A Recommend living place for relaxing travel in Toronto

By Haowen Wang

1. Introduction: About the Problem

Toronto can be a famous place for tourists from all over the world, and it has many tourist attractions in the old Toronto (East/ Central/ Downtown/ West Toronto). So when tourists want to have a visit to old Toronto, they have to choose a place to live. For tourists who want to experience the life of Toronto residents, variable venues can be necessary. This recommend system is designed to help tourists to choose the best place to live in Toronto. To some extent, better place means better quality of holiday.

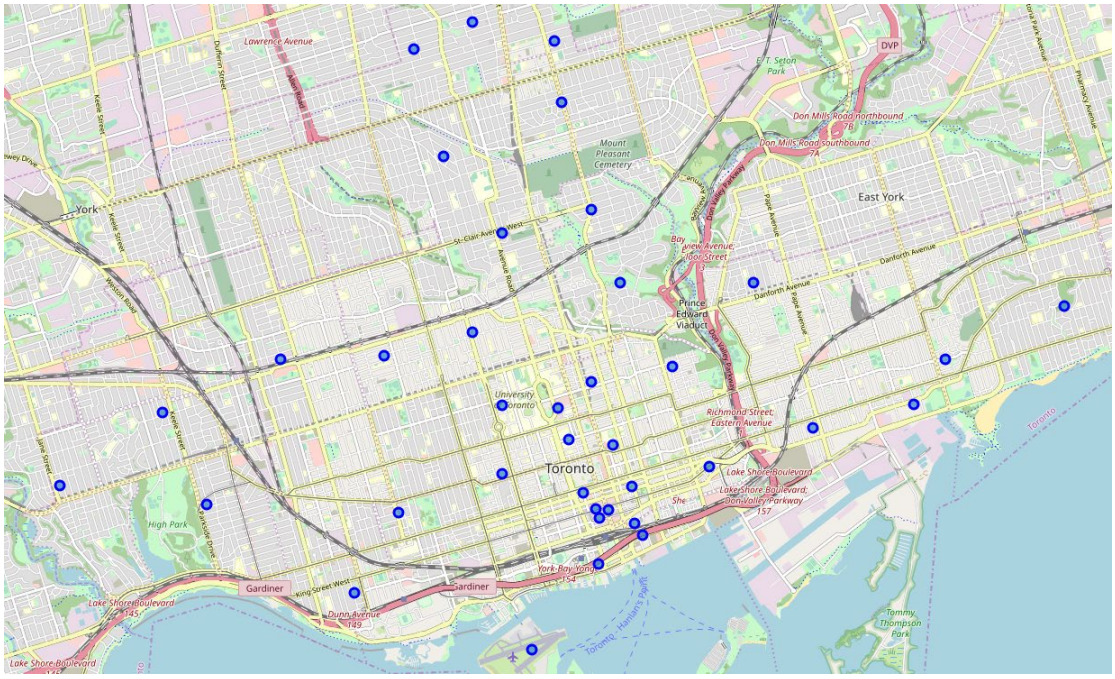
For example, if the place tourists choose is far away from the recreational facilities, it means that they need to spend a lot of time for commuting. For people who do not want to live in the downtown, the bus station and subway station can be important so that they can guarantee them to go to other place more convenient.

The tourists should choose place to stay where is closest to much more different venues so that they can experience the life of local residents more conveniently. Restaurant and commuting station can be important because they are related to daily life.

2. Data acquisition and cleaning

We will need location data information about venues inside old Toronto borough, in fact, we have acquired borough data in Toronto ahead. And what I need to do is to filter the old Toronto borough data (East/ Central/ Downtown/ West Toronto). Then by using the Foursquare API, I got different venues' data in different neighborhoods in old Toronto borough. These data covers the location of the venues, the name of the venues and its neighborhoods. And the most important, the venues' Category and the venues' summary.

We will first filter the tourists place where is popular, because they are the most attractive to tourists, then we will use the K-Means algorithm to analyze the data so that we can find the best neighborhood for tourists to live, where can guarantee the most variable of venues in old Toronto. We will do research about the clusters and analyze the data we got.
Then we could create a map focus on the old Toronto.

