

2.2 Centrum en spreiding

Inleiding

Een dataset bestaat meestal uit zoveel gegevens dat je ze moet groeperen in klassen en/of in overzichtelijke figuren. Je wilt ze vaak karakteriseren door een paar eenvoudig te bepalen getallen die aangeven waar het centrum van de dataset zit en hoeveel de gegevens gespreid liggen. De belangrijkste centrummaten ken je wel, denk maar aan modus en gemiddelde, maar ze worden nu toch nog even op een rijtje gezet. Ook worden enkele bijbehorende spreidingsmaten besproken.

Je leert in dit onderwerp

- een reeks waarnemingen samen te vatten met centrummaten, zoals gemiddelde, mediaan en modus;
- een reeks waarnemingen samen te vatten met spreidingsmaten, zoals spreidingsbreedte en interkwartielafstand;
- een reeks waarnemingen samen te vatten in een boxplot en deze boxplot te interpreteren.

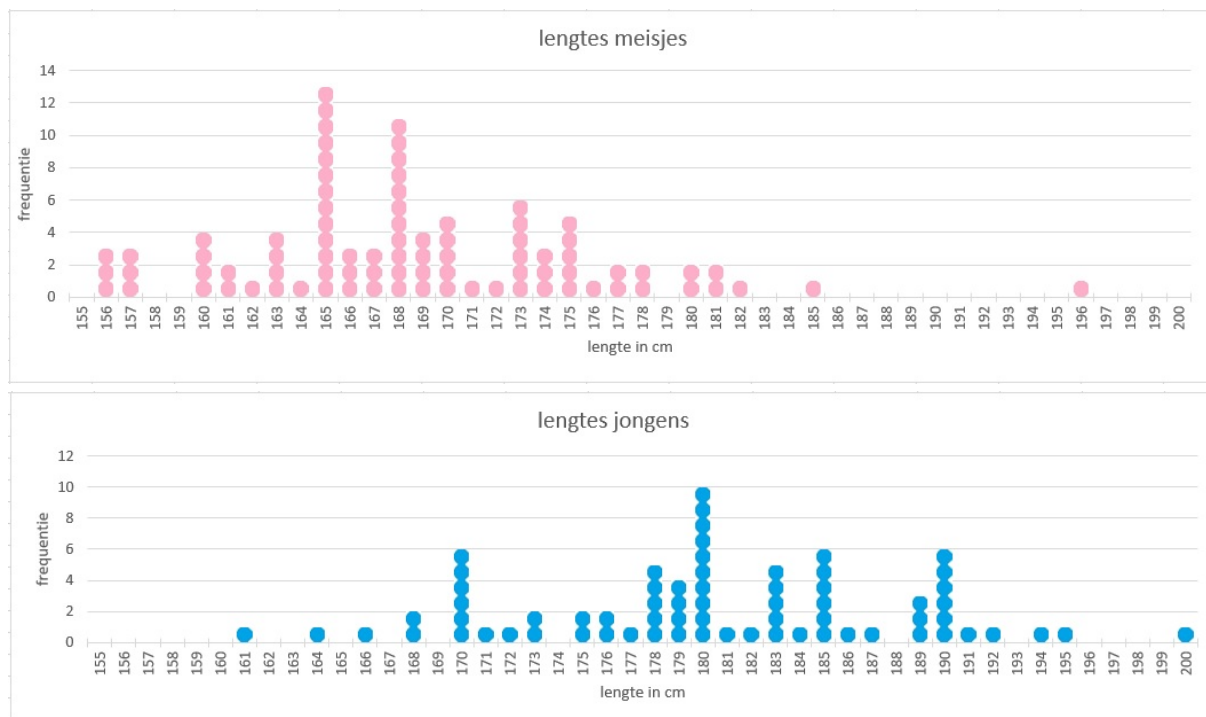
Voorkennis

- de begrippen data, populatie, steekproef, aselekt en representatief, kwantitatief en kwalitatief, absolute en relatieve frequentie, discrete en continue variabele, klassenbreedte, klassenmidden en klassengrens;
- dotplots en staafdiagrammen en andere diagrammen interpreteren.

Verkennen

Opgave V1

Bekijk de dotplots gemaakt vanuit de dataset **Gegevens 154 havo 4-leerlingen**.



Figuur 1

- a Waar zou je bij beide deelgroepen het midden van de frequentieverdeling plaatsen? Licht je antwoord toe.

- b** De mediaan is de lengte die op de helft van de verdeling zit, dus waar 50% van de lengtes onder zit (en dus ook 50% erboven). Bepaal de mediaan van de lengtes van de meisjes. Doe dat ook bij de jongens.
- c** Bij welke deelgroep zijn de gegevens het meest verspreid? Licht je antwoord toe.
- d** Laat bij de dotplot van de meisjes de twee grootste scores weg. Maakt dat veel verschil voor de mediaan? En voor de spreiding van de verdeling?
- e** Beantwoord dezelfde vragen als bij d voor de jongens.
- f** Bereken voor de lengte van de jongens de spreidingsbreedte. En voor de meisjes.
- g** De beide spreidingsbreedtes verschillen nauwelijks. Vind je dat de spreiding van de lengtes van de jongens en de meisjes vrijwel even groot is?
Je kunt een dataset in groepen van 25% verdelen, dus vier kwarten met even veel data. Deze groepen hebben de volgende vijf grenzen: het minimum, het eerste kwartiel Q_1 , de mediaan, het derde kwartiel Q_3 en het maximum.
- h** Welke lengtes hebben de 25% kleinste jongens?
- i** Laat met een berekening zien dat 25% van de langste meisjes inderdaad de lengtes 173 tot en met 197 cm heeft.
- j** Bepaal nu zowel voor de jongens als voor de meisjes uit de dotplots de grenzen van de vier kwarten. Verwerk dit in een schets.

Uitleg 1

Om snel belangrijke informatie over een grote dataset te krijgen, probeer je de dataset met een paar maten (getallen) samen te vatten. Je gebruikt hiervoor centrummaten en spreidingsmaten.

De bekende centrummaten zijn:

- de modus: de meest voorkomende waarde (bij een klassenindeling zeg je: modale klasse);
- de mediaan: de middelste waarde;
- het gemiddelde: alle waarden opgeteld en gedeeld door het totaal aantal waarden.

Een bekende maat voor de spreiding is de spreidingsbreedte. Bekijk je bijvoorbeeld een frequentieverdeling van lengtes van meisjes, dan is de spreidingsbreedte het verschil tussen de langste lengte en de kortste lengte van de meisjes. Als het langste meisje 197 cm is en het kleinste meisje 156 cm, is de spreidingsbreedte $197 - 156 = 41$ cm.

Bekijk wel eerst of centrummaten zinnig zijn voor je dataset. Bij de variabele *geboortjaar* is het gemiddelde bijvoorbeeld niet zinvol. En de spreidingsbreedte van het *profiel* van havo 4-leerlingen is ook onzinnig.

In het practicum kun je vinden hoe je centrum- en spreidingsmaten kunt berekenen met de grafische rekenmachine en ook hoe dit gaat met Excel of VUstat.

Opgave 1

Bekijk de dotplots van de lengtes van de jongens en de meisjes gemaakt vanuit de dataset **Gegevens 154 havo 4-leerlingen**. Deze opgave kun je met Excel maken.

Zie de figuur in **Opgave V1**.

- a** Maak van de lengtes van de jongens een overzicht met de drie centrummaten en de spreidingsbreedte. Je kunt dit ook met de grafische rekenmachine doen.
- b** Waarom zou je bij de meisjes de waarden 196 een uitschieter noemen? Heb je bij de jongens ook uitschieters?
- c** Welke centrummaat wordt door deze uitschieters sterk beïnvloed?
- d** Vind je het verantwoord om uitschieters weg te laten bij het samenvatten van een frequentieverdeling? Geef argumenten voor en tegen.

Opgave 2

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

- a Waarom kun je vanuit de frequentietabel de gemiddelde lengte niet precies uitrekenen, maar alleen nog 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?

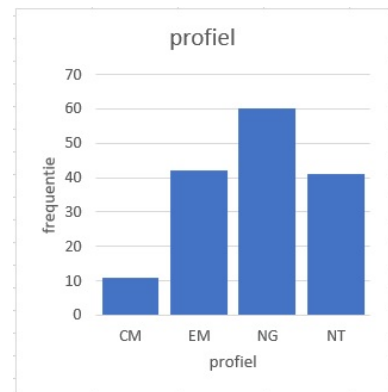
nettoweekloon (in €)	aantal werknemers
400 – < 450	2
450 – < 500	3
500 – < 550	4
550 – < 600	8
600 – < 650	3
650 – < 700	2
700 – < 750	2
750 – < 800	1

Tabel 1

Opgave 3

Bekijk het staafdiagram van de profielkeuzes van 154 havo 4-leerlingen.

- a Waarom kun je geen spreidingsbreedte vaststellen?
- b Je kunt wel vaststellen welk profiel de meeste jongens of de meeste meisjes hebben. Waarom kun je dat toch geen centrummaat noemen?
- c Bestaat er een zinvolle centrummaat voor deze statistische variabele?
- d “Het staafdiagram van de profielkeuzes van de jongens is veel schever dan dat van de meisjes.” Waarom is zo’n uitspraak hier niet zinvol?



Figuur 2

Uitleg 2

Bekijk de boxplot van de verdeling van de (gehele) eindcijfers van een groep examenkandidaten. In zo’n boxplot zijn de resultaten van deze groep in vier delen verdeeld. Die delen heten kwartielen.

Een boxplot heeft dus vijf grenzen:

- Linkergrens met het laagste getal.
- Rechtergrens met het hoogste getal.
- Middelste grens Q_2 , de mediaan.
- De tweede grens Q_1 tussen de linker grens en Q_2 ; de mediaan van de eerste helft.
- De vierde grens Q_3 tussen Q_2 en de rechtergrens; de mediaan 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.

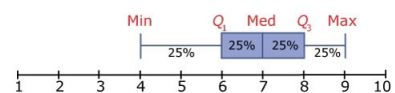
In het eerste kwartiel zitten de eerste 25%-waarden. Dit kwartiel is de rechtergrens van die 25%-waarden, dus het cijfer 6.

De mediaan (de middelste) is de rechtergrens van 50% van de waarden, dus het cijfer 7.

Het derde kwartiel is de rechtergrens van 75% van de waarden, dat is het cijfer 8.

De ‘interkwartielafstand’ is het verschil tussen het eerste kwartiel (Q_1) en het derde kwartiel (Q_3), dus $Q_3 - Q_1 = 8 - 6 = 2$.

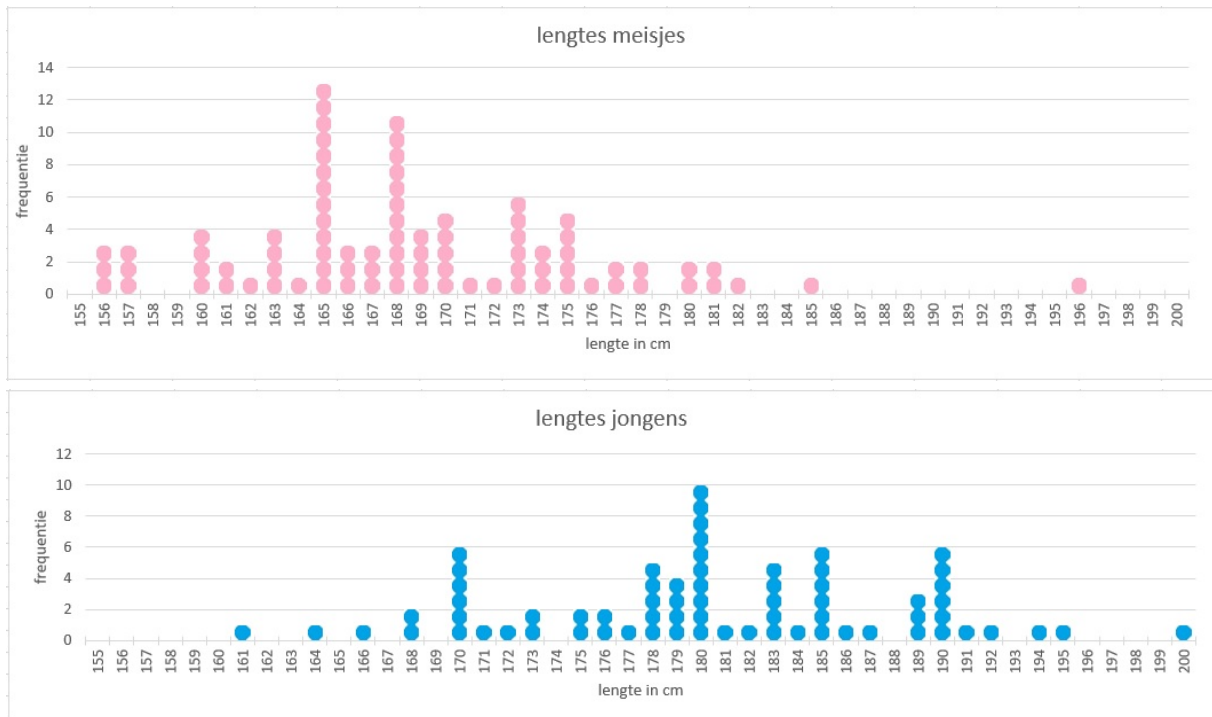
Een waarde die meer dan 1,5 keer de interkwartielafstand onder het eerste kwartiel of boven het derde kwartiel zit, wordt opgevat als een uitschieter.



Figuur 3

Opgave 4

Bekijk de dotplots van de lengtes van de jongens en de meisjes gemaakt vanuit de dataset **Gegevens 154 havo 4-leerlingen**. Deze opgave kun je met Excel maken.



Figuur 4

- a** Bepaal van de lengtes van de jongens het minimum, het maximum, de kwartielen en de mediaan. Hoe groot is de interkwartielafstand?
Lees in **Uitleg 2** nog eens wanneer je een gegeven een uitschieter noemt.
- b** Laat zien dat bij de jongens de waarden 161 en 200 cm uitschieters zijn.
- c** Laat deze uitschieters weg en maak een nieuw overzicht van de drie centrummaten en de twee spreidingsmaten.
- d** Welke spreidingsmaat wordt door deze uitschieters sterk beïnvloed en welke niet?
- e** Welke centrummaat wordt door deze uitschieters sterk beïnvloed?
- f** Vind je het verantwoord om uitschieters weg te laten bij het samenvatten van een frequentieverdeling? Geef argumenten voor en tegen.

Opgave 5

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

Waarom kun je met alleen een klassenindeling geen nauwkeurige boxplot maken?

Theorie en voorbeelden

Om te onthouden

Een frequentieverdeling kun je samenvatten met **centrummaten**:

- het **gemiddelde**, het evenwichtspunt van de verdeling (niet zichtbaar in een boxplot);
- de **mediaan**, de middelste waarde van de verdeling, op de helft van de boxplot;
- de **modus**, de meest voorkomende waarde (niet zichtbaar in een boxplot).

Daarnaast gebruik je **spreidingsmaten**:

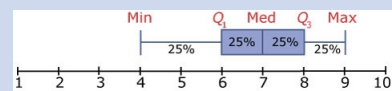
- de **spreidingsbreedte**, maximum – minimum (zichtbaar in een boxplot);
- de **interkwartielafstand**, $Q_3 - Q_1$ (zichtbaar in een boxplot).

Een frequentieverdeling kun je ook samenvatten in kwarten. Je krijgt de 5-getallensamenvatting.

- het **minimum**, de laagste waarde;
- het **eerste kwartiel** Q_1 , de bovengrens van het eerste kwart;
- de **mediaan** Q_2 , de bovengrens van het tweede kwart (dus precies op de helft);
- het **derde kwartiel** Q_3 , de bovengrens van het derde kwart;
- het **maximum**, de hoogste waarde.

De Q is afkomstig van het Engels woord ‘quartile’.

Je kunt de kwarten of kwartielen zichtbaar maken met een **boxplot**. De ‘box’ is het gebied tussen Q_1 en Q_3 . Elk kwartiel bevat 25% van de waarnemingen.



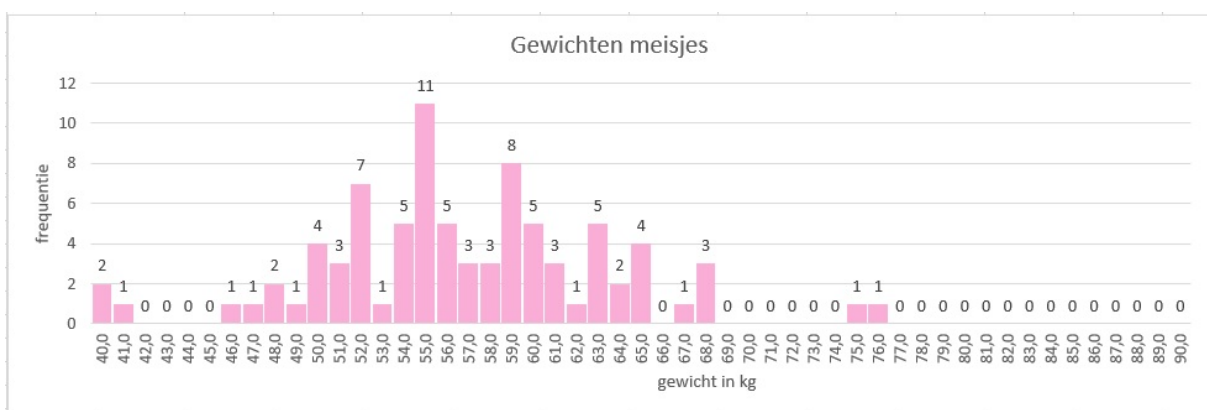
Figuur 5

Als je veel data hebt, is een boxplot overzichtelijker dan een dotplot of een staafdiagram. Met een boxplot verlies je echter ook informatie. In een boxplot kun je namelijk alleen nog het kwart aanwijzen waarin een bepaalde waarde zit.

Een **uitschieter** is een waarde die meer dan 1,5 keer de interkwartielafstand onder het eerste kwartiel of boven het derde kwartiel zit. Een uitschieter zie je niet in een boxplot, maar wel in een dotplot en ook in een staafdiagram.

Voorbeeld 1

Bekijk het staafdiagram met de gewichten van de 84 meisjes uit de dataset [Gegevens 154 havo 4-leerlingen](#).



Figuur 6

Bereken de mediaan en het gemiddelde van de gewichten in één decimaal nauwkeurig. Bereken ook de spreidingsbreedte en de kwartielafstand. Ga na welke van deze centrum- en spreidingsmaten het meest zinvol is. Je kunt de centrummaten handmatig uitrekenen of met de grafische rekenmachine.

Antwoord

De mediaan verdeelt de gewichten in twee gelijke delen (ze staan al op volgorde). Omdat van 84 meisjes het gewicht bekend is, neem je hiervoor het gemiddelde van het 42^e en het 43^e gewicht. Het 42^e gewicht is 56 kg en het 43^e ook, dus de mediaan is 56 kg.

Het gemiddelde gewicht bereken je vanuit de ruwe data. Je kunt het ook berekenen vanuit het gegeven staafdiagram. Houd dan wel rekening met de frequenties. Ga na, dat je $\approx 56,8$ kg krijgt.

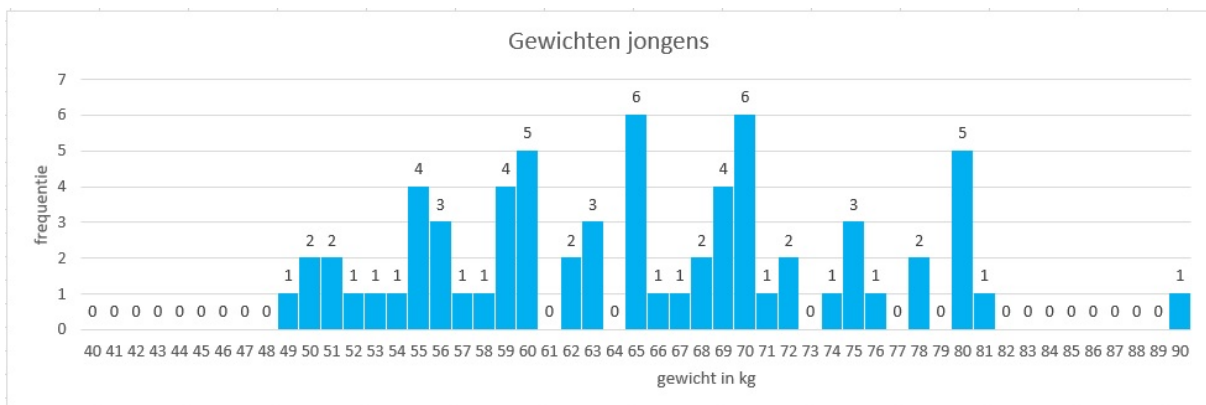
De spreidingsbreedte is hier $76 - 40 = 36$ kg.

Voor de kwartielfafstand moet je beide kwartielen Q_1 en Q_3 bepalen. Q_1 verdeelt de eerste helft van de gewichten weer in twee gelijke delen en is dus het gemiddelde van het 21^e en het 22^e gewicht. Dus $Q_1 = 52$ kg. En op dezelfde manier is $Q_3 = 60,5$ kg. De kwartielfafstand is daarom $60,5 - 52 = 8,5$ kg.

Hoe zinvol zijn nu al die maten? De modale lengte zegt niet veel over de verdeling. In dit geval zit die lengte nog redelijk in het midden, maar dat is toeval. Juist de waarden die meer in het midden zitten, komen weinig voor. De mediaan is een zinvolle maat, 50% van de lengtes zit eronder en 50% zit erboven. Ook het gemiddelde is een zinvolle maat: het is het evenwichtspunt van de verdeling. De kwartielfafstand is als maat voor de spreiding geschikter dan de spreidingsbreedte: die laatste maat wordt nogal eens bepaald door de uitschieters bij deze verdeling. Dat geldt voor de kwartielfafstand niet.

Opgave 6

Bekijk het staafdiagram met de gewichten van de jongens uit de dataset **Gegevens 154 havo 4-leerlingen**.

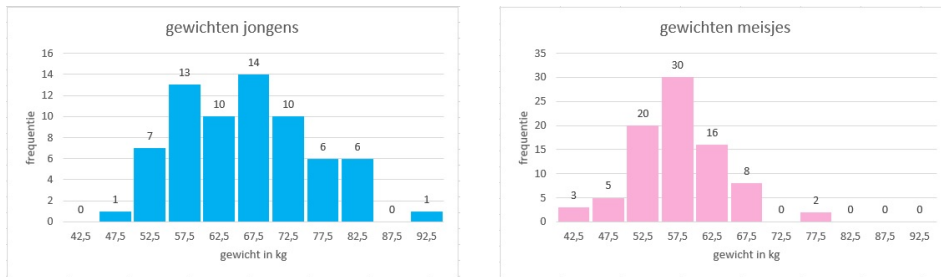


Figuur 7

- Bereken de mediaan en het gemiddelde van de gewichten van de jongens.
- Waarom is de modus niet vast te stellen?
- Bepaal de spreidingsbreedte en de kwartielfafstand.
- Er is bij de jongens één uitschieter. Welke centrummaat en/of spreidingsmaat verandert het sterkst als je deze uitschieter weglaat?
- Veranderen de centrum- en/of de spreidingsmaten als je alle absolute frequenties omrekent naar relatieve frequenties?
- Hoeveel wegen de 25% lichtste jongens?
- Hoeveel procent van de jongens weegt meer dan 78 kg?

Opgave 7

Bekijk de frequentieverdelingen van de gewichten van jongens en meisjes. Ze zijn gegroepeerd in klassen. De vraag is of je bij een indeling in klassen de centrummaten nog kunt berekenen.



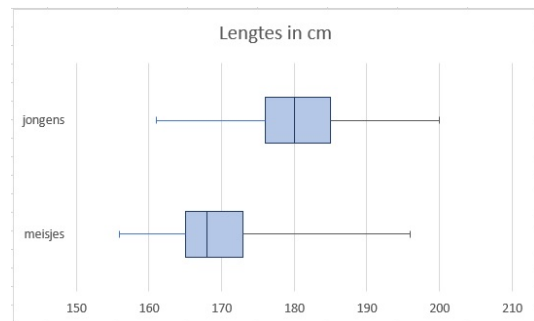
Figuur 8

- a. Waarom kun je vanuit deze frequentieverdelingen de mediaan niet meer vaststellen? In welke klasse zit de mediaan bij de meisjes? En bij de jongens?
- b. Maak bij deze klassenindeling frequentietabellen voor de gewichten van de jongens en de meisjes en voeg daaraan de klassenmiddens toe. Doe dit zowel in Excel als met je grafische rekenmachine. (Bekijk in het **Practicum** hoe je dit kunt doen.)
- c. Maak deze frequentieverdelingen met je grafische rekenmachine.
- d. Waarom kun je met deze klassen het gemiddelde alleen nog maar schatten?
- e. Geef een schatting van het gemiddelde met behulp van de klassenmiddens voor zowel de jongens als voor de meisjes. Doe dit zowel met behulp van Excel als met behulp van de grafische rekenmachine.
- f. Wijken je antwoorden af van die in **Voorbeeld 1**?

Voorbeeld 2

Je kunt een dataset met de lengte van meisjes en jongens in groepen van 25% verdelen, dus vier kwarten met evenveel data. Deze groepen hebben de volgende vijf grenzen: het minimum, het eerste kwartiel, de mediaan, het derde kwartiel en het maximum. De boxplots maken dat goed zichtbaar. De beide middelste kwarten vormen de box.

Welke lengtes hebben de 25% kleinste jongens?



Figuur 9

Antwoord

In de bovenste boxplot zie je de lengte van de 25% kleinste jongens: van 161 tot 176 cm.

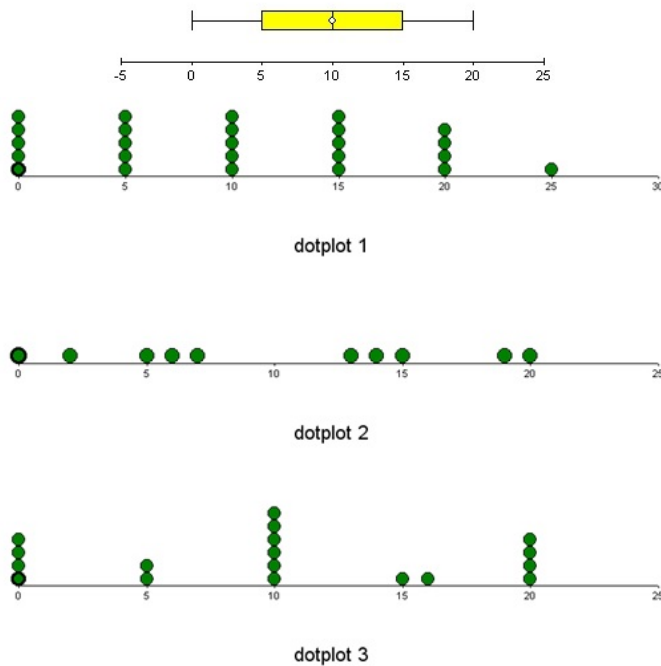
Opgave 8

Bekijk de boxplots in **Voorbeeld 2**.

- a. Hoeveel procent van de meisjes is langer dan 165 cm?
- b. Hoeveel procent van de meisjes is langer dan het derde kwartiel? En hoeveel procent heeft een lengte tussen de mediaan en het derde kwartiel?
- c. Waaraan kun je zien dat meer dan 75% van de meisjes kleiner is dan de langste van de 25% kleinste jongens?
- d. Kun je aan de boxplot zien hoe de data binnen de box verspreid zijn?

Opgave 9

Bekijk de boxplot en een aantal dotplots. De dotplots verschillen erg van vorm.



Figuur 10

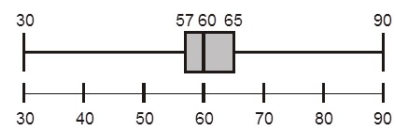
- a Verander bij elke dotplot één waarneming van plaats zodat de dotplot de verdeling van de boxplot goed weergeeft.
- b Verzin zelf een dotplot die goed weergegeven wordt door deze boxplot.

Verwerken

Opgave 10

Voor een toets kun je maximaal honderd punten scoren. Je ziet hoe een groep van veertig personen de toets heeft gemaakt. Van de scores is ook een samenvatting gemaakt in de vorm van een boxplot.

59 - 57 - 53 - 60 - 63 - 58 - 77 - 33 - 50 - 59
 58 - 75 - 62 - 54 - 53 - 78 - 59 - 68 - 65 - 62
 57 - 60 - 80 - 47 - 90 - 30 - 60 - 35 - 57 - 87
 63 - 65 - 63 - 58 - 65 - 70 - 73 - 58 - 63 - 55



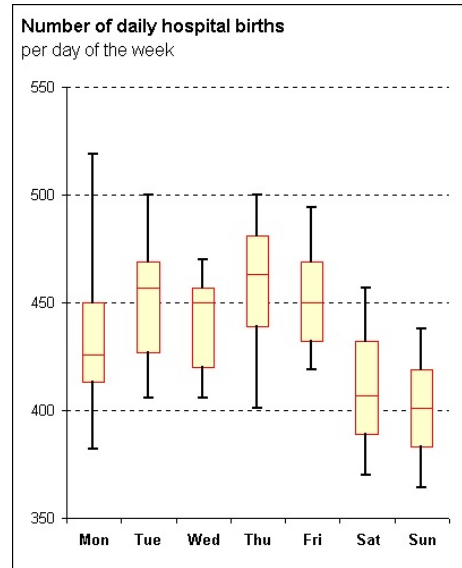
Figuur 11

- a Hoe groot is de gemiddelde score op één decimaal nauwkeurig?
- b Lees de kwartielen af uit de boxplot.
- c Welke centrummaat vat de data het beste samen?
- d Maak bij deze gegevens een frequentietabel met klassen $25- < 35$, $35- < 45$, etc. Maak met behulp van deze frequentietabel een schatting van het gemiddelde.
- e Waarom is dit een handige klassenindeling als het eindcijfer wordt berekend door de score door 10 te delen en je snel wilt kunnen vaststellen hoeveel onvoldoendes (eindcijfer lager dan 5,5) er zijn?

Opgave 11

Bekijk de boxplots van het aantal geboortes in ziekenhuizen per dag voor de verschillende dagen van de week.

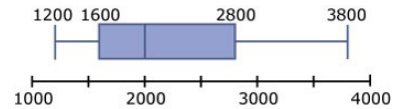
- a Op welke dag van de week is de spreidingsbreedte van het aantal geboortes in ziekenhuizen het grootst? Waarom kun je de dagen niet goed vergelijken met behulp van de spreidingsbreedtes?
- b Welke conclusie kun je trekken uit deze boxplots? Er zijn meerdere conclusies.
- c Hoeveel procent van de zondagen zijn er minder dan vierhonderd geboortes in ziekenhuizen?
- d Benader het gemiddelde aantal bevallingen op donderdagen.



Figuur 12

Opgave 12

In een bedrijf met 120 medewerkers is het modale salaris ongeveer € 1600,00 per maand. Het gemiddelde salaris is € 1800,00 per maand. Het hoogste salaris is dat van de algemeen directeur. Deze boxplot vat de verdeling van de salarissen samen.



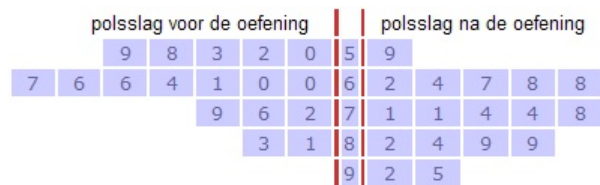
Figuur 13

Bereken steeds het modale salaris en het gemiddelde salaris. En wat gebeurt er met de spreidingsbreedte en de interkwartielafstand als

- a alle medewerkers een loonsverhoging krijgen van 3%?
- b alle medewerkers een maandelijkse toeslag krijgen van € 200,00?
- c het salaris van de algemeen directeur met € 840,00 per maand verhoogd wordt?

Opgave 13

Als je in de sportzaal een tijdje een bepaalde oefening hebt gedaan, gaat je polsslag omhoog. In het tweezijdige steelbladdiagram vind je wat data. Van elke sporter werd één keer voor en één keer na de oefening de polsslag gemeten. De tientallen staan in de stam (rood), de eenheden op een blad (zwart).



Figuur 14

- a Waarom zegt de modale polsslag hier weinig over het centrum van de verdeling? Is de modale polsslag een zinvol getal?
- b Bereken de gemiddelde polsslag voor en ook na de oefening in één decimaal nauwkeurig. Is het gemiddelde bij deze gegevens een bruikbare centrummaat om te vergelijken?
- c Is het wel handig om de polsslag voor en na de oefening apart in beeld te brengen?

Opgave 14

Bekijk de dataset met de gegevens over de **Sportprestaties van 74 brugklassers**. Werk in Excel.

- a Bereken voor het vergooien alle centrummaten en alle spreidingsmaten vanuit de ruwe data.
- b Waarom kun je dit altijd beter vanuit de ruwe data doen dan vanuit een klassenindeling?
- c Probeer conclusies te trekken over het vergooien. Gebruik daarbij de centrum- en de spreidingsmaten. Welke centrum- en spreidingsmaten zijn hier zinvol?

Toepassen

Opgave 15: Hematocrietwaarde

Uit de wielersport komen regelmatig berichten over dopinggebruik. Wielrenners lijken naar verboden middelen te grijpen om hun prestaties te verhogen. Een van de meest genoemde stoffen is erythropëetine, kortweg EPO. Dit middel bevordert de aanmaak van rode bloedlichaampjes, waardoor de zuurstoftransportfunctie van het bloed wordt vergroot. Je gaat hierdoor beter presteren.

De hematocrietwaarde is de hoeveelheid rode bloedlichaampjes als percentage van de totale hoeveelheid bloed. Die hematocrietwaarde stijgt als een wielrenner EPO gaat gebruiken.

Bij een wielervedstrijd in 1997 heeft men de hematocrietwaarde van een aantal wielrenners gemeten. De meetresultaten staan in de tabel.

hematocrietwaarde	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57
frequentie	2	3	5	11	15	8	4	3	2	1	0	0	1	2	0	0	1

Tabel 2

- a Bereken de gemiddelde hematocrietwaarde van deze wielrenners.

Ook in 1998 en 1999 heeft men bij deze wielervedstrijd van een aantal wielrenners de hematocrietwaarde gemeten. In 1998 was de gemiddelde hematocrietwaarde 45,9. De hematocrietwaarden uit 1999 zijn verwerkt in deze boxplot.



Figuur 15

- b Toon aan dat, op grond van de boxplot, de gemiddelde hematocrietwaarde in 1999 groter was dan in 1998.

(bron: examen havo wiskunde A in 2002, eerste tijdvak)

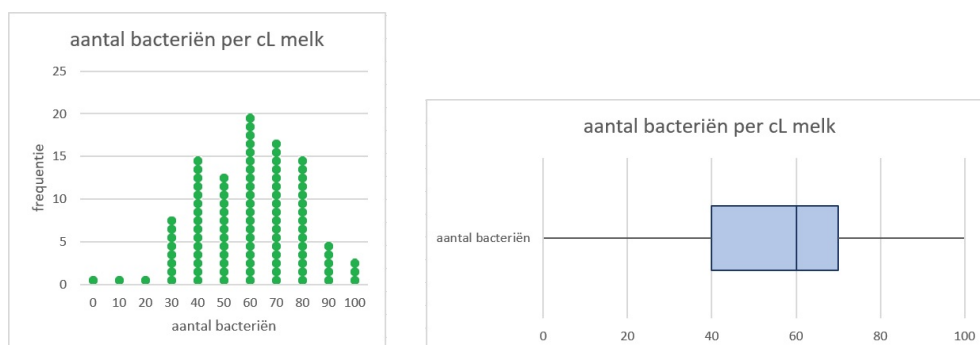
Opgave 16: Bacteriën in melk

Een bedrijf haalt elke dag melk bij honderd boeren in de regio op. Voordat de melk in de transporttank gaat, wordt bij elke boer een monster van de melk genomen. In het lab wordt de melk onderzocht op het voorkomen van bacteriële vervuiling. Daartoe wordt in elk monster het aantal fecale bacteriën per centiliter geteld. De totale hoeveelheid melk die elke dag wordt opgehaald, heet een 'dagproductie'.

De dienst die verantwoordelijk is voor de kwaliteitsbewaking, stelt als eis een maximum van 100 fecale bacteriën per milliliter. In de tabel staan honderd waarden gegeven die het lab in een bepaalde dagproductie heeft gevonden. De laborante heeft berekend dat het gemiddelde 59 bacteriën per centiliter is. Ze heeft voor haar analyse ook nog twee representaties gemaakt.

aantal bacteriën	freq
0	1
10	1
20	1
30	8
40	15
50	13
60	20
70	17
80	15
90	5
100	3

Tabel 3



Figuur 16

- a De laborante heeft het gemiddelde aantal bacteriën in de honderd monsters berekend. Is dit ook het gemiddelde van de gehele dagproductie? Licht je antwoord toe.
- b Het laboratorium is alert op uitschieters omdat die op bijzonderheden kunnen wijzen. Vaak zijn dat meetfouten. Het lab hanteert een eigen vuistregel voor uitschieters: waarden die liggen buiten het interval mediaan + 1,5 · interkwartielafstand en - 1,5 · interkwartielafstand. Ga na hoeveel uitschieters er zijn.
- c De laborante besluit de twee monsters die het laagste scoorden opnieuw te meten. Dat levert in beide gevallen een score van vijftig bacteriën op en daarmee een nieuwe dataset. Beredeneer dat in deze nieuwe set het gemiddelde groter en de spreidingsbreedte kleiner is dan in de oorspronkelijke dataset.
- d Beredeneer wat er verandert aan de mediaan en aan de interkwartielafstand.

(bron: voorbeeldopgaven syllabus, 2014)

Testen

Opgave 17

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

- 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 met je grafische rekenmachine een staafdiagram van de relatieve frequenties.
- d In welke klasse zit de mediaan? Kun je precies zeggen hoe groot die mediaan is?
- e Bereken een schatting van het gemiddelde.

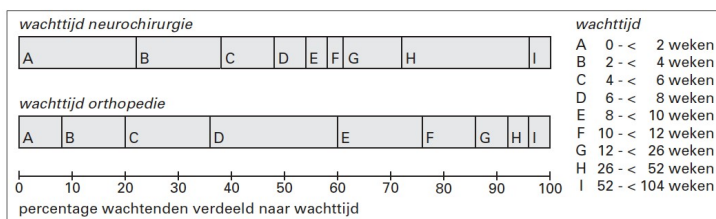
lengte regenworm (cm)	aantal
0,0– < 3,0	4
3,0– < 6,0	8
6,0– < 9,0	17
9,0– < 12,0	22
12,0– < 15,0	23
15,0– < 18,0	17
18,0– < 21,0	6
21,0– < 24,0	2
24,0– < 27,0	1

Tabel 4

Opgave 18

In de gezondheidszorg staan veel mensen op een wachtlijst, bijvoorbeeld voor een behandeling in een ziekenhuis of voor een plaats in een verzorgingshuis. De regering wil dat minder mensen op een wachtlijst staan. Ook wil men dat mensen die op een wachtlijst staan, minder lang hoeven te wachten. Voor alle specialismen, waaronder neurochirurgie en orthopedie, heeft men de wachtlijsten in kaart gebracht. Men heeft bekeken hoeveel mensen op een wachtlijst staan en hoelang ze moeten wachten. In de figuur staat een weergave van de wachttijd bij neurochirurgie en bij orthopedie. In de figuur kun je bijvoorbeeld over de wachtenden op een behandeling bij neurochirurgie aflezen:

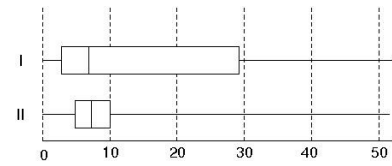
- bijna 40% van de mensen is binnen vier weken aan de beurt en
- meer dan 25% moet 25 weken of langer wachten.



Figuur 17

- a Hoeveel procent van de wachtenden bij neurochirurgie moet tussen de vier en de tien weken wachten? Licht je antwoord toe.

- b Beide afdelingen kunnen op grond van de gegevens in de figuur beweren dat ze het beter doen dan de andere afdeling. Noem voor elke afdeling een argument dat ze kunnen aanvoeren.
- c De gegevens uit de figuur zijn verwerkt in twee boxplots. Geef aan welke boxplot bij welke afdeling hoort.



Figuur 18

Practicum

Als je met niet al te grote datasets te maken hebt, kun je met de **grafische rekenmachine** de gegevens verwerken, analyseren en presenteren. Hoe dat gaat zie je in de practica:

- [Statistiek en de TI84](#)
- [Statistiek en de TIInspire](#)
- [Statistiek en de Casio cfx-9850](#)
- [Statistiek en de HPprime](#)
- [Statistiek en de NumWorks](#)

Met **Excel** (een spreadsheetprogramma, een rekenblad) werken is bij statistiek eigenlijk onontbeerlijk. Je kunt er grote hoeveelheden gegevens in kwijt. Die gegevens kun je ordenen en presenteren. Bekijk de eerste vier delen van het practicum:

- [Data presenteren](#)

Je kunt ook data analyseren en presenteren met de app 'Data analyse' van **VUstat**. Daarin kun je eigen databestanden vanuit Google-Drive toevoegen, maar er zijn ook diverse datasets beschikbaar. Ga hiervoor naar:

- [Data analyse VUstat](#)



© 2021

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.

Email: f.spijkers@math4all.nl

Met de Math4All maatwerkdienst kunnen complete readers worden samengesteld en toetsen worden gegenereerd. Docenten kunnen bij a.f.otten@xs4all.nl een gratis inlog voor de maatwerkdienst aanvragen.
