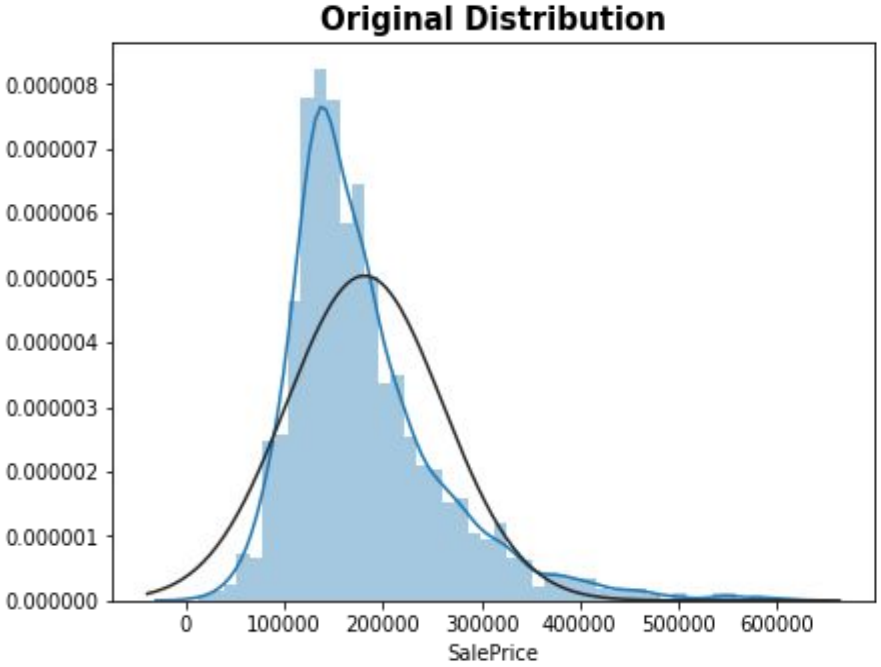
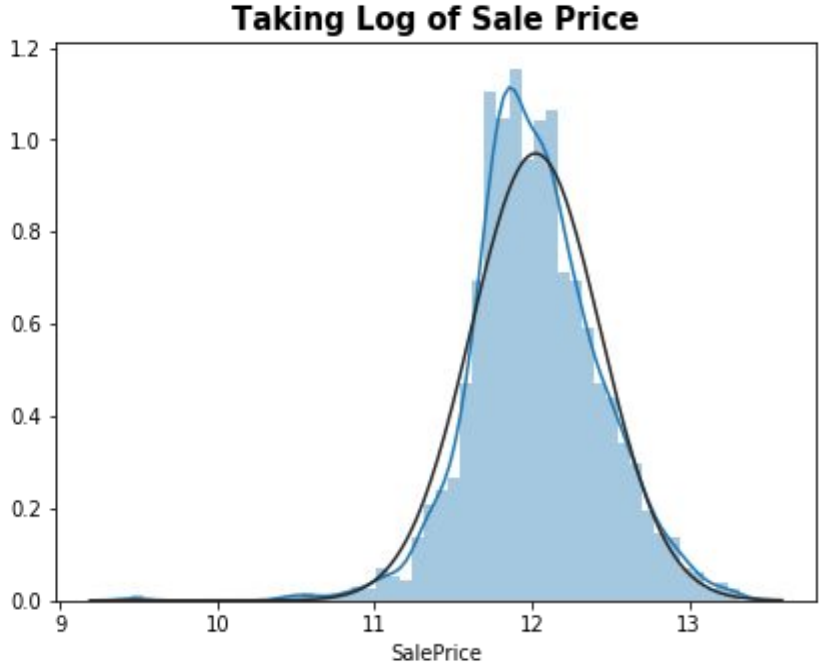
AMES HOUSING

