Fachhochschule Köln
Fakultät für Wirtschafts- und Rechtswissenschaften
Prof. Dr. Arrenberg
Raum 221, Tel. 39 14
jutta.arrenberg@fh-koeln.de

Übungen zur Vorlesung Wirtschaftsstatistik

Lineare Regression

Aufgabe 7.1

Ein Unternehmen möchte den Einfluss seiner Werbemaßnahmen auf den erzielten Umsatz quantifizieren. Hierfür werden die jährlichen Ausgaben für Werbung und die jährlich erzielten Umsätze über einen Zeitraum von acht Jahren erfasst. Die Daten sind der folgenden Tabelle zu entnehmen:

Jahr	Werbung in 1000 €	Umsatz in 100 000 €
2006	6	9
2007	9	10
2008	10	13
2009	11	12
2010	12	10
2011	12	16
2012	13	15
2013	15	25

a) Berechnen und interpretieren Sie eine geeignete statistische Maßzahl zur Messung des linearen Zusammenhangs zwischen den jährlichen Werbeausgaben und dem Umsatz.

b) Angenommen das Unternehmen plant, 14 000 € für Werbung auszugeben. Mit welchem Umsatz kann es dann rechnen?

c) Angenommen das Unternehmen plant, einen Umsatz von 2 600 000 € zu erzielen. Wie viel Euro sind dann für Werbung auszugeben?

d) Für wie verlässlich halten Sie Ihre Ergebnisse aus b) und c)? (Begründung!)

Aufgabe 7.2

Ein Automobil-Konzern veröffentlichte in seinem Geschäftsbericht für das Jahr 2013 u. a. folgende Kennzahlen:

	Berichtsjahr 2006	2008	2010	2012
Umsatz (Mio GE)				
Inland	23 682	36 360	34 326	34 504
Ausland	41 670	39 955	42 260	53 615
Produktion (Tsd. Automobile)				
Inland	1 783	1 814	1 411	1 526
Ausland	1 165	1 424	1 608	2 069
Produktivität = produzierte Automobile je Mitarbeiter				
Inland	11,1	10,9	9,4	10,7
Ausland	12,9	12,9	15,6	18,1

a) Prognostizieren Sie die Produktion des Konzerns im Ausland für das Jahr 2017 mittels einer geeigneten statistischen Methode.

b) Nehmen Sie kritisch zur Aussagekraft des Prognosewertes Stellung.

c) Beurteilen Sie die Güte Ihres Prognosemodells mit Hilfe der Berechnung einer geeigneten Maßzahl.

Aufgabe 7.3
Hinweis: Diese Aufgabe bringt wirklich Spaß!
Im Netz liegen Applets aus meinem Statistik-Projekt. Bearbeiten Sie das Applet „Raten von Korrelationen“. Sie gelangen zu dem Applet wie folgt:

http://campus.uni-muenster.de/fileadmin/einrichtung/imib/lehre/skripte /biomathe/bio/eda/korr.html

Klicken Sie auf „neue Plots“. Es erscheinen vier Streudiagramme und vier Korrelationswerte. Ordnen Sie die Korrelationswerte den jeweiligen Streudiagrammen zu.

Aufgabe 7.4
Im Netz liegen Applets aus meinem Statistik-Projekt. Bearbeiten Sie das Applet „KQ Regressionsgerade“. Sie gelangen zu dem Applet wie folgt:

1. Loggen Sie sich ein unter *web.neuestatistik.de/ inhalte_web/content/start.html*
2. Klicken Sie auf *Medien Galerie*
3. Klicken Sie auf *Regression*
4. Klicken Sie auf Java-Applet KQ Regressionsgerade

Die Steigung ist $b = \widehat{\alpha}$ und der Achsenabschnitt ist $a = \widehat{\beta}$. Das Bestimmtheitsmaß ist $B = R^2$. Zeichnen Sie per Maus-Klick einige Punkte in das Streudiagramm und beobachten Sie dabei den Wert des Bestimmtheitsmaßes.

Aufgabe 7.5
Im Netz liegen Applets aus meinem Statistik-Projekt. Bearbeiten Sie das Applet Regressionsgerade“. Sie gelangen zu dem Applet wie folgt:

1. Loggen Sie sich ein unter *web.neuestatistik.de/ inhalte_web/content/start.html*
2. Klicken Sie auf *Medien Galerie*
3. Klicken Sie auf *Regression*
4. Klicken Sie auf Java-Applet Regressionsgerade

Die Steigung ist $b = \widehat{\alpha}$ und der Achsenabschnitt ist $a = \widehat{\beta}$. Es sind Punkte im Streudiagramm eingezeichnet. Verschieben Sie mit der Maus Punkte in dem Streudiagramm, so dass sich das Vorzeichen der Steigung verändert. Gelingt Ihnen das?

