THE BATTLE OF THE NEIGHBORHOODS

An optimal location for a new Italian restaurant

Xueyue Xiao

June 27, 2019

- Mr. Panini is an Italian restaurant owner in California, and recently he would like to expand his business to NYC. He would like to open a new restaurant **in Manhattan**, but he doesn't know where should him to set up his new business.

- TARGET: help Mr. Panini to find an optimal location or block for his business.

BUSINESS PROBLEM

- Since Mr. Panini is new to NYC, I will first create a map of NYC and Manhattan to him. Given the map, he is supposed to have a big picture about what does NYC and Manhattan look like, how the venues distribute.

- Mr. Panini is going to open an Italian restaurant. His main competitors are existed Italian restaurant. Therefore, I will list all the Italian restaurant in Manhattan with their name, coordinates, rating.

- In order to show a clear picture, I will create a map to show how Italian restaurants are distributed in each neighborhood.

- I will cluster the restaurants to narrow down the location and calculate the average rating.

- By choosing the lowest average rating to be our target cluster, I will choose the neighborhood with most Italian restaurants. Comparison will make Mr. Panini's restaurant stand out.

- Finally, the location will be determined according to the Italian restaurants in this neighborhood.

STRATEGY

④ The source of data is from New York data json file given by the course and from Foursquare API.

④ **New York data json file** provides me with the borough, neighborhood, latitude and longitude of each neighborhood in NYC. This data source allows me to create the map of NYC and Manhattan.

