# Comparison of 3 Medium Size French cities profiles:

# Toulouse, Bordeaux and Nantes

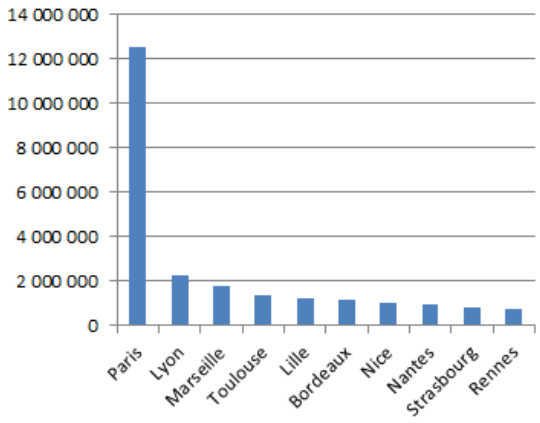
## 1. INTRODUCTION

### 1.1. Background

In March 2017, the population of France officially reached 67,000,000 people. As in many countries in the world in the last decades, French population has tended to densify around a limited set of big cities which, as a consequence, concentrate the highest level of economic activities of the country.

The table below [1] lists the ten largest metropolitan areas in France, based on their population at the 2015 census.

| Rank | Metropolitan Area | Population (2015) |
|---|---|---|
| 1 | Paris | 12,532,901 |
| 2 | Lyon | 2,291,763 |
| 3 | Marseille | 1,752,938 |
| 4 | Toulouse | 1,330,954 |
| 5 | Lille | 1,215,769 |
| 6 | Bordeaux | 1,184,708 |
| 7 | Nice | 1,005,891 |
| 8 | Nantes | 949,316 |
| 9 | Strasbourg | 780,515 |
| 10 | Rennes | 719,840 |




It appears clearly that Paris is, by far, the densest area of the French territory (representing up to 19% of the French population) followed by Lyon and Marseille respectively in second and third position. With a population of approximately 1 Million inhabitants, metropolitan areas from the 4$^{th}$ to the 10$^{th}$ rank can all be considered, at the French scale, as illustrations of mid-size cities.

For French people specificities of Paris, Lyon and Marseille are often well known. This is due to the fact that those cities get strong national focus thanks to news, cinema, national sport events (i.e. Soccer) or even lessons provided during history / geography class at school.

Mid-size cities get a lower attention and French people would probably have more difficulties to underline differences or similarities between cities like Toulouse or Nantes for example.

### 1.2. Problem

As we did with New York and Toronto, my idea is to perform data exploration of a set of French midsize cities and to determine whether the resulting segmentation and clustering reflects some similarities or specificities.

Data exploration will be performed for the cities of Toulouse, Bordeaux and Nantes. Apart from the size criteria, selection of those cities is also correlated to the availability of proper location datasets on open platforms (details provided in the next chapter). As a complement to venues data already tested in the labs, exploration will embed economic data (i.e. houses prices).

*Toulouse*

*Nantes*

*Bordeaux*

Studies will be limited to the administrative limits of the cities: suburbs will not be taken into account.

## 2. DATA SOURCES

Cities neighborhoods clustering performed for this project will mix economic data (houses prices) and information related to list of venues available near each city streets.

### 2.1. Economic data

Thanks to an open data platform managed by French government, nationwide houses sell prices for 2018 are freely available at: https://www.data.gouv.fr [2]. List of houses sold in the cities of Toulouse, Nantes and Bordeaux will be extracted from this datasets in order to get an estimate of average houses prices (Eur / $m^2$) for every street.

Unfortunately, location information contained in the previous dataset is limited to: street names, post code and city names. In order to perform proper data exploration, it will be necessary to collect latitude and longitude coordinates of every street of the cities of Toulouse, Nantes and Bordeaux. Hopefully datasets linking cities streets names to latitude & longitude are freely available from http://bano.openstreetmap.fr. Note that the French territory is divided into 101 administrative areas called “départements”, each of them being designated by a specific number. One file per administrative area is being accessible from openstreetmap.

Cities taken into consideration for this project are all belonging to distinct administrative areas: Toulouse (31), Nantes (44) and Bordeaux (33). This means that 3 specific datasets will have to be downloaded from openstreepmap [3][4][5]. After cleaning, this data will be merged to houses prices data in order to build specific data frames for every city containing: location information (street names, latitude & longitude) and corresponding average houses prices.

### 2.2. Venues List

Foursquare API will be used to list the most common venues near the streets of the 3 cities taken into consideration for this project [6]. Venues data collected with this API will be merged with houses prices data to be used as inputs to K-means clustering algorithm.

## 3. METHODOLOGY

### 3.1. Houses Prices Data frame

Thanks to an open data platform managed by French government, nationwide houses sell prices for 2018 are freely available from https://www.data.gouv.fr. List of houses sold in France in 2018 can then be recorded in a single csv file and converted to a data frame.

In order to process this large volume of data, I went through a set of preprocessing steps.

First, I started by building 3 specific data sets (one per city) containing information only related to the cities of Toulouse, Nantes and Bordeaux. Then I went through a series of data cleaning operations: Nan values removal, dropping columns containing irrelevant data for the project, data type conversion. Using surface and price data stored in 2 distinct columns, I built a new column showing, for every street of the cities, mean houses price per square meter.

