IBM APPLIED DATA SCIENCE

CAPSTONE PROJECT

OPENING A NEW INDIAN RESTAURANT IN NEW YORK CITY, USA

By: Karpagameenakshi Gurumoorthy

MAY 2020 – JUNE 2020

1. Introduction and Background of the Business Problem:

INTRODUCTION:

New York City (NYC) is the most popular, densely populated major city in the United States of America. It is composed of five boroughs namely Brooklyn, Queens, Manhattan, the Bronx, and Staten Island which were consolidated into a single city in the year 1898. Presently, NYC has become an America's city of immigration as it welcomed people from all over the world. As a result of immigration, NYC has become the hub of diverse culture. Consequently, the construction of several restaurants, shopping malls have increased. Taking account of the people population from India in NYC, this Capstone Project explores the best locations for the construction of Indian Cuisine Restaurants.

BUSINESS PROBLEM:

Commencing up a new restaurant is not a straightforward job. The initial and foremost step involved in opening a new restaurant is deciding the best location, as it will determine whether a restaurant will be prosperous in the business or it would be a failure. Thus, the main object of this Capstone Project is to examine and select the best locations in NYC, to begin a new Indian restaurant with the help of data science and the data obtained from Foursquare API.

TARGET AUDIENCE:

➢ The main objective of this project is to help the business person who wants to invest or open a new Indian Restaurant in New York City.
➢ The other objective is to find the perfect location for setting up the restaurant that helps the investor to have more income.

2. Data Section:

DATA SOURCES:

The data sources used for the Capstone Project are,

➢ In this project to determine the highest Indian Population in the United States of American, I utilized the dataset from the site WorldAtlas "https://www.worldatlas.com/articles/top-10-us-metropolitan-areas-with-the-highest-population-of-indians.html". From the dataset, we can conclude that NYC stands first with highest number of Indian population.

➢ Along with the population dataset, I am using the site "https://www.worldatlas.com/articles/the-world-s-most-popular-tourist-attractions.html" to obtain the most popular tourist attraction spots in New York City. Thus utilizing the

above two dataset, I chose Times Square in New York City as a target destination for opening the restaurant.

➢ Foursquare API, a location data provider is utilized for exploring the Indian Restaurants present around Times Square. Also, the same dataset is used for obtaining the data about the Indian Restaurants present in a particular postal code or zip code.

➢ With the above datasets, depending upon the total amount of restaurants in a particular postal code, a perfect location for setting up the Indian Restaurant around Times Square can be determined. To know the neighborhoods in which a new Indian Restaurant is about to open, I am using the dataset from the site "https://www.health.ny.gov/statistics/cancer/registry/appendix/neighborhoods.htm".

So, from the above data sources we can decide the City, Neighborhood in which an Indian Restaurant could be opened with high profit.