

IBM Applied Data Science Capstone

# The battle of Neighborhoods

## Where to eat in Paris?

---



Leyla Villarroel, January 2021

---

---

## Introduction

Paris is one of the world's most iconic cities and the third most visited city in the world [1]. The French capital is a global center of finance, fashion, arts, science, gastronomy, and history. From museums and historical boulevards to legendary restaurants and high-end boutiques, this city can be very intimidating and overwhelming to visit.

## Business Problem

Paris is a multicultural city, where people from different parts of the world live. Therefore, there is a wide variety of food offers and from many nationalities. It may be overwhelming to choose where to eat and what kind of food to eat. This is why it would be really useful to classify the arrondissements of Paris by the nationality of the food their restaurants offer. This way, as a tourist, you would know what kind of food each arrondissement offers and it would be easier to choose an arrondissement to eat

## Target Audience

This project is particularly useful for tourists who want to know where to eat some kind of specific food. This way they will know what arrondissement offers which kind of food.

## Data

For this project I will use three sources of data:

1. Foursquare API
2. Geopy package
3. Paris arrondissements' [2]

From sources 2 and 3, the following DataFrame will be obtained.

---

	Arrondissement	Latitude	Longitude
0	1	48.864671	2.330138
1	2	48.867534	2.343925
2	3	48.864425	2.359444
3	4	48.856137	2.357145
4	5	48.846095	2.348502
5	6	48.852813	2.342799
6	7	48.854757	2.324789
7	8	48.876250	2.324441
8	9	48.875781	2.327347
9	10	48.873690	2.348787

Table 1: coordinates of arrondissements.

From source 1 and the DataFrame above, the following DataFrame will be obtained.

	Arrondissement	Arrondissement Latitude	Arrondissement Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	1	48.864671	2.330138	Librairie Galignani	48.864989	2.328570	Bookstore
1	1	48.864671	2.330138	Grand Bassin Rond	48.862880	2.329295	Fountain
2	1	48.864671	2.330138	Balagan	48.865432	2.329680	Israeli Restaurant
3	1	48.864671	2.330138	Le Dalí	48.865333	2.328137	French Restaurant
4	1	48.864671	2.330138	Ladurée	48.866121	2.328449	Dessert Shop
5	1	48.864671	2.330138	Place des Pyramides	48.863924	2.332224	Plaza
6	1	48.864671	2.330138	Balenciaga	48.865927	2.330284	Women's Store
7	1	48.864671	2.330138	Hôtel Le Meurice	48.865333	2.328137	Hotel
8	1	48.864671	2.330138	Angelina	48.865090	2.328443	Tea Room
9	1	48.864671	2.330138	Hôtel Le Pradey	48.864459	2.331654	Hotel

Table 2: data of venues in each arrondissement.

## Methodology

### Data extraction and processing

The data was extracted from three sources:

1. The Paris arrondissements' from wikipedia.
2. The latitude and longitude of each arrondissement from the geopy package
3. The restaurants from the Foursquare API.

The processing of the data was to filter the venues obtained from the Foursquare API to get only the restaurants. Then, filter the most common types of restaurants.

### Exploratory analysis

---

Obtain the distribution of the restaurants between arrondissements, the distribution of the type of restaurants and the quantity of each type in each arrondissement.

## Clustering by type of food offered

I decided to cluster the arrondissement by the type of food they offer, this way I could get the arrondissement that are alike and help the tourist choose an arrondissement to eat. The algorithm used for clustering was kMeans with euclidean distance, because it's simple and easy to use. In this algorithm it is necessary to know the number of partitions to be made, therefore I use the elbow method to find the right number of partitions.

