Battle of Neighbourhoods:IBM Capstone Project

Introduction

As part of my capstone project, I decided to compare two neighbourhoods in New York in depth to find out how they compare to each other using location data from Foursquare API. The aim of this project is to find the ten most popular venues in each of the neighbourhoods then clustering them for future interested parties for examples data analysts, location data managers, tourists etc.

New York City often called the big Apple, comprises of 5 boroughs sitting where the Hudson River meets the Atlantic Ocean. It is the largest city in the United States with a long history of international immigration having a population of 8.149 million as of 2019. Its iconic sites include skyscrapers such as the Empire State building and sprawling Central Park. New York City's demographics show that it is a large and ethnically diverse metropolis, thus New York city is known worldwide as a cultural melting pot. While other states have had immigration surges, none have compared to the diversity and sheer number of immigrants that have made their way to the City. The varied cultures combined to create a great diversity for itself, the city is the world's Financial epicentre home to NYSE and NASDAQ, Communications, Real Estate, Insurance, Technology, Entertainment, Healthcare etc.

Manhattan: Originally the only borough in the city has the smallest land area with a population of 1.6mil diverse (example neighbourhoods of Soho and Harlem), rich culture, leading centre for performing arts and iconic landmarks: Central park, Skyscrapers, Empire State building, historic Cathedrals, Wallstreet, Grand Central station etc.

Travelling using the subway, walking or taxis etc it is divided into three grids: Uptown, Downtown and Midtown, roads are Avenues (north to south) and Streets (east to west).

Bronx: 4th largest, the northernmost and only borough on American mainland, population of 1.3mil largely residential, has vibrant neighbourhoods some areas became the symbol of urban decay that happened in the 1960s to 80s. Home of Hip Hop music, the Yankee Stadium, New York Botanical Gardens.

Data
To accomplish the comparison between these two boroughs of Manhattan and Bronx in New York, Foursquare API will be used to gather data on the ten most popular venues in each of the neighbourhoods mentioned above.Data scrapping and cleaning by using python packages.

Neighbourhood and Borough data for New York was provided from the lab in week 3 of the course https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-Skillabs/newyork_data.json

Gathering the Data

New York data on borough, neighbourhood is extracted using python packages to open the json file and create a python dataframe then adding the data to it.

```
!wget -q -O 'newyork_data.json' https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json
print('Data downloaded!')

   Data downloaded!
```

```
with open('newyork_data.json') as json_data:
    newyork_data = json.load(json_data)
```

```
newyork_data
```

```
7]: {'type': 'FeatureCollection',
    'totalFeatures': 306,
    'features': [{'type': 'Feature',
     'id': 'nyu_2451_34572.1',
     'geometry': {'type': 'Point',
      'coordinates': [-73.84720052054902, 40.89470517661]},
     'geometry_name': 'geom',
     'properties': {'name': 'Wakefield',
      'stacked': 1,
```

```
ny_data = newyork_data['features']
```

```
column_names = ['Borough', 'Neighborhood', 'Latitude', 'Longitude']
ny_df = pd.DataFrame(columns=column_names)
ny_df
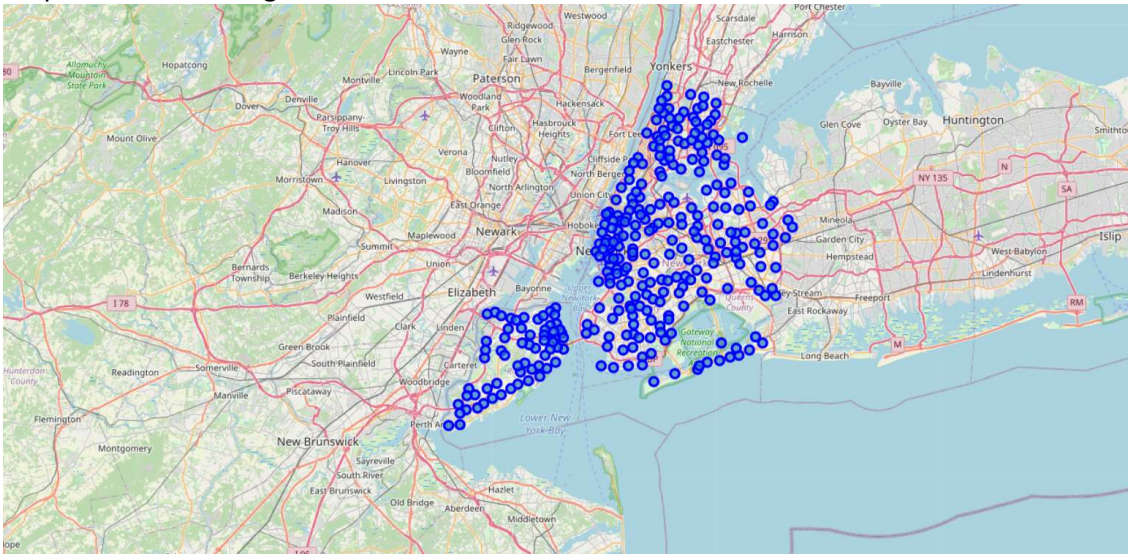```

```
9]:
        Borough  Neighborhood  Latitude  Longitude
```

```
for data in ny_data:
    borough = neighborhood_name = data['properties']['borough']
    neighborhood_name = data['properties']['name']

    neighborhood_latlon = data['geometry']['coordinates']
    neighborhood_lat = neighborhood_latlon[1]
    neighborhood_lon = neighborhood_latlon[0]

    ny_df = ny_df.append({'Borough': borough,
                          'Neighborhood': neighborhood_name,
                          'Latitude': neighborhood_lat,
                          'Longitude': neighborhood_lon}, ignore_index=True)
```

Map of New York Neighbourhoods

As the Project is comparing two boroughs these are extracted from the New York data and then can be used to gathering venues data from the API and sorting the data.

