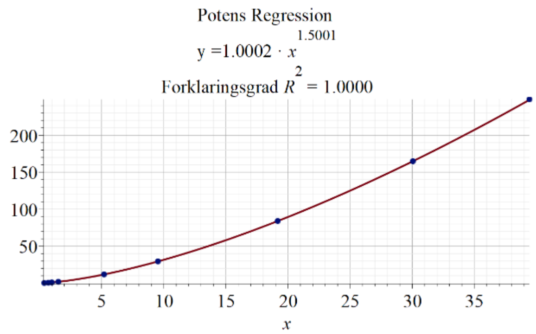
ISBN 9788770668781

Løsninger til øvelser i kapitel 9

Øvelse 9.1

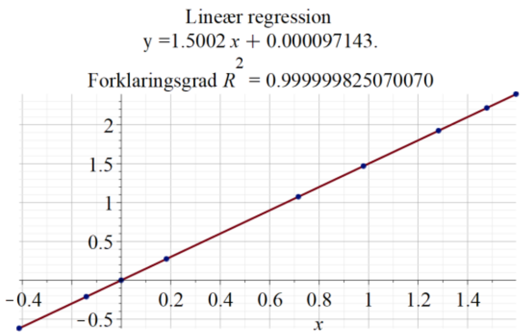
Potensregression – vi bemærker, at konstanten ”skulle” være 1, ikke 1.002, og eksponenten 1.5, ikke 1.5001



Linearisering: Først omregnes til log-værdier i begge søjler:

-0,412	-0,618
-0,141	-0,211
0,000	0,000
0,183	0,274
0,716	1,074
0,980	1,469
1,283	1,924
1,478	2,217
1,596	2,395

Så gennemføres lineær regression på denne tabel – hvor vi bemærker i forlængelse af ovst., at den ”skulle” være $y = 1.5x$



Øvelse 9.2

a)

Vi har fået opgivet halvakserne på de 6 inderste planeter: T = [0.39, 0.72, 1, 1.52, 5.2, 9.54]

Den sidste, Saturns halvakse skaleres op til 100. Så bliver de øvrige:

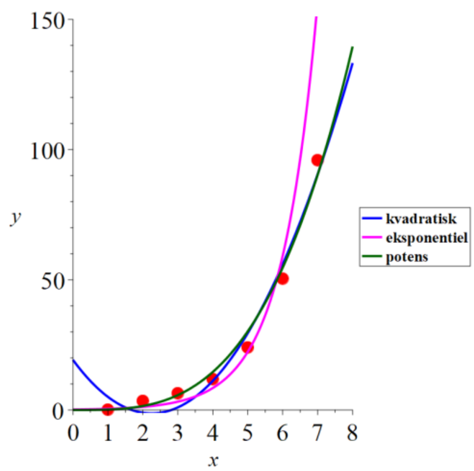
$$\frac{100}{9.54} \cdot T = [4.088,\ 7.547,\ 10.48,\ 15.93,\ 54.51,\ 100.0]$$

b)

Trækkes den første, Merkurs halvakse fra får vi listen:

[0., 3.459, 6.392, 11.842, 50.422, 95.912]

De første tal i listen kunne ligne en følge af fordoblinger. Men efter 11,8 går det galt. På den anden side – var der her en planet med halvakse ca. 25, så passer systemet godt.



Indsættes en sådan værdi, og har vi således tallene 1..7 på *x*-aksen, får vi følgende grafbillede af tre mulige modeller:

Bedømt ud fra R^2 er den kvadratiske bedst (R^2=0,98), dernæst potens (R^2=0,97), og sidst eksponentiel (R^2=0,86

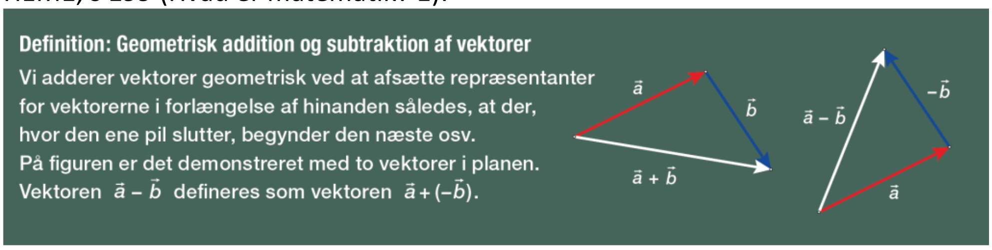
c)

