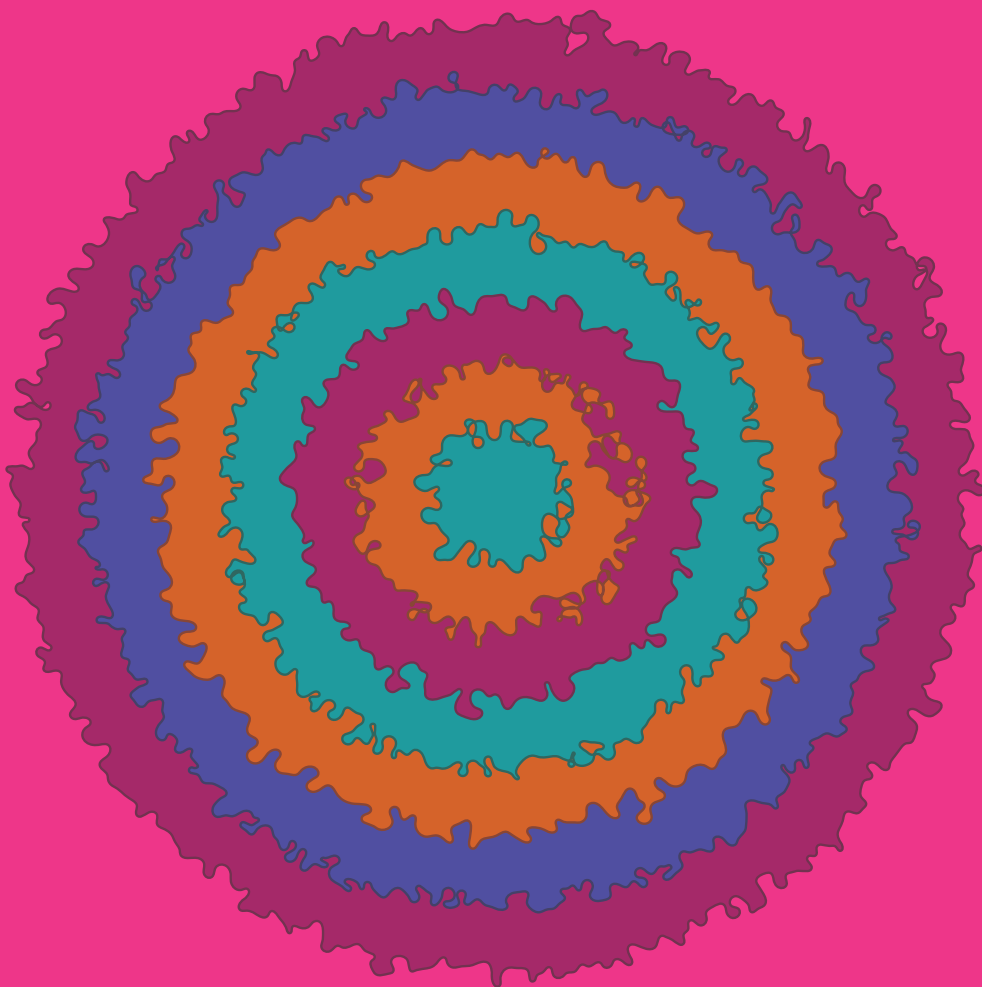


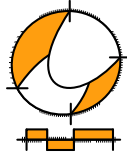
# Wiskunde / PGA

2 HAVO / VWO

## 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



---

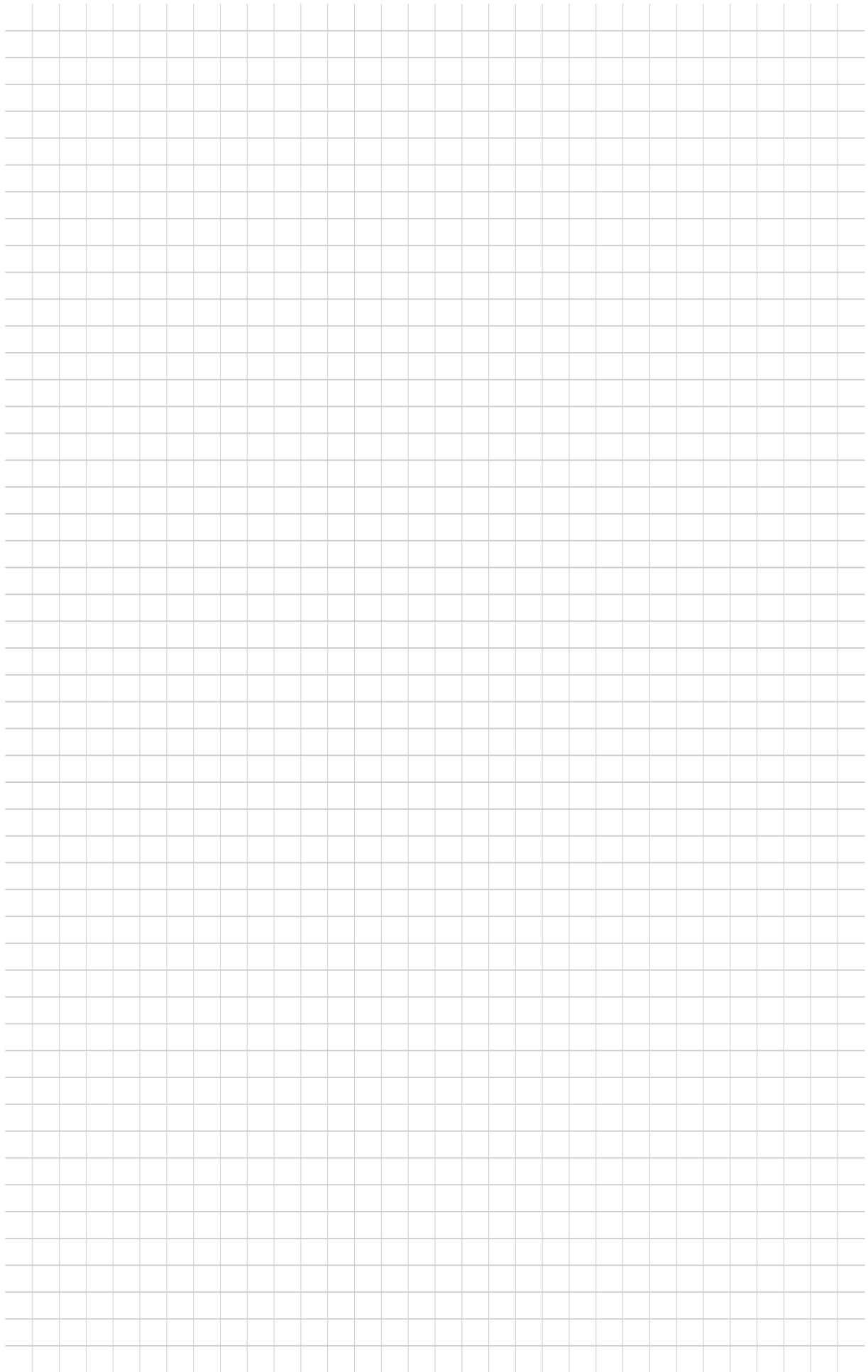
# 1

---

## Statistiek

1.1	Centrummaten	6
1.2	Spreidingsmaten	13
1.3	Klassenindeling	21
1.4	Schattingen	30
1.5	Statistische uitspraken	38
1.6	Totaalbeeld	44





## 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.

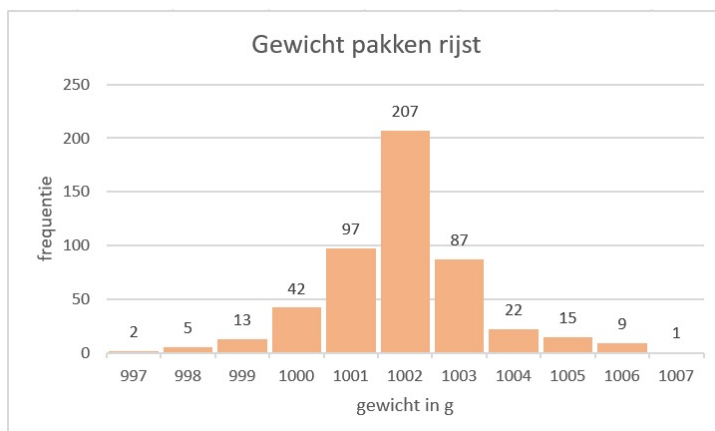
- Erik haalde voor zijn wiskundetoetsen de cijfers 8, 5, 6, 7, 10, 3, 7, 7 en 8.
- De directeur van een onderneming krijgt € 8000 per maand. De vier adjunctdirecteuren ontvangen maandelijks € 3000. De overige dertig werknemers verdienen elk € 1900 per maand.
- Van de docenten van een scholengemeenschap 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

- Bepaal de modus en de mediaan.
- Hoeveel pakken werden er elke werkdag gecontroleerd?

- c Hoeveel procent van de gecontroleerde pakken rijst had een te laag gewicht?
- d Bereken 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 uitbetaald. De mediaan van die prijzen blijkt € 1500 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 en 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.

★★★ **Opgave 1.7**

In klas V2A zitten 20 leerlingen. Voor een toets Engels scoorde de klas gemiddeld 6,6. Ook klas V2B maakte die toets, maar hier was het gemiddelde cijfer 8,1. Het gemiddelde cijfer van de twee klassen samen was 7,5.

- a Leg uit waarom het gemiddelde niet gelijk is aan  $\frac{6,6+8,1}{2} = 7,35$ .
- b Hoeveel leerlingen zitten er in klas V2B?
- c De jongens in klas V2A scoorden gemiddeld 6,4 en de meisjes 6,9. Hoeveel jongens zitten er in klas V2A?

## Toepassen

### ★ ★ Opgave 1.8: Centrummaten met een spreadsheet

Een spreadsheet (rekenbladprogramma) kan je helpen bij het maken van frequentietabellen en het berekenen van gemiddelden. Bekijk het [Practicum](#). Daarin werk je met het een bestand dat is gemaakt voor de spreadsheet waarmee je werkt.

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

- a Maak zelf een frequentietabel van de gehele eindcijfers voor wiskunde van deze groep leerlingen in 4 vwo.
- 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 1.9: 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 in 4 vwo.
- 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: Statistiek met een spreadsheet

Een **spreadsheet of rekenbladprogramma** kan centrummaten voor je berekenen vanuit de data. In de volgende practica leer je werken met statistiek in een spreadsheet. Download wel eerst het bijbehorende bestand "Gegevens 154 leerlingen".

- [Data presenteren met behulp van MS-Excel](#)
- [Data presenteren met behulp van Open Office Calc](#)
- [Data presenteren met behulp van Google spreadsheets](#)

# 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** De mediaan is 6,5.
- d** Nieuwe gemiddelde 6,6, modus is 7 en mediaan is 7.

**1.2 a** Het gemiddelde is  $\approx 6,8$ .

- b** De modus is € 1900.
- c** De modus is het openbaar vervoer.
- d** Het gemiddelde is 23 °C.

**1.3 a** De modus is 1002 gram en de mediaan is 1002 gram.

- b** 25 pakken per werkdag.
- c** 4%.
- d** 1 pak per dag.
- e** 8160 pakken.

**1.4 a** Van 45 doosjes lucifers.

- b** 1599
- c**  $\approx 35,5$  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** Er zijn één of meerdere enorm hoge prijzen uitgekeerd.

**1.6** Al je voor het laatste proefwerk een 8,7 of hoger haalt.

**1.7 a** De aantallen in beide klassen zijn verschillend.

- b** 30 leerlingen.
- c** 12 jongens.

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

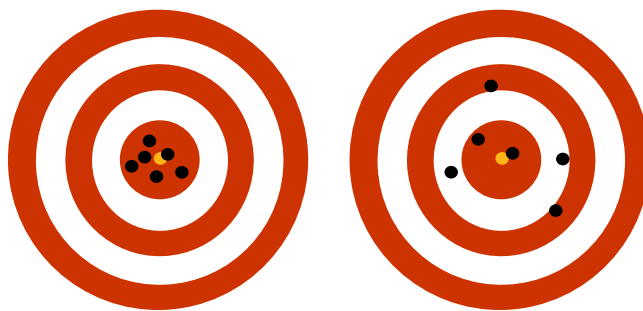
- b** Vergelijk je resultaten met die van de andere leerlingen. Het gemiddelde is ongeveer 7,5 en de mediaan en de modus zijn beide 8.
- c** Vergelijk je resultaten met die van de andere leerlingen.
- d** Vergelijk je resultaten met die van de andere leerlingen.

**1.9 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?

### Inleiding

Ook al hebben twee groepen gegevens dezelfde centrummaat, dan nog kunnen er grote verschillen zijn. De spreiding 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 en interpreteren als spreidingsmaten;
- een serie gegevens samenvatten in een boxplot.

### Voorkennis

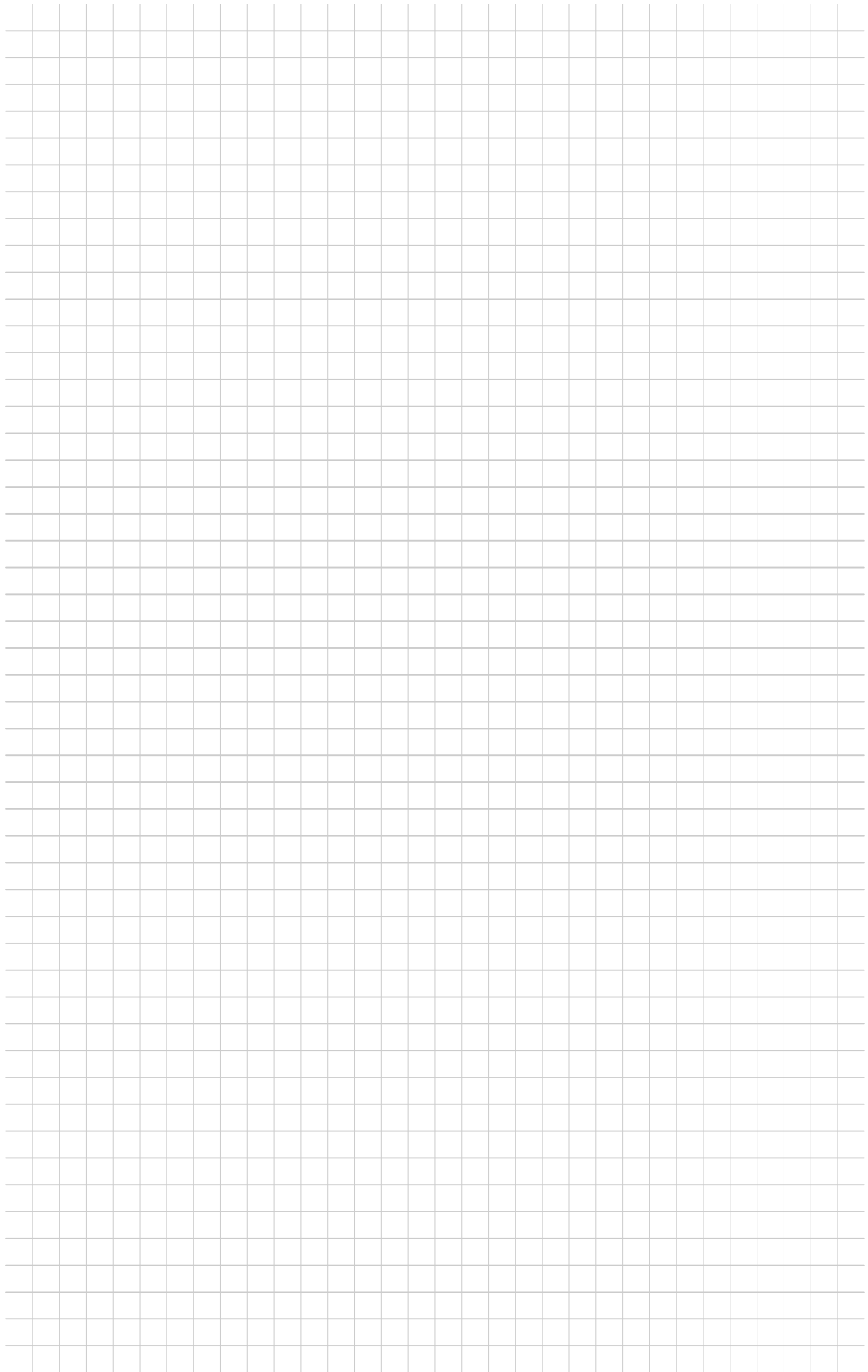
- (gewogen) gemiddelde, modus en mediaan van een hoeveelheid gegevens berekenen;
- werken met (relatieve) frequenties en (relatieve) frequentietabellen;
- werken met kruistabellen.

### Voor de leerling

Je krijgt in groepjes één of meer opdrachten waarmee je de theorie die bij dit onderdeel hoort zelf gaat opbouwen. Het gaat om het leren werken met 'spreidingsbreedte' en '(inter)kwartielafstand' als spreidingsmaat van een serie gegevens, een dataset. Verder leer je die spreiding weergeven in een boxplot en door boxplots te vergelijken uitspraken te doen.

Maak eigen aantekeningen en uiteindelijk voor je zelf een theorie-overzicht.

### Aantekeningen



## Theorie

### Om te onthouden

A large grid of graph paper with a light blue background and a thin grey grid pattern, intended for taking notes on the theory of measures of dispersion.

## Verwerken

### ★ Opgave 2.1

Op 1 januari 2021 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 2.1**

Hetzelfde gebeurde bij een even grote groep meisjes:

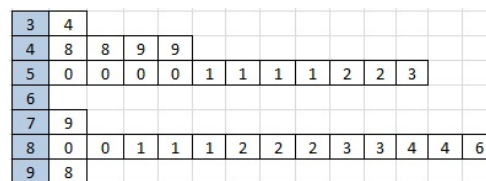
2400	2500	2500	2600	2700	2800	2800	2800	2900	2900
3000	3000	3100	3100	3200	3200	3200	3300	3300	3350
3400	3400	3500	3500	3600	3700	3800	3900	4000	4000

**Tabel 2.2**

- 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 2.2

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



**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 2.3

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 2.3**

- 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 2.4**

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

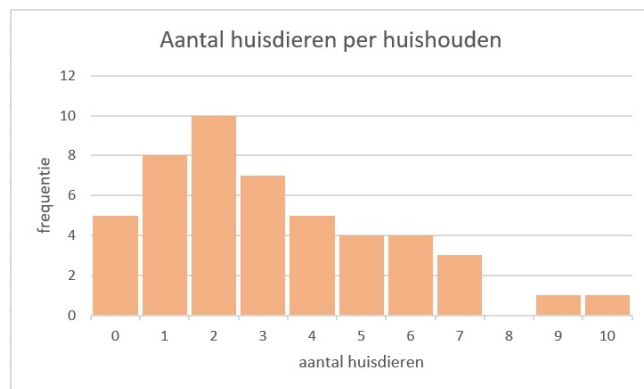
72	50	63	59	63	51	56	64	58	74
52	69	42	51	54	58	51	60	60	59
50	67	61	45	66	49	64	63	57	56

**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.



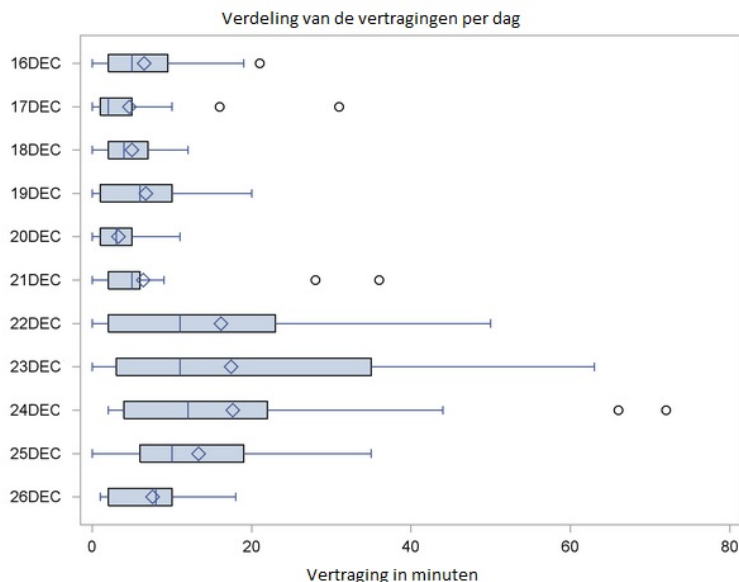
**Figuur 2.3**

- 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?

\*\*\*

## Opgave 2.6

Op een bepaald vliegveld is een aantal opeenvolgende dagen bijgehouden hoeveel vertraging de vluchten van die dag hadden. Elke dag vinden er dezelfde vluchten plaats, in totaal 320 vluchten per dag. De echte uitschieters zijn buiten de boxplots gehouden en aangegeven door rondje. Ook is het gemiddelde aantal minuten vertraging elke dag aangegeven met een symbool in de vorm van een ruit.



**Figuur 2.4**

- Hoeveel uitschieters waren er op 17 december? Veranderen de mediaan en de kwartielen als deze uitschieters wel worden meegerekend?
- Waarom zie je dat bij het berekenen van het gemiddelde de uitschieters wel zijn meegeteld?
- Welke dag kende de minste vertragingen?
- Welke dag was de kleinste vertraging het grootst?
- Op welke dagen was meer dan een kwart van de vluchten meer dan 20 minuten vertraagd?
- Hoe kan het dat sommige dagen de gemiddelde vertraging nogal veel groter is dan de mediaan?

## Toepassen

\*\*

### Opgave 2.7: Boxplots met een spreadsheet

Een spreadsheet kan je helpen bij het berekenen van centrum- en spreidingsmaten en het maken van boxplots. Bekijk het [Practicum](#) en het download het bestand 'Gegevens 154 Leerlingen'.

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.

- Maak zelf een tabel met de kwartielen, de mediaan en maximum en minimum van de gehele eindcijfers voor wiskunde van deze groep leerlingen in 4 vwo.
- 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 2.8: Lengtes en gewichten vergelijken**

Werk met hetzelfde 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.

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

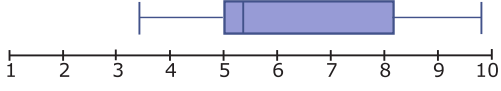
**Practicum: Statistiek met een spreadsheet**

Een **spreadsheet of rekenbladprogramma** kan centrum- en spreidingsmaten voor je berekenen vanuit de data.

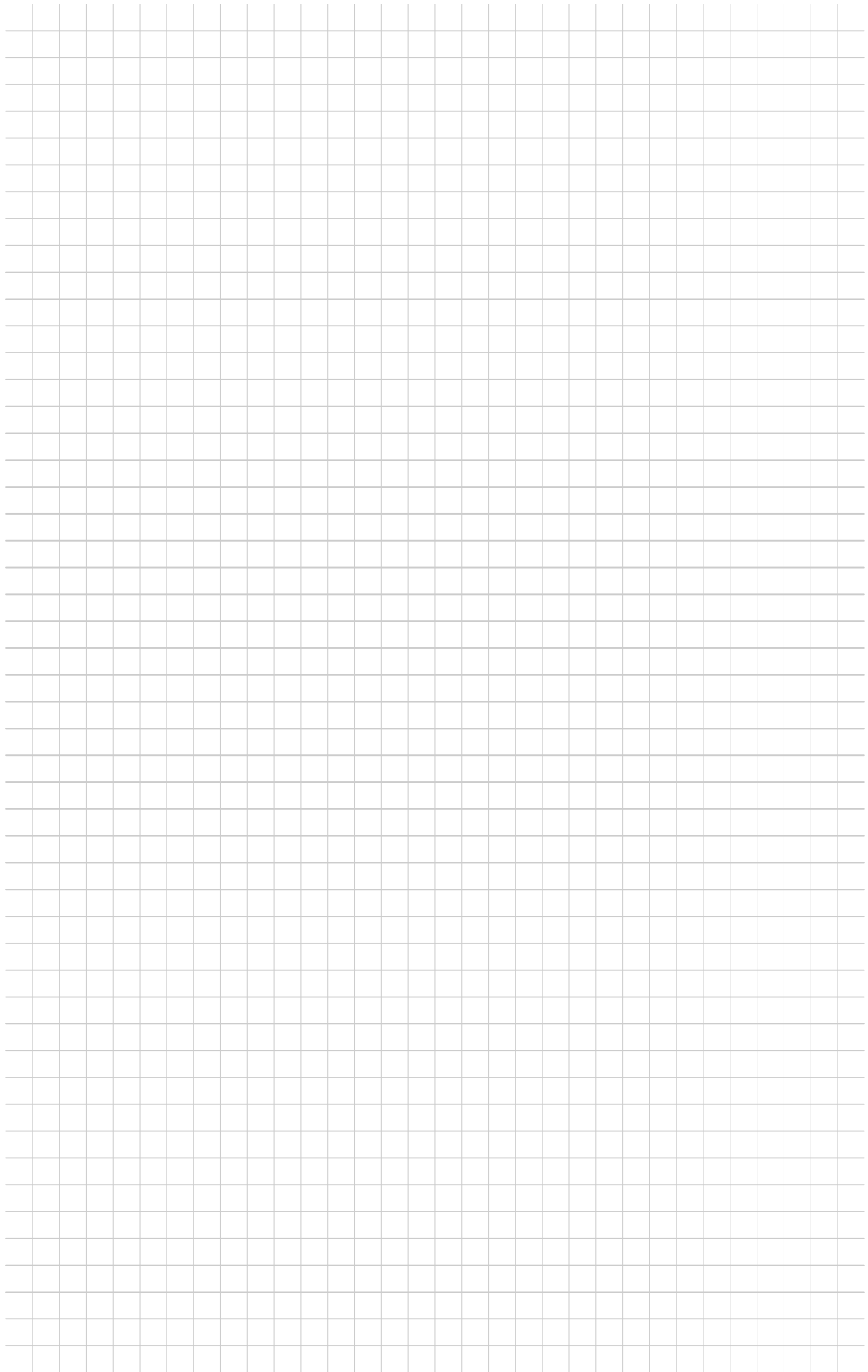
In de volgende practica leer je werken met statistiek in een spreadsheet. Je kunt er ook boxplots in maken. Download wel eerst het bijbehorende bestand 'Gegevens 154 leerlingen'.

- [Data presenteren met behulp van MS-Excel](#)
- [Data presenteren met behulp van Open Office Calc](#)
- [Data presenteren met behulp van Google spreadsheets](#)

## Antwoorden

- 2.1 a** Spreidingsbreedte: jongens 3000, meisjes: 1600.
- b** Jongens: minimum 1850 gram,  $Q_1 = 3150$  gram, mediaan 4000 gram,  $Q_3 = 4180$  gram en maximum 4850 gram.  
Meisjes: minimum 2400 gram,  $Q_1 = 2800$  gram, mediaan 3200 gram,  $Q_3 = 3500$  gram en maximum 4000 gram.
- c** 50%.
- d** 50%.
- 2.2 a** Zie de figuur.
- 
- b** De duidelijke scheiding tussen voldoende en onvoldoende zie je het best in het steelbladdiagram. De boxplot laat de spreiding beter zien.
- 2.3 a** Van beide teams zijn zowel de gemiddelde leeftijd als de mediaan 23.
- b** Team 1: minimum 18,  $Q_1 = 20,5$ , mediaan 23,  $Q_3 = 26$  en maximum 28 jaar.  
Team 2: minimum 21,  $Q_1 = 22$ , mediaan 23,  $Q_3 = 24$  en maximum 26 jaar.
- 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 dichter bij elkaar liggen.
- 2.4 a** Het is handig om eerst een frequentietabel te maken.  
De mediaan is 85,5 en de interkwartielafstand is 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, interkwartielafstand: 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** Twee. Als je deze uitschieters meerekent veranderen de mediaan en de kwartielen niet of nauwelijks.
- b** Het gemiddelde op bijvoorbeeld 17 december ligt dicht bij het derde kwartiel vanwege de twee uitschieters. Datzelfde geldt voor 21 december.
- c** 20 december, toen was er maximaal ongeveer 12 minuten vertraging.
- d** Op 24 december.
- e** Op 22, 23 en 24 december.
- f** Op die dagen zitten de vertragingen relatief dicht bij de maximale vertraging.
- 2.7 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.
- 2.8 a** Vergelijk je resultaten met die van de andere leerlingen.
- b** Vergelijk je resultaten met die van de andere leerlingen.







## Theorie

### Om te onthouden

A large grid of graph paper with a light blue background and a grid of thin grey lines. The grid is intended for students to write down their theory notes.

## 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.

- Je hebt op de huishoudbeurs aan 500 bezoekers hun leeftijd gevraagd. De jongste bezoeker was 17 en de oudste was 81.
- Je hebt aan die zelfde 500 bezoekers van de huishoudbeurs hun lievelingskleur gevraagd. Tussen de antwoorden die de bezoekers gaven zaten 12 verschillende kleuren.
- 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

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 3.1

- Wat was het grootste getal dat hij had kunnen gooien? Leg uit.
- Bereken de gemiddelde score in één decimaal nauwkeurig.
- Maak een klassenindeling van de scores. Neem als klassen  $5- < 8$  en  $8- < 11$ , enzovoorts.
- 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?
- Bereken nu met behulp van de klassenindeling de gemiddelde score.
- In hoeveel procent van de worpen scoorde Johan meer dan 13 maar minder dan 23?
- Waarom zou score 18 zoveel vaker voorkomen dan score 28?

### ★ Opgave 3.3

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

- Welk klassenmidden heeft de klasse  $45- < 50$ ? Licht je antwoord toe.
- Bereken het gemiddelde gewicht in één decimaal nauwkeurig.
- 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 3.2



★ **Opgave 3.4**

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

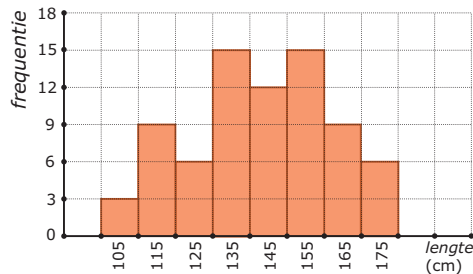
31	29	18	16	24	47	12	32	52	10	26
22	12	53	49	25	21	59	51	32	16	27
18	29	31	37	17	42	31	39	19	33	28
44	16	29	22	17	19	54	27	30	28	51

**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 3.1**

- 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 2021 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.

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

## Toepassen

### ★ ★ Opgave 3.7: Lengtes in klassen in een spreadsheet

Een spreadsheet (rekenbladprogramma) 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 "Gegevens 154 Leerlingen".

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

- a Maak zelf de frequentieverdeling van de lengtes van deze groep leerlingen in 4 vwo.
- b Maak een staafdiagram.
- c 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 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.

- a Maak voor de lengtes van de jongens en de meisjes afzonderlijke frequentietabellen en schat met behulp daarvan hun gemiddelde lengtes. Kun je conclusies trekken?
- b 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: Statistiek met een spreadsheet

Een **spreadsheet of rekenbladprogramma** 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 de volgende practica leer je werken met statistiek in een spreadsheet:

- [Data presenteren met behulp van MS-Excel](#)
- [Data presenteren met behulp van Open Office Calc](#)
- [Data presenteren met behulp van Google spreadsheets](#)

## Antwoorden

- 3.1 a** Ja, een klassenindeling is hier verstandig; eerste klasse bijvoorbeeld  $15- < 25$ .  
**b** Nee, je kunt geen klassenindeling maken bij kleuren, dat gaat alleen bij getallen.  
**c** Ja, een klassenindeling kan wel; eerste klasse bijvoorbeeld:  $1- < 3$ .

- 3.2 a** 30 ogen.  
**b**  $\approx 17,9$  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**  $\approx 18,1$  ogen.  
**f** In 76,7% van de worpen.  
**g** Omdat de score 18 op meer manieren met vijf dobbelstenen te maken is dan de score 28.

- 3.3 a** 47,5  
**b**  $\approx 64,6$  kg.  
**c** Ongeveer 80%.

- 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** De leeftijden 50,51,52,53,54, dus het klassenmidden is 52.  
**c** Ongeveer 28 jaar.  
**d** Ongeveer 56,8% van de bezoekers.
- 3.5 a**  $100- < 110$   
**b** 75 leerlingen.  
**c** 143 cm.

d Je krijgt nu 142 cm.

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 3700 gram en meisjes  $\approx$  3233 gram.

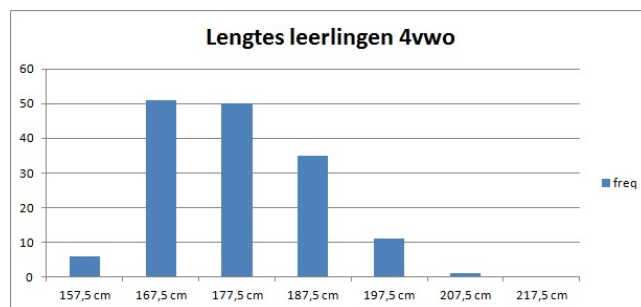
c Ongeveer 67,7%.

d Ongeveer 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 Schattingen

### Inleiding

Je hebt leren werken met klassenindelingen en de bijbehorende frequentietabellen. Voordeel ervan is het krijgen van enig overzicht als je met veel ruwe data te maken hebt. Nadeel is dat je alle centrummaten en spreidingsmaten alleen nog maar kunt schatten.

#### Je leert in dit onderwerp

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

#### Voorkennis

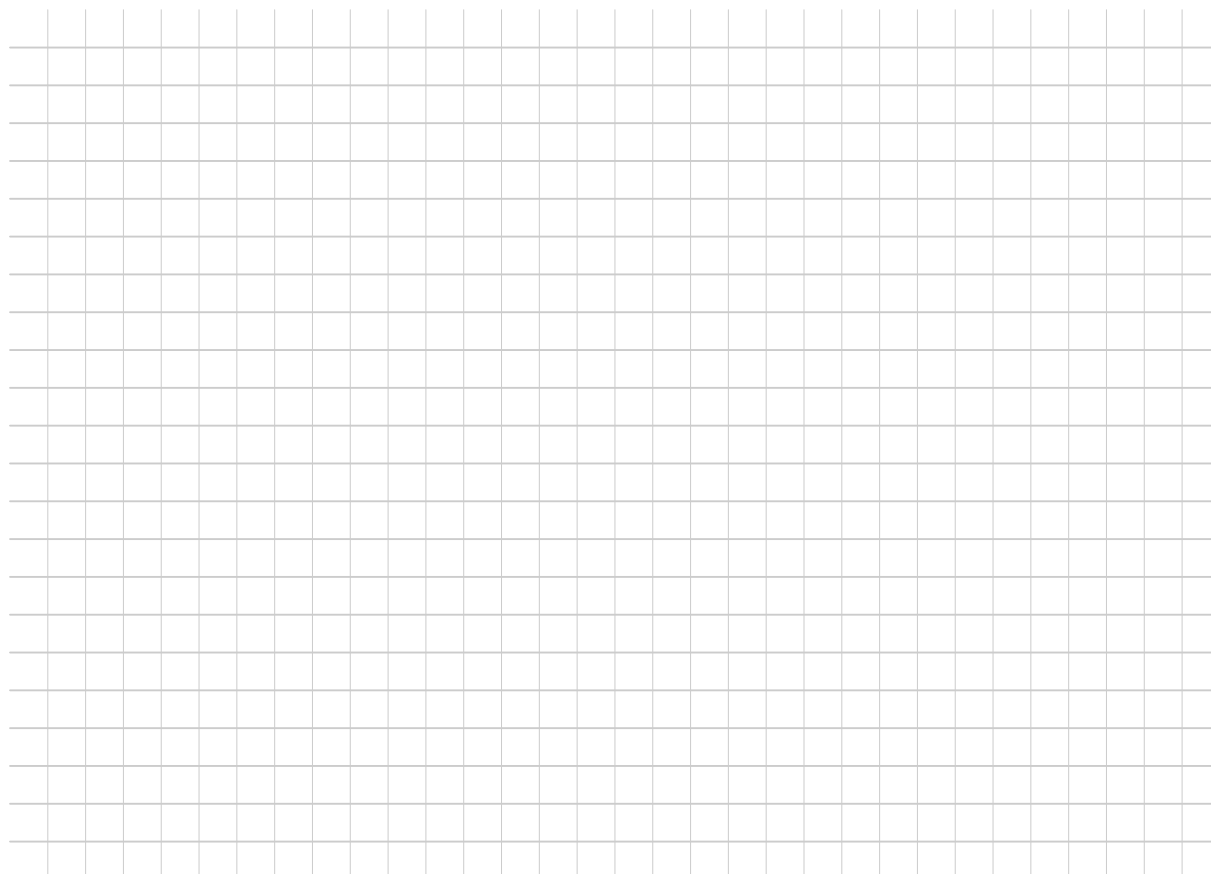
- (gewogen) gemiddelde, modus en mediaan van een hoeveelheid gegevens berekenen;
- een frequentietabel met klassenindeling en geschikte klassenbreedte maken bij een verzameling ruwe data;
- werken met kruistabellen.

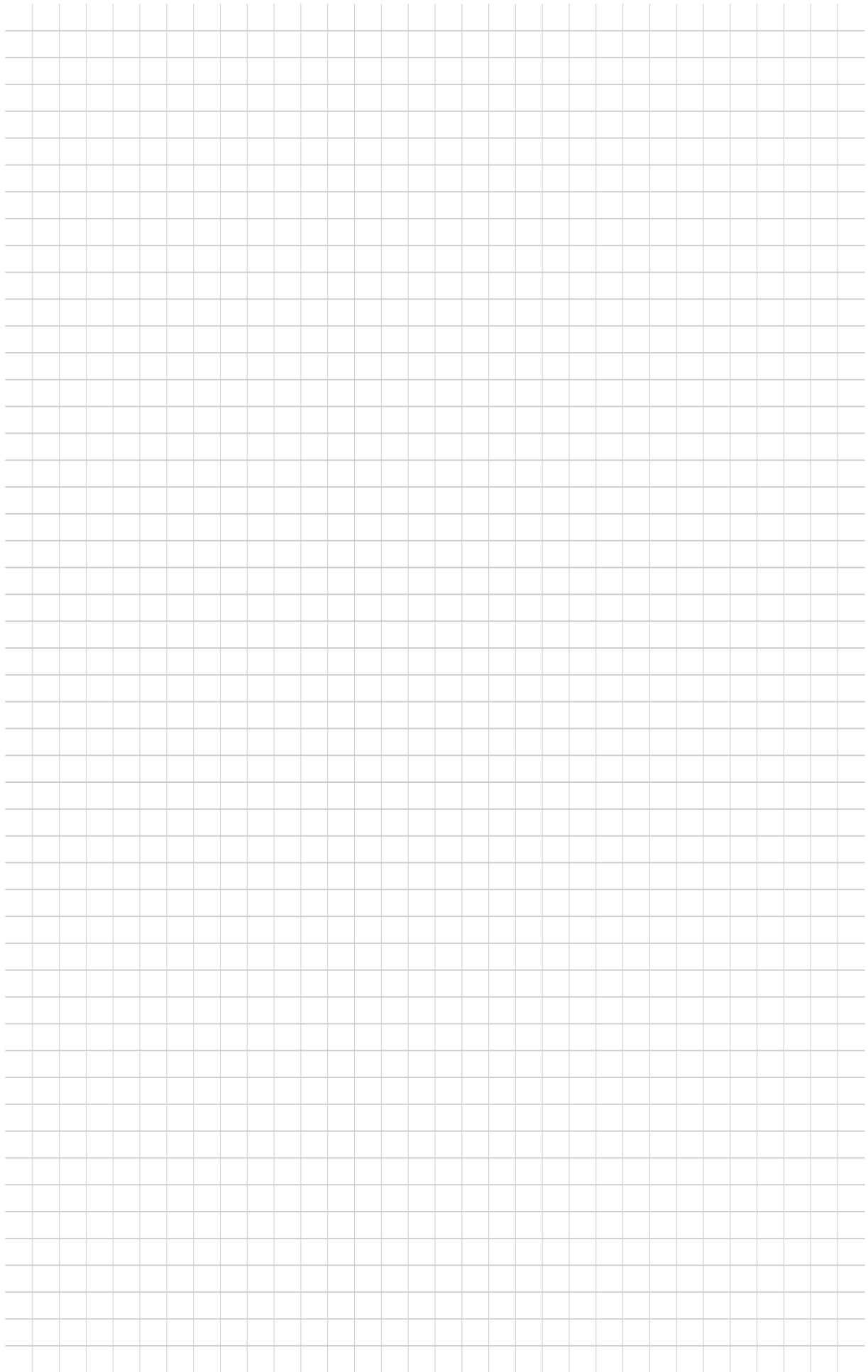
### Voor de leerling

Je krijgt in groepjes één of meer opdrachten waarmee je de theorie die bij dit onderdeel hoort zelf gaat opbouwen. Het gaat om het maken van schattingen van de centrummaten bij een klassenindeling van een dataset en het maken van een bijpassende boxplot.

Maak eigen aantekeningen en uiteindelijk voor je zelf een theorie-overzicht.

### Aantekeningen







## Theorie

### Om te onthouden

A large grid of graph paper with a light blue background and a grid of thin grey lines. The grid is intended for students to write down theory or important points related to the topic of estimation.



## 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 4.1

- 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

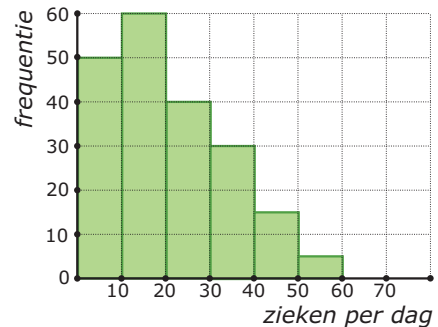
Bedenk in de volgende situaties wat het klassenmidden is van de gegeven klasse.

- De cijfers die klas 2C voor scheikunde heeft gehaald. Klasse 7– < 8.
- Het aantal zieke leerlingen per schooldag. Klasse 10– < 15.
- De lichaamslengtes van een groep kinderen. Klasse 160– < 170.
- Het aantal bezoekers per dag van een winkeltje. Klasse 30– < 45.

### ★ Opgave 4.3

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 4.4**

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.

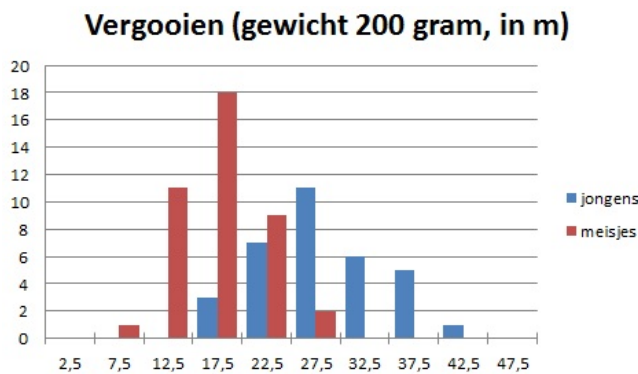
- a Bepaal de modale klasse.
- b Hoe schat je het gemiddelde? Bereken deze schatting.
- c Schat de mediaan.
- d Schat de kwartielen en teken een bijpassend boxplot.

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

Tabel 4.2

★ **Opgave 4.5**

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.



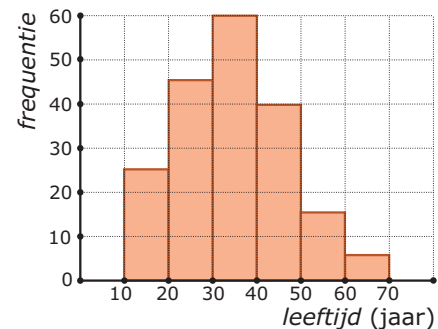
Figuur 4.2

- 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.

★ **Opgave 4.6**

In het staafdiagram zie je de leeftijdsverdeling van een groep reizigers die met bussen op weg is naar de Costa del Sol.

- a Maak bij het staafdiagram een frequentieverdeling. Neem dezelfde klassen als in het staafdiagram.
- b Bepaal de spreidingsbreedte.
- c Schat de gemiddelde leeftijd van de reizigers in gehele jaren.
- d Geef een schatting van de mediaan. Beredeneer je schatting.
- e Geef een schatting van de kwartielen en bereken de kwartielafstand.
- f Hoeveel procent van de buspassagiers was jonger dan de gemiddelde leeftijd?
- g De reisorganisatie geeft korting aan alle reizigers die ouder zijn dan 55 jaar of jonger dan 15 jaar. Geef een schatting van het percentage reizigers dat korting kreeg.
- h Welke klasse is de modale klasse?
- i Tijdens de vakantie zijn twaalf van de reizigers jarig. Verandert hierdoor de gemiddelde leeftijd in gehelen?



Figuur 4.3

## Toepassen

### ★ ★ Opgave 4.7: Uren huiswerk per week

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 data) en je hebt diagrammen getekend. Voor een boxplot gebruik je ook het liefst de ruwe data, 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.

- 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 passend lijndiagram van de jongens en meisjes in één figuur.
- c Probeer een conclusie te trekken.

### ★ ★ Opgave 4.8: Uren huiswerk per week en boxplots

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

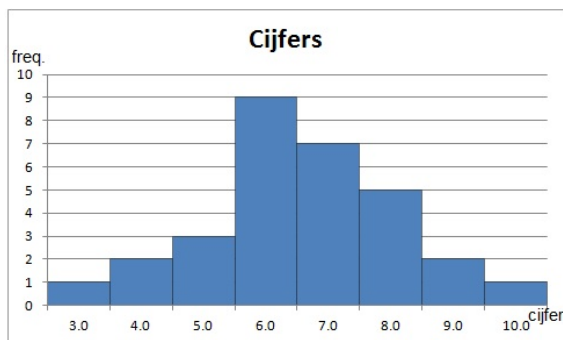
- a Maak nu twee boxplots naast elkaar voor het aantal uren huiswerk dat de jongens per week besteden en het aantal uren huiswerk dat de meisjes per week besteden.
- b Probeer opnieuw conclusies te trekken, betrek ook de gemiddelden erbij.

# Antwoorden

4.1 a 7,0

b Omdat je anders heel veel staafjes van lengte 1 krijgt. En daaruit kun je nauwelijks wat zinnigs aflezen.

c Zie de figuur.



d 5,5– < 6,5

e 6,5

f  $Q_1 \approx 5,7$ .

4.2 a Klassenmidden is 7,5.

b Klassenmidden is 12.

c Klassenmidden = 165 cm.

d Klassenmidden is 37.

4.3 a 200

b  $\approx 18$

c Afgerond op gehelen 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.

4.4 a De modale klasse (meest voorkomende) is: 170– < 180.

b Gebruik de klassenmiddens, het geschatte gemiddelde is:  $\approx 172,5$  cm.

c De geschatte mediaan is 173 cm.

d Minimum 150,  $Q_1 \approx 167$ , mediaan 173,  $Q_3 \approx 177$  en maximum 200.

4.5 a Jongens: minimum: 15,  $Q_1 \approx 24,0$ , mediaan  $\approx 28,0$ ,  $Q_3 \approx 33,5$ , maximum 45.

Meisjes: minimum 5,  $Q_1 \approx 14,0$ , mediaan  $\approx 17,5$ ,  $Q_3 \approx 20,5$ , maximum 30.

b Jongens: 28,5 m; meisjes: 17,5 m.

c Ja, de gemiddelden verschillen nogal.

4.6 a Zie de tabel.

klasse	frequentie
10– < 20	25
20– < 30	45
30– < 40	60
40– < 50	40
50– < 60	15
60– < 70	5
totaal	190

b 60

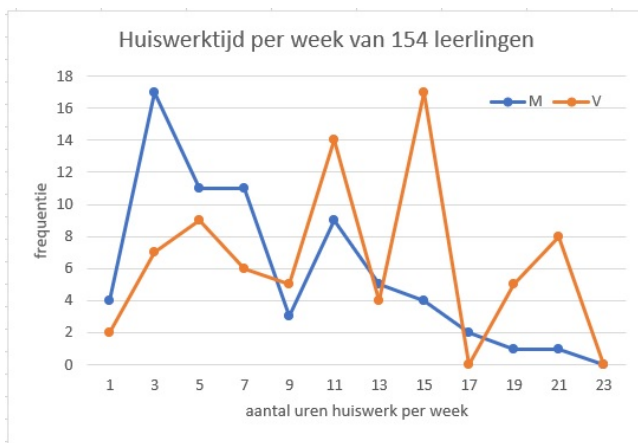
c Afgerond op gehelen is dat 34 jaar.

- d Dus de mediaan is  $\approx 34$  jaar.
- e  $Q_1 \approx 25$  en  $Q_3 \approx 43$ , dus de kwartielafstand is ongeveer 18 jaar.
- f Ongeveer 51%.
- g Ongeveer 13%.
- h De klasse  $30- < 40$ .
- i Als minstens één van de jarigen in een hogere leeftijdsklasse komt zal de gemiddelde leeftijd in gehelen veranderen.

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

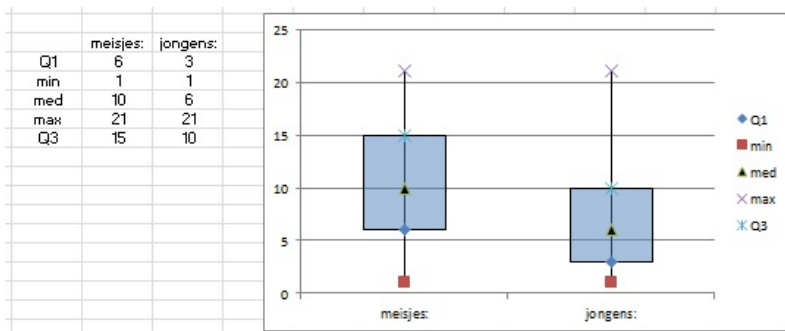
huiswerkuren per week			M	V
laagste	midden	hoogste		
0	1	1,9	4	2
2	3	3,9	17	7
4	5	5,9	11	9
6	7	7,9	11	6
8	9	9,9	3	5
10	11	11,9	9	14
12	13	13,9	5	4
14	15	15,9	4	17
16	17	17,9	2	0
18	19	19,9	1	5
20	21	21,9	1	8
22	23	23,9	0	0
			68	77

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.

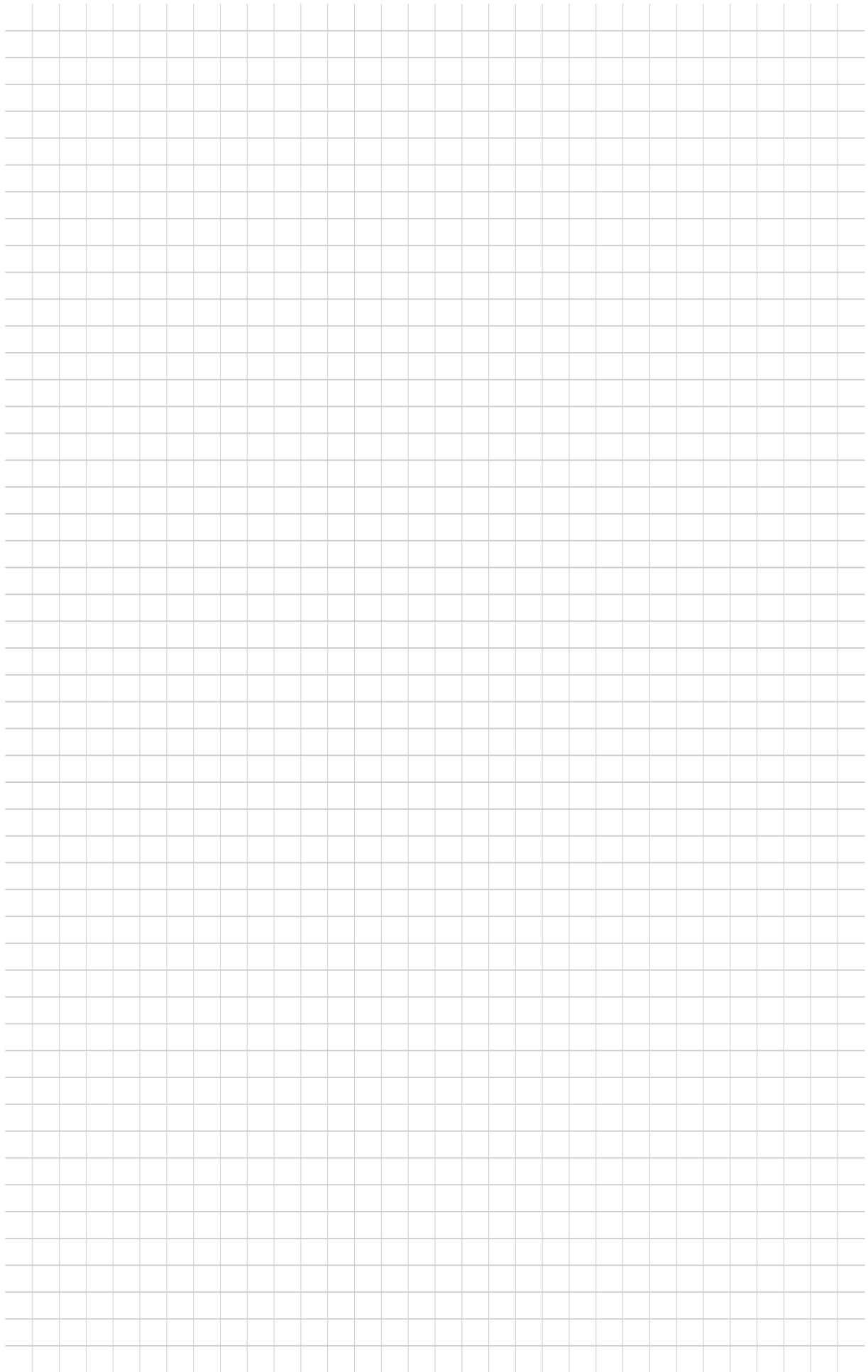
4.8 a Hier zie je de boxplots met de gegevens die Excel berekent.



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...







## 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.

- Waar moet een goede steekproef aan voldoen?
- 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.

- “De boeren in Europa die kippeneieren produceren zitten in een crisis.”
- “Er zijn veel mensen in Nederland van plan om komend jaar hun huis te gaan verbouwen.”
- “Nederlandse jongens zijn beter in atletiek dan Nederlandse meisjes.”
- “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 vwo diploma behaald hebben en daarna een hbo studie gaan doen in plaats van naar de universiteit te gaan. Je hebt al wat rondgevraagd en krijgt het vermoeden dat vooral veel meisjes deze keuze maken, en dat veel minder jongens dat doen.

Je hebt inmiddels met 25 studenten gesproken die een hbo opleiding zijn gestart nadat ze hun vwo diploma gehaald hebben. 15 van deze studenten zijn vrouw en 10 zijn man. Je schrijft als kop boven je artikel: “Meisjes met een vwo diploma kiezen veel vaker voor een hbo opleiding dan jongens met een vwo diploma.”

- Is dit een statistische uitspraak? Leg uit waarom wel of niet.
- Via de website van het [Centraal Bureau voor de Statistiek](#) kun je aan meer informatie komen voor je artikel. Bekijk de statistieken over instroom, doorstroom, en uitstroom in het onderwijs.

Je kunt kiezen bij Geslacht voor het totaal aantal, het aantal mannen of het aantal vrouwen.

Je kunt kiezen bij Periode voor meerdere jaren.

Je kunt kiezen bij Onderwijspositie basisjaar voor 6 vwo met/zonder diploma.

Hoe zou je deze gegevens gebruiken voor je artikel?

## Toepassen

### ★★ Opgave 5.4: Een eigen onderzoeksvraag bedenken

Bekijk het [Practicum](#) en het bestand ‘Gegevens 154 Leerlingen’. Dit bestand bevat informatie over een groep 4 vwo leerlingen. Deze gegevens ga je gebruiken om een statistische onderzoeksvraag te beantwoorden die jij zelf bedenkt.

De enige voorwaarde is dat het een onderzoek moet zijn waarbij je ‘op zoek gaat naar het percentage van een bepaalde populatie dat een zekere eigenschap heeft.’

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

Hieronder kun je zien welke stappen je moet zetten om een goed resultaat te krijgen.

- a Bedenk een eigen statistische onderzoeksvraag.
- b Stel een plan van aanpak op. Daarin beschrijf je hoe de 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?

**Practicum: Statistiek met een spreadsheet**

Een **spreadsheet of rekenbladprogramma** 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 de volgende practica leer je werken met statistiek in een spreadsheet:

- [Data presenteren met behulp van MS-Excel](#)
- [Data presenteren met behulp van Open Office Calc](#)
- [Data presenteren met behulp van Google spreadsheets](#)

## 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** Je kunt voor mannen en vrouwen afzonderlijk nagaan hoeveel leerlingen er van het vwo naar het hbo gaan.  
Een voorzichtige conclusie die je kunt trekken is dat er inderdaad meer vrouwen van vwo naar hbo gaan dan mannen. Je zult nog wel na moeten gaan of dit niet kan komen doordat er meer vrouwen een vwo diploma halen dan mannen.  
Wat je ook voorzichtig kunt concluderen is dat er door de jaren heen een daling te zien is (die de laatste jaren stagneert) in het aantal mannen en vrouwen dat van vwo naar hbo gaat. Maar ook hierbij zul je eerst moeten checken of het aantal mensen dat een vwo diploma haalt niet net zo hard mee daalt.
- 5.4 a** Eigen antwoord.  
**b** Eigen antwoord.  
**c** Vraag aan je docent of je de juiste conclusie(s) hebt getrokken.
- 5.5 a** Laat je onderzoeksvraag door je docent goedkeuren.  
**b** Laat je plan van aanpak door je docent goedkeuren.  
**c** Doen.  
**d** Doen.  
**e** Laat de resultaten van je onderzoek en je conclusie aan je docent zien.  
**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 schoenmaten van 124 volwassen mannen.

aantal eieren	39	40	41	42	43	44
frequentie	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 mediaan en de kwartielafstand.
- 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 **Opgave 6.3**. 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

Conciërge Martin heeft elke schooldag van drie brugklassen het aantal telaarcomers 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 gemeten. In de frequentietabel zie je de resultaten. Je gaat de gegevens van beide velden vergelijken.

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**

- 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?

★ **Opgave 6.10**

Gebruik de gegevens over de regenwormen uit de vorige opgave.

- 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.

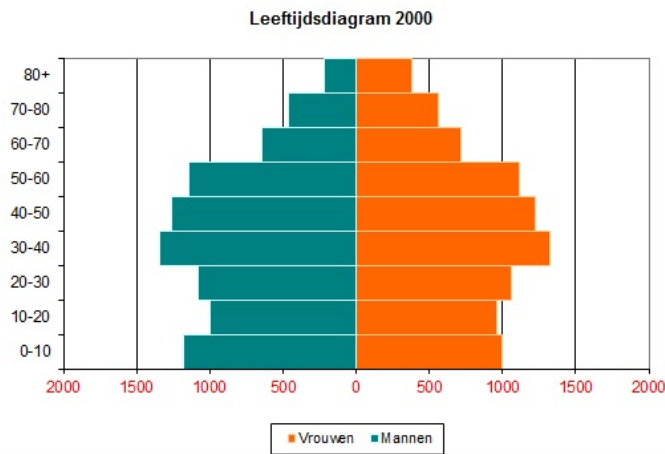
★ **Opgave 6.11**

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

Hier zie je een zogenaamde **bevolkingspiramide**, een leeftijdsdiagram. Daarin 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 0-10-jarigen in 2000 uit ongeveer 1.200.000 jongens en ongeveer 1.000.000 meisjes bestond.

- Om wat voor soort diagrammen gaat het hier?
- Hoeveel mannen en vrouwen van 10– < 20 waren er in Nederland in 2000?
- 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?
- Wat was in 2000 de modale leeftijdscategorie?
- Schat de gemiddelde leeftijd van de Nederlanders in 2000.
- Het woord ‘bevolkingspiramide’ is wat verouderd. Waarom is dat woord niet meer zo van toepassing als vroeger, denk je?

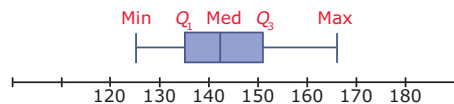
### ★★ Opgave 6.13: Bevolkingspiramide vervolg

Bekijk de bevolkingspiramide van de voorgaande opgave nog eens.

- Schat de mediaan van de leeftijden van de Nederlandse mannen in 2000. Doe hetzelfde voor de vrouwen.
- 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.
- 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  $\frac{5608}{136} \approx 41,2$  eieren per dag.
- 6.2 a** De spreidingsbreedte is:  $44 - 39 = 5$ .
- b** Voor de mediaan bekijk je het 62e en het 63e getal. Die zijn beide 42, dus de mediaan is 42.  
De kwartielen zijn:  $Q_1 = 41$  (het 31e en het 32e getal zijn beide 41) en  $Q_3 = 43$  (het 92e en het 93e getal zijn beide 43). De kwartielaafstand is:  $43 - 39 = 4$ .
- c** Doen. 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 en de frequenties te gebruiken.  
Als je aanneemt dat in de klasse  $150- < 160$  alle getallen van 150,00... tot en met 159,99... zitten, is het klassenmidden 155. Dan geldt iets dergelijks ook voor de andere klassen.  
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** 3,2 leerlingen per dag.
- 6.7 a** 41
- b** 142 cm.
- c**  $Q_1 = 135$  en  $Q_3 = 151$ . De kwartielaafstand is  $151 - 135 = 16$  cm.
- d** Zie de figuur.

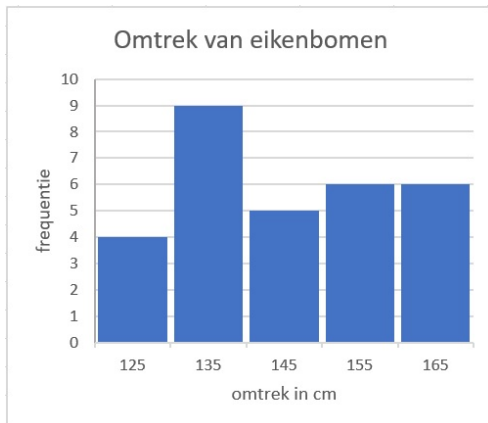




6.8 a Zie de tabel.

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

b Zie de figuur.



c 40%.

d 40%.

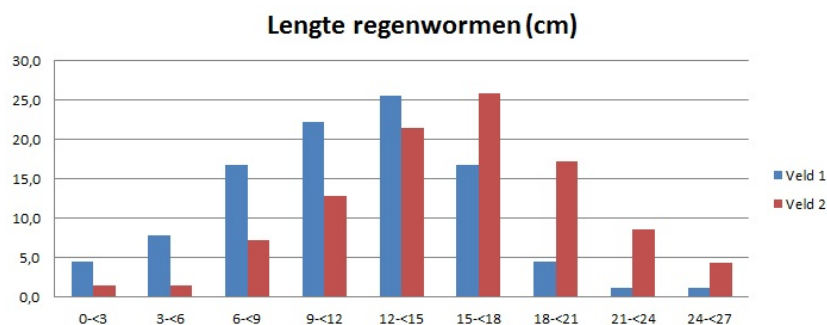
e Je vindt dan ongeveer 145,3. Dat is hoger dan het werkelijke gemiddelde want de meeste bomen hebben een omtrek die lager is dan het klassenmidden van de klasse waar ze in terecht komen.

6.9 a Omdat in die eerste klasse alle getallen vanaf 0 tot aan 3 kunnen voorkomen.

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 Veld 1 12– < 15 en veld 2 15– < 18.

b Veld 1: minimum is 0, eerste kwartiel is 8,4, mediaan is 11,9, derde kwartiel is 14,9 en maximum is 27.

Veld 2: minimum is 0, eerste kwartiel is 12, mediaan is 15,5, derde kwartiel is 19 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. Het gemiddelde is dan ongeveer 38,2 jaar.
- f** Zo'n 100 jaar geleden zag zo'n leeftijdsdiagram er nog wel als een piramide uit. Er werden toen meer kinderen per koppel man/vrouw geboren en elke later leeftijdsgroep vielen er meer mensen af. Ouder worden dan zo'n 50 jaar was toen zeer uitzonderlijk.
- 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 gemiddelde berekenen.	1.1 <input type="checkbox"/> 1.2 <input type="checkbox"/> 1.3 <input type="checkbox"/> 1.4 <input type="checkbox"/> T6.6 <input type="checkbox"/>	1.6 <input type="checkbox"/> 1.8 <input type="checkbox"/> 1.9 <input type="checkbox"/> T6.12 <input type="checkbox"/> T6.13 <input type="checkbox"/>	1.7 <input type="checkbox"/>
	Modus, mediaan en gemiddelde interpreteren als centrummaten van die gegevens.	1.1 <input type="checkbox"/> 1.2 <input type="checkbox"/> 1.3 <input type="checkbox"/> 1.4 <input type="checkbox"/> 1.5 <input type="checkbox"/> T6.6 <input type="checkbox"/>	1.6 <input type="checkbox"/> 1.8 <input type="checkbox"/> 1.9 <input type="checkbox"/> T6.12 <input type="checkbox"/> T6.13 <input type="checkbox"/>	1.7 <input type="checkbox"/>
<b>2</b>	<b>Spreidingsmaten</b>	★	★★	★★★
	De spreidingsbreedte en de kwartielafstand van een serie gegevens berekenen en interpreteren als spreidingsmaten.	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"/> T6.7 <input type="checkbox"/>	2.7 <input type="checkbox"/> 2.8 <input type="checkbox"/>	2.6 <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"/> T6.7 <input type="checkbox"/>	2.7 <input type="checkbox"/> 2.8 <input type="checkbox"/>	2.6 <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"/> T6.8 <input type="checkbox"/>	3.7 <input type="checkbox"/> 3.8 <input type="checkbox"/> T6.13 <input type="checkbox"/>	
	Bij zo'n klassenindeling diagrammen maken.	3.5 <input type="checkbox"/> T6.8 <input type="checkbox"/>	3.7 <input type="checkbox"/> T6.12 <input type="checkbox"/> T6.13 <input type="checkbox"/> T6.13 <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"/> T6.8 <input type="checkbox"/>	3.8 <input type="checkbox"/> T6.12 <input type="checkbox"/> T6.13 <input type="checkbox"/>	
<b>4</b>	<b>Schattingen</b>	★	★★	★★★
	Bij een klassenindeling centrummaten en spreidingsmaten schatten.	4.1 <input type="checkbox"/> 4.2 <input type="checkbox"/> 4.3 <input type="checkbox"/> 4.4 <input type="checkbox"/> 4.6 <input type="checkbox"/> T6.9 <input type="checkbox"/> T6.10 <input type="checkbox"/>	4.7 <input type="checkbox"/> 4.8 <input type="checkbox"/>	
	Bij een klassenindeling een boxplot maken.	4.1 <input type="checkbox"/> 4.4 <input type="checkbox"/> 4.5 <input type="checkbox"/> T6.10 <input type="checkbox"/>	4.7 <input type="checkbox"/> 4.8 <input type="checkbox"/>	
<b>5</b>	<b>Statistische uitspraken</b>	★	★★	★★★
	Een statistische uitspraak herkennen.	5.3 <input type="checkbox"/>	5.4 <input type="checkbox"/> 5.5 <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"/> T6.11 <input type="checkbox"/>	5.5 <input type="checkbox"/>	
	Aangeven of een steekproef representatief is.	5.1 <input type="checkbox"/> 5.2 <input type="checkbox"/> 5.3 <input type="checkbox"/> T6.11 <input type="checkbox"/>	5.5 <input type="checkbox"/>	
	Een eenvoudig statistisch onderzoek opzetten.	5.3 <input type="checkbox"/>	5.4 <input type="checkbox"/> 5.5 <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)

