Should you renovate your house or not?

Hello!

I am Neslihan Bisgin

Self paced student

https://github.com/nxbisgin/

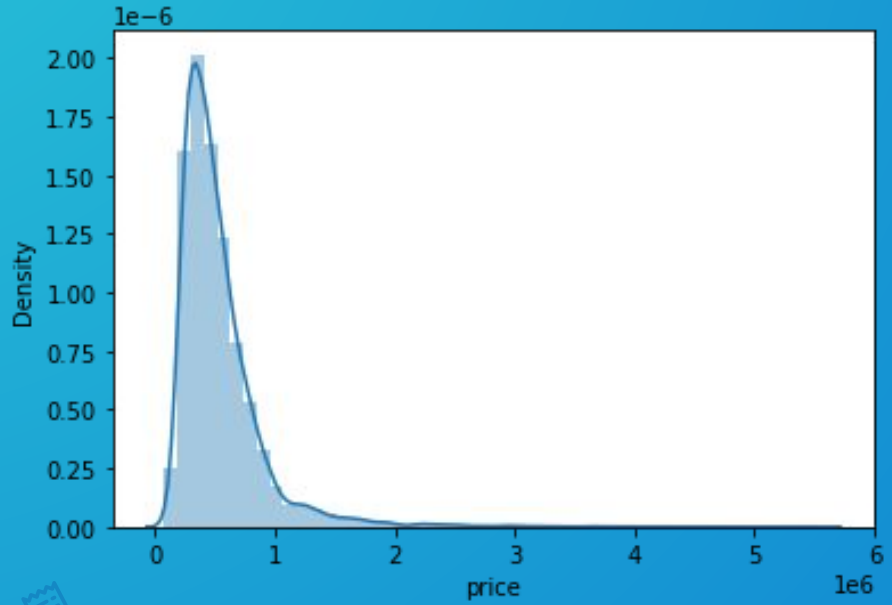
1.
Business and Data Understanding

Providing advice for home buyers/sellers

Kent County House Prices Data

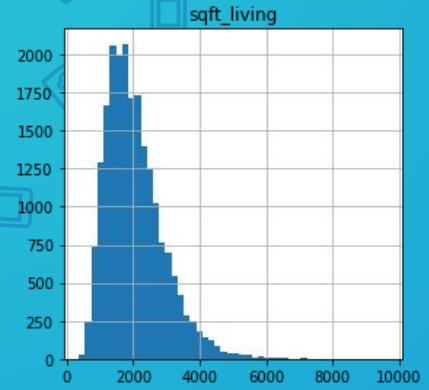
- Price
- Sale Date, Year Built, Renovation Date
- Number of bedrooms, bathrooms, floors
- Waterfront, view
- Condition, grade
- Sqft of living area, lot, basement, above and sqft of living area and lot for 15 nearest neighbors
- Zip Code, latitude, longitude

