2 hours @Paris :

A good place for after work by district

1. Introduction

1.1 Background

Paris is the most visited city over the world by the international traveler, but in fact, the name of Paris means two different geographic definitions for the local people:

- For the Paris city heart : the real city heart 105KM² with 20 districts (in French, arrondissement)
- For the Big Paris city: the 20 districts and the department 92/93/94 (in French, Paris petite couronne)

The most of CBDs around Paris city heart are located in the 92/93/94 department. For the people who work around Paris, even the urban is well developed, the city heart remain always the most popular place for their afterwork/team building: because of the historical place, more choice for café, bar and restaurant, much more approach to the theater, cinema, park and the Seine river…

1.2 Problem

The 92/93/94 departments make a circle around Paris, thus the data can help the people who works there to know where to go for the after work/team building, according the city heart district that they choose, probably it can be

- the district next to where they work
- a place approach to somewhere specifically chosen

The project aims to locate, a good place for the after work place by district, in order to meet business request. Here what I think as 3 examples:

- For who works at La Defense, choose a best noted bar nearby Arc de Triomphe.
- For who wishes to organize a team building in a park then to restaurant along the Seine river , where to go.
- For business partners leave Paris how to organize a 2 hours small trip with a drink nearby CDG airport or 4 main railway stations.

2. Data

2.1 Data request

To solve the problem, we will need the data described as below:

- List of neighborhood in Paris: the first step is to list the districts from 1-20, based on the Paris city heart location. Within this scope, the city heart was defined and then all of the interest place could be found as the landmarks.
- List of latitude and longitude which coordinates to the interest place. This step need plot the map.
- Base on the latitude and longitude, the venue data is essential
- The last step is to perform the clustering, in order to give the conclusion.

2.2 Data source

- The Wikipedia source:

https://en.wikipedia.org/wiki/Arrondissements_of_Paris#Arrondissements

- The Paris city open data :

https://opendata.paris.fr/page/home/

This is a very interesting and powerful open source data base for Paris city. Many developer works on it and the user can get the all needed format data on it, as for this project, I use the city's GEOJSON.

- The foursquare :

The foursquare API can be used for the venue data. The site is one of the largest database for more than 100Millions places worldwide and it is used by over 130 000 developers. In this project, the Paris city venue data will be explored based on the interest place by district.

2.3 Data preparation and cleaning:

As described in the previous section, the data from 1) Wikipedia, 2) Paris open data website and 3) the foursquare need to be downloaded and selected for different uses.

2.3.1 From the Wikipedia site, the objective of this project is that take the maximum data possible from there, at least:

Name of district, the area, position compared to the Seine river (left or right), name of interest place that can be visited (at least one by district)

For the Wikipedia website give us necessary information, but there are also useless value for this project, so the idea is to set a data frame just by number of district, name, the popular landmarks, and find its latitude and longitude.

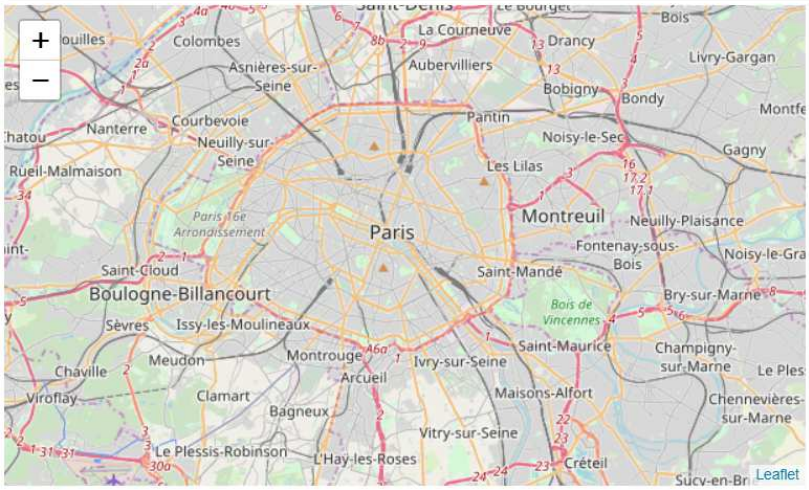
This step is the end of the date preparation of Wikipedia data. Here is an example:

```
Cathédrale Notre-Dame de Paris, 48.85293705, 2.3500501225,
Arc de Triomphe, 48.8737791, 2.29503722603767,
Centre Pompidou, 48.8605918, 2.35247431766534,
Sacré-Cœur Basilica, 48.88680575, 2.34301534488351,
Eiffel Tower, 48.8582602, 2.29449905431968,
```