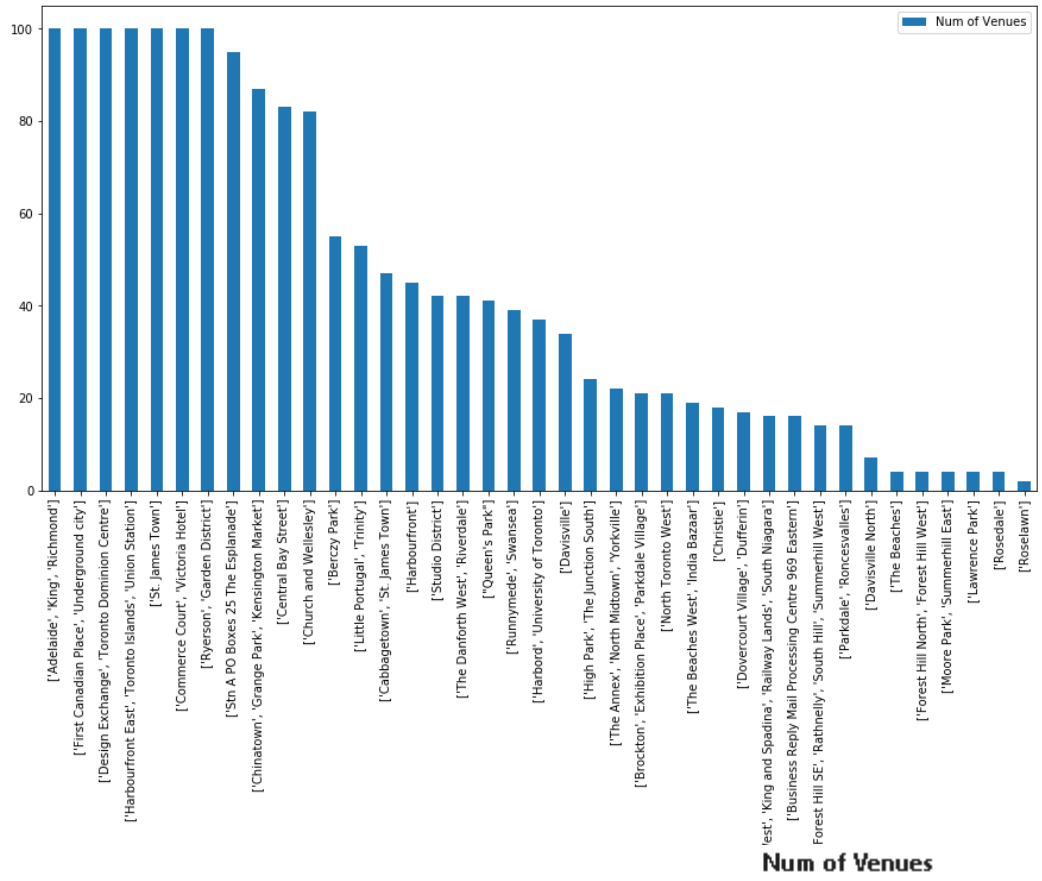


3. Methodology and Data analysis

We explore some summary information about Neighborhoods inside old Toronto so that it can help us have a better understanding of the data.

3.1. The top 10 Neibourhood which has the most Venues

We can have the top ten neighbourhoods which has the most venues, thus we found that the top 6 have the same number as 100, so we choose 6 cluster to do the K-Means cluster algorithm later. The data tell us that it is possible to do this.

