Good positions to live or open a restaurant in HaiKou

Xiang Fu 05/09/2020

Section 1: Introduction

My hometown: Haikou, the southern city in China, is an international tourism city. Haikou is the capital and most populous city of the Chinese province of Hainan. As of 2018, Haikou has the second best air quality among major cities nationally, according to the 2005 statistical book issued by the National Bureau of Statistics, Haikou scored the highest among China's main cities in air quality. Its beautiful sea-views and wonderful weather attracted millions of adventurous people from the globe to tour and live in this city.

The city received over 6 million tourists each year. The city earned approximately three billion yuan (361 million US dollars) from the tourism industry since 2002, and keep fast increasing each year. so many businessmen may want to know which neighborhood is best to live in or open a new restaurant, the project here is deemed to answer the questions and try to provide an alternative solution.

Section 2: Data Collection

The project uses three datasets: *Haikou_GDP , Haikou_NBH* and *Haikou_POP*.

Haikou_GDP includes economic and geographic data of all 4 districts in Haikou (Haikou is a young city, which has just built up for 32 years until 2020). It has 8 variables:

Table 1 Haikou GDP data

	District	Postal code	Population (Tenth of Thousands)	Area (Square Kilometres)	GDP (Tenth of Thousands RMB)	GDP Per Capital (RMB)	salary (RMB)	Home Price (RMB)
0	LongHua	570100.0	66.98	300.6	6873900.0	102626.157062	34220.0	14664
1	XiuYing	570300.0	38.78	511.5	2202000.0	56781.846313	30428.0	16826
2	QiongShan	571100.0	51.17	939.9	1700600.0	33234.316983	31405.0	13669
3	MeiLan	570200.0	70.28	552.8	3640000.0	51792.828685	31503.0	16263

This data lists all of the Haikou's 4 districts. And covers the population, GDP, Salary and Home Price. I will convert some data for better analysis such as use population/area to get the density of population.

Haikou_NBH lists 18 prominent/popular neighborhoods in Haikou as well as the districts they are located in. It has three variables: *District*, *Neighborhood, Latitude* and *Longitude*.

Table 2 Haikou Neighborhood data

	District	Neighborhood	latitude	longitude
0	LongHua	Haikou Gymnasium	20.033267	110.317063
1	LongHua	Hainan Exhibition Convention Center	20.033332	110.310497
2	LongHua	Jinniuling Park	20.012162	110.312838
3	LongHua	Xiuying Fort	20.021276	110.307811
4	LongHua	Yusha Capital City	20.027786	110.317278
5	LongHua	People's Park	20.033424	110.341123
6	XiuYing	CHANGLIU	20.011966	110.192909
7	XiuYing	HAIXIUZHEN	19.992670	110.279999
8	XiuYing	Haikou Bay	20.028924	110.277252
9	XiuYing	HOUHAI	20.058145	110.188431
10	QiongShan	DINGCUN	19.977550	110.338411
11	QiongShan	DAOKECUN	20.010427	110.331246
12	QiongShan	Temple of the Five Lords	20.007195	110.356431
13	QiongShan	PANDANCUN	20.009864	110.365142
14	MeiLan	HAIDIAN	20.049942	110.337175
15	MeiLan	Baishamen	20.066214	110.316566
16	MeiLan	Mingzhu Square	20.025723	110.310838
17	MeiLan	XINBU	20.073899	110.353644

Although these 18 neighborhoods are not as much as those in large cities, they are still very welcome by the local people. The data here is limited because of the difficulty during data collection, but those most important places were include.

Haikou_POP contains the following information:

Table 3 Haikou Population data

	Neighborhood	Population	0-14 Years	15-64 Years	>65 Years
0	Neighborhood	Population	0-14 Years	15-64 Years	>65 Years
1	Total	2.2721e+06	16.78	76.04	7.18
2	XiuYing	387800	18.3	74.85	6.85
3	LongHua	669800	16.88	76.45	6.67
4	QiongShan	511700	18.09	74.49	7.42
5	MeiLan	702800	14.9	77.43	7.67

The population is the same as that in *Haikou_GDP*, but the ratios of ages are very useful, they are good indicators for special purpose like find where to open children's cloth store.

All of the data source listed as below (some of them are chinese and I translated them into English in my dataset):

Geography: https://en.wikipedia.org/wiki/Haikou

GDP : https://zh.twgreatdaily.com/0m-p_GwBJleJMoPMNI3m.html

Home price: http://hk.cityhouse.cn/market/

Area: https://zh.wikipedia.org/wiki/%E6%B5%B7%E5%8F%A3%E5%B8%82

Salary:

longhua :

http://www.haikou.gov.cn/xxgk/szfbjxxgk/jhzj/qzfgzbg/201907/t20190723_1428721.html

XiuYing :

https://www.hainan.gov.cn/hn/zwgk/zfgzbg/dszf/2018sx/201804/t20180408_2600533.html

QiongShan:

https://www.hainan.gov.cn/hn/zwgk/zfgzbg/dszf/2018sx/201804/t20180409_2600878.html

MeiLan: http://www.hainan.gov.cn/hn/zwgk/zfgzbg/dszf/201804/t20180408_2600506.html

Please Note that it is extremely hard to find one single source that provides related data, but for me, as a local citizen lived there for over 20 year, I can assure above data could reflect the accuracy of data.

