Ames Housing Project

Created by:Lubna Asmi

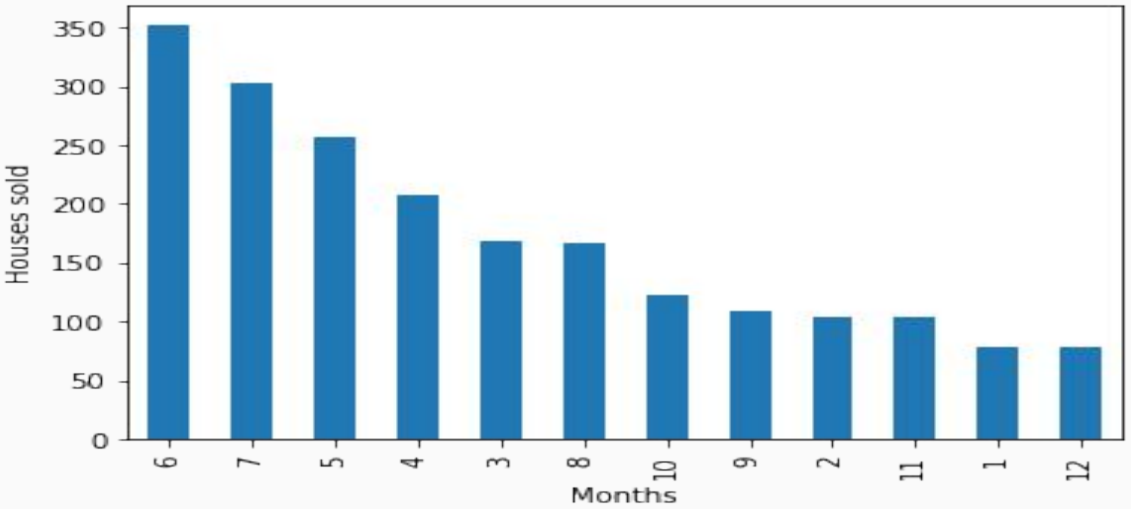
Problem Statement

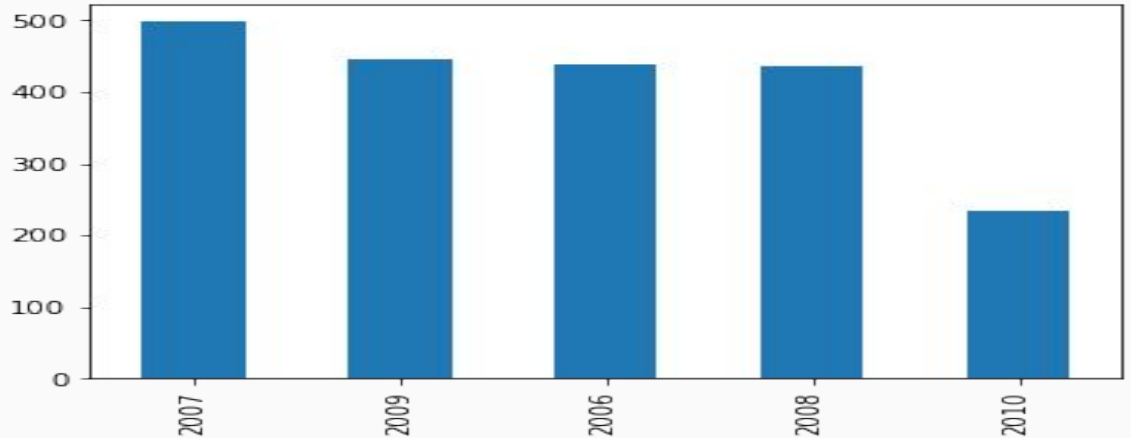
Data Analysis on Ames Housing Dataset.

Building Linear Regression model to predict house Sale Price.

DATA ANALYSIS:

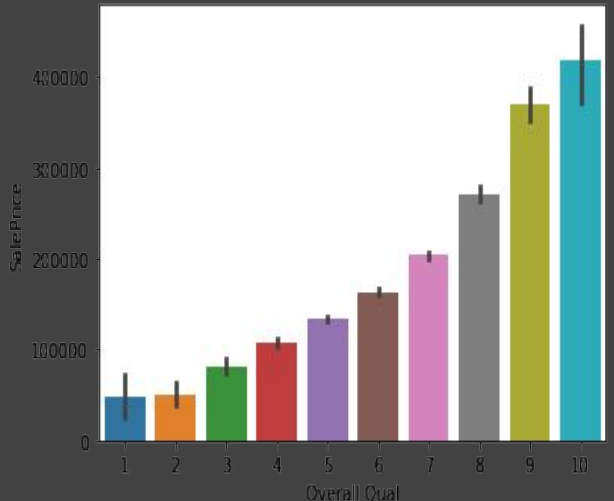
Houses sold max in June,July and least in Dec,Jan.
-Houses sold max in 2007 (498 houses) and least in 2010 (234 houses)

