Crime worsened but got slightly better during the pandemic
A comprehensive overview of Toronto's crime landscape from 2014 to 2020

Ka Chun Mo 1004764873

06 February 2022

Abstract

Most Canadians consider Canada to be a safe place to live in. However, they consider Toronto to be less safer than the smalls towns and rural areas in Canada. This paper found that Toronto's crime worsed from 2014 to 2019 but it got slightly better during the pandemic. This report can also act as a guide for individuals and families to learn about the most/least occurred criminal offences and the safest and the least safest neighborhoods in Toronto.

1 Introduction

In 2019, the Economist Intelligence Unit (EIU) released a Safe Cities Index ranking Toronto the 6th out of 60 cities. (Toronto 2020)The EIU ranked the cities by their digital security, health security, infrastructure safety, and personal safety. Most Canadians considered Canada to be a safe place to live in. However, there were regional variations where people who lived in small towns and rurals areas felt safer than people who lived in big cities. In particular, residents who lived in Toronto reported that they felt less safe comparing to those who lived in small towns and rural areas. (*Canadians' Perceptions of Personal Safety and Crime* 2017) Canadians think Toronto is more dangerous than the other cities in the rest of the country. In reality, Toronto is actually one of the safest cities in Canada. (*Canadians Think Toronto Is Way More Dangerous Than It Really Is, Thanks to Media Attention* 2017) With the contradiction of Canadians' perception and real-world data of how safe Toronto is, it is crucial that we look more closely to Toronto's safety over the years and explore various aspects of the dangers in Toronto.

In this report, we will dive into the dataset, Major Crime Indicators, from . . . to gain a general sense of the criminal offence landscape in Toronto. We will investigate the most/least occurred crime categories in section . . . and the locations with the most/least crime occurred. We will also look into crime counts over the years, months, and days of the month in section . . . to determine if there is a pattern of the number of criminal offence over time. This report will serve as a high-level overview of the crimes and safety in Toronto.

2 Data

2.1 Data Source

This report uses a dataset that includes all Major Crime Indicators (MCI) occurrences reported to the Toronto Police Service in Toronto from the Open Data Toronto Portal. Some of the Major Crime Indicators are Assault, Robbery, Theft, Unlawfully In Dwelling-House. This data is also gathered from the perspective of the victims. One crime occurrence can correspond to multiple rows since multiple incidents or categories can occur in a single occurrence.

2.2 Data Collection & Methodology

This dataset contains all the Major Crime Indicators occurrences that were reported to the Toronto Police services. It includes the MCI occurrences in 140 neighborhoods in Toronto. However, the coordinate fields

of some occurrences can be blank since their locations cannot be verified. The dataset also contains some occurrences outside the city of Toronto.

We use R, a Statistical Computing Language, to analyze this dataset (R Core Team 2021).

2.3   Data Characteristics

2.4   Key Features

The dataset contains 242879 observation with 28 attributes such as index, offence type, reported year, occurrence year, neighborhood, longitude, and latitude. It includes most crimes including Assault, Robbery, Theft, but it excludes sexual violations. The attributes that we use in this report include offence category, occurrence date, reported date, neighborhoods, etc.. There are some variables that are similar and we will discuss that in section 4.

2.5   Limitations

There are also limitations on this dataset. For instance, the location of the crime occurrences were offset to the nearest road intersection. The Toronto Police Service did this to protect the privacy of the individuals that are involved in the occurrence. Because of this, the numbers by Division and Neighborhood may not reflect the accurate number of occurrences in those regions. Moreover, some crime occurrences do not have a location because the location of the occurrences are unknown. As a result, coordinate fields of those occurrences might be blank. On the other hand, this dataset does not include all the crime occurrences that are in Toronto since Statistics Canada did not include the crime occurrences if the police investigation did not occur or attempted.

