Data Science Capstone Project

Real State Agent Helper. A tool for everyone

1. Introduction. Business Problem

1.1. *Problem description and Bussiness understanding*

A real State Agent generally has to be able to provide all the possible information about the characteristics of an accommodation, like the area, number of rooms, if there is a garden or not?, is there a balcony? a garden? etc. They possessed this kind of information easily.

However, agents sometimes must face specific client demands about the characteristics of the different locations. This information is not so easy to access and summarize in order to satisfy the demands and provide useful information for the clients.

Examples of different real state clients demands concerning locations are:

- How many primary, high schools are there for my kids? Are there parks around?
- Are there many bars and restaurants around in order to enjoy the night life?
- How about cultural locations. I love to go to the teather and cinemas. Are There many options around?

The general question to ask is: How can we provide useful information about locations regarding a specific real state client demand. In other words, Is the future accommodation suitable for a client specific needs?

We can identify different kind of needs. We start to define three (3) kind of needs or scenarios:

- Family environments an cultural environment: is the location suitable for kids and families?. Are there any parks, schools, is it safe? Are there museums, cinemas and theatres?
- Nightlife environments: is the location suitable for a person that enjoys go out frequently. How many bars, restaurants, discos are there?
- Service facilities: how many service facilities are there? such as hotels, banks, train stations, spa, gyms?

This solution also compares how any particular location is similar to another concerning these scenarios

1.2. *Project environment. Analytic approach*

Where to implement this solution

This solution is suitable for any kind of city that possess many kind of different locations in a defined area. For this project the city of Paris will be explored.

- Why Paris? It is one of the cities in which each location has its particularity and provides different kind of needs

Analytic Approach:

For this project, a clustering machine learning solution will be needed in order to group the different neighborhoods for the 4 defined scenarios

In the next section we are going to explain the different requirement for the data and how to collect it, in order to provide the answer to our problem

2. Data

2.1. *Data requirements*

In order to solve the problem we have to collect, combine and analyze two kind of informations/datasets:

- *Demographic information:* for Paris, or any other city, we have to determine and investigate how it is divided. Is it divided only by neighborhoods? or any zones?. After some web investigation we can say that the city of Paris is divided by (arrondissements) wich are the districts and "quartiers" which we can translate by neighborhoods.

- *Location venues:* we need to get a dataset that provides all the different kinds of venues for a particular location. For this project we will use the foursquare location data that we use in this capstone module

2.2. *Data sources and collection*

- For the demographic information we can use two sources or methods:
  - Webscraping of websites like Wikipedia: There are many wikepedia sites that provides tables of the different district of Paris and its name like https://fr.wikipedia.org/wiki/Liste_des_quartiers_administratifs_de_Paris, or
  - A public dataset of the Paris cityhall website https://opendata.paris.fr/explore/dataset/quartier_paris/table/. that contains all the "quartier administratifs" (official neighborhoods) of the city. This public data source offers the opportunity to obtain the dataset in different formats (csv, json, excel, etc). We are going to get the csv file. This file we will provide the district number, neighboorhood name and coordinates (lat and long). **This is the method that I will use since it provides all the necessary information**

- For the venues:
  - As we said we are going to use the Foursquare dataset. We are going to use the API and we are going to use the explore url request:

    https://api.foursquare.com/v2/venues/explore?client_id=CLIENT_ID&client_secret=CLIENT_SECRET&ll=LATITUDE,LONGITUDE&v=VERSION&limit=LIMIT

- The latitude and longitude information is getted in the demographic dataset

2.3. Data Cleaning

The dataset obtained by the French public data source (https://www.data.gouv.fr/fr/datasets/r/a3b31fdc-85dc-4aeb-94c6-a8b57aebef77) has the following structure:

Out[2]:

	n_sq_qu	c_qu	c_quinsee	l_qu	c_ar	n_sq_ar	perimetre	surface	geom_x_y	geom
0	750000014	14	7510402	Saint-Gervais	4	750000004	2678.340923	4.220282e+05	48.8557186509,2.35816233385	{"type": "Polygon", "coordinates": [[[2.363764...
1	750000025	25	7510701	Saint-Thomas-d'Aquin	7	750000007	3827.253353	8.265594e+05	48.8552632694,2.32558765258	{"type": "Polygon", "coordinates": [[[2.322133...
2	750000038	38	7511002	Porte-Saint-Denis	10	750000010	2736.292954	4.721136e+05	48.873617661,2.35228289495	{"type": "Polygon", "coordinates": [[[2.355344...
3	750000001	1	7510101	Saint-Germain-l'Auxerrois	1	750000001	5057.549475	8.690007e+05	48.8606501352,2.33491032928	{"type": "Polygon", "coordinates": [[[2.344593...
4	750000073	73	7511901	Villette	19	750000019	5191.018830	1.285705e+06	48.8876610888,2.37446821213	{"type": "Polygon", "coordinates": [[[2.370498...

