Identifying Locations for Coffee Shops Investors

Karim Haikal

March 17, 2020

1- Introduction: Business Problem

Identifying a good location to start a coffee shop is a key parameter towards the success of such investment. Thus in this study, we will try to find an optimal location in order to start a Coffee shop in Beirut. Specifically, this report will be targeted to stakeholders interested in opening a Coffee Shop in Beirut, Lebanon.

Since the main customer target of the coffee shop would be university students, we would try to detect locations that are not crowded with coffee shops but as close as possible to as much universities in the area.

We will use our data science powers to generate a few most promising neighborhoods based on this criterion. Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

2- Data

Based on the definition of our problem, we selected two main factors that would influence our decision:

- Number of existing coffee shops in the neighborhood
- Proximity distance to universities in the neighborhood which would be reflected by the proximity to aggregate center of all the existing universities

We decided to use regularly spaced grid of locations, centered around the aggregate center of the existing universities, to define our neighborhoods.

As for the data sets, the following data sources will be needed to extract/generate the required information:

- number of universities and coffee shops in every neighborhood will be obtained using Foursquare API
- Distance between our neighborhoods and the aggregate center of universities
- Number of coffee shops in each neighborhood
- Coordinate of Beirut will be obtained using **Nominatim from geopy.geocoders**

3- Methodology

In this project we will direct our efforts on detecting areas of Beirut that have low coffee shops density and high university density. We will limit our analysis to an area of approximately 2 km around the city center.

In our first step we have collected the required data: location of every coffee shop within 2 km from Beirut center. We have also identified all universities existing in this area (according to Foursquare). Moreover, we have segmented our city with different neighborhoods in order to assess the best location.

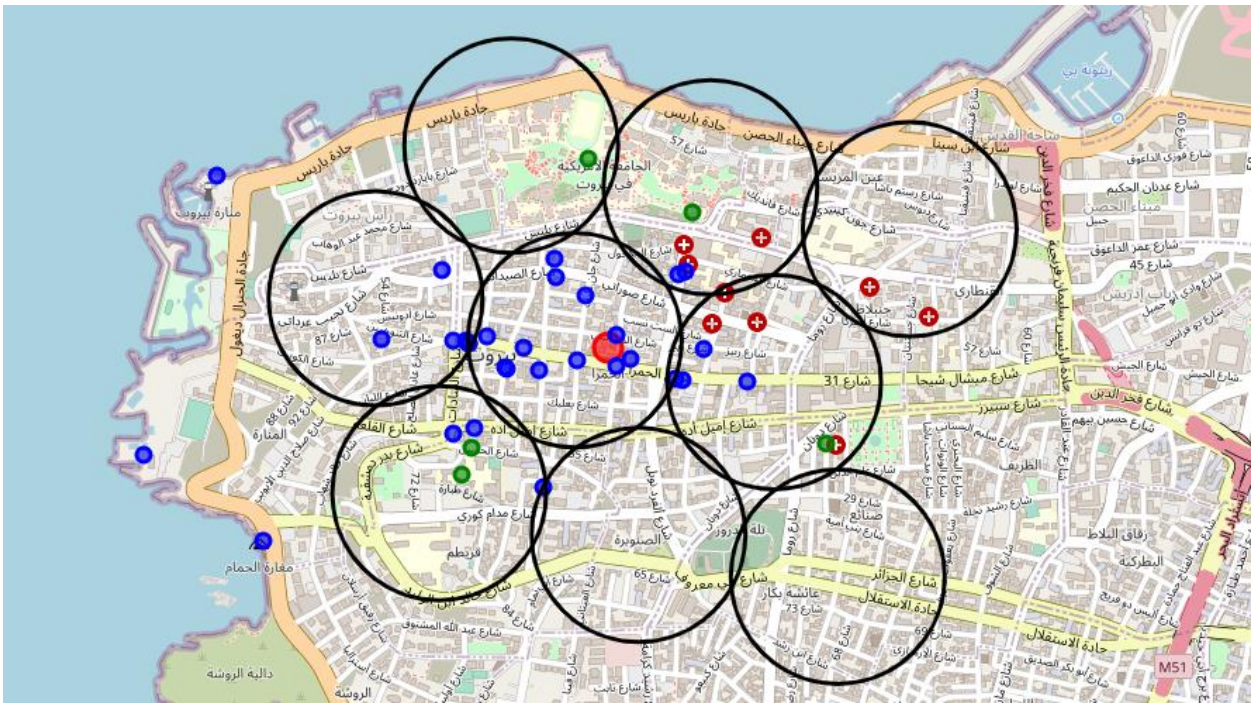

Figure 1: Neighborhood distribution

In the second step of our analysis will be calculating and exploring each neighborhood based on its proximity to the aggregate center of the universities and the density of its existing coffee shops.

Table 1: Main evaluation parameter for each neighborhood

	Latitude	Longitude	X	Y	Distance from center	Number of coffee shops
0	33.892429	35.476874	2.408347e+06	3.945394e+06	655.743852	2
1	33.891373	35.482951	2.408947e+06	3.945394e+06	556.776436	0
2	33.890317	35.489028	2.409547e+06	3.945394e+06	953.939201	0
3	33.897346	35.474931	2.408047e+06	3.945913e+06	700.000000	4
4	33.896290	35.481009	2.408647e+06	3.945913e+06	100.000000	12
5	33.895234	35.487087	2.409247e+06	3.945913e+06	500.000000	4
6	33.901208	35.479067	2.408347e+06	3.946433e+06	655.743852	0
7	33.900151	35.485145	2.408947e+06	3.946433e+06	556.776436	2
8	33.899095	35.491223	2.409547e+06	3.946433e+06	953.939201	0

In third and final step we will conduct an evaluation process in order to assess the best neighborhood for investing. The decision formula will be based on two main factors:

- Proximity to the aggregate center of universities
- Low density of other competitors

The formula will assign a weight of 50% to the selection of each category.

