Battle of Neighborhoods

Capstone Project

Business Problem

- Residents of Toronto Downton neighborhoods started looking for better places to live, as global COVID pandemic impacted busy financial center.
- In order to help with new neighborhood search, we developed a model based on Toronto geolocation and Foursquare venues. The model contains venues which are highly attractable during pandemic, such as parks, dog playgrounds, outdoor venues, and medical centers.

Data source

- Toronto neighborhoods listing with geolocations from Wikipedia.
- Listing of Toronto amenities and facilities from Foursquare API.
- We will focus on amenities which allow social distancing, considering specific of current COVID pandemic and focus of this research.

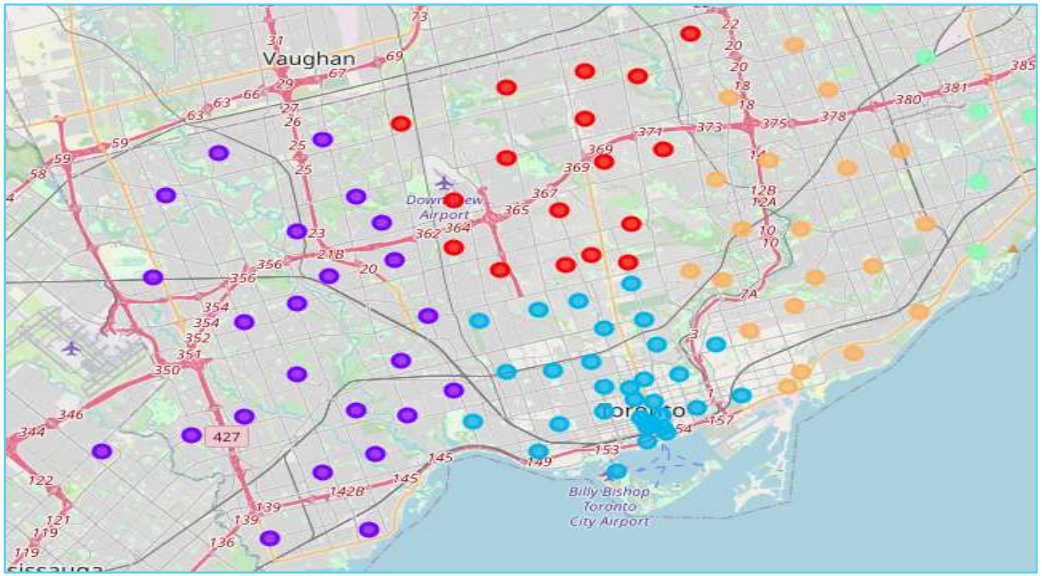
Methodology

- Set up Toronto neighborhood data frame with geolocations.
- Clean the data, from blanked and repeating entries.

	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M3A	North York	Parkwoods	43.753259	-79.329656
1	M4A	North York	Victoria Village	43.725882	-79.315572
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

Methodology

- Analyze neighborhood clusters with folium maps



Methodology

- Create Toronto venues data frame from Foursquare, and merge it with Toronto geolocation neighborhoods.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Parkwoods	43.753259	-79.329656	Brookbanks Park	43.751976	-79.332140	Park
1	Parkwoods	43.753259	-79.329656	TTC stop #8380	43.752672	-79.326351	Bus Stop
2	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
3	Parkwoods	43.753259	-79.329656	Corrosion Service Company Limited	43.752432	-79.334661	Construction & Landscaping
4	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena

Methodology

- Drop venues, which are not COVID friendly, such as restaurants and entertainment centers. Keep only parks, medical centers and grocery stores.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Agincourt	Skating Rink	Train Station	Discount Store	Grocery Store	Golf Course
1	Alderwood, Long Branch	Skating Rink	Train Station	Discount Store	Grocery Store	Golf Course
2	Bathurst Manor, Wilson Heights, Downsview North	Supermarket	Pharmacy	Train Station	Discount Store	Grocery Store
3	Bayview Village	Train Station	Dog Run	Harbor / Marina	Grocery Store	Golf Course
4	Bedford Park, Lawrence Manor East	Pharmacy	Grocery Store	Discount Store	Golf Course	Garden

Methodology

- Perform correlation and descriptive statistics analysis to learn any data patterns.

