Introduction

In a city restaurants,bakeries, cafes, bars, etc. might get their supplies all from the same contractor. Assume that this contractor was looking to improve their service. The contractors are looking for a place to build a new warehouse where they can stock it locally within the same borough so that their customers get fresher products and in turn allow older customers to get a better quality of service. With a local warehouse then customers will be able to get their products delivered quicker. There would be less delays allowing which would give the contractor a better reputation as well as potentially more clients. The contractor should build the warehouse where it is close to its customers so that it can reduce the costs of transportation and net a bigger profit. The recommender system to be designed will find and sort the best neighborhoods for this task. For this system we will be looking at the borough of Scarborough in Toronto.

Data

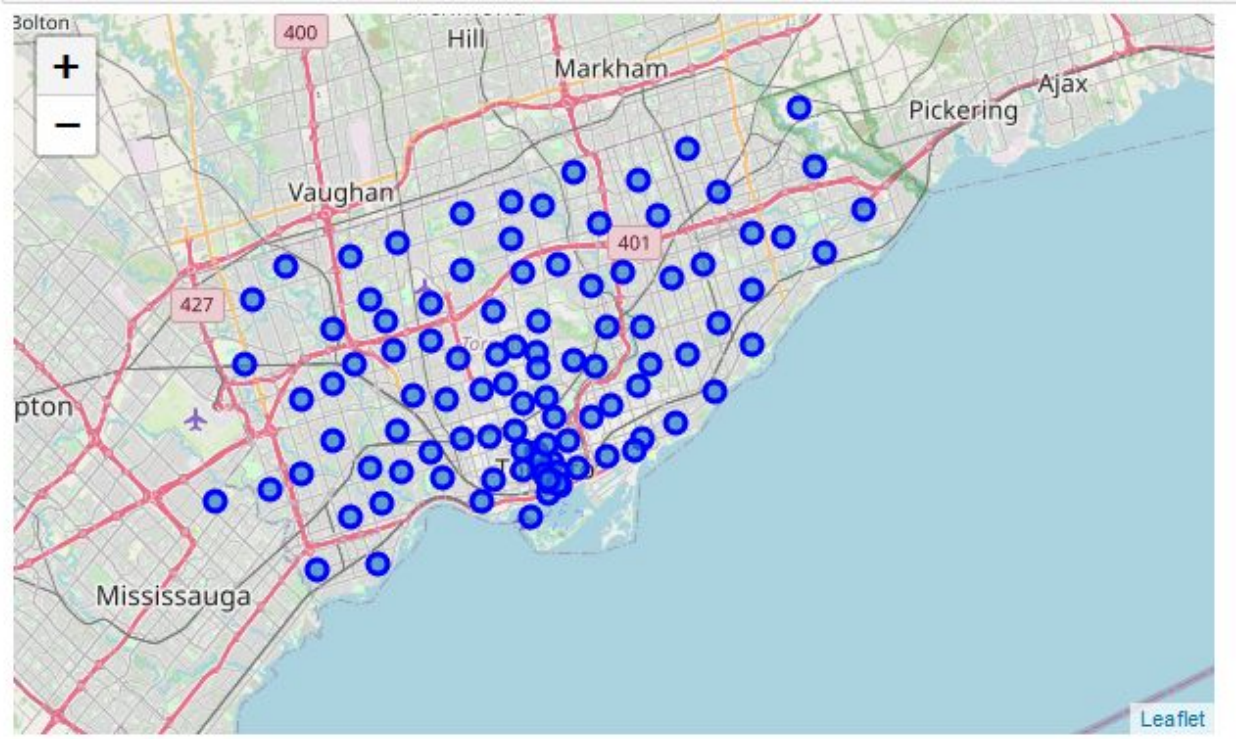
Some of the data we need are the locations of the neighborhoods in Toronto. We will acquire the postal codes from this link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. We will acquire the latitude and longitude location of each neighborhood from: http://cocl.us/Geospatial_data. Using these two sources we'll beagle to put together a list to go through and by using Foursquare we will be able to access different venues in each area. The locations of these venues will help us find the distance to the center of the neighborhood. The data gained from Foursquare will also help us determine whether the venue is popular as well as what category each venue would fall under.

