ImmoEliza Business plan

By Graciela Lopez Rosson, Michel Ombessa, MinhHien & Anne Jungers

TABLE OF CONTENT

I. OUR MISSION

GOAL :

→ To make ImmoEliza the biggest real estate company in Belgium !!

HOW?

- → They can become the biggest company by knowing and selling the most valuable properties in Belgium

FINAL PURPOSE :

→To create a machine learning model to predict prices on Belgium's sales.

- → Once they know where and what are the most valuable properties they can pick out them to sell them
- → With this model, ImmoEliza will increase its turnover of about 300% and will become the first real estate player in Belgium

II. MAIN QUESTIONS

- Identify the target variable
- Identify the variables to correlate with the target variable
- Identify the most valuable properties per region:
  - Wallonia
  - Bruxelles
  - Flanders

III. ANALYSIS

1). Intro analysis

- For our analysis we used a cleaned dataset called: ‘immo_eliza_dataset.csv’

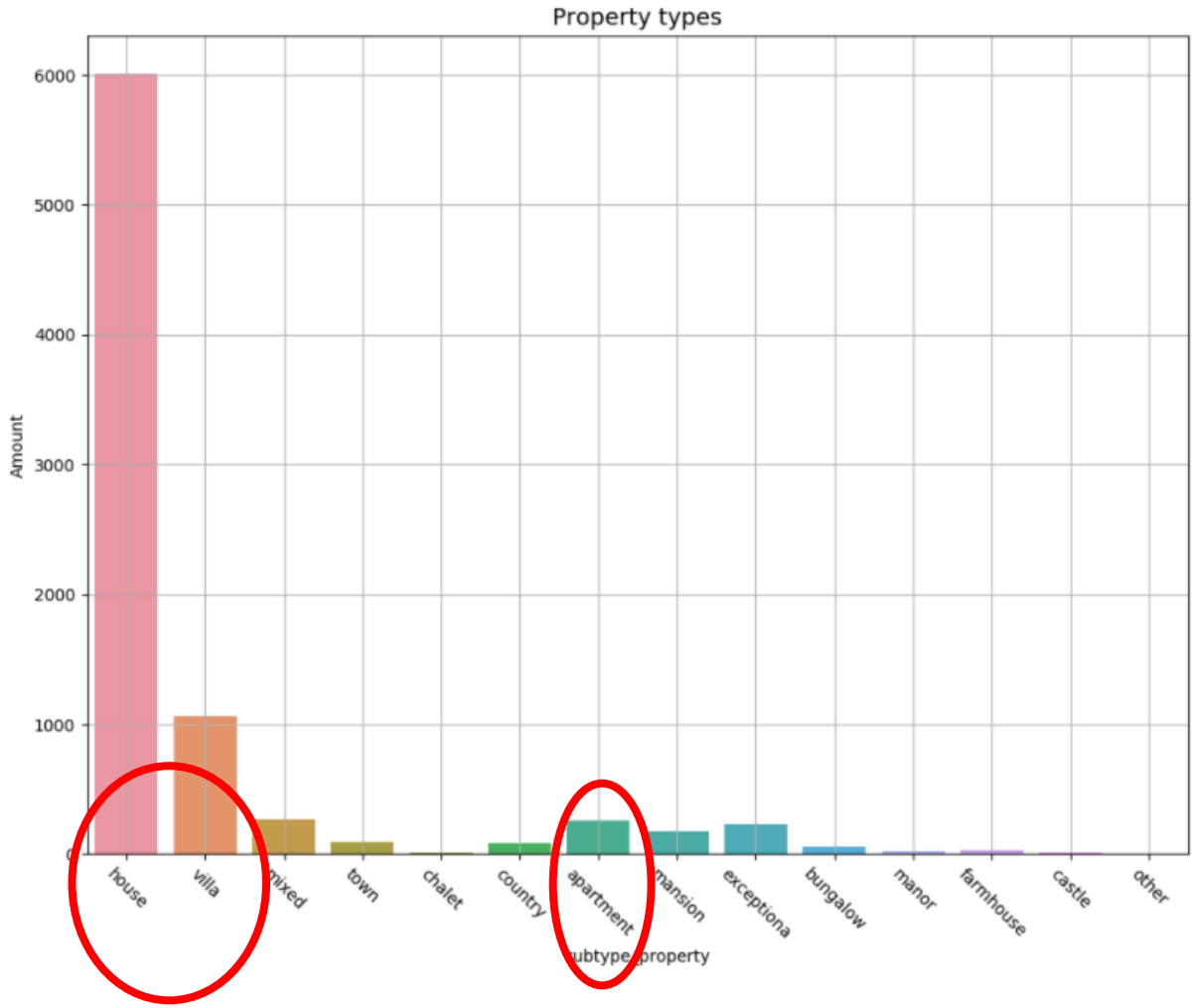
- The target variable to correlate with other variables is: THE PRICE

