IDENTIFYING POTENTIAL LOCATION FOR CHINESE RESTAURANT IN TORONTO

The Battle of Neighborhoods - Applied Data Science Capstone by IBM/Coursera

Godwin E. AKPAN

February 2020

Identifying potential location for Chinese Restaurant in Toronto

The Battle of Neighborhoods - Applied Data Science Capstone by IBM/Coursera

---

Having completed "IBM Data Science Professional Certificate" course, and been equipped with the skills and the tools to use location data to explore a geographical location, I was tasked with a capstone project to come up with an idea to leverage the Foursquare location data to explore or compare neighborhoods or cities of my choice or to come up with a problem that I can use the Foursquare location data to solve. I chose to explore the city of Toronto in Ontario, Canada for a possible best location to open Chinese restaurant by a businessperson who relies on expert (data scientist) recommendation.

Introduction: Business Problem

Toronto, the most populous city in Canada attracts both visitors and investors because of its many neighbourhoods with diverse businesses, culture and arts.

This project aims at finding best possible location for a restaurant within the neighbourhoods. This is to particularly assist a businessperson who is interested in opening a Chinese restaurant in Toronto, Ontario. Meanwhile there are a lot of restaurants in Toronto. We will try to identify locations with less or no restaurants, especially neighbourhoods with no Chinese restaurants.

Using data science approaches, we will generate a few most likely neighbourhoods, clearly expressing advantages of each area in order that the businessperson can choose the best possible location.

Data

Neighbourhoods and boroughs in Toronto were obtained from the Wikipedia page: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. To have the Wikipedia page in a structured format for exploring and clustering the neighbourhoods in the city of Toronto, **BeautifulSoup package** (https://beautiful-soup-4.readthedocs.io/en/latest/) was used to scrape the page, wrangle and clean the data, and read into pandas dataframe. The geographical coordinates of each Toronto's postal code area were obtained from csv file link: http://cocl.us/Geospatial_data. **Foursquare API** was used to obtain number of restaurants and their type alongside location in every neighbourhood.

Methodology

This project is about identifying areas (neighborhoods) with low density of Chinese restaurants within neighborhoods in the city of Toronto. Required data was collected through the scraping of Wikipedia page using BeautifulSoup package, read into Pandas dataframe. Geopy.geocoders was used to convert

addresses in each Toronto's postal code area into geographical coordinates (Table 1). Foursquare API was used to explore and segment the neighborhoods, and extract information on venues and category of each venue in each neighborhood.

Having identified boroughs with Chinese restaurants through UNIQUE-VENUE-CATEGORIZATION, analysis and exploration of 'Chinese restaurant density' across boroughs (neighborhoods) in Toronto was performed. Both bar graphs and scatter plots (violin plot) were used to visualise the density of Chinese restaurants in each borough.

K-means clustering was used to group the neighborhoods into 5 and visualised on map using **folium**. Counts of unique venues (including Chinese restaurant) were produced by means of bar graphs. Detailed analysis is linked to https://github.com/Jedidiah24/Capstone-Exercise/blob/master/Capstone%20-%20The%20Battle%20of%20Neighborhoods.ipynb.

Table 1: Dataframe of first five rows of neighborhoods with geographical coordinates

```
In [11]: df2 = pd.merge(df_data_table, df1, on='Postcode')
         df2.head()
```

Out[11]:

	Postcode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge,Malvern	43.806686	-79.194353
1	M1C	Scarborough	Port Union,Rouge Hill,Highland Creek	43.784535	-79.160497
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476

Results and Discussion

In trying to find best possible location(s) for Chinese restaurant within Toronto neighbourhoods, to assist a businessperson who is interested in opening a Chinese restaurant in Toronto, Ontario, 11 boroughs, 103 neighborhoods and 2213 venues were identified (Tables 2-4). The boroughs were Central Toronto, Downtown Toronto, East Toronto, East York, Etobicoke, Mississauga, North York, Queen's Park, Scarborough, West Toronto and York. Apart from Queen's Park, 267 unique venue categories were identified in other 10 boroughs (Table 5; Figure 1). Chinese restaurant was predicted to exist in the neighborhoods of 5 boroughs: Central Toronto, Downtown Toronto, Etobicoke, North York and Scarborough (Table 6; Figure 2).

