

BY QBICLE.COM



**IMMO ELIZA**

**Project : BREPP\_AI**

Belgium Real Estate Price Predictor - Powered by AI

**SPRINT 2 Report: Data Visualisation**

**16/02/2024**

# AGENDA

- Context
- The dataset and its first output
- The approach
- Features/Price charting
- Global correlation matrix
- Principal component analysis
- Conclusion & Next Steps

# CONTEXT

OBJECTIVE : Design a price predictor for the Belgian real estate market based on a machine learning model

4 SPRINTS :

- DATA COLLECTION AND PREPROCESSING
- DATA VISUALIZATION
- DEVELOP THE MACHINE LEARNING MODEL
- DEPLOYMENT THROUGH API

# A dataset reworked to be quantitative

- 25K unique properties
  - 44% houses
  - 56% apartments
- 28 actionable parameters
  - of which 24 are now quantitative (cluster to boolean)
- Rather high NaN values (25% average)

# 6 categories for 28 parameters

# Quantitative

# Qualitative

The diagram illustrates the hierarchical structure of real estate features, categorized into six main groups:

- Geoloc**: Price, locality, province, region
- General**: type, subtype, year, new, condition, frontage, tenement, flood safety
- Energy**: energy class, heating, double glazing
- Access**: elevator, disabled
- Interior**: living surface, furnished, bedrooms, bathrooms, kitchen, open fire
- Exterior**: plot surface, garden, terrace, parking, pool

# First outputs : real estate market facts

Property  
Features →

- Price Mean/Median: 475/340 KEUR
- Age Mean/Median : 1976 /1978
- 19% considered as new (2021+)
- 23% of the goods to be renovated
- 33% in good energy class (B or better)
  - 95% have double glazing
- 5% in a flooding zone
- 28% accessible (disabled people)
- 6% tenement property

# First outputs : real estate market facts

- 50% parking
- 40% terrace
- 20% garden
- 2.5% swimming pool

← Exteriors  
Features

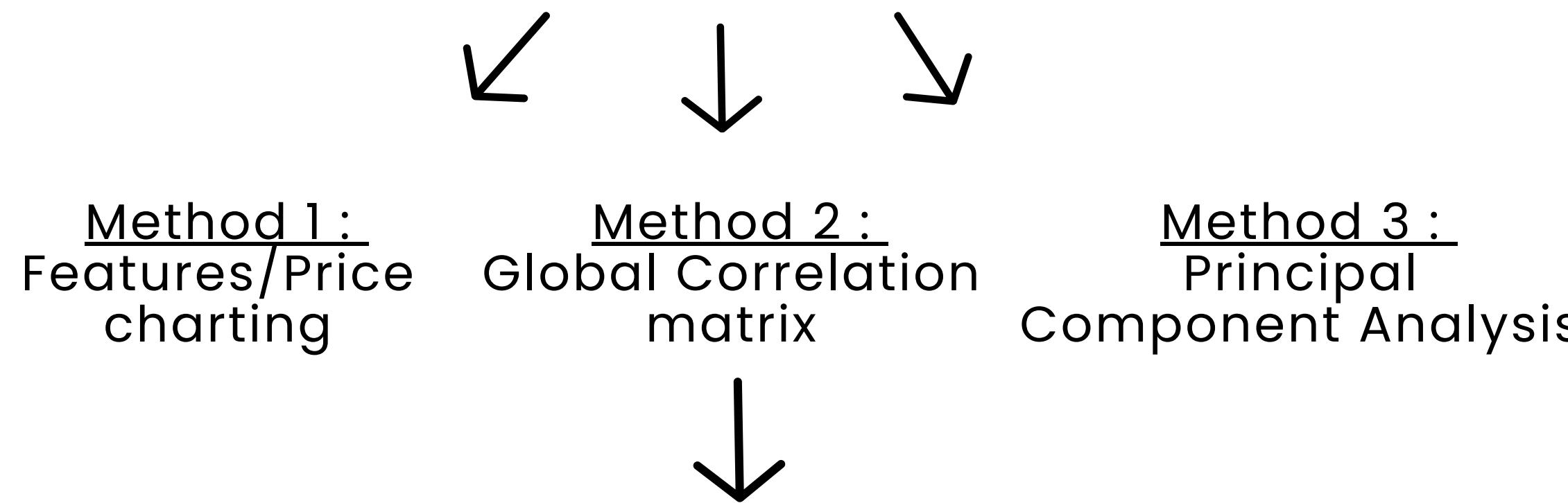
Interiors  
Features →

- Average living surface 170 m<sup>2</sup>
- Avg bedroom/bathroom: 3.0/1.4
- 95% of equipped kitchen
- 3% of open fire
- 2.5% furnished

