

# MACHINE LEARNING

HOUSING PRICE

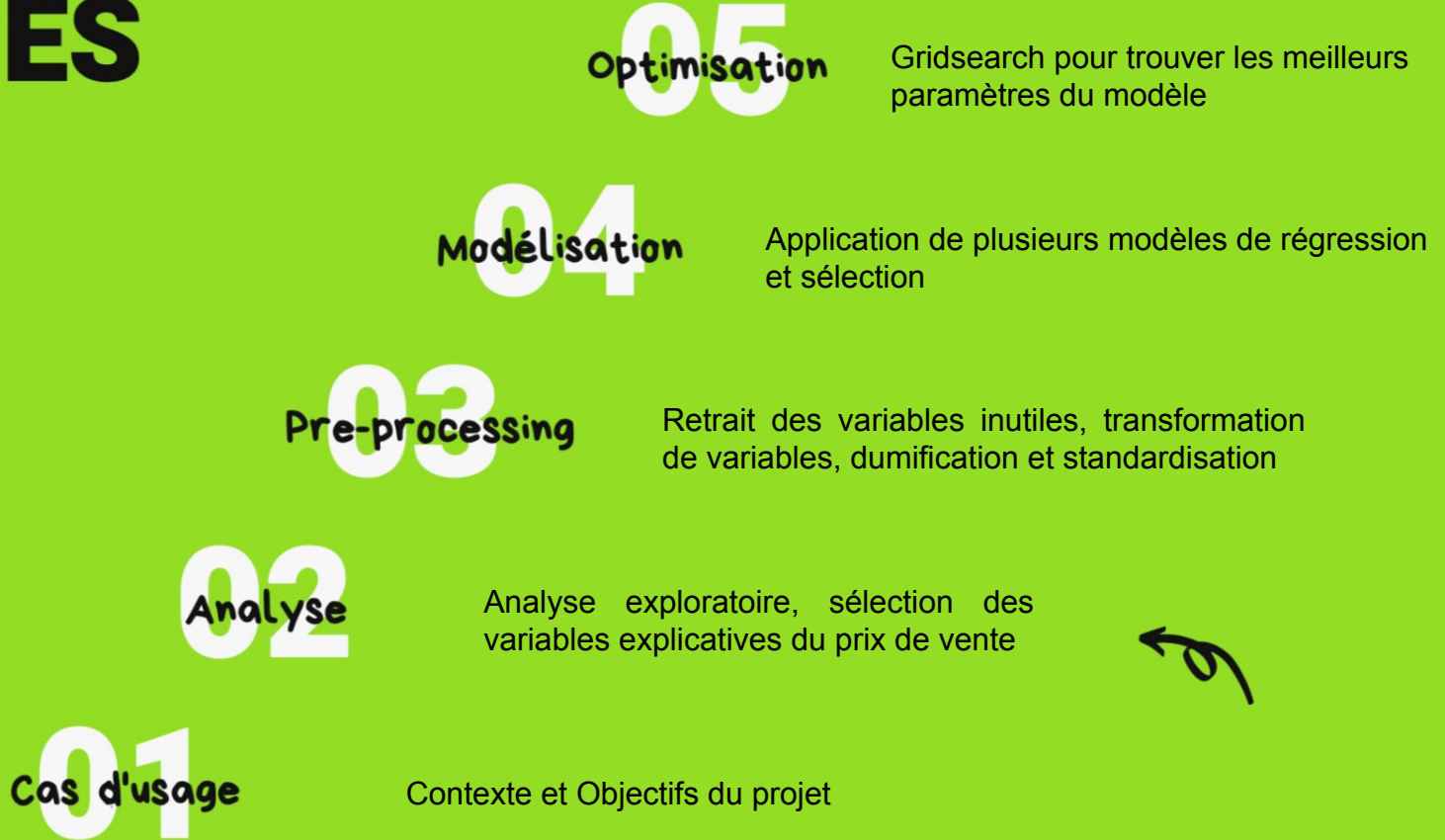


Alexis Da Mota - Rodolphe Lefevre - Palita Marco  
Célina Abidi - Saif Eddine BEN HALIMA



Quel est le prix idéal  
pour cette maison ?