3   Data Exploration

3.1   Offence categories

Table 1 shows the top 6 most occurred offence categories and the top 6 least occurred offence categories. We can see that "Assault," "B&E" (Breaking and Entering), and "Theft Of Motor Vehicle" are the top 3 most occurred offence categories, whereas "Set/Place Trap/Intend Death/Bh," "Traps Likely Cause Bodily Harm," and "Theft Of Utilities Over" are the top 3 least occurred offence categories. These two tables give us a sense of the most and least frequent offence in Toronto. However, if we examine the offence categories more closely, we can see that some offences (e.g. Assault, B&E) have sub-categories in the dataset. This might lead us to over-estimate or under-estimate the number of offences in each category. We will discuss more about the consequence of this in Section 4.

Category	Count	Category	Count
Assault	88919	Set/Place Trap/Intend Death/Bh	1
B&E	42464	Traps Likely Cause Bodily Harm	1
Theft Of Motor Vehicle	29158	Theft Of Utilities Over	4
Assault With Weapon	21540	B&E - M/Veh To Steal Firearm	11
Robbery - Mugging	7500	B&E - To Steal Firearm	11
B&E W'Intent	6279	Unlawfully Causing Bodily Harm	13

Table 1: Top 6 Most/Least Occurred Offence Categories

3.2   Crime Location

Table 2 shows the top 6 most occurred locations and the top 6 least occurred locations of crime in Toronto. We can see that "Waterfront Communities-The Islan," "Church-Yonge Corridor, and"Bay Street Corridor"

are the top 3 most occurred locations, whereas "Lambton Baby Poin," "Yonge-St.Clair," and "Maple Leaf" are the top 3 least occurred offence locations. These two tables give us a sense of the most and least frequent locations of crime in Toronto. If we look up locations on a map, we can see the neighborhoods in downtown Toronto have more crime comparing to other regions in Toronto. In particular, Waterfront communities in Toronto is notorious with surprising incidents. For example, the infamous ice condos have multiple incidents right near the buildings where there are thousands of residents. For individuals or families that place safety as top priority, they can get a sense of where in Toronto would be a safe place to live.

Location	n	Location	n
Waterfront Communities-The Island	8912	Lambton Baby Point	374
Church-Yonge Corridor	7855	Yonge-St.Clair	435
Bay Street Corridor	7237	Maple Leaf	447
West Humber-Clairvill	6520	Woodbine-Lumsden	454
Moss Park	5911	Guildwood	472
Kensington-Chinatown	4816	Markland Wood	515

Table 2: Top 6 Most/Least Occurred Offence Locations

3.3 Crime Counts

3.3.1 By Year

From 2014 to 2019, there is a trend of increasing number of offences. However, there is a dip in 2020. 2020 is the first year since 2014 which the number of offences is decreasing. This can due to the COVID-19 pandemic. Since people are more cautious and most people avoid going out in the public during the pandemic, there are less people going out to commit offences.
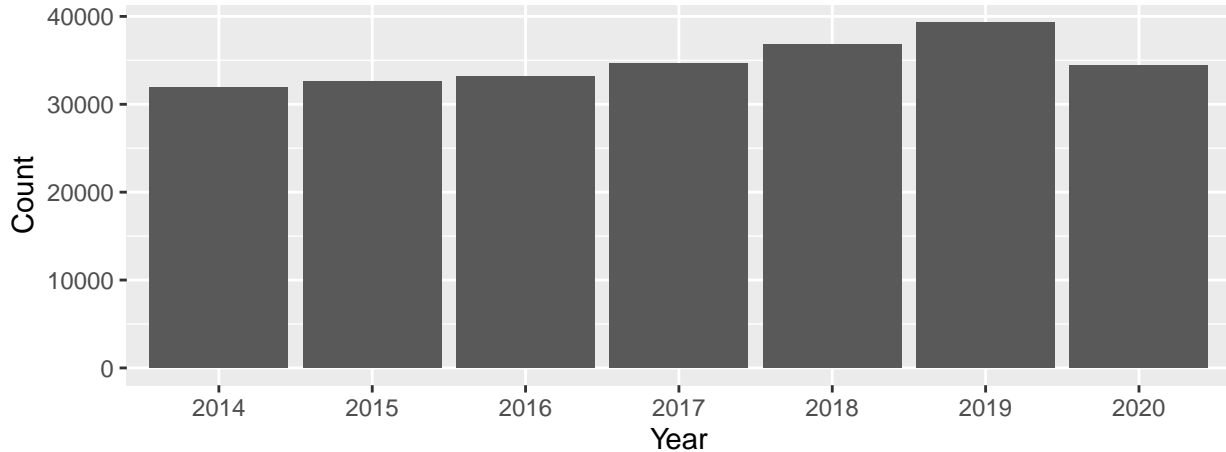

