Paris: a world cuisine capital ?

Coursera Applied Data Science Capstone Project by IBM

Alexandre MARTAN

January 25, 2020

Abstract:

This study takes place as part of the Coursera Applied Data Science Capstone Project by IBM meant to apply theoretical knowledge acquired all along previous courses.
The context of this project is to cluster the city of Paris into several groups based on the cuisine served in its restaurants. Results from this analysis could interest anyone who would like to open a restaurant in Paris and would like to find the adequate area, or could be used by tourist when traveling around Paris to find a neighborhood with the food they would like to eat. Moreover, public institutions could find some characteristics in these clusters which could help them providing the best environment for their inhabitants.
The methodology consists in implementing an unsupervised machine learning algorithm, the k-mean algorithm, after determining the optimal number of cluster k. To find this k, we experimented two different methods: first the elbow method which results to be not reliable, then we implement the silhouette method. We then decided to cluster the 80 neighborhoods of Paris into 8 groups.
Results of this project provide a map of Paris with its neighborhood assigned to eight different clusters based on the main cuisine.

Note:
This report is based on a Jupyter Notebook available at the same GitHub repository, in addition of the slides of the presentation.

Table des matières

1. Introduction

1.2 Background

Paris, capital of France, is a model for its culture and life habits. With all its museums and architecture landmarks, the city of Light offers plenty of strolls for everyone and everything, including the inescapable French specialty: the cuisine.
In fact, Paris is acclaimed for its restaurants and its gourmet cuisine, but also for its bistros and bakeries, scattering all around the city.
However, with today's influence of Paris metropolitan area (*Grand Paris*), counting more than seven millions inhabitants (INSEE, 2016, link in French), and receiving millions of tourists each year, Paris basically owns every national cuisine in the world.
With such cosmopolitan area, does one small village of indomitable French dishes still holds out against the invaders tastes?

1.3 Problem

The idea of this project is to segment the Parisian neighborhoods into major clusters depending on the origin of their cuisine.
This origin will be determined by the surrounding restaurants, thanks to an unsupervised machine learning algorithm, namely 'k-mean clustering', and leveraging the Foursquare location data.

1.4 Stakeholders

The results of this analysis can benefit several actors:

- City mappers and institutions: based on the proposed clusters, these actors would be able to visualize if minorities are concentrating their activities in specific neighborhoods thus proposing them adequate accommodations;
- Restaurant owners: knowing the main cuisine of each Parisian neighborhood, someone who wants to open a restaurant in Paris could determine in which area this restaurant could profit;
- Tourists and residents: thanks to this clustering, choosing a place to eat will become easier as long as you know what kind of food you want, then just wander around the beautiful streets of a specialized area and pick the restaurant you're interested in!

2. Data Presentation

2.1 Data sources

- **Paris Data**

Link : https://opendata.paris.fr/explore/dataset/quartier_paris/information/
License : [Open Database License (ODbL)]()

This official dataset provides all the administrative districts of Paris with their surfaces, their localizations and their shapes. These administrative districts consist in a subdivision of the main boroughs (*"Arrondissements"*) into four parts. Therefore, this dataset proposes 80 neighborhoods. In this project we download a CSV file and a GeoJSON file from this source.

With these information, we can display the map of the Figure 1 thanks to the Folium library.

Figure 1:Map of Paris with administrative districts

- **Foursquare Places API**

Link : https://developer.foursquare.com/docs

Foursquare API is a location data provider based on REST API. Given a specific localization, Foursquare API will retrieve all the nearby venues with their categories, such as restaurant or museums.
The most interesting benefits of this API is the hierarchy in their categories : as we will focus on restaurant, Foursquare already know the subdivision of this broad category. A list of this hierarchy can be found here.

For each query, Foursquare API returns a JSON file containing a lot of information, such as name of the venue, its category, its localization and so on.

2.2 Data cleaning

Data provided by our sources represent a lot of information, but sometimes too much or too messy to be useful in the project. In our case, all the details from Paris Data website are not relevant, in fact we only need to retrieve the name of each neighborhood with their localizations and shapes. Other information, such as its surface, its perimeter or its national ID are irrelevant in our scope, thus could be deleted.
About the localization, we need to make our dataset coherent regarding of all our sources. While using Foursquare API, we are not able to retrieve venues inside a polygon surface but only a

square, determined from its south-western point to its north-eastern point. Hence the need to extract only these points inside all the coordinates provided by Paris Data.
Once the desired data rearranged as we would like too, we can begin the analysis to answer our goal, namely clustering the neighborhoods of Paris.

