Covid-19 Cases in City of Toronto Neighborhoods

Masoud Bozorg Bigdeli

July 2020

Introduction

- Over the past few months of the covid-19 outbreak, it is well established that an effective way to control the spread of novel coronavirus is by testing the symptomatic and potential cases, followed quarantining the positive cases.

- Although testing is critical, healthcare sector has a defined capacity and constrained resources, rendering the idea of testing an entire nation over and over for a period of time impractical.

- The need for a more focused testing strategy was highlighted by the researchers recently, which would lead to efficient use of the resources available at hand.

- Analyzing covid-19 cases at neighborhood level could present a more focused point of view for defining an effective testing strategy making best use of available resources.

Data

- Covid-19 neighborhood data made available via City of Toronto municipality is used in this study. To view the data source click here.
- The coordinates of each neighborhood are obtained from Nominatim geolocation service through GeoPy library in Python. To view the GeoPy documentation click here.
- Toronto neighborhoods demographic data is extracted from Wellbeing Toronto application. To view the data source click here.
- Foursquare venues data are used to see what type of venues are most popular in the neighborhood with highest numbers of covid-19 cases
- The data extracted from the above data sources were cleaned and merged to study the potential variables having effect on number of cases in each neighbourhood.

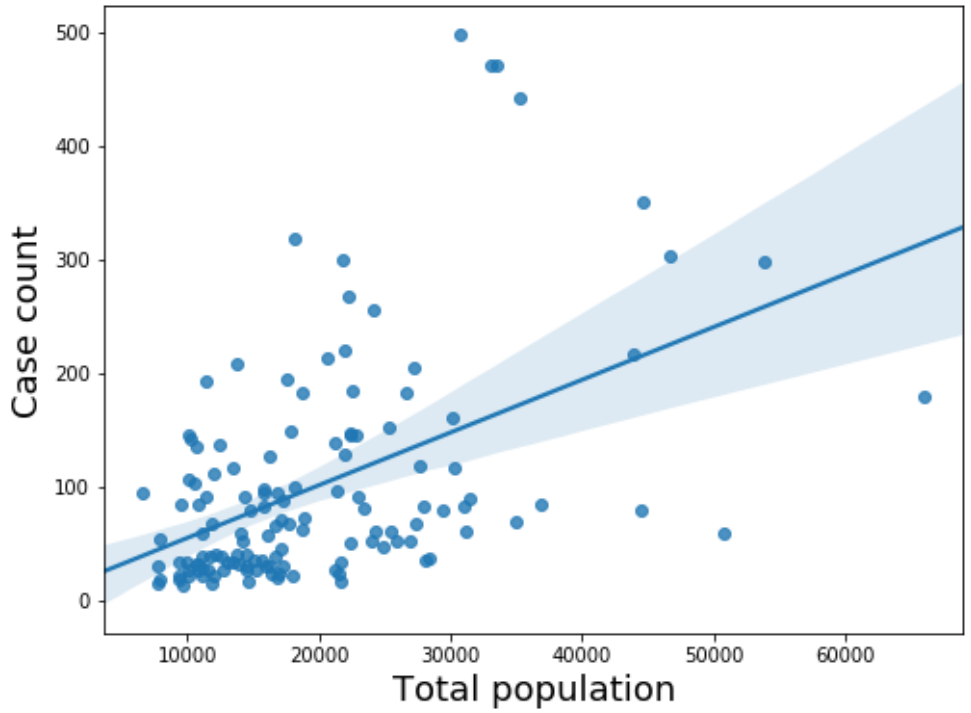
Data

- In the following you can see an example of dataframes made in this study by joining resulting data collected from different sources.

