IBM data science capstone peoject

COURSERA - 2020

Business problem & data

Businessman want to know the most suitable place to open a new restaurant in Paris.
Let's explore the capital of France to determine the best choice.
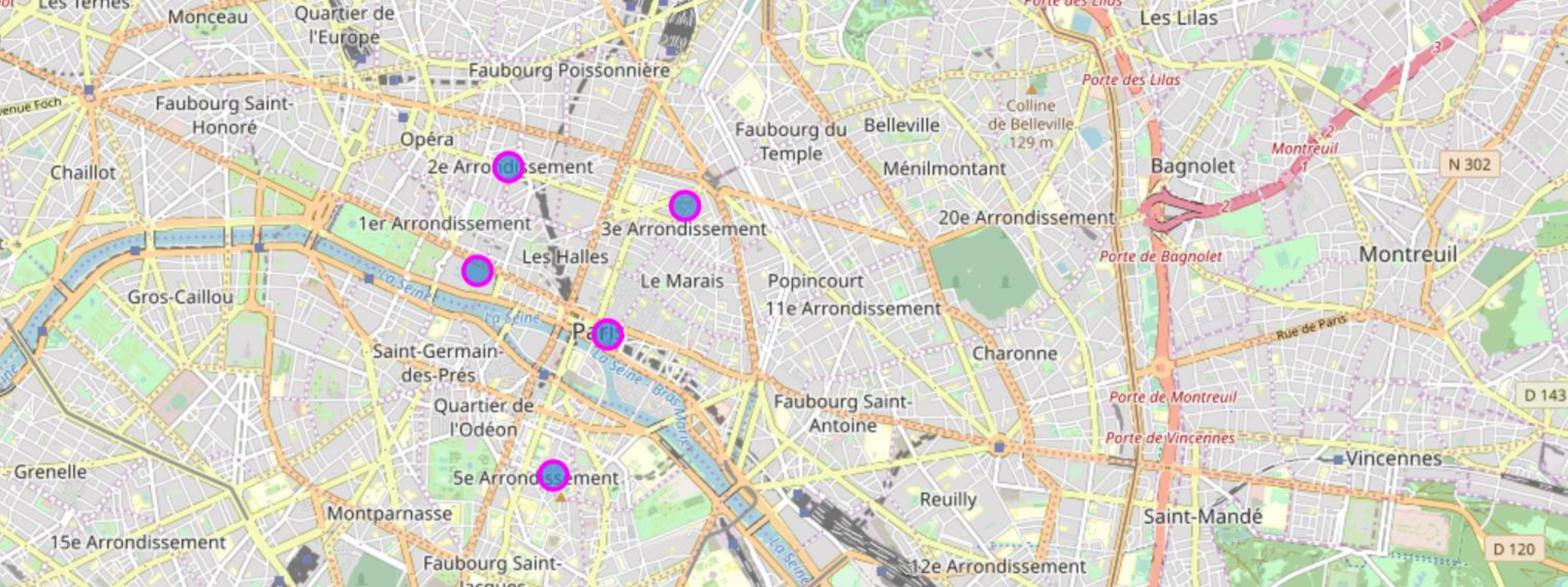
The data we will use is from:

1. Foursquare (venues, coordinate values…)

2. This wikipedia page: https://fr.wikipedia.org/wiki/Arrondissements_de_Paris (to get the list of Paris districts)

3. This webpage: https://www.journaldunet.com/patrimoine/prix-immobilier/paris/ville-75056 (to get the price of land in different districts of Paris)

	Nom	Superficie(ha)	Population	Densité1 (hab./km²)	Latitude	Longitude	Prix du m²
0	Louvre	183	16 545	9 041	48.861147	2.338028	10 000 €/m2
1	Bourse	99	20 796	21 006	48.868630	2.341474	9 750 €/m2
2	Temple	117	35 049	29 956	48.865900	2.361100	11 310 €/m2
3	Hôtel-de-Ville	160	27 146	16 966	48.856500	2.352400	11 380 €/m2
4	Panthéon	254	59 333	23 359	48.846200	2.346400	11 890 €/m2

