Property Value
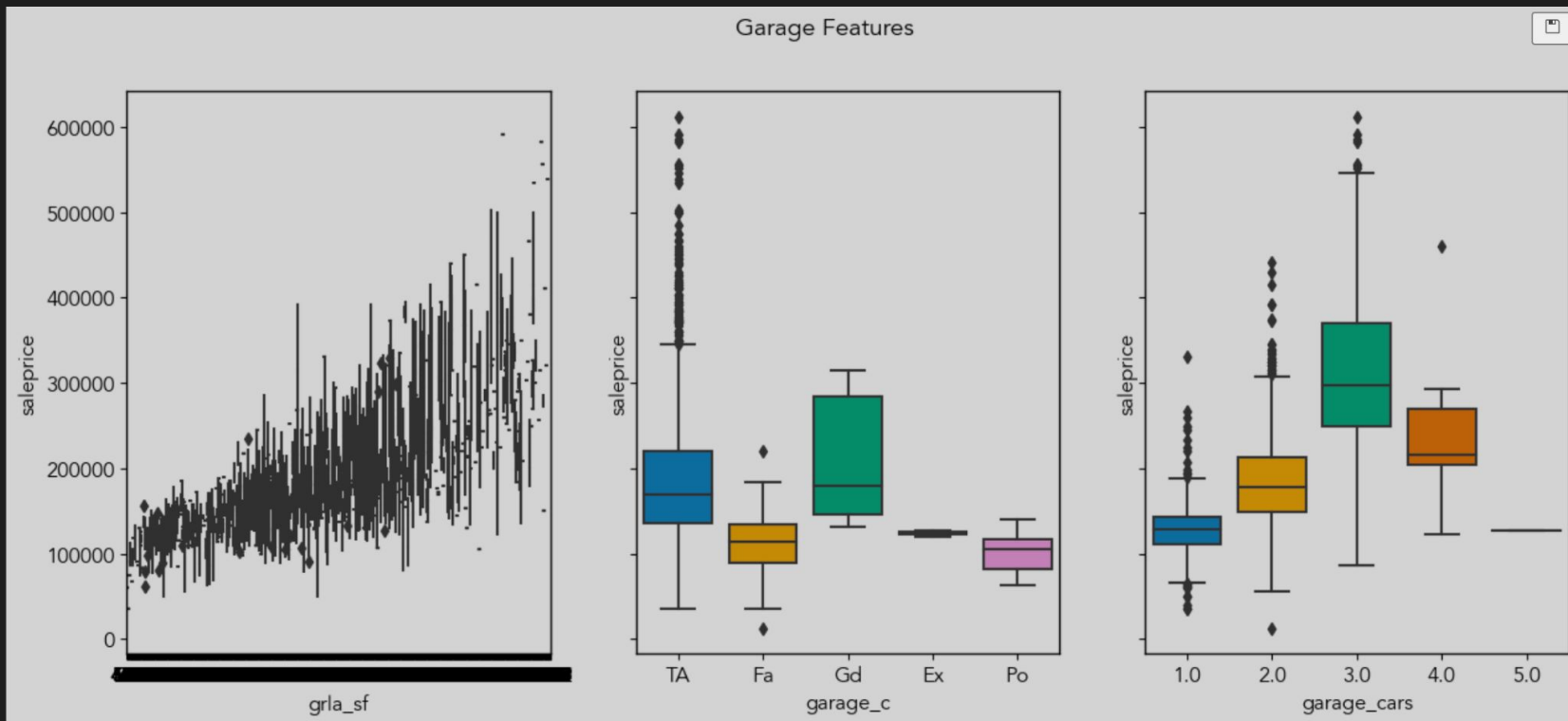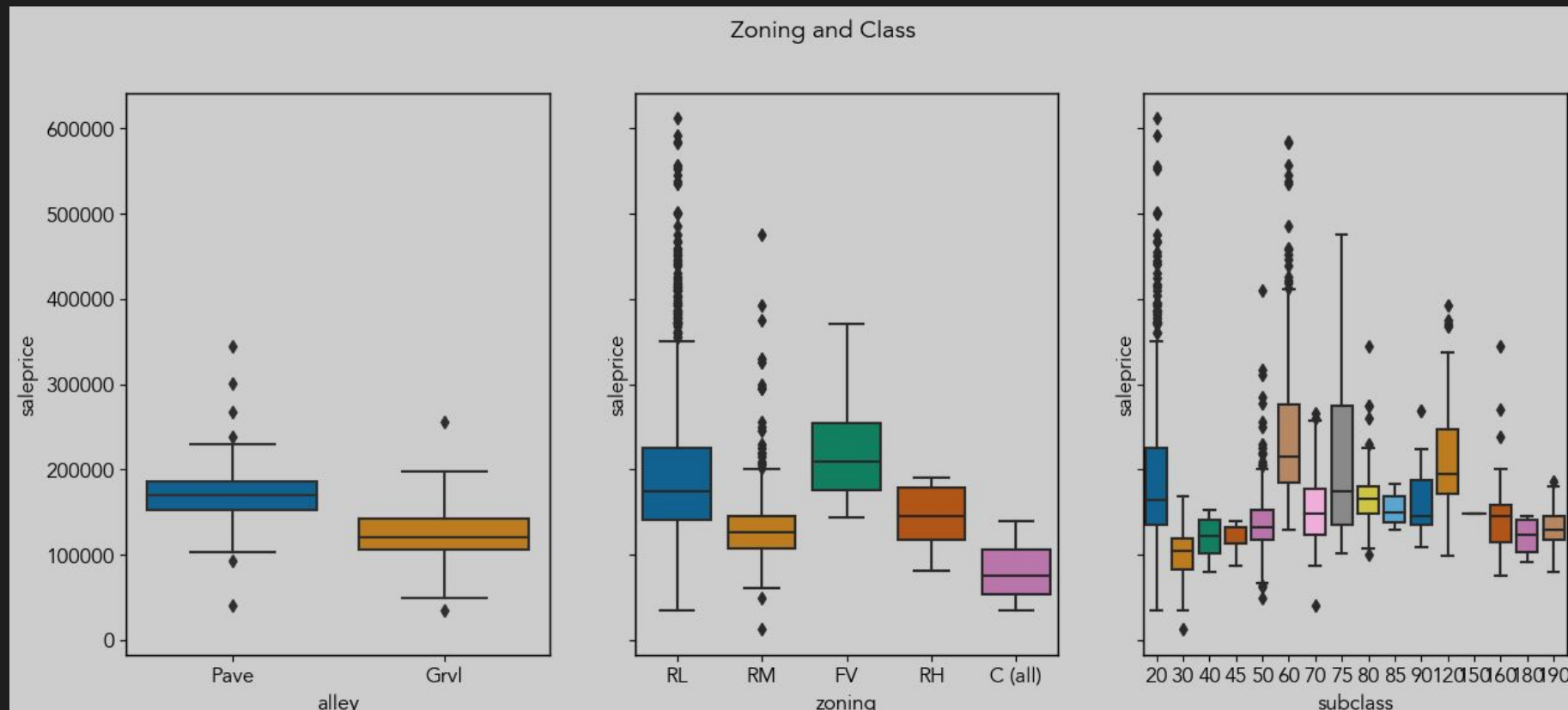
