Home Value Estimator and Recommendations for Real Estate Investors

A regression model based on the Ames Housing Dataset, proposing to a local real estate company

Melody Wu, February 22, 2021

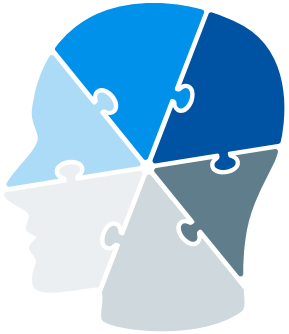
How much is a home worth?

Usually...

But, actually...

Our Goals

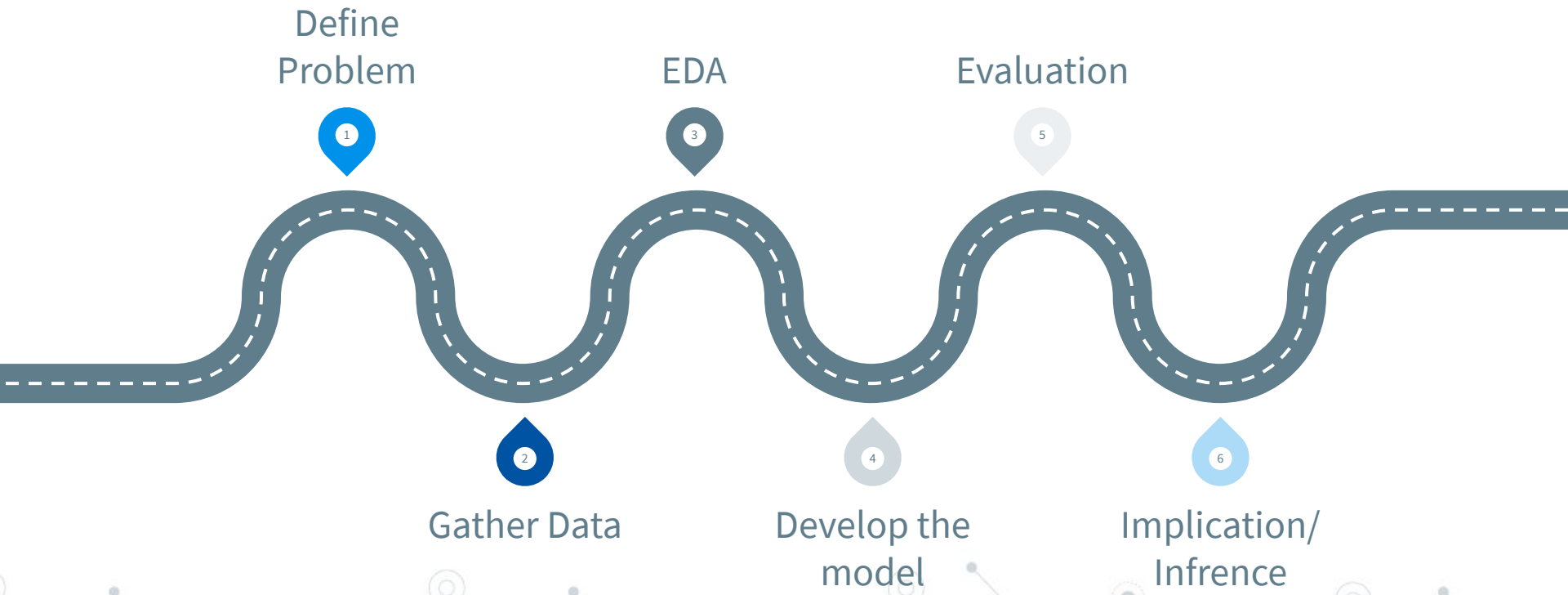
Insights



Benchmark

Optimization