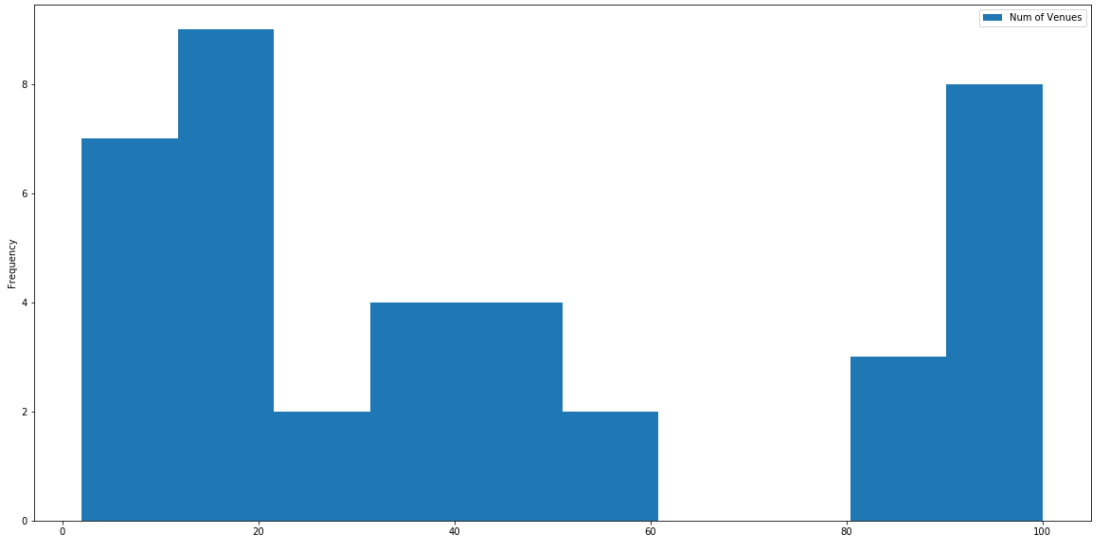


Neighbourhood	Num of Venues
['Adelaide', 'King', 'Richmond']	100
['First Canadian Place', 'Underground city']	100
['Design Exchange', 'Toronto Dominion Centre']	100
['Harbourfront East', 'Toronto Islands', 'Union Station']	100
['St. James Town']	100
['Commerce Court', 'Victoria Hotel']	100
['Ryerson', 'Garden District']	100
['Stn A PO Boxes 25 The Esplanade']	95
['Chinatown', 'Grange Park', 'Kensington Market']	87
['Central Bay Street']	83

3.2. The hist plot of Venues

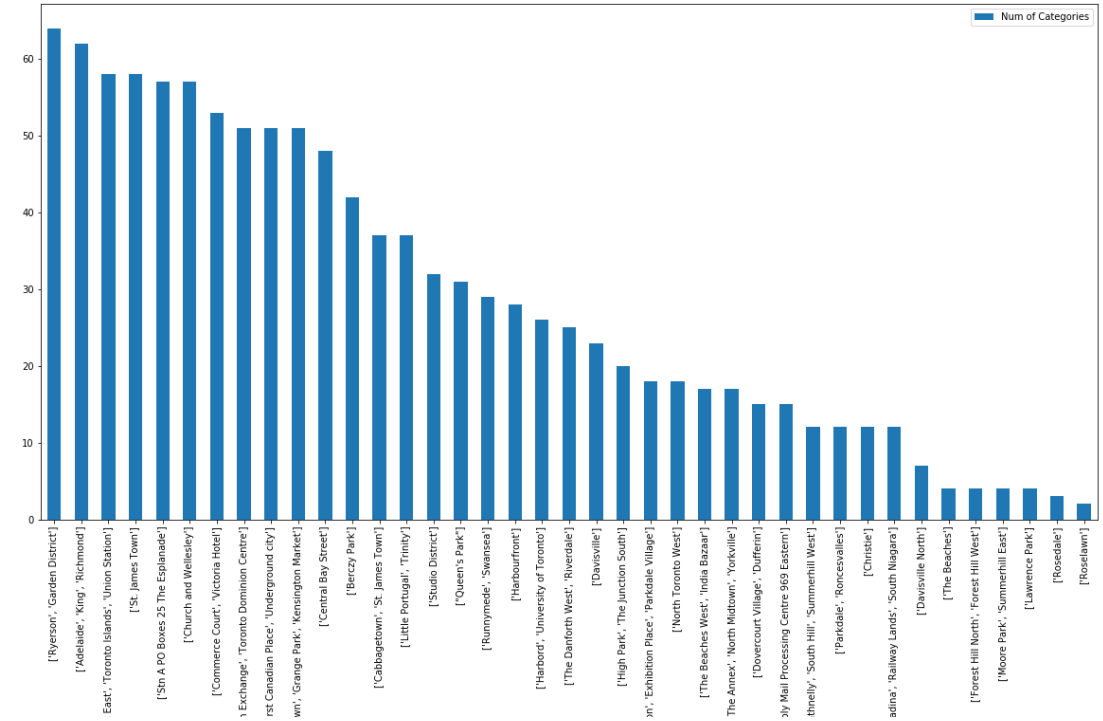
The hist figure describes the frequency of neighbourhood, it is clear that the neighbourhoods which have the most venues neighborhoods which are at the top and the bottom are the most. So it tells us that the neighborhood can have a feature of cluster, it is suitable for us to use the K-Means algorithm

to do later research.



