



# Real estate Resale Needs Analysis

Rachael Bryant  
June 21, 2022



# Summary



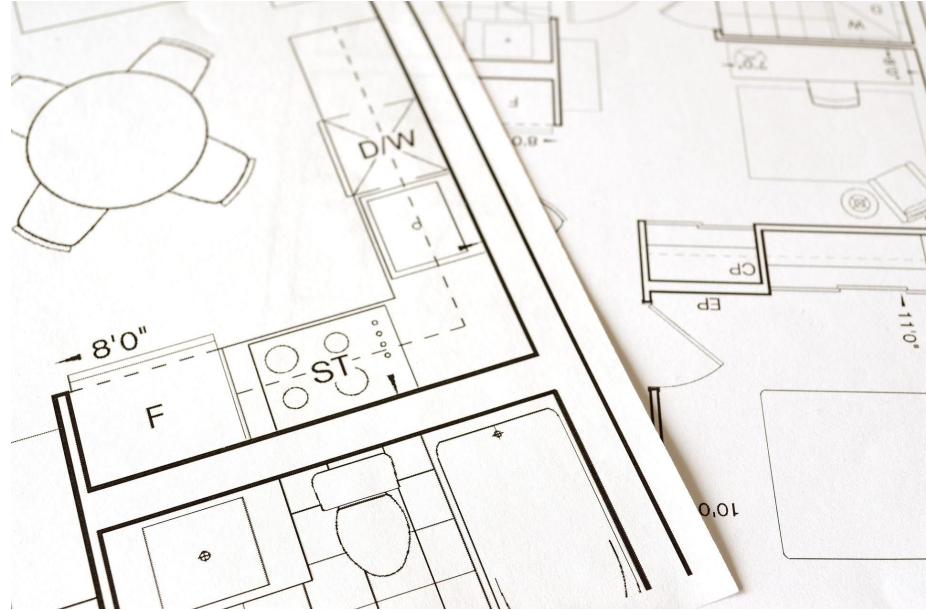
Descriptive and inferential analysis of the data from King County House Sales depicted correlations between aspects of house renovations that are more profitable.

# Outline

- Business Problem/Data
- Modeling
- Regression results
- Recommendations
- Next Steps
- Thank you

# Business Problem/Data

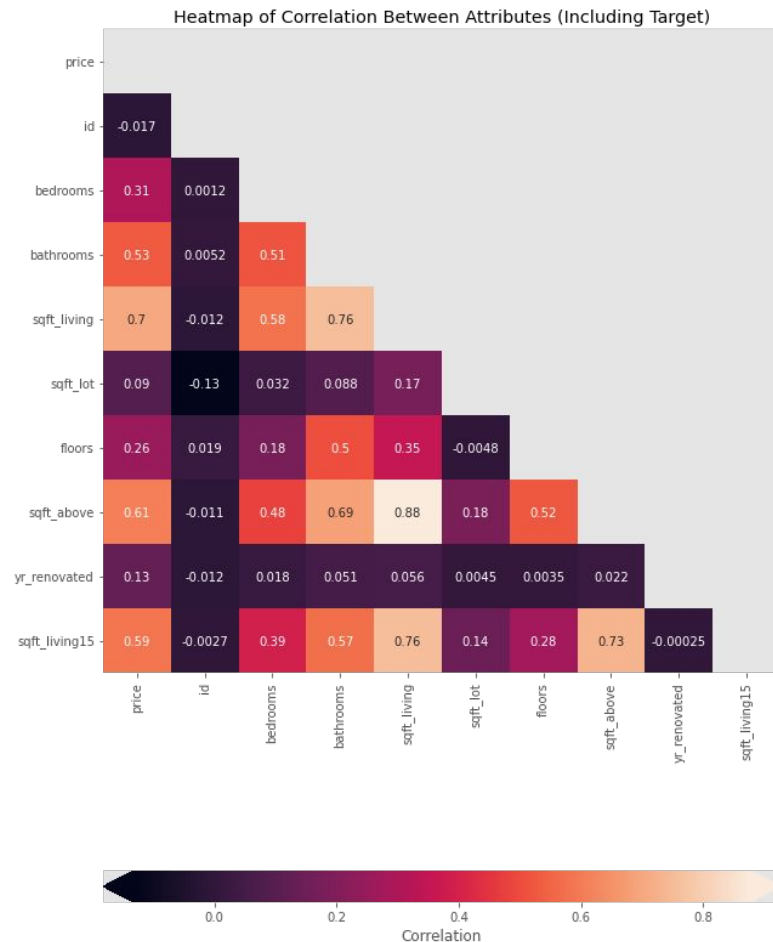
- What features of home renovations in King County lead to the most resale profit.
- The Data utilized is the King County House Sales dataset provided by csv file.



# Molding

My base model was founded upon finding out that the square feet of the living space was the most correlated with the prices of the houses.

I then tried two more models in order to improve the model and account for more variance.



