

# **The Battle of Neighborhoods: Greater Boston Area**

Thu Le  
July 30, 2021

## **1. Introduction/Business Problem**

### **1.1 Background**

The United States has been the destination for immigrants around the world. More than 100 years ago, Chinese Americans settled in Boston, Massachusetts and built a community in Chinatown. Since then, there has been a large Asian American community in Greater Boston. Even though asian have moved to many other parts of Massachusetts, most of Asian American still live in suburban towns near Boston. And the asian immigrants have made an influence over american's food. As a result, there are a lot of asian restaurants in Massachusetts, each belonging to different categories such as chinese cuisine, japanese cuisine, vietnamese cuisine, taiwan cuisine, etc.

### **1.2 Problems**

In this project, we will use Foursquare API to analyze different neighborhoods and find a good location to open a bubble tea business. With major bubble tea franchises such as GongCha, Kungfu tea,... it is inevitable that there is a growing demand for this beverage. We look for a location which is closer to a lot of asian restaurants, but also the neighborhood should not be crowded with bubble tea shops to avoid competition.

Greater Boston Area was chosen in this project due to the following reasons:

- Sixteen cities and towns in Greater Boston in 2016 had at least 4,000 Asian Americans. Many asian american are still living close to the suburban Greater Boston Area.
- Massachusetts is a destination for international students, in which China and India were the leading countries of origin for international students. Therefore, there is an ever-growing demand for asian restaurants.

### **1.3 Interest**

Obviously, this project would be useful for investors who are interested in opening a bubble tea shop, or asian coffee shop, or even a particular asian restaurant such as vietnamese restaurant, thai restaurant, etc. Not only so, international students or travellers will be interested in this analysis if they want to find a neighborhood full of asian food.

## **2. Data acquisition and cleaning**

## 2.1 Data sources

Massachusetts data that contains a list of states, counties, cities, zip code, latitudes, longitudes and estimated population can be found on <https://data.healthcare.gov/resource/52wv-g36k.json>. Some of the counties in this dataset was missing; therefore, we have to look for the missing data in another dataset - “List of Massachusetts area code” downloaded from [https://www.downloadexcelfiles.com/us\\_en/download-list-us-zip-codes-massachusetts-state#.YQNVVo5Kg2w](https://www.downloadexcelfiles.com/us_en/download-list-us-zip-codes-massachusetts-state#.YQNVVo5Kg2w).

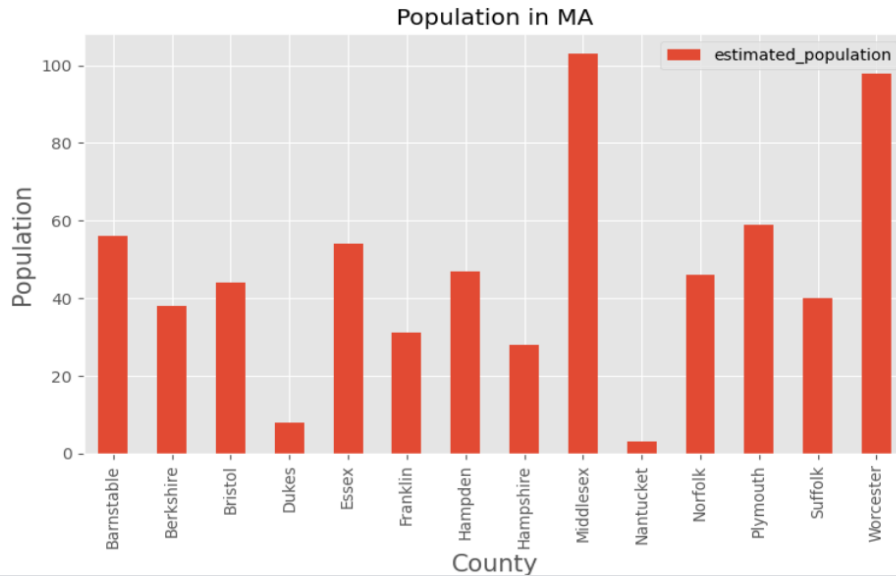
The next dataset is the asian restaurants and bubble tea shops in “Suffolk” county, where Boston - the capital and most populous city of Massachusetts - belongs to. The latitude and longitude of the cities in Suffolk counties will be passed through Foursquare; and as result, Foursquare API will return the venues in Suffolk county. We will then filter the asian restaurants and bubble tea shops for our purpose. We will also use mapping visualization and look for the neighborhood with dense asian restaurants but this neighborhood will not be crowded with bubble tea shops to avoid competition.