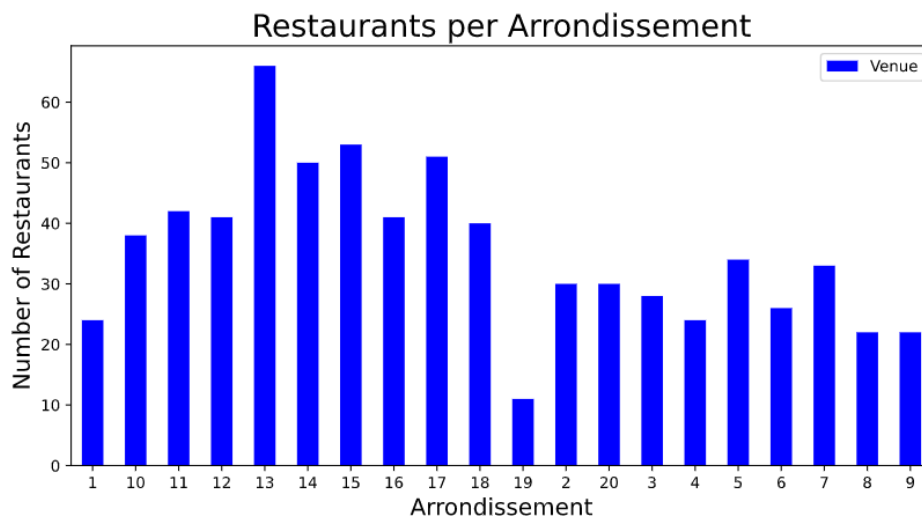
## Results

The results of the exploratory analysis are shown below:

### Quantity of restaurants per arrondissement:

The number of restaurants in each arrondissement varies between 11 and 66, the mean is 35.3 restaurants per arrondissement, and the median is 33.5.

The arrondissement with more restaurants is the 13th and the one with less restaurants is the 19th.



Graph 1: number of restaurants per arrondissement.

---

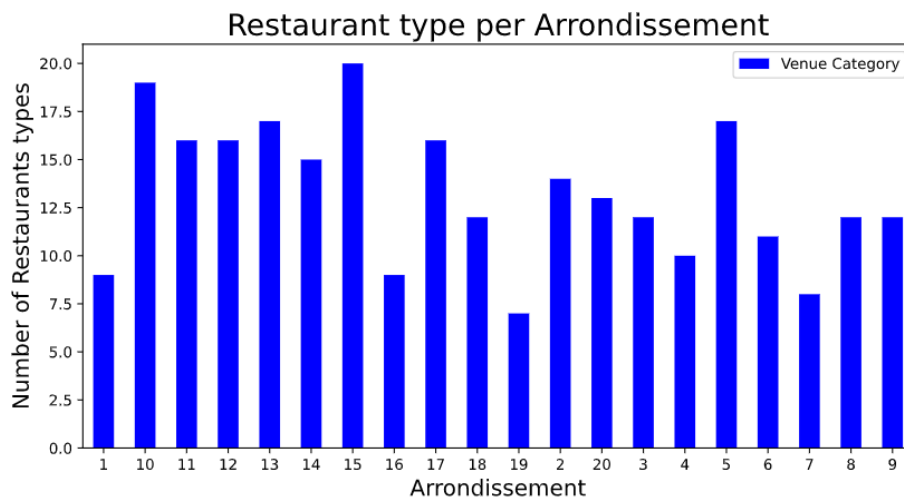
### Quantity of restaurants types per arrondissement:

The number of types of restaurants in each arrondissement varies between 7 and 20, the mean is 13.25 and the median is 12.5 types of restaurants per arrondissement.

The arrondissement with more diversity of restaurants is the 15th and the one with less diversity is the 19th.

The least diverse arrondissement is also the one with less restaurants, which can be a factor that influences the limited types of restaurants out there.

Other arrondissements with little diversity in the type of food they offer are 1st, 7th and 16th. This behavior is no longer explained by the number of restaurants in the arrondissement, since these have twice as many restaurants as the 19th and only have 1 extra type of food offered than this arrondissement.

