Best location for a car rental agency in Paris

Table of content

[**1.** **Introduction** 2](#_Toc48674267)

[**2.** **Data descritption** 2](#_Toc48674268)

[**3.** **Methodology** 2](#_Toc48674269)

[**4.** **Results** 5](#_Toc48674270)

[**5.** **Discussion** 6](#_Toc48674271)

[**6.** **Conclusion** 6](#_Toc48674272)

[**7.** **References** 7](#_Toc48674273)

1. **Introduction**
   1. **Description of the problem and a discussion of the background**

Paris is the most densely populated capital city in Europe and the fourth in the world (21,498/km² - 53,000/sq mi). As a consequence the real estate prices are amongst the highest in the world ( 8th in the world with 14,017.63 $ per square meter).

Paris being a city where relatively few people possess a car (due to excessive taxes and parking fees). So a car rental agency can be a juicy business because a lot of residents may need occasionally a car

For this project let's put ourselves in the shoes of an entrepreneur looking to open a car rental agency. So for the location we have to find a balance between the real estate prices and the presence of competitors.

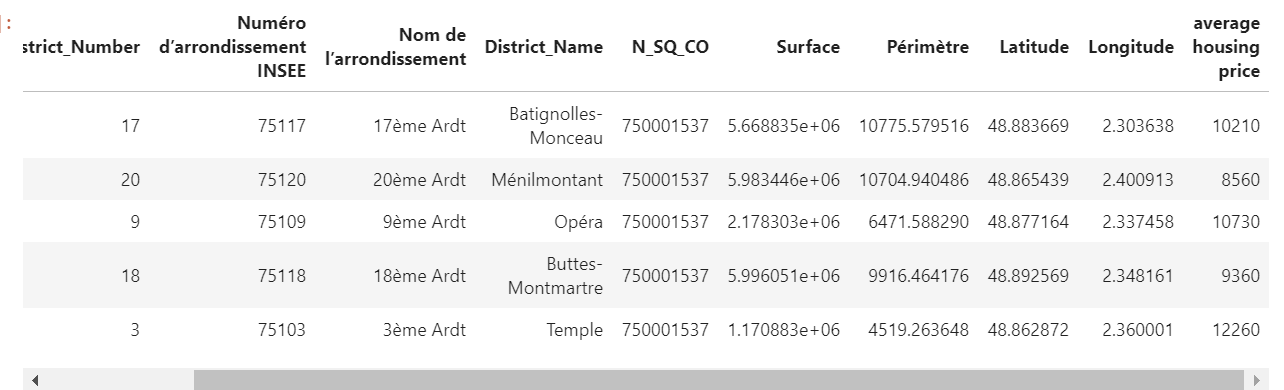
1. **Data descritption**

To solve this problem we will use the following data :

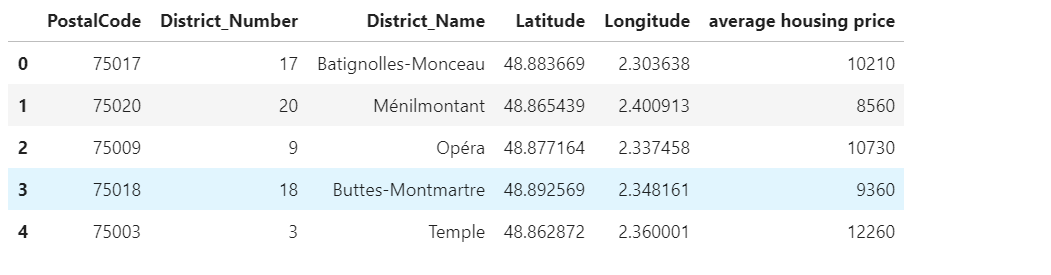
* An excel file (Arrondissements\_Paris.xlsx) downloaded from "Open platform for French public data" containing the list of all the Paris districts with their coordinates. [1]
* Foursquare API to get the car rental agencies for each district [2]
* The real estate prices for all Paris districts [3]

1. **Methodology**
   1. **Loading and exploring dataset**

I gathered for all the districts the coordinates and the housing prices in one file "Arrondissements\_Paris.xlsx" :

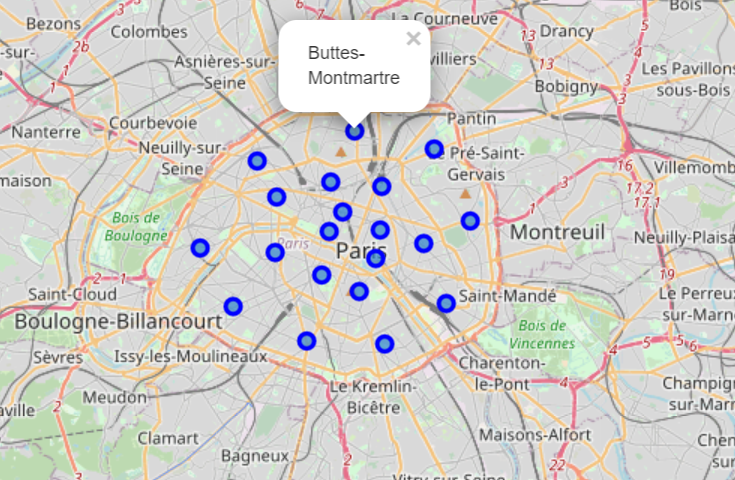


We will keep only the following columns : PostalCode,District\_Number,District\_Name and the coordinates :



