

Find the best neighborhood in Paris to setup offices for a startup

1. Introduction

1.1 Background

As a founder of a new startup, a friend of mine is looking to setup a new office in Paris. The company is growing up very fast and the need of hiring new people is critical. He would like to hire talented people who have been newly graduated. He's looking as well to offer internship for students. Therefore, it's important for the company to be attractive to students and young peoples. One way to attract talents is to have nice offices located in a neighborhood with a good quality of life.

1.2 Problem

The problem is to find the best neighborhood in Paris to setup the offices. The neighborhood must be attractive for young people in order to give them some interest on the company. This project aims to find the best places by selecting neighborhood on criteria important for students and newly graduated people. I focused on the following criteria as I supposed they were important for the targeted audience. A neighborhood is supposed to be attractive to young people when there is a lot of:

- Restaurants and Food venues. Students like to go outside to grab their meals
- Nightlife Spots and Bars to chill out with friends and meet new people
- Outdoors and Recreational venues to relax and playing sports

2. Data acquisition and cleaning

2.1 Data sources

In order to have the best dataset for running analysis, data is collected from different sources. The first source to get Paris neighborhood information is from the online platform <http://opendata.paris.fr>. This service provides public and general information on the city of Paris.

To expand the first data set I used Wikipedia to get population of the different neighborhoods. https://fr.wikipedia.org/wiki/Liste_des_quartiers_administratifs_de_Paris

Then I used Foursquare to get information on the venues. I used 2 services. The first one to get venues for each neighborhood and the second one to get categories and parent categories. There are multiple level of categories and I used the <https://developer.foursquare.com/docs/resources/categories> page to get the Foursquare Venue Category Hierarchy. In my analyses I used only the top categories (for instance 'Arts & Entertainment', 'College & University').

Finally, I used another web site to get the Young Ratio. We define the Young Ratio as the percentage of people with age range between 20 and 39. We get data from the French Institute of Statistics and Economic Studies <https://www.insee.fr/fr/statistiques/zones/4177137>.

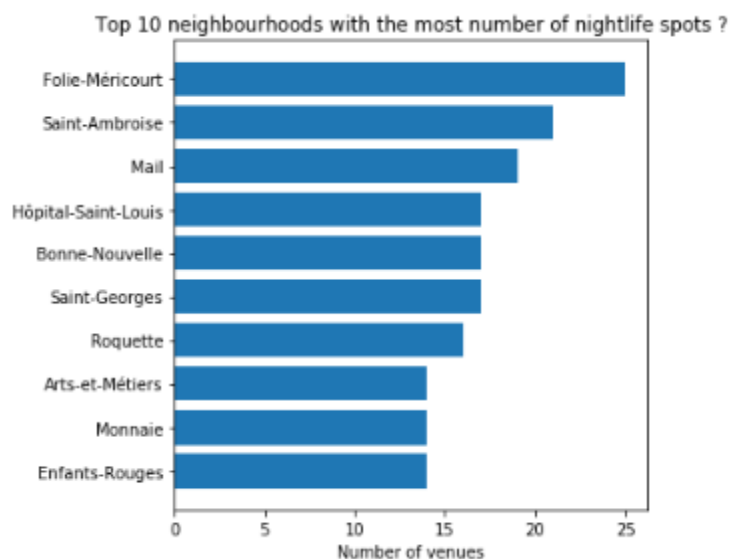
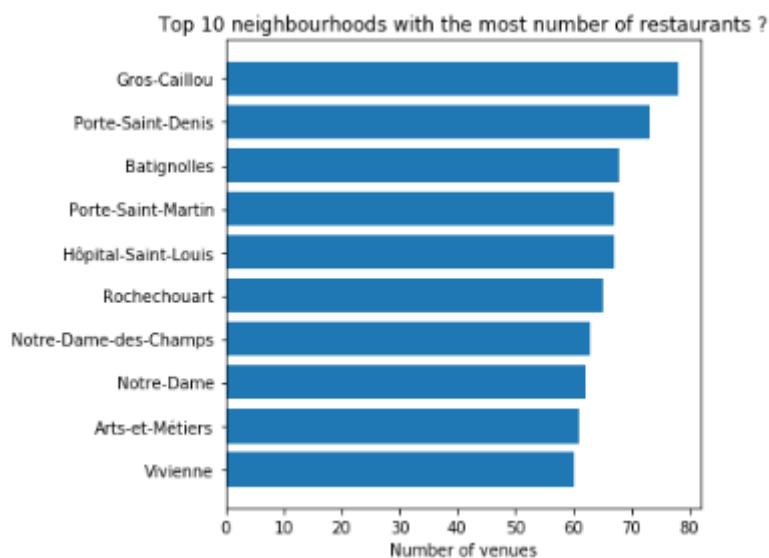
2.2 Data cleaning and feature selection

