Homocides in Toronto from 2004-2020*

-Justin Teng

06 February 2022

Abstract

Understanding the datasets of homicides in The City of Toronto Neighborhoods are crucial for determining the distribution of police resources among the neighborhoods. From we obtain this data of homicide cases in Toronto between the years of 2004-2020. We find that except the years between 2010-2015, we have relatively high number of homicide cases. Our findings have implication for allocation of police surveillance, police patrols, and the management of weapons registration.

1 Introduction

First paragraph is going to be motivational and broad.

Second Paragraph is about what was done and what was found.

Third Paragraph about implication.

The remainder of this paper is: Section 2 explains the data. Section 4 covers results

2 Data

Paragraph or two introducing the dataset broadly.

Read and clean the required data for my exploratory data analysis.

```r
library(tidyverse)
library(knitr)

homicide <-
  read.csv(here::here("inputs/data/homocide.csv"))

# Clean out the unnecessary variables
cleaned_homicide_data <-
  homicide |>
  select(Occurrence_year,
         Homicide_Type,
         Neighbourhood)
```

Then show an extract of the dataset (Table 1) .

---

*Code and data are available at: https://github.com/justinteng1999/starter_folder-main

Table 1: First ten rows of a dataset of a dataset that shows homicide occurrence

Year of Occurrences,	Neighborhood of Occurences	Type of Homicide
2004	Yonge-St.Clair (97)	Other
2004	Woburn (137)	Shooting
2004	Malvern (132)	Shooting
2004	Dovercourt-Wallace Emerson-Junction (93)	Shooting
2004	Rouge (131)	Shooting
2004	Downsview-Roding-CFB (26)	Stabbing
2004	Downsview-Roding-CFB (26)	Shooting
2004	Mount Olive-Silverstone-Jamestown (2)	Other
2004	Malvern (132)	Shooting
2004	Clairlea-Birchmount (120)	Shooting

```
cleaned_homicide_data |>
  select(Occurrence_year, Neighbourhood, Homicide_Type) |>
  slice(1:10) |>
  kable(
    caption = "First ten rows of a dataset of a dataset that shows homicide occurrence",
    col.names = c("Year of Occurrences,",
                  "Neighborhood of Occurences", "Type of Homicide"),
    digits = 1,
    booktabs = TRUE,
    linesep = ""
  )
```

Paragraph or two more about Table 1.

We are interested in the number of homicide occurrence each year from 2004-2020, and the neighbourhoods with high number of homicide cases, as well as the type of homicide in each case. This is the measure of how homicide cases are distributed and has implications for allocation of police resources.

(Figure 1) shows the number of homicide occurrence each year from 2004-2020.

```
tibble_occurrence_year <-
  tibble(cleaned_homicide_data) |>
  count(Occurrence_year)

tibble_neighbourhood <-
  tibble(cleaned_homicide_data) |>
  count(Neighbourhood)

tibble_homicide_type <-
  tibble(cleaned_homicide_data) |>
  count(Homicide_Type)
```

