# The Battle of the Neighborhoods - Week 2

## Introduction & Business Problem :

### Problem Background:

The City of New York, is the city with most quantity of people in the United States. One of its major characteristics is that it is multicultural and a lot of people from different countries access to the chance of the american dream. It provides great business opportinities and cultural experience as well. Also, the city is a major center for different sectors such as financial, retail, international commerce, tourism, fashion, arts (and others) in the US.

This also means that the market flow is extremely competitive and the investment needed for entrepreneur businessin one of the highest worldwide. That's why every possinle new business venture have to analyse the return of investment to see if they could survive with that high costs. In this lab, we are going to expose some insights base on data anlysis available with the goal of being part of the decision that every entrepreneur o any company is going to make.

### Problem Description:

New York is well known for its cuisine, that has a unique influence by a great variety of international cuisine. Also, regarding to restaurants, the level of competition in this city is high. For possible new restaurants, depending on the core value of the company, the main differences or similarities with other brands every new entrepreneur should evaluate in high level where the restaurant should be located.

The strategy plan can consider the quantity of people that lives nearby, the quality of other restaurants in the zone, the type of food, the area mood, etc.

Also, additional variables should be analyzed befare any decision is made:

Demographics, Population, possible provider (Freshmarkets, Wholesale locations) concurrence of the place competitors rate Segmentation of the Borough Untapped or Saturated markets etc

Eventhough well funded XYZ Company Ltd. is going to select the best location fit to start its first venture.If this is successful they can replicate the same in other locations. As they are building their brand, the results of the performance of this first restaurant will be highly impactfull

### Target Audience:

To recommend the correct location, XYZ Company has appointed me to lead of the Data Science team. The objective is to recommend to the management which NY neighborhood will be best fit to open their restaurant.

This would be of interest of anyone looking to open a food related company in New York

# Success Criteria:

The success criteria of this project will be a good recommendation of borough/Neighborhood choice to XYZ Company Ltd based on Lack of such restaurants in that location and nearest suppliers of ingredients, eith a high level of significance.

## Data we will use:

Data : The city of principal analysis in this project : New York City.

We will be using the following datasets for analysing New York city

**Data 1 :**

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

This dataset exists for free on the web. Link to the dataset is
: https://geo.nyu.edu/catalog/nyu_2451_34572

**Data 2**

Second data which will be used is the DOHMH Farmers Markets and Food Boxes dataset. In this we will be using the data of Farmers Markets.

https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2

Website-https://www.grownyc.org/greenmarketco/foodbox GrowNYC's Fresh Food Box Program is a food access initiative that enables under-served communities to purchase fresh, healthy, and primarily regionally grown produce well below traditional retail prices.

A farmers' market is often defined as a public site used by two or more local or regional producers for the direct sale of farm products to consumers. In addition to fresh fruits and vegetables, markets may sell dairy products, fish, meat, baked goods, and other minimally processed foods.

**Data 3**

For the below analysis we will get data from the web as given below :

New York Population New York City Demographics Cuisine of New York city https://en.wikipedia.org/wiki/New_York_City https://en.wikipedia.org/wiki/Economy_of_New_York_City https://en.wikipedia.org/wiki/Portal:New_York_City https://en.wikipedia.org/wiki/Cuisine_of_New_York_City https://en.wikipedia.org/wiki/List_of_Michelin_starred_restaurants_in_New_York_City

**Data 4**

Newyork city geographical coordinates data will be utilized as input from Foursquare API, that will be leveraged to provision venues information for each neighborhood.We will use the Foursquare API to explore neighborhoods in New York City. The below is image of the Foursquare API data.
Data 3 : For the below analysis we will get data from wikipedia as given below :

New York Population New York City Demographics Cuisine of New York city https://en.wikipedia.org/wiki/New_York_City https://en.wikipedia.org/wiki/Economy_of_New_York_City https://en.wikipedia.org/wiki/Portal:New_York_City https://en.wikipedia.org/wiki/Cuisine_of_New_York_City https://en.wikipedia.org/wiki/List_of_Michelin_starred_restaurants_in_New_York_City

We will have lattittudes and postal codes as follows

| | Postalcode | Borough | Neighborhood | Latitude | Longitude |
|---|---|---|---|---|---|
| 0 | M3A | North York | Parkwoods | 43.753259 | -79.329656 |
| 1 | M4A | North York | Victoria Village | 43.725882 | -79.315572 |
| 2 | M5A | Downtown Toronto | Regent Park , Harbourfront | 43.654260 | -79.360636 |
| 3 | M6A | North York | Lawrence Manor , Lawrence Heights | 43.718518 | -79.464763 |
| 4 | M7A | Downtown Toronto | Queen's Park , Ontario Provincial Government | 43.662301 | -79.389494 |
| 5 | M9A | Etobicoke | Islington Avenue | 43.667856 | -79.532242 |
| 6 | M1B | Scarborough | Malvern , Rouge | 43.806686 | -79.194353 |