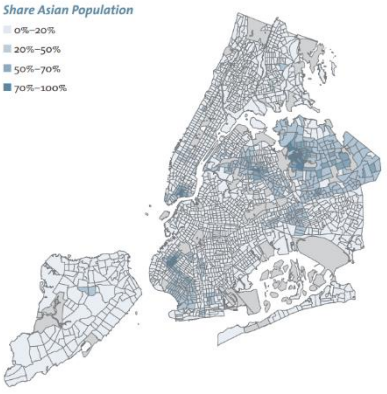
IBM Data Science
Applied Data Science Capstone

Bengali Restaurant in New York City

Introduction

New York City has an extraordinarily diverse population. It is the largest city in the United States with a long history of international immigration. It is one of the few cities in the country in which four different racial/ethnic groups each make up at least 10 percent of the population. According to the 2010 decennial census, 33 percent of New York City residents are white, 26 percent are Hispanic, 26 percent are black, and 13 percent are Asian. While the diversity of New York City's population is not reflected in many of the city's neighborhoods, over the past 20 years the racial and ethnic makeup of the city's neighborhoods (census tracts) has increasingly come to look more like that of the city itself. Between 2000 and 2010 the city as a whole became more Hispanic and Asian. While the population at large has declined the past few years, the city's Asian population has, meanwhile, become more concentrated.[1]



According to the 2000 US Census, there are approximately 57,000 Bangladeshis in the United States and the New York metropolitan area is home to the largest Bangladeshi population. From 1990 to 2000, New York City's Bangladeshi population increased by 471% (from 4,955 to 28,269). About 85% of Bangladeshi New Yorkers were foreign-born, and 77% of these immigrants came to the US from 1990 to 2000. Some 30% of all Bangladeshi immigrants in the city in 2000 were naturalized US citizens – compared with 45% of all foreign-born city residents.[2]

Business Problem

The location of food service business will impact its success nearly as much as the menu. If the restaurant is in the wrong place, we cannot attract the number of customers we will need in order to stay in business. The same is true if our location is inundated with competitors, or has poor visibility, or is hard to find. There are many things that must be considered as we look for a location in which to open our business.

The objective of the Capstone project is to analyze and select the best locations in the city of New York to open a new Bengali/Bangladeshi restaurant, using data science methodology and instruments such as data analysis and visualization.

[1] Racial and Ethnic Makeup of NYC Neighborhoods, The Furman Center for Real Estate & Urban Policy
[2] AAFNY's "Census Profile: New York City's Bangladeshi American Population"

Target Audience

The project is useful to developers and investors looking to open or invest in a Bangladeshi restaurant in the city of New York. Overall, New York is a great place to open a restaurant with an ethnical cuisine as New York is the most diverse city in the world. With its diverse culture, comes diversity in the food items.

We decided to focus on Bangladeshi cuisine in our project because there are around 100,000 Bangladeshi New Yorkers. About 85% of Bangladeshi New Yorkers were foreign-born, and 77% of these immigrants came to the US from 1990 to 2000. So, the Bangladeshi population is expected to rise considerably with its 3rd and 4th generation in the 2020s.

Data

To solve the problem, we will need the following data:

- New York City data containing the neighborhoods and boroughs.
- Latitude and longitude coordinates of those neighborhoods.
- Venue data, particularly data related to restaurants.

New York City data containing the neighborhoods and boroughs will be obtained from the open data source: **https://cocl.us/new_york_dataset**.

Next, we will get the geographical coordinates of the neighborhoods (latitude and longitude) using the **Python Geocoder** package. Example dataframe from the New York Dataset on the right side.

	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

We will use the **Foursquare API** to get the venue data for the neighborhoods defined at the previous step. Foursquare has one of the largest databases of 105+ million places, 900+ venue categories, 30+ attribute fields, and over 150,000 developers use this application. Among the many Foursquare API provided categories of the venue data, we are particularly interested in the restaurant data (e.g. restaurant names, location, ratings, likes) to solve the business problem defined above. Example data retrieved from Foursquare API below.

	Borough	Neighborhood	ID	Name	Likes	Rating	Tips
34	Queens	Jamaica Hills	0	0	0	0.0	0
35	Manhattan	Sutton Place	4a63bfb4f964a520b3c51fe3	Chola Eclectic Indian Cuisine	154	8.3	56
36	Manhattan	Sutton Place	4fe4fb50c2eee335e4fea69d	Moti Mahal Delux	185	8.4	82
37	Manhattan	Sutton Place	3fd66200f964a52099e91ee3	Dawat	35	7.7	33
38	Queens	Sunnyside Gardens	4c48da9f3013a59356c5f0e1	Saffron Garden	17	7.6	16