

Capstone Project – The Battle of Neighborhoods

Introduction/Business Problem

The city of Toronto is the financial capital of Canada. It has diverse cultures and enormous potential for new businesses. As a global hub of commerce, a lot of people come to the city to look for new opportunities in the market every year. However, great opportunity also means cruel competition. To reduce the risk of failure and explore the benefit as much as we can, careful analysis needs to be done before investing in the real market.

Meanwhile, bubble tea, a popular Taiwanese tea-based drink has become famous throughout the world. Its perfect combination of tea, milk and sugar, together with its different toppings satisfies people's cravings from different continents. A growing number of bubble tea shops have popped up in Toronto in the last couple of years. More and more people enjoy this asian taste during their weekday breaks or weekend free time. We are interested in opening a new bubble tea shop in the city of Toronto and would like to enter the market with a smooth start. Therefore, it is important for us to know about the potential competitors around different boroughs and find the best location for our new shop.

In this project, we will investigate the distribution of bubble tea shops in Toronto and pick the best location to open a new bubble tea shop.

Data

The first dataset we are using is based on the Wikipedia page:

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M, which contains postal codes of different boroughs and neighborhoods of Toronto. We are going to extract the table and save the information into a pandas dataframe. The table in the webpage looks like below,

Postcode ↕	Borough ↕	Neighbourhood ↕
M1A	Not assigned	Not assigned
M2A	Not assigned	Not assigned
M3A	North York	Parkwoods
M4A	North York	Victoria Village
M5A	Downtown Toronto	Harbourfront
M5A	Downtown Toronto	Regent Park
M6A	North York	Lawrence Heights
M6A	North York	Lawrence Manor
M7A	Queen's Park	Not assigned
M8A	Not assigned	Not assigned
M9A	Etobicoke	Islington Avenue

The second dataset has the geographical coordinates of each postal code and we are using the dataset provided in the capstone project week 3: http://cocl.us/Geospatial_data. Just like what we did in project week 3, we combined the geographical data with the postal codes, boroughs and neighborhoods table. The snippet of the final table is shown below,

	Postcode	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	Harbourfront,Regent Park	43.654260	-79.360636
3	M6A	North York	Lawrence Heights,Lawrence Manor	43.718518	-79.464763
4	M7A	Queen's Park	Queen's Park	43.662301	-79.389494

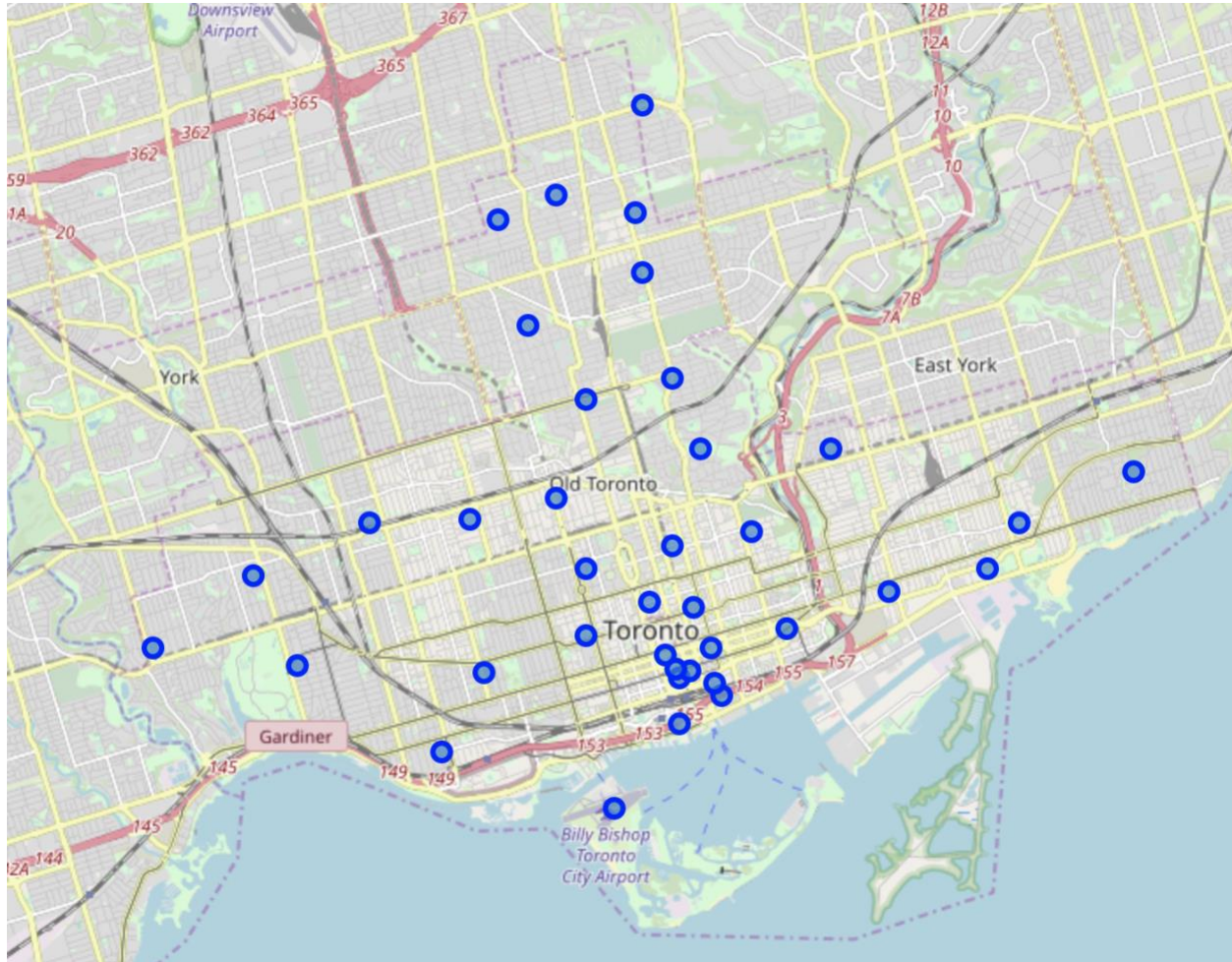
In addition to the above two datasets, we are utilizing the FourSquare api to retrieve the bubble tea shops information around each interested neighborhood. Based on all these information, we will pick the best location for our new bubble tea shop.