Which features predict property value?
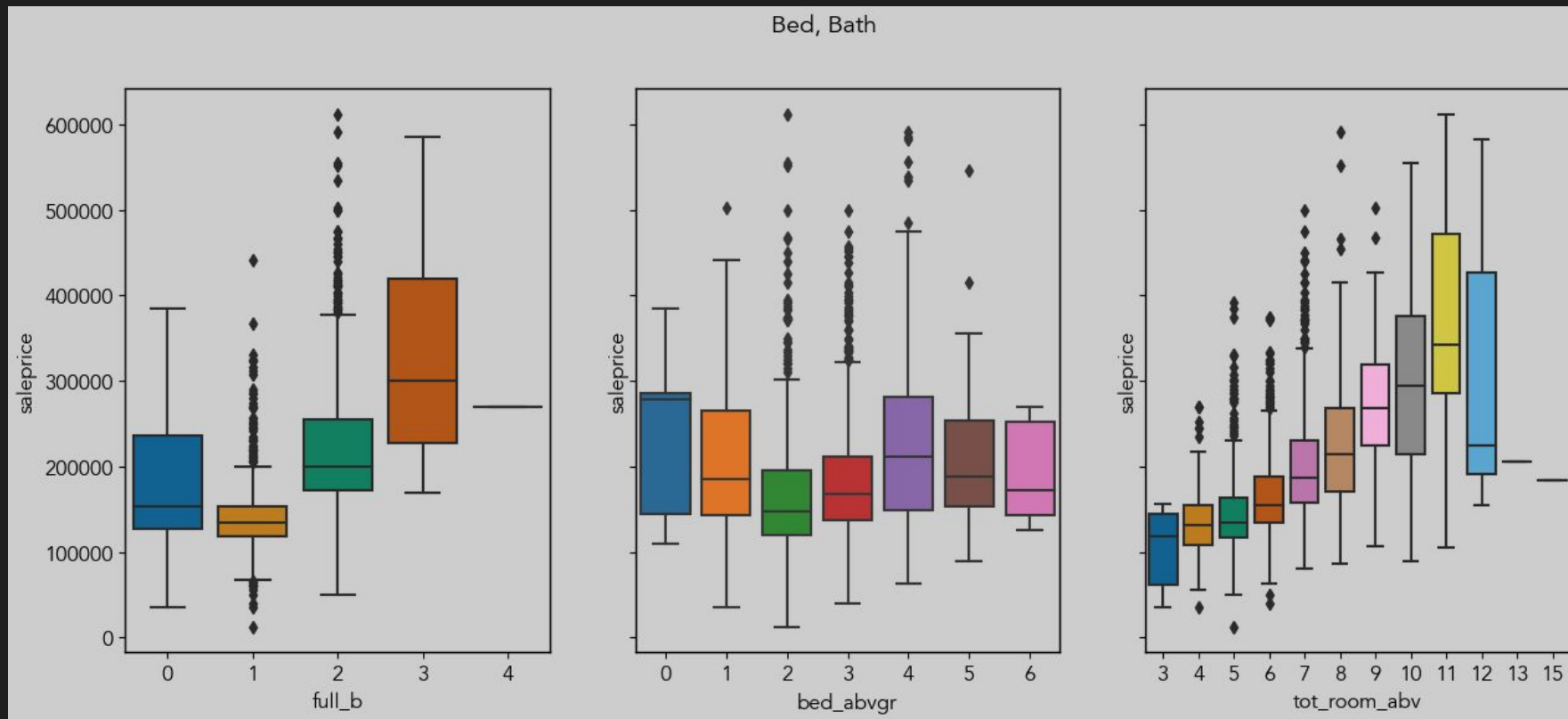
Garage Features: Not Linear

