1.3 Resultaten vergelijken

Inleiding

Grote verzamelingen statistische gegevens zijn (ondanks het gebruik van tabellen en diagrammen) vaak nogal onoverzichtelijk. Vergelijken van frequentieverdelingen is ook niet altijd in één oogopslag mogelijk. Daarom worden de resultaten van een statistisch onderzoek vaak samengevat in een aantal getallen die snel informatie geven over frequentieverdelingen, zoals gemiddelden en spreiding van gegevens.

Je leert in dit onderwerp

- een reeks waarnemingen samen te vatten met centrummaten: gemiddelde, mediaan en modus;
- een reeks waarnemingen samen te vatten met spreidingsmaten: variantie, spreidingsbreedte, kwartielafstand en standaardafwijking;
- een reeks waarnemingen samen te vatten in een boxplot.

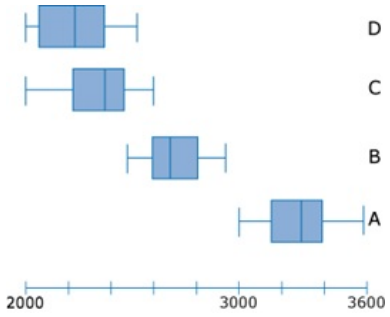
Voorkennis

- gegevens in een frequentietabel (met zowel absolute als relatieve als cumulatieve frequenties) verwerken;
- een (cumulatief) histogram en een (cumulatieve) frequentiepolygoon maken bij een frequentietabel.

Verkennen

Opgave V1

Bij lampen wordt onderzoek gedaan naar de branduren. Bekijk de boxplots van branduren van lampen. Die laten het resultaat van steekproeven van vier typen lampen zien. Elke figuur verdeelt voor een bepaald type lamp de waargenomen branduren in vier gelijke delen, die kwartielen heten. Het eind van het tweede kwartiel is de mediaan.



Figuur 1

a Vergelijk de branduren van type A met de andere types. Gebruik daarbij de figuren, maar ook bijvoorbeeld de mediaan en de kwartielen.

b Zijn er nog andere manieren om de branduren van verschillende lampen te vergelijken?

Uitleg

Twintig soldaten geven hun kledingmaat door aan het magazijnbeheer, opdat ze de juiste maat overhemden krijgen. Zo'n verdeling van de kledingmaten kun je samenvatten met behulp van centrummaten en spreidingsmaten.

37	38	38	39	39
39	40	40	40	41
41	42	42	42	42
42	43	43	43	44

Tabel 1

De volgende centrummaten kun je berekenen:

- De 'gemiddelde' kledingmaat $\overline{x}$ van de groep is $\overline{x} = 40{,}75$.
- De 'modus' is 42, dat is de maat die het meest voorkomt.
- De 'mediaan' is 41, dat is hier het gemiddelde van de twee maten die in het midden van de rij maten op volgorde staan.

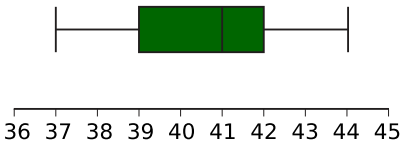
De volgende spreidingsmaten kun je bepalen:

- De 'spreidingsbreedte' is 7, dat is het verschil tussen de grootste en kleinste maat.
- De 'interkwartielafstand' is 3: dat is het verschil tussen het eerste en derde kwartiel.
  - Het 'eerste kwartiel' (ook wel afgekort als Q_1) is 39:
    dat is de mediaan van de eerste helft van de getallen.

– Het ‘derde kwartiel’ (Q_3) is 42:
dat is de mediaan van de tweede helft van de getallen.

Bestaan de waarnemingen uit een oneven aantal waarden, dan wordt de mediaan van de hele set niet meegenomen om Q_1 en Q_3 te berekenen.

De dataset kan overzichtelijk worden weergegeven in een boxplot.



Figuur 2

Dergelijke centrummaten en spreidingsmaten zijn natuurlijk vooral zinvol bij grote datasets. Dan laat je ze berekenen door een spreadsheetprogramma zoals Excel. Zie het **Practicum**.

Opgave 1

Bekijk de **Uitleg**. In de tabel zie je de cijfers van een wiskundetoets van twee parallelklassen.

cijfers klas A						**cijfers klas B**				
6,7	6,4	4,9	3,8	4,0		4,0	6,2	4,9	3,9	5,9
5,6	5,8	6,8	8,2	4,7		7,3	4,7	6,7	7,6	9,4
3,4	8,5	4,1	6,9	7,3		8,3	5,7	7,2	8,7	7,1
6,1	7,5	6,7	6,2	3,4		7,0	6,5	7,4	5,0	4,8
7,9	4,5	8,3				7,7	6,5	4,9	8,8	6,3

Tabel 2

a Waarom heeft het geen zin om van beide klassen het modale cijfer te vergelijken?

b Bepaal van beide klassen de mediaan.

c Zegt de mediaan iets over welke klas beter heeft gescoord?

d Bereken van beide klassen het gemiddelde cijfer.

e Welke van beide klassen heeft het hoogste gemiddelde? Kun je nu zonder meer zeggen dat die klas ook beter heeft gescoord?

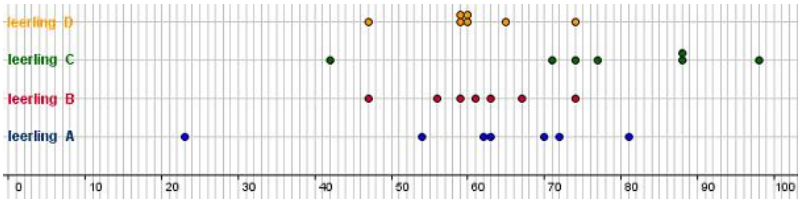
Opgave 2

Je ziet de SE-cijfers (schoolexamen) van enkele leerlingen aan het eind van havo 5. Hun eindcijfer SE is het gemiddelde van deze cijfers.

	se1	se2	se3	se4	se5	se6	se7
A	7,2	6,3	7,0	2,3	6,2	8,1	5,4
B	5,9	7,4	5,6	6,7	6,1	6,3	4,7
C	8,8	9,8	7,4	8,8	5,7	7,3	4,2
D	7,5	6,0	6,0	6,5	6,1	6,1	4,7

Figuur 3

Elk SE-cijfer telt even zwaar mee. In de figuur is voor elke leerling elk SE-cijfer aangegeven door een bolletje op een getallenlijn (de komma in het cijfer is weggelaten).



Figuur 4

a De leerlingen A en B hebben hetzelfde gemiddelde. Toch is hun cijferbeeld nogal verschillend. Hoe komt dat?

b De spreiding van de cijfers van leerling A en leerling C is vrijwel hetzelfde. Waarin verschilt hun cijferbeeld vooral?

c De cijfers van de leerlingen B en D hebben dezelfde spreidingsbreedte. Is de spreiding van hun cijfers ook hetzelfde?

Een andere maat voor de spreiding vind je door te kijken hoe ver elk cijfer van het gemiddelde af ligt. Bereken van elk cijfer het verschil met het gemiddelde. Je ziet die verschillen voor leerling A.

	se1	se2	se3	se4	se5	se6	se7
A	7,2	6,3	7,0	2,3	6,2	8,1	5,4
	1,1	0,2	0,9	-3,8	0,1	2,0	-0,7

Figuur 5

d Bereken het gemiddelde van deze verschillen. Verbaast het antwoord je? Licht je antwoord toe.

Het gemiddelde van deze verschillen is geen goede spreidingsmaat. Dat zit hem in de mintekens. Door te kwadrateren vallen die mintekens weg. Maak voor leerling A een lijst van de kwadraten van de verschillen.

e Bereken daarvan het gemiddelde. Heb je nu een goede spreidingsmaat?

Door het kwadrateren wordt het getal dat je bij e hebt gevonden, nogal groot. Dat los je op door de wortel uit dit getal te nemen. Je krijgt dan de standaardafwijking σ ('sigma' is de Griekse letter 's') van de set cijfers.

f Ga na dat de standaardafwijking voor leerling A ongeveer 1,73 is.

g Bereken ook voor leerling B de verschillen van de cijfers met het gemiddelde. Bereken vervolgens het gemiddelde van de kwadraten van die verschillen en de standaardafwijking.

Opgave 3

Jos heeft zes toetsen voor wiskunde gemaakt.
Zijn cijfers zijn: 5,4; 8,2; 7,3; 7,9; 5,8 en 5,8.

a Hoe groot is de spreidingsbreedte van deze cijfers?

b Bereken de standaardafwijking van deze cijfers. Rond af op twee decimalen.

c Bereken de standaardafwijking van deze cijfers met je grafische rekenmachine. Rond af op twee decimalen.

Theorie en voorbeelden

Om te onthouden

Neem een dataset $x_1, x_2, \ldots, x_n$ en veronderstel dat de waarden op volgorde van klein naar groot staan. Dan zijn er een aantal bekende **centrummaten**.

- De **modus** is de waarneming met de hoogste frequentie. Deze is vooral geschikt voor kwalitatieve variabelen.
- De **mediaan** is het middelste waarnemingsgetal. Is het aantal even, dan is de mediaan het gemiddelde van de middelste twee.
- Het **gemiddelde** bereken je door alle waarnemingsgetallen x_i bij elkaar op te tellen en te delen door het totale aantal. Dit schrijf je als:

  $\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n}$ met $\sum_{i=1}^{n} x_i = x_1 + x_2 + \ldots + x_n$

  of als $\overline{x} = \frac{\sum_{i=1}^{n} x_i \cdot f_i}{n}$ als je met de frequentie f_i van elk waarnemingsgetal rekening houdt.

  De Griekse hoofdletter Σ ('sigma') is het somteken.

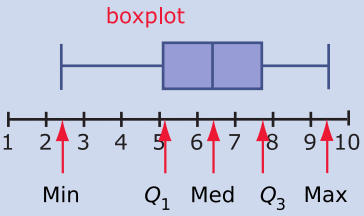
Centrummaten alleen zeggen nog weinig, er hoort steeds een spreidingsmaat bij. Er zijn drie **spreidingsmaten**.

- De **spreidingsbreedte** is het verschil tussen het hoogste en laagste waarnemingsgetal.
- De **interkwartielafstand** is het verschil tussen het eerste en derde kwartiel. Hierbij is het eerste kwartiel de mediaan van de eerste helft van de waarnemingsgetallen en het derde kwartiel die van de tweede helft. Bestaan de waarnemingen uit een oneven aantal waarden, dan wordt de mediaan van de hele set niet meegenomen om Q_1 en Q_3 te berekenen.

- De **standaardafwijking** (of **standaarddeviatie**) is een maat die aangeeft hoe dicht de spreiding rondom het gemiddelde zit. De Griekse (kleine) letter σ ('sigma') is het teken voor standaardafwijking. De standaardafwijking vind je met:

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2}{n}} \text{ of met } \sigma = \sqrt{\frac{\sum_{i=1}^{n}(x_i-\overline{x})^2 \cdot f_i}{n}}.$$

Met de spreidingsbreedte, mediaan en interkwartielafstand kun je een boxplot maken waarmee je gemakkelijk kunt overzien hoe de waarnemingsgetallen in vier even grote groepen zijn verdeeld.



Figuur 6

Voorbeeld 1

Bekijk het steelbladdiagram van toetscijfers in een klas.
Bepaal de modus, het gemiddelde en teken de boxplot bij dit steelbladdiagram. Geef ook de spreidingsbreedte en de interkwartielafstand.

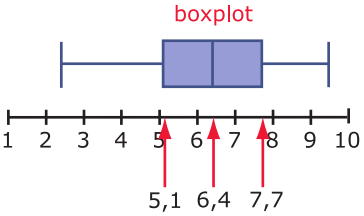
2	4
3	9 9
4	4 4
5	0 0 1 5 5 9 9
6	2 4 4 4 6 6 8 9
7	1 3 7
8	2 5 6 6 6 6
9	5

Figuur 7

Antwoord

- De modus is 8,6 en de modale klasse is $6{,}0 - < 7{,}0$.
- De mediaan is 6,4 (het gemiddelde van de twee middelste cijfers), het eerste kwartiel is 5,1 en het derde kwartiel is 7,7.
- De spreidingsbreedte is $9{,}5 - 2{,}4 = 7{,}1$ als je naar de werkelijke cijfers kijkt, of $10{,}0 - 2{,}0 = 8{,}0$ als je naar de klassenindeling kijkt.
- De interkwartielafstand is $7{,}7 - 5{,}1 = 2{,}6$.
  Het is de breedte van de box van de boxplot.

De boxplot verdeelt alle cijfers in vier delen met elk 25% van deze cijfers.



Figuur 8

Opgave 4

Bekijk het steelbladdiagram in **Voorbeeld 1**.
Stel dat er nog een cijfer bij komt van een leerling die de toets nog moest inhalen. Hij heeft een 8,0 gehaald.

a Wat is nu de mediaan?

b Omdat er nu 31 cijfers zijn, kun je de cijfers niet in twee even grote helften verdelen. Voor het bepalen van de kwartielen laat je daarom de mediaan buiten beschouwing. Je krijgt dan weer twee even grote helften.
Bepaal het eerste en derde kwartiel.

c Als de docent bepaalt om iedereen er 0,5 punten bij te geven, wat gebeurt er dan met de boxplot?

Opgave 5

Een bedrijf heeft 25 werknemers in vaste dienst met een volledige werkweek. De netto weeklonen van deze werknemers zijn in de frequentietabel in klassen verwerkt. De weeklonen zijn verdeeld in klassen met een breedte van 50. De ruwe data zijn niet bekend.

netto weekloon (in €)	aantal werknemers
400– < 450	2
450– < 500	3
500– < 550	4
550– < 600	8
600– < 650	3
650– < 700	2
700– < 750	2
750– < 800	1

Tabel 3

a Waarom kun je vanuit de frequentietabel het gemiddelde netto weekloon niet precies uitrekenen, maar alleen schatten?

b Bepaal de klassenmiddens en bereken hiermee dit geschatte gemiddelde.

c Welke klasse is de modale klasse?

d In welke klasse zit de mediaan?

e Hoe groot is de spreidingsbreedte?

f Stel er komt tijdelijk een interim-directeur in dienst met een weekloon van € 1200,00. Welke centrummaten zouden dan veranderen door deze uitschieter?

Voorbeeld 2

Bekijk de tabel met **gegevens van 36 mannen**.

gegevens van 36 mannen											
nr	*leeftijd* (j)	*lengte* (cm)	*gewicht* (kg)	*nr*	*leeftijd* (j)	*lengte* (cm)	*gewicht* (kg)	*nr*	*leeftijd* (j)	*lengte* (cm)	*gewicht* (kg)
i	*L*	*l*	*G*	*i*	*L*	*l*	*G*	*i*	*L*	*l*	*G*
1	20	180	95	13	46	180	95	25	67	171	68
2	23	184,5	91	14	47	177,5	75	26	67	170	81
3	23	180,5	90	15	51	196	86	27	68	169	97
4	29	169	91	16	52	188,5	93	28	68	176	72
5	30	176	80	17	53	170	85	29	71	171	84
6	33	182	82	18	54	178,5	77	30	71	180,5	75
7	36	182,5	85	19	56	172,5	74	31	71	172	77
8	36	176,5	74	20	61	164	85	32	74	171	77
9	38	180	73	21	61	181	89	33	75	183	95
10	40	190,5	112	22	61	173	95	34	75	172,5	74
11	41	184	89	23	61	181	89	35	82	173	66
12	44	179	75	24	63	181,5	100	36	84	175	90

Tabel 4

Je wilt van alle drie de series waarnemingsgetallen zowel de drie centrummaten als de drie spreidingsmaten bepalen.

Gebruik je Excel, dan gebruik je de statistische functies om de centrummaten en de spreidingsmaten te (laten) berekenen.

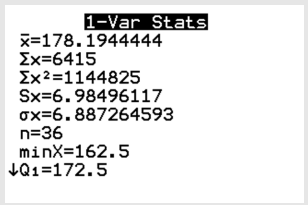
Werk je met de hand of met de grafische rekenmachine, dan maak je meteen geschikte frequentieverdelingen. Je kunt dan de centrummaten en de spreidingsmaten alleen nog schatten vanuit de klassenmiddens. Zie het **Practicum**.

Antwoord

Voor de modus en de mediaan zijn de sorteerfuncties in Excel erg handig. Verder kun je gemakkelijk optellen en kolommen maken met de waarden van een waarnemingsgetal maal zijn frequentie, enzovoort.

Werk je met de grafische rekenmachine, maak dan eerst een tabel van relatieve frequenties met bijbehorende klassenmiddens. Laat vervolgens de machine de centrum- en spreidingsmaten bepalen. Zo krijg je voor de lengte deze gegevens.

Bij de klassenindelingen zijn centrum- en de spreidingsmaten niet exact bepaald, omdat je door indelen in klassen de echte gegevens kwijt bent geraakt. Je krijgt zo schattingen van deze waarden.



Figuur 9

Opgave 6

Bekijk in **Voorbeeld 2** de tabel met gegevens van 36 mannen en in het antwoord de tabel met centrummaten en spreidingsmaten.

a Hoe wordt het gemiddelde berekend?

b Hoe wordt de spreidingsbreedte (variatiebreedte) berekend?

c Hoe wordt de kwartielafstand berekend?

d Ga na dat de modale leeftijd, de modale lengte en het modale gewicht correct zijn.

e Bereken de kwartielen en teken de bijpassende boxplots.

f Hoe wordt de standaardafwijking berekend?

Opgave 7

Bekijk het bestand **gegevens van 36 vrouwen**.

gegevens van 36 vrouwen											
nr	*leeftijd* (j)	*lengte* (cm)	*gewicht* (kg)	nr	*leeftijd* (j)	*lengte* (cm)	*gewicht* (kg)	nr	*leeftijd* (j)	*lengte* (cm)	*gewicht* (kg)
i	*L*	*l*	*G*	*i*	*L*	*l*	*G*	*i*	*L*	*l*	*G*
1	21	181	85	13	53	168,5	64	25	70	166	65
2	22	171	55	14	53	165	64	26	71	159	93
3	30	176	55	15	53	179	77	27	73	154	76
4	32	176,5	70	16	57	174	76	28	73	153	54
5	34	158,5	61	17	62	179	80	29	74	160	60
6	43	160	80	18	64	163,5	80	30	75	163,5	79
7	49	179	75	19	65	160	92	31	75	153	93
8	50	166	74	20	65	182	78	32	76	159	81
9	50	171	84	21	65	154	74	33	77	165	68
10	50	166,5	65	22	67	162	64	34	78	163	89
11	51	171	80	23	68	160	63	35	80	166,5	60
12	52	171	71	24	69	163	75,5	36	80	157	72

Tabel 5

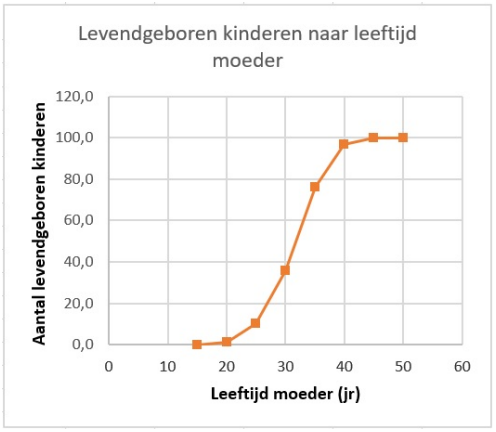
a Bepaal vanuit de basisgegevens de centrummaten en de kwartielen van de leeftijden, de lengtes en de gewichten van vrouwen.

b Bepaal ook de bijbehorende spreidingsmaten en teken de boxplots.

Voorbeeld 3

Bekijk de cumulatieve relatieve frequentiepolygoon bij de klassenindeling van levendgeborenen naar leeftijd van de moeder.

Maak een bijpassende boxplot.



Figuur 10

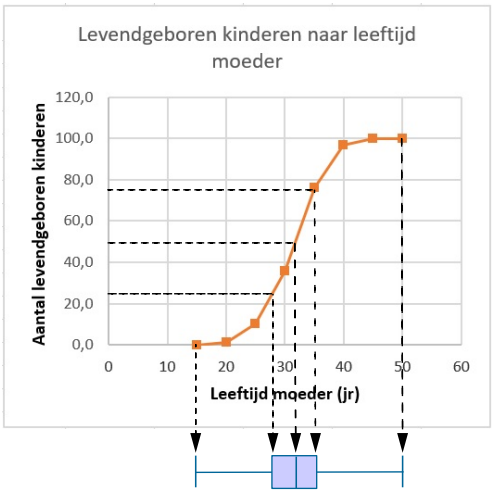
Antwoord

Lees bij 50% de mediaan af, bij 25% het eerste kwartiel en bij 75% het derde kwartiel. Het minimum en maximum zitten bij 0% en 100%.

Dit geeft:

- het eerste kwartiel $Q_1 = 28$
- de mediaan $Q_2 = 31{,}5$
- het derde kwartiel $Q_3 = 35$

Voor de duidelijkheid kun je beter de stippellijnen laten staan.



Figuur 11

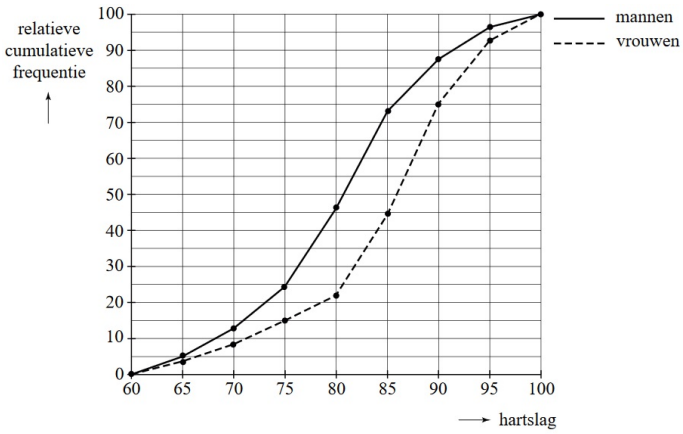
Opgave 8

Bestudeer in **Voorbeeld 3** hoe je bij een klassenindeling een boxplot maakt.

Bekijk vervolgens de tabel met gegevens van 36 mannen. Maak bij de klassenindeling van de kolom ‘lengte’ een cumulatieve relatieve frequentiepolygoon en een boxplot.

Opgave 9

Bepaal vanuit de cumulatieve frequentiepolygonen de vijf getallen die nodig zijn om de frequentieverdeling in een boxplot samen te vatten.



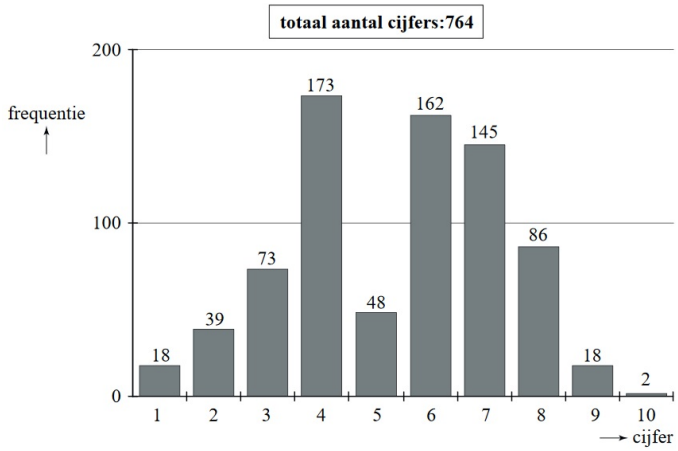
Figuur 12

(bron: examen vwo wiskunde A in 2008, tweede tijdvak)

Verwerken

Opgave 10

In januari 2008 verscheen in de 'NRC' een artikel over de becijfering van een tentamen Recht. In de figuur is de verdeling van de cijfers voor dat tentamen weergegeven.



Figuur 13

Uit de gegevens volgt dat het gemiddelde van de tentamencijfers 5,4 was en de standaardafwijking 1,9.

Bereken zelf het gemiddelde en de standaardafwijking. Rond je antwoord af op twee decimalen.

(bron: examen vwo wiskunde A in 2012, eerste tijdvak)

Opgave 11

Je hebt de waarnemingsgetallen 16,18,22,24,26,26,28,30 en 36.

a Teken een boxplot.

b Doe dat nog eens als je bij alle getallen 4 optelt.

c En ook als je van alle getallen 40 aftrekt.

d Doe het nog eens als je alle getallen door 2 deelt.

e Welk resultaat krijg je als je alle getallen met 3 vermenigvuldigt?

f Beschrijf wat er met de boxplot gebeurt als bij alle waarnemingsgetallen een getal wordt opgeteld of ervan afgetrokken wordt.

g Beschrijf wat er met de boxplot gebeurt als alle waarnemingsgetallen met een getal worden vermenigvuldigd of door een getal worden gedeeld.

Opgave 12

Voor een practicum biologie worden regenwormen gevangen. De lengte van die regenwormen vind je in de tabel.

lengte regenworm (cm)	*aantal*
0,0– < 3,0	4
3,0– < 6,0	8
6,0– < 9,0	17
9,0– < 12,0	22
12,0– < 15,0	23
15,0– < 18,0	17
18,0– < 21,0	6
21,0– < 24,0	2
24,0– < 27,0	1

Tabel 6

a Kijk naar de manier waarop de klassen zijn gemaakt. Hoe nauwkeurig zijn de regenwormen gemeten? Bij welke klasse hoort een regenworm die 3,0 cm lang is?

b Welke klasse is de modale klasse?

c Teken een histogram van de cumulatieve relatieve frequenties. Teken in dezelfde figuur de cumulatieve frequentiepolygoon.

d In welke klasse zit de mediaan? Kun je precies zeggen hoe groot die mediaan is? Schat de mediaan met behulp van de cumulatieve frequentiepolygoon.

e Bereken het gemiddelde en de standaardafwijking.

Opgave 13

De cijfers voor een toets wiskunde van dertig leerlingen worden voor alle leerlingen met 0,5 verhoogd. De cijfers hebben in de oude situatie een gemiddelde van 5,3 en een standaardafwijking van 1,2.

Bereken het gemiddelde en de standaardafwijking van de cijfers wiskunde na de verhoging.

Opgave 14

Een supermarkt laat onderzoek verrichten naar de besteding per klant en naar de hoeveelheid tijd die een klant aan de kassa nodig heeft om af te rekenen. Er worden op verschillende tijdstippen tellingen gehouden. Je ziet de resultaten.

kassatijd (min)	*aantal klanten*
0– < 1	22
1– < 2	75
2– < 3	58
3– < 4	25
4– < 5	16
5– < 6	4

kassabon (€)	*aantal klanten*
0– < 50	24
50– < 100	61
100– < 150	75
150– < 200	25
200– < 250	12
250– < 300	2
300– < 350	1

Tabel 7

a Bepaal bij beide tabellen de modus, de mediaan, het eerste en het derde kwartiel en het gemiddelde.

b Hoe groot is de standaardafwijking bij beide verdelingen?

c Teken bij beide tabellen een boxplot.

De supermarkt heeft een weekomzet van € 150000,00. Een caissière mag 38 uur per week werken.

d Hoeveel caissières moet de supermarkt in dienst nemen als er vanwege de wisselende winkeldrukte een overcapaciteit van 25% wordt aangehouden?

Opgave 15

Op elk uur van een dag is de temperatuur bepaald. De uren van middernacht tot 12 uur 's middags worden aangegeven met am (het Latijnse 'ante meridiem' (am) betekent 'voor het middaguur'), en de uren van 12 uur 's middags tot middernacht met pm ('post meridiem' (pm)).

uur	1	2	3	4	5	6	7	8	9	10	11	12
am	16,1	15,8	15,4	15,1	14,8	15,0	16,1	17,4	18,5	19,4	20,3	21,1
pm	21,9	22,6	22,7	22,5	21,9	21,0	19,8	18,8	18,1	17,5	17,0	16,6

Tabel 8

a Verwerk deze gegevens in een dubbel steelbladdiagram.

b Maak boxplots van elk dagdeel afzonderlijk en van de totale dag.

c Bereken voor beide dagdelen afzonderlijk het gemiddelde en de standaardafwijking.

d Bereken het gemiddelde en de standaardafwijking van alle metingen van die dag.

e Geef een verklaring voor de verschillen die je vindt. (Dit is bivariate statistiek. Je bekijkt twee variabelen en hun eventuele samenhang/verschil.)

Toepassen

In 1947 zijn bij 5001 vrouwelijke klanten van de Bijenkorf vijftien lichaamsmaten opgemeten. Vervolgens is gekeken welke van deze maten het meest bruikbaar zijn om een **maatsysteem voor kleding** op te baseren. Bekijk een deel van de uitkomst van het onderzoek in het bestand **Statistiek Bijenkorf 1947**.

Je kunt met Excel centrum en spreiding van de gegevens berekenen en boxplots maken.

Soms is er sprake van een **uitschieter**, een waarde die wel erg veel afwijkt.

In een boxplot is een uitschieter een waarde die meer dan 1,5 keer de interkwartielafstand onder het eerste kwartiel of boven het derde kwartiel zit.

Opgave 16: Lengte en gewicht

Bekijk bij **Toepassen** de gegevens over lengte en gewicht van de 5001 gemeten vrouwen.

a Bepaal de modale lengte en bereken de gemiddelde lengte.

b Bepaal de mediaan en de kwartielen van de lengte en teken een bijpassende boxplot.

c Zijn er bij deze lengtes uitschieters? Licht je antwoord toe.

d Verdeel de lengtes in klassen van 5 cm, te beginnen bij $135-<140$. Maak bij de nieuwe frequentieverdeling een relatieve cumulatieve frequentiepolygoon. Welke vorm heeft de polygoon?

e Bepaal opnieuw de gemiddelde lengte, de mediaan en de kwartielen. Wijken de resultaten veel af van de antwoorden bij a en b?

Opgave 17: Mouwlengte en kniehoogte

In de dataset in **Toepassen** vind je ook gegevens over de *mouwlengte* en de *kniehoogte* van de 5001 vrouwen.

a Bepaal modus, mediaan en het gemiddelde van de mouwlengtes m_i.

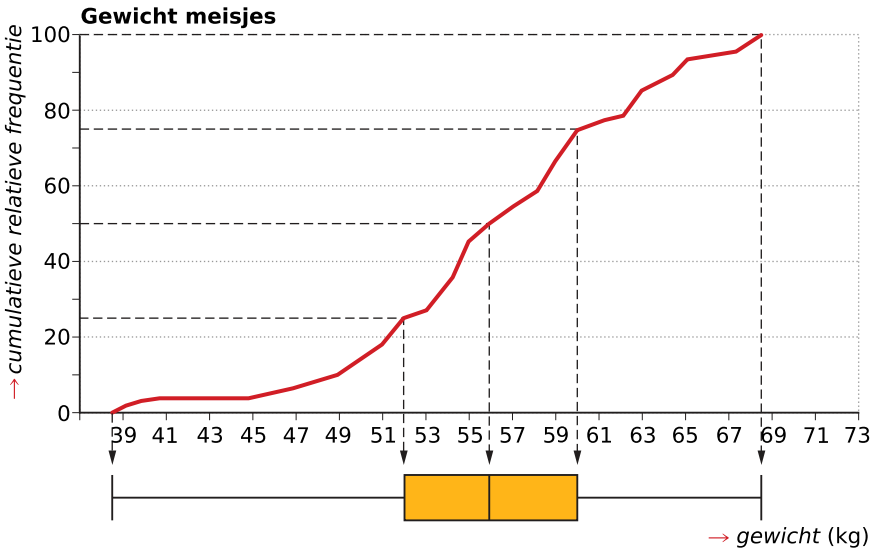
b Bepaal de spreidingsbreedte, de interkwartielafstand en de standaardafwijking van de mouwlengtes m_i.

c Er is één nogal afwijkende mouwlengte. Is hier sprake van een uitschieter?

Opgave 18: Gewichten van 84 meisjes en 68 jongens

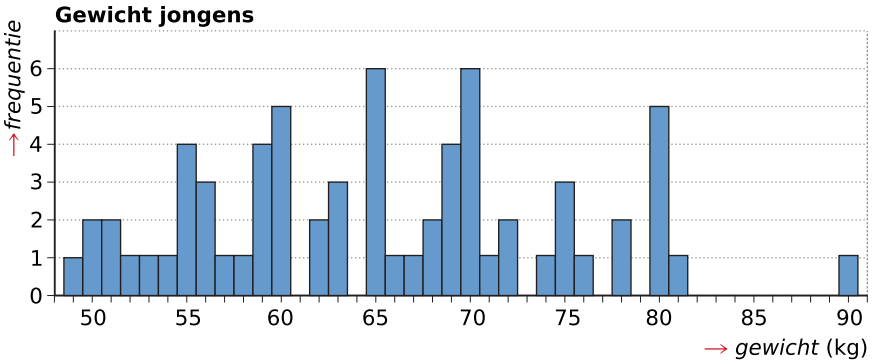
Bekijk de relatieve cumulatieve frequentiepolygoon met de gewichten van een groep van 84 meisjes met de bijbehorende boxplot.

- Het gemiddelde gewicht is 56,8 kilogram.
- De modus van het gewicht is 55 kilogram.



Figuur 14

Ook de gewichten van een vergelijkbare groep jongens zijn bekend. Bekijk het bijbehorende histogram.



Figuur 15

Om beide groepen afzonderlijk te bestuderen en onderling te kunnen vergelijken, moet je nog een aantal gegevens berekenen en diagrammen maken. Bedenk eerst goed welke gegevens en diagrammen je nog nodig hebt.

Bereken alle maten en maak de diagrammen. Beschrijf de opbouw van de meisjesgewichten en de jongensgewichten afzonderlijk en ook in vergelijking met elkaar. Doe dat op basis van alle maatgegevens en diagrammen en verwijs waar nodig in je argumentatie. Beargumenteer ook waarom jezelf de uitschieters juist wel of niet in de boxplots zou verwerken.

Testen

Opgave 19

Op een feestje zijn acht personen aanwezig. Je ziet een tabel met gegevens over de feestgangers.

naam	leeftijd in jaar	lengte in cm	favoriete drankje	gewicht in kg	zakgeld in € per week	vervoer-middel
Jan	15	187	cola	71	€ 15	fiets
Leo	16	178	colal	69	€ 20	fiets
Elske	15	172	sinas	66	€ 17	bus
Daphne	17	171	cola	67	€ 20	scooter
Mart	19	195	cassis	76	€ 19	auto
Durk	57	180	bier	74	€ 50	auto
Henk	17	187	cola	75	€ 18	fiets
Leo	15	179	cola	70	€ 18	brommer

Figuur 16

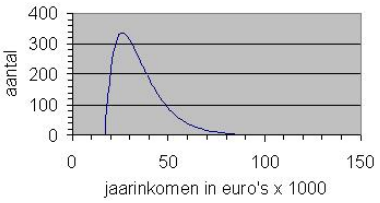
a Welke centrummaat zou je gebruiken om voor elke kolom de feestgangers te typeren?

b Welke spreidingsmaat zou je zo mogelijk gebruiken om voor elke kolom de feestgangers te typeren?

c Hoe zou je de doorsneefeestganger omschrijven?

Opgave 20

In de grafiek vind je de jaarinkomens van de werknemers van een grote fabriek.



Figuur 17

a Wat verdient de doorsneewerknemer van deze fabriek? Welke centrummaat heb je gekozen en waarom?

b Welke centrummaat is groter, de mediaan of de modus? Leg uit waar je dat aan kunt zien.

Opgave 21

Een groep leerlingen wordt tijdens de zwemles gevraagd om zo lang mogelijk hun adem in te houden onder water. Je ziet hier de bijbehorende tijden (in seconden).

17; 39; 17; 21; 26; 21; 31; 17; 37; 43
36; 17; 15; 29; 21; 31; 35; 23; 18; 17
26; 21; 28; 23; 22; 16; 37; 33; 27

a Bepaal de drie centrummaten van deze gegevens.

b Bepaal de spreidingsbreedte en de standaardafwijking.

c Teken een boxplot bij deze gegevens.

d Maak een klassenindeling met de eerste klasse $15- < 20$.

e Bereken het gemiddelde en de standaarddeviatie bij deze klassenindeling.

f Teken een bijpassende cumulatieve relatieve frequentiepolygoon. Bepaal daarmee de mediaan bij deze klassenindeling.

Practicum

Met de volgende practica kun je de statistische functies van de grafische rekenmachine doornemen. Onder andere staat er in hoe je gegevens in lijsten invoert en statistische diagrammen maakt. En ook hoe je de verschillende statistische centrummaten en spreidingsmaten bepaalt.

- **Statistiek met de TI84**
- **Statistiek met de TInspire**
- **Statistiek met de Casio fx-CG50**
- **Statistiek met de HPprime**
- **Statistiek en de NumWorks**

De statistische functies van Excel vind je in het volgende practicum. Er staat meer in dan op dit moment nodig is, maar onder andere kun je er nog eens in vinden hoe je diagrammen maakt. En ook hoe je de verschillende statistische centrummaten en spreidingsmaten bepaalt.

- **Data presenteren**

© 2024

Deze paragraaf is een onderdeel van het Math4All wiskundemateriaal.

Math4All stelt het op prijs als onvolkomenheden in het materiaal worden gemeld en ideeën voor verbeteringen in de content of dienstverlening kenbaar worden gemaakt. Klik op in de marge bij de betreffende opgave. Uw mailprogramma wordt dan geopend waarbij het emailadres en onderwerp al zijn ingevuld. U hoeft alleen uw opmerkingen nog maar in te voeren.

Email: f.spijkers@math4all.nl

Met de Math4All Foliostraat kunnen complete readers worden samengesteld en toetsen worden gegenereerd. Docenten kunnen bij a.f.otten@math4all.nl een gratis inlog voor de maatwerkdienst aanvragen.