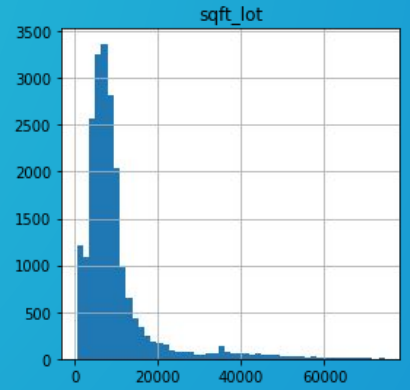
Price

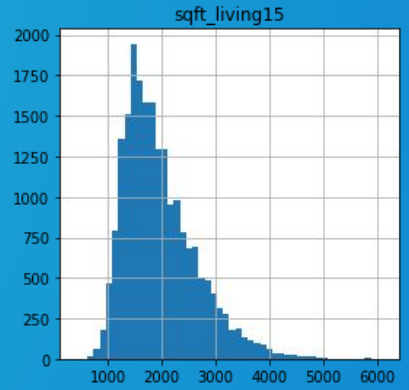


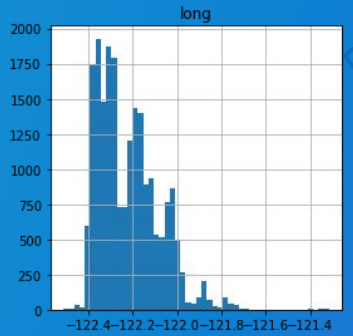
Min: $78k
Max: $7.7M
Avg: $540k

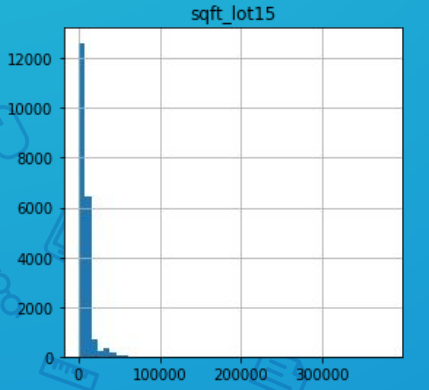
Continuous Variables

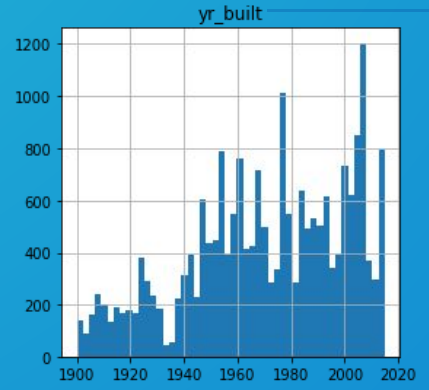


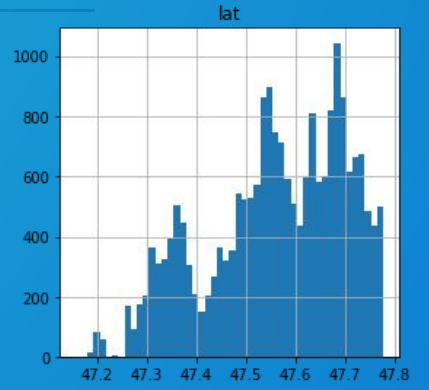




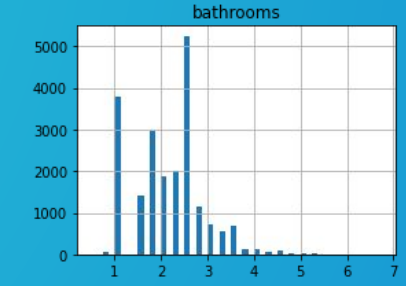


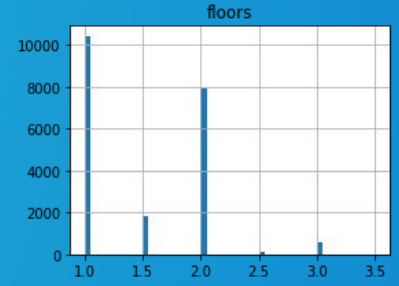


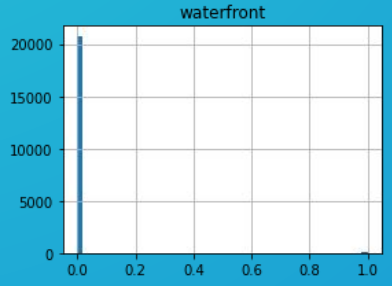


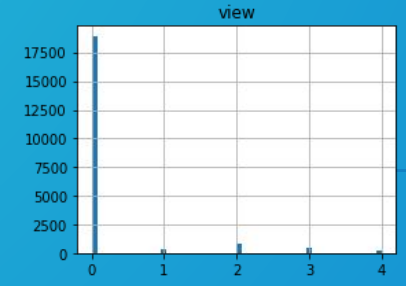


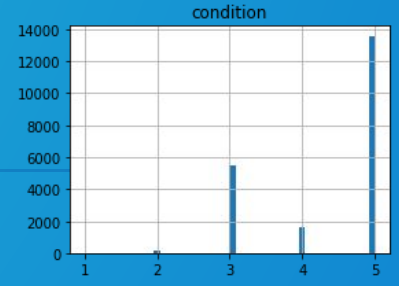
Discrete Variables

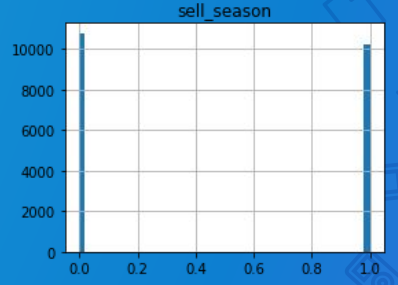


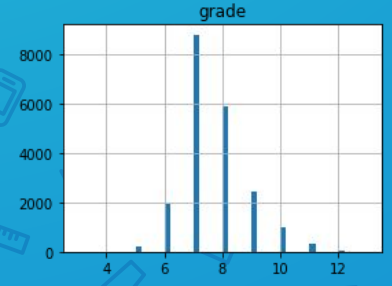


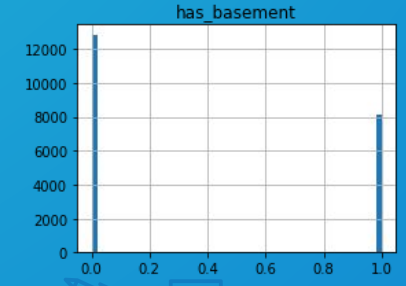


