A Recommendation System for Groceries Contractor

Synopsis

- Part 1: **Problem Description**

There is a groceries contractor in one of the boroughs of Toronto (Scarborough). This contractor provides places such as: Different types of Restaurants, Bakery, Breakfast Spot, Brewery and Café with fresh and high-quality groceries. The contractor wants to build a warehouse for the groceries it buys from villagers and farmers inside the borough, so that they will support more customers and also bring better "Quality of Service" to the old customers.

Synopsis

- Part 2: **Data We Need**

a) **We will need geo-locational information about that specific borough and the neighborhoods in that borough. We assume it is "Scarborough" in Toronto. This is easily provided for us by the contractor, because the contractor has already made up his mind about the borough.**

Scarborough / Coordinates

43.7764° N, 79.2318° W

Synopsis

- Part 2: **Data We Need**

b) We will need data about different venues in different neighborhoods of that specific borough. In order to gain that information we will use "Foursquare" locational information. A typical request from Foursquare will provide us with the following information:

	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

```
arborough_venues.tail()
```

	Postal Code	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Summary	Venue Category	Distance
396	M1W	Steeles West, L'Amoreaux West	43.799525	-79.318389	Divine Wok Restaurant	This spot is popular	Chinese Restaurant	957
397	M1W	Steeles West, L'Amoreaux West	43.799525	-79.318389	Buddy Cafe	This spot is popular	Chinese Restaurant	973
398	M1W	Steeles West, L'Amoreaux West	43.799525	-79.318389	Birchwood Plaza	This spot is popular	Shopping Mall	977
399	M1W	Steeles West, L'Amoreaux West	43.799525	-79.318389	Olympian Swimming	This spot is popular	Gym Pool	978
400	M1W	Steeles West, L'Amoreaux West	43.799525	-79.318389	Dumpling & Szechuan Cuisine（川流不息店）	This spot is popular	Chinese Restaurant	989

Main Article

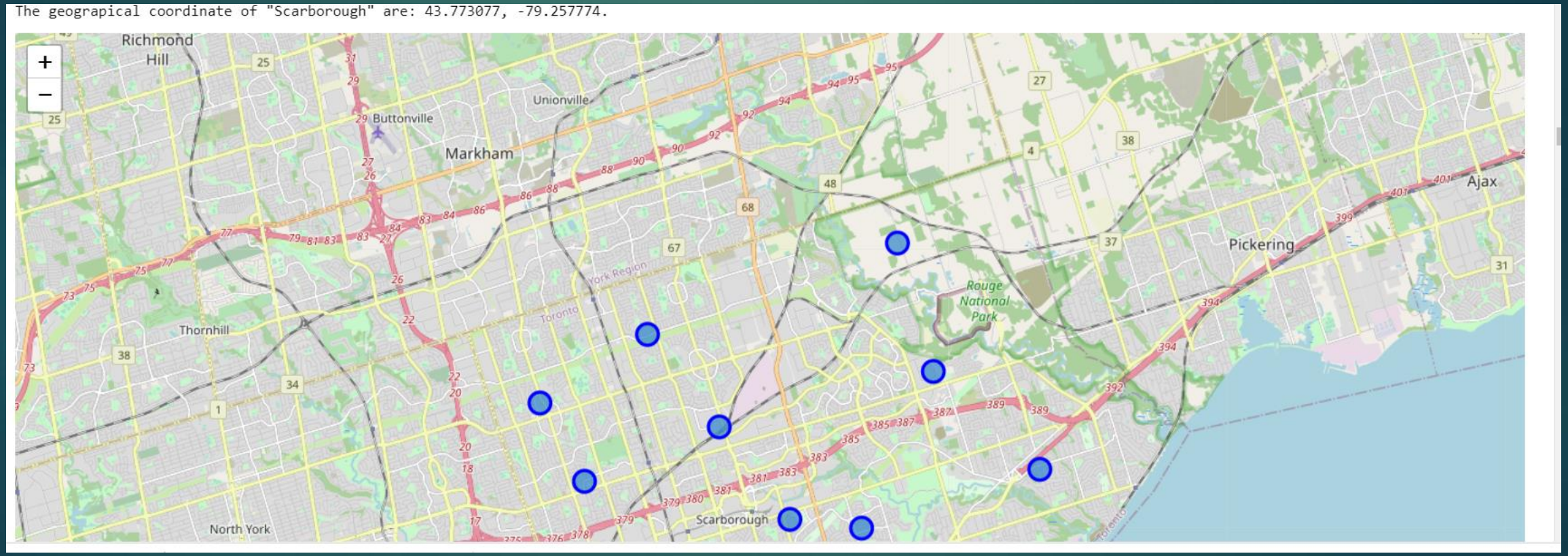
- Part 1: Identifying Postal Codes (and then Neighborhoods) in "Scarborough"

