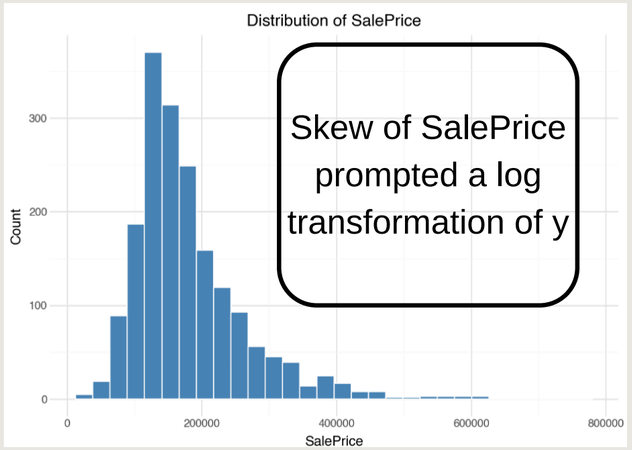
Predictive Analysis Regression - GSB 544

Step 1 - Data Cleaning

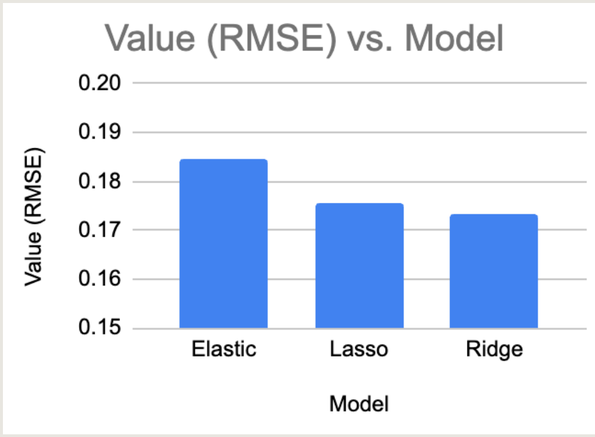
- Check for missing values
- Check for outliers
- Dummify variables

Step 2 - Log Transformation



Skew of SalePrice prompted a log transformation of y

Step 3 - Model Evaluation



The Ridge model had the lowest CV RMSE of 0.173

Step 4 - Parameter Tuning and Cross Validation

```
alphas = {'ridge_regression__alpha': np.array
([0.001, 0.01, 0.1, 1, 10])}
```

```
gscv_ridge = GridSearchCV(ridge_pipeline,
param_grid=alphas, cv = 5,
scoring='neg_mean_squared_error')
```

Step 5 - Conclusion and Results

Model Selection Based On:

- Different models
- Different features
- Different parameters

Accuracy on Competition: 0.15988

Step 6 - Interpretation

Benefits of building a house price predictive model:

- Trend Recognition
- Market Fluctuation Analysis
- Comparisons To Other Markets
- Understand Economic Climate