JSC370 final report: COVID19 analysis
Wei Yu

1 Introduction

Our dear campus, University of Toronto St. George has only suspended its in-person courses 5 times during its hundreds of years of history: the Great Fire of 1890 at University College, World War I, World War II, a heavy snowstorm in early 2019, and the COVID-19 pandemic. Beside the heavy snowstorm in early 2019 which only caused the university to close for one day, all the other four led to great influence on campus life, where we witnessed the history. What needs attention is that COVID-19, not only has an important and serious influence in our campus life, but also has made a world-wide large influence for its duration and propagation. Specifically in this project, I focused on local data of Ontario to gain more insight for the virus and its influence for different people. As advised and guided by the course website, I got covid19 data from Toronto open data, an open data initiative by the City of Toronto government and whose website was launched at the Toronto Innovation Showcase forum. I also used some data from https://www.ontario.ca/page/covid-19-coronavirus, where official data were officially posted by the Government of Ontario. I want to answer the following questions: 1.When and where were people more likely to get covid19? 2. What builds the ‘‘luck‘‘(not getting covid or mild severity if get covid19 or survive from extremely severe symptoms): age (what range), vaccination status (not vaccinated? partially vaccinated? fully vaccinated?), life habits (by source of infection, for example, if behaviors involving crowding contributes a lot, then ‘‘otaku‘‘becomes an ‘‘advantage‘‘.

2 Methods

2.1 Data Collection

As advised and guided by the course website, some data were from Toronto open data about covid19 cases, *with* columns *including_id*, *Assigned_ID*, Outbreak Associated, Age Group, Neighbourhood Name, FSA, Source of Infection, Classification, Episode Date, Reported Date, Client Gender, Outcome, Ever Hospitalized, Ever in ICU, *EverIntubated* from https://open.toronto.ca/dataset/covid-19-cases-in-toronto/. I want to explore the relationship between the severity of covid19 of each patient and the age, source of infection, neighbourhood, etc. (I plan to evaluate the severity by if ever hospitalized, if ever in ICU, if ever intubated, which I want to assign some value by danger level and use the sum to be the severity of each patient). However, I do not think that gives me the full picture especially since I still need to filter out some data like I will only use confirmed data without the ‘‘probable‘‘ones, and Toronto open data does not offer much other very meaningful data that I can use. Thus, I want to use some more data at https://www.ontario.ca/page/covid-19-coronavirus, from which https://data.ontario.ca/ has extracted some data but still the raw data and information are fragmented. I want to combine the information from different data including the hospital admissions, data change in cases by districts (and I want to compare the difference between districts with the average vaccination rates group by districts), data of patients requiring long-term care after getting covid etc.

2.2 Data cleaning wrangling

I filter out the cases lacking critical information, like I filtered out the cases with "No Information" for the source of infection. Most analysis for the midterm project is qualitative analysis, with visualization of graphs, but there is also quantitative analysis used here like with chi-squared test, tables, and directly printed-out values. Since data from different sources can offer a fuller picture, I merge some dataset together for further analysis (usually merge by date). For example, I merge the proportion of hospitalization and icu-treatment that are due to covid for each date with the death data at each date to analyze for the relationship between the dominance of covid19 and number of death cases. Also, I created new columns for level scores for multiple usage. For data points with special non numeric values, I use the representative numeric value to represent and replace those values for convenience of later visualization and analysis. For example, when dealing with the long-term care outbreak data, I want to use values of $Total\_LTC\_HCW\_Cases$ to create an evaluation for the severity of the outbreak and I want to investigate the relationship between $Total\_LTC\_Resident\_Cases$ and $Total\_LTC\_Resident\_Deaths$, but all of them have values of < 5, which makes it hard to plot. Thus, for simplicity I use the representative medium 2.5 of that group (1, 2, 3, 4) to replace the value of < 5. Then I change those character columns $Total\_LTC\_HCW\_Cases$, $Total\_LTC\_Resident\_Cases$ and $Total\_LTC\_Resident\_Deaths$ into numeric. To evaluate the severity of each outbreak, I created a new column in the long-term care outbreak dataset named *hcw_level* denoted by 6 to 1 as the most serious to the lightest. The *hcw_level* is defined in the way that: if the number of active COVID-19 positive cases among staff is no fewer than 50, denote the outbreak as the most serious, by 6; if the number of active COVID-19 positive cases among staff is fewer than 50 and no fewer than 40, *hcw_level* is set as 5; if the number of active COVID-19 positive cases among staff is fewer than 40 and no fewer than 30, *hcw_level* is set as 4; if the number of active COVID-19 positive cases among staff is fewer than 30 and no fewer than 20, *hcw_level* is set as 3; if the number of active COVID-19 positive cases among staff is fewer than 20 and no fewer than 10, *hcw_level* is set as 2; otherwise, *hcw_level* is set as 1. Details are included in later analysis.

