



# ImmoAnalysis

Patrick Brunswyck  
Jose Roldan  
Amaury van Kesteren



# The mission

## Data Operations

- Data acquisition
- Data cleaning

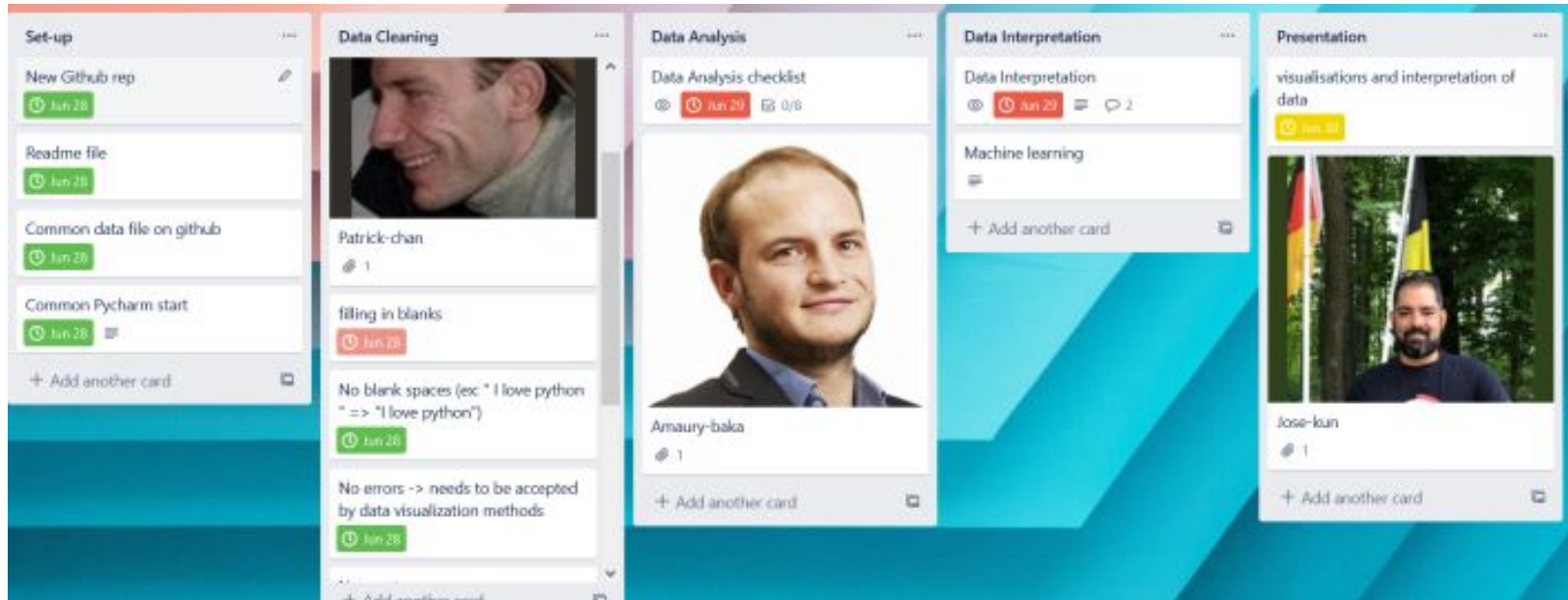
## Data Analysis

- Target Variable
- Correlations
- Influences
- Quality vs Quantity
- Percentages

## Data Interpretation

- Outliers
- Variables importance
- Most Expensive municipalities
- Less Expensive municipalities

# The project management



# Data Operations





# Data Operations

## Data Acquisition

house\_list\_houses\_dataset.csv

## Data Cleaning

1. **NaN values:**
  - a. `df[column].mean()`
2. **Drop empty [*“price”*] rows**
  - a. Target variable = Price
  - b. Accurate values only
  - c. 35 out of 10000
3. **See fig.1**

