

Einfache lineare Regression

Übung 1: Ermitteln der Regressionsfunktion



Übung 1: Zusammenhang zwischen Körpergröße und Gewicht

Größe cm	Gewicht kg
170	55
167	56
187	84
169	61
176	70
184	83
202	93
190	88
175	65
167	59
158	50
176	65
179	77
182	80
191	82
181	78
183	81
177	82
172	78
169	64
166	59
179	69
189	80
170	56
174	76
168	58

Daten „Regression_Beiispiel.csv“ (s. Moodle)

Aufgabe: Nehmen Sie einen linearen Zusammenhang zwischen Größe und Gewicht an und ermitteln Sie die Regressionsgerade zur Erklärung des Gewichts (hier = abhängiges Merkmal) über die Größe (hier = unabhängiges Merkmal).

1. Berechnen Sie Steigung a und Achsenabschnitt b der Regressionsgeraden.

$$a = \frac{\text{Kovarianz}(x,y)}{\text{Varianz}(x)}$$

$$b = \bar{y} - a * \bar{x}$$

2. Welches Gewicht sagt das Modell für eine Körpergröße von a) 160 cm und b) 195 cm voraus?

Einfache lineare Regression

Lösung zur Übung 1: Ermitteln der Regressionsfunktion



x = Größe (cm)	y = Gewicht (kg)	de_mean(x)	de_mean(y)	de(x)*de(y)	de(x)*de(x)
170	55	-7	-16	112,19	48,46
167	56	-10	-15	150,57	99,23
187	84	10	13	129,34	100,77
169	61	-8	-10	80,53	63,39
176	70	-1	-1	1,07	0,92
184	83	7	12	83,65	49,54
202	93	25	22	547,96	626,92
190	88	13	17	220,15	170,00
175	65	-2	-6	12,00	3,85
167	59	-10	-12	120,69	99,23
158	50	-19	-21	400,38	359,54
176	65	-1	-6	5,88	0,92
179	77	2	6	12,00	4,16
182	80	5	9	44,76	25,39
191	82	14	11	152,80	197,08
181	78	4	7	27,80	16,31
183	81	6	10	59,69	36,46
177	82	0	11	0,42	0,00
172	78	-5	7	-34,16	24,62
169	64	-8	-7	56,65	63,39
166	59	-11	-12	132,80	120,16
179	69	2	-2	-4,31	4,16
189	80	12	9	106,96	144,92
170	56	-7	-15	105,23	48,46
174	76	-3	5	-14,47	8,77
168	58	-9	-13	117,53	80,31
mean	176,96	71,12		101,08	92,19

Kovarianz(x,y) Varianz(x)

$$a = \frac{\text{Kovarianz}(x,y)}{\text{Varianz}(x)}$$

$$b = \bar{y} - a * \bar{x}$$

Aufgabe 1:
Berechnung a und b:

$$a = 101,08 / 92,19 = 1,096$$

$$b = 71,12 - 1,096 * 176,96$$

$$= -122,91$$

Regressionsfunktion

$$y = 1,096 * x - 122,91$$

Einfache lineare Regression

Lösung zur Übung 1: Ermitteln der Regressionsfunktion



Aufgabe 2:

$$\text{a) } x = 160 \rightarrow y = 1,096 * 160 - 122,91 = 52,5$$

Bei einer Größe von 160 cm wird ein Gewicht von 52,5 kg vorhergesagt.

$$\text{b) } X = 195 \rightarrow y = 1,096 * 195 - 122,91 = 90,9$$

Bei einer Größe von 195 cm wird ein Gewicht von 90,9 kg vorhergesagt.

Einfache lineare Regression

Übung 2: Ermitteln der Güte



Übung 2: Güte des Modells Gewicht – Körpergröße

Größe cm	Gewicht kg
170	55
167	56
187	84
169	61
176	70
184	83
202	93
190	88
175	65
167	59
158	50
176	65
179	77
182	80
191	82
181	78
183	81
177	82
172	78
169	64
166	59
179	69
189	80
170	56
174	76
168	58

Aufgabe:

Berechnen Sie das Bestimmtheitsmaß zur ermittelten Regressionsgeraden und interpretieren Sie diesen Wert.

Formel:

$1 - \text{mittleres Fehlerquadrat} / \text{Varianz}(y)$

Einfache lineare Regression

Lösung zur Übung 2: Ermitteln der Güte



y = Gewicht (kg)	Prognose y	Abw = y - Prog	Abw **2	de_mean(y)**2
55	63,48	-8,48	71,95	259,71
56	60,19	-4,19	17,58	228,47
84	82,12	1,88	3,53	166,01
61	62,39	-1,39	1,92	102,32
70	70,06	-0,06	0,00	1,24
83	78,83	4,17	17,37	141,24
93	98,57	-5,57	31,01	478,94
88	85,41	2,59	6,70	285,09
65	68,96	-3,96	15,72	37,40
59	60,19	-1,19	1,42	146,78
50	50,33	-0,33	0,11	445,86
65	70,06	-5,06	25,61	37,40
77	73,35	3,65	13,32	34,63
80	76,64	3,36	11,29	78,94
82	86,51	-4,51	20,32	118,47
78	75,54	2,46	6,04	47,40
81	77,74	3,26	10,65	97,71
82	71,16	10,84	117,56	118,47
78	65,68	12,32	151,90	47,40
64	62,39	1,61	2,60	50,63
59	59,10	-0,10	0,01	146,78
69	73,35	-4,35	18,93	4,47
80	84,31	-4,31	18,62	78,94
56	63,48	-7,48	55,99	228,47
76	67,87	8,13	66,13	23,86
58	61,29	-3,29	10,82	172,01
71,12		mean	26,81	137,64

