

# Capstone- IBM Applied Data Science by IBM

## Feasibility of Indian or Italian style Restaurant in Clermont-Ferrand, France

### Table of Contents

*Introduction*

*Data*

*Methodology*

*Results*

*Discussion*

*Conclusion*

# 1. Introduction: Business Problem

## 1.1 Background

Clermont-Ferrand is a university city in central France, bordered by the volcanic Chaîne des Puys mountains. Near the fountains and statues of Place de Jaude square are the Gothic Notre-Dame-de-l'Assomption Cathedral, constructed from lava stone, and the Romanesque Basilica of Notre-Dame du Port, with mosaics.

It's metropolitan area has a population of 467178 as of census of 2011.

Clermont-Ferrand hosts the [Clermont-Ferrand International Short Film Festival](#) (Festival du Court-Métrage de Clermont-Ferrand), one of the world's leading international festivals for short films. It is also home to the corporate headquarters of [Michelin](#), the global tyre company founded there more than 100 years ago. Clermont-Ferrand's most famous public square is the Place de Jaude with the statue of Vercingetorix in the centre. Because of its various universities and also HQ of the global company Michelin the city has a sizeable population of expats. The city centre is bustling with its number of galleries, café's, bakeries and restaurants.

## 1.2 Problem

In this project we will try to find an optimal location for a restaurant. The choice of the type of restaurant would be an Indian or Italian cuisine restaurant

Since there are a sizeable number of restaurants in the city of Clermont-Ferrand, we will try to detect locations that are not already crowded with restaurants.

We are also particularly interested in areas with no Italian or Indian restaurants in vicinity. However, we would also prefer locations as close to city centre as possible, assuming that first two conditions are met.

We will use the data science tools and methods to explore and generate a few most promising neighbourhoods based on these criteria.

Advantages of each area will then be clearly expressed so that best possible final location can be chosen by stakeholders.

## 1.3 Interest

Specifically, this report will be targeted to stakeholders interested in opening a restaurant in Clermont Ferrand, France.

# 2. Data

## 2.1 Data Sources

We decided to use regularly spaced grid of locations, centred around city centre, to define our neighbourhoods.

Following data sources will be needed to extract/generate the required information:

1. Coordinate of Clermont Ferrand city centre will be obtained using **geopy library** of well-known Clermont Ferrand location, Place de Jaude.
2. Centres of candidate areas will be generated algorithmically and approximate addresses of centres of those areas will be obtained using **geopy library**
3. Number of restaurants and their type and location in every neighbourhood will be obtained using **Foursquare API**

Based on definition of our problem, factors that will influence our decision are:

- Number of existing restaurants in the neighbourhood (any type of restaurant)
- Number of and distance to our choice of restaurants in the neighbourhood, if any
- Distance of neighbourhood from city centre

