Exploring the Restaurants in Toronto

Jinling W. Flinn

1. Introduction

The city of Toronto is Canada's largest city, the provincial capital of Ontario, and the fourth largest city in North America [1] with a population of 2,731,571 in 2016 [2]. Toronto is also very diverse. It is one of the most multicultural cities in the world [3] with more than 50% of residents belong to a visible minority population groups [4]. In 2016, 16.7% of immigrants were born in Americas, 23.6% in Europe, 6.1% in Africa, 53.3% in Asia, and 0.3% in Oceania [5]. Toronto is the financial capital of Canada, home to the Toronto Stock exchange, and the headquarters of many large corporations. It is also one of Canada's leading tourism cities. 43 Million people visited Toronto in 2017[6].

As such a dynamic metropolis, Toronto is very attractive to various investors. In this report, the restaurants (café and fast food exclude) in Toronto will be explored. The potential clients are the people who are interested to invest a restaurant in Toronto. The findings in this report will help them to decide where and what type of restaurants they should invest in Toronto by answering a few questions: (1) which neighborhood is the most popular one (with the most venues) in Toronto? (2) which neighborhood has the most restaurants? (3) what are the relation between the number of venues and the number of restaurants in each neighborhood? (4) due to its diverse population what are the most popular cuisine in Toronto? and (5) any other recommendations to the types of restaurants should be invested?

2. Data

2.1 Data Sources

The location data of the venues in Toronto and the details of Toronto restaurants can be acquired from Foursquare API. In order to fetch the location data via Foursquare, the geographical coordinates of each neighborhood in Toronto are required. The coordinates

can be obtained by using the Geocoder package or a cvs file (Geospatial_Coordinates.csv from http://cocl.us/Geospatial_ data). In this project， the cvs file was used because the Geocoder package can be very unreliable. In this file, the latitude and longitude coordinates of each postal code are provided. However, there is lack of neighborhood names in this file. This information can be found from Wikipedia page (https://en.wikipedia.org/wiki/List_ of_postal_codes_of_ Canada:_M), which has a list of postal codes with borough name and neighborhood name in Canada.

2.2 Data Preparation

A notebook was created to scrape the Wikipedia page (https://en.wikipedia.org/wiki/List_ of_postal_codes_of_Canada:_M) to obtain the data in the table of postal codes and load it into a Python pandas DataFrame called "neighborhoods". The DataFrame has three columns: Postcode, Borough, and Neighborhood. To wrangle the data, the cells with a borough that is "Not assigned" were removed. If a cell has a borough but the neighborhood is "Not assigned", the neighborhood was assigned the same value as the borough. For the postal code has more than one neighborhood, the neighborhoods were combined into one cell and the names of the neighborhoods were separated with a comma. The result was save in a DataFrame called "toronto_df" with the columns of Postcode, Borough, and Neighborhood.

The Geospatial_Coordinates.csv was loaded into a DataFrame called "geo_coor". It has three columns: Postal Code, Latitude, and Longitude. By inner joining the DataFrame "toronto_df" with the DataFrame "geo_coor" and dropping the column of Postal Code, two new columns (Latitude and Longitude) were added to toronto_df. Since this project focuses on Toronto area, only the Boroughs that contain the word "Toronto" in the name were considered and saved in DataFrame "toronto_df2".

For each neighborhood, the latitude and longitude coordinates were used the acquire venues within a radius of 500 meters via Foursquare. The venue information was saved in the DataFrame "toronto_venues", which has 8 columns (Neighborhood, Neighborhood Latitude, Neighborhood Longitude, Venue, Venue ID, Venue Latitude, Venue Longitude,

and Venue category). For the venues of interested (such as the restaurants in the area), the Venue ID can be used to request the details of venues by Foursquare, such as price tier, rating, number of likes, number of okes, and number of dislikes for the Toronto restaurants.

3. Methodology

In this study, data visualization methods and descriptive statistic data analysis were used to explore the data of Toronto venues. Data visualization methods in this study included plotting the maps with labels and markers, using the scatter plot to show the relation between variables, plotting histogram to show the distribution, and plotting word clouds to show the frequency of occurrence.

To study the relation of variables, the correlation matrix was calculated. The Pearson Correlation coefficient with p-value was reported. For the variables showed linear relation, a linear regression model was used to calculate the slope and the intercept. The coefficient R^2 and residual plot were used to evaluate the linear regression model.

4. Results

4.1 Neighborhoods in Toronto

By scraping the table from Wikipedia page (https://en.wikipedia.org/wiki/List_ of_postal_codes_of_Canada:_M), 210 neighborhoods were obtained. After data wrangle, there were 103 unique postcodes (the “Neighborhood” cells with the same postcode were combined), and 39 of them were in Toronto area which were listed in Table 1 and shown in Figure 1.

Table 1. Neighborhoods in Toronto (DataFrame "toronto_df2")

	Postcode	Borough	Neighborhood	Latitude	Longitude
0	M5A	Downtown Toronto	Harbourfront	43.654260	-79.360636
1	M9A	Downtown Toronto	Queen's Park	43.667856	-79.532242
2	M5B	Downtown Toronto	Ryerson, Garden District	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
4	M4E	East Toronto	The Beaches	43.676357	-79.293031
5	M5E	Downtown Toronto	Berczy Park	43.644771	-79.373306
6	M5G	Downtown Toronto	Central Bay Street	43.657952	-79.387383
7	M6G	Downtown Toronto	Christie	43.669542	-79.422564
8	M5H	Downtown Toronto	Adelaide, King, Richmond	43.650571	-79.384568
9	M6H	West Toronto	Dovercourt Village, Dufferin	43.669005	-79.442259
10	M5J	Downtown Toronto	Harbourfront East, Toronto Islands, Union Station	43.640816	-79.381752
11	M6J	West Toronto	Little Portugal, Trinity	43.647927	-79.419750
12	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188
13	M5K	Downtown Toronto	Design Exchange, Toronto Dominion Centre	43.647177	-79.381576
14	M6K	West Toronto	Brockton, Exhibition Place, Parkdale Village	43.636847	-79.428191
15	M4L	East Toronto	The Beaches West, India Bazaar	43.668999	-79.315572
16	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817
17	M4M	East Toronto	Studio District	43.659526	-79.340923
18	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
19	M5N	Central Toronto	Roselawn	43.711695	-79.416936
20	M4P	Central Toronto	Davisville North	43.712751	-79.390197
21	M5P	Central Toronto	Forest Hill North, Forest Hill West	43.696948	-79.411307
22	M6P	West Toronto	High Park, The Junction South	43.661608	-79.464763
23	M4R	Central Toronto	North Toronto West	43.715383	-79.405678
24	M5R	Central Toronto	The Annex, North Midtown, Yorkville	43.672710	-79.405678
25	M6R	West Toronto	Parkdale, Roncesvalles	43.648960	-79.456325
26	M4S	Central Toronto	Davisville	43.704324	-79.388790
27	M5S	Downtown Toronto	Harbord, University of Toronto	43.662696	-79.400049
28	M6S	West Toronto	Runnymede, Swansea	43.651571	-79.484450
29	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160
30	M5T	Downtown Toronto	Chinatown, Grange Park, Kensington Market	43.653206	-79.400049
31	M4V	Central Toronto	Deer Park, Forest Hill SE, Rathnelly, South Hi...	43.686412	-79.400049
32	M5V	Downtown Toronto	CN Tower, Bathurst Quay, Island airport, Harbo...	43.628947	-79.394420
33	M4W	Downtown Toronto	Rosedale	43.679563	-79.377529
34	M5W	Downtown Toronto	Stn A PO Boxes 25 The Esplanade	43.646435	-79.374846
35	M4X	Downtown Toronto	Cabbagetown, St. James Town	43.667967	-79.367675
36	M5X	Downtown Toronto	First Canadian Place, Underground city	43.648429	-79.382280
37	M4Y	Downtown Toronto	Church and Wellesley	43.665860	-79.383160
38	M7Y	East Toronto	Business Reply Mail Processing Centre 969 Eastern	43.662744	-79.321558