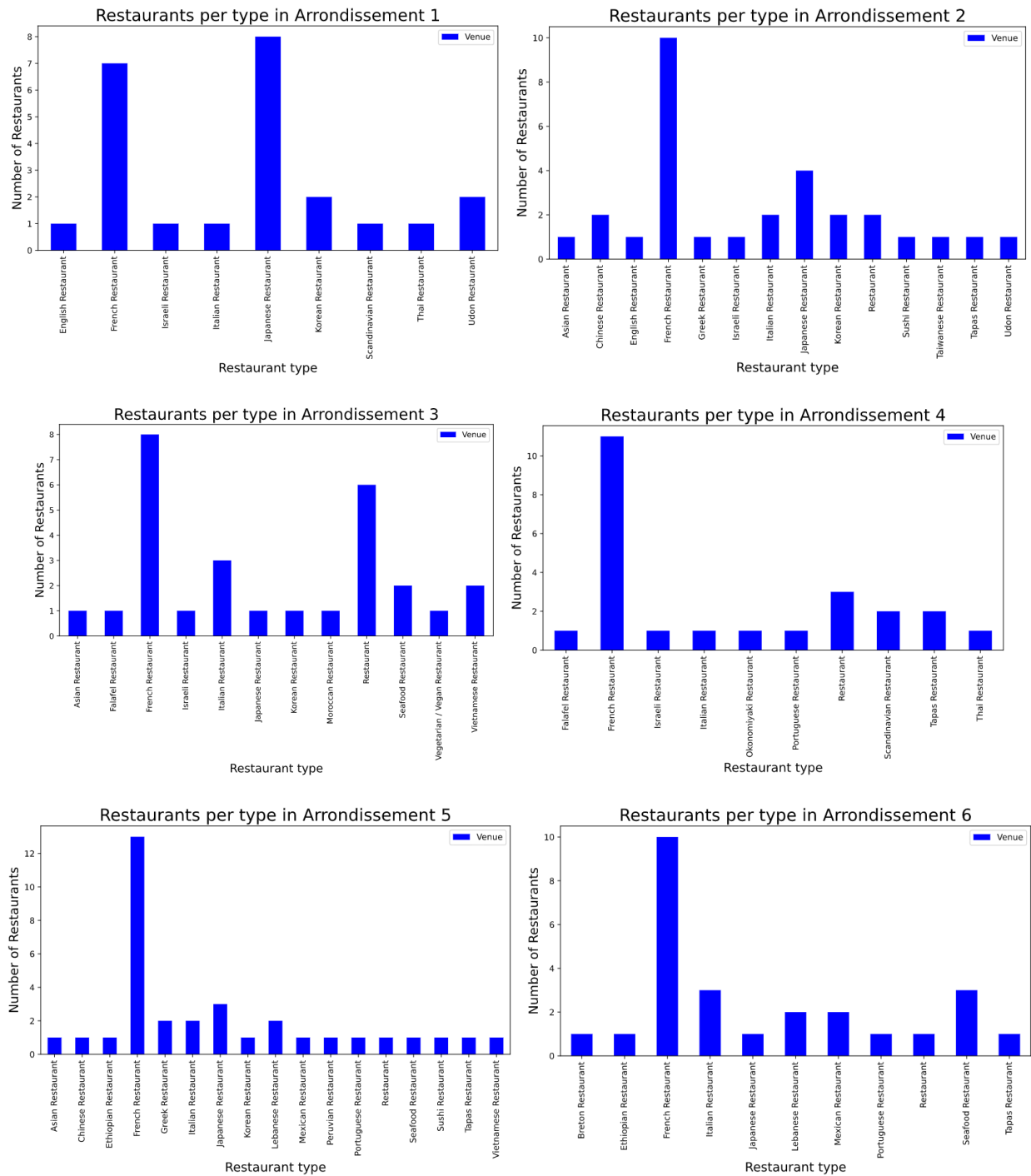


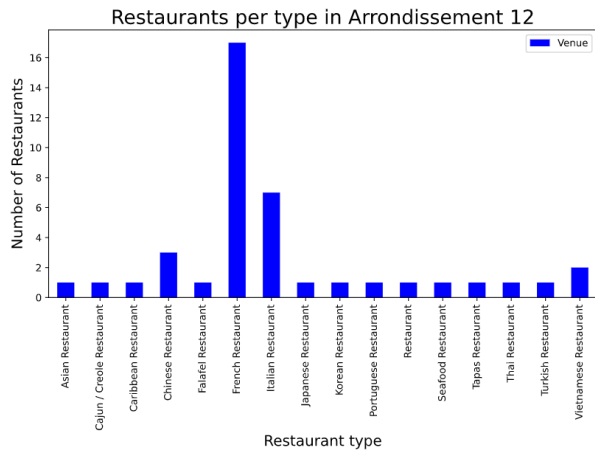
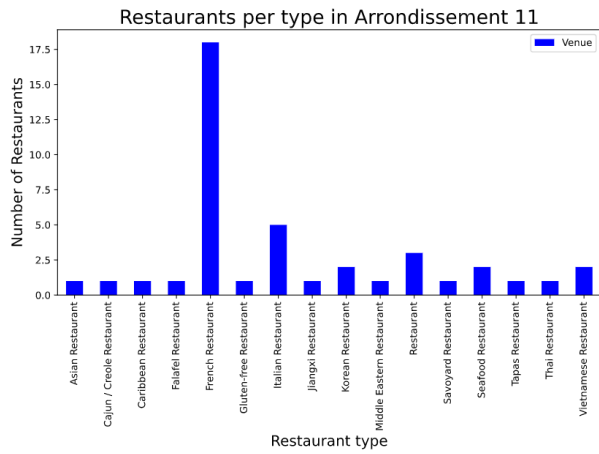
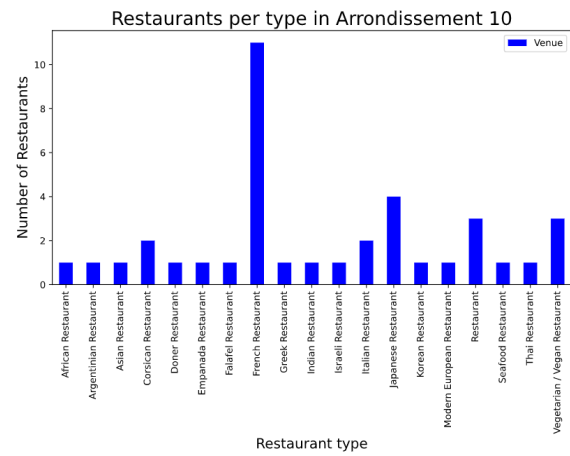