Finally, for every city, we get a data frame containing: street name (Voie), city name (Commune) and price (in Eur) per m² (Price_m2)

| | Voie | Commune | Price_m2 |
|---|---|---|---|
| 0 | A BARBE TORTE | NANTES | 3728.560426 |
| 1 | ABBE DE L EPEE | NANTES | 29381.947131 |
| 2 | ABEL GANCE | NANTES | 2449.345100 |
| 3 | ADOLPHE MOITIE | NANTES | 3770.275922 |
| 4 | AGUESSE | NANTES | 5289.855072 |

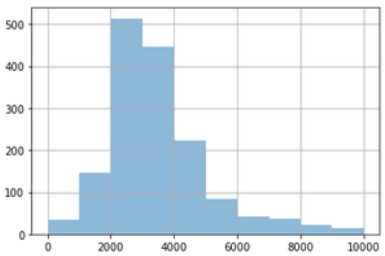
Then I applied additional data cleaning operations to this new data frame. First I noticed that some price per m² values were probably wrong or corrupted. As illustrated below for the city of Bordeaux, it appeared that some price / m² values weren't coherent at all.

```
count      1119.000000
mean       7411.815597
std       16816.618090
min           0.000805
25%        3933.464036
50%        4945.422266
75%        6460.866175
max      440653.488501
Name: Price_m2, dtype: float64
```

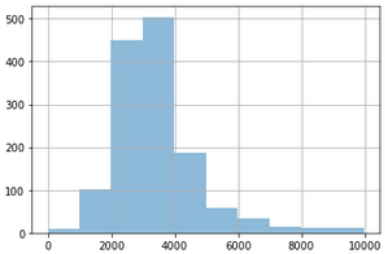
No one would reasonably pay more than 440 kEur for a single square meter! So I decided to suppress all rows containing price/m² above 10 kEur for the cities of Toulouse and Nantes and above 20 kEur for the city of Bordeaux.

Then I compared median prices per square meter between the 3 cities. It appears that median prices for Toulouse and Nantes cities are close to each other (respectively 3185 Eur / m² and 3216 Eur / m²) but the one related to the city of Bordeaux is much higher (4835 Eur / m²).

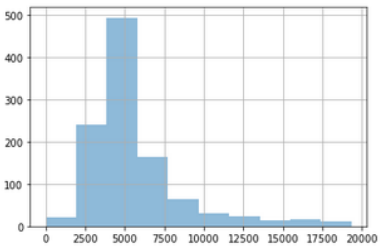
Then I noticed that the way prices are distributed is not homogeneous between the 3 cities (see histograms below):



Price per m² - Toulouse



Price per m² - Nantes



Price per m² - Bordeaux

As my ambition was to be able to compare patterns between the 3 cities, I needed to perform additional data manipulation steps in order to normalize price information.

Finally, using median prices, I decided to segment price per square meter data into 5 categories:

| Category | Description |
|---|---|
| Low | <= 0.4*median |
| Medium Low | > 0.4*median & <=0.8 * median |
| Medium | > 0.8*median & <=1.2 * median |
| Medium High | > 1.2*median & <=1.8 * median |
| High | > 1.8*median |

An example of the resulting data frame for the city of Bordeaux is given below:

| | Voie | Commune | Price_m2 | Houses_L | Houses_ML | Houses_M | Houses_MH | Houses_H |
|---|---|---|---|---|---|---|---|---|
| 1 | ACHARD | BORDEAUX | 8299.734327 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| 2 | ADRIEN BAYSSELANCE | BORDEAUX | 7509.405597 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| 3 | AGEN | BORDEAUX | 3980.925068 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 4 | ALBERT | BORDEAUX | 3680.622151 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 5 | ALBERT 1ER | BORDEAUX | 12545.375075 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |

### 3.2. Getting location data

As location data (latitude & longitude) is not contained in the datasets built during the previous steps, I had to collect such information from another data source: http://bano.openstreetmap.fr. Considering the fact that, on this platform, location data are segmented by French administrative areas, I had to download 3 data sets (Toulouse, Nantes and Bordeaux belong to distinct areas).

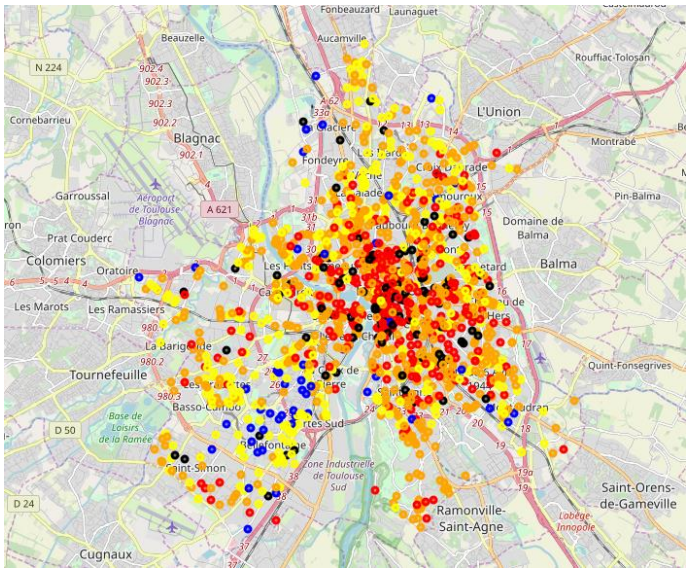
Then I had to do data cleaning: suppressing irrelevant data, suppressing French specific characters (accents), dropping Nan values, suppressing duplicated rows and unnecessary columns.

