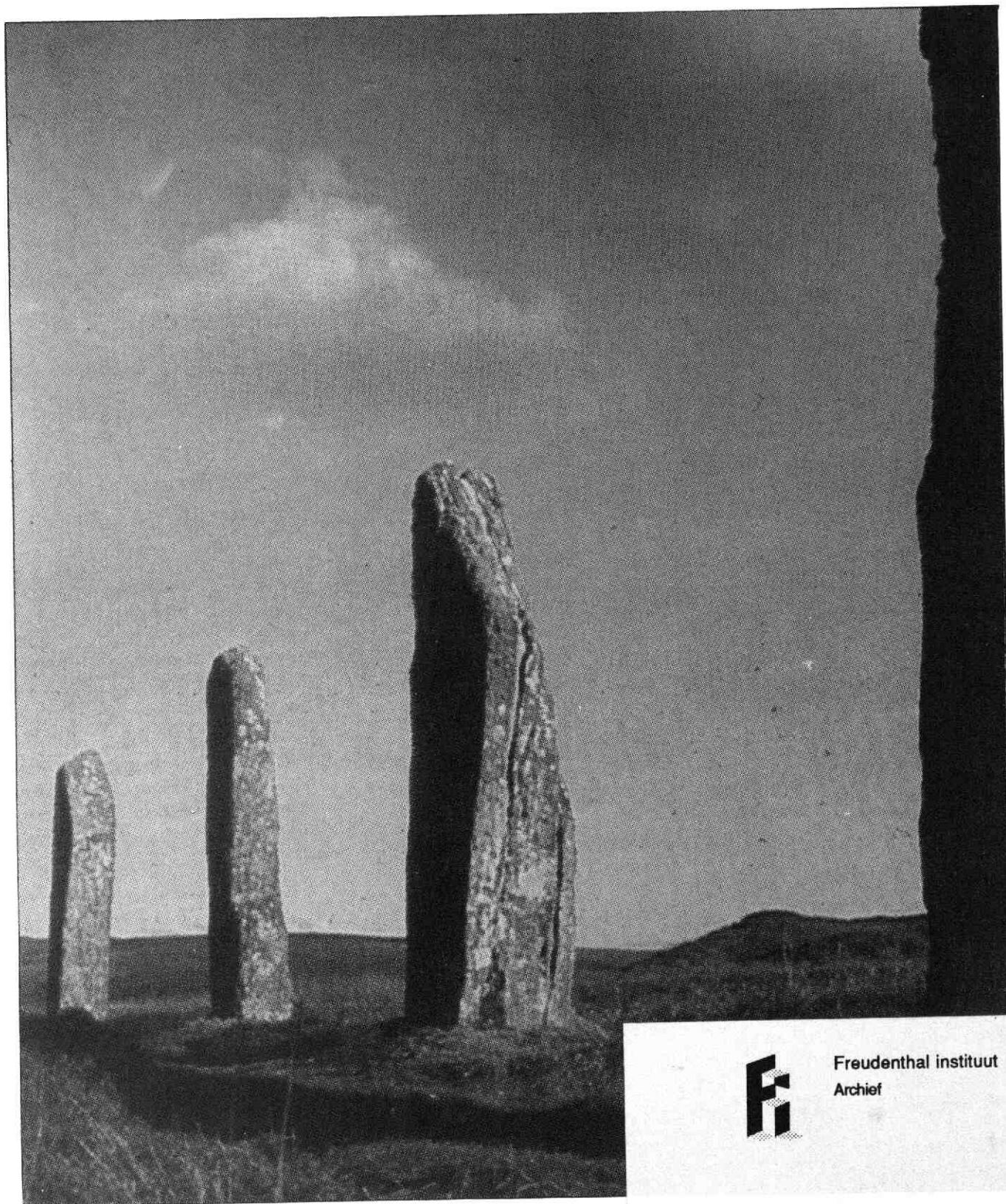




Grafische verwerking

<https://hdl.handle.net/1874/10232>



Freudenthal instituut
Archief

GRAFISCHE VERWERKING

**GRAFISCHE
VERWERKING**



Tiberdreef 4 - 3561 GG Utrecht

Omslagfoto:

*Geen versteend staafdiagram, maar
'standing stones' die deel uitmaken
van de Ring van Brogar, op de
Orkneys.*

GRAFISCHE VERWERKING

Een produktie ten behoeve van de
experimenten in het kader van de
Herverkaveling Eindexamenprogramma's
Wiskunde I en II V.W.O.

Samenstelling: Jan de Lange Jzn
Martin Kindt

Vormgeving: Ellen Hanepen

© 1983; 3e herziene versie

Utrecht, juni 1983.

*Met dank aan Yvonne Taverne voor het belangeloos ter beschikking
stellen van haar ORKNEY-onderzoeksmateriaal.*

INHOUDSOPGAVE

1. KRITISCH KIJKEN	pag. 1
2. STAAF- EN CIRKELDIAGRAMMEN	11
3. HISTOGRAMMEN EN POLYGONEN	19
4. GRAFISCHE VOORSTELLINGEN	29
5. CENTRUMMATEN	37
6. SPREIDINGSMATEN	43
7. KLASSEBREEDTE	49
8. LORENTZ-KROMMEN	57
9. DRIEHOEKSGRAFIEKEN	61
10. DRIE DIMENSIES EN CORRELATIE	67
11. LEVENSVERWACHTING EN BEVOLKINGSGROEI IN DE DERDE WERELD	75
BIJLAGE	82
12. SAMENVATTING	85

1

KRITISCH KIJKEN

De waarheid spreken kan op vele manieren. Je hoeft de krant er maar op na te lezen om te zien dat dit juist is. Vooral in de politiek zie je nogal eens tegengestelde meningen. En beide partijen zeggen het gelijk aan hun zijde te hebben.

Een fraai (?) voorbeeld uit het voorjaar van 1979, toen Carter nog president van de Verenigde Staten was:

De citaten:

"De defensieuitgaven zijn *gedaald* tot het niveau van voor de Vietnamoorlog."

Harold Brown
Ministerie van Defensie
21 februari 1979.

"Ik ben verheugd u te kunnen meedelen dat onder president Carter het werkelijke uitgavenpeil voor defensie steeds is *gestegen*."

Edward R. Jayne,
8 maart 1979.

Wie spreekt de waarheid?

Harold Brown en Edward Jayne lijken allebei fatsoenlijke heren, die hun uitspraken heus wel kunnen 'bewijzen' met getallenmateriaal. Hetzelfde getallenmateriaal? Dat zal haast wel. Maar waarom komen ze dan tot heel verschillende conclusies? Omdat het ze beter uitkomt, dat spreekt.

In dit hoofdstukje zullen we eens bekijken hoe je kunt manipuleren met getallen.



Jimmy Carter



Harold Brown

Toen Harold Brown werd gevraagd om zijn uitspraak:

"de defensieuitgaven zijn gedaald tot het niveau van vòòr de Vietnamoorlog"

met feiten te staven, kwam hij met de volgende gegevens aanzetten:

- voor het jaar 1980 vertoont het bedrag dat voor defensie is uitgetrokken weliswaar een werkelijke toename van 3,1% vergeleken met 1979, maar het is daarmee slechts 4,6% van het Bruto Nationaal Produkt*, en dat is het laagste sinds 1940.
- het bedrag voor defensie is 23% van het totaal van alle (Federale) uitgaven, en dat is, afgezien van 1978, het laagste sinds 1940.

Aangezien Brown het niet doet, zullen wij je maar vertellen om hoeveel dollar het nu eigenlijk gaat.

In 1979 besteedden de Amerikanen zo'n 114,7 miljard dollar aan defensie en voor 1980 stond er 126 miljard dollar op de begroting.

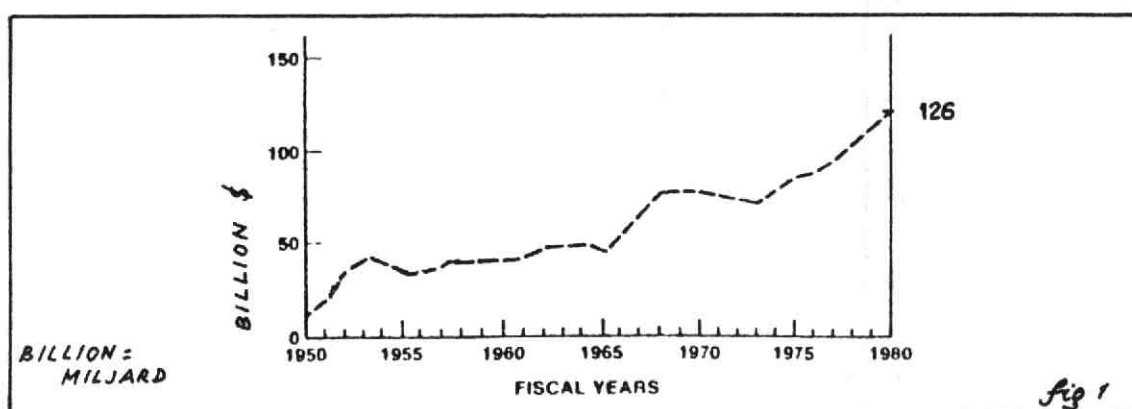
 * B.N.P. = De totale waarde van goederen en diensten geproduceerd door de binnenlandse economie gedurende een jaar.

- » 1. Vergelijk deze getallen eens met de uitspraken van Brown. Hoe zit dat met die 3,1% werkelijke toename? Wat vind je van de argumenten van Harold Brown?

Tot zover Brown. Nu naar senator Edward Jayne. Die zei:

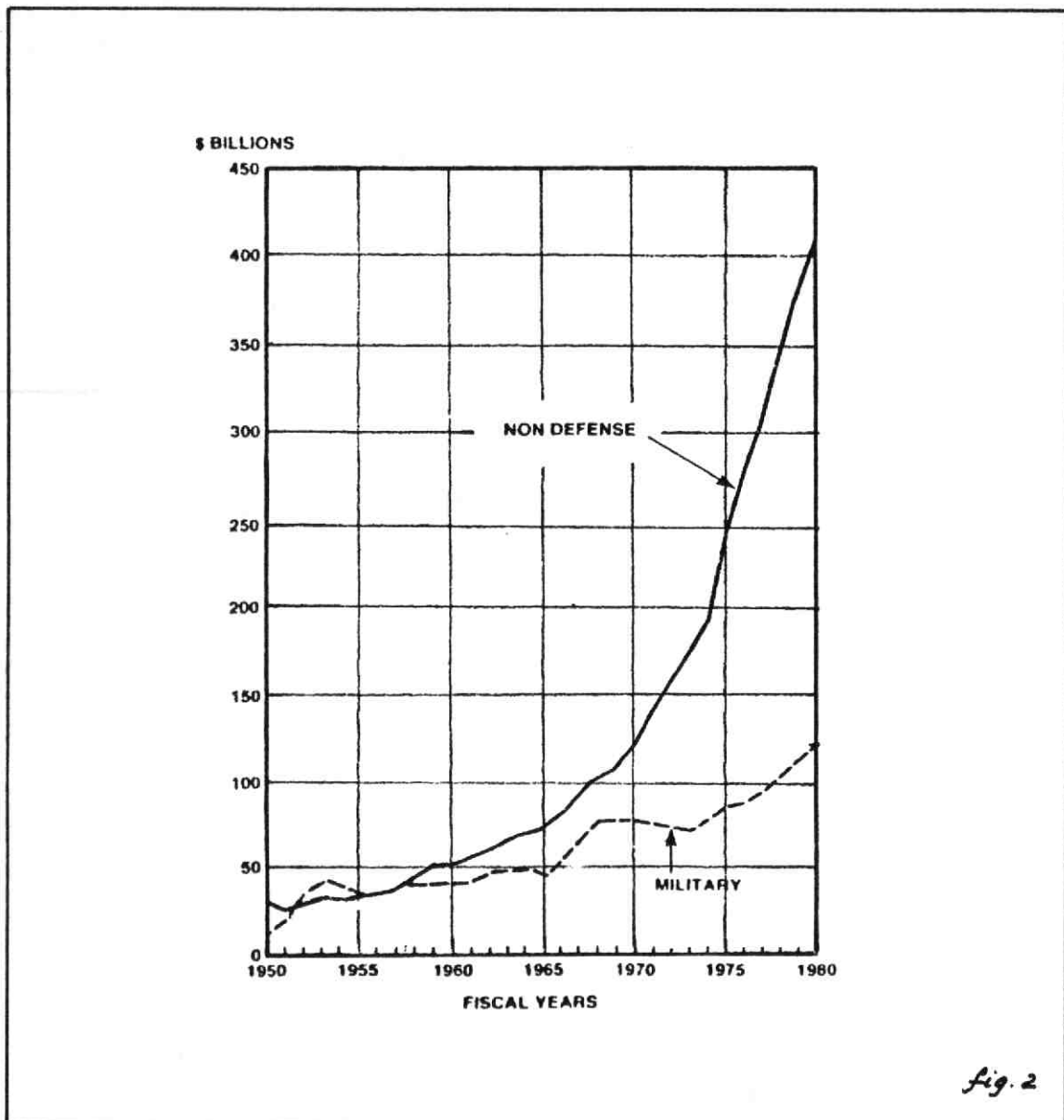
"het werkelijke uitgavenpeil voor defensie is steeds gestegen (onder Carter)."

Hij baseert zich daarbij mede op de volgende grafiek:



Je ziet in de grafiek hierboven de defensieuitgaven in miljarden dollars (langs de verticale as), zoals die in de betreffende jaren (langs de horizontale as) op de begroting te vinden waren.

- » 2. Wat vind je van de uitspraak van Jayne?
- » 3. Kun je aan de hand van bovenstaande grafiek een schatting maken van het *extra* bedrag dat de Koreaanse oorlog en de Vietnamoorlog maximaal per jaar hebben gekost?
- » 4. Brown zei: "de defensieuitgaven zijn gedaald tot het niveau van vòòr de Vietnamoorlog."
Klopt dat met bovenstaande grafiek? Vind je dat Brown 'liegt'?



Hierboven nogmaals de grafiek van de defensieuitgaven, maar in hetzelfde plaatje ook de grafiek van de overige (= niet-defensie) uitgaven.

Jayne zegt dat de defensieuitgaven zijn gestegen. Dat klopt met de grafiek. Jayne kijkt naar de "absolute getallen" voor defensie en die zijn sinds 1973 steeds gestegen. Brown daarentegen zei niets over deze werkelijke getallen, maar vergeleek die met andere uitgaven; hij werkte met "relatieve getallen", uitgedrukt in percentages. Laten we dat ook eens doen aan de hand van bovenstaande twee grafieken.

- » 5. Vul de volgende tabel in met behulp van de grafieken van de vorige bladzijde:

	DEFENSIE UITGAVEN IN MILJARDEN \$	NIET-DEFENSIE UITGAVEN IN MILJ. \$	DEFENSIE UITGAVEN als % van NIET- DEFENSIE-UITGAVEN
1955			
1960			
1965			
1970			
1975			
1980			

- » 6. Teken een grafiek waarbij je horizontaal de tijd uitzet (1955-1960-1965-1970-1975-1980) en vertikaal het percentage dat de defensie-uitgaven bedragen van de niet-defensieuitgaven.
Mag je de zes punten van de grafiek zó maar door lijntjes verbinden?
Kies zelf geschikte schaalverdelingen langs de assen.
- » 7. Vergelijk de gevonden resultaten met de uitspraken van Brown en Jayne.
- » 8. Eén van de zaken die je in aanmerking kan nemen is de rol van de inflatie. Misschien is dat woord al eerder gevallen in de discussie tot nog toe. Zo ja, bij welke vraag en met welk resultaat?

Eén van de mogelijkheden om de inflatierol 'uit te schakelen' is om alle bedragen, zoals die op de begrotingen stonden, te herleiden tot "1980 Dollars".

- » 9. Wat zal bedoeld zijn met "1980 dollars"?



Dit plaatje toont de afnemende waarde van de dollar sinds 1950. Het plaatje is getekend in 1970 zodat de 46 cent voor 1980 slechts een voorspelling is. In werkelijkheid was de 1950 dollar in 1980 minder dan 40 cent!

Dit verschijnsel - inflatie - heeft natuurlijk grote invloed op begrotingen. Als de begroting voor defensie in 1950 40 miljard bedraagt, zal die in 1980 ruim 100 miljard moeten zijn om op hetzelfde reële niveau te blijven!

» 10. Dit artikeltje komt uit de Volkskrant van 11-2-1983. Daaruit blijkt dat er in de afgelopen 40 jaar een inflatie van ruim 540.000% is geweest in Argentinië.

Argentijnse inflatie 540.115 percent

Van onze correspondent

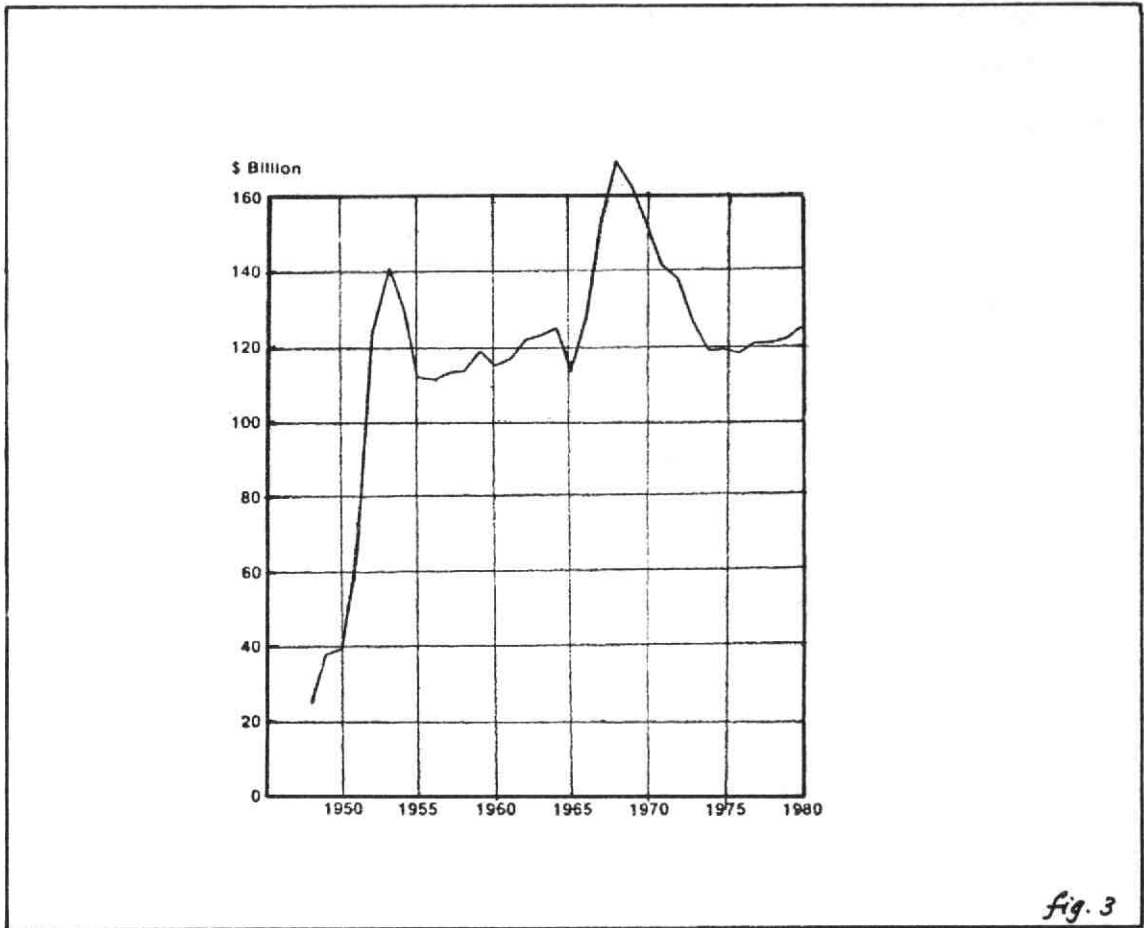
MEXICO-STAD — De inflatie in Argentinië gedurende de afgelopen 40 jaar bedraagt 540.115,064 percent. Dit heeft het in Buenos Aires verschijnende blad *Tiempo Argentino* becijferd na een analyse van de prijsontwikkeling tijdens de eerste regering Peron in 1944. Om het inflatiecijfer in beeld te brengen zou een grafiek nodig zijn met een breedte van 80 centimeter en een hoogte van 54 kilometer.

De grootste bijdrage aan deze ontwikkeling is geleverd door het bewind van generaal Videla (1976-1981) met een totaal van ruim 9000 percent en een jaargemiddelde van 208. De tweede plaats is voor de weduwe Peron, die absoluut gezien echter hoger scoort: 893,5 percent in één jaar en acht maanden. In de zeven maanden die de huidige president, generaal Bignone aan de macht is bedraagt de inflatie 8,3 percent, terwijl het totaal over 1982 op iets meer dan 200 uitkwam.

a. Bekijk de laatste zin van het vetgedrukte gedeelte. Wat is de schaalverdeling van de grafiek van de inflatie?

b. Klopt de volgende zin? "De grootste bijdrage aan deze ontwikkeling is geleverd door het bewind van generaal Videla (1976 - 1981) met een totaal van ruim 9000 percent en een jaargemiddelde van 208."

Verklaar!



Hierboven de grafiek die de defensieuitgaven van de U.S.A. weergeeft door de jaren heen, waarbij het bedrag voor defensie is uitgedrukt in "1980 dollars".

Je ziet duidelijk dat de Koreaanse oorlog en de Vietnamoorlog voor extra hoge defensieuitgaven hebben gezorgd.

Tussen die twee oorlogen (dus tussen 1955 en 1965) kun je het verloop van de uitgaven redelijk benaderen door een rechte lijn. Deze 'trend' zet zich voort na de Vietnamoorlog (na 1973).

» 11. Probeer deze 'trend' in de grafiek zo nauwkeurig mogelijk te tekenen met één rechte lijn en tracht met behulp van deze trendlijn een voorspelling te doen over het bedrag dat in de VS aan defensie besteed zal worden in 1990. (Uiteraard in "1980 dollars").

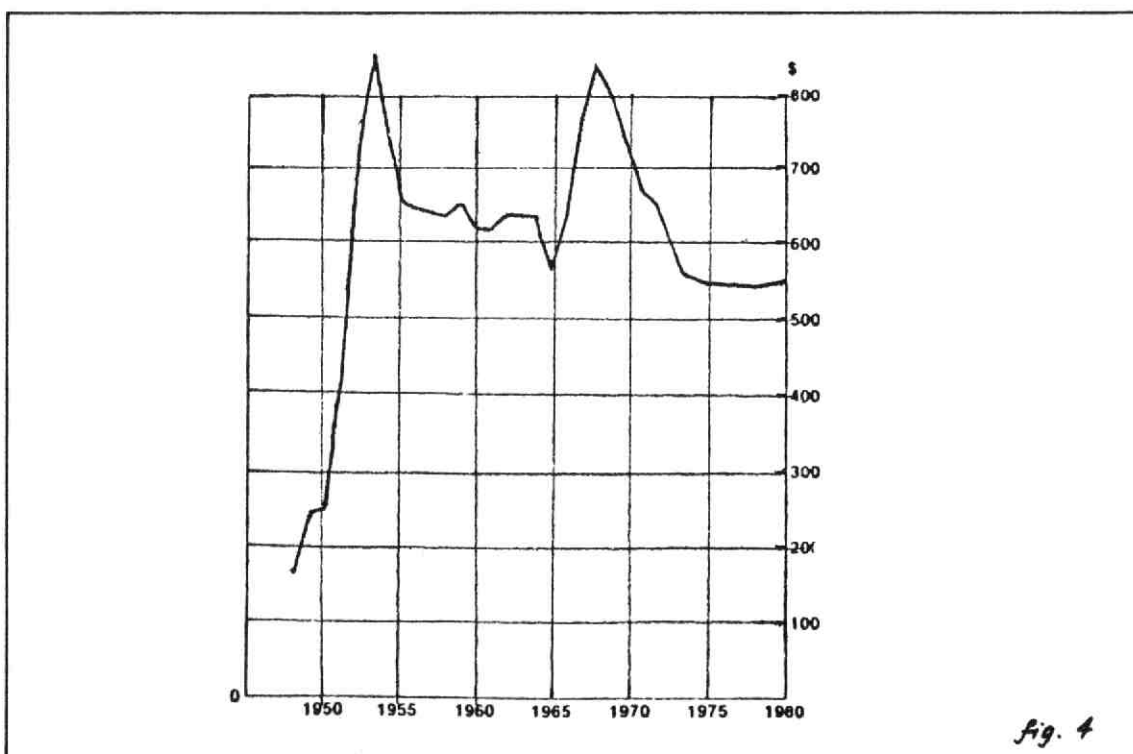
» 12. Als de inflatie in de jaren 80 per jaar gemiddeld zo'n 10% is, voor welk bedrag zal defensie in 1990 dan op de begroting staan?

» 13. Wie heeft gelijk: Brown of Jayne?

Denk niet dat met de grafiek van fig. 3 het laatste woord is gesproken. De vindingsrijkheid van politici is groot. Kijk maar eens kritisch naar de volgende redenering:

In de laatste 35 jaar is de bevolking van de Verenigde Staten flink toegenomen. En, zeggen vele deskundigen, moet je dus de defensieinspanningen vergroten. En, is de redenering, je moet dus niet naar de *totale* voor inflatie gecorrigeerde - kosten kijken, maar naar de kosten per hoofd van de bevolking.

Die grafiek ziet er zó uit:



» 14. Teken ook hier de "trendlijn" van de laatste 25 jaar zo goed mogelijk.

» 15. Wat vind je van het argument dat je per hoofd van de bevolking moet kijken?

» 16. Wie heeft er gelijk, Brown of Jayne?

Eén ding is duidelijk: eenvoudig is het niet om te bepalen wie er gelijk heeft. De ene keer kun je je gelijk proberen te bewijzen met behulp van de *absolute* getallen, een andere keer kan het goed uitkomen om met *relatieve* getallen te werken. In dit laatste geval heb je dan vaak nog keus ook.

Zo vergeleek Brown de defensieuitgaven met het Bruto Nationaal Produkt en met de totale Federale uitgaven. Bovendien zagen we de defensieuitgaven ook nog vergeleken met de niet-defensieuitgaven en werden ze per hoofd van de bevolking uitgerekend. Keus genoeg kortom.

Dit soort praktijken wordt nogal eens toegepast en heus niet alleen in de politiek. Getallen en grafieken worden steeds vaker gebruikt, zowel in wetenschappelijke publikaties als in de pers. Hopelijk is nu al duidelijk dat zulke zaken kritisch bekeken moeten worden. Zeker omdat er nog veel meer kanten aan zo'n discussie zitten. We noemen één voorbeeld:

Stel dat een begroting van een land wordt verhoogd om de werkloosheid te bestrijden.