Methodology

In this project, since the dataset contains data of the whole Great Toronto Area, which consists of multiple surrounding areas and cities around central Toronto. We choose to focus on the

center of Toronto. Therefore, we filter the dataset to only contain boroughs: Downtown Toronto, East Toronto, West Toronto and Central Toronto.

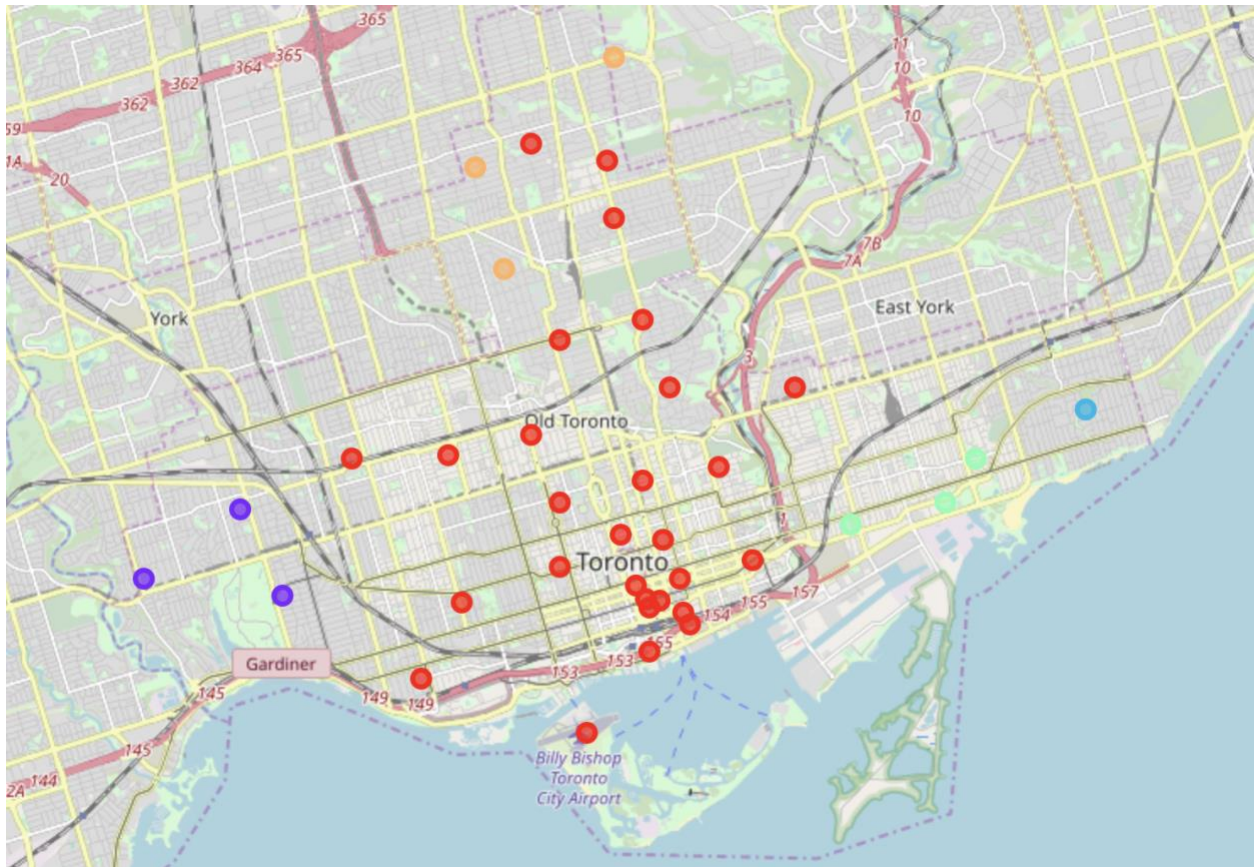
We then visualize all the neighborhoods on the map using FourSquare api.



