

The Battle of the Neighborhoods

OCTOBER 30

Authored by: Luka Sarić



Contents

Introduction.....	3
Background	3
Problem	3
Interest	3
Data acquisition and cleaning	4
Data sources	4
Data cleaning	4
Feature selection	4
Methodology	5
Results.....	6
Discussion.....	15
Conclusion	16
Acknowledgement	16

Introduction

Background

Paris is gigantic city, with thousands of years of history. Poets, artists, playwrights, writers, journalists, and more have all written about their love of this city...and it's hard not to fall in love with Paris. As we all know, Paris is one of the most visited places in the world, so opening any kind of business there can be quite a challenge.

People from all over the world come to Paris to see everything it has to offer and, of course, people enjoy visiting restaurants that serves their favorite cuisine.

Problem

Data that we will research is based on restaurants in Paris, mainly Italian. Our data will help us find the best place to open an Italian restaurant near the city center. Most of our data is geodata, which contains information about longitudes, latitudes, addresses, and all of them are vital for our project since we are trying to find the least populated place with other restaurants that is also crowded with lots of people.

Interest

The main goal is to find the best place in Paris that is near the city center to open an Italian restaurant. Of course, important thing is that there are not as many restaurants around since it would create competition. So, we are looking for a place that is not too crowded with other restaurants, especially Italian.

Data acquisition and cleaning

Data sources

We have collected our data from the Foursquare API mainly since it had information about latitude, longitude and addresses that we needed. We have also used <https://france-geojson.gregoireddavid.fr/> in order to obtain information about districts in Paris and label them on our map.

Data cleaning

Since most of our data came in correct format, there wasn't much of a cleaning job left to do. There were some redundant information about addresses, but since it's was a json file, we have selected through our code what exactly we want to extract from the request.

Feature selection

We have collected data for 364 addresses in Paris with attributes such as latitude, longitude. We have created two new features in order to calculate distances in meters. Than we have included that feature that showed distance for each address from the city center.

Methodology

In this project we will direct our efforts on detecting areas of Paris that have low restaurant density, particularly those with low number of Italian restaurants. We will limit our analysis to area ~6km around city center.

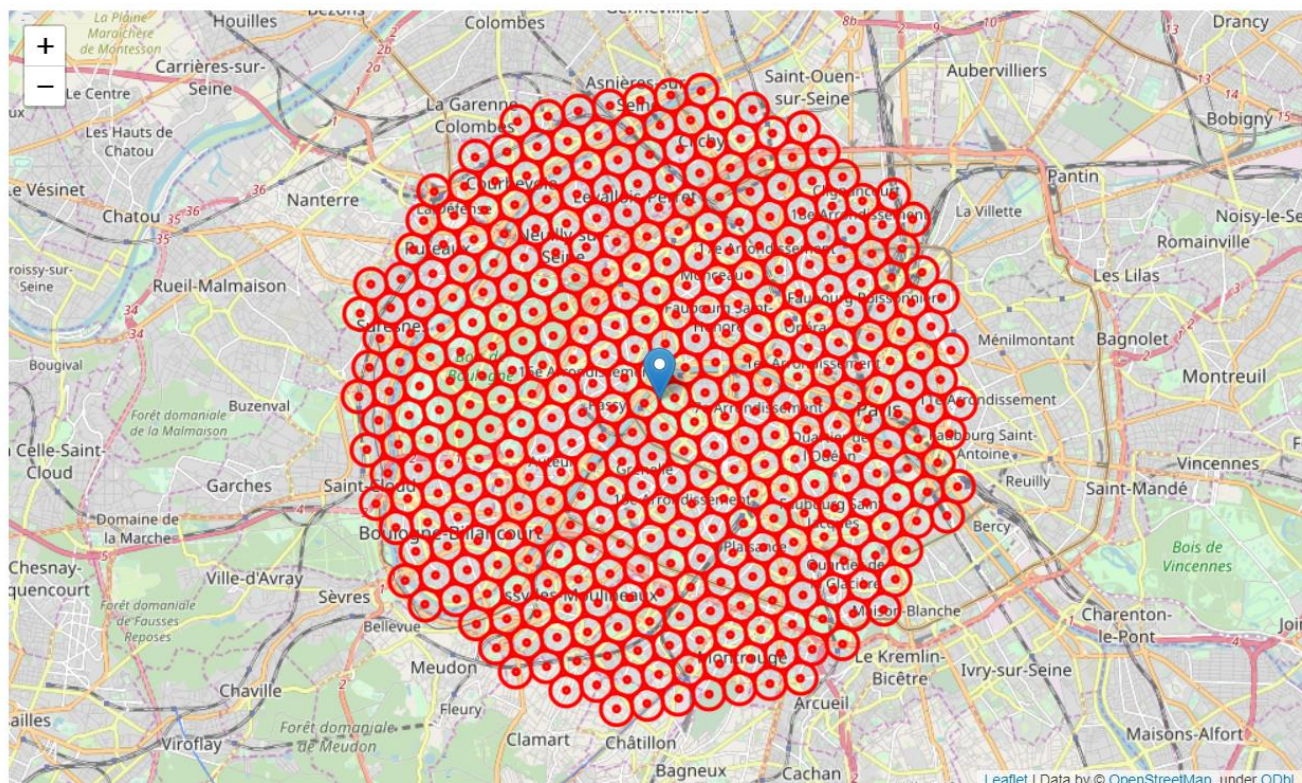
First, we have identified the exact location that we choose as the center point, which is the Eiffel Tower, Paris, France. We have decided to split the neighborhood into equal parts, in the 6km range from our center point. Using the Foursquare, we have identified the density of restaurants in general and also, Italian restaurants.

