

# How to find the best place to live (France)

Oct 19.

## 1. Introduction – Problem - Interest

Hello, my name is Julien. I'm living in Paris and I will probably move to Lyon in the next years. Therefore, I need to create a tool to choose carefully where I want to live (because Lyon is the 2nd biggest city in France, and there are many cities/neighborhoods around Lyon).

I will then create a tool for decision making. The final customer will be me in this case, but we can standardize the tool in order to help other people.

I will focus my studied location to Chassieu city (and I will suppose that my work is located there). It's a small city in the Est suburb of Lyon.

Interest: A real estate company would love my tool to help people choose where they ideally can live! That will help people to imagine their life in this new house and to buy it!

## 2. Data sources

1. First, I will use an online file with information about cities in France: City Code, longitude, latitude, Name of the city, etc...

With it, I will be able to refine this list to cities which are less than 25km away than Chassieu.

Link: <https://public.opendatasoft.com/explore/dataset/code-insee-postaux-geoflar/export/?flg=fr>

2. I will use google API to get the time of travel between cities, therefore I created a google developer account and use Matrix distance API (which can give the time in car between 2 locations).

I will refine my study to cities which are 25 min away from Chassieu (in car).

3. I will use a geojson file of cities in France to create a choropleth map and then plot cities with color regarding the time of travel from Chassieu.

Link: <https://github.com/gregoireddavid/france-geojson>

4. I will also use foursquare to get venues in these cities, in order to cluster (same analysis than NYC lab, but with some modifications to fit to my data).

I will then plot all data on a same map to make a choice.

I tried also to import real estate prices but unfortunately, all websites were protected (it is sensitive data).

About Data cleaning, I had to be careful on some points:

- File 1 (information about cities in France): Some postcodes were 4 digits instead of 5 digits. I needed to clean that because geojson file was 5 digits only.
- For Foursquare data, I realized France isn't a big fan of this application. So, the number of venues was quite low. I needed to remove cities with low venues (less than 5) to a special cluster.

### 3. Methodology section

The Methodology is made by 3 steps:

- Refining the perimeter around the targeted city by distance (25km)
- Then refining the perimeter around the targeted city by travel time in car (25km)
- We will explore foursquare venues on these targeted cities and cluster them using K-mean clustering. We will have to select the good number of cluster to obtain meaningful clusters.

### 4. Results

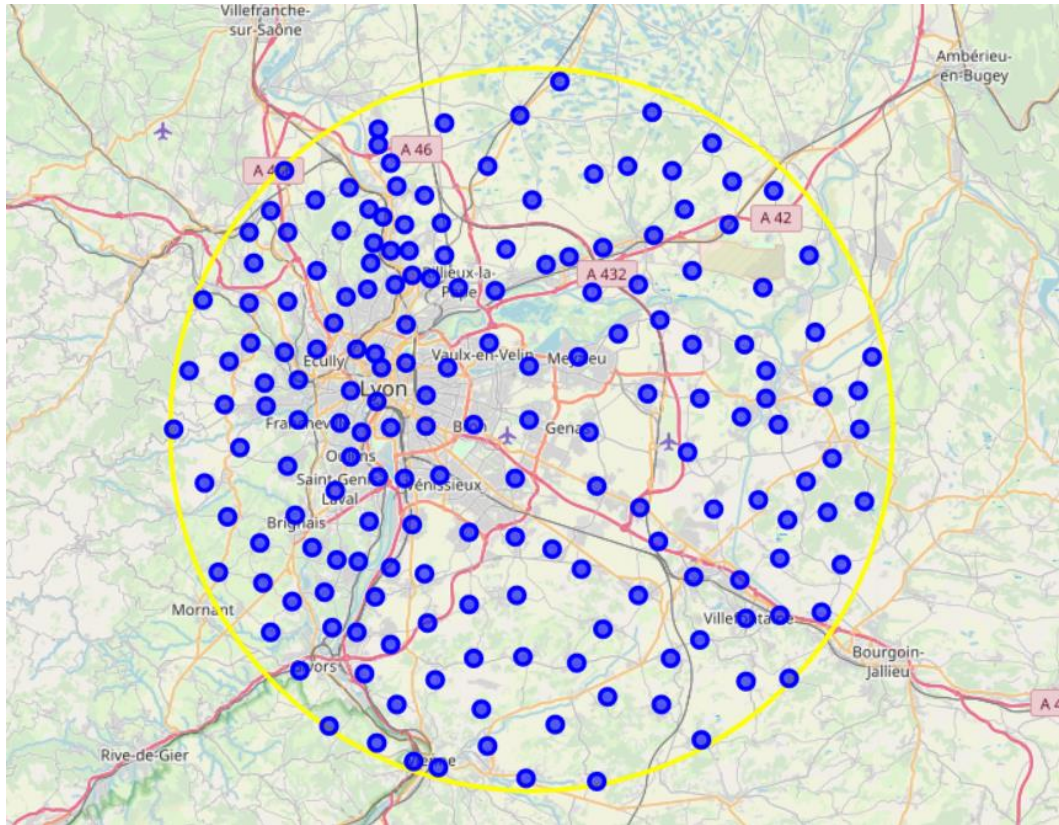
I selected a target city => the city where I will probably work.

