Capstone Project - The Battle of the Neighborhoods (Week 2)

Applied Data Science Capstone by IBM/Coursera

Table of contents

Introduction: Business Problem

The objective of this project is to develop a tool that will help 2 immigrants identify the best location for opening a new ethinc resturant in New York City. The first immigrant is Mr. Lucino, a Polish immigrant, who wants to open a new Polish resturant and Mrs. Jundi, a Lebanese immigrant, who wants to open a new lebanese resturant. The new immigrants are overwhelmed looking for the best location in the NY, they want this tool to help them narrow down their options by determining the best neighborhood to open their resturants.

Data

The tool will use NY city neighborhoods and boroughs database used in Week 3 lab for the list of neighborhoods and their coordinates. The database was exported as CSV file and used here. Foursquare API will be used to get the venues information. The tool will correlate all the venues categories to identify venues that open in the same proximity then it will identify neighborhoods where the new resturant category (Lebanese or Polish) is lacking.

The factors that will be considered are:

- **Availablility of other venues that open within proximity of a Lebanese and Polish resturants.**
- **Number of existing Lebanese or Polish resturants in the neighborhood.**

First, lets download all the dependencies we need.

In [1]:

```
import numpy as np # library to handle data in a vectorized manner

import pandas as pd # library for data analsysis
pd.set_option('display.max_columns', None)
pd.set_option('display.max_rows', None)

import json # library to handle JSON files

!conda install -c conda-forge geopy --yes
from geopy.geocoders import Nominatim # convert an address into latitude and longitude va
lues

import requests # library to handle requests
from pandas.io.json import json_normalize # tranform JSON file into a pandas dataframe

# Matplotlib and associated plotting modules
import matplotlib.cm as cm
```

```
import matplotlib.colors as colors

# import k-means from clustering stage
from sklearn.cluster import KMeans

import folium # map rendering library

print('Libraries imported.')
```

```
Collecting package metadata (current_repodata.json): ...working... done
Solving environment: ...working... done

# All requested packages already installed.

Libraries imported.
```

Then we upload NY neighboorhoods data from CSV file.

In [2]:

```
hoods = pd.read_csv('NY Neighborhoods.csv')
hoods.head(10)
```

Out[2]:

	Unnamed: 0	Borough	Neighborhood	Latitude	Longitude
0	0	Bronx	Wakefield	40.894705	-73.847201
1	1	Bronx	Co-op City	40.874294	-73.829939
2	2	Bronx	Eastchester	40.887556	-73.827806
3	3	Bronx	Fieldston	40.895437	-73.905643
4	4	Bronx	Riverdale	40.890834	-73.912585
5	5	Bronx	Kingsbridge	40.881687	-73.902818
6	6	Manhattan	Marble Hill	40.876551	-73.910660
7	7	Bronx	Woodlawn	40.898273	-73.867315
8	8	Bronx	Norwood	40.877224	-73.879391
9	9	Bronx	Williamsbridge	40.881039	-73.857446

Here we define our Foursquare API credentials.

In [3]:

```
CLIENT_ID = 'W5Y2SJWEORJOICYFVOD4PQNIWW3DO0NKLT302I21GMAM3CQB' # your Foursquare ID
CLIENT_SECRET = 'K3ILU01A43TQRFNXAABVNDFLKK5INNLFXXFII2GN0PK15PNF' # your Foursquare Secr
et
VERSION = '20180605' # Foursquare API version
LIMIT = 100 # A default Foursquare API limit value
```

We create getNearbyVenues function that we will use to retreive venues information for each neighborhood.

In [4]:

```
def getNearbyVenues(names, latitudes, longitudes, radius=500):

    venues_list=[]
    for name, lat, lng in zip(names, latitudes, longitudes):
        print(name)

        # create the API request URL
        url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret=
{}&v={}&ll={},{}&radius={}&limit={}'.format(
            CLIENT_ID,
            CLIENT_SECRET,
```

```
            VERSION,
            lat,
            lng,
            radius,
            LIMIT)

        # make the GET request
        results = requests.get(url).json()["response"]['groups'][0]['items']

        # return only relevant information for each nearby venue
        venues_list.append([(
            name,
            lat,
            lng,
            v['venue']['name'],
            v['venue']['location']['lat'],
            v['venue']['location']['lng'],
            v['venue']['categories'][0]['name']) for v in results])

    nearby_venues = pd.DataFrame([item for venue_list in venues_list for item in venue_l
ist])
    nearby_venues.columns = ['Neighborhood',
                  'Neighborhood Latitude',
                  'Neighborhood Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

We use getNearbyVenues to get the list of top 100 venues in each neighborhood in NY.

In [5]:

```
venues = getNearbyVenues(names=hoods['Neighborhood'],
                                   latitudes=hoods['Latitude'],
                                   longitudes=hoods['Longitude']
                                  )
```

```
Wakefield
Co-op City
Eastchester
Fieldston
Riverdale
Kingsbridge
Marble Hill
Woodlawn
Norwood
Williamsbridge
Baychester
Pelham Parkway
City Island
Bedford Park
University Heights
Morris Heights
Fordham
East Tremont
West Farms
High  Bridge
Melrose
Mott Haven
Port Morris
Longwood
Hunts Point
Morrisania
Soundview
Clason Point
Throgs Neck
Country Club
Parkchester
```

Westchester Square
Van Nest
Morris Park
Belmont
Spuyten Duyvil
North Riverdale
Pelham Bay
Schuylerville
Edgewater Park
Castle Hill
Olinville
Pelham Gardens
Concourse
Unionport
Edenwald
Bay Ridge
Bensonhurst
Sunset Park
Greenpoint
Gravesend
Brighton Beach
Sheepshead Bay
Manhattan Terrace
Flatbush
Crown Heights
East Flatbush
Kensington
Windsor Terrace
Prospect Heights
Brownsville
Williamsburg
Bushwick
Bedford Stuyvesant
Brooklyn Heights
Cobble Hill
Carroll Gardens
Red Hook
Gowanus
Fort Greene
Park Slope
Cypress Hills
East New York
Starrett City
Canarsie
Flatlands
Mill Island
Manhattan Beach
Coney Island
Bath Beach
Borough Park
Dyker Heights
Gerritsen Beach
Marine Park
Clinton Hill
Sea Gate
Downtown
Boerum Hill
Prospect Lefferts Gardens
Ocean Hill
City Line
Bergen Beach
Midwood
Prospect Park South
Georgetown
East Williamsburg
North Side
South Side
Ocean Parkway
Fort Hamilton
Chinatown
Washington Heights
Inwood

Hamilton Heights
Manhattanville
Central Harlem
East Harlem
Upper East Side
Yorkville
Lenox Hill
Roosevelt Island
Upper West Side
Lincoln Square
Clinton
Midtown
Murray Hill
Chelsea
Greenwich Village
East Village
Lower East Side
Tribeca
Little Italy
Soho
West Village
Manhattan Valley
Morningside Heights
Gramercy
Battery Park City
Financial District
Astoria
Woodside
Jackson Heights
Elmhurst
Howard Beach
Corona
Forest Hills
Kew Gardens
Richmond Hill
Flushing
Long Island City
Sunnyside
East Elmhurst
Maspeth
Ridgewood
Glendale
Rego Park
Woodhaven
Ozone Park
South Ozone Park
College Point
Whitestone
Bayside
Auburndale
Little Neck
Douglaston
Glen Oaks
Bellerose
Kew Gardens Hills
Fresh Meadows
Briarwood
Jamaica Center
Oakland Gardens
Queens Village
Hollis
South Jamaica
St. Albans
Rochdale
Springfield Gardens
Cambria Heights
Rosedale
Far Rockaway
Broad Channel
Breezy Point
Steinway
Beechhurst

Bay Terrace
Edgemere
Arverne
Rockaway Beach
Neponsit
Murray Hill
Floral Park
Holliswood
Jamaica Estates
Queensboro Hill
Hillcrest
Ravenswood
Lindenwood
Laurelton
Lefrak City
Belle Harbor
Rockaway Park
Somerville
Brookville
Bellaire
North Corona
Forest Hills Gardens
St. George
New Brighton
Stapleton
Rosebank
West Brighton
Grymes Hill
Todt Hill
South Beach
Port Richmond
Mariner's Harbor
Port Ivory
Castleton Corners
New Springville
Travis
New Dorp
Oakwood
Great Kills
Eltingville
Annadale
Woodrow
Tottenville
Tompkinsville
Silver Lake
Sunnyside
Ditmas Park
Wingate
Rugby
Park Hill
Westerleigh
Graniteville
Arlington
Arrochar
Grasmere
Old Town
Dongan Hills
Midland Beach
Grant City
New Dorp Beach
Bay Terrace
Huguenot
Pleasant Plains
Butler Manor
Charleston
Rossville
Arden Heights
Greenridge
Heartland Village
Chelsea
Bloomfield
Bulls Head

