The Multifaceted Nature of Urban Crime: A Case Study of Toronto*

Abdul Ahad Qureshi

25 January 2024

This work presents a novel approach to analyzing crime statistics, focusing on the simulation and exploration of the dataset.Our findings reveal significant trends in age and gender among arrested individuals, as well as the variability of crime types across different neighborhoods. These insights not only enhance our understanding of urban crime dynamics but also demonstrate the power of data science in transforming public safety data into actionable intelligence.

Introduction

This paper specifically addresses the analysis of crime statistics, with a focus on the "Police Annual Statistical Report - Arrested and Charged Persons" dataset made available by the Toronto Open Data portal. The importance of this analysis lies in its potential to reveal underlying trends and correlations within urban crime data, which can be pivotal for law enforcement and public policy formulation.

The analysis presented here is two-fold. Firstly, we develop a simulation model using R programming to create a synthetic yet realistic representation of crime data. This model not only aids in understanding the structure and nuances of the actual data but also serves as a crucial tool for hypothesis testing and preliminary analysis. Secondly, we implement a script for downloading and processing the real dataset from the Toronto Open Data portal. This script ensures reproducibility and efficiency in data handling, forming a solid foundation for subsequent in-depth analysis

The remainder of this paper is organized as follows. We begin with a detailed description of the methodology employed for both the simulation and the actual data retrieval. This is followed by a presentation of the results from our exploratory data analysis, highlighting key

*Code and data supporthing this analysis is available at: https://github.com/abdulahadq29/The-Multifaceted-Nature-of-Urban-Crime-A-Case-Study-of-Toronto.git findings such as age and gender trends among arrested individuals, and the distribution of crime types across neighborhoods. We then discuss the implications of these findings, drawing connections to existing literature and potential applications in policy and law enforcement strategies. Finally, the paper concludes with a summary of the key insights, limitations of the current study, and suggestions for future research directions in the field of urban crime data analysis.

Data

The dataset central to this analysis is sourced from the "Police Annual Statistical Report - Arrested and Charged Persons," available on the Toronto Open Data portal(Gelfand 2022). Data was collected and analyzed using the statistical programming software R (R Core Team 2023), with additional support from `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `dplyr` (Wickham et al. 2023), `readr` (Wickham, Hester, and Bryan 2023).

This dataset represents a comprehensive aggregation of individuals arrested and charged in Toronto, detailing various attributes such as age, gender, division, neighborhood, and crime category. The ethical implications of handling such sensitive data, including privacy concerns and the potential for statistical bias, have been rigorously considered. This data is publicly released by law enforcement agencies, ensuring compliance with ethical standards and public transparency.

The dataset's richness allows for a multifaceted analysis of crime trends in Toronto. It provides insights not only into the prevalence of various crime types but also into demographic trends among those arrested and charged. To fully explore this dataset, we utilized R programming (R Core Team 2023), specifically employing ggplot2 for visualization (Wickham 2016)

Using ggplot2, we constructed several plots to visualize the data in its raw, non-summarized form. For instance, a line graph (Figure 2) depicts the trend in genders of arrested individual. Another plot (Figure 4) shows the distribution of arrests accoeding to the felony, highlighting higher incidences of specific crime types.

Results

Trend in Arrests by Gender

From Figure 1, we can infer the following:

Male Dominance in Arrests: The data shows a significantly higher number of males being arrested compared to females in Toronto for the period covered. This is a common trend in many regions where males tend to have higher arrest rates.The male arrest count is approaching the 500,000 mark, while the female count looks to be roughly between 100,000 and 200,000.

