Lösungen

Lösung zu a) Die Daten scheinen sich näherungsweise entlang einer Geraden „positioniert" zu sein. Das lässt schon recht gut vermuten, dass ein linearer Zusammenhang vorhanden sein könnte.

Lösung zu b) Wenn ein Anstieg/eine Abnahme der Deutschen Bank Aktie auch zu einem Anstieg/einem Sinken des DAX-Index führt, so könnte man den Aktienkurs nutzen, um Vorhersagen für den DAX zu treffen. Der DAX ist ja eine Indexzahl, die aus 30 ausgewählten deutschen Aktien gebildet wird. Dies ist aber erst möglich, wenn die Kurse der Aktien feststehen. Hat man mit einer ausgewählten Aktie einen guten Indikator für das Verhalten (Anstieg oder Sinken) des DAX, so könnte das z.B. eine Entscheidungshilfe für den Kauf von sehr risikobehaftete Optionsscheine sein. Bei Optionsscheinen geht es darum (ähnlich wie beim Roulette) darauf zu spekulieren, ob der DAX steigt oder fällt.

Lösung zu c) Rein optisch sicher die beiden oberen, da die Daten sich eher entlang einer Geraden zu orientieren scheinen. Prinzipiell ist die Frage aber unsinnig, das es sich immer um dieselben Daten handelt. Dies unterstreicht aber umso mehr, dass Deutungen aufgrund graphischer Darstellungen mit Vorsicht zu genießen sind.

Lösung zu d):

		1.Schritt		**2.Schritt**	
Kurswert	Dax	KW-MW	DAX-MW	(KW-MW)/SA	(DAX-MW)/SA
24	4210	-9	-30	**-1,09**	**-1,39**
31	4260	-2	20	**-0,24**	**0,93**
44	4250	11	10	**1,33**	**0,46**
Mittelwert (MW)					
33	4240				
Standardabweichung (SA)					
8,29	21,60				

Lösung zu e)

		1.Schritt		**2.Schritt**	
Kurswert	Dax	KW-MW	DAX-MW	(KW-MW)/SA	(DAX-MW)/SA
56	5240	-10,5	74,75	**-1,30**	**1,32**
62	5100	-4,5	-65,25	**-0,56**	**-1,15**
71	5122	4,5	-43,25	**0,56**	**-0,76**
77	5199	10,5	33,75	**1,30**	**0,60**
Mittelwert (MW)					
66,5	5165,25				
Standardabweichung (SA)					
8,08	56,69				

Lösung zu f)

Immer gilt Mittelwert = 0 und Standardabweichung = 1.

Da von allen Daten der Mittelwert abgezogen wurde (also insbesondere auch vom Mittelwert dieser Daten), verteilen sie sich symmetrisch um 0. Da „Minus" gerechnet wurde, folgt Mittelwert – Mittelwert = 0, also liegt ihre neue Mitte bei 0.

Da alle diese „zentrierten" Daten durch ihre mittlere Abweichung (Streuung) von 0 geteilt wurden (also insbesondere auch die Standardabweichung dieser Daten), beträgt die Abweichung der neuen Daten von 0 durchschnittlich 1, da Standardabweichung / Standardabweichung =1 ergibt.

Der folgende rechnerische Nachweis ist sicherlich nicht einfach, aber erforderlich, um die zuvor beschriebene Idee noch präziser zu begründen! Er sollte von denjenigen gründlich durchdacht werden, die sich für den Leistungskurs interessieren – mehr noch: Wer sich für den LK Mathematik interessiert, der sollte die folgende Rechnung auch interessant finden!
Der rechnerische Nachweis für den Mittelwert kann folgendermaßen erfolgen:
Der Mittelwert von n transformierten Daten $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ berechnet sich durch:
$\frac{1}{n}\left(\tilde{x}_1 + \tilde{x}_2 + \ldots + \tilde{x}_n\right) = \frac{1}{n}\left(\frac{(x_1 - \bar{x})}{s_x} + \frac{(x_2 - \bar{x})}{s_x} + \ldots + \frac{(x_n - \bar{x})}{s_x}\right)$. Das lässt sich weiter umformen:

$$
\begin{aligned}
\frac{1}{n}\left(\frac{(x_1 - \bar{x})}{s_x} + \frac{(x_2 - \bar{x})}{s_x} + \ldots + \frac{(x_n - \bar{x})}{s_x}\right) &= \frac{1}{n \cdot s_x}\left((x_1 - \bar{x}) + (x_2 - \bar{x}) + \ldots + (x_n - \bar{x})\right) \\
&= \frac{1}{n \cdot s_x}\left(x_1 + x_2 + \ldots + x_n - \bar{x} - \bar{x} - \ldots - \bar{x}\right) \\
&= \frac{1}{n \cdot s_x}\left(x_1 + x_2 + \ldots + x_n - n \cdot \bar{x}\right) \\
&= \frac{1}{s_x}\left(\frac{1}{n}(x_1 + x_2 + \ldots + x_n - n \cdot \bar{x})\right) \\
&= \frac{1}{s_x}\left(\frac{1}{n}(x_1 + x_2 + \ldots + x_n) - \frac{1}{n} n \cdot \bar{x}\right) \\
&= \frac{1}{s_x}\left(\bar{x} - \bar{x}\right)
\end{aligned}
$$

= 0. Und das war die Behauptung!

