Opening a new Vietnamese restaurant in New York City

Course: Capstone Project - Data Science Specialization

Student: Tra Nguyen

Date: 1st February 2019

Introduction/ Business problem

Asian cuisines recently become more and more popular in the States. Considering it as a business opportunity, Mr X., who owns a very famous Vietnamese restaurant in Washington DC, wants to open a new branch in New York City. However, he does not know in which neighborhood the restaurant would gain the most attention from customers and face the least competition from other restaurants. Therefore, I, as a data science student, would like to help him to find out the best neighborhood for his business by conducting an analysis using methods that I have learned throughout the course.

Data

There are 5 boroughs and 306 neighborhoods in New York City. For this project, a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the the latitude and longitude coordinates of each neighborhood, is used (here is the link to the dataset (https://geo.nyu.edu/catalog/nyu_2451_34572)
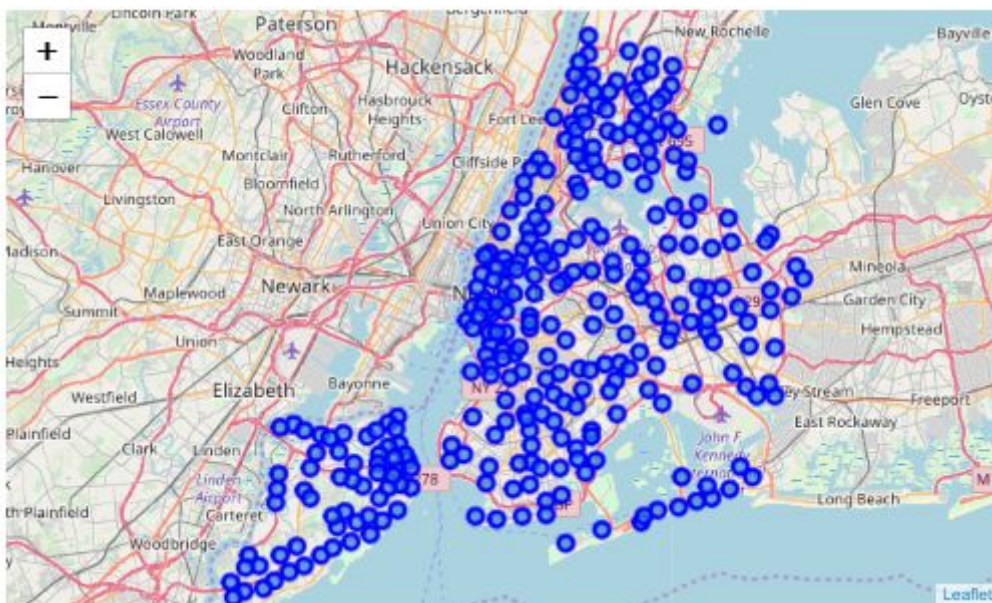

Figure 1: The map of New York City with neighborhoods superimposed on top

Then the Foursquare API is utilized to explore and segment the neighborhoods based on the density of existing Vietnamese restaurants in each neighborhood.

Methodology

First the dataset is downloaded, then converted into dataframe format with names of boroughs, neighborhoods, their corresponding latitudes & longitudes. The dataframe has 5 boroughs and 306 neighborhoods.
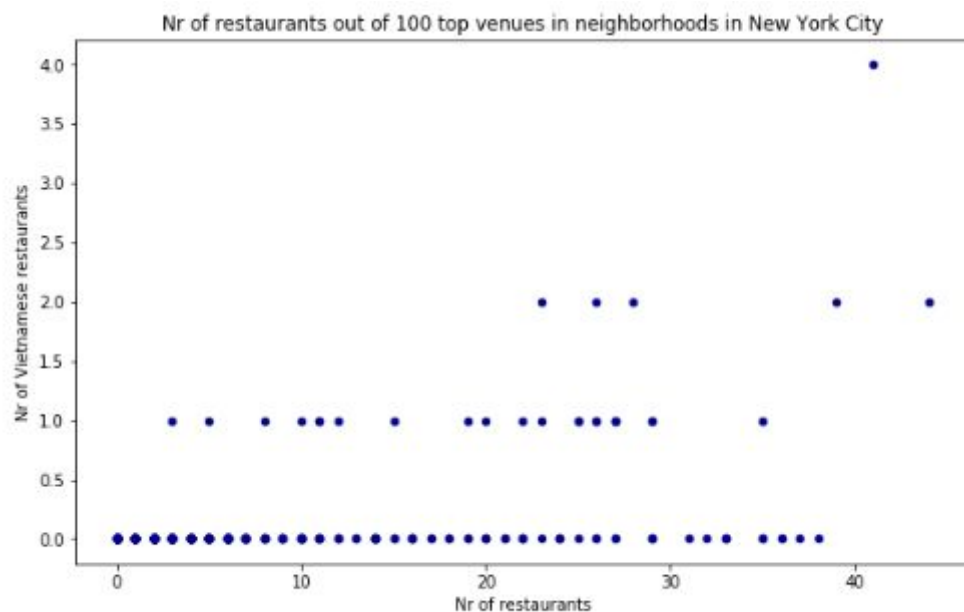
Then Foursquare API is used to get the list of top 100 venues in each neighborhoods (radius = 500m). Only information about the number of restaurants and the number of Vietnamese restaurants are extracted to create a new dataframe.

