

Coursera IBM Applied Data Science Capstone

Opening a Korean Restaurant in the New York City

Final Report

by

Muhammad Imran

December 2020

Introduction

New York city is the largest metropolitan area of the United States. It is a melting pot of culture s all around the world. The city's cultural pluralism is well reflected in its food culture. New York has one of the world's most dynamic and competitive restaurant scenes of the world. One report from earlier in the decade showed that 80% of New York restaurants close in the first five years. This exemplify the competitiveness of running a restaurant business in NYC.

This final project will explore the most suitable location for opening a Korean restaurant in NYC. Location is one of the most important factors behind the long term success of any business. It is even more important for a restaurant.

Business Problem

As it has been discussed in the introduction that location is one the most important aspect of opening a restaurant, an investor must answer this question before deciding to start her business. The investor must figure out which place has the most demand for a Korean restaurant. By employing different techniques of Data Science such as Data Analysis and Visualization,

this project will try to answer which area of New York city is the most suitable for opening a new Korean Restaurant.

Target Audience

Korean restaurants have been growing at 3% rate annually in the US. Since NYC is a cosmopolitan city its inhabitants have a wide ranging palettes and they prefer ethno cuisines over many fast food or traditional western choices. So there are many advantages of opening a Korean restaurant in New York city. As a result, this project will be useful for any investor who would like to leverage these advantages and invest in a Korean restaurant in NYC.

Data

There are 3 types of data that we will need for this project:

1. Data regarding the boroughs and neighbourhoods of NYC.
2. Latitude and longitude coordinates of those neighbourhoods. This is essential for plot the map and get the venue data.
3. Restaurant Venue data. This data will be used to perform further analysis of the neighbourhoods.

The data of NYC's boroughs and neighbourhoods are collected from the open source dataat: https://cocl.us/new_york_dataset. Later, the geographical coordinates of the neighbourhoods will be extracted using the python geocoder package.

We utilize foursquare API to get the data about the venues.

Methodology

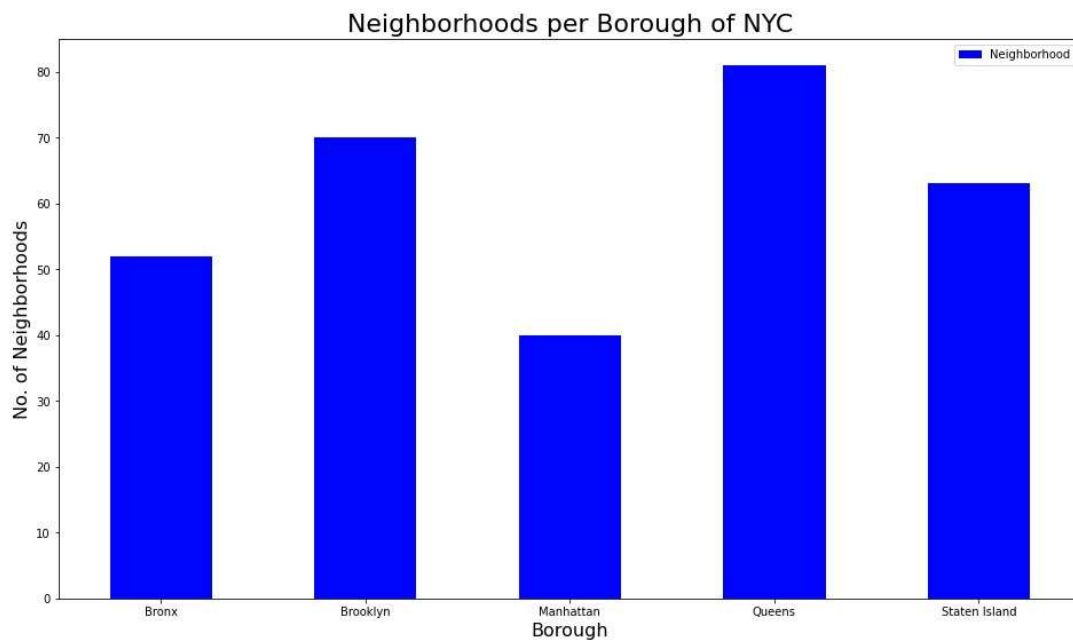
- First, data about the boroughs and neighbourhoods of Ney York was collected from https://cocl.us/new_york_dataset. This data was cleaned and processed into a dataframe.
- Then Foursquare was used to locate and filter data about all Korean Restaurants in NYC. Later, data about Ratings, Tips and Likes by users were extracted from squarespace and added to the dataframe.

