Identifying clusters of Indian Restaurants in Toronto

Monir Moniruzzaman

July 4, 2019

1. Introduction/Background

Toronto is a popular destination for exploring various attractive areas/places of interest e.g. Casaloma, Toronto Zoo, Royal Ontario Museum etc. Toronto is also known for its restaurants with cuisine from all over the world. Indian restaurants are very popular among the visitors in Toronto. It is therefore intriguing to identify clusters of Indian restaurants so visitors interested in Indian cuisine can go in one of those clusters to try different authentic Indian foods. The findings of this project would also be of interest for someone who wants to open a new Indian restaurant in Toronto. In this project, I will use Foursquare API to call for Indian restaurants and analyze them to identify potential clusters of Indian restaurant. I will also try to visualize the data from Foursquare using the visualization library, Folium.

2. Data Sources

This project will utilize data from Foursquare. Foursquare is a technology company that built a massive dataset of location data. What is interesting about Foursquare is that they were very smart about building their dataset. They actually crowd-sourced their data and had people use their app to build their dataset and add venues and complete any missing information they had in their dataset. Currently its location data is the most comprehensive out there, and quite accurate that it

powers location data for many popular services like Apple Maps, Uber, Snapchat, Twitter and many others, and is currently being used by over 100,000 developers, and this number is only growing.

The data returned from Foursquare API includes many information about restaurants including latitude, longitude, distance from search point, city, province, complete address, business category etc.


```
{'meta': {'code': 200, 'requestId': '5d1cf90a5d891b00397a5ecd'},
 'response': {'venues': [{'id': '4b7ccc72f964a520e3a52fe3',
    'name': 'Banjara Indian Cuisine',
    'location': {'address': '164 Eglinton Ave E',
     'crossStreet': 'at Redpath Ave',
     'lat': 43.7078104847312,
     'lng': -79.3932956275409,
     'labeledLatLngs': [{'label': 'display',
       'lat': 43.7078104847312,
       'lng': -79.3932956275409}],
     'distance': 2389,
     'cc': 'CA',
     'city': 'Toronto',
     'state': 'ON',
     'country': 'Canada',
     'formattedAddress': ['164 Eglinton Ave E (at Redpath Ave)',
      'Toronto ON',
      'Canada']},
```


Figure 1: Example of raw Foursquare data in json format

After numerous cleaning and filtering, the dataframe with Indian restaurants in Toronto looks as following table:

	name	categories	address	cc	city	country	crossStreet	distance	formattedAddress	labeledLatLngs	lat	lng	postalCode	state	id
0	Banjara Indian Cuisine	Indian Restaurant	164 Eglinton Ave E	CA	Toronto	Canada	at Redpath Ave	2389	[164 Eglinton Ave E (at Redpath Ave), Toronto ...	[{'label': 'display', 'lat': 43.7078104847312,...	43.707810	-79.393296	NaN	ON	4b7ccc72f964a520e3a52fe3
2	Aroma Fine Indian Restaurant	Indian Restaurant	287 King St. W	CA	Toronto	Canada	at John St.	4659	[287 King St. W (at John St.), Toronto ON M5V ...	[{'label': 'display', 'lat': 43.64646252150344...	43.646463	-79.389644	M5V 1J5	ON	4aef8854f964a5201cd921e3
4	Utsav Indian Cuisine	Indian Restaurant	69 Yorkville Ave.	CA	Toronto	Canada	NaN	2018	[69 Yorkville Ave., Toronto ON M5R 1B8, Canada]	[{'label': 'display', 'lat': 43.67115433553916...	43.671154	-79.390802	M5R 1B8	ON	4d7c569786cfa1438065c8a0
5	Bhoj Indian Cuisine	Indian Restaurant	21 Davenport Rd	CA	Toronto	Canada	NaN	1933	[21 Davenport Rd, Toronto ON M5R 1H2, Canada]	[{'label': 'display', 'lat': 43.67276456431629...	43.672765	-79.389063	M5R 1H2	ON	56c743c7498e7033ff262ce2
6	Banjara Indian Cuisine	Indian Restaurant	796 Bloor St W	CA	Toronto	Canada	at Crawford St	3148	[796 Bloor St W (at Crawford St), Toronto ON M...	[{'label': 'display', 'lat': 43.66291638533112...	43.662916	-79.421911	M6G 1L8	ON	4adb969ef964a520332921e3
7	Indian Roti House	Indian Restaurant	256 Queens Quay W	CA	Toronto	Canada	btwn Rees St & Lower Simcoe S	5539	[256 Queens Quay W (btwn Rees St & Lower Simco...	[{'label': 'display', 'lat': 43.63906038875002...	43.639060	-79.385422	M5J 1B5	ON	50b79b94e4b0a577af25a83f
10	Indian Biriyani House	Indian Restaurant	181 Dundas St W	CA	Toronto	Canada	W of Chestnut St	3796	[181 Dundas St W (W of Chestnut St), Toronto O...	[{'label': 'display', 'lat': 43.65511996683289...	43.655120	-79.386645	M5G 1C7	ON	4afd920ff964a520ad2822e3

Figure 2: Cleaned restaurants data for Toronto

Visualizing the data

Using folium library, we could show the location of the Indian restaurants in Toronto where the red dot represents Toronto downtown.

