Analyzing Toronto Police Statistics between 2014 to 2022*

Sirui Tan

January 22, 2024

In this paper, we investigated the crime dynamics in Toronto over the past nine years using data from the opendatatoronto Library. Employing sophisticated statistical tools from the R language, our analysis revealed discernible patterns, portraying Toronto's crime rate as a quadratic function with fluctuations tied to significant events like the 2020 city-wide closure due to the COVID-19 outbreak. By delving into specific crime categories, we identified a predominant focus on property infringement and threats to personal safety. Notably, our findings highlight the necessity for comprehensive societal efforts to address the surge in juvenile crimes observed in 2022. This research offers valuable insights into the complex interplay of societal and environmental factors influencing urban safety trends, contributing significantly to our understanding of crime dynamics in Toronto.

*Code and data are available at: https://github.com/siru1366/starter-folder.git.

1 Context

1.1 Introduction

1.2 Data

1.3 discussion

1.3.1 Yearly Trends: Statistics on the Occurrence of Five Major Crime Categories

1.3.2 Spatial and Temporal Analysis of Crime Incidents Across Toronto Neighborhoods Over the Years

1.3.3 Temporal Variation in Crime Incidents Across Different Age Cohorts

1.3.4 Demographic Analysis of Crime Incidents: A Cross-Sectional Study on Gender and Age Groups

1.3.5 Gender-Based Distribution of Crime Types: An In-Depth Analysis Across Different Offenses

1.4 Reference

2 Introduction

Criminal statistics, particularly within the domain of policing, serve as a vital instrument in comprehending and addressing patterns of criminal behavior within society. Toronto Police statistics, as an exemplar in this context, offer a meticulous portrayal of the demographic dynamics and criminal typologies characterizing arrests within the city from the temporal span of 2014 to 2022. Such statistical analyses provide a nuanced understanding of the intricate interplay between law enforcement efforts and the evolving landscape of criminal activities.

In the preliminary stage of this research, an examination was conducted to assess the fluctuations in Toronto's crime rates spanning the years 2014 to 2022. The findings depicted a distinctive pattern characterized by an initial decrease, followed by a subsequent rise, with a minimum observed in 2020. Acknowledging the necessity for a more nuanced comprehension, the study proceeds to a more in-depth analysis, incorporating four additional variables: gender, age group, crime type, and region. These variables are systematically paired, enabling a comprehensive exploration of their interdependencies, offering insights beyond the broad trends observed in the overall crime rate. In the data section, a clear presentation of the data is provided, organized by year. Transitioning to the discussion section, a more intricate analysis unfolds as the data undergoes further examination through the amalgamation of five

distinct sets of variables. The subsequent shift towards a more granular analysis, as proposed, aligns with the pursuit of deeper insights.

3 Data

The data employed in this paper was sourced from the opendatatoronto Library of the City of Toronto (Gelfand (2022)). Specifically, the dataset utilized is titled Police Annual Statistical Report - Arrested and Charged Persons Toronto Police Services (2023) from the Toronto Police Services. Data cleaning and analysis were conducted using the open-source statistical programming language R (R Core Team (2022)), leveraging functionalities from the tidyverse (Wickham et al. (2019)) suite, including ggplot2 (Wickham (2016)), dplyr (Wickham et al. (2023)), readr (Wickham, Hester, and Bryan (2024)), tibble (Müller and Wickham (2023)),stringr (Wickham (2023)) ,janitor (Firke (2023)) and knitr (Xie (2023)). The detailed procedures for data extraction and cleaning are expounded upon in the subsequent subsections.