Second step in our analysis will be calculation and exploration of 'restaurant density' across different areas of Paris. We will use heatmaps to identify a few promising areas close to center with low number of restaurants in general (and hopefully no Italian restaurants nearby) and focus our attention on those areas.

In third and final step we will focus on most promising areas that we have discovered in the previous step 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 four restaurants in radius of 250 meters, and we want locations without Italian restaurants in radius of 200 meters. We will present map of all such locations but also create clusters (using k-means clustering) of those locations to identify general zones / neighborhoods / addresses which should be a starting point for final 'street level' exploration and search for optimal venue location by stakeholders.

Results

First we collected information about the exact location of the Paris. After that we have sliced the Paris into little neighborhoods, 364 of them to be precise, so we could easily identify the best spots for the restaurant. The popup represents the Eiffel Tower.



Next thing was getting the information about address in that area.

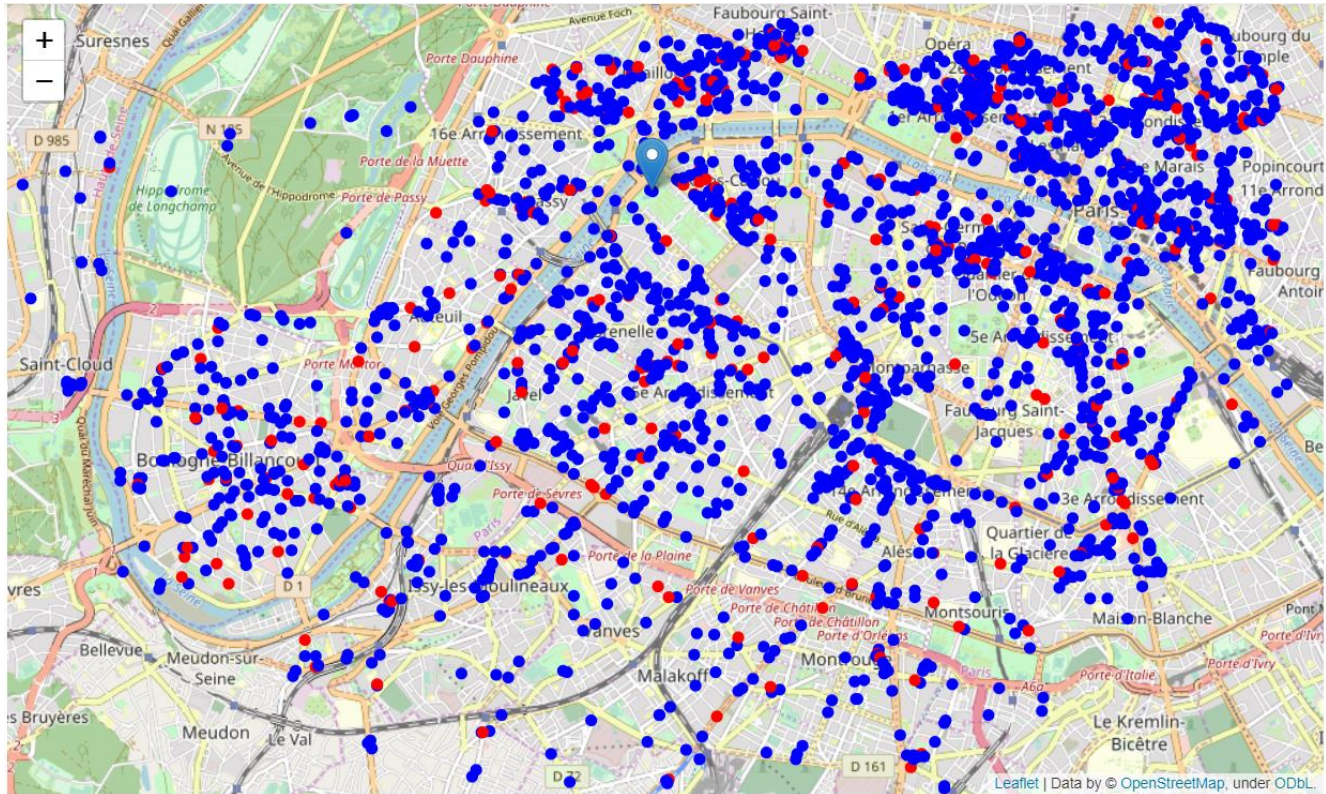
We have done that by using the Foursquare API and placed that data inside a data frame so we could easily manipulate with it. We have converted latitude and longitude to x and y values so we could count the distances from the city center. Those are the last three columns we have additionally added.

