Data

To solve the problem, we will need the following data:

New York City data containing the neighborhoods and boroughs. Latitude and longitude coordinates of those neighborhoods. This is required to plot the map and get the venue data. Venue data, particularly data related to restaurants. We are going to use this data to perform further analysis of the neighborhoods.

Data Source and methods to extract them

New York City data containing the neighborhoods and boroughs will be obtained from the open data source: https://cocl.us/new_york_dataset. After it, we will get the geographical coordinates of the neighborhoods (latitude and longitude) using Python Geocoder package.

Finally, we will use Foursquare API to get the venue data for the neighborhoods defined at the previous step. Foursquare has one of the largest databases of 105+ million places and over 125,000 developers use this application. Foursquare API provides many categories of the venue data; we are particularly interested in the restaurant data to solve the business problem defined above.

This project will require using of many data science skills, from web scrapping (open source dataset), working with API (Foursquare), data cleaning, data wrangling, to map visualization (Folium).

Methodology

- Data will be collected from https://cocl.us/new_york_dataset and cleaned and processed into a dataframe.
- FourSquare be used to locate all venues and then filtered by African restaurants. Ratings, tips, and likes by users will be counted and added to the dataframe.
- Data will be sorted based on rankings.
- Finally, the data be will be visually assessed using graphing from Python libraries.

Problem Statement
- What is / are the best location(s) for African cuisine in New York City?
- In what Neighborhood and/or borough should the investor open an African restaurant to have the best chance of being successful?
- Where would I go in New York City to have the best African food?