

juli 1992

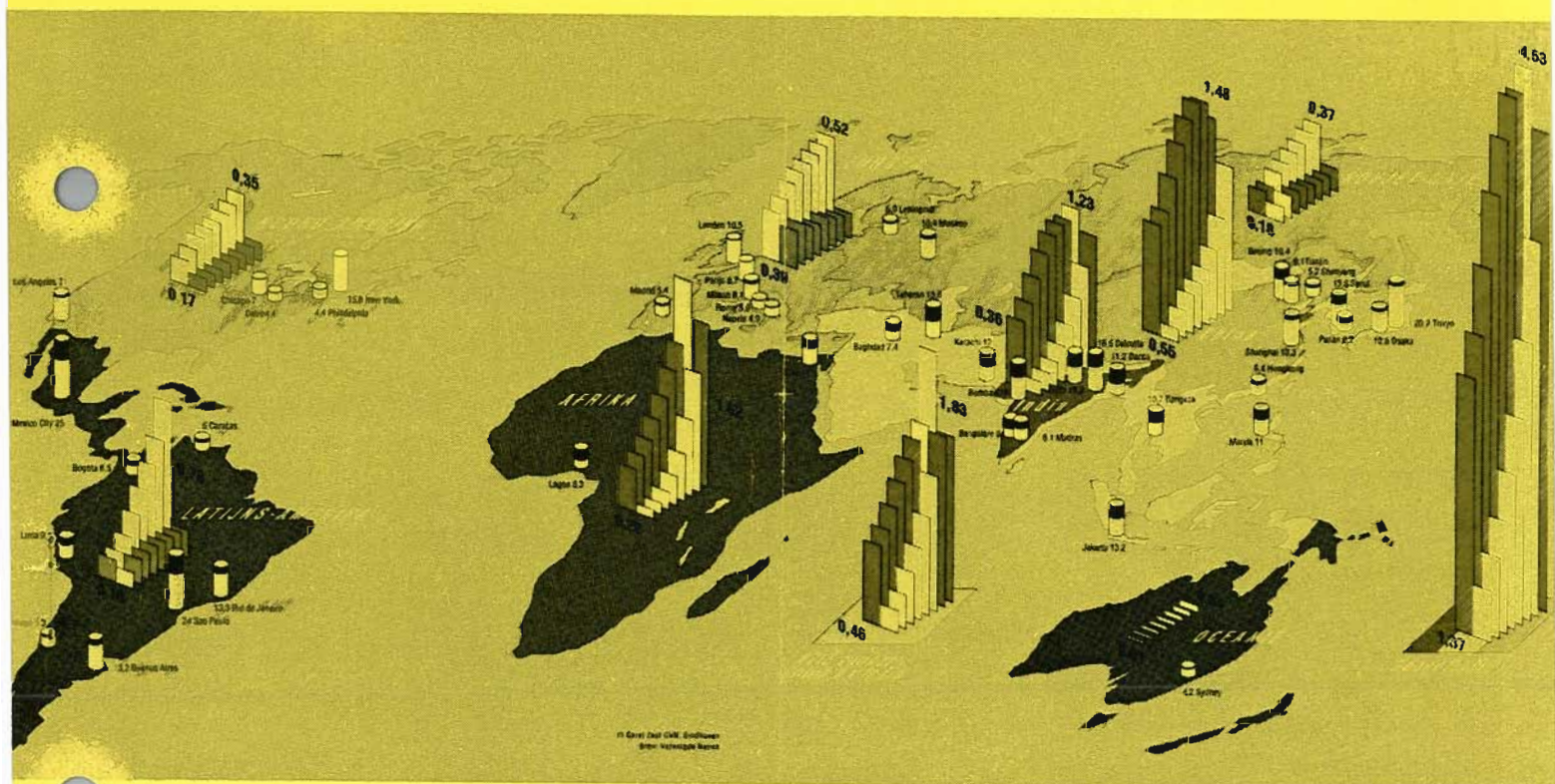
experimentele versie



Freudenthal instituut
Oerarchief

Statistiek

Leerlingenmateriaal



Publikatie van het team W12-16
onder verantwoordelijkheid van de
Commissie Ontwikkeling Wiskundeonderwijs

Ontwerp: Heleen Verhage, i.s.m. Gerrit van den Heuvel en Hans Krabbendam.

Inhoudsopgave

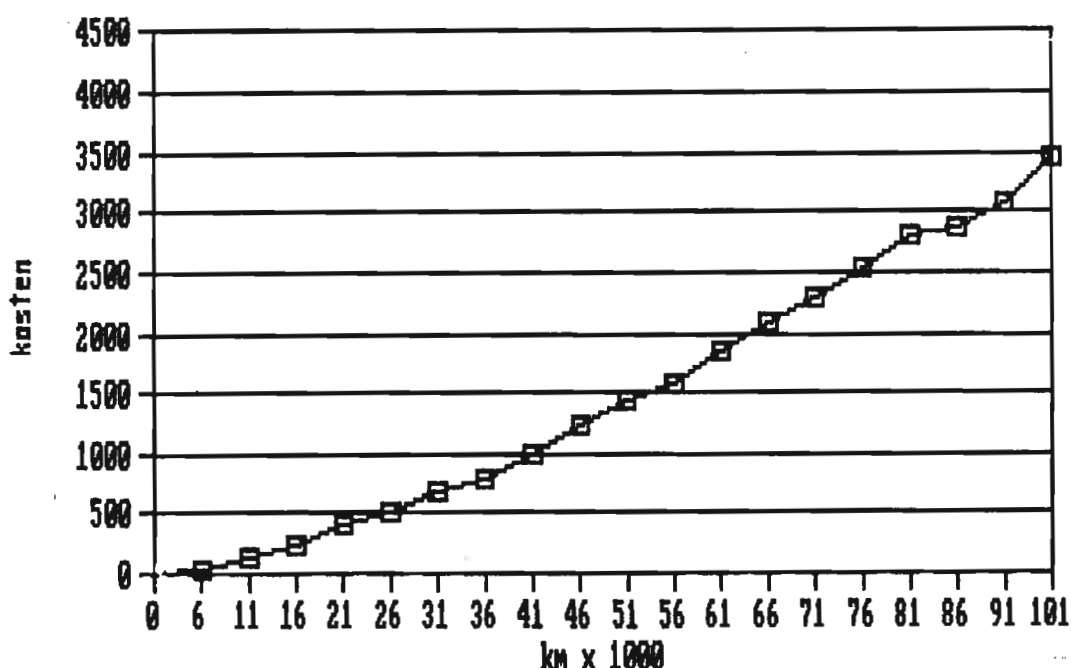
Hoofdstuk 1	Grafieken en tabellen in de statistiek	3
Hoofdstuk 2	Het stam-bladdiagram	12
Hoofdstuk 3	Beschrijven en voorspellen	18
Hoofdstuk 4	Het verwerken van gegevens	23
Hoofdstuk 5	De boxplot	29
Werkbladen		41

1. a. Welke drie typen auto's worden er onderscheiden?
 b. Wat zijn de kosten bij 6000 km voor de drie typen auto's?
 c. En bij 101.000 km?
2. In de grafiek moet de linker auto de steilste klim maken. Is dat de 'onderhoudsgevoelige auto' of de 'onderhoudsvriendelijke auto' uit het bijschrift?
3. De onkosten voor de drie verschillende typen auto's zijn in één grafiek getekend. Waarom zal men dat gedaan hebben?
4. a. Hoeveel bedragen volgens deze grafiek de onderhouds- en reparatiekosten van een 'vriendelijke' auto die 51.000 km heeft gereden?
 b. Hoeveel cent per km is dat?

Omdat de grafiek 'ruimtelijk' getekend is, gaat het aflezen voor de andere auto's moeilijk.

5. a. Hoe kun je de kosten bij 51.000 km voor de gemiddelde auto aflezen? (Tip: gebruik lineaal of geodriehoek).
 b. Hoeveel kost de gemiddelde auto per km aan onderhoud en reparatie?
 c. Bepaal ook de gemiddelde kosten per km voor de onderhoudsgevoelige auto.

De grafiek van de vorige bladzijde is een variant op de gewone *lijngrafiek*, maar dan mooi uitgevoerd door een tekenaar. Zonder franje ziet de grafiek voor de gemiddelde auto er zo uit:



6. Schets (je hoeft dus niet alle punten precies te tekenen) op het werkblad zelf de grafieken voor de de twee andere typen auto's.

Hoofdstuk 1 - Grafieken en tabellen in de statistiek

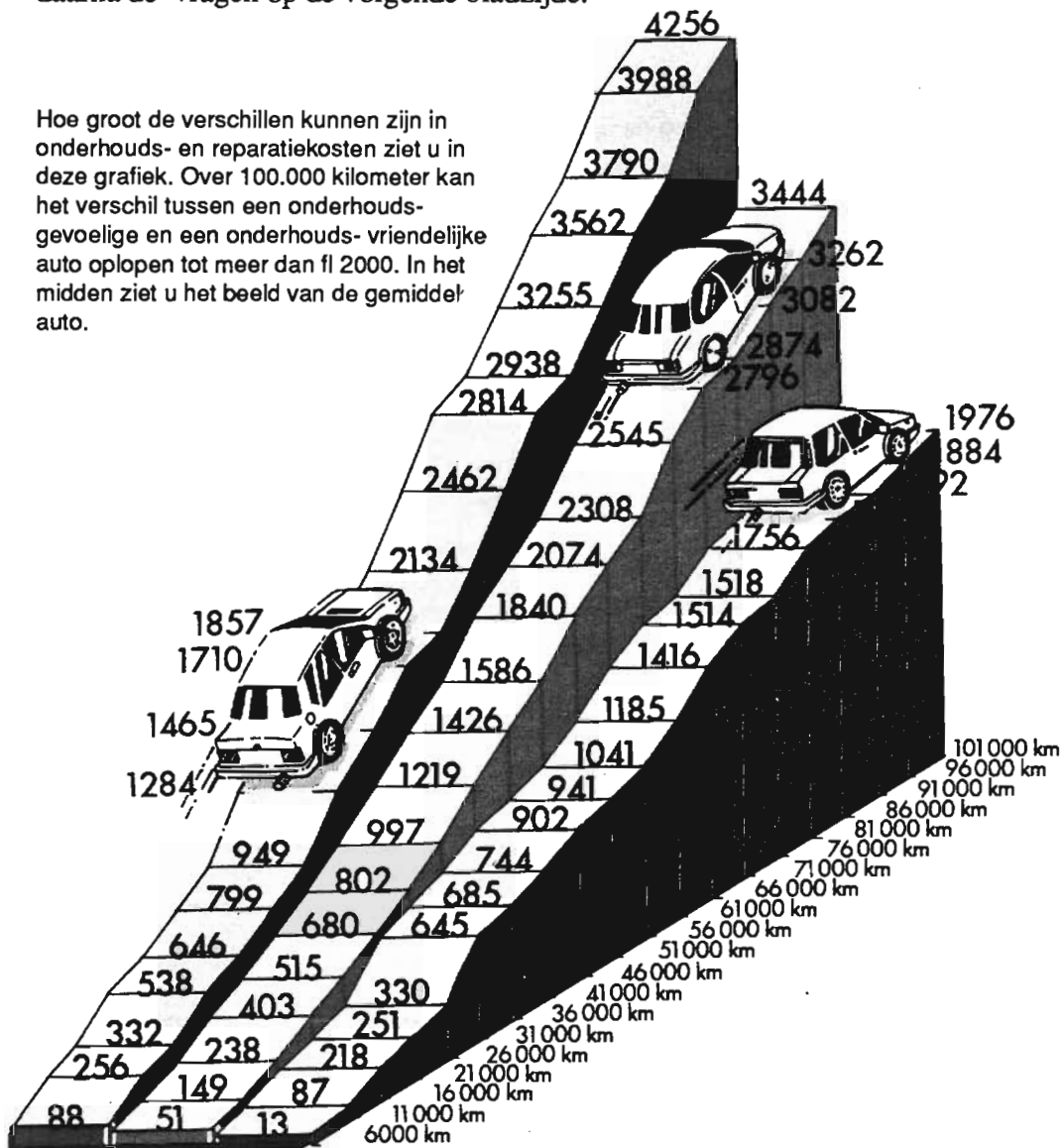
Dit hoofdstuk gaat over grafieken en tabellen in de statistiek. Grafieken en tabellen kom je overal tegen, vooral in kranten en tijdschriften. Het zijn vaak mooie plaatjes, maar wat je er uit af kunt lezen, is niet altijd meteen duidelijk. De bedoeling van grafieken en tabellen is, om grote hoeveelheden gegevens kort en overzichtelijk weer te geven.

Consumentenstatistiek

De Consumentenbond onderzoekt de kwaliteit van allerlei producten. De resultaten van de testen worden gepubliceerd in de *Consumentengids*, het tijdschrift van de Consumentenbond. Daar kun je als consument je voordeel mee doen.

De Consumentenbond vergelijkt bijvoorbeeld geregeld de kosten van onderhoud en reparaties van auto's. Hieronder staat een grafiek met de uitkomsten van zo'n onderzoek. Bekijk de grafiek en lees de toelichting die er bij staat. Beantwoord daarna de vragen op de volgende bladzijde.

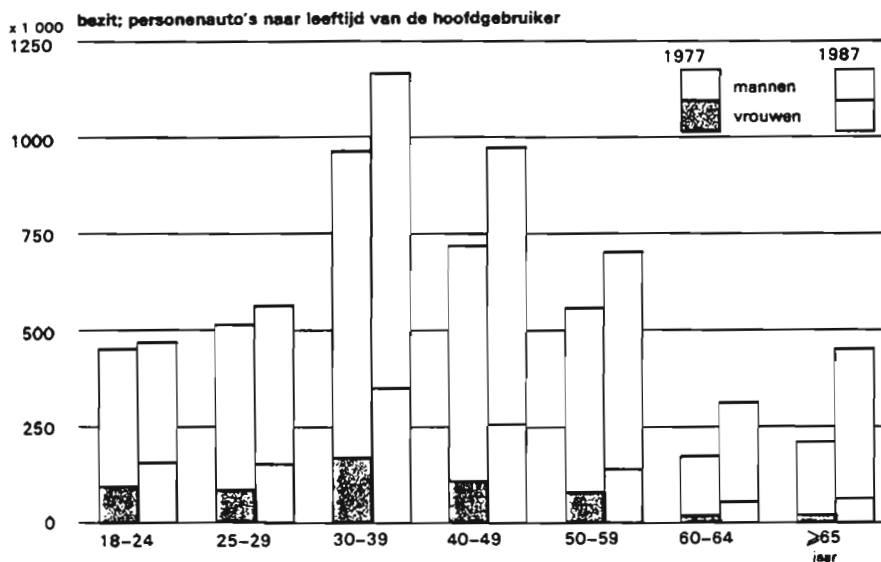
Hoe groot de verschillen kunnen zijn in onderhouds- en reparatiekosten ziet u in deze grafiek. Over 100.000 kilometer kan het verschil tussen een onderhouds-gevoelige en een onderhouds-vriendelijke auto oplopen tot meer dan fl 2000. In het midden ziet u het beeld van de gemiddelde auto.



Het Centraal Bureau voor de Statistiek

Het CBS (Centraal Bureau voor de Statistiek) in Voorburg verzamelt en bewerkt alle mogelijke gegevens over ons land. Het CBS geeft elk jaar het *Statistisch Zakboek* uit. In het Statistisch Zakboek staan tabellen en grafieken over alle mogelijke onderwerpen. De grafiek hieronder komt uit het Statistisch Zakboek van 1988. Het is een *staafdiagram* van het autobezit in Nederland in 1977 en 1987.

In het staafdiagram staat veel informatie tegelijk: over verschillende jaren ('77 en '87), verschillende leeftijdscategorieën en voor mannen en vrouwen apart.



