

COURSERA CAPSTONE : BAKERY IN PARIS REGION

FRANCOIS LEDEIN

31 JULY, 2020

1 INTRODUCTION

1.1 CONTEXT AND BACKGROUND

Paris is the capital city of France, and the biggest town of the country. It has 12 millions inhabitants in the city and suburbs, in an area of around 17 000 km². This city is the economic center of France, and a very touristic region.

Bakeries are typically french, so there is a lot of bakeries in Paris, or other venues that sell bread and pastries, like traditionnal bakeries. Set up a new bakery is a big investment in money and time, and should be done in a suitable place, to ensure a good revenue .

1.2 BUSINESS PROBLEM

Data can help find candidates places for a new bakery in Paris region. This includes the location of currently existing bakeries or bread and pastries sellers. This project aims to find some candidates locations to establish a new bakery.

1.3 INTEREST

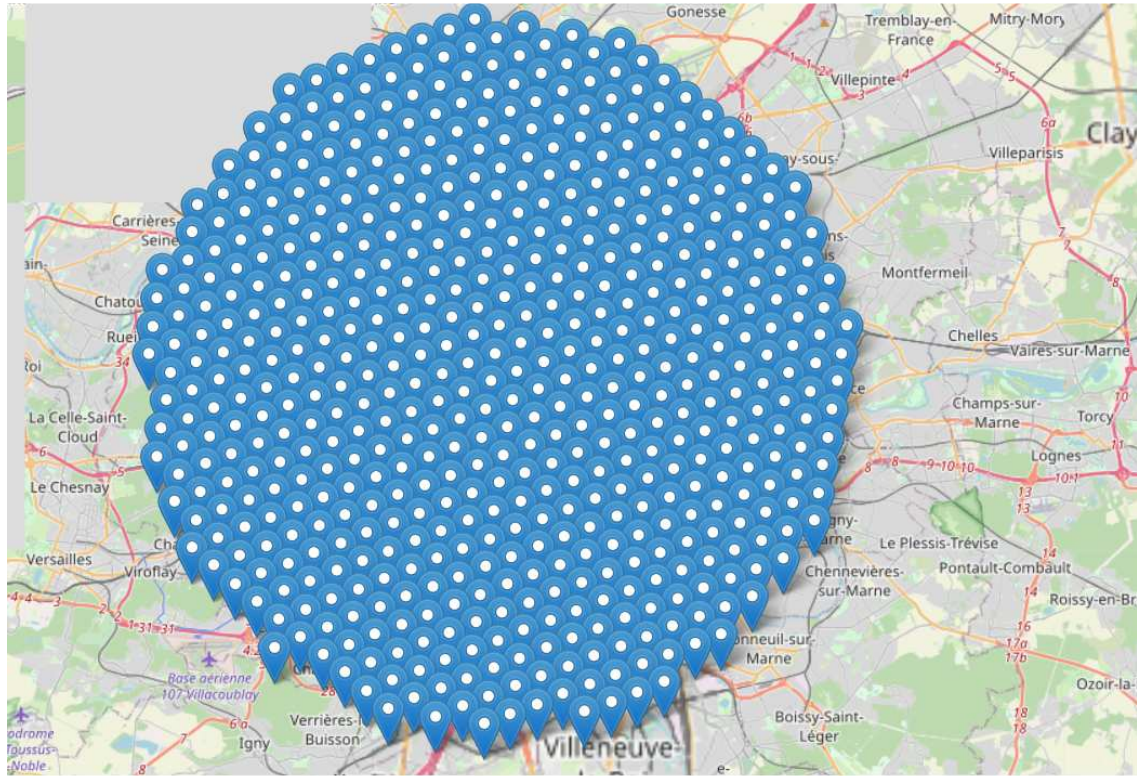
This project can help investors, to help them find a suitable locatino for their project. It can helps future managers of the nex bakery, to decide wether or not to go. Public managers (mayors) can e interested for example to convince stakeholder to open a new bakery in their town, and then improve local shops.

