Can Predicting House Prices in Ames, IA Help First Time Home Buyers?

Curtis M Hope Hill

Background

- **Millennials and Housing.**
- Near constant discussion over the last decade + about housing and Millenials
- Home Prices have continued to grow at a staggering rate
- Wages and Salaries for Millenials have remained largely stagnant

Problem Statement

- Millennials have a low home ownership rate compared to previous generations.
- This impacts Millennials finances as renting an apartment or home can be more expensive than a mortgage.
- By understanding and accurately predicting home prices in Ames, IA we may be able to identify important features that impact the price of a home.
- Examining these features may illuminate high value features that aren't necessary for first-time buyers.

Data Science Workflow

- Initial EDA and Data cleaning
- Modeling with Linear Regression
  - Train, Test, Split
  - Regularization
- Three Model Comparison:
  - Initial Naive Model
  - Common Features Model
  - Complex Features Model
- Conclusions and Interpretations

Naive Model

- Focus on:
  - Bedrooms, Bathrooms
  - Overall and Exterior Qualities and Condition
  - Utilities
  - Garage
  - Total Square Footage
- Accuracy Scores:
  - Train .832
  - Test : .805
- RMSE or Dollars off of Sale Price
  - Train: 33,122.28
  - Test: 32729.8

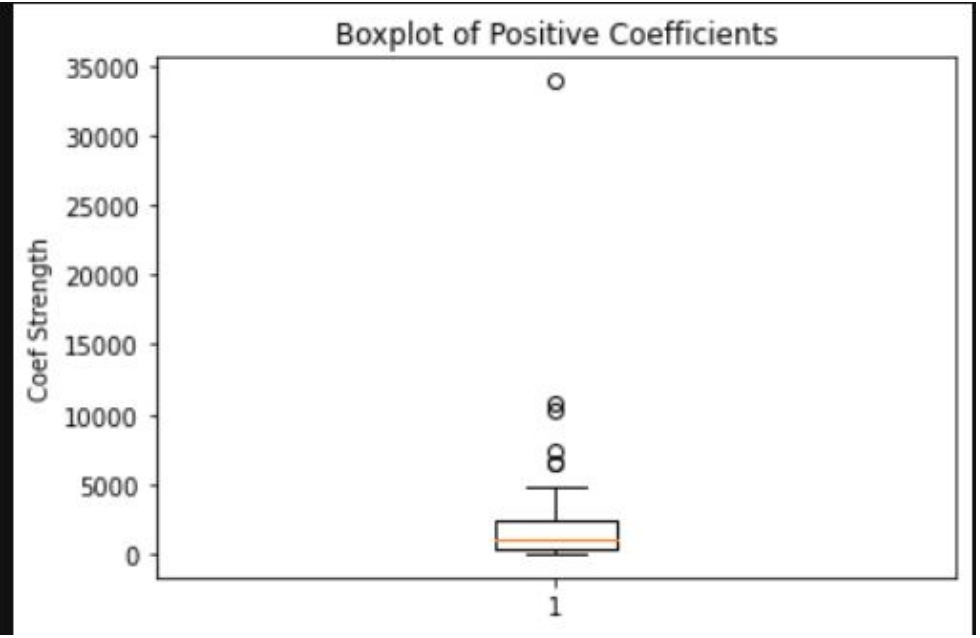
Common Features Model

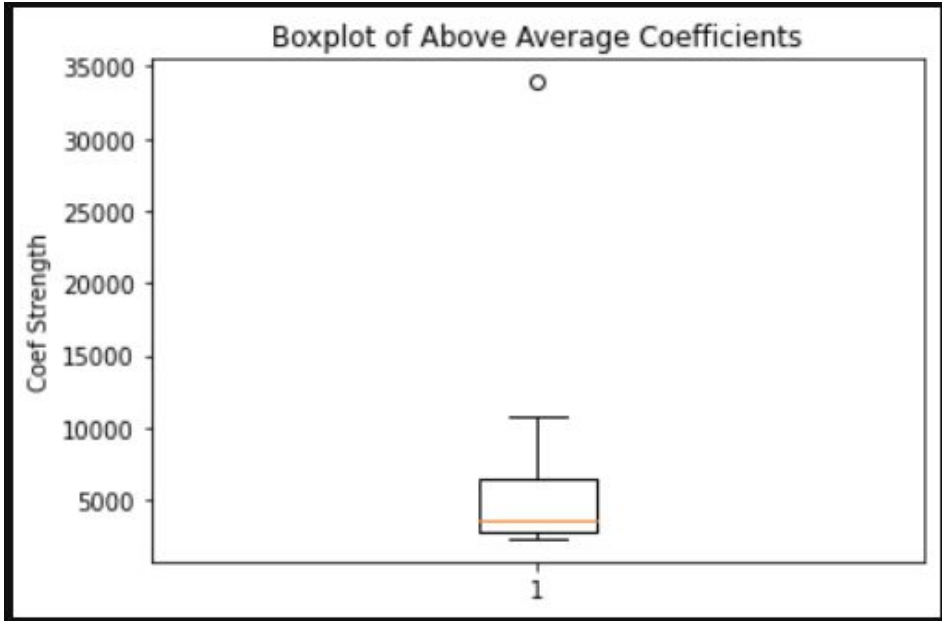
- Based on Features in Common Real Estate Listing (Zillow and MLS Real Estate)
  - Bedrooms
  - Bathrooms
  - Square Feet
  - Lot Size
  - Property Type
- Accuracy Scores
  - Train : .777
  - Test : .734
- RMSE or Dollars off of Sale Price
  - Train: 37948.81
  - Train : 39166.69
- Worse when compared to the Naive Model