Table 2: Boroughs and their respective number of neighborhoods

```
In [16]: #count Borough and Neighborhood
         df2.groupby('Borough').count()['Neighborhood']
```

```
Out[16]: Borough
         Central Toronto      9
         Downtown Toronto    19
         East Toronto         5
         East York            5
         Etobicoke           11
         Mississauga          1
         North York          24
         Queen's Park         1
         Scarborough         17
         West Toronto         6
         York                 5
         Name: Neighborhood, dtype: int64
```

Table 3: Venues size and first five rows of venues and their categories

```
In [34]: # check output
         print(venues.shape)
         venues.head()
```

```
(2213, 7)
```

Out[34]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Scarborough	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Scarborough	43.806686	-79.194353	Interprovincial Group	43.805630	-79.200378	Print Shop
2	Scarborough	43.784535	-79.160497	Chris Effects Painting	43.784343	-79.163742	Construction & Landscaping
3	Scarborough	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
4	Scarborough	43.784535	-79.160497	Affordable Toronto Movers	43.787919	-79.162977	Moving Target

Table 4: Count of venues in each borough (neighbourhood)

```
In [35]: venues.groupby('Neighborhood').count()
```

Out[35]:

	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Neighborhood						
Central Toronto	115	115	115	115	115	115
Downtown Toronto	1296	1296	1296	1296	1296	1296
East Toronto	122	122	122	122	122	122
East York	77	77	77	77	77	77
Etobicoke	74	74	74	74	74	74
Mississauga	12	12	12	12	12	12
North York	242	242	242	242	242	242
Scarborough	84	84	84	84	84	84
West Toronto	173	173	173	173	173	173
York	18	18	18	18	18	18

Table 5: Unique venue categories within neighborhoods of 10 boroughs in Toronto

```
to_grouped
```

```
(10, 267)
```

Out[52]:

	Neighborhoods	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	...	Tr St
0	Central Toronto	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.017391	...	0.0
1	Downtown Toronto	0.000000	0.000772	0.000772	0.000772	0.000772	0.001543	0.002315	0.001543	0.013117	...	0.0
2	East Toronto	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.024590	...	0.0
3	East York	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0
4	Etobicoke	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0
5	Mississauga	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.083333	...	0.0
6	North York	0.004132	0.000000	0.004132	0.000000	0.000000	0.000000	0.000000	0.000000	0.008264	...	0.0
7	Scarborough	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.011905	...	0.0
8	West Toronto	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0
9	York	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0

10 rows × 267 columns

```
toronto_map
```

```
The geograpical coordinates of Toronto are 43.653963, -79.387207.
```

Out[22]:

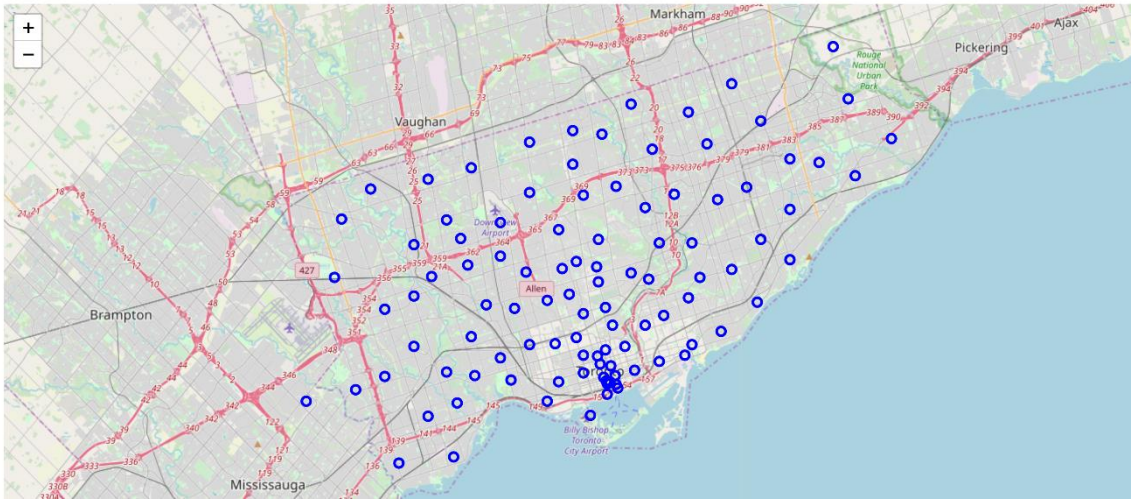

Figure 1: Geospatial distribution of neigborhoods across boroughs in Toronto

Table 6: Density of Chinese restaurant in each borough (neighbourhood)

```
In [56]: to_chinese
```