Finally I merged those new datasets with the ones containing houses prices to build a new pandas data frame linking street names, houses prices and latitude & longitudes coordinates, see example for the city of Nantes below:

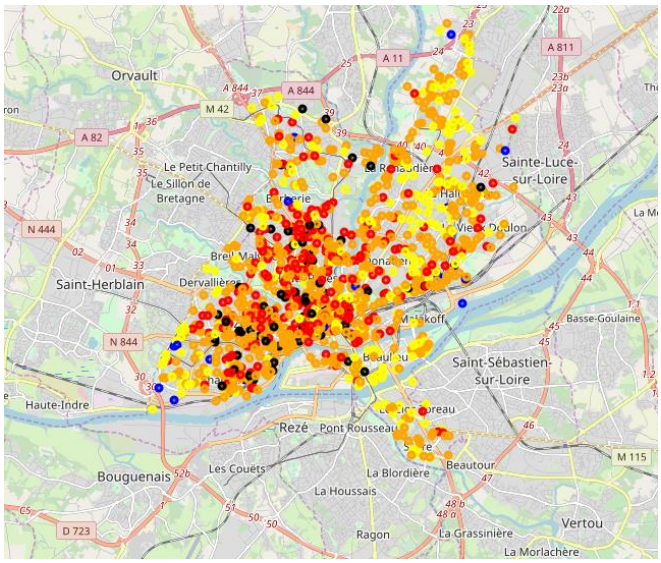
| Voie | Commune | Price_m2 | Houses_L | Houses_ML | Houses_M | Houses_MH | Houses_H | Latitude | Longitude |
|---|---|---|---|---|---|---|---|---|---|
| ABEL GANCE | NANTES | 2449.345100 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 | 47.247670 | -1.532205 |
| ADOLPHE MOITIE | NANTES | 3770.275922 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 47.222162 | -1.555209 |
| AGUESSE | NANTES | 5289.855072 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 | 47.212044 | -1.598330 |
| AINO AALTO | NANTES | 2985.507680 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 47.241994 | -1.509593 |
| ALAIN COLAS | NANTES | 3564.157706 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 | 47.210472 | -1.527245 |

### 3.3. Visualization on a Map

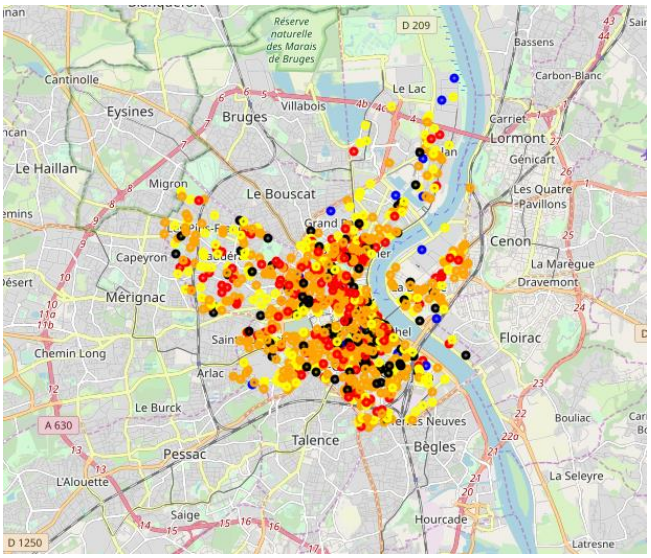
Thanks to those new data, I used the Folium library to visualize how house prices are distributed on the cities maps. I applied a color dictionary to distinguish belongings to house prices categories: **Blue** for Low, **yellow** for Medium Low, **Orange** for Medium, **Red** for Medium High and **Black** for High.




*House prices visualization for the city of Toulouse*




*House prices visualization for the city of Nantes*




*House prices visualization for the city of Bordeaux*

### 3.4. Getting venues data from Foursquare API

Like it was done in the labs, I used the Foursquare API to explore venues near every streets of each of the cities. Once collected, I stored venues data into 3 csv files.

