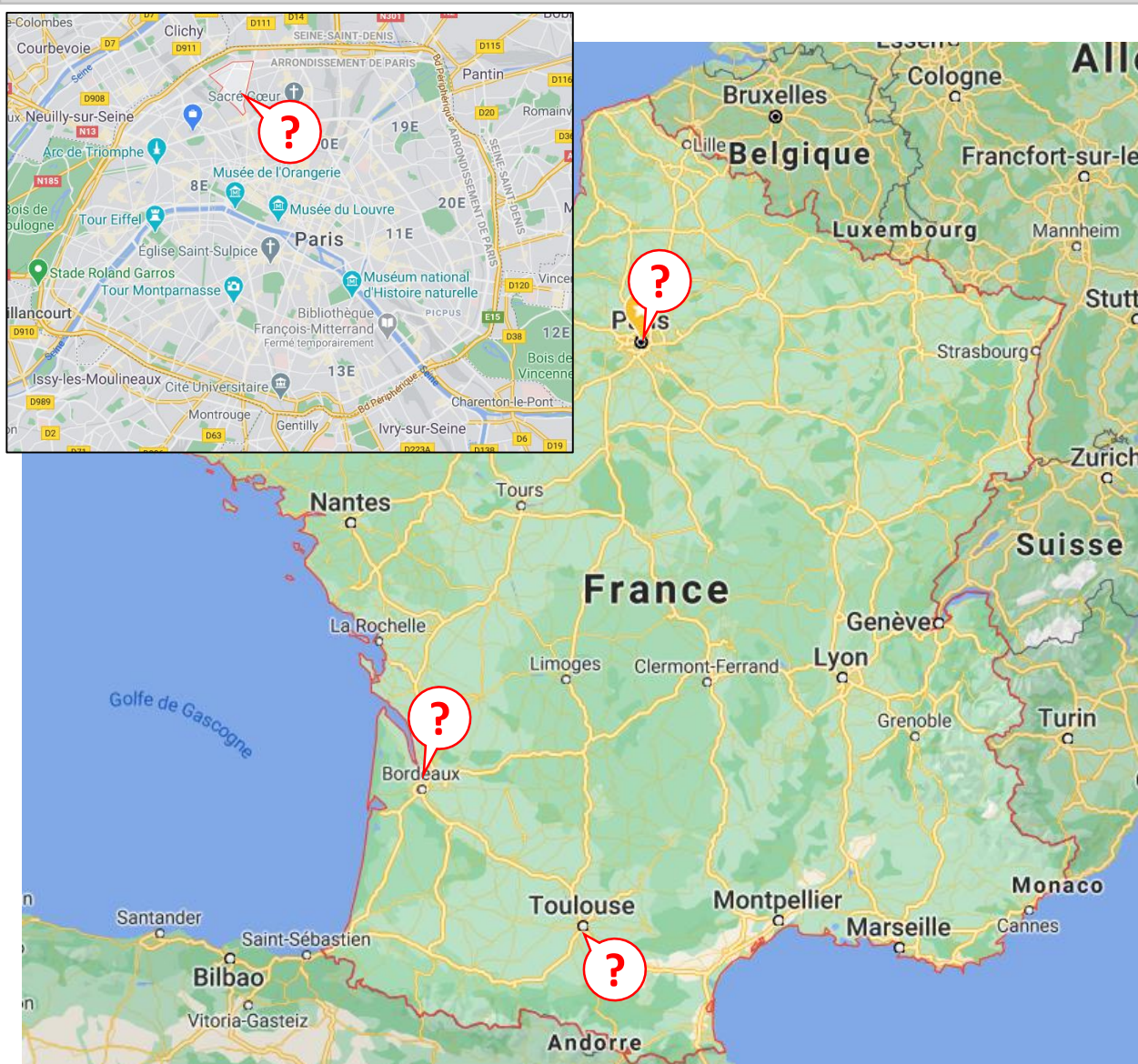


BUSINESS PROBLEM

Where to move out?



WHERE TO MOVE OUT?

- Assume you are leaving in a neighborhood
 - You enjoy living in this place & its way of life
 - But you want to move out for whatever reason:
 - Want to buy your first home
 - Looking for a bigger house
 - Looking for cheaper real estate prices
 - Looking for somewhere closer to your new job...
- You are looking for a list of neighborhoods similar to yours with their real estates prices

Here we will assume you are living in the Epinettes in Paris for example, & you are ready to move to another neighborhood in Paris, Bordeaux or Toulouse

We will provide:

- List of neighborhoods from those 3 cities that are similar to the Epinettes based on their venues
- Their top 5 most common venues
- Their real estate prices

METHODOLOGY & DATA SOURCES

1. GET THE LIST OF NEIGHBORHOODS FOR THE 3 CITIES & THEIR COORDINATES

- After some data mining I have been able to find open source data for Paris & Toulouse neighborhoods
- As for Bordeaux I had to go on google maps to get the coordinates & create an excel database
 - Source « quartier_paris.csv »: <https://www.data.gouv.fr/fr/datasets/recensement-population-2013-grands-quartiers-population/>
 - Source « quartier_toulouse.csv »: <https://data.toulouse-metropole.fr/explore/dataset/recensement-population-2016-grands-quartiers-diplomes/information/>
 - Source: « quartier_bordeaux.xlsx »: from google maps
- **154 neighborhood in our data:** 80 in Paris, 60 in Toulouse, 14 in Bordeaux

