

DATA SCIENCE

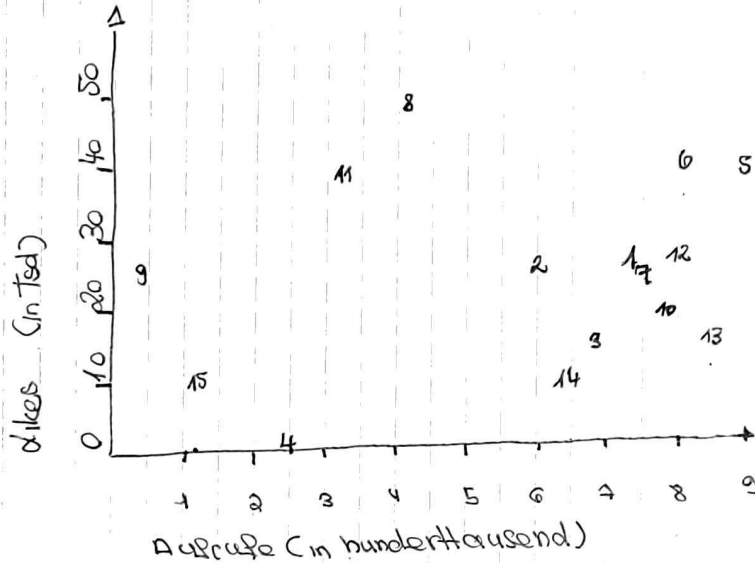
Praktikum 2

Fabian Gottling & Larissa Rychlau

D1

- a) Statistische Einheiten: "Videos"
Merkmale: "Likes", "Aufrufe"
Skalenniveau: kardinal (Verhältnisskala)

b)



"Likes" und "views" korrelieren.

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

$$\bar{X} \text{ Aufrufe: } 580,93 \approx 581$$

$$\bar{Y} \text{ Likes: } 23,6 \approx 24$$

$$r = \frac{8797}{13100,42}$$

$$r = 0,67$$

X	NR	X ²	NR	Likes	Aufrufe
183	23409		25	-24 = -1	734 - 581 = 153
28	784		25	-24 = -1	608 - 581 = 28
1077	1.162.081		13	-24 = -11	670 - 581 = 89
7458	55.621.264		2	-24 = -22	242 - 581 = -339
4560	20.793.600		39	-24 = -15	885 - 581 = 304
3712	13.778.944		40	-24 = -16	818 - 581 = 237
0	0		24	-24 = 0	751 - 581 = 170
3356	15.649.336		47	-24 = 23	409 - 581 = -172
522	272.484		25	-24 = 1	68 - 581 = -513
1152	1327.104		18	-24 = -6	773 - 581 = 192
3556	12.645.136		38	-24 = 14	328 - 581 = -253
446	19.8916		20	-24 = 2	804 - 581 = 223
2730	7.452.900		14	-24 = -10	854 - 581 = 273
1020	1.040.400		9	-24 = -15	649 - 581 = 68
6454	41.654.116		10	-24 = -14	120 - 581 = -461
	171.622.577				

- d) Nicht zwingend, eine nicht-niedrig/überdurchschnittliche Like-Anzahl ist aber wahrscheinlich.

D2

a)
$$p = 1 - \frac{G \sum d_i^2}{n(n^2-1)} = 1 - \frac{G \sum (R_{xi} - R_{yi})^2}{n(n^2-1)}$$

$$\Rightarrow p = 1 - \frac{G \cdot 8}{9(9^2-1)} = 1 - \frac{48}{792} = 0,03$$

- b) Die Zuordnung der nicht-numerischen Werte ist rein subjektiv. Eine andere Bewertung führt zu anderen Ergebnissen

- c) Es muss eine Ordinale Sortierung geben bzw. es müssen Objektstandards (z.B. bpm) für die Sortierung gefunden werden

NR	(x ²)	Test	
1	1	0	0
1	2	1	1
1	2	1	1
2	2	0	1
2	2	0	1
2	2	0	1
3	3	0	1
3	1	2	1
3	5	4	9
		<u>16</u>	<u>9</u>
8		26	16

D3

a)