```
bronx_data = ny_df[ny_df['Borough'] == 'Bronx'].reset_index(drop=True)
bronx_data.head()
```

2]:

	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

```
manhattan_data = ny_df[ny_df['Borough'] == 'Manhattan'].reset_index(drop=True)
manhattan_data.head()
```

1]:

	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Foursquare API explore call to gather venues data and sort the data into a dataframe. The API call returned results across the boroughs; the Bronx results were a third of the Manhattan responses although having a larger land area indicates it isn't as densely populated as Manhattan.

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

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

        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)

        results = requests.get(url).json()["response"]['groups'][0]['items']

        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_list])
    nearby_venues.columns = ['Neighborhood',
                  'Neighborhood Latitude',
                  'Neighborhood Longitude',
                  'Venue',
                  'Venue Latitude',
                  'Venue Longitude',
                  'Venue Category']

    return(nearby_venues)
```

```
bronx_venues = getNearbyVenues(names=bronx_data['Neighborhood'],
                               latitudes=bronx_data['Latitude'],
                               longitudes=bronx_data['Longitude']
                               )
```

```
print(bronx_venues.shape)
bronx_venues.head()
```

(1196, 7)

	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

```
manhattan_venues = getNearbyVenues(names=manhattan_data['Neighborhood'],
                                   latitudes=manhattan_data['Latitude'],
                                   longitudes=manhattan_data['Longitude']
                                   )
```

```
print(manhattan_venues.shape)
manhattan_venues.head()
```

(3243, 7)

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Astral Fitness & Wellness Center	40.876705	-73.906372	Gym
4	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop

Bronx

Sorting the data results and calculating the mean of the API call responses to get the 10 most popular venues in the Bronx neighbourhoods.

```
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = bronx_grouped['Neighborhood']

for ind in np.arange(bronx_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(bronx_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Allerton	Deli / Bodega	Pizza Place	Supermarket	Discount Store	Spa	Donut Shop	Food	Gas Station	Martial Arts School	Check Cashing Service
1	Baychester	Discount Store	Donut Shop	Bank	Supermarket	Pet Store	Men's Store	Fast Food Restaurant	Fried Chicken Joint	Sandwich Place	Mattress Store
2	Bedford Park	Diner	Mexican Restaurant	Deli / Bodega	Pizza Place	Chinese Restaurant	Spanish Restaurant	Sandwich Place	Fried Chicken Joint	Donut Shop	Pub
3	Belmont	Italian Restaurant	Deli / Bodega	Pizza Place	Bakery	Mexican Restaurant	Dessert Shop	Grocery Store	Bank	Diner	Shoe Store
4	Bronxdale	Chinese Restaurant	Deli / Bodega	Bank	Pizza Place	Eastern European Restaurant	Performing Arts Venue	Park	Spanish Restaurant	Mexican Restaurant	Supermarket

Then clustering Bronx neighbourhoods by most popular venues.

```
9]: kclusters = 5

    bronx_grouped_clustering = bronx_grouped.drop('Neighborhood', 1)

    kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(bronx_grouped_clustering)

    kmeans.labels_[0:10]

t[99]: array([2, 2, 2, 2, 0, 2, 2, 2, 0, 2], dtype=int32)

1]: neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

    bronx_merged = bronx_data
    bronx_merged = bronx_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

    bronx_merged.head()
```

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bronx	Wakefield	40.894705	-73.847201	2	Pharmacy	Deli / Bodega	Ice Cream Shop	Dessert Shop	Laundromat	Sandwich Place	Donut Shop	Peruvian Restaurant	Nightclub	Optical Shop
1	Bronx	Co-op City	40.874294	-73.829939	2	Accessories Store	Baseball Field	Deli / Bodega	Restaurant	Donut Shop	Pizza Place	Fast Food Restaurant	Chinese Restaurant	Pharmacy	Park
2	Bronx	Eastchester	40.887556	-73.827806	2	Deli / Bodega	Bus Station	Caribbean Restaurant	Diner	Chinese Restaurant	Bowling Alley	Food & Drink Shop	Metro Station	Seafood Restaurant	Fast Food Restaurant
3	Bronx	Fieldston	40.895437	-73.905643	1	Plaza	Music Venue	River	Bus Station	Accessories Store	Park	Music Store	Nightclub	Optical Shop	Outlet Store
4	Bronx	Riverdale	40.890834	-73.912585	1	Park	Bus Station	Home Service	Medical Supply Store	Gym	Bank	Food Truck	Plaza	Playground	Pharmacy

Manhattan

The same process is done with Manhattan neighbourhood data to get the 10 most popular venues.

