# Predicting Sale Prices for Homes in Ames, IA

## Multiple Linear Regression Model
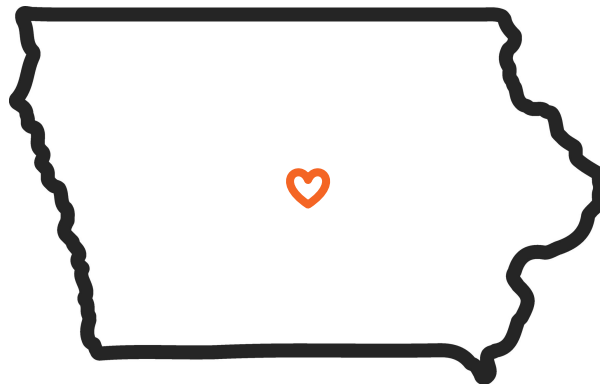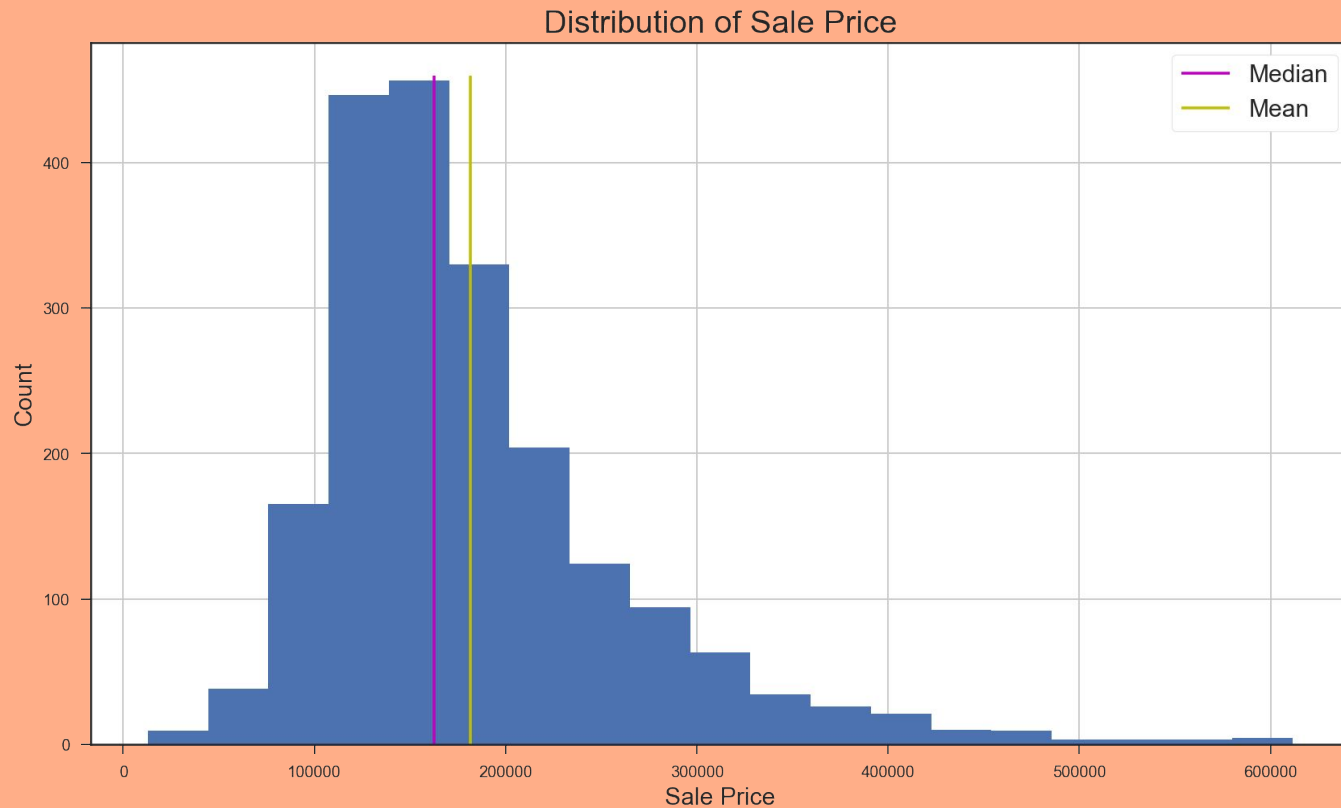
By: Tofer Kim

