IBM Data Science Capstone Project

The battle for restaurants in the Paris neighborhoods

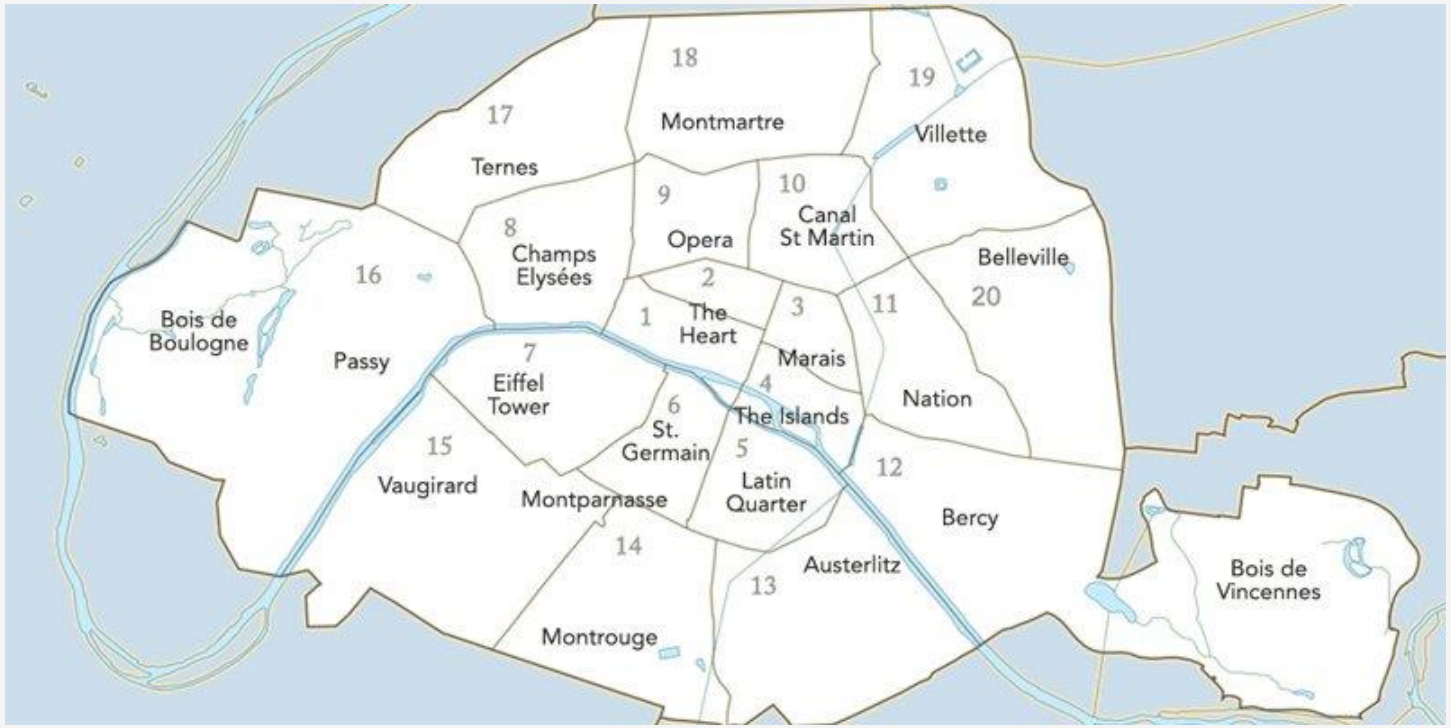
Philippe Bottier

May 18, 2021

Business problem

The problem to be solved is totally imaginary.

My company PhB Data Consulting, specializing in data analysis, was contacted by the Paris Tourist Office to provide them with an analysis that would allow them to advise tourists on the types of restaurants they might find during their visits to the Parisian neighborhoods.



Methodology

To conduct this study, I list below the different steps of the methodology that I followed :

1. Data Acquisition with the JSON library
2. Preparing, Cleaning and Merging Data with the PANDA library
3. Venues Acquisition with the Foursquare API
4. One Hot Encoding with the PANDA library
5. Clustering K-means with sklearn.cluster library
6. Cluster Analysis
7. Creation of a clusters map with the Folium library

1 Data Acquisition with the JSON library

From the Open Data site of the city of Paris :

- Paris arrondissements : https://opendata.paris.fr/explore/dataset/arrondissements/export/
- Paris neighbourhoods : https://opendata.paris.fr/explore/dataset/quartier_paris/export/

