Analyzing the Effects of Lockdowns in Toronto Areas Amid COVID-19

SHIDONG GUI

22/12/2020

Abstract

With the second wave of COVID-19 outbreak, Toronto areas have entered strict lockdown on November 23rd.(City of Toronto, 2020) In fact, as early as March 17th, Ontario premier Doug Ford announced state of emergency and warned citizens of Ontario to be aware of the newly discovered pandemic. However, are these lockdowns and procedure effective enough to reduce the growth curve of the COVID-19? The goal of this analysis is to discover whether the lockdowns has positively helped controlling the increase rate of COVID-19 and consequently reduce the fatal rate.

Keywords

COVID-19, *Pandemic*, *Observational Study*, *Logistic Regression*, *Toronto Public Health*

Introduction

The current pandemic of COVID-19 has created numerous disastrous incidents and global panic. Since the first presumptive case appeared on the 25th of January, premier Doug Ford had first announced state of emergency in March 17th and then initiated the lockdown in Toronto areas on the 20th of November. (Nielsen, 2020) Furthermore, the Ontario government had enforced the law that face masks or face covering are mandatory in enclosed public places in July 7th. (City of Toronto, 2020) Though the government continuously appeal the citizens of Ontario to practice social distancing, as of December 20th, the newly daily reported cases is at a scaring number of 646. (City of Toronto, 2020) Thus, it is worth the consideration whether the lockdowns or the restrictions have done their jobs. That is, the efficient lockdowns should be able to reduce the fatal rate and slow the growth rate of confirmed cases.

On way of to determine the fatal and growth rate of the patients who have confirmed with positive test result is to calculate the rates and numbers from the data set, which is published by the Toronto Public Health. Nevertheless, when requiring to discover the effects of the lockdowns, the dissection of the data is necessary. In other words, two or more groups need to be extracted from the whole data set in order to perform the compare and contrast of the periods before and after the lockdowns.

Three seperate data sets will be developed from the original data in order to investigate how efficient the first lockdown was and where the ongoing lockdown is leading all of the citizens. In the Methodology section, the detailed information of the dataset will be introduced and the models that are used to analyze the results of the lockdowns will be listed and explained. In the Results and Discussion sections, results of the analysis will be included and any weakness or further concerns, suggestions will be mentioned as well.

Methodology

Data

The data used to perform the analysis is obtained from the City of Toronto Open Data Portal. As described on the official website, Òthis data set contains demographic, geographic, and severity information for all confirmed and probable cases reported to and managed by Toronto Public Health since the first case was reported in January 2020.Ó (City of Toronto, 2020) In this data set, there are variables including the age group, gender, living neighbourhood, source of infection, episode date, reported date, outcomes and detailed hospitalized information(if applicable) of all the patients who have confirmed the infection of COVID-19 since the date of January. However, this analysis will only use the reported date as the partition of the lockdown dates. Mainly because the symptoms of COVID-19 can take up to 14 days to appear after the exposure to the COVID-19. (Government of Canada, 2020) Three important periods(cutting points) has been selected to perform

the analysis: the first lockdown (declare of state emergency) on the 17th in March, the period in between the first lockdown and the second lockdown, which in from the 17th of March to the 20th of November and finally the second lockdown that initiated on the 23th of November till the most recently recorded case, which is in Decemeber 14th. (Nielsen 2020 and City of Toronto 2020) Also, as mentioned in the introduction, the original data has been dissected into three groups: the group of patients that confirmed with COVID-19 during the first lockdown, the group of patients that caught COVID-19 in between the two lockdowns and the patients who tested positive after the second lockdown.

Table 1 | Since the full data that records the COVID-19 cases in Toronto areas contains almost 50,000 observations, it is more reasonable to create a cleaner and more manageable data set. This data set represents the individuals, who had confirmed positive results of COVID-19 after the first presumptive case in Toronto and before the first lockdown(March 17th). Also, this more focused data includes certain selected variables like age group, neighbourhood name, source of infection and etc. which are the critical conditions that will determine the likelihood of an individualÕs exposure to COVID. Furthermore, for the purpose of model building, this data set has modified both ÒClient GenderÓ and ÒOutcomeÓ columns in the original data set into ÒgenderÓ and ÒoutcomeÓ. In the ÒgenderÓ column, Ò1Ó indicates a male individual and Ò0Ó indicates a female individual. As for the ÒoutcomeÓ column, Ò1Ó indicates a resolved COVID-19 case and Ò0Ó indicates a fatal case.

```
formattable(head(in_between_lockdowns))
```