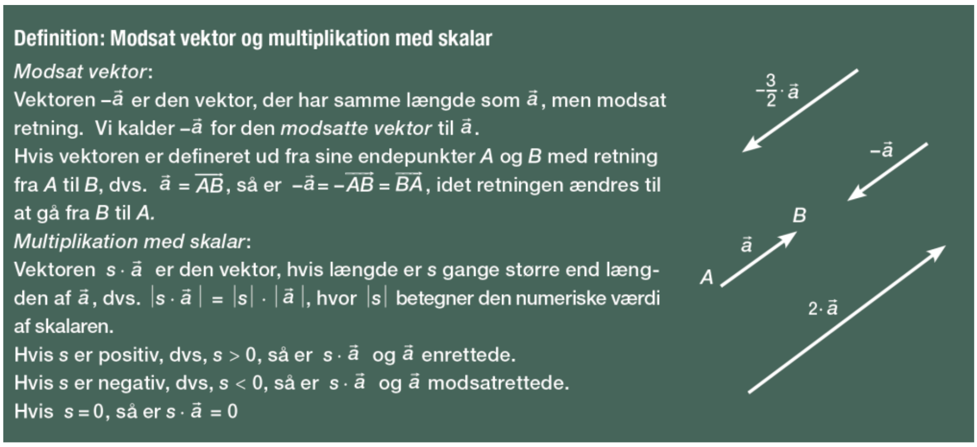
Svaret er som i b): ca. 25 enheder (+ Merkurs halvakse)

d)

Efter Saturn skulle de næste planeter være ca. 200 enheder (+ Merkurs halvakse), ca. 400 enheder (+ Merkurs halvakse).

ISBN 9788770668781

<table>
<tr><td></td><td>e)
Mellem Merkur og Venus: Det må være i 1,75 enheder (+ Merkurs halvakse), da det dobbelte skal være ca. 3,5.
f)
Vi anvender Fit-kommandoen i Maple:
$$\mathrm{Fit}(a+b\cdot k^t, N, A, t) = 2.708 + 1.059 \cdot 1.908^t$$
Vi sammenligner funktionsværdierne med de faktiske planetafstande:
Model: [4.729, 6.563, 10.06, 16.74, 29.49, 53.81, 100.2]
Empiri: [4.088, 7.547, 10.48, 15.93, 29, 54.51, 100.0]
hvor vi har tilføjet tallet 29, der repræsenterer ”en planet” mellem Mars og Jupiter (hvor asteroidebæltet er).
- Første indtryk: Det er en forbløffende god model.
- Men: Der er kun 6 empiriske værdier her.
- Da Uranus blev opdaget, viste det sig, at den passede udmærket ind i systemet. Men da Neptun blev opdaget, og desværre passede rigtig dårligt ind i systemet, var formlen skudt i stykke.
- Der er ingen teoretisk - fysisk argumentation for denne ”lov”</td></tr>
<tr><td>Øvelse 9.3</td><td>a)
Når ligningssystemet er bragt på en form som det trekantede – en form der hedder Echelon-form - løser vi det nedefra:
$$60\cdot x_4 = 6300 \Rightarrow x_4 = \tfrac{6300}{60} = 105$$
$$16x_3 + 4x_4 = 2212 \Rightarrow 16x_3 + 4\cdot 105 = 2212 \Rightarrow 16x_3 = 1792 \Rightarrow x_3 = \tfrac{1792}{16} = 112$$
osv.
b)
Enhver operation vi har udført på ligningssystemet har en omvendt operation. Så vi kan komme tilbage fra det trekantede ligningssystem til det oprindelige. En løsning til det trekantede er derfor også en løsning til det oprindelige.</td></tr>
<tr><td>Øvelse 9.4</td><td>Svaret på denne øvelse indgår i projekt 9.2.</td></tr>
<tr><td>Øvelse 9.5</td><td>a)
HEM1, s 199 (Hvad er matematik? 1):

Definition: Geometrisk addition og subtraktion af vektorer
Vi adderer vektorer geometrisk ved at afsætte repræsentanter for vektorerne i forlængelse af hinanden således, at der, hvor den ene pil slutter, begynder den næste osv.
På figuren er det demonstreret med to vektorer i planen.
Vektoren $\vec{a}-\vec{b}$ defineres som vektoren $\vec{a}+(-\vec{b})$.