# The triple down approach

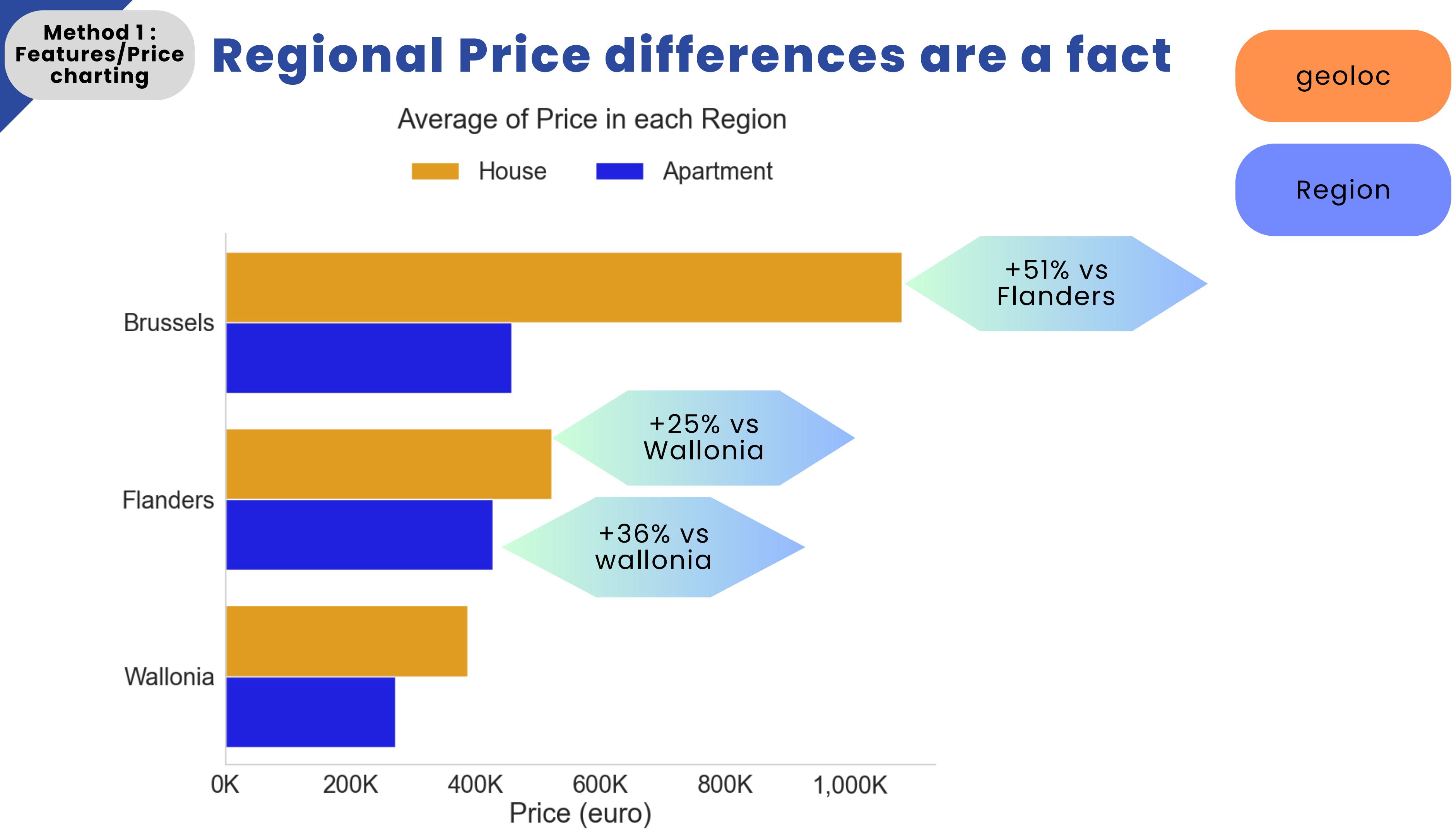
**28 Features**

**3 correlation methods used in synergy**



**1 main output : The 5(+4) max impact features on price**

WHAT YOUR BUYERS WANT  
WHAT WILL MAKE YOUR PRICING RELIABLE



Method 1:  
Features/Price  
charting

# Driven by provincial disparities...

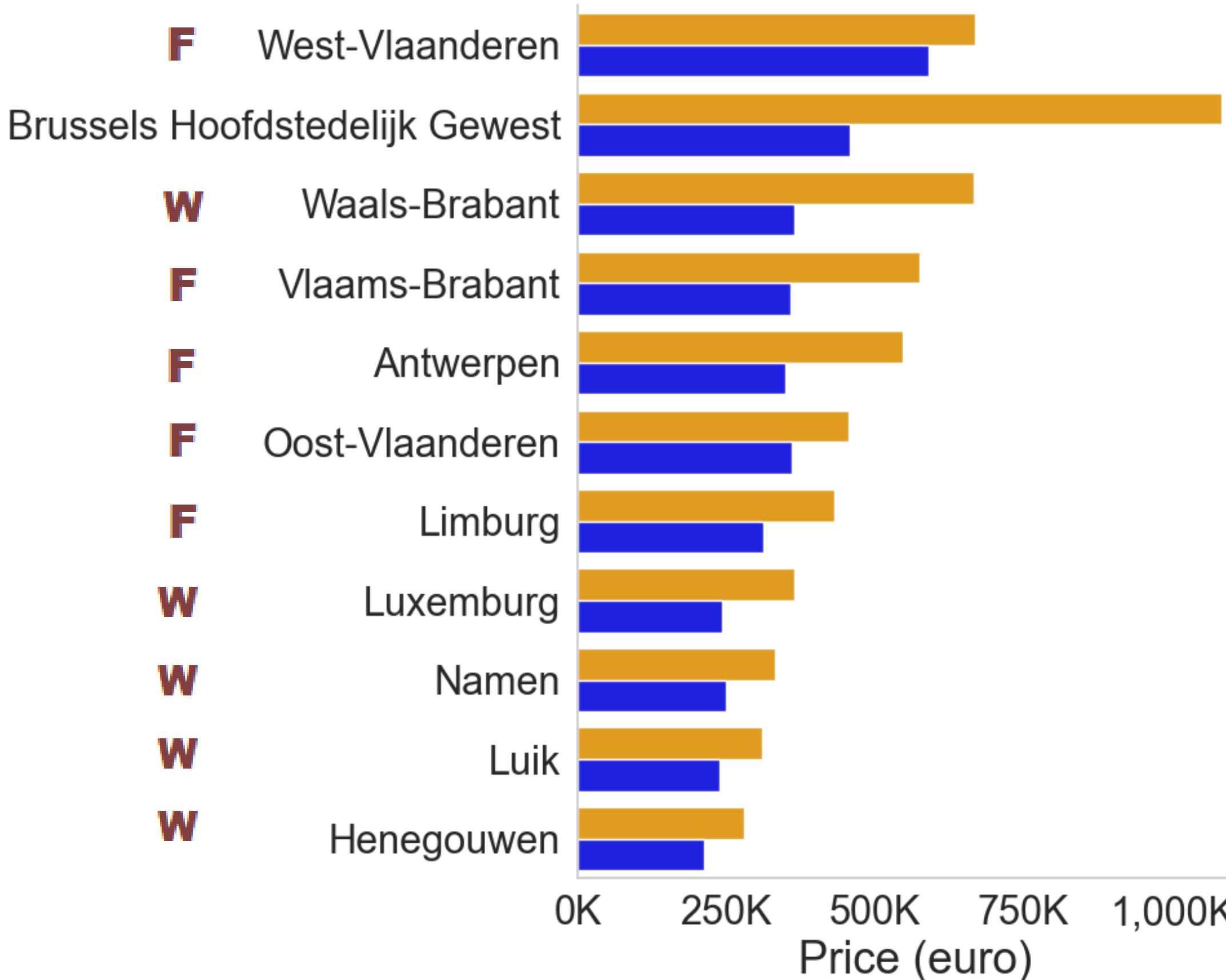
Average of Price in each Provinces

House Apartment

geoloc

Province

**W** for Wallonia  
**F** for Flanders



Method 1:  
Features/Price  
charting

# ... and confirmed by local disparities as well : the big 10X stretch

geoloc

Locality

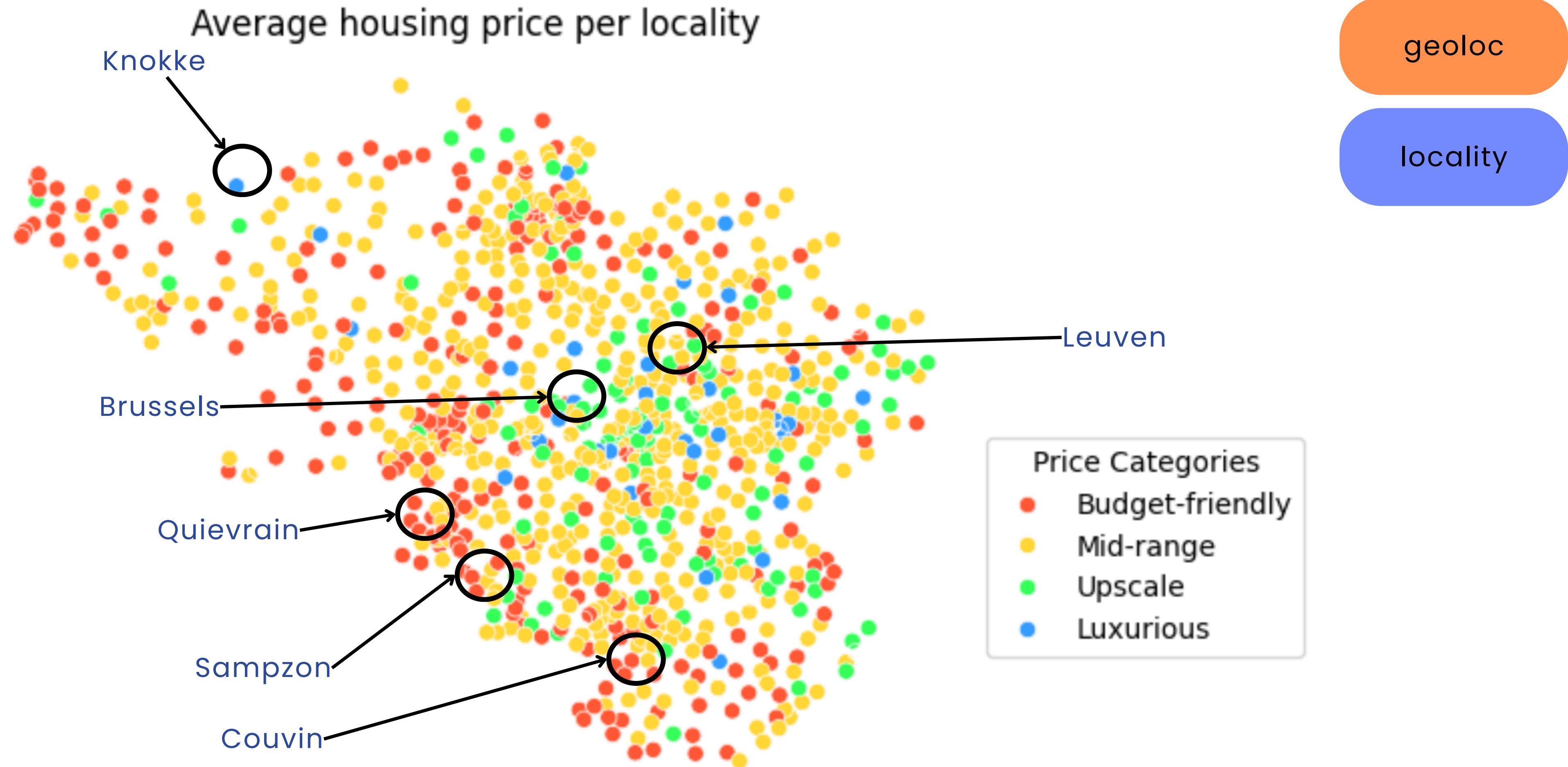
Top 3 Cheapest localities			
Postal code	Locality	province	Price (per m2)
5660	Couvin	Namen	1,008
7120	Sampzon	Henegouwen	1,018
7380	Quievrain	Henegouwen	1,046

Top 3 Most Expensive localities			
Postal code	Locality	province	Price (per m2)
8300	Knocke-Heist	West-Vlaanderen	11,393
8301	Knocke-Heyst	West-Vlaanderen	7,713
3000	Leuven	Vlaams-Brabant	6,718

\* Only Localities with 10 or more listings were selected for this analysis

Method 1:  
Features/Price  
charting

# All this reflects in locality heatmap BUT...



Method 1:  
Features/Price  
charting

# ... but localities impact price even on narrower geography as well

geoloc

Locality

Top 3 Most Cheapest localities in Brussels		
Postal code	Locality	Price (per m2)
1081	Koekelberg	2,628
1080	Molenbeek-Saint-Jean	2,618
1130	Haren	2,520

Top 3 Most Expensive localities in Brussels		
Postal code	Locality	Price (per m2)
1050	Ixelles	4,748
1180	Uccle	4,531
1000	Bruxelles	4,406