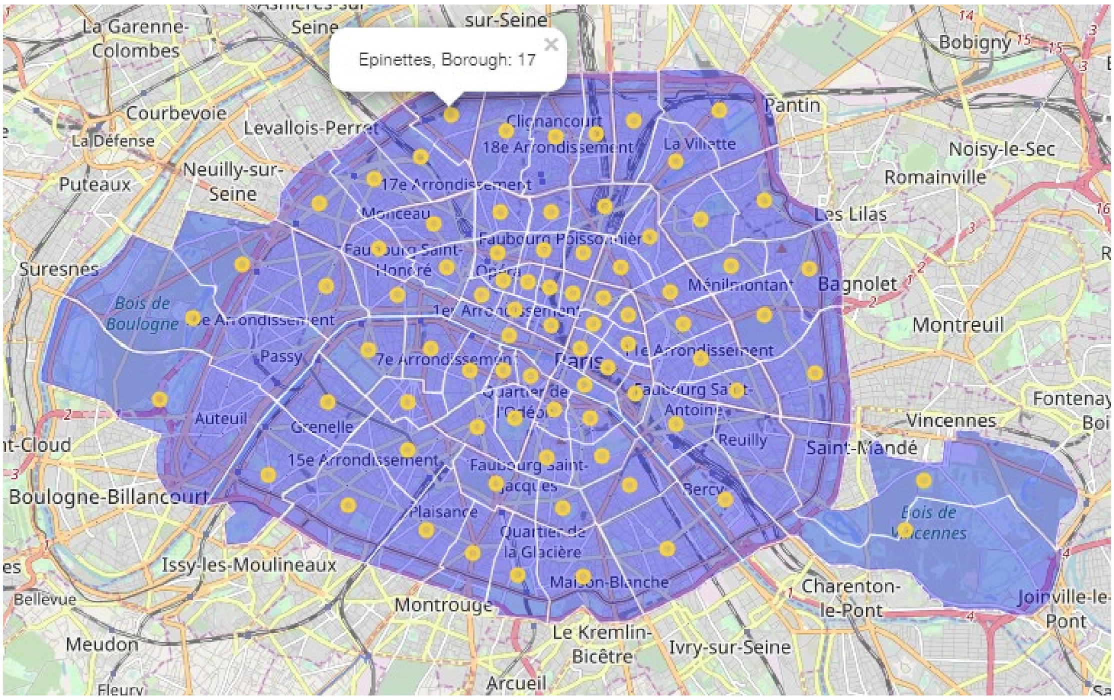


Figure 2:Parisian neighbourhoods analysed in this project

Concerning the Foursquare API data, some functions need to be defined to query venues and extracting the desired information. As our project tends to cluster Paris based on its cuisine, we prefer retrieving only restaurant. Moreover, a limit has to be set for the number of retrieved venues: we choose to set this limit to 100 restaurants.

2.3 Category selection

Once querying the Foursquare API, we are returned 6991 restaurants all around Paris, divided into 123 categories. However, an overview of these categories shows a high redundancy in our data. A perfect example is illustrated by the category called *Restaurant* without further explanations, thus irrelevant in our scope. Others irrelevant categories exists, such as *Cafeteria*, *Gluten-Free Restaurant* or *Vegetarian / Vegan Restaurant*, therefore we decide to remove all of them from our dataset.
Staying on the category names, some of them are subdivision of another, for example we can find Chinese regional cuisine like *Jiangxi, Cantonese* or even *Shandong*. Going so deeply could mislead the clustering algorithm as it would consider these region as a same level than a continent (Asian cuisine represent its own category). Therefore, we regroup some of the category into higher parents to increase weights of low represented country.
Finally, looking at the occurrence for each category, a numerous amount of French restaurant appears, shown by the figure 3.

```
Venue Category
French Restaurant        1831
Italian Restaurant        558
Bakery                    388
Café                      380
Japanese Restaurant       363
                         ...
English Restaurant          1
Fondue Restaurant           1
Provençal Restaurant        1
German Restaurant           1
Cuban Restaurant            1
Name: Venue, Length: 123, dtype: int
```

```
Venue Category
Italian Restaurant        562
Japanese Restaurant       475
Brasserie                 391
Bakery                    388
Pizza Place               220
                         ...
Cuban Restaurant            1
German Restaurant           1
Fondue Restaurant           1
Venezuelan Restaurant       1
Provençal Restaurant        1
Name: Venue, Length: 81, dtype: int
```

Figure 3: Total number of venues in Paris. On the left: before selection, on the right: after selection

With more than 25% of the overall number of restaurants, we can definitely state that French Cuisine is by far the most represented cuisine in Paris. However, this high rate of restaurants could also mislead the algorithm, as all the neighbourhood will tend to be filled by French restaurant. In addition, this *French Restaurant* category is regrouping more detailed categories, such as *Brasseries* or *Creperie*. Thus, we decide to also remove each venue with *French Restaurant* category, but we still keep in our dataset categories about regional French food.

After this selection, we now have 81 categories instead of 123, but only 4008 venues out of the 6991 at the beginning.
However, we can stay confident in our dataset as no neighborhood present a too low number of venues after rearrangement. With a minimum of 14 for only one of them (*Goutte d'Or,* but already only had 24 before selection), most of the other neighborhoods still have more than 40 venues..