```
Carnegie Hill
Noho
Civic Center
Midtown South
Richmond Town
Shore Acres
Clifton
Concord
Emerson Hill
Randall Manor
Howland Hook
Elm Park
Remsen Village
New Lots
Paerdegat Basin
Mill Basin
Jamaica Hills
Utopia
Pomonok
Astoria Heights
Claremont Village
Concourse Village
Mount Eden
Mount Hope
Sutton Place
Hunters Point
Turtle Bay
Tudor City
Stuyvesant Town
Flatiron
Sunnyside Gardens
Blissville
Fulton Ferry
Vinegar Hill
Weeksville
Broadway Junction
Dumbo
Manor Heights
Willowbrook
Sandy Ground
Egbertville
Roxbury
Homecrest
Middle Village
Prince's Bay
Lighthouse Hill
Richmond Valley
Malba
Highland Park
Madison
Bronxdale
Allerton
Kingsbridge Heights
Erasmus
Hudson Yards
Hammels
Bayswater
Queensbridge
Fox Hills
```

let's examine the shape and data of venues dataframe

In [6]:

```
print(venues.shape)
venues.head(5)
```

```
(10177, 7)
```

Out[6]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Wakefield	40.894705	-73.847201	Lollipops Gelato	40.894123	-73.845892	Dessert Shop
1	Wakefield	40.894705	-73.847201	Rite Aid	40.896649	-73.844846	Pharmacy
2	Wakefield	40.894705	-73.847201	Walgreens	40.896528	-73.844700	Pharmacy
3	Wakefield	40.894705	-73.847201	Carvel Ice Cream	40.890487	-73.848568	Ice Cream Shop
4	Wakefield	40.894705	-73.847201	Dunkin'	40.890459	-73.849089	Donut Shop

Let's see how many unique venue categories there are

In [7]:

```
print('There are {} uniques categories.'.format(len(venues['Venue Category'].unique())))
```

```
There are 439 uniques categories.
```

Methodology

We will use the venue categories to create 2 venue profiles of each neighborhood. One for the count of the venues and one for the mean. The mean profile will be used to corrolate then venues categories to identify venues that open in the same proximity of Lebanese and Polish restaurants. The count profile will be used to identify the neighborhoods where Lebanese and Polish restaurants are lacking. The neigborhoods will be ranked depending on their profile corrolation and availability of Polish or Lebanese restaurants. Two maps will be generated with markers that indicate the neighborhoods rankings.

Analysis

Now, we will create a new dataframe with the neighborhood and the number of venues in each category

In [8]:

```
hotVenues = pd.get_dummies(venues[['Venue Category']], prefix="", prefix_sep="")

# add neighborhood column back to dataframe
hotVenues['Neighborhood'] = venues['Neighborhood']

# move neighborhood column to the first column
#fixed_columns = [toronto_onehot.columns[-1]] + list(toronto_onehot.columns[:-1])
#toronto_onehot = toronto_onehot[fixed_columns]
mid = hotVenues['Neighborhood']
hotVenues.drop(labels=['Neighborhood'], axis=1,inplace = True)
hotVenues.insert(0, 'Neighborhood', mid)
hotVenues.head()
```

Out[8]:

	Neighborhood	ATM	Accessories Store	Acupuncturist	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Arcade	Are Restaura
0	Wakefield	0	0	0	0	0	0	0	0	0	
1	Wakefield	0	0	0	0	0	0	0	0	0	
2	Wakefield	0	0	0	0	0	0	0	0	0	
3	Wakefield	0	0	0	0	0	0	0	0	0	
4	Wakefield	0	0	0	0	0	0	0	0	0	

Then we create two grouped datasets, one grouped by the mean of the venues count (meanHoods) and one by the count of venues (countHoods). meanHoods will be used to correlate venues and count to determine neighborhoods with low Polish or Lebanese resturants.

In [9]:

```
meanHoods = hotVenues.groupby('Neighborhood').mean().reset_index()
print(meanHoods.shape)
meanHoods.head()
```

```
(302, 439)
```

Out[9]:

	Neighborhood	ATM	Accessories Store	Acupuncturist	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Arcade	Arep Restaura
0	Allerton	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0
1	Annadale	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0
2	Arden Heights	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0
3	Arlington	0.0	0.0	0.0	0.0	0.0	0.0	0.25	0.0	0.0	0
4	Arrochar	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.0	0.0	0

In [10]:

```
countHoods = hotVenues.groupby('Neighborhood').sum().reset_index()
print(countHoods.shape)
countHoods.head()
```

```
(302, 439)
```

Out[10]:

	Neighborhood	ATM	Accessories Store	Acupuncturist	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Antique Shop	Arcade	Arep Restaura
0	Allerton	0	0	0	0	0	0	0	0	0	
1	Annadale	0	0	0	0	0	0	0	0	0	
2	Arden Heights	0	0	0	0	0	0	0	0	0	
3	Arlington	0	0	0	0	0	0	1	0	0	
4	Arrochar	0	0	0	0	0	0	0	0	0	

Then we create normalized dataframes for the correlation of venues with Lebanese and Polish resturants.

In [207]:

```
from sklearn import preprocessing

venuesCorr = meanHoods.corr()
lebCorr = venuesCorr[["Lebanese Restaurant"]]
polCorr = venuesCorr[[ "Polish Restaurant"]]


# Normalizing the dataframes

lebArr = lebCorr.values #returns a numpy array
min_max_scaler = preprocessing.MinMaxScaler()
lebArr = min_max_scaler.fit_transform(lebArr)
lebCorr["Weight"] = lebArr
```