Out[56]:

	Neighborhoods	Chinese Restaurant
0	Central Toronto	0.008696
1	Downtown Toronto	0.009259
2	East Toronto	0.000000
3	East York	0.000000
4	Etobicoke	0.013514
5	Mississauga	0.000000
6	North York	0.016529
7	Scarborough	0.059524
8	West Toronto	0.000000
9	York	0.000000

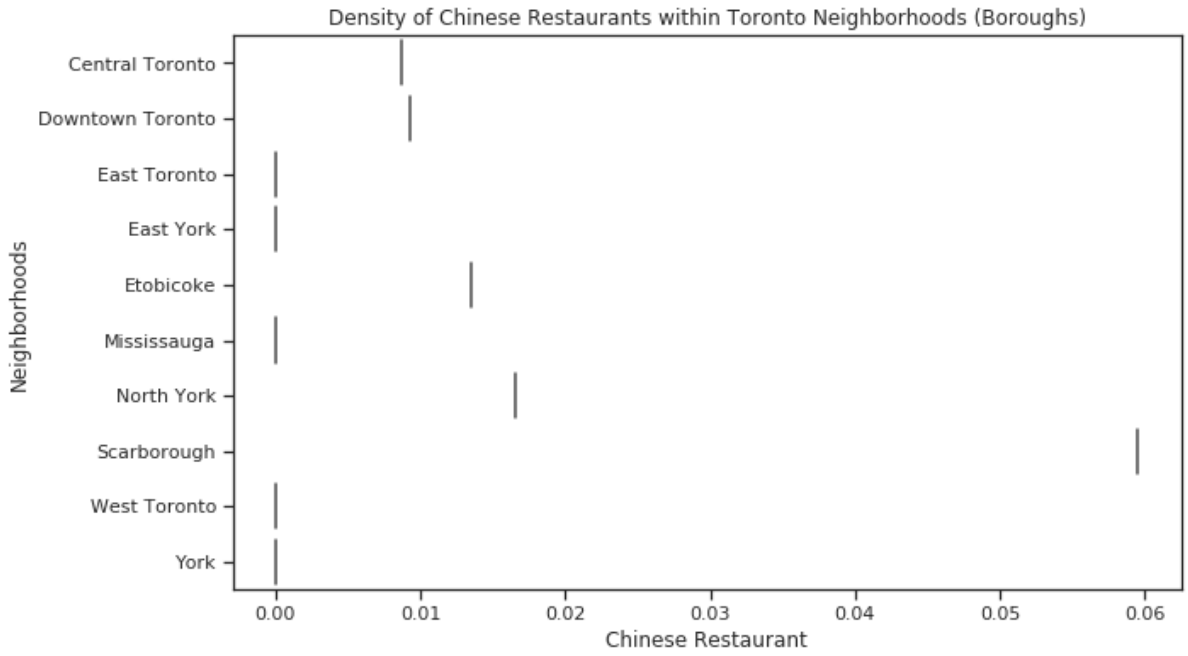

Figure 2: Density of Chinese restaurants within Toronto neighborhood (borough)

Analysis shows that similarities in venue category occur in 5 clusters across neighborhoods in Toronto (Figure 3). The first cluster - Cluster 0 (red) exists amongst the lone neighborhood in Mississauga (Table 2) with 10 most common venues more of food places, with very limited number of (or no) Chinese restaurant (Tables 6 and 7; Figures 2 and 4). The second cluster - Cluster 1 (purple) exists in York which has 5 neighborhoods (Table 2; Figure 3) with 10 most common venues (diverse), present in all 5 neighborhoods (Table 8; Figure 5). Analysis shows that there are very limited number of (on no) Chinese restaurant in Cluster 1 (Table 6; Figure 2).

```
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```

Out[65]:

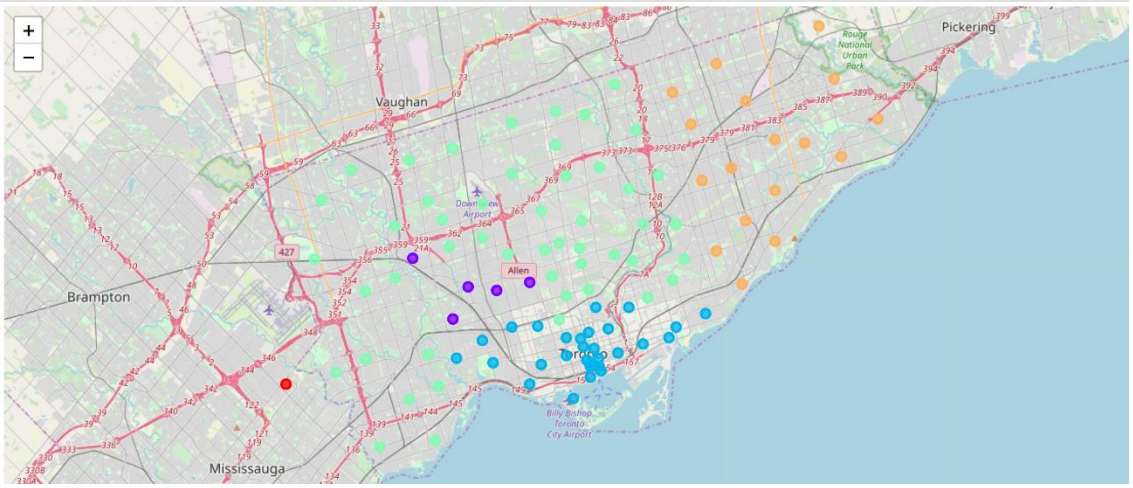

Figure 3: Clusters of similarities in venue category across neighborhoods in Toronto

Table 7: Ten most common venue categories in the first cluster - Cluster 0 (red)

