

# Identifying Neighborhoods within Manhattan Borough of New York city to locate a new Mexican restaurant

September 20, 2020

## Introduction

With over 20 million residents, New York City is one of the most populous cities in the world. Out of the 5 major boroughs, Manhattan is the most densely populated borough and is also the financial, cultural and administrative center of New York.

## Business Problem

In this project we have been approached by a Mexican restaurant group to help them identify high potential neighborhoods within Manhattan where they can locate their next restaurant.

Based on an initial discussion with the group, the initial criteria for identifying target neighborhoods is as follows:

1. Neighborhoods need to have high density of restaurants
2. Neighborhoods need to have a low presence of Mexican restaurants

The primary reasoning for the criteria was that it is better to be in a place where there are customers already present rather than try to attract them to a different place. Further having found a busy neighborhood, the next item is to make sure that there is little competition in the Mexican cuisine arena.

## Data Acquisition & Preparation

Based on our problem definition, we obtained the neighborhood & borough geo data from the New York City website. For restaurant specific information we used the FourSquare API to obtain the information.

The following data sources were used to gather the data:

- NYC neighborhood geo data - [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)
- Restaurant information from Foursquare API
- Geo data to visualize maps - <https://data.cityofnewyork.us/City-Government/Borough-Boundaries/tqmj-j8zm>

The advantage of using the data from the City and FourSquare was that data was cleaned and compiled thoroughly. The main focus was on extracting the relevant information and analyzing it.

## Methodology

We followed a three step process to find a solution to this problem:

Step 1 - We identified the top 25 Neighborhoods based on the total number of restaurants. Once we had a list of the neighborhoods, we counted the number of Mexican Restaurants in each of these neighborhoods.

Step 2 - To ensure that we identified areas with the most potential, we clustered these neighborhoods using the K-means algorithm.

Step 3 - Finally to get a better sense of the geography, we mapped these clusters and were able to identify the boundaries within which to focus our search for a site.

## Analysis

Step 1(a): Identified the top 25 Neighborhoods based on the total number of restaurants. Below is a representative list of the Top 10 Neighborhoods with the Total number of restaurants.

	Neighborhood	Total Restaurants
0	East Village	76
1	Flatiron	74
2	Greenwich Village	73
3	Chinatown	72
4	West Village	72
5	Noho	69
6	Soho	65
7	Little Italy	65
8	Midtown South	61
9	Midtown	59
10	Clinton	57

Step 1 (b): Counted the number of Mexican restaurants in each of these selected neighborhoods. Below is a list of neighborhoods with a breakdown of Total Restaurants and Mexican restaurants.

	Neighborhood	Total Restaurants	Mexican Restaurants
0	East Village	76	7
1	Flatiron	74	3
2	Greenwich Village	73	1
3	Chinatown	72	4
4	West Village	72	4
5	Noho	69	5
6	Soho	65	2
7	Little Italy	65	1
8	Midtown South	61	0
9	Midtown	59	2
10	Clinton	57	4

