

Data Science Capstone Project

Toronto Public Libraries - Ensuring Their Success by Identifying Nearby Businesses That Influence People to Visit Neighborhood Libraries

By: Alexander Martinez

1. Introduction

1.1 Background

Based on several studies, the presence of "3rd places" strengthens communities. The home (first place), workspace / school (second place), and the library (or church or coffee shops) so called 3rd places help provide socially accepting and safe environments. The library can be considered as a place where people meet and social relationships made. Neighborhoods that have "famous" 3rd places like libraries tend to add economic and safety values to their neighborhoods.

At the end of this project, we would know what establishments are conducive to library foot traffic. The stakeholders would be the communities around the libraries. By providing the communities the information, they can continue to support their libraries more productively.

1.2 Problem

This project will determine popularity of Public Libraries based on foot traffic (using Foursquare Ratings). The stakeholders in this case are people who are living in or around these neighborhoods. And as such, this is not really a business problem but a social issue. It is imperative to ensure that people continue to visit libraries to help strengthen their communities. To further assist in increasing foot traffic to the libraries, we will determine top venues that would possibly contribute to the success of these libraries.

1.3 Interest

As stated earlier, a library success is also a community's success. People who visit libraries come from a wide spectrum of races, financial backgrounds, levels of education and varied interests. Having diverse people come to a "meeting place" can generate a sense of economic stability and social well-being for its residents.

2. Data Acquisition

2.1 Data Sources

We will determine which libraries are having considerable foot traffic and those that have low visits. We will limit our studies to libraries located within 20 kilometer radius from the geo center of Toronto. We will also get the top 10 venues that are close to the libraries.

Foursquare will be providing the data which would be the library locations, top venues, latitude / longitude data and ratings .

2.2 Data Processing

Data from Foursquare will be processed for both the libraries and the top venues and determine which libraries have ratings thus more likely to get good foot traffic and those that have poor or no rating that will probably need to be revitalized. We will recommend / suggest establishments of venues that would influence people to "check-in" and rate its neighbor libraries.

A list of libraries and venues and their ratings would have to be retrieved. Together with these, we need the latitude and longitude information. We would explore the top venues based on these geographic coordinates.

2.3 Feature Selection

Initial dataset would contain the name of the libraries, the top venues and geographical coordinates of the libraries and these establishments. Also needed would be the venue IDs to retrieve the ratings. .

Final data would include Library names, top venues' category, geographical coordinates of the libraries and top venue. The ratings would serve to determine which libraries need to be revitalized and which venues would help in the process.

2.4 Sample Data

Below is just a sample screenshots of data that would be used.

2.4.1 List of Toronto Public Libraries with Geographical Locations

Toronto Public Library - Toronto Reference Library 43.671795384146186 -79.38694401269817
 Toronto Public Library 43.652631397709854 -79.38329463125922
 Toronto Public Library - Lillian H. Smith Branch 43.65813719643491 -79.39837207880689
 Toronto Public Library - St. Lawrence Branch 43.6500480748644 -79.3683296509197
 Toronto Public Library (St. James Town) 43.668790439960134 -79.37499810757755
 Toronto Public Library - Palmerston Branch 43.66507382915182 -79.41397750802938
 Toronto Public Library - Bloor Gladstone Branch 43.66009688239916 -79.43417281113102
 Toronto Public Library - College/Shaw Branch 43.65494116588014 -79.42016707568195
 Toronto Public Library - Parkdale Branch 43.641248107585604 -79.43271440599969
 Toronto Public Library (Fort York Branch) 43.63917212520274 -79.40044547871639

2.4.2 Top Venues related to one of the libraries

Sansotei Ramen 43.671795384146186 -79.38694401269817
 Tim Hortons 43.652631397709854 -79.38329463125922
 Starbucks 43.65813719643491 -79.39837207880689
 Thor Espresso Bar 43.6500480748644 -79.3683296509197
 Colette Grand Café 43.668790439960134 -79.37499810757755

