

11. Der Entwickler eines Filesystems untersucht, ob es einen Zusammenhang zwischen der Größe und der Anzahl der Dateien in einem Verzeichnis gibt. Er ermittelt die mittlere Dateigröße und die Datei-anzahl für sechs Verzeichnisse.

Anzahl	3	15	60	89	167	193
Größe [MB]	7.3	8.2	9.5	3.2	0.7	1.4

- (a) Berechne die Kovarianz, das Bestimmtheitsmaß und den Korrelationskoeffizienten. Interpretiere das Ergebnis.
- (b) Zeichne einen Scatterplot (Größe über Anzahl), berechne die Regressionsgerade und zeichne diese dort ein.
12. Ein Stapel CD-RWs wird überprüft und festgestellt, wie oft jede CD-RW (wieder-)beschrieben wurde (x -mal) und wie viele Blöcke defekt sind (y Blöcke).

$x \backslash y$	0	1	2	3	4
1	18	2			
2	15	6	1		
3	12	5	8	3	
4	4	6	6	3	
5		1	5	2	3

Gib die vollständige Kontingenztafel (inklusive Randhäufigkeiten) mit absoluten sowie relativen Häufigkeiten an. Berechne \bar{x} , \bar{y} , s_x , s_y , $s_{x,y}$, $r_{x,y}$ und interpretiere die Ergebnisse.

13. Eine zweidimensionale Stichprobe $\{(x_1, y_1), \dots, (x_n, y_n)\}$ wird mit $y \approx f(x) = ax^2$ modelliert. Gib allgemein den quadratischen Fehler (SSE) dieses Modells an. Minimiere diesen durch die Wahl von a mittels Ableitung nach a und Nullsetzen. Berechne a konkret für folgende Stichprobe, zeichne einen Scatterplot und zeichne die Regressionskurve $y = ax^2$ ein.

x_i	1.1	1.9	2.4	3.3	4.5
y_i	1.4	7.1	7.3	19.5	25.2