

# Check Statistical Hypotheses of the Regression

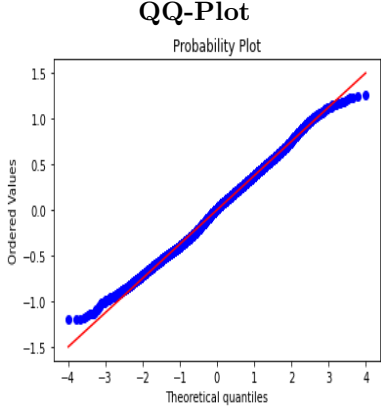
## Linearity:

**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  $p = 0.877 > 0.05 = \alpha$ . 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:

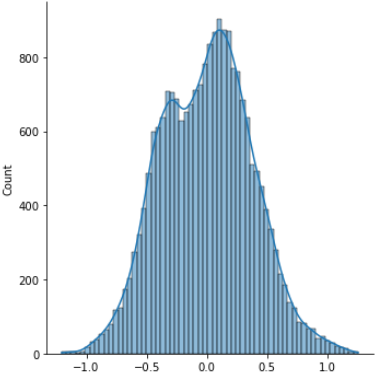


I will, also, use D'Agostino Test for Normality:

**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.

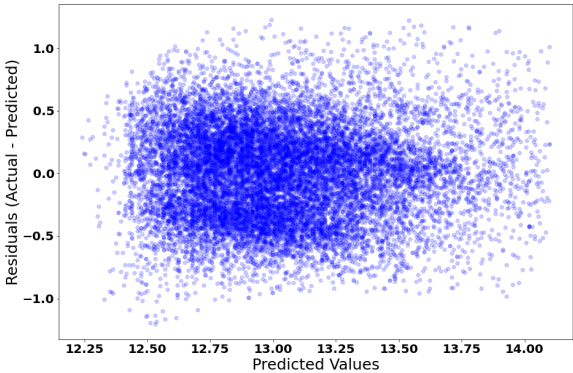
DISTRIBUTIONS PLOT OF RESIDUALS



**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.



Now, I will use 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  $p = 0.12609 \geq 0.05 = \alpha$ . Thus, we don't have enough evidence to reject the Null Hypothesis and we conclude that Breusch-Pagan Test tells us, that we have don't have heteroscedasticity.

**Conclusion:** From the Residual-vs-Predicted values plot and Breusch-Pagan Test, I conclude that we don't have Heteroscedasticity in our model.

## Overall Conclusion:

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