Statistical Analysis of COVID-19 Cases In Toronto (Rough Draft)

Anonymous

09/12/2020

Abstract

COVID-19 is a big issue we are pacing in 2020, so the report is about analyzing the COVID-19 cases happening in Toronto.The data is collected by 'Toronto Open Data', and from this data we created graphs and visualizations to show COVID trends and how people are exposed to the diseases in our daily life situations. The report focuses on different age group, how people mostly get infected, see the differences between male and female and distributions of viruses in the neighbourhoods.

Keywords

1. Ratio
2. COVID-19
3. data
4. bar graph
5. Trend

Introduction

People say we are living in a 'new normal', where everyone's lifestyle has changed after the big pandemic called COVID-19 viruses.COVID-19 is hurting economies, regardless of income level impacting whole global perspective (2020). People have lost jobs, and markets have closed and it has changed most of people's life style. So we are going to look at how people are exposed to this viruses, and view it in statistical methods. Using 'About COVID-19 Cases in Toronto' data from the 'Toronto Open Data' website. The data contains gender, source of infection, classification, Reported date, etc.The goal of this project is to create an analysis of this data to see how people mostly got infected, and see the different graphs and data to analyze the coronavirus in toronto right now.

COVID-19 is a global pandemic but in this report we are going to look at local data, Toronto.Toronto had one lockdown in early 2020, and after the second wave Ontario government has decided to process the lockdown again. Coronavirus is spreading in very large number and it has been spreading constantly. So we are going to look at how people are getting infected and what is the ratio of male and female on getting infected.

First we are going to clean the raw data, then select variables that are going to be used in the analysis.

Data

'About COVID-19' data is collected from 'Toronto Open Data' website. The raw data includes number of people that has been infected or still in infection, with many different categories, such as 'Assigned ID', 'Age

Group', 'Client Gender', 'Source of Infection', etc. Since there are too many variables, and it makes it harder to look into the data, we cleaned some variables that were not as important as other variables.

The cleaned data look like as the following:

```
## # A tibble: 6 x 10
##   `Age Group` `Neighbourhood ~ `Source of Infe~ Classification `Episode Date`
##   <chr>       <chr>            <chr>            <chr>          <date>
## 1 50 to 59 Y~ Willowdale East  Travel           CONFIRMED      2020-01-22
## 2 50 to 59 Y~ Willowdale East  Travel           CONFIRMED      2020-01-21
## 3 20 to 29 Y~ Parkwoods-Donal~ Travel           CONFIRMED      2020-02-05
## 4 60 to 69 Y~ Church-Yonge Co~ Travel           CONFIRMED      2020-02-16
## 5 60 to 69 Y~ Church-Yonge Co~ Travel           CONFIRMED      2020-02-20
## 6 50 to 59 Y~ Newtonbrook West Travel           CONFIRMED      2020-02-24
## # ... with 5 more variables: `Reported Date` <date>, `Client Gender` <chr>,
## #   Outcome <chr>, `Currently Hospitalized` <chr>, source_of_infection <chr>
```

* Variable includes 'Age Group', 'Neighbourhood Name', 'Source of Infection', 'Classification', 'Episode Date', 'Reported Date', 'Client Gender', 'Outcome', and 'Currently Hospitalized.'

Model

First we are going to create the ratio of different gender that are exposed to the virus. Since the variables in the data is categorical variables, and the status is already fixed as people who already reported as infected we are going to create bar graphs see the frequencies.

