

Coefficient of Determination

The coefficient of determination, R-square, is a measure of the proportion of variability in the response variables explained by the predictor variables in the analysis. It's calculated as the model sum of squares divided by the total sum of squares.

$$R^2 = \frac{SS_M}{SS_T}$$

R-square is bounded between 0 and 1. The value is close to 0 if the predictor variables do not explain much variability in the data, and close to 1 if the predictor variables explain a relatively large proportion of variability in the data. Thus, it's ideal that the variability, due to differences between the groups (or the explained variability), makes up a larger proportion of the total variability than the random error within the groups (or the unexplained variability). This indicates that the model is a relatively good fit to the data.

In this case, R-square=0.1579 indicates that the four levels of heating quality explain 15.79% of the variability in the response variable SalePrice. The other 84.21% represents unexplained variability, or residual variation. Although values of R-square closer to 1 are preferred, judging the magnitude of R-square depends on the context of the problem.