④ **Foursquare API** allows me to get the Venue in each neighborhood, its latitude, longitude, category, ID and rating. This data source allows me to obtain the list of Italian restaurants, conduct calculation of the rating, and create the map with Italian restaurants.
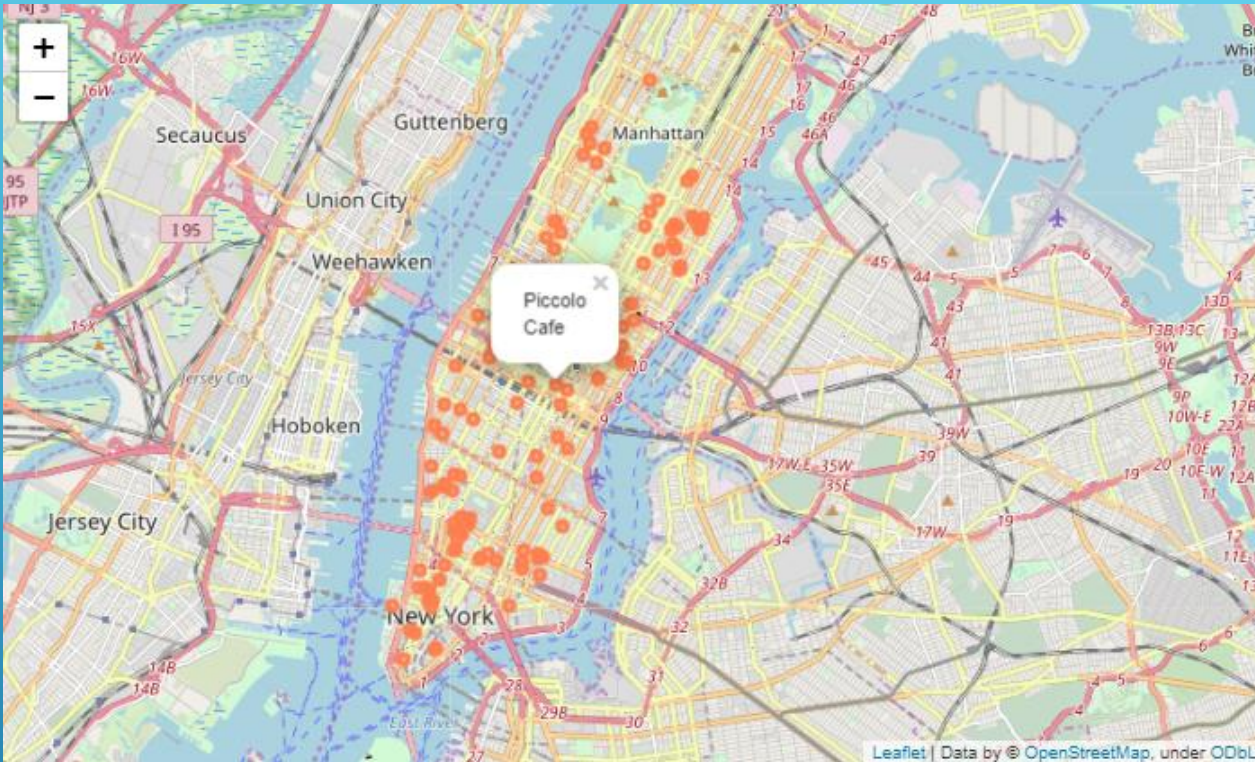