These files are in GeoJSON format.

```
import json
geo_borough = json.load(open("arrondissements.geojson")) # Paris arrondissements
geo_neighbourhood = json.load(open("quartier_paris.geojson")) # Paris neighborhoods
```

```
geo_borough["features"][0]['properties']
```

```
: {'n_sq_co': 750001537,
   'perimetre': 4519.26364836,
   'l_ar': '3ème Ardt',
   'surface': 1170882.82818778,
   'geom_x_y': [48.86287238, 2.3600009859],
   'n_sq_ar': 750000003,
   'l_aroff': 'Temple',
   'c_arinsee': 75103,
   'c_ar': 3}
```

```
geo_neighbourhood["features"][0]['properties']
```

```
: {'n_sq_qu': 750000015,
   'n_sq_ar': 750000004,
   'geom_x_y': [48.851585175, 2.36476795387],
   'c_qu': 15,
   'surface': 487264.93707154,
   'l_qu': 'Arsenal',
   'perimetre': 2878.55965556,
   'c_quinsee': 7510403,
   'c_ar': 4}
```

From these objects, I have created 2 Data Frames with Panda library :

Borough data frame

	BoroughNumber	Borough	Latitude	Longitude	BoroughPerimetre
0	3	Temple	48.862872	2.360001	4519.263648
1	7	Palais-Bourbon	48.856174	2.312188	8099.424883
2	13	Gobelins	48.828388	2.362272	11546.546526
3	17	Batignolles-Monceau	48.887327	2.306777	10775.579516
4	20	Ménilmontant	48.863461	2.401188	10704.940486

Neighborhood data frame

	BoroughNumber	NeighNumber	Neighbourhood	NeighLatitude	NeighLongitude	NeighGeoloc	NeighPerimetre
0	4	15	Arsenal	48.851585	2.364768	[48.851585175, 2.36476795387]	2878.559656
1	5	18	Jardin-des-Plantes	48.841940	2.356894	[48.8419401934, 2.35689388962]	4052.729521
2	10	39	Porte-Saint-Martin	48.871245	2.361504	[48.8712446509, 2.36150364735]	3245.891413
3	11	43	Roquette	48.857064	2.380364	[48.8570640408, 2.38036406173]	4973.010557
4	12	46	Picpus	48.830359	2.428827	[48.8303592424, 2.42882681508]	18261.910318
...	...	...	...	...	...	...	...

2 Preparing, Cleaning and Merging Data with the PANDA library

After cleaning and merging the data frames, the new data frame created will be used for the rest of the study:

```
borough_neighbourhoodDF= bouroughDF.set_index('BoroughNumber').join(neighbourhoodDF.set_index('BoroughNumber')).reset_index()
borough_neighbourhoodDF.head()
```

	BoroughNumber	Borough	Latitude	Longitude	BoroughPerimetre	NeighNumber	Neighbourhood	NeighLatitude	NeighLongitude	NeighGeoloc	NeighPerimetre
0	1	Louvre	48.862563	2.336443	6054.936862	3	Palais-Royal	48.864660	2.336309	[48.8646599781, 2.33630891897]	2166.839239
1	1	Louvre	48.862563	2.336443	6054.936862	1	Saint-Germain-l'Auxerrois	48.860650	2.334910	[48.8606501352, 2.33491032928]	5057.549475
2	1	Louvre	48.862563	2.336443	6054.936862	2	Halles	48.862289	2.344899	[48.8622891081, 2.34489885831]	2606.417128
3	1	Louvre	48.862563	2.336443	6054.936862	4	Place-Vendôme	48.867019	2.328582	[48.8670185906, 2.32858166493]	2147.817602
4	2	Bourse	48.868279	2.342803	4554.104360	6	Vivienne	48.869100	2.339461	[48.8691001998, 2.33946074375]	2058.472959

```
borough_neighbourhoodDF.shape
```

```
(80, 11)
```

3 Venues Acquisition with the Foursquare API

After getting venues suggestions and filtering places that deal only food and drink the new data frame created looks like this :

```
df_food = paris_venues [paris_venues ['Venue Category'].str.contains('Restaurant|Bar|Snack|Food|Pizza')].reset_index(drop=True)
df_food.index = np.arange(1, len(df_food)+1)
```

