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. One way to do so, is by studying the spread of covid-19 at a greater level of granularity. Therefore, instead of viewing total counts in states and provinces, study the number of cases at neighborhood levels. Such approach could present a more focused point of view for defining an effective testing strategy making best use of available resources.

Data

To study the number of covid-19 cases in Toronto at neighborhood level, covid-19 neighborhood data made available via City of Toronto municipality [this link] is used here. Moreover, the age distribution of cases, and the severity by age is acquired from the same source. The coordinates of each neighborhood are obtained from Nominatim geolocation service through GeoPy library in Python [this link]. Moreover, Toronto neighborhoods population data is extracted from Wellbeing Toronto application [this link] to see if there is a relationship between the number of covid-19 cases in each neighborhood with population, level of income and age. Finally, Foursquare venues data are used to see what type of venues are most popular in the neighborhood with highest numbers of covid-19 cases. See Tables 1 to 6 in the following.

Table 1. Top 10 neighborhoods in terms of covid-19 cases

Neighbourhood ID	Neighbourhood Name	Case Count	Rate per 100,000 people
NaN	Missing Address/Postal Code	613	NaN
25.0	Glenfield-Jane Heights	499	1636.548490
1.0	West Humber-Clairville	471	1413.904899
2.0	Mount Olive-Silverstone-Jamestown	471	1429.265036
26.0	Downsview-Roding-CFB	443	1263.836586
27.0	York University Heights	420	1522.125177
14.0	Islington-City Centre West	351	798.362334
113.0	Weston	318	1767.452201
131.0	Rouge	303	651.668961
24.0	Black Creek	299	1375.534802

Table 2. Geolocation data added to neighborhoods data using GeoPy library

Neighbourhood ID	Neighbourhood Name	Case Count	Rate per 100,000 people	Latitude	Longitude
25.0	Glenfield-Jane Heights	499	1636.548490	43.6535	-79.3839
1.0	West Humber-Clairville	471	1413.904899	43.6824	-79.4808
2.0	Mount Olive-Silverstone-Jamestown	471	1429.265036	43.6535	-79.3839
26.0	Downsview-Roding-CFB	443	1263.836586	43.7493	-79.4622
14.0	Islington-City Centre West	351	798.362334	43.6488	-79.549
113.0	Weston	318	1767.452201	43.7002	-79.5162
131.0	Rouge	303	651.668961	43.8049	-79.1658
24.0	Black Creek	299	1375.534802	43.6954	-79.4855
137.0	Woburn	298	557.165560	43.7598	-79.2253
6.0	Kingsview Village-The Westway	267	1213.636364	43.6995	-79.5563

Table 3. Age distribution of covid-19 cases in the City of Toronto

	Age Group	% of Total Case Count	Case Count
0	0-19	6.122038	912
1	20-29	14.298181	2130
2	30-39	14.063234	2095
3	40-49	14.076660	2097
4	50-59	15.808552	2355
5	60-69	10.706854	1595
6	70-79	7.068537	1053
7	80-89	10.263812	1529
8	90+	7.377324	1099
9	Unknown	0.214808	32

Table 4. Severity of covid-19 cases by age group in the City of Toronto

	Age Group	ICU Cases	Deaths	Hospitalized Cases	Intubated Cases
0	0-19	3	1	13	1
1	20-29	11	1	51	7
2	30-39	12	1	85	7
3	40-49	40	9	147	24
4	50-59	96	34	301	77
5	60-69	109	102	337	80
6	70-79	86	208	355	63
7	80-89	39	398	383	27
8	90+	4	366	174	1
9	Unknown	0	0	0	0

Table 5. Demographics data acquired from Wellbeing Toronto at neighborhoods level