```
num_top_venues = 10

indicators = ['st', 'nd', 'rd']

# create columns according to number of top venues
columns = ['Neighborhood']
for ind in np.arange(num_top_venues):
    try:
        columns.append('{}{} Most Common Venue'.format(ind+1, indicators[ind]))
    except:
        columns.append('{}th Most Common Venue'.format(ind+1))

# create a new dataframe
neighborhoods_venues_sorted = pd.DataFrame(columns=columns)
neighborhoods_venues_sorted['Neighborhood'] = manhattan_grouped['Neighborhood']

for ind in np.arange(manhattan_grouped.shape[0]):
    neighborhoods_venues_sorted.iloc[ind, 1:] = return_most_common_venues(manhattan_grouped.iloc[ind, :], num_top_venues)

neighborhoods_venues_sorted.head()
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Battery Park City	Park	Coffee Shop	Hotel	Gym	Clothing Store	Memorial Site	Boat or Ferry	Pizza Place	Shopping Mall	Burger Joint
1	Carnegie Hill	Coffee Shop	Café	Bar	Yoga Studio	Pizza Place	Wine Shop	Bookstore	Gym	Cosmetics Shop	French Restaurant
2	Central Harlem	African Restaurant	American Restaurant	Gym / Fitness Center	French Restaurant	Bar	Seafood Restaurant	Chinese Restaurant	Café	Cocktail Bar	Tapas Restaurant
3	Chelsea	Coffee Shop	Bakery	Art Gallery	French Restaurant	Italian Restaurant	American Restaurant	Seafood Restaurant	Ice Cream Shop	Wine Shop	Nightclub
4	Chinatown	Chinese Restaurant	Bakery	Cocktail Bar	American Restaurant	Hotpot Restaurant	Spa	Dessert Shop	Salon / Barbershop	Boutique	Sandwich Place

Then Manhattan neighbourhoods are clustered by the most popular venues.

```
kclusters = 5

manhattan_grouped_clustering = manhattan_grouped.drop('Neighborhood', 1)

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(manhattan_grouped_clustering)

kmeans.labels_[0:10]
```
```
]: array([2, 2, 2, 2, 2, 2, 3, 2, 2], dtype=int32)
```
```
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

manhattan_merged = manhattan_data

manhattan_merged = manhattan_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

manhattan_merged.head()
```
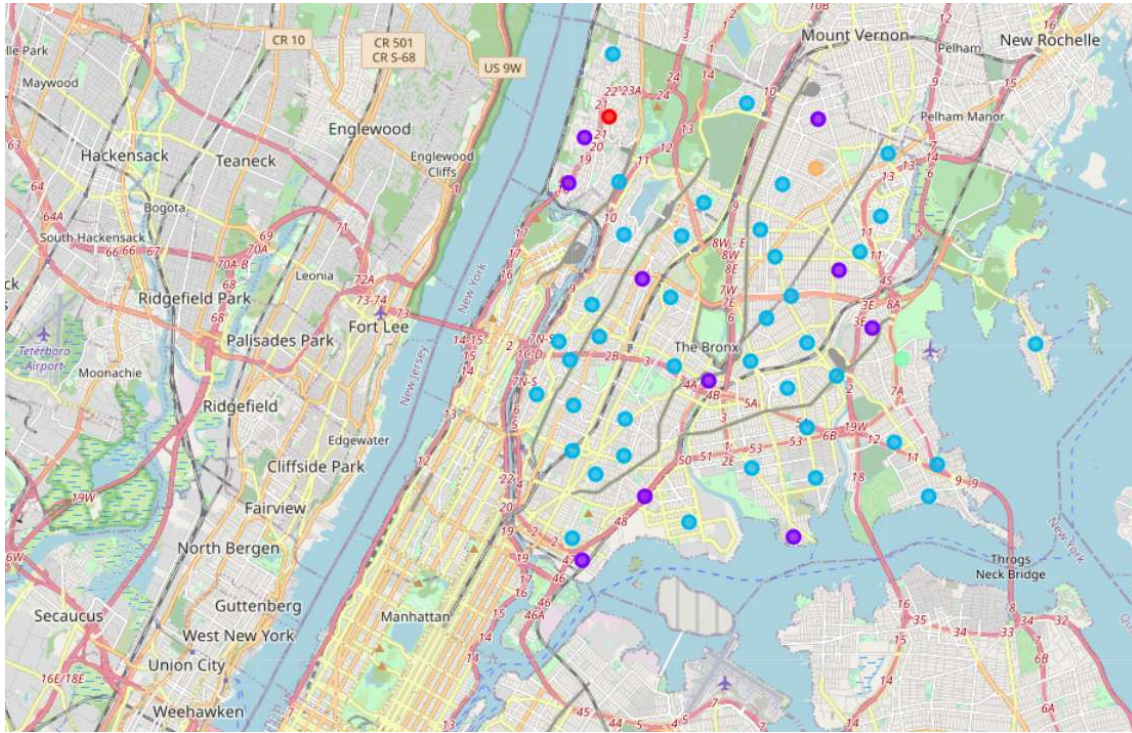
	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	2	American Restaurant	Gym	Sandwich Place	Yoga Studio	Diner	Steakhouse	Clothing Store	Coffee Shop	Seafood Restaurant	Deli / Bodega
1	Manhattan	Chinatown	40.715618	-73.994279	2	Chinese Restaurant	Bakery	Cocktail Bar	American Restaurant	Hotpot Restaurant	Spa	Dessert Shop	Salon / Barbershop	Boutique	Sandwich Place
2	Manhattan	Washington Heights	40.851903	-73.936900	3	Café	Chinese Restaurant	Bakery	Bank	Mobile Phone Shop	Latin American Restaurant	Grocery Store	Pizza Place	Deli / Bodega	Park
3	Manhattan	Inwood	40.867684	-73.921210	3	Café	Mexican Restaurant	Restaurant	Chinese Restaurant	Deli / Bodega	Lounge	Caribbean Restaurant	Park	Bakery	Wine Bar
4	Manhattan	Hamilton Heights	40.823604	-73.949688	3	Pizza Place	Café	Coffee Shop	Mexican Restaurant	Deli / Bodega	School	Park	Bakery	Indian Restaurant	Latin American Restaurant

