Chartres

Classification of the Neighborhoods in Chartres.

Table of contents

Introductory section

This report is provided for Ursula Pezeu who will move to a small town called Chartres near Paris in next February. She wants to open a B&B in Chartres and wish to know where exactly is the best place to be in Chartres. Chartres is a town of about 40 000 people surrounded by a suburb of about 200 000 people. This city is well known fot its magnificient cathedral which attract hundreds of tourists every day. In order to decide where she wants to have her B&B Ursula wants to know where are the biggest number of venues in Chartres.

To answer this question we will segment Chartres in different Neighborhoods and classify those Neighboorhoods according to different caracteristics including the number of venues per Neighborhoods. The result will be in the form of various maps summarizing the characteristics of the Neighborhoods.

Data section

The data we will be using consists of:

1) The number of inhabitants, flats, houses, tenants, owners per Neighborhoods accessible on this page: https://immobilier-chartres.nestenn.com/quartiers-chartres

2) The geographical delimitation of the Neighborhoods accessible as a geoJson file on this page: https://public.opendatasoft.com/explore/dataset/contours-iris/export/?refine.nom_com=Chartres

3) The Venues per Neighborhoods accessible through foursquare

We want to classify the different Neighborhoods according to their precise characteristics. Therefore we will import all the data in a pandas DataFrame in order to get the venues, inhabitants, flats, houses, tenants and owners per Neighborhoods easily accessible.

Once we have this dataFrame available we will be able to easily slice through the dataFrame to produce different maps of Chartres and classify the Neighborhoods according to various characteristics in order for our client to construct a precise idea of the characteristics of the city.

The aim of this report is to allow our client to decide where it would be the best idea to create her Bed and Breakfast. We are convinced that maps created with folium and choropleth will be the perfect fit for our client's needs.

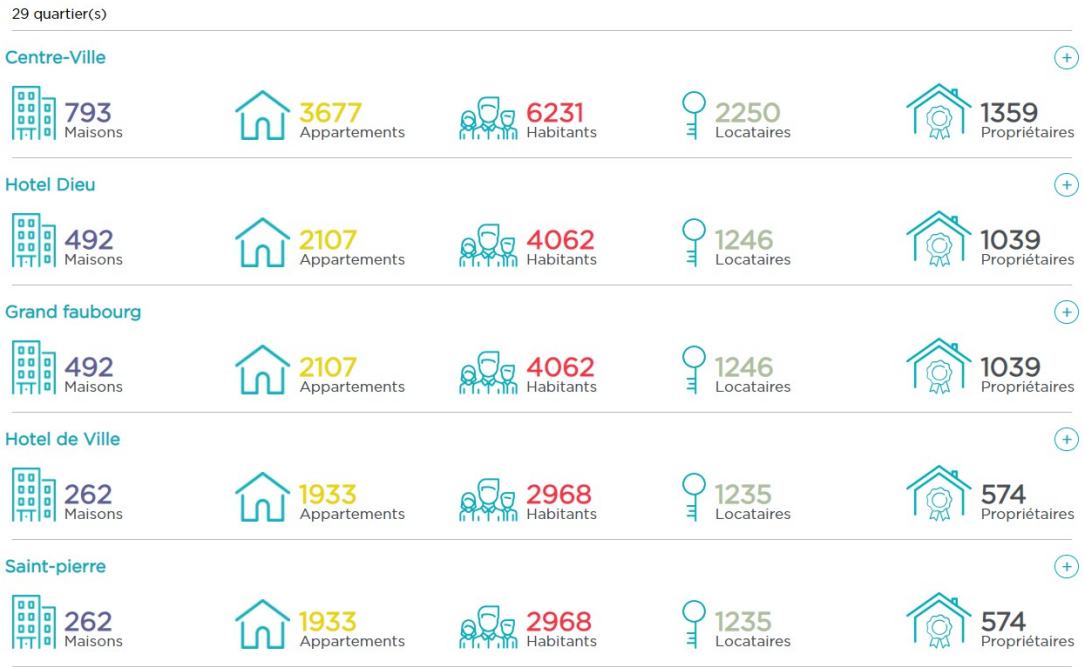
Methodology and Result Section

In the case of this report it was way more effective to fusion the Methodology and the result section since one depends directly on the other and it is better understandable to be able to read the results in the same time as the methodology.

