

My title*

My subtitle

Siming Shan

07 February 2022

Abstract

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

In January 2022, two NYPD police officers Jason Rivera, 22 years old, and Wilbert Mora, 27 years old, were shot and killed during their investigation to a domestic incident.

2 Data

2.1 Raw Dataset Introduction

I utilized an Annual Victims of crime report(data 2021) from Toronto Open Data portal, using the `opendatatoronto` package(Gelfand 2020) to have a better understanding of the relationship between peace officers and the communities in Toronto. This data was published by Toronto Police Services, and the last time the data was refreshed on Aug 12, 2021. There are 854 observations in the raw data, each observation has 12 variables - `id`, `index`, `reportedyear`, `category`, `subtype`, `assaultsubtype`, `sex`, `agegroup`, `agecohort`, `count`, `objectid`, `geometry`. This dataset collects victim of crime which crime was reported between 2014 and 2020, and the victim of the crimes are classified as peace officers, other, and unknown. Subtypes of assault against peace officers include Aggravated Peace Officer, Assault Peace Officer, Assault Peace Officer Weapon/Bodily Harm, and Assault Resist Arrest. Since I am only interested in crimes against peace officers, I cleaned the data and extracted essential observations by using `R`(R Core Team 2020), `tidyverse`(Wickham et al. 2019), `dplyr`(Wickham 2021), and `janitor`(Firke 2021).

2.2 Cleaning Process And Cleaned Dataset

The columns “`index`” and “`geometry`” of the dataset are all `NA` and `Null`, and the column “`ObjectId`” is redundant with “`id`,” thus I removed these columns as they don’t provide useful information for analysis. Then I modified each column name to make them more organized by using the package `janitor`(Firke 2021). Finally, I extracted the observation that only peace officers were victims. The cleaned dataset has 168 observations, each observations has 9 variables:

- `id`: Unique row identifier
- `reported_year`: Year crime was reported (from 2014 to 2020)
- `category`: Crime category
- `subtype`: Crime category subtype
- `assault_subtype`: Breakdown of assault subtypes
- `sex`: Sex of identified victim
- `age_group`: Age group of identified victim, adult or youth

*Code and data are available at: <https://github.com/SimingShan/STA304-Project-1>

- age_cohort: Age cohort of identified victim
- count: Count of identified victims

A brief view of the dataset made by using knitr(Xie 2021) is shown below:

Table 1: Cleaned Dataset

id	reported_year	category	subtype	assault_subtype	sex	age_group	age_cohort	count
1	2014	Crimes Against the Person	Assault	Aggravated Peace Officer	M	Adult	25-34	2
2	2014	Crimes Against the Person	Assault	Aggravated Peace Officer	M	Adult	55-64	1
3	2014	Crimes Against the Person	Assault	Assault Peace Officer	F	Adult	18-24	1
4	2014	Crimes Against the Person	Assault	Assault Peace Officer	F	Adult	25-34	5
5	2014	Crimes Against the Person	Assault	Assault Peace Officer	F	Adult	35-44	3
6	2014	Crimes Against the Person	Assault	Assault Peace Officer	F	Adult	45-54	1

Table 1: A brief view of the cleaned dataset.

2.3 Simple Assault Is The Most Common Crime Against Peace Officers

In this dataset, there are four types of assault against peace officers - aggravated assault, assault, weapon/bodily harm, resist arrest. Assault, also known as simple assault, includes minor injuries, touching, and threatening word or actions, while aggravated includes serious injuries or any conduct involving weapon(Reaves 2019). Weapon/bodily harm means hurts and injuries to a person through weapon or physical contact. Resist arrest means attacking peace officers while being arrested. To analysis which subtype of assault is the most common one against peace officers, I made Figure 1 by using ggplot2 package(Wickham 2016).

