HOUSING PRICE DETERMINANTS.

THE CASE OF HOUSING DATA IN KING COUNTY, WAHINGTON STATE, USA.

Overview

- IN THIS PRESENTATION I WILL BE INVESTIGATING THE DETERMINANTS OF HOUSING PRICES/ RELATIONSHIPS USING DATA FROM KING COUNTY.
- WE SHALL DEMONSTRATE RELATIONSHIPS USING LINEAR REGRESSION AND MAKE PREDICTIONS ABOUT PRICE FROM THE DATA.
- WELCOME TO THE STATISTICAL TRIP…
- MAY THE ODDS BE IN YOUR FAVOUR…

Business Understanding

- AS A REAL ESTATE BUSINESS, IT IS IMPORTANT TO UNDERSTAND MARKET FACTORS AND IN ORDER TO MAKE BETTER DECISIONS AND TO STAY AFLOAT.
- HOUSE PRICES CAN BE AFFECTED BY A NUMBER OF FACTORS FROM QUALITY, QUANTITY AND LOCATION.
- ONCE WE UNDERSTAND HOW STRONG THE RELATIONSHIP IS, WE CAN BE ABLE TO PREDICT TRENDS AND KNOW HOW TO PRIORITIZE INVESTMENT SECTORS.

Data Understanding

- THE DATA COLLECTED HAS ENOUGH FACTORS THAT MIGHT INFLUNCE PRICE SUCH AS :
- SIZE OF PROPERTY, LOCATION OF THE PROPERTY, AMENITIES, QUALITY OF CONSTRUCTION, AGE AND STATE OF THE PROPERTY.

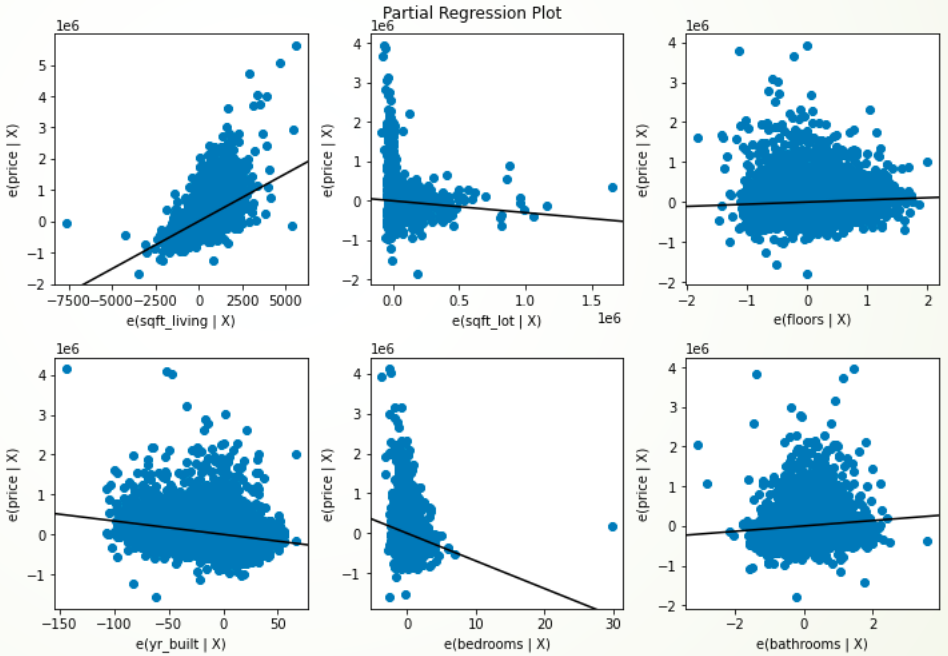
Modeling

- I used Linear Regression modelting and foun the base relationship to be as follows:
- * The model is statistically significant overall, with an F-statistic p-value well below 0.05 =
- * The model explains about 49.3% of the variance in price
- * The model coefficients (`const` and `sqft_living`) are both statistically significant, with t-statistic p-values well below 0.05
- * If a house had 0 sqft living we would expect price to be about -43988.89
- * An increase of 1 square foot living space, would lead to an increase of about 280.86 in price

Regression Results

- From the model, we can see that the regression line we found was
- $price_{est}$ = 6559487.332 + 303.767sqft living – 0.302sqft_lot + 54213.630floor −3367.618 yr_built + −68999.961bedrooms + 67469.023 bathrooms
- The model is statistically significant overall, with an F-statistic p-value well below 0.05 =
- The model explains about 49.3% of the variance in price
- The model coefficients (`const', 'sqft_living', 'sqft_lot','floors','yr_built','bedrooms', 'bathrooms'd `) are all statistically significant, with t-statistic p-values well below 0.05
- If a house is had 0 X we would expect price to be about 6559487.332
- An increase of 1 in X, would lead to an increase of about 6559487.332 in price

