# Final Project – The Battle of Neighborhoods

## Introduction

This project is the final project in the IBM Data Science Professional Certificate specialization course in Coursera. The goal is to define a problem statement and solve the problem using Machine Learning Methodology. Additionally, FourSquare data should be used to solve the problem.

## Business Problem

A friend of mine is moving to New York and he is interested to know the sales price of houses / properties in and around New York City, so he can decide which area he can move to. He is interested to know the price variation based on the locality (Neighborhood, the different venues around the Neighborhood, etc).

## Data Selection

I will be using couple of data sets to solve the problem:

* Average sales price of properties in New York City, to get the average Sales Price of properties sold in NY city. This data is openly available [here](https://opendata.cityofnewyork.us/).
* Geo location data to get the list of Neighborhoods
* Data from FourSquare that will provide the different venues available in that location.