IBM DATA SCIENCE CAPSTONE

Café Location Research

INTRODUCTION

Business Problem

A client wants to open a café in Seoul as part of her retirement plan; she hired us to pick out the optimal location for her new business. The client requested that we pay special attention to the rent cost.

According to Hyundai Research Institute, coffee is a 5.7 billion dollar industry in South Korea. An average adult drinks 353 cups of coffee per year. Despite the high demand, coffee is a risky business, because the market is already saturated. Our client is one of many who wants to run a café after retirement.

So where should we set up a cafe in a city saturated with big and small coffee shops?

We consulted the advices given by commerical property experts in Seoul. Five suggestions were given:

1. Although setting up a shop in a large commercial district guarantees revenue to an extent, the cost is too high. It is better to target the periphery of a growing commercial district with significant foot traffic and cheaper rent.
2. It is a bad idea to set up a shop near malls, which provide better facilities than surrounding shops. People will likely walk past the cafe.
3. Avoid a location with a competitor within 200-300m radius.
4. Office complex or university nearby will boost the coffee business.
5. Restuarants can be a plus as Koreans tend to go to a cafe after the restuarant.[1]

Objectives

Based on these advices and the client's preference, we devised a list of factors to consider:
1. The cafe should be situated at the periphery of a large commericial district or a growing commerical district. Therefore, the cafe cannot be near a mall.
2. There cannot be another coffee shop within 200-300m radius.
3. Neighbourhood near a university or an office complex is a good candidate.

Our research objective is to narrow down the location candidates by picking out a district which:
- has a university.
- has less coffee shops.
- has less malls.

[1] Sung-joon Huh, "Where to set up a cafe according to three commerical propery experts," https://biz.chosun.com/site/data/html_dir/2011/12/29/2011122901750.html.

COMPANY NAME

Solution

Based on the above criteria, we recommend the following three districts as good locations for a new café: "Seongbuk-gu," "Dongdaemun-gu", "Seongdong-gu".

Data

We extracted data from four different sources.

1. A list of districts in Seoul from Wikipedia: https://en.wikipedia.org/wiki/List_of_districts_of_Seoul.
2. Foursquare API for locating malls and universities in Seoul
3. Geopy for retrieving the geological locations
4. "2020 Seoul Commercial Area Information for Small Business Market Promotion" published by the Ministry of SMEs and Startups in South Korea for data on all the cafes located in Seoul.

METHODOLOGY

DISTRICTS IN SEOUL

First, we created a list of districts in Seoul with geographic coordinate information. The list was scraped from Wikipedia with BeautifulSoup. Then, we used Geopy to add the geographic coordinate information to each district. The resulting list consisted of 25 different districts.

We extracted the district names from the addresses, since we are recommending districts ideal for setting up a new café.

	District	Population	Area	Population density	Gu	Latitude	Longitude
0	Dobong-gu	355712	20.70 km²	17184/km²	도봉구	37.668600	127.046600
1	Dongdaemun-gu	376319	14.21 km²	26483/km²	동대문구	37.574200	127.039500
2	Dongjak-gu	419261	16.35 km²	25643/km²	동작구	37.512100	126.939500
3	Eunpyeong-gu	503243	29.70 km²	16944/km²	은평구	37.602400	126.929300
4	Gangbuk-gu	338410	23.60 km²	14339/km²	강북구	37.639500	127.025500
5	Gangdong-gu	481332	24.59 km²	19574/km²	강동구	37.530000	127.123700
6	Gangnam-gu	583446	39.50 km²	14771/km²	강남구	37.517700	127.047300
7	Gangseo-gu	591653	41.43 km²	14281/km²	강서구	37.550900	126.849700
8	Geumcheon-gu	258030	13.02 km²	19818/km²	금천구	37.456500	126.895400
9	Guro-gu	457131	20.12 km²	22720/km²	구로구	37.495200	126.887700
10	Gwanak-gu	531960	29.57 km²	17990/km²	관악구	37.478200	126.951800
11	Gwangjin-gu	377375	17.06 km²	22120/km²	광진구	37.538400	127.082800
12	Jongno-gu	165344	23.91 km²	6915/km²	종로구	37.580310	126.983079
13	Jung-gu	136227	9.96 km²	13677/km²	중구	37.563656	126.997510
14	Jungnang-gu	423411	18.50 km²	22887/km²	중랑구	37.606300	127.093000
15	Mapo-gu	395830	23.84 km²	16604/km²	마포구	37.566571	126.901532
16	Nowon-gu	586056	35.44 km²	16536/km²	노원구	37.654000	127.056700
17	Seocho-gu	454288	47.00 km²	9666/km²	서초구	37.483500	127.032200
18	Seodaemun-gu	320861	17.61 km²	18220/km²	서대문구	37.579075	126.936786
19	Seongbuk-gu	475961	24.58 km²	19364/km²	성북구	37.590000	127.016500
20	Seongdong-gu	303891	16.86 km²	19364/km²	성동구	37.563500	127.036500
21	Songpa-gu	671794	33.88 km²	19829/km²	송파구	37.514500	127.105800
22	Yangcheon-gu	490708	17.40 km²	28202/km²	양천구	37.517100	126.866300
23	Yeongdeungpo-gu	421436	24.53 km²	17180/km²	영등포구	37.526200	126.895900
24	Yongsan-gu	249914	21.87 km²	11427/km²	용산구	37.532300	126.990000
25	Seoul	10369593	605.21 km²	17134/km²	서울	37.566679	126.978291