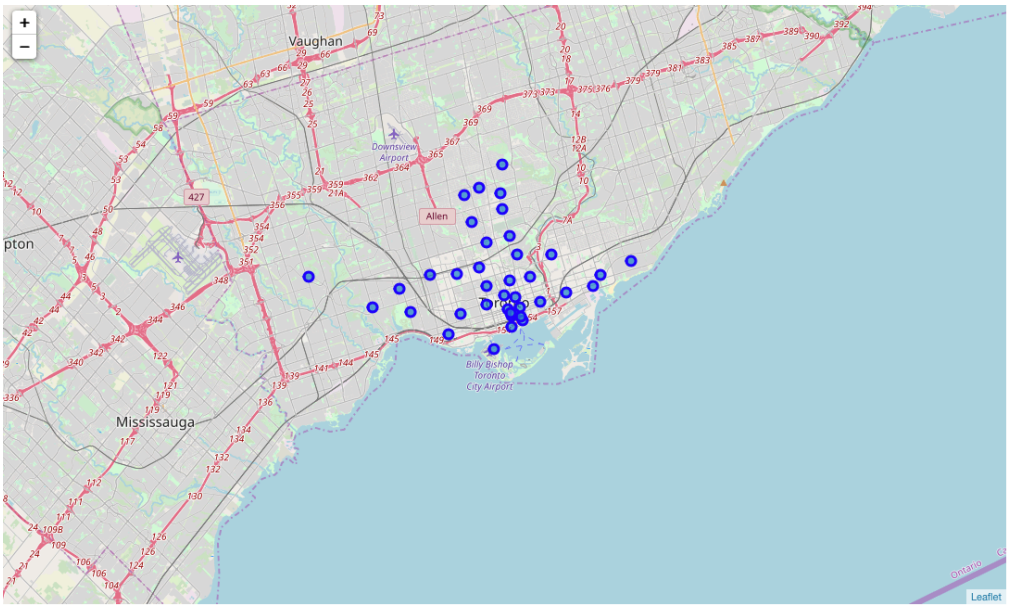


Figure 1. Map of Toronto with Neighborhoods Marked

4.2 Venues in Toronto

To explore the venues around the 39 neighborhoods in Toronto, venues explore requests were submitted via Foursquare to fetch the location data of the venues within 500 meters from the coordinates of each neighborhoods in Table 1. Due to the limitation of the Foursquare account, a maximum limit of venues in each neighborhood was set to 100. There were 1670 venues obtained and saved in DataFrame toronto_venues. Table 2 shows the first 5 rows in this DataFrame. The neighborhood Queen's Park with a postcode of M9A did not have any venues. Therefore, only 38 neighborhoods were studied in this study.

Table 2. First 5 Rows in DataFrame toronto_venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category
0	Harbourfront	43.65426	-79.360636	Roselle Desserts	54ea41ad498e9a11e9e13308	43.653447	-79.362017	Bakery
1	Harbourfront	43.65426	-79.360636	Tandem Coffee	53b8466a498e83df908c3f21	43.653559	-79.361809	Coffee Shop
2	Harbourfront	43.65426	-79.360636	Cooper Koo Family YMCA	574c229e498ebb5c6b257902	43.653191	-79.357947	Gym / Fitness Center
3	Harbourfront	43.65426	-79.360636	Body Blitz Spa East	50760559e4b0e8c7babe2497	43.654735	-79.359874	Spa
4	Harbourfront	43.65426	-79.360636	Morning Glory Cafe	4ae5b91ff964a520a6a121e3	43.653947	-79.361149	Breakfast Spot

Among the 1670 venues, there were 232 unique venue categories. The venues that have the word "restaurant" in the "Venue Category" columns were considered the Toronto

restaurants in this study. There were a total of 398 restaurants in Toronto.

4.3 Total Number of Venues vs Total Number of Restaurants in Each Neighborhood

Figure 2 showed the distribution of the venues. The red color denoted the neighborhood with a larger number of venues and the blue color denoted the neighborhood with a smaller number of venues.

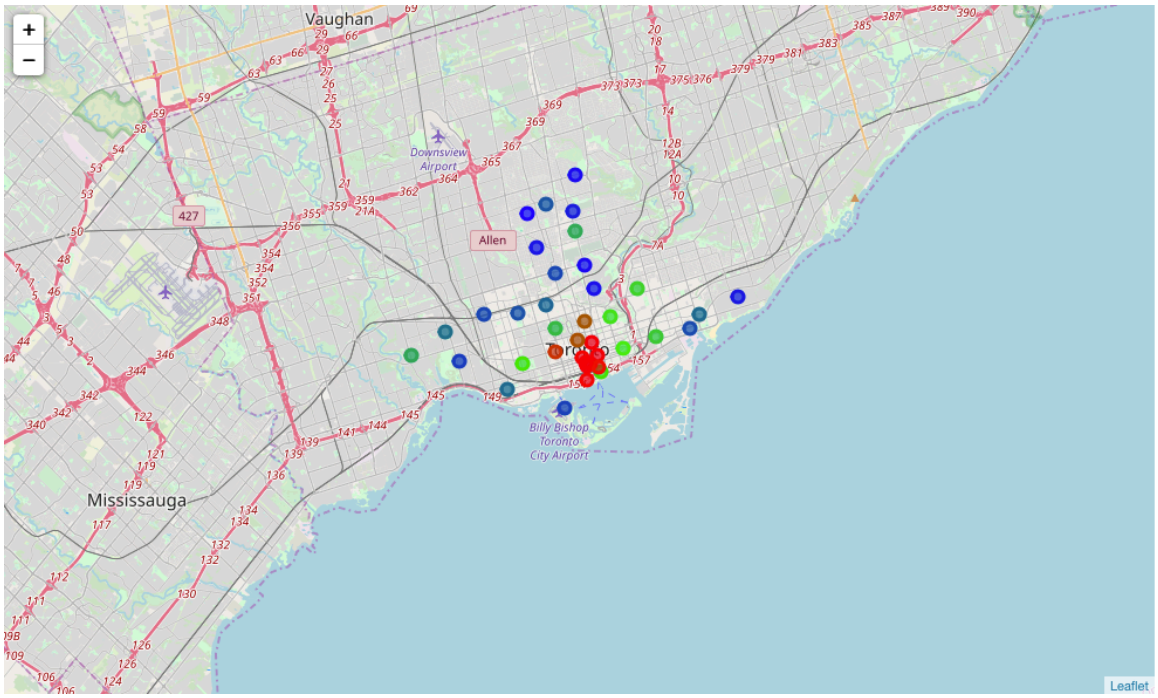


Figure 2. Total Number of Venues in Each Neighborhood

Figure 3 showed the distribution of the restaurants. The red color represented the neighborhood with a larger number of restaurants and the blue color represents the neighborhood with a smaller number of restaurants.

