

Looking for new opportunities for a Chinese restaurant in New York City

Shuang Wu

09.06.2020

1. Introduction

1.1 Background

This project researches the optimization of Chinese restaurants in New York. Different branches of Chinese food have arrived in the U.S more than 150 years old firstly with the railway workers and early migrations, it has been popular ever since. According to an article from New York Times in 2018, more than 60% of Americans can apply chopsticks in their daily life. The project presumes that a group of investors are eager to involve into this old yet inspiring business behind a long history and colorful culture.

1.2 Data

The project extracts aforementioned N.Y.C neighborhoods data containing their Geolocations from https://cocl.us/new_york_dataset for its accessibility. N.Y.C neighborhood boundaries acquired from the data source, which is a pre-made json file: <https://gist.github.com/benjycook/b6ec3c7b5cf05f10ccdb11e3c3497d10> Other relevant data comes from Foursquare API and Libraries in Python.

1.3 Methodology

- The project collects from https://cocl.us/new_york_dataset then processes into a data frame.
- The project supplements its data frame by using Foursquare within its call limits, to locate and filter venues and its rating, tips and likes generated from users.
- Data will be sorted based on rankings
- Finally, the project visualizes the sorted data.

1.1 Objectives

- What is the best upcoming location for a Chinese restaurant in NYC?
- Amongst different options, which Neighborhoods will optimize the chance to have a successful restaurant?

2. Analysis

2.1 NYC Neighborhoods

The project starts with acquiring the neighborhood data from given public link that was provided in our courses. The columns sections are divided and cleaned into “Borough”, “Neighborhood”, “Latitude”, and “Longitude”. The data returns with 306 different Neighborhoods in NYC’s five major boroughs. The latitude and longitude of each neighborhoods help the project to clarify that which locations are good for the business. As the figure 1 presents, the analysis return with Chinese restaurants in each districts in NYC, in which, Queens have the largest number of the restaurants, Brooklyn second and Manhattan follows.

