The Battle of the Neighborhoods

## INTRODUCTION / BUSINESS PROBLEM

**PROBLEM BACKGROUND:**

The NBCNEWS article <https://www.nbcnews.com/news/asian-america/anti-asian-hate-crimes-increased-nearly-150-2020-mostly-n-n1260264> stated that “Anti-Asian hate crimes increased by nearly 150% in 2020, mostly in N.Y. and L.A.”

“The analysis revealed a surge in cities such as New York, where anti-Asian hate crimes rose from three in 2019 to 28 in 2020, a 833 percent increase.”

**PROBLEM DESCRIPTION:**

Ever since the start of the pandemic, Asians have been the target of hate crimes. There is a need to increase awareness in the Asian community and help them protect and empower themselves. The organizers would like to conceptualize the program as appropriate to the community of Manhattan in New York City first. Restaurants would be ideal to aim first as they can also help promote to their customers as well.

**TARGET AUDIENCE:**

Asian restaurants in the Manhattan borough of New York City

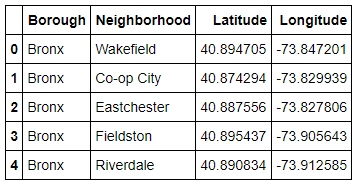
**SUCCESS CRITERIA:**

The success criteria of the project will be to provide to the organizers the spread of Asian Restaurants in Manhattan and to provide comparative data on their composition so the organizers can custom-built the program and tailor fit as appropriate for increase adaptability and make it more effective.

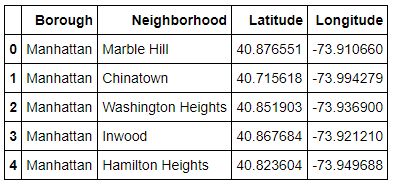
## DATA

We will be looking at the New York City Neighborhoods data.

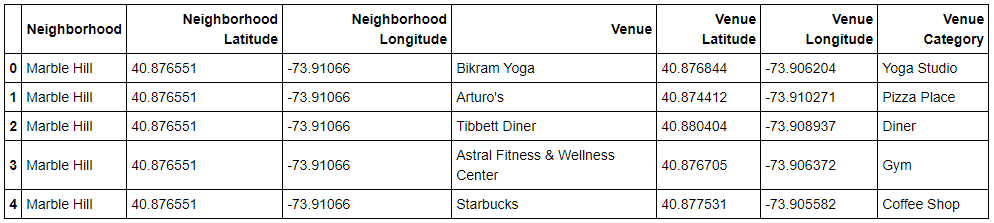
These data will be loaded from <https://cf-courses-data.s3.us.cloud-object-storage.appdomain.cloud/IBMDeveloperSkillsNetwork-DS0701EN-SkillsNetwork/labs/newyork_data.json>



Then, we will explore the neighborhoods in Manhattan. Python Geocode package will be used to get the latitude and longitude coordinates for those neighborhoods.



FourSquare API (<https://foursquare.com/>) will be used to get nearby venues at the different Manhattan neighbourhoods.



The program will be catered to the Asian community and would like to start with Asian restaurants; therefore, we will extract the different Asian-related restaurant venues. We will analyze where they are located and break them into clusters to further cater the program to be developed depending on the distinguishing characteristics of the clusters.



## METHODOLOGY

**Business Understanding:**

The program development team would like to deploy a program to help the Asian Restaurants in the Manhattan neighbourhoods to increase awareness on and protection from Anti-Hate crimes. We will provide them insights to the Asian restaurants in the area and what would help best cater the program for successful deployment and adaption.

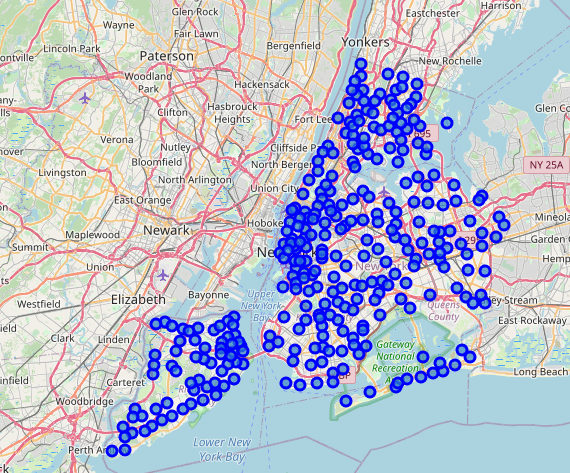
**Analytic Approach:**

New York city has a total of 5 boroughs and 306 neighborhoods. This project will cluster the neighbourhoods of Manhattan.

Data Exploration Analysis:

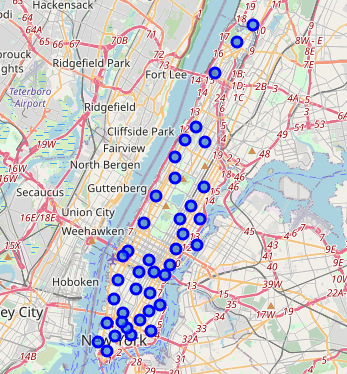
1. New York City Geographical Coordinates data (Python Geocoder package)
   1. Visualization (Folium)

