ImmoEliza Data Analysis

Team Tensorflow

Umrah Javed, Kivac Gunduz, Vincent Palau, Aditya Chugh

Cleaning and Assumptions

- Source of data: ImmoWeb
- Data cleaned using Python with Pandas
- Empty values replaced with assumed values

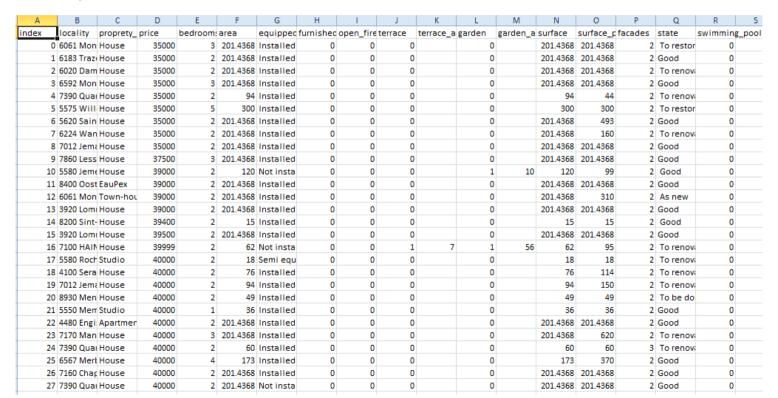
rsurface	surface_p	facades	state s	W
		2	To restore	
		2		
			To renovate	e
94	44	2	To renovate	e
300	300	2	To restore	
	493			
	160		To renovate	e



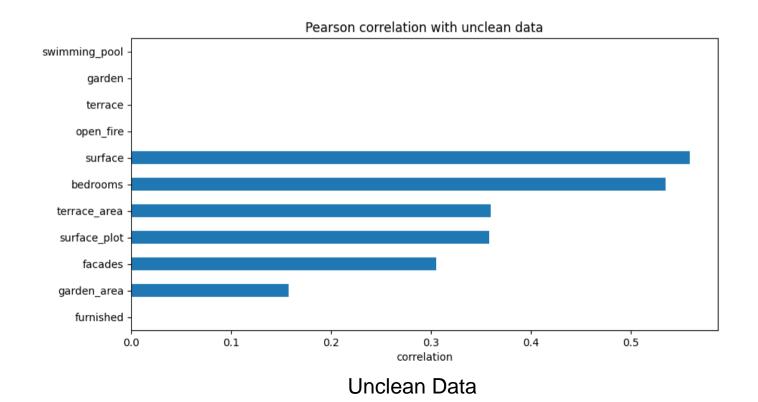
surface	surface_p	facades	state :
201.4368	201.4368	2	To restore
201.4368	201.4368	2	Good
201.4368	201.4368	2	To renova
201.4368	201.4368	2	Good
94	44	2	To renova
300	300	2	To restore
201.4368	493	2	Good
201.4368	160	2	To renova

How many rows and columns?

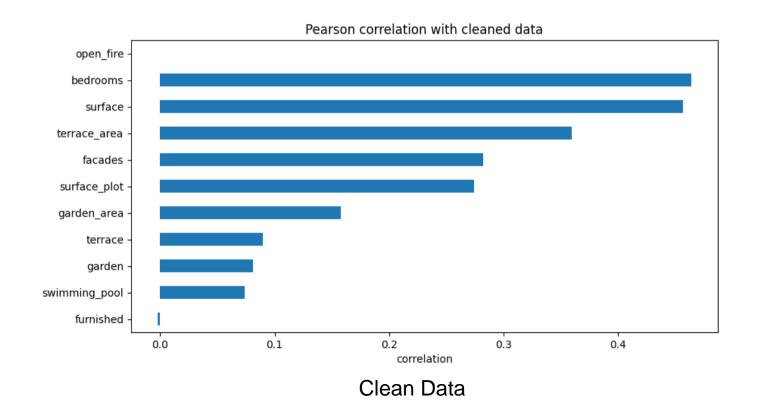
- 8505 rows by 19 columns



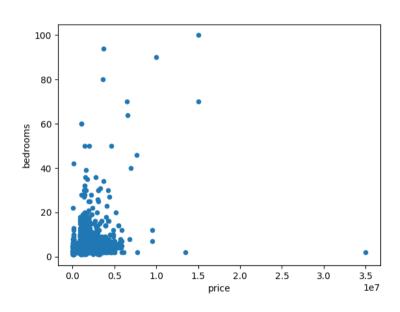
Correlation between the variables and the target