2 DATA ACQUISITION AND CLEANING

2.1 DATA SOURCES

The Goal is to obtain a list of all bakeries in Paris Region, with their locations ans adress. We will use Foursquare API to achieve this.

Forusquare can give up to 50 venues for a unique API call, so we need to make multiple API call. To do that we will first create a grid of points, regularly spaced by 1200 m, in a radius of 15 km from Paris Center.



Then for each of these point, call Foursquare API to search for Bakeries.

2.2 DATA CLEANING

Duplicates are stored in a dictionary by foursquare ID, to avoid duplicates.

Each venue is a bakery, because we search foursquare by category.

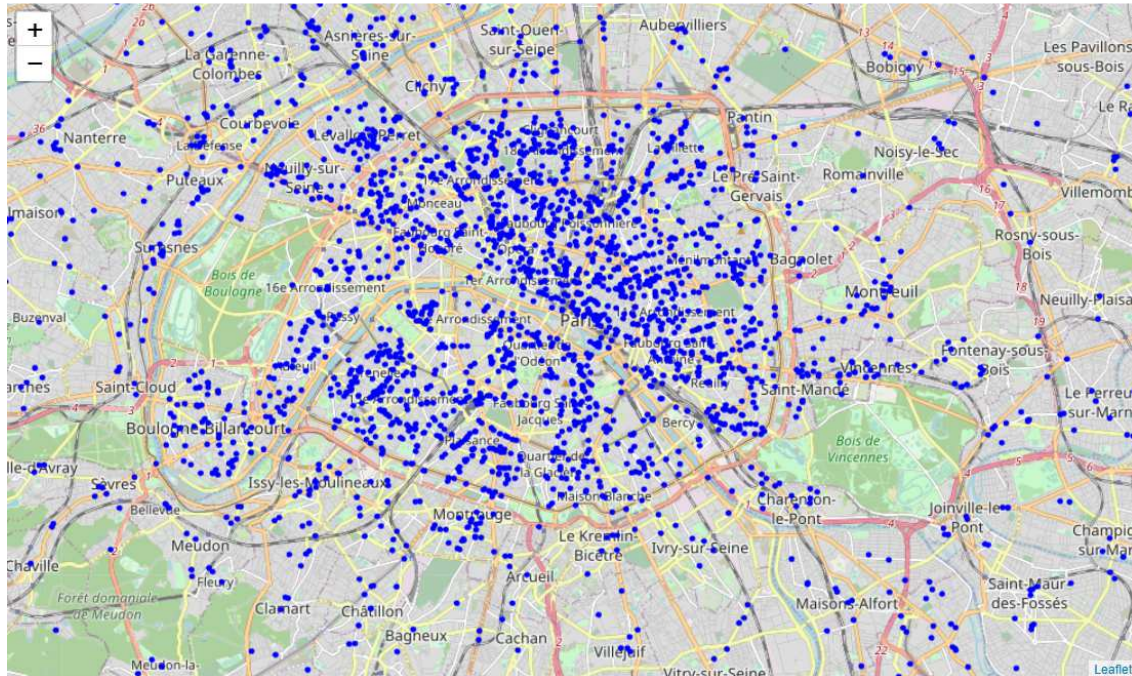
Each venue has a location and a name.

Some venues don't have an adress. So we choose not to use the foursquare adress and drop the data.