# Data Description

- Observations from 2,051 houses sold in Ames, IA between 2006 and 2010
- 82 descriptive features including, PID, square footage, neighborhood, various quality/condition scores, etc.
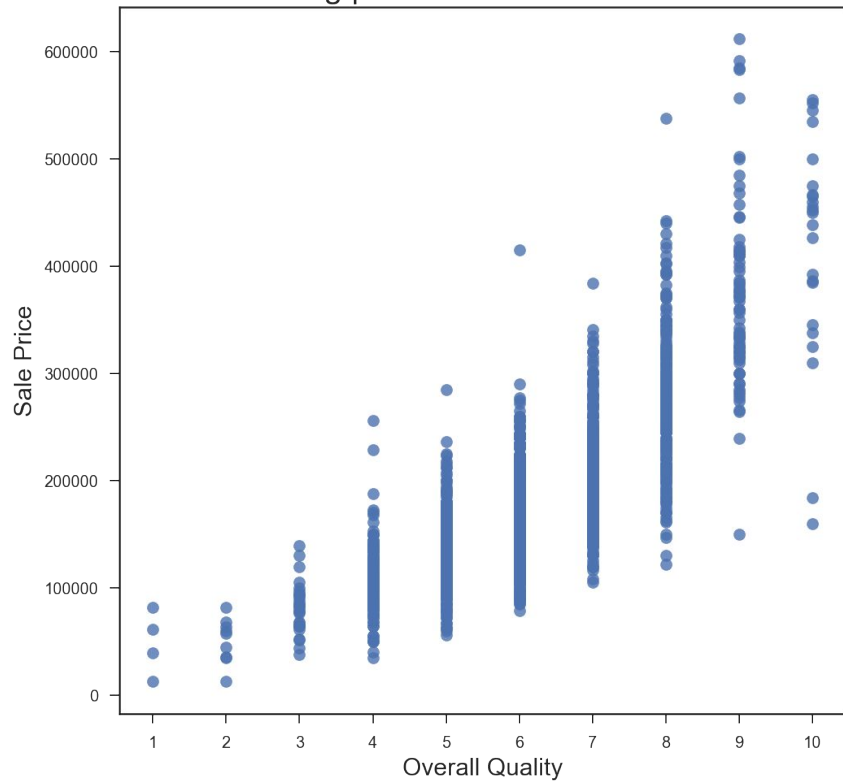
**Median:**
**$162,500.00**

**Mean:**
**$181,469.70**

Distribution of Sale Price

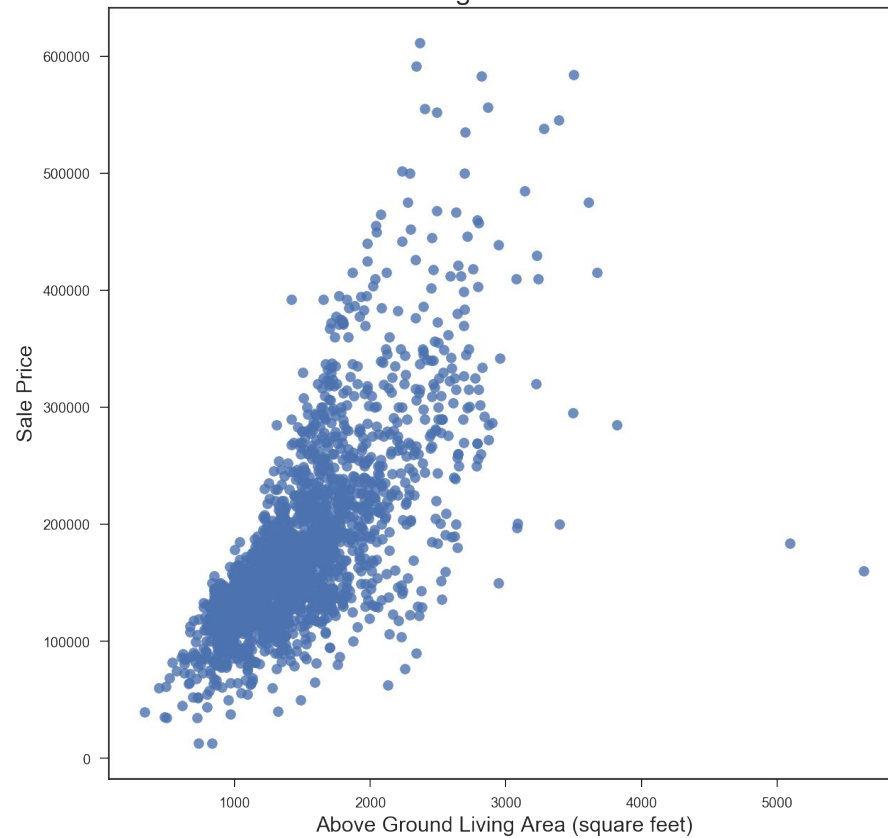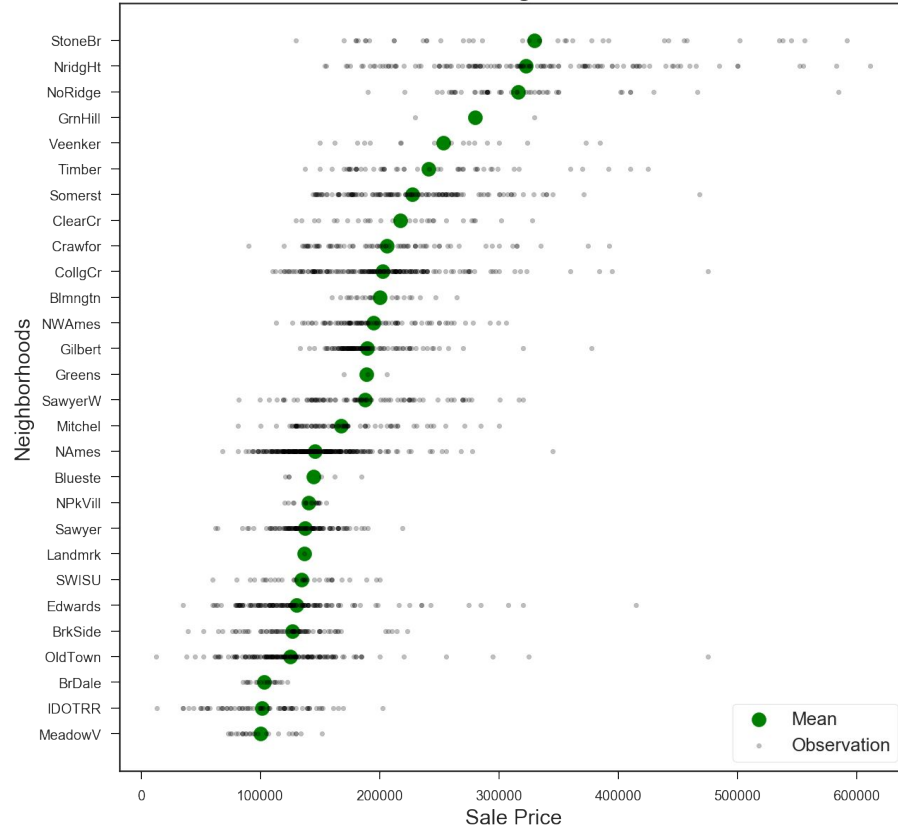The Overall Quality of a home is a strong predictor of its Sale Price

