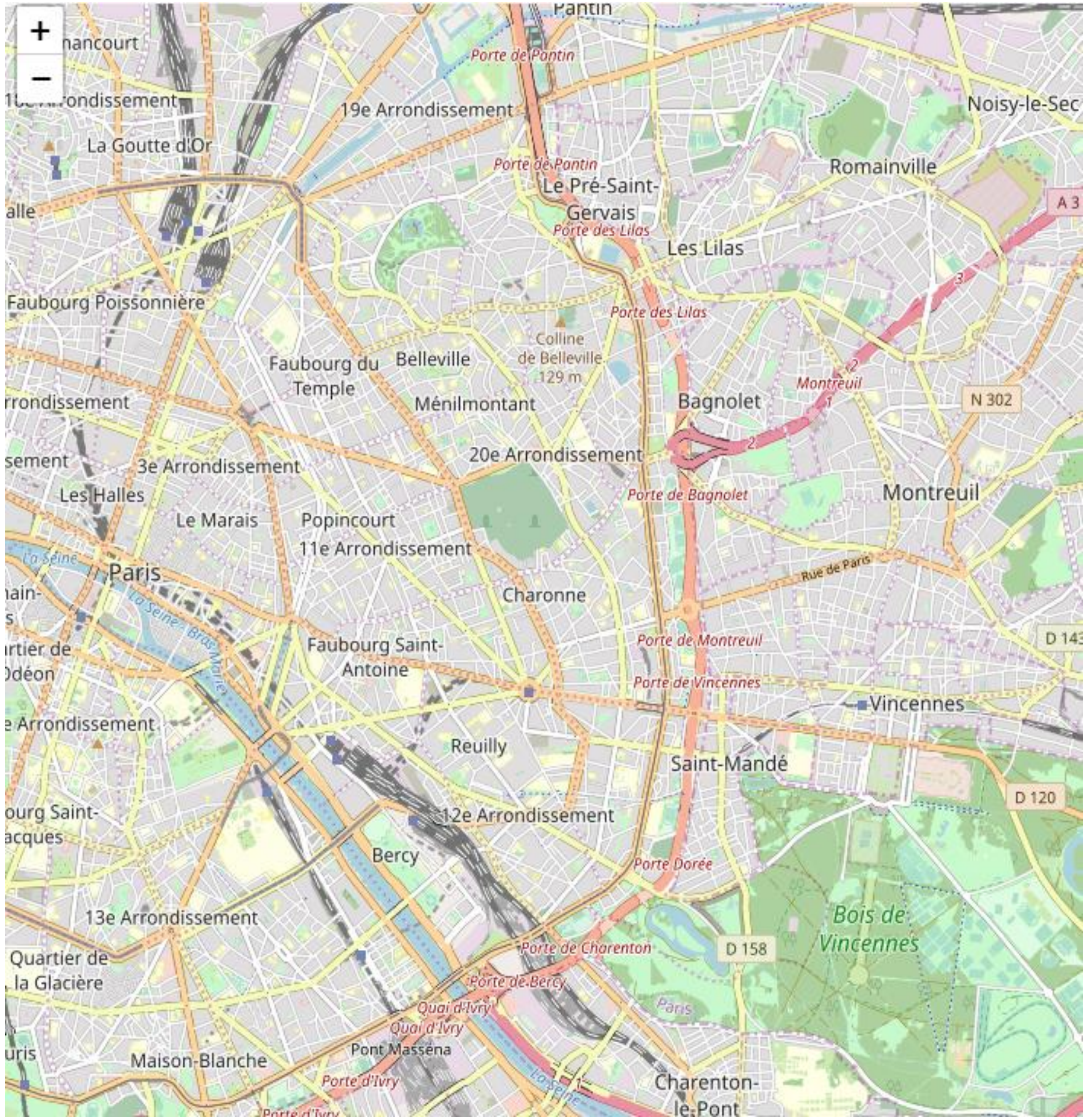
Paris – City of Light

Coursera Capstone

Paul Preda

January 2020

Adjust local public policies in order to increase the attractiveness of neighborhoods East Paris



Help Paris Mayor and his colleagues for East Paris

- To increase the attractiveness of their cities…
- considering that the real estate price/square meter is a good indicator of this attractiveness

Is there any link…

- Between the number and nature of interest venues in each neighborhood…
- … and the attractiveness of this one?