2. GET THE REAL-ESTATE PRICE FOR EACH NEIGHBORHOOD

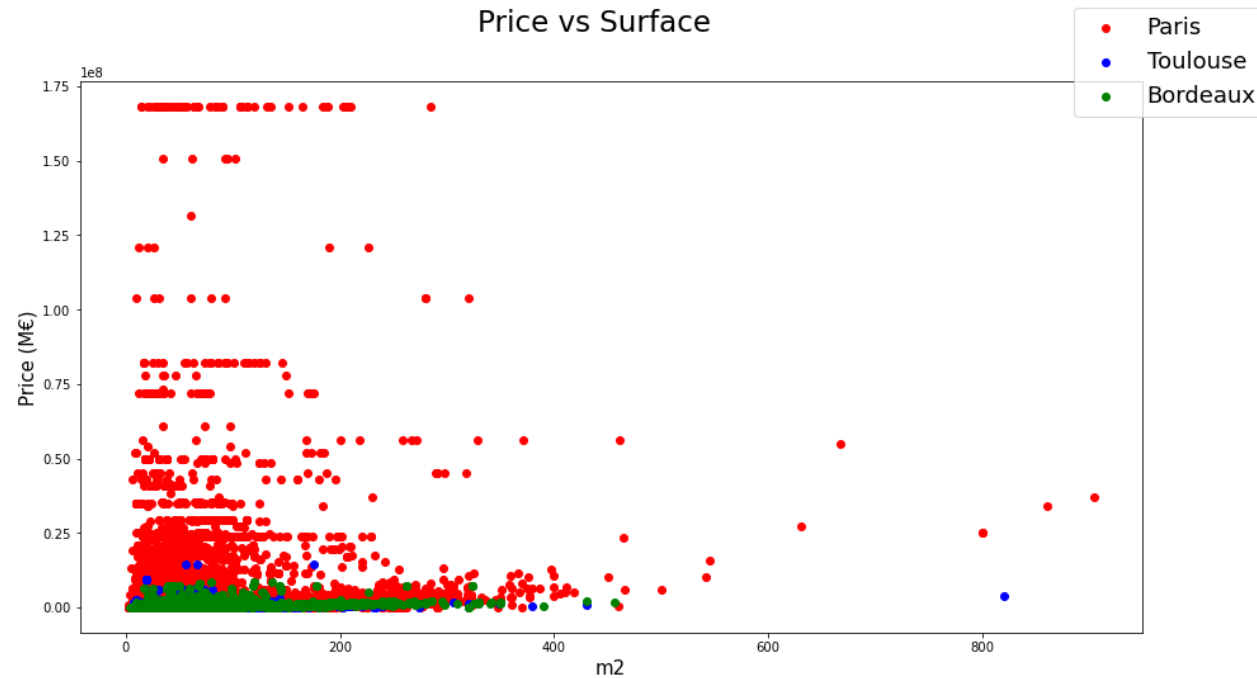
- Open source data available with all real estate transactions for 2019: **database of 50012 transactions for our 3 cities**
 - Source « 31.csv », « 75.csv », « 33.csv »: <https://cadastre.data.gouv.fr/data/etalab-dvf/latest/csv/2019/>
- Prices have slightly increased in 2020 in those 3 cities but it is marginal
- Filter the « exotic » or « luxury » properties & get an average price/m2 for each neighborhood

3. COMPARE & CLUSTER THE NEIGHBORHOODS BASED ON THEIR VENUES

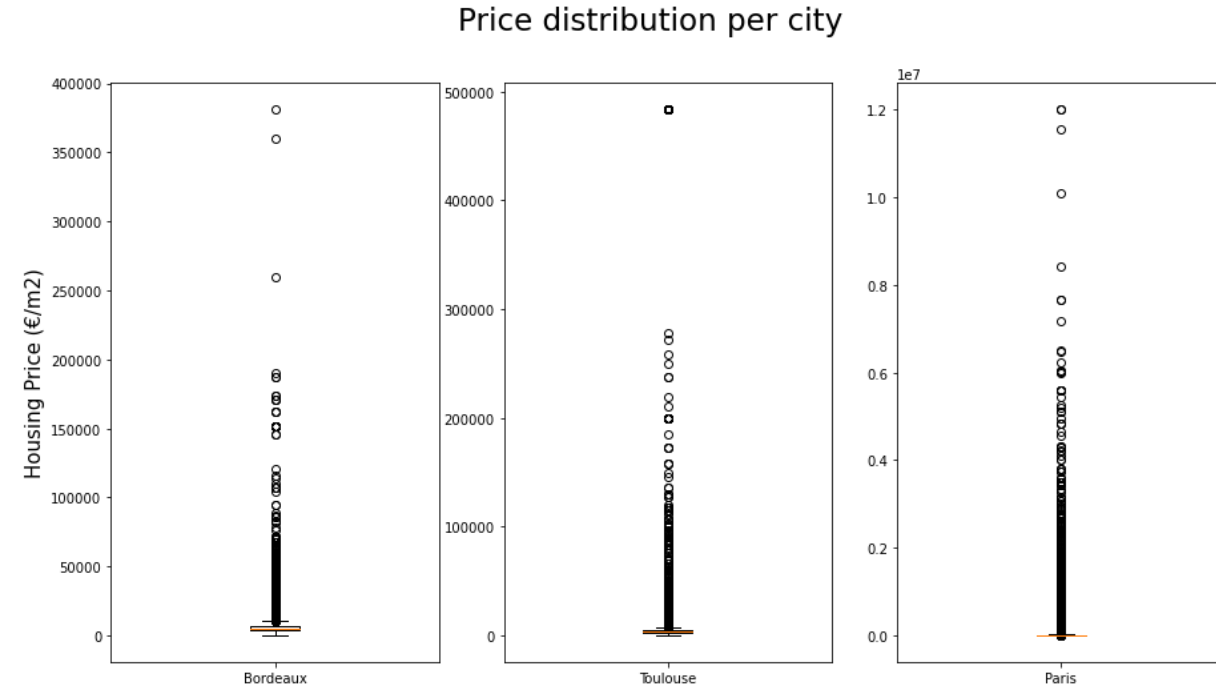
- Get the venues for each neighborhood
 - Source: Foursquare API
- Get the 5 most common venue categories for each neighborhood
- Cluster the neighborhoods using kmeans clustering

→ **Result: list of neighborhoods from the same cluster with their 5 most common venue categories & their real estate price**

REAL ESTATE PRICE DISTRIBUTION PER CITY



We can't really observe a trend when looking at price vs surface...