2.3.2 From Paris open data website, the key is to use the arrondissement GEOJSON, means that the objectif, is to get the outcome in a Paris map with the location of interest place. The data preparation can be done in 2 steps:

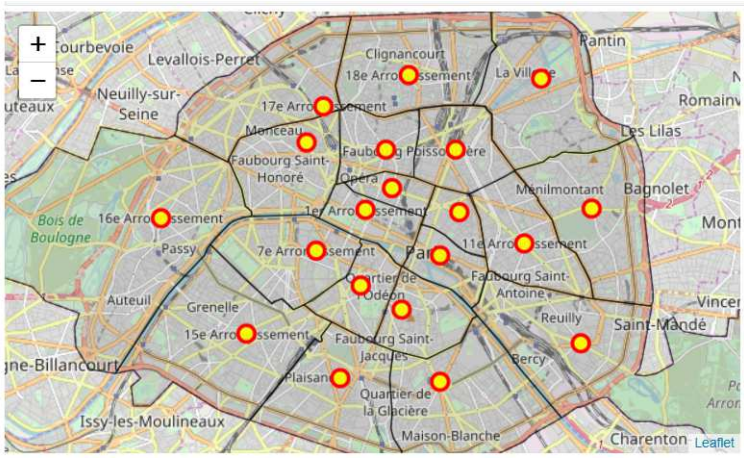
- Pick the GEOJONSO to show the city heart map

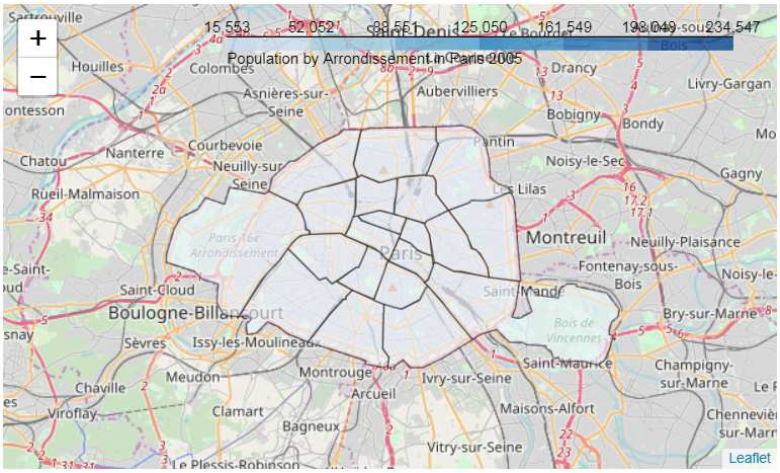
```
paris_lat = 48.8566101
paris_long = 2.3514992

paris_map = folium.Map(location=[paris_lat, paris_long], zoom_start = 12)
paris_map
```



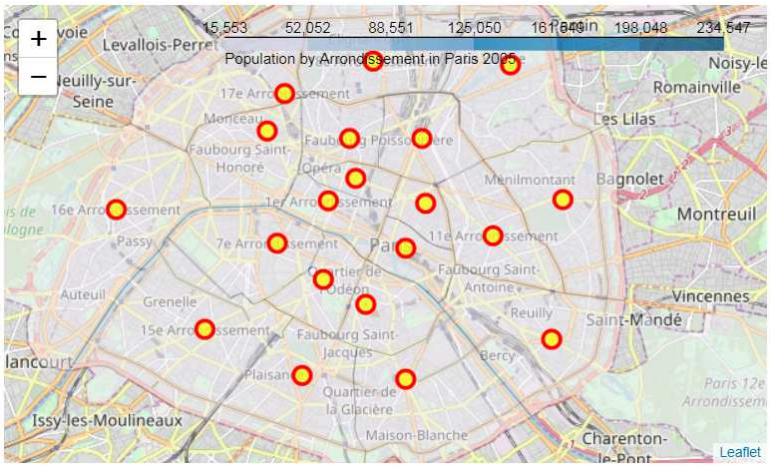
- Coordinate the geo data from the previous step to locate the land markers on it