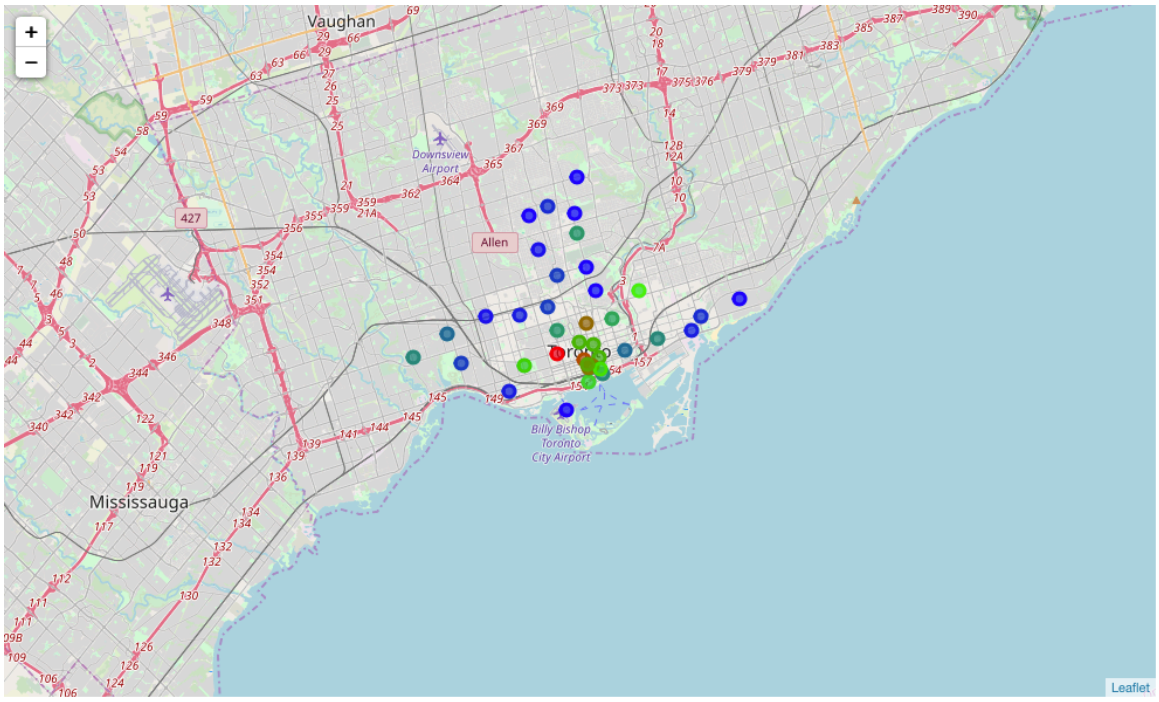


Figure 3. Total number of Restaurants in Each Neighborhood

The total number of venues and the total number of restaurants were shown in Figure 4. In general, the neighborhood with a larger total number of venues also had more restaurants.

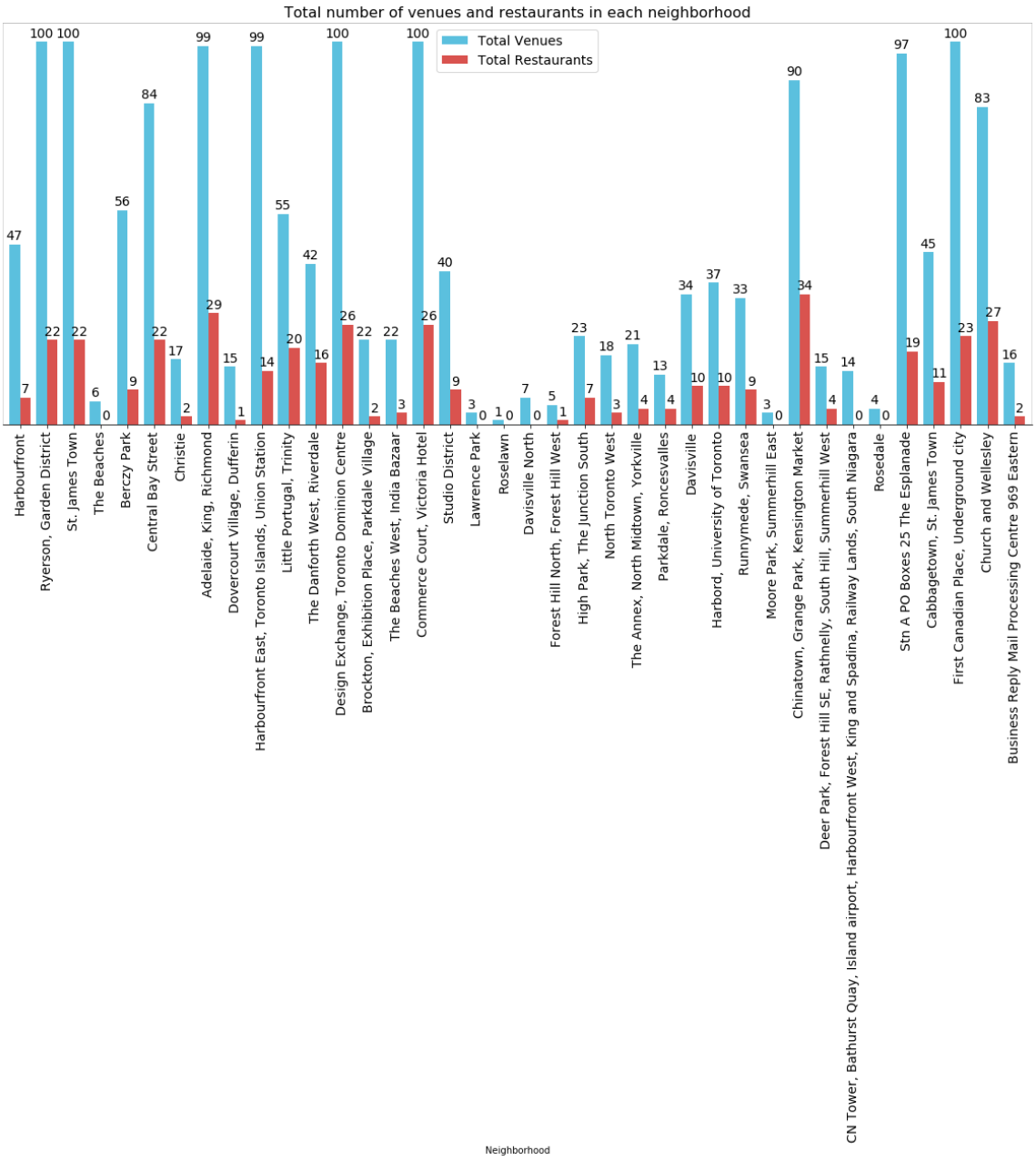


Figure 4. Number of Venues vs Number of Restaurants

The financial district of Toronto has a high concentration of banks and brokerage firms on Bay Street. We were wondering if a neighborhood with a high density of business firms would have more restaurants in percentage of the total venues in the same neighborhood. To answer this question, the "Percent Restaurant" was calculated as the ratio of "Total Restaurants" to "Total Venues" in each corresponding neighborhood (shown in the last column of Table 3). It represented how many percentages of the venues

in each neighborhood were restaurants.