Lösung zu Aufgabe 7.1:
X=Werbeausgaben (in 1 000 Euro pro Jahr)
Y=Jahres-Umsatz (in 100 000 Euro)

x_i	y_i	$x_i \cdot y_i$	x_i^2	y_i^2
6	9			
9	10			
10	13			
11	12			
12	10			
12	16			
13	15			
15	25			
88	110	1 288	1 020	1 700

a) $b_1 = \dfrac{8 \cdot 1\,288 - 88 \cdot 110}{8 \cdot 1\,020 - 88 \cdot 88} = \dfrac{624}{416} = 1{,}5$

$b_2 = \dfrac{624}{8 \cdot 1\,700 - 110 \cdot 110} = \dfrac{624}{1\,500} = 0{,}416$

$r = +\sqrt{1{,}5 \cdot 0{,}416} = 0{,}789\,936\,7$

d.h. es liegt positive mittelstarke Korrelation vor.

b) Gesucht: $a_1 + b_1 \cdot 14 =?$

$a_1 = \dfrac{110 - 1{,}5 \cdot 88}{8} = -2{,}75$

$-2{,}75 + 1{,}5 \cdot 14 = 18{,}25$

d.h. es ist mit einem Umsatz von 1 825 000 Euro zu rechnen.

c) Gesucht: $a_2 + b_2 \cdot 26 =?$

$a_2 = \dfrac{88 - 0{,}416 \cdot 110}{8} = 5{,}28$

$5{,}28 + 0{,}416 \cdot 26 = 16{,}096$

d.h. es ist mit Ausgaben für Werbung in Höhe von 16 000 Euro zu rechnen.

d) Da es sich bei der Berechnung unter b) um eine Interpolation handelt und da der lineare Zusammenhang beinahe stark ist, hat der Prognosewert Verlass. Hingegen ist der Prognosewert unter c) ein extrapolierter Wert, auf den kein Verlass ist.

Lösung zu Aufgabe 7.2:
X = Zeit
Y = Produktion im Ausland
Regressionsgerade $f(x) = a_1 + b_1 x$

Jahr	x_i	y_i	$x_i\,y_i$	x_i^2
2006	0	1 165	0	0
2008	2	1 424	2 848	4
2010	4	1 608	6 432	16
2012	6	2 069	12 414	36
$\sum$	12	6 266	21 694	56

$$b_1 = \frac{4 \cdot 21\,694 - 12 \cdot 6\,266}{4 \cdot 56 - 12 \cdot 12} = \frac{11\,584}{80} = 144{,}8$$

$$a_1 = \frac{6\,266 - 144{,}8 \cdot 12}{4} = \frac{4\,528{,}4}{4} = 1\,132{,}1$$

a) Jahr 2017$\widehat{=}$ Zeitpunkt 11
$f(11) = 1\,132{,}1 + 144{,}8 \cdot 11 = 2\,724{,}9$
d.h. für das Jahr 2017 ergibt sich ein Prognosewert von 2 724 900 Automobilen.

b) Der Zeitpunkt $x = 11$ liegt weit weg (Extrapolation) von den Zeitpunkten 0,2,4,6, so dass eine Prognose für diesen Zeitpunkt mittels der Regressionsgeraden gewagt ist.

c)

Jahr	y_i	y_i^2
2006	1 165	1 357 225
2008	1 424	2 027 776
2010	1 608	2 585 664
2012	2 069	4 280 761
$\sum$	6 266	10 251 426

Korrelationskoeffizient:

$$\begin{aligned} r_{xy} &= \frac{4 \cdot 21\,694 - 12 \cdot 6\,266}{\sqrt{(4 \cdot 56 - 12 \cdot 12)(4 \cdot 10\,251\,426 - 6\,266 \cdot 6\,266)}} \\ &= \frac{11\,584}{\sqrt{80 \cdot 1\,742\,948}} = \frac{11\,584}{11\,808{,}2954} = 0{,}9810 \end{aligned}$$

d.h. die Regressionsgerade beschreibt die Daten sehr gut.