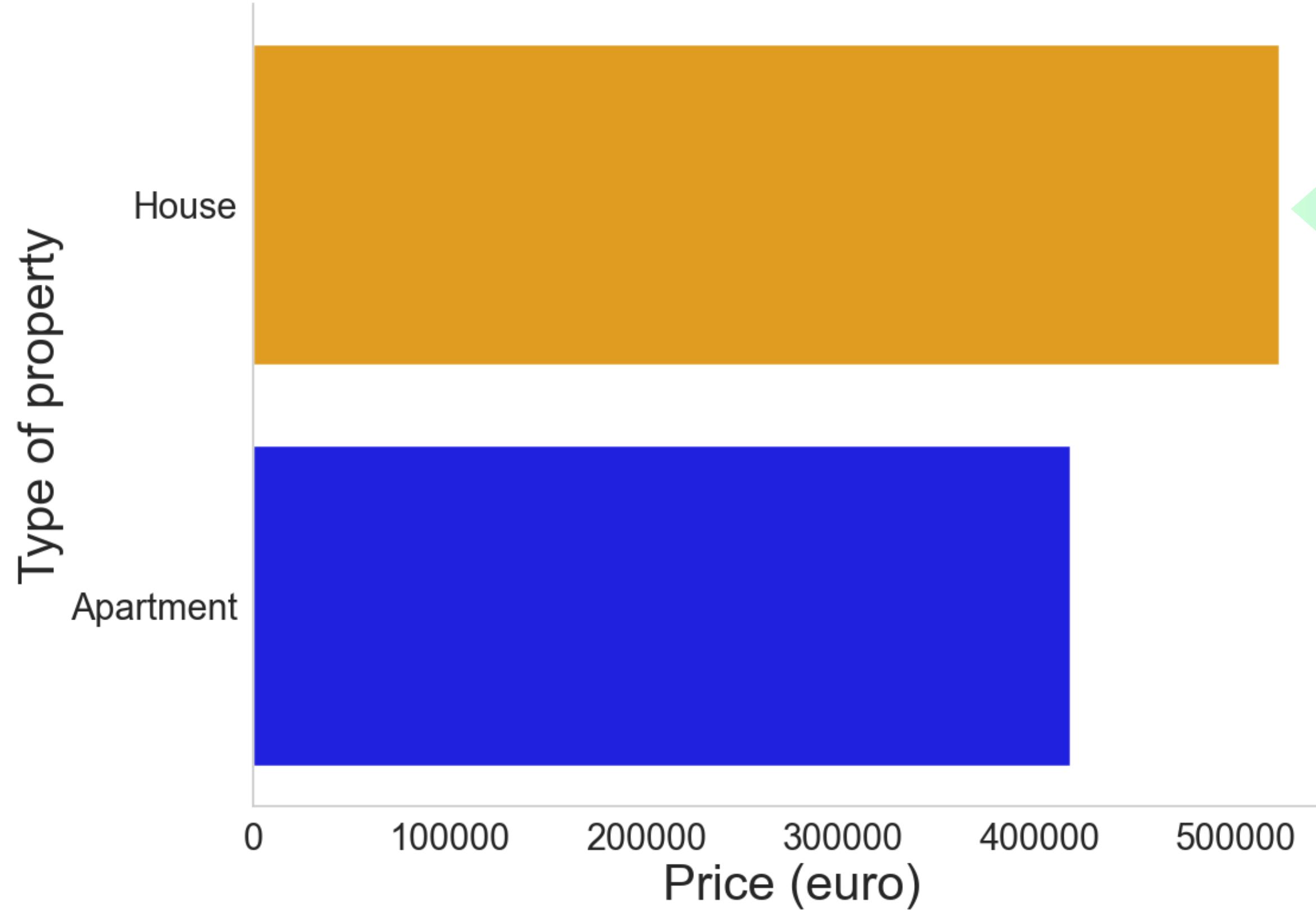
Method 1:  
Features/Price  
charting

# Property Type: A serious impact

general

Type

Property Type Vs Price



Method 1:  
Features/Price  
charting

# Features : The 9 silver bullets ?

general

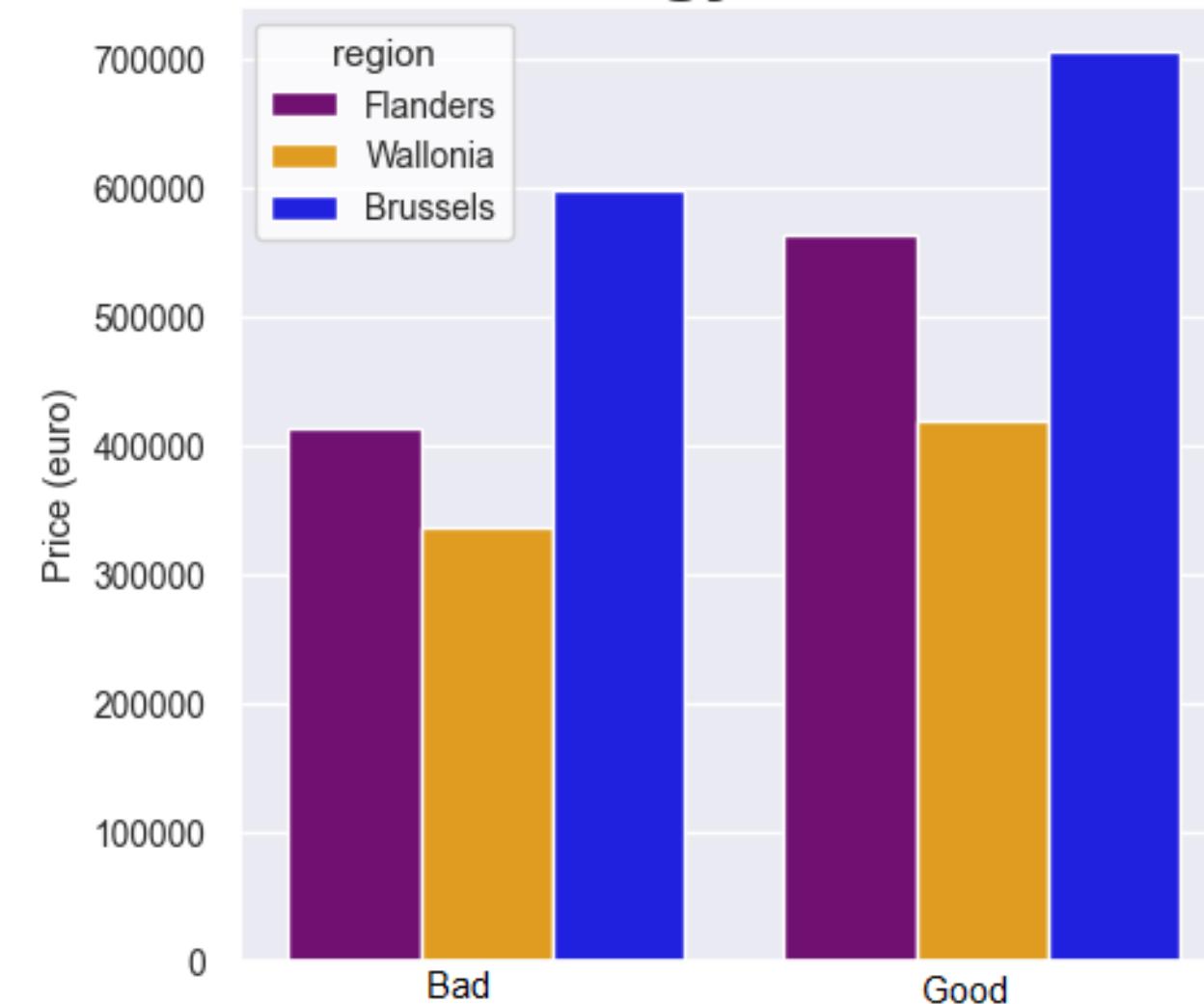
energy class

Condition

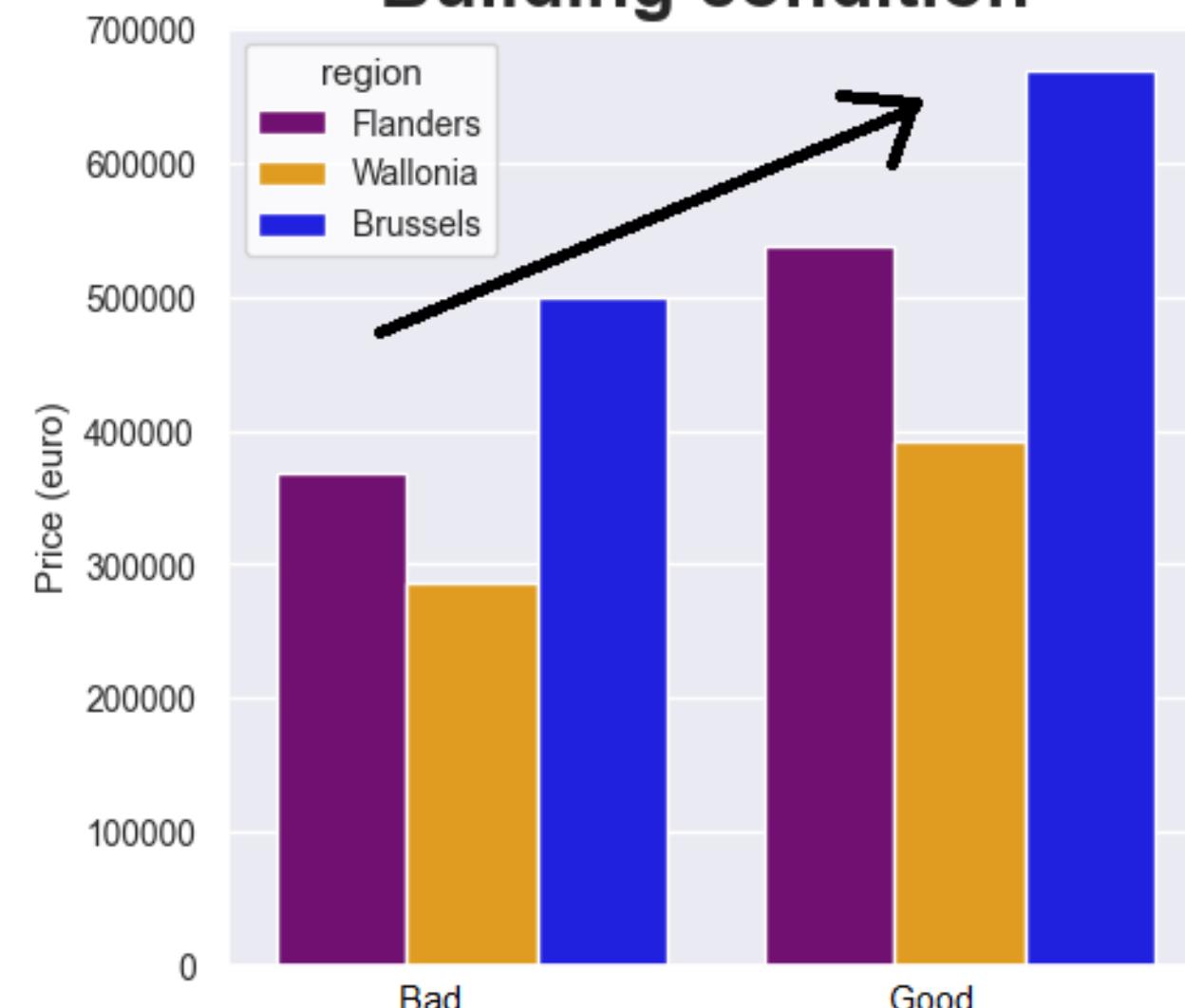
Tenement