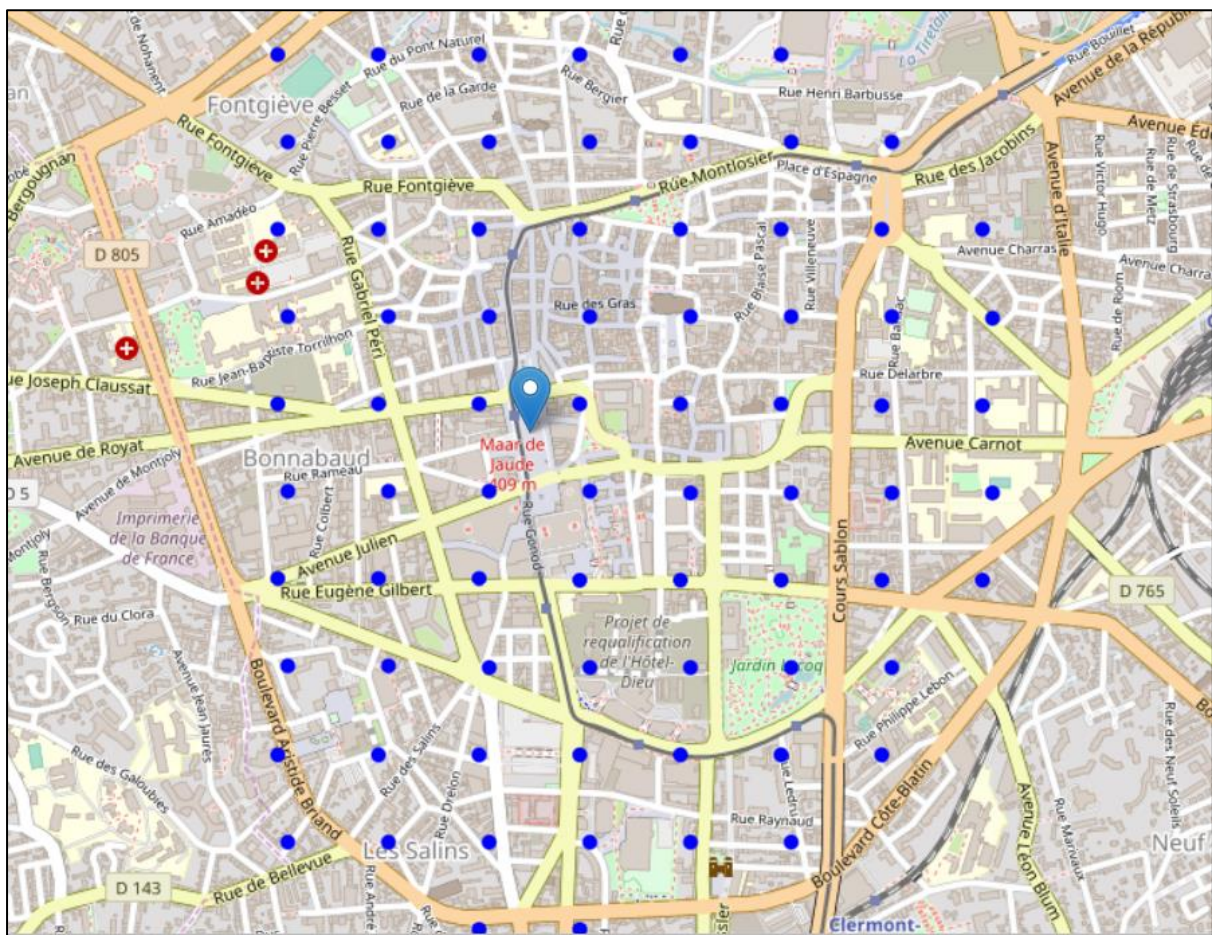
## 2.2 Data Cleaning

Since we are using the geopy library to obtain geographical data there is not much of data cleaning to be done. It is a matter of narrowing down the neighbourhood areas as well as the restaurant categories near our candidate area.

Geopy library has been used to determine the co-ordinates of areas within certain distance from the city centre. For narrowing down the neighbourhood area, the steps to create the grid of candidate areas have been iterated as many times to arrive at an optimum grid allocation.

The latitude and longitudes of candidate areas have been arrived at by spacing the neighbourhood areas 200 meters apart from each other and these areas are within 1 km from the city centre (Place de Jaude).

After finalizing the distance, area of the neighbourhoods we get a total of 79 candidate neighbourhoods generated.



## 2.3 Feature selection

In the above feature list obtained we have the co-ordinates of the areas and the distance of the area from the city centre. We need to have the address of the location to have a more meaningful identification and analysis. So, to this feature list we add the address of all the co-ordinates so

obtained. The geopy library has been used to obtain the addresses. After reverse transformation of the co-ordinates we have the neighbourhood data with complete addresses. This is copied to a data frame as well as uploaded to a pickle file.

	Address	Latitude	Longitude	X	Y	Distance from center
0	(55, Boulevard Jean Jaurès, Les Salins, Clermo...	45.767590	3.080837	506285.588846	5.068229e+06	985.471599
1	(3, Rue du Tonnet, Les Salins, Clermont-Ferran...	45.767588	3.083410	506485.588846	5.068229e+06	985.471599
2	(Berlioz, Rue Berlioz, Chamalières, Clermont-F...	45.769152	3.075953	505905.588846	5.068402e+06	939.116090
3	(Rue de Vallières, Les Salins, Clermont-Ferran...	45.769150	3.078525	506105.588846	5.068402e+06	854.364694
4	(32, Avenue Marx Dormoy, Les Salins, Clermont-...	45.769148	3.081097	506305.588846	5.068402e+06	811.134410
5	(Coubertin, Place Pierre de Coubertin, Les Sal...	45.769146	3.083669	506505.588846	5.068402e+06	816.050875
6	(Résidence Pierre de Coubertin, 18, Rue Abbé d...	45.769145	3.086241	506705.588846	5.068402e+06	868.296626
7	(25, Rue Vermeuouse, Saint-Jacques, Clermont-F...	45.769143	3.088814	506905.588846	5.068402e+06	960.176562
8	(62, Boulevard Aristide Briand, Chamalières, C...	45.770711	3.075697	505885.588846	5.068575e+06	807.417976
9	(Regensburg, Rue de La Rochefoucault, Les Sali...	45.770709	3.078270	506085.588846	5.068575e+06	701.372789

I have used the *Explore* call to the Foursquare API to obtain the venues nearby the candidate neighbourhoods. From the venues selection we will include in our selection only those venues that have ‘restaurant’ in their category name. Apart from this we will also include all sub categories of *Italian and Indian restaurants* in the neighbourhood.