In my case, it was Chassieu, a small city in the Est suburb of Lyon.

This city is close to 3 departments (= State in US). I filtered them on my data (full list of cities in France, 37000 cities) to have a smaller dataset.

Then, with geopy, I calculate the distance in km between Chassieu and remaining cities. I then selected only cities closer than 25km (see yellow circle).

The map below presents all the cities in a 25 km range from Chassieu:

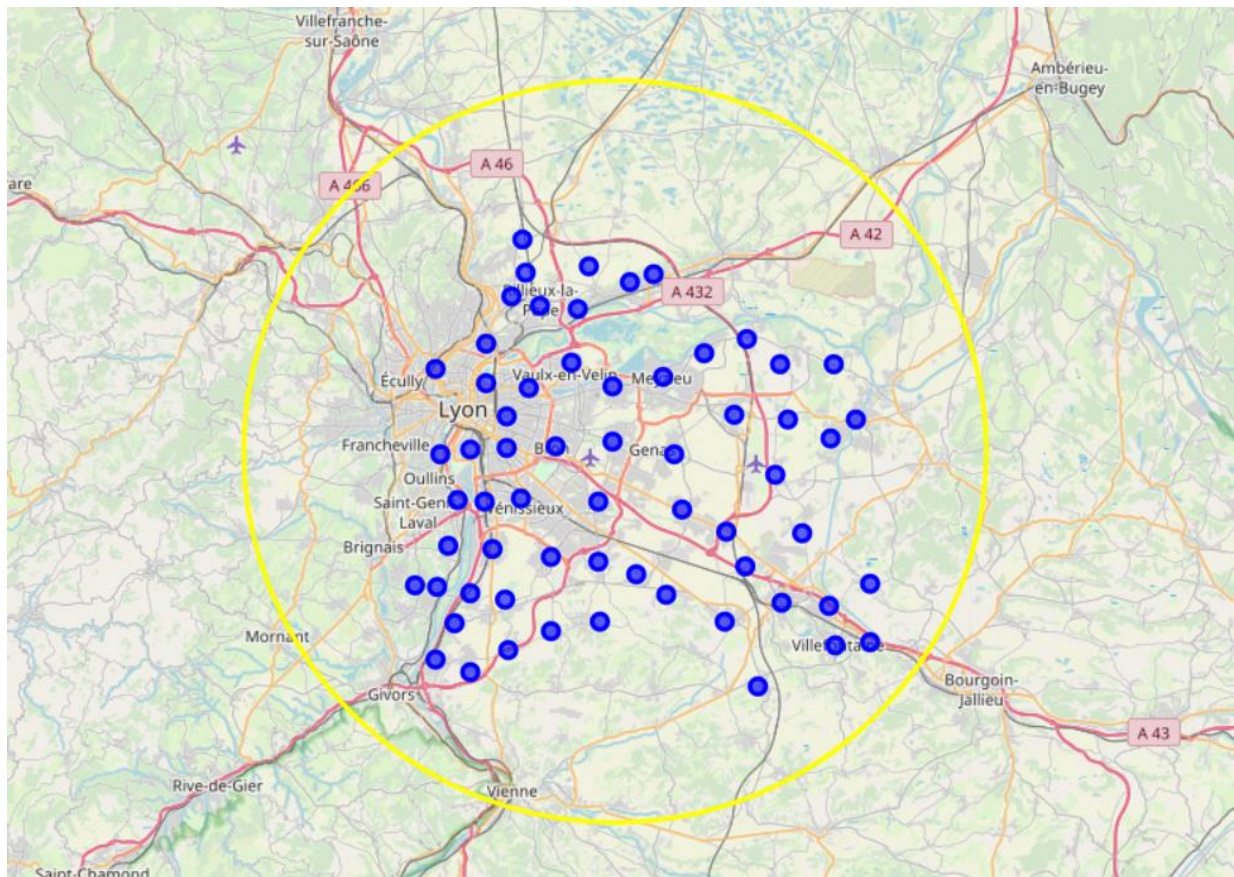


Map 1: Cities in a range of 25km from Chassieu

On these selected cities, I used Google API to calculate the time of road. For each city, I made an average between time to go to work, and time to get back to work. I then selected only cities which are at less than 25 min from the work.

