Should you renovate your house or not?

Neslihan Bisgin

Self paced student

https://github.com/nxbisgin/

OUTLINE

1. Business and Data Understanding
   a. Problem Statement
   b. Data and Preparation
2. Modeling and Regression Results
   a. Multiple Linear Regression Model
   b. Most Effective Features
3. Recommendations
   a. Price Prediction
   b. Renovation

1.
Business and Data Understanding

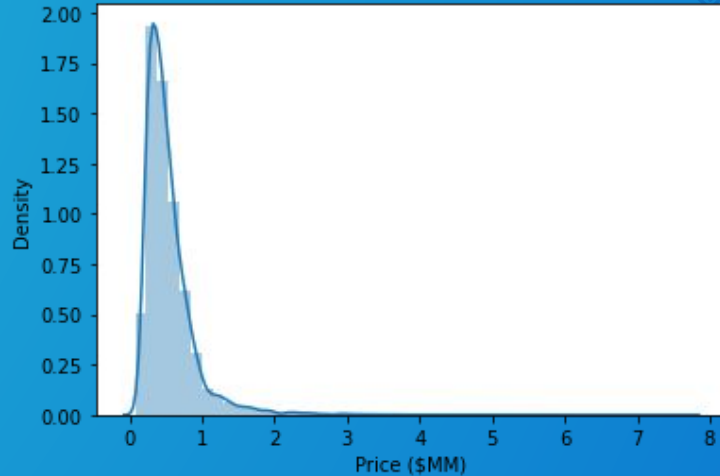Providing advice for home buyers/sellers

Business Problem

Home Sellers in Kent County

Will renovating my house increase the value of it? If so, which features should I focus on renovation?

Data

King County House Prices Data:

- Total of 21,597 houses
- 21 features for each house
- Price:
  - Min: $78 K
  - Max: $7.7 M
  - Avg: $540 K

Data Preparation

- Filled in missing values
  - ex: If no value for "waterfront", there is no waterfront view

- Removed extreme values, left with 20,932 houses
  - ex: Houses with more than 6 bedrooms

- Created new features that are derived from existing ones
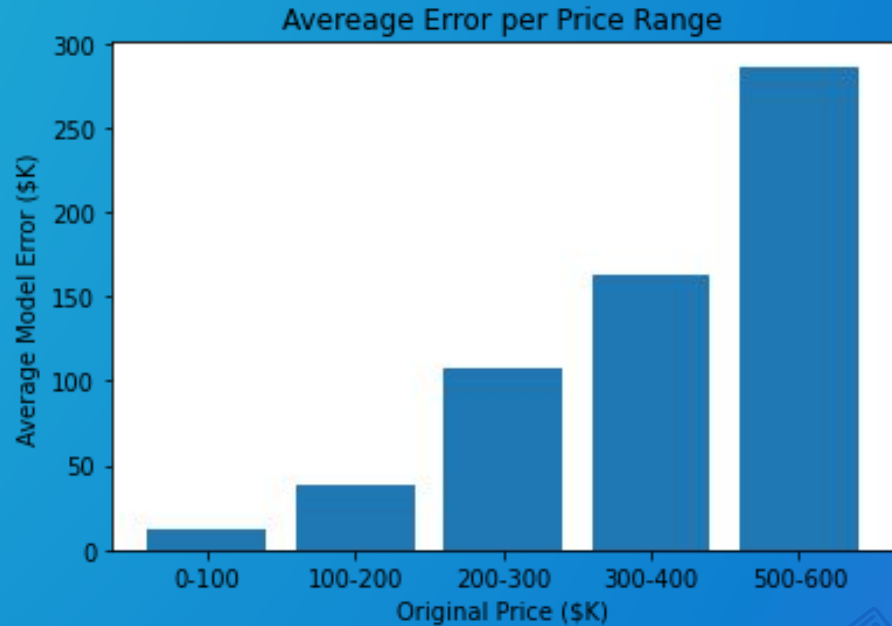 - ex: Given the year built, find the age of house

2.
Modeling and Regression Results

Modeling the house prices with given features

Model Performance

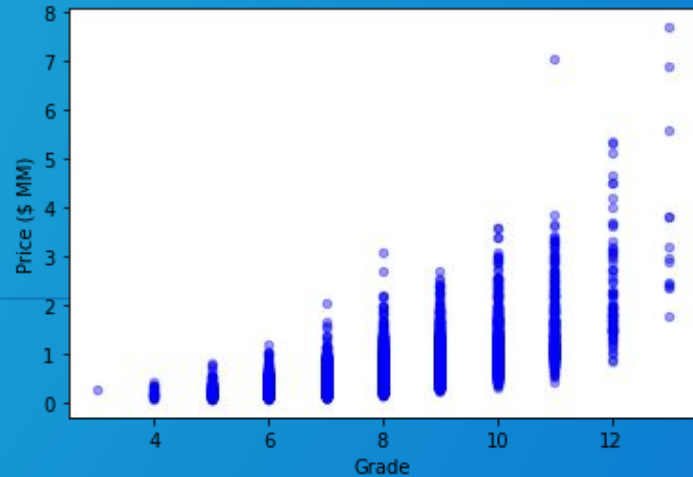- Multiple Linear Regression
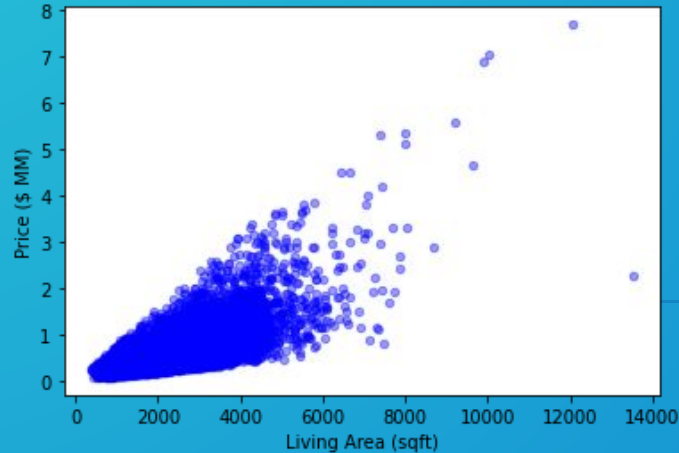- Model can explain ~56% of variability in data

Most Important Feature: Grade

- Price increases most with the increase in the **grade** of the house.
- If the grade increases by 1 unit, the price increases by 28%.

Second Most Important Feature: Living Area

- Second most effective feature on the price is the **living area** of the house.
- If the living area of the house increases by 10%, the price will increase by 1%.

3.
Recommendations

Advice for buyers/sellers

Which features are most effective of the price?

- Overall Grade of the house determined by King County: Related to construction and design of the house. Ranges from 1 to 13, from very poor to excellent.

- Living Area: Square footage of living space in the house.

Should the house be renovated?

- I would highly recommend renovating the house if the grade of the house is increased or more living area is added by renovation.

Next Steps

Zip code: I believe prices should be mostly affected by the location of the house. If zip code information is included in the model, the accuracy will increase.

Age of house: I believe age of house should also be important for predicting the price but could not be used in the model due to non-linear relation. Grouping ages may help in improving the model..

Thanks!

Any questions?

You can find the details of the project:
https://github.com/nxbisgin/phase-2-project-Kent-County-House-Price-Prediction-Linear-Regression/