The Simple Linear Regression Model

In simple linear regression, the goal is to identify the equation that characterizes the linear association between the predictor variable and the response variable, and use the model to then estimate the response for a given value of the predictor.

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

In the linear regression equation, Y is the response variable, which, in this case, is SalePrice. X is the predictor variable, which is Lot_Area. β_0 is the intercept parameter, which corresponds to the value of the response variable when the predictor is 0. In other words, this is where the regression line crosses the Y axis. β_1 is the slope parameter, which is the average change in Y for a 1-unit change in X. ε is the error term that represents the variation of Y around the line.

The regression line is the expected value, or mean of Y (at any given X), which equals $\beta_0 + \beta_1 X$.

The intercept is often of less interest than the slope. Sometimes, the intercept corresponds to an impossibility. For example, in a regression of height on weight, the intercept would indicate the height of someone who weighs nothing. Other times, when X=0 is a possible value, it can be outside the range of actual data points. Therefore, be cautious when you interpret the regression relationship outside the range of your data.

Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression

Close