2.4.3 Library Ratings and Venue ID

	Name	Venue ID	Longitude	Latitude	Rating
0	Toronto Public Library	4c8938c8944e224b52e72285	-79.383295	43.652631	6.4
1	Toronto Public Library - Toronto Reference Lib...	4b5f2e80f964a52088ab29e3	-79.386944	43.671795	9.1
2	Toronto Public Library - Lillian H. Smith Branch	4ae6010ff964a520f7a321e3	-79.398372	43.658137	8.2
3	Toronto Public Library - St. Lawrence Branch	4b51e5aff964a5203c5a27e3	-79.368330	43.650048	5.5
4	Toronto Public Library (St. James Town)	4b807beef964a5209d7630e3	-79.374998	43.668790	7.2
5	Toronto Public Library - Palmerston Branch	4b26b348f964a520b97f24e3	-79.413978	43.665074	0.0
6	Toronto Public Library - College/Shaw Branch	4d9b3f1bb4fa37044e7b980d	-79.420167	43.654941	5.8
7	Toronto Public Library - Bloor Gladstone Branch	4b80365bf964a520e25c30e3	-79.434173	43.660097	8.3

2.4.4 List of Top Venues relative to Library Locations

	Library Name	Library Rating	Venue Name	Category	Venue id	Venue Latitude	Venue Longitude
0	Toronto Public Library	6.4	Downtown Toronto	Neighborhood	5227bb01498e17bf485e6202	43.65323167517444	-79.38529600606677
1	Toronto Public Library	6.4	Nathan Phillips Square	Plaza	4ad4c05ef964a520a6f620e3	43.65227047322295	-79.38351631164551
2	Toronto Public Library	6.4	Eggspensation Bell Trinity Square	Breakfast Spot	537773d1498e74a75bb75c1e	43.65314383888587	-79.38198016678167
3	Toronto Public Library	6.4	Old City Hall	Monument / Landmark	4ad4c05ef964a5208ef620e3	43.652008800876125	-79.3817442232328
4	Toronto Public Library	6.4	Indigo	Bookstore	4b2a6eb8f964a52012a924e3	43.65351471121164	-79.38069591056922
5	Toronto Public Library	6.4	M Square Coffee Co	Coffee Shop	54132b3b498ee9ca9332e189	43.65121797253777	-79.38355459932247

2.4.5 Venue Rating

Venue Name	Venue Rating	Category
Downtown Toronto	9.3	Landmark
Nathan Phillips Square	9.0	Landmark
Eggspensation Bell Trinity Square	8.7	Cafe
Old City Hall	8.5	Landmark
Indigo	8.5	Bookstore
Four Seasons Hotel Toronto	8.8	Hotel
Sabai Sabai	8.7	Restaurant

3.0 Methodology

As mentioned earlier, the objective is determine which libraries are in need of more foot traffic and which venues can possibly help boost visit to these libraries. Several factors are necessary for the libraries to play a successful role as economic and social stimuli for their communities. As such, we would be dealing with objective data and a little subjective as well. The key venues needed would have to be of different categories so as to attract different people of varied interests which is an essential factor in library success.

The study was limited to a 5 KM radius from the geographical center of Toronto and limited to 20 libraries. About 30% of the libraries do not have any rating which would probably suggest that people are not enthusiastic or interested in those libraries. In Foursquare lingo, there is not enough sentiment data to do a rating calculation. Instead of not including these in the study, I decided to include them and clearly classify them as low foot traffic. We will review the other 70% of libraries that have ratings and from these identify the establishments that would help.

We are going to use k-means clustering to identify similarities among the libraries.

4.0 Results

Fig. 1 is a bar graph of the ratings of the different libraries. 7 of the libraries have no ratings on Foursquare.

