

Seafoods restaurant in Paris (Mohammed Bouayoun)



Figure 1: Paris

1. Introduction and Business Problem

1.1 Background

Paris is the capital and most populous city of France, with a population of 2,148,271 residents in an area of 105 square kilometers (41 square miles). Since the 17th century, Paris has been one of Europe's major centers of finance, diplomacy, commerce, fashion, science and the arts. The City of Paris is the center and seat of government of the Ile-de-France, or Paris Region, which has an estimated official 2020 population of 12,278,210, or about 18 percent of the population of France.

The Paris Region had a GDP of €709 billion (\$808 billion) in 2017. According to the Economist Intelligence Unit Worldwide Cost of Living Survey in 2018, Paris was the second most expensive city in the world, after Singapore, and ahead of Zürich, Hong Kong, Oslo and Geneva. Another source ranked Paris as most expensive, on a par with Singapore and Hong Kong, in 2018.

1.2 Business Problem



Figure 2: Seafood

Fresh fish is a highly perishable food product, so it must be eaten promptly or discarded; it can be kept for only a short time. In many countries, fresh fish are filleted and displayed for sale on a bed of crushed ice or refrigerated. Fresh fish is most commonly found near bodies of water, but the advent of refrigerated train and truck transportation has made fresh fish more widely available inland.

The closest sea to Paris is 194 km away, it's necessary to be close to the fish shops to get fresh fish for our new restaurants.

How I choose the right seafood location ?

What is a good location and how to identify the best business premises ?

Choosing the right business location is a key strategic issue. The success of the business depends on it.

The best locations are often already taken or extremely difficult to obtain. The amount of the lease fee can quickly skyrocket.

Identifying the right commercial premises starts with fieldwork: observation, location scouting and surveys of local residents or merchants.

Here is some list of essential criteria to consider when choosing a commercial location :

- What are the characteristics of the individuals frequenting the neighbourhood (residents, transients, employees, workers...) ? Age, sex, activity, purchasing power, intentions, behaviour ?
- Does attendance vary according to the days of the week, the time of day or the season ?
- What are the demographic trends in the neighbourhood ?
- The presence of competitors nearby is not a negative point, on the contrary it can be stimulating.

1.3 Interest

The objective of this project is to find the best locations to open a new seafood restaurant near fish shop to get fresh fish and seafood.

Paris is a city where there are various peoples in the world, many of them prefer fish to meat, such as Muslims, Jews and Asians.

2. Data

2.1 IRIS Data

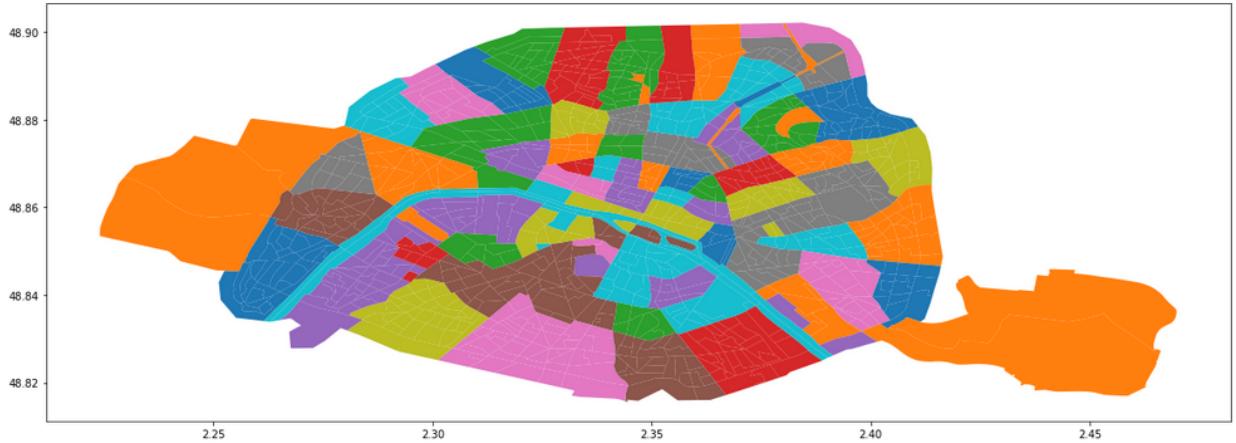


Figure 3: IRIS

In order to prepare for the dissemination of the 1999 population census, INSEE had developed a division of the territory into homogeneously sized grids called IRIS2000. This acronym stood for “Ilots Regrouped for Statistical Information” and referred to the target size of 2,000 inhabitants per elementary grid cell.

Since then, IRIS has been the basic building block for the dissemination of infra-municipal data. It must comply with geographical and demographic criteria and have unambiguously identifiable contours that are stable over time.

Municipalities with at least 10,000 inhabitants and a high proportion of municipalities with 5,000 to 10,000 inhabitants are divided into IRIS. This division constitutes a partition of their territory. France has approximately 16,100 IRIS, 650 of which are in the overseas departments.

By extension, in order to cover the whole territory, each of the communes not divided into IRIS is considered as one IRIS.

There are three types of IRIS:

- Housing IRIS: their population is generally between 1,800 and 5,000 inhabitants. They are homogeneous in terms of habitat type and their boundaries are based on the major cuts in the urban fabric (main roads, railways, waterways, etc.).
- Activity IRISs: they have more than 1,000 employees and have at least twice as many salaried jobs as the resident population.
- Miscellaneous IRIS: these are large specific areas with few inhabitants and a large surface area (leisure parks, port areas, forests, etc.).

