

The International Student Guide of Choosing a Borough in Toronto



Theo Blank
February 12, 2020

Table of Contents:

Introduction:	2
Background:	2
Problem:	2
Interest:	3
Data Collection and Cleaning:	3
Data Collection:	3
Data Cleaning:	3
Feature Selection	4
Data Analysis:	6
Analyzing Venues in Each Borough:	6
Getting 10 Most Occurence Venue Types per Borough:	7
Clustering:	8
Results:	9
Cluster Map:	9
Result Table:	9
Recommendation:	10
Conclusion:	12

1. Introduction:

1.1. Background:

According to the Canadian Bureau for International Education the influx of international students in Canada has increased by approximately 52% from 2014 to 2017. Toronto is one of the cities attracting a lot of international students, and one of the hardest parts of moving to a new country as a student is finding the right accommodation for your studies. How do you know if the ones offered by the University of your choice are overpriced, or badly located? Which neighborhoods offer a wide selection of good venues while still having affordable real estate?

To answer such a question, this report will be analyzing which neighborhoods of Toronto are best suited for University students based on the surrounding venues and the average prices of 1 bedroom apartments.

1.2. Problem:

This project will give insight to the following research question:

a. What neighborhoods of Toronto are best suited for international University students?

The following sub questions will help reach the answer for the research question:

- a. What are the most occurring venue types for each borough?
- b. Which neighborhoods have the most diverse types of venues?
- d. Which neighborhood has the cheapest price?

1.3. Interest:

The results of this report can be useful for international students coming to Toronto and wondering which neighborhood will have the best combination of affordable pricing and good selection of venues. This data can prove useful to others as well, such as people looking for work or internships. However since I am using the average price of 1-bedroom apartments, the interest is limited to people residing alone.

2. Data Collection and Cleaning:

2.1. Data Collection:

For collecting the Borough's of Toronto, I web-scraped a [wikipedia page](#) that included all of Toronto's postal codes along with boroughs and neighborhoods. I checked the sources from the wikipedia page to make sure that the data was accurate and up to date. For web-scraping, I used a tool called Beautiful Soup, a Python library designed for pulling data out of HTML and XML files. The 1-bedroom average price data was pulled out of a [table](#) made by the [Canada Mortgage and Housing Corporation](#). I made my own table out of the data leaving out unnecessary information such as boroughs that I did not include in my research (they were too far from Downtown Toronto). Last, I used the [Foursquare API](#) to find nearby venues and their location for each borough.

2.2. Data Cleaning:

The first step of data cleaning was creating two tables: one for the boroughs of Toronto and one for the 1-Bedroom apartment average price using the data scraped from the previously mentioned websites. On the table featuring Toronto's boroughs, I created two new columns and fit the latitude and longitude to each borough. After that I

merged the two tables together to create the final cleaned dataframe, shown in the figure below:

Figure 1: Cleaned Data Frame of Toronto's Boroughs and Their Average Studio Price

	Borough	Latitude	Longitude	1-Bedroom Apt. Avg. Price (\$)
0	Central Toronto	43.690773	-79.403162	1522
1	Downtown Toronto	43.652503	-79.383558	1721
2	East Toronto	43.680800	-79.294400	1289
3	East York	43.691200	-79.341700	1199
4	Etobicoke	43.620500	-79.513200	1261
5	North York	43.761500	-79.411100	1281
6	Scarborough	43.776400	-79.231800	1216
7	West Toronto	43.658315	-79.442783	1339
8	York	43.695700	-79.450400	1301

2.3. Feature Selection

Due to the fact that I chose to analyze boroughs rather than neighborhoods, I wanted to make sure to include various venues around each borough to have an adequate idea of the venues available around any particular borough. I set the limit of venues to 100 calls and a radius to 1 km (any more than that would defeat the purpose of choosing an accommodation based on venues around a chosen location). The resulting table shows the amount of venues received by the Foursquare API for each borough:

Table 1: Number of Venues Per Borough in Toronto

Borough	# of Venues
<i>Central Toronto</i>	100
<i>Downtown Toronto</i>	100
<i>East Toronto</i>	100
<i>West Toronto</i>	100
<i>East York</i>	100
<i>Etobicoke</i>	92
<i>North York</i>	100
<i>Scarborough</i>	79
<i>York</i>	75

The feature selected will be a table including the boroughs along with their most common nearby venues and the average 1-bedroom apartment price.