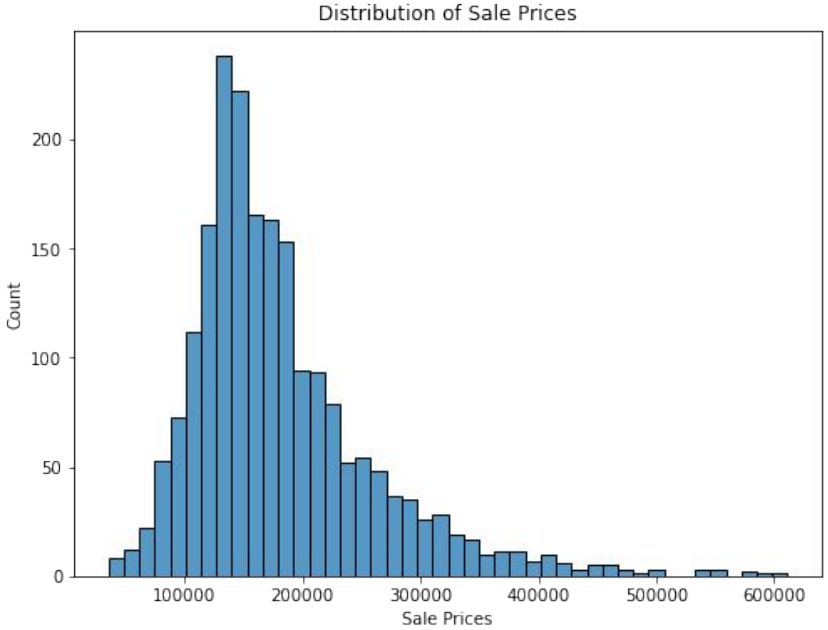
Project Roadmap

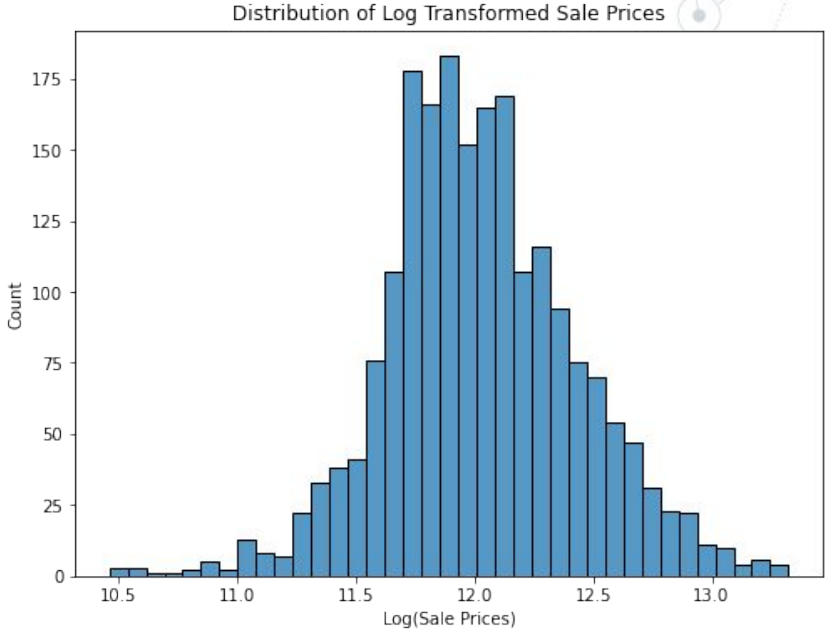


Dataset

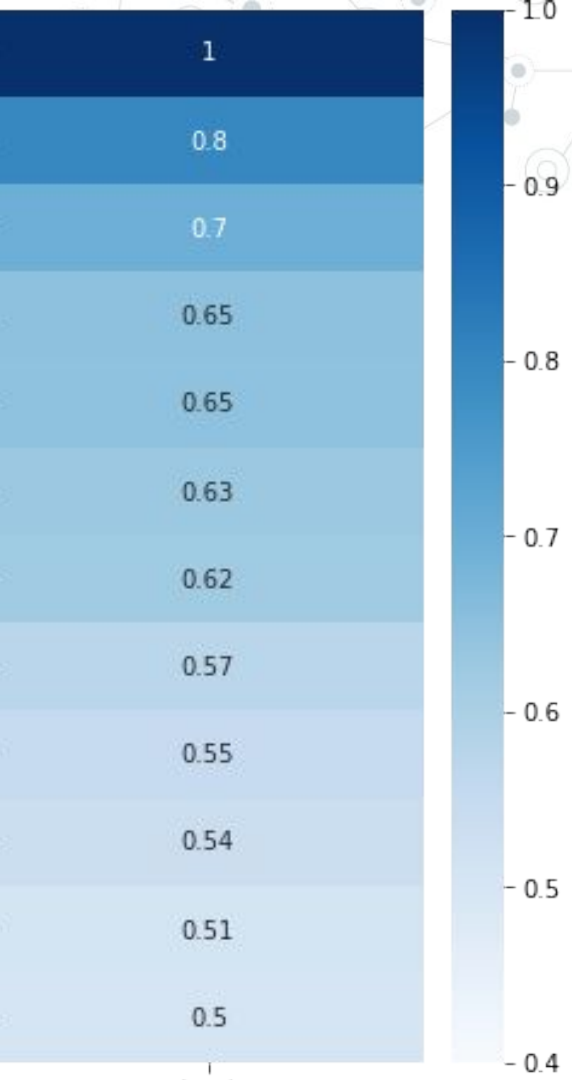
- Ames Iowa Housing Data (2006 - 2010)
- Observations and features:
  - 2051 observations for training
  - 876 observations for testing
  - 34 numeric variables (20 continuous + 14 discrete)
  - 46 categorical variables (23 nominal, 23 ordinal)