Figure 1: Crime Count By Year

3.3.2 By Month

From figure , it is difficult to conclude that there is a trend in the number of offences by month. We can see that there seems to be an increasing number of crime from the middle to late months (July to October) of a year.
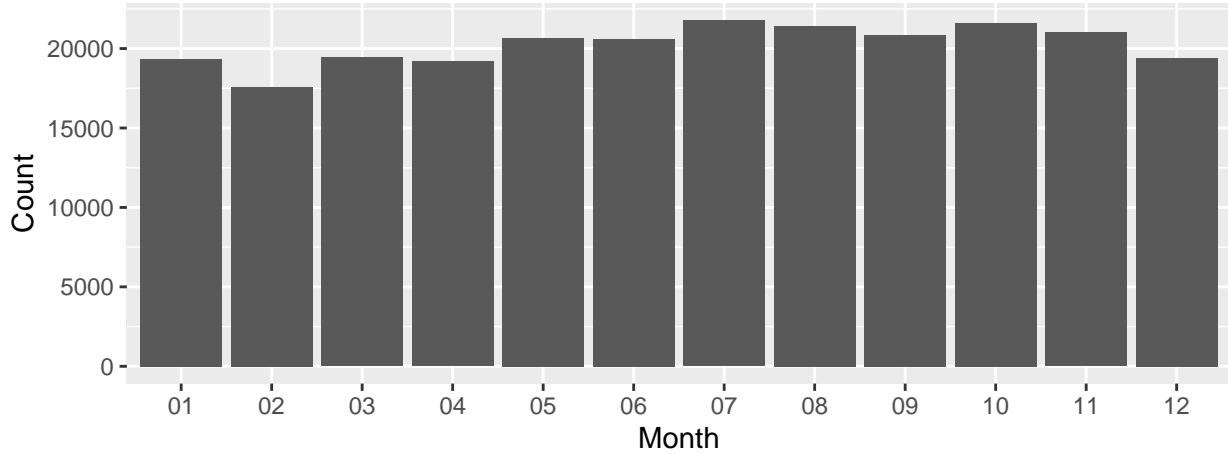
Figure 2: Crime Count By Month

3.3.3 By Day of the Month

The number of offences does not seem to change by the day of the month. We can see that offences seem to occur the most from the 15th - 20th of the month. There are also less offences towards the end of the month (29th - 31st). However, we can tell that this is due to around half of the month do not have a 31st day and February's do not have the 30th and the 31st and some 29th.
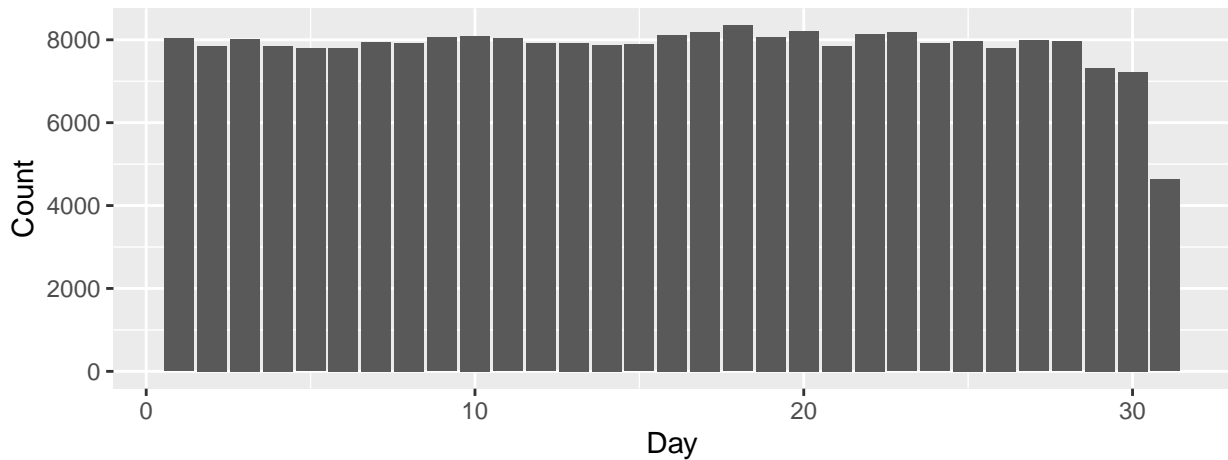