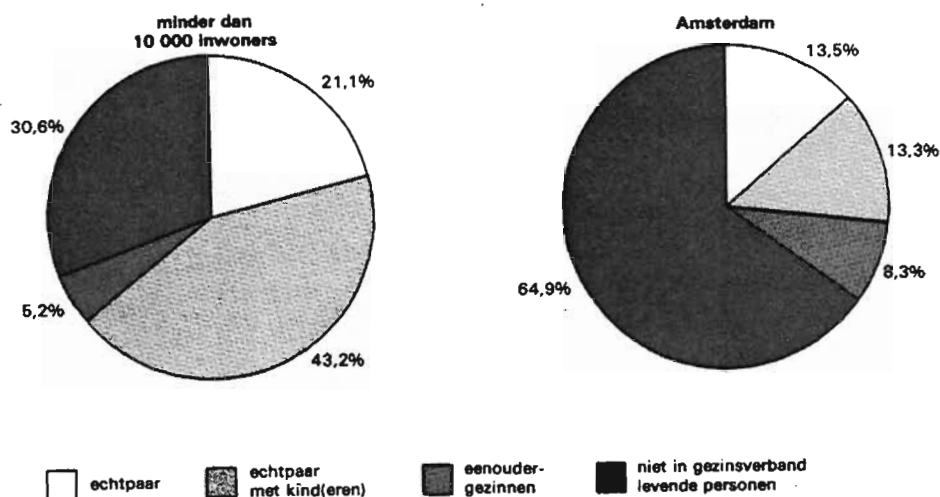
7. a. Wat valt er allemaal bij de assen af te lezen?
Gebruik werkblad 2.
b. Kleur (op het werkblad) de staven die het autobezit van vrouwen in 1987 aangeven.
8. a. Kijk eens naar de staven voor de 30-39 jarigen. Welke *informatie* kun je hier allemaal uithalen? Probeer zoveel mogelijk dingen op te schrijven.
b. In 1987 zijn er 1,2 miljoen vrouwen van 30-39 jaar en ook 1,2 miljoen mannen van die leeftijd. Welk deel van hen heeft een auto?
9. a. Is het totale autobezit sinds 1977 toe- of afgenomen?
b. Welke groepen bezitten in 1987 meer/minder auto's dan in 1977: jongeren, ouderen, mannen, vrouwen?
10. Vul de zinnen hieronder aan tot een bewering die klopt met de grafiek:
 - a. - "Het autobezit onder mannen van 18 tot 30 jaar is"
 - b. - "Het autobezit onder vrouwen van 30 tot 50 jaar is"
 - c. - "Het autobezit onder bejaarden is"

Gebruik werkblad 2.

11. Teken zelf een staafdiagram met alléén het aantal mannelijke autobezitters in 1987, verdeeld naar leeftijd.

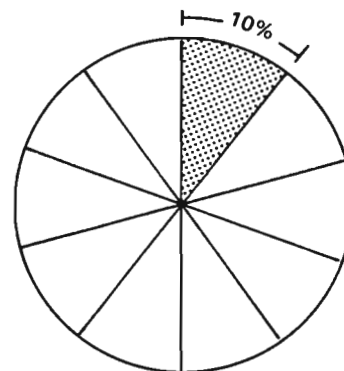
Gezinssamenstelling

De cirkeldiagrammen hieronder komen ook uit het Statistisch Zakboek. Ze brengen de samenstelling van gezinnen (per 1-1-1987) in beeld. Links voor de kleine gemeenten (minder dan 10.000 inwoners), rechts voor de stad Amsterdam.



12. Zoals je ziet zijn de twee cirkeldiagrammen heel verschillend. Kun je drie verschillen noemen? Probeer de verschillen ook te verklaren.

Voor het tekenen van een cirkeldiagram is de '100%-cirkel' een handig hulpmiddel. De cirkelschijf is dan niet in graden, maar in procenten verdeeld. Elke 'taartpunt' is 10%. Hiernaast staat zo'n 100%-cirkel.



13. Hoeveel graden is de hoek die bij 10% hoort?

Gezinssamenstelling van heel Nederland.

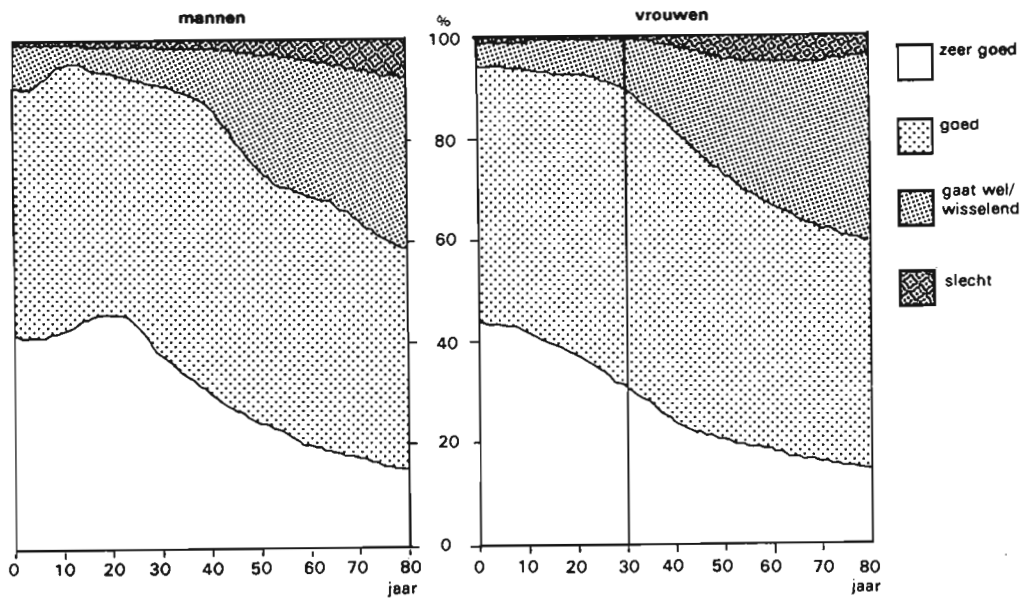
type gezin	percentage
echtpaar	19,0 %
echtpaar met kinderen	32,9 %
éénoudergezinnen	6,3 %
niet in gezinsverband levend	41,8 %

Gebruik werkblad 3.

14. a. Teken nu zelf een cirkeldiagram (geodriehoek!) bij deze tabel. Teken zoals hierboven gedaan is: begin middenboven met de echtparen en draai rechtsom.
b. Kleur de cirkelsegmenten ook in.
15. Vergelijk het diagram dat je getekend hebt met de twee cirkeldiagrammen hierboven. Kun je weer drie verschillen vinden?

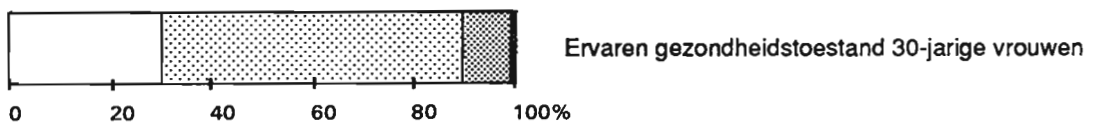
Gezondheidstoestand

Het CBS heeft de 'ervaren gezondheidstoestand naar leeftijd' onderzocht. Het resultaat van dit onderzoek staat in de grafieken hieronder.



16. Wat zouden ze bedoelen met 'ervaren gezondheidstoestand'?
17. a. Bedenk een manier om gegevens te verzamelen over de 'ervaren gezondheidstoestand'.
b. Welke problemen kom je dan tegen?
18. a. Op welke leeftijd voelen volgens de grafiek mannen zich het gezondst?
b. En vrouwen?
19. Hoe is het volgens de grafiek gesteld met de gezondheid van jonge kinderen?
20. a. Tot welke leeftijd voelt meer dan 80% van de mannen zich goed tot zeer goed gezond?
b. En hoe zit dat bij de vrouwen?

In de grafiek van de vrouwen is bij de leeftijd van 30 jaar een lijn getrokken. Een dwarsdoorsnede van de grafiek als het ware. Met behulp van die lijn kun je een *reepdiagram* voor de 30-jarige vrouwen tekenen:



Gebruik werkblad 3.

21. Teken zelf zo'n reep voor de vrouwen van 50 jaar.

Lekkerbekjes

In de Consumentengids staat in een test over lekkerbekjes (bepaalde soort gebakken vis) geschreven:

Lekkerbekjes, meestal gebakken kabeljauw of wijting, doen hun naam geen eer aan. Ze hebben een veel te dikke en vette bakkorst, ze bezorgen je een mond vol graten, de geur, kleur of smaak wekt weerzin op, of de gebruikte frituurolie deugt niet. Daarbij doet het er weinig toe of je ze op de markt of in de viswinkel koopt. Jammer, want vis kan een belangrijke bijdrage leveren aan het dagelijks menu.

Bij het artikel staat deze tabel:

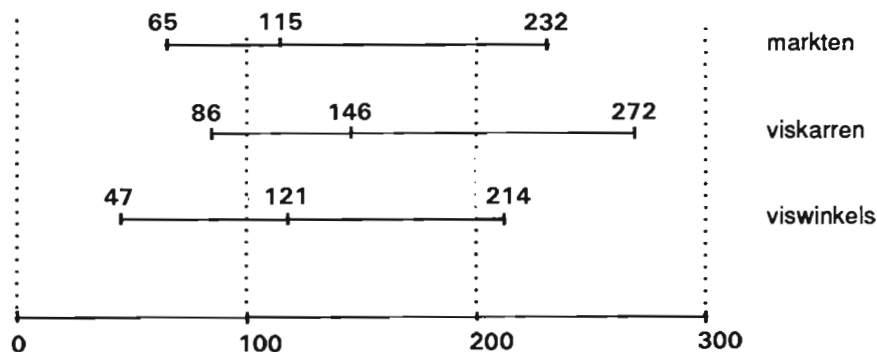
LEKKERBEKJES						
VERKOOPPUNT	LAAGSTE-GEN.-HOOGSTE GEWICHT PER STUK (G)	LAAGSTE-GEN.-HOOGSTE PRIJS PER KG (P)	GOED-REDELIJK-SLECHT VAKTECHNISCH	GOED-REDELIJK-SLECHT KLEUR, GEUR EN SMAAK	GOED-REDELIJK-SLECHT FRIJUROLIE	AANTAL ONBEROEPBAAR
OP 14 MARKTEN	65-115-232	14,20-18,20-24,20	2- 4- 8	2-10-2	8-4-2	0
BIJ 8 VISKARREN ¹⁾	86-146-272	13,80-18,20-24,50	2- 6- 0	0- 5-3	4-2-2	0
IN 18 VISWINKELS ²⁾	47-121-214	15,00-18,70-26,80	7- 7- 4	4-13-1	10-3-5	3
40 VERKOOPPUNTEN	47-125-272	13,80-18,40-26,80	11-17-12	6-28-6	22-9-9	3

1 Inclusief viskiosken 2 Inclusief supermarkten

CONSUMENTENGIDS APRIL 1989

22. a. Wat is volgens de tabel het gemiddelde gewicht van een lekkerbekje?
b. En wat is de gemiddelde kiloprijs?
23. Een controleur koopt een aantal lekkerbekjes en weegt die. Het resultaat: 232, 95, 185, 261, 103, 145, 163, 207, 133 en 178 gram.
Wat is het gemiddelde gewicht?
24. Voor het gewicht van een lekkerbekje staan in de eerste kolom steeds *drie* getallen.
Zoek in de tabel de getallenreeks 65-115-232 op.
a. Wat betekenen deze drie getallen?
b. Welke reeks hoort bij de viskarren?
c. En bij de viswinkels?

Van die reeksen van drie getallen kun je een grafiek maken, zo:



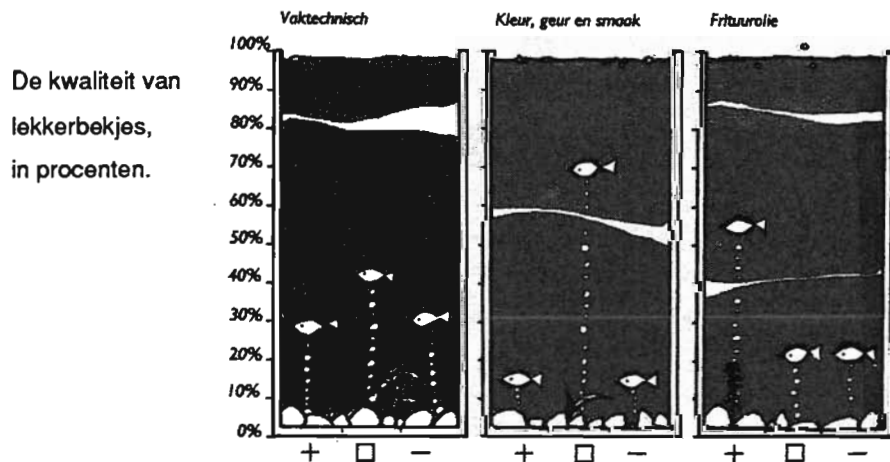
25. Gebruik werkblad 4.
Teken er zelf het lijnstukje voor 'alle verkooppunten' bij.

De grafiek met de lijnstukjes geeft dezelfde informatie als de rijtjes van drie getallen. Namelijk informatie over 'het gemiddelde gewicht' en over 'de spreiding in gewicht' bij de verschillende verkooppunten.

26. Wat zal er bedoeld worden met 'de spreiding in gewicht'?
27. Bij welke van de drie typen verkooppunten is:
- het gemiddelde gewicht het kleinst?
 - het gemiddelde gewicht het grootst?
 - de spreiding in gewicht het kleinst?
 - de spreiding in gewicht het grootst?
28. a. Wat kun je zeggen over het gemiddelde gewicht bij *alle verkooppunten*, vergeleken met de drie typen verkooppunten apart?
 b. Vergelijk ook de spreiding in gewicht bij alle verkooppunten met de spreiding bij de drie typen afzonderlijk. Conclusie?
29. Kijk nog eens naar de metingen van de controleur bij opgave 23. Die vissen zijn bij één type verkooppunt gekocht. Waar zal dat geweest zijn, op markten, bij viskarren of in viswinkels?

Gebruik werkblad 4.

30. In de kolom 'prijs per kg' (zie de tabel) staan ook rijtjes van drie getallen. Teken er een grafiek van.
31. In de kolom 'vaktechnisch' staan ook rijtjes van drie getallen.
- Wat stellen die getallen voor?
 - Wat voor een grafiek kun je daar bij maken?
32. Probeer uit te zoeken wat de grafiek hieronder voorstelt, en wat die te maken heeft met de tabel. (Tip: kijk naar de visjes in de viskommen).



Tijdsbesteding

In het Statistisch Zakboek van het CBS staan behalve grafieken ook veel tabellen.

Een tabel over de tijdsbesteding van Nederlanders per week:

† Gemiddelde tijdsbesteding per week van personen van 12 jaar en ouder, oktober 1985

	participanten			duur ¹		
	mannen	vrouwen	totaal	mannen	vrouwen	totaal
	%			uren:minuten		
Persoonlijke behoeften (eten, slapen e.d.)	100	100	100	74:11	76:28	75:27
Verzorging van kinderen	30	47	39	4:15	8:53	7:17
Huishoudelijk werk	89	99	94	7:13	24:02	16:57
Boodschappen	83	96	90	2:36	4:00	3:26
Klussen en karweitjes	68	46	56	4:51	2:31	3:48
Betaald werk	66	38	50	35:58	20:31	29:35
Onderwijs en vorming	40	41	41	16:42	13:34	14:57
Participatie en hulpverlening	43	57	51	4:26	3:41	3:58
Sport, hobby en spel	75	73	74	6:54	4:29	5:35
Mediagebruik	100	100	100	20:16	16:05	17:58
Uitgaan	78	75	77	5:04	4:01	4:30
Sociale relaties	96	99	97	10:13	12:34	11:32
Reizen	98	98	98	8:01	7:05	7:30
Steekproefaantal (abs.)	1 847	2 277	4 125			

Gemiddelde tijdsbesteding van personen die de activiteit hebben verricht (participanten).

33. Als je een tabel tegen komt, is belangrijk om te kijken wat de getallen in de tabel voorstellen. Je kunt een tabel immers alleen gebruiken als je hem ook begrijpt!
- In de eerste kolom staat onder andere het getal 83. Wat betekent dit getal?
 - Op sommige plaatsen staat het getal 100. Wat wil dit zeggen?
 - Wat betekent het getal 29:35 in de laatste kolom?
34. Hoeveel tijd *per dag* besteden we gemiddeld aan 'persoonlijke behoeften'?
35. a. Bij welke activiteiten vind je grote verschillen in tijdsbesteding tussen mannen en vrouwen?
b. Noem ook een paar activiteiten waarbij er géén verschillen zijn.