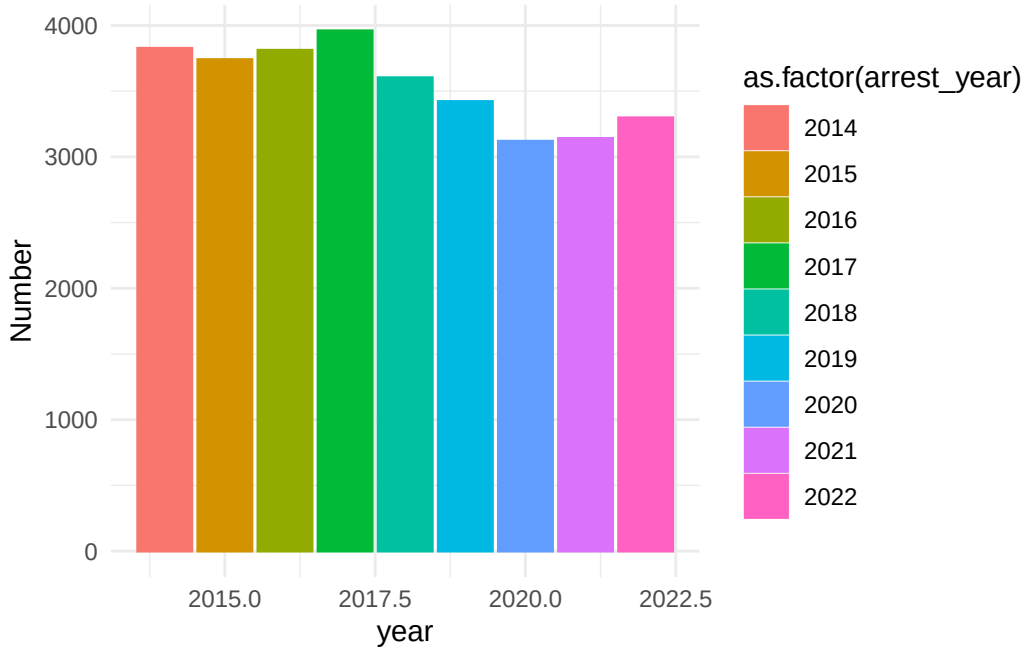

Figure 1: Relationship between crimes and year

Overall,upon analyzing the planes Figure 1 representation, discernible fluctuations in Toronto's crime rate over the years are apparent, resembling the characteristic curve of a quadratic function with an upward orientation. The nadir in 2020 coincides notably with the implementation of a city-wide closure policy in response to the COVID-19 outbreak.@HODGKINSON2022101881 indicated that amidst the COVID-19 epidemic, Toronto

experienced a notable reduction in its crime rate, attributable to a myriad of governmental control policies .

4 discusssion

An examination solely focused on the aggregate count of criminal incidents per annum yields a somewhat overarching perspective. To cultivate a more nuanced understanding, it becomes imperative to conduct a more granular analysis that delves into the intricate details and specific categories of criminal activities. In the discussion section, a dual-stage scrutiny of Toronto offender data is conducted.

4.1 Yearly Trends: Statistics on the Occurrence of Five Major Crime Categories

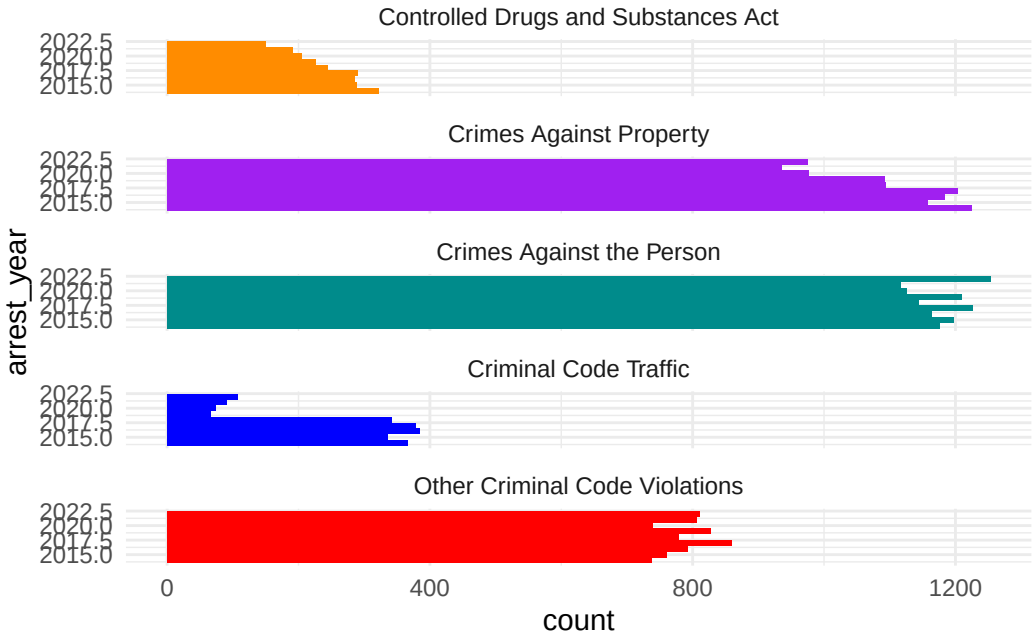

Figure 2: Yearly Trends—Statistics on the Occurrence of Five Major Crime Categories

The analysis of criminal conduct in Toronto reveals a discernible classification into five distinct categories. Drawing upon data Figure 2 spanning the preceding nine years, a prominent pattern emerges, underscoring that the primary forms of criminality center around property infringement and threats to personal safety. This nuanced observation underscores the significance of addressing these specific dimensions within the context of Toronto's criminal landscape.