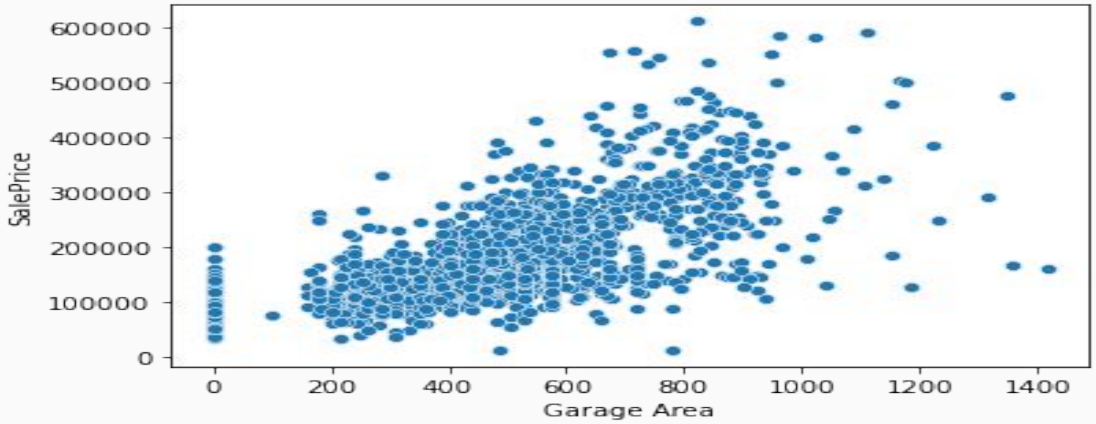


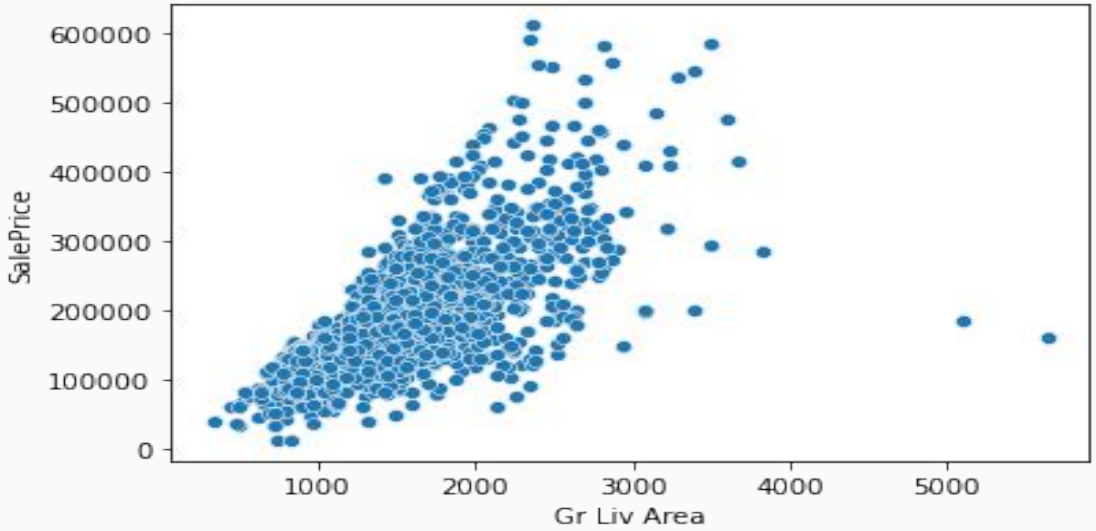


Linear Relation:

-Sales Price and Garage Area.
-Sales Price and Gr Living Area.
-Sales Price and total Basement Sq__f.
-SalePrice and Overall quality.