	Neighbourhood	Neighbourhood Id	Low Income Families	Healthy Food Index	Total Area	Total Population	Pop 0 - 4 years	Pop 5 - 9 years	Pop 10 - 14 years	Pop 15 - 19 years	...	Pop 55 - 59 years	Pop 60 - 64 years	Pop 65 - 69 years	Pop 70 - 74 years	Pop 75 - 79 years	Pop 80 - 84 years	Pop 85 - 89 years	Pop 90 - 94 years	Pop 95 - 99 years	Pop 100 years and over
0	West Humber-Clairville	1	3790	23.82	30.09	33312	1540	1720	1790	2325	...	2195	1795	1595	1185	885	700	615	160	50	5
1	Mount Olive-Silverstone-Jamestown	2	5010	37.57	4.60	32954	2190	2500	2415	2585	...	1955	1520	1285	885	630	465	300	70	10	0
2	Thistletown-Beaumond Heights	3	1080	42.26	3.40	10360	540	600	595	650	...	660	535	490	375	335	320	350	100	20	5
3	Rexdale-Kipling	4	1110	23.31	2.50	10529	560	515	565	635	...	870	650	520	350	295	270	300	85	15	0
4	Elms-Old Rexdale	5	1190	24.71	2.90	9456	540	605	660	690	...	730	525	415	305	235	180	145	40	5	0
5	Kingsview Village-The Westway	6	2800	35.69	5.10	22000	1425	1485	1325	1500	...	1330	1220	1015	750	660	595	575	170	20	0
6	Willowridge-Martingrove-Richview	7	1610	30.46	5.50	22156	1175	1230	1155	1275	...	1600	1305	1245	945	920	910	885	240	30	10
7	Humber Heights-Westmount	8	820	32.82	2.80	10948	440	490	515	525	...	790	720	580	485	495	530	950	375	70	15
8	Edenbridge-Humber Valley	9	900	46.41	5.50	15535	565	705	845	885	...	1230	1155	910	635	550	520	665	200	55	0
9	Princess-Rosethorn	10	460	39.52	5.20	11051	450	585	730	780	...	1045	810	630	435	335	315	325	115	10	0

Table 6. Merged demographics and geolocation dataframes

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
74.0	North St. James Town	182	977.706151	43.6694	-79.3727	3450	34.03	0.4	18690	3025	3805	4040	2645	2520	1520	655	400	80	46725.000000
104.0	Mount Pleasant West	160	539.483445	43.7084	-79.3901	3220	41.60	1.3	30095	3205	5690	7285	4005	3435	2715	1775	1530	455	23150.000000
75.0	Church-Yonge Corridor	89	283.982131	43.6799	-79.3887	3860	33.50	1.4	31430	2310	9560	7235	3870	3935	2555	1280	615	70	22450.000000
72.0	Regent Park	31	286.957327	43.6607	-79.3605	2360	42.52	0.6	10840	2380	2240	2005	1540	1460	785	295	115	20	18066.666667
73.0	Moss Park	214	1043.596996	43.6546	-79.3697	3540	42.76	1.4	20590	2060	4330	4960	3165	2965	1965	775	300	70	14707.142857
76.0	Bay Street Corridor	53	205.450246	43.6673	-79.3885	10050	24.83	1.8	25920	3080	10020	5170	2145	2165	1540	1050	650	100	14400.000000
61.0	Taylor-Massey	82	522.859147	43.6976	-79.3012	2350	38.19	1.1	15785	3675	2235	2830	2380	2060	1340	645	505	115	14350.000000
84.0	Little Portugal	36	231.377338	43.6474	-79.4311	1300	37.58	1.2	15675	1820	3910	3975	1935	1465	1135	810	520	105	13062.500000
78.0	Kensington-Chinatown	99	551.685706	43.6534	-79.4002	3690	37.57	1.5	18090	2055	5320	3115	1920	2000	1515	1005	990	170	12060.000000
88.0	High Park North	51	230.123635	43.6574	-79.471	1810	40.24	1.9	22450	3635	3430	4590	3290	2810	2280	1215	940	260	11815.789474

Table 7. Top 5 most common venues in each neighborhood acquired from Foursquare venues

