# Determining an Upscale Restaurant Location in Manhattan

IBM Coursera Capstone Final Assessment

*Alexa Hijazi*

# Determining Restaurant Location

- Restaurant Variables to Consider:
  - High cost menu
- What to look for in a location
  - Proximity to an Office
    - Koo *et al*. (1999) has found that high food costs at restaurants are less bothersome to individuals whose primary purpose is that of business, rather than family meal outings.
  - Proximity to a Hotel
    - Kunst *et al.* (2019) found that nearly 57% of individuals travelling on business trips selected a hotel based on its proximity to restaurants and bars.

# Data Sources

- **Neighborhood boundaries:**
  - Obtained from the NYU Spatial Data Repository.
  - Contains details of neighborhood names, location (latitude and longitude), and the bureau each neighborhood resides in. This will then be narrowed down to focus specifically on the borough of Manhattan. The neighborhood locations will be utilized when creating cluster locations in Foursquare API.

- **Neighborhood venues:**
  - Data for each neighborhood will be obtained from the foursquare database.
  - Data will be "called" in the notebook using foursquare credentials, and the location of interest latitude and longitude. https://geo.nyu.edu/catalog/nyu_2451_34572

# Methods

- The json file from the NYU spatial repository was imported and transformed into a pandas data frame
- Geopy function was used to extract the latitude and longitude, the preliminary map was constructed to ensure the correct location was utilized
- Foursquare credentials were connected, and the Get request was utilized to import the URL file which was then translated into a pandas data frame. Columns were defined to better extract venue categories to isolate those of interest
- Separate data frames were constructed for Hotels and Offices, these were combined into a single data frame
- Clusters of neighbhorhoods were created using kmeans clusters

# Methods 2

- An initial frequency count was conducted to see where the highest frequency of hotels & offices were in Manhattan

```
In [262]: print('There are {} hotels and offices on record in New York.'.
          format(len(Office_Hotel_List['Venue'].unique())))

          There are 62 hotels and offices on record in New York.
```

```
In [261]: Office_Hotel_List['Neighborhood'].value_counts()

Out[261]: Midtown               7
          Midtown South         5
          Battery Park City     5
          Financial District    4
          Civic Center          4
          Chelsea               4
          Hudson Yards          4
          Upper East Side       3
          Noho                  3
          Gramercy              3
          Murray Hill           3
          Clinton               3
          Tribeca               2
          Little Italy          2
          Turtle Bay            2
          Soho                  2
          Lincoln Square        2
          Tudor City            2
          Carnegie Hill         1
          Upper West Side       1
          Sutton Place          1
          Flatiron              1
          Greenwich Village     1
          Chinatown             1
          Name: Neighborhood, dtype: int64
```

The highest frequency of hotels and offices are in Midtown and Midtown South

# Methods 3

- We then created a visual using folium. Adding our markers of interest to the map, Hotels and Offices for the purpose of this study

```
In [240]: # create map
map_clusters = folium.Map(location=[40.7896239, -73.9598939], zoom_start=1
1)

# 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 clusters to the map
markers_colors = []
for lat, lon, poi, cluster in zip(manhattan_merged['Latitude'], manhattan_
merged['Longitude'], manhattan_merged['Neighborhood'], manhattan_merged['C
luster Labels']):
    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)

#add points of interest markers to the map (Hotels & Offices)
for lat, long, points, cat in zip(Office_Hotel_List['Venue Latitude'], Off
ice_Hotel_List['Venue Longitude'],
                                  Office_Hotel_List['Venue'], Office_Hotel_
List['Venue Category']):
        label = '{}, {}'.format(points, cat)
        label = folium.Popup(label, parse_html=True)
        folium.CircleMarker(
            [lat, long],
            radius = 7,
            popup = label,
            color = 'yellow',
            fill = True,
            fill_color = 'blue',
```

# Methods 4

- We then analyzed all 5 clusters.

