

Analyzing Crime Trends in Toronto: A Descriptive Study Using Police Annual Statistical Reports (2014-2022)

Zixi Song

2024-01-22

Table of Contents

1. [Introduction](#)
2. [Data](#)
 - [Source and Context](#)
 - [Data Overview](#)
 - [Descriptive Analysis](#)
3. [References](#)

Introduction

Toronto is the fourth largest city in North America and has a relatively low crime rate compared to other major cities, although crime rates have shown different trends in recent years. For example, in 2021, Toronto saw a decrease in some types of crime, such as break-ins, which may have been influenced by changes in residential occupancy patterns during the COVID-19 pandemic (Moreau, 2022). On the contrary, online fraud and cybercrime have increased, highlighting the constant evolution of criminal activity in the digital age (Statistics Canada, 2022). In addition, the number of shootings in 2019 hit a record high (Wikipedia, 2024). Recently, there has been a wave of highly publicized violent crimes, including mass shootings and subway stabbings, that have upset residents (Landau, 2022). The Toronto Police Service's data portal shows a 17.2 per cent increase in overall major crime in 2022, which includes increases in assaults, auto thefts, break-ins, robberies, sexual violence and thefts. Despite the rise in crime, the homicide rate is down 16% from 2021 (Landau, 2022). These trends provide important context for the current study, underscoring the need for ongoing and comprehensive analysis of crime data to inform effective policing strategies and public safety policies in Toronto. In

contemporary urban security and law enforcement, the role of data-driven analytics is critical. The research presented in this paper draws on an extended data set provided by the Annual Police Statistical Report - Reporting Crime, from the Open Data Toronto Portal (City of Toronto, 2024). The dataset is a comprehensive crime-related statistics database, an initiative of the Toronto Police Service (TPS), covering the number of different crime types in different areas of the City of Toronto from 2014 to 2022. This study will focus on descriptive analysis and data visualization to provide a basis for subsequent data analysis and utilization. The main part of this paper is the Data section. This section begins by discussing the data sources used for analysis and their background, as well as the tools used in the analysis process. Next, a statistical summary is used to give the reader an initial understanding of the data set. Finally, some data visualization methods are used to more intuitively present the data and trends within it.

Data

Source and Context

The Data used in this paper - “Police Annual Statistical Report - Reported Crimes” - is from the Open Data Toronto Portal through the R library `opendatatoronto` (Gelfand 2022). This dataset contains the Annual Statistical Report (ASR) of the Toronto Police Service (TPS) from 2014 to 2022 (Toronto Police Service, 2024). The ASR is a comprehensive overview of police-related statistics, including reported crimes, crime victims, personal searches, firearms, traffic accidents, personnel, budgets, communications, public complaints, regulatory interactions, and other administrative information. The crime data comes from a variety of sources, such as the TPS Records Management System, the TPS Communications System, the Office of the Director of Independent Police Review and the Attorney General’s Department, among others. It aims to improve understanding of policing, increase transparency and accountability, and support evidence-based decision-making to safeguard public safety in Toronto. ASR data is hosted on the TPS Public Safety Data Portal, a platform for sharing and exploring police open data (Toronto Police Service, 2024). ASR data is also available on the City of Toronto Open Data Portal, an open data repository from various city departments and agencies (City of Toronto, 2024; <https://open.toronto.ca/dataset/police-annual-statistical-report-reported-crimes/>). The ASR data is updated annually, usually in the first quarter of the following year. ASR data is revised and corrected as new information becomes available or errors are discovered. Data in this paper were collected and analyzed using statistical analysis tool R (R Core Team 2024). The following libraries and packages are used in the analysis: `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016).

