

# Searching a Location in New York City for a Chinese Restaurant

Jiang Li

## 2. Data acquisition and description

### 2.1 Data sources

The geographic data of NYC is from the IBM Developer Skills Network. This dataset contains the 5 boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood. The second dataset is about the demographics of NYC, which is introduced by Wikipedia at [https://en.wikipedia.org/wiki/Demographics\\_of\\_New\\_York\\_City](https://en.wikipedia.org/wiki/Demographics_of_New_York_City). Finally, the third data set about the property prices in NYC can be found in Kaggle datasets at <https://www.kaggle.com/new-york-city/nyc-property-sales>.

### 2.2 Description

The geographic dataset from the IBM Developer Skills Network is used for locations visualizing with Folium package, and also used for locations searching and exploring with Foursquare API. The demographic dataset from Wikipedia is used to analyze the population distribution and composition of NYC in order to choose a suitable location with a large number of potential customers to open a Chinese restaurant. In order to reduce the cost of opening a Chinese restaurant, we must consider the issue of housing rent. The Kaggle dataset shows the property sales in NYC and this dataset is used to estimate the housing rent of a location in NYC. These three datasets are used and analyzed together to search a proper location for opening a Chinese restaurant where is convenient and accessible for most people especially for the Asian people living in NYC and is also with acceptable housing rent.