Exploratory Data Analysis (EDA)





Variable of interest: Sales Price (in US Dollar)



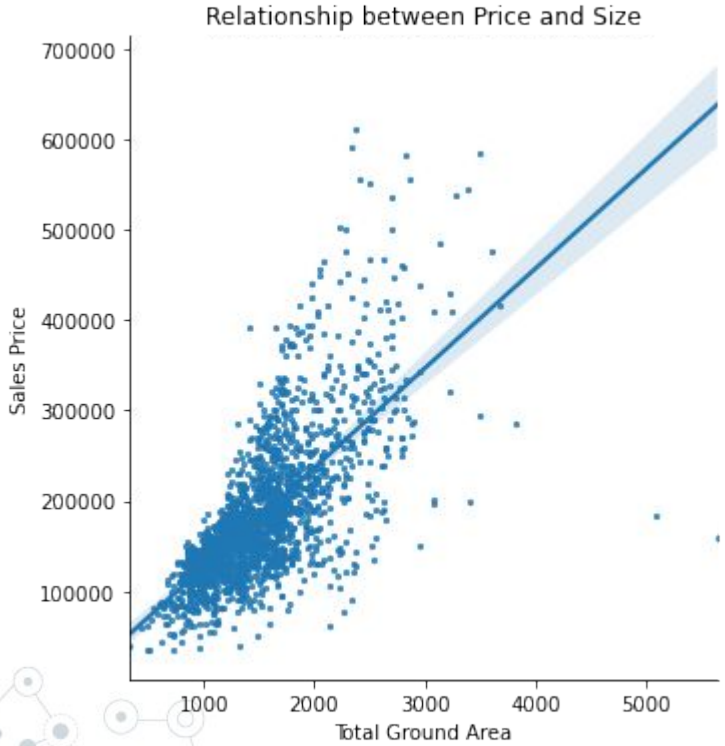
Size:

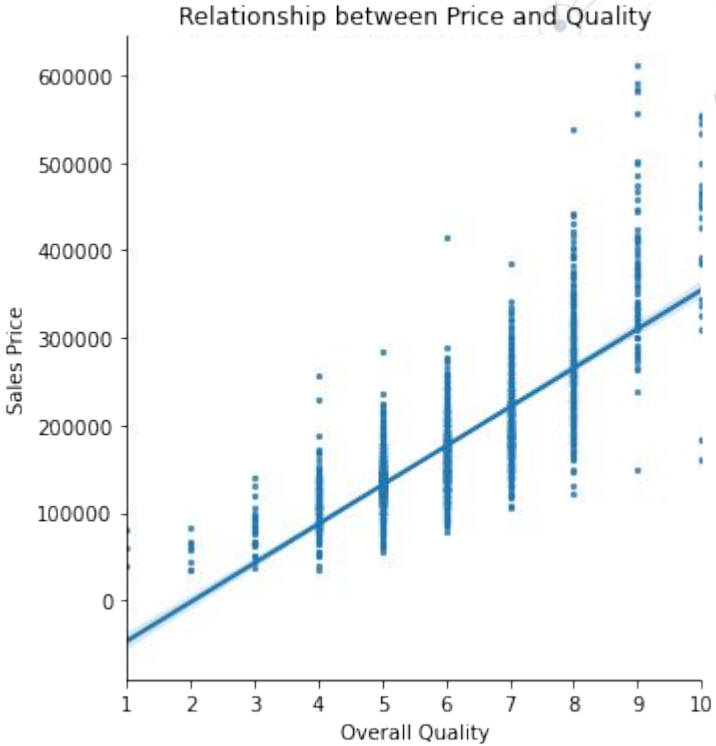
- Above ground living area square feet (0.7)
- Garage Area (0.65)
- Total square feet of basement area (0.63)
- First Floor square feet (0.62)