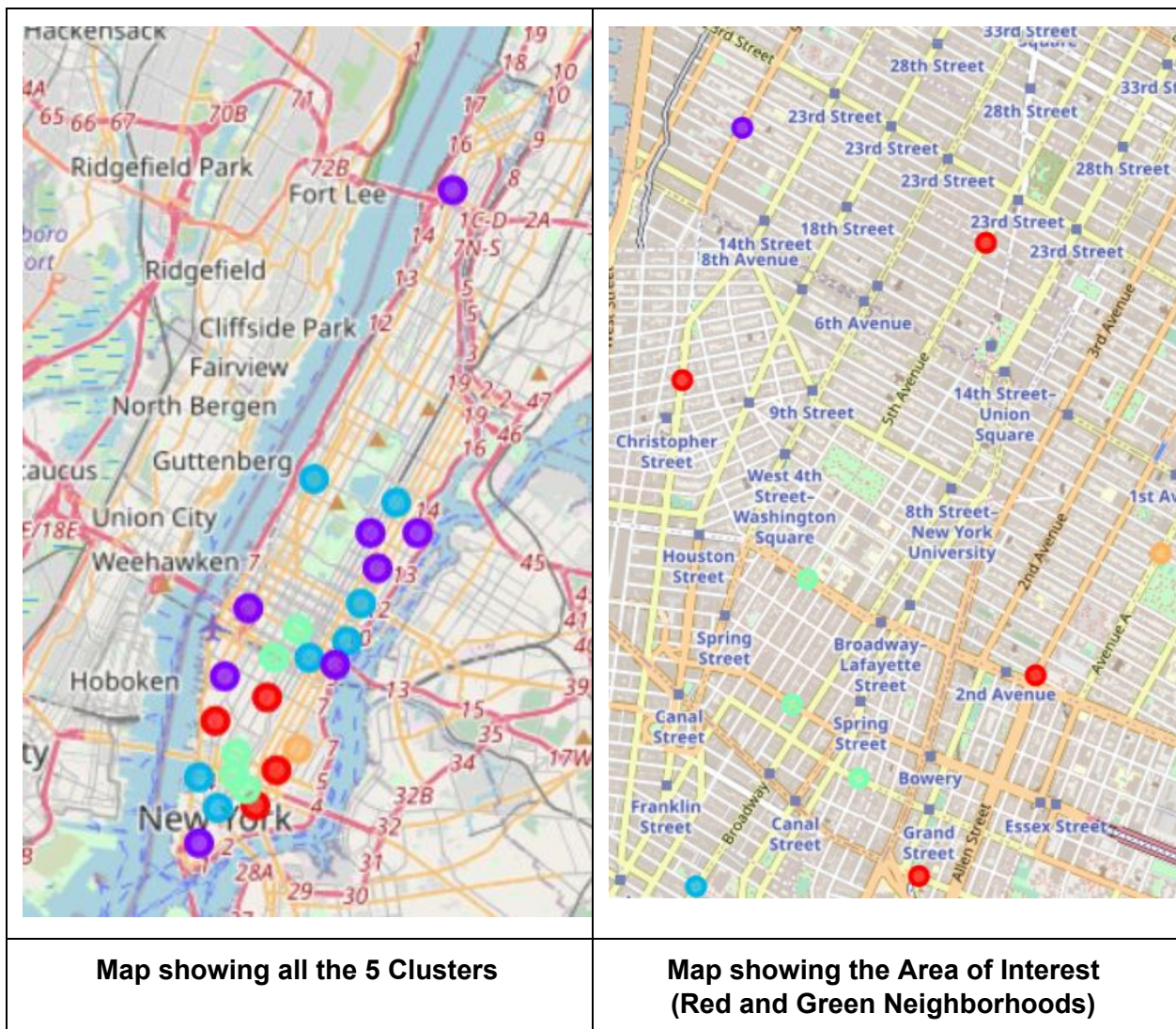
Step 2: Since our goal was to identify neighborhoods with higher density of restaurants in general but with fewer Mexican options available, we used the K-means clustering, a popular unsupervised Machine Learning algorithm to group these neighborhoods into 5 groups.

	Cluster Labels	Neighborhood
0	4	East Village
1	0	Flatiron
2	3	Greenwich Village
3	0	Chinatown
4	0	West Village
5	0	Noho
6	3	Soho
7	3	Little Italy
8	3	Midtown South
9	3	Midtown
10	1	Clinton
11	1	Lenox Hill
12	1	Chelsea
13	2	Murray Hill
14	2	Civic Center
15	1	Yorkville

After a close examination of these clusters, it was determined that Cluster 3 and Cluster 0 were the Clusters with the target neighborhoods since both these clusters grouped restaurants with relatively higher number of restaurants and fewer Mexican restaurants.

Step 3: The next step was to visually see where these clusters were located on a map to give a better idea of what the boundaries would look like.

We used Folium, a popular mapping tool, to map the results.



## **Results & Discussion**

Our analysis shows that there are 40 neighborhoods in the Manhattan Borough in New York City. Each neighborhood has its own unique character. For our purpose of identifying the neighborhoods with the most promise for locating a new Mexican restaurant, we chose to focus on the top 25 neighborhoods with the most number of restaurants. The rationale behind picking these neighborhoods was that it is easier to cater to an existing client base, in this case customers who visiting the area to eat, then to locate in neighborhoods with fewer restaurants.

After we picked these neighborhoods, we wanted to prioritise the neighborhoods with less Mexican options.

## **Conclusion**

Based on our analysis of the 40 neighborhoods in the Manhattan, NY area, we would recommend the 9 neighborhoods located between 23rd Street and Canal Street for further exploration for locating a new Mexican restaurant. Additional considerations like demographics, site visibility and access, rent and traffic count will play a critical role in ultimately finalizing the site for the new restaurant location.