# Explore Toronto crime data: what is the relationship between crimes, ages, and gender?*

Ben Li

January 23, 2024

A number of crimes are crucial in evaluating the city's safety in Toronto. This report aims to examine the Toronto crime report data from 2014 to 2022, including exciting variables: year, crime type, sex, and age. By exploring the trend for the number of crimes in Toronto, it shows an increasing trend overall. With further analysis to find out the relationship between the number of times and age and sex, we find men and women have a similar probability of committing crimes, and adults commit the dominant crimes.

## Table of contents

---

*Code and data supporting this analysis is available at: https://github.com/UofT-DailinLi/sta-302-assignemnt-1.git

# 1 Introduction

The Toronto Crime Report records all crimes from 2014 to 2022 and analyzes the data to provide recommendations to help reduce crime rates. This report looks at the year of the crime, type of case, age, and gender. In section 2, the organized database is presented with 1110 crimes from 2014 to 2022 and a visualization of the change in the number of crimes per year from 2014 to 2022 through Figure 1. In section 3, figure 2 reveals that the predominant type of crime out of the 1110 is ASSAULT. However, figure 3 overrides common sense by revealing no significant difference between the number of crimes committed by males and the number of crimes committed by females, proving that the probability of committing a crime by males and females is not that different. Figure 4 demonstrates that out of the 1110 crimes committed, adults accounted for a significant share of crimes with a caseload of approximately 780. Further details are provided in sections 2 and 3.

# 2 Data

Data used in this paper are gathered from from Open Data Toronto Portal through the library `opendatatoronto`(Gelfand 2022). Data was cleaned and analyzed using the open source statistically programming language R (R Core Team 2022), and additional packages from `tidyverse` (Wickham et al. 2019), `ggplot2` (Wickham 2016), `janitor` (Firke 2021), and `knitr` (Xie 2014). Details of the data extraction and cleaning processes are discussed in the subsections below.

Table 1: First Ten Rows of Toronto Crime Data

| year | crime_type | sex | age_group |
|------|------------|-----|-----------|
| 2014 | Sexual Violation | M | Adult |
| 2014 | Sexual Violation | M | Adult |
| 2014 | Sexual Violation | M | Adult |
| 2014 | Sexual Violation | M | Adult |
| 2014 | Sexual Violation | M | Adult |
| 2014 | Sexual Violation | M | Adult |
| 2014 | Sexual Violation | M | Child |
| 2014 | Sexual Violation | M | Unknown |
| 2014 | Sexual Violation | M | Youth |
| 2015 | Assault | F | Unknown |

Table 1 shows the first ten rows of cleaned Toronto crime data. Variable "year" indicates the specific year when the crime happened, and variable "crime_type" shows the crime types, including four main types: "Sexual Violation," "Assault," "Robbery," and "Other." Variable

"sex" involves male, female and unknown. Variable "age_group" contains four types: "Adult," "Child," "Unknown," and "Youth."

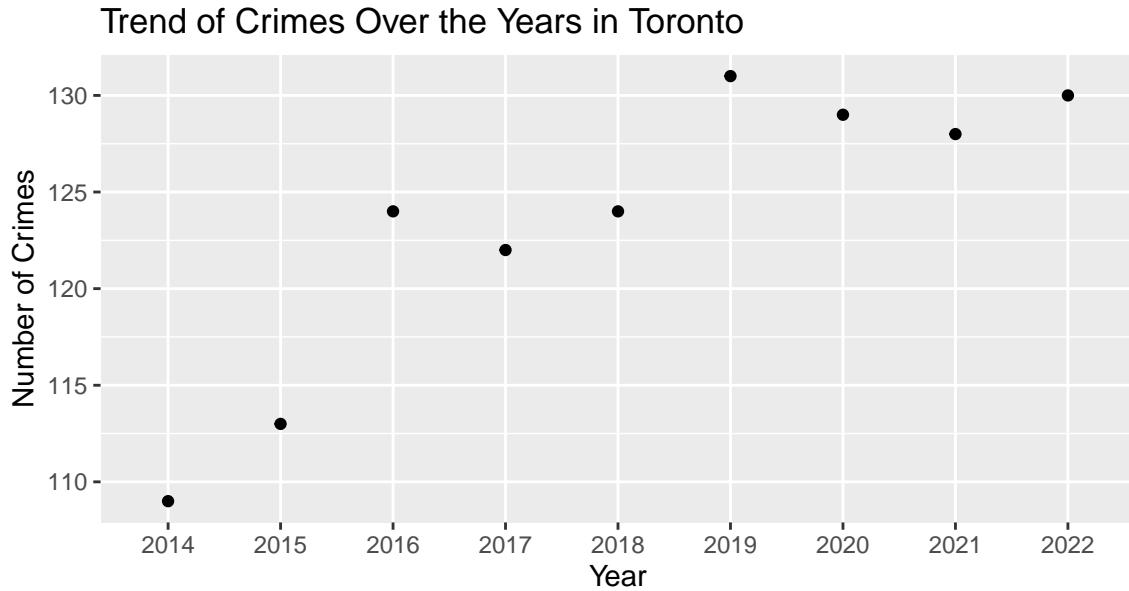## Trend of Crimes Over the Years in Toronto



