Exploring New York City to Open an Indian Restaurant

IBM Data Science Capstone Project

Abhishek Mishra

June 2020

1. INTRODUCTION

For this Capstone Project, I am using the hypothetical scenario for a concept Indian Entrepreneur who wants to open an Indian Restaurant in New York City (NYC). It might present a good opportunity for an Indian American already living in NYC and are well versed with the Places and the Neighborhoods. As Indian cuisine is quite popular with Americans and Indian Americans alike, there are already many restaurants most of which are a Franchise or a family owned business.

The New York City region is home to the largest Indian American population among metropolitan areas by a significant margin and represents the second-largest metropolitan Asian national diaspora both outside of Asia and within the New York City metropolitan area.

Ambience, menu, hygiene and of course taste are all important factors to be kept in mind before getting into the Hospitality Industry but these are all problems that can be tackled internally by the person(s) in charge. The location of a restaurant is also of utmost importance regardless of the history of a business or the taste of the food. If people don't come in to eat then none of the preparations matter. That is the problem I am tackling in this project.

2. PROBLEM STATEMENT

The objective is to find a suitable location(s) to open an Indian Restaurant in New York City, USA. This project makes use of various Data Science and Machine Learning methodologies (k-means Clustering) to provide a Solution to the client. The project aims to provide a Solution to the Question : 'Where should you consider opening an Indian Restaurant in New York City?'

3. DATA

3.1 Data

I have used the following Data for the completion of the project :

- List of Boroughs and Neighborhoods in NYC - This gives the coordinates of all the neighborhoods and is used to call the Foursquare API.
- List of Places and Venues in NYC - This contains data about all the nearby venues like Restaurants, Bars, Gym etc.
- Demographics of American Indians in New York City - Vital to understand the distribution of the target audience in NYC.
- Latitude and Longitude Data of the neighborhood(s) - To plot and visualize our data.

3.2 Data Sources

- New York City Neighborhoods Data from NYU website [1].
- Nearby Venues Data created using Foursquare API [2].
- The Demographics Data is scraped from Wikipedia [3].
- Latitude and Longitude values are obtained using the Geocoder package in python.

BIBLIOGRAPHY

[1] https://geo.nyu.edu/catalog/nyu_2451_34572

[2] https://foursquare.com/developers/apps

[3] https://en.wikipedia.org/wiki/Indians_in_the_New_York_City_metropolitan_region