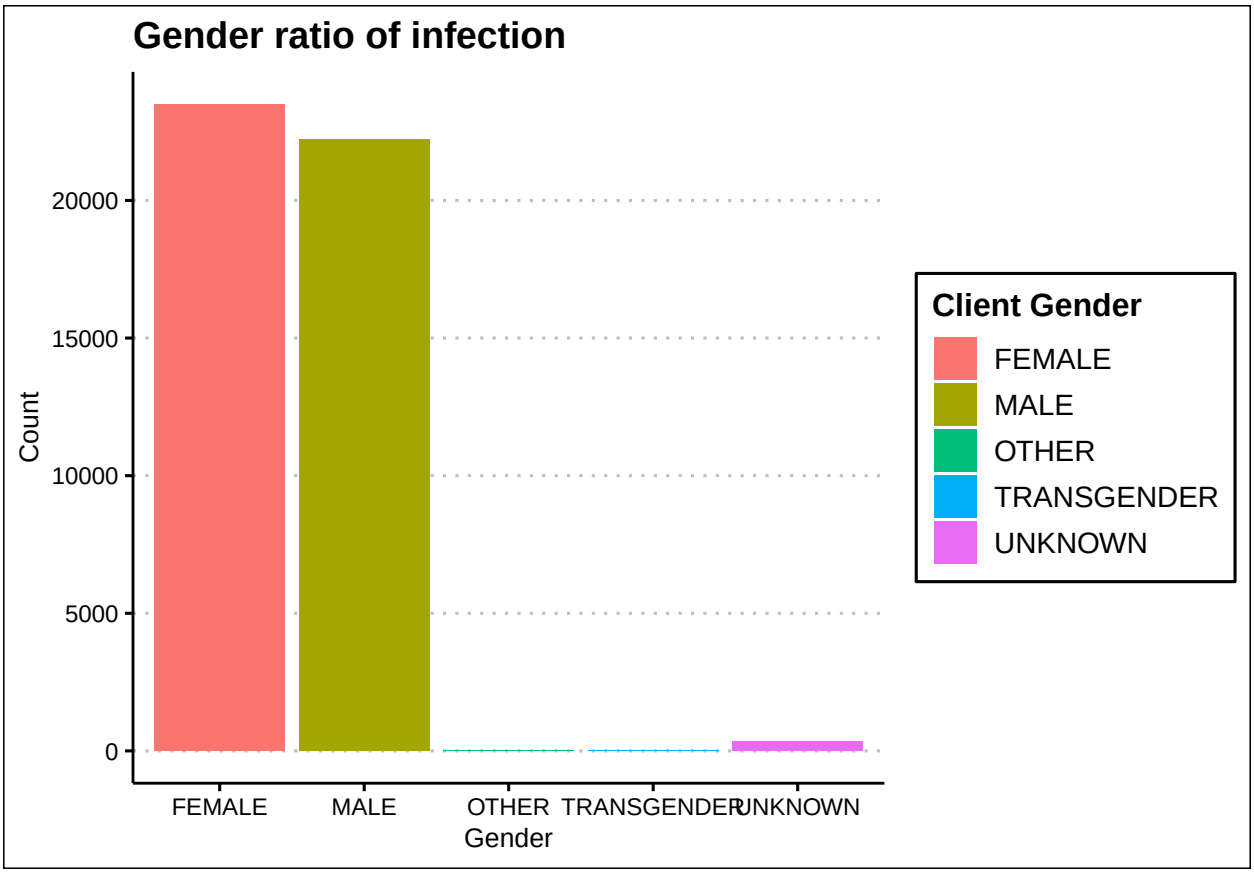