## 2.2 Data cleaning

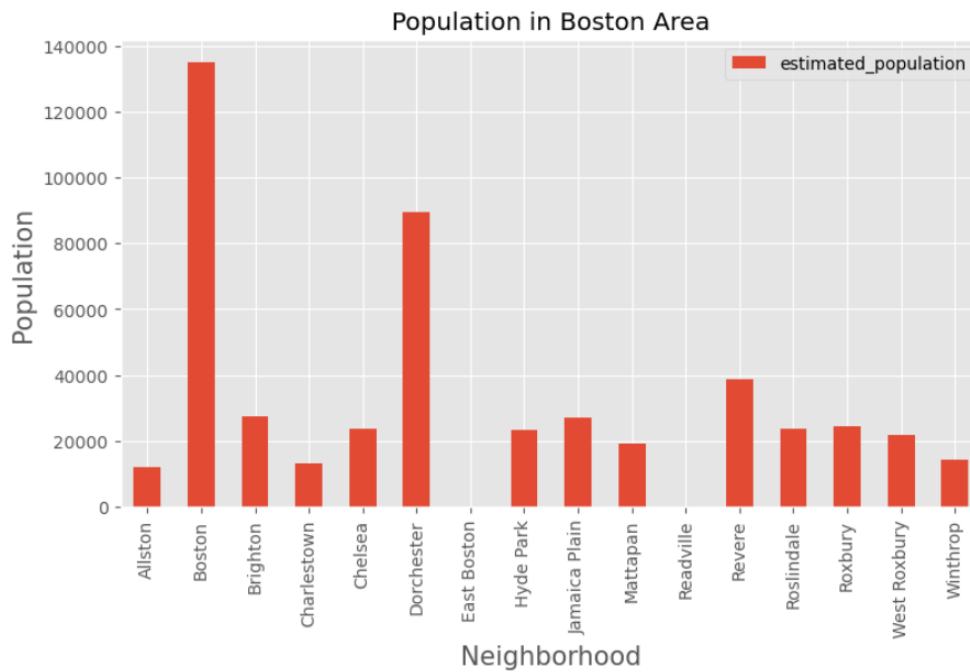
Data downloaded or scraped from multiple sources were combined into one table; therefore, there are some missing data which need cleaning. Moreover, some venues were not correctly classified; for example, some asian restaurants were categorized as “restaurants”, or some bubble tea shops were categorized as “desserts”. Foursquare returns the findings based on a specific latitude and longitude in the center of the neighborhood, causing some missing data.