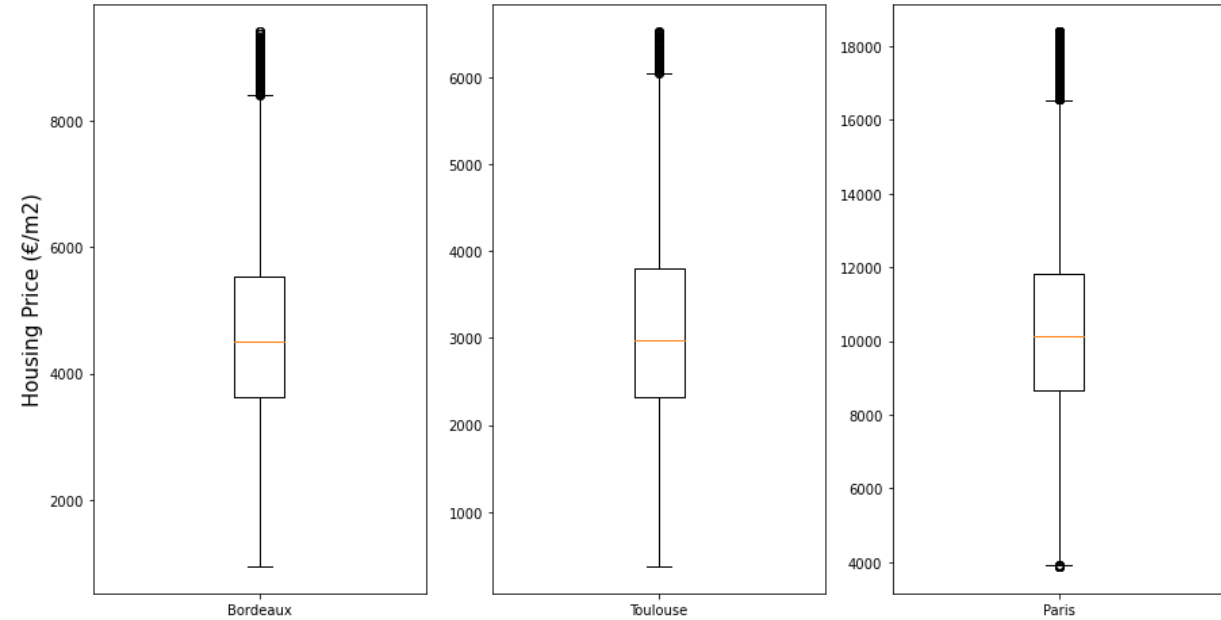


... Because there are a lot of outliers – luxury properties with a very high price/m2

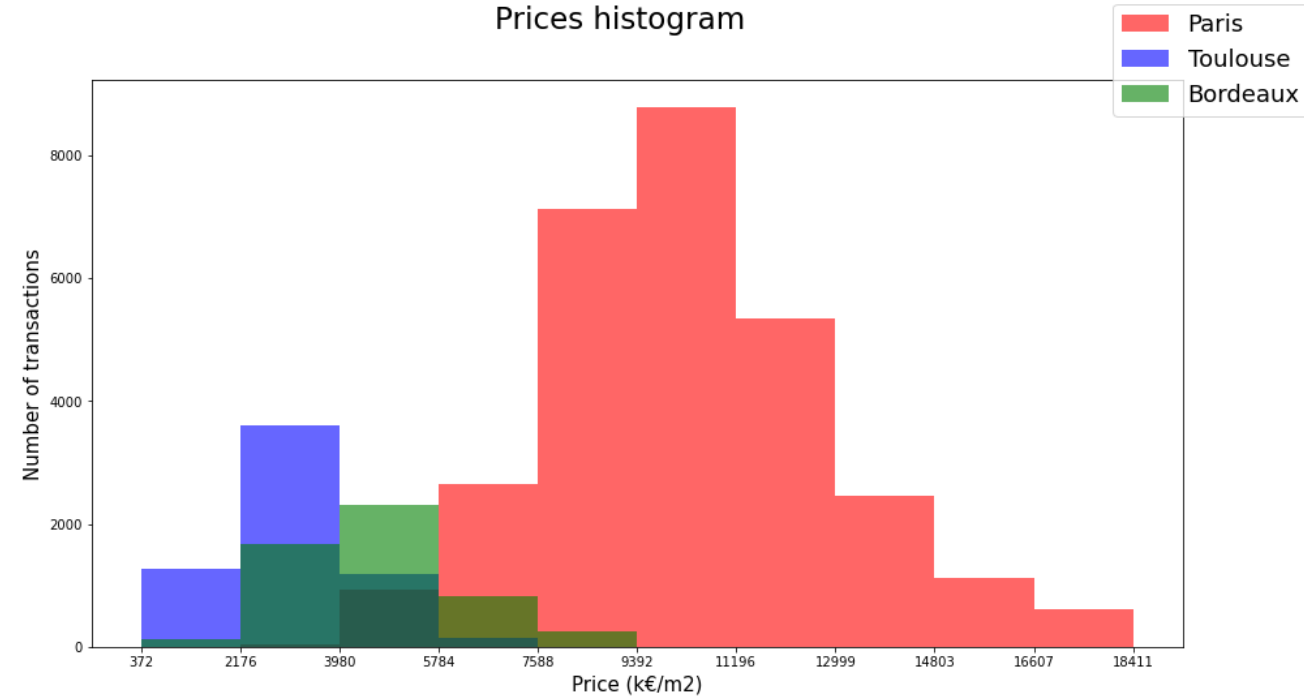
→ We will remove the outliers as we are not interested in extraordinary transactions
→ We don't want them to impact our results

