**Neighborhood Recommendation Analysis of a New Chinese Restaurant in Downtown Toronto**

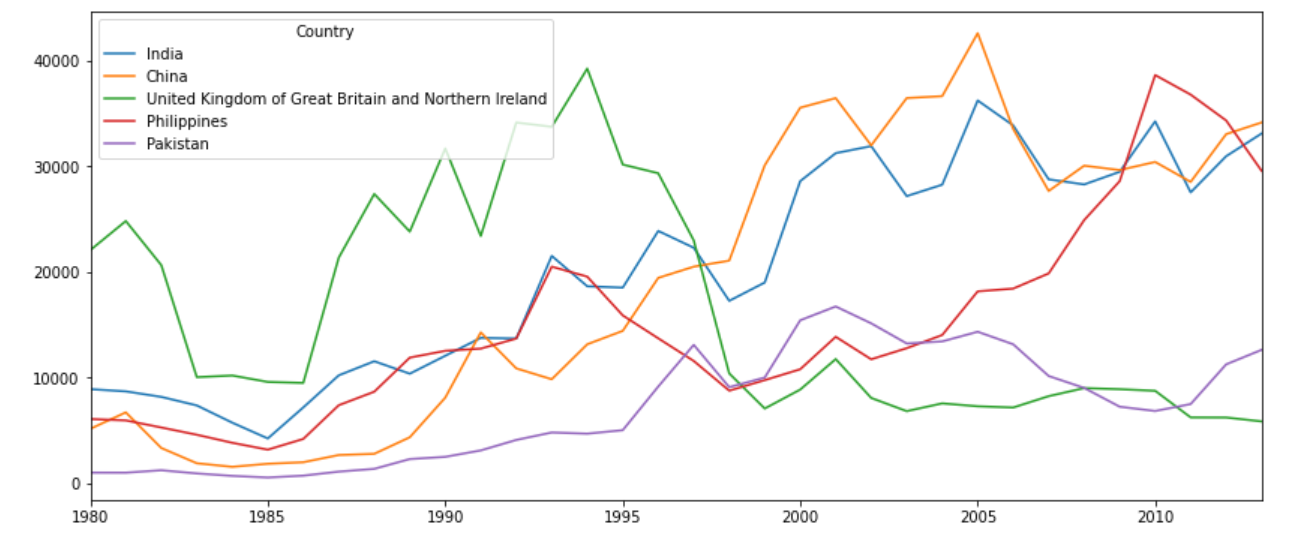
**Anthony Yang**

**IBM Data Science Professional Certificate Capstone**

**Introduction to business problem and who would be interested in this project:**

This is a real life issue, my Chinese uncle and aunt sent their kids to universities in Canada, and after the graduation, their son and daughter choose to stay and work in Toronto Canada, with the missing of their kids, my uncle decide to move to Canada to spend more time with their kids, both my uncle and aunt is in their 50s, they still want to make a live on their own, so they are thinking about open a Chinese Restaurant in Downtown Toronto, where their kids work and live. They are discussing about this at family’s spring festival dinner, so I think maybe I can do some exploratory research with this topic using data science skills I learned on Coursera.

Not just my uncle’s family faced this problem, with the immigration wave from China to Canada, especially the older generation at their 40s or 50s, they can speak only limited English and have little experience living abroad, I believe open a Chinese Restaurant is a good choice to make a living on their own. And where to open a new Chinese restaurant is a realistic question. In fact, I drew a line graph using the immigration data to show the Top 5 countries with most immigration people to Canada, you can see China is No.1 after 2000s.