Neighbourhood ID	Neighbourhood Name	Case Count	Rate per 100,000 people	Latitude	Longitude	Low Income Families	Healthy Food Index	Total Area	Total Population	Pop 0 - 19	Pop 20 - 29	Pop 30 - 39	Pop 40 - 49	Pop 50 - 59	Pop 60 - 69	Pop 70 - 79	Pop 80 - 89	Pop 90+	Density	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	Number_of_Venues
113.0	Weston	318	1767.452201	43.7002	-79.5162	2650	41.45	2.60	18140	4095	2570	2580	2430	2755	1940	910	690	170	6976.923077	Train Station	Coffee Shop	Soccer Field	Sandwich Place	Laundromat	19
8.0	Humber Heights-Westmount	193	1762.879065	43.6861	-79.5288	820	32.82	2.80	11385	1970	1135	1155	1295	1610	1300	980	1480	460	4066.071429	Bank	Pizza Place	Café	Restaurant	Plaza	9
25.0	Glenfield-Jane Heights	499	1636.548490	43.6535	-79.3839	4420	41.05	5.20	30715	8015	4150	3630	3875	4270	2930	2095	1550	200	5906.730769	Clothing Store	Coffee Shop	Hotel	Diner	Restaurant	78
115.0	Mount Dennis	209	1537.556095	43.687	-79.4896	1930	29.91	2.10	13705	3265	1960	1965	1835	2105	1300	705	460	110	6526.190476	Coffee Shop	Furniture / Home Store	Pizza Place	Grocery Store	Bus Line	9
140.0	Guildwood	145	1462.135727	43.7552	-79.1982	440	15.57	3.80	10120	1830	940	890	1295	1570	1345	1020	1030	200	2663.157895	Train Station	Storage Facility	Baseball Field	Yoga Studio	Dumpling Restaurant	4
2.0	Mount Olive-Silverstone-Jamestown	471	1429.265036	43.6535	-79.3839	5010	37.57	4.60	33080	9690	5055	4435	4590	4145	2805	1515	765	80	7191.304348	Clothing Store	Coffee Shop	Hotel	Diner	Restaurant	78
112.0	Beechborough-Greenbrook	94	1429.223050	43.695	-79.4717	950	25.57	1.80	6665	1545	905	900	880	970	695	395	300	75	3702.777778	Furniture / Home Store	Italian Restaurant	Video Store	Bar	Auto Garage	10
1.0	West Humber-Clairville	471	1413.904899	43.6824	-79.4808	3790	23.82	30.09	33525	7375	5905	4380	4205	4670	3390	2070	1315	215	1114.157527	Gas Station	Home Service	Coffee Shop	Park	Wine Shop	6
29.0	Maple Leaf	142	1404.411037	43.7123	-79.4902	880	14.42	2.50	10215	2095	1290	1265	1410	1515	915	825	775	125	4086.000000	Bakery	Basketball Court	Yoga Studio	Filipino Restaurant	Doctor's Office	2
24.0	Black Creek	299	1375.534802	43.6954	-79.4855	3680	41.84	3.40	21875	6190	3280	2725	2750	2835	1925	1230	785	155	6433.823529	Fast Food Restaurant	Construction & Landscaping	Coffee Shop	Playground	Yoga Studio	4

Results

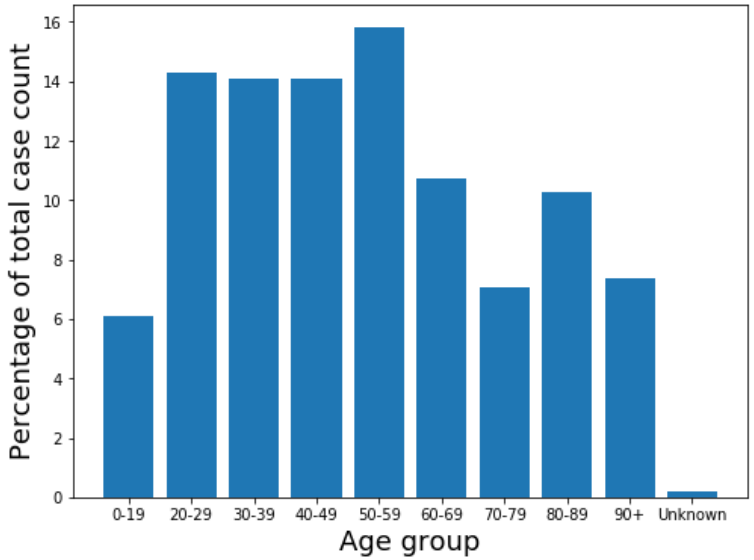
- It is observed that number of covid-19 cases have a direct relationship with the population of neighborhoods. A general increasing trend can be evidenced from figure below.

