Capstone Project – Open a coffee shop in Paris

Table of contents:

1. Introduction

1.1 Background

I've decided to study the coffee shops in Paris. Why Paris and why the coffee shops?
First of all, I'm french and when I go in Paris, I always see the coffee shops crowded.
Paris is the center of French economy, politics, traffic and culture. Paris is one of the most multicultural cities. It's the capital and the largest city of France. Indeed, Parisian coffees serve as a center of social and culinary life in Paris. They have existed since the 17th century, and serve as the meeting place, neighborhood hub, conversation matrix, meeting spot, and networking source, a place to relax or to refuel - the social and political pulse of the city. Parisian coffees show the Parisian way of sitting undisturbed for a couple of hours, watching things happening and people going by.

According to the SOFRES survey 13% of the French population visit several times a week a coffee shop (which 4% visit almost daily). Also, the survey shows the kind of people who frequent coffee shops. They are primarily men, young people and workers. So, we can see that all sections of the population go to the coffee shops and especially active people.

1.2 Problem

If someone is looking to open a coffee shop in Paris, it is interesting to study many features very important. Before creating any business, we have to look at it three different standpoints: marketing, financial and operational.
One of the most important feature is the location. Where will coffee shop be? So, the mainly problem is: *Where can I open a coffee shop in Paris?*

Here our target are the active people. We can study three important points:
 – the rental cost according to the place
 – the place activity
 – the potential customers attracted by coffee shops
 – the number of coffee shops in each borough to study the potential competition in each borough

1.3 Interest

This study in very interesting for all people who looking to open a coffee shop in Paris. Also, I think it can be interesting for people who just looking for open a business in Paris because in this study I'll give many information about boroughs in Paris. These information are interesting for someone who looking to open other businesses like restaurant, store and other in Paris.

2. Data acquisition and cleaning

2.1 Data sources

I need the average commercial rental price in Paris.
I've got the data in this website: https://www.localcommercial.net/estimerloyer/38080/paris-20-arrondissement.html
Then, I created a csv file that contains postal code and average commercial rental price.

With foursquare API, I've got:

- The coffee shop with their address, postal code, name, longitude and latitude.
- The nearby venues in Paris to see many information:
Where are the active places in Paris?
What are the most common venues in terms of borough?
What are the boroughs where coffee is a most common venue?
When coffee is a most common venue in a borough, what are the other most common venues?
What is the kind of people who frequent each borough?

I've got the features to create a map with the borough of Paris with a geoJSON file: arrondissements.geojson downloaded in this page :
https://opendata.paris.fr/explore/dataset/arrondissements/export/?disjunctive.c_ar&disjunctive.c_arinsee&disjunctive.l_ar&location=13,48.85156,2.32327

2.2 Data creation

I created a dataframe with information about coffee shops in Paris with foursquare API. I've got one hundred coffee shops with the "get venue recommendations endpoint". I think it's interesting to get the coffees with recommendations to take example of business that work. So, I didn't take coffee without recommendations. I think that to open a business that work, we have to take example of good coffee shops. With the csv with average commercial rental price, I created a dataframe which contains the average commercial rental price depending on the borough in Paris.
I merged this two dataframes to one dataframe named *merged_inner*. I created another dataframe named *dfcount* which contains the number of coffee shops for each borough in Paris. I added to the *dfcount* the column with the average commercial rental price. I did this because I was looking to see if there had any relationship between the rental price and the number of coffee shop in the place.
I also created a df with the geographic information about the borough in Paris with the geojson file. I got these information to have a good vision of each borough in a choropleth map.

To answer to the problem *Where can I open a coffee shop in Paris*? an essential thing is to study each borough. To open a business you have to know many things about each place, like people who go in this place to study the potential clients of my business. Who is the target, what kind of clients can be attracted to my business according to the place ? To study each borough, I created a dataframe named *paris_venues*. I got this data with foursquare API, I got all venues in each borough. Then I got the most common venues in each borough to have an idea about people who haunting each place and see where the business could work.

I created a dataframe named *venues_sorted* with the ten most common venues in each borough. I join this dataframe with the *merged_inner* dataframe where we have all information about each coffee to see what are the top ten of most common places where we have as much coffee shops as possible.

2.3 Feature selection

 With foursquare API I got many features but I selected only the interesting features for my analyze. So, when I've got the hundred coffee shops with recommendations in Paris, I selected the name of a coffee, the latitude, the longitude, the postal code and the address. I also selected the city if I have another city by error. The geographic information about these coffees allows us to view there in the map and to link with other dataframe when I've got features about each borough. I added the average commercial rental price by m2 to this df for each borough. I divided by twelve the rent to get the rent by month because I think that it's more significant than by year.