Results

Map of Bronx neighbourhood clusters

Blue Cluster is the most dominant and spreads across the borough.it has different types of residential, tourist and sporting venues like banks, basketball court, marina, music venue, bowling alley, night clubs etc

```
#Blue Cluster
bronx_merged.loc[bronx_merged['Cluster Labels'] == 2, bronx_merged.columns[[1]] + list(range(5, bronx_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Co-op City	Pizza Place	Accessories Store	Basketball Court	Donut Shop	Restaurant	Fast Food Restaurant	Pharmacy	Park	Bus Station	Discount Store
2	Eastchester	Bus Station	Caribbean Restaurant	Diner	Deli / Bodega	Seafood Restaurant	Platform	Bus Stop	Bowling Alley	Pizza Place	Chinese Restaurant
5	Kingsbridge	Pizza Place	Sandwich Place	Bar	Mexican Restaurant	Latin American Restaurant	Bakery	Supermarket	Burger Joint	Spanish Restaurant	Pharmacy
6	Woodlawn	Pub	Deli / Bodega	Food & Drink Shop	Pizza Place	Bakery	Park	Grocery Store	Food Truck	Pharmacy	Bar
7	Norwood	Pizza Place	Park	Chinese Restaurant	Pharmacy	Bank	American Restaurant	Caribbean Restaurant	Bus Station	Spanish Restaurant	Fried Chicken Joint
8	Williamsbridge	Nightclub	Soup Place	Bar	Caribbean Restaurant	Accessories Store	Music Venue	Other Great Outdoors	Outlet Store	Paper / Office Supplies Store	Park
9	Baychester	Donut Shop	Mexican Restaurant	Electronics Store	Discount Store	Fast Food Restaurant	Sandwich Place	Pizza Place	Convenience Store	Pet Store	Fried Chicken Joint
10	Pelham Parkway	Italian Restaurant	Pizza Place	Sandwich Place	Food	Smoke Shop	Bus Station	Gourmet Shop	Donut Shop	Frozen Yogurt Shop	Chinese Restaurant
11	City Island	Seafood Restaurant	Thrift / Vintage Store	Harbor / Marina	Bar	Spanish Restaurant	Café	Music Venue	Park	History Museum	French Restaurant
12	Bedford Park	Diner	Pizza Place	Mexican Restaurant	Chinese Restaurant	Sandwich Place	Deli / Bodega	Baseball Field	Bus Station	Pub	Burger Joint
13	University Heights	Pizza Place	Sandwich Place	Bakery	Fried Chicken Joint	Grocery Store	Supermarket	Shoe Store	Pharmacy	History Museum	Bank
14	Morris Heights	Deli / Bodega	Recreation Center	IT Services	Pizza Place	Grocery Store	Pharmacy	Spanish Restaurant	Bank	Park	Nail Salon
16	East Tremont	Pizza Place	Puerto Rican Restaurant	Spanish Restaurant	Donut Shop	Breakfast Spot	Supermarket	Shoe Store	Mobile Phone Shop	Chinese Restaurant	Bank
18	High Bridge	Pharmacy	Pizza Place	Chinese Restaurant	Spanish Restaurant	Latin American Restaurant	Bus Station	Gym	Electronics Store	Deli / Bodega	Food
19	Melrose	Pizza Place	Discount Store	Sandwich Place	Grocery Store	Department Store	Mexican Restaurant	Diner	Deli / Bodega	Clothing Store	Paper / Office Supplies Store
20	Mott Haven	Pizza Place	Donut Shop	Gym	Grocery Store	Latin American Restaurant	Spanish Restaurant	Burger Joint	Storage Facility	Pharmacy	Fish & Chips Shop
23	Hunts Point	Bank	Farmers Market	Café	Gourmet Shop	Grocery Store	Home Service	BBQ Joint	Spanish Restaurant	Pizza Place	Juice Bar
24	Morrisania	Bus Station	Discount Store	Donut Shop	Fast Food Restaurant	Pizza Place	Metro Station	Grocery Store	Bowling Alley	Deli / Bodega	Sandwich Place
25	Soundview	Chinese Restaurant	Discount Store	Bus Station	Burger Joint	Fried Chicken Joint	Breakfast Spot	Grocery Store	Basketball Court	Pharmacy	Bus Stop
27	Throgs Neck	Deli / Bodega	Asian Restaurant	Coffee Shop	Baseball Field	Bar	Liquor Store	Pizza Place	American Restaurant	Italian Restaurant	Juice Bar
29	Parkchester	Supermarket	Pizza Place	Women's Store	Kids Store	Restaurant	Chinese Restaurant	Shipping Store	Sandwich Place	Deli / Bodega	Plaza
30	Westchester Square	Fast Food Restaurant	Mexican Restaurant	Donut Shop	Baseball Field	Pharmacy	Sandwich Place	Pizza Place	Supermarket	Mobile Phone Shop	Health & Beauty Service
31	Van Nest	Pizza Place	Bus Station	BBQ Joint	Supermarket	Mobile Phone Shop	Board Shop	Coffee Shop	Hookah Bar	Bus Stop	Deli / Bodega
32	Morris Park	Pizza Place	Burger Joint	Bakery	Deli / Bodega	Fast Food Restaurant	Sandwich Place	Pharmacy	Coffee Shop	Grocery Store	Buffet
33	Belmont	Italian Restaurant	Pizza Place	Deli / Bodega	Bakery	Sandwich Place	Dessert Shop	Grocery Store	Bank	Bar	Shoe Store
35	North Riverdale	Pizza Place	Bank	Chinese Restaurant	Italian Restaurant	Burger Joint	Bus Station	Coffee Shop	Donut Shop	Grocery Store	Gym / Fitness Center
37	Schuylerville	Diner	Pizza Place	Mexican Restaurant	Pharmacy	Bank	Italian Restaurant	Hookah Bar	American Restaurant	Japanese Restaurant	Donut Shop
38	Edgewater Park	Italian Restaurant	Pizza Place	Coffee Shop	Chinese Restaurant	Juice Bar	Bar	Deli / Bodega	Park	Spa	Sports Bar
39	Castle Hill	Bank	Diner	Park	Pizza Place	Pharmacy	Market	Accessories Store	Nightclub	Other Great Outdoors	Outlet Store
40	Olinville	Caribbean Restaurant	Deli / Bodega	Laundromat	Supermarket	Food	Chinese Restaurant	Bakery	Playground	Pool	Outlet Store
42	Concourse	Grocery Store	Spanish Restaurant	Fried Chicken Joint	Italian Restaurant	Bakery	Fast Food Restaurant	Caribbean Restaurant	Chinese Restaurant	Pharmacy	Food
43	Unionport	Latin American Restaurant	Donut Shop	Ice Cream Shop	Juice Bar	Mobile Phone Shop	Diner	Deli / Bodega	Dance Studio	Cuban Restaurant	Comfort Food Restaurant
45	Claremont Village	Grocery Store	Bus Station	Bakery	Pizza Place	Discount Store	Food	Pharmacy	Gym	Caribbean Restaurant	Restaurant
46	Concourse Village	Sandwich Place	Pizza Place	Fried Chicken Joint	Pharmacy	Supermarket	Deli / Bodega	Mexican Restaurant	Donut Shop	Bus Station	Park
47	Mount Eden	Pharmacy	Fast Food Restaurant	Deli / Bodega	Supermarket	Grocery Store	Chinese Restaurant	Pizza Place	Food Truck	Intersection	Bus Station
48	Mount Hope	Grocery Store	Deli / Bodega	Donut Shop	Pizza Place	Sandwich Place	Chinese Restaurant	Spanish Restaurant	Supermarket	Mobile Phone Shop	Ice Cream Shop
49	Bronxdale	Chinese Restaurant	Deli / Bodega	Pizza Place	Breakfast Spot	Supermarket	Mexican Restaurant	Gym	Performing Arts Venue	Bank	Spanish Restaurant
50	Allerton	Pizza Place	Deli / Bodega	Supermarket	Discount Store	Chinese Restaurant	Electronics Store	Spa	Bus Station	Gas Station	Pharmacy
51	Kingsbridge Heights	Pizza Place	Bus Station	Mexican Restaurant	Chinese Restaurant	Coffee Shop	Food Truck	Deli / Bodega	Sandwich Place	Café	Check Cashing Service

Purple Cluster is spread across the borough and the second most dominant cluster.it has mixed venues with notable places like recording studio, scenic lookout, intersection and storage facility which differentiate it from other clusters.

