**Location optimization for establishing a new Chinese restaurant in Toronto**

Quan Jinji

April 12, 2021

**1. Introduction**

* 1. **Background**

Toronto is the capital city of the Canadian province of Ontario. With a recorded population of 2,731,571,it is the most populous city in Canada and the fourth most populous city in North America. There are 299,460 Chinese People live in City of Toronto. Which makes 11.1% of the total population. Also the Chinese population has a tendency to grow continuously. Thus, It is a big business opportunity for a Chinese restauranteur to open a Chinese restaurant in the city of Toronto.

* 1. **Problem**

The problem I want to solve is to help a Chinese restauranteur to find a optimal place in city of Toronto for establishing his restaurant. As a invester, It is always important to find a place where has the lowest risk and highest probability of success. In this project, I will scan throughout the Toronto neighborhood, classify the areas with the Chinese population,number of restaurants，median income.

* 1. **Interest**

This project is targeted to investors who would like to establish a new Chinese restaurant in city of Toronto.

**2. Data acquisition and cleaning**

**2.1 Data sources**

#### In order to get data such as Chinese population,median income,latitude and longitude, number of Chinese restaurant. I will scrap and get data from following:

1.City of Toronto Neighbourhood Profiles   
<https://open.toronto.ca/dataset/neighbourhood-profiles/>

2.Latitude and longitude of neighborhood in Toronto can be gotten through usage of python library Geocoder. Refer to the following link for more information.  
<https://geocoder.readthedocs.io/>

3.Foursquare API to explore the famous venues especially Chinese restaurant in the neighborhood of Toronto. One will need to register his developer account from the following URL in order to access to Foursquare API.  
<https://developer.foursquare.com/>

**2.2 Data cleaning**

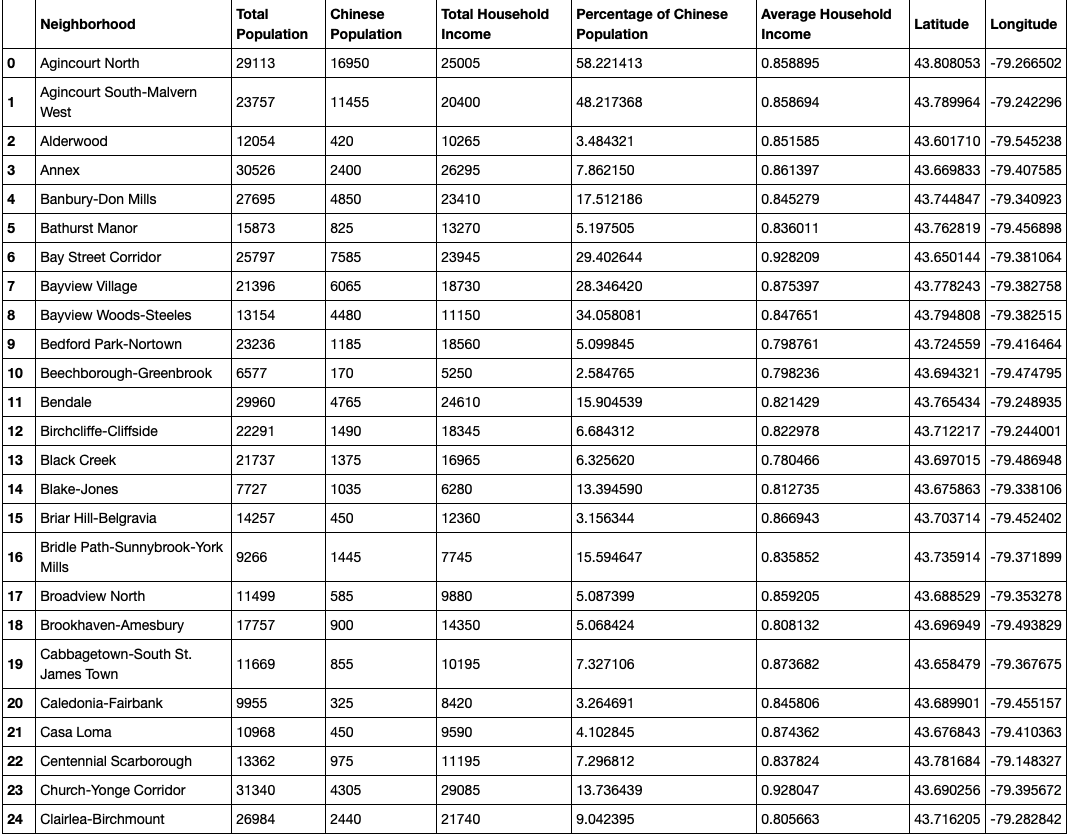
****

#### From 2383rows and 141columns I chose row[3],[1602],[945] to see Neighborhoods’ total population; Chinese population; total income in Toronto City.