Coffee shop	Name of coffee shop
Coffee Latitude	Geographical information about the coffee shop
Coffee longitude	Geographical information about the coffee shop
Postal Code	The postal code of the coffee shop
City	The city of the coffee shop
Address	The address of the coffee shop
average commercial rental price by m2	The average commercial rental price by month by m2 of the coffee shop

When I created the *paris_venues* dataframe I selected the name, the latitude, the longitude, the postal code, the city, the address and the venue category.

Venues	The name of venue
Venues Latitude	The geographic location of venue
Venues Longitude	The geographic location of venue
Venues Postal code	The borough of venue
Venues City	The city of venue
Venues Address	The address of venue
Venues Category	The category name of venues (example: coffee shop)

I've selected the geographic information about the venues to see what is the most activity in each borough, what kind of people go in each borough, what are the boroughs where coffee shops work well, where are the active place?

3. Exploratory Data Analysis

3.1. Coffee shop

I grouped data of merged inner by postal code and count the number of coffee shop in each borough and sort value by ascending order. I've noticed that three boroughs don't have coffee shop with recommendations, seven boroughs contain minus by five coffee shops and the others contain more than five coffee shops with recommendations.

	Postal Code	Coffee shop
0	75016	1
1	75015	1
2	75014	1
3	75017	2
4	75019	3
5	75018	4
6	75002	5
7	75009	6
8	75007	6
9	75011	7
10	75006	7
11	75005	8
12	75004	8
13	75008	8
14	75003	9
15	75001	9
16	75010	11

The boroughs with the most coffee shop are Paris 10th, 1e, 3th, 8th, 4th, 5th.
The three boroughs without coffee shop recommendations are Paris 12th, 13th, 20th.
The borough with only one coffee shop with recommendation are Paris 16th, 15th, 14th.

Let's try to explain this observation with, first of all, the average commercial rental price of each borough.

3.2 Average commercial rental price

Let's see in a barchart the average commercial rental price by borough.
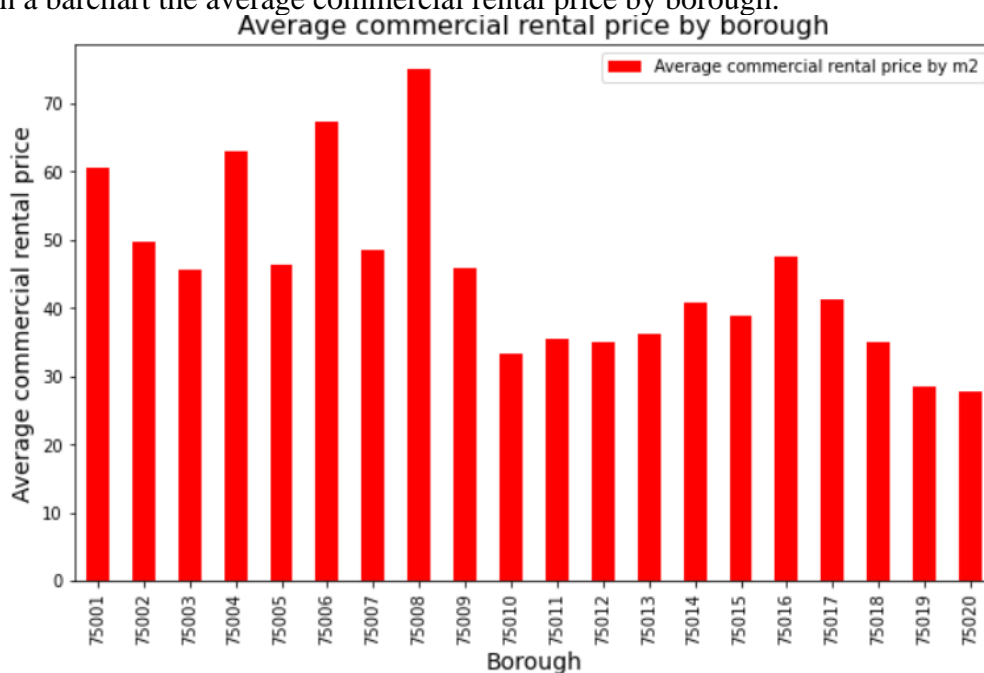
We can see that the first nine boroughs have a higher commercial rental price than Paris 10 to Paris 20.

