**Applied Data Science Capstone Project**

**The Battle of Neighbourhoods**

(Week - 4&5)

**Authored By:**

**Gopal Sahu**

1. **Problem Statement:**

New York is one of the cities where a huge number of Indian reside. With growing number of Indian/Sub-continent population, the number of restaurants is not sufficient or very sparse across the city. Also people from various other countries like to try out authentic Indian food which is very unique. An Indian food chain sees this as an opportunity and wants to expand outside India and wants to find out the suitable areas for opening restaurants. The problem statement is to find recommended place to open new branch of the food chain.

Target Audience:

Restaurant chains, Hotel owners

Stake holders:

Food Authority of USA, Sub-continent people who like Indian food, Restaurant Owners, Shopping malls who lease space for hotels

1. **Data:**

First I will get the location data of New York and its neighbourhood. Then I will use these location data from foursquare to get Indian Restaurants in the neighbourhood.

* Neighbourhood location co-ordinates from NYU (JSON)
* Google Maps data for venues
* Foursquare location data of various restaurants (JSON)

(Obtained from foursquare.com API calls)

Once the neighbourhood is marked with their similarities, the clusters where the first 10 common venue include restaurants are suitable for opening new restaurant.

* **Title: -** New York neighbourhood data

**Type of data: -** JSON

**Duration: -** 2014

**Description of the dataset: -**

All the area and the respective neighbourhood data with location properties.

**Source:-** [**https://geo.nyu.edu/catalog/nyu\_2451\_34572**](https://geo.nyu.edu/catalog/nyu_2451_34572)

* **Title: -** Google Maps Geocoding API

**Type of data: -** JSON

**Duration: -** N/A

**Description of the data: -** Location coordinates obtained by Gmaps API calls.

Location Information obtained from Price Paid Dataset is used to obtain the location coordinates from Google Maps.

**Source: -** Google Cloud Platform/ Google Maps

* **Title: -** Foursquare location data

**Type of data: -** JSON

**Duration: -** N/A

**Description of the data: -**

Location coordinates obtained by Foursquare API calls. To determine the proximity of various amenities as per the client’s requirement, Foursquare location data is used.

**Source: -** <https://foursquare.com/>

1. **Methodology:**

The whole methodology remains same as in New York venue exploration. However the analysis will include the category of venue in foursquare search for Indian restaurants.

1. **Results:**

The cluster similar to ‘Upper West Side’ i.e. cluster 1 are favourable for Indian restaurants.

Also the cluster 4, where there are more cafes and theatres present, there is an opportunity for **take away** branch.

1. **Discussion:**

Also the areas where there are less restaurants but has theatres can be potential areas.

1. **Conclusion:**

As there is less data related to Indian restaurants, we need to get search trend data related to areas, which will improve the analysis.