In [241]:
```
manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 0, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

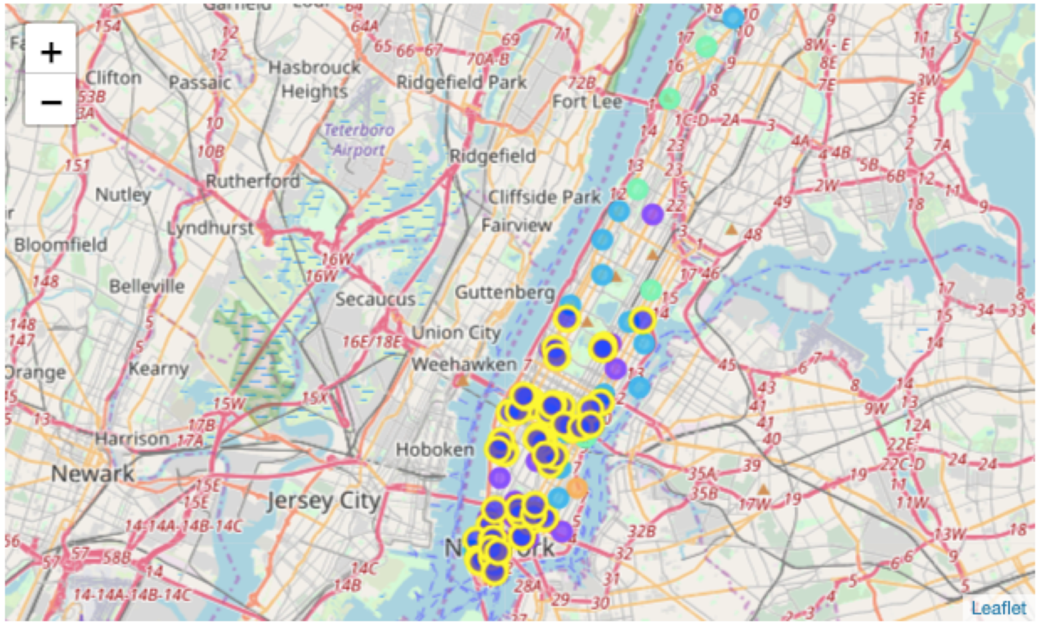
Out[241]:

| | 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 M Comn Vei |
|---|---|---|---|---|---|---|---|---|
| **33** | Midtown South | Korean Restaurant | Hotel | Hotel Bar | Japanese Restaurant | Dessert Shop | Coffee Shop | Cosme Shop |

# Results

- Highest Frequency of Hotels and Offices observed in Midtown and Midtown South Neighborhood
- Cluster analysis using Folium visualizing the number of Hotels and Offices in proximity to the neighborhood clusters defined.

Out[240]:



# Discussion

- Highest frequency of Offices and Hotels in Midtown South and Midtown Neighborhoods.
- This would suggest that due to the nature of our clients proposed menu pricing, these would be the most optimal neighborhoods.
- Other factors that can contribute to restaurant success is its proximity to competitors (Parsa *et al*., 2011). As there is a lower frequency f restaurants in Midtown south, that would be our first choice.

# Conclusion and Limitations

- *Limitations and Future Research*
  - We found data relating to Office locations were limited, and likely limited the generalizability of these results. Future studies should look to incorporate a larger data set to determine the frequency of office locations in the borough of Manhattan.
  - Additionally. Other factors contribute to the overall success of a restaurant’s outcome. While proximity to a hotel and office are factors in determining the success of a restaurant, other factors include vicinity to competition (Parsa *et al.,* 2011). Future studies should incorporate other factors which can contribute to the failure of a restaurant on the basis of location.
- Conclusion
  - Most ideal location for an upscale restaurant is either in Midtown South or Midtown. As there is a lower concentration of restaurants in Midtown South than in Midtown, we would suggest the most ideal location for the new restaurant location is in Midtown.

# References

H. G. Parsa, John Self, Sandra Sydnor-Busso & Hae Jin Yoon (2011) Why Restaurants Fail? Part II - The Impact of Affiliation, Location, and Size on Restaurant Failures: Results from a Survival Analysis, Journal of Foodservice Business Research, 14:4, 360-379, DOI: 10.1080/15378020.2011.625824

Pillsbury, R. (1987) FROM HAMBURGER ALLEY TO HEDGEROSE HEIGHTS: TOWARD A MODEL OF RESTAURANT LOCATION DYNAMICS, The Professional Geographer, 39:3, 326-344, DOI: 10.1111/j.0033-0124.1987.00326.x

Koo, L.C., Tao, F.K. and Yeung, J.H., 1999. Preferential segmentation of restaurant attributes through conjoint analysis. *international Journal of Contemporary Hospitality management*.

Kunst, A. (2017) How important are nearby restaurants and bars when choosing a hotel for a business trip? *Statista*. https://www.statista.com/statistics/719823/importance-of-restaurant-proximity-to-hotels-to-business-travelers-us/