More we are close to the Parisian center more the rental price is high.
Now let's see the number of coffee shops according to the average commercial rental price.

	Postal Code	Coffee shop	Average commercial rental price by m2
0	75020	1	27.833333
1	75016	1	47.666667
2	75015	1	38.833333
3	75014	1	40.833333
4	75017	2	41.250000
5	75019	3	28.500000
6	75018	4	34.916667
7	75006	4	67.333333
8	75002	6	49.750000
9	75007	6	48.500000
10	75001	7	60.666667
11	75009	7	45.833333
12	75005	8	46.250000
13	75011	8	35.500000
14	75004	9	62.916667
15	75003	9	45.583333
16	75008	10	75.000000
17	75010	11	33.333333

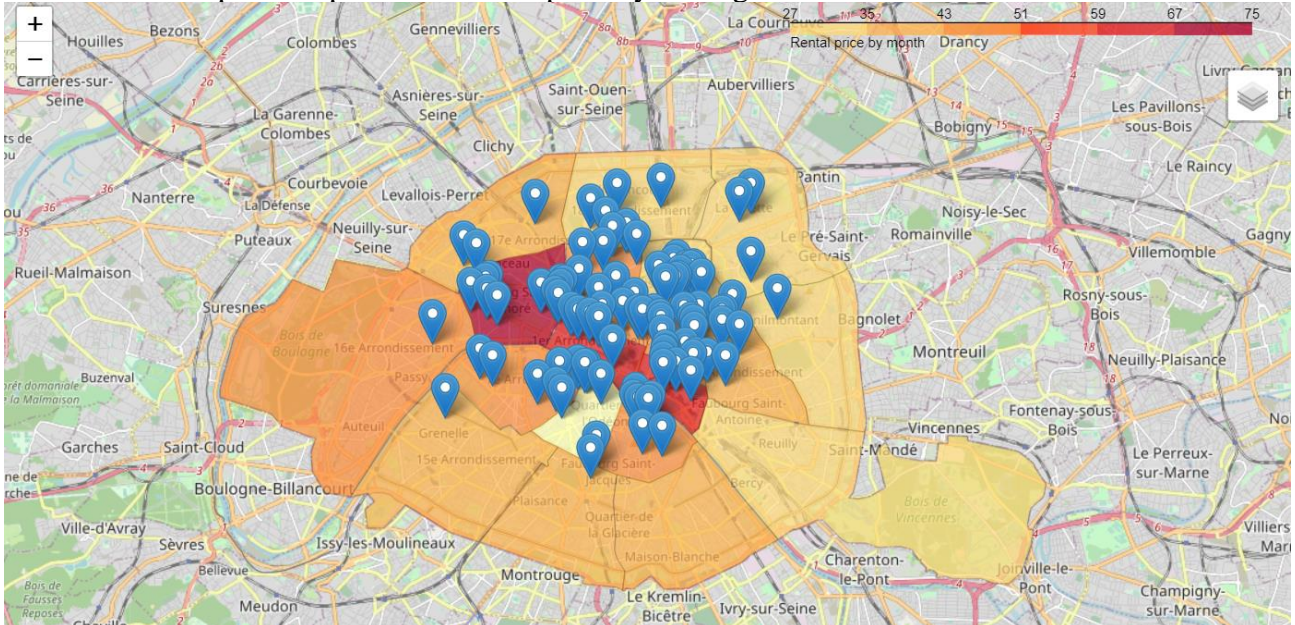We can see that the borough with the most coffee shop is Paris 10 with the third lowest commercial rental price.
Let's see the other boroughs with the most coffee shop with the scatter plot.

