The Battle of the Neighborhoods-Toronto

JPaul Tuyikunde

June 23, 2021

1. Introduction

1.1 Background

The battle of the neighborhoods is the project aimed to find difference in neighborhoods across some cities, as to indicate the best neighborhoods for different purposes such as living, working, business openings etc. I took this Battle now to city of Toronto, which is the capital city of Canadian province of Ontario, most populous city in Canada and fourth one in North America. It is recognized as multicultural and cosmopolitan city and center of business, finance and arts. According to the city of Toronto, the neighborhoods are changing from 140 to 158 in number, where 34 new ones replaced old 16 neighborhoods. These neighborhood profiles were developed to help government and community agencies with their local planning, by providing socio-economic data at a meaningful geographic area. To open a business or finding the best spot for your office in this amazing city requires to do a research on where and what to consider before taking a decision.

1.2 Problem

The project aims to provide recommendations to stakeholders who are interested in opening a restaurant or a business in neighborhoods of Toronto.

2. Data Acquisition

2.1 Data sources

The neighborhoods data was found by scraping the Toronto city website here. The data was clean, so no other cleaning process was required.

The neighborhoods data was used to pull new such as venues, latitude and longitudes of the venues, the category of each venue by the help of FourSquare API. The resulted data was clean.

Table 1: snippet of dataset

	Neighborhood	Venue	Venue_Latitude	Venue_Longitude	Venue_Category
0	Etobicoke West Mall	Centennial Park	43.656154	-79.587540	Park
1	Etobicoke West Mall	Tim Hortons	43.644742	-79.567681	Coffee Shop
2	Etobicoke West Mall	The Beer Store	43.641313	-79.576925	Beer Store
3	Etobicoke West Mall	Porta Via	43.663449	-79.589638	Sandwich Place
4	Etobicoke West Mall	Best for Bride	43.635767	-79.539916	Women's Store

3. Exploratory data analysis

Calculating Value counts on our dataset, there is large number of sandwich places and coffee shops in Toronto as well as banks as it is center of finance and business.

Table 2: Top five Categories.

	Venue_Category
Sandwich Place	98
Coffee Shop	78
Bank	63
Pizza Place	56
Park	49

Within the data collected, there is 483 restaurants in our dataset. With different types of cuisines, some of them are Thai, American, Persian, Turkish, Greeks restaurants, etc.

In addition, there is more than 1300 other business other than restaurants in the data set.

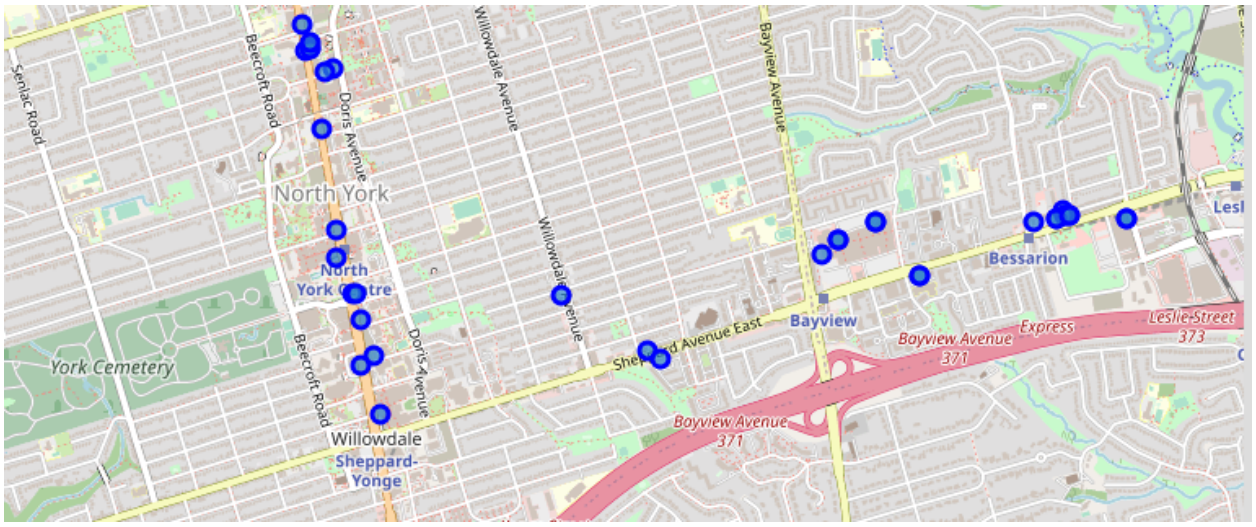

Figure 1: North York Neighborhood

4. K-Means clustering

In this project, the use of clustering was employed to cluster the restaurants venues. Cluster is a group of objects that are similar to other objects in cluster and dissimilar to data points in other clusters.