Een andere tabel over de tijdsbesteding van mensen:

	student/scholier	huisvrouw	'senior'	werknemer	tweeverdieners
Persoonlijke verzorging	11:17	11:00	11:41	10:15	10:05
Kinderverzorging	0:04	0:55	0:04	0:17	0:01
Huishoudelijk werk	0:35	4:25	2:47	1:35	1:06
Boodschappen doen	0:14	0:37	0:29	0:22	0:19
Betaald werk	0:20	0:07	0:09	4:36	5:27
Onderwijs, vorming	4:47	0:07	0:06	0:13	0:11
Reizen	1:15	0:52	0:47	1:08	1:09
Vrije tijd	5:28	5:57	7:57	5:34	5:42

36. Vergelijk de twee tabellen met elkaar en noem een paar verschillen.

37. a. Bedenk een vraag die je alléén met de eerste tabel kunt beantwoorden.
 b. Bedenk ook een vraag die je alléén met de tweede tabel kunt beantwoorden.

38. Tabellen kun je vaak verduidelijken door er een grafiek bij te maken.
 Als jij een grafiek zou moeten maken bij de kolom 'student/scholier' uit de tweede tabel, welk type grafiek zou je dan kiezen: lijngrafiek, cirkeldiagram, staafdiagram, of nog iets anders?

In de volgende opgaven lopen we nog eens precies langs hoe je een *cirkeldiagram* bij die kolom kunt maken.

Om te beginnen: als je ergens een grafiek van maakt, is het goed altijd eerst te controleren of de gegevens wel kunnen kloppen.

39. Ga na of het totaal aantal uren bij student/scholier inderdaad op 24 uren (het aantal uren in een etmaal) uitkomt.

Om het cirkeldiagram te kunnen maken, moeten de uren en minuten omgerekend worden in *procenten*. Een tabel is daarbij een handig hulpmiddel.

Gebruik werkblad 4.

40. De tabel hieronder staat ook op het werkblad, maak hem daar af.

	pers. verzorg.	kinder- verzorg.	huish. werk	bood- schappen	betaald werk	onderw. vorming	reizen	vrije tijd	totaal
uren:minuten	11:17	0:04	0:35	0:14	0:20	4:47	1:15	5:28	24:00
minuten	677								1440
procenten	47								100

41. Teken een cirkeldiagram bij deze tabel. Je kunt weer de 100%-cirkel gebruiken.
 Schrijf bij de taartpunten wat ze voorstellen, en zet ook een titel bij de grafiek.

Extra oefening:

42. Kies een andere kolom uit en maak daar ook een grafiek bij.

Hoofdstuk 2 - Het stam-blad diagram

Dit hoofdstuk gaat over het ordenen en verwerken van gegevens in een tabel. De tabellen die aan de orde komen zijn: turftabel, frekwentietabel en stam-bladdiagram. Turftabel en frekwentietabel ken je waarschijnlijk al, het stam-bladdiagram is nieuw.

Verjaardagen

De gegevens die we om te beginnen gaan bekijken zijn de verjaardagen van een groep kinderen. De verjaardagen zijn:

4 maart, 24 december, 8 september, 26 juni, 1 februari, 18 november, 22 januari, 14 april, 20 juni, 24 september, 30 september, 11 maart, 27 februari, 4 maart, 14 november, 2 oktober, 28 april, 3 januari, 15 maart, 30 november, 30 september, 1 juli, 15 juni, 12 april, 12 januari, 9 september, 5 maart, 13 juni.

1. a. In welke maand zijn er de meeste kinderen jarig?
- b. In welke maand is er niemand jarig?
- c. Zijn er ook kinderen op dezelfde dag jarig?

Het is niet moeilijk om deze vragen te beantwoorden, maar het is wel een heel gezocht in het lijstje met verjaardagen. Dat komt doordat de verjaardagen door elkaar staan. We noemen dit wel *ongeordende gegevens*. Een manier om de gegevens te ordenen, is het maken van een *turftabel*. Uit de turftabel kun je een *frekwentietabel* maken.

2. a. Neem de turftabel hieronder over in je schrift en maak hem af (een streepje stelt een verjaardag voor).
- b. Maak daarna de frekwentietabel.

Turftabel

<i>maand</i>	<i>aantal</i>
januari	
februari	/
maart	/
april	
mei	
juni	/
juli	
augustus	
september	/
oktober	
november	
december	/
totaal	

Frekwentietabel

<i>maand</i>	<i>aantal</i>
januari	
februari	
maart	
april	
mei	
juni	
juli	
augustus	
september	
oktober	
november	
december	
totaal	

3. a. Kijk nog eens naar vraag 1. Welke van deze vragen a, b en c kun je met alléén de frekwentietabel (zonder de oorspronkelijke gegevens) nog steeds beantwoorden?
- b. Welke vraag kun je niet meer beantwoorden?

Het stam-blad diagram

Bij het opstellen van de frekwentietabel gaat dus *informatie verloren*. De precieze verjaardagen kun je er immers niet meer uit aflezen.

Er bestaat een andere manier van 'turven' waarbij er géén informatie verloren gaat:

januari	22, 03, 12
februari	01, 27
maart	04, 11, 04, 15, 05
april	14, 28, 12
mei	
juni	26, 20, 15, 13
juli	01
augustus	
september	08, 24, 30, 30, 09
oktober	02
november	18, 14, 30
december	24

Dit heet wel een *stam-bladdiagram*. Dit woord wordt ook wel samengetrokken tot *stamgram*. De linker kolom met de maanden is de 'stam', de dagen zijn de 'bladeren'. Een rijtje bladeren heet een 'tak'.

4. a. Ga na dat in dit stam-bladdiagram inderdaad precies alle verjaardagen van de kinderen staan.
 - b. Wat is de 'tak' van maart?
5. Hieronder staat nog een stam-bladdiagram van de verjaardagen. Wat is het verschil met het vorige diagram?

januari	03, 12, 22
februari	01, 27
maart	04, 04, 05, 11, 15
april	12, 14, 28
mei	
juni	13, 15, 20, 26
juli	01
augustus	
september	08, 09, 24, 30, 30
oktober	02
november	14, 18, 30
december	24

Televisiespelletje

Bij een televisiespelletje moeten de deelnemers de prijs van een koffiezetapparaat raden. De mensen die meedoen schatten de bedragen:

98, 75, 109, 89, 95, 64, 119, 100, 89, 82, 69, 95, 90, 105, 120

De stam van het bijbehorende stam-bladdiagram met alvast één blad:

6	
7	
8	
9	8
10	
11	
12	

- 6 a. Wat stellen de getallen in de stam hier voor?
- b. Maak het volledige stambladdiagram, met de getallen in volgorde van grootte.

Proefwerkcijfers

Een stam-bladdiagram van proefwerkcijfers. Links staan de cijfers voor Wiskunde, rechts voor Engels.

<i>Wiskunde</i>		<i>Engels</i>
$\frac{1}{2}$ 0 -	4	
$\frac{1}{2}$ $\frac{1}{2}$ + 0 0	5	- +
$\frac{1}{2}$ + 0 - -	6	- - 0 + + $\frac{1}{2}$
$\frac{1}{2}$ $\frac{1}{2}$ 0	7	- 0 0 + $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$
$\frac{1}{2}$ $\frac{1}{2}$ - -	8	0 + + $\frac{1}{2}$ $\frac{1}{2}$
+ 0	9	+
0	10	

7. a. Wat zal 6 | + bij Engels betekenen?
- b. Wat was het hoogste cijfer bij wiskunde?
8. a. Hoeveel leerlingen hebben een onvoldoende voor Wiskunde?
- b. En hoeveel voor Engels?
- c. Hoeveel leerlingen hebben minstens een acht voor wiskunde?
9. Bij een van beide proefwerken was een leerling ziek. Was dat bij wiskunde of bij Engels?
10. Vergelijk de proefwerkcijfers voor wiskunde en Engels eens met elkaar. Voor welk vak zijn de resultaten het beste, volgens jou?

Vulkanen

Het idee van het stam-bladdiagram is eigenlijk heel eenvoudig. Toch is het gebruik van dit diagram pas iets van de laatste tijd. Het eerste stam-bladdiagram dat gepubliceerd is, staat hieronder. Het brengt de hoogte van vulkanen in beeld. De 'voet' is een maat die veel in Engels sprekende landen wordt gebruikt.

0 / 9 = 900 feet

0		98766562
1		97719630
2		69987766544422211009850
3		876655412099551426
4		9998844331929433361107
5		97666666554422210097731
6		898665441077761065
7		98855431100652108073
8		653322122937
9		377655421000493
10		0984433165212
11		4963201631
12		45421164
13		47830
14		00
15		676
16		52
17		92
18		5
19		39730

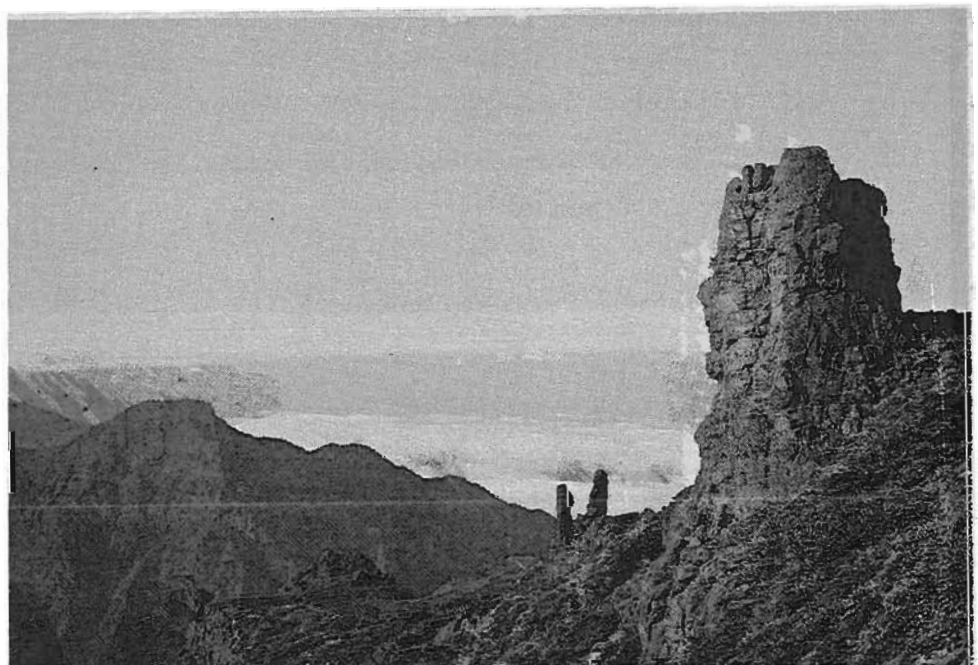
Stem-and-leaf displays:
heights of 218 volcanoes, unit 100 feet.

19 / 3 = 19,300 feet

11. a. Weet je hoeveel centimeter een 'voet' is?
b. Bij het diagram staat als uitleg: $19 / 3 = 19.300$ voet. Wat wordt hier mee bedoeld?

12. a. Hoeveel vulkanen van 4300 voet zijn er?
b. Hoe hoog is de hoogste vulkaan?
c. En de laagste?
d. Hoeveel vulkanen zijn er tussen 10.000 en 12.000 voet?

*De vulkaan
Caldera de
Taburiente,
La Palma.*



13. Om een idee te krijgen van de hoogte in meters van de vulkanen, is een omrekeningstabel handig. Neem de tabel over in je schrift en maak hem af:

hoogte in voeten	1000	2000	3000	4000	5000	10.000	15.000	20.000
hoogte in meters	315							

Bij de volgende opgaven kijken we alleen naar de vulkanen van 10.000 voet en hoger, het zijn er 51:

10	0984433165212
11	4963201631
12	45421164
13	47830
14	00
15	676
16	52
17	92
18	5
19	39730

14. Maak een nieuw stam-bladdiagram, met daarin deze vulkanen precies op volgorde van grootte:

10	0112233445689
11
12
..	
..	

15. Schat met behulp van dit stam-bladdiagram de gemiddelde hoogte van deze 51 vulkanen.
16. Bereken de gemiddelde hoogte van deze 51 vulkanen.
(Tip: met z'n tweeën kun je het rekenwerk verdelen.)

Het *gemiddelde* is een maat voor het 'midden' van een serie waarnemingen. Er bestaan ook andere maten, bijvoorbeeld de *mediaan*. De mediaan is de middelste waarneming uit een serie.

17. Wat is de mediaan van de 51 vulkanen?
18. Vergelijk gemiddelde en mediaan met elkaar. Kun je het verschil verklaren?

Lengtes meten

Met de hele klas:

19. a. Verzamel de lengtes van alle leerlingen uit jouw klas.
b. Maak een stam-bladdiagram van deze lengtes.
c. Wat is de mediaan?
d. En wat het gemiddelde?
20. Maak nog een stambladdiagram, maar nu voor jongens en meisjes apart.
Je kunt de stam in het midden zetten, en dan links de jongens en rechts de meisjes (zoals het is gedaan bij de 'Proefwerkcijfers').
21. Vergelijk de lengtes van jongens en meisjes met elkaar. Zijn er verschillen?

Misschien kwam je bij opgave 20 tot de ontdekking dat je bij het verzamelen van de lengtes geen onderscheid hebt gemaakt tussen jongens en meisjes. Gevolg: je moet de lengtes opnieuw verzamelen, nu met de aanduiding voor het geslacht erbij. Als je gegevens verzamelt, is het belangrijk van te voren precies te bedenken wat je wilt weten. Anders verzamel je óf te veel, óf te weinig gegevens.

Dienstregeling

22. Maak twee niet al te makkelijke proefwerkvragen bij deze dienstregeling.

Från HAKBERGET
via Radiohuset - Linnégatan - Stureplan - Kungsboplan
till HORNSBERG

MÅNDAG - FREDAG	
0	05
5	23 41 59
6	01 19 37 55
7	08 19 29 40 50
8	01 11 22 32 43 53
9	04 20 36 52
10	08 24 39 57
11	13 24 33 42 51
12	00 09 18 27 36 45 54
13	03 12 21 30 39 48 57
14	06 15 24 33 42 51
15	00 09 18 27 36 45 53
16	02 12 21 31 40 50 59
17	10 20 31 42 53
18	09 25 43
19	00 15 30
20	} 00 30
21	
22	
23	



42

Dienstregeling van een stadsbus in Stockholm. De dienstregeling heeft dezelfde vorm als een stam-bladdiagram.

Hoofdstuk 3 - Beschrijven en voorspellen

Dit hoofdstuk gaat over de groei van de wereldbevolking. Je komt grafieken tegen die daar een *beschrijving* van geven en grafieken die een *voorspelling* in beeld brengen. Bij voorspellen probeer je op grond van gegevens uit het verleden iets te zeggen over de toekomst.

De wereldbevolking

Hieronder staat een tabel van de wereldbevolking in miljarden.

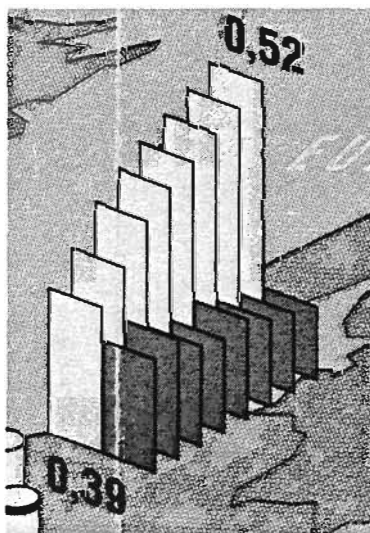
jaar	± 1630	± 1820	± 1890	1930	1950	1960	1968	1975	1981	1987
wereldbevolking in miljarden	$\frac{1}{2}$	1	$1\frac{1}{2}$	2	$2\frac{1}{2}$	3	$3\frac{1}{2}$	4	$4\frac{1}{2}$	5

- a. Zet de gegevens uit deze tabel in een grafiek.
b. Probeer op grond van deze gegevens de wereldbevolking in het jaar 2000 te voorspellen.
c. Kun je ook een voorspelling doen voor het jaar 2100? Waarom wel/niet?

Voor de opgaven die nu volgen heb je de 'verstedelijkingskaart' nodig die achterin het boek zit. De kaart is afkomstig van de Verenigde Naties.

- a. Wat betekent 'verstedelijking'?
b. Op de kaart komen twee typen grafieken voor. Welke zijn dat en wat is het nut van deze grafieken?

Op de verstedelijkingskaart staat in Europa de volgende grafiek getekend:



- a. Wat stellen de witte staven voor? En de grijze?
b. Wat betekenen de getallen 0,39 en 0,52?