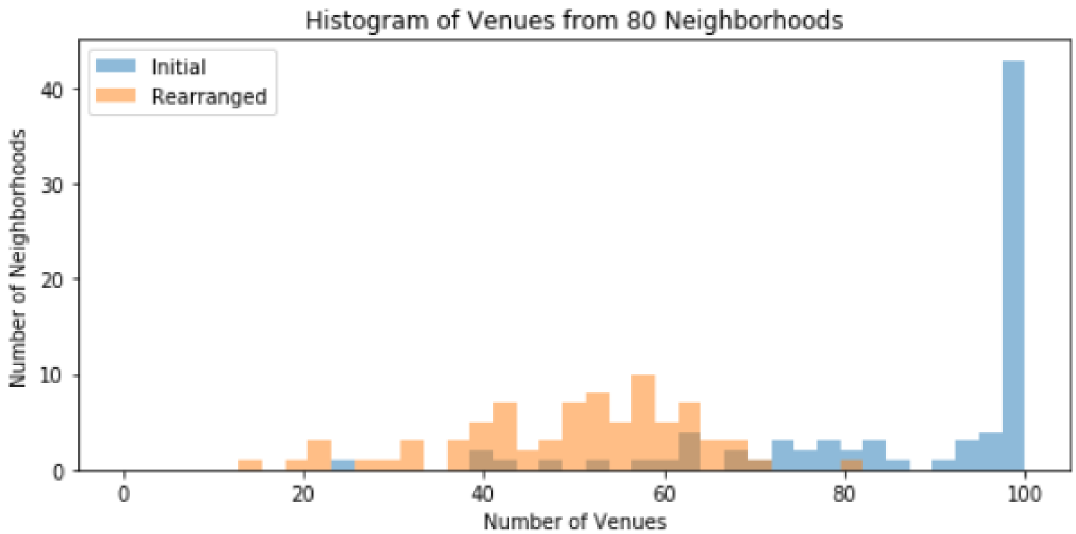


Figure 4: Histogram of the number of venues per neighborhoods before and after selection

3. Methodology

3.1 Determining the optimal number of clusters

When using the k-means algorithm, the number of cluster K shall be manually chosen. If most of the time, the data analysis provides some hints on this number, in this case of clustering a whole city based on its neighborhood cuisines, this number K is difficult to determine. However, several methods can provide an idea of this optimal number of cluster, such as the elbow method.

3.1.1 Elbow method

This method looks at the percentage of variance in function of the number of clusters K. In fact, the representative curve is supposed to present a clear disruption when the number of cluster is optimal, as more clusters would not bring more information.

Using this method, we can plot the Figure 5 hereunder:

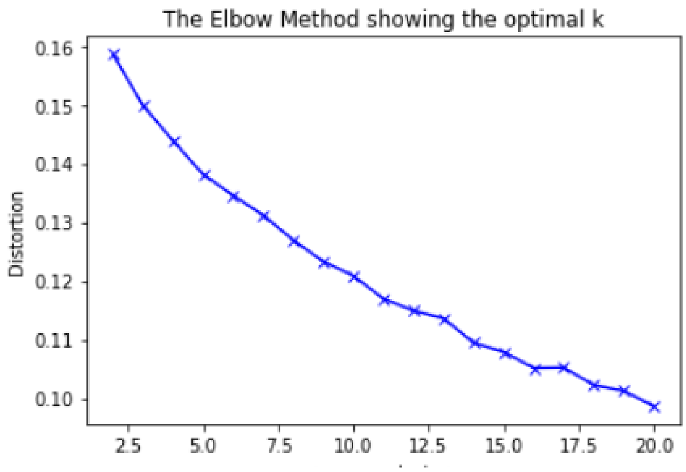


Figure 5:Distortion in function of the number of clusters K

However, this figure is not the one we expected: instead of a curve in two parts with a disruptive point, we can see a slowly decreasing curve. Based on that figure we can only understand that the more clusters we choose, the better the results should be. Yet choosing 20 clusters when having only 81 different categories does not fulfill our requirement to cluster Paris into few clusters. Another method need to be used.

3.1.2 Silhouette method

To tackle the issue raised by the previous method, we decide to change method and use the silhouette method, known to be more reliable.
This method is a graphical representation of how well each object has been classified, based on the computation of a distance metric.

Using this method, we first decide to plot the figure 6, representing the average silhouette score in function of the number of clusters. In this graph, the highest the average score is, the better the clustering is. However, we need to compare each elements of each cluster to one another and not just rely on the average score, therefore the figure 6 provides an overview to limit the number of clusters to analyses.