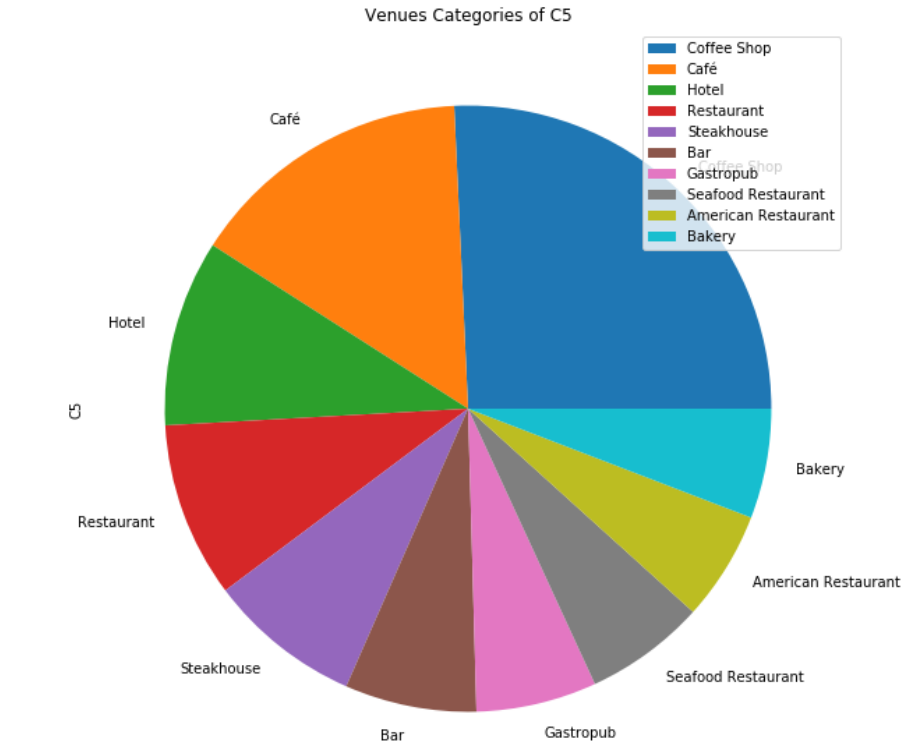
3.3. The top 10 Neibourhood which have the most Venues Categories

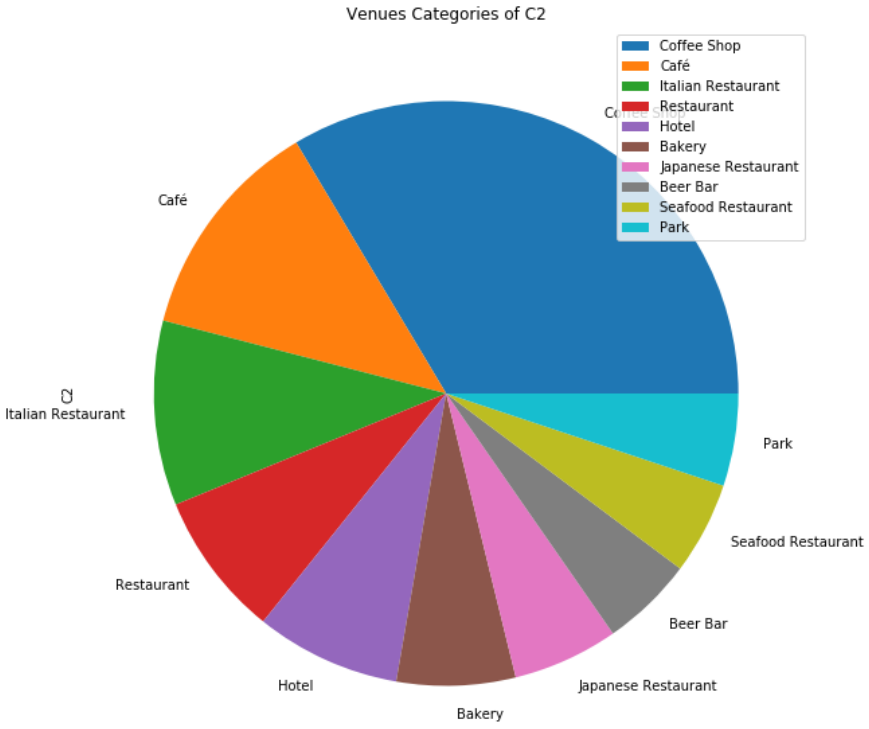
This figure told us that there are different categories in this area rather than many venues of the same category. So it guarantee us to do a meaningful job.