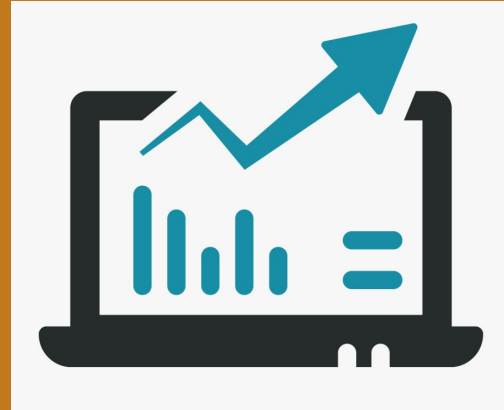
	area	price	building_state	facades	bedrooms	kitchen_equipped	furnished	open_fire	locality
0	123.000000	200000.0	good	3.000000	2.0	installed	False	1.144966	3294
1	255.000000	698000.0	just renovated	4.000000	4.0	installed	False	1.144966	2070
2	234.000000	499000.0	as new	2.970242	5.0	undefined	False	1.144966	8800
3	219.000000	717000.0	as new	2.970242	4.0	undefined	False	1.144966	8000
4	156.000000	429000.0	good	4.000000	4.0	installed	False	1.144966	8800
5	140.000000	279000.0	good	2.000000	5.0	installed	False	1.144966	9000
6	256.000000	210000.0	good	3.000000	4.0	installed	False	1.144966	7080
7	850.000000	849000.0	good	2.000000	9.0	undefined	False	1.144966	3300
8	161.000000	295500.0	as new	2.970242	4.0	hyper equipped	False	1.144966	5590
9	61.000000	42000.0	to renovate	2.970242	1.0	undefined	False	1.144966	1460
10	160.000000	229000.0	good	2.000000	3.0	installed	False	1.144966	6230

DataSet Cleaned

land_surface	terrace	terrace_surface	swimming_pool	property_type	property_subtype	garden	garden_surface
429.0	False	0	False	house	house	True	350
1210.0	False	0	False	house	villa	False	0
744.0	False	0	False	house	villa	False	0
760.0	False	0	False	house	house	False	0
419.0	False	0	False	house	villa	True	90
85.0	True	40	False	house	house	False	0
201.0	False	0	False	house	house	False	0
615.0	True	20	False	house	mixed	False	0
533.0	False	0	False	house	house	True	631
194.0	False	0	False	house	house	False	0
480.0	True	10	False	house	house	True	400

