Where to open a new stadium in Shanghai, China

Coursera Project: Capstone

By: Zhe Chen
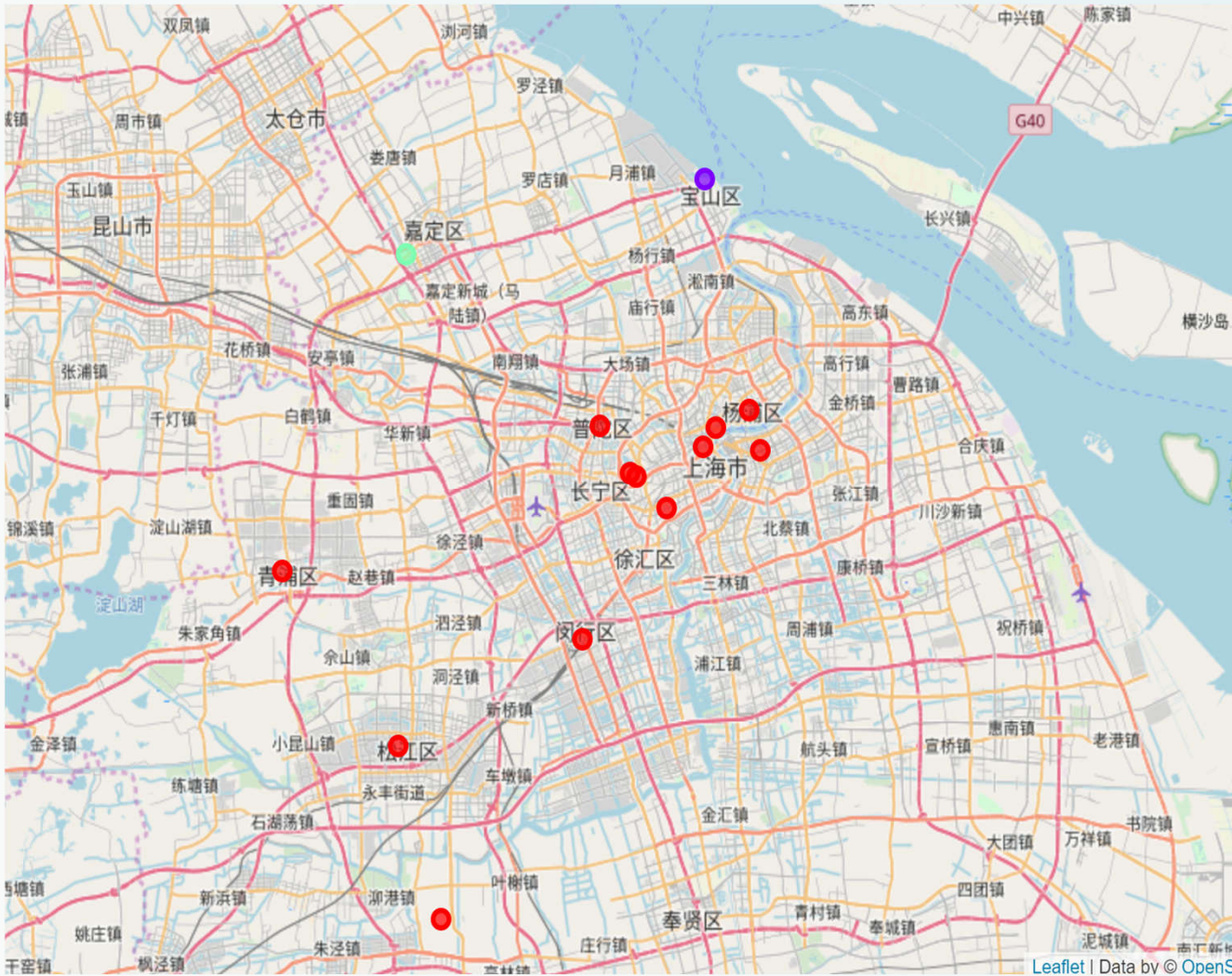
Jan. 2020

0 1 DATA

- List of districts in Shanghai (Wikipedia)

- The coordinates of these districts (Geocoder package)

- Venue data related to the current stadium (foursquare API)

Methodology

- Web scraping Wikipedia page for districts list
- Get latitude and longitude coordinates using Geocoder
- Use Foursquare API to get venue data
- Group data by district and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by "stadium"
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

Results

- Cluster 0: Districts with low number of stadiums

- Cluster 1: Districts with moderate number to no existence of stadiums

- Cluster 2: Districts with high concentration of stadiums

- one stadium sitting in the Baoshan district and there're two in the Jiading district
- The rest districts have zero.
- According to calculation, cluster 0 has no stadium
- The stadiums in cluster 2 are likely meet the residents' demand in that district

0 5 Recommendations

- The stadium developers to open new stadiums in the districts in cluster 0.

- The stadium developers are advised to avoid the districts in cluster 2.

THANK YOU FOR LISTENING