4. a. Hoeveel mensen woonden er in Europa in 1950 in de stad en hoeveel op het platteland?
b. Wat is de voorspelling voor 2025?
5. Hoe zal men aan de gegevens voor de jaren 1950 - 1980 gekomen zijn? En aan de gegevens voor de jaren daarna (van 1990 tot 2025)?
6. Vul met behulp van de kaart de volgende tabel in:

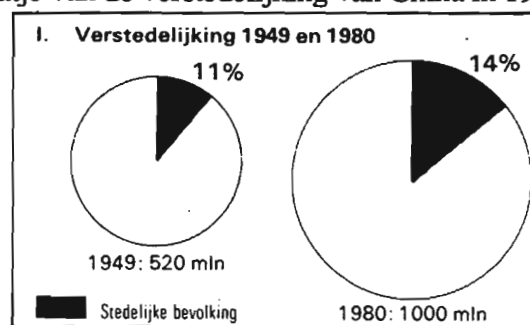
werelddeel	inwoners in 1950	inwoners in 2025 (verwachte aantal)
Noord-Amerika		
Latijns-Amerika		
Europa		
Afrika		
Azië		
Oceanië		
totaal		

7. In welk werelddeel neemt de bevolking het hardste toe?

De verstedelijking in China

Er komen steeds meer mensen in de grote steden te wonen. Je hebt dat kunnen zien op de verstedelijkingskaart.

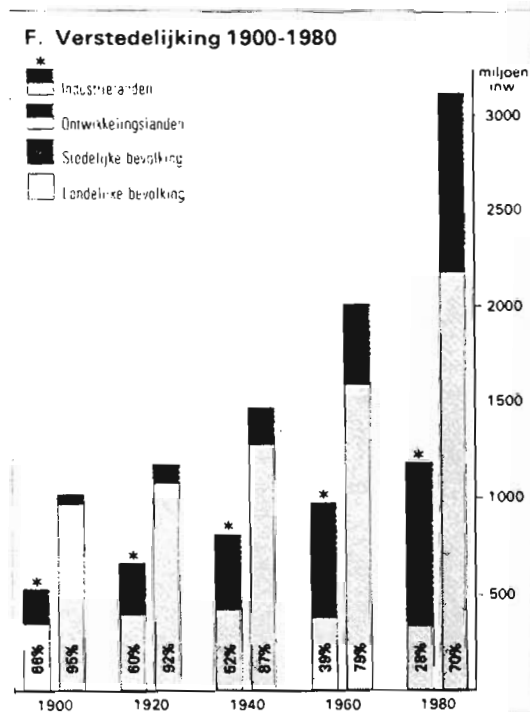
Een plaatje van de verstedelijking van China in 1949 en in 1980:



8. Hoeveel Chinezen woonden er in 1949 in de steden? En in 1980?
9. Waarom is het rechter cirkelcirdiagram groter getekend dan het linker?

Verstedelijking over de hele wereld

De verstedelijking is al heel lang aan de gang. Een grafiek van de verstedelijking van de hele wereld voor de periode van 1900 tot 1980:



10. Geef commentaar op de volgende bewering:
 "In 1980 woonden er in de ontwikkelingslanden net zoveel mensen in de steden als in de industrielanden"

Bij de volgende opgaven heb je de verstedelijkingskaart weer nodig.

11. Op de verstedelijkingskaart staat geschreven:
 "Vanaf het jaar 2015 woont meer dan de helft van de wereldbevolking in de steden"
 Klopt deze bewering met de grafieken op de kaart?
12. Kun je met behulp van de gegevens uit de verstedelijkingskaart het staafdiagram hierboven uitbreiden met staven voor 1950? Doe het ook voor 2025.

De verstedelijkingskaart geeft ook informatie over grote steden. Het plaatje van Mexico City staat hiernaast nog eens afgedrukt.

Mexico City 25

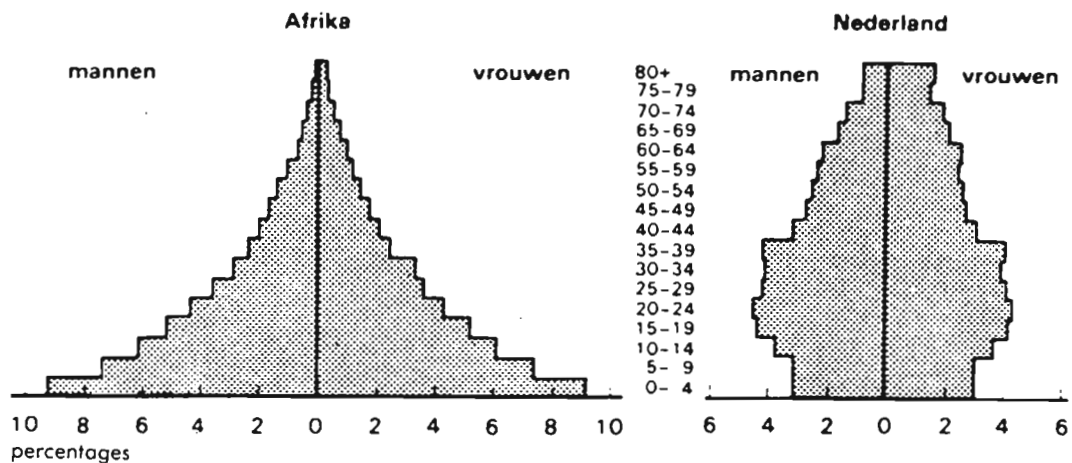


13. Wat kun je uit dit plaatje aflezen?

14. Er zijn steden waarvan het aantal inwoners van 1985 tot 2000 ongeveer verdubbelt. Welke steden zijn dat?
15. Maak een top-tien van de grootste steden in het jaar 2000.

Bevolkingspiramides

De bevolking van een land kun je tekenen in de vorm van een *bevolkingspiramide*. Hieronder zie je twee bevolkingspiramides, één voor Afrika en één voor Nederland. De grafieken zijn uit 1986.

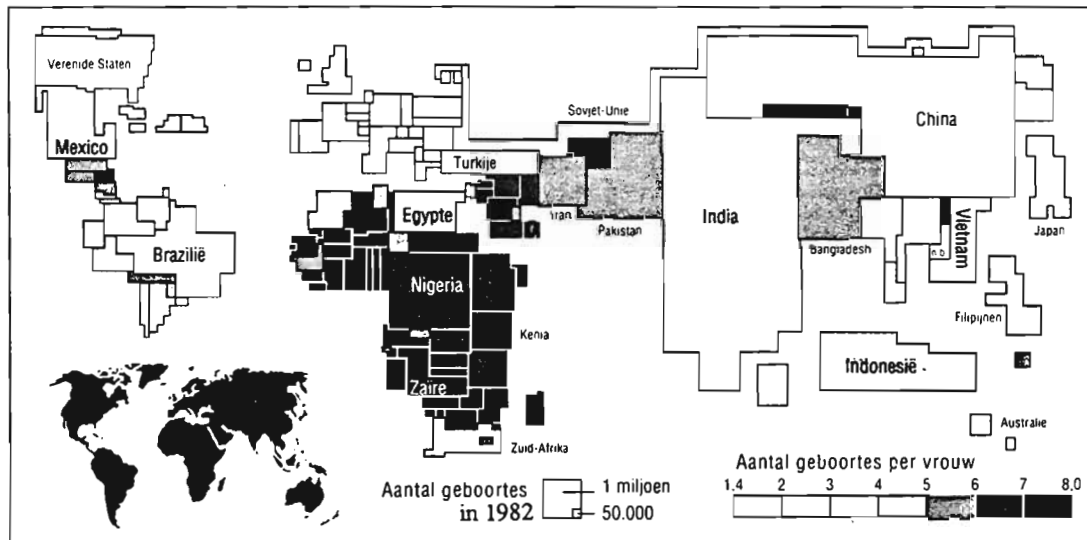


16. Waarom zouden deze grafieken *bevolkingspiramides* heten? Is het eigenlijk wel zo'n goede naam?
17. De grafiek voor Afrika ziet er heel anders uit dan die voor Nederland. Kun je een paar verschillen noemen?
18. Hoeveel procent van de Afrikaanse bevolking is jonger dan 20 jaar? En hoe zit dat voor Nederland?
19. Welk percentage van de bevolking is 65+? Zoek het uit voor Nederland en voor Afrika.
20. Wie worden er het oudst, mannen of vrouwen? Vergelijk weer Nederland met Afrika.
21. Hoe zullen de piramides voor het jaar 2000 eruit zien? Probeer ze te schetsen.

Geboortes per land

Bij de kaart hieronder is de oppervlakte van de landen bepaald door het aantal geboortes per jaar in dat land. Links onder in de hoek staat een 'gewone' kaart.

AANTALLEN GEBOORTES PER LAND EN PER VROUW (1985)



De oppervlakte van de landen wordt bepaald door het aantal geboortes per jaar. De tint geeft aan het gemiddeld aantal geboortes per vrouw

22. Noem een paar landen die in verhouding heel groot zijn getekend. Wat betekent dat?
23. Welke landen zijn in verhouding heel klein getekend? Noem er een paar.
24. Wat betekenen de kleuren (wit-grijs-zwart) op de kaart?

In dit hoofdstuk heb je gezien dat er allerlei verschillende grafieken gebruikt worden om iets te laten zien over de (wereld)bevolking. Afhankelijk van wat benadrukt moet worden, wordt een type grafiek gekozen. Je hebt voorbeelden gezien van:

- grafieken om landen (of werelddelen) met elkaar te vergelijken,
- grafieken die de groei van de wereldbevolking laten zien,
- grafieken die de verstedelijking uitbeelden,
- grafieken die de leeftijdsopbouw van een land aangeven,
- een grafiek die iets zegt over de geboortes per land.

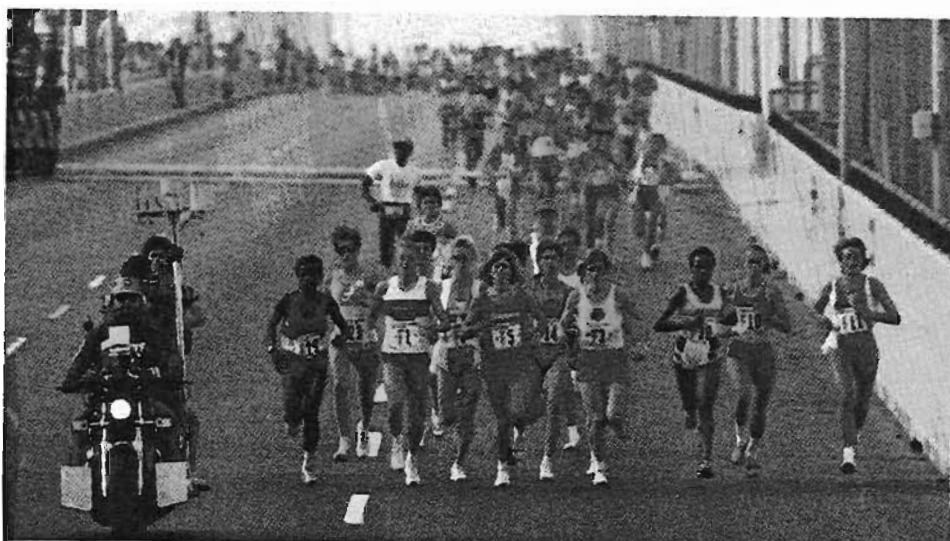
In de meeste landen wordt wel iets van bevolkingsstatistiek bij gehouden. Het is voor een land immers belangrijk om te weten hoe de bevolking zich in de toekomst zal ontwikkelen. Als het op *voorspellen* van de toekomstige bevolking aankomt, zijn naast de geboortecijfers van een land ook de sterftcijfers erg belangrijk.

Hoofdstuk 4 - Het verwerken van gegevens

Statistiek gaat over het verwerken van gegevens en het trekken van conclusies daaruit. In dit hoofdstuk herhalen we enkele bekende begrippen zoals gemiddelde, modus en mediaan. Ook kom je de frequentietabel en het histogram weer tegen. Nieuw is de *puntenband*.

Marathonloop

Een marathon is een hardloopwedstrijd over ruim 42 km. Om die te kunnen lopen moet je goed getraind zijn!



- 1 Hoe lang zou jij doen over 42 km lopen? En op de fiets?

Bij een bepaalde marathon hebben de deelnemers de volgende tijden gelopen:

2.26	2.31	2.33	2.37	2.44	2.48	3.01	3.03	3.07	3.09
3.15	3.15	3.16	3.18	3.20	3.25	3.29	3.35	3.45	3.45
3.47	3.50	3.56	3.57	3.58	3.59	4.05	4.09	4.15	4.15
4.25	4.25	4.27	4.29	4.33	4.37	4.42	4.42	4.42	5.01
5.09	5.17	5.31	5.46	6.32					

- 2a Hoeveel deelnemers waren er?
- b De winnaar liep de marathon in 2 uur en 26 minuten. Wat was zijn gemiddelde snelheid?

- c Hoe lang deed de hekkeluiters (= degene die het laatste aan kwam) over de marathon?
Hoeveel km per uur is dat gemiddeld?

- 3 Op een zeker moment is de helft van de lopers binnen.
a Welke tijd heeft de loper die op dat moment over de finish gaat?

Waarnemingen worden vaak geordend van klein naar groot. Bij de looptijden is de volgorde van binnenkomst de ordening. De middelste waarneming heeft een speciale naam, deze wordt de *mediaan* genoemd.

- 4 Wat is de mediaan van de marathontijden?
- 5 Er waren bij deze wedstrijd zes deelnemers van de sportclub Grande Vitesse. Hun tijden waren:
2.37 3.03 3.25 3.58 4.37 5.17
- a Wat was de gemiddelde tijd van deze zes lopers?

Bij een oneven aantal waarnemingen is de *mediaan* de middelste waarneming. Als het aantal waarnemingen even is, zoals bij lopers van Grande Vitesse, is dit geen goede afspraak. Er is dan immers niet één middelste waarneming. In dat geval is de afspraak om naar de middelste *twee* waarnemingen te kijken. De mediaan is dan het gemiddelde van die twee getallen.

- b Wat is de mediaan van de tijden van de lopers van Grande Vitesse?
- 6 De gemiddelde looptijd van *alle* 45 deelnemers aan de marathon is 3 uur en 59 minuten. Ga dat na met een berekening en probeer die overzichtelijk op te schrijven.
TIP: Doe deze opgave niet alleen, maar met z'n tweeën of zelfs met z'n vieren. Je kunt dan het werk verdelen. Bedenk voordat je aan het rekenen slaat eerst hoe je het aan wilt pakken!

Mediaan en gemiddelde worden wel *centrummaten* genoemd. Ze geven iets aan over het centrum (midden) van een serie waarnemingen.

Er is nog een derde centrummaat, de *modus*. Dat is de waarneming die het vaakst voorkomt.

- 7 Wat is de modus van de marathontijden?
Zegt die inderdaad iets over het centrum? Waarom wel/niet?

Frequentietabel en histogram

In de vorige paragraaf is de uitslag van een marathonwedstrijd getypeerd met behulp van centrummaten. Maar er zijn meer manieren om naar de uitslag te kijken.

Je kunt de gegevens van de marathonwedstrijd ook verwerken in een *frequentietabel*.

Daar kun je dan vervolgens een *histogram* bij tekenen.

De waarnemingen worden dan ingedeeld in *klassen*. Dat heet het maken van een *klasse-indeling*.

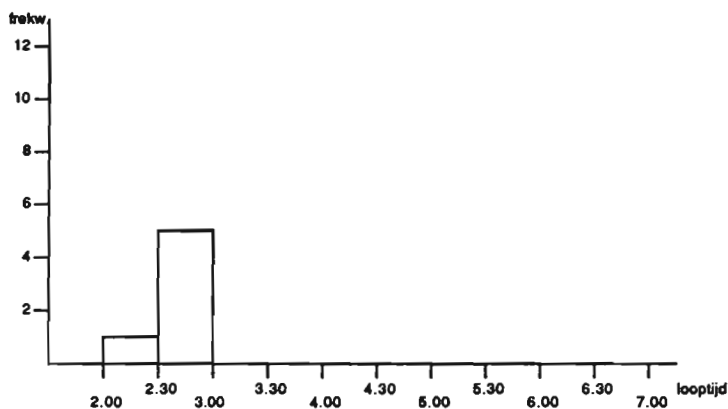
- 8a Hoeveel klassen krijg je als de klassebreedte 30 minuten is en de laagste klasse begint bij 2.00 uur?
- b Maak (in je schrift) de frequentietabel af:

<u>tijd</u>	<u>aantal lopers</u>
2.00 - 2.29	1
2.30 - 2.59	5
....	

- c In welke klasse zitten de meeste lopers?

Bij een frequentieverdeling heet de klasse met het hoogste aantal waarnemingen de *modale klasse*. Als je de frequentietabel eenmaal hebt, kun je er een histogram bij tekenen.