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Agincourt North	Bank	Bakery	Chinese Restaurant	Ice Cream Shop	Sandwich Place
1	Agincourt South-Malvern West	Chinese Restaurant	Asian Restaurant	Shopping Mall	Rental Car Location	Hong Kong Restaurant
2	Alderwood	Pizza Place	Pub	Coffee Shop	Skating Rink	Gym
3	Annex	Pizza Place	Park	Gym	Bistro	Thai Restaurant
4	Banbury-Don Mills	Park	Botanical Garden	Intersection	Coffee Shop	Trail
5	Bathurst Manor	Park	Convenience Store	Baseball Field	Playground	Dumpling Restaurant
6	Bay Street Corridor	Coffee Shop	Café	Clothing Store	French Restaurant	Sushi Restaurant
7	Bayview Village	Bank	Fish Market	Sandwich Place	Outdoor Supply Store	Persian Restaurant
8	Bayview Woods-Steeles	Dog Run	Yoga Studio	Eastern European Restaurant	Dive Bar	Doctor's Office
9	Bedford Park-Nortown	Construction & Landscaping	Rental Car Location	Yoga Studio	Dumpling Restaurant	Distribution Center

Methodology

The geolocation of neighborhoods is displayed in Figure 1 in a folium map. The geolocation data is merged with the demographics data from Wellbeing Toronto (Table 2 merged with Table 5). Using this new dataframe (Table 6) it is possible to analyze the effect of different demographics parameters on the covid-19 case count. For example, one can study the effect of total population vs case counts in each neighborhood. This is discussed in more detail in the results section of this report. Population density is another column that is added to this data. This additional feature can be calculated by dividing the "Total Population" by "Total Area". Moreover, the most common venues in each neighborhood is determined in Foursquare API using their geolocation (Table 7). This table is then merged with Table 6, which has the case counts and demographics data to add another dimension to the factors affecting covid-19 case counts in Toronto neighborhoods (Table 8).

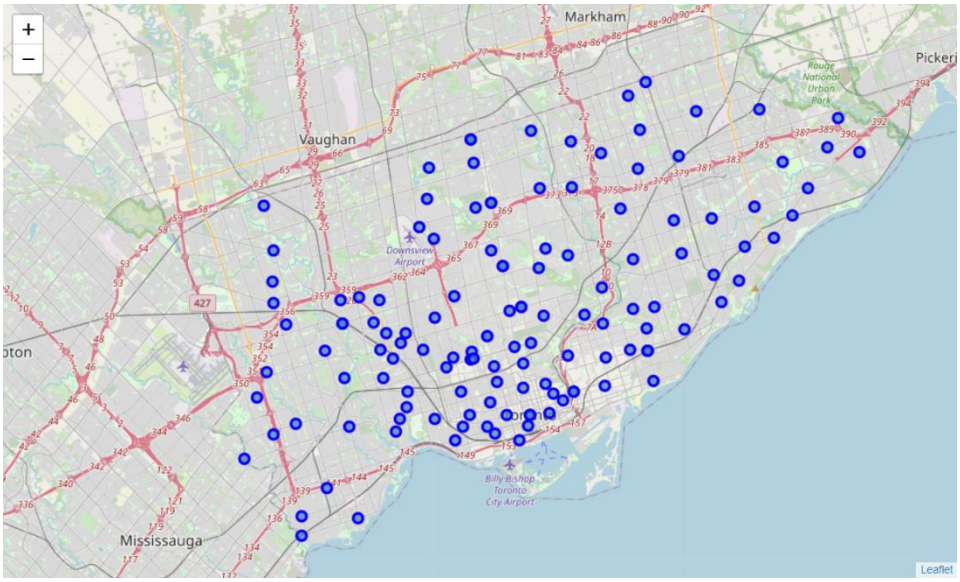

Figure 1. Geolocation of neighborhoods of the City of Toronto in a folium map

Table 8. Merged venues, demographics and neighborhood case counts dataframes

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

From Fig. 2, number of covid-19 case counts increases with total population in each neighborhood.