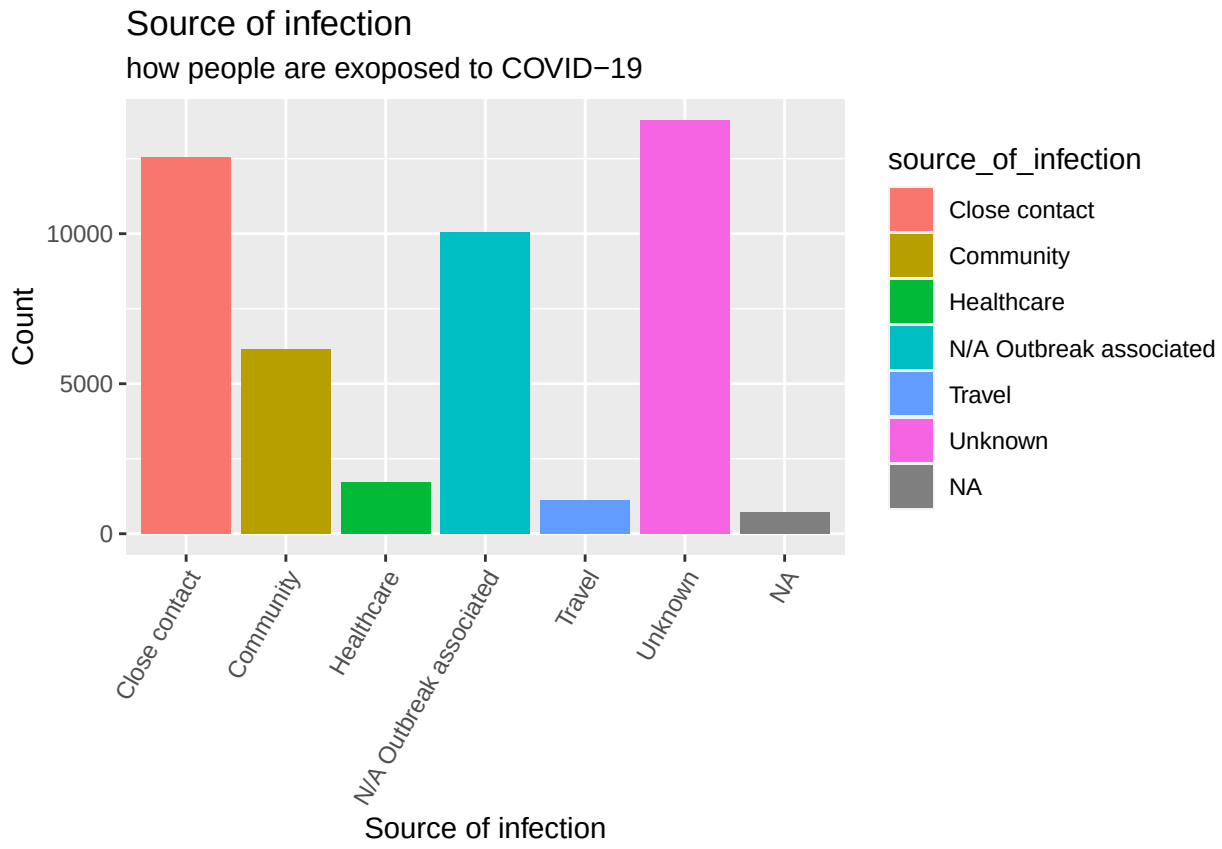