Figure 3: Crime Count By Day

3.4 Speed of Reporting offences

The speed of reporting offences can determine how soon polices respond to the situation and how soon the police can start investigation. From 2014 to 2017, the means days between the time of the offence occurred and the time of reporting the offence has shortened throughout the years. However, the mean days started increasing since 2017. There is another increasing trend from 2019 to 2020. Between 2019 to 2020, the increasing mean days can be due to the COVID-19 pandemic with police being less responsive because of the inconvenience COVID-19 brought.
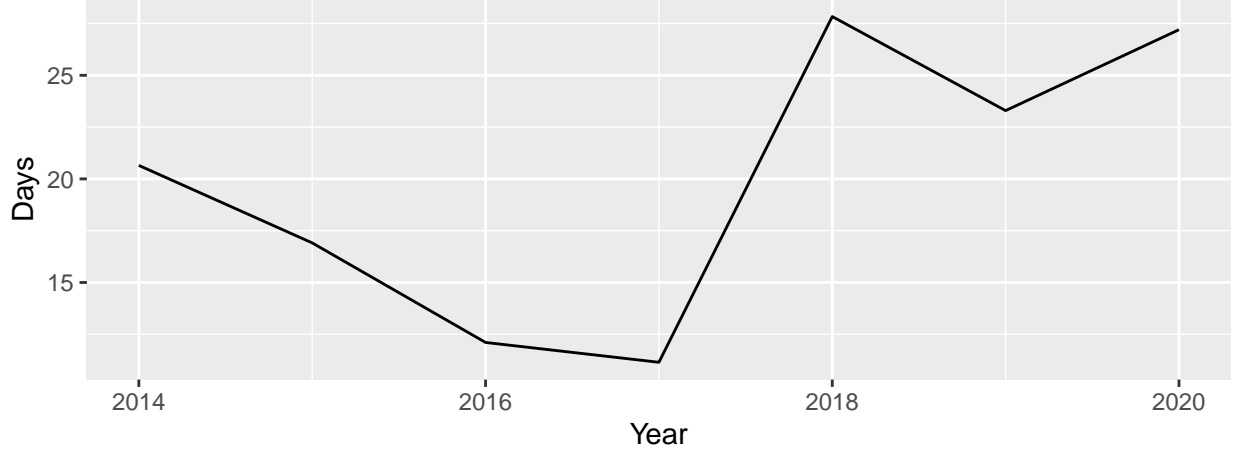
Figure 4: Mean Days between Reported Time and Occurence Time

4 Discussion

From the analysis above, we have a rough understanding of the major crime in Canada. This section will focus on the limitations identified when doing data exploration and analysis.

4.1 Offence Names

The "offence" offence column shows the offence categories categorized by Toronto Police. Using the column, it gives us a sense of the most and least occurred criminal offences in Toronto. However, we can see that there can be sub-categories for each category. For example, the category "Assault" can have sub-categories "Assault - Force/Thrt/Impede" and "Assault Peace Officer Wpn/Cbh." It can be difficult for the people who analyze the data to deal with this situation because they are not expert of criminal offence and they would not know if it is appropriate to put all these similar categories as one category. It can pose a challenge for the data analyst to consider what to do with these similar categories. In this report, we did not combine these categories together. However, if we want to dive deeper into this attribute, we will have to consult with expert in criminal offence to learn about what to do with these data.