Figure 3: Location of Indian restaurants in Toronto

3. Methodology: k-means clustering

The primary purpose of *k*-means algorithm is to split a bunch of data points into a smaller number of groups or clusters. In doing so, the clustering technique minimizes the within-group sum of squares among the variables considered in the cluster analysis. See for more details about the clustering algorithm in Hartigan & Wong (1979).

Elbow method was used to identify suitable number of clusters in the data by means of k-means clustering algorithm. We can see there are 2/3 potential clusters of Indian restaurants in Toronto area.

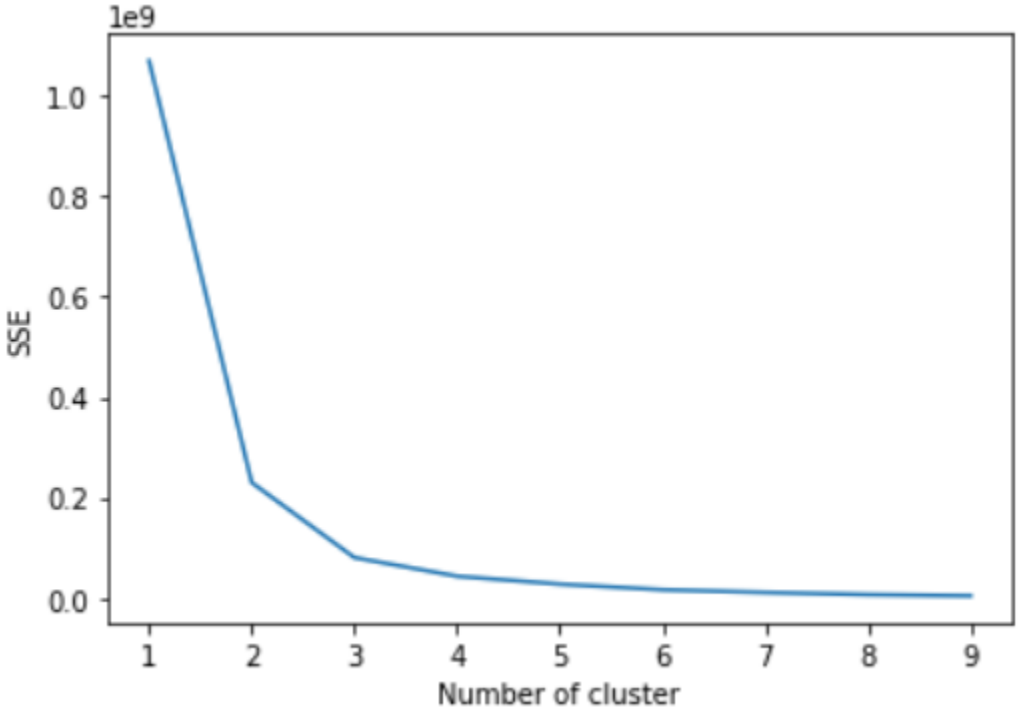

Figure 4: Identifying number of clusters using Elbow method

4. Results

After running k-means clustering with three clusters, we attached the cluster labels with each row in the restaurant data to identify which cluster each restaurants belong to.