	Address	Latitude	Longitude	X	Y	Distance from center
0	France	48.805421	2.283475	-432670.16604	5.484312e+06	5992.495307
1	Avenue De Paris	48.806315	2.291446	-432070.16604	5.484312e+06	5840.376700
2	France	48.807208	2.299417	-431470.16604	5.484312e+06	5747.173218
3	45 avenue Marx Dormoy	48.808101	2.307389	-430870.16604	5.484312e+06	5715.767665
4	35 Rue Molière	48.808993	2.315361	-430270.16604	5.484312e+06	5747.173218
...
359	30 Rue Chanzy	48.907520	2.273587	-431470.16604	5.495744e+06	5747.173218
360	Rue de la Station	48.908416	2.281574	-430870.16604	5.495744e+06	5715.767665
361	France	48.909312	2.289561	-430270.16604	5.495744e+06	5747.173218
362	Place Marguerite Durand	48.910206	2.297549	-429670.16604	5.495744e+06	5840.376700
363	150 Boulevard du Général Leclerc (Allée de l'E...	48.911101	2.305537	-429070.16604	5.495744e+06	5992.495307

After we got all the data we need about addresses around the Paris, next thing was locating the restaurants on those locations. Again, we have used the Foursquare API in order to retrieve information about their positions. First thing we observed was how many restaurants there are in general and how many of them were Italian.

```
Total number of restaurants: 2666
Total number of Italian restaurants: 324
Percentage of Italian restaurants: 12.15%
Average number of restaurants in neighborhood: 6.403846153846154
```

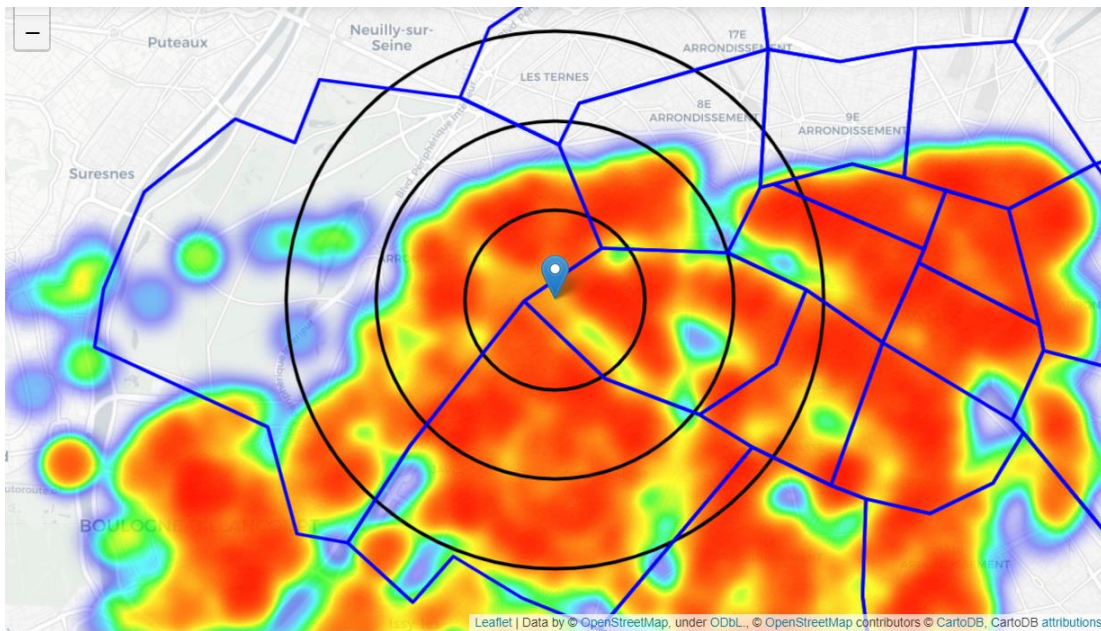

In the next step, we have presented all restaurants in that area on the map. For the Italian restaurants, we used the red dot and for the other restaurants, we have used the blue color.