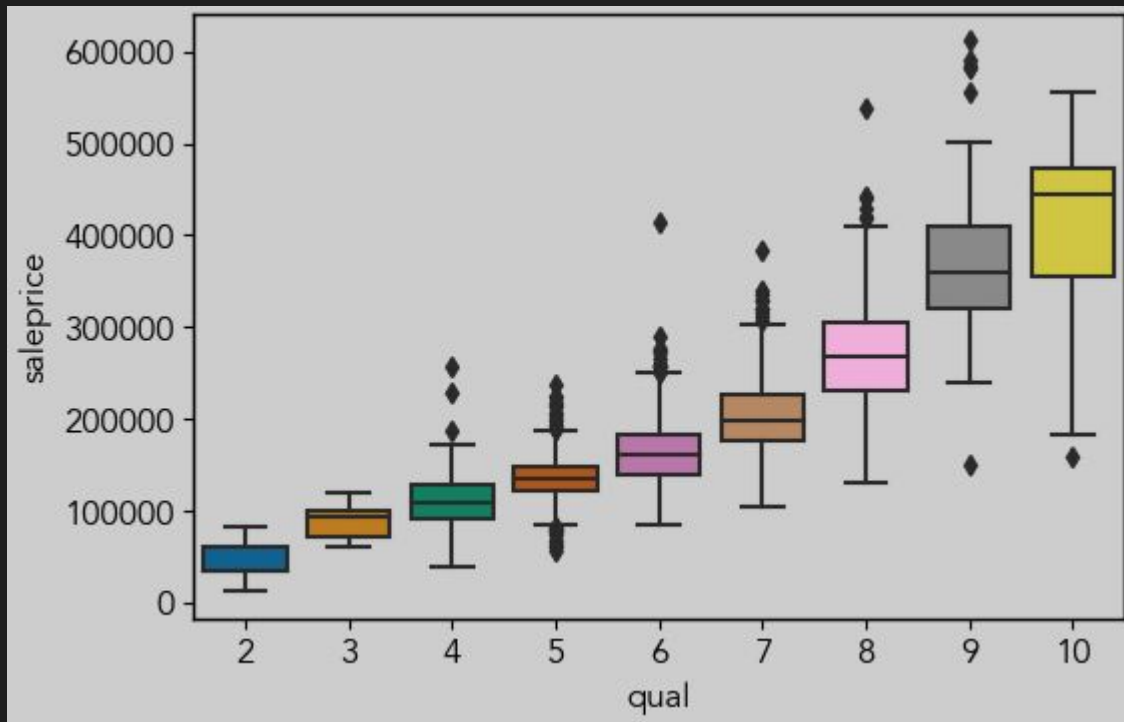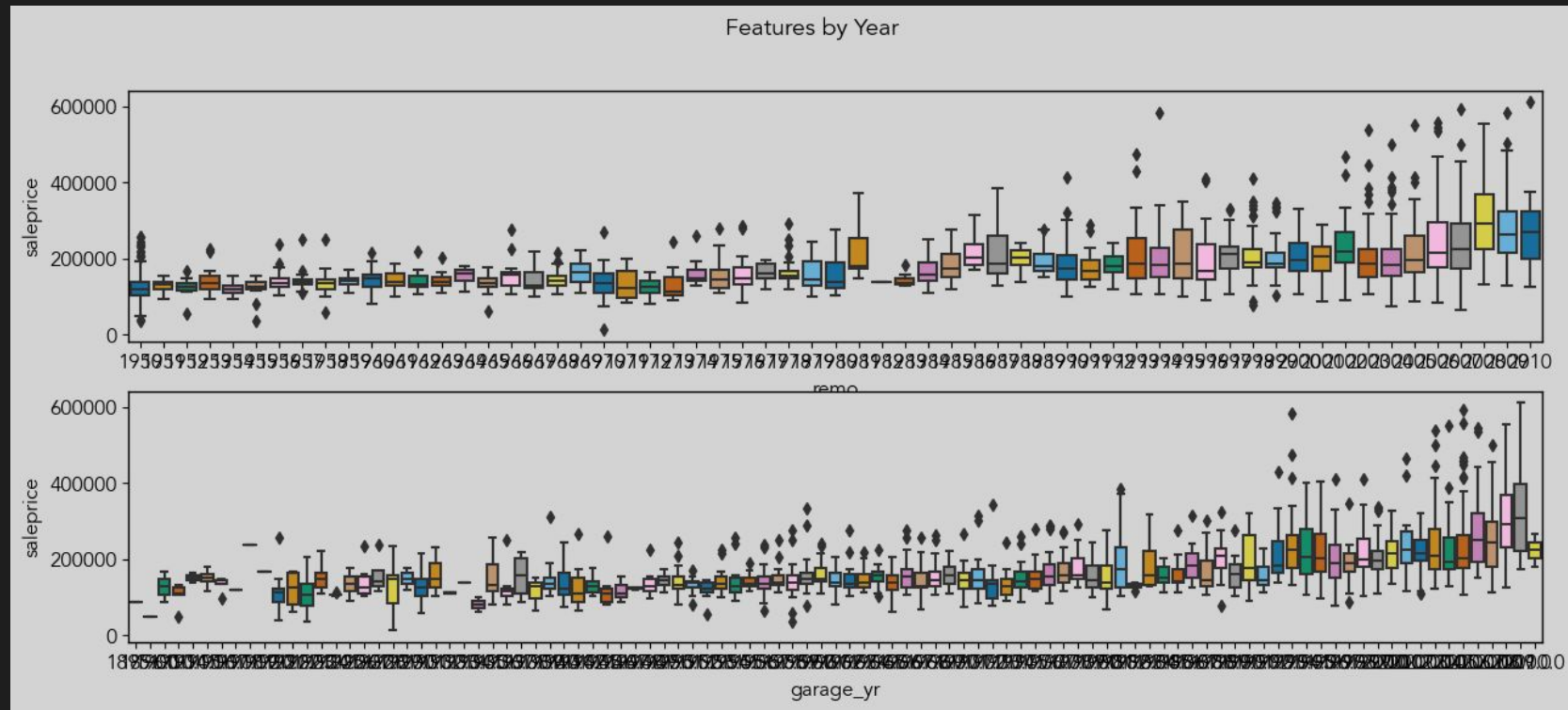# Categorical: Lot & CLass

