

## 2 Verwerken van data

2

### Voorkennis Statistiek

#### Bladzijde 48

- 1 Als je kijkt naar de hoogtetoename van de staven dan lijkt er sprake te zijn van een afvlakking: de hoogtetoename van de derde naar de vierde en van de vierde naar de vijfde staaf is immers lang niet zo groot als die van de eerste naar de tweede en van de tweede naar de derde staaf. De eerste drie staven gaan echter over gehele jaren, terwijl de laatste twee staven over kwartalen gaan. Je mag ze daarom niet zo vergelijken.

#### Bladzijde 49

- 2 a De Telegraaf heeft 560 duizend abonnees en de Volkskrant heeft 230 duizend abonnees.  
Dus  $\frac{560}{230} \approx 2,4$  keer zoveel.
- b De Telegraaf heeft 560 duizend abonnees.  
Samen hebben de Volkskrant, NRC en Trouw  $230 + 190 + 110 = 530$  duizend abonnees.  
Dat is  $\frac{560 - 530}{530} \times 100\% \approx 5,7\%$  meer.
- 3 a In 2004 was de omzet 420 miljoen en in 2014 was de omzet 180 miljoen.  
De procentuele verandering is  $\frac{180 - 420}{420} \times 100\% \approx -57,1\%$ .  
Dus een afname van 57,1%.
- b De procentuele verandering in de periode 2004-2008 bedroeg  $\frac{280 - 420}{420} \times 100\% \approx -33,3\%$ .  
De procentuele verandering in de periode 2008-2014 bedroeg  $\frac{180 - 280}{280} \times 100\% \approx -35,7\%$ .  
Dus in de periode 2008-2014 is de omzet procentueel het meest afgenomen.
- 4 a De stip van 2014 ligt vier keer zo hoog als de stip van 2008.  
b De toename is  $\frac{83 - 80}{80} \times 100\% \approx 3,8\%$ .  
c Er had een scheurlijn gebruikt moeten worden.

#### Bladzijde 50

- 5 a 8% heeft bloedgroep B, dus je verwacht  $0,08 \cdot 250 = 20$  mensen met bloedgroep B.  
b Sector A heeft 42%.  
Dus  $0,42 \times 360^\circ \approx 151^\circ$ .
- 6 a Het is vaak minder nauwkeurig doordat het vereenvoudigd is.  
b 15% komt overeen met 16,5 miljoen.  
Dus totale beroepsbevolking is  $\frac{16,5}{0,15} = 110$  miljoen.  
c Italië heeft  $0,40 \cdot 24,6 = 9,84$  miljoen vakbondsleden.  
Duitsland heeft  $0,30 \cdot 37,6 = 11,28$  miljoen vakbondsleden.  
Dus Duitsland heeft de meeste vakbondsleden, namelijk  $\frac{11,28 - 9,84}{9,84} \times 100\% \approx 14,6\%$  meer dan Italië.
- 7 a De lengte (en de breedte) van het biljet bij 2014 is vier keer zo groot als bij het biljet van 2013.  
b De oppervlakte van het biljet bij 2014 is 16 keer zo groot als bij het biljet van 2013. Daardoor lijkt het of de winst 16 keer zo groot is.

**Bladzijde 51**

- 8 a De relatieve frequentie van bloedgroep A is  $\frac{12}{28} \times 100\% \approx 42,9\%$ .
- b Bij de drie leerlingen met bloedgroep B hoort 15%.  
Dus  $\frac{3}{0,15} = 20$  leerlingen in totaal.
- c Bij de 48 000 Surinamers in Den Haag hoort 9,4%.  
Dus het aantal inwoners van Den Haag is  $\frac{48000}{0,094} \approx 511\,000$ .

**Bladzijde 53**

- 9 a totale frequentie =  $3 + 3 + 7 + 0 + 3 + 2 + 2 = 20$ , dus het onderzoek duurde 20 dagen.
- b gemiddelde =  $\frac{3 \times 1 + 7 \times 2 + 3 \times 4 + 2 \times 5 + 2 \times 6}{20} = 2,55$   
mediaan =  $\frac{10^{\text{e}} \text{ getal} + 11^{\text{e}} \text{ getal}}{2} = \frac{2 + 2}{2} = 2$   
modus = 2, want 2 heeft de grootste frequentie.
- 10 a totale frequentie =  $2 + 4 + 5 + 3 + 3 + 2 + 1 = 20$   
gemiddelde =  $\frac{4 \times 1 + 5 \times 2 + 3 \times 3 + 3 \times 4 + 2 \times 5 + 1 \times 32}{20} = 3,85$   
mediaan =  $\frac{10^{\text{e}} \text{ getal} + 11^{\text{e}} \text{ getal}}{2} = \frac{2 + 2}{2} = 2$   
modus = 2
- b De klas wijst op de mediaan die voor beide klassen gelijk is.
- c Het gemiddelde is gevoelig voor uitschieters, de mediaan en modus zijn dat niet.

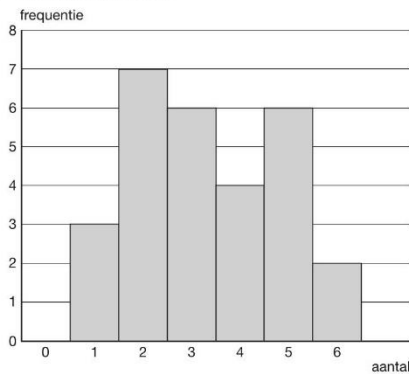
**2.1 Frequentieverdelingen**

**Bladzijde 54**

- 1 a Er zijn 6 leerlingen met 2 broers of zussen.
- b Elke leerling is ondervraagd, dus  $5 + 12 + 6 + 2 + 2 + 1 = 28$  leerlingen.
- c De relatieve frequentie is  $\frac{12}{28} \times 100\% \approx 42,9\%$ .

**Bladzijde 55**

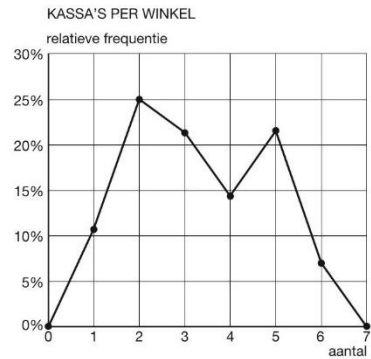
- 2 a KASSA'S PER WINKEL



b De totale frequentie is  $3 + 7 + 6 + 4 + 6 + 2 = 28$ .

$$\text{relatieve frequentie} = \frac{\text{frequentie}}{28} \times 100\%$$

aantal	relatieve frequentie
1	10,7
2	25,0
3	21,4
4	14,3
5	21,4
6	7,1



c Er zijn 10 winkels met minder dan drie kassa's.

$$\text{Dus } \frac{10}{28} \times 100\% \approx 35,7\% \text{ van de winkels heeft minder dan drie kassa's.}$$

