Reducing the Risk of COVID-19 Infections in Vulnerable ages*

Mohid Sharif

04 February 2022

Abstract

COVID-19 has taken over the world for more than two years now and it is important to study and understand in which circumstances cases rise and fall. I studied and analyzed data obtained about COVID-19 cases from various communities in Toronto specifying age, gender, and source of infection of the patient. I found that younger age groups are more likely to be infected in their household while the elderly are more likely to be infected in community settings. My findings can be used to identify circumstances in which outbreaks are less likely for certain age groups.

1 Introduction

COVID-19 has taken over the world for more than two years now. Thousands of cases have been recorded in this data set and we can use this to our advantage. Analyzing the data set we can compare variables such as age groups and gender to isolate sources of infection which are more likely to infect certain groups and find changes to those circumstances to reduce the likeliness of infection in those groups. Using this technique we can keep making adjustments to what should and should not be allowed during the pandemic to ensure cases don't rise uncontrollably.

I first analysed the data set to understand how cases had been recorded. Every case in the data set records the source of infection, age group, and gender of the patient. These are only a few of the variables but they are also the most important and they are the ones I mainly looked at in this data study. I then created two graphs to compare variables in the data set. The first graph is a comparison of source of infection between age groups, I used these graphs to visualize which age groups were more likely to be infected in certain settings. I found that younger ages (0-19 years) were more likely to be infected in households while elderly people are more likely to be infected in community settings or healthcare institutes. The second graph compares the source of infection between gender groups, I used these graphs to analyze which gender groups were more likely to be infected in certain settings. I found that both males and females had almost identical data for how they were being infected, however there were twice as many females being infected in healthcare institutes.

I can use my data to isolate situations in which infections are more likely for certain groups and how we can change that to reduce infections. Firstly younger ages are more likely to get infected in their homes, in this case they are likely infected by a family member. Family members should in this case sanitize regularly while entering and exiting the house and reduce outdoor activity to necessities. Secondly elderly are more likely to be infected in healthcare institutes because of regular visits for their health. For this healthcare institutes should do their best to isolate the non-infected elderly from the infected reducing their risk for infection. Lastly females are more likely then men to be infected in healthcare institutes due to the fact that women dominate nursing, otherwise both males and females are just as likely to be infected in any setting.

The remainder of this paper is: Section 2. R Markdown automatically makes the sections lower case and adds a dash to spaces to generate labels, for instance, Section 5.1.

*Code and data are available at: https://github.com/MohidSharif/COVID19_data_paper

Table 1: Outbreaks in Agincourt North

Source of Infection	Age Group	Gender
Travel	30 to 39 Years	FEMALE
Outbreaks, Other Settings	50 to 59 Years	FEMALE
Outbreaks, Congregate Settings	20 to 29 Years	MALE
Community	50 to 59 Years	MALE
Outbreaks, Other Settings	30 to 39 Years	MALE
Outbreaks, Healthcare Institutions	30 to 39 Years	FEMALE
Close Contact	30 to 39 Years	FEMALE
Outbreaks, Healthcare Institutions	19 and younger	MALE
Outbreaks, Healthcare Institutions	50 to 59 Years	FEMALE
Outbreaks, Healthcare Institutions	20 to 29 Years	FEMALE

2 Data

Paragraph or two introducing the dataset broadly.

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

```
## Warning: package 'knitr' was built under R version 4.1.2
```

(Figure 1) shows the relationship between age and source of infection.

Talk more about it.

Also bills and their average (Figure 3). (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.

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 Wickham et al. (2019) 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.