Section 3: Methodology

In order to answer questions that interest tourists and newcomers to the city, methods such as exploratory data analysis are used. The best indicator for this purpose is to find the venues.

3.1 Use folium and FourSquare API to get venues' list of the neighborhoods.

With the geographic (latitude and longitude) information of all neighborhoods in Haokou at handy, we can create a map of Haikou with all 18 prominent neighborhoods highlighted.

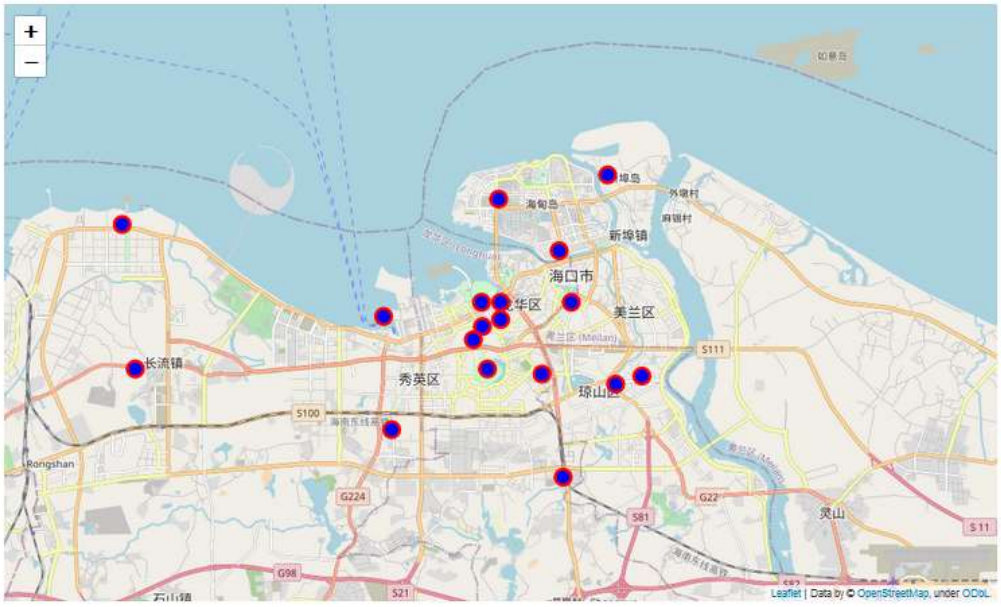

Figure 1. Haikou Venues Map by using folium library and FourSquare API

Then the next step is to use Foursquare API to explore neighborhoods in Haikou. Foursquare API enables developers to get trending venues around a location. By constructing a URL, setting parameters and sending a request to the API, the following results are returned:

Table 4 Venues data returned by FourSquare

	District	Neighborhood	NBH Latitude	NBH Longitude	Venue(V)	V Latitude	V Longitude	V Category
0	LongHua	Haikou Gymnasium	20.033267	110.317063	Pizza Hut (必胜客)	20.030633	110.315275	Pizza Place
1	LongHua	Haikou Gymnasium	20.033267	110.317063	McDonald's (麦当劳)	20.028276	110.315507	Fast Food Restaurant
2	LongHua	Haikou Gymnasium	20.033267	110.317063	Starbucks (星巴克)	20.029340	110.315886	Coffee Shop
3	LongHua	Haikou Gymnasium	20.033267	110.317063	观澜湖（电影公社）	20.034981	110.316828	Multiplex
4	LongHua	Haikou Gymnasium	20.033267	110.317063	万达影城	20.032488	110.322245	Multiplex
...	...	...	...	...	...	...	...	...
75	MeiLan	Baishamen	20.066214	110.316566	美丽沙食府	20.067508	110.311310	Chinese Restaurant
76	MeiLan	Baishamen	20.066214	110.316566	阿南鲜鲍鱼汤	20.063843	110.321640	Seafood Restaurant
81	MeiLan	XINBU	20.073899	110.353644	Howard Johnson New Port Resort Haikou (海口皇海湾蒙生...	20.071403	110.351469	Resort
82	MeiLan	XINBU	20.073899	110.353644	Novotel Haikou	20.072424	110.357940	Hotel
83	MeiLan	XINBU	20.073899	110.353644	maan coffee	20.068875	110.352330	Coffee Shop

68 rows × 8 columns

This answers the key questions for newcomers: what venues can I walk by when I live in a neighborhood? Notice that each venue' latitude and longitude is also generated by Foursquare API. Therefore, they will be applied for showing more detailed in the map.

Next, with venues data, I found that there are 32 kinds of venues, then I use the code to calculate the total numbers of each category and sort them with the 'District'.

