

Finding the best location for a healthy food spot in Paris

Study Report

1. Introduction/Business Problem

The basis of this study is to help the customer - a group of entrepreneurs and investors planning to open a new chain of healthy food restaurants in Paris - to find the best location for their first spot.

The main aim of the enterprise is to promote healthy food and lifestyle and make it popular among the young people.

So, the main target audience is supposed to be teenagers, students and office workers.

The client is also interested in expanding the company by concluding contracts with public and private institutions, such as schools, universities, private enterprises, etc. - offering them special conditions to attract students and employees on a continuous basis.

The objective of the study is to locate and recommend to the customer, which arrondissement of Paris will be the best choice to start their business.

It could also be interesting for other potential investors looking to start a new food business for a similar target audience in Paris.

2. Data

The necessary information required for the study will come from the following sources:

Open data from <https://opendata.paris.fr/explore/dataset/arrondissements/> to get the geo information about administrative districts of Paris - arrondissements:

- Arrondissements number
- Arrondissements name
- Geographic coordinates (latitude, longitude)

Foursquare API (<https://developer.foursquare.com/>) to collect information on the target audience venues: high schools, universities and offices, as well as other venues/competitors in the district.

After that we will leverage the data in order to find the most appropriate location to launch the client's business.

3. Methodology

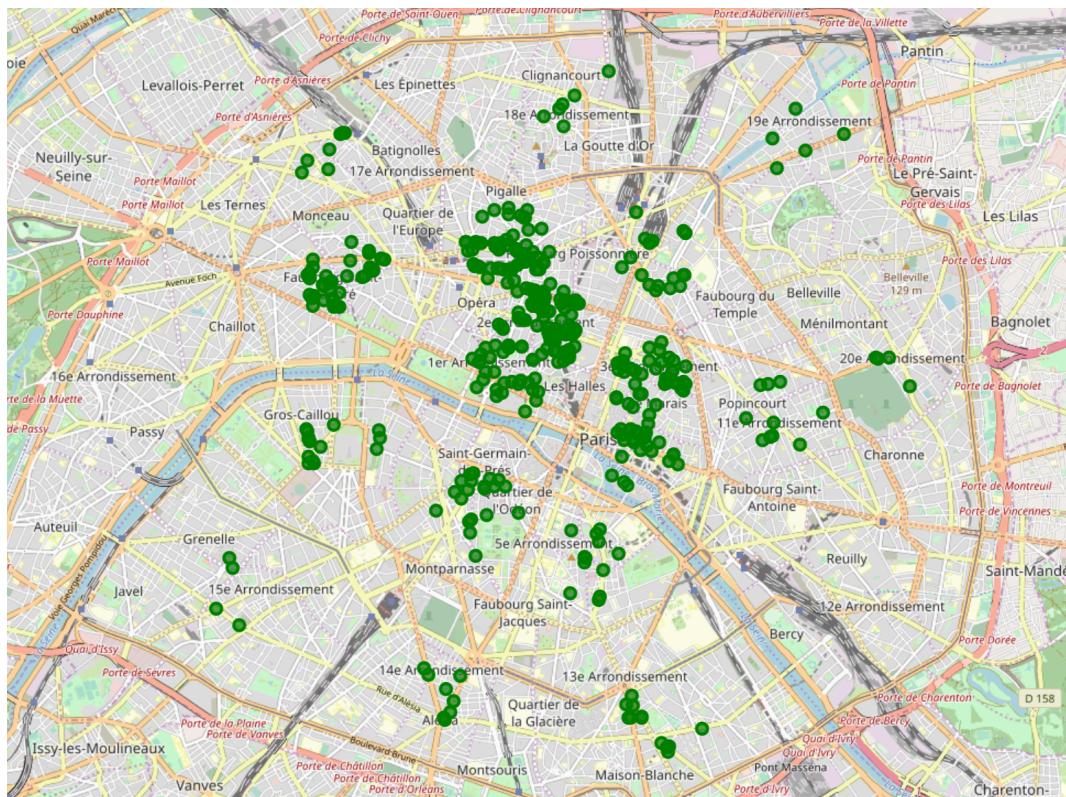
In order to find the best target location, we will explore the Paris districts by segmenting the data and conducting descriptive analysis using Panda. The Foursquare API will be used to get the data regarding potential competitors and target audiences.

