King County Real Estate Consulting

Jake Sherman

Overview

- Provide local homeowners advice on how best to renovate their home.

- Additional square foot adds $96 to home value.

- Each unit of condition adds about $41,000 to home value.

Business Value

- Linear Regression informs clients how unit changes in certain independent variables will affect sales price.

	date	price	bedrooms	bathrooms	sqft_living	sqft_lot	floors	waterfront	condition	grade	sqft_above	sqft_basement
0	2014	221900.0	3	1.00	1180	5650	1.0	0	3	7	1180	0.0
1	2014	538000.0	3	2.25	2570	7242	2.0	0	3	7	2170	400.0
2	2015	180000.0	2	1.00	770	10000	1.0	0	3	6	770	0.0
3	2014	604000.0	4	3.00	1960	5000	1.0	0	5	7	1050	910.0
4	2015	510000.0	3	2.00	1680	8080	1.0	0	3	8	1680	0.0

Methodology

SOURCED DATA FROM KC
HOUSING DATASET

PERFORMED LINEAR
REGRESSION ON THE DATA SET

ITERATED PARAMETERS TO
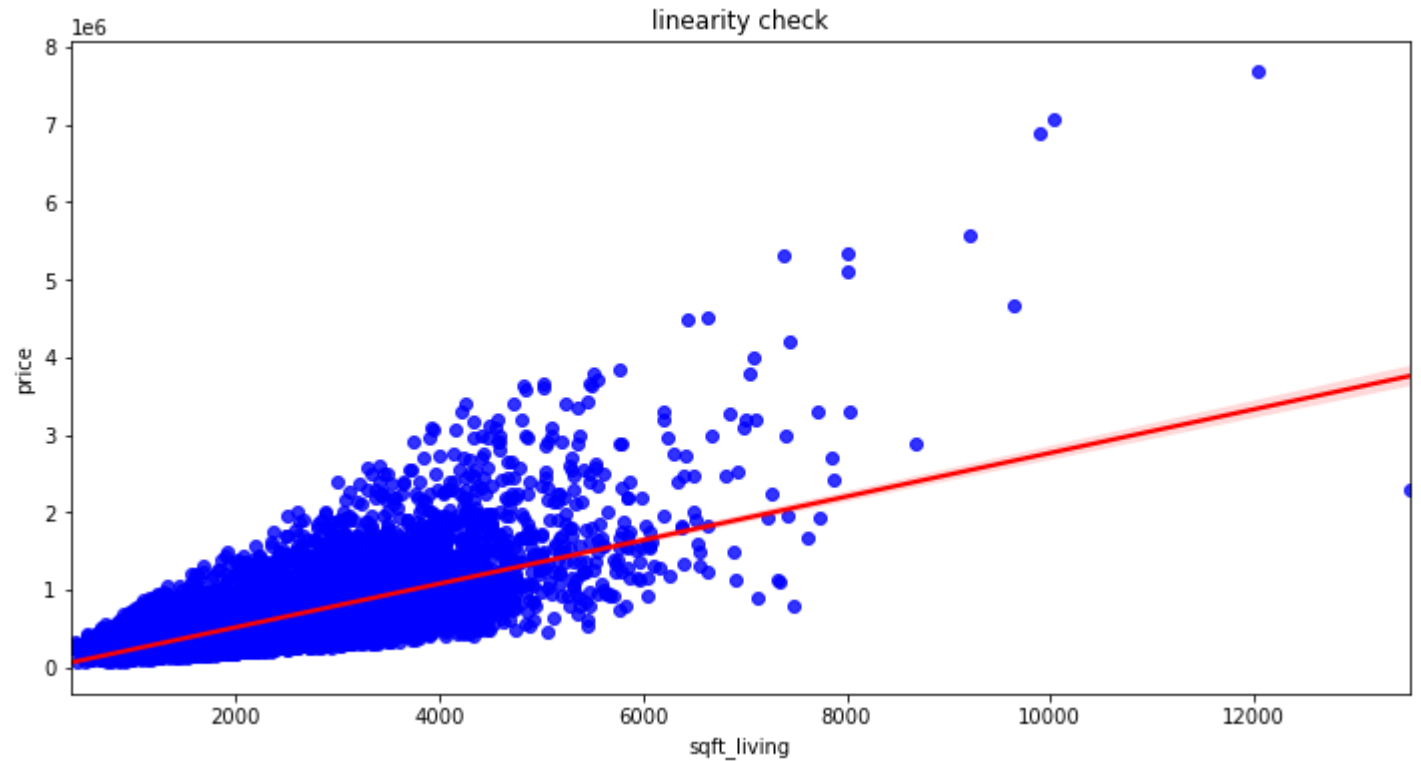OPTIMIZE RMSE

Correlation Breakdown

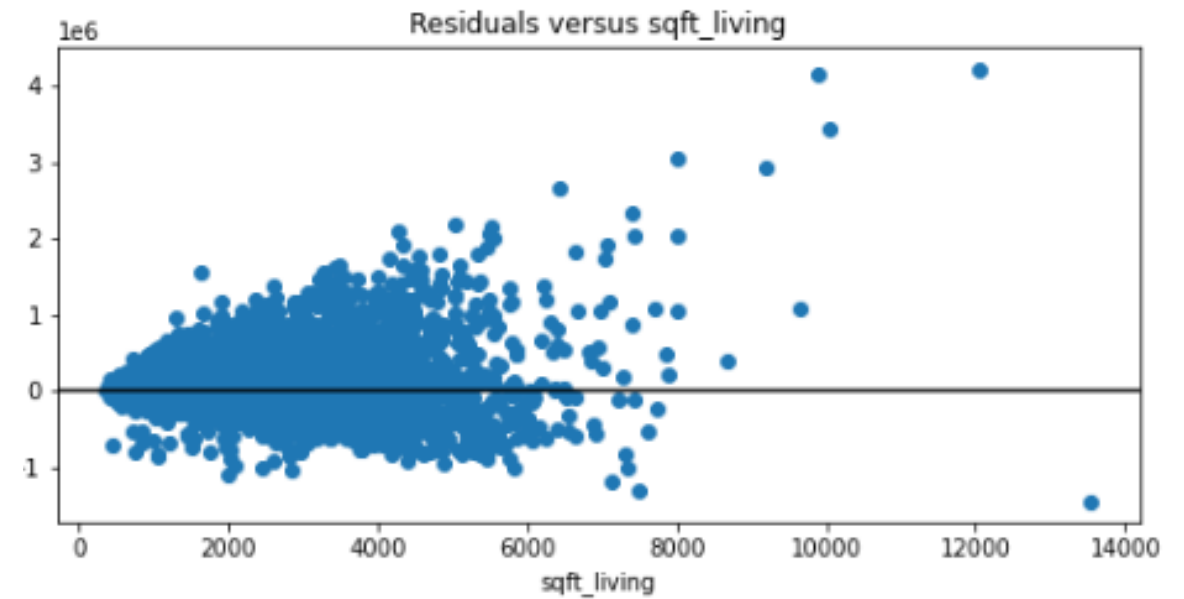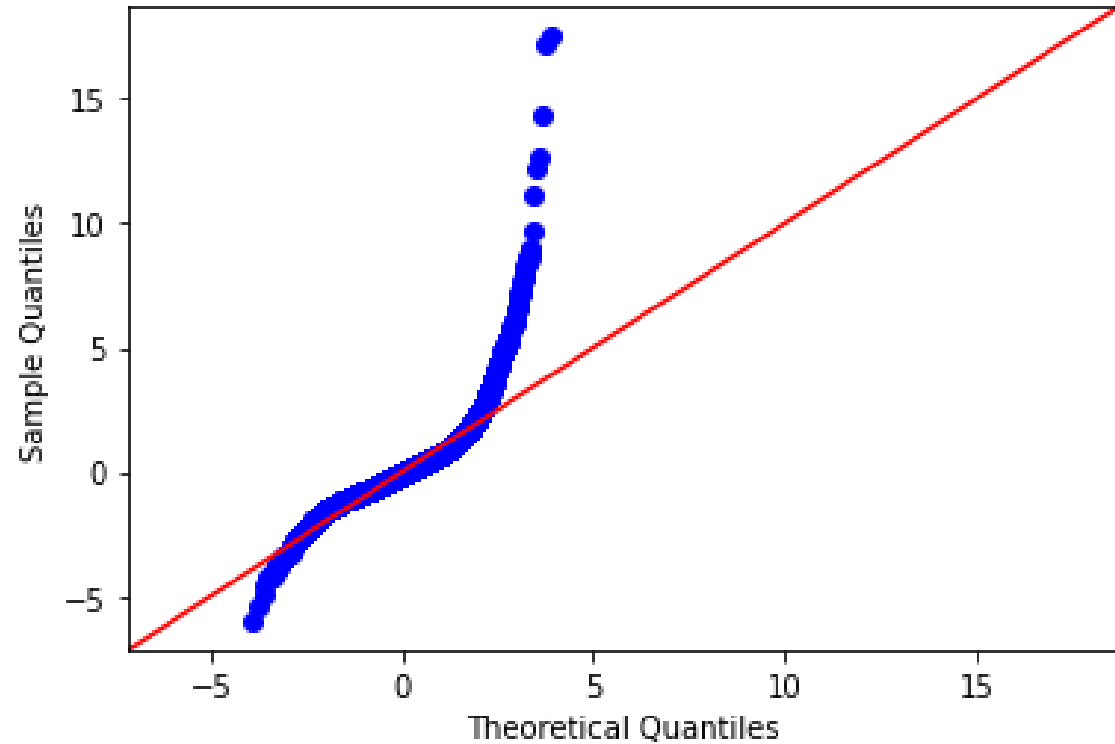NOT A LOT OF STRONG
CORRELATIONS WITH PRICE

