**Analysis of business conditions of boroughs in New York and Toronto**

*Zhongxiang Wu, Mar.2020*

1. **Background and description of the problem**

Each borough in the same or different cities has different life styles and business environment. Some boroughs are quite and suitable for living while some boroughs are fast-paced with prosperous business. In this section, a borough comparison has been conducted to investigate the differences and characteristics between two famous cities in New York and Toronto respectively.

Through the analysis, we can find the venues characteristics as well as the venue preference of consumers in each borough of these two cities. The analysis results can provide a good reference for people to choose a familiar living environment, and for business owners to select compatible business opportunity in one borough.

People looking for an living place, and business owner can be benefit from this analysis.

1. **Data acquisition and description**

Neighborhood data of New York and Toronto, which obtained from data portal of NYU and Wikipedia, includes the data of neighborhood name, borough name, latitude, and longitude.

Venue data can be fetched according to the neighborhood information (e.g. Neighborhood name, latitude and longitude) via Foursquare explore API. As a result, the detailed information of venues in different borough can be obtained for further analysis such as similarity analysis.

1. **Methodology description**

In data preprocessing phase, each borough will be linked with a relevant venue list that contains the name, category and geographical coordinate information via Foursquare API. In addition, the one-hot coding process is implemented to quantify the categorical data, and then the data is aggregated by borough and calculate the average values.

In order to identify the data pattern and cluster the data, the dimensionality of dataset is reduced via principle component analysis (PCA), and the output of PCA is used for data clustering and visualization. Lastly, the borough is clustered based on K-means methodology.

1. **Results**

The PCA result is shown as diagram 1 when the component amount equals 2. Most of boroughs from Toronto except downtown Toronto are positioned closely, this indicates the venue situation of each borough is very similar. Boroughs of city New York are diverse and different with each other, and not same as the boroughs of Toronto either. Long Euclidean distance between borough Manhattan and all the rest boroughs demonstrates its unique business characteristics.

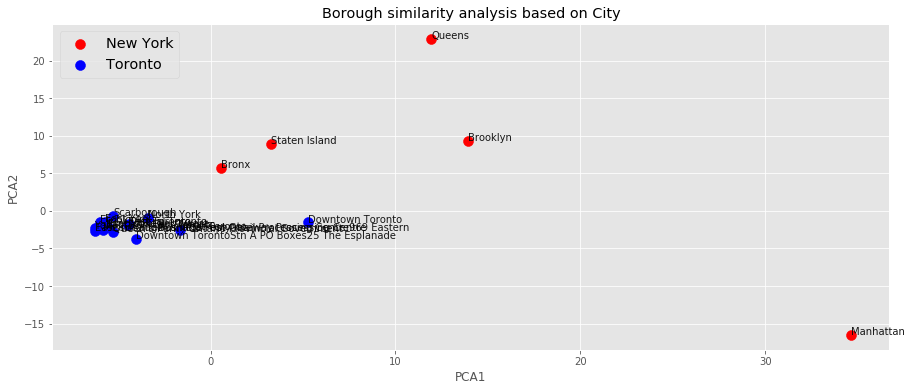


Diagram 1. PCA results (components = 2)

The result of K-Means is shown as diagram 2 when K is four. 4 borough clusters of New York and Toronto are obtained, they’re described as below:

* Cluster 1: All borough of Toronto except Downtown Toronto(City Toronto)
* Cluster 2: Downtown Toronto(City Toronto), Bronx and Staten Island(City New York)
* Cluster 3: Queens, and Brooklyn(City New York)
* Cluster 4: Manhattan (City New York)

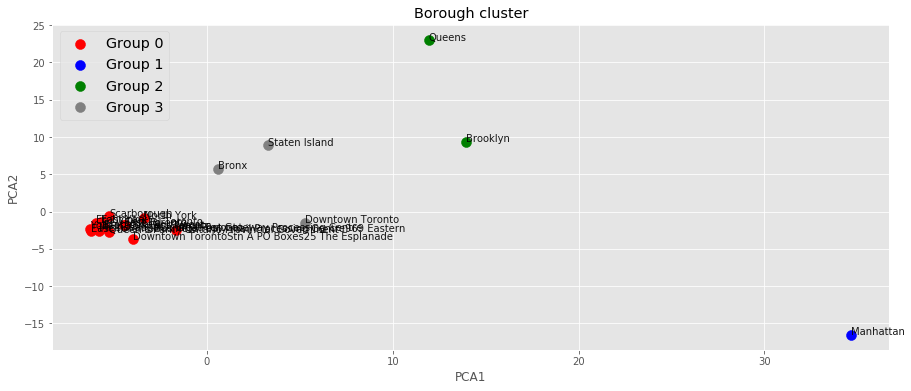


Diagram 2. K-means cluster results (K=4)

From the cluster results, we found that almost all the boroughs of Toronto are clustered as one group and they’re quite different with other ones in New York, where TOP3 most common venues are park, coffee shop and rental car location; Regarding borough Manhattan, it’s totally different with all other boroughs no matter in Toronto or New York, most common venues are related to coffee shop and restaurant; Borough Staten Island and Bronx from city New York and Downtown Toronto from city Toronto are clustered as one group, pizza place and coffee shop are much favorite venues; Lastly, Queens and Brooklyn from city New York are clustered as one category as Pizza and bar are most preferred by consumers there.

1. **Discussion**

According to the clustering analysis, we can make conclusions that (1) most boroughs in city Toronto are quite relax and closed to nature as many parks are there, and adequate rental parking lots can also demonstrate its convenience for people in city Toronto to travel. The food in most boroughs of city Toronto are not complicated, pizza, sandwich and fast food are most favorite there. In summary, it will be a good choice for the people who likes the nature and relax lifestyle to live in most boroughs of city Toronto except downtown Toronto, and business owner can focus on light meal/fast food; (2) A large variety of restaurants is a characteristic of borough Manhattan, and we can feel the prosperous business and luxurious lifestyle; (3) Pizza and Italian restaurant are very popular in borough Bronx, Staten Island and downtown Toronto. Besides, the bus stop, pharmacy and grocery store are easy to find in these boroughs. As a result that the borough Bronx and Staten Island will be a good opinion for the people who enjoy urban life with much conveniences. Fast food will be potential business opportunities in these places. (4) Bars in Queens and Brooklyn are popular place for consumers, and people there are fond of pizza, deli and bodega. We can guess that people in these two boroughs enjoy social life, and featured bars and pizzerias could be potential business opportunities.

1. **Conclusions**

Business situation analysis of each borough in two well-known cities New York and Toronto has been conducted in this research, and the different characteristic of each borough has been investigated. The analysis based on the data includes the neighborhood information of New York and Toronto, and the venue information from Foursquare API. The raw data are processed via data integration and one hot coding, and un-supervisory machine learning techniques includes principle component analysis (PCA) and K-Means have been utilized to identify the data pattern and cluster the data. Finally, 4 borough clusters are determined and different kinds of life and business styles of these boroughs are learned from the analysis. This analysis result will be the useful information for people to choose the favorite living place, and it can also support business owner to search potential business opportunities in these boroughs.