If yes, how can the Mayor…

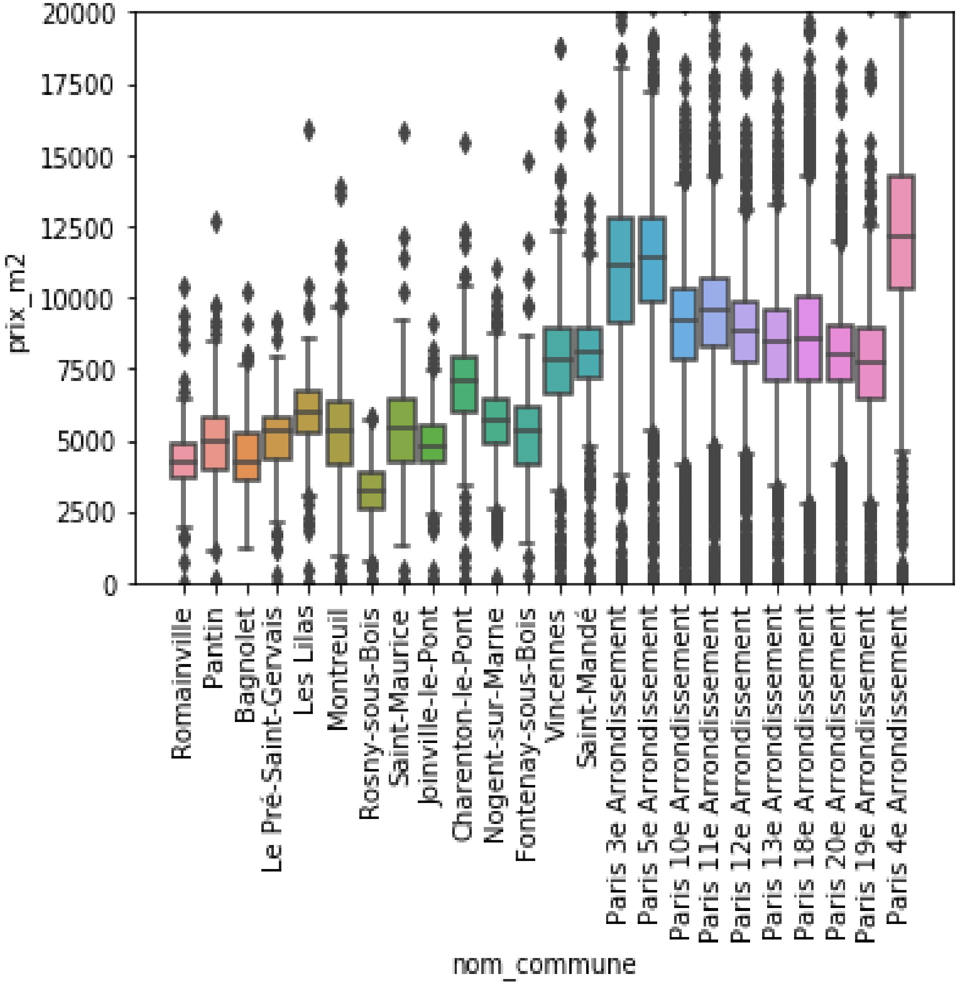
- … adjust the public policies in order to increase this attractiveness?

Data acquisition and cleaning – Real Estate transactions

Postal code	City Name	Number of final transactions
75003	Paris 3e Arrondissement	760
75004	Paris 4e Arrondissement	524
75005	Paris 5e Arrondissement	824
75010	Paris 10e Arrondissement	1627
75011	Paris 11e Arrondissement	2561
75012	Paris 12e Arrondissement	1675
75013	Paris 13e Arrondissement	1576
75018	Paris 18e Arrondissement	3626
75019	Paris 19e Arrondissement	1829
75020	Paris 20e Arrondissement	2123
93100	Montreuil	1035
93110	Rosny-sous-Bois	472
93170	Bagnolet	312
93230	Romainville	205
93260	Les Lilas	275
93310	Le Pré-Saint-Gervais	198
93500	Pantin	643
94120	Fontenay-sous-Bois	430
94130	Nogent-sur-Marne	572
94160	Saint-Mandé	407
94220	Charenton-le-Pont	427
94300	Vincennes	849
94340	Joinville-le-Pont	206
94410	Saint-Maurice	211