→ We want to know what are the variables that most influence the price

- The tested variables are:

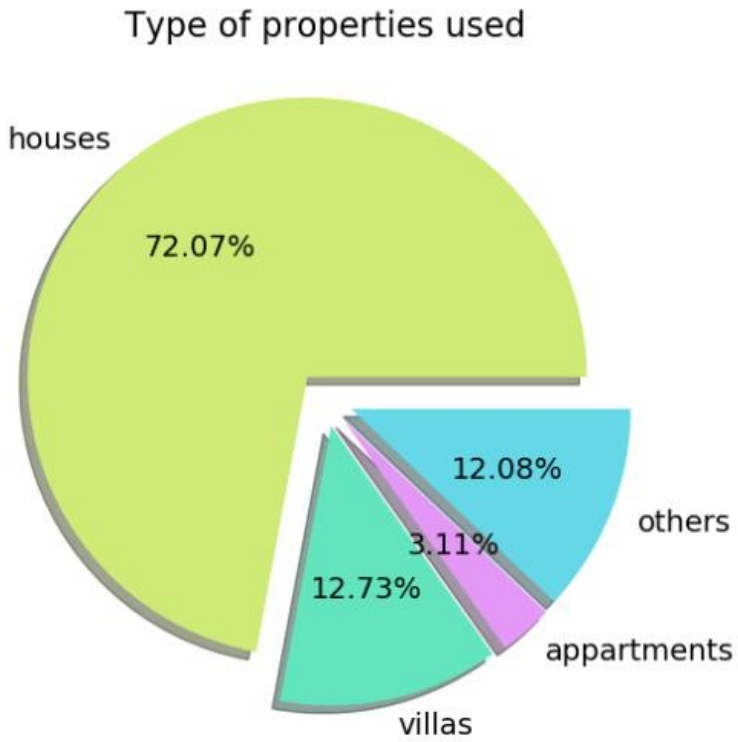
- Area, Number of bedrooms, location, type of properties, subtype

Focus on three types of properties: House, villa, apartment



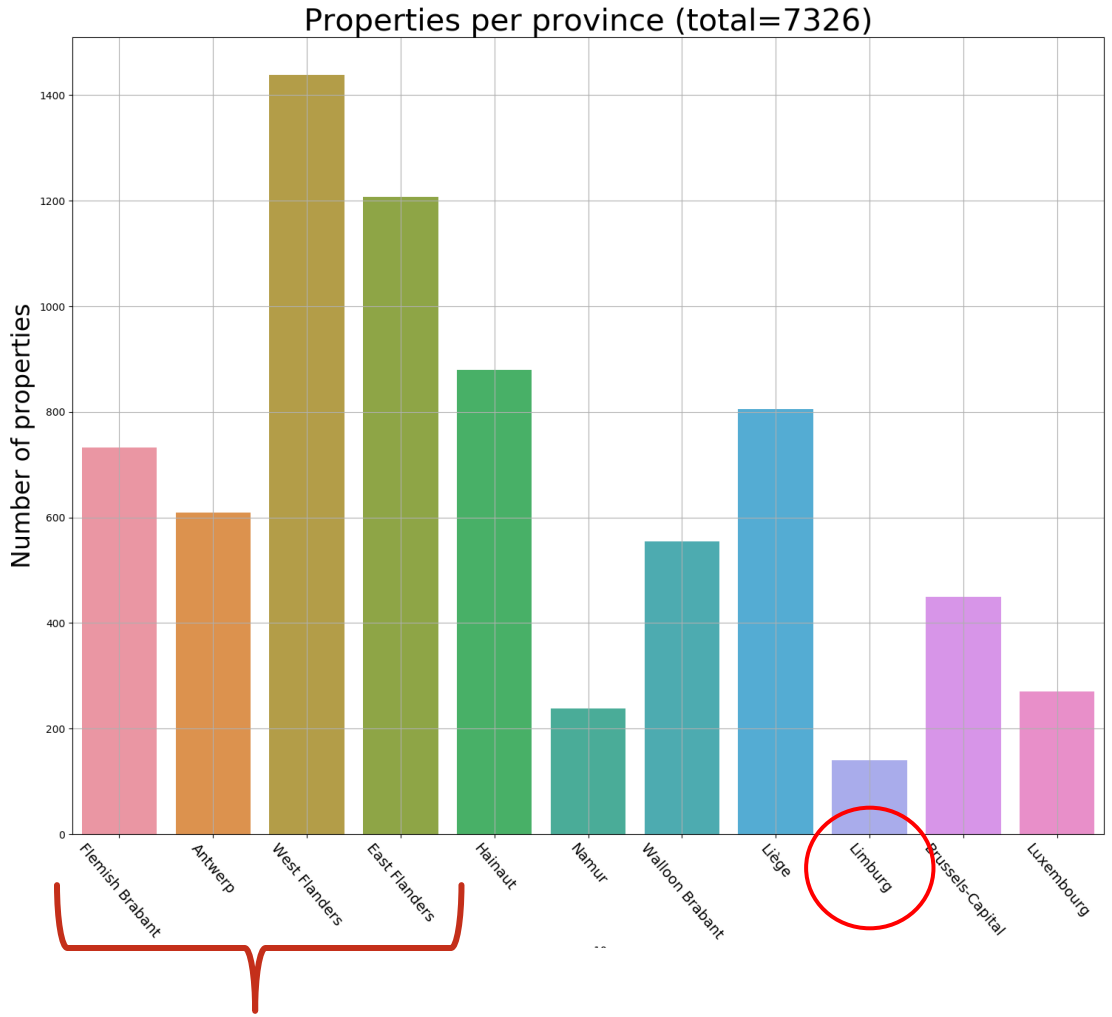
Reason:
- houses, villas and apartment represent the biggest part of the market