Table 5 detailed list to show all the categories of the venues

	Neighborhood	District	Asian Restaurant	BBQ Joint	Bar	Boat or Ferry	Café	Chinese Restaurant	Clothing Store	Coffee Shop	...	Pizza Place	Resort	Restaurant	Seafood Restaurant	Shopping Mall	Snack Place
7	Haikou Gymnasium	LongHua	0	0	0	0	0	0	0	1	...	1	0	0	0	1	0
8	Hainan Exhibition Convention Center	LongHua	0	0	0	0	0	0	1	0	...	1	0	0	0	0	0
9	Jinniuling Park	LongHua	0	0	0	0	0	0	0	0	...	0	0	1	1	0	0
11	People's Park	LongHua	1	0	0	0	0	0	0	0	...	1	0	0	0	3	0
14	Xiuying Fort	LongHua	0	0	0	0	0	0	0	0	...	1	0	0	0	0	0
15	Yusha Capital City	LongHua	0	0	1	0	0	0	0	0	...	0	0	0	0	0	0
0	Baishamen	MeiLan	0	0	0	0	1	1	0	0	...	0	0	0	1	0	0
3	HAIDIAN	MeiLan	0	0	0	0	0	0	0	1	...	0	0	0	0	0	1
13	XINBU	MeiLan	0	0	0	0	0	0	0	1	...	0	1	0	0	0	0
1	DAOKECUN	QiongShan	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
2	DINGCUN	QiongShan	0	0	0	0	0	0	0	0	...	0	0	0	1	1	0
10	PANDANCUN	QiongShan	0	1	0	0	0	0	0	0	...	0	0	0	0	0	1
12	Temple of the Five Lords	QiongShan	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
4	HAIXIUZHEN	XiuYing	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
5	HOUHAI	XiuYing	0	0	0	0	1	0	0	0	...	0	0	0	0	1	0
6	Haikou Bay	XiuYing	0	0	0	3	0	0	0	0	...	0	0	0	0	0	0
total	Total	Total	1	1	1	3	2	1	1	3	...	4	1	1	3	6	2

17 rows × 34 columns

3.2 Prepare the data for analysis

For the target of this project, the plot figures are needed to show readers where to select.
It is good to sum up the above data to get the total number of venues for each district for next step's plotting.

	Venues
LongHua	31
XiuYing	11
QiongShan	15
MeiLan	11

Not matter for new visitors or businessmen, it is best to separate the venue data for each district and apply to the next steps.

Table 6 summary of the venues for each district

District	Neighborhood	Venues
XiuYing	HOUHAI	6
	Haikou Bay	4
	HAIXIUZHEN	1
QiongShan	DINGCUN	5
	PANDANCUN	4
	Temple of the Five Lords	4
	DAOKECUN	2
MeiLan	Baishamen	4
	HAIDIAN	4
	XINBU	3
LongHua	People's Park	12
	Haikou Gymnasium	7
	Jinniuling Park	5
	Xiuying Fort	3
	Hainan Exhibition Convention Center	2
	Yusha Capital City	2

At last, get the GDP, salary, home price and population data ready as the parameter for analysis.

Table 7 summary of the General information

	District	GDP Per Capital (RMB)	salary (RMB)	Home Price (RMB)	Venues	Population	0-14 Years	15-64 Years	>65 Years	Area (Square Kilometres)
0	LongHua	102626.157062	34220.0	14664	31	2228.21	376.122	1703.47	148.622	300.6
1	XiuYing	56781.846313	30428.0	16826	11	758.162	138.744	567.484	51.9341	511.5
2	QiongShan	33234.316983	31405.0	13669	15	544.42	98.4855	405.538	40.3959	939.9
3	MeiLan	51792.828685	31503.0	16263	11	1271.35	189.431	984.403	97.5122	552.8

Section 4: Results

The results are expected and intuitive. The very central part of Haikou (LongHua District) can be interpreted as the focused location. In real, it is Haikou's political and economic center surrounded by the rest of other districts. Therefore, for newcomers, they are recommended to tour in the center at first. Its pros is surrounded by many venues so that is benefit to live or open a restaurant; while its cons, is by its most population of all, so that it might be very crowded, expensive and with busy traffic in most of the time. The bar charts below help to illustrate this point:

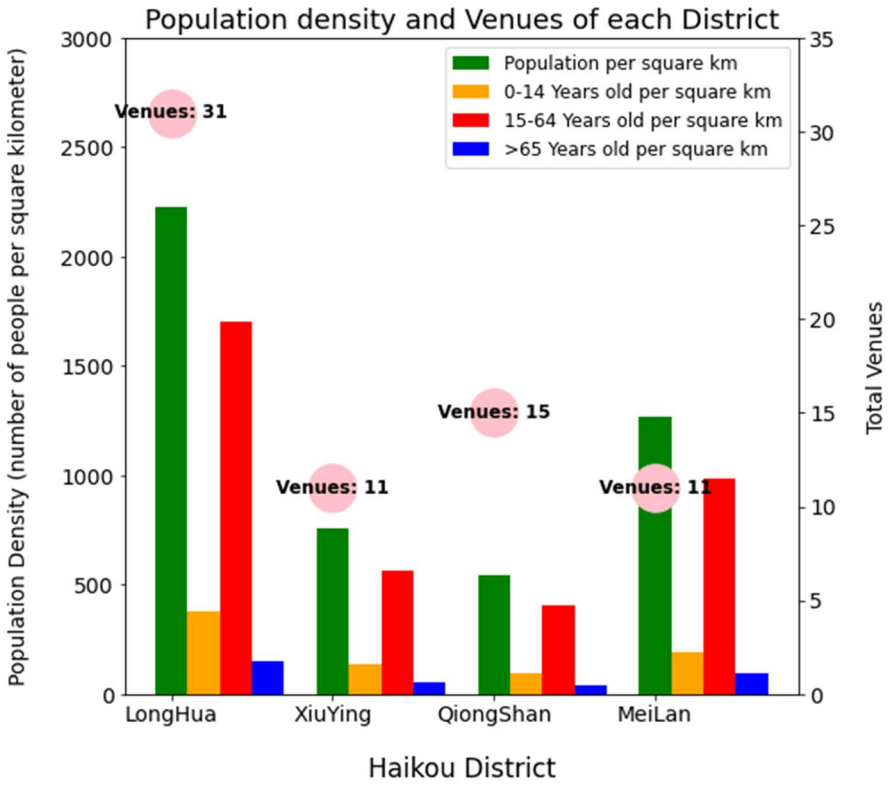

Figure 2 Population vs. Venue number of each district

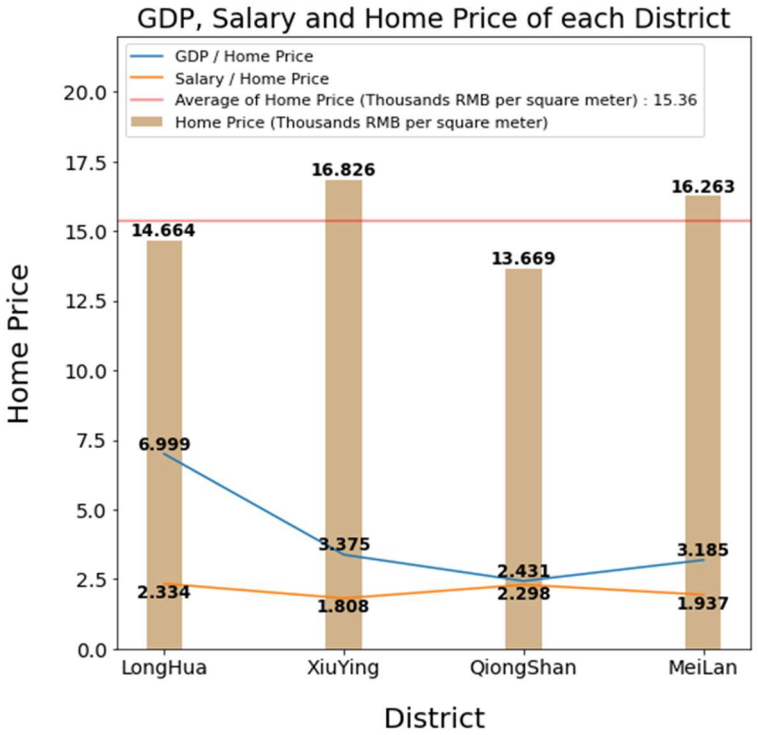

Figure 3 GDP and Salary vs. Home Price of each district

From these charts, it can be easily seem that LongHua has the most venues, most population, and highest GDP per capita. Meanwhile, its home price is relatively lower and salary is second highest, meaning that it would be extremely good for a starter to afford the cost of daily living or begin a business. But it should be also noted that the home price represents how welcome the place is, like XiuYing, which has the highest home price, is the community of many rich men because it has beautiful beaches and new houses. So that if the newcomers have enough budgets and owning personal vehicles (go to work in LongHua from XiuYing), XiuYing is the best choice.

Next, for different purpose, the detailed venues for each district were exhibited by the plotting,