```
#Purple Cluster
bronx_merged.loc[bronx_merged['Cluster Labels'] == 1, bronx_merged.columns[[1]] + list(range(5, bronx_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Wakefield	Pharmacy	Deli / Bodega	Laundromat	Donut Shop	Sandwich Place	Dessert Shop	Ice Cream Shop	Lake	Pizza Place	Paper / Office Supplies Store
4	Riverdale	Bus Station	Park	Plaza	Playground	Home Service	Gym	Medical Supply Store	Baseball Field	Bank	Accessories Store
15	Fordham	Fast Food Restaurant	Shoe Store	Donut Shop	Mobile Phone Shop	Clothing Store	Bank	Pharmacy	Supplement Shop	Supermarket	Spanish Restaurant
17	West Farms	Park	Bus Stop	Donut Shop	Bus Station	Convenience Store	Chinese Restaurant	Pizza Place	Basketball Court	Coffee Shop	Playground
21	Port Morris	Brewery	Restaurant	Furniture / Home Store	Storage Facility	Grocery Store	Peruvian Restaurant	Spanish Restaurant	Donut Shop	Distillery	Latin American Restaurant
22	Longwood	Deli / Bodega	Diner	Latin American Restaurant	Donut Shop	Sandwich Place	Mexican Restaurant	Grocery Store	Fast Food Restaurant	Pet Store	Outlet Store
26	Clason Point	Park	Recording Studio	South American Restaurant	Pool	Playground	Convenience Store	Bus Stop	Boat or Ferry	Home Service	Grocery Store
34	Spuyten Duyvil	Park	Scenic Lookout	Grocery Store	Tennis Court	Tennis Stadium	Thai Restaurant	Bank	Intersection	Pharmacy	Pet Store
36	Pelham Bay	Bank	Fast Food Restaurant	Diner	Italian Restaurant	Gym / Fitness Center	Donut Shop	Convenience Store	Latin American Restaurant	Asian Restaurant	Supermarket
41	Pelham Gardens	Pharmacy	Bus Station	Spanish Restaurant	Boat or Ferry	Grocery Store	Bank	Sandwich Place	Playground	Mobile Phone Shop	Locksmith

