

**IBM – Coursera**  
**Data Science Specialization**

**The Capstone Project**  
**The best neighbourhood in New York to setup an**  
**Indian Style Restaurant**

Clement Emerson ([clementemerson@gmail.com](mailto:clementemerson@gmail.com))

May 2020

## Contents

Part I The Problem .....	3
Introduction: .....	3
Business Problem:.....	3
Target audience of this project:.....	3
Part II Data .....	4
Location Data .....	4
Data Source .....	4

## Part I The Problem

### Introduction:

Indian population in New York, US is always increasing. This is evident from the wiki page

[https://en.wikipedia.org/wiki/Indians\\_in\\_the\\_New\\_York\\_City\\_metropolitan\\_region](https://en.wikipedia.org/wiki/Indians_in_the_New_York_City_metropolitan_region)

Rank ↕	Borough ↕	City ↕	Indian Americans ↕	Density of Indian Americans per square mile ↕	Percentage of Indian Americans in municipality's population ↕
1	Queens (2014) <sup>[33]</sup>	New York City	144,896	1,326.5	6.2
2	Brooklyn (2012)	New York City	25,270	357.9	1.0
3	Manhattan (2012)	New York City	24,359	1,060.9	1.5
4	The Bronx (2012)	New York City	16,748	398.6	1.2
5	Staten Island (2012)	New York City	6,646	113.6	1.4
	Total (2014) <sup>[31]</sup>	New York City	227,994	753.4	2.7

Even the data we have is updated as of 2014, the Indian population in NY is keep on increasing. So, there is a great opportunity to open any Indian style business in NY, especially Indian style restaurants. From the above figure, we can easily say, that, we can open a restaurant in Queens borough. But exactly where? Which neighbourhood of Queens borough is the best? There may be some other neighbourhood in some other borough which is also good for the same. We need a clearly defined approach to find the best places for setting up the Indian style restaurants. Also, the approach, that we use to find the best neighbourhood, should follow proven scientific methods. Indian restaurant and Pizza place are not same. So, the approach to find the best location for an Indian style restaurant is also different from the approach to find best location to setup a Pizza place. In this report, I am carefully drafted such a method using machine learning and finally finds a few places which are best to setup an Indian style restaurant. Even this report is New York centric, this can be easily extended to other cities also, to find the best neighbourhood in that city to setup an Indian style restaurant.

### Business Problem:

I wish to keep this simple and clear. I am trying to find the answer for the following question: **Where would an Indian hotelier open an Indian style restaurant in New York, USA?**

### Target audience of this project:

This project is directly useful for the investors and Indian hoteliers who are interested in setting up an India style restaurant in New York, USA. Deciding the location for setting up any business is very crucial. And the first and foremost questions in the mind of investors. This project helps them in

deciding the location. This project finds all neighbourhoods which holds good for this purpose. So, if the investor or hotelier already has a business in one of the neighbourhoods, he can extend the business to other places listed in the project results.

## Part II Data

### Location Data

I need the below data to carry out my study on Indian restaurants.

- List of neighbourhoods in NYC
- At max, 100 venues in each neighbourhood
- Venue category

### Data Source

All data related to neighbourhoods in NYC are located in the following URL, [https://geo.nyu.edu/catalog/nyu\\_2451\\_34572](https://geo.nyu.edu/catalog/nyu_2451_34572). The URL has a geo JSON, in which borough, neighbourhood name, its latitude and longitude and a lot of other information is there. The data in the geo JSON is helpful to draw map. The data of our interest is, borough name, neighbourhood name, latitude and longitude. The data set has 5 boroughs and 306 neighbourhoods.

For the remaining data, we can rely on foursquare location APIs. You can find the URL for the endpoints here <https://developer.foursquare.com/docs/places-api/endpoints/>. An account needs to be created in foursquare.com to use their APIs. Foursquare provides free as well as premium APIs. For this project, I am using the free APIs.

With the neighbourhood data in our hand (latitude and longitude), we can get the venue information and the venue category information from foursquare.com. The following API is used to get the venue related information.

[https://api.foursquare.com/v2/venues/explore?&client\\_id=CLIENT\\_ID&client\\_secret=CLIENT\\_SECRET&v=20180605&ll=40.89470517661,-73.84720052054902&radius=500&limit=100](https://api.foursquare.com/v2/venues/explore?&client_id=CLIENT_ID&client_secret=CLIENT_SECRET&v=20180605&ll=40.89470517661,-73.84720052054902&radius=500&limit=100)

The above API will return information about 100 venues which are in 500m radius from the latitude and longitude mentioned in the API. This data is used for the learning purposes.