# Regression Results

- The best model still only accounts for 52% variability observed in the target variable is explained by the regression model.
- We cannot have too much confidence in the models

## OLS Regression Results

```
=====
Dep. Variable:      price  R-squared:      0.552
Model:              OLS   Adj. R-squared:    0.552
Method:             Least Squares  F-statistic:    4437.
Date:               Mon, 20 Jun 2022  Prob (F-statistic):    0.00
Time:               07:41:03  Log-Likelihood:    -2.9872e+05
No. Observations:   21597  AIC:              5.974e+05
Df Residuals:       21590  BIC:              5.975e+05
Df Model:           6
Covariance Type:    nonrobust
=====
```

```
=====
              coef  std err          t  P>|t|  [0.025   0.975]
-----
Intercept  1.028e+05  9175.960   11.204   0.000   8.48e+04  1.21e+05
sqft_living 256.6538    1.900   135.083   0.000  252.930  260.378
cyl_YES     5.31e+05  2.49e+04   21.360   0.000   4.82e+05   5.8e+05
yr_EXCELLENT 2.878e+05  1.85e+04   15.583   0.000   2.52e+05  3.24e+05
yr_FAIR     4.909e+04  1.56e+04    3.152   0.002   1.86e+04  7.96e+04
yr_GOOD     8.156e+04  1.34e+04    6.095   0.000   5.53e+04  1.08e+05
yr_NONE    -1.189e+05  7989.878  -14.880   0.000  -1.35e+05  -1.03e+05
=====
```

```
=====
Omnibus:      13485.396  Durbin-Watson:      1.976
Prob(Omnibus): 0.000  Jarque-Bera (JB):    489867.377
Skew:         2.450  Prob(JB):              0.00
Kurtosis:     25.812  Cond. No.              3.75e+04
=====
```

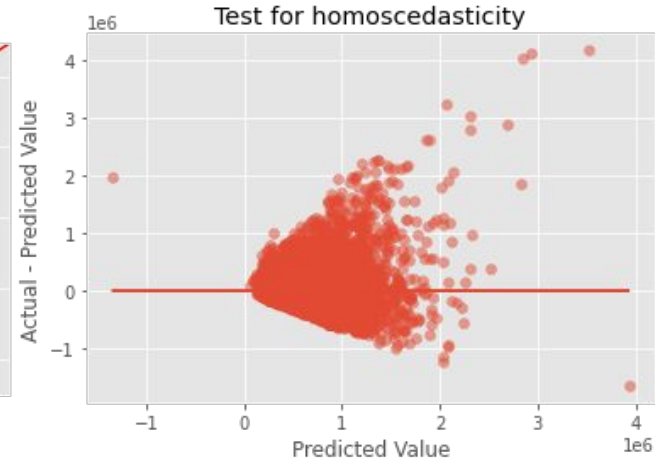
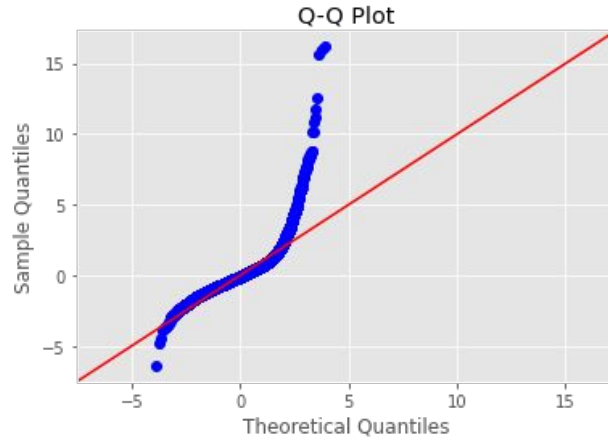
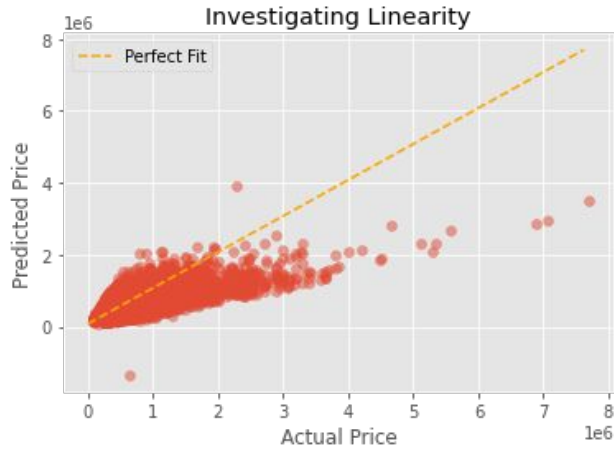
## Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The condition number is large, 3.75e+04. This might indicate that there are strong multicollinearity or other numerical problems.

# Regression Results

- The three tests of multicollinearity shown below also depict that the models are not reliable.



# Recommendations

This analysis leads to a small correlation between price and the aspects of the houses:

- I would tentatively advise the local real estate office that renovations for more sqft of space within the house and additions of bathrooms and bedrooms could gain more of a profit than other potential renovations.

## Next Steps

Further analyses could yield additional insights to further aid in the local real estate office at King County:

- Perhaps collecting a wider pool of data would add in adding variance and leading to a stronger correlation between the values as well as looking into other aspects of what renovations entail, such as the expense for particular jobs, trends in home making magazines and many other aspects that could lead to different renovation recommendations for the most profit in selling the houses.



# Thank You!

Email: [rachaelbryant94@gmail.com](mailto:rachaelbryant94@gmail.com)

GitHub: @rabrya0072

LinkedIn: [linkedin.com/in/rachaelbryant/](https://www.linkedin.com/in/rachaelbryant/)