Methodology

Using Pandas I read and extracted the data relevant from the initial two sources. The data was the postal code, borough, neighborhood, latitude and longitude.

	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	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494

With this data I used folium to create a visualization of the map of Toronto with its neighborhoods labeled.



By going through this list along with using the Foursquare api we were able to find the venues in each neighborhood as well as the relevant category.

	Postal Code	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Summary	Venue Category	Distance
0	M1B	Malvern, Rouge	43.806686	-79.194353	Harvey's	This spot is popular	Restaurant	807
1	M1B	Malvern, Rouge	43.806686	-79.194353	Wendy's	This spot is popular	Fast Food Restaurant	600
2	M1B	Malvern, Rouge	43.806686	-79.194353	Wendy's	This spot is popular	Fast Food Restaurant	387
3	M1B	Malvern, Rouge	43.806686	-79.194353	RBC Royal Bank	This spot is popular	Bank	906
4	M1B	Malvern, Rouge	43.806686	-79.194353	Caribbean Wave	This spot is popular	Caribbean Restaurant	912

Due to the nature of the unique categories we have to manually exclude the irrelevant ones in order to ensure we don't lose any potentially data we want to keep. We then make an updated data frame to look at the amount each neighborhood contains.

	American Restaurant	Asian Restaurant	BBQ Joint	Bakery	Breakfast Spot	Burger Joint	Cajun / Creole Restaurant	Cantonese Restaurant	Caribbean Restaurant	C R
Neighborhood										
Agincourt	1	1	1	2	1	0	0	1	2	8
Birch Cliff, Cliffside West	0	0	0	0	0	0	0	0	0	0
Cedarbrae	0	1	0	3	0	1	0	0	1	1
Clarks Corners, Tam O'Shanter, Sullivan	0	0	0	0	0	0	0	1	1	1
Cliffside, Cliffcrest, Scarborough Village West	0	0	0	0	0	1	1	0	0	0

Results

Since we know how many relevant venues each neighborhood has we can add them up. Using KMeans clustering we can divide the borough up to finalize our data and see which region would be the best for a warehouse.

	Bakery	Breakfast Spot	Diner	Fish Market	Food & Drink Shop	Grocery Store	Noodle House	Pizza Place	Sandwich Place	Total Restaurants	Total Joints
G1	2.000000	1.000000	0.000000	0.0	0.00	1.000000	1.000000	2.000000	1.000000	22.000000	1.000000
G3	1.000000	0.000000	0.000000	0.0	0.00	0.666667	0.666667	1.333333	1.000000	12.333333	1.333333
G5	2.000000	0.500000	0.000000	0.5	0.00	1.000000	0.000000	1.500000	0.000000	8.500000	2.500000
G2	0.750000	0.250000	0.000000	0.0	0.25	1.000000	0.000000	1.500000	1.000000	7.000000	0.500000
G4	0.333333	0.166667	0.166667	0.0	0.00	0.333333	0.000000	0.833333	0.333333	2.166667	0.333333

We can also see which neighborhoods are in each group and so the contractors know where they should focus their business to reduce travel and maximize quality.

	Neighborhood	Group
0	Agincourt	1
1	Birch Cliff, Cliffside West	4
2	Cedarbrae	5
3	Clarks Corners, Tam O'Shanter, Sullivan	3
4	Cliffside, Cliffcrest, Scarborough Village West	4
5	Dorset Park, Wexford Heights, Scarborough Town...	3
6	Golden Mile, Clairlea, Oakridge	4
7	Guildwood, Morningside, West Hill	2
8	Kennedy Park, Ionview, East Birchmount Park	2
9	Malvern, Rouge	2
10	Milliken, Agincourt North, Steeles East, L'Amo...	3
11	Rouge Hill, Port Union, Highland Creek	4
12	Scarborough Village	4
13	Steeles West, L'Amoreaux West	2
14	Wexford, Maryvale	5
15	Woburn	4

Discussion

Based on the results we can see group 1 would be the best choice while group 4 would be the worst choice based on our criteria. The algorithm used is straightforward. If we wanted a more accurate recommendation we would have to take into consideration the ingredients each restaurant uses which in turn would complicate the whole process.

Conclusion

By utilizing Foursquare we were able to solve our problem in recommending a location based on the number of restaurants and the location of nearby neighborhoods.