2.3 Data Exploration

I used R and packages & functions in R.

2.3.1 Preliminary Results

To begin with, I add a new column to denote the danger level of each patient after getting covid, and the danger level is calculated based on if the patient has been hospitalized, in icu, or incubated. By taking the first glance, I see more "dangerous" cases for age groups with higher ages. Thus, I perform a chi-squared test to examine if there is a significant difference between different age groups for the danger of infection of covid19 (credits and many thanks to Professor Franklin): chi-squared test for difference of the danger score of infection of covid between different age groups. Null hypothesis: H0: there is no significant difference between different age groups for the danger level of infection of covid19; alternative hypothesis: Ha: there is significant difference between different age groups for the danger level of infection of covid19.

```
        Pearson's Chi-squared test

data:  table(case_data$Age.Group, case_data$danger_score)
X-squared = 33846, df = 27, p-value < 2.2e-16
```

Figure 1

Since the p-value is too small, I can strongly reject the null hypothesis and take the alternative hypothesis. Thus, there is significant difference for the danger level

of infection of covid19 between different age groups.

Since the danger score is defined by me, to getting rid of objective factors, I perform chi-square tests for each of the factors contributing to the danger level, i.e., the three standards that measure the danger which the patients go through after infection of covid19: if ever hospitalized, if ever in icu, and if ever intubated. 1. test for the difference for if ever hospitalized between different age groups: Null hypothesis: H0: there is no significant difference between different age groups for the possibility of ever being hospitalized after infection of covid19; alternative hypothesis: Ha: there is significant difference between different age groups for the possibility of ever being hospitalized after infection of covid19. 2.test for the difference for if ever in ICU between different age groups: Null hypothesis: H0: there is no significant difference between different age groups for the possibility of ever in ICU after infection of covid19; alternative hypothesis: Ha: there is significant difference between different age groups for the possibility of ever in ICU after infection of covid19. 3. test for the difference for if ever intubated between different age groups: Null hypothesis: H0: there is no significant difference between different age groups for the possibility of ever intubating after infection of covid19; alternative hypothesis: Ha: there is significant difference between different age groups for the possibility of ever intubating after infection of covid19.

Results of Pearson's Chi-squared test

	X-squared	df	p-value
for ever hospitalized	31566	9	$< 2.2e-16$
for ever in icu	4718.5	9	$< 2.2e-16$
for ever intubated	2927	9	$< 2.2e-16$

Since p values for all of the three p values are very small, I can strongly reject the null hypotheses and take the alternative hypotheses. Thus, there is significant difference for the possibility of being hospitalized, in icu and incubated after infection of covid19 between different age groups, i.e. different age groups have different level of risk/danger after infection of covid19.

Since age does make a difference on the influence of covid, I will do more analysis on age. Since each age group is a range of age, I want to represent each group by one age value for convenience of analysis. For each group in the middle, I choose the representative age value as the "middle", i.e., the average of the upper bound and lower bound of each group; for the youngest group of people, i.e. "19 and younger", I choose 19 as the representative value and for the group of people with largest age, i.e. "90 and older", I represent the age by 90, since the actual lower and upper bound for this two groups are unknown, like 0-year-old babies are not that likely to be included in this investigation and if the upper bound for the age group with largest group is extremely big (like 200 years old), then the middle age will not be representative.

Filter the data to remove NA in the representative age (from checking, the NA is and only is due to missing data at Age.Group, which is very important information for analysis here and should not be missing. The missing data here makes it look very unreliable.)

After wrangling the age group value into more convenient for analysis, I summarize the (representative) age value for different value of danger score (I calculated under the standard of if ever hospitalized, if ever in icu, if ever intubated).