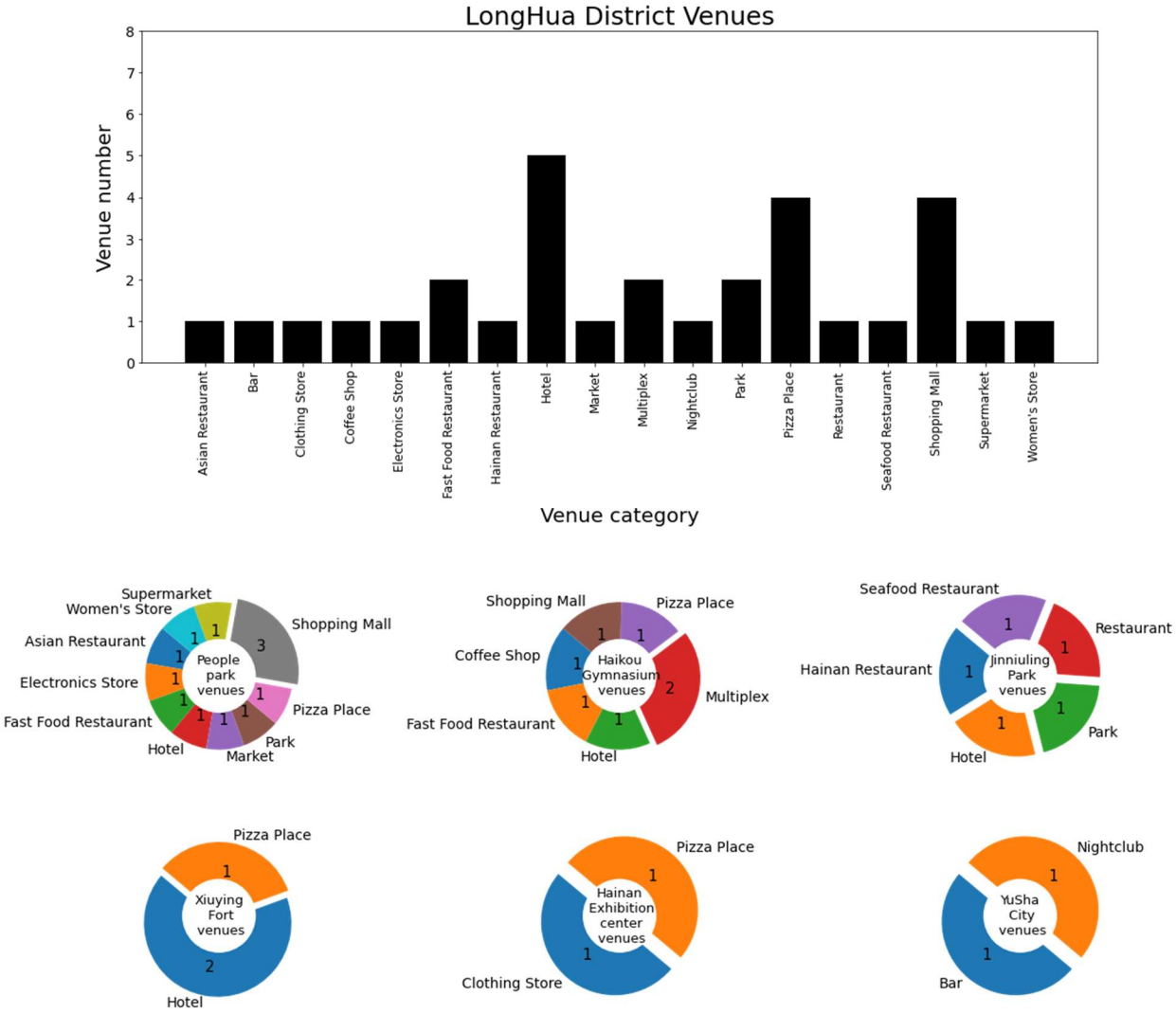

Figure 4 Venues' information of LongHua district

LongHua is the busiest district in terms of GDP per capita ranking, and people in this area tend to earn the most and have relatively less stress to live. In fact, it is very good for many newcomers who want to live here while having fun during the weekend. Because there are many places to go and visit, and the public transportation is convenient in Haikou. So as the economic and community center, LongHua is the first choice for those whom want to begin a life or open restaurants.

The other district are also potential good choice, there may be good places to live, or less competitive in business:

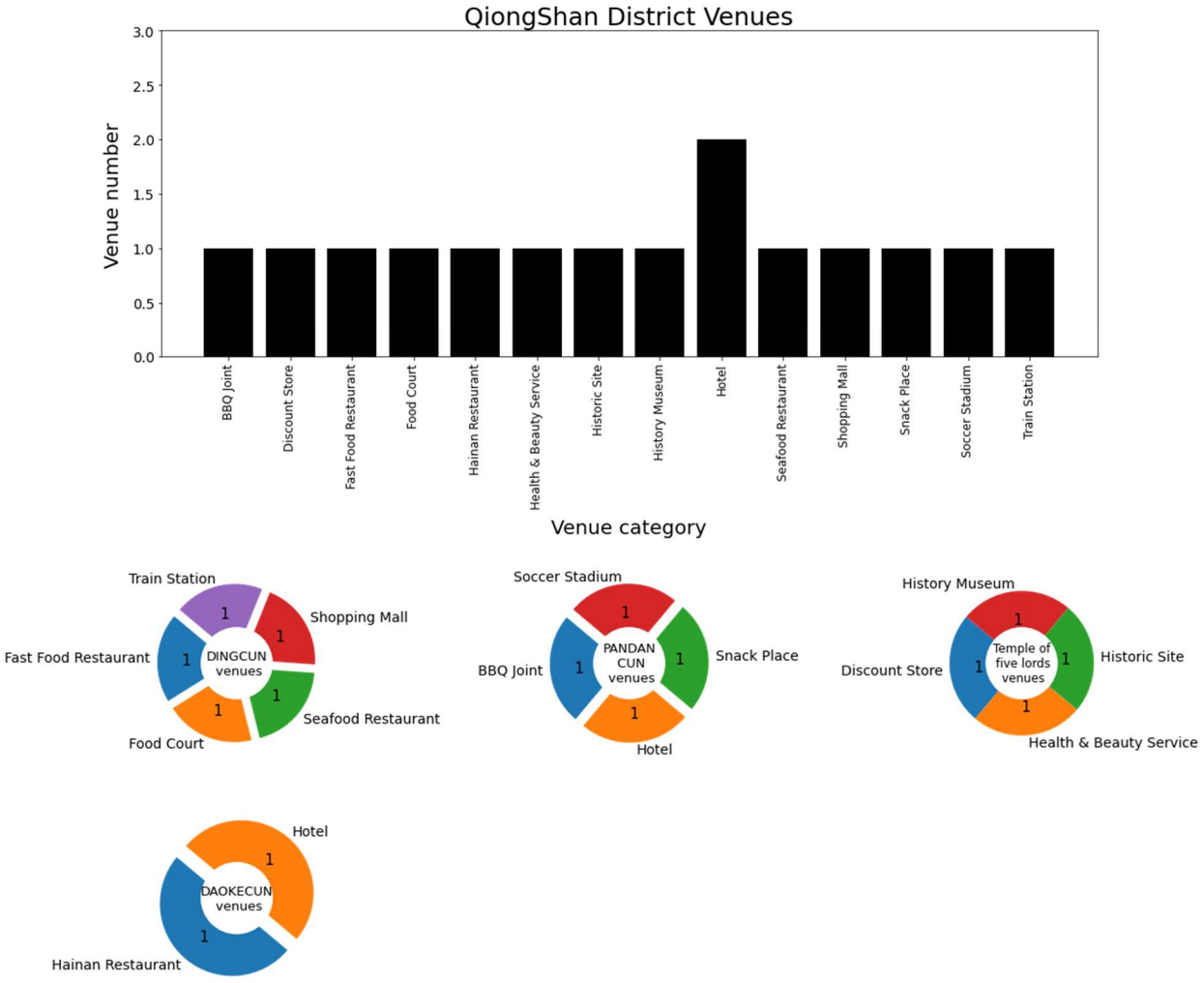

Figure 5 Venues' information of QiongShan district

QiongShan has the biggest area, there are many undeveloped lands, it may open many industrial factories in the future, if there is some data, it worth to see if it is good to open a factory.