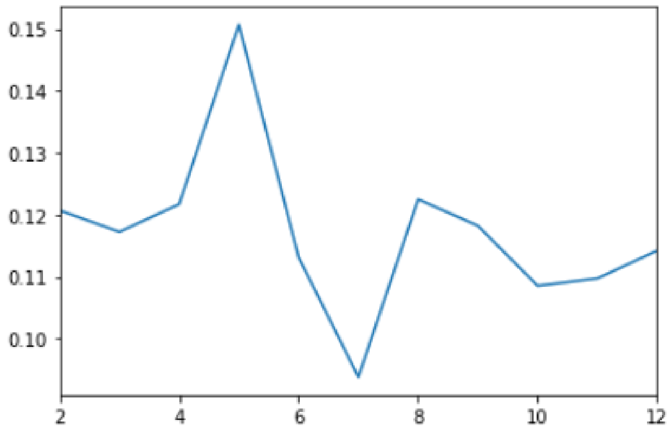
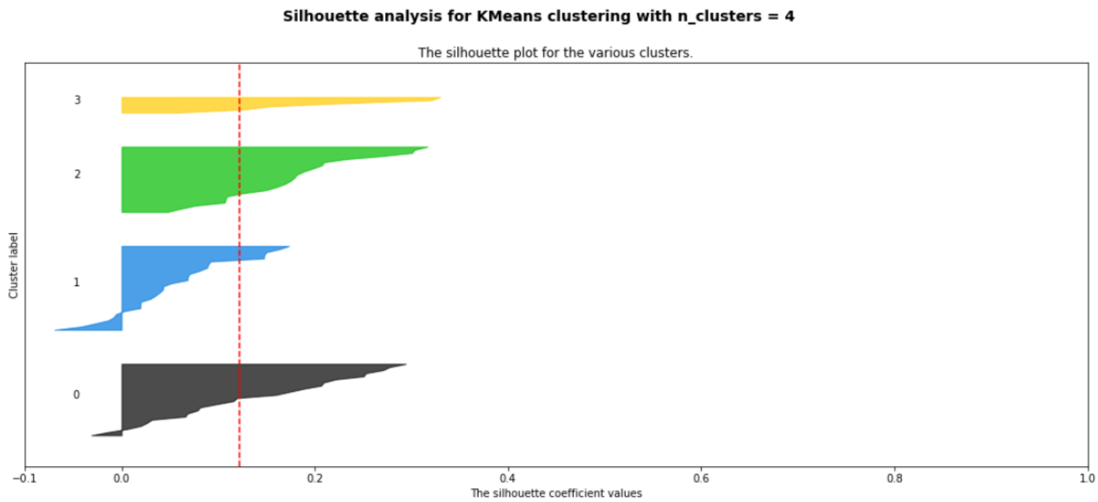


Figure 6: Average silhouette score in function of the number of clusters

With the figure 6, we can understand a good choice would be to choose 5 clusters, howver 4 and 8 or 9 appear to be a great solution too. To make our choice, we can plot the graphs of the figure 7.

As a brief explanation, in the graph of the Figure 7, a negative value indicates the item was assigned to an incorrect cluster. The red line is the average silhouette score. The aim is to find a graph with the lower rate of negative values and all the cluster passing by the average silhouette score. In addition, avoiding unique item in a cluster is better.