Basic statistics related to data collected from Foursquare are given in the table below:

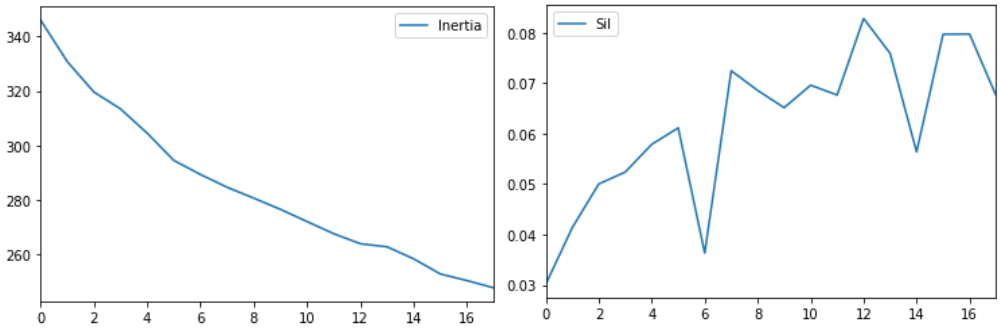
| City | Number of venues | Distinct venues categories |
|---|---|---|
| Toulouse | 9615 | 220 |
| Nantes | 9141 | 188 |
| Bordeaux | 8952 | 200 |

As my objective was to be able, at the end, to perform comparison between clusters of different cities I merged venues categories collected from the 3 cities into a single vector. This approach was a way to guarantee that K means centroids vectors defined for each of the 3 cities would have similar dimensions. The resulting “distinct venues” vector is made of 300 distinct rows.

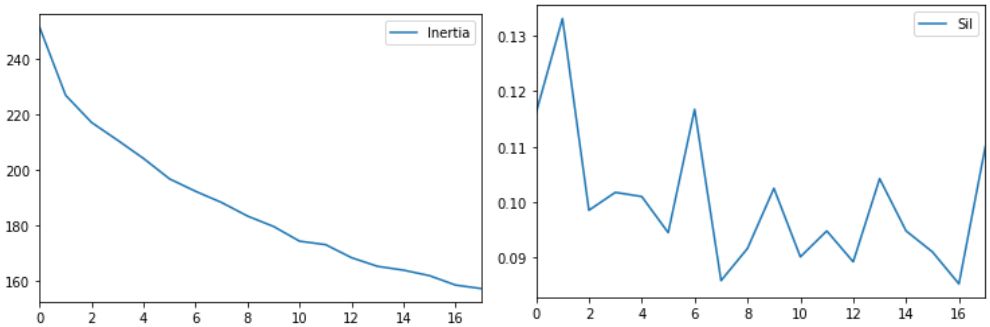
Next steps for me were to prepare data for K-means clustering: applying one hot encoding to venues, merging venues data with houses prices data, calculating the mean of the frequency of occurrence of each venue category.

### 3.5. K-means clustering

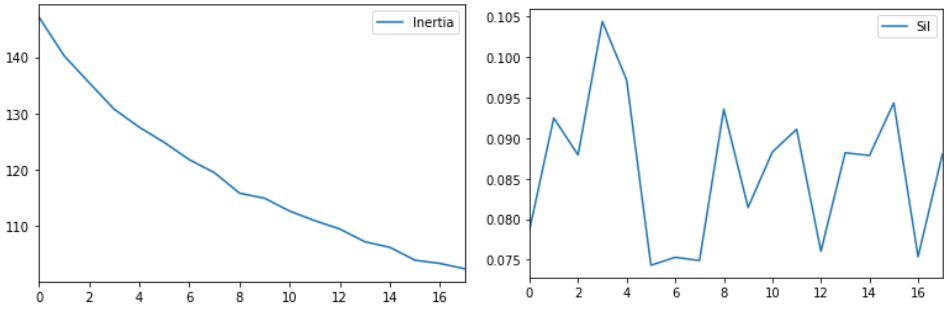
Then, K-means algorithm was applied to the 3 data frames containing houses prices and venues information (street by street segmentation). In order to select the proper number of clusters for each city, I ran the K-means algorithm multiple times and used Inertia and Silhouette metrics to select the proper values of K for the different cities.




*Inertia & Silhouette for the city of Toulouse*




*Inertia & Silhouette for the city of Nantes*




*Inertia & Silhouette for the city of Bordeaux*

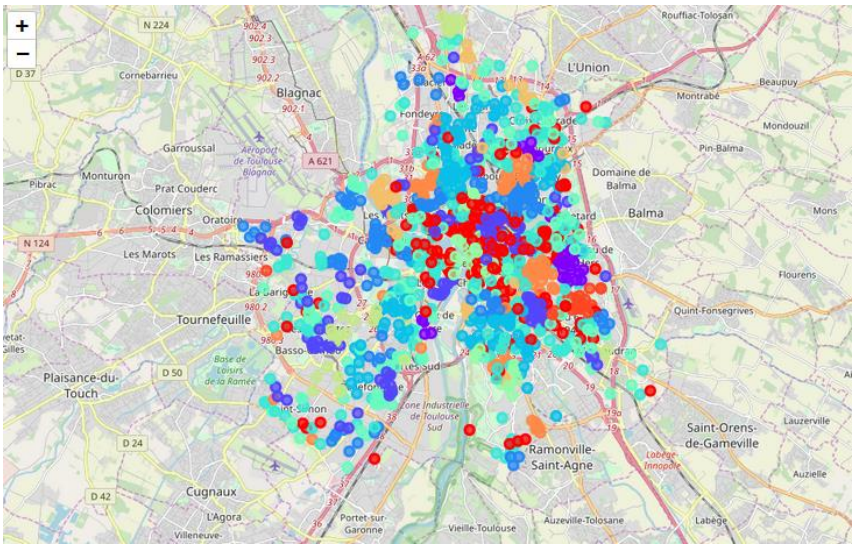
Number of clusters retained for the different cities is given in the table below

| City | Value of K |
|---|---|
| Toulouse | 12 |
| Nantes | 6 |
| Bordeaux | 3 |