Figure 1: Trend of total Crimes Over the Years in Toronto

Figure 1 shows an overall increase trend for the number of crimes from 2014 to 2022 in Toronto. The lowest number of crimes was around 108 in 2014. The highest number of crimes was around 133 in 2019. From the figure, we can see a considerable jump from 2016 to 2016. Even though the overall trend is increasing, the gap between the highest and lowest is about 25 crimes a year. This indicates Toronto did not do an excellent job of maintaining society's safety, but there might be other factors that affect the number of crimes, such as the total population in Toronto, the economy, the number of homeless, etc.

## 3  Result

Figure 2 illustrates the distribution of different types of crimes in Toronto. The type of crime, "assault," is much higher than the other three. The other three types of crime, "robbery," "sexual violation," and "others," share similar amounts of crimes. But assault is almost three times as the other three types of crimes. This indicates that the Toronto government should take more action to reduce the assault crime rate.

Figure 3 shows the distribution of crimes by sex in Toronto. Based on human common sense, males tend to have more crime probability than females. Figure 3 shows there is a considerable difference between males and females. Males have around 510 crimes in total from 2014 to 2022 in Toronto; females have around 460 crimes from 2014 to 2022 in Toronto, and around
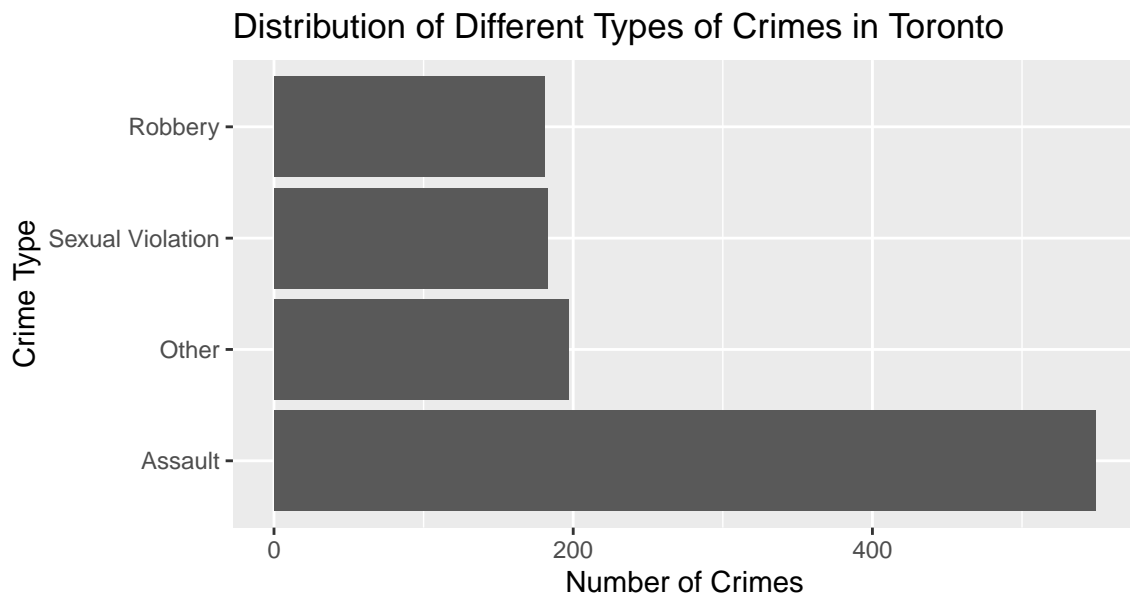
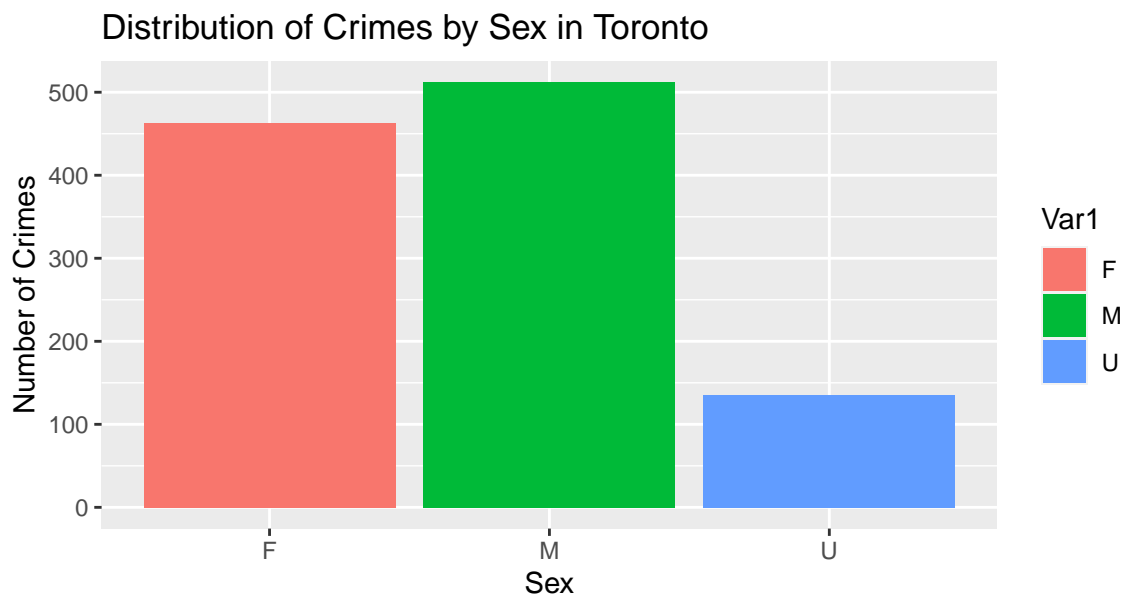Figure 2: Distribution of Different Types of Crimes in Toronto