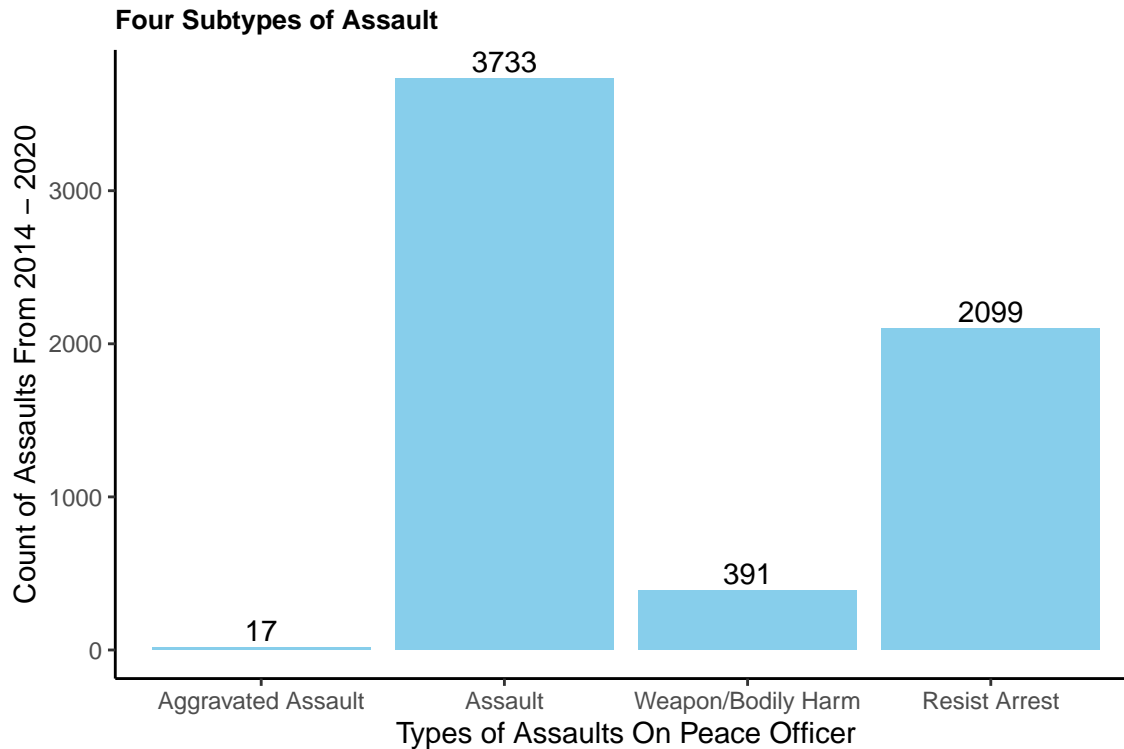


Figure 1: Total number of cases of each subtype of assault on peace officers from 2014 to 2020.

In figure 1, x-axis presents all four subtypes of assault on peace officers, y-axis represents the total number of each subtype of assault from 2014 to 2020. The figure shows the fact that simple assault is the most common assault on peace officers in Toronto, counting for 59.8% of all assault, resist arrest is the second most common assault, counting for 33.6% of all assault, while weapon/bodily harm and aggravated assault counted for only 6.3% and 0.3% of all assault respectively.

2.4 Younger peace officers more likely to be assaulted

According to the chart of age distribution of police officers provided by Statistics Canada, there are 2.33% of police officers are between the ages of under 20 and 24, 28.7% of police officers are between the ages of 25 and 34, 37.5% of police officers are between the ages of 35 and 44, and 31.45% of police officers are over 45 years old (Statistics Canada 2015). To investigate whether a certain age cohort of peace officers are more likely to be assaulted, I made a table (Table 2) and a plot (Figure 2) by using knitr (Xie 2021) and ggplot2 (Wickham 2016) packages to present and compare the age distribution of police officers and age distribution of victims.