Der rechnerische Nachweis für die Standardabweichung kann folgendermaßen erfolgen:
Die Standardabweichung von n transformierten Daten $\tilde{x}_1, \tilde{x}_2, \ldots, \tilde{x}_n$ berechnet sich gemäß der Formel für die Standardabweichung durch $\sqrt{\frac{1}{n}\left((\tilde{x}_1 - \tilde{\tilde{x}})^2 + (\tilde{x}_2 - \tilde{\tilde{x}})^2 + \ldots + (\tilde{x}_n - \tilde{\tilde{x}})^2\right)}$.
Hierbei ist $\tilde{\tilde{x}}$ der Mittelwert der transformierten Werte. Wie wir aber in der vorherigen Rechnung gezeigt haben, ist dieser Mittelwert = 0. Der Term für die Standardabweichung vereinfacht sich also zu : $\sqrt{\frac{1}{n}\left((\tilde{x}_1 - 0)^2 + (\tilde{x}_2 - 0)^2 + \ldots + (\tilde{x}_n - 0)^2\right)} = \sqrt{\frac{1}{n}\left(\tilde{x}_1^{\,2} + \tilde{x}_2^{\,2} + \ldots + \tilde{x}_n^{\,2}\right)}$

Diesen Term kann man umschreiben zu: $\sqrt{\frac{1}{n}\left(\left(\frac{x_1 - \bar{x}}{s_x}\right)^2 + \left(\frac{x_2 - \bar{x}}{s_x}\right)^2 + \ldots + \left(\frac{x_n - \bar{x}}{s_x}\right)^2\right)}$

$$
\begin{aligned}
&= \sqrt{\frac{1}{n \cdot s_x^{\,2}}\left((x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2\right)} \\
&= \frac{1}{s_x} \cdot \sqrt{\frac{1}{n}\left((x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2\right)} \\
&= \frac{1}{s_x} \cdot s_x
\end{aligned}
$$

= 1. Und das war die Behauptung!

Lösung zu g)

1.	2.	3.	4.	5.	6.	7.
r= 1	r= 0,83	r= 0,45	r= 0,00	r= -0,45	r= -0,83	r= -1

1. und 7. perfekt
2. und 6. stark
3. und 5. schwach
4. kein linearer Zusammenhang

Zur Orientierung:

0 = kein linearer Zusammenhang
0-0,5 = schwach linearer Zusammenhang
0,5-0,8 = mittlerer linearer Zusammenhang
0,8-1 = starker linearer Zusammenhang
1 = perfekt linearer Zusammenhang

Lösung zu h)

r = ± 2 ist nicht möglich! Genauer: r kann weder größer als 1 noch kleiner als –1 werden, es gilt also: -1 ≤ r ≤ +1

Das lässt sich folgendermaßen rechnerisch erklären:

1. In Aufgabe f) haben wir begründet, dass die Standardabweichung der transformierten Daten 1 ist, also $\sqrt{\frac{1}{n}\left(\tilde{x}_1^{\,2}+\tilde{x}_2^{\,2}+...+\tilde{x}_n^{\,2}\right)}=1$, umgeformt heißt das $\left(\tilde{x}_1^{\,2}+\tilde{x}_2^{\,2}+...+\tilde{x}_n^{\,2}\right)=n$, bzw. $\left(\tilde{y}_1^{\,2}+\tilde{y}_2^{\,2}+...+\tilde{y}_n^{\,2}\right)=n$.
2. Wenn man mit Hilfe der 1. und 2. binomischen Formel $(\tilde{x}_i \pm \tilde{y}_i)^2$ berechnet,
   - so liefert der „mittlere Teil" der binomischen Formeln die Produkte $\tilde{x}_i \cdot \tilde{y}_i$, die wir für den Korrelationskoeffizienten benötigen und
   - die ersten und dritten Glieder der binomischen Formeln die Quadrate $\tilde{x}_i^{\,2}$ und $\tilde{y}_i^{\,2}$, über deren Summen wir schon etwas aus 1. wissen und
   - für beide Formeln gilt durch das Quadrieren natürlich immer $(\tilde{x}_i \pm \tilde{y}_i)^2 \geq 0$.

