Housing Sale-Price Predictions

Using Our Machine Learning Model to target housing sale prices for your real estate company.

The Problem:

Initial: What features and conditions of a house on the market most reliably predict what the house will sell for?

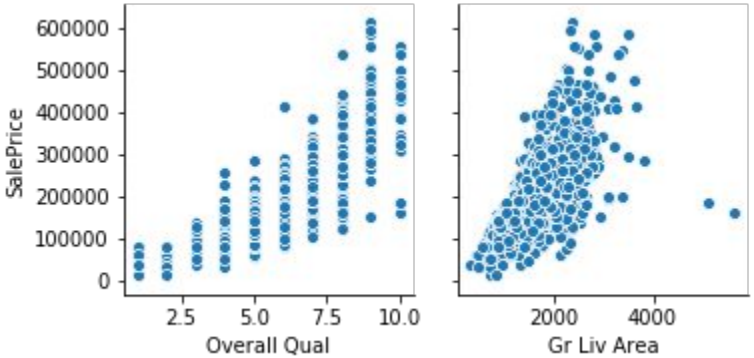
Using data from the Ames, Iowa housing market from 2006 to 2010, we analyzed the metrics across each sale.

Although there are many factors that may go into deciding the sale-price of a given unit, we wanted to build a model able to *target the most relevant factors*.

So: What feature set is most important in creating a model that predicts housing sale-prices with minimal errors in it's predictions?

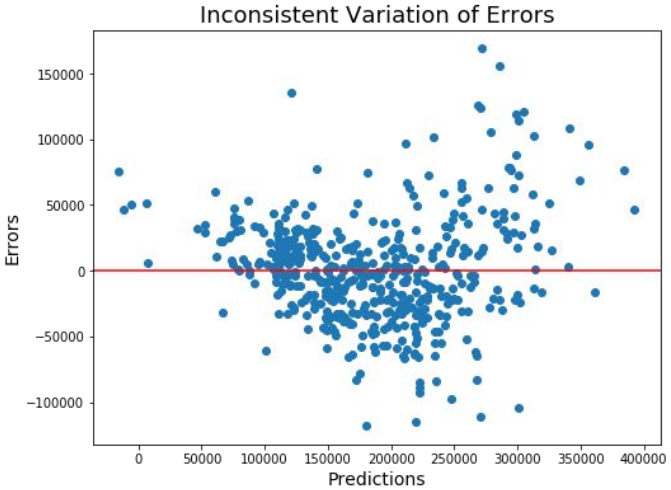
What we do different: Building on Basic

Basic models may use only readily available quantitative metrics to determine housing prices



Common metrics to predict sale price are often size and quality of the unit. Although these do have (intuitively) a linear relationship to final sale-price, they don't paint a complete picture.

The goal is to find the conditions that best predict sale-prices - we must look beyond surface-level quantitative analysis to build models that minimize errors in predicted sale-prices.



This Basic Model can only account for 72% of variation of sale prices

Our Method: Minimizing Error in Predictions

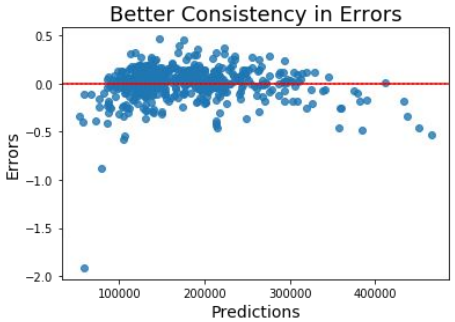
Feature Engineering:

We read in housing data from a given market, clean it, and return complex features and variables to create a model of best prediction.