(Figure 2) shows the top 10 neighbourhoods with most number of homicide cases

(Figure 3) shows number of each homicide type

Our data is of homicide cases (Figure 4).

```
## Warning: It is deprecated to specify 'guide = FALSE' to remove a guide. Please
## use 'guide = "none"' instead.
```
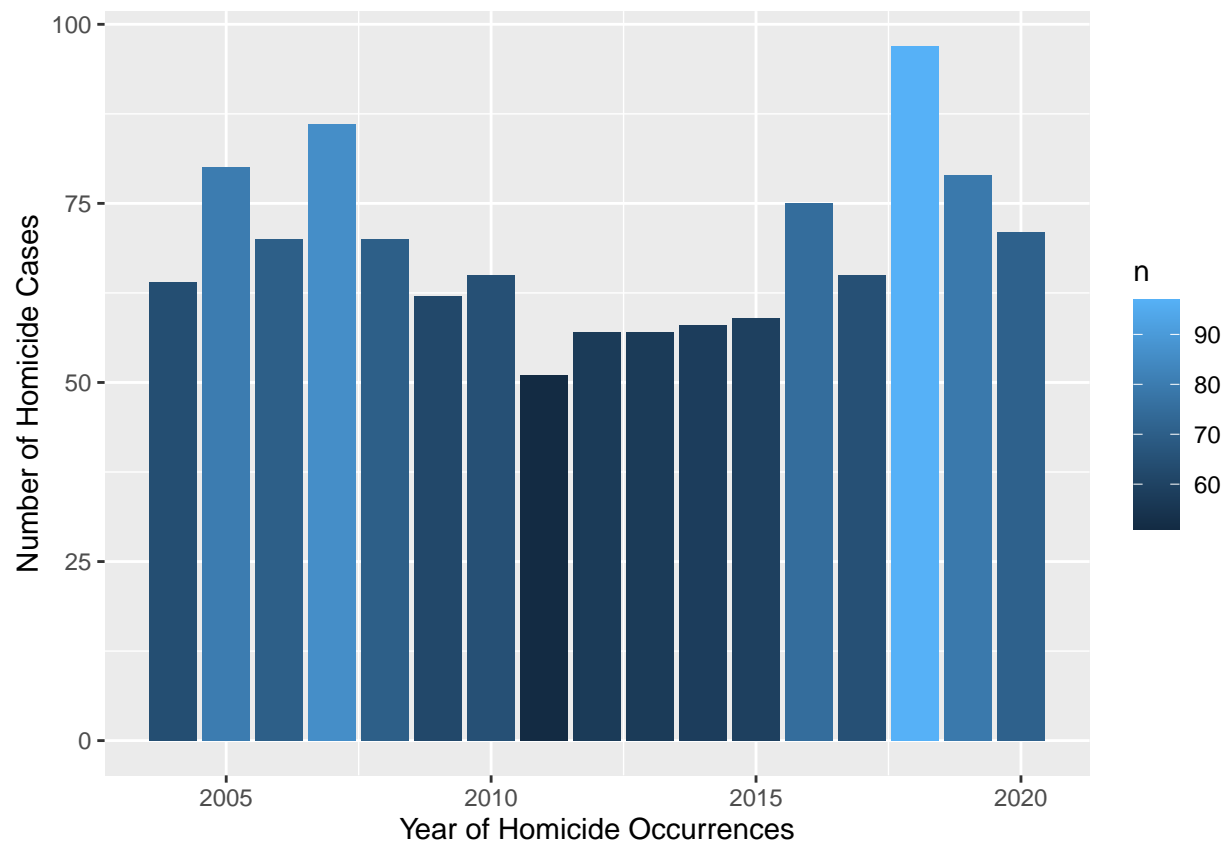
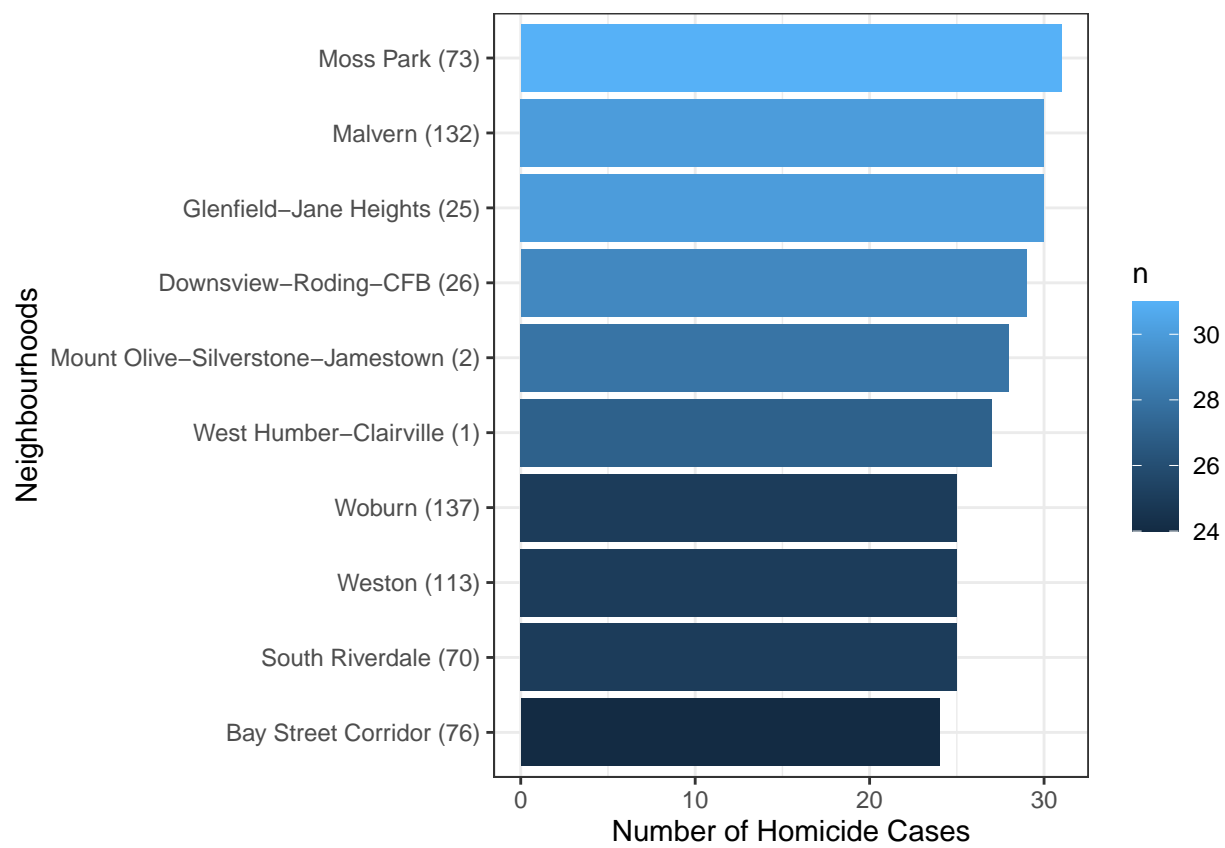Figure 1: Homicide Occurrences each Year(2004-2020)

