# BATTLE OF THE NEIGHBORHOODS

LOC QUACH

INTRODUCTION

- One of the most common businesses is that of a restaurant. A major reason for this is because it is a business that have a relatively low barrier of entry. For this project, I am assuming that a client is interested in finding a good Toronto neighborhood to open a Japanese restaurant in.

- While it is not difficult to set up a restaurant, it is not easy to successfully run one. The client can serve fantastic Japanese food and provide unparalleled service but the restaurant may still fail if it is in a poor location. For instance, the client may face an uphill battle if the restaurant in a neighborhood with many similar restaurants.

- Because a good location is vital to the success of the client's venture, I will analyze Toronto's geographic data to find an ideal neighborhood for the hypothetical client's new Japanese restaurant.

DATA

- For this project, I will aggregate and analyze data from three sources: Toronto neighborhood data from Wikipedia, latitude and longitude data from Geocoder, and venue data from Foursquare.

- Firstly, I imported postal code, borough, and neighborhood data from Wikipedia. I then merge the resulting dataframe with latitude and longitude data from Geocoder. Here is a map of Toronto with all its neighborhoods.
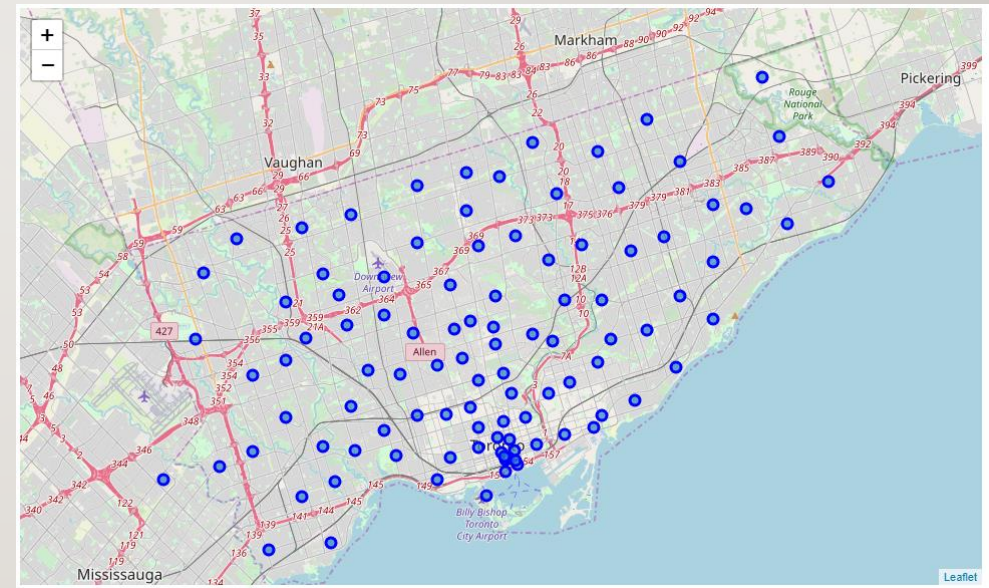
DATA

	Postalcode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848

Here are the first ten observations of the resulting geographic data.

DATA

- I next acquire venue data from Foursquare. I proceed to merge the venue data with the latitude and longitude data. Here are the first few observations of the new geographic dataframe.

	name	categories	lat	lng
0	Toronto Pan Am Sports Centre	Athletics & Sports	43.790623	-79.193869
1	African Rainforest Pavilion	Zoo Exhibit	43.817725	-79.183433
2	Toronto Zoo	Zoo	43.820582	-79.181551
3	Polar Bear Exhibit	Zoo	43.823372	-79.185145
4	Australasia Pavilion	Zoo Exhibit	43.822563	-79.183286

METHODOLOGY

- Restaurants are some of the most common venues Toronto. This suggests that restaurants do well in Toronto overall. Particularly interesting for the client is the fact that many neighborhoods have an abundance of Asian restaurants, indicating that there is demand for Asian food. But it also means that there will be many entrenched competitors for the client. To solve this conundrum, I will use the k-means clustering algorithm to identify clusters where Asian food is in particularly high demand. To avoid the market oversaturation challenge, I will only consider neighborhoods within these clusters where there are few Asian restaurants.

RESULTS

- The k-means algorithm is an unsupervised algorithm, meaning that the modeler cannot define the criteria for the algorithm. The modeler can simply interpret the clusters that the algorithm generates.

- I will present two clusters where the characteristics are easy to distinguish.