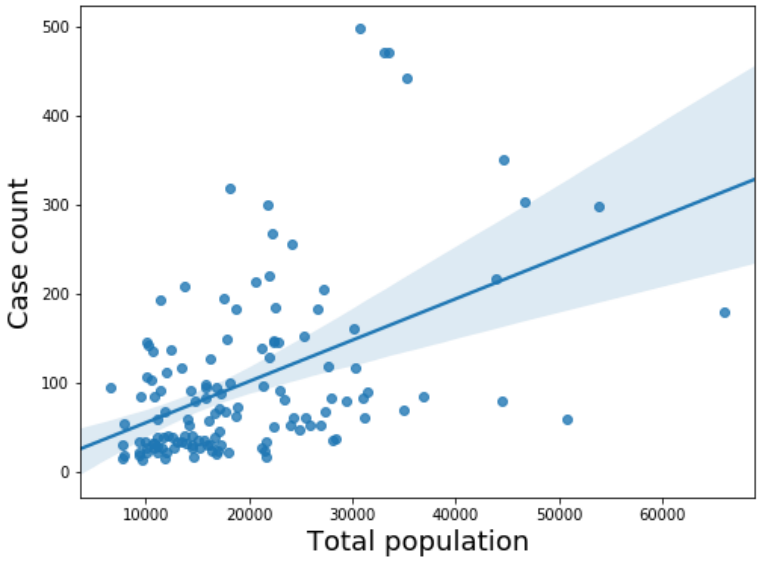

Figure 2. Total population of each neighborhood vs covid-19 case counts

From Fig. 3, it can be observed that distribution of covid-19 cases counts is higher among age groups between 20 to 59 years.

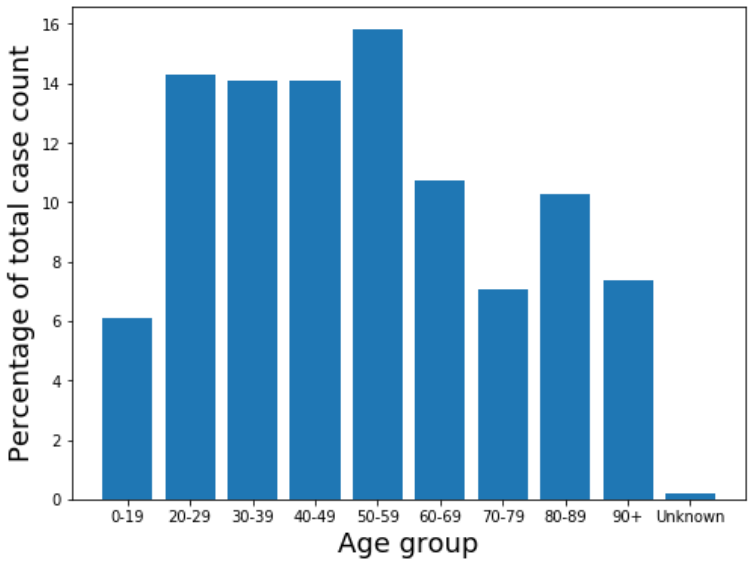

Figure 3. Distribution of covid-19 case counts among age groups

Bar plot in Fig. 4 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. In order to define suitable preventive measures to