Table 1: Client Gender ratio

Gender	percentage
FEMALE	50.9663146
MALE	48.2007678
OTHER	0.0281977
TRANSGENDER	0.0173524
UNKNOWN	0.7873674

The graph shows the total count of each gender, and the Table 1 shows the ratio(percentage) of the gender who have been infected.It shows that 51 percent of the infected gender are female and 48 percent is male, this shows that female are likely more exposed to the infection in Toronto.

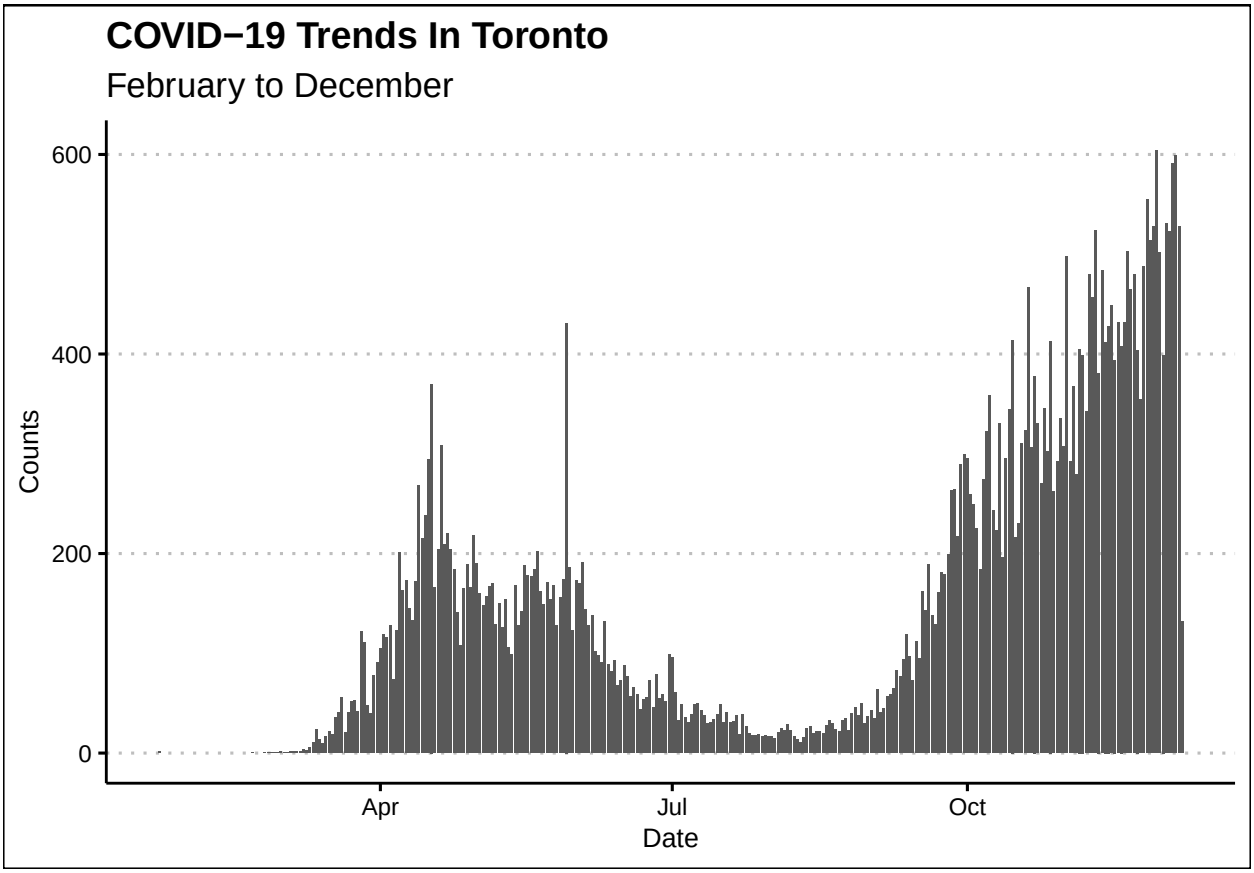
Then we will look at how people get infected mostly, and how easily people are exposed to the COVID-19 virus. Like the first method, we will create a bar graph of different source of infections.



The source of infection includes close contact, community, health care, travel, N/A Outbreak Associated, and unknown. Then we have created the COVID-19 graphs in Toronto to see the trend of the spread of the viruses.