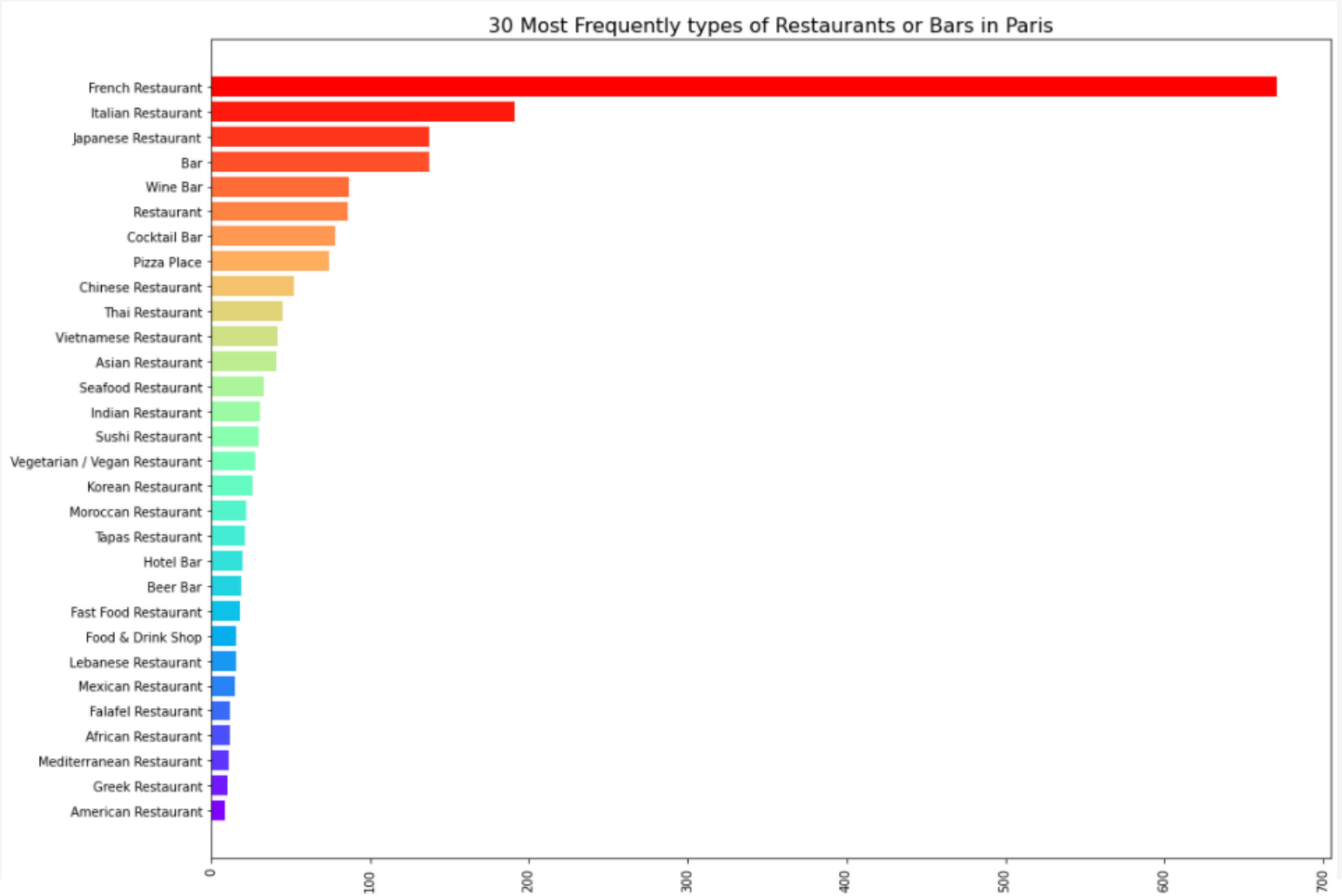
```
df_food
```

	Neighbourhood	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
1	Palais-Royal	48.86466	2.336309	Udon Bistro Kunitoraya (Kunitoraya)	48.865884	2.336782	Udon Restaurant
2	Palais-Royal	48.86466	2.336309	Sanukiya	48.864713	2.333805	Udon Restaurant
3	Palais-Royal	48.86466	2.336309	Brasserie Réjane	48.865486	2.334824	Restaurant
4	Palais-Royal	48.86466	2.336309	Verjus Bar à Vins	48.866306	2.337471	Wine Bar
5	Palais-Royal	48.86466	2.336309	Restaurant Kunitoraya	48.866116	2.336467	Japanese Restaurant
...	...	...	...	...	...	...	...
2232	Charonne	48.85476	2.407430	Le Magnolia	48.858992	2.405873	French Restaurant
2233	Charonne	48.85476	2.407430	Domino's Pizza	48.852447	2.403890	Pizza Place
2234	Charonne	48.85476	2.407430	McDonald's	48.853252	2.410679	Fast Food Restaurant
2235	Charonne	48.85476	2.407430	Royal Kebab	48.853461	2.409690	Kebab Restaurant
2236	Charonne	48.85476	2.407430	Pizzeria du glacier de Venise	48.853790	2.411600	Pizza Place

2236 rows × 7 columns

3 Venues Acquisition with the Foursquare API

By exploring the data, we realize that French restaurants are widely represented in Paris. :



4 One Hot Encoding with the PANDA library

One Hot Encoding is a technique that turns categorical variables into Boolean variables. After applying this technique to the database previously created, I grouped the different categories of restaurants by district of Paris. Then, for each of these neighborhoods, I limited the data to the 10 most common types of restaurants.

