Luxury furniture in Brussels (Belgium) - Marketing campaign

Contents

[Luxury furniture in Brussels (Belgium) - Marketing campaign 1](#_Toc1809949)

[1. Introduction 1](#_Toc1809950)

[2. Data 1](#_Toc1809951)

[3. Methodology 2](#_Toc1809952)

[3.1. K-Means 2](#_Toc1809953)

[4. Results 2](#_Toc1809954)

[4.1. Marketing Campaign 2](#_Toc1809955)

[4.2. Launch of store in particular location 3](#_Toc1809956)

[5. Discussion 4](#_Toc1809957)

[5.1. Marketing Campaing 4](#_Toc1809958)

[5.2. Store location 4](#_Toc1809959)

[6. Conclusion 4](#_Toc1809960)

# Introduction

In this section, we will provide a description of the problem, a discussion of the background and who would be interested in this project.

An international seller of luxury furnishing wants to:

* launch a marketing campaign in Belgium
* create a few stores in Belgium (if it estimates the market is not packed yet of similar stores)

For that purpose, it wants first to identify on a map where its potential customers are clustered (i.e. spot the highest number of potential customer per postal code).

The potential customers share the following features:

* Age: 18-44 years
* High propensity to buy luxury furniture’s assessed based on housing price
* Owner of the housing (these are assumed to be more likely to buy luxury furniture’s)

Second, for the launch of its stores, it wants to avoid locations where similar stores already exist. It will accept to create a store in a location only if the number of similar stores per potential customer (within a certain radius) is below a certain threshold.

In this project, we will show on a map where its potential clients are mostly concentrated. To this end, we will color postal codes based on a scale of low number of potential clients (Dark Blue) to high number of potential clients (Dark Red). Postal codes for which the assessment could not be made will be marked in Grey.

# Data

Let’s have a look at the data we will be using to solve the problem and their source.

The number of residence and the percentage of owner per postal code will be used to determine the number of resident per postal code. Combining this data with the percentage of 18-44 years per postal code among total population, we will estimate a **proxy** of the percentage of **owner of the age of 18-44 years** per postal code.

The average price of housing per postal code will be used to identify potential clients which are likely to buy luxury furtnitures. We assume a positive correlation between housing price and the likelihood of buying luxury furnitures.

Foursquare will be used to have a rough estimate of the number of existing furniture’s stores in specific locations and whether the location is already packed of similar stores.

Finally, Geospatial Coordinates of postal code in Belgium will be used to plot postal codes on the map in Python.

|  |  |  |
| --- | --- | --- |
| Data | Data file | Source |
| - Number of residence  - Percentage of owner (per postal code) | logement occupé par proprio\_2011.csv  logement occupé par proprio\_2011.json | Statbel (2011)[[1]](#footnote-1) |
| - Percentage of 18-44 years per postal code among total population | TF\_SOC\_POP\_STRUCT\_2015\_Age.xlsx | Statbel |
| Average Price (per postal code) | immo\_by\_municipality\_2010-2019.xlsx | Statbel |
| *Housing features (per region only)* | *ImmoFeatures\_SdB\_Rooms\_Noccupants\_2011.xlsx* | *Statbel* |
| *Average renting price (per postal code in Brussels only) – won’t be used* | *Loyer\_Brussels\_2017.json*  *Loyer\_Brussels\_2017.xls* | *Statbel* |
| Geospatial Coordinates of postal code in Belgium | Geospatial\_Coordinates\_Belgium.csv |  |
| Number of furniture stores per region/postal code |  | Foursquare |

# Methodology

This section is the main component of the report. Exploratory analysis, inferential statistical testing and/or machine learning methods used in the context of the project are discussed and described.

# K-Means

As explained in the introduction, we are trying to segment the population based on their propensity to buy expensive furnitures.

For this purpose we are non-supervised machine learning technique called K-means in order to segment the population based on their features contained in a dataset we prepared in the pre-processing phase.

Note that in the dataset, we do not have any indication on whether their propensity to buy is high or low. This is the reason why we use unsupervised ML technique.

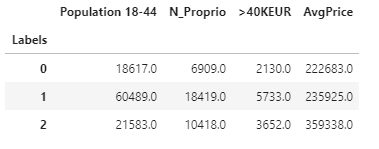
The explanatory variables we will look at are:

* 'Population 18-44': the number of persons in the neighbor(i.e; postal code) which have an age we deem is comprised in our target range
* 'N\_Proprio': the number of housing owners: we assume housing owners are wealthy people which are likely to invest in high-end furniture (people that rent a place are less likely to buy expensive furniture as we assume they are less likely to stay in their housing for long term)
* ,'>40KEUR: the number of people that earn more than 40k€/year in the neighbor. We assume people that earn less are less likely to buy expensive furniture
* ','AvgPrice': a measure to assess the propensity and capacity of the population to buy expensive furntirues is the avgprice of their house
  + we assume the higher the price, the bigger the house and hence the more furniture will be bought.
  + We assume the higher the price, the richer are the inhabitants