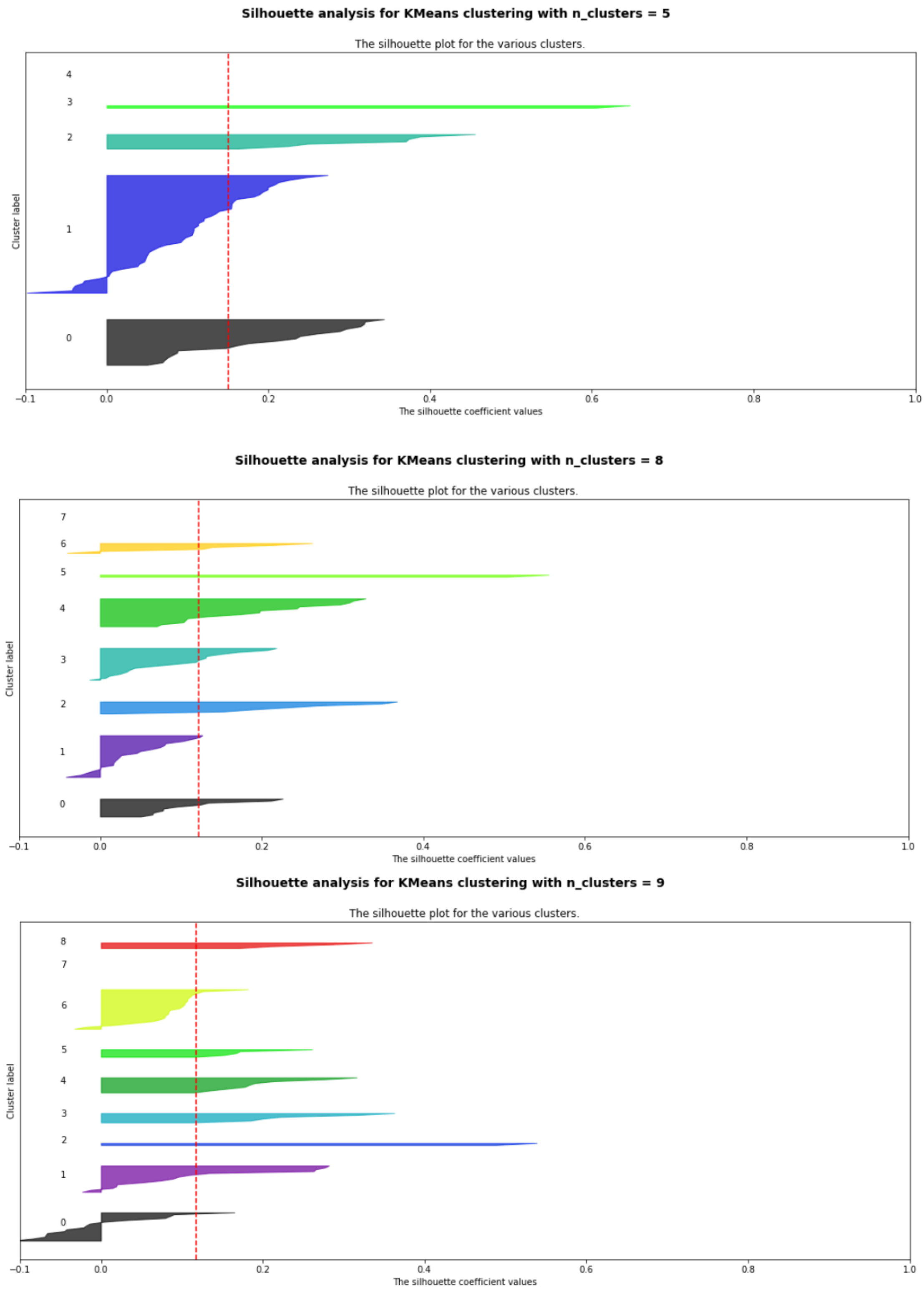


Figure 7: 4 graphs used by the Silhouette method

Based on these previous graphs, we choose to cluster Parisian neighborhoods into 8 groups. This number seems to provide a good average silhouette score, several groups of similar size and not too many incorrect assignments.

3.2 k-means algorithm

To cluster the Parisian neighborhoods into 8 groups, we select the k-means algorithm thanks to its simplicity, either to understand it and to implement it.
After randomly choosing the 8 first seeds, we obtained the results plotted on the map of the figure 8. This map is interactive thanks to the Folium library and can be found on the notebook which was the support of this report.

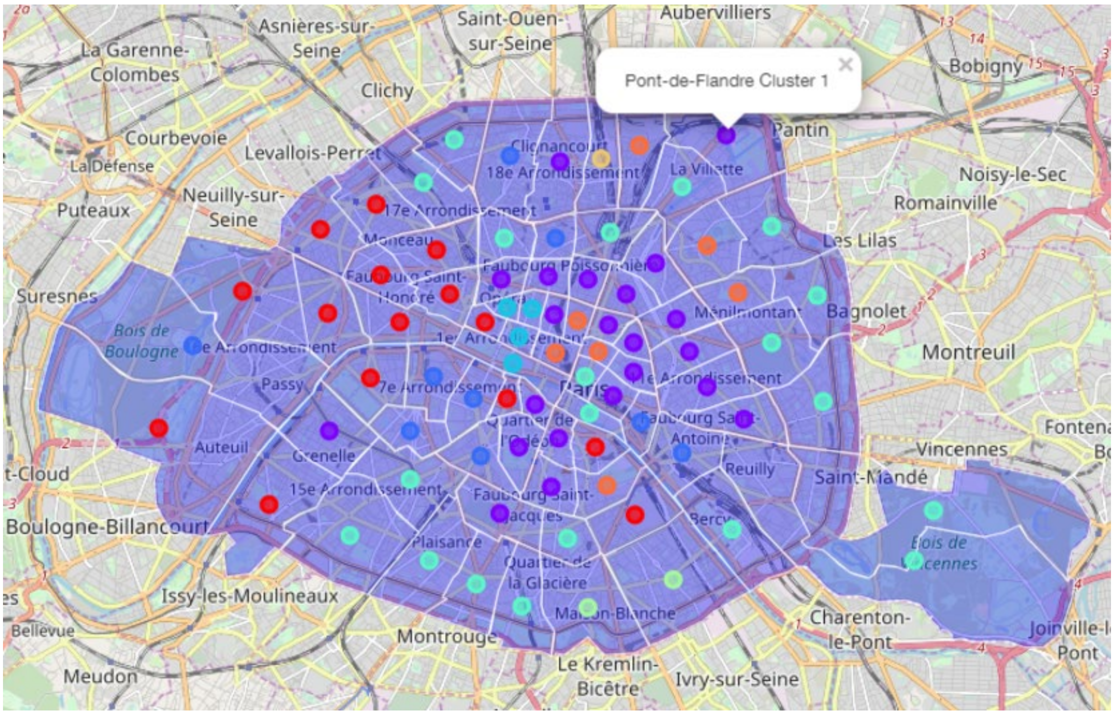


Figure 8: Map of Paris divided into 8 clusters

4. Results

4.1 Presentation of the eight clusters

In the next paragraphs, we will show the composition of each clusters and propose a classification name, namely the main region of cuisine, which group them.