## Features impact on Price

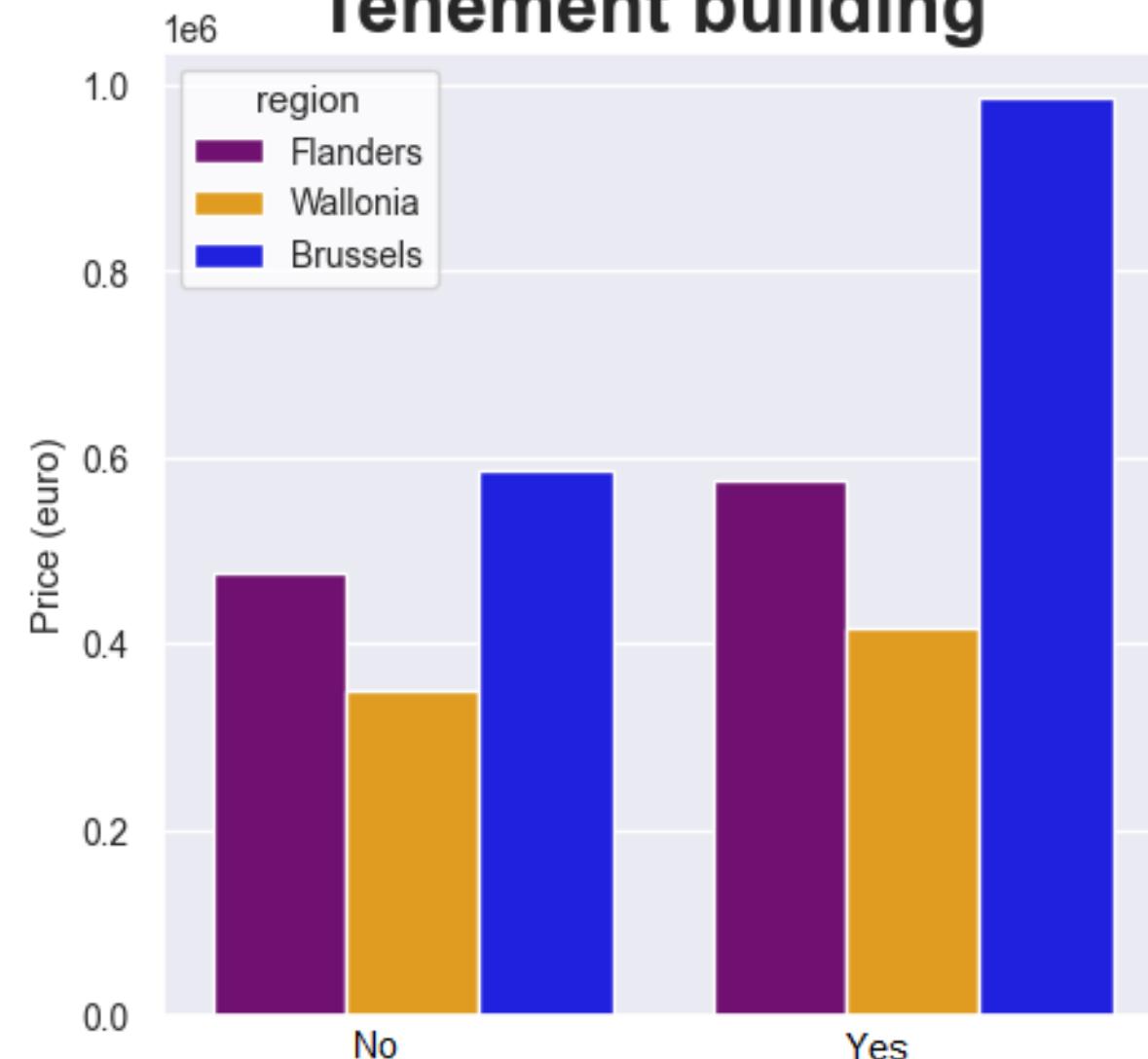
Energy class



Building condition



Tenement building



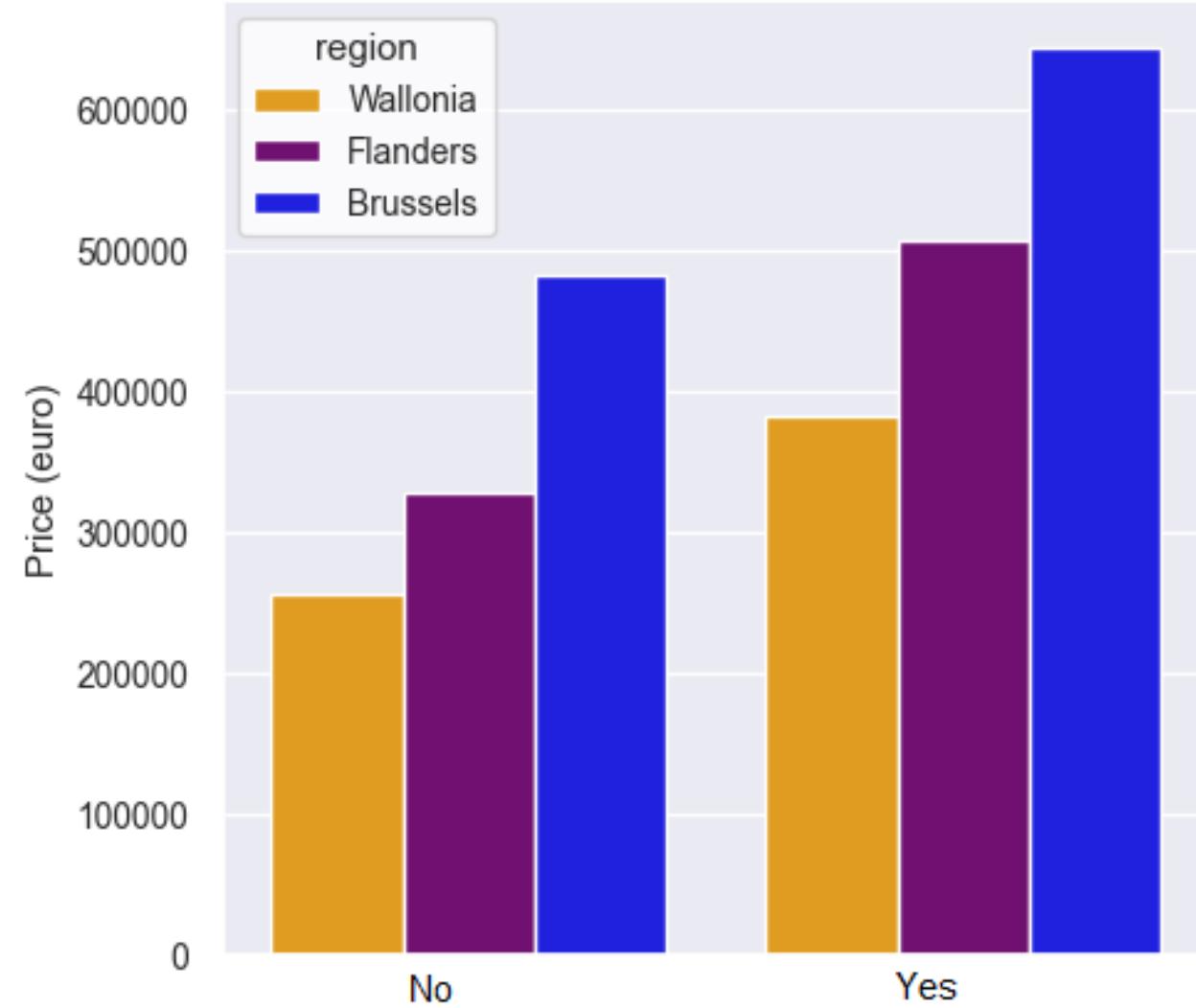
**Method 1:  
Features/Price  
charting**

# Features : The 9 silver bullets ?

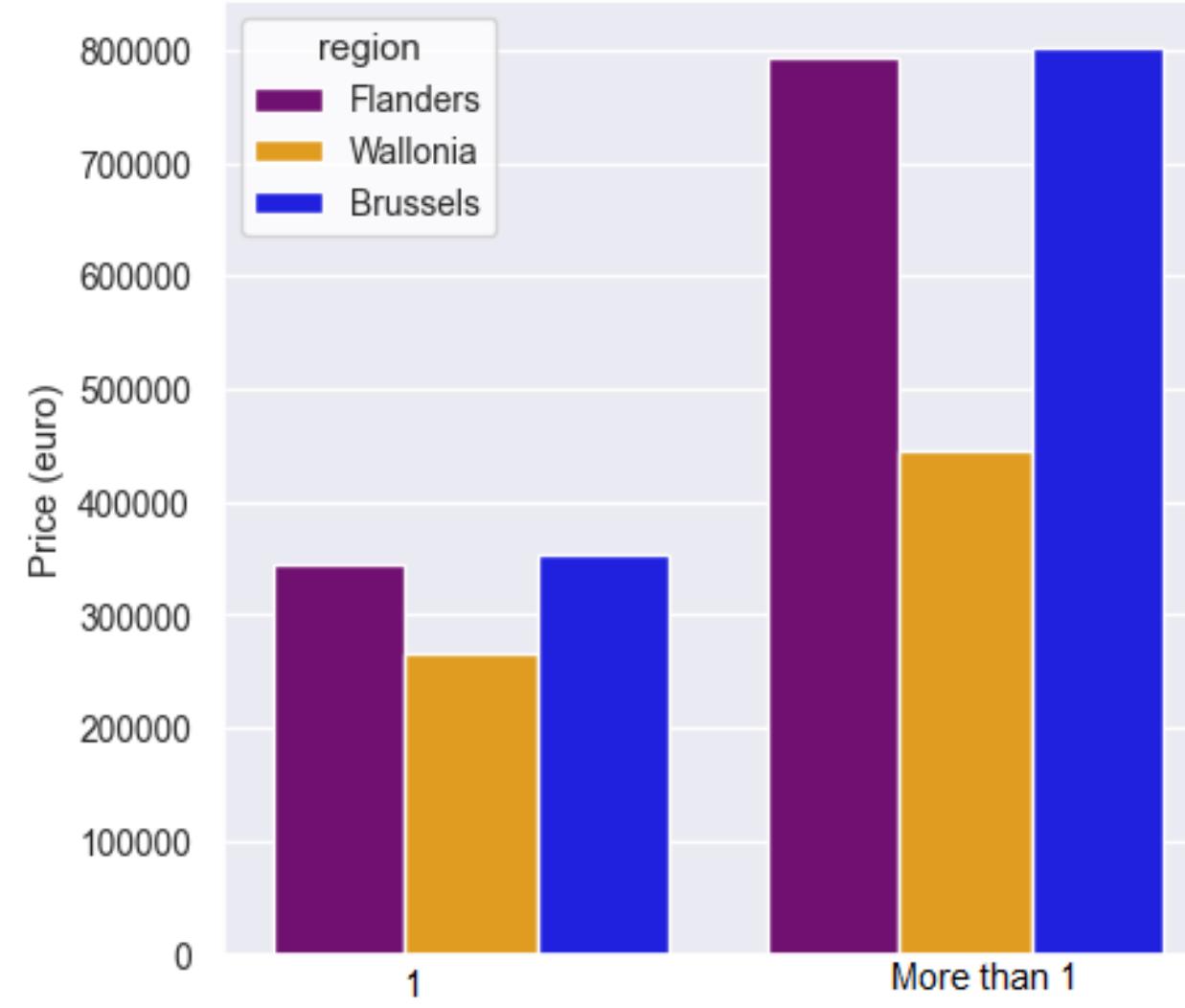
- furnished      Bedrooms
- Interior      Bathrooms
- Surface
- Open fire
- kitchen

