AMES IOWA HOUSING DATA ANALYSIS: PREDICTING SALES PRICE

Presented by:

Tanner Pitts

WHY IT MATTERS

It is important to understand what drives your customers in any given market. In this analysis, we will look at what key features drive up the cost of a home and then attempt to predict the cost of a home based off these features.

Reasons why this is important:

- **focused construction**
- **trend predictions**
- **loss prevention in purchasing**

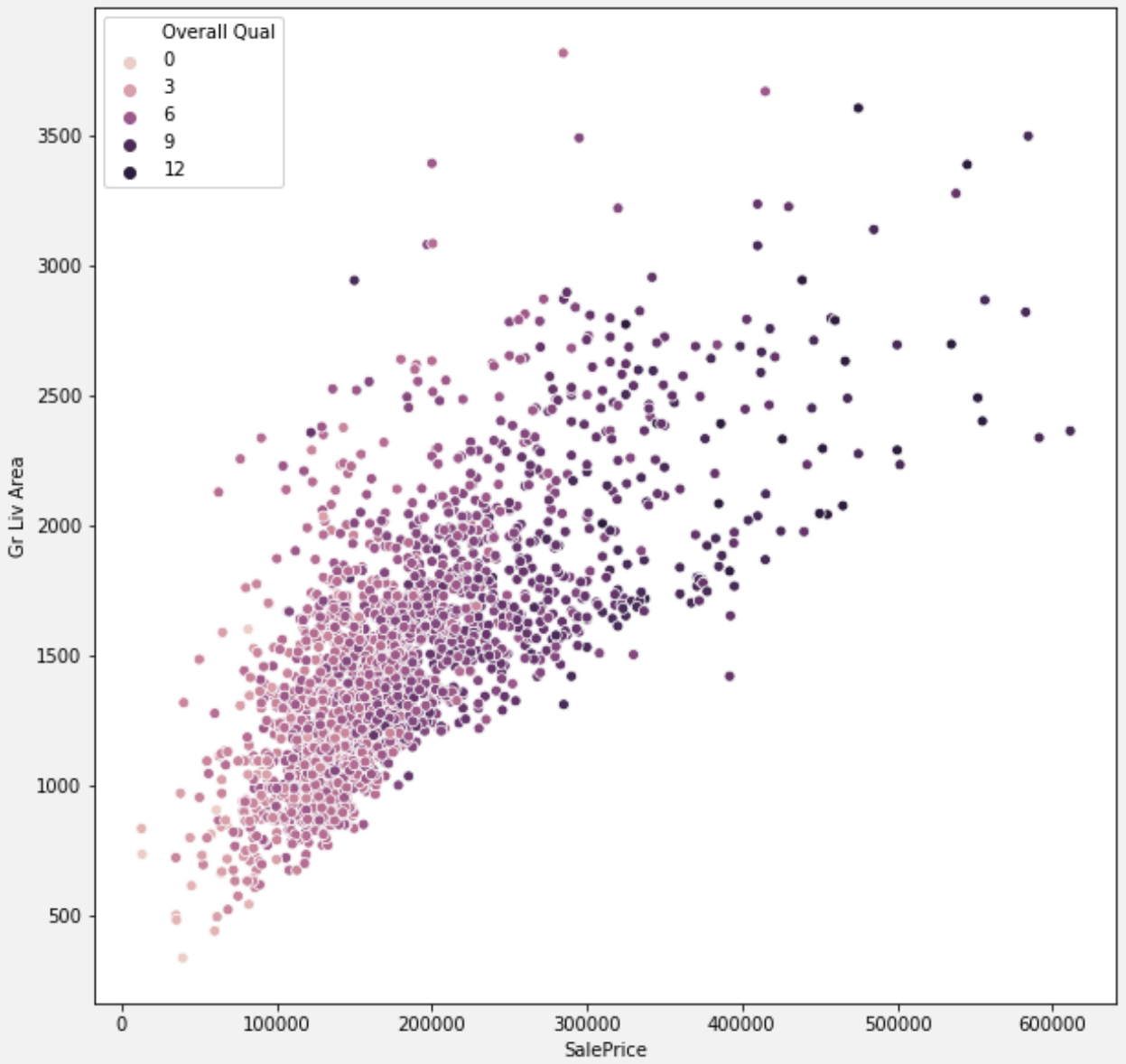
THE DATA

Observations: 2048

Variable	Units	Definition
Gr Liv Area	Continuous (Ft^2)	Size of living area
Lot Area	Continuous (Ft^2)	Size of lot
Overall Qual	Ordinal (1-10)	Quality of material used
Overall Cond	Ordinal (1-10)	Condition the house is in
Half Bath	Discrete	Number of half bathrooms
Bedroom AbvGr	Discrete	Number of bedrooms (Basement not included)

