1. **Introduction**

## Background

Beijing is the capital, and one of the four direct-controlled municipalities of the People's Republic of China. Beijing is governed as a municipality under the direct administration of the central government with 16 urban, suburban, and rural districts (boroughs).

China’s house prices rise sharply during the last decades, especially in Beijing. The house price in Beijing is mounting to astronomical figures and shows no signs at present of ceasing to rise. However, since Beijing is the center of China, it has over 21 million residents within an administrative area of 16,410.5 km2. In 2016 the Global Cities Business Alliance said Beijing was the most expensive city to rent an apartment in, with average monthly rent costing 1.2 times the average monthly salary. The housing problem in Beijing is severe.

## Problem

For individuals working in Beijing, they need to buy or rent an apartment far from their working place since the house price in center are extremely expensive. Some of them even choose to live in rural districts of Beijing, as if they are accessible by public transport. Thus, knowing the price map of houses is important.

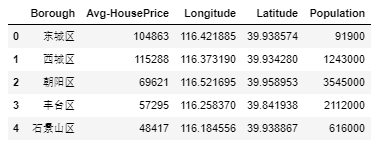
Moreover, for an investor, they prefer the districts where there is a lower real estate cost, dense population, and the type of business they want to install is less intense.

When we consider all these problems, we can create a map and information chart where the real estate index is placed on Beijing and each district is clustered according to the venue density.

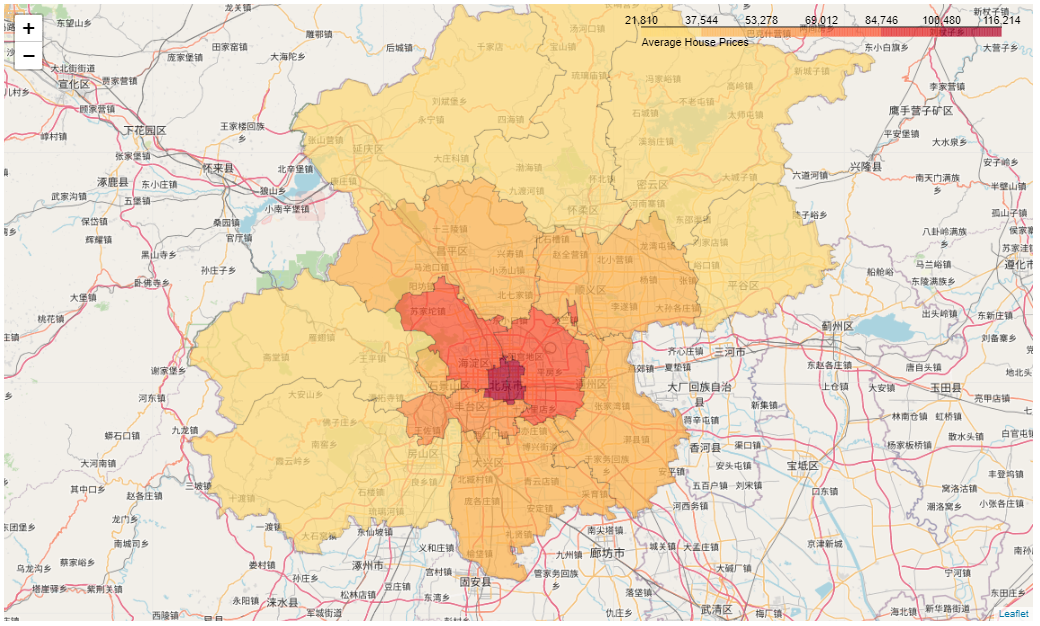
# Data acquisition

The data of administrative division of the Beijing is from the following Wikipedia page, <https://en.wikipedia.org/wiki/List_of_administrative_divisions_of_Beijing>. I used Forsquare API to get the most common venues of given district of Beijing. I collected latest per square meter Housing Sales Price (HSP) Averages for each district of Beijing from housing retail web page, <https://newhouse.fang.com/house/fangjia/>.