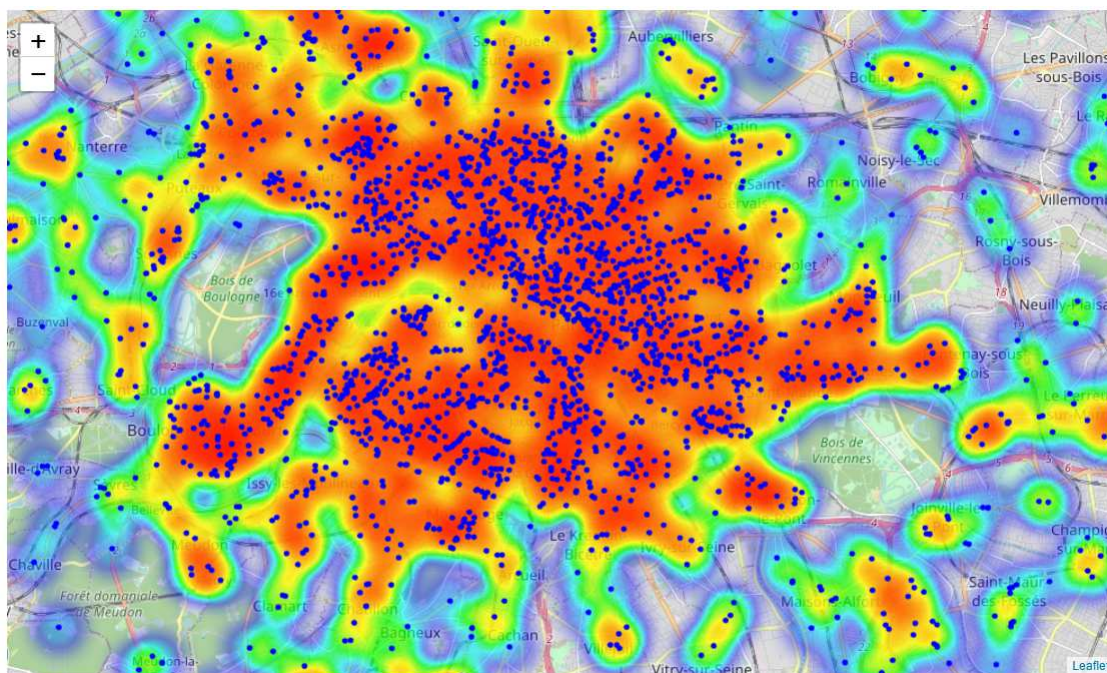
3 EXPLORATORY DATA ANALYSIS

3.1 MAP OF THE BAKERIES.

We put all the results on a Paris map.



And showing with a heatmap the distribution of bakeries.



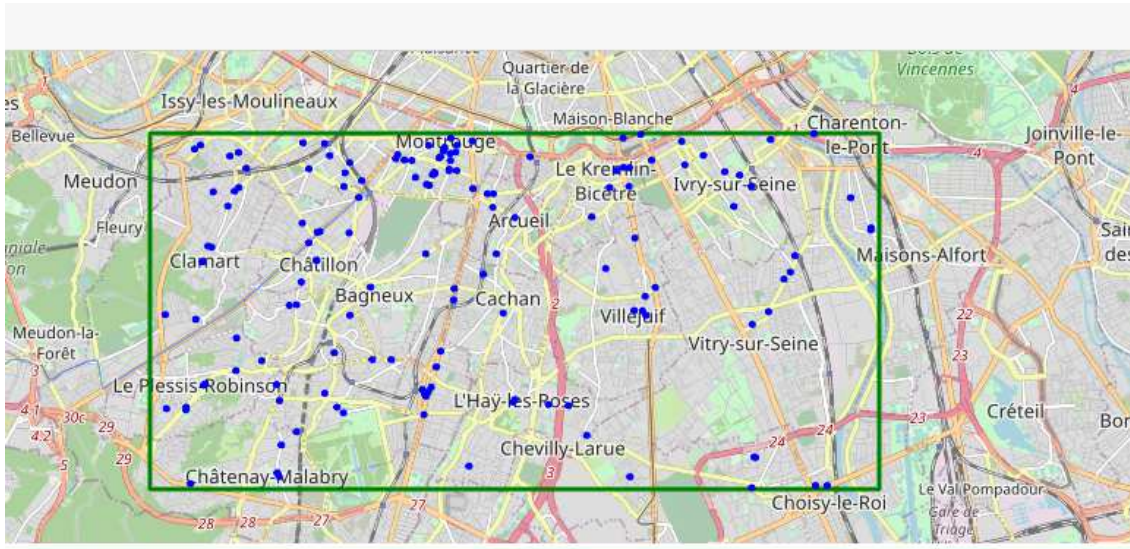
As a result, we can see that most of the bakeries are in Paris City.

