

# Ames Housing Pricing Model

Larry Jackelen, Iowan Data Scientist

# Problem Statement

Our firm wants to diversify our portfolio by getting into the real estate business in Ames, Iowa. Like any instrument, we want to know if our return will make the investment worthwhile.

Using available data, build a model for estimating the sale price of a home. Which types of attributes are good indicators of value and which detract?

# The Dataset

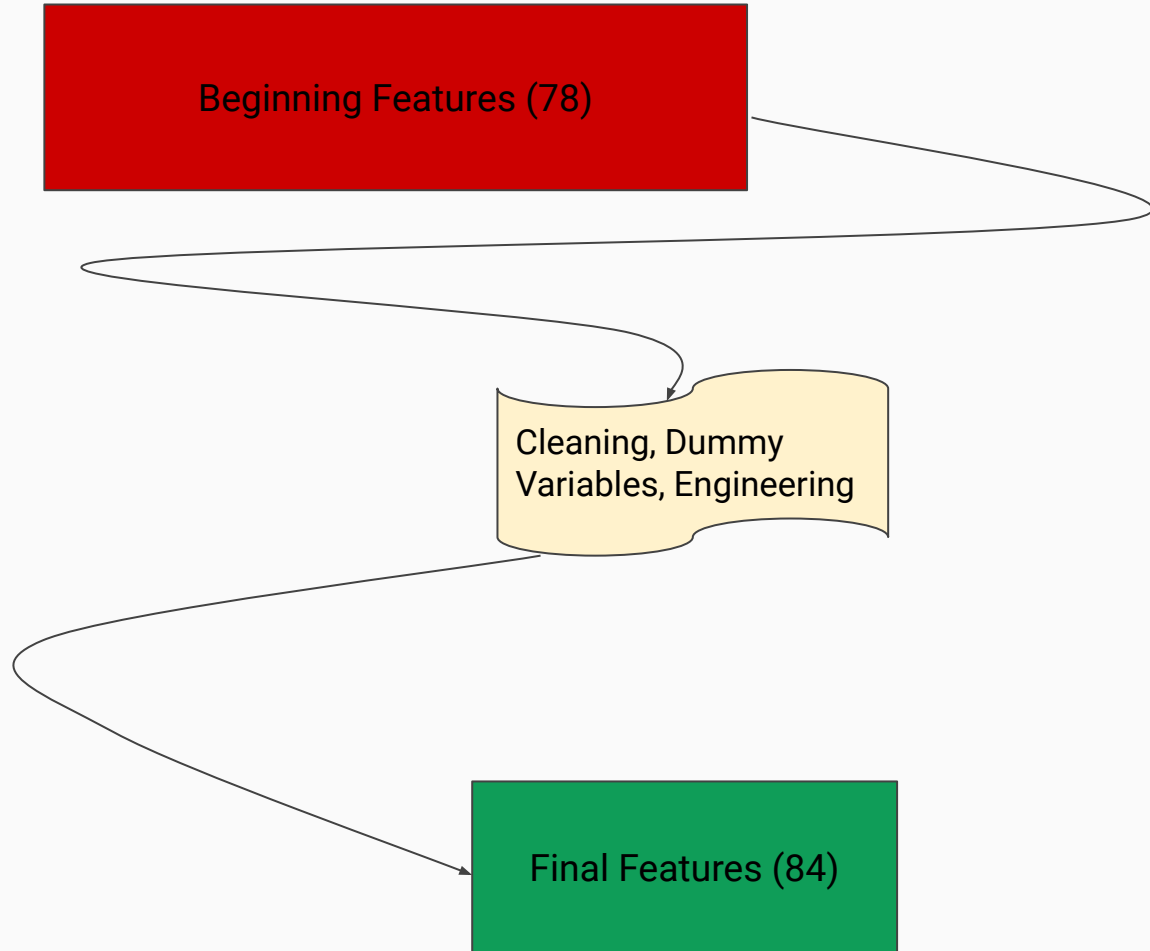
Source: Ames, Iowa Assessor's Office

- 2,051 records of individual residential properties sold in Ames from 2006 to 2010
- 78 attributes per record to divine our model from
- Sale price data for each sale spanning all wallet sizes: \$12k to +\$600k

# Feature Selection

All 78 variables were cleaned, examined, and analyzed to determine the best set of indicators for our model

In the end, after creating helpful variables from the data, among other things, 84 variables were sent to modelling

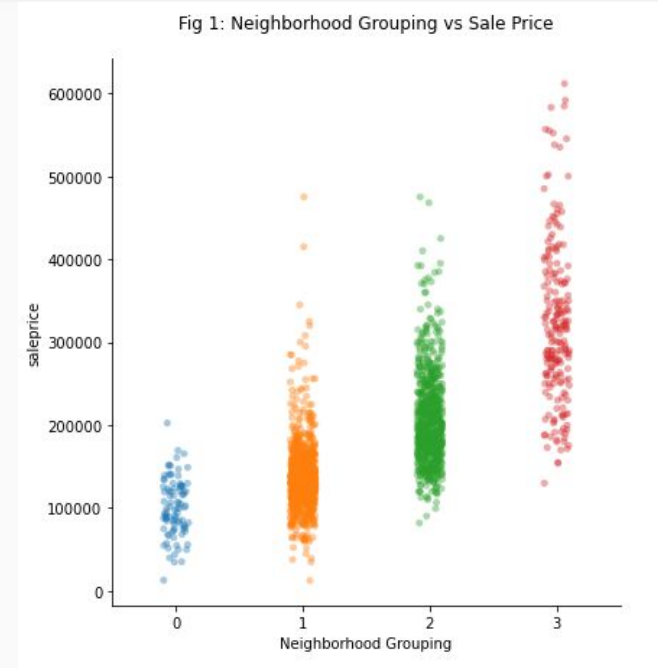


# EDA Callouts: Neighborhoods

28 total neighborhoods available for consideration

Clustered each neighborhood by its mean sale price compared to the total dataset mean and standard deviation

Resulted in 4 distinct groupings of neighborhoods for our model



# Get Out(lier) of Town!

A critical part of data cleaning is the finding and treating of outliers.

