SIMILAR TO HONG KONG?

Capstone Project

Abstract

Discover the similarity of Hong Kong with other cities in Asia such as Singapore and Taipei. Provide information and analysis for people to find alternative choices of Hong Kong.

By Cloud Li

Problem Introduction

During the seventh month of protest in Hong Kong, some Hong Kongers have thought of leaving their home town. Immigration companies in Hong Kong told news reporters that they have received 60% to 70% increase in the inquiry. News published by "The New York Times" at 3rd January 2020 titled as – Hong Kong Considers the Future: 'If You Can Afford It, Leave'.

While people/investors are considering living aboard, 'where to go' becomes a hot topic. Articles on the Internet emphasize on country's immigration policy, taxes, labour supply/career development and living expense through potential places.

This project leverages the Foursquare location data to explore the similarity of different cities/countries, specifically different districts in each place. This report could be a reference for Hong Kong investors who are finding new places for their capital, or people who want to immigrant to other places through investment immigration policy on deciding what business or where to start.

Data

This project compares Hong Kong with Singapore, Shanghai, Beijing, and Taipei. Considering Hong Kong is a well-developed international financial hub, and playing the role of connecting mainland China with foreign countries. Financial institutions might consider Singapore, Shanghai, and Beijing as alternatives. Including Taipei because it is a popular vacation choice among Hong Kongers, especially youngsters.

First of all, raw data includes districts in each of the above cities/countries, which is collected through web scraping. After pre-processing and cleaning, I used a function called nominatim,

which is imported from geopy.geocoders, for collecting latitude and longitude coordinate for districts in each city/country. The following chart shows the processed data for Hong Kong.

HK

	District	Region	Latitude	Longitude
0	Central and Western,Hong Kong	Hong Kong Island	22.285039	114.158382
1	Wan Chai,Hong Kong	Hong Kong Island	22.279015	114.172483
2	Sham Shui Po,Hong Kong	Kowloon	22.328190	114.160854
3	Kowloon City,Hong Kong	Kowloon	22.330160	114.189937
4	Kwun Tong,Hong Kong	Kowloon	22.312937	114.225610
5	Wong Tai Sin,Hong Kong	Kowloon	22.341654	114.193859
6	Yau Tsim Mong,Hong Kong	Kowloon	22.307404	114.165526
7	Kwai Tsing,Hong Kong	New Territories	22.341007	114.104285
8	Sai Kung,Hong Kong	New Territories	22.382249	114.272828
9	Sha Tin,Hong Kong	New Territories	22.381056	114.188879
10	Tai Po,Hong Kong	New Territories	22.449402	114.171133
11	Tsuen Wan,Hong Kong	New Territories	22.371661	114.113470
12	Tuen Mun,Hong Kong	New Territories	22.390827	113.973169
13	Yuen Long,Hong Kong	New Territories	22.442646	114.030434

After having five data frames with coordinates, Foursquare API will be used to get venues that are nearby each district. These data will then be processed to get common venues in each district. K-Means, one of the clustering technique in a package called Scikit-learn, could cluster different locations. The result of K-Means would be discussed for the conclusion of this project.

The below graph of "HK_Venues" demonstrates an identical result of using the Foursquare API to get gear venues. One of the limitations of Foursquare API is that the exploring result for places in mainland China, such as Beijing and Shanghai, is not as good as other places. The "beijing_Venues" shows that in exploring Beijing, the API is only able to return a result of seven rows. Although I tried to find a Chinese API to replace Foursquare on searching mainland China, to the best I know, there was not such an API existing. Therefore, this project will focus on analyzing the similarity of Hong Kong and Singapore and Taipei.

HK_Venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Central and Western,Hong Kong	22.285039	114.158382	Apple ifc mall	22.284602	114.159072	Electronics Store
1	Central and Western,Hong Kong	22.285039	114.158382	Four Seasons Hotel Hong Kong (香港四季酒店)	22.286554	114.156929	Hotel
2	Central and Western,Hong Kong	22.285039	114.158382	ifc mall (國際金融中心商場)	22.284613	114.158156	Shopping Mall
3	Central and Western,Hong Kong	22.285039	114.158382	Pool Terrace (池畔餐廳)	22.286218	114.157583	Snack Place
4	Central and Western,Hong Kong	22.285039	114.158382	The Spa at Four Seasons	22.286279	114.157623	Spa
...	...	...	...	...	...	...	...
63	Yuen Long,Hong Kong	22.442646	114.030434	Aya (彩)	22.441018	114.030366	Ramen Restaurant
64	Yuen Long,Hong Kong	22.442646	114.030434	亞玉豆腐花	22.444035	114.031049	Dessert Shop
65	Yuen Long,Hong Kong	22.442646	114.030434	Ichininmae (一人前)	22.442218	114.031530	Noodle House
66	Yuen Long,Hong Kong	22.442646	114.030434	Man Fong Café (萬芳冰室)	22.440765	114.031362	Hong Kong Restaurant
67	Yuen Long,Hong Kong	22.442646	114.030434	發記腸粉粥品	22.442320	114.031150	Chinese Restaurant

68 rows × 7 columns

beijing_Venues

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Beijing City ,Beijing	39.906217	116.391276	Tian'anmen Tower (天安门城楼)	39.906562	116.391582	Historic Site
1	Beijing City ,Beijing	39.906217	116.391276	端门	39.908694	116.391192	Historic Site
2	Beijing City ,Beijing	39.906217	116.391276	Tian'anmen Square (天安门广场)	39.904848	116.391306	Plaza
3	Beijing City ,Beijing	39.906217	116.391276	Zhongshan Park (Zhongshan Park 中山公园)	39.908765	116.388602	Park
4	Beijing City ,Beijing	39.906217	116.391276	中山音乐堂 Forbidden City Concert Hall	39.909374	116.389439	Concert Hall
5	Miyun County ,Beijing	40.375103	116.837184	KFC (肯德基)	40.374122	116.839876	Fast Food Restaurant
6	Miyun County ,Beijing	40.375103	116.837184	密云鼓楼	40.375220	116.838240	Theme Park