## Features impact on Price

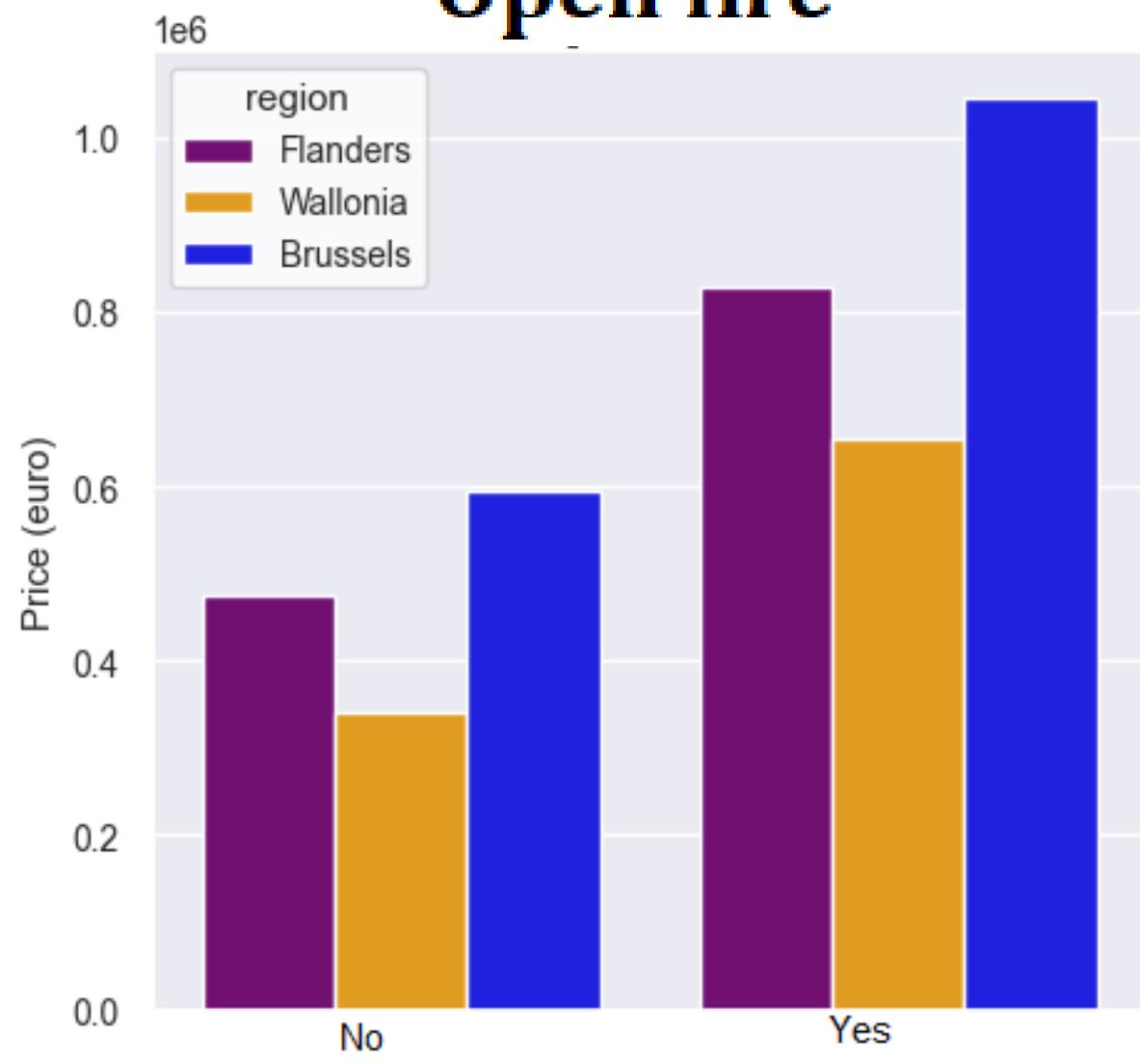
**Kitchen equipped**



**Bathrooms total nb**



**Open fire**

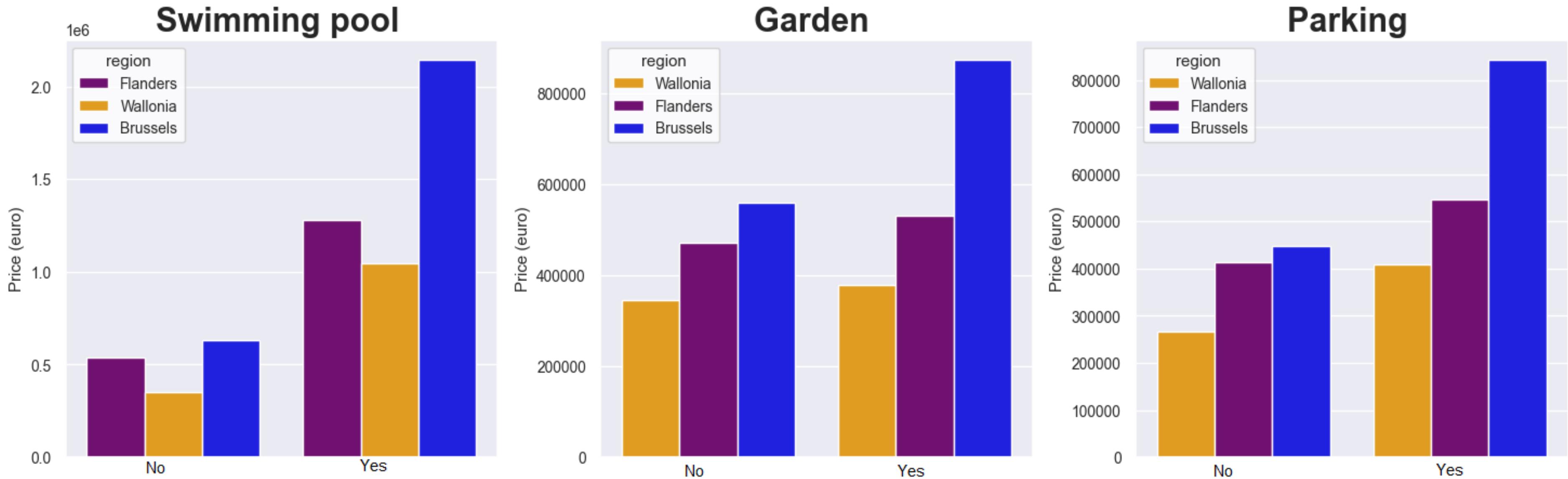


Method 1:  
Features/Price  
charting

# Features : The 9 silver bullets ?

- Exterior
- Plot surf.
- parking
- Surf. tot
- garden
- pool

## e Features impact on Price



## **Method 2 : Features correlation matrix**

# 8 bullets showing up in the correlation matrix

- Living surface, bedrooms & bathroom nb (0.4)
  - Swimming pool (0.32)
  - Terrace and garden surface (0.25)
  - Num of frontages > house vs appment (0.17)
  - Parking (0.14)
  - Energy class (0.11)
  - Building conditions (0.12)
  - Open fire (0.12)