PRICE DISTRIBUTION PER CITY AFTER OUTLIERS REMOVAL

Price distribution per city

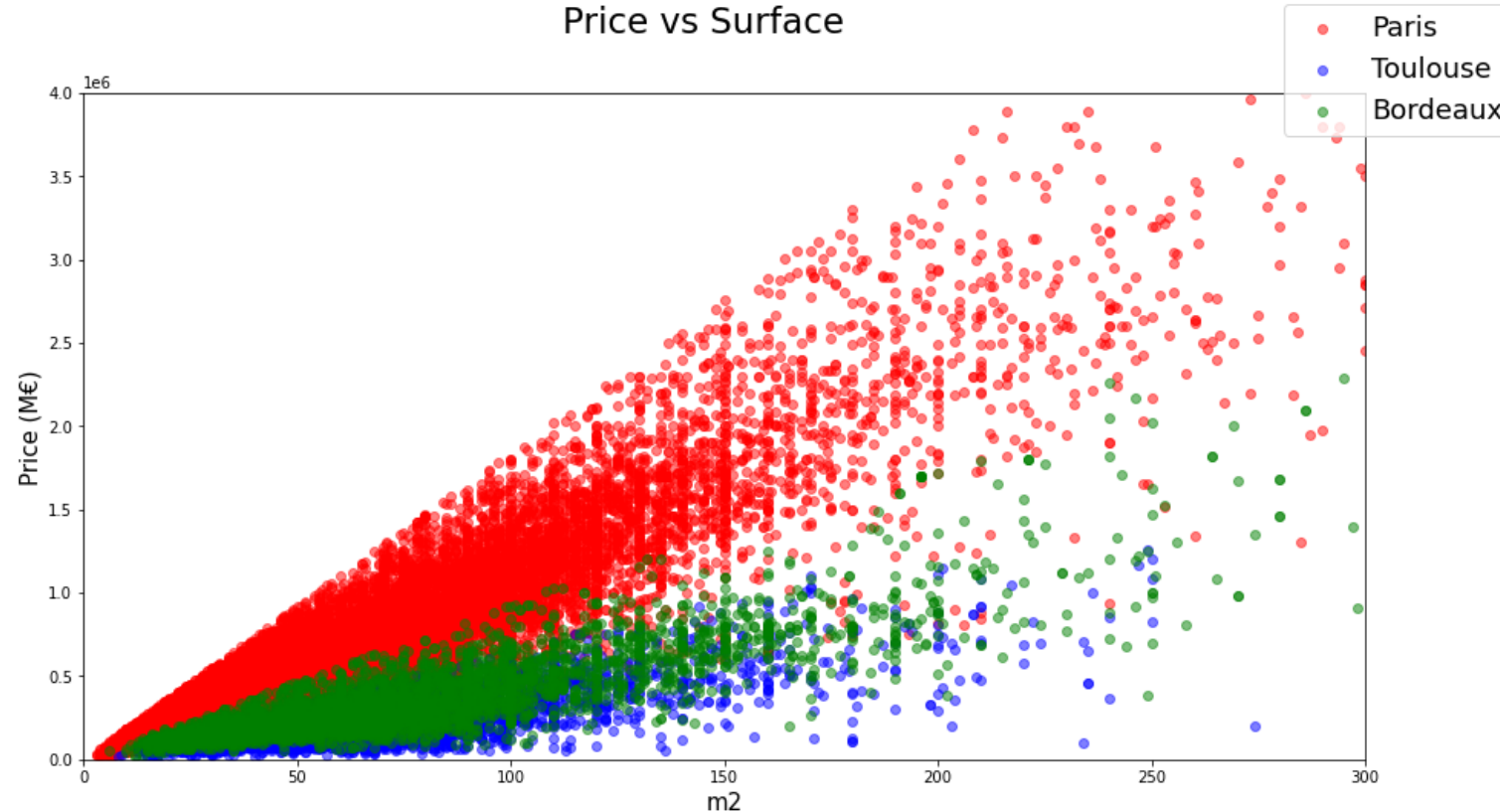


Prices histogram



- Dataset reduced from 50012 to 40435 transactions

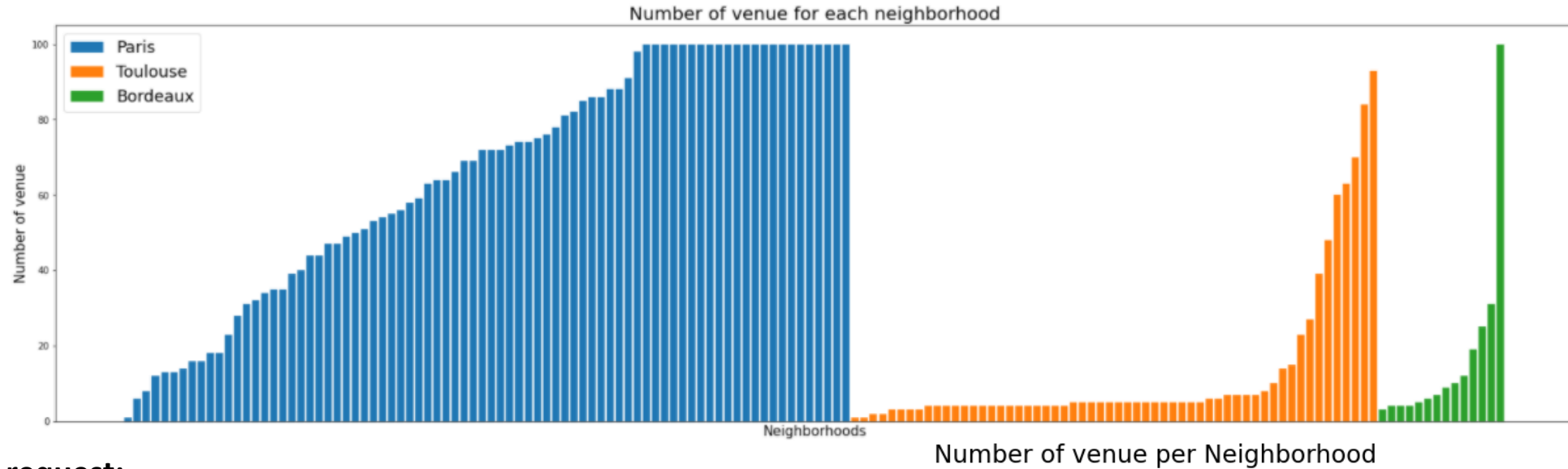
PRICE VS SURFACE – TRANSACTIONS FILTERED



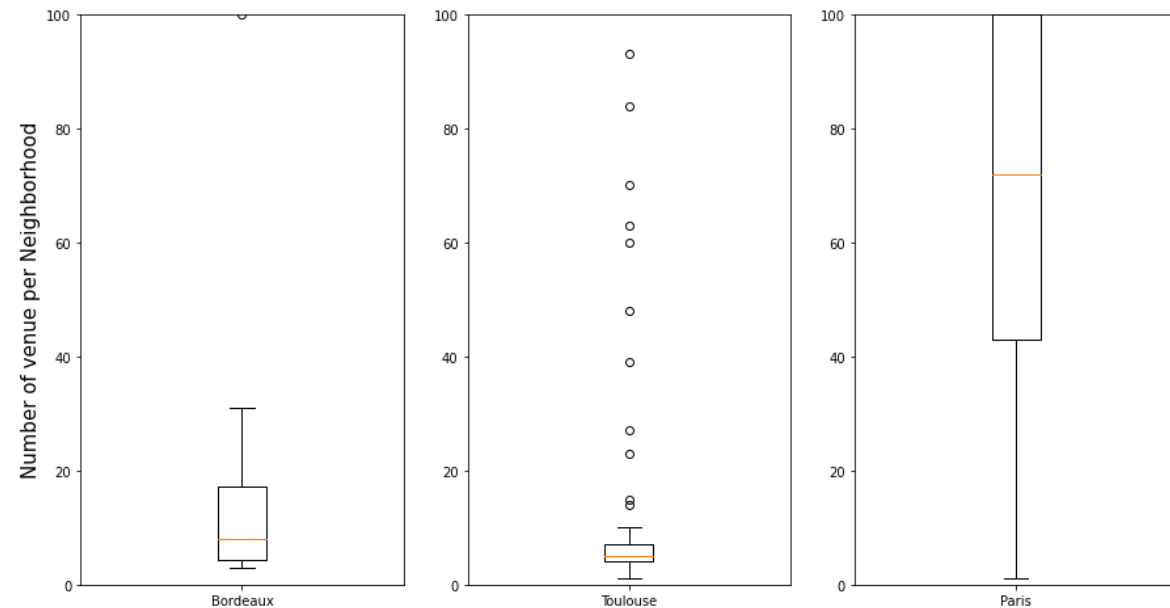
We can now observe expected trends – Dataset ready to be used

- Price increase linearly with surface once « exotic » properties have been removed
 - Paris is more expensive than Bordeaux & Toulouse
- For a given surface there is still a price spread for each city depending on the neighborhood