```
lebCorr = lebCorr.drop(columns = "Lebanese Restaurant")
lebCorr = lebCorr.drop("Lebanese Restaurant") #remove Lebanese Restaurant form the list




polArr = polCorr.values #returns a numpy array

polArr = min_max_scaler.fit_transform(polArr)
polCorr["Weight"] = polArr
polCorr = polCorr.drop(columns = "Polish Restaurant")
polCorr = polCorr.drop("Polish Restaurant")#remove Polish Restaurant form the list

print(lebCorr)
print(polCorr)
```

```
                                            Weight
ATM                                       0.044142
Accessories Store                         0.128230
Acupuncturist                             0.044142
Adult Boutique                            0.322458
Afghan Restaurant                         0.042205
African Restaurant                        0.043398
American Restaurant                       0.054035
Antique Shop                              0.040049
Arcade                                    0.043319
Arepa Restaurant                          0.038612
Argentinian Restaurant                    0.035121
Art Gallery                               0.028397
Art Museum                                0.041968
Arts & Crafts Store                       0.038151
Arts & Entertainment                      0.042302
Asian Restaurant                          0.049422
Athletics & Sports                        0.034779
Auditorium                                0.044142
Australian Restaurant                     0.039620
Austrian Restaurant                       0.042277
Auto Garage                               0.043186
Auto Workshop                             0.042604
Automotive Shop                           0.042016
BBQ Joint                                 0.030472
Baby Store                                0.044142
Bagel Shop                                0.041094
Bakery                                    0.087846
Bank                                      0.020975
Bar                                       0.036849
Baseball Field                            0.034996
Baseball Stadium                          0.043519
Basketball Court                          0.036221
Bath House                                0.044142
Beach                                     0.035412
Beach Bar                                 0.044142
Beer Bar                                  0.358172
Beer Garden                               0.177488
Beer Store                                0.197035
Big Box Store                             0.039033
Bike Rental / Bike Share                  0.041055
Bike Shop                                 0.042621
Bike Trail                                0.270431
Bistro                                    0.036709
Board Shop                                0.037319
Boat or Ferry                             0.037916
Bookstore                                 0.026079
Boutique                                  0.249256
Bowling Alley                             0.037673
Boxing Gym                                0.034578
Brazilian Restaurant                      0.038855
Breakfast Spot                            0.027317
Brewery                                   0.037064
Bridal Shop                               0.044142
Bridge                                    0.040617
Bubble Tea Shop                           0.081596
```

Buffet	0.038824
Building	0.038797
Burger Joint	0.054590
Burmese Restaurant	0.044142
Burrito Place	0.125179
Bus Line	0.038996
Bus Station	0.019441
Bus Stop	0.025509
Business Service	0.040763
Butcher	0.039670
Cafeteria	0.044142
Café	0.042422
Cajun / Creole Restaurant	0.037683
Camera Store	0.041302
Campground	0.044142
Candy Store	0.061405
Cantonese Restaurant	0.040405
Caribbean Restaurant	0.032785
Caucasian Restaurant	0.042314
Cha Chaan Teng	0.044142
Check Cashing Service	0.035700
Cheese Shop	0.035225
Chinese Restaurant	0.042478
Chocolate Shop	0.039620
Church	0.042425
Circus	0.044142
Climbing Gym	0.039922
Clothing Store	0.133704
Club House	0.040841
Cocktail Bar	0.051688
Coffee Shop	0.078034
College Academic Building	0.043600
College Arts Building	0.044142
College Basketball Court	0.044142
College Bookstore	0.044142
College Cafeteria	0.042972
College Gym	0.044142
College Theater	0.044142
Colombian Restaurant	0.044142
Comedy Club	0.132329
Comfort Food Restaurant	0.043281
Comic Shop	0.043039
Community Center	0.042987
Concert Hall	0.040020
Construction & Landscaping	0.042488
Convenience Store	0.026354
Cooking School	0.044142
Cosmetics Shop	0.075550
Costume Shop	0.044142
Coworking Space	0.044142
Creperie	0.156969
Cuban Restaurant	0.113025
Cultural Center	0.044142
Cupcake Shop	0.038637
Cycle Studio	0.030646
Czech Restaurant	0.044142
Dance Studio	0.036350
Daycare	0.043638
Deli / Bodega	0.011237
Department Store	0.104187
Dessert Shop	0.123662
Dim Sum Restaurant	0.038346
Diner	0.017229
Discount Store	0.019723
Distillery	0.043155
Dive Bar	0.039453
Dive Shop	0.044142
Doctor's Office	0.039175
Dog Run	0.041190
Donut Shop	0.000000
Dosa Place	0.044142
Drugstore	0.044142

Dry Cleaner	0.041654
Dumpling Restaurant	0.037591
Duty-free Shop	0.044142
Eastern European Restaurant	0.089698
Electronics Store	0.081470
Empanada Restaurant	0.042660
English Restaurant	0.720229
Entertainment Service	0.042793
Ethiopian Restaurant	0.038166
Event Service	0.044142
Event Space	0.034977
Exhibit	0.043523
Factory	0.044142
Falafel Restaurant	0.033795
Farm	0.041679
Farmers Market	0.031076
Fast Food Restaurant	0.015633
Field	0.044142
Filipino Restaurant	0.038541
Fish & Chips Shop	0.040686
Fish Market	0.036370
Fishing Store	0.044142
Flea Market	0.039622
Flower Shop	0.032407
Food	0.031839
Food & Drink Shop	0.034692
Food Court	0.035969
Food Stand	0.042866
Food Truck	0.027102
Fountain	0.042533
French Restaurant	0.153229
Fried Chicken Joint	0.017528
Frozen Yogurt Shop	0.100796
Fruit & Vegetable Store	0.041155
Furniture / Home Store	0.065127
Gaming Cafe	0.277542
Garden	0.037384
Garden Center	0.038022
Gas Station	0.030996
Gastropub	0.032361
Gay Bar	0.039785
General Entertainment	0.042555
German Restaurant	0.038265
Gift Shop	0.027922
Gluten-free Restaurant	0.044142
Go Kart Track	0.044142
Golf Course	0.042401
Gourmet Shop	0.179342
Greek Restaurant	0.124724
Grocery Store	0.024493
Gym	0.074505
Gym / Fitness Center	0.093956
Gym Pool	0.044142
Gymnastics Gym	0.042235
Halal Restaurant	0.040864
Harbor / Marina	0.038380
Hardware Store	0.150019
Hawaiian Restaurant	0.039024
Health & Beauty Service	0.086386
Health Food Store	0.036335
Heliport	0.043186
Herbs & Spices Store	0.044142
High School	0.044142
Hill	0.044142
Himalayan Restaurant	0.040875
Historic Site	0.041647
History Museum	0.096318
Hobby Shop	0.097706
Home Service	0.037572
Hookah Bar	0.033752
Hostel	0.042161
Hot Dog Joint	0.040842

Hotel	0.058950
Hotel Bar	0.035928
Hotel Pool	0.044142
Hotpot Restaurant	0.040292
Ice Cream Shop	0.056292
Indian Restaurant	0.193357
Indie Movie Theater	0.088736
Indie Theater	0.038941
Indonesian Restaurant	0.044142
Indoor Play Area	0.042793
Insurance Office	0.044142
Intersection	0.041536
Irish Pub	0.087557
Israeli Restaurant	0.040085
Italian Restaurant	0.138534
Japanese Curry Restaurant	0.041366
Japanese Restaurant	0.044288
Jazz Club	0.147327
Jewelry Store	0.035970
Jewish Restaurant	0.042283
Juice Bar	0.051735
Karaoke Bar	0.035383
Kebab Restaurant	0.042347
Kids Store	0.034323
Korean BBQ Restaurant	0.044142
Korean Restaurant	0.034339
Kosher Restaurant	0.040463