# ETAPES



# CONTEXTE & AGENCE IMMOBILIÈRE OBJECTIFS



## Contexte

- Historique de 1460 biens vendus
- Croître le chiffre d'affaire en optimisant le prix de vente des prochaines acquisitions



## Objectifs

- Déterminer les caractéristiques à privilégier lors de nouveaux achats de biens pour augmenter le chiffre d'affaire
- Développer un outil permettant une estimation simple et rapide du prix de vente optimal de nos propriétés



# ANALYSE EXPLORATOIRE



## Objectif

Confirmer ou infirmer des hypothèses



## Méthodologie

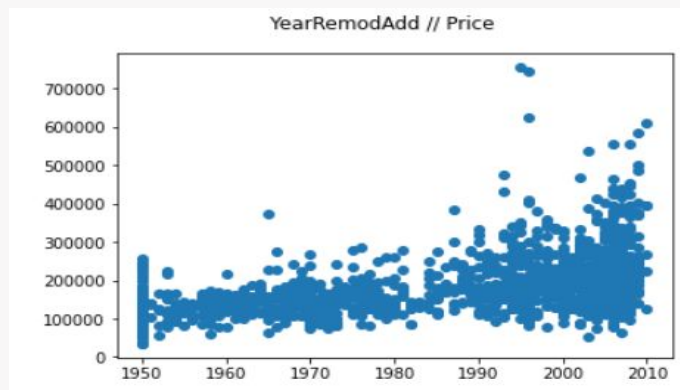
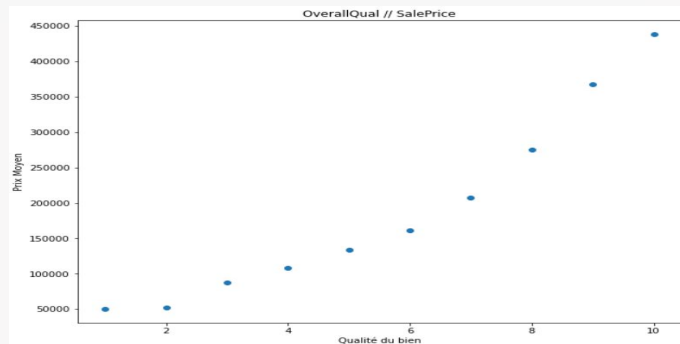
Mise en commun des résultats de :

- graphiques de corrélation  $\text{prix} = f(\text{variable})$
- répartition des valeurs pour chaque variable
- nombre de valeurs manquantes

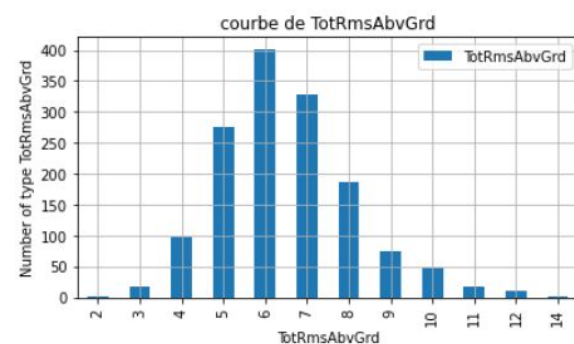
# ANALYSE EXPLORATOIRE



Recommandations/observations  
Graphs en appui



Répartition des valeurs des colonnes

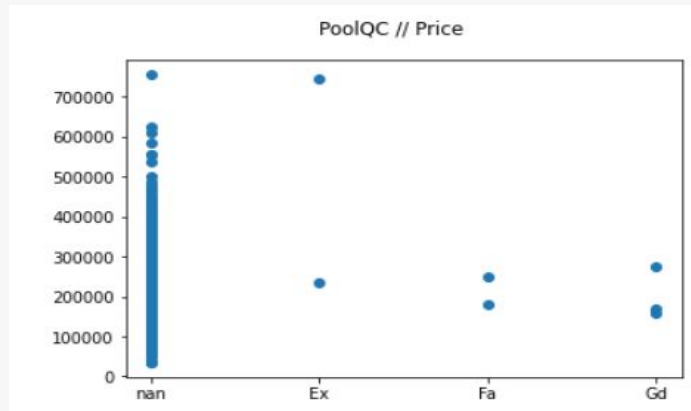


# ANALYSE EXPLORATOIRE



Recommandations/observations  
Graphs en appui

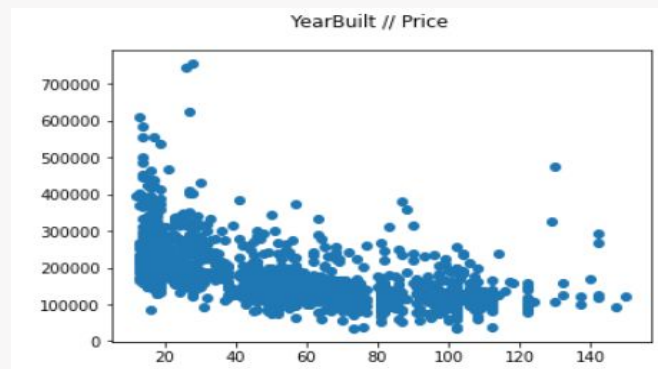
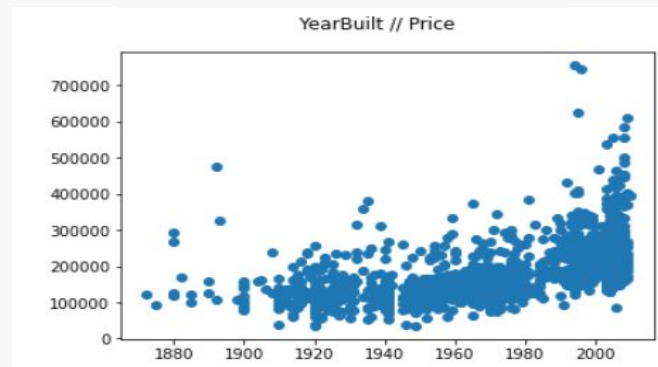
Piscine - répartition mauvaise sur le graphe