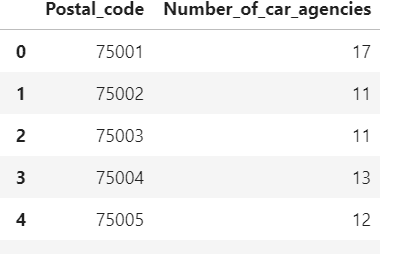
* 1. **Paris districts visualization**

Visualization of the Paris 20 districts using a folium map :

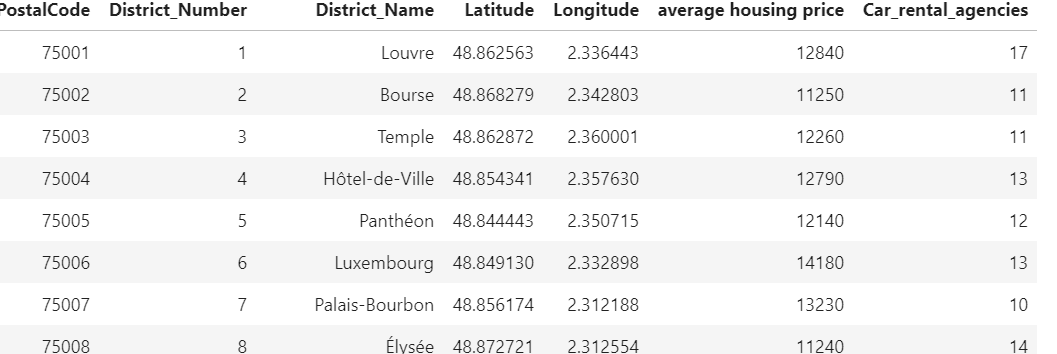


* 1. **Retrieving car rental agencies coordinates using the Foursquare API :**

After getting all the needed data from Foursquare we created a data frame with the number of car rental agencies per district :



Then we merge it with the data frame containing all the districts coordinates and the housing prices :



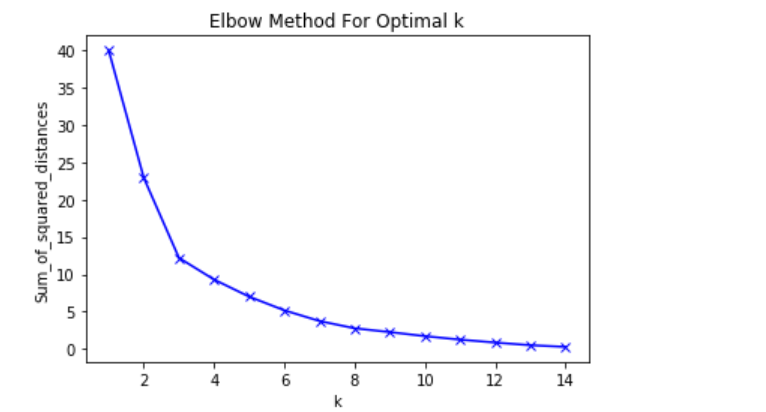
The total number of car agencies in Paris is 229.

* 1. **Clustering Dsitricts with Kmeans method**

First we extracted the relevant data ( the column "Car\_rental\_agencies" which contains the number of car rental agencies for each district and the column “average housing price”).

After we normalized the data, we used unsupervised learning K-means algorithm to cluster the districts because the data is unlabeled.

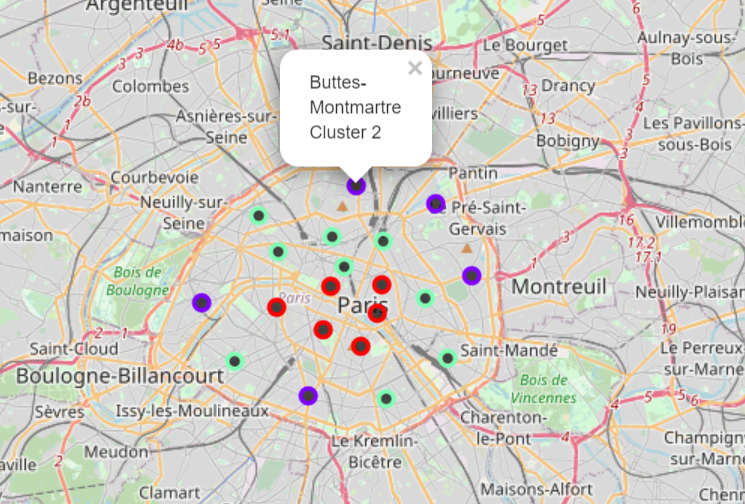
We determined the optimal k using the elbow method :



