8.8: Chapter Key Terms

Key Term	Definition
a is the symbol for the Y-Intercept	Sometimes written as b_0, because when writing the theoretical linear model β_0 is used to represent a coefficient for a population.
b is the symbol for Slope	The word coefficient will be used regularly for the slope, because it is a number that will always be next to the letter "x." It will be written as b_1 when a sample is used, and β_1 will be used with a population or when writing the theoretical linear model.
Bivariate	two variables are present in the model where one is the "cause" or independent variable and the other is the "effect" of dependent variable.
Linear	a model that takes data and regresses it into a straight line equation.
Multivariate	a system or model where more than one independent variable is being used to predict an outcome. There can only ever be one dependent variable, but there is no limit to the number of independent variables.
R^2 – Coefficient of Determination	This is a number between 0 and 1 that represents the percentage variation of the dependent variable that can be explained by the variation in the independent variable. Sometimes calculated by the equation $R^2 = \frac{SSR}{SST}$ where SSR is the "Sum of Squares Regression" and SST is the "Sum of Squares Total." The appropriate coefficient of determination to be reported should always be adjusted for degrees of freedom first.
Residual or "error"	the value calculated from subtracting $y_0 - \hat{y}_0 = e_0$. The absolute value of a residual measures the vertical distance between the actual value of y and the estimated value of y that appears on the best-fit line.
R – Correlation Coefficient	A number between −1 and 1 that represents the strength and direction of the relationship between "X" and "Y." The value for "r" will equal 1 or −1 only if all the plotted points form a perfectly straight line.
Sum of Squared Errors (SSE)	the calculated value from adding up all the squared residual terms. The hope is that this value is very small when creating a model.
X – the independent variable	This will sometimes be referred to as the "predictor" variable, because these values were measured in order to determine what possible outcomes could be predicted.
Y – the dependent variable	Also, using the letter "y" represents actual values while $\hat{y}$ represents predicted or estimated values. Predicted values will come from plugging in observed "x" values into a linear model.