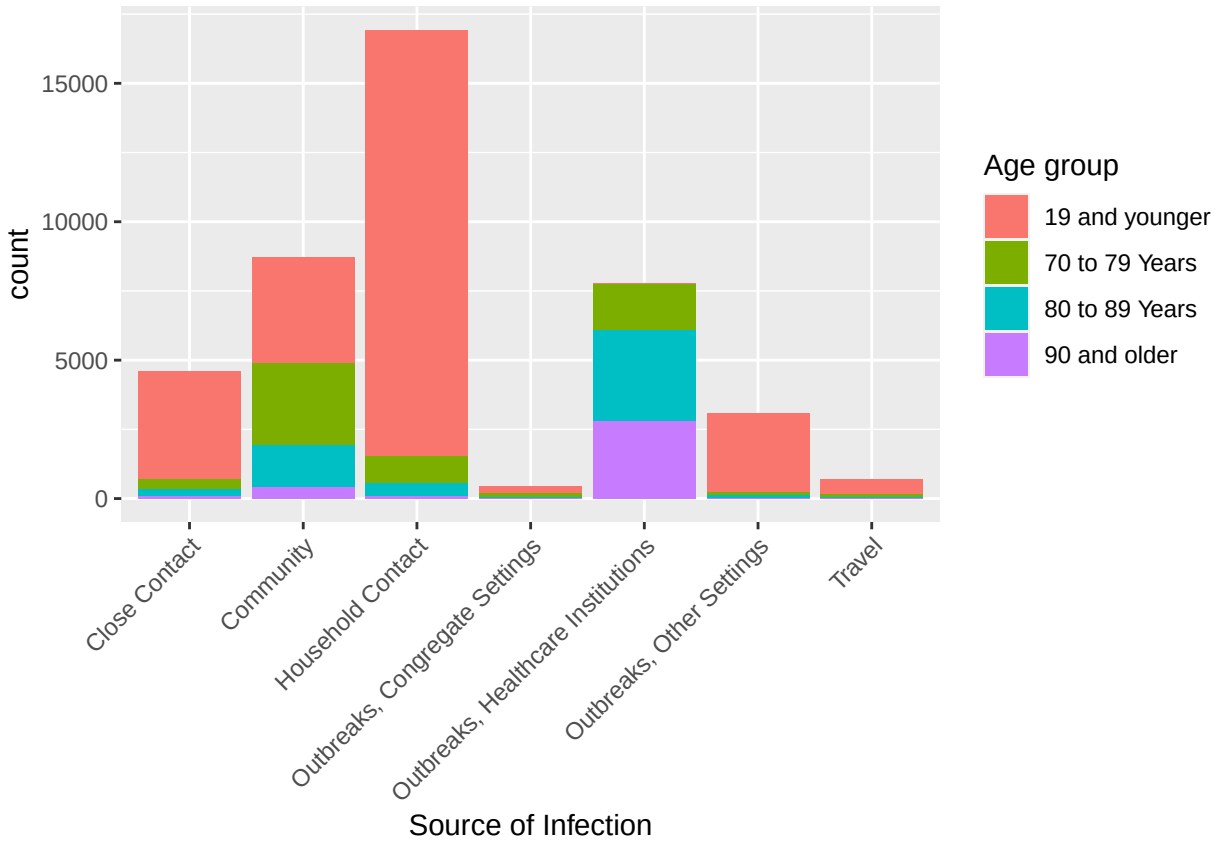

Figure 1: Source of Infections by Age group

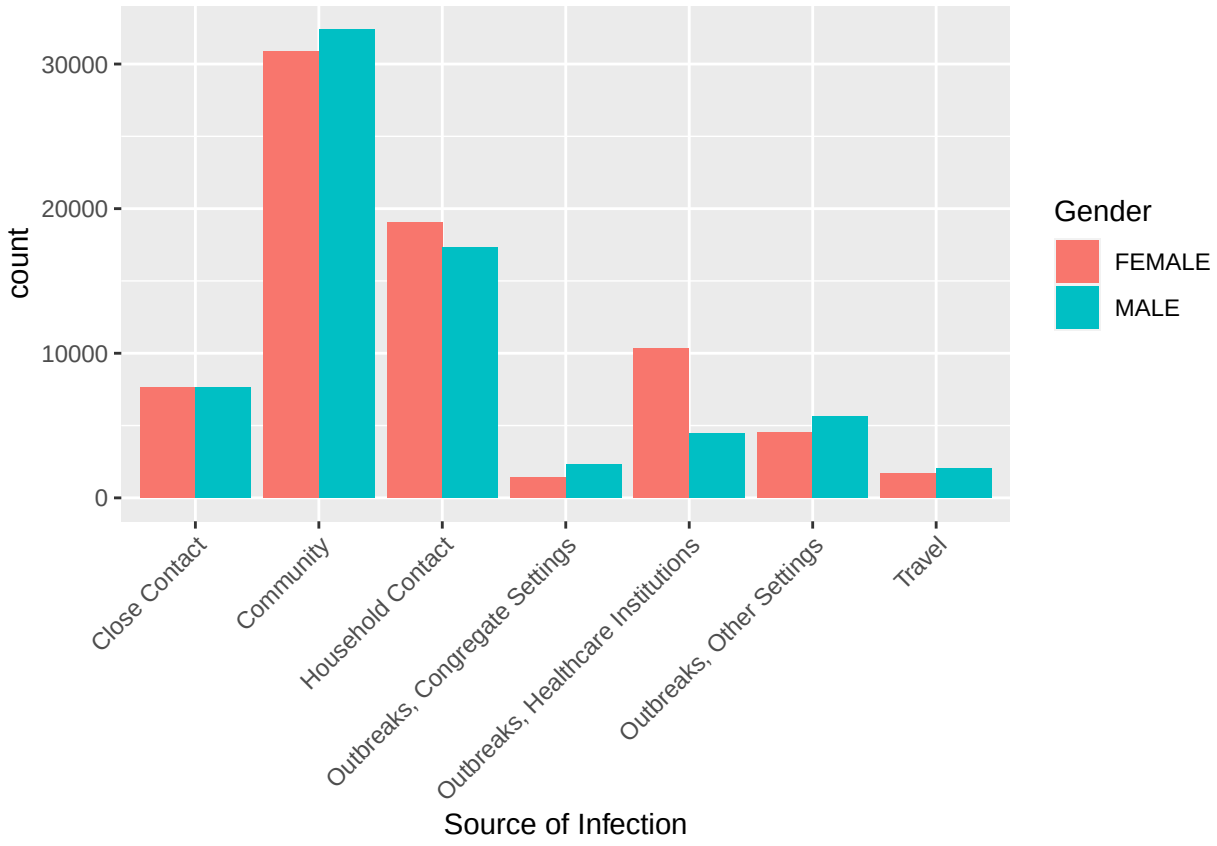

Figure 2: Source of Infections by Gender