Kaggle Challenge

Target Variable: Sale Price





Target Variable Outliers

Lot Frontage	330
Garage Yr Blt	114
Mas Vnr Area	22
Bsmt Half Bath	2
Bsmt Full Bath	2
Garage Cars	1
BsmtFin SF 1	1
BsmtFin SF 2	1
Bsmt Unf SF	1
Total Bsmt SF	1
Garage Area	1
Lot Area	0

dtype: int64

Pool QC	2040
Misc Feature	1984
Alley	1909
Fence	1649
Fireplace Qu	999
Garage Cond	114
Garage Qual	114
Garage Finish	114
Garage Type	113
Bsmt Exposure	57
BsmtFin Type 2	55
BsmtFin Type 1	54
Bsmt Cond	54
Bsmt Qual	54
Mas Vnr Type	22

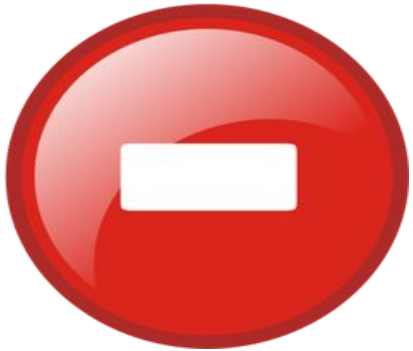
Missing Values:

Feature Engineering

- **Total SF**
- **Total Bath**
- **Total Porch**

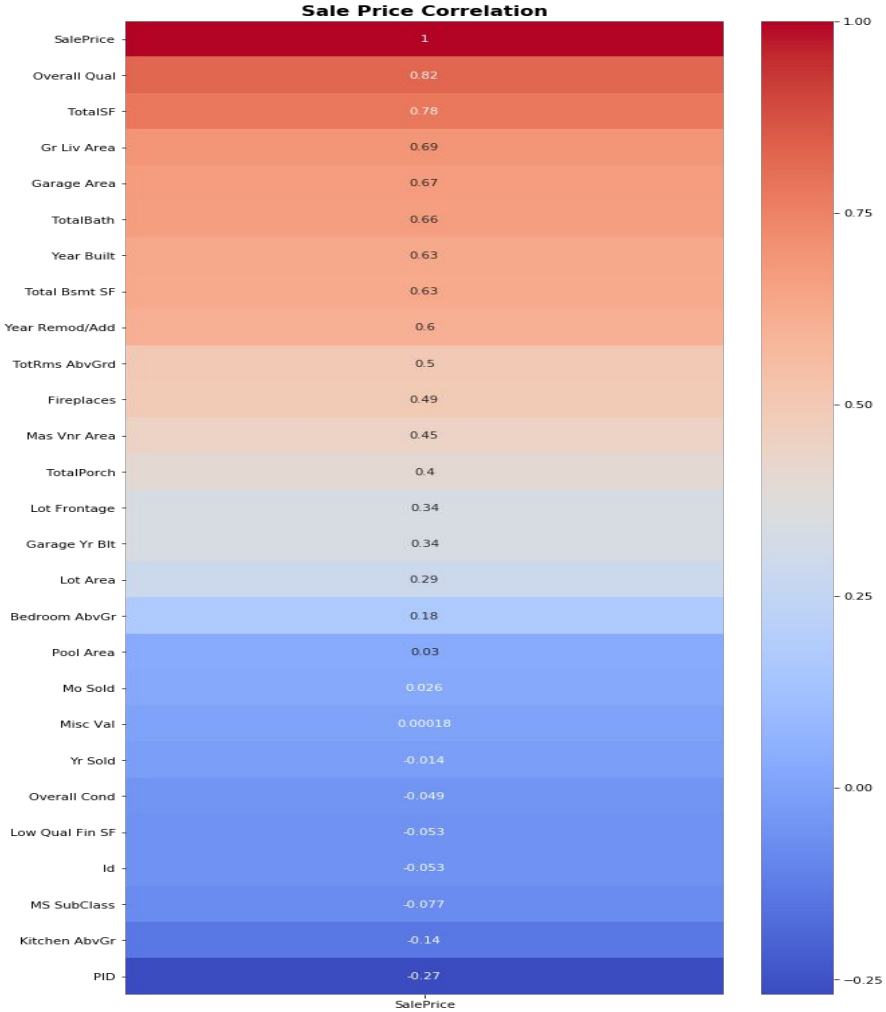
- 1st Flr SF
- 2nd Flr SF
- Bsmt Full Bath
- Bsmt Half Bath
- Full Bath
- Half Bath