Figure 2: Top 10 Neighbourhoods with Most Homicide Occurrences (2004-2020)
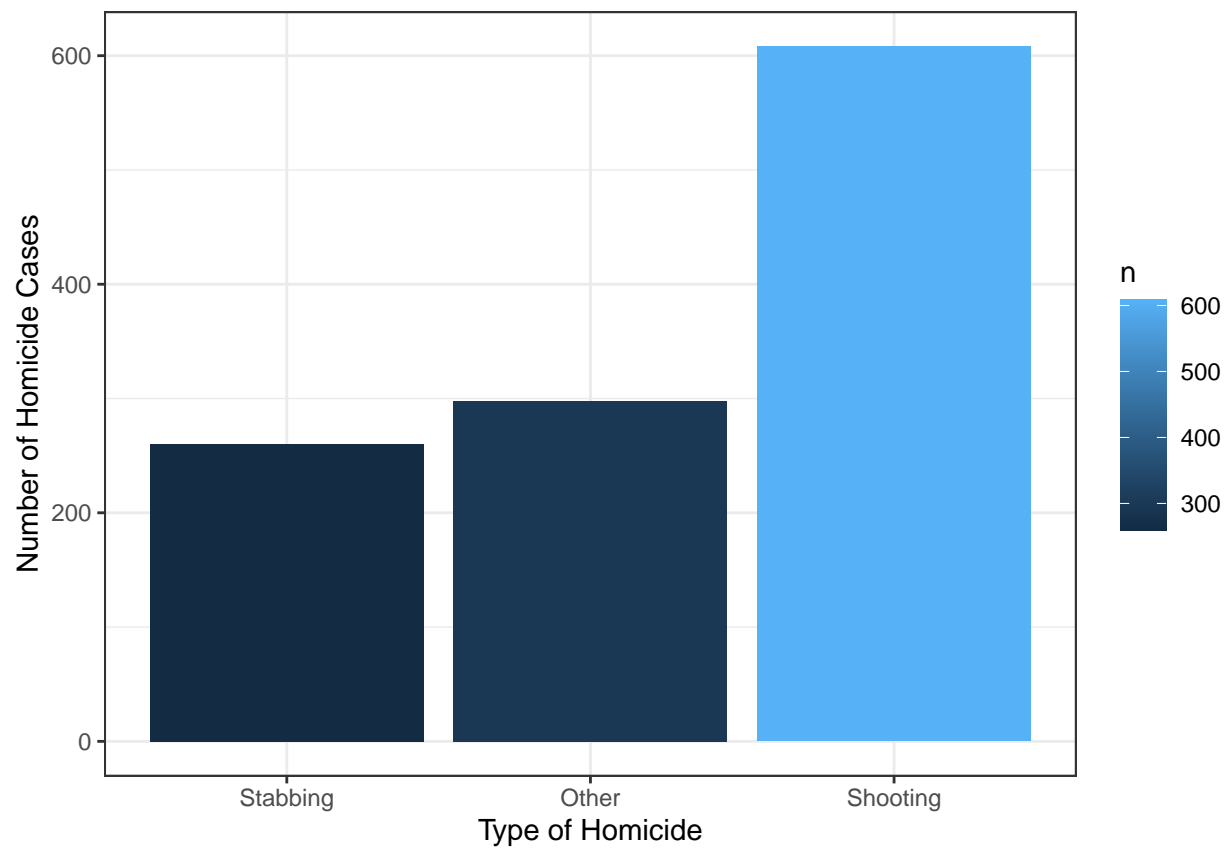
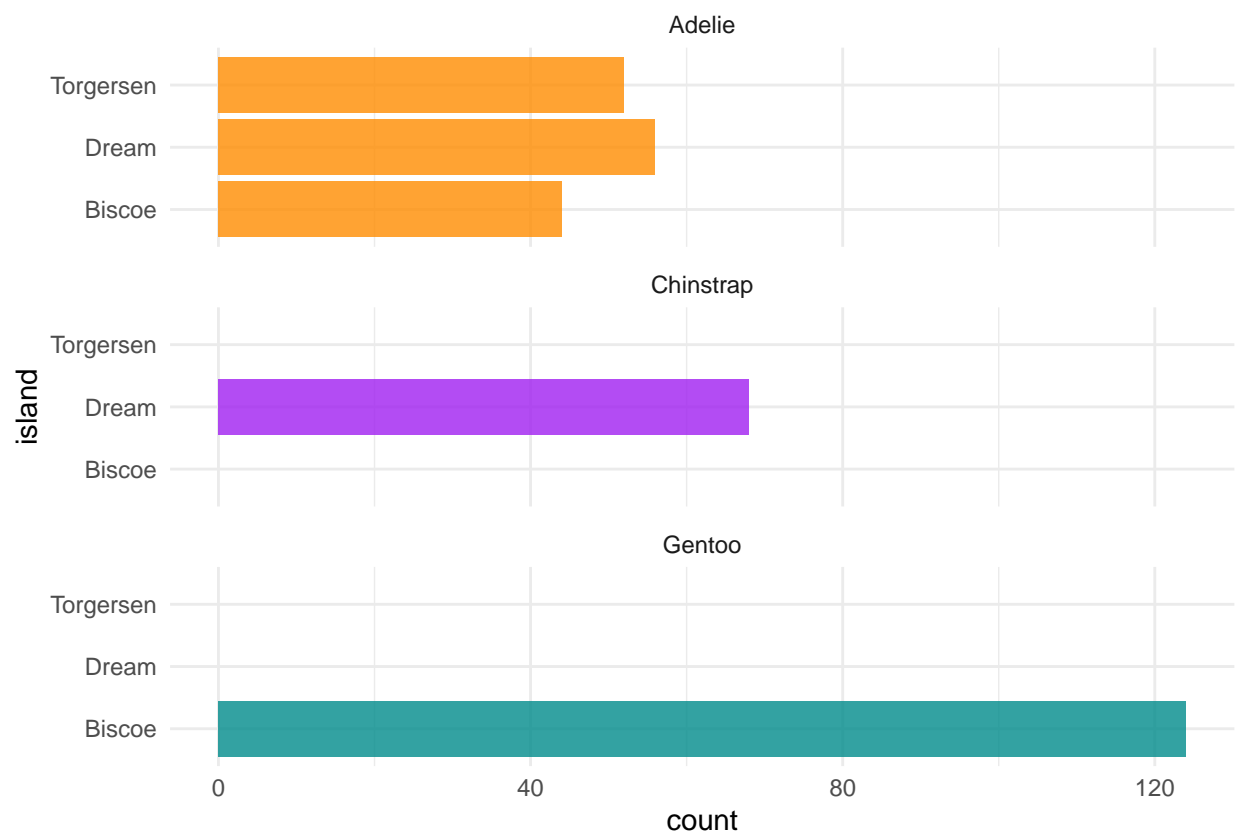Figure 3: Type of Homicide Cases