In the plot above the elbow is at k=3 indicating the optimal k for this dataset is 3

1. **Results**
   1. **Visualization of the resulting clusters**

To visualize the clusters we used a Folium map :

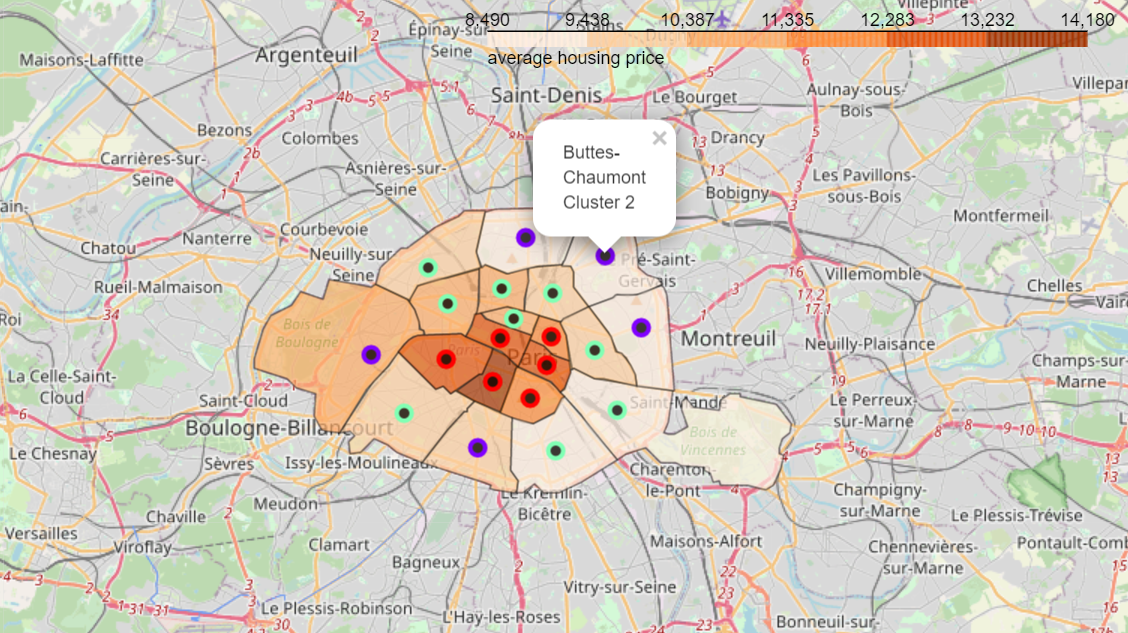


We can label the clusters as follows :

* Cluster 0 (green markers) : High number of car rental agencies and low housing prices
* Cluster 1(red markers) : High number of car rental agencies and high housing prices
* Cluster 2(purple markers ) : low number of car rental agencies and low housing prices

We get a json file file 'arrondissements.geojson' with the paris district coordinates we created a choropleth map containing :

* The district name
* The cluster name
* The average housing price



1. **Discussion**

Our results show that there are relatively few car rental agencies in Paris (229) so our first intuition that Paris has relatively few car rental agencies has been confirmed.

It seems wise to chose a district from the cluster 0 with low density of car rental agencies and low housing prices such as "Buttes-Montmartre","Buttes-Chaumont","Ménilmontant","Passy" or "Observatoire" ( The purple markers in the above map).

Among these districts "Passy" and "Observatoire" would be a better choice since they draw an important business clientele.

These recommendations have to be considered only as a starting point, others factors have to be taken into account such as the road traffic and the availability of building lands in the concerned districts

1. **Conclusion**

The aim of this project was to identify the best location to open a car rental agency that is to say a district with few car rental agencies and low housing prices.

By using KMeans clustering method with data coming from Foursquare API we clustered Paris districts into 3 clusters:

* Cluster 0 : High number of car rental agencies and low housing prices
* Cluster 1 : High number of car rental agencies and high housing prices
* Cluster 2 : low number of car rental agencies and low housing prices

We highlighted that the cluster 0 was the better choice.

This cluster contains the following district : "Buttes-Montmartre","Buttes-Chaumont","Ménilmontant","Passy" or "Observatoire"

Besides of this recommendation the final decision must take into account some more factors such as the road traffic and the availability of building lands.

1. **References**

* [1] [Paris districts coordinates](https://www.data.gouv.fr/fr/datasets/arrondissements-1/)
* [2] [Foursquare API](https://developer.foursquare.com/)
* [3] [Average housing prices in Paris](https://droit-finances.commentcamarche.com/faq/7409-immobilier-a-paris-prix-au-m2-des-arrondissements)