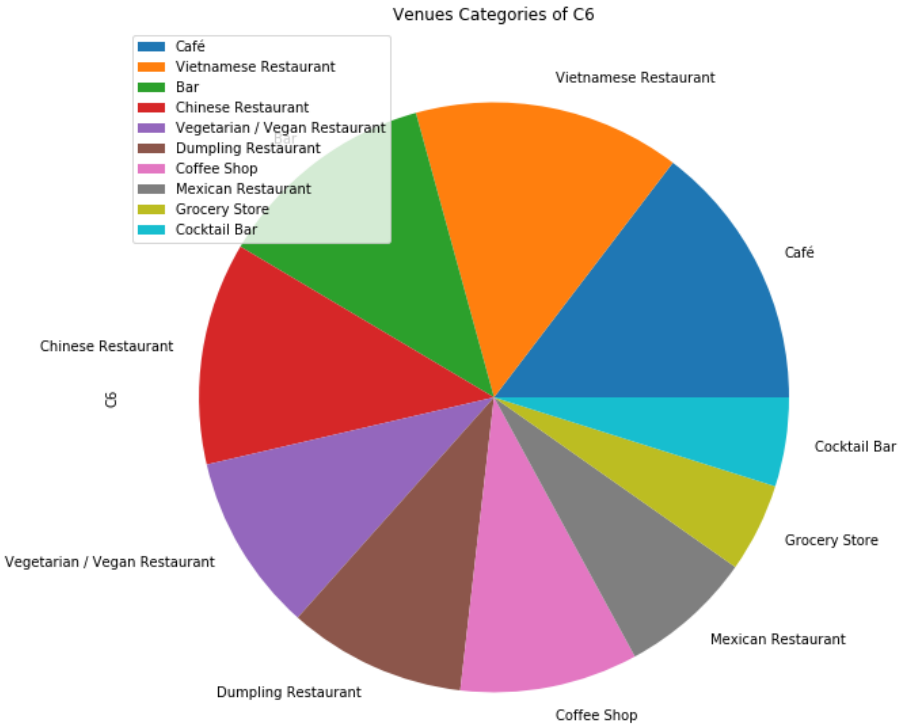


4. Results and Discussion

We cluster the data in 6 clusters, Because 6 Neighborhoods has the same most Venues Categories and the top 3 clusters are C5, C2 and C6







Discussion:

In fact, we can find that Coffee Shops and Cafe are popular in all 6 clusters, Hotel is easier to find both in C5 and C2. If tourists want to go to the park, they are more likely to consider C2, and for those who like Chinese food including dumplings, C6 can be a better choice. Seafood Restaurant can be easier to find in C5 and C2. Overall speaking, C6 is more suitable for tourists to have dinner while C5 and C2 can be easy to find a hotel to stay.

5. Conclusion

The Best Neighborhoods:

especially for those want to live in hotel and eat seafood restaurant

Neighborhood	Group
['Adelaide', 'King', 'Richmond']	5
['First Canadian Place', 'Underground city']	5
['Design Exchange', 'Toronto Dominion Centre']	5
['Commerce Court', 'Victoria Hotel']	5

The Second Best Neighborhoods:

especially for those want to live in hotel and eat Japanese restaurant and go to park

Neighborhood	Group
['Central Bay Street']	2
['Harbourfront East', 'Toronto Islands', 'Union Station']	2
['Stn A PO Boxes 25 The Esplanade']	2
['St. James Town']	2

The Third Best Neighborhoods:

especially for those want to eat Chinese Restaurant, and this Neighborhood is also popular with other food

Neighborhood	Group
['Chinatown', 'Grange Park', 'Kensington Market']	6