Number of people from Top 5 countries immigrate to Canada

Data Source: [International migration flows to and from selected countries - The 2015 revision](http://www.un.org/en/development/desa/population/migration/data/empirical2/migrationflows.shtml) by United Nations

**Data that will be used to solve the problem and the source of the data:**

The data will be used in this exploratory analysis include:

1. Data of nearby venues of all the neighborhoods from downtown Toronto, includes “venue name”, “venue type”, “venue geographic coordinate” from FourSquare.com
2. Data of ratings of all the Chinese restaurant venues from FourSquare.com
3. Use geopy library to get the latitude and longitude values of Downtown Toronto

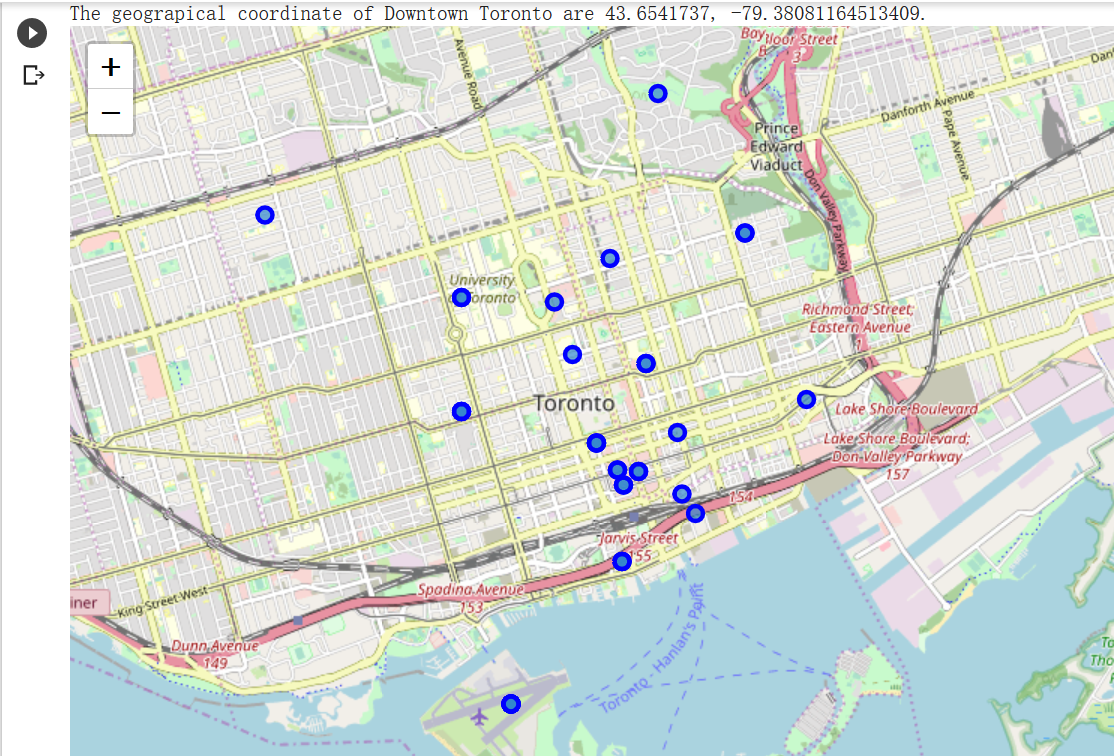
**Exploratory Analysis Methodology:**

1. Firstly, as I never been to Toronto before and basically know nothing about it, I used a machine learning modeling k-means clustering to cluster the different neighbors of Downtown Toronto, which gives me some whole picture of the difference between different Downtown Toronto’s neighborhoods.
2. Secondly I use the data from Foursquare to get the Chinese restaurant information of different neighborhood and analyses them to know about the location of these restaurants, location information can means the market saturation of this neighborhood and the peoples dietary habit of this neighborhood.
3. Finally I used the data from Foursquare to get the rating of these restaurants, rating can mean the competition level of certain area and also the peoples degree of satisfaction of the existing restaurant