- Public data corresponding to real estate transactions done in 2018 in 3 French Departments
- Downloaded from the public site https://www.data.gouv.fr/fr/datasets/demandes-de-valeurs-foncieres-geolocalisees/
- Choose data corresponding to 24 cities (10 Paris districts and 14 cities nearby East Paris)
- Group the elements per transaction; keep a **single line per transaction**
- Keep only simple sales transactions concerning the apartments
- Calculate the price / square meter for each transaction
- Final dataset of **23000 real estate transactions** done in **2018** in the **24 cities** of our choice, with the associated **price / square meter**

Analyze the price per square meter in the 24 cities

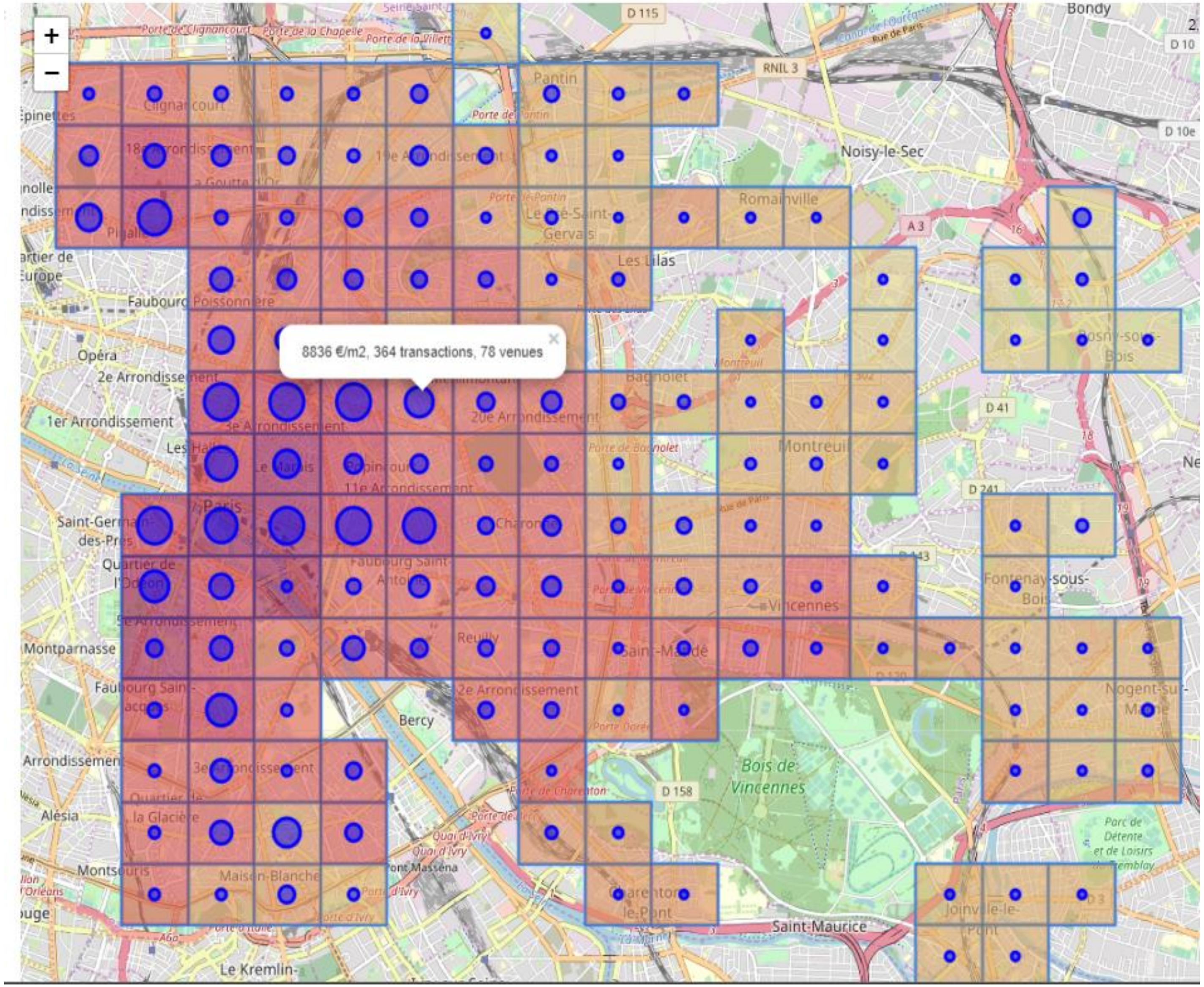


- Important discrepancies between the 24 cities
- Median price varying from 3000€/m² to 12000€/m²
- High dispersion between min and max prices in the studied cities
- Many transaction outliers that we dropped out

