

# Capstone Project - The Battle of Neighborhoods(Week1)

## Introduction

New York City is a major populated area and people are coming to this great city from all over the world. There are many languages spoken in New York, making it the most linguistically diverse city in the world. English remains the most widely spoken language, although there are areas in the outer boroughs in which up to 25% of people speak English as an alternate language, and/or have limited or no English language fluency. English is least spoken in neighborhoods such as Flushing, Sunset Park, and Corona

With its diverse culture, comes diverse food items. There are many restaurants in New York City, each belonging to different categories like Chinese, Indian, French etc

So as part of this project, will list and visualize all major parts of New York City that has great Indian restaurants. Although there are many Asian restaurants in the Neighborhood, some boroughs hardly have Indian restaurant. In my opinion, one would show interest in starting an Indian restaurant in the Brooklyn borough neighborhood based on the results shown (results section and code will show in the next week's assignment) after all the methodologies are done.

## Data

For this project, will take data from week 3 of this course (Neighborhoods in New York City)

New York City data contains many Boroughs, Neighborhoods and along with their latitude and longitude

Data source : [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)

Description : This data set contains the required information. And will use this data set to explore various neighborhoods and boroughs of the New York City

Indian restaurants in each neighborhood of New York City

Data source : Foursquare API

Description : By using this API we will get all the venues in each neighborhood. We can filter these venues to get Indian restaurants

## Methodology

Download and Explore the data set by using the given link [https://cocl.us/new\\_york\\_dataset](https://cocl.us/new_york_dataset)

**Neighborhood has a total of 5 boroughs and 306 neighborhoods. In order to segment the neighborhoods and explore them, will essentially need a dataset that contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood**

Use geopy library to get the latitude and longitude values of New York City and Folium is a great visualization library to reveal the name of the neighborhood and its respective borough

Using Foursquare API to explore the neighborhoods and segment them and will find venues for each neighborhood

**Filter out venues that are Indian Restaurants**

**Run k-means to cluster the neighborhood and examine the clusters**