

# Predicting Home Prices

How understanding the predictions can help you make money



### **Agenda**

- What data are we looking at?
- What do we see in the data?
- How does the model work?
- How can you use these insights to increase the value of your home?



# **Data Summary**

#### What's inside the data set?

#### What data are we looking at?

- Houses Sold in Ames, IA from 2006-2010
- 2930 homes, 82 variables
- Objective quantitative data
  - Square footage
  - Lot area
- Subjective qualitative data
  - Exterior quality
  - Kitchen quality
  - Overall quality/condition



#### Any issues with the data?

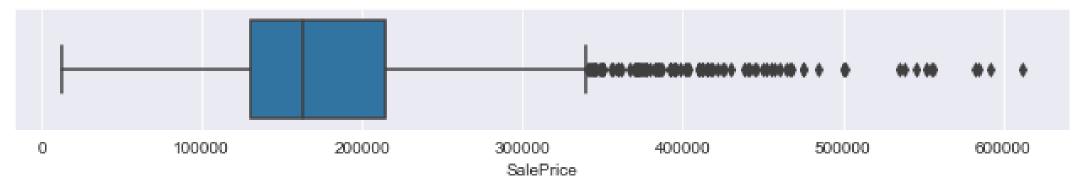
- For subjective home features like exterior quality, we don't know how the rating was derived.
- A few outliers with large square footage and partial sales, also a garage that was built in 2077!
- A lot of 'missing' data that implied that the homes didn't have those features.
   E.g. Pools

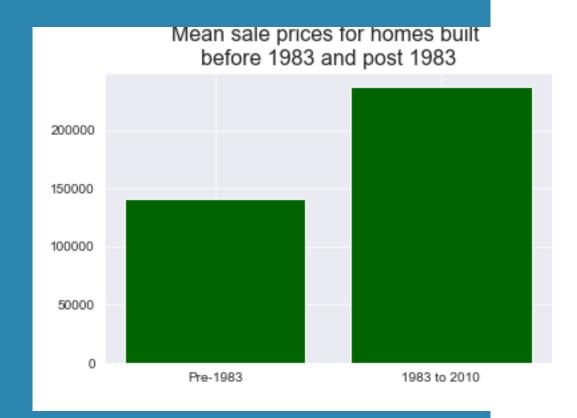


Let's look at some visualizations!



#### Distribution of Sales Prices



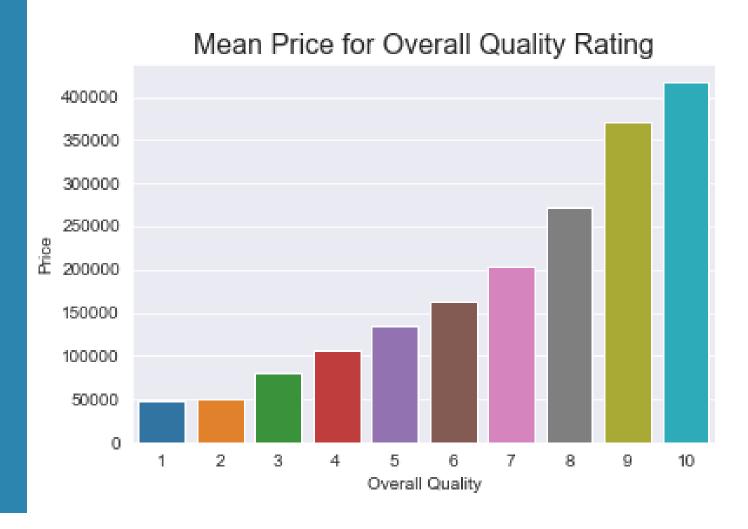


# Age of Home and Sale Price

• Homes prices in Ames, IA jump dramatically when the home has been built after 1983.

#### **Overall Quality**

• Though Overall Quality is a subjective measure, it seems to be a very strong indicator for house price.





## About the models

#### What was the process?



We chose a few different models for our purposes—some are more interpretable than others

We fed all the housing data into our models and let the computer crunch the numbers