Lake	0.042875
Latin American Restaurant	0.042558
Laundromat	0.038737
Laundry Service	0.042277
Lawyer	0.039981
Leather Goods Store	0.521138
Library	0.041355
Lingerie Store	0.213058
Liquor Store	0.033098
Locksmith	0.042159
Lounge	0.027779
Malay Restaurant	0.042335
Marine Terminal	0.044142
Market	0.039272
Martial Arts School	0.036492
Massage Studio	0.041110
Mattress Store	0.033877
Medical Center	0.044142
Mediterranean Restaurant	0.195922
Memorial Site	0.044142
Men's Store	0.033777
Metro Station	0.029240
Mexican Restaurant	0.037101
Middle Eastern Restaurant	0.080654
Mini Golf	0.044142
Miscellaneous Shop	0.035297
Mobile Phone Shop	0.021484
Modern European Restaurant	0.042286
Molecular Gastronomy Restaurant	0.044142
Monument / Landmark	0.040001
Moroccan Restaurant	0.042277
Motorcycle Shop	0.041954
Movie Theater	0.036880
Moving Target	0.034973
Multiplex	0.044142
Museum	0.037228
Music School	0.042347
Music Store	0.039811
Music Venue	0.038022
Nail Salon	0.035483
Nature Preserve	0.044142
New American Restaurant	0.088108
Newsstand	0.040838
Nightclub	0.038510
Nightlife Spot	0.044142

Non-Profit	0.042751
Noodle House	0.035808
North Indian Restaurant	0.042338
Office	0.039722
Opera House	0.042902
Optical Shop	0.077327
Organic Grocery	0.038554
Other Great Outdoors	0.039268
Other Nightlife	0.043883
Other Repair Shop	0.041089
Outdoor Gym	0.044142
Outdoor Sculpture	0.040480
Outdoors & Recreation	0.038992
Outlet Mall	0.044142
Outlet Store	0.043015
Paella Restaurant	0.042280
Pakistani Restaurant	0.044142
Paper / Office Supplies Store	0.036538
Park	0.045128
Pastry Shop	0.044142
Pedestrian Plaza	0.043293
Performing Arts Venue	0.031948
Perfume Shop	0.044142
Persian Restaurant	0.720229
Peruvian Restaurant	0.034512
Peruvian Roast Chicken Joint	0.044142
Pet Café	0.041191
Pet Service	0.041874
Pet Store	0.023637
Pharmacy	0.002479
Photography Studio	0.043628
Physical Therapist	0.044142
Piano Bar	0.042277
Pie Shop	0.040009
Pier	0.043710
Piercing Parlor	0.042277
Pilates Studio	0.228089
Pizza Place	0.022899
Platform	0.044142
Playground	0.036676
Plaza	0.032191
Poke Place	0.039028
Polish Restaurant	0.039708
Pool	0.042024
Pool Hall	0.040962
Portuguese Restaurant	0.044142
Post Office	0.041428
Pub	0.061474
Public Art	0.043028
Puerto Rican Restaurant	0.042288
Racetrack	0.043164
Ramen Restaurant	0.109007
Record Shop	0.159487
Recording Studio	0.043951
Recreation Center	0.043661
Rental Car Location	0.035927
Rental Service	0.043417
Residential Building (Apartment / Condo)	0.036380
Resort	0.042720
Rest Area	0.043985
Restaurant	0.053996
River	0.043686
Rock Climbing Spot	0.044142
Rock Club	0.039604
Roller Rink	0.044142
Romanian Restaurant	0.044142
Roof Deck	0.037149
Russian Restaurant	0.040068
Sake Bar	0.042122
Salad Place	0.030333
Salon / Barbershop	0.068817
Sandwich Place	0.028423

Scandinavian Restaurant	0.044142
Scenic Lookout	0.031036
School	0.041333
Sculpture Garden	0.043877
Seafood Restaurant	0.107538
Shabu-Shabu Restaurant	0.044142
Shanghai Restaurant	0.038328
Shipping Store	0.032170
Shoe Repair	0.720229
Shoe Store	0.070900
Shop & Service	0.041880
Shopping Mall	0.036058
Skate Park	0.041608
Skating Rink	0.039963
Ski Area	0.044142
Smoke Shop	0.084384
Smoothie Shop	0.044142
Snack Place	0.159812
Soccer Field	0.042052
Social Club	0.043131
Soup Place	0.043881
South American Restaurant	0.034158
South Indian Restaurant	0.044142
Southern / Soul Food Restaurant	0.038277
Souvlaki Shop	0.044142
Spa	0.055646
Spanish Restaurant	0.021132
Speakeasy	0.036187
Sporting Goods Shop	0.096727
Sports Bar	0.036203
Sports Club	0.042255
Sri Lankan Restaurant	0.044142
Stadium	0.044142
State / Provincial Park	0.044142
Stationery Store	0.042344
Steakhouse	0.076068
Storage Facility	0.039298
Street Art	0.042278
Strip Club	0.521138
Supermarket	0.017622
Supplement Shop	0.031157
Surf Spot	0.043661
Sushi Restaurant	0.094909
Swiss Restaurant	0.044142
Szechuan Restaurant	0.236580
Taco Place	0.025998
Tailor Shop	0.039821
Taiwanese Restaurant	0.039328
Tanning Salon	0.044142
Tapas Restaurant	0.125586
Tattoo Parlor	0.037314
Tea Room	0.033817
Tech Startup	0.044142
Tennis Court	0.059921
Tennis Stadium	0.042317
Tex-Mex Restaurant	0.044142
Thai Restaurant	0.048269
Theater	0.031413
Theme Park	0.042604
Theme Park Ride / Attraction	0.041851
Thrift / Vintage Store	0.034814
Tibetan Restaurant	0.042387
Tiki Bar	0.040359
Toll Plaza	0.044142
Tourist Information Center	0.042959
Toy / Game Store	0.035531
Track	0.041058
Trail	0.039195
Train	0.044142
Train Station	0.031212
Tree	0.044142
Turkish Restaurant	0.040374

Udon Restaurant	0.521138
Used Bookstore	0.040935
Vape Store	0.042390
Varenyky restaurant	0.044142
Vegetarian / Vegan Restaurant	0.117276
Venezuelan Restaurant	0.044142
Veterinarian	0.042416
Video Game Store	0.031843
Video Store	0.036578
Vietnamese Restaurant	0.194102
Volleyball Court	0.044142
Warehouse Store	0.043021
Waste Facility	0.044142
Waterfront	0.043096
Weight Loss Center	0.043319
Whisky Bar	0.040134
Wine Bar	0.056396
Wine Shop	0.051665
Wings Joint	0.034706
Women's Store	0.028755
Yemeni Restaurant	0.044142
Yoga Studio	0.089596

	Weight
ATM	0.056480
Accessories Store	0.051219
Acupuncturist	0.056480
Adult Boutique	0.050589
Afghan Restaurant	0.053799
African Restaurant	0.055451
American Restaurant	0.025275
Antique Shop	0.076541
Arcade	0.055341
Arepa Restaurant	0.048825
Argentinian Restaurant	0.043993
Art Gallery	0.034685
Art Museum	0.067409
Arts & Crafts Store	0.068008
Arts & Entertainment	0.053932
Asian Restaurant	0.036485
Athletics & Sports	0.112897
Auditorium	0.056480
Australian Restaurant	0.050220
Austrian Restaurant	0.053898
Auto Garage	0.055156
Auto Workshop	0.054352
Automotive Shop	0.053538
BBQ Joint	0.045405
Baby Store	0.056480
Bagel Shop	0.130814
Bakery	0.184234
Bank	0.041872
Bar	0.065711
Baseball Field	0.043819
Baseball Stadium	0.055618
Basketball Court	0.045515
Bath House	0.056480
Beach	0.044395
Beach Bar	0.056480
Beer Bar	0.099008
Beer Garden	0.043802
Beer Store	0.107021
Big Box Store	0.061494
Bike Rental / Bike Share	0.052206
Bike Shop	0.054374
Bike Trail	0.050856
Bistro	0.046192
Board Shop	0.047036
Boat or Ferry	0.047862
Bookstore	0.053060
Boutique	0.073023
Bowling Alley	0.047526
Boxing Gym	0.067631

Brazilian Restaurant	0.049161
Breakfast Spot	0.033190
Brewery	0.051298
Bridal Shop	0.056480
Bridge	0.051601
Bubble Tea Shop	0.052422
Buffet	0.049119
Building	0.049081
Burger Joint	0.043047
Burmese Restaurant	0.056480
Burrito Place	0.061436
Bus Line	0.068665
Bus Station	0.033666
Bus Stop	0.118077
Business Service	0.051803
Butcher	0.050290
Cafeteria	0.056480
Café	0.060822
Cajun / Creole Restaurant	0.047539
Camera Store	0.052549
Campground	0.056480
Candy Store	0.049049
Cantonese Restaurant	0.051308
Caribbean Restaurant	0.030035
Caucasian Restaurant	0.053950
Cha Chaan Teng	0.056480
Check Cashing Service	0.044794
Cheese Shop	0.044137
Chinese Restaurant	0.000000
Chocolate Shop	0.050221
Church	0.054104
Circus	0.056480
Climbing Gym	0.050639
Clothing Store	0.037403
Club House	0.051911
Cocktail Bar	0.085291
Coffee Shop	0.038003
College Academic Building	0.055730
College Arts Building	0.056480
College Basketball Court	0.056480
College Bookstore	0.056480
College Cafeteria	0.054860
College Gym	0.056480
College Theater	0.056480
Colombian Restaurant	0.056480
Comedy Club	0.053103
Comfort Food Restaurant	0.055288
Comic Shop	0.054953
Community Center	0.054882
Concert Hall	0.050774
Construction & Landscaping	0.054191
