New York City's Night Life

Clustering and Segmentation of Neighborhood

Intruduction

New York City is the most populous city in the United States, home to the headquarters of the United Nations and an important center for international diplomacy. It just might be the most diverse city on the planet, as it is home to over 8.6 million people and over 800 languages. It has many important sites for human development but also many places to keep the city alive at night.

New York City has the most known and important night clubs, shows and many more night events in the entire world. This is why this project will be based on the night life of New York City, so that anybody could see the segmentation of the night life of one of the most important cities in the world.

Problem Description

New York City received a ninth consecutive annual record of approximately 65.2 million tourists in 2018, counting not just overnight visitors but anyone visiting for the day from over 50 miles away, including commuters. Overall the city welcomed 37.9 million visitors who stayed overnight in 2018, of which 13.6 million were international.

Major destinations include the Empire State Building, Ellis Island, the Statue of Liberty on Liberty Island, Broadway theatre productions, Central Park, Times Square, Coney Island, the Financial District, museums, and sports stadiums.

Other major visitor activities include luxury shopping along Fifth and Madison Avenues; entertainment events such as the Tribeca Film Festival; Randalls Island music festivals such as Governors Ball, Panorama and Electric Zoo; and free performances in Central Park at Summerstage and Delacorte Theater. Many New York City ethnic enclaves, such as Jackson Heights, Flushing, and Brighton Beach are major shopping destinations for first and second generation Americans.

Due to this extended tourism there is a need to segment and classify neighborhoods with its night life, this means that every tourist will have acces to where they whan to experience in an easy way.

Target Audience

This Capstone Project is targeted to tourist who wants to know which neighborhoods have the highest night life, this could affect the choice of the Hotel or Airbnb.

Data Collection

Foursquare API

Foursquare API, a location data provider, will be used to make RESTful API calls to retrieve data about venues in different neighborhoods. This is the link to Foursquare Venue Category Hierarchy. Venues retrieved from all the neighborhoods are categorized broadly into 'Arts & Entertainment', 'College & University', 'Event', 'Food', 'Nightlife Spot', 'Outdoors & Recreation', etc. An extract of an API call is as follows:

https://developer.foursquare.com/docs

New York City Dataset

New York City neighborhood data will be obtained by the open source data:

https://cocl.us/new_york_dataset

With this data we'll obtain the Borough, Neighborhood and the latitude and longitude in New York City. This will be done using the Python Geocoder package.

Methodology

Import libraries

```
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 # uncomment this line if you haven't completed the Foursquare API lab
from geopy.geocoders import Nominatim # convert an address into latitude and longitude values

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

#!conda install -c conda-forge folium=0.5.0 --yes # uncomment this line if you haven't completed the Foursquare API lab
import folium # map rendering library

print('Libraries imported.')
```

Load Data

```
!wget -q -O 'newyork_data.json' https://cocl.us/new_york_dataset
print('Data downloaded!')
```

```
Data downloaded!
```

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

```
newyork_data
```

Tranform Data into Data Frame

```
# define the dataframe columns
column_names = ['Borough', 'Neighborhood', 'Latitude', 'Longitude']

# instantiate the dataframe
neighborhoods = pd.DataFrame(columns=column_names)
```

```
neighborhoods
```

Borough	Neighborhood	Latitude	Longitude

```
for data in neighborhoods_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]

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

```
neighborhoods.head()
```

	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

Geopy for Latitude and Longitude

