Capstone Project
Find the best place to open a Restaurant in Toronto

Business Problem

- In the city of Toronto, we are looking to open a new restaurant so the problem is where to open this restaurant.

- I would think of opening the restaurant beside an interesting place with a lot of visitors like a university, a gym or an airport.

- Also, it should not be a quiet place like a residential place for example to not making noise to the neighbors.

Data Source

- Wikipedia page with Toronto neighborhood data.

- https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

- This wikipedia page provides a list of postal codes in Canada staring with the letter "M" which means they are located within the city of Toronto in the province of Ontario.
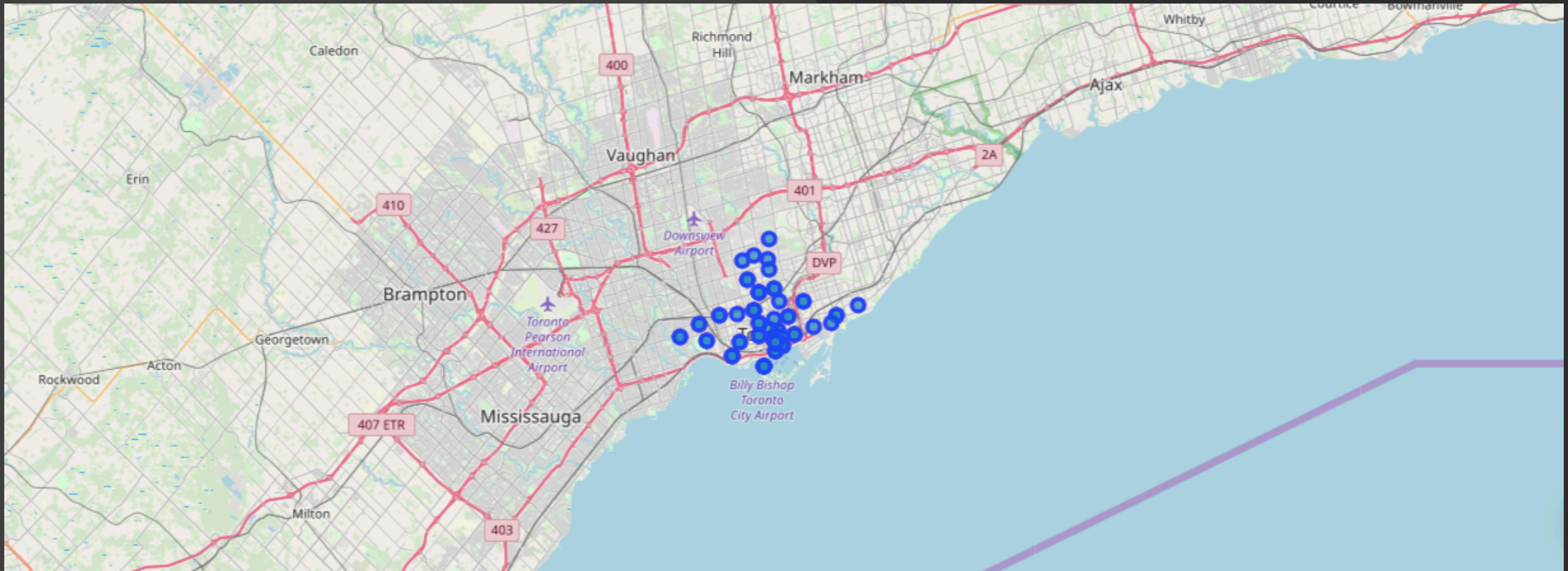
Data Cleaning

- We will scrape this Wikipedia page and wrangle the data we need, clean it and load it to a pandas dataframe to have a structured format.

- Then, we will use a readily-available csv file with all the geographical coordinates of Canada postal codes to assign each neighborhoud the longitude and latitude using the Geocoder package.

-

Data Filtration

- We will filter to get the data for Toronto Neighborhouds only.

- Then, we will use Foursquare to explore the venues using the neighborhouds coordinates and find the venues in interest for this project like universities, gyms or airports. This exploration will guide us with the best place to open a restaurant.

Toronto Map

Dataset after Cleaning and filtration

	Postcode	Borough	Neighbourhood	Latitude	Longitude
0	M5S	Downtown Toronto	University of Toronto	43.662696	-79.400049
1	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636
2	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
3	M6K	West Toronto	Parkdale Village	43.636847	-79.428191
4	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
5	M6P	West Toronto	High Park	43.661608	-79.464763
6	M6R	West Toronto	Parkdale	43.648960	-79.456325
7	M4T	Central Toronto	Moore Park	43.689574	-79.383160
8	M5T	Downtown Toronto	Grange Park	43.653206	-79.400049
9	M4V	Central Toronto	Deer Park	43.686412	-79.400049

Toronto Venues

- There are 128 unique venues extracted from Foursquare for the specific university and parks neighborhoods.