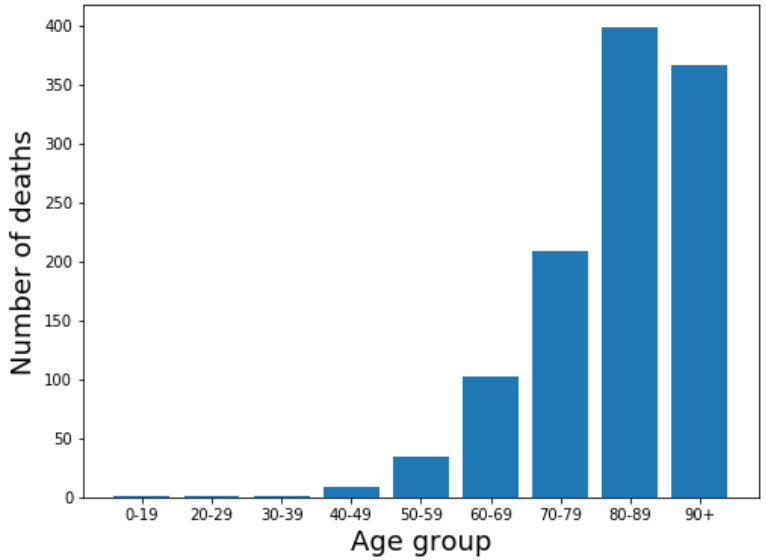


Results

- Distribution of covid-19 cases counts is higher among age groups between 20 to 59 years. We'll refer to this group as middle aged or working population in this study.

- Moreover, distribution of number of deaths among different age groups shows that number of covid-19 related deaths increases exponentially with age. High number of deaths occurred for age groups between 60 to 90+ years. We'll refer to this group as seniors in this study.

