How much do improvements contribute to home prices?

By David Richter, Flatiron School, May 2022

Improvements and Home Price

- Many people who want to sell their home want to know whether it's worth it to invest in improvements.
- Improvements obviously add value to a home
- On the other hand, improvements cost money!
- So how do we determine when these improvements are worth their expense?
- The answer: Linear Regression

How Linear Regression Can Help Us

- To determine whether we should make improvements we need to first determine their value.
- The value of an improvement can be defined as the amount or percentage that a given improvement contributes to the price of a home.
- A Linear regression model for home prices uses a range of variables that may contribute to the price of a home to try to predict home prices. If we design our model carefully, it will also help us to estimate how much these variables individually contribute to price.

Our Regression Model

- The linear regression model that I'm going to discuss today has been built and trained on data from King County, WA, so it is only valid for predicting home prices in this area. It takes in a number of variables, including:
    - Square footage
    - Lot Size
    - # of Bathrooms
    - Zip Code
    - Sale Month
    - Grade (quality of the construction)
   - Condition (relative to age and grade)

Predictor Variables

- We're only interested in how some of our predictor variables contribute to the overall price.
- This is because many variables are outside of our homeowners' control., such as zip code and lot size.
- The variables that I'll focus on are all within the homeowner's control, specifically:
  - Grade: this is the quality of the house's construction, and can be affected by extensive remodeling
  - Square footage: which can be increased through remodeling
  - Bathrooms: which can be added to a home through remodeling
  - Condition: The superficial quality of the home, which can be improved through repairs and renovation.

Further details on these termscan be found at https://info.kingcounty.gov/assessor/esales/Glossary.aspx?type=r

Model Performance

- Before we move on to our model's predictions, we'll look briefly at it's performance.
- This scatter plot represents our model's performance when tested trained on one set of data and tested on another set of data.
- When tested on data that it hadn't been trained on, the model accounted for over 88% of the data's variance.
- On the right, we can see how actual home prices fall above or below the prices predicted by the model (the red line).
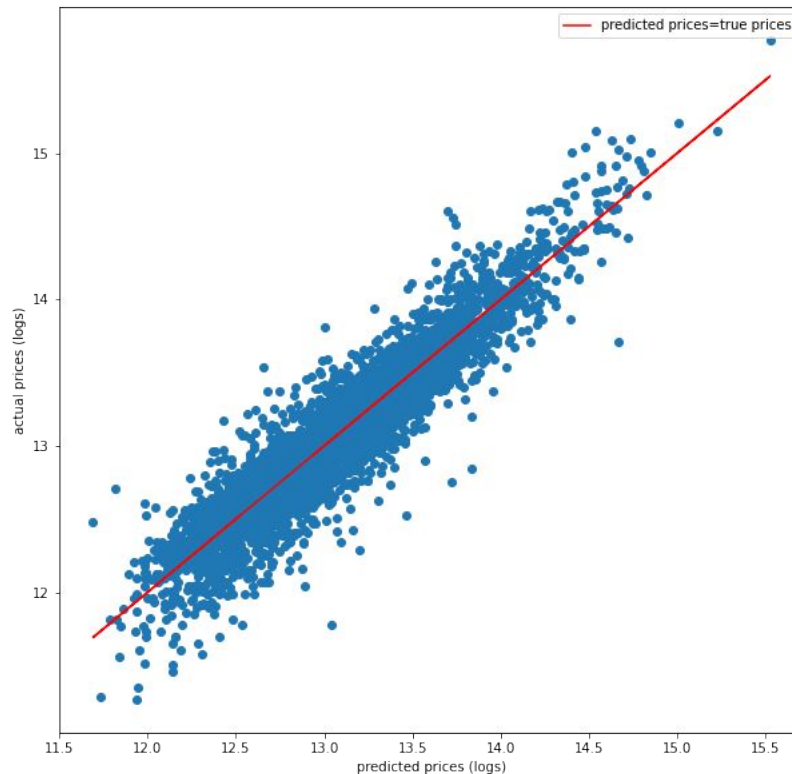
Effects of Predictor Variables

Below, we can see how changes in our predictor variables affect the model's predictions.

According to our model, we have 97.5% confidence that:

- 1 point increase in home's grade corresponds to: minimum 10% appreciation in value
- 1 point increase in a home's condition corresponds to: minimum 4% appreciation in value
- 1 additional bathroom corresponds to: minimum 4% appreciation in value
- 1% increase in living area corresponds to: minimum 0.38% appreciation in value

Putting Our Model to Use

Here's an example of how are model can be put to use:

- Let's say we currently value a home at $500,000.
- Because the model predicts with 97.5% confidence that an added bathrooms adds at least 4% in value to the home, the predicted value in dollars of an extra bathroom would be a minimum of $20,000.
- If we incorporate our 97.5% confidence percentage into the predicted value, we get an expected value of $19,500 for an additional bathroom.
- This expected value represents the maximum that a homeowner should spend on adding a bathroom. If it can't be done for less than this, there's a good change that the improvement won't pay off!

Thank you!

I hope you enjoyed my presentation.

For more details on my linear regression model please see this project's github page at https://github.com/DavidKRichter/dsc-real-estate-project

You can email me with any questions at d.richte@gmail.com