DATA DESCRIPTION
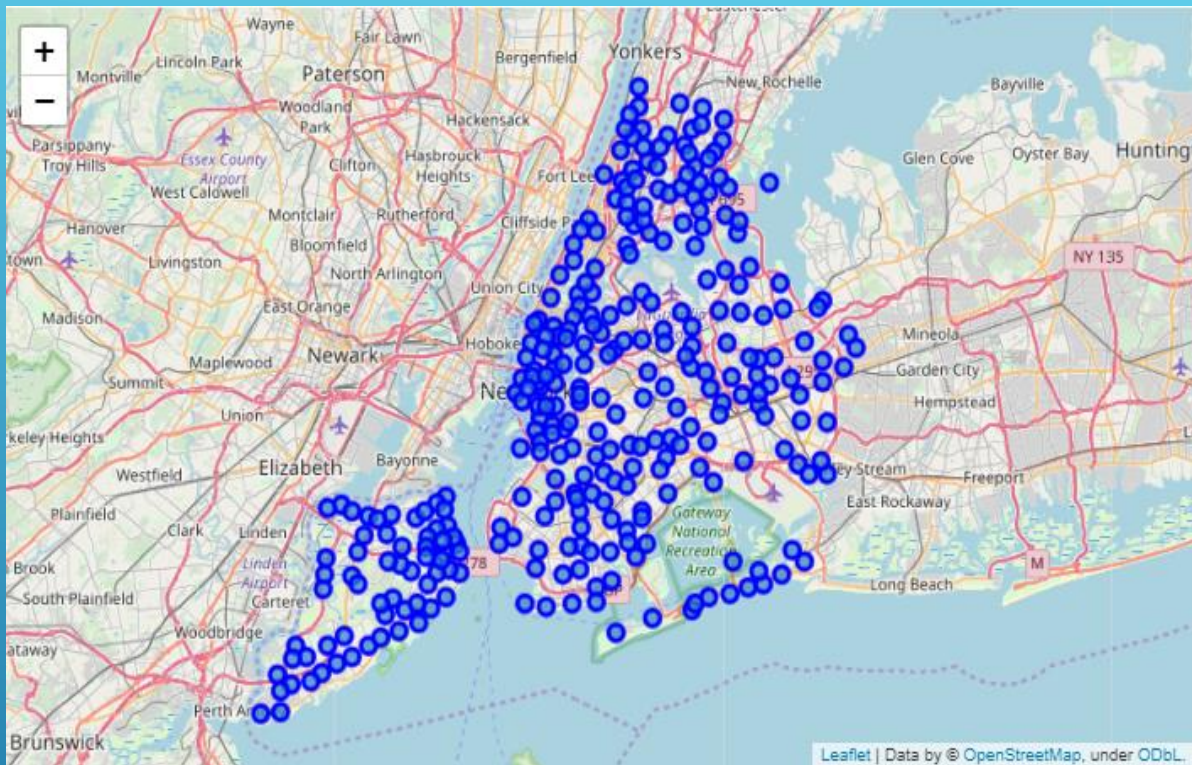
④ There are **5 boroughs** in NYC: Bronx, Brooklyn, Manhattan, Queens, Staten Island.
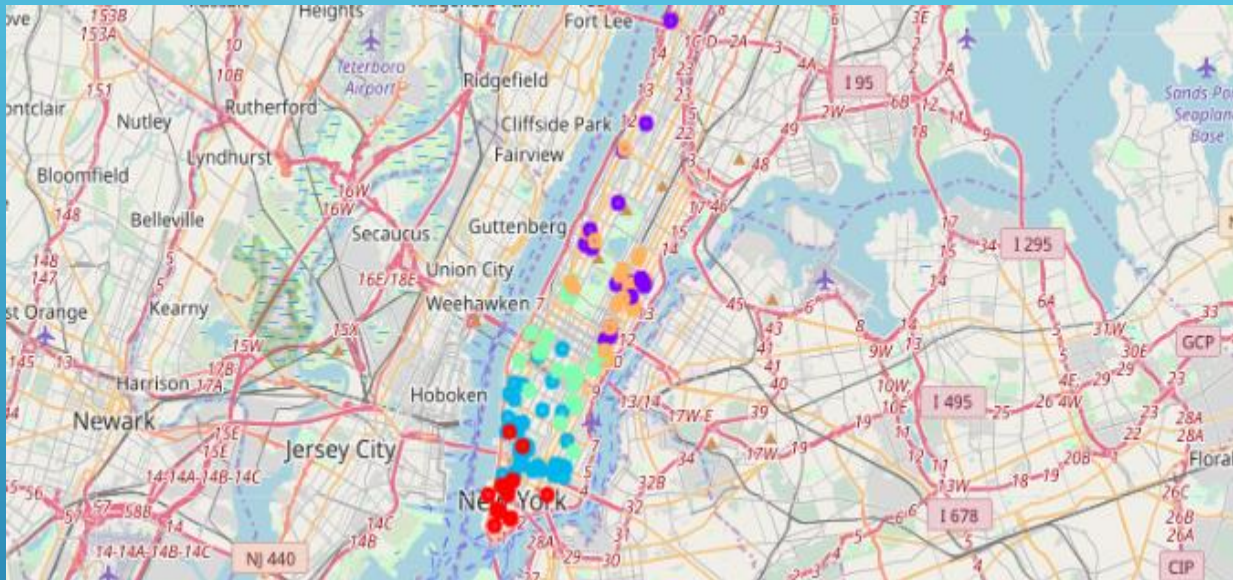
BIG PICTURE OF NYC
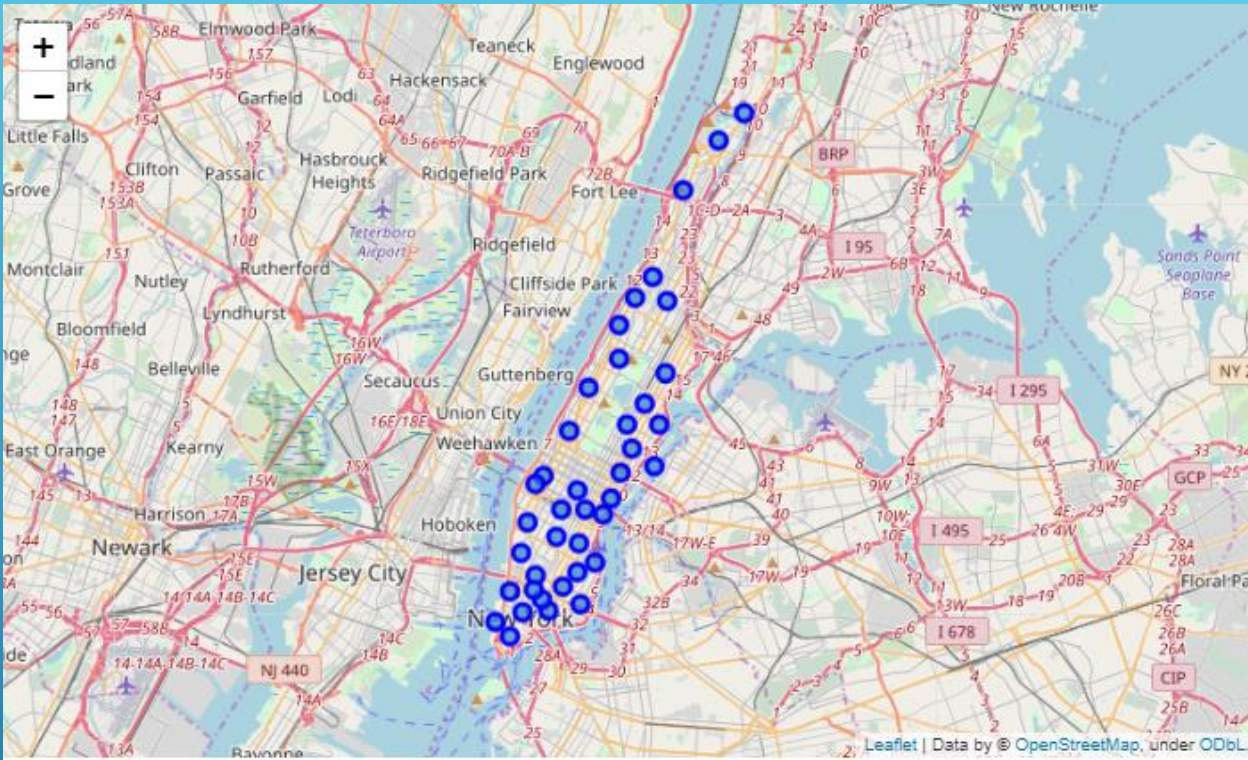
④ There are **40 neighborhoods** in Manhattan.