Rental price in terms of number of coffee shop in Paris
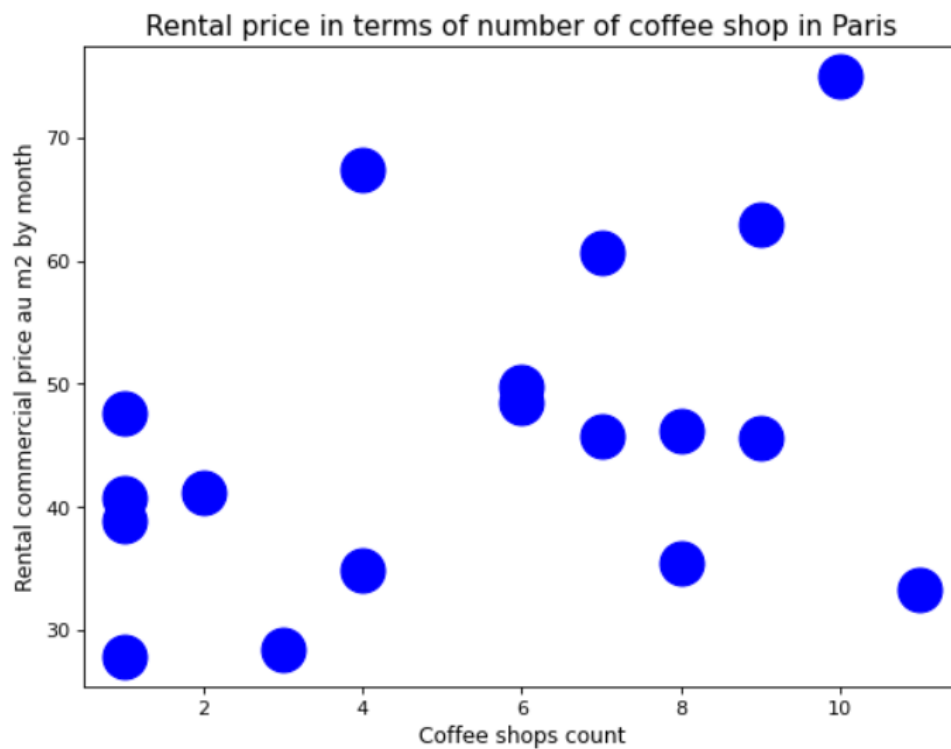
We see that the places with many coffee shops have a middle rent and higher rent and the place with a little number of coffees have generally low rent.

So, we can say that there is a relationship between the number of coffee shops by borough and the commercial rental price by borough.

I created a choropleth map with the rental price by borough.

We can clearly see again that the borough with highest and middle rent contains most coffee shop with recommendations than the other borough. We can also see that the coffee shops are concentrated at the center of Paris. More we are close to Parisian suburb and more the rent is low. So, we can say that the coffee shops work better in the boroughs at the Parisian center than the borough close to Parisian suburb.

Now let's study the kind of people that attend each borough, to have an idea about the potential clientele of coffee shop.

3.3 Analyze of each borough

When I created a dataframe *paris_venues* with the nearby venues in Paris. I obtained ten thousand of results, a huge number!
I counted the number of venues by borough to see the active places.

Venues Postal Code	Venues	Venues Latitude	Venues Longitude	Venues City	Venues Address	Venues Category
75001	1300	1300	1300	1300	1300	1300
75002	700	700	700	700	700	700
75003	1200	1200	1200	1200	1200	1200
75004	1800	1800	1800	1800	1800	1800
75005	1000	1000	1000	1000	1000	1000
75006	1800	1800	1800	1800	1800	1800
75007	200	200	200	200	200	200
75009	100	100	100	100	100	100
75010	600	600	600	600	600	600
75011	1200	1200	1200	1200	1200	1200
75012	100	100	100	100	100	100

We can see that they are five very active boroughs which are: Paris 6th, 4th, 1e, 3th, 11th. These five boroughs have many coffee shops according to my 3.1 point about coffee shops. That seems to be coherent, more an area is active, more there will be many coffee shops in this area.

Let's see the top ten of most common venues by borough.

	Venues Postal Code	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	75001	Plaza	Art Museum	Historic Site	Women's Store	Exhibit	Fountain	Pedestrian Plaza	Garden	Udon Restaurant	Toy / Game Store
1	75002	Hotel	Plaza	Wine Bar	Historic Site	Beer Bar	Italian Restaurant	Pastry Shop	Church	Exhibit	Electronics Store
2	75003	Sandwich Place	Bubble Tea Shop	Bookstore	French Restaurant	Cocktail Bar	Coffee Shop	Supermarket	Farmers Market	Wine Bar	Vietnamese Restaurant
3	75004	Plaza	Souvenir Shop	Seafood Restaurant	Israeli Restaurant	Café	Cocktail Bar	Ice Cream Shop	Hotel	Restaurant	Church
4	75005	Portuguese Restaurant	Indie Movie Theater	French Restaurant	Comic Shop	Botanical Garden	Bookstore	Bistro	Bakery	Monument / Landmark	Dessert Shop
5	75006	Wine Bar	Plaza	Seafood Restaurant	Garden	Restaurant	French Restaurant	Electronics Store	Ice Cream Shop	Cocktail Bar	Clothing Store
6	75007	Art Gallery	Art Museum	Church	French Restaurant	Fountain	Farmers Market	Exhibit	Electronics Store	Dessert Shop	Dance Studio
7	75009	Hotel	Indie Movie Theater	Women's Store	Church	Fountain	Farmers Market	Exhibit	Electronics Store	Dessert Shop	Dance Studio
8	75010	Vegetarian / Vegan Restaurant	Hotel	Bakery	Breakfast Spot	Cocktail Bar	Cheese Shop	Women's Store	Clothing Store	Fountain	Farmers Market
9	75011	Italian Restaurant	Bakery	Speakeasy	Dance Studio	French Restaurant	BBQ Joint	Trattoria/Osteria	Tapas Restaurant	Coffee Shop	Bookstore
10	75012	Asian Restaurant	Women's Store	Church	French Restaurant	Fountain	Farmers Market	Exhibit	Electronics Store	Dessert Shop	Dance Studio