So will have the following feature list

- List of all restaurants in the neighbourhood area
- List of Italian restaurants in the neighbourhood area
- List of Indian restaurants in the neighbourhood area.
- List of restaurants around locations

For this feature list venues within a radius of 500 meters from the candidate neighbourhood area is explored.

<div>List of all restaurants</div> <div><div>(<div>'5a2ae756e185f33e5e4683cb', 'A La Bonne Heure', 45.77092134, 3.0811042, '10 Avenue Marx Dormoy, 63000 Clermont-Ferrand, France', 498, False 06305.9519244365, 5068598.992463984)</div><div>(<div>'55211497498eb6846070637c', 'L'universal', 45.771499, 3.083177, 'France', 471, False, 506467.04792490683, 5068663.337426941)</div><div>(<div>'4e27feaa1f6e88a1545d3d4f', 'De L'eau à la Bouche', 45.771427, 3.083565, 'France', 476, False, 506497.22345323395, 5068655.36950919)</div><div>(<div>'52ee2e6b498e71d257936b27', 'Au p'tit marmiton', 45.76938533659876, 3.072494502934678, 'France', 418, False, 505636.6920506263, 5068427.695185)</div><div>(<div>'4fed8560e4b072533f1b687e', 'alfred', 45.7727826335402, 3.081596226069066, 'Rue du Puits Artesien (Rue Lagarlaye), 63000 Clermont-Ferrand, nce', 472, False, 506343.9965299406, 5068805.825563097)</div><div>(<div>'4d4da41f819554813022da93', 'Brasserie du Jardin', 45.7708275104889, 3.0871489366727873, '55 boulevard François Mitterrand, 63000 Clermont-rand, France', 227, False, 506775.94903479517, 5068589.062259003)</div><div>(<div>'5149f229e4b000e635293079', 'Grill Palace', 45.765504594235736, 3.0918818070298175, '15 Rue Des Plats, France', 469, False, 507144.615703805067998.085841427)</div><div>(<div>'4c9c7e8dd3c2b60cd8ecb7bc', 'Un Assaggio', 45.773966414060226, 3.0769536428780477, 'Clermont-Ferrand, France', 448, False, 505982.9153723635068936.988580587)</div><div>(<div>'4be1b45eedb0f473365a615', 'Restaurant L'Odevie', 45.77374172409866, 3.0814112947485057, '1 rue Eugène Gilbert, 63000 Clermont-Ferrand, Fr e', 388, False, 506329.5098735292, 5068912.368195534)</div><div>(<div>'4f3ab173e4b05505712b7b41', 'N°3', 45.77378874448363, 3.0850155035131723, 'Rue Georges Clemenceau, 63000 Clermont-Ferrand, France', 446, Fa e, 506609.7218810226, 5068917.883908532)</div></div></div></div></div></div></div></div></div></div></div></div>	<div>List of Italian restaurants</div> <div><div>(<div>'4c9c7e8dd3c2b60cd8ecb7bc', 'Un Assaggio', 45.773966414060226, 3.0769536428780477, 'Clermont-Ferrand, France', 448, True, 505982.91535068936.988580587)</div><div>(<div>'4f3ab173e4b05505712b7b41', 'N°3', 45.77378874448363, 3.0850155035131723, 'Rue Georges Clemenceau, 63000 Clermont-Ferrand, France', 4, 506609.7218810226, 5068917.883908532)</div><div>(<div>'4d25b0e6bcc1224b1be5921b', 'Pizzeria Italia', 45.77513116580566, 3.083911291825762, 'mail d'Allagnat (5 place Louis Aragon), 63000 C errand, France', 416, True, 506523.7159306328, 5069066.939517381)</div><div>(<div>'4cb0de4039458cfa8acd06a0', 'Bistro Vénitien', 45.77871170588981, 3.0837842263383792, '32 rue des Gras, 63000 Clermont-Ferrand, Franc True, 506513.4202989816, 5069464.736105854)</div><div>(<div>'5179784ee4b0b9ec8469f127', 'Il Visconti', 45.77858645389464, 3.087191235264401, 'France', 469, True, 506778.2977487672, 5069451.1035...</div></div><div>Total: 5</div></div></div></div></div></div>
---	---