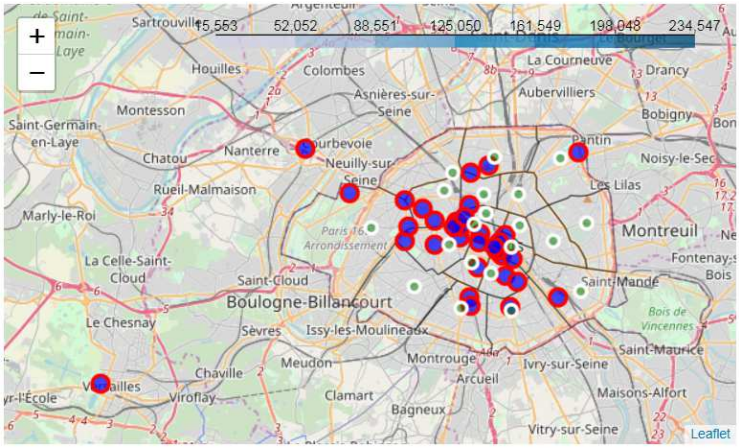
[20]:





2.3.3 The data from Foursquare is the last data set to be merged with the cleaned data. The objective is to show among the 20 district, which is the most popular one and within the scope, and then get the venue data. This project will give a simple ranking by district, to show the most popular place of each district, for example, the most popular place for 1st arrondissement is the art and museum.





An example for a top venue for the 4th district:

```
----4th----
                venue  freq
0     French Restaurant  0.11
1          Cocktail Bar  0.08
2           Coffee Shop  0.08
3                 Plaza  0.08
4      Tapas Restaurant  0.06
5  Provençal Restaurant  0.03
6        Ice Cream Shop  0.03
7          Dessert Shop  0.03
8             Bookstore  0.03
9         Boat or Ferry  0.03
```

2.4 Data clustering:

The previous steps show the all necessary data, then this step goes straight for the aim of this project: give out the suggestion by clustering the venue by district. The test is done with the K means methodology.

```
# K Means method to cluster the Paris city into 5: set and run
kclusters = 5
arrond_grouped_clustering = arrond_LL.drop('Arrondissement', 1)
kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(arrond_grouped_clustering)

kmeans.labels_[0:10]
```

```
array([0, 0, 2, 2, 3, 3, 3, 0, 0, 2], dtype=int32)
```

3. Discussion:

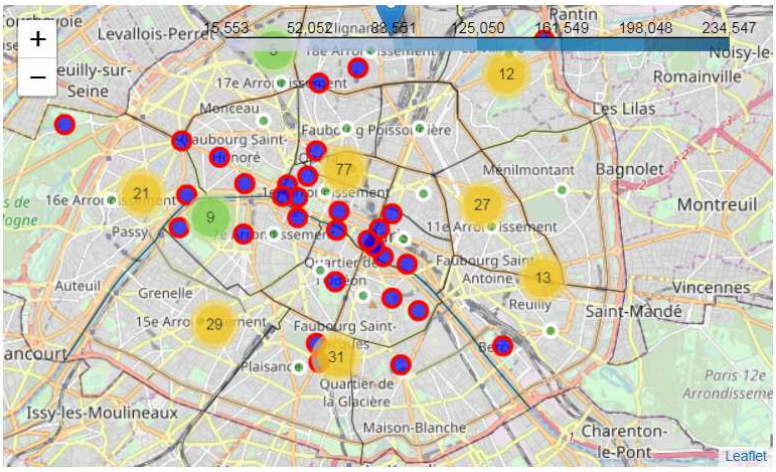
On the map we can see that if the 2 hours choice is made among the 1st to 8 arrondissement, the difference is slight, in another word, from 1 to 8 arrondissement, we can pick ant district and every choice must have at least one top noted interest place with the bar or café or restaurant nearby.

From 9 to 20 arrondissement, the choice could be limited:

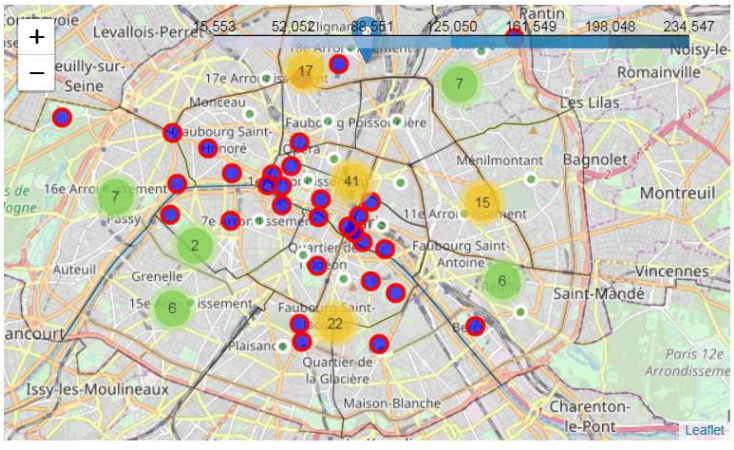
- for the 12th, the best place should be around the bois de Vincennes,
- for the 13th, the choice could be made for the Asia food and his culture,
- for 16th or 17th, the 2 hours place could be the axe of Champs Elysee or the Bois de Boulogne,
- for the 18th, the choice should be done around the Sacré Coeur and Montmartre,