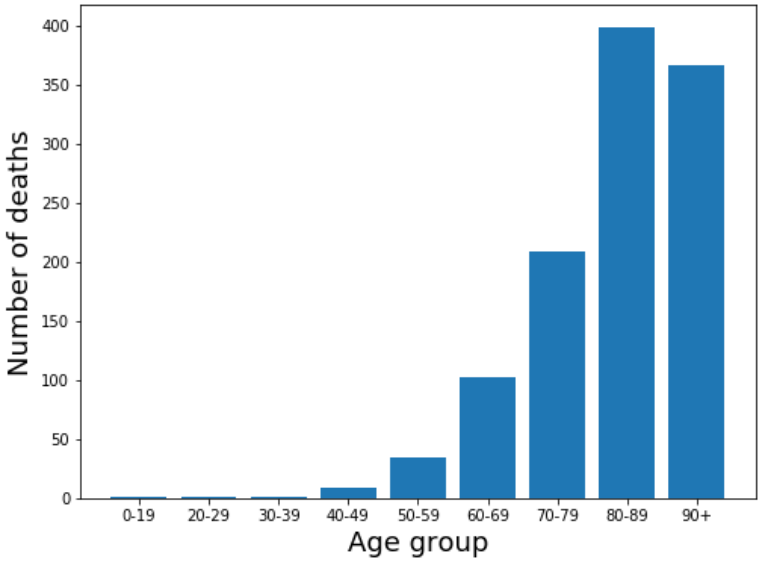

Figure 4. Number of covid-19 related deaths by age group

protect this age group from contracting the novel coronavirus, we need to know what neighborhoods have the highest number of senior citizens. Three age groups have been defined as young (0 to 19 years old), middle aged (20 to 59 years old) and senior (60+ years old) by adding up the smaller age group categories from Wellbeing Toronto data. Figure 5 shows the neighborhoods hosting a high number of seniors. As we can see from this figure, in 3 neighborhoods, namely L'Amoreaux, Woburn and Islington, more than 30,000 seniors live. Strict measures should be considered in such neighborhoods to protect seniors from covid-19.

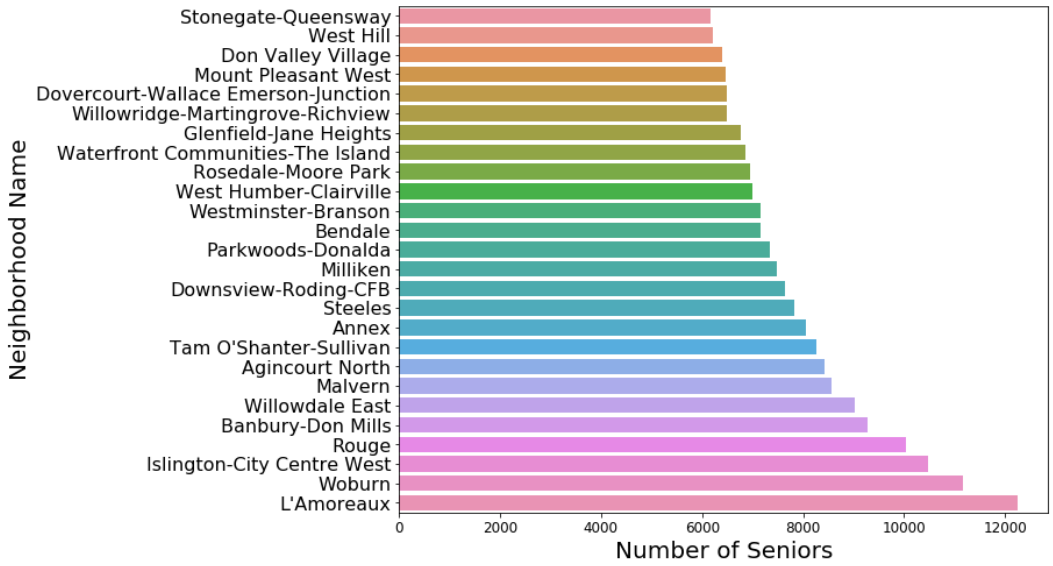

Figure 5. Neighborhoods hosting more than 6000 seniors

Fig. 6 depicts the covid-19 case counts against population age considering the distribution observed in Fig. 3. From this figure we can see the number of cases are significantly higher among ages 20 to 59 years old, which we can consider as working age population.