Convenience Store	0.031857
Cooking School	0.056480
Cosmetics Shop	0.079047
Costume Shop	0.056480
Coworking Space	0.296614
Creperie	0.067346
Cuban Restaurant	0.056281
Cultural Center	0.056480
Cupcake Shop	0.048859
Cycle Studio	0.037799
Czech Restaurant	0.056480
Dance Studio	0.054003
Daycare	0.055783
Deli / Bodega	0.053899
Department Store	0.036361
Dessert Shop	0.027936
Dim Sum Restaurant	0.048456
Diner	0.040358
Discount Store	0.026705
Distillery	0.055114
Dive Bar	0.049989

Dive Shop	0.056480
Doctor's Office	0.049605
Dog Run	0.052394
Donut Shop	0.059872
Dosa Place	0.056480
Drugstore	0.056480
Dry Cleaner	0.065684
Dumpling Restaurant	0.047411
Duty-free Shop	0.056480
Eastern European Restaurant	0.108052
Electronics Store	0.047251
Empanada Restaurant	0.054428
English Restaurant	0.056480
Entertainment Service	0.054612
Ethiopian Restaurant	0.048208
Event Service	0.056480
Event Space	0.043794
Exhibit	0.055623
Factory	0.056480
Falafel Restaurant	0.042158
Farm	0.053071
Farmers Market	0.038393
Fast Food Restaurant	0.017017
Field	0.056480
Filipino Restaurant	0.048727
Fish & Chips Shop	0.051697
Fish Market	0.045722
Fishing Store	0.056480
Flea Market	0.050224
Flower Shop	0.056952
Food	0.039449
Food & Drink Shop	0.184568
Food Court	0.057690
Food Stand	0.054714
Food Truck	0.102210
Fountain	0.054253
French Restaurant	0.086910
Fried Chicken Joint	0.025586
Frozen Yogurt Shop	0.047014
Fruit & Vegetable Store	0.075717
Furniture / Home Store	0.060019
Gaming Cafe	0.052944
Garden	0.047126
Garden Center	0.204202
Gas Station	0.038282
Gastropub	0.083190
Gay Bar	0.050449
General Entertainment	0.054284
German Restaurant	0.217457
Gift Shop	0.050063
Gluten-free Restaurant	0.056480
Go Kart Track	0.056480
Golf Course	0.054070
Gourmet Shop	0.033743
Greek Restaurant	0.091162
Grocery Store	0.116557
Gym	0.040374
Gym / Fitness Center	0.020524
Gym Pool	0.056480
Gymnastics Gym	0.098818
Halal Restaurant	0.051942
Harbor / Marina	0.048504
Hardware Store	0.047948
Hawaiian Restaurant	0.088363
Health & Beauty Service	0.050145
Health Food Store	0.045674
Heliport	0.055156
Herbs & Spices Store	0.056480
High School	0.056480
Hill	0.056480
Himalayan Restaurant	0.051957
Historic Site	0.053027

History Museum	0.047839
Hobby Shop	0.050875
Home Service	0.047386
Hookah Bar	0.042098
Hostel	0.053738
Hot Dog Joint	0.051912
Hotel	0.110032
Hotel Bar	0.045109
Hotel Pool	0.056480
Hotpot Restaurant	0.051150
Ice Cream Shop	0.112883
Indian Restaurant	0.040959
Indie Movie Theater	0.048006
Indie Theater	0.049280
Indonesian Restaurant	0.056480
Indoor Play Area	0.054612
Insurance Office	0.056480
Intersection	0.052872
Irish Pub	0.053818
Israeli Restaurant	0.050865
Italian Restaurant	0.098000
Japanese Curry Restaurant	0.052637
Japanese Restaurant	0.033375
Jazz Club	0.047194
Jewelry Store	0.045168
Jewish Restaurant	0.053906
Juice Bar	0.046213
Karaoke Bar	0.044356
Kebab Restaurant	0.053995
Kids Store	0.052848
Korean BBQ Restaurant	0.056480
Korean Restaurant	0.075160
Kosher Restaurant	0.051388
Lake	0.054727
Latin American Restaurant	0.026728
Laundromat	0.048999
Laundry Service	0.053898
Lawyer	0.050721
Leather Goods Store	0.053898
Lebanese Restaurant	0.053898
Library	0.052622
Lingerie Store	0.043502
Liquor Store	0.023997
Locksmith	0.053735
Lounge	0.044104
Malay Restaurant	0.053979
Marine Terminal	0.056480
Market	0.054054
Martial Arts School	0.051048
Massage Studio	0.052283
Mattress Store	0.042271
Medical Center	0.056480
Mediterranean Restaurant	0.269534
Memorial Site	0.056480
Men's Store	0.042132
Metro Station	0.035852
Mexican Restaurant	0.112697
Middle Eastern Restaurant	0.258932
Mini Golf	0.056480
Miscellaneous Shop	0.044236
Mobile Phone Shop	0.073112
Modern European Restaurant	0.053910
Molecular Gastronomy Restaurant	0.056480
Monument / Landmark	0.050748
Moroccan Restaurant	0.053898
Motorcycle Shop	0.053451
Movie Theater	0.063045
Moving Target	0.043788
Multiplex	0.056480
Museum	0.046910
Music School	0.053995
Music Store	0.050486

Music Venue	0.048009
Nail Salon	0.061394
Nature Preserve	0.056480
New American Restaurant	0.351403
Newsstand	0.051907
Nightclub	0.060656
Nightlife Spot	0.056480
Non-Profit	0.054554
Noodle House	0.510459
North Indian Restaurant	0.053983
Office	0.050361
Opera House	0.054763
Optical Shop	0.042410
Organic Grocery	0.097051
Other Great Outdoors	0.049733
Other Nightlife	0.056121
Other Repair Shop	0.052254
Outdoor Gym	0.176547
Outdoor Sculpture	0.051410
Outdoors & Recreation	0.387289
Outlet Mall	0.056480
Outlet Store	0.054921
Paella Restaurant	0.053902
Pakistani Restaurant	0.176547
Paper / Office Supplies Store	0.045954
Park	0.054941
Pastry Shop	0.056480
Pedestrian Plaza	0.055305
Performing Arts Venue	0.055940
Perfume Shop	0.056480
Persian Restaurant	0.056480
Peruvian Restaurant	0.050422
Peruvian Roast Chicken Joint	0.056480
Pet Café	0.052396
Pet Service	0.053341
Pet Store	0.067553
Pharmacy	0.069859
Photography Studio	0.055769
Physical Therapist	0.056480
Piano Bar	0.053898
Pie Shop	0.050759
Pier	0.055882
Piercing Parlor	0.053898
Pilates Studio	0.048761
Pizza Place	0.090120
Platform	0.056480
Playground	0.084121
Plaza	0.133504
Poke Place	0.049401
Pool	0.053548
Pool Hall	0.052078
Portuguese Restaurant	0.056480
Post Office	0.052724
Pub	0.037223
Public Art	0.054938
Puerto Rican Restaurant	0.053914
Racetrack	0.055126
Ramen Restaurant	0.035541
Record Shop	0.132620
Recording Studio	0.056216
Recreation Center	0.055814
Rental Car Location	0.053747
Rental Service	0.055476
Residential Building (Apartment / Condo)	0.085908
Resort	0.054511
Rest Area	0.056262
Restaurant	0.055984
River	0.055849
Rock Climbing Spot	0.056480
Rock Club	0.050199
Roller Rink	0.056480
Romanian Restaurant	0.056480

Roof Deck	0.046800
Russian Restaurant	0.050841
Sake Bar	0.053684
Salad Place	0.037365
Salon / Barbershop	0.047997
Sandwich Place	0.119716
Scandinavian Restaurant	0.056480
Scenic Lookout	0.038338
School	0.052592
Sculpture Garden	0.056113
Seafood Restaurant	0.027198
Shabu-Shabu Restaurant	0.056480
Shanghai Restaurant	0.073351
Shipping Store	0.039908
Shoe Repair	0.056480
Shoe Store	0.035407
Shop & Service	0.053348
Shopping Mall	0.051274
Skate Park	0.052973
Skating Rink	0.050696
Ski Area	0.056480
Smoke Shop	0.040559
Smoothie Shop	0.056480
Snack Place	0.044218
Soccer Field	0.053586
Social Club	0.055080
Soup Place	0.056119
South American Restaurant	0.049052
South Indian Restaurant	0.056480
Southern / Soul Food Restaurant	0.048362
Souvlaki Shop	0.056480
Spa	0.041255
Spanish Restaurant	0.024628
Speakeasy	0.045468
Sporting Goods Shop	0.052390
Sports Bar	0.045490
Sports Club	0.053868
Sri Lankan Restaurant	0.056480
Stadium	0.056480
State / Provincial Park	0.056480
Stationery Store	0.053992
Steakhouse	0.040701
Storage Facility	0.049774
Street Art	0.053900
Strip Club	0.053898
Supermarket	0.065231
Supplement Shop	0.104899
Surf Spot	0.055815
Sushi Restaurant	0.120070
Swiss Restaurant	0.056480
Szechuan Restaurant	0.118512
Taco Place	0.246654
Tailor Shop	0.050498
Taiwanese Restaurant	0.049816
Tanning Salon	0.056480
Tapas Restaurant	0.039556
Tattoo Parlor	0.047029
Tea Room	0.109784
Tech Startup	0.056480
Tennis Court	0.051918
Tennis Stadium	0.053953
Tex-Mex Restaurant	0.056480
Thai Restaurant	0.069489
Theater	0.049438
Theme Park	0.054352
Theme Park Ride / Attraction	0.053308
Thrift / Vintage Store	0.043568
Tibetan Restaurant	0.054051
Tiki Bar	0.051244
Toll Plaza	0.056480
Tourist Information Center	0.054842
Toy / Game Store	0.061374

