



## **The Battle of the Neighborhoods - Week 1-B**

### **Data**

As mentioned in the document The Battle of the Neighborhoods\_Week1-a, the city of New York and its boroughs will be the object of this analysis.

New York City is often referred to collectively as the five boroughs, and in turn, there are hundreds of distinct neighborhoods throughout the boroughs, many with a definable history and character to call their own. If the boroughs were each independent city, four of the boroughs (Brooklyn, Queens, Manhattan, and The Bronx) would be among the ten most populous cities in the United States (Staten Island would be ranked 37<sup>th</sup>).

The following dataset will be used to analyze the demographic aspects of the city of New York: [https://geo.nyu.edu/catalog/nyu\\_2451\\_34572](https://geo.nyu.edu/catalog/nyu_2451_34572)

Since we were going to on building a new restaurant, we should also analyze the suppliers. In this case, the dataset for the Farmers market will also be utilized: <https://data.cityofnewyork.us/dataset/DOHMH-Farmers-Markets-and-Food-Boxes/8vwk-6iz2>

Other information related to population, demographics and from the city itself, I'll be using the Wikipedia: [https://en.wikipedia.org/wiki/New\\_York\\_City](https://en.wikipedia.org/wiki/New_York_City)

This project will also use Four-square API to analyze geographical coordinates and information of other venues for each neighborhood analyzed.

Python packages:

- Pandas - Library for Data Analysis
- NumPy – Library to handle data in a vectorized manner
- JSON – Library to handle JSON files
- Geopy – To retrieve Location Data
- Requests – Library to handle http requests
- Matplotlib – Python Plotting Module
- Sklearn – Python machine learning Library
- Folium – Map rendering Library