We're not going to use all these variables, we're going to do a hard clean up and take only the variables we're interested in. The density of the population, foreigners and immigrants will help us find a location at our restaurant in addition to the Foursquare API to find our conquerors and fishmongers.

we can find this dataset in opendatasoft.com :

https://public.opendatasoft.com/explore/dataset/iris-demographie/download/?format=geojson&refine.nom_ept=Paris&refine.dep=75&timezone=Europe/Berlin&lang=fr

Attribute Definition

- c_ir : IRIS Code
- c_typeir : Type of IRIS: Habitat (H), Activity (A), Miscellaneous (D), Other (Z)
- c12_pop1_1 : Number of persons aged 15 years or over Farmers
- c12_pop1_2 : Number of persons aged 15 years or more Craftsmen, Traders, Entrepreneurs
- c12_pop1_3 : Number of persons aged 15 years and over Executives and Senior Professionals
- c12_pop1_4 : Number of persons aged 15 or over Intermediate occupations
- c12_pop1_5 : Number of persons 15 years of age or older Employees
- c12_pop1_6 : Number of persons 15 years or older Workers
- c12_pop15p : Number of employed persons 15 years of age or older
- com : Department code followed by the commune or municipal district number for Paris Lyon and Marseille
- denspop12 : population density
- dep : Department Code
- geometry : IRIS geometry
- iris : Department code followed by the commune or municipal district number followed by the IRIS number
- iris_num : IRIS number
- l_ir : IRIS label
- libcom : Name of the municipality or municipal district for Paris Lyon and Marseille
- libris : "Wording of the IRIS within the municipality or municipal district for Paris Lyon and Marseille
- n_qu : Neighbourhood number
- n_sq_ir : Iris number
- nom_ept : Name of the city
- objectid : Object Id
- p12_f0014 : female population between 0 and 14 years old
- p12_f1529 : female population between 15 and 29 years old
- p12_f3044 : female population between 30 and 40 years old
- p12_f4559 : female population between 45 and 59 years old
- p12_f6074 : female population between 60 and 73 years old
- p12_f75p : population women over 75 years old
- p12_h0014 : male population between 0 and 14 years old
- p12_h1529 : male population between 15 and 29 years old
- p12_h3044 : male population between 30 and 40 years old
- p12_h4559 : male population between 45 and 59 years old
- p12_h6074 : male population between 60 and 73 years old
- p12_h75p : male population over 75 years old
- p12_pop : total population in 2012
- p12_pop_et : foreign population
- p12_pop_im : immigrant population
- tx12_pop_e : Foreign population rate
- tx12_pop_i : Foreign population rate of immigrants
- typ_iris : Type of IRIS
- uu2010 : Department code or "00" for urban units extending over several departments or even across the border followed by a code on a position indicating the population size and then a serial number

within the size

2.2 Foursquare API

With Foursquare API we have seafoods restaurants and fish shop.

2.4 Source of data

All data will be cleaned, analysed and visualised.

- The IRIS dataset : opendatasoft.com :
https://public.opendatasoft.com/explore/dataset/iris-demographie/download/?format=geojson&refine.nom_ept=Paris&refine.dep=75&timezone=Europe/Berlin&lang=fr
- Foursquare API
- Geocoder
- fish shop : fidh_paris.csv :
https://github.com/mbouayoun/Coursera_Capstone/blob/master/fishs_paris.csv

we can use other dataset if necessary such as fish shops : data.gouv.fr : https://www.data.gouv.fr/fr/datasets/r/c627e81a-a353-462d-a471-c2e645a9d14c

3. Methodology

I download and extract the information about IRIS dataset of Paris, I use the public.opendatasoft.com website :

https://public.opendatasoft.com/explore/dataset/iris-demographie/download/?format=geojson&refine.nom_ept=Paris&refine.dep=75&timezone=Europe/Berlin&lang=fr

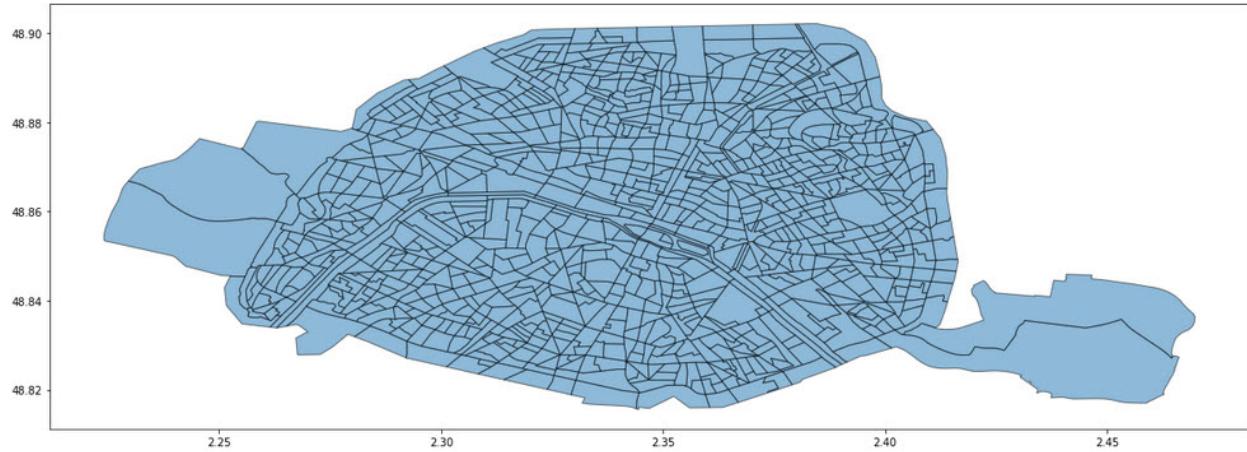


Figure 4: IRIS

The dataset is geoson file a format for encoding a variety of geographic data structures. GeoJSON supports the following geometry types: Point, LineString, Polygon, MultiPoint, MultiLineString, and MultiPolygon. Geometric objects with additional properties are Feature objects. Sets of features are contained by FeatureCollection objects.

In this project we are going to find the places close to the fish sellers and far from the conquerors for our new seafood restaurant.