	Postal Code	Borough	Neighborhood	Latitude	Longitude
6	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
12	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	43.784535	-79.160497
18	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
22	M1G	Scarborough	Woburn	43.770992	-79.216917
26	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
32	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
38	M1K	Scarborough	Kennedy Park, Ionview, East Birchmount Park	43.727929	-79.262029
44	M1L	Scarborough	Golden Mile, Clairlea, Oakridge	43.711112	-79.284577
51	M1M	Scarborough	Cliffside, Cliffcrest, Scarborough Village West	43.716316	-79.239476
58	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848
65	M1P	Scarborough	Dorset Park, Wexford Heights, Scarborough Town...	43.757410	-79.273304
71	M1R	Scarborough	Wexford, Maryvale	43.750071	-79.295849
78	M1S	Scarborough	Agincourt	43.794200	-79.262029
82	M1T	Scarborough	Clarks Corners, Tam O'Shanter, Sullivan	43.781638	-79.304302
85	M1V	Scarborough	Milliken, Agincourt North, Steeles East, L'Amo...	43.815252	-79.284577
90	M1W	Scarborough	Steeles West, L'Amoreaux West	43.799525	-79.318389
95	M1X	Scarborough	Upper Rouge	43.836125	-79.205636

Main Article

- Part 1: Identifying Postal Codes (and then Neighborhoods) in "Scarborough"

The geograpical coordinate of "Scarborough" are: 43.773077, -79.257774.



Main Article

- Part 2: Connecting to Foursquare and Retrieving Locational Data for Each Venue in Every Neighborhood

After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 1000 meter. It means that we have asked Foursquare to find venues that are at most 1000 meter far from the center of the neighborhood.

Main Article

- Part 3: Processing the Retrieved Data and Creating a DataFrome for All the Venues inside the Scarborough

When the data is completely gathered, we will perform processing on that raw data to find our desirable features for each venue. Our main feature is the category of that venue. After this stage, the column "Venue's Category" wil be One-hot encoded and different venues will have different feature-columns. After On-hot encoding we will integrate all restaurant columns to one column "Total Restaurants" and all food joint columns to "Total Joints" column.

Main Article

- **Part 3: Processing the Retrieved Data and Creating a DataFrome for All the Venues inside the Scarborough**

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

Now, the dataset is fully ready to be used for machine learning (and statistical analysis) purposes.

Neighborhood	Bakery	Breakfast Spot	Diner	Fish Market	Food & Drink Shop	Grocery Store	Noodle House	Pizza Place	Sandwich Place
Agincourt	2	1	0	0	0	0	1	2	2
Birch Cliff, Cliffside West	0	0	1	0	0	0	0	0	0
Cedarbrae	2	0	0	0	0	1	0	1	0
Clarks Corners, Tam O'Shanter, Sullivan	0	0	0	0	0	1	1	2	2
Cliffside, Cliffcrest, Scarborough Village West	0	0	0	0	0	0	0	2	0
Dorset Park, Wexford Heights, Scarborough Town Centre	2	0	0	0	0	0	0	1	1
Golden Mile, Clairlea, Oakridge	2	0	0	0	0	1	0	1	1
Guildwood, Morningside, West Hill	0	0	0	0	1	1	0	3	1
Kennedy Park, Ionview, East Birchmount Park	0	0	0	0	0	2	0	1	1
Malvern, Rouge	0	0	0	0	0	0	0	0	1
Milliken, Agincourt North, Steeles East, L'Amoreaux East	2	0	0	0	0	1	2	2	0
Rouge Hill, Port Union, Highland Creek	0	1	0	0	0	0	0	0	0
Scarborough Village	0	0	0	0	0	1	0	1	1
Steeles West, L'Amoreaux West	2	1	0	0	0	2	0	2	1
Wexford, Maryvale	1	1	0	1	0	2	0	2	0
Woburn	0	0	0	0	0	0	0	0	0

Main Article

- **Part 4: Applying one of Machine Learning Techniques (K-Means Clustering)**

```
# import k-means from clustering stage
from sklearn.cluster import KMeans

# run k-means clustering
kmeans = KMeans(n_clusters = 5, random_state = 0).fit(scarborough_onehot)
```

	Bakery	Breakfast Spot	Diner	Fish Market	Food & Drink Shop	Grocery Store	Noodle House	Pizza Place	Sandwich Place	Total Sum
G1	1.5	1.00	0.00	0.5	0.0	2.0	0.00	2.00	0.50	7.50
G5	2.0	0.00	0.00	0.0	0.0	1.0	2.00	2.00	0.00	7.00
G3	2.0	0.25	0.00	0.0	0.0	0.5	0.25	1.25	1.00	5.25
G4	0.0	0.00	0.00	0.0	0.2	1.0	0.20	1.80	1.00	4.20
G2	0.0	0.25	0.25	0.0	0.0	0.0	0.00	0.00	0.25	0.75

Decision Making and Reporting Results

Now, we focus on the centers of clusters and compare them for their "Total Restaurants" and their "Total Joints". The group which its center has the highest "Total Sum" will be our best recommendation to the contractor. {Note: Total Sum = Total Restaurants + Total Joints.} This algorithm although is pretty straightforward yet is strongly powerful.

Decision Making and Reporting Results

Result:

Best Group is G1;

Second Best Group is G5; ¶

Third Best Group is G3;

Decision Making and Reporting Results

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

Decision Making and Reporting Results

	Neighborhood	Group
0	Agincourt	3
2	Cedarbrae	3
5	Dorset Park, Wexford Heights, Scarborough Town...	3
6	Golden Mile, Clairlea, Oakridge	3

The best place is located in

'Neighborhood': 'Agincourt',

'Neighborhood Latitude': 43.7942003,

'Neighborhood Longitude': -79.26202940000002

ALTERNATIVE: Dorset Park, Wexford Heights, Scarborough Town,Golden Mile, Clairlea, Oakridge

THANKYOU