**Neighbourhoods Prices and Venues Analysis in Paris**

# **Introduction**

## **Business Problem**

Paris is the capital and most populous city of France, with a population of 2,148,271 residents in an area of 105 square kilometres. It is one of the most attracting places in the world and also the city where I live for now. The city is divided into 20 boroughs and each borough has its own four neighbourhoods. Paris is a very interesting city and one of Europe's major centres of finance, diplomacy, commerce, fashion, science and arts.

Paris is a very beautiful city, but it can be a nightmare for people to find the best place to live. House Prices are very high and can increase until 12.000 € per m2. Also, as I said before, Paris is a big city divided into neighbourhoods with some similar and others very different. Finding a new place to live in Paris is really exciting but with all this information people can be lost. Consequently, the idea of this work is to help people finding the best place to live according to what they like for the best price.

For now, it is really painful for people to have access to all this information. With our work people have access to house prices according to the borough. Each neighbourhood is categorize depending on nearby places. Finally, they can visualize it very quickly on a map. For people who want to live in Paris, they save time and money.

## **Data**

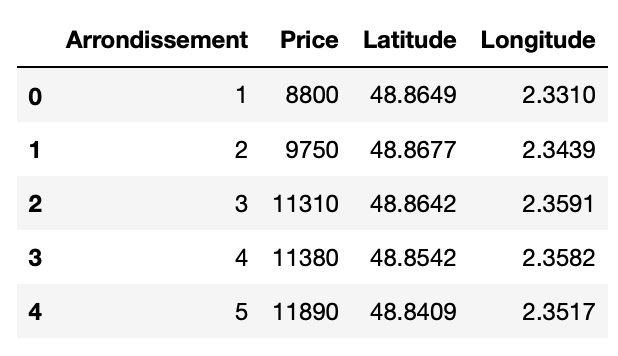
I used several sources of information to build this project:

* The French government provides area geographic data. Paris has twenty boroughs. I used the GeoJson file to delimit each borough and build the choropleth map of Paris [1]
* I created a database with all coordinates (latitude and longitude) and prices for each borough. I couldn’t find this information directly from a file on the web. For the coordinates, I used google maps by searching nearby place for each borough [2]. For the price, I used a french real estate web site which calculate the mean price for each borough on the last semester [3].
* I created a database with all coordinates (latitude and longitude) and prices for each neighbourhood. I couldn’t find this information directly from a file on the web. For the coordinates, I used google maps by searching nearby place for each neighbourhood. I used this information to build the choropleth map. This allow to have more information on the map and going deeper on the search for helping people finding the best neighbourhood [4].
* Finally, I used foursquare API to find social venues nearby each neighbourhood. For that, I make a request for each neighbourhood with the coordinate from the previous database I just created. This will help me to find the most common venue in each neighbourhood and categorize them.

## **Methodology**

All my project was implemented on a notebook where each step is well detailed. As I mentioned in the introduction, my work has two main goals. The first is to visualize on a map the prices for each borough. The second is to classify each neighbourhood according to the most common nearby venues.

The first step was to create a database with coordinates (center of the neighbourhood) and prices for each arrondissement. You can see the first five rows of this dataframe just below:



Then, we have to delimiter each borough on a map. For that, we downloaded a GeoJSON file published by the French Government containing all the coordinates information. We used the very powerful “Folium” library from python to edit the map. By combining this information, we can display each borough where the price is represented by a color from yellow to red as you can see on the legend.

Une image contenant texte, carte

Description générée automatiquement

The second step is classifying each neighbourhood according to the most common place. To analyse deeper, I decided to represent 4 neighbourhood per arrondissement. It allows to catch different types of social venues in the same borough. Just below you can find the first rows of the database I created with all this information.

Une image contenant capture d’écran

Description générée automatiquement

From this table, we can use the “CircleMarker” method from “Folium” library to represent each neighbourhood by a blue circle on the map. By clicking on the point, you can directly see his arrondissement and neighbourhood information.

Une image contenant texte, carte

Description générée automatiquement

Next, we have to find nearby venues for each point on the map. To do that, I use the foursquare API that returns venues on a specify radius area from the point. I choose a radius of 500 meters and a limit of 100 venues. Neighbourhoods are approximatively separate of 500 meters each other’s. The limit of 100 venues will be enough to have a global idea of the venues present in the neighbourhood. That is why I chose this parameter. I found 296 uniques categories for all the neighbourhoods.

## **Result**

We can then find the top 10 most common venue for each neighbourhood as you can see below:

**Une image contenant capture d’écran

Description générée automatiquement**

We have some common venue categories in boroughs. For this reason I used unsupervised learning K-means algorithm to cluster the boroughs. K-Means algorithm is one of the most common cluster method of unsupervised learning. I divided my neighbourhoods into 5 different clusters.

**Une image contenant capture d’écran

Description générée automatiquement**

In this table, you can see the cluster labels for each neighbourhood according to their nearby venues.

Finally, I added the cluster information on the map. Each cluster is represented by a different color. By clicking on the point you can see the neighbourhood and cluster information.

Une image contenant texte, carte

Description générée automatiquement

## **Discussion**

This map shows that Paris is a heterogeneous city. We can find very different neighbourhood on the same borough. Consequently, it is interesting for people who like their neighbourhood and have to move out for different reason (work, house prices …) to know that they can find similar neighbourhood in other boroughs.

As expected, house prices are more expensive in the centre and lower in the border. However, it is interesting to notice that we can find similar neighbourhood in the centre and in the border. For people actually leaving in the centre and who wants to reduce the houses prices while keeping the same social venues it could be an interesting alternative. An improvement would be to add the suburb of Paris and analyse how the neighbourhoods are classified.

## **Conclusion**

The goal of this study was to help people finding the best place to live according their venues preferences for the best price. For that, we built five different cluster according to the most common venue in each neighbourhood. Each cluster is represented on the map with a different colour. Therefore, someone who want to move out just have to find the cluster he is interested in and look at each point of the same colour on the map.

**Appendix**

[1] [Borough Paris](https://www.data.gouv.fr/fr/datasets/arrondissements-1/)

[2] [Google Maps](https://www.google.fr/maps)

[3] [Real Estate Price in Paris](https://www.journaldunet.com/patrimoine/prix-immobilier/paris/ville-75056)

[4] [Foursquare API](https://developer.foursquare.com/)