Restaurants around location 11: Restaurant l'Odevie, N°3, Un Assaggio, Le Wylder  
 Restaurants around location 12: Restaurant l'Odevie, Los Dos Hermanos, Pizzeria Italia, N°3, Starfood, Le Wylder  
 Restaurants around location 13: Restaurant l'Odevie, Los Dos Hermanos, N°3, Starfood, A La Bonne Heure, 231 East Street  
 Restaurants around location 14: N°3, Starfood  
 Restaurants around location 15:  
 Restaurants around location 16: alfred, Restaurant l'Odevie, Un Assaggio, Au p'tit marmiton, Le Madras  
 Restaurants around location 17: Restaurant l'Odevie, L'Auguste, Un Assaggio, Le Wylder  
 Restaurants around location 18: Restaurant l'Odevie, Pizzeria Italia, Los Dos Hermanos, L'Auguste, N°3, Le Pont de Jaude, Le Wylder, Un Assaggio, Starfood  
 Restaurants around location 19: Restaurant l'Odevie, Los Dos Hermanos, Pizzeria Italia, N°3, L'Auguste, Le Pont de Jaude, Konnichiwa, Starfood, Le Wylder  
 Restaurants around location 20: Restaurant l'Odevie, Los Dos Hermanos, Pizzeria Italia, N°3, L'Auguste, Le Pont de Jaude, Starfood, Konnichiwa, 231 East Street, A La Bonne Heure  
 Restaurants around location 21: Los Dos Hermanos, Pizzeria Italia, N°3, Starfood, Konnichiwa  
 Restaurants around location 22: Starfood  
 Restaurants around location 23: alfred, Restaurant l'Odevie, Un Assaggio, Le Madras  
 Restaurants around location 24: Restaurant l'Odevie, Pizzeria Italia, L'Auguste, Le Pont de Jaude, Le Faisan Doré, Un Assaggio, Le Wylder, Le Bouchon de Jaude  
 Restaurants around location 25: Restaurant l'Odevie, Pizzeria Italia, Los Dos Hermanos, Le Boeuf Café, L'Auguste, La Gourmandine, N°3, Le Pont de Jaude, Le Faisan Doré, Lard et la Manière  
 Restaurants around location 26: Restaurant l'Odevie, Los Dos Hermanos, Pizzeria Italia, Le Sisisi, Le Boeuf Café, L'Auguste, N°3, Le Pont de Jaude, Le Faisan Doré, Konnichiwa

### 3. Methodology

In this project we will direct our efforts to find areas in Clermont Ferrand with low restaurant density and also areas with less Italian and/or Indian restaurants. We will limit our area of exploration and analysis to within 1 km of the city centre, Place de Jaude.

In the first step we have gathered data about all the restaurants within 1 km of the city centre, Place de Jaude. We have also identified the Italian and Indian restaurants amongst the list of restaurants obtained in the neighbourhood. We also know which restaurants are situated in the vicinity of the candidate neighbourhood.

Our second step will be the analysis of the data so collected and narrowed down. In this step we will do some exploratory analysis of the restaurant density of the neighbourhood area of Clermont - Ferrand. We will use **Heat maps** to identify promising areas with low restaurant density and also no *Italian or Indian* restaurant in the vicinity. We can then focus on these areas so highlighted

In third and final step we will focus on most promising areas and within those create clusters of locations that meet some basic requirements established in discussion with stakeholders: we will take into consideration locations with no more than two restaurants in radius of 200 meters, and we want locations without Italian and/or Indian restaurants in radius of 250 meters.