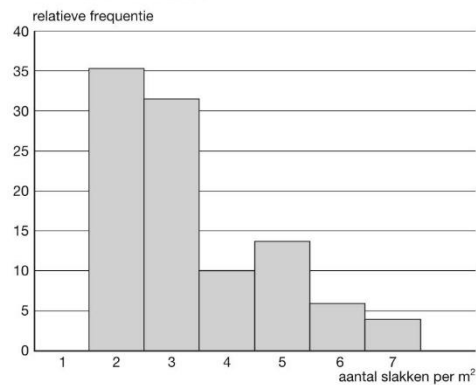
Er zijn 25 winkels met minstens twee kassa's.

$$\text{Dus } \frac{25}{28} \times 100\% \approx 89,3\% \text{ van de winkels heeft minstens twee kassa's.}$$

3 a

aantal slakken per m <sup>2</sup>	frequentie	relatieve frequentie
2	18	35,3%
3	16	31,4%
4	5	9,8%
5	7	13,7%
6	3	5,9%
7	2	3,9%
	51	

SLAKKEN VOLKSTUUNTJE



b Oppervlakte volkstuintje is  $18 + 16 + 5 + 7 + 3 + 2 = 51 \text{ m}^2$ .

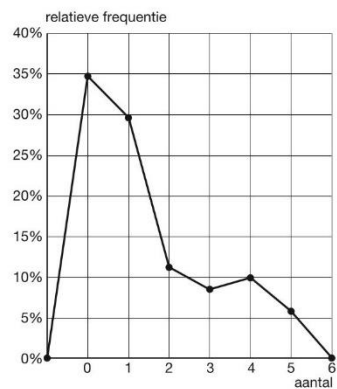
c Het totale aantal slakken is  $18 \times 2 + 16 \times 3 + 5 \times 4 + 7 \times 5 + 3 \times 6 + 2 \times 7 = 171$ .

4 a Het aantal adressen is  $148 + 126 + 48 + 36 + 42 + 26 = 426$ .

b

aantal	relatieve frequentie
0	34,7%
1	29,6%
2	11,3%
3	8,5%
4	9,9%
5	5,9%

POSTSTUKKEN PER ADRES



c Minder dan drie poststukken bij  $148 + 126 + 48 = 322$  adressen.

Dat is  $\frac{322}{426} \times 100\% \approx 75,6\%$  van de adressen.

d Hij heeft  $126 \times 1 + 48 \times 2 + 36 \times 3 + 42 \times 4 + 26 \times 5 = 628$  poststukken bezorgd.

5 Het geeft geen beter overzicht, want elk waarnemingsgetal heeft frequentie 1.

**Bladzijde 57**

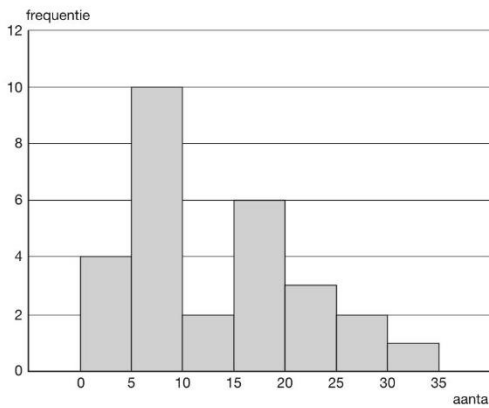
6 a WERKNEMERS PER WINKEL

aantal	frequentie
0 -< 5	4
5 -< 10	10
10 -< 15	2
15 -< 20	6
20 -< 25	3
25 -< 30	2
30 -< 35	1

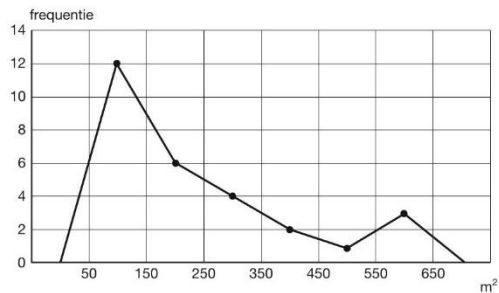
c VLOEROPPERVERLAKTE

m <sup>2</sup>	frequentie
50 -< 150	12
150 -< 250	6
250 -< 350	4
350 -< 450	2
450 -< 550	1
550 -< 650	3

b WERKNEMERS PER WINKEL



d VLOEROPPERVERLAKTE



- 7 a Het aantal 16 komt 2 keer voor.
- b Het kleinste aantal is 3 werknemers.
- c Het aantal 8 komt het vaakst voor.
- d De klassen zijn 0 -< 10, 10 -< 20, 20 -< 30 en 30 -< 40.

**Bladzijde 58**

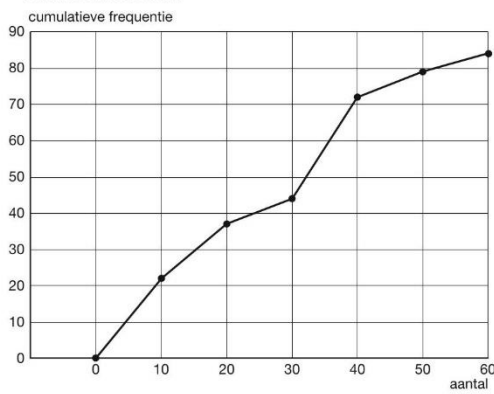
- 8 a Tellen geeft 16 mannen en 16 vrouwen. Dus 16 echtparen.
- b Er was 1 man van 33 jaar en 2 vrouwen waren 35 jaar oud.
- c Aantal personen van 40 jaar of ouder is  $3 + 2 + 6 = 11$ .
- d Aantal mannen jonger dan 30 jaar is  $1 + 5 = 6$ .  
Aantal vrouwen jonger dan 30 jaar is  $4 + 4 = 8$ .

**Bladzijde 59**

**9 a ARTIKELEN PER KLANT**

aantal	cumulatieve frequentie
0 –< 10	22
10 –< 20	37
20 –< 30	44
30 –< 40	72
40 –< 50	79
50 –< 60	84

**b ARTIKELEN PER KLANT**



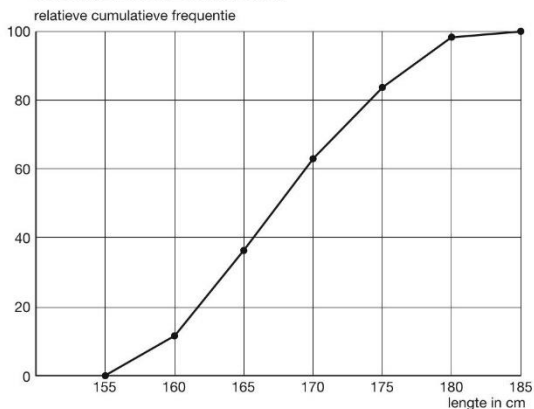
c 44 klanten kochten minder dan 30 artikelen.

d  $7 + 5 = 12$  klanten kochten minstens 40 artikelen.