	Neighbourhood	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	Amérique	French Restaurant	Health Food Store	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fondue Restaurant
1	Archives	French Restaurant	Italian Restaurant	Bar	Japanese Restaurant	Cocktail Bar	Tapas Restaurant	Falafel Restaurant	Thai Restaurant	Pizza Place	Cajun / Creole Restaurant
2	Arsenal	French Restaurant	Vegetarian / Vegan Restaurant	Cocktail Bar	Italian Restaurant	Thai Restaurant	Tapas Restaurant	Wine Bar	Southwestern French Restaurant	Bar	Brazilian Restaurant
3	Arts-et-Métiers	French Restaurant	Wine Bar	Cocktail Bar	Chinese Restaurant	Italian Restaurant	Restaurant	Bar	Vietnamese Restaurant	Japanese Restaurant	Moroccan Restaurant
4	Auteuil	French Restaurant	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fondue Restaurant	Food
...	...	...	...	...	...	...	...	...	...	...	...
74	Sorbonne	French Restaurant	Bar	Wine Bar	Italian Restaurant	Lebanese Restaurant	Seafood Restaurant	Asian Restaurant	Chinese Restaurant	Cocktail Bar	Ethiopian Restaurant
75	Ternes	French Restaurant	Italian Restaurant	Seafood Restaurant	Japanese Restaurant	Hotel Bar	Moroccan Restaurant	American Restaurant	Asian Restaurant	Breton Restaurant	Cocktail Bar
76	Val-de-Grâce	French Restaurant	Bar	Asian Restaurant	Wine Bar	Pizza Place	Chinese Restaurant	Mexican Restaurant	Lebanese Restaurant	Beer Bar	Korean Restaurant
77	Villette	Bar	French Restaurant	Food Truck	Italian Restaurant	Fast Food Restaurant	Asian Restaurant	Beer Bar	Seafood Restaurant	Japanese Restaurant	Middle Eastern Restaurant
78	Vivienne	French Restaurant	Japanese Restaurant	Wine Bar	Cocktail Bar	Udon Restaurant	Korean Restaurant	Food & Drink Shop	Israeli Restaurant	English Restaurant	Taiwanese Restaurant

79 rows × 11 columns

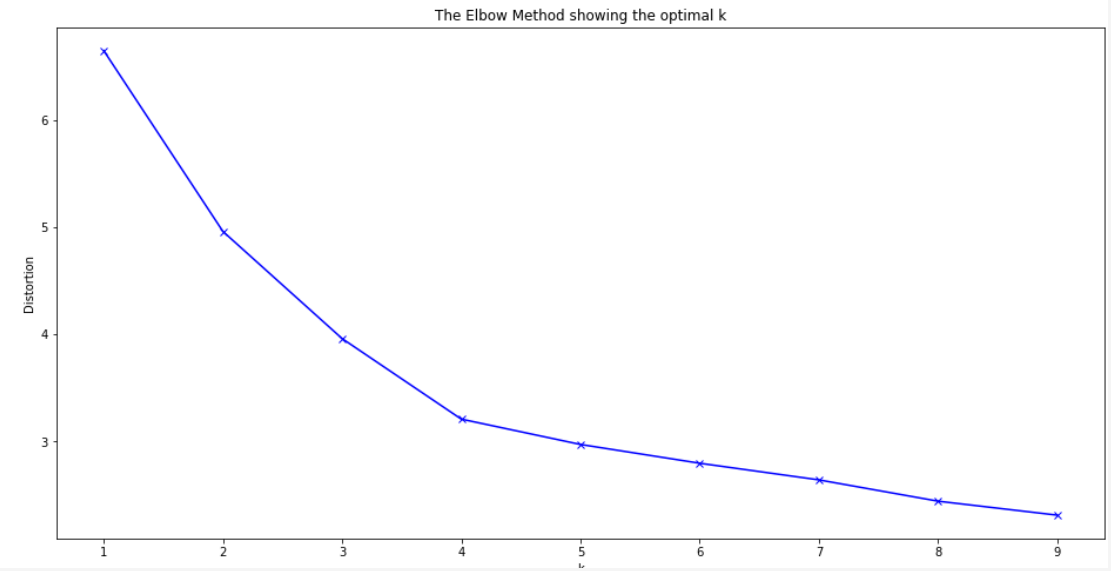
This last data frame will make it possible to execute an unsupervised machine learning algorithm, more precisely a k-means clustering algorithm from the scikit-learn package.