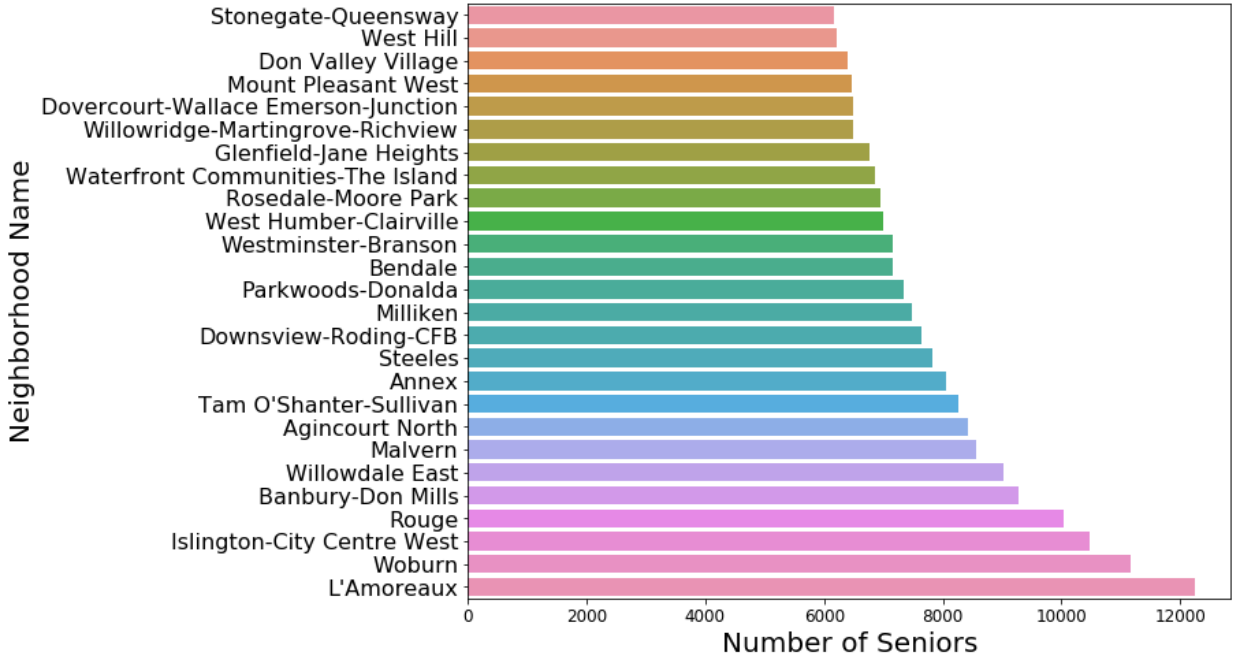




Results

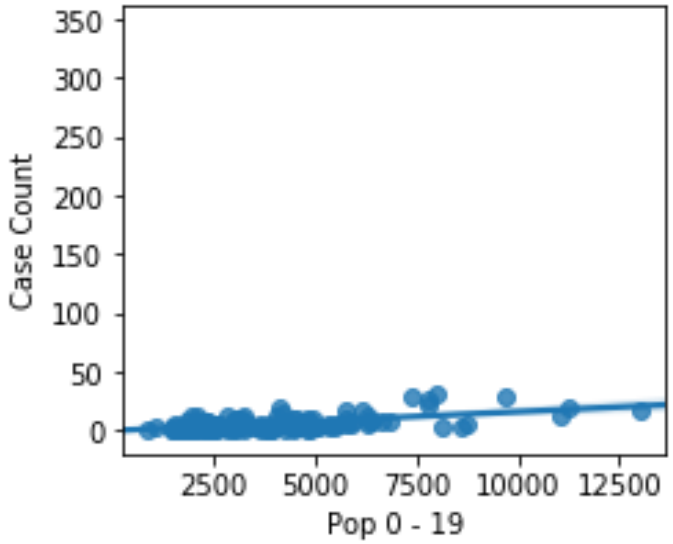
- Number of seniors are higher specifically in 3 neighborhoods, namely L'Amoreaux, Woburn and Islington, where more than 30,000 seniors live in total.

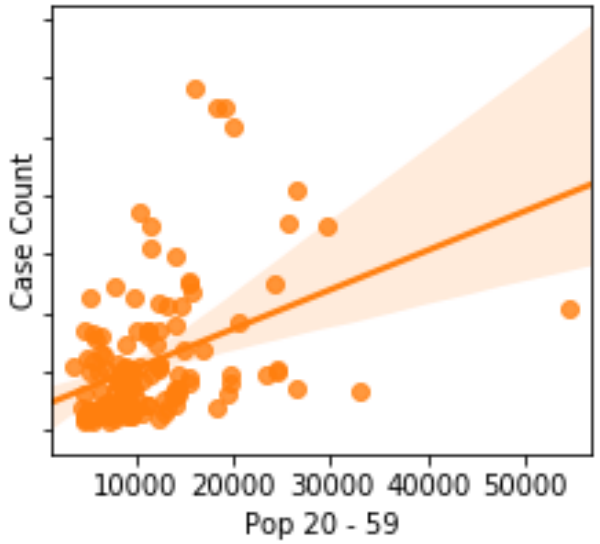
- Strict measures should be considered in such neighborhoods to protect seniors from covid-19.

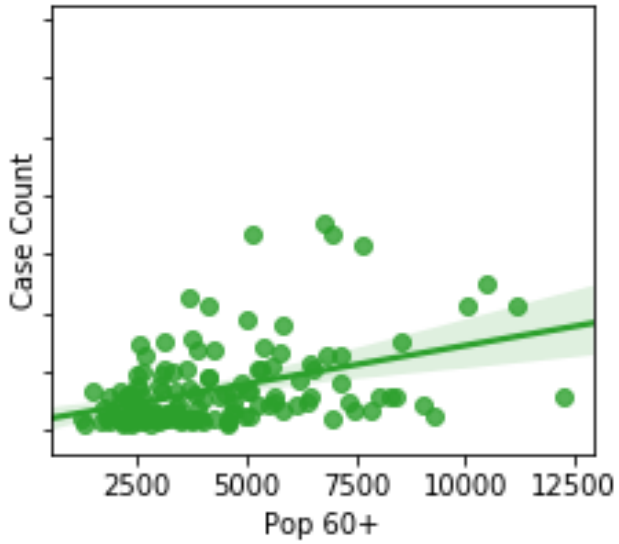


Results

- Covid-19 case counts against population age shows that the number of cases are significantly higher among ages 20 to 59 years old, which we can consider as working age population.

