Beyond the Myth: A Comparative Analysis of Media Crime Reports and Actual Crime Statistics in Toronto*

Kelly Lyu

2024-01-25

This report offers crucial insights into Toronto's crime trends, which are indispensable for the daily decision-making processes of its residents. By examining crime data collected by Open Data Toronto, it accomplishes two primary objectives: firstly, it evaluates changes in police efficiency regarding crime resolution and analyzes crime patterns, including locations and types of crimes; secondly, it contrasts these actual crime statistics with how they are reported in the media. The findings indicate a decline in police effectiveness from 2014 to 2022, highlighting theft under $5000 as the most common crime, crime predominantly occurring in downtown Toronto. In contrast, media coverage implies an increase in police efficiency and inaccurately focuses on murder as the leading crime, wrongly implying that North York is as affected by crime as downtown.

Table of contents

*Code and data are available at: https://github.com/kelly-Lyu/Reported_Crime.git.

1. Introduction

Every Halloween, the phenomenon of 'Halloween Sadism'—characterized by malevolent acts against children, such as the adulteration of candy or deliberate infliction of harm—emerges prominently in media narratives (Best 2008). These narratives often emphasize the perceived dangers posed by strangers to children during this holiday growing up; this stark contrast between ominous warnings and the absence of such incidents in my own experiences prompted a deeper investigation. The findings were enlightening: the ostensibly pervasive threat of Halloween Sadism is largely a mythological construct supported by scant empirical evidence (Best 2008). The few existing cases show only a tenuous connection to this supposed phenomenon, casting doubt on the narrative perpetuated by the media.

This disparity between media portrayal and reality intrigued a broader inquiry: How accurately does the media depict criminal activity? With many relying on media as their primary source of crime-related knowledge, the impact of these portrayals on public perception and conduct is profound (Stroman and Seltzer 1985). Currently, there is no data on the detailed difference between media portray and real crime statistics, especially in Toronto. Employing the Open Data Toronto dataset, this paper analyzes reported crime statistics from 2014 to 2022, contrasting them with media representations to explain the state of crime in Toronto.

The remaining part of the paper proceeds as follows: data, graph, discussion, and conclusion. The data section begins by focusing on the dataset sourced from the City of Toronto's Open Data Toronto Library, detailing the explanation of the variable in the dataset, as well as the data cleaning and analysis processes undertaken. The graph section gives an illustration of crime types, their geographical distribution, and trends in crime resolution rates, supported by visual graphs. Notably, it was found that police efficiency in solving crime decreased, and the most common crime type is theft under $5000; crime primarily occurred in Toronto's downtown area. The Discussion section discusses the discrepancies between the crime data and media reports. Finally, the paper outlines its key findings, underscoring the gap between media representations and actual crime statistics in Toronto, stating the limitations of these studies, and advising future focus.

2. Data

Data Source and Collection

We utilized a detailed dataset from Toronto Police Services featuring comprehensive crime records (Gelfand 2024). This dataset includes every reported crime in Toronto, with 33,343 entries across key columns like ID, division, subtype, category, report year, crime count, and cleared count. Its latest update was on January 14, 2024, and we accessed it through the City of Toronto's Open Portal on January 18, 2024.

The dataset's compliance with the Municipal Freedom of Information and Protection of Privacy Act (MFIPPA) aligns it with established legal standards. This legal compliance confirms that the dataset is managed and released under recognized regulations, concentrating on data accuracy and access - key elements that indicate its adherence to legal norms and enhance its trustworthiness. As this dataset separates theft into two categories by theft money amount, every crime is categorized precisely, minimizing subjective interpretation.

Compared with alternative datasets, this one stands out for its complete coverage of crime-related data, especially its inclusion of cleared crime counts, which is pivotal for our analysis. It is the most recently updated dataset with a 100% quality score (Gelfand 2024). Nonetheless, it's essential to consider potential biases: the dataset might include crimes reported in Quebec by Toronto residents because instead of specifying the location of the crime, it sets the location of the police division, which could distort the data. Furthermore, issues like under reporting in sensitive cases such as sexual assault need to be considered, which causes the dataset to not encompass all instances of criminal activity (Yung 2015).

Data collection and analysis were executed using the statistical programming software R (R Core Team 2020), bolstered by the indispensable assistance of tidyverse (Wickham et al. 2019), ggplot2 (Wickham et al. 2021), lubridate (Grolemund and Wickham 2011), knitr (Xie 2021) and kableExtra (Zhu et al. 2021). A more exhaustive cleaning and analytical procedures will be supplied in subsequent sections of this paper.