	Cluster Labels	name	categories	address	cc	city	country	crossStreet	distance
0	0	Banjara Indian Cuisine	Indian Restaurant	164 Eglinton Ave E	CA	Toronto	Canada	at Redpath Ave	2389
2	0	Aroma Fine Indian Restaurant	Indian Restaurant	287 King St. W	CA	Toronto	Canada	at John St.	4659
4	0	Utsav Indian Cuisine	Indian Restaurant	69 Yorkville Ave.	CA	Toronto	Canada	NaN	2018
5	0	Bhoj Indian Cuisine	Indian Restaurant	21 Davenport Rd	CA	Toronto	Canada	NaN	1933
6	0	Banjara Indian Cuisine	Indian Restaurant	796 Bloor St W	CA	Toronto	Canada	at Crawford St	3148
7	0	Indian Roti House	Indian Restaurant	256 Queens Quay W	CA	Toronto	Canada	btwn Rees St & Lower Simcoe S	5539
10	0	Indian Biriyani House	Indian Restaurant	181 Dundas St W	CA	Toronto	Canada	W of Chestnut St	3796
11	0	Indian Street Food Co.	Indian Restaurant	1701 Bayview	CA	Toronto	Canada	Eglinton	3116
13	0	Indian Roti House	Indian Restaurant	NaN	CA	Toronto	Canada	NaN	3093
14	0	Maja Indian Cuisine	Indian Restaurant	345 Bloor	CA	Toronto	Canada	NaN	2546
15	2	Tich - Modern Indian Cuisine	Indian Restaurant	2314 lakeshore Blvd. east	CA	Etobicoke	Canada	NaN	10358
16	2	Karaikudi Chettinad South Indian Restaurant	Indian Restaurant	1225 Kennedy Rd	CA	Toronto	Canada	at Forbes (Between Lawrence and Ellesmere)	12691
17	0	Marigold Indian Bistro	Indian Restaurant	552 Mount Pleasant Rd.	CA	Toronto	Canada	NaN	2072
18	0	Earth Indian Express	Indian Restaurant	NaN	CA	Toronto	Canada	NaN	1984
19	1	Indian Hero Restaurant	Indian Restaurant	8920 Highway 50	CA	Brampton	Canada	NaN	22478

Figure 5: Results of cluster analysis

Visualizing the clusters

Finally, let us visualize the clusters on the Toronto map.

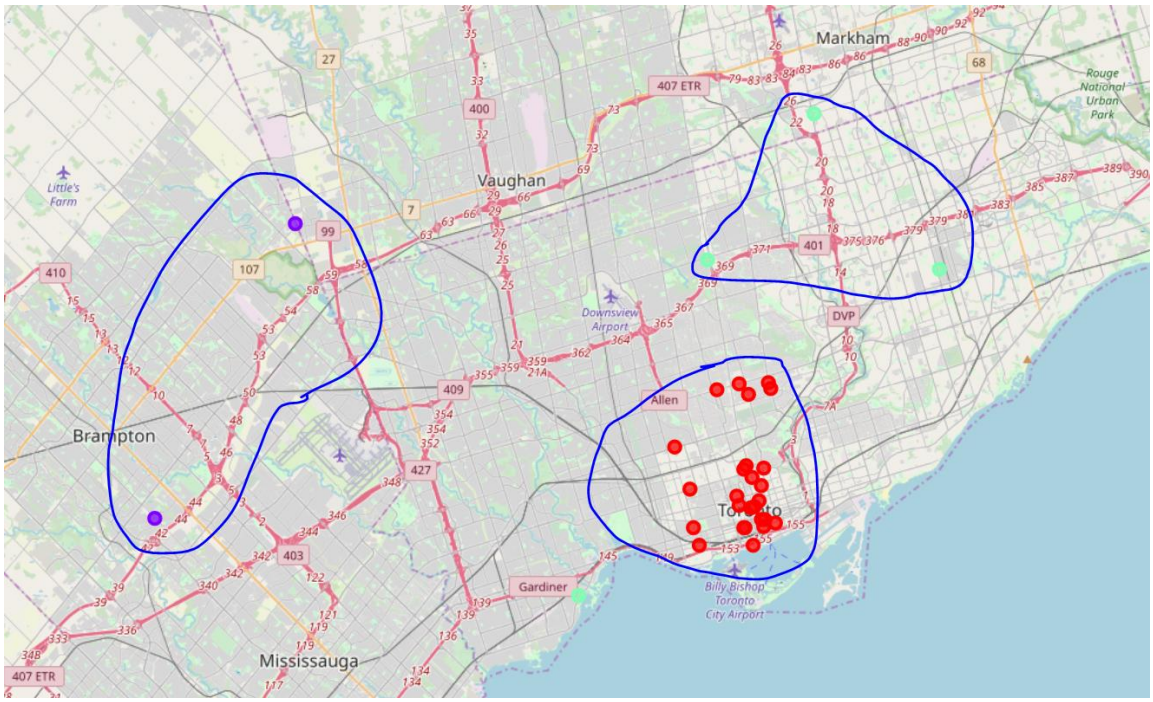

Figure 6: Map of 3 clusters of Indian restaurants in Toronto

5. Conclusions and further directions

- The map shows that there are only one major cluster of Indian restaurants in Toronto.
- Two other clusters identified in this study have only few restaurants in them and are further away from downtown.
- So we can conclude that if an Indian food lover visitor in Toronto would like to try authentic Indian foods, should go to cluster 0 (zero) where there are many options to choose from.
- For new business (Indian restaurant), it would be competitive to establish a new Indian restaurant in cluster zero. Additional analyses would be required to confirm where exactly a new Indian restaurant could be established in the vicinity.