We can see three boroughs with coffee shop in the top of most common venues. In Paris 3th, Coffee shops are the 6th Common venue. We can see others most common venues in this borough which are sandwich place, bubble tea shop, bookstore, restaurant and supermarket. The kind of people who must are in this borough are students and workers.

In Paris 4th the coffees are the 5th most common venue. The other most common venues are plaza, souvenir shop, restaurant, hotel and church. Many historic place and shop where we can generally find tourists.

In Paris 11th the coffee shops are the 9th most common venue. The other most common venues are the restaurant, dance studio, bookstore. People going in the dance studio are probably resident of Paris because generally, the tourists don't stay long enough to take dance courses. The restaurant or bookstore can be frequented by workers or students.
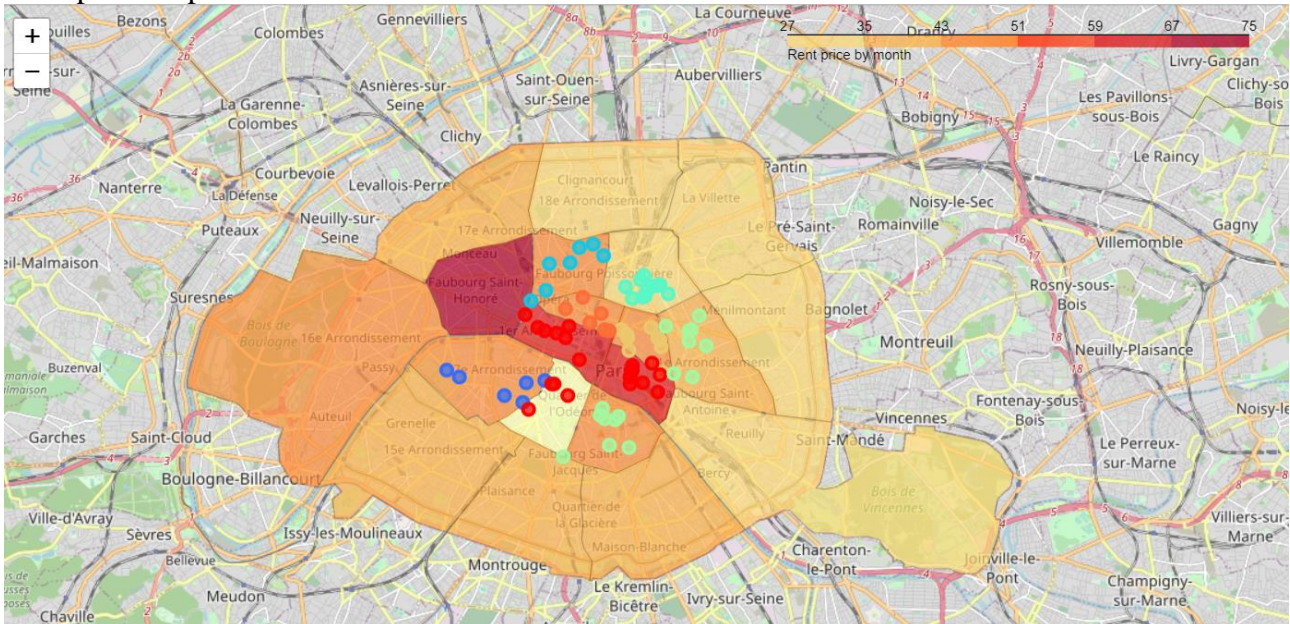
So, we can conclude that in most cases the kind of client attracted by coffee shops are students and workers.