RESULTS

	Neighborhood	Cluster Labels	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
23	York Mills West	2.0	Park	Bank	Convenience Store	Women's Store	Eastern European Restaurant	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant
25	Parkwoods	2.0	Park	Food & Drink Shop	Women's Store	Dumpling Restaurant	Discount Store	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Electronics Store
40	East Toronto	2.0	Park	Metro Station	Convenience Store	Women's Store	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant	Electronics Store
74	Caledonia-Fairbanks	2.0	Park	Women's Store	Market	Fast Food Restaurant	Grocery Store	Dessert Shop	Event Space	Ethiopian Restaurant	Empanada Restaurant	Electronics Store
90	The Kingsway, Montgomery Road, Old Mill North	2.0	River	Park	Women's Store	Drugstore	Diner	Discount Store	Dog Run	Doner Restaurant	Donut Shop	Dumpling Restaurant
98	Weston	2.0	Park	Women's Store	Eastern European Restaurant	Discount Store	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Dumpling Restaurant	Electronics Store

This neighborhood cluster is dominated by recreational facilities and retailers.
However, this cluster does not satisfy the first criteria, which is numerous Asian restaurants.

RESULTS

	Neighborhood	Cluster Labels	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
6	East Birchmount Park, Ionview, Kennedy Park	8.0	Discount Store	Department Store	Bus Station	Coffee Shop	Dumpling Restaurant	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Women's Store
13	Clarks Corners, Sullivan, Tam O'Shanter	8.0	Pharmacy	Pizza Place	Shopping Mall	Bank	Italian Restaurant	Thai Restaurant	Chinese Restaurant	Noodle House	Fried Chicken Joint	Fast Food Restaurant
15	L'Amoreaux West	8.0	Fast Food Restaurant	Chinese Restaurant	Pharmacy	Grocery Store	Pizza Place	Sandwich Place	Breakfast Spot	Coffee Shop	Bubble Tea Shop	Thrift / Vintage Store
24	Willowdale West	8.0	Pizza Place	Pharmacy	Discount Store	Coffee Shop	Drugstore	Diner	Dog Run	Doner Restaurant	Donut Shop	Eastern European Restaurant
89	Alderwood, Long Branch	8.0	Pizza Place	Gym	Pool	Sandwich Place	Athletics & Sports	Pub	Skating Rink	Coffee Shop	Pharmacy	Dumpling Restaurant
99	Westmount	8.0	Pizza Place	Middle Eastern Restaurant	Sandwich Place	Discount Store	Chinese Restaurant	Coffee Shop	Intersection	Diner	Dog Run	Doner Restaurant
101	Albion Gardens, Beaumond Heights, Humbergate, ...	8.0	Fast Food Restaurant	Fried Chicken Joint	Sandwich Place	Discount Store	Pizza Place	Beer Store	Japanese Restaurant	Pharmacy	Grocery Store	Gift Shop

RESULTS

- Restaurants dominate the above cluster of neighborhoods. Asian cuisine is in the top 10 in most neighborhoods, fulfilling the first criteria and suggesting that a new Asian restaurant may do well in this cluster of neighborhoods. The Willowdale neighborhood is the only neighborhood in this cluster without an Asian restaurant in the top 10. This neighborhood is a good prospect for the client to build a new Japanese restaurant.

- Further reviewing Willowdale West shows that the neighborhood only has three venues, 1 coffee shop, one discount store, and 1 pizza place. There is a dearth of sitdown restaurants. Willowdale West is in a neighborhood cluster where Asian food is popular but has no Asian restaurants itself. This neighborhood is where the client should build a new Japanese restaurant.

DISCUSSION

- The k-means algorithm is a very useful algorithm for solving problems such as identifying a good neighborhood to place a new restaurant. However, it does have limitations. Because it is an unsupervised algorithm, the user could not define criteria. This resulted in many clusters that were not useful.

- There are likely other algorithms that could generate similar if not better results without the limitation listed. Finding an algorithm better suited for this analysis could be an area of further research.

CONCLUSION

- To find a good neighborhood to build a new Japanese restaurant in, I utilized Toronto neighborhood data from Wikipedia, longitude and latitude data from Geocoder and venue data from Foursquare. I apply the k-means algorithm to identify a neighborhood that satisfied two criteria:
  1) In a neighborhood cluster with many Asian restaurants, indicating that Asian food is in demand.
  2) In a neighborhood that has few or zero Asian restaurants, indicating that competition is limited.

- The Willowdale West neighborhood satisfied both criteria. It is in a neighborhood cluster where there are Asian restaurants and host no Asian restaurant of its own; and therefore, is an ideal neighborhood for the client to build a new Japanese restaurant.