

Coursera Capstone IBM Applied Data Science Capstone

Opening a New Coffee shop around Universities in Beijing

By: Ruiqi Su

July 14th 2020

Business Problem

Background

- Although Coffee is not that main beverage in China, but for modern society, Coffee shops is still an important kind of fast-retailing in big city.
- My undergraduate studying was in China University of Mining and Technology, Beijing. Most of Coffee shops here are Starbucks, Costa, and they are too expensive for our students. So this project is assuming that I am going to invest a Coffee shop or Chain brand Coffee for students in this area.

Business Question:

- The objective of this capstone project is to analyse and select the best locations in the city of Beijing near Universities to open a new coffee shop.
- This project aims to provide solutions to answer the business question: In Beijing and region near Universities, if I am looking to open a new coffee, where is the best place?

Data

Data Reuired

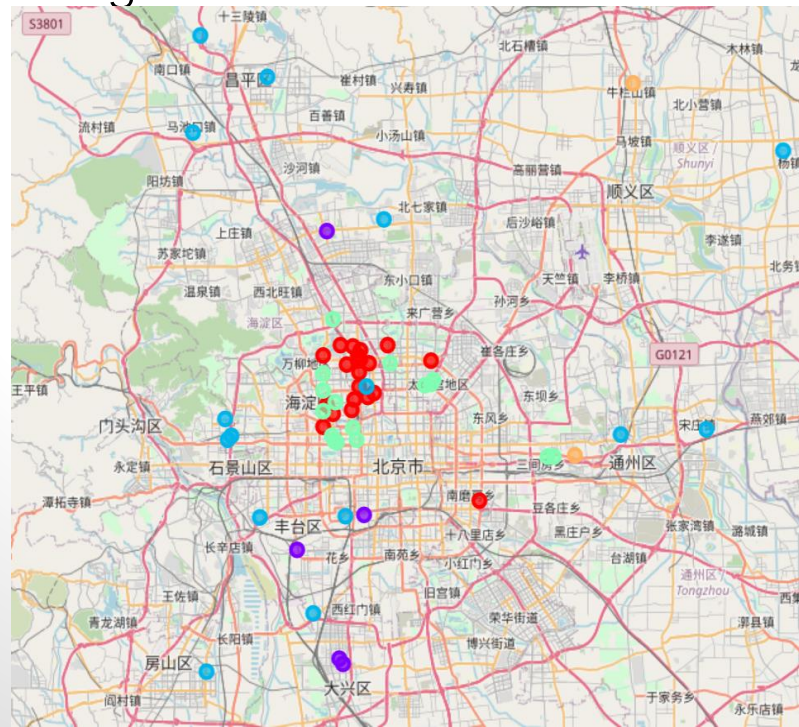
- List of Universities in Beijing. This defines the scope of this project which is confined to the city of Beijing, the capital of P.R.China.
- Latitude and longitude coordinates of those Universities. This is required in order to plot the map and also to get the venue data.
- Venue data, particularly data related to Coffee shops. We will use this data to perform clustering on the neighbourhoods
- **Business Question:**
 - Sources of data Wikipedia page for neighbourhoods (https://en.wikipedia.org/wiki/List_of_universities_and_colleges_in_Beijing)
 - GaoDe API Geofor latitude and longitude coordinates
 - Foursquare API for venue data

Methodology

- Web scraping Wikipedia page for neighbourhoods/Universities list
- Get latitude and longitude coordinates using Gaode API
- Use Foursquare API to get venue data
- Group data by neighbourhood and taking the mean of the frequency of occurrence of each venue category
- Filter venue category by Coffee Shop
- Perform clustering on the data by using k-means clustering
- Visualize the clusters in a map using Folium

Results

- Cluster 0: Neighbourhoods with moderate number of Coffee Shop
- Cluster 1: Neighbourhoods with relatively high concentration of Coffee Shop
- Cluster 2: Neighbourhoods with low number to no existence of Coffee Shop
- Cluster 3: Neighbourhoods with moderate concentration of Coffee Shop
- Cluster 4: Neighbourhoods with high concentration of Coffee Shop



Observation

As we can see, because most of Universities in Beijing are located in Haidian District, a large numbers of coffee shops are located in this region, with with moderate number in cluster 0 and cluster 2. But does it mean we have a great opportunity to open new coffee shops in this area? Not really. Coffee shops here are concentratede in office buildings and shopping malls, that are not too far from univeristy. Brands like Starbucks, Costa, Pcific Coffee have craved up the market, but mainly focused on white-collar not too much on students. Since students in University cannot pay that expensive coffee like 30-40 yuan per cup, so there are still potential for this sub-market. Luckin coffee has been always focused on it, but since Covid-19 and there are no students in campus, so a lot of School Luckin Coffee shops have been closed. However, there is still a large investment opportunity in this area for opening studentent Coffee shop.

Outside Haidian District, there are still several university cluster but they are relatively remote and they have two caes as below. First is remote but in center of that area; Second is remote but not in center of that area. In First case, like purpple dot(Cluster1) and orange dot(Cluster 4), investors should use the same strategy as in Haidian District. In second case, you should not focus on students but also other groups like withe-collars etc.

Conclusion

- Answer to business question: The neighbourhoods in cluster 1 are the most preferred locations to open a new Student Coffee shop. It must mainly focused on students in Universities around.
- Outside Haidian District, there are still several university cluster but they are relatively remote and they have two cases as below. First is remote but in center of that area; Second is remote but not in center of that area. In First case, like purple dot(Cluster1) and orange dot(Cluster 4), investors should use the same strategy as in Haidian District. In second case, you should not focus on students but also other groups like white-collars etc.
- Findings of this project will help the relevant stakeholders to capitalize on the opportunities on high potential locations while avoiding overcrowded areas in their decisions to open a Coffee Shop in this area.