Significantly, an observable downward trajectory characterizes the incidence of offenses falling within the purview of the Controlled Drugs and Substances Act.

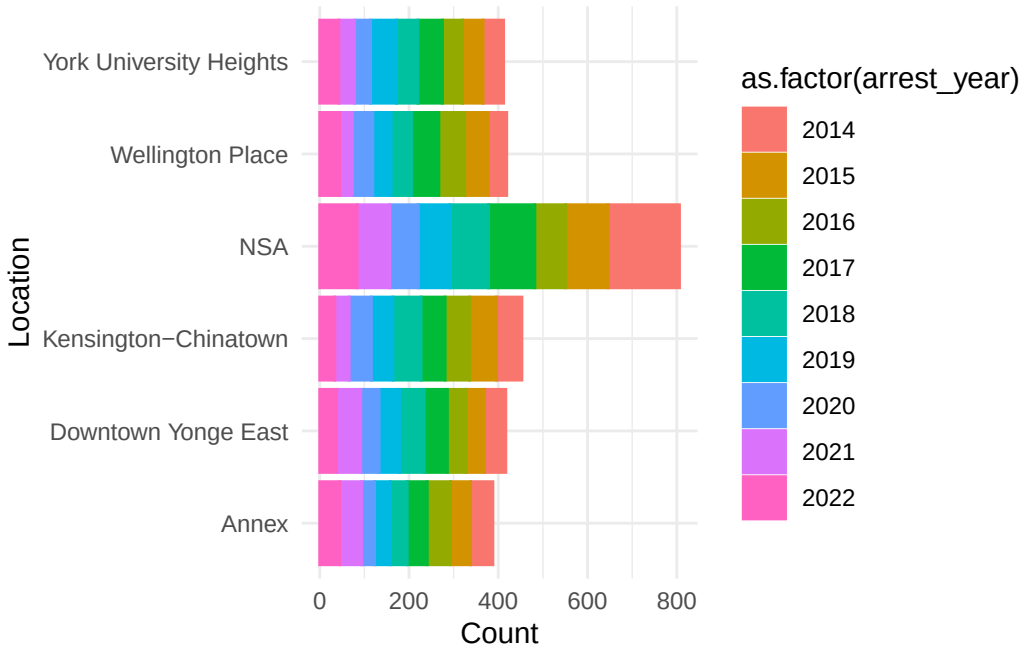

Figure 3: Spatial and Temporal Analysis of Crime Incidents Across Toronto Neighborhoods Over the Years

4.2 Spatial and Temporal Analysis of Crime Incidents Across Toronto Neighborhoods Over the Years

Beyond the classification based on the nature of the crime, the geographical context of criminal incidents also holds paramount significance in crime analysis. The primary dataset encompasses an extensive list of 158 distinct streets in Toronto, rendering it operationally unwieldy for presentation purposes. In the interest of visual clarity and expediency, a pragmatic approach is adopted, focusing specifically on the six blocks that emerge as the most frequent locations for criminal activities. This judicious selection facilitates a more focused and manageable exploration of spatial patterns within the broader landscape of criminal occurrences. Figure 3 The neighborhood known as NSA (Neighborhood Security Analysis) in Toronto registers the highest incidence of criminal activities, exhibiting a prevalence that is nearly twofold compared to the area occupying the second position in the ranking.

4.3 Temporal Variation in Crime Incidents Across Different Age Cohorts

The cartography of offenders holds a pivotal role in comprehensive crime analysis, particularly when delineating them based on age categories. This stratification allows for the identification of specific age cohorts warranting heightened scrutiny and focused attention. This meticulous