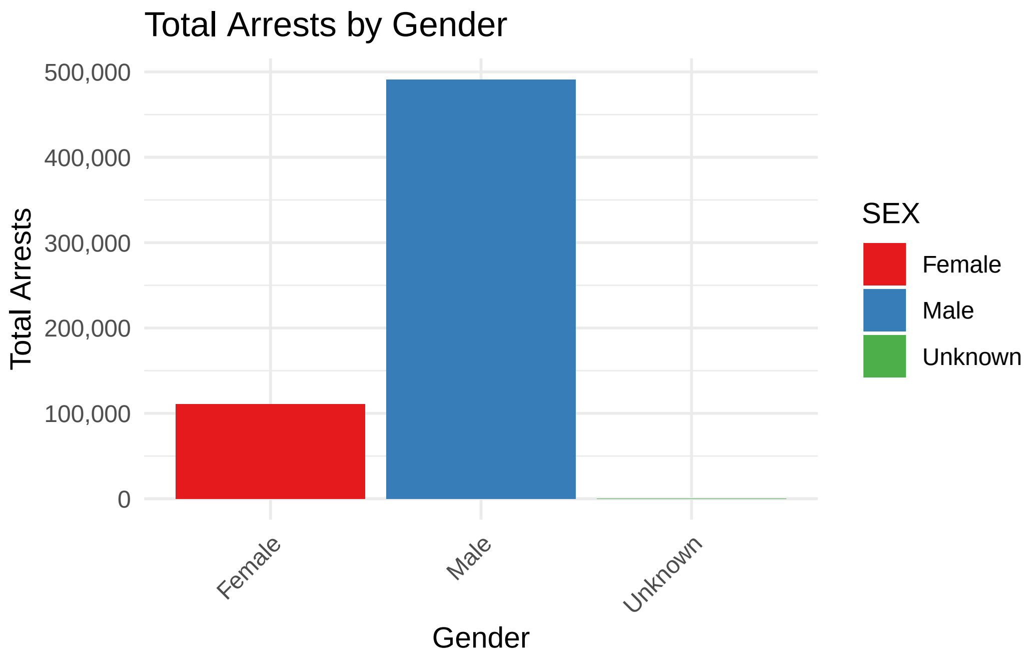

Figure 1: Total Arrests by Gender

Gender Disparities: The disparities between the genders could lead to discussions about various factors that might contribute to these differences, such as social, economic, and behavioral factors.

Trend in Arrests by Gender over the years

Figure 2 illustrates the total number of arrests by gender (Female, Male, Unknown) from the year 2014 to 2022. Here are the key observations:

Male Arrest Trends: The blue line, representing male arrests, shows a general decline from 2014 until around 2019 or 2020, after which there is a slight increase. The number of male arrests remains consistently higher than that of females throughout the period.

Female Arrest Trends: The red line, indicating female arrests, remains relatively flat and much lower than that of male arrests. There appears to be a slight decline over the years, but the change is not as pronounced as with the male arrests.

Overall Arrests Decrease: Both male and female arrests seem to show a decline over the given period, with male arrests showing a small recovery in the last observed year.

Data Points: The data points are recorded annually and show a distinct pattern for each gender, which can be valuable for analyzing trends and policy impacts over time.

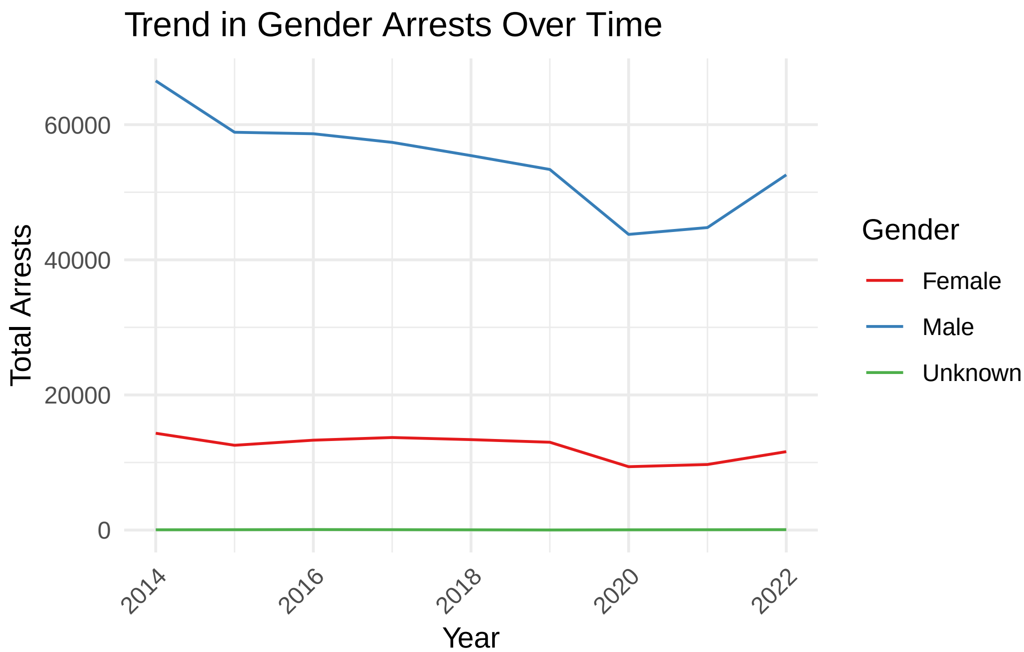

Figure 2: Trend in Total Arrests by Gender

Scale and Variability: The y-axis scale goes up to 60,000 arrests, which allows us to observe the variability in the number of arrests across genders clearly. The scale also suggests that the maximum number of male arrests in a year is around 60,000, while the maximum for females is considerably lower, around 10,000.

Trend in Arrests by Age Groups

Figure 3 illustrates the number of arrests across different age groups from 2014 to 2022. Each line represents an age cohort.