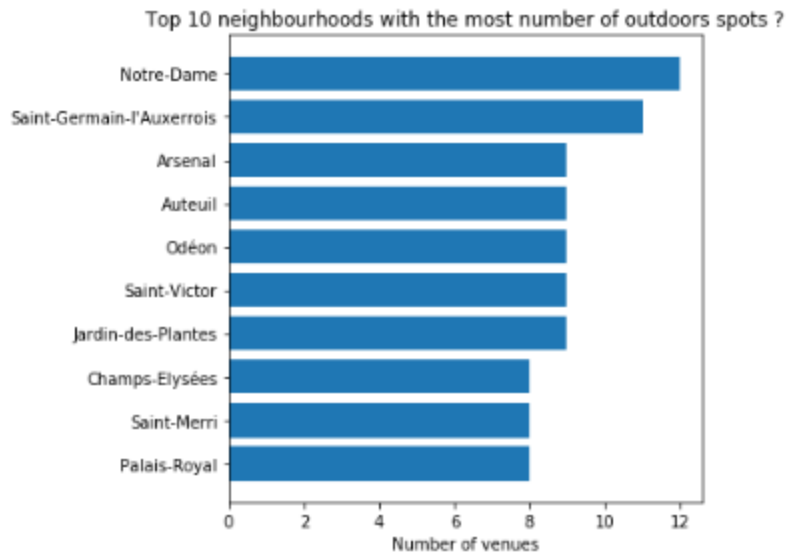
Globally the data quality from the different sources was good and there was no need to spend a lot of effort on cleaning the data. There was not too much missing data. When it was the case, I handled missing data by filling empty values with the mean of the columns.

3. Exploratory Data Analysis

3.1 Bar Graphs of top 10 neighbourhoods by venue categories

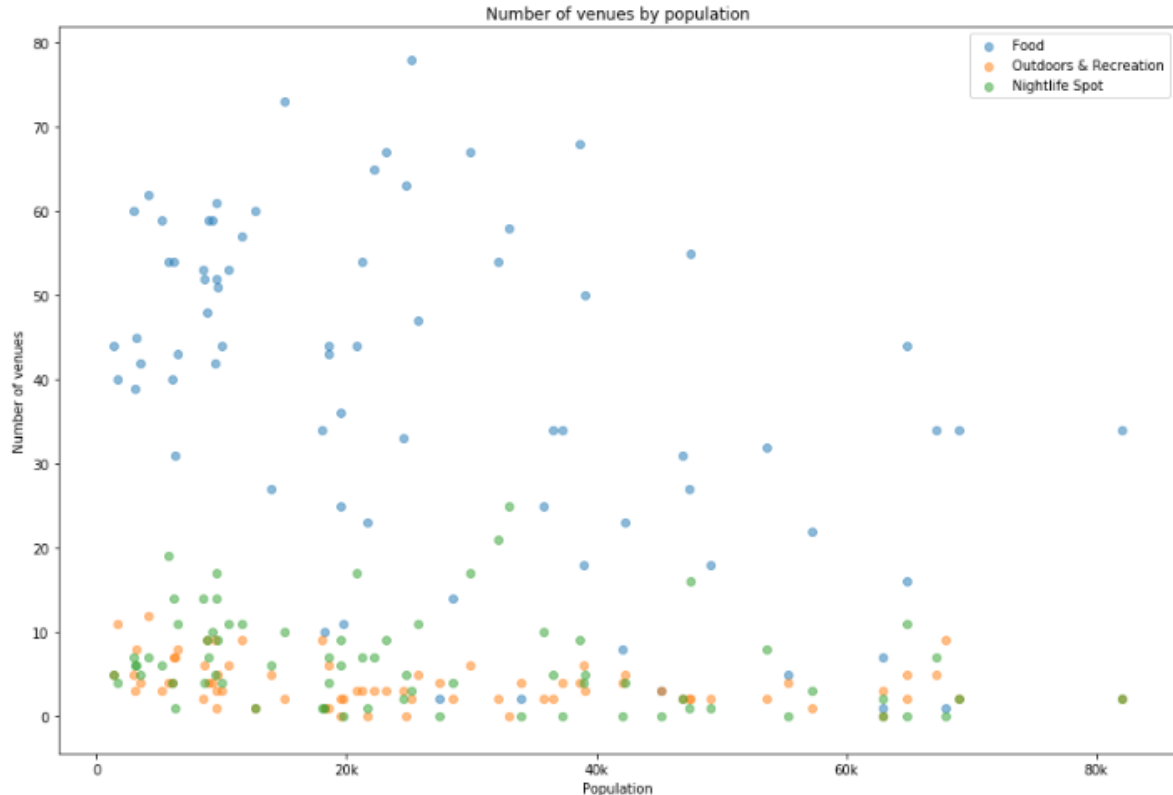
After collecting and preparing data, I ran a first analysis to display the Top 10 neighbourhoods with the greatest number of venues for the 3 categories: Food, Nightlife Spots and Outdoors Spots. First thing I learned is that there is an uneven distribution of venue categories. The Food venues are more common than the other categories. I couldn't find a neighborhood which is included in the 3 graphs. This could have been a perfect neighborhood for young people.