Data Cleaning

Upon importing the dataset into our analytical environment, we clean start by selecting key columns pertinent to our research: division, subtype, report year, count cleared, and crime count. Next, we introduced two new columns to heighten our analysis. The first one, 'unsolved_crime,' quantifies the total unresolved crimes for each data entry, calculated as the difference between the total reported crimes and those cleared. The second, 'clearance_rate,' provides a crucial metric, representing the proportion of cleared crimes relative to the total crime count. Both these columns were transformed into integer values to ensure data clarity and precision. Then, We converted all column names to lowercase to improve their readability. These enhancements were integrated into the dataset, creating a refined version aptly labeled 'clean.' Table 1 below shows the first six rows of the clean data. This file serves as a comprehensive record for future studies, embodying a clear and well-structured representation of Toronto's crime data.

Table 1: Preview of First 6 Rows of Cleaned Data

division	subtype	report_year	count_cleared	count_	clearance_rate	unsolved_crime
D11	Other	2014	9	22	0	13
D11	Theft Over $5000	2014	1	1	1	0
D11	Other	2014	1	1	1	0
D11	Robbery-Financial	2014	1	1	1	0
D11	Break & Enter-House	2014	13	23	0	10
D11	Theft Over $5000	2014	1	1	1	0

Table 2: Yearly Summary of Clearance Rates

Year	Total	Average Clearance Rate
2014	882	0.2277893
2015	844	0.2202505
2016	796	0.2061108
2017	758	0.1979629
2018	668	0.1823642
2019	591	0.1633048
2020	527	0.1477432
2021	510	0.1444350
2022	477	0.1336883

Table 2 presents a 'Yearly Summary of Clearance Rates' spanning from 2014 to 2022, showing the summary statistics of the mutate variable. A significant observation from this data is the noticeable downward trend in the average clearance rates for reported crimes. In 2014, the average clearance rate was recorded at 0.228, which has markedly declined to 0.134 by 2022. This reduction in clearance rates is more effectively illustrated through visual representation.

3. Graph

Assessing the Effectiveness of Crime Solving

This analysis focuses on police effectiveness. This line graph figure outlines the average crime clearance rate from 2014 to 2022. With years marked on the x-axis and the clearance rate on the y-axis, it shows an initial clearance rate of just above 0.225 in 2014, followed by a consistent decline. The rate experienced a notable drop in 2017 and continued to decrease steadily, reaching just under 0.15 by 2022.

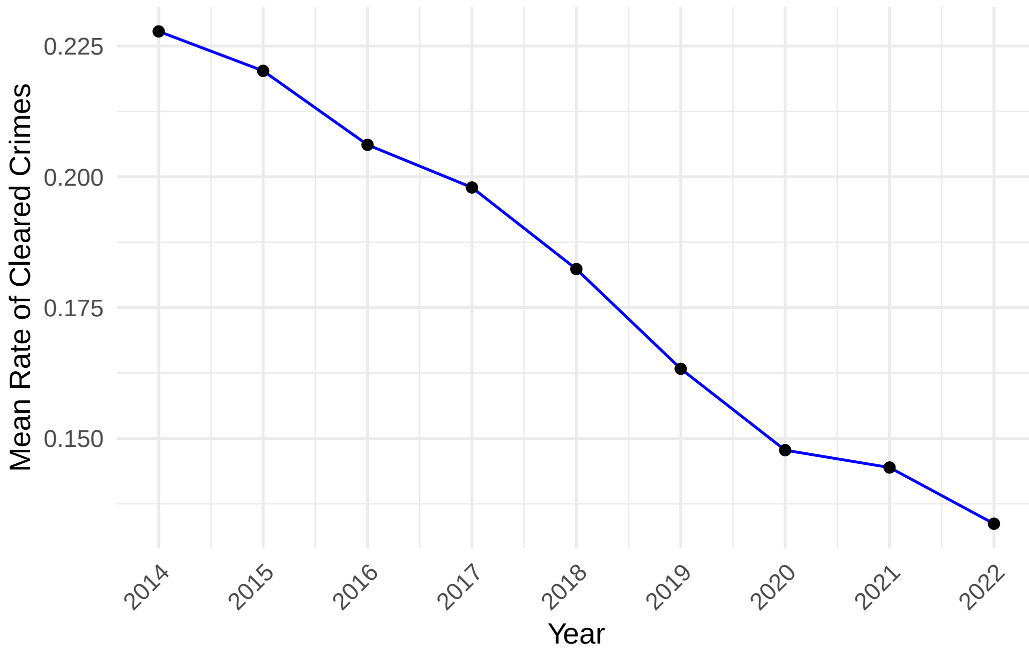

Figure 1: Mean Rate of Cleared Crimes Over Years

Discussion 1

This Figure 1 and the above table indicate a continuing decrease in police efficiency in solving crimes, with no immediate signs of improvement anticipated in the coming year. Conversely, research indicates that public perception of police effectiveness remains largely positive and may be affected by media. An analysis of leading Toronto media outlets, such as The Toronto Sun, The Toronto Star, and The Globe and Mail, reveals a notable trend. Despite statistics indicating a decrease in police efficiency, these publications increasingly portray police performance positively (Porter 1995). For instance, The Toronto Sun reported 125 positive references to police actions compared to only 13 negative mentions. These data clarify how police efficacy differs between the media portrayed and the actual statistics.