Results

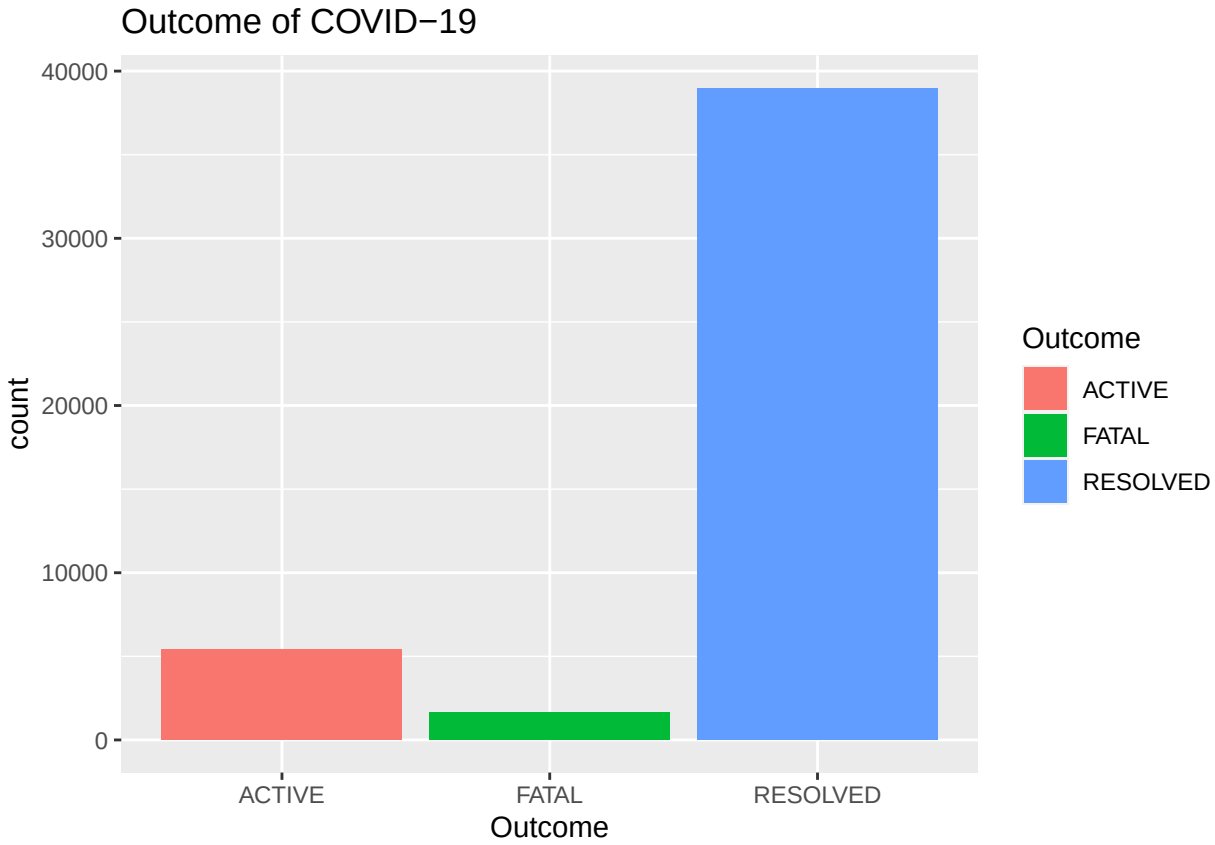
This is the graph for COVID-19 trend in early 2020 to late 2020:



The COVID-19 trend graphs shows that during July to late September it had less than 50 counts, but starting October the COVID-19 has increased extremely, about 10 times more higher than July. If the graphs continues the virus will spread more, or this can be the peak of the spread and it can decrease if people be aware of it, and keep the social distancing.

Table 2: Outcome ratio

Outcome	Ratio
Active	0.1180184
Fatal	0.0362666
Resolved	0.8457150



The Outcome of the Virus shows that about 85% of people who were infected are resolved, 12% are still in active and 4% of people are recorded fatal.This tells us that most of the people get healed from the COVID-19 in Toronto, and hospitality is doing a great job for people that are sick.

Summary

The report explains different analysis of COVID-19. For gender, it is almost 5:5 ratio of male and female on infection. People get mostly infected by close contact and community meeting. Also people in health care services are most exposed to the virus so even if they prevent the chance, they can still get infected from the COVID-19. COVID-19 is currently spreading in a very fast speed. The lock down and vaccine can reduce the spread of the virus there are more possibility of going down. In appendix there are table for COVID-19 in each neighbourhoods for people that are interested for their neighbourhood.

Conclusion

It is important for people to keep the social distancing and follow the rules government has made. There are vaccine available in many country now, and by keeping the social distancing and preventing infections this will bring back our old normal. The pandemic has taken up whole 2020, and this will be the time for people to change and develop our weaknesses and needs.

Weakness & Next Steps:

The data only contains categorical variables and it only includes people that are already exposed to the COVID-19 virus, so finding relationship between the variables were not possible. For next step, we can collect data for people not exposed to the COVID-19 and find relationship between people's contact tracings.

References

Impact of Coronavirus: 1. Coronavirus: The economic impact – 10 July 2020. (n.d.). Retrieved December 17, 2020, from https://www.unido.org/stories/coronavirus-economic-impact-10-july-2020

CSV Data: 2. Open Data Dataset. (n.d.). Retrieved December 10, 2020, from https://open.toronto.ca/dataset/covid-19-cases-in-toronto/

