



CAPSTONE PROJECT

IBM DATA SCIENCE

By ABAY MIRMANOV

INTRODUCTION

New York is the most populous city in the US. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km²), New York City is also the most densely populated major city in the United States. Located at the southern tip of the U.S. state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass. With almost 20 million people in its metropolitan statistical area and approximately 23 million in its combined statistical area, it is one of the world's most populous megacities.

This Capstone project explores the best locations for Japanese restaurants throughout the city of New York. As New York is the most diverse city in the world (800 languages are spoken in New York), it has a long tradition of different ethnical restaurants. Now when the idea of a healthy lifestyle conquered the minds of people all over the country, Japanese restaurants became extremely popular, as they offer a healthy alternative to regular American eating habits. That's why Potentially the owner of the new East European restaurant can have great success and consistent profit. However, as with any business, opening a new restaurant requires serious considerations and is more complicated than it seems from the first glance. In particular, the location of the restaurant is one of the most important factors that will affect whether it will have success or a failure. So our project will attempt to answer the questions "What is the best locations to open a Japanese restaurant?"

DATA

Data required

- List of borough and neighborhoods of New York
- Latitude and longitude coordinates of those neighborhoods
- Venue data, particularly data related to restaurants

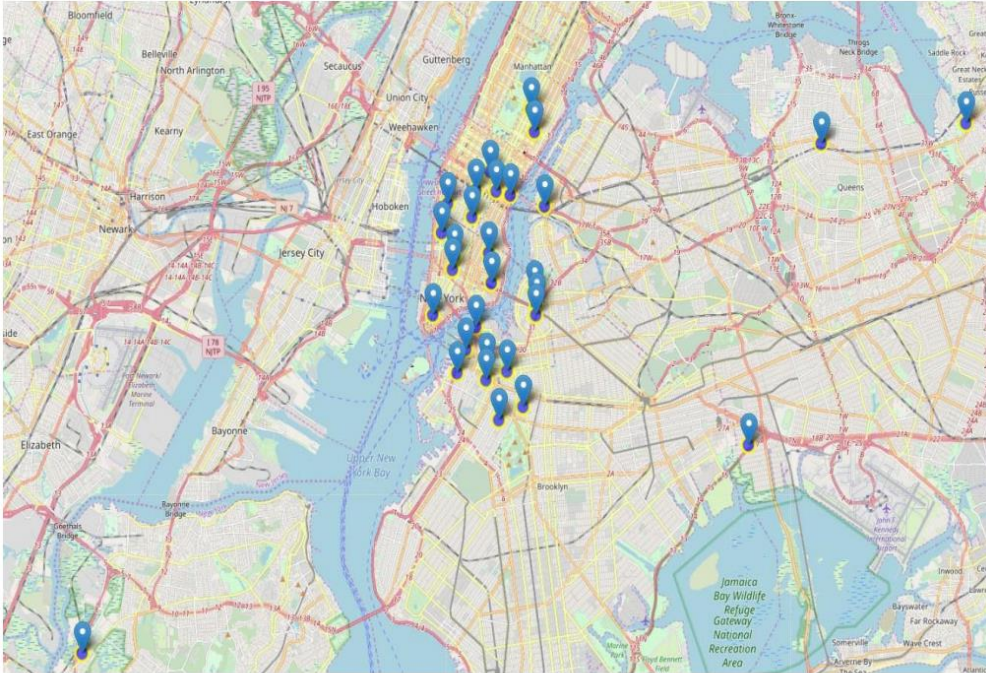
Sources of data

- Open source dataset of New York:
https://cocl.us/new_york_dataset
- Geocoder package for latitude and longitude coordinates
- Foursquare API for venue data

METHODOLOGY

- Collect the data from https://cocl.us/new_york_dataset
Clean and process the data into a data frame.
- Using FourSquare, locate all venues and then filter by Japanese restaurants. Ratings, tips, and likes by users will be counted and added to the data frame.
- Sort the data based on rankings.
- Visualize the data using graphing from Python libraries.

RESULTS



- **Manhattan** has the highest average rating for Japanese Restaurants.
- **Manhattan** has the highest number of Japanese restaurants comparing to the other boroughs.
- Two neighborhoods with the highest average rating for Japanese Restaurants are **Cobble Hill** and **North Side** (both located in **Brooklyn**).

DISCUSSION



Based on the results of our analysis, **Manhattan** and **Brooklyn** are the best locations for Japanese cuisine in NYC.

Brooklyn is a more promising place to open a new Japanese restaurant.

Neighborhoods **Cobble Hill** and **North Side** are the best place in Brooklyn to open a new Japanese restaurant.

CONCLUSION

- Answer to business question: Where in the city of New York, should the investor open a Japanese Restaurant?
 - Best rated Japanese Restaurant in NYC is OOTOYA Restaurant in Manhattan.
- In a fast moving world, there are many real life problems or scenarios where data can be used to find solutions to those problems. Like seen in the example above, data was used to cluster neighborhoods in New York based on the most common venues in those neighborhoods. Similarly, data can also be used to solve other problems, which most people face in metropolitan cities.



THANKS

Does anyone have any questions?

abaymir@gmail.com
+7 707 6912893