**Introduction**

When you come to Shanghai, you must visit the French concession, a variety of garden houses, tall plane trees on both sides of the road, and cafes and bars hidden in every corner. However, only a few well-known streets are known to everyone, and other streets with similar styles may not be known to everyone. Therefore, by clustering similar streets, more tourism information can be provided for tourists

**Data**

To consider the problem we can list the datas as below:

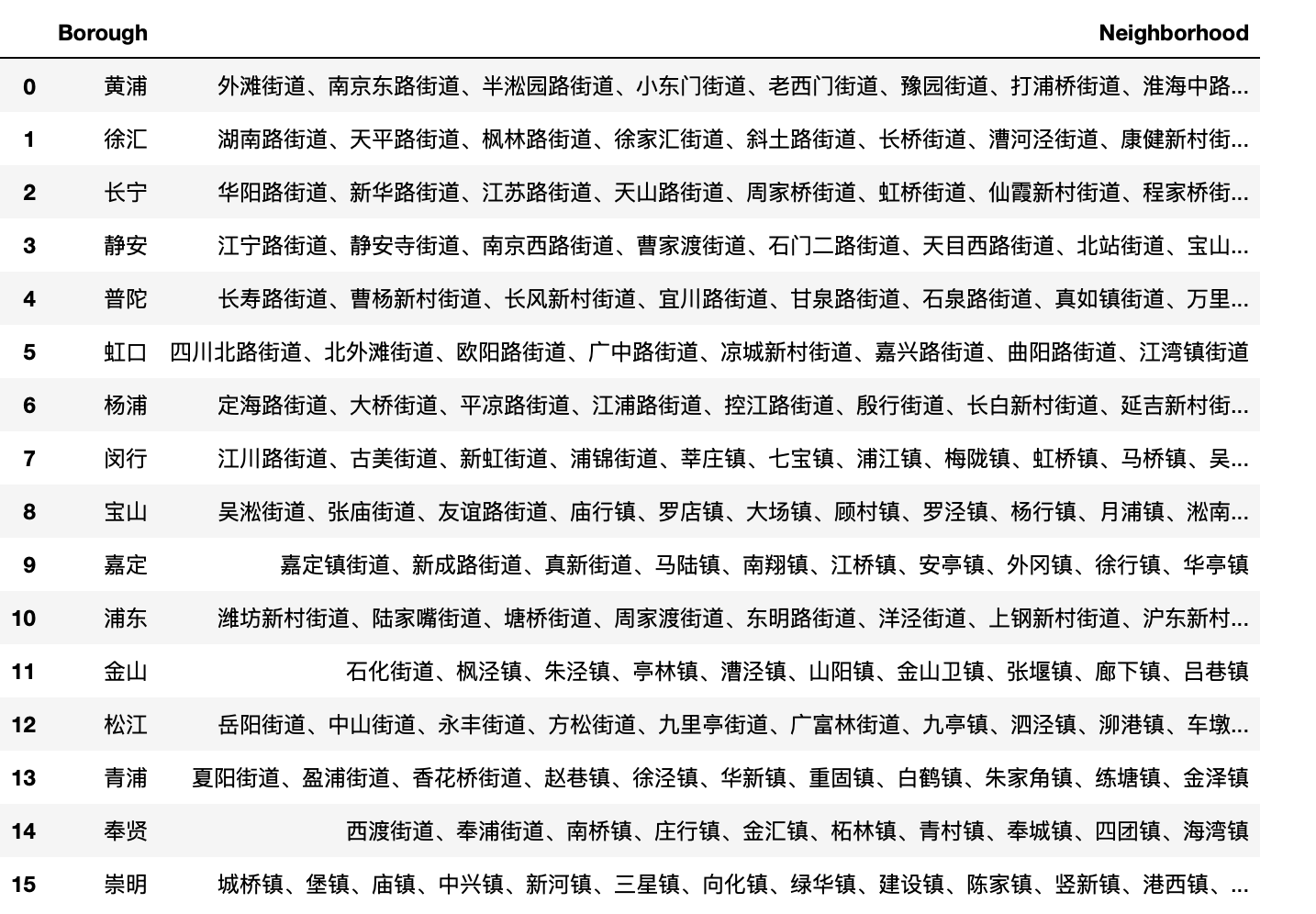
* List of neighbourhoods in Shanghai. This defines the scope of this project.
* Latitude and longitude coordinates of those neighbourhoods. This is required in order to plot the map and also to get the venue data
* Venue data, particularly data related to shopping malls. We will use this data to perform clustering on the neighbourhoods