Table 3. Number of Venues vs Nubmer of Restaurants

	Neighborhood	Total Venues	Total Restaurants	Percent Restaurant
0	Adelaide, King, Richmond	99	29	0.292929
1	Berczy Park	56	9	0.160714
2	Brockton, Exhibition Place, Parkdale Village	22	2	0.090909
3	Business Reply Mail Processing Centre 969 Eastern	16	2	0.125000
4	CN Tower, Bathurst Quay, Island airport, Harbo...	14	0	0.000000
5	Cabbagetown, St. James Town	45	11	0.244444
6	Central Bay Street	84	22	0.261905
7	Chinatown, Grange Park, Kensington Market	90	34	0.377778
8	Christie	17	2	0.117647
9	Church and Wellesley	83	27	0.325301
10	Commerce Court, Victoria Hotel	100	26	0.260000
11	Davisville	34	10	0.294118
12	Davisville North	7	0	0.000000
13	Deer Park, Forest Hill SE, Rathnelly, South Hi...	15	4	0.266667
14	Design Exchange, Toronto Dominion Centre	100	26	0.260000
15	Dovercourt Village, Dufferin	15	1	0.066667
16	First Canadian Place, Underground city	100	23	0.230000
17	Forest Hill North, Forest Hill West	5	1	0.200000
18	Harbord, University of Toronto	37	10	0.270270
19	Harbourfront	47	7	0.148936
20	Harbourfront East, Toronto Islands, Union Station	99	14	0.141414
21	High Park, The Junction South	23	7	0.304348
22	Lawrence Park	3	0	0.000000
23	Little Portugal, Trinity	55	20	0.363636
24	Moore Park, Summerhill East	3	0	0.000000
25	North Toronto West	18	3	0.166667
26	Parkdale, Roncesvalles	13	4	0.307692
27	Rosedale	4	0	0.000000
28	Roselawn	1	0	0.000000
29	Runnymede, Swansea	33	9	0.272727
30	Ryerson, Garden District	100	22	0.220000
31	St. James Town	100	22	0.220000
32	Stn A PO Boxes 25 The Esplanade	97	19	0.195876
33	Studio District	40	9	0.225000
34	The Annex, North Midtown, Yorkville	21	4	0.190476
35	The Beaches	6	0	0.000000
36	The Beaches West, India Bazaar	22	3	0.136364
37	The Danforth West, Riverdale	42	16	0.380952

The top 10 neighborhoods with the largest number of venues and restaurants were listed in Table 4 (a) and (b) respectively.

Table 4. Top 10 Neighborhoods with the Largest Number of Venues (a) and Restaurants (b)

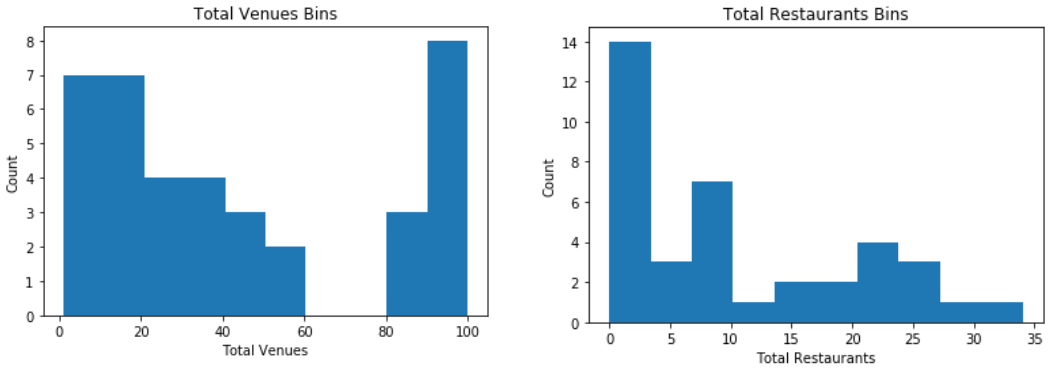
	Neighborhood	Total Venues
14	Design Exchange, Toronto Dominion Centre	100
31	St. James Town	100
30	Ryerson, Garden District	100
10	Commerce Court, Victoria Hotel	100
16	First Canadian Place, Underground city	100
0	Adelaide, King, Richmond	99
20	Harbourfront East, Toronto Islands, Union Station	99
32	Stn A PO Boxes 25 The Esplanade	97
7	Chinatown, Grange Park, Kensington Market	90
6	Central Bay Street	84

(a)

	Neighborhood	Total Restaurants
7	Chinatown, Grange Park, Kensington Market	34
0	Adelaide, King, Richmond	29
9	Church and Wellesley	27
14	Design Exchange, Toronto Dominion Centre	26
10	Commerce Court, Victoria Hotel	26
16	First Canadian Place, Underground city	23
31	St. James Town	22
30	Ryerson, Garden District	22
6	Central Bay Street	22
23	Little Portugal, Trinity	20

(b)

The histogram for the total number of venues and the total number of restaurants were shown in Figure 5.



(a) (b)

Figure 5. Histogram for the total number of venues (a) and the total number of restaurants (b)

4.4 Relation Between Total Number of Venues and Total Number of Restaurants

The results from Table 3 were plotted in Figure 6 showing the relation between "Total Venues" and "Total Restaurants" (a), between "Total Venues" and "Percent Restaurants"

(b), and between “Total Restaurants” and “Percent Restaurants” (c).

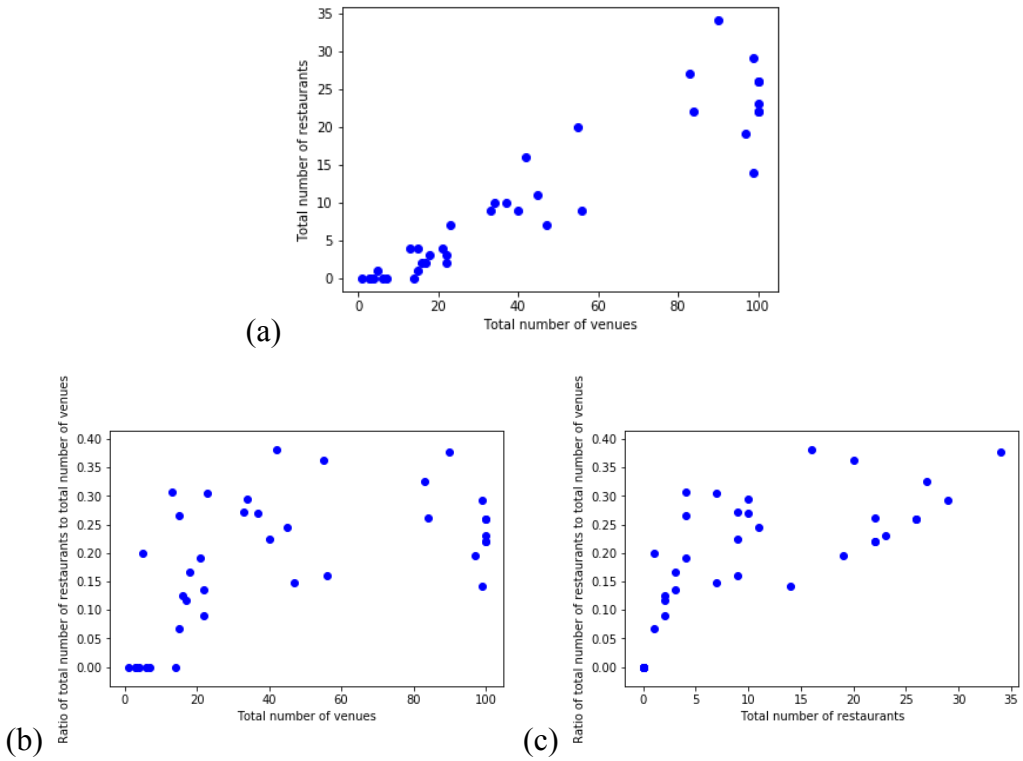


Figure 6. Relation between Total Number of Venues, Total Number of Restaurants, and Ratio of Total Restaurants to Total Venues

From Figure 6, it was observed that there was a positive correlation between the total number of venues and the total number of restaurants. It was also observed that some neighborhoods did reach the maximum limit of the total venues requested which was 100. To eliminate the effect of this limitation, the neighborhoods with 100 total venues were removed when the correlation was calculated. Note that these neighborhoods were kept in this study for the rest analysis.

The correlation matrix was shown in Table 5. From Table 5, it was observed that there was a strong positive correlation between Total Venues and Total Restaurants. The Pearson Correlation coefficient is 0.907626 with a p-value of less than 0.001.