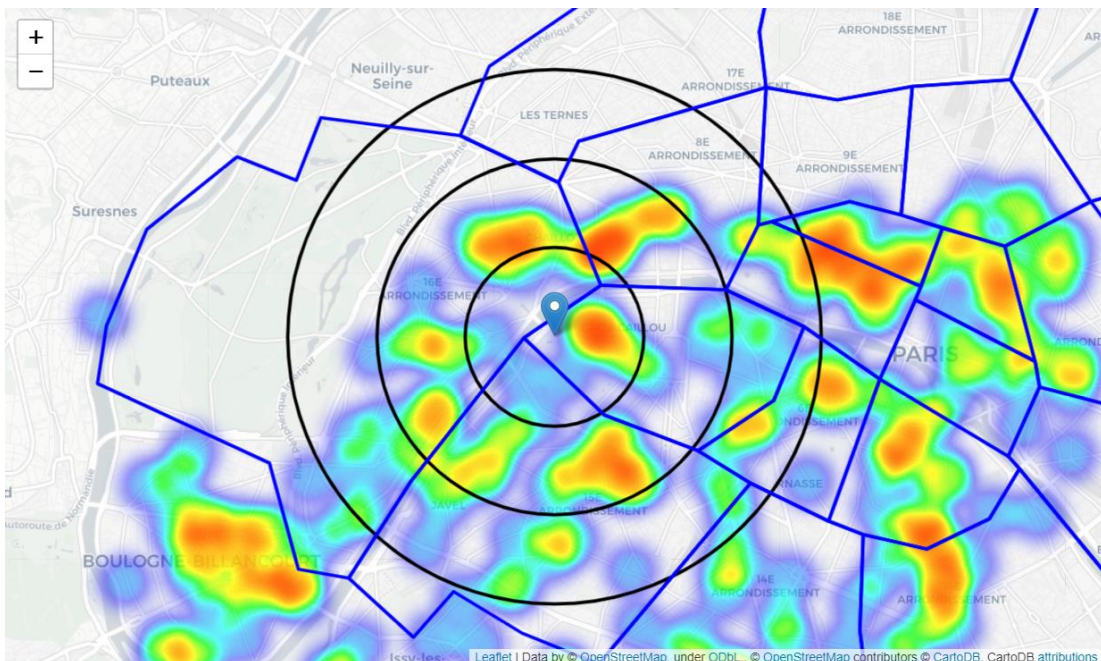
Now, we needed to see how many restaurants are in each one of the areas that we have created earlier. In the next data frame, beside all the basic information about addresses, we have added the number of restaurants in that area.

	Address	Latitude	Longitude	X	Y	Distance from center	Restaurants in area
0	France	48.805421	2.283475	-432670.16604	5.484312e+06	5992.495307	1
1	Avenue De Paris	48.806315	2.291446	-432070.16604	5.484312e+06	5840.376700	1
2	France	48.807208	2.299417	-431470.16604	5.484312e+06	5747.173218	3
3	45 avenue Marx Dormoy	48.808101	2.307389	-430870.16604	5.484312e+06	5715.767665	0
4	35 Rue Molière	48.808993	2.315361	-430270.16604	5.484312e+06	5747.173218	3
5	60 avenue Marx Dormoy	48.809885	2.323333	-429670.16604	5.484312e+06	5840.376700	1
6	C.C La Vache Noire (Place de la Vache Noire)	48.810776	2.331306	-429070.16604	5.484312e+06	5992.495307	9
7	92140 Clamart	48.808639	2.270346	-433570.16604	5.484832e+06	5855.766389	2
8	France	48.809534	2.278317	-432970.16604	5.484832e+06	5604.462508	1
9	France	48.810428	2.286288	-432370.16604	5.484832e+06	5408.326913	1

Now we have created a heatmap / density of restaurants and tried to extract some meaningful information from that. Also, we have showed borders of Paris boroughs on our map and a few circles indicating distance of 1km, 2km and 3km from Eiffel Tower.



So, now we see the heatmap or the density of all the restaurants in the desired area. The next thing we wanted to explore was the density of only Italian restaurants.



This map is not so 'hot' (Italian restaurants represent a subset of ~12% of all restaurants in Paris) but it also indicates higher density of existing Italian restaurants directly east and north from Eiffel Tower, with closest pockets of low Italian restaurant density positioned west, south-west and south from city center.

Based on this we will now focus our analysis on areas **south-west and south from Paris center** - we will move the center of our area of interest and reduce it's size to have a radius of **2.5km**. This places our location candidates mostly in boroughs **Paris 15e Arrondissement, Paris 16e Arrondissement and Paris 7e Arrondissement**. The Paris 15e Arrondissement is not very interesting to our stakeholders since it's mostly a residential zone. On the other hand the Paris 16e Arrondissement is a much better choice since it's nearer to the Seine and Eiffel Tower. Another interesting area is Paris 7e Arrondissement, but it's a little bit crowded. So, the best choice would be **Paris 16e Arrondissement**, and we will proceed analyzing that borough in the further sections.