```
address = 'New York City, NY'

geolocator = Nominatim(user_agent="ny_explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print('The geograpical coordinate of New York City are {}, {}.'.format(latitude, longitude))
```

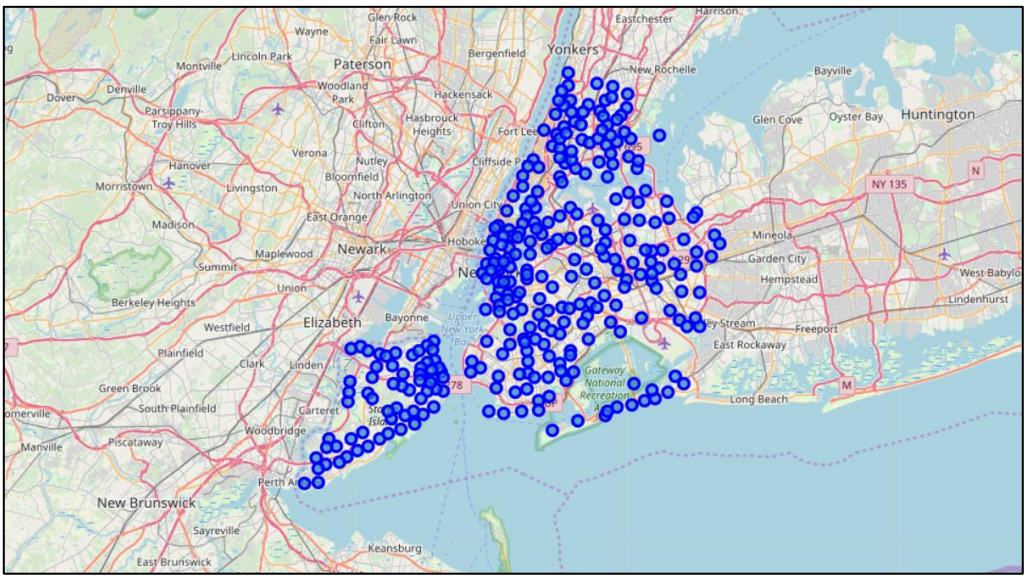
```
The geograpical coordinate of New York City are 40.7127281, -74.0060152.
```

Map of NYC Neighborhoods

```
# create map of New York using latitude and longitude values
map_newyork = folium.Map(location=[latitude, longitude], zoom_start=10)

# add markers to map
for lat, lng, borough, neighborhood in zip(neighborhoods['Latitude'], neighborhoods['Longitude'], neighborhoods['Borough'], neighborhoods['Neig
    label = '{}, {}'.format(neighborhood, borough)
    label = folium.Popup(label, parse_html=True)
    folium.CircleMarker(
        [lat, lng],
        radius=5,
        popup=label,
        color='blue',
        fill=True,
        fill_color='#3186cc',
        fill_opacity=0.7,
        parse_html=False).add_to(map_newyork)

map_newyork
```



Clustering

```
# set number of clusters
kclusters = 5

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

# run k-means clustering
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(manhattan_grouped_clustering)

# check cluster labels generated for each row in the dataframe
kmeans.labels_[0:10]
```

```
array([3, 1, 0, 1, 1, 0, 0, 4, 4, 3], dtype=int32)
```

```
# add clustering labels
neighborhoods_venues_sorted.insert(0, 'Cluster Labels', kmeans.labels_)

manhattan_merged = manhattan_data

# merge toronto_grouped with toronto_data to add latitude/longitude for each neighborhood
manhattan_merged = manhattan_merged.join(neighborhoods_venues_sorted.set_index('Neighborhood'), on='Neighborhood')