Quality:

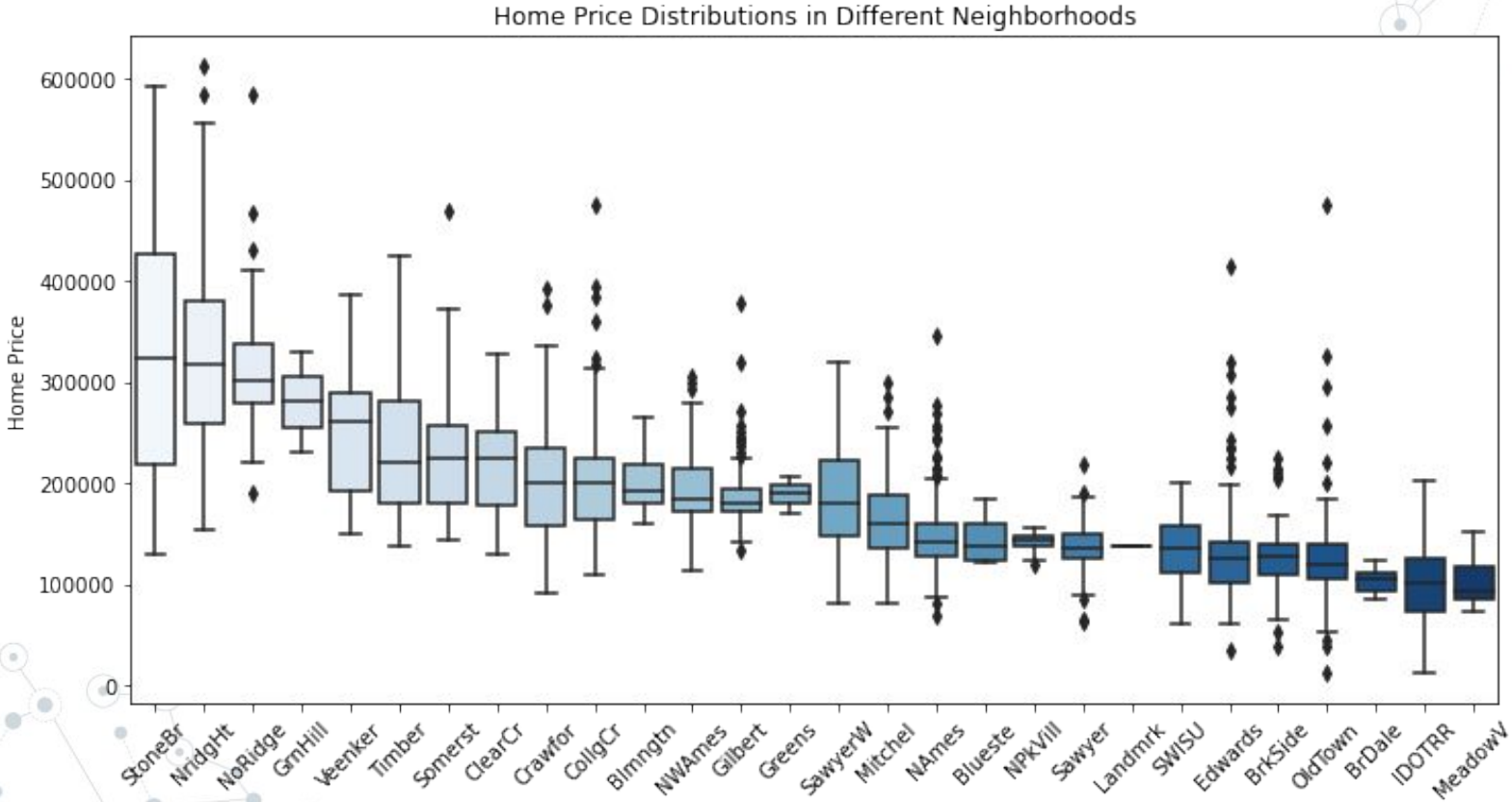
- Overall Quality (0.8)
- Original construction date (0.55)
- Remodel date (0.55)
- Qualitative Evaluations
  - Foundation
  - Basement
  - Heating/central air
  - Kitchen