A tibble: 4 × 6

danger_score <dbl>	min_age <dbl>	max_age <dbl>	mean_age <dbl>	median_age <dbl>	mode_age <dbl>
0	19	90	41.80824	34.5	24.5
1	19	90	67.94384	74.5	84.5
6	19	90	63.15704	64.5	64.5
16	19	90	63.75254	64.5	64.5

4 rows

Figure 2

Thus, it can be viewed that for the team with no danger (i.e. the people who have not ever in hospital, icu or intubated after getting covid) whose mean age (about 42), median age (about 34.5) are both young, and the most number of people who were brought no

danger by covid after infection had representative age even younger, as only 24.5, i.e. in the age group of "20 to 29 Years", which should be the most safe group among all the age groups. For all the non-zero danger levels caused by covid, the mean, median and mode of representative age are all large, i.e., all kinds of "danger" caused by covid prefer to seek the senior citizens. From the summary above, I think it is necessary for the senior citizens to receive extra protection and care for medical and health to go through the potential risk of the left covid viruses. At the same time, I think it worth noticing that all kinds of level can happen to all the age groups since the min and max for the representative age at each group are all 19 and 90, which means that even young people with age "19 and younger" (no kids are counted here since the data were collected starting at 2020) who should be safe can have the most danger score, i.e. having a super dangerous bundled journey of being hospitalized, ICU-treated and intubated, and senior citizens with age "90 and older" can have little influence under infection of covid: even without need to see a doctor. Thus, this should remind us that although they are the senior citizens who should be given the most protection and care, protection measures should not be totally relieved from any age group of people.

The communities with the top 30 number of covid-19 cases are:

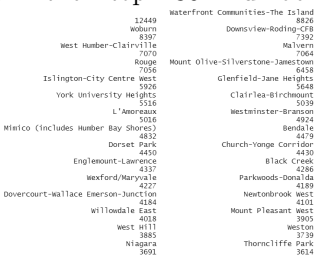

Figure 3

For the most frequent ones, Waterfront Communities-The Island have a large portion of senior citizens, Woburn is a family oriented neighbourhood, and other communities are with similar traits. Thus, it can be viewed that the communities with the highest frequency of covid19 cases are either with a larger portion of vulnerable people or with higher interaction inside. Thus, probably people who live alone are safer to keep away from covid19.

2.4 Results

In Figure 4, the red color denotes how the proportion of covid19 cases in the hospitalized cases change by time and the green color denotes how the proportion of covid19 cases in the ICU-treated cases change by time.

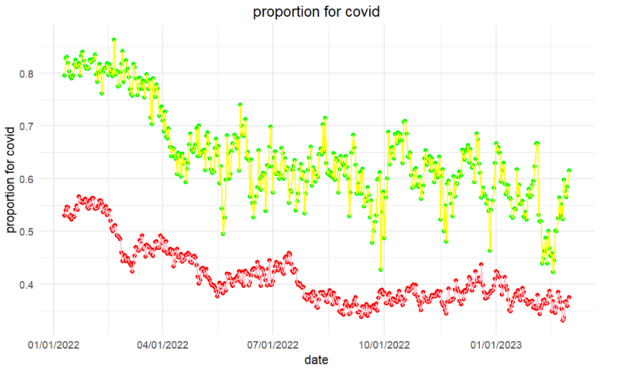

Figure 4

The overall trend is decreasing for both hospital and icu for covid. The lower the proportion is, the more "normal" covid becomes: which means people are back to normal at life and things are better. Things are better as time pass by overally and the highest portion of covid for hospital happened at around Febrary, 2022, and dropped after but the proportion of icu treatment for covid did not drop did not drop around April. I guess it was due to the cancel of some travel restrictions at Canada border: As of February 28, 2022 at 12:01 a.m. EST, instead of quarantine, fully-vaccinated travellers, were only be randomly selected for arrival testing. My hypothesis is that they were likely to get covid, which is however not dangerous to them, but the possible propogation of possible viruses carried by them can be dangerous to vurnerable people.

Only consider the observations with known source

of infection, i.e., filter out all the observations without known source of infection.