- Then the dataframe was sorted based on different criteria.
- Numpy and Pandas packages of python was used for data analysis. For data visualization, matplotlib and folium packages were used.

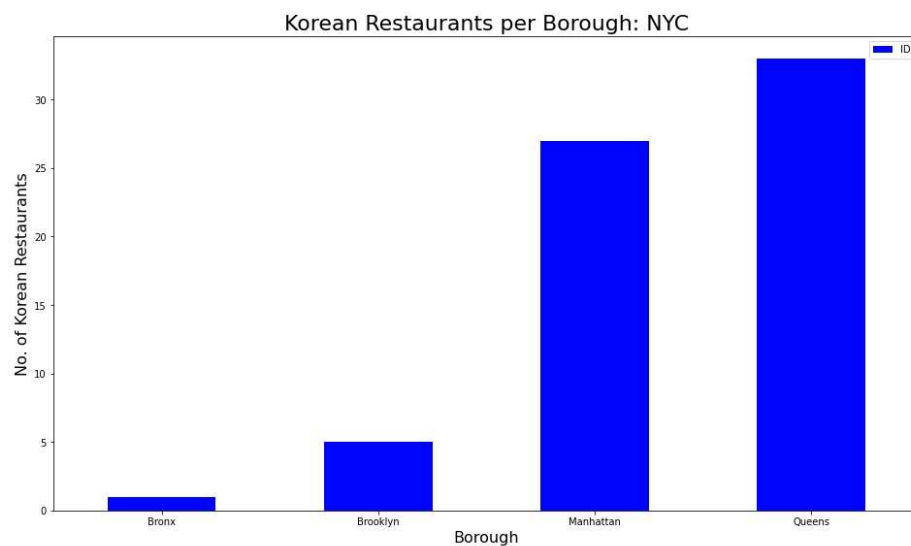
Results

The results of our analysis shows:

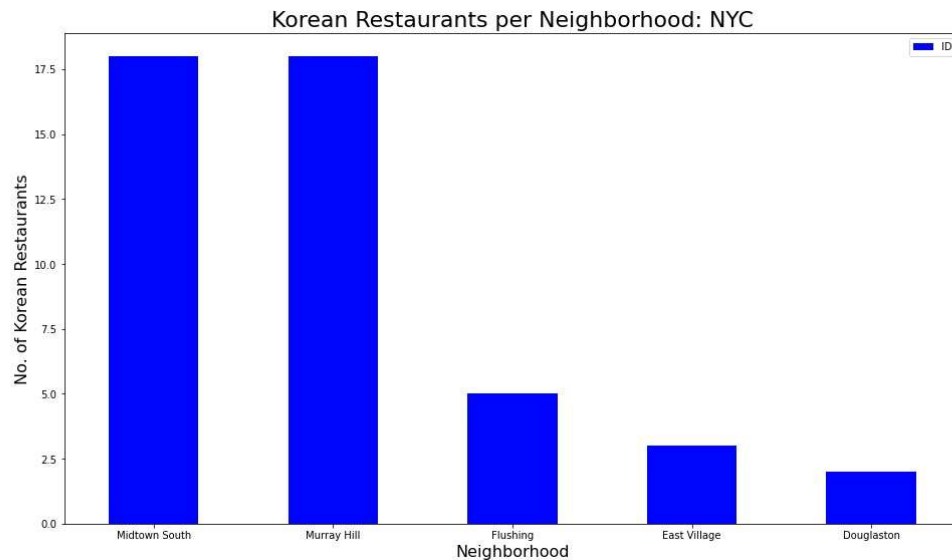
1. Queens has the highest number of neighbourhood