By using the K-Means clustering, the data was divided into non-overlapping subsets without any cluster-internal structure.

After performing one-hot encoding to the restaurants, I imported the **KMeans** from **sklearn** library, and chooses the number of clusters to be four.

The resulted labels were inserted to the dataframe to perform the rest of the analysis based on those labels.

Table 3: Categories with cluster-labels

Neighborhood	Venue	Venue_Latitude	Venue_Longitude	Venue_Category	Cluster_labels
Etobicoke West Mall	Mrakovic	43.666641	-79.578850	Eastern European Restaurant	2
Etobicoke West Mall	Taste of Thailand Cuisine	43.635928	-79.540785	Thai Restaurant	2
Etobicoke West Mall	Bravo Bistro	43.659420	-79.603604	Eastern European Restaurant	2
Etobicoke West Mall	Anatolia Restaurant	43.644596	-79.532810	Turkish Restaurant	2
Etobicoke West Mall	Astoria Shish Kebob House	43.621795	-79.570540	Greek Restaurant	2

5 rows × 53 columns

The results of clustering was then plotted on map using folium as shown below.

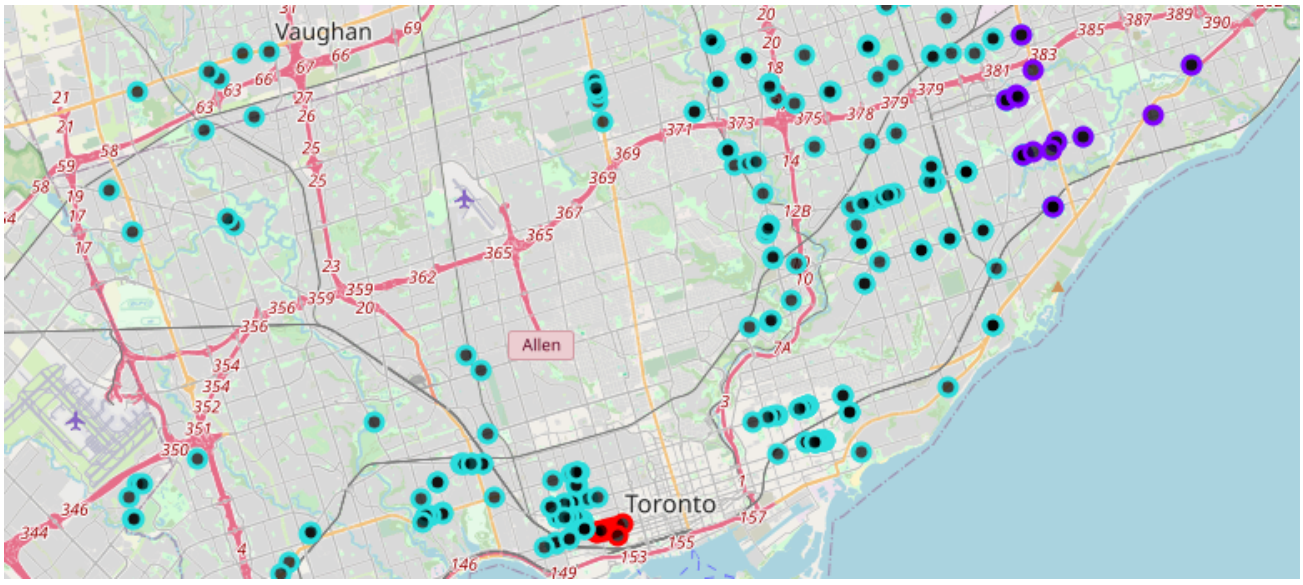

Figure 2: Clusters of Restaurants

5. Results and Discussion

Through this analysis, we found that there is high density of Restaurants in East of Eglington Avenue and Lawrence Avenue and there is low density of Restaurants in North of Etobicoke. We also found that according to clusters of Restaurants, that there is some areas with high competition and low competition in Toronto. Within cluster1, there is large number of American Restaurants but again their number keep getting low in the other Clusters.

	Venue_Category
American Restaurant	15
Restaurant	10
Seafood Restaurant	5
Italian Restaurant	5
Thai Restaurant	5
Middle Eastern Restaurant	4
Mexican Restaurant	3
New American Restaurant	2
Chinese Restaurant	2
Mediterranean Restaurant	2
Sushi Restaurant	2
Vegetarian / Vegan Restaurant	2
Dumpling Restaurant	1
Asian Restaurant	1

Figure 3: Cluster 1 restaurants

The Fast Food Restaurants Seems to be in all clusters, which shows another major competition in opening the Fast food type of restaurants.

