

# Capstone Project

## Introduction

Nowadays, choosing optimal locations while opening commercial sites lays a critical role of whether the business would success. Stakeholders are trying to find the best location in the city to locate their businesses in order to attract more targeted consumers and to gain more revenue. Shanghai is one of the four municipalities in China, and it's the second most populous urban area in the world. Because Shanghai is a global center for finance, more and more stakeholders are attracted to expanding their businesses in Shanghai. However, in order to make sure the location of a commercial site won't have negative impact on growth of revenue, it's critical to optimize a location that attracts the most customers and has the best opportunity of growth.

This project will try to find an optimal location (district) to open a shopping mall in Shanghai, China. Since there are so many shopping malls, detecting locations that are not already crowded but still attractive is the main task here. By analyzing data and visualizing maps, I'll come up with two to three promising district areas based on the criteria. The results would benefit the stakeholders who seek opportunity to open a shopping mall in Shanghai.

## Data

Based on above descriptions, the data will be used are:

- Geographical coordinates of districts in Shanghai
- Population and density in each district
- Number of shopping malls and their locations in each district

In order to obtain the data, following data sources will be needed:

- Wikipedia category website
- Beautiful soup python library
- Geolocator library
- Foursquare API