Our company is interested in making money in *aggregate*, so removing a few records to increase the models overall accuracy is key!

Note how the “line” now appears clearer with outliers removed.

Other notable outliers removed:

- lot frontage
- lot area
- houses with 6+ baths!

Fig 4. 1st floor Sqft. Outliers

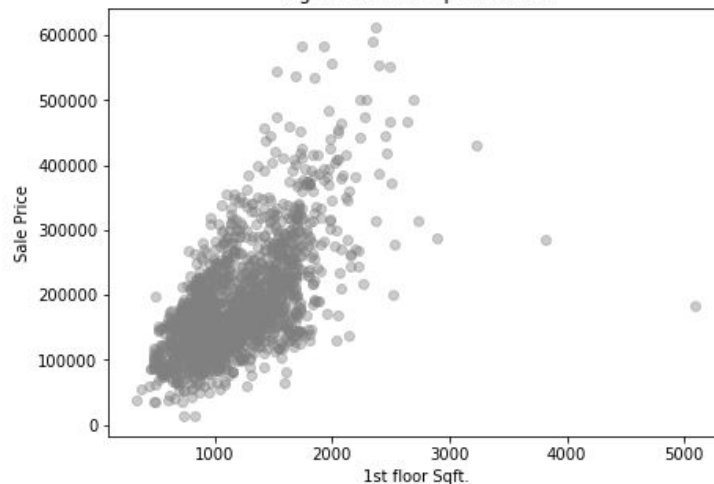
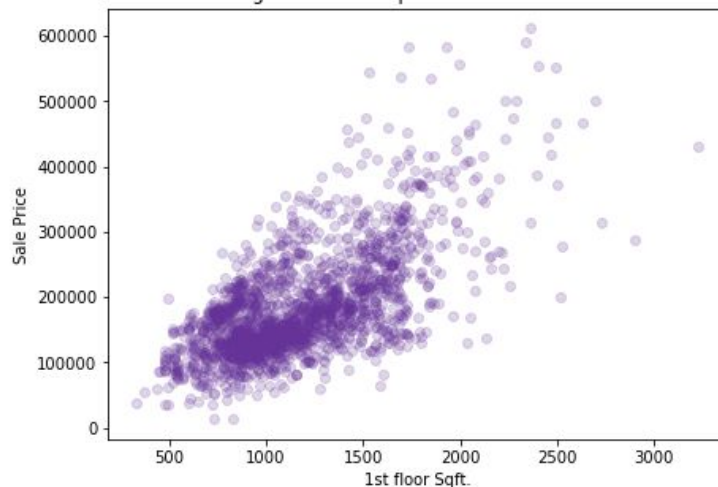


Fig 5. 1st floor Sqft. Outliers Removed



# The Model

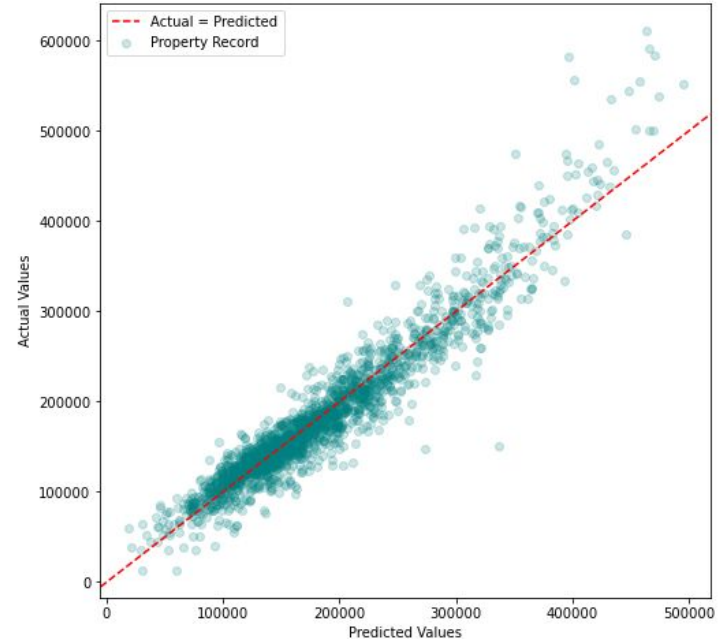
Started with 84 variables

Whittled down to 57 variables

Finished with a model accuracy with a  
95% Confidence Interval of:

**$0.91 \pm 0.02$**

Fig 6. Model Performance



# How can we apply this?

Use the model!

However, for homes not in a registrar, focus on these data points when boots are on the ground:

## Things to look for:

1. Neighborhood Grouping 3
  - Membership here adds \$33k in Sale Price
2. Above Grade Living Area
  - Every 500 sqft add \$25k in Sale Price
3. Floating Village Residential Zones
  - Adds \$15k in Sale Price
  - Second closest zone is Residential Low Density (adds \$4k)

## Things to avoid:

1. Basements
  - Decreased value of \$23k
  - Require at least 1,100 basement square footage to make up for it
2. Older homes
  - Homes over 100 years old go for \$20k less
3. Remodeled Homes
  - Remodeled homes go for \$3.6k less

\* All values are with holding everything else constant



# Thank you!

Contact me:

The cubicle by the men's room  
12 Corn Dr  
Ames, IA