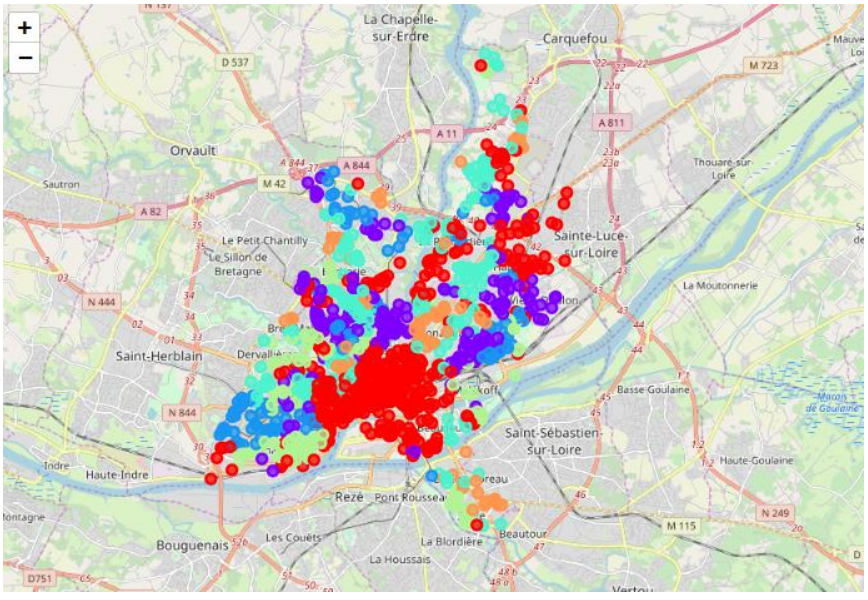
# 4. RESULTS

## 4.1. Clusters visualization

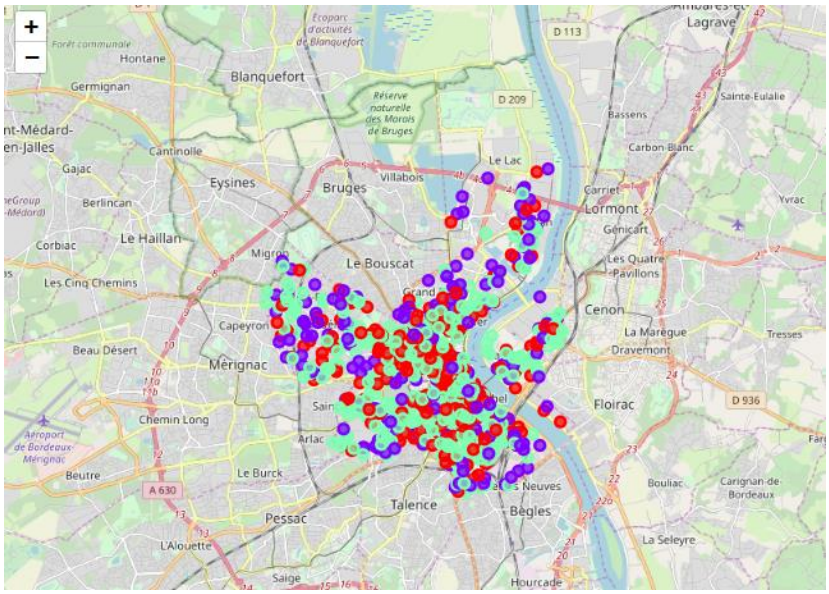
Again, I used the Folium library to visualize the repartition of obtained clusters on each city maps.




*Visualization of the 12 clusters obtained for the city of Toulouse*




*Visualization of the 6 clusters obtained for the city of Nantes*

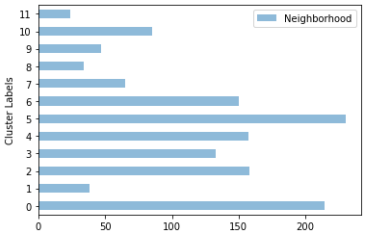



*Visualization of the 3 clusters obtained for the city of Bordeaux*

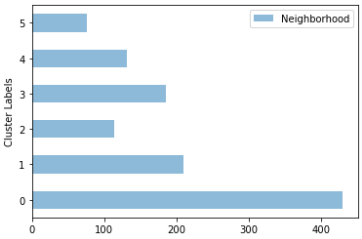
It appears that the spatial repartition of clusters differs from one city to the other. While we can notice that some clustering areas are quite dense for the cities of Toulouse and Nantes (i.e. dense red area close to the city center of Nantes), repartition of clusters for the city of Bordeaux is closer to a homogeneous distribution of single nodes over the city map.