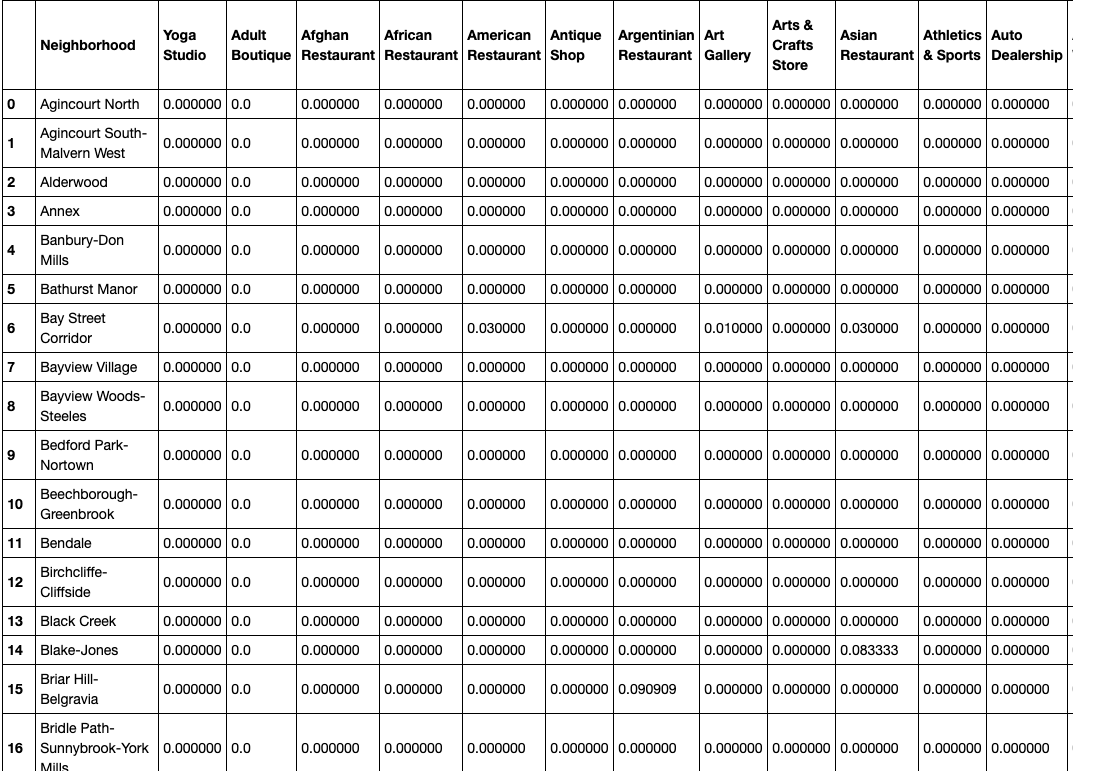
#### Then put these data into a new Dataframe-neighborhood.

**3. Methodology**

**3.1 Geocoding API and Foursquare API**



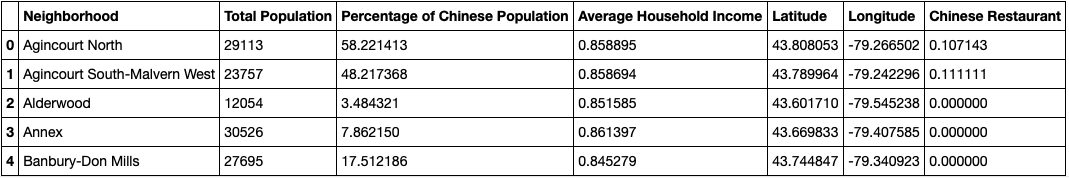
#### Using Geocoding API to get the Neighborhoods’ latitude and longitude. Insert into the Dataframe.