- Black circle : fish sellers
- Big green circle : location for a new seafood restaurant
- Others circles : Conquerors restaurant

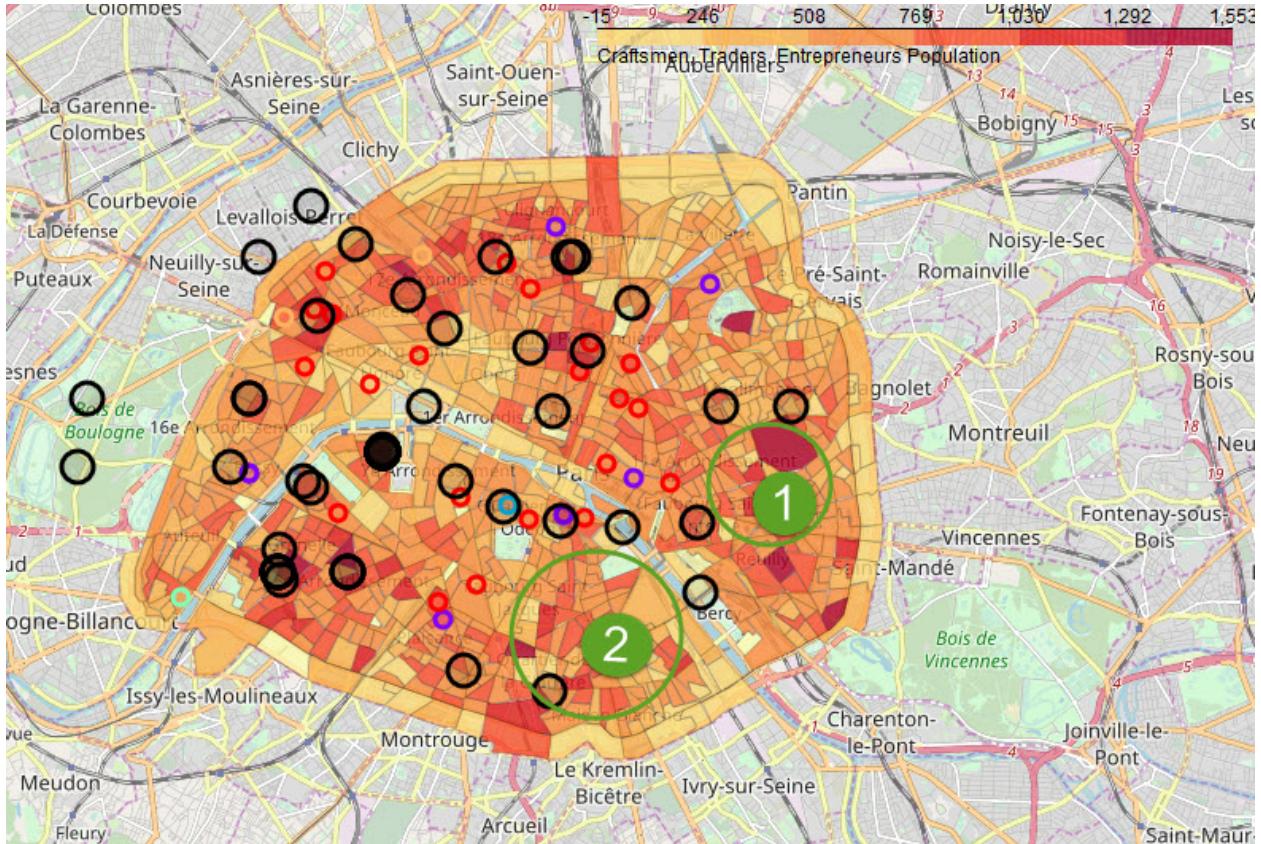


Figure 5: location

Here are the steps we are going to follow:

- Download the IRIS dataset
- Clean the IRIS dataset
- Analyze the IRIS dataset
- Visualize the IRIS dataset
- Utilize the Foursquare API to explore the neighborhoods and seafood restaurant in Paris
- Merge cleaned IRIS dataset with seafoods restaurant

- Segment neighborhoods using k-means clustering to locate seafood restaurant
- Overlay clusters with restaurants, fish shops and activity people
- Take places that are far from restaurants and near to fish shops according to the density of the people

4 Analysis

4.1 Cleaning data

Our *iris_paris* dataset contains several columns we don't need. We will create a new dataset *clean_paris* that contains only the columns we need.

We will also delete places like rivers and parks where we can't install a restaurant.

The surface of the IRIS polygons are small which forces us to choose the search radius at 160 meters, if the limit is 500 we will have several polygons next to one place.

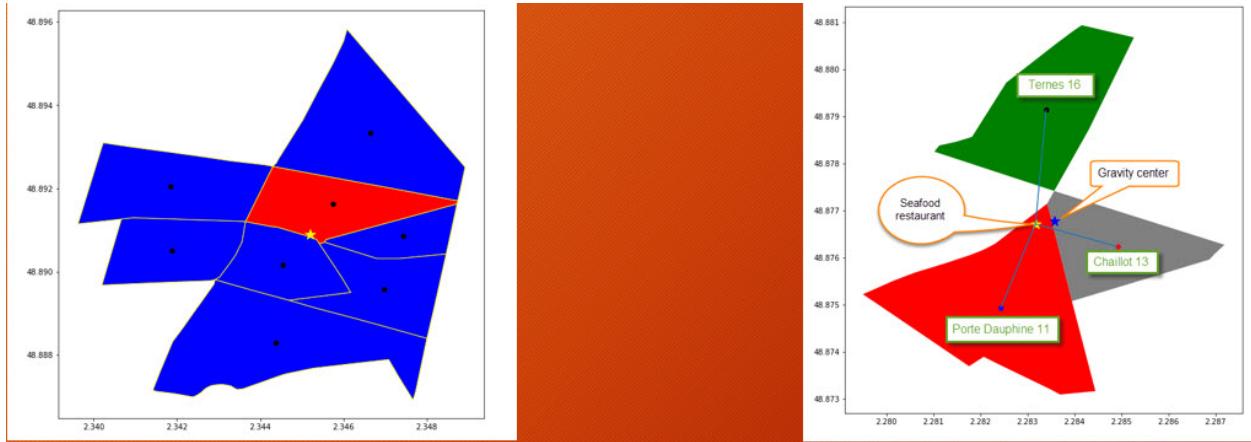


Figure 6: LIMIT > 160

4.2 Exploratory Data Analysis

Here we can see the population density according to their activities:

- Farmers
- Craftsmen, Traders, Entrepreneurs
- Executives and Senior Professionals
- Intermediate occupations

Their histograms show that they follow a normal law flattened to the right.