manhattan_merged.head() # check the last columns!
```

	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	3	Coffee Shop	Sandwich Place	American Restaurant	Kids Store	Bank	Supplement Shop	Gym	Miscellaneous Shop	Donut Shop	Shopping Mall
1	Manhattan	Chinatown	40.715618	-73.994279	1	Chinese Restaurant	American Restaurant	Cocktail Bar	Salon / Barbershop	Spa	Optical Shop	Bakery	Vietnamese Restaurant	Hotpot Restaurant	Asian Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	4	Café	Bakery	Grocery Store	Mobile Phone Shop	Spanish Restaurant	Deli / Bodega	Chinese Restaurant	Mexican Restaurant	Tapas Restaurant	Coffee Shop
3	Manhattan	Inwood	40.867684	-73.921210	4	Mexican Restaurant	Lounge	Café	Pizza Place	Restaurant	Bakery	Park	Chinese Restaurant	American Restaurant	Frozen Yogurt Shop
4	Manhattan	Hamilton Heights	40.823604	-73.949688	4	Pizza Place	Café	Park	Coffee Shop	Mexican Restaurant	Yoga Studio	Bakery	Sandwich Place	Cocktail Bar	School

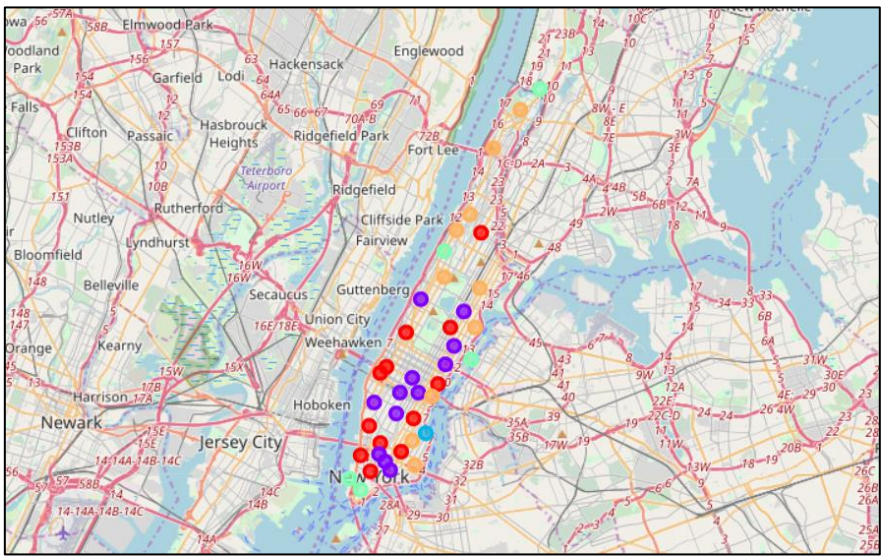
Cluster Map

```
# create map
map_clusters = folium.Map(location=[latitude, longitude], zoom_start=11)

# set color scheme for the clusters
x = np.arange(kclusters)
ys = [i + x + (i*x)**2 for i in range(kclusters)]
colors_array = cm.rainbow(np.linspace(0, 1, len(ys)))
rainbow = [colors.rgb2hex(i) for i in colors_array]

# add markers to the map
markers_colors = []
for lat, lon, poi, cluster in zip(manhattan_merged['Latitude'], manhattan_merged['Longitude'], manhattan_merged['Neighborhood'], manhattan_merged['Cluste
    label = folium.Popup(str(poi) + ' Cluster ' + str(cluster), parse_html=True)
    folium.CircleMarker(
        [lat, lon],
        radius=5,
        popup=label,
        color=rainbow[cluster-1],
        fill=True,
        fill_color=rainbow[cluster-1],
        fill_opacity=0.7).add_to(map_clusters)