0.80

Pearson Correlation

There are 2 houses that are
much larger than the rest

Distribution of Sale Prices vary greatly between Neighborhoods

# Ridge

Train Score:

94.9%

Test Score:

91.8%

CV Score:
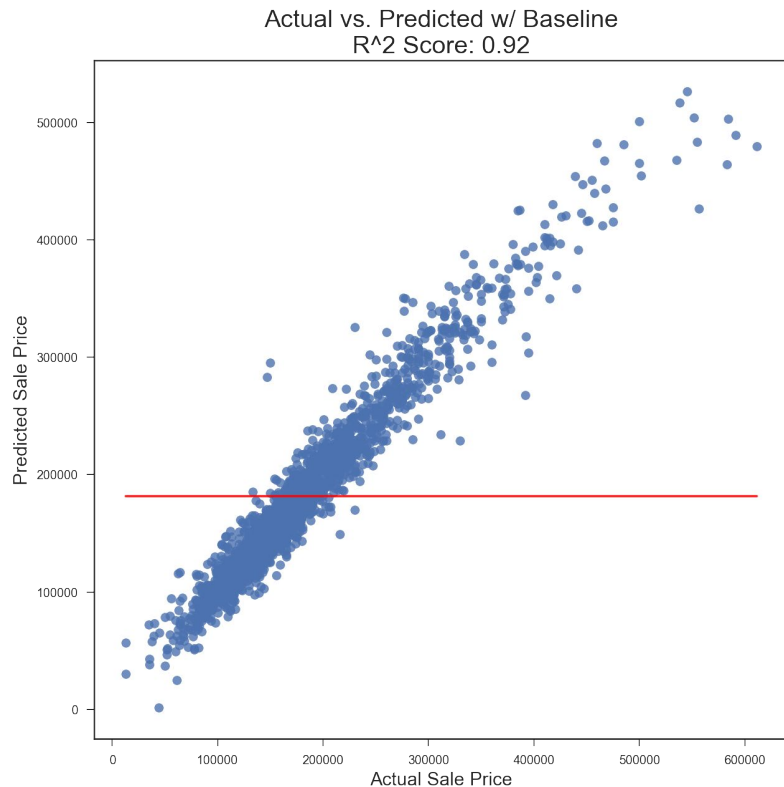
91.3%

# Lasso

Train Score:

93.5%

Test Score:

91.1%

CV Score:

91.4%

# Conclusion

**Actual vs. Predicted w/ Baseline**
**R^2 Score: 0.92**



- Final model accounts for 92.1% of variability within our target variable, Sale Price

- Lasso method considers 201 predictors:

|  | coefs |
|---|---|
| kitchen_qual gr_liv_area | 27084.565073 |
| gr_liv_area | -20079.108042 |
| exter_qual gr_liv_area | 19343.502479 |
| kitchen_qual | -15095.018678 |
| exter_qual | -15001.064306 |
| exter_qual neighborhood | 12587.415458 |
| bsmt_qual gr_liv_area | 11271.068923 |
| bsmt_qual | -11202.491960 |

# Limitations

- Coefficients are difficult to interpret.

- This model should not be used to predict prices for homes sold outside of Ames, IA , before 2006 or after 2010.

- Better optimization with different polynomial features and hyperparameter combinations.

- Explore classification methods to impute values and categorical features.

- Did the listed price vary from sale price in any meaningful way? Were there any homes that went up for sale but never sold?

# Thank you!