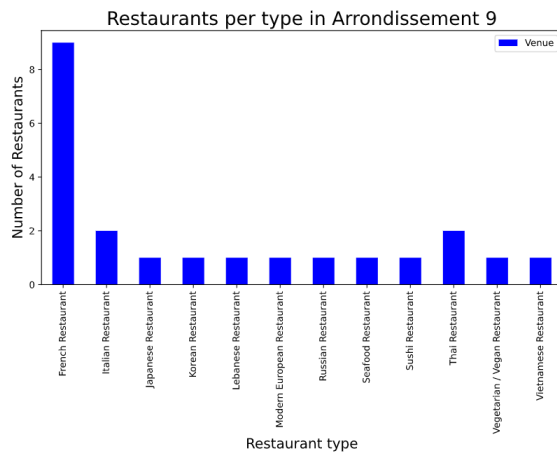
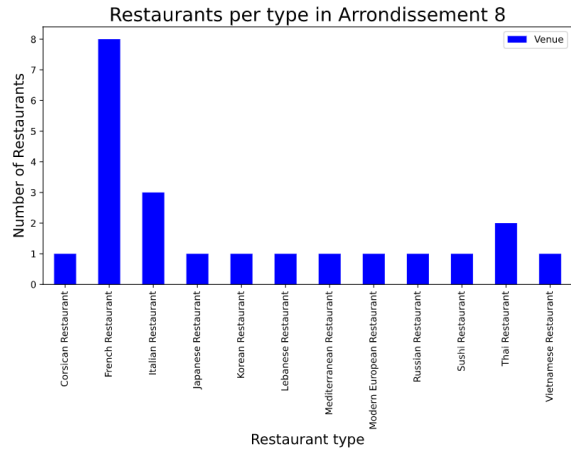
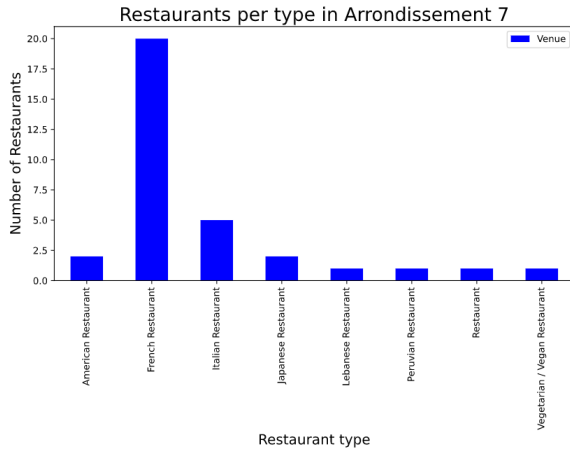
Graph 2: number of restaurants types per arrondissement.