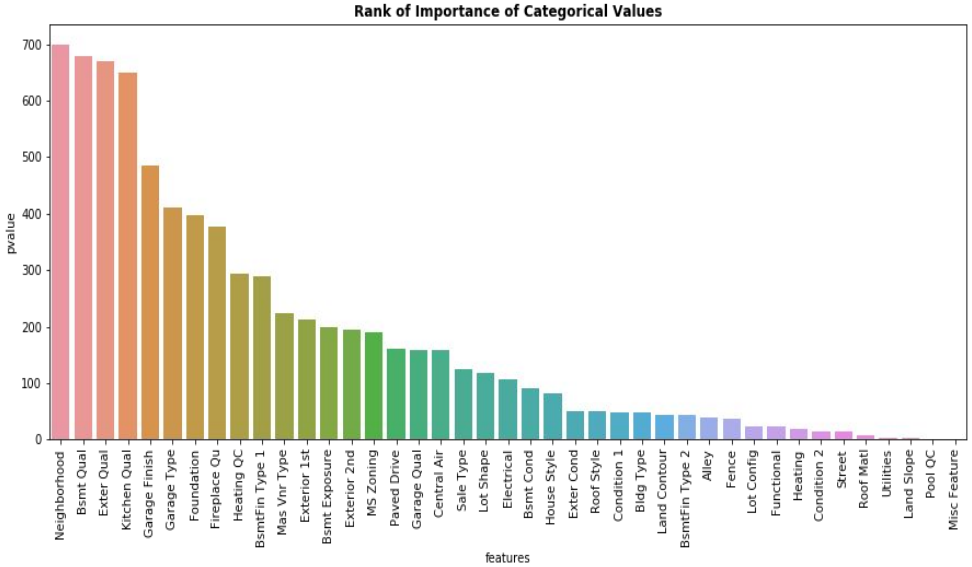
- Wood Deck SF
- Open Porch SF
- Enclosed Porch
- 2Ssn Porch
- Screen Porch



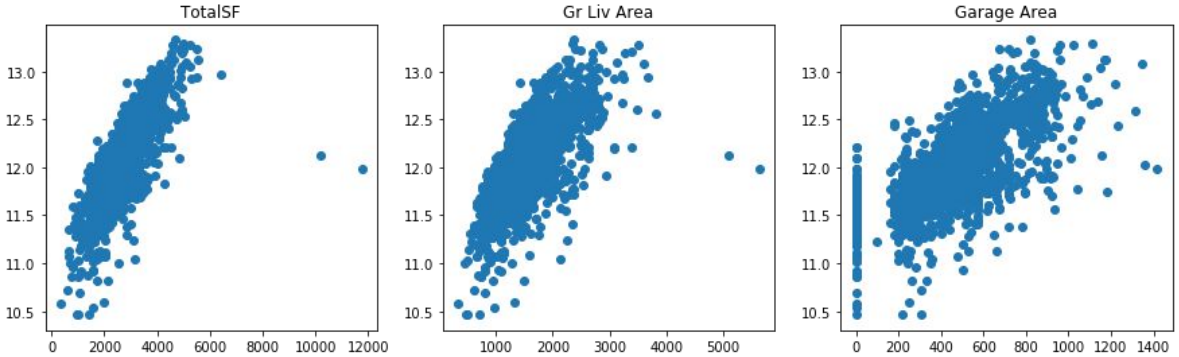
Garage Cars
BsmtFin SF 1
BsmtFin SF 2
Bsmt Unf SF
Garage Cond

Correlation





Outliers:



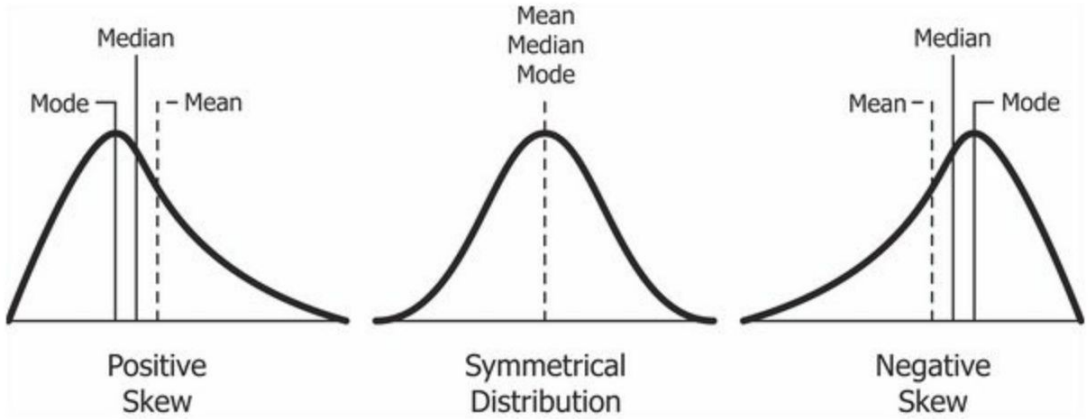
Total SF > 5949.5

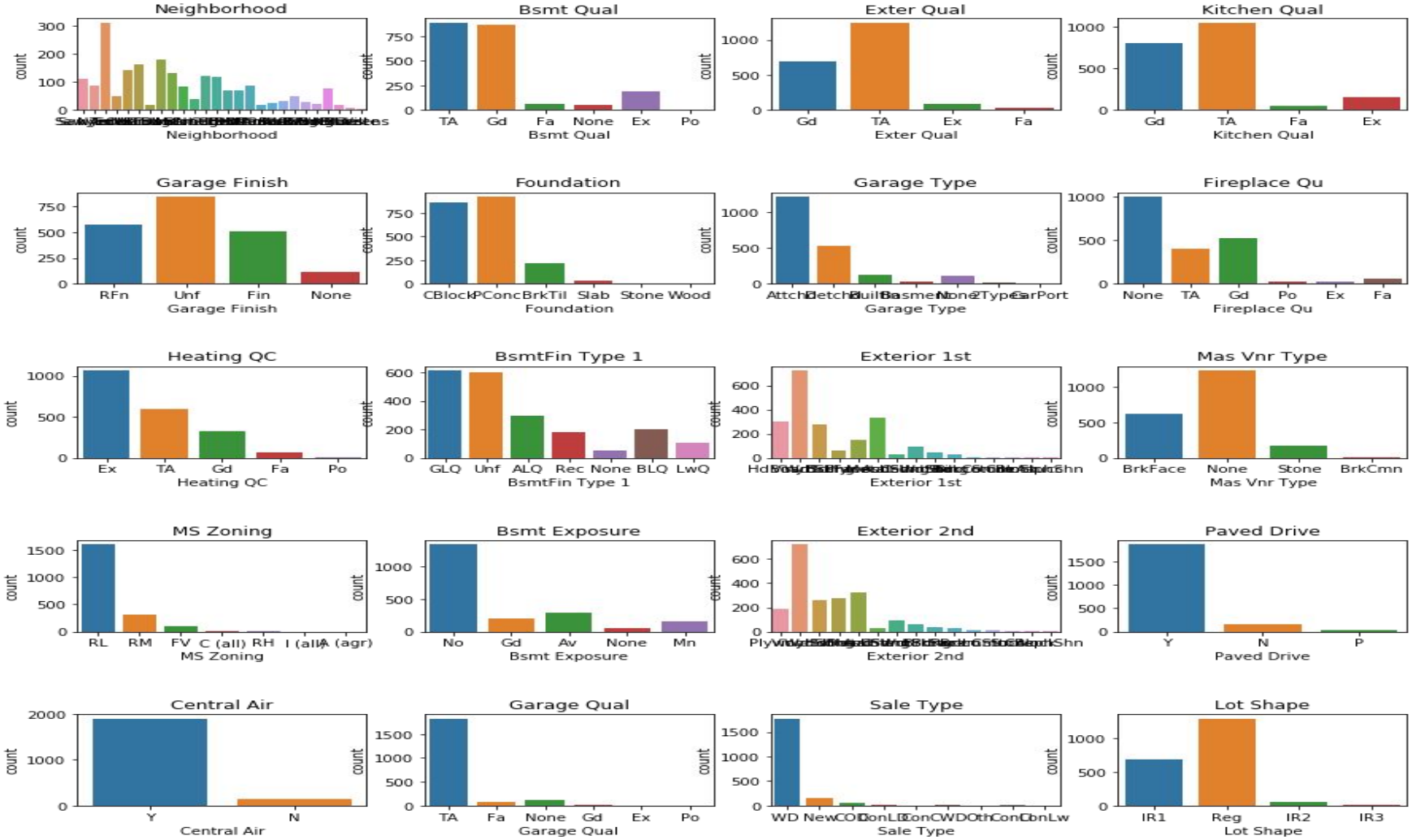
Mas Vnr Area > 636.5