4.1.1 Cluster 0: When Italy meets Japan

With nearly only Italian Restaurant as a 1st most common venue then followed by Japanese Restaurant as the 2nd most common place, this cluster actually regroups the most numerous place inside Paris.

Without being wrong we can state this cluster is Italian/Japanese.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
4	Chaillot	0	Italian Restaurant	Brasserie	Japanese Restaurant	Thai Restaurant	Lebanese Restaurant
12	Auteuil	0	Italian Restaurant	Japanese Restaurant	Brasserie	Bakery	Middle Eastern Restaurant
13	Saint-Victor	0	Italian Restaurant	Japanese Restaurant	Bakery	Brasserie	Burger Joint
14	Madeleine	0	Italian Restaurant	Brasserie	Japanese Restaurant	Thai Restaurant	Asian Restaurant
16	Porte-Dauphine	0	Italian Restaurant	Japanese Restaurant	Bakery	Chinese Restaurant	Pizza Place
20	Gros-Caillou	0	Italian Restaurant	Brasserie	Bakery	Japanese Restaurant	Pizza Place
28	Salpêtrière	0	Italian Restaurant	Japanese Restaurant	Chinese Restaurant	Bakery	Brasserie
29	Place-Vendôme	0	Japanese Restaurant	Italian Restaurant	Bakery	Brasserie	Tapas Restaurant
32	Javel	0	Italian Restaurant	Japanese Restaurant	Bakery	Brasserie	Persian Restaurant
38	Champs-Elysées	0	Italian Restaurant	Bakery	Japanese Restaurant	Steakhouse	Brasserie
43	Saint-Germain-des-Prés	0	Italian Restaurant	Japanese Restaurant	Brasserie	Greek Restaurant	Chinese Restaurant
48	Plaine de Monceaux	0	Italian Restaurant	Bakery	Japanese Restaurant	Asian Restaurant	Mediterranean Restaurant
61	Faubourg-du-Roule	0	Italian Restaurant	Japanese Restaurant	Bakery	Brasserie	Thai Restaurant
62	Europe	0	Italian Restaurant	Japanese Restaurant	Bakery	Asian Restaurant	Thai Restaurant
69	Ternes	0	Italian Restaurant	Bakery	Japanese Restaurant	Asian Restaurant	Moroccan Restaurant

Table 1: Results of the Cluster 0

4.1.2 Cluster 1: The indomitable Gauls

Surprisingly, even after removing most of the French Restaurant, the category *Brasserie* and *Bakery* are still part of the most important ones, followed by *Creperie*, famous Brittany specialty.

Even more interesting, this cluster is the biggest of the eight.

This cluster can then be assigned as French.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	Pont-de-Flandre	1	Brasserie	Japanese Restaurant	Asian Restaurant	Burger Joint	Pizza Place
7	Faubourg-Montmartre	1	Italian Restaurant	Thai Restaurant	Brasserie	Japanese Restaurant	Pizza Place
11	Grenelle	1	Italian Restaurant	Brasserie	Bakery	Pizza Place	Japanese Restaurant
23	Porte-Saint-Martin	1	Pizza Place	Italian Restaurant	Brasserie	Burger Joint	Indian Restaurant
24	Roquette	1	Brasserie	Japanese Restaurant	Italian Restaurant	Pizza Place	Vietnamese Restaurant
27	Folie-Méricourt	1	Brasserie	Italian Restaurant	Pizza Place	Vietnamese Restaurant	Bakery
36	Monnaie	1	Italian Restaurant	Brasserie	Japanese Restaurant	Pizza Place	Creperie
37	Odéon	1	Brasserie	Italian Restaurant	Japanese Restaurant	Bakery	New American Restaurant
42	Enfants-Rouges	1	Brasserie	Japanese Restaurant	Italian Restaurant	Burger Joint	Bakery
45	Saint-Ambroise	1	Brasserie	Italian Restaurant	Bakery	Pizza Place	Japanese Restaurant
47	Montparnasse	1	Creperie	Brasserie	Japanese Restaurant	Italian Restaurant	Pizza Place
49	Val-de-Grâce	1	Italian Restaurant	Brasserie	Japanese Restaurant	Creperie	Chinese Restaurant
53	Saint-Gervais	1	Italian Restaurant	Brasserie	Bakery	Creperie	Japanese Restaurant
55	Porte-Saint-Denis	1	Pizza Place	Brasserie	Burger Joint	Bakery	Indian Restaurant
59	Arts-et-Métiers	1	Chinese Restaurant	Japanese Restaurant	Italian Restaurant	Brasserie	Vietnamese Restaurant
60	Archives	1	Japanese Restaurant	Italian Restaurant	Thai Restaurant	Creperie	Brasserie
63	Sainte-Marguerite	1	Italian Restaurant	Brasserie	Pizza Place	Japanese Restaurant	Vietnamese Restaurant
67	Hôpital-Saint-Louis	1	Brasserie	Indian Restaurant	Chinese Restaurant	Italian Restaurant	Pizza Place
71	Mail	1	Italian Restaurant	Bakery	Thai Restaurant	Burger Joint	Japanese Restaurant
74	Clignancourt	1	Brasserie	Italian Restaurant	Pizza Place	Bakery	Vietnamese Restaurant
76	Sorbonne	1	Bakery	Italian Restaurant	Creperie	Brasserie	Burger Joint
78	Chaussée-d'Antin	1	Brasserie	Italian Restaurant	Japanese Restaurant	Burger Joint	Chinese Restaurant

