Capstone Project: The Battle of the Neighborhoods

YogaBear: mindful beginning, one day at a time

The Background

YogaBear is a Chinese yoga company that has been very successful in mainland China. Over the past 10 years, the company have grown from one small yoga studio in Shanghai to over 200 studios spread across China. On the back of this domestic success, they begin to set their eyes on overseas market.

Their initial research has brought them to two of the biggest cities in North America, namely New York and Toronto.

(1) New York, the most populous city in the United States (US) with an estimated population of 8,398,748 as of 2018, is the most densely populated major city in the US. A global power city, New York has been described as the cultural, financial, and media capital of the world, and exerts a significant impact upon education, tourism, art, fashion, and sports.

(2) Toronto, the provincial capital of Ontario, is the most populous city in Canada with a population of 2,731,571 as of 2016. Like New York, it is also an international center of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world.

Because both are major metropolitans with extremely high cost of doing business, any new business venture in these cities deserve a proper and careful analysis. In this project, we are going to do an analysis to assess the prospect of opening a yoga studio there. The insights derived from our analysis will help YogaBear decide which city will be the right market to go into as their first major overseas investment. *So, without further ado, let's get into it.*

The Critical Success Factors

The practice of yoga has experienced an exponential growth over the past couple of decades and it continues to attract more and more people. Naturally, as more and more people practice yoga, more and more yoga studios are opening. However, opening a yoga studio is not without its challenges. Good location, good surrounding venues, the competition and the ability to maintain and expand student base are all very important factors.

(1) The competition - who's around the corner?

It is common sense to check the competition before going into any market. The point being that it is a good practice to scope out areas where the competition is already established. What is the point of opening a new studio if the market is already oversaturated by the existing players in the area. Bottom line is: research your market, avoid competition if possible, then open with a big bang to quickly establish yourself as the main player in the area.

(2) Good location - where should we go?

Whenever you want to open a new shop (in this case, a yoga studio), it makes sense to pick a good location. Based on their experience in China, it is always good practice to open the first studio in the CBD area. If this becomes successful, it will usually become easier for them to open new studios in other areas in the city.

(3) Good surrounding venues - who is around us?

In China, yoga studios go hand in hand with parks and coffee shops. Having a park nearby is apparently an important success factor for a yoga studio. People often like to do yoga outdoor, and having a public park nearby means they can do so conveniently. Surprisingly, it is also important to have coffee shops or cafe nearby. Usually, people who like to drink coffee tend to live a high pace life and this sometimes can be a very stressful life. YogaBear found that these people tend to benefit the most and are more likely to sign up to yoga than others.

The Data

In this project, we will be using the New York dataset and the Toronto data set from the Wikipedia page to obtain neighborhoods information about New York and Toronto. For Toronto, we will also use the latitude and longitude data in the Geospatial_Coordinates.csv file to complement our data. For this project, we will focus our analysis on the CBD areas in Manhattan and Toronto.

Then, we will use the foursquare location data to locate nearby venues information. We will check the most common venues, the least common venues, how many yoga studios already exist in the areas, how many parks, coffee shops and cafes in those areas. Based on our analysis we will then decide which city will be the right market to go into as YogaBear's first major investment overseas. We will also use charts and maps to plot our analysis and proposed locations.

Methodology

Manhattan, New York

From the data that we obtained above, we pre-processed the data into a pandas DataFrame that includes neighborhoods in Manhattan, including their latitudes and longitudes. We focus our analysis on the Manhattan neighborhood as we would like to search for a place for our yoga studio in the CBD area.

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Table 1. The first 5 neighborhoods in Manhattan (In total, there are 40 neighborhoods in Manhattan that we analyze for this project)

After deciding on which neighborhoods we would like to analyze, we used the foursquare location data to locate nearby venues information for those neighborhoods.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop
4	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop

Table 2. The first 5 venues in Marble Hill, Manhattan

We then perform further analysis on the data and check the top 10 most common and least common venues in Manhattan.

	Neighborhood	1st Least Common Venue	2nd Least Common Venue	3rd Least Common Venue	4th Least Common Venue	5th Least Common Venue	6th Least Common Venue	7th Least Common Venue	8th Least Common Venue	9th Least Common Venue	10th Least Common Venue
0	Battery Park City	Accessories Store	Optical Shop	Opera House	Office	Noodle House	Non-Profit	Nightclub	Newsstand	New American Restaurant	Nail Salon
1	Carnegie Hill	Accessories Store	Opera House	Office	Noodle House	Non-Profit	Nightclub	Newsstand	Nail Salon	Music Venue	Music School
2	Central Harlem	Accessories Store	Outdoors & Recreation	Outdoor Sculpture	Other Great Outdoors	Organic Grocery	Optical Shop	Opera House	Office	Paella Restaurant	Noodle House
3	Chelsea	Accessories Store	Opera House	Non-Profit	Newsstand	Nail Salon	Music Venue	Music School	Optical Shop	Museum	Moroccan Restaurant
4	Chinatown	Accessories Store	Non-Profit	Nightclub	Newsstand	Nail Salon	Music School	Movie Theater	Office	Moroccan Restaurant	Molecular Gastronomy Restaurant

	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	Battery Park City	Park	Coffee Shop	Hotel	Gym	Clothing Store	Memorial Site	BBQ Joint	Women's Store	Gourmet Shop	Italian Restaurant
1	Carnegie Hill	Coffee Shop	Pizza Place	Café	Yoga Studio	French Restaurant	Japanese Restaurant	Italian Restaurant	Bakery	Gym	Bookstore
2	Central Harlem	Chinese Restaurant	Bar	Seafood Restaurant	African Restaurant	American Restaurant	Cosmetics Shop	Art Gallery	French Restaurant	Spa	Gym / Fitness Center
3	Chelsea	Coffee Shop	Italian Restaurant	Bakery	Ice Cream Shop	American Restaurant	Art Gallery	Nightclub	Hotel	Theater	Market
4	Chinatown	Chinese Restaurant	Cocktail Bar	American Restaurant	Vietnamese Restaurant	Salon / Barbershop	Spa	Bakery	Optical Shop	Hotpot Restaurant	Ice Cream Shop
5	Civic Center	Gym / Fitness Center	Coffee Shop	Hotel	French Restaurant	Spa	Italian Restaurant	Bakery	Sandwich Place	Cocktail Bar	Yoga Studio

Table 3. The 10 most common venues and the 10 least common venues in Manhattan

From our analysis, we found that the least common venues in Manhattan are Accessories Store and Indie Theater. We also found that there are 35 yoga studios already exist in Manhattan alone, so probably not a very place to open up another one from the competition perspective.

We then perform the k-means clustering algorithm to cluster the Manhattan neighborhoods into 5 clusters. We want to see how the neighborhoods in Manhattan are clustered together and what the most common venues are available in the most clustered neighborhoods.
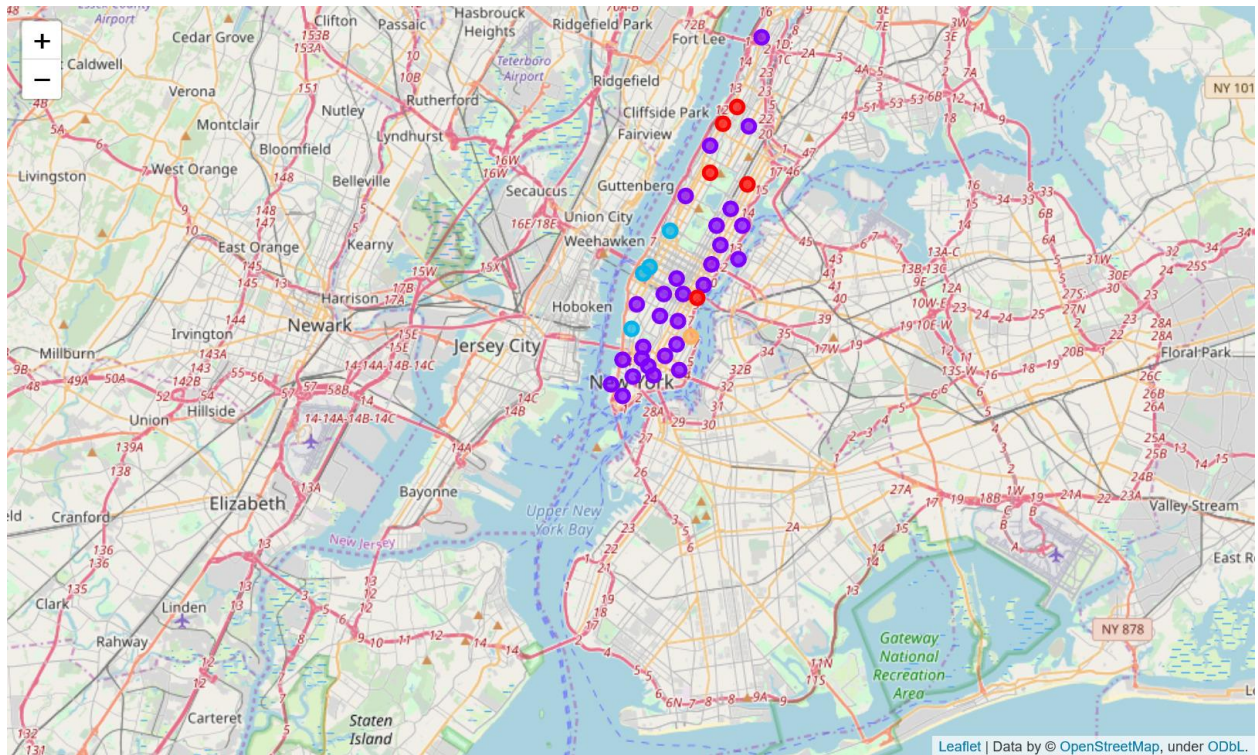
Map 1. K-means clustering performed on Manhattan neighborhoods

Our finding shows that the biggest cluster is around Chinatown, Upper West Side and Lower East Side of Manhattan. We can see that there is a lot of gym/fitness center in these areas and a number of yoga studios, making Manhattan not the most attractive place to open up another yoga studio. But don't get discouraged. Let's turn our attention to our other city, Toronto.

Toronto, Toronto
Similarly, we pre-processed the Toronto data into a pandas DataFrame that includes neighborhoods in Toronto (with name Toronto included), including their latitudes and longitudes.

	Postcode	Borough	Neighborhood	Latitude	Longitude
0	M5A	Downtown Toronto	Harbourfront	43.654260	-79.360636
1	M7A	Downtown Toronto	Queen's Park	43.662301	-79.389494
2	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
4	M4E	East Toronto	The Beaches	43.676357	-79.293031

Table 4. The first 5 neighborhoods in Toronto (In total, there are 39 neighborhoods that we analyze for this project)

After deciding on which neighborhoods we would like to analyze, we then used the foursquare location data to locate nearby venues information for those neighborhoods and determine the top 10 most common and least common venues in those neighborhoods.

	Neighborhood	1st Least Common Venue	2nd Least Common Venue	3rd Least Common Venue	4th Least Common Venue	5th Least Common Venue	6th Least Common Venue	7th Least Common Venue	8th Least Common Venue	9th Least Common Venue	10th Least Common Venue
0	Adelaide, King, Richmond	Yoga Studio	Gift Shop	Spa	Gourmet Shop	Southern / Soul Food Restaurant	Grocery Store	Wings Joint	Snack Place	Harbor / Marina	Health & Beauty Service
1	Berczy Park	Yoga Studio	Latin American Restaurant	Light Rail Station	Lingerie Store	Lounge	Mac & Cheese Joint	Market	Massage Studio	Mediterranean Restaurant	Men's Store
2	Brockton, Exhibition Place, Parkdale Village	Yoga Studio	Liquor Store	Lounge	Mac & Cheese Joint	Market	Massage Studio	Mediterranean Restaurant	Men's Store	Mexican Restaurant	Middle Eastern Restaurant
3	Business Reply Mail Processing Centre 969 Eastern	Gym	Lingerie Store	Liquor Store	Lounge	Mac & Cheese Joint	Market	Massage Studio	Mediterranean Restaurant	Men's Store	Mexican Restaurant

	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	Adelaide, King, Richmond	Coffee Shop	Café	Bar	Thai Restaurant	Steakhouse	Bakery	Cosmetics Shop	Sushi Restaurant	Restaurant	Burger Joint
1	Berczy Park	Coffee Shop	Seafood Restaurant	Cheese Shop	Steakhouse	Bakery	Farmers Market	Café	Beer Bar	Cocktail Bar	Department Store
2	Brockton, Exhibition Place, Parkdale Village	Café	Breakfast Spot	Coffee Shop	Bakery	Climbing Gym	Burrito Place	Stadium	Italian Restaurant	Restaurant	Intersection
3	Business Reply Mail Processing Centre 969 Eastern	Yoga Studio	Auto Workshop	Skate Park	Smoke Shop	Spa	Burrito Place	Farmers Market	Fast Food Restaurant	Restaurant	Recording Studio
4	CN Tower, Bathurst Quay, Island airport, Harbo	Airport Service	Airport Lounge	Airport Terminal	Boutique	Harbor / Marina	Rental Car Location	Coffee Shop	Plane	Boat or Ferry	Bar

Table 5. The 10 most common venues and the 10 least common venues in Toronto

Our analysis shows that Yoga Studio and Gym are the two least common venues in Toronto. There is only 11 yoga studios and 25 gyms in Toronto (as opposed to 35 yoga studios and 58 gyms in Manhattan). This looks more promising compared to Manhattan.

In fact, our k-means clustering show more spread out clusters in Toronto neighborhoods, with the biggest cluster centered around Downtown and West Toronto. We can see that whilst there are a lot of parks and coffee shops/cafes in these areas, there is not as many gyms or yoga studios. So, our comparison of the two cities, show that Toronto is the more appropriate city for a new YogaBear studio.

The next step is to decide on which location in Toronto we would like to open it. We sorted the neighborhoods in Toronto based on the number of yoga studio in each neighborhood and then we picked the neighborhoods that do not have any yoga studios. We then check how many parks, cafes and coffee shops these places have and recommend the 3 most appropriate locations for our yoga studio.
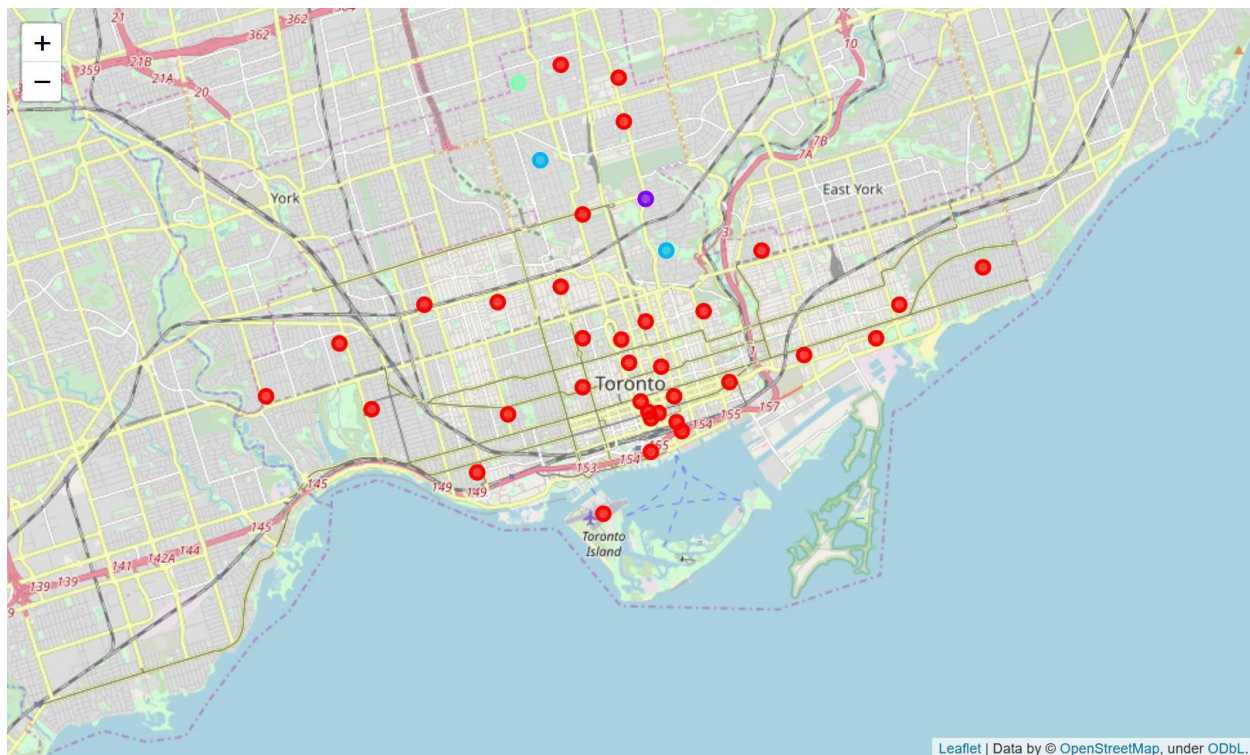
Map 2. K-means clustering performed on Toronto neighborhoods

	Neighborhood	Yoga Studio	Park	Coffee Shop	Café
20	Harbourfront East, Toronto Islands, Union Station	0	2	12	4
33	Stn A PO Boxes 25 The Esplanade	0	2	11	4
32	St. James Town	0	2	6	6
7	Chinatown, Grange Park, Kensington Market	0	2	4	5
8	Christie	0	2	1	3
37	The Beaches West, India Bazaar	0	2	1	0
28	Rosedale	0	2	0	0
10	Commerce Court, Victoria Hotel	0	1	11	7
1	Berczy Park	0	1	4	2
5	Cabbagetown, St. James Town	0	1	3	2

Table 6. Top 10 neighborhoods with no yoga studios and surrounded with parks, cafes and coffee shops.

The above table shows top 10 neighborhoods that do not have yoga studios, but surrounded by parks and cafes/coffee shops. As mentioned by YogaBear earlier, they determined that they are two of the most critical success factors when opening new yoga studios in China and they hope to emulate their success here in Toronto by applying the same formula.

Observations, Recommendations and Result

Based on our analysis above, we proposed the following neighborhoods for our new yoga studio:

	Postcode	Borough	Neighborhood	Latitude	Longitude	Yoga Studio	Park	Coffee Shop	Café
0	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418	0	2	6	6
1	M5J	Downtown Toronto	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752	0	2	12	4
2	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846	0	2	11	4

These three neighborhoods in Downtown Toronto do not have any yoga studios open yet and they are surrounded by the parks and cafes/coffee shops. Based on their experience, YogaBear determined that the above are critical success factors for opening new yoga studios (as shown in the barplots below).
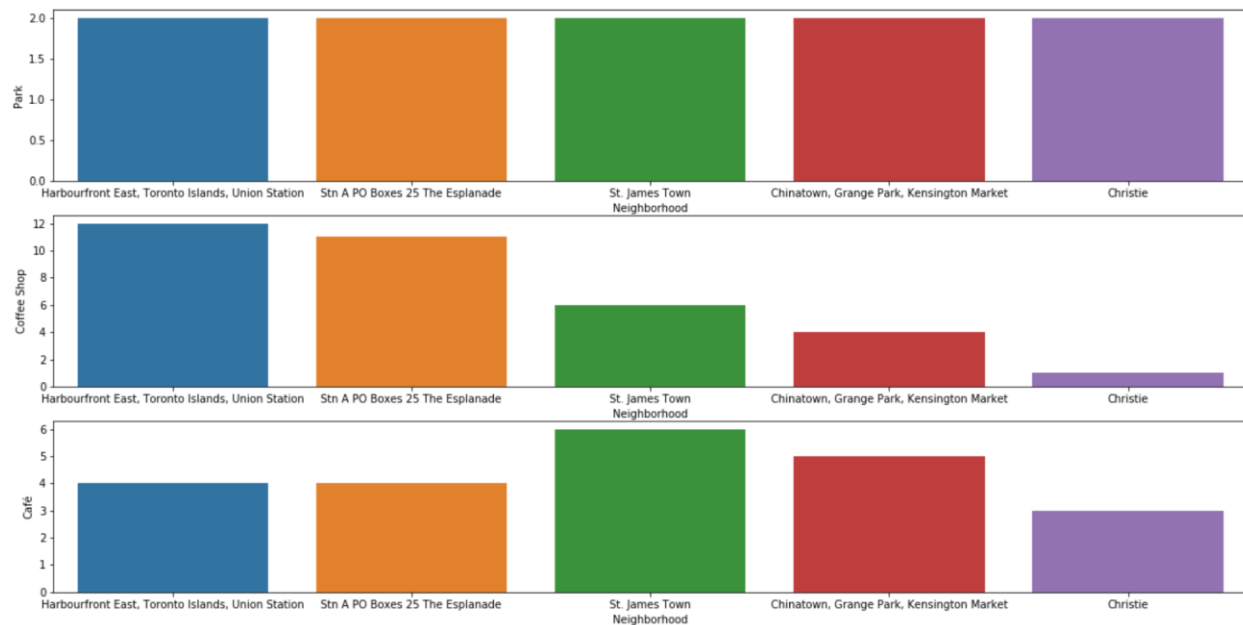

Figure 1. Barplots of top 5 neighborhoods in Toronto that have no yoga studios, showing the number of parks, cafes and coffee shops in those neighborhoods

Conclusion

We have performed our data pre-processing in this exercise to ensure we include the appropriate data in our analysis. We have also performed a number of exploratory data analysis and use foursquare data to check the locations and surrounding venues for our recommendation. We have performed k-means clustering to check the clusters in the neighborhoods and also some data analysis to check our competitions and supporting surrounding venues.

Our analysis shows that Toronto is the better city to open the new studio. It is the most populous city in Canada, does not have many yoga studios or gyms in the area, and the locations are perfect for opening a new yoga studio as they are surrounded by parks and cafes.

Finally, let's plot our proposed locations on the map to see exactly where they are. For comparison, we will also show the locations of the neighborhoods that are not recommended for our new yoga studio (the bottom 8 in the above table).
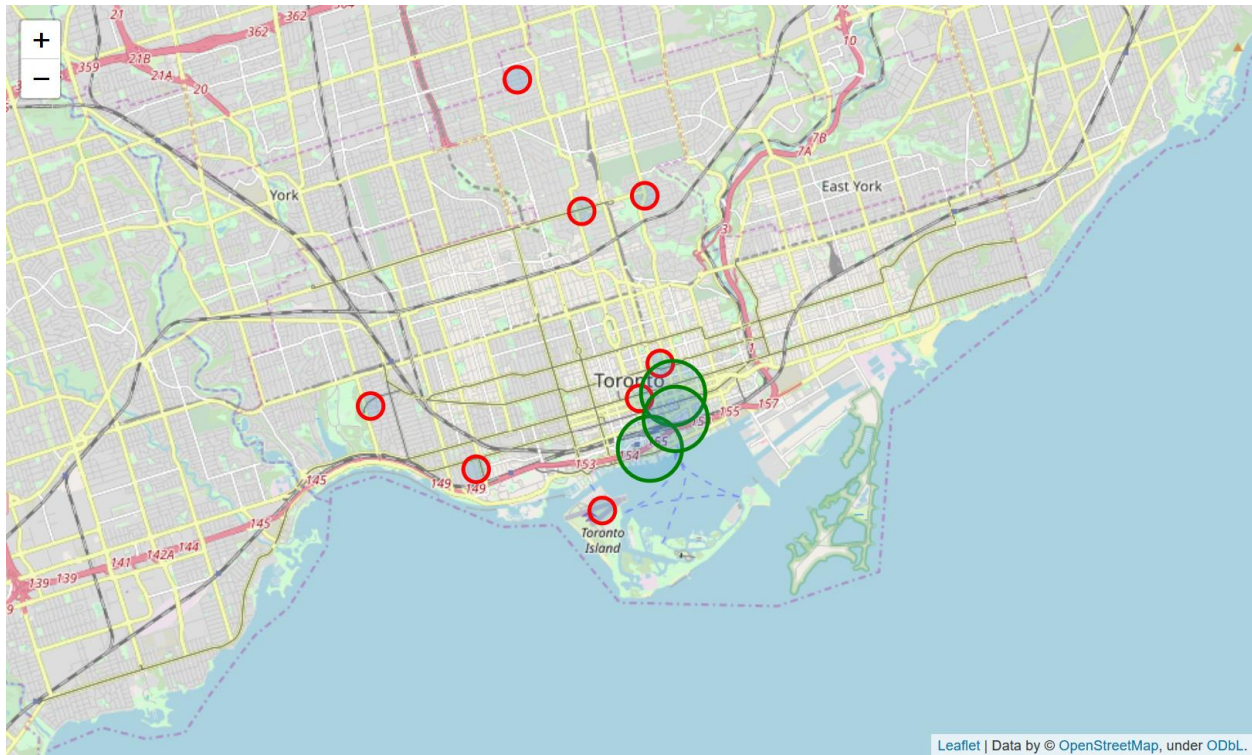
Map 3. Map of our proposed locations. The green markers above are the recommended places for our new yoga studio. YogaBear can choose one of them to begin their first overseas investment. The smaller red markers show the places we do not recommend, at least at this stage.

This concludes our report. Thank you for reading this report.

The End