- We will only focus on venues containing "Restaurant" word like the example below.

	Neighborhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	University of Toronto	43.662696	-79.400049	Yasu	43.662837	-79.403217	Japanese Restaurant
1	University of Toronto	43.662696	-79.400049	Rasa	43.662757	-79.403988	Restaurant
2	University of Toronto	43.662696	-79.400049	Piano Piano	43.662949	-79.402898	Italian Restaurant
5	University of Toronto	43.662696	-79.400049	Cafe Cancan	43.662735	-79.403447	French Restaurant
11	University of Toronto	43.662696	-79.400049	Akai Sushi	43.662470	-79.404946	Sushi Restaurant

Number of restaurants in each neighborhood

- To have a better insight, we will count the number of restaurants existing in each neighborhood.
- Reason: when the neighborhood have less number of restaurants, it will be more prone to open more restaurants around.

	Neighborhood	Count
0	Berczy Park	9
1	Deer Park	4
2	Grange Park	33
3	High Park	6
4	Moore Park	1
5	Parkdale	4
6	Parkdale Village	4
7	Regent Park	6
8	University of Toronto	9

Number of restaurants in each neighborhood grouped by type

To be able to think of which type of restaurant we can open, we calculated the count of restaurants per type per neighborhood.

Neighborhood	Venue Category	Count
Berczy Park	Comfort Food Restaurant	1
	French Restaurant	1
	Greek Restaurant	1
	Italian Restaurant	1
	Restaurant	1
	Seafood Restaurant	2
	Thai Restaurant	1
	Vegetarian / Vegan Restaurant	1
Deer Park	American Restaurant	1
	Restaurant	1
	Sushi Restaurant	1
	Vietnamese Restaurant	1

One hot encoding

	Neighborhood	American Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant	Doner Restaurant	Dumpling Restaurant	Eastern European Restaurant	...	Italian Restaurant	Japanese Restaurant	Mexican Restaurant	Ramen Restaurant	Restaurant
0	University of Toronto	0	0	0	0	0	0	0	0	0	...	0	1	0	0	0
1	University of Toronto	0	0	0	0	0	0	0	0	0	...	0	0	0	0	1
2	University of Toronto	0	0	0	0	0	0	0	0	0	...	1	0	0	0	0
5	University of Toronto	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0
11	University of Toronto	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0

Prepare data for Clustering

We grouped the data by each neighborhood and taking the mean of number of restaurants in each neighborhood.

	Neighborhood	American Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant	Doner Restaurant	Dumpling Restaurant	Eastern European Restaurant	...	Italian Restaurant	Japanese Restaurant	Mexican Restaurant	Ramen Restaurant	Restaurant	S Rest
0	Berczy Park	0.00	0.000000	0.000000	0.000000	0.111111	0.00	0.000000	0.000000	0.00	...	0.111111	0.000000	0.000000	0.000000	0.111111	0.
1	Deer Park	0.25	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.00	...	0.000000	0.000000	0.000000	0.000000	0.250000	0.
2	Grange Park	0.00	0.000000	0.060606	0.151515	0.060606	0.00	0.030303	0.121212	0.00	...	0.030303	0.030303	0.121212	0.030303	0.000000	0.
3	High Park	0.00	0.166667	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.00	...	0.166667	0.000000	0.333333	0.000000	0.000000	0.
4	Moore Park	0.00	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.00	...	0.000000	0.000000	0.000000	0.000000	1.000000	0.
5	Parkdale	0.00	0.000000	0.000000	0.000000	0.000000	0.25	0.000000	0.000000	0.25	...	0.250000	0.000000	0.000000	0.000000	0.250000	0.
6	Parkdale Village	0.00	0.000000	0.250000	0.000000	0.000000	0.00	0.000000	0.000000	0.00	...	0.250000	0.000000	0.000000	0.000000	0.250000	0.
7	Regent Park	0.00	0.000000	0.000000	0.000000	0.000000	0.00	0.000000	0.000000	0.00	...	0.166667	0.000000	0.333333	0.000000	0.333333	0.
8	University of Toronto	0.00	0.000000	0.000000	0.111111	0.111111	0.00	0.000000	0.000000	0.00	...	0.111111	0.222222	0.000000	0.000000	0.222222	0.