5 Clustering K-means with scikitlearn.cluster library

1. **Ellbow method** to set the value of the optimum k : the graph shows that the optimal value of k is 4.

```
plt.figure(figsize=(16,8))
plt.plot(K, distortions, 'bx-')
plt.xlabel('k')
plt.ylabel('Distortion')
plt.title('The Elbow Method showing the optimal k')
```

```
Text(0.5, 1.0, 'The Elbow Method showing the optimal k')
```



2. **k-means** with k = 4 :

```
from sklearn.cluster import KMeans

k_clusters = 4

#drop the Neighbourhood column to work with numerical values only
df_k_clustering = df_grouped.drop('Neighbourhood', 1)

KM = KMeans(n_clusters=k_clusters, random_state=0)
```

3. The **final data frame** created with clusters labels added

```
#adding the labels to the top10 df
df_venues_sorted.insert(0, 'Cluster Labels', KM.labels_)
df_venues_sorted
```

	Cluster Labels	Neighbourhood	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	2	Amérique	French Restaurant	Health Food Store	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fondue Restaurant
1	3	Archives	French Restaurant	Italian Restaurant	Bar	Japanese Restaurant	Cocktail Bar	Tapas Restaurant	Falafel Restaurant	Thai Restaurant	Pizza Place	Cajun / Creole Restaurant
2	0	Arsenal	French Restaurant	Vegetarian / Vegan Restaurant	Cocktail Bar	Italian Restaurant	Thai Restaurant	Tapas Restaurant	Wine Bar	Southwestern French Restaurant	Bar	Brazilian Restaurant
3	3	Arts-et-Métiers	French Restaurant	Wine Bar	Cocktail Bar	Chinese Restaurant	Italian Restaurant	Restaurant	Bar	Vietnamese Restaurant	Japanese Restaurant	Moroccan Restaurant
4	2	Auteuil	French Restaurant	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fondue Restaurant	Food
...	...	...	...	...	...	...	...	...	...	...	...	...
74	0	Sorbonne	French Restaurant	Bar	Wine Bar	Italian Restaurant	Lebanese Restaurant	Seafood Restaurant	Asian Restaurant	Chinese Restaurant	Cocktail Bar	Ethiopian Restaurant
75	0	Ternes	French Restaurant	Italian Restaurant	Seafood Restaurant	Japanese Restaurant	Hotel Bar	Moroccan Restaurant	American Restaurant	Asian Restaurant	Breton Restaurant	Cocktail Bar
76	3	Val-de-Grâce	French Restaurant	Bar	Asian Restaurant	Wine Bar	Pizza Place	Chinese Restaurant	Mexican Restaurant	Lebanese Restaurant	Beer Bar	Korean Restaurant
77	3	Villette	Bar	French Restaurant	Food Truck	Italian Restaurant	Fast Food Restaurant	Asian Restaurant	Beer Bar	Seafood Restaurant	Japanese Restaurant	Middle Eastern Restaurant
78	3	Vivienne	French Restaurant	Japanese Restaurant	Wine Bar	Cocktail Bar	Udon Restaurant	Korean Restaurant	Food & Drink Shop	Israeli Restaurant	English Restaurant	Taiwanese Restaurant

79 rows × 12 columns

6 Cluster Analysis

Thus, four clusters of types of restaurants were created by the K-Means model for the city of Paris. The clusters were named according to the frequency of the types of restaurants that appear the most among the top 3 most common places in each of the groups.

Cluster 1 (0) - French and Italian Restaurants

```
cluster_1_1stMostCommonVenue = cluster_1['1st Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_1_2ndMostCommonVenue = cluster_1['2nd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_1_3rdMostCommonVenue = cluster_1['3rd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
print(cluster_1_1stMostCommonVenue);print(cluster_1_2ndMostCommonVenue);print(cluster_1_3rdMostCommonVenue)
```

```
                  frequency
French Restaurant         36
                          frequency
Italian Restaurant               17
Bar                               4
Japanese Restaurant               3
Wine Bar                          2
Pizza Place                       2
Vegetarian / Vegan Restaurant     2
Chinese Restaurant                1
Food Truck                        1
Portuguese Restaurant             1
Greek Restaurant                  1
Vietnamese Restaurant             1
Corsican Restaurant               1
                          frequency
Japanese Restaurant               6
Italian Restaurant                5
Wine Bar                          4
Pizza Place                       4
Bar                               3
Vietnamese Restaurant             2
Thai Restaurant                   2
Hotel Bar                         2
Restaurant                        2
Cocktail Bar                      2
Seafood Restaurant                1
Moroccan Restaurant               1
Vegetarian / Vegan Restaurant     1
Fast Food Restaurant              1
```