## 3.Explanatory Data Analysis

### 3.1. Population of Neighborhood in Massachusetts and Suffolk County



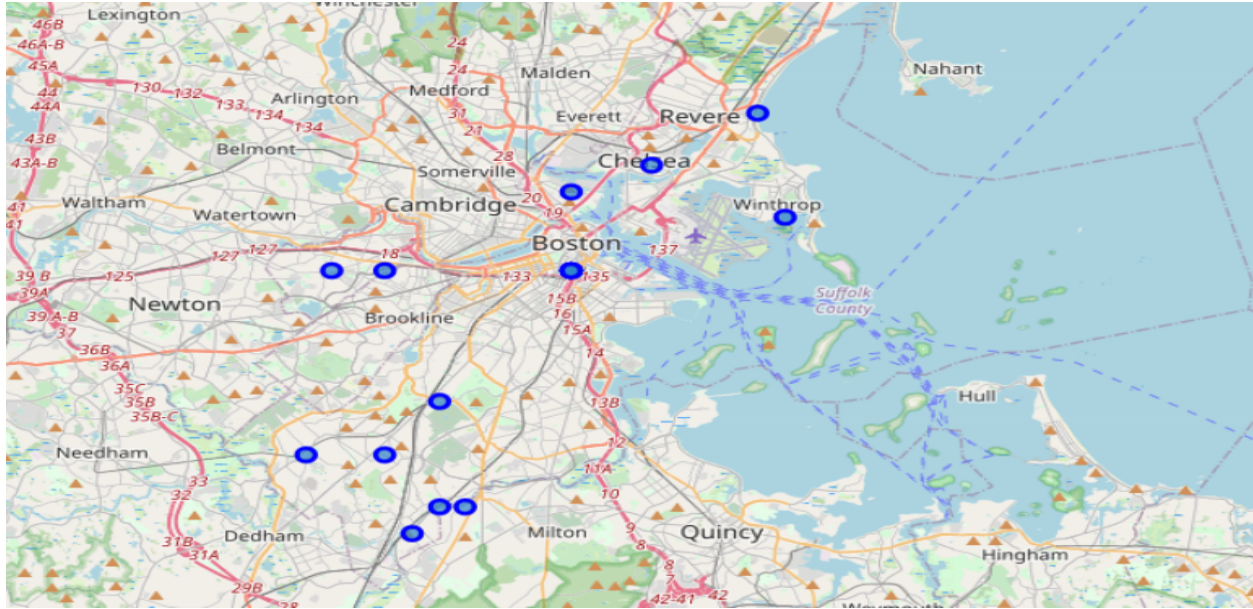
Suffolk is not the most populous county in Massachusetts; instead, Worcester and Middlesex are the most populous ones.



Meanwhile, Boston is the most condensed area; followed by Dorchester, Revere, Roxbury and Brighton.

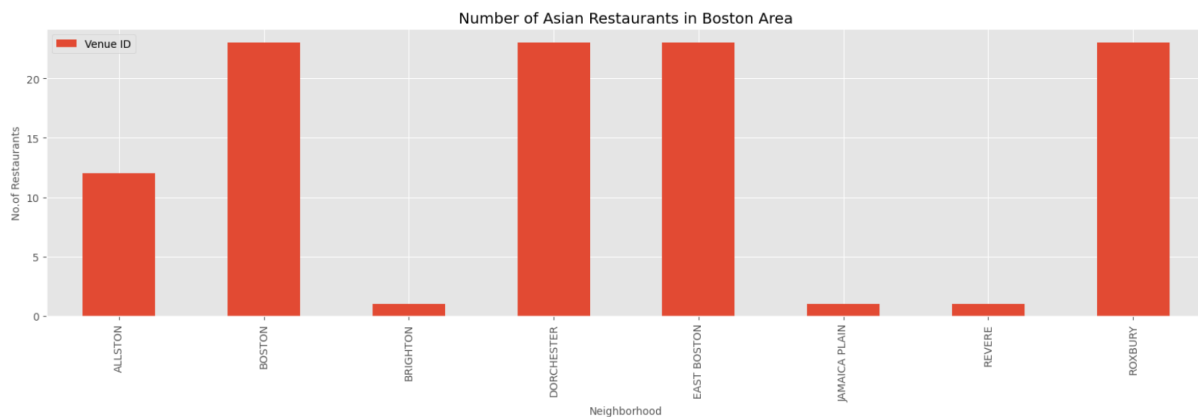
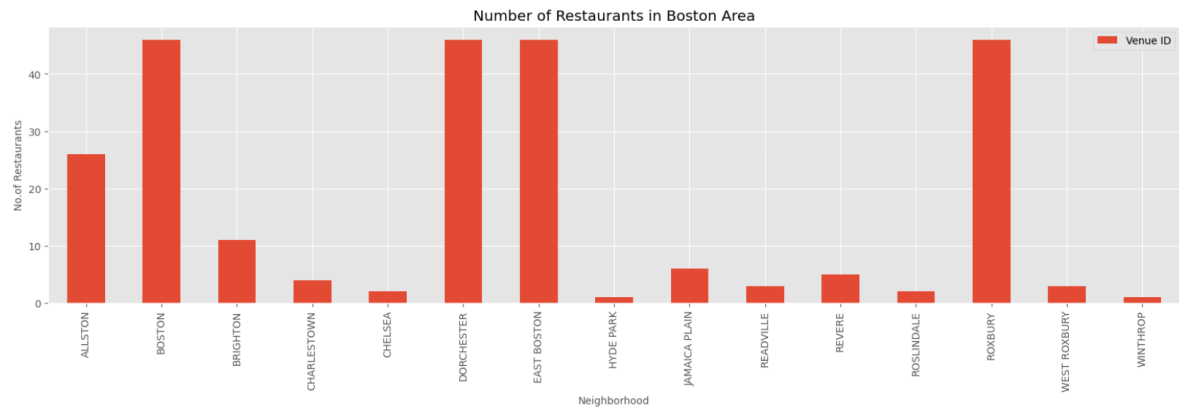
### 3.2 Zip Code Distribution Across Suffolk County