Finally the new dataframe is visualized using scatter plot by matplotlib. A list of recommended neighborhoods where the owner should open his new restaurant is also presented and highlighted with red circle marker on map using folium.

Results

	Neighborhood Latitude	Neighborhood Longitude	Nr of restaurants	Nr of Vietnamese restaurants
count	306.000000	306.000000	306.000000	306.000000
mean	40.701362	-73.946254	8.562092	0.117647
std	0.097498	0.121202	9.908983	0.419854
min	40.505334	-74.246569	0.000000	0.000000
25%	40.617778	-74.000062	2.000000	0.000000
50%	40.702969	-73.932009	4.000000	0.000000
75%	40.766645	-73.857525	11.750000	0.000000
max	40.908543	-73.708847	44.000000	4.000000

Figure 2: Basic statistic for the number of restaurants out of top 100 venues in 306 neighborhoods in New York City

Figure 3: Scatter plot of the number of restaurants out of top 100 venues in neighborhoods in New York City

As we can see from the scatter plot and statistic above, most of neighborhoods in New York city don't have any Vietnamese restaurant in their top 100 venues. Besides, in some neighborhoods, the competition among restaurants in general is fiercer than in the rest due to higher density.
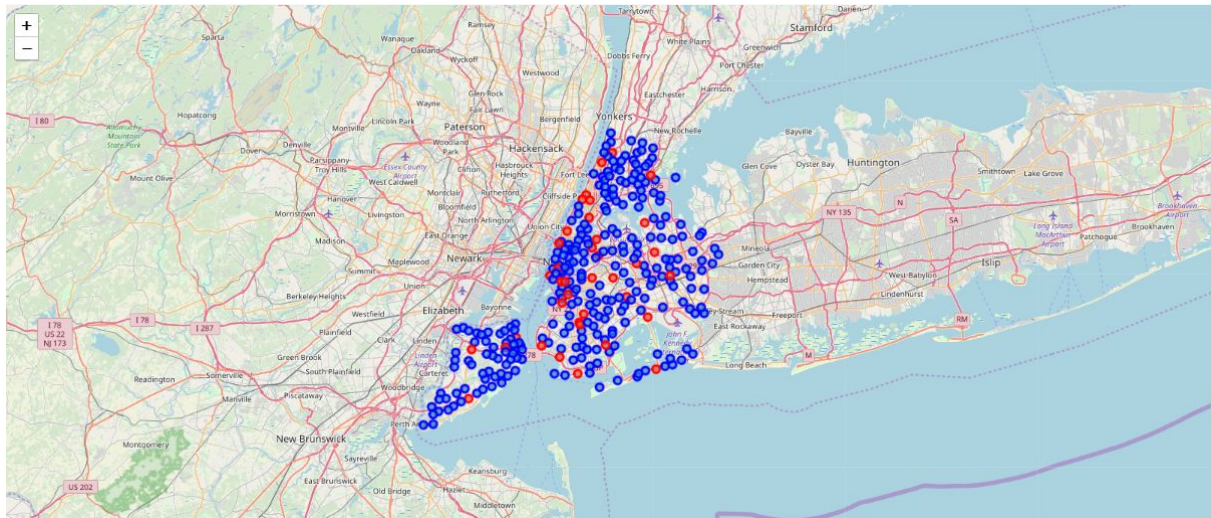
Figure 4: The map of New York City with neighborhoods superimposed on top. The red markers highlight the recommended neighborhoods where the owner should open his new restaurant

Discussion

Since many neighborhoods in New York City do not have any Vietnamese restaurant in their top 100 venues, It would be a business opportunity for Mr X. who wants to open a new branch here. However we should consider also the competition on the market by excluding all neighborhoods that already have too many famous restaurants (>30) and the ones that have too few restaurants as top venues as the need of eating out seems to be low (<10). In short, I would recommend a list of neighborhoods that have no Vietnamese restaurant and have from 10 to 20 other restaurants in top 100 venues.

The list includes 37 neighborhoods, namely 'Kingsbridge', 'Pelham Bay', 'Brighton Beach', 'Gowanus', 'Cypress Hills', 'Bath Beach', 'Boerum Hill', 'Prospect Lefferts Gardens', 'Prospect Park South', 'East Williamsburg', 'Fort Hamilton', 'Inwood', 'Hamilton Heights', 'Manhattanville', 'Central Harlem', 'East Harlem', 'Lincoln Square', 'Soho', 'Battery Park City', 'Howard Beach', 'Long Island City', 'Sunnyside', 'Ridgewood', 'Rego Park', 'College Point', 'Rockaway Beach', 'Queensboro Hill', 'Ravenswood', 'Stapleton', 'Eltingville', 'Ditmas Park', 'Bulls Head', 'Mill Basin', 'Jamaica Hills', 'Fulton Ferry', 'Dumbo', 'Hudson Yards'

Conclusion

In short, if Mr. X wants to open a new Vietnamese restaurant in New York City, he can choose one from 37 neighborhoods where there is no Vietnamese restaurant, the need of eating out is relatively high and the competition on the market is relatively low.