A Comprehensive Analysis of Toronto's Major Crime Indicators from 2014 to 2024*

Data from open data toronto displaying all 2014-2024 Major Crime Indicators from Toronto Police Service

Kuiyao Qiao

April 25, 2024

This study analyzes a comprehensive dataset of Major Crime Indicators (MCIs) reported in Toronto, Canada, from 2014 to 2024 to identify patterns, trends, and distributions of criminal activity across the city. Using statistical analysis and data visualization techniques, we examine the frequency and characteristics of major crime. Our findings reveal significant temporal and spatial patterns, including an overall increase in crime rates from 2014 to 2019 and a subsequent resurgence in recent years, with assault being the most prevalent crime category. The results highlight the importance of data-driven approaches to understanding urban crime dynamics and inform targeted strategies for law enforcement, resource allocation, and public safety initiatives in large urban centers.

Table of contents

*Code and data are available at: https://github.com/QPP123/Toronto-Crime-Stats-Over-Years

1 Introduction

Crime is a chronic concern in major cities around the world, affecting public safety, quality of life and socio-economic well-being. Effective law enforcement and crime prevention strategies rely on a thorough understanding of crime patterns, trends and distribution across different dimensions such as time, place and type of crime. The City of Toronto is the largest city in Canada, and its Major Crime Indicators (MCIs) data are collected and made publicly available by the Toronto Police Service. This dataset provides insight into the nature and dynamics of criminal activity in a large urban center over an extended period of time.

This paper provides a comprehensive analysis of Toronto's MCI data from 2014 to 2024(City of Toronto 2024), covering a decade of reported crimes such as assault, auto theft, burglary, robbery, and theft over $5,000. While previous studies have examined crime patterns in Toronto and other cities, there was a need for an up-to-date, in-depth analysis using the latest data and employing powerful statistical and visualization techniques. Such analysis can reveal emerging trends, seasonal variations, and geographic hotspots that can inform resource allocation, policy decisions, and community engagement efforts aimed at enhancing public safety.

We conducted an extensive exploratory analysis of the MCI dataset, focusing on the frequency, distribution, and temporal patterns of reported crime in the Toronto region. Using the R programming language and various data visualization libraries, the data were cleaned and processed to generate summary statistics and produce informative tables, charts, and maps. The results of the study reveal some clear trends, such as an overall increase in reported crime from 2014 to 2019, a possible sharp decline in 2020 due to the COVID-19 pandemic, and a rebound in crime in recent years. Assault was also identified as the most prevalent crime category, with peak crime times in the evening and at night, specific site types, and police zoning districts with higher concentrations of crime.

The insights derived from this analysis have important implications for law enforcement strategies, resource allocation and public safety initiatives in Toronto. By understanding temporal and spatial patterns of crime, authorities can develop targeted interventions, community policing efforts and crime prevention programs.

The data in this post was compiled using the open source statistical programming language R (R Core Team 2023), library arrow(Contributors 2023), dplyr(Wickham et al. 2022), readr(Wickham, Hester, and Bryan 2022), ggplot2(Wickham 2016), rmarkdown(Allaire et al. 2023), here(Müller 2020), knitr(Xie 2014), kableExtra(Zhu 2021), forcats(Wickham 2023), tidyr(Wickham and Henry 2023), gridExtra(Auguié 2023) and scales(Wickham and Seidel 2023)to clean up.

2 Data

2.1 Source Data

The "Major Crime Indicators" dataset was obtained from the City of Toronto Open Data Portal, a publicly accessible data portal provided by the City of Toronto. The data is collected and maintained by the Toronto Police Service (TPS). The TPS compilation of reported major crime incidents covers reported major crime incidents within the City of Toronto categorized as assaults, break-ins, auto thefts, robberies, and thefts over $5,000 (excluding sexual assaults). The data is provided at the crime level, meaning that one incident may have multiple associated rows, depending on the MCIs used to categorize the incident. The purpose of the data is to provide an accessible and transparent view of crime trends in the city to inform public safety strategies and community awareness.