Conclusion

- The analyze per city is too coarse
- Use a more granular approach, based on geographical small neighborhoods

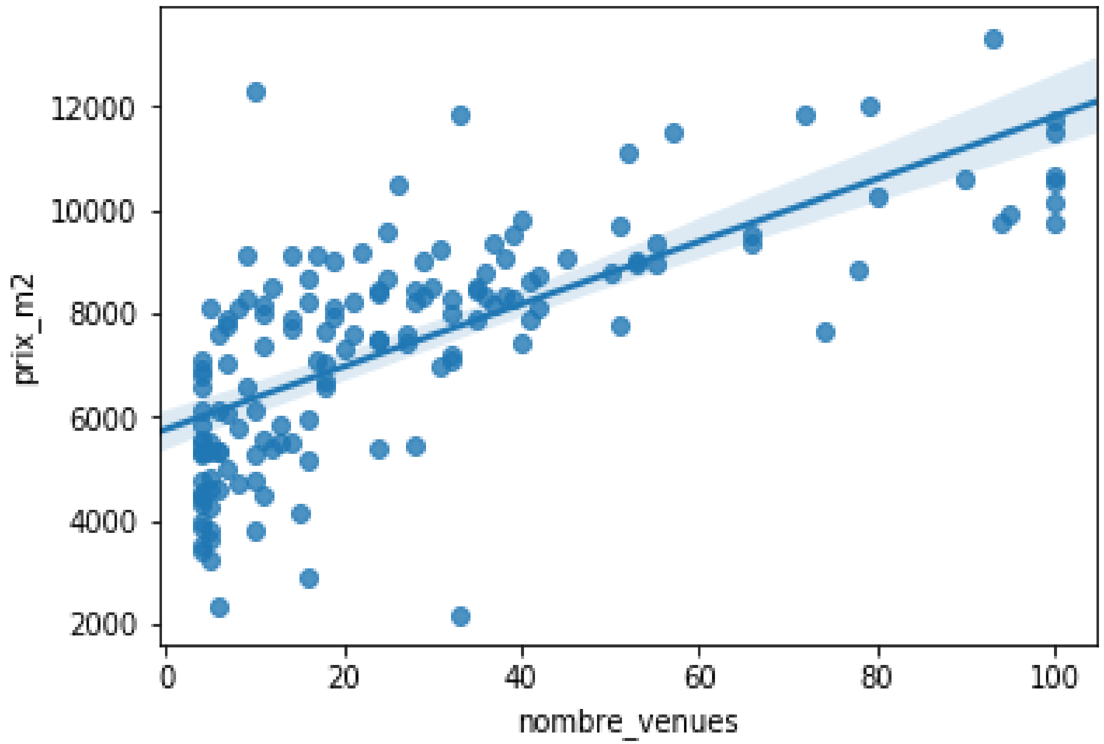
Data collection and analysis:
use small geographical zones and collect venues of interest



- Segment the geographical zone of interest in a grid of 18*18 square tiles
- Calculate the mean price/square meter for the transactions in each tile
- Collect the Foursquera venues of interest in each tile (radius of 400m): more than 4400 venues collected
- Keep only the tiles with more than 20 transactions and at least 4 interest venues
- Represent on the same map:
  - Mean prices/m² as choropleth
  - The number of venues of interest as circle marker radius

There seem to be a correlation between the number of venues and the mean price/m²