Wir berechnen zuerst $\sum_{i=1}^{n}(\tilde{x}_i-\tilde{y}_i)^2$: Es gilt

$$0 \leq \sum_{i=1}^{n}(\tilde{x}_i-\tilde{y}_i)^2 = \sum_{i=1}^{n}(\tilde{x}_i^{\,2}-2\tilde{x}_i\tilde{y}_i+\tilde{y}_i^{\,2}) = \sum_{i=1}^{n}\tilde{x}_i^{\,2}-2\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i+\sum_{i=1}^{n}\tilde{y}_i^{\,2} = n-2\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i+n.$$

Lassen wir den mittleren Teil weg, so gilt $0 \leq n-2\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i+n$ oder $2\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i \leq 2n$ oder noch kürzer $\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i \leq n$.

Das bedeutet, dass wenn man die Produkte aller x- und y-Werte der transformierten Werte bildet, deren Summe nie größer wird als die Anzahl der jeweiligen Datenpaare. Teilt man nun beide Seiten durch n, dann gilt: $\frac{1}{n}\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i \leq 1$, also kurz r ≤ 1.

Nun berechnen wir $\sum_{i=1}^{n}(\tilde{x}_i+\tilde{y}_i)^2$: Es gilt

$$0 \leq \sum_{i=1}^{n}(\tilde{x}_i+\tilde{y}_i)^2 = \sum_{i=1}^{n}(\tilde{x}_i^{\,2}+2\tilde{x}_i\tilde{y}_i+\tilde{y}_i^{\,2}) = \sum_{i=1}^{n}\tilde{x}_i^{\,2}+2\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i+\sum_{i=1}^{n}\tilde{y}_i^{\,2} = n+2\sum_{i=1}^{n}\tilde{x}_i\tilde{y}_i+n.$$

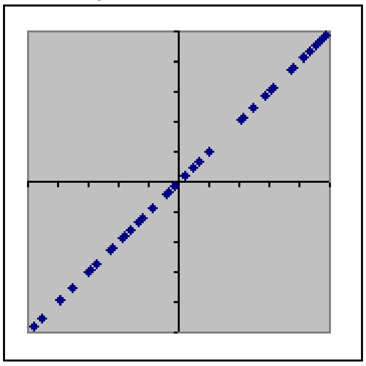
Lassen wir den mittleren Teil weg, so gilt $0 \le n + 2\sum_{i=1}^{n} \tilde{x}_i \tilde{y}_i + n$ oder $-2\sum_{i=1}^{n} \tilde{x}_i \tilde{y}_i \le 2n$ oder noch kürzer $-\sum_{i=1}^{n} \tilde{x}_i \tilde{y}_i \le n$.

Das bedeutet, dass wenn man die Produkte aller x- und y-Werte der transformierten Werte bildet, deren Summe nie größer wird als die Anzahl der jeweiligen Datenpaare. Teilt man nun beide Seiten durch n, dann gilt: $-\frac{1}{n}\sum_{i=1}^{n} \tilde{x}_i \tilde{y}_i \le 1$, also kurz $-r \le 1$. Damit dies gilt, muss $r \ge 1$ sein. Also gilt insgesamt $-1 \le r \le +1$.

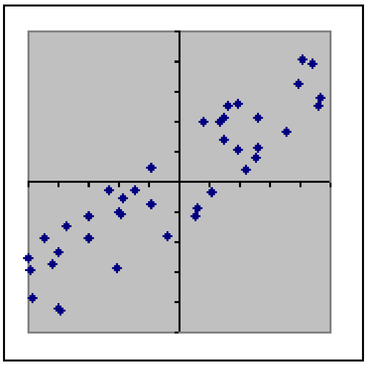
Lösung zu i)

r = 0,83, das entspricht einem starken linearen Zusammenhang. Somit könnte man zumindest für den betrachteten Zeitraum vermuten, dass die Entwicklung der Deutschen Bank Aktienwerte gut den Trend des DAX voraussagen. Langfristige Prognosen dieser Art würden stets neue Analysen dieser (und noch besserer) Art erforderlich machen.