As we can see from Follium, the zip code around Suffolk county is not close. This may raise a concern that the findings will be returned for another neighborhood.

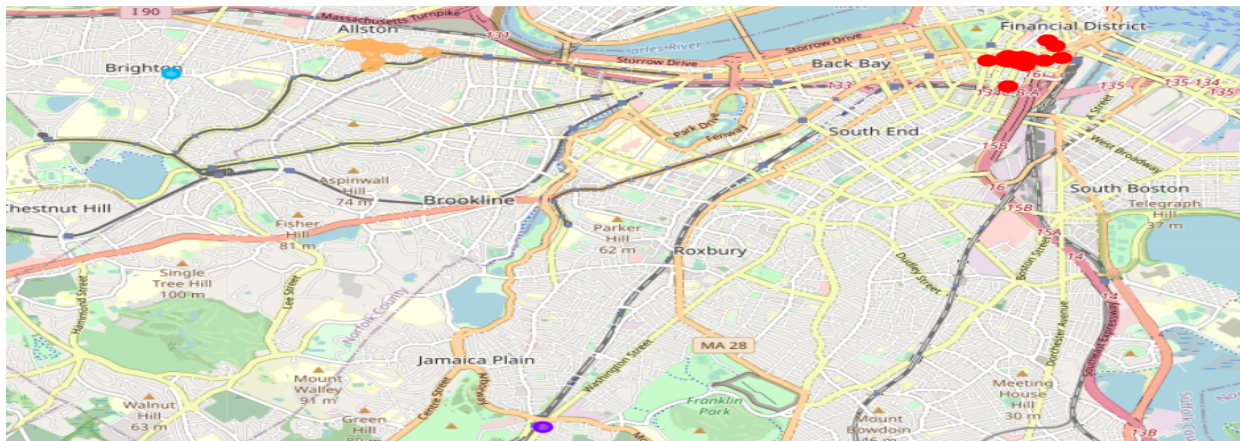


### 3.3 Foursquare Venue Data

Using Foursquare to obtain the venue data and plot the graph, we find that Dorchester, Boston, East Boston and Roxbury have the largest number of restaurants as well as asian restaurants in Suffolk County. Allston is the fifth one on the list.