- We start to delete some of the columns that are not necessary and change the name of the columns in order to have a clean dataset. We split also the column "geom_x_y" in order to have the columns Latitude and Longitude:

Out[8]:

	Neighborhood	District	Latitude	Longitude
0	Saint-Gervais	4	48.855719	2.358162
1	Saint-Thomas-d'Aquin	7	48.855263	2.325588
2	Porte-Saint-Denis	10	48.873618	2.352283
3	Saint-Germain-l'Auxerrois	1	48.860650	2.334910
4	Villette	19	48.887661	2.374468
5	Val-de-Grâce	5	48.841684	2.343861
6	Necker	15	48.842711	2.310777
7	Père-Lachaise	20	48.863719	2.395273
8	La Chapelle	18	48.894012	2.364387
9	Europe	8	48.878148	2.317175

2.4. Data formatting

- Using the Foursquare API and the explore method mentioned in section 2.2 we have got the following dataframe

Out[14]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Saint-Gervais	48.855719	2.358162	Tasca	48.856686	2.356374	Portuguese Restaurant
1	Saint-Gervais	48.855719	2.358162	Miznon	48.857201	2.358957	Israeli Restaurant
2	Saint-Gervais	48.855719	2.358162	Murciano Boulangerie et Patisserie	48.856984	2.359789	Bakery
3	Saint-Gervais	48.855719	2.358162	Aux Merveilleux de Fred	48.855686	2.356369	Dessert Shop
4	Saint-Gervais	48.855719	2.358162	Autour du Saumon	48.855587	2.357802	Scandinavian Restaurant
5	Saint-Gervais	48.855719	2.358162	Jardin de l'Hôtel de Sens	48.853842	2.358404	Garden
6	Saint-Gervais	48.855719	2.358162	Florence Kahn	48.857242	2.359057	Deli / Bodega
7	Saint-Gervais	48.855719	2.358162	Grom	48.856737	2.356933	Ice Cream Shop
8	Saint-Gervais	48.855719	2.358162	Vingt Vins d'Art	48.855214	2.357940	Wine Bar
9	Saint-Gervais	48.855719	2.358162	Comme à Lisbonne	48.856767	2.356462	Café

- In order to use this information for the clustering algorithm we apply the one hot enconding and grouped all the venues for each neighborhood:

	Neighborhood	Accessories Store	African Restaurant	American Restaurant	Antique Shop	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	...	Vegetarian / Vegan Restaurant	Venezuelan Restaurant	Video Game Store	Vietnamese Restaurant	Wine Bar	Wine Shop	Women's Store	Yoga Studio	Zoo	Zoo Exhibit
0	Amérique	0.0	0.0	0.00000	0.000000	0.0	0.00	0.00	0.00	0.00	...	0.00000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.0	0.0
1	Archives	0.0	0.0	0.00000	0.000000	0.0	0.00	0.04	0.02	0.00	...	0.00000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.0	0.0
2	Arsenal	0.0	0.0	0.00000	0.000000	0.0	0.00	0.00	0.00	0.00	...	0.03125	0.0	0.0	0.000000	0.015625	0.000000	0.000000	0.00	0.0	0.0
3	Arts-et-Métiers	0.0	0.0	0.00000	0.000000	0.0	0.01	0.01	0.00	0.00	...	0.01000	0.0	0.0	0.040000	0.050000	0.000000	0.000000	0.00	0.0	0.0
4	Auteuil	0.0	0.0	0.00000	0.000000	0.0	0.00	0.00	0.00	0.00	...	0.00000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.0	0.0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
75	Sorbonne	0.0	0.0	0.00000	0.000000	0.0	0.01	0.00	0.00	0.01	...	0.00000	0.0	0.0	0.010000	0.030000	0.000000	0.000000	0.01	0.0	0.0
76	Ternes	0.0	0.0	0.03125	0.000000	0.0	0.00	0.00	0.00	0.00	...	0.00000	0.0	0.0	0.015625	0.015625	0.000000	0.000000	0.00	0.0	0.0
77	Val-de-Grâce	0.0	0.0	0.00000	0.000000	0.0	0.00	0.00	0.00	0.00	...	0.00000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.0	0.0
78	Villette	0.0	0.0	0.00000	0.000000	0.0	0.00	0.00	0.00	0.00	...	0.00000	0.0	0.0	0.000000	0.000000	0.000000	0.000000	0.00	0.0	0.0
79	Vivienne	0.0	0.0	0.00000	0.010101	0.0	0.00	0.00	0.00	0.00	...	0.00000	0.0	0.0	0.000000	0.070707	0.010101	0.010101	0.00	0.0	0.0

80 rows × 303 columns

2.5. *Feature Selection*

As we have said in the introduction. This tool will consider three kind of scenarios or needs. Each scenario has a certain number of features (in this case the venues) that are going to be the variables to use in order to segment the data and define the family of similar neighborhoods.