I Use Foursquare API to get information about some venues around Paris (*seafood_paris*) and csv file *fish_paris* to find fish sellers.

After merging *clean_data* and *seafood_paris* dataset and grouping, we use K-Mean Algorithm.

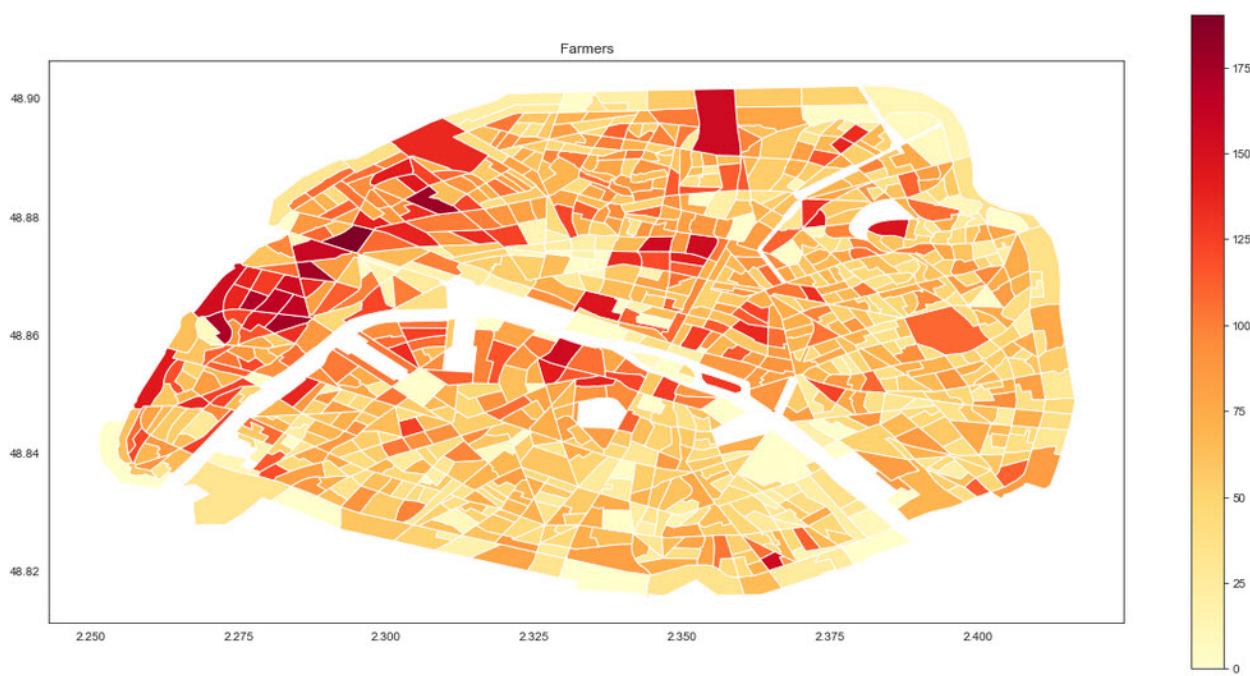