```
In [67]: merged.loc[merged['Cluster Labels'] == 0, merged.columns[[1] + list(range(5, merged.shape[1]))]]
```

Out[67]:

	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	10 M C V
90	Canada Post Gateway Processing Centre	Hotel	Coffee Shop	American Restaurant	Middle Eastern Restaurant	Mediterranean Restaurant	Intersection	Burrito Place	Fried Chicken Joint	Gym	S Pl

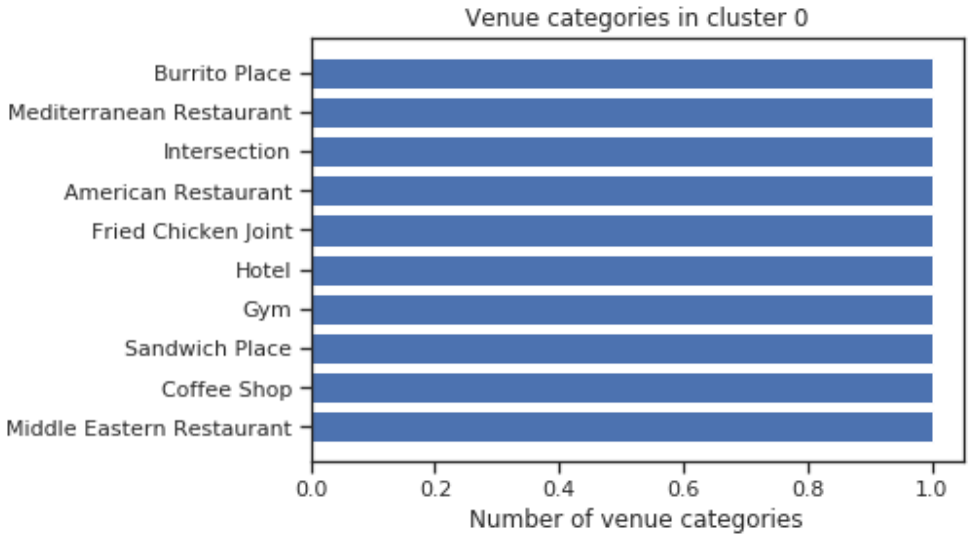

Figure 4: Count of ten most common venue categories in the neighborhoods of the first cluster (Cluster 0)

Table 8: Ten most common venue categories in the second cluster - Cluster 1 (purple)