Table 1: Top 5 rows as an example

_id	REPORT_YEAR	DIVISION	CATEGORY	SUBTYPE	COUNT_	COUNT_CLEARED
1	2014	D11	Crimes Against the Person	Other	22	9
2	2014	D11	Crimes Against Property	Theft Over \$5000	1	1
3	2014	D11	Crimes Against the Person	Other	1	1
4	2014	D11	Crimes Against the Person	Robbery-Financial	1	1
5	2014	D11	Crimes Against Property	Break & Enter-House	23	13

Data Overview

The dataset “Police Annual Statistical Report - Reported Crimes” contains 32000 rows. **Table 1** shows the first five lines of data from the dataset as an example. **Table 2** explains each variable in the data set. The dataset contains six valid variables and a list of indexes. It’s worth explaining the difference between COUNT_ and COUNT_CLEARED, the former being the total number of reports, the latter indicates the total number of crimes that have been resolved through the arrest of a suspect, charges being laid, or the case being otherwise closed. It reflects the number of reported crimes for which there has been a conclusion that satisfies legal or police requirements. Due to the small amount of data, no problematic data (such as missing values, incorrect data, etc.) was found after manual review, so the data cleaning step was skipped in this study.

Table 2: The explanation of the dataset attributions

Attributions	Description
REPORT_YEAR	The year the crime happened, ranging from 2014 to 2022.
DIVISION	The regional divisions of the Toronto Area.
CATEGORY	A primary classification of crimes consisting of 6 distinct categories.
SUBTYPE	A secondary classification of crimes consisting of 15 distinct sub-categories.
COUNT_	Total number of this sub-type crimes in that year.
COUNT_CLEARED	Total number of this sub-type crimes in that year that are identified as cleared.

Descriptive Analysis

Descriptive analysis involves summarizing and interpreting data to extract meaningful patterns and trends. It employs statistical techniques to present a clear, concise picture of data sets, crucial for understanding underlying structures without making inferential or predictive assumptions (Dowdy, Wearden & Chilko, 2011). In this part, statistical summary and data visualization will be used to interpret the data.

Fig 1 shows the total number of reported crimes and cleared crimes over years from 2014 to 2022. In the figure, we can see that the total number of reported crimes rises from 2014 till 2019, reaching the peak of close to 150000. After that, it dropped sharply these years, falling to 81557 in 2022. This is an interesting point to explore further: why the crime rate drops so quickly in the recent years?

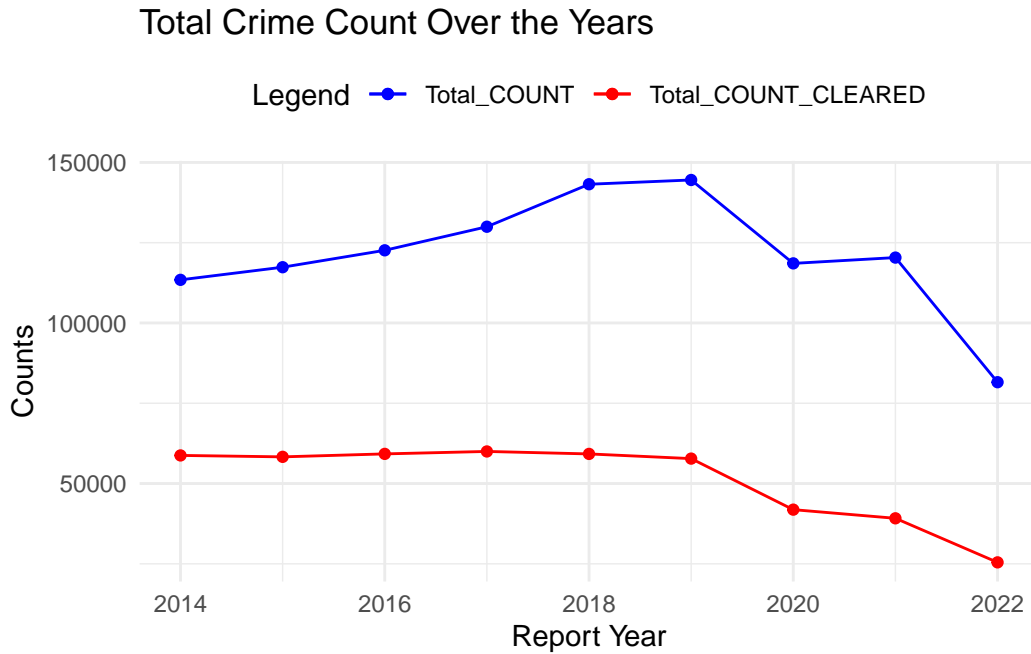


Figure 1: The total number of reported crimes and cleared crimes over years from 2014 to 2022.