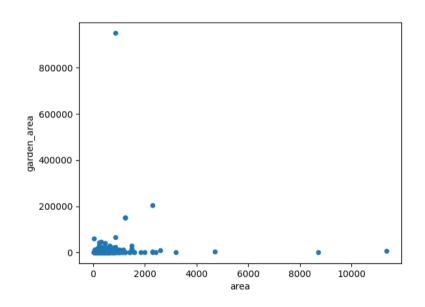


Correlation between the variables and the target



Outliers





Outliers according to bedrooms

Outliers according to garden area

5 most important variables

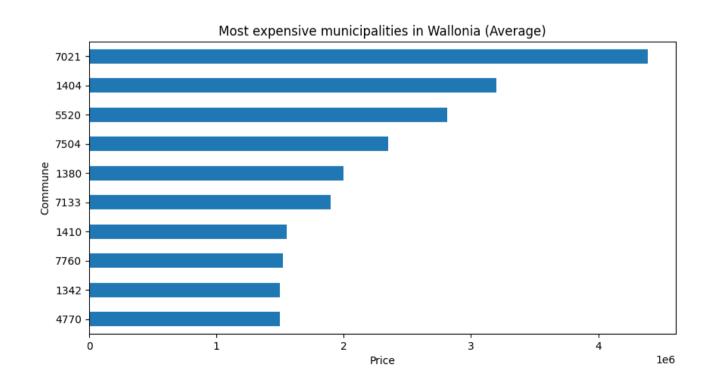
- Price
- Surface
- Bedrooms
- Terrace area
- Surface of the plot

Most expensive municipalities in Belgium

	locality	proprety_type	price l
3144	3150 Haacht	Apartment block	35000000
3803	3150 Haacht	House	1950000
7688	3150 Tildonk	Manor house	1375000

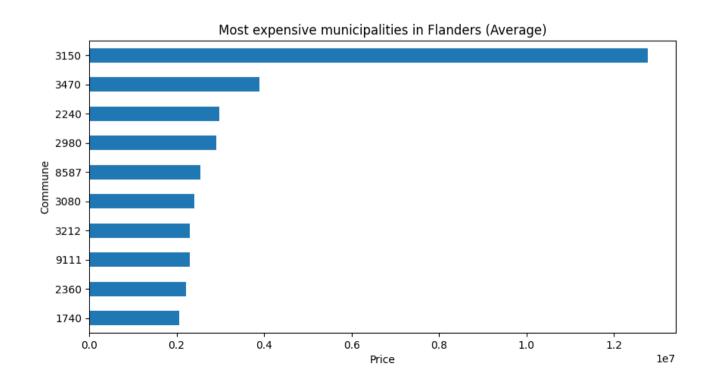
Most expensive municipalities in Wallonia (Average)

- Havré



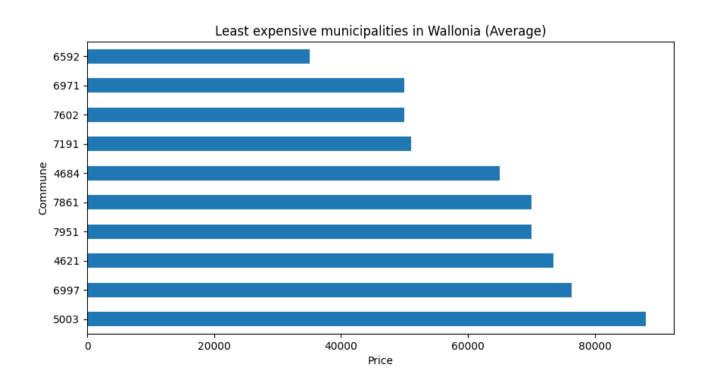
Most expensive municipalities in Flanders (Average)

- Haacht



Least expensive municipalities in Wallonia (Average)

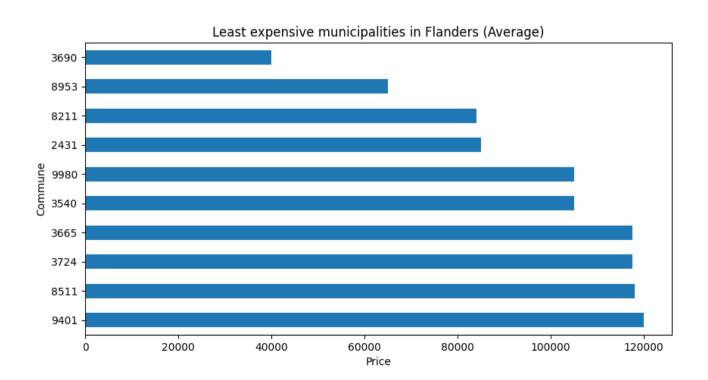
- Imbrechies



Data Interpretation

Least expensive municipalities in Flanders (Average)

- Overijse



Challenges

- Time :D

To Go Further

- Deeper Analysis