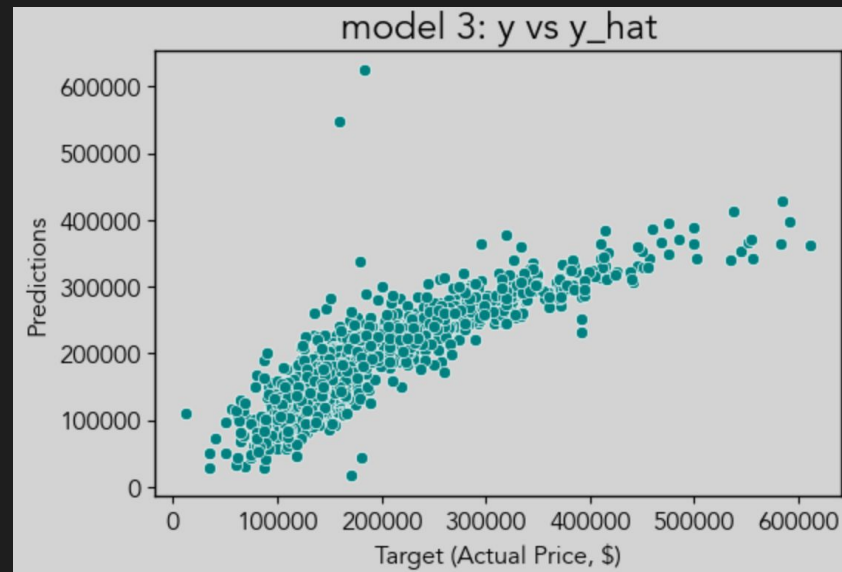Zoning and Class

Categorical or Numerical ?

Categorical: Quality

Deviation increases over time

Models 1, 2, 3: 20 Features
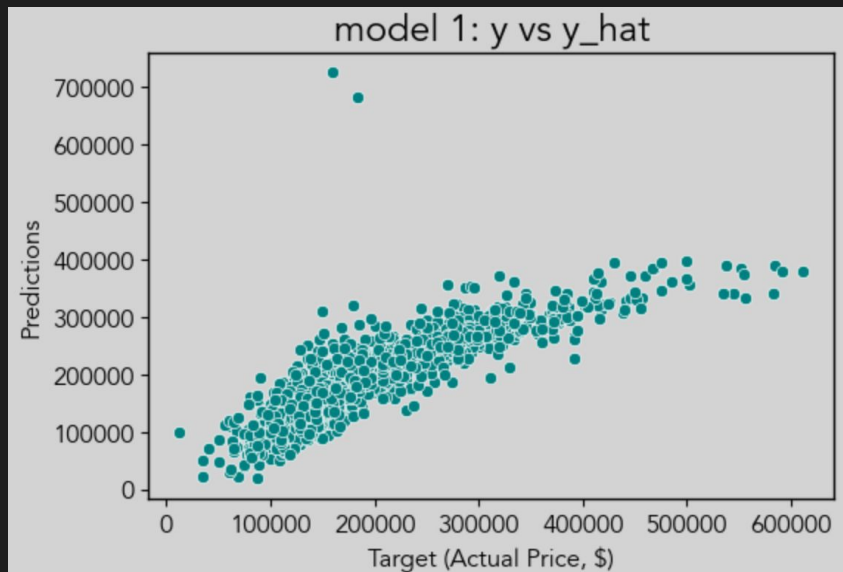
model 1: y vs y_hat

model 3: y vs y_hat

Model	Feature Engineering	R2	R2, X-val	RMSE	Slope	Intercept