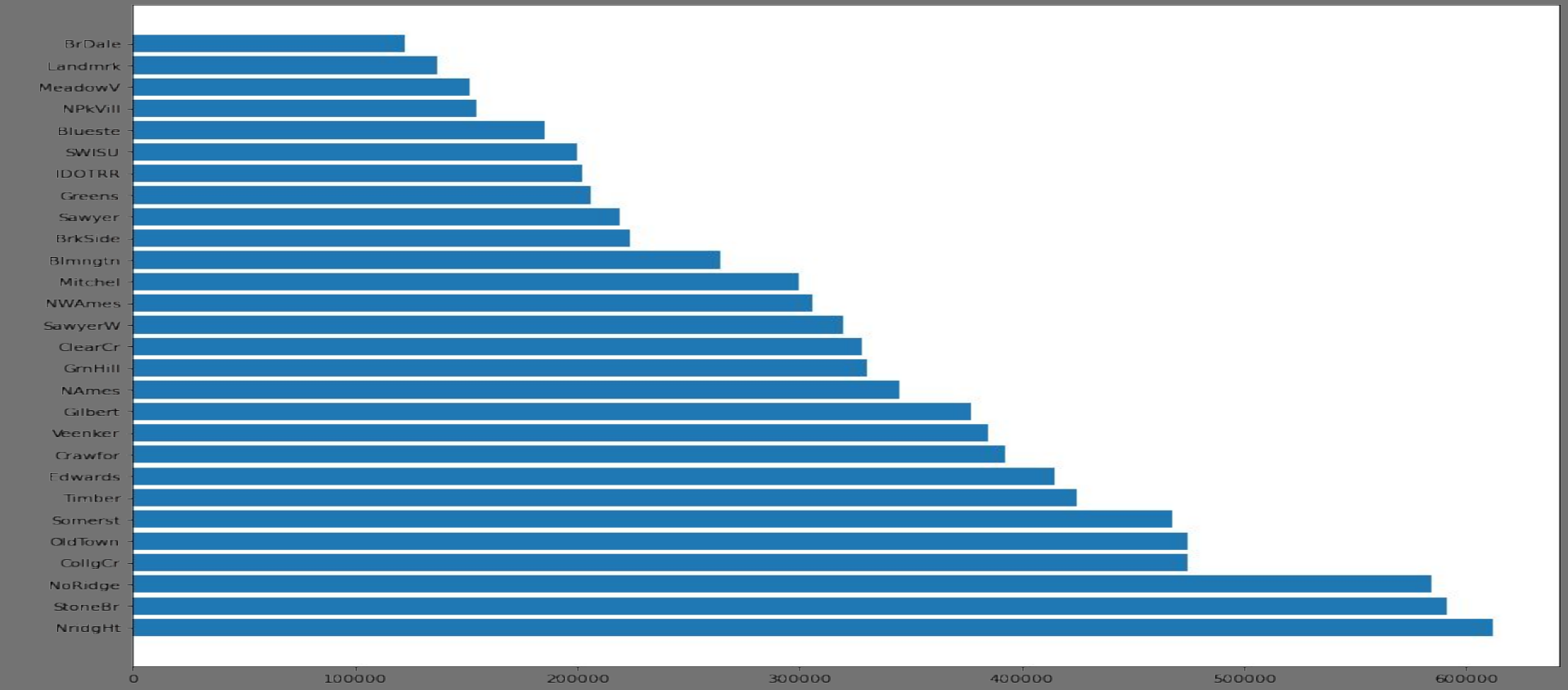


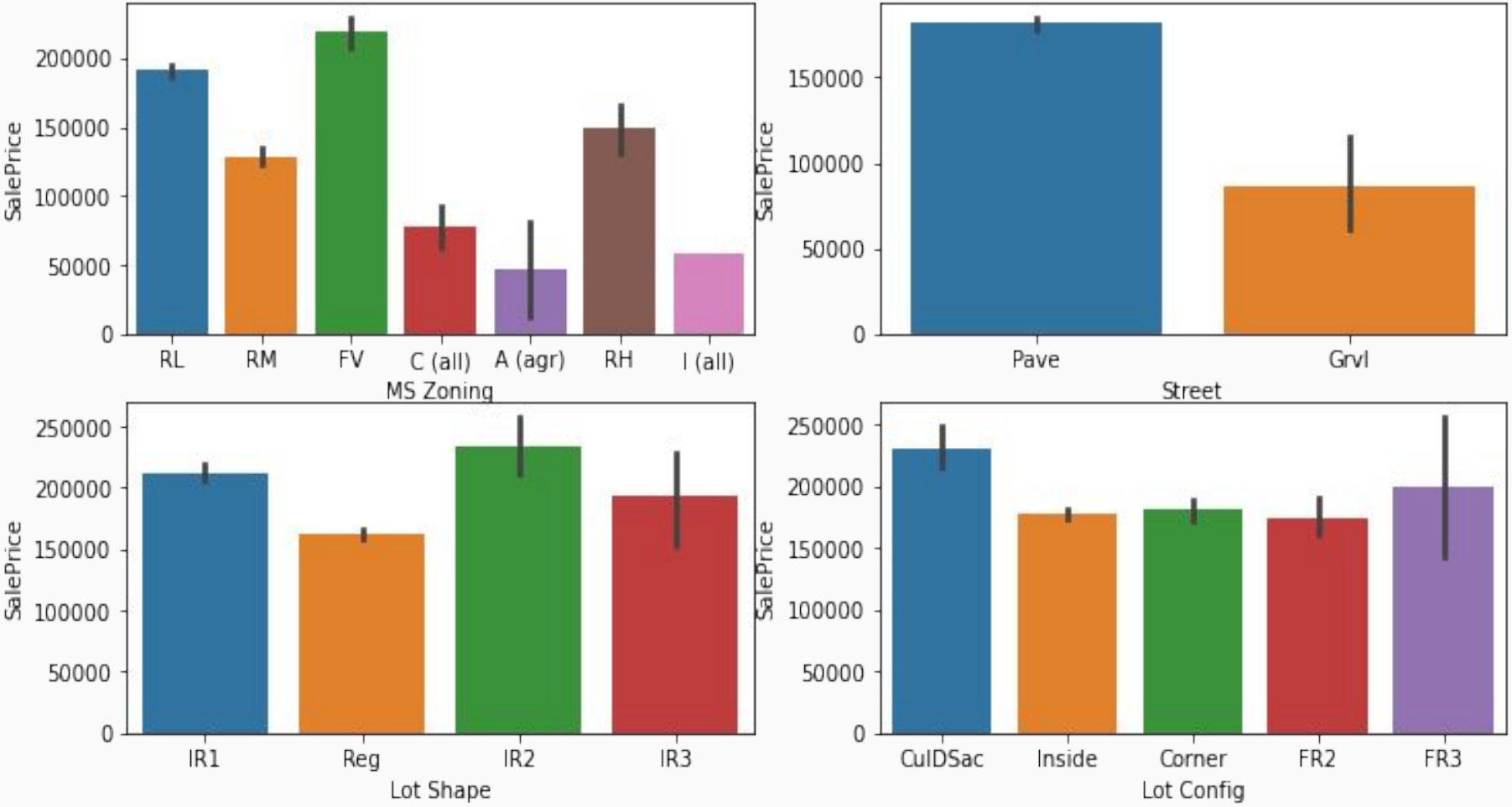


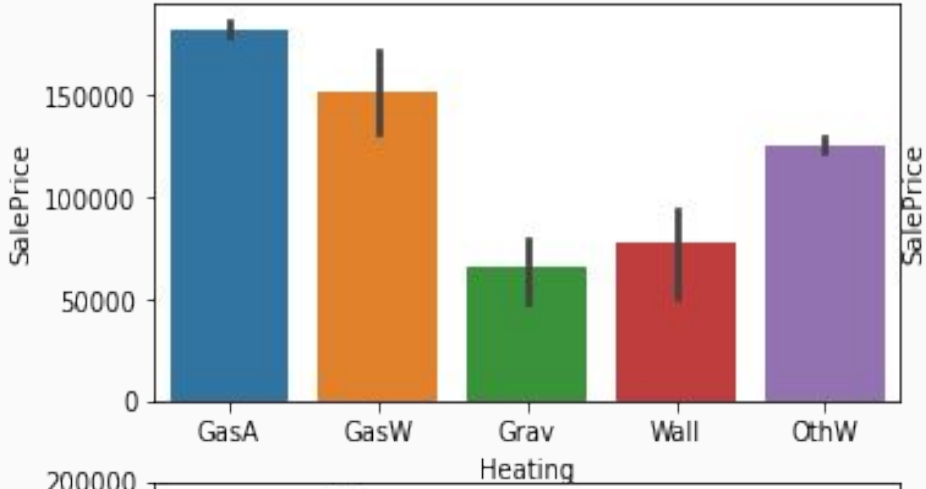