To have a better understanding of the proportion notion → circular diagram!!



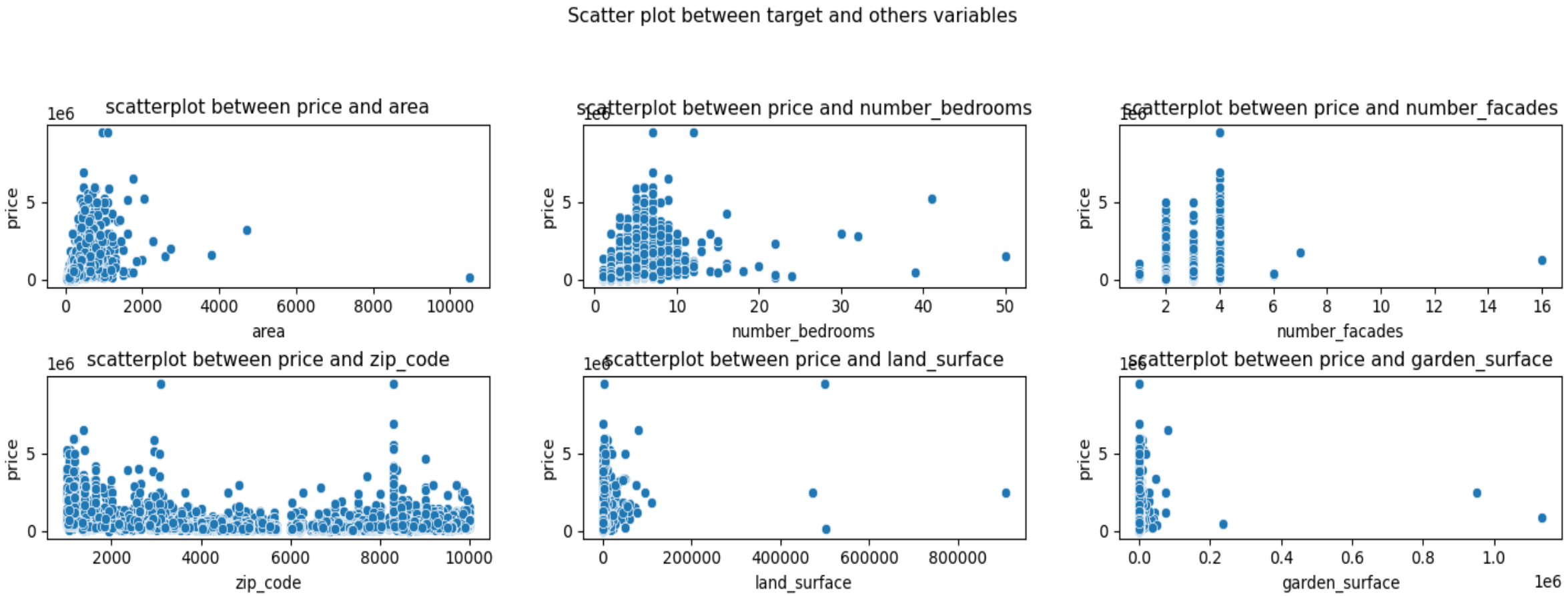
As you can clearly see here, the “other properties” represent only 12%

