# Battle of neighborhoods - Report

Alex Ye

August 9, 2019

## 1 Introduction

Due to a recent raise of vandalism in Paris, the police department of Paris
contracted us to make a report on the situation of Paris. To address that we'll
first find a list of the highest vandalism rate neighborhood and attempt to find
a correlation between vandalism and the vicinity characteristics.

## 2 Data

For that we'll use :

- Foursquare API

- Paris anomalie dataset

- Paris arrondissement

I take only last 3 years of the anomalies dataset which is constituted of signals
from diverse parisian citizen.
Assuming all the venues stayed the same for that period, since Foursquare only
gives us info about the venues currently in place. Or there is an option but i
am not aware.
The anomalies dataset is composed of the following 10 categories of signals.
Let's explain a bit each categories to give more context. Categories of interest
are indicated with a star.

- Graffitis, tags, affiches et autocollants: public vandalism *

- Voirie et espace public: Street deterioration

- Mobiliers urbains: urban furniture deterioration *

- Propreté: Piss on the street *

- Éclairage / Électricité: malfunctions of public lights

- Objets abandonnés: cumbersome objects left on the street

- Arbres, végétaux et animaux: dangerous tree or presence of rats, maintenance issue

- Eau: flood, water related issues

- Autos, motos, vélos...: abandonned/wreck vehicles *

- Activités commerciales et professionnelles: flyers with fraudulous use of Paris' logo or colors

```
Graffitis, tags, affiches et autocollants     77214
Autos, motos, vélos...                          9447
Mobiliers urbains                               9316
Propreté                                        5071
```

Figure 1: Number of each type of signal considered
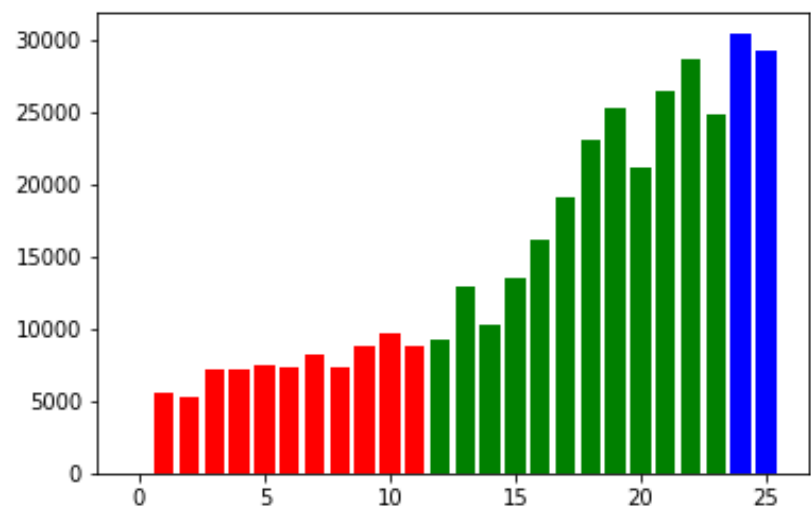


Figure 2: Signals per month, years in different colors
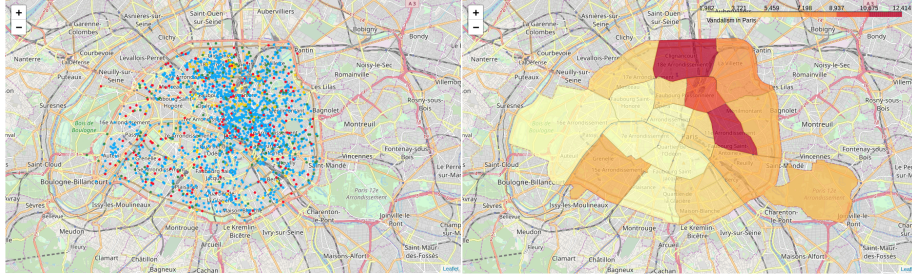
If we plot these data on a map

Figure 3: All signals



Figure 4: All signals by borough

As we can see most of the signal comes from northern and eastern side of Paris.