cluster_1

	Neighbourhood	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	...	BoroughNumber
2	Arsenal	French Restaurant	Vegetarian / Vegan Restaurant	Cocktail Bar	Italian Restaurant	Thai Restaurant	Tapas Restaurant	Wine Bar	Southwestern French Restaurant	Bar	...	4
5	Batignolles	French Restaurant	Italian Restaurant	Bar	Restaurant	Pizza Place	Japanese Restaurant	Turkish Restaurant	Chinese Restaurant	Cocktail Bar	...	17
10	Chaillot	French Restaurant	Italian Restaurant	Japanese Restaurant	Chinese Restaurant	Seafood Restaurant	Bar	Cambodian Restaurant	Cantonese Restaurant	Food & Drink Shop	...	16
11	Champs-Elysées	French Restaurant	Italian Restaurant	Japanese Restaurant	Peruvian Restaurant	Hotel Bar	Thai Restaurant	Cocktail Bar	Seafood Restaurant	Alsatian Restaurant	...	8
							Vegetarian /					
68	Saint-Thomas-d'Aquin	French Restaurant	Italian Restaurant	Bar	Vietnamese Restaurant	Hotel Bar	Peruvian Restaurant	Pizza Place	Restaurant	Cocktail Bar	...	7
69	Saint-Victor	French Restaurant	Wine Bar	Italian Restaurant	Japanese Restaurant	Falafel Restaurant	Bar	Portuguese Restaurant	Turkish Restaurant	Peruvian Restaurant	...	5
72	Sainte-Marguerite	French Restaurant	Italian Restaurant	Bar	Wine Bar	Thai Restaurant	Ethiopian Restaurant	Korean Restaurant	Beer Bar	Restaurant	...	11
74	Sorbonne	French Restaurant	Bar	Wine Bar	Italian Restaurant	Lebanese Restaurant	Seafood Restaurant	Asian Restaurant	Chinese Restaurant	Cocktail Bar	...	5
75	Ternes	French Restaurant	Italian Restaurant	Seafood Restaurant	Japanese Restaurant	Hotel Bar	Moroccan Restaurant	American Restaurant	Asian Restaurant	Breton Restaurant	...	17

36 rows × 21 columns

6 Cluster Analysis

Cluster 2 (1) - Pizza Place and International Cuisine

```
cluster_2_1stMostCommonVenue = cluster_2['1st Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_2_2ndMostCommonVenue = cluster_2['2nd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_2_3rdMostCommonVenue = cluster_2['3rd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
print(cluster_2_1stMostCommonVenue);print(cluster_2_2ndMostCommonVenue);print(cluster_2_3rdMostCommonVenue)
```

```
             frequency
Pizza Place          1
          frequency
Wine Bar          1
            frequency
Food Truck          1
```

```
cluster_2
```

	Neighbourhood	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	...	BoroughNumber	Borough
53	Porte-Dauphine	Pizza Place	Wine Bar	Food Truck	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	...	16	Passy

1 rows × 21 columns

6 Cluster Analysis

Cluster 3 (2) - French Restaurants and Wine Bars

```
cluster_3_1stMostCommonVenue = cluster_3['1st Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_3_2ndMostCommonVenue = cluster_3['2nd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_3_3rdMostCommonVenue = cluster_3['3rd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
print(cluster_3_1stMostCommonVenue);print(cluster_3_2ndMostCommonVenue);print(cluster_3_3rdMostCommonVenue)
```

```
                  frequency
French Restaurant         6
                  frequency
Wine Bar                  2
Health Food Store         1
Italian Restaurant        1
Snack Place               1
Asian Restaurant          1
                           frequency
Wine Bar                           3
Eastern European Restaurant        2
Japanese Restaurant                1
```

```
cluster_3
```

	Neighbourhood	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	...	BoroughNumber	Borough
0	Amérique	French Restaurant	Health Food Store	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	...	19	Buttes-Chaumont
4	Auteuil	French Restaurant	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fondue Restaurant	...	16	Passy
6	Bel-Air	French Restaurant	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	Fondue Restaurant	...	12	Reuilly
17	Ecole-Militaire	French Restaurant	Asian Restaurant	Wine Bar	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	Falafel Restaurant	Fast Food Restaurant	...	7	Palais-Bourbon
32	Invalides	French Restaurant	Italian Restaurant	Japanese Restaurant	Cocktail Bar	Restaurant	Vegetarian / Vegan Restaurant	Food	Doner Restaurant	Eastern European Restaurant	...	7	Palais-Bourbon
41	Muette	French Restaurant	Snack Place	Wine Bar	Food & Drink Shop	Doner Restaurant	Eastern European Restaurant	Empanada Restaurant	English Restaurant	Ethiopian Restaurant	...	16	Passy