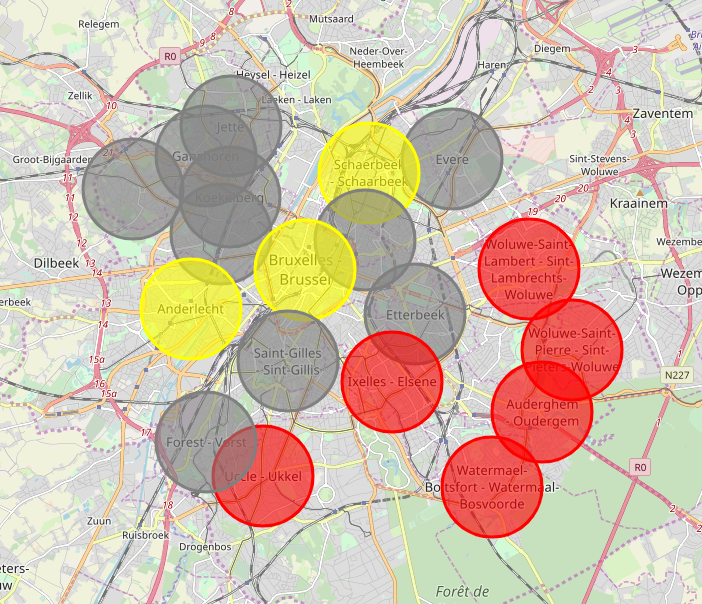
# Results

# Marketing Campaign

The results of the clusters are the following



* **Label\_0 (low potential)** : lowest number of young people and do not seem really wealthy: lowest # of 18-44y people, lowest number of proprio lowest # of salary >40kEUR and lowest avg price of houses
* **Label\_1 (med-high potential – YELLOW CIRCLE)**: Highest number of young people and relatively wealthy (based on avg price of houses)): 1st for young people, first for # owners , first for # of salary >40kEUR but people seem to live in average in less pricy homes ( we assume therefore less propensity to buy expensive furnitures than people in Cluster labelled 2).
* **Label\_2 (high potential – RED CIRCLE)**: Wealthy (based on avgprice of houses and # of >40kEUR salary) but few young people: above Label 0 for # of young people and number of owners, second for # of salary >40kEUR and first for avg price house



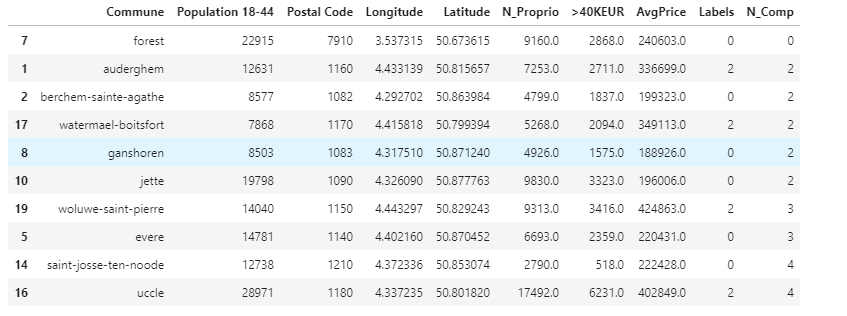
# Launch of store in particular location

In the table below, we have sorted the Commune (= postal codes) in ascending orders of the number of furniture store in a radius of 5KM (N\_Comp). The results are based on a query on foursquare.

Please note that the query results seem much lower than expected. I assume foursquare is not complete for ming queries for furniture store (compared to restaurants bar etc..) Anyhow…. We will assume the results reflect the reality…

What we see that forest has no furniture store in a radius of 5km (0 in column N\_Comp). However, its cluster category is 0 (**low potential).**

Auderghem and Watermael-boistfort have few potential competitors (only 2) and their cluster is labelled 2. Also there are located close to other Cluster 2 neighborhoods. We will therefore prefer these two neighbors to launch our luxury furniture store



# Discussion

Based on these observations, it is difficult to make a recommendation between Label\_1 and Label\_2. On one hand Label\_2 seems to include the most wealthy people, but on the other hand, there is a lower amount of people earning >40k than for Label\_1.

# Marketing Campaing

For the marketing campaign, we can afford to target several locations.

Therefore, we will target all neigbors located in Red and Yellow zone.

# Store location

For the launch of our store, we cannot afford to target several locations. We have to choose one location. So which on to choose?

* All in all, we still recommend to target Label\_2 neighborhood - as they seem to be the wealthiest and have still a significant amount of population earning >40kEUR .
* We recommend to target a location which is close to other Label\_2Neigborhood
* We recommend to target a location which has low number of competitor in radius 5km

As a result we decide to launch the store in Auderghem, which has a relatively low number of competitor (only 2) and is clustered in the Label\_2 (high potential –red) zone.

# Conclusion

As a conclusion, we think there is a market in Brussels to launch 1 store in Auderghem neighborhood.

Also, we think it is worth making a marketing campaign in the neigbors labelled in Red and Yellow or the following reasons:

* In the red clusters, 18-44y population seems to live in houses with higher prices
* In the Yellow clusters, there is the highest number of 18-44y population with relatively important number of them which earns a salary > 40kEur/year. The avg price of house is quite

For the reasons discussed in Section 5.2, we decide to launch the luxury furtniture store in Auderghem.

1. <https://bestat.statbel.fgov.be/bestat/crosstable.xhtml?view=a14f782c-353f-4f1b-97b3-995b8a435b69> [↑](#footnote-ref-1)