HEM1 s. 200:

Definition: Modsat vektor og multiplikation med skalar
Modsat vektor:
Vektoren $-\vec{a}$ er den vektor, der har samme længde som $\vec{a}$, men modsat retning. Vi kalder $-\vec{a}$ for den modsatte vektor til $\vec{a}$.
Hvis vektoren er defineret ud fra sine endepunkter A og B med retning fra A til B, dvs. $\vec{a}=\overrightarrow{AB}$, så er $-\vec{a}=-\overrightarrow{AB}=\overrightarrow{BA}$, idet retningen ændres til at gå fra B til A.
Multiplikation med skalar:
Vektoren $s\cdot\vec{a}$ er den vektor, hvis længde er s gange større end længden af $\vec{a}$, dvs. $|s\cdot\vec{a}| = |s|\cdot|\vec{a}|$, hvor $|s|$ betegner den numeriske værdi af skalaren.
Hvis s er positiv, dvs. $s>0$, så er $s\cdot\vec{a}$ og $\vec{a}$ enrettede.
Hvis s er negativ, dvs. $s<0$, så er $s\cdot\vec{a}$ og $\vec{a}$ modsatrettede.
Hvis $s=0$, så er $s\cdot\vec{a}=0$

b)
HEM1, s. 199: Associative og Kommutative lov</td></tr>
</table>

ISBN 9788770668781

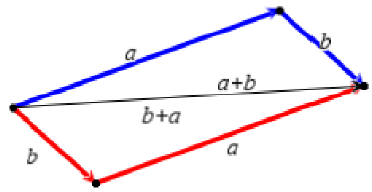
Sætning 1: Regneregler for addition

For alle vektorer $\vec{a}$, $\vec{b}$ og $\vec{c}$ gælder:

1. Vektoradditionen $\vec{a} + \vec{b}$ er uafhængig af, hvilke repræsentanter vi vælger for vektorerne. Vi siger, at addition er *veldefineret*.
2. $\vec{a} + \vec{b} = \vec{b} + \vec{a}$ Vi siger, at den *kommutative* lov gælder for vektoraddition.
3. $\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$ Vi siger, at den *associative* lov gælder for vektoraddition.

Argument for kommutative lov:

De to vektorer bestemmer et parallelogram, hvor modstående sider er parallelle, dvs. har samme retning, og har samme længde. Da vektorer kan afsættes hvor som helst, repræsenterer de modstående sider samme to vektorer $\vec{a}$ og $\vec{b}$. Diagonalen i parallelogrammet kan derfor fremkomme både som $\vec{a}+\vec{b}$ og som $\vec{b}+\vec{a}$, så de to repræsenterer samme vektor.



(*Bemærk*: Beviset for hele sætning 1 ligger på website for HEM1, og kan tilgås via en QR-kode hvor sætningen er formuleret, s 199.)

HEM1, s. 201: *Distributive lov*

Sætning 3: Regneregler for skalarmultiplikation

For alle vektorer $\vec{a}$ og $\vec{b}$, og alle skalarer s og t gælder de distributive love:

1. $s \cdot (\vec{a} + \vec{b}) = s \cdot \vec{a} + s \cdot \vec{b}$
2. $(s + t) \cdot \vec{a} = s \cdot \vec{a} + t \cdot \vec{a}$

Argument for punkt 1:

Beviset bygger på teorien for ensvinklede trekanter, som er gennemgået i HEM1, kapitel 6, afsnit 5 (s. 211ff). Resultatet fremkommer ved at skalere trekanten dannet af $\vec{a}$, $\vec{b}$ og $\vec{a}+\vec{b}$ op med faktoren *s*, og anvende definitionen på vektoraddition: Ved at skalere op har vi stadig en trekant. De to af siderne er $s\cdot\vec{a}$ og $s\cdot\vec{b}$. Ved vektoraddition af disse fås den tredje side. Men den tredje side er jo netop $s\cdot\left(\vec{a}+\vec{b}\right)$. Dette giver formlen.