#### With the Foursquare API we can also retrieve all the venues in each Neighborhoods.

#### We are going to only focus on the Chinese Restaurants.



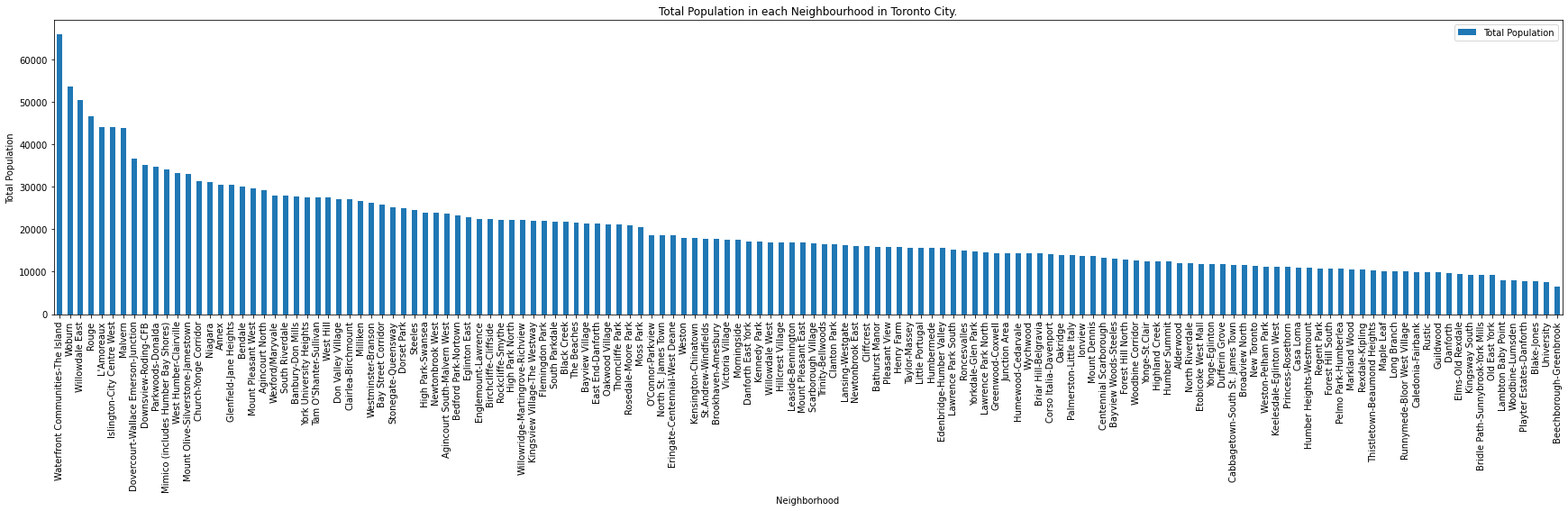
Add it to the Dataframe-neighborhood.

**3.2 Data Visualization**

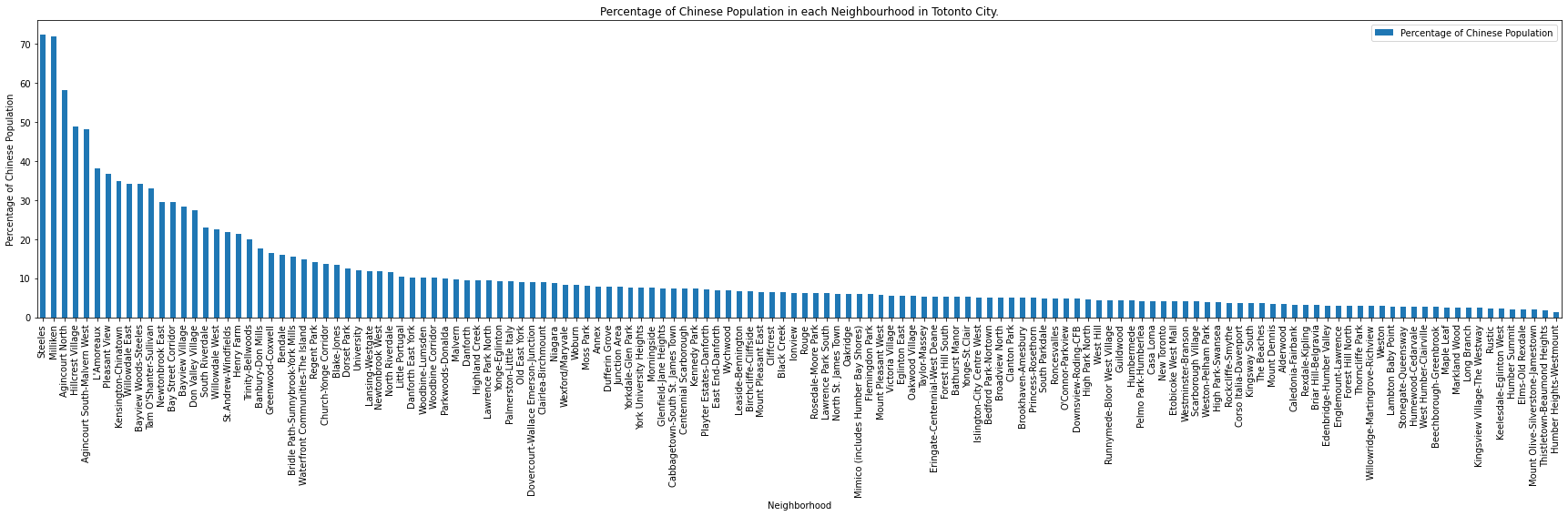
Mean Frequency of Chinese Restaurants in each Neighbourhood in Toronto City.