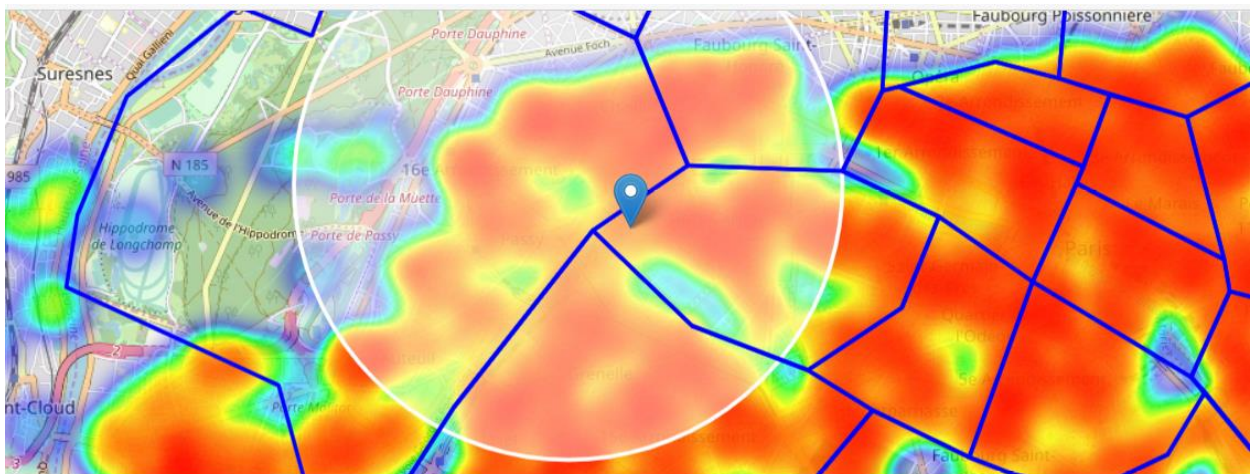
Paris 16e Arrondissement is a large district that occupies most of the West of Paris, extending east-west between the bends of the Seine from the Jardins of Trocadero immediately facing the Eiffel Tower to the expansive Bois du Boulogne (which occupies a larger part of the 16th's territory), and north-south from the Etoile to the southern border of Paris. It is known to be the residence of choice for affluent Parisians, and for hosting numerous internationally famous events, such as the Roland Garros French Open tennis tournaments, as well as the home stadium of the Paris Saint-Germain football club.

Some of the reviews from TripAdvisor:

“It is one of the richest arrondissements of Paris and this is easy to guess from the many beautiful residential buildings, nice restaurants and shops. Except some very touristic areas (like Trocadero) it is less crowded by tourists. Also, the Bois de Boulogne and many parks, lakes, sports fields are in the same district.”

“The area we visited in the 16th Arrondissement had the best views of the Eiffel Tower, that is not situated in the Seine. As we ventured further along though, there wasn’t much to do and the people there weren’t too friendly. We don’t have time for stuck up people so we left.”

Now we have created a white circle around the area that we are interested in so that we could concentrate on important parts of the city for us.



Okay, now we will calculate two most important things for each location candidate: number of restaurants in vicinity (we'll use radius of 250 meters). The collected data will be placed inside a dataframe for easier manipulating later.

	Latitude	Longitude	X	Y	Restaurants nearby	Distance to Italian restaurant
0	48.849770	2.259063	-433620.16604	5.489528e+06	8	366.525137
1	48.849919	2.260392	-433520.16604	5.489528e+06	7	288.198547
2	48.850069	2.261722	-433420.16604	5.489528e+06	5	227.543686
3	48.850218	2.263051	-433320.16604	5.489528e+06	5	201.230852
4	48.850367	2.264381	-433220.16604	5.489528e+06	4	221.836838