Table 5. Correlation Matrix of Total Venues, Total Restaurants, and Percent Restaurants

	Total Venues	Total Restaurants	Percent Restaurant
Total Venues	1.000000	0.907626	0.564633
Total Restaurants	0.907626	1.000000	0.733205
Percent Restaurant	0.564633	0.733205	1.000000

Figure 7 showed the linear regression plot of these two variables. The slope was 0.27 and the intercept was -1.19. The coefficient R^2 was 0.8238.

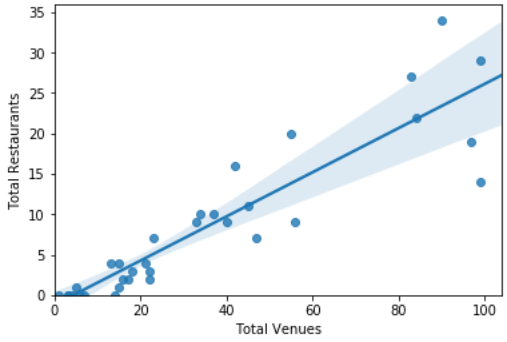


Figure 7. Linear Regression Plot of Total Venues and Total Restaurants

The residuals of the linear regression were shown in Figure 8. As can be seen from Figure 8, the residuals were evenly distributed. Based on the previous analysis, a linear regression model was a good fit to describe the relation between Total Venues and Total Restaurants.

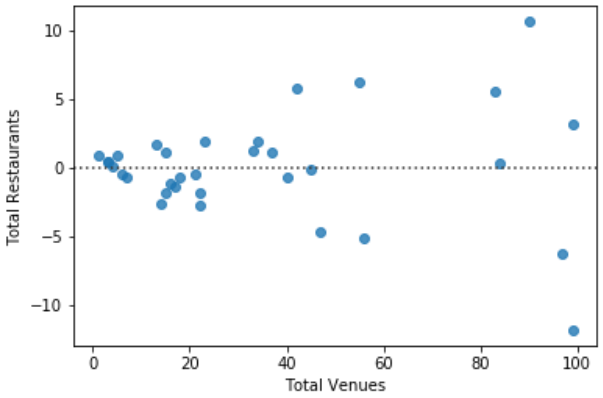


Figure 8. Residual Plot

4.5 Frequency of Occurrence

Figure 9 showed the word clouds of "Venue Category" in DataFrame toronto_venues. As can be seen, the category "Coffee Shop" and "Café" were very popular.

Figure 9. Word Clouds for Venue Category of Toronto Venues

The 10 most common venues in the top 10 neighborhoods that sorted by the total number of venues were listed in Table 6(a). The 10 most common venues in the top 10 neighborhoods that sorted by the total number of restaurants were listed in Table 6(b).

Table 6. 10 Most Common Venues in the top 10 Neighborhoods

(a) Sorted by the total number of venues

	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	10th Most Common Venue
14	Design Exchange, Toronto Dominion Centre	Coffee Shop	Hotel	Café	Italian Restaurant	Restaurant	American Restaurant	Bar	Bakery	Deli / Bodega	Gastropub
31	St. James Town	Café	Coffee Shop	Restaurant	Hotel	Breakfast Spot	Cocktail Bar	Beer Bar	Cosmetics Shop	Bakery	Diner
30	Ryerson, Garden District	Coffee Shop	Clothing Store	Café	Cosmetics Shop	Middle Eastern Restaurant	Pizza Place	Fast Food Restaurant	Plaza	Bookstore	Sporting Goods Shop
10	Commerce Court, Victoria Hotel	Coffee Shop	Café	Hotel	Restaurant	Italian Restaurant	American Restaurant	Steakhouse	Gym	Seafood Restaurant	Deli / Bodega
16	First Canadian Place, Underground city	Coffee Shop	Café	Hotel	Restaurant	Gym	Steakhouse	Burger Joint	Seafood Restaurant	Deli / Bodega	Asian Restaurant
0	Adelaide, King, Richmond	Coffee Shop	Café	Bar	Steakhouse	Burger Joint	Restaurant	Sushi Restaurant	Asian Restaurant	Thai Restaurant	Salad Place
20	Harbourfront East, Toronto Islands, Union Station	Coffee Shop	Aquarium	Hotel	Italian Restaurant	Café	Scenic Lookout	Brewery	Fried Chicken Joint	Restaurant	History Museum
32	Stn A PO Boxes 25 The Esplanade	Coffee Shop	Restaurant	Café	Japanese Restaurant	Seafood Restaurant	Hotel	Beer Bar	Cocktail Bar	Sandwich Place	Breakfast Spot
7	Chinatown, Grange Park, Kensington Market	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Café	Bar	Coffee Shop	Dumpling Restaurant	Chinese Restaurant	Bakery	Mexican Restaurant	Cocktail Bar
6	Central Bay Street	Coffee Shop	Italian Restaurant	Café	Sandwich Place	Ice Cream Shop	Burger Joint	Middle Eastern Restaurant	Juice Bar	Japanese Restaurant	Salad Place

(b) Sorted by the total number of restaurants

	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	10th Most Common Venue
7	Chinatown, Grange Park, Kensington Market	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Café	Bar	Coffee Shop	Dumpling Restaurant	Chinese Restaurant	Bakery	Mexican Restaurant	Cocktail Bar
0	Adelaide, King, Richmond	Coffee Shop	Café	Bar	Steakhouse	Burger Joint	Restaurant	Sushi Restaurant	Asian Restaurant	Thai Restaurant	Salad Place
9	Church and Wellesley	Coffee Shop	Japanese Restaurant	Gay Bar	Sushi Restaurant	Restaurant	Mediterranean Restaurant	Gym	Hotel	Gastropub	Café
14	Design Exchange, Toronto Dominion Centre	Coffee Shop	Hotel	Café	Italian Restaurant	Restaurant	American Restaurant	Bar	Bakery	Deli / Bodega	Gastropub
10	Commerce Court, Victoria Hotel	Coffee Shop	Café	Hotel	Restaurant	Italian Restaurant	American Restaurant	Steakhouse	Gym	Seafood Restaurant	Deli / Bodega
16	First Canadian Place, Underground city	Coffee Shop	Café	Hotel	Restaurant	Gym	Steakhouse	Burger Joint	Seafood Restaurant	Deli / Bodega	Asian Restaurant
31	St. James Town	Café	Coffee Shop	Restaurant	Hotel	Breakfast Spot	Cocktail Bar	Beer Bar	Cosmetics Shop	Bakery	Diner
30	Ryerson, Garden District	Coffee Shop	Clothing Store	Café	Cosmetics Shop	Middle Eastern Restaurant	Pizza Place	Fast Food Restaurant	Plaza	Bookstore	Sporting Goods Shop
6	Central Bay Street	Coffee Shop	Italian Restaurant	Café	Sandwich Place	Ice Cream Shop	Burger Joint	Middle Eastern Restaurant	Juice Bar	Japanese Restaurant	Salad Place
23	Little Portugal, Trinity	Bar	Asian Restaurant	Restaurant	Pizza Place	Men's Store	New American Restaurant	Coffee Shop	Café	Vegetarian / Vegan Restaurant	Vietnamese Restaurant

The word clouds for the "Venue Category" of Toronto restaurants were shown in Figure 10. It was observed that the "Restaurant" and "Italian Restaurant" were the top 2 categories.

Figure 10. Word Clouds for Venue Category of Toronto Restaurants

Table 7 listed the number of each category in a descending order. It was consisted with what was found from word clouds. The Italian Restaurant was the most popular cuisine in Toronto.