Highest Arrest Rates: The 25 to 34 age group (green line) consistently has the highest number of arrests across the entire timeframe, although there is a noticeable dip around 2020.

Trends Over Time:

The <18 age group (red line) shows a decrease over time, with a slight increase around 2021 but remains the group with the lowest number of arrests. The 18 to 24 age group (orange line) and the 25 to 34 age group both show a decline up to around 2020, followed by a recent increase. The 35 to 44 age group (light green line) and the 45 to 54 age group (light blue line) show slight fluctuations but remain relatively stable. The 55 to 64 age group (dark blue line)

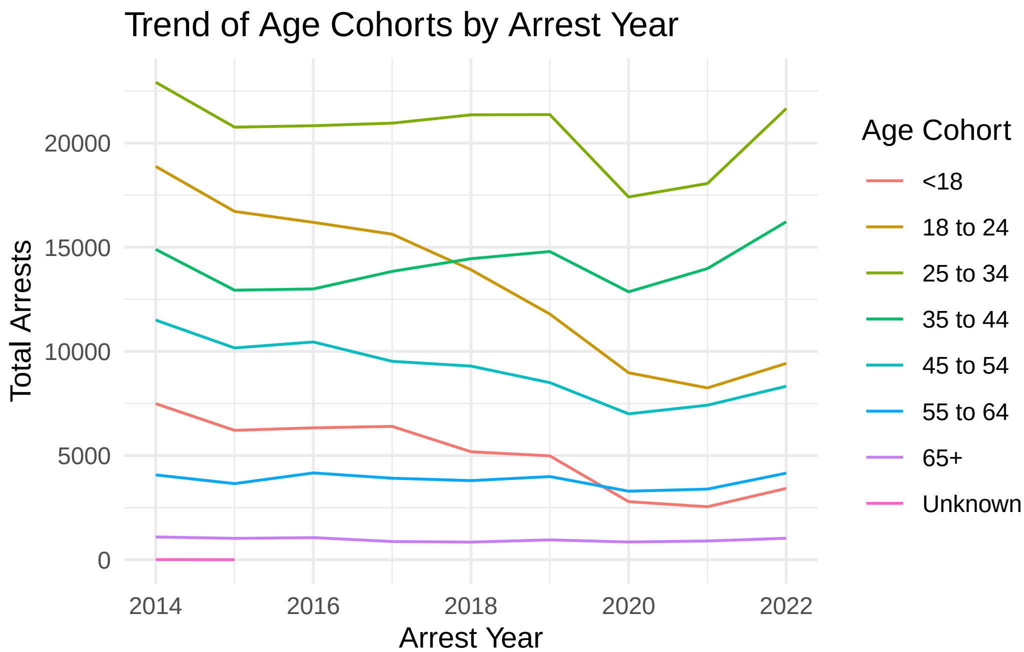

Figure 3: Trend in Age Cohorts of Arrested Individuals

and the 65+ age group (purple line) show an overall increase in arrests, with the 65+ group showing the most significant rise towards 2022.

Interpretation of Changes: The increase in arrests among older individuals (especially noticeable in the 65+ cohort) could suggest changes in policing, demographic changes in the population, or other socio-economic factors affecting older populations.

Contextual Factors: The dip seen around 2020 for several cohorts could be related to the global COVID-19 pandemic, which may have affected arrest rates due to lockdowns, changes in police enforcement, or changes in criminal activity patterns.

Data Quality: The consistent low or zero level for the 'Unknown' category suggests that the data collection on ages is thorough and that most arrest records include accurate age information.

Trend in Arrests by Felony Type

Figure 4 depicts the following:

Controlled Drugs and Substances Act (CDSA): The pink line shows a significant decrease in arrests from 2014 to around 2018, followed by a dramatic increase that peaks sharply in 2022.