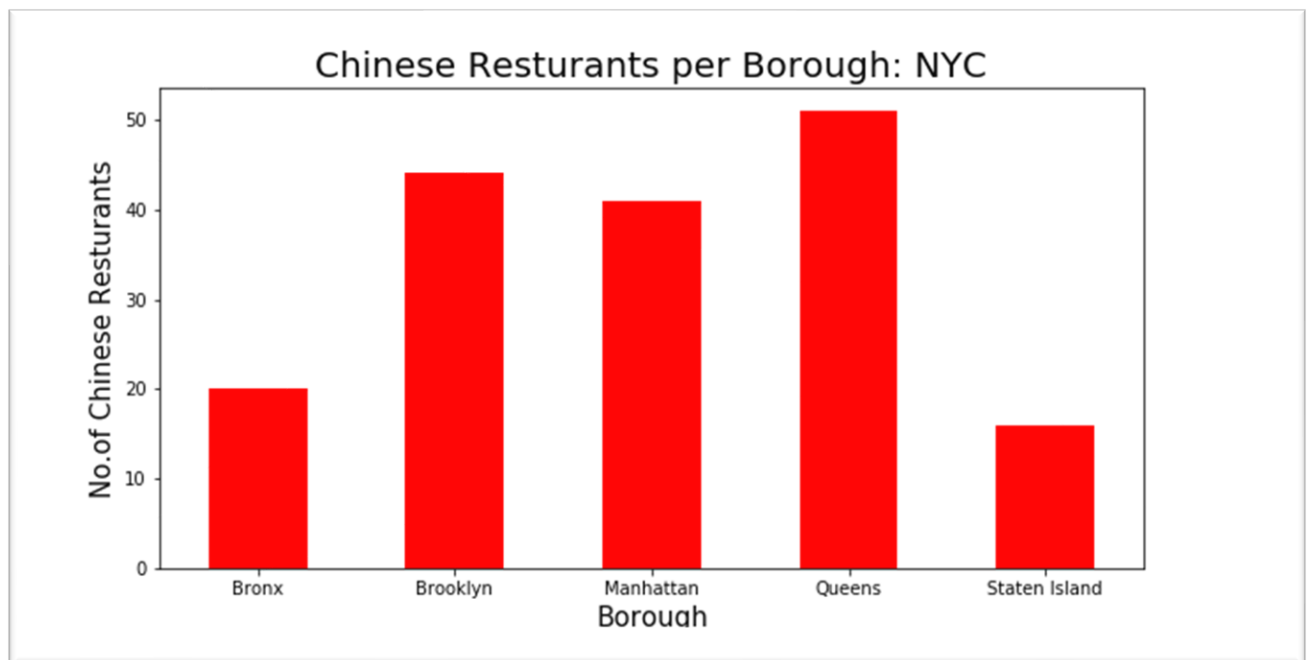


Figure 1. Chinese Restaurants per Borough in NYC

At next, for choosing the optimum location, the project intends to find out in which neighborhoods locate the most Chinese restaurants. As figure 2 shows, there is very little surprise

that the most Chinese restaurants are placed inside Chinatown, which given historical reasons, was the first large Chinese community in New York. But the question is, is China town our ideal location because its high concentration of data?

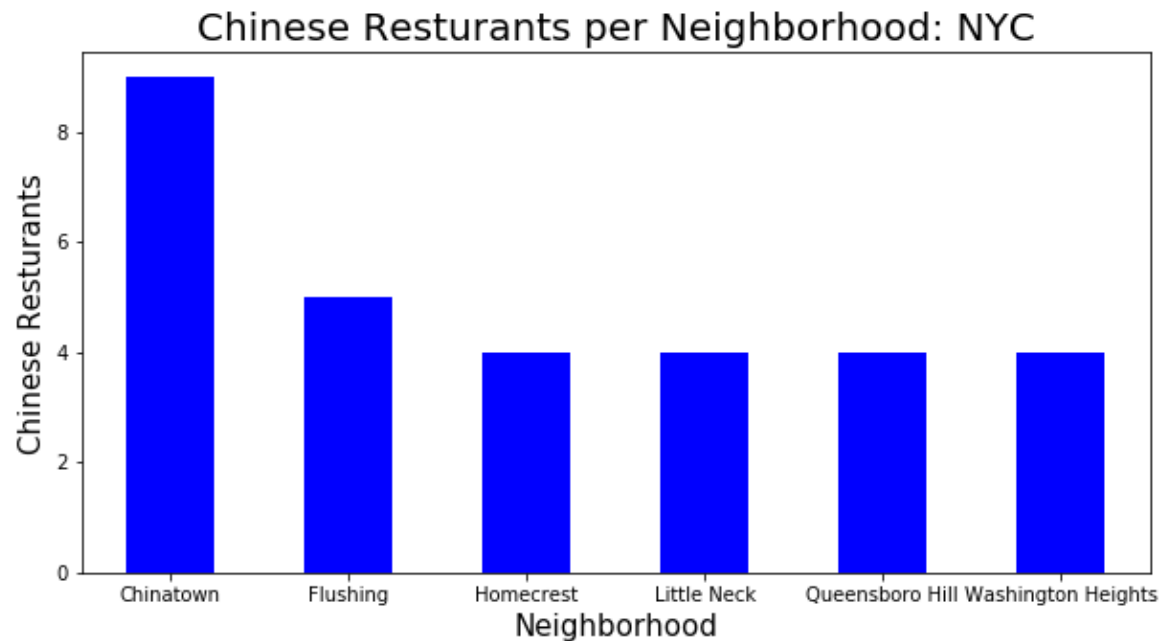


Figure 2. Chinese Restaurants per Neighborhood in NYC

2.2 Ratings

At this stage, the problem becomes clear, although we can not go for a district or a neighborhood because it is highly packed with the same category of restaurants. So, we turn in API calls to extra ratings. To be advised here is that, Foursquare although is free, it limits the call for rating, likes etc as premium feature, so we will have to deal with what we have here. So the codes returned with 172 rows of data with features that are intended to be found: Borough, Neighborhood, ID, Name, Likes, Rating and Tips. From these numbers we are able to discover the average rating of Chinese restaurants for each Borough in New York, as the figure illustrated below:

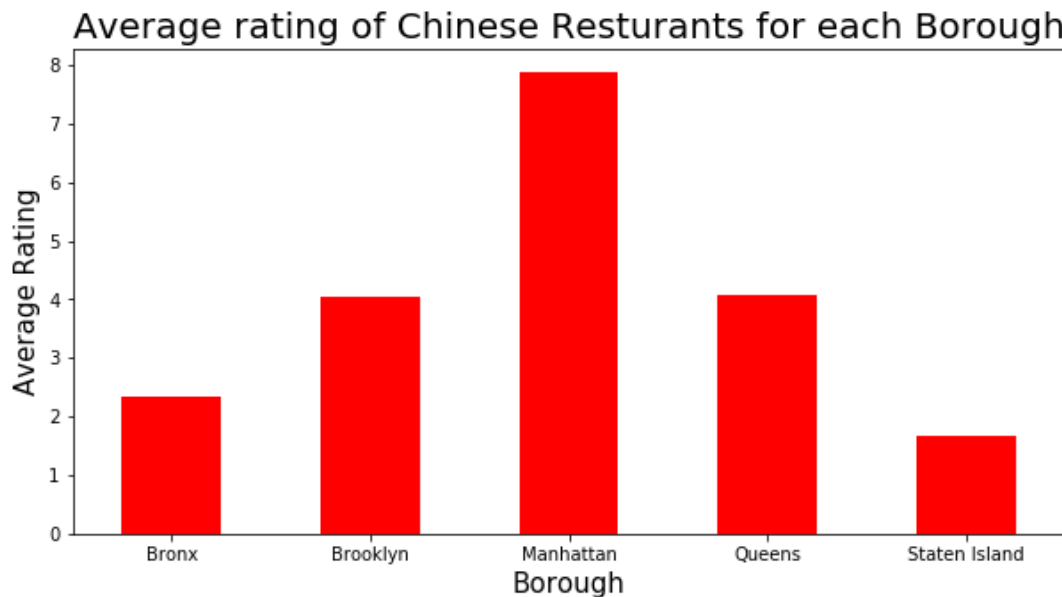


Figure 3 Average Rating

Therefore, it is obvious that Manhattan has the highest rate from our given data even though the number of restaurants is not the greatest. However, it is still recommended to take a look at a bigger picture: at here, we firstly draw out of all places that have the rating higher than 7.0 in town, then combine it with the neighborhood data, in the end the placement of restaurants and their geodata have matched.

2.3 Visualization

With returned data of each neighborhood, borough and their location and rating, now we could initiate the visualization in order to present a better view of the whole dataset. At first, the project pinpoints all the found restaurant locations with rating higher than 7.0 over the map of New York, then combining with rating data from the API calls, we have the result of this analysis. The reason we only employee high rating data is for the sake of clarification.