So, for each scenario we have defined the following parameters (venues)

- For the family, cultural friendly neighbourhoods we use the following features:

```
['Bookstore','Garden','Park','Theater','Art Museum','Historic Site','Pedestrian Plaza','Art Gallery','Indie Movie Theater',
 'Museum','Concert Hall','Comedy Club','Arts & Crafts Store','Music Venue','Science Museum','Church','Cultural Center',
 'Middle Eastern Restaurant','Playground']
```

- For the nightlife scenario the features to use are:

```
['French Restaurant','Italian Restaurant','Bar','Japanese Restaurant','Café','Bistro','Plaza','Wine Bar',
 'Restaurant','Coffee Shop','Pizza Place','Cocktail Bar','Sandwich Place','Thai Restaurant','Ice Cream Shop',
 'Chinese Restaurant','Indian Restaurant','Vietnamese Restaurant','Tea Room','Burger Joint','Seafood Restaurant',
 'Asian Restaurant','Creperie','Korean Restaurant','Sushi Restaurant','Dessert Shop','Salad Place',
 'Vegetarian / Vegan Restaurant','Pub','Beer Bar','Tapas Restaurant','Hotel Bar','Moroccan Restaurant','Gastropub',
 'Steakhouse','Mexican Restaurant','Diner','Brasserie','Lebanese Restaurant','Breakfast Spot','Fast Food Restaurant',
 'Greek Restaurant','Falafel Restaurant','Food & Drink Shop','Mediterranean Restaurant','Argentinian Restaurant','Juice Bar',
 'African Restaurant','Ethiopian Restaurant','American Restaurant','Nightclub','Noodle House','Liquor Store','Ramen Restaurant',
 'Turkish Restaurant','Udon Restaurant','Cajun / Creole Restaurant','Lounge','Portuguese Restaurant','Scandinavian Restaurant',
 'Bubble Tea Shop','Corsican Restaurant','Food Truck','Fountain','Israeli Restaurant','Movie Theater','Peruvian Restaurant',
 'Basque Restaurant','Fish & Chips Shop','New American Restaurant','Southwestern French Restaurant']
```

- For the services, facilities scenario the features to use are

```
['Hotel','Bakery','Supermarket','Pastry Shop','Clothing Store','Cheese Shop','Gym / Fitness Center','Boutique',
 'Cosmetics Shop','Chocolate Shop','Gourmet Shop','Convenience Store','Spa','Farmers Market','Grocery Store','Wine Shop',
 'Bagel Shop','Deli / Bodega','Furniture / Home Store','Candy Store','Beer Store','Bike Rental / Bike Share','Bus Stop',
 'Jewelry Store','Miscellaneous Shop','Pool','Gym','Cupcake Shop','Perfume Shop','Department Store','Multiplex',
 'Shoe Store','Electronics Store','Hostel','Metro Station','Record Shop','Tailor Shop','Toy / Game Store',
 'Accessories Store','Souvlaki Shop','Train Station','Tram Station','Yoga Studio']
```

3. Methodology

The goal of this project is to propose to the clients the neighbourhoods that will satisfy certain need. For that we need the grouped the similar neighbourhoods that share the same frequency of venues for each of the category.

Since we want to talk about grouping or segment different elements based on similarities or dissimilarities we will use and unsupervised machine learning technique called Clustering

The algorithm that we are going to use is the K-means algorithm:

K-Means is a type of partitioning clustering, that divides the data into K non-overlapping subsets or clusters without any cluster internal structure or labels.

With this algorithm neighbourhoods within a cluster are very similar, and neighbourhoods across different clusters are very different or dissimilar

The algorithm works in this way:

- First we chose and place K centroids, one of each cluster. In our case we are going to choose these centroids randomly
- Second, the distance between each point (neighbourhood) and the centroid is calculated
- Third, we assign each data point (neighbourhood) to its closest centroid, creating a cluster
- Fourth, we recalculate the position of the centroids
- And finally we iterate the second, third and fourth step until the centroids do not move

There are another type of clustering algorithm such as DBSCAN but it is more suitable for searching anomalies between the elements. We do not want to leave any neighbourhood behind and alone without a cluster

3.1. Clustering the city considering all the venues

In this part we are going to segment our city, without taking into consideration the 3 scenarios or needs that we define. That means we are going to group the neighborhoods by taking into account all the venues collected