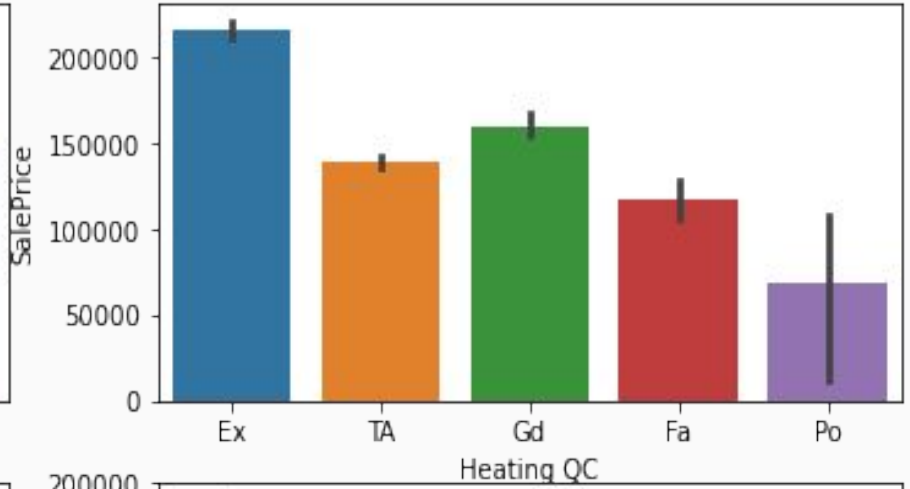
Neighborhood:

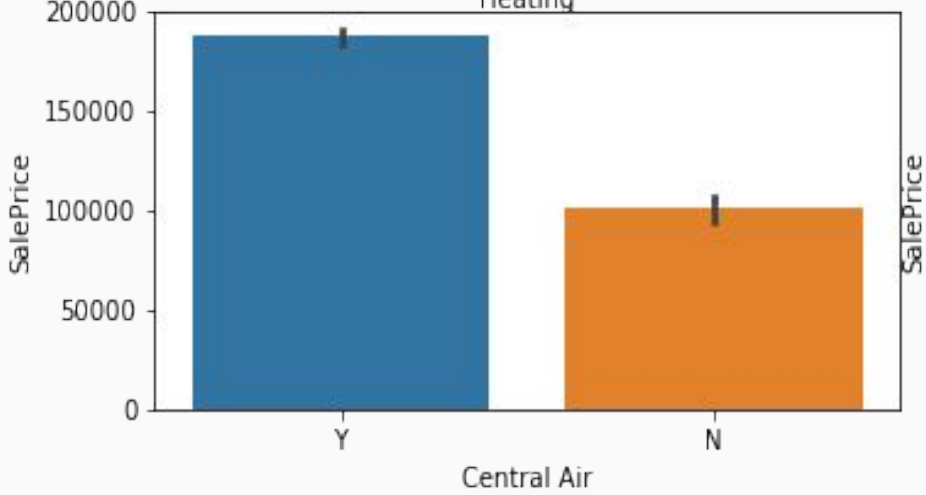
Northridge Heights got maximum Sale Price and minimum in Briardale.Total sales is low in Landmark.

