COVARIANCE AND CORRELATION

Covariance of two variables:

$$Cov(X,Y) = \frac{\sum_{i=1}^{n}(X - \bar{X}) \cdot (Y - \bar{Y})}{n-1}$$

Correlation coefficient:

$$r = \frac{Cov(X,Y)}{\sqrt{s_X^2 \cdot s_Y^2}}$$

$$r = \frac{Cov(X,Y)}{s_X \cdot s_Y}$$

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - X)^2 \cdot \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

$$\boxed{r = \frac{SS_{XY}}{\sqrt{SS_X \cdot SS_Y}}}$$

Where:

$$SS_X = \sum \left(X_i - \bar{X} \right)^2 = \sum X_i^2 - \frac{\left(\sum X_i \right)^2}{n}$$

$$SS_Y = \sum \left(Y_i - \bar{Y} \right)^2 = \sum Y_i^2 - \frac{\left(\sum Y_i \right)^2}{n}$$

$$SS_{XY} = \sum \left(X_i - \bar{X} \right)\left(Y_i - \bar{Y} \right) = \sum X_i Y_i - \frac{\left(\sum X_i \right)\left(\sum Y_i \right)}{n}$$

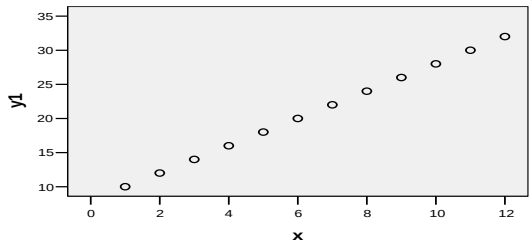
$$r = \frac{\sum \left[\left(Y_i - \bar{Y} \right)\left(X_i - \bar{X} \right) \right]}{\sqrt{\sum \left(Y_i - \bar{Y} \right)^2 \sum \left(X_i - \bar{X} \right)^2}} \cdot$$

$$r = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{n}}{\sqrt{\left[\sum X_i^2 - \frac{\left(\sum X_i\right)^2}{n}\right]\left[\sum Y_i^2 - \frac{\left(\sum Y_i\right)^2}{n}\right]}}$$

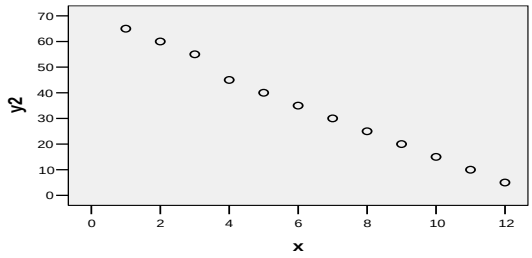
or

$$\boxed{r = \frac{SS_{XY}}{\sqrt{SS_X \cdot SS_Y}}}$$

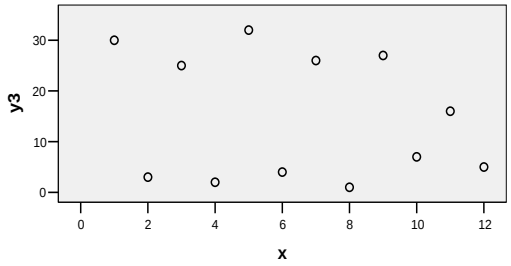
(a) $r = +1$



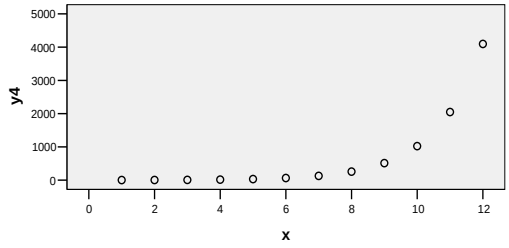
(b) $r = -1$



(c) $r = -0{,}1$


y3
30
20
10
0
0
2
4
6
8
10
12
x

(d) $r = +0,6$


y4
5000
4000
3000
2000
1000
0
0
2
4
6
8
10
12
x