Age Group	Neighbourhood Name	Source of Infection	Reported Date	Client Gender	Outcome	Classification	gender	outcome
30 to 39 Years	Guildwood	Close contact	20201110	FEMALE	RESOLVED	CONFIRMED	0	1
50 to 59 Years	Bedford Park-Nortown	Close contact	20200318	MALE	RESOLVED	CONFIRMED	1	1
40 to 49 Years	Bedford Park-Nortown	Close contact	20200318	FEMALE	RESOLVED	CONFIRMED	0	1
60 to 69 Years	Niagara	Healthcare	20200318	MALE	RESOLVED	CONFIRMED	1	1
50 to 59 Years	Palmerston-Little Italy	Travel	20200318	MALE	RESOLVED	CONFIRMED	1	1
20 to 29 Years	Church-Yonge Corridor	Travel	20200318	FEMALE	RESOLVED	CONFIRMED	0	1

Table 2 | This is the data set the represents the time period in between two lockdowns. All variables and indicators remain the same to Table 1. However, as listed above, the reported date for the cases has narrowed down to March 17th to November 20th.

```
formattable(head(second_lockdown))
```

Age Group	Neighbourhood Name	Source of Infection	Reported Date	Client Gender	Outcome	Classification	gender	outcome
19 and younger	Leaside-Bennington	Close contact	20201130	MALE	RESOLVED	CONFIRMED	1	1
19 and younger	Thorncliffe Park	Community	20201206	FEMALE	ACTIVE	CONFIRMED	0	0
50 to 59 Years	Glenfield-Jane Heights	Community	20201129	FEMALE	RESOLVED	CONFIRMED	0	1
20 to 29 Years	Mount Pleasant East	Healthcare	20201126	FEMALE	RESOLVED	CONFIRMED	0	1
20 to 29 Years	Weston	Unknown/Missing	20201127	MALE	RESOLVED	CONFIRMED	1	1
20 to 29 Years	Trinity-Bellwoods	Community	20201130	MALE	RESOLVED	CONFIRMED	1	1

Table 3 | Similarly, this the data set represents the time period of second lockdown. Again, all the variables/categories remain the same. The main difference is the time frame, which this data records the confirmed cases from November 23rd (the actual effective date) to the most recent recorded date,December 14th.

Model

The model selected is the linear regression. Though the outcome need to serve the purpose of determining the efficiency of the lockdowns, it is to essential to discover whether any variables stand out that will effect the degrees of individualÕs exposure to COVID-19. Three different linear regressions were performed based on each of the data sets mentioned in the Data section. Variables such as age group, gender, neighbourhood name and source of infection were used for the regressions. Of course, the dependent variable (or the y variable) should be the outcome, that is, whether that specific case had been resolved or led to fatality.

The equations that were used for the construction of the models are as following:

$$1)\ \hat{outcome} = \beta_0 + \beta_1 \hat{20to29Years} + \ldots + \beta_8 \hat{90andolder} + \beta_9 \hat{gender} + \beta_{10} \hat{AgincourtSouth - MalvernWest} + \ldots + \beta_{148} \hat{Yorkdale - GlenPark} + \beta_{149} \hat{Community} + \ldots + \beta 154 \hat{Unknown/Missing} + \epsilon$$

