The Battle of Neighborhoods

Baiju Thomas

1. Introduction

   Migration is happening from country to country for decades for better ppertunities. Globalization is one of the reasons people move to new places for better job opportunities Migration is a huge event for most migrants. As immmigartion increases, people will look for more oppertunities in terms of accomodation and comfort. . For example, a donut lover who usually needs a cup of coffee and donut every morning from Tim hortans will prefer a place with a cafe nearby. I would like to use the Foursquare location data to analyze the neighborhoods of four different cities New York and Toronto. These cities are the economic or political center of the countries. I wish to find similar neighborhoods among these cities and provide some useful information for people who are considering moving among Canada and USA
2. Data
   The neighborhood data of New York is provided here:
   https://geo.nyu.edu/catalog/nyu_2451_34572
   The neighborhoods of Toronto are extracted from here:
   https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

3. Data Preparation , Acqusition and Analysis
   The New York data is available online with the neighborhood name and location information (latitude and longitude). For Toronto, only neighborhood name is
   provided, I will be using the Nominatim library to obtain the location information by searching each neighborhood name.

   The Foursquare API was used to search for nearby venues of each neighborhood in radius of 1000 meters. Only the venues name and venues category (i.e. café, restaurant, school, etc.) are extracted.

   According to the total number of venues in each category, I will be using hierarchical agglomerative clustering method to compare neighborhoods among cities. Unlike kmeans or other machine learning clustering methods, hierarchical agglomerative clustering doesn't require the number of clusters at the beginning. Moreover, it could also tell us whether the dataset is good for clustering at first glance.

4 Results

4.1 Total number of neighborhoods in each cluster

Let's look at the total number of neighborhoods in each cluster to make sure they are properly Clustered

Cluster	0	1	2	3	4
Total number of neighborhoods	39	159	131	32	26

Cluster 0

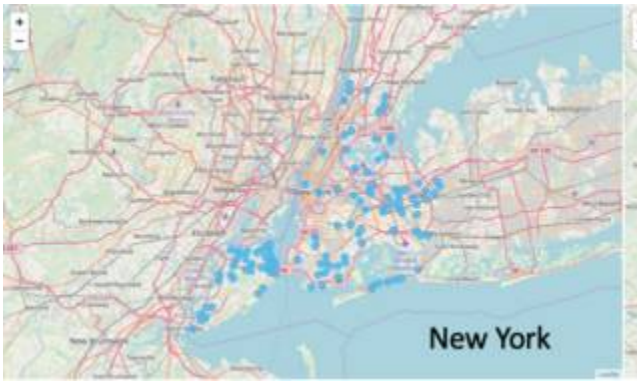
Cluster 0 contains the neighborhoods in New York with a lot of nearby pizza places, fast food restaurants and Caribbean restaurants.

Cluster 1

Cluster 1 is basically the residential area of each city. Each neighborhood has access to fast food restaurants and outdoor parks. There are also a lot of restaurants, convenience stores and supermarkets.

Cluster 2

Cluster 2: New York and Toronto neighborhoods with pizza places, banks, mobile phone shops nearby.



Cluster 3

Cluster 3: Most are New York neighborhoods with a lot of Italian and pizza places.

Discussion

Overall, the results are not surprising. For those who are looking for a good residential area to settle down, neighborhoods in cluster 1 will be a good consideration. These neighborhoods are not crowded as city centers, but has a number of outdoor parks, restaurants and supermarkets.

Conclusion

I hope this could provide some useful information for people who are considering moving among Canada and USA.