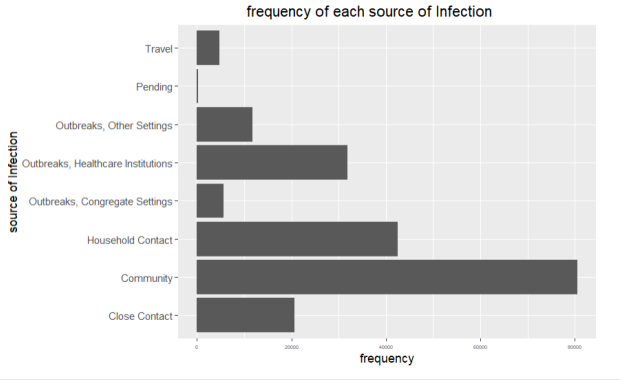

Figure 5

From the output of codes, only 0.02389096 of the cases have Source.of.Infection as "Travel". Thus, travelling only consists of a small portion of the source of infection (only about 2%), so the travelling restriction shouldn't be a huge problem. Thus, I think the proportion of cases for covid in both hospital and icu in around February is not due to the relief of the travel restriction, but due to the seasonality (credit and many thanks to Professor Franklin!): it can be viewed that in winter a larger portion of cases treated in hospital and icu is for covid. From the boxplot, it can be viewed that it is most likely to get covid infection in the community, i.e., avoiding crowding in community activities is a good way to stay from infection.

The above restriction relief is for people who are fully-vaccinated and I think it makes sense since they should be protected from vaccinations. I will then investigate the level of protection given by the vaccinations.

In the Figure 6, red color denotes cases of people who are ‘‘unvaccinated‘‘for covid19, defined as not having any dose, or between 0-13 days after administration of the first dose of a COVID-19 vaccine; green color denotes cases of people who ‘‘have been partially vaccinated for covid-19‘‘, defined as 14 days or more after the first dose of a 2-dose series COVID-19 vaccine, or between 0-13 days after administration of the second dose; blue color denotes cases of people who ‘‘have been fully vaccinated for covid-19‘‘, defined as 14 days or more after receipt of the second dose of a 2-dose series COVID-19 vaccine. The graph on the left shows how the frequency of hospitalized cases in Ontario change by time and more specifically for the cases with the most danger, the graph on the right shows how the frequency of the ICU-treated cases in Ontario change by time. Please note that though some curves with the same color (which means for the same group of vaccination status of people) look somewhat similar on the left and on the right, they still differ much in frequency, since the range for each unit length is much larger for the visualization of the full picture of hospitalization.

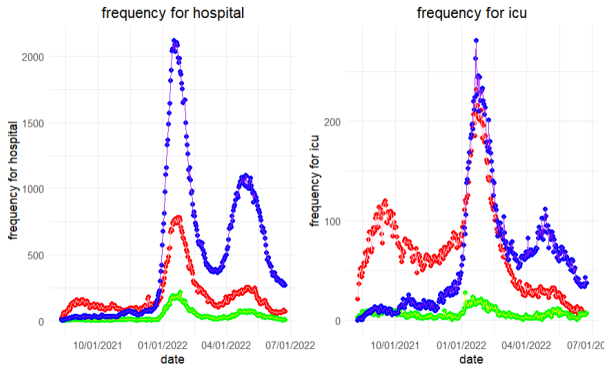

Figure 6

Thus, looking like partially vaccinated are the most safe, but I do not think it is necessary that partial vaccination works better than full vaccination since the current data cannot rule out the possibility that people who are fully vaccinated include those who are the most vulnerable since those people have to seek protection from vaccination as much as possible, and another possibility is that some people may go to hospital because of feeling ill after getting the third vaccination: specifically, since some people who are the most vulnerable to covid viruses show seriously bad response to the third dose, which sent them into hospital and was

counted "for the reason of covid", the effect of the fully vaccination include the individual difference covering those people, so I cannot conclude that partial vaccination is the best, better than fully vaccination. Thus, I think although it can be viewed from the graphs above that partially vaccinated people go to hospital / icu the least frequently, I cannot conclude that partial vaccination is better than fully vaccination. However, I think it is safe to conclude that partial vaccination is better than no vaccination. I will further research on the effect of different kinds of vaccination in my final project.

This graph also verify the seasonality of covid19: there is a huge jump in frequency for both hospital and icu in the winter, which probably means that covid viruses love cold weather and can survive better in lower temperature. Also, according to this graph, even in winter, the covid19-burst season, frequency for both hospital and icu for the people who are partially vaccinated are increased by the fewest among the three groups. Thus, I advise people, who are not sure if the third dose is suitable for them especially if they have some chronic health problem, to receive partial vaccination, which can effectively protect them.