New York neighbourhood visualization



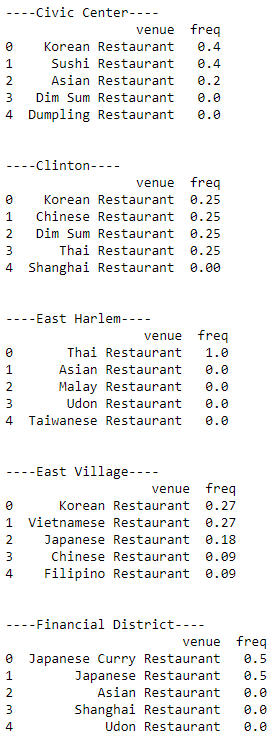
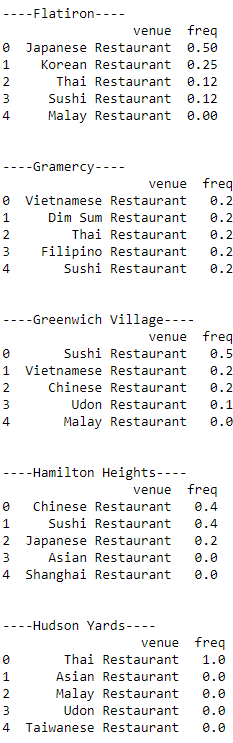
1. Manhattan Geographical Coordinates data (Python Geocode package)
   1. Visualization (Folium)

Manhattan neighbourhood visualization

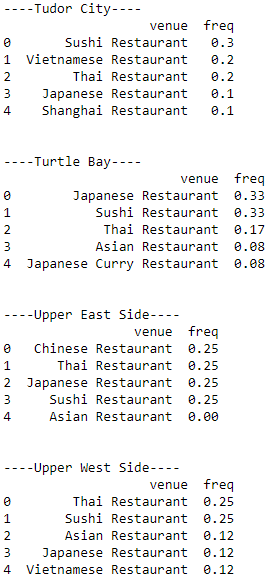
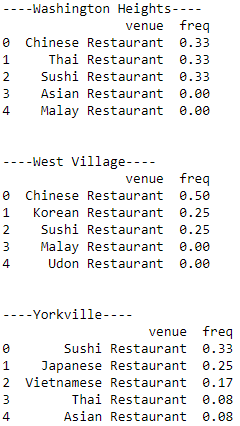


1. Foursquare API calls are made to get the top 100 venues in a radius of 500 meters.
2. Get the subset of the venue categories related to Asian restaurants.
3. Check the top 5 venues in the different Manhattan neighbourhoods.

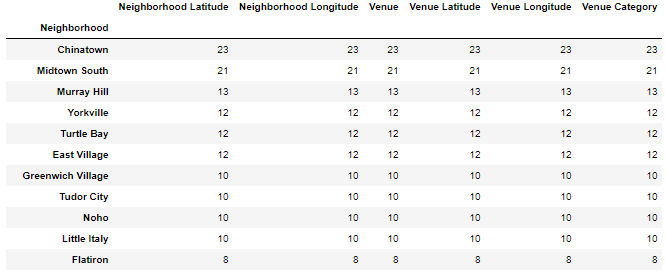
**TOP 5 VENUES:**

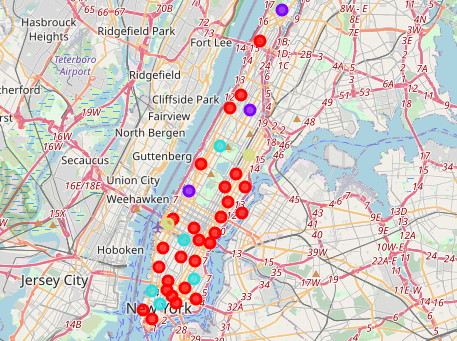
 

**Manhattan Neighborhoods with the most Asian Restaurants**



1. Analyze each neighbourhood by grouping the rows by neighbourhood and taking the mean of the frequency of occurrence of each venue category
2. Identify the best number of cluster for K-means clustering using silhouette\_score from sklearn.metrics.
3. Use K-means clustering algorithm to partition into 4 clusters.

**Visualization of 4 Clusters**

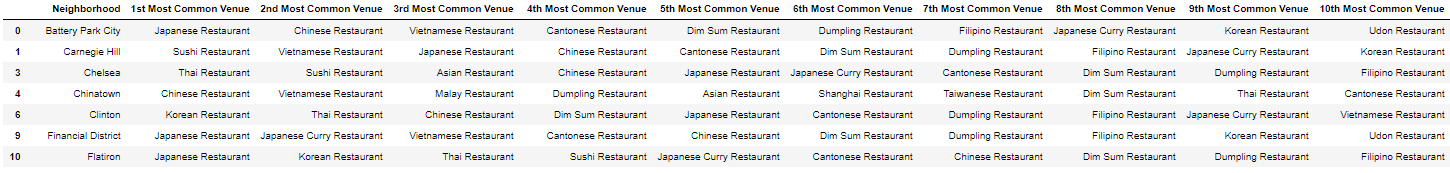


1. Compare the different clusters

## RESULTS

**Categorize the neighbourhoods into 4 clusters**

Cluster 0



Cluster 1



Cluster 2



Cluster 3



## DISCUSSION

* The project can target the top neighbourhoods with the most Asian restaurants.
* The top Asian restaurants per neighbourhood can give insights to the customization of the program
* Cluster 0 shows that the most common are Japanese restaurants
* Cluster 1 shows that the most common are Chinese Restaurants
* Cluster 2 shows that most common are Korean Restaurants
* Cluster 3 shows that the most common are Thai Restaurants

## CONCLUSION

We recommend to the program team to start with Chinatown and Midtown South where they can find most of the Asian restaurants.

The can cater the program to be more adaptable to the different clusters: Japanese, Chinese Korean and Thai.