Data

After scraping the webdata form the two webpages mentioned previously, we can manipulate the data to finally have the above dataframe.
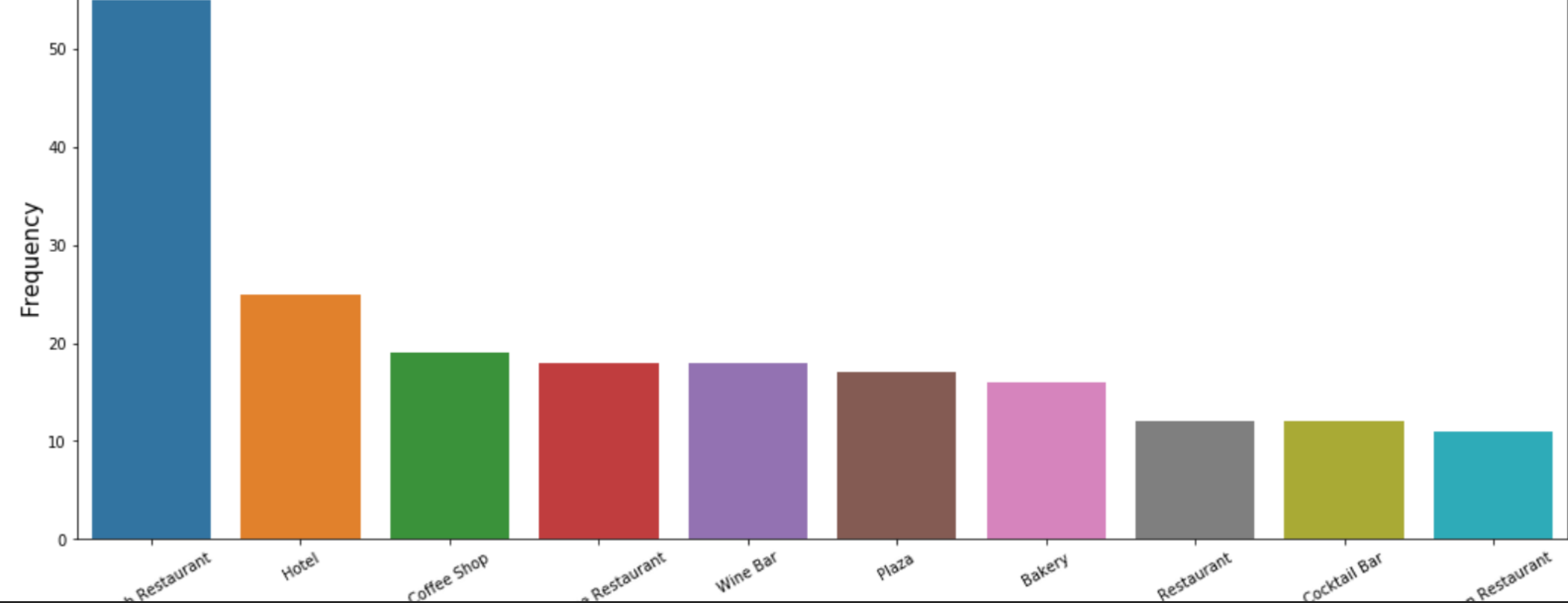
We will work on the first 5 districts of Paris.

Districts

We are seeing a picture of the five districts that we will explore:

Louvre, Bourse, Temple, Hôtel-de-Ville and Panthéon

10 Most Frequently Occuring Venues in 5 Major Districts of Paris

We can see that the most popular are fench restaurants and hotels

Restaurants

Here, each color represents the popular restaurants in a different district.

	Nom	Superficie(ha)	Population	Densité1 (hab./km²)	Latitude	Longitude	Prix du m²
0	Louvre	183	16 545	9 041	48.861147	2.338028	10 000 €/m2
1	Bourse	99	20 796	21 006	48.868630	2.341474	9 750 €/m2
2	Temple	117	35 049	29 956	48.865900	2.361100	11 310 €/m2
3	Hôtel-de-Ville	160	27 146	16 966	48.856500	2.352400	11 380 €/m2
4	Panthéon	254	59 333	23 359	48.846200	2.346400	11 890 €/m2

Data

After scraping the webdata form the two webpages mentioned previously, we can manipulate the data to finally have the above dataframe.

We will work on the first 5 districts of Paris.