From figure 6, we can clearly see that the frequency of cases for both normal hospitalization and icu-treatment for all groups of people have somewhat similar trends and show strong seasonality: with frequency burst in winter. Let's go back to figure 4, and we can clearly see that the proportion of cases due to covid19 also peak in winter, with the highest "dominance" shown in January and February in 2022. Thus, we can clearly see that covid-19 viruses are not only prevalent with the most dominance in winter, but also the most dangerous in winter. Although the proportion for covid at January this year is not that high, I do not think it disobey such law: the cold temperature arrived late this year and was still relatively warm at that time this year, which further supports the relationship between coldness (the main factor in the seasonality) and the frequency for hospitalization and icu-treatments.

The above analysis has shown the seasonality of danger of covid-19 according to the frequency for hospitalization and icu-treatments, and I want to further investigate on the extreme danger of the most severe cases with direct death data, with I merge with the hospitalization and icu data by date to give a fuller picture.

Here I define a new variable called *total_death_level* to denote the danger level by number of death cases on each day, with numbers 1 to 6, lightest to the most serious. More specifically, level 1 represents that the number of occurrences of death cases in Ontario does not surpass 10, level 2 represents that the number of occurrences of death cases in Ontario surpasses 10 and does not surpass 20, level 3 represents that the number of occurrences of death cases in Ontario surpasses 20 and does not surpass 30, level 4 represents that the number of occurrences of death cases in Ontario surpasses 30 and does not surpass 40, level 5 represents that the number of occurrences of death cases in Ontario surpasses 40 and does not surpass 50, level 6 represents that the number of occurrences of death cases in Ontario surpasses 50.

Here is a screenshot of the interactive visualization investigating the relationship between the dominance (proportion taken) of covid-19 in hospitalization and icu-treatments and the number of death cases in each day between 2022/1/10 to 2023/3/2:

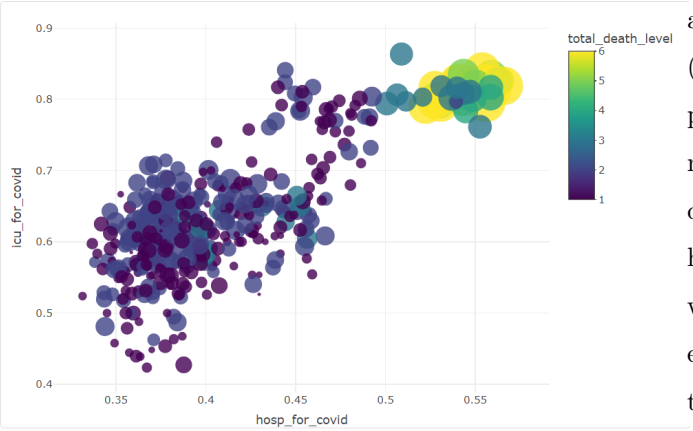

Figure 7

It can be viewed from Figure 7 that the proportion of icu cases that are due to covid19, the proportion of hospitalization cases that are due to covid19, and the number of death cases on that day are pairwisely positively correlated with each other. It can also be viewed that the highest death level happens when about 55% of the hospitalization cases are due to covid19 and about 80% of icu cases are due to covid19. One additional important thing that can be deducted from the graph I want to point out is the importance of effective self-treatment solution: it can be viewed that for those days whose covid19 dominance for icu cases are around 80%, there are still significant difference for the number of death cases between them: some cases even have *total_death_level* of 1, which means that it is possible for number of death cases to be no more than 10 in Ontario even if covid19 cases take up up to or even more than 80 of icu treatments! However, it is worth noticing that those cases happen at and only happen at when the proportion of covid19 taken cases are relatively low! Is it a coincidence? I do not think so: let's look at the horizontal line fixing *icu_for_covid* (proportion of icu cases that are due to covid19) around 0.8, the *total_death_level* (number of death cases on that day) increases as *hosp_for_covid* (proportion of hospitalization cases that are due to covid19) increases and ALL the days with the death cases more than 40 (*total_death_level* 5 and 6) have *hosp_for_covid* (proportion of hospitalization cases that are due to covid19) more than 0.5. Thus, the number of occurrences of death can be effectively controlled if the proportion of hospitalization for covid19 can be controlled! That's why I think it suggest the unignorable importance of effective self-treatment solution that people can make themselves feel better or even cure themselves so that they do not need to go to hospital to receive hospitalization treatments: this can add more hope or even much more hope for survival to those people who struggle with the most severe symptoms of covid19 and have to be treated to icu. One possible reason I can come up with is that more resources can be put into the most vulnerable icu patients when there are fewer covid19 patients waiting for hospitalization. Also, a smaller proportion of hospitalization due to covid19 reduces risk of cross infection: the various variants are so sneaky that even the most strict measures can hardly ensure 100% complete safety and they can be so dangerous for the most vulnerable patients. Thus, I strongly advise future study to put more focus on medical solutions of self-treatment with which people can cure themselves and do not need to go to hospital if they face noticeable but not that serious symptoms: it can contribute so much for the most dangerous cases though those medical solutions themselves may not be able to help with the most severe symptoms directly.