6 rows × 21 columns

6 Cluster Analysis

Cluster 4 (3) - French Restaurants

```
cluster_4_1stMostCommonVenue = cluster_4['1st Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_4_2ndMostCommonVenue = cluster_4['2nd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
cluster_4_3rdMostCommonVenue = cluster_4['3rd Most Common Venue'].value_counts()[0:topn].to_frame(name='frequency')
print(cluster_4_1stMostCommonVenue);print(cluster_4_2ndMostCommonVenue);print(cluster_4_3rdMostCommonVenue)
```

```
                      frequency
French Restaurant            23
Bar                           5
Japanese Restaurant           3
Indian Restaurant             2
Cocktail Bar                  2
Italian Restaurant            1
                      frequency
French Restaurant             8
Bar                           7
Japanese Restaurant           7
Wine Bar                      3
Italian Restaurant            3
Restaurant                    2
Asian Restaurant              2
Chinese Restaurant            1
Cocktail Bar                  1
Sushi Restaurant              1
Thai Restaurant               1
                      frequency
Italian Restaurant            7
Wine Bar                      7
Japanese Restaurant           5
Cocktail Bar                  4
French Restaurant             3
Chinese Restaurant            2
Restaurant                    2
Food Truck                    1
Portuguese Restaurant         1
Vegetarian / Vegan Restaurant 1
Asian Restaurant              1
Bar                           1
African Restaurant            1
```

cluster_4

	Neighbourhood	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	...	BoroughNumber	Bo
1	Archives	French Restaurant	Italian Restaurant	Bar	Japanese Restaurant	Cocktail Bar	Tapas Restaurant	Falafel Restaurant	Thai Restaurant	Pizza Place	...	3	Te
3	Arts-et-Métiers	French Restaurant	Wine Bar	Cocktail Bar	Chinese Restaurant	Italian Restaurant	Restaurant	Bar	Vietnamese Restaurant	Japanese Restaurant	...	3	Te
7	Belleville	Bar	French Restaurant	Japanese Restaurant	Italian Restaurant	Pizza Place	African Restaurant	Chinese Restaurant	Cocktail Bar	Indian Restaurant	...	20	Mé
8	Bercy	French Restaurant	Italian Restaurant	Japanese Restaurant	Restaurant	Bar	Beer Bar	Cambodian Restaurant	Chinese Restaurant	Doner Restaurant	...	12	Re
9	Bonne-Nouvelle	Cocktail Bar	French Restaurant	Wine Bar	Japanese Restaurant	Restaurant	Italian Restaurant	Chinese Restaurant	Bar	Thai Restaurant	...	2	Bo
71	Sainte-Avoie	French Restaurant	Restaurant	Chinese Restaurant	Wine Bar	Italian Restaurant	Vietnamese Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant	Health Food Store	...	3	Mé
73	Salpêtrière	Indian Restaurant	French Restaurant	Italian Restaurant	Corsican Restaurant	Chinese Restaurant	Bar	Mediterranean Restaurant	Sushi Restaurant	Wine Bar	...	13	Bu Mo
76	Val-de-Grâce	French Restaurant	Bar	Asian Restaurant	Wine Bar	Pizza Place	Chinese Restaurant	Mexican Restaurant	Lebanese Restaurant	Beer Bar	...	5	
77	Villette	Bar	French Restaurant	Food Truck	Italian Restaurant	Fast Food Restaurant	Asian Restaurant	Beer Bar	Seafood Restaurant	Japanese Restaurant	...	19	
78	Vivienne	French Restaurant	Japanese Restaurant	Wine Bar	Cocktail Bar	Udon Restaurant	Korean Restaurant	Food & Drink Shop	Israeli Restaurant	English Restaurant	...	2	

