Mexican Restaurants New York

Rikard Johansson

January 2021

1 Introduction

1.1 Background

Millions of people live in New York making it one of the most lucrative markets in the world. New York is very ethnically diverse which may lead to a large variety in food preferences. The Mexican food cuisine is highly appreciated giving room for restaurants such as Chipotle to establish an enormous customer base. The project will look in to what is the market opportunity for opening a Mexican restaurant in the New York area and what area in New York would be suitable.

1.2 Problem

One of the most difficulties with this will be to figure out if the concentration of good restaurants in a specific area is due to denser population or a specific food preference for the people in that specific area.

2 Data acquisition and cleaning

2.1 Data sources

The different Neighborhoods for New York can be found here : New York Data set. The different neighborhoods and boroughs will be stored in a data set. To get the different Mexican restaurants in the different neighborhoods calls will be made to Foursquare.

2.2 Data cleaning

What is needed from the data for this project is :

- Borough
- Neighborhood
- Latitude
- Longitude

This data is needed for calls to the Foursquare API. All the data from Foursquare gives us the data for the Mexican restaurants in the different New York neighborhoods. There is a total of 306 Neighborhoods in the New York area. Figure 1 shows the first four from the data set.

[4]:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Figure 1: New York Neighborhoods

3 Data analysis

3.1 Visualization of the data

There was a total of 50 restaurants in the data set. Using the different matlib plots can give us some more understanding of the data.

	Likes	Rating	Tips
count	50.000000	50.000000	50.000000
mean	45.940000	7.286000	18.920000
std	57.595851	1.701021	21.573642
min	0.000000	0.000000	0.000000
25%	9.250000	6.900000	3.000000
50%	24.000000	7.600000	8.000000
75%	54.000000	8.100000	35.250000
max	250.000000	9.100000	88.000000

Figure 2: Data for restaurants

The mean rating of the restaurants is about 7.3 which is quite high. We can see that majority of the rating ranges from 8.1 to 6.9. There are a couple of zeros in the data set that will be filtered out later. The restaurants can also be shown distributed across the different neighborhoods. In the box-plot it is shown that

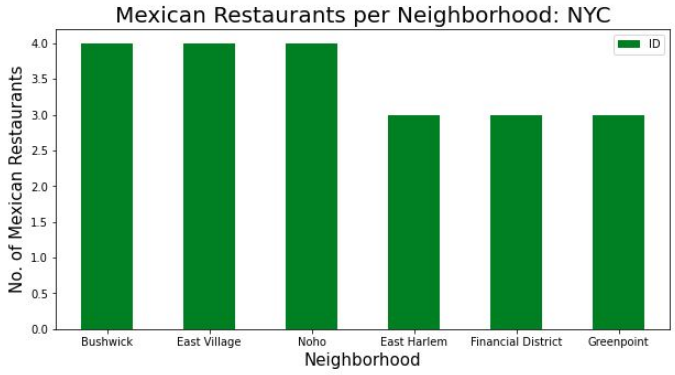

Figure 3: Mexican restaurants per neighborhood

Bushwick, EastVillage and Noho has most restaurants.

The restaurants plotted on a New York map. There is a high concentration in the lower part of Manhattan, western Brooklyn and spread around Bronx. There are not a lot of restaurants between these two clusters, there could be that we are missing those in our data or the fact that they do not exist. This would need further investigation a reason could be that rates might be higher in some areas making it hard to make a profit running a restaurant business in specific areas.

Figure 4: Restaurants with 3 or higher rating

References

https://github.com/rikardbelowaverage/Capstone/blob/master/new_york_data.ipynb
https://cocl.us/new_york_dataset
https://foursquare.com/developers/apps