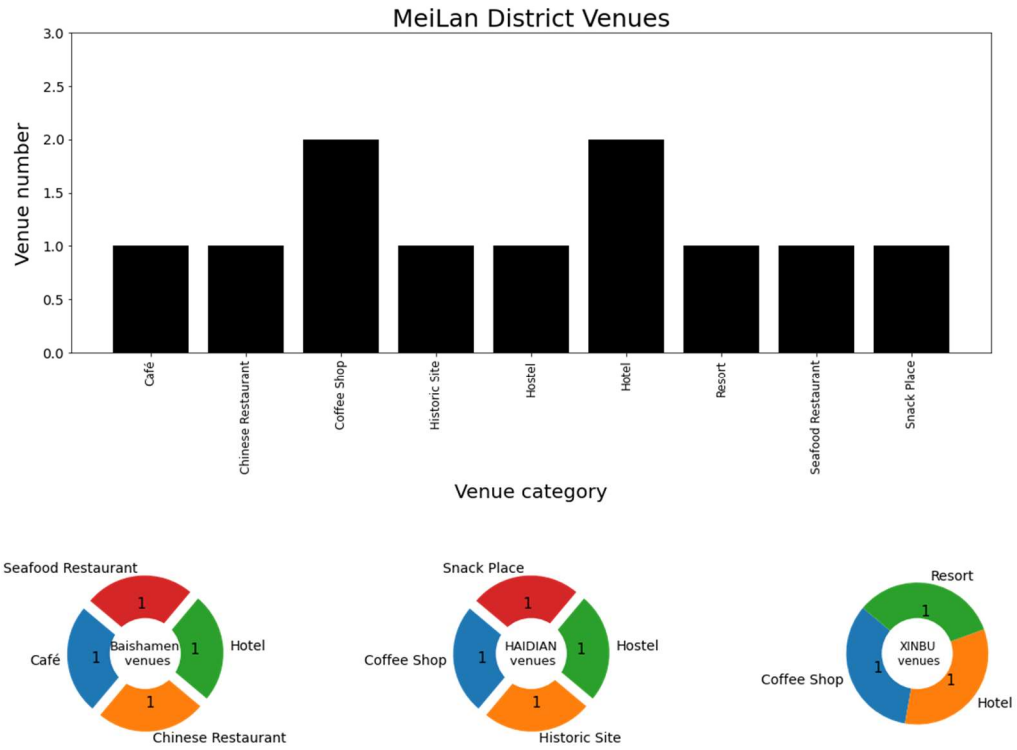

Figure 6 Venues' information of MeiLan district

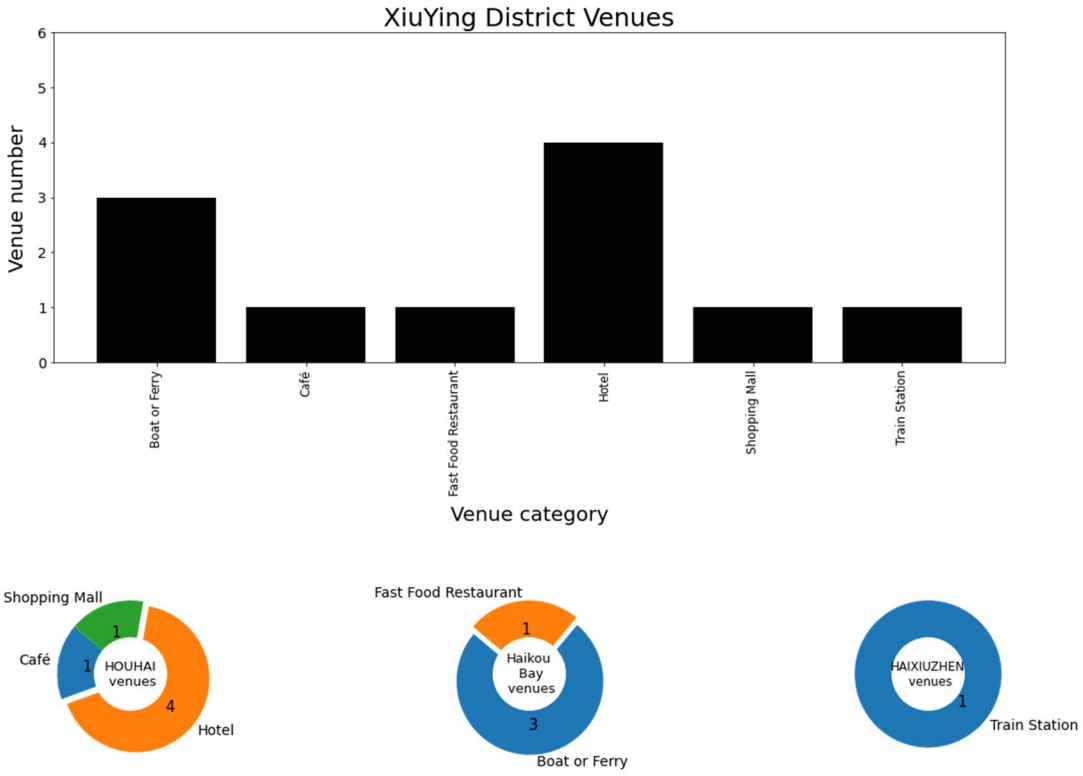

Figure 7 Venues' information of XiuYing district

MeiLan is the potential good place, for its young ages and GDP increasing very fast, and XiuYing, as described above, has the best houses. Both of them are also nice location to do development.

From the exploratory data analysis in the result, one can easily locate favorite venues in Haikou and have a rough idea on which neighborhoods are touristy and which district is suitable for a newcomer to live in.

The project should consider different purposes for different people, so the small program also included to show all venues in the map. Use folium and FourSquare again, the color circle separates the venues of different district (LongHua: Red, XiuYing: Blue, QiongShan: Orange, MeiLan: Green):

```
# or input 0 for the all the venues in the map:
Haikou_check(0,12)
```

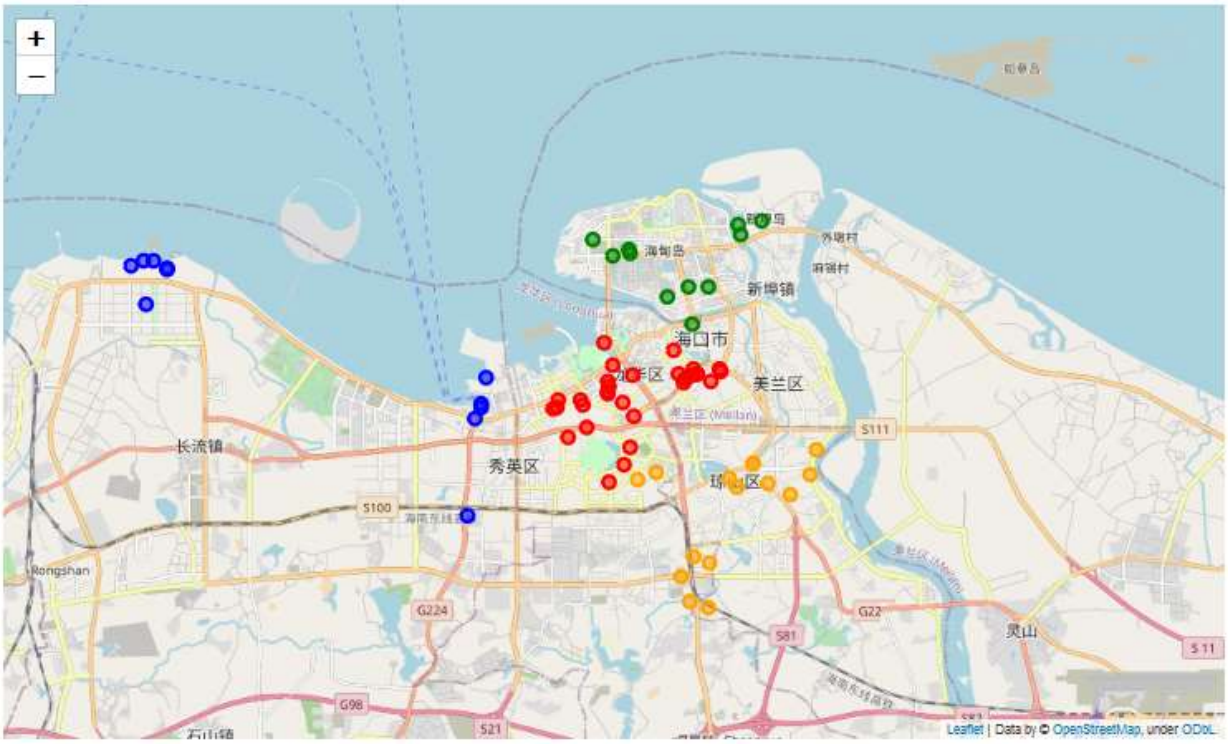

Figure 8 information of each venue in the program

This map can be easily access to find special category, location and name of each venue. So as to fulfill different needs. For example, if the person wants to find the information of Hotel, the 'Hotel' index could show him/her the locations and names of the hotels in the map (see below). Other categories can also be found from this map.

```
# Draw map using the index, input the the venue type and the zoom_size of the map into the parenthess
Haikou_check('Hotel',12)
```

