data to gain actually meaningful, accurate insights from it, instead of generalizing all major crime into one measurement.

Table 1 shows the exact record of total MCI occurrences per month over the years, while Figure 4 plots this information for more general comparison.

Table 1: Crime per month over the years

Month	2018	2019	2020	2021
January	2920	2879	3033	2377
February	2545	2599	2861	2047
March	2883	2857	2913	2363
April	2632	2949	2294	2220
May	3095	3202	2314	2800
June	2954	3328	2569	2744
July	2861	3451	3040	2829
August	3023	3249	3006	2995
September	2863	3155	2878	2969
October	3104	3451	2714	3097
November	3083	3174	2457	2843
December	2992	2902	2437	2729
<b>Total</b>	<b>34955</b>	<b>37196</b>	<b>32516</b>	<b>32013</b>

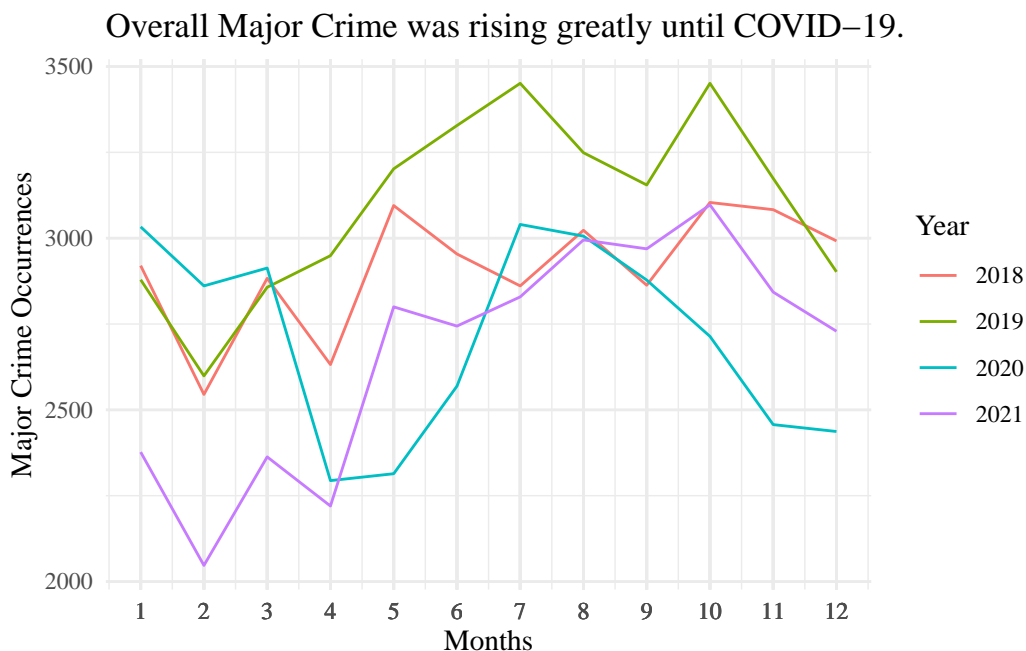


Figure 4: Comparing monthly crime occurrences over the years

Looking at either, it is noticeable that crime rates were greatly increasing through 2018 and 2019, though there are the regular month-related fluctuations (more crime in the summer and October, less otherwise). 2018 had a total occurrence count of 34,955, and then 2019 lead with thousands more at 37,196. Both had the most crime recorded in October at 3,104 and 3,451 occurrences respectively. This is supported by news coverage of the time, as stories arose discussing that the average Canadian felt that they were perceiving more and more crime during this period (Korzinski 2022).

We see in early 2020 that this was continuing, as January, February, and March of that year saw noticeably more crime than the two previous years. Then, April shows a stark, unprecedented dip in overall crime,

recording at ~700 less occurrences than the same month the year before in 2019. This persisted in May 2020 that had an even larger contrast with ~900 less occurrences than 2019. This coincides with the beginning of the COVID-19 pandemic, as Ontario instated social distancing and lock-down measures throughout these months, limiting societal activity in general (Rodrigues 2020). Crime begins to rise again in the summer of 2020 as Ontario began their reopening plan (CBC News 2020), and then declined again from September onward as this initiative was put on pause. Then, we see the lowest number of crime occurrences in early 2021 (most noticeably February with ~2,000 occurrences, aligning with the strictest Ontario lockdown in these months where all non-essential businesses were shut down (Davidson 2020)). As this lock-down loosened later in the year (May onward), crime rate also began increasing, though it had not reached its heights of 2019. The highest crime rate month in 2021, also October, was still ~400 less than 2019.

