Best Coffee Spots of Toronto

By Alex Dardick

I. Introduction

Today we are going to discover the best locations in Toronto to open up a new coffee shop. This information will help any young entrepreneur figure out where exactly the best spots are to target opening up a new coffee shop. When finding a good location, it's important to keep in mind where your competition is and what areas your target market will be attracted to. So what we're going to do is discover where the locations of venues that would host people who would want to drink coffee more often than others and find out what neighborhoods host the highest balance of these locations to coffee shops that already exist around those parts. Some of these locations will include schools, colleges, libraries, office buildings, and malls. We will use K Clustering to group neighborhoods as either a risk neighborhood, a possible neighborhood, or a great neighborhood.

II. Data

For this project, we'll simply be using Foursquare to figure out which areas have the highest concentration of desired locations and which areas have the highest concentration of competitors. We'll be able to use the number of good venues that host people who want to drink coffee, the number of venues that represent competition, and the ratio of these two numbers within a certain radius for each neighborhood in order to cluster Toronto up accordingly. This way, we'll be able to see what kinds locations are within a walking distance to each neighborhood, and whether or not they would be supportive of a new coffee shop business. I'll have to use a CSV file I have of latitudes and longitudes of postal codes throughout Toronto in order to accurately find close enough venues to our neighborhoods.

III. Methodology

A. Retrieving Our Data:

The first thing we want to get all of our data together, so I created a beautiful soup object to scrape a Wikipedia page of all neighborhoods/boroughs/postal codes of Toronto and set it to a dataframe like this:

	Postal Code	Borough	Neighborhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Queen's Park	Ontario Provincial Government
...	...	...	...
98	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North
99	M4Y	Downtown Toronto	Church and Wellesley
100	M7Y	East TorontoBusiness reply mail Processing Cen...	Enclave of M4L
101	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...
102	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...

103 rows × 3 columns

This was then the time to merge each neighborhood with its coordinates from the CSV file I had. Since the dataframe above and csv file I had both have the postal codes as attributes, I was able to merge them based on that like so:

	Postal Code	Borough	Neighborhood	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	Queen's Park	Ontario Provincial Government	43.662301	-79.389494

B. Finding Nearby Venues:

Now, at this point, I wanted to run an algorithm that would return all of the venues within a certain radius of each neighborhood. After some research I figured 500 meters is a reasonable radius for each neighborhood. The reason for this is because we want people to be able to walk to our coffee shop from where they are at instead of worrying about driving, and 500 meters came out to about the maximum reasonable walking distance. We're going to have to use Foursquare data in order to properly retrieve the data we're looking for, so after putting in our credentials and running the algorithm, we get the following dataframe:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue Name	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	KFC	43.754387	-79.333021	Fast Food Restaurant
2	Parkwoods	43.753259	-79.329656	GTA Restoration	43.753396	-79.333477	Fireworks Store
3	Parkwoods	43.753259	-79.329656	Variety Store	43.751974	-79.333114	Food & Drink Shop
4	Victoria Village	43.725882	-79.315572	Victoria Village Arena	43.723481	-79.315635	Hockey Arena
5	Victoria Village	43.725882	-79.315572	Tim Hortons	43.725517	-79.313103	Coffee Shop
6	Victoria Village	43.725882	-79.315572	Portugril	43.725819	-79.312785	Portuguese Restaurant
7	Victoria Village	43.725882	-79.315572	Eglinton Ave E & Sloane Ave/Bermondsey Rd	43.726086	-79.313620	Intersection

Now that we have each of the venues around all of our neighborhoods, we can now one hot encode our dataframe by the 'Venue Category' attribute in order to quantify each venue by category. We can then group up each neighborhood and find out how much of each Venue Category every neighborhood has. We can use a sum function here on the groupby function to get:

	Neighborhood	Yoga Studio	Accessories Store	Airport	Airport Food Court	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	Antique Shop	...	Trail	Train Station
0	Agincourt	0	0	0	0	0	0	0	0	0	...	0	0
1	Alderwood, Long Branch	0	0	0	0	0	0	0	0	0	...	0	0
2	Bathurst Manor, Wilson Heights, Downsview North	0	0	0	0	0	0	0	0	0	...	0	0
3	Bayview Village	0	0	0	0	0	0	0	0	0	...	0	0
4	Bedford Park, Lawrence Manor East	0	0	0	0	0	0	0	1	0	...	0	0
...	...	...	...	...	...	...	...	...	...	...	...	...	...