Analysis 1: Simple Linear Regression

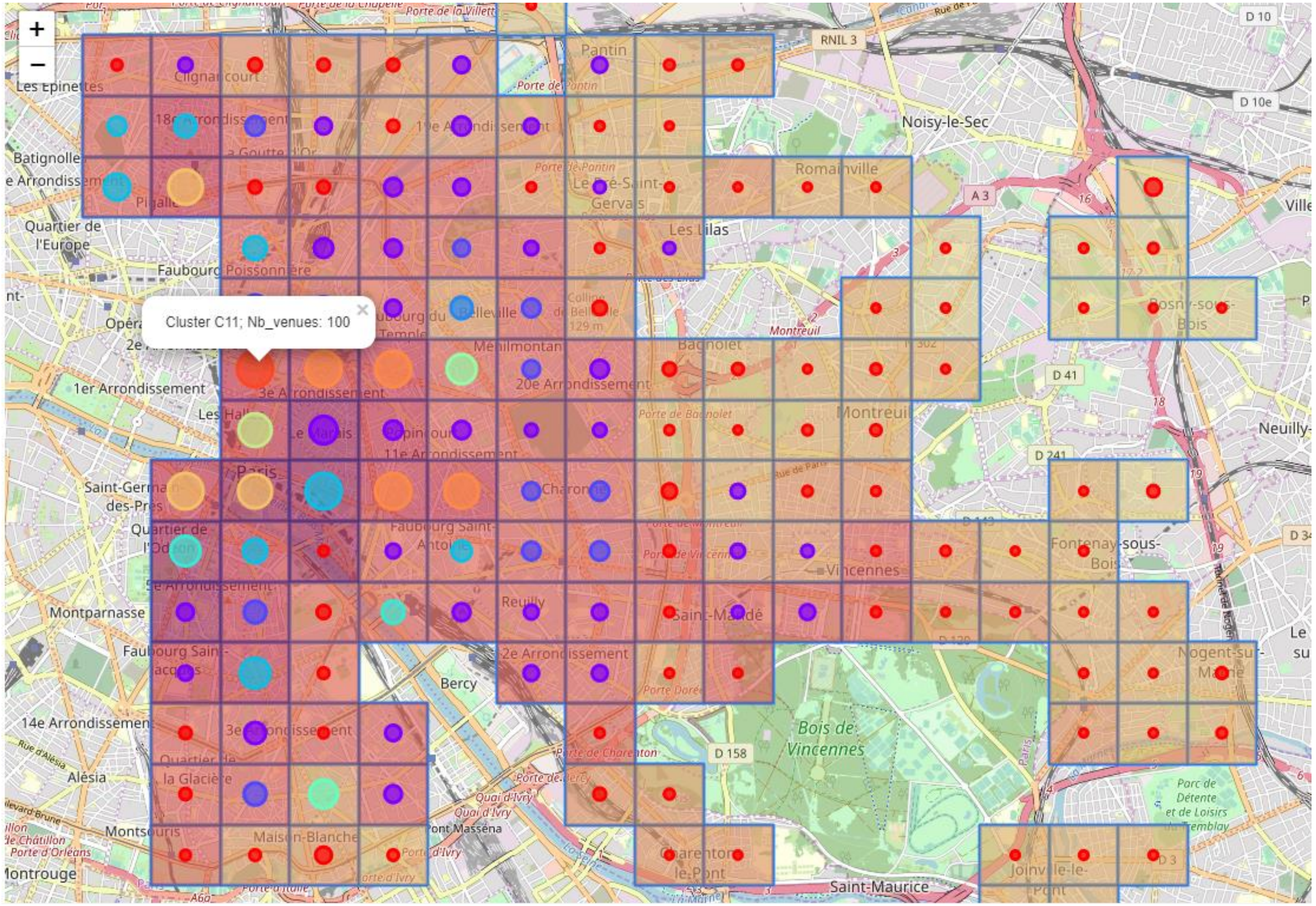


Simple linear regression:
price/m² = f(number of venues)

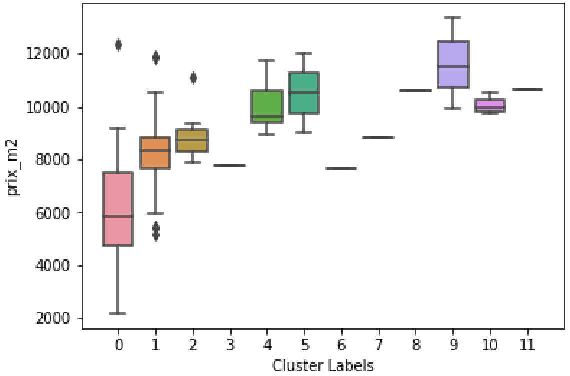
- Look for a simple linear regression between the average price/m² and the number of venues in each tile
- Poor correlation:

Variance score of 0.47

Analysis 2: Cluster the tiles by categories of venues



- Use KMeans algorithm to segment the tiles in 12 different clusters
- Clustering based on the number of venues per category present in each tile



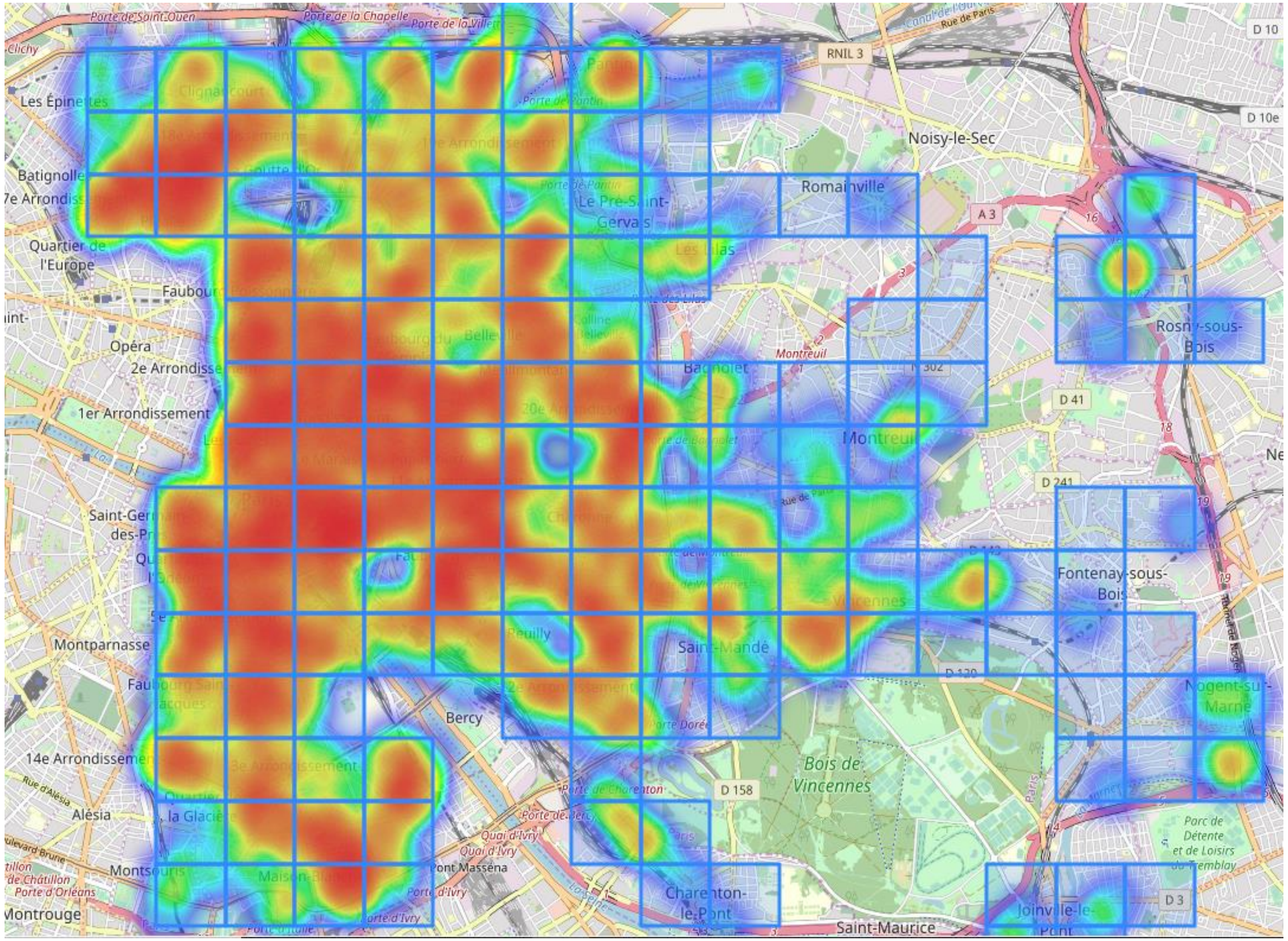
- Analysis shows a weak correlation between the clusters and mean price/m2

