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 Buyers in Kent County

Given the features of the house, is the price reasonable?

Which features are most predictive of the price?

Home Sellers in Kent County

What is the value of my house with the given features?

Will renovating my house increase the value of it?

Kent County House Prices Data:

- Total of 21,597 houses
- 21 features for each house

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
- Converted string values to integers, did necessary transformations to get the data ready for modeling.

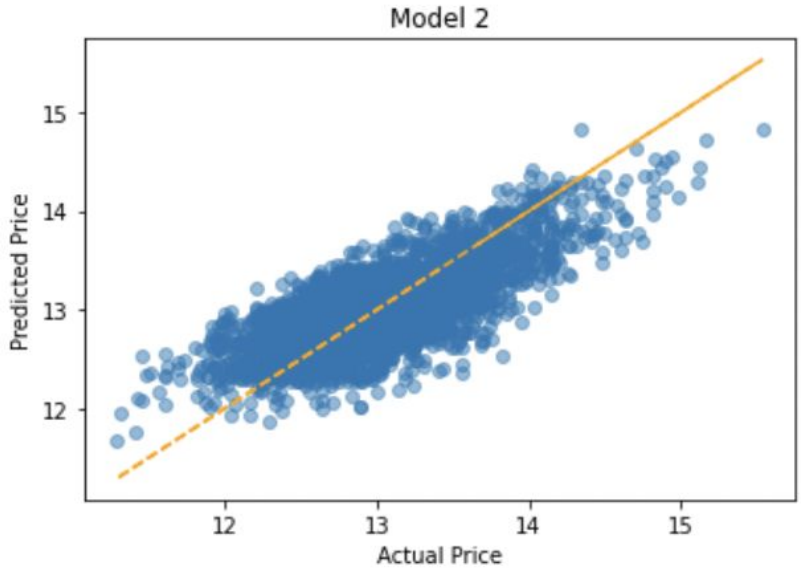
2.
Modeling and Regression Results

Modeling the house prices with given features

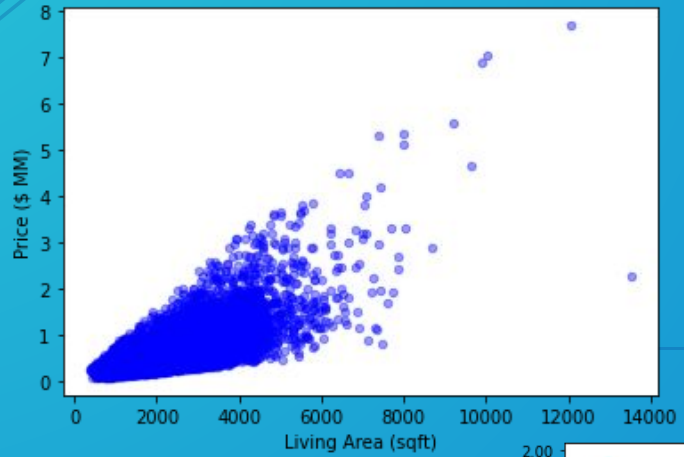
Model Performance

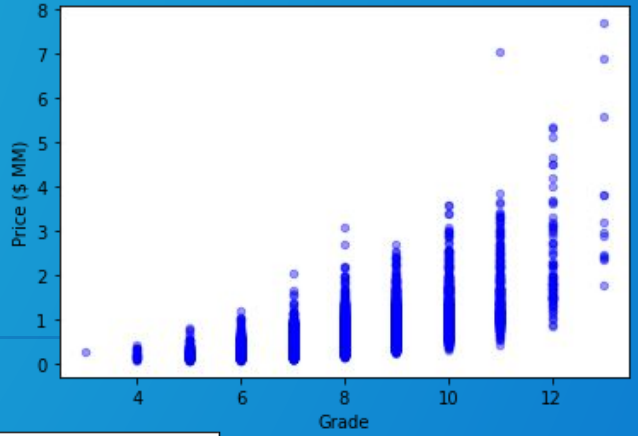
- Multiple Linear Regression
- Model can explain ~56% of variability in data
- Price increases most with the increase in:
  1. Grade of the house
  2. Living area of the house
-

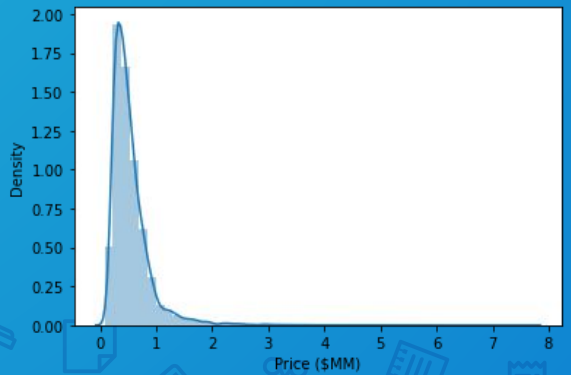
```
Train score for Model : 0.5714089891539308
Test score for Model : 0.5562907717946383
```



Most Important Features For Price







Min: $78k
Max: $7.7M
Avg: $540k

3.
Recommendations

Advice for buyers/sellers

How can I help?

Buyer

Given the features of the house, is the price reasonable?

I can predict the house price with 56% accuracy.

How can I help?

Buyer

Which features are most predictive 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.

How can I help?

Buyer

Should I renovate my house?

- 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: Information about zip code is not included in the model. I may need to go deeper into zip codes to incorporate zip codes.

Age of house: I believe age of house should be important for predicting the price but could not be used in the model due to non-linear relation. I may use binning to incorporate age into 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/