FOURSQUARE API: UNIVERSITIES AND MALLS IN SEOUL

We used Foursquare API to create a list of universities in Seoul, since placing the café near one is part of our game plan. There are nearly 20 universities. After cleaning up the data, we were able to narrow down to 12 districts based on the presence of one or more universities.

	District	Population	Area	Population density	Gu	Latitude	Longitude	University
0	Dongdaemun-gu	376319	14.21 km²	26483/km²	동대문구	37.574200	127.039500	2
1	Dongjak-gu	419261	16.35 km²	25643/km²	동작구	37.512100	126.939500	2
2	Eunpyeong-gu	503243	29.70 km²	16944/km²	은평구	37.602400	126.929300	1
3	Gwanak-gu	531960	29.57 km²	17990/km²	관악구	37.478200	126.951800	1
4	Gwangjin-gu	377375	17.06 km²	22120/km²	광진구	37.538400	127.082800	2
5	Mapo-gu	395830	23.84 km²	16604/km²	마포구	37.566571	126.901532	1
6	Nowon-gu	586056	35.44 km²	16536/km²	노원구	37.654000	127.056700	1
7	Seocho-gu	454288	47.00 km²	9666/km²	서초구	37.483500	127.032200	1
8	Seodaemun-gu	320861	17.61 km²	18220/km²	서대문구	37.579075	126.936786	4
9	Seongbuk-gu	475961	24.58 km²	19364/km²	성북구	37.590000	127.016500	2
10	Seongdong-gu	303891	16.86 km²	19364/km²	성동구	37.563500	127.036500	2