Poor variance score: 0.37

Analysis 2: Qualitative analysis per venue Cluster

Cluster label	Mean number of venues	Price per square meter	Description	Most common venues					
0 -	10	Low	Out of Paris, few venues, most utilities	0.63 Supermarket	0.52 Hotel	0.46 French Restaurant	0.31 Bakery	0.31 Plaza	0.25 Café
1 -	30	Medium	Paris or close to Paris mix area	3.08 French Restaurant	1.83 Hotel	1.25 Bar	1.19 Café	1.14 Italian Restaurant	1.14 Japanese Restaurant
2 -	44	Medium	Paris residential area	5.92 French Restaurant	4.50 Bar	1.75 Hotel	1.50 Bistro	1.50 Pizza Place	1.33 Coffee Shop
3 -	51	Medium	Paris – Bars and Asian cusine	9.00 Bar	6.00 Chinese Restaurant	6.00 Vietnamese Restaurant	4.00 French Restaurant	3.00 Supermarket	2.00 Dim Sum Restaurant
4 -	62	High	Central touristic area	11.62 French Restaurant	3.88 Hotel	3.62 Italian Restaurant	2.62 Bakery	2.00 Café	1.62 Bar
5 -	66	High	Central touristic area	9.00 Hotel	6.00 French Restaurant	3.50 Sandwich Place	2.00 Indie Movie Theater	2.00 Nightclub	2.00 Cocktail Bar
6 -	74	Medium	Chinese neighborhood	16.00 Vietnamese Restaurant	11.00 Asian Restaurant	9.00 Thai Restaurant	5.00 Chinese Restaurant	4.00 French Restaurant	3.00 Supermarket
7 -	78	Medium	Paris Republique – Bars & Bistros	21.00 Bar	3.00 Pizza Place	3.00 French Restaurant	3.00 Restaurant	2.00 Middle Eastern Restaurant	2.00 Bistro
8 -	90	High	Central touristic area	8.00 French Restaurant	5.00 Bakery	4.00 Hotel	3.00 Furniture / Home Store	3.00 Cosmetics Shop	3.00 Plaza
9 -	96	High	Central touristic area	19.00 French Restaurant	4.00 Plaza	3.67 Hotel	3.33 Ice Cream Shop	3.00 Bakery	3.00 Italian Restaurant
10 -	98	High	Central touristic area	12.75 French Restaurant	6.50 Bar	4.25 Bistro	3.75 Hotel	3.50 Wine Bar	2.75 Restaurant
11 -	100	High	Central "chic" area	9.00 Cocktail Bar	6.00 French Restaurant	6.00 Wine Bar	5.00 Hotel	5.00 Bakery	4.00 Thai Restaurant

Analysis 3: Multiple linear regression per category of venues



Heat map of venues having a positive correlation with the attractiveness of the neighborhoods

- Use a multiple linear model regression with regularization
- Search for a multiple linear model able to predict price/m² = f(venue categories)

Variance score: 0.58

27 categories of venues identified as having a strong positive impact on their neighborhood

Results and discussion: what public policies?

11 public or semi public type of venues on which mayors can act in order to improve the attractiveness of their cities:

Venue category	Correlation coefficient
Bike Rental / Bike Share	90.629154
Plaza	96.167862
Metro Station	104.286308
Pedestrian Plaza	124.636646
Historic Site	133.162531
Science Museum	186.993403
Garden	207.822985
Museum	223.934375
Park	225.227832

Private type of venues that mayors can encourage the installation:

Venue category	Correlation coefficient
Sushi Restaurant	95.987113
Burger Joint	104.464146
Pastry Shop	108.642213
Vegetarian / Vegan Restaurant	113.688283
Coffee Shop	116.578945
Falafel Restaurant	119.387966
Gourmet Shop	124.331205
Gastropub	127.436126
Chinese Restaurant	134.736091
French Restaurant	146.000185
Japanese Restaurant	163.238495
Bistro	163.795533
Gym	193.036590
Hotel	210.575776
Café	289.155960

Conclusions and further developments

- There is a moderate correlation between the type of interest venues available in a neighborhood and the attractiveness of this one (measured via the mean real estate price)
- Some public or semi public type of venues have a clear positive impact on their environment:
  - Park and gardens
  - Museum and historical sites
  - Plazas
- Public policies can directly act on the above, but also encourage private venues as cafés, quality restaurants…
- This study is just a first step, as many other factors need to be considered:
  - Quality of schools
  - Criminality rate
  - Socio economic dynamics
  - Tax level etc.