4.2 Data Collection

As mentioned in section 2.5, the location of the criminal offence were classified to the nearest road intersection. As a result, the division and the neighborhood of each offence might not reflect the actual location where the offence occurred. In addition, the coordinate attributes, 'Long' and 'Lat' can be blank as a result of missing information of the location. Offence data is also not gathered for those that the police investigation never occurred or attempted. Thus, the data does not reflect an accurate picture of the major crimes in Toronto but it does give us a general sense of the criminal offence landscape.

5 Conclusion

In conclusion, this paper gave us a general sense of the criminal offence landscape in Toronto from 2014 to 2020. criminal offence has a trend of worsening from 2014 to 2019. However, the number of criminal offences decreased in 2020 which is also the year when the pandemic started. The number of criminal offences does not seem to change much throughout the months of the year and the days of the month. The paper also mentioned the most/least occurred criminal offences and the locations where there are the most/least criminal offences over the years. This can be helpful information to the individuals and families that put safety as their top priorities so they can make well-informed decisions to choose a safe neighborhood to live in.

6 Introduction

You can and should cross-reference sections and sub-sections. For instance, Section 2. R Markdown automatically makes the sections lower case and adds a dash to spaces to generate labels, for instance.

7 Data

Our data is of penguins (Figure 5).

```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```
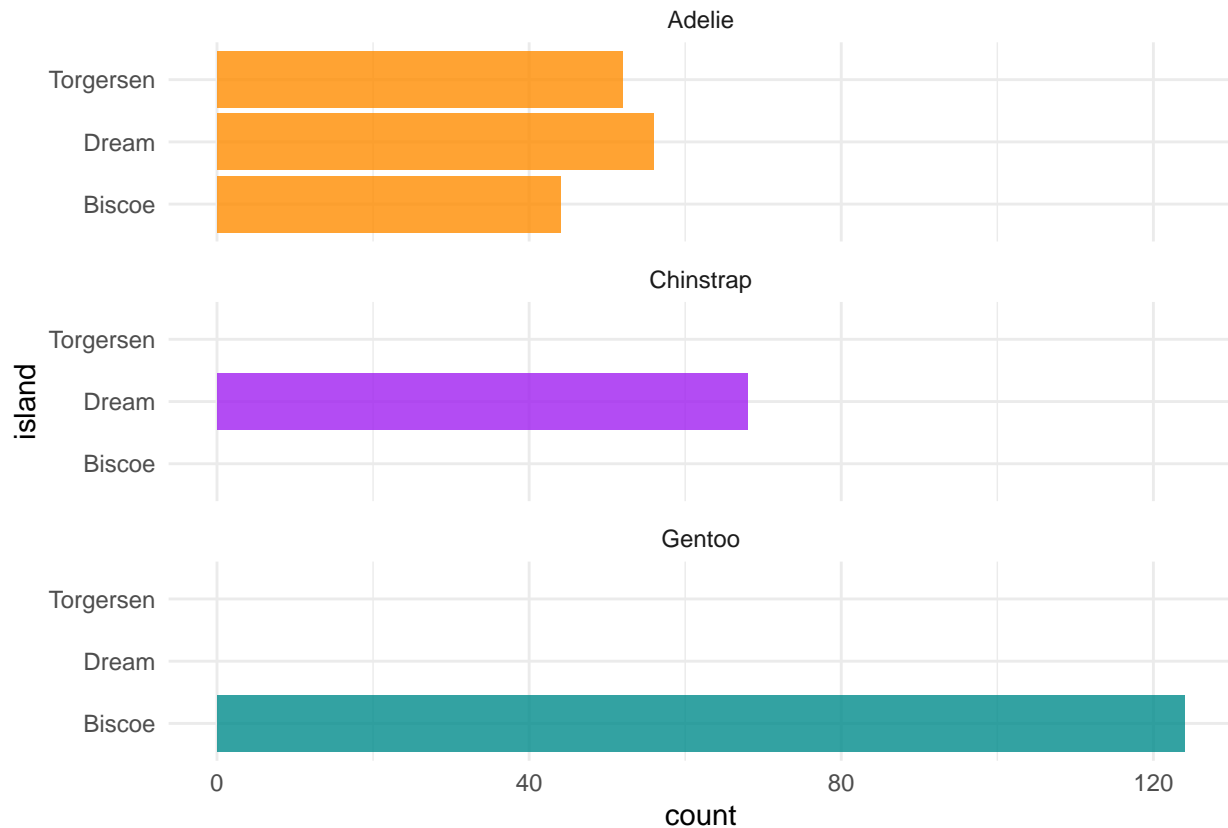

Figure 5: Bills of penguins

Talk more about it.

Also bills and their average (Figure 6). (Notice how you can change the height and width so they don't take the whole page?)