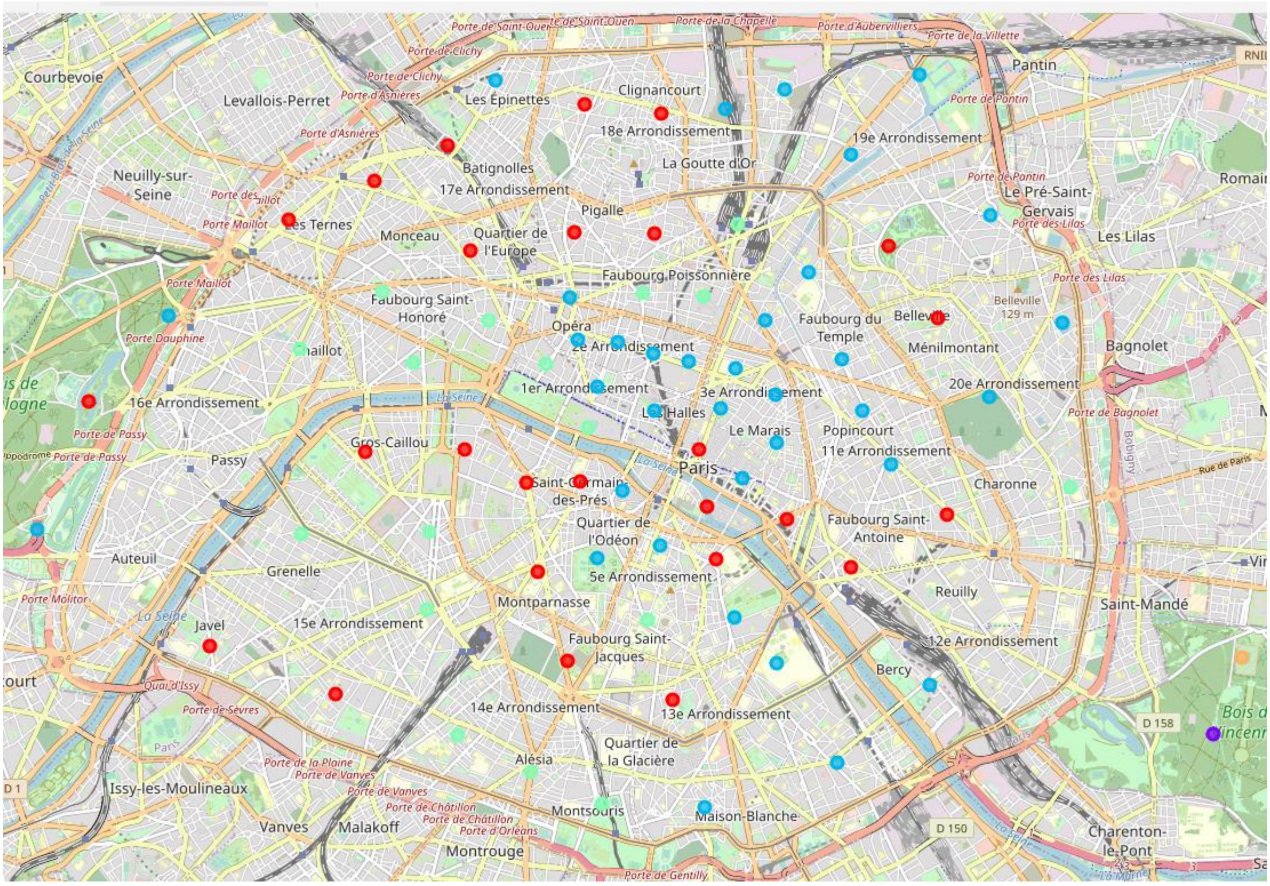
In general, we are going to get 5 different clusters for all 80 neighboorhods for each scenario

The code used to perform the K_means algorithm is the following:

```
# We apply the K_means algorithm clustering

kclusters = 5
paris_clust_general = paris_grouped.drop('Neighborhood', 1)

#apply k-means algorithm

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(paris_clust_general)
kmeans.labels_[0:10]
```

```
]: array([2, 2, 0, 2, 2, 0, 4, 0, 2, 2], dtype=int32)
```

- After applying the algorithm, we have got these clusters:



- Cluster 1: we can say that in this cluster the most important venues are french restaurants, bar and hotels. This neighbourhoods seem to be very visited by locals and tourist.
- Cluster 2: we only one neighbourhood in this cluster. The most important venues are shops and services. We can see there are open spaces like pedestrian plazas and parks. It is not a crowded area.
- Cluster 3: this cluster is similar to the first cluster. However the difference with the first cluster is that we see more varieties of services, shops and open spaces like parks and gardens.
- Cluster 4: this cluster is and hybrid of the first and third cluster. We can find a lot of restaurants and hotels like the first cluster. This cluster has also many shops and services facilities like the third cluster but less open spaces.

- Cluster 5: As Cluster 2. This cluster has one neighbourhood. It is like cluster 2 because there are many open spaces like playgrounds and plazas. However, there are more restaurants and bars than the cluster 2 and the shops are different

3.2. Clustering the city considering the family, cultural venues

How about if you love to live in a family environment that offers parks, gardens and many cultural activities such as museums and concerts?