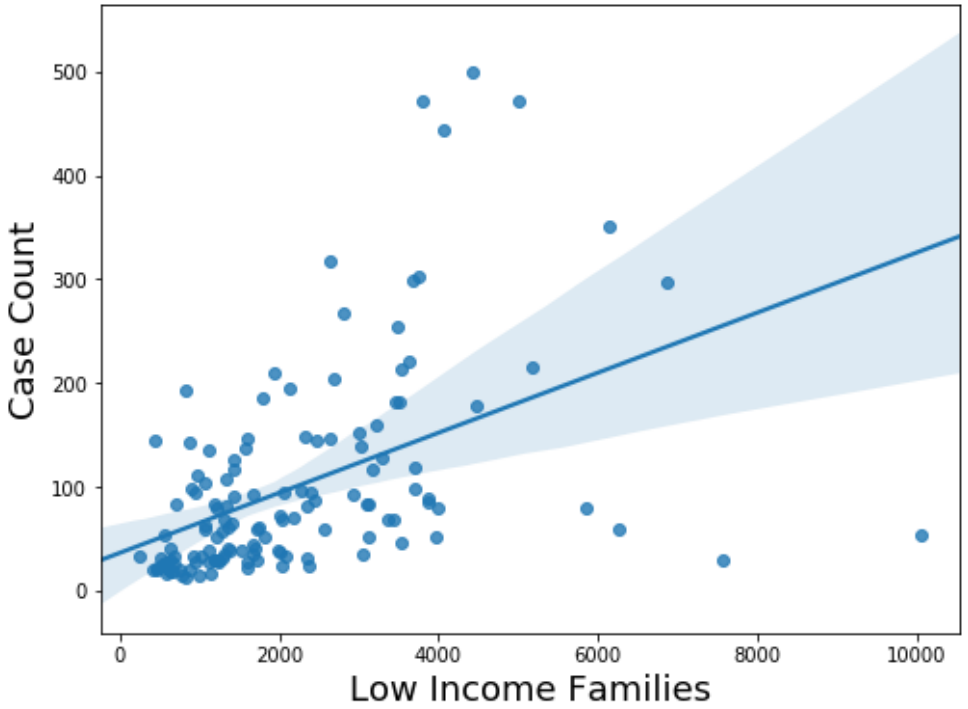






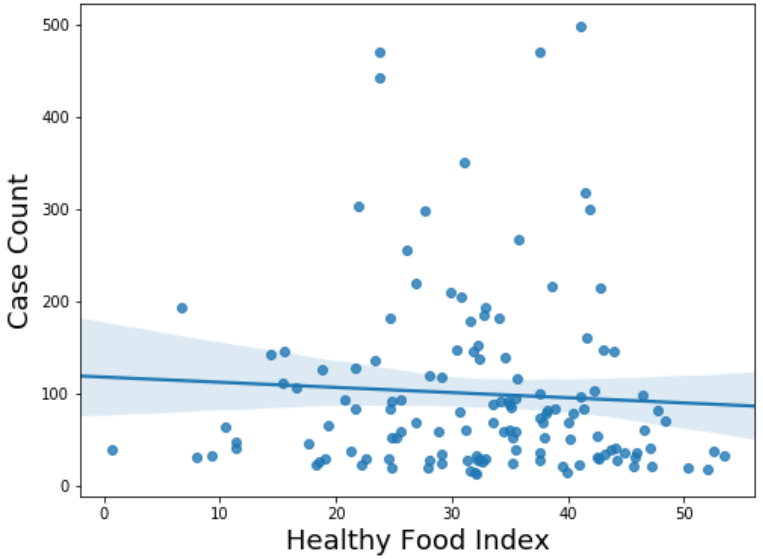
Results

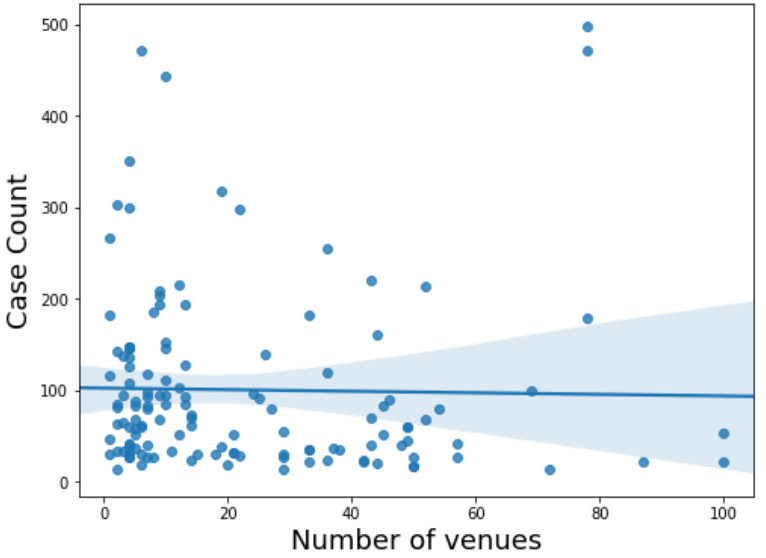
- It is noticed that number of covid-19 cases are higher among neighborhoods hosting higher number of low income families.