```
In [69]: merged.loc[merged['Cluster Labels'] == 1, merged.columns[[1] + list(range(5, merged.shape[1]))]]
```

Out[69]:

	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
79	Humewood-Cedarvale	Park	Hockey Arena	Convenience Store	Pizza Place	Caribbean Restaurant	Restaurant	Market	Sandwich Place	Brewery
80	Caledonia-Fairbanks	Park	Hockey Arena	Convenience Store	Pizza Place	Caribbean Restaurant	Restaurant	Market	Sandwich Place	Brewery
81	Silverthorn,Mount Dennis,Keelesdale,Del Ray	Park	Hockey Arena	Convenience Store	Pizza Place	Caribbean Restaurant	Restaurant	Market	Sandwich Place	Brewery
82	The Junction North,Runnymede	Park	Hockey Arena	Convenience Store	Pizza Place	Caribbean Restaurant	Restaurant	Market	Sandwich Place	Brewery
83	Weston	Park	Hockey Arena	Convenience Store	Pizza Place	Caribbean Restaurant	Restaurant	Market	Sandwich Place	Brewery

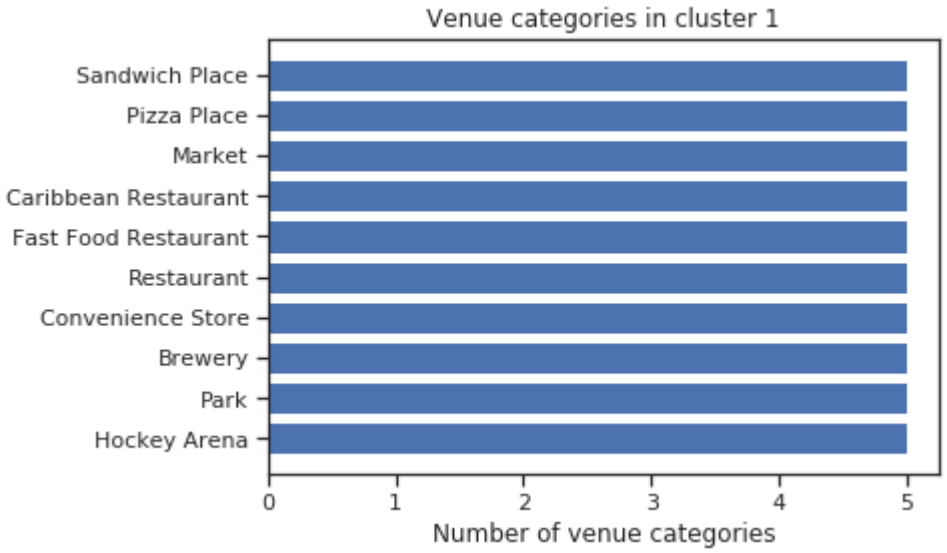

Figure 5: Count of ten most common venue categories in the neighborhoods of the second cluster (Cluster 1)

The third cluster - Cluster 2 (blue) exists in Downtown Toronto, East Toronto and West Toronto (Figure 3). They have a total of 30 neighborhoods (Table 2) with 10 most common venues more of food places (Figure 6; Table 9). Café, restaurant, coffee shop, Italian restaurant and bakery found in all 30 neighborhoods, while bar is found in about 24 neighborhoods, park in about 23, and hotel, Japanese restaurant and seafood restaurant in about 18 neighborhoods (Figure 6). Analysis shows that few Chinese restaurants exist in Downtown Toronto, and probably very few or none in East Toronto and West Toronto (Table 6; Figure 2).

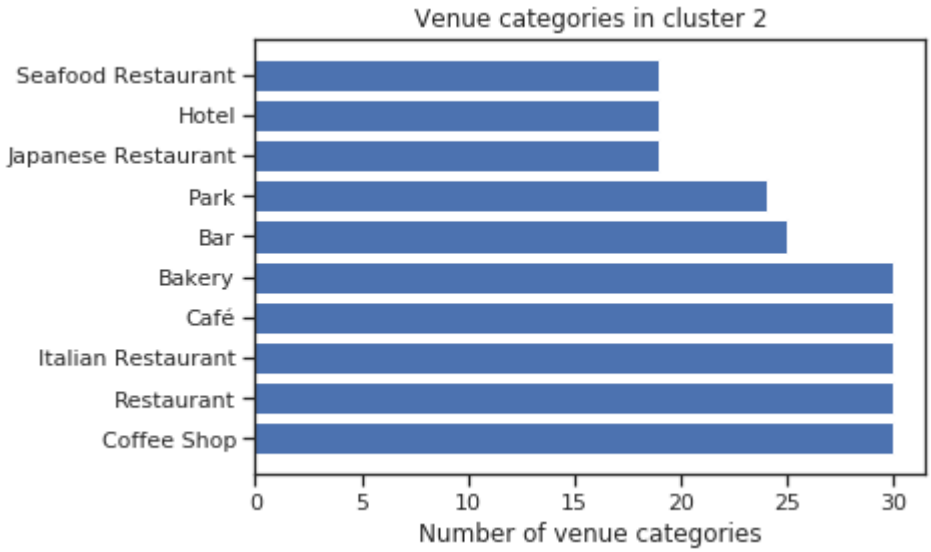

Figure 6: Count of ten most common venue categories in the neighborhoods of the third cluster (Cluster 2)

Table 9: Ten most common venue categories in the third cluster - Cluster 2 (blue)