As I analyzed before, crowding in community activities is a great source for infection of covid19. However, personal healthcare workers are expensive to afford and thus many people who have difficulties living alone have to choose long-term care homes and live in communities. However, long-term care homes have the risk for covid19 outbreak considering many people here are vulnerable to the viruses. Probably because of the

potential rapid spread inside LTC homes, not only the residents there have to bear the risks, the staff there are exposed to the risks too. Specifically, I measure the covid19 outbreak frequency for long-term care for different cities of Ontario.

Sort by frequency, I get that the cities in Ontario with the most covid19 outbreaks of long term care homes are:

"Toronto"	"Scarborough"	"Ottawa"	"Mississauga"	"Etobicoke"	"London"
"Brampton"	"North York"	"Hamilton"	"Windsor"	"Kitchener"	"Burlington"
"Woodbridge"	"Barrie"	"Richmond Hill"	"Oshawa"	"Niagara Falls"	"Oakville"
"Cambridge"	"Newmarket"	"Thunder Bay"	"Sudbury"	"Waterloo"	"Sarnia"
"Markham"	"Brantford"	"Peterborough"	"St. Catharines"	"St Catherines"	"Cornwall"
"Kingston"	"Guelph"	"Welland"	"Sault Ste. Marie"	"Stouffville"	"Nepean"
"Maple"	"Dundas"	"Whitby"	"Leamington"	"Ajax"	"Aurora"
"Stoney Creek"	"Kanata"	"Orillia"	"Belleville"	"Simcoe"	"Vaughan"
"Fort Erie"	"Pickering"				

Figure 8

Then, I group the cities with the most frequent such outbreaks into 5 communities by locations: Capital, GTA (Greater Toronto Area), NO (Northern Ontario), SWO (Southwestern Ontario), GH (Golden Horseshoe), and I group all the other cities with fewer outbreaks into a group named "Covid19-proof cities" since they are relatively safe from such covid19 outbreaks and get communities as:

Capital	Covid19-proof cities	Golden Horseshoe	Greater Toronto Area	Northern Ontario	Southwestern Ontario
5101	31905	10120	31533	3934	7535

Figure 9

Visualize the distribution of the outbreak counts in figure 9 by barplot, I get:

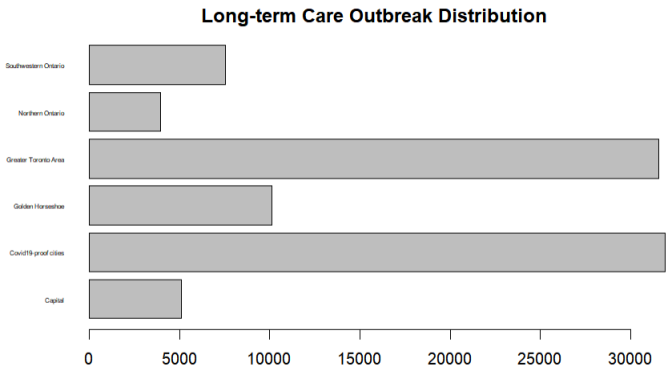

Figure 10

According to Figure 10, outbreaks are dense at pupulation-dense regions, especially GTA, whose frequency is almost the same as the total frequency of all the out breaks of all LTC homes not in the cities most dangerous for such outbreaks. Then, I make visualizations for each long term care outbreak. For each of the interactive visualizations (and screenshots of them in the report), I only contain the data for the 5 communities mentioned above containing the 50 Ontario cities with the highest 50 frequency of long term care outbreaks. For all the visualizations below, *hcw_level* is a new variable defined by me to evaluate the severity of each outbreak. *hcw_level* is denoted by 6 to 1 as the most serious to the lightest. The *hcw_level* is defined in the way that: if the number of active COVID-19 positive cases among staff is no fewer than 50, denote the outbreak as the most serious, by 6; if the number of active COVID-19 positive cases among staff is fewer than 50 and no fewer than 40, *hcw_level* is set as 5; if the number of active COVID-19 positive cases among staff is fewer than 40 and no fewer than 30, *hcw_level* is set as 4; if the number of active COVID-19 positive cases among staff is fewer than 30 and no fewer than 20, *hcw_level* is set as 3; if the number of active COVID-19 positive cases among staff is fewer than 20 and no fewer than 10, *hcw_level* is set as 2; otherwise, *hcw_level* is set as 1.