Hence the formula is the following:

$$Evaluation\ score = \left(1 - \frac{Number\ of\ competitors}{Maximum\ number\ of\ competitors}\right) * 50\% + \left(\frac{Minimun\ Distance\ from\ the\ Aggregate\ center}{Distance\ from\ the\ Aggregate\ center}\right) * 50\%$$

Thus the nearest neighborhood for the aggregate center universities will have the highest score regarding the proximity parameter and the neighborhood with least number of coffee shops will have the highest score regarding competition density parameter.

4- Results

Our analysis shows that the optimal area to start a coffee shop is in the south of the city as shown in the following map:

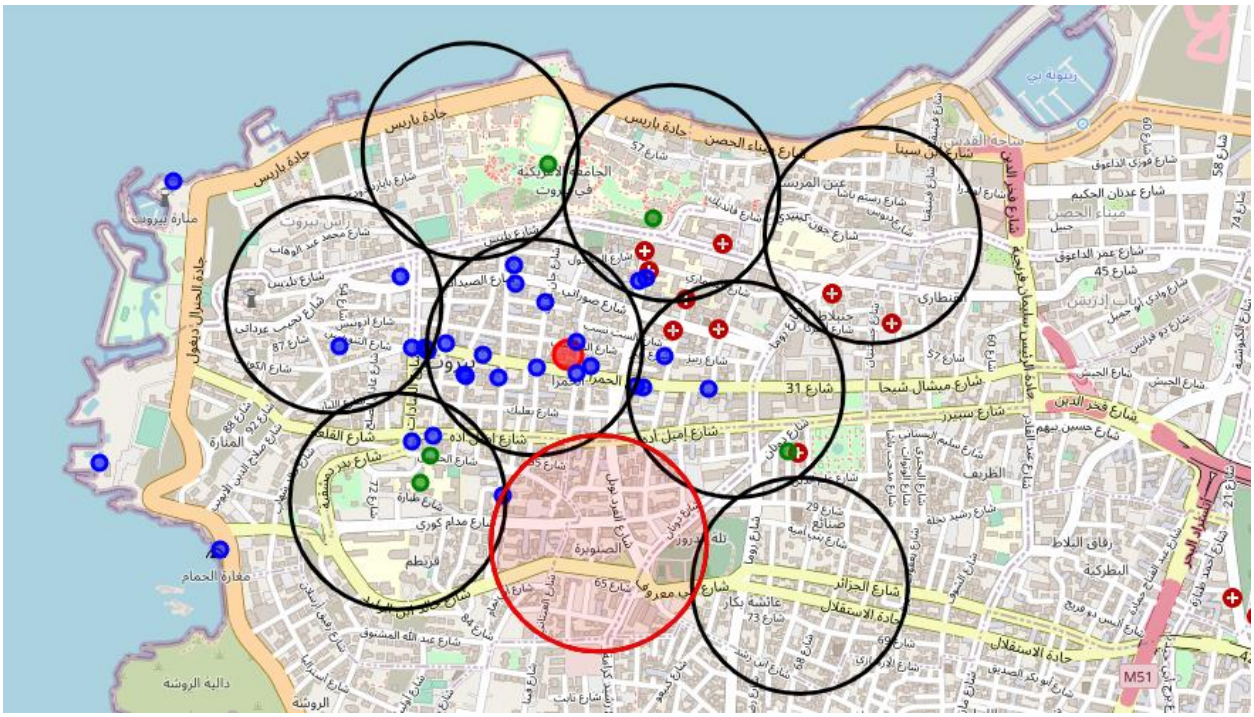

Figure 2: Highlight of the Neighborhood with the highest evaluation score

In fact, this neighborhood has the highest evaluation score based on the suggested formula:

Table 2: List of the top Neighborhoods for investment

	Latitude	Longitude	X	Y	Distance from center	Number of coffee shops	Evaluation Scoring
0	33.891373	35.482951	2.408947e+06	3.945394e+06	556.776436	0	0.589803
1	33.901208	35.479067	2.408347e+06	3.946433e+06	655.743852	0	0.576249
2	33.899095	35.491223	2.409547e+06	3.946433e+06	953.939201	0	0.552414
3	33.890317	35.489028	2.409547e+06	3.945394e+06	953.939201	0	0.552414
4	33.900151	35.485145	2.408947e+06	3.946433e+06	556.776436	2	0.506469

5- Discussion

After assessing the obtained results, we can notice that the decision making procedure is highly affected by the suggested evaluation formula. Thus changing the weight percentages of any of the parameters or adding new evaluation criteria can have impacts on the results. For this reason, it is very important to identify the weight of each parameter taken into consideration.

6- Conclusion

The Purpose of this project was to identify Beirut areas close to the aggregate center of universities with low number of coffee shops competitors in order to aid stakeholders in narrowing down the search for

optimal location for a new coffee shop. We calculated the coffee shop density distribution from Foursquare data and the aggregate center of the existing universities. We have segmented our city in multiple neighborhoods and assessed the competition density in each one along with the proximity of the neighborhood center to the aggregate center of the universities.

The final decision for an on optimal coffee shop location will be made by stakeholders based on additional specific characteristics of neighborhoods and locations in the recommended zone, taking into consideration additional factors like attractiveness of each location, levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighborhood etc.