Øvelse 9.6

a)

$|\vec{a}|=\sqrt{5^2+12^2+0^2}=13$ $\quad |\vec{b}|=\sqrt{1^2+2^2+(-4)^2}=\sqrt{21}$ $\quad |\vec{c}|=\sqrt{(-3)^2+6^2+2^2}=7$

b)

Undersøg selv dit værktøj

Øvelse 9.7

a)

Ifølge sætning 5 s. 204 i HEM1 har *Forbindelsesvektoren* $\overrightarrow{AB}$ koordinaterne $\begin{pmatrix} b_1-a_1 \\ b_2-a_2 \\ b_3-a_3 \end{pmatrix}$.

Ved at indsætte i formlen i sætning 1 får vi formlen

b)

$A(2,3,6)$ og $B(5,-4,-2)$. *Forbindelsesvektoren* $\overrightarrow{AB}$ har længden:

$$\left|\overrightarrow{AB}\right|=\sqrt{(5-2)^2+(-4-3)^2+(-2-6)^2}=\sqrt{3^2+7^2+8^2}=\sqrt{122}$$

Øvelse 9.8

a)

L angiver søjlen med vægt-data. Deskriptorerne kan beregnes med formlerne direkte i regnearket. Her er data ført ind i værktøjsprogrammet (Maple) og beregnet med kommandoer dér:

middelværdi(L) = 88.28636363

spredning(L) = 11.6359277256300

ISBN 9788770668781

varians(L)= 135.394814036084

b)

Interval med 1 sprednings bredde: $]76.65;99.92[$. Andel af obs: 68,2%

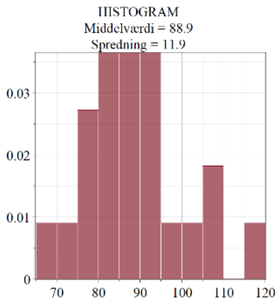
Interval med 2 spredningers bredde: $]65.01;111.56[$. Andel af obs: 95.5%

Interval med 3 spredningers bredde: $]53.38;123.19[$. Andel af obs: 100%

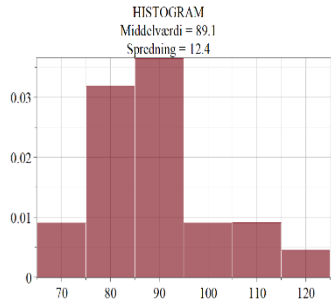
c)

(tegnet i Maple):

Histogram med intervalbredde 5

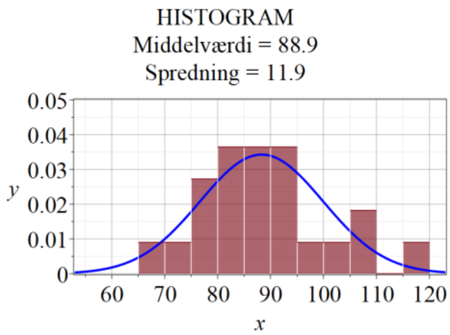


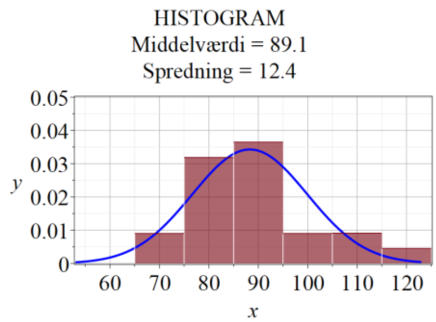
Histogram med intervalbredde 10



d)

Tegner med overlejret normalfordeling med middelværdi 88,29 og spredning 11,64:





Man kan godt sige, at vægten med tilnærmelse er normalfordelt. Det er tydeligt af det grafiske billede og af tabellerne, at der ikke er fuld symmetri: Der er en større tendens mod høje tal end lave tal. Der er ikke så mange letvægtshåndboldspillere

Øvelse 9.9

Middeltallet af talsættet $\{z_1,z_2,\ldots,z_n\}$ er $\bar{z}=\dfrac{z_1+z_2+\ldots+z_n}{n}$.