We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighbourhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

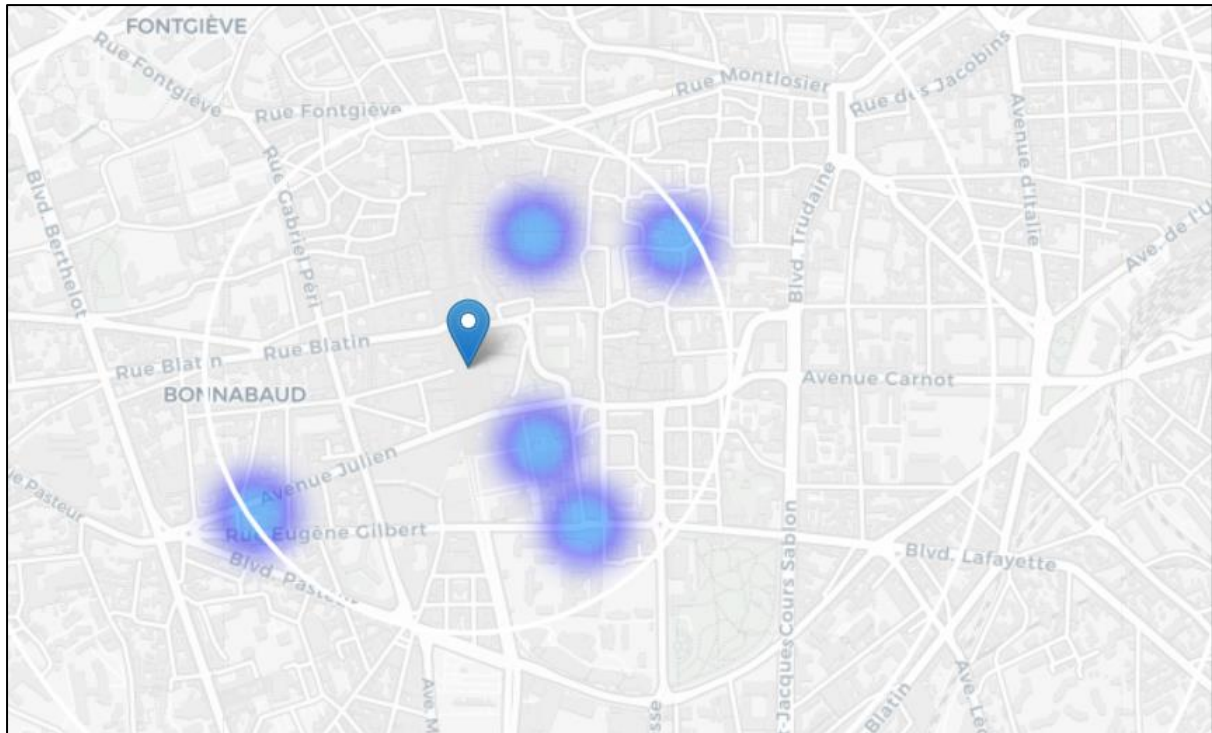
### 4. Analysis

Let us perform some basic exploratory analysis. We will start by taking a count of the restaurants in the candidate neighbourhoods identified.

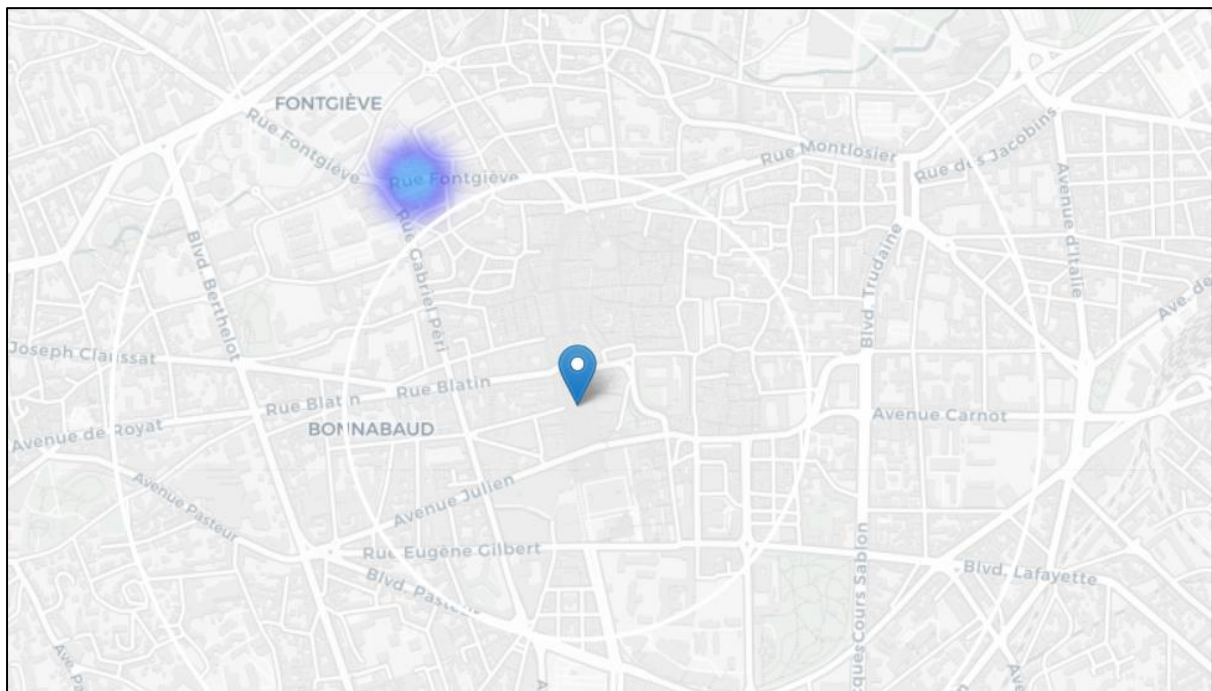
Then we will calculate the distance to every Italian and/or Indian restaurant from every neighbourhood area.