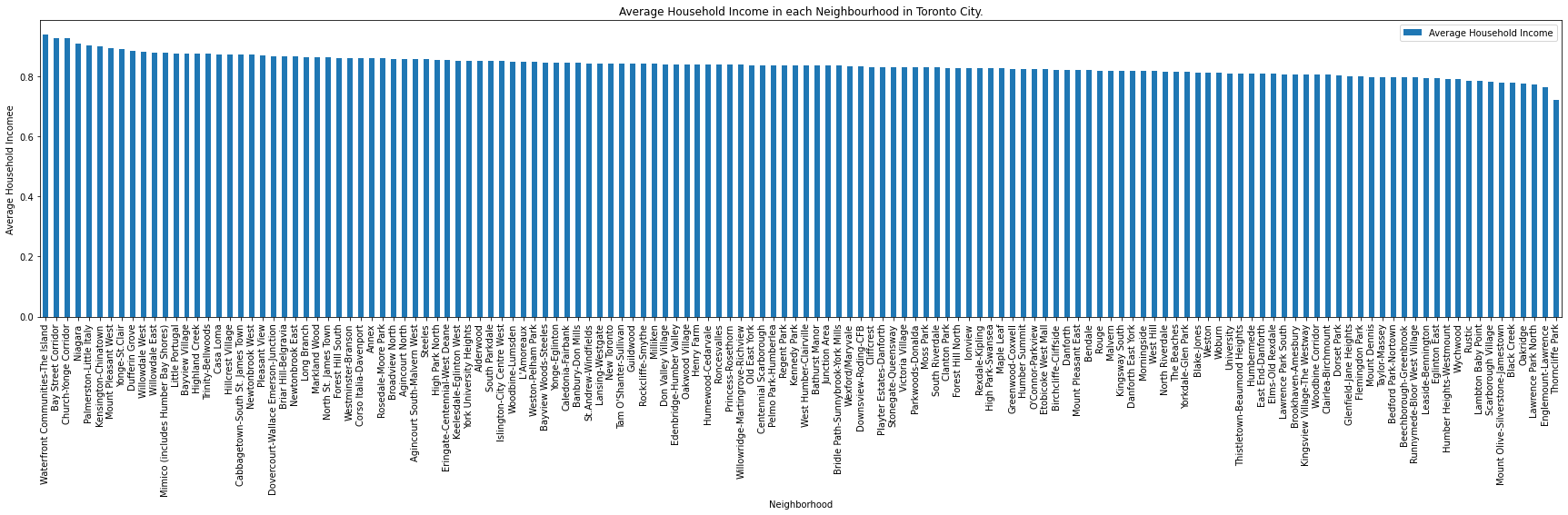
Total Population in each Neighbourhood in Toronto City.



Percentage of Chinese Population in each Neighbourhood in Totonto City.