map_clusters
```



Results

Cluster 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	Grocery Store	Mobile Phone Shop	Spanish Restaurant	Deli / Bodega	Chinese Restaurant	Mexican Restaurant	Tapas Restaurant	Coffee Shop
3	Inwood	Mexican Restaurant	Lounge	Café	Pizza Place	Restaurant	Bakery	Park	Chinese Restaurant	American Restaurant	Frozen Yogurt Shop
4	Hamilton Heights	Pizza Place	Café	Park	Coffee Shop	Mexican Restaurant	Yoga Studio	Bakery	Sandwich Place	Cocktail Bar	School
5	Manhattanville	Coffee Shop	Deli / Bodega	Italian Restaurant	Park	Seafood Restaurant	Mexican Restaurant	Japanese Curry Restaurant	Café	Bike Trail	Sushi Restaurant
7	East Harlem	Mexican Restaurant	Thai Restaurant	Bakery	Latin American Restaurant	Deli / Bodega	Cuban Restaurant	Pizza Place	Beer Bar	Taco Place	Gas Station
9	Yorkville	Italian Restaurant	Coffee Shop	Gym	Bar	Sushi Restaurant	Deli / Bodega	Bakery	Wine Shop	Diner	Japanese Restaurant
19	East Village	Bar	Wine Bar	Ice Cream Shop	Chinese Restaurant	Mexican Restaurant	Cocktail Bar	Pizza Place	Vegetarian / Vegan Restaurant	Italian Restaurant	Coffee Shop
20	Lower East Side	Art Gallery	Coffee Shop	Pizza Place	Cocktail Bar	Bakery	Café	Ramen Restaurant	Park	Japanese Restaurant	Shoe Store
25	Manhattan Valley	Coffee Shop	Bar	Yoga Studio	Mexican Restaurant	Café	Thai Restaurant	Pizza Place	Indian Restaurant	Playground	Furniture / Home Store
36	Tudor City	Café	Park	Mexican Restaurant	Pizza Place	Deli / Bodega	Coffee Shop	Diner	Greek Restaurant	Garden	Asian Restaurant

Cluster 2

	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
6	Central Harlem	African Restaurant	Bar	American Restaurant	Seafood Restaurant	French Restaurant	Chinese Restaurant	Tapas Restaurant	Spa	Cosmetics Shop	Beer Bar
8	Upper East Side	Italian Restaurant	Exhibit	Coffee Shop	Bakery	Art Gallery	Juice Bar	Gym / Fitness Center	French Restaurant	American Restaurant	Hotel
13	Lincoln Square	Theater	Café	Plaza	Concert Hall	Italian Restaurant	Performing Arts Venue	French Restaurant	Gym / Fitness Center	American Restaurant	Indie Movie Theater
14	Clinton	Theater	Italian Restaurant	Gym / Fitness Center	Coffee Shop	American Restaurant	Hotel	Wine Shop	Spa	Sandwich Place	Gym
18	Greenwich Village	Italian Restaurant	Clothing Store	Sushi Restaurant	Café	Indian Restaurant	Seafood Restaurant	French Restaurant	Dessert Shop	Gourmet Shop	Sandwich Place
21	Tribeca	American Restaurant	Park	Italian Restaurant	Spa	Café	Wine Shop	Wine Bar	Greek Restaurant	Men's Store	Coffee Shop
24	West Village	Italian Restaurant	New American Restaurant	Cosmetics Shop	Park	Cocktail Bar	Wine Bar	American Restaurant	Coffee Shop	Theater	Bakery
27	Gramercy	Bar	Italian Restaurant	Mexican Restaurant	Pizza Place	Grocery Store	Thai Restaurant	Playground	Bagel Shop	Diner	Comedy Club
31	Noho	Italian Restaurant	Cocktail Bar	French Restaurant	Mexican Restaurant	Gift Shop	Bookstore	Rock Club	Coffee Shop	Pizza Place	Grocery Store
32	Civic Center	Italian Restaurant	Gym / Fitness Center	Coffee Shop	Sandwich Place	French Restaurant	Hotel	Yoga Studio	Cocktail Bar	Spa	Park
35	Turtle Bay	Italian Restaurant	Steakhouse	Sushi Restaurant	Coffee Shop	Wine Bar	Ramen Restaurant	French Restaurant	Park	Japanese Restaurant	Café

Cluster 3

	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	Chinatown	Chinese Restaurant	American Restaurant	Cocktail Bar	Salon / Barbershop	Spa	Optical Shop	Bakery	Vietnamese Restaurant	Hotpot Restaurant	Asian Restaurant