Firstable we discussed extensively with our client to define precisely her needs and the question she wanted to answer. We decided together that we should try to get as much informations about each Neighborhoods as possible and sequentially come back to her to check if the informations we extracted are valuable for her or not.

1) Exploratory Data Analysis

The first step in the search was to define which are the Neighborhoods in the city. Therefore we realized a search with the keywords « Chartres » and « Quartiers » which is the french Translation of the word Neighborhoods. This search led us to https://immobilier-chartres.nestenn.com/quartiers-chartres which is a summary of the number of inhabitants, flats, houses, tenants and owners in Chartres.



These informations are valuable and therefore we scraped the page using Beautiful Soup. To finaly have a dataset recollecting these information :

	Quartier	nbMaison	nbHabitants	nbLocataires	nbProprietaires
0	Centre-Ville	3677	6231	2250	1359
1	Hotel Dieu	2107	4062	1246	1039
2	Grand faubourg	2107	4062	1246	1039
3	Hotel de Ville	1933	2968	1235	574
4	Saint-pierre	1933	2968	1235	574
5	Cathedrale	1426	2131	737	511

Figure 1: Number of houses flats, tenants, owners and inhabitants per Neighborhood

2) Localisations

Using geopy we could get the longitude and latitude of each Neighborhood except 5 which were mistakes by the geopy program. We decided to exclude those 5 Neighborhoods and obtained the folowing dataFrame :

	Quartier	nbMaison	nbHabitants	nbLocataires	nbProprietaires	latitude	longitude
0	Centre-Ville	3677	6231	2250	1359	48.042346	-1.706558
1	Hotel Dieu	2107	4062	1246	1039	48.440960	1.480725
2	Grand faubourg	2107	4062	1246	1039	48.443634	1.482033
3	Hotel de Ville	1933	2968	1235	574	48.443801	1.489084
4	Saint-pierre	1933	2968	1235	574	48.448406	1.480907
5	Cathedrale	1426	2131	737	511	48.447812	1.487910

Figure 2: Number of houses flats, tenants, owners and inhabitants per Neighborhood with their coordinates

With these informations we can already display the different Neighborhoods on a folium map.

