Stat Support Activity: Correlation Coefficient Analysis

Notes:

- When two different quantitative variables with different units are related or associated, we say there is <u>correlation</u> between the variables.

- The <u>Correlation Coefficient</u> (r) is a number between -1 and $+1$ that determines the strength and direction of the correlation.

- If the Correlation Coefficient (r) is close to zero on the number line, then the two quantitative variables have not correlation.

- If the Correlation Coefficient (r) is close to $+1$ on the number line, then the two quantitative variables have a strong positive correlation.

- A positive correlation also indicates that the two quantitative variables are <u>directly</u> <u>related</u>. As one variable (x) increases, the other variable (y) also tends to increase. The scatterplot of the data will have a spread pattern that generally goes up from left to right.

- If the Correlation Coefficient (r) is close to -1 on the number line, then the two quantitative variables have a strong negative correlation.

- A negative correlation also indicates that the two quantitative variables are <u>inversely</u> <u>related</u>. As one variable (x) increases, the other variable (y) decreases. As one variable (x) decreases, the other variable (y) increases. The scatterplot of the data will have a spread pattern that generally goes down from left to right.

- Coefficient of Determination (r^2) : This number is calculated by squaring the Correlation Coefficient (r) and will also need to be converted to a percentage. It is the proportion or percentage of variability in the response variable (y) that can be explained by the linear relationship with the explanatory variable (x). Note: The Correlation Coefficient (r) is NOT a proportion so should NEVER be converted to a percentage.

- If r is close to +1 (like r = +0.893) then there is a Strong, Positive Correlation (line going up from left to right (positive slope) and the points in scatterplot are close to line) ,
(r ≈ +0.6, +0.7, +0.8 , +0.9 usually indicate pretty strong positive correlation)

- If r is close to −1 (like r = −0.916) then there is a Strong Negative Correlation (line going down from left to right (negative slope) and the points in the scatterplot are close to the line)
(r ≈ −0.6, −0.7, −0.8 , −0.9 usually indicate pretty strong negative correlation)

- If r close to zero (like +0.037 or −0.009) then there is a No linear correlation. Points in the scatterplot do not follow any linear pattern. There still could be a nonlinear curved pattern though. (r ≈ ±0.1, ±0.0 usually indicate no linear correlation)

- If r ≈ ±0.2, ±0.3 usually indicate very weak linear correlation. There is some linear pattern but the points are very far from the regression line.

Stat Support Activity: Correlation Coefficient Analysis

- If r ≈ ±0.4, ±0.5 usually indicate moderate linear correlation. There is a linear pattern and points are only moderately close to the regression line.

Directions: Fill out the following table.

	Correlation Coefficient (r)	Coefficient of Determination Proportion (r^2)	Coefficient of Determination Percentage (r^2)	Strength of the Correlation (<u>S</u>trong, <u>M</u>oderate, <u>W</u>eak, or <u>No</u> Correlation	If there is some correlation, what is the direction of correlation? (<u>P</u>ositive or <u>N</u>egative)	If there is some correlation, are the two quantitative variables <u>directly</u> related or <u>inversely</u> related?
1.	+0.273					
2.	−0.875					
3.	0					
4.	+0.581					
5.	−0.118					
6.	+0.924					
7.	−0.512					
8.	+0.041					
9.	−0.236					
10.	+0.701					
11.	−0.933					