Hieronder is alvast een begin gemaakt:

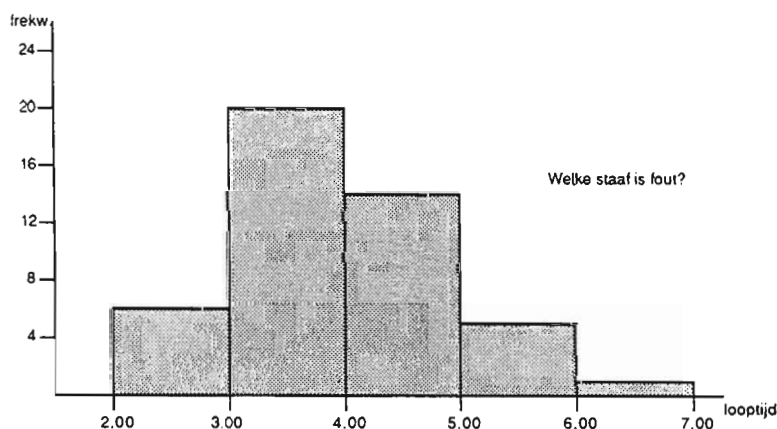


- 9a Maak het histogram verder af.
- b Hoe lees je uit het histogram de modale klasse af?
- 10a Geef met een kruisje op de horizontale as aan waar de mediaan ligt. Teken ook een verticale lijn door dit punt. Het histogram is nu in twee stukken verdeeld.
- b Wat kun je zeggen over de grootte van deze twee stukken?

Meer histogrammen

Bij de tabel en het histogram is een klassebreedte van 30 minuten als uitgangspunt genomen. De laagste klasse begon bij 2.00 uur. Dit kan ook anders.

- 11 Neem een klassebreedte van 30 min. en laat de laagste klasse beginnen bij 2.15 uur.
- Hoeveel klassen krijg je dan?
 - Maak de frequentietabel en teken het histogram.
 - Wat is nu de modale klasse?
 - Wijkt dit histogram erg af van het histogram van opgave 9?
- 12 Het histogram hieronder hoort bij de marathontijden, maar nu is een klassebreedte van 60 min genomen. Helaas heeft één staaf de verkeerde hoogte gekregen. Welke staaf is dat? Verbeter die.



- 13 Vergelijk de histogrammen van opgave 9 en opgave 12 met elkaar. De twee histogrammen hebben op de horizontale as dezelfde schaalverdeling. De schaalverdeling op de verticale as is echter *niet* hetzelfde.
- Wat is het verschil? Waarom zou dat gedaan zijn?
 - Welk histogram vind je duidelijker, het histogram met klassebreedte van 30 min of het histogram met een klassebreedte van 60 min? Waarom?

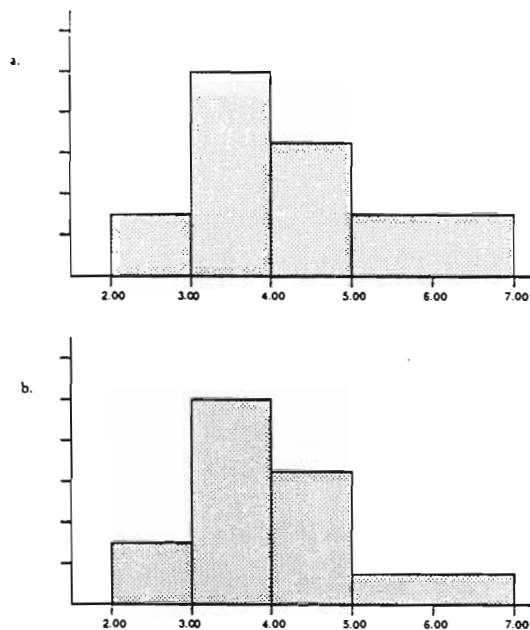
Bij één serie waarnemingen zijn dus verschillende histogrammen mogelijk.

Als je een histogram tekent, moet je eerst beslissingen nemen over:

- de klassebreedte, rekening houdend met het aantal klassen dat je dan krijgt,*
- de klassengrenzen, in het bijzonder de ondergrens van de eerste klasse.*

Hoe meer klassen het histogram heeft, hoe gedetailleerder informatie je eruit kunt aflezen. Vanwege de overzichtelijkheid zal een histogram meestal tussen de 5 en de 10 klassen hebben. Een enkele keer kan het aantal klassen groter zijn, maar meer dan 20 klassen zul je niet vaak tegen komen.

Soms kom je wel eens een histogram tegen waarbij de klassen niet allemaal even breed zijn. Hieronder staan er twee, weer bij de marathontijden.



- 14a Vergelijk de twee histogrammen met elkaar. Wat is het verschil?
 b Schrijf in de staven van de histogrammen het bijbehorende aantal waarnemingen.
 c Welk histogram vind je het beste? Waarom?

Bij histogrammen met ongelijke klassebreedtes is de afspraak dat de *deoppervlakte* van de staven de maat is voor het aantal waarnemingen. Het is dan niet meer mogelijk frequenties bij de verticale as te zetten. Je kunt de frequenties wel noteren in of vlak boven de staven.

- d Welk histogram is volgens deze afspraak het goede? Had jij dat ook bij c?

Alle punten op een rij

Een goed gekozen histogram geeft een duidelijk beeld van een serie waarnemingen.

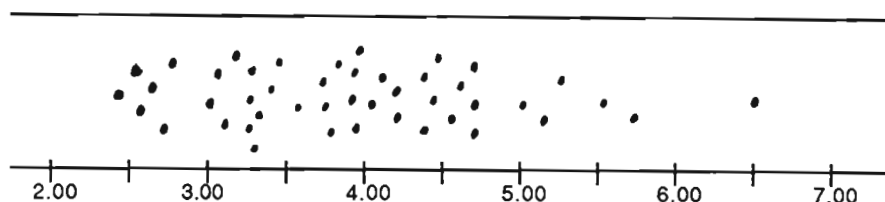
Maar: de losse waarnemingen zijn er niet meer precies in terug te vinden.

In deze paragraaf komt een plaatje aan bod waarin wél alle waarnemingen afzonderlijk staan.

Nogmaals de tijden van de 45 marathonlopers:

2.26	2.31	2.33	2.37	2.44	2.48	3.01	3.03	3.07	3.09
3.15	3.15	3.16	3.18	3.20	3.25	3.29	3.35	3.45	3.45
3.47	3.50	3.56	3.57	3.58	3.59	4.05	4.09	4.15	4.15
4.25	4.25	4.27	4.29	4.33	4.37	4.42	4.42	4.42	5.01
5.09	5.17	5.31	5.46	6.32					

In het volgende plaatje staan alle looptijden van de marathonwedstrijd ingetekend:



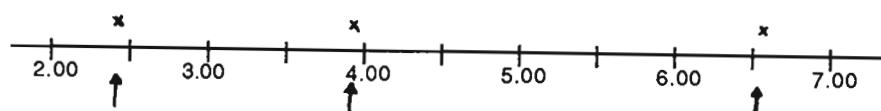
Zo'n plaatje heet wel een *puntenband*.

De punten zijn op verschillende hoogtes getekend om ze niet over elkaar heen te laten vallen. Verder heeft die hoogte geen betekenis.

- 15a Zoek in de puntenband de snelste, de langzaamste en de middelste loper op. Zet een rondje om hun looptijd.
- b Hoeveel lopers waren er volgens het plaatje tussen de drie en vier uur onderweg?
Klopt je antwoord met de oorspronkelijke gegevens?
- 16a Na hoeveel tijd is een kwart van de lopers binnen?
- b Na hoeveel tijd is driekwart van de lopers binnen?

De drie omcirkelde tijden (laagste waarneming, hoogste waarneming en mediaan) geven samen al een aardig beeld van de wedstrijd.

Je kunt er een plaatje van maken door ze op een 'looptijdenlijn' te plaatsen, zo:



- 17 Iemand zegt: "Hè? De middelste waarneming moet toch midden tussen de kleinste en de grootste liggen?"
Probeer uit te leggen hoe dit zit.

Oefenopgave

- 18 Een rijtje proefwerkcijfers:
6.8, 8.2, 5.5, 7.4, 6.5, 5.2, 7.3, 9.1, 8.4, 4.9, 6.7,
7.4, 6.6, 6.3, 7.8, 9.6, 4.1, 5.3, 7.5, 6.6, 5.7, 4.5
- a Teken een puntenband van de proefwerkcijfers.
- b Zet de cijfers op volgorde van grootte.
- c Teken het laagste cijfer, het hoogste cijfer en de mediaan op een getallenlijn.
- d Bereken het gemiddelde. Is het gemiddelde groter of kleiner dan de mediaan?

Hoofdstuk 5 - De boxplot

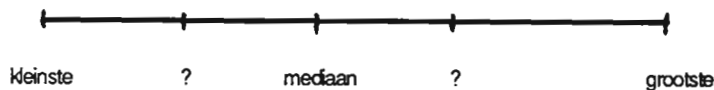
In de statistiek zoekt men naar simpele plaatjes die veel informatie geven.

Een puntenband geeft veel informatie, maar is nogal bewerkelijk om te tekenen.

Het lijnstukje kleinste-mediaan-grootste is snel getekend, maar bevat minder informatie.

Een nieuw plaatje in dit hoofdstuk is de *boxplot*.

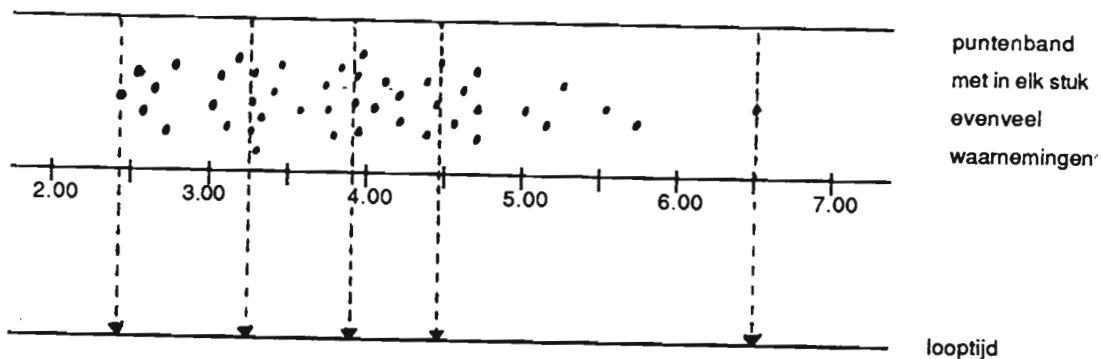
Als uitbreiding op het lijnstukje met drie punten, is een plaatje met vijf punten bedacht. Om te beginnen het drietal kleinste-mediaan-grootste, en dan nog twee andere. Zo iets:



19 Heb je al een vermoeden welke die andere twee punten zullen zijn?

De afspraak is om de reeks waarnemingen in vier stukken te verdelen, met in elk stuk evenveel getallen. We nemen de puntenband van de marathonlooptijden uit het vorige hoofdstuk om te kijken hoe dit werkt.

- *deel de puntenband in vieren, met in elk stuk evenveel waarnemingen:*



- *markeer de vijf getallen die je zo vindt op de getallenlijn.*

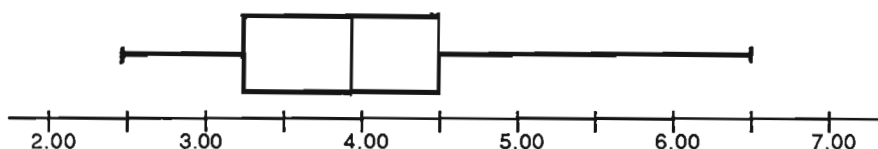
Er zijn dus vijf getallen van belang:

- de snelste tijd : 2:26 (kleinste)
- de tijd waarop een kwart binnen is : 3:15 (eerste kwartiel)
- de middelste tijd : 3:56 (de mediaan)
- de tijd waarop driekwart binnen is : 4:29 (derde kwartiel)
- de langzaamste tijd : 6:32 (grootste)

20a Je ziet hier boven (in de rechter kolom) een nieuw woord staan. Welk woord is dat en wat zal het betekenen?

b Het tweede kwartiel bestaat ook. Wat zal dat zijn?

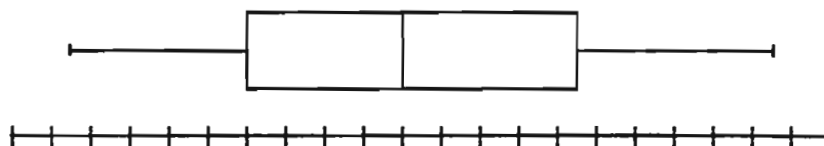
Niet vergissen: een *kwartiel* is één getal, dus niet een interval of zo.
 Bij de vijf getallen wordt de volgende tekening gemaakt:



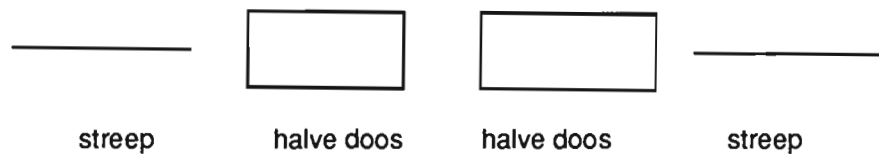
Zo'n tekening heet een *box-plot*. (De getallenlijn hoort er ook bij, anders kun je de box-plot immers niet aflezen.)

- 21 Aan de marathonwedstrijd deden 45 lopers mee.
- Wat kun je als je alleen de box-plot hebt, zeggen over de looptijd van de loper die als 15e binnenkwam?
 - Verzin een uitspraak met: '50% van de lopers heeft een tijd gelopen'
- 22a Kijk nog eens terug naar vraag 15a. Kun je die vraag beantwoorden met alléén de box-plot?
- Kun je vraag 15b beantwoorden met alléén de box-plot?
 - Kun je vraag 16 beantwoorden met alléén de box-plot?

De algemene vorm van een box-plot is:



Een box-plot bestaat eigenlijk uit vier stukken:

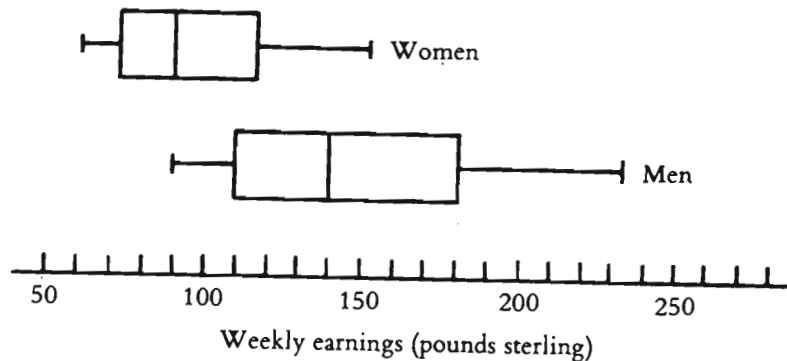


Elk stuk bevat een kwart (25%) van de waarnemingen. In de doos in het midden zit dus 50% van de waarnemingen.

- 23 Teken een box-plot bij de getallen:
 kleinste 12, eerste kwartiel 15, mediaan 21, derde kwartiel 23, grootste 40.

Salarissen

Box-plots worden vooral gebruikt om twee of meer *groepen met elkaar te kunnen vergelijken*. Bijvoorbeeld de inkomens van werkende mannen en vrouwen in Engeland:



De bedragen zijn per week en in pound sterling (Engelse ponden).
In de krant kun je opzoeken hoeveel een Engels pond ongeveer waard is.

- 24 Vul de juiste getallen in:
- 'Een Engelse vrouw verdient maximaal pond per week.'
 - 'Een Engelse man verdient ten minste pond per week.'
- 25 Kijk naar de middelste 50% van beide groepen. Wat zijn de salariegrenzen van die groepen?
- 26 Hieronder staan drie beweringen. Geef op elke bewering commentaar, uitgaande van de box-plots die hierboven gegeven zijn.
- Bewering A: 'alle vrouwen verdienen in Engeland minder dan de mannen'
 - Bewering B: 'vrouwen verdienen in Engeland in het algemeen minder dan mannen'
 - Bewering C: 'vrouwen worden slechter betaald dan mannen'
- 27 Kloppen de volgende uitspraken met de box-plot?
- '50% van de mannen verdient meer dan het maximum weeksalaris van de vrouwen.'
 - 'vrijwel alle mannen verdienen meer dan de 50% laagst betaalde vrouwen.'
- 28 Kun je zelf nog twee uitspraken bij de box-plots verzinnen?