### 4.2. Clusters distribution

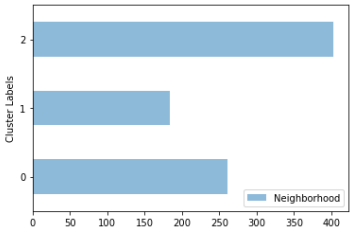
It seemed interesting to me to determine, for each city, the number of neighborhoods (one neighborhood = one city street) contained in each cluster. This distribution is illustrated by the horizontal bar plots below:




*NB of streets / cluster- Toulouse*




*NB of streets / cluster- Nantes*




*NB of streets / cluster- Bordeaux*

It appears that the 2 biggest groups identified for the city of Toulouse are clusters 0 and 5 (each of them containing more than 200 streets) followed by a group made of medium size clusters (2, 3, 4 and 6).

Clustering applied to the city of Nantes shows a big cluster (Number 0) followed by 5 other groups of similar size.

Finally, the biggest cluster identified for the city of Bordeaux is number 2 (containing 400 streets) while the size of the 2 other clusters is quite homogeneous.

### 4.3. Clusters comparison

The last step of my project was to identify if similarities are existing between the clusters of the cities of Toulouse, Nantes and Bordeaux. To do this, I decided to measure Euclidian distance between centroids of the different clusters. This was made possible because I was vigilant to get centroids vectors of similar size.

Clusters will be considered as similar if their inter centroids distance (Euclidian) is close to zero.

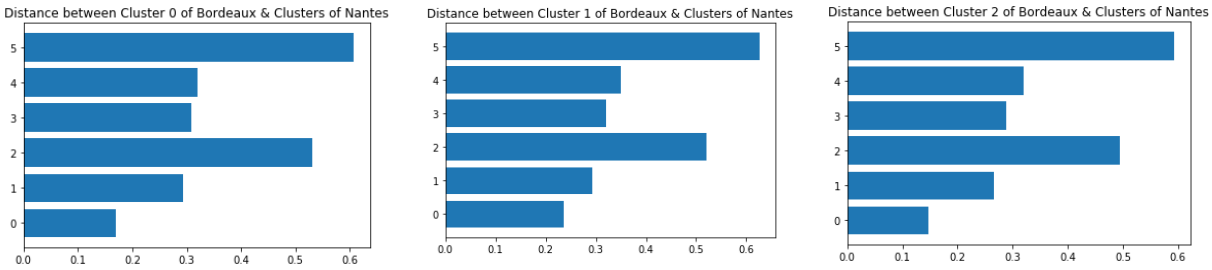
As an example, the array illustrated below measures the distance between Bordeaux Cluster 2 and the different clusters of the city of Nantes. It appears that cluster 2 of Bordeaux seems quite similar to cluster 0 defined for Nantes.

```
dist_bordeaux_nantes[2,:]

array([0.14710909, 0.2667333 , 0.49529797, 0.28877924, 0.3195258 ,
       0.5923125 ])
```

Bordeaux to Nantes

Bar plots given below measure inter centroids Euclidian distances between the different city clusters of Bordeaux and Nantes.




*Bordeaux Cluster 0 - Nantes* *Bordeaux Cluster 1 - Nantes* *Bordeaux Cluster 2 - Nantes*

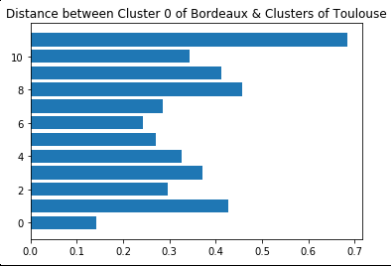
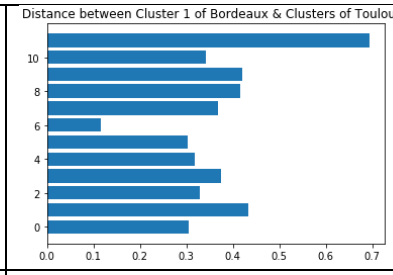
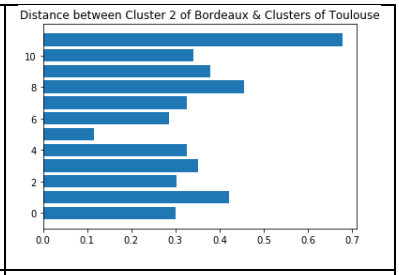
Detailed distances between Bordeaux and Nantes clusters are reported in the numpy array below:

```
[[0.17002824 0.29292572 0.53110303 0.30798524 0.32007478 0.60604286]
 [0.23621138 0.29317079 0.52011049 0.32020786 0.34922881 0.62610063]
 [0.14710909 0.2667333  0.49529797 0.28877924 0.3195258  0.5923125 ]]
```

It appears that cluster 2 of Bordeaux and cluster 0 of Nantes are the closest. Note that Clusters 0 of Bordeaux and Nantes are also close to each other.