Het totale bedrag van de begroting gaat daardoor van 100 miljard naar 110 miljard. De defensie-inspanning in dat land was 5 miljard. "Dus" zeggen defensiespecialisten "moet dat nu 5,5 miljard worden", want dan blijft het bedrag voor defensie *relatief* hetzelfde, terwijl het *absoluut* toeneemt. Kortom, het is dus niet erg logisch om voor defensie een 'vast' percentage vast te leggen.

» 17. In een land staat in het jaar 1980 op de begroting voor defensie 30 miljoen. De totale begroting omvat 500 miljoen. Een jaar later staat er voor defensie 35 miljoen op de begroting, op een totaal van 605 miljoen.

De inflatie in de periode tussen de twee begrotingen is 10%.

- Je houdt een lezing voor een pacifistisch gezelschap en wilt uitleggen dat de defensieuitgaven toch maar mooi gedaald zijn het laatste jaar. Hoe?

- Je houdt een toespraak op de K.M.A. en vertelt ze vol trots dat de defensie-inspanningen aanmerkelijk zijn toegenomen. Hoe?



De bultrug is een erg speels en luidruchtig dier. Veel reden om opgewekt te zijn hebben deze walvissen echter niet, want zij worden ernstig met uitsterven bedreigd. Alleen de Eskimo's mogen er nog tien per jaar vangen. Verder geldt er een algeheel vangverbod.

2

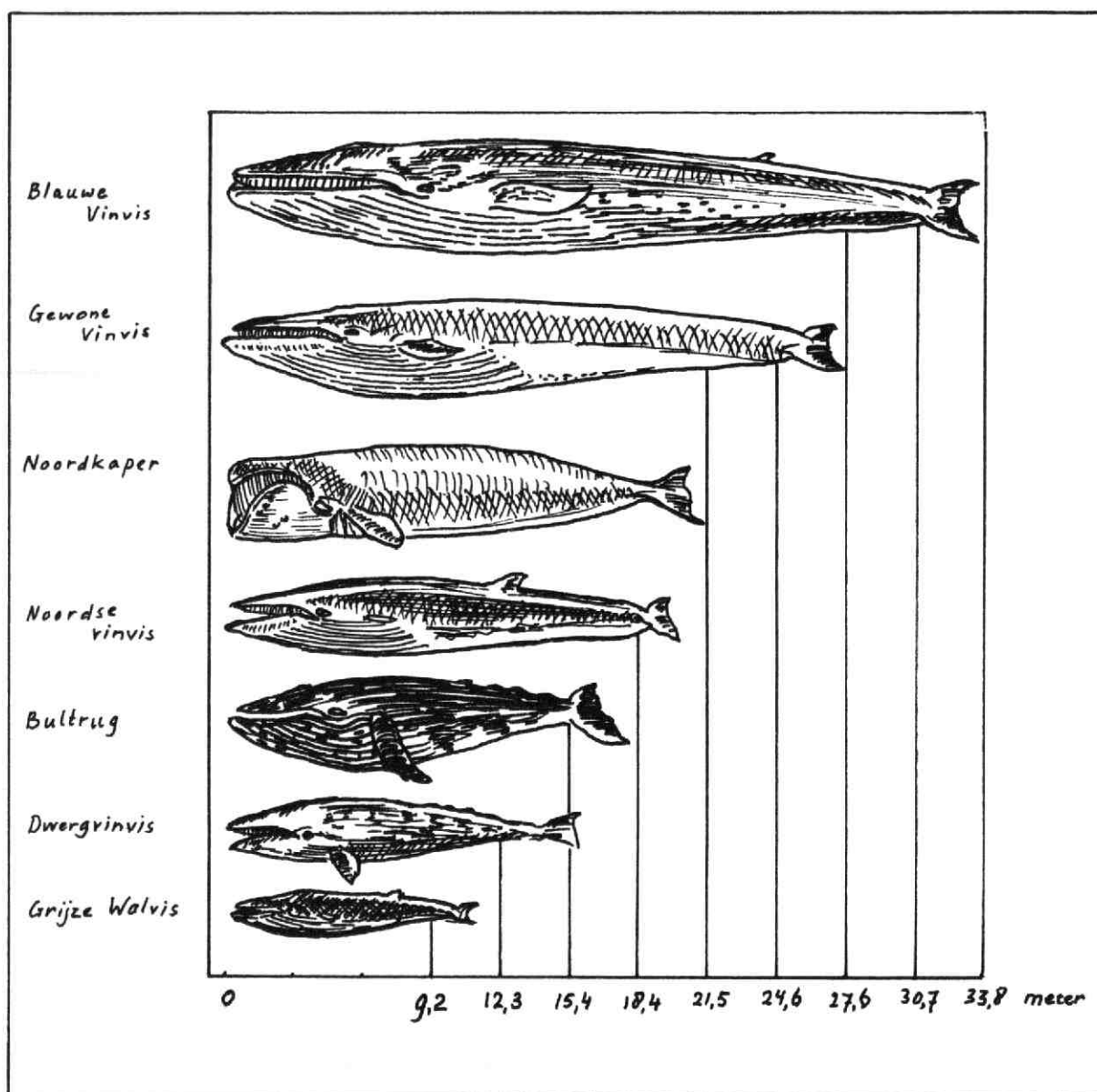
STAAF- EN CIRKELDIAGRAMMEN

Veel diersoorten worden bedreigd. Daarover maken veel mensen en instanties zich zorgen. Regelmatig worden verontrustende cijfers prijsgegeven om te laten zien hoe slecht het er met bepaalde dieren voorstaat. De walvissen vormen een voorbeeld van zo'n bedreigde diersoort. Jacht op walvissen is nog steeds toegestaan. Walvisolie, -vlees, -bot en andere delen worden verwerkt in kunstmest, soep, cosmetica, margarine, zeep en vooral honde- en kattevoer.

Regelmatig worden - uitermate moeilijk uit te voeren en daarom ook niet erg betrouwbare - tellingen uitgevoerd.

De gegevens voor 1972 waren in *tabelvorm*:

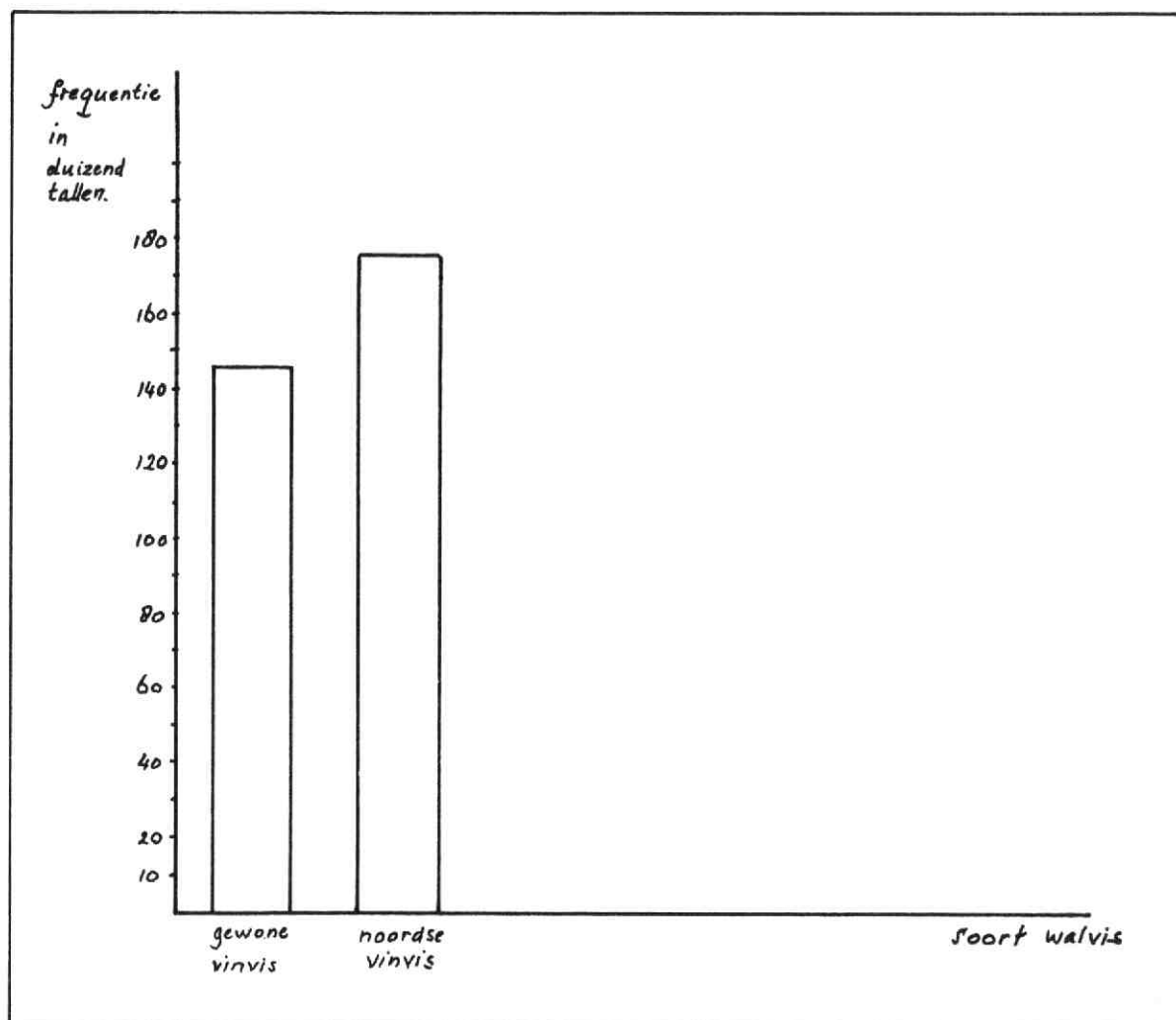
SOORT WALVIS	FREQUENTIE 1972 (aantal)
Gewone Vinvis	145.000
Noordse Vinvis	175.000
Blauwe Vinvis	11.000
Dwerg Walvis	9.000
Bultrug	6.300
Grijze Walvis	11.000
Dwerg Vinvis	150.000



We kunnen de tabel ook 'in beeld brengen' met een grafiek. Eén van de manieren om dat te doen is het *staafdiagram*. De hoogte van de staaf geeft de frequentie aan. De breedte van de staaf doet er vaak niet toe.

Op de volgende bladzijde is een gedeelte van de tabel van blz.11 als staafdiagram getekend.

» 18. Neem het getekende deel in je schrift over en maak het staafdiagram af.



Langs de vertikale as staan de (absolute) frequenties of aantallen. Langs de horizontale as de walvissoorten.

» 19. Doet de volgorde van de walvissoorten langs de horizontale as er toe?

Opmerking:

Vaak wordt het staafdiagram zo getekend dat de staven van links naar rechts steeds kleiner worden, maar dit kan natuurlijk *alléén* maar als de volgorde van de zaken langs de horizontale as er niet toe doet.

Er zijn - ruwe - schattingen van de oorspronkelijke aantallen walvissen.
In tabelvorm:

SOORT WALVIS	FREQUENTIE (oorspronkelijk)
Gewone vinvis	428.000
Noordse vinvis	210.000
Blauwe vinvis	156.000
Dwerg walvis	7.500
Bultrug	110.000
Grijze walvis	20.000
Dwerg vinvis	150.000

» 20. Maak een staafdiagram van deze tabel en wel zó dat de staaflengte van links naar rechts afneemt.

Als je de twee tabellen en staafdiagrammen met elkaar vergelijkt is het lastig om te zien welke diersoort het meest bedreigd is. Dit is o.a. te verhelpen door met *relatieve frequenties* te gaan werken.

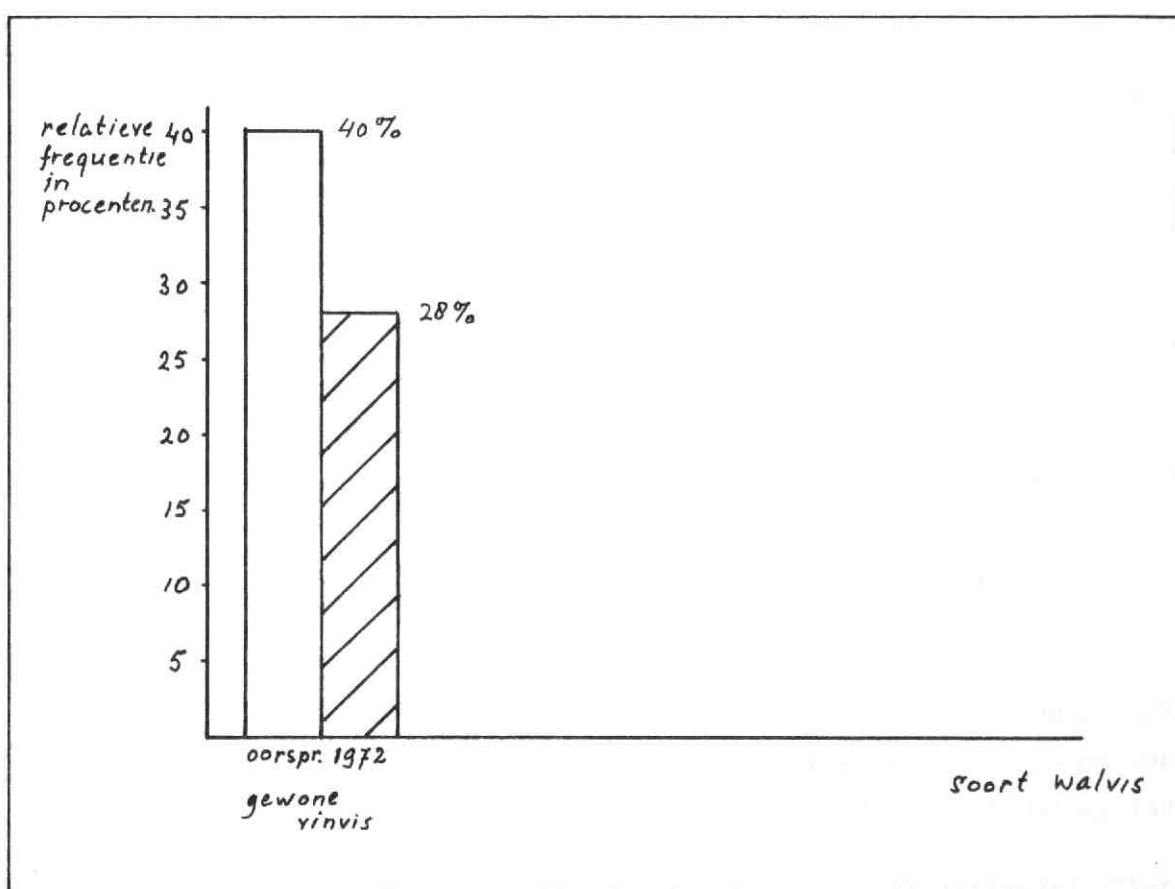
$$\text{relatieve frequentie} = \frac{\text{aantal van bepaalde soort}}{\text{totale aantal}}$$

Dit kun je in procenten uitdrukken. Hoe?

» 21. Maak de volgende relatieve frequentieverdelingstabel af (1972):

SOORT WALVIS	FREQUENTIE	REL.FREQUENTIE	IN PROCENTEN
Gewone vinvis	145.000	$\frac{145.000}{507.300} = 0,29$	rel.freq. x 100% = 29%
Noordse visvis	175.000		
Blauwe vinvis	11.000		
Dwerg walvis	9.000		
Bultrug	6.300		
Grijze walvis	11.000		
Dwerg vinvis	150.000		
TOTAAL	N = 507.300		

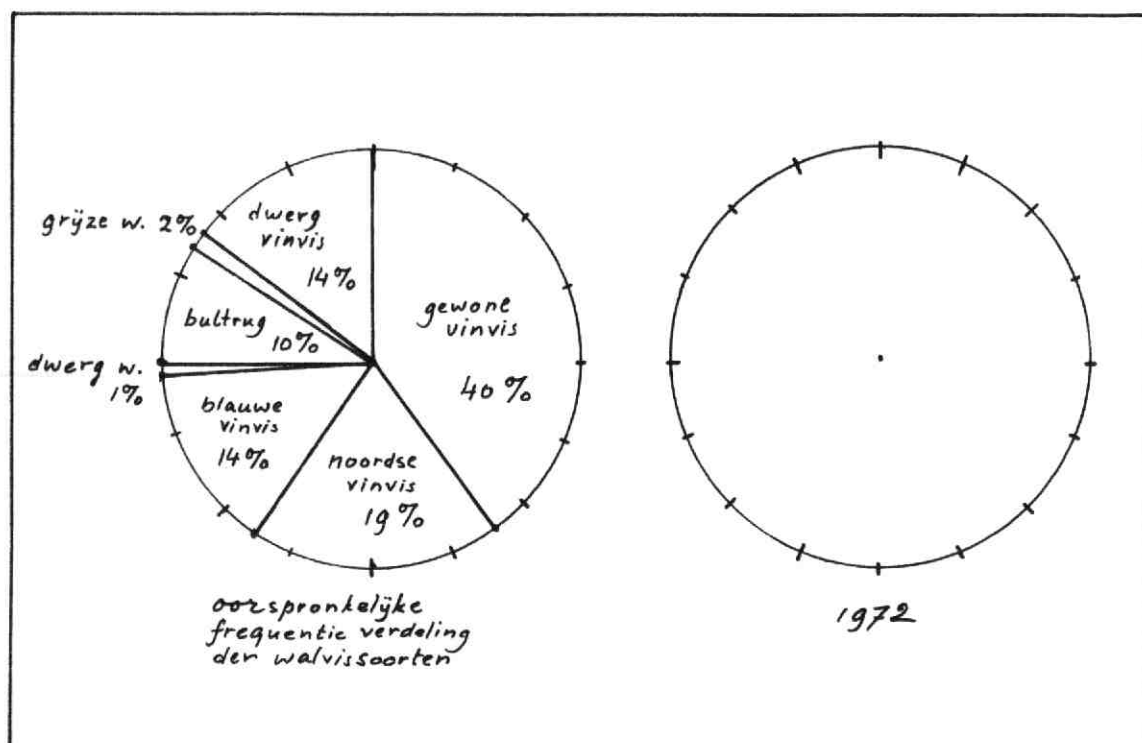
- » 22. Maak eenzelfde tabel voor het oorspronkelijke aantal walvissen.
- » 23. Hoe groot is het aantal nu levende walvissen als percentage van het oorspronkelijke aantal?
- » 24. Maak het volgende staafdiagram af, waarin de relatieve frequenties van de zeven walvissoorten wordt vergeleken (de oorspronkelijke steeds links, de 1972-getallen rechts).



- » 25. Kun je aan de hand van » 23 en » 24 vaststellen welke walvissen het meest ernstig in hun bestaan worden bedreigd?

Een andere - vooral in kranten nogal veel gebruikte - manier om relatieve frequentieverdelingen aanschouwelijk te maken is met behulp van een cirkeldiagram. We hebben de tabel van de oorspronkelijke frequentieverdeling van de walvissen als voorbeeld in beeld gebracht (zie pag. 14). De totale opper-

vlakke van de cirkel komt overeen met 100%. De oppervlakte van de cirkelsectoren komen overeen met de relatieve frequenties.



» 26. Maak zelf het cirkeldiagram voor de relatieve frequentieverdeling zoals die is in 1972.

Ruwweg gesproken waren er oorspronkelijk twee keer zoveel walvissen als nu. Om dat in de diagrammen weer te geven, wordt het linker diagram groter getekend. Vaak maakt men de *straal* dan tweemaal zo groot.

» 27. Vergelijk de oppervlakte van de twee cirkeldiagrammen in dat geval. Kun je een 'eerlijker' voorstelling van zaken bedenken?

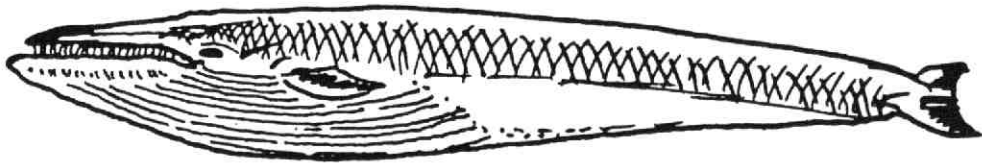
Hiernaast zie je een plaatje waarin gepoogd is het dreigende uitsterven van walvissen duidelijk in beeld te brengen zonder frequentietabellen, staafdiagrammen of cirkeldiagrammen. Bekijk de plaat goed en tracht er achter te komen hoe deze tot stand is gekomen (alle gegevens staan in dit hoofdstuk) en geef enkele voordelen en vooral nadelen van deze manier van tekenen. Kijk vooral kritisch naar de 'eerlijkheid' van de voorstelling.

DE WALVIS

naar de haaien?.....



Blauwe Vinvis



Gewone Vinvis



Noordse Vinvis



Bultrug



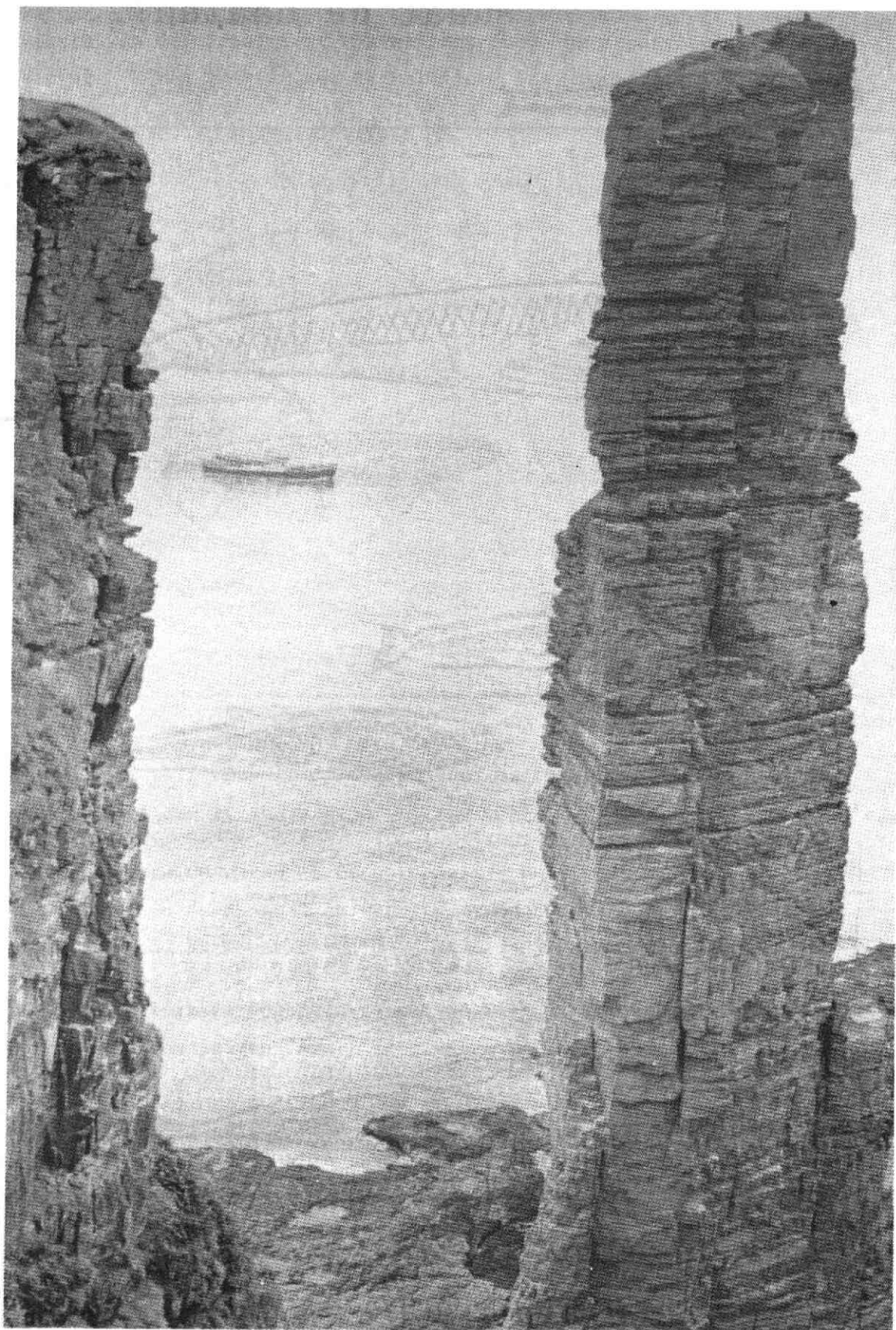
Dwerg Vinvis



Grijze Walvis

TOEN

NU



Rotsen voor de kust van Hoy op de Orkneys

3

HISTOGRAMMEN EN POLYGONEN

De Orkneys is een eilandengroep ten noorden van Schotland. North Ronaldsay, Eday en Papa Westray behoren er toe, zoals je kunt zien op de kaart op blz. 20. Er wonen niet zoveel mensen op deze groene zacht glooiende eilanden, met hier en daar klifkusten. In 1981 was het totale aantal inwoners zo'n 16.500.

De gegevens over de bevolking op drie eilanden: North Ronaldsay, Eday en Papa Westray in tabelvorm:

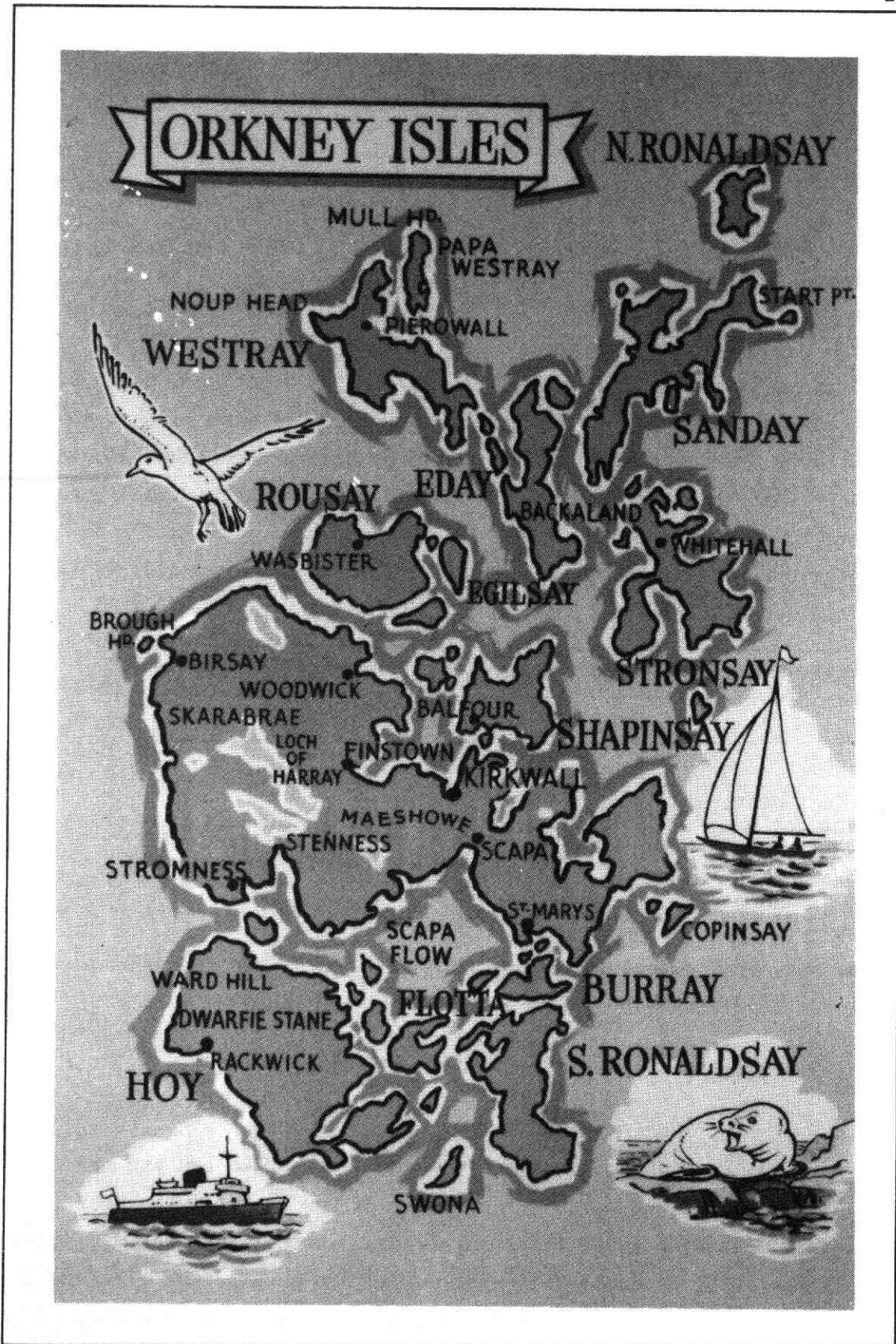
NAAM EILAND	BEVOLKING '71	BEVOLKING '81
North-Ronaldsay	134	109
Eday	170	143
Papa Westray	110	89

(Absolute aantallen).

- » 28. Maak een 'dubbel staafdiagram' van de relatieve frequenties van de bevolking van de drie eilanden in 1971 en 1981 (zoals bij » 24).
- » 29. Welke van de eilanden ontvolkt het snelst?

In de winter van 1980/'81 was er op de Orkneys sprake van een mazelen-epidemie. De ziektegevallen werden van week tot week bijgehouden.

De eerste meldingen (twee gevallen) kwamen binnen in de week die begon op 25 oktober 1980. De daaropvolgende week werden geen nieuwe ziektegevallen gemeld. Vanaf 8 november (week 3) was het echt raak, zoals blijkt uit de volgende absolute frequentietabel:



week	1 25/10	2 1/11	3 8/11	4 15/11	5 22/11	6 29/11	7 6/12	8 13/12
aantal nieuwe patienten	2	0	26	14	47	44	67	34

week	9 20/12	10 27/12	11 3/1	12 10/1	13 17/1	14 24/1	15 31/1
aantal nieuwe patienten	10	0	6	11	2	0	1

Om een goed beeld van het ziekteverloop te krijgen kun je weer een staafdiagram tekenen. Alleen is de verdeling op de horizontale-tijd-as nu wèl aan allerlei voorwaarden verbonden.

Zo zou het erg gek zijn om hier de verschillende staven van plaats te verwisselen omdat er een 'natuurlijke ordening' is door de tijd.

Bovendien zullen in het algemeen in gevallen als deze de staven altijd 'tegen elkaar aangelegd' worden: bij 'ieder' tijdstip behoort een aantal zieken.

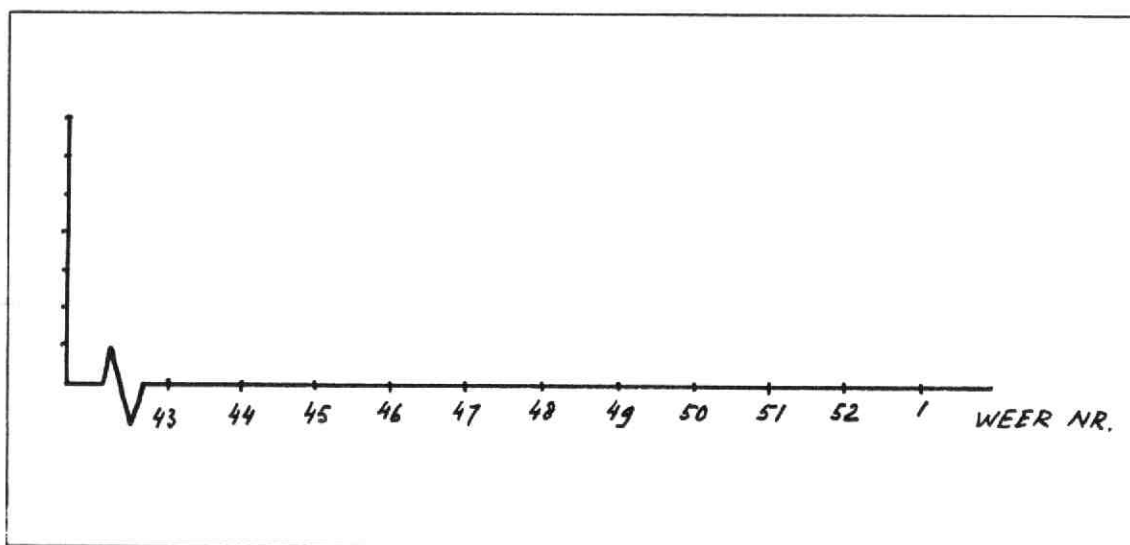
Als een staafdiagram aan bovenstaande voorwaarden voldoet, noemen we het meestal een HISTOGRAM.

» 28. Teken het histogram van de bovenstaande tabel. Neem langs de horizontale as de tijd, met als eenheid: de week.

In het eerste hoofdstuk van dit boekje zag je geen histogrammen, maar grafieken die bestonden uit rechte lijnstukjes. Zo'n grafiek heet een *frequentie-polygoon*.

Een FREQUENTIE-POLYGOON krijg je heel eenvoudig uit een histogram door de middens van de toppen van opeenvolgende staven met rechte lijnen te verbinden.

Voordat je dit gaat proberen nog een opmerking. De week van 25 oktober noemden we week 1, omdat toen het eerste ziektegeval binnenkwam. Van 1980 was het echter de 43e week. Het is echter nogal onhandig om de eerste 42 hokjes niets te tekenen.



Dit los je op door een 'zaagtand' in de horizontale as te tekenen. Zoiets kan desgewenst ook in de verticale as gedaan worden.

- » 29. Teken een frequentie-polygoon van het mazelenverloop in de winter van 1980/'81, te beginnen in de 43e week van 1980.

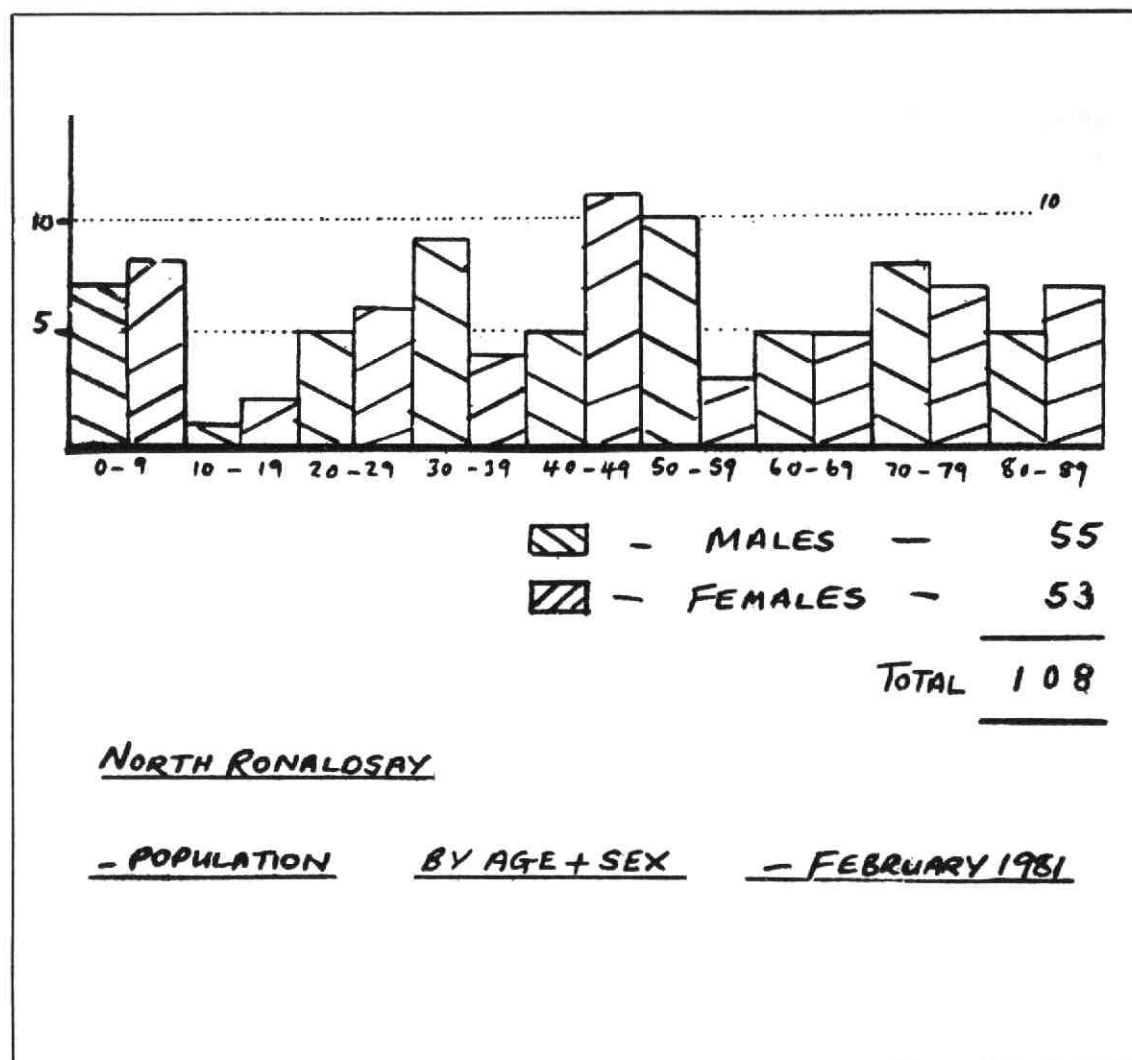
Een voordeel van een weergave met een frequentie-polygoon is dat de 'trend': toename of afname, duidelijker gesuggereerd wordt.

- » 30. In welke periode nam het aantal ziektegevallen het snelst toe?

- » 31. Waarom zullen we van de walvissengrafiek uit opgave » 18 *nooit* een polygoon maken?

Bovendien kun je heel goed meerdere polygonen in een plaatje tekenen om ze te vergelijken. Kijk maar eens naar fig. 2 van het eerste hoofdstukje. Als die grafieken getekend waren als histogrammen, was met name het gedeelte tussen 1950 en 1960 vrijwel niet te ontwarren geweest.

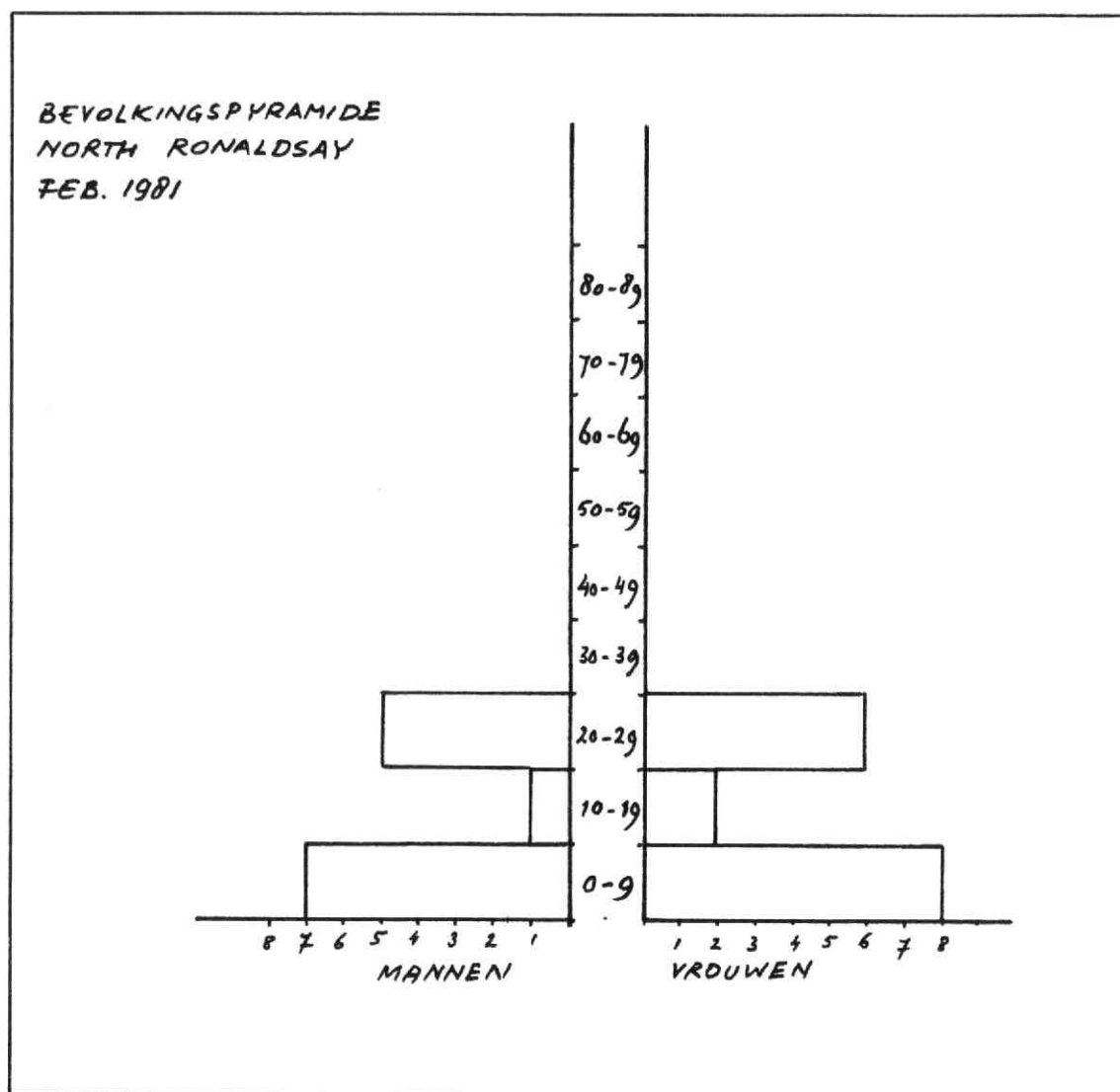
Een histogram kun je ook 'dubbel' tekenen zoals b.v. dit voorbeeld van het al eerder genoemde eilandje North-Ronaldsay:



Je ziet dat de bevolking is ingedeeld in klassen, in dit geval leeftijdsklassen van 10 jaar 'breed'.

- » 32. Wat zijn nadelen van een veel fijnere klassenindeling, b.v. van 1 jaar 'breed'?
- » 33. Maak een soortgelijk histogram als hierboven, maar met leeftijdsklassen van 20 jaar 'breed'. Wat voor nadeel blijkt nu?

Als het om de bevolkingsopbouw gaat wordt vaak voor een afwijkende vorm van het histogram gekozen: de bevolkingspiramide.



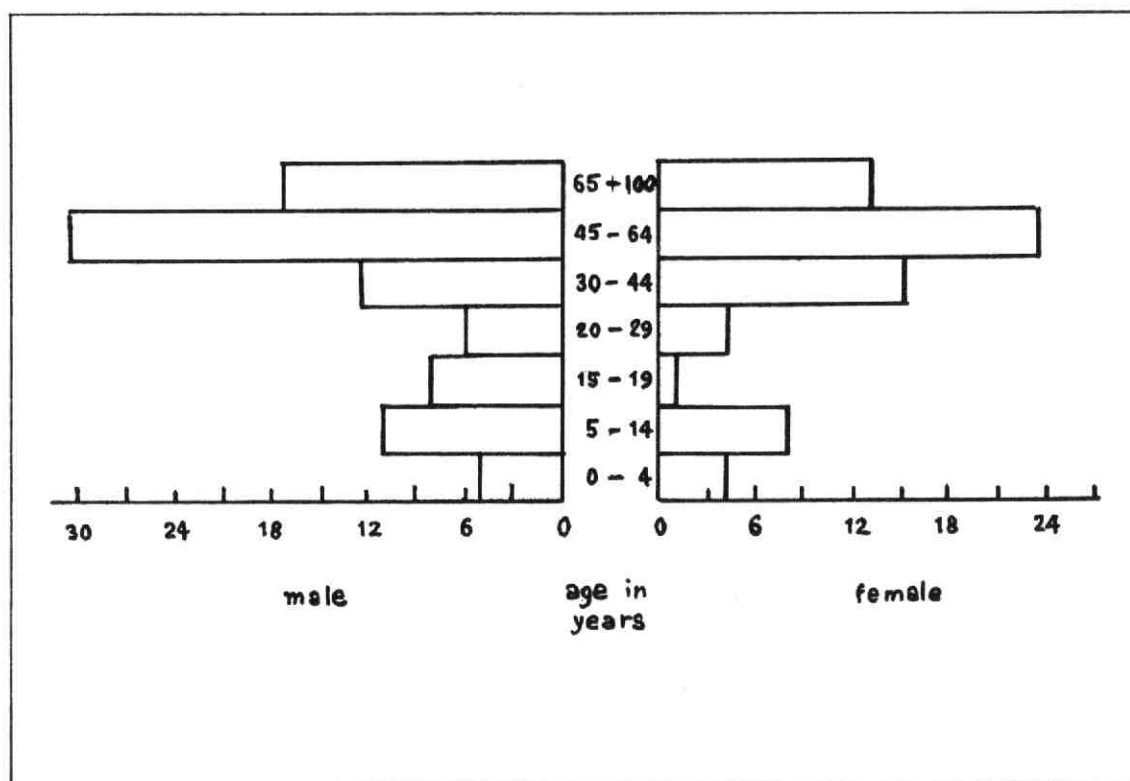
» 34. Maak de piramide af (in je schrift).

Zo'n piramide geeft snel en duidelijk informatie over de opbouw van de bevolking.

» 35. Kun je aan de hand van deze piramide verklaren waarom men de ontwikkelingen voor North-Ronaldsay enige jaren terug met veel zorg tegemoet zag, terwijl het er nu weer wat zonniger uitziet?

Juist omdat de allerjongste groep voor zo'n eiland een heel belangrijke groep vormt, kiest men daarvoor wel eens een smallere klassebreedte. Ook de groep van 15-19 jaar is erg belangrijk, omdat die groep vaak voor verdere scholing van de eilanden vertrekt.

Vandaar dat het niet veel verbazing behoeft te wekken de volgende piramide te vinden:



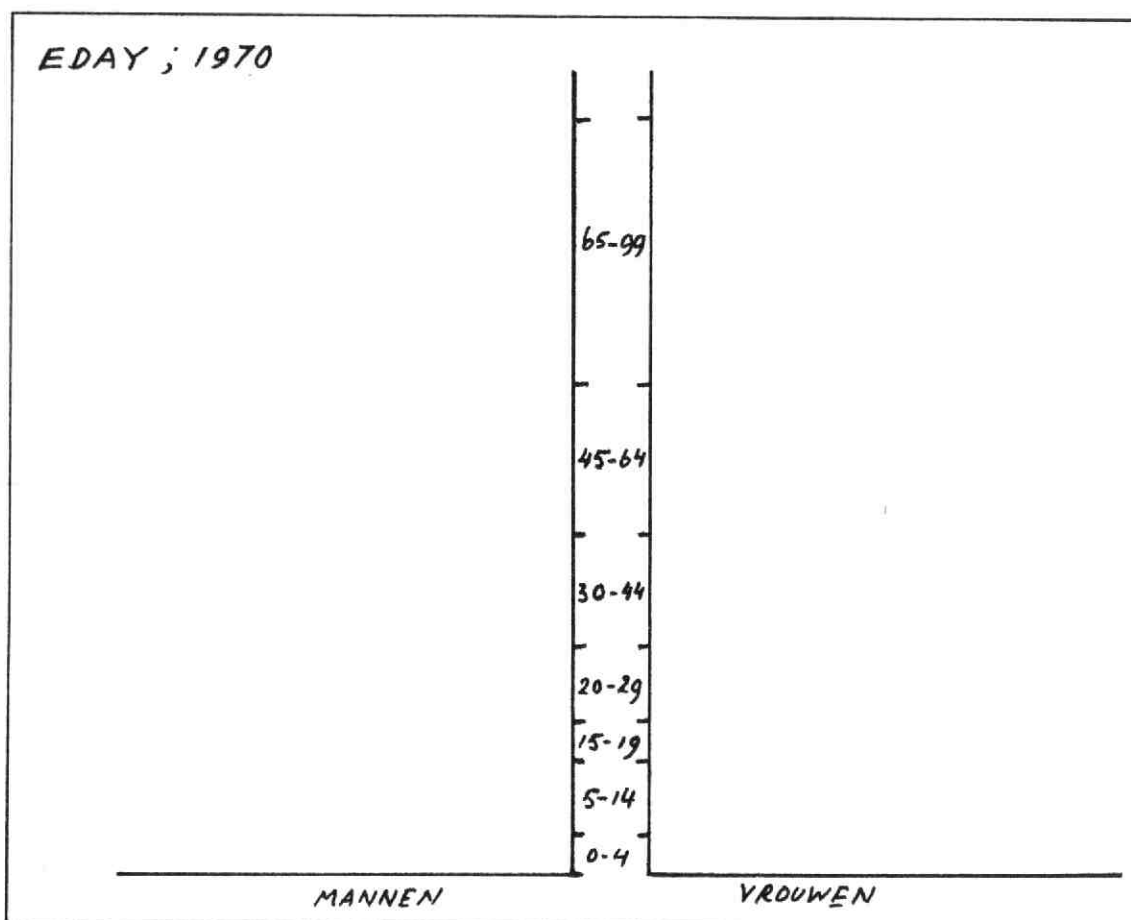
Deze opbouw is van Eday en wel in 1970. Duidelijk is de zorgelijke opbouw van de bevolking. Alhoewel

» 36. Geeft bovenstaande bevolkingspiramide een eerlijk beeld van de bevolkingsopbouw?

Verklaar je antwoord zo goed mogelijk.

Een manier om een eerlijker voorstelling te geven is om het aantal mensen in de klasse te vergelijken met de breedte van die klasse en om de verschillen in de breedte van die klassen ook langs de as te laten blijken.

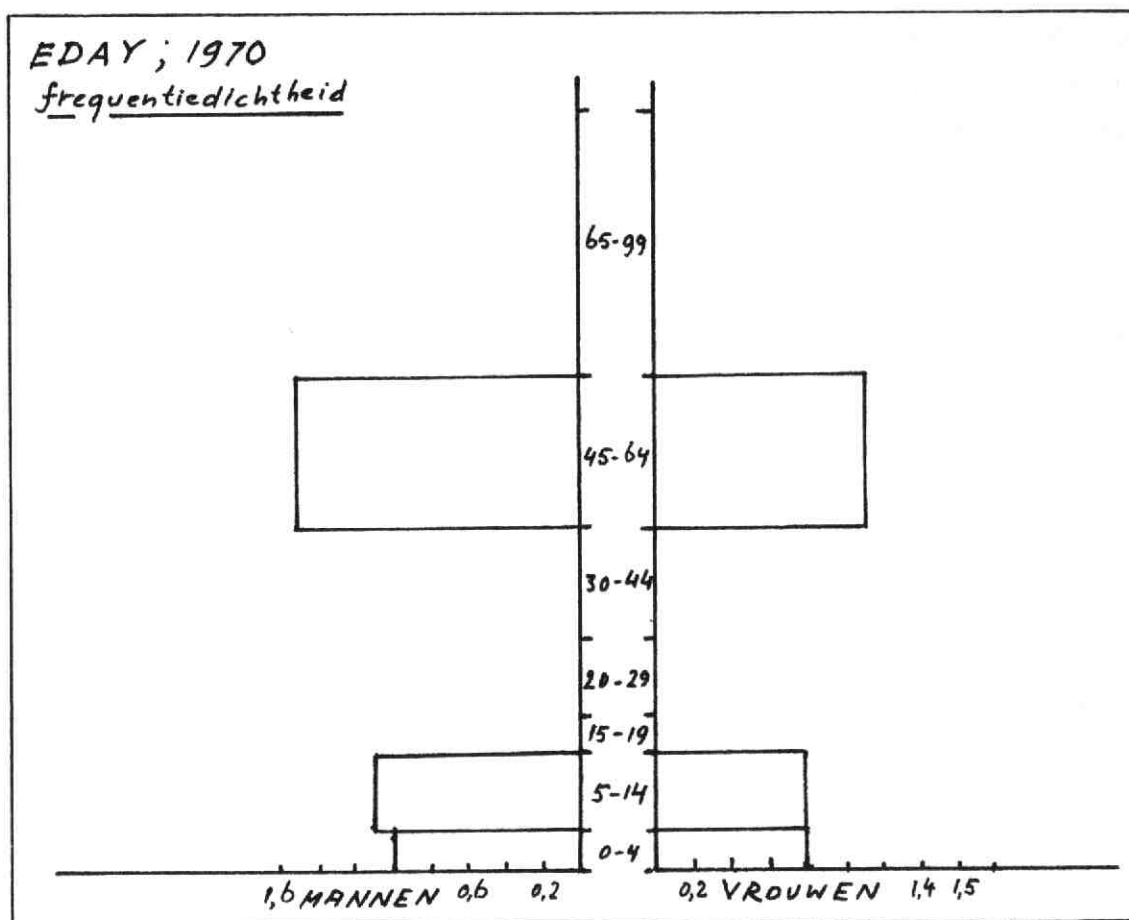
Dus de as-verdeling wordt dan:



$$\text{De frequentiedichtheid} = \frac{\text{frequentie}}{\text{klassebreedte}}$$

0 - 4 : mannen : 5	dus dichtheid:	$\frac{5}{5} = 1$
vrouwen : 4		$\frac{4}{5} = 0,8$
5 - 14 : mannen : 11		$\frac{11}{10} = 1,1$
vrouwen : 8		$\frac{8}{10} = 0,8$
45 - 65 : mannen : 31		$\frac{31}{20} = 1,55$
vrouwen : 23		$\frac{23}{20} = 1,15$

» 37. Bereken de overige frequentiedichtheden zelf en maak de nieuwe piramide af op de volgende bladzijde.



Je ziet grote verschillen met de oorspronkelijke piramide. Ook weer een staaltje van bewuste of onbewuste manipulatie.

Tenslotte:

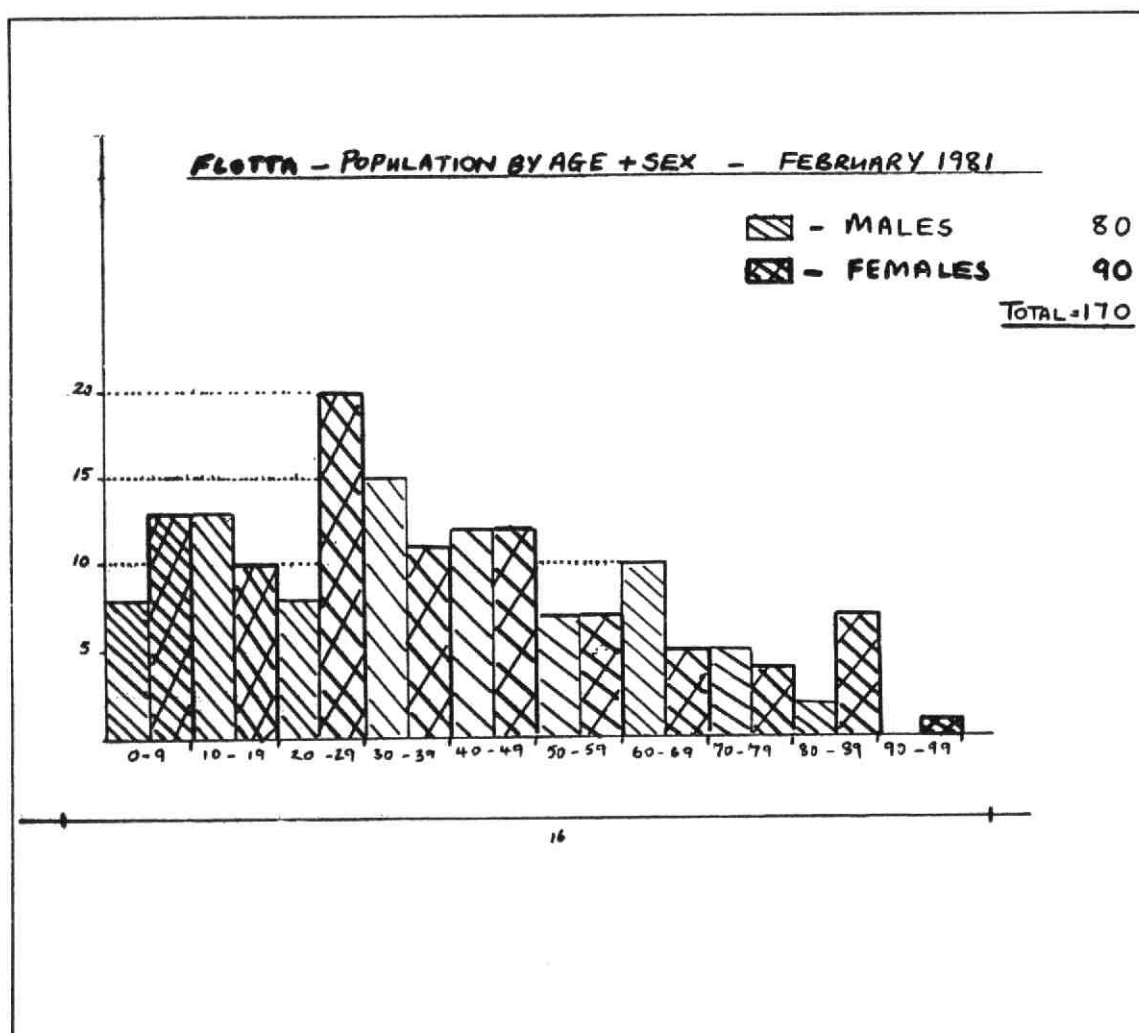
De cijfers van 1981 voor Eday:

leeftijdsgroep	0-9	10-19	20-39	40-59	60-100
mannen	10	8	15	22	22
vrouwen	6	7	16	18	19

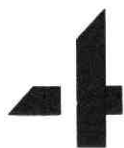
- » 38. Maak de bevolkingspiramide op dezelfde manier als de eerste van Eday in 1970. (Zie blz. 26).
- » 39. Daarna die met de frequentiedichtheden zoals hierboven.
- » 40. Hoe komt het dat je twee frequentiedichthedenpiramides goed kunt vergelijken ondanks geheel verschillende klasse-indelingen?

Op één van de kleinste eilanden van de Orkneys, FLOTTA, is het sinds enige jaren een drukte van belang. Dat komt omdat er een pijpleiding met aardolie vanuit de Noordzee hier aan land komt. Dat heeft voor de Orkney-eilanden economisch een grote betekenis.

Het bevolkingshistogram van Flotta zag er in februari 1982 zó uit:



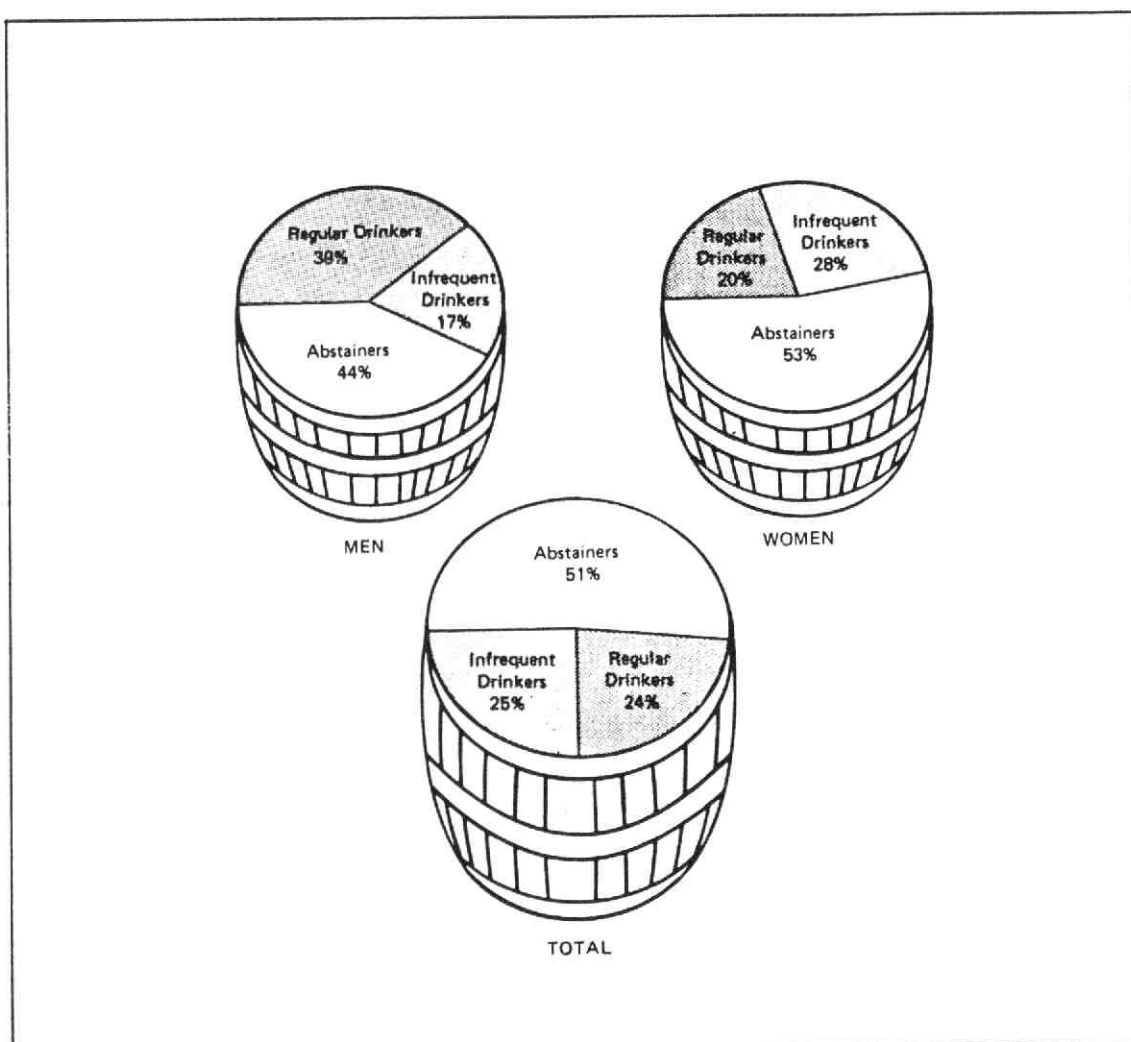
- » 41. Welk belangrijkste verschil kun je constateren tussen dit histogram (en bijbehorende bevolkingspiramide) en die van andere eilanden zoals North Ronaldsay en Eday? Kun je dat verklaren?



GRAFISCHE VOORSTELLINGEN

In de eerste drie hoofdstukjes kwamen verschillende vormen van grafische verwerking ter sprake, zoals staafdiagrammen, histogrammen, polygonen, cirkeldiagrammen, bevolkingspiramides en in het geval van de uitstervende walvissen (blz. 12) een plaatjesdiagram, ook wel pictogram genoemd. Ook werd hier en daar al aangestipt dat je erg moet oppassen bij het kiezen van de manier waarop je die plaatjes tekent en hoe al of niet bewust misbruik gemaakt kan worden van grafieken.

In dit hoofdstukje kijken we naar een serie heel verschillende "grafische voorstellingen" en wat je eraan kunt zien en niet kunt zien.



» 42. Een variatie op het cirkeldiagram.

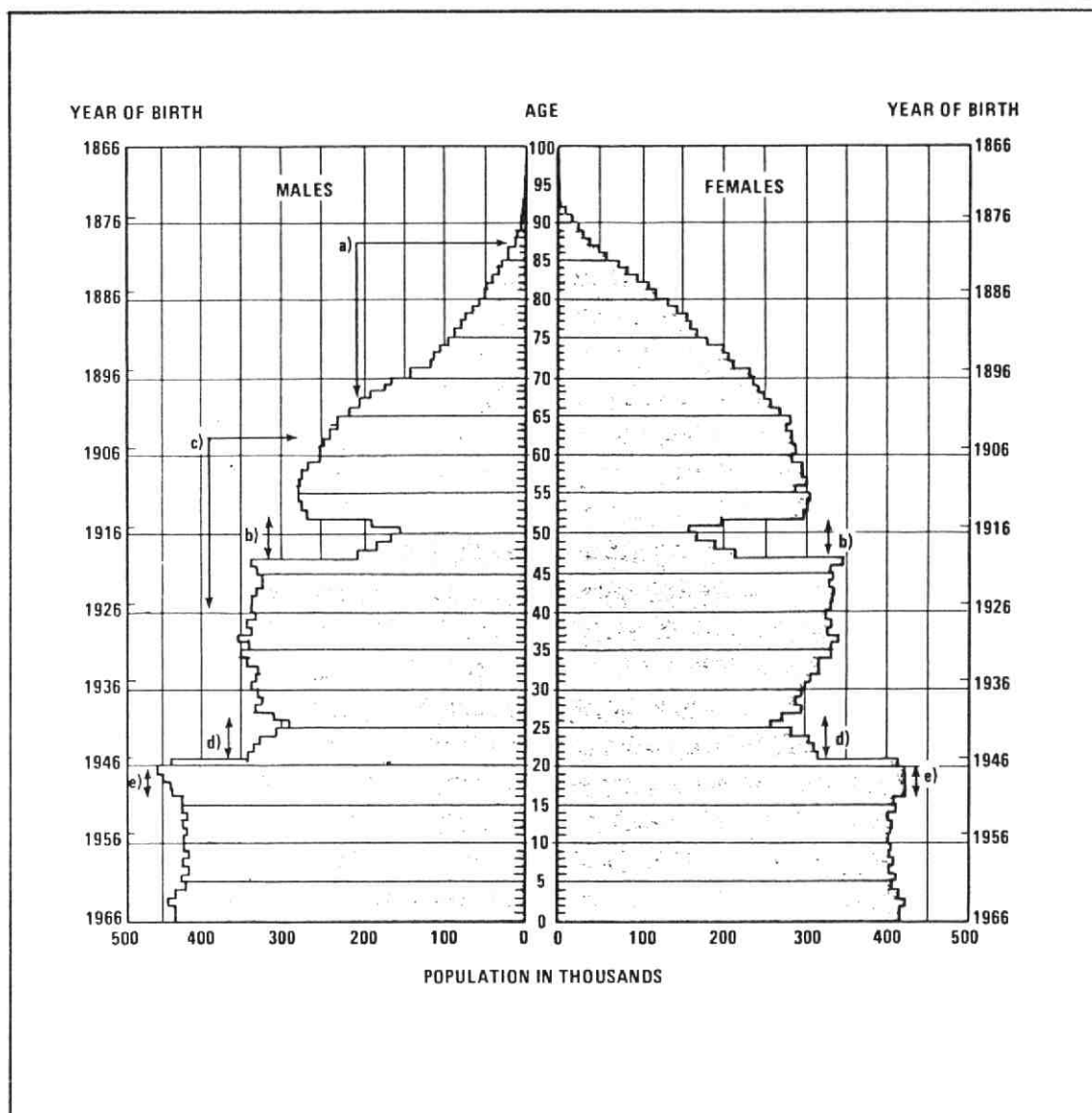
Je ziet drankgewoonten van 65⁺-ers in New York. Daartoe werden zo'n 963 mensen ondervraagd: 483 mannen en 480 vrouwen.

a. Je ziet dat de vaten voor mannen en vrouwen even groot getekend zijn. Het 'totale vat' is wat groter.

Is het totale vat twee keer zo groot als de andere?

b. In New York wonen meer vrouwen dan mannen in de onderzochte leeftijdsgroep. Hoe kun je aan de getallen in het 'gezamenlijke vat' zien of daar rekening mee is gehouden?

Hoe zou men dat kunnen doen?



» 43. De bevolkingspiramide van Frankrijk op nieuwjaarsdag 1967.

- Bij a) en c) zie je duidelijk een 'tekort' aan mannen.

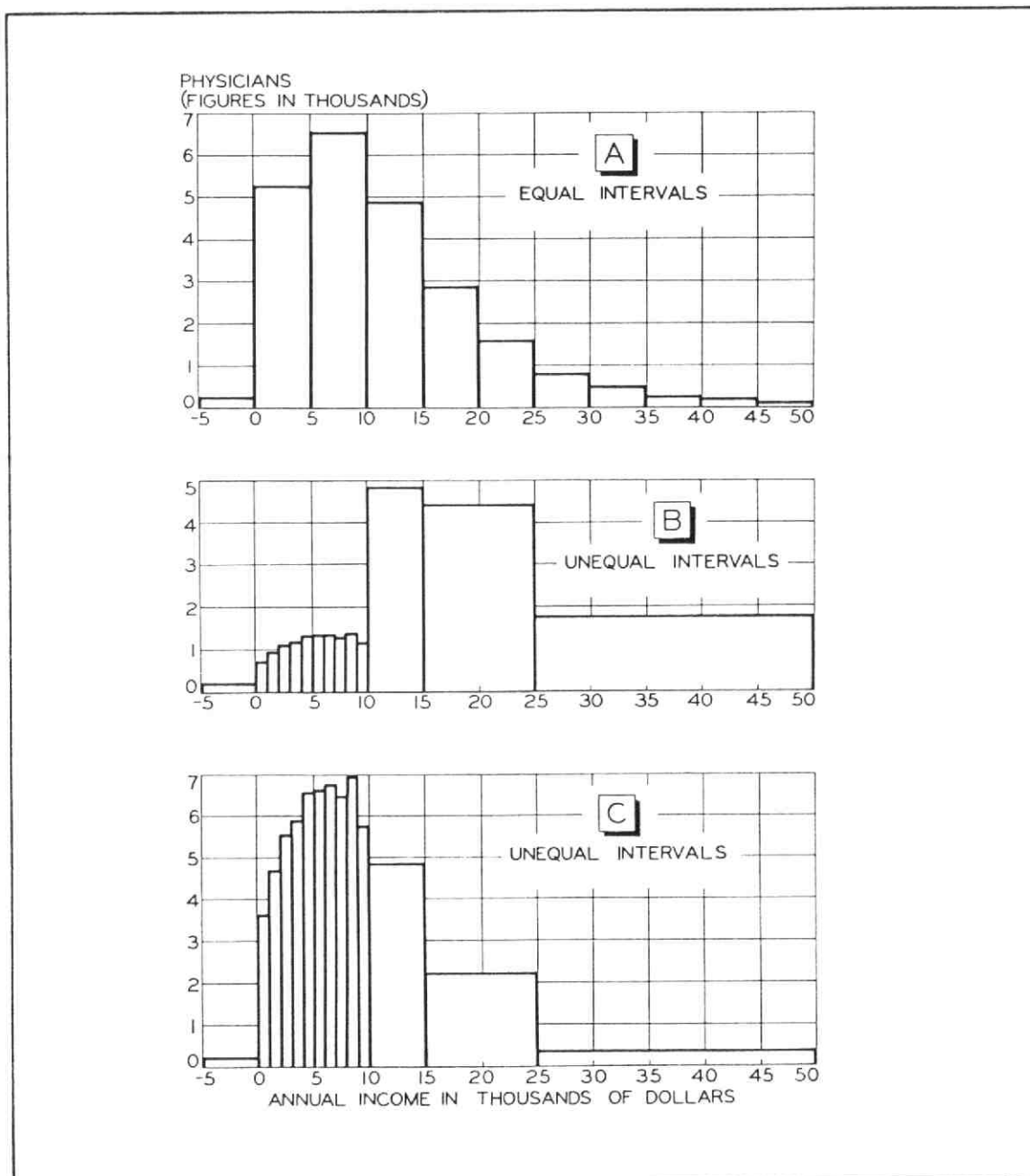
Hoe komt dat denk je?

- Bij b) en d) zie je 'gaten' aan beide kanten.

Kun je dat ook verklaren?

- Bij e) zie je een 'overschot' aan 18-19 jarigen.

Verklaar dat.



- » 44. Het inkomen van artsen in de V.S. in de jaren '40 in beeld gebracht met behulp van drie histogrammen. Eén met gelijke klassebreedte, twee met ongelijke klassebreedte.
- De eerste is goed getekend. Bij de tweede en derde zie je histogrammen met ongelijke klassebreedten.
- Verklaar waarom deze er zó uitzien en wat een bezwaar van ieder van deze voorstellingen is.
- Welke van de twee is in feite een frequentiedichthedengrafiek?

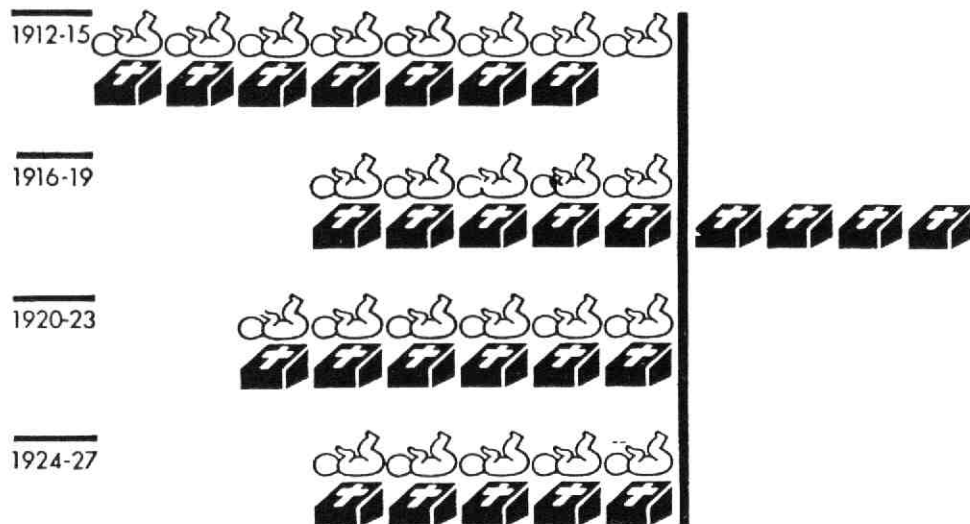
In juni 1983 kwam de Rabobank met een nieuw soort hypotheek (= geldlening): de Rente Stabiel Hypotheek. Een hypotheek die "de schokken opvangt".



De Rabobank Rente Stabiel Hypotheek vangt 30 jaar lang de rente-schokken op.

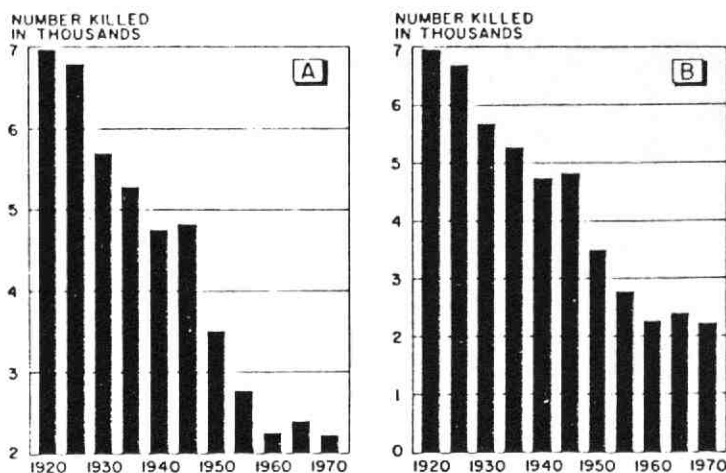
- » 45. In advertenties zoals deze, werd met grafieken duidelijk gemaakt hoe voordelig die rente zou zijn geweest, als hij in januari 1970 zou zijn afgesloten op de toen geldende rente van 9%.
- Verklaar aan de hand van de grafiek - voor zover mogelijk - hoe deze nieuwe hypotheekvorm 'werkt'.
 - Heeft de klant in de 13 jaar van de grafiek echt voordeel gehad van deze nieuwe hypotheek?
 - Bij dalingen van rente zal ook pas na een daling van 2% de stabiel hypotheek worden aangepast.
Geef een voorbeeld (grafisch) waarbij de bank 30 jaar lang méér rente krijgt met deze stabiel hypotheek dan met de gewone hypotheek.

GEBURTEN UND STERBEFÄLLE IN WIEN



Jedes Kind = 20.000 Lebendgeburten Jeder Sarg = 20.000 Sterbefälle

- » 46. De geboorten en sterfgevallen in Wenen nogal beeldend in een pictogram weergegeven.
- Wat geeft de vertikale streep aan en wat beduiden die vier doodskisten rechts van die streep?
Kun je dat ook verklaren?
 - Wat wordt bij deze manier van tekenen duidelijker: een geboorteoverschot of een overlijdensoverschot?
 - Teken twee polygonen in één plaatje met horizontaal de tijd en vertikaal de aantallen geboorten en sterftegevallen.
Verklaar welke voor- en/of nadelen deze voorstelling heeft vergeleken met het pictogram.



» 47. Met de trein reizen is ook niet altijd even veilig geweest zoals blijkt uit bovenstaande histogrammen, waarin weergegeven is hoeveel mensen om het leven zijn gekomen bij spoorwegongelukken in de V.S. De linker suggereert dat het reizen nu wel héél veilig is.

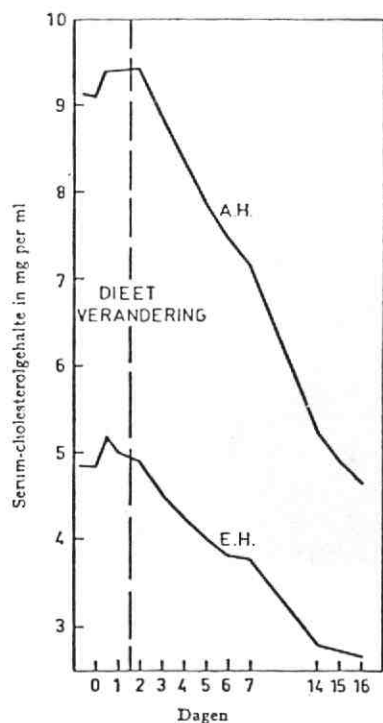
a. Hoe is dat effect bereikt?

b. Hoeveel mensen kwamen er in 1970 per dag om bij spoorwegongelukken?

Tevens suggereert de linker grafiek een veel snellere daling dan de rechter.

c. Hoe is dat effect bereikt?

» 48.



Hiernaast zie je zo spectaculair mogelijk het effect van een dieetverandering op het cholesterolgehalte in beeld gebracht. Vooral bij persoon A.H. is de daling sensationeel. Bij z'n vrouw E.H. is de daling wat minder sprekend, maar daar staat tegenover dat haar gehalte tot minimale proporties lijkt te dalen. Teken deze grafiek opnieuw zó dat duidelijk is:

a. waar het nul-niveau zit.

b. dat de daling veel minder spectaculair is dan nu gesuggereerd wordt.



5

CENTRUMMATEN

Een histogram of polygoon kan gebruikt worden om grote hoeveelheden gegevens overzichtelijk weer te geven. Vaak wordt echter een nog kortere 'samenvatting' gegeven.

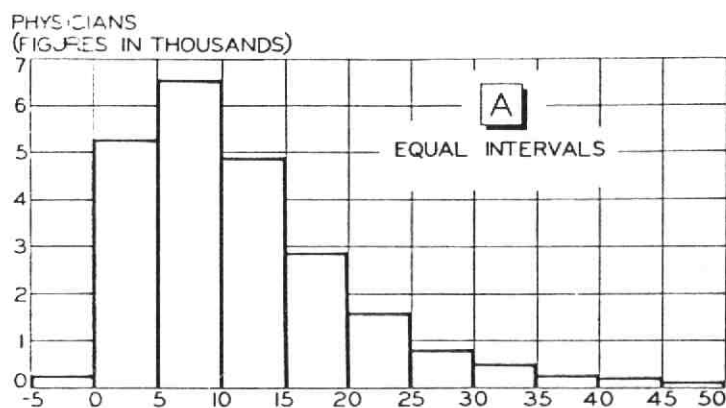
Zo kan de vraag:

"Hoeveel verdienen Amerikaanse artsen?"

beantwoord worden met:

"Gemiddeld zo'n US \$ 11.500,-- (1949).

Het bijbehorende histogram zagen we al eerder. (blz. 32).



- » 49. Teken een histogram van (fictieve) inkomens van artsen in een ander land S. In S verdienen de artsen gemiddeld ook US \$ 11.500,-- per jaar, maar extreem hoge en extra lage lonen komen er niet voor: er is minder 'spreiding' in lonen.

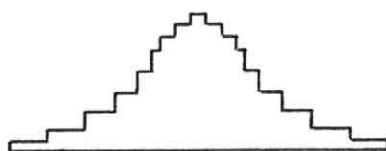
Dit en het volgende hoofdstuk gaan over *gemiddelde* en *spreiding*.
Het gemiddelde in het voorbeeld van blz. 37 is het rekenkundig gemiddelde:

Rekenkundig gemiddelde: tel de waarden van de waarnemingsgetallen op en deel dit door het aantal waarnemingsgetallen.

Of in formule: het gemiddelde van de getallen x_1, x_2, \dots, x_n is:

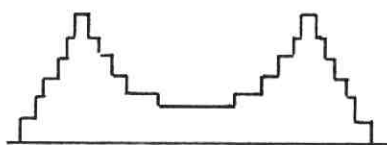
$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} = \frac{x_1 + x_2 + \dots + x_n}{n}$$

- » 50. a. Bereken het gemiddelde van: 60, 59, 67
b. Bereken het gemiddelde van: 560, 559, 567
c. Bereken het gemiddelde van: 0, -1, 7
d. Welke samenhang zit er tussen de antwoorden op de vragen a, b en c?
- » 51. 1350 artsen verdienen (gemiddeld) US \$ 5.500,-- per persoon.
1350 artsen verdienen (gemiddeld) US \$ 6.500,-- per persoon.
Hoeveel verdienen deze 2700 artsen ieder (gemiddeld)?
- » 52. 1400 artsen verdienen ieder (gemiddeld) US \$ 8.500,--.
1200 artsen verdienen ieder (gemiddeld) US \$ 9.500,--.
Hoeveel verdienen deze 2600 artsen ieder (gemiddeld)?
- » 53. a. Een frequentie-histogram heeft het volgende beeld:



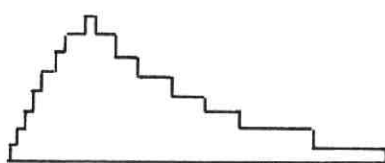
Waar zal het gemiddelde van de waarnemingen liggen?

b. Dezelfde vraag voor de figuur hieronder:



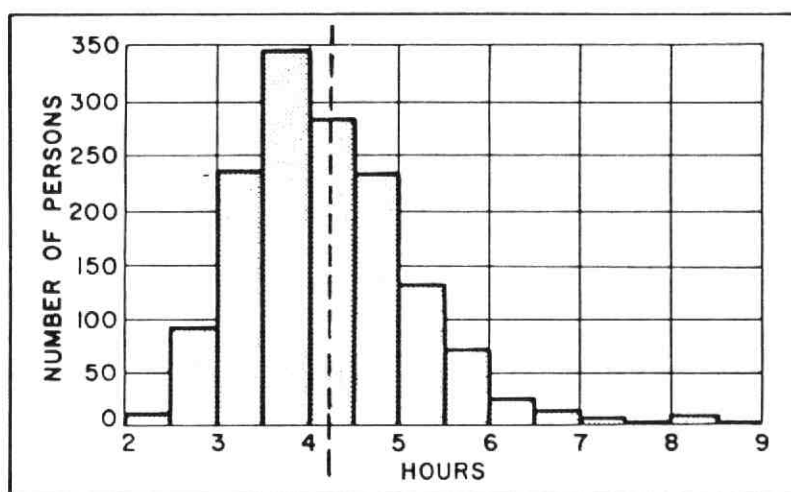
c. Inkomensverdelingen hebben vaak een verdeling als het histogram hieronder.

Waar zal het gemiddelde liggen t.o.v. de plaats van het maximum?



Bij vraag c van opgave » 53 was het niet eenvoudig om het gemiddelde nauwkeurig aan de hand van de grafiek aan te geven.

De grafiek lijkt enigszins op onderstaand histogram:



Dit histogram geeft de resultaten weer van de jaarlijkse Honolulu (Hawaii) marathon van 12 december 1976. De *gemiddelde* tijd van de 1447 deelnemers was 4 uur en 15 minuten (over de 42,195 km). De gemiddelde tijd is in het histogram aangegeven met een stippellijn.

Misschien verwacht je dat de helft van de deelnemers nu sneller dan 4 uur en 15 minuten heeft gelopen, maar dat is niet zo.

822 deelnemers liepen sneller dan 4 u. 15 min. , dat is dus bijna 57%. En 43% liep dus langzamer dan het gemiddelde.

Om uit te zoeken hoe het komt dat die verdeling niet fifty-fifty is, organiseren we een oefenmarathon met maar vier deelnemers.

De eerste drie deelnemers zijn binnengekomen na 3 uur, 3 u. 30 min. en 3 u. 30 min. (afgerond op een kwartier). Er wordt met spanning op de vierde deelnemer gewacht.

» 54. Teken een histogram (klasse: $[2,45 - 3,15]$, $[3,15 - 3,45]$ enz.) van deze marathon voor de volgende gevallen:

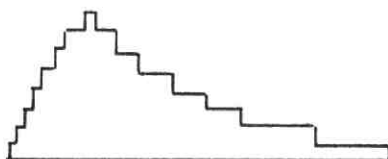
- a. de vierde deelnemer finisht na 4 uur.
- b. de vierde deelnemer finisht na 5 uur.
- c. de vierde deelnemer finisht na 6 uur.
- d. de vierde deelnemer finisht na 7 uur.

Bereken in ieder van deze gevallen de gemiddelde tijd.

Uit het voorbeeld van de marathon met vier deelnemers blijkt duidelijk dat een 'scheef' histogram het gemiddelde 'meetrekt', waardoor het gemiddelde niet altijd te gebruiken is.

In het laatste geval: 3, 3,5; 3,5; 7 waren immers 75% van de deelnemers sneller dan het gemiddelde!

Iets dergelijks heb je als de grafiek aan de 'staart naar rechts' heeft:



Een ander soort centrummaat kan in zo'n geval beter zijn.

De waarnemingen worden gerangschikt van klein naar groot.
De *mediaan* is de waarde van de middelste waarneming.

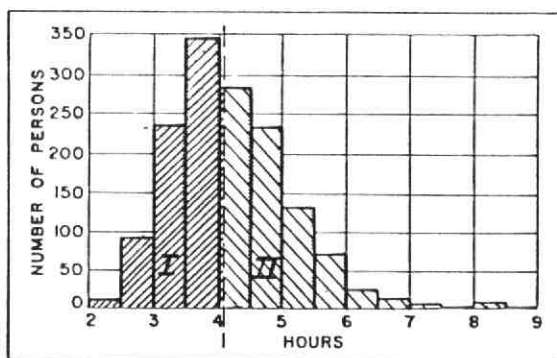
Vb: 2 3 4 4 5. De mediaan is 4.
↑

Als het om een even aantal waarnemingen gaat, b.v. in het marathon-voorbeeld, is er geen middelste. We nemen dan het gemiddelde van de twee rond het midden:

3	3,5	3,5	4	mediaan 3,5
3	3,5	3,5	5	mediaan 3,5
3	3,5	3,5	6	mediaan 3,5
3	3,5	3,5	7	mediaan 3,5

De mediaan laat zich dus niet van de wijs brengen door zijn staart.

De oppervlakte van een histogram wordt door de mediaan in twee even grote delen verdeeld.



oppervlakte I =
oppervlakte II

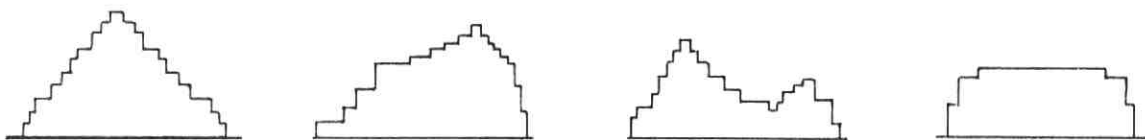
» 55. Verklaar:

Bij een histogram met een staart naar rechts is het gemiddelde groter dan de mediaan.

» 56. Inkomensverdelingshistogrammen hebben vaak een staart naar rechts.

Kijk maar naar het voorbeeld uit het begin van dit hoofdstuk. Er wordt meestal bij inkomens naar de mediaan gekeken en niet naar het gemiddelde. Verklaar dit.

» 57. Bij welk van onderstaande histogrammen is de mediaan groter of kleiner dan het gemiddelde?



» 58. Dezelfde vraag voor het histogram van Flotta. (blz. 28).

Aan een experiment doen 20 mensen mee die een bepaalde opdracht uit moeten voeren binnen één uur. De tijd van ieder van de deelnemers wordt genoteerd. Vier mensen krijgen hun werk niet af, Voor hen wordt één uur als tijd genoteerd.

De resultaten voor wat betreft de tijdsduur per deelnemer:

- twee deelnemers in een kwartier;
- drie deelnemers in een half uur;
- vijf deelnemers in drie kwartier;
- zes deelnemers in één uur;
- vier deelnemers krijgen de proef niet af. (Er wordt één uur genoteerd).

» 59. Verklaar waarom het gemiddelde geen goede indruk van de tijdsduur van het experiment is.

Hoe groot is de mediaan in dit geval?

De onderzoeker die het experiment begeleidt heeft het gevoel - gezien de resultaten van de vier deelnemers die de proef niet binnen de gestelde tijd afkregen - dat een verstandige centrummaat hier ongeveer één uur zou moeten zijn.

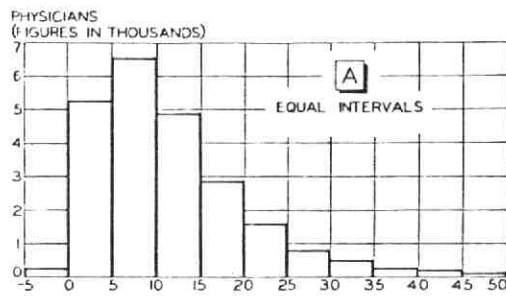
De centrummaat die hij voor ogen heeft is: de modus.

Modus: de waarde met de hoogste frequentie.



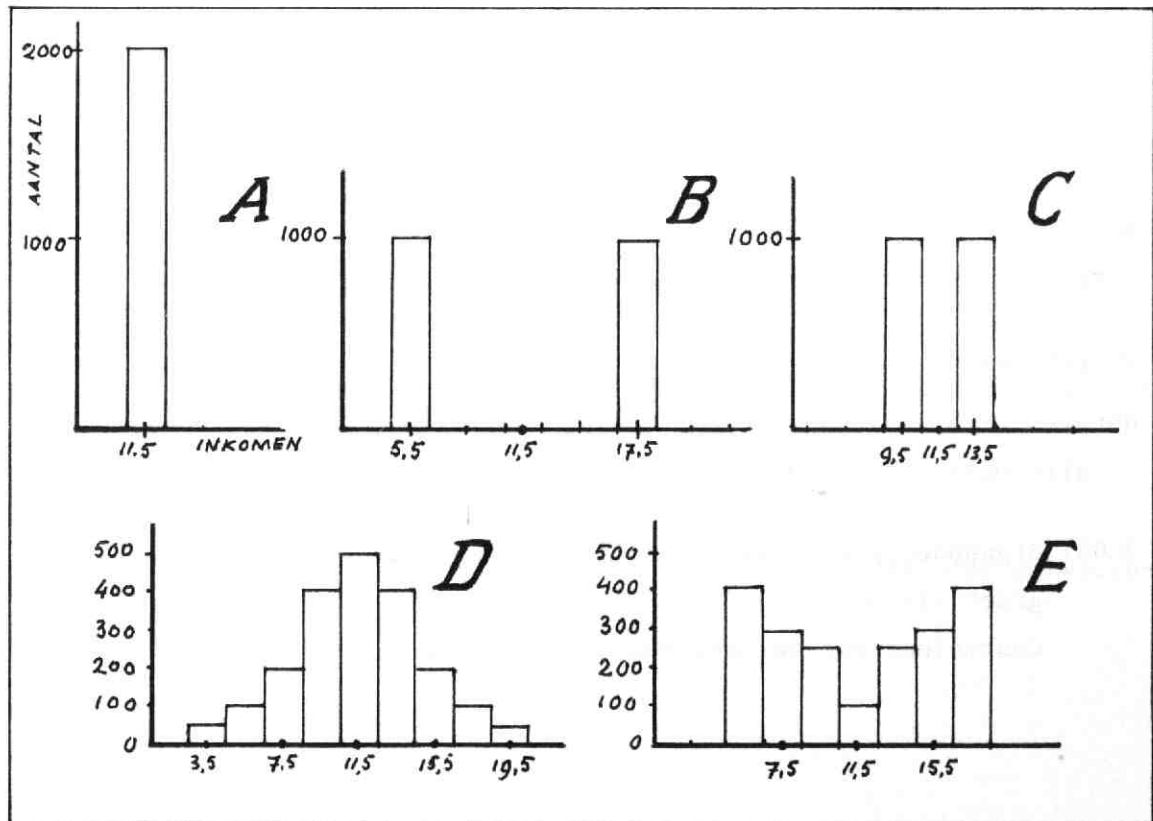
SPREIDINGSMATEN

Nog eens de inkomensverdeling van de Amerikaanse artsen:



Het gemiddelde inkomen was US \$ 11.500,-- per jaar.

In enkele fictieve landen, die hetzelfde gemiddelde hebben, is de verdeling heel anders:



Alle vijf verdelingen hebben hetzelfde gemiddelde.

» 60. Bepaal in elk van de vijf gevallen de mediaan. Wat is je conclusie?

» 61. Bepaal in elk van de vijf gevallen de modus. Wat is je conclusie?

Als het gemiddelde niet voldoende is om een verdeling te karakteriseren, wordt er vaak een *maat* voor de *spreiding* van de waarneming bij vermeld.

» 62. a. In welk van de vijf gevallen is er nauwelijks sprake van enige spreiding?

b. In welk geval is de spreiding het grootst?

Een voor de hand liggende maat voor de spreiding is de 'afwijking van het gemiddelde'.

Bijvoorbeeld in geval B:

1000 waarnemingen wijken ieder - 6000 af van het gemiddelde
(5.500 - 11.500)

1000 waarnemingen wijken ieder + 6000 af van het gemiddelde
(17.500 - 11.500)

De gemiddelde afwijking: $\frac{1000 \cdot -6000 + 1000 \cdot 6000}{2000} = 0.$

» 63. Waarom geldt altijd: de gemiddelde afwijking is gelijk aan 0?

De gemiddelde afwijking is dus niet zo'n geslaagde maat voor de spreiding! Beter zal het gaan als we de gemiddelde absolute afwijking nemen:

De gemiddelde *absolute* afwijking: $\frac{1000 \cdot 6000 + 1000 \cdot 6000}{2000} = 6000.$

Hetgeen prima overeenkomt met het histogram:

alle waarnemingen wijken immers 6000 van het gemiddelde af!

» 64. Afgaande op het histogram lijkt de spreiding bij B driemaal zo groot als bij C.

Controleer met een berekening of dit klopt.

De formule voor de gemiddelde absolute afwijking:

$$\text{g.a.a.} = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n}$$

\bar{x} = rekenkundige gemiddelde; n = aantal waarnemingen;

x_i = i^{de} waarneming.

- » 65. Bepaal de g.a.a. van de verdelingen van D en E en geef een schatting van de g.a.a. van de inkomensverdeling van de artsen.

Alhoewel de g.a.a. een goede maat is voor de spreiding, wordt er weinig gebruik van gemaakt. De g.a.a. is uit wiskundig-statistisch oogpunt niet zo goed hanteerbaar.

Wèl goed hanteerbaar, maar op dit moment wat gekunsteld overkomend, is een andere spreidingsmaat: de *standaard afwijking* of *standaarddeviatie* (S.D.).

In plaats van de *absolute waarde* van de afwijking van het gemiddelde nemen we nu het *kwadraat*.

Dus in voorbeeld B:

$$\frac{1000 \cdot (11.500 - 5.500)^2 + 1000 \cdot (17.500 - 11.500)^2}{2000} =$$

$$\frac{1000 \cdot 6000^2 + 1000 \cdot 6000^2}{2000} = 6000^2.$$

- » 66. Controleer dat het gemiddelde van de gekwadrateerde afwijkingen van C gelijk is aan: 2000^2 .
- » 67. Is het volgende juist:
gemiddelde gekwadrateerde afwijking B = 3 x gemiddelde gekwadrateerde afwijking C?

Alhoewel we zien dat de spreiding van B driemaal zo groot is als die van C (en dit ook tot uiting kwam in de g.a.a.) blijkt het niet te kloppen voor het gemiddelde van de gekwadraterde afwijking:

gemiddelde gekwadraterde afwijking B = 9 x gemiddelde gekwadraterde afwijking.

Dit kunnen we corrigeren door de $\sqrt{\text{g.g.a.}}$ te nemen.

Kijk maar:

$$\sqrt{6000^2} = 3 \cdot \sqrt{2000^2}$$

Dit is nu juist de *standaarddeviatie*.

In formule:

$$\text{S.D.} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

- » 68. Bepaal de SD van de verdelingen van D en E en geef een schatting van de SD van die van de artsen.
- » 69. Elk van de volgende serie waarnemingen heeft een gemiddelde van 50. Zet de series in volgorde van spreiding; dus de serie met de grootste SD eerst, enz.
- I 0, 20, 40, 50, 60, 80, 100
- II 0, 1, 2, 50, 98, 99, 100
- III 0, 48, 49, 50, 51, 52, 100
- » 70. Welk van de twee waarnemingsseries heeft - geschat - de grootste SD? Controleer je antwoord door de SD te berekenen.
- I 5, 5, 6, 7, 7, 8
- II 3, 4, 6, 7, 7, 9
- » 71. Bepaal gemiddelde en SD van de volgende twee series:
- I 11, 13, 14, 15, 17
- II 66, 68, 69, 70, 72

» 72. Er zijn in Nederland regelmatig discussies over salarisverhogingen in centen of procenten.

In centen: iedereen krijgt hetzelfde bedrag per maand erbij b.v.
f 100,--.

In procenten: iedereen krijgt eenzelfde percentage salarisverhoging
b.v. 5%.

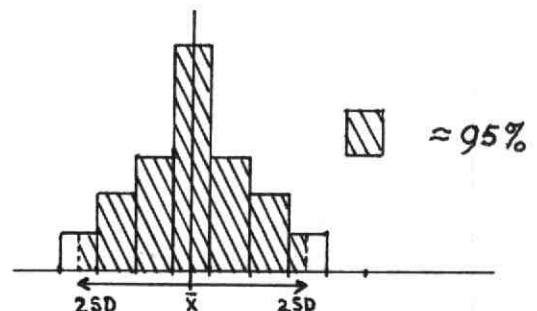
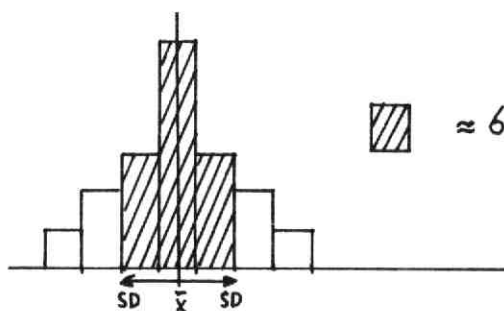
a. Wat voor effect heeft een verhoging 'in centen' op het gemiddelde? En op de SD?

b. Dezelfde vragen voor de verhoging 'in procenten'.

De betekenis van de Standaard Deviatie ligt niet alleen in het feit dat het een redelijke maat is om de spreiding van de waarnemingen aan te geven, maar de SD geeft ook aan hoe *ver* de getallen *afliggen* van het gemiddelde. Meestal is het zo dat ongeveer $\frac{2}{3}$ (68%) van de waarnemingen minder dan $1 \cdot SD$ van het gemiddelde verschilt. En 95% heeft een afwijking van minder dan $2 \cdot SD$ van het gemiddelde.

Voorbeeld:

6	6	6	6	6	6	gemiddelde: 6
5	5	5	7	7	7	SD : 1,5
4	4	8	8	3	9	$2 \cdot SD$: 3,0



Op deze statistische vuistregel wordt later (in de kansrekening) teruggekomen.

SAMENVATTING

- De Standaard Deviatie is een maat voor de spreiding.
- De Standaard Deviatie is als volgt te berekenen:

$$\sqrt{\text{gemiddelde (afwijking van gemiddelde)}^2}$$

of:

$$\sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

- Ongeveer 68% van de waarnemingen heeft een afwijking van minder dan 1•SD van het gemiddelde.
- Ongeveer 95% van de waarnemingen heeft een afwijking van minder dan 2•SD van het gemiddelde.

7

KLASSEBREEDTE

In vorige hoofdstukjes heb je verschillende bevolkingspiramides gezien. Daarbij was de bevolking in leeftijdsgroepen ingedeeld. Soms keurig, alle groepen even breed, b.v. 10 jaar.

Soms, zoals in het voorbeeld van Eday, was de klassebreedte van ieder van de klassen erg verschillend.

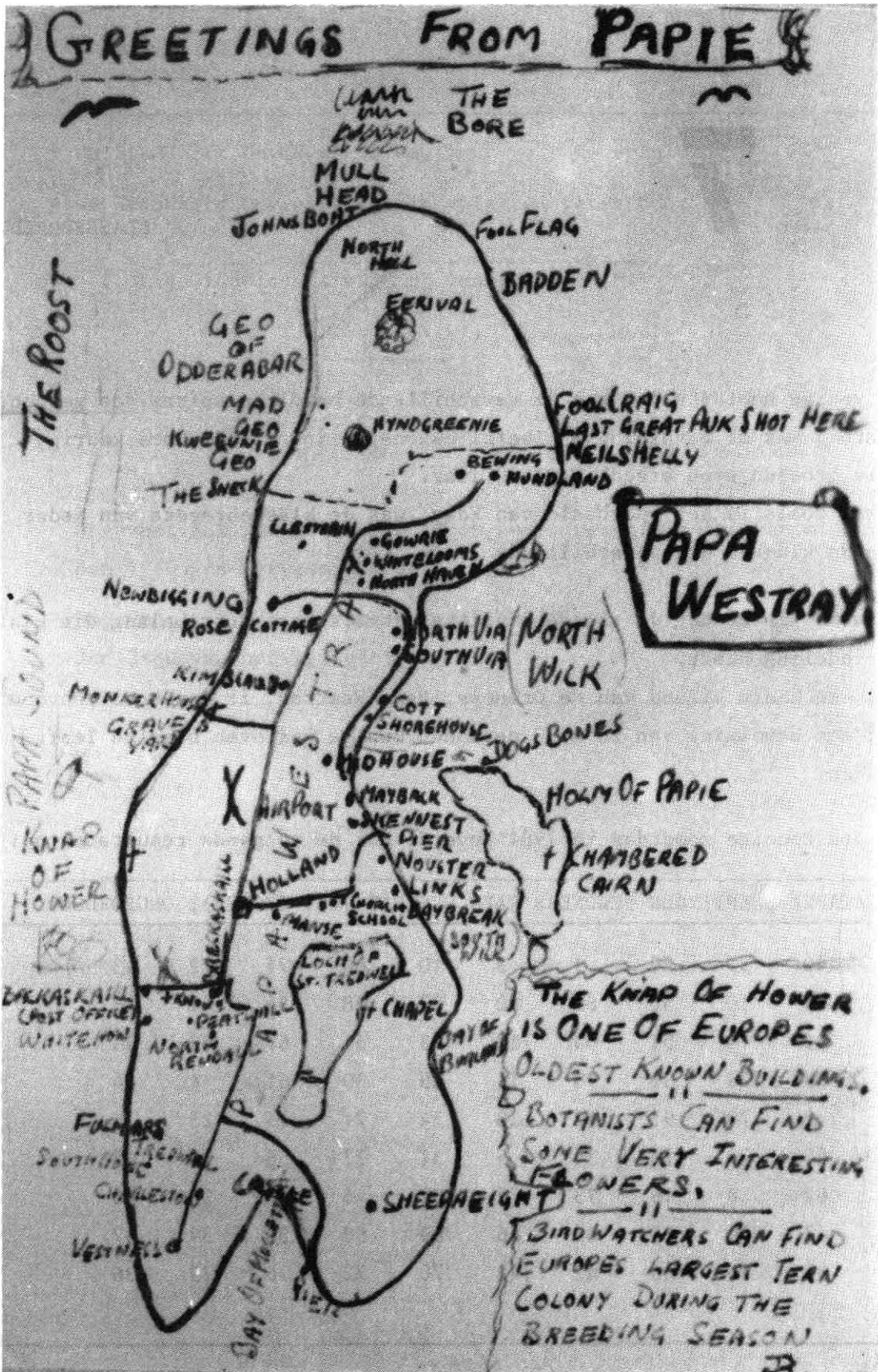
De vraag is nu hoe je aan de grafieken komt en wat voor belang die klasse-indeling heeft.

Bij een klein eiland van de Orkneys, Papa Westray, is dat heel eenvoudig. Bij een bevolking van zo'n 90 personen kun je wel even naar de leeftijd vragen.

Op een zonnige zomerdag in 1981 leverde dat de volgende resultaten op:

DATATABEL LEEFTIJDEN BEWONERS PAPA WESTRAY 1981 (ZOMER); ONGEORDEND

40	51	20	1	3	10	60	71	72	43
42	10	50	85	80	27	28	54	41	18
19	11	51	64	87	13	51	67	68	75
73	23	30	73	14	25	30	61	4	64
8	70	97	62	28	54	71	16	23	6
55	78	78	34	5	31	77	56	55	92
82	8	16	73	77	57	44	44	35	36
47	49	58	69	66	69	58	58	59	37
39	35	46	42	42	70	72	76	79	36



Papa Westray mag dan een klein eilandje zijn, het heeft wel een eigen postbode die ook nog deze ansichtkaart ontwierp.

Misschien wat verwarrend zo'n tabel. Daar kan wat aan gedaan worden. Eén idee is bij iedere leeftijd te kijken hoeveel maal die voorkomt in de tabel. Dus de frequentie van 1, van 2, enz. te bepalen.

» 73. Waarom is dat in dit geval niet zo'n briljante gedachte?

Een beter idee lijkt om de leeftijden maar direct in te delen in *klassen*. Iedere klasse bestaat uit een aantal opeenvolgende waarden. Vaak worden daarbij de volgende gedachten in het achterhoofd gehouden:

1. de klassebreedte wordt *in principe* voor iedere klasse gelijk genomen;
2. het totale aantal klassen wordt in het algemeen tussen de 6 en 20 genomen.

Voordat we de frequentietabel voor de gegroepede leeftijden gaan maken kijken we eerst naar een geschikte groepsbreedte.

Vijf (jaar) lijkt een goede keus, gezien ook de grafieken van het vorige hoofdstuk. Kleiner dan vijf kwam daar niet voor en bovendien kun je de breedte heel gemakkelijk groter maken, aan de hand van je tabel, door twee naast elkaar liggende klassen tot één klasse van tien jaar te maken bijvoorbeeld.

Als de breedte 5 wordt krijgen we in principe 20 klassen, als volgt:

klasse	klassegrenzen	waarden
0 - 4	$0 \leq l < 5$	0, 1, 2, 3, 4
5 - 9	$5 \leq l < 10$	5, 6, 7, 8, 9
10 - 14	$10 \leq l < 15$	10, 11, 12, 13, 14
⋮	⋮	⋮
95 - 99	$95 \leq l < 100$	95, 96, 97, 98, 99

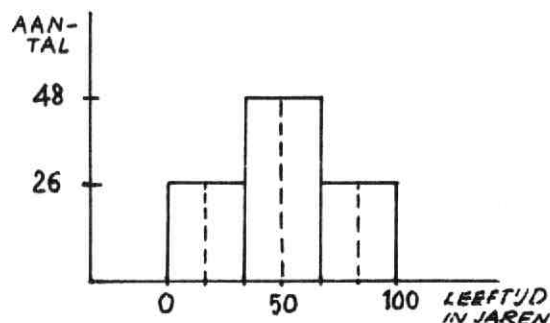
- » 74. Maak een 'turftabel' van de gegevens van Papa Westray met de klasse-indeling van blz. 51; direct aansluitend een frequentietabel.
- » 75. Maak ook een frequentietabel met klassebreedten van 10, 20 en 50.
- » 76. Maak in één plaatje frequentie-polygonen van de tabel op blz. 51 en van bovenstaande twee tabellen en beoordeel de resultaten.

De vraag is wat het kiezen van een klassebreedte voor effect heeft op het gemiddelde en standaarddeviatie.

De gemiddelde leeftijd uitgaande van de tabel op blz. 49 is 42,31 jaar.

- » 77. Dit lijkt erg precies, maar dat is maar schijn. Waarom?

Bij een klassebreedte van 33,3 jaar ziet het histogram er zó uit:



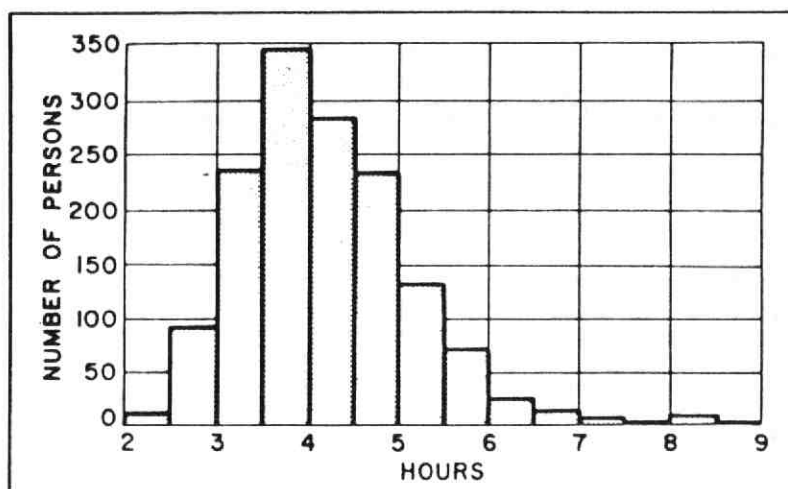
- » 78. Wat is de gemiddelde leeftijd uitgaande van dit histogram?

De berekening van het gemiddelde en standaarddeviatie bij zulke grote aantallen waarnemingen is een tijdrovende zaak. Automatische gegevensverwerking maakt het mogelijk om zowel tabellen, grafieken, gemiddelden, als standaarddeviatie snel op papier of scherm te krijgen.

Als de computer het juiste standaardprogramma bevat, kun je volstaan met het invoeren van de waarnemingen en vervolgens de gewenste klassebreedte. Daarna kun je in hoog tempo de verschillende histogrammen, gemiddelden en S.D. met elkaar en de originele gegevens vergelijken.

- » 79. Verwerk de Papa Westray-gegevens met de computer. Probeer verschillende klassebreedten en besluit welke de beste is. Kijk nauwkeurig of er verschillen zijn tussen je computergrafieken en de reeds door jou getekende bij opgave » 76.

- » 80. Je weet dat het gemiddelde van de Hawaii-marathon (12 december '76) 4 u. 15 min. was en dat er 1447 deelnemers waren.
Schat - vrij nauwkeurig - aan de hand van de grafiek de S.D.



- » 81. Tijdens een marathonloop (over ruim 42 km) worden door de deelnemers de volgende tijden gemaakt: (in uren en minuten).

2.26	3.35	4.33
2.31	3.45	4.37
2.33	3.45	4.42
2.37	3.47	4.42
2.44	3.50	4.42
2.48	3.56	5.01
3.01	3.57	5.09
3.03	3.58	5.17
3.07	3.59	5.31
3.09	4.05	5.46
3.15	4.09	6.32
3.15	4.15	
3.16	4.15	
3.18	4.25	
3.20	4.26	
3.25	4.27	
3.29	4.29	

Voer de gegevens in de computer in en maak verschillende staafdiagrammen, te weten met klassebreedten van:

10 min., 15 min., 30 min., 1 uur; bereken het gemiddelde en de S.D. Vergelijk de resultaten met die van een marathon die op 12 december 1976 op Honolulu werd gehouden.

Regelmatig wordt ook gewerkt met een iets andere tabel: de SOMFREQUENTIE-TABEL. Deze tabel kun je snel uit die van opgave $\gg 74$ als volgt halen:

SOMFREQUENTIE TABEL INWONERAANTAL PAPA WESTRAY:

0 - 4	...	→	Hier komt hetzelfde in als bij $\gg 52$.
5 - 9	...	→	Hier komt het aantal inwoners in dat in deze (5-9) klasse zit <i>plus</i> het aantal uit de klasse (0-4)
10 - 14	...	→	Hier komt het aantal inwoners in dat in deze (10-14) klasse zit <i>plus</i> het aantal uit de voorgaande klassen: (0-4) en (5-9). enz.

$\gg 82$. Waarom volgt uit het voorgaande onmiddellijk dat de somfrequentie voor de laatste klasse: (95,99) gelijk is aan 90?

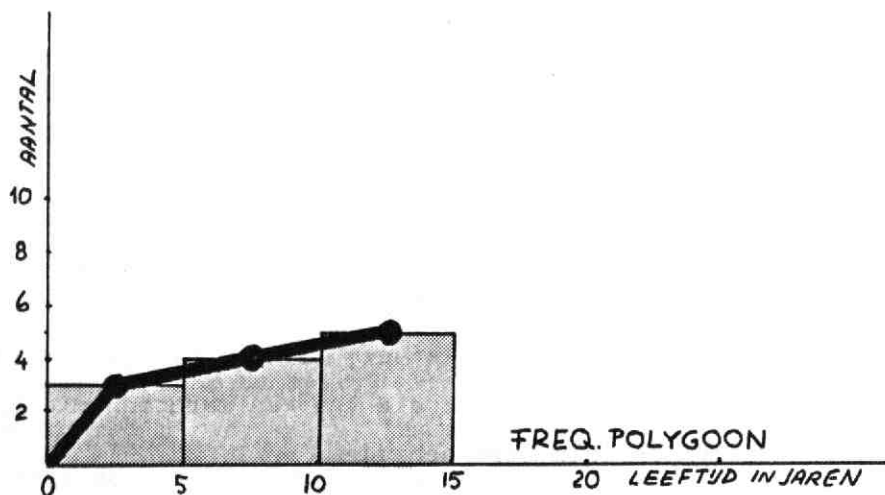
$\gg 83$. Maak de somfrequentietabel van Papa Westray.

De vraag is of je van zo'n somfrequentietabel gemakkelijk een histogram of polygoon kunt maken.

Dat kan inderdaad, maar er is een *belangrijk verschil* tussen een 'gewoon' frequentiepolygoon en een somfrequentiepolygoon.

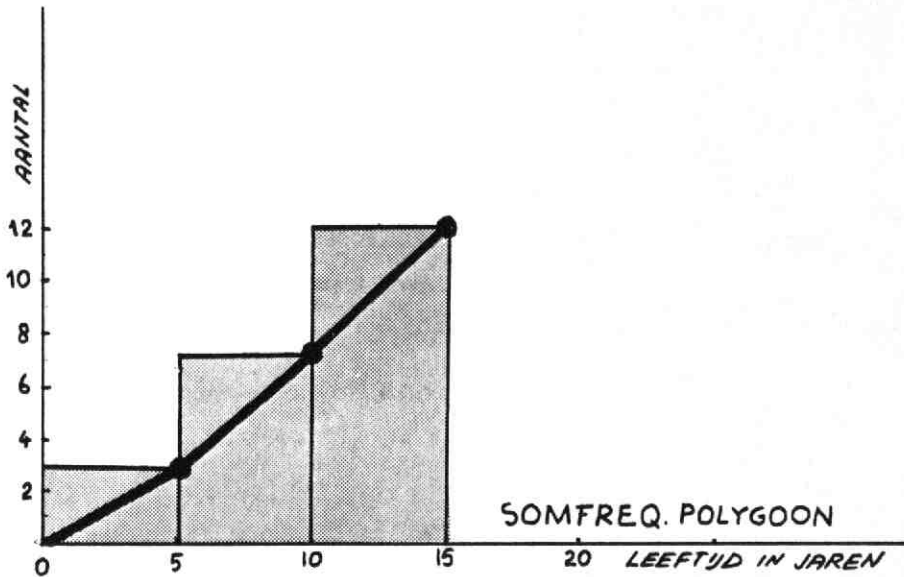
Bij een frequentiepolygoon zoals in $\gg 76$ zette je het punt *midden boven* de klasse.

Dus:



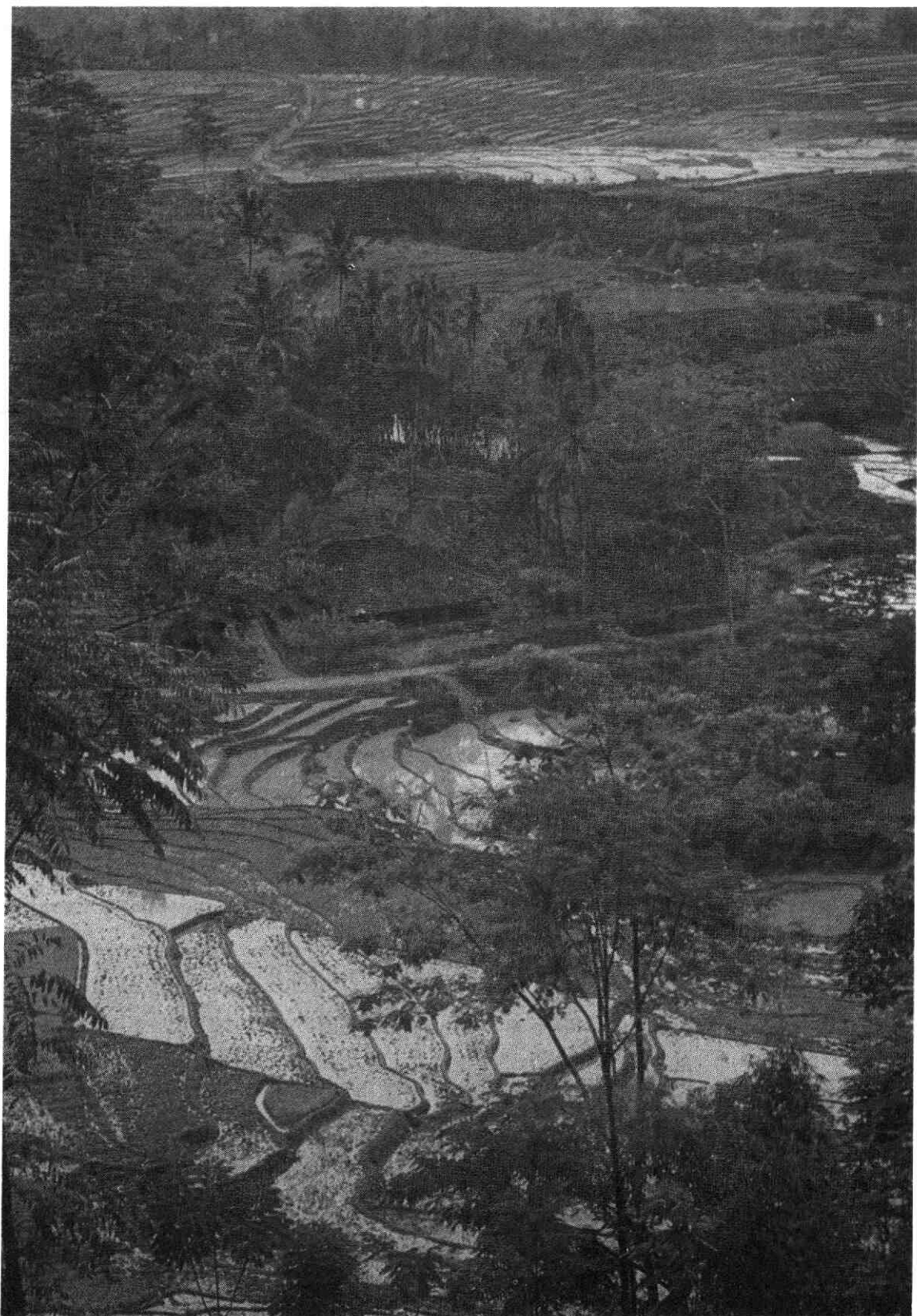
Bij een somfrequentiepolygoon gebeurt dat niet. Daar wordt de waarde van de somfrequentie *boven de klassegrens* gezet.

Te beginnen boven de bovengrens van de eerste klasse:



- » 84. Maak de somfrequentie-polygoon van Papa Westray af.
- » 85. Kun je aan de hand van de somfrequentiepolygoon bepalen hoeveel personen jonger dan 30 zijn in Papa Westray?
Hoeveel % van de bevolking is dat?

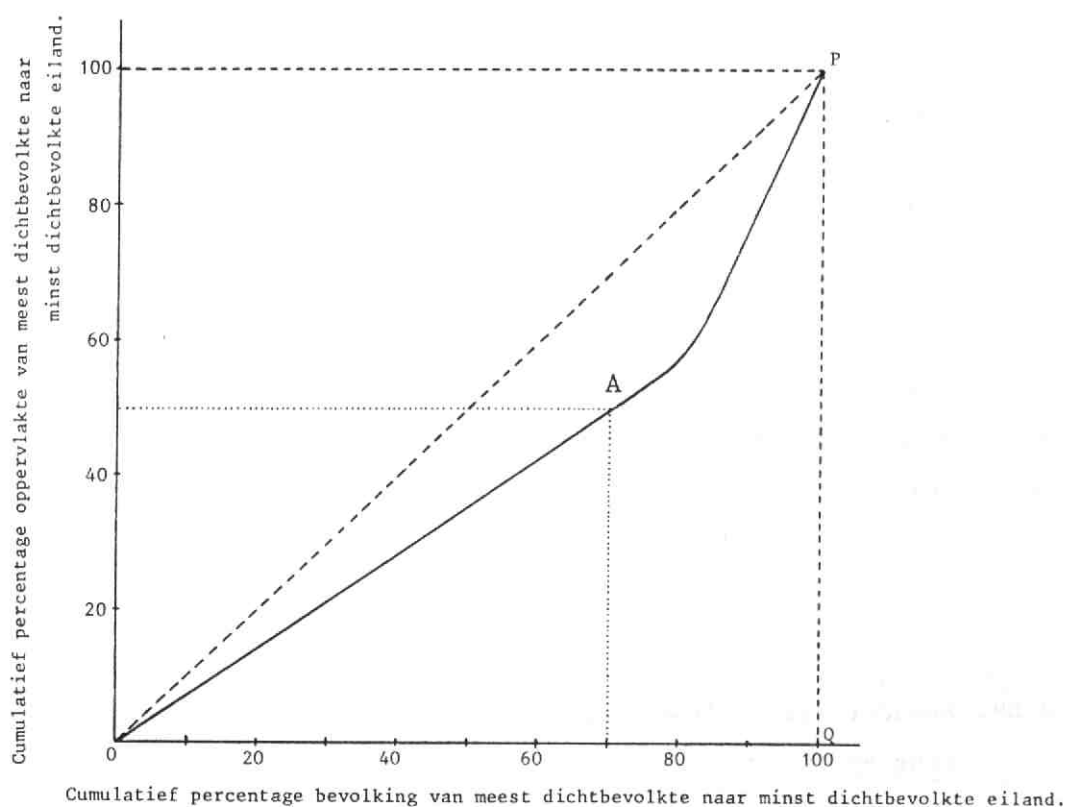
Over een bijzonder soort somfrequentie- of cumulatieve grafiek gaat het volgende hoofdstuk.



Sawahs op Java, het eiland van Indonesië waar 'vrijwel' de gehele Indonesische bevolking woont (zie laatste gedeelte van hoofdstuk 8).



LORENTZ-KROMMEN



Om te onderzoeken of de bevolking van een land of gebied gelijkmatig over dat land verdeeld is, wordt een bijzonder soort cumulatieve- of somfrequentiegrafiek getekend: de *Lorenz-kromme*. Hierboven zie je zo'n kromme voor de Orkneys (de ononderbroken kromme). Je kunt hieruit o.a. aflezen dat 70% van de bevolking op nog geen 50% van de totale oppervlakte van de eilanden woont (punt A).

Zo'n Lorenz-kromme wordt altijd zó getekend dat hij aanvankelijk vrij vlak loopt en naar het eind toe steeds steiler. Dat houdt in dat je eerst naar het eiland kijkt met de grootste bevolkingsdichtheid, daarna naar het op één na dichtstbevolkte eiland enz.

» 86. Verklaar dit.

» 87. Als de grafiek zou samenvallen met de stippellijn OP, wat kun je dan zeggen over de verdeling van de bevolking over de eilanden?

» 88. Hoe zal de Lorenz-kromme eruit zien van een land waarbij bijna de hele bevolking woont in enkele grote steden?

Teken de Lorenz-kromme van Uruguay en Joegoslavië in één tekening als je het volgende weet:

- In Uruguay woont 60% van de bevolking op 4% van alle oppervlakte.
- In Joegoslavië woont 40% van de bevolking op 34% van de oppervlakte en 70% op 65%.

Conclusies?

Terug naar de Orkneys. De kromme en de diagonaal sluiten een gebied in dat kleiner is naarmate de verdeling gelijkmatiger is. Als we de oppervlakte van dat gebied aanduiden met opp. OAP, dan is de GINI-coëfficiënt G als volgt te definiëren:

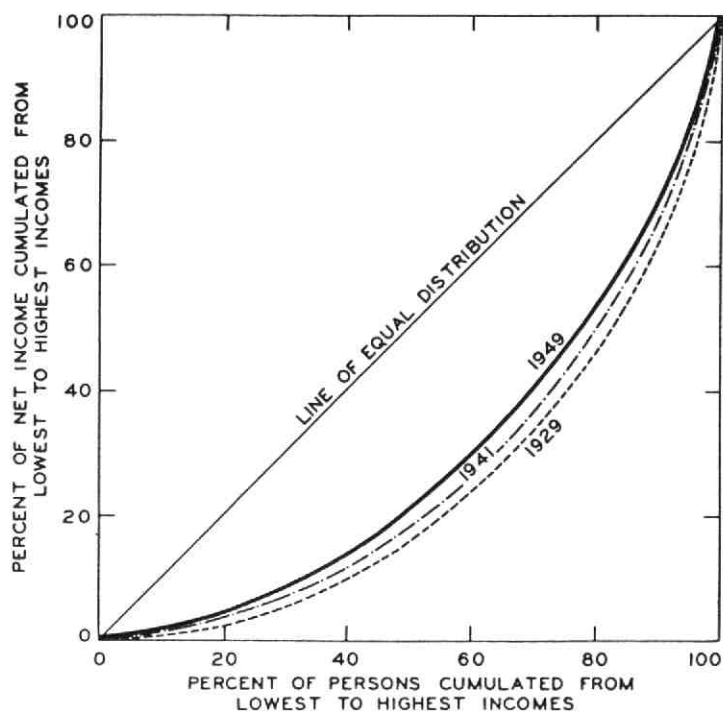
$$G = \frac{\text{opp. OAP}}{\text{opp. } \triangle OQP}$$

» 89. Benader uit de tekeningen de Gini-coëfficiënten van de Orkneys, Uruguay en Joegoslavië.

Tussen welke waarden ligt de Gini-coëfficiënt?

Lorenz-krommen zie je ook in andere situaties.

Op de volgende bladzijde zie je nogmaals de verdiensten van artsen in de V.S. in de jaren 1929, 1941 en 1949 als Lorentz-krommen afgebeeld.



» 90. Wat volgt er uit het feit dat de Lorenz-kromme in de loop der jaren steeds dichterbij de diagonaal kruipt?

» 91. Teken in je grafiek (in je schrift) ook de Lorenz-kromme van de inkomensverdeling in Nederland in 1975 als je het volgende weet:

- 30% van de bevolking verdient 15% van het totale inkomen;
- 50% van de bevolking verdient 30% van het totale inkomen;
- 80% van de bevolking verdient 60% van het totale inkomen.

Conclusies?

» 92. Maak een Lorenz-kromme van de verdeling van de bevolking van Indonesië over het totale grondgebied op grond van de volgende gegevens (de eilanden staan al in afnemende bevolkingsdichtheidvolgorde):

EILAND(EN)	OPP. IN KM ²	INWONERS IN MILJOENEN ('80)
Djawa (Java) en Madura	132.000	91
Bali en Nusa Tenggara	74.000	8
Sumatera (Sumatra)	474.000	28
Sulawesi (Celebes)	190.000	10
Maluku (Molukken)	75.000	2
Kalimantan (Borneo)	540.000	7
West Irian (Nw Guinea)	413.000	1

- » 93. Maak ook een Lorenz-kromme van de verdeling van de bevolking *van de buitengewesten* van Indonesië (d.w.z. heel Indonesië, uitgezonderd Java).
- » 94. Wat vind je van het idee van de Indonesische regering om in de jaren tachtig per jaar 1 miljoen Javanen naar Borneo en West-Irian te laten emigreren?

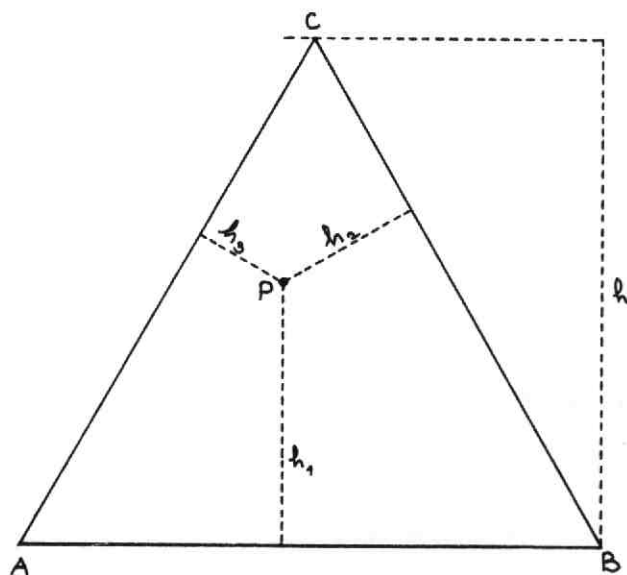


DRIEHOEKSGRAFIEKEN

Een tamelijk bijzondere manier om percentages te verwerken is met een driehoeksgrafiek. Om deze grafiek te kunnen begrijpen eerst wat meetkunde:

Als P een punt binnen een gelijkzijdige driehoek is, dan is de som van de afstanden van P tot de zijden van de driehoek gelijk aan de hoogte van de driehoek.

In een tekening:



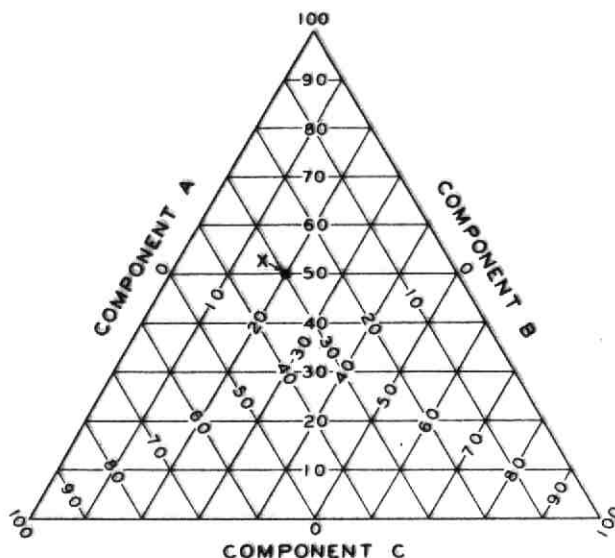
Als ABC gelijkzijdig is, dan is $h_1 + h_2 + h_3 = h$.

» 95. Toon dit aan.

Als we nu *drie* variabelen, b.v. x , y , z hebben die bij elkaar steeds dezelfde constante c opleveren (dus $x + y + z = c$) kunnen we gebruik maken van

een driehoeksgrafiek. Vaak worden x , y en z in % uitgedrukt en geldt:
 $c = 100\%$.

Een driehoeksgrafiek ziet er in principe als volgt uit:



Het betreft hier de samenstelling van een chemische stof met drie componenten A, B en C.

Stof X bestaat, volgens de grafiek, uit:

20% van stof A; 30% van stof B en 50% van stof C.

» 96. Waar zou stof Y in de grafiek komen te staan als de samenstelling van Y is:

35% van stof A; 25% van stof B en 40% van stof C?

» 97. Wat vind je duidelijker: zo'n grafiek of de omschrijving: 20%A, 30% B en 50% C?

In de V.S. is enkele jaren geleden een onderzoek gedaan onder High School studenten.

Eén van de vragen was:

aan de jongens:

Stel dat je kans hebt om een avondje uit te gaan met of een "cheerleader"*) of het knapste (studie) meisje van de klas, of met het mooiste meisje.

Welke zou je kiezen?

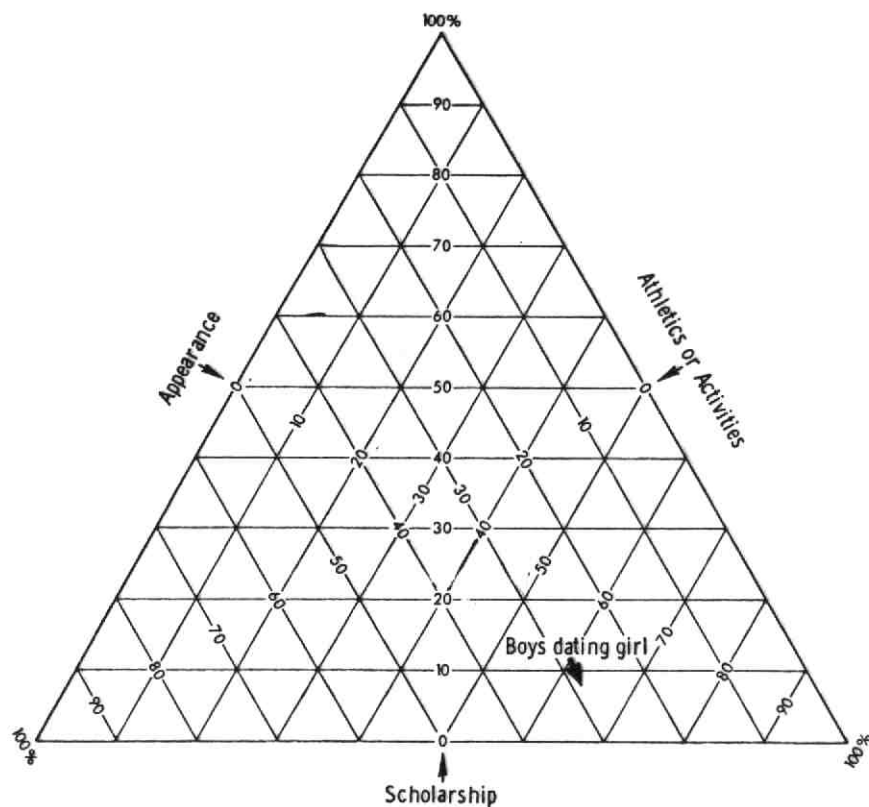
*) Cheerleader: meisje dat 'in woord en gebaar' het publiek bij sportwedstrijden oproept om de ploeg aan te moedigen.

Aan de meisjes:

Stel dat je kans hebt om een avondje uit te gaan met een hele goede sportman, de knapste jongen van de klas, of de mooiste.

Welke zou je kiezen?

Het resultaat voor de jongens werd als volgt grafisch verwerkt:



» 98. Hoe groot is het percentage jongens dat het uiterlijk schoon het belangrijkste vindt?

Hoeveel procent gaf voorkeur aan het slimste meisje?

» 99. De meisjes gaven de volgende voorkeuren te kennen:

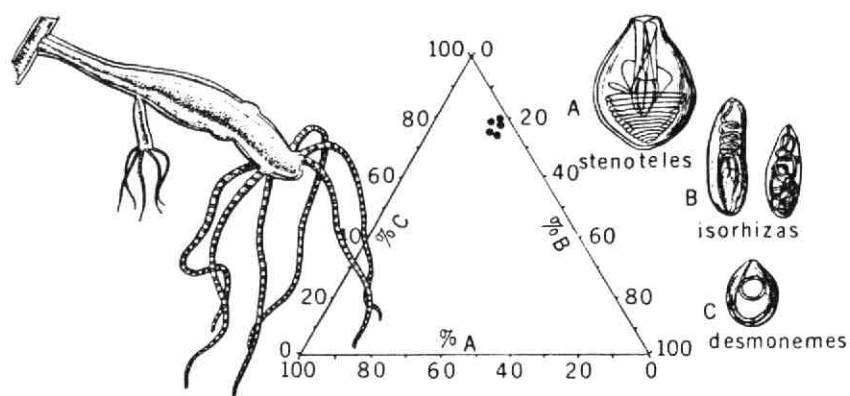
- 47% vindt uiterlijk het belangrijkste;
- 17% het knap (studie) zijn;
- 36% het sportief zijn.

Geef dit punt in de grafiek aan.

Bij driehoeksgrafieken is het soms erg moeilijk om te zien wat bedoeld wordt. Bij de voorbeelden tot nog toe was misverstand bijna uitgesloten. De volgende grafiek is echter al moeilijker te lezen.

In de tentakels van een zoet-water-poliepsoort komen drie soorten draadcellen voor:

de tamelijk grote stenoteles (A), de isorhizas (B) en de desmonemes (C). Alle drie soorten werden geteld in vijf tentakels van een vrouwelijk exemplaar. De grafiek leverde het volgende resultaat:



» 100. Je ziet vijf stippen in de driehoek. Eén voor iedere tentakel. Rond de 76% van de cellen is van type C. Is dat duidelijk in deze grafiek? Hoeveel % van de cellen is van type A en hoeveel % van type B?

Ook niet echt duidelijk is de volgende driehoeksgrafiek. Het betreft hier een grafiek uit een artikel over archeologische opgravingen in Engeland. En wel in het bijzonder een onderzoek naar de verspreiding van botten en skeletten van dieren ter grootte van een koe of paard.

Er werden drie 'componenten' onderscheiden:

grote botten, stukken skelet (SKULL) (P)

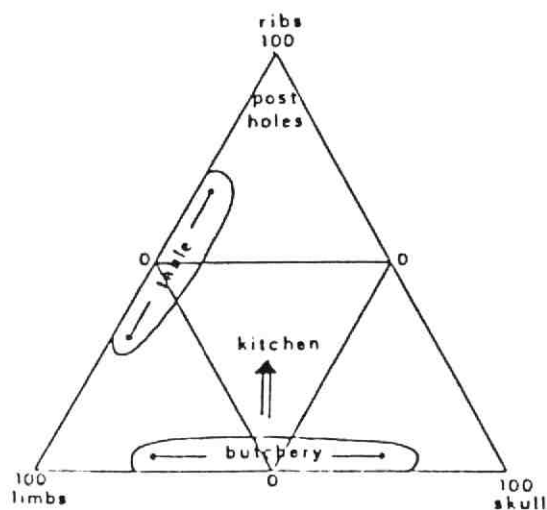
botten van poten (LIMBS) (Q)

kleinere botten o.a. robben (RIBS) (R)

Je zou verwachten - volgens de onderzoekers (Halstead e.a.; Cambridge)- dat P en Q botten het meest te vinden zijn in het slachtgedeelte van het

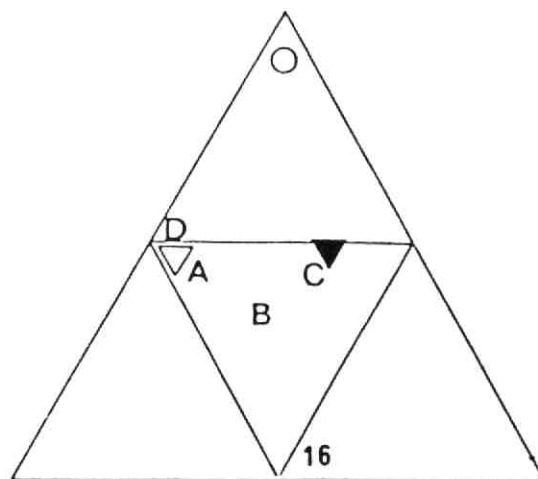
huis of nederzetting, dat in de keuken alle drie soorten te vinden zijn, dat rond de eettafel Q en R te vinden zijn, en de afgegeten botjes (R) verdwijnen in de afvalbakken (postholes).

Grafisch:



De te verwachten verdeling van botten over verschillende plaatsen.

Na vele afgravingen kwam de volgende grafiek als resultaat tot stand:



De werkelijke verdeling van botten op de vindplaatsen A, B, C en D en de vier andere plaatsen:

▼, ▼, ○ en 16.

- » 101. Wat vind je van deze grafiek? Wat ontbreekt er zoal aan? Wat kan beter?
- » 102. Welke van de acht vindplaatsen zal waarschijnlijk een afvalplaats zijn? En welke een slachtplaats? En welke de eetkamer? Keuken?

10

DRIE DIMENSIES EN CORRELATIE

In dit hoofdstukje kijken we naar enkele drie-dimensionale grafieken.

Bij een onderzoek (Navratil 1962) werd gekeken naar het verband tussen *lengte* en *hoofdomtrek* van pas geboren babies.

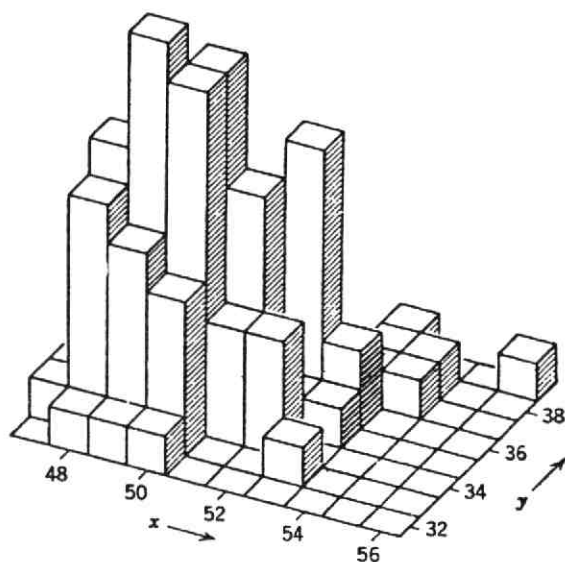
Na 100 babies had Dr. Navratil het volgende tabelletje bij elkaar geturfd:

		hoofdomtrek in cm									
		↓									
lengte in cm	→	47	48	49	50	51	52	53	54	55	56
39											
38											
37											
36							### 				
35				###	### 	### 					
34			### 	### ###	### 						
33			### 	###							
32											A

Je ziet duidelijk dat je een 'normaal' hoofd aantreft op de diagonaal van links onder naar rechts boven.

» 103. Aan welke kant van die diagonaal zitten de mensen met een relatief groot hoofd (waterhoofd)?

Van bovenstaande tabel kun je nu een drie-dimensionaal histogram maken. Je vat dan het aantal waarnemingen in één hokje op als de 'hoogte' van het staafje.



» 104. Zet de letter A uit de tabel op de juiste plaats in het histogram-
Wat voor nadeel heeft dit histogram in vergelijking met de tabel?

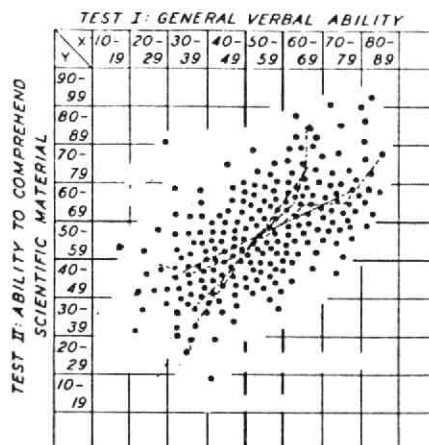
Bij een ander onderzoek keek men naar het verband tussen:

- spreekvaardigheid;
- vaardigheid in het begrijpen van wetenschappelijk materiaal.

Daartoe werden 231 eerstejaars studenten getest van het College of Engineering van de University of Washington.

De studenten kregen twee 'proefwerken' te maken. Voor iedere test waren maximaal 100 punten te halen. Als een student in de spreekvaardigheidstest nu 30 punten haalde en 80 in de wetenschapstest, werd dat resultaat als een punt (30,80) in een rooster gezet. Dat werd met de resultaten van alle 231 studenten gedaan.

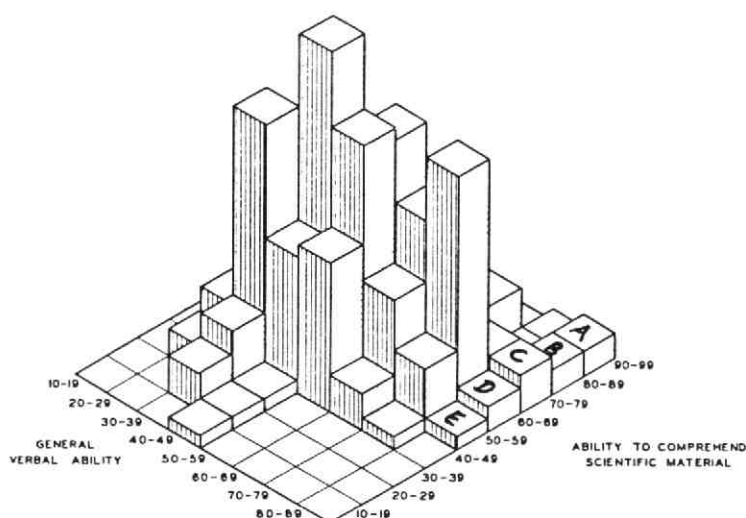
Het volgende resultaat:



De punten vormen een 'wolk' van linksonder naar rechtsboven.

Aan de hand van dit puntendiagram kun je een drie-dimensionale grafiek tekenen door het aantal punten te tellen in ieder roostervierkantje en dat aantal als hoogte van een 'wolkenkrabber' te nemen.

Dit levert de volgende grafiek:



- » 105. Laat zien dat de blokjes A, B, C, D en E niet precies kloppen met het puntendiagram.
- » 106. Is er enig verband (correlatie) tussen de twee gemeten grootheden: spreekvaardigheid en opnemen van wetenschappelijk materiaal?
- » 107. In welk roosterhokje bevinden zich de meeste punten?
- » 108. Kun je aan de hand van dit experiment ook uitspraken doen over Nederlandse studenten?
- » 109. Maak zelf een puntendiagram met daarin resultaten van twee proefwerken. Langs de X-as die van een wiskundeproefwerk, langs de Y-as die van een natuurkundeproefwerk.

leerling	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
wiskunde	9	4	6	9	8	3	5	3	8	5	6	6	5	8	9	6	3	7	7
natuurkunde	8	5	4	9	7	5	5	3	5	7	5	8	6	8	6	6	4	6	7

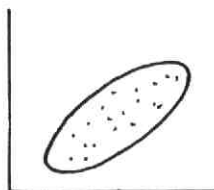
- » 110. Verbaast het verloop van de grafiek je?
- » 111. Doe dezelfde opgave eens voor de laatste resultaten van het proefwerk Wiskunde en Engels in je eigen klas. Kijk of er net zo'n sterk verband (correlatie) tussen Engels en Wiskunde bestaat als tussen Wiskunde en Natuurkunde.

In de vorige opgaven hebben we gekeken naar het verband tussen twee variabelen.

Dat waren resp.:

- | | |
|----------------------------------|---------------------------|
| - lengte in cm (van een persoon) | - hoofdomtrek in cm; |
| - spreekvaardigheidstest | - wetenschapstest; |
| - resultaten Wiskunde | - resultaten Natuurkunde; |
| - resultaten Wiskunde | - resultaten Engels. |

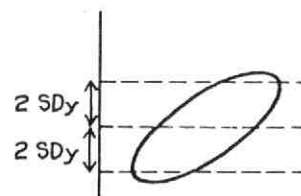
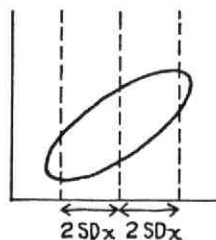
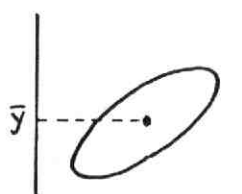
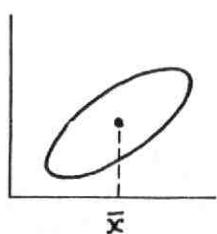
De grafieken zagen eruit als een wolk van punten. Die wolken hadden steeds een wat andere vorm. Die bij Wiskunde-Natuurkunde zag er ongeveer zó uit:



Er is een duidelijke correlatie tussen de Wiskunde- en Natuurkundecijfers. Hoe kun je dat nu in een *getal* uitdrukken?

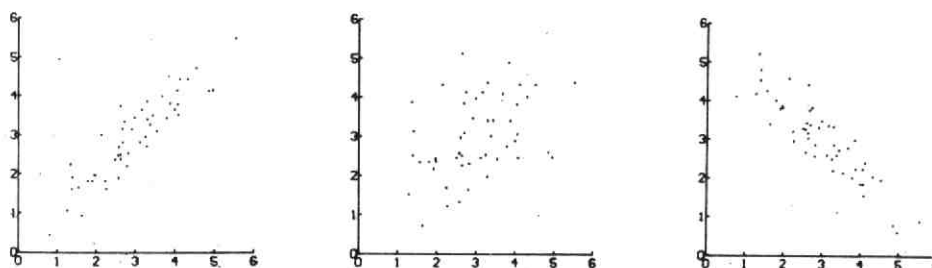
Vier mogelijkheden hebben we al:

1. het gemiddelde van X (Wiskunde);
2. het gemiddelde van Y (Natuurkunde);
3. de SD van de X;
4. de SD van de Y.



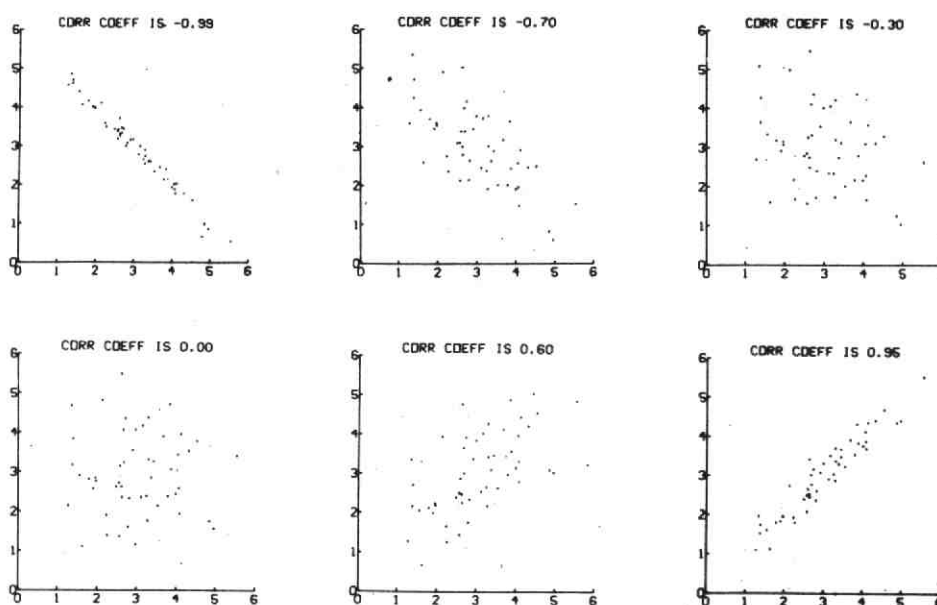
Jammergenoeg zijn we er nog niet met deze vier kenmerken.

De volgende drie grafieken hebben dezelfde gemiddelden (3) en SD's (1) en zijn duidelijk heel anders:



De afspraak is nu dat hoe meer de punten op één lijn lijken te liggen, des te hoger is de *correlatie-coëfficiënt*.

Kijk maar eens goed naar de volgende voorbeelden:



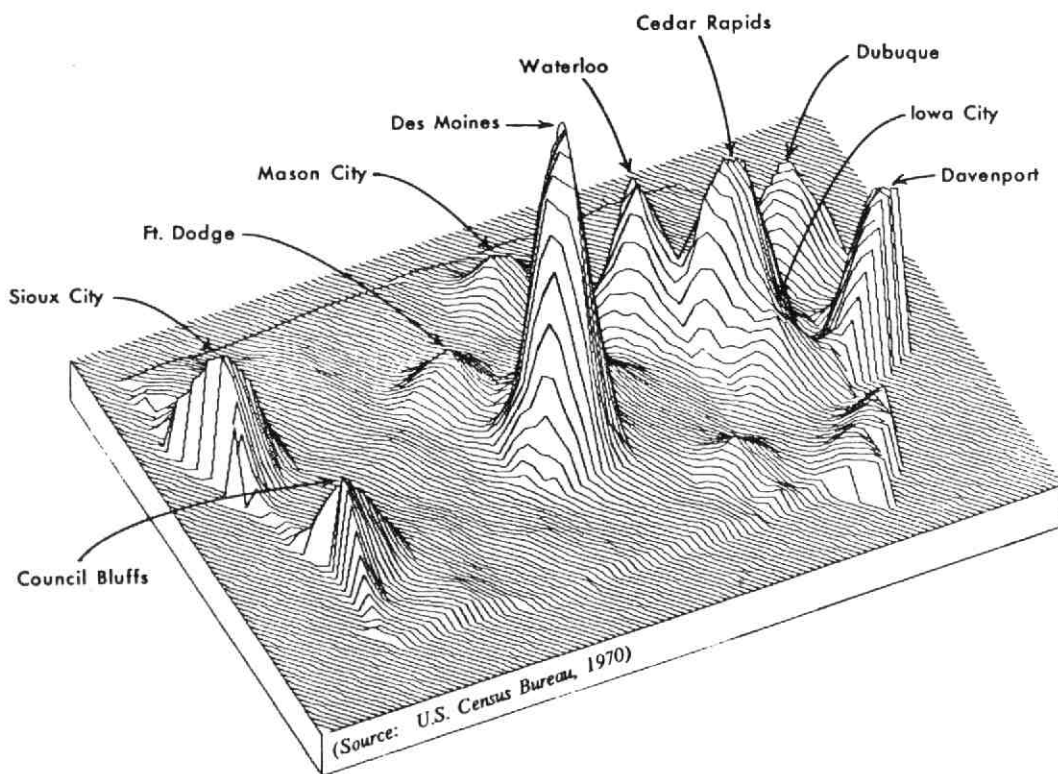
De correlatiecoëfficiënt ligt altijd tussen -1 en 1 .

» 112. Schat de correlatiecoëfficiënten (in één decimaal achter de komma) van de:

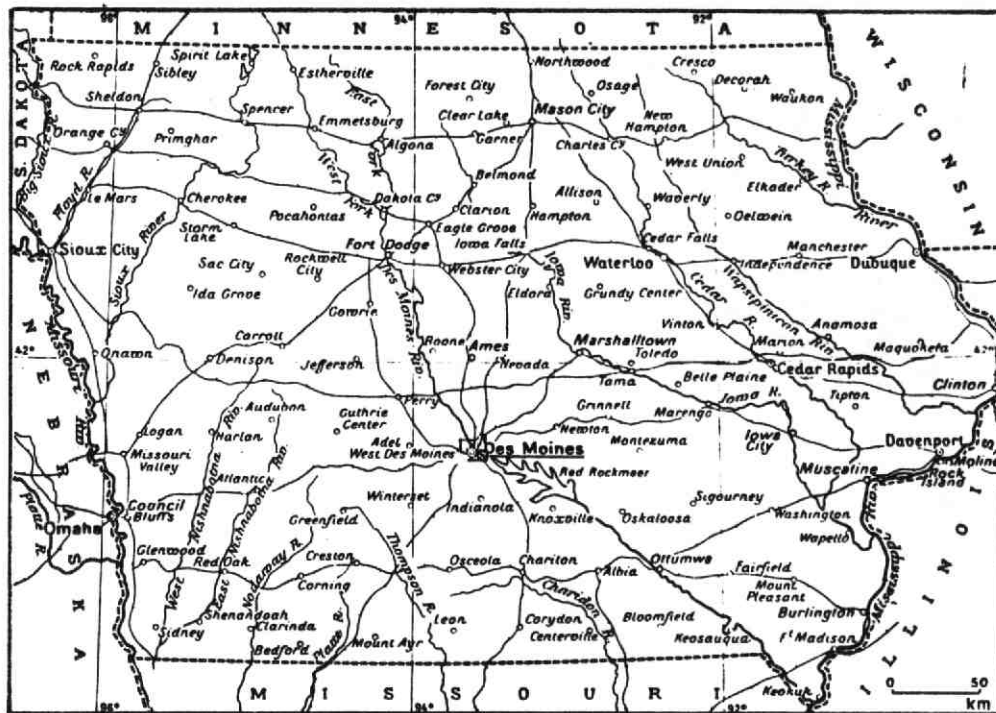
- Spreekvaardigheidstest / Wetenschapstest.
- Wiskunde / Natuurkunde.
- Wiskunde / Engels.
- En van de drie grafieken boven aan deze bladzijde.

Computers kunnen heel gemakkelijk drie-dimensionale histogrammen tekenen. Of, wat vaak nog mooier is: drie-dimensionale grafieken.

Zoals deze:



Een door een computer getekende grafiek met daarin de grootte van de bevolking van de Amerikaanse staat IOWA.



» 113. Op de kaart vind je zo'n 150 km ten zuidoosten van Des Moines het stadje Ottumwa. Geef dit plaatsje aan in de grafiek.

» 114. Teken de grenzen van de staat IOWA in de grafiek in.



11

LEVENSVRWACHTING EN BEVOLKINGSGROEI IN DE DERDE WERELD

Dit hoofdstukje bevat een (ingekort en vertaald) artikel uit het populair wetenschappelijk blad

Scientific American (mei '82). Het laat zien hoe een artikel door grafieken verduidelijkt kan worden, mits op de juiste wijze gebruikt. Opvallend zijn daarbij de eerste grafieken die heel erg globaal en onnauwkeurig zijn en toch de essentie van het verhaal weergeven op een zeer beeldende manier.

LEVENSV ERWACHTING EN BEVOLKINGSGROEI IN DE DERDE WERELD

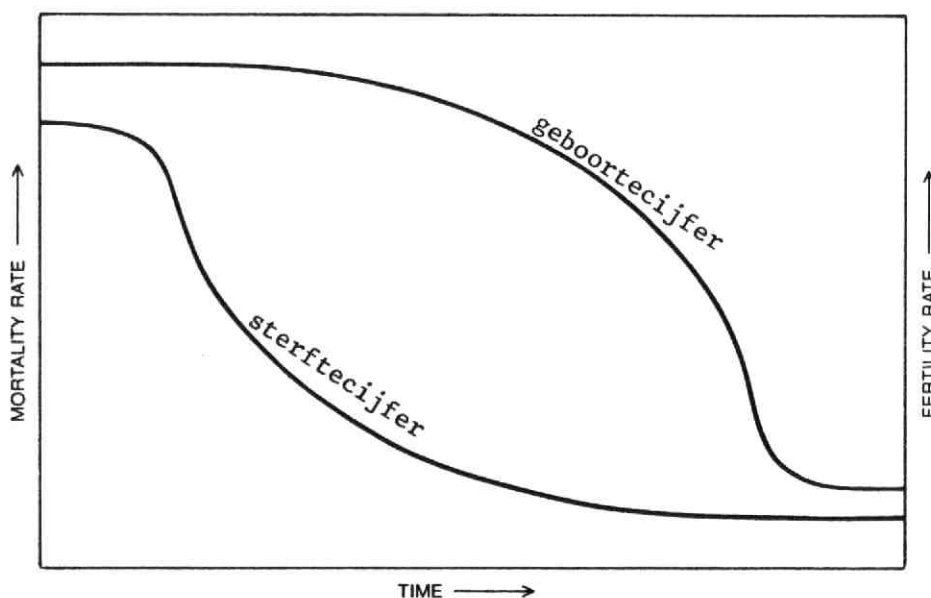
(naar een artikel van Gwatkin en Brandel; *Scientific American*, mei 1982).

Op dit moment wordt verwacht dat de bevolking van de derde wereld zal verdrievoudigen in de volgende eeuw en daarna stabiel zal blijven.

De snelle groei van de bevolking ontstaat doordat er minder mensen sterven en de vruchtbaarheid nog erg groot is. Deze fase wordt wel de 'demografische overgang' genoemd. Eerst zijn het sterfte- en geboortecijfer groot, na de 'overgang' zijn sterfte- en geboortecijfer klein.

De ellende is alleen dat éérst de sterftecijfers afnemen, terwijl de vruchtbaarheid hoog blijft. Pas later, soms tientallen jaren later, begint het geboortecijfer af te nemen.

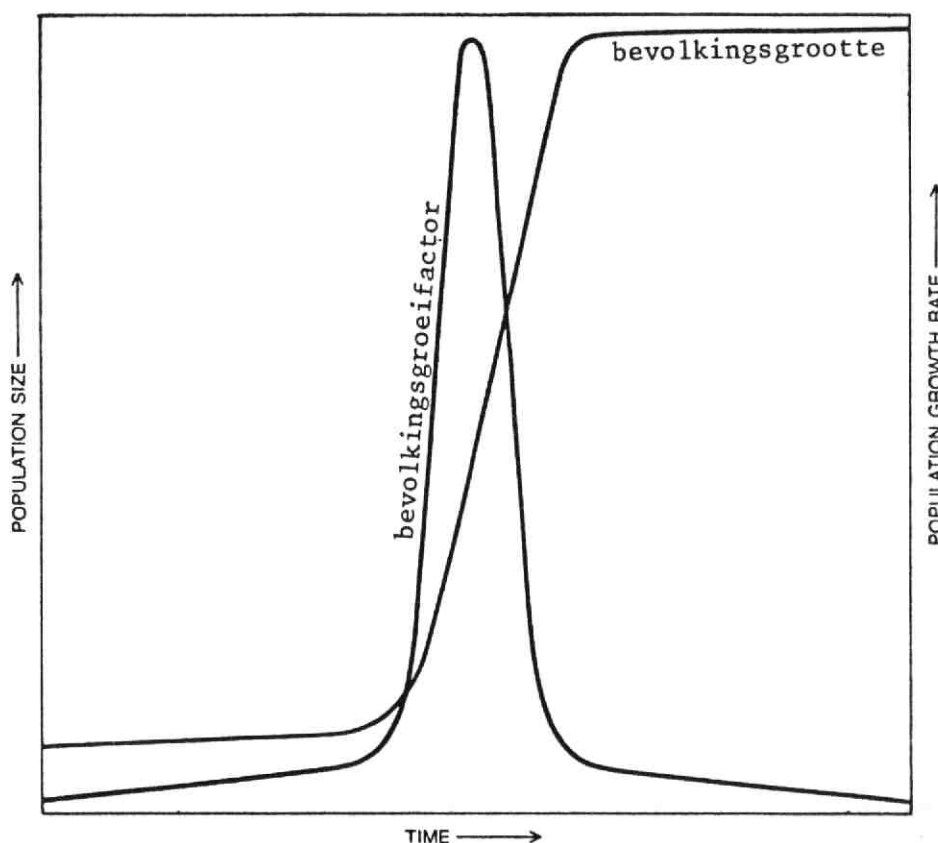
Schematisch:



- » 115. Een 'gekke' grafiek. Géén schaalverdeling langs de assen. Wat is klaarblijkelijk het enige doel van zo'n grafiek?
- » 116. Waarom is het aannemelijk dat de geboortegrafiek bóven de sterftegrafiek loopt?

Alle moderne landen hebben deze 'overgang' al achter de rug. In de meeste ontwikkelingslanden zit men in het midden van de grafiek: het sterftecijfer is snel achteruit gelopen, het geboortecijfer nog niet.

Het effect van het doorlopen van de overgangsfase wordt duidelijk aan de hand van de volgende grafiek:

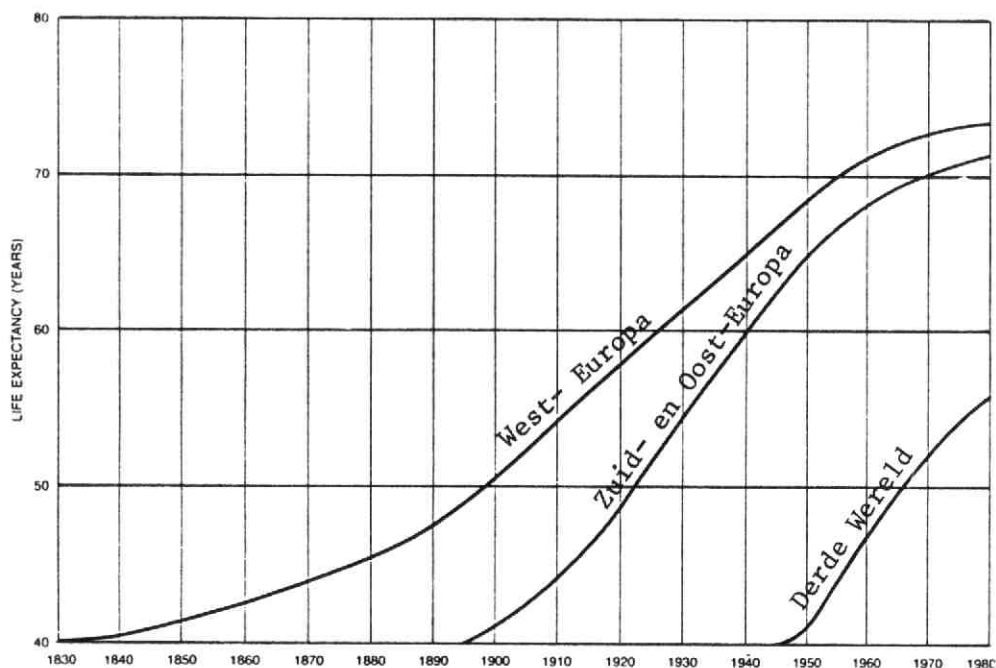


» 117. Verklaar het verloop van de bevolkingsgroefactor met behulp van de vorige grafiek.

» 118. Verklaar dat de bevolkingsgroottegrafiek daar direct uit volgt.

Een begrip dat algemeen gebruikt wordt bij het bestuderen van bevolkingsgroei, is het gemiddeld aantal te verwachten levensjaren. In 1830 was dat maar 40 jaar voor West-Europeanen. Nu is dat ruim 70 jaar. Uit de geschiedenis blijkt dat de toename van de levensverwachting van 40 naar 50 jaar samenvalt met dit begin van de demografische overgangsfase.

In Zuid- en Oost-Europa begon deze overgang veel later en in de derde wereldlanden is deze fase pas kort geleden begonnen, zoals blijkt uit de volgende grafiek:



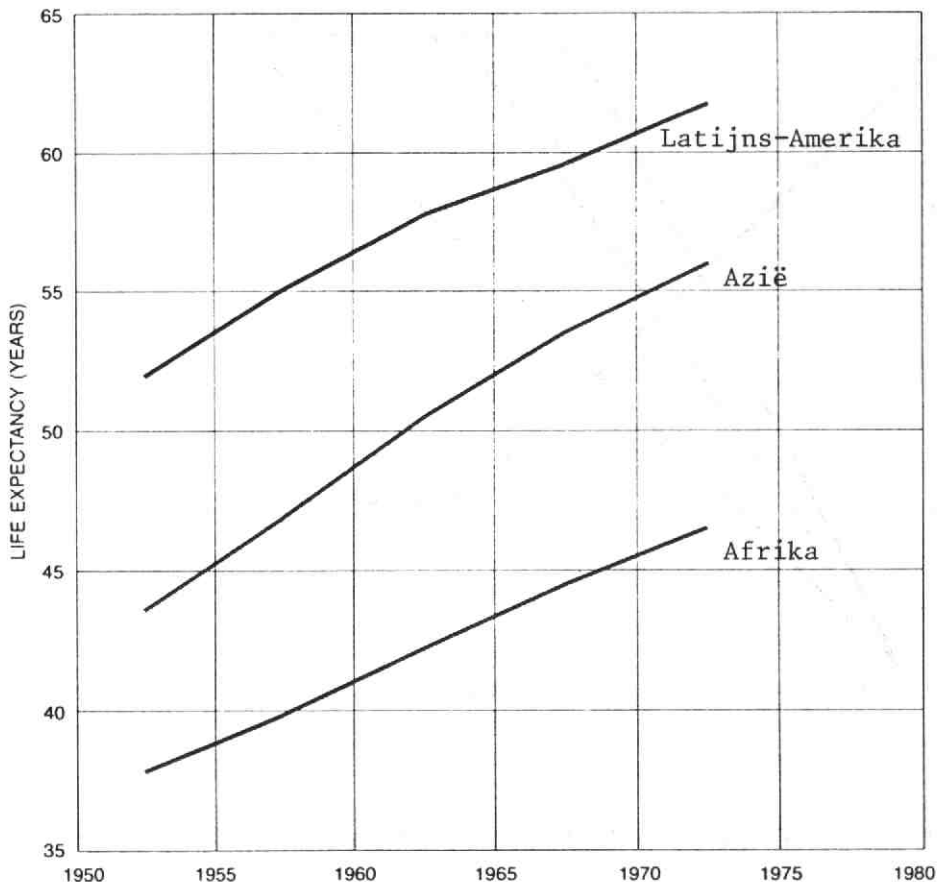
» 119. Hoe lang duurde het in West-Europa voor de levensverwachting was toegenomen van 40 naar 50 jaar?

» 120. En in Zuidoost-Europa? En in de Derde Wereld?

Het toenemen van de levensverwachting in derde wereldlanden ging het snelst van de drie onderzochte gebieden.

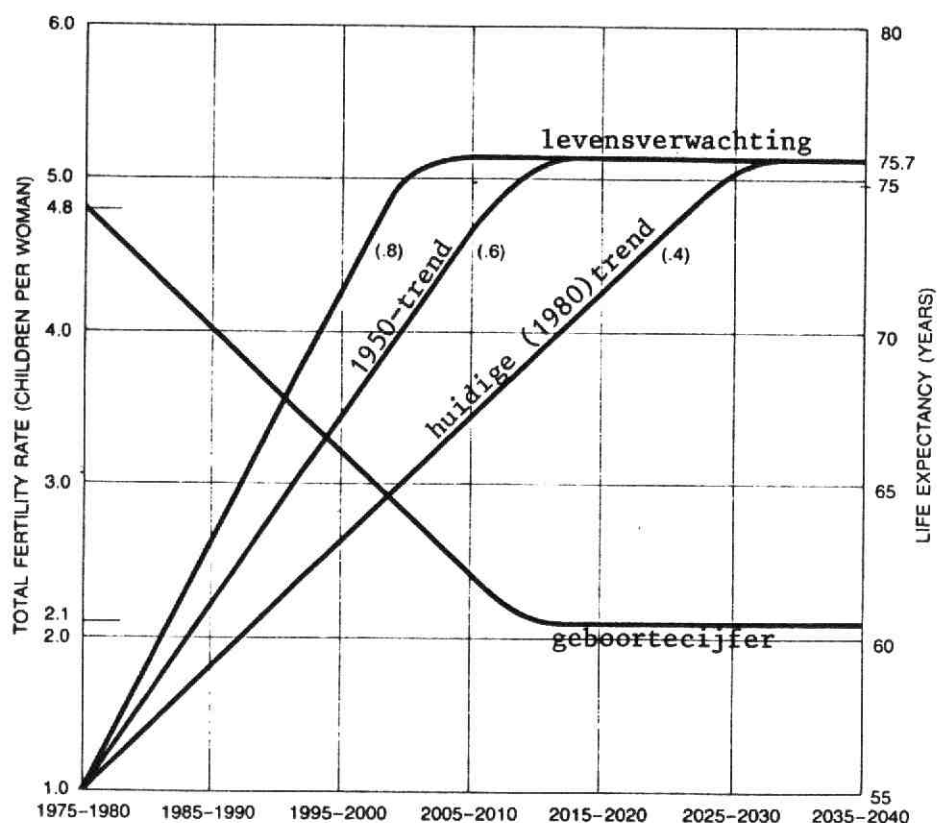
» 121. Hoe zit het met de toename van 50 naar 60 jaar? Gaat dat ook sneller in de Derde Wereld?

Overigens zijn er grote verschillen tussen de verschillende derde wereld-landen, zoals blijkt uit deze grafiek:



- » 122. In welk van de drie gebieden is de levensverwachting het grootst?
- » 123. In welk gebied neemt de levensverwachting het snelst toe?
- » 124. Vergelijk de toename met de toename uit de vorige grafiek voor alle derde wereldlanden. Deze laatste loopt veel steiler. Hoe kan dat?

De auteurs van het artikel voorspellen de toekomstige bevolking van de derde wereldlanden, en wel op drie manieren: één uitgaande van de huidige trend, één uitgaande van de wat sneller stijgende trend van de jaren 50 en één die precies tweemaal zo groot is als de huidige. Verder zijn ze ervan uitgegaan dat het geboortecijfer (aantal kinderen per vrouw) lineair zal afnemen van de huidige 4,8 tot 2,1. Bij 2,1 blijft de bevolking constant.

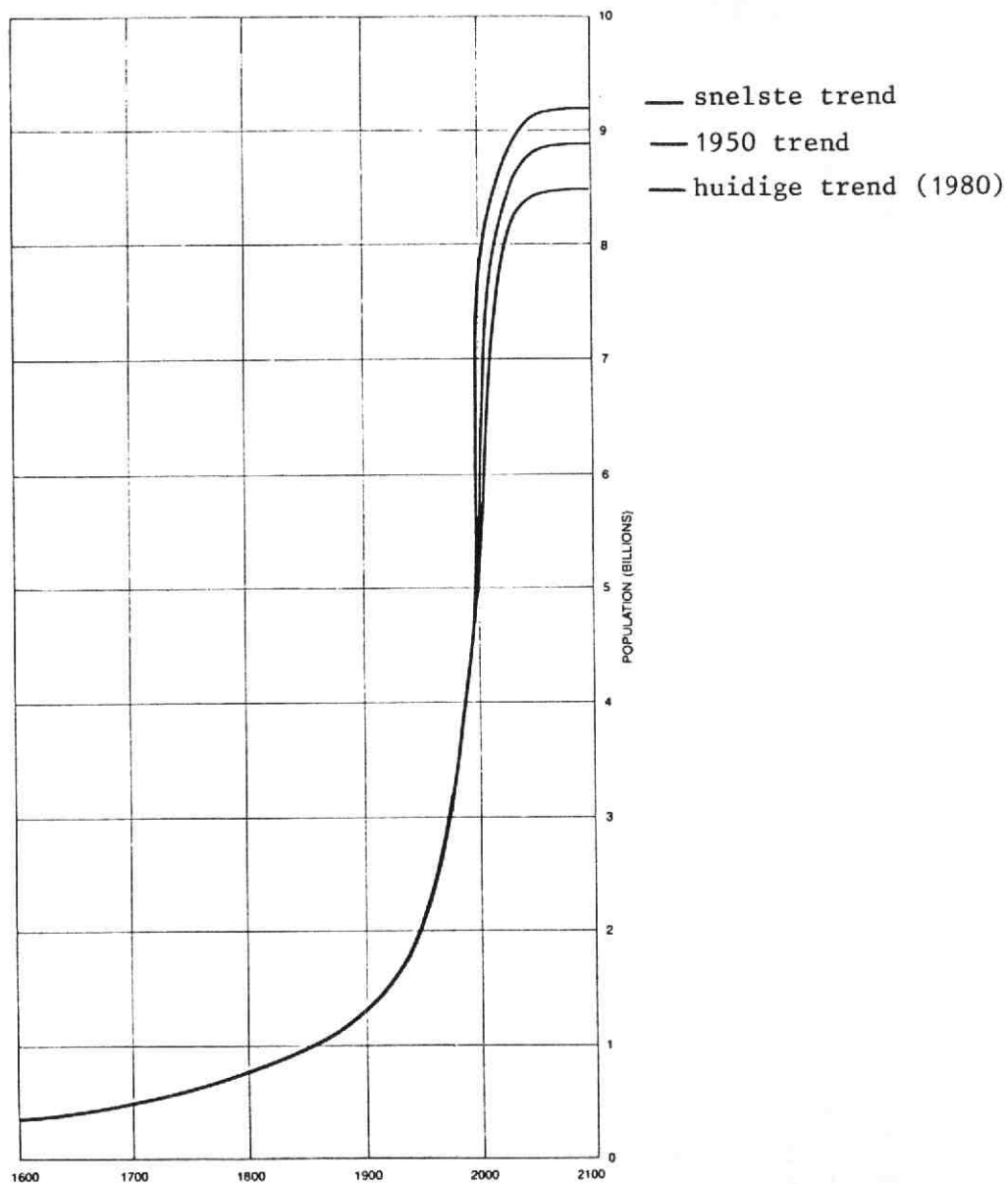


- » 125. Klopt die lineaire afname van de vruchtbaarheid met de grafiek van blz. 76?
- » 126. Waarom zal bij 2,1 de bevolking constant blijven en niet bij 2,0 kinderen per vrouw?

De meest rechtse van de drie grafieken is degenen die ervan uitgaat dat de huidige (1980) trend doorgaat.

- » 127. In welk jaar zal de levensverwachting van 75,7 jaar bereikt worden?
- » 128. De middelste gaat uit van de snellere vooruitgang uit de jaren 50. Vind je het aannemelijk dat dat gebeurt? (Kijk naar de grafiek op blz. 78).

Tenslotte de laatste grafiek die een voorspelling doet over de grootte van de bevolking van de derde wereldlanden, uitgaande van de drie eerder vermelde mogelijkheden:



- » 129. Je ziet dat de voorspelling (ondanks de tamelijk grote trendverschillen) nogal dicht bij elkaar liggen. Tussen welke grenzen?
- » 130. Klopt de verdrievoudiging van de derde wereldbevolking uit het begin van het artikel?
- » 131. Welk gedeelte van de grafiek is niet goed getekend? Hoe had dit probleem voorkomen kunnen worden?

Indonesië bezig met grote 'volksverhuizing'



Minister Zain

Door onze correspondent
WILLEM VAN KEMENADE

JAKARTA, juni — Veel grote landen kampen met ongelijke spreiding van de bevolking over hun grondgebied, maar nergens is het probleem zo extreem als in Indonesië. Op het eiland Java, dat met 192.000 vierkante kilometer nog geen zeven procent van de Indonesische landoppervlakte van bijna twee miljoen vierkante kilometer uitmaakt, wonen 92 miljoen mensen, d.w.z. 62 procent van de totale bevolking van 147 miljoen.

Vandaar dat „transmigratie“, de overplanting van grote aantallen Javanen naar andere delen van de archipel, een centraal, zoniet dominerend thema in de nationale ontwikkeling is geworden.

Transmigratie (Indonesisch: transmigrasi) is een uit het Nederlands afkomstig woord. De Nederlanders zijn er in 1905 mee begonnen. Van 1905 tot 1949 zijn er 30.569 gezinnen, samen 227.884 personen, van Java naar Sumatra overgebracht, zegt de Indonesische minister van arbeid en transmigratie, Harun Al-Rasjid Zain. Van 1950 tot 1979 steeg het aantal naar 204.425 gezinnen met gemiddeld vijf personen. Het totale aantal transmigranten uit Java heeft in 75 jaar dus slechts 1,2 miljoen mensen bedragen.

Vraag: *Is de strijd tegen de overbevolking van Java dan niet tot mislukken gedoemd of al verloren?*

Zain: „Sinds het derde vijfjarenplan — 1974-'84 — hebben we niet alleen aan schaalvergroting gedaan maar ook een integrale aanpak toege-

past. We brengen de mensen niet meer alleen over naar een ander eiland, maar koppelen die overbrenging systematisch aan de regionale ontwikkeling van het betrokken buitengewest. Ons streefgetal voor de transmigratie in de periode van het derde vijfjarenplan is een half miljoen gezinnen. In het vierde plan willen we omhoog naar driekwart of één miljoen. Veel hangt af van de ervaring die we nu in het veld opdoen.”

Fouten

Minister Zain zegt dat er aanvankelijk nogal wat fouten zijn gemaakt bij het voorbereidende werk. „Uit tijds-overwegingen schakelden wij, bij het opruimen van oerwoud het allerzwaarste materiaal in, totdat bodemkundig onderzoek uitwees dat het uitrukken van bomen de toplaag van de grond zodanig beschadigde dat bebouwing aanzienlijk geremd werd. Nu worden de bomen, die vaak 80 meter hoog zijn en twee meter diameter hebben, met de cirkelzaag vlak boven de wortels afgezaagd en moet het terrein verder met handkracht gereedgemaakt worden. Vaak gaat daar veel tijd in zitten. Het tempo van de vestigingen wordt geheel bepaald door de hoeveelheid land die gereedkomt en de infrastructuur. Tien, twintig jaar geleden hebben we mensen bij wijze van spreken diep in de binnenlanden van Borneo geplaatst en ze produceerde een goede oogst, maar die was niet te verkopen. Nu is de be-

nadering integraal. De toegang moet eerst verzekerd zijn.”

Het is nog niet duidelijk welke herverdeling van de bevolking Indonesië uiteindelijk wenst te bereiken. Minister Zain heeft een grote kaart naast zijn bureau hangen waarop voor het jaar 2001 een bevolkingscijfers van 237 miljoen staat aangegeven, waarvan 139 miljoen op Java. Demografen hebben de optimale bevolkingsgrootte voor Java echter op 70 miljoen gesteld.

Betekent het dan dat er tientallen miljoenen Javanen zullen worden verplaatst? „Ja, dat is zo“, zegt minister Zain. „De bevolking op Java groeit nog steeds met twee procent per jaar (...). We kunnen niet direct in de toekomst springen en de ideale bevolkingsgrootte bereiken, maar we kunnen wel het probleem verzachten.”

Hij vervolgt dat de ontlasting van Java slechts de derde doelstelling van het transmigratiebeleid is. De eerste twee zijn verhoging van de voedselproductie en ontwikkeling van de buitengewesten.

Exodus

Behalve de systematische exodus uit Java is er nog steeds een stroom naar Java toe, vooral naar Jakarta. „Maar de balans slaat nu door in het voordeel van Java“ zegt de minister. „Verbetering van de onderwijsfaciliteiten op de andere eilanden heeft de trek naar Java afgeremd.”

Vraag: Het Indonesische deel van Borneo is zo groot als Frankrijk maar er wonen nog geen zeven miljoen mensen. Hoeveel mensen wil de regering uiteindelijk op Borneo hebben?

Zain: „Streefcijfers zijn niet te geven omdat we pas drie jaar met de grootschalige aanpak bezig zijn. Duitse experts doen al vijf jaar onderzoek in Oost-Borneo om de beste landbouwgebieden uit gigantische stukken oerwoud te selecteren. De Fransen maken bodemkaarten van Zuid-Borneo. (Nederlanders doen dit op Zuid-Celebes en Oost-Sumatra.) De topografie van Borneo is zo totaal anders dan die van Java. Java is het vruchtbaarst van alle eilanden omdat de bodem vulkanisch is. Je hebt hier op Java twee, drie oogsten per jaar.”

„Op Borneo moet je al meteen na de eerste oogst kunstmest gebruiken. De ontwikkeling per hectare is op Borneo het duurst. We moeten zoveel mogelijk fouten zien te vermijden en de allerbeste middelen toepassen, geen grond meer verwoesten etc. Vervolgens moeten we het huidige „persproces”, d.w.z. mensen aansporen om Java te verlaten, vervangen door een „zuigproces”, dus zoveel industrie- en andere projecten opzetten zodat de mensen spontaan naar Borneo transmigreren.”

Irian Jaya

Vraag: Onlangs was vice-president Adam Malik in Irian Jaya (voormalig Nieuw

Guinea) en zei in Merauke dat er negen miljoen Javanen geplaatst moeten worden. Op welke termijn dan wel?

Minister Zain zegt dat de aantallen op wetenschappelijke wijze vastgesteld moeten worden. Er is al een begin gemaakt met transmigratie naar Irian Jaya maar op veel kleinere schaal dan bijvoorbeeld naar Sumatra omdat de meest elementaire infrastructuur op Irian ontbreekt. Alles moet daar per vliegtuig.

Op de vraag of het de bedoeling is om Irian volledig te „Javaniseren” antwoordt Zain lachend: „Dat kan zo gezegd worden maar het is natuurlijk niet helemaal de bedoeling, want het gaat niet alleen om mensen van Java maar ook van Bali en Nusa Tenggara Bharat (de verzamelnaam voor Soembawa en Lombok) die ook plaatselijk overbevolkt zijn. Van de 27 provincies zijn er acht „leveranciers” (de vijf provincies op Java, Bali en NTB) en de rest is afnemers”.

In Nederland heeft men zich in de jaren vijftig en zestig nogal druk gemaakt over het zelfbeschikkingsrecht en de eigen identiteit van de Papoea's. We vragen de minister of daar nog iets van overblijft als er uiteindelijk miljoenen Javanen neerstrijken.

„U moet het probleem niet zwart-wit zien. We wijzen gebieden aan die niet al te dicht bij de autochtone bevolkingsconcentraties liggen. De nederzettingen van transmigranten liggen doorgaans tien tot vijftig kilometer van de oorspronkelijke dessa's.”

Volgens minister Zain is er nog nauwelijks sprake van echte cultuurbotsingen tussen transmigranten en autochtonen geweest. „Er zijn natuurlijk verschillen en problemen, maar we hebben allerlei specialisten van universiteiten ingeschakeld om daar inzicht in te krijgen”.

Schokbrekers

Ook zijn er speciale regels in de transmigratie-wetgeving die als schokbrekers moeten dienen. Zain zegt dat tien procent van het budget voor elk project moet worden aangewend voor verbetering van de infrastructuur van de omringende autochtone bevolking of tien procent van de naburige autochtonen in het project moeten worden opgenomen met dezelfde rechten als de transmigranten. „In Irian zijn we geneigd om tot 20 procent te gaan”.

Een van de moeilijkste transmigranten-groepen zijn Balinezen. Religie (Hindoeïsme), cultuur en gewoonten zijn bij de Balinezen één en daarom passen zij hun cultuur moeilijker aan dan anderen. Zij worden dan ook alleen in grote homogene groepen overgebracht, die dan op Sumatra en Celebes aparte kleine Hindoe maatschappijtjes vormen.

Glunderend van plezier zegt de minister dat zij de Balinezen het liefst neerzetten in gebieden waar veel wilde zwijnen zitten. „Die Javanen is Moslin en eet geen varkensvlees, maar de Balinezen vreten die zwijnen met huid en haar op”.

Bevolkings spreiding in Indonesië

Gebied	Oppervlakte	Percentage van totale landoppervlakte	Bevolking in 1980 (miljoenen)	Percentage van totale bevolking
Java/Madura	132.187 km ²	6.95	91.281	61.93
Sumatra	473.606	24.86	27.981	18.99
Kalimantan (Borneo)	539.460	28.32	6.721	4.56
Sulawesi (Celebes)	189.216	9.93	10.377	7.04
Bali	5.561	0.30	2.470	1.68
Irian Jaya (Nieuw Guinea)	421.981	22.16	1.146	0.78
Overige eilanden	142.558	7.48	7.406	5.02
TOTAAL	1.904.569	100.	147.384	100.

Bron: Centraal Bureau voor de Statistiek, Indonesië

12

SAMENVATTING

Dit boekje bevat nauwelijks een duidelijke wiskundelijn. Ook staan er weinig formules in. Toch zijn het geen onbelangrijke zaken die aan de orde komen. De titel 'Grafische Verwerking' zegt eigenlijk al waar het om gaat: bij onderzoek komen reeksen cijfers te voorschijn. Vaak in zulke hoeveelheden dat het moeilijk is overzicht te krijgen. Door middel van grafieken - en we vatten dat woord dan ruim op - is het vaak mogelijk orde in de chaos te scheppen. Het gevaar van grafieken is echter dat je ze op je 'eigen' manier kunt tekenen en daarmee iets suggereren dat misschien niet helemaal 'eerlijk' is. Bovendien moet je zelf ook enigszins in staat zijn grafieken te lezen en te beoordelen. Over al dit soort zaken gaat