2.2 Survey Methodology

The Major Crime Indicators dataset is not based on a survey but on administrative data collected by the Toronto Police Service (TPS). Administrative data refers to information collected routinely as part of an organization's operations, such as reporting and recording crime incidents. When a crime is reported to the police, the responding officer(s) gather information about the incident, such as the type of crime, location, date, and time, which is then recorded in the TPS's incident reporting system. This information is later processed and categorized according to standardized crime definitions and classifications, such as the Uniform Crime Reporting (UCR) system. The resulting dataset provides a comprehensive record of reported crimes in Toronto, allowing for analyzing crime patterns and trends over time. However, it is essential to note that this data collection method differs from a survey, where data is collected through questionnaires or interviews with a sample of individuals. The Major Crime Indicators dataset relies on the systematic and consistent recording of crime incidents by the TPS rather than the direct responses of individuals.

2.3 Data overview

This dataset contains Toronto Major Crime Incident (MCI) reports from 2014 to 2024, including detailed records of more than 384,000 crime incidents, with 2024 data available as of March 31st. Each record contains data fields for the year, month, day, and hour the crime was reported and occurred, the police precinct, location type, and premises type involved in the crime, and the specific category of MCI. The data includes the following key variables:

REPORT_YEAR, REPORT_MONTH, REPORT_DAY: time the crime was reported

OCC_YEAR, OCC_MONTH, OCC_DAY, OCC_HOUR: time of the crime

LOCATION_TYPE, PREMISES_TYPE: nature of the place where the crime occurred

MCI_CATEGORY: the Major Crime Indicator (MCI) category of the offence

DIVISION: police precinct where the crime was reported

These variables can be summarized into three categories:

Time-related variables (REPORT_YEAR, OCC_YEAR, etc.): Numeric.

Location-related variables (DIVISION, LOCATION_TYPE, etc.): Categorical.

Crime type variables (MCI_CATEGORY): Categorical.

OFFENCE could be simplified and divided into five kinds of MCI_CATEGORY.Table 1 Table 2

Table 1: Distribution of Offences by Major Crime Category

MCI_CATEGORY	OFFENCE	Count
Assault	Assault	139442
Assault	Assault With Weapon	34719
Assault	Assault Bodily Harm	8806
Assault	Assault Peace Officer	6469
Assault	Assault - Resist/ Prevent Seiz	3466
Assault	Aggravated Assault	2924
Assault	Discharge Firearm With Intent	2360
Assault	Discharge Firearm - Recklessly	1707
Assault	Pointing A Firearm	1383
Assault	Assault Peace Officer Wpn/Cbh	758
Assault	Administering Noxious Thing	631
Assault	Assault - Force/Thrt/Impede	435
Assault	Use Firearm / Immit Commit Off	216
Assault	Disarming Peace/Public Officer	133
Assault	Crim Negligence Bodily Harm	131

Table 1: Distribution of Offences by Major Crime Category

MCI_CATEGORY	OFFENCE	Count
Assault	Air Gun Or Pistol: Bodily Harm	32
Assault	Aggravated Aslt Peace Officer	28
Assault	Aggravated Assault Avails Pros	19
Assault	Unlawfully Causing Bodily Harm	15
Assault	Set/Place Trap/Intend Death/Bh	2
Assault	Traps Likely Cause Bodily Harm	2
Assault	Hoax Terrorism Causing Bodily	1
Auto Theft	Theft Of Motor Vehicle	61216
Break and Enter	B&E	60678
Break and Enter	B&E W'Intent	9033
Break and Enter	Unlawfully In Dwelling-House	2287
Break and Enter	B&E Out	117
Break and Enter	B&E - To Steal Firearm	16
Break and Enter	B&E - M/Veh To Steal Firearm	13
Robbery	Robbery - Mugging	9151
Robbery	Robbery With Weapon	6351
Robbery	Robbery - Other	5768
Robbery	Robbery - Business	5341
Robbery	Robbery - Swarming	2572
Robbery	Robbery - Home Invasion	1370
Robbery	Robbery - Vehicle Jacking	1320
Robbery	Robbery - Purse Snatch	1139
Robbery	Robbery - Financial Institute	996
Robbery	Robbery - Delivery Person	307
Robbery	Robbery - Taxi	267
Robbery	Robbery - Atm	119
Robbery	Robbery - Armoured Car	53
Robbery	Robbery To Steal Firearm	2
Theft Over	Theft Over	7222
Theft Over	Theft From Motor Vehicle Over	3161
Theft Over	Theft From Mail / Bag / Key	1496
Theft Over	Theft Over - Shoplifting	631
Theft Over	Theft Over - Bicycle	182
Theft Over	Theft Over - Distraction	152
Theft Over	Theft - Misapprop Funds Over	39
Theft Over	Theft Of Utilities Over	9

Distribution of Offences by Major Crime Category

Table 2: Number of Crimes for MCI Category

MCI Category	Number of Crimes
Assault	203679
Break and Enter	72144
Auto Theft	61216
Robbery	34756
Theft Over	12892

PREMISES_TYPE(Table 3):,Apartment,Commercial,Educational,House,Other,Outside and Transit.

