

CAPSTONE PROJECT – THE BATTLE OF NEIGHBORHOODS – REPORT

1. INTRODUCTION

The City of New York, is the most populous city in the United States. It is diverse and is the financial capital of the USA. It is multicultural. It provides a lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major centre for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theatre, fashion, and the arts in the United States. This also means that the market is highly competitive. As it is a highly developed city so the cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analysed carefully. The insights derived from analysis will give a good understanding of the business environment which helps in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

BUSINESS PROBLEM

The City of New York is famous for its excellent cuisine. Its food culture includes an array of international cuisines influenced by the city's immigrant history. Sushi restaurants have become so popular in the United States now it seems that there is one on every corner, not only in major cities but also in smaller cities. Starting a sushi restaurant can be a great business opportunity, but you need to distinguish yourself from others to enjoy long-term success.

If you plan a real restaurant that can demand higher prices for fresh fish, delivered daily from Japan, focus on neighbourhoods and outlets that already attract a sophisticated Japanese client. If you plan a cheap buffet restaurant, points to the masses looking for affordable high-traffic locations with large shopping centres and other local points of interest.

My client wants to open his business in Manhattan area, so I focus on that borough during my analysis. We define potential neighbourhood based on the number of sushi bars which are operating right in each neighbourhood. Manhattan has full potential but also is a very challenging district to open a business because of high competition. New sushi bar should be open in an area that inadequate neighbourhood in this way the bar can attract more customers. Therefore, this analysis necessary to ensure that we have enough customers and that we are not so close to other sushi places.

2. DATA

Data 1: Neighbourhood has a total of 5 boroughs and 306 neighbourhoods. In order to segment the neighbourhoods and explore them, we will essentially need a dataset that contains the 5 boroughs and the neighbourhoods that exist in each borough as well as the the latitude and longitude coordinates of each neighbourhood. This dataset exists for free on the web. Link to the dataset is: https://geo.nyu.edu/catalog/nyu_2451_34572

BOROUGH

NEIGHBORHOOD

LATITUDE

LONGITUDE

Data2: New York city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighbourhood will use the Foursquare API to explore neighbourhoods in New York City. The below is image of the Foursquare API data.

In addition, Sushi category Id 4bf58dd8d48988d1d2941735 is used for retrieving data from Foursquare API.

3. METHODOLOGY

SEGMENTATION AND CLUSTERING OF NEIGHBORHOODS IN NEW YORK CITY

Read the data and Define data frame

```
Out[6]:
```

| | 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 |

Above, I have done convert addresses into their equivalent latitude and longitude values. Then we will use the Foursquare API to explore neighborhoods in Manhattan, New York. After that, explore function to get sushi restaurant categories in each neighborhood.

| Neighborhood | Neighborhood | Latitude | Neighborhood | Longitude | \ |
|--------------|-----------------------|----------------|-----------------|---------------------|---|
| 0 | Marble Hill | 40.876551 | | -73.910660 | |
| 1 | Chinatown | 40.715618 | | -73.994279 | |
| 2 | Chinatown | 40.715618 | | -73.994279 | |
| 3 | Chinatown | 40.715618 | | -73.994279 | |
| 4 | Chinatown | 40.715618 | | -73.994279 | |
| | Venue | Venue Latitude | Venue Longitude | Venue Category | |
| 0 | Planet Tokyo | 40.886233 | -73.989479 | Sushi Restaurant | |
| 1 | Nakaji | 40.715791 | -73.996855 | Sushi Restaurant | |
| 2 | Shinsen | 40.715608 | -73.996611 | Japanese Restaurant | |
| 3 | Whole Foods Sushi Bar | 40.724000 | -73.992277 | Sushi Restaurant | |
| 4 | Douzo | 40.719069 | -73.990781 | Japanese Restaurant | |

```
Out[33]: (1104, 7)
```