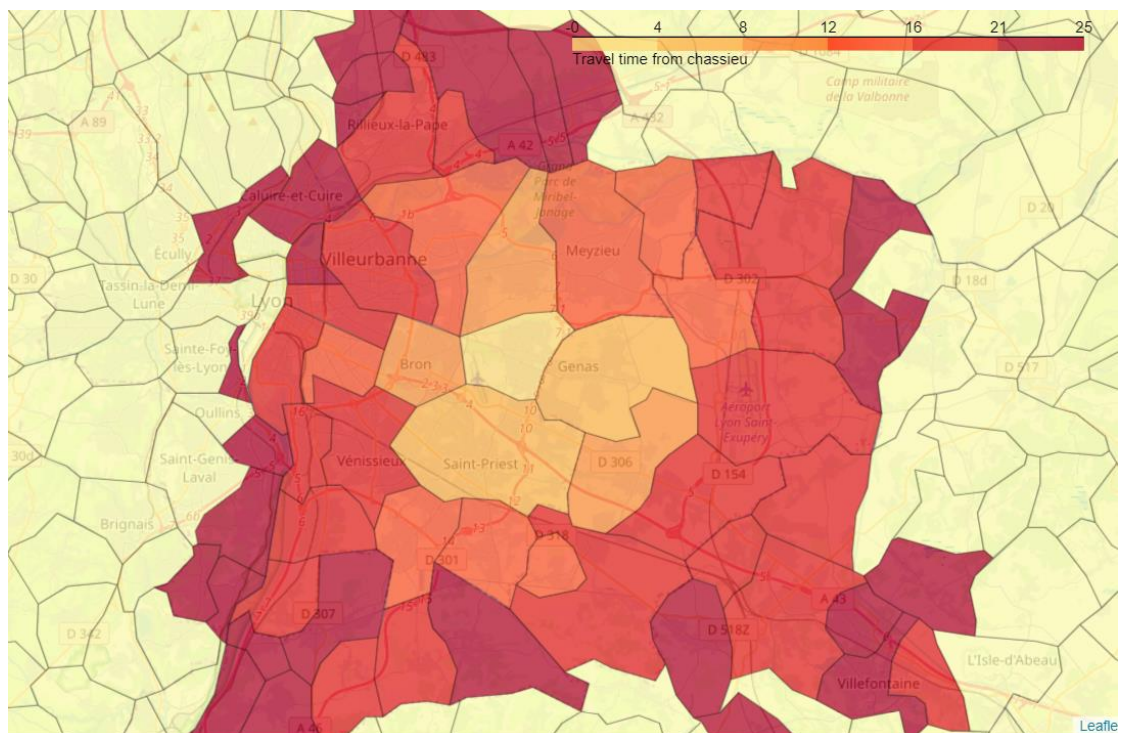
This is giving next map:





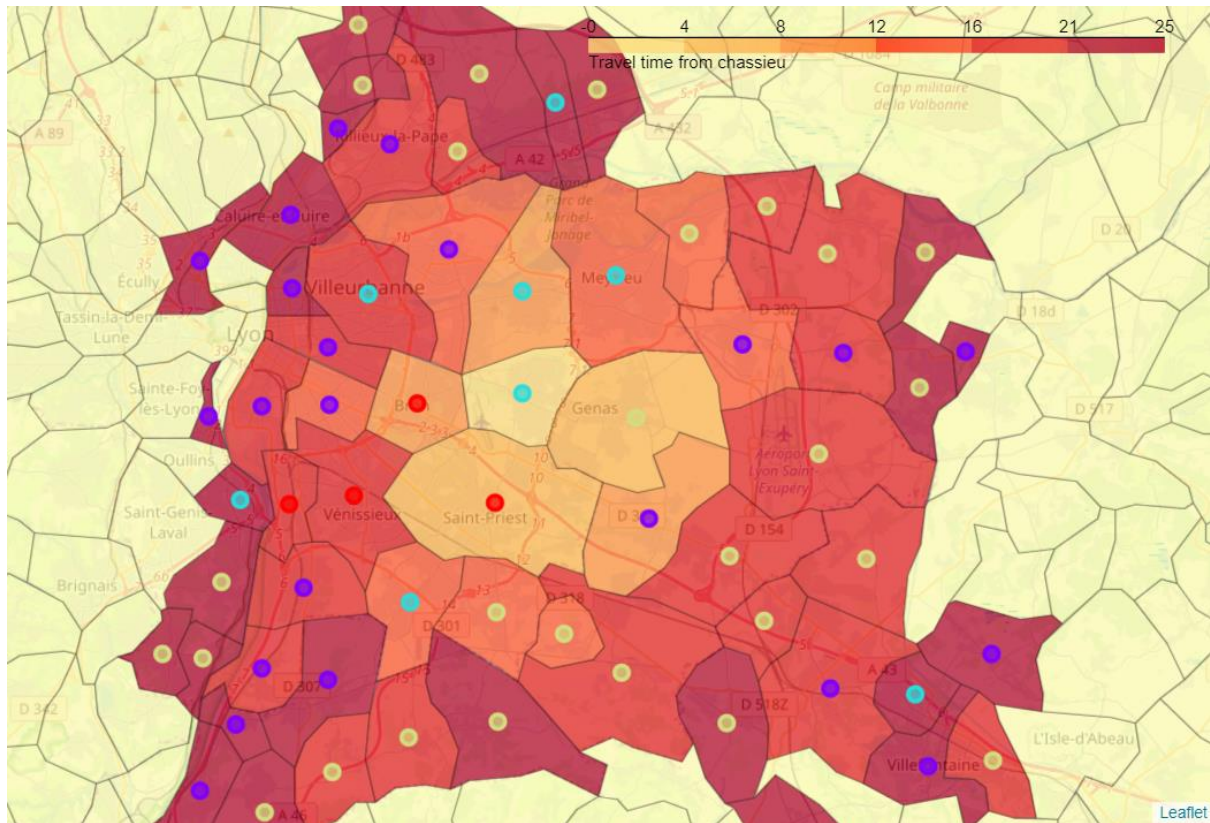
Map 2: Cities in a range of 25min from Chassieu

I then used geojson file to create a choropleth map:



Map 3: Cities in a range of 25min from Chassieu, colored by time from Chassieu

I finally used Foursquare to get venues and cluster them in 4 clusters:



Map 4: Map 3 with 4 clusters

## 5. Discussion section

We got 4 clusters:

Red cluster: Suburb with lot of public transports

- This cluster is composed of Tram station, fast food and supermarket / shop
- It's a suburb cluster with public transport as Tram.
- It would be nice to live. And it is also very close from Chassieu.

Light blue cluster: Suburb without public transport

- It's almost the same cluster than red cluster but there is less public transport.
- It is then a suburb cluster with less public transport. That would also be cool to live there. It's quite close to Chassieu.

### Purple cluster: City center

- This cluster is principally composed of Bar, Hotel, Business, Shop.
- It's clearly a cluster of a city center, in this case, Lyon's center.
- There are some cities of this cluster in noise in the country side (because small villages can have a bar and some shops).

### Yellow cluster: Villages / Countryside

- It's the cluster where there are less than 5 venues.
- This cluster is representing small villages. It could be cool to live there in order to have our own house with a lower price.
- But this is also further to our work and we won't have close activities and close shops.

We also got a map (Map 3) with the time of travel for each city in the region. This is a mandatory feature to choose where to leave.

I recommend selecting:

- Yellow cluster if you prefer countryside, having a house, and a small budget.
- Purple cluster if you love city, going to bar and cities activities.
- Light blue cluster or red cluster if you want to be close to the city center with a smaller price of real estate. Red is preferred if you don't have a car.

## 6. Conclusion

In conclusion, this analysis is giving us the possibilities of location to explore. I'm not really sure of which cluster we will choose.

On the first hand, countryside has big advantages, like having a house, for a lower price. But it's also a long travel to work... On the other, commodities from the 2 suburb clusters are appreciable. But it will be more expensive!

What is sure is that I will look at the choropleth map of travel time closely because I want to minimize travel time to work.

The next step of this analysis is exporting average price of flat/apartment for each cities in order to facilitate the choice. I tried to do that but this information is protected on many websites.

Also, Foursquare is not very popular in France, so results are not very accurate and meaningful. This analysis would have been better in US for example.

-