**Sources of data and methods to extract them:** Shanghai civil affairs bureau page contains a list of neighbourhoods in Shanghai, with a total of 106 neighbourhoods. We will use web scraping techniques to extract the data from the, with the help of Python requests and beautifulsoup packages. Then we will get the geographical coordinates of the neighbourhoods using Python Geocoder package which will give us the latitude and longitude coordinates of the neighbourhoods. **(web scraping)**

After that, we will use Foursquare API to get the venue data for those neighbourhoods.  **(working with API)**

Foursquare API will provide many categories of the venue data in order to help us to solve problem put forward **(data cleaning, data wrangling, to machine learningand map visualization)**

## Methodology

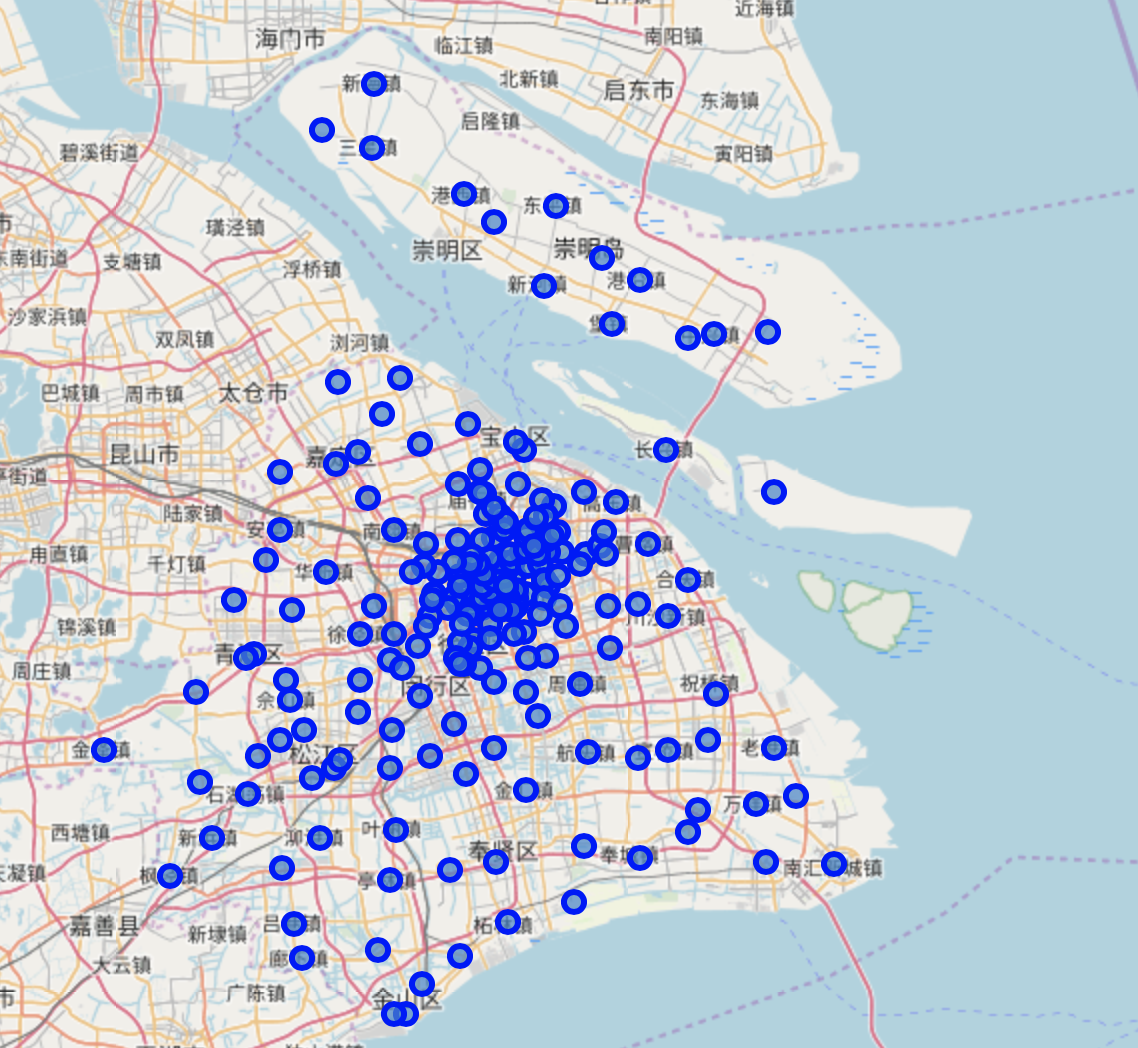
1. **Scrap data from page into a DataFrame**
2. **Use geocode API to get the latitude and longitude values**