C. What's good and what's bad?

Now at this point, we have to figure out exactly what we're looking for. We want to be able to decide what venues are beneficial to having near our coffee shop, and what venues are detrimental to the success of our new coffee shop? Let's look at all of the unique venues around Toronto:

```
toronto_venues['Venue Category'].unique()
```

executed in 17ms, finished 11:38:31 2021-08-22

```
array(['Park', 'Fast Food Restaurant', 'Fireworks Store',
       'Food & Drink Shop', 'Hockey Arena', 'Coffee Shop',
       'Portuguese Restaurant', 'Intersection', 'Pizza Place', 'Bakery',
       'Distribution Center', 'Restaurant', 'Spa', 'Pub',
       'Gym / Fitness Center', 'Historic Site', 'Breakfast Spot',
       'Chocolate Shop', 'Performing Arts Venue', 'Farmers Market',
       'Dessert Shop', 'French Restaurant', 'Mexican Restaurant',
       'Theater', 'Yoga Studio', 'Event Space', 'Café', 'Boutique',
       'Furniture / Home Store', 'Vietnamese Restaurant',
       'Clothing Store', 'Accessories Store', 'Miscellaneous Shop',
       'Italian Restaurant', 'Beer Bar', 'Sushi Restaurant', 'Creperie',
       'Fried Chicken Joint', 'Hobby Shop', 'Burrito Place', 'Diner',
       'Japanese Restaurant', 'Smoothie Shop', 'Bank', 'Sandwich Place',
       'Gym', 'College Auditorium', 'Bar', 'Caribbean Restaurant',
       'Baseball Field', 'Athletics & Sports', 'Gastropub', 'Pharmacy',
       'Pet Store', 'Flea Market', 'Plaza', 'Comic Shop',
       'Ramen Restaurant', 'Burger Joint', 'Music Venue',
       'Electronics Store', 'Movie Theater', 'Thai Restaurant',
       'Steakhouse', 'Shopping Mall', 'Art Gallery', 'Tanning Salon',
       'Pool', 'Middle Eastern Restaurant', 'Construction & Landscaping',
       'Discount Store', 'Grocery Store', 'Bike Shop',
       'Sporting Goods Shop', 'Beer Store', 'Supermarket',
       'Dim Sum Restaurant', 'Asian Restaurant', 'Chinese Restaurant',
       'Skating Rink', 'Curling Ice', 'Bus Stop', 'Cosmetics Shop',
       'American Restaurant', 'BBQ Joint', 'Church', 'Hotel',
       'Camera Store', 'New American Restaurant',
       'Vegetarian / Vegan Restaurant', 'Field', 'Trail', 'Liquor Store',
       'Shopping Plaza', 'Rental Car Location', 'Donut Shop',
       'Medical Center', 'Health Food Store', 'Neighborhood', 'Museum',
       'Seafood Restaurant', 'Fountain', 'Cocktail Bar', 'Concert Hall',
       'Fish Market', 'Basketball Stadium', 'Cheese Shop', 'Jazz Club',
       'Greek Restaurant', "Women's Store", 'Korean BBQ Restaurant',
       'Indian Restaurant', 'Sports Bar', 'Fish & Chips Shop',
       'Department Store', 'Brewery', 'Modern European Restaurant',
       'Bubble Tea Shop', 'Tea Room', 'Art Museum', 'Falafel Restaurant',
       'Office', 'Candy Store', 'Baby Store', 'Nightclub',
       'Hakka Restaurant', 'Gas Station', 'Golf Course',
       'Mediterranean Restaurant', 'Dog Run', 'Deli / Bodega',
       'Bridal Shop', 'Ice Cream Shop', 'Mobile Phone Shop',
       'Warehouse Store', 'Speakeasy', 'Lounge', 'Smoke Shop',
       'General Travel', 'Monument / Landmark', 'Food Court',
       'Playground', 'Business Service', 'Toy / Game Store',
       'Salon / Barbershop', 'Juice Bar', 'Video Game Store',
       'Massage Studio', 'Convenience Store', 'Salad Place', 'Lake',
```

```
'IT Services', 'Roof Deck', 'Bistro', 'Aquarium', 'Dance Studio',
'Korean Restaurant', 'Wine Bar', 'Cuban Restaurant', 'Record Shop',
"Men's Store", 'Airport', 'Fruit & Vegetable Store', 'Bookstore',
'Tibetan Restaurant', 'Train Station', 'Climbing Gym', 'Stadium',
'Bus Line', 'Metro Station', 'Bus Station', 'Soccer Field',
'Board Shop', 'Gluten-free Restaurant', 'Tailor Shop',
'Basketball Court', 'Motel', 'Food Truck', 'Gay Bar',
'Comfort Food Restaurant', 'Stationery Store',
'Latin American Restaurant', 'Coworking Space', 'Butcher',
'Food Service', 'General Entertainment', 'Farm', 'College Stadium',
'Arts & Crafts Store', 'Swim School', 'Health & Beauty Service',
'Garden', 'Jewelry Store', 'Antique Shop',
'Cajun / Creole Restaurant', 'Irish Pub', 'Auto Garage',
'History Museum', 'Shoe Repair', 'Flower Shop',
'Eastern European Restaurant', 'Gift Shop', 'Gourmet Shop',
'College Gym', 'College Arts Building', 'Noodle House',
'Indie Movie Theater', 'Lawyer', 'Organic Grocery',
'Arepa Restaurant', 'Belgian Restaurant', 'Dumpling Restaurant',
'Thrift / Vintage Store', 'Light Rail Station', 'Airport Lounge',
'Harbor / Marina', 'Airport Food Court', 'Airport Terminal',
'Plane', 'Airport Service', 'Sculpture Garden', 'Boat or Ferry',
'Drugstore', 'Truck Stop', 'Taiwanese Restaurant', 'Market',
'Snack Place', 'River', 'Theme Restaurant', 'Escape Room',
'Martial Arts School', 'Skate Park', 'Garden Center',
'Auto Workshop', 'Recording Studio', 'Wings Joint',
'Supplement Shop', 'Hardware Store', 'Kids Store'], dtype=object)
```