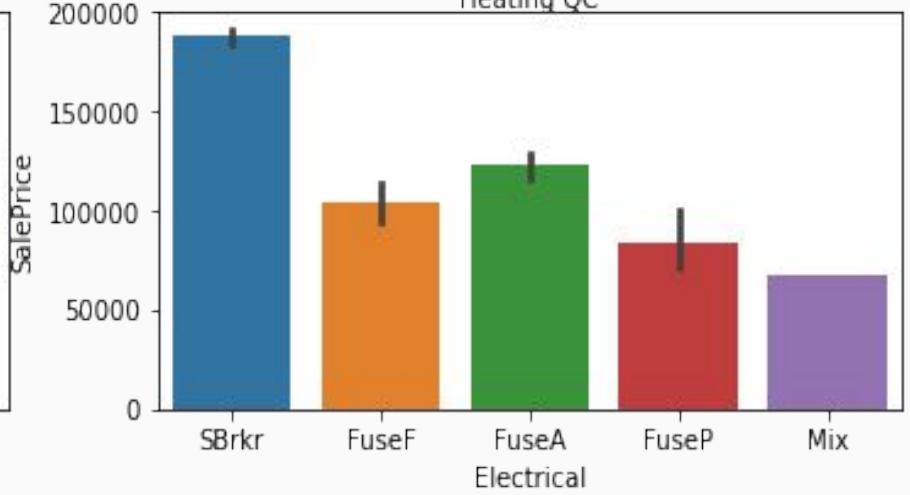


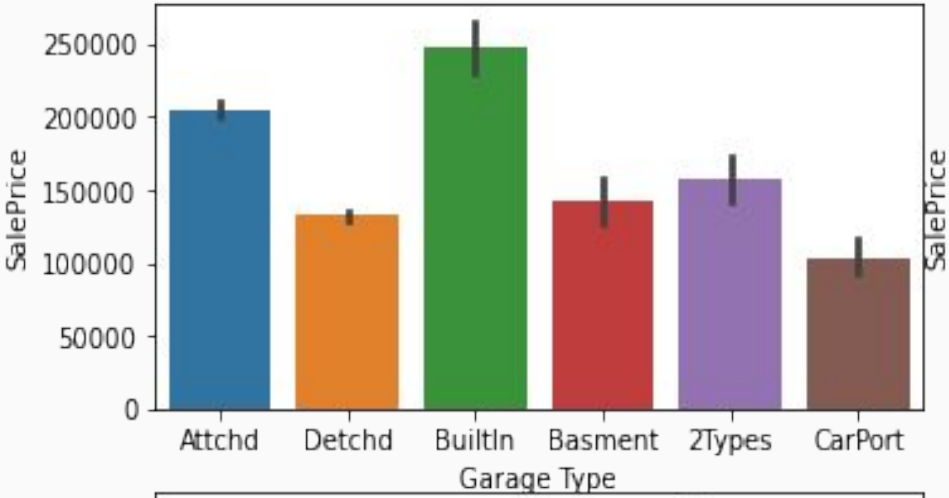

SalePrice
200000
150000
100000
50000
0
RL
RM
FV
C (all)
A (agr)
RH
I (all)
MS Zoning
SalePrice
150000
100000
50000
0
Pave
Grvl
Street
SalePrice
250000
200000
150000
100000
50000
0
IR1
Reg
IR2
IR3
Lot Shape
SalePrice
250000
200000
150000
100000
50000
0
CulDSac
Inside
Corner
FR2
FR3
Lot Config