Table 2: Results of the Cluster 1

4.1.3 Cluster 2: Not so far from each other

In this cluster, Italian cuisine comes back to the top followed by some bakeries and other brasseries, typically French.

This cluster can be assigned as Italian/French

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Notre-Dame-des-Champs	2	Italian Restaurant	Bakery	Brasserie	Japanese Restaurant	Creperie
3	Muette	2	Italian Restaurant	Bakery	Japanese Restaurant	Brasserie	Pizza Place
6	Invalides	2	Brasserie	Bakery	Italian Restaurant	Japanese Restaurant	Pizza Place
10	Ecole-Militaire	2	Italian Restaurant	Bakery	Brasserie	Spanish Restaurant	Japanese Restaurant
17	Grandes-Carrières	2	Italian Restaurant	Brasserie	Bakery	Pizza Place	Spanish Restaurant
21	Arsenal	2	Italian Restaurant	Bakery	Gastropub	Tapas Restaurant	Pizza Place
33	Rochechouart	2	Bakery	Italian Restaurant	Japanese Restaurant	Pizza Place	Burger Joint
54	Saint-Thomas-d'Aquin	2	Italian Restaurant	Bakery	Brasserie	Japanese Restaurant	Pizza Place
58	Quinze-Vingts	2	Italian Restaurant	Bakery	Pizza Place	Vietnamese Restaurant	Creperie

Table 3: Results of the Cluster 2

4.1.4 Cluster 3: *Itadakimasu! (Bon appetite!)*

With only 4 items, this clusters represents the highest rate of Japanese restaurant, each neighborhood having 30% of their restaurant serving Japanese Cuisine.

The cluster 3 is thus assigned as Japanese.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
8	Gaillon	3	Japanese Restaurant	Italian Restaurant	Korean Restaurant	Burger Joint	Asian Restaurant
41	Vivienne	3	Japanese Restaurant	Brasserie	Italian Restaurant	Bakery	Thai Restaurant
56	Saint-Germain-l'Auxerrois	3	Japanese Restaurant	Bakery	Italian Restaurant	Pizza Place	Brasserie
79	Palais-Royal	3	Japanese Restaurant	Italian Restaurant	Bakery	Brasserie	Korean Restaurant

Table 4: Results of the Cluster 3

4.1.5 Cluster 4: Across the world in one street

Similar to the cluster 3 from the main presence of Japanese restaurant as 1st most common restaurant, their frequencies is a lot lower than the other cluster but has more variety. Another main cuisine in this cluster can be seen with the French usual brasseries and bakeries.