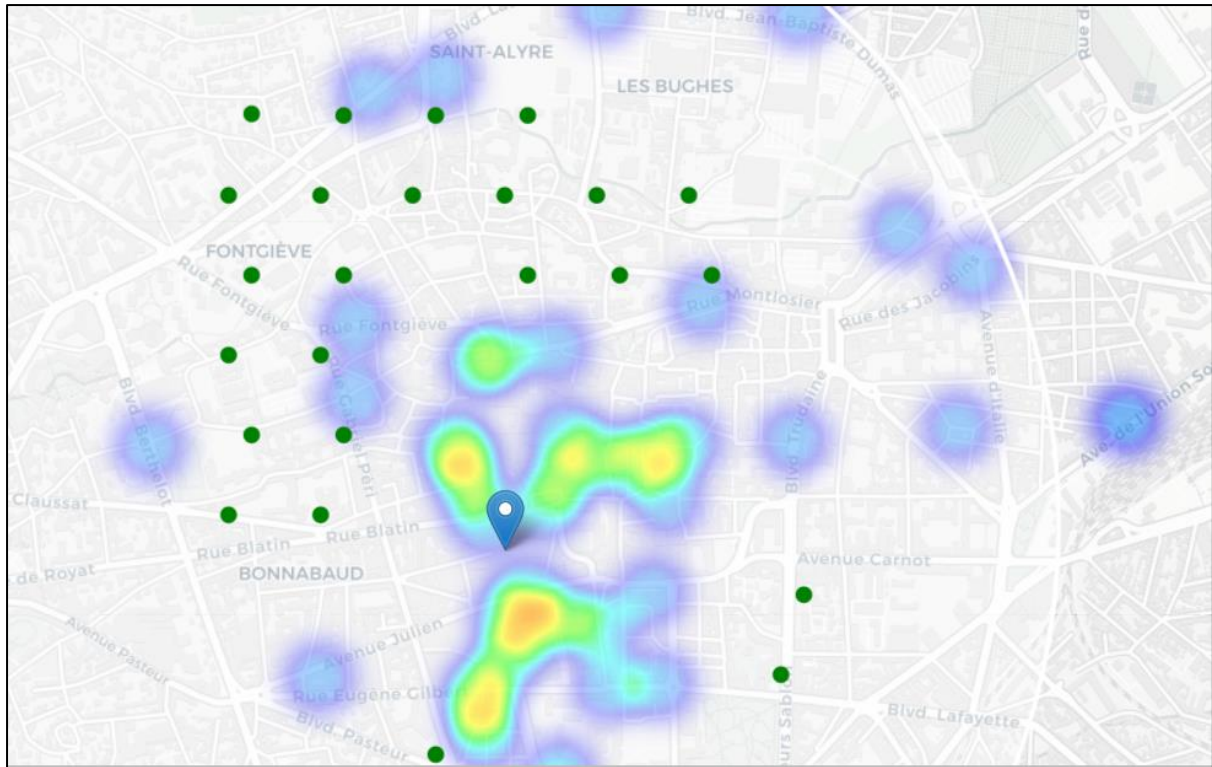


## Indian



The HeatMap of Italian restaurants is not very 'hot' and in case of the Indian restaurant category we see that the map is pretty cool.

To further narrow down the feasible addresses and also to take care of the specification which is, not more than 3 restaurants in the area and no Italian and/or Indian restaurants within a radius of 300 mts, we created a grid of location candidates in the new region of interest of Bonnabaud. The neighbourhood areas are close to each other at 200 mts apart.