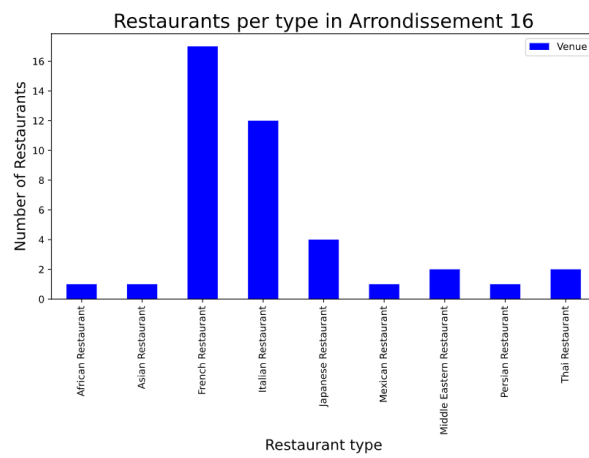
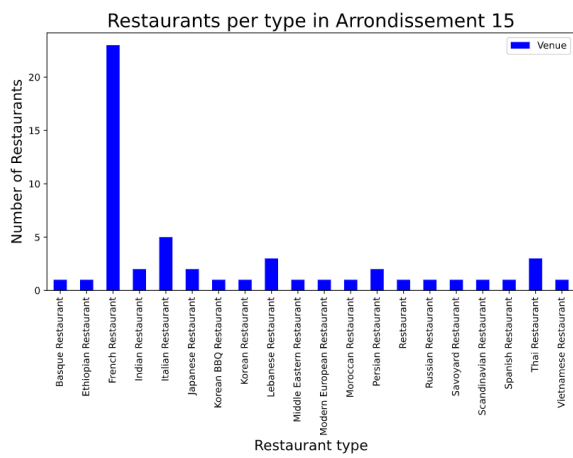
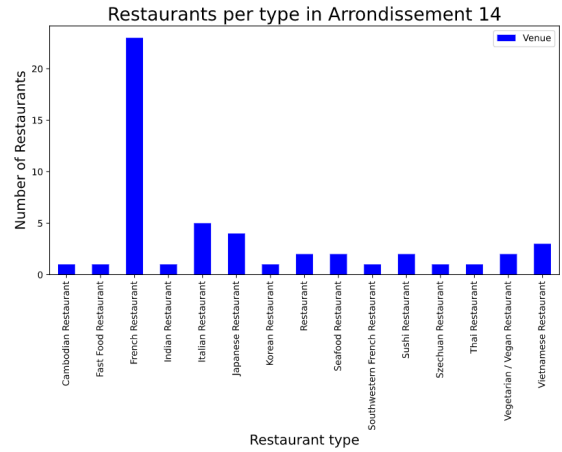
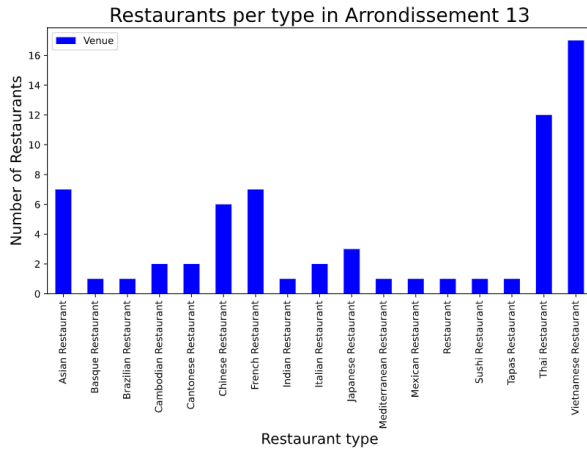
### Number of restaurants per type in each arrondissement:

The most common type of restaurant is French in almost all arrondissement, except the 13th. Furthermore, in most districts more than 35% of the restaurants are French. Those who don't meet this rule are: 1st, 2nd, 3th, 10th, and 13th. The second most common type of restaurants is Japanese.