Bordeaux to Toulouse

Bar plots given below measure inter centroids Euclidian distances between the different city clusters of Bordeaux and Toulouse.

| | | |
|---|---|---|
|   |   |   |
| *Bordeaux Cluster 0 - Toulouse* | *Bordeaux Cluster 1 - Toulouse* | *Bordeaux Cluster 2 - Toulouse* |

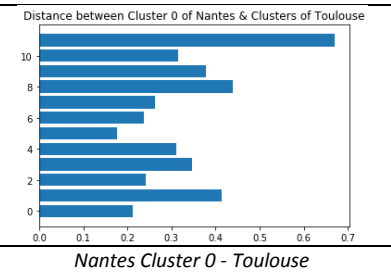
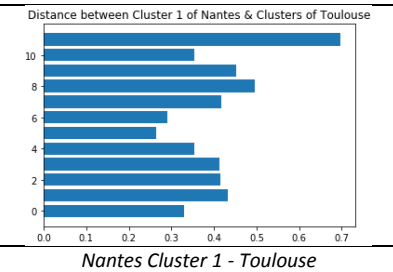
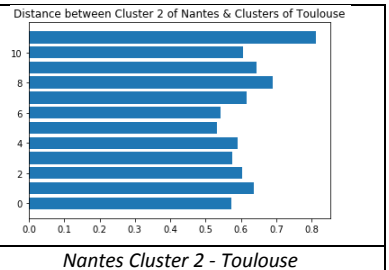
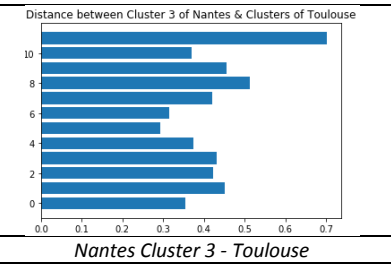
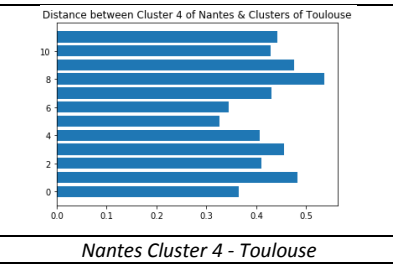
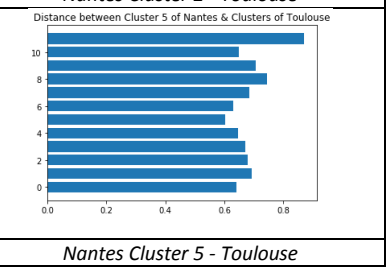
Detailed distances between Bordeaux and Toulouse clusters are reported in the numpy array below:

```
Distance between clusters of Bordeaux and Toulouse arrays
[[0.14168713 0.42638545 0.29734887 0.37204503 0.32613901 0.27121412
  0.24364566 0.28559836 0.45712473 0.41178741 0.34255309 0.68324224]
 [0.3048055  0.43295936 0.32998458 0.37383596 0.31760542 0.30248239
  0.11668045 0.36828753 0.41568827 0.42081483 0.34304218 0.69323313]
 [0.30057735 0.4212734  0.30344083 0.35077005 0.32496757 0.11482059
  0.28622894 0.3267444  0.45439904 0.3797324  0.34064161 0.67748579]]
```

It appears that cluster 2 of Bordeaux and cluster 5 of Toulouse are the closest. Note that cluster 1 of Bordeaux and Cluster 6 of Toulouse are also really close to each other.

## Nantes to Toulouse

Finally, inter cluster centroids distances for the cities of Nantes and Toulouse are illustrated by the diagrams below.

| | | |
|---|---|---|
|   |   |   |
| *Nantes Cluster 0 - Toulouse* | *Nantes Cluster 1 - Toulouse* | *Nantes Cluster 2 - Toulouse* |
|   |   |   |
| *Nantes Cluster 3 - Toulouse* | *Nantes Cluster 4 - Toulouse* | *Nantes Cluster 5 - Toulouse* |

Detailed distances between Nantes and Toulouse clusters are reported in the numpy array below:

```
Distance between clusters of Nantes and Toulouse arrays
[[0.21221461 0.41391515 0.24245469 0.34618536 0.31034336 0.17713935
  0.23683115 0.26355877 0.43869887 0.37828784 0.31501181 0.66940947]
 [0.33054774 0.43278539 0.41497542 0.4132317  0.35468857 0.26503781
  0.2910292  0.4162905  0.49657387 0.45208336 0.35347569 0.69639062]
 [0.57234348 0.63752819 0.6046826  0.57475678 0.59086221 0.53140151
  0.54136508 0.61722565 0.68940776 0.64427511 0.60585494 0.81168802]
 [0.35445408 0.45189675 0.422665   0.4311263  0.3736848  0.29233363
  0.3157696  0.42011839 0.51304252 0.45661938 0.36944194 0.70190652]
 [0.36477653 0.48310453 0.41087769 0.45511391 0.4069293  0.32698255
  0.34387098 0.42998537 0.53599531 0.47626518 0.42920948 0.44256876]
 [0.64094091 0.69402668 0.67882376 0.67177188 0.64597027 0.60198605
  0.62957575 0.68561387 0.74427264 0.70643217 0.6495357  0.86940617]]
```

It appears that cluster 0 of Nantes and cluster 5 of Toulouse are the closest. All other clusters have low affinity.