Categorizing Crime Types

In Figure 2, we present a point chart illustrating the variations in crime rates across different crime types over the course of several years. On the y-axis, various crime subtypes are listed, while the x-axis quantifies incident numbers. 'Theft Under $5000' consistently shows high frequency, highlighting its prevalence. On the other hand, 'Attempted Murder' and 'Robbery - Financial' are rare, as indicated by their short bars. Subtypes like Assault, Auto Theft, and Break-and-Enter occupy intermediate positions.

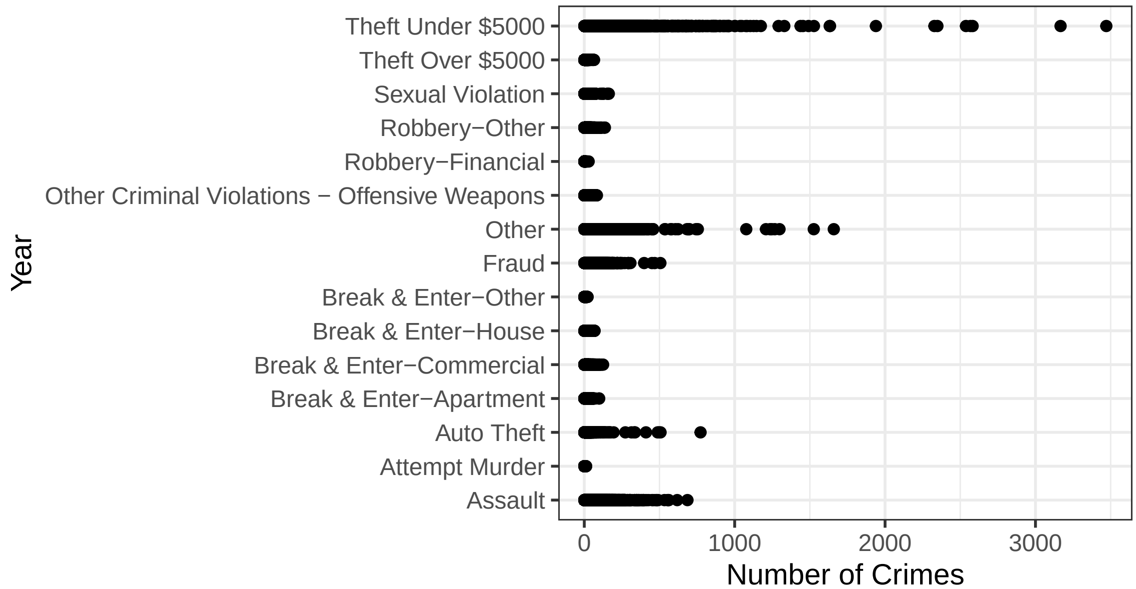

Figure 2: Number of Crimes Reported in Toronto by Subtype

Discussion 2

Public perception often diverges from the data-driven reality of crime, a point highlighted by recent studies (Siegel and McCormick 2019). High-profile crimes such as murder and aggravated assault, which dominate news cycles and capture public attention, are perceived to be more common than they are (Siegel and McCormick 2019). In reality, statistical data shows that these violent crimes occur less frequently than non-violent offenses, like theft under $5000, which predominate in crime reports. This contrast points to a significant insight: The enduring impactful crimes in society are often the more common, yet less sensational, offenses such as petty theft, despite the media's focus on more severe but rarer crimes.

Categorizing Reporting Divisions

Figure 3 presents a graph that tracks crime incidents across different divisions from 2014 to 2022. The x-axis is segmented with labels ranging from D11 to D53 and an NSA category to differentiate the police divisions, while the y-axis measures the reported number of crimes. Utilizing a stacked bar chart, this visualization provides a detailed comparative analysis of crime trends within each division throughout the nine-year timeframe. The data visually identifies five police divisions— D51, D52, and D55—with the highest crime counts, suggesting these areas are crime hotspots. In contrast, divisions D11, D12, and D13 have comparatively lower crime figures. The remaining divisions fall in between these two extremes.

