Comparison of 3 Medium Size French cities profiles:

Toulouse, Bordeaux and Nantes

2. DATA SOURCES

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

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 [1]. 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 / m²) 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 [2][3][4]. 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 [5]. Venues data collected with this API will be merged with houses prices data to be used as inputs to K-means clustering algorithm.

2.3. References

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

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

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

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

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