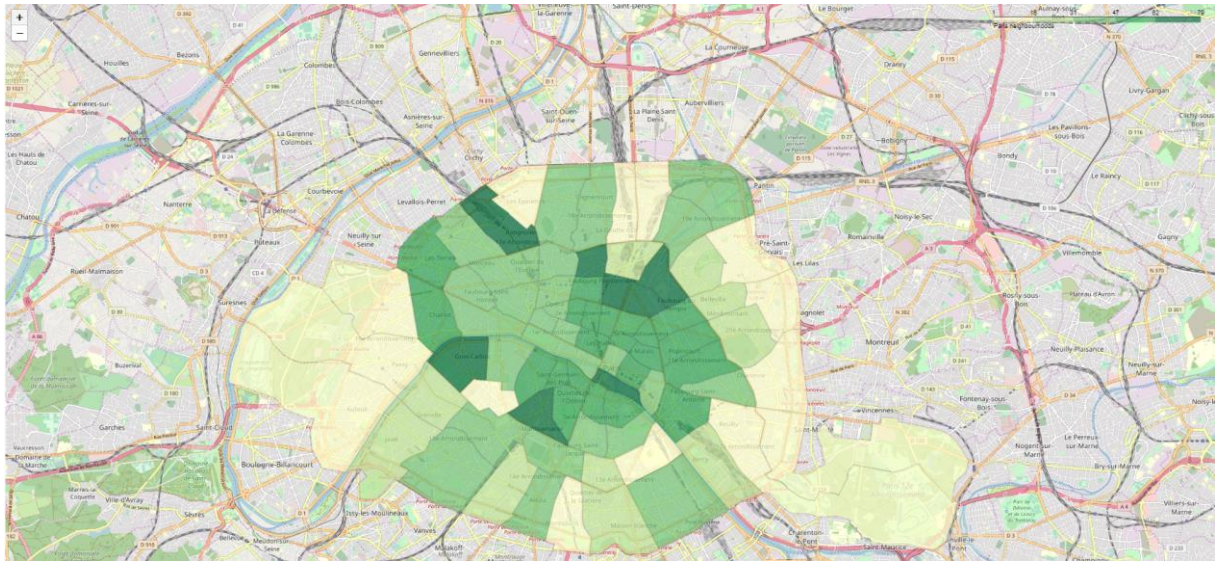
3.2 Scatter plot of number of venues by population

I tried to evaluate the relationship between number of venues and population. I hypothesized that there was a strong correlation between the 2 variables and the greater the population, the greater the number of venues. In fact, it was almost the opposite. It seems from the graph that there is no relationship between the 2 features. The points are evenly distributed on the graph. For the Food category only, we can barely distinguish that for neighborhood with a small population, the number of venues is greater. Probably in this lively neighborhood the number of food, shopping & service venues are high and there is not to many housing.