- 1) The arrondissements data was pulled into from the Opendata Paris JSON File to create a dataframe containing the arrondissements names and numbers and their geographic coordinates

ArrNumber	ArrName	Arrondissement	Latitude	Longitude
1	1er Ardt	Louvre	48.862563	2.336443
2	2ème Ardt	Bourse	48.868279	2.342803
3	3ème Ardt	Temple	48.862872	2.360001
4	4ème Ardt	Hôtel-de-Ville	48.854341	2.357630
5	5ème Ardt	Panthéon	48.844443	2.350715
6	6ème Ardt	Luxembourg	48.849130	2.332898
7	7ème Ardt	Palais-Bourbon	48.856174	2.312188
8	8ème Ardt	Élysée	48.872721	2.312554
9	9ème Ardt	Opéra	48.877164	2.337458
10	10ème Ardt	Entrepôt	48.876130	2.360728
11	11ème Ardt	Popincourt	48.859059	2.380058
12	12ème Ardt	Reuilly	48.834974	2.421325
13	13ème Ardt	Gobelins	48.828388	2.362272
14	14ème Ardt	Observatoire	48.829245	2.326542
15	15ème Ardt	Vaugirard	48.840085	2.292826
16	16ème Ardt	Passy	48.860392	2.261971
17	17ème Ardt	Batignolles-Monceau	48.887327	2.306777
18	18ème Ardt	Buttes-Montmartre	48.892569	2.348161
19	19ème Ardt	Buttes-Chaumont	48.887076	2.384821
20	20ème Ardt	Ménilmontant	48.863461	2.401188

- 2) For each arrondissement, information about the potential competitors venues have been collected from Foursquare. Since there's no special separate category for a healthy food spot or restaurant, the following categories were used as the ones that can have some intersections with the target category:

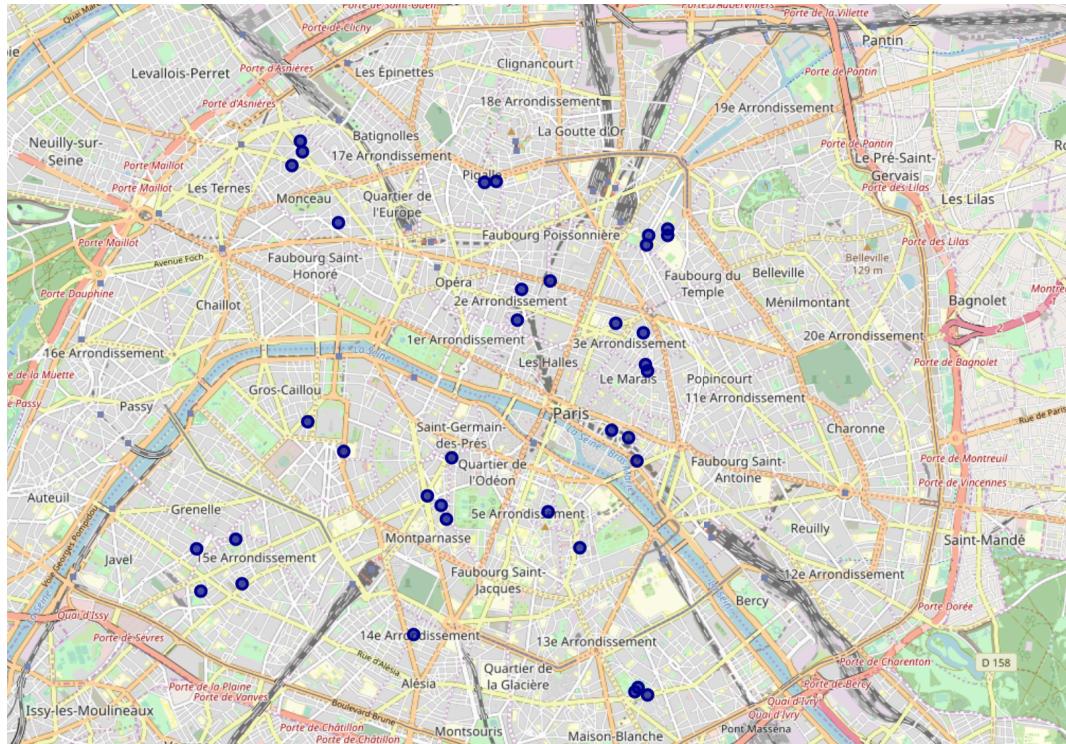
- Salad Place
- Sandwich Place
- Soup Place
- Vegetarian / Vegan Restaurant
- Health Food Store