Smaller Clusters (Red, Green, Orange)

They only have one neighbourhood each which could mean these neighbourhoods are different from the rest of the borough if not then with a different cluster size they could have been included in different cluster results.

Red Cluster: Neighbourhood of Fieldston

```
#Red cluster
bronx_merged.loc[bronx_merged['Cluster Labels'] == 0, bronx_merged.columns[[1] + list(range(5, bronx_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
3	Fieldston	Plaza	River	Bus Station	Medical Supply Store	Accessories Store	Peruvian Restaurant	Nightclub	Other Great Outdoors	Outlet Store	Paper / Office Supplies Store

Green Cluster: Neighbourhood of Country Club

```
#Green Cluster
bronx_merged.loc[bronx_merged['Cluster Labels'] == 3, bronx_merged.columns[[1] + list(range(5, bronx_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
28	Country Club	Sandwich Place	Playground	Athletics & Sports	Accessories Store	Peruvian Restaurant	Nail Salon	Nightclub	Other Great Outdoors	Outlet Store	Paper / Office Supplies Store

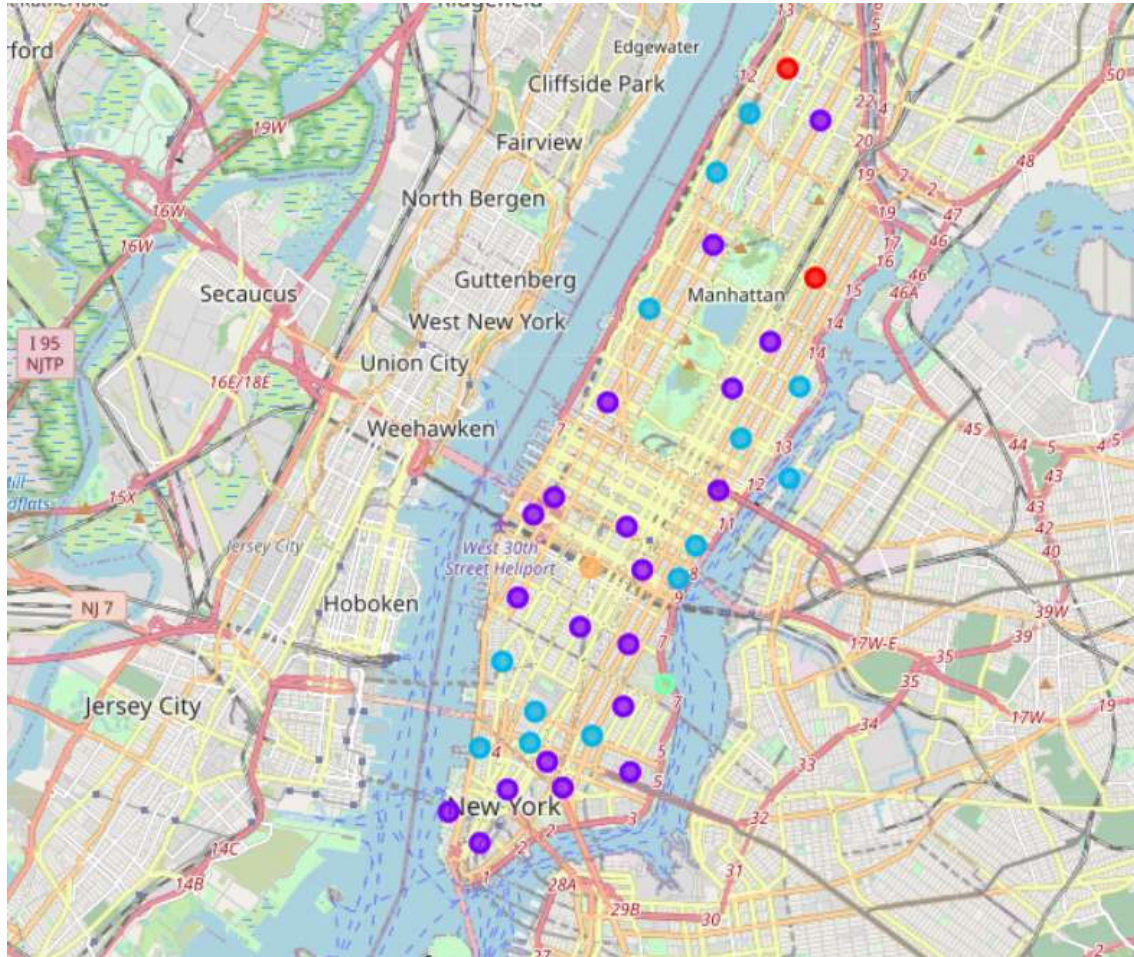Orange Cluster: Neighbourhood of Edenwood

```
#Orange Cluster
bronx_merged.loc[bronx_merged['Cluster Labels'] == 4, bronx_merged.columns[[1] + list(range(5, bronx_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
44	Edenwald	Fried Chicken Joint	Supermarket	Grocery Store	Accessories Store	Music Venue	Nail Salon	Nightclub	Other Great Outdoors	Outlet Store	Paper / Office Supplies Store

Map of Manhattan neighbourhood clusters

Purple Cluster is the most dominant and the central hub of the city with night life, art and culture venues example exhibit, public art, memorial site, theatre and night clubs. There are lots of coffee shops which indicates lots of tourists coming to the venues hence a high traffic area.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Marble Hill	Gym	Sandwich Place	Yoga Studio	Department Store	Steakhouse	Shopping Mall	Clothing Store	Coffee Shop	Seafood Restaurant	Diner
1	Chinatown	Bakery	Chinese Restaurant	Cocktail Bar	Hotpot Restaurant	American Restaurant	Spa	Dessert Shop	Salon / Barbershop	Ice Cream Shop	Boutique
6	Central Harlem	Public Art	Gym / Fitness Center	Seafood Restaurant	Bar	Fried Chicken Joint	Art Gallery	French Restaurant	Chinese Restaurant	African Restaurant	American Restaurant
8	Upper East Side	Exhibit	Italian Restaurant	Gym / Fitness Center	Coffee Shop	Bakery	Juice Bar	French Restaurant	American Restaurant	Spa	Hotel
13	Lincoln Square	Plaza	Performing Arts Venue	Theater	Concert Hall	Café	Bakery	Indie Movie Theater	Gym / Fitness Center	Wine Shop	Gym
14	Clinton	Theater	American Restaurant	Italian Restaurant	Coffee Shop	Gym / Fitness Center	Gym	Hotel	Sandwich Place	Wine Shop	Spa
15	Midtown	Hotel	Coffee Shop	Steakhouse	Clothing Store	Theater	Sandwich Place	Bookstore	Bakery	Sporting Goods Shop	Indian Restaurant
16	Murray Hill	Coffee Shop	Sandwich Place	Hotel	American Restaurant	Japanese Restaurant	Gym / Fitness Center	Sushi Restaurant	Burger Joint	Gym	Pub
17	Chelsea	Coffee Shop	Bakery	French Restaurant	American Restaurant	Art Gallery	Nightclub	Italian Restaurant	Wine Shop	Seafood Restaurant	Ice Cream Shop
19	East Village	Bar	Mexican Restaurant	Pizza Place	Wine Bar	Speakeasy	Cocktail Bar	Italian Restaurant	Vegetarian / Vegan Restaurant	Korean Restaurant	Coffee Shop
20	Lower East Side	Chinese Restaurant	Pizza Place	Coffee Shop	Café	Bakery	Sandwich Place	Art Gallery	Ramen Restaurant	Convenience Store	Diner
22	Little Italy	Bakery	Café	Bubble Tea Shop	Mediterranean Restaurant	Ice Cream Shop	Tea Room	Chinese Restaurant	Italian Restaurant	Sandwich Place	Hotel
25	Manhattan Valley	Coffee Shop	Mexican Restaurant	Pizza Place	Bar	Szechuan Restaurant	Playground	Indian Restaurant	Thai Restaurant	Clothing Store	Bubble Tea Shop