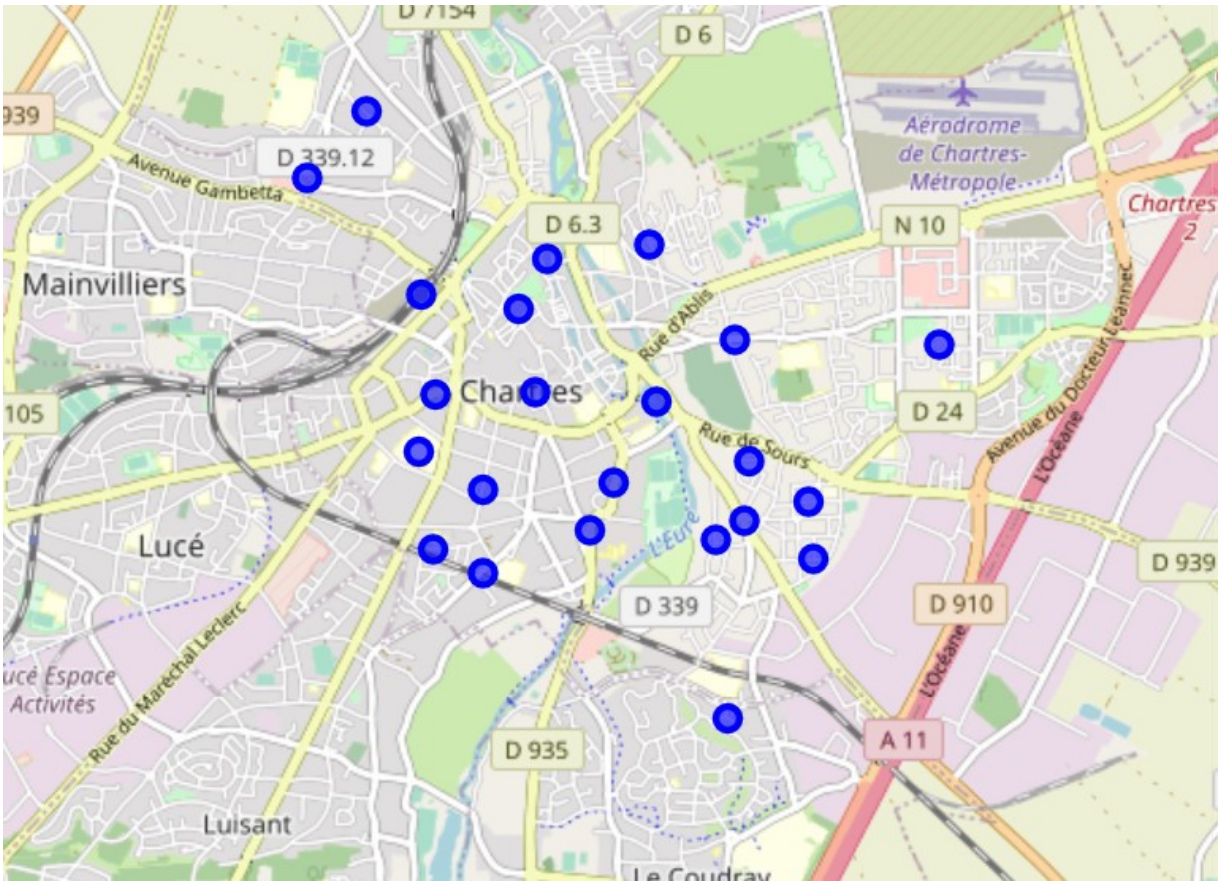


Figure 3: The location of Neighborhoods in Chartres

3) Grouping

With Foursquare we can obtain all the venues classified by Neighborhood and count the number of venues by Neighborhood :

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Centre-Ville	48.042346	-1.706558	Ozan Kebab	48.041819	-1.706296	Kebab Restaurant
1	Centre-Ville	48.042346	-1.706558	optique de la marionnais	48.041942	-1.705520	Optical Shop
2	Centre-Ville	48.042346	-1.706558	Les Blés d'Or	48.041984	-1.702462	French Restaurant
3	Hotel Dieu	48.440960	1.480725	Best Western Le Grand Monarque Hotel Chartres	48.444390	1.482929	Hotel
4	Hotel Dieu	48.440960	1.480725	Théâtre de Chartres	48.442809	1.485891	Theater
5	Hotel Dieu	48.440960	1.480725	La Cour	48.444353	1.482920	Brasserie
6	Hotel Dieu	48.440960	1.480725	Place des Épars	48.444234	1.483841	Plaza
7	Hotel Dieu	48.440960	1.480725	Subway	48.444429	1.484396	Sandwich Place
8	Hotel Dieu	48.440960	1.480725	Point Fort Fichet B . S . C Concessionnaire Br...	48.441940	1.476158	Miscellaneous Shop

Figure 4: Venues per Neighberhood

Neighborhood	Count
Bas Bourgs	5
Bas Menus	1
Beaulieu	5
Cathedrale	28
Centre-Ville	3
Chanzy	4
Clos Brette	3
Comtesses	1
Coteau d'Aboville	4
Croix Bonnard	2
Grand faubourg	18

Figure 5: Number of venues per Neighborhood

In order to construct a choropleth map we need the GeoJson file of the Neighborhoods in Chartres : We found it on this page : https://public.opendatasoft.com/explore/dataset/contours-iris/export/?refine.nom_com=Chartres and were able to see the different Neighborhoods on a folium map:

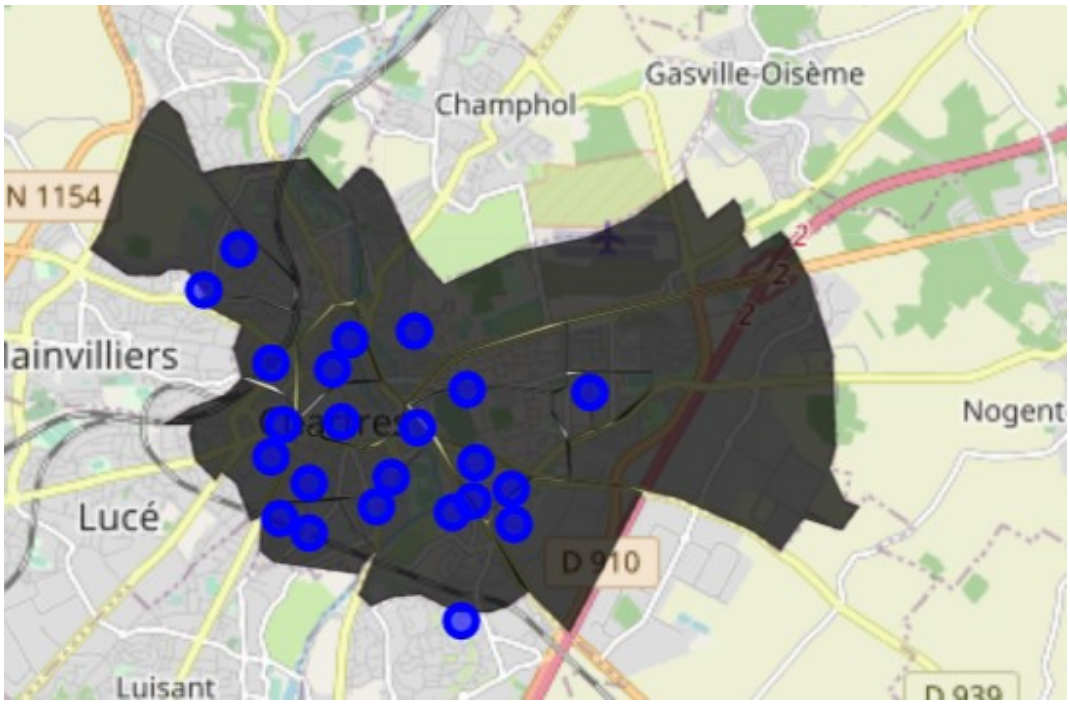


Figure 6: The geoJson representation of the Neighberhoods

4) Refining the data

Our problem was that the geoJson file did not separate the Neighborhoods in the exact same way than our data we scraped from the internet.

Therefore we had to refine our dataFrame in order to show the Neighborhoods which are in the geoJson file :

	Count	Quartier
Cathedrale Saint-André	40.0	Cathedrale Saint-André
Bas Menus Rechèvres	5.0	Bas Menus Rechèvres
Hotel de Ville Saint-pierre	37.0	Hotel de Ville Saint-pierre
Beaulieu Clos Brette	8.0	Beaulieu Clos Brette
Saint-Chéron	4.0	Saint-Chéron
Saint-jean Rechèvres Bourgneuf	4.0	Saint-jean Rechèvres Bourgneuf

Figure 7: The number of venues per melted Neighborhoods

With these informations we can provide already a map showing the number of venues per Neighborhoods in Chartres.

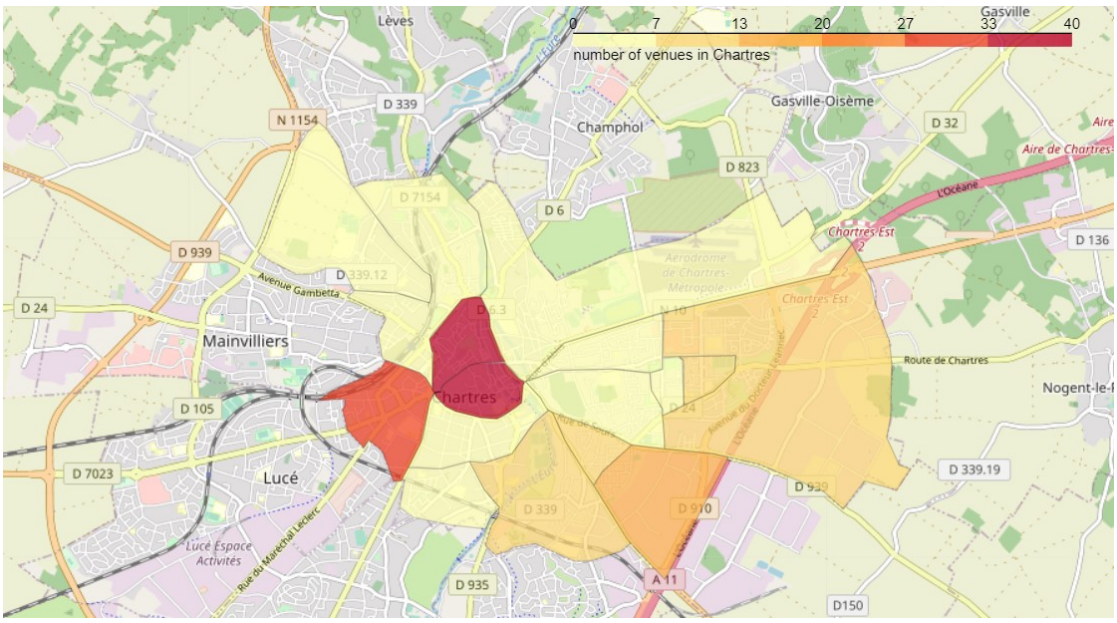


Figure 8: Map presenting the number of venues per melted Neighborhoods

5) Cost Analysis

An other important information that would direct the choice of the client is the price of the houses in the different Neighborhoods. Therefore we went on this website which gives the price of each Neighborhoods : https://www.seloger.com/prix-de-l-immo/vente/centre/eure-et-loir/chartres/280085.htm