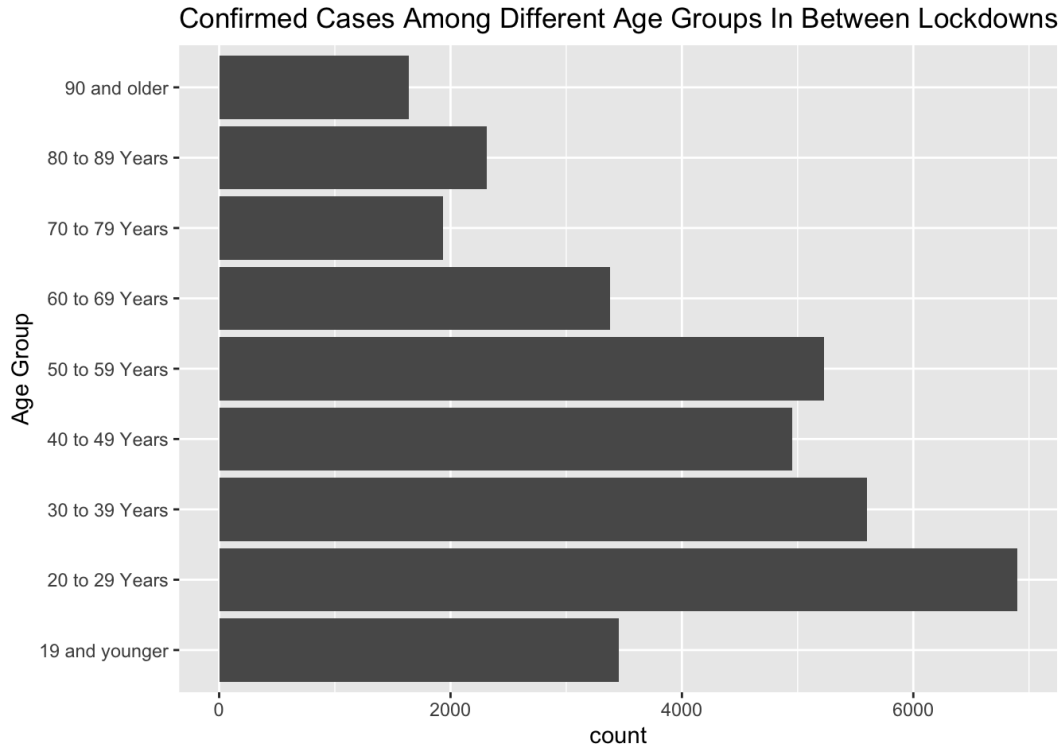
$\underbrace{\hat{\quad}}_{AgeGroup}$ $\underbrace{\hat{\quad}}_{Neighbourhood}$ $\underbrace{\hat{\quad}}_{SourceofInfection}$

Where β_0 represents the interception of the regressions and βs represent the slope of each variables. Noting that since there are many neighbourhoods, age groups and sources of infection, the equation is not fully extended. From the results of the regression, there is not a single variable that stands out (meaning having a conspicuous p-value). Thus, it is concluded that though individuals who were exposed to and confirmed with COVID_19 come from various age groups, living environments and have differenct sources of infection, the result indicate that they all have almost equal chances of catching the pandemic.

Results

This bar graph shows that during the first lockdown in Toronto areas, though citizens who are in the Ò60 to 60 YearsÓ appeared to have high positive rates among senior groups, most of the confirmed cases came from the youth groups.

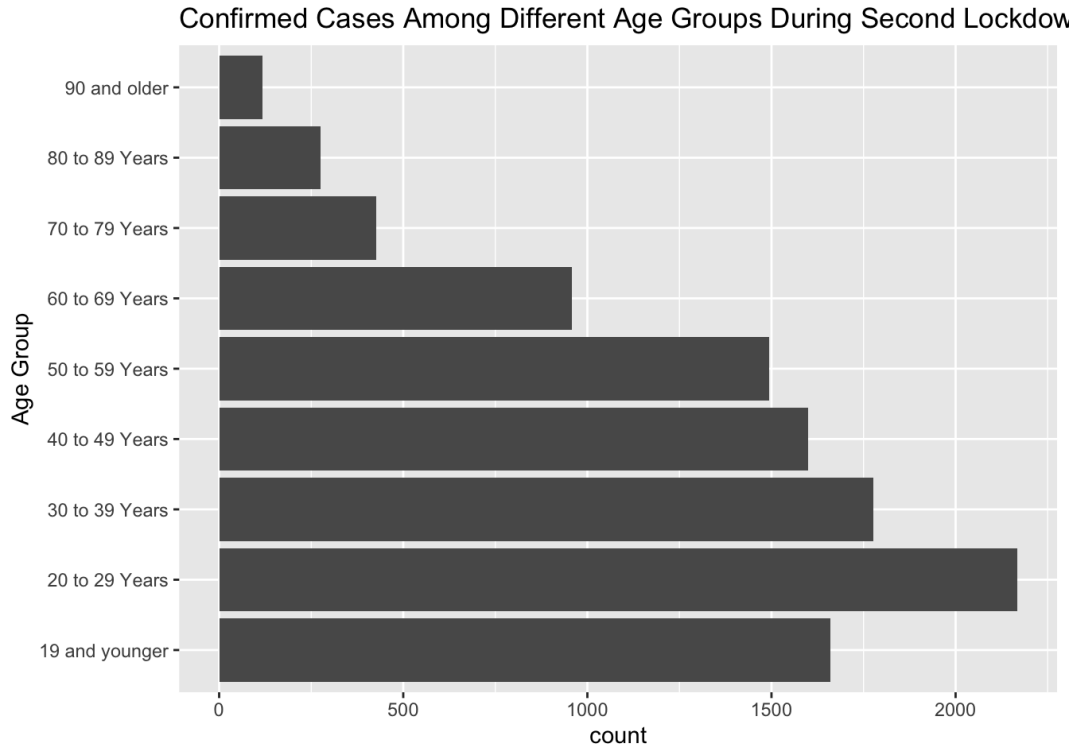
```
ggplot(data = in_between_lockdowns, aes(x=`Age Group`)) + geom_bar() + labs(title = "Confirmed Cases Among Differe
nt Age Groups In Between Lockdowns") + coord_flip()
```