10	Lenox Hill	Coffee Shop	Italian Restaurant	Sushi Restaurant	Pizza Place	Cocktail Bar	Café	Burger Joint	Sporting Goods Shop	Gym	Gym / Fitness Center
12	Upper West Side	Italian Restaurant	Coffee Shop	Wine Bar	Bar	Bakery	Mediterranean Restaurant	Café	Gym / Fitness Center	Ice Cream Shop	Indian Restaurant
15	Midtown	Hotel	Food Truck	Coffee Shop	Theater	Clothing Store	Sporting Goods Shop	Bakery	Café	Japanese Restaurant	Bookstore
16	Murray Hill	Coffee Shop	Sandwich Place	American Restaurant	Japanese Restaurant	Hotel	Gym / Fitness Center	Italian Restaurant	Chinese Restaurant	Sushi Restaurant	Bar
17	Chelsea	Coffee Shop	Bakery	Italian Restaurant	Ice Cream Shop	Wine Shop	Theater	American Restaurant	Hotel	Nightclub	Bookstore
22	Little Italy	Italian Restaurant	Café	Bubble Tea Shop	Bakery	Mediterranean Restaurant	Sandwich Place	Pizza Place	Cocktail Bar	Clothing Store	Salon / Barbershop
23	Soho	Clothing Store	Boutique	Women's Store	Art Gallery	Shoe Store	Italian Restaurant	Sporting Goods Shop	Bakery	Mediterranean Restaurant	Men's Store
30	Carnegie Hill	Coffee Shop	Pizza Place	Cosmetics Shop	Yoga Studio	Bakery	Gym	Bookstore	Café	Japanese Restaurant	Wine Shop
33	Midtown South	Korean Restaurant	Hotel	Hotel Bar	Japanese Restaurant	Coffee Shop	Dessert Shop	American Restaurant	Gym / Fitness Center	Cocktail Bar	Scenic Lookout
34	Sutton Place	Gym / Fitness Center	Italian Restaurant	American Restaurant	Furniture / Home Store	Gym	Coffee Shop	Beer Garden	Bakery	Grocery Store	Beer Bar

Cluster 4

	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	Coffee Shop	Sandwich Place	American Restaurant	Kids Store	Bank	Supplement Shop	Gym	Miscellaneous Shop	Donut Shop	Shopping Mall
11	Roosevelt Island	Deli / Bodega	Park	Sandwich Place	Coffee Shop	Bridge	Farmers Market	Greek Restaurant	Metro Station	Supermarket	Bubble Tea Shop
26	Morningside Heights	Park	Bookstore	Coffee Shop	American Restaurant	Food Truck	Sandwich Place	Burger Joint	Deli / Bodega	Pizza Place	Ice Cream Shop
28	Battery Park City	Park	Coffee Shop	Hotel	Wine Shop	Clothing Store	Gym	Women's Store	Memorial Site	Boat or Ferry	Pizza Place
29	Financial District	Coffee Shop	American Restaurant	Pizza Place	Hotel	Gym	Italian Restaurant	Wine Shop	Food Truck	Steakhouse	Gym / Fitness Center

Discussion

By looking at the results in this analysis we can obtain the following information over nightlife in distinct neighborhoods.

East Village

-1st most common venue bar

-2nd most common venue wine bar

- 6th most common venue cocktail bar

Gramercy

-1st most common venue Bar

Nojo

- 2nd most common venue cocktail bar

- 7th most common venue Rock Club

Upper West Side

- 3rd most common venue Wine Bar

- 4th most common venue Bar

West Village

- 5th most common venue cocktail Bar

- 6th most common venue Wine Bar

- 9th most common venue Theater

Conclusion

By making this analysis it is concluded that the Neighborhood with the highest nightlife is East Village followed up by West Village then Upper West Side, then Gramercy and then Nojo. This are the top 5 neighborhoods for nightlife in New York. The cluster with the lowest

nightlife is Cluster 4 which will indacate that nightlife tourist will tend to avoide those neighborhoods.

This analysis could be duplicated for any other type of venues such as hotels, restaurants, etc. In this case was made for nightlife such as clubs, bars, Concerthalls etc. to help tourist with their neighborhood selection.