```
In [71]: merged.loc[merged['Cluster Labels'] == 2, merged.columns[[1] + list(range(5, merged.shape[1]))]]
```

Out[71]:

	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 Cor Ven
46	The Beaches	Greek Restaurant	Coffee Shop	Italian Restaurant	Brewery	Café	Ice Cream Shop	Park	Pub	Res
47	Riverdale,The Danforth West	Greek Restaurant	Coffee Shop	Italian Restaurant	Brewery	Café	Ice Cream Shop	Park	Pub	Res
48	The Beaches West,India Bazaar	Greek Restaurant	Coffee Shop	Italian Restaurant	Brewery	Café	Ice Cream Shop	Park	Pub	Res
49	Studio District	Greek Restaurant	Coffee Shop	Italian Restaurant	Brewery	Café	Ice Cream Shop	Park	Pub	Res
50	Business Reply Mail Processing Centre 969 Eastern	Greek Restaurant	Coffee Shop	Italian Restaurant	Brewery	Café	Ice Cream Shop	Park	Pub	Res
60	Rosedale	Coffee Shop	Café	Restaurant	Hotel	Bakery	Italian Restaurant	Japanese Restaurant	Bar	Sea Res
61	St. James Town,Cabbagetown	Coffee Shop	Café	Restaurant	Hotel	Bakery	Italian Restaurant	Japanese Restaurant	Bar	Sea Res
62	Church and Wellesley	Coffee Shop	Café	Restaurant	Hotel	Bakery	Italian Restaurant	Japanese Restaurant	Bar	Sea Res
63	Harbourfront	Coffee Shop	Café	Restaurant	Hotel	Bakery	Italian Restaurant	Japanese Restaurant	Bar	Sea Res
64	Garden District,Ryerson	Coffee Shop	Café	Restaurant	Hotel	Bakery	Italian Restaurant	Japanese Restaurant	Bar	Sea Res
65	St. James Town	Coffee	Café	Restaurant	Hotel	Bakery	Italian	Japanese	Bar	Sea

The fourth cluster - Cluster 3 (green) exists in Etobicoke, North York, Central Toronto and East York (Figure 3) with a total of 54 neighborhoods (Table 2). The 10 most common venues are more of food

places (Table 10; Figure 7). Pizza place and coffee shop are predicted to exist in about 48 neighborhoods. Café and sandwich place in about 45 neighborhoods. Fast food restaurant in about 40 neighborhoods. Park, grocery store and clothing store in above 33 neighborhoods. Shopping mall and Japanese restaurant in about 24 neighborhoods. Analysis shows that there are a considerable number of Chinese restaurants in North York, few in Etobicoke, fewer in Central Toronto, and very few or none in East York (Table 6; Figure 2).

Table 10: Ten most common venue categories in the fourth cluster - Cluster 3 (green)

```
In [73]: merged.loc[merged['Cluster Labels'] == 3, merged.columns[[1] + list(range(5, merged.shape[1]))]]
```

Out[73]:

	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 Co Ver
17	Hillcrest Village	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
18	Henry Farm,Fairview,Oriole	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
19	Bayview Village	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
20	Silver Hills,York Mills	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
21	Willowdale,Newtonbrook	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
22	Willowdale South	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
23	York Mills West	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
24	Willowdale West	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
25	Parkwoods	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
26	Don Mills North	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
27	Don Mills South,Flemingdon Park	Coffee Shop	Clothing Store	Fast Food Restaurant	Pizza Place	Japanese Restaurant	Park	Sandwich Place	Shopping Mall	Grc Sto
	Bathurst Manor,Wilson	Coffee	Clothing	Fast Food	Pizza	Japanese		Sandwich	Shopping	Grc

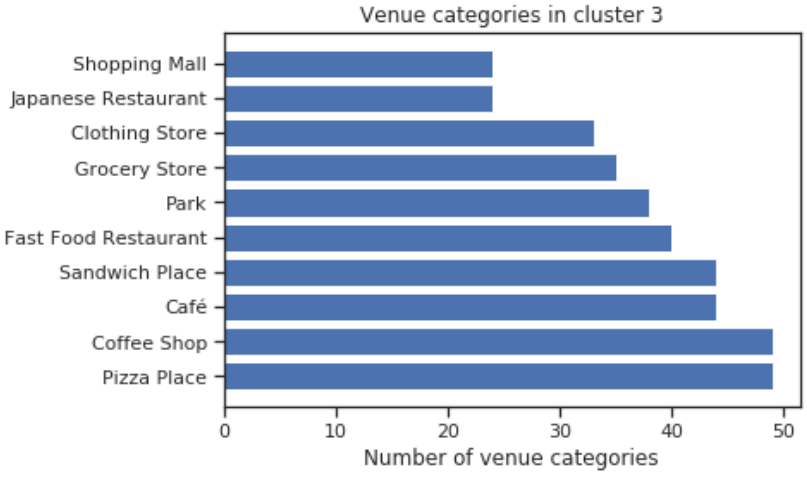

Figure 7: Count of ten most common venue categories in the neighborhoods of the fourth cluster (Cluster 3)

The fifth cluster - Cluster 4 (orange) exists in Scarborough (Figure 3) which has 17 neighborhoods (Table 2) with diverse 10 most common venues (including Chinese restaurant) in all 17 neighborhoods (Table 11; Figure 8). Analysis shows that cluster 4 has the highest number of Chinese restaurants, perhaps due to high Chinese population in Scarborough (Table 6; Figure 2).

Table 11: Ten most common venue categories in the fifth cluster - Cluster 4 (orange)