| This bar graph reflects the fact that in

between the first and the second lockdowns, confirmed cases among senior groups had actually reduced. When comparing the Ò60 to 69 YearsÓ category with the previous graph, a decrease of proportion of confirmed cases among each age groups is discovered. The category of Ò20 to 29 YearsÓ in the age groups occupies the most proportion of confirmed cases and the category of Ò30 to 39 YearsÓ takes over the second largest proportion.

```
ggplot(data = second_lockdown, aes(x=`Age Group`)) + geom_bar() + labs(title = "Confirmed Cases Among Different Age Groups During Second Lockdown") + coord_flip()
```



| This bar graph shows the counts of confirmed cases after the second lockdown. Though very similar to the second bar graph, it is obvious that the proportion of positive COVID-19 among the age group of Ò19 and youngerÓ has increased tremendously.

From the first branched data set, there were total of 130 confirmed COVID-19 cases. Among these cases, 128 were resolved and only 2 of them led to fatality. Thus, using the formula $\frac{FATAL\ CASES}{TOTAL\ CONFIRMED\ CASES}$, the fatal rate during the first lockdown was roughly 1.538%.

From the second branched data set, the total number of confirmed cases had increased to a total of 35,391 in between the lockdowns. 33748 of these cases were resolved and 1,601 of the confirmed cases resulted in fatality. However, there remains 42 active cases. By adopting the same formula, a 4.523% of fatality rate was calculated.

Lastly, from the third data set, which recorded the counting, second lockdown, there are total of 10,475 COVID-19 cases and 5,077 are still active. The fatality number so far stays at 47. Using the formula mentioned above, the fatality rate is at 0.449%. However, this is not the final fatality rate since the second lockdown has not ended yet.

Discussion

Summary

In this analysis, the original data obtained from the Toronto Public Health was rearranged into three different branches. Each of the branches represents a

certain time period (first lockdown, in between two lockdowns and the second lockdown) and the analyses are performed based on these separated data sets accordingly. The models built based on these data sets adopted the linear regression method. Though the regressions were not mainly used to determine whether the lockdowns are effective, they contributed to the conclusion that the difference between individualsÕ backgrounds do not significantly affect their chances of exposure to the virus. After calculating the fatality rate of three data sets, it is clear that the fatality rate reached its peak in between the first and the second lockdown. Though the fatality rate of the second lockdown remains fluid, it has drastically decreased comparing to the time period in between the lockdowns.

Conclusions

The fatality rate from the first lockdown was around 1.538%, 4.523% for the time period in between the first and the second lockdown and 0.449% for the second lockdown. Although the second lockdown is still ongoing, the decreasing percentage from 4.523 to 0.449 indicates that the second lockdown is much more efficient and effective than the first one. One of the reasons might be the Stage 2 reopening in Toronto areas initiated on 24th of June. The reopening might be a method to relief the high pressure on many of the local, small business, but it could be problematic and difficult for social distancing. Furthermore, the enforced law of mandatory face masks/covering wearing did not execute till the July 7th. There were a gap of 14 days in between, but the symptoms of COVID-19 take up to 14 days to reveal. Many patients who were exposed to COVID without knowing it might have already spreaded the virus in public areas in these two weeks.