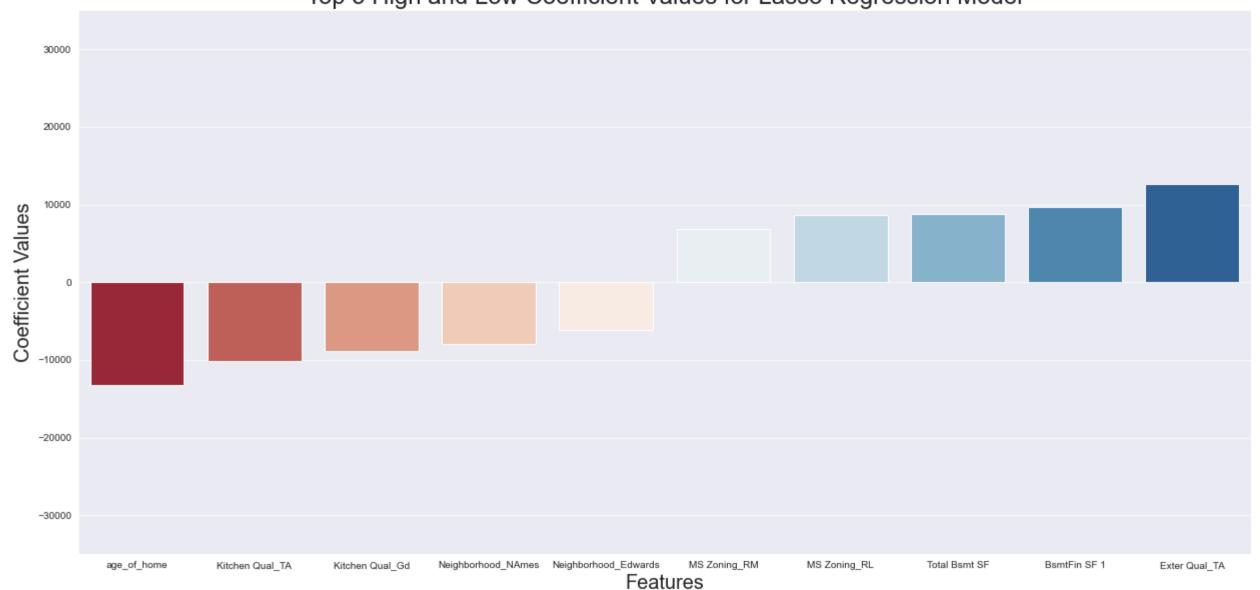




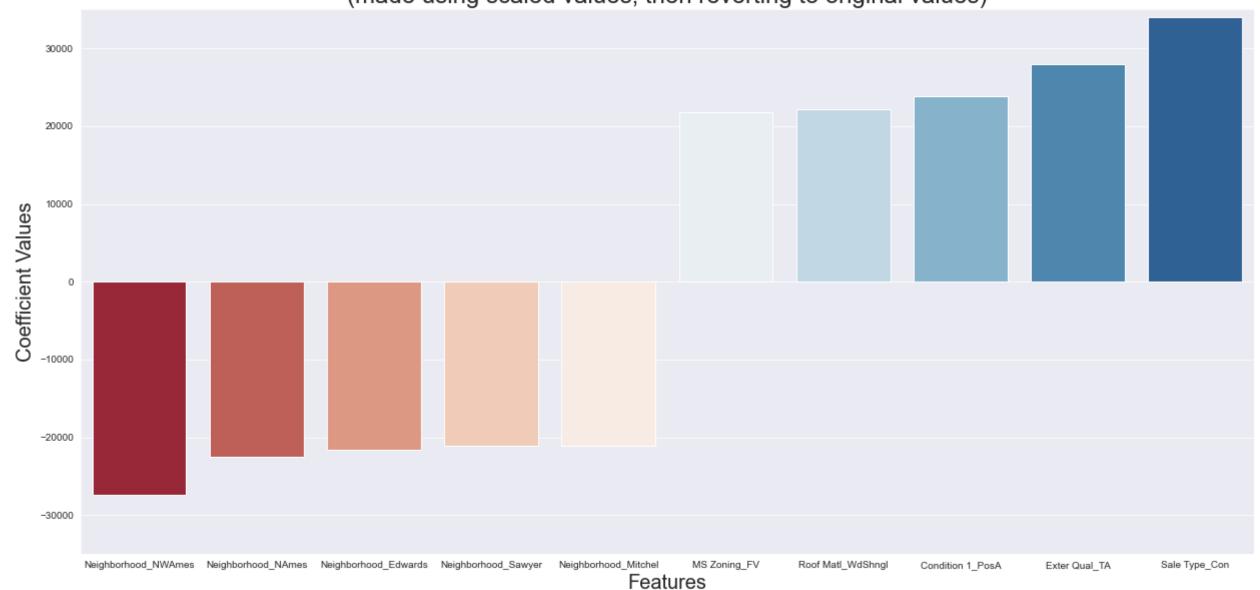
Our models found features of homes that are strongly correlated with home price And now we get to share those results with you!



Top 5 High and Low Coefficient Values for Lasso Regression Model



Top 5 High and Low Coefficient Values for Linear Regression Model (made using scaled values, then reverting to original values)





#### What did we discover?