```
In [75]: merged.loc[merged['Cluster Labels'] == 4, merged.columns[[1] + list(range(5, merged.shape[1]))]]
```

Out[75]:

	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 I Com Venu
0	Rouge,Malvern	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
1	Port Union,Rouge Hill,Highland Creek	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
2	Guildwood,Morningside,West Hill	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
3	Woburn	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
4	Cedarbrae	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
5	Scarborough Village	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
6	East Birchmount Park,Ionview,Kennedy Park	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
7	Golden Mile,Oakridge,Clairlea	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
8	Cliffcrest,Scarborough Village West,Cliffside	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
9	Cliffside West,Birch Cliff	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar
10	Wexford Heights,Dorset Park,Scarborough Town C...	Chinese Restaurant	Breakfast Spot	Fast Food Restaurant	Bakery	Coffee Shop	Pizza Place	Thai Restaurant	Bank	Phar

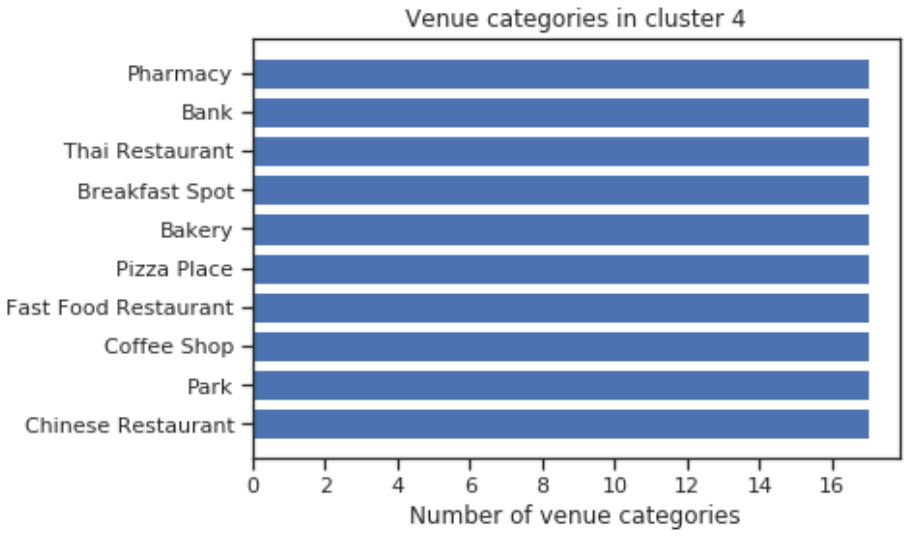

Figure 8: Count of ten most common venue categories in the neighborhoods of the fifth cluster (Cluster 4)