We also created a cluster of all the candidate areas to have 10 zones of interest. These 10 addresses would be the starting point for further analysis and decision making.

Use reverse coding to obtain addresses of the 10 areas from the co-ordinates

```

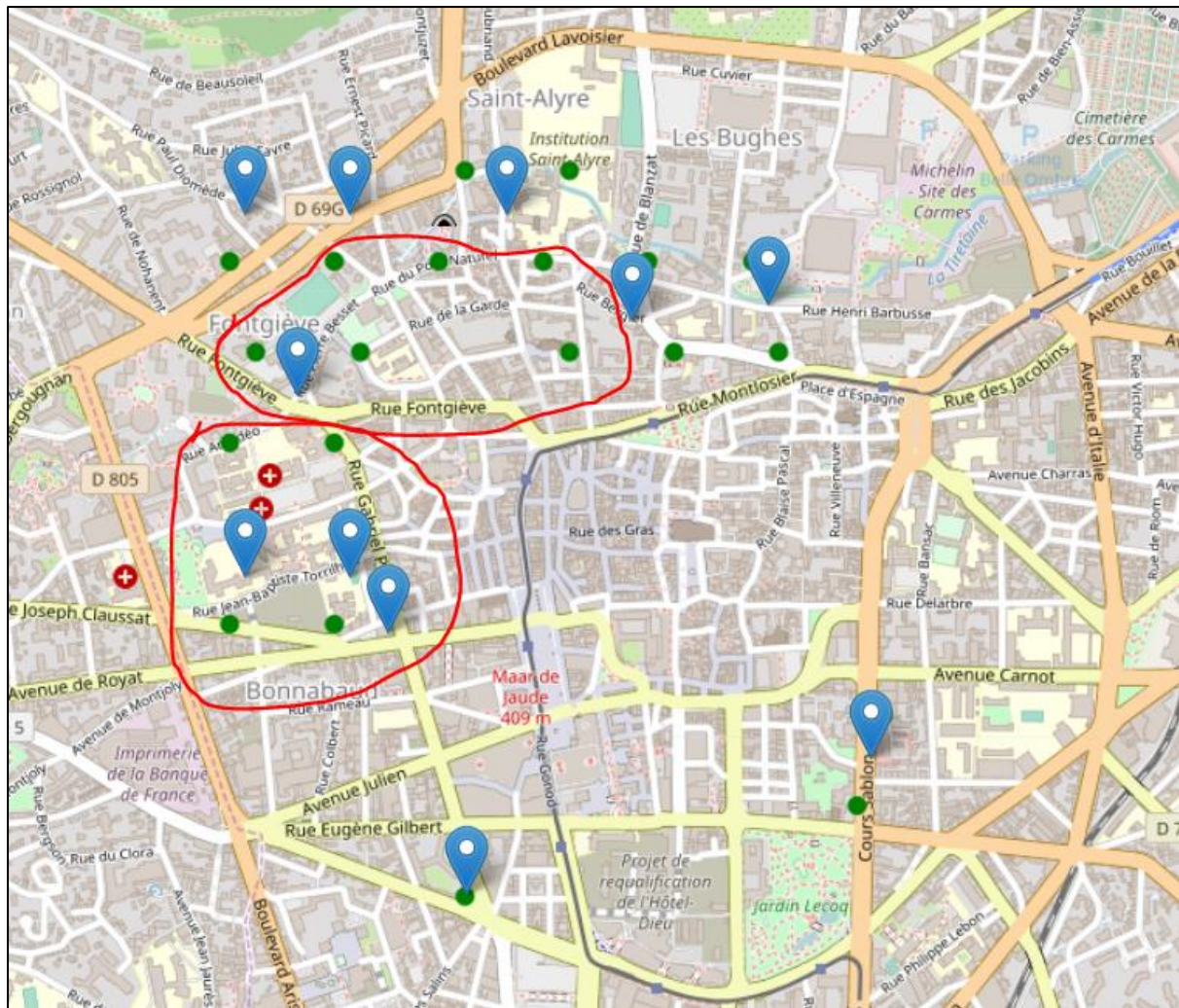
=====
Addresses of centers of areas recommended for further analysis
=====

16, Rue Bergier, Ecoquartier de Tremonteix, Saint-Alyre, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France
e de Jaude
Fontgîève, Rue Pierre Besset, Ecoquartier de Tremonteix, Fontgîève, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France
m from Place de Jaude
Hélicoptère Gazelle, Cours Sablon, Saint-Jacques, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France =>
ude
Central téléphonique, Boulevard Lavoisier, Ecoquartier de Tremonteix, Fontgîève, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France
=> 1.01km from Place de Jaude
Lycée professionnel Amédée Gasquet, Rue Antoine Menat, Ecoquartier de Tremonteix, Bonnabaud, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 6
=> 0.83km from Place de Jaude
Résidence Arverne, Boulevard Pasteur, Les Salins, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France =>
ude
Institution Saint-Alyre, Place Sainte-George, Ecoquartier de Tremonteix, Saint-Alyre, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, F
=> 0.87km from Place de Jaude
Rue Béranger, Ecoquartier de Tremonteix, Fontgîève, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France
Jaude
Parking des Trois Ponts, Rue Henri Barbusse, Les Bughes, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, France métropolitaine, 63000, France
e de Jaude
Centre Hospitalier Psychiatrique Sainte-Marie, Rue Jean-Baptiste Torrilhon, Ecoquartier de Tremonteix, Bonnabaud, Clermont-Ferrand, Puy-de-Dôme, Auvergne-Rhône-Alpes, Fra
0, France
=> 0.56km from Place de Jaude