1	All Numeric	0.7241	0.7109	41317	19.94344	-1364941
2	Standard Scalar	0.7241	0.7109	41317	19.94344	-1364941
3	Included `frontage` but dropped all null, 1556 observations	0.7398	0.7174	41875	-24.2397	-1123871

Models 4, 5: 20 Features
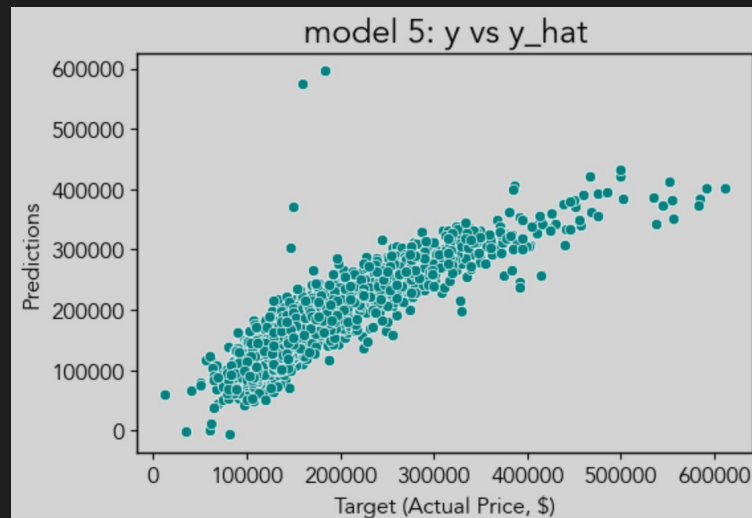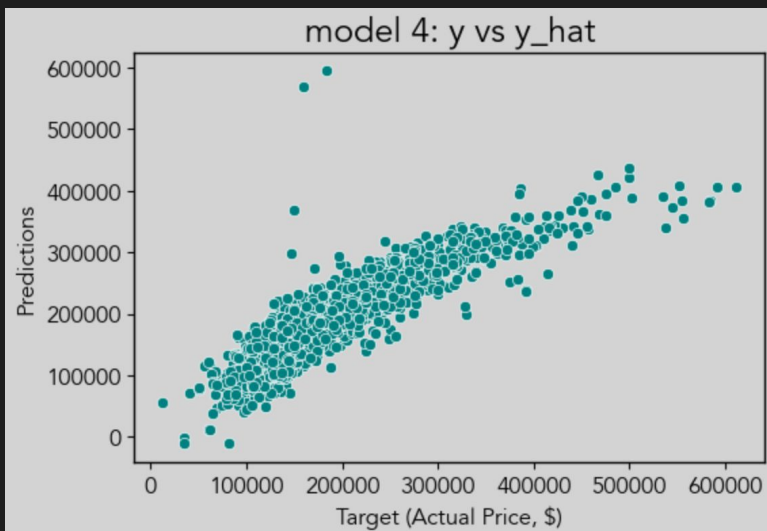
model 4: y vs y_hat