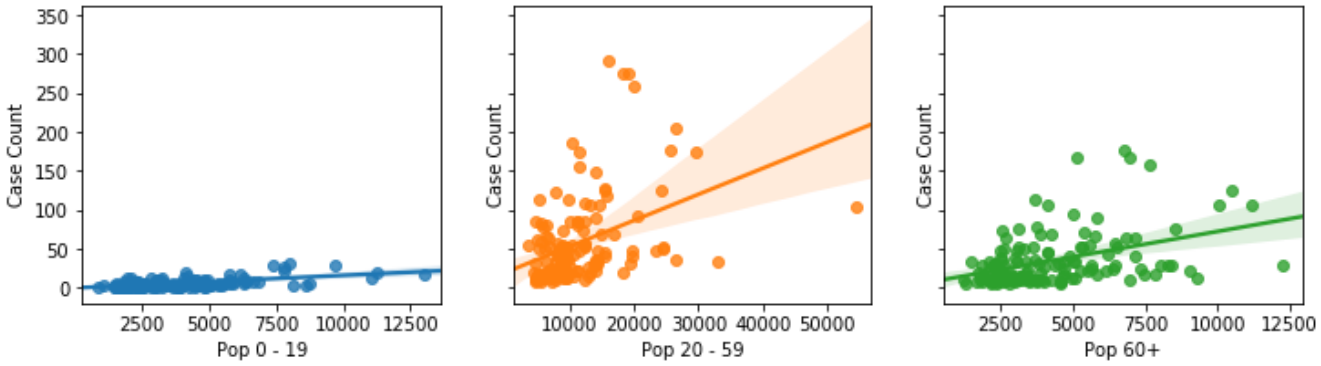

Figure 6. Population age in each neighborhood vs covid-19 case counts

Moreover, Fig 7. demonstrates that number of covid-19 cases are higher among neighborhoods hosting higher number of low income families.

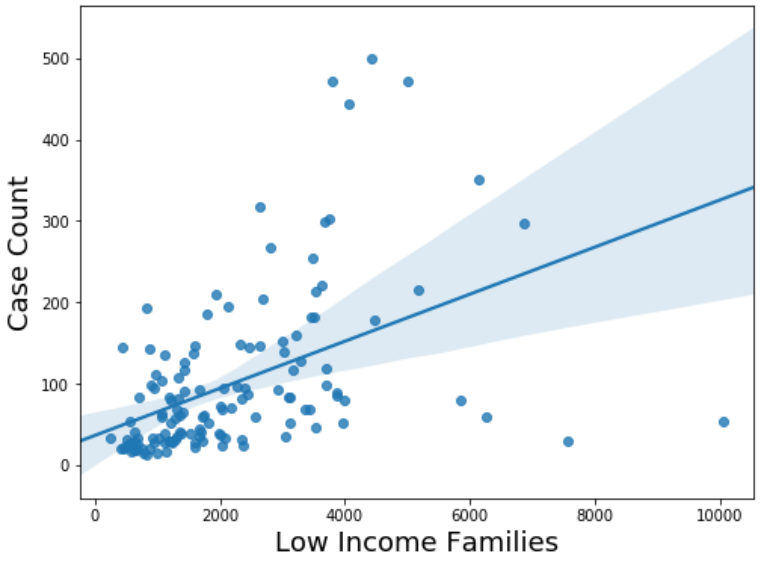

Figure 7. Number of low income families in each neighborhood vs covid-19 case counts