For this scenario we are going to use the following venues as parameters to use in the K_means algorithm:

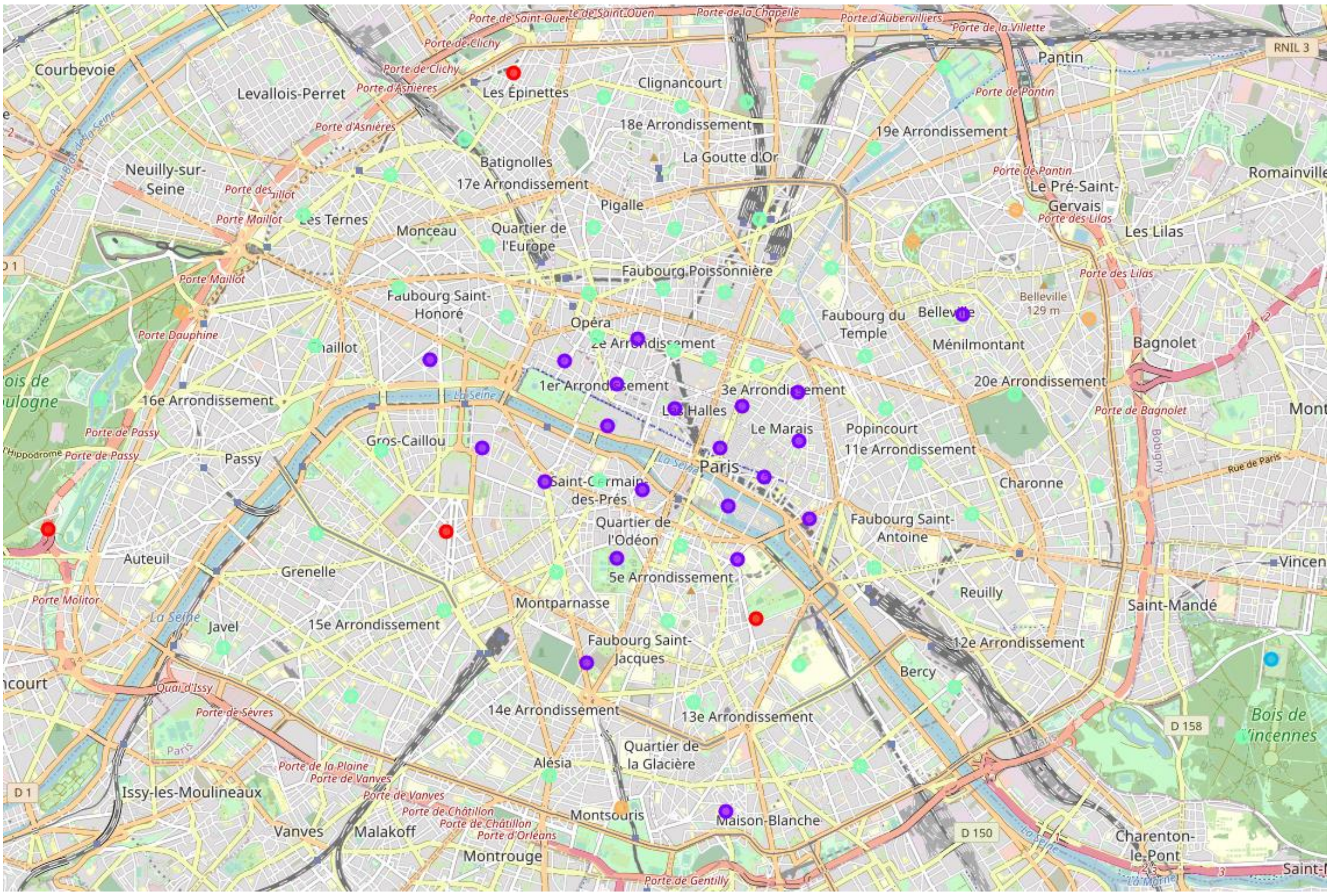
```
=['Bookstore','Garden','Park','Theater','Art Museum','Historic Site','Pedestrian Plaza','Art Gallery','Indie Movie Theater',
 'Museum','Concert Hall','Comedy Club','Arts & Crafts Store','Music Venue','Science Museum','Church','Cultural Center',
 'Middle Eastern Restaurant','Playground']
```

We apply as well the K_means algorithm with 5 clusters:

```
#Let's now reaply the k-means algorithm

paris_clust_family_cultural = paris_family_cultural.drop('Neighborhood', 1)

#apply k-means algorithm

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(paris_clust_family_cultural)
kmeans.labels_[0:10]
```

```
5]: array([4, 1, 1, 3, 0, 3, 2, 1, 3, 3], dtype=int32)
```

- After applying the algorithm, we have got these clusters:



- Cluster 1: we have found 4 neighbourhoods. If you love gardens and general museums these neighbourhoods are four you
- Cluster 2: this is the zone for the art galleries and art museums
- Cluster 3: this cluster has only one neighbourhood. We do not find a lot museums nearby. But this neighbourhood is good if you like open spaces, concerts and general cultural activities.
- Cluster 4: this is the cluster with the most number of neighbourhoods. Here you can find many cultural shops like bookstores, as well as general museums and theaters. It has many cultural centers too and gardens. In general a very versatile cluster a typical downtown cluster
- Cluster 5: If you are looking for parks and great open areas these are the neighbourhoods for you.