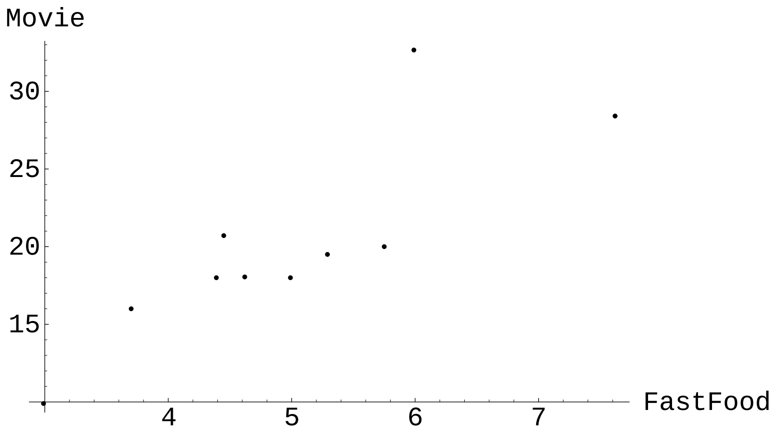
Fachhochschule Köln
Fakultät für Wirtschafts- und Rechtswissenschaften
Prof. Dr. Arrenberg
Raum 221, Tel. 39 14
jutta.arrenberg@fh-koeln.de

Vorlesung Wirtschaftsstatistik

Arbeitsblatt Lineare Regression

Beispiel (Berenson et al., page 82)
In der nachfolgenden Tabelle stehen die Preis in US$ für einen Fast-Food-Hamburger und zwei Kinokarten im Jahr 2005 in verschiedenen Weltstädten:

City	Hamburger	Kinokarten
Tokyo	5,99	32,66
London	7,62	28,41
New York	5,75	20,00
Sydney	4,45	20,71
Chicago	4,99	18,00
San Francisco	5,29	19,50
Boston	4,39	18,00
Atlanta	3,70	16,00
Toronto	4,62	18,05
Rio de Janeiro	2,99	9,90
Köln		



Streudiagramm

a) Tragen Sie die Werte für Köln ins Streudiagramm ein.

b) Legen Sie durch die Punktewolke eine Gerade, welche die Daten sehr gut beschreibt.

Fachhochschule Köln
Fakultät für Wirtschafts- und Rechtswissenschaften
Prof. Dr. Arrenberg
Raum 221, Tel. 39 14
jutta.arrenberg@fh-koeln.de

Übungen zur Vorlesung Wirtschaftsstatistik
Arbeitsblatt: Lineare Regression

Beispiel
Gibt es bei den Bundestagswahlen einen Zusammenhang zwischen dem Stimmanteil für die CDU/CSU und der insgesamten Wahlbeteiligung?
Aus den Jahren 1949 bis 2013 liegen folgende Daten vor:

Jahr	Wahlbeteiligung (in %)	CDU/CSU (in %)	SPD (in %)
1949	78.5	31.0	29.2
1953	85.5	45.2	28.8
1957	87.8	50.2	31.8
1961	87.7	45.3	36.2
1965	86.8	47.6	39.3
1969	86.7	46.1	42.7
1972	91.1	44.9	45.8
1976	90.7	48.6	42.6
1980	88.6	44.5	42.9
1983	89.1	48.8	38.2
1987	84.4	44.3	37.0
1990	77.8	43.8	33.5
1994	79.0	41.5	36.4
1998	82.2	35.1	40.9
2002	79.1	38.5	38.5
2005	77.7	35.2	34.2
2009	70,8	33,8	23,0
2013	71,5	41,5	25,7

wahlen.sav

Stellen Sie ein Modell auf, das den Stimmenanteil der CDU/CSU erklärt durch die Höhe der Wahlbeteiligung.
Lösung:
X=Wahlbeteiligung (in Prozent)
Y= Stimmenanteil der CDU/CSU
Modell: $\boxed{\text{CDU/CSU-Stimmenanteil}} \approx a_1 + b_1 \cdot \boxed{\text{Wahlbeteiligung}}$
$b_1 = 0{,}656$ d.h. steigt die Wahlbeteiligung um einen Prozentpunkt, so steigt der Stimmenanteil für die CDU/CSU um etwa 0,7 Prozentpunkte.
$a_1 = -11{,}917$
$r = 0{,}727$ d.h. positive fast starke Korrelation; d.h. je höher die Wahlbeteilgung, desto höher ist der Stimmenanteil für die CDU/CSU.