Antag nu, at $z_i=y_i-a\cdot x_i$, indsæt:

$$\begin{aligned}\bar{z}&=\frac{(y_1-a\cdot x_1)+(y_2-a\cdot x_2)+\ldots+(y_n-a\cdot x_n)}{n}\\&=\frac{y_1+y_2+\ldots+y_n-a\cdot x_1-a\cdot x_2-\ldots-a\cdot x_n}{n}\\&=\frac{y_1+y_2+\ldots+y_n-(a\cdot x_1+a\cdot x_2+\ldots+a\cdot x_n)}{n}\\&=\frac{y_1+y_2+\ldots+y_n-a\cdot(x_1+x_2+\ldots+x_n)}{n}\\&=\frac{y_1+y_2+\ldots+y_n}{n}-a\cdot\frac{(x_1+x_2+\ldots+x_n)}{n}=\bar{y}-a\cdot\bar{x}\end{aligned}$$

parentesregler og brøkregler

hvor vi har brugt en del parentesregler og brøkregler. Så $\bar{z}=\bar{y}-a\cdot\bar{x}$

Øvelse 9.10

a)

Sætning 3: Den bedste rette linje $y=\hat{a}\cdot x+\hat{b}$, hvor $\hat{b}=\bar{y}-\hat{a}\cdot\bar{x}$. Indsæt denne i y:

ISBN 9788770668781

L&R Uddannelse EGMONT

$$y=\hat{a}\cdot x+\hat{b}=\hat{a}\cdot x+\bar{y}-\hat{a}\cdot\bar{x}=\hat{a}\cdot(x-\bar{x})+\bar{y}$$

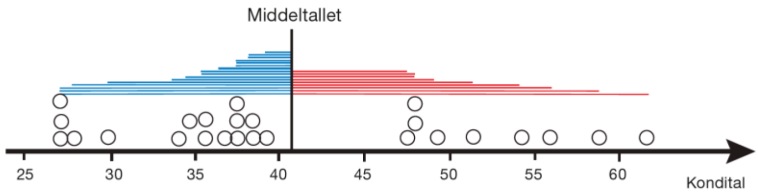
b)

Indsæt $\bar{x}$ i $y=\hat{a}\cdot(x-\bar{x})+\bar{y}$: $y=\hat{a}\cdot(\bar{x}-\bar{x})+\bar{y}=\bar{y}$, dvs. punktet $(\bar{x},\bar{y})$ ligger på linjen

c)

I HEM1, kap 2 om beskrivende statistik, argumenteres på s 70 for, at middeltallet kan opfattes som *tyngdepunkt* for et datasæt. Argumentet gennemføres både geometrisk og beregningsmæssigt.

Geometrisk:



Middeltallet splitter datasættet i to ulige store delsæt med 16 observationer (blå) til venstre for middeltallet og 9 observationer (røde) til højre for middeltallet. Men de 9 observationer ligger længere væk fra middeltallet, så tilsammen vægter de lige så meget som de 16 observationer, der ligger til venstre. Vi siger at fordelingen er højre-skæv med en lang hale af observationer til højre.

Det er disse forholdsvis store kondital, der trækker gennemsnittet i vejret i forhold til medianen.

Beregningsmæssigt:

Eksempel: Middeltallet som balancepunkt for afstandene (især for A-niveau)

Den følgende udregning kan gennemføres helt tilsvarende for ethvert datasæt.
Ud fra definitionen er middeltallet m bestemt som det tal, der opfylder:

$$25\cdot m=27{,}1+27{,}1+\ldots+58{,}7+61{,}7$$
$$m+m+\ldots+m=27{,}1+27{,}1+\ldots+58{,}7+61{,}7$$

Vi flytter nu alle tal, der er mindre end m, over på venstre side og parrer dem hver for sig sammen med et m. Resten af m'erne flyttes over på højre side:

$$(m-27{,}1)+(m-27{,}1)+\ldots+(m-39{,}1)=(47{,}4-m)+(47{,}9-m)+\ldots+(61{,}7-m)$$

Alle parenteser repræsenterer nu afstanden fra en observation hen til middeltallet. Ligningen siger, at summen af afstandene fra de observationer, der ligger til venstre for middeltallet, er lig med summen af afstandene fra de observationer, der ligger til højre for middeltallet.

Da denne udregning kunne gennemføres for ethvert datasæt, er konklusionen generel:
Middeltallet er balancepunktet for afstandene til observationerne.

Øvelse 9.11

a)

Ifølge øvelse 9.9 kan middeltal af en flerleddet størrelse udregnes led for led, så:

$$\hat{y}=\hat{a}\cdot x+\hat{b}\Rightarrow\bar{\hat{y}}=\overline{\hat{a}\cdot x+\hat{b}}\Rightarrow\bar{\hat{y}}=\overline{\hat{a}\cdot x}+\bar{\hat{b}}=\hat{a}\cdot\bar{x}+\hat{b}\,,$$

da middeltallet af en konstant - som $\hat{a}$ - er konstanten selv.

b)

Øvelse 9.12

Resultaterne af a), b), c) er alle givet i øvelsen

d)

Ved at lægge formlerne i sætning 3 ind i regnearket med data, får vi: $y=-5{,}2668x+115{,}3678$

e)

Residualplottet er givet i øvelsen, og kommentaren er givet som billedtekst til plottet.