36 rows × 21 columns

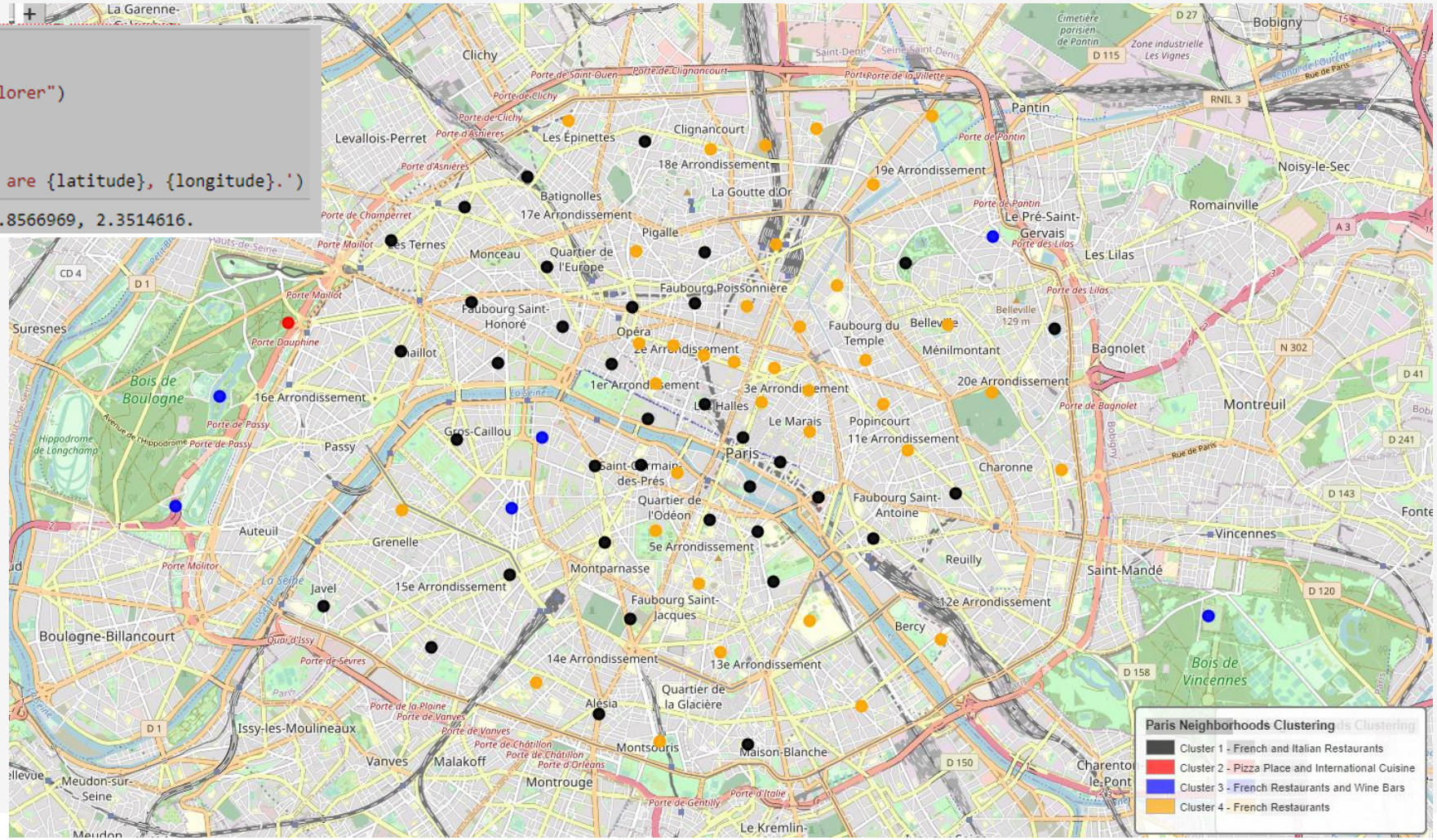
7 Creation of a clusters map with the folium library

The geographic coordinates of Paris were extracted using the GeoPy library in Python.
They were used to draw the map of Paris with colored districts according to the types of restaurants present.

```
address = 'Paris'

geolocator = Nominatim(user_agent="Paris_Explorer")
location = geolocator.geocode(address)
latitude = location.latitude
longitude = location.longitude
print(f'The geograpical coordinates of Paris are {latitude}, {longitude}.')
```

```
The geograpical coordinates of Paris are 48.8566969, 2.3514616.
```



Discussion and Conclusion

Discussion

- French restaurants are widely represented in Paris. Indeed, there are nearly 680.
- In 3 out of 4 clusters reproduced by the K-Means model, there is a large number of French restaurants in the first most frequent place of each neighbourhood.
- Only group 2 (Pizza Place and International Cuisine) does not have a French restaurant. This cluster is made up solely of the Porte Dauphine district located in the 16th arrondissement of Paris.
- We also note that only the Picpus district, located in the 12th arrondissement of Paris, does not have a restaurant.

Conclusion

- The objective of the project presented at the beginning of this presentation has been achieved.
- Indeed, my company has delivered as planned to the Tourist Office of the city of Paris a tool which now allows them to advise tourists on the types of restaurants they could find during their visits to the Parisian districts.