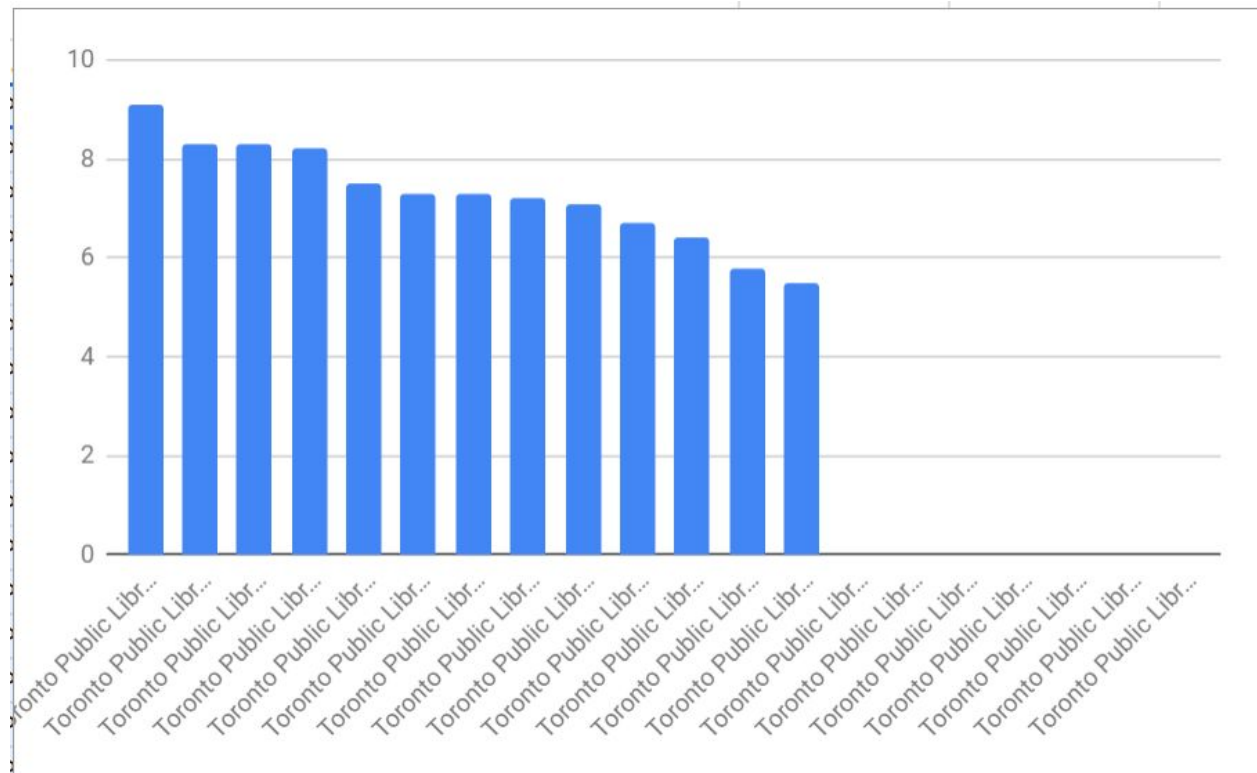


Fig. 1 Chart showing Library Ratings

The pie chart below (Fig. 2) shows the different categories of popular venues around 500 meters of each library. From the chart, it is evident that those with ratings have a big chunk of establishments nearby as cafe (or coffee shops) then restaurants. Italian restaurants are in good demand around these libraries. It is also interesting to note that bars are also near these libraries.