SQUARE FOOTAGE OF LIVING
SPACE WAS THE STRONGEST AT .7

Baseline
Root Mean Squared Error and Linearity

- Initial RMSE: $241,000

- Average Price of home: $540,300

- R2 Score: .57

- Add regression line

linearity check

Baseline Assumptions

Iterations

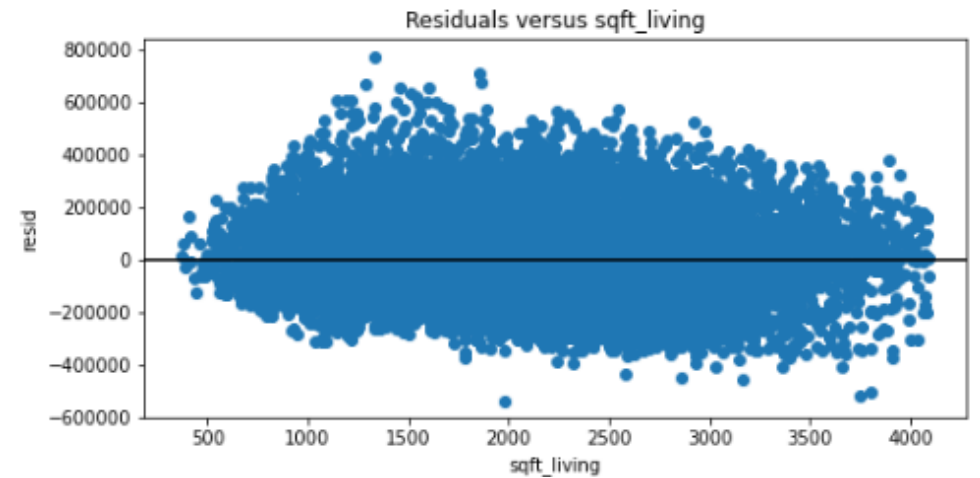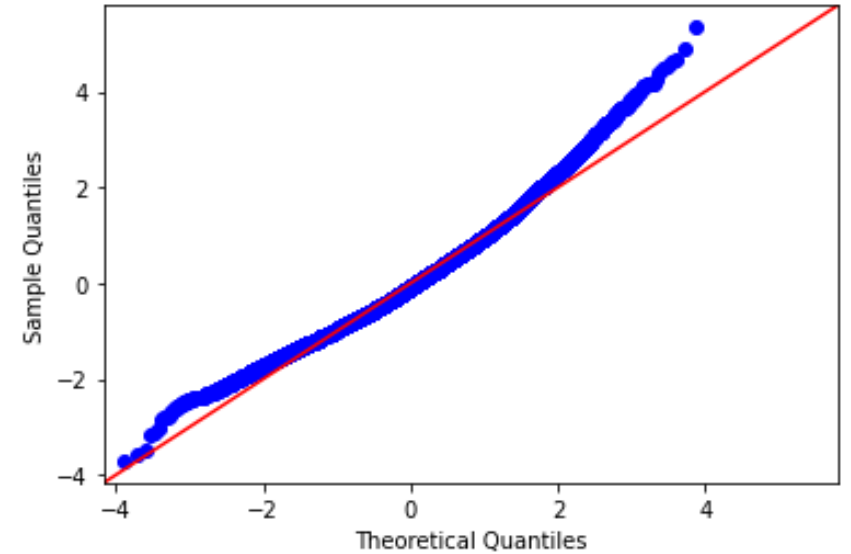| Remove	Outliers
Create	Categorical Variables
Scaling	Feature Scaling

Results

- Improved normality, and homoscedasticity

- No high p-values (above .05)

- RMSE is now down to **$144,361** from $241,000

- R2 score: .471

Residuals versus sqft_living

Conclusion

- Each addition Sqft= +$96.

- Each unit of increased condition = +$40,910.

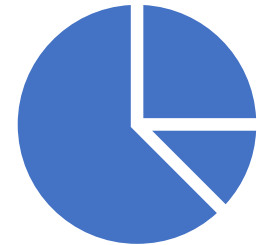- Full renovation from 1 to 5 =+$163,640.

Limitations

Data did not have an ideal linearity which affects predictability.
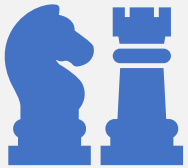
Different zip codes weren't accounted for.

Data such as view, condition, and grade are subjective.

Next steps

SEGMENT DATA TO FIND IMPROVE GOODNESS OF FIT (R2).

INCORPORATE MORE DATA AVAILABLE TO REFINE CORRELATIONS.

ANALYZE COSTS OF IMPROVEMENT TO MEASURE PROFITABILITY.

Thank You!

Github.com/jsherman918