Histogram properties per province



Flanders has the biggest amount of properties the in dataset

Scatterplots to find correlations between the main variable “price” and the other variables



IV. RESULTS

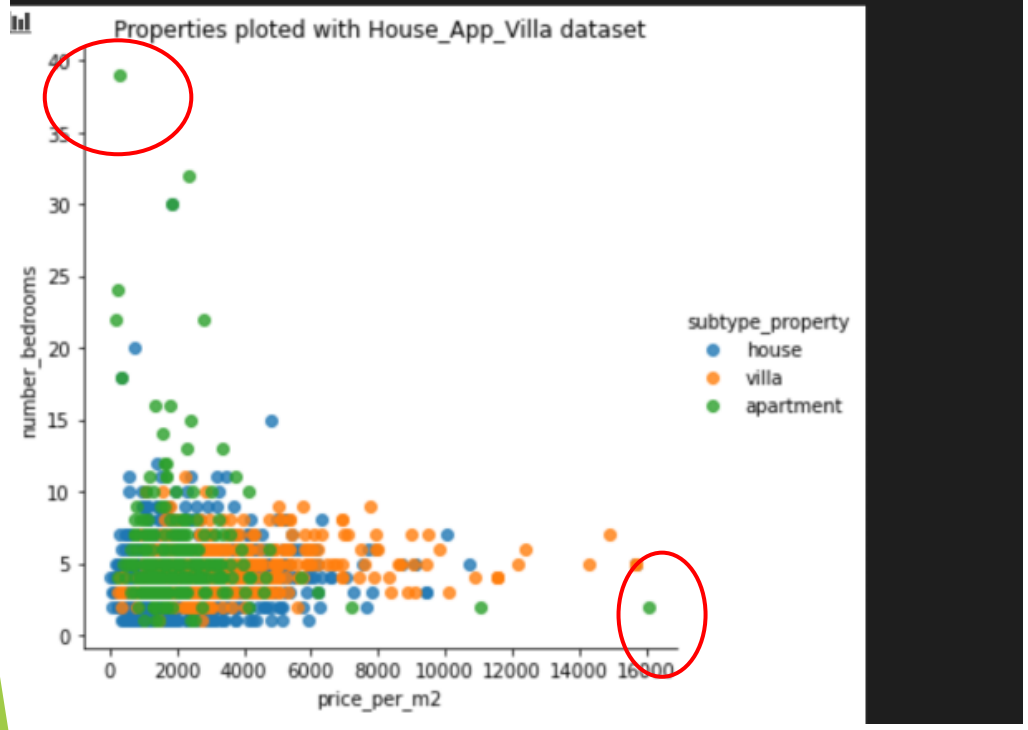
- Are there any outliers? If yes, which ones and why?
- Which variables you deleted and why ?
- Which 5 variables are the most important and why?
- What are the most expensive province in Belgium, Flanders & Wallonia?
- What are the less expensive regions in Belgium, Flanders & Wallonia?

Outliers? If yes, which ones and why?

```
sns.lmplot(y='number_bedrooms', x='price_per_m2', data=interesting_prop,
           fit_reg=False,
           hue='subtype_property')

plt.title("Properties ploted with House_App_Villa dataset")
```

```
ext(0.5, 1.0, 'Properties ploted with House_App_Villa dataset')
```