The results showed that lockdowns can be effective but adopting lockdown as the one and only method for preventing the spread of COVID-19 is not enough. The second lockdown enforces more restrictions and results in a much lower fatality rate. People who refused to wear facial covers or self-isolating after the contraction of the virus should take in consideration that lockdowns and procedures that are established by the government take some time to reflect their functions. Also, from the bar graphs we can tell that more and more youth have tested positive for COVID-19. Ideally, schools or other teenager-related asscociation or social groups should raise the awareness among young citizens. Young people may have a higher possibility of self-healing after contracted with the virus, but they can carry the virus back home and put the seniors at home in higher risks.

Weakness & Next Steps

One obvious weakness is the data set that obtained from the Toronto Public Health. Since the Toronto Public Health department only records those who reported to hospitals or other health institutions, the validity of the data is highly debatable. Some of the patients may ignore their symptoms and continue their daily lives without reporting to health departments voluntarily. As for some others, they may have exposed themselves to the virus without knowing the fact they have contracted COVID since many of the infected patients showed no symptoms at the beginning.

One of core limitations of this is analysis is the lack of variety in the data. Currently, the whole analysis is based on one data set that provided by the Toronto Public Health. When obtaining more data sets that are related to the recording of COVID-19 cases in Toronto areas, more models can be constructed and analyses will be more accurate. Another limitation is that this data set has some missing values. For example, some resources of infection remains unknown and some values in the dataset.

The next step is to continue practicing social distancing as holiday season comes near. More people will attempt to go out or reunite with families. However, as mentioned in the conclusion section, the second lockdown acts more effectively than the first one. Also, if self-reportings and self-isolations are more encouraged by the authorities, the lockdowns may be able to prevent the spread and the growth of COVID-19 to the largest extents.

References

Canada, Public Health Agency of. ÒGovernment of Canada.Ó Canada.ca, / Gouvernement Du Canada, 20 Nov. 2020, www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/symptoms.html.

City of Toronto, City of Toronto. ÒCOVID-19: Lockdown Guide for Toronto Residents.Ó City of Toronto, 18 Dec. 2020, www.toronto.ca/home/covid-19/covid-19-reopening-recovery-rebuild/covid-19-guide-for-toronto-residents/.

City of Toronto. ÒCOVID-19: Status of Cases in Toronto.Ó City of Toronto, 16 Dec. 2020, www.toronto.ca/home/covid-19/covid-19-latest-city-of-toronto-news/covid-19-status-of-cases-in-toronto/.

Ellis, Laura. ÒCreate Stylish Tables in R Using Formattable.Ó Little Miss Data, Little Miss Data, 25 Sept. 2018, www.littlemissdata.com/blog/prettytables.

Kabacoff, Robert I. ÒSubsetting Data.Ó Quick-R: Subsetting Data, 2017, www.statmethods.net/management/subset.html. Nielsen, Kevin. ÒA Timeline of the Novel Coronavirus in Ontario.Ó Global News, Global News, 17 Dec. 2020, globalnews.ca/news/6859636/ontario-coronavirus-timeline/.

Prison RodeoPrison Rodeo 10111 gold badge11 silver badge33 bronze badges, et al. ÒFirst-Line Paragraph Indenting in PDFs Using R Markdown.Ó Stack Overflow, 19 Jan. 2018, stackoverflow.com/questions/29460112/first-line-paragraph-indenting-in-pdfs-using-r-markdown. RStudio. Code Chunks, rmarkdown.rstudio.com/lesson-3.html.

Toronto Public Health. ÒAbout COVID-19 Cases in Toronto.Ó Opendatatoronto, 16 Dec. 2020, open.toronto.ca/dataset/covid-19-cases-in-toronto/. user3922483user392248312322, gold badges22, silver badges77 bronze badges, et al. ÒR - Remove Dashes from a Column with Phone Numbers.Ó Stack Overflow, 1 Sept. 1963, stackoverflow.com/questions/25294493/r-remove-dashes-from-a-column-with-phone-numbers.