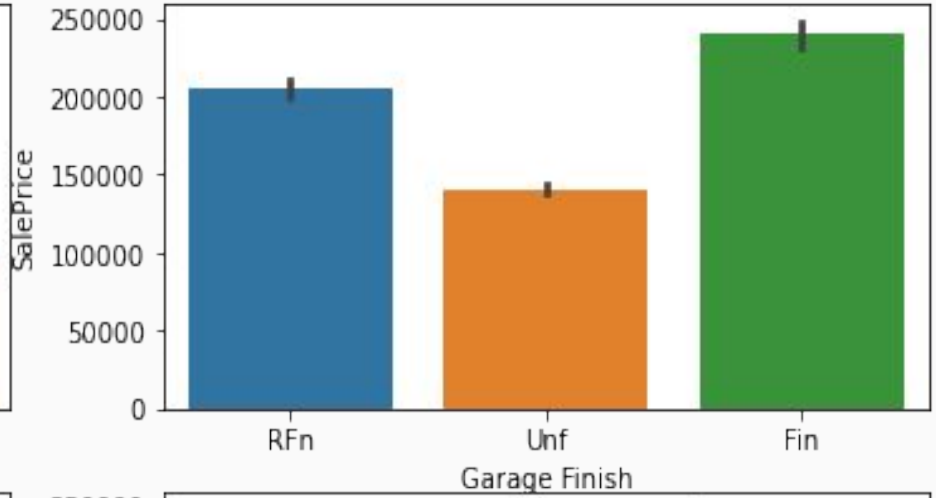

SalePrice
150000
100000
50000
0
GasA
GasW
Grav
Wall
OthW
Heating

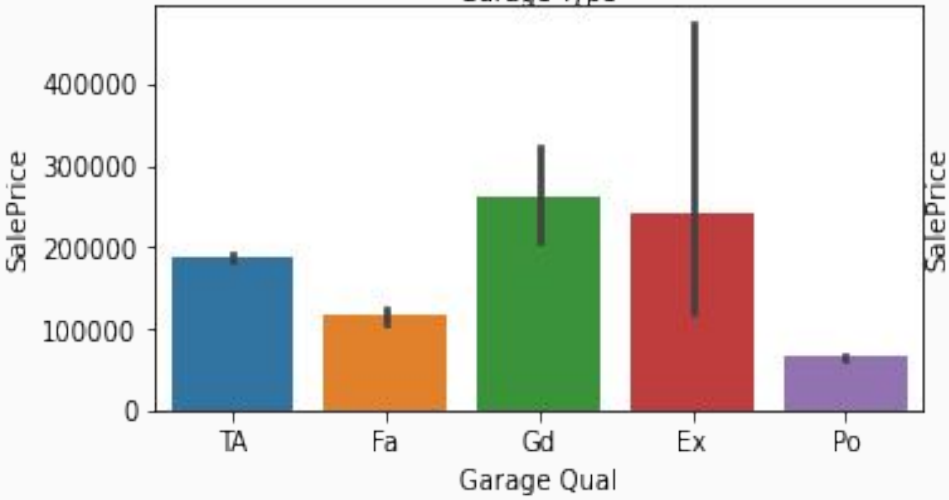

SalePrice
200000
150000
100000
50000
0
Ex
TA
Gd
Fa
Po
Heating QC

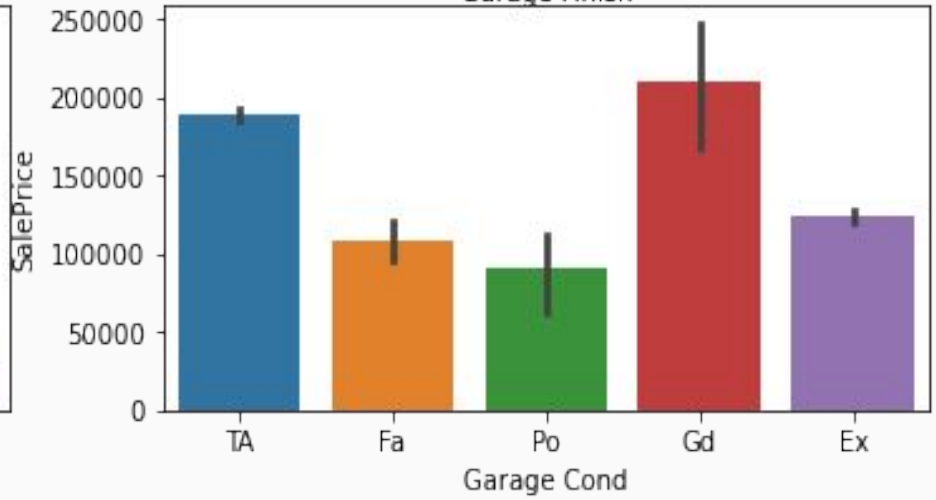

SalePrice
200000
150000
100000
50000
0
Y
N
Central Air


SalePrice
200000
150000
100000
50000
0
SBrkr
FuseF
FuseA
FuseP
Mix
Electrical


SalePrice
250000
200000
150000
100000
50000
0
Attchd
Detchd
BuiltIn
Basment
2Types
CarPort
Garage Type


SalePrice
250000
200000
150000
100000
50000
0
RFn
Unf
Fin
Garage Finish


SalePrice
400000
300000
200000
100000
0
TA
Fa
Gd
Ex
Po
Garage Qual


SalePrice
250000
200000
150000
100000
50000
0
TA
Fa
Po
Gd
Ex
Garage Cond

MODELLING:

Building a linear regression model to predict house Sale Price.