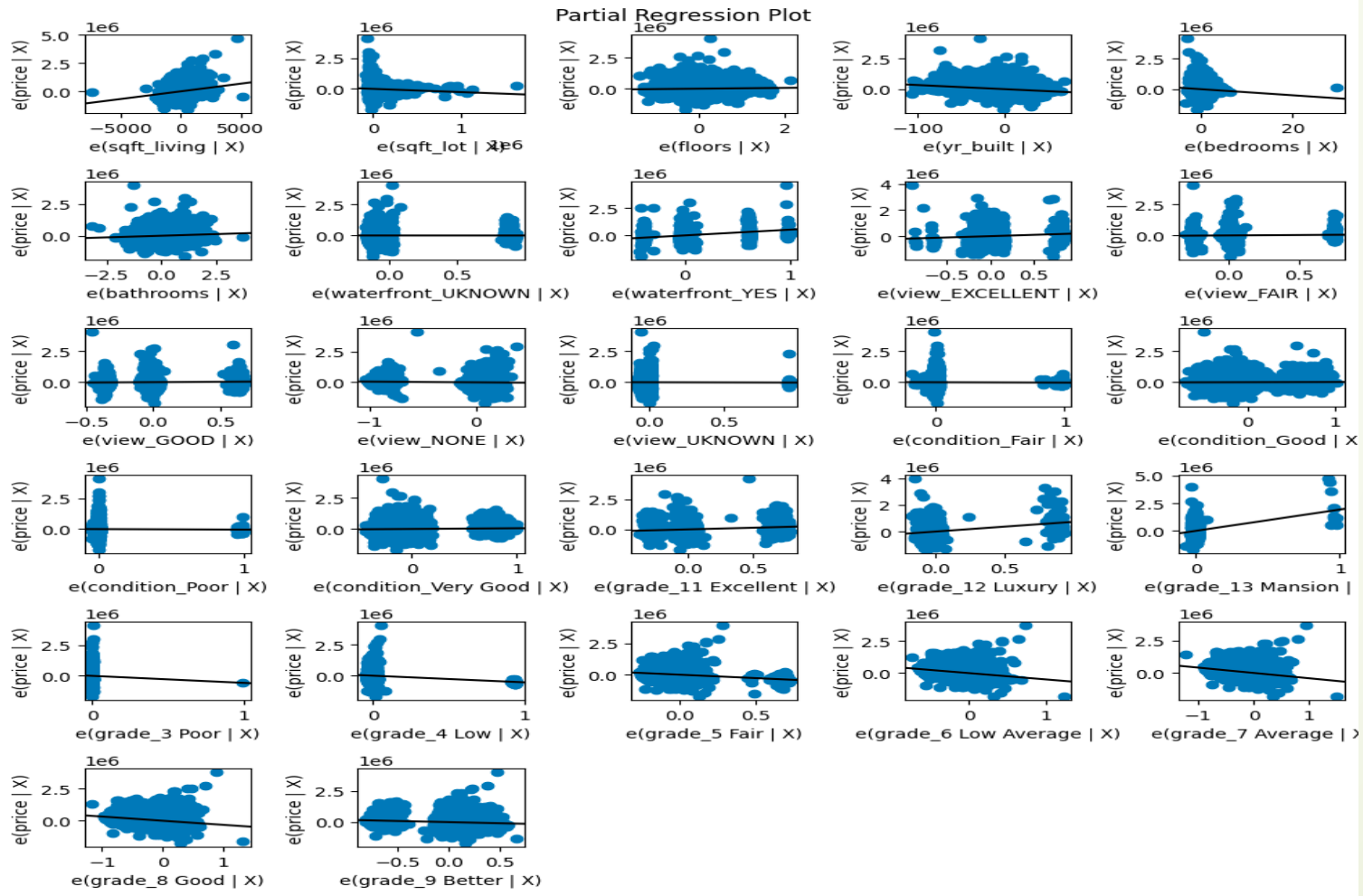
Regression Results



RESULTS

- OLS Regression Results
==========
===== Dep. Variable: price R-squared: 0.277 Model: OLS Adj. R-squared: 0.277 Method: Least Squares F-statistic: 8256. Date: Fri, 30 Sep 2022 Prob (F-statistic): 0.00 Time: 21:03:18 Log-Likelihood: -3.0389e+05 No. Observations: 21597 AIC: 6.078e+05 Df Residuals: 21595 BIC: 6.078e+05 Df Model: 1 Covariance Type: nonrobust
==========
===== coef std err t P>|t| [0.025 0.975] ---------- ---------- const 8712.6518 6224.769 1.400 0.162 -3488.356 2.09e+04 bathrooms 2.512e+05 2765.054 90.863 0.000 2.46e+05 2.57e+05
==========
===== Omnibus: 17251.570 Durbin-Watson: 1.960 Prob(Omnibus): 0.000 Jarque-Bera (JB): 882735.889 Skew: 3.452 Prob(JB): 0.00 Kurtosis: 33.550 Cond. No. 7.76
==========
===== Notes: [1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Regression Results.



Recommendations.

- * If all other factors were 0, price would be 7071780.6
- * Luxury and mansion categories affect the price more.
- * Houses with very good conditions are more likely to increase price
- * Having a Waterfront property is statistically more significant than not having at alfa = 0.05
- * the number of floors and bathrooms has a great effect in the price

Next Steps

- If all other factors were 0, price would be 7071780.6
- * Luxury and mansion categories affect the price more.
- * Houses with very good conditions are more likely to increase price
- * Having a Waterfront property is statistically more significant than not having at alfa = 0.05
- * the number of floors and bathrooms has a great effect in the price

Thank you for your time…