	Athletics & Sports	Bakery	Baseball Field	Beach	Bus Line	Bus Station	Bus Stop	Curling Ice	Department Store
Athletics & Sports	1.000000	0.036051	-0.008427	-0.025282	-0.051177	-0.036735	0.254850	0.555525	-0.051752
Bakery	0.036051	1.000000	-0.061966	0.030629	0.097458	0.164543	-0.057182	-0.041716	-0.069869
Baseball Field	-0.008427	-0.061966	1.000000	-0.016380	-0.033159	-0.021867	-0.022453	-0.016380	-0.031688
Beach	-0.025282	0.030629	-0.016380	1.000000	-0.021767	-0.015624	-0.0		
Bus Line	-0.051177	0.097458	-0.033159	-0.021767	1.000000	0.190643	-0.0		
Bus Station	-0.036735	0.164543	-0.021867	-0.015624	0.190643	1.000000	-0.0		
Bus Stop	0.254850	-0.057182	-0.022453	-0.014739	-0.029836	-0.021417	1.0		
Curling Ice	0.555525	-0.041716	-0.016380	-0.010753	-0.021767	-0.015624	0.4		
Department Store	-0.051752	-0.069869	-0.031688	-0.022011	-0.044557	0.686015	-0.0		
Discount Store	0.009177	-0.063749	-0.022162	-0.030229	-0.061193	0.418774	-0.0		
Dog Run	-0.033884	-0.055911	-0.021954	-0.014411	-0.029173	-0.020940	-0.0		

	Athletics & Sports	Bakery	Baseball Field	Beach	Bus Line	Bus Station	Bus Stop	Curling Ice	D
count	94.000000	94.000000	94.000000	94.000000	94.000000	94.000000	94.000000	94.000000	
mean	0.006218	0.019547	0.009733	0.000180	0.010106	0.002996	0.004179	0.001520	
std	0.025642	0.048848	0.061946	0.001748	0.048405	0.019988	0.029561	0.014735	
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
50%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
75%	0.000000	0.010000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	
max	0.142857	0.250000	0.500000	0.016949	0.250000	0.166667	0.250000	0.142857	

Methodology

- Finalize model to pull data for specific customer needs and preferences based on mean frequency. For example, if someone looking for a neighborhood with park, here is a top selection.

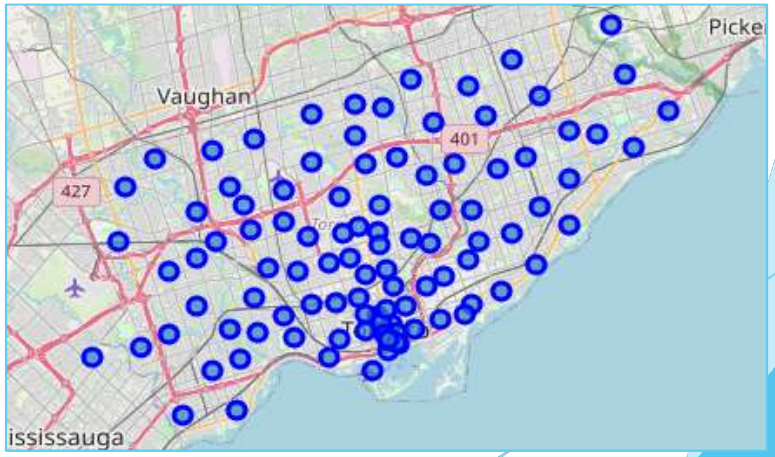
	Neighborhood	Park
88	Willowdale, Newtonbrook	1.000000
10	Caledonia-Fairbanks	0.500000
65	Rosedale	0.500000
93	York Mills West	0.500000
50	Milliken, Agincourt North, Steeles East, L'Amo...	0.333333
81	The Kingsway, Montgomery Road, Old Mill North	0.333333
26	East Toronto, Broadview North (Old East York)	0.250000
30	Forest Hill North & West, Forest Hill Road Park	0.250000
44	Kingsview Village, St. Phillips, Martin Grove ...	0.250000
46	Lawrence Park	0.250000

RESULTS

- As a result of this project, we build a model of Toronto neighborhoods with COVID friendly venues.
- The model is based on the Toronto geolocation data, and Foursquare venues.
- The model allows find suitable neighborhood based on specific preferences, such as park hospitals, dog playgrounds, grocery stores and proximity to public transit.

DISCUSSION

- I observed that there is no positive / negative correlations between current data set.
- Another thing is concentration of 103 neighborhoods. We can clearly see that neighborhoods are heavily concentrated around Toronto's downtown core, and less spread toward to city borders.



CONCLUSION

- Residents of Toronto Downton neighborhoods started looking for better places to live, as global COVID pandemic impacted busy financial center.
- In order to help with new neighborhood search, we developed a model based on Toronto geolocation and Foursquare venues. The model contains venues which are highly attractable during pandemic, such as parks, dog playgrounds, outdoor venues, and medical centers.
- Models allows to factor specific individual needs of each person seeking to relocate to a new neighborhood and help them find optimal living solution.
- For example, a dog owner currently living in Toronto downtown would be interested in Hillcrest Village neighborhood with the highest mean playground rating among 103 Toronto neighborhoods.

	Neighborhood	Dog Run
37	Hillcrest Village	0.200000
59	Parkdale, Roncesvalles	0.071429
15	Church and Wellesley	0.012500
0	Agincourt	0.000000
1	Alderwood, Long Branch	0.000000