CITY OF
Ames™
cityofames.org

Ames Iowa
Housing Price Analysis

Jared Delora-Ellefson
Data Scientist

Problem Statement

Given the available data (2006-10):

- **What factors most affect Property Sale Prices?**
- **Is it possible to accurately predict Property Sale Prices based on these factors?**
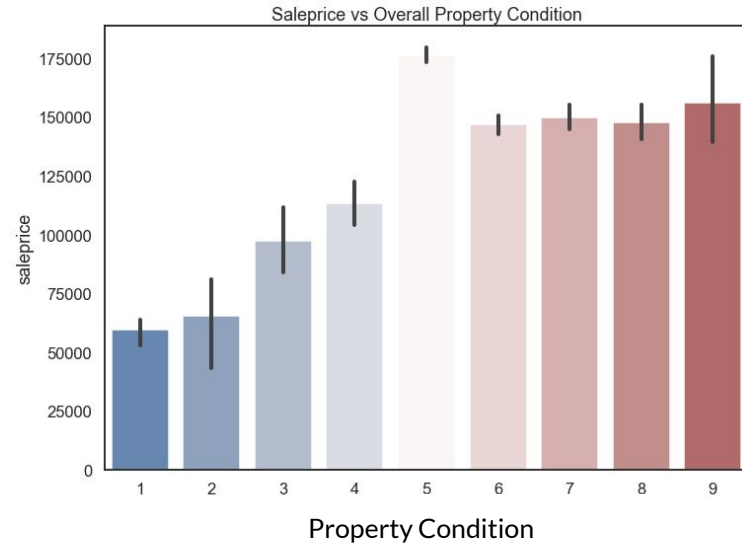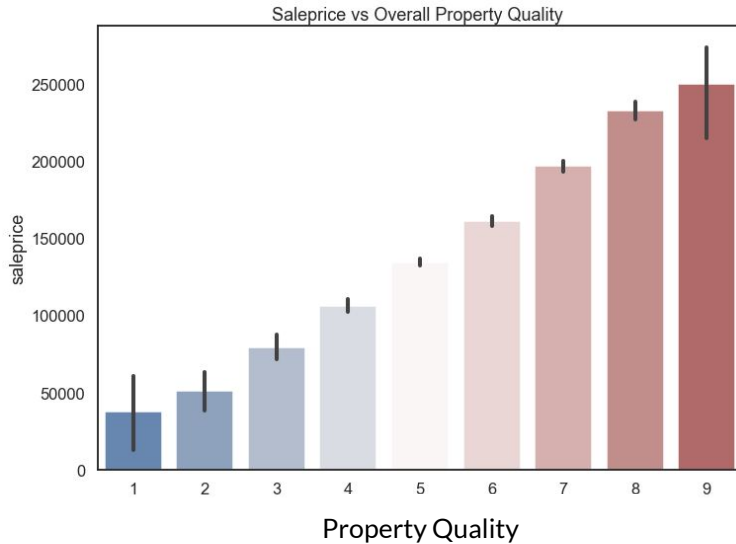
Analysis of Current Data

- Lot Area Size and Sale Price

- Property Sale Price Correlates only loosely with Lot Area
- Property Overall Quality is reflected in the shading
- Overall Quality appears to correlate well with Sale Price