```
Track                                          0.052211
Trail                                          0.049632
Train                                          0.056480
Train Station                                  0.038582
Tree                                           0.056480
Turkish Restaurant                             0.051264
Udon Restaurant                                0.053898
Used Bookstore                                 0.052041
Vape Store                                     0.189293
Varenyky restaurant                            0.056480
Vegetarian / Vegan Restaurant                  0.104575
Venezuelan Restaurant                          0.056480
Veterinarian                                   0.054091
Video Game Store                               0.039455
Video Store                                    0.046009
Vietnamese Restaurant                          0.060633
Volleyball Court                               0.056480
Warehouse Store                                0.054928
Waste Facility                                 0.056480
Waterfront                                     0.055032
Weight Loss Center                             0.055341
Whisky Bar                                     0.108735
Wine Bar                                       0.053097
Wine Shop                                      0.037333
Wings Joint                                    0.043418
Women's Store                                  0.035181
Yemeni Restaurant                              0.056480
Yoga Studio                                    0.067771
```

```
<ipython-input-207-95fb74ce3c21>:13: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_g
uide/indexing.html#returning-a-view-versus-a-copy
  lebCorr["Weight"] = lebArr
<ipython-input-207-95fb74ce3c21>:23: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_g
uide/indexing.html#returning-a-view-versus-a-copy
  polCorr["Weight"] = polArr
```

Lets explore the top 10 venues for each type of restaurants.

In [208]:

```
topleb = lebCorr.nlargest(10,"Weight")


toppol = polCorr.nlargest(10,"Weight")


print(topleb)
print(toppol)
```

```
                       Weight
Persian Restaurant   0.720229
Shoe Repair          0.720229
English Restaurant   0.720229
Leather Goods Store  0.521138
Strip Club           0.521138
Udon Restaurant      0.521138
Beer Bar             0.358172
Adult Boutique       0.322458
Gaming Cafe          0.277542
Bike Trail           0.270431
                             Weight
Noodle House               0.510459
Outdoors & Recreation      0.387289
New American Restaurant    0.351403
```