The 2 hours program therefore could be 1 or 2 activities among the choice below:

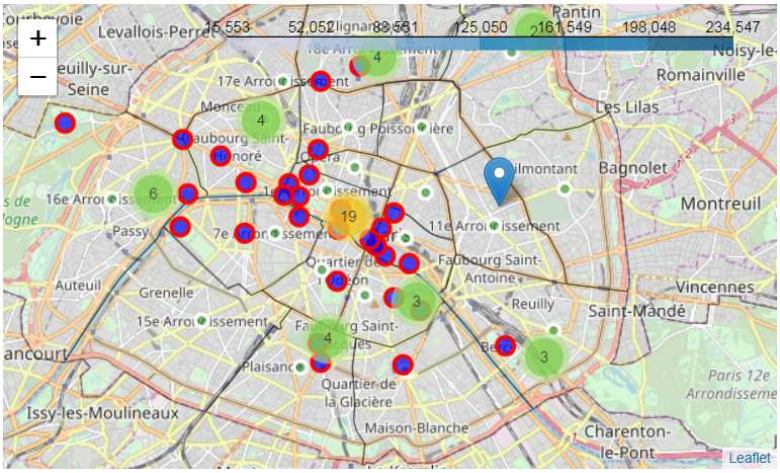
2 hours @ French restaurant



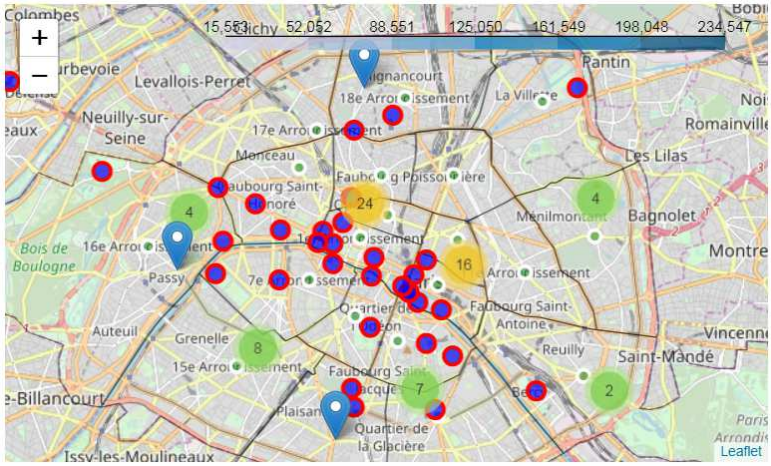
2 hours to have International food



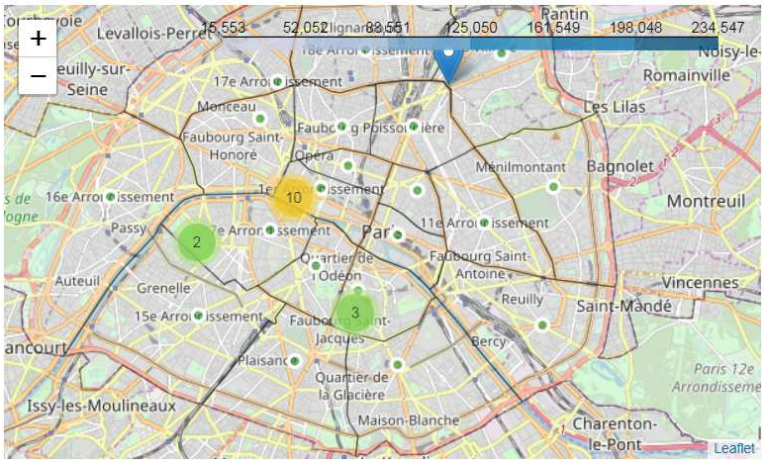
2 hours @ Museum



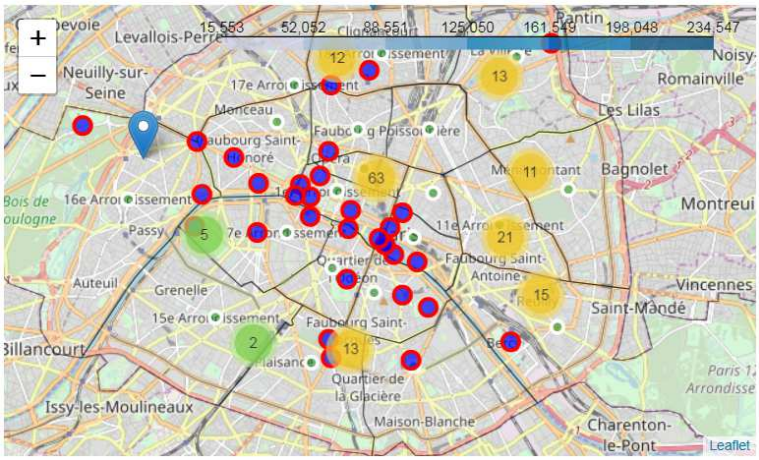
Choice of Dessert



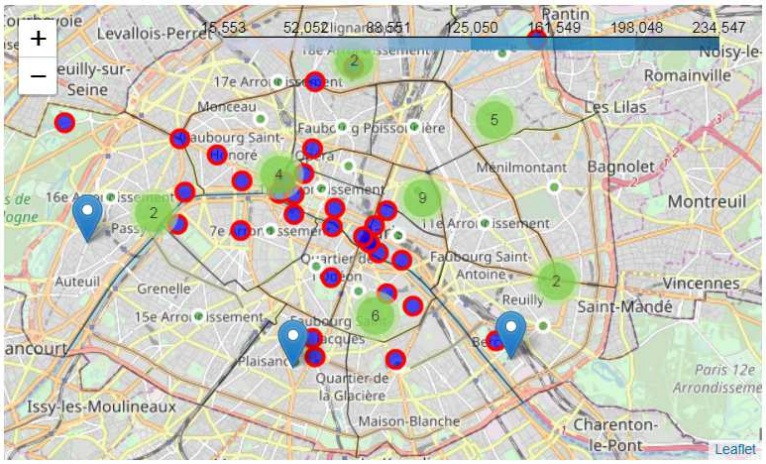
Choice of Historical site, like monument



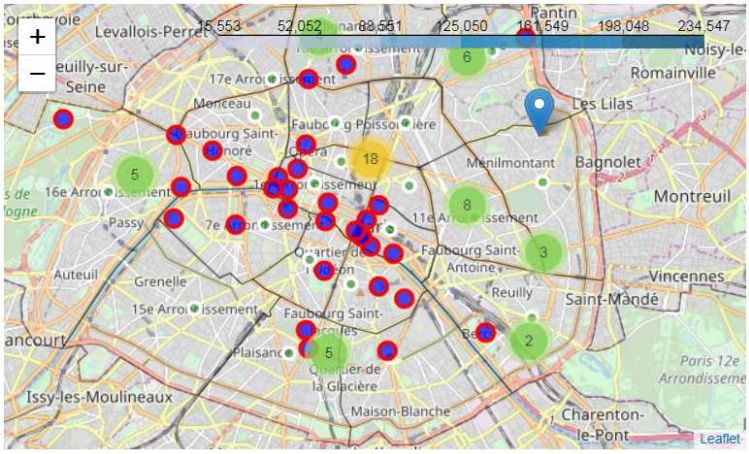
To have a Drink