As we can see, there are many considerations to keep in mind. What venues are direct competition? What venues host people that regularly drink coffee? Which places would people be more likely to be hanging out around? Many things to consider, but I came down to the following lists of good venues and bad venues.

Good Venues: 'Office', 'Airport', 'College Arts Building', 'College Gym', 'Coworking Space', 'Metro Station', 'Bus Station', 'Train Station', 'IT Services', 'Dance Studio', 'Business Service', 'Art Museum', 'Museum', 'Shopping Plaza', 'Shopping Mall', 'Construction & Landscaping', 'Plaza', 'College Auditorium', 'Theater', 'Bookstore', 'Stationery Store', 'Truck Stop', 'History Museum', 'Auto Garage', 'Breakfast Spot'

Bad Venues: 'Coffee Shop', 'Café', 'Tea Room', 'Lounge', 'Food Court', 'Airport Food Court'

D.Create our new DataFrame:

At this point, I want to create a dataframe solely based on the information I'm looking for:

1) How many good venues are within walking distance of each neighborhood?
2) How many bad venues are within walking distance of each neighborhood?
3) What is the ratio of the good venues to the bad ones?

So I'm going to define a function that sums all the entries for a given row for a given list of attributes. This way I can run this function on two occasions; once to find the total good venues for each row, and another time to find all the bad venues for each row:

	Neighborhood	# of Good Venues	# of Bad Venues	Ratio of good/bad venues
0	Agincourt	1	1	NaN
1	Alderwood, Long Branch	0	1	NaN
2	Bathurst Manor, Wilson Heights, Downsview North	1	2	NaN
3	Bayview Village	0	1	NaN
4	Bedford Park, Lawrence Manor East	0	3	NaN

Now I'm going to write a function that will divide the good venues by the bad venues to find the ratio for each row in order to fill out the last attribute:

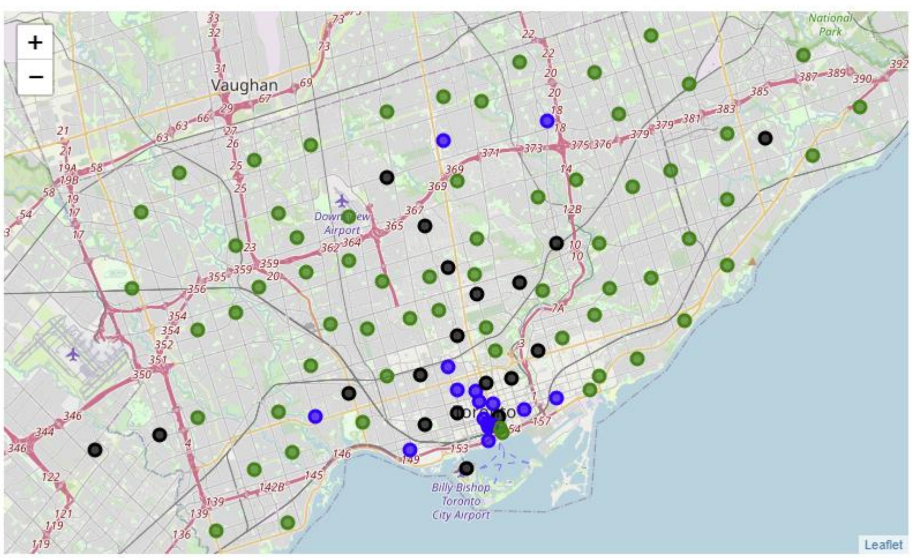
	Neighborhood	# of Good Venues	# of Bad Venues	Ratio of good/bad venues
0	Agincourt	1	1	1
1	Alderwood, Long Branch	0	1	0
2	Bathurst Manor, Wilson Heights, Downsview North	1	2	0.5
3	Bayview Village	0	1	0
4	Bedford Park, Lawrence Manor East	0	3	0

E. K-Clustering:

At this point we can finally run our K-Clustering algorithm! I want to only define 3 clusters for the purpose of hopefully getting a good, a bad, and an ugly cluster. Once we drop the ‘Neighborhood’ column, we can run the kmeans algorithm to fit our dataframe above. Afterwards, we’re going to combine the cluster results, along with our location dataframe from earlier, to our toronto_coffee dataframe above:

	Postal Code	Borough	Neighborhood	Latitude	Longitude	# of Good Venues	# of Bad Venues	Ratio of good/bad venues	Cluster Label
0	M3A	North York	Parkwoods	43.753259	-79.329656	0	0	0	0.0
1	M4A	North York	Victoria Village	43.725882	-79.315572	0	1	0	0.0
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	3	7	0.428571	1.0
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	0	1	0	0.0
4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494	2	8	0.25	1.0

We do this for the purpose of being able to call whatever information we want at this point from a single dataframe. Once we drop all the rows that have Null values, we are going to make folium map of the cluster we created.