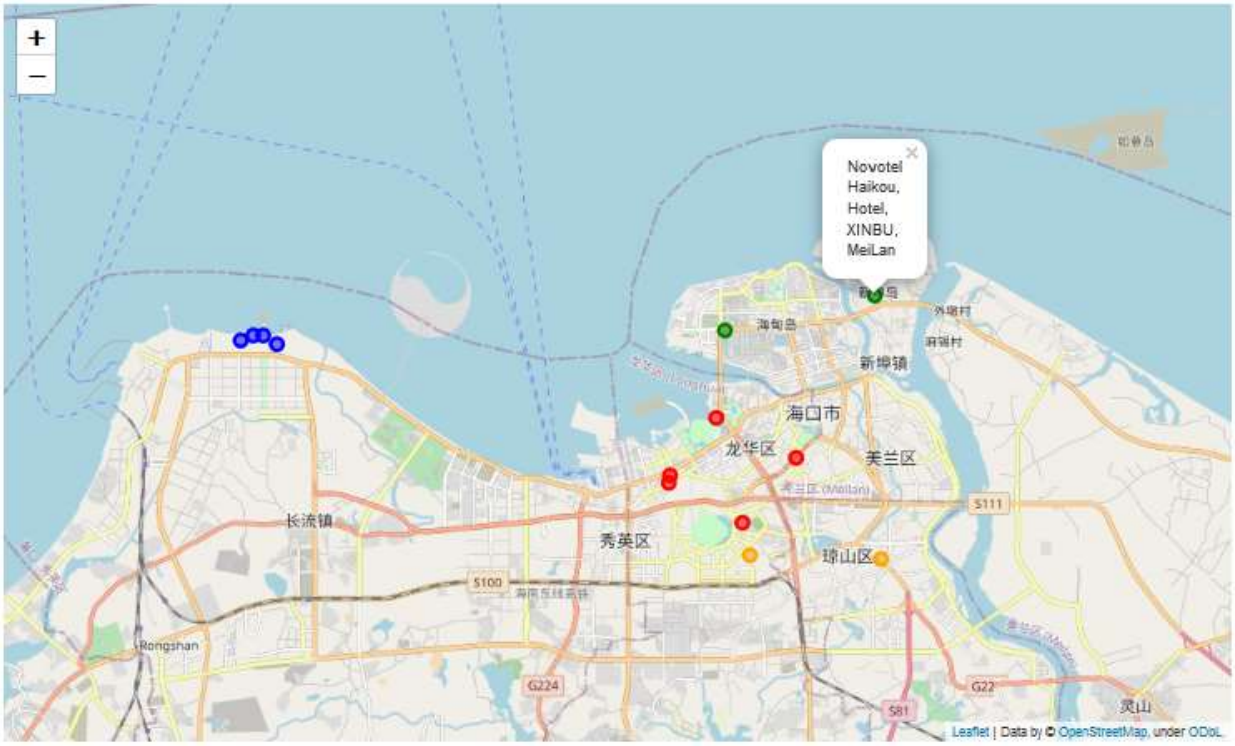

Figure 9 example of finding the hotel in the program

Section 5: Discussion and Conclusion

Before conclusion, it is essential to mention potential problems of this project. Firstly, like has already been discussed in the Data section, all the data of each district is not consistent; even though they are all in 2018, they are not in a time series, it is hard to do prediction by a single year. Therefore I didn't do the clustering in this project. And those data are not from the same source, and the time is too limited to do verification for all of them. Secondly, venues in Haikou generated by Foursquare API are problematic and are limited in numbers. Many locations only return renowned venues (e.g. KFC and McDonald's for fast food category, and Starbucks for café category). It fails to return less renowned or "Eastern" venues (like Chinese restaurant, which is most popular in Hakou). For example, the venue numbers are very limited for each neighborhood of District, in reality, there are actually hundreds of 'Chinese name' venues around every neighborhood, however, the API only return few. It may also because Haikou is a young

city and there are few data in the system, and remote areas in Haikou has never been a focus of Foursquare API, and it fails to take into account location information of many less well-known venues. Finally, reproduced counting. Foursquare API is unable to categorize venues that belong to the same category. For instance, Coffee shops and café are essentially the same.

In conclusion, this project still presents many venues information for each neighborhood over Haikou city, categorizes all prominent neighborhoods, and visualizes them in the. The project recommends newcomers to begin lives or business in LongHua District and consider others depending on the budgets.