## Missing:

- Kitchen equipped,
  - Tenement building

# Correlating variables

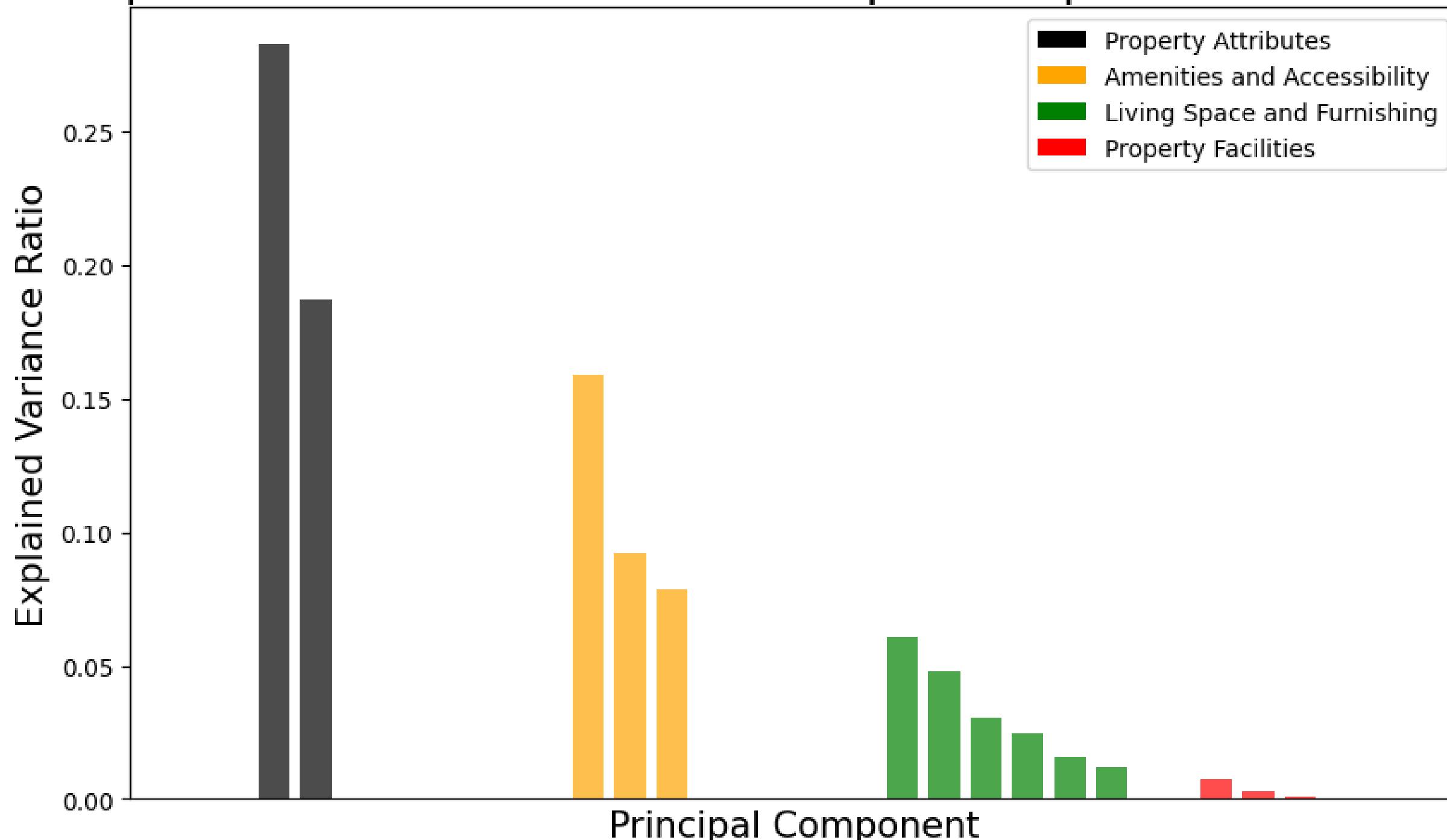
A thick black arrow pointing to the right, indicating the direction of the next step.

- Construction year /Building condition / Access / Elevator / Terrace / Energy Class / Living surface/ Number of Bedroom(<0),
  - Number of Bathroom / Number of Bedroom / Living surface
  - Terrace surface / Living surface, Swimming pool, Number of bedroom, Garden
  - Parking to number of frontage

# Highest correlation to Price

# Quantifying attributes' impact on pricing for houses

Explained Variance Ratio of Principal Components for Houses



**Construction year:** 22.78%  
**Building condition:** 11.84%

**Energy class :** 9.11%  
**Double glazing:** 7.08%  
**Accessible if disabled:** 6.42%

**Living surface (sqm):** 6.04%  
**Furnished:** 5.30%  
**Nb of Bedrooms:** 4.96%  
**Nb of Bathrooms:** 4.31%  
**Kitchen equipped:** 4.00%  
**Open fire:** 3.52%



# Conclusions : 5 crucial parameters

- **Geographical location** (locality/province/region):
  - Cultural, economic factors and density are key influencers.
- **Type of property:** Houses/Apartments
  - Necessity for separate modeling due to disparities?
- **Energy efficiency rating:**
  - Increasing global and political significance.
- **Year of construction/building condition:**
  - Influence on housing quality and renovation costs.
- **Living area/number of bedrooms/number of bathrooms:**
  - Property dimensions and layout crucial for market value.

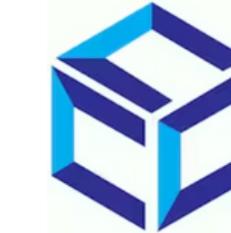
# Conclusions : 4 desirable parameters



- **Parking space:**
  - Still a practical necessity for buyers
- **Swimming pool:**
  - Swimming pool indicate a premium for luxury amenities
- **Equipped kitchen:**
  - Equipped kitchens are the standard
- **Open fire:**
  - Synonym of comfort and ambiance

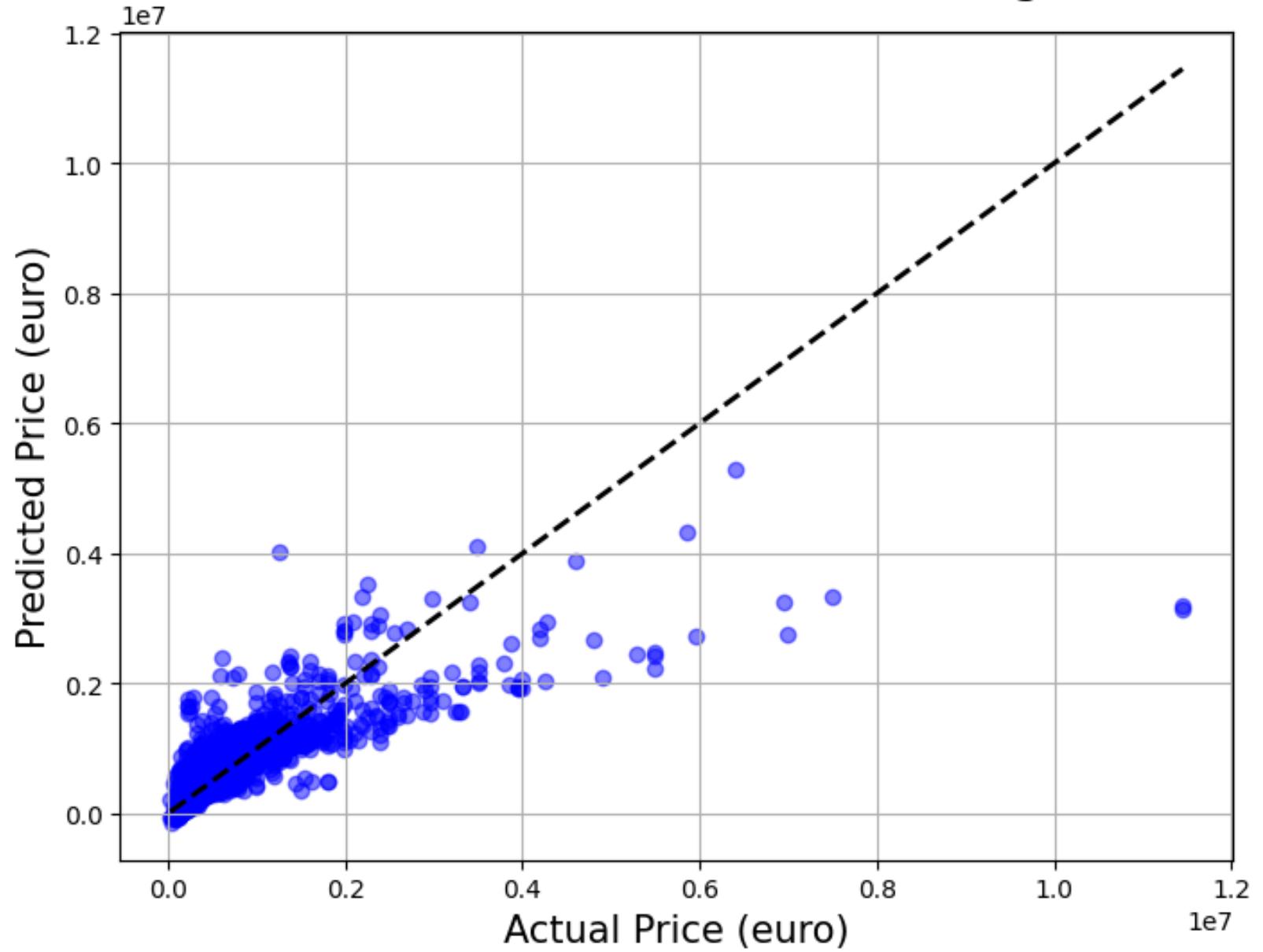
# Next Step : Price Prediction

BY QBICLE.COM

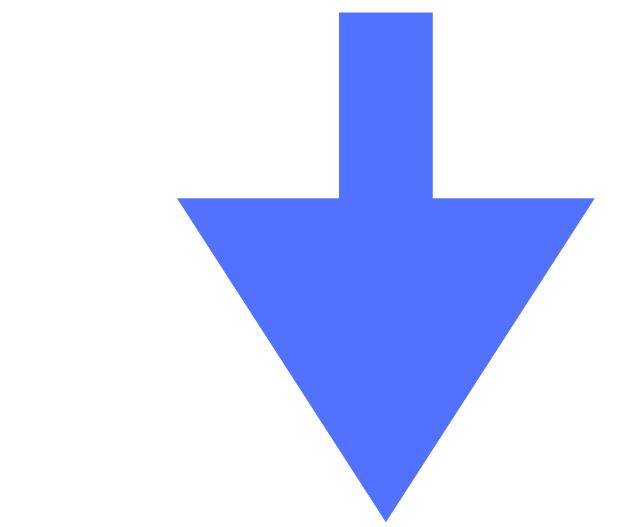


Ready for sprint 3: Create the machine learning model

Actual vs Predicted Prices (Linear regression)



Regression Model Performance:  
R-squared: 0.664  
Mean Absolute Error: 133,798.43



Room for improvement

**IMMO ELIZA  
PROJECT: BREPP\_AI**

BY QBICLE.COM



THANK YOU !

&

SEE YOU NEXT WEEK !