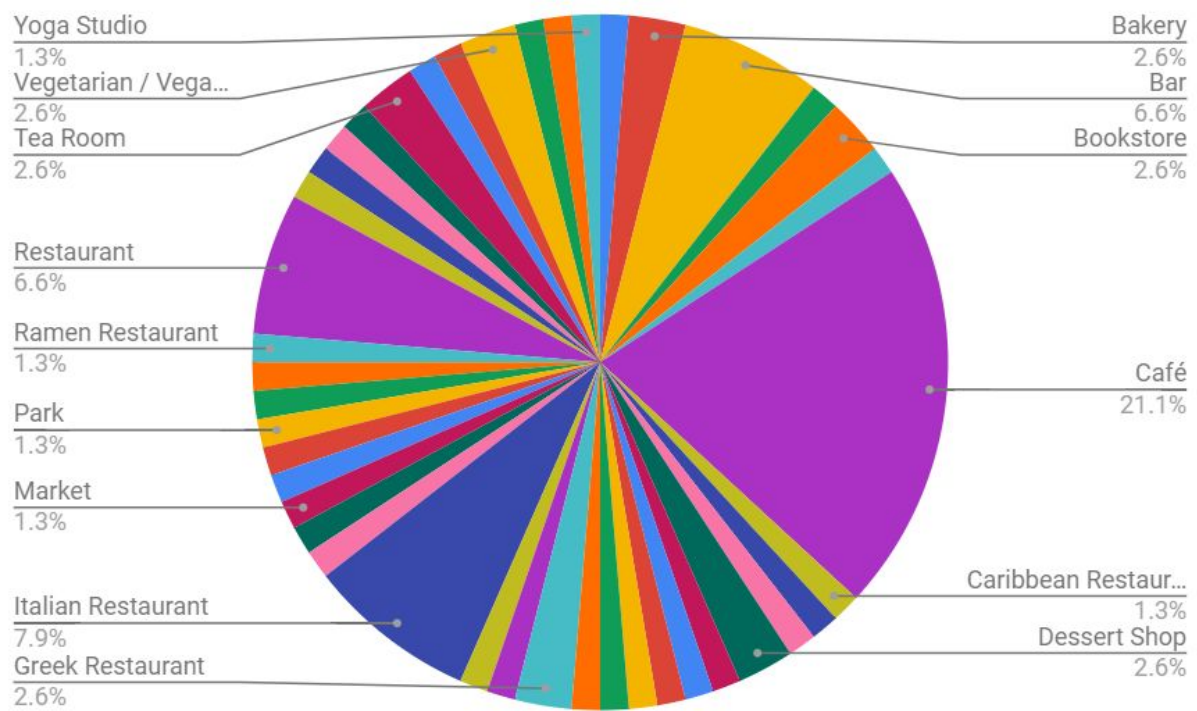


Fig. 2 Distribution of Simplified Categories for Libraries with Ratings

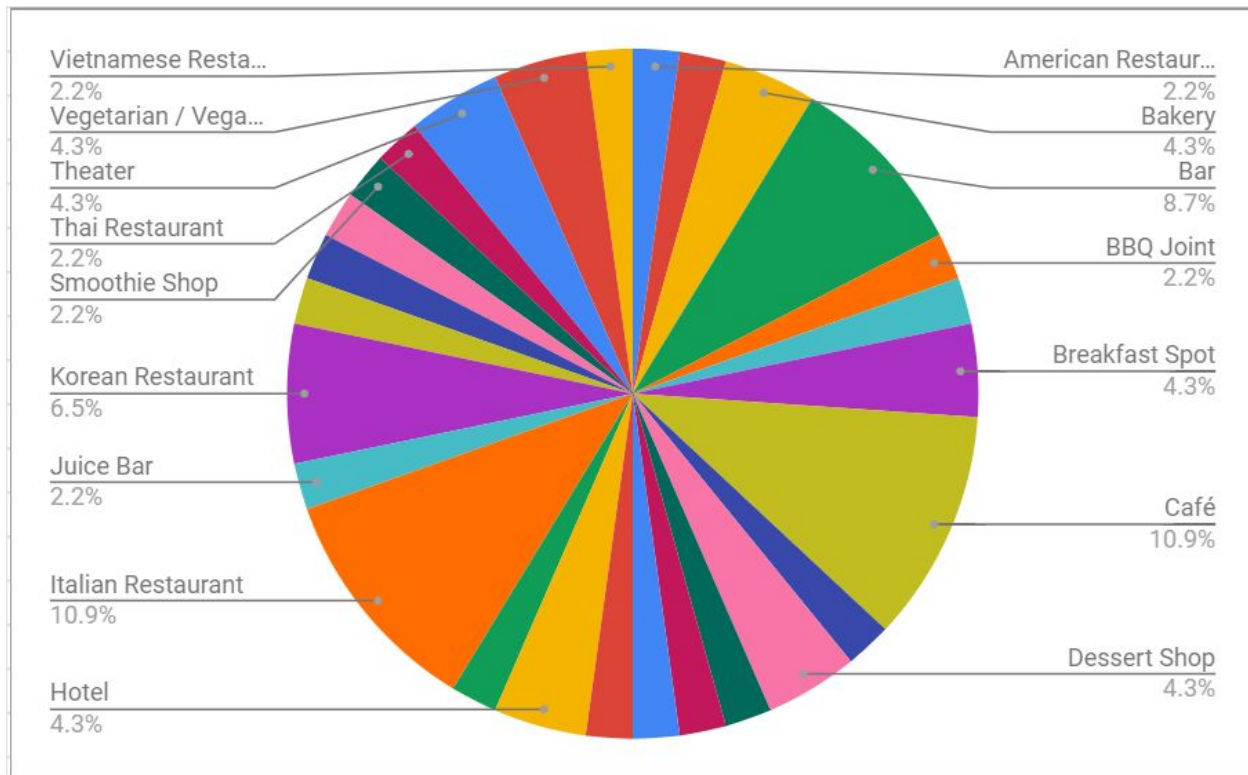


Fig. 3 Distribution of Simplified Categories for Libraries with No Ratings

The above pie chart represents top venues for libraries with no rating. The top categories are the same for both with ratings and none (Cafe, Italian Restaurant, Restaurants, and Bar).

