Recommend Academic Institute to Open a Food/Restaurant Business

Applied Data Science Capstone

IBM Data Science Professional Certificate

Coursera

Submitted by:

Mehedi Masud

November 2018

Table of Contents

1. Introduction

One of my friends came to Canada as a new immigrant who had experience in food/restaurant business in his home country. He was wondering what to do in Canada. He asked me about the possibility to start a food corner/restaurant/coffee shop near an academic institute particularly in the Toronto city. My friend requested me to help him to suggest a suitable location or institute and type of restaurant/coffee shop he should start with. So he requested me to perform feasibility study by visiting all the academic institutes in Toronto. I told him I can recommend you proper place and suitable shops using my Data Science knowledge instead of visiting all the locations physically.

Objective and Target Audience: Opening a food/restaurant business is all about location, location, location. However, not every business is suitable for every location, and vice versa. It comes down to a combination of concept and ideal customer. So the objective of this project is to recommend someone who is interested to open a food/restaurant business nearby a college or university in the Toronto city. *Mainly, the outcome of this project is to recommend popular type of restaurants among the students nearby by each academic areas. Based on the results one can decide the type of restaurant he should open and in which institute.*

2. Data Description

The data that are required by this project is described below with exampls.

1- First we need to identify college and universities in the Toronto city. We collect all the college and university names from following link

  https://en.wikipedia.org/wiki/Category:Universities_and_colleges_in_Toronto

2. We filtered out some non interested institutes. After filtering we identified 21 institutes that need to be explored. Following table shows the names of the institutes.

	InstituteName
0	Centennial College
1	Business College
2	Canadian Forces College
3	Canadian Memorial Chiropractic College
4	Ewart College
5	George Brown College
6	Glendon College
7	The Glenn Gould School
8	Humber College
9	Islamic Institute of Toronto
10	Keele Campus (York University)
11	College and Seminary
12	The Michener Institute
13	OCAD University
14	Oxford College
15	Ryerson University
16	Seneca College
17	University of Toronto
18	Tyndale University College and Seminary
19	University of Guelph-Humber
20	York University

3. Then, we need geo-locational information (latitude and longitude) for each institute in the Toronto city. The samples are shown below that are obtained using "geolocator" package.

SL	institute	Latitude	Longitude
0	Centennial College	43.785792	-79.227810
1	Business College	43.656394	-79.382597
2	Canadian Forces College	43.742528	-79.413474
3	Canadian Memorial Chiropractic College	43.802370	-79.371136
4	Ewart College	43.686855	-79.473009
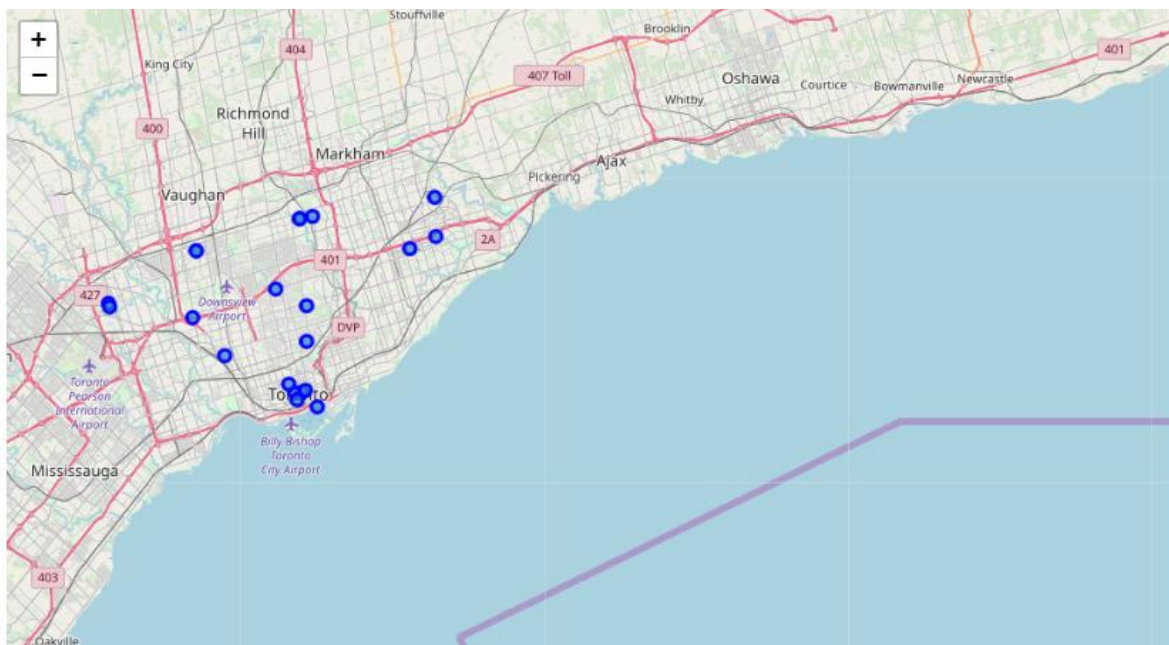
Map of Institutes using latitude and longitude in Toronto City

4. We identified different venues nearby of each institute, and extract venues' data searching by venue category with

categoryId='4d4b7105d754a06374d81259' //Food category ID.

