

Capstone Project – The Battle of Neighborhoods

Report

Introduction/Business Problem

New York City is experiencing a Chinese food renaissance. There are a lot of spots in the city with the Chinese food available.

In this project we will try to find an optimal location for a restaurant. Specifically, this report will be targeted to stakeholders interested in opening a Chinese restaurant in Manhattan, New York City, USA.

Since there are lots of restaurants in New York City, and Manhattan specifically, we will try to detect locations that are not already crowded with restaurants. We are also interested in areas with no Chinese restaurants in nearby.

We will use our data science powers to compare different neighborhoods in Manhattan. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

The main purpose of this project is answer following questions:

1. Visualize which parts of the city have Chinese Restaurants
2. Identify the best area for Chinese Cuisine in Manhattan
3. Identify the areas which lack the Chinese Restaurants

Answering this three questions will help to answer the main problem question:

1. What is best place to open new Chinese Restaurant in Manhattan?

Data Section

To find the solution to the main problem the following data will be used:

1. Source: https://cocl.us/new_york_dataset

This New York City Neighborhood Names point file was created as a guide to New York City's neighborhoods that appear on the web resource, "New York: A City of Neighborhoods."

2. Source: Foursquare API

The data source will be used to find location of the Chinese restaurants in each neighborhood of Manhattan.

Above data sources will be needed to extract/generate the required information:

- location for the NYC center
- center addresses of candidate neighborhoods
- number of restaurants and their type and location in every neighborhood will be obtained using Foursquare API

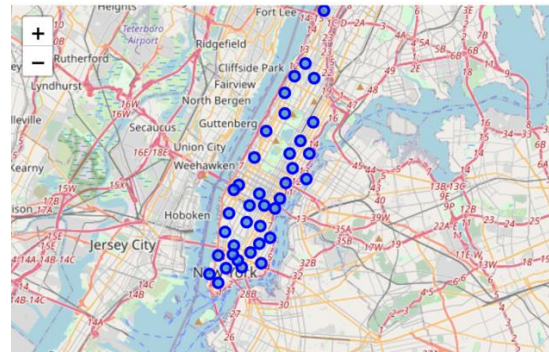
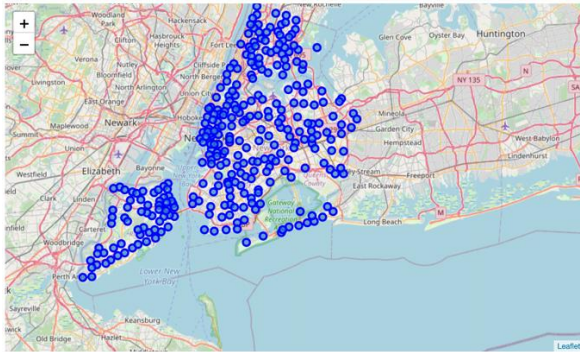
Methodology

We start by collecting data from a https://cocl.us/new_york_dataset website

Obtained dataframe has the following format

	Borough	Neighborhood	Latitude	Longitude
301	Manhattan	Hudson Yards	40.756658	-74.000111
302	Queens	Hammels	40.587338	-73.805530
303	Queens	Bayswater	40.611322	-73.765968
304	Queens	Queensbridge	40.756091	-73.945631
305	Staten Island	Fox Hills	40.617311	-74.081740

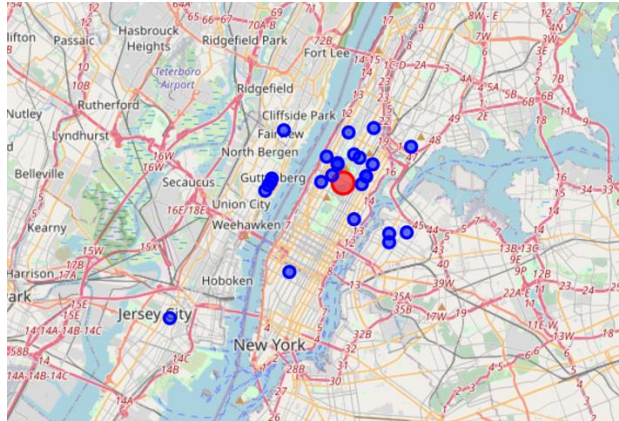
Mao shown below illustrates all of the New York City and Manhattan neighborhoods respectively



Next, we extract the data about Chinese restaurants in Manhattan form the Foursquare API. Finally, after data wrangling procedure the following dataframe containing name, latitude, longitude, and distance from the Manhattan center was obtained.

	name	categories	lat	lng	distance
0	Best Time Chinese & Tex-Mex Food	Chinese Restaurant	40.788643	-73.948524	964
1	U Like Chinese	Chinese Restaurant	40.798554	-73.963403	1037
2	Hong Kong Chinese	Chinese Restaurant	40.792676	-73.945675	1245
3	Fu Wing Chinese	Chinese Restaurant	40.792273	-73.945705	1231
4	No. 1 Chinese Restaurant	Chinese Restaurant	40.801317	-73.949707	1559
6	Acupuncture & Chinese Healing by R. Brown	Acupuncturist	40.801959	-73.970011	1616
7	Empire III Chinese & Pan Asian Cuisine	Chinese Restaurant	40.802902	-73.953270	1580
8	No. 1 Chinese Kitchen	Chinese Restaurant	40.789793	-74.005196	3818
9	Wing Sing Chinese Restaurant	Chinese Restaurant	40.798323	-73.941637	1817
10	Harvest Chinese Thai Cuisine	Asian Restaurant	40.771962	-73.953103	2047
11	New Aroma Chinese	Chinese Restaurant	40.813038	-73.956325	2623
12	U Like Chinese Restaurant	Chinese Restaurant	40.793198	-73.966782	703
13	Eastern Chinese Restaurant	Chinese Restaurant	40.765613	-73.930882	3622
14	Chinese East	Chinese Restaurant	40.747429	-73.993916	5503
19	Foliage Garden Chinese Restaurant	Chinese Restaurant	40.785829	-74.008721	4136
20	Hong Kong Chinese Restaurant	Chinese Restaurant	40.815205	-73.940559	3280
21	Hwa Ying Chinese Restaurant	Chinese Restaurant	40.798274	-73.963267	1004
22	No. 1 Peking Chinese Restaurant	Chinese Restaurant	40.725567	-74.068329	11595
23	Jade Garden Chinese Tex Mexican Restaurant	Chinese Restaurant	40.791532	-74.004271	3746

