

Analysis of Toronto Crime Trends in 2014*

Tayedza Chikumbirike

19 April 2023

In this paper, we are interested in Major Crime Indicators (MCI) occurrences in Toronto. We obtain data on offenses that took place in 2014 from Open Data Toronto which was collected by the Toronto Police Service Public Safety Department. We find the most common MCI occurrence and the relationship between MCI occurrence, premises type, and occurrence day of the week. This work has relevance given the recent increases in criminal offenses in Toronto. It is important to develop an understanding of such data to inform areas of improvement in order to increase levels of safety.

1 Introduction

Crime—the act of completing an offense— has been present in human societies throughout history. It is a large indicator of societal organization. Toronto, the largest city in Canada, experiences a wide range of crime-related challenges, including but not limited to violence, property crime, robbery, and drug-related offenses. This can have a significant impact on individuals, communities, and societies as a whole as this does have a significant impact on public safety, economic prosperity, and social cohesion.

Being able to understand the causes and consequences of crime, what types of crimes are more frequent in certain areas, locations, and days, as well as develop effective strategies to prevent crime is extremely critical to ensure public safety. While Toronto’s crime rate is relatively low in comparison to other major North American cities, it continues to be a significant issue that requires ongoing attention and action.

The ability to analyze data on crime to inform government officials what areas require improvement is crucial and this paper proposes a way for this type of data to be explored. Our findings show that assault is the most common MCI occurrence as there was always a reported incidence every day at every location. We also find that there could be a potential relationship

*Code and data are available at: <https://github.com/tayedzac/crime.git>

between house auto thefts and house break and enters as the data suggests that these may have occurred at the same time.

This paper holds a total of 4 sections. In section 2, we explore the data source, data cleaning, and data visualization through a series of figures and tables. Section 3 contains a graph displaying how the 3 variables of focus—MCI offense, day of the week, and premises type—interact with each other. Finally, in section 4 we go into further analysis explaining what the results mean as well as addressing any weaknesses and next steps.

For this implementation, we use R (R Core Team 2022) for all data cleaning and analysis and use R packages *tidyverse* (Wickham et al. 2019), *janitor* (Firke 2023), *tidyr* (Wickham, Vaughan, and Girlich 2023), *dplyr* (Wickham et al. 2023), and *ggplot* (Wickham 2016), to create the figures.

2 Data

2.1 Data Source

The data used in this paper was obtained from Open Data Toronto (Gelfand 2022), which is an official open source of data provided by divisions and agencies throughout Toronto. The Toronto Police Service Public Safety Department collected data in 2014 regarding all Major Crime Indicators (MCI) occurrences. MCI includes crimes such as Assault (excluding sexual assault), Breaking and Entering, Auto Theft, and Theft Over (theft over \$5000).

2.2 Data Cleaning

Each entry in the dataset contains detailed information regarding the offense. This includes an event id, occurrence date, report date, report time, division ID, offense information, and location information. Through the data cleaning process, we chose to keep the following variables to guide this analysis: `occ_dow` (occurrence day of the week), `premises_type` (the premises type), and `offense` (the offense type). Whilst the original dataset takes into account MCI occurrences that were reported and not found, this analysis primarily focuses on reported occurrences that were found and verified by the Toronto Police.

2.3 Data Visualization

We are interested in exploring the total number of occurrences of each crime, what locations are more likely to have crimes as well as what days of the week crimes take place the most. First, Table 1 below shows the number of occurrences that took place for each MCI category.

Table 1: Number of Occurences per Crime

Crime Type	Number of Occurrences
Assault	16557
Auto Theft	3589
Break and Enter	7096
Robbery	3746
Theft Over	987

Looking at Table 1, the assault category has the most occurrences in Toronto with 16557 offenses in 2014. Breaking and entering is the second most common crime than robbery, auto theft, and theft over follow.