In order to retrieve food categories venues information, we used "Foursquare" API.

url='https://api.foursquare.com/v2/venues/search?categoryId={}&limit={}&radius={}&client_id={}&client_secret={}&ll={},{}&v={}'.format(categoryId,limit,radius,CLIENT_ID, CLIENT_SECRET, latitude, longitude, VERSION)

One result is shown below:

{u'name': u"Sid's Southside Cafe", u'categories': [{u'pluralName': u'College Cafeterias', u'primary': True, u'name': u'College Cafeteria', u'shortName': u'Cafeteria', u'id': u'4bf58dd8d48988d1a1941735', u'icon': {u'prefix': u'https://ss3.4sqi.net/img/categories_v2/education/cafeteria_', u'suffix': u'.png'}}], u'hasPerk': False, u'location': {u'distance': 160, u'labeledLatLngs': [{u'lat': 43.66201648509592, u'lng': -79.39850649514088, u'label': u'display'}], u'city': u'Toronto', u'cc': u'CA', u'country': u'Canada', u'state': u'ON', u'formattedAddress': [u'100 St. George St (University of Toronto)', u'Toronto ON', u'Canada'], u'crossStreet': u'University of Toronto', u'address': u'100 St. George St', u'lat': 43.66201648509592, u'lng': -79.39850649514088}, u'id': u'4ba247e6f964a5206ee937e3', u'referralId': u'v-1540363652'}
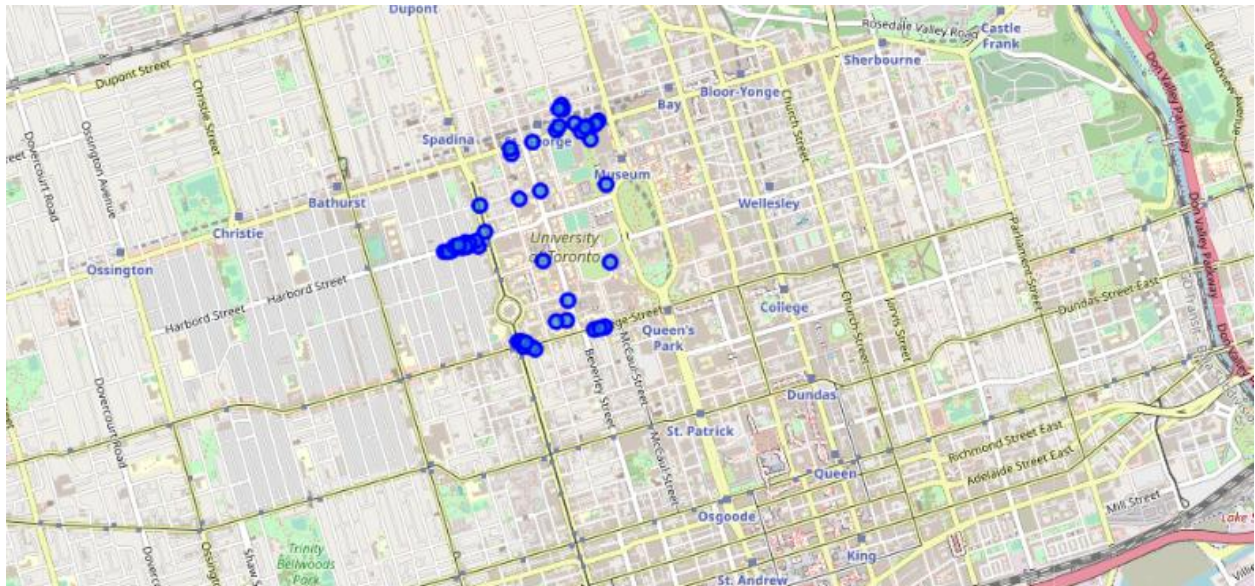
Map shows food venues near University of Toronto

5. From the results we extract venue names, venue category and locations for each institute. The partial result near University of Toronto is shown below:

SL	name	categories	lat	lng
0	Sid's Southside Cafe	College Cafeteria	43.662016	-79.398506
1	Starbucks	Coffee Shop	43.664942	-79.400100
2	Second Cup	Coffee Shop	43.665356	-79.398749
3	The Green Beet Cafe	Vegetarian / Vegan Restaurant	43.661996	-79.394181

6. After getting all relevant data we explore venues for each institute and perform some statistical analysis and machine learning (Clustering) to recommend proper venue category and place.