Table 3: Number of Crimes for Premises Type

Premises Type	Number of Crimes
Outside	103168
Apartment	90580
Commercial	76582
House	70117
Other	23012
Transit	11587
Educational	9641

REPORT_YEAR and OCC_YEAR: represent the number of reported crimes from 2014 through 2024, and crimes that occurred from 2013 through 2024.(Table 4)

Table 4: Annual Reported Crimes from 2014 to 2024

Year	Number_of_Crimes
2014	32461
2015	32884
2016	33532
2017	35140
2018	37357
2019	40124
2020	35179
2021	35133
2022	41739
2023	49348
2024	11790

Data for 2024 are available through March 31st.

This dataset provides valuable information on the incidence and distribution of significant crimes in Toronto over ten years. The dataset allows for analyzing crime trends, seasonal patterns, and the relationship between crime and location type. This information is helpful for law enforcement resource allocation, crime prevention strategies and public safety policy decisions.

##Data Limitations

The dataset has some limitations that should be considered when interpreting the results: Unfounded incidents: The data does not include occurrences deemed unfounded through police investigation, meaning that the reported offence did not occur or was not attempted. Location verification: Some incidents may have unverified locations or coordinates outside the City of Toronto. Crime categorization: The MCI categories are broad and may not provide detailed insight into specific types of crimes within each category. Reporting bias: The data only includes reported crimes, which may not represent the true prevalence of crime due to underreporting. Partial data for 2024: The dataset only contains data up to March 31st, 2024, which limits the analysis for that year.