```
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
```

Talk way more about it.

8 Model

Here's a dumb example of how to use some references: In paper we run our analysis in R (R Core Team 2021). We also use the `tidyverse` which was written by Wickham et al. (2019) If we were interested in baseball data then Friendly et al. (2020) could be useful.
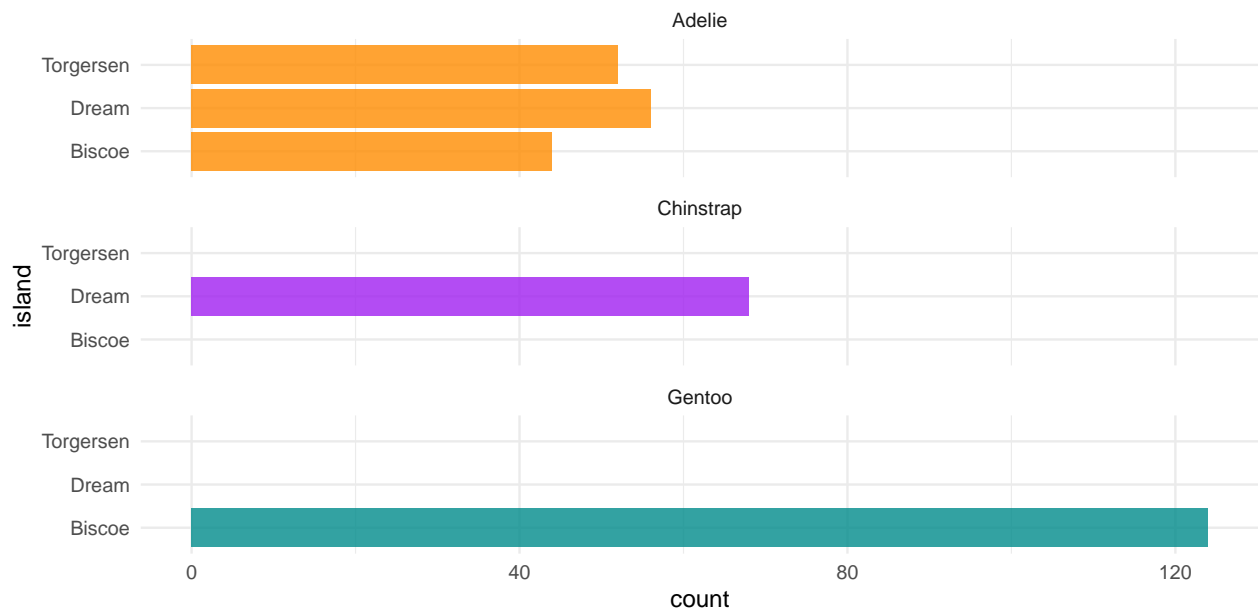
Figure 6: More bills of penguins

We can use maths by including latex between dollar signs, for instance θ.

9   References

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Canadians' Perceptions of Personal Safety and Crime. 2017. https://www150.statcan.gc.ca/n1/daily-quotidien/171212/dq171212b-eng.htm.

Canadians Think Toronto Is Way More Dangerous Than It Really Is, Thanks to Media Attention. 2017. https://nationalpost.com/news/canada/canadians-think-toronto-is-way-more-dangerous-than-it-really-is-thanks-to-media-attention-poll.

Friendly, Michael, Chris Dalzell, Martin Monkman, and Dennis Murphy. 2020. *Lahman: Sean 'Lahman' Baseball Database.* https://CRAN.R-project.org/package=Lahman.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Toronto, City of. 2020. *2019 Safe Cities Index.* https://www.toronto.ca/city-government/data-research-maps/toronto-progress-portal/world-rankings-for-toronto/2019-safe-cities-index/.

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.