

Capstone Project

The Battle of the Neighborhood

Introduction: Business Problem

In the project, we try to know whether specific streetlook will affect the real estate price in **Taipei**, Taiwan.

Since Taipei is a **Mixed Residential Commercial City**, there are no obvious boundary between the commercial area and residential area. Also, people live in Taipei tend to rent or buy house near the center of the city.

To find if there are some similarities or relationship between the neighborhoods and real estate price, we use the data science to separate the same type of district and compare with the real estate price.

Data

Based on our business problem, factors that we will need are:

- Neighborhood looklike of districts
- Real estate price distribution

To conduct our research, the data source are below:

- Neighborhood looklike will be extracted from the **Foursquare API**
- Divisions in Taipei will be extracted from **Wikipedia**
- The coordinates of each district and the real estate price distribution are from the **Open Source Database of Government**

Out[4]:

	PostalCode	Division	Chinese_Name
0	100	Zhongzheng District	中正區
1	103	Datong District	大同區
2	104	Zhongshan District	中山區
3	105	Songshan District	松山區
4	106	Daan District	大安區
5	108	Wanhua District	萬華區
6	110	Xinyi District	信義區
7	111	Shilin District	士林區
8	112	Beitou District	北投區
9	114	Neihu District	內湖區
10	115	Nangang District	南港區
11	116	Wenshan District	文山區

We get all the 11 districts in Taipei City

Out[5]:

	PostalCode	Division	Chinese_Name	Latitude	Longitude
0	100	Zhongzheng District	中正區	25.032405	121.519884
1	103	Datong District	大同區	25.063424	121.513042
2	104	Zhongshan District	中山區	25.069699	121.538160
3	105	Songshan District	松山區	25.059991	121.557588
4	106	Daan District	大安區	25.026770	121.543445
5	108	Wanhua District	萬華區	25.028590	121.497986
6	110	Xinyi District	信義區	25.030621	121.571670
7	111	Shilin District	士林區	25.125467	121.550847
8	112	Beitou District	北投區	25.148068	121.517799
9	114	Neihu District	內湖區	25.083706	121.592383
10	115	Nangang District	南港區	25.036009	121.609757
11	116	Wenshan District	文山區	24.988579	121.573608

The Real estate Price Data

Out[41]:

	PostalCode	Division	Chinese_Name	500	501~1000	1001~1500	1501~2000	2001~2500	2501~3000	3001~4000	4001~5000	5001~7000	7001~9000	90
0	100	Zhongzheng District	中正區	18	51	80	60	59	40	52	22	24	10	3
1	103	Datong District	大同區	31	71	62	50	37	25	21	13	12	1	1
2	104	Zhongshan District	中山區	57	219	245	151	119	75	66	45	63	14	13
3	105	Songshan District	松山區	12	45	81	93	76	44	64	31	12	4	1
4	106	Daan District	大安區	16	31	92	92	99	89	141	61	71	23	12
5	108	Wanhua District	萬華區	56	128	89	72	31	21	26	14	1	1	0
6	110	Xinyi District	信義區	22	38	100	122	90	54	42	17	22	6	7
7	111	Shilin District	士林區	23	64	143	124	61	47	43	24	31	2	2
8	112	Beitou District	北投區	30	125	142	108	55	35	50	26	21	3	3
9	114	Neihu District	內湖區	6	106	223	215	130	96	58	56	75	12	8
10	115	Nangang District	南港區	3	30	69	55	45	27	33	28	13	10	8
11	116	Wenshan District	文山區	26	102	205	134	85	66	67	21	10	0	0