Now that we see where all of our clustered neighborhoods are, let's get an idea of what kind of venues are in the clusters themselves.

Cluster 1:

	index	Postal Code	Borough	Neighborhood	Latitude	Longitude	# of Good Venues	# of Bad Venues	Ratio of good/bad venues	Cluster Label
0	0	M3A	North York	Parkwoods	43.753259	-79.329656	0	0	0	0.0
1	1	M4A	North York	Victoria Village	43.725882	-79.315572	0	1	0	0.0
2	3	M6A	North York	Lawrence Manor, Lawrence Heights	43.718518	-79.464763	0	1	0	0.0
3	6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353	0	0	0	0.0
4	7	M3B	North York	Don Mills North	43.745906	-79.352188	0	1	0	0.0
5	8	M4B	East York	Parkview Hill, Woodbine Gardens	43.706397	-79.309937	1	1	1	0.0
6	10	M6B	North York	Glencairn	43.709577	-79.445073	0	0	0	0.0
7	11	M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...	43.650943	-79.554724	0	0	0	0.0

- We can see cluster one has generally very few bad venues per neighborhood, but it also has very few good venues as well generally per neighborhood. This isn't a bad cluster, but maybe the other cluster fare a bit better.

Cluster 2:

	index	Postal Code	Borough	Neighborhood	Latitude	Longitude	# of Good Venues	# of Bad Venues	Ratio of good/bad venues	Cluster Label
0	2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636	3	7	0.428571	1.0
1	4	M7A	Queen's Park	Ontario Provincial Government	43.662301	-79.389494	2	8	0.25	1.0
2	9	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	3	6	0.5	1.0
3	24	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383	2	10	0.2	1.0
4	30	M5H	Downtown Toronto	Richmond, Adelaide, King	43.650571	-79.384568	1	7	0.142857	1.0
5	33	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556	2	5	0.4	1.0
6	36	M5J	Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	43.640816	-79.381752	4	4	1	1.0
7	42	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	2	7	0.285714	1.0
8	43	M6K	West Toronto	Brockton, Parkdale Village, Exhibition Place	43.636847	-79.428191	2	5	0.4	1.0
9	48	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	1	6	0.166667	1.0
10	54	M4M	East Toronto	Studio District	43.659526	-79.340923	3	5	0.6	1.0
11	59	M2N	North York	Willowdale South	43.770120	-79.408493	2	5	0.4	1.0
12	74	M5R	Central Toronto	The Annex, North Midtown, Yorkville	43.672710	-79.405678	1	5	0.2	1.0
13	80	M5S	Downtown Toronto	University of Toronto, Harbord	43.662696	-79.400049	5	5	1	1.0
14	81	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450	1	5	0.2	1.0
15	97	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280	0	9	0	1.0

- Cluster 2 certainly has way more good venues per neighborhood than cluster, but they also have far more bad venues as well, so this cluster seems a bit tricky to work with given the sheer amount of competition in these areas.

Cluster 3:

	index	Postal Code	Borough	Neighborhood	Latitude	Longitude	# of Good Venues	# of Bad Venues	Ratio of good/bad venues	Cluster Label
0	13	M3C	North York	Don Mills South	43.725900	-79.340923	1	2	0.5	2.0
1	15	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	1	4	0.25	2.0
2	17	M9C	Etobicoke	Eringate, Bloordale Gardens, Old Burnhamthorpe...	43.643515	-79.577201	1	2	0.5	2.0
3	22	M1G	Scarborough	Woburn	43.770992	-79.216917	0	2	0	2.0
4	23	M4G	East York	Leaside	43.709060	-79.363452	2	3	0.666667	2.0
5	25	M6G	Downtown Toronto	Christie	43.669542	-79.422564	0	4	0	2.0
6	28	M3H	North York	Bathurst Manor, Wilson Heights, Downsview North	43.754328	-79.442259	1	2	0.5	2.0
7	37	M6J	West Toronto	Little Portugal, Trinity	43.647927	-79.419750	1	2	0.5	2.0
8	41	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188	1	2	0.5	2.0
9	55	M5M	North York	Bedford Park, Lawrence Manor East	43.733283	-79.419750	0	3	0	2.0
10	69	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763	1	2	0.5	2.0
11	73	M4R	Central Toronto	North Toronto West	43.715383	-79.405678	0	3	0	2.0
12	76	M7R	MississaugaCanada Post Gateway Processing Centre	Enclave of L4W	43.636966	-79.615819	0	3	0	2.0
13	79	M4S	Central Toronto	Davisville	43.704324	-79.388790	0	4	0	2.0
14	84	M5T	Downtown Toronto	Kensington Market, Chinatown, Grange Park	43.653206	-79.400049	0	4	0	2.0
15	86	M4V	Central Toronto	Summerhill West, Rathnelly, South Hill, Forest...	43.686412	-79.400049	0	2	0	2.0
16	87	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420	1	2	0.5	2.0