Mais avec des maisons similaire ayant ou non une piscine, on a une moyenne de prix:

- maison sans piscine: 180k
- maison avec piscine: 280k

Changement des valeurs



Plus ancien → moins chère

# RECOMMANDATIONS

01

Piscines

Acquérir des biens  
**avec piscines**  
augmenterait en  
moyenne le prix de  
vente de 100 000  
€.

02

Chambres

Acquérir des  
biens avec **4**  
**chambres au**  
**maximum.**  
Au-delà, le prix de  
vente diminue.

03

Ancienneté

Acquérir des biens  
avec **construit**  
**après 1990.**

04

M<sup>2</sup>/Pièces

Acquérir des biens  
avec dont la  
**surface par pièce**  
**est importante.**

# PRE-PROCESSING



01

-----

02

-----

03

-----

04

-----

05

-----

## Supression

des variables non  
explicatives du prix  
de vente : liste des  
variables

## Enrichissement

Nom de colonnes

## Transform

Transformation des  
variables restantes

## Catégorisation

des variables  
explicatives : Nominale,  
ordinaire, numérique

## Dummification, standardisation

```
to_drop = ["Fence", "YrSold", "MoSold", "Condition2", "KitchenAbvGr", "3SsnPorch",  
"PoolArea", "MiscVal", "SaleCondition", "Utilities", "Heating", "HeatingQC", "Street",  
"Alley", "GarageCond", "BsmtCond", "MiscFeature", "YearRemodAdd", "RoofMatl",  
"LandContour", "LotConfig", "Id", "Electrical", "MasVnrArea", "LotFrontage",  
"GarageYrBlt", "GarageQual", "PoolQC"]
```



# MODEL ET CROSS-VALIDATION

- Cross validation paramètre :
- Division 4
- Métriques: mean square error
- Les modèles utilisés:
  - Régression Linéaire
  - Decision tree
  - Random Forest

## Régression Linéaire

```
array([-1.12125384e-01, -1.61692698e+19,  
-7.46847867e+19, -6.97993870e+18])
```

## Random Forest

```
array([-0.14403395, -0.2282756 ,  
-0.17103888, -0.20346908])
```

## Decision tree

```
aarray([-0.26003228, -0.34972067,  
-0.19561401, -0.25529931])
```

# OPTIMISATION



Meilleur modèle : Random Forest

Méthode d'optimisation : Gridsearch

```
parameters = {  
    'n_estimators':[80, 100, 150], 'max_depth': [4, 5, 6, 8],  
    'min_samples_split': [2,3]}
```

## Résultats:

meilleur score : -0.14250268084417836

## Meilleur estimateur :

```
RandomForestRegressor(max_depth=8, min_samples_split=3,  
n_estimators=150)
```

```
result.best_score_  
-0.14250268084417836
```

```
rclf.best_estimator_  
RandomForestRegressor(max_depth=8,  
min_samples_split=3, n_estimators=150)
```

# PRODUIT FINAL



**Maquette**

ACPRS Immobilier

Accueil	Client	Habitation	Consultation	Estimation	...	...	...
---------	--------	------------	--------------	------------	-----	-----	-----

Type de Logement :

Choisir ▼

Duplex ☐

... ☐

... ☐

... ☐

Date de Construction :

Date de Rénovation \* :

Surface (m<sup>2</sup>) :

Fondation :

Choisir ▼

Style du Logement :

Choisir ▼

Proximité :

Choisir ▼

Suivant ►

# PRODUIT FINAL



Maquette


ACPRS Immobilier

Accueil	Client	Habitation	Consultation	Estimation	...	...	...
---------	--------	------------	--------------	------------	-----	-----	-----

Type de Chauffage :

Date de Construction :

Climatisation centrale :  
Oui ☐ Non ☐

ESTIMER LE BIEN 

Valeur du bien estimé :

Choisir ▼

Gaz ☐

... ☐

... ☐

... ☐

**MERCI POUR**  
**VOTRE ATTENTION**

HOUSE PRICING

N'hésitez si vous avez des questions