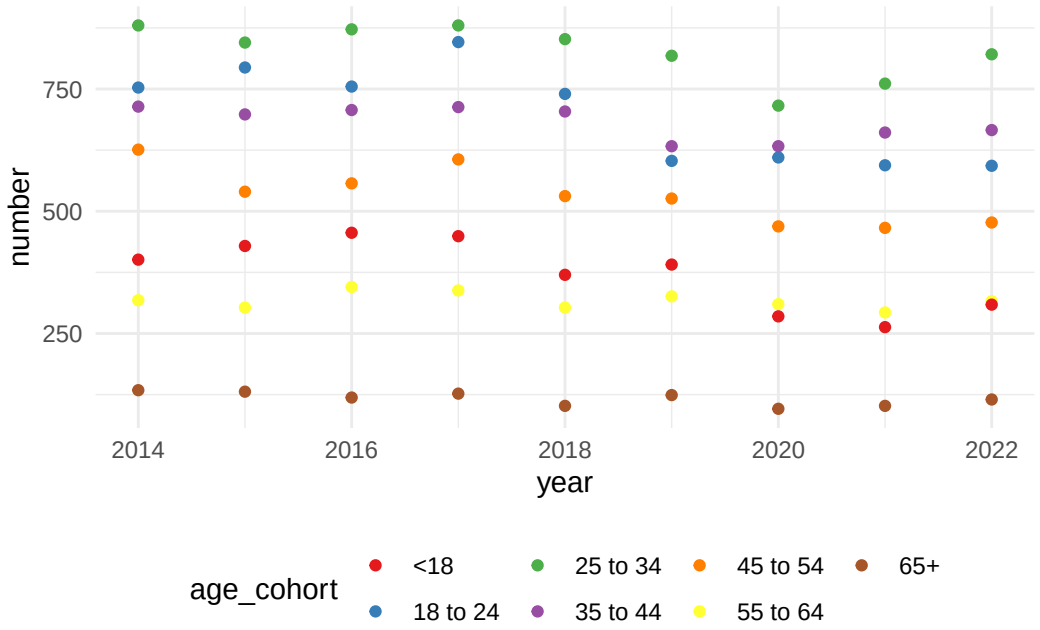

Figure 4: Temporal Variation in Crime Incidents Across Different Age Cohorts

categorization aids in discerning nuanced patterns within the demographic landscape of offenders, contributing substantively to a more refined understanding of the interplay between age dynamics and criminal behavior. Over the course of the past nine years, the predominant demographic among offenders in Toronto has consistently manifested within the age bracket spanning from 25 to 44 years. Figure 4 It is crucial to highlight that the incidence of juvenile crimes in Toronto has shown an elevation in the year 2022 when juxtaposed against the preceding year.

4.4 Demographic Analysis of Crime Incidents: A Cross-Sectional Study on Gender and Age Groups

The consideration of gender, alongside age, constitutes a pivotal aspect of our analytical framework. The chart Figure 5 provides insights into the gender distribution across various age groups among offenders. Regardless of the age cohort, the proportion of female offenders consistently falls significantly below that of their male counterparts. This trend is particularly pronounced in the juvenile and senior age groups, where the number of female offenders represents only about one-quarter of the male demographic. In the middle-aged group, this figure hovers at approximately one-third of the male population.

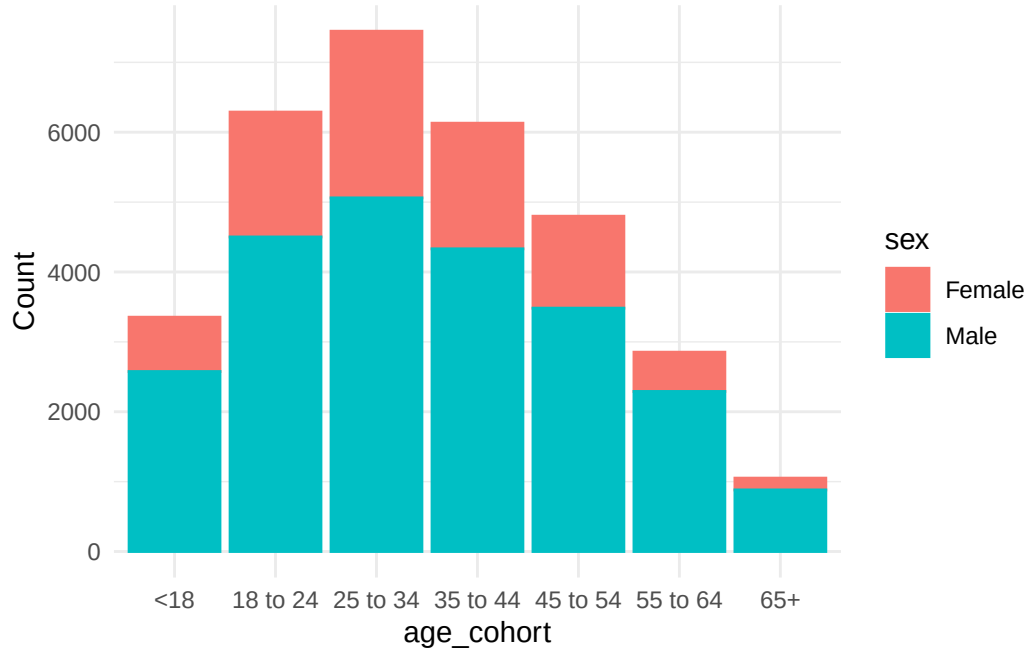

Figure 5: Demographic Analysis of Crime Incidents—A Cross-Sectional Study on Gender and Age Groups