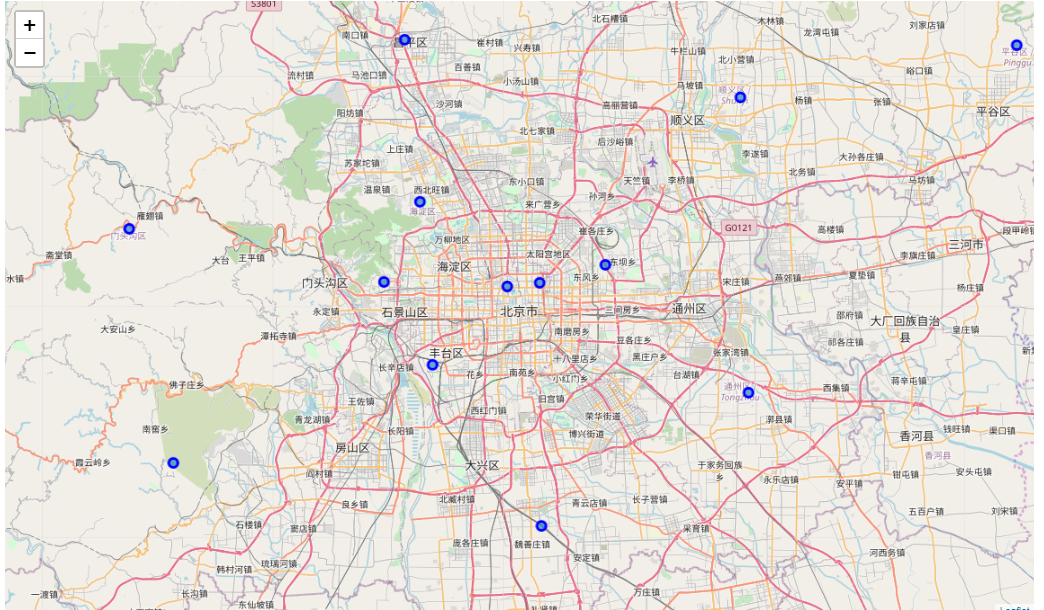
My master data which has the main components Borough, Average House Price, Latitude and Longitude and Population information of the city.



1. **Methods**

I used python folium library to visualize

I used python folium library to visualize geographic details of Beijing and its boroughs and I created a map of Beijing with boroughs superimposed on top. I used latitude and longitude values to get the visual as below:

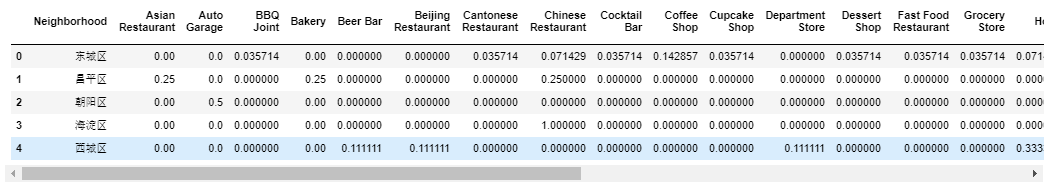


I utilized the Foursquare API to explore the boroughs and segment them. I designed the limit as 50 and the radius 500 for each borough from their given latitude and longitude information. Here is the head of the merged table of boroughs and venues.



There are 29 unique categories which can be curated from all the returned venues.

Here is the mean of the frequency of occurrence of each category.



Here shows the list of top 10 venue category for each borough.