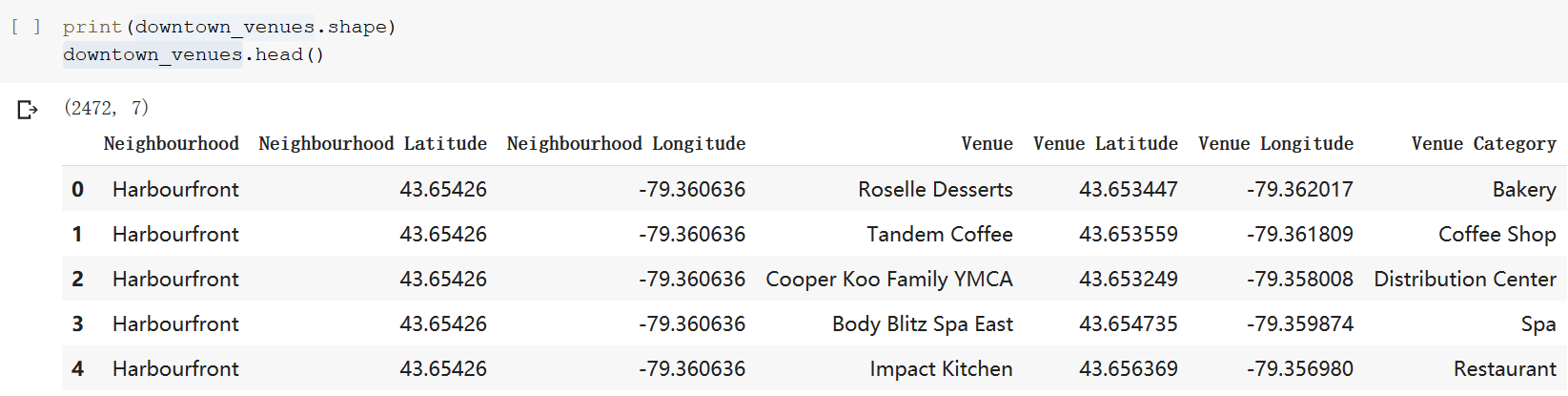
**Results section:**

Here is the process and the result of K-means clustering:

1. After some data wrangling we get the neighborhood of Downtown Toronto:



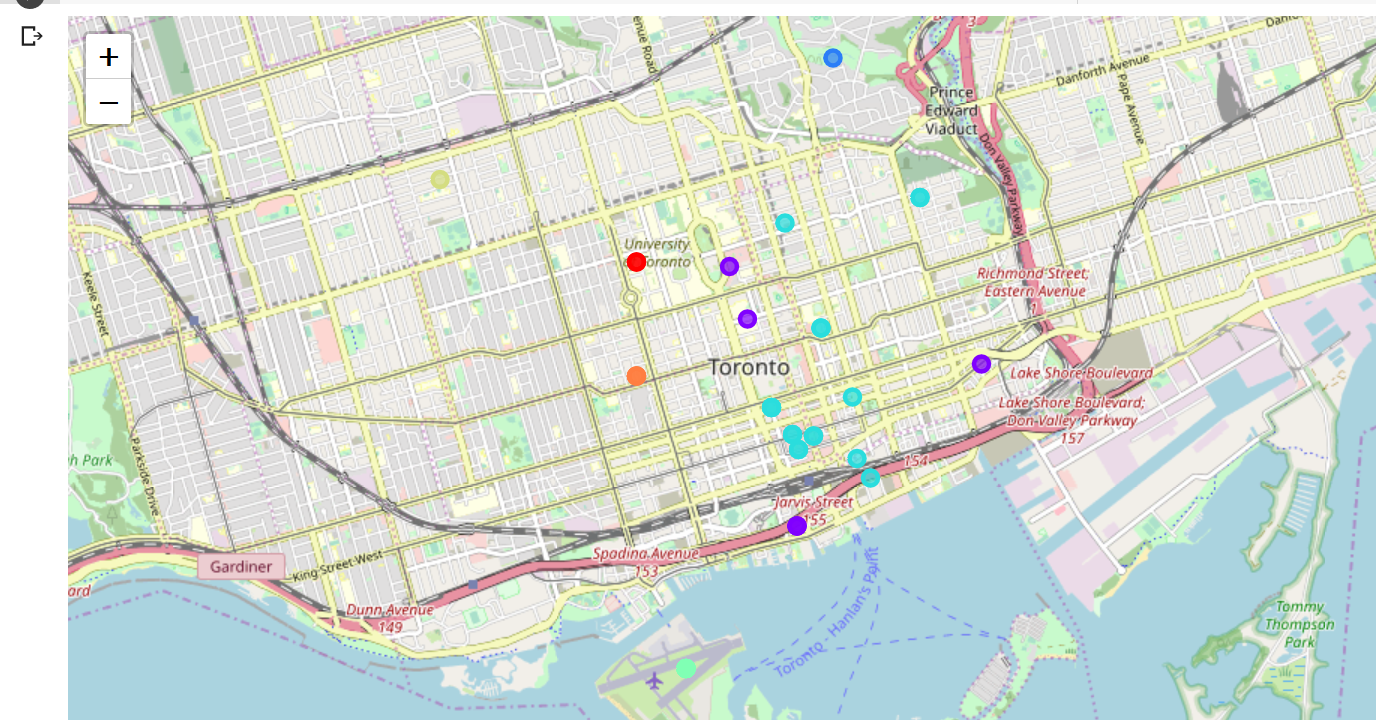
1. Then we use the data from Foursquare to get the venues from downtown Toronto:



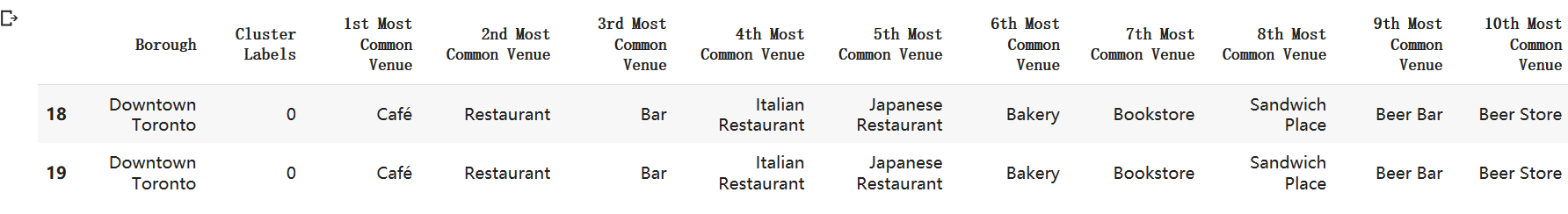
1. Then use some one-hot encoding technique we can get the frequency of different venue category in different Neighborhood and top 10 venue category of different Neighborhood



1. Then we use K-means clustering to get the cluster of these neighborhood, after several test, I set the kclusters = 7, with folium visualization tool I get the visualization of these clusters on map



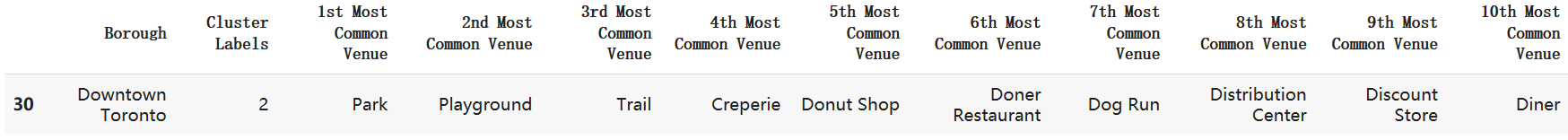
1. Then I got the top 10 venue categories of each clusters to understand the pattern and why ML will cluster these neighborhoods together
   1. Cluster 1:



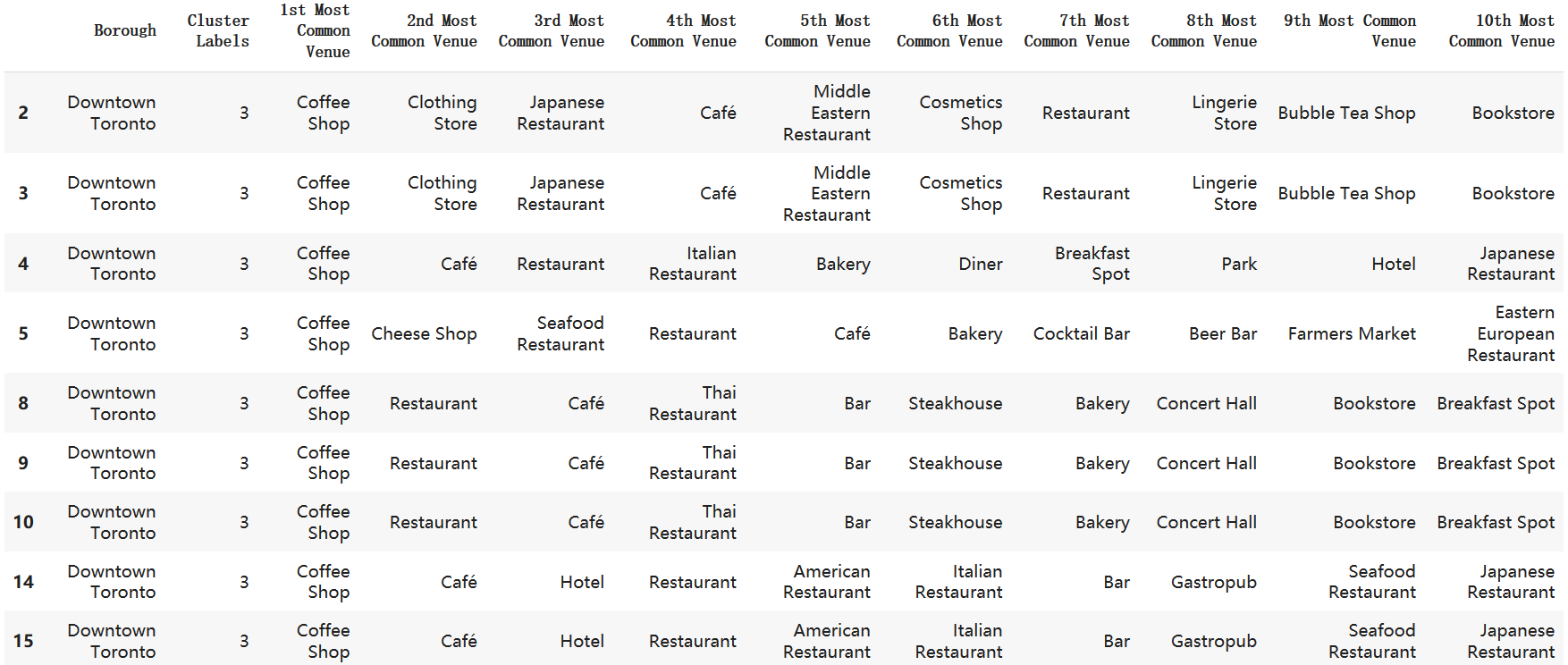
* 1. Cluster 2:



* 1. Cluster 3:



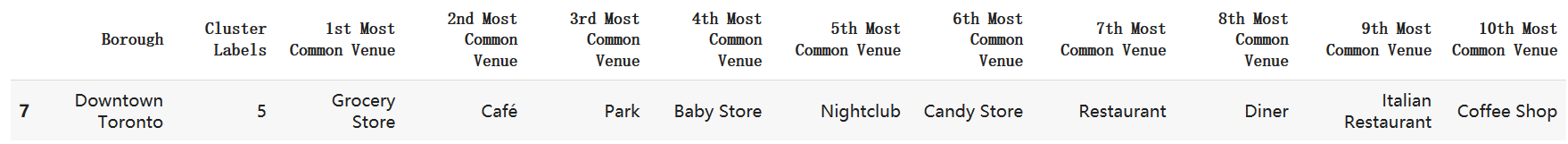
* 1. Cluster 4:



* 1. Cluster 5:



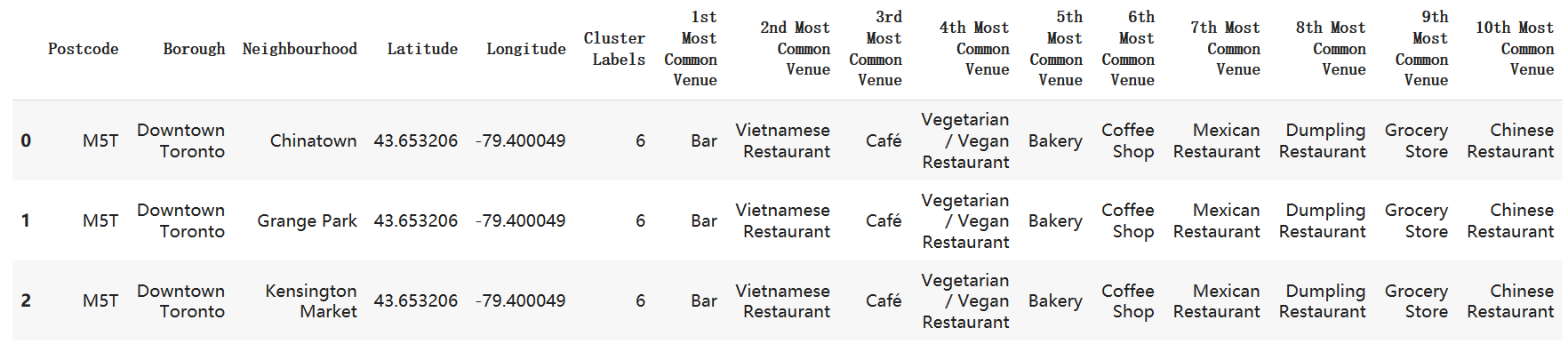
* 1. Cluster 6:



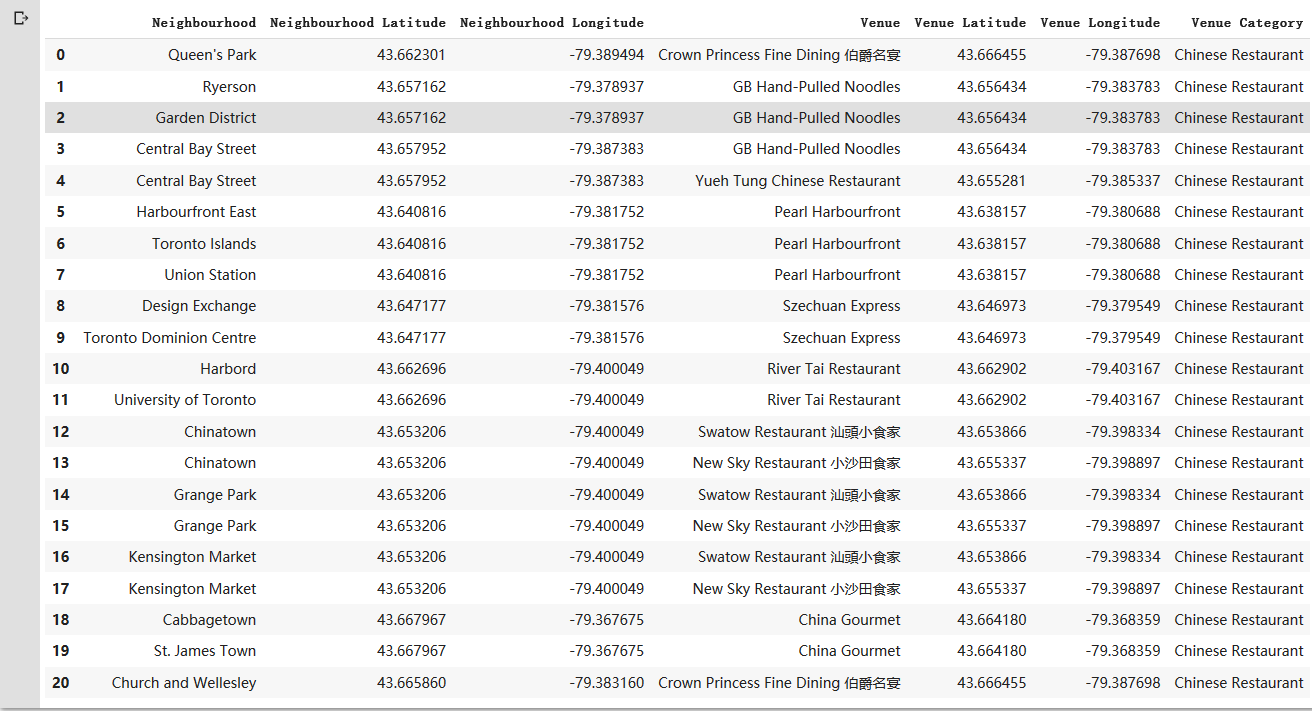
* 1. Cluster 7:



So as you can see , Cluster 7 is different from other clusters cause the Chinese Restaurant is belong to the Top 10 common venues here, which is the orange point on the map, close to Chinatown and reasonable, here we get a whole picture of neighborhood in Downtown Toronto, just from the ML result we can consider neighborhoods in Cluster 6 as a potential location: Chinatown, Grange Park, and Kensington Market



1. Then we do some deeper research about all the Chinese Restaurant in Downtown Toronto



There are 20 different combinations between neighborhoods and Chinese Restaurant, and there are totally 9 Chinese restaurant in Downtown Toronto, some neighborhoods are expose to 2 Chinese Restaurant and others only expose to 1, the neighborhood not in this chart is expose to 0.

1. Then we got the ratings of these restaurant from Foursquare, here is the result, I assume the rating above 7.5 is Good, 7.0-7.5 is medium and below 7.0 is not good



**Discussion section:**

Here are some observations and recommendations

1. Not every neighborhood in Downtown Toronto has Chinese Restaurant, let’s assume there is a reason: the citizen lived there don’t like Chinese Food, and maybe there used to be a Chinese Restaurant there but closed because of unsuccess business, so we only look at the neighborhoods with existing Chinese restaurant here
2. There are several Neighborhoods have 2 Chinese Restaurant and rating Good, like: Central Bay Street, Chinatown, Grange Park and Kensington Market, this indicates the neighbors there like Chinese food, and also the competition level is pretty high, so the recommendation for these neighborhoods is if the people want to open Chinese Restaurant have faith in their cooking and maybe have a successful restaurant back in China, these neighborhoods are recommended
3. There are also several Neighborhoods only have 1 Chinese Restaurant and rating not good, like Cabbagetown, St. James Town, Harbord and University of Toronto, so the recommendation is if the people who want to open a Chinese Restaurant don’t have much experience before and only want to have a try, these neighborhoods’ competition level is relatively low and people who lives there don’t really happy with the existing Chinese Restaurant.

**Conclusion section:**

So in conclusion,

1. The neighborhoods recommended for people who have successful experience before and prefer the potential market size to open a Chinese Restaurant is Central Bay Street, Chinatown, Grange Park and Kensington Market.
2. The neighborhoods recommended for people who don’t have much successful experience before and prefer the low competition environment to open a Chinese Restaurant is Cabbagetown, St. James Town, Harbord and University of Toronto.