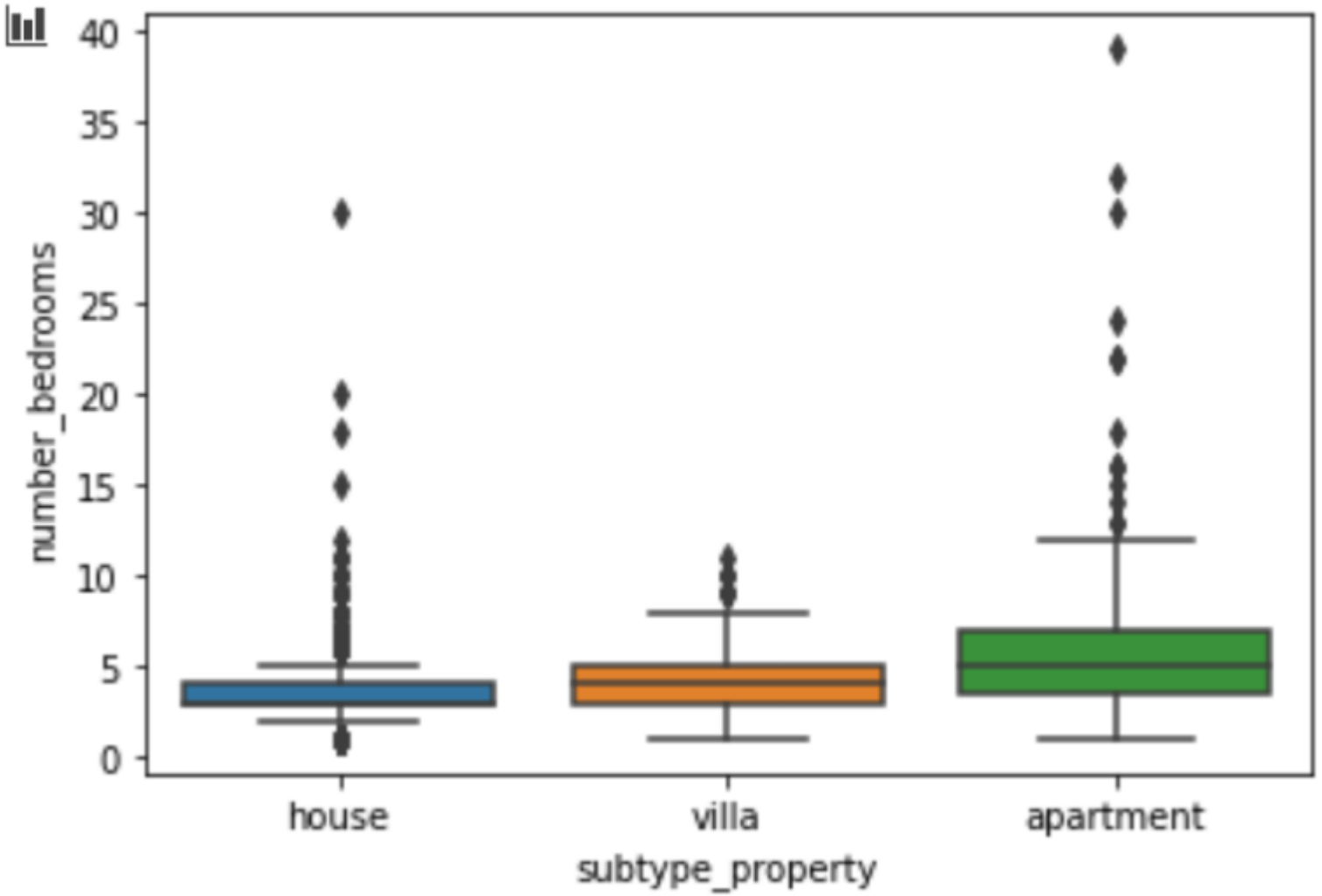
Here is an example of outliers.

Our interpretation:

- An apartment with almost 40 bedrooms seems to be an error of classification of
- If it is effectively like this it is probably difficult to sell

- On the opposite of the plot, an apartment 1 room with a very high price per m2 is also very unlikely, so it is also probably an error, or except if this apartment is located in a very exceptional situation

Other example of outlier



Which variables you deleted and why ?

We removed the following variables of the dataset:

- State of the building
- Number of facades,
- Fully equipped kitchen
- Open fire
- Surface of the land
- Terrace
- Swimming pool
- Furnished
- Garden
- Garden surface

Why:

- Not enough relevant
- No influence on the price
- Not enough data for the variable
- We wanted to bring clarity in our analysis by only testing few variables

Which 5 variables are the most important and why?