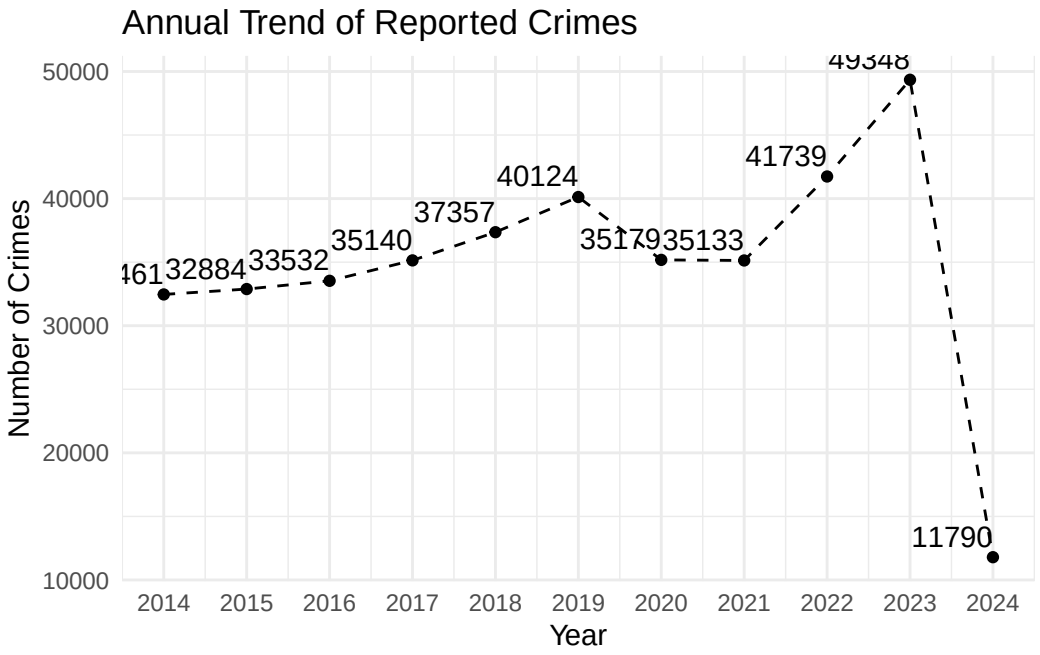
2.4 Data cleaning

During the cleaning process, several steps were undertaken to ensure the dataset's utility for analysis while maintaining the integrity of the original data. Fields that did not contribute to the crime analysis, such as EVENT_UNIQUE_ID, various detailed date fields (REPORT_DATE, OCC_DATE, REPORT_DOY, OCC_DOY), day of the week information, and redundant neighbourhood structure identifiers (HOOD_158, NEIGHBOURHOOD_158, HOOD_140, NEIGHBOURHOOD_140), were removed. This focused on the most impactful and relevant data fields, such as crime type, occurrence time, and location specifics. Handling missing data was a crucial aspect of the cleaning process. The script assessed the presence of null or missing values across critical variables and applied appropriate methods to handle these cases, ensuring no significant loss of data integrity. No imputation was deemed necessary; instead, records with incomplete information critical to the analysis were excluded to preserve the accuracy of crime trend interpretations. No new variables were constructed during the cleaning process, as the emphasis was on analyzing the data in its most straightforward and interpretable form. This approach maintains the authenticity of the reported data, focusing analysis on the core elements provided by the Toronto Police Service. This cleaning protocol streamlined the dataset for effective analysis and ensured that the focus remained sharply on the variables most likely to yield meaningful insights into crime patterns within Toronto.

3 Results

The Results section provides an in-depth analysis of the Toronto Police Service's Major Crime Indicators (MCIs) for the ten-year period from 2014 to 2024. The data was scrutinized to identify crime patterns, distribution and trends. Below, we detail these findings through summary statistics, tables, charts and illustrative text.

Annual Number of Reported Crimes and Growth Rate (Table 4, Table 5, and Figure 1): The number of reported crimes increased steadily from 2014 to 2019, with a significant decrease in 2020 (-12.32% growth rate), which may be attributed to the COVID-19 pandemic. Subsequently, the number of crimes rises sharply in 2022 and 2023, with growth rates of 18.80% and 18.23%, respectively. Figure 1 visually illustrates this trend, showing a clear upward trajectory in recent years. Note that data for 2024 is only available up to March 31, which explains the lower number of crimes and negative growth rate in that year.



(a) Data for 2024 are available through March 31st.

Figure 1: Total Reported Crimes Per Year

Table 5: Annual Growth Rate of Crimes

Table 5: Annual Growth Rate of Reported Crimes from 2014 to 2024

Year	Number_of_Crimes	Growth_Rate
2014	32461	0.0000000
2015	32884	1.3031022
2016	33532	1.9705632
2017	35140	4.7954193
2018	37357	6.3090495
2019	40124	7.4069117
2020	35179	-12.3242947
2021	35133	-0.1307598
2022	41739	18.8028349
2023	49348	18.2299528
2024	11790	-76.1084542

Data for 2024 are available through March 31st.

Monthly Crime Frequency Figure 2: shows the monthly distribution of crime for 2014, 2017, 2020 and 2023. The data shows a relatively even distribution of crime rates across the months, with a slight increase during the summer months (June through August). This seasonal trend appears to be consistent across the years studied.