**10 a BEVOLKINGSONDERZOEK VROUWEN**

lengte in cm	frequentie	cumulatieve frequentie	relatieve cumulatieve frequentie
155 –< 160	538	538	11,8%
160 –< 165	1135	1673	36,6%
165 –< 170	1218	2891	63,2%
170 –< 175	941	3832	83,8%
175 –< 180	657	4489	98,2%
180 –< 185	83	4572	100%

**b** BEVOLKINGSONDERZOEK VROUWEN

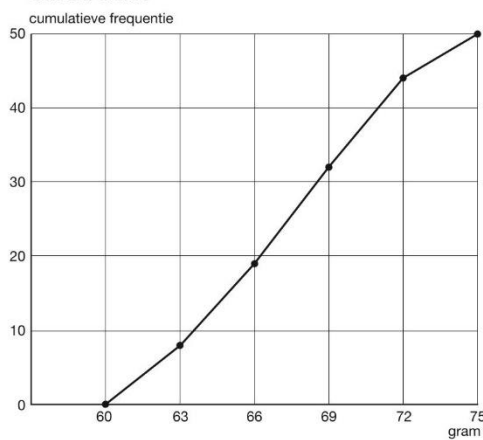


**Bladzijde 60**

**11** GEWICHT EIEREN

gram	frequentie	cumulatieve frequentie
60 – < 63	8	8
63 – < 66	11	19
66 – < 69	13	32
69 – < 72	12	44
72 – < 75	6	50

GEWICHT EIEREN



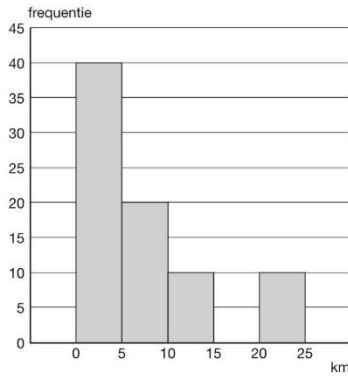
**12** De frequentie die je bij de rechtergrens afleest heeft dan de betekenis ‘aantal met deze waarde of lager’. Zie bijvoorbeeld de polygoon in opgave 11. In de klasse 60 – < 63 zitten 8 eieren. De waarnemingen in deze klasse liggen verspreid tussen 60 en 63 gram. Als het punt bij het klassenmidden 61,5 uitgezet zou zijn, zou het net lijken of alle 8 eieren minder dan 61,5 gram zouden wegen.

- 13** a Lees af: 80 leerlingen.  
 b Lees af: 60 leerlingen.  
 c Lees af: 40 leerlingen.  
 d  $70 - 60 = 10$  leerlingen

e

afstand in km	frequentie
0 –< 5	40
5 –< 10	20
10 –< 15	10
15 –< 20	0
20 –< 25	10

AFSTAND HUIS-SCHOOL



- 14 a Lees af: minder dan 100 seconden geeft 60%.  
Dat zijn  $0,60 \cdot 300 = 180$  klanten.
- b Lees af: minder dan 200 seconden geeft 85%.  
Dus  $0,85 \cdot 300 = 255$  klanten.
- c 2,5 minuut = 150 seconden  
Lees af: minstens 150 seconden geeft:  $100\% - 70\% = 30\%$ .  
Dus  $0,30 \cdot 300 = 90$  klanten.
- d Lees af: minstens 100 seconden, maar minder dan 200 seconden geeft  $85\% - 60\% = 25\%$ .  
Dus  $0,25 \cdot 300 = 75$  klanten.
- e WACHTTIJD BIJ KASSA

aantal seconden	frequentie
0 –< 50	60
50 –< 100	120
100 –< 150	30
150 –< 200	45
200 –< 250	30
250 –< 300	15

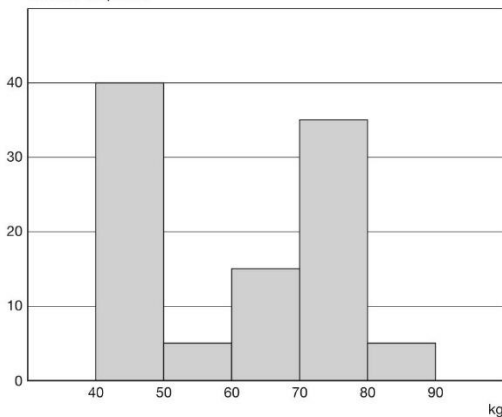
**Bladzijde 61**

- 15 a Van 10:00 uur tot 20:00 uur is  $10 \cdot 4 = 40$  kwartieren.
- b Lees bij 10 klanten af 20%.  
Dus  $0,20 \cdot 40 = 8$  kwartieren.
- c Lees bij 20 klanten af 80%.  
Dus  $0,80 \cdot 40 = 32$  kwartieren.
- d Lees bij 15 klanten af 40%.  
Dus 40% van de kwartieren kwamen minder dan 15 klanten. Dan kwamen er dus  $100\% - 40\% = 60\%$  van de kwartieren minstens 15 klanten binnen.  
Dat zijn  $0,60 \cdot 40 = 24$  kwartieren.

- 16 a** Van 8.00 tot 20.00 uur is 12 uur.  
 Vijf dagen, dus het is gedurende  $5 \cdot 12 = 60$  uur bijgehouden.  
 Lees af:  $A$  en 30 klanten geeft 50%.  
 Dus  $0,50 \cdot 60$  uur = 30 uur.
- b** Lees af:  $B$  en 40 klanten geeft 20%.  
 Dus minstens 40 klanten is 80%.  
 Dat is  $0,80 \cdot 60 = 48$  uur.
- c** Lees af:  $B$  en 50 klanten geeft 30%.  
 30% van 5 dagen is 1,5 dagen.  
 Het kan dus hooguit 1,5 dag geweest zijn.
- d** Dat is niet waar. Dat de grafiek bij  $B$  minder snel stijgt op het begin betekent dat een klein percentage van de uren er weinig klanten waren. Het kwam bij  $B$  dus vaker voor dat er veel mensen tankten.
- 17 a**
- Lees bij I en 50 kg af 40%. Dus  $0,40 \cdot 200 = 80$  bomen.
  - Lees bij II en 60 kg af 25%. Dus  $0,25 \cdot 160 = 40$  bomen.
  - Lees bij I en 50 kg af 40% en bij 70 kg 60%. Dus  $60\% - 40\% = 20\%$  en  $0,20 \cdot 200 = 40$  bomen.
- b** Perceel II. Omdat de grafiek daar op het einde het meest toeneemt zie je dat bij de hogere opbrengsten de grootste percentages horen.
- c** OPBRENGST PERCEEL I

kg	relatieve cumulatieve frequentie	relatieve frequentie
40 – < 50	40	40
50 – < 60	45	5
60 – < 70	60	15
70 – < 80	95	35
80 – < 90	100	5

OPBRENGST PERCEEL I  
 relatieve frequentie



## 2.2 Centrummaten en variabelen

### Bladzijde 63

- 18 a** Nee, daarvoor moet je de metingen kunnen optellen en delen. Je hebt daarvoor getallen nodig. Wat is immers het gemiddelde van bijvoorbeeld fiets en bus? Dat heeft geen betekenis.
- b** De modus kan wel, dat is hier de bus.