Appendix

Table 3: Covid-19 cases in different neighbourhoods

Var1	Freq
Agincourt North	355
Agincourt South-Malvern West	291
Alderwood	145
Annex	290
Banbury-Don Mills	193
Bathurst Manor	398
Bay Street Corridor	186
Bayview Village	136
Bayview Woods-Steeles	195
Bedford Park-Nortown	272
Beechborough-Greenbrook	202
Bendale	722
Birchcliffe-Cliffside	437
Black Creek	882
Blake-Jones	60
Briar Hill - Belgravia	256
Bridle Path-Sunnybrook-York Mills	53
Broadview North	90
Brookhaven-Amesbury	454
Cabbagetown-South St. James Town	228
Caledonia-Fairbank	148
Casa Loma	78
Centennial Scarborough	126
Church-Yonge Corridor	339
Clairlea-Birchmount	545
Clanton Park	181
Cliffcrest	250
Corso Italia-Davenport	174
Danforth	96
Danforth-East York	117

Table 4: Covid-19 cases in different neighbourhoods

Var1	Freq
Don Valley Village	242
Dorset Park	703
Dovercourt-Wallace Emerson-Junction	346
Downsview-Roding-CFB	1146
Dufferin Grove	143
East End-Danforth	200
Edenbridge-Humber Valley	237
Eglinton East	474
Elms-Old Rexdale	243
Englemount-Lawrence	472
Eringate-Centennial-West Deane	198
Etobicoke West Mall	241
Flemingdon Park	526
Forest Hill North	153
Forest Hill South	62
Glenfield-Jane Heights	1147
Greenwood-Coxwell	120
Guildwood	232
Henry Farm	247
High Park-Swansea	181
High Park North	135
Highland Creek	283
Hillcrest Village	135
Humber Heights-Westmount	290
Humber Summit	381
Humbermede	539
Humewood-Cedarvale	93
Ionview	152
Islington-City Centre West	717
Junction Area	166
Keelesdale-Eglinton West	181

Table 5: Covid-19 cases in different neighbourhoods

Var1	Freq
Kennedy Park	366
Kensington-Chinatown	234
Kingsview Village-The Westway	671
Kingsway South	86
L'Amoreaux	584
Lambton Baby Point	56
Lansing-Westgate	138
Lawrence Park North	87
Lawrence Park South	90
Leaside-Bennington	123
Little Portugal	222
Long Branch	77
Malvern	1130
Maple Leaf	446
Markland Wood	87
Milliken	261
Mimico (includes Humber Bay Shores)	422
Morningside	464
Moss Park	367
Mount Dennis	405
Mount Olive-Silverstone-Jamestown	1449
Mount Pleasant East	83
Mount Pleasant West	359
New Toronto	146
Newtonbrook East	90
Newtonbrook West	533
Niagara	321
North Riverdale	75
North St. James Town	379
O'Connor-Parkview	189
Oakridge	295

Table 6: Covid-19 cases in different neighbourhoods

Var1	Freq
Oakwood Village	366
Old East York	87
Palmerston-Little Italy	94
Parkwoods-Donalda	365
Pelmo Park-Humberlea	303
Playter Estates-Danforth	92
Pleasant View	163
Princess-Rosethorn	102
Regent Park	172
Rexdale-Kipling	293
Rockcliffe-Smythe	411
Roncesvalles	113
Rosedale-Moore Park	129
Rouge	1271
Runnymede-Bloor West Village	47
Rustic	232
Scarborough Village	518
South Parkdale	468
South Riverdale	146
St.Andrew-Windfields	98
Steeles	207
Stonegate-Queensway	160
Tam O'Shanter-Sullivan	395
Taylor-Massey	251
The Beaches	81
Thistletown-Beaumond Heights	449
Thorncliffe Park	673
Trinity-Bellwoods	128
University	95
Victoria Village	283
Waterfront Communities-The Island	688
West Hill	442

Table 7: Covid-19 cases in different neighbourhoods

Var1	Freq
West Hill	442
West Humber-Clairville	1260
Westminster-Branson	618
Weston	745
Weston-Pellam Park	163
Wexford/Maryvale	376
Willowdale East	264
Willowdale West	149
Willowridge-Martingrove-Richview	368
Woburn	1356
Woodbine-Lumsden	55
Woodbine Corridor	74
Wychwood	150
Yonge-Eglinton	72
Yonge-St.Clair	67
York University Heights	989
Yorkdale-Glen Park	469