Capstone Project

Data Sources

Neighborhoods Data

In this section, I will discuss the data sources that will be used for this project. New York city has 306 neighborhoods spread out among 5 boroughs. This New York city neighborhoods data will be downloaded from https://geo.nyu.edu/catalog/nyu_2451_34572. This is a json file that contains neighborhood data in the features key that looks like this:


```
{'type': 'Feature',
 'id': 'nyu_2451_34572.1',
 'geometry': {'type': 'Point',
  'coordinates': [-73.84720052054902, 40.89470517661]},
 'geometry_name': 'geom',
 'properties': {'name': 'Wakefield',
  'stacked': 1,
  'annoline1': 'Wakefield',
  'annoline2': None,
  'annoline3': None,
  'annoangle': 0.0,
  'borough': 'Bronx',
  'bbox': [-73.84720052054902,
   40.89470517661,
   -73.84720052054902,
   40.89470517661]}}
```


The neighborhood name, borough, latitude and longitude data will be extracted from this json file and loaded into a pandas dataframe that looks like this:

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Foursquare API

The Foursquare API provides location based experiences with diverse information about venues, users, photos, and check-ins. We will use this API to get information about the venues in the various neighborhoods. The neighborhood coordinates from the neighborhoods dataframe will be used with the Foursquare API to analyze the neighborhoods. The response will be parsed and loaded in a dataframe that looks like this:

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
Marble Hill	40.876551	-73.91066	Rite Aid	40.875467	-73.908906	Pharmacy

This data will be analyzed to find an ideal location for the vegan restaurant.