model 5: y vs y_hat

Model	Feature Engineering	R2	R2, X-val	RMSE	Slope	Intercept
1	All Numeric	0.7241	0.7109	41317	19.94344	-1364941
2	Standard Scalar	0.7241	0.7109	41317	19.94344	-1364941
3	Included `frontage` but dropped all null, 1556 observations	0.7398	0.7174	41875	-24.2397	-1123871
4	High Corr. features, and 4 categoricals, dummified	0.7776	0.7665	37100	4135	-100827
5	Only categorical feature is `zoning`	0.7748	0.7701	37333	3180	-109936

Ridge + Lasso: Benchmark
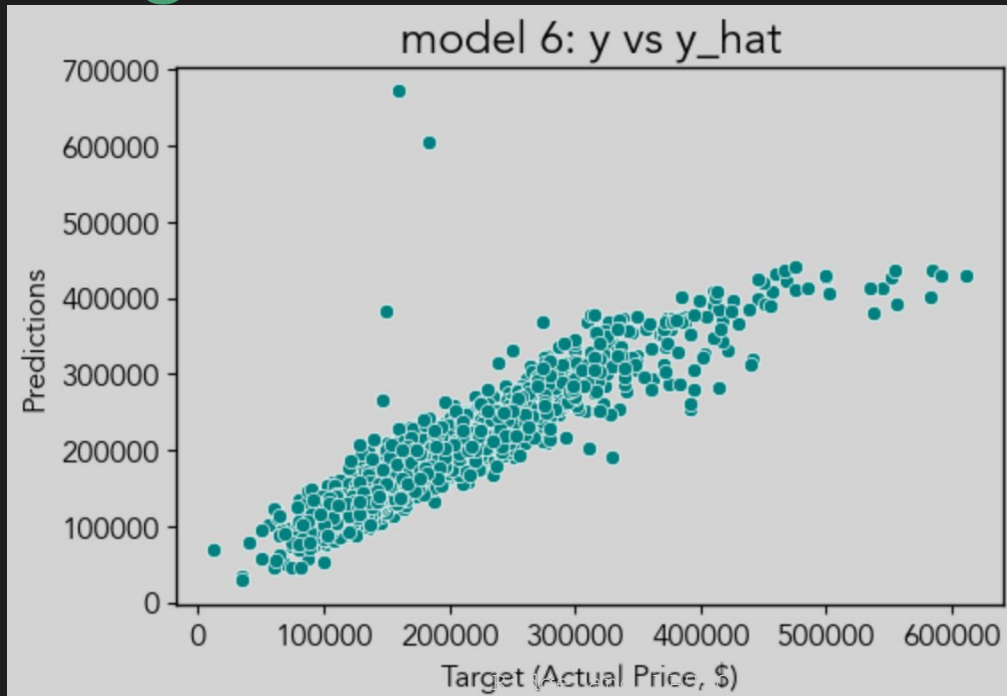
model 6: y vs y_hat

NUMERIC:

Basement Area
1st Floor Area
2nd Floor Area
Basement Full Bath
Full Bath Total
Half Bath Total
Total Room Above Ground
Fireplaces
Garage Year
Car Garage
Garage Area
Wood Deck Area
Open Porch Area
Enclosed Porch
Screened Porch
Pool Area
Month of Sale

r2	0.7244	Ridge Train	0.737051
rmse	41301	Ridge Test	0.691658
intercept	-0.000001	LASSO Train	0.740855
slope	1.89	LASSO Test	0.689215