Øvelse 9.13

Dette er projekt 9.12

Øvelse 9.14

Planen udspændt af $\boldsymbol{u}$ og $\boldsymbol{v}$ betegnes: span($\boldsymbol{u}$, $\boldsymbol{v}$). Den består af alle *linearkombinationer*:

$$s\cdot\boldsymbol{u}+t\cdot\boldsymbol{v}$$

ISBN 9788770668781

L&R Uddannelse EGMONT

	a) Vi har af foregående side: $\boldsymbol{x}_\perp = \boldsymbol{x} - \overline{x}\cdot\boldsymbol{d} = \boldsymbol{x} + (-\overline{x})\cdot\boldsymbol{d}$, med det er jo en vektor i planen udspændt af ***x*** og ***d***. Så $\boldsymbol{x}_\perp$ ligger i span(***x*** **,** ***d***) b) En linearkombination: $s\cdot\boldsymbol{x}_\perp + t\cdot\boldsymbol{d}$ ligger også i span(***x*** , ***d***), fordi: $s\cdot\boldsymbol{x}_\perp + t\cdot\boldsymbol{d} = s\cdot\left(\boldsymbol{x} - \overline{x}\cdot\boldsymbol{d}\right) + t\cdot\boldsymbol{d} = s\cdot\boldsymbol{x} - s\cdot\overline{x}\cdot\boldsymbol{d} + t\cdot\boldsymbol{d} = s\cdot\boldsymbol{x} + \left(t - s\cdot\overline{x}\right)\cdot\boldsymbol{d}$. altså det er skrevet som en linearkombination af ***x*** og ***d*****.** Og dermed en vektor i span(***x*** , ***d***). c) Ligningen $\boldsymbol{x}_\perp = \boldsymbol{x} - \overline{x}\cdot\boldsymbol{d}$ kan omskrives til: $\boldsymbol{x} = \boldsymbol{x}_\perp + \overline{x}\cdot\boldsymbol{d}$, altså ***x*** er skrevet som en linearkombination af $\boldsymbol{x}_\perp$ og ***d***, dvs. ***x*** ligger i span($\boldsymbol{x}_\perp$, ***d***). d) En linearkombination: $s\cdot\boldsymbol{x} + t\cdot\boldsymbol{d}$ ligger også i span($\boldsymbol{x}_\perp$, ***d***), fordi: $s\cdot\boldsymbol{x} + t\cdot\boldsymbol{d} = s\cdot\left(\boldsymbol{x}_\perp + \overline{x}\cdot\boldsymbol{d}\right) + t\cdot\boldsymbol{d} = s\cdot\boldsymbol{x}_\perp + s\cdot\overline{x}\cdot\boldsymbol{d} + t\cdot\boldsymbol{d} = s\cdot\boldsymbol{x}_\perp + \left(s\cdot\overline{x} + t\right)\cdot\boldsymbol{d}$ altså det er skrevet som en linearkombination af $\boldsymbol{x}_\perp$ og ***d*****.** Og dermed en vektor i span($\boldsymbol{x}_\perp$, ***d***). *Den samlede konklusion er derfor:* *Alt i span(**x** , **d**) ligger i span($\boldsymbol{x}_\perp$, **d**), og alt i span($\boldsymbol{x}_\perp$, **d**) ligger i span(**x** , **d**), dvs.:* *span(**x** , **d**) = span($\boldsymbol{x}_\perp$, **d**)*
Øvelse 9.15	Vi henter Anscombes data og udfører lineær regression på alle 4. Du kan finde et resultat heraf via **website**. Du vil finde de overraskende resultater, at de 4 meget forskellige datasæt, med vidt forskellige grafer, har samme middelværdi, samme spredning, samme lineære regressionslinje, samme forklaringsgrad, samme korrelationskoefficient. Men det betyder omvendt, at tal som r^2 ikke kan anvendes isoleret til at konkludere noget om styrken af en sammenhæng mellem givne data.