Size and Quality

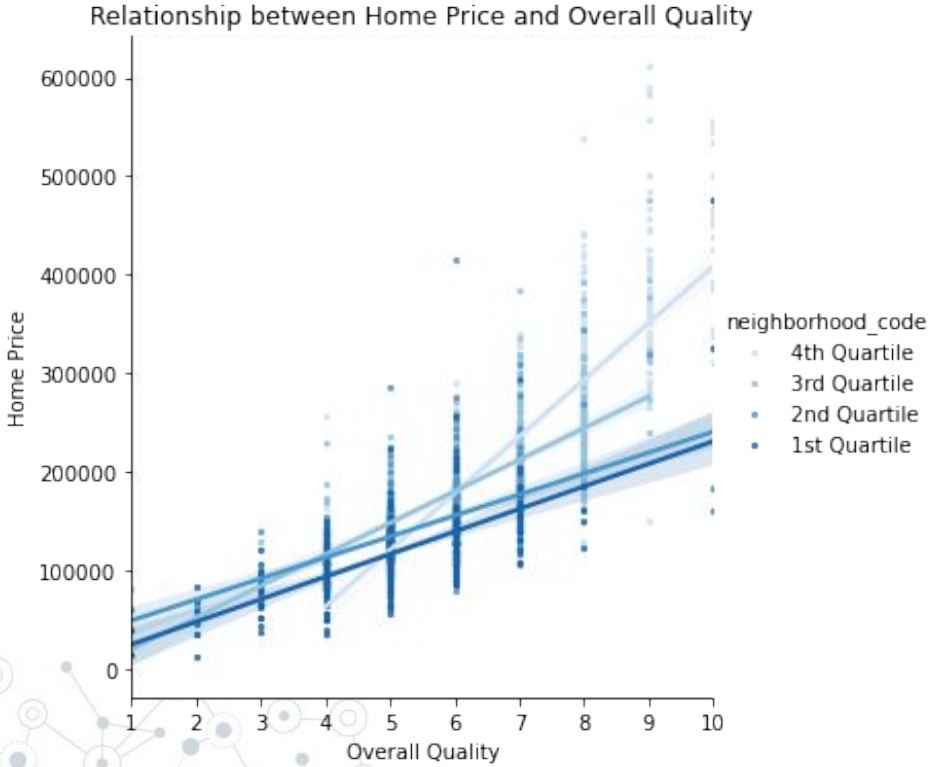




Neighborhood Matters



Neighborhood Matters



Model Development and Evaluations

Model	Training R-Square	Testing R-square
Model 1: Linear Regression (selected numeric only)	0.837	0.844
Model 2: LASSO Regularized Regression (all numeric)	0.908	0.854
Model 3: Ridge Regularized Regression (all numeric)	0.882	0.854
Model 4: Linear Regression (selected numeric + categorical)	0.861	0.858

Key features will increase home values:

Neighborhood:

- Stone Brook
- Northridge Heights
- Northridge
- Green Hills
- Veenker
- Timberland
- Somerset

Size and Quality

- Overall quality
- Square footage
- Heating quality
- Kitchen quality
- New vs. old (remodel year)
- Regular lot

Add-ons

- Central air
- Fireplace
- Basement exposure

Strategies and Considerations

1. Check the **neighborhoods** (safety, school, convenience)
2. **Size matters** - both above ground, basement, garage
3. **Overall quality** is more important than other single-item qualities
4. Quick fix and high potential return: improve evaluation from Good to **Excellent** for **kitchen** and **heating**
5. **Remodel**, remodel, remodel!
6. Key features to install: **central air, fireplace**

Areas of Improvements

- Qualitative study
  - In-depth interviews/focus group discussion with buyers, sellers, investors and to understand their needs, priorities, and how and why decisions are made (i.e. what is special in the neighborhoods)
  - Ethnography and on-site visits
  - Expects talks (developers, inspectors)
- Data and Methods
  - Subgroup analysis
    - for extremely large houses
    - Different neighborhoods (fixed/random effects, interactions)
  - Sample quality (i.e., weights, time of collection)
- Different models for prediction and inference

Thank you!

Any questions?

You can find me at:

melody.jwu@gmail.com