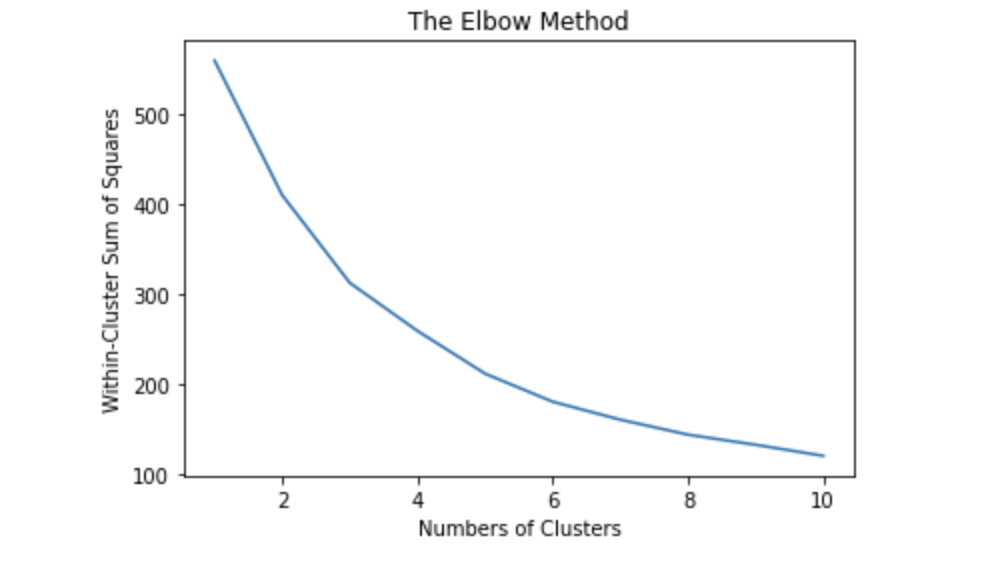


Average Household Income in each Neighbourhood in Toronto City.