Figure 7: Farmers



Figure 8: Craftsmen, Traders, Entrepreneurs

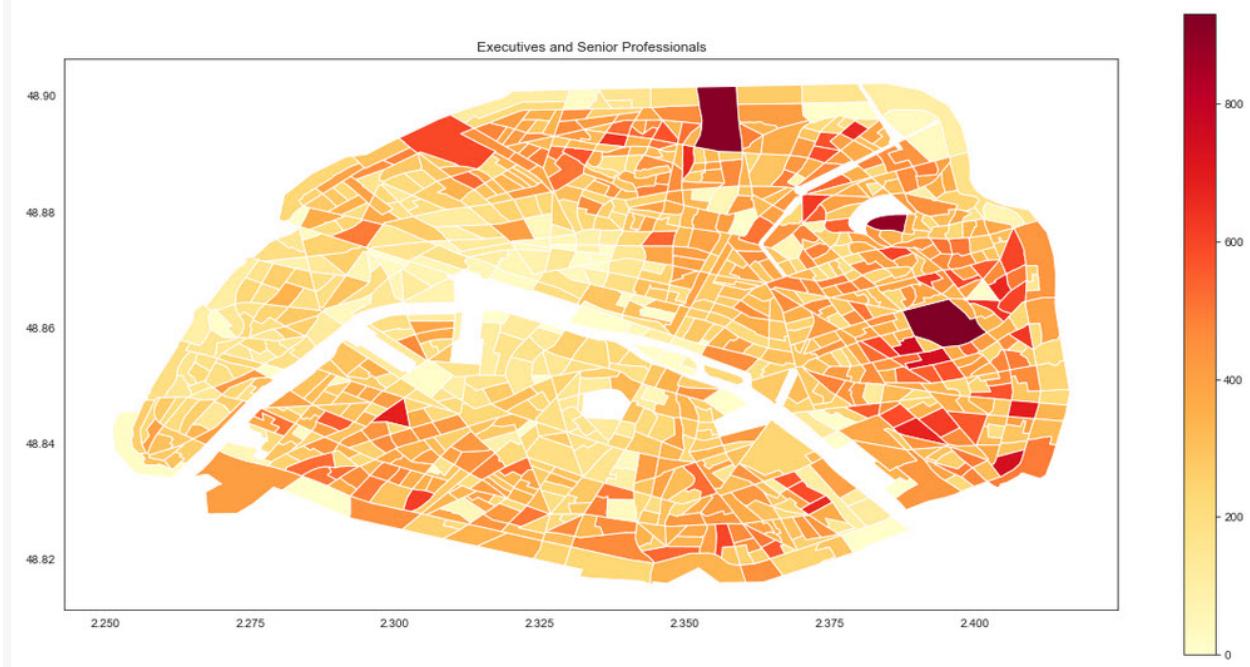


Figure 9: Executives and Senior Professionals

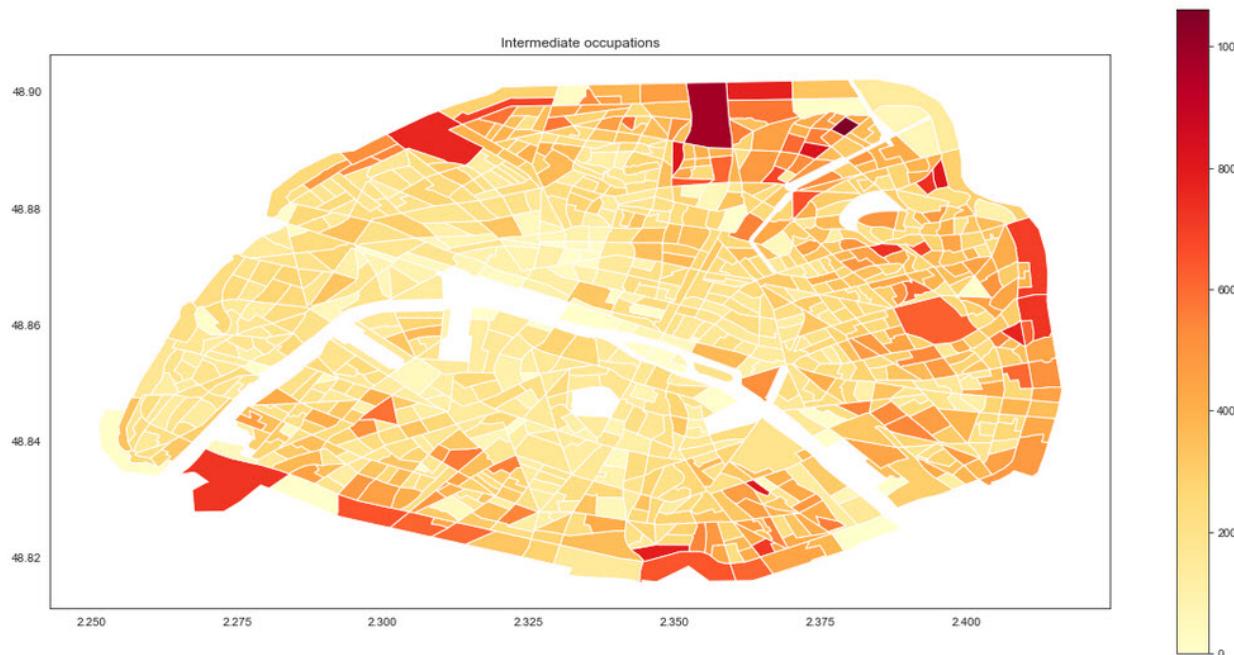


Figure 10: Intermediate