Ridge + Lasso: Model 2
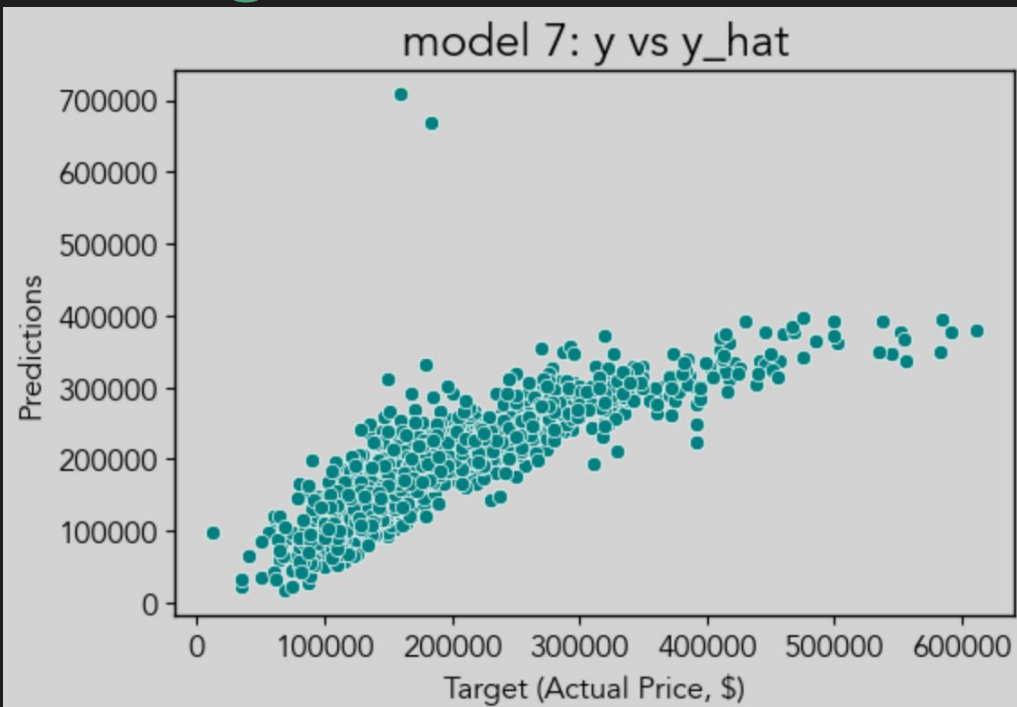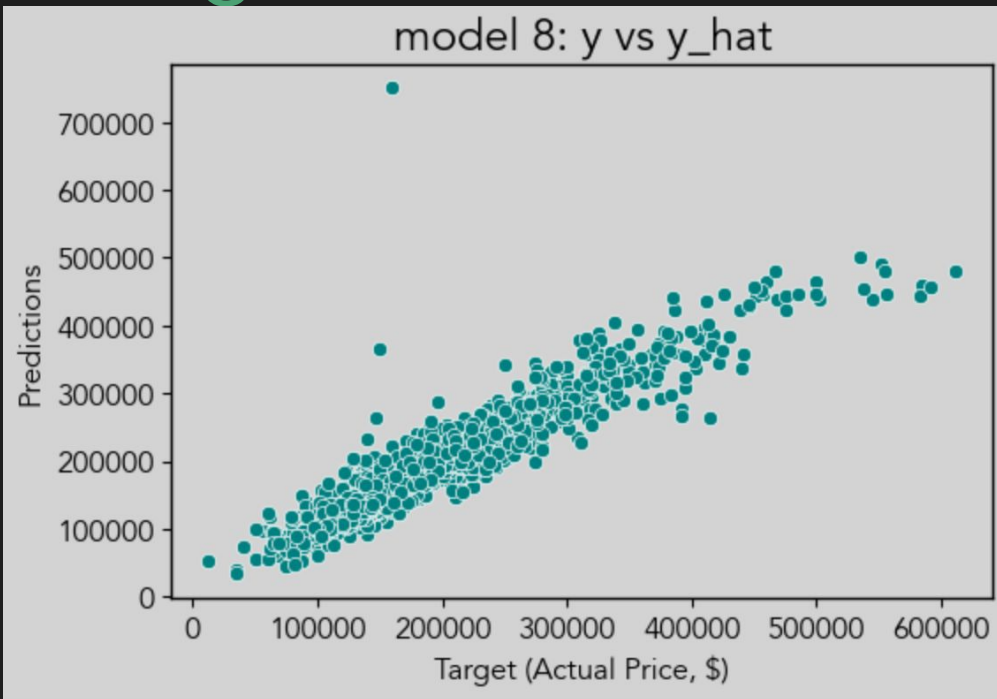

model 7: y vs y_hat

NUMERIC

Basement Area
1st Floor Area
2nd Floor Area
Basement Full Bath
Full Bath Total
Half Bath Total
Total Room Above Ground
Fireplaces
Garage Year
Car Garage
Garage Area
Wood Deck Area
Open Porch Area
Enclosed Porch
Screened Porch
Pool Area
Month of Sale