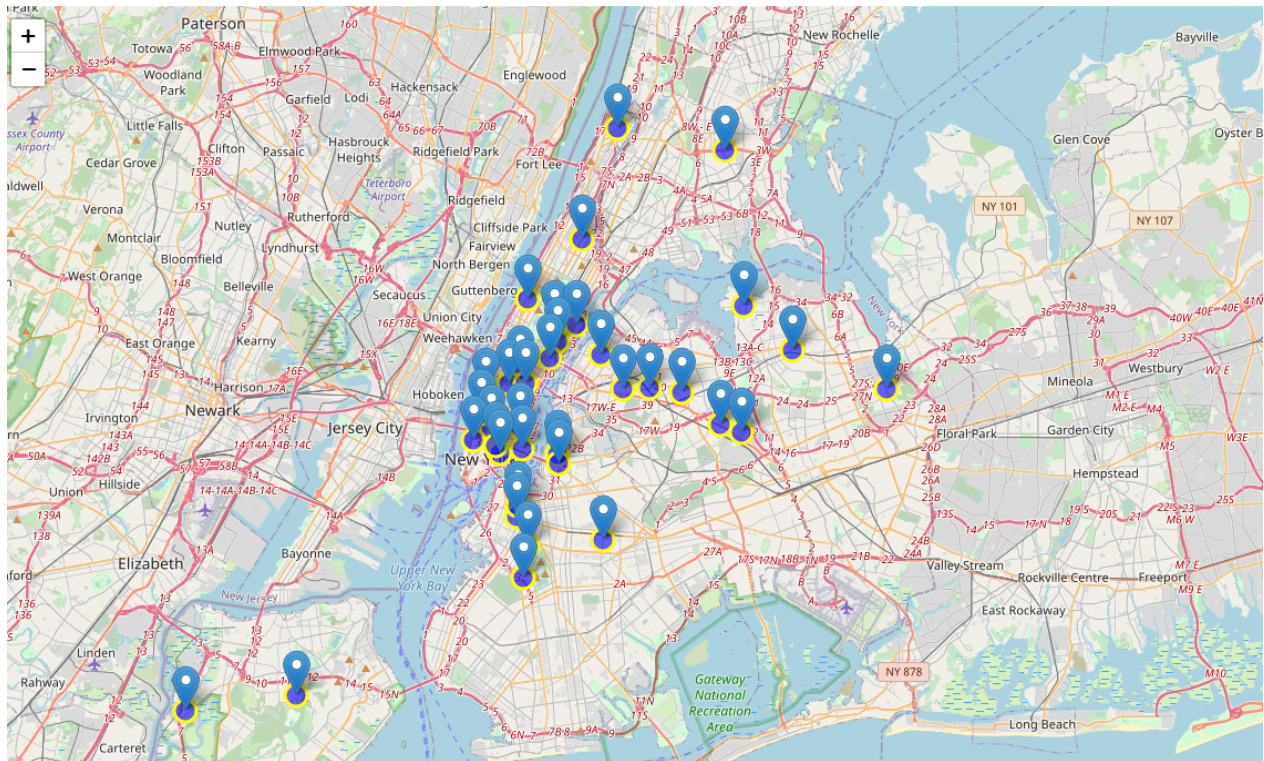


Figure 4 High rating establishments

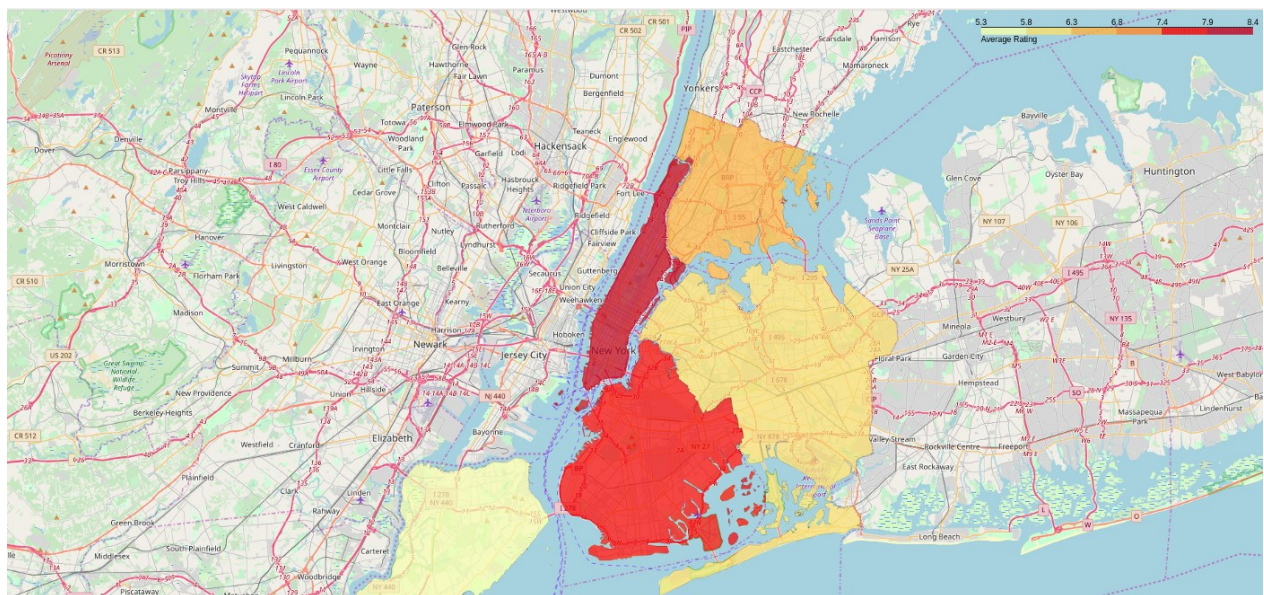


Figure 5 Rating by Colors

3. Conclusion and Findings

This project discovers that even though Chinatown is traditionally Chinatown is the place where Chinese restaurants collectively gather, however Manhattan has the most concentrated high-rating restaurants. Queens have the largest number of them. However, the rating is merely average. The project indicates that the investors to avoid Chinatown and Flushing neighborhoods as it is already a blue ocean there. Manhattan provides good opportunities for a business to success there. Brooklyn would be another potential option as the average ratings are between 4-5 out 10, the number of restaurants is lesser than in Queens.