General Assembly

STUDY GUIDE

CORRELATION AND COVARIANCE

Key Terms

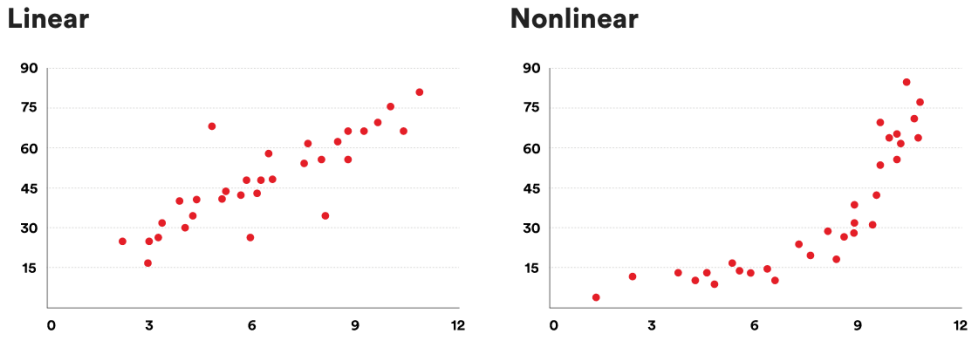
Relationship: When one variable may be influencing another. We can never use the word "cause" in statistics.

Cheat Sheet

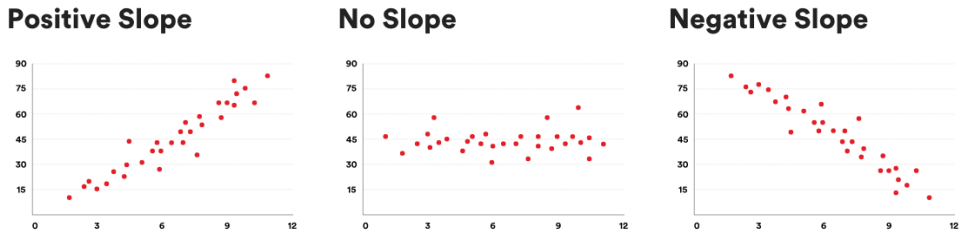
1. Correlation and Covariance

	Correlation Coefficient	**Coefficient of Determination**	**Covariance**
Designation	r	r^2	COV(X,Y)
Possible Values	-1 to 1	0 to 1	$-\infty$ to $+\infty$
Description	The strength and direction of a relationship	How much of the variance in y can be explained by the X variable	How the variables "move" together
Excel Function	CORREL()	Square the result of CORREL()	COVARIANCE.P() or COVARIANCE.S()

2. Understanding Scatterplots
    - » Scatterplots are a great way to visualize the relationship between two quantitative variables.
    - » There are few patterns you can spot within a scatterplot. These patterns are described in terms of:
    - » **Linearity**: Whether a data pattern is linear (straight) or nonlinear (curved).



» **Slope**: The direction of change in variable Y when variable X increases. If variable Y also increases, then the slope is positive; but if variable Y decreases, the slope is negative.



» **Strength**: The degree of "scatter" in the plot. If the dots are widely spread, the relationship between variables is weak. If the dots are concentrated around a line, the relationship is strong.