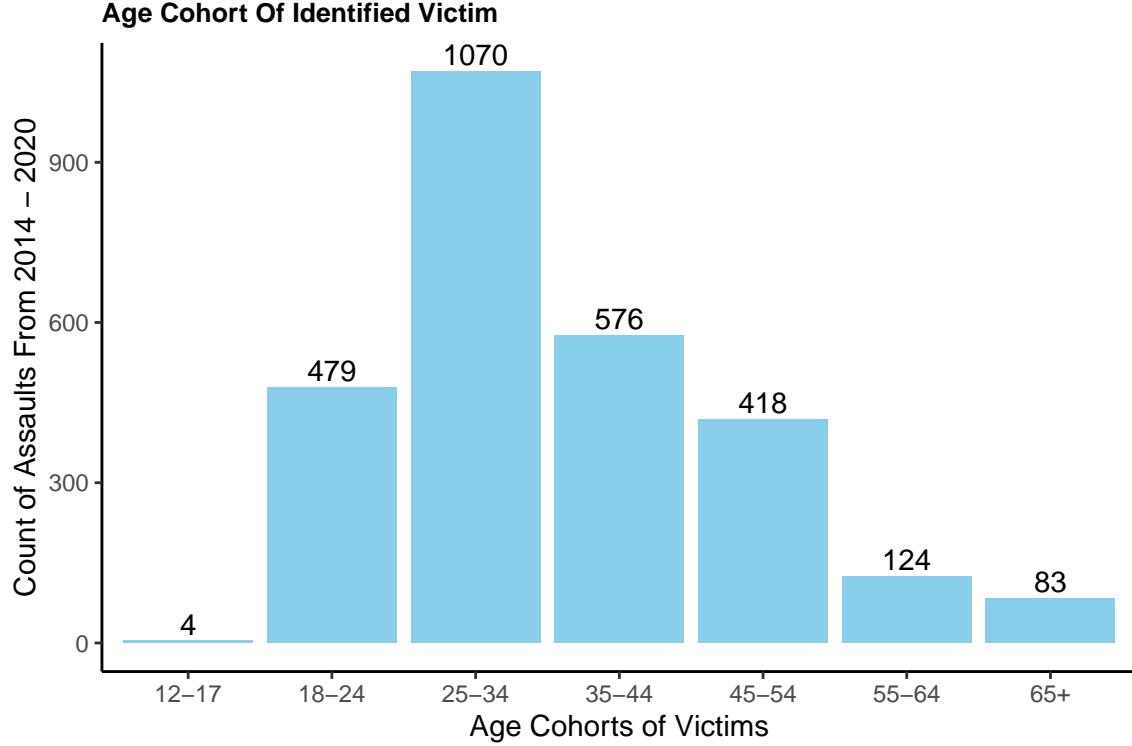


Figure 2: Total number of victims of each age cohort from 2014 to 2020.

Table 2: Age Distribution Comparison

age_cohort	age_count	police_age_percentage	victim_age_percentage	net_percentage
12-17	4	0.02	0.15	-0.13
18-24	479	2.31	17.39	-15.08
25-34	1070	28.66	38.85	-10.19
35-44	576	37.53	20.92	16.61
45-54	418	27.07	15.18	11.89
55-64	124	3.63	4.50	-0.87
65+	83	0.75	3.01	-2.26

Table 2: Age distribution of peace officers vs age distribution of victims from 2014 to 2020.

Figure 2 shows the age distribution of victims, x-axis represents different age cohorts of victim, y-axis represents total number of assault from 2014 to 2020. The plot is highly right skewed, which means most of the victims of assault are younger peace officers, however, It doesn't mean younger peace officers are more likely being assaulted since there are probably more younger peace officers in the first place. To address this issue, I produced Table 2 to compare the age distribution of peace officers and victims. I calculate the net percentage by subtracting the distribution of victim age from the distribution of police age, the smaller the net percentage, the more likely the peace officers in corresponding age cohort will be assaulted. the Table 2 shows the net percentages for the 18-24 and 25-34 age cohorts are -15.08% and -10.19% respectively, and the net percentages for the 35-44 and 45-54 are 16.61% and 11.89% respectively, and the net percentages for other age cohorts are close to 0. As a result, younger peace officers are more likely to be assaulted.

3 Conclusion

Reference

- data. 2021. “Police Annual Statistical Report - Victims of Crime.” Toronto Police Services. <https://open.toronto.ca/dataset/police-annual-statistical-report-victims-of-crime/>.
- Firke, Sam. 2021. *Janitor: Simple Tools for Examining and Cleaning Dirty Data*. <https://cran.r-project.org/web/packages/janitor/index.html>.
- Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. <https://cran.r-project.org/web/packages/opendatatoronto/index.html>.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Reaves, Walter. 2019. “What Is the Difference Between Simple Assault and Aggravated Assault in Texas?” <https://www.reaveslegal.com/faqs/what-is-the-difference-between-simple-assault-and-aggravated-assault-in-texas-.cfm#:~:text=Simple%20assault%20encompasses%20minor%20injuries,and%20threatening%20words%20or%20behavior.&text=Aggravated%20assault%20involves%20serious%20injuries,could%20be%20considered%20aggravated%20assault>.
- Statistics Canada, Government of Canada. 2015. “Age Distribution of Police Officers, Canada, 2011.” <https://www150.statcan.gc.ca/n1/pub/85-225-x/2012000/longdesc-ct005-eng.htm>.
- Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <http://ggplot2-book.org>.
- . 2021. *Dplyr: A Grammar of Data Manipulation*. <https://cran.r-project.org/web/packages/dplyr/index.html>.
- 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. 2021. *Knitr: A General-Purpose Package for Dynamic Report Generation in r*. <https://cran.r-project.org/web/packages/knitr/index.html>.