Capstone Project: Location for future Vietnamese restaurant

TRAN Trung Duc

November 15, 2020

1 Introduction of business problem

Paris is one of the most dynamic cities in Europe. Before the pandemic, every day, millions of people go to the town to work, study, see each other, and enjoy its multi-cultural gastronomy. Therefore, the food and beverage service is an attractive market in the long term, although, at this moment, the pandemic hurts this sector. After the epidemic, I will open a Vietnamese restaurant in Paris. We focus on serving students and young working people. Therefore, the ideal location should be near to universities or co-working spaces. However, the concurrence from other restaurants and other Vietnamese restaurants should be taken into consideration. In this project, I use the "power" of data science to find the ideal location for my future restaurant.

2 Data collection and exploitation

As aforementioned in the previous section, to make a decision, I take into consideration the followings factors:

- There are universities or co-working spaces near to the candidate location.
- Number of restaurants near to the candidate location.
- Number of Vietnamese restaurants near to the candidate location.

For this purpose, I use the following data:

- The location of universities and co-working spaces in Paris. I call these locations the interesting points.
- The number of restaurants (including Vietnamese restaurant) near the universities and co-working spaces. These numbers represent the competition around the interesting points

This data is collected from the Foursquare site. The detail of these data is presented in the remains of this section.

2.1 Universities and coworking spaces

Firstly, I request the Foursquare site to get the locations marked as "University" in Paris. The response gives me a total of 103 results. Each result contains the information on the name, latitude, longitude and venue category of locations marked as University (For example, Figure 1 shows five first results).

	Id	Name	category	latitude	longitude
0	4b522597f964a5205d6b27e3	UPMC – Université Pierre et Marie Curie	University	48.846936	2.354886
1	4e8e9b0adab454671ece48ca	Cours de Civilisation Française de la Sorbonne	University	48.851714	2.347674
2	5384a19b498e22b2a690b1d1	NYU in Paris	University	48.850427	2.346816
3	4cebf24bbaa6a1cd7a10416c	Université Paris I – Panthéon-Sorbonne	University	48.847046	2.343903
4	4b5eb4acf964a520809629e3	Maison de la Recherche	University	48.852440	2.341689

Figure 1: Five first results from the Foursquare response

However, the Foursquare response could contain the locations which are considered quasi similar to "University". I need to check if these locations are suitable for my business. For this purpose, I take a look at the statistical description of results (see Figure 2). Through this description, I found that besides the university category, the result also contains the building marked under other categories: College Auditorium, College Science Building, Student Center. These buildings are all university buildings and have students. Therefore I still consider them interesting points.

category	Id	Name	latitude	longitude
College Auditorium	1	1	1	1
College Science Building	1	1	1	1
General College & University	3	3	3	3
Student Center	1	1	1	1
University	97	97	97	97

Figure 2: Different categories related to “university”

Secondly, I request Foursquare to get the location marked under the “Co-working space” category. Similar to the previous request, the result could contain the location under other categories. I check the statistic description of the result and find some no-related location categories, as shown in Figure 3. They are Business Center, Café, and Car Wash. Because there could be not many target clients for my business in these locations, I remove them from my interesting point list.

category	Id	Name	latitude	longitude
Business Center	1	1	1	1
Café	1	1	1	1
Car Wash	1	1	1	1
Coffee Shop	4	4	4	4
Coworking Space	75	75	75	75
Office	1	1	1	1

Figure 3: Different categories related to “co-working space”

Finally, I have a list of interesting points, including universities and co-working spaces in Paris. These locations are shown in Figure 4

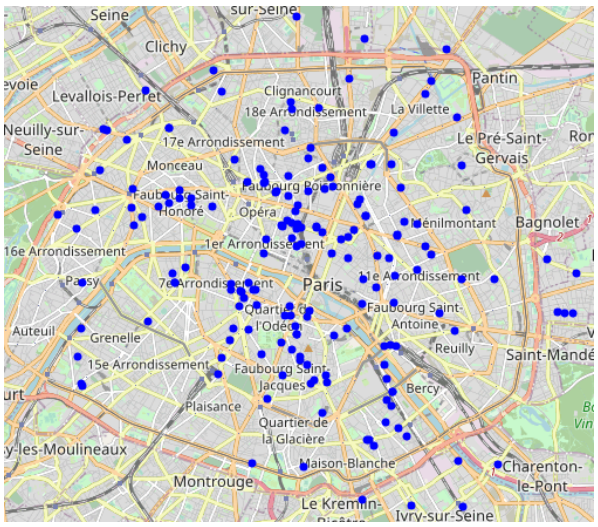

Figure 4: The interesting points on paris map

2.2 Competition data

To measure the competition, I determine the number of Vietnamese restaurants around each interesting point. For each interesting point, I send a request to the Foursquare site to get Vietnamese restaurants within 200m and 800m. The statistic description of all results is shown in Figure 5.

	latitude	longitude	800m_Viet_restaurant	200m_Viet_restaurant
count	178.000000	178.000000	178.000000	178.000000
mean	48.860822	2.345902	7.803371	0.859551
std	0.019297	0.031338	6.815618	1.191990
min	48.813901	2.273655	0.000000	0.000000
25%	48.847867	2.328738	3.000000	0.000000
50%	48.862526	2.345046	5.000000	0.000000
75%	48.873892	2.367483	11.000000	1.000000
max	48.908361	2.425920	36.000000	5.000000

Figure 5: Statistical description for competition

According to the table, there could be 36 Vietnamese restaurants within 800m from an interesting point. This huge number shows very high competition. Figure 7 shows the interesting points that have more than 18 Vietnamese restaurants. The high competition points are in the 13th district because there are many Vietnamese people there. The other competition points locate in the center of the city - the 1st district. It's understandable. Because in the city center, the restaurant density is very high. But I'm amazed by the increased competition in the east of Paris. I don't know much about this area.