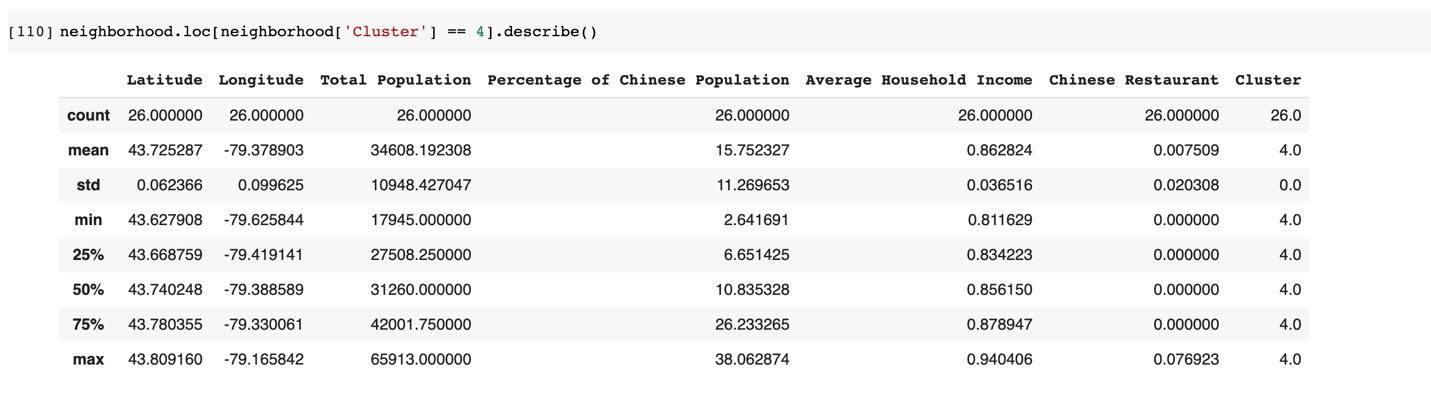
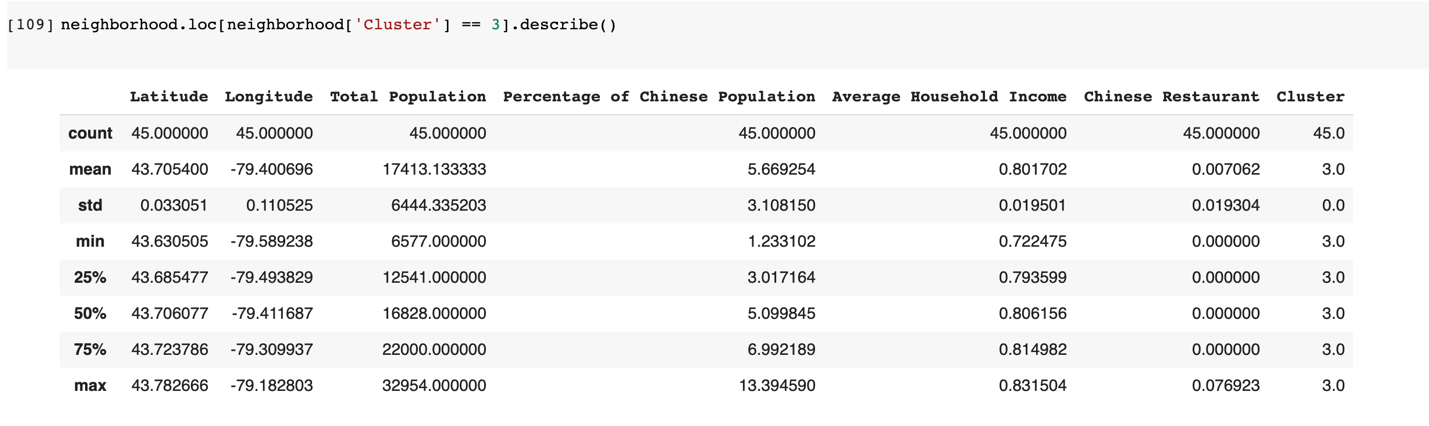
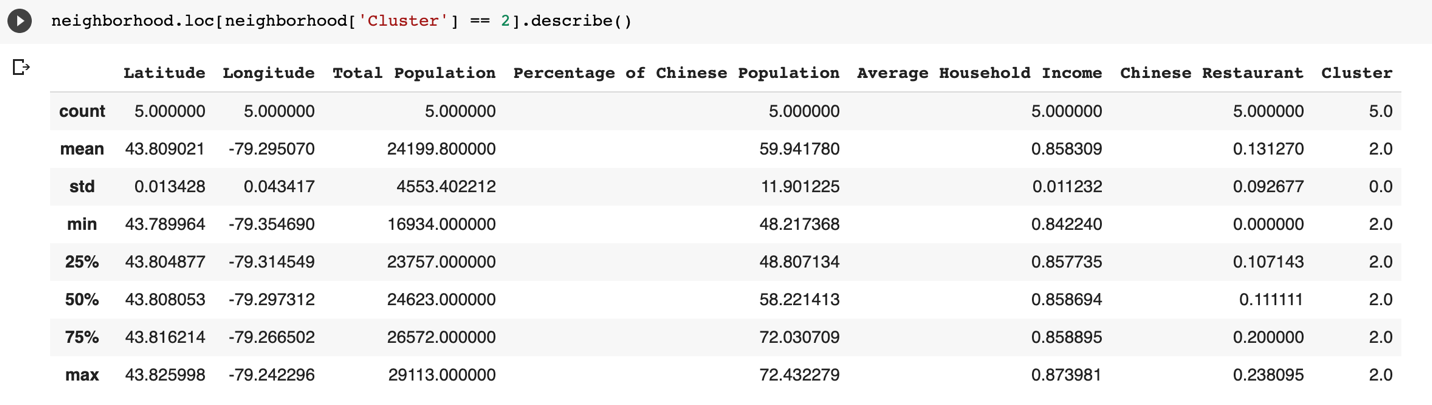
