Market Price Prediction Model

Ames, Iowa

Problem Statement

Create a linear regression model that can predict the market value of real estate in Ames, Iowa as accurately as possible.

Description of Data

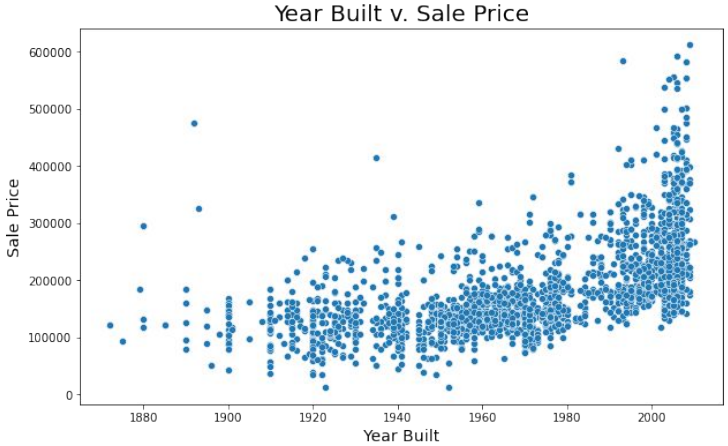
- Real estate sold during 2007-2010 in Ames, Iowa
- 80 columns of property characteristics
- Target column: SalePrice

Feature Selection

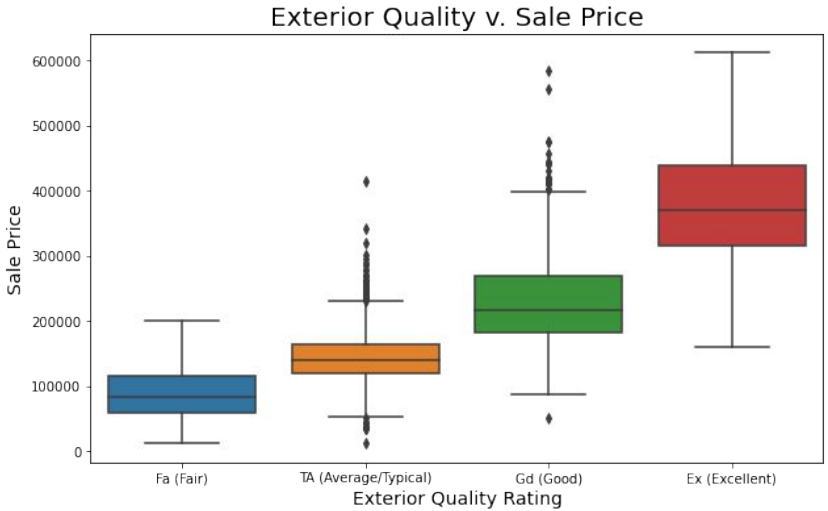
- Preliminary Plots
  - Scatterplots for Numerical Features
  - Boxplots for Categorical Features