RANGE OF OUR DATA

- Minimum price: $12,789.00
- Maximum price: $611,657.00
- Range: $598,868.0
- Mean price: $181,495.16
- Standard Dev: $79,311.90



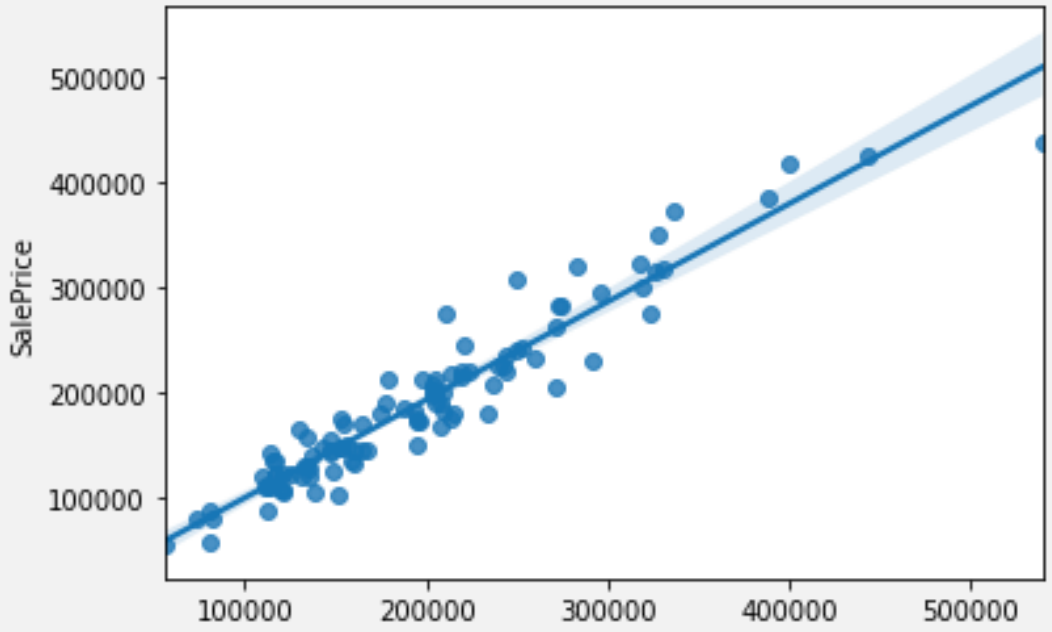
CLEANING THE DATA

- Variable that had to be removed: Pool QC, Fence, Misc. Feature, and Alley
- Variables changed from categorical to Dummies: Year Sold, Sales Type, Neighborhood, MS SubClass, Exter Qual, Foundation, Paved Drive, and Lot Config
- Variables changed into unique dummies: Heating, Zoning, and Electrical
- Filled in missing values with 0: Basement sqft, Garage Area, and Garage Cars

THE MODEL

Linear Regression:

- Pros:
  - Interpretive Outputs
  - Simple model
- Cons:
  - Linearity Assumption
  - Independence of Variables



INTERPRETATIONS

Key Factors

- Positive
  - Neighborhood
  - Type of Sale
  - Half Bath $1,358
  - Living Area $69.27/sqft
  - Car space $1,222.66/car
- Negative
  - Neighborhood
  - Type of Dwelling
  - Exterior Material
  - Full Bath $-122.69
  - Bedrooms $-7135.38

IS OUR MODEL GOOD?

Key Metrics:

R-squared score: 0.913

RMSE(Root Mean Squared Error) : 23374.98 (Key Focus)

Concerns:

Very specific to Ames Iowa

Non Linear relationship