Number of Restaurants as Top-Most Venues
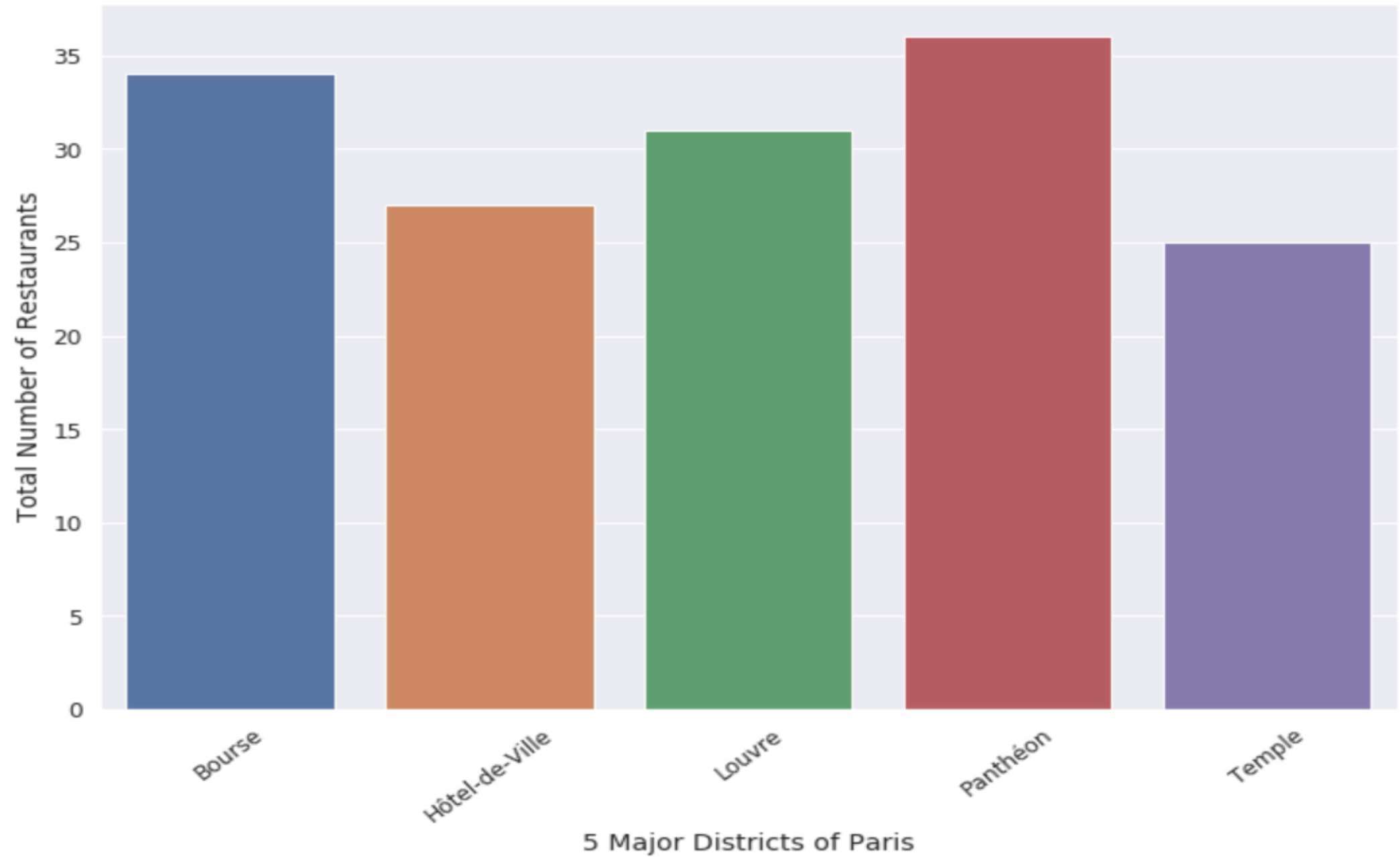
```
%%%%%%%%%Louvre%%%%%%%%
                 Venue  Freq
0     French Restaurant  0.10
1                 Hotel  0.09
2                 Plaza  0.07
3   Japanese Restaurant  0.06
4              Wine Bar  0.04
```

```
%%%%%%%%%Bourse%%%%%%%%
                 Venue  Freq
0     French Restaurant  0.11
1   Japanese Restaurant  0.07
2              Wine Bar  0.06
3                 Hotel  0.05
4                Bakery  0.04
```

```
%%%%%%%%%%Temple%%%%%%%%
                 Venue  Freq
0            Coffee Shop  0.07
1      French Restaurant  0.06
2                  Hotel  0.05
3               Wine Bar  0.04
4                 Bakery  0.04
```

```
%%%%%%%%%Panthéon%%%%%%%%
                 Venue  Freq
0     French Restaurant  0.16
1                 Hotel  0.06
2                 Plaza  0.04
3                Bakery  0.04
4                   Bar  0.04
```
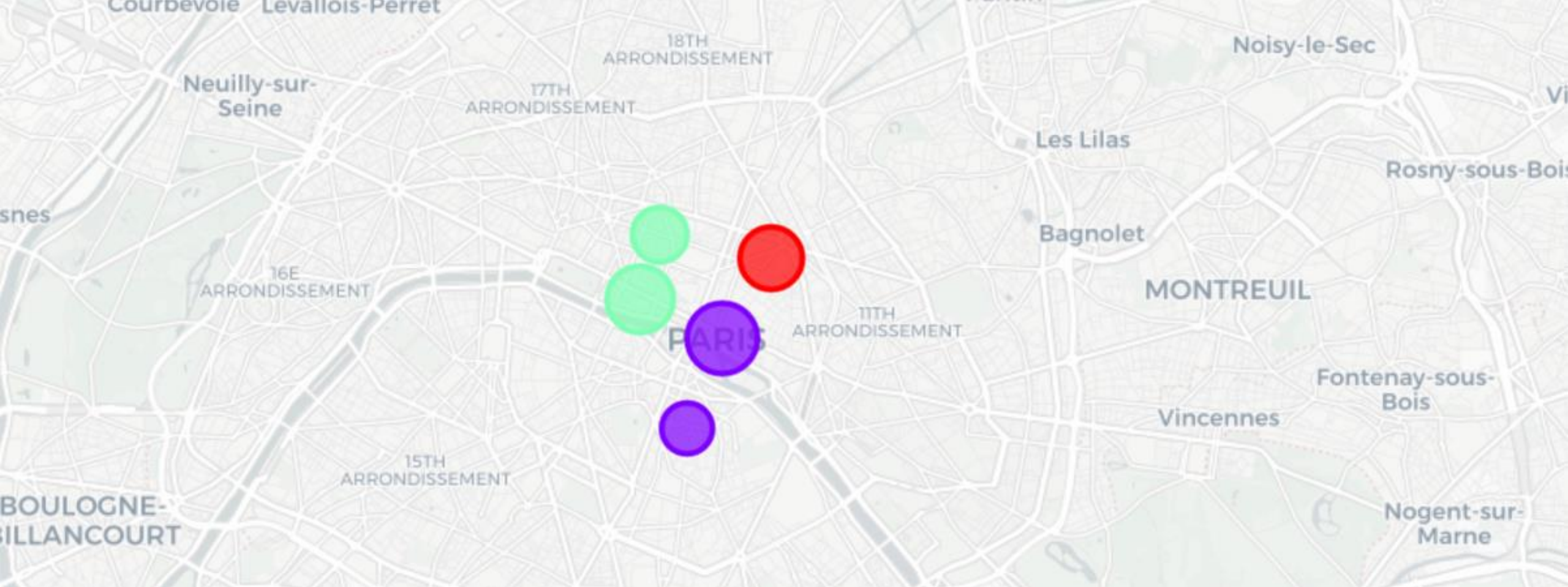
```
%%%%%%%%%Hôtel-de-Ville%%%%%%%%
                 Venue  Freq
0     French Restaurant  0.12
1        Ice Cream Shop  0.05
2                 Plaza  0.05
3           Art Gallery  0.04
4            Restaurant  0.03
```