- 3) Target audience data was taken from Foursquare as well: high schools, colleges, universities and offices locations were defined.

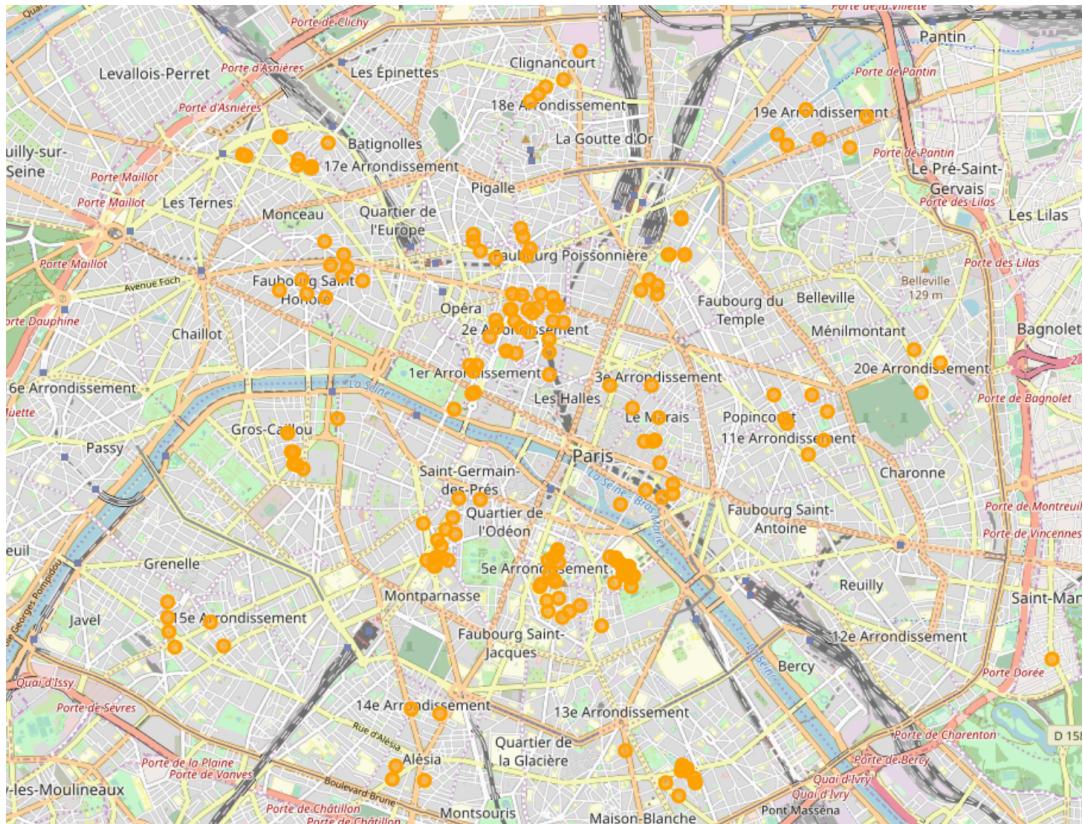
Schools:

ArrName	Arrondissement	Arrondissement Latitude	Arrondissement Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2ème Ardt	Bourse	48.868279	2.342803	Lycée Jean-Baptiste Lulli	48.869871	2.343470	High School
2ème Ardt	Bourse	48.868279	2.342803	Cours Lafayette	48.866667	2.342775	High School
2ème Ardt	Bourse	48.868279	2.342803	Lycée Privé Edgar Poe	48.870815	2.348110	High School
3ème Ardt	Temple	48.862872	2.360001	Lycée François Truffaut	48.861896	2.363554	High School
3ème Ardt	Temple	48.862872	2.360001	Lycée Simone Weil	48.861239	2.363970	High School



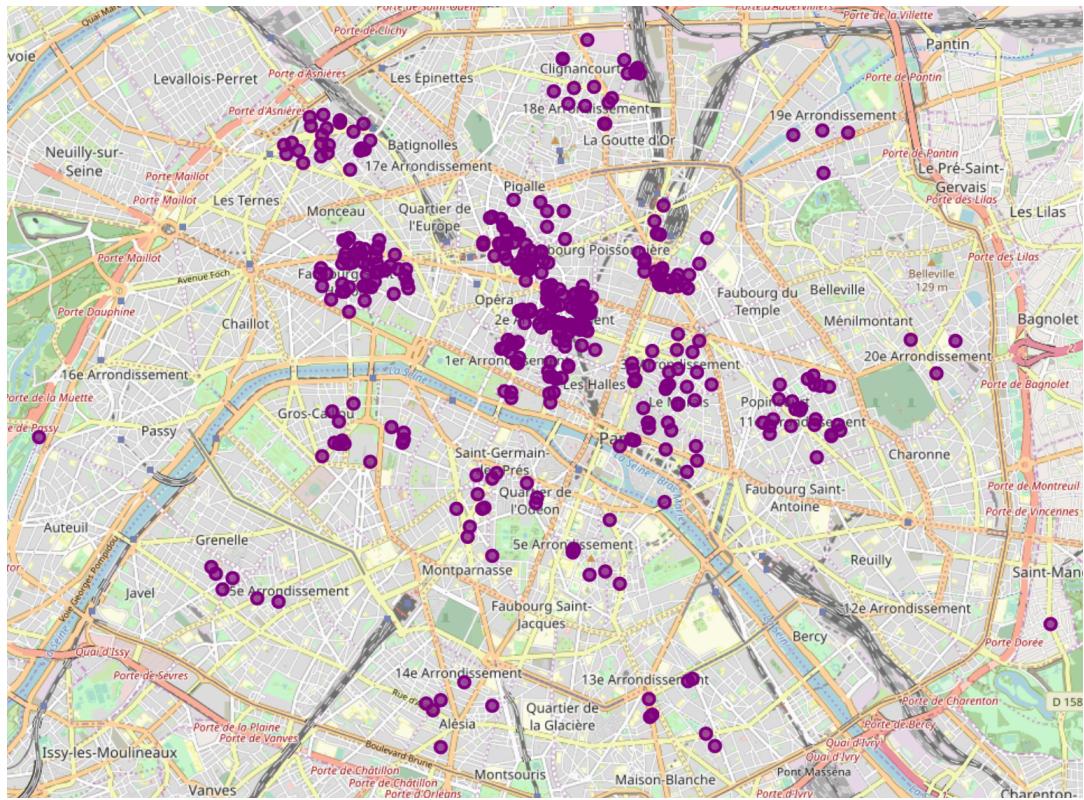
Colleges and universities:

ArrName	Arrondissement	Arrondissement Latitude	Arrondissement Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2ème Ardt	Bourse	48.868279	2.342803	Institut National d'Histoire de l'Art (INHA)	48.866680	2.338986	General College & University
2ème Ardt	Bourse	48.868279	2.342803	École de la Chambre Syndicale de la Couture	48.868478	2.342351	University
2ème Ardt	Bourse	48.868279	2.342803	EEMI	48.868944	2.341118	University
2ème Ardt	Bourse	48.868279	2.342803	Esmod Isem	48.867723	2.345213	University
2ème Ardt	Bourse	48.868279	2.342803	EEMI Passage Des Panoramas	48.870299	2.342846	University



Offices:

ArrName	Arrondissement	Arrondissement Latitude	Arrondissement Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2ème Ardt	Bourse	48.868279	2.342803	FCINQ	48.868473	2.341914	Office
2ème Ardt	Bourse	48.868279	2.342803	Webloyalty France webloyalty.fr	48.868120	2.344235	Office
2ème Ardt	Bourse	48.868279	2.342803	Autorité des Marchés Financiers (AMF)	48.869440	2.340892	Office
2ème Ardt	Bourse	48.868279	2.342803	Red Bull HQ	48.866644	2.342369	Office
2ème Ardt	Bourse	48.868279	2.342803	M&C Saatchi Corporate	48.868480	2.341972	Office



- 4) For each arrondissement, the sums of the target institutions and potential competitors places were computed.
- 5) For each of these four categories, a weight has been defined - positive or negative, depending on the influence on the result (the payment capacity of the target audiences were taken into consideration as well):
 - -1 for potential competitors spots
 - 1 for schools
 - 1.5 for colleges and universities
 - 2 for offices
- 6) Finally, a score was computed for each arrondissement as the weighted sum of the number of venues in each of the four categories (potential competitors, schools, colleges and universities, offices).

4. Results

The 2nd arrondissement - Bourse - turned out to be the district with the best score (110), having a great number of potential customers: college and university students as well as workers from the nearby offices, and a low level of competition.

ArrName	Arrondissement	Score
2ème Ardt	Bourse	110.0
8ème Ardt	Élysée	71.0
17ème Ardt	Batignolles-Monceau	64.0
5ème Ardt	Panthéon	57.0
10ème Ardt	Entrepôt	54.0
11ème Ardt	Popincourt	46.5
18ème Ardt	Buttes-Montmartre	45.5
9ème Ardt	Opéra	44.0
6ème Ardt	Luxembourg	32.5
7ème Ardt	Palais-Bourbon	26.0
1er Ardt	Louvre	25.5
13ème Ardt	Gobelins	23.0
15ème Ardt	Vaugirard	21.0
4ème Ardt	Hôtel-de-Ville	18.0
3ème Ardt	Temple	13.0
14ème Ardt	Observatoire	12.5
19ème Ardt	Buttes-Chaumont	12.0
20ème Ardt	Ménilmontant	6.5
12ème Ardt	Reuilly	3.5
16ème Ardt	Passy	2.0

5. Discussion

From the results discovered and presented, the following observations and recommendations can be made:

- Based on the criteria given by the customer and the data, the main district recommendation would be for the 2nd arrondissement – Bourse. With a total score of 110, this area is a prime location for a flagship healthy food spot/restaurant.
- A secondary recommendation can be made for the 8th arrondissement – Élysée, ranking in the top five total score assessment, as well as having less competition.
- For more accurate results, a deeper analysis can be done, taking into consideration more categories (e.g. breakfast spots as potential competitors or related areas, such as yoga clubs, in order to find the target audience).
- Within each arrondissement, a distance between all the venues can also been taken into account in order to find a location with the greatest number of potential customers.

6. Conclusion

In conclusion, the scope of present analysis is to some degree limited. The restaurant industry is ever changing, and the information used may be dated due to relying on user information via

Foursquare. However, the model created can easily be replicated more than once with constantly updated data via the Foursquare API. As already mentioned, for more accurate and refined results, a deeper analysis is required. With the data analysed and scoring system defined by the client's targets, we stand by the recommendations made.