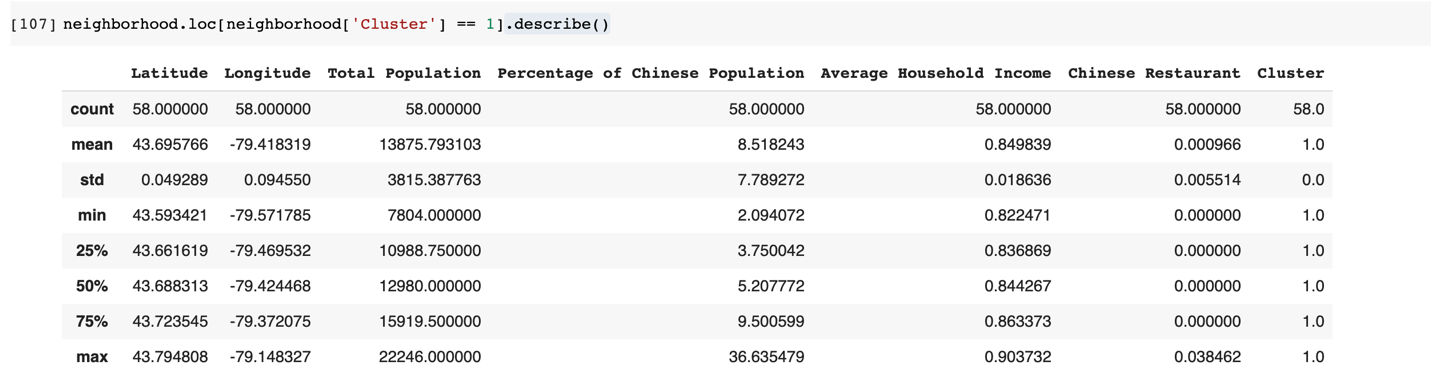
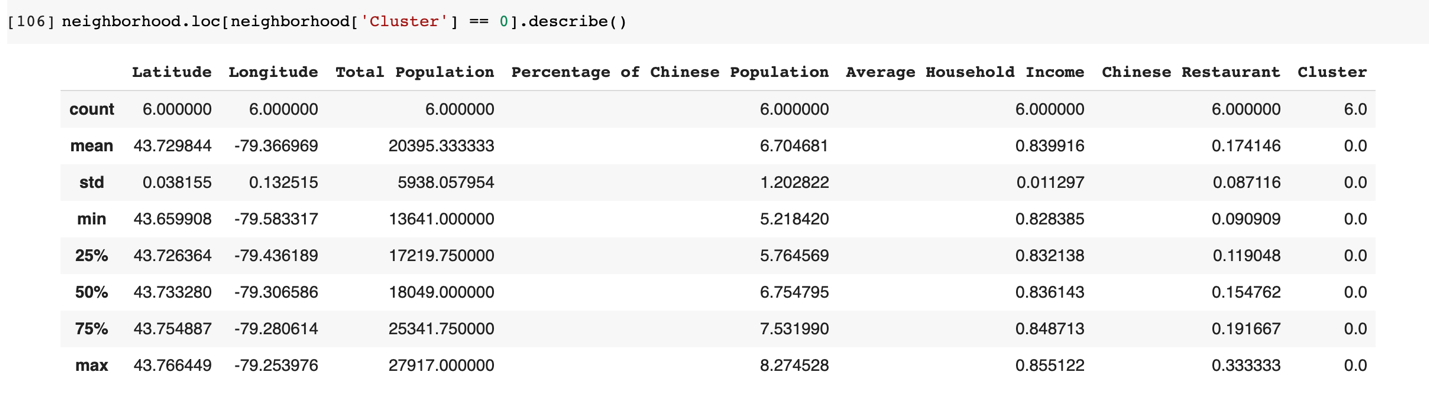
**3.3 Clustering**