Results

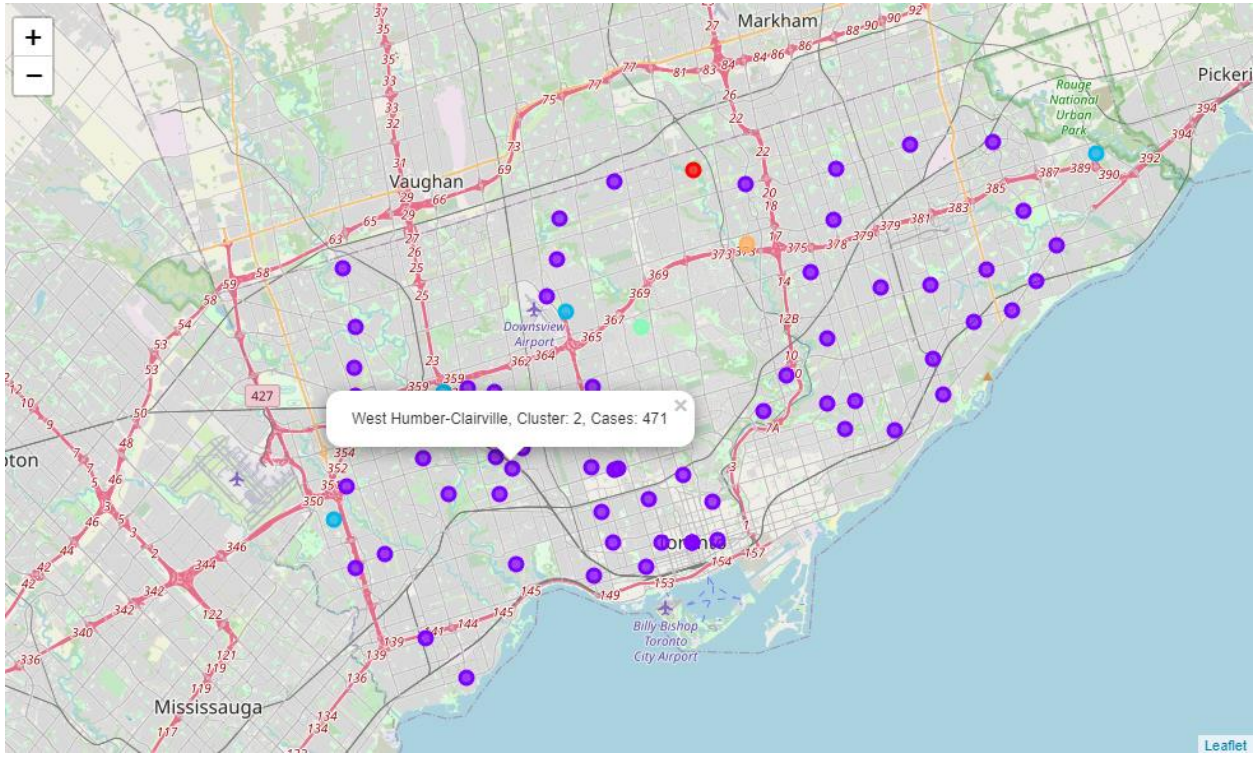
- Healthy food index and number of venues in each neighborhood did not show a clear trend with number of covid-19 cases.