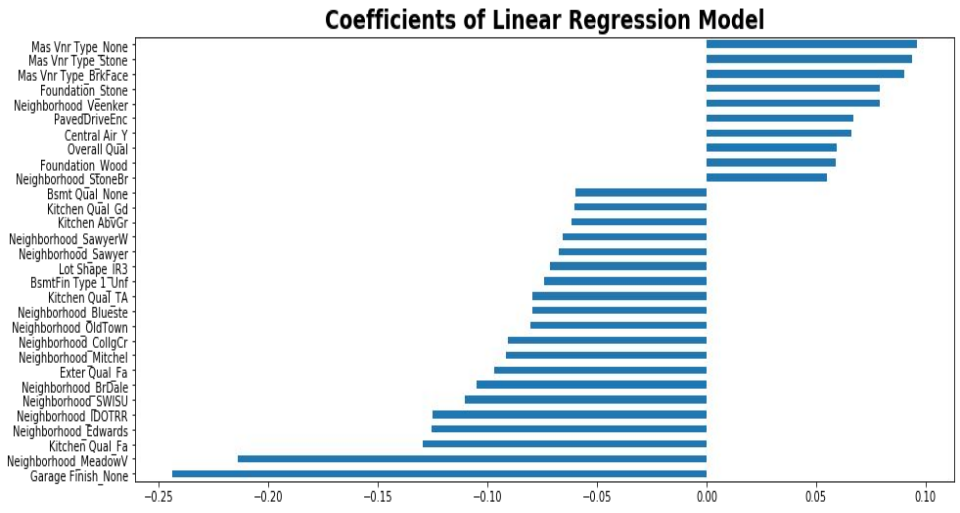
Total Porch > 924.25

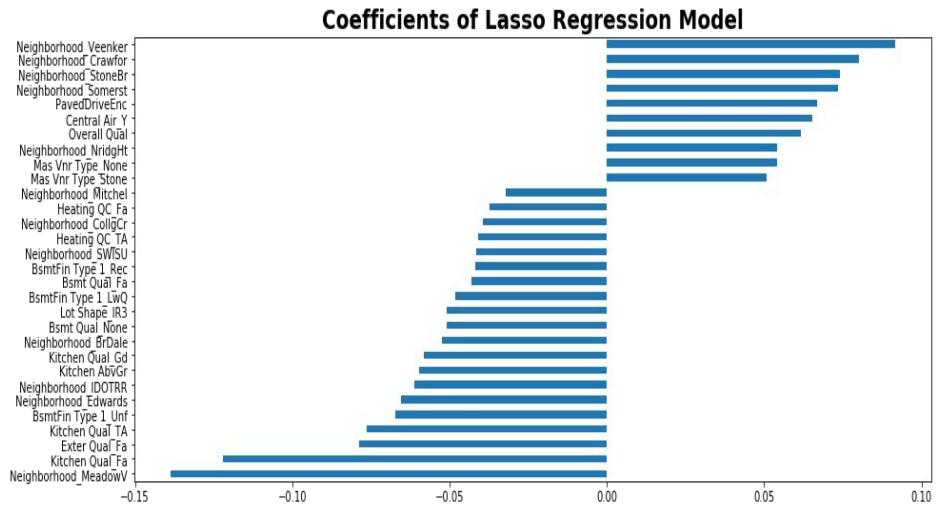
Skewed Features:

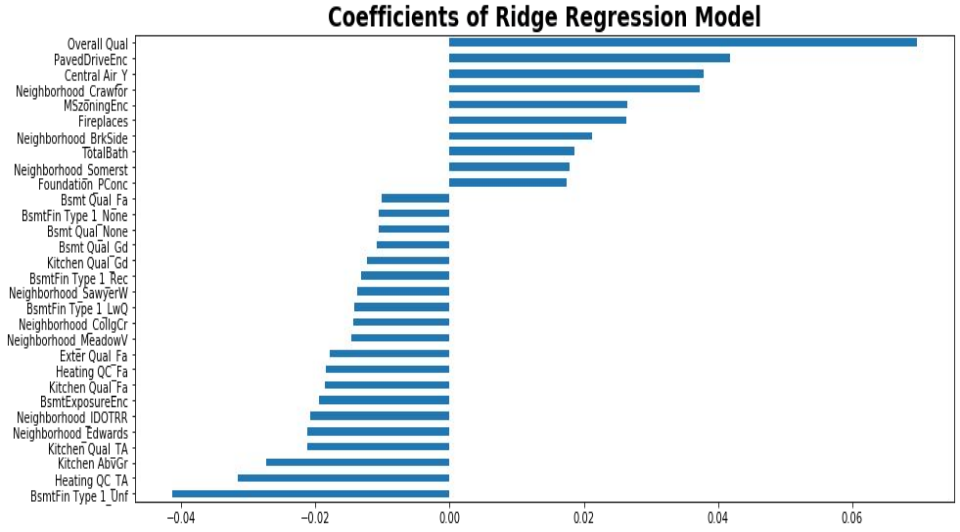
Low Qual Fin SF	11.486985
Lot Area	9.202783
Kitchen AbvGr	4.342167
Mas Vnr Area	2.610229
Lot Frontage	1.835751
TotalSF	1.683826
Total Bsmt SF	1.397878
TotalPorch	1.388651
MS SubClass	1.386901
Gr Liv Area	1.281737
TotRms AbvGrd	0.842221
Fireplaces	0.725669
TotalBath	0.469416
Bedroom AbvGr	0.371753
Garage Area	0.195903
Overall Qual	0.177343
SalePrice	0.074681
PID	0.069386
Id	-0.010003
Year Remod/Add	-0.451544
Year Built	-0.607855
Garage Yr Blt	-3.857371




Neighborhood
count
Neighborhood
Bsmt Qual
TA Gd Fa None Ex Po
Bsmt Qual
Exter Qual
Gd TA Ex Fa
Exter Qual
Kitchen Qual
Gd TA Fa Ex
Kitchen Qual
Garage Finish
RFn Unf Fin None
Garage Finish
Foundation
CBlock PConc BrkTil Slab Stone Wood
Foundation
Garage Type
Attchd Detchd BuiltIn Basment None 2Types CarPort
Garage Type
Fireplace Qu
None TA Gd Po Ex Fa
Fireplace Qu
Heating QC
Ex TA Gd Fa Po
Heating QC
BsmtFin Type 1
GLQ Unf ALQ Rec None BLQ LwQ
BsmtFin Type 1
Exterior 1st
Exterior 1st
Mas Vnr Type
BrkFace None Stone BrkCmn
Mas Vnr Type
MS Zoning
RL RM FV C (all) RH I (all) A (agr)
MS Zoning
Bsmt Exposure
No Gd Av None Mn
Bsmt Exposure
Exterior 2nd
Exterior 2nd
Paved Drive
Y N P
Paved Drive
Central Air
Y N
Central Air
Garage Qual
TA Fa None Gd Ex Po
Garage Qual
Sale Type
WD New COD ConLD ConLI ConLw CWD Oth ConLw
Sale Type
Lot Shape
IR1 Reg IR2 IR3
Lot Shape







Method	R^2_train	R^2_test	RMSE
Linear Regression	0.9064	0.9081	0.1176
Ridge	0.9066	0.9008	0.1221
Lasso	0.9158	0.9084	0.1174
Elastic Net	0.9084	0.9084	0.1174(a little higher than lasso