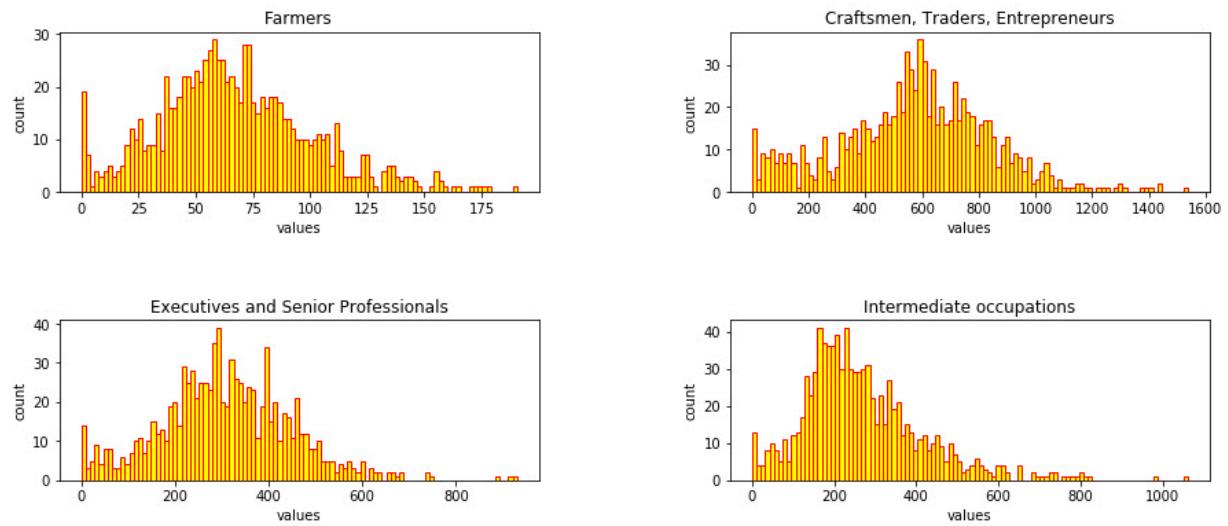


Figure 11: Density

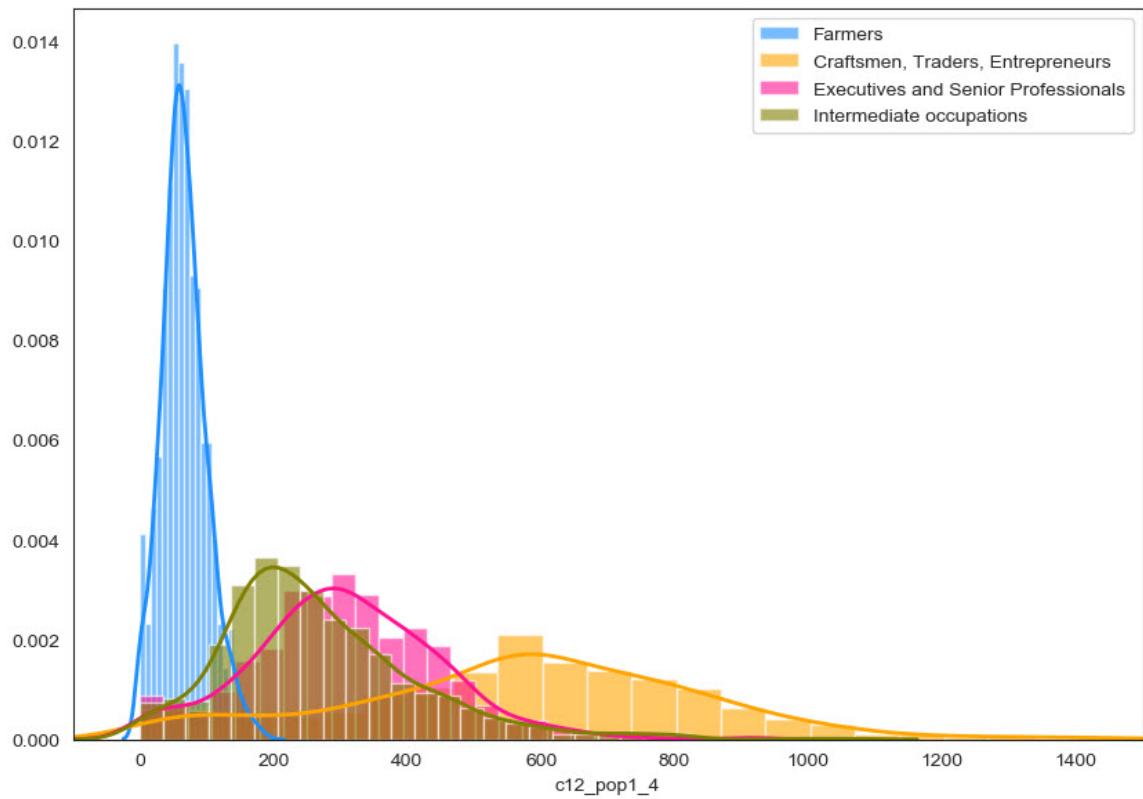
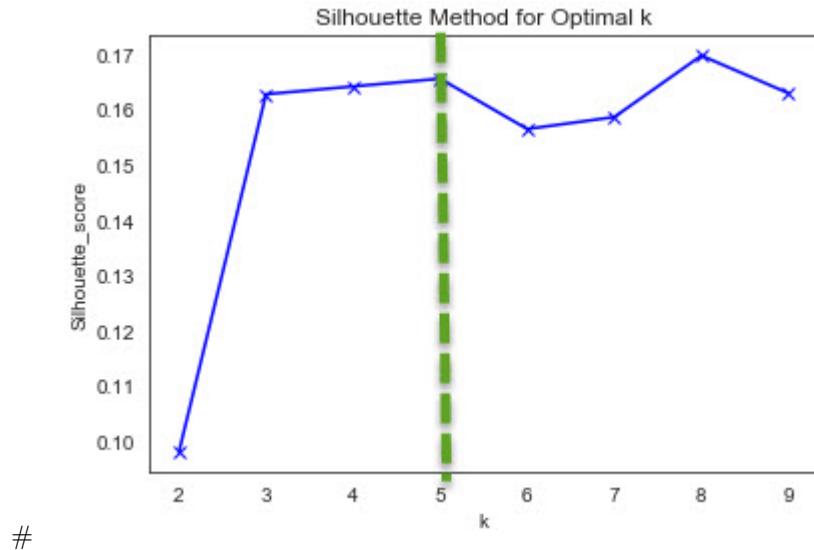


Figure 12: Density2

To find best number of cluster I use silhouette Method : (k=5)



#

And final resulting clusters visualization:

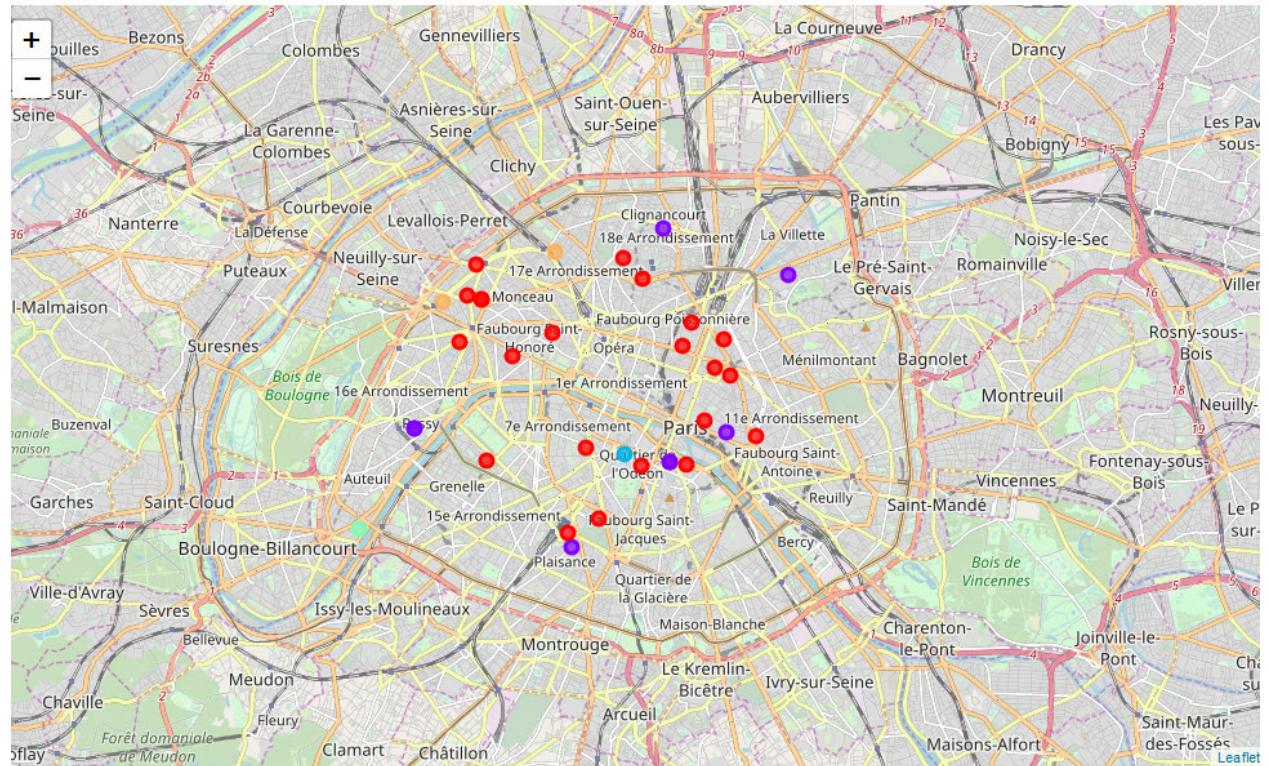


Figure 13: clusters

5. Result

K-meand find 5 clusters

- Cluster 0: It is the largest cluster where you can find all the people who work in Paris.
- Cluster 1: it's mostly travellers who go to hotels...
- Cluster 2: it's the people who love shows, there's the odeon theatre and film studios.
- Cluster 3: less urbanized, people who escape from the center of Paris.
- Cluster 4: people who like to go for a walk in Paris.

For Executives and Senior Professionals density we find 2 big places (green circle).

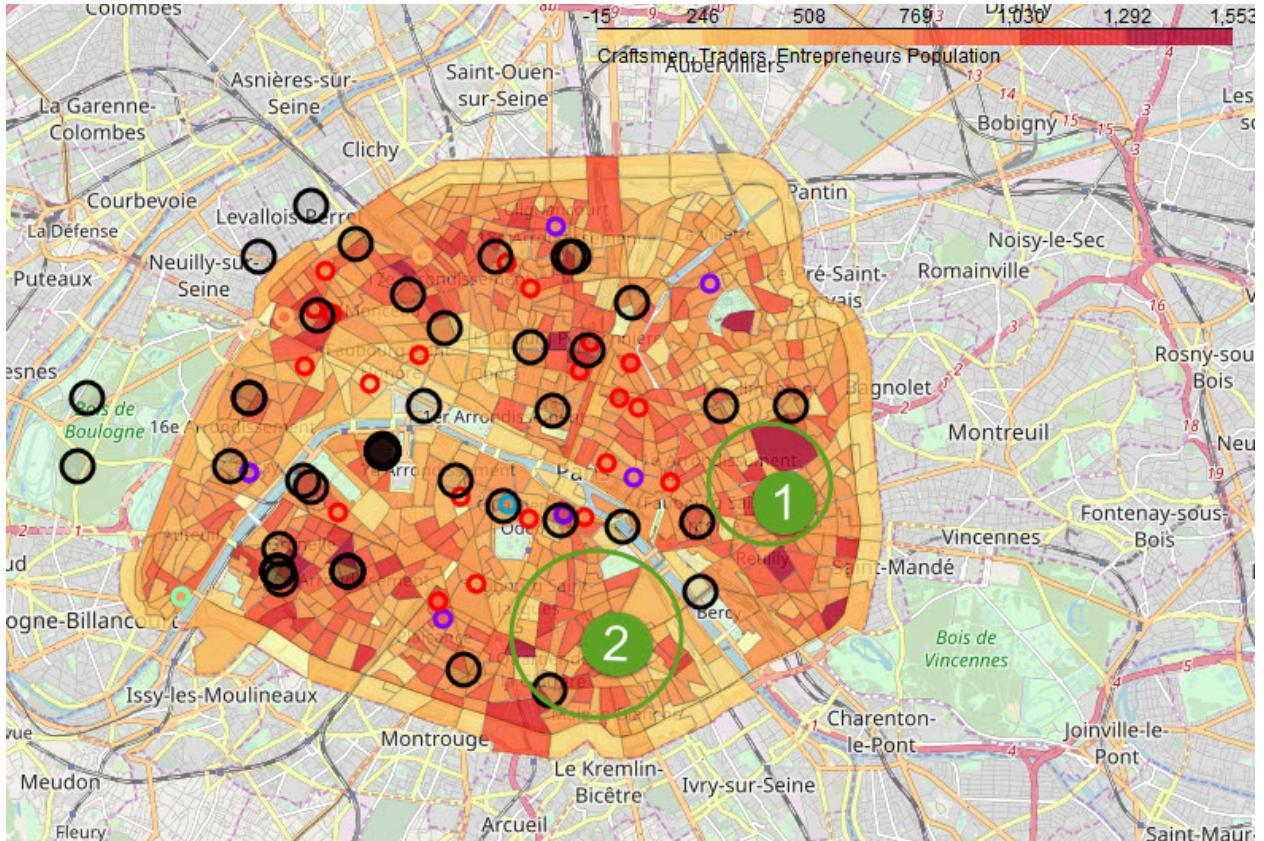


Figure 14: Executives and Senior Professionals

for immigrants we find also 2 places :