### 4.4. Deeper Clusters comparison

Let's have a closer look at the content of clusters having high similarities. Let's take the example of Nantes cluster 0 and Toulouse cluster 5.

| | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | Commune | Price_m2 | Houses_L | Houses_ML | Houses_M | Houses_MH | Houses_H |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | Soccer Stadium | Food Truck | French Restaurant | Shopping Mall | Farmers Market | NANTES | 2449.345100 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 1 | Bakery | Bar | Café | Farmers Market | Beer Garden | NANTES | 3770.275922 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 3 | Supermarket | Rental Car Location | French Restaurant | Soccer Stadium | Zoo | NANTES | 2985.507680 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 8 | Hotel | Construction & Landscaping | Department Store | Food | Farmers Market | NANTES | 4633.152174 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| 12 | French Restaurant | Bike Rental / Bike Share | Plaza | Tram Station | Jewelry Store | NANTES | 3146.002758 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1139 | Bar | Ice Cream Shop | Trail | Indian Restaurant | French Restaurant | NANTES | 2192.810595 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 1141 | Construction & Landscaping | Gym | Pizza Place | Flower Shop | Falafel Restaurant | NANTES | 1458.620690 | 0.0 | 1.0 | 0.0 | 0.0 | 0.0 |
| 1142 | Bar | Plaza | French Restaurant | Dessert Shop | Lounge | NANTES | 3263.204747 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 1144 | Plaza | Bus Stop | French Restaurant | Wine Bar | Bistro | NANTES | 5337.837838 | 0.0 | 0.0 | 0.0 | 1.0 | 0.0 |
| 1145 | French Restaurant | Plaza | Bakery | Bus Stop | Gym | NANTES | 6964.818182 | 0.0 | 0.0 | 0.0 | 0.0 | 1.0 |

*Above : Content of Nantes – cluster 0*

| | 1st Most Common Venue | 2nd Most Common Venue | 3rd Most Common Venue | 4th Most Common Venue | 5th Most Common Venue | Commune | Price_m2 | Houses_L | Houses_ML | Houses_M | Houses_MH | Houses_H |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 3 | Hotel | Pub | Tram Station | Bus Stop | Ramen Restaurant | TOULOUSE | 3102.086745 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 4 | Plaza | Pub | Metro Station | Gastropub | Gym | TOULOUSE | 3719.427095 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 11 | Bakery | French Restaurant | Brewery | Metro Station | Restaurant | TOULOUSE | 2834.903520 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 12 | French Restaurant | Bar | Pizza Place | Japanese Restaurant | Comedy Club | TOULOUSE | 3637.552083 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 15 | Supermarket | Park | Shopping Mall | Funeral Home | Zoo | TOULOUSE | 3560.975610 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 1307 | Farm | Train Station | Bus Stop | Zoo | Event Service | TOULOUSE | 2993.472680 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 1308 | Motorcycle Shop | Gym | Italian Restaurant | Chinese Restaurant | Farmers Market | TOULOUSE | 3579.710145 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 1312 | French Restaurant | Restaurant | Dance Studio | Deli / Bodega | Sandwich Place | TOULOUSE | 2848.677249 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 1315 | Japanese Restaurant | Furniture / Home Store | Pool | Restaurant | Dance Studio | TOULOUSE | 2691.765152 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |
| 1334 | Business Service | Grocery Store | Gym | Flower Shop | Flea Market | TOULOUSE | 3125.085872 | 0.0 | 0.0 | 1.0 | 0.0 | 0.0 |

*Above : Content of Toulouse – cluster 5*

Well, to be honest, it's not easy to identify similarities at the first sight. Having a look at venues and houses prices may help to define this cluster as middle class residential areas (houses prices being mainly medium or medium high).

## 5. DISCUSSION

As stated as a presumption this project shows that similar patterns can be identified between French midsize cities thanks to K-Means clustering algorithm. The large number of columns (over 300) characterizing the clusters added to the large number of samples (rows) linked to each cluster highlight that inter-clusters comparison cannot be performed by a single human brain.

The usage of mathematical metrics appears as an absolute necessity to get an objective way of measuring similarities between clusters. In our case, Euclidian distance between clusters centroids was used as the main indicator of similarities.

The methodology applied to the project can easily be extended to other cities to identify existing similarities between other French main cities.

As a next step to this project, getting a deeper focus on each cluster could contribute to assign labels to each of them (i.e. "residential area", "shopping area", etc.) with the objective, at the end, to have a more explicit vision of the inner structure of the cities.

## 6. CONCLUSION

Running this project was really exciting for me as it offered me real opportunities to apply what I learned during this class on real problems and real data.

This project gives precious inputs to investors as it highlights that similarities are existing between cities. This could help them to better target their business by taking benefits of economies of scales by running similar projects on similar city clusters.

## 7. REFERENCES

[1] – Wikipedia - Demographic of France -
https://en.wikipedia.org/wiki/Demographics_of_France#Demographic_statistics

[2] – List of houses sold in 2018 on French Metropolitan area
https://www.data.gouv.fr/fr/datasets/r/1be77ca5-dc1b-4e50-af2b-0240147e0346

[3] – Location data for the city of Toulouse (administrative area n°31)
http://bano.openstreetmap.fr/data/bano-31.csv

[4] – Location data for the city of Nantes (administrative area n°44)
http://bano.openstreetmap.fr/data/bano-44.csv

[5] – Location data for the city of Bordeaux (administrative area n°33)
http://bano.openstreetmap.fr/data/bano-33.csv

[6] – Foursquare API – Venues list
https://api.foursquare.com/v2/venues/explore?