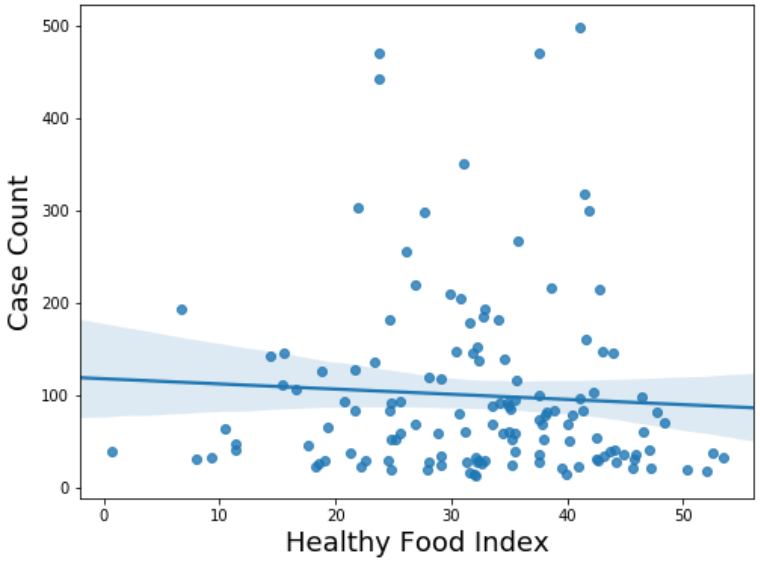

Figure 8. Healthy food index vs covid-19 case counts

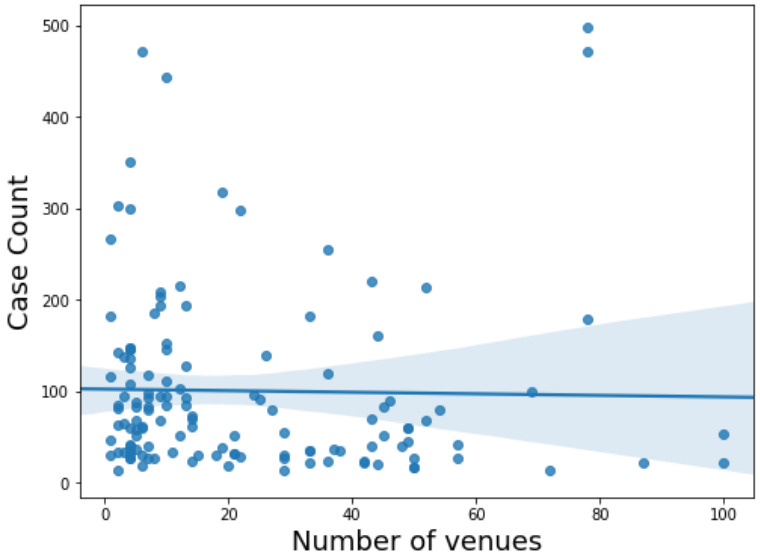

Figure 9. Number of venues around each neighborhood vs number of covid-19 cases

Healthy food index from Fig. 8, does not show a clear trend with number of covid-19 cases in neighborhoods. Also, number of venues around each neighborhood, in Fig. 9 acquired using Foursquare API, does not show a relationship with number of case counts.

Finally, using K-means algorithm the top 40 neighborhoods, sorted by covid-19 case counts, are clustered into 5 clusters, Fig. 10. 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 (see Fig. 11), *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.