```

Based on the addresses obtained with the data on the distance to the city center we can say that the clusters in the Bonnabaud and Fontgieve areas are more feasible.





## 5. Result and Discussion

Our analysis shows that there are quite a number of restaurants near the city centre, about 52 restaurants. Our area of analysis was 1 km from the city centre. We also have areas of low restaurants density that are close to the city centre. These are in the area of west of the city centre.

So our efforts were to identify areas close to the city centre but with a lesser density of restaurants.

Our HeatMap of the restaurants in the area shows high concentration of restaurants in the North, North east and south east of the city centre.

The HeatMap of Italian restaurants is not very 'hot' and in case of the Indian restaurant category we see that the map is pretty cool.

There is no visible density of restaurants close to the city centre and also within 1 km from the city centre. There is just 1 restaurant in the Fontgieve quarters which is about 700 meters from the city centre.

The total presence of Italian restaurants is only 9% and for India restaurants it is just 1% of the total number of restaurants.

Taking into account the specification of having not more than 3 restaurants in the radius of 250 mts and no Italian/Indian restaurant in the vicinity of 300 mts, we put focus to the Bonnabaud and

Fontgieve quartier of the city. This area is not far from the city center. It is about 500 meters from Place de Jaude.

We also have a low density of restaurants in this area.

Analysis of the map created and also of websites relating to the Bonnabaud areas shows us that Rue Blatin and Avenue Julien are two of the main streets in this area leading to the city centre. Rue Fontgieve in the Fontgieve quartiers is the main street which also leads to the city centre. These streets are lined with residences as well as shopping areas and flagship shops leading to the city centre.

We then defined new grid location area which would be around Bonnabaud. For the new areas we had the neighbourhoods closer to each other at 50m apart and 500 meters from the city centre.

The candidate locations were then clustered to create centres of zones containing good locations.

The result of all this is that we have 10 addresses having our candidate areas to explore further. These 10 addresses should be taken only as the starting point for more detailed analysis. This could mean analysis on various factors like street-level exploration, area, cost etc.

The cluster zones as well as the candidate areas generated are feasible for having an Indian and/or Italian restaurant. The final decision would be up to the stakeholders based on various external factors and considerations.

## 6. Conclusion

Purpose of this project was to identify areas close to city center of Clermont-Ferrand with low number of restaurants (particularly Italian and Indian restaurants) in order to aid stakeholders in narrowing down the search for optimal location for a new Italian and/or Indian restaurant.

By calculating restaurant density distribution from Foursquare data, we have first identified areas that justify further analysis, and then generated extensive collection of locations which satisfy some basic requirements regarding existing nearby restaurants. Clustering of those locations was then performed in order to create major zones of interest (containing greatest number of potential locations) and addresses of those zone centres were created to be used as starting points for final exploration by stakeholders.

The cluster zones as well as the candidate areas generated are feasible for having an Indian and/or Italian restaurant. The final decision would be up to the stakeholders based on various external factors and considerations.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighbourhoods and locations in every recommended zone, taking into consideration additional factors like attractiveness of each location (proximity to park or water), levels of noise / proximity to major roads, real estate availability, prices, social and economic dynamics of every neighbourhood etc.