Figure 3: Distribution of Crimes by Sex in Toronto

125 crimes from 2014 to 2022 in Toronto. Figure 3 indicates males and females have similar changes to commit crimes.
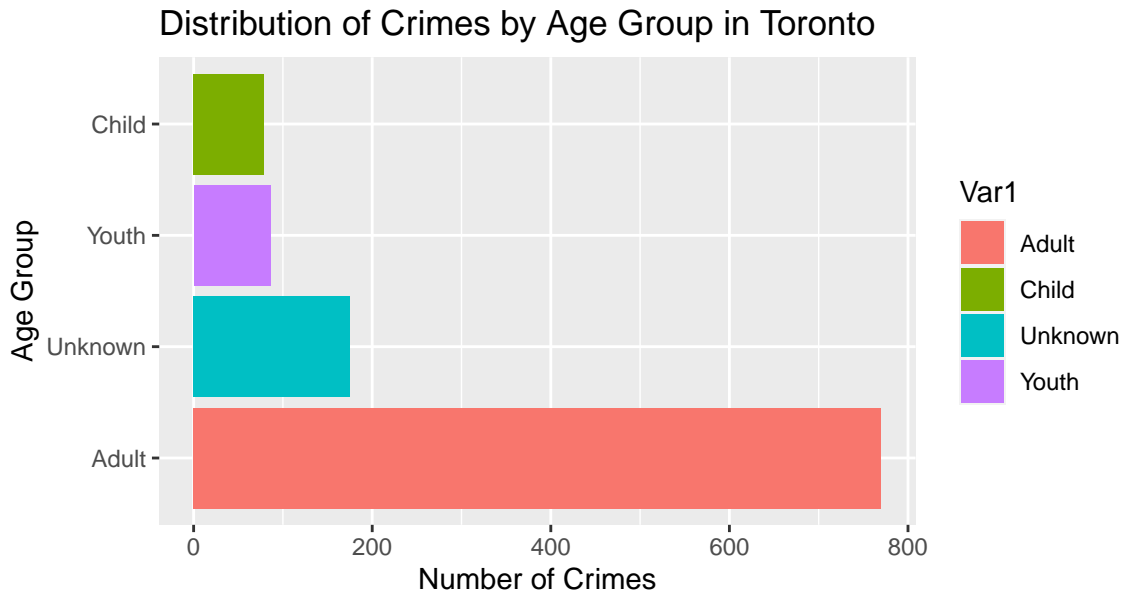


Figure 4: Distribution of Crimes by Age Group in Toronto

Figure 4 visualizes the distribution of crime by age group in Toronto from 2014 to 2022. The dominant age group is adults. The number of crimes of adults in Toronto from 2014 to 2022 is about 780 cases. The second largest group is unknown, which means the police did not find out the criminal. Children and youth have a similar share of the number of crimes. Both age groups have around 90 cases of crimes.

# 4 Discussion

# Reference

Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data.* https://CRAN.R-project.org/package=janitor.

Gelfand, Sharla. 2022. *Opendatatoronto: Access the City of Toronto Open Data Portal.* https://CRAN.R-project.org/package=opendatatoronto.

R Core Team. 2022. *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.

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. http://www.crcpress.com/product/isbn/9781466561595.