Table 7. The Number of Each Restaurant Category

Restaurant	53	Gluten-free Restaurant	4
Italian Restaurant	44	Modern European Restaurant	3
Japanese Restaurant	26	Falafel Restaurant	2
Seafood Restaurant	23	Ethiopian Restaurant	2
American Restaurant	21	Cuban Restaurant	2
Vegetarian / Vegan Restaurant	21	Brazilian Restaurant	2
Sushi Restaurant	20	Korean Restaurant	2
Thai Restaurant	18	Portuguese Restaurant	2
Greek Restaurant	15	Hotpot Restaurant	1
Asian Restaurant	14	Colombian Restaurant	1
Chinese Restaurant	12	Molecular Gastronomy Restaurant	1
Mexican Restaurant	11	Cajun / Creole Restaurant	1
French Restaurant	10	Malay Restaurant	1
Vietnamese Restaurant	10	Eastern European Restaurant	1
Middle Eastern Restaurant	9	Southern / Soul Food Restaurant	1
Fast Food Restaurant	9	Afghan Restaurant	1
New American Restaurant	8	Belgian Restaurant	1
Indian Restaurant	8	Theme Restaurant	1
Comfort Food Restaurant	7	Filipino Restaurant	1
Ramen Restaurant	6	German Restaurant	1
Latin American Restaurant	5	Doner Restaurant	1
Caribbean Restaurant	5	Taiwanese Restaurant	1
Dumpling Restaurant	4	Dim Sum Restaurant	1
Mediterranean Restaurant	4	Empanada Restaurant	1

4.6 Toronto Restaurants

For each Toronto restaurants, the Venue ID was used to request the detail information via Foursquare. The price tier, the rating, the number of likes, the number of okes, and the number of dislikes were collected. For all the restaurants in this study, the number of okes and the number of dislikes were "False". Since it did not provide any meaningful information, they were removed. The first 5 rows of the DataFrame toronto_restaurants were shown in Table 8. For the price tier, level 1, 2, 3, and 4 represented cheap, moderate, expensive, and very expensive, respectively.

Table 8. The First 5 Rows of DataFrame toronto_restaurants

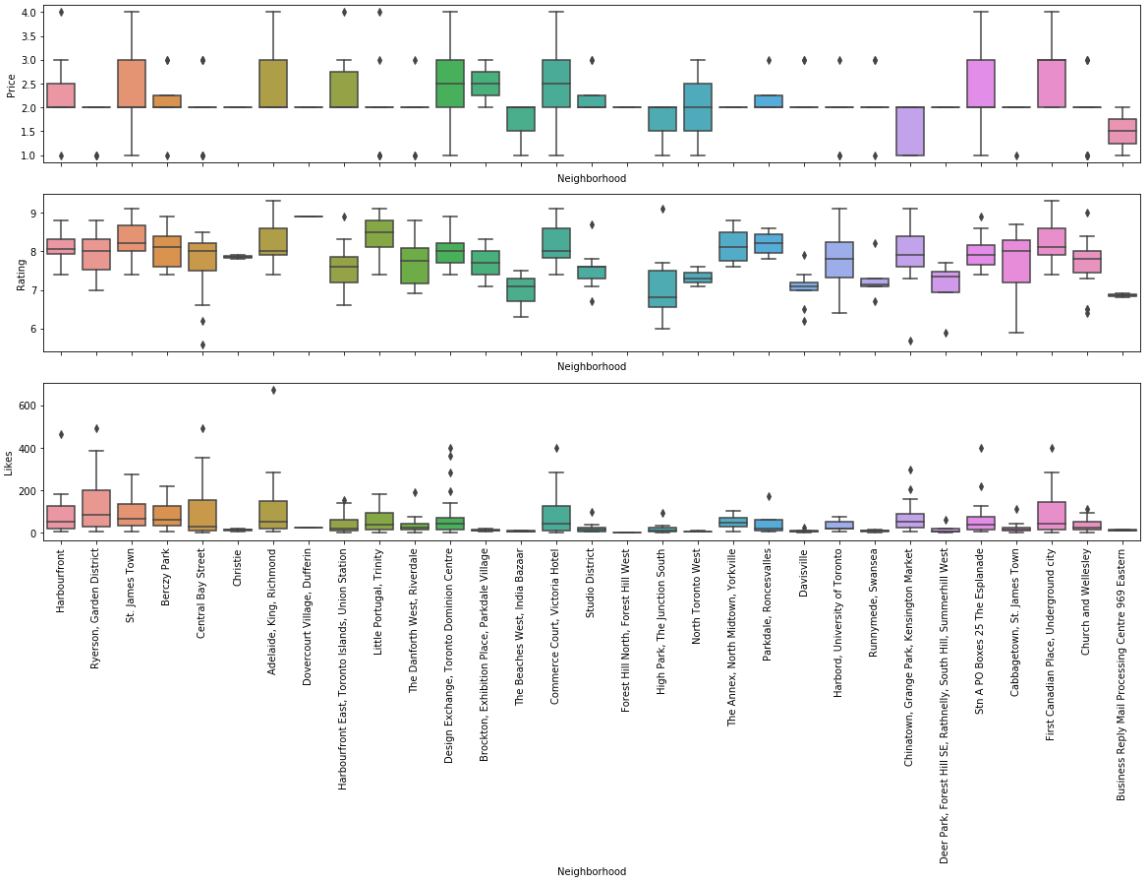
	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue ID	Venue Latitude	Venue Longitude	Venue Category	Price	Rating	Likes
0	Harbourfront	43.65426	-79.360636	Impact Kitchen	5612b1cc498e3dd742af0dc8	43.656369	-79.356980	Restaurant	2.0	8.8	51.0
1	Harbourfront	43.65426	-79.360636	Cluny Bistro & Boulangerie	53a22c92498ec91fda7ce133	43.650565	-79.357843	French Restaurant	4.0	8.4	183.0
2	Harbourfront	43.65426	-79.360636	El Catrin	51ddecee498e1ffd34185d2f	43.650601	-79.358920	Mexican Restaurant	3.0	8.1	465.0
3	Harbourfront	43.65426	-79.360636	Cocina Economica	5542ab36498e2f92a8c248f2	43.654959	-79.365657	Mexican Restaurant	1.0	7.9	67.0
4	Harbourfront	43.65426	-79.360636	Izumi	50e88fc6e4b007fcbb57aae8	43.649970	-79.360153	Asian Restaurant	2.0	8.0	18.0

The descriptive statistic results were shown in Table 9. Figure 11 showed the results

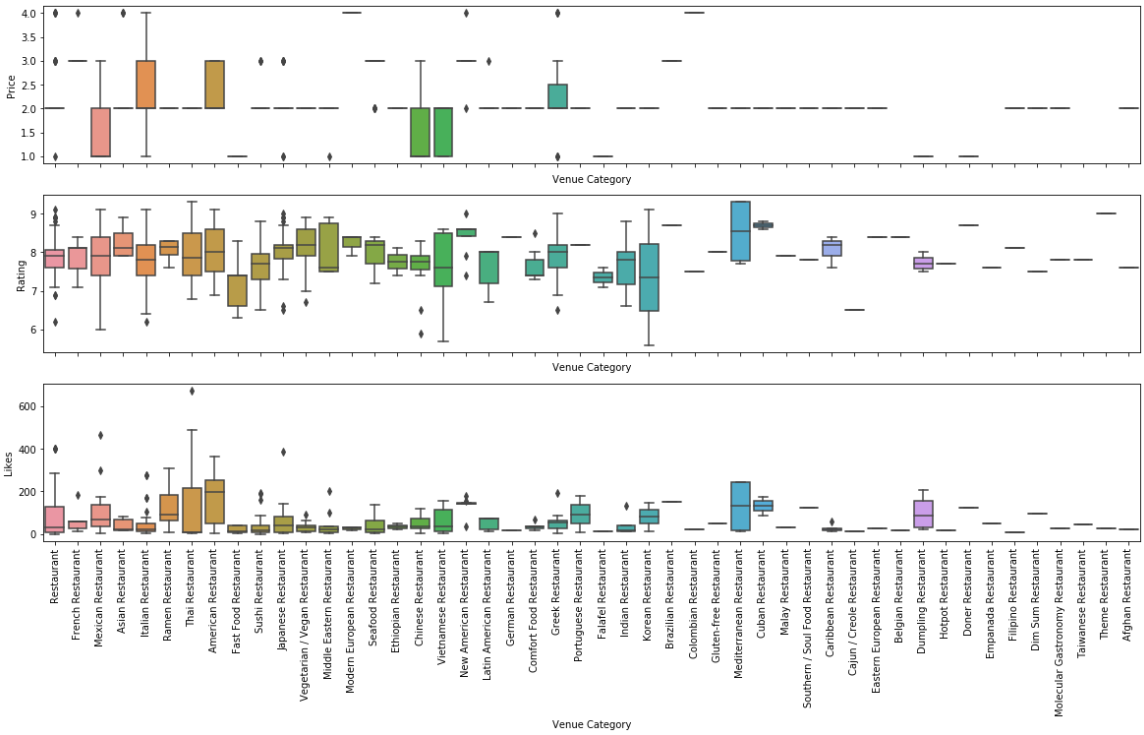
grouped by neighborhood (a) and by category (b).