- 7 features selected

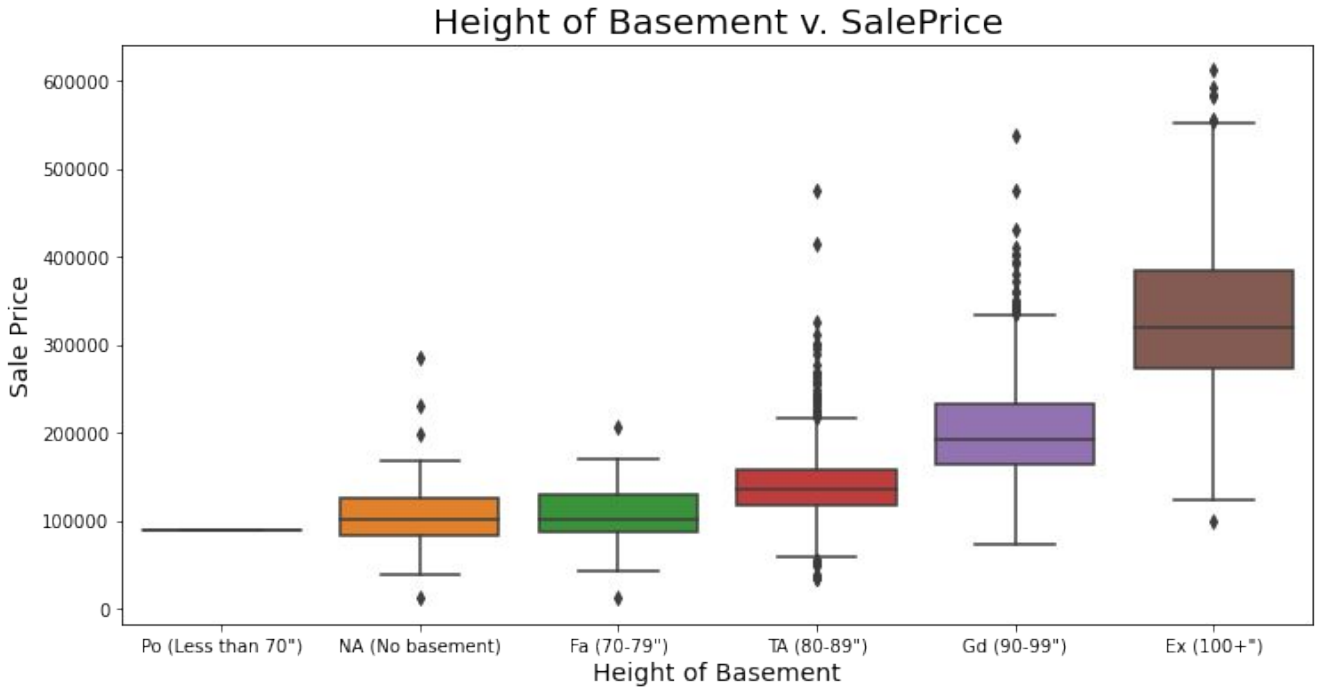
Numerical Features



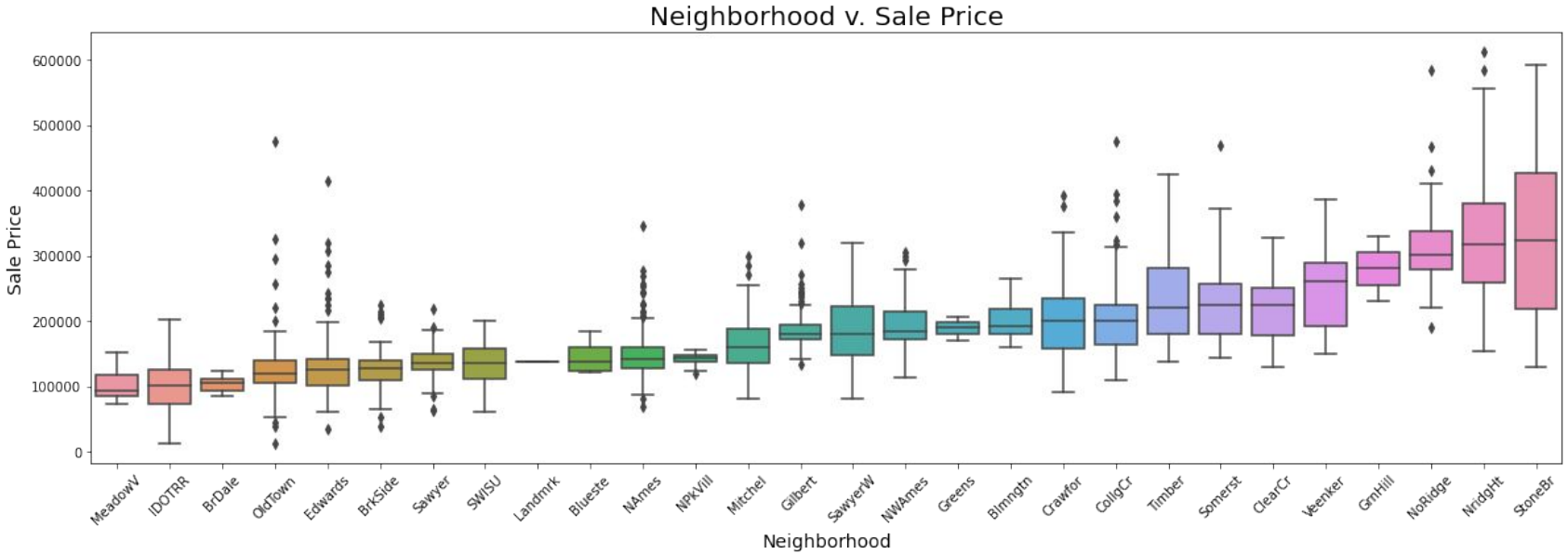
Categorial Features



Categorical Features



Categorical Features



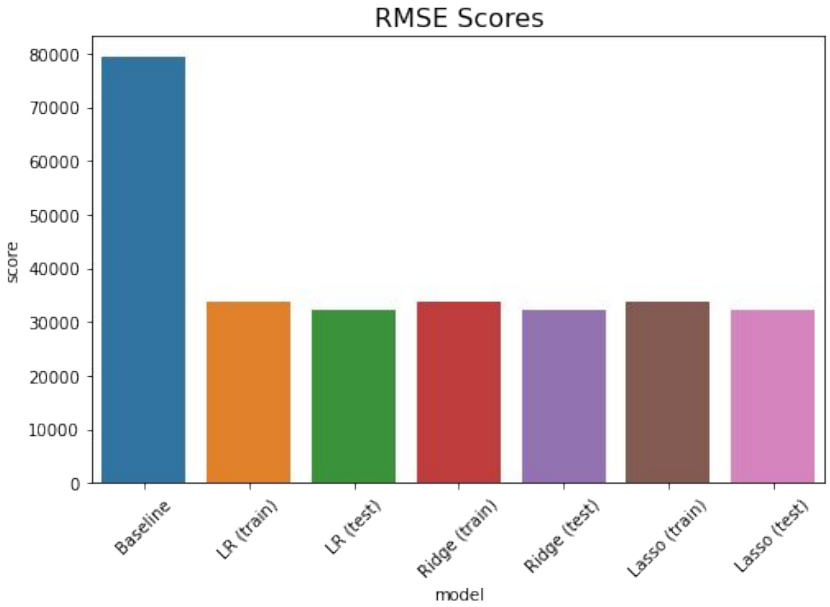
Preprocessing

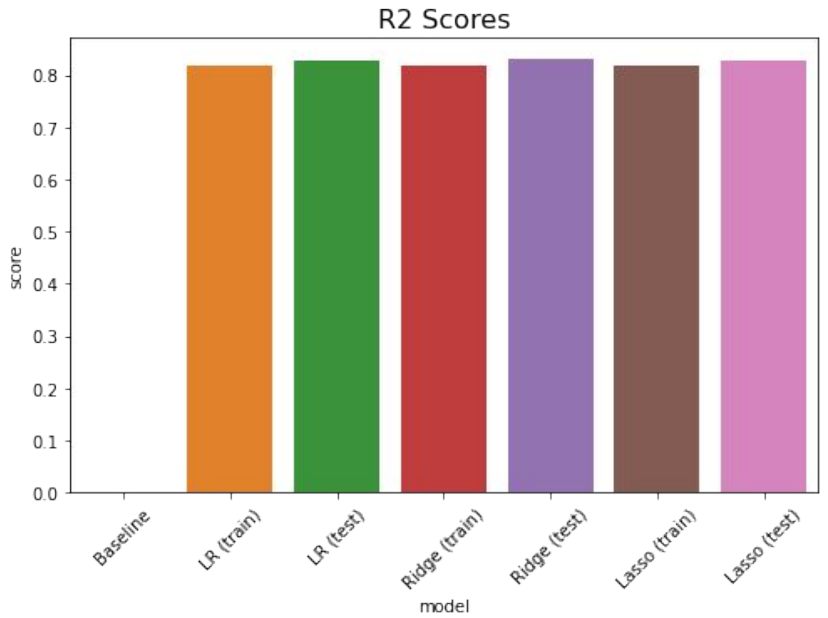
- Data cleaning/missing values resolution
- Replaced Poor - Excellent ratings with numerical (0-5)
- Dummy variables for Neighborhood
- Train-test Split (75%/25%)
- Scaling

Modeling

- 3 models built
  - Linear Regression
  - Ridge
  - Lasso
- Compared against baseline null model (mean of sale price)

Performance of Models





Conclusion/Recommendations

---

1. Created good starter model
2. Next steps:
    a. Address outliers
    b. Add more features
    c. Feature engineering