This analysis has shown that the holistic crime rate fluctuated greatly in accordance with COVID-19-related circumstances. Now, though, looking deeper into how occurrence trends look for specific MCI types, we turn to Figure 5. This graph shows the number of monthly occurrences over the years but broken down to each MCI type. Unlike Figure 4, which overlaps each year on top of each other to see direct comparisons by month, Figure 5 uses a linear timeline of months from January 2018 until December 2021 as the x-axis. By doing so and overlaying the occurrence counts for each MCI type, we can see the long-term trends that they've followed and compare them against each other.

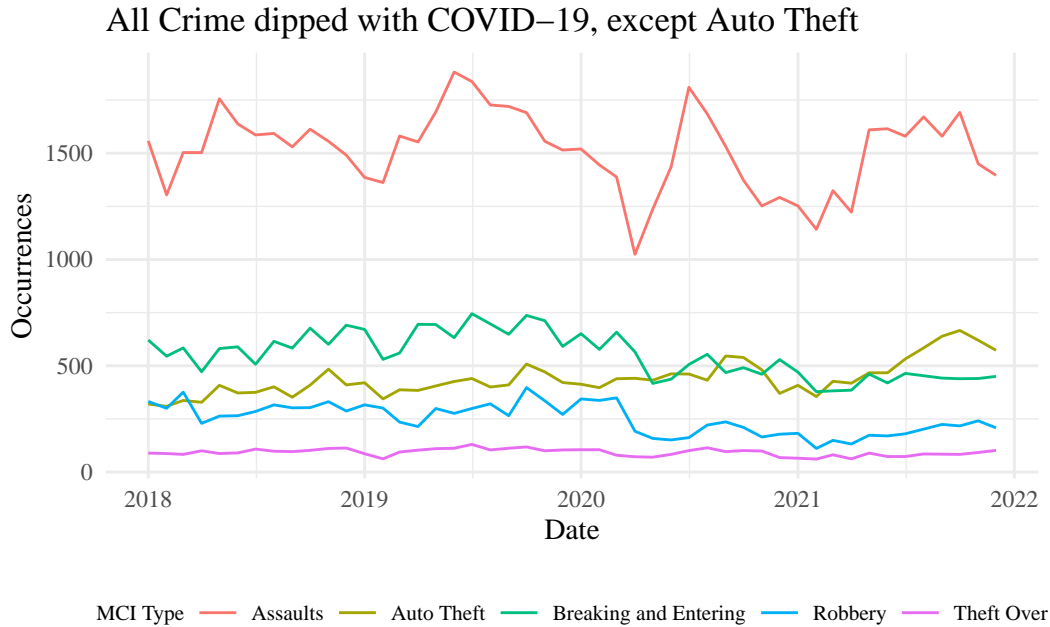


Figure 5: Different MCI type crime occurrences over time.

The most noticeable visual point is the major dip in Assaults in the first third of 2020, in line with the overall trends in Figure 4 and Table 1. Looking closer, we see that this also happened with Breaking and Entering (B&E) and Robbery, and very slightly for Theft Over. What is interesting here is that Auto Theft is the one MCI category that was unaffected by this initial COVID dip. In fact, Figure 5 shows that it increased through 2020, clearly beating out its 2018 and 2019 numbers. Even Assault, which did see a major recovery in rate in mid-2020, did not surpass its peak from 2019. Auto Theft continued to rise, and, in late 2021, hit the highest monthly records seen in all four years.