TOP 5 VENUES FOR EACH DISTRICT

WE CAN NOW CLUSTER THESE DISTRICTS INTO 3 CLUSTERS:

The 3 Clusters

Let's now examine each cluster:

Cluster 1

	Densité1 (hab./km²)	Latitude	Longitude	Prix du m²	Cluster Label	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	9 041	48.861147	2.338028	10 000 €/m2	2	French Restaurant	Hotel	Plaza	Japanese Restaurant	Wine Bar	Bakery	Restaurant	Cocktail Bar	Coffee Shop	Italian Restaurant
1	21 006	48.868630	2.341474	9 750 €/m2	2	French Restaurant	Japanese Restaurant	Wine Bar	Hotel	Cocktail Bar	Bakery	Bookstore	Restaurant	Coffee Shop	Women's Store

In this cluster, the most popular venue is French restaurants.

This is not a good sign because there might too much competition for a newcomer.

Cluster 2

	Densité1 (hab./km²)	Latitude	Longitude	Prix du m²	Cluster Label	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
3	16 966	48.8565	2.3524	11 380 €/m2	1	French Restaurant	Plaza	Ice Cream Shop	Art Gallery	Historic Site	Coffee Shop	Burger Joint	Restaurant	Cosmetics Shop	Park
4	23 359	48.8462	2.3464	11 890 €/m2	1	French Restaurant	Hotel	Bar	Pub	Plaza	Bakery	Coffee Shop	Creperie	Indie Movie Theater	Ice Cream Shop

In this cluster, the most popular venue is French restaurants, too.

However, land price here is significantly higher than the cluster 1.

Cluster 3

	Densité1 (hab./km²)	Latitude	Longitude	Prix du m²	Cluster Label	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
2	29 956	48.8659	2.3611	11 310 €/m2	0	Coffee Shop	French Restaurant	Hotel	Italian Restaurant	Bakery	Wine Bar	Cocktail Bar	Restaurant	Sandwich Place	Breakfast Spot

In this cluster, the most popular venue is coffee shops.

This is a good sign because there might not be a lot of competition to compete with.

The land price is on the high side, however.

Conclusion

If the businessman is planning for a big investment (ie: a high end competitive restaurant) then "Bourse" may be the best option (good population too).

If the the businessman wants to play it safe, then he ought open the new restaurant in "Temple" because there is less competition.