27	Gramercy	Bar	Bagel Shop	Pizza Place	American Restaurant	Italian Restaurant	Grocery Store	Playground	Spa	Cocktail Bar	Mexican Restaurant
28	Battery Park City	Park	Coffee Shop	Hotel	Memorial Site	Playground	Gym	Clothing Store	BBQ Joint	Burger Joint	Beer Garden
29	Financial District	Coffee Shop	Pizza Place	Cocktail Bar	American Restaurant	Bar	Italian Restaurant	Gym / Fitness Center	Hotel	Wine Shop	Dessert Shop
30	Carnegie Hill	Coffee Shop	Café	Yoga Studio	Bookstore	French Restaurant	Cosmetics Shop	Gym	Pizza Place	Wine Shop	Gym / Fitness Center
32	Civic Center	Coffee Shop	Spa	French Restaurant	American Restaurant	Gym / Fitness Center	Hotel	Park	Cocktail Bar	Falafel Restaurant	Sandwich Place
34	Sutton Place	Gym / Fitness Center	Gym	Furniture / Home Store	Italian Restaurant	Park	American Restaurant	Pizza Place	Bar	Coffee Shop	Vegetarian / Vegan Restaurant
38	Flatiron	American Restaurant	Japanese Restaurant	Italian Restaurant	New American Restaurant	Gym / Fitness Center	Wine Shop	Mediterranean Restaurant	Spa	Cosmetics Shop	Coffee Shop
39	Hudson Yards	American Restaurant	Gym / Fitness Center	Hotel	Italian Restaurant	Café	Restaurant	Park	Nightclub	Gym	Coffee Shop

Blue Cluster is the second most dominant, this cluster is residential, and offices evidenced by coffee shops, gym, boutique, food truck, barbershop/salon. It has lots of different types of restaurants thus diverse.

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
5	Manhattanville	Coffee Shop	Chinese Restaurant	Mexican Restaurant	Seafood Restaurant	Food Truck	Deli / Bodega	Italian Restaurant	Bus Station	Café	Diner
9	Yorkville	Italian Restaurant	Gym	Coffee Shop	Bar	Deli / Bodega	Sushi Restaurant	Japanese Restaurant	Wine Shop	Mexican Restaurant	Park
10	Lenox Hill	Pizza Place	Italian Restaurant	Coffee Shop	Cocktail Bar	Sushi Restaurant	Café	Gym	Gym / Fitness Center	Burger Joint	Salad Place
11	Roosevelt Island	Park	Metro Station	Supermarket	Bridge	Bubble Tea Shop	Gym	Greek Restaurant	Noodle House	Soccer Field	Outdoors & Recreation
12	Upper West Side	Italian Restaurant	Bakery	Indian Restaurant	Bar	Sushi Restaurant	Mediterranean Restaurant	Coffee Shop	Café	Wine Bar	Breakfast Spot
18	Greenwich Village	Italian Restaurant	Clothing Store	Sushi Restaurant	Indian Restaurant	Boutique	Coffee Shop	Dessert Shop	Gourmet Shop	Sandwich Place	Bubble Tea Shop
21	Tribeca	Park	American Restaurant	Italian Restaurant	Wine Bar	Café	Spa	Bakery	Men's Store	Gym / Fitness Center	Cocktail Bar
23	Soho	Clothing Store	Italian Restaurant	Shoe Store	Bakery	Coffee Shop	Sporting Goods Shop	Women's Store	Salon / Barbershop	Art Gallery	Boutique
24	West Village	Italian Restaurant	American Restaurant	Cocktail Bar	New American Restaurant	Park	Ice Cream Shop	Wine Bar	Coffee Shop	Bakery	Boutique
26	Morningside Heights	Coffee Shop	American Restaurant	Bookstore	Park	Café	Burger Joint	Deli / Bodega	Food Truck	Mexican Restaurant	Tennis Court
31	Noho	Italian Restaurant	Mexican Restaurant	Pizza Place	Coffee Shop	French Restaurant	Sushi Restaurant	Hotel	Wine Bar	Art Gallery	Grocery Store
35	Turtle Bay	Coffee Shop	Italian Restaurant	Sushi Restaurant	Park	Japanese Restaurant	Seafood Restaurant	Deli / Bodega	Ramen Restaurant	Steakhouse	Garden
36	Tudor City	Park	Café	Mexican Restaurant	Deli / Bodega	Coffee Shop	Pizza Place	Greek Restaurant	Gym	Diner	Restaurant

Smaller Clusters (Red, Green, Orange)

Green and Orange are not as ethnically diverse as the red cluster, it could also be as only had one neighbourhood each, thus could have different results if the cluster size was different.

Red Cluster: Neighbourhoods of Washington Heights, Inwood, Hamilton Heights, East Harlem. This cluster has different type of restaurants such as Thai, Mexican, French, Chinese and Indian restaurants to mention a few clearly indicating that it's a mixed ethnic neighbourhood.