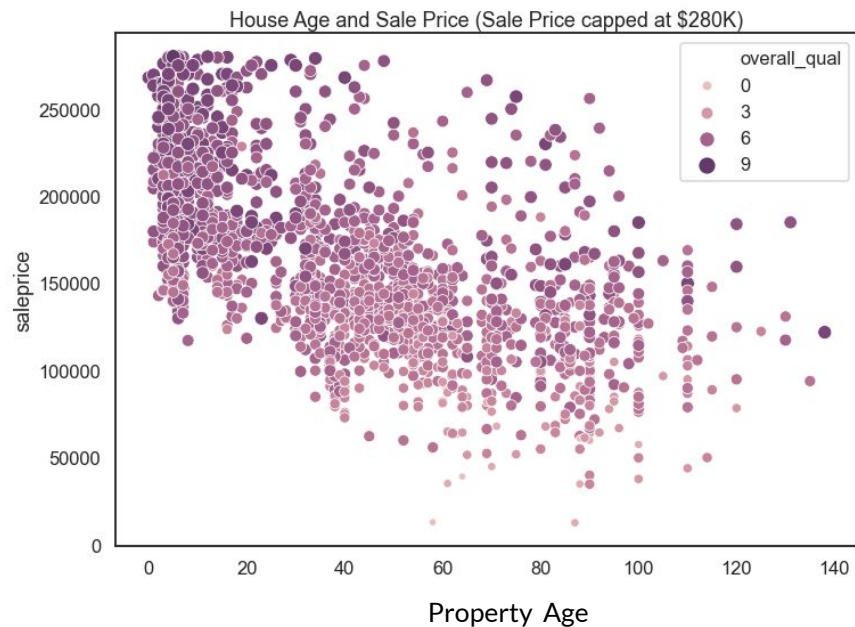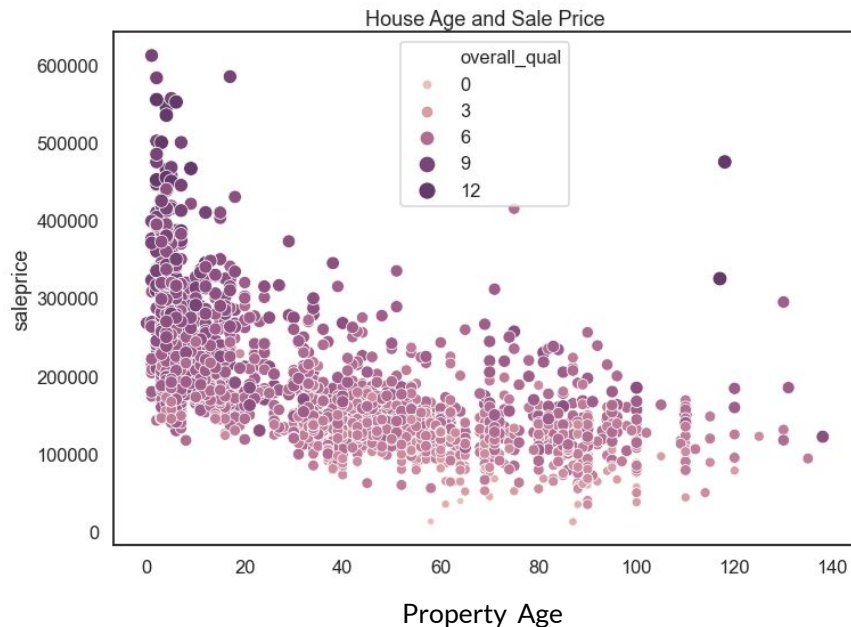Sale Price vs Lot Area (Property Quality in Shading)

Property Quality/Condition and Sale Price

- **Property Quality correlates very well with Sale Price**
- **Property condition correlates with Overall Condition for only poor condition properties**

Property Age and Sale Price

House Age and Sale Price

House Age and Sale Price (Sale Price capped at $280K)

- **There are a lot of outliers in the Sale Prices.**
- **Only Sale Prices of less than $280K were considered in the model to eliminate these outliers.**

CITY OF
Ames

Property Feature Correlations

Correlation Map of Property Features

Property Quality, Full Baths, and Garage Size are the variables that most closely correlate to Sale Price.

Property Model Features

- **Lot Area**
- **Property Quality**
- **Property Condition**
- **Central A/C**
- **Full Baths**
- **Garage Capacity**
- **Property Age**
- **An interaction term: Property Quality × Lot Area**

CITY OF Ames™

Model Metrics

	Coef	P>|t|
Dep. Variable: saleprice		**R-squared:** 0.768
lot_area	0.0868	0.922
overall_qual	1.353e+04	0.000
overall_cond	4856.9474	0.000
central_air	6657.4231	0.021
full_bath	1.585e+04	0.000
garage_cars	9711.9172	0.000
house_age	-386.1123	0.000
interaction1	0.5540	0.000

A score less than 0.5 is desired

These scores should be close to each other.

- Training score (R-Squared) **0.76**
- Testing score (R-Squared): **0.74**
- Check the RMSE for Training data: **24070**
- Check the RMSE for the Pred data: **24500**

These scores should be as low as possible.

- **Lot Area did not turn out to be important**
- **The model has moderate bias, and low variance. More complexity could be added.**

CITY OF Ames

Next Steps

- **The accuracy of the model can be increased by adding more features**
- **More advanced techniques should be employed to improved the results**
- **Add Neighborhood Socioeconomic Considerations, e.g. Average Income, Unemployment Percentages, etc.**

CITY OF Ames™