Fig 2 shows the percentage of different of categories of crimes from 2014 to 2022. In the figure, we can see that the crimes against property occupied the largest share of total crimes, nearly 64%. The second major crime type is against the person, which is 21%. Then comes other criminal code violation (11%). The top three accounted 96% of the total.

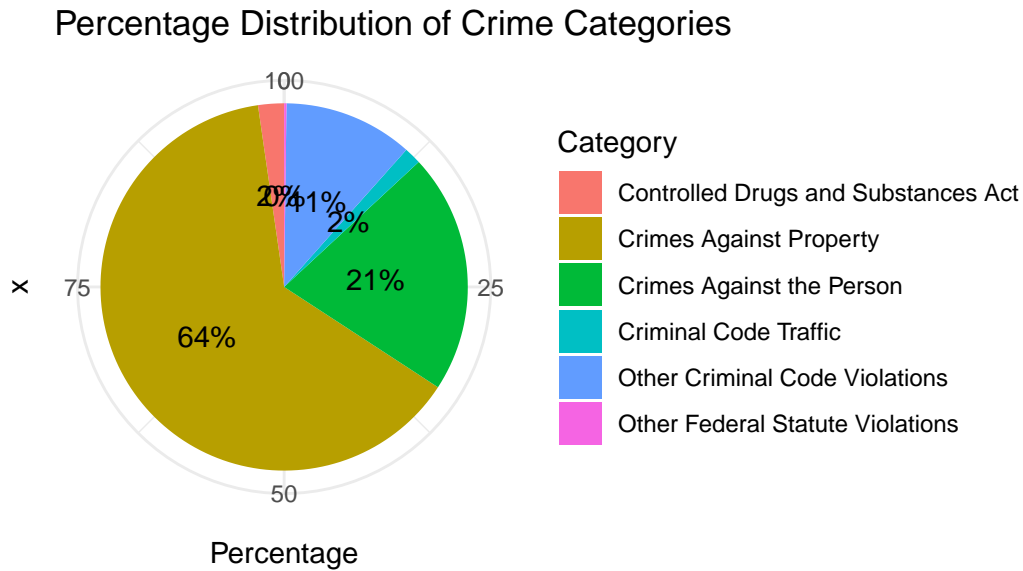


Figure 2: The percentage of different of categories of crimes from 2014 to 2022.

References

- City of Toronto. (2024). Open Data. Retrieved January 25, 2024, from <https://www.toronto.ca/city-government/data-research-maps/open-data/>
- Dowdy, S., Wearden, S., & Chilko, D. (2011). Statistics for research. John Wiley & Sons.
- Gelfand, Sharla. 2022. Opendatatoronto: Access the City of Toronto Open Data Portal. <https://CRAN.R-project.org/package=opendatatoronto>.
- Landau, J. (2022, December 21). Toronto residents questioning the city’s safety amid wave of shocking crimes. BlogTO. Retrieved January 25, 2024, from <https://www.blogto.com/city/2022/12/crime-wave-toronto-questioning-safety/>
- Moreau, G. (2022, August 2). Police-reported crime statistics in Canada, 2021. Statistics Canada. Retrieved January 25, 2024, from <https://www150.statcan.gc.ca/n1/pub/85-002-x/2022001/article/00013-eng.htm>
- R Core Team. 2024. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Statistics Canada. (2022, October 18). Impact of cybercrime on Canadian businesses, 2021. Retrieved January 25, 2024, from <https://www150.statcan.gc.ca/n1/daily-quotidien/221018/dq221018b-eng.htm>
- Toronto Police Service. (2024). Retrieved January 25, 2024, from <https://www.tps.ca/>
- 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 Golemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. <https://doi.org/10.21105/joss.01686>. Wikipedia contributors. (2024). Crime in Toronto. In Wikipedia. Retrieved January 25, 2024, from https://en.wikipedia.org/wiki/Crime_in_Toronto