For all the visualizations below (Figure 11, Figure 12, Figure 13), ‘‘beds‘‘denote the total number of licensed beds in the Long-Term Care (LTC) home, $Total\_LTC\_HCW\_Cases$ denotes the number of active COVID-19 positive cases among staff associated with the Long-Term Care home, $Total\_LTC\_Resident\_Cases$ denotes the number of active Cumulative COVID-19 positive cases among residents of the Long-Term Care home, $Total\_LTC\_Resident\_Deaths$ denotes the cumulative number of COVID-19 related resident deaths in a Long- Term Care home, and and proportion denotes the proportion of deaths in the COVID-19 related resident cases in a Long- Term Care home.

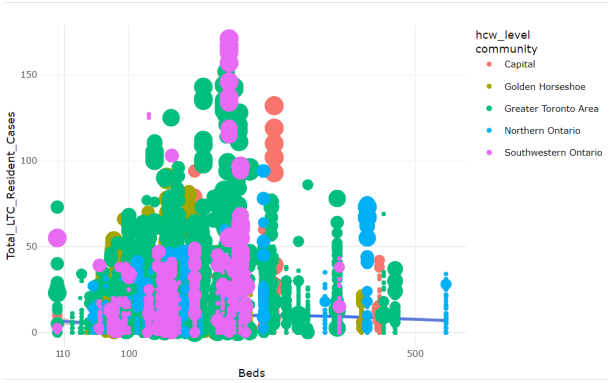

Figure 11

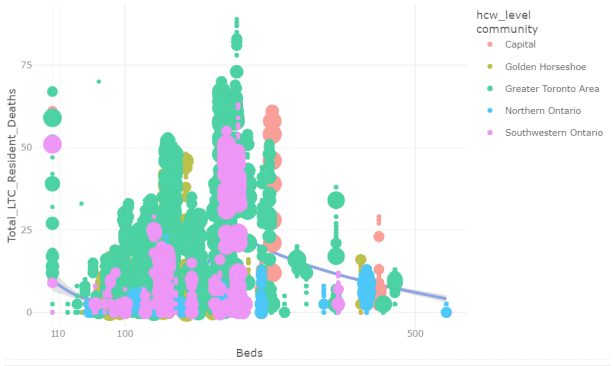

Figure 12

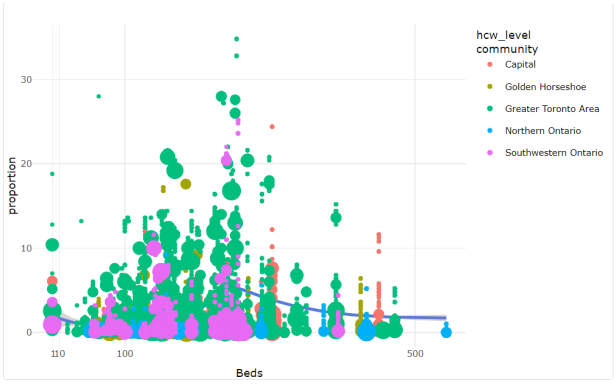

Figure 13

From Figure 11, Figure 12, Figure 13, we can view that the trend of change in the number of active Cumulative COVID-19 positive cases and deaths among residents of the Long-Term Care home and the proportion of deaths in the COVID-19 related resident cases in a Long- Term Care home are similar when the total number of licensed beds in the Long-Term Care (LTC) home changes in the same way: higher for medium size of number of licensed beds in the Long-Term Care (LTC) home and lower if number of licensed beds in the Long-Term Care (LTC) home is low or high enough. This means that small and large enough LTC homes are safer than medium size LTC home in term of such covid19 outbreaks. Also, from the figures it can be viewed that it is relatively safer to choose an LTC home in North-Western Ontario since the proportion for deaths of the covid19-related resident cases in such outbreaks are all low. However, I do not mean that LTC homes in North-Western Ontario are 100% safe, especially considering that the the number of active Cumulative COVID-19 positive cases among residents of the Long-Term Care home can also be high considering Figure 11, but it is worth-noticing that such infections here are not that dangerous as like places such as Greater Toronto Area. My explanation for the difference is due to that Northwestern Ontario is vast and sparsely populated while GTA is highly urbanized and extremely densely populated.