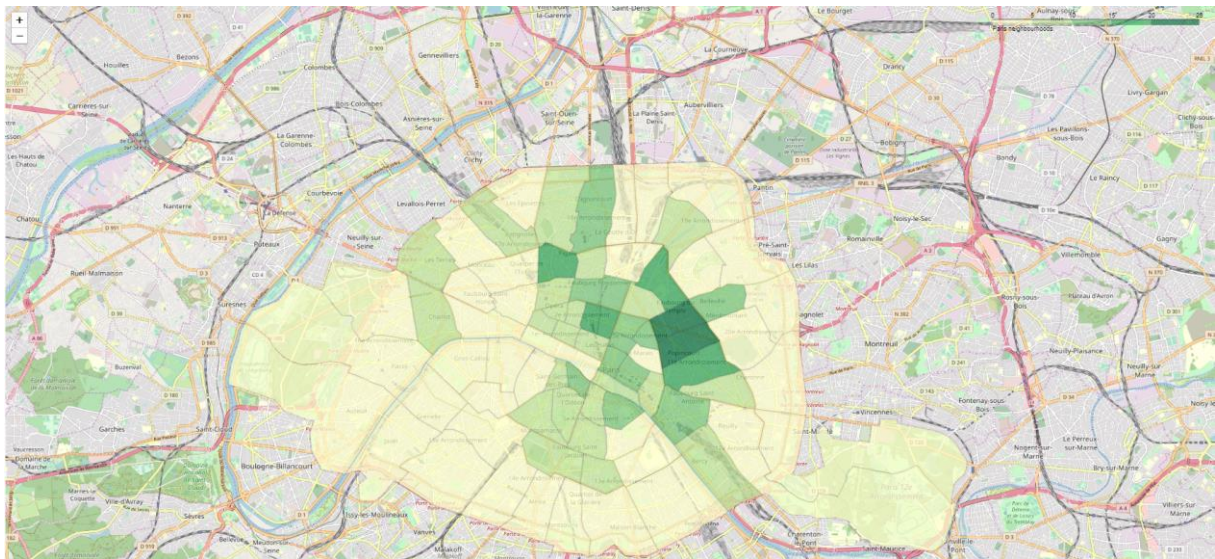
3.3 Map with Restaurants

I plotted on a map the number of restaurants (food category) for each neighborhood. We can see that restaurants are mainly located on the center with few exceptions.



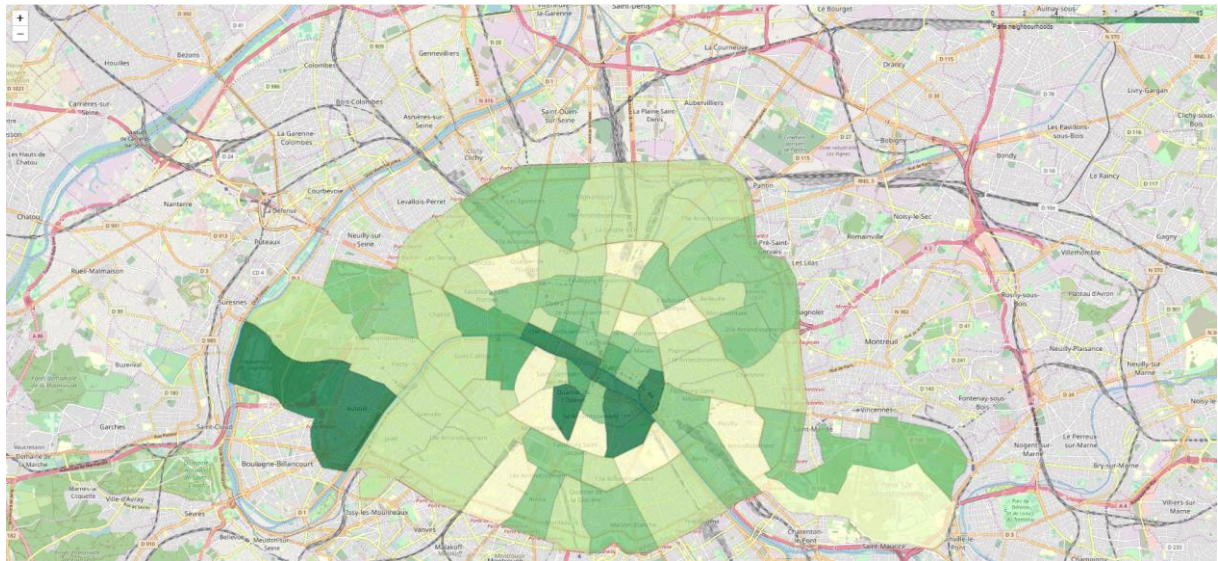
3.4 Map with Nightlife Spots

I plotted on a map the number of nightlife spots for each neighborhood. It seems that nightlife spots are concentrated on specific neighborhoods mainly located in east Paris.