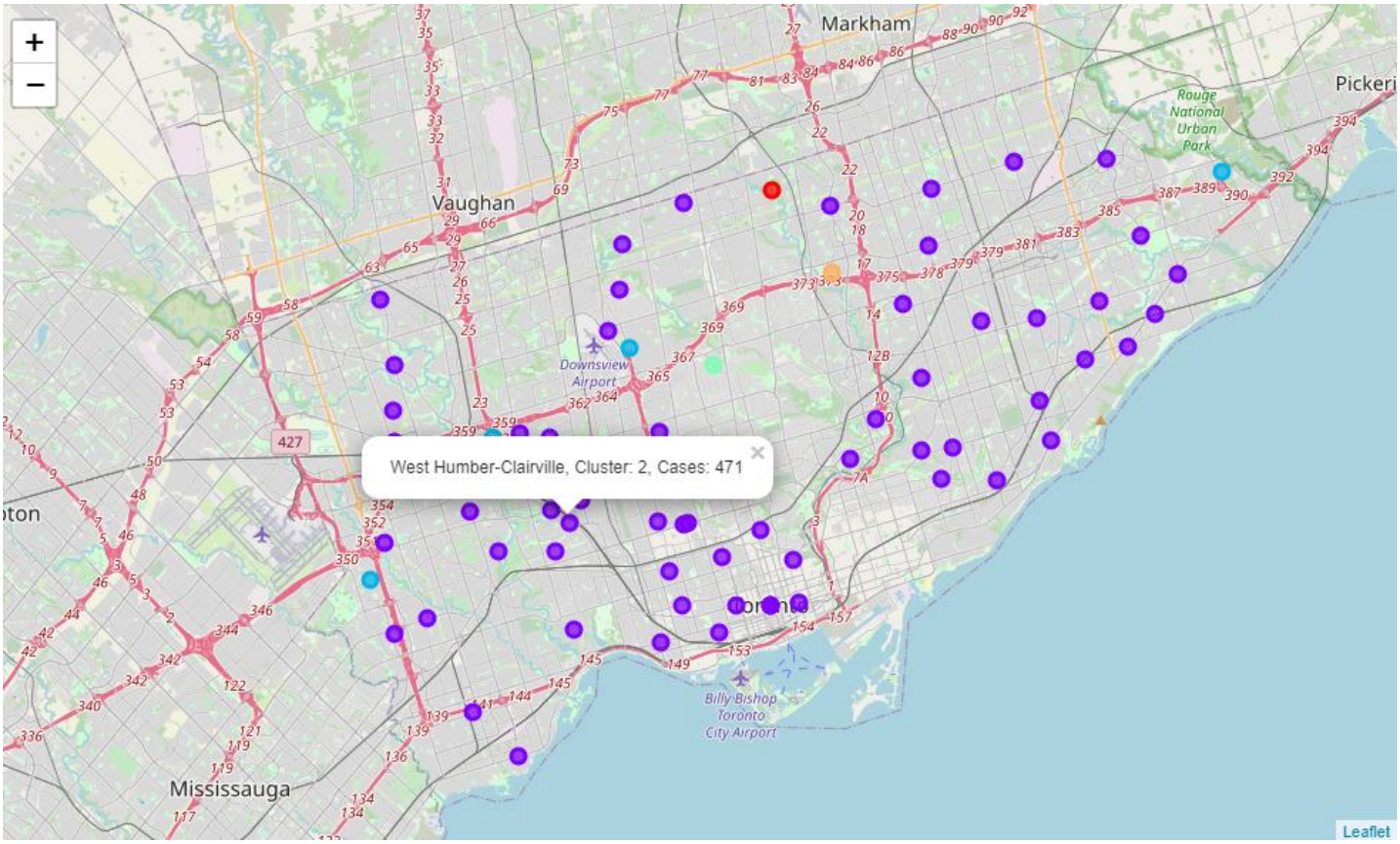

Figure 10. Clustering neighborhoods into 5 clusters using K-means algorithm

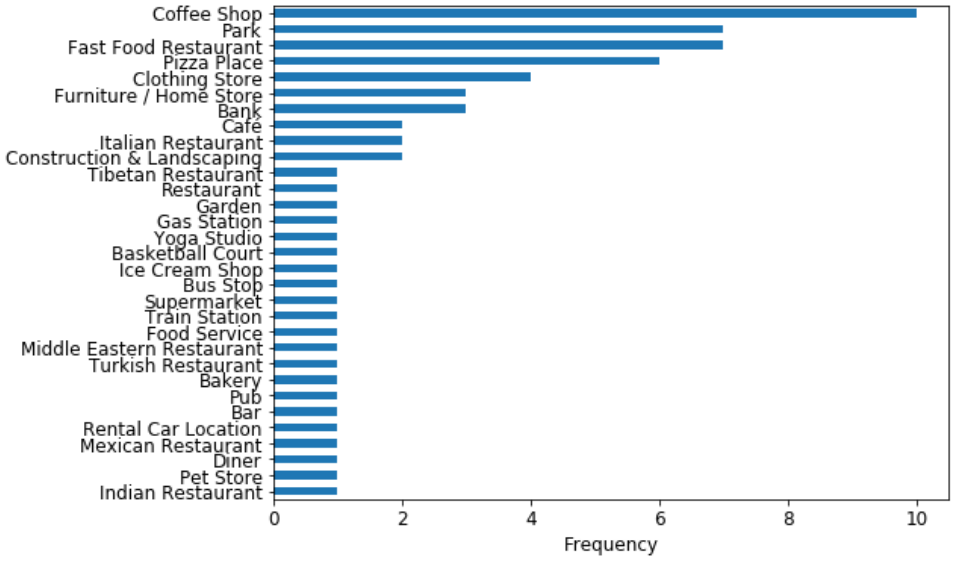

Figure 11. Frequency of venues in neighborhoods within cluster 2

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