Fig. 1

# Data Analysis



**Target  
Variable**



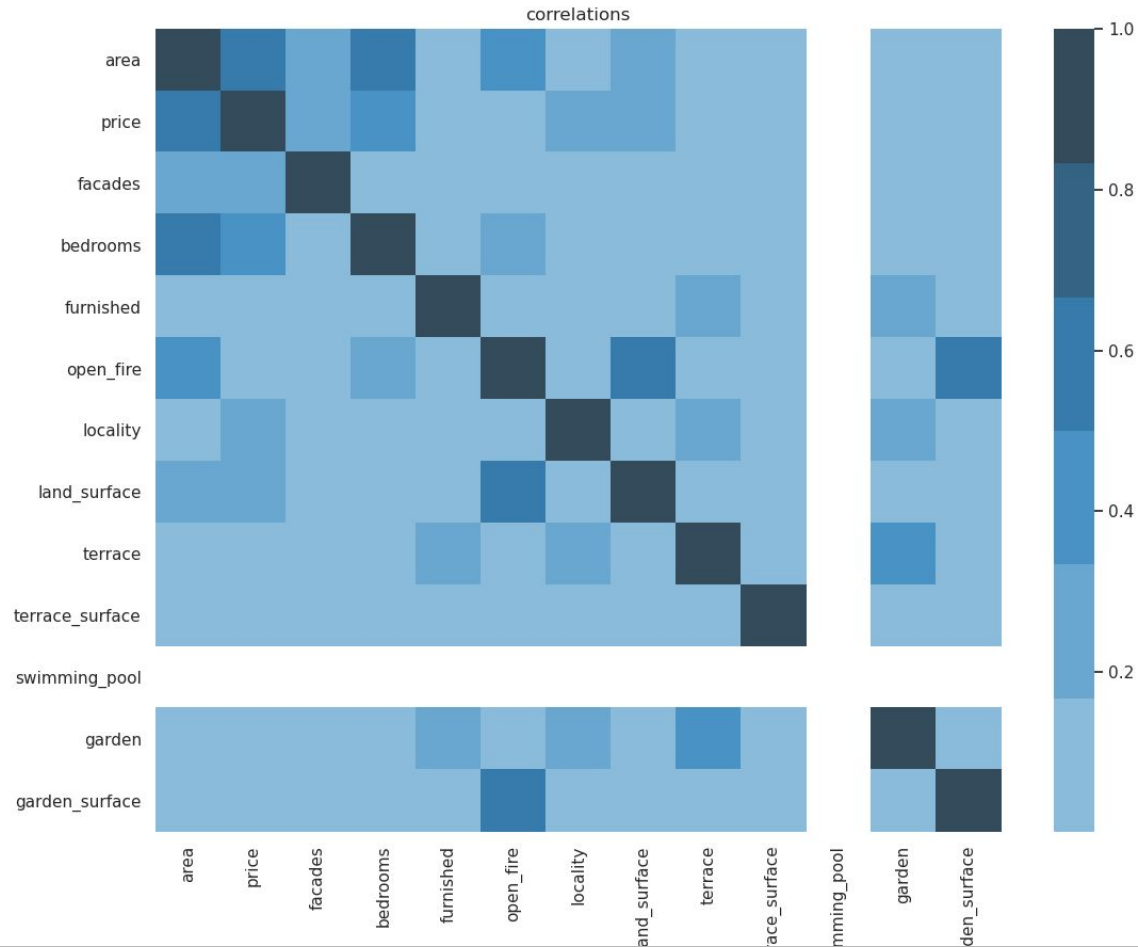
['price']

```
1 price 10058 non-null float64
```



