**New Asian Recreational Food Hall & Venues Data Analysis of Queens**

Elena Jayne

November 18, 2019

**1. Introduction**

Asian elements become more and more popular in the business industry. Our business investors are eager to build a new recreational food hall in Queens, NY with an Asian theme. It's not only for Asians but also for attracting locals and all the people who like to spend money to enjoy the vibes with food and entertainment. A place for people to hang out with friends, to get drinks, desserts and various trendy food types. All in one special block. The problem for my investors is that they want to find a perfect neighborhood for them to invest. They need me to use data analysis to help them to find the location which have the most Asian customers who have the highest tendency to spent money to enjoy the recreational dining and relaxing at business place. Our target customers are those who would often spend money on high priced drink, bubble tea. The reason we decided to find out potential customers by using behavioral related spending pattern is that our business is targeting people whoever are interested in trendy food and comforting food such as Korean chicken wings, Hawaii poke bowl, Asian Street food, Vietnamese pho, noodle soup, brahmi, bubble tea, shaved ice and ice cream, Asian desserts, bar and drinks.

**2. Data**

The data I used in this project is obtained from <https://geo.nyu.edu/catalog/nyu_2451_34572>. It contains the data of all the neighborhoods and boroughs in New York City. It belongs to New York University Spatial Data Repository. The dataset is free and downloaded to the server.

Th data I obtained from New York University Spatial Data Repository is in json format. I ran a wget to load the dataset. I extracted the dataset contains 5 boroughs and the neighborhoods. All the relevant data was in the feature key, which is basically a list of the neighborhoods. After the “features” was defined, I transformed the data into a nested Python dictionaries into a pandas data frame. After all the data frame was created and filled with one row at a time, it showed the neighborhood has a total of 5 boroughs and 306 neighborhoods. Then I extracted the latitude and longitude coordinates of each neighborhood by using geopy library. In order to an instance of the geocoder, I defined a user\_agent and named it as “ny\_explorer”. Then the last step was to create a map of New York with neighborhood superimposed on top.

A picture containing text, map

Description automatically generated

Figure 1. Map of New York with neighborhood superimposed on top

Creating the map of the New York City, I used folium library. Folium is for visualization. I simplify the map and segment and clustered the neighborhoods in Queens. I sliced the original data frame and created a new data frame of Queens. Then I got the geographical coordinates of Queens only. The next process was utilizing the Foursquare API to explore the neighborhoods and segment them.

A picture containing text, map

Description automatically generated

Figure 2. Map of Queens with neighborhood superimposed on top

I defined Foursquare credentials and version. As I did with all of Queens. By exploring the first neighborhood in the data frame, which is Astoria. I got its latitude and longitude values. I also got the top 100 venues that are in Astoria within a radius of 500 meters. Using the Foursquare API and then I cleaned the json and structure it into pandas data frame. 100 venues were returned by Foursquare.

**3.Methodology**

I started from exploring the neighborhoods in Queens. I created new data frame called queens\_venues. It returned all the neighborhood names of Queens. I grouped all the venues by the neighborhoods they belong to. There are 270 unique categories. Then I created the new data frame and display the top 10 venues for each neighborhood.

A screenshot of a cell phone

Description automatically generated

Figure 3.

The following part was clustering the neighborhoods. Ran K-means to cluster the neighborhood into 5 clusters. Then I created a new data frame that includes the cluster as well as the top 10 venues for each neighborhood in Queens. I used folium to visualize the resulting clusters.

The last part was examining clusters. Determined the discriminating venue categories that distinguish each other. There are 5 clusters in total.

**4.Results**

**A close up of a map

Description automatically generated**

Figure 4.

Cluster 1 (label 0 red) has multiple neighborhoods, including commercial business area , restaurants, bank, pharmacy, grocery store, supermarket, and bakery.

Cluster 2 (label 1 purple) only has one neighborhood, Somerville, park area.

Cluster 3 (label 2 blue) has multiple neighborhoods, including residential area, deli, gym, bar, restaurant, bus station and bakery.

Cluster 4 (label 3 light green) only has one neighborhood, Jamaica Estates, including bus station and rich residential area.

Cluster 5 (label 4 orange) only has one neighborhood, Neponset, including beach and residential area.

**5.Discussion**

I tried to find the neighborhood which has bubble tea shop as its first most common venue. There was no return. Then I repeated this process four times until I got two targeted neighborhoods. One was Flushing and another one was Whitestone. Flushing has bubble tea shop as its third most common venue. Whitestone has bubble tea shop as its fourth most common venue. Interestingly after I did further research on google. It turned out both neighborhoods have large Asian populations.

**6. Conclusion**

The neighborhood in Queens has the most bubble tea shops is Flushing. The second highest common venue for bubble tea shop is Whitestone. However, Flushing is in cluster 1(cluster label 0 red color) commercial area, Whitestone is in cluster 3(cluster label 2 blue color) residential area. What we are looking for is a commercial oriented neighborhood. Therefore, our new Asian recreational food hall should be placed in Flushing since it is a neighborhood which has most of the customers who are the same targeted customers of bubble tea shops.