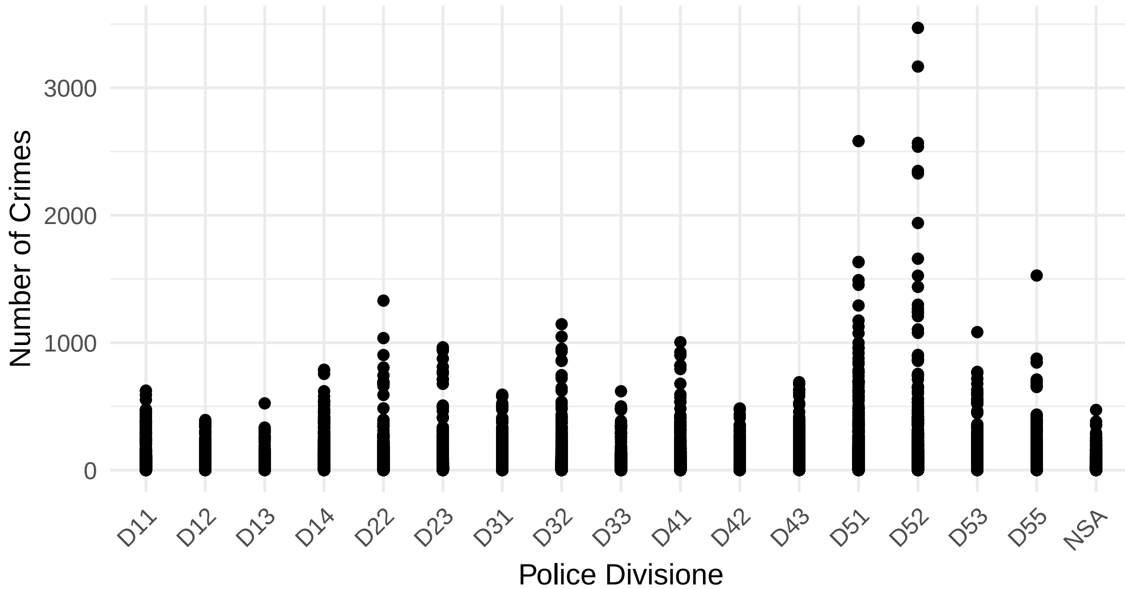

Figure 3: Number of Crimes Reported in Toronto by Police Division

Discussion 3

This graph tracks crime incidents across Toronto's divisions, stressing a higher crime presence in divisions D51, D52, and D55. The map clarifies the location of these divisions, with a notable concentration in downtown Toronto (D51, D52, D55) versus a lower occurrence in North York (D32). This data contrasts with Toronto media reports, which often portray North York's crime levels as similar to downtown's (Cinnamon, n.d.). Our analysis demonstrates that North York experiences fewer crimes than the downtown area. This variance points to a media bias, suggesting North York is more crime-ridden than it is, and underscores the gap between media portrayal and actual crime statistics.

4. Conclusion

This study analyzes various types of crimes, their geographic distribution across Toronto, the most frequently reported crimes, and trends in crime resolution rates. A key observation is the diminishing efficacy of police in addressing crimes, notably theft under $5,000, which is the most common crime in downtown Toronto. A critical issue identified is the disparity between the actual crime statistics and their representation in media reports. These findings highlight the necessity for accurate crime statistics dissemination and advocate for strategies to align media reports with factual data. Nonetheless, the study's credibility is somewhat compromised by omitting 'Crimes Against the Person' cases where the victim is unnamed, potentially distorting the overall crime picture.

5. References

Best, Joel. 2008. "Halloween Sadism: The Evidence." *Department of Sociology and Criminal Justice. Http://Udspace. Udel. Edu/Handle/19716/726, Accessed March* 8: 2019.

Cinnamon, Lindi Jahiu Jonathan. n.d. "Media Coverage and Territorial Stigmatization: An Analysis of Crime News Articles and Crime Statistics in Toronto."

Gelfand, Sharla. 2024. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

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/.

Porter, Julie Elizabeth. 1995. "Media Portrayal of Police: A Content Analysis of the" Toronto Star", the" Toronto Sun" and" Now"(ontario)."

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Siegel, L. J., and C. McCormick. 2019. *Criminology in Canada: Theories, Patterns, and Typologies*. Nelson Education Limited. https://books.google.ca/books?id=Wg3HvwEACAAJ.

Stroman, Carolyn A, and Richard Seltzer. 1985. "Media Use and Perceptions of Crime." *Journalism Quarterly* 62 (2): 340–45.

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, Winston Chang, Lionel Henry, Thomas Lin Pedersen, Kohske Takahashi, Claus Wilke, Kara Woo, Hiroaki Yutani, and Dewey Dunnington. 2021. *Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics*. https://CRAN.R-project.org/package=ggplot2.

Xie, Yihui. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.

Yung, Corey Rayburn. 2015. "Concealing Campus Sexual Assault: An Empirical Examination." *Psychology, Public Policy, and Law* 21 (1): 1.

Zhu, Hao, Thomas Travison, Timothy Tsai, Will Beasley, Yihui Xie, GuangChuang Yu, Stéphane Laurent, et al. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=dplyr.