Map 1: Map of Toronto's Boroughs Unclustered



3. Data Analysis:

3.1. Analyzing Venues in Each Borough:

The first step of the exploratory data analysis was to create a One-Hot Encoding of the surrounding venues for each borough. One-Hot Encoding is making categorical values such as venue names into numerical values (0 for False and 1 for True).

Figure 2: One-Hot Encoding of Surrounding Venues for Each Borough (846 rows, 172 columns)

	Borough	Afghan Restaurant	American Restaurant	Antique Shop	Argentinian Restaurant	Art Gallery	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	BBQ Joint	...
0	Central Toronto	0	0	0	0	0	0	0	0	0	...
1	Central Toronto	0	0	0	0	0	0	0	0	0	...
2	Central Toronto	0	0	0	0	0	0	0	0	0	...
3	Central Toronto	0	0	0	0	0	0	0	0	0	...
4	Central Toronto	0	0	0	0	0	0	0	0	0	...

Next, I transformed the Data Frame to show at what frequency each venue name shows up for each borough:

Figure 3: Venue Frequency Table for Each Borough (9 rows, 172 columns)

	Borough	Afghan Restaurant	American Restaurant	Antique Shop	Argentinian Restaurant	Art Gallery	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	BBQ Joint	...
0	Central Toronto	0.00	0.030000	0.0	0.0	0.00	0.00	0.00000	0.01	0.020000	...
1	Downtown Toronto	0.00	0.020000	0.0	0.0	0.02	0.01	0.01000	0.00	0.010000	...
2	East Toronto	0.00	0.000000	0.0	0.0	0.00	0.00	0.00000	0.00	0.030000	...
3	East York	0.01	0.020000	0.0	0.0	0.00	0.00	0.01000	0.00	0.010000	...
4	Etobicoke	0.00	0.021739	0.0	0.0	0.00	0.00	0.01087	0.00	0.021739	...

3.2. Getting 10 Most Occurrence Venue Types per Borough:

In order to prepare for clustering, I used the frequency table to create a new Data Frame that consists of the 10th most common venue surrounding each Borough in sequence. This table concludes the exploratory data analysis and allows the execution of clustering the boroughs.

Figure 4: 10 Most Occurrence Venues Types per Borough

	Borough	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	West Toronto	Café	Coffee Shop	Bar	Gastropub	Restaurant	Bakery	Cocktail Bar	Park	Brewery	Italian Restaurant
1	Downtown Toronto	Coffee Shop	Café	Hotel	Theater	Steakhouse	Gastropub	Pizza Place	Japanese Restaurant	Restaurant	Plaza
8	York	Coffee Shop	Grocery Store	Furniture / Home Store	Sandwich Place	Bakery	Pizza Place	Italian Restaurant	Fast Food Restaurant	Playground	Bank
6	Scarborough	Coffee Shop	Pharmacy	Sandwich Place	Gas Station	Park	Indian Restaurant	Fast Food Restaurant	Gym	Wings Joint	Chinese Restaurant
2	East Toronto	Coffee Shop	Pub	Bakery	Breakfast Spot	Park	Beach	BBQ Joint	Pharmacy	Thai Restaurant	Japanese Restaurant
4	Etobicoke	Coffee Shop	Restaurant	Italian Restaurant	Sandwich Place	Bakery	Convenience Store	Sushi Restaurant	Grocery Store	Fast Food Restaurant	Burrito Place
3	East York	Greek Restaurant	Café	Coffee Shop	Pizza Place	Gastropub	Ice Cream Shop	Bakery	Ethiopian Restaurant	Yoga Studio	Burger Joint
0	Central Toronto	Italian Restaurant	Park	Café	Coffee Shop	Restaurant	American Restaurant	Middle Eastern Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Indian Restaurant
5	North York	Korean Restaurant	Coffee Shop	Bubble Tea Shop	Pizza Place	Japanese Restaurant	Grocery Store	Ramen Restaurant	Bank	Sandwich Place	Fried Chicken Joint