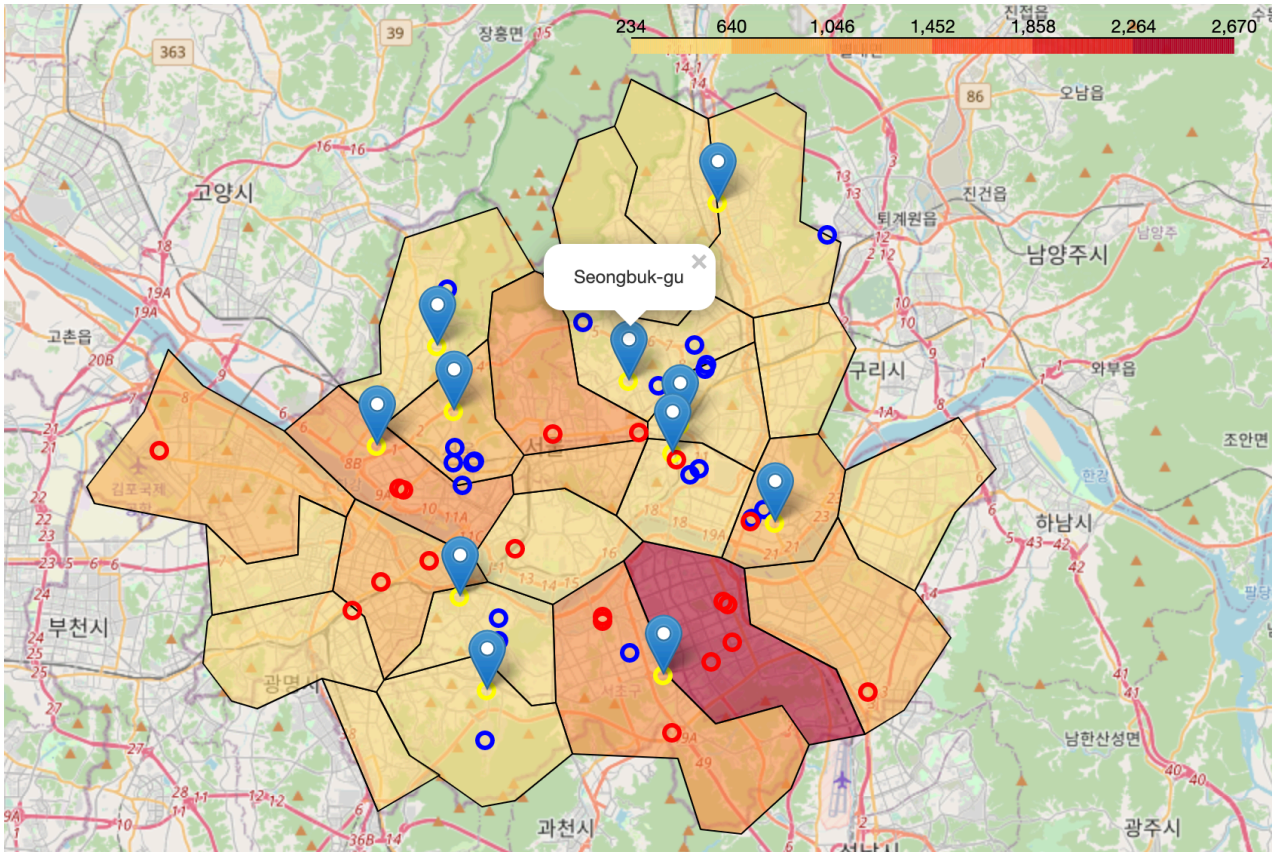
We repeated the same process to find all the malls in Seoul, because we want to avoid making the mistake of having a mall as the café's competitor. With a free Foursquare developer account, it was impossible to use FourSquare API to get data on *all* the cafés in Seoul. There are literally 18220 cafés in Seoul! So we turned to the "2020 Seoul Commercial Area Information for Small Business Market Promotion" published as csv file by the Ministry of SMEs and Startups in South Korea. After cleaning the data, we were able to find out how many cafés are in each district. Ideally, the café should be set up in a district with less café density.

시군구명	상호명
강남구	2646
서초구	1321
마포구	1233
종로구	1121
영등포구	986
중구	914
송파구	891
서대문구	772
강서구	763
광진구	759
용산구	612
강동구	556
성북구	542
구로구	537
은평구	531
노원구	506
관악구	493
동대문구	492
성동구	467
양천구	456
동작구	408
금천구	372
중랑구	305
도봉구	279
강북구	258

Finally, we used the dataframes generated for universities, malls, and cafes to create maps. First, a choropleth map, based on the number of café in each district, allowed us to quickly rule out the districts overly saturated with cafés. Second, blue markers for universities were added to the map as well as pop-up labels for the corresponding districts. Next, red markers for the malls were added to mark the locations to avoid.

RESULTS

The resulting map provided information on café density and locations of universities and malls.



DISCUSSION

We looked for three qualities when picking out the candidates, based on the map.

1. Districts with less competitors (shaed as yellow, not orange or red).
2. Districts with one or more univeristies (displayed as blue markers).
3. Districts with no mall (displayed as red marker).

Based on these criteria, we recommend the eastern districts, not too far from the center of Seoul: "Seongbuk-gu," "Dongdaemun-gu", "Seongdong-gu".

CONCLUSION

Using various tools, we were able to strategically narrow down the location for the client's new café to three districts. Although our recommendation saved our client a lot of time, it is only the starting point.

Further research could zoom in on the three districts to locate subway stations, measure the foot traffic, compare the rents, and assess the environment. Finding the ideal spot for a new business requires even more details that can only be gained by walking the street.