CATEGORICAL

Month of Sale
Fireplace
Zoning

r2	0.7307	Ridge Train	0.742372
rmse	40 825	Ridge Test	0.699286
intercept	-0.000001	LASSO Train	0.747558
slope	1.909096e+01	LASSO Test	0.694796

FEATURES:20

Ridge + Lasso: Model 3

model 8: y vs y_hat

NUMERIC	CATEGORICAL
Basement Area	Month of Sale
1st Floor Area	Fireplace
2nd Floor Area	Zoning
Basement Full Bath	Subclass
Full Bath Total	Quality
Half Bath Total	Total Room Above
Total Room Above	Neighborhood
Fireplaces	
Garage Year	
# Car Garage	
Garage Area	
Wood Deck Area	
Open Porch Area	
Enclosed Porch	
Screened Porch	
Pool Area	
Month of Sale	

```
      r2  0.8758          Ridge Train  0.899396
    rmse  27 718           Ridge Test  0.827300
intercept -5000           LASSO Train  0.907058
   slope  2.409            LASSO Test  0.810956
```

FEATURES:25

Ridge + Lasso: Model 4

model 8: y vs y_hat

NUMERIC

Basement Area
1st Floor Area
2nd Floor Area
Basement Full Bath
Full Bath Total
Half Bath Total
Total Room Above
Fireplaces
Garage Year
Car Garage
Garage Area
Wood Deck Area
Open Porch Area
Enclosed Porch
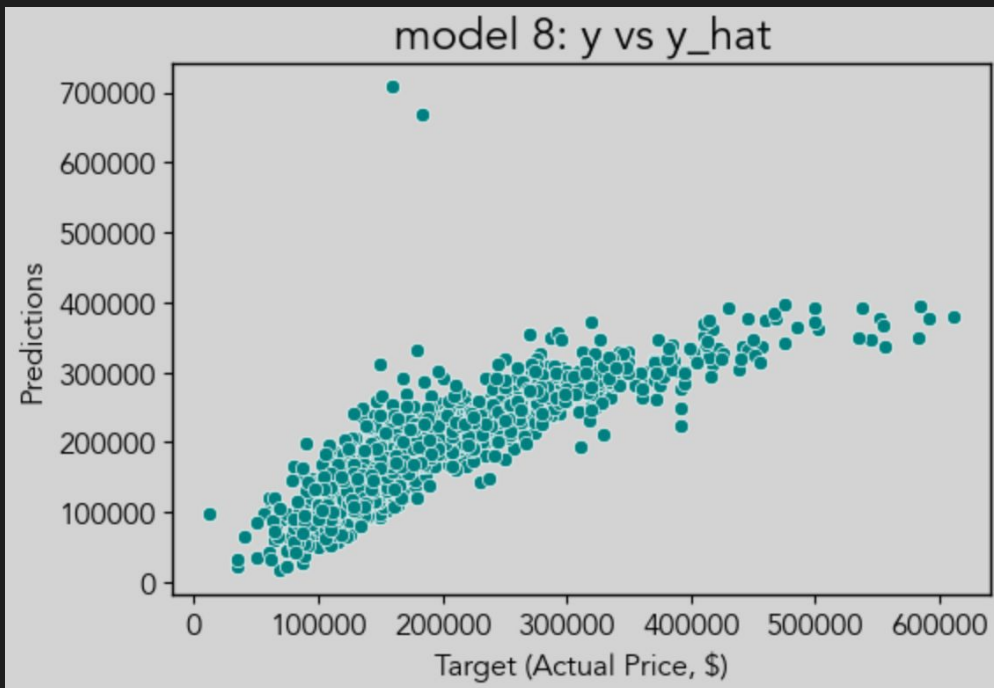Screened Porch
Pool Area
Month of Sale

CATEGORICAL

Month of Sale
Fireplace
Zoning
Subclass
Quality
Total Room Above
Neighborhood
Full Bath Total

FEATURES:25

Conclusion

- HIGH-QUALITY CATEGORICAL IS SIGNIFICANT