Mahak, Maarten, Sylvain

**IMMO ELIZA  
PROJECT: BREPP\_AI**

BY QBICLE.COM

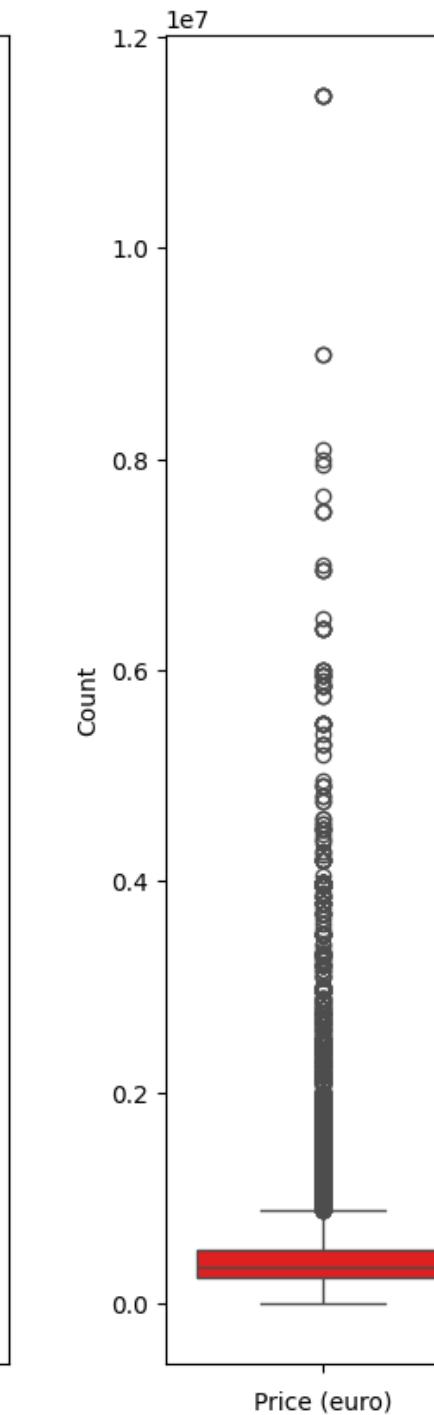
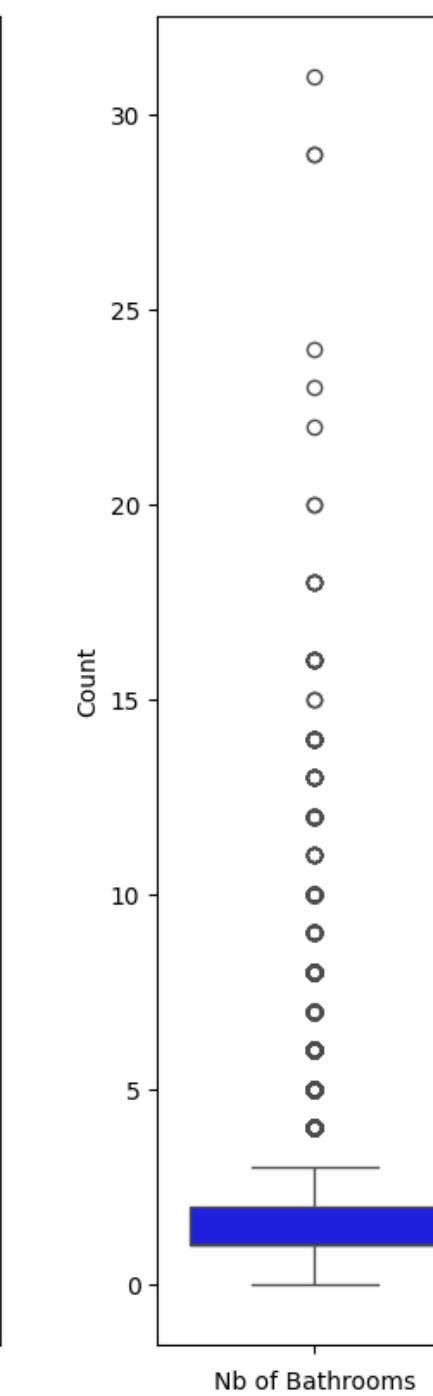
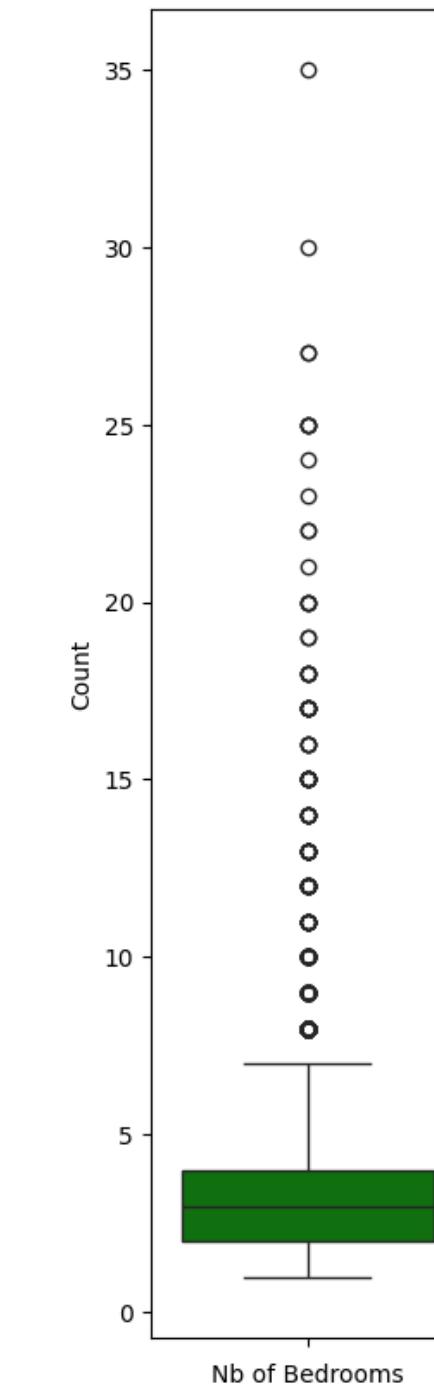
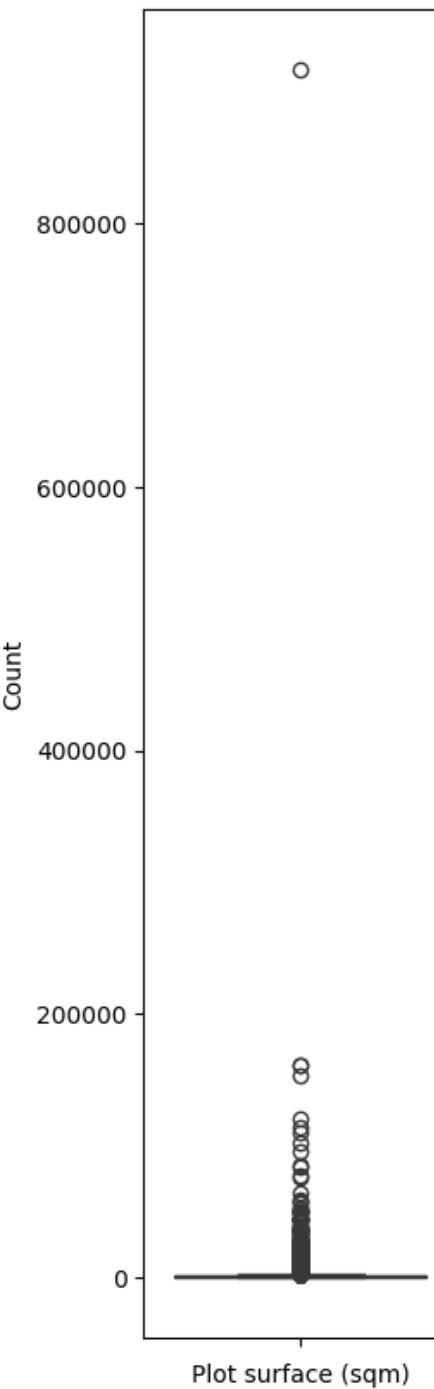
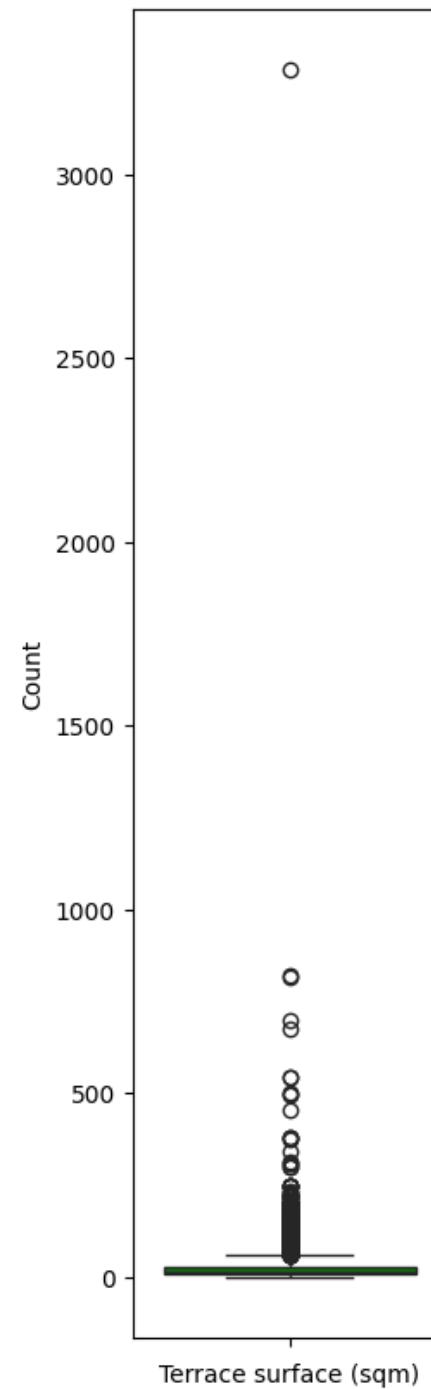
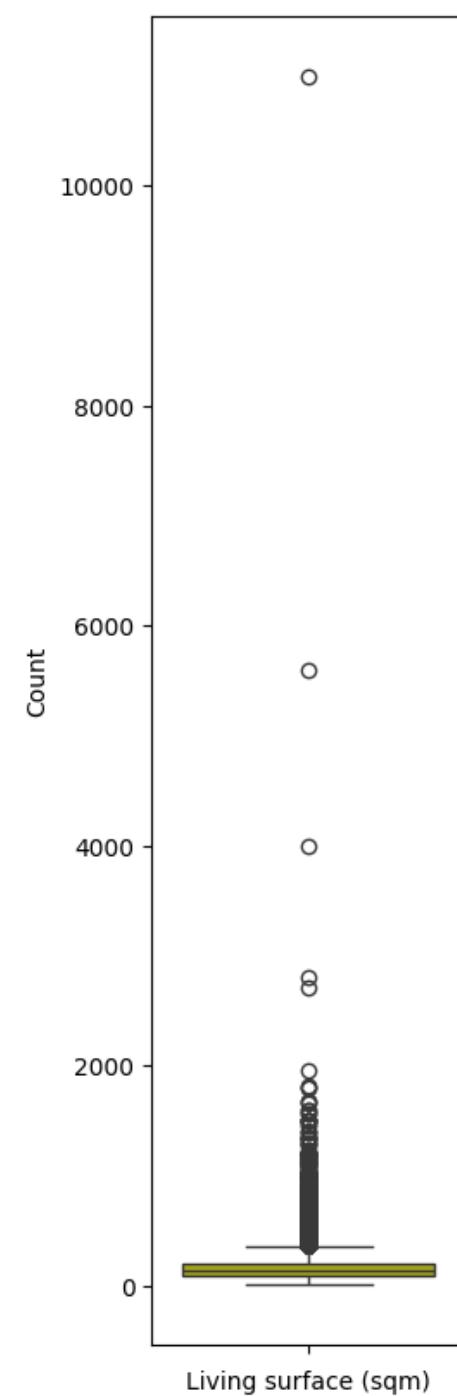