There are also regions with a lot of bakeries in the east suburb, in west and south-west suburb (Boulogne City).

The north-east region and south region seem to have less bakeries.

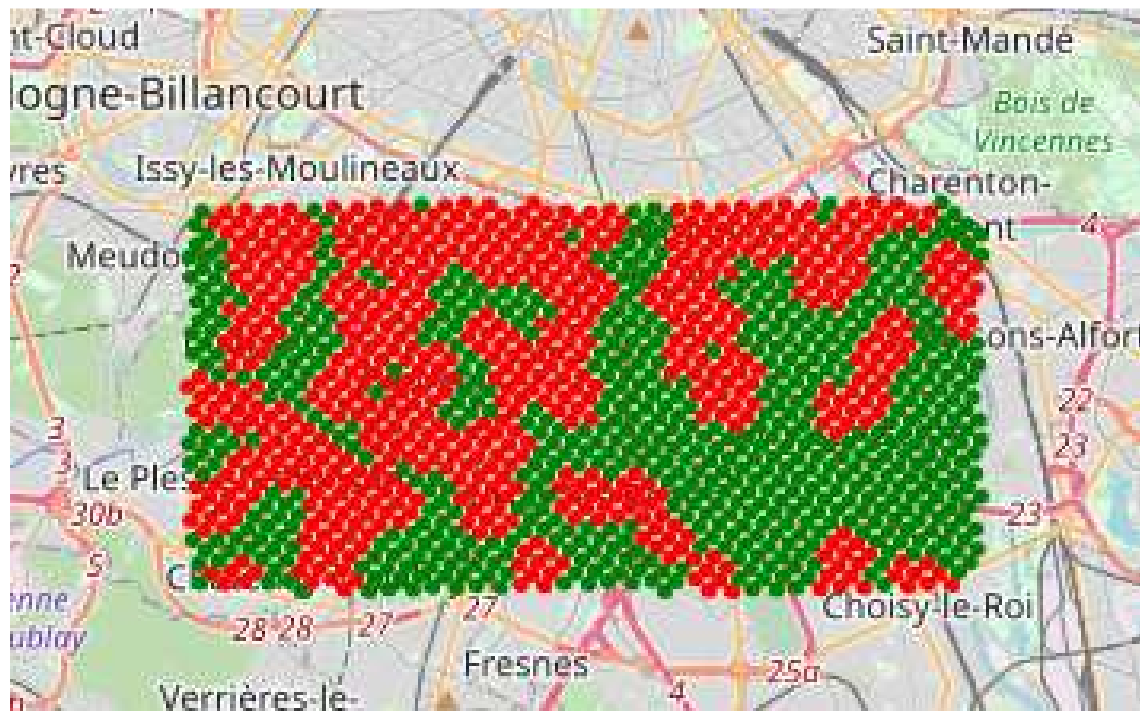
3.2 ZONE RESTRICTION

We choose to focus on south suburb : in the two regions with less bakeries, south is richer than north east suburb. Here is the restricted zone with existing bakeries :



