Model

$$\ln(price) = 10.2082 + 0.3618 \cdot \text{waterfront} - 0.0160 \cdot \text{bedrooms} - 0.0153 \cdot \text{bathrooms} + 0.1400 \cdot \text{sqft\_living}^{0.3}$$
$$+ 0.0088 \cdot \text{floors} + 0.1494 \cdot \text{view}^{0.5} + 0.0105 \cdot \text{grade}^2 + 0.1187 \cdot \ln(\text{sqft\_living15})$$

# Check Statistical Hypotheses of the Regression

Linearity:

The Null Hypothesis: The model is linearly predicted by the feature,
The Alternative Hypothesis: The model is not linearly predicted by the feature.
Our p-value for this model is $p = 0.933 > 0.05 = \alpha$. Thus, we don't have enough evidence to reject **The Null Hypothesis** and we conclude that our model satisfies Linearity Assumption.

Normality Assumption for Errors

To check Normality, I used the following checks:
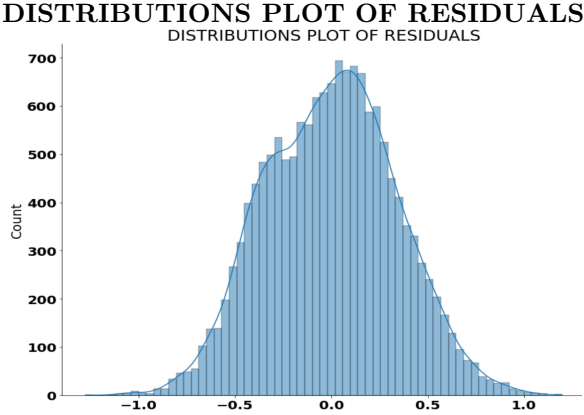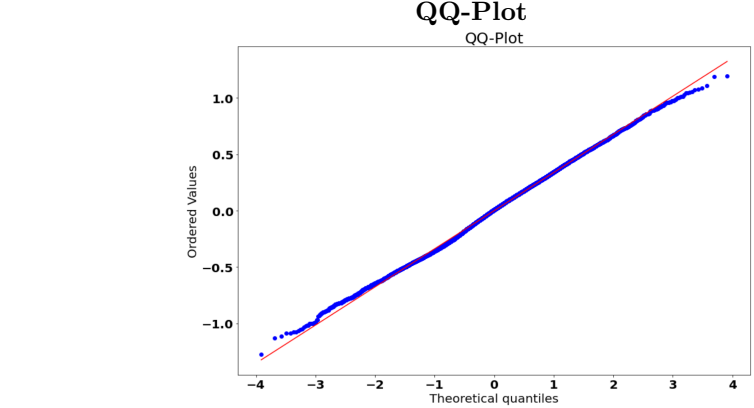


I also, used D'Agostino Test for Normality:
The Null Hypothesis: The Residuals are normally distributed,
The Alternative Hypothesis: The Residuals are not normally distributed.
Our p-value for this model is $p = 0.000 < 0.05 = \alpha$. Thus, we have enough evidence to reject the Null Hypothesis and conclude that D'Agostino Test tells us, that residuals are not normally distributed.

Conclusion: Based on QQ-Plot, Distributions Plot, and D'Agostino Test, I conclude that the Distribution of Errors is not far away from Normal. Also, since we have a lot of observations Normality Assumption doesn't play a critical role, since Central Limit Theorem will apply in this case.

Constant Error Variance

To if heteroscedasticity is present in the model, I will use Residual-vs-Predicted values plot and Breusch-Pagan test. I look at at the Residual-vs-Predicted values plot first.
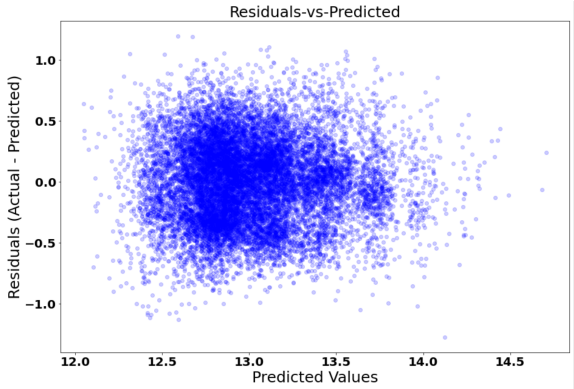
I used Breusch-Pagan Test:
The Null Hypothesis: Homoscedasticity is present,
The Alternative Hypothesis: Homoscedasticity is not present (i.e. heteroscedasticity exists).
Our p-value for this model is $p = 0.000 < 05 = \alpha$. Thus, we have enough evidence to reject the Null Hypothesis and we conclude from Breusch-Pagan Test, that we have don't have heteroscedasticity.
Conclusion: From the Residual-vs-Predicted values plot and Breusch-Pagan Test, I conclude that we have some Heteroscedasticity in our model, but it is not very bad.

Overall Conclusion:

I conclude that our model almost satisfies statistical assumptions for the regression model.