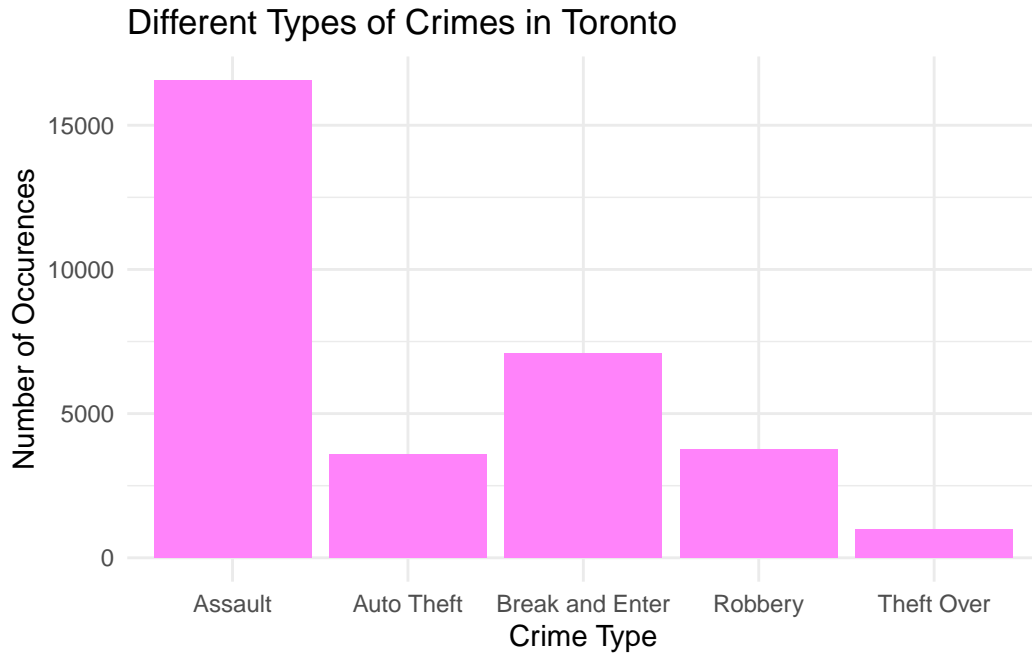


Figure 1: Occurences of the different types of crimes in Toronto.

Figure 1 is a graph displaying the values from Table 1. The graph demonstrates the drastic difference in the occurrence of assaults in Toronto in comparison to other MCI offenses. By exploring the other variables, we can further understand where and when these occurrences take place.

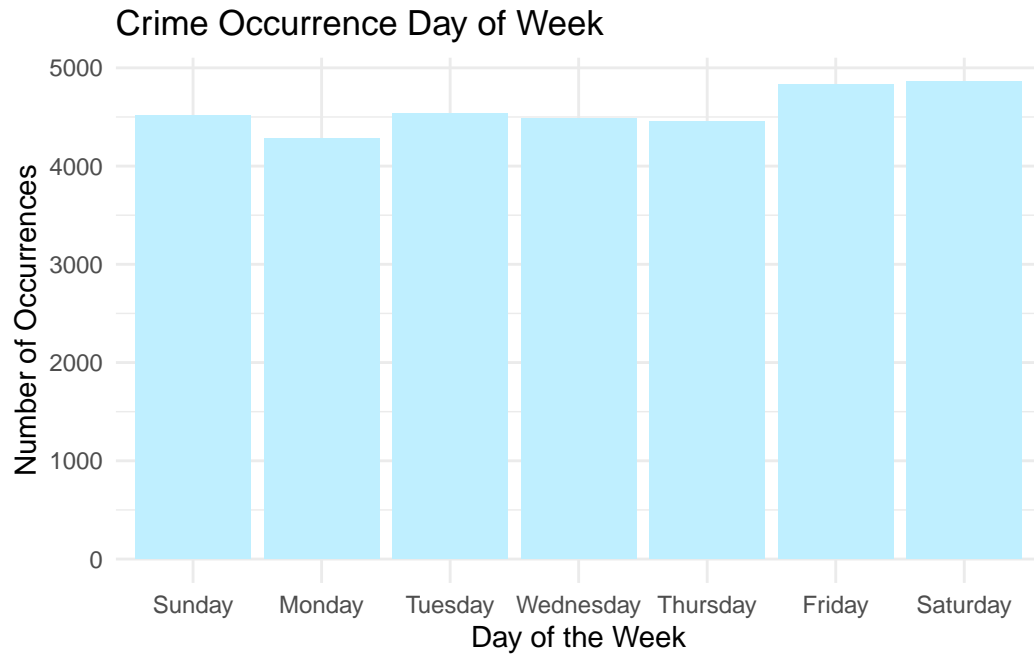


Figure 2: Days when crimes most likely took place.

Figure 2 displays the number of occurrences of crimes in general that took place on specific days. We find that MCI offenses were most common on Fridays and Saturdays with just under 5000 occurrences and Tuesdays and Sundays following with approximately 4500 occurrences. However, there is no drastic difference in occurrences which indicates that an MCI offense was likely to occur every single day in 2014.