- This cluster has less bad venues than cluster 2 generally but more than cluster 1. This would work well but unfortunately there aren't many good venues in any of these neighborhoods. So it seems cluster 1 is the best option for discovering plausible neighborhoods.

F. Our Good Cluster:

Now that we know which cluster has the best chance for finding a good neighborhood, we can sort it to find which neighborhoods have the highest number of good venues out the whole cluster. We quickly discover only 6 neighborhoods have more than 1 good venue while maintaining 1 or less bad venue. So we'll go ahead and separate this section of the dataframe out into it's own dataframe, and get a list of all the neighborhoods that show up in this dataframe:

	level_0	index	Postal Code	Borough	Neighborhood	Latitude	Longitude	# of Good Venues	# of Bad Venues	Ratio of good/bad venues	Cluster Label
0	48	75	M6R	West Toronto	Parkdale, Roncesvalles	43.648960	-79.456325	3	1	3	0.0
1	13	20	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	2	1	2	0.0
2	43	67	M4P	Central Toronto	Davisville North	43.712751	-79.390197	2	0	2	0.0
3	46	71	M1R	Scarborough	Wexford, Maryvale	43.750072	-79.295849	2	0	2	0.0
4	25	44	M1L	Scarborough	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577	2	0	2	0.0
5	24	40	M3K	North York	Downsview East	43.737473	-79.464763	2	0	2	0.0

```
['Parkdale, Roncesvalles',
 'Berczy Park',
 'Davisville North',
 'Wexford, Maryvale',
 'Golden Mile, Clairlea, Oakridge',
 'Downsview East']
```

At this point, we can now create a dataframe that hosts all venues that contain a neighborhood within this list. After this dataframe is created, we can then filter it so that all venues that have a category that is in our 'good_venues' list is sent to a new dataframe so that we have all the good venues around our 6 eligible neighborhoods. We can do the same to create a dataframe of all the bad venues around our eligible neighborhoods.

Good Venues:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue Name	Venue Latitude	Venue Longitude	Venue Category
550	Downsview East	43.737473	-79.464763	Toronto Downsview Airport (YZD)	43.738883	-79.470111	Airport
550	Downsview East	43.737473	-79.464763	Toronto Downsview Airport (YZD)	43.738883	-79.470111	Airport
639	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577	Warden Subway Station	43.711229	-79.279602	Metro Station
640	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577	Warden Station Bus Loop	43.711241	-79.279576	Bus Station
229	Berczy Park	43.644771	-79.373306	Hockey Hall Of Fame (Hockey Hall of Fame)	43.646974	-79.377323	Museum
884	Wexford, Maryvale	43.750072	-79.295849	Wexford Heights Plaza	43.746136	-79.293782	Shopping Mall
841	Davisville North	43.712751	-79.390197	Windowrama by Paul	43.712185	-79.395317	Construction & Landscaping
944	Parkdale, Roncesvalles	43.648960	-79.456325	A Good Read	43.649470	-79.450339	Bookstore
885	Wexford, Maryvale	43.750072	-79.295849	Scarborough Garage Door Repair	43.751288	-79.301508	Auto Garage
945	Parkdale, Roncesvalles	43.648960	-79.456325	Aris Grill	43.650091	-79.450396	Breakfast Spot
946	Parkdale, Roncesvalles	43.648960	-79.456325	Butler's Pantry	43.650087	-79.450458	Breakfast Spot
240	Berczy Park	43.644771	-79.373306	Eggspectation	43.646526	-79.375134	Breakfast Spot
835	Davisville North	43.712751	-79.390197	Homeway Restaurant & Brunch	43.712641	-79.391557	Breakfast Spot