```
#Red Cluster
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
2	Washington Heights	Café	Bakery	Pizza Place	Grocery Store	Bank	Mobile Phone Shop	New American Restaurant	Deli / Bodega	Sandwich Place	Tapas Restaurant
3	Inwood	Mexican Restaurant	Café	Restaurant	Bakery	Pizza Place	Lounge	Spanish Restaurant	Wine Bar	Caribbean Restaurant	Chinese Restaurant
4	Hamilton Heights	Pizza Place	Coffee Shop	Café	Mexican Restaurant	Park	Deli / Bodega	Yoga Studio	Bakery	Liquor Store	Indian Restaurant
7	East Harlem	Mexican Restaurant	Bakery	Thai Restaurant	Spa	Latin American Restaurant	Sandwich Place	Deli / Bodega	New American Restaurant	French Restaurant	Gas Station

Green Cluster: Neighbourhood of Stuyvesant Town

```
#Green Cluster
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 3, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
37	Stuyvesant Town	Park	Bar	Coffee Shop	Gas Station	Fountain	Heliport	Baseball Field	Bistro	Farmers Market	Boat or Ferry

Orange Cluster: Neighbourhood of Midtown South

```
#Orange Cluster
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 4, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
33	Midtown South	Korean Restaurant	Hotel	Dessert Shop	Cosmetics Shop	American Restaurant	Coffee Shop	Cocktail Bar	Japanese Restaurant	Ramen Restaurant	Hotel Bar

Other limitations, the API results feature venues it does not cover office spaces which could highlight local foot traffic means it doesn't give a full picture of all the popular tourist attractions/venues in the area.

Conclusion:

Blue and purple clusters are the most dominant in both, most of the neighbourhoods in the two boroughs have similar popular venues like coffee shops, night clubs and varied restaurants.

Determinants of where venues are depends on industries and local services and government policies etc.

API limitations on results – data used to request venues, limits of API data, etc.