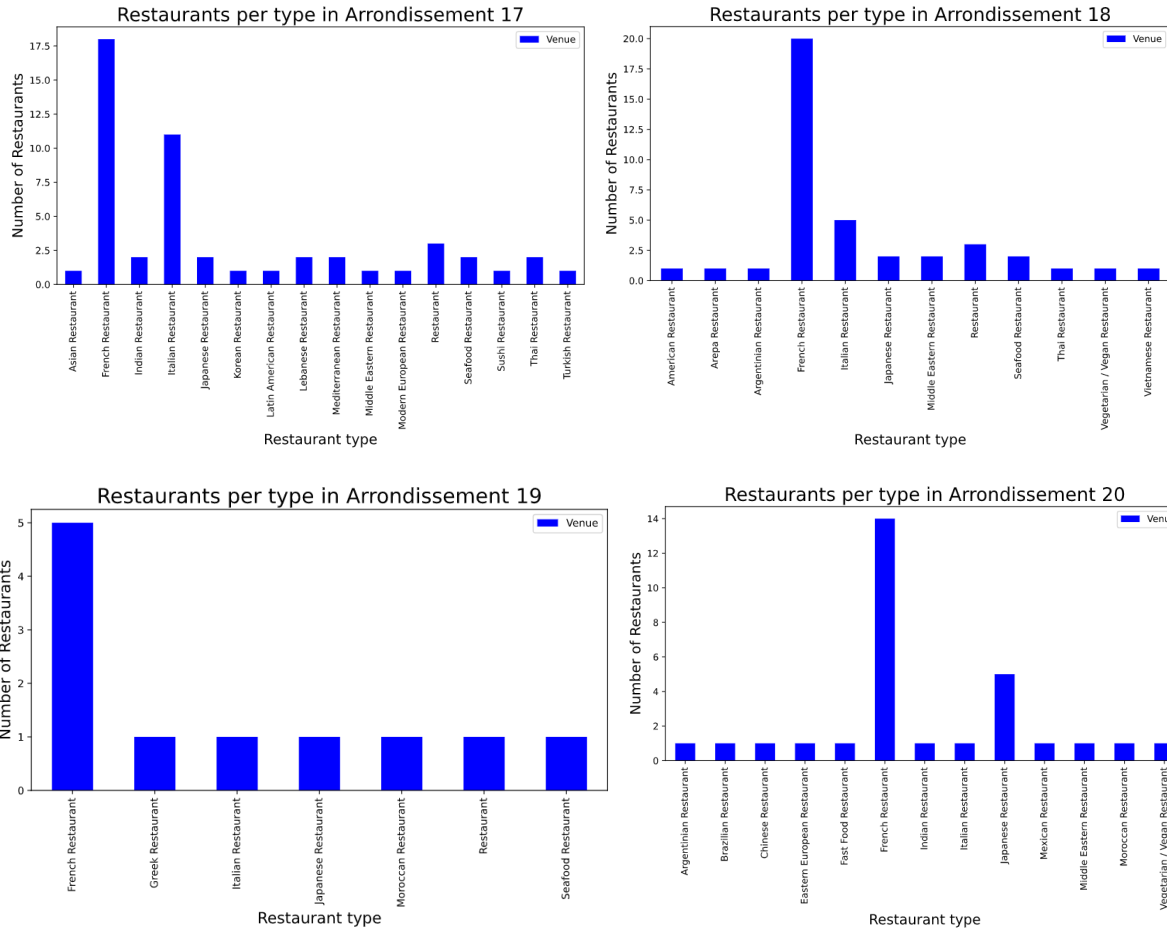
There are a lot of types of restaurants that are not very popular and only have one or two restaurants per district.