Figure 4: Bills of penguins

Talk more about it.

Also bills and their average (Figure 5). (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.
```
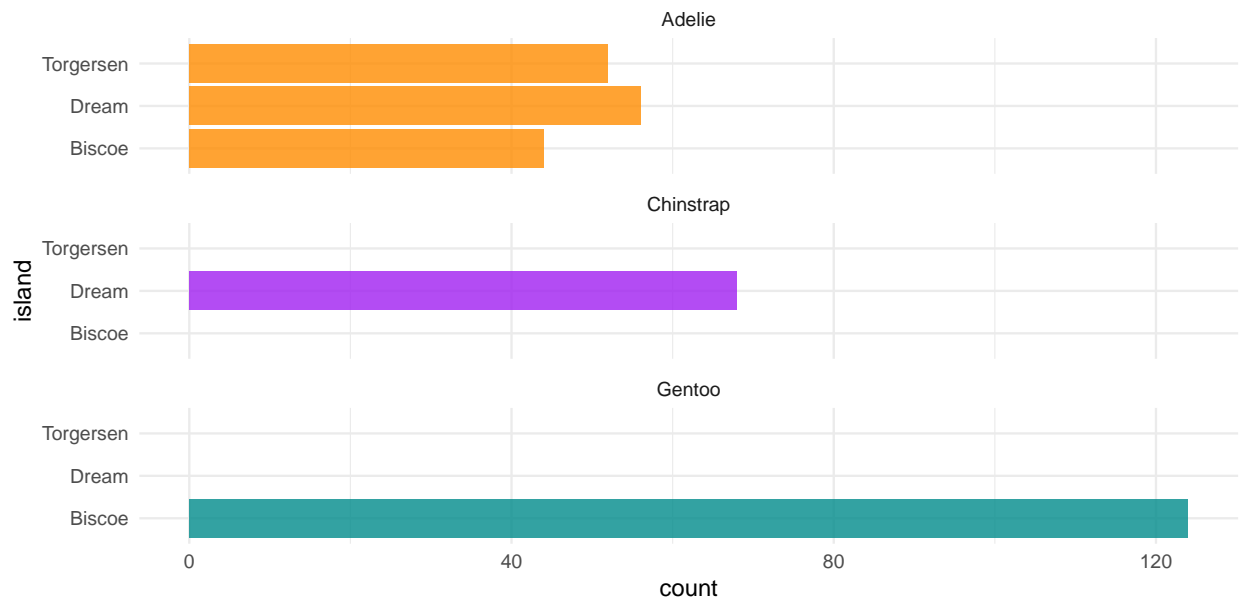
Figure 5: More bills of penguins

Talk way more about it.

3 Model

$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{Pr(y)} \tag{1}$$

Equation (1) seems useful, eh?

Here's a dumb example of how to use some references: In paper we run our analysis in R (R Core Team 2020). We also use the tidyverse which was written by (**thereferencecanbewhatever?**) If we were interested in baseball data then Friendly et al. (2020) could be useful.

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

4 Results

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A   Additional details

References

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. 2020. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.