4. **Cluster Coffee shops**

4.1 **Creation of cluster**

I ran k-means to cluster the coffee shop into eight clusters. We can see each cluster in the following choropleth map.

Let's examine each cluster.

4.2 **Analyse of clusters**

Cluster 0:

It's represented by the red circles in the map of cluster.
The coffee shop in this cluster are in the borough with highest commercial rental price concentrated at the center of Paris. The boroughs are Paris 4th, Paris 6th and Paris 1e.

The most common venues in those boroughs are restaurant, plaza, museum, historic site, clothing store, game store, wine bar, etc... I think that it is very active place with all kinds of people (tourists, workers, students, etc...).

Cluster 2:

It's represented by the dark blue circles in the map. The cluster is concentrated in Paris 7th. The commercial rental price is middle, and the tourists must frequent this borough. Indeed, the most common venues are art gallery and art museum, fountain, exhibit.

Cluster 3:

It's represented by light blue circles in the map. The cluster is concentrated in Paris 9th. The commercial rental price is middle. The place must be mostly frequented by many tourists. Indeed,

the first most common venue are the hotel. Among the most venue we also find Theater, church, women's store, exhibit...

Cluster 4:

It's represented by cyan circles in the map. The cluster is concentrated in Paris 11th and Paris 5th. The commercial rental prices are low and middle. The boroughs must be mostly frequented by workers and students. The most common venues are restaurant, bookstore. Among the most common venues in Paris 11th we find the coffee shop.

Cluster 5:

It's represented by light green circles in the map. The cluster is concentrated in Paris 10th. The borough must be mostly frequented by tourists because the first most common venue is the hotel. We can also see an interesting thing: the people in this borough are big consumer because we find in the most common venue, bakery, bar, women's store, clothing store, etc...

Cluster 6:

It's represented by the yellow circles in the map. The cluster is concentrated in Paris 3th. The commercial rental price is middle. The borough must be mostly frequented by students and workers as I said (see 3.3 point). Among the most common venues we find coffee shops.

Cluster 7:

It's represented by the orange circles in the map. The cluster is concentrated in Paris 2th. The commercial rental price is middle high. The borough must be mostly frequented by tourists. The most common venues are hotel, plaza, historic site, etc...

To summarize my study, I created a board which combine all information to choose a borough where it's interesting to open a coffee depending on your budget, your target (the clientele), the activity of each borough, the most coffee venue by borough, the competition with other coffee (the coffee shop number by borough), the activity of the borough (active place or no), the borough's cluster. You can see that I choose only the first eleven borough because the other borough I consider that I don't have enough information about other boroughs. There is not enough coffee shop, it's not active borough so I think it's not interesting to open a coffee shop in this borough.

	Price	Potential client	Coffee Venue	Coffee number	Active	Cluster
75001	High	All kind	x	High	Very active	0
75002	Middle	tourists	x	Middle	Active	7
75003	Middle	Students/ workers	✓	High	Very active	6
75004	High	All kind	✓	Middle	Very very active	0
75005	Middle	Students/ workers	x	Middle	Very active	5

75006	Middle	All kind	x	Middle	Very very active	0
75007	Middle	tourists	x	middle	Not very active	2
75008	Very high	---------------	x	high	---------------	---------------
75009	Middle	tourists	x	middle	Not very active	3
75010	Middle low	tourists	x	high	active	4
75011	Middle low	Students/ workers	✓	middle	Very active	5

5. Conclusion

To conclude, the best area to open a coffee shop for me is Paris 11. Indeed, according to the above board, it's the best compromise. The rental price is middle low, the potentiel clients are the workers and students, one of the ten most venue is the coffee shop, the number of coffee shop with recommendations is middle so there will not have highly competition and the place is very active so there is many people whose frequented this borough. It's only a suggestion according to the parameters studied in this report. I think that each person has a different target. Someone can have a high budget and wants to attract more tourists, so it's not the best choice to choose Paris 11 for this person.
I think that with this study someone who wants to open a coffee shop in Paris can answer, depending on his budget, target, etc, to the problem "*where can I open a coffee shop in Paris?*".
So, business owner I let you this choice and I hope that this study will can help you!

6. Future direction

To have an idea more accurate I could study what kind of coffee open according to the place. For example, the reviews of each coffee shop to see what type of coffee shop works depending on the clientele and the area. We can study the old of client attracted according to the kind of coffee (working coffee with wifi, cosy or other...).