

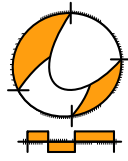
# Wiskunde / PGA

2 VMBO

## Statistiek

ConTeXt College





© 2024

Het auteursrecht op dit lesmateriaal berust bij Stichting Math4All. Math4All is derhalve de rechthebbende zoals bedoeld in de hieronder vermelde creative commons licentie.

Het lesmateriaal is met zorg samengesteld en getest. Stichting Math4All aanvaardt geen enkele aansprakelijkheid voor onjuistheden en/of onvolledigheden in de module. Ook aanvaardt Math4All geen enkele aansprakelijkheid voor enige schade, voortkomend uit (het gebruik van) dit lesmateriaal

Voor deze module geldt een Creative Commons Naamsvermelding Niet Commercieel 3.0 Nederland Licentie. (zie <http://creativecommons.org/licenses/by/3.0>).

Dit lesmateriaal is open, gratis en vrij toegankelijk lesmateriaal afkomstig van Stichting Math4All en is speciaal ontwikkeld voor het vak wiskunde in het voortgezet onderwijs. Het lesmateriaal op de website [www.math4all.nl](http://www.math4all.nl) is afgestemd op kerndoelen wiskunde, tussendoelen wiskunde en eindtermen voor de vakken wiskunde A, B en C. Dit lesmateriaal is mediumneutraal ontwikkeld en op diverse manieren te bekijken en te gebruiken. Voor informatie en vragen kunt u contact opnemen via [info@math4all.nl](mailto:info@math4all.nl). Ook houden we ons altijd aanbevolen voor suggesties, verbeteringen en/of aanvullingen.

---

Het lesmateriaal in dit katern is gebaseerd op het materiaal dat je kunt vinden op de Math4All website [www.math4all.nl](http://www.math4all.nl). In de tekst staan dan ook regelmatig verwijzingen naar die website. Waar je precies moet zijn op die website kun je zien in de kopregel van iedere pagina.

Ieder hoofdstuk bestaat uit een aantal paragrafen en wordt steeds afgesloten met een paragraaf *Totaalbeeld* waar de leerstof wordt samengevat en/of herhaald.

## **PGA**

PGA staat voor 'probleemgestuurde aanpak'. Je werkt dan onder begeleiding van je docent in kleine groepjes aan wiskundige problemen en samen bouw je de theorie op en maak je er een overzicht van.

De PGA wordt ondersteund door verwerkings- en toepassingsopgaven waarmee je kunt nagaan of je de stof beheerst. Deze opgaven worden op drie niveaus aangeboden. De niveau aanduiding vind je terug in de marge.

- ★ het basale niveau, dat iedereen zou moeten behalen
- ★ ★ een iets pittiger niveau, waarin iets meer uitdaging zit en die je alleen hoeft te maken als je er genoeg tijd voor hebt
- ★ ★ ★ een bijzondere toepassing of een echt pittige opgave die je alleen maakt als de rest veel te gemakkelijk voor je was



---

# 1

---

## Statistiek

<b>1.1</b>	<b>Centrummaten</b>	<b>6</b>
<b>1.2</b>	<b>Spreidingsmaten</b>	<b>15</b>
<b>1.3</b>	<b>Klassenindeling</b>	<b>25</b>
<b>1.4</b>	<b>Schatten</b>	<b>35</b>
<b>1.5</b>	<b>Statistische uitspraken</b>	<b>44</b>
<b>1.6</b>	<b>Totaalbeeld</b>	<b>51</b>

# 1.1 Centrummaten

## Inleiding

Als je beschikt over een hele serie gegevens, zoals alle rapportcijfers van alle leerlingen in een bepaald leerjaar van alle vakken, dan heb je een enorme brij aan getallen. Hoe krijg je daar enig overzicht over? Aicha gaat een poging doen. Eerst gebruikt ze de eindcijfers van enkele klassen van jaren geleden.

BIJ		ne		en		fa		ak		gs	
leerling	geslacht	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE
1	v	6,7	7	4,4	4	5,6	6	6,6	7	6,8	7
2	v	5,6	6	5,3	5	6,1	6	7,1	7	6,8	7
3	m	8,1	8	6,7	7	5,8	6	7,2	7	7,6	8
4	m	8,5	9	5,1	5	6,1	6	6,1	6	6,1	6
5	m	4,9	5	9,7	10	6,6	7	8,0	8	7,5	8
6	v	6,2	6	9,4	9	7,2	7	6,6	7	7,8	8

Figuur 1.1

Zoals ze weet kunnen frequentietabellen en diagrammen helpen. Maar soms is een enkel getal genoeg...

### Je leert in dit onderwerp

- van een hoeveelheid gegevens de modus, de mediaan en het (gewogen) gemiddelde berekenen;
- modus, mediaan en gemiddelde interpreteren als centrummaten van die gegevens.

### Voorkennis

- (gewogen) gemiddelde van een hoeveelheid gegevens berekenen;
- werken met (relatieve) frequenties en (relatieve) frequentietabellen.



## Theorie

### Om te onthouden





## Verwerken

### Opgave 1.1

Dit zijn de rapportcijfers die in een bepaalde klas voor economie zijn behaald:

7	4	6	6	5	5	7	6	7	9	10	6	8	7
8	6	5	7	5	8	3	7	8	6	6	10	5	7

Tabel 1.1

- Maak een frequentietabel van de rapportcijfers voor economie.
- Bepaal de modus.
- Bepaal de mediaan.

Twee leerlingen die tijdens de toets ziek waren maken een inhaaltoets. Ze halen allebei een zeven.

- Bereken het nieuwe gemiddelde, de nieuwe modus en de nieuwe mediaan.

### ★ Opgave 1.2

Een centrummaat moet een goede indruk geven van de waarnemingen die je bestudeert. Kies in de volgende gevallen een geschikte centrummaat, en bereken deze maat.

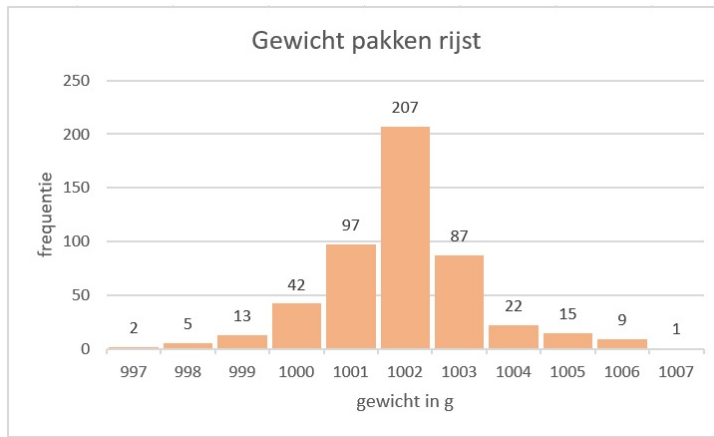
- Erik haalde voor zijn wiskundetoetsen de cijfers 8, 5, 6, 7, 10, 3, 7, 7 en 8.
- De directeur van een onderneming krijgt € 8000,00 per maand. De vier adjunctdirecteuren ontvangen maandelijks € 3000,00. De overige dertig werknemers verdienen elk € 1900,00 per maand.
- Van de docenten van een scholengemeenschap in Overdal komen er 16 te voet naar school, 20 komen met de fiets, 25 met de auto en 37 met het openbaar vervoer.
- De resultaten van een meting van de maximumtemperaturen in een week in juli in graden Celsius waren:

dag	ma	di	wo	do	vr	za	zo
temperatuur in °C	22	23	24	23	23	23	23

Tabel 1.2

### ★ Opgave 1.3

In een supermarkt worden pakken rijst verkocht met een gewicht van 1 kg. Elke werkdag wordt een vast aantal pakken gecontroleerd op afwijkend gewicht. In het staafdiagram zijn de resultaten verwerkt van 20 werkdagen.



**Figuur 1.2**

- a Bepaal de modus en de mediaan.
- b Hoeveel pakken werden er elke werkdag gecontroleerd?
- c Hoeveel procent van de gecontroleerde pakken rijst had een te laag gewicht?
- d Bereken in één decimaal nauwkeurig het gemiddelde aantal pakken rijst per dag waarvan het gewicht te laag is.
- e In totaal zijn in het afgelopen jaar 8500 pakken rijst verkocht. Geef een schatting van het aantal pakken dat minstens het juiste gewicht had.

★ **Opgave 1.4**

Iemand heeft van een aantal doosjes lucifers geteld hoeveel lucifers er in zitten. De resultaten staan in de tabel hiernaast.

- a Van hoeveel doosjes lucifers is het aantal lucifers geteld?
- b Hoeveel lucifers zijn er in totaal geteld?
- c Hoeveel lucifers zitten er gemiddeld in een doosje? Geef je antwoord in één decimaal nauwkeurig.
- d Waarom wist je voordat je de berekening bij c uitvoerde dat het gemiddelde lager moest zijn dan de modus 36 ?

aantal lucifers	frequentie
31	1
32	3
33	4
34	5
35	7
36	8
37	11
38	6

**Tabel 1.3**

★ **Opgave 1.5**

Een loterij heeft op een dag aan prijzen gemiddeld € 20000,00 uitbetaald. De mediaan van die prijzen blijkt € 1500,00 te zijn.

Geef een verklaring voor het grote verschil tussen die twee centrummaten.

★★ **Opgave 1.6**

Een school kent drie rapportperiodes die alle drie even zwaar tellen.

Voor wiskunde zijn er de eerste periode drie toetsen gegeven die elk 1 keer, 1 keer en 3 keer meetellen.

In de tweede periode zijn er in dat vak twee toetsen gegeven die 2 keer en 3 keer meetellen.

In de derde en laatste periode zijn er twee toetsen die 1 keer meetellen en een proefwerk

dat drie keer meetelt.

Je staat voor wiskunde een 7,2 en alleen het laatste proefwerk moet nog worden gemaakt. De cijfers voor alle toetsen worden in één decimaal nauwkeurig bepaald.

Kun je nog een 8 halen als eindcijfer? Laat met een berekening zien welk cijfer je dan voor het laatste proefwerk moet halen.

## Toepassen

Excel kan je helpen bij het maken van frequentietabellen en het berekenen van gemiddelden. Bekijk het **Practicum**. Daarin werk je met het Excelbestand **Gegevens154Leerlingen.xls** met gegevens van 154 vierde klassers.

Denk er wel om dat je op een pc of een tablet dit Excel-bestand eerst moet downloaden. Er wordt van uit gegaan dat je wel eens eerder met Excel hebt gewerkt en al weet hoe je met cellen werkt, met formules in cellen werkt, absoluut en relatief kopiëren begrijpt, etc.

### ★ Opgave 1.7: Cijfer voor wiskunde

Gebruik het gegevensbestand uit **Toepassen**.

- a Maak zelf een frequentietabel van de gehele eindcijfers voor wiskunde van deze groep leerlingen.
- b Zoek uit hoe je van een rij gegevens het gemiddelde, de modus en de mediaan door Excel kunt laten berekenen. Het gaat met behulp van formules. Zoek een geschikte plek in het Excel-bestand om deze centrummaten neer te zetten.
- c Bereken het gemiddelde voor wiskunde nog eens met behulp van de frequentietabel.
- d Maak vervolgens afzonderlijke groepen van de jongens en de meisjes door te sorteren. Bereken de centrummaten voor deze afzonderlijke groepen en vergelijk ze met elkaar. Maak ook frequentietabellen en staafdiagrammen van de afzonderlijke groepen.

### ★ Opgave 1.8: Geboortemaanden

Werk met hetzelfde Excel-bestand als in de voorgaande opgave. Zet de gegevens over de geboortemaanden op een afzonderlijk werkblad.

- a Maak een frequentietabel van de geboortemaanden van deze groep leerlingen.
- b Welke centrummaat is hier zinvol?
- c Probeer een conclusie te trekken uit de frequentietabel. Of maak er eerst een staafdiagram van en probeer dan een conclusie te trekken.

## Practicum: Werken met Excel

Excel kan zelf frequentietabellen voor je maken. Je hoeft dan niet met de hand te tellen.

In het Excelbestand **Gegevens154Leerlingen.xls** vind je enkele gegevens van 154 vierde klassers. Je kunt daarbij een frequentieverdeling voor bijvoorbeeld 'cijfwis', het cijfer voor wiskunde van deze leerlingen, maken. Dat gaat als volgt:

- Maak eerst een kopie van het werkblad met gegevens en verwijder alle kolommen behalve 'cijfwis'. Dat wordt dan kolom A.
- Maak eerst een kolom met cijfers van (bijvoorbeeld) 3, 4, 5, t/m 10 in de cellen C2 t/m C9. Geef die het opschrift 'cijfwis'.
- Dan zijn de cellen D2 t/m D9 bestemd voor de frequenties. Selecteer die allemaal.
- Klik nu in de formulebalk en zet daar in: =INTERVAL(A2:A155;C2:C9) en doe [ctrl][shift][enter].

Als het goed is krijg je nu de juiste frequenties. Excel telt dus in de cellen A2 t/m A155 (de cijfers voor wiskunde) hoeveel er steeds voorkomen van de waarden in de cellen C2 t/m C9, de verschillende cijfers die voorkomen.

Maar misschien wil je de jongens en de meisjes vergelijken? Dan laat je Excel eerst sorteren op de kolom 'geslacht'. Dat doe je door die kolom te selecteren en 'Sorteren en filteren' te kiezen. Sorteert van laag naar hoog en klik op OK als Excel je vraagt of de selectie moet worden uitgebreid (want de andere kolommen moeten ook worden meegenomen bij het sorteren). Nu heb je de jongens en de meisjes als groepen bij elkaar staan en kun je hun centrummaten vergelijken met of zonder frequentietabellen.

**LET OP:**

Je moet het Excel-bestand Gegevens154Leerlingen.xls eerst downloaden om er in te kunnen werken.

Je kopieert het dan naar de computer waarop je werkt.

# Antwoorden

**1.1 a** Zie de tabel.

cijfer	2	3	4	5	6	7	8	9	10
aantal	0	1	1	5	7	7	4	1	2

- b** Er is geen modus, de 6 en de 7 komen even vaak voor.
- c** Zet alle cijfers op volgorde van klein naar groot en kies de middelste. Er zijn 28 cijfers behaald dus het 14e en 15e cijfer staan in het midden. Dit zijn een 6 en een 7, het gemiddelde daarvan is 6,5. Dus de mediaan is 6,5.
- d** Er waren eerst 28 cijfers, maar met deze twee erbij zijn er nu dus 30 cijfers.

$$\text{Nieuwe gemiddelde: } \frac{1 \cdot 3 + 1 \cdot 4 + 5 \cdot 5 + 7 \cdot 6 + 9 \cdot 7 + 4 \cdot 8 + 1 \cdot 9 + 2 \cdot 10}{30} = \frac{198}{30} = 6,6.$$

De modus is nu 7 want dat cijfer komt nu het vaakst voor.

Nu er 30 cijfers zijn, is de mediaan het gemiddelde van het 15e en 16e cijfer. Dus het gemiddelde van 7 en 7 dus de mediaan is 7.

**1.2 a** Het gemiddelde ligt hier voor de hand.  $\frac{61}{9} \approx 6,8$ .

- b** Hier is de modus geschikter voor een goede indruk. € 1900,00 komt het meest voor.
- c** Een gemiddelde van auto's en fietsen berekenen kan niet, en van klein naar groot rangschikken is ook een probleem. De modus is wel te bepalen: het openbaar vervoer.
- d** Hier ligt het gemiddelde voor de hand, en dat is  $\frac{22+5 \cdot 23+24}{7} = 23$  °C.

**1.3 a** De modus is in dit geval het gewicht dat het vaakst voor komt. Dat is 1002 gram.

Er zijn 500 pakken gewogen dus de mediaan is het gemiddelde van het 250ste en 251ste pak. Deze wegen beide 1002 gram. Dus de mediaan is 1002 gram.

- b** In totaal werden er 500 pakken gecontroleerd in 20 dagen.  $\frac{500}{20} = 25$  pakken per werkdag.
- c** Als een pak minder dan 1000 gram weegt, is dat te weinig. Dat is het geval bij 20 van de 500 pakken. Dus  $\frac{20}{500} \cdot 100 = 4\%$  van de pakken.
- d** 20 pakken in 20 dagen hadden een te laag gewicht. Dat is dus 1 pak per dag.
- e** Als 4% van de pakken te licht was, dan had de andere 96% van de pakken het juiste gewicht (ervan uitgaande dat een pak zwaarder van 1000 gram mag zijn). En 96% van 8500 is  $0,96 \cdot 8500 = 8160$  pakken.

**1.4 a** Van 45 doosjes lucifers.

**b**  $1 \cdot 31 + 3 \cdot 32 + 4 \cdot 33 + 5 \cdot 34 + 7 \cdot 35 + 8 \cdot 36 + 11 \cdot 37 + 6 \cdot 38 = 1599$  lucifers.

**c**  $\frac{1599}{45} \approx 35,5$  lucifers.

**d** Omdat er het aantal doosjes met minder dan 36 lucifers groter is dan het aantal doosjes met meer dan 36 lucifers.

**1.5** Dat grote verschil kan ontstaan als er één of meerdere enorm hoge prijzen zijn uitgekeerd. Bijvoorbeeld: Twee prijzen van € 1000,00, twee van € 1500,00 en één van € 95000,00.

**1.6** Om een 8 te halen moet het gemiddelde minimaal een 7,5 zijn. Een cijfer dat één keer telt kun je zien als één cijfer, een cijfer dat twee keer telt als twee cijfers, etc.

In totaal zijn er dan dus  $1 + 1 + 3 + 2 + 3 + 1 + 1 + 3 = 15$  cijfers te behalen. Als je daarvoor een 7,5 gemiddeld wilt halen moet je in totaal  $15 \cdot 7,5 = 112,5$  punten halen.

Voor de eerste 12 cijfers heb je in totaal  $12 \cdot 7,2 = 86,4$  punten gehaald.

Voor het laatste cijfer dat drie keer telt moet je dus  $112,5 - 86,4 = 26,1$  punten halen.

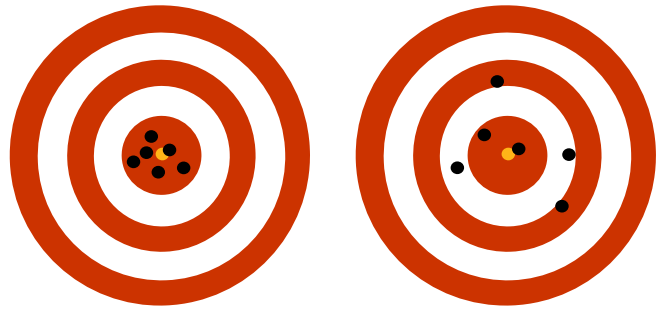
Dus voor het laatste proefwerk moet je een  $\frac{26,1}{3} = 8,7$  of hoger halen.

- 1.7 a** Vergelijk je resultaten met die van de andere leerlingen.
- b** Vergelijk je resultaten met die van de andere leerlingen. Het gemiddeld is ongeveer 7,5 en de mediaan en de modus zijn beide 8.
- c** Vergelijk je resultaten met die van de andere leerlingen.
- d** Vergelijk je resultaten met die van de andere leerlingen.
- 1.8 a** Vergelijk je resultaten met die van de andere leerlingen.
- b** Alleen de modus. In dit geval is januari de modus.
- c** Vergelijk je resultaten met die van de andere leerlingen. Opvallend is dat van deze leerlingen er maar weinig in februari en augustus zijn geboren. Heb je daar een verklaring voor?

## 1.2 Spreidingsmaten

### Inleiding

Aicha ziet dat twee groepen gegevens met eenzelfde centrummaat nog flink van elkaar kunnen verschillen. De spreiding van de gegevens kan erg verschillend zijn. Je ziet dat hier met de schoten op deze roos. Van de tweede serie is de spreiding groter.



Figuur 2.1

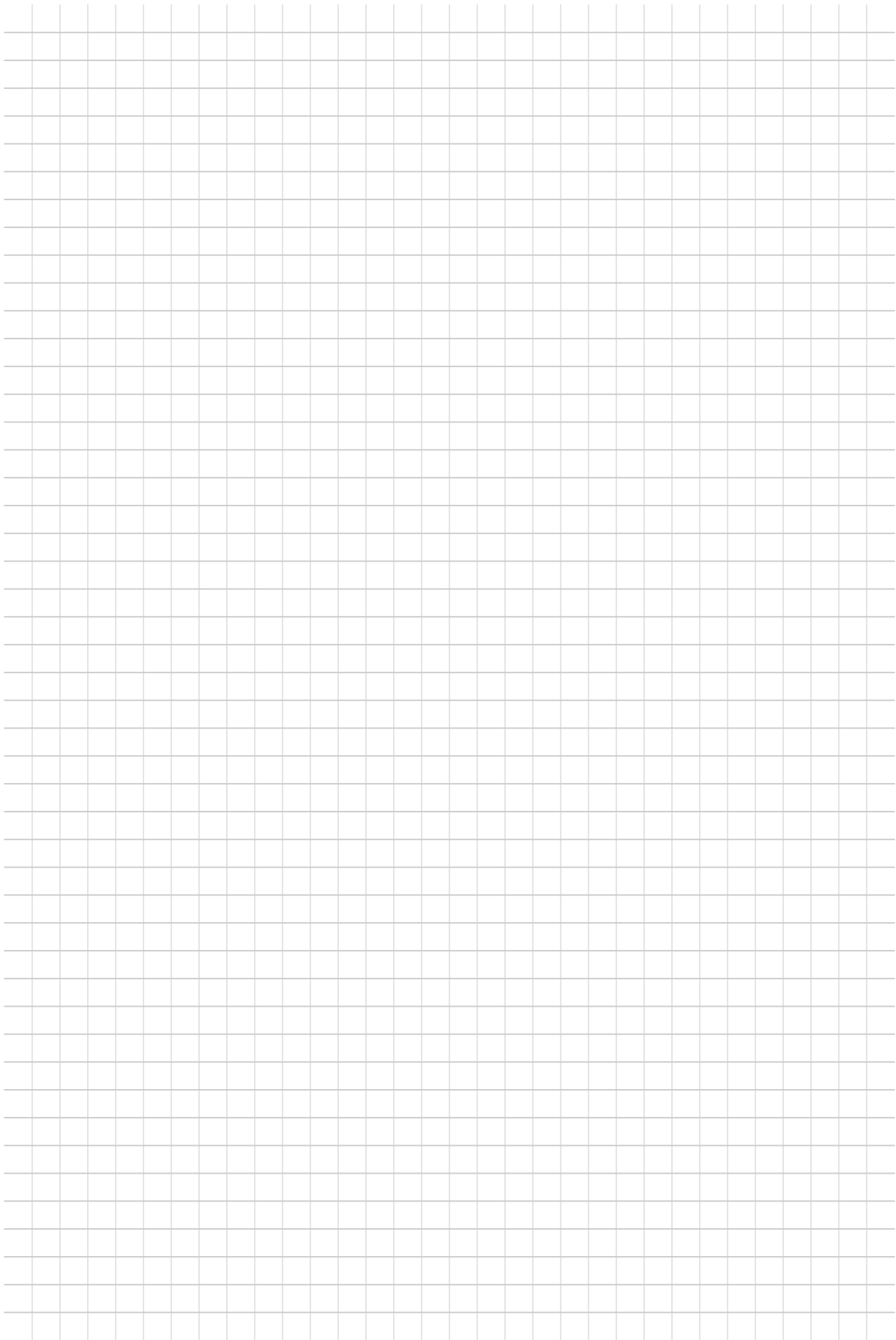
### Je leert in dit onderwerp

- de spreidingsbreedte en de kwartielafstand van een serie gegevens berekenen;
- spreidingsbreedte en kwartielafstand interpreteren als spreidingsmaten van die gegevens;
- een serie gegevens samenvatten in een boxplot.

### Voorkennis

- (gewogen) gemiddelde, modus en mediaan van een hoeveelheid gegevens berekenen;
- werken met (relatieve) frequenties en (relatieve) frequentietabellen.

## Aantekeningen

A large grid of graph paper with 20 columns and 30 rows of small squares, intended for taking notes or drawing.





## Theorie

### Om te onthouden

A large grid of graph paper with 20 columns and 40 rows, intended for taking notes on the theory.

## Verwerken

### ★ Opgave 2.1

Dit is een steelbladdiagram van de cijfers voor een scheikundetoets in een vijfde klas.

3	4																			
4	8	8	9	9																
5	0	0	0	0	1	1	1	1	2	2	3									
6																				
7	9																			
8	0	0	1	1	1	2	2	2	3	3	4	4	6							
9	8																			

Figuur 2.2

- Maak er een boxplot bij.
- Welk diagram geeft het beste beeld van de verdeling van de cijfers, het steelbladdiagram of de boxplot?

### ★ Opgave 2.2

Van twee honkbalteams zijn alle leeftijden genoteerd:

team 1	18	20	21	22	23	23	24	28	28
team 2	21	22	22	22	23	23	24	24	26

Tabel 2.1

- Bepaal van beide teams de gemiddelde leeftijd en de mediaan.
- Teken bijbehorende boxplots.
- Krijg je hierdoor een goede indruk van het verschil in leeftijdsopbouw tussen beide teams? Leg dat uit.

### ★ Opgave 2.3

Op 1 januari 2013 werden de volgende geboortegewichten van een groep jongens gemeten. De resultaten in grammen:

1850	2100	2500	2600	2700	2800	2900	3150	3250	3300
3300	3400	3400	3500	4000	4000	4000	4050	4090	4100
4150	4160	4180	4200	4200	4400	4500	4700	4800	4850

Tabel 2.2

Hetzelfde gebeurde bij een even grote groep meisjes:

2400	2500	2500	2600	2700	2800	2800	2800	2900	2900
3000	3000	3100	3100	3200	3200	3200	3300	3300	3350
3400	3400	3500	3500	3600	3700	3800	3900	4000	4000

Tabel 2.3

- Bepaal de spreidingsbreedte bij de jongens en bij de meisjes.
- Maak de boxplots die bij deze gegevens horen.
- Hoeveel procent van de meisjes woog tussen de 2800 en 3500 gram?
- Hoeveel procent van de jongens woog meer dan het zwaarste meisje?

★ **Opgave 2.4**

Een van de leerlingen van de vijfde klas heeft de gewichten van al zijn klasgenoten opgeschreven. Het resultaat in kilogrammen is als volgt:

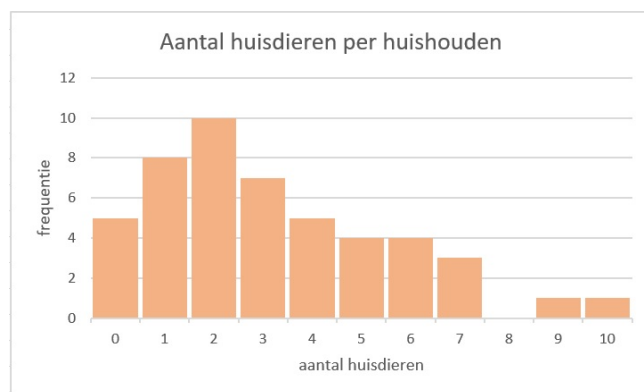
72	50	63	59	63	51	56	64	58	74
52	69	42	51	54	58	51	60	60	59
50	67	61	45	66	49	64	63	57	56

**Tabel 2.4**

- a Bereken bij deze gegevens de mediaan en de interkwartielafstand.
- b Hoeveel procent van deze leerlingen weegt minder dan 63 kg volgens je antwoord bij a? Is dat in werkelijkheid ook zo?

★ **Opgave 2.5**

Aan een groep kinderen uit klas 2 is gevraagd hoeveel huisdieren er in hun huishouden zijn. De resultaten zie je in het staafdiagram.



**Figuur 2.3**

- a Hoeveel kinderen zijn er ondervraagd?
- b Bereken het gemiddelde aantal huisdieren per huishouden.
- c Bereken de mediaan en de kwartielen.
- d Bepaal de spreidingsbreedte en de interkwartielafstand.
- e Teken de bijbehorende boxplot.
- f Hoeveel huisdieren hebben de 25% huishoudens met de meeste huisdieren?

**Toepassen**

Excel kan je helpen bij het berekenen van centrum- en spreidingsmaten en het maken van boxplots. Bekijk het [Practicum](#) en het bestand [Gegevens154Leerlingen.xls](#).

Denk er wel om dat je op een pc dit Excel-bestand eerst moet downloaden. Er wordt van uit gegaan dat je wel eens eerder met Excel hebt gewerkt en al weet hoe je met cellen werkt, met formules in cellen werkt, absoluut en relatief kopiëren begrijpt, etc.

### ★ Opgave 2.6: Cijfer voor wiskunde

Gebruik het gegevensbestand uit [Toepassen](#).

- Maak zelf een tabel met de kwartielen, de mediaan en maximum en minimum van de gehele eindcijfers voor wiskunde van deze groep leerlingen.
- Maak een bijbehorend boxplot.
- Maak vervolgens afzonderlijke groepen van de jongens en de meisjes door te sorteren. Bereken de centrum- en spreidingsmaten voor deze afzonderlijke groepen en vergelijk ze met elkaar. Maak ook boxplots van de afzonderlijke groepen.

### ★ Opgave 2.7: Lengtes

Werk met hetzelfde Excel-bestand als in de voorgaande opgave. Zet de gegevens over de lengtes van de jongens en de meisjes op een afzonderlijk werkblad. Doe hetzelfde met de gewichten van de jongens en de meisjes.

- Maak voor de lengtes van de jongens en de meisjes afzonderlijke boxplots. Kun je conclusies trekken?
- Maak voor de gewichten van de jongens en de meisjes afzonderlijke boxplots. Kun je conclusies trekken?

## Practicum: Werken met Excel

Excel kan centrummaten en spreidingsmaten voor je berekenen. Je hebt daarvoor geen frequentietabellen nodig, die zijn alleen erg handig bij het maken van diagrammen.

In het Excelbestand [Gegevens154Leerlingen.xls](#) vind je enkele gegevens van 154 vierde klassers.

Je kunt daarbij van bijvoorbeeld 'cijfwis', het cijfer voor wiskunde van deze leerlingen, het minimum, het eerste kwartiel, de mediaan, het derde kwartiel en het maximum laten berekenen.

Wil je die ook nog in een boxplot weergeven, dan ga je (vanaf Excel2010) zo te werk:

- Maak eerst een kolom met onder elkaar Q1, minimum, mediaan, maximum en Q3 in die volgorde. Geef de kolom ernaast het opschrift 'cijfwis'.
- In de cellen van die kolom maak je de formules =KWARTIEL(H2:H155;1), =MIN(H2:H155), =MEDIAAN(H2:H155), =MAX(H2:H155) en =KWARTIEL(H2:H155;3) als 'cijfwis' in de kolom H staat.
- Nu selecteer je beide kolommen en het opschrift en via Invoegen kies je een lijndiagram met markeringen.
- Je krijgt nu een lijndiagram. Klik met de rechtermuisknop op de figuur en kies 'Gegevens selecteren'. Klik daar op 'Schakelen tussen rij en kolom'. Je krijgt dan de vijf waarden boven elkaar.
- Bij 'Hulpmiddelen voor grafieken' kun je bij 'Indeling' kiezen voor 'Omhoog/omlaag balken weergeven' en 'Hoog/laag lijnen'. Je krijgt dan je boxplot, dat je nog even zo mooi moet maken als je zelf wilt. (Rechtermuisknop op elke markering en opmaak kiezen.)

Misschien wil je de jongens en de meisjes vergelijken? Dan laat je Excel eerst sorteren op de kolom 'geslacht'. Dat doe je door die kolom te selecteren en 'Sorteren en filteren' te kiezen. Sorteert van laag naar hoog en klik op OK als Excel je vraagt of de selectie moet worden uitgebreid (want de andere kolommen moeten ook worden meegenomen bij het sorteren). Nu heb je de jongens en de meisjes als groepen bij elkaar staan en kun je hun centrummaten en spreidingsmaten vergelijken en twee boxplots in één figuur maken.



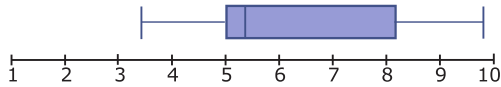
**LET OP:**

Je moet het Excel-bestand Gegevens154Leerlingen.xls eerst downloaden om er in te kunnen werken.

Je kopieert het dan naar de computer waarop je werkt.

## Antwoorden

- 2.1 a** Laagste cijfer is 3,4 gram,  $Q_1 = 5,0$  gram, de mediaan is 5,3 gram,  $Q_3 = 8,2$  gram en het maximumgewicht is 9,8 gram.



- b** De duidelijke scheiding tussen voldoende en onvoldoende zie je het best in het steelblad-diagram. De boxplot laat de spreiding beter zien.
- 2.2 a** Van beide teams zijn zowel de gemiddelde leeftijd als de mediaan 23.
- b** Team 1: minimumleeftijd is 18 jaar,  $Q_1 = 20,5$  jaar, de mediaan is 23 jaar,  $Q_3 = 26$  jaar en het maximumleeftijd is 28 jaar.  
Team 2: minimumleeftijd is 21 jaar,  $Q_1 = 22$  jaar, de mediaan is 23 jaar,  $Q_3 = 24$  jaar en het maximumleeftijd is 26 jaar.  
Tekenen de bijbehorende boxplots.
- c** Ja, je kunt goed het verschil tussen de minimumleeftijden en de maximumleeftijden zien. Verder is het verschil tussen beide kwartielafstanden een duidelijke aanwijzing dat bij team 2 de leeftijden veel dicht bij elkaar liggen.
- 2.3 a** Spreidingsbreedte jongens:  $4850 - 1850 = 3000$ .  
Spreidingsbreedte meisjes:  $4000 - 2400 = 1600$ .
- b** Jongens: minimumgewicht is 1850 gram,  $Q_1 = 3150$  gram, de mediaan is 4000 gram,  $Q_3 = 4180$  gram en het maximumgewicht is 4850 gram.  
Meisjes: minimumgewicht is 2400 gram,  $Q_1 = 2800$  gram, de mediaan is 3200 gram,  $Q_3 = 3500$  gram en het maximumgewicht is 4000 gram.  
Tekenen nu de boxplots met deze gegevens.
- c** Tussen  $Q_1$  en  $Q_3$  zit altijd 50% van de waarnemingen. Dus 50% van de meisjes woog tussen de 2800 en 3500 gram.
- d** Het zwaarste meisje woog 4000 gram. Dat is precies de mediaan bij de jongens. Tussen de mediaan en de grootste waarneming zit 50% van de waarnemingen. Dus 50% van de jongens weegt meer dan het zwaarste meisje.
- 2.4 a** Het is handig om eerst een frequentietabel te maken.

gewicht	frequentie
42	1
45	1
49	1
50	2
51	3
52	1
54	1
56	2
57	1
58	2
59	2
60	2
61	1
63	3
64	2
66	1
67	1
69	1
72	1
74	1
totaal	30

Er zijn 30 getallen, dus de mediaan is het gemiddelde van het 15e en 16e getal. Dus van 58 en 59, de mediaan is dus 85,5. Het aantal is even, dus mediaan doet mee bij splitsing.

Voor  $Q_1$  kijk je naar het midden van de eerste 15 getallen, dus naar het 8e getal. Dit is 51.

Voor  $Q_3$  kijk je naar het midden van de laatste 15 getallen, dus naar het 8e getal van het tweede deel. Dit is 63.

De interkwartielafstand is  $63 - 51 = 12$ .

- b** Dat zou 75% moeten zijn. In werkelijkheid zijn 20 van de 30 leerlingen lichter, dat is minder dan 75%. Aan de andere kant zijn maar 7 leerlingen zwaarder en dat is nog geen 25%. Dit heeft allemaal met afrondingen te maken en het kleine aantal gegevens.

**2.5 a** 48

- b** Ongeveer 3,2 dus 3.

- c** Mediaan: 3.

Eerste kwartiel: 1.

Derde kwartiel: 5.

- d** Spreidingsbreedte:  $10 - 0 = 10$ .

Interkwartielafstand:  $5 - 1 = 4$ .

- e** De vijf getallen zijn: laagste 0,  $Q_1 = 1$ , mediaan 3,  $Q_3 = 5$  en hoogste 10.

- f** 5 t/m 10 huisdieren.

**2.6 a** Vergelijk je resultaten met die van de andere leerlingen.

- b** Vergelijk je resultaten met die van de andere leerlingen.

- c** Vergelijk je resultaten met die van de andere leerlingen.
- 2.7 a** Vergelijk je resultaten met die van de andere leerlingen.
- b** Vergelijk je resultaten met die van de andere leerlingen.



## 1.3 Klassenindeling

### Inleiding

Aicha bekijkt de cijfers voor geschiedenis van klas B1H op één decimaal afgerond.

Dan zijn bijna alle gegevens verschillend.

Het maken van een frequentietabel levert nu weinig overzicht op.

Daarvoor kun je beter de cijfers groeperen in zogenaamde 'klassen'.

Daarover gaat dit onderdeel.

B1H		ne		en		fa		ak		gs	
leerling	geslacht	RE	RE	RE	RE	RE	RE	RE	RE	RE	RE
1	v	6,7	7	4,4	4	5,6	6	6,6	7	6,8	7
2	v	5,6	6	5,3	5	6,1	6	7,1	7	6,8	7
3	m	8,1	8	6,7	7	5,8	6	7,2	7	7,6	8
4	m	8,5	9	5,1	5	6,1	6	6,1	6	6,1	6
5	m	4,9	5	9,7	10	6,6	7	8,0	8	7,5	8
6	v	6,2	6	9,4	9	7,2	7	6,6	7	7,8	8

Figuur 3.1

### Je leert in dit onderwerp

- een frequentietabel met klassenindeling en geschikte klassenbreedte maken bij een verzameling ruwe data;
- bij zo'n klassenindeling diagrammen maken;
- bij een klassenindeling het gemiddelde schatten.

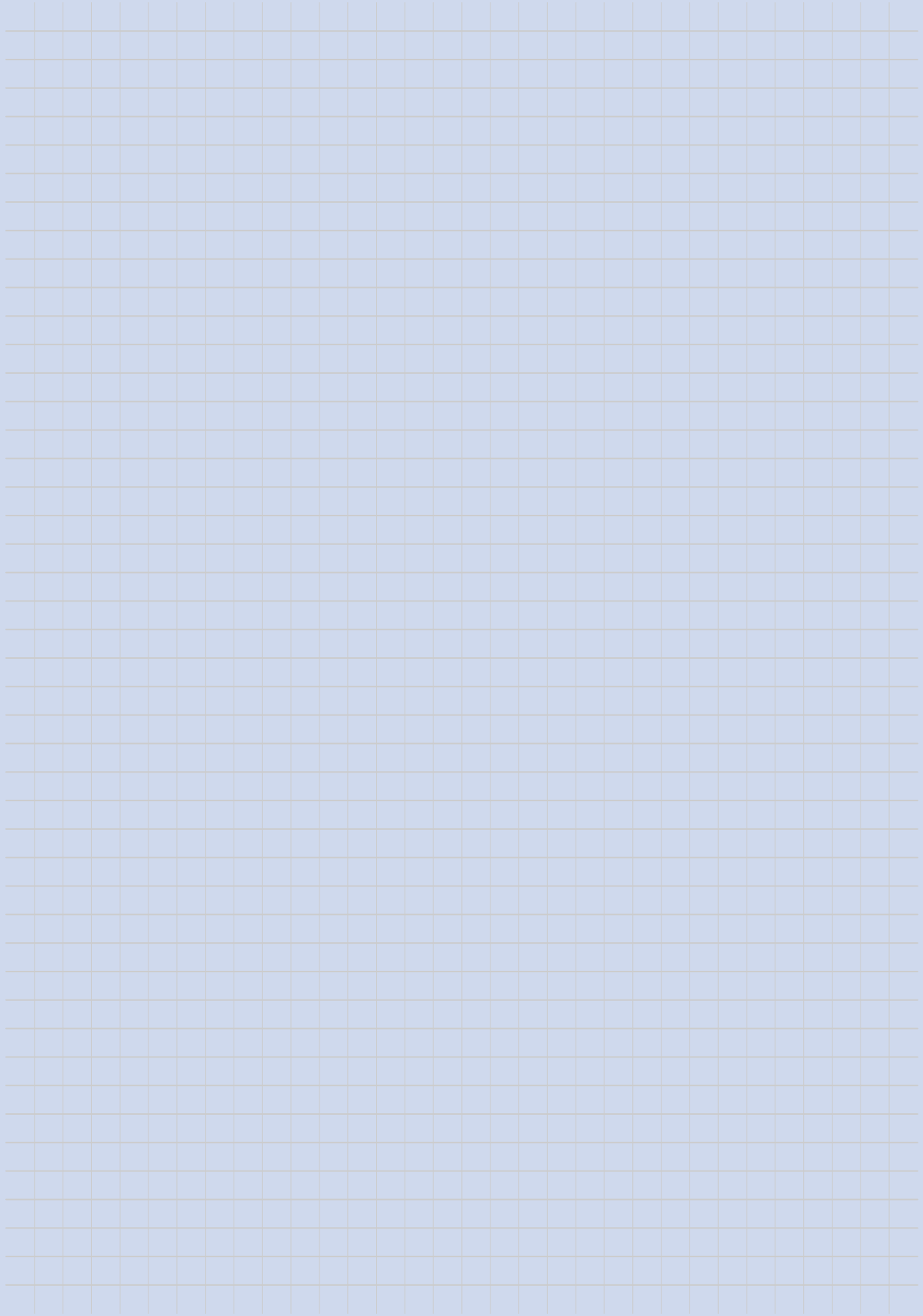
### Voorkennis

- (gewogen) gemiddelde, modus en mediaan van een hoeveelheid gegevens berekenen;
- werken met (relatieve) frequenties en (relatieve) frequentietabellen.



## Theorie

### Om te onthouden



A large grid of graph paper with light blue lines on a white background, intended for taking notes or drawing diagrams.

## Verwerken

### ★ Opgave 3.1

Geef in de volgende situaties aan of een klassenindeling maken mogelijk/zinvol is, leg uit waarom. Zo ja, geef aan wat jij als eerste klasse zou kiezen.

- a Je hebt op de huishoudbeurs aan 500 bezoekers hun leeftijd gevraagd. De jongste bezoeker was 17 en de oudste was 81.
- b Je hebt aan die zelfde 500 bezoekers van de huishoudbeurs hun lievelingskleur gevraagd. Tussen de antwoorden die de bezoekers gaven zaten 12 verschillende kleuren.
- c Je hebt aan die zelfde 500 bezoekers van de huishoudbeurs gevraagd hoe vaak ze nu in totaal de huishoudbeurs al bezocht hebben. Het laagste aantal was 1 keer en het hoogste aantal was 15 keer.

### ★ Opgave 3.2

Je ziet hier een klassenindeling van de gewichten van een groep personen.

- a Welk klassenmidden heeft de klasse  $45- < 50$ ? Licht je antwoord toe.
- b Bereken het gemiddelde gewicht in één decimaal nauwkeurig.
- c Hoeveel procent van deze groep mensen weegt minder dan 70 kg?

gewichtsklasse	frequentie
45– < 50	1
50– < 55	5
55– < 60	8
60– < 65	15
65– < 70	19
70– < 75	9
75– < 80	3

Tabel 3.1

### ★ Opgave 3.3

Johan gooit zestig keer met vijf dobbelstenen en noteert telkens de som van de vijf aantallen ogen. Zijn scores zijn als volgt:

16	22	21	16	20	24	20	23	16	22
17	15	17	26	23	20	17	19	21	23
17	22	18	14	18	15	16	18	17	18
20	22	18	18	25	14	13	12	15	17
22	14	18	20	12	20	15	18	15	24
17	12	23	14	11	14	21	16	18	5

Tabel 3.2

- a Wat was het grootste getal dat hij had kunnen gooien? Leg uit.
- b Bereken de gemiddelde score in één decimaal nauwkeurig.
- c Maak een klassenindeling van de scores. Neem als klassen  $5- < 8$  en  $8- < 11$ , enzovoorts.
- d Welke scores horen tot de klasse  $5- < 8$ ? Wat is dus het klassenmidden? En waarom kun je nu niet gewoon het gemiddelde van de klassengrenzen nemen?
- e Bereken nu met behulp van de klassenindeling de gemiddelde score.

- f In hoeveel procent van de worpen scoorde Johan meer dan 13 maar minder dan 23?
- g Waarom zou score 18 zoveel vaker voorkomen dan score 28?

★ **Opgave 3.4**

Bij een cabaretvoorstelling is aan 44 bezoekers hun leeftijd gevraagd. Het resultaat is als volgt:

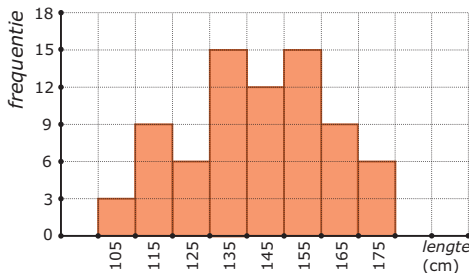
31	29	18	16	24	47	12	32	52	10	26
22	12	53	49	25	21	59	51	32	16	27
18	29	31	37	17	42	31	39	19	33	28
44	16	29	22	17	19	54	27	30	28	51

Tabel 3.3

- a Maak een klassenindeling met als klassen  $10- < 15$ , enzovoorts. Voeg de absolute frequenties toe..
- b Welke leeftijden horen tot de klasse  $50- < 55$ ? Wat is dus het klassenmidden?
- c Schat met behulp van de klassenindeling de gemiddelde leeftijd van deze bezoekers.
- d Hoeveel procent van de ondervraagden was jonger dan 30?

★ **Opgave 3.5**

Je ziet hier een verdeling van de lengtes van de vrouwelijke leerlingen van een balletschool. Op de horizontale as staan de klassenmiddens in cm.



Figuur 3.2

- a Schrijf de eerste klasse van de indeling op.
- b Van hoeveel leerlingen is de lengte gemeten?
- c Schat met behulp van de klassenindeling de gemiddelde lengte van deze leerlingen in cm nauwkeurig.
- d Maak een nieuwe klassenindeling met klassen  $100- < 120$ , enzovoorts. Bereken opnieuw de gemiddelde lengte van deze leerlingen, maar nu vanuit de nieuwe klassenindeling. Verklaar het verschil.

★ **Opgave 3.6**

Op 1 januari 2003 werden de volgende geboortegewichten van een groep jongens gemeten. De resultaten in grammen:

1850	2100	2500	2600	2700	2800	2900	3150	3250	3300
3300	3400	3400	3500	4000	4000	4000	4050	4090	4100
4150	4160	4180	4200	4200	4400	4500	4700	4800	4850

**Tabel 3.4**

Hetzelfde gebeurde bij een even grote groep meisjes:

2400	2500	2500	2600	2700	2800	2800	2800	2900	2900
3000	3000	3100	3100	3200	3200	3200	3300	3300	3350
3400	3400	3500	3500	3600	3700	3800	3900	4000	4000

**Tabel 3.5**

Omdat je de geboortegewichten van deze twee groepen wilt vergelijken maak je een klassenindeling en de bijbehorende frequentietabellen. Neem klassen met een breedte van 500 gram. De eerste klasse begint bij 1500.

- Maak een geschikte klassenindeling met bijbehorende absolute frequenties.
- Schat de gemiddelden van beide groepen met behulp van de klassenindeling bij a.
- Hoeveel procent van de jongens woog meer dan het bij b gevonden gemiddelde van de meisjes?
- Hoeveel procent van de jongens woog meer dan het zwaarste meisje?

## Toepassen

Excel kan centrummaten en spreidingsmaten voor je berekenen vanuit de ruwe data, je hebt daarvoor geen frequentietabellen nodig. Maar frequentietabellen zijn handig bij het maken van diagrammen. En vooral als je over veel gegevens beschikt is groeperen in klassen erg handig. Bekijk het [Practicum](#) en het bestand [Gegevens154Leerlingen.xls](#).

Denk er wel om dat je op een pc dit Excel-bestand eerst moet downloaden. Er wordt van uit gegaan dat je wel eens eerder met Excel hebt gewerkt en al weet hoe je met cellen werkt, met formules in cellen werkt, absoluut en relatief kopiëren begrijpt, etc.

### ★ Opgave 3.7: Lengtes in klassen

Gebruik het gegevensbestand uit [Toepassen](#).

- Maak zelf de frequentieverdeling van de lengtes van deze groep vierde klassers.
- Maak een staafdiagram.
- Maak vervolgens afzonderlijke groepen van de jongens en de meisjes door te sorteren. Maak frequentietabellen voor deze afzonderlijke groepen en vergelijk ze met elkaar.

### ★ Opgave 3.8: Gemiddelden schatten

Werk met hetzelfde Excel-bestand als in de voorgaande opgave. Je hebt de gegevens over de lengtes van de jongens en de meisjes op een afzonderlijk werkblad gezet. Doe hetzelfde met de gewichten van de jongens en de meisjes.

- Maak voor de lengtes van de jongens en de meisjes afzonderlijke frequentietabellen en schat met behulp daarvan hun gemiddelde lengtes. Kun je conclusies trekken?
- Maak voor de gewichten van de jongens en de meisjes afzonderlijke frequentietabellen en schat met behulp daarvan hun gemiddelde lengtes. Kun je conclusies trekken?

## Practicum: Werken met Excel

Excel kan centrummaten en spreidingsmaten voor je berekenen vanuit de ruwe data, je hebt daarvoor geen frequentietabellen nodig. Maar frequentietabellen zijn handig bij het maken van diagrammen. En vooral als je over veel gegevens beschikt is groeperen in klassen erg handig.

In het Excelbestand [Gegevens154Leerlingen.xls](#) vind je enkele gegevens van 154 vierde klassers.

Je vindt daarin bijvoorbeeld 'lengte', de lengtes van deze leerlingen.

Wil je daar een frequentietabel bij maken, dan bedenk je eerst een geschikte klassenindeling. Maak een kopie van het werkblad en verwijder daarin de kolommen D, E, F, H, I, J, K, L, M en ook de uitleg naast de tabel kan wel weg. Je gegevens komen nu in kolom D.

Nu ga je (in Excel2010) zo te werk:

- Laat Excel eerst het minimum en het maximum van de lengtes in de cellen D2 t/m D155 berekenen. Je kunt dan je klassenindeling kiezen, zorg voor ongeveer tien klassen. Bijvoorbeeld 155– < 160, enzovoorts.
- Maak eerst een kolom met de kleinste getallen in de tien klassen in (bijvoorbeeld) F3 t/m F13, zet de klassenmiddens ernaast en daarnaast weer de grootste getallen in de klassen. Bij de klasse 155– < 160 worden dat 150, 157,5 en 159.
- De cellen I3 t/m I13 zijn bestemd voor de frequenties. Selecteer die allemaal.
- Klik nu in de formulebalk en zet daar in: =INTERVAL(D2:D155;H3:H13) en doe [ctrl][shift][enter].

Als het goed is krijg je nu de juiste frequenties. Excel telt dus in de cellen D2 t/m D155 hoeveel er steeds maximaal gelijk zijn aan de grootste waarde in de klasse en groter zijn dan de grootste waarde van de vorige klasse.

Misschien wil je de jongens en de meisjes vergelijken? Dan laat je Excel eerst sorteren op de kolom 'geslacht'. Dat doe je door die kolom te selecteren en 'Sorteren en filteren' te kiezen. Sorteert van laag naar hoog en klik op OK als Excel je vraagt of de selectie moet worden uitgebreid (want de andere kolommen moeten ook worden meegenomen bij het sorteren). Nu heb je de jongens en de meisjes als groepen bij elkaar staan en kun je hun centrummaten en spreidingsmaten vergelijken en twee boxplots in één figuur maken.

**LET OP:** Je moet het Excel-bestand Gegevens154Leerlingen.xls eerst downloaden om er in te kunnen werken. Je kopieert het dan naar de computer waarop je werkt.

# Antwoorden

**3.1 a** Ja, een klassenindeling is hier verstandig, de leeftijden variëren van 17 tot 81 dus zonder klassenindeling krijg je een erg grote en onoverzichtelijke tabel en/of grafiek.

Als eerste klasse zou je bijvoorbeeld kunnen nemen:  $15- < 25$ . Meerdere antwoorden mogelijk.

**b** Nee, je kunt geen klassenindeling maken bij kleuren, dat gaat alleen bij getallen. Het enige wat je zou kunnen doen is zeggen dat bijvoorbeeld limoengroen en mosgroen allebei onder groen vallen. Maar aangezien er maar 12 verschillende kleuren genoemd zijn is dit niet echt nodig, je kunt hier prima een overzichtelijke tabel of grafiek bij maken.

**c** Ja, een klassenindeling kan wel. Je zou dan als eerste klasse bijvoorbeeld kunnen nemen:  $1- < 3$ , maar ook een klassenindeling als  $0,5- < 1,5$ , etc., is mogelijk en dan komt het op hetzelfde neer als geen klassenindeling maken.

**3.2 a** 47,5, het gemiddelde van de klassengrenzen. Dat komt omdat voor de gewichten alle waarden vanaf 45 tot aan 50 mogelijk zijn.

**b**  $\frac{3875}{60} \approx 64,6$  kg.

**c**  $\frac{48}{60} \approx 0,80$ , dus 80%.

**3.3 a**  $5 \cdot 6 = 30$  ogen.

**b**  $\frac{1074}{60} \approx 17,9$  ogen.

**c** Zie de tabel.

klasse	freq.
5- < 8	1
8- < 11	0
11- < 14	5
14- < 17	15
17- < 20	17
20- < 23	14
24- < 26	7
26- < 29	1

**d** Tot de klasse  $5- < 8$  horen alleen de getallen 5, 6 en 7. Het klassenmidden is dus 6. Je kunt niet het gemiddelde van de klassengrenzen nemen want niet alle getallen vanaf 5 tot 8 zijn mogelijk, het is nu de middelste van alleen gehele getallen.

**e**  $\frac{1086}{60} \approx 18,1$  ogen.

**f** Dit gaat het makkelijkst met de frequentietabel:  $\frac{46}{60} \approx 0,767$  dus in 76,7% van de worpen.

**g** Omdat de score 18 op meer manieren met vijf dobbelstenen te maken is dan de score 28.



**3.4 a** Zie de tabel.

leeftijdsklasse	freq.
10– < 15	3
15– < 20	9
20– < 25	4
25– < 30	9
30– < 35	7
35– < 40	2
40– < 45	2
45– < 50	2
50– < 55	3
55– < 60	1

- b** Tot de klasse 50– < 55 horen de leeftijden 50 tot en met 54,999....  
De klassengrenzen zijn 50 en 55, dus het klassenmidden is 52.
- c**  $\frac{1231}{44} \approx 27,98$ , dus ongeveer 28 jaar.
- d** Dit gaat het makkelijkst met de frequentietabel:  $\frac{25}{44} \approx 0,568$  dus 56,8% van de bezoekers.

**3.5 a** 100– < 110

- b** 75 leerlingen.
- c**  $\frac{10725}{75} = 143$  cm.
- d** Je krijgt nu  $\frac{10650}{75} = 142$  cm. Het verschil ontstaat doordat bij het samenvoegen van twee klassen geen rekening is gehouden met de verschillen in aantal binnen de afzonderlijke klassen.

**3.6 a** Zie de tabel.

	jongens	meisjes
gewichtsklasse	abs.freq.	abs.freq.
1500– < 2000	1	0
2000– < 2500	1	1
2500– < 3000	5	9
3000– < 3500	6	12
3500– < 4000	1	6
4000– < 4500	12	2
4500– < 5000	4	0
totaal	30	30

- b** Jongens:  $\frac{111000}{30} = 3700$  gram.  
Meisjes:  $\frac{97000}{30} \approx 3233$  gram.

c  $\frac{20}{30} = 0,667$  en dat is 67,7%.

d Het zwaarste meisje woog 4000 gram.

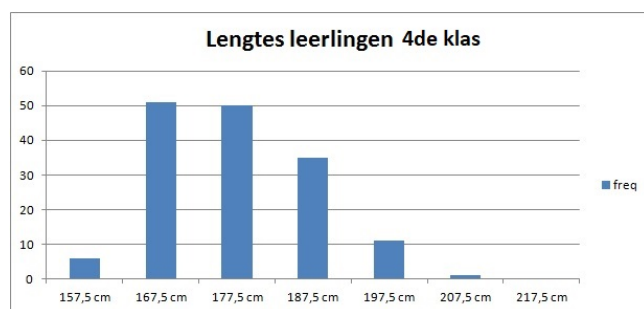
Er zijn 13 jongens die meer dan 4000 gram wegen.

Dus:  $\frac{13}{30} \cdot 100 \approx 43,3\%$ .

3.7 a Zie de tabel.

lengte klasse	frequentie
150– < 159	6
160– < 169	51
170– < 179	50
180– < 189	35
190– < 199	11
200– < 209	1
209– < 219	0

b Zie de figuur.



c Zie de tabel.

lengte klasse	freq meisjes	freq jongens
150– < 159	6	0
160– < 169	46	5
170– < 179	26	24
180– < 189	6	29
190– < 199	1	10
200– < 209	0	1
209– < 219	0	0

3.8 a Vergelijk je resultaten met die van de andere leerlingen.

b Vergelijk je resultaten met die van de andere leerlingen.

## 1.4 Schatten

### Inleiding

Aicha heeft leren werken met klassenindelingen en de bijbehorende frequentietabellen.

In plaats van werken met een brij van onafgeronde rapportcijfers, werkt ze nu met overzichtelijke frequentietabellen waarin alle cijfers zijn afgerond op gehele getallen, zoals deze voor het vak Frans. Dat zijn dus eigenlijk klassenindelingen.

Voordeel ervan is het krijgen van goed overzicht als je met veel ruwe gegevens te maken hebt.

Nadeel is dat je alle centrummaten en spreidingsmaten alleen nog maar kunt schatten.

cijfer	frequentie
4	1
5	4
6	7
7	6
8	6
9	6

Tabel 4.1

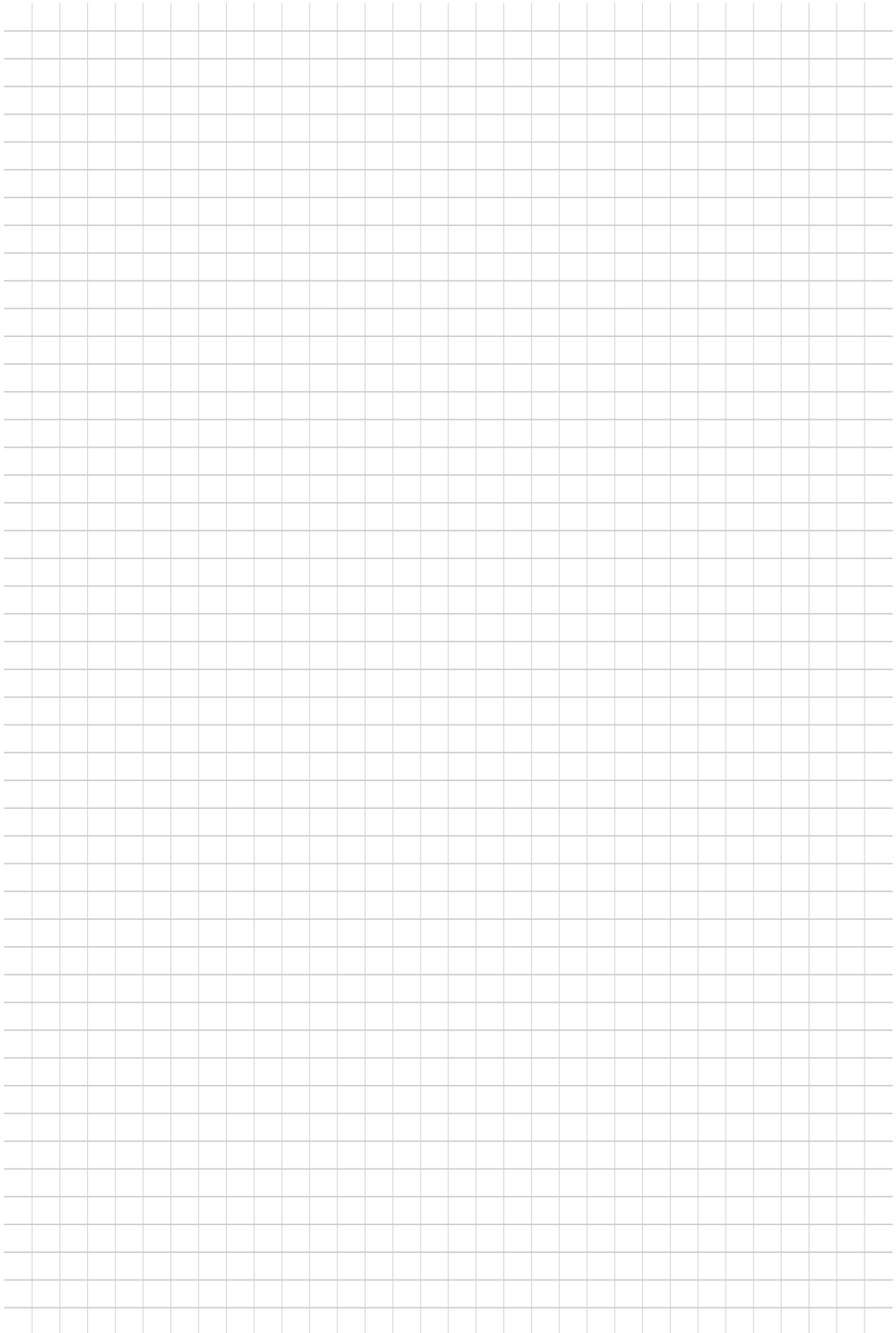
### Je leert in dit onderwerp

- bij een klassenindeling de centrummaten en de spreidingsmaten schatten;
- bij een klassenindeling een boxplot maken.

### Voorkennis

- (gewogen) gemiddelde, modus en mediaan van een hoeveelheid gegevens berekenen;
- een frequentietabel met klassenindeling en geschikte klassenbreedte maken bij een verzameling ruwe data.

## Aantekeningen

A large grid of graph paper with 28 columns and 40 rows, intended for taking notes.

## Theorie

### Om te onthouden

A large grid of graph paper with a light blue background and a fine grid of light gray lines, intended for taking notes.

## Verwerken

### ★ Opgave 4.1

In een klas zijn voor een toets de volgende cijfers behaald:

6,5	9,3	5,2	6,1	6,2	6,2	2,6	3,6	9,6	7,9
7,4	8,4	7,1	4,4	9,0	8,3	8,3	4,8	8,2	6,5
6,5	5,8	6,4	5,4	6,0	6,2	7,2	5,8	6,3	6,9

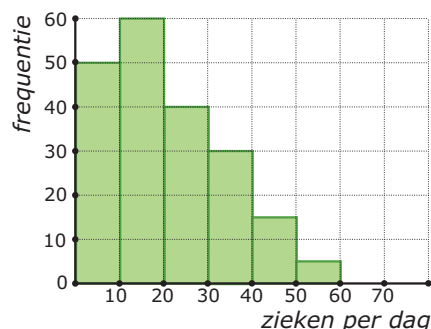
Tabel 4.2

- Bepaal de variatiebreedte.
- Als je een staafdiagram van deze gegevens wilt maken is het verstandig om eerst een klassenindeling te maken. Waarom is dat?
- Maak een klassenindeling van de resultaten en teken het staafdiagram. Welke klassen liggen hier voor de hand?
- Bepaal de modale klasse.
- In welke klasse ligt de mediaan? Hoeveel wordt de mediaan naar schatting?
- Hoeveel wordt het eerste kwartiel naar schatting?

### ★ Opgave 4.2

Op een school is gedurende een jaar elke schooldag bijgehouden hoeveel leerlingen er ziek waren. Je ziet de gegevens hier verwerkt in een staafdiagram. De laatste klasse bevat in feite alle dagen dat het aantal zieken 50 of meer was.

- Hoeveel schooldagen waren er dat jaar?
- Schat de mediaan.
- Schat het gemiddelde aantal zieke leerlingen per dag in gehelen. Neem voor de laatste klasse als klassenmidden 55.
- Je hebt voor de laatste klasse als klassenmidden 55 genomen. Voor welke centrummaat is die keuze van belang? Licht je antwoord toe.



Figuur 4.1

### ★ Opgave 4.3

De lengtes in cm van 60 leerlingen zijn in klassen ingedeeld. Je kunt nu het gemiddelde alleen nog schatten want je weet de ruwe data niet.

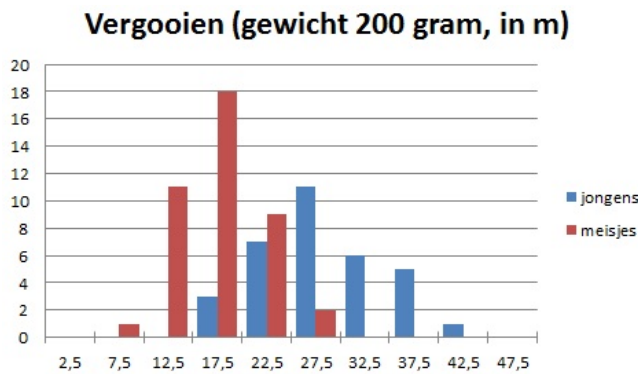
- a Bepaal de modale klasse.
- b Hoe schat je het gemiddelde? Bereken deze schatting.
- c Schat de mediaan.
- d Schat de kwartielen en teken een bijpassend boxplot.

lengteklasse	frequentie
150– < 160	3
160– < 170	17
170– < 180	34
180– < 190	4
190– < 200	2
totaal	60

Tabel 4.3

★ **Opgave 4.4**

Hier zie je in één staafdiagram de resultaten van het gooien met een 200 grams kogel van een groep van 74 brugklasleerlingen. Er zijn afzonderlijke diagrammen voor de jongens en de meisjes. Alle geworpen afstanden zijn tot op 0,5 m nauwkeurig gemeten. Elke klasse is door het klassenmidden aangegeven.



Figuur 4.2

- a Maak bij beide staafdiagrammen een boxplot.
- b Maak een schatting van het gemiddelde van de jongens en de meisjes afzonderlijk.
- c Kun je concluderen dat jongens van die leeftijd verder gooien dan meisjes? Licht je antwoord toe.

**Toepassen**

Je hebt de belangrijkste vaardigheden met MS-Excel in de voorgaande onderdelen voorbij zien komen. Je hebt centrummaten en spreidingsmaten berekend (dat doe je het liefst vanuit de ruwe gegevens) en je hebt diagrammen getekend. Voor een boxplot gebruik je ook het liefst de ruwe gegevens, maar voor een staafdiagram of een lijndiagram of een cirkeldiagram werk je meestal met een klassenindeling. In de volgende opgaven werk je met het bestand [Gegevens154Leerlingen.xls](#).

In dit bestand vind je een tabel met het aantal uren huiswerk per week. Je gaat de jongens en de meisjes van deze vierde klas met elkaar vergelijken.

Denk er wel om dat je op een pc het bijbehorende Excel-bestand eerst moet downloaden.

★ **Opgave 4.5: Uren huiswerk per week**

Gebruik het gegevensbestand uit **Toepassen**.

- a Maak een geschikte frequentieverdeling van het aantal uren dat de jongens aan hun huiswerk besteden. Doe dit ook voor de meisjes.
- b Maak met Excel een bijpassend lijndiagram van de jongens en meisjes in één figuur.
- c Probeer een conclusie te trekken.

★ **Opgave 4.6: Uren huiswerk per week en boxplots**

Werk met hetzelfde Excel-bestand als in de voorgaande opgave.

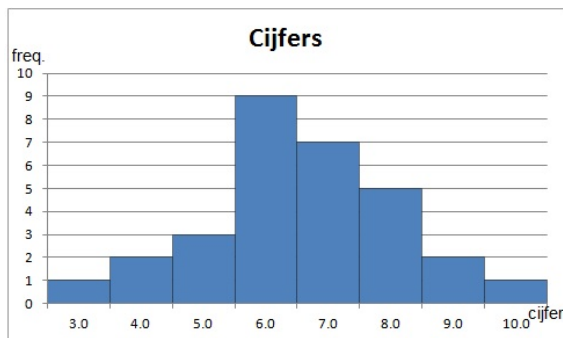
- a Maak nu twee boxplots naast elkaar voor het aantal uren dat de jongens per week aan hun huiswerk besteden en het aantal uren dat de meisjes per week aan hun huiswerk besteden.
- b Probeer opnieuw conclusies te trekken, betrek ook de gemiddelden erbij.



# Antwoorden

- 4.1 a**  $9,6 - 2,6 = 7,0$
- b** Omdat je anders heel veel staafjes van lengte 1 krijgt. En daaruit kun je nauwelijks wat zinnigs aflezen.
- c** Zie de tabel hieronder. Je neemt de klassen  $2,5 - < 3,5$ , enzovoorts. Maak een staafdiagram bij deze tabel.

klasse	freq.
$2,5 - < 3,5$	1
$3,5 - < 4,5$	2
$4,5 - < 5,5$	3
$5,5 - < 6,5$	9
$6,5 - < 7,5$	7
$7,5 - < 8,5$	5
$8,5 - < 9,5$	2
$9,5 - < 10,5$	1



- d**  $5,5 - < 6,5$
- e** Precies op de grens van de klassen  $5,5 - < 6,5$  en  $6,5 - < 7,5$ . Dus die zou je schatten als 6,5.
- f**  $Q_1 \approx 5,5 + 2 \cdot \frac{1}{9} \approx 5,7$ .
- 4.2 a**  $50 + 60 + 40 + 30 + 15 + 5 = 200$
- b** Voor het 100e en 101e getal van de hele reeks moet je het gemiddelde van het 50e en 51e getal uit de klasse  $10 - < 20$  schatten. Dat is  $10 + 10 \cdot \frac{50,5}{60} \approx 18$ .
- c**  $\frac{5 \cdot 50 + 15 \cdot 60 + 25 \cdot 40 + 35 \cdot 30 + 45 \cdot 15 + 55 \cdot 5}{200} = 20,75$ , dus afgerond op gehele is dat 21 zieken per dag.
- d** Dat is alleen belangrijk voor het gemiddelde, de mediaan en de modale klasse veranderen niet als je dat klassenmidden anders kiest.
- 4.3 a** De modale klasse (meest voorkomende) is:  $170 - < 180$ .
- b** Je schat het gemiddelde door de klassenmiddens te gebruiken.

Het geschatte gemiddelde is:  $\frac{10350}{60} \approx 172,5$ .

c De mediaan is het gemiddelde van het 30e en 31e getal, dus van  $170 + \frac{10}{34} \cdot 10 \approx 172,9$  en  $170 + \frac{11}{34} \cdot 10 \approx 173,2$ . De geschatte mediaan is daarom 173.

d Het eerste kwartiel is het gemiddelde van het 15e en 16e getal, dus van  $160 + \frac{12}{17} \cdot 10 \approx 167,1$  en  $160 + \frac{13}{17} \cdot 10 \approx 167,6$ . Het eerste kwartiel is dus ongeveer 167.

Zo is het derde kwartiel ongeveer 177.

Met het minimum van 150 en het maximum van 200 kun je het boxplot tekenen.

**4.4 a** Geschatte gegevens afgerond op 0,5 m:

Voor de 33 jongens:

minimum: 15;

eerste kwartiel, dus de 8e en 9e jongen:  $20 + 5 \cdot \frac{5,5}{7} \approx 24,0$ ;

mediaan, dus de 17e jongen:  $25 + 5 \cdot \frac{7}{11} \approx 28,0$ ;

derde kwartiel, dus de 25e en 26e jongen:  $30 + 5 \cdot \frac{4,5}{6} \approx 33,5$ ;

maximum 45.

Voor de 41 meisjes:

minimum 5;

eerste kwartiel 14,0;

mediaan 17,5;

derde kwartiel 20,5;

maximum 30.

Teken bijpassende boxplots.

**b** Jongens: 28,5 m (afgerond op 0,5 m).

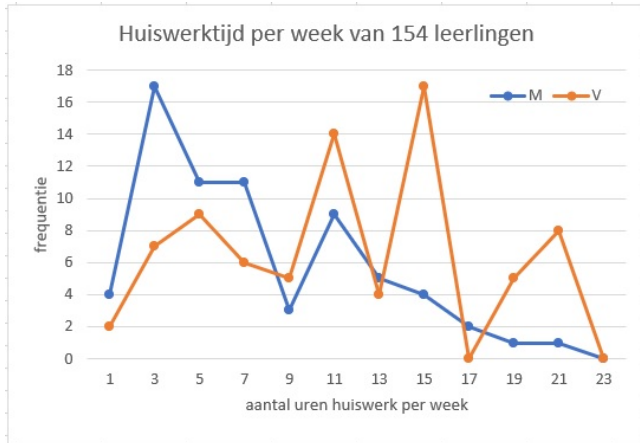
Meisjes: 17,5 m.

**c** Voor deze groep kun je die conclusie wel trekken. De gemiddelden verschillen nogal: deze jongens gooien gemiddeld nogal wat verder dan deze meisjes. Aan de boxplots zie je dat meer dan 75% van de jongens verder gooit dan het derde kwartiel bij de meisjes.

**4.5 a** Zie tabel in Excel. Niet alle leerlingen hebben hun huiswerktijd ingevuld.

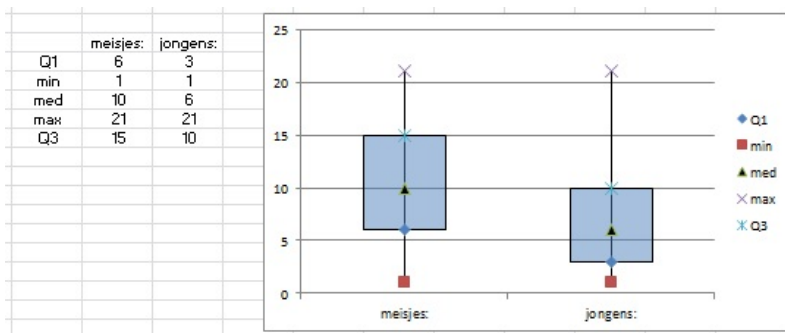
huiswerkuren per week				
laagste	midden	hoogste	M	V
0	1	1,9	4	2
2	3	3,9	17	7
4	5	5,9	11	9
6	7	7,9	11	6
8	9	9,9	3	5
10	11	11,9	9	14
12	13	13,9	5	4
14	15	15,9	4	17
16	17	17,9	2	0
18	19	19,9	1	5
20	21	21,9	1	8
22	23	23,9	0	0
			68	77

b Zie de figuur.



c Over het algemeen besteden meisjes meer tijd aan het huiswerk dan jongens. Want aan de rechterkant van de grafiek (waar de vele uren huiswerk zitten) ligt de lijn van de meisjes boven die van de jongens, en juist aan de linkerkant (waar weinig huiswerkuren zitten) ligt de lijn van de jongens boven die van de meisjes.

4.6 a Hier zie je de boxplots met de gegevens die Excel berekent.



b Vergelijk je resultaten met die van de andere leerlingen.

Het lijkt er op dat het grootste deel van de jongens veel minder tijd aan huiswerk besteedt dan de meisjes. 75% van de jongens besteedt minder tijd dan de mediaan van de meisjes. Maar er zijn uitschieters. Je zou kunnen besluiten om opnieuw te kijken als je de echte uitschieters weglaat...

## 1.5 Statistische uitspraken

### Inleiding

Je hebt inmiddels allerlei tabellen en diagrammen leren maken om statistische gegevens overzichtelijker weer te geven.

Die gegevens krijg je door statistisch onderzoek.

Aicha vindt het hoog tijd worden om zelf eens zo'n onderzoek te doen.

Maar wat zijn eigenlijk 'statistische gegevens' - ook wel data genoemd - precies?

En wat is een vraag die je moet beantwoorden met behulp van statistiek? Waarom doe je eigenlijk statistiek, waarom is dit nodig?



Figuur 5.1

### Je leert in dit onderwerp

- een statistische uitspraak herkennen;
- bij een statistisch onderzoek de begrippen populatie en steekproef herkennen;
- aangeven of een steekproef representatief is;
- een eenvoudig statistisch onderzoek opzetten.

### Voorkennis

- (gewogen) gemiddelde, modus en mediaan van een hoeveelheid gegevens berekenen;
- een frequentietabel met klassenindeling en geschikte klassenbreedte maken bij een verzameling ruwe data;
- werken met diverse tabellen en er verschillende diagrammen bij maken;
- bij een klassenindeling de centrummaten en de spreidingsmaten schatten;
- bij een klassenindeling een lijndiagram, een staafdiagram en een boxplot maken.



## Aantekeningen

A large grid of graph paper for taking notes, consisting of approximately 25 columns and 40 rows of small squares.



## Theorie

### Om te onthouden

A large grid of graph paper with 20 columns and 30 rows, intended for taking notes on the theory of statistics.



## Verwerken

### ★ Opgave 5.1

Bij statistisch onderzoek wil je een uitspraak doen over een grote groep (de populatie) op grond van de resultaten van een steekproef uit die groep.

- a Waar moet een goede steekproef aan voldoen?
- b Welke twee soorten statistisch onderzoek ken je op dit moment? Geef van beide een voorbeeld.

### ★ Opgave 5.2

Als je onderzoek wilt doen naar de hobby's van Nederlanders van 18 jaar en ouder, dan zijn belangrijke kenmerken van de populatie: geslacht, leeftijd, regio, en opleiding.

Bedenk bij elk van de volgende uitspraken om welke populatie het gaat en welke kenmerken voor die populatie belangrijk zijn om na te gaan of de steekproef representatief is.

- a “De boeren in Europa die kippeneieren produceren zitten in een crisis.”
- b “Er zijn veel mensen in Nederland van plan om komend jaar hun huis te gaan verbouwen.”
- c “Nederlandse jongens zijn beter in atletiek dan Nederlandse meisjes.”
- d “Vrouwen gaan minder vaak naar de kapper dan mannen.”

### ★ Opgave 5.3

Je bent journalist en wilt een artikel schrijven over studenten. Je bent vooral nieuwsgierig naar mensen die een mbo studie gaan doen en voor welke sector ze dan kiezen. Je hebt al wat rondgevraagd en krijgt het vermoeden dat vooral veel meisjes voor de zorgsector kiezen, en dat veel minder jongens dat doen.

Je hebt inmiddels met 25 studenten gesproken die een mbo opleiding in de zorg zijn gestart. 15 van deze studenten zijn vrouw en 10 zijn man. Je schrijft als kop boven je artikel: “Meisjes kiezen veel vaker voor een mbo opleiding in de zorg dan jongens.”

- a Is dit een statistische uitspraak? Leg uit waarom wel of niet.  
Zoek ook op het internet of je meer gegevens kunt vinden over de keuze voor de sectoren in het mbo. En ook of dit te maken heeft met het geslacht. Kijk eerst welke sectoren er zijn via [de website van de MBO-raad](#).
- b Hoe zou je een onderzoek opzetten waarmee je een artikel over de sectorkeuze in het mbo gaat schrijven?

## Toepassen

Hier vind je het Excel bestand [Gegevens154Leerlingen](#). Dit bestand bevat informatie over een groep vierde klassers. Deze gegevens ga je gebruiken om een statistische onderzoeksvraag te beantwoorden die jij zelf bedenkt.

### ★ Opgave 5.4: Een eigen onderzoeksvraag bedenken

Gebruik het bestand uit [Toepassen](#).

- a Bekijk de gegevens in het Excel bestand en formuleer een goede statistische onderzoeksvraag.
- b Gebruik de gegevens uit het Excel bestand om één duidelijke grafiek te maken die past bij je onderzoeksvraag.



- c Trek aan de hand van je grafiek conclusies en beantwoord je onderzoeksvraag.

★ **Opgave 5.5: Een eigen statistisch onderzoek doen**

Het is nu tijd om zelf een statistisch onderzoek te doen. Kies een onderwerp gebaseerd op de twee soorten statistische onderzoeksvragen waar je tot nu toe mee hebt kennisgemaakt. Bedenk vervolgens hoe je het onderzoek kunt uitvoeren, hoe je de gegevens kunt verzamelen en overzichtelijk weergeven. En bedenk tenslotte hoe je conclusies kunt trekken en verwoorden.

- a Bedenk een eigen statistische onderzoeksvraag.
- b Stel een plan van aanpak op. Daarin beschrijf je hoe je steekproef wordt samengesteld en hoe je het onderzoek gaat aanpakken.
- c Voer je onderzoek uit en verzamel alle gegevens (bijvoorbeeld in MS-Excel).
- d Bedenk een manier om je gegevens goed te presenteren, denk aan tabellen en grafieken.
- e Trek een conclusie.
- f Kijk terug op je onderzoek en op de aanpak. Wat had beter gekund?



# Antwoorden

- 5.1 a** De steekproef moet representatief zijn voor de populatie, dus een goed beeld geven van de populatie. Verder moeten de deelnemers aan die steekproef willekeurig worden gekozen.
- b** Bij de éne soort wil je bijvoorbeeld weten hoeveel procent van alle Nederlanders kleurenblind is. Bij de andere soort vergelijk je twee groepen, bijvoorbeeld de levensduur van twee soorten lampen.
- 5.2 a** Populatie: Alle boeren in Europa die kippeneieren produceren.  
Kenmerken: het land waarin de boer woont (hoeveel procent van de kippenboeren woont in Nederland, Duitsland, België, etc), maar ook de regio per land waarin de boer woont, leeftijd van de boeren, leeftijd van het bedrijf/boerderij, de grootte van het bedrijf, soort eieren (scharreleieren, vrije uitloop eieren, etc).
- b** Populatie: Alle inwoners van Nederland die een eigen huis hebben. (ervanuitgaande dat je een huurhuis niet gaat verbouwen).  
Kenmerken: Regio, inkomen, bouwjaar huis, oppervlakte huis, leeftijd, aantal kinderen, huwelijkse staat.
- c** Populatie: Alle jongens en meisjes die aan atletiek doen in Nederland.  
Kenmerken: geslacht, leeftijd, niveau, regio, aantal jaren ervaring, favoriete atletiek onderdeel, lengte, gewicht, welke atletiekvereniging.
- d** Populatie: Alle mannen en vrouwen ter wereld die wel eens naar de kapper gaan.  
Kenmerken: geslacht, land, regio, afkomst, soort haar (lang, kort, krullen, kroeshaar, etc), inkomsten.
- 5.3 a** Het is een statistische uitspraak, je kunt dit nagaan door veel gegevens te verzamelen en jongens en meisjes met elkaar te vergelijken. Die 25 mensen die je tot nu toe hebt ondervraagd zijn alleen nog lang niet genoeg.
- b** Eigen antwoord, waarin je een uitgebreide beschrijving van een onderzoek geeft.  
Je kunt voor mannen en vrouwen afzonderlijk proberen op te zoeken hoeveel leerlingen er naar het mbo gaan en voor welke sector ze dan kiezen. Gebruik de sectoren die de MBO-raad onderscheidt. Je hebt dan achtergrondinformatie. Maar je wilt natuurlijk ook weten hoe dit bij de keuzes van huidige vmbo-scholieren zit. Of misschien zijn er ook havo-leerlingen die naar het MBO willen?  
Houd daarom een enquête (bijvoorbeeld onder vierde klassers op vmbo/havo) over de keuzes die zij denken te gaan maken. Probeer daaruit conclusies te trekken. Beschrijf ook hoe goed je steekproef is.
- 5.4 a** Eigen antwoord, vraag aan je docent of je een goede onderzoeksvraag hebt.  
Voorbeelden:  
Hoeveel procent van de meisjes heeft een wiskundecijfer boven de 7?  
Hoeveel procent van de jongens gaat met behoorlijk tot veel plezier naar school?  
Hoeveel procent van de jongens is langer dan 170 cm?  
Zijn jongens langer dan meisjes?  
etc.
- b** Eigen antwoord. Vraag aan je docent of je het juiste type grafiek hebt gekozen en of je de grafiek goed hebt gemaakt.  
Let op: Niet ieder grafiektype past bij iedere onderzoeksvraag.  
Voorbeeld:



Hoeveel procent van de jongens gaat met behoorlijk tot veel plezier naar school? Je zou een cirkeldiagram kunnen kiezen, je bent hier geïnteresseerd in de sectoren 'behoorlijk' en 'veel'. Het zou erg raar zijn om hier een lijndiagram bij te maken. Dat wekt de indruk dat er ook nog keuzes zijn tussen behoorlijk en veel in.

**c** Vraag aan je docent of je de juiste conclusie(s) hebt getrokken.

**5.5 a** Kies één van onderstaande twee soorten onderzoeksvragen en bedenk van die soort zelf een onderzoeksvraag:

- onderzoek waarbij je op zoek gaat naar het percentage van een bepaalde populatie dat een zekere eigenschap heeft;
- onderzoek waarbij je twee groepen in een populatie met elkaar vergelijkt.

Laat je onderzoeksvraag door je docent goedkeuren.

- b**
- Bedenk eerst wat de populatie is die bij je onderzoeksvraag hoort.
  - Welke kenmerken van deze populatie zijn belangrijk? (geslacht, leeftijd, regio, etc) Probeer hier percentages van te vinden op internet.
  - Bedenk hoe je een steekproef gaat samenstellen.
  - Welke vragen ga je stellen aan de steekproef? (Vraag ook naar de kenmerken zodat je achteraf kunt controleren hoe representatief je steekproef was).

Laat je plan van aanpak door je docent goedkeuren.

**c** Doen.

**d** Manieren om gegevens te presenteren:

- (Frequentie)tabel met of zonder klassenindeling
- Staafdiagram
- Lijndiagram
- Boxplot
- Cirkeldiagram
- Steelbladdiagram
- etc.

Niet iedere manier van presenteren past bij ieder onderzoek, bedenk goed wat bij jouw onderzoek past en waarom.

**e** Laat de resultaten van je onderzoek en je conclusie aan je docent zien.

De conclusie moet leiden tot een antwoord op je onderzoeksvraag.

**f** Schrijf voor jezelf op wat goed ging en wat beter kon, bespreek dit daarna met je docent.

## 1.6 Totaalbeeld

### Samenvatten

#### Begrippenlijst

- frequentietabel — centrummaat — modus — mediaan — gemiddelde
- boxplot — kwartiel — spreidingsbreedte — (inter)kwartielafstand
- klassenindeling — klassengrenzen, klassenmidden, klassenbreedte
- schatten
- statistisch onderzoek — populatie — steekproef — representatief

#### Activiteitenlijst

- gegevens samenvatten in frequentietabellen en beschrijven met behulp van centrummaten;
- gegevens samenvatten in frequentietabellen en beschrijven met behulp van spreidingsmaten en boxplots;
- klassenindelingen gebruiken, het gemiddelde schatten;
- centrummaten en spreidingsmaten schatten vanuit een klassenindeling;
- kennismaken met statistisch onderzoek, populatie en representatieve steekproef

#### ★ Opgave 6.1

In deze frequentietabel zie je de resultaten van tellingen van het aantal eieren dat dagelijks in een hok vol kippen werd geraapt.

aantal eieren	35	36	37	38	39	40	41	42	43	44
frequentie	1	2	3	6	12	15	27	35	23	12

Tabel 6.1

Deze gegevens kun je samenvatten met behulp van een drietal centrummaten.

- Welk aantal eieren per dag is de modus?
- Welk aantal eieren per dag is de mediaan?
- Bereken het gemiddelde aantal eieren per dag.

#### ★ Opgave 6.2

In deze frequentietabel zie je de resultaten van tellingen van het aantal eieren dat dagelijks in een hok vol kippen werd geraapt.

aantal eieren	35	36	37	38	39	40	41	42	43	44
frequentie	1	2	3	6	12	15	27	35	23	12

Tabel 6.2

Bij deze gegevens kun je behalve de drie centrummaten ook een paar spreidingsmaten vaststellen.

- Bepaal de spreidingsbreedte.
- Bepaal de kwartielafstand.

c Teken een bijpassend boxplot.

★ **Opgave 6.3**

De lengtes in cm van 60 leerlingen zijn in klassen ingedeeld. Je kunt nu het gemiddelde alleen nog schatten want je weet de ruwe data niet.

- a Bepaal de modale klasse.
- b Hoe schat je het gemiddelde? Bereken deze schatting.

lengteklasse	frequentie
150– < 160	3
160– < 170	17
170– < 180	34
180– < 190	4
190– < 200	2
totaal	60

Tabel 6.3

★ **Opgave 6.4**

Gebruik de frequentietabel van de vorige opgave. Door de mediaan en de kwartielen te schatten kun je een boxplot maken.

- a Schat de mediaan.
- b Schat de kwartielen en teken een bijpassend boxplot.

★ **Opgave 6.5**

Bij statistisch onderzoek wil je een uitspraak doen over een grote groep (de populatie) op grond van de resultaten van een steekproef uit die groep.

- a Waar moet een goede steekproef aan voldoen?
- b Welke twee soorten statistisch onderzoek ken je op dit moment? Geef van beide een voorbeeld.

**Testen**

★ **Opgave 6.6**

Martin heeft elke schooldag van drie brugklassen het aantal telaatkomers genoteerd.

aantal leerlingen te laat	0	1	2	3	4	5	6	7
frequentie	5	6	6	8	4	2	3	6

Tabel 6.4

- a Bepaal de modus en de mediaan.
- b Hoeveel weken hield Martin het te laat komen bij?
- c Bereken het gemiddelde in één decimaal nauwkeurig.

★ **Opgave 6.7**

Tijdens een milieuonderzoek noteerde men in een eikenbos van 30 bomen de omtrek in centimeters. De omtrek is telkens gemeten op een hoogte van 1 m boven de grond.

125	126	128	129	131	131	135	135	135	135	135	138	139	142	142
142	142	143	150	150	151	151	151	155	160	161	165	165	165	166

Tabel 6.5

- a Bepaal de spreidingsbreedte.
- b Bepaal de mediaan.
- c Bepaal beide kwartielen en bereken de kwartielafstand.
- d Teken de bijbehorende boxplot.

★ **Opgave 6.8**

Hier zie je nog een keer de omtrek van 30 bomen in centimeters.

125	126	128	129	131	131	135	135	135	135	135	138	139	142	142
142	142	143	150	150	151	151	151	155	160	161	165	165	165	166

Tabel 6.6

- a Maak bij deze waarnemingsgetallen een frequentietabel. Neem een klassenbreedte van 10 cm en als ondergrens van de eerste klasse 120 cm.
- b Maak een bijpassend staafdiagram.
- c Van hoeveel procent van deze bomen is de omtrek groter dan 150 cm?
- d Schat hoeveel procent van deze bomen een omtrek heeft die groter is dan de gemiddelde omtrek.
- e Je kunt de gemiddelde omtrek ook schatten vanuit de frequentieverdeling. Welke waarde vind je dan voor het gemiddelde? En waarom is dat hoger dan het werkelijke gemiddelde?

★ **Opgave 6.9**

Voor een biologiepracticum zijn op twee velden regenwormen gevangen en is hun lengte in gehele cm gemeten. In de frequentietabel zie je de resultaten. Je gaat de gegevens van beide velden vergelijken.

- a Leg uit waarom het klassenmidden van de eerste klasse 1,5 is.
- b Schat de gemiddelde lengte van de regenwormen op veld 1. Doe hetzelfde voor veld 2.
- c Als je de twee staafdiagrammen bij deze verdelingen wilt vergelijken dan kun je het beste alle frequenties omrekenen naar relatieve frequenties. Waarom is dat? Teken vervolgens beide staafdiagrammen in één figuur.
- d Hoeveel procent van de regenwormen op veld 1 is 15 cm of langer? En op veld 2?

regenwormen		
lengte (cm)	veld 1	veld 2
0-<3	4	1
3-<6	7	1
6-<9	15	5
9-<12	20	9
12-<15	23	15
15-<18	15	18
18-<21	4	12
21-<24	1	6
24-<27	1	3
	90	70

Figuur 6.1

★ **Opgave 6.10**

Gebruik de gegevens over de regenwormen uit de vorige opgave. Elke regenworm is gemeten in gehele cm nauwkeurig.

- a Welke klasse is de modale klasse op veld 1? En op veld 2?
- b Maak bij beide velden een boxplot van de verdeling van de lengtes van de regenwormen.
- c Kun je een conclusie trekken? Motiveer je antwoord.

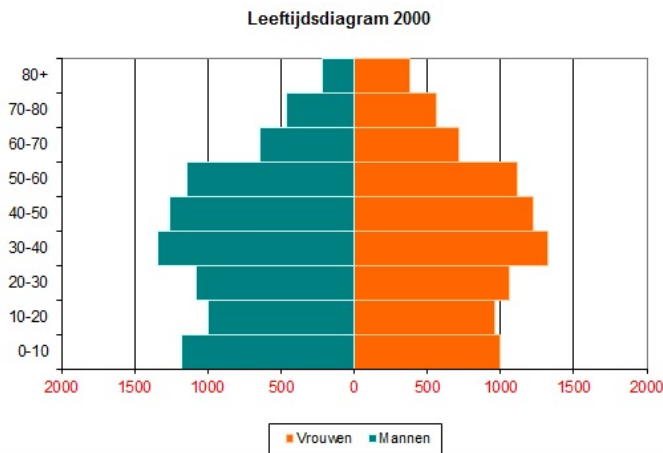
★ **Opgave 6.11**

Stel je voor dat je wilt weten hoeveel procent van de scholieren in het voortgezet onderwijs in Nederland met de fiets naar school gaat. Je gaat dit uitzoeken door middel van een steekproef.

- a Hoe ga je die steekproef samenstellen? Noem minstens twee punten waar je rekening mee moet houden.
- b Zou je dit onderzoek via het internet kunnen uitvoeren? Is dat een goede aanpak?
- c Kun je een betere aanpak verzinnen? Motiveer je antwoord.

**Toepassen**

Aicha ziet bij aardrijkskunde een zogenaamde **bevolkingspiramide**. Hierin is de bevolkingsopbouw van Nederland in het jaar 2000 weergegeven. Het lijkt wel een staafdiagram, maar dan op zijn kant.



Figuur 6.2

**Opgave 6.12: Bevolkingspiramide in 2000**

In **Toepassen** zie je een bevolkingspiramide, een leeftijdsdiagram van de Nederlandse bevolking in het jaar 2000. Alle aantallen zijn duizendtallen.

Je kunt er bijvoorbeeld in aflezen dat de groep 0-4-jarigen in 2000 uit ongeveer 1.200.000 jongens en ongeveer 1.000.000 meisjes bestond.

- a Om wat voor soort diagrammen gaat het hier?
- b Hoeveel mannen en vrouwen van 10– < 20 waren er in Nederland in 2000?
- c In de jaren na de Tweede Wereldoorlog werden er nogal veel kinderen geboren. Dat werd de ‘baby-boom’ genoemd. Hoe vind je dat in de bevolkingspiramide van 2000 terug?
- d Wat was in 2000 de modale leeftijdscategorie?
- e Schat de gemiddelde leeftijd van de Nederlanders in 2000.

**Opgave 6.13: Bevolkingspiramide vervolg**

Bekijk de bevolkingspiramide van de voorgaande opgave nog eens.

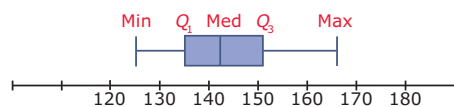
- a Schat de mediaan van de leeftijden van de Nederlandse mannen in 2000. Doe hetzelfde voor de vrouwen.
- b Teken boxplots voor zowel de mannen als de vrouwen in één figuur. Kun je een conclusie trekken?

Via de website **CBS: bevolkingspiramide** kun je zien hoe de bevolkingsopbouw in Nederland voor dit jaar er uitziet.

- c Beschrijf het verschil van de bevolkingsopbouw in Nederland in dit jaar met die in 2000.

# Antwoorden

- 6.1 a** De modus (meest voorkomende) is 42 eieren per dag.
- b** De mediaan (middelste) is 42, het gemiddelde van het 68e en 69e getal.
- c** Het gemiddelde is  $\frac{35 \cdot 1 + 36 \cdot 2 + \dots + 44 \cdot 12}{1 + 2 + 3 + \dots + 12} = \frac{5608}{136} \approx 41,2$  eieren per dag.
- 6.2 a** De spreidingsbreedte is:  $44 - 35 = 9$ .
- b** De kwartielen zijn:  $Q_1 = 39$  en  $Q_3 = 43$ . De kwartielafstand is:  $43 - 39 = 4$ .
- c** Je weet alle vijf getallen die je er voor nodig hebt. Geef ze bij je figuur aan.
- 6.3 a** De modale klasse (meest voorkomende) is:  $170- < 180$ .
- b** Je schat het gemiddelde door de klassenmiddens (155, 165, etc.) en de frequenties te gebruiken.  
Het geschatte gemiddelde is:  $\frac{10350}{60} \approx 172,5$ .
- 6.4 a** De mediaan is het gemiddelde van het 30e en 31e getal, dus van  $170 + \frac{10}{34} \cdot 10 \approx 172,9$  en  $170 + \frac{11}{34} \cdot 10 \approx 173,2$ . De geschatte mediaan is daarom 173.
- b** Het eerste kwartiel is het gemiddelde van het 15e en 16e getal, dus van  $160 + \frac{12}{17} \cdot 10 \approx 167,1$  en  $160 + \frac{13}{17} \cdot 10 \approx 167,6$ . Het eerste kwartiel is dus ongeveer 167.  
Zo is het derde kwartiel ongeveer 177.  
Met het minimum van 150 en het maximum van 200 kun je het boxplot tekenen.
- 6.5 a** De steekproef moet representatief zijn voor de populatie, dus een goed beeld geven van de populatie. Verder moeten de deelnemers aan die steekproef willekeurig worden gekozen.
- b** Bij de éne soort wil je bijvoorbeeld weten hoeveel procent van alle Nederlanders kleurenblind is. Bij de andere soort vergelijk je twee groepen, bijvoorbeeld de levensduur van twee soorten lampen.
- 6.6 a** Allebei 3.
- b** 40 dagen, dus 8 weken.
- c**  $\frac{6 \cdot 1 + 6 \cdot 2 + 8 \cdot 3 + 4 \cdot 4 + 2 \cdot 5 + 3 \cdot 6 + 6 \cdot 7}{40} = 3,2$  leerlingen per dag.
- 6.7 a**  $166 - 125 = 41$
- b** 142 cm.
- c**  $Q_1 = 135$  en  $Q_3 = 151$ . De kwartielafstand is  $151 - 135 = 16$  cm.
- d** Zie de figuur.

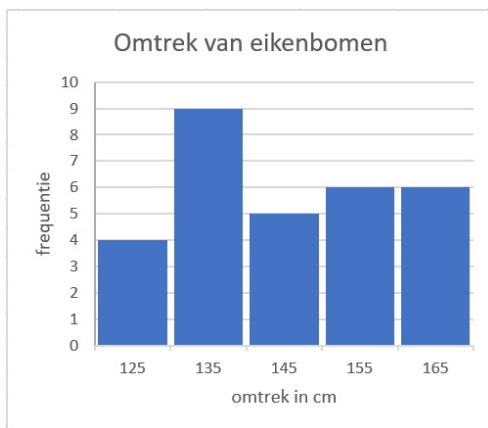




6.8 a Zie de tabel.

klasse	frequentie
120– < 130	4
130– < 140	9
140– < 150	5
150– < 160	6
160– < 170	6
totaal	30

b Zie de figuur.



c  $\frac{12}{30} = 0,4$  dus 40%.

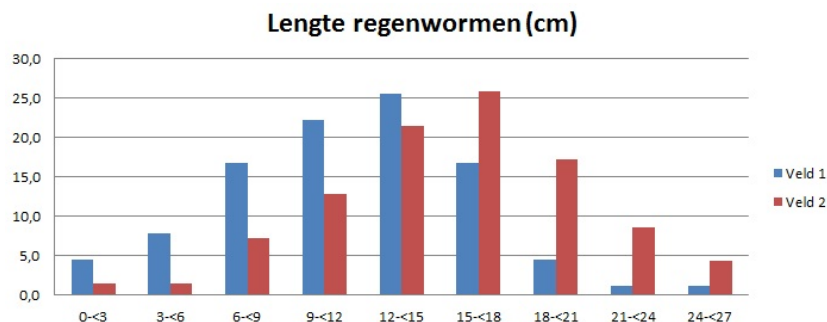
d De gemiddelde omtrek is 144,1 cm (gebruik de ruwe data). Er zijn 12 bomen met een grotere omtrek en dat is 40%.

e Je vindt dan ongeveer 145,3. Dat is hoger dan het werkelijke gemiddelde want de meeste bomen hebben een omtrek die lager is dan het klassenmidden van de klasse waar ze in terecht komen.

6.9 a Omdat  $\frac{1+3}{2} = 1,52$ .

b Veld 1: het gemiddelde is 11,6 cm.  
Veld 2: het gemiddelde is ongeveer 15,5 cm.

c Er zijn verschillende aantallen regenwormen gemeten op deze velden, dus om eerlijk te kunnen vergelijken reken je alles om naar procenten. Zie de figuur.



d Op veld 1 is dat 23,3% en op veld 2 is dat 55,7%.

6.10 a Op veld 1 is dat de klasse 12– < 15.  
Op veld 2 is dat de klasse 15– < 18.

- b** Het gaat nu om alleen maar gehele getallen in elke klasse.  
 Veld 1: minimum is 0, eerste kwartiel is 9, mediaan is 11, derde kwartiel is 14 en maximum is 27.  
 Veld 2: minimum is 0, eerste kwartiel is 12, mediaan is 15, derde kwartiel is 20 en maximum is 27.
- c** De regenwormen op veld 2 zijn langer dan die op veld 1. 75% van de regenwormen op veld 2 is langer dan de mediaan van veld 1.

**6.11 a** Eigen antwoord.

In ieder geval moet je er om denken dat de verdeling over de verschillende schooltypen overeen komt met de werkelijkheid. En verder moet je een goede verdeling over de leeftijdsklassen maken. En tenslotte een goede verdeling over het land. En zo kun je wellicht nog wel wat verzinnen.

- b** Of dit een goede aanpak is, is de vraag. Weliswaar lijkt elke scholier wel een smartphone te hebben en zou hij/zij zo'n vraag via internet kunnen beantwoorden, maar gaat iedereen dat ook werkelijk doen? En is de groep die dat doet wel een goede steekproef? En hoe kun je rekening houden met de zaken die je bij a hebt genoemd?
- c** Eigen antwoord.

**6.12 a** Een leeftijdsdiagram bestaat uit twee staafdiagrammen, één voor de mannen en één voor de vrouwen.

- b** Ongeveer 1.000.000 mannen en ongeveer 950.000 vrouwen.
- c** De staven die bij de leeftijdsgroep 30– < 40 en 40– < 50 horen springen er duidelijk uit.
- d** 30– < 40
- e** Maak eerst een frequentietabel zoals die hieronder met de klassenmiddens er in. Het gemiddeld is dan ongeveer 38,2 jaar.

leeftijd	aantal
5	2200
15	1950
25	2150
35	2650
45	2500
55	2250
65	1350
75	1000
85	600
	16650

**6.13 a** Werk met een frequentietabel zoals die hieronder. De mediaan van de mannen is ongeveer 36 jaar. En die van de vrouwen is ongeveer 38 jaar.

leeftijd	mannen	vrouwen
5	1200	1000
15	1000	950
25	1100	1050
35	1350	1300
45	1250	1200
55	1150	1100
65	650	700
75	450	550
85	200	400
	8350	8250

- b** Mannen: minimum leeftijd 0, eerste kwartiel 18, mediaan 36, derde kwartiel 43, maximum leeftijd 90 (mag ook hoger).  
 Vrouwen: minimum leeftijd 0, eerste kwartiel 21, mediaan 38, derde kwartiel 45, maximum leeftijd 90 (mag ook hoger).

Het verschil is niet erg groot, vrouwen worden gemiddeld wat ouder dan mannen (dat zie je in de bevolkingspiramide) en daarom liggen de leeftijden voor de vrouwen wat hoger dan voor de mannen.

- c Eigen antwoord.

# Leerdoelentabel

In het  achter de opgave kun je aangeven hoe je de opgave hebt gemaakt:

✓ goed gemaakt — **S** wel begrepen maar een slordige fout gemaakt — **H** hulp nodig gehad — **G** samen met groepje goed gemaakt — **X** fout gemaakt en niet goed begrepen — **N** niet bekeken

<b>1</b>	<b>Centrummaten</b>	★	★★	★★★
	Van een hoeveelheid gegevens de modus, de mediaan en het (gewogen) gemiddelde berekenen.	1.3 <input type="checkbox"/> 1.4 <input type="checkbox"/> 1.7 <input type="checkbox"/> 1.8 <input type="checkbox"/> T6.6 <input type="checkbox"/>	1.6 <input type="checkbox"/>	
	Modus, mediaan en gemiddelde interpreteren als centrummaten van die gegevens.	1.2 <input type="checkbox"/> 1.3 <input type="checkbox"/> 1.4 <input type="checkbox"/> 1.5 <input type="checkbox"/> 1.7 <input type="checkbox"/> 1.8 <input type="checkbox"/> T6.6 <input type="checkbox"/>		
<b>2</b>	<b>Spreidingsmaten</b>	★	★★	★★★
	De spreidingsbreedte en de kwartielafstand van een serie gegevens berekenen.	2.1 <input type="checkbox"/> 2.2 <input type="checkbox"/> 2.3 <input type="checkbox"/> 2.4 <input type="checkbox"/> 2.5 <input type="checkbox"/> 2.6 <input type="checkbox"/> 2.7 <input type="checkbox"/> T6.7 <input type="checkbox"/>		
	Spreidingsbreedte en kwartielafstand interpreteren als spreidingsmaten van die gegevens.	2.1 <input type="checkbox"/> 2.2 <input type="checkbox"/> 2.3 <input type="checkbox"/> 2.4 <input type="checkbox"/> 2.5 <input type="checkbox"/> 2.6 <input type="checkbox"/> 2.7 <input type="checkbox"/> T6.7 <input type="checkbox"/>		
	Een serie gegevens samenvatten in een boxplot.	2.1 <input type="checkbox"/> 2.2 <input type="checkbox"/> 2.3 <input type="checkbox"/> 2.5 <input type="checkbox"/> 2.6 <input type="checkbox"/> 2.7 <input type="checkbox"/> T6.7 <input type="checkbox"/>		
<b>3</b>	<b>Klassenindeling</b>	★	★★	★★★
	Een frequentietabel met klassenindeling en geschikte klassenbreedte maken bij een verzameling ruwe data.	3.1 <input type="checkbox"/> 3.2 <input type="checkbox"/> 3.3 <input type="checkbox"/> 3.4 <input type="checkbox"/> 3.5 <input type="checkbox"/> 3.6 <input type="checkbox"/> 3.7 <input type="checkbox"/> 3.8 <input type="checkbox"/> T6.8 <input type="checkbox"/> T6.9 <input type="checkbox"/> T6.10 <input type="checkbox"/>		
	Bij zo'n klassenindeling diagrammen maken.	3.3 <input type="checkbox"/> 3.5 <input type="checkbox"/> 3.7 <input type="checkbox"/> 3.8 <input type="checkbox"/> T6.8 <input type="checkbox"/> T6.9 <input type="checkbox"/> T6.10 <input type="checkbox"/>		
	Bij een klassenindeling het gemiddelde schatten.	3.2 <input type="checkbox"/> 3.3 <input type="checkbox"/> 3.4 <input type="checkbox"/> 3.5 <input type="checkbox"/> 3.6 <input type="checkbox"/> 3.7 <input type="checkbox"/> 3.8 <input type="checkbox"/> T6.8 <input type="checkbox"/>		
<b>4</b>	<b>Schatten</b>	★	★★	★★★
	Bij een klassenindeling de centrummaten en de spreidingsmaten schatten.	4.1 <input type="checkbox"/> 4.2 <input type="checkbox"/> 4.3 <input type="checkbox"/> 4.4 <input type="checkbox"/> 4.5 <input type="checkbox"/> 4.6 <input type="checkbox"/> T6.10 <input type="checkbox"/>		
	Bij een klassenindeling een boxplot maken.	4.1 <input type="checkbox"/> 4.2 <input type="checkbox"/> 4.3 <input type="checkbox"/> 4.4 <input type="checkbox"/> 4.5 <input type="checkbox"/> 4.6 <input type="checkbox"/> T6.10 <input type="checkbox"/>		



5

<b>Statistische uitspraken</b>	★	★★	★★★
Een statistische uitspraak herkennen.	5.1 <input type="checkbox"/> 5.2 <input type="checkbox"/> 5.3 <input type="checkbox"/> 5.4 <input type="checkbox"/> 5.5 <input type="checkbox"/> T6.11 <input type="checkbox"/>		
Bij een statistisch onderzoek de begrippen populatie en steekproef herkennen.	5.1 <input type="checkbox"/> 5.2 <input type="checkbox"/> 5.3 <input type="checkbox"/> 5.4 <input type="checkbox"/> 5.5 <input type="checkbox"/> T6.11 <input type="checkbox"/>		
Aangeven of een steekproef representatief is.	5.3 <input type="checkbox"/> 5.4 <input type="checkbox"/> T6.11 <input type="checkbox"/>		
Een eenvoudig statistisch onderzoek opzetten.	5.3 <input type="checkbox"/> 5.4 <input type="checkbox"/> T6.11 <input type="checkbox"/>		

**Het lesmateriaal in deze reader is gebaseerd op het materiaal dat ook op de Math4All website staat.**

**De reader is gegenereerd met de Math4All maatwerkdienst. De inhoud en de volgorde van de onderwerpen in deze reader zijn gekozen door docenten van het ConText College.**

**Stichting Math4All**



[www.math4all.nl](http://www.math4all.nl)