Kwartielen bepalen

Nogmaals het rijtje proefwerkcijfers (zie opgave 18), nu op volgorde gezet:

4.1, 4.5, 4.9, 5.2, 5.3, 5.5, 5.7, 6.3, 6.5, 6.6, 6.6,
6.7, 6.8, 7.3, 7.4, 7.4, 7.5, 7.8, 8.2, 8.4, 9.1, 9.6

Je hebt hier al een puntenband van getekend. Nu richten we ons op het tekenen van de box-plot bij deze cijfers. Daarvoor heb je de vijf box-plot-getallen nodig. Drie getallen heb je zo te pakken: kleinste 4.1; mediaan 6.65; grootste 9.6

Blijft over het bepalen van 1e en 3e kwartiel. Die kun je uit de gegevens vinden door te bedenken:

- eerste kwartiel = de mediaan van de eerste helft
- derde kwartiel = de mediaan van de tweede helft

eerste helft 4.1, 4.5, 4.9, 5.2, 5.3, 5.5, 5.7, 6.3, 6.5, 6.6, 6.6,

tweede helft 6.7, 6.8, 7.3, 7.4, 7.4, 7.5, 7.8, 8.2, 8.4, 9.1, 9.6

29a Bepaal de kwartielen.

b Teken de box-plot.

c Probeer drie uitspraken te verzinnen met:

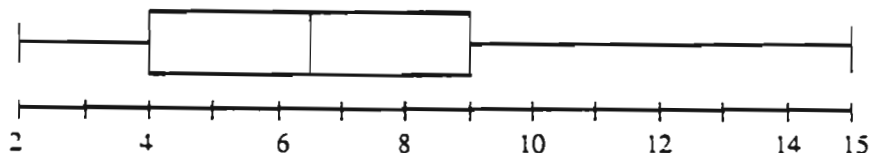
'50% van de leerlingen heeft een proefwerkcijfer

30 Een rijtje gegevens:

15, 7, 11, 3, 9, 3, 10, 5, 2, 7, 3, 8, 6, 4, 6, 7, 5, 2

Wat is het eerste kwartiel, de mediaan, het derde kwartiel?

31 Kan de onderstaande box-plot bij de gegevens van opgave 30 horen?



32 Verzin een rijtje getallen waarvoor deze box-plot goed is.

Temperaturen vergelijken

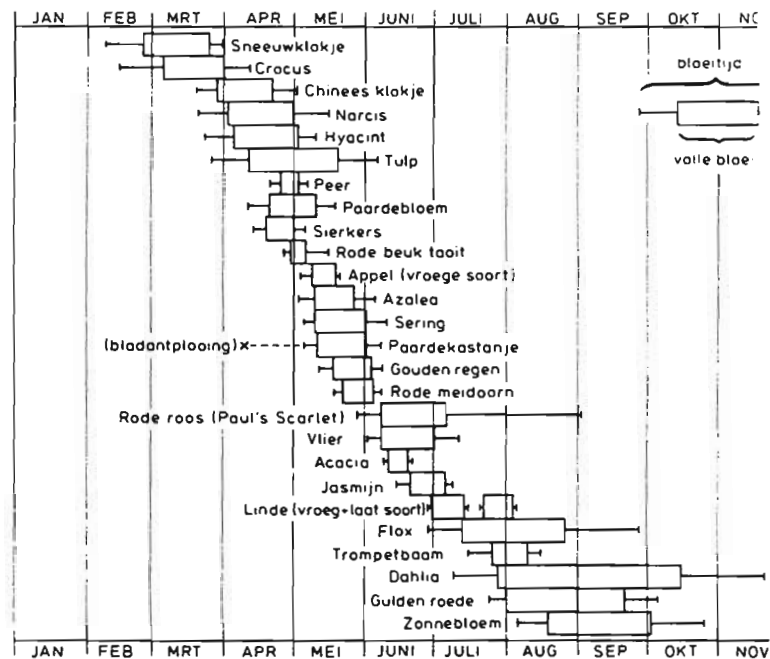
In twee opeenvolgende jaren registreerde het KNMI de volgende temperaturen in de maand mei:

dagnr	temp '82	temp '83	dagnr	temp '82	temp '83
1	10.1	15.7	17	23.6	16.3
2	11.6	12.5	18	19.4	16.5
3	13.5	10.1	19	19.3	15.8
4	12.6	13.8	20	17.9	17.0
5	11.1	15.8	21	20.2	12.7
6	9.1	17.5	22	18.1	15.0
7	10.6	16.0	23	15.5	16.9
8	11.3	14.7	24	16.9	11.3
9	14.4	15.9	25	20.0	10.8
10	16.0	14.2	26	26.3	10.1
11	17.2	11.1	27	24.9	10.8
12	18.8	13.4	28	18.0	12.4
13	19.6	16.0	29	20.5	12.3
14	22.8	16.2	30	23.8	16.2
15	25.6	16.1	31	27.2	22.8
16	24.3	13.1			

33 Probeer zoveel mogelijk verschillende manieren te bedenken om de twee meimaanden met elkaar te vergelijken en maak daar een mooie bladzijde in je schrift van.

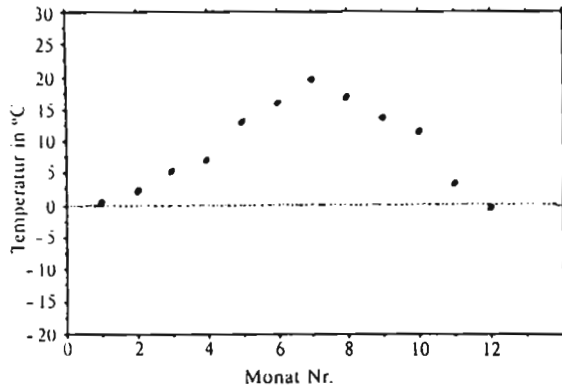
Zijn dit ook box-plots?

Fig. 67. Gemiddelde bloeiperioden van diverse plantesoorten voor De Bilt, zoals waargenomen door G. W. Th. M. de Bont over een periode van 25 jaar. Van jaar op jaar kunnen de bloeidata twee weken eerder of later vallen dan in de grafiek aangegeven. In het algemeen voltrekt de bloei zich wel ieder jaar in dezelfde volgorde.



Een heel jaar temperaturen

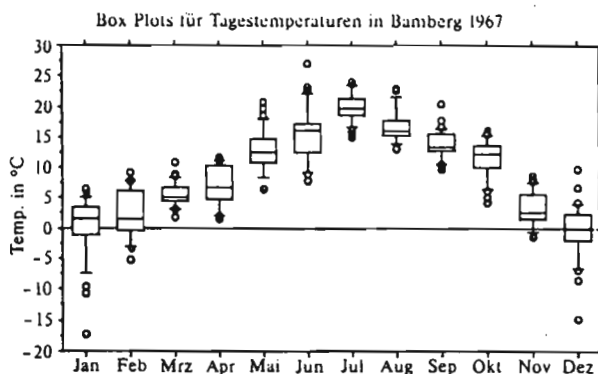
Als je een heel jaar lang elke dag de temperatuur meet, kun je daar een grafiek van tekenen. Dat kan op verschillende manieren. Hieronder zie je een grafiek van de gemiddelde maandtemperatuur in 1967 in het Duitse stadje Bamberg.



Tagestemperaturen in Bamberg 1967

- 34a Welke maand was het (volgens deze grafiek) het warmst? Hoe warm was het toen?
 b En welke maand was het het koudst?
 c Kun je iets zeggen over de warmste dag van het jaar? En over de koudste dag?
 d Kun je iets zeggen over temperatuurverschillen binnen één maand?

De grafiek hieronder gaat ook over de temperatuur in Bamberg. Nu heeft men voor elke maand een box-plot getekend. De box-plot is anders getekend dan je gewend bent, want de strepen lopen niet door tot de uiteinden. In plaats daarvan heeft men de drie laagste en de drie hoogste waarnemingen apart met een rondje aangegeven.

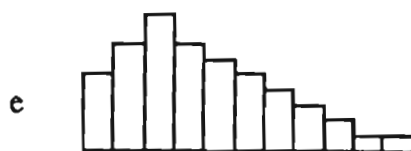
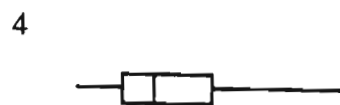
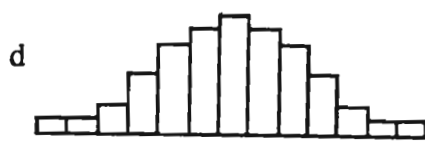
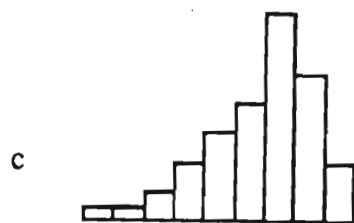
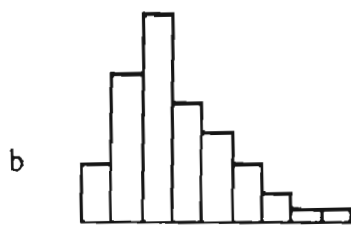
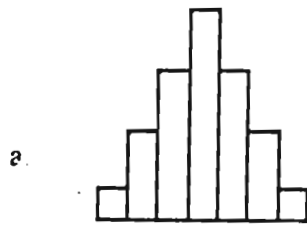


Tagestemperaturen in Bamberg 1967
 Zusammenfassung nach 12 Monaten: Boxplots

- 35 Beantwoord de vragen van hiervoor nog een keer, maar nu aan de hand van deze grafiek.
- 36 Vergelijk de twee grafieken eens met elkaar:
 - Welke grafiek is het makkelijkste te begrijpen?
 - Welke grafiek bevat de meeste informatie?

Histogram en box-plot vergeleken

37 Hieronder staan vijf histogrammen en vijf box-plots. Ze horen twee-aan-twee bij elkaar. Welk histogram hoort bij welke box-plot?



Symmetrisch of scheef

Bij het beantwoorden van de vorige vraag heb je vast en zeker gebruik gemaakt van *symmetrie* in de grafieken. Histogram symmetrisch? Dan is de box-plot die daar bij hoort ook symmetrisch!

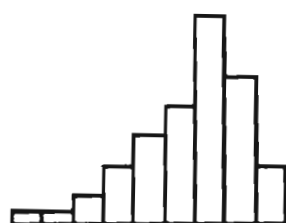
- 38a Bij een histogram kun je altijd een box-plot tekenen. Gaat het omgekeerde ook op?
 b Iemand zegt: 'Hoge staven in het midden van een histogram betekent een grote doos in het midden van de bijbehorende box-plot'.
 Welke denkfout wordt hier gemaakt?

- 39 In het kader van een medisch experiment wordt een aantal cavia's ingespoten met een tuberculose bacil. Men kijkt na hoeveel dagen de cavia's overlijden. Na 43 dagen gaat de eerste dood en de sterkste houdt het 598 dagen vol.
 De volledige gegevens zijn:

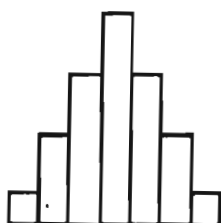
43	45	53	56	56	57	58	66	67
73	74	79	80	80	81	81	81	82
83	83	84	88	89	91	91	92	92
97	99	99	100	100	101	102	102	102
103	104	107	108	109	113	114	118	121
123	126	128	137	138	139	144	145	147
156	162	174	178	179	184	191	198	211
214	243	249	329	380	403	511	522	598

- a Maak bij deze gegevens een histogram. Kies eerst zelf een klassebreedte die je geschikt lijkt, rekening houdend met het aantal klassen je dan krijgt. Komen er ook lege klassen voor?
 b Teken bij deze gegevens de box-plot.

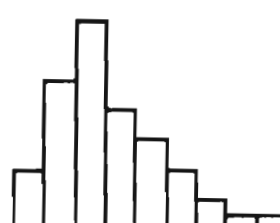
Een histogram kan symmetrisch zijn of scheef:



scheef naar links



symmetrisch

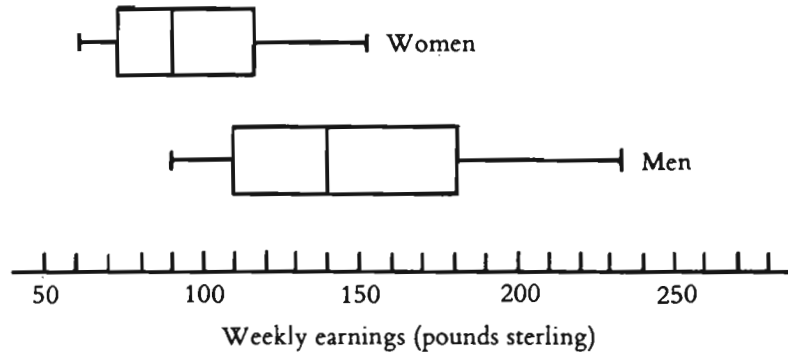


scheef naar rechts

- 40 Is de verdeling van de levensduur van de cavia's scheef naar links of scheef naar rechts?
 41 Hoe kun je aan een box-plot zien of een verdeling scheef naar links of scheef naar rechts is?

Nogmaals salarissen

Hieronder staan nogmaals de box-plots van de salarissen in Engeland:



42 Wat kun je zeggen over de scheefheid van de inkomensverdelingen voor vrouwen en mannen?

43 Een lastige vraag:

Probeer eens of je een box-plot voor de mannen en vrouwen samen kunt maken!

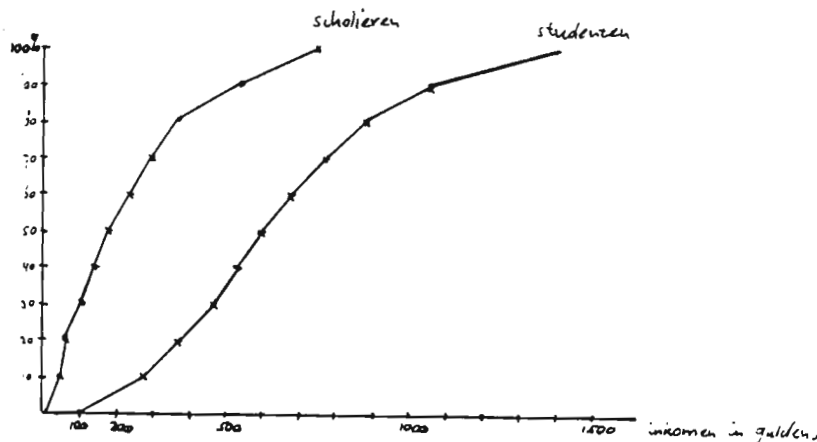
Neem voor het gemak even aan dat de getekende box-plots betrekking hebben op 1000 vrouwen en 1000 mannen.

De volgende vragen helpen je een stukje op weg:

- wat is het verschil tussen het hoogste en het laagste salaris voor de hele groep?
- wat kun je zeggen over de mediaan van de hele groep?

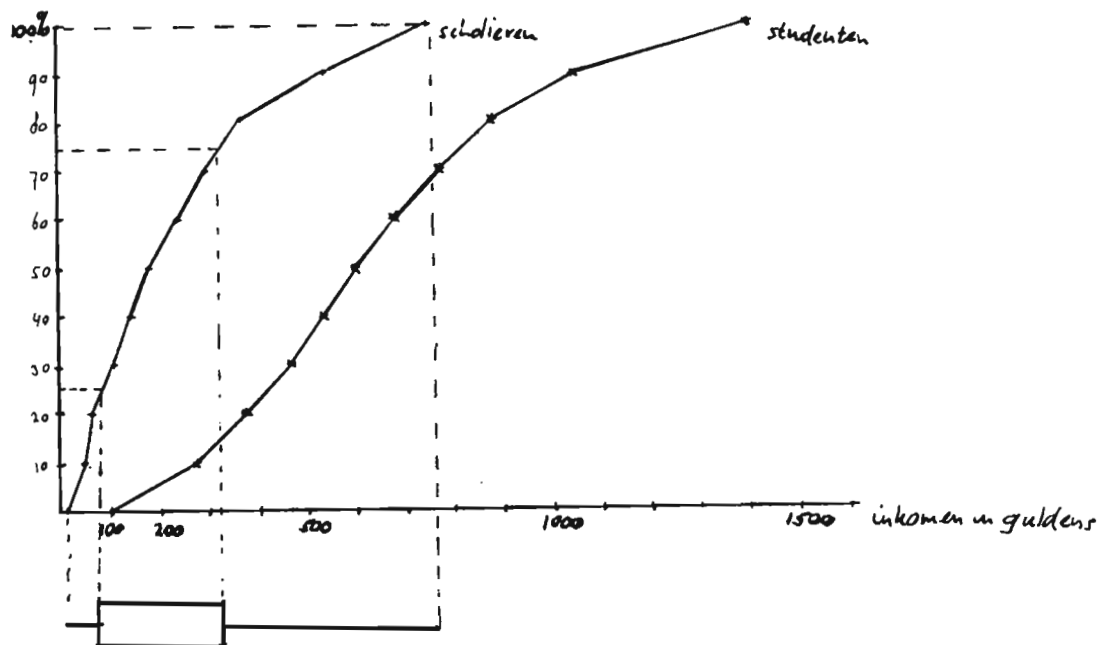
Inkomsten van scholieren

Elk jaar wordt er een onderzoek gedaan onder de Nederlandse jeugd. Men kijkt dan onder andere naar de maandelijkse inkomsten van scholieren (13-17 jaar) en studenten (18-21 jaar). De grafiek hieronder brengt dat in beeld.