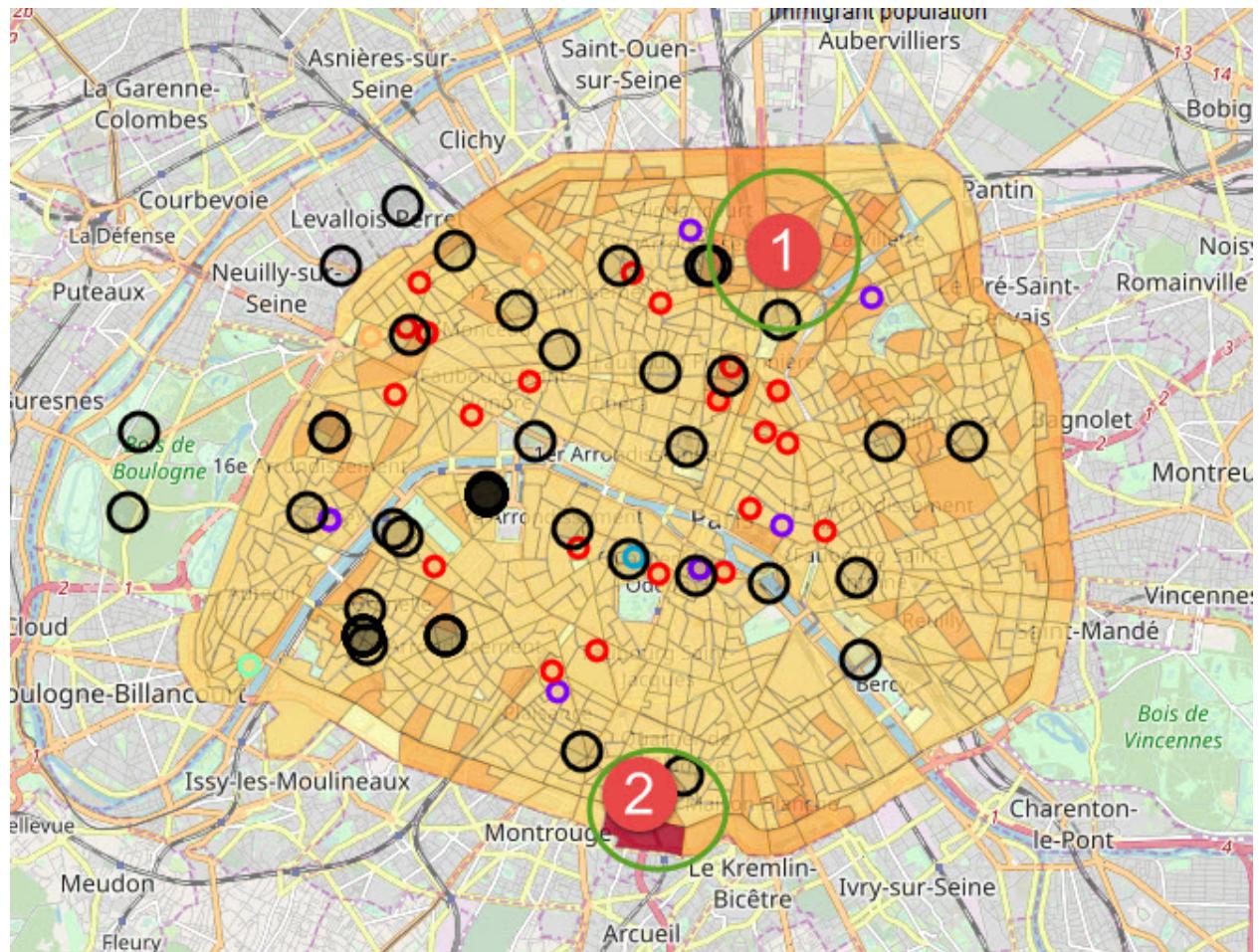


Figure 15: Immigrant

6. Conclusion

The best choice for new seafood restaurant depends on the demographic trends in the neighbourhood and the characteristics of the individuals frequenting the neighbourhood (residents, transients, employees, workers...) we can use others criteria for choosing a good commercial location.

- The flow of passers-by (or cars).

The flow of passers-by is likely to vary greatly from one street to another, or from one sidewalk to another. The turnover of a shop can be divided or multiplied by 3 depending on whether it is located in a given place or 50 metres away.

Are the sidewalks wide enough to allow passers-by to stop and stroll ?

Are there any obstacles or dangers that draw the eye of passers-by to another point ?

What is the direction of traffic for passers-by? Where do they go ?

How fast do cars go by? Do drivers have the possibility to turn their eyes away or slow down ?

- Ease of access.

Do the target customers travel by car, on foot or by public transport instead ?

Is the location easily accessible? Is it possible to park nearby ?

Are there public transportation stops nearby ?

- The attractiveness of the business environment.

The presence of competitors nearby is not a negative point, on the contrary it can be stimulating.

The type of activity of the nearby shops must be consistent with the targeted activity. For example, a cheese maker would be well advised to set up next to a wine merchant.

The presence of a single large sign in the vicinity can be enough to energize an entire neighborhood.

Is the street dynamic favourable or unfavourable ?

- Visibility.

Is the premises visible from afar? from up close? from the other side of the street ?

Is the storefront easily noticeable ?

How long is the window? A window that is too narrow will be almost invisible.

Are there elements that tend to hide the window or sign? Street furniture, trees, vehicles, signs.

A street corner location often proves to be interesting.

- The exhibition.

Does the room receive natural light or direct sunlight? This can be decisive if customers enjoy eating on the terrace, for example.

- The surface area and interior design of the room.

The size of the room must be adapted to the activity.

Guests will not feel comfortable in a room that is too cramped or bicornuous.

The sales area must be appropriate, pleasant, and consistent with the surface area of the stock.

- Future urban developments.

It is essential to find out about any future works or developments on the street in question.

7. Discussions

In the *fish_paris* dataset we have the vote of the customers that can help us to use the recommendation system. We can be close to a restaurant with a bad rating and far away if it has a good rating.