3.4 Map with Outdoors & Recreation

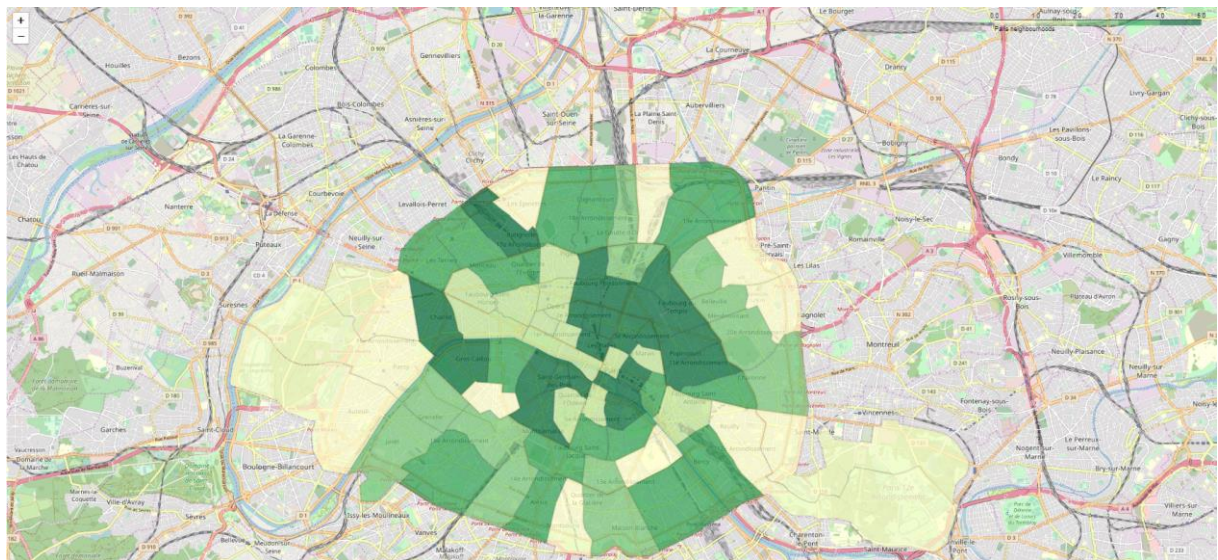
I plotted on a map the number of Outdoors & Recreation venues for each neighborhood. It seems that there are a lot of outdoors & recreation venues in the center of Paris except for one neighborhood.



4. Clustering Modeling

4.1 Map with Clusters

I then ran k-means to cluster the neighborhood into 5 clusters in order to find the ideal neighborhood for young people. It seems that the most interesting neighborhoods are colored in dark green (cluster n° 4). There are mainly located in center of Paris.