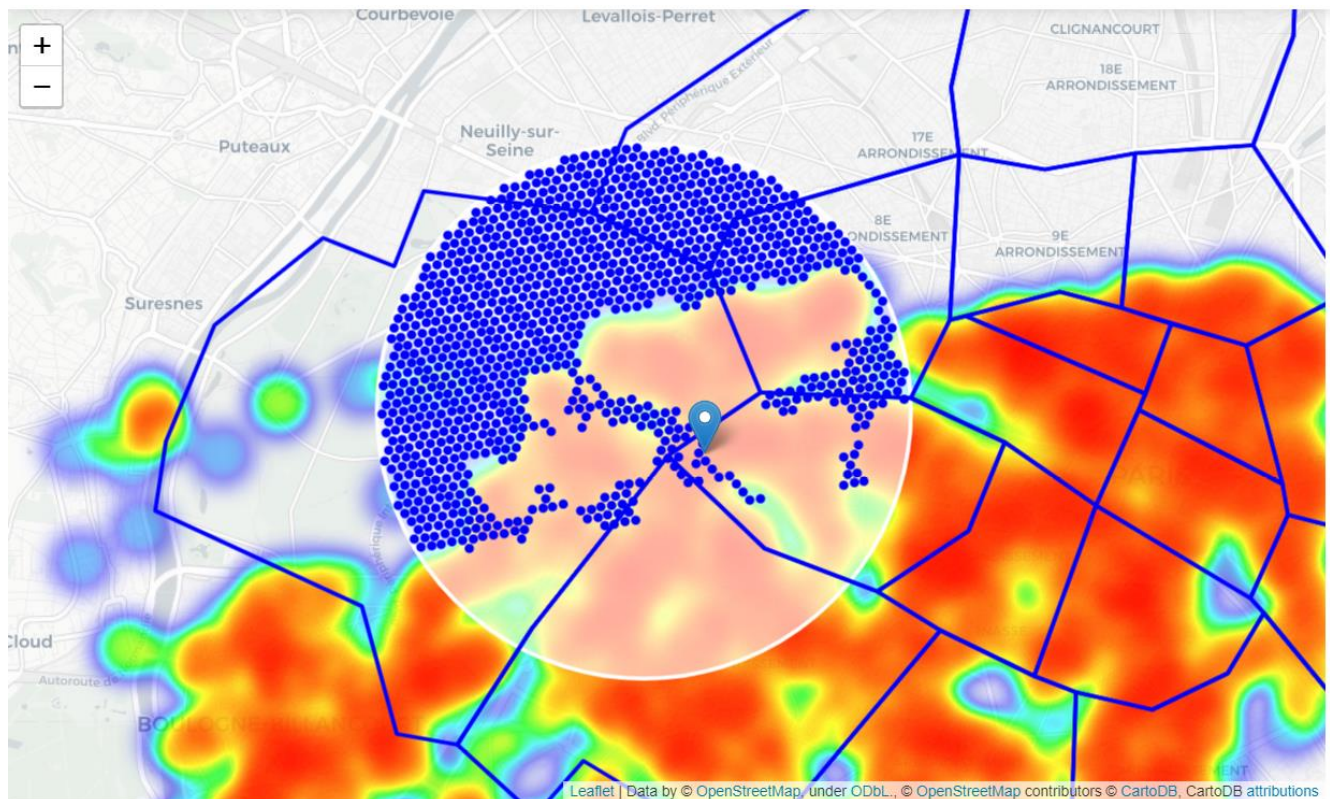
After we have obtained information about restaurants that are nearby, we have filtered those locations by criteria: we're interested only in locations with no more than four restaurants in radius of 250 meters, and no Italian restaurants in radius of 200 meters. We have gained information about number of restaurants that fullfil either one of criteria, but we are the most interested in the ones that fill out both.