```
Coworking Space              0.296614
Mediterranean Restaurant     0.269534
Middle Eastern Restaurant    0.258932
Taco Place                   0.246654
German Restaurant            0.217457
Garden Center                0.204202
Vape Store                   0.189293
```

From this we can confirm that Lebanese Resturants are usually in neighborhoods that has Persian and English Restaurants as well as Shoe repair and Leather Goods shops. While Polish Restaurants are in neighborhoods with Noodle Houses, Outdoors and Recreation, New American restaurants and Coworking spaces.

Now we will correlate the retaurants profiles with venues means in each neighborhood to see how similar the neighboorhood venues are to the venues that open near Lebanese and Polish restaurants.

In [209]:

```
lebHoods = meanHoods.drop(["Lebanese Restaurant"],1)#remove Lebanese Restaurant since the
value will be 1
lebHoods = lebHoods.T # transpose the dataframe to match the shape of the restaurant prof
ile matrix
lebHoods.columns = lebHoods.iloc[0]
lebHoods = lebHoods.drop("Neighborhood", axis = 0)# make Neighborhoods as teh column head
ers

lebHoods = pd.DataFrame(lebHoods)# convert to pandas DataFrame

lebHoodsCorr = pd.DataFrame(index =lebHoods.columns, columns = ["Value"] ) # create a da
taframe for the neighborhoods correlations with Lebanese restaurants profiles

for i in range(len(lebHoods.columns)):
    lebHoodsCorr.loc[lebHoods.columns[i]] = lebHoods[lebHoods.columns[i]].astype('float
64').corr(lebCorr["Weight"].astype('float64')) #correlate the neighborhoods with the prof
iles


polHoods = meanHoods.drop(["Polish Restaurant"],1)#remove Polish Restaurant since the va
lue will be 1
polHoods = polHoods.T # transpose the dataframe to match the shape of the restaurant prof
ile matrix
polHoods.columns = polHoods.iloc[0]
polHoods = polHoods.drop("Neighborhood", axis = 0)# make Neighborhoods as teh column head
ers

polHoods = pd.DataFrame(polHoods)# convert to pandas DataFrame

polHoodsCorr = pd.DataFrame(index =polHoods.columns, columns = ["Value"] ) # create a da
taframe for the neighborhoods correlations with Polish restaurants profiles

for i in range(len(polHoods.columns)):
    polHoodsCorr.loc[polHoods.columns[i]] = polHoods[polHoods.columns[i]].astype('float
64').corr(polCorr["Weight"].astype('float64')) #correlate the neighborhoods with the prof
iles
```

let's explore the top 10 neighborhoods for Lebanese restaurants

In [210]:

```
lebHoodsCorr.sort_values("Value", ascending=False).head(10)
```

Out[210]:

	Value
Neighborhood	
Sutton Place	0.293417
Greenwich Village	0.228352

Neighborhood	Value
Soho	0.103279
Noho	0.082809
West Village	0.062787
Upper West Side	0.050469
Astoria	0.050225
Civic Center	0.048836
Egbertville	0.044713
Carroll Gardens	0.042252

let's explore the top 10 neighborhoods for Polish restaurants

In [211]:

```
polHoodsCorr.sort_values("Value", ascending=False).head(10)
```

Out[211]:

	Value
Neighborhood	
Forest Hills Gardens	0.388636
Arrochar	0.293602
Ridgewood	0.221355
Astoria	0.164868
West Village	0.150217
Greenpoint	0.137565
East Harlem	0.132904
Kensington	0.132519
Noho	0.126034
Little Italy	0.12485

let's normalize the two correlation dataframes

In [213]:

```
lebArr = lebHoodsCorr.values #returns a numpy array
min_max_scaler = preprocessing.MinMaxScaler()
lebArr = min_max_scaler.fit_transform(lebArr)
lebHoodsCorr["Value"] = lebArr


polArr = polHoodsCorr.values #returns a numpy array
polArr = min_max_scaler.fit_transform(polArr)
polHoodsCorr["Value"] = polArr
```

At this point, we have identified the neighborhoods that are correlated with the restaurants. Now we need to identify the availability of Lebanese and Polish Restaurants in the neighborhoods.

we will do this by creating two normalized availability dataframes for the count of lebanese and polish restaurants. Higher availability values for a neighborhood indicates higher number of lebanse or polish restaurants.

In [214]:

```
lebCount = countHoods[["Neighborhood","Lebanese Restaurant"]].copy()
lebCount = lebCount.set_index("Neighborhood")
lebArr = lebCount.values #returns a numpy array
min_max_scaler = preprocessing.MinMaxScaler()
```

```
lebArr = min_max_scaler.fit_transform(lebArr)
lebCount["Avail"] = lebArr
lebCount = lebCount.drop(["Lebanese Restaurant"],axis = 1)


polCount = countHoods[["Neighborhood","Polish Restaurant"]].copy()
polCount = polCount.set_index("Neighborhood")
polArr = polCount.values #returns a numpy array
min_max_scaler = preprocessing.MinMaxScaler()
polArr = min_max_scaler.fit_transform(polArr)
polCount["Avail"] = polArr
polCount = polCount.drop(["Polish Restaurant"],axis = 1)
```

Finally, we will calculate a score for each neighborhood based on their venues profile and availability of the restaurants types. The score is calculated by multiplyting the value of correlation by (1-availability). Neighborhoods with higher number of restaurant type will score lower.

In [235]:

```
lebHoodsRank = pd.DataFrame(columns = ["Score"], index = lebCount.index) # create a data frame for the neighborhoods Scores for Lebanese Restaurants
for i in range(len(lebCount.index)):
    lebHoodsRank.loc[lebCount.index[i]] = lebHoodsCorr.loc[lebCount.index[i]]["Value"] * (1-lebCount.loc[lebCount.index[i]]["Avail"])



polHoodsRank = pd.DataFrame(columns = ["Score"], index = polCount.index) # create a data frame for the neighborhoods Scores for Polish Restaurants
for i in range(len(polCount.index)):
    polHoodsRank.loc[polCount.index[i]] = polHoodsCorr.loc[polCount.index[i]]["Value"] * (1-polCount.loc[polCount.index[i]]["Avail"])
```

lets explore the top 10 neighborhoods for Lebanese restaurants and compare it with the correlations rank to see the effect of availability.

In [242]:

```
lebHoodsExp = pd.DataFrame(columns = ["Score","Corr","Avail"], index = lebCount.index)
lebHoodsExp["Score"] = lebHoodsRank["Score"]
lebHoodsExp["Corr"] = lebHoodsCorr["Value"]
lebHoodsExp["Avail"] = lebCount["Avail"]

lebHoodsExp.sort_values("Avail", ascending = False).head()
```

Out[242]:

	Score	Corr	Avail
Neighborhood			
Greenwich Village	0.0	0.829073	1.0
Sutton Place	0.0	1.000000	1.0
Allerton	0.035489	0.035489	0.0
Ozone Park	0.040313	0.040313	0.0
Olinville	0.10001	0.100010	0.0

We can see that accounting for the availability has exclueded Sutton Place and Greenwich Village even though they are highly correlated with Lebanese Restaurants profile.

In [244]:

```
polHoodsExp = pd.DataFrame(columns = ["Score","Corr","Avail"], index = polCount.index)
polHoodsExp["Score"] = polHoodsRank["Score"]
polHoodsExp["Corr"] = polHoodsCorr["Value"]
polHoodsExp["Avail"] = polCount["Avail"]
```