Methodology

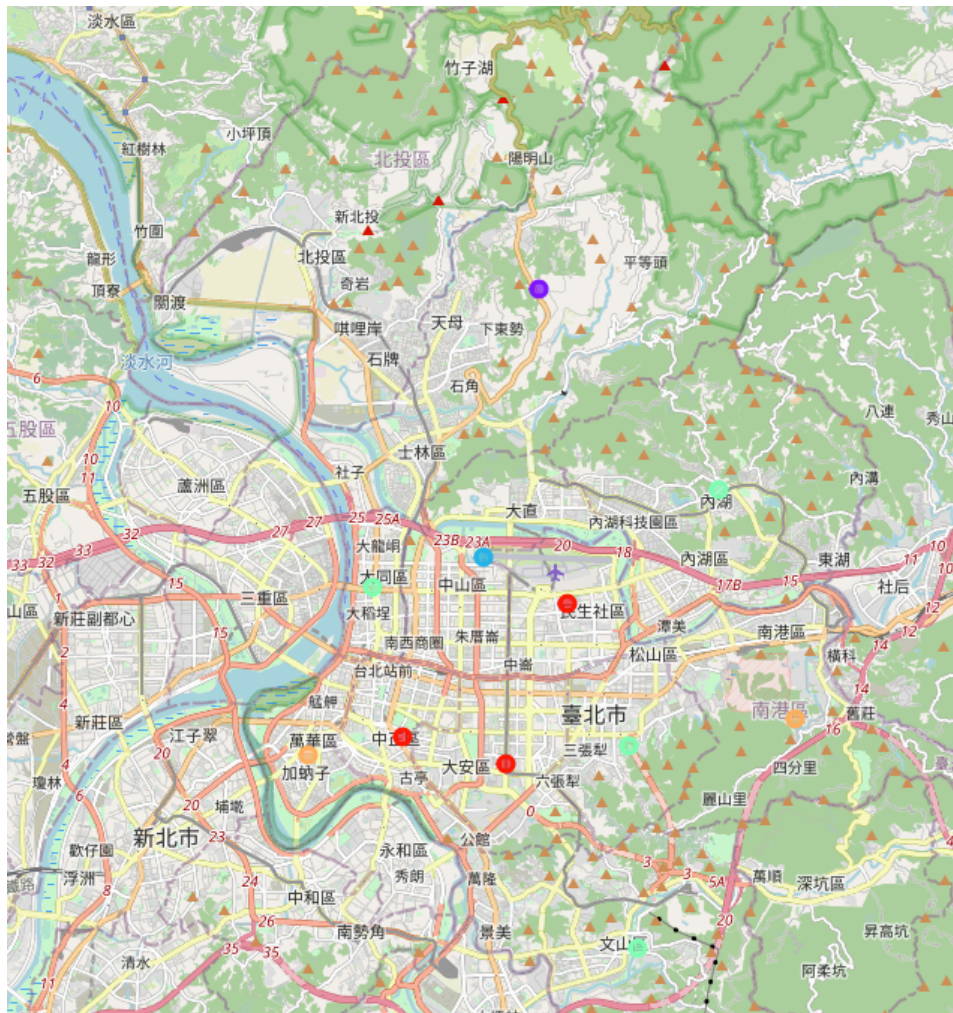
First, we use the Foursquare API to determine the neighborhood looklike in our districts in Taipei, using frequency as factor to calculate.

And we use machine learning clustering to find the similar community.

Then we use the clustering to cluster the realestate price distribution and get another cluster.

Compare these 2 clusters to find the similar districts.

Analysis



Result and Discussion

According to the clustering, in the similar neighborhood, we've got:

- Cluster 0: Zhongzheng, Songshan and Daan: Cafe, Coffee Shop, Dumpling Restaurant
- Cluster 1: Shilin: Asian Restaurant, Trail, Airport Terminal
- Cluster 2: Zhongshan: Convenience Store, Fish Market, Farmers Market
- Cluster 3: Datong, Xinyi, Neihu and Wenshan: Convenience Store, Chinese Restaurant, Coffee Shop
- Cluster 4: Wanhua and Nangang: Convenience Store, Supermarket, Bus Station

And the price clusters are:

- Cluster 0: Daan (Highest average)
- Cluster 1: Xinyi, Shilin, Beitou and Wenshan (Lowest average)
- Cluster 2: Zhongzheng, Datong, Songshan, Wanhua, Nangang
- Cluster 3: Zhongshan
- Cluster 4: Neihu

According to the result above, we can find there are similar communities like:

- ZhongZheng and Songshan
- Xinyi and Wenshan
- Wanhua and Nangang
- Zhongshan is a special one

Conclusion ¶

It is interesting to find the community featuring and price relationship in the city I am currently living. But the Foursquare API seems not have much data in Taiwan, there are only about 250 venues in the Taipei city, so the bias may influence the result.

According to my experience, Daan District is the Top1 district that rich people want to live in Taipei, with large parks and convenient transportation. Otherwise Wanhua is facing old age and poverty problem (which can see from the average realestate price), and the street look is also old and more complex. But the neighborhood cluster is in the same group with Nangang, which is rapidly developing in Taipei.

But we do find some interesting point:

1. Xinyi and Wenshan are close and the street look is quite similar, the realestate price are quite similar too (except for the CBD in Xinyi)
2. Zhongshan is a really special space, it has not only the older and complex area like Wanhua, but the comfortable area like Daan, so the price distribution is extreme.