**Bladzijde 64**

- 19 a Mediaan, vanwege de uitschieter.  
 b Modus, kwalitatieve gegevens.  
 c Gemiddelde of mediaan.  
 d Modus of mediaan, vanwege de uitschieter (een meetfout).
- 20 a kwalitatief  
 b kwantitatief  
 c kwalitatief  
 d kwantitatief  
 e kwalitatief
- 21 a De variabele 'aantal keer fastfood gegeten afgelopen week'.  
 Deze variabele is kwantitatief.  
 b De totale frequentie is  $2 + 4 + 6 + 5 + 4 + 4 + 3 + 2 = 30$ .  

$$\text{gemiddelde} = \frac{2 \cdot 0 + 4 \cdot 1 + 6 \cdot 2 + 5 \cdot 3 + 4 \cdot 4 + 4 \cdot 5 + 3 \cdot 6 + 2 \cdot 7}{30} = \frac{99}{30} = 3,3$$
  

$$\text{modus} = 2$$
  

$$\text{mediaan} = \frac{15^{\text{e}} \text{ getal} + 16^{\text{e}} \text{ getal}}{2} = \frac{3 + 3}{2} = 3$$
  
 c Van klein naar groot: modus, mediaan, gemiddelde.  
 d De totale frequentie is nu  $30 + 4 = 34$ .  
 $3,5$  keer fastfood per week geeft dus totaal  $34 \cdot 3,5 = 119$  keer fastfood. Het totaal was 99 keer fastfood.  
 De 4 leerlingen aten dus  $119 - 99 = 20$  keer fastfood.  
 De 4<sup>e</sup> leerling at dus  $20 - 3 \cdot 6 = 2$  keer fastfood afgelopen week.

**Bladzijde 65**

- 22 De totale frequentie is  $4 + 7 + 12 + 14 + 5 + 8 = 50$ .  
 De laagst mogelijke uitkomst krijg je door te rekenen met de linkergrenzen, dat geeft  

$$\text{gemiddelde} = \frac{4 \cdot 50 + 7 \cdot 55 + 12 \cdot 60 + 14 \cdot 65 + 5 \cdot 70 + 8 \cdot 75}{50} = 63,3 \text{ kg.}$$
  
 De hoogst mogelijke uitkomst krijg je door te rekenen met de rechtergrenzen, dat geeft  

$$\text{gemiddelde} = \frac{4 \cdot 55 + 7 \cdot 60 + 12 \cdot 65 + 14 \cdot 70 + 5 \cdot 75 + 8 \cdot 80}{50} = 68,3 \text{ kg.}$$
  
 Het gewicht 64,1 kg ligt tussen deze uitkomsten en kan dus kloppen.

**Bladzijde 66**

- 23 a discreet  
 b continu  
 c continu  
 d discreet  
 e continu  
 f discreet
- 24 a Discrete variabele, omdat prijzen in supermarkten op centen worden afgerond.  
 b De totale frequentie is  $26 + 17 + 21 + 8 + 13 + 2 = 87$ .  
 Door klassenmiddens te gebruiken krijg je  

$$\text{gemiddelde} = \frac{26 \cdot 50 + 17 \cdot 150 + 21 \cdot 250 + 8 \cdot 350 + 13 \cdot 450 + 2 \cdot 550}{87} \approx 217 \text{ cent.}$$
  
 c De mediaan is het middelste waarnemingsgetal. Bij 87 waarnemingsgetallen is dit de 44<sup>e</sup> meting.  
 Dus in de klasse  $200 - < 300$ .  
 d De modale klasse is de klasse  $0 - < 100$ , want die klasse heeft de grootste frequentie.
- 25 a Omdat de leeftijd in jaren gemeten is, gaat het hier om een discrete variabele.  
 b De totale frequentie is  $18 + 3 + 40 + 9 = 70$ .  
 Het kleinst mogelijke gemiddelde krijg je door te rekenen met de laagst mogelijke leeftijd per klasse,  
 dat geeft gemiddelde = 
$$\frac{18 \cdot 25 + 3 \cdot 35 + 40 \cdot 45 + 9 \cdot 55}{70} \approx 41 \text{ jaar.}$$
  
 Het grootst mogelijke gemiddelde krijg je door juist te rekenen met de hoogst mogelijke leeftijd per  
 klasse, dat geeft gemiddelde = 
$$\frac{18 \cdot 34 + 3 \cdot 44 + 40 \cdot 54 + 9 \cdot 64}{70} \approx 50 \text{ jaar.}$$
  
 c mediaan = 
$$\frac{35^{\text{e}} \text{ getal} + 36^{\text{e}} \text{ getal}}{2}$$
. Zowel het 35<sup>e</sup> als het 36<sup>e</sup> getal ligt in de klasse  $45 - < 55$ .  
 De mediaan ligt dus in de klasse  $45 - < 55$ .

- d In de klasse  $45 < 55$  vallen 10 leeftijden, namelijk 45, 46, 47, 48, 49, 50, 51, 52, 53 en 54.  
 Doordat deze klasse frequentie 40 heeft komt een leeftijd uit deze klasse minstens 4 keer voor.  
 De leeftijd 36 jaar valt in de klasse  $35 < 45$  met frequentie 3 en komt dus hoogstens 3 keer voor.

### 2.3 Spreidingsmaten

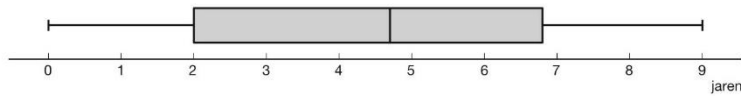
#### Bladzijde 68

- 26 Bij de mediaan van de eerste groep hoort een relatieve cumulatieve frequentie van 25%.  
 Bij de mediaan van de tweede groep hoort een relatieve cumulatieve frequentie van 75%.

#### Bladzijde 69

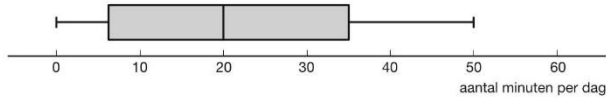
- 27 a Lees af:  $Q_1 = 2$ , mediaan = 4,7 en  $Q_3 = 6,8$ .

b OUDERDOM AUTO



- 28 Lees af:  $Q_1 = 6,25$ , mediaan = 20 en  $Q_3 = 35$ .  
 Aflazen bij 0% en 100% geeft respectievelijk 0 en 50.