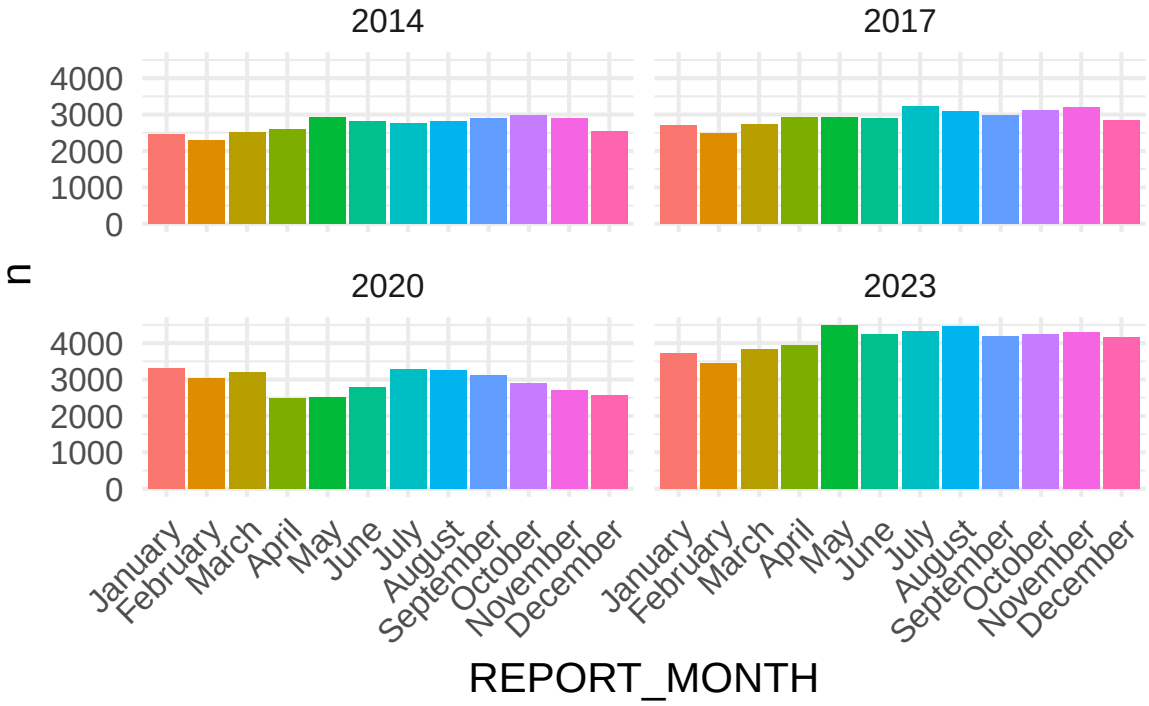


Figure 2: Monthly Crime Frequencies for 2014

Frequency of crime by time period Figure 3: Crime rates were highest between 6 pm and midnight, with the highest rate (7.0%) occurring at 9 pm. The highest rate (7.0%) occurred during the early morning hours (2 am to 1 pm). The frequency of crime is lowest in the early morning hours (2 am2 am to 6 am). This indicates that criminal activity is more likely to occur in the evening and nighttime hours.

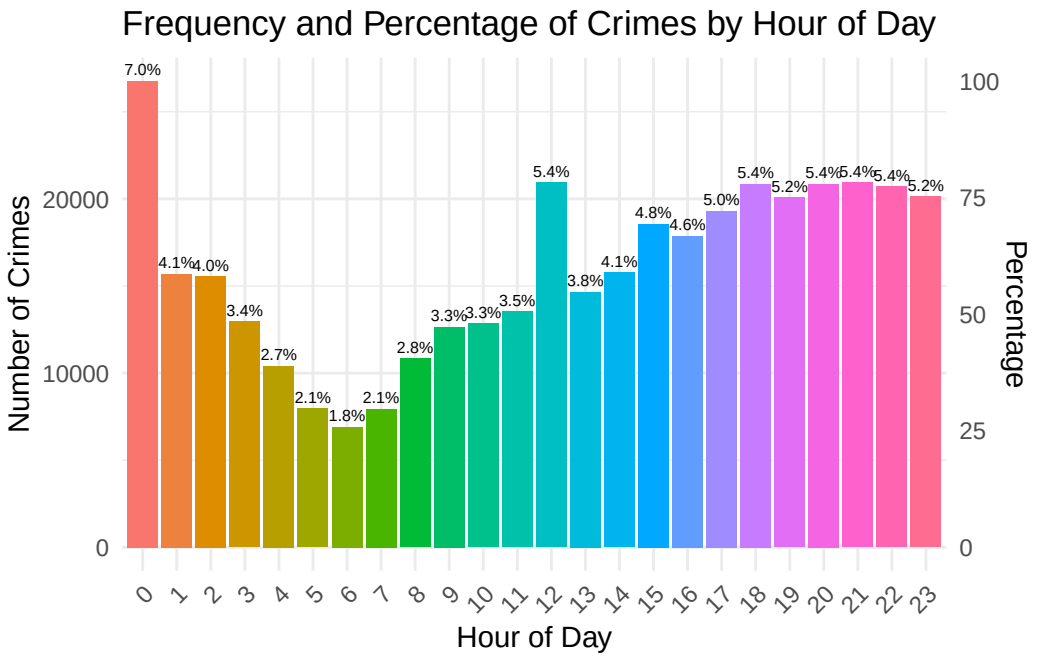


Figure 3: Frequency of Crimes by hour of the day

Distribution of Crimes by Major Crime Type (Table 1 and Figure 4): The data show that assault was the most common type of crime, accounting for 52.9% of all reported crimes. This was followed by auto theft (18.8%), breaking and entering (15.9%), robbery (9%), and theft over $5,000 (3.4%). Within the assault category, assault (139,442) and assault with a weapon (34,719) were the most common.

