

IBM Applied Data Science Capstone Project

Coursera

Opening a new Indian Cuisine Restaurant in New York City, USA

Dataset Details

Submitted By:

Supriyo Chowdhury

March 2020



Data Section

To solve the problem we need the following data:

1. List of neighbourhoods in New York City.
2. Longitude and latitude of those neighbours. This is required in order to plot maps
3. Venue data, particularly data related to Indian Cuisine restaurants. This data will be used perform the clustering on the neighbours.

Data Source

For this project we need the following data:

1. New York City data that contains list Boroughs, Neighbourhoods along with their latitude and longitude.
 - a. **Data Source:** https://cocl.us/new_york_dataset
 - b. **Description:** This data set is the most important dataset that contains the information of the neighbourhoods. It also contains the latitude and longitude also.
2. List of Indian restaurants in each neighbourhood of New York City:
 - a. **Data Source:** Foursqaure API
 - b. **Description:** This data provides all the venues in each neighbourhood.
3. GeoSpace Dataset:
 - a. **Data Source:** <https://data.cityofnewyork.us/CityGovernment/Borough-Boundaries/tqmj-j8zm>
 - b. **Description:** This will help us to visualize the choropleth map.