The Restaurants that are based on countries culture are scattered in Toronto. We can see some like *Greek, Chinese, Italian, Persian, Cuban, Afghan* Restaurants, etc..., those restaurants do not seem to have a large number among clusters and for someone looking to open restaurants based on his/her country, it is worth a shot.

In addition, we can see that Indian and Italian Restaurants are popular ones on those types. However, trying to open them in cluster1, you can have the less competition compared to other clusters.

According to the map of services, there is big number of services and other business other than restaurants on North-East of University of Toronto, which in fact could be the best setup for *offices* for stakeholders interested in locating their offices in Toronto. Those businesses available on that side includes *bars, brewery, cafes, clothing stores and bakeries*. There is also *parks, stables and banks*.

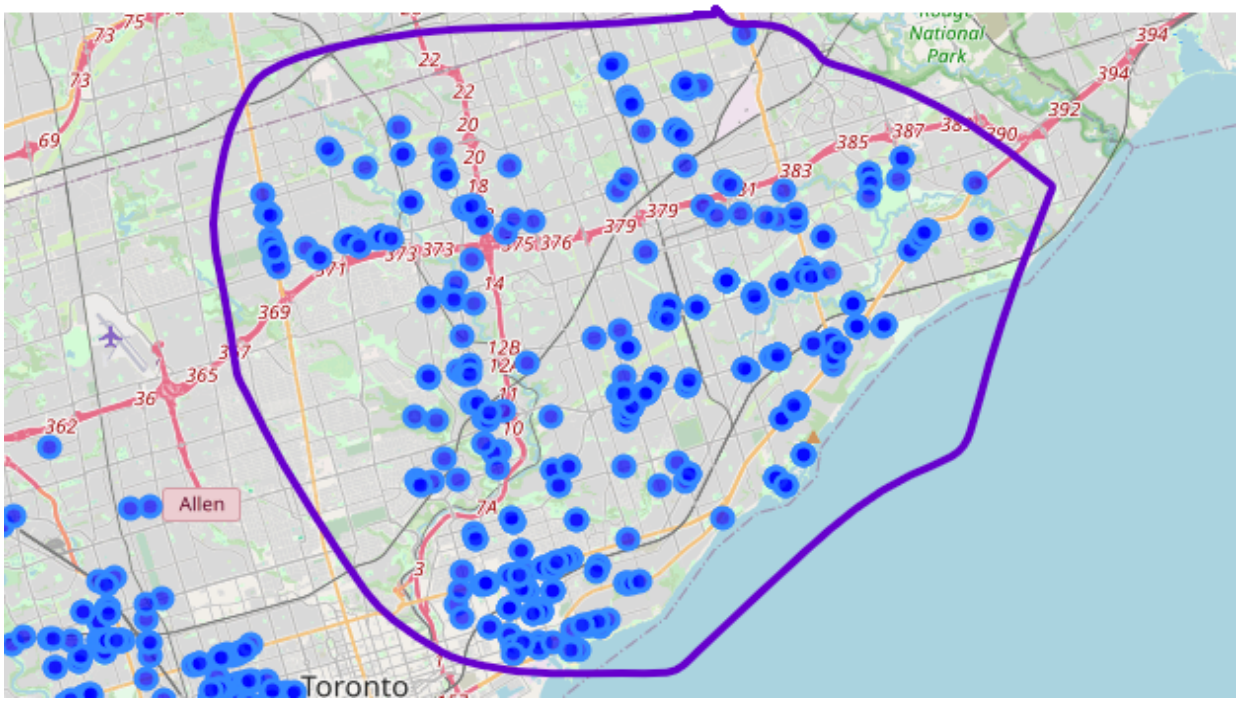

Figure 4: High-density business area.

6. Conclusion

The project's purpose was to gather data on neighborhoods of Toronto performing analysis on data about the restaurants and services to aid stakeholders who want to open mainly restaurants or other services, to locate the best area in Toronto's Neighborhoods.

By gathering data from Toronto city website about neighborhoods, and use them to gather venues in those neighborhoods using FourSquare API. Even though some data were unavailable on some neighborhoods, the available ones helped in clustering the neighborhood's venues and finding the best areas suited for opening restaurants and types of restaurants suited for each clusters.

After all the decision is up to stakeholder, based on his/her specific characteristics of neighborhoods in the recommended areas, also by considering some additional factors such as culture based business, attractiveness of neighborhoods, services available, prices and major roads etc.