Table 9. Descriptive Statistic Results of DataFrame toronto_restaurants

	Neighborhood Latitude	Neighborhood Longitude	Venue Latitude	Venue Longitude	Price	Rating	Likes
count	398.000000	398.000000	398.000000	398.000000	382.000000	392.000000	398.000000
mean	43.656581	-79.387467	43.656885	-79.387648	2.185864	7.901786	69.309045
std	0.013319	0.026194	0.013284	0.026088	0.690779	0.645402	93.510610
min	43.636847	-79.484450	43.636951	-79.487067	1.000000	5.600000	0.000000
25%	43.648198	-79.400049	43.648571	-79.394850	2.000000	7.575000	12.000000
50%	43.651571	-79.382280	43.651813	-79.381784	2.000000	7.900000	32.000000
75%	43.661608	-79.378937	43.662446	-79.378326	2.000000	8.300000	78.000000
max	43.715383	-79.315572	43.716785	-79.315081	4.000000	9.300000	672.000000



(a)



(b)

Figure 11. Descriptive Statistic Results of DataFrame toronto_restaurants

Figure 12 showed the average of the price tier, the rating, and the number of likes grouped by neighbourhood and venue category. Due to lack of data (sparse matrix), a lot of the cells in this figure were in red color (denoted a value of 0).

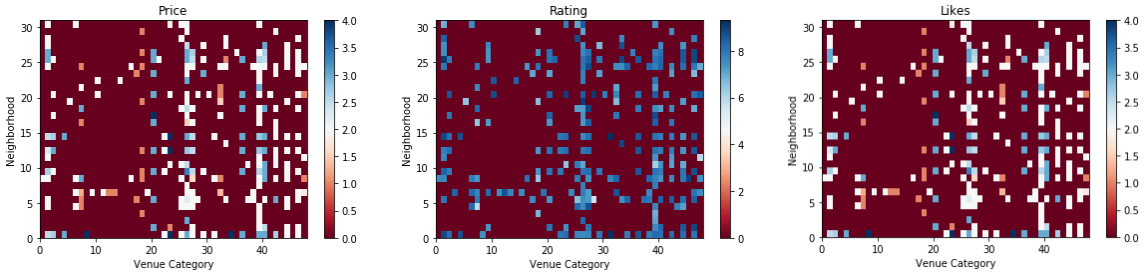


Figure 12. Average of Price Tier, Rating, and Number of Likes Grouped by Neighborhood and Venue Category

For all the 398 Toronto restaurants in this study, 43 of them were in price tier 1, 244 of them were in price tier 2, 76 of them were in price tier 3, 19 of them were in price tier 4, and 16 of them did not have price information available. As can be seen, the price tier 2 (moderate) was the most popular price tier.

4.7 Italian Restaurants

There were a total of 44 Italian restaurants were in this study. They were located in 23 different neighborhoods in Toronto. Table 10 showed the average price, average number of likes, average rating, minimal rating, and maximal rating of each neighborhood that was sorted by the average rating.

Table 10. Average Price, Average Likes, Average Rating, Minimal Rating, and Maximal Rating of Italian Restaurants in Each Neighborhood Sorted by Average Rating

	Neighborhood	Ave Price	Ave Likes	Ave Rating	Min Rating	Max Rating
21	The Beaches West, India Bazaar	2.000000	10.000000	7.100000	7.1	7.1
7	Davisville	2.500000	8.500000	7.100000	7.1	7.1
12	Harbourfront East, Toronto Islands, Union Station	2.500000	26.750000	7.250000	6.6	7.9
10	Harbord, University of Toronto	1.500000	38.000000	7.350000	6.4	8.3
11	Harbourfront	2.000000	20.000000	7.400000	7.4	7.4
20	Studio District	2.500000	23.000000	7.450000	7.1	7.8
3	Central Bay Street	2.250000	64.000000	7.550000	6.2	8.2
9	First Canadian Place, Underground city	3.000000	45.000000	7.600000	7.6	7.6
0	Adelaide, King, Richmond	3.000000	45.000000	7.600000	7.6	7.6
2	Cabbagetown, St. James Town	2.000000	24.500000	7.650000	7.1	8.2
19	Stn A PO Boxes 25 The Esplanade	3.000000	14.000000	7.650000	7.4	7.9
16	Runnymede, Swansea	2.000000	9.000000	7.750000	7.3	8.2
5	Christie	2.000000	6.000000	7.800000	7.8	7.8
15	Parkdale, Roncesvalles	3.000000	6.000000	7.800000	7.8	7.8
4	Chinatown, Grange Park, Kensington Market	2.000000	71.000000	7.900000	7.9	7.9
8	Design Exchange, Toronto Dominion Centre	2.750000	41.000000	7.975000	7.6	8.5
17	Ryerson, Garden District	2.000000	138.500000	8.050000	8.0	8.1
6	Commerce Court, Victoria Hotel	3.333333	114.000000	8.066667	7.6	8.7
18	St. James Town	3.000000	129.666667	8.100000	7.6	8.7
1	Brockton, Exhibition Place, Parkdale Village	3.000000	21.000000	8.300000	8.3	8.3
14	Little Portugal, Trinity	2.000000	20.000000	8.500000	8.5	8.5
22	The Danforth West, Riverdale	2.000000	24.000000	8.550000	8.3	8.8
13	High Park, The Junction South	2.000000	35.000000	9.100000	9.1	9.1

For the 44 Italian restaurants, one of them was in price tier 1, 26 in price tier 2, 13 in price tier 3, and 4 in price tier 4. The price tier 2 (moderate) was the most popular price

tier for Italian restaurants, which was the same as the observation made for Toronto restaurants in Session 4.6.

For the 26 Italian restaurants in the price tier 2, they were located in 17 different neighborhoods. The detail information was summarized in Table 11. In this table, "Number P2 Italian" represented the number of Italian restaurants in the price tier 2 in each neighborhood, and "Percent P2 Italian Restaurant" was the ratio of "Number P2 Italian" to "Total Venues".

