Arbeitsblatt lineare Regression

Formeln:

Arithmetisches Mittel $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ (Gleiches Vorgehen für die Y-Werte)

Varianz $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ (Gleiches Vorgehen für die Y-Werte)

Standardabweichung $s = \sqrt{s^2}$

Kovarianz $\text{cov}(x,y) = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n}$

Lineare Regression $b = \text{cov}/s^2_x$ $a = \bar{y} - b * \bar{x}$

Aufgaben:

1) Gegeben sind die folgenden vier Punkte:
   x-Werte 0, 2, 3, 5
   y-Werte 8, 3, 1, -2

a) Bestimme die Mittelwerte von x und y, die Varianzen und die Standardabweichung s, die Kovarianz und die Gleichung der Regressionsgeraden $\hat{Y} = a + b * x$.

b) Beschreibe die Gerade mithilfe der slope und des intercept.

2) Gegeben sei die Regressionsgleichung $\hat{Y} = 4 - 0{,}5X$. Welche Aussage kann auf dieser Grundlage für die Korrelation zwischen X und Y getroffen werden?

   Gegeben sei die Regressionsgleichung $\hat{Y} = 2 + 4X$. Welche Aussage kann auf dieser Grundlage für die Korrelation zwischen X und Y getroffen werden?

3) Gegeben sei eine lineare Regression mit a = 12 und b = -1. Wie lautet der vorausgesagte Wert für $\hat{Y}$ an der Stelle X = 5?

4) Zeichne die Regressionsgerade $\hat{Y} = 2 + 2X$ in ein Koordinatensystem, in dem der Wert 0 auf der X-Achse <u>nicht</u> gleichzeitig die Y-Achse darstellt.

5) Beschreibe in eigenen Worten die Funktionsweise der linearen Regression und erkläre das Prinzip der „Ordinary-Least-Squares“.