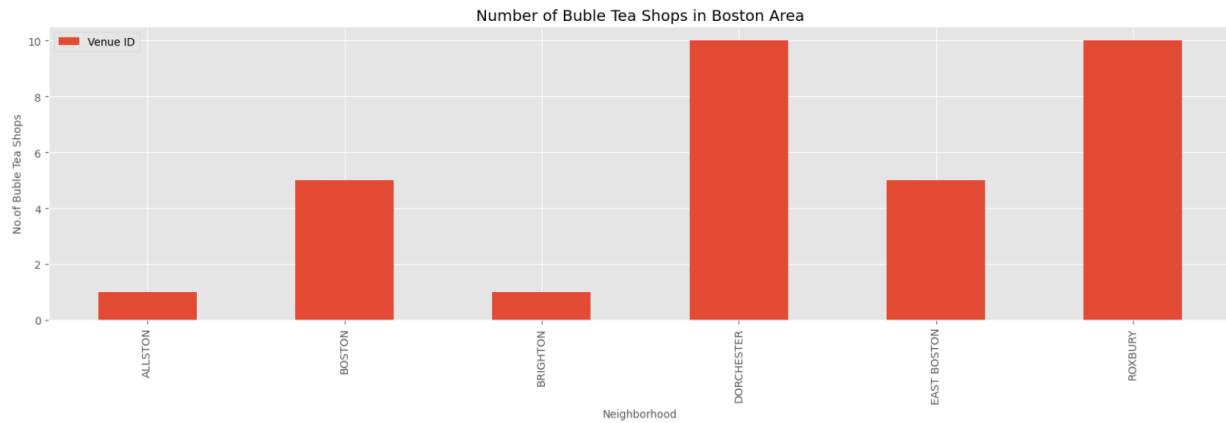


From those two graphs, we now focus on only 5 cities: Boston, Dorchester, East Boston, Roxbury and Allston in order to find a suitable location for the bubble tea shops. Next, we create a dataframe that includes the frequency cluster as well as the top 10 venues for each neighborhood in Suffolk county, and visualize the resulting clusters. The map proves that Boston (in red) and Allston (in yellow) have the most common restaurant venues.



Until now, we have two locations: Boston and Allston to decide whether we should open the bubble tea store. This would be confirmed with the graph for Bubble Tea Shops in Suffolk County. It shows that Boston, Dorchester, East Boston and Roxbury are crowded with Bubble Tea Shop. Meanwhile, Allston and Brighton each has only small number of tea shops. This concludes our analysis. In order to avoid aggressive competition, Allston would be the best choice we should pick for a bubble tea shop.

|              | Neighborhood Latitude | Neighborhood Longitude | Venue ID | Venue | Venue Latitude | Venue Longitude | Venue Category |
|--------------|-----------------------|------------------------|----------|-------|----------------|-----------------|----------------|
| Neighborhood |                       |                        |          |       |                |                 |                |
| ALLSTON      | 1                     | 1                      | 1        | 1     | 1              | 1               | 1              |
| BOSTON       | 5                     | 5                      | 5        | 5     | 5              | 5               | 5              |
| BRIGHTON     | 1                     | 1                      | 1        | 1     | 1              | 1               | 1              |
| DORCHESTER   | 10                    | 10                     | 10       | 10    | 10             | 10              | 10             |
| EAST BOSTON  | 5                     | 5                      | 5        | 5     | 5              | 5               | 5              |
| ROXBURY      | 10                    | 10                     | 10       | 10    | 10             | 10              | 10             |



#### 4. Result and Discussion

Our analysis shows that there is a large number of asian restaurants in Suffolk County. After removing some missing data retrieved from Foursquare, it showed that there are 153 asian restaurants in this county. Highest density of asian restaurants are located in 5 cities: Boston, Dorchester, East Boston, Roxbury and Allston. The location candidate - Suffolk County- was then clustered to create zones of interest which contain the most common venues. It is interesting to note that Allston is not as populous as the other 4 cities, but the number of asian restaurants is not small. This could be hypothesized that there are a lot of asian, especially international students living in Allston and nearby city - Brighton, because there are some big universities in this nearby neighborhood such as Boston University or Boston College; etc. Meanwhile, only 4 cities - Boston, Roxbury, East Boston and Roxbury have quite a significant number of bubble tea shops. Therefore, Allston will be a good location for investors to open the bubble tea stores. With low competition and high density of asian restaurants, it is inevitable that opening a bubble tea business will bring great profit in future; because bubble tea would be a good dessert place to visit while waiting for the restaurant's seating or after the meals. Recommended zones should be considered as a starting point for a more detailed analysis. In a more detailed analysis, we could include other factors such as proximity to public transportation, ratings of the asian restaurants and bubble tea shops, or real estate availability; etc.

#### 5. Conclusion

Purpose of this project is to identify a location in Suffolk county with high density of asian restaurant and low density of bubble tea shops in order to aid the stakeholders to find a location for a new tea shop. We choose Suffolk county for analysis because asian americans and international asian students are mostly living in the suburban areas of the Great Boston Area. By calculating the density of asian restaurants and bubble tea shops using Foursquare and visualization from Follium map, we identify Boston and Allston as our favourite locations. However, even though Boston is the most populous city of Suffolk Area, if we open a new tea shop in Boston, we have to compete with the fiercely big bubble tea franchises such as GongCha

or Kungfu Tea. Therefore, Allston is still a better location for the stakeholders to open a tea shop.