3.3. Clustering the city considering the nightlife activities

How about if you love to go out to have diner or lunch and you love have drinks with friends and party. This study tells you all

For this scenario we are going to use the following venues as parameters to use in the K_means algorithm:

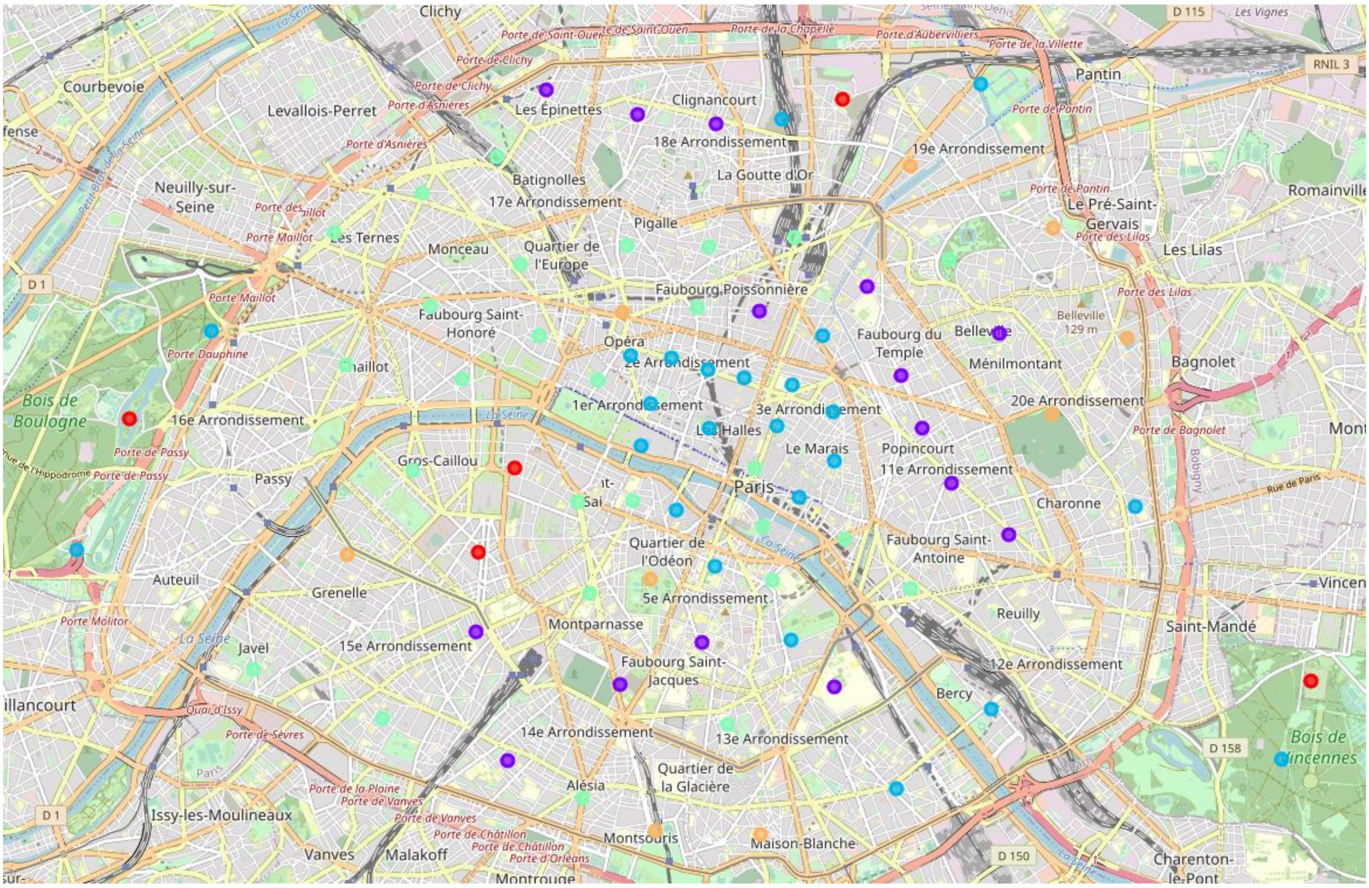
```
paris_grouped[['Neighborhood','French Restaurant','Italian Restaurant','Bar','Japanese Restaurant','Café','Bistro','Plaza','Wine Bar',
       'Restaurant','Coffee Shop','Pizza Place','Cocktail Bar','Sandwich Place','Thai Restaurant','Ice Cream Shop',
       'Chinese Restaurant','Indian Restaurant','Vietnamese Restaurant','Tea Room','Burger Joint','Seafood Restaurant',
       'Asian Restaurant','Creperie','Korean Restaurant','Sushi Restaurant','Dessert Shop','Salad Place',
       'Vegetarian / Vegan Restaurant','Pub','Beer Bar','Tapas Restaurant','Hotel Bar','Moroccan Restaurant','Gastropub',
       'Steakhouse','Mexican Restaurant','Diner','Brasserie','Lebanese Restaurant','Breakfast Spot','Fast Food Restaurant',
       'Greek Restaurant','Falafel Restaurant','Food & Drink Shop','Mediterranean Restaurant','Argentinian Restaurant','Juice Bar',
       'African Restaurant','Ethiopian Restaurant','American Restaurant','Nightclub','Noodle House','Liquor Store','Ramen Restaurant',
       'Turkish Restaurant','Udon Restaurant','Cajun / Creole Restaurant','Lounge','Portuguese Restaurant','Scandinavian Restaurant',
       'Bubble Tea Shop','Corsican Restaurant','Food Truck','Fountain','Israeli Restaurant','Movie Theater','Peruvian Restaurant',
       'Basque Restaurant','Fish & Chips Shop','New American Restaurant','Southwestern French Restaurant']]
```

