

Patterns in Types Of Restaurants in New York

VARSHA GADEKAR

NOV 11, 2020

1. INTRODUCTION

1.1 Background

New York City (NYC), often called simply New York, is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km²), it is one of the world's most populous megacities. Situated on one of the world's largest natural harbors, New York City is composed of five boroughs, each of which is a county of the State of New York.

The restaurant industry is essential to New York City's social and economic fabric. From small, family-owned restaurants and food carts to four-star world-famous establishments, the City abounds with restaurants that offer cuisines from every corner of the globe. Restaurants lend vitality to each neighborhood and are integral to Manhattan's central business districts and the City's vibrant tourism industry, attracting millions of visitors each year who spend a total of \$46 billion annually.

1.2 Problem

The project aims to find out if there is any pattern in the type of restaurants (including cafés) located in every borough and neighborhood of New York.

- Are some cuisines / types of restaurants more represented in certain areas than in others?
- Is it possible to group districts or neighborhoods together based on the sole criteria of food?

The project will also focus on two major boroughs in New York for detailed Mexican Restaurant venues based on user ratings and likes.

1.3 Interest

Since eating out and sipping on a cup of coffee is a major aspect of people's life, what if you are new to New York or planning a trip to New York. Wouldn't it be great to know beforehand the boroughs and neighborhoods of New York based on what we can eat there or which area to visit as per our taste of eating.

Why Mexican? Mexican food is as American as apple pie. In fact, there are roughly 60,000 Mexican restaurants in the U.S., representing approximately 9% of all the restaurants in the country. Of equal importance, Mexican food is delicious, nutritious, and relatively inexpensive, especially when you compare dining in a Mexican restaurant with dining at steakhouses or seafood restaurants.

"Americans love Mexican food, and this popular cuisine type appeals to palates all across the nation", said Brad T. Bloom, Director of Sales at CHD Expert North America.

2. Data

2.1 Data Sources

Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. This dataset exists for free on the web. Here is the link to the dataset: https://geo.nyu.edu/catalog/nyu_2451_34572. Geopy library was used to get the latitude and longitude values of New York City. Next, Foursquare API was utilized to explore all the neighborhoods and segment them. The API was also used to get details of venue

2.2 Data Cleaning

Data downloaded was cleaned and saved in a dataframe. The file contained json data so it was converted into a dataframe. Only required features from data were saved in the data frame. Then I checked if the data has all 5 boroughs and 306 neighborhoods. The data was further divided into two different data frames consisting data of boroughs for further analysis.

2.3 Feature Selection

Each record in json file contain following features:

```
{'type': 'Feature',  
  'id': 'nyu_2451_34572.1',  
  'geometry': {'type': 'Point',  
    'coordinates': [-73.84720052054902, 40.89470517661]},  
  'geometry_name': 'geom',  
  'properties': {'name': 'Wakefield',  
    'stacked': 1,  
    'annoline1': 'Wakefield',  
    'annoline2': None,  
    'annoline3': None,  
    'annoangle': 0.0,  
    'borough': 'Bronx',  
    'bbox': [-73.84720052054902,  
      40.89470517661,  
      -73.84720052054902,  
      40.89470517661]}}
```

For the project only the following features were selected :

- *borough* -> Borough
- *Name* -> Neighborhood
- *Coordinates* -> Latitude, Longitude

Shape of data after cleaning -> (306, 4)

