

4/20/2021

Scatterplots, Correlation, Coefficient of Determination

Correlation – is a measure of the strength of a linear relationship among two variables (numerical). If the slope of the best-fit line is positive, the correlation will be positive. If the slope of the best-fit line is negative, the correlation will be negative. The correlation is a value between -1 and 1. Closer to -1 or 1 is a strong correlation, and closer to 0 represents a weak linear correlation.

In general, we can think of correlations as being strong, moderate or weak.

$|r| > 0.7$ or 0.8 the correlation is considered strong

$0.4 < |r| < 0.7$ the correlation is considered moderate

$|r| < 0.4$ the correlation is considered to be weak.

Coefficient of Determination r^2 or R^2 is a measure how much the relationship between two variables can explain the variability of the y-variable.

Is a value between 0 and 1, and it can be interpreted as a percent:

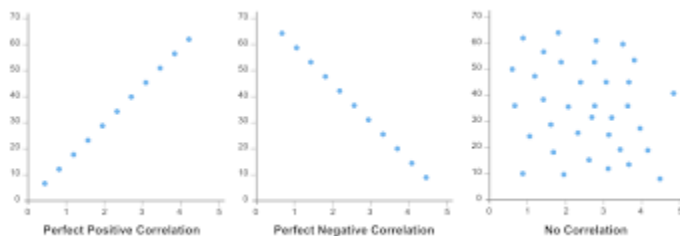
Suppose the $R^2 = 0.75$, then we would that the variability in the y-variable can be explained (reduced) but knowing about the relationship to the x-variability. If we know the value of x, we can better predict the value of y. Knowing the value of x reduces the uncertainty in our prediction of y by 75%.

$R^2 > 0.5$ or 0.6 is a strong relationship

$0.2 < R^2 < 0.5$ is a moderate relationship

$R^2 < 0.2$ is a weak relationship

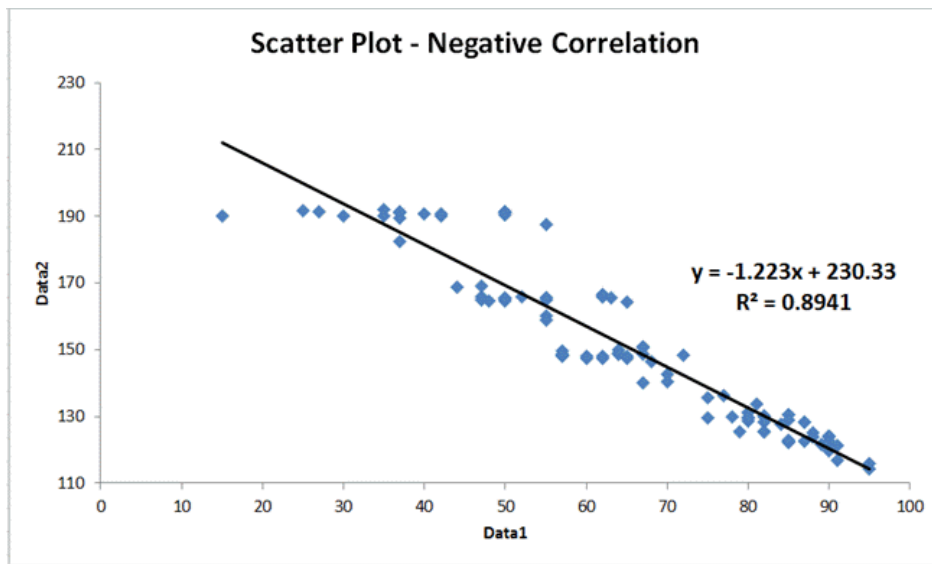
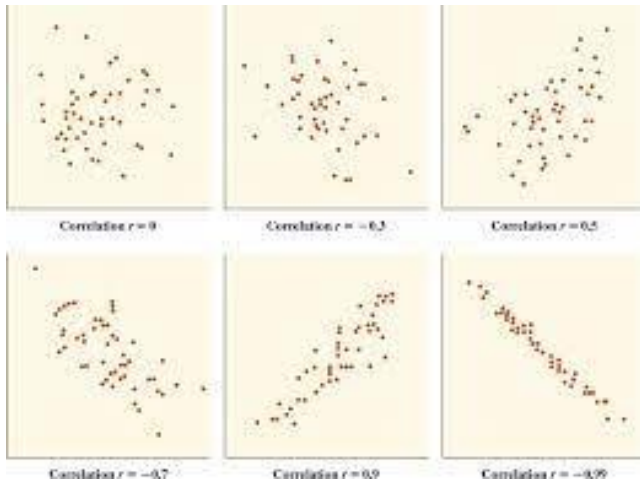
The Greek population variable for correlation is ρ .



$r = 1$

$r = -1$

$r = 0$



$$r = -0.94556 = -\sqrt{R^2}$$

$$y = 0.0241x - 413.28$$

$$R^2 = 0.7359$$

x is cost, and y is units

The variability in units produced can be reduced by 73.59% because of the relationship with cost.

For each dollar increase in cost the number of units produced increases by 0.0241.

Final Exam:

Some emphasis on this last material for HW 7 because it hasn't been tested before

The rest of the exam (most of it) will be based on Exam 1 and Exam 2.