Researching current events records at the time yields that lock down measures had led, at the time, to car manufacturers shutting down, stiffening the supply of available cars on the market (Khan 2022). This, coupled with the constant demand for cars and car parts, meant that car prices went through a major increase, creating more incentive for Auto Theft. While COVID-19 lock downs may have brought major

crimes overall, it indirectly affected the specific category of Auto Thefts by causing an imbalance in the car market. This effect persisted into an ongoing crisis of the proliferation of Auto Theft.

If we were to have only looked at general crime rates as we did in Table 1 and Figure 4 without the application of Figure 5, we may have arrived at an incomplete understanding that COVID-19 decreased all major crime. Only by dividing out our analysis by MCI type are we able to notice these details.

With the aim of bettering our society, and thus seeking policies that reduce crime, data on the topic should be used as a measure of ongoing circumstances. Beyond that, this paper has shown that simply looking at an overall crime rate is likely not enough in such an important context, and has exemplified that in multiple cases. The available data on crime occurrences can be very helpful to gain insights from and inform the way our society is organized, and so it should be treated with nuanced approaches that do not wrongfully overgeneralize the information.

## References

- CBC News. 2020. “Ontario Puts ‘Pause’ on Further Loosening of Public Health Measures as Covid-19 Numbers Rise.” CBC/Radio Canada. <https://www.cbc.ca/news/canada/toronto/covid-19-coronavirus-ontario-september-8-update-1.5715518>.
- Davidson, Sean. 2020. “Ontario Will Enter Strict Lockdown on Dec. 26, Nearly All Non-Essential Businesses to Close.” *Toronto*. CTV News. <https://toronto.ctvnews.ca/ontario-will-enter-strict-lockdown-on-dec-26-nearly-all-non-essential-businesses-to-close-1.5239810>.
- 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>.
- Government of Canada, Statistics Canada. 2021. “After Five Years of Increases, Police-Reported Crime in Canada Was down in 2020, but Incidents of Hate Crime Increased Sharply.” *The Daily* <https://www150.statcan.gc.ca/n1/daily-quotidien/210727/dq210727a-eng.htm>.
- Grolemund, Garrett, and Hadley Wickham. 2011. “Dates and Times Made Easy with lubridate.” *Journal of Statistical Software* 40 (3): 1–25. <https://www.jstatsoft.org/v40/i03/>.
- Khan, Coco. 2022. “Why Are Used Cars so Expensive Now? We Ask an Expert.” *The Guardian*. Guardian News; Media. <https://www.theguardian.com/lifeandstyle/2022/jul/22/why-are-used-cars-so-expensive-now-we-ask-an-expert>.
- Korzinski, David. 2022. “Half of Canadians Say Crime Is Rising in Their Communities, as Confidence in Police, Courts Wanes.” *Angus Reid Institute*. <https://angusreid.org/justice-system-confidence-2020/>.
- Major Crime Indicators. 2022. City of Toronto Open Data Portal, Toronto Police Services. <https://open.toronto.ca/dataset/major-crime-indicators/>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Rodrigues, Gabby. 2020. “Ontario Government Declares State of Emergency Amid Coronavirus Pandemic.” *Global News*. Global News. <https://globalnews.ca/news/6688074/ontario-doug-ford-coronavirus-covid-19-march-17/>.
- The Hill. 2020. *Rep. @AOC: “Crime Is a Symptom of a Diseased Society That Neglects Its Most Marginalized People, and We Do Not Solve That Problem with Police.”*. <https://twitter.com/thehill/status/1282502759388819457?lang=en>.
- Toronto Police Services Analytics and Innovation. 2023. *Public Safety Data Portal: Open Data Documentation*. <https://torontops.maps.arcgis.com/sharing/rest/content/items/c0b17f1888544078bf650f3b8b04d35d/data>.
- 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. 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.
- . 2016. *Bookdown: Authoring Books and Technical Documents with R Markdown*. Boca Raton, Florida: Chapman; Hall/CRC. <https://bookdown.org/yihui/bookdown>.
- . 2023. *Tinytex: Helper Functions to Install and Maintain TeX Live, and Compile LaTeX Documents*. <https://github.com/rstudio/tinytex>.
- Zhu, Hao. 2021. *kableExtra: Construct Complex Table with ‘Kable’ and Pipe Syntax*. <https://CRAN.R-project.org/package=kableExtra>.