```
polHoodsExp.sort_values("Avail", ascending = False).head(10)
```

Out[244]:

Neighborhood	Score	Corr	Avail
Greenpoint	0.0	0.479657	1.0
Arrochar	0.401521	0.803042	0.5
Downtown	0.178483	0.356967	0.5
Ridgewood	0.326656	0.653312	0.5
Forest Hills Gardens	0.5	1.000000	0.5
Allerton	0.074009	0.074009	0.0
Ocean Hill	0.289871	0.289871	0.0
Paerdegat Basin	0.112729	0.112729	0.0
Ozone Park	0.100603	0.100603	0.0
Olinville	0.092924	0.092924	0.0

For the Polish restaurant, Greenpoint has been excluded while Arrochar, Downtown, Ridgewood and Forest Hills Gardens scores have been lowered since they aleady have Polish restaurants.

Lets explore the top 10 neighborhoods scores for Lebanese restaurants.

In [251]:

```
lebHoodsExp.sort_values("Score",ascending = False).head(10)
```

Out[251]:

Neighborhood	Score	Corr	Avail
Soho	0.500507	0.500507	0.0
Noho	0.446732	0.446732	0.0
West Village	0.394136	0.394136	0.0
Upper West Side	0.361777	0.361777	0.0
Astoria	0.361135	0.361135	0.0
Civic Center	0.357485	0.357485	0.0
Egbertville	0.346655	0.346655	0.0
Carroll Gardens	0.340191	0.340191	0.0
Flatiron	0.338216	0.338216	0.0
Midtown	0.338191	0.338191	0.0

We can see that Soho, Noho and West Village are teh best location to open a new Lebanese Restaurant because they are highly correlated with the profile of a Lebanese restaurant and currently does not have any lebanese restaurants.

Now, lets explore the top 10 neighborhoods scores for Lebanese restaurants.

In [250]:

```
polHoodsExp.sort_values("Score",ascending = False).head(10)
```

Out[250]:

	Score	Corr	Avail
Neighborhood			
Astoria	0.536243	0.536243	0.0
West Village	0.505879	0.505879	0.0
Forest Hills Gardens	0.5	1.000000	0.5
East Harlem	0.469997	0.469997	0.0
Kensington	0.469199	0.469199	0.0
Noho	0.455759	0.455759	0.0
Little Italy	0.453306	0.453306	0.0
Auburndale	0.448815	0.448815	0.0
Bushwick	0.423651	0.423651	0.0
Cobble Hill	0.422321	0.422321	0.0

This tell us that the best neighborhoods for a new Polish restaurant are Astoria, West Village, and Forest Hills Gardens because they are highly correlated with the Polish restaurant profile and with the exception of Forest Hills Gardens have no Polish restaurants.

Now Lets create a map to visually represent the neighborhoods score. First of all, we will create a dataframe with the scores and the Neighborhoods longitudes and latitudes.

In [311]:

```
mapHoods = pd.DataFrame(columns = ["Neighborhood","Latitude","Longitude","Leb Score","Le
b Color","Pol Score", "Pol Color"])

mapHoods[["Neighborhood","Latitude","Longitude"]] = hoods[["Neighborhood","Latitude","Lon
gitude"]]
mapHoods = mapHoods.set_index("Neighborhood")
mapHoods["Leb Score"] = lebHoodsExp["Score"]
mapHoods["Pol Score"] = polHoodsExp["Score"]
mapHoods = mapHoods.reset_index()

mapHoods["Leb Color"] = pd.cut(mapHoods["Leb Score"], bins=4, labels=['red',  'orange','
yellow' ,'green'])
mapHoods["Pol Color"] = pd.cut(mapHoods["Pol Score"], bins=4, labels=['red',  'orange','
yellow' ,'green'])

mapHoods.head()
```

Out[311]:

	Neighborhood	Latitude	Longitude	Leb Score	Leb Color	Pol Score	Pol Color
0	Wakefield	40.894705	-73.847201	0.108154	red	0.218721	orange
1	Co-op City	40.874294	-73.829939	0.073231	red	0.187542	orange
2	Eastchester	40.887556	-73.827806	0.056022	red	0.099989	red
3	Fieldston	40.895437	-73.905643	0.141728	orange	0.220873	orange
4	Riverdale	40.890834	-73.912585	0.121452	red	0.19621	orange

Then we will create the map for Lebanese Restaurants

In [312]:

```
map_newyork = folium.Map(location=[40.710318150000006, -74.00496837048613], zoom_start=1
0)

for lat, lng, neighborhood, score, clr in zip(mapHoods['Latitude'], mapHoods['Longitude']
```

```
, mapHoods['Neighborhood'],mapHoods['Leb Score'],mapHoods['Leb Color'] ):
    label = '{},Score: {}'.format(neighborhood, score)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=3,
        popup=label,
        color= clr,
        fill=True,
        fill_color=clr,
        fill_opacity=0.7,
        parse_html=False).add_to(map_newyork)

map_newyork
```

Out[312]:

Make this Notebook Trusted to load map: File -> Trust Notebook

We will create a map for Polish Restaurants.

In [313]:

```
map_newyork = folium.Map(location=[40.710318150000006, -74.00496837048613], zoom_start=1
0)

for lat, lng, neighborhood, score, clr in zip(mapHoods['Latitude'], mapHoods['Longitude']
, mapHoods['Neighborhood'],mapHoods['Pol Score'],mapHoods['Pol Color'] ):
    label = '{},Score: {}'.format(neighborhood, score)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=3,
        popup=label,
        color= clr,
        fill=True,
        fill_color=clr,
        fill_opacity=0.7,
        parse_html=False).add_to(map_newyork)

map_newyork
```

Out[313]:

Make this Notebook Trusted to load map: File -> Trust Notebook

Results and Discussion

Lebanese Restaurant

By examining the neighborhoods rankings we can see that Soho, Noho and West Village are the best neighbohoods to open a new Lebanese Restaurant because they are highly correlated with the profile of a Lebanese restaurant and currently does not have any lebanese restaurants. In addition, by examining the map, we can see that Manhattan is the best Borough to open a Lebanese Restaurant.

Polish Restaurant

Neighborhoods rankings tell us that the best neighborhoods for a new Polish restaurant are Astoria, West Village, and Forest Hills Gardens because they are highly correlated with the Polish restaurant profile and with the exception of Forest Hills Gardens have no Polish restaurants. By examining the map, we can see that for Polish Restaurants the best borough is also Manhattan with few neighborhoods in other boroughs such as Bushwick in Brooklyn, Bayside in Queens and Fox Hills in Staten Island.

Conclusion

The objective of this project is to develop a tool that will help a Lebanese and a Polish immigrants immigrants identify the best neighborhood for opening a new ethnic resturant in New York City. The project utilized New York neighborhood information used in Week 3 as well as Foursquare API to collect neighborhoods venues information. Profiles for Lebanese and Polish restuarnts were creatare Astoria, West Village, and Forest Hills Gardensed to correlate the neighborhoods venues with the profiles. Then a score for each neighborhoods was calculated based on its correlation with the restaurant profile and availability of the type of resturants. finally, maps were created to visualize the neighborhoods and their scores to get a birds eye view of the results.

The project concluded that Manhattan is the best Borough for both types of restaurants. For Lebanese Restaurants, the best neighborhoods are Soho, Noho and West Village (neighborhoods in west Manhattan). For Polish Restaurants,the best neighborhoods are Astoria, West Village, and Forest Hills Gardens.