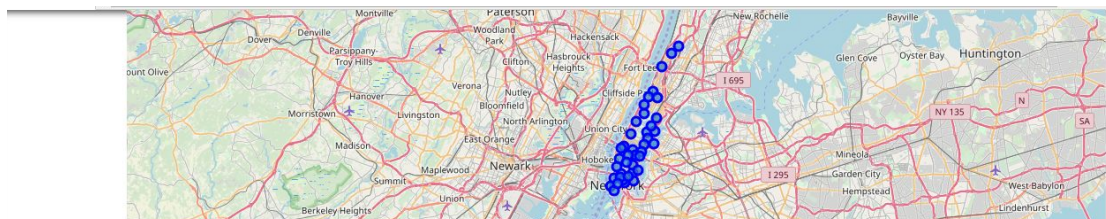


Fig 3.1: Sushi bars in Manhattan, New York

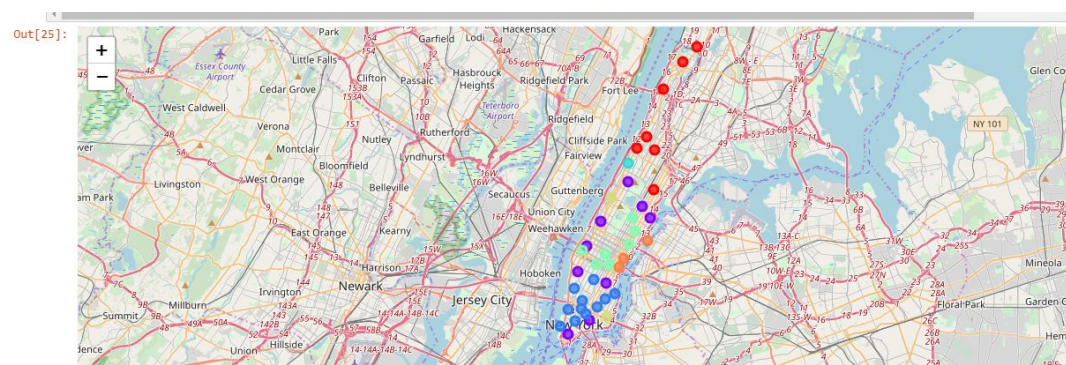
| | Neighborhood | Asian Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant | Noodle House | Poke Place | Ramen Restaurant | Restaurant | Sake Bar | Sandwich Place | Seafood Restaurant | Smoothie Shop |
|---|--------------|------------------|--------|--------------------|--------------|---------------|---------------|---------------------|---------------------|--------------|------------|------------------|------------|----------|----------------|--------------------|---------------|
| 0 | Marble Hill | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | Chinatown | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | Chinatown | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | Chinatown | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | Chinatown | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

Then use this feature to group the neighborhoods into clusters K-means clustering algorithm will be use to complete this task. And also, the Folium library to visualize the neighborhoods in Manhattan and its emerging clusters.

Out[22]:

| | 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 | Sushi Restaurant | Japanese Restaurant | Noodle House | Theme Restaurant | Vegetarian / Vegan Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |
| 1 | Carnegie Hill | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Chinese Restaurant | Noodle House | Bakery | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 2 | Central Harlem | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 3 | Chelsea | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Vegetarian / Vegan Restaurant | Smoothie Shop | Seafood Restaurant | Sandwich Place | Sake Bar | Restaurant | Ramen Restaurant |
| 4 | Chinatown | Sushi Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |

| | 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 | 0 | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 1 | Manhattan | Chinatown | 40.715618 | -73.994279 | 1 | Sushi Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 2 | Manhattan | Washington Heights | 40.851903 | -73.936900 | 0 | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 3 | Manhattan | Inwood | 40.867684 | -73.921210 | 0 | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 4 | Manhattan | Hamilton Heights | 40.823604 | -73.949688 | 0 | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |



4. RESULTS

K-mean Cluster Using K-mean to clustering data area with less number of sushi bars

K=7

CLUSTER 0:

Out[26]:

| | 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 | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 2 | Washington Heights | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 3 | Inwood | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 4 | Hamilton Heights | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 5 | Manhattanville | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 6 | Central Harlem | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 7 | East Harlem | Sushi Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |

CLUSTER 1:

In [27]: `manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 1, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]`

Out[27]:

| | 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 | Sushi Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 9 | Yorkville | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Chinese Restaurant | Noodle House | Bakery | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 13 | Lincoln Square | Sushi Restaurant | Japanese Restaurant | Smoothie Shop | Chinese Restaurant | Grocery Store | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Cocktail Bar | Deli / Bodega |
| 14 | Clinton | Sushi Restaurant | Japanese Restaurant | Poke Place | Chinese Restaurant | Cocktail Bar | Asian Restaurant | Seafood Restaurant | Sandwich Place | Sake Bar | Restaurant |
| 17 | Chelsea | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Vegetarian / Vegan Restaurant | Smoothie Shop | Seafood Restaurant | Sandwich Place | Sake Bar | Restaurant | Ramen Restaurant |
| 25 | Manhattan Valley | Sushi Restaurant | Hawaiian Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |
| ... | ... | Sushi | Japanese | Vegetarian / | Chinese | ... | ... | ... | ... | ... | Hawaiian |