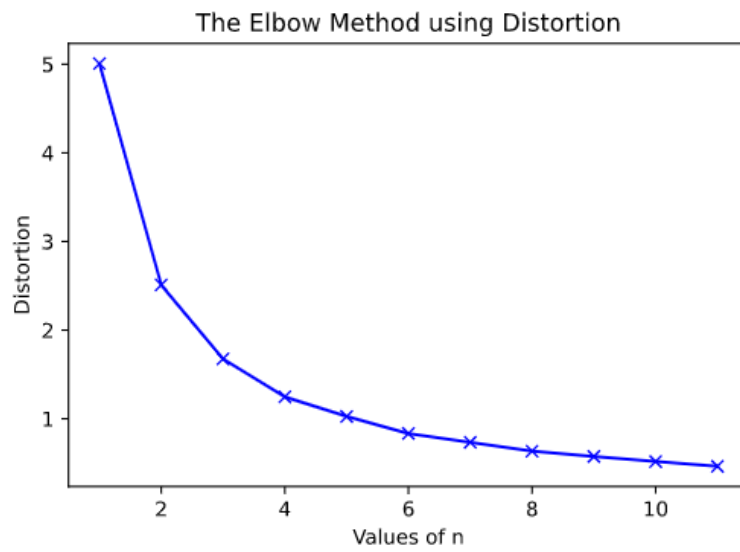




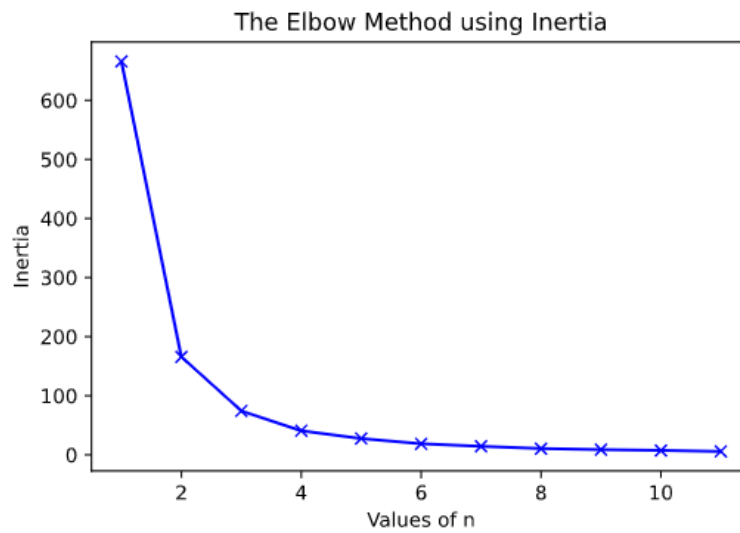
graph 3: restaurants per type per arrondissement.

## The elbow method

The results of the elbow method realized to determine the number of partitions for the clustering is shown below. The number of partitions selected is 3, because in the graphs 4 and 5 is shown that the elbow might be 2 or 3, and to get more information I selected 3 as the n.



Graph 4: The elbow method using Distortion metric for n between 1 and 11.



Graph 4: The elbow method using Inertia metric for n between 1 and 11.

---

## Arrondissements clustered

Finally, the result of the clustering are 3 groups:

- Cluster 0 - purple (1st, 2nd, 3rd, 5th, and 10th): This cluster is in the center of Paris. The first and second most common restaurants here are French and Japanese. Except in the 3rd Arrondissement, where the Japanese restaurants are not so common. Another characteristic is that the 2nd, 3rd and 4th Arrondissements have Italian Restaurant as the third or fourth most common restaurant. Also, this Arrondissement has less French participation (less than 40%) than cluster 2.
- Cluster 1 - blue (13th): It has only one Arrondissement, the 13th. This Arrondissement may be classified as Asian, because it has plenty of restaurants from this continent. Unlike the other clusters the French participation is smaller (near 10%)
- Cluster 2 - red (4th, 6th, 7th, 8th, 9th, 11th, 12th, 14th, 15th, 16th, 17th, 18th, 19th, and 20th): This is the bigger cluster. The first and second most common restaurants are French and Italian. Other common restaurants are Seafood, Thai, Vietnamese and Japanese.

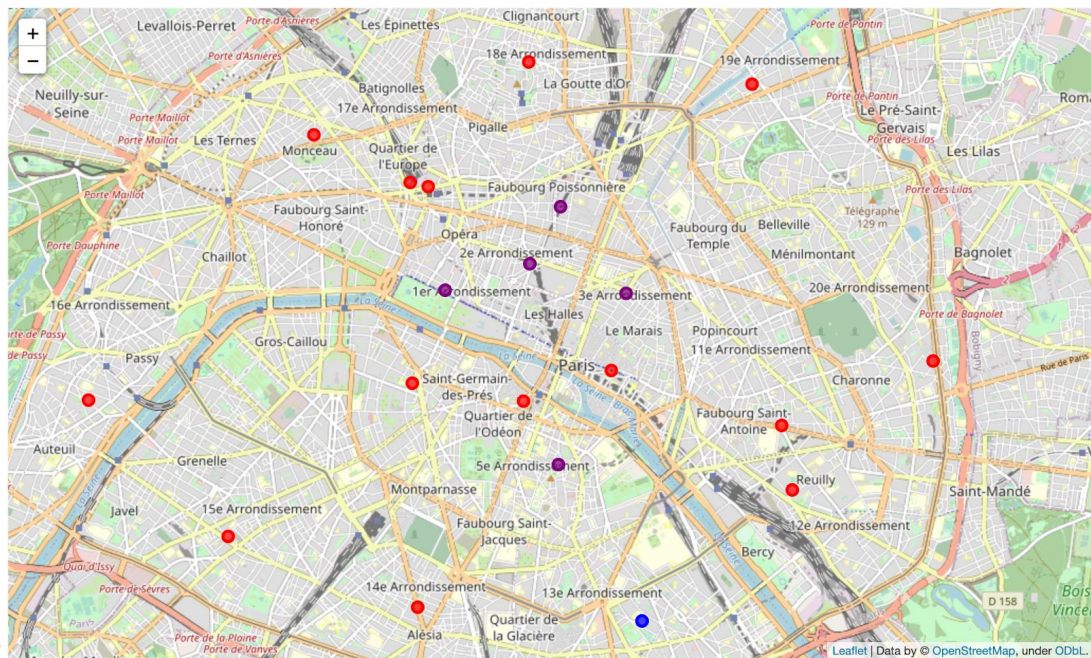


Image 1: Paris arrondissements' clustered in 3 groups.