Results

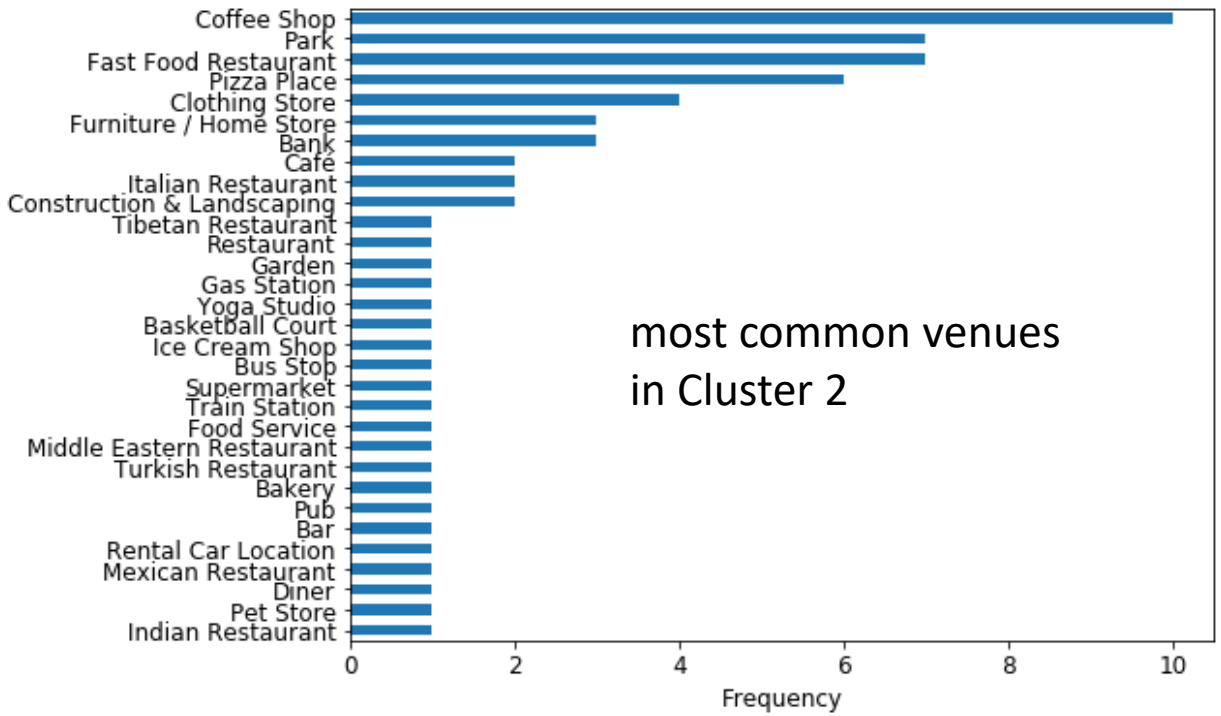
- Finally, using K-means algorithm the top 40 neighborhoods, sorted by covid-19 case counts, are clustered into 5 clusters



Results

- The most common venues in each cluster based on frequency are *Dog Run* in cluster 1, *Coffee Shop* and *Fast Food Restaurant* in cluster 2 (bar plot below), *Park* in cluster 3, *Seafood Restaurant* in cluster 4, and *Tennis Court* in cluster 5.

- From this clustering process, it is observed that, cluster 2 corresponds to neighborhoods with high number of venues and case counts. Hence, suitable preventive measures should put in place in these neighborhoods.



Discussion

- From the data analysis conducted in the results section, it is observed that number of covid-19 case counts in each neighborhood can be related to neighborhoods population, number of low income families, number of seniors and working population in each neighborhood, as well as the types of venues available within each neighborhood.

- Such parameters together with other variables being tracked on a daily basis could help the decision makers to choose the best strategy for each neighborhood to control the outbreak efficiently.

- Decisions made for a neighborhood with high number of seniors population, might have to be different from the decisions made for a neighborhood with high number of working population.

- Similarly, neighborhoods with high number of restaurants require preventive measures different from those with high number of parks and recreation facilities.

Conclusion

- The covid-19 case counts are studied in this work for neighborhoods in the City of Toronto.
- Total population of each neighborhood, number of low income families, working population, and type of venues were recognized to be parameters that could have an effect on the number of covid-19 cases in each neighborhood.
- This approach to the covid-19 data could present a more focused point of view for defining an effective strategy in tackling the outbreak at neighborhood level.

References

[1] COVID-19: Status of Cases in Toronto, link: https://www.toronto.ca/home/covid-19/covid-19-latest-city-of-toronto-news/covid-19-status-of-cases-in-toronto/

[2] GeoPy's documentation, link: https://geopy.readthedocs.io/en/stable/#nominatim

[3] Wellbeing Toronto, link: https://map.toronto.ca/wellbeing