However, what if some people have no choice but to live in GTA? I subset the dataframe with GTA data and perform further analysis:

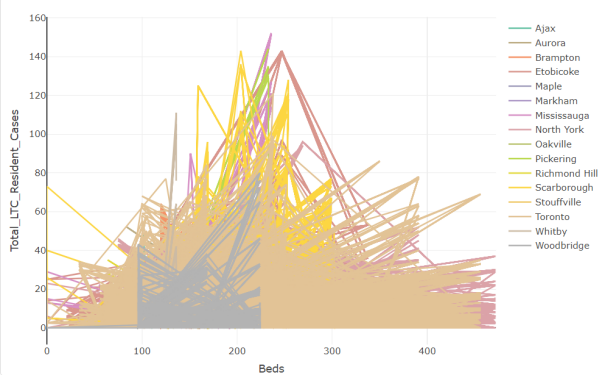

Figure 14

The triangle shape of the distribution further support my argument before that it is relatively safer to choose a long term care home that is small or large enough rather than a medium size one. According to Figure 14 (need to play with the interactive visu-

alization to more clearly see the results), if one requiring long term care has to live in GTA, at least avoid Toronto and Scarborough if possible. My city choice advice can also be justified with a direct barplot:

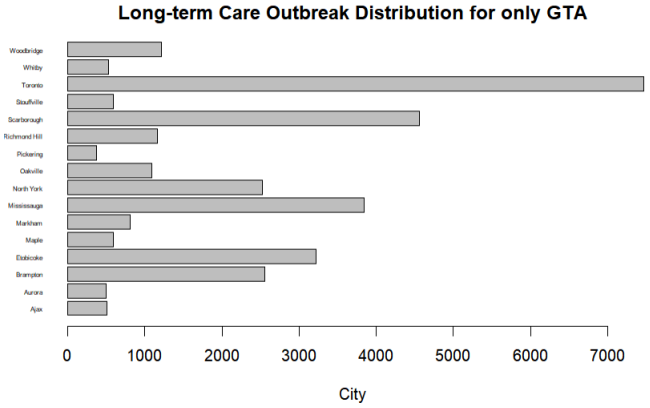

Figure 15

where Toronto and Scarborough also have more counts.

3 Conclusions and Summary

1. When and where were people more likely to get covid19?

People are more likely to get covid19 in winter, in community (crowded) activities and if they live in communities with more interpersonal interactions. For patients who have to receive long term care so that they have to live in community, I still advice to choose LTC homes that are located in sparsely populated areas if possible. For the vulnerable people, avoid our densely populated GTA if possible. Also, people who have to receive long-term care can have a smaller risk for covid19 if they choose to live in large enough or small enough long term care home: the medium size LTC homes are comparatively more dangerous for such risk.

2. What builds the ‘‘luck‘‘(not getting covid or mild severity if get covid19 or survive from extremely severe symptoms):

There is significant difference between danger level of infection (according to both the danger score, and each of the separate measuring standards including if ever hospitalized, if ever in icu and if ever intubated). Younger people (especially from the age group of between 20 and 29, like us) are more likely to stay away from covid, since all groups of non-zero danger level of infection of covid19 have the median, mean and mode of representative age to fall in range of age of senior citizens. However, every kind of age group has the possibility for both no risk and the highest danger risk, so I suggest protection measurements to maintain for all age groups but emphasize on the group of senior citizens. It looks like partial vaccination status is the safest status generally, according to frequency for hospital and icu, but I do not think it can be concluded that partial vaccination is necessarily better than full vaccination, but I think it is safer to crowd generally, especially to those with basic health conditions. People who live alone or do not participate in crowded community activities are safer too, which means ‘‘otaku ” may become an ‘‘advantage‘‘. Also, luck can come from everyone as a unit: developing the medical self-cure solutions for medium-level symptoms that otherwise need hospitalization can boost the hope for patients who struggle in ICUs to survive even if those medical solutions cannot cure them directly.

References

[1] https://open.toronto.ca/dataset/covid-19-cases-in-toronto/

[2] https://www.ontario.ca/page/covid-19-coronavirus

[3] https://data.ontario.ca/