The first step is Feature Scaling for independent variables. Then use K-means clustering algorithm, as the picture shows, the best k is 5.

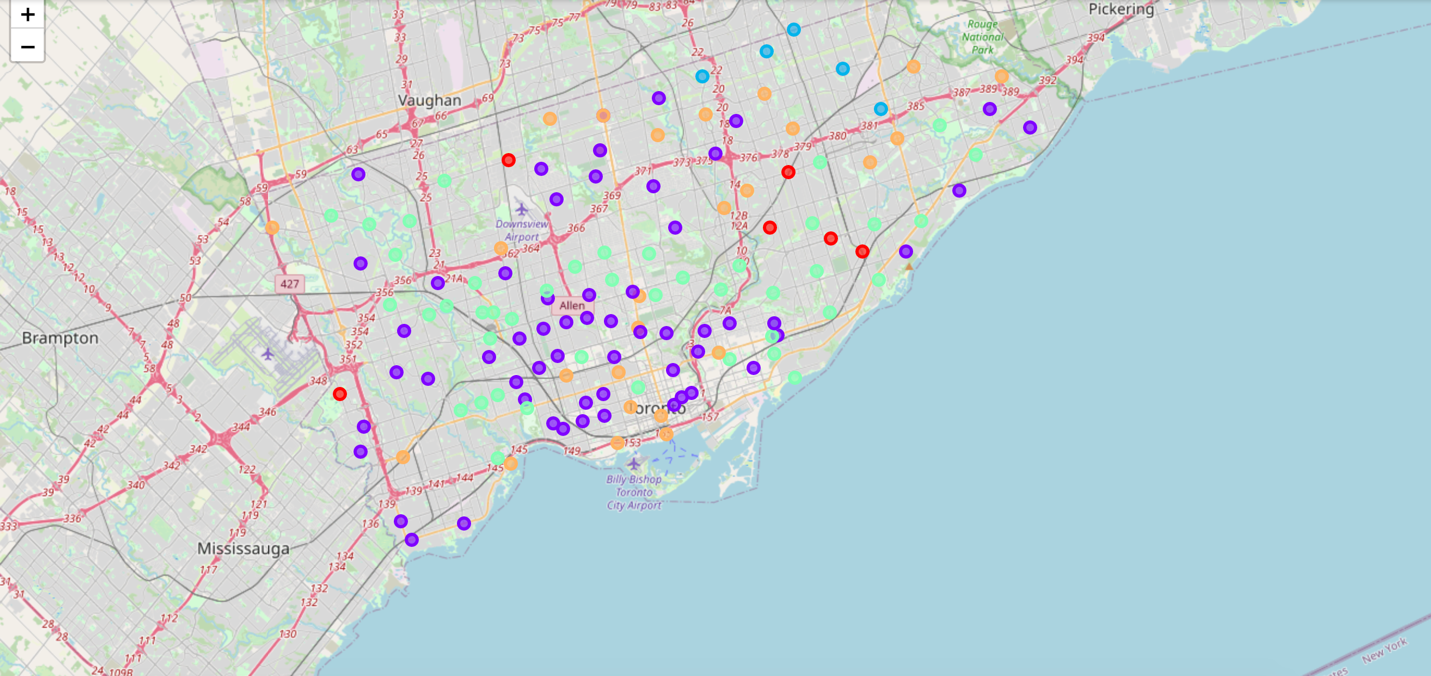


**4. Results**

The Neighborhoods are separated into 5 cluster groups shown below.

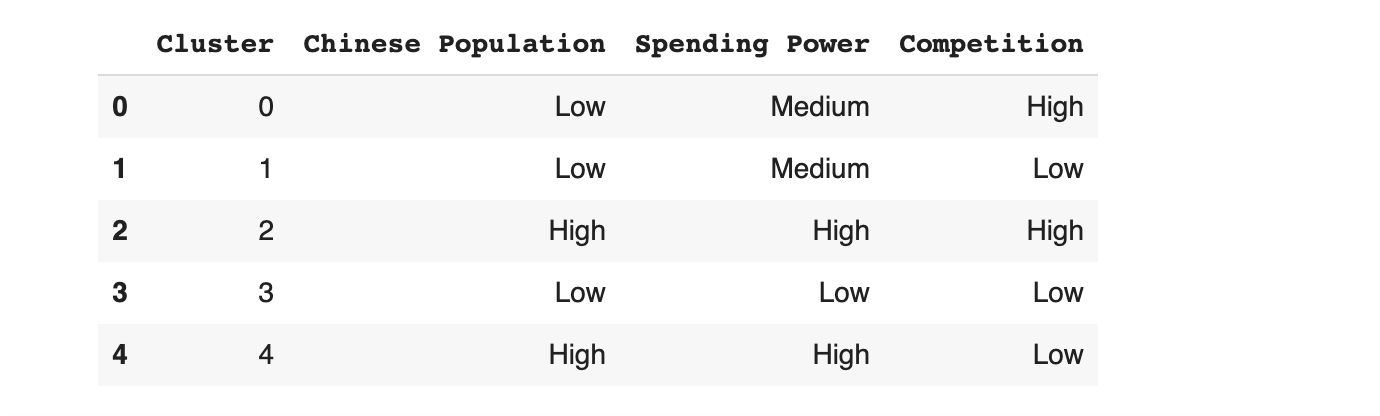
****

By using Folium, a map showing Toronto City can be generated with different cluster indicated by different colors.

****

**5. Discussion**

The analysis above is summarized in a table below.



For leading business to success: Chinese Population should be as high as possible, as well as the Spending Power, Competition should be as low as possible.

**6.Conclusion**

As a result, cluster 4 is to be prioritized by the invester who is going to eastablish a Chinese Restaurant in Toronto City. Accoring to the Map, north-east and south are the best place in Toronto to start a new Chinese Restaurant.