MANHATTAN

④ There are **125** Italian restaurants in Manhattan with average rating **8.41**

DISTRIBUTION OF ITALIAN RESTAURANTS IN MANHATTAN

④ According to the venue coordinates and rating, I used unsupervised learning l-means algorithm to cluster the venues. k-means is one of the most common clustering algorithm which is very easy to understand. I set k as **5** to get appropriate amount of clusters.

K-MEANS ALGORITHM

Number of Italian Restaurants in the neighborhood
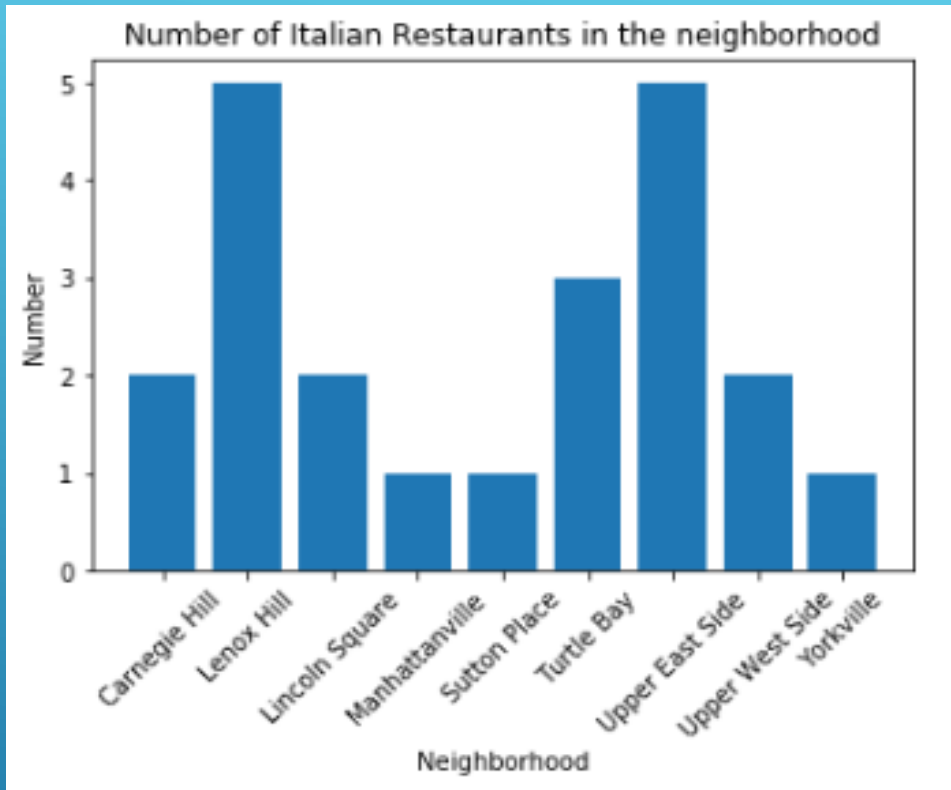
- ④ Divided the restaurants based on the neighborhood which can help me narrow down the area.

- ④ Lenox Hill and Upper East Side have 5 Italian restaurants.

DETERMINE THE LOCATION

④ Chose the neighborhood with lower average rating, which turns out to be Upper East Side.

④ The center of five restaurants is (40.77, -73.95), which is 1187 Lexington Avenue.

RESULT

④ Manhattan is the most popular borough in NYC, with the highest density and a variety of restaurants.

④ Separate Manhattan into five clusters using k-mean method to figure out the average rating in each cluster. The restaurants at this Neighborhood have average lowest rating. When people are surrounded by low rating restaurants, people would tend to choose the new one. As long as the new restaurant can offer tasty dishes, it is highly likely that the new one can receive higher rating.

④ When I look inside the cluster, the method I used is to compare the number of restaurants in each neighborhood. Upper East Side has the lowest rating. When people conduct comparing, Mr. Panini's restaurant can stand out.

DISCUSSION

④ The purpose of this project is to help Mr. Panini find an optimal location to move his business from California to New York.

④ Thanks for Cousera capstone project provides me with NYC data file so I can easily get the coordinates of neighborhood and borough in NYC. Through data from Foursquare API, I am able to cluster the Italian restaurants into five cluster and calculate the average rating.

④ Clustering helps me to narrow down the area. This project determine the optimal location only based on the competitors' category, locations and ratings.

④ For Mr. Panini, he is supposed to consider more factors include price, other type of restaurants, food quality, menu, population density, real estate availability and etc.

CONCLUSION