back up

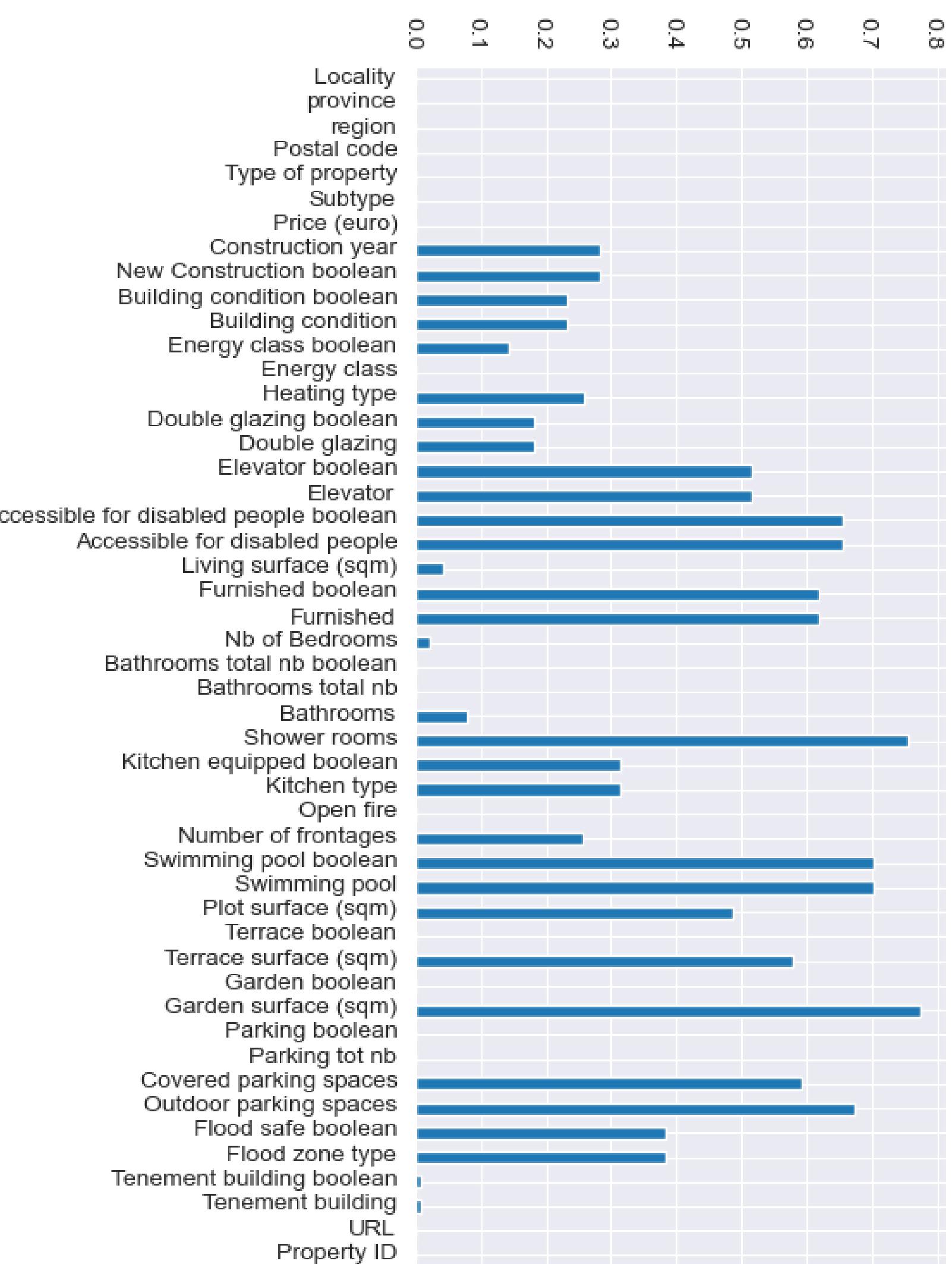
Mahak, Maarten, Sylvain

**IMMO ELIZA  
PROJECT: BREFP\_AI**

# Parameters with Outliers



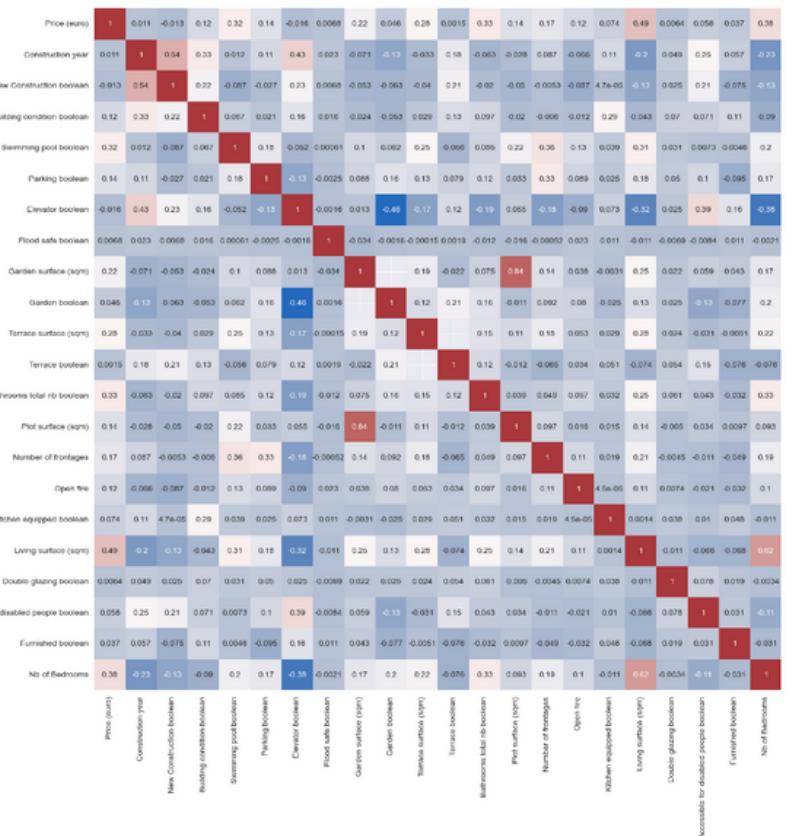
# Nan



# Intervariable correlation

## Examples of intervariable correlation

- construction year /building condition / access / elevator / terrace / nrj / living surface/ bedroom nb(<0),
- bathroom nb / bedroom nb / living surface
- terrace surface / living surface, swimming pool, nb of bedroom, garden
- parking to nb of frontage



# Conclusion: Impactful Parameters

BY QBICLE.COM



## MAIN PARAMETER

- Type of Building : Houses or Appartement >> 2 arms
- Geoloc (locality/province/region)
- Energy class
- Construction year /building condition
- Living surface /nb of bedrooms / nb bathrooms

## SIMPLE PRICE CORREL

- Price Analysis :
- Energy Class
- Good Building condition
- Tenement Buildings
- Equipped kitchen
- No. of Bathroom /No. of bedrooms/Living Surface
- Open Fire
- Swimming Pool
- Garden
- Parking

## PRINCIPAL COMPONENT ANALYSIS

- Construction year: 22.78%
- Good Building condition : 11.84%
- Energy class (A+,A & B) : 9.11%
- double glazing and
- Accessible for disabled people: 6.42

## More crispy recommendation

- cocktail in the pool ?
- kitchen equiped chicken happy !
- Belgium Hygge (said "hooga" is of Danish) > light a fire
- Belgium car lover > parking

## • correlation MAP

### 0.10 - 0.15

- Good building condition
- Parking present
- Plot surface
- Open fire

### 0.16 - 0.30

- Garden surface
- Number of frontage
- Terrace surface

### Above 0.30

- Number of bathroom
- Number of bedroom
- Living surface
- Swimming pool

**IMMO ELIZA  
PROJECT: BREFP\_AI**