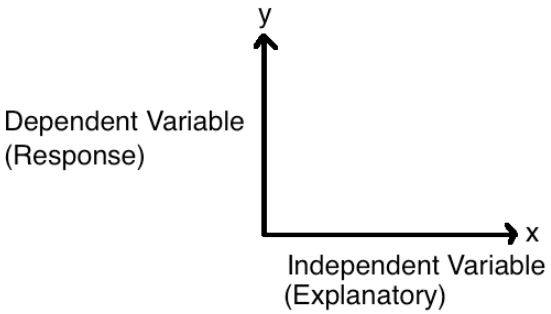
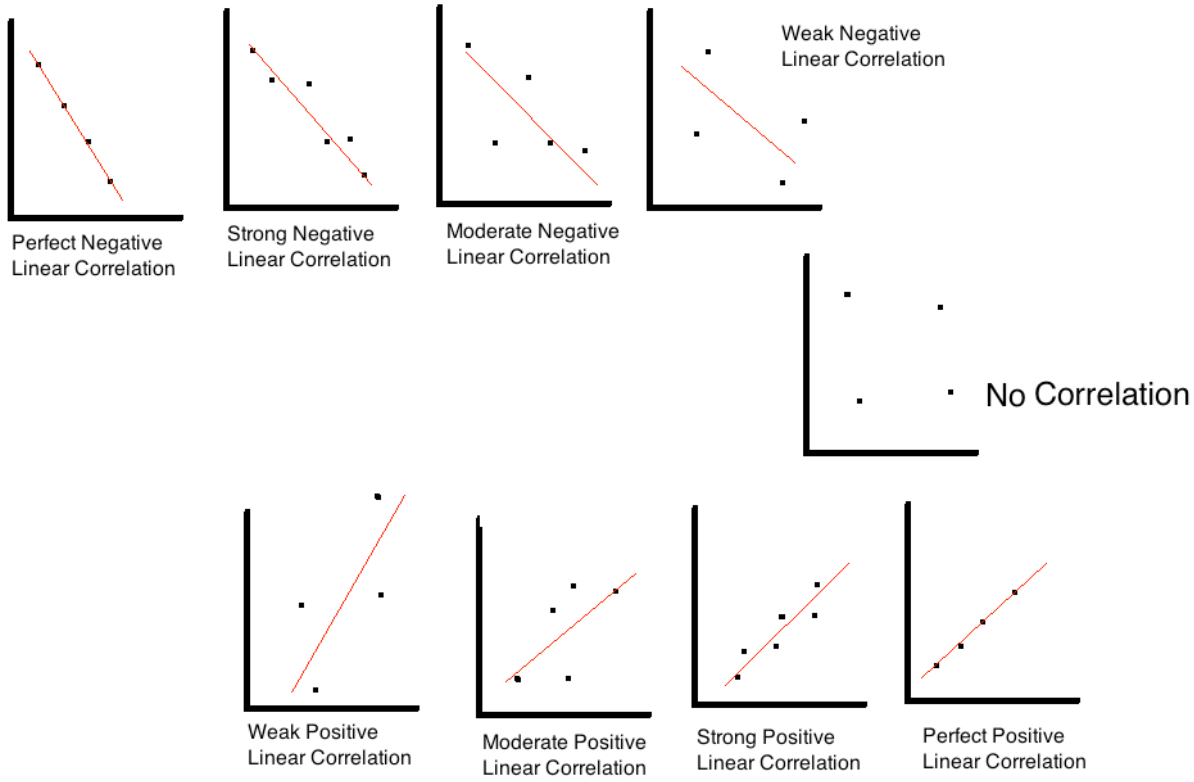
Lesson: Linear Correlation and Regression

Scatter Plot
- graph that describes relationship between two variables



Line of Best Fit
- passes as close as possible to all the points on a scatter plot and represents the relationship between two variables

Linear Correlation
- changes in one variable that are proportional to changes in another
- weak, moderate, or strong (strength of the relationship)
- positive or negative (direction of the relationship)



Coefficient of Correlation
- the strength between the two variables

x	y
$(-2)^2$	24
$+(-1)^2$	35
$+(0)^2$	72
$+(1)^2$	48
$+(2)^2$	96
Σx^2	

$n = 5$

$-1 \le r \le 1$

$\Sigma xy = (-2)(24) + (-1)(35) + (0)(72) + (1)(48) + (2)(96)$

total # of data

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2]\,[n\sum y^2 - (\sum y)^2]}}$$

Coefficient Correlation

Negative ⟷ Positive

Strong	Moderate	Weak	Weak	Moderate	Strong
$-1 < r < -\frac{2}{3}$	$-\frac{2}{3} \le r < -\frac{1}{3}$	$-\frac{1}{3} \le r < 0$	$0 < r \le \frac{1}{3}$	$\frac{1}{3} < r \le \frac{2}{3}$	$\frac{2}{3} < r < 1$

-1 (perfect) | $-\frac{2}{3}$ (-0.67) | $-\frac{1}{3}$ (-0.33) | 0 (No Correlation) | $\frac{1}{3}$ (0.33) | $\frac{2}{3}$ (0.67) | 1 (perfect)