Prepare data for Clustering (Cont'd)

We sorted the venues in a descending order and create a new dataframe to display the top 10 venues for each neighborhood

	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	Berczy Park	Seafood Restaurant	Thai Restaurant	Restaurant	Comfort Food Restaurant	Italian Restaurant	Greek Restaurant	French Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Dumpling Restaurant
1	Deer Park	Vietnamese Restaurant	Sushi Restaurant	Restaurant	American Restaurant	Seafood Restaurant	Thai Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant
2	Grange Park	Vegetarian / Vegan Restaurant	Chinese Restaurant	Vietnamese Restaurant	Dumpling Restaurant	Mexican Restaurant	Caribbean Restaurant	Comfort Food Restaurant	Italian Restaurant	Doner Restaurant	Filipino Restaurant
3	High Park	Mexican Restaurant	Thai Restaurant	Cajun / Creole Restaurant	Italian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Falafel Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant
4	Moore Park	Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant	Doner Restaurant	Dumpling Restaurant
5	Parkdale	Restaurant	Cuban Restaurant	Italian Restaurant	Eastern European Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant
6	Parkdale Village	Caribbean Restaurant	Restaurant	Italian Restaurant	Falafel Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant
7	Regent Park	Restaurant	Mexican Restaurant	Italian Restaurant	French Restaurant	Vietnamese Restaurant	Falafel Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant
8	University of Toronto	Restaurant	Japanese Restaurant	Sushi Restaurant	Chinese Restaurant	Comfort Food Restaurant	Italian Restaurant	French Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Cajun / Creole Restaurant

K-Means Clustering

	Postcode	Borough	Neighborhood	Latitude	Longitude	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
0	M5S	Downtown Toronto	University of Toronto	43.662696	-79.400049	0	Restaurant	Japanese Restaurant	Sushi Restaurant	Chinese Restaurant	Comfort Food Restaurant	Italian Restaurant	French Restaurant	Vietnamese Restaurant	Eastern European Restaurant	Cajun / Creole Restaurant
1	M5A	Downtown Toronto	Regent Park	43.654260	-79.360636	2	Restaurant	Mexican Restaurant	Italian Restaurant	French Restaurant	Vietnamese Restaurant	Falafel Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant
2	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Seafood Restaurant	Thai Restaurant	Restaurant	Comfort Food Restaurant	Italian Restaurant	Greek Restaurant	French Restaurant	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Dumpling Restaurant
3	M6K	West Toronto	Parkdale Village	43.636847	-79.428191	4	Caribbean Restaurant	Restaurant	Italian Restaurant	Falafel Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Chinese Restaurant	Comfort Food Restaurant	Cuban Restaurant
5	M6P	West Toronto	High Park	43.661608	-79.464763	2	Mexican Restaurant	Thai Restaurant	Cajun / Creole Restaurant	Italian Restaurant	Fast Food Restaurant	Vietnamese Restaurant	Falafel Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant
6	M6R	West Toronto	Parkdale	43.648960	-79.456325	4	Restaurant	Cuban Restaurant	Italian Restaurant	Eastern European Restaurant	Vietnamese Restaurant	Fast Food Restaurant	Cajun / Creole Restaurant	Caribbean Restaurant	Chinese Restaurant	Comfort Food Restaurant

K-Means Clustering (Cont'd)

Results and Discussions

- Our analysis shows that Toronto has a lot of places where we can invest and open new restaurants in.

- Toronto has a very big area but we focused our attention only to places which contain greatest number of location candidates. Like universities and parks.

- After applying our filtration, we found 9 places where they met this criteria and we focused our analysis to these 9 neighborhoods only.

- We clustered these 9 neighborhoods based on the top 10 different restaurants in each neighborhood into 5 clusters.

- These clusters represent the recommended places where we can open a new restaurant.

Conclusion

- The purpose of this project is identify the best place to open a new restaurant in Toronto in order to help stackholders and investors to narrow down their search criteria and locations to open the new restaurant.

- We used a dataset for Toronto neighborhoods and we used Foursquare to extract the nearby venues. We only focused on locations near universities and parks.

- At the end, we used a clustering technique kmeans to identify the locations as longitude and latitude that are best from our point of view to open a new restaurant.

- Our analysis provides a recommendation for the investors and should be considered only as a starting point for investors to facilitate the process of searching for locations and recommend the best places to attract people to this new restaurant.