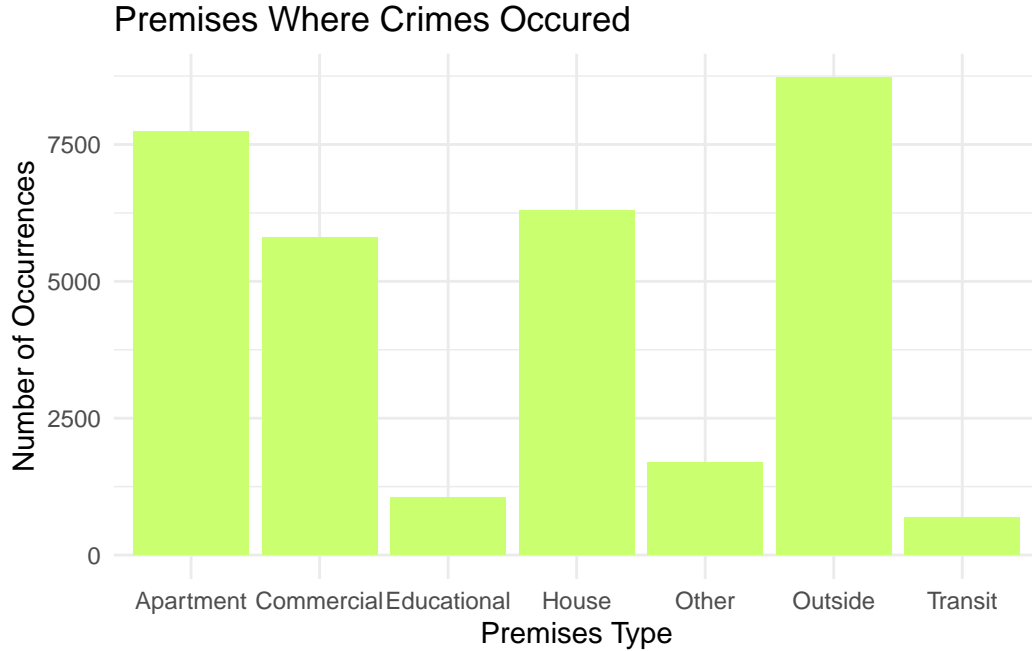


Figure 3: Premises types where crimes occurred.

Lastly, Figure 3 explores different types of premises in which crimes take place. Whilst the reports do include specific information in regard to the location where the incidents occurred, the premises types help categorize the locations into general groups which is useful for data collection purposes. The seven categories include apartment, commercial (including hotels, motels, B&B, etc.), educational, house, other (including corporate places, warehouses, etc.), outside, and transit. The major of MCI occurrences took place outside whilst the least occurrences took place in transit. In the results section of this paper, we visualize how the MCI category, day of the week, and premises type influence each other and go into further analysis in the discussion section.

3 Results



Figure 4: Comparing Crime Type, Premises Type, and Day of Week.

In this section, we display the relationship between MCI occurrence type, day of the week, and premises type. With Figure 4, we visualize the frequency of these incidences in a more specific context. For example, we see that there are more assault occurrences in educational premises on weekdays in comparison to weekends which makes sense as there are more students and faculty present on those days. Next, we will explore the results of each MCI occurrence in more depth in the discussion section.

4 Discussion

4.1 Assault Analysis

As seen in both Table 1 and Figure 1, assault is the most common MCI occurrence. Referring to Figure 4, there are consistent assaults in apartments every single day of the week. This makes sense because occupancy rates by household during this time consisted of 61.6% of people living in apartments (Toronto, n.d.). We find the same consistency when looking at assaults that took place outside and in houses. By looking at the break and enter data, we can infer that there can be a relationship with the house assaults as the points look similar. Interestingly, assaults were less common to take place in transit on weekends than on weekdays. This may be because fewer people are using transit as people won't need to go to work or school. Overall, assault is likely to be reported in all premise types every day of the week, thus making it the most common MCI offense.

4.2 Auto Theft Analysis

Auto theft is the second to least common MCI occurrence in 2014. Figure 4 shows a consistent trend of auto thefts outside every single day. This is due to the fact that cars are more accessible. However, it is also important to note the relationship between house auto thefts and house break and enters as the data suggests that these could have occurred simultaneously. Furthermore, auto thefts almost never occurred on educational premises.

4.3 Break and Enter Analysis

Break and Enter primarily occurred in apartments, commercial, and houses every day of the week. All other premises are less likely to have a break and enter occurrence because they're either an institution, transportation, or outdoors. It is also interesting to note the slight increase in break and enter occurrences in educational premises on Fridays.

4.4 Robbery Analysis

As the third most common form of crime, robbery almost always occurred outdoors. However, there is still a record of robbery in every premise type every single day. A notable observation is that the second most likely place for robbery to take place every day is commercial. This is due to the fact that the commercial premises are all locations that are easily accessible to the public.

4.5 Theft Over Analysis

Theft over is the least common MCI occurrence that happened in Toronto at that time. Table 1 shows that there was a total of 987 occurrences. Some notable observations seen on Figure 4 are that theft over most commonly took place in commercial premises as well as outdoors. Outdoor theft over was more likely to happen on Fridays and commercial theft over was likely to happen every day but more likely on Tuesdays, Thursdays, and Fridays.

4.6 Weaknesses and Next Steps

A weakness of this exploration is the dataset is 9 years old. It would be interesting to explore a more recent dataset and see if the patterns remain the same or if there are any major differences. Furthermore, it would be interesting to explore the neighborhoods in which these offenses occurred. This way we can truly develop an understanding of what scenarios would compromise one's safety.

References

- Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*.
- Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://CRAN.R-project.org/package=opendatatoronto>.
- R Core Team. 2022. *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. *Housing Occupancy Trends*. <https://www.toronto.ca/city-government/data-research-maps/research-reports/planning-development/housing-occupancy-trends/>.
- 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 Grolmund, 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*.
- Wickham, Hadley, Davis Vaughan, and Maximilian Girlich. 2023. *Tidyr: Tidy Messy Data*.