3.3. Clustering:

As a machine learning algorithm I chose K-means clustering as it was the best fit for the type of data I was dealing with. Clustering will categorize each borough and cluster them into different groups depending on how similar/dissimilar they are from one another. I used K = 3 and produced the following Data Frame:

Figure 5: 10 Most Occurrence Venues Types per Borough Clustered

	Borough	Latitude	Longitude	1-Bedroom Apt. Avg. Price (\$)	Cluster	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
0	Central Toronto	43.690773	-79.403162	1522	0	Italian Restaurant	Park	Café	Coffee Shop	Restaurant	American Restaurant	Middle Eastern Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Indian Restaurant
8	York	43.695700	-79.450400	1301	1	Coffee Shop	Grocery Store	Furniture / Home Store	Sandwich Place	Bakery	Pizza Place	Italian Restaurant	Fast Food Restaurant	Playground	Bank
6	Scarborough	43.776400	-79.231800	1216	1	Coffee Shop	Pharmacy	Sandwich Place	Gas Station	Park	Indian Restaurant	Fast Food Restaurant	Gym	Wings Joint	Chinese Restaurant
4	Etobicoke	43.620500	-79.513200	1261	1	Coffee Shop	Restaurant	Italian Restaurant	Sandwich Place	Bakery	Convenience Store	Sushi Restaurant	Grocery Store	Fast Food Restaurant	Burrito Place
5	North York	43.761500	-79.411100	1281	1	Korean Restaurant	Coffee Shop	Bubble Tea Shop	Pizza Place	Japanese Restaurant	Grocery Store	Ramen Restaurant	Bank	Sandwich Place	Fried Chicken Joint
7	West Toronto	43.658315	-79.442783	1339	2	Café	Coffee Shop	Bar	Gastropub	Restaurant	Bakery	Cocktail Bar	Park	Brewery	Italian Restaurant
1	Downtown Toronto	43.652503	-79.383558	1721	2	Coffee Shop	Café	Hotel	Theater	Steakhouse	Gastropub	Pizza Place	Japanese Restaurant	Restaurant	Plaza
2	East Toronto	43.680800	-79.294400	1289	2	Coffee Shop	Pub	Bakery	Breakfast Spot	Park	Beach	BBQ Joint	Pharmacy	Thai Restaurant	Japanese Restaurant
3	East York	43.691200	-79.341700	1199	2	Greek Restaurant	Café	Coffee Shop	Pizza Place	Gastropub	Ice Cream Shop	Bakery	Ethiopian Restaurant	Yoga Studio	Burger Joint

4. Results:

4.1. Cluster Map:

After running the K-means clustering, the algorithm returned 3 clusters of Toronto's boroughs. In the first cluster (0) only one borough was attached: Central Toronto. The second cluster (1) contains York, Scarborough, Etobicoke and North York. The third cluster (2) contains most of Toronto's Old Town boroughs: West Toronto, Downtown Toronto, East Toronto and East York.

Map 2: Map of Toronto's Boroughs Clustered



4.2. Result Table:

The following table can be said to be the results itself, however I extracted the information from the table to recommend what I thought was the boroughs that represented the best features from their clusters:

Figure 6: Final Data Frame Clustered

	Borough	1-Bedroom Apt. Avg. Price (\$)	Cluster	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
0	Central Toronto	1522	0	Italian Restaurant	Park	Café	Coffee Shop	Restaurant	American Restaurant	Middle Eastern Restaurant	Sushi Restaurant	Vegetarian / Vegan Restaurant	Indian Restaurant
8	York	1301	1	Coffee Shop	Grocery Store	Furniture / Home Store	Sandwich Place	Bakery	Pizza Place	Italian Restaurant	Fast Food Restaurant	Playground	Bank
6	Scarborough	1216	1	Coffee Shop	Pharmacy	Sandwich Place	Gas Station	Park	Indian Restaurant	Fast Food Restaurant	Gym	Wings Joint	Chinese Restaurant
4	Etobicoke	1261	1	Coffee Shop	Restaurant	Italian Restaurant	Sandwich Place	Bakery	Convenience Store	Sushi Restaurant	Grocery Store	Fast Food Restaurant	Burrito Place
5	North York	1281	1	Korean Restaurant	Coffee Shop	Bubble Tea Shop	Pizza Place	Japanese Restaurant	Grocery Store	Ramen Restaurant	Bank	Sandwich Place	Fried Chicken Joint
7	West Toronto	1339	2	Café	Coffee Shop	Bar	Gastropub	Restaurant	Bakery	Cocktail Bar	Park	Brewery	Italian Restaurant
1	Downtown Toronto	1721	2	Coffee Shop	Café	Hotel	Theater	Steakhouse	Gastropub	Pizza Place	Japanese Restaurant	Restaurant	Plaza
2	East Toronto	1289	2	Coffee Shop	Pub	Bakery	Breakfast Spot	Park	Beach	BBQ Joint	Pharmacy	Thai Restaurant	Japanese Restaurant
3	East York	1199	2	Greek Restaurant	Café	Coffee Shop	Pizza Place	Gastropub	Ice Cream Shop	Bakery	Ethiopian Restaurant	Yoga Studio	Burger Joint

4.3. Recommendation:

Cluster 0: (Mostly restaurants, parks and cafés)

Central Toronto:

As the most expensive borough in Toronto, it's positioning allows close access to all other boroughs but features mostly restaurants and parks. Therefore if you are looking for good meals and open air activities, Central Toronto would be a good option for you.

Cluster 1: (Mostly coffee shops, Grocery Stores and restaurants)

York:

York fits into the middle price range for studio accomodation and features useful everyday shops such as grocery stores and furniture/home stores. It also has

a few selection of restaurants, banks and bakeries. York is also located close to Central and Downtown Toronto, making it the ideal Borough if you can not afford Central Toronto. If you are looking for a reasonably affordable borough with a lot of convenient facilities, York might be for you.

Scarborough:

Being the most expensive borough of Cluster 1, Scarborough is also located to the far east of Toronto, further away from Downtown than other boroughs. On the other hand Scarborough has more parks and gas stations than York, which is convenient if you own a car and enjoy the outdoors. With coffee shops being the most common venue, Scarborough trades the convenience of errand facilities for pharmacies and sandwich places. At last it also offers gyms and a few restaurant venues.

Cluster 2: (Mostly coffee shops, pubs/bars and cafés)

Downtown Toronto:

Downtown Toronto is the most expensive of all boroughs but offers the widest variety of venue choices. If you have a lot of visitors during the year they will be happy to know that this borough offers a wide variety of Hotels. While at their stay they can enjoy leisure venues such as Theaters, Steakhouses, Gastropubs and restaurants. It is noteworthy to mention that Downtown Toronto does not have a wide variety of grocery stores or parks.

East Toronto:

East Toronto is significantly cheaper than Downtown Toronto but still has a central location. East Toronto features a lot of pubs and coffee shops, making it the

night-life choice from the other boroughs. East Toronto also features beaches and BBQ joints, which are not very frequent or absent in other boroughs. There is a good selection of bakeries and pharmacies and a decent selection of restaurants. In conclusion East Toronto makes a great central location for students on a medium budget

5. Conclusion:

While the data chosen was accurate there are a few limitations to the accuracy of the results. Since I chose to analyze boroughs instead of neighborhoods, the area surrounding a borough is larger, which could result in overlapping venues from two different boroughs or exclusive venues which are too far away from the coordinates of the borough center point. In that case, the results shown are to be considered accurate for up to 1 km around the coordinates center point of any boroughs analyzed in this report. Nevertheless, whether you are an international student or looking for work/internships, this report should help you decide where to begin looking for apartments on your next real estate agency visit.