I also used k-means clustering to identify what's common with these libraries. The clusters that were developed includes both libraries with rating and no rating.

Below is a map showing the clustered libraries showing a mix of rated and non-rated libraries.

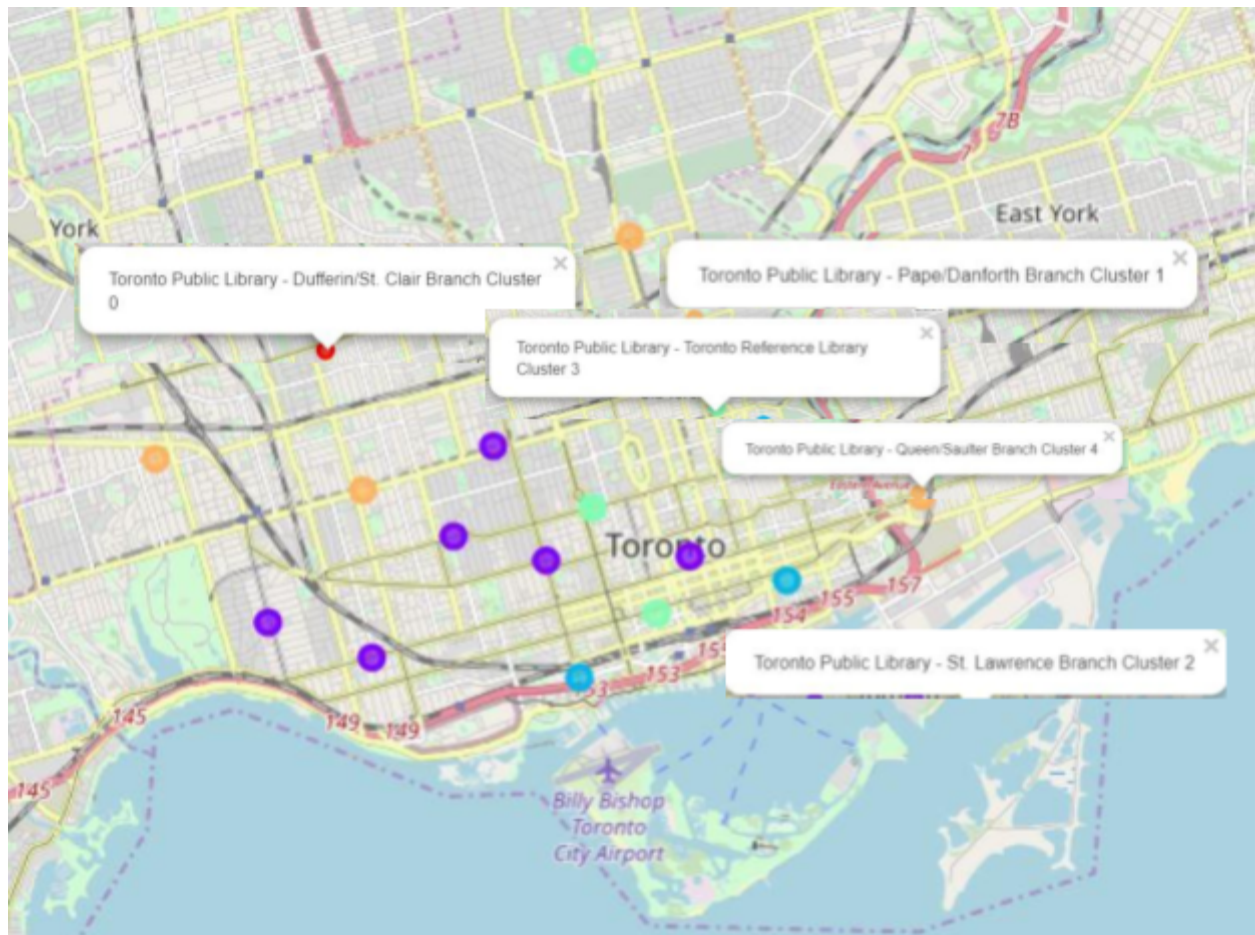


Fig. 4 Map showing Clustering of Libraries

5.0 Discussion

There are venue similarities between these libraries regardless whether they are rated or not. This was shown with the pie charts generated and with the k-means clustering. Fig. 5 shows the results of the clustering. Only one cluster was isolated but still showed restaurants, although specializing more in foreign flavors. It can also be that these libraries are reasonably closed to one another thus showing the same top venues. This “colocation” may have attributed to some of them having the same results as far as the top venues are concerned.

	Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
18	Toronto Public Library - Dufferin/St. Clair Br...	Italian Restaurant	Breakfast Spot	Sushi Restaurant	Brazilian Restaurant	Coffee Shop

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

	Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Toronto Public Library	Plaza	Bookstore	Monument / Landmark	Restaurant	Coffee Shop
5	Toronto Public Library - Palmerston Branch	Korean Restaurant	Café	Health Food Store	Vegetarian / Vegan Restaurant	Dessert Shop
6	Toronto Public Library - College/Shaw Branch	Nightclub	Café	Bookstore	Music Store	Tapas Restaurant
9	Toronto Public Library - Parkdale Branch	French Restaurant	Arts & Crafts Store	BBQ Joint	Hawaiian Restaurant	Thrift / Vintage Store
11	Toronto Public Library (Sanderson Branch)	Filipino Restaurant	Art Gallery	Park	Vegetarian / Vegan Restaurant	Italian Restaurant
12	Toronto Public Library - Pape/Danforth Branch	Pizza Place	Greek Restaurant	Caribbean Restaurant	Diner	Donut Shop
17	Toronto Public Library (High Park Branch)	Bakery	American Restaurant	Eastern European Restaurant	Coffee Shop	Food & Drink Shop

```
library_merged.loc[library_merged['Cluster Labels'] == 2, library_merged.columns[[0] + list(range(5, library_merged.shape[1]))]]
```

	Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	Toronto Public Library - St. Lawrence Branch	Coffee Shop	Grocery Store	Restaurant	Bar	Liquor Store
4	Toronto Public Library (St. James Town)	Coffee Shop	Market	Steakhouse	Diner	Restaurant
8	Toronto Public Library (Fort York Branch)	Coffee Shop	Yoga Studio	Café	Diner	Ramen Restaurant

```
library_merged.loc[library_merged['Cluster Labels'] == 3, library_merged.columns[[0] + list(range(5, library_merged.shape[1]))]]
```

	Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Toronto Public Library - Toronto Reference Lib...	Hotel	Gourmet Shop	Italian Restaurant	Comic Shop	Furniture / Home Store
2	Toronto Public Library - Lillian H. Smith Branch	Supermarket	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Coffee Shop	Gaming Cafe
10	Toronto Public Library - Northern District Branch	Italian Restaurant	Bookstore	Plaza	Coffee Shop	Tea Room
14	Toronto Public Library Bookmobile	Theater	Hotel	Concert Hall	Gym / Fitness Center	Comedy Club
15	Toronto Public Library (Merril Collection)	Supermarket	Vietnamese Restaurant	Vegetarian / Vegan Restaurant	Coffee Shop	Gaming Cafe

```
library_merged.loc[library_merged['Cluster Labels'] == 4, library_merged.columns[[0] + list(range(5, library_merged.shape[1]))]]
```

	Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
7	Toronto Public Library - Bloor Gladstone Branch	Bar	Café	Vegetarian / Vegan Restaurant	Bakery	Beer Bar
13	Toronto Public Library - Deer Park Branch	Café	Yoga Studio	Tea Room	Bagel Shop	Chiropractor
16	Toronto Public Library (Annette Street)	Bar	Café	Grocery Store	Arts & Crafts Store	Italian Restaurant
19	Toronto Public Library - Queen/Saulter Branch	Bar	Hotel	Restaurant	Middle Eastern Restaurant	Café

Fig. 5 Clustering Result

6.0 Conclusion

Based on these observations, I would have to conclude that foot traffic to libraries are not greatly influenced by establishments around them.

The study actually started out with the idea of finding which venues would help in establishing libraries as a third space. Those that can help stimulate economic growth and at the same time become a social center to strengthen the communities.

The results of the study indicate that the venues near the libraries do not extensively contribute to their foot traffic. What would encourage foot traffic to libraries would then be other factors. It can be proximity to schools, transportation stations, or government institutions but that would be for another study.