Table 11. Neighborhoods with Price Tier 2 Italian Restaurants

	Neighborhood	Ave Price	Ave Likes	Ave Rating	Min Rating	Max Rating	Number of P2 Italian	Total Venues	Total Restaurants	Percent Restaurant	Italian Restaurant	Percent P2 Italian Restaurant
0	Cabbagetown, St. James Town	2.0	24.5	7.650000	7.1	8.2	2	45	11	0.244444	2	0.044444
1	Central Bay Street	2.0	67.0	7.333333	6.2	8.1	3	84	22	0.261905	4	0.035714
2	Chinatown, Grange Park, Kensington Market	2.0	71.0	7.900000	7.9	7.9	1	90	34	0.377778	1	0.011111
3	Christie	2.0	6.0	7.800000	7.8	7.8	1	17	2	0.117647	1	0.058824
4	Davisville	2.0	10.0	7.100000	7.1	7.1	1	34	10	0.294118	2	0.029412
5	Design Exchange, Toronto Dominion Centre	2.0	42.0	7.900000	7.7	8.1	2	100	26	0.260000	4	0.020000
6	Harbord, University of Toronto	2.0	58.0	8.300000	8.3	8.3	1	37	10	0.270270	2	0.027027
7	Harbourfront	2.0	20.0	7.400000	7.4	7.4	1	47	7	0.148936	1	0.021277
8	Harbourfront East, Toronto Islands, Union Station	2.0	28.0	7.466667	6.8	7.9	3	99	14	0.141414	4	0.030303
9	High Park, The Junction South	2.0	35.0	9.100000	9.1	9.1	1	23	7	0.304348	1	0.043478
10	Little Portugal, Trinity	2.0	20.0	8.500000	8.5	8.5	1	55	20	0.363636	1	0.018182
11	Runnymede, Swansea	2.0	9.0	7.750000	7.3	8.2	2	33	9	0.272727	2	0.060606
12	Ryerson, Garden District	2.0	138.5	8.050000	8.0	8.1	2	100	22	0.220000	2	0.020000
13	Stn A PO Boxes 25 The Esplanade	2.0	6.0	7.400000	7.4	7.4	1	97	19	0.195876	2	0.010309
14	Studio District	2.0	37.0	7.100000	7.1	7.1	1	40	9	0.225000	2	0.025000
15	The Beaches West, India Bazaar	2.0	10.0	7.100000	7.1	7.1	1	22	3	0.136364	1	0.045455
16	The Danforth West, Riverdale	2.0	24.0	8.550000	8.3	8.8	2	42	16	0.380952	2	0.047619

5. Discussion

In this study, the questions listed in the Introduction session were answered as follows:

(1) The most popular neighborhoods (with the most venues) in Toronto were #10 Commerce Court, Victoria Hotel, #14 Design Exchange, Toronto Dominion Centre, #16 First Canadian Place, Underground city, #30 Ryerson, Garden District, and #31 St. James Town. Noted that all of these neighborhoods had a total number of 100 venues, which was the maximal limit for the Foursquare

request. Therefore, the results might potentially change without such a limitation constraint.

(2) The neighborhood has the largest number of restaurants was #7 Chinatown, Grange Park, Kensington Market. There were a total of 34 restaurants in this neighborhood, followed by 29 restaurants in #0 Adelaide, King, Richmond, and 27 restaurants in #9 Church and Wellesley. The histogram in Figure 5 showed that 14 neighborhoods had only 0-3 restaurants.

(3) There was a strong positive correlation between the number of venues and the number of restaurants in each neighborhood. The Pearson Correlation coefficient was 0.907626 with a p-value of less than 0.001. This relation could also be observed by comparing Figure 2 and Figure 3, as well as from Figure 4.

Furthermore, the number of venues and the number of restaurants in each neighborhood had a linear relationship, fitted in a linear regression model with a slope of 0.27 and an intercept of -1.19. The coefficient R^2 and residual plot were used to verify that it was a good fit by using the linear regression model.

Since there was a strong positive correlation between the number of venues and the number of restaurants and the correlation with the percentage of restaurants was not strong, the neighborhoods with more venues and lower percentage of restaurants will be a good choice. The neighborhoods with more venues will bring more traffic and potential customers, which will support more restaurants. With a lower percentage of restaurants, the restaurant market in these neighborhoods is not saturated. Therefore, the restaurants in these neighborhoods are more likely to survive.

(4) Due to its diverse population, it would be useful to learn what is the most popular cuisine in Toronto. In this studied, the most popular cuisine was the Italian restaurant. There were 44 Italian restaurants in Toronto, followed by 26 Japanese

restaurants and 23 seafood restaurants.

However, this was only a preliminary study. Further investigation might be needed by carefully examing the venue categories. In this study, it was assumed that the restaurants would have the word “restaurant” in the venue category, which did not include the venues such as “Steakhouse” and “Pizza Place”. Moreover, some categories might need to be combined, such as “Japanese Restaurant” and “Sushi Restaurant”, as well as “Dim Sum Restaurant”, “Dumpling Restaurant”, and “Asian Restaurant”. This will require further information on the details of each restaurant. No decision could be made based on the current information.

(5) In this study, the data was also grouped by the neighborhood and the venue category in Session 4.6. Due to the limited number of samples in this study (sparse matrix), no conclusion can be made for the effect of different factors on the price tier, the rating, and the number of likes. No inferential statistic data analysis was performed.

The recommendations to the types of restaurants should be invested includes considering the most popular Italian restaurants especially with a moderate price (price tier 2), considering the neighborhoods with more venues (more potential customers) and less percentage of restaurants (the market is not saturated), and avoiding the neighborhoods with a similar restaurant (to the one planning to invest) which has a high rating (avoid the high competition). If the neighborhood chosen by the investors does not have any similar restaurants, the possible reasons need to be found. For example, certain cuisine might not be welcomed due to the cultural and religion background of the residents or the people associated with that neighborhood. If that is the case, this neighborhood should not be considered.

6. Conclusion

In this study, the venues in Toronto especially the Toronto restaurants were investigated. It was found that there was a linear relationship between the number of venues and the number of restaurants in each neighborhood, Italian restaurants were the most popular cuisine in Toronto, and price tier 2 (moderate) was the most popular price range. Moreover, the descriptive statistic was performed for the price tier 2 Italian restaurants in each neighborhood. Table 11 summarized the total number of venues, the total number of restaurants, the total number of Italian restaurants, the total number of price tier 2 Italian restaurants, percentage of restaurants, percentage of price tier 2 Italian restaurants, average price, average number of likes, average rating, minimal rating, and maximal rating to provide the investors some insights for each neighborhood.

The recommendations to the investors and the limitation of this study were discussed in the Discussion session. In the future work, there were other factors need to be considered once further information of the investors (such as budget, previous experience and expertise in certain fields) is available. Also, the decisions need to be made based on the current available locations. For example, there might not be a venue available to be rent or purchased in the ideal neighborhood, or it may require a bid. The rent and other cost (such as renovation fees) as well as all the pros and cons of each potential location need to be weighted.

References

[1] www.toronto.ca, Accessed on 12/14/2019

[2] https://en.wikipedia.org/wiki/Toronto, Accessed on 12/14/2019

[3] "Diversity – Toronto Facts – Your City". City of Toronto. Archived from the original on April 6, 2015. https://web.archive.org/web/20150406180836/http://www1.toronto.ca/wps/portal/contentonly?vgnextoid=dbe867b42d853410VgnVCM10000071d60f89RCRD&vgnextchannel=57a12cc817453410VgnVCM10000071d60f89RCRD, Accessed on 12/14/2019

[4] "Focus on Geography Series, 2016 Census: Toronto, City (CSD) – Ontario: Immigration and Ethnocultural diversity". Statistics Canada. Archived from the original on November 8, 2017, Accessed on 12/14/2019

[5] http://www12.statcan.ca/census-recensement/2016/as-sa/fogs-spg/Facts-CSD-eng.cfm?TOPIC=7&LANG=eng&GK=CSD&GC=3520005, Accessed on 12/14/2019

[6] Toronto welcomed a record 43 million visitors in 2017, https://www.cbc.ca/news/canada/toronto/toronto-welcomed-a-record-43-million-visitors-in-2017-1.4501626, Posted 01/24/2018, Accessed on 12/14/19