****

1. **Merging**

****

1. **Create a map of Shanghai with neighborhoods(township-level divisions) superimposed on top**

****

1. **Filter only boroughs at the center of city**

****

1. **Use the Foursquare API to explore the neighborhoods**

****

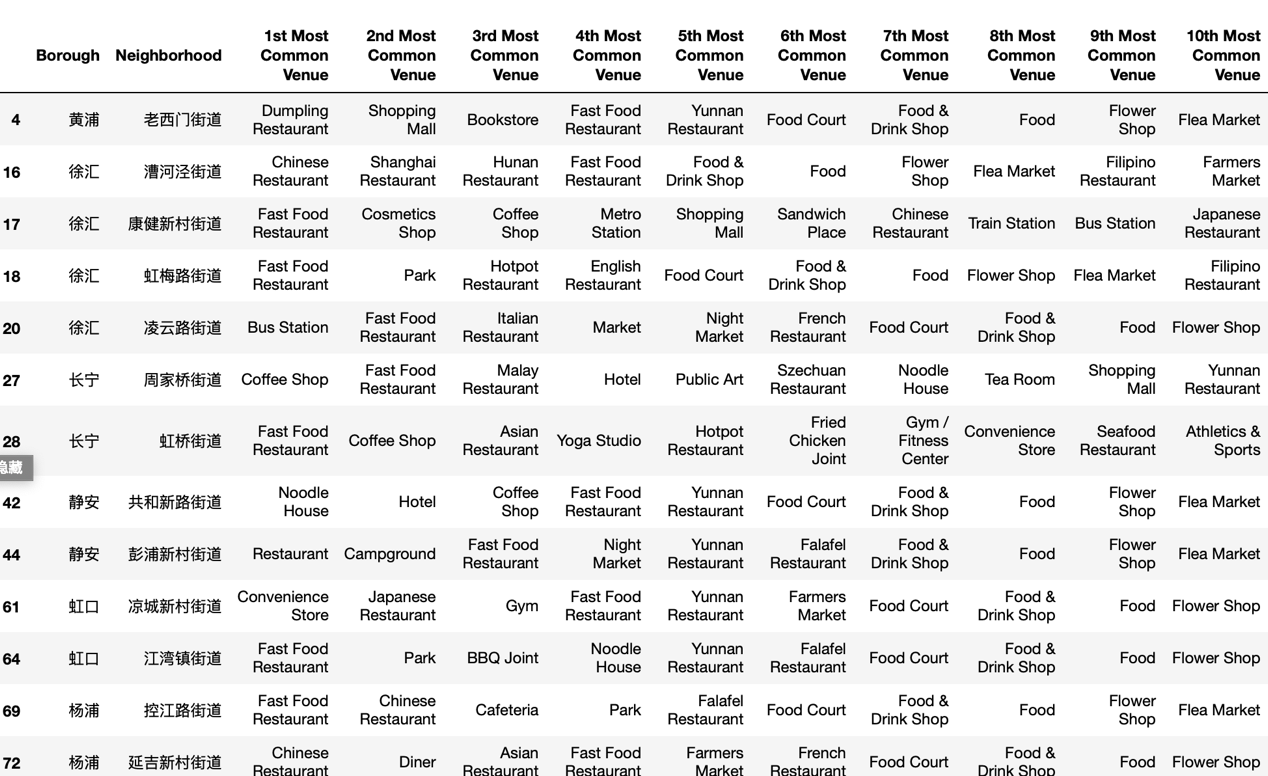
1. **Analyze Each Area**

****

**Result & Discussion**

****

****

****

**We can see that in the Cluster 3, the neighborhood style is closer to the French concession style, which means that in addition to the French concession, you can visit these places at the same time**

*At the same time, we find the Cluster 1, the community style is close to the traditional Chinese culture area, and we can go to these places to understand the living conditions of ordinary Shanghai residents*