We apply as well the K_means algorithm with 5 clusters:

```
#Let's now reaply the k-means algorithm

paris_clust_nightlife = paris_nightlife.drop('Neighborhood', 1)

#apply k-means algorithm

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(paris_clust_nightlife)
kmeans.labels_[0:10]
```

```
]: array([4, 2, 3, 2, 2, 3, 0, 1, 2, 2], dtype=int32)
```

- After applying the algorithm, we have got these clusters:



- Cluster 1: we have found 5 neighbourhoods. This first cluster is for restaurant lovers only. We can find mostly French restaurants but many other options to eat too. However, no bars nearby

- Cluster 2: in this second cluster we find also many French restaurants and a well variety of international restaurants. The difference with the first cluster is that we found a lot of bars too. It seems the neighbourhood to have fun
- Cluster 3: the neighbourhoods of this cluster has a very good balance between restaurants and bars
- Cluster 4: as cluster number 1, neighbourhoods on cluster 4 offers mostly restaurants. No so many bars. However, we found more italian restaurants here
- Cluster 5: this cluster has a versatile food option: french, italian and japanese restaurants, bistros and more calm venues such as cafés.

3.4. Clustering the city considering the service venues

How about if you want that all the services venues are close. You like to shop, to go to gym, spas, and have all the facilities near. Look this scenario!

These venues were defined in the following list:

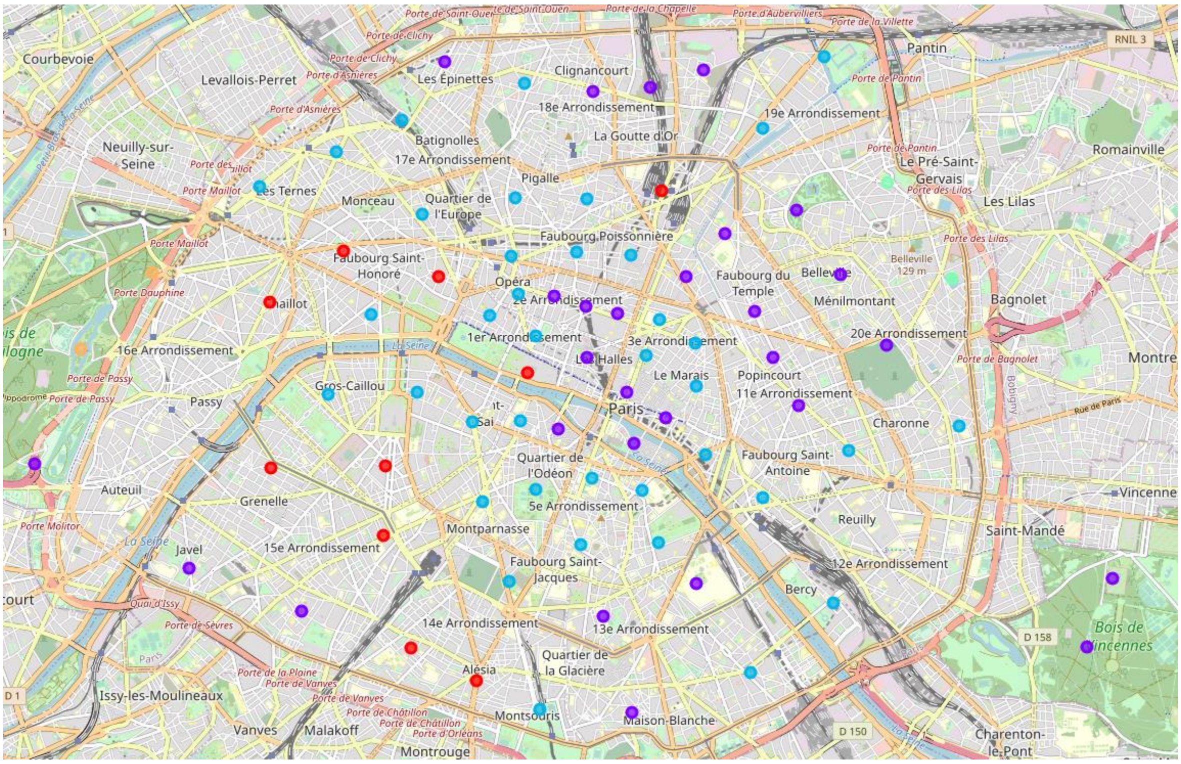
```
paris_grouped[['Neighborhood','Hotel','Bakery','Supermarket','Pastry Shop','Clothing Store','Cheese Shop','Gym / Fitness Center','Boutique',
        'Cosmetics Shop','Chocolate Shop','Gourmet Shop','Convenience Store','Spa','Farmers Market','Grocery Store','Wine Shop',
        'Bagel Shop','Deli / Bodega','Furniture / Home Store','Candy Store','Beer Store','Bike Rental / Bike Share','Bus Stop',
        'Jewelry Store','Miscellaneous Shop','Pool','Gym','Cupcake Shop','Perfume Shop','Department Store','Multiplex',
        'Shoe Store','Electronics Store','Hostel','Metro Station','Record Shop','Tailor Shop','Toy / Game Store',
        'Accessories Store','Souvlaki Shop','Train Station','Tram Station','Yoga Studio']]
```