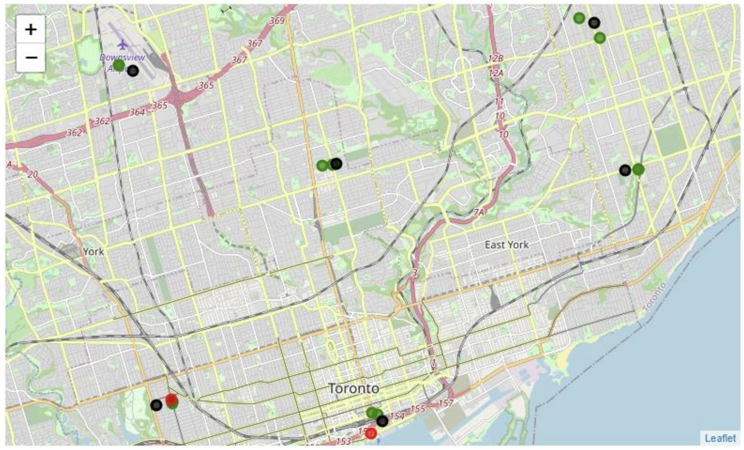
Bad Venues:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue Name	Venue Latitude	Venue Longitude	Venue Category
940	Parkdale, Roncesvalles	43.648960	-79.456325	Reunion Island Coffee Bar	43.650463	-79.450610	Coffee Shop
253	Berczy Park	43.644771	-79.373306	Mos Mos	43.641640	-79.377552	Coffee Shop

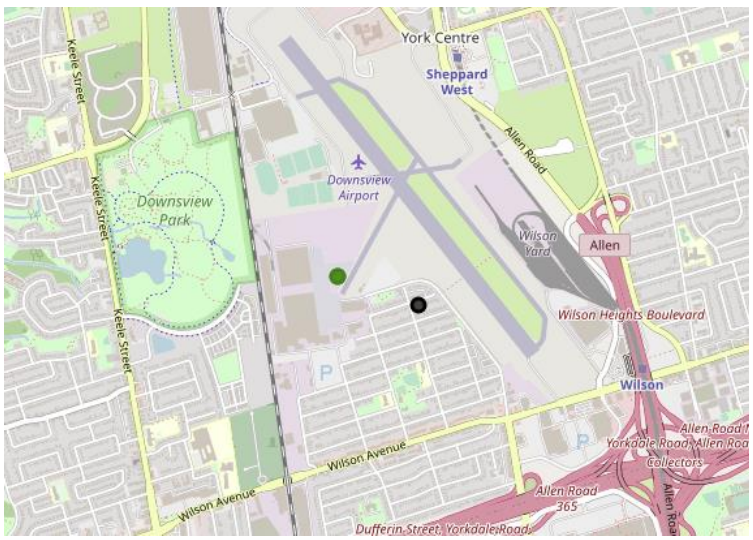
Now that we have these dataframes, we can now plot our 6 neighborhoods on a folium map along with every good venue and bad venue surrounding it!

IV. Results

After running our folium map, this is what we get.



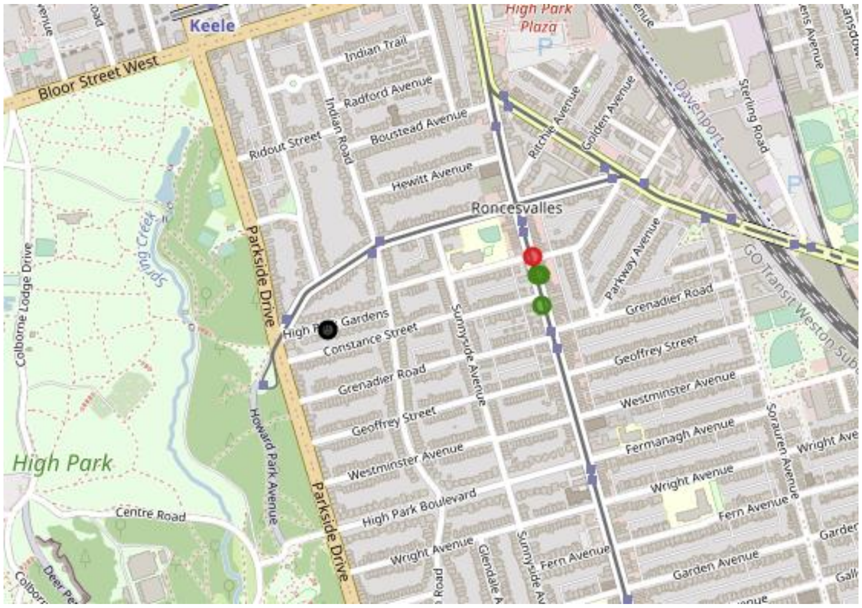
Downsview East:



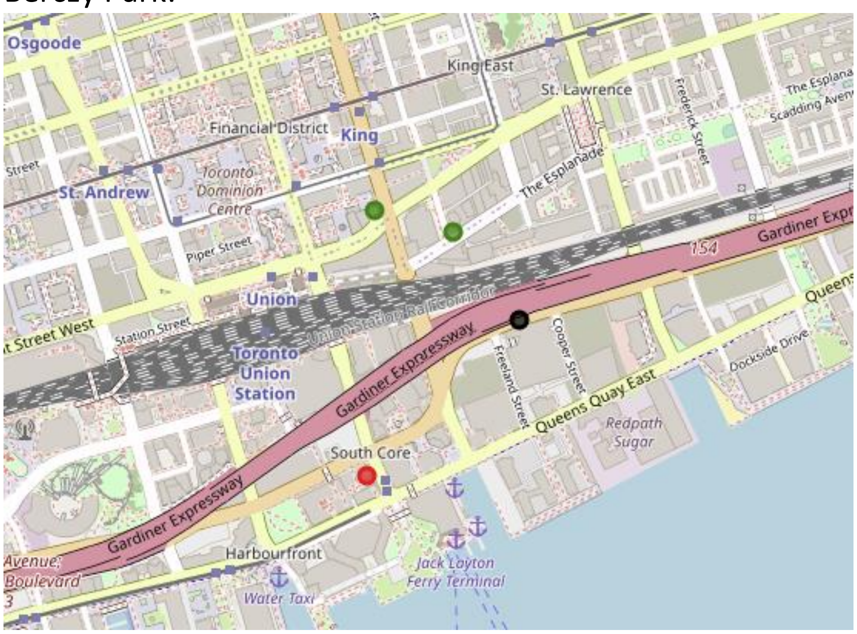
- After inspection, we can see there's only one venue that shows up, and after looking back at my dataframe, this makes sense. Somehow 3 datapoints had accidently got repeated in my Toronto_coffee dataframe:

550	Downsview East	43.737473	-79.464763	Toronto Downsview Airport (YZD)	43.738883	-79.470111	Airport
551	Downsview East	43.737473	-79.464763	Ttc Bus #120 - Plewes Rd	43.734898	-79.464221	Bus Stop
552	Downsview East	43.737473	-79.464763	Ancaster Park	43.734706	-79.464777	Park
550	Downsview East	43.737473	-79.464763	Toronto Downsview Airport (YZD)	43.738883	-79.470111	Airport
551	Downsview East	43.737473	-79.464763	Ttc Bus #120 - Plewes Rd	43.734898	-79.464221	Bus Stop
552	Downsview East	43.737473	-79.464763	Ancaster Park	43.734706	-79.464777	Park

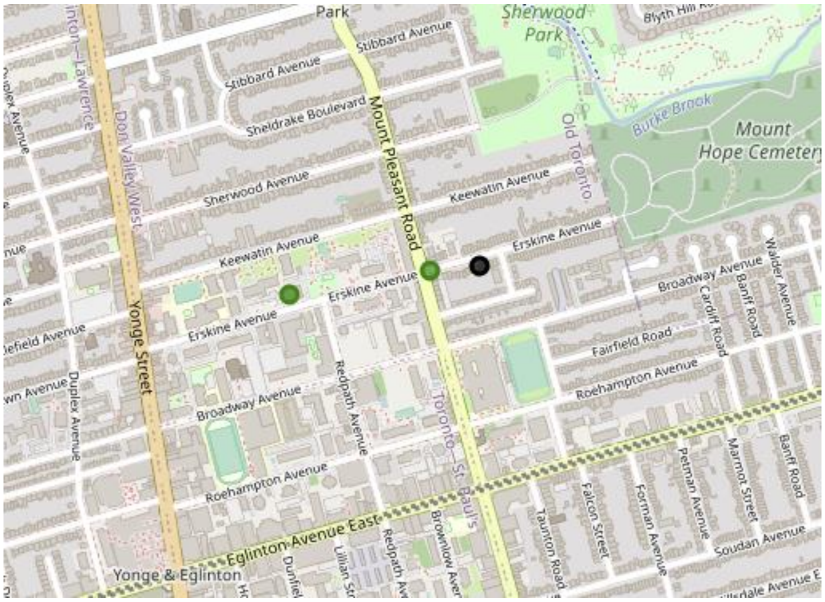
Parkdale, Roncesvalles:



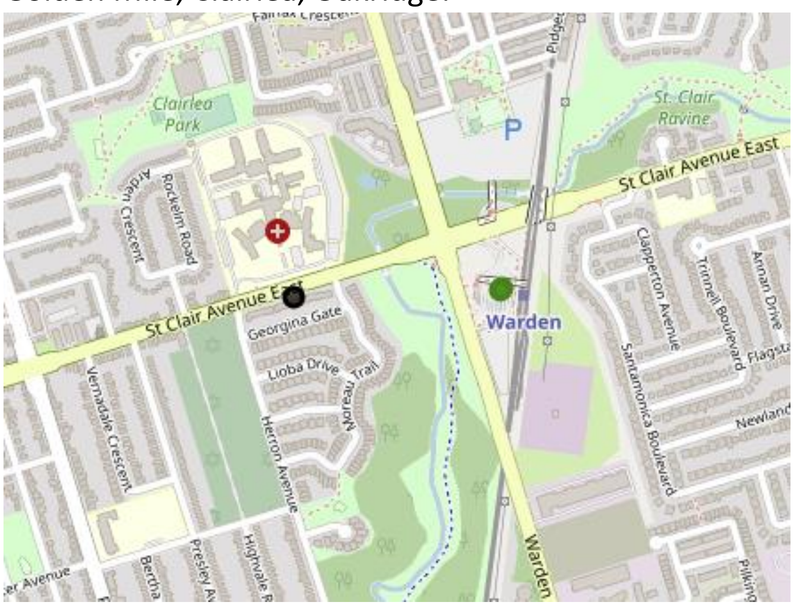
Berczy Park:



Davisville North:



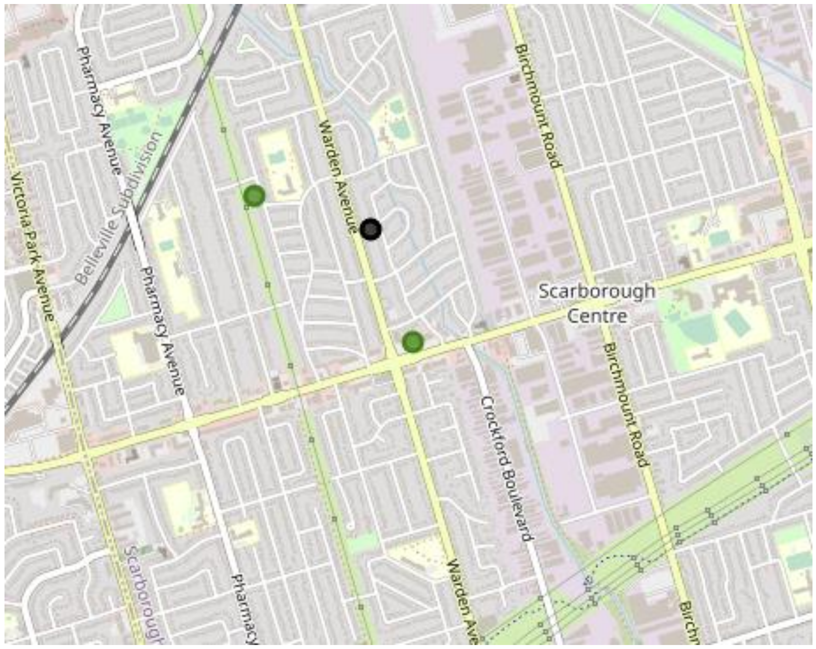
Golden Mile, Clairlea, Oakridge:



- This point looks like it has only 1 good venue just like Downsview East, however, we can see that this bus station is in the same exact location as a metro station, so it counts as 2 venues.

639	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577	Warden Subway Station	43.711229	-79.279602	Metro Station
640	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577	Warden Station Bus Loop	43.711241	-79.279576	Bus Station

Wexford, Maryvale:



V. Discussion

As we can see from our individual maps, we only have 2 neighborhoods with multiple surrounding good venues in distinct locations and no bad venues. We have 2 neighborhoods with a surrounding coffee shop and multiple good venues. We have Golden Mile, Clairlea, and Oakridge next to the bus station and metro station. Finally, we have the neighborhood next to the airport which wasn’t supposed to be included in this list in the first place. Since we have 2 neighborhoods with distinct coffee shops there already, we’ll go ahead and put them to the side. However, if the coffee shops surrounding those neighborhoods (Berczy Park and Parkdale, Roncesvalles) are not doing so well, then perhaps further research is necessary to get a good idea if those neighborhoods would be good to start up in.

With Davisville North being next to a breakfast spot and a construction & landscaping business, we can definitely add this neighborhood to a short list of good neighborhoods to consider. Wexford, Maryvale is another great option to consider with its proximity to a shopping mall and an auto garage. All of these businesses are those whose workers get up early or provide a place for people to gather at and attract customers (like the shopping mall). We can consider opening up in Golden Mile, Clairlea, and Oakridge. However, the bus station and metro station may not be considered as a strong enough market for opening a coffee shop in this neighborhood, but we can still have it open for consideration.

VI. Conclusion

We can see that in order to maximize our chances for success when opening a coffee shop, we have a handful of options to consider. Our first options we should be targeting are the neighborhoods of Davisville North and Wexford, Maryvale. After these options are considered, olden Mile, Clairlea, and Oakridge would be a good natural progression to make. With a metro station and bus station nearby, you'll get plenty of foot traffic in the surrounding area daily. When these options are considered, we can move on to the neighborhoods of Berczy Park and Parkdale, Roncesvalles. These may be a good consideration if the coffee shops in the area aren't favored for some reason. If the shops are merely unsuccessful, then these neighborhoods should probably just be ignored.

There are still neighborhoods that are unexplored that only host one good local venue in walking distance, and perhaps further research into the city of Toronto would reveal different ways of segmenting the city to find an even better means of finding a good neighborhood location. We could also explore possible datasets with income data in order to find out what areas of the city have the highest income and sort by that. The segmenting I've done is just one example of how we can explore the city and many other factors can be considered for future analysis.