Arrondissement	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	48.864671	2.330138	0	Japanese Restaurant	French Restaurant	Korean Restaurant	Udon Restaurant	Israeli Restaurant
2	48.867534	2.343925	0	French Restaurant	Japanese Restaurant	Chinese Restaurant	Italian Restaurant	Restaurant
3	48.864425	2.359444	0	French Restaurant	Restaurant	Italian Restaurant	Vietnamese Restaurant	Seafood Restaurant
5	48.846095	2.348502	0	French Restaurant	Japanese Restaurant	Greek Restaurant	Italian Restaurant	Lebanese Restaurant
10	48.873690	2.348787	0	French Restaurant	Japanese Restaurant	Restaurant	Vegetarian / Vegan Restaurant	Corsican Restaurant

Table 3: Cluster 0 most common restaurants.

Arrondissement	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
13	48.829351	2.362052	1	Vietnamese Restaurant	Thai Restaurant	Asian Restaurant	French Restaurant	Chinese Restaurant

Table 4: Cluster 1 most common restaurants.

Arrondissement	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
4	48.856137	2.357145	2	French Restaurant	Restaurant	Scandinavian Restaurant	Tapas Restaurant	Portuguese Restaurant
6	48.852813	2.342799	2	French Restaurant	Italian Restaurant	Seafood Restaurant	Lebanese Restaurant	Mexican Restaurant
7	48.854757	2.324789	2	French Restaurant	Italian Restaurant	American Restaurant	Japanese Restaurant	Peruvian Restaurant
8	48.876250	2.324441	2	French Restaurant	Italian Restaurant	Thai Restaurant	Vietnamese Restaurant	Lebanese Restaurant
9	48.875781	2.327347	2	French Restaurant	Italian Restaurant	Thai Restaurant	Vietnamese Restaurant	Seafood Restaurant
11	48.850298	2.384783	2	French Restaurant	Italian Restaurant	Restaurant	Vietnamese Restaurant	Korean Restaurant
12	48.843369	2.386362	2	French Restaurant	Italian Restaurant	Chinese Restaurant	Vietnamese Restaurant	Seafood Restaurant
14	48.830851	2.325746	2	French Restaurant	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Seafood Restaurant
15	48.838435	2.294916	2	French Restaurant	Italian Restaurant	Lebanese Restaurant	Thai Restaurant	Indian Restaurant
16	48.852952	2.272285	2	French Restaurant	Italian Restaurant	Japanese Restaurant	Thai Restaurant	Middle Eastern Restaurant
17	48.881367	2.308817	2	French Restaurant	Italian Restaurant	Restaurant	Seafood Restaurant	Mediterranean Restaurant
18	48.889073	2.343700	2	French Restaurant	Italian Restaurant	Restaurant	Seafood Restaurant	Japanese Restaurant
19	48.886722	2.379901	2	French Restaurant	Restaurant	Seafood Restaurant	Greek Restaurant	Moroccan Restaurant
20	48.857126	2.409257	2	French Restaurant	Japanese Restaurant	Eastern European Restaurant	Chinese Restaurant	Vegetarian / Vegan Restaurant

Table 5: Cluster 2 most common restaurants.

## Discussion

After the analysis, I would say that even though Paris is a multicultural and diverse city, the main gastronomic menu is French. If you like to eat French food you could go to any arrondissement, but in those belonging to Cluster 2 you will certainly find a French restaurant. On the other hand, if you like Asian cuisine you should go to 13th

---

arrondissement. Finally, if you like Japanese food you should go to the arrondissements belonging to Cluster 0.

## **Conclusion and Future Work**

In conclusion Paris could be divided in 3 according to the food that its restaurants offer. The first group for people who like Japanese and Italian food, the second group for people who like Asian food, and the last group for people who like French food.

Even though Paris is a multicultural and diverse city, the main gastronomic menu offered is French, so you could find French food in any arrondissement.

## **References**

[1] France tourism statistics 2020. <https://www.condorferries.co.uk/france-tourism-statistics>

[2] Arrondissements of Paris. [https://en.wikipedia.org/wiki/Arrondissements\\_of\\_Paris](https://en.wikipedia.org/wiki/Arrondissements_of_Paris)