- Area
- Number of bedrooms
- Location
- Type of property
- Subtype

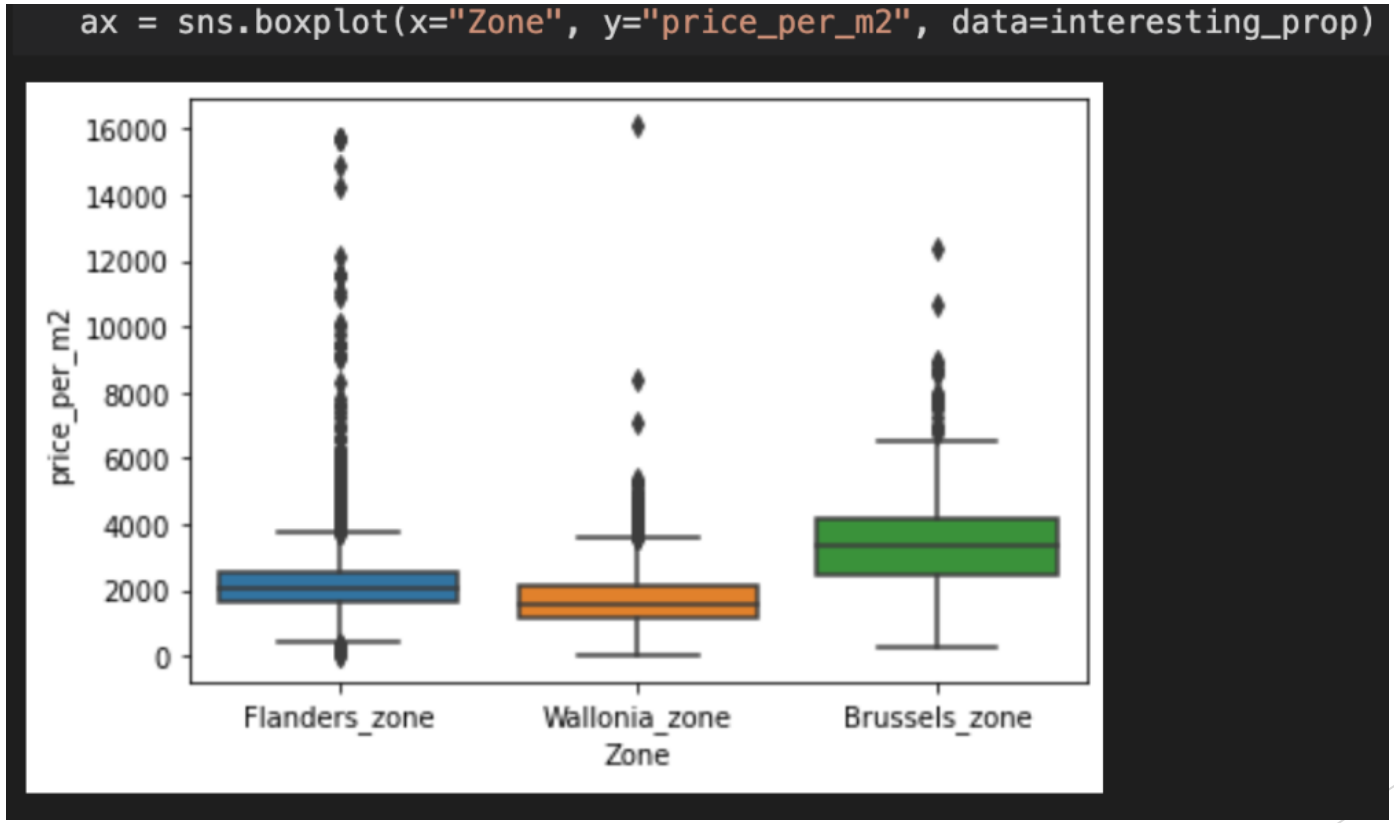
- WHY???

→ Because these variables were the most correlated with the price our main variable