# 3    Methodology

We know there's a disparity of vandalism in the distribution between borough. Is it possible to distinguish groups out of this ? We'll cluster all signals into k demographic, then according to the type of vandalism, we'll then try to deduce who they are and what are their motivation.
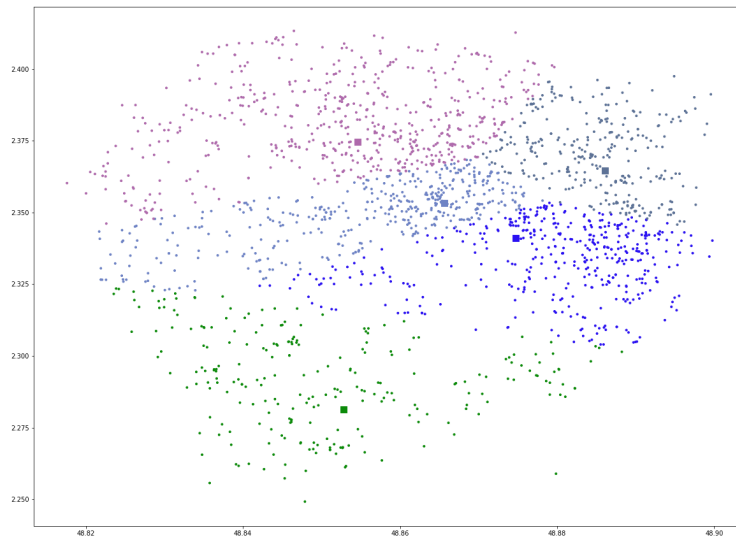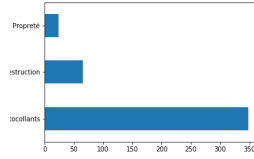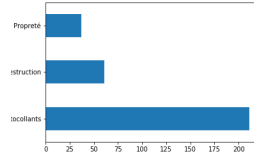


Figure 5: clusters k = 5
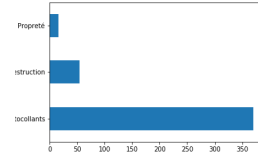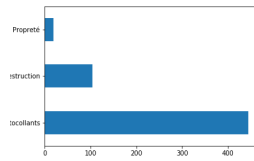
Figure 6: A



Figure 7: B
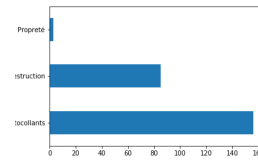


Figure 8: C



Figure 9: D



Figure 10: E

The proportions doesn't allow us to draw clear geospatial limits as there's little difference between clusters. We can't say the type of crime is geographically distinct. Where vandalism happen, it happens in about the same proportion. The categories "Autos, motos, vélos..." and "Mobiliers urbains" seems very close. Let's look if they are correlated, if it's the case we can group them as one class.
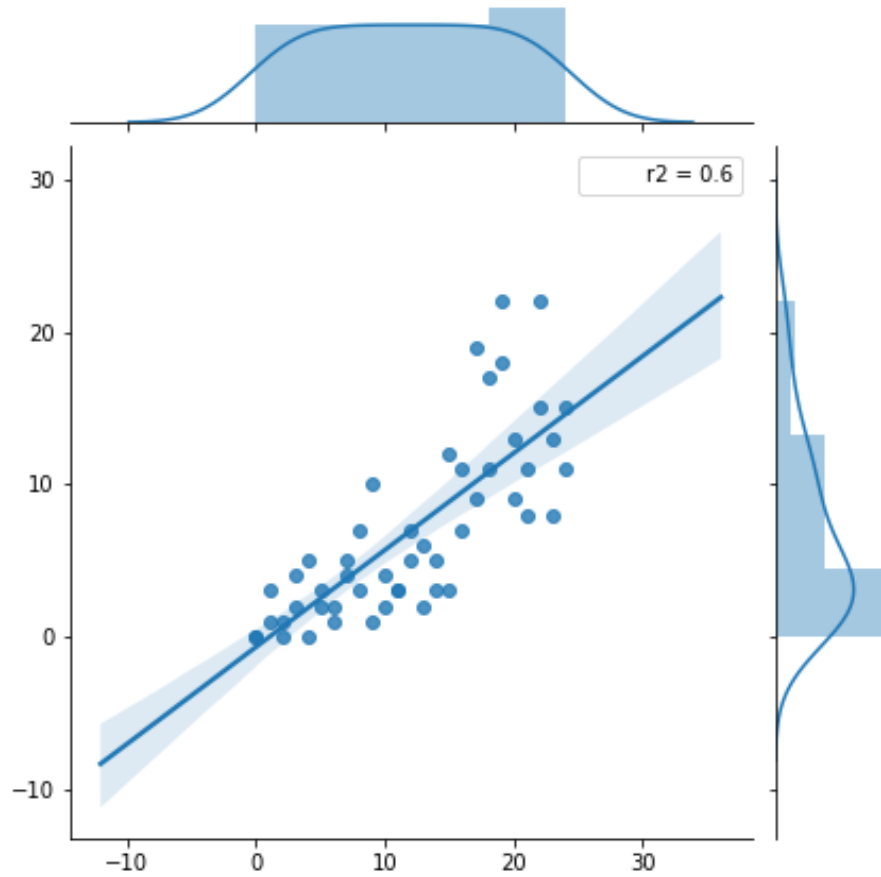
Figure 11: regplot