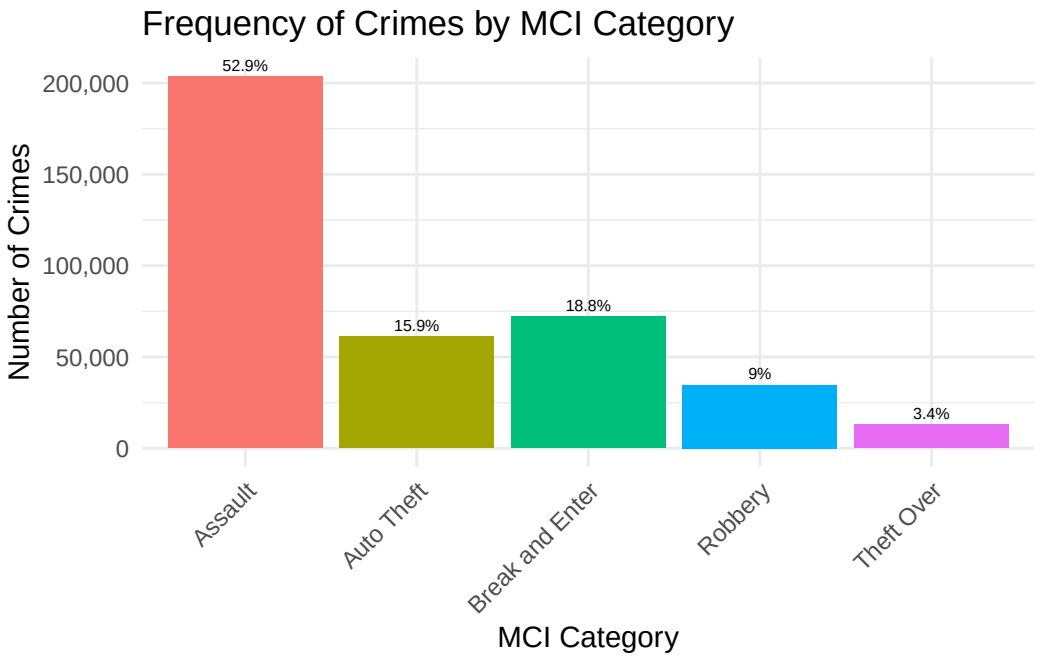


Figure 4: Frequency of Crimes by MCI Category

Number of Crimes by Offense Classification (Table 2): This table summarizes the total number of crimes in each major category. The highest number of crimes was assault, with 203,679, while the lowest number was theft, with 12,892.

Number of crimes by type of place (Table 3 and Figure 5): The type of place where the most criminal cases occurred was outdoors (103,168), followed by apartments (90,580) and businesses (76,582). Educational institutions had the lowest number of reported cases, 9,641. Table 5 shows the frequency distribution by type of establishment, with apartments (26.82%), commercial properties (23.55%), and residences (18.23%) having the highest incidence of crime.