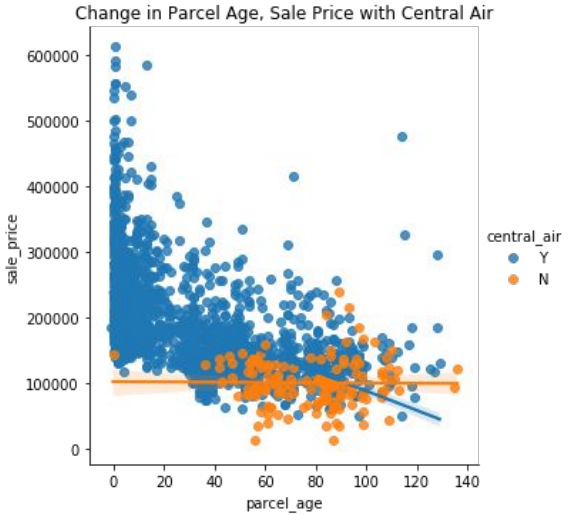
This helps us minimize the error of our predictions to the actual Sale-Prices

Modeling:

Running targets through transformations to better account for inconsistencies in prediction variances.



Our model finds finding the relationships among disparate data points to find deeper meaning.

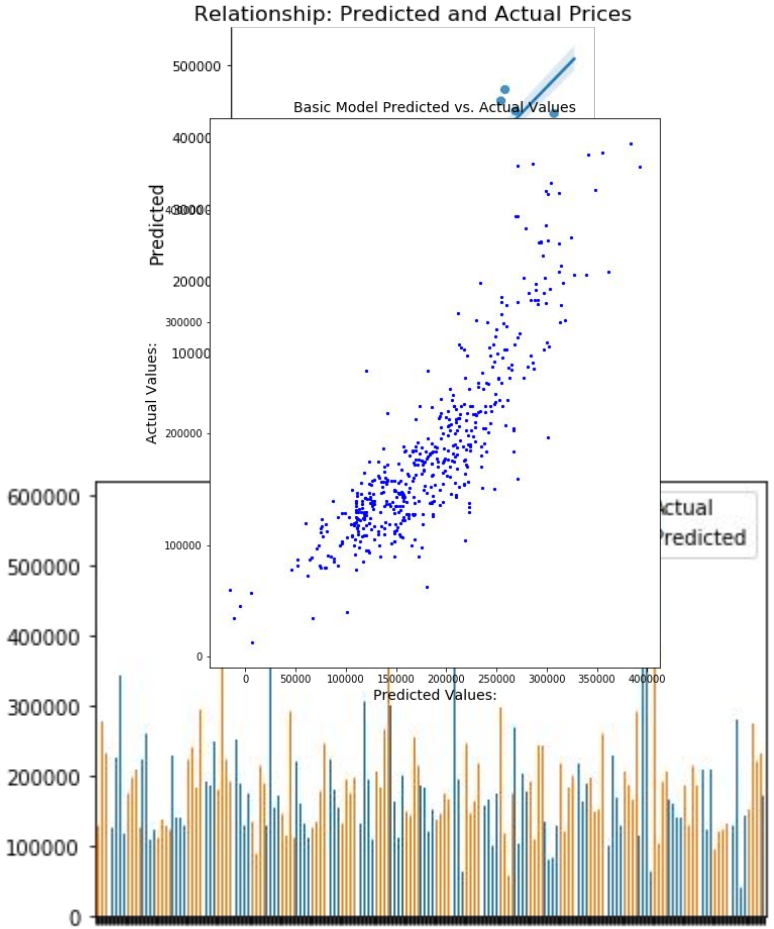


Our Findings:

Although certain “default variables” like quality ratings and total square feet of living space are important factors in the sale-price for a home, there are several other variables our model takes into account:

- **Age** of House, Additions, and Remodeling
- **Quality Rating** of garage and basement
- **Unit Size**
- **Proximity** to external factories: main roads, etc.

We find that these features help maximize accuracy in our model’s prediction..



83.74%

Our model was able to predict almost 84% of the variation in house sale-price using our machine learning methods and feature engineering.

This is a 14.7% increase in sale-price predictability over basic modeling.

Thank You