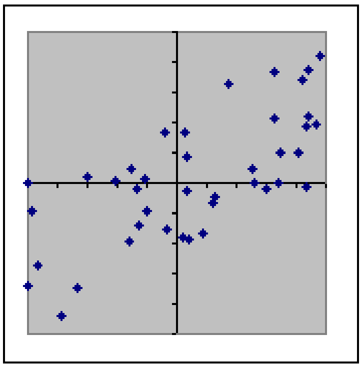
Lösung zu j)



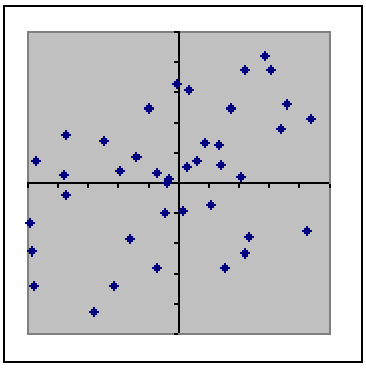
Korrelationskoeffizient = 1,0



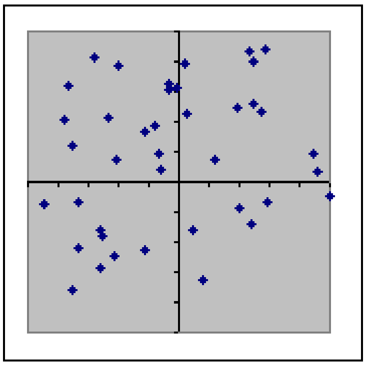
Korrelationskoeffizient = 0,9



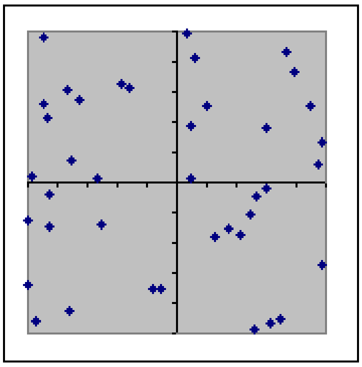
Korrelationskoeffizient = 0,7



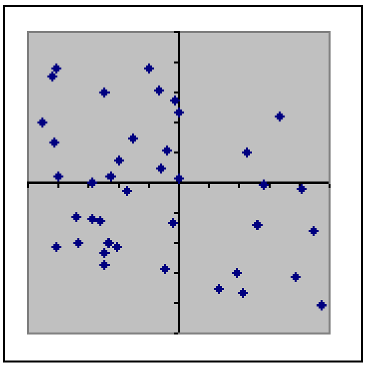
Korrelationskoeffizient = 0,4



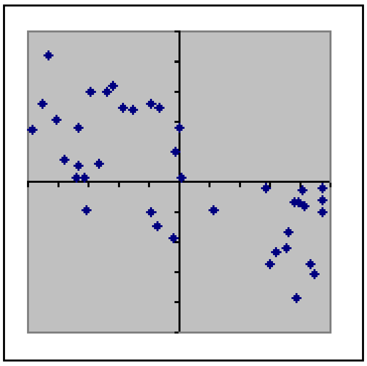
Korrelationskoeffizient = 0,2



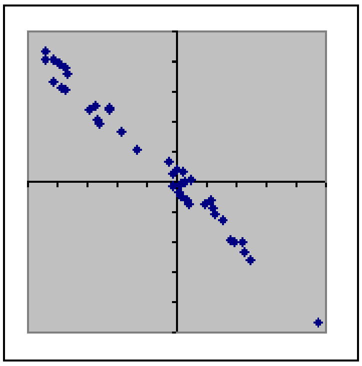
Korrelationskoeffizient = 0,0



Korrelationskoeffizient = –0,3



Korrelationskoeffizient = ––0,7



Korrelationskoeffizient = –0,99

Lösung zu k)

Datum	Deutsche Bank-Aktie (€)	DAX (Index-Werte)
03.04.2007	63	4375
04.04.2007	74	4525
05.04.2007	81	4650
06.04.2007	78	4600
07.04.2007	55	4250
10.04.2007	62	4125
11.04.2007	96	4700
12.04.2007	75	4575
13.04.2007	64	4450
14.04.2007	67	4725
17.04.2007	48,5	4150
18.04.2007	73	4575
19.04.2007	79	4925
20.04.2007	55,5	4175
21.04.2007	56	4175
24.04.2007	53,5	4150
25.04.2007	60	4400
26.04.2007	60	4400
27.04.2007	50	4325
28.04.2007	53	4175
03.05.2007	57	4125

Steuung Deutsche Bank 11,9578341
Streuung DAX 229,795828
Korrelationskoeffizent
r = 0,82695757

Steigung der Originalgeraden:
r * sy / sx = 15,891791

Mittelwert Deutsche Bank
64,7857143
Mittelwert DAX
4407,14286

Lösung zu l)

Der obigen Grafik kann man den Punkt $\left(\overline{x} \;/\; \overline{y}\right)$ = (64,786 / 4407,143) entnehmen.

Man such die Gleichung einer linearen Funktion, deren Steigung 15,893 man kennt und einen Punkt, durch den die Gerade verlaufen soll. Daher lautet der Ansatz:

$$4407{,}143 = 15{,}892 \cdot 64{,}786 + b$$

Aufgelöst nach b ergibt sich b ≈ 3377,6. Passt haargenau!