# Correlations

correlations between price & other columns



# Influences





# Qualitative vs Quantitative

# Missing Values

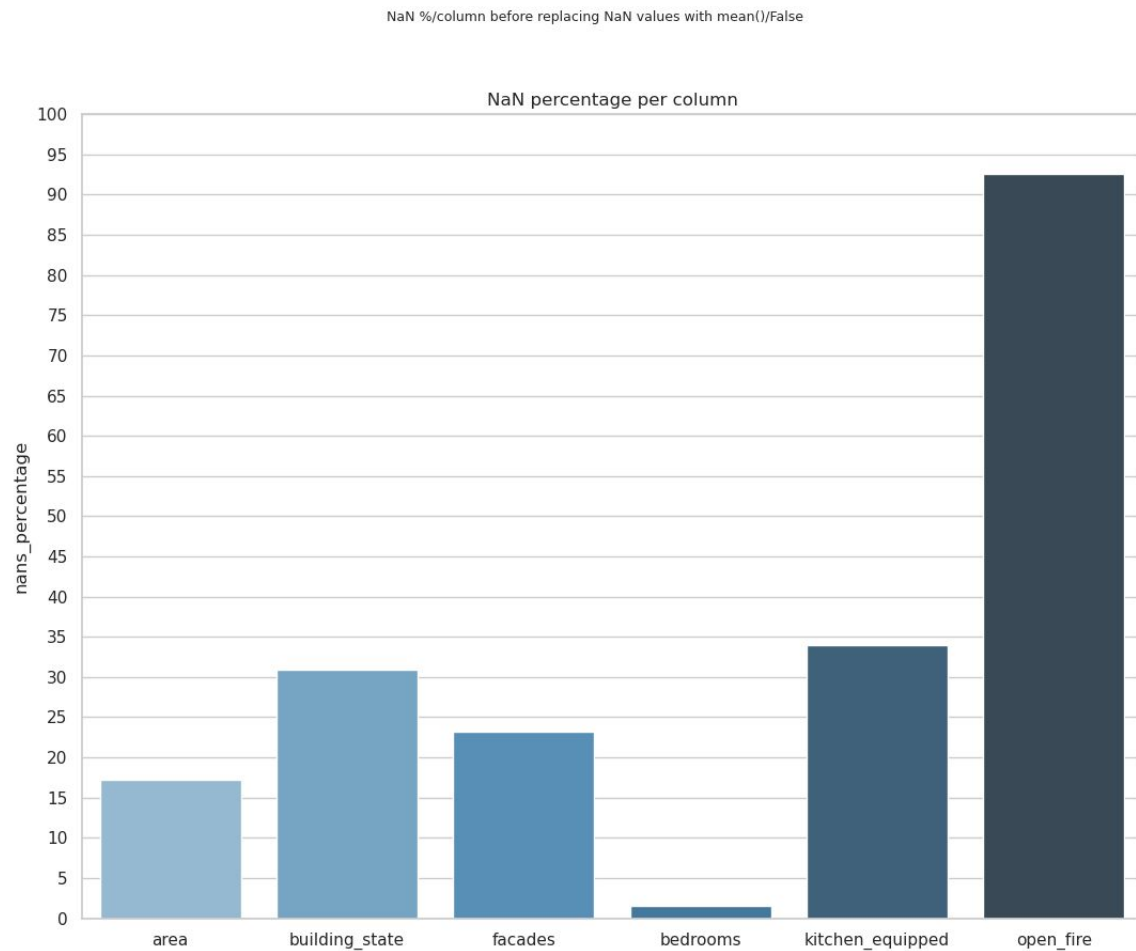


Fig. 1

# Data Interpretation



# Outliers ↑



Castle

**€9,500,000**

12 bdr. · 1100 m<sup>2</sup>  
3090 Overijse



LAP  
LATOUR & P



```
area          1100.0
price         9500000.0
building_state as new
facades       4.0
bedrooms      12.0
kitchen_equipped installed
furnished     False
open_fire     1.144966
locality      3090
land_surface  500000.0
terrace       False
terrace_surface 0
swimming_pool True
property_type house
property_subtype castle
garden        False
garden_surface 0
Name: 2090, dtype: object
```

```
area          945.0
price         9500000.0
building_state as new
facades       4.0
bedrooms      7.0
kitchen_equipped undefined
furnished     False
open_fire     1.144966
locality      8300
land_surface  2348.0
terrace       False
terrace_surface 0
swimming_pool True
property_type house
property_subtype house
garden        False
garden_surface 0
Name: 4956, dtype: object
```