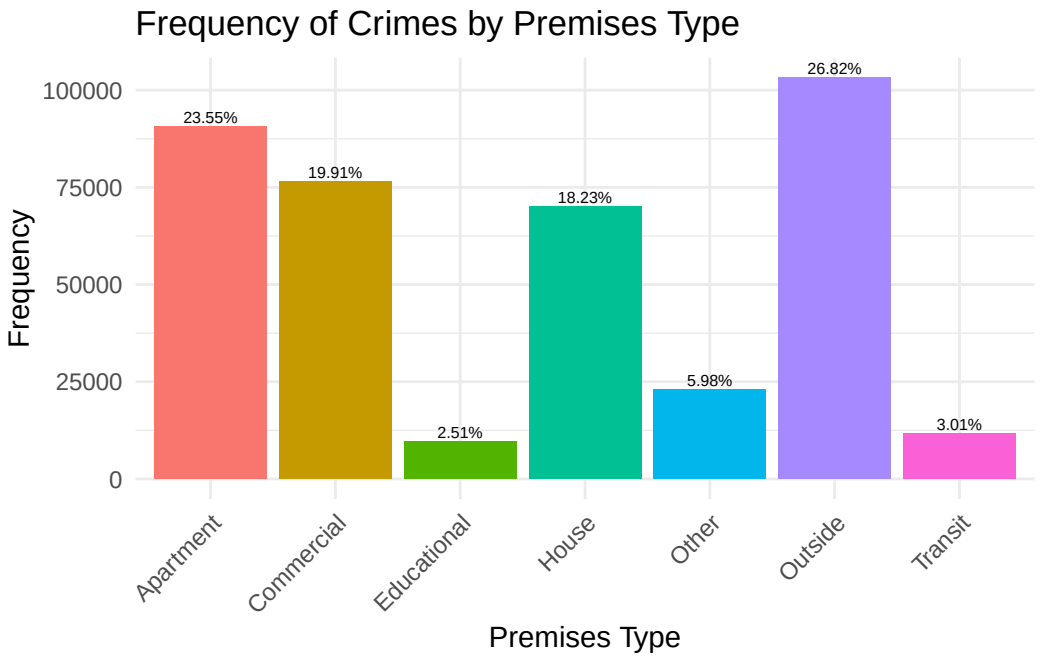


Figure 5: Frequency of Crimes by Premises Type

Frequency of crime by police Division Figure 6: D14 Precinct had the highest crime rate at 8.2%, followed by D43 Precinct (7.9%) and D51 Precinct (7.3%).D53 Precinct had the lowest crime rate at 1.0%, followed closely by D54 Precinct (1.6%).D54 Precinct had the lowest crime rate at 1.0%, followed closely by D54 Precinct (1.6%).D54 Precinct had the lowest crime rate at 1.0%, followed closely by D54 Precinct (1.6%). This information helps to identify high-crime areas and allocate police resources accordingly.

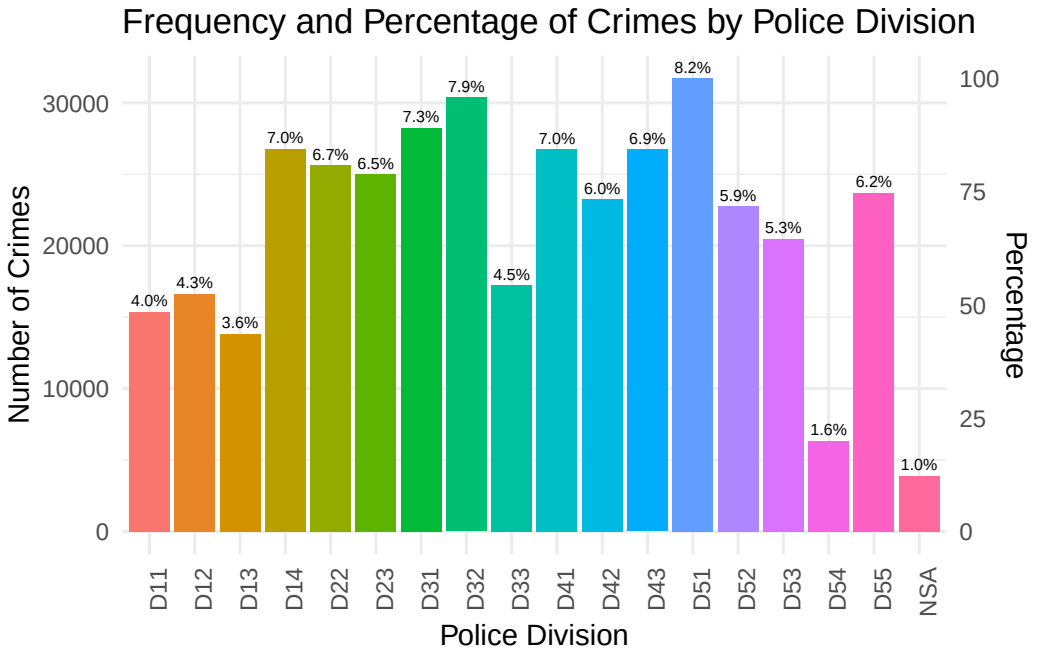


Figure 6: Frequency of Crimes by Division

These results provide valuable insight into the patterns and trends of major crime in Toronto over the past decade. The data and visualizations presented provide a comprehensive overview of crime trends in Toronto, highlighting key aspects such as the increase in crime reports over the years, seasonal variations, peak crime times, and the types of venues most affected. The data emphasize the prevalence of "strikes" and the concentration of crime in specific locations and times. This information can be used to develop crime prevention strategies, resource allocation and public safety initiatives.