Locations with no more than four restaurants nearby: 1085

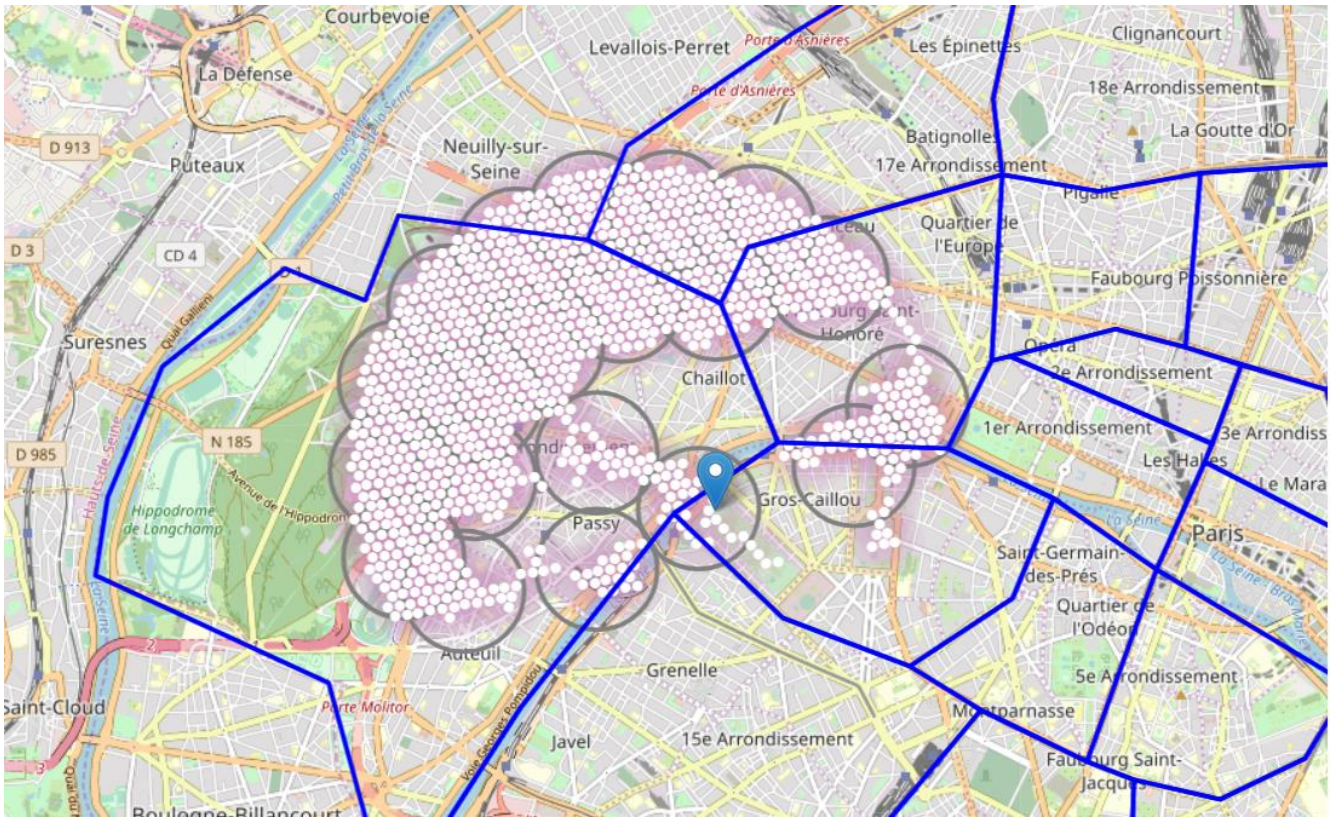
Locations with no Italian restaurants within 200m: 1287

Locations with both conditions met: 1063

What we did next is that we have created another heatmap in order to show the locations that fullfil both of the criteria.



After that we have clusteres those locations in order to create a center for each zone that contains “good location”. Those zones, their centers and addresses will be the final result of our analysis. The popup is the location of the Eiffel Tower.



This concludes our analysis. We have created 20 addresses, shown below, representing centers of zones containing locations with low number of restaurants and no Italian restaurants nearby, all zones being fairly close to city center (all less than 4km from Eiffel Tower, and about half of those less than 2km from Eiffel Tower). Although zones are shown on map with a radius of ~500 meters (grey circles), their shape is actually very irregular and their centers/addresses should be considered only as a starting point for exploring area neighborhoods in search for potential restaurant locations. Most of the zones are located in **Paris 16e Arrondissement** boroughs, as we have mostly focused on that area, which we have identified as interesting due to being popular with tourists, fairly close to city center and well connected by public transport.