3. Methodology to Explore Institutes

I. First part is to identify, gather and process required data to explore institutes. The steps were discussed before, here we summarize:

- Identify names of academic institutes in the Toronto City. We only consider colleges and universities

- Identify location information (latitude and longitude) for each institute. We used "geolocator" package for this purpose

- Indentify different venues nearby of each institute, and extract venues data searching by venue category with "Food Category ID", our choice of interest. In order to retrieve food categories venues information, we used "Foursquare" API.

- Extract venue names, venue category and locations for each institute from the API results

II. Second part is to do exploratory and machine learning data analysis and . For this purpose we performed following analysis:

- Number of different venues of food categories for each institute. A sample is shown below

institute	Venue_Category	
Business College	American Restaurant	2
	Bakery	1
	Breakfast Spot	2
	Bubble Tea Shop	2
	Burger Joint	1
	Café	1

- Identify Top 20 venues of food categories among all institutes. See sample below.

	institute	Venue_Category	Frequency
247	The Michener Institute	Coffee Shop	15
8	Business College	Coffee Shop	12
143	OCAD University	Coffee Shop	12
200	Ryerson University	Coffee Shop	11
290	University of Toronto	Coffee Shop	8
185	Oxford College	Restaurant	6
238	Seneca College	Vietnamese Restaurant	6

- Find top 10 venues of food categories in institutes in Toronto Area. See sample below.

	Venue_Category	Frequency
22	Coffee Shop	98
65	Restaurant	37
17	Café	34
59	Pizza Place	30
33	Fast Food Restaurant	26

- Visualize the statistics. See figure below.

Number of venues per Venue Category in the institutes in Toronto
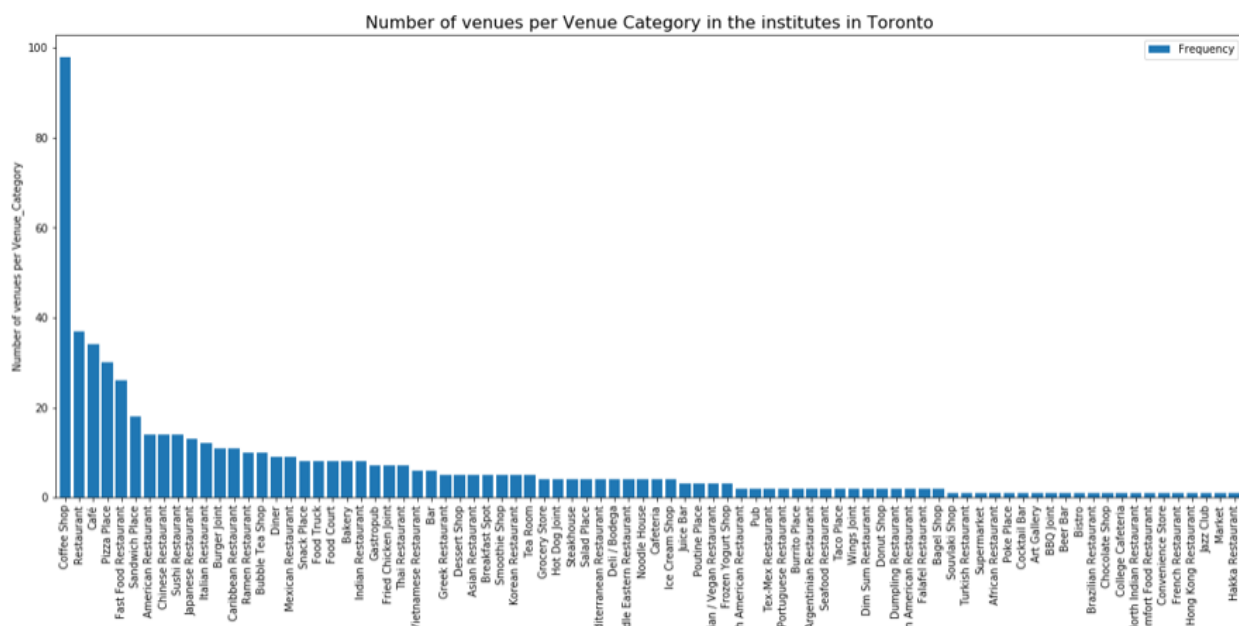
- Find number of Coffee shops in each institute

	institute	Venue_Category	Frequency
247	The Michener Institute	Coffee Shop	15
8	Business College	Coffee Shop	12
143	OCAD University	Coffee Shop	12
200	Ryerson University	Coffee Shop	11
290	University of Toronto	Coffee Shop	8
319	York University	Coffee Shop	6