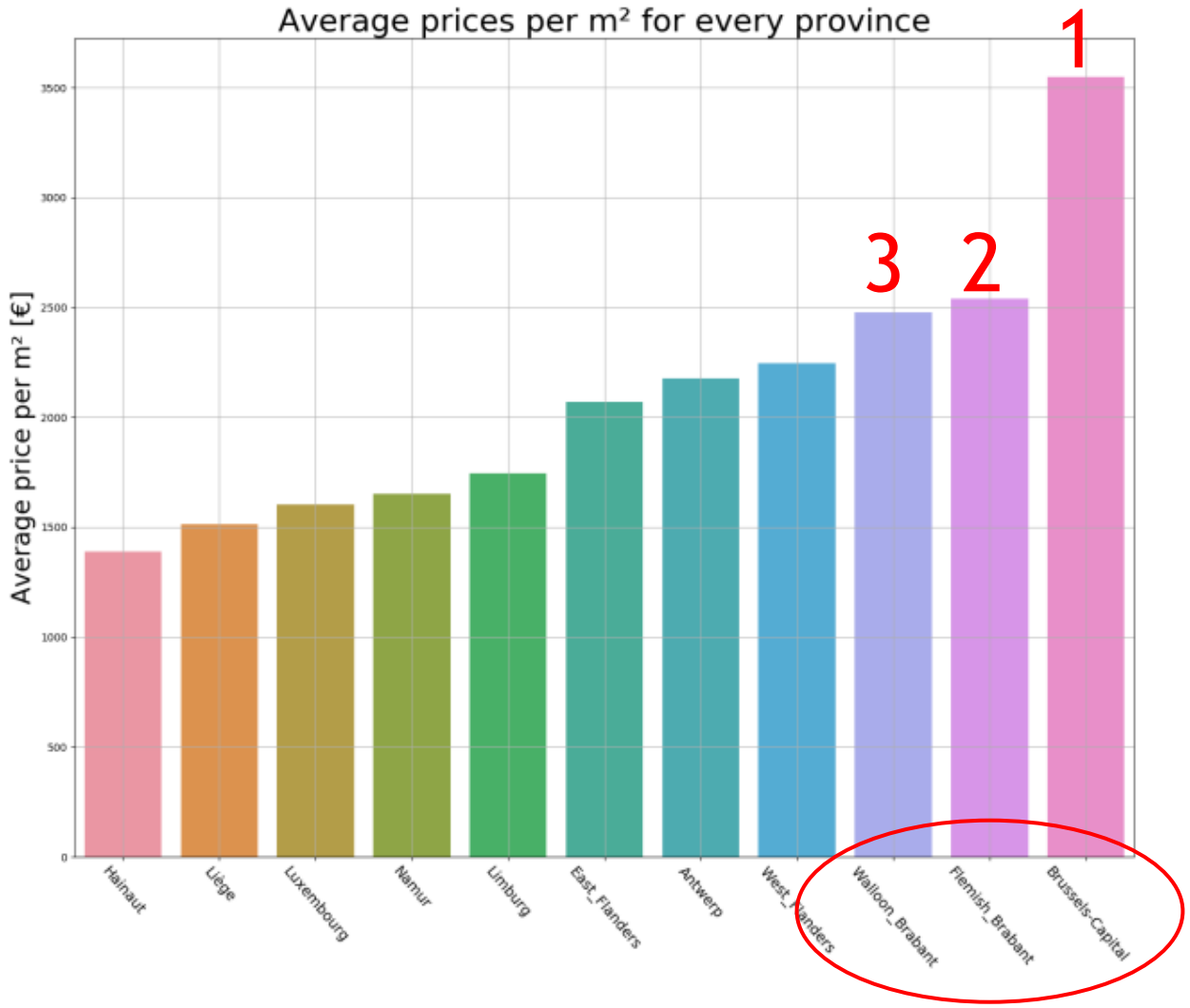
→ They explain the best why a price is higher or not

What are the most expensive regions in Belgium?

As you can see the prices are the highest in first: Brussels, Flanders and Wallonia