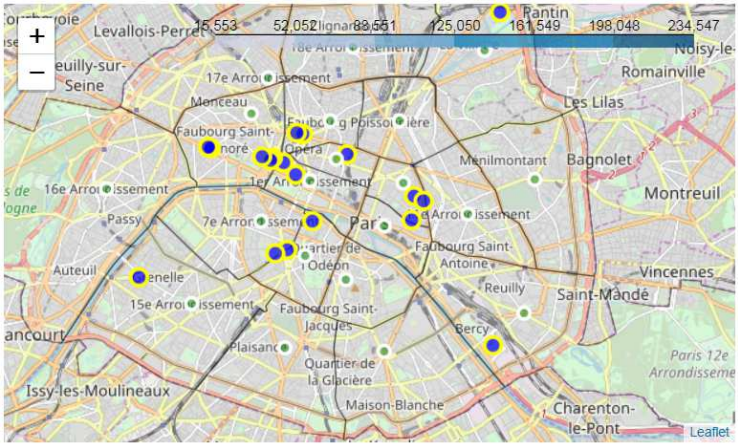
To have a Coffee



Take some Fast food



2 hours to visit a Fashion store (for who finish the training or seminar)



4. Conclusion:

Paris city heart is a small area and it is very easy to make a choice for 2 hours afterwork. The topic of the afterwork could be very varied, from interest place to shopping, from French food to drink. The best choice is in the 1-8 arrondissement, that means during 2 hours, the after work can be organized in 2 activities as the places are concentrated.