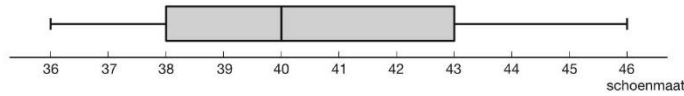
MOBIEL GEBRUIK LEERLINGEN



#### Bladzijde 70

- 29 Voer in lijst 1 = {36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46} en lijst 2 = {6, 12, 18, 11, 6, 4, 15, 21, 8, 3, 1}.  
 1-Var Stats (TI) of 1VAR (Casio) geeft  $\min X = 36$ ,  $Q_1 = 38$ ,  $\text{Med} = 40$ ,  $Q_3 = 43$  en  $\max X = 46$ .

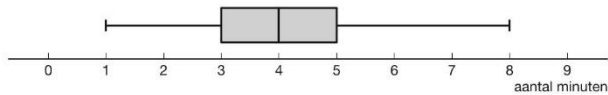
LEERLINGEN 4 HAVO



#### Bladzijde 71

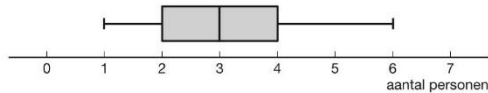
- 30 a Voer in lijst 1 = {1, 2, 3, 4, 5, 6, 8} en lijst 2 = {3, 5, 8, 10, 8, 7, 2}.  
 1-Var Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 4,1$ . De mediaan is 4 en de modus is 4.  
 b  $\min X = 0$ ,  $Q_1 = 3$ ,  $\text{Med} = 4$ ,  $Q_3 = 5$  en  $\max X = 8$ .

TE LAAT IN DE LES OP DONDERDAG 17 SEPTEMBER



- 31 a Voer in lijst 1 = {1, 2, 3, 4, 5, 6} en lijst 2 = {7, 19, 10, 12, 4, 1}.  
 1-Var Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 2,8$ . De mediaan is 3 en modus is 2.  
 b  $\min X = 1$ ,  $Q_1 = 2$ ,  $\text{Med} = 3$ ,  $Q_3 = 4$  en  $\max X = 6$ .

OMVANG HUISHOUDEN



- 32 a Gemeente West Maas en Waal ligt op  $Q_1$ . Boven  $Q_1$  ligt nog 75% van de gemeenten. Dat zijn dus  $0,75 \cdot 56 = 42$  gemeenten met meer inwoners dan West Maas en Waal.  
 b Bij elk deel van de boxplot hoort 25% van de gemeenten. Dat zijn  $0,25 \cdot 56 = 14$  gemeenten.  
 Van het 3<sup>e</sup> deel ligt  $\frac{10000}{13000} = \frac{10}{13}$  deel rechts van 30 000.  
 Dus  $\frac{10}{13} \cdot 14 + 14 \approx 25$  gemeenten met meer dan 30 000 inwoners.  
 c Van het 2<sup>e</sup> deel ligt  $\frac{2000}{9000} = \frac{2}{9}$  deel links van 20 000. Dus  $\frac{2}{9} \cdot 14 + 14 \approx 17$  gemeenten.  
 d Als schatting voor het gemiddelde aantal inwoners van een gemeente in een deel gebruiken we voor elk deel het gemiddelde van de linker- en rechtergrens.  
 Dat geeft  $\frac{18000 + 27000}{2} \cdot 14 + \frac{27000 + 40000}{2} \cdot 14 = 784000$  inwoners.  
 e Het gemiddelde aantal inwoners van het eerste deel is  $\frac{1500 + 18000}{2} = 9750$ .  
 In totaal zijn dit dus  $14 \cdot 9750 = 136500$  inwoners.  
 Dus het gevraagde percentage is  $\frac{136500}{1960000} \times 100\% \approx 7,0\%$ .

**Bladzijde 72**

- 33 a De mediaan is bij elke klas 3 km.  
 b Nee, de boxplots zijn duidelijk verschillend, maar de medianen zijn gelijk.  
 c De middelste 50% afstanden liggen in klas 4B dicht bij de mediaan dan in klas 4A.  
 d De spreiding is het grootst in klas 4A.  
 De spreiding is het kleinst in klas 4C.

**Bladzijde 73**

- 34 a De spreidingsbreedte is bij alle drie de leerlingen  $70 - 30 = 40$ .  
 b Bij A is de kwartielafstand  $44 - 35 = 9$   
 Bij B is de kwartielafstand  $43 - 37 = 6$ .  
 Bij C is de kwartielafstand  $55 - 32 = 23$ .  
 c Bij C is de spreiding het grootst omdat de kwartielafstand het grootst is.  
 d De spreidingsbreedte verandert, deze wordt  $80 - 30 = 50$ .  
 De kwartielafstand blijft 9.  
 e De spreidingsbreedte is gevoelig voor uitschieters.

- 35 Eerst 4H-C, dan 4H-A en als laatste 4H-B.

**Bladzijde 74**

- 36 Voer in lijst 1 = {1, 2, 3, 4, 5, 6, 7, 8} en lijst 2 = {4, 15, 19, 28, 25, 9, 4, 2}.  
 1-VAR Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 4,0$  personen per huishouden en  $\sigma \approx 1,5$  personen per huishouden.  
 37 Voer in lijst 1 = {7.45, 7.12, 6.99, 7.01, 6.9, 7.58, 7.63, 7.05, 7.41}.  
 1-VAR Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 7,22$  m en  $\sigma \approx 0,25$  m.

**Bladzijde 75**

- 38 Voer in lijst 1 = {150, 160, 170, 180} en lijst 2 = {19, 186, 244, 51}.  
 1-VAR Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 167$  cm en  $\sigma \approx 7$  cm.  
 39 a Het meest waarschijnlijk is 8 cm.  
 b Het meest waarschijnlijk is 1,8.  
 40 a Bij frequentie 50 geldt mediaan =  $\frac{25^{\text{e}} \text{ getal} + 26^{\text{e}} \text{ getal}}{2} = \frac{5,1 + 5,1}{2} = 5,1$ .  
 Dus de mediaan is 5,1 kg.  
 Mediaan 1<sup>e</sup> groep = 13<sup>e</sup> getal = 5,0 en mediaan 2<sup>e</sup> groep = 38<sup>e</sup> getal = 5,2.  
 De kwartielafstand is  $5,2 - 5,0 = 0,2$  kg en  
 de spreidingsbreedte is grootste getal - kleinste getal =  $5,4 - 4,8 = 0,6$  kg.  
 b Bij  $\sigma = 0,3$  zou  $2\sigma = 0,6$  zijn. Maar  $2\sigma$  kan niet gelijk zijn aan de spreidingsbreedte en die is hier 0,6.  
 c Voer in lijst 1 = {4.8, 4.9, 5, 5.1, 5.2, 5.3, 5.4} en lijst 2 = {2, 4, 10, 18, 12, 3, 1}.  
 1-VAR Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 5,09$  kg en  $\sigma \approx 0,12$  kg.



