# 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

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 but 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.

xxx

# Results

The results of the research are the following

Xxx

# Discussion

Discussion section where you discuss any observations you noted and any recommendations you can make based on the results

# Conclusion

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