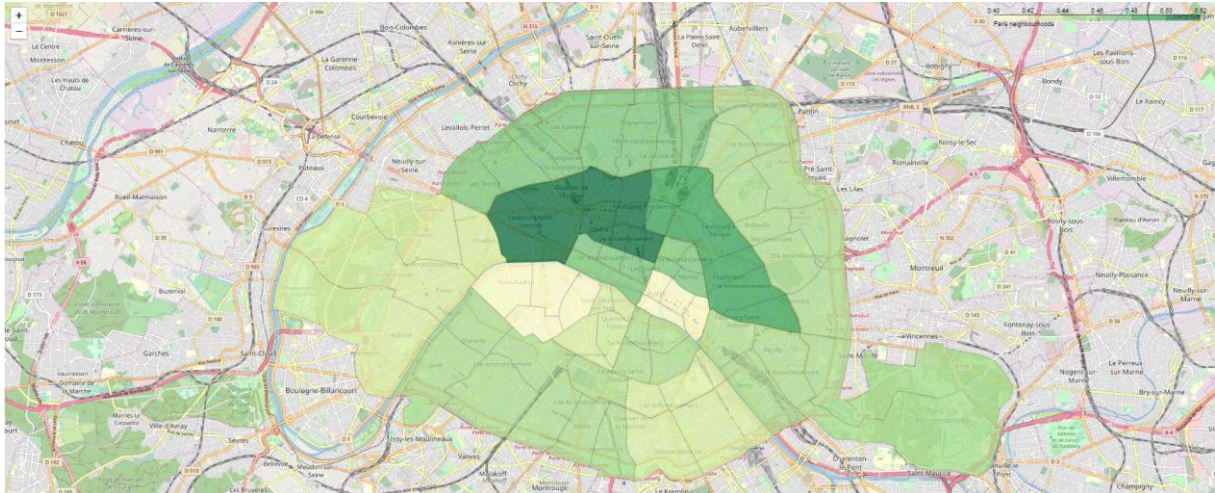
4.2 Most Common Venues by Cluster

I created a new dataframe and displayed the top 5 venues for each neighborhood in order to qualify the clusters. Again, we can see that the Food venues is globally the 1st Most Common Venues for each neighborhood. For the cluster 4, the Nightlife Spot venues are ranked 2nd or 3rd.

Cluster Labels	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	0	Amérique	Food	Outdoors & Recreation	Shop & Service	Travel & Transport
49	0	Picpus	Shop & Service	Food	Travel & Transport	Professional & Other Places
47	0	Parc-de-Montsouris	Food	Outdoors & Recreation	Travel & Transport	Arts & Entertainment
41	0	Muette	Outdoors & Recreation	Food	Travel & Transport	Shop & Service
35	0	La Chapelle	Food	Outdoors & Recreation	Shop & Service	Nightlife Spot
62	0	Saint-Fargeau	Food	Shop & Service	Outdoors & Recreation	Travel & Transport
19	0	Epinettes	Food	Shop & Service	Outdoors & Recreation	Travel & Transport
17	0	Ecole-Militaire	Food	Travel & Transport	Outdoors & Recreation	Shop & Service
74	0	Salpêtrière	Food	Shop & Service	Travel & Transport	Outdoors & Recreation
54	0	Porte-Dauphine	Outdoors & Recreation	Shop & Service	Food	Travel & Transport
6	0	Bel-Air	Outdoors & Recreation	Food	Travel & Transport	Shop & Service
4	0	Auteuil	Outdoors & Recreation	Arts & Entertainment	Shop & Service	Professional & Other Places
75	1	Sorbonne	Food	Shop & Service	Nightlife Spot	Travel & Transport
63	1	Saint-Georges	Food	Travel & Transport	Nightlife Spot	Arts & Entertainment
24	1	Gaillon	Food	Shop & Service	Travel & Transport	Arts & Entertainment
46	1	Palais-Royal	Food	Shop & Service	Travel & Transport	Outdoors & Recreation
22	1	Faubourg-du-Roule	Food	Shop & Service	Travel & Transport	Arts & Entertainment
58	1	Quinze-Vingts	Food	Shop & Service	Nightlife Spot	Travel & Transport
50	1	Place-Vendôme	Food	Shop & Service	Travel & Transport	Arts & Entertainment
1	1	Archives	Food	Shop & Service	Arts & Entertainment	Outdoors & Recreation
36	1	Madeleine	Food	Shop & Service	Travel & Transport	Outdoors & Recreation
45	1	Odéon	Food	Travel & Transport	Shop & Service	Outdoors & Recreation
13	1	Chaussée-d'Antin	Food	Shop & Service	Travel & Transport	Arts & Entertainment
65	1	Saint-Germain-l'Auxerrois	Food	Shop & Service	Arts & Entertainment	Outdoors & Recreation
11	1	Champs-Élysées	Food	Shop & Service	Arts & Entertainment	Travel & Transport
68	1	Saint-Merri	Food	Shop & Service	Arts & Entertainment	Nightlife Spot
34	2	Javel	Food	Travel & Transport	Shop & Service	Outdoors & Recreation
38	2	Maison-Blanche	Food	Shop & Service	Outdoors & Recreation	Travel & Transport
57	2	Père-Lachaise	Food	Shop & Service	Outdoors & Recreation	Nightlife Spot
42	2	Necker	Food	Travel & Transport	Shop & Service	Outdoors & Recreation
...
20	3	Europe	Food	Travel & Transport	Shop & Service	Nightlife Spot
73	3	Sainte-Marguerite	Food	Shop & Service	Nightlife Spot	Travel & Transport
14	3	Clignanecourt	Food	Nightlife Spot	Shop & Service	Travel & Transport
77	3	Val-de-Grâce	Food	Nightlife Spot	Travel & Transport	Arts & Entertainment
27	3	Grandes-Carrières	Food	Nightlife Spot	Outdoors & Recreation	Shop & Service
69	4	Saint-Thomas-d'Aquin	Food	Shop & Service	Travel & Transport	Arts & Entertainment
66	4	Saint-Gervais	Food	Shop & Service	Nightlife Spot	Outdoors & Recreation