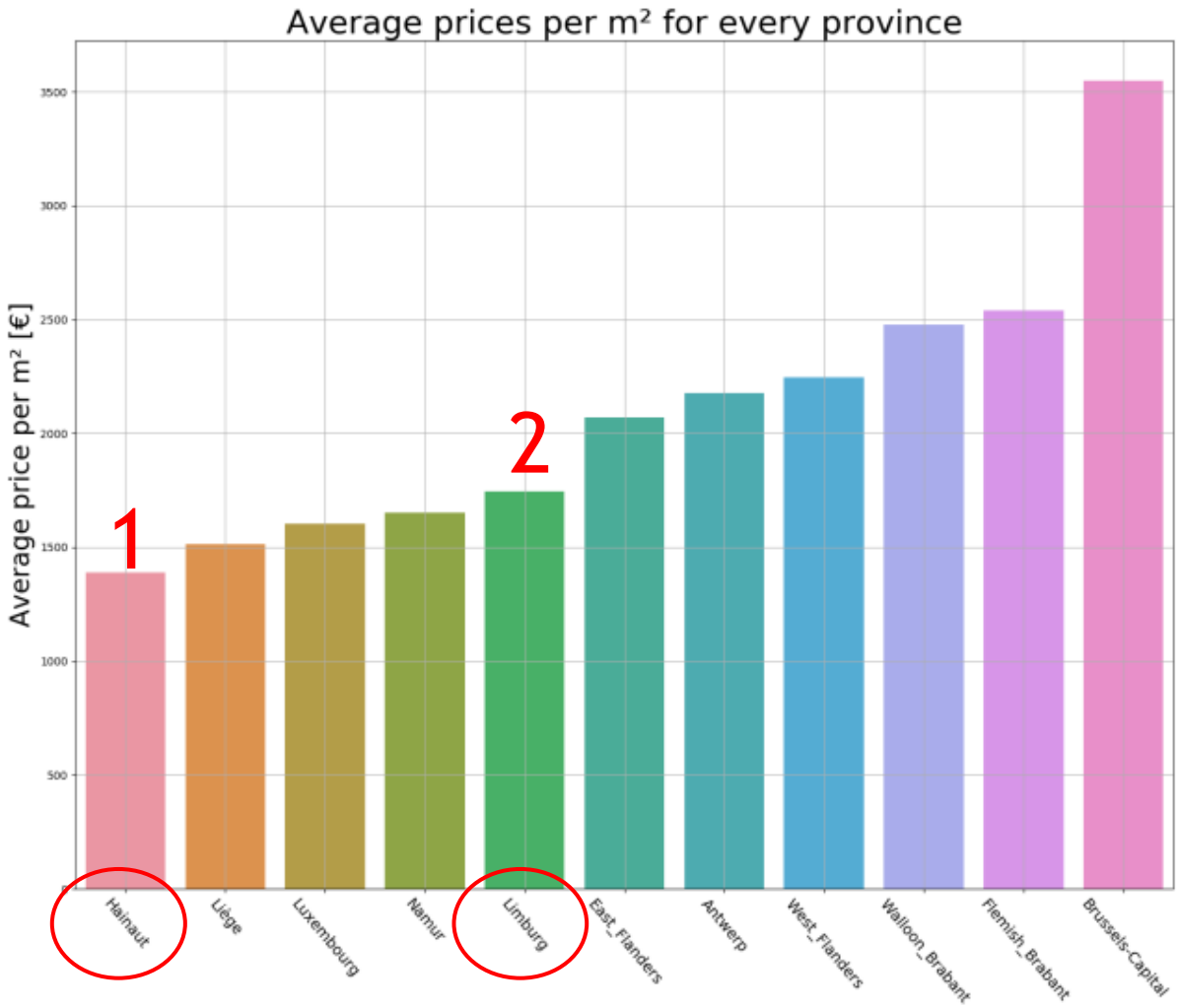


What are the most expensive provinces in Belgium, Flanders and Wallonia?



- The most expensive province in Belgium is Brussels
- The most expensive province in Flanders in is Flemish Brabant
- The most expensive province in Wallonia is Walloon Brabant

Less expensive provinces in Belgium, Wallonia and Brussels



- The less expensive province in Flanders in is Limburg
- The less expensive province in Wallonia is Hainaut

How we found our median price for the histogram

zone10	min1	max1	mean1	median1
Antwerp	201.005025	11078.651685	2177.070391	2122.641509
Brussels-Capital	238.853503	12395.833333	3548.071672	3333.333333
East_Flanders	190.000000	6250.000000	2071.180879	2037.037037
Flemish_Brabant	17.605634	7000.000000	2538.752947	2360.000000
Hainaut	145.228216	16086.956522	1388.349423	1262.863593
Limburg	287.081340	3453.947368	1746.861609	1720.779221
Liège	101.851852	3297.520661	1512.440645	1500.000000
Luxembourg	41.322314	3500.000000	1602.022434	1545.454545
Namur	137.362637	3593.750000	1650.457608	1625.467547
Walloon_Brabant	169.491525	8372.093023	2478.798258	2380.952381
West_Flanders	58.478947	15701.570681	2246.562176	1970.103093

V. CONCLUSIONS

- After analysis of our data, we can argue that the most expensives properties are located in the following regions: Brussels, Flemish Brabant and Walloon Brabant
- The cheapest provinces are located in : Hainaut and Namur

- It is difficult to advice ImmoEliza because we should need other data to determine where it is the best to invest

Thanks for your attention