Next, we use the explore function of FourSquare api to find bubble tea shops around each neighborhood. We set the radius to be 3000 meters and the limit to be 100 since the further the distance of the store, the less possible people are going to buy a cup of bubble tea. After retrieving all the bubble tea shops, we follow the same steps of capstone project week 3 to encode the neighborhood,

	Neighborhood	Asian Restaurant	Bakery	Bubble Tea Shop	Café	Chinese Restaurant	Dessert Shop	Frozen Yogurt Shop	Juice Bar	Tea Room	Vietnamese Restaurant
0	Adelaide,King,Richmond	0.014493	0.014493	0.826087	0.028986	0.014493	0.014493	0.000000	0.028986	0.028986	0.028986
1	Berczy Park	0.015873	0.031746	0.777778	0.031746	0.031746	0.015873	0.000000	0.031746	0.031746	0.031746
2	Brockton,Exhibition Place,Parkdale Village	0.000000	0.000000	0.818182	0.000000	0.000000	0.000000	0.000000	0.000000	0.090909	0.090909
3	Business Reply Mail Processing Centre 969 Eastern	0.200000	0.200000	0.400000	0.000000	0.200000	0.000000	0.000000	0.000000	0.000000	0.000000
4	CN Tower,Bathurst Quay,Island airport,Harbourf...	0.034483	0.000000	0.724138	0.034483	0.000000	0.000000	0.000000	0.068966	0.068966	0.068966
5	Cabbagetown,St. James Town	0.018182	0.036364	0.781818	0.036364	0.036364	0.018182	0.000000	0.036364	0.018182	0.018182
6	Central Bay Street	0.014085	0.028169	0.802817	0.028169	0.028169	0.014085	0.000000	0.028169	0.028169	0.028169
7	Chinatown,Grange Park,Kensington Market	0.014493	0.014493	0.826087	0.028986	0.014493	0.014493	0.000000	0.028986	0.028986	0.028986
8	Christie	0.033333	0.000000	0.800000	0.033333	0.033333	0.000000	0.000000	0.000000	0.066667	0.033333
9	Church and Wellesley	0.014085	0.028169	0.802817	0.028169	0.028169	0.014085	0.000000	0.028169	0.028169	0.028169
10	Commerce Court,Victoria Hotel	0.015385	0.030769	0.784615	0.030769	0.030769	0.015385	0.000000	0.030769	0.030769	0.030769
11	Davisville	0.000000	0.000000	0.833333	0.000000	0.000000	0.000000	0.000000	0.000000	0.166667	0.000000
12	Davisville North	0.000000	0.000000	0.833333	0.000000	0.000000	0.000000	0.000000	0.000000	0.166667	0.000000

We then use k-means to cluster the neighborhoods to 5 clusters and visualize the 5 clusters using the Folium library,



To reduce the risk of failure of opening a new bubble tea shop, we choose to open the shop in the neighborhood with fewer competitors. Meanwhile, we would like to know the other types of venues around these neighborhoods, which can help us better understand the preferences of the people nearby. For example, if people nearby enjoy more casual life style or enjoy coffee more often, it is more possible that they will buy a cup of bubble tea. Also, since bubble tea originates from Taiwan and it is very popular now in Asia, it makes sense that Asian people in

Toronto are more likely to buy bubble tea. We pick the best location for the new bubble tea shop based on all the above information.