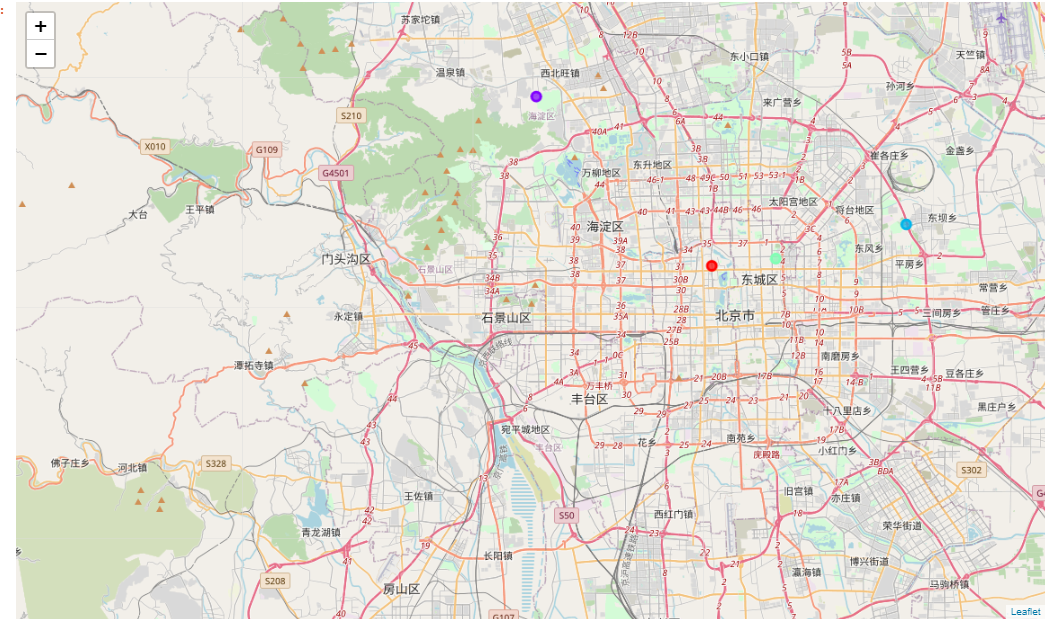
There are some common venue categories in boroughs. In this reason I used unsupervised learning K-means algorithm to cluster the boroughs.

Here is the merged table with cluster labels for each borough.

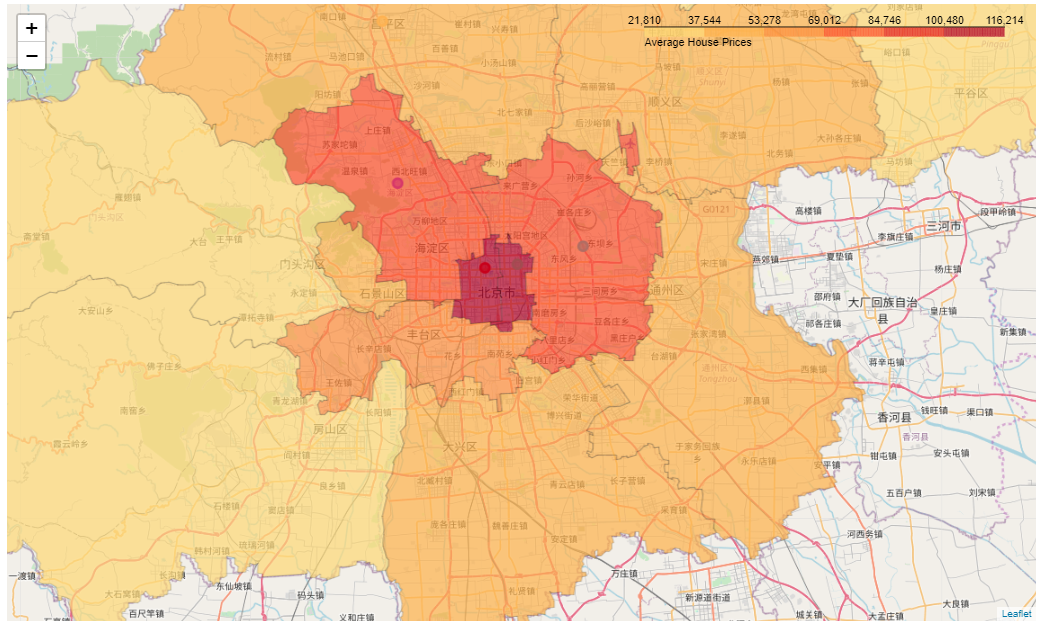


# Results

Here is the clustered map of boroughs of Beijing.



Here is the choropleth map containing the following information for each borough: borough name, cluster name, house prices, and the top 3 number of venues.



# Discussion

As can be found from the tables and maps above, the house price in the center of Beijing is extremely high. And the business activities in these areas are prosperous. If an investor wants to start his or her business in Beijing, it is better to choose these areas with higher population density and prosperous business activities though the house price is much higher.

# Conclusion

Individuals and investors can use these data to make their decisions.