House prices

2024-03-05

Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	...	MiscFeature	MiscVal	MoSold	YrSold	SaleType	SaleCondition	SalePrice
0	60	RL	65.0	8450	Pave	...	NaN	0	2	2008	WD	Normal	208500
1	20	RL	80.0	9600	Pave	...	NaN	0	5	2007	WD	Normal	181500
2	60	RL	68.0	11250	Pave	...	NaN	0	9	2008	WD	Normal	223500
3	70	RL	60.0	9550	Pave	...	NaN	0	2	2006	WD	Abnorml	140000
4	60	RL	84.0	14260	Pave	...	NaN	0	12	2008	WD	Normal	250000

	Variable 1	Variable 2	Corrélation
4	GarageCars	GarageArea	0.882475
1	YearBuilt	GarageYrBlt	0.825667
3	GrLivArea	TotRmsAbvGrd	0.825489
2	TotalBsmtSF	1stFlrSF	0.819530
0	OverallQual	SalePrice	0.817185

PoolQC	nombre total de données manquantes
MiscFeature	2786
Alley	2690
Fence	2322
MasVnrType	1750
FireplaceQu	1415
LotFrontage	476
GarageYrBlt	155
GarageFinish	155
GarageQual	155
GarageCond	155
GarageType	153
BsmtExposure	81
BsmtCond	81
BsmtQual	80
BsmtFinType2	79
BsmtFinType1	78
MasVnrArea	22
MSZoning	4
BsmtHalfBath	2

Feature	Pourcentage de données manquantes
PoolQC	99.792100
MiscFeature	96.534997
. . .	. . .
BsmtFinType1	2.702703
MasVnrArea	0.762301
MSZoning	0.138600
BsmtHalfBath	0.069930

Feature	Count
BsmtCond	81
BsmtExposure	81
BsmtQual	80
BsmtFinType2	79
BsmtFinType1	78
Foundation	..
Heating	Ø
HeatingQC	Ø
CentralAir	Ø
SaleCondition	Ø

	Modèle	Moyenne CV RMSE	RMSLE Ensemble Données	% Prédictions Correctes ±5%
0	XGBRegressor	0.128751	0.093918	45.550105
1	LGBMRegressor	0.110928	0.069386	68.044849
2	Ridge	0.110566	0.087891	50.735809
3	Lasso	0.111788	0.087409	50.805886
4	GradientBoostingRegressor	0.109919	0.073125	84.723196
5	RandomForestRegressor	0.136842	0.086273	56.131745
6	SVR	0.107887	0.088619	57.603364
7	ElasticNet	0.111487	0.087443	51.016118
8	DecisionTreeRegressor	0.176293	0.089538	50.105116