Results

The 5 clusters we get,

Cluster 1:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Harbourfront,Regent Park	Bubble Tea Shop	Tea Room	Juice Bar	Café	Bakery	Vietnamese Restaurant	Dessert Shop	Chinese Restaurant	Frozen Yogurt Shop	Asian Restaurant
1	Ryerson,Garden District	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Chinese Restaurant	Café	Bakery	Dessert Shop	Asian Restaurant	Frozen Yogurt Shop
2	St. James Town	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Chinese Restaurant	Café	Bakery	Dessert Shop	Asian Restaurant	Frozen Yogurt Shop
4	Berczy Park	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Chinese Restaurant	Café	Bakery	Dessert Shop	Asian Restaurant	Frozen Yogurt Shop
5	Central Bay Street	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Chinese Restaurant	Café	Bakery	Dessert Shop	Asian Restaurant	Frozen Yogurt Shop
6	Christie	Bubble Tea Shop	Tea Room	Vietnamese Restaurant	Chinese Restaurant	Café	Asian Restaurant	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Bakery
7	Adelaide,King,Richmond	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Café	Dessert Shop	Chinese Restaurant	Bakery	Asian Restaurant	Frozen Yogurt Shop
8	Dovercourt Village,Dufferin	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant
9	Harbourfront East,Toronto Islands,Union Station	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Café	Dessert Shop	Chinese Restaurant	Bakery	Asian Restaurant	Frozen Yogurt Shop
10	Little Portugal,Trinity	Bubble Tea Shop	Vietnamese Restaurant	Tea Room	Juice Bar	Chinese Restaurant	Café	Asian Restaurant	Frozen Yogurt Shop	Dessert Shop	Bakery

(Note: the first group is relatively large, we show the snippet of this cluster)

Cluster 2:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
21	High Park,The Junction South	Bubble Tea Shop	Tea Room	Frozen Yogurt Shop	Vietnamese Restaurant	Juice Bar	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant
24	Parkdale,Roncesvalles	Bubble Tea Shop	Tea Room	Frozen Yogurt Shop	Vietnamese Restaurant	Juice Bar	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant
27	Runnymede,Swansea	Tea Room	Frozen Yogurt Shop	Bubble Tea Shop	Vietnamese Restaurant	Juice Bar	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant

Cluster 3:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	The Beaches	Bubble Tea Shop	Asian Restaurant	Vietnamese Restaurant	Tea Room	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Chinese Restaurant	Café	Bakery

Cluster 4:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
14	The Beaches West,India Bazaar	Chinese Restaurant	Bubble Tea Shop	Asian Restaurant	Vietnamese Restaurant	Tea Room	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Café	Bakery
16	Studio District	Bubble Tea Shop	Chinese Restaurant	Bakery	Vietnamese Restaurant	Tea Room	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Café	Asian Restaurant
37	Business Reply Mail Processing Centre 969 Eastern	Bubble Tea Shop	Chinese Restaurant	Bakery	Asian Restaurant	Vietnamese Restaurant	Tea Room	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Café

Cluster 5:

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
17	Lawrence Park	Bubble Tea Shop	Tea Room	Vietnamese Restaurant	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant
18	Roselawn	Bubble Tea Shop	Tea Room	Vietnamese Restaurant	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant
20	Forest Hill North,Forest Hill West	Tea Room	Bubble Tea Shop	Vietnamese Restaurant	Juice Bar	Frozen Yogurt Shop	Dessert Shop	Chinese Restaurant	Café	Bakery	Asian Restaurant

As we can see, cluster 2, 3, 4 and 5 neighborhoods have relatively fewer bubble tea shops. Therefore, there are fewer competitors nearby. We then further investigate the other types of venues of these neighborhoods to further understand the preferences of people visiting that area. We get the top 10 common venues of cluster 2, 3, 4 and 5 neighborhoods, which are shown below,

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Business Reply Mail Processing Centre 969 Eastern	Grocery Store	Café	Park	Convenience Store	Coffee Shop	Nightclub	Athletics & Sports	Baby Store	Diner	Restaurant
1	Forest Hill North,Forest Hill West	Coffee Shop	Hotel	Aquarium	Café	Italian Restaurant	Bakery	Scenic Lookout	Sporting Goods Shop	Pizza Place	Brewery
2	High Park,The Junction South	Coffee Shop	Bakery	Café	Park	Mexican Restaurant	Pub	Breakfast Spot	Farmers Market	Event Space	Electronics Store
3	Lawrence Park	Coffee Shop	Café	Steakhouse	Thai Restaurant	Bar	American Restaurant	Restaurant	Breakfast Spot	Asian Restaurant	Hotel
4	Parkdale,Roncesvalles	Coffee Shop	Clothing Store	Cosmetics Shop	Café	Middle Eastern Restaurant	Bubble Tea Shop	Ramen Restaurant	Japanese Restaurant	Italian Restaurant	Ice Cream Shop
5	Roselawn	Pharmacy	Bakery	Supermarket	Music Venue	Middle Eastern Restaurant	Café	Brewery	Bar	Bank	Park
6	Runnymede,Swansea	Coffee Shop	Hotel	Café	Restaurant	Italian Restaurant	Clothing Store	Cocktail Bar	Gastropub	Breakfast Spot	Cosmetics Shop
7	Studio District	Coffee Shop	Café	Ice Cream Shop	Italian Restaurant	Sandwich Place	Burger Joint	Indian Restaurant	Spa	Salad Place	Middle Eastern Restaurant
8	The Beaches	Health Food Store	Other Great Outdoors	Trail	Pub	Deli / Bodega	Eastern European Restaurant	Donut Shop	Discount Store	Diner	Dessert Shop
9	The Beaches West,India Bazaar	Coffee Shop	Cocktail Bar	Beer Bar	Seafood Restaurant	Cheese Shop	Café	Farmers Market	Bakery	Steakhouse	Clothing Store

Based on the above results, we choose neighborhoods of cluster 2 and 5 as the better location to open a new bubble tea shop.

Discussion

In this section, we discuss the observations we noted and recommendations we make based on the results.

Cluster 2, 3, 4, 5 neighbourhoods have relatively fewer bubble tea shops. There are fewer competitors nearby. We exclude cluster 1 neighborhoods since there are already a lot of bubble tea shops and the chance of surviving through the competition is relatively low.

Cluster 2 neighbourhoods:

1. High Park, The Junction South
2. Parkdale, Roncesvalles
3. Runnymede, Swansea

There are coffee shops, hotels, bakeries nearby, which indicate people visiting these areas usually buy drinks. These places are good choice to open bubble tea shops.

Cluster 3 neighbourhoods:

1. The Beaches

There is health food store, outdoors, trails, pubs nearby. People visiting this place tend to be more active and pay more attention on their health. Since bubble tea is not considered as a healthy drink. People may not buy bubble tea in this place.

Cluster 4 neighbourhoods:

1. The Beaches West, India Bazaar
2. Studio District
3. Business Reply Mail Processing Centre 969 Eastern

There are coffee shops, cafes in these locations which means people enjoy coffee at these places. However, there are also beer bars, nightclubs, western style food restaurants at these places, which could be an indicator that people visiting these places may enjoy adult style drinks more often and they may prefer western style food than Taiwanese bubble tea.

Cluster 5 neighbourhoods:

1. Forest Hill North, Forest Hill West
2. Lawrence Park
3. Roselawn

There are coffee shops, hotels, supermarkets and Asian style restaurants in these places. People visiting these places may spend more time and enjoy a cup of bubble tea.

In summary, cluster 2 and cluster 5 neighbourhoods are better choices to open a new bubble tea shop. It's better if we can obtain more demographics data and crime data of these locations since people from Asia tends to buy bubble tea more often than western people and safety of these locations is also an important factor.

Conclusion

In this project, we investigate the borough Downtown Toronto, East Toronto, West Toronto and Central Toronto to find a location for our new bubble tea shop. With the explore function of FourSquare api, we are able to retrieve the existing bubble tea shops information and based on that information, we have found better neighbourhoods for the new shop to survive.

With the limitation of demographics data and crime data, we predict the optimum location with only geographical data and the venues data. It is possible to further improve the results if we can get more data about the areas we are interested in.