Übung 2: Güte des Modells

$$R^2 = 1 - (26,8/137,6) = 0,805$$

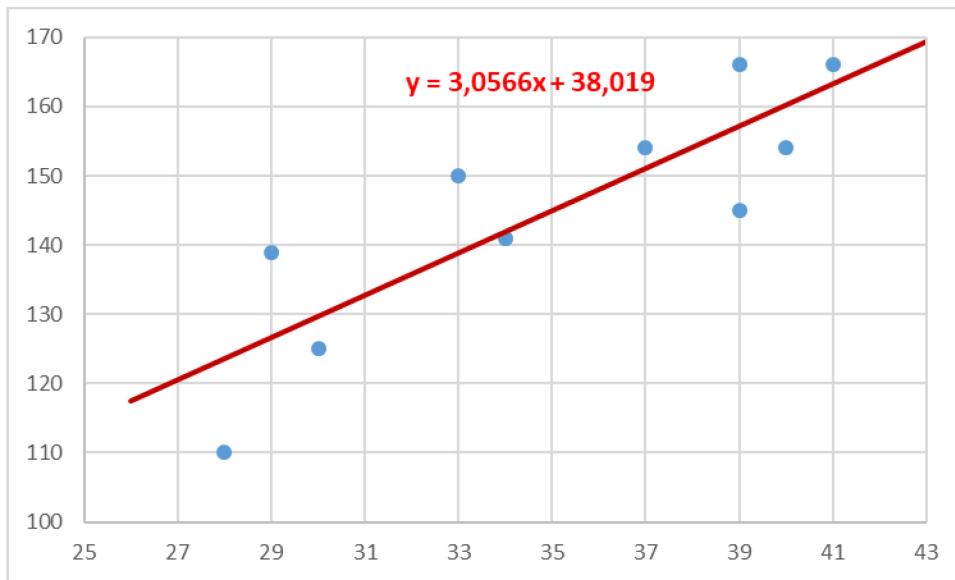
80,5% der Varianz in den beobachteten Werten für das Gewicht können durch das Modell erklärt werden.

Einfache lineare Regression

Umsetzung in XLS



Regressionsgerade in xls berechnen und anzeigen



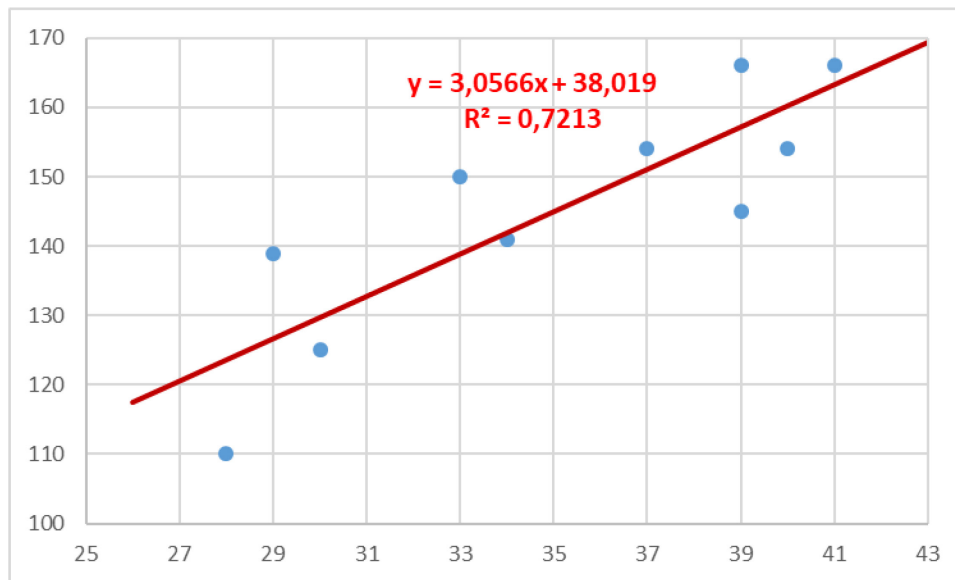
1. Punktdiagramm aus x und y erstellen
2. Diagrammelement hinzufügen: „Trendlinie linear“
3. Trendlinie formatieren: „Formel anzeigen“

Einfache lineare Regression

Umsetzung in XLS



Bestimmtheitsmaß in xls berechnen und anzeigen



1. In bestehendem Chart:
Trendlinie formatieren:
„Bestimmtheitsmaß
anzeigen“