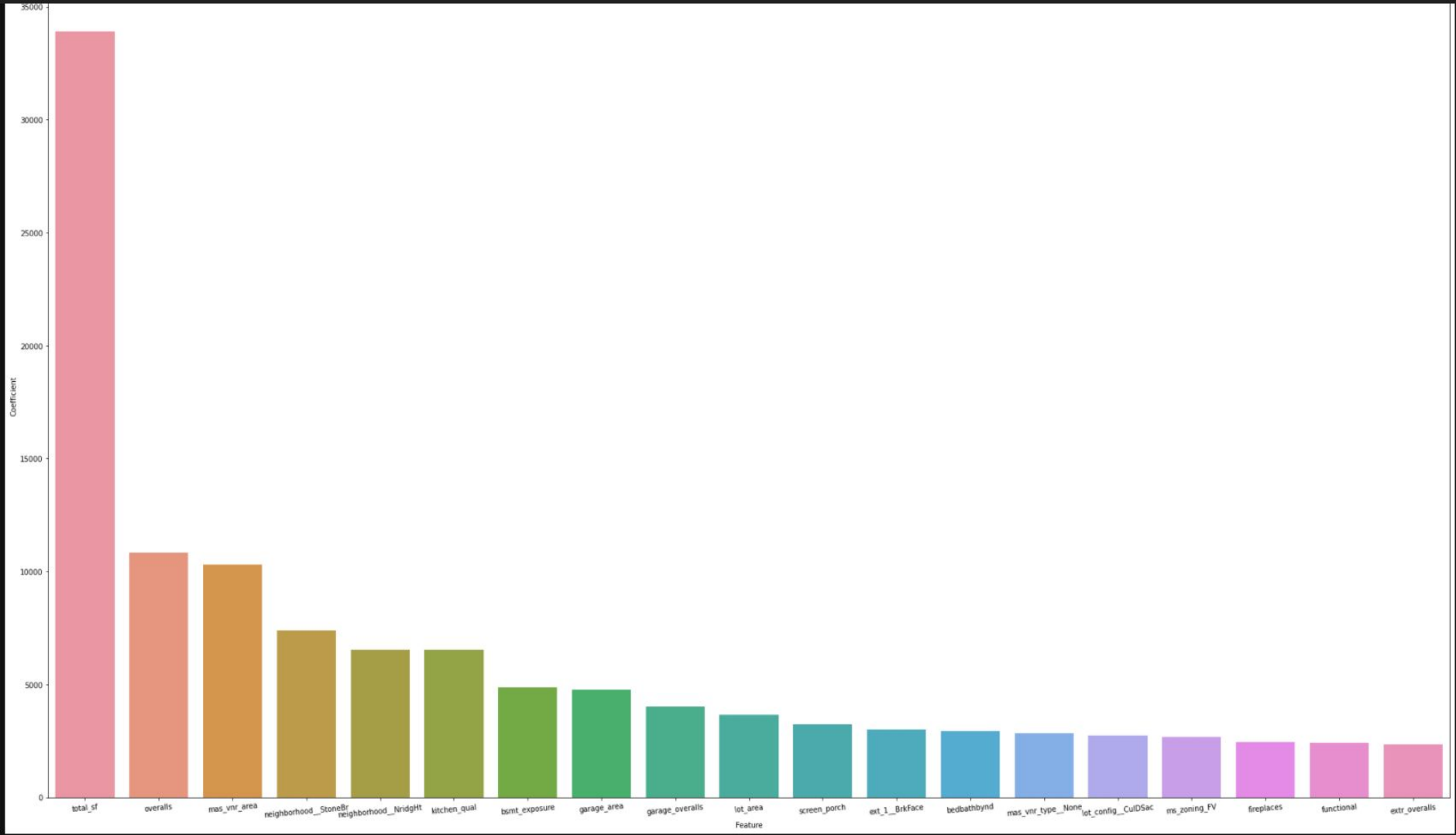
Complex Features Model

- Kitchen Sink Method
- Used Ridge and Lasso to Regularize the Data
  - Lasso to Find important Coefficients
- Accuracy
  - Train: 0.9233
  - Test 0.893
- RMSE
  - Train: 22649.01
  - Test: 24240.98

Positive Coefficients






35000
30000
25000
20000
15000
10000
5000
0
Coefficient
total_sf
overalls
mas_vnr_area
neighborhood__StoneBr
neighborhood__NridgHt
kitchen_qual
bsmt_exposure
garage_area
garage_overalls
lot_area
screen_porch
ext_1__BrkFace
bedbathbynd
mas_vnr_type__None
lot_config__CulDSac
ms_zoning_FV
fireplaces
functional
extr_overalls
Feature

Conclusions from top Coefficients

We can group these features by:

1. Total Square Footage of the home
1. Varying Neighborhood Locations for the home
1. Quality of important parts of the home
    * Kitchen most important
    * Basement and the garage are next most important
    * Having a Brick Exterior
1. Size of the lot
1. Number of rooms in the home
1. Suburban location (Cul De Sac location or A Floating Village Residential Area)

Two Important Features were not in the previous models:

1. Screen Porch
2. Fireplace

Questions?