**The Battle of Neighbourhoods**

## 1. Introduction

In this project we will try to find an optimal location for weekend activities. Specifically, this report will be targeted to people who are newly moved to **Shanghai**, China.

**The location should have many choices of restaurants, gyms, coffee shops or bars and shopping possibilities**. Except that there should be possibilities to go **clubbing** in the night and have **metro station nearby** for good transportation.

This is an interesting task to analyse with valid questions for anyone newly moving to Shanghai. The same method can be applied to explore other big cities as well. This case is also applicable for anyone interested in finding a new apartment to rent.

## 2. Data

List of neighbourhoods of Shanghai with their geodata (latitude and longitude)

Venues for each Shanghai neighbourhood (then can be clustered for restaurants, bars, cinemas etc.)

location of nearby subway metro stations, as needed

The data will be used as follows:

Use Foursquare and geopy data to map top 20 venues for all neighbourhoods and clustered in groups

Use foursquare and geopy data to map the location of subway metro stations

These Data will answer the key questions to decide: which neighbourhoods contain all requirements (restaurant, bar, cinema, shops and clubs)? and find out which metro stations are nearby?

## 3. Methodology

In this project we will cluster each neighbourhood and limit our analysis to area ~5km from the neighbourhood center.

In first step we have collected the required data: location and top within 6km from neighbourhood center.

In next step we will focus on clusters of locations: We will create clusters (using k-means clustering) and map the clusters. Afterwards we will have a dataframe to explore the most common venues in each neighbourhood from each cluster.

## 4. Results

The total analyzed neighborhood in Shanghai:



Neighborhood after clustering:



The cluster ‘0’ – target cluster of interest:



## 5. Discussion

Our analysis has identified top 20 venues in each neighbourhood and then clustered neighbourhood in 8 clusters. By Looking into each cluster, we found that cluster one fulfilled our needs for looking for an area for weekend activities. The top venue in this cluster contains restaurants, coffee shops, shopping malls, and bars for late night activities. In both neighbourhoods from Pudong New Area District, there are metro stations nearby (within 6km), which narrows down our search.

## 6. Conclusion

Purpose of this project was to identify in areas in Shanghai for weekend activities. Area with couple restaurants, bars, shopping malls and metro station. By identifying top venues 6km from neighbourhood center using Foursquare data and then clustering each neighbourhood, we have identified first the Pudong New Area District as one of the top Borough. And Xintiandi and Zhangjiang Town as our promising candidate neighbourhood.

Final decision on optimal location will be made by stakeholders based on specific characteristics of neighbourhoods and locations in every recommended zone, taking into consideration additional attractive venues such as museums, art galleries, theme park or gyms etc.