Therefore, we choose to assign this cluster as Japanese/French.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Petit-Montrouge	4	Brasserie	Italian Restaurant	Japanese Restaurant	Bakery	Vietnamese Restaurant
5	Epinettes	4	Japanese Restaurant	Brasserie	Italian Restaurant	Pizza Place	Bakery
9	Amérique	4	Japanese Restaurant	Bakery	Pizza Place	Brasserie	Fast Food Restaurant
15	Saint-Fargeau	4	Bakery	Japanese Restaurant	Brasserie	Italian Restaurant	Pizza Place
18	Saint-Merri	4	Bakery	Brasserie	Burger Joint	Hot Dog Joint	Portuguese Restaurant
19	Notre-Dame	4	Japanese Restaurant	Brasserie	Bakery	Thai Restaurant	Creperie
25	Picpus	4	Italian Restaurant	Japanese Restaurant	Brasserie	Bakery	Thai Restaurant
26	Plaisance	4	Japanese Restaurant	Bakery	Brasserie	Pizza Place	Indian Restaurant
31	Charonne	4	Japanese Restaurant	Pizza Place	Brasserie	Italian Restaurant	Fast Food Restaurant
34	Bercy	4	Japanese Restaurant	Bakery	Italian Restaurant	Brasserie	Burger Joint
40	Croulebarbe	4	Japanese Restaurant	Italian Restaurant	Thai Restaurant	Bakery	Brasserie
44	Saint-Vincent-de-Paul	4	Indian Restaurant	Japanese Restaurant	Bakery	Brasserie	Italian Restaurant
46	Bel-Air	4	Japanese Restaurant	Bakery	Italian Restaurant	Brasserie	Pizza Place
50	Necker	4	Japanese Restaurant	Italian Restaurant	Korean Restaurant	Bakery	Brasserie
51	Père-Lachaise	4	Bakery	Japanese Restaurant	Brasserie	Italian Restaurant	Pizza Place
57	Villette	4	Japanese Restaurant	Bakery	Brasserie	Italian Restaurant	Chinese Restaurant
64	Parc-de-Montsouris	4	Italian Restaurant	Japanese Restaurant	Pizza Place	Bakery	Indian Restaurant
65	Saint-Lambert	4	Bakery	Italian Restaurant	Japanese Restaurant	Pizza Place	Lebanese Restaurant
70	Batignolles	4	Japanese Restaurant	Italian Restaurant	Pizza Place	Steakhouse	Chinese Restaurant
77	Saint-Georges	4	Italian Restaurant	Japanese Restaurant	Bakery	Chinese Restaurant	Brasserie

Table 5: Results of the Cluster 4

4.1.6 Cluster 5: Asian flavors

Composed by only two neighborhoods, this cluster is really different from all the others as its main venues are Asian restaurants.

This cluster can be labeled Asian.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
39	Maison-Blanche	5	Vietnamese Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Bakery
73	Gare	5	Vietnamese Restaurant	Chinese Restaurant	Asian Restaurant	Thai Restaurant	Italian Restaurant

Table 6: Results of the Cluster 5

4.1.7 Cluster 6: Eat on the go

As the smallest cluster with only a unique element, this place also has a strong specificity: its main venues are places to eat quick or to take-away.

This cluster is a Fast-Food place.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
75	Goutte-d'Or	6	Fast Food Restaurant	Pizza Place	Brasserie	Indian Restaurant	Mexican Restaurant

Table 7: Results of the Cluster 6

4.1.8 Cluster 7: A complex link

Based on the frequency of the venues in these Neighborhoods, a distinction is more difficult for the cluster 7. In fact, if a common point is with at least 10% of these venues are Chinese restaurant, each neighborhood has their own specificities.

This cluster need to be considered as a mix cuisine place.

	Neighborhood	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
22	Jardin-des-Plantes	7	Chinese Restaurant	Japanese Restaurant	Italian Restaurant	Bakery	Greek Restaurant
30	Combat	7	Chinese Restaurant	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Bakery
35	Halles	7	Bakery	Pizza Place	Italian Restaurant	Chinese Restaurant	Brasserie
52	La Chapelle	7	Bakery	Indian Restaurant	Asian Restaurant	Thai Restaurant	Brasserie
66	Sainte-Avoie	7	Chinese Restaurant	Bakery	Italian Restaurant	Thai Restaurant	Burger Joint
68	Belleville	7	Chinese Restaurant	Bakery	Vietnamese Restaurant	Pizza Place	Thai Restaurant
72	Bonne-Nouvelle	7	Bakery	Pizza Place	Turkish Restaurant	Thai Restaurant	Italian Restaurant

Table 8: Results of the Cluster 7

4.2 Summary

This paragraph proposes a single table to retrieve the key characteristics of each cluster.

Cluster	Main region of cuisine
0	Italian – Japanese
1	French
2	Italian - French
3	Japanese
4	Japanese – French
5	Asian
6	Fast-Food
7	Mix

5. Discussion

Thanks to the results of this study, we can know see that Parisian neighborhoods can be regrouped into eight different clusters based on the kind of food served in their restaurants.

If some of the cluster only possess few items, they are probably the most recognizable with unique characteristics, contrary to bigger clusters mixing different culture.

However, this study had to simplify some aspects of the data, either to be able to use a tool (Foursquare API) which resulted as mixing several neighborhoods together, or to avoid a too frequent kind of cuisine, the French one, which could have biased the results.

6. Conclusion

In this study, we divided the 80 Parisian neighborhoods into 8 clusters represented by a special type of cuisine thanks to an unsupervised machine learning algorithm: the k-means algorithm.

As targeted; this survey can be used by someone who would like to open a restaurant in the best place to target the adequate costumers, but can also be used by tourists who would like to wander in a neighborhood with a special meal. Finally, public institutions could use this information to understand how the different minority are living in Paris.

In the last case, it would be interesting to had more socio-economical parameter when clustering the neighborhoods in order to take in account the real-estate price, the average income of the area and so on.