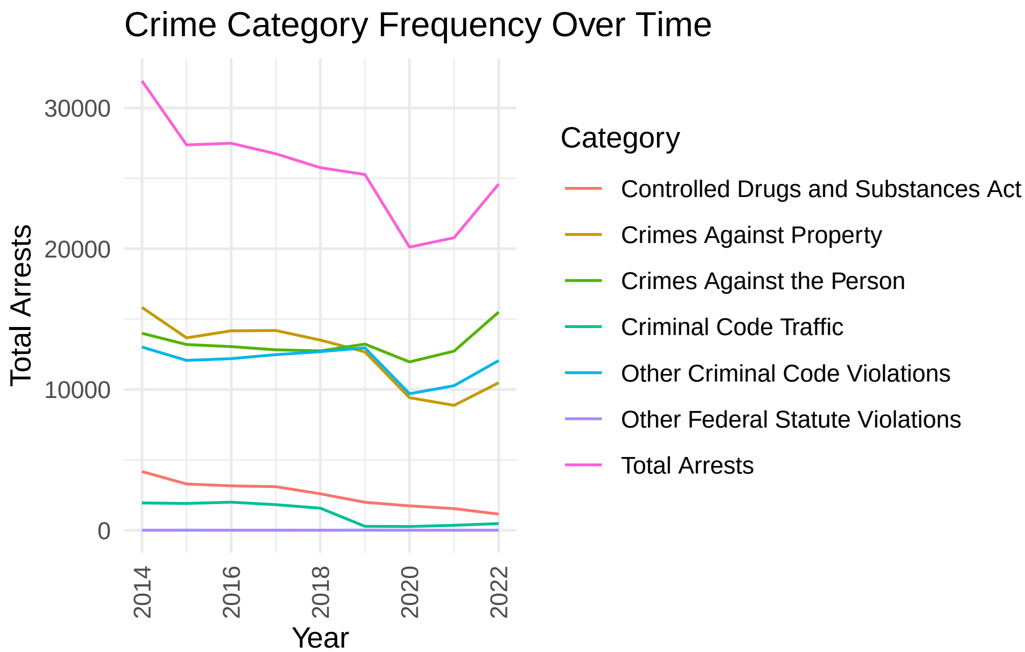

Figure 4: Total Arrests categorised into type of crime

Crimes Against Property: The orange line shows a gradual decline over the years, with a small uptick around 2022.

Crimes Against the Person: The green line remains relatively stable, with minor fluctuations but no clear trend upwards or downwards.

Criminal Code Traffic: The light blue line displays a slight decline over time with some fluctuation.

Other Criminal Code Violations: The dark blue line shows a general decrease in arrests, especially notable from 2014 to around 2019, after which it levels off.

Other Federal Statute Violations: The purple line shows an increase from 2014, peaks around 2017, then decreases slightly but remains relatively high compared to the starting point.

Total Arrests: The magenta line, which presumably sums all the categories, shows an overall downward trend with a significant dip around 2020, then a sharp rise back up towards 2022.

Key observations and implications:

The peak in CDSA-related arrests in 2022 could reflect changes in law enforcement focus, legal reforms, or shifts in drug-related activity.

The decline in property crime arrests could be related to a variety of factors, including better security measures, economic changes, or shifts in policing.

The stability of crimes against the person suggests a consistent rate of these types of crimes over time or consistent reporting and law enforcement practices.

The decline in traffic-related arrests and other Criminal Code violations may suggest improved compliance with laws, changes in law enforcement tactics, or legislative changes.

The overall decline in total arrests by 2020 could be attributed to the COVID-19 pandemic, which might have led to fewer interactions between the public and the police, as well as changes in crime patterns during lockdowns.

The recent increase in total arrests could indicate a return to pre-pandemic levels of law enforcement activity or a reaction to changes in crime rates.

Discussion

The discussion section of this paper synthesizes the results derived from the analysis of arrest data in Toronto, taking into account the various dimensions of age, gender, and crime type. The observed trends reveal that while male arrests outnumber female arrests, this discrepancy is shrinking over time, suggesting potential shifts in crime participation or law enforcement focus. The gender disparities in arrests, while significant, are shown to be in a state of flux, potentially reflecting deeper societal changes or evolving enforcement policies. The increase in arrests among older populations, especially noted in individuals over 65, raises questions about the interplay between aging demographics and criminal activity. The surge in violations of the Controlled Drugs and Substances Act demands an examination of the socio-political catalysts, while the downtrend in property crimes could be indicative of more effective preventive measures or economic developments. These findings underscore the complexity of urban crime and highlight the importance of a nuanced approach to understanding its dynamics.

Conclusion

In conclusion, "The Multifaceted Nature of Urban Crime: A Case Study of Toronto" underscores the intricate interplay of demographic factors and crime typologies within the urban tapestry of crime. The nuanced portrait of arrest patterns in Toronto, distinguished by gender and age disparities, as well as variations across crime categories, serves as a critical resource for policymakers and law enforcement agencies. This case study, by delineating the contours of urban criminality, offers empirical foundations upon which evidence-based policies and targeted crime prevention strategies can be formulated. The insights gleaned from this analysis hold the potential to steer policy-making towards more effective, equitable, and adaptive approaches to managing urban crime, thereby enhancing the efficacy of public safety initiatives and contributing to the broader discourse on urban governance and social justice.

References

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://sharlagelfand.github.io/opendatatoronto/.

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.

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, Kirill Müller, and Davis Vaughan. 2023. *Dplyr: A Grammar of Data Manipulation*. https://dplyr.tidyverse.org.

Wickham, Hadley, Jim Hester, and Jennifer Bryan. 2023. *Readr: Read Rectangular Text Data*. https://CRAN.R-project.org/package=readr.