CLUSTER 2:

In [28]: `manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 2, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]`

Out[28]:

| | 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 |
|----|-------------------|-----------------------|-----------------------|-------------------------------|-------------------------------|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 18 | Greenwich Village | Sushi Restaurant | Japanese Restaurant | Sake Bar | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |
| 19 | East Village | Sushi Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 21 | Tribeca | Sushi Restaurant | Noodle House | Japanese Restaurant | Theme Restaurant | Vegetarian / Vegan Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |
| 22 | Little Italy | Sushi Restaurant | Japanese Restaurant | Noodle House | Vegetarian / Vegan Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 23 | Soho | Sushi Restaurant | Japanese Restaurant | Noodle House | Theme Restaurant | Vegetarian / Vegan Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |
| 24 | West Village | Sushi Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Sake Bar | Asian Restaurant | Seafood Restaurant | Sandwich Place | Smoothie Shop | Restaurant | Ramen Restaurant |
| 28 | Battery Park City | Sushi Restaurant | Japanese Restaurant | Noodle House | Theme Restaurant | Vegetarian / Vegan Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |

CLUSTER 3:

In [29]: `manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 3, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]`

Out[29]:

| | 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 |
|----|---------------------|-----------------------|-----------------------|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 26 | Morningside Heights | Sushi Restaurant | Hawaiian Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Japanese Restaurant |

CLUSTER 4:

In [30]: `manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 4, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]`

Out[30]:

| | 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 |
|----|-----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-------------------------------|-------------------------------|-----------------------|------------------------|
| 8 | Upper East Side | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 10 | Lenox Hill | Sushi Restaurant | Asian Restaurant | Japanese Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |
| 15 | Midtown | Sushi Restaurant | Asian Restaurant | Seafood Restaurant | Japanese Restaurant | Ramen Restaurant | Bakery | Vegetarian / Vegan Restaurant | Sandwich Place | Sake Bar | Restaurant |
| 16 | Murray Hill | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Restaurant | Bakery | Chinese Restaurant | Ramen Restaurant | Vegetarian / Vegan Restaurant | Sake Bar | Sandwich Place |

CLUSTER 5:

```
In [31]: manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 5, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

Out[31]:

| | 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 |
|----|-----------------|-----------------------|-----------------------|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 12 | Upper West Side | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Grocery Store | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Hawaiian Restaurant |
| 20 | Lower East Side | Sushi Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant |

CLUSTER 6:

```
In [32]: manhattan_merged.loc[manhattan_merged['Cluster Labels'] == 6, manhattan_merged.columns[[1] + list(range(5, manhattan_merged.shape[1]))]]
```

Out[32]:

| | 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 |
|----|------------------|-----------------------|-----------------------|-----------------------|-------------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 11 | Roosevelt Island | Sushi Restaurant | Asian Restaurant | Noodle House | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store | Hawaiian Restaurant | Japanese Restaurant |
| 35 | Turtle Bay | Sushi Restaurant | Japanese Restaurant | Asian Restaurant | Steakhouse | Seafood Restaurant | Bakery | Chinese Restaurant | Cocktail Bar | Deli / Bodega | Grocery Store |
| 36 | Tudor City | Sushi Restaurant | Asian Restaurant | Japanese Restaurant | Vegetarian / Vegan Restaurant | Steakhouse | Smoothie Shop | Seafood Restaurant | Sandwich Place | Sake Bar | Restaurant |

Based on data frame analysis above Cluster 5 (Upper West Side) and Cluster 3 (Morningside Heights) areas are the best places to open a new sushi bar business.

5. DISCUSSIONS

In this section, I would be discussing the observations I have noted and the recommendation that I can make based on the results.

This analysis is performed on limited data. This may be right or may be wrong.

But if good amount of data is available there is scope to come up with better results.

- There is high competition in Midtown and Soho so it is very risky to open business in these areas.
- Central Harlem has also potential, where it is closest to the Morningside Heights area.
- It can be done more detailed analysis by adding other factors such as transportation, demographics of inhabitants.

Finally, Foursquare proved to be a good source of data but frustrating at times. Despite having a Developer account, I regularly exceeded my hourly limit locking me out for the day.

6. **CONCLUSION**

The Goal of the project is met but there are more further improvements and development as mentioned below. Now we are able to develop an application that supports a business in an unknown location. A venue with low risk and competition can be identified like Sushi restaurants.