Model1

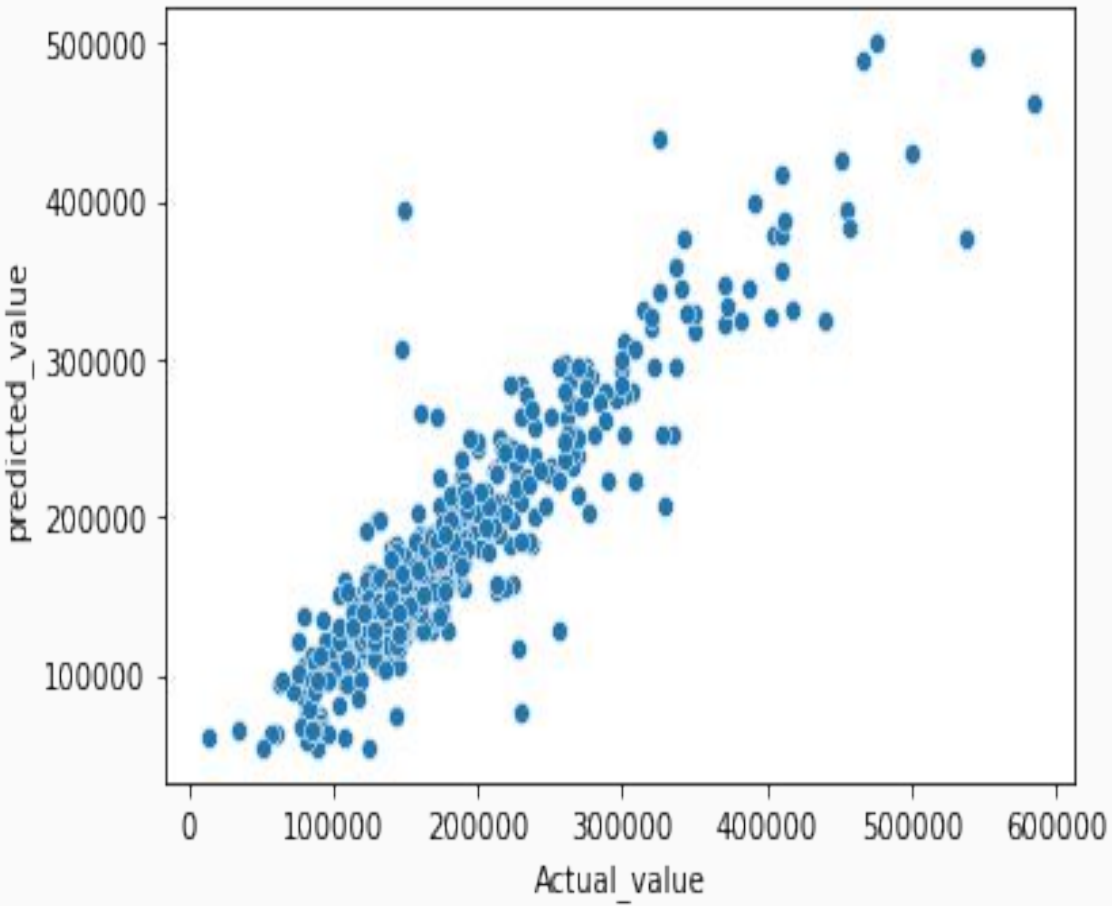
cross_val_score:85.49%

R2 for train_data:86.31%

R2 for test_data:84.25%

Rmse:32695

One unit increase in Gr Liv Area and Garage Area Sale Price will increase by 14275.
Total Bsmt SF and Garage Area- 5565.



RIDGE And LASSO:

Ridge_R2 for train_data:86.3%

Ridge_R2 for test_data:84.2%

Lasso_r2 for train_data:86.3%

Lasso_r2 for test_data:84.2%

Model4

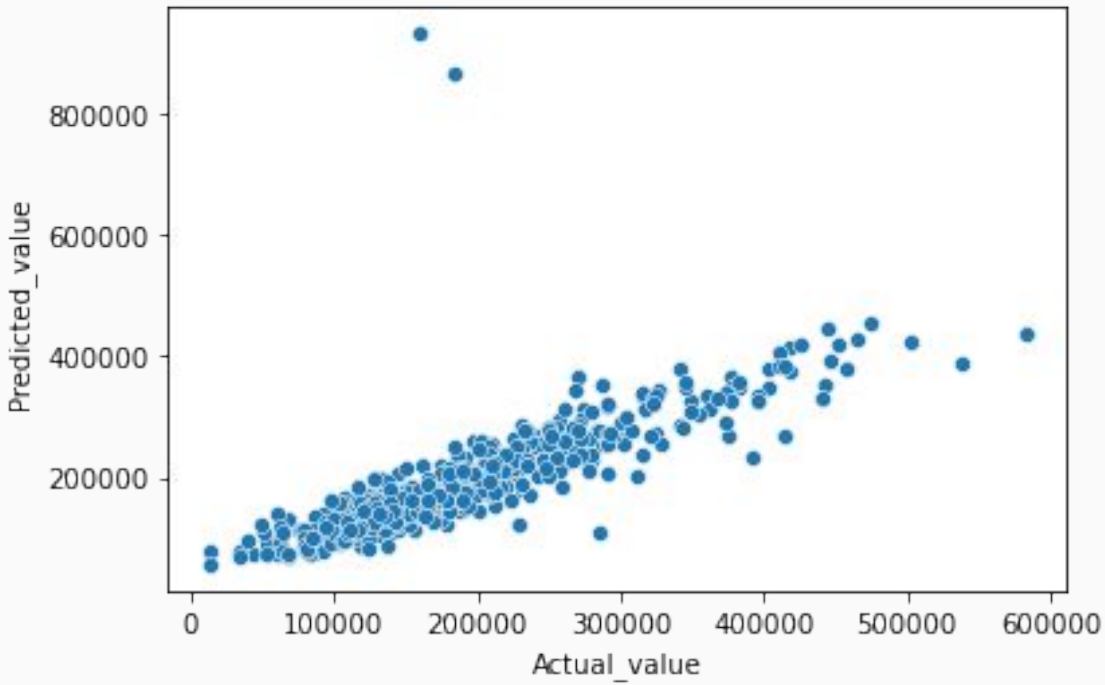
cross_val_score:85.49%

R2 for train_data:85.89%

R2 for test_data:85.83%

Rmse:31928

-Increase in overall quality,Sale price will increase by 15498.
-Availability of Central Air will increase the sale price by 1905 and unavailability will decrease by 1905.
-increase in Overall quality and availability of central air,Sale price will increase by 3802.



Recommendation:

Increasing Overall quality of the house by using high quality products.

Providing Central Air availability to houses will increase the sale price.

Older Age Houses got less Sale Price,Remodelling at proper time may lead to Price increase.

THANKS