4 Discussion

4.1 Summary

In this study, a comprehensive analysis of key crime indicators for Toronto from 2014 to 2024 was conducted using data from the Toronto Police Service. The study analyzes the frequency and distribution of crime across a number of dimensions, including crime type, place type, time and location. The data is visualized through tables, graphs and charts to identify patterns and trends in criminal activity over the past decade.

4.2 Insight into crime patterns:

A key finding from the analysis is that assault is the most common type of major crime in Toronto, accounting for more than half of all reported incidents. This finding highlights the need for targeted interventions and resources to address violent crime and ensure public safety. Another important revelation is the concentration of crime in specific locations, such as condominiums, commercial properties and outdoor spaces. This knowledge can help law enforcement agencies and policymakers to allocate resources more efficiently and develop location-specific crime prevention strategies.

4.3 Time trends in crime

The study also revealed temporal patterns of criminal activity in Toronto. The data shows consistent seasonal trends, with slightly higher crime rates during the summer months. In addition, the hourly distribution of crime shows higher rates in the evening and at night, particularly between 6 p.m. and midnight. These findings suggest that crime prevention efforts should be strengthened during these high-risk hours in order to curb criminal activities and improve public safety.

4.4 Limitations and shortcomings

While this study provides valuable insights into crime patterns in Toronto, its limitations must be recognized. First, the exclusion of unfounded crime reports, while standard in crime data analysis, may underestimate the true number of crime reports. The analysis relies solely on reported crime, which may not provide a comprehensive picture of criminal activity due to underreporting. In addition, the broad categorization of crime types and place types may mask more nuanced patterns within each category. The study also does not consider potential confounding factors, such as socioeconomic conditions or population density, which may influence crime rates in different areas.

##Future directions Based on this study, future research could explore the potential factors that contribute to the observed crime patterns. This may require a combination of demographic, socioeconomic, and environmental data to identify potential correlates of criminal activity. Additionally, qualitative research (e.g., interviews with law enforcement officers and community members) could provide valuable contextual information and insight into the drivers of crime in specific locations. Future work could also focus on assessing the effectiveness of existing crime prevention strategies and identifying best practices for reducing crime rates in high-risk areas. Longitudinal studies are necessary to assess the effectiveness of specific crime prevention strategies affected by our findings. Such studies could track the impact of increased police patrols during identified peak crime periods and changes in urban planning on crime rates.

In summary, this study provides a comprehensive overview of key crime indicators in Toronto and identifies major patterns and trends in criminal activity over the past decade. The findings emphasize the importance of a data-driven approach to crime prevention and highlight the need for targeted interventions in high-risk areas and time periods. By exploring the complex factors behind crime patterns on the basis of this research, policymakers and law enforcement agencies can develop more effective strategies to promote public safety and reduce the impact of crime on communities.

References

Allaire, JJ, Yihui Xie, Jonathan McPherson, Javier Luraschi, Kevin Ushey, Aron Atkins, Hadley Wickham, Joe Cheng, Winston Chang, and Richard Iannone. 2023. *Rmarkdown: Dynamic Documents for r*. https://CRAN.R-project.org/package=rmarkdown.

Auguié, Baptiste. 2023. *gridExtra: Miscellaneous Functions for "Grid" Graphics*. https://CRAN.R-project.org/package=gridExtra.

City of Toronto. 2024. "Major Crime Indicators." City of Toronto Open Data Portal. 2024. https://open.toronto.ca/dataset/major-crime-indicators/.

Contributors, Apache Arrow. 2023. *Arrow: Integration to 'Apache Arrow'*. https://CRAN.R-project.org/package=arrow.

Müller, Kirill. 2020. *Here: A Simpler Way to Find Your Files*. https://CRAN.R-project.org/package=here.

R Core Team. 2023. *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.

———. 2023. *Forcats: Tools for Working with Categorical Variables (Factors)*. https://CRAN.R-project.org/package=forcats.

Wickham, Hadley, Romain Francois, Lionel Henry, and Kirill Muller. 2022. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Wickham, Hadley, and Lionel Henry. 2023. *Tidyr: Tidy Messy Data*. https://CRAN.R-project.org/package=tidyr.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2022. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.

Wickham, Hadley, and Dana Seidel. 2023. *Scales: Scale Functions for Visualization*. https://CRAN.R-project.org/package=scales.

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.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.