- 44 Een uitspraak bij deze grafiek is:
 '30 procent van de scholieren heeft maandelijkse inkomsten van ten hoogste'
 a Welk getal moet op de stippeltjes staan?
 b Hoeveel procent van de studenten heeft maandelijkse inkomsten van méér dan 1000 gulden?

De inkomsten van de scholieren kunnen ook weergegeven worden in een box-plot. Die box-plot kun je afleiden uit de gegeven grafiek. Dat gaat zo:



- 45a Kun je uitleggen hoe de box-plot getallen zijn gevonden?
 b De mediaan ontbreekt nog, teken die zelf in.

46 Maak op deze manier ook de box-plot voor de maandelijkse inkomsten van studenten.

Uit het scholierenonderzoek is ook gebleken dat het *gemiddelde* maandinkomen van scholieren 248 gulden bedraagt.

- 47 De zestienjarige Peter heeft maandelijks 200 gulden te besteden. Hij vindt dat zelf nogal weinig. Zijn moeder zegt dat dat heus wel meevalt.
 - Bedenk een argument waarmee Peter moeder kan overtuigen.
 - Bedenk een argument waarmee moeder Peter kan overtuigen.

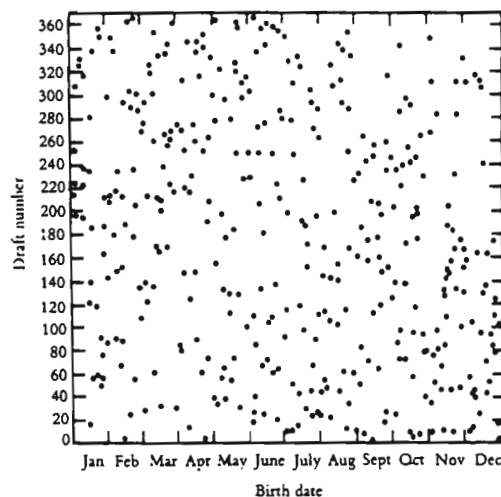
Waar gebeurd verhaal

In 1970 werd in Amerika door middel van loting uitgemaakt wie er in militaire dienst moest. In die tijd betekende dat ook dat je als soldaat uitgezonden kon worden naar Vietnam, waar toen oorlog was. Men wilde volgens het toeval geboortedata uitkiezen. De mannen met de uitgekozen geboortedata en geboren tussen 1943 en 1952 zouden opgeroepen worden. De trekking van de geboortedata verliep als volgt.

Alle dagen van het jaar werden stuk voor stuk op een briefje geschreven en de briefjes werden in een balletje gestopt. De balletjes gingen in een grote ton.

Het eerste balletje werd getrokken en alle soldaten met de geboortedatum van dit balletje kregen nummer 1. Daarna werd het volgende balletje getrokken, de soldaten met die datum kregen nummer 2. Enzovoort, tot alle 366 balletjes getrokken waren. De mannen zouden in volgorde van de nummers opgeroepen worden. Eerst de nummers 1, als er meer soldaten nodig waren de nummers 2, enzovoort. Met een laag nummer was je dus snel aan de beurt. De gelukkigen met een hoog nummer zouden waarschijnlijk helemaal niet opgeroepen worden.

In de figuur hieronder zijn de nummers uitgezet tegen de maanden.

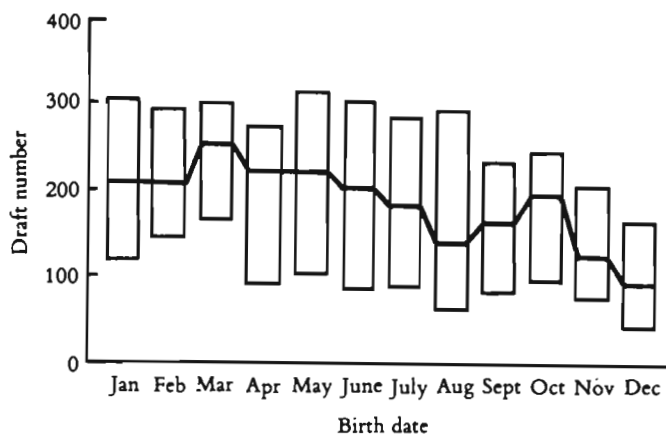


Het lijkt erop dat de nummers 1 t/m 5 terecht zijn gekomen in februari, april, september, oktober en december. De precieze volgorde is niet goed af te lezen.

Ogenschijnlijk lijkt het alsof de stippen willekeurig over de maanden verdeeld zijn.

Een Amerikaanse journalist heeft echter ontdekt dat de verdeling over de maanden niet louter toevallig is. December heeft teveel lage nummers gekregen.

Het bleek dat men de balletjes onvoldoende geschud had. Daardoor lagen de balletjes van december in verhouding teveel bovenop, waardoor ze een grotere kans hadden om getrokken te worden en dus lage nummers kregen. Pech voor de mannen die in december geboren waren!

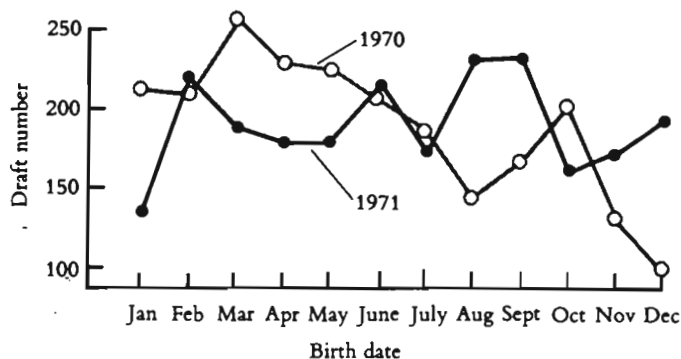


Maand-medianen en -kwartielen voor de loterij van 1970

Het bovenstaande plaatje laat zien dat de data in december inderdaad lage nummers hadden. Per maand is de doos van de box-plot getekend (de strepen zijn dus weggelaten). De medianen zijn vervolgens met elkaar verbonden. Nu is duidelijk te zien dat december inderdaad veel lage nummers heeft.

Zo zie je hoe je door het kiezen van een geschikt plaatje méér informatie uit waarnemingen af kunt leiden!

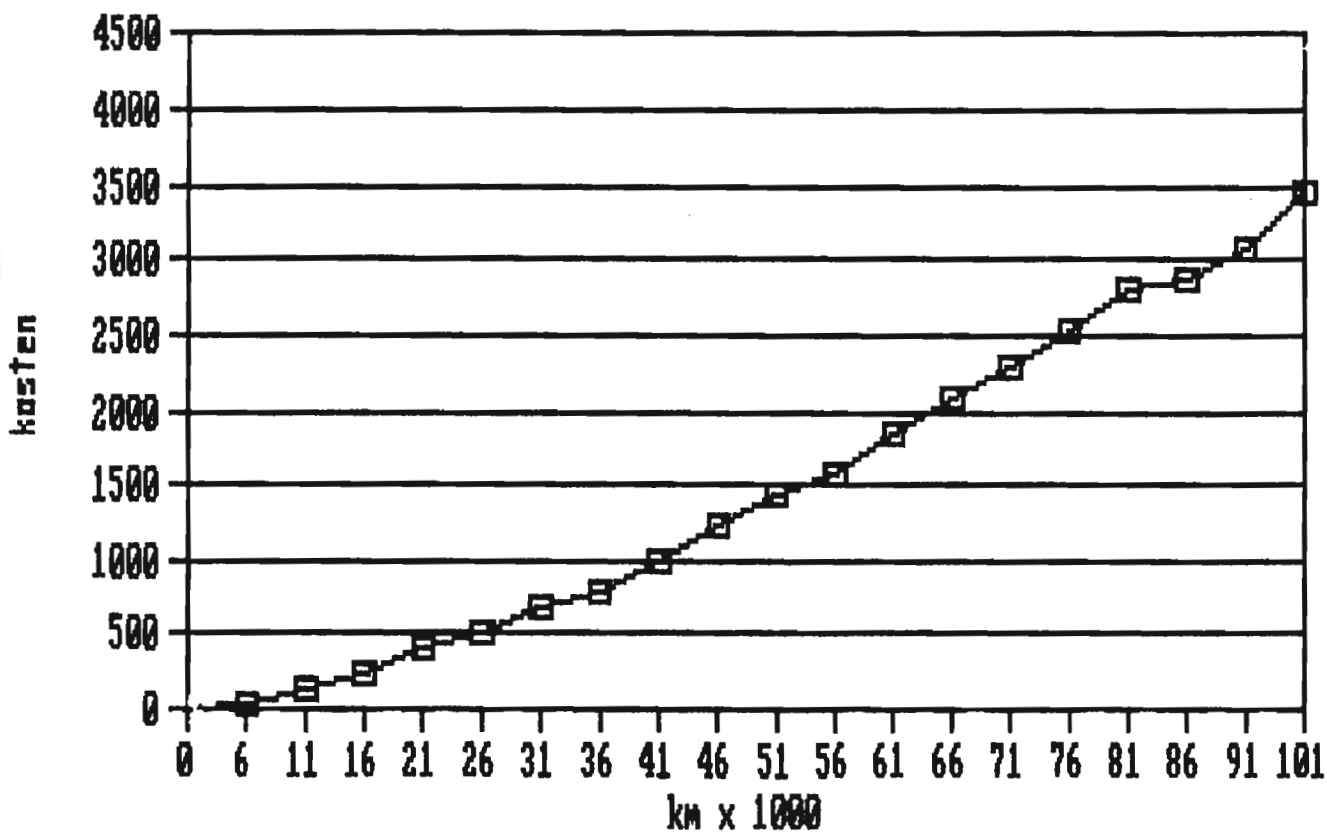
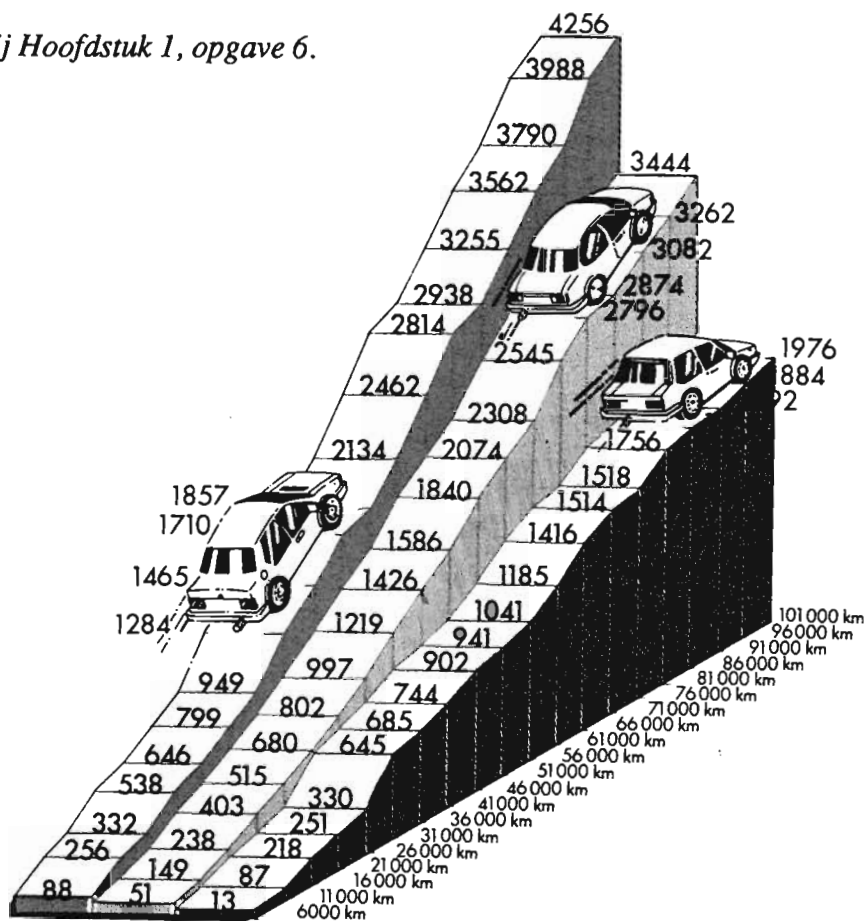
Het jaar daarop, in 1971, heeft men de lotingsprocedure veranderd. Nu werden twee tonnen met balletjes gebruikt, één met de geboortedata en één met de getallen 1 t/m 366. De twee tonnen werden goed geschud. Daarna werd uit beide tonnen een balletje getrokken. De geboortedatum uit de ene ton kreeg het nummer van het balletje uit de andere ton. In de figuur hieronder staan de medianen per maand voor 1970 en 1971. De spreiding blijkt ook in 1971 vrij groot te zijn, maar het extreme resultaat van december 1970 komt niet meer voor. De loting was nu wel eerlijk.



Maand-medianen voor de loterijen van 1970 en 1971

Werkblad 1

Bij Hoofdstuk 1, opgave 6.

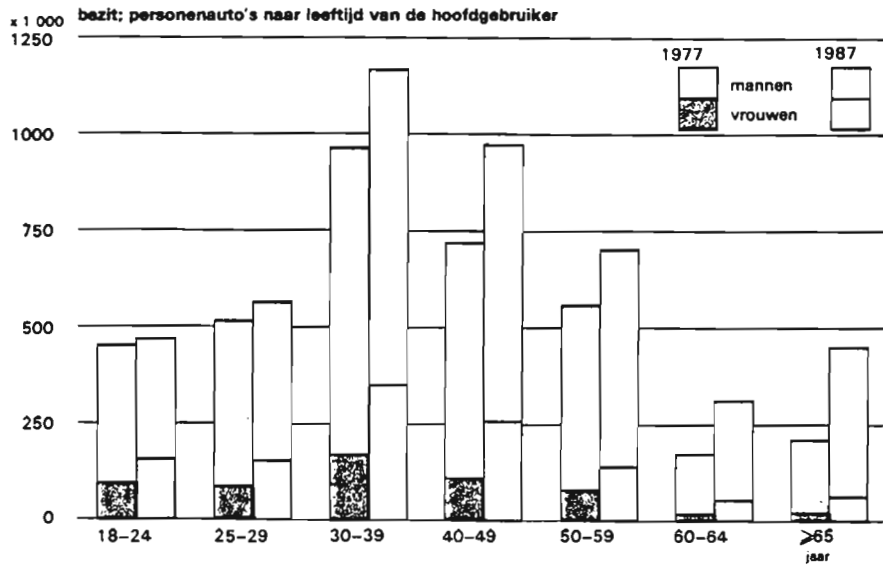


Eén auto is alvast getekend. Schets de grafieken voor de twee andere auto's.

Werkblad 2

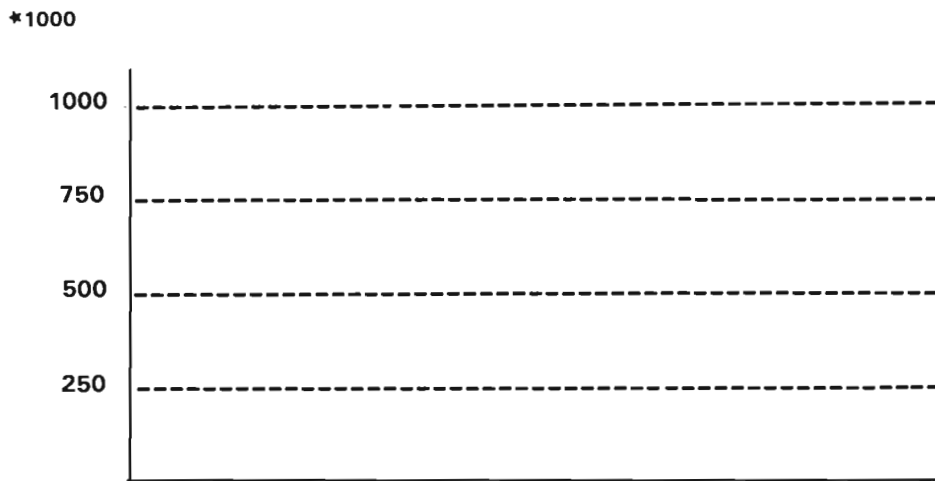
Bij Hoofdstuk 1, opgave 7.

Kleur de staven die het autobezit van vrouwen in 1987 aangeven.



Bij Hoofdstuk 1, opgave 11.

Teken een staafdiagram met alléén het aantal mannelijke autobezitters in 1987.



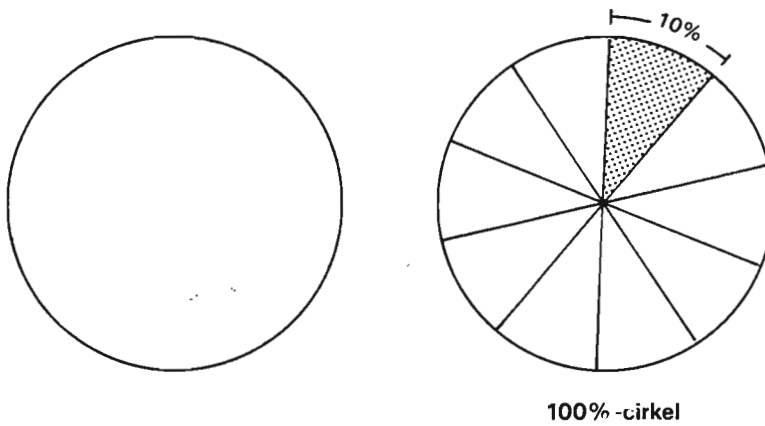
Werkblad 3

Bij Hoofdstuk 1, opgave 14.

Teken een cirkeldiagram bij de tabel.

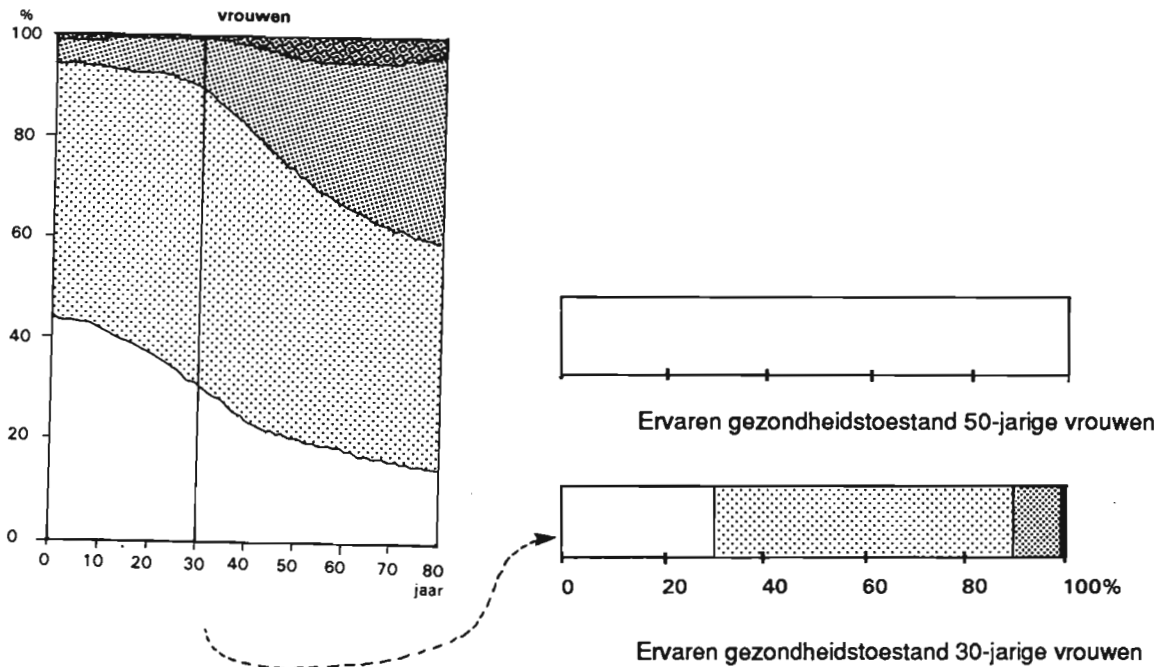
Gezinssamenstelling van heel Nederland.

type gezin	percentage
echtpaar	19,0 %
echtpaar met kinderen	32,9 %
éénoudergezinnen	6,3 %
niet in gezinsverband levend	41,8 %



Bij Hoofdstuk 1, opgave 21.

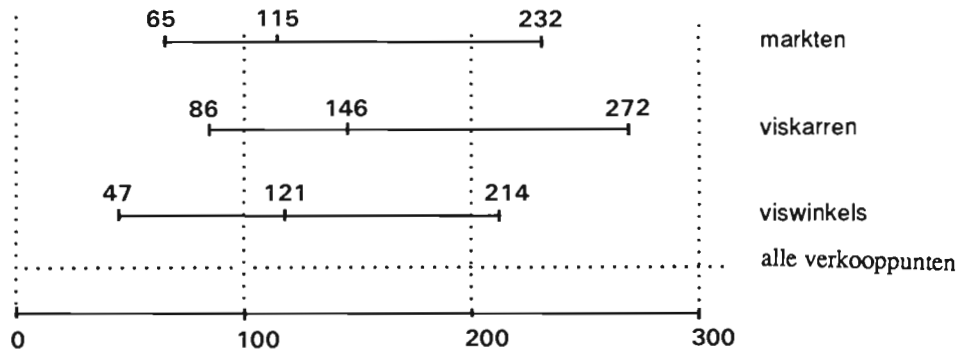
Teken een reepdiagram voor de vrouwen van 50 jaar.



Werkblad 4

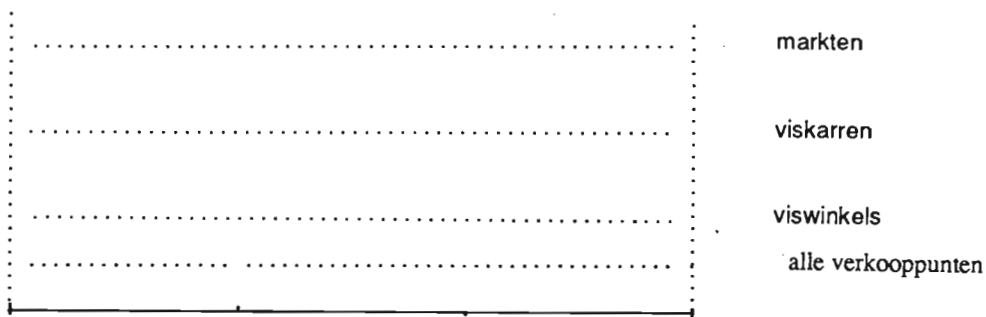
Bij Hoofdstuk 1, opgave 25.

Teken het lijnstukje voor 'alle verkooppunten'.



Bij Hoofdstuk 1, opgave 30.

Teken de lijnstukjes die horen bij de 'prijs per kg' (zie de tabel over de lekkerbekjes).



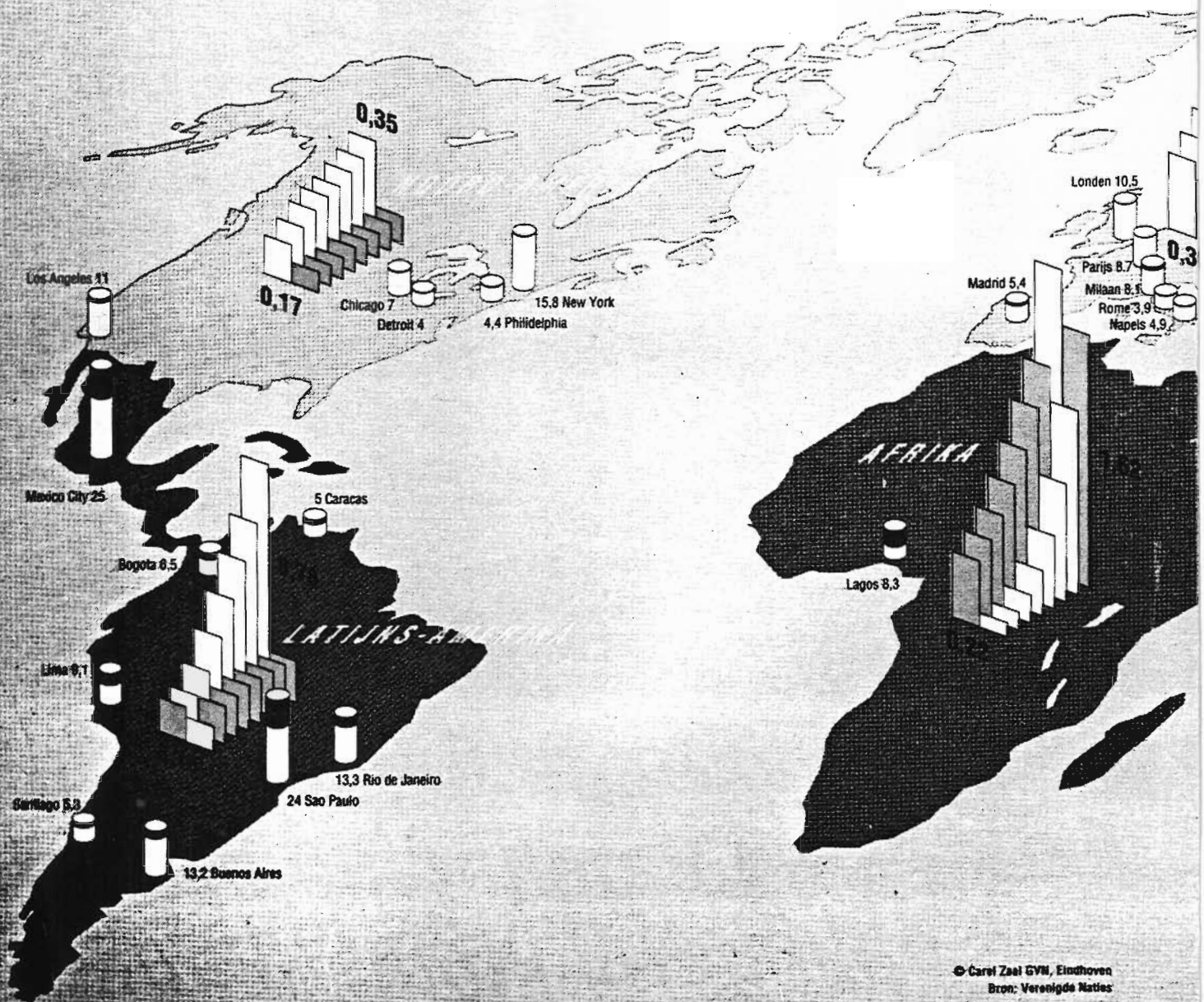
Bij Hoofdstuk 1, opgave 40.

Maak de tabel af en teken er een cirkeldiagram bij.

	pers. verzorg.	kinder- verzorg.	huish. werk	bood- schappen	betaald werk	onderw. vorming	reizen	vrije tijd	totaal
uren:minuten	11:17	0:04	0:35	0:14	0:20	4:47	1:15	5:28	24:00
minuten	677								1440
procenten	47								100

VERSTEDELIJKINGSKAART

Vanaf het jaar 2015 woont meer dan de helft van de wereldbevolking in de steden



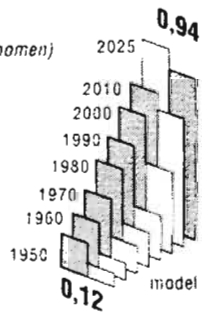
© Carel Zaai GVM, Eindhoven
Bron: Verenigde Naties

De groei van de stedelijke bevolking t.o.v. de plattelandsbevolking (1950 -2025)

Per macro-regio
(China en India zijn ook apart opgenomen)

- Stedelijke bevolking
- Plattelandsbevolking

De getallen geven aan hoe groot de **totale** bevolking is in 1950 en in 2025 in miljarden inwoners



De 46 grootste steden in 1985 en hun groei tot het jaar 2000

Voorbeeld: de stad Lahore in Pakistan

6.2 Lahore

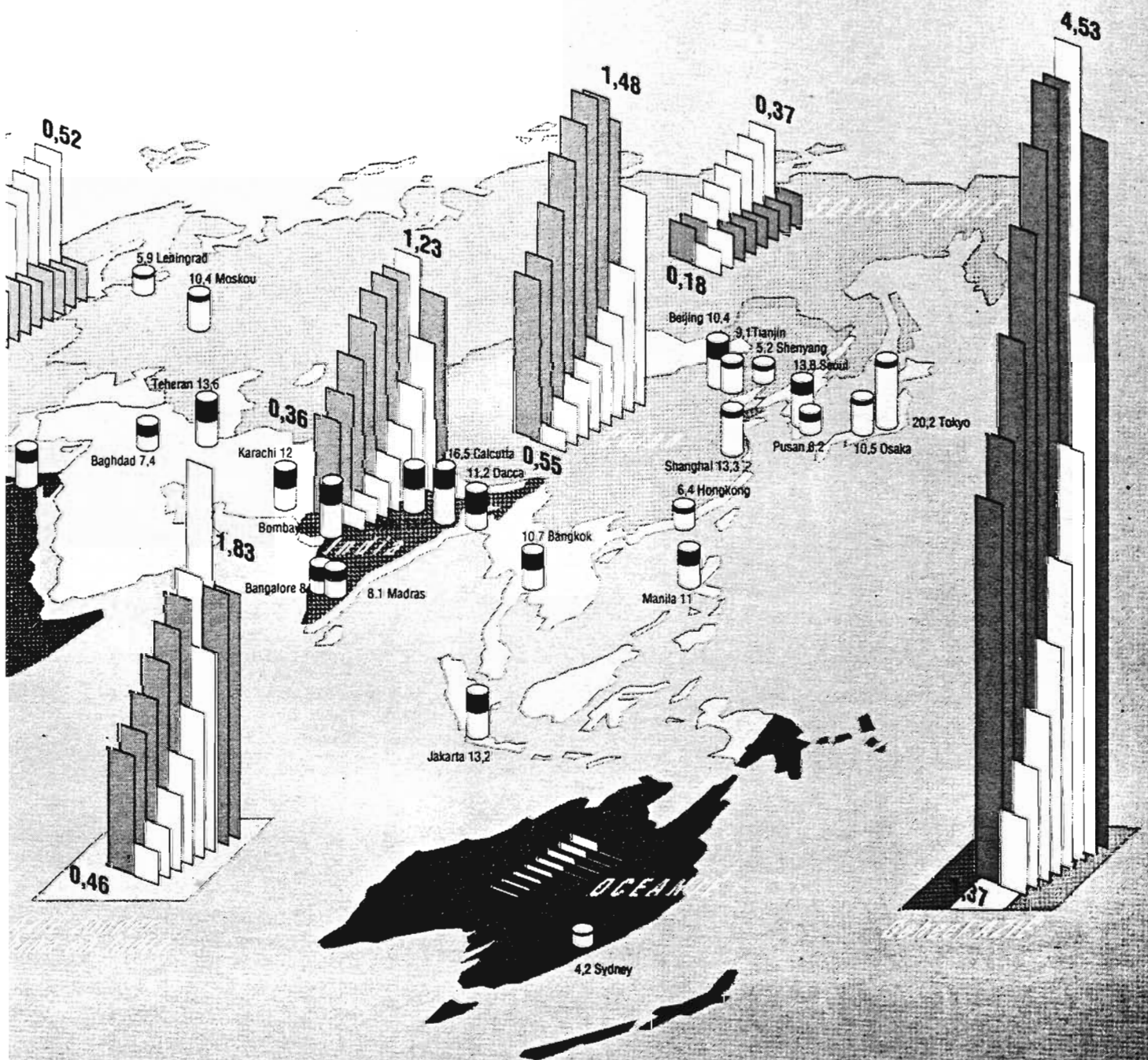


Aantal inwoners in het jaar 2000 in miljoenen



Toename aantal inwoners van 1985 - 2000

Aantal inwoners in 1985



archief FI

02.01.91

Statistiek

Leerlingenmateriaal

Verhage, H. , G. van den Heuvel , H.