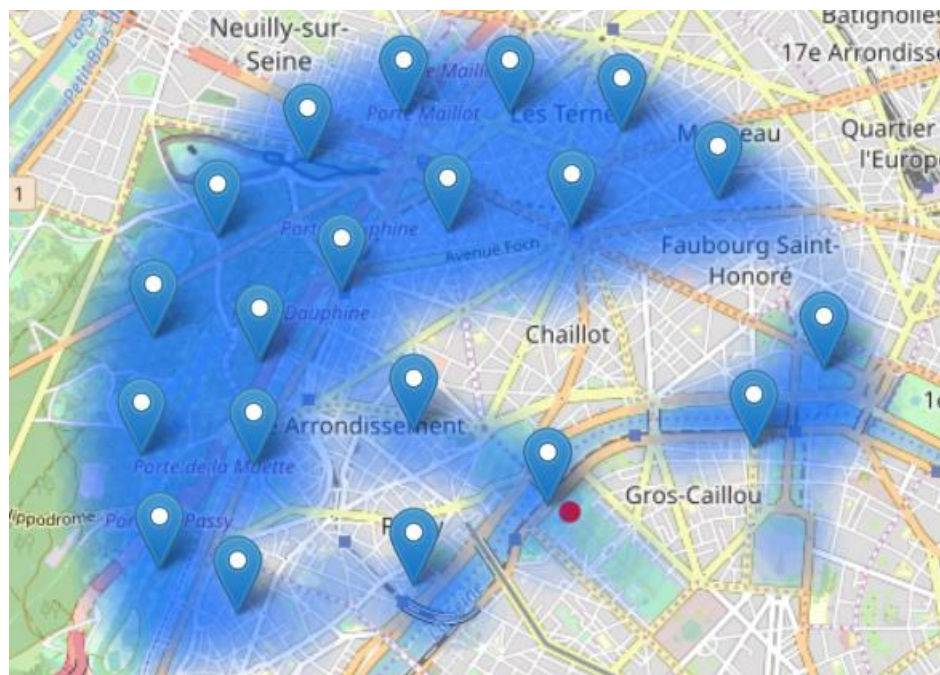
=====

Addresses of centers of areas recommended for further analysis

=====

57 boulevard de Montmorency (Bois de Boulogne)	=> 2.6km from Eiffel Tower
France	=> 1.9km from Eiffel Tower
8 avenue Dutuit	=> 1.8km from Eiffel Tower
Chemin de ceinture du Lac Inférieur	=> 2.9km from Eiffel Tower
58 avenue de Wagram	=> 2.4km from Eiffel Tower
5 avenue Anatole France (Tour Eiffel, 1er étage)	=> 0.1km from Eiffel Tower
14 rue de Monceau	=> 2.2km from Eiffel Tower
14 Boulevard de la Tour, Maubourg	=> 1.2km from Eiffel Tower
21 rue de Sablonville (Rue Montrosier)	=> 2.7km from Eiffel Tower
1 boulevard Lannes	=> 2.0km from Eiffel Tower
22 rue Cortambert	=> 1.1km from Eiffel Tower
Porte de la Muette	=> 2.0km from Eiffel Tower
Place Charles de Gaulle (Avenue des Champs-Élysées)	=> 1.8km from Eiffel Tower
8 avenue du Mahatma Gandhi	=> 2.8km from Eiffel Tower
Avenue du Mahatma Gandhi (Bois de Boulogne)	=> 2.8km from Eiffel Tower
Bois de Boulogne	=> 2.8km from Eiffel Tower
France	=> 2.5km from Eiffel Tower
Maison de la Radio (116 avenue du Président Kennedy)	=> 1.1km from Eiffel Tower
10 Square du Dr Blanche	=> 2.2km from Eiffel Tower
Chemin de ceinture du Lac Inférieur	=> 2.2km from Eiffel Tower

The map below shows the selected 20 locations that would be best for opening an Italian restaurant in Paris.



Discussion

Our analysis shows that although there is a great number of restaurants in Paris (~2000 in our initial area of interest which was 12x12km around Eiffel Tower), there are pockets of low restaurant density fairly close to city center. Highest concentration of restaurants was detected east and south-east from Eiffel Tower, so we focused our attention to areas with closest pockets of low Italian restaurant density positioned west, north-west and south from city center. That area was especially ****Paris 16e Arrondissement****. We have decided to go for that one since it's the closest one to the city center and one of the biggest in Paris. It has lots of tourists and good transport. We have been able to identify 20 addresses that would be good for opening an Italian restaurant.

After directing our attention to this more narrow area of interest (covering approx. 5x5km west, north-west and south from Eiffel Tower) we first created a dense grid of location candidates (spaced 100m appart); those locations were then filtered so that those with more than four restaurants in radius of 250m and those with an Italian restaurant closer than 200m were removed.

Those location candidates were then clustered to create zones of interest which contain greatest number of location candidates. Addresses of centers of those zones were also generated using reverse geocoding to be used as markers/starting points for more detailed local analysis based on other factors.

Result of all this is 20 zones containing largest number of potential new restaurant locations based on number of and distance to existing venues - both restaurants in general and Italian restaurants particularly. This, of course, does not imply that those zones are actually optimal locations for a new restaurant! Purpose of this analysis was to only provide info on areas close to Paris center but not crowded with existing restaurants (particularly Italian) - it is entirely possible that there is a very good reason for small number of restaurants in any of those areas, reasons which would make them unsuitable for a new restaurant regardless of lack of competition in the area. Recommended zones should therefore be considered only as a starting point for more detailed analysis which could eventually result in location which has not only no nearby competition but also other factors taken into account and all other relevant conditions met.

Conclusion

Main goal of this project was to identify best places for opening a new Italian restaurant in a city that already has a huge amount of restaurants. Throughout this project we were exploring best possible values and finally we came to a conclusion and found 20 places that filled all the requirements. We have split the Paris into boroughs to easily identify the boroughs that fill the requirements. After that, we calculated the density of restaurants in general and Italian restaurants using the Foursquare API in those areas and based on that we have chosen the borough that was most promising. 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 centers were created to be used as starting points for final exploration by stakeholders.

Final decision on optimal restaurant location will be made by stakeholders based on specific characteristics of neighborhoods 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 neighborhood etc.

Acknowledgement

I would like to thank everyone of the professor and people who helped me during this journey. I've learned quite a lot and I'm looking forward to applying it in the real life projects.