**Datasets used for New York restaurants cuisine analysis project**

1] Foursquare api from foursquare city guide platform

2] New York City neighborhoods data

3] Geopy package used to convert address to coordinates.

1] Foursquare api from foursquare city guide platform:

We use the foursquare api to acquire the top 20 restaurants in each neighborhood of new York and their cuisine types.

The syntax for fetching for one neighborhood is:

url = 'https://api.foursquare.com/v2/venues/search?client\_id={}&client\_secret={}&ll={},{}&v={}&query={}&radius={}&limit={}'.format(CLIENT\_ID, CLIENT\_SECRET, lat, lon, VERSION, search\_query, radius, LIMIT)

results = requests.get(url).json()

This returns a json object with detailed information of each of the top restaurants. But this project just needs the name and cuisine from this json object.

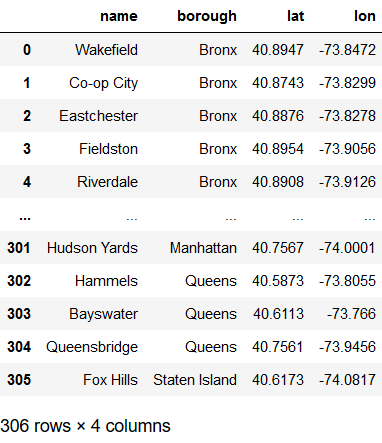
The name is in results.response.venues[].name

The cuisine is in results.response.venues[].categories[0].shortname

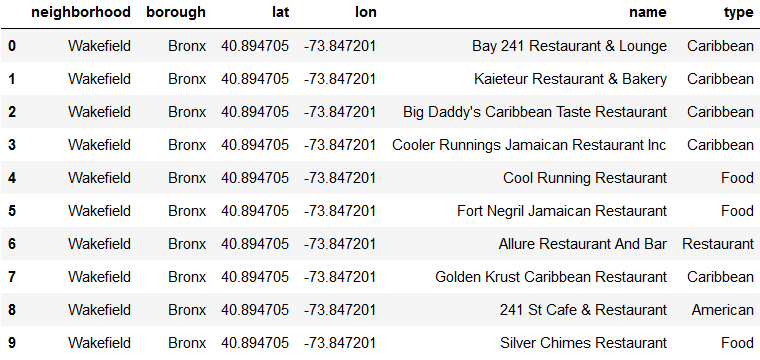
2] New York City neighborhoods data

We download the dataset from https://geo.nyu.edu/catalog/nyu\_2451\_34572 or https://cocl.us/new\_york\_dataset to fetch the names, borough, coordinates of every New York neighborhood.

The dataset looks like this:



We use the dataset acquired from foursquare api and the above to form the following dataset:

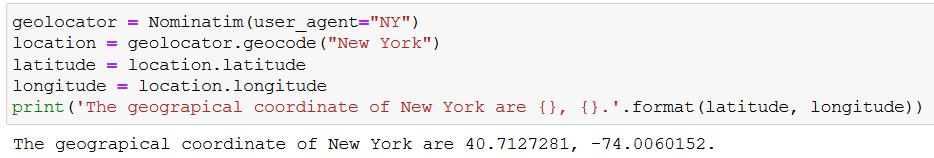


The dataset will further be compressed by grouping rows by neighborhood and one hot encoding it to get no of restaurants in each neighborhood by their cuisines and train the kmeans model to cluster the neighborhoods.

3] Geopy package

The geopy package is used to convert an address to its relevant coordinated we can use in folium map plotting or in foursquare api.

The syntax to convert address “New York” to its coordinates is:



These three are what will be used to get all the data needed for the project.