Sadly we couldn't use BeautifulSoup on this page because of acces authorization and we had to take the information manually. After a while we could compute the information in the form of a map.

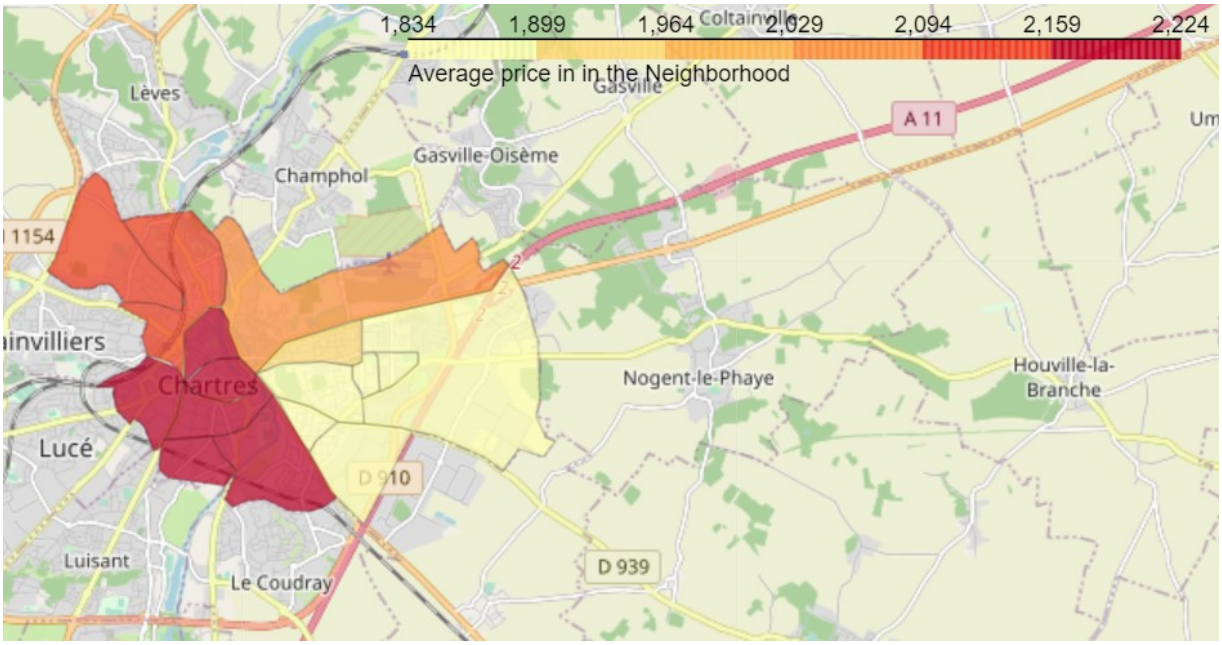


Figure 9: The average price of houses in Neighberhoods

With these pieces of information we could make a first advice for the place to open the B&B in Chartres in order to be in a cheap area and close to the places where a lot is happening.

6) Clustering

Finally we want to classify the different Neighborhoods according to their 10 most frequent venues. This helps us understanding the city better and let us decide which Neighborhood could be more attractive than others.

We can do that using our already collected informations with a kmean algorithm: this provides us with the following map where each color represents a cluster of similar neighborhoods:

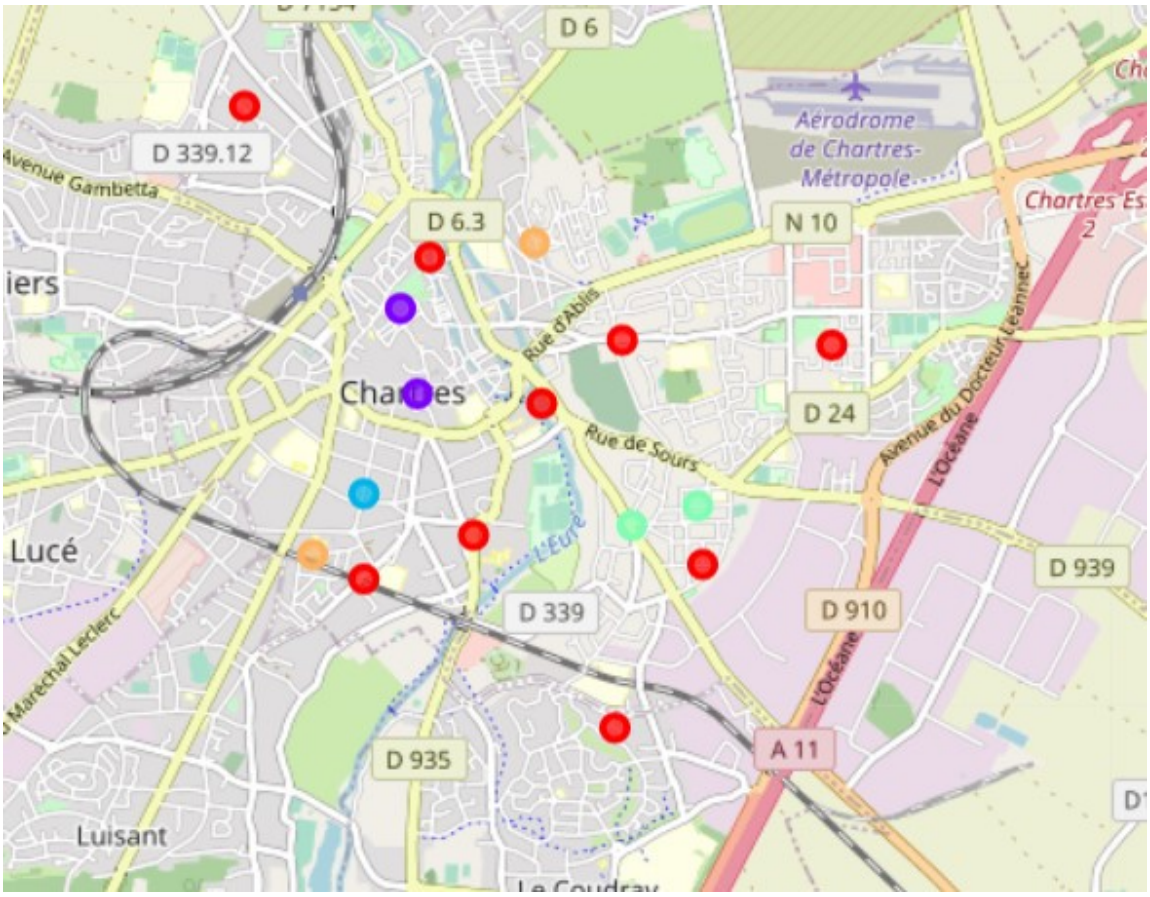


Figure 10: A kmean classification of the neighberhoods

Discussion

The results provided an insights of the popularity and the prices of the different Neighborhoods. As one could expect there is a relationship between the price and popularity of the places.

Nevertheless this point is not a major issue since our client wants to be in a 10 minutes by car area around the interesting places. This area offers a lot of possibilitys.

We highlighted numerous features of interests in Chartres which could help our client get a better idea of the project she is into. Thanks to this analysis we could make a significant difference and that was very valuable.

Some further discussions remains necessary with our client to decide wheter or not the provided information is sufficient and whether or not she can safely step into the realisation of her project. Of course this project is subject to randomness since she needs to find a place that is available and that there will be other criterias to take into account like the surface of the place, the quality of the house, the accessibility by car or other transportation.

Thanks to our analysis and according to the information we collected the best place which combines low prices with high interest seems to be Beaulieu and the surroundings.

Conclusion

In this topic we started by clarifying the needs of our clients and made clear the road we were to follow. We used some data scientist tools to scrape the internet in the search for specific data helping us solving the problem of our client.

We managed to put together informations from various places and to organize those informations into pandas to create highly visual information with the help of folium and choropleth. All the information collected has not been yet translated into maps and could be the focus of further works on this topic.

After collecting, organizing and presenting the data in the simpler possible form we could produce a recommandation to our client according to the question that had been asked. We of course remain accessible for any further work or details about this topic.

References :

Coursera Platform : IBM Data Science Professionnal Certificate