APPLIED DATA SCIENCE CAPSTONE

Project: The Battle of Neighborhoods

Opening a new shopping mall in Beijing

Yvette Lu

2020/10/03

Introduction

A large shopping mall is undoubtedly a good place for entertainment. In the mall, people can relax and enjoy themselves. Within a short walk, they can reach a variety of shops, such as grocery stores, restaurants, cinemas, supermarkets, fruit shops and even indoor amusement parks. Shopping centers can meet the needs of many types of consumers. However, the location of a shopping center is very important for the profit of a store. It should be located in a prominent and convenient place where there is a lot of people. Although there are many concentrated and huge shopping centers in the center of Beijing, there are still many places where residents need shopping centers closer to them. Besides, the location of shopping malls plays an important role in the profits of dealers and real estate developers.

● Business Problem

The goal of the project is to use a data-science approach, using various algorithms, such as clustering, to select a location in Beijing for real estate developers or overseas investors to build a new shopping mall.

● Who would be interested

In recent years, shopping malls in Beijing have been mushrooming, and their development has been stable and orderly. This is due to the strong strength and advanced concepts of developers, as well as the superior location of site selection, and the continuous optimization and adjustment of developers by relying on large-scale residential communities. In the words of Jianhui Han, the vice President of China Urban Commercial Network Construction management Federation, Beijing's commercial project status could be characterized by 'rapid development, fierce competition, survival of the fittest and wide traps' [1]. Some industry watchers said, Although commercial properties are relatively concentrated in the urban center of Beijing, the economic development is stable and mature, old projects are being renovated and upgraded, and the sub-center of the city is becoming more and more perfect, with the feature of multi-center becoming increasingly obvious.

Commercial property still has a broad space for development and is still the focus of many large commercial real estate developers and brand retailers[2].

Data

● Description of data

We might need the following data to solve the problem above:
1. A list of neighborhoods in Beijing, especially the prominent neighbors (In this project, Ethnic enclaves might not be included).
2. The latitude and longitude coordinates of these neighborhoods.
3. The site data necessary related to the mall.

- **How to use the data**

    A list of major neighbourhoods in Beijing, including neighbors like Qianmen and Sanlitun, etc. could be crawled through the Wikipedia page [3]. There are 41 communities in total.

    By using the beautifulsoap package, we can crawl data from the Wikipedia and use the Python Geocoder package to get the geographical coordinates of our neighbours, which can give us the latitude and longitude coordinates of our neighbours.

    After that, we'll use Foursquare APl to get the location data for these communities. By using this API, we can get the data we want, especially the mall data, then data cleansing and machine learning techniques would be applied, then it's time to visualize the data using Folium.

References

[1] Chu Qian, From "Wolf" to "cat" -- Ten years of Changes in Beijing shopping Malls [J].China Market, 2008(4):36-36.

[2] Ninghui Chen, List of Shopping Centers to be opened in Beijing 2020: Dongba Wanda Plaza, Joy in the Spring Breeze... Lianshang Retail research Center. *http://www.linkshop.com.cn/web/archives/2020/444113.shtml*

[3] Wikipedia page: *https://en.wikipedia.org/wiki/Neighborhoods_in_Beijing*