NEW



Exceptional property

**€9,500,000**

7 bdr. · 945 m<sup>2</sup>  
8300 Knokke-Heist

CAMBIER  
DE NIL  
KNOKKE - ZOUTE

# Outliers ↓

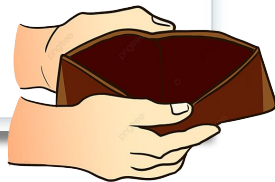


House

**Reserve price : €5,000**

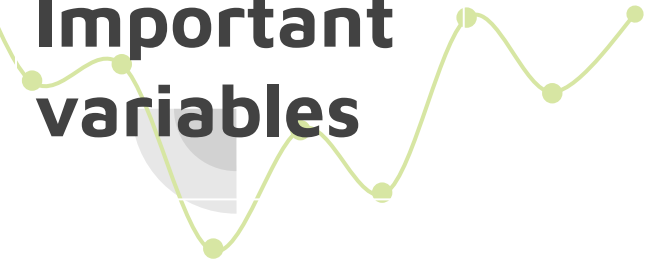
2 bdr. · 121 m<sup>2</sup>

6870 Saint-Hubert



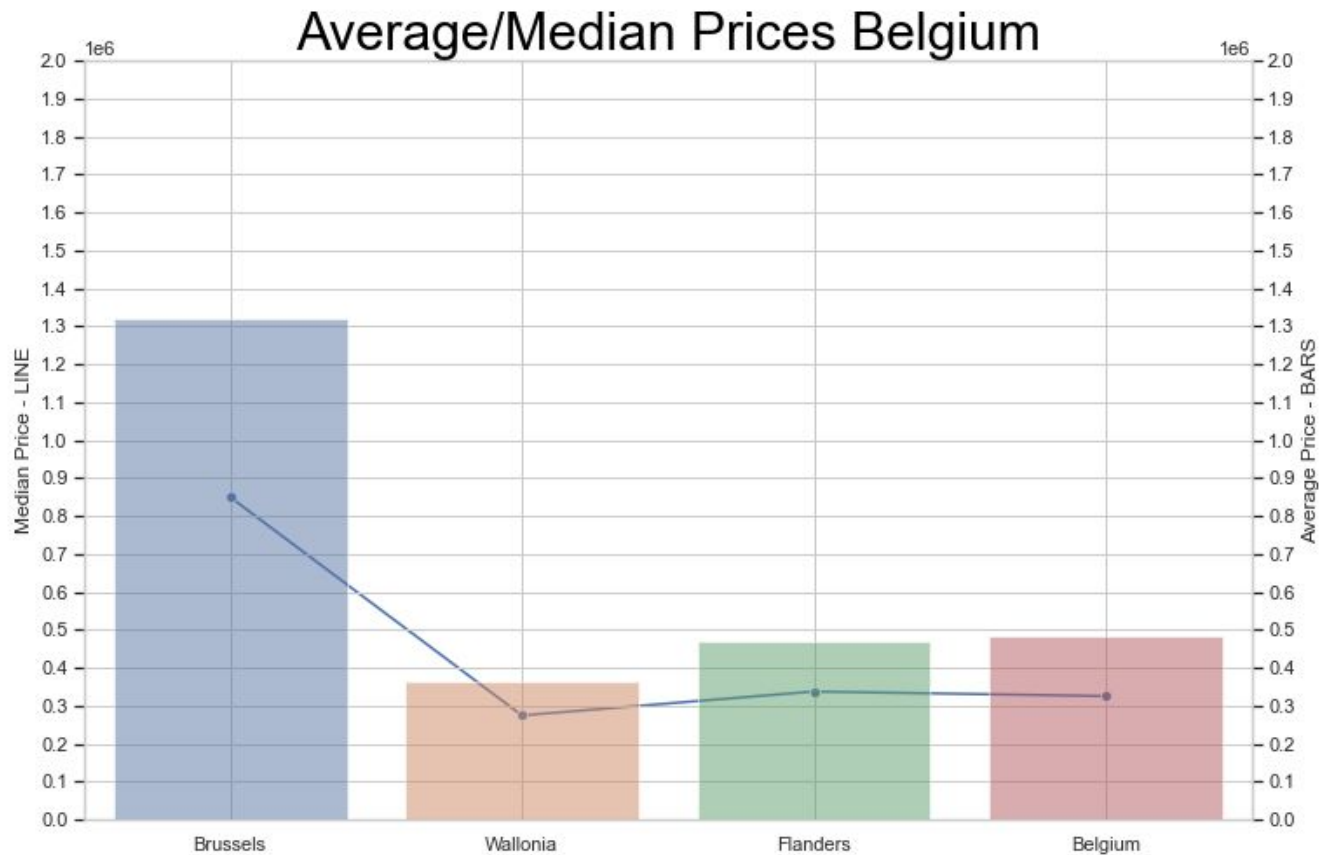
```
area          121.0
price          5000.0
building_state to restore
facades        2.0
bedrooms       2.0
kitchen_equipped undefined
furnished      False
open_fire      1.144966
locality       6870
land_surface   285.0
terrace        False
terrace_surface 0
swimming_pool  False
property_type  house
property_subtype house
garden         False
garden_surface 0
Name: 6399, dtype: object
```

# Important variables





# Prices in Belgium



# Conclusions