The following map illustrated the location of Chinese restaurants with blue dots



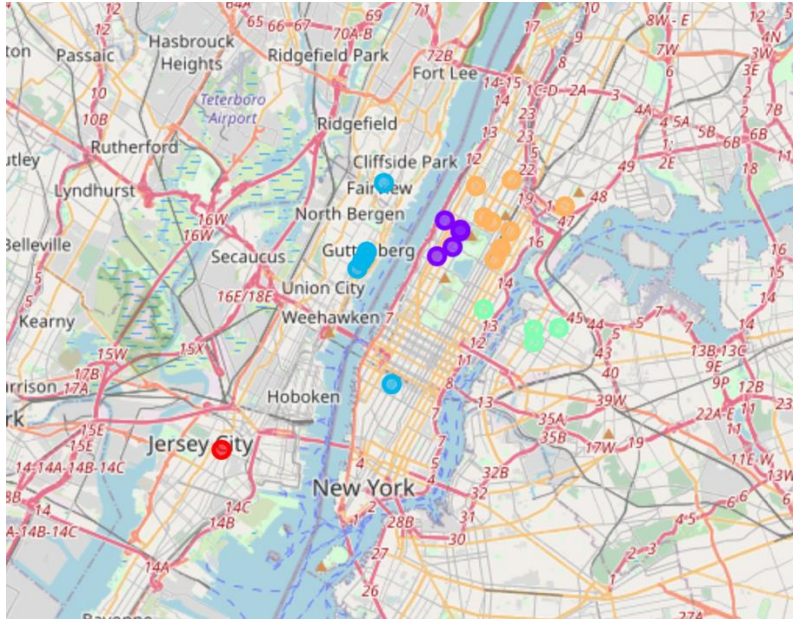
Finally, the kMeans clustering machine learning algorithm was implemented in order to cluster restaurants based on their location in the city

Final dataframe with the Cluster labels is shown below

Cluster Labels		name	categories	lat	lng	distance
4	Best Time Chinese & Tex-Mex Food	Chinese Restaurant	40.788643	-73.948524	964	
1		U Like Chinese	Chinese Restaurant	40.798554	-73.963403	1037
4	Hong Kong Chinese	Chinese Restaurant	40.792676	-73.945675	1245	
4	Fu Wing Chinese	Chinese Restaurant	40.792273	-73.945705	1231	
4	No. 1 Chinese Restaurant	Chinese Restaurant	40.801317	-73.949707	1559	
1	Acupuncture & Chinese Healing by R. Brown	Acupuncturist	40.801959	-73.970011	1616	
4	Empire III Chinese & Pan Asian Cusine	Chinese Restaurant	40.802902	-73.953270	1580	
2	No. 1 Chinese Kitchen	Chinese Restaurant	40.789793	-74.005196	3818	
4	Wing Sing Chinese Restaurant	Chinese Restaurant	40.798323	-73.941637	1817	
3	Harvest Chinese Thai Cuisine	Asian Restaurant	40.771962	-73.953103	2047	
4	New Aroma Chinese	Chinese Restaurant	40.813038	-73.956325	2623	
1	U Like Chinese Restaurant	Chinese Restaurant	40.793198	-73.966782	703	
3	Eastern Chinese Restaurant	Chinese Restaurant	40.765613	-73.930882	3622	
2	Chinese East	Chinese Restaurant	40.747429	-73.993916	5503	
2	Foliage Garden Chinese Restaurant	Chinese Restaurant	40.785829	-74.008721	4136	
4	Hong Kong Chinese Restaurant	Chinese Restaurant	40.815205	-73.940559	3280	
1	Hwa Ying Chinese Restaurant	Chinese Restaurant	40.798274	-73.963267	1004	
0	No. 1 Peking Chinese Restaurant	Chinese Restaurant	40.725567	-74.068329	11595	
2	Jade Garden Chinese Tex Mexican Restaurant	Chinese Restaurant	40.791532	-74.004271	3746	
4	Good Choice Chinese Restaurant	Chinese Restaurant	40.806516	-73.947477	4028	

Results

The following map clearly demonstrates the results of the algorithm



Based on this result, we can state that most of Chinese restaurants are located close to the center of Manhattan and north part of it.

Discussion

Based on the result of the project the following recommendations for the stakeholders can be made:

- 1) The best location for new Chinese restaurant is considered to be the Southern part of Manhattan

The Project also took Jersey City into considerations. So, the center of Jersey City was found to be the best place for the potential Chinese restaurant.

Conclusion

This Capstone Project was aimed to help stakeholders to identify the best location for the new Chinese restaurant in Manhattan, New York City, USA.

The Kmeans algorithm was implemented in order to solve this problem. Finally, the Southern part of Manhattan seems to be the most reasonable location.

Some improvements can be made in algorithms, and probably solving this problem by different machine learning techniques may result in better solution.