-

- Find number of different restaurants in each institute and count different restaurants in all institutes.

	institute	Venue_Category	Frequency
185	Oxford College	Restaurant	6
238	Seneca College	Vietnamese Restaurant	6
246	The Michener Institute	Chinese Restaurant	4
23	Business College	Restaurant	4
229	Seneca College	Fast Food Restaurant	4

	Venue_Category	count
27	Restaurant	37
11	Fast Food Restaurant	26
30	Sushi Restaurant	14
6	Chinese Restaurant	14
1	American Restaurant	14
18	Japanese Restaurant	13
17	Italian Restaurant	12

- Analyze venues in each institute (the top 10 venues for each Institute). Samples shown below.

	institute	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 Common V
0	Business College	Coffee Shop	Restaurant	Diner	Ramen Restaurant	Bubble Tea Shop	American Restaurant	Breakfast Spot	
1	Canadian Forces College	Coffee Shop	Wings Joint	Donut Shop	College Cafeteria	Comfort Food Restaurant	Convenience Store	Deli / Bodega	
2	Canadian Memorial Chiropractic College	Korean Restaurant	Diner	Ice Cream Shop	Coffee Shop	Sandwich Place	Pizza Place	Food Court	
3	Centennial College	Coffee Shop	Fast Food Restaurant	Café	Restaurant	Japanese Restaurant	Bistro	Hot Dog Joint	
4	College and Seminary	Coffee Shop	Chinese Restaurant	Sandwich Place	Pizza Place	Korean Restaurant	Food Court	Deli / Bodega	

- Cluster Institutes based on venues. We used K-Means clustering algorithm and considered 5 clusters to find the best clusters. Sample shown below.

	institute	Cluster
0	Business College	1
1	Canadian Forces College	4
2	Canadian Memorial Chiropractic College	4
3	Centennial College	4
4	College and Seminary	4
5	Ewart College	2

4. Results

I. Exploratory analysis results

- Coffee Shop is more popular in academic institute

- Restaurant is second most popular

- "The Michener Institute" is good choice to open Coffee Shop

- For restaurant category "general" and "First food restaurant" is more popular

- "Seneca College" and "Humber College" is good to open a restaurant

II. Clustering analysis results

- Best cluster is C3 which consist of following institutes.

	institute	Cluster
10	Keele Campus (York University)	3
19	University of Toronto	3
20	York University	3

- The second best cluster is C1 which consist of following institutes.

	institute	Cluster
0	Business College	1
11	OCAD University	1
13	Ryerson University	1
16	The Michener Institute	1

5. Discussions

The objective of this project is to recommend someone who is interested to open a food corner/restaurant/coffee shop nearby a college or university in the Toronto city. To do the job we performed several exploratory analysis and applied K-Mean clustering algorithm. Following points are observed.

- Students like more Coffee shop and Cafe in an academic institute. The "Michener Institute" has most "Coffee shop"

- Among restaurants "First-food" restaurant is very popular.

- Among types if restaurant "Chinese" and "American" Restaurants are very popular

- " Vietnamese Restaurant " is popular in Seneca College

- Among cluster, cluster with institutes "Keele Campus", "Toronto University" and "York University" is the best cluster.

6. Recommendation

- Opening a "Coffee shop" is a good choice in an academic institute.

- The "Michener Institute" is recommended to open a "Coffee shop"

- Second recommendation is to open a "general" restaurant or "Fast Food" restaurant.

- "Seneca College" and "Humber College" are good places to open a general or first food restaurant

- If someone wants to open a " Vietnamese Restaurant " recommendation is "Seneca College"

- In general one is recommended to consider the cluster with institutes "Keele Campus", "Toronto University" and "York University" is the best cluster.

7. Conclusion

In this project we identified academic institutes in the Toronto city and explore venues considering venues with food category. The objective project is to recommend someone who is interested to open a food corner/restaurant/coffee shop nearby a college or university in the Toronto city. Mainly, the outcome of this project is to recommend popular type of restaurants among the students nearby by each academic areas. Based on the results one can decide the type of restaurant he should open and in which institute.