2. Queens also has the highest number of Korean restaurants in NYC.



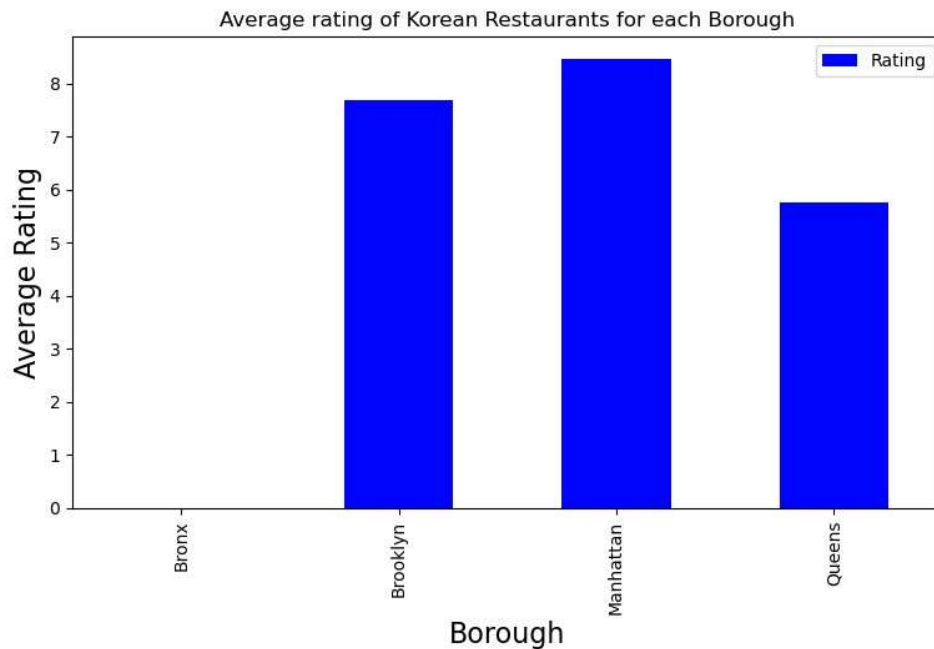
3. The Murray Hill in Queens and Midtown South neighbourhoods of Manhattan has the highest number of Korean restaurants.



4. Manhattan has the highest average ratings for Korean restaurants, followed by Brooklyn. Top 3 three neighbourhoods with highest average ratings all are from Manhattan. Tribeca takes the top spot for highest average ratings.

| | Neighborhood | Average Rating |
|----|-------------------|----------------|
| 19 | Tribeca | 8.900000 |
| 5 | East Village | 8.800000 |
| 3 | Clinton | 8.500000 |
| 12 | Midtown South | 8.455556 |
| 6 | Flatiron | 8.400000 |
| 16 | Prospect Heights | 8.400000 |
| 20 | West Village | 8.200000 |
| 14 | North Side | 8.200000 |
| 11 | Manhattan Valley | 8.100000 |
| 18 | Sunnyside Gardens | 7.800000 |

Top Ten neighbourhoods ranked by average ratings of Korean Restaurants



Discussion

According to our analysis, Queens have the most number of Korean restaurants. Manhattan comes in second place followed distantly by Brooklyn. Staten Island doesn't have any Korean restaurant according to Foursquare data while Bronx has only one. Although Queens have the most number of Korean restaurants, when it comes to rating it's Manhattan that comes out on the top, followed by Brooklyn. Among the 9 neighbourhoods with average rating of 8.0 or more, 7 is located in Manhattan. The other two is in Brooklyn. Because of Manhattan's highly rated restaurants it might be tempting to open a new Korean restaurant in that borough. But we have to keep in mind that it also has a lot more Korean restaurants than Brooklyn. On the other hand, restaurants in Queens don't have good average ratings. So in my opinion Brooklyn is the best place to open a new Korean restaurant. While restaurants here are not as highly rated as in Manhattan, it has less competition and it has a lot more growth potential.

But I would like to add that, the above analysis is highly contingent on the accuracy of Foursquare data. To make a more informed decision, more data from other verified sources need to be incorporated into the analysis.

One limitation that was encountered during the project was the limitation on the number of API calls to Squarespace. Since, a free Sandbox tier was used to make API calls it had a limit on how many calls can be made per day. Another limitation was the squarespace itself. The whole project was completely depended on data provided by Squarespace. For further analysis, it is recommended to explore other data sources. Also a paid tier of squarespace should be used.

Conclusions

The project was executed in multiple phases. In the first phase the business problem, its audience and data sources were identified. Then data was collected, cleaned and analysed to solve the business problem. During analysis various data science methods and instruments were applied. The results were also visualized through charts and maps. The question that was asked in the business problem was, which is the best location in the New York city to open a new Korean restaurant? At the end of our project we found the answer and that is: Brroklyn.