Generally, analysis show that among 78 venues for targeting food and hangout places, Chinese restaurant comes at position 15 (Figure 9). These Chinese restaurants are mainly in Scarborough, North York, Etobicoke, Downtown Toronto and Central Toronto (Table 6; Figure 2). Neighborhoods in East York, East Toronto, Mississauga, West Toronto and York may most likely be potential areas for a businessperson to open Chinese restaurant, as they probably have very limited number of or no Chinese restaurant at all (Table 6; Figures 2and 3).

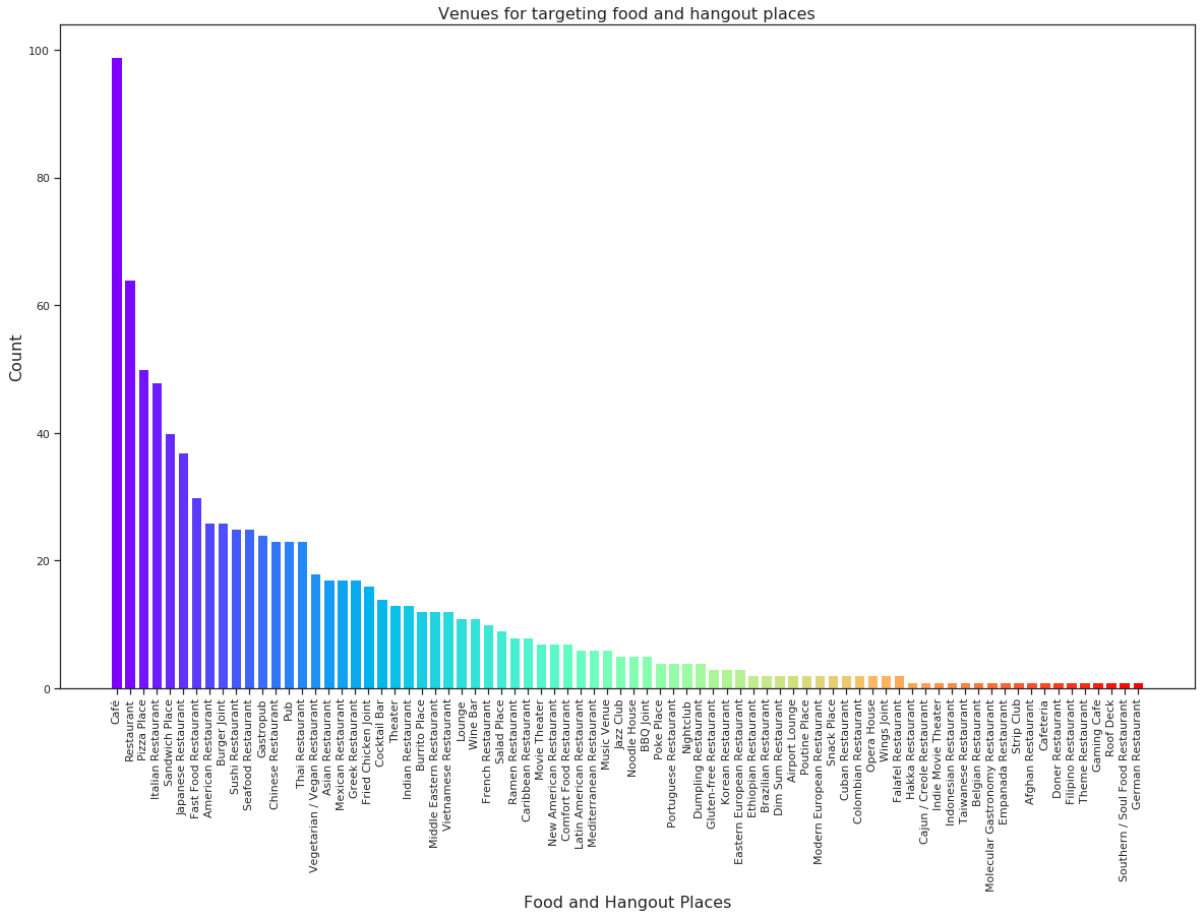

Figure 9: Count of venues for seventy-eight (78) targeting food and hangout places

Conclusion

Toronto is actually a city of neighborhoods. The aim of this project was to identify neighborhood(s) with less number of (or no) Chinese restaurant, to assist a businessperson who is willing to own a Chinese restaurant in Toronto. The businessperson will be at liberty to choose any neighborhood within the 5 boroughs with very limited number of (or no) Chinese restaurant for his Chinese restaurant business.

Results of this work may be of interest to individuals who wish to patronise Chinese restaurant in Toronto. It can give a quick overview of neighborhoods with potential Chinese restaurants, for example, Scarborough and North York.

However, this study can be further improved with the help of more data and different machine learning techniques. For example, population data and heatmap of possible suitable locations can improve the study greatly. Nonetheless, the results of this study have demonstrated to be of immense use within the context of choosing neighborhoods with no Chinese restaurant for Chinese restaurant business.