Important Considerations

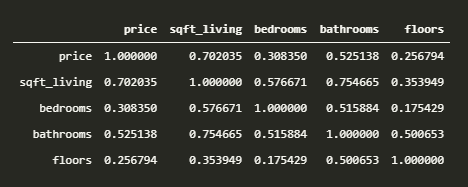
There are a few critical assumptions about your data set that must be true to proceed with a regression analysis:

1. The variables must be truly independent (using a Chi-square test).

2. The data must not have different error variances (this is called heteroskedasticity (also spelled heteroscedasticity)).

3. The error terms of each variable must be uncorrelated. If not, it means the variables are serially correlated.

Correlation of dependent and independent Variable:



Linear Regression model for **price** and **sqft\_living**

1. Get most useful descriptive statistics for each column in the data frame – number of observations, mean, standard deviation, and so on.

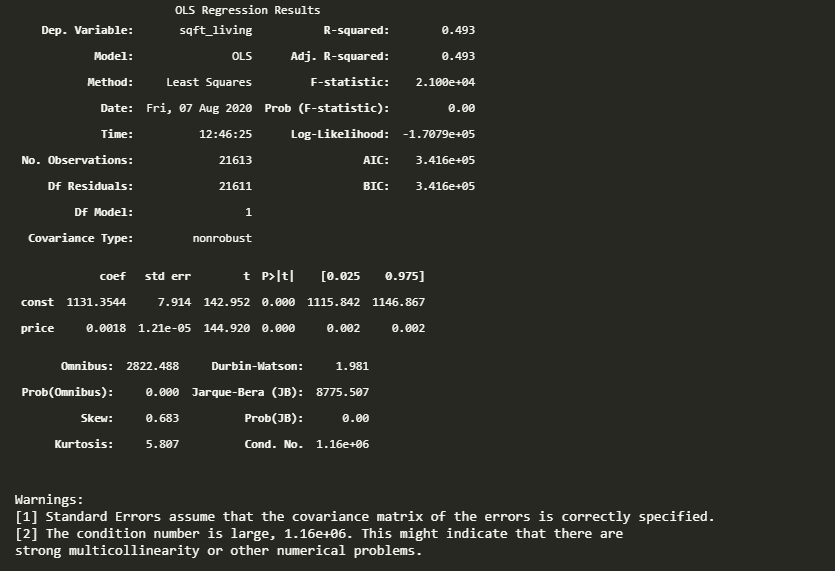


1. ordinary least squares regression(OLS)

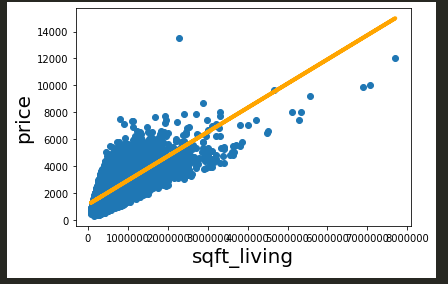
According to linear regression equation:

ŷ = β0+ β1\* x

y = 0.0018 \*df\_x(independent variable) + 1131.35



1. The R-Square value is... 0.5014186290572902
2. The RMSE value is........ 645.8415059351391
3. Plotting the Regression line

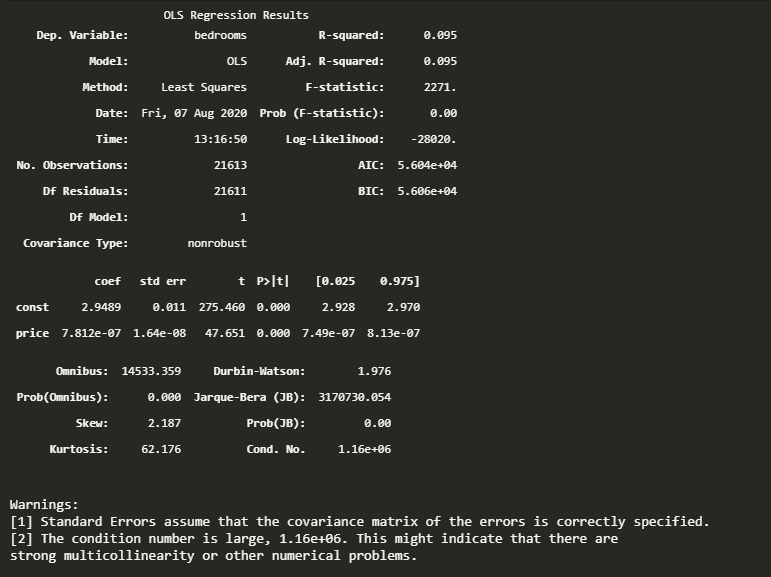


Linear Regression model for **price** and **bedrooms:**

1. ordinary least squares regression(OLS)