We apply as well the K_means algorithm with 5 clusters:

```
#Let's now reaply the k-means algorithm

paris_clust_services = paris_services.drop('Neighborhood', 1)

#apply k-means algorithm

kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(paris_clust_services)
kmeans.labels_[0:10]
```

```
]: array([3, 2, 2, 2, 1, 2, 1, 1, 2, 1], dtype=int32)
```

- After applying the algorithm, we have got these clusters:



- Cluster 1: the neighbourhoods of this cluster are suitable for tourism due to the variety of hotels
- Cluster 2: The neighbourhoods of this cluster provide an interesting variety of services. We can find hotels, many supermarkets, and bakeries. It is a typical downtown neighbourhood
- Cluster 3: This cluster is a typical downtown cluster too. It is very suitable for tourists as cluster 1 due to the variety of hotels. The difference from cluster 1 is that cluster 3 seem to offer more bakeries and supermarkets
- Cluster 4: the two neighbourhoods of this cluster offers many services, but they are not suitable for visitors due to the lack of hotels offers
- Cluster 5: like Cluster 4, the neighbourhoods of this cluster are not suitable for tourist. It is not a typical downtown cluster because we do not find supermarkets and bakeries

4. Results and discussion

This tool helps real state agents and clients identify the most suitable neighborhoods to live considering the general and specific needs

This tool has stablished 4 types of pool of clusters. Bellow we can describe the results:

a) General clustering: for this clustering exercise we have considered all the venues. We did not distingish a category. We have found the following:

- In this first analysis we have seen that there are neigborhoods more crowded than others.
- We identify neighborhoods that offers more open spaces than others (playgrounds, gardens and plazas)
- We identiy crowded areas that offers several kind of restaurant venues and services like shops

b) Family and cultural clustering: in this exercise we consider only the venues that are considered

- We identify cluster of neighbourhoods that offers open spaces like gardens and museums
- We identify a cluster in which art is essential in the form of art museums and galleries
- There are other neighbourhoods in which we do not find art and museums but it offers other cultural activities like concerts
- There is a cluster in which we found many bookstores and theaters

c) Nightlife clustering: bars and restaurants. In this exercise we have found:

- Neighbourhoods that offers more options to eat than to drink (not so many bars)
- We have found also the opposite. Clusters that offers a very good variety of bars and nightclubs
- Neighbourhoods that offers the two options: a more calm venues like restaurants and cafés and more "loud" venues like bars and nightclubs

d) Services clustering: In this exercise we have found:

- There are neighbourhoods more suitable for the tourism which offers a good variety of hotels
- There are neighbourhoods that offers more "downtown" kind of life with a good variety of supermarket, shops and bakeries
- There are neighbourhoods less crowded that do not offer venues like supermarket, groceries stores and bakeries

5. Conclusions

This tool helps real state agents to aim and search the most suitable neighbourhoods considering many criteria’s. We can have many specific client’s demands about the characteristics of the different locations.

This tool collects, summarizes and segment the group of neighbourhoods that offer similar venues.

With this tool we can get:

- The neighbourhoods more crowded and less crowded
- The neighbourhoods that offers more open spaces like gardens and parks
- The neighbourhoods that offers the amount and kind of cultural offer like art galleries, museums and theatres
- The neighbourhoods that offers more restaurants and cafés.
- The neighbourhoods that are more suitable for tourism
- The neighbourhoods that are more suitable for nightlife activities (bars and nightclubs)

Finally, as we use a cluster algorithm. The neighbourhoods are grouped. This could help a client if they want to go to a neighbourhood that is similar to the current one or a total different one.