- 50 a Wat voor vlekken is bekeken? Andere wasmiddelen verdrijven wellicht 99% van dezelfde vlekken.  
 b 31 leerlingen van een schoolklas is geen representatieve steekproef voor de hele Nederlandse jeugd.  
 c In een nieuwbouwwijk wonen vaak andere mensen (jonger, meer welgesteld) dan in een oude stadskern.  
 d Er staat niet bij hoeveel en welke andere middelen er getest zijn.  
 e Het is interessanter om te weten hoe de regenval over een jaar is gespreid.  
 f De artsen kunnen de brillen cadeau gekregen hebben, of ze hebben een andere reden (mode, status) om de brillen te bezitten.  
 g De levensduur van een fietsband hangt niet af van de gebruikte fietspomp.

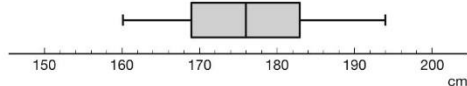
#### Bladzijde 82

- 51 a Uit tabel I zou je kunnen concluderen dat mannen meer kans hebben de ziekte op te lopen dan vrouwen.  
 b Uit tabel II blijkt dat 50% van de rokers en slechts 10% van de niet-rokers de ziekte heeft opgelopen, onafhankelijk van het geslacht. Niet het geslacht maar het roken is van invloed.
- 52 a Het gewicht van de zwaarste jongen is 75 kg.  
 Het gewicht van het lichtste meisje is 45 kg.  
 b Het verschil van het hoogste en laagste gewicht bij de jongens is 24 kg.  
 c Het gewicht 53 kg komt het meeste voor bij de meisjes. De modus is dus 53 kg.

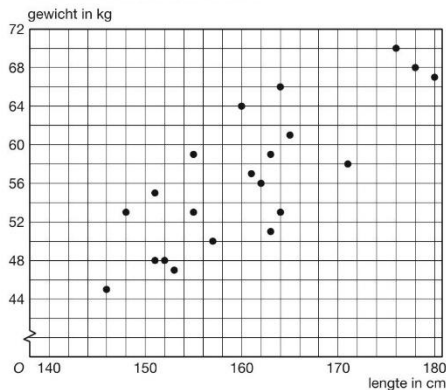
#### Bladzijde 83

- 53 a Lees af 194 cm.  
 Nee, de zwaarste jongen is 192 cm.  
 b De spreidingsbreedte is  $194 - 160 = 34$  cm.  
 c De modus is 165 cm (komt 3 keer voor).  
 mediaan = lengte 13<sup>e</sup> jongen = 176 cm  
 d Voer in lijst 1 = {160, 165, 165, 165, 166, 168, 169, 170, 170, ..., 192, 194}.  
 1-VAR Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 176$  cm en  $\sigma \approx 9$  cm.  
 e 1-VAR Stats (TI) of 1VAR (Casio) geeft  $Q_1 = 168,5$  cm, mediaan = 176 cm en  $Q_3 = 182,5$  cm.

LENGTE VAN 25 JONGENS



- 54 a GEWICHT EN LENGTE 21 MEISJES



- b Voer in lijst 1 = {61, 53, 57, ..., 51, 55}.  
 1-VAR Stats (TI) of 1VAR (Casio) geeft  $\bar{x} \approx 57$  kg en  $\sigma \approx 7$  kg.

#### Bladzijde 84

- 55 a De hoogst gemeten waarde is  $22^\circ\text{C}$  en de bijbehorende waarde volgens het model is  $17^\circ\text{C}$ .  
 b Het laagste maandgemiddelde volgens het model is  $-3^\circ\text{C}$ . In werkelijkheid was het deze maand gemiddeld  $3^\circ\text{C}$ .  
 c spreidingsbreedte gemeten =  $22 - -7 = 29$   
 spreidingsbreedte model =  $18 - -3 = 21$   
 Het scheidt  $\frac{29 - 21}{21} \times 100\% \approx 38,1\%$ .

- d Ja, de puntenwolk ligt redelijk dicht bij de lijn  $G = M$ .
- e Ja, de gemeten waarden beginnen te laag, doordat  $b = -1,334$ . Verdere schattingen zijn ook te laag, doordat de richtingscoëfficiënt minder dan 1 is.

## 2.5 Grafische verwerking met Excel

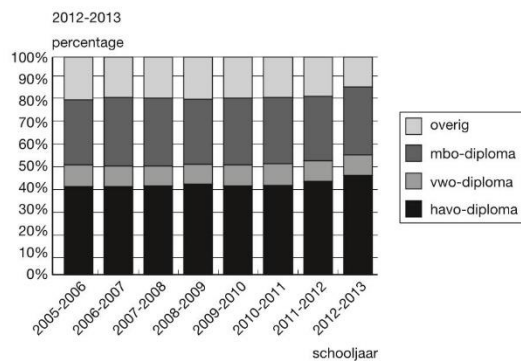
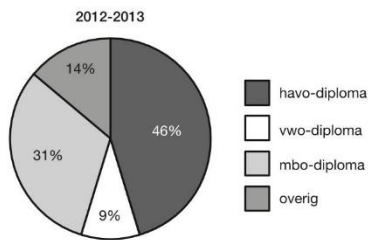
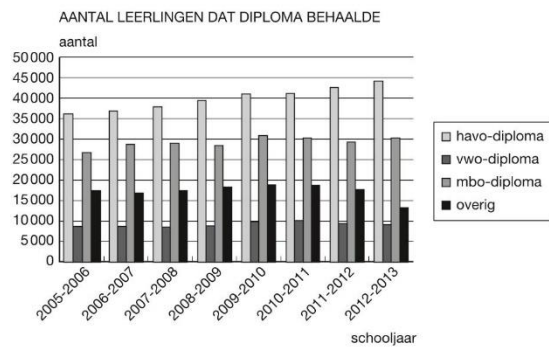
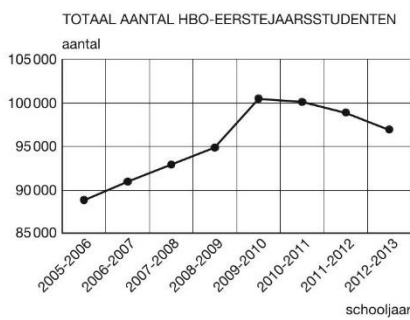
### Bladzijde 86

- 56 a In het boek staat per abuis laptop. Dit moet zijn smartphone, ook in het opschrift van figuur 2.35.
- b \*

### Bladzijde 87

- 57 \*
- 58 \*

- 59 Voorbeelden van een lijndiagram, staafdiagram, cirkeldiagram en stapeldiagram:



### Bladzijde 88

- 60 Voor de kwantitatieve gegevens geeft Excel de onderstaande waarden.