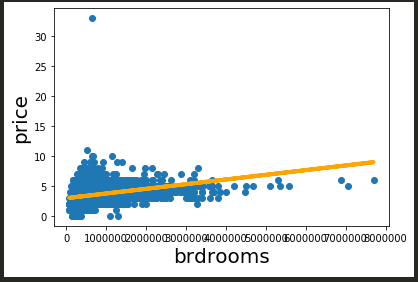
According to linear regression equation:

ŷ = β0+ β1\* x

 y = (7.812e-07\*df\_x) + 2.9489

2. The R-Square value is... 0.08655218927185626

3. The RMSE value is........ 0.8661949294458479

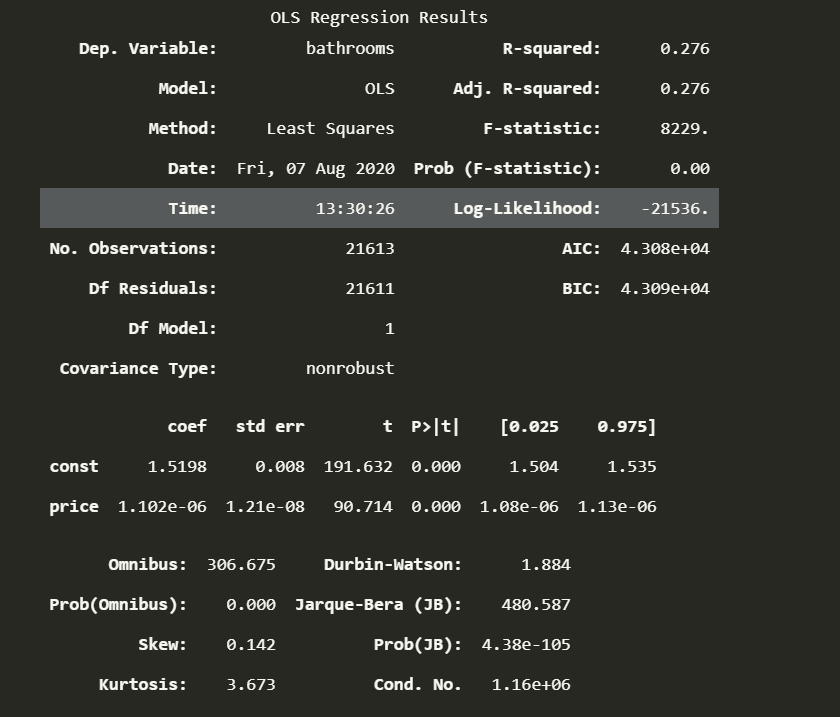
4. Plotting the Regression line

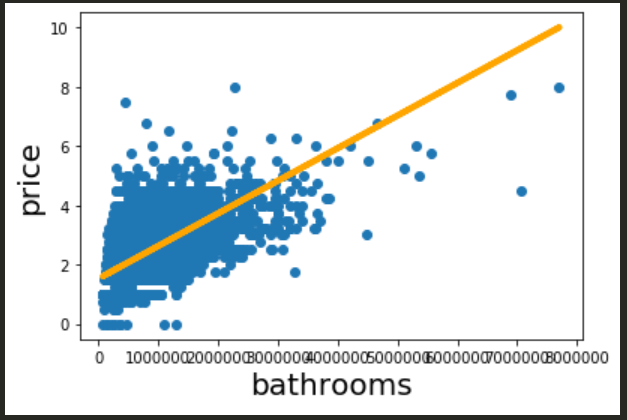
Linear Regression model for **price** and **bathrooms:**

1. Ordinary least squares regression(OLS)

According to linear regression equation:

ŷ = β0+ β1\* x

 y = (1.102e-06\*df\_x) + 1.5198

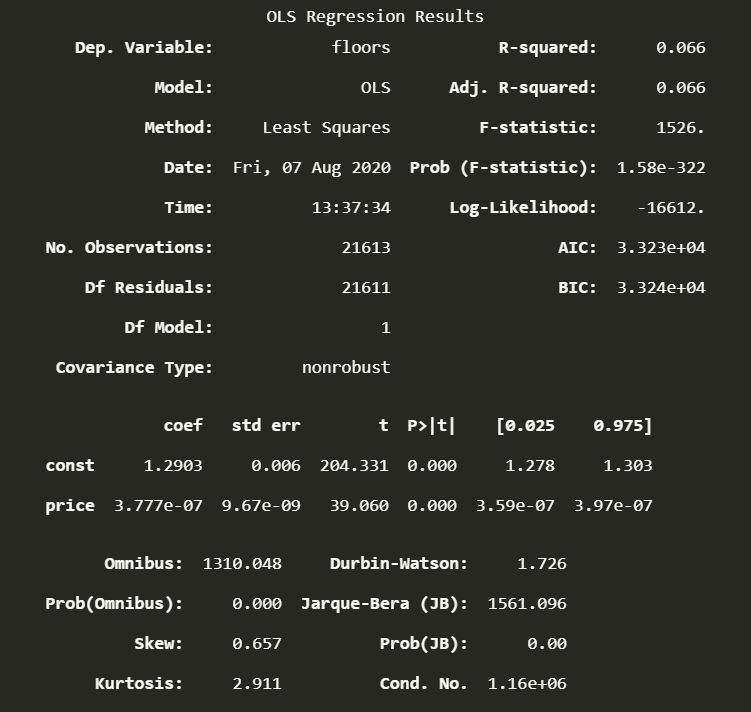
1. The R-Square value is... 0.28669446818708044
2. The RMSE value is........ 0.6504653050644241
3. Plotting the Regression line

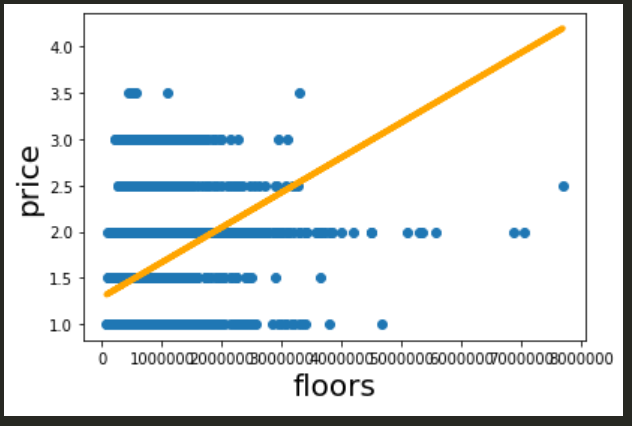
Linear Regression model for **price** and **floors:**

1. Ordinary least squares regression(OLS)

According to linear regression equation:

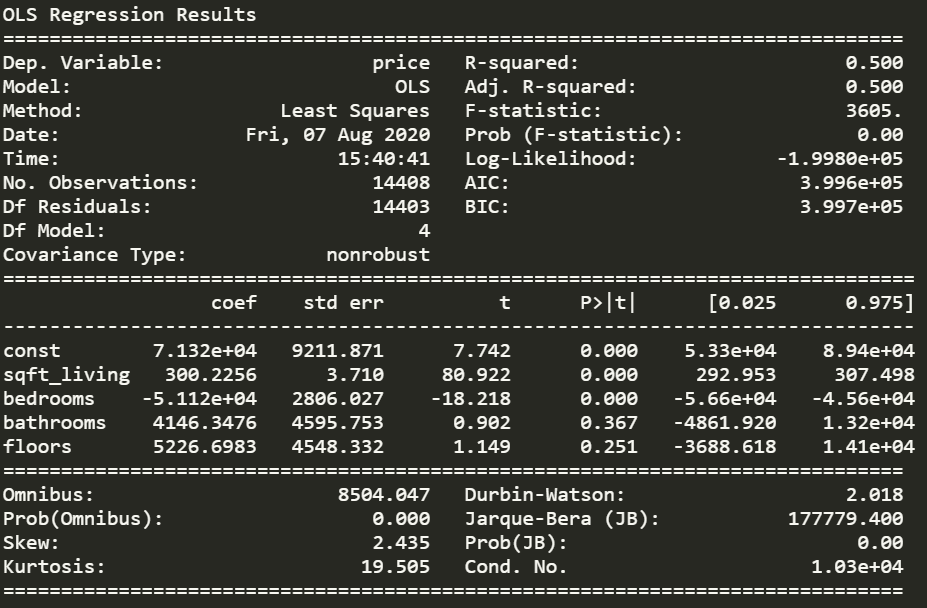
ŷ = β0+ β1\* x

 y= (3.777e-07\*df\_x) + 1.2903

1. The R-Square value is... 0.06472556030115528
2. The RMSE value is........ 0.5181200272227835
3. Plotting the Regression line

Multiple Linear regression :

OLS regression result:



According to the r-square value **sqft\_living** variable are more significant