3.3 REFINED ANALYSIS

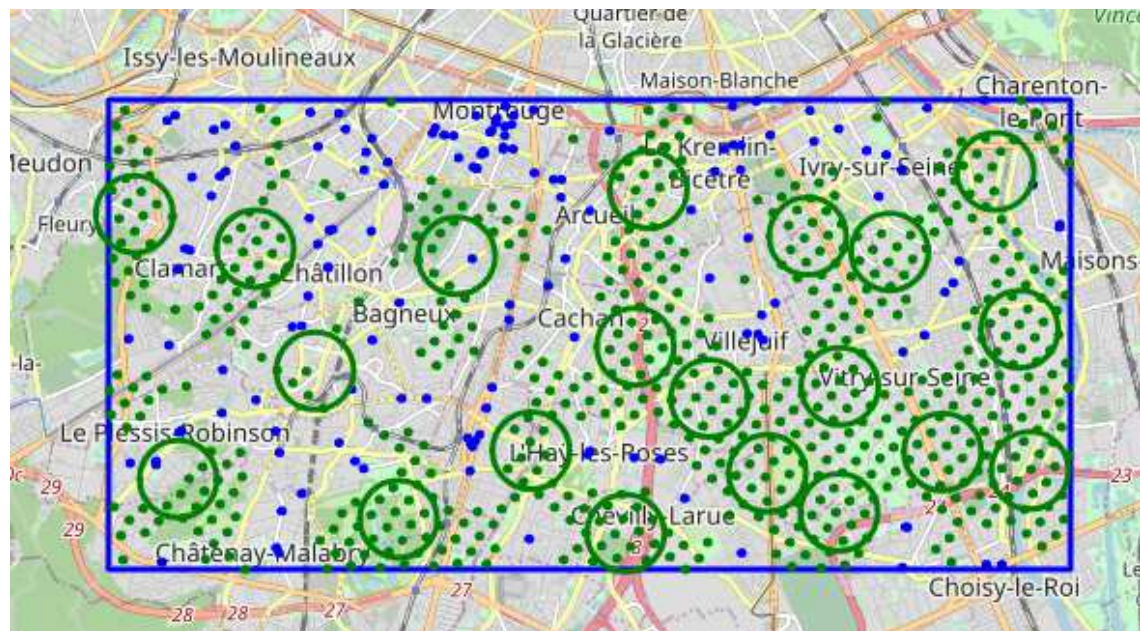
In the restricted region, we create a new grid of points, regularly spaced by 300m. For each of these points we calculate the nearest bakery, and then filter the points with no bakeries within a 500m radius. The result is this one : red points with a bakery within the 500m, Green one when there is no bakery within 500m. We can identify some regions that are interesting.



4 MODELING

Using only the green points of the previous map, The goal is to identify some addresses to look further at.

We use k-mean clustering algorithm, from scikit learn library. 20 clusters are identified.



Each cluster as a center, with latitude and longitude. We use Google map API to reverse geocode the adress of the centers. We obtain a list of 20 adresses that fits all our criteria : in the south suburb, and with no bakery within 500 m.

	latitude	longitude	adresses
0	48.791115	2.343770	148 Rue Gabriel Péri, 94230 L'Haÿ-les-Roses, F...
1	48.802196	2.388969	16 Rue Jean-Baptiste Renoult, 94200 Ivry-sur-S...
2	48.806606	2.254747	43 Avenue Adolphe Schneider, 92140 Clamart, Fr...
3	48.776600	2.413875	Seine Bridge, A86, 94600 Choisy-le-Roi, France
4	48.788246	2.286838	14 Rue André Neyts, 92260 Fontenay-aux-Roses, ...
5	48.771688	2.379975	14 Rue Jean Mermoz, 94320 Thiais, France
6	48.809479	2.346087	30 Rue Labourse, 94250 Gentilly, France
7	48.801727	2.311990	29 Rue de Verdun, 92220 Bagneux, France
8	48.775592	2.262536	47 Rue du Moulin Fidel, 92350 Le Plessis-Robin...

5 CONCLUSION

Using data science tools, we were able to locate all existing bakeries in Paris area. Some vizualization shox us that east suburb ans south suburb are less crowded with bakeries than other areas. We chose to focus on south suburb. In the chosen region, We determine 20 adresses without any bakerie within 500m.

Theses adresses are candidates for a new installation of a bakery. There may be a very good reason why there is no bakery in the area, or example there is a park or forest, an hospital.

The location study has to continue by going inplace, on the adress and look at the neighborhood attractiveness : population, existing venues, proximity of a road or a railwaxy station, real estate prices, etc...

The aim of the project was to determine addresses suitable to open a new bakery. This goal is achieved.

6 FUTURE ?

Regarding the business problem, the 20 adresses determines with data science have to be furthered analysed with other criteria : attractiveness of the neighborhood, population, real estate prices..

Regarding the data science project, new criteria can be add to the process. For example excluding reserved area like cimeteris, forests, hospitals. Second example we can add the population of each neighborhood, to give locations where there is less bakeries by 10 000 inhabitants than other locations.