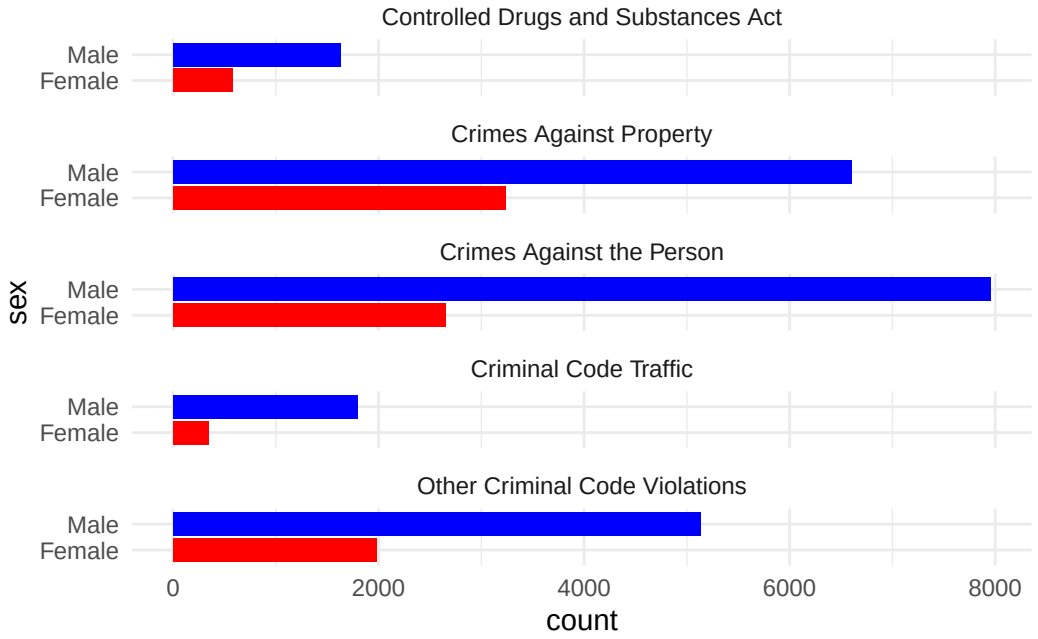

Figure 6: Gender-Based Distribution of Crime Types—An In-Depth Analysis Across Different Offenses

4.5 Gender-Based Distribution of Crime Types: An In-Depth Analysis Across Different Offenses

By employing various screening methodologies, we can discern the gender distribution among offenders across distinct crime categories. Figure 6 Consistently, the representation of women in the five principal crime types in Toronto is notably lower than that of men. Specifically, the male-to-female ratio is most proximate in offenses against property, standing at approximately 2:1. In contrast, the male-to-female ratio diverges considerably in the category of criminal traffic offenses, where it reaches approximately 5:1.This variance may be attributed to the demographic composition of drivers, where the proportion of female drivers is noticeably smaller than that of their male counterparts.For instance,the representation of women among Canadian train drivers stands at a mere three percent (Hanson (2021)).

Reference

Firke, Sam. 2023. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. https://CRAN.R-project.org/package=janitor.

Gelfand. 2022. "OpenDataToronto Library." https://www.toronto.ca/city-government/data-research-maps/open-data/.

Hanson, Natasha. 2021. "The Intersections of Global Capital and Family Rhythms in Truck Driving: Elucidating the Canadian Trucking Industry Labour Crisis." *Applied Mobilities* 6 (2): 153–68. https://doi.org/10.1080/23800127.2020.1809969.

Müller, Kirill, and Hadley Wickham. 2023. *Tibble: Simple Data Frames*. https://github.com/tidyverse/tibble.

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 Police Services. 2023. "Police Annual Statistical Report - Arrested and Charged Persons." 2023. https://open.toronto.ca/dataset/police-annual-statistical-report-arrested-and-charged-persons/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

———. 2023. *Stringr: Simple, Consistent Wrappers for Common String Operations*. https://stringr.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. 2024. *Readr: Read Rectangular Text Data*. https://readr.tidyverse.org.

Xie, Yihui. 2023. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. https://yihui.org/knitr/.