

# Capstone Project - The Battle of Neighborhoods (Week 1)

## Background:

New York City is one of the busiest cities in the world and one of the world's major commercial, financial and cultural centers.. Being the most populous as well as the most densely populated city in the US, lots of people find great business opportunities here.

An entrepreneur is looking to open a coffee shop in one of New York's neighborhoods. He does not understand the area and the availability of coffee shops in each neighborhood. So he wants to investigate the area and identify the best place to open a new coffee shop.

## Data Requirements and Sources

### 1. Dataset 1:

The below link contains information about the boroughs and the different neighbourhoods under each borough of NYC along with its latitude and longitude coordinates.

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

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

### 2. Dataset 2:

The below link contains information about the demographics of NYC and segmentation of ethnic races across different boroughs.

[https://en.wikipedia.org/wiki/Demographics\\_of\\_New\\_York\\_City](https://en.wikipedia.org/wiki/Demographics_of_New_York_City)

### 3. Dataset 3:

The Foursquare API will be used to obtain the geographical location data for the NYC Area.

These will be used to explore the venues in the neighbourhoods of NYC. The venues will provide the categories needed for the analysis, such as the coffee shop.

## Methodology

The first two datasets will be analysed by considering the venues within the neighbourhood of NYC postal code areas. The combined data will be used to determine the highest density of people in each Borough. Then I will evaluate the best Borough of coffee shops and determine

the best cluster. Finally, I will visualise the best borough with the coffee shop to visually see spread. Due to Foursquare restrictions, the number of venues will be limited to 100 venues.