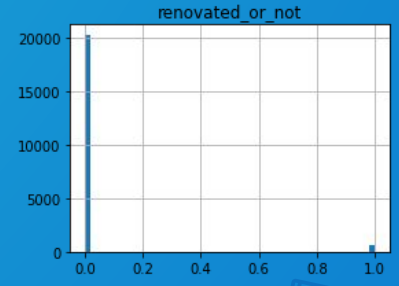




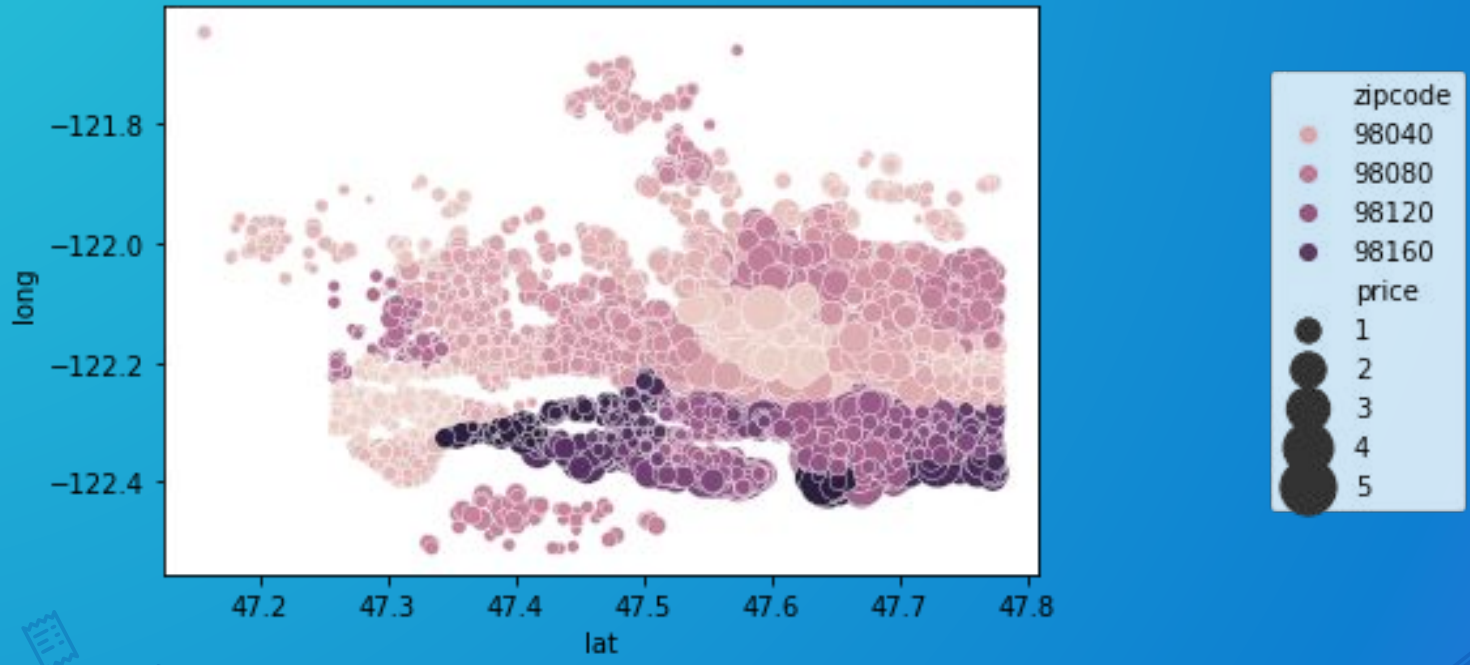








Zip Code


−121.8
−122.0
−122.2
−122.4
long
47.2
47.3
47.4
47.5
47.6
47.7
47.8
lat
zipcode
98040
98080
98120
98160
price
1
2
3
4
5

How can I help?

Buyer

Given the features of the house, is the price reasonable? Which features are most predictive of the price?

Seller

Will renovating my house increase the value of it?

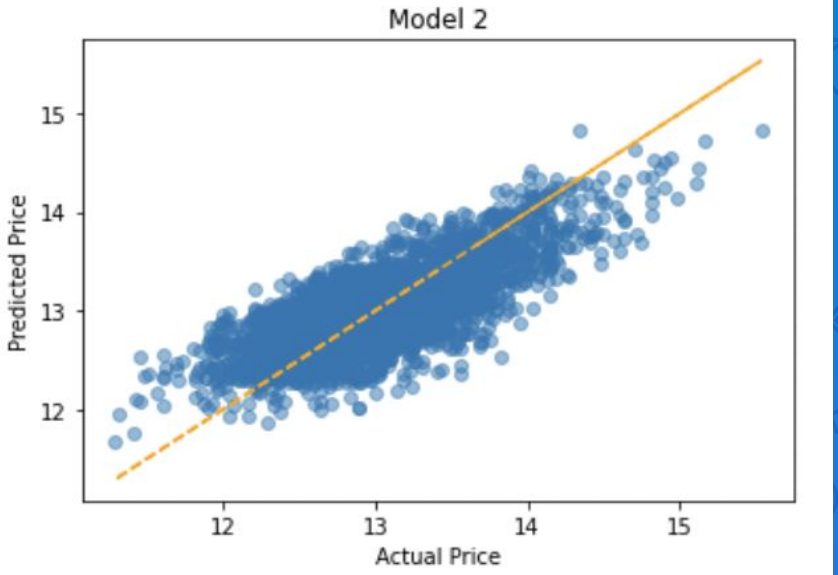
2.
Modeling and Regression Results

Modeling the house prices with given features

Model Performance

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

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



Model Performance

- Multiple Linear Regression
- Using features:
  - Grade
  - Living area
  - Has basement or not
  - On waterfront or not
  - Renovated or not
  - Number of floors

Price ~= 1.28*Grade +
1.15*Living area +
1.07*Has basement +
1.05*On waterfront +
1.04*Renovated +
1.008*Number of floors

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?

- Renovating the house will increase the log price of the house by 0.0423, which means the new price will be 1.04 times the old price. If the house is worth $100k, it will become $104k after renovation. Since the renovation will certainly cost more than $4k, I would not recommend renovations.

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/