70	4	Saint-Victor	Food	Nightlife Spot	Outdoors & Recreation	Arts & Entertainment
72	4	Sainte-Avoie	Food	Shop & Service	Nightlife Spot	Arts & Entertainment
64	4	Saint-Germain-des-Prés	Food	Shop & Service	Travel & Transport	Nightlife Spot
61	4	Saint-Ambroise	Food	Nightlife Spot	Shop & Service	Arts & Entertainment
39	4	Monnaie	Food	Shop & Service	Nightlife Spot	Outdoors & Recreation
59	4	Rochechouart	Food	Travel & Transport	Shop & Service	Nightlife Spot
3	4	Arts-et-Métiers	Food	Nightlife Spot	Shop & Service	Arts & Entertainment
5	4	Batignolles	Food	Travel & Transport	Nightlife Spot	Shop & Service
9	4	Bonne-Nouvelle	Food	Nightlife Spot	Shop & Service	Arts & Entertainment
10	4	Chailiot	Food	Travel & Transport	Shop & Service	Nightlife Spot
18	4	Enfants-Rouges	Food	Shop & Service	Nightlife Spot	Arts & Entertainment
21	4	Faubourg-Montmartre	Food	Travel & Transport	Nightlife Spot	Shop & Service
23	4	Folie-Méricourt	Food	Nightlife Spot	Shop & Service	Travel & Transport
29	4	Gros-Cailou	Food	Travel & Transport	Shop & Service	Nightlife Spot
30	4	Halles	Food	Shop & Service	Nightlife Spot	Arts & Entertainment
31	4	Hôpital-Saint-Louis	Food	Nightlife Spot	Outdoors & Recreation	Shop & Service
37	4	Mail	Food	Nightlife Spot	Shop & Service	Arts & Entertainment
43	4	Notre-Dame	Food	Outdoors & Recreation	Shop & Service	Nightlife Spot
44	4	Notre-Dame-des-Champs	Food	Travel & Transport	Shop & Service	Arts & Entertainment
55	4	Porte-Saint-Denis	Food	Nightlife Spot	Arts & Entertainment	Travel & Transport
56	4	Porte-Saint-Martin	Food	Shop & Service	Nightlife Spot	Arts & Entertainment
60	4	Roquette	Food	Nightlife Spot	Shop & Service	Arts & Entertainment
79	4	Vivienne	Food	Shop & Service	Nightlife Spot	Outdoors & Recreation

4.3 Map with Young Ratio

From the French Institute of Statistics and Economic Studies we get the actual Young Ratio (percentage of people with age range between 20 and 39). We can see that for neighborhoods in the center of Paris the ratio is higher. This corroborates our hypothesis however we can see there are some exceptions and the 2 maps (with cluster and with young ratio) doesn't match very well.



5. Conclusions

I would suggest my friend to setup his offices in the neighborhoods of the cluster n°4. However, by using the actual young ratio we can see that is not always right and neighborhoods with a high young ratio may be labelled with a different cluster. This could be explained by many reasons:

- I hypothesized that the most attractive categories for young people are Food, Nightlife Spots and Outdoors & Recreation. This is maybe not the case and maybe young people are sensitive to other categories
- I have excluded important criteria like price of housing which are probably very important for young people to select a neighborhood
- The number of venues from Foursquare was low for some neighborhoods which could have an impact on our analyses. Furthermore, the Food category was overrepresented.

6. Future directions

In order to better predict the best neighborhood, I would try to run a study in order to collect important criteria for young people when selecting a neighborhood. As well I would try to augment the data set with price of housing. Finally I would try another venues provider (like google map) to get more accurate venues information.