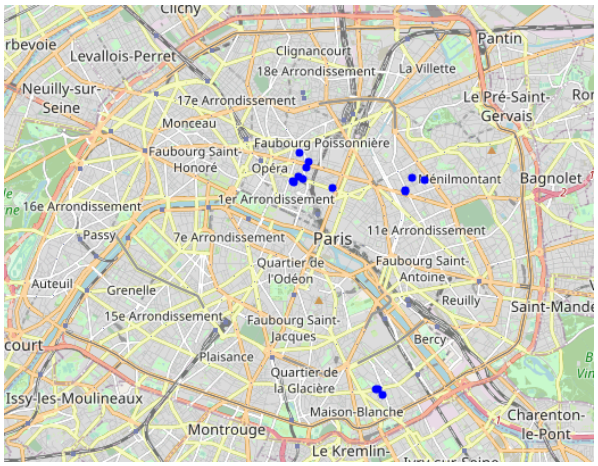

Figure 6: The high competition point

3 Methodology

In this analysis, I aim to discover the areas in Paris near universities and co-working spaces. Because there, I could find my potential clients: students and young people. However, I have considered the competition from other restaurants in these areas (especially from other Vietnamese restaurants). For this purpose, I conduct my analysis based on the following workflow:

- Firstly, I determine the number of Vietnamese restaurants around interesting points(universities and co-working spaces). These numbers represent the competition between existing Vietnamese restaurants. I remove the interesting points with a high level of competition (more than three Vietnamese restaurants within 800m and one within 200m).
- Secondly, I cluster the remaining interesting points into different candidate zones. These zones have a radius of 800m. Inside these zones, the competition between Vietnamese restaurants is low, and the opportunity to reach the target client is still high.

- Thirdly, I consider the competition from other restaurants inside the candidate zones by determining their number of restaurants (all kinds of restaurants). Then, I remove the zone having a high number of restaurants (more than 57 restaurants)
- Finally, I determine the center address of each remaining zones. This address is the starting point to do more further analysis and searches.

4 Result and analysis

In the end, there are 8 addresses that represent the center of the potential zone. Each potential zone is a circle with a radius of 800m. The identified zones locate in four corners of Paris: Nord, South, west, and south. The west zones (16th district) and the east zones (Montreuil borough) are affluent areas of Paris with many offices.

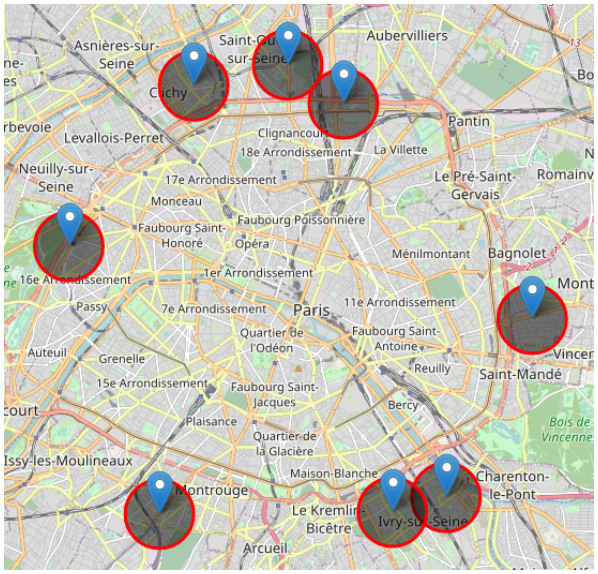

Figure 7: The high competition point

5 Discussion

As the result above, I identify eight potential zones for my future restaurant. These zones are quite far from the city center. It is normal because the competition is very high in the city center with many restaurants, including Vietnamese restaurants. Although there could be fewer activities and habitats in these zones than in the city center, the competition in these zones is lower than in the city center (especially, some interesting points are having more than 30 Vietnamese restaurants around in 1st and 13th district). Moreover, there are co-working spaces or universities within the identified zone. Therefore, I could find the locations for my future restaurant in these zones.

6 Conclusion

In this project, I look for a potential location for my new Vietnamese restaurant in Paris. I focus on areas near universities or coworking spaces, where many young people are my target client. For this purpose, firstly, I identify all universities and coworking space in Paris. I call them as interesting points. Then, I count Vietnamese restaurants near these points (within 800m around and 200m around). The more restaurants there are, the higher competition I have. Therefore, I remove the interesting point with high competition from my list. Next, the remaining interesting points are clustered into 16 zones with a radius of 800m. I call them the potential zone where I could find a good location. Besides the Vietnamese restaurants, I take into account also the competition from other restaurants. I count the number of restaurants presented in the potential zones. The zone has more than 58 restaurants are removed. In the end, it remains eight potential zones with low competition but near enough to

the interesting points. The address of the zone center is the starting point to do further analysis and search. Among the identified zone, I prefer to do more analysis in the "rich zone" such as the 16th district and Montreuil borough firsts.