	A	B	C	D	E	F
1		gewicht in kg	aantal keer sporten per week	zakgeld per maand in euro's	afstand huis-school in km	gezinsomvang
2	gemiddelde	62,7	2,1	16,60	10,1	4
3	standaardafwijking	9,7	1,7	6,80	7,0	1,3
4	kleinste	46,1	0,0	6,00	1,0	2
5	Q1	57,0	0,8	11,50	4,8	3
6	mediaan	61,1	2,0	16,50	8,0	4
7	Q3	69,1	3,3	20,50	14,3	5
8	grootste	81,2	5,0	32,00	24,0	7

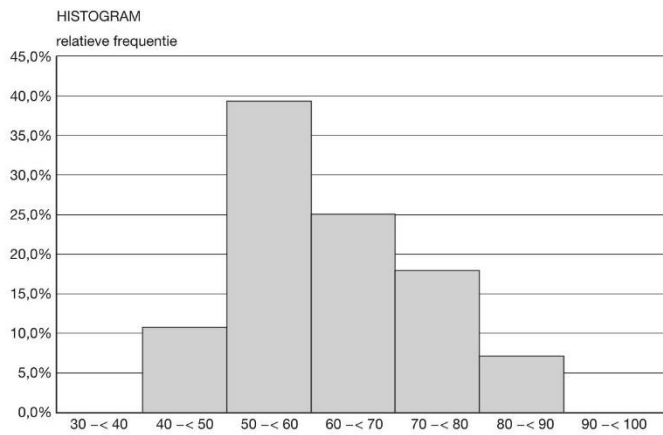
61

	A	B	C	D	E	F	G
1	ondergrens	bovengrens	klasse	klassenmidden	frequentie	cum. freq	rel. cum. freq.
2	30	40	30 -< 40	35	0	0	0,0%
3	40	50	40 -< 50	45	3	3	10,7%
4	50	60	50 -< 60	55	11	14	50,0%
5	60	70	60 -< 70	65	7	21	75,0%
6	70	80	70 -< 80	75	5	26	92,9%
7	80	90	80 -< 90	85	2	28	100,0%
8	90	100	90 -< 100	95	0	28	100,0%

62 a \*  
b \*

c

klasse	rel freq
30 -< 40	0,0%
40 -< 50	10,7%
50 -< 60	39,3%
60 -< 70	25,0%
70 -< 80	17,9%
80 -< 90	7,1%
90 -< 100	0,0%

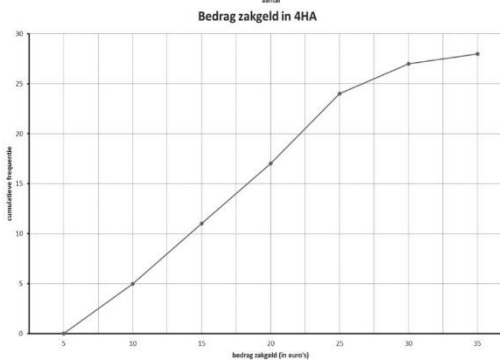
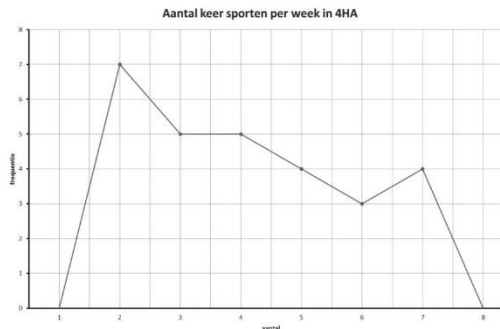


Bladzijde 89

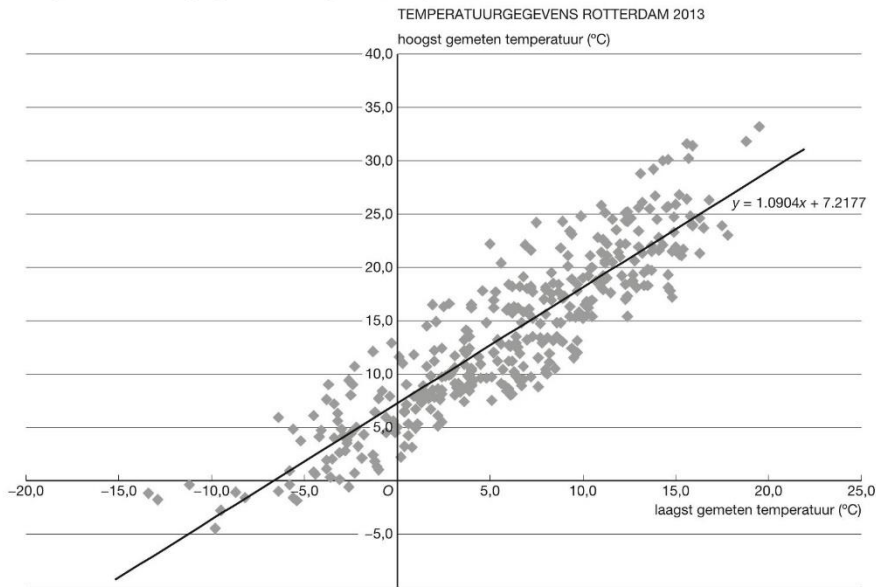
63

	A	B
1	aantal keer sporten per week	frequentie
2		0
3	0	7
4	1	5
5	2	5
6	3	4
7	4	3
8	5	4
9		0

	A	B	C	D
1	ondergrens	bovengrens	frequentie	cuml. freq.
2	0	5	0	0
3	5	10	5	5
4	10	15	6	11
5	15	20	6	17
6	20	25	7	24
7	25	30	3	27
8	30	35	1	28



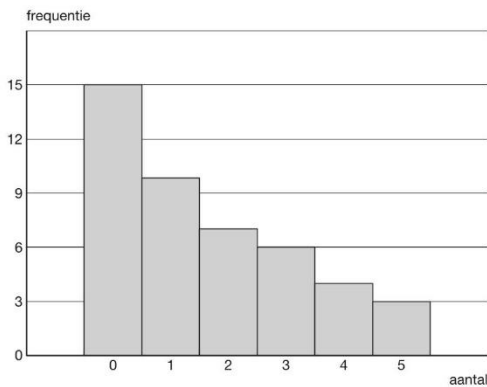
- 64 Excel geeft als formule voor de best passende lijn  $y = 1,0904x + 7,2177$  (zie onderstaande figuur).  
 Bij de laagst gemeten temperatuur  $18^\circ\text{C}$  hoort  $x = 18$ .  
 De bijbehorende hoogst gemeten temperatuur is  $y = 1,0904 \cdot 18 + 7,2177 \approx 27^\circ\text{C}$ .



### Diagnostische toets

#### Bladzijde 90

- 1 a PAKKEN MET TE LAAG GEWICHT

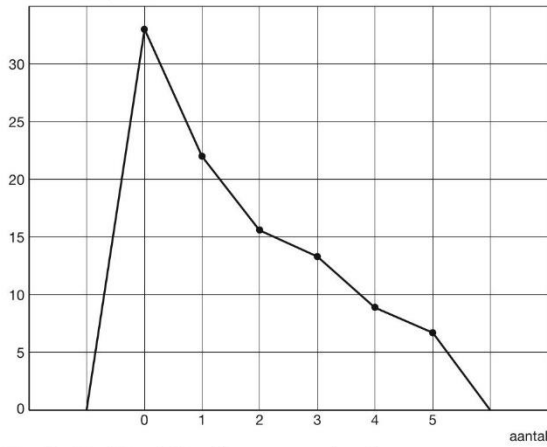


b

aantal	frequentie	rel. freq.
0	15	33,3%
1	10	22,2%
2	7	15,6%
3	6	13,3%
4	4	8,9%
5	3	6,7%
	45	



PAKKEN MET TE WEINIG GEWICHT  
relatieve frequentie

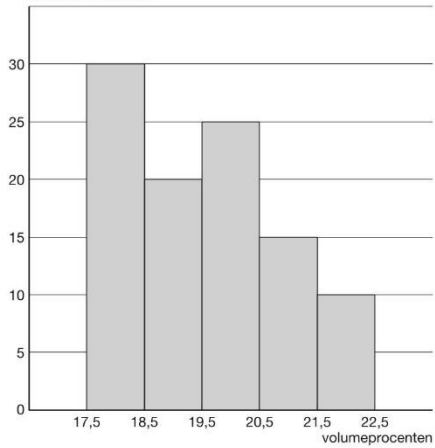


- c Er zijn  $45 \cdot 20 = 900$  pakken gecontroleerd.  
Hiervan hadden er  $1 \cdot 10 + 2 \cdot 7 + 3 \cdot 6 + 4 \cdot 4 + 5 \cdot 3 = 73$  te weinig gewicht.  
Dat is  $\frac{73}{900} \times 100\% \approx 8,1\%$ .

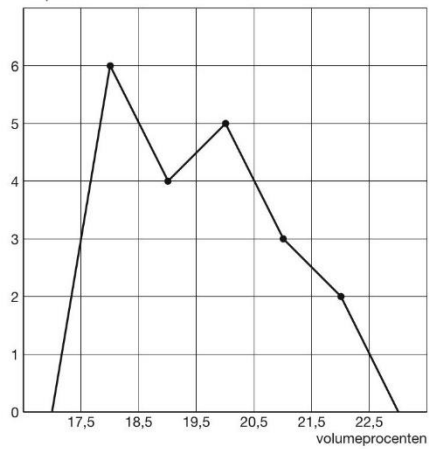
2 a

zuurstofgehalte		frequentie	rel. freq.
17,5 – < 18,5		6	30%
18,5 – < 19,5		4	20%
19,5 – < 20,5		5	25%
20,5 – < 21,5		3	15%
21,5 – < 22,5		2	10%
		20	

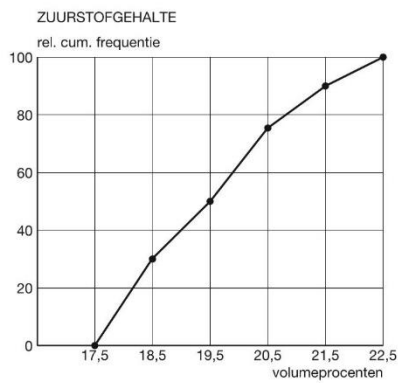
b ZUURSTOFGEHALTE  
relatieve frequentie



**c** ZUURSTOFGEHALTE  
frequentie



**d** De relatieve cumulatieve frequenties zijn 30%, 50%, 75%, 90%, 100%.



**e** zuurstofgehalte

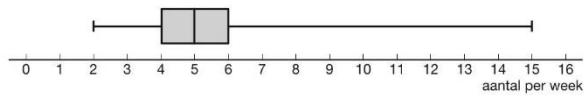
17	5	6					
18	2	2	2	4	5	6	7
19	3	6					
20	3	3	4	4	6	8	9
21	9						
22	3						

- 3** **a** Serie is een kwalitatieve variabele.  
Exemplaren verkocht is een kwantitatieve variabele.  
Gemiddelde waardering is een kwalitatieve variabele.
- b** Een kwalitatieve variabele kan niet worden gemiddeld.  
Als we al een gemiddelde zouden willen bepalen, dan zouden we moeten weten hoeveel stemmen er op elke game zijn uitgebracht.
- c** Bijvoorbeeld de variabele 'geschikt vanaf leeftijd'. Deze variabele is discreet.  
Of bijvoorbeeld 'gemiddelde speelduur'. Deze variabele is continu.

**Bladzijde 91**

- 4 a** Voer in lijst 1 = {2, 3, 4, 5, 6, 7, 8, 10, 15} en lijst 2 = {4, 6, 10, 18, 4, 4, 2, 3, 1}.  
1-Var Stats (TI) of IVAR (Casio) geeft  $\bar{x} \approx 5,2$  en  $\sigma \approx 2,3$ .  
Dus het gemiddelde is 5,2 keer uitrukken per week en standaardafwijking is 2,3 keer uitrukken per week.
- b** 1-Var Stats (TI) of IVAR (Casio) geeft  $\min X = 2$ ,  $Q_1 = 4$ ,  $\text{Med} = 5$ ,  $Q_3 = 6$  en  $\max X = 15$ .

UITRUKKEN BRANDWEER



- c** spreidingsbreedte =  $15 - 2 = 13$   
kwartielfstand =  $6 - 4 = 2$
- d** Er is een uitschieter aan de rechterkant. Dat kun je zien aan het lange lijnstuk van  $Q_3$  tot  $\max X$ . Deze uitschieter zorgt ervoor dat het gemiddelde hoger is dan de mediaan.
- 5 a** 1 In de ochtend hebben mensen minder trek in ice tea.  
2 Als het warm is hebben meer mensen trek in ice tea.  
3 Alleen leerlingen uit de onderbouw worden ondervraagd. Dit is niet een goede afspiegeling van de doelgroep.
- b** Omdat het een warme zomerdag is zullen jongeren die op het strand zijn de smaken misschien meer waarderen.
- 6 a** aantal jaren ouder dan vijftig: ratio  
opleidingsniveau: ordinaal  
hobby's: nominaal  
aantal keer sporten per week: ratio
- b** aantal jaren ouder dan vijftig: gemiddelde of mediaan  
opleidingsniveau: mediaan of modus  
hobby's: modus  
aantal keer sporten per week: gemiddelde of mediaan
- 7 a** Voer in lijst 1 = {8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 24} en  
lijst 2 = {2, 2, 2, 3, 8, 3, 3, 3, 1, 1, 1, 1}.  
1-Var Stats (TI) of IVAR (Casio) geeft  $\bar{x} = 12,8$  en mediaan = 12.
- b** De onderste boxplot, omdat hierbij de kwartielfstand het kleinst is.
- c** Wiebe kan het beste een spreidingsdiagram gebruiken met de lengte van de voornaam op de ene as en de lengte van de achternaam op de andere as.