Distribution of number of venues per neighborhood

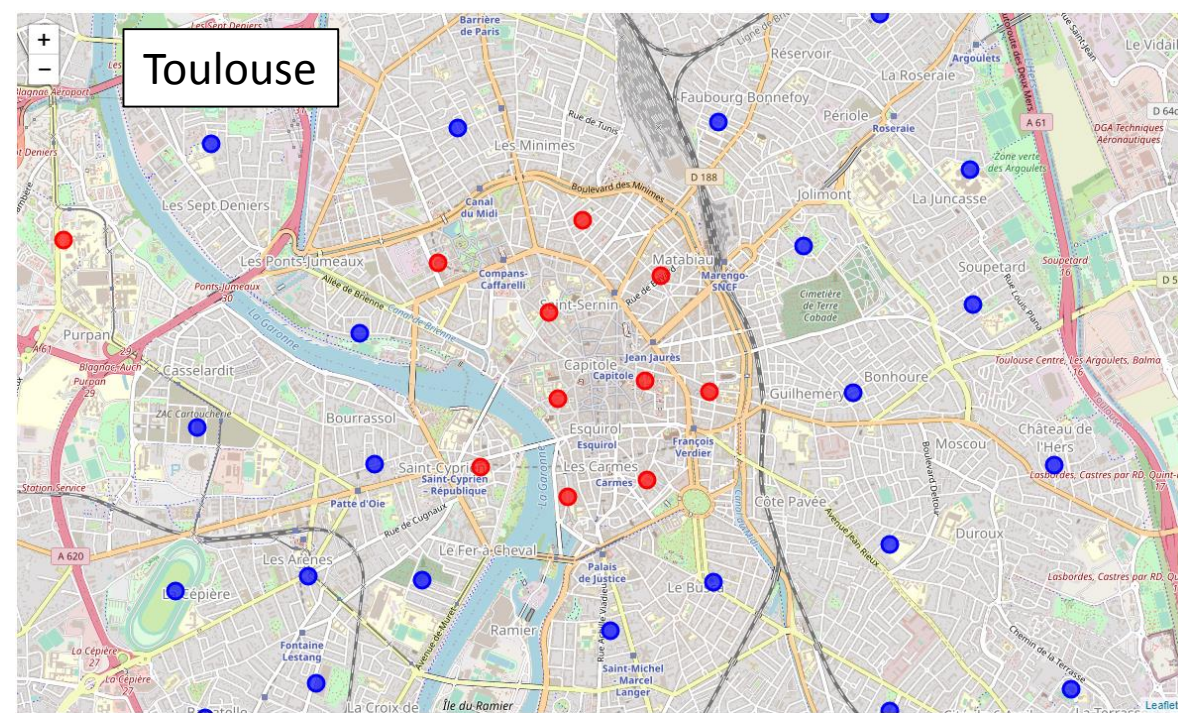
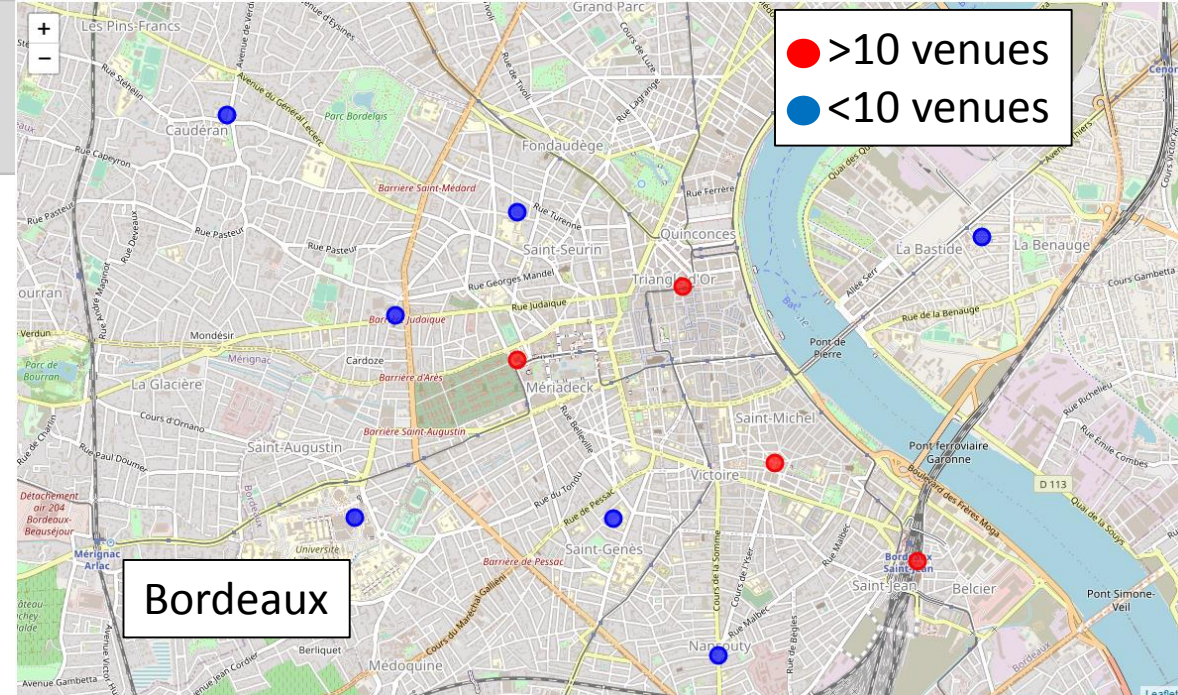
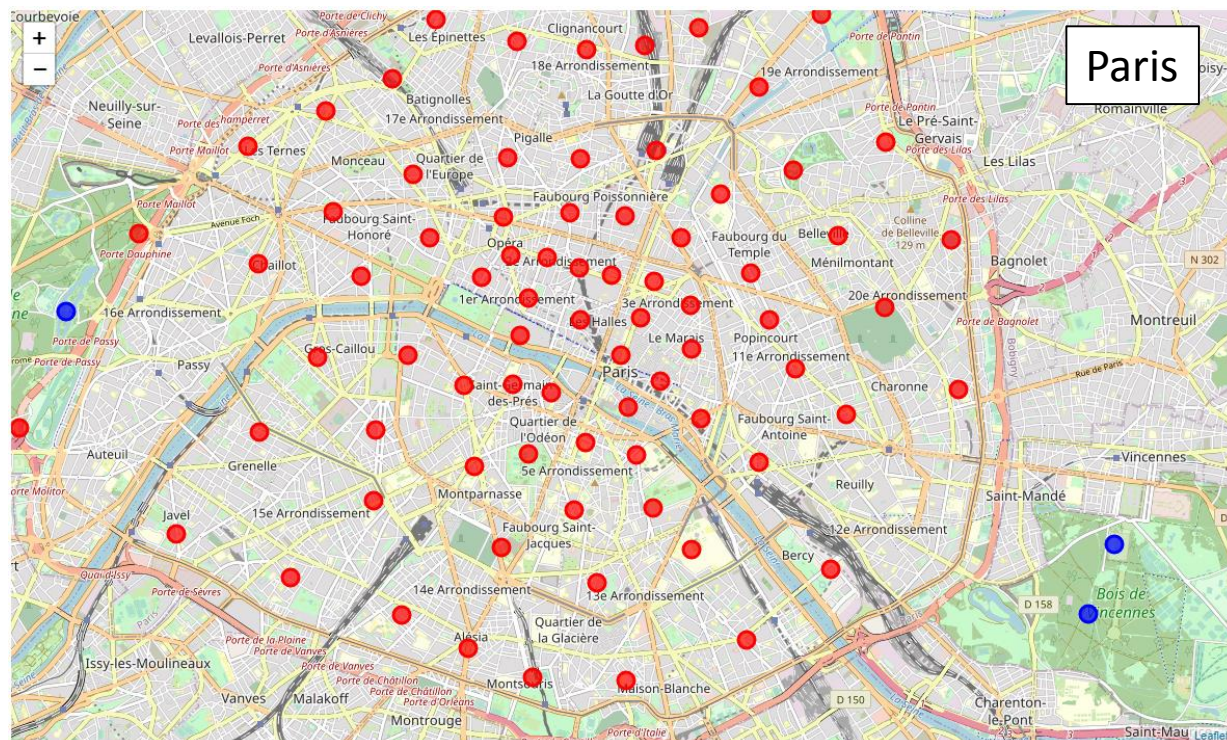


- **Foursquare request:**
 - Venues in a radius of 500m for each neighborhoods
 - Max 100 venues per neighborhoods
- **6236 venues for Paris, Toulouse & Bordeaux**
- **Distribution are very different:**
 - Several neighborhoods in Paris have more than 100 venues referenced under Foursquare
 - Many neighborhoods from Bordeaux & Toulouse have a small number of Foursquare referenced venues



Number of venues per neighborhood

Map display

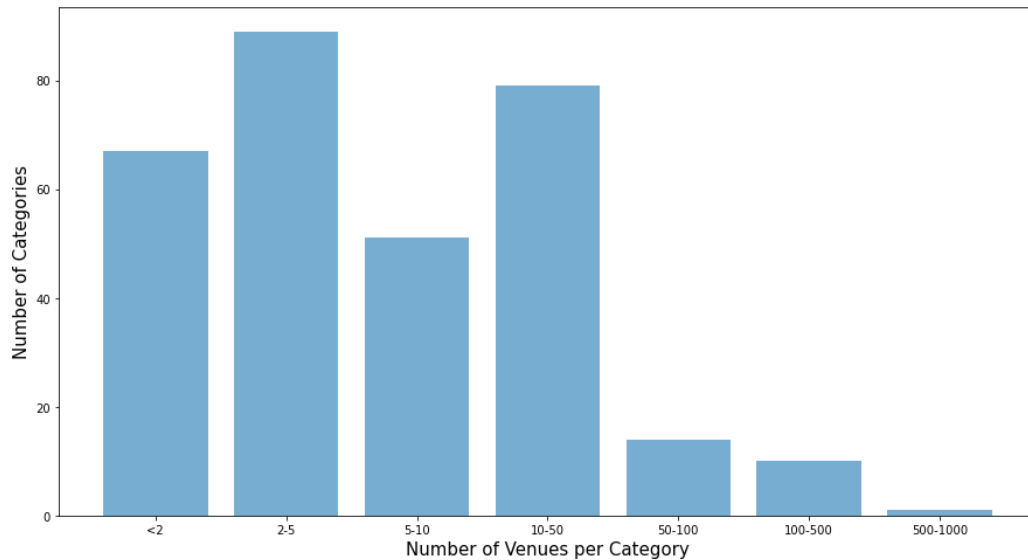


- Higher density of venues in city centers
- Paris « intra-muros » can be considered as one huge city center
- Some neighborhoods from Bordeaux or Toulouse are essentially residential
- **But there is also a bias:** higher proportion of venues referenced under Forsquare in Paris than in Toulouse or Bordeaux: Paris is the capital, more touristic, more international, more cosmopolitan

Data preparation for kmeans clustering

- **Filter: Remove neighborhoods with less than 10 venues from Foursquare Request**
 - Either they are not representative: not enough venues to characterise the neighborhood
 - Or there is actually a very small number of venues & in this case we are not interested in those to move in
- Number of neighborhoods down from 154 to 93
- **Number of venues down from 6236 to 5954**
- 60% of neighborhoods concentrates 95% of the Foursquare request results

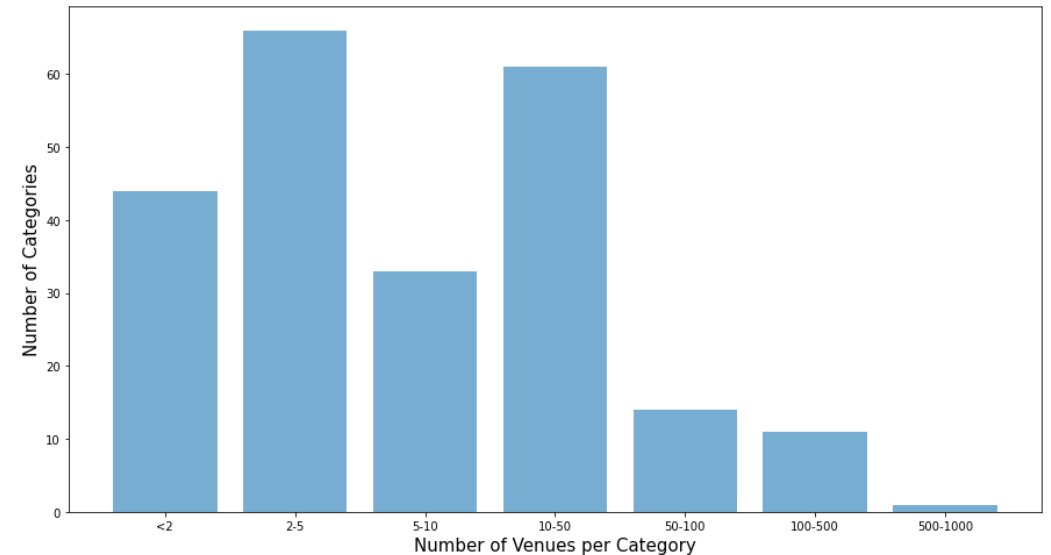
Number Venues per category histogram



Grouping
Some
Categories

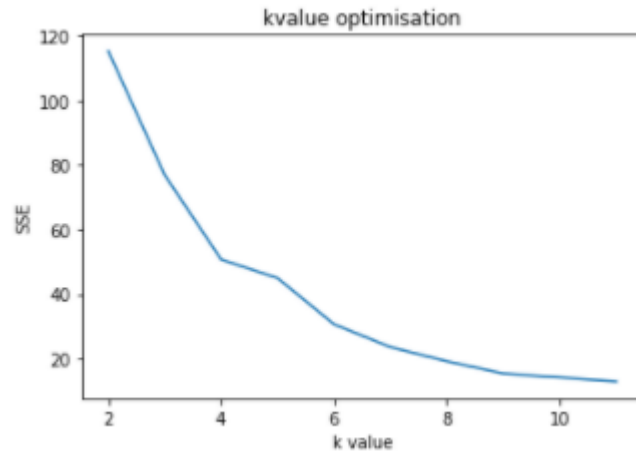


Number Venues per category histogram

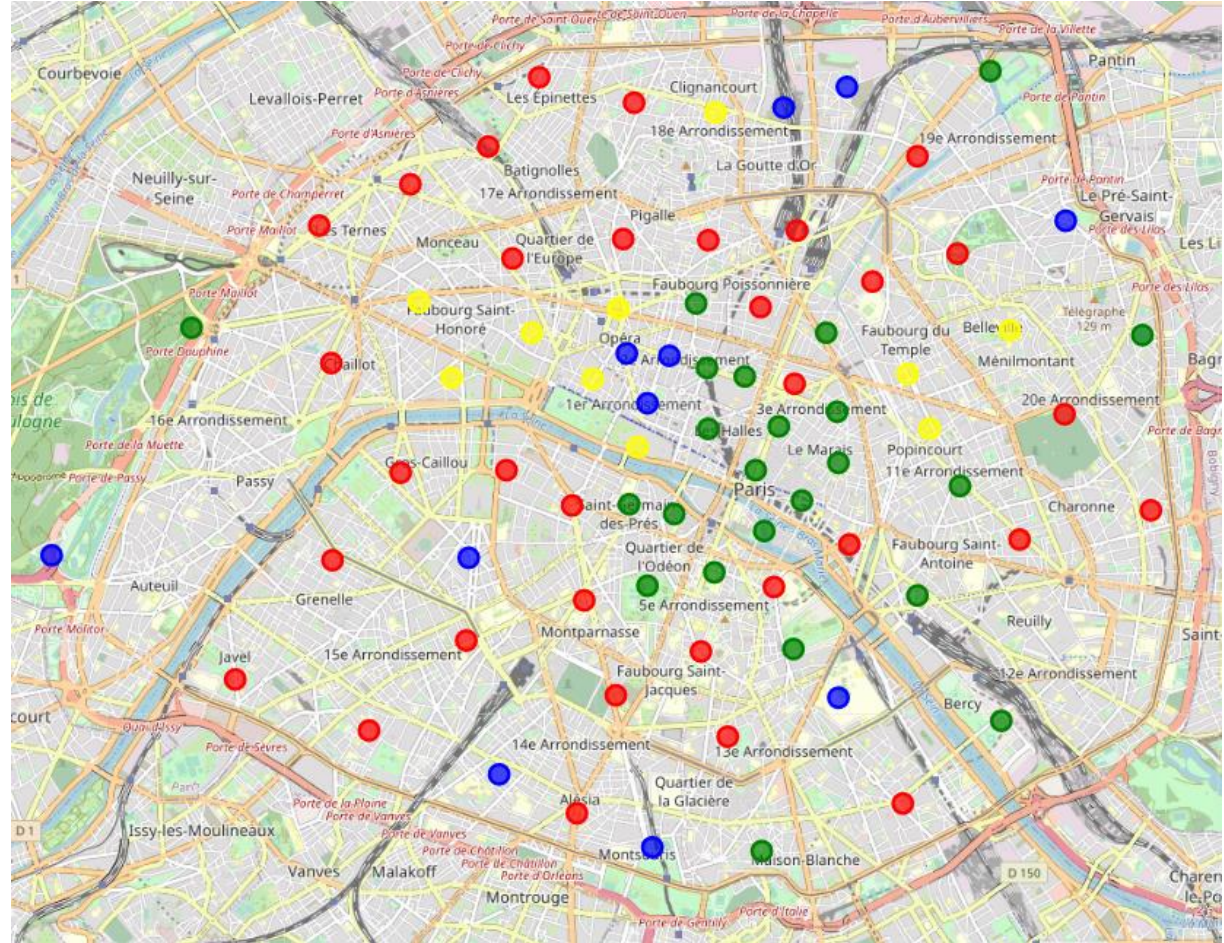


- **There are 319 different categories of venue**
- 87 of them are different types of restaurant → this is too much detailed & some similar restaurant may fall in different categories if they have been described slightly differently
- **We reduce the restaurant categories to 9**
- Number of different categories down to 239

Kmeans clustering



K=4 is the optimised value for clustering



Paris example: each color is a different cluster

Result with k = 4

The list is quite long – lets run with more clusters

Neighborhood	Latitude	Longitude	City	arrondissement	prix_m2	1st Most common venue	2nd Most common venue	3rd Most common venue	4th Most common venue	5th Most common venue	Cluster
Saint Jean - Belcier	44.83	-0.56	Bordeaux	Bordeaux	4577.33	Cluster	Hotel	Nightclub	Supermarket	Asian_Restaurant	1
SAINT-GEORGES	43.60	1.45	Toulouse	Toulouse	4515.94	French Restaurant	Plaza	Hotel	Bar	Asian_Restaurant	1
MATABIAU	43.61	1.45	Toulouse	Toulouse	4380.18	Cluster	Hotel	French Restaurant	Bar	Asian_Restaurant	1
SAINT-AUBIN - DUPUY	43.60	1.45	Toulouse	Toulouse	4173.31	French Restaurant	Hotel	Bar	European_Restaurant	North_American_Restaurant	1
COMPANS	43.61	1.43	Toulouse	Toulouse	3741.14	French Restaurant	Hotel	Sandwich Place	Bar	Pizza Place	1
Arsenal	48.85	2.36	Paris	4	0.00	French Restaurant	Hotel	European_Restaurant	Modern_Food_Restaurant	Asian_Restaurant	1
Plaine de Monceaux	48.89	2.30	Paris	17	0.00	Cluster	French Restaurant	European_Restaurant	Hotel	Bakery	1
Porte-Saint-Denis	48.87	2.35	Paris	10	0.00	Cluster	Asian_Restaurant	Hotel	French Restaurant	Bakery	1
Père-Lachaise	48.86	2.40	Paris	20	0.00	Cluster	Asian_Restaurant	Bar	Bakery	Bistro	1
Rochechouart	48.88	2.34	Paris	9	0.00	French Restaurant	Hotel	Bakery	Pizza Place	Asian_Restaurant	1
Saint-Georges	48.88	2.33	Paris	9	0.00	Cluster	Hotel	French Restaurant	European_Restaurant	Asian_Restaurant	1
Notre-Dame-des-Champs	48.85	2.33	Paris	6	0.00	French Restaurant	Hotel	North_American_Restaurant	Bakery	Asian_Restaurant	1
Saint-Lambert	48.83	2.30	Paris	15	0.00	French Restaurant	Asian_Restaurant	Hotel	French_Regions_Restaurant	Café	1
Saint-Thomas-d'Aquin	48.86	2.33	Paris	7	0.00	French Restaurant	Hotel	Café	Asian_Restaurant	Coffee Shop	1
Saint-Victor	48.85	2.35	Paris	5	0.00	French Restaurant	Asian_Restaurant	European_Restaurant	Bakery	Hotel	1
Saint-Vincent-de-Paul	48.88	2.36	Paris	10	0.00	Cluster	Asian_Restaurant	French Restaurant	Hotel	African_Restaurant	1
Sainte-Marguerite	48.85	2.39	Paris	11	0.00	French Restaurant	Asian_Restaurant	Bar	European_Restaurant	Hotel	1
Ternes	48.88	2.29	Paris	17	0.00	French Restaurant	European_Restaurant	Asian_Restaurant	Hotel	Modern_Food_Restaurant	1
Val-de-Grâce	48.84	2.34	Paris	5	0.00	Cluster	Asian_Restaurant	French Restaurant	Bar	Hotel	1
Petit-Montrouge	48.83	2.33	Paris	14	0.00	Cluster	Hotel	Asian_Restaurant	French Restaurant	European_Restaurant	1
Necker	48.84	2.31	Paris	15	0.00	Cluster	French Restaurant	Asian_Restaurant	Hotel	Café	1
Arts-et-Métiers	48.87	2.36	Paris	3	0.00	Cluster	French Restaurant	Asian_Restaurant	Hotel	Wine Bar	1
Europe	48.88	2.32	Paris	8	0.00	French Restaurant	Hotel	Asian_Restaurant	Pizza Place	European_Restaurant	1
Batignolles	48.89	2.31	Paris	17	0.00	French Restaurant	Hotel	European_Restaurant	Asian_Restaurant	Bar	1
Chaillot	48.87	2.29	Paris	16	0.00	Hotel	French Restaurant	Asian_Restaurant	European_Restaurant	Art Museum	1
Charonne	48.85	2.41	Paris	20	0.00	Cluster	Supermarket	Asian_Restaurant	Bar	Hotel	1
Combat	48.88	2.38	Paris	19	0.00	French Restaurant	Pool	Restaurant	European_Restaurant	Asian_Restaurant	1
Croulebarbe	48.83	2.35	Paris	13	0.00	Cluster	Asian_Restaurant	French Restaurant	Bar	Sandwich Place	1
Epinettes	48.89	2.32	Paris	17	0.00	Cluster	Asian_Restaurant	French Restaurant	Restaurant	Bakery	1
Gare	48.83	2.37	Paris	13	0.00	Cluster	Asian_Restaurant	Hotel	Sandwich Place	French Restaurant	1

Where to move if you want to leave the Epinettes

And what are the real estate prices

	Latitude	Longitude	City	arrondissement	prix_m2	1st Most common venue	2nd Most common venue	3rd Most common venue	4th Most common venue	5th Most common venue	Cluster	Cluster6
Neighborhood												
Saint Jean - Belcier	44.83	-0.56	Bordeaux	Bordeaux	4577.33	Cluster	Hotel	Nightclub	Supermarket	Asian_Restaurant	1	0
MATABIAU	43.61	1.45	Toulouse	Toulouse	4380.18	Cluster	Hotel	French_Restaurant	Bar	Asian_Restaurant	1	0
Arts-et-Métiers	48.87	2.36	Paris	3	0.00	Cluster	French_Restaurant	Asian_Restaurant	Hotel	Wine Bar	1	0
Plaine de Monceaux	48.89	2.30	Paris	17	0.00	Cluster	French_Restaurant	European_Restaurant	Hotel	Bakery	1	0
Val-de-Grâce	48.84	2.34	Paris	5	0.00	Cluster	Asian_Restaurant	French_Restaurant	Bar	Hotel	1	0
Saint-Vincent-de-Paul	48.88	2.36	Paris	10	0.00	Cluster	Asian_Restaurant	French_Restaurant	Hotel	African_Restaurant	1	0
Saint-Georges	48.88	2.33	Paris	9	0.00	Cluster	Hotel	French_Restaurant	European_Restaurant	Asian_Restaurant	1	0
Père-Lachaise	48.86	2.40	Paris	20	0.00	Cluster	Asian_Restaurant	Bar	Bakery	Bistro	1	0
Porte-Saint-Denis	48.87	2.35	Paris	10	0.00	Cluster	Asian_Restaurant	Hotel	French_Restaurant	Bakery	1	0
Petit-Montrouge	48.83	2.33	Paris	14	0.00	Cluster	Hotel	Asian_Restaurant	French_Restaurant	European_Restaurant	1	0
Charonne	48.85	2.41	Paris	20	0.00	Cluster	Supermarket	Asian_Restaurant	Bar	Hotel	1	0
Necker	48.84	2.31	Paris	15	0.00	Cluster	French_Restaurant	Asian_Restaurant	Hotel	Café	1	0
Montparnasse	48.84	2.33	Paris	14	0.00	Cluster	French_Restaurant	Hotel	European_Restaurant	Asian_Restaurant	1	0
Grenelle	48.85	2.29	Paris	15	0.00	Cluster	Hotel	French_Restaurant	Asian_Restaurant	Bistro	1	0
Gare	48.83	2.37	Paris	13	0.00	Cluster	Asian_Restaurant	Hotel	Sandwich Place	French_Restaurant	1	0
Epinettes	48.89	2.32	Paris	17	0.00	Cluster	Asian_Restaurant	French_Restaurant	Restaurant	Bakery	1	0
Croulebarbe	48.83	2.35	Paris	13	0.00	Cluster	Asian_Restaurant	French_Restaurant	Bar	Sandwich Place	1	0
Villette	48.89	2.37	Paris	19	0.00	Cluster	Hotel	Bar	Café	French_Restaurant	1	0