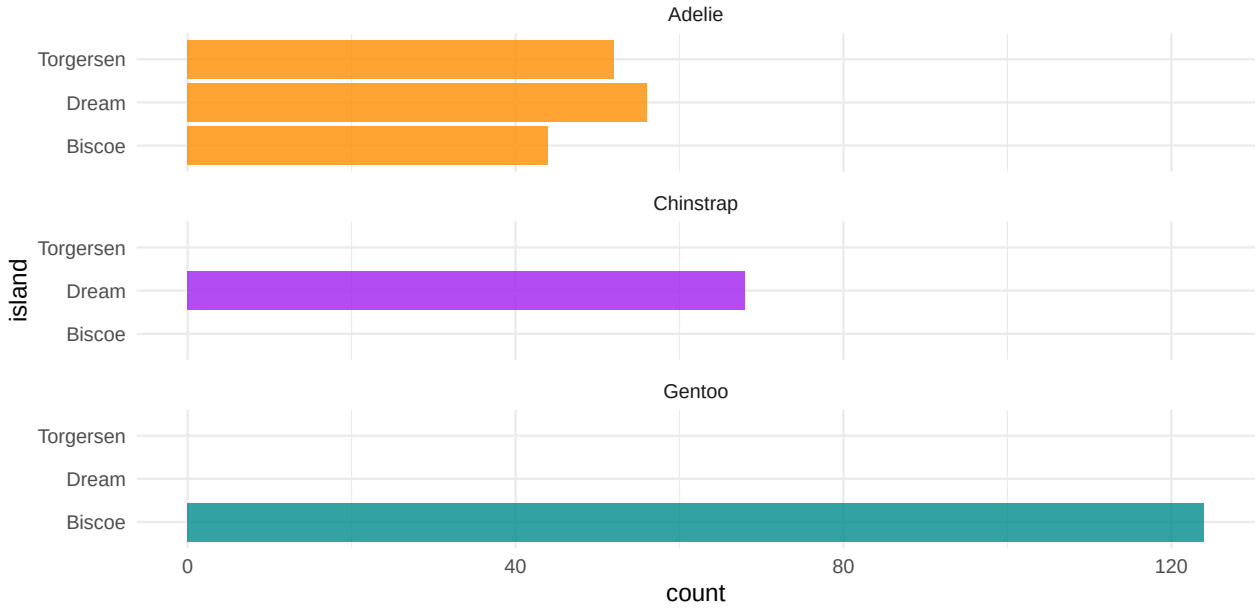

Figure 3: More bills of penguins

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/.

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.