R2 = 0.6 is a somewhat strong correlation, we'll group these two in one category, as both describe destruction of vehicles or furnitures, let's name the group 'Destruction'.
So we end up with 3 categories

- Graffitis, tags, affiches et autocollants

- Propreté

- Destruction

We can also confirm that all types of vandalism are raising on about the same scale.
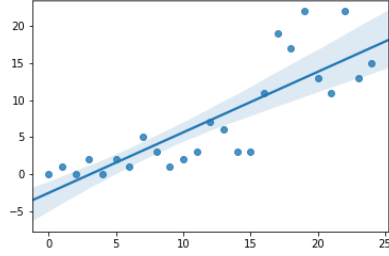
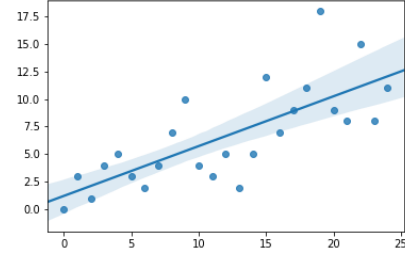Figure 12: Autos, motos, vélos... / Destruction



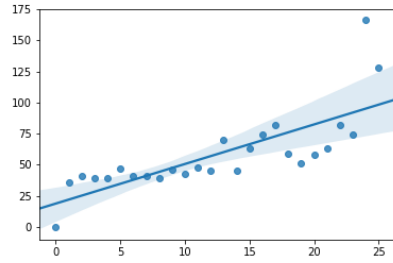Figure 13: Mobiliers urbains / Destruction



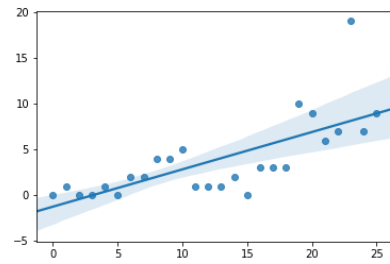Figure 14: Graffitis, tags, affiches et autocollants



Figure 15: Propreté

For each location we'll make a search with Foursquare API and find the venues in vicinity. What are these venues, what logical conclusion could this finding lead us to ?
We found that the most common venue near (within 15 meters) each and every types of signals are 'French Restaurant'

```
[('French Restaurant', 4005),
 ('Bar', 3106),
 ('Café', 1968),
 ('Bakery', 1860),
 ('Hotel', 1728)]
```

Figure 16: Top 5 most frequent venues near signal

There's no significant difference between categories of signals. The top 2 venues never change, the third to fifth can switch places with a slight difference. We'll first look at the top venue 'French Restaurant'.
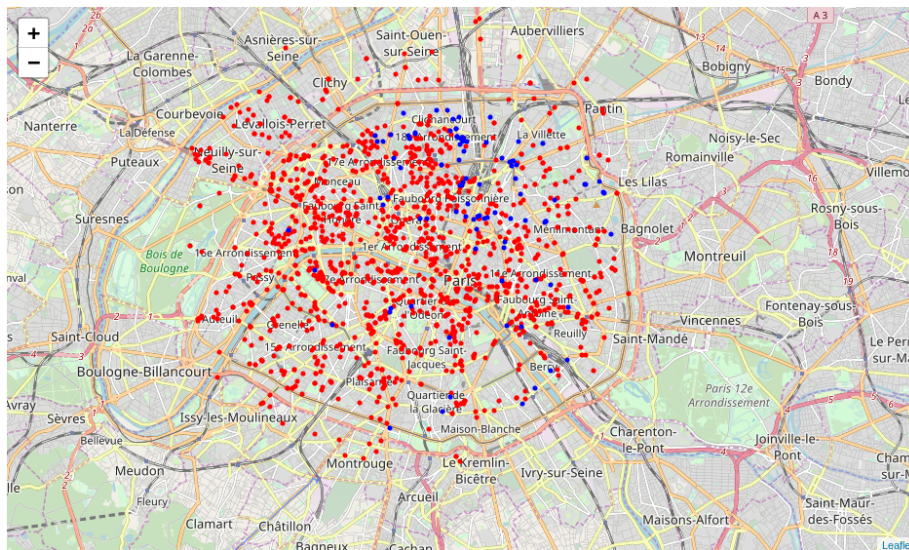Let's plot these data on the map and see what it can tell us.
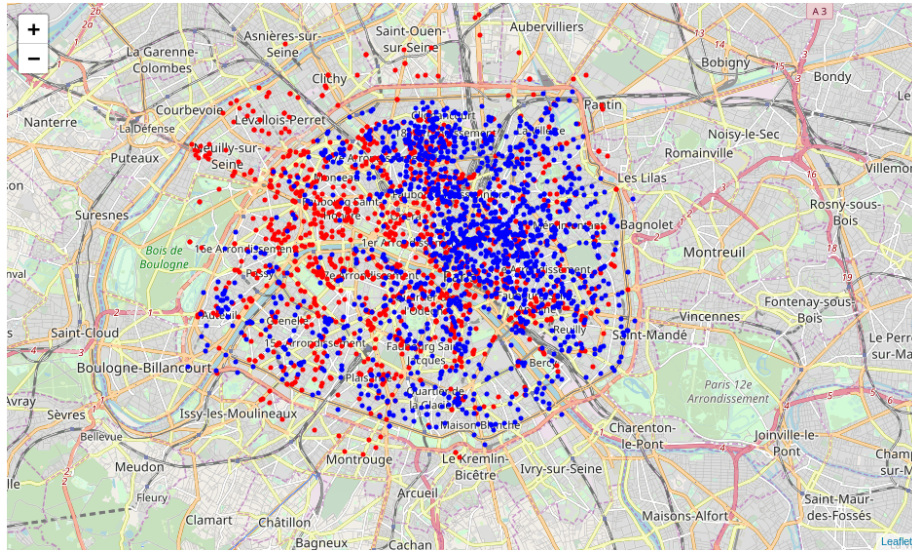


Figure 17: Blue = "Propreté", Red = 'French Restaurant"

Figure 18: Blue = "Graffitis, tags, affiches et autocollants", Red = 'French Restaurant"
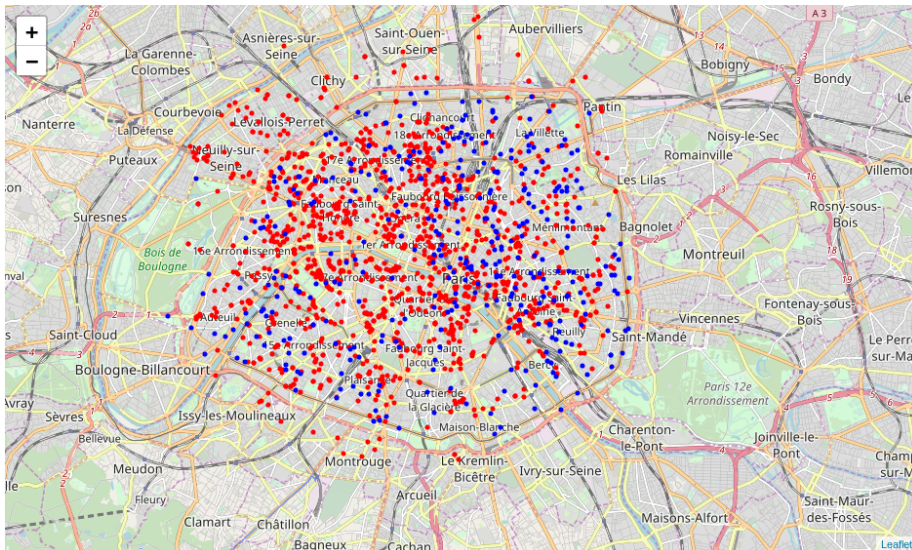


Figure 19: Blue = "Destruction", Red = 'French Restaurant"

This doesn't look promising, visually speaking the French Restaurants are less present in the north east area. There's most likely no correlation between the two, the result was likely due to high density of 'French Restaurant' in Paris, pointed out by the other most common venues in the top 5.

8

A last idea to explore, July 2018 has seen the world championship of football, French became the world champions. We'd like to look at that as it is said that football event often lead stores to close due to high risk of damage to establishments.
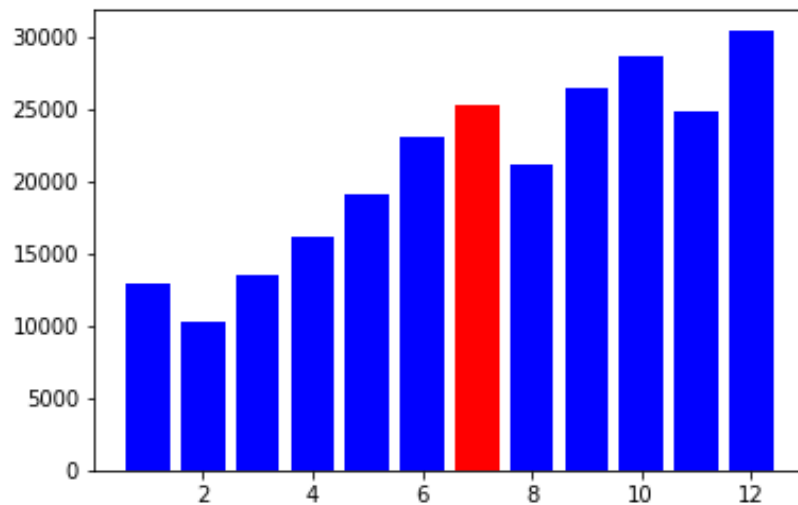


Figure 20: Signals per month during 2018

As we can see the football competition did actually raise the bar, for the summer period at least. Then declined right after, until another thing takes place around September.
We can easily assume it's the strong dissatisfaction toward Macron's policies (yellow jackets, black blocks and Benalla's doings). That would explain the sudden raise of vandalism.

# 4    Results

First of all, we can say that vandalism from all kind have been raising steadily and they are more common in the north-east areas of Paris.
The geospatial analysis can't explain the recent raise. But socially speaking it is pretty clear that it was majorly due to social event happening.

# 5    Conclusion

By the numbers we've got, we could tend to say as French restaurant are always near vandalized spots, so French Restaurant must attract vandals, therefore gen-

erate vandalism. But it can also mean that customers of French restaurants are more attentive to the surroundings and tends to signal when there's anomalies. In the end we can't really conclude to any precise causes but we do know for sure that vandalism are much more common on the north-east areas of Paris. Therefore these are the areas that need more Police surveillance.

# 6  Discussion

Note, i am not necessarily trying to denounce yellow vests' actions, but i do regret that they've been heard only after they started breaking stuff in the city. That's my opinion only, i think the problem lies within the French government that doesn't listen to pacifists.
The dataset used might not be suited or incomplete for this case. We have to keep in mind these are signals gathered by the crowd and not professionnal dataset. So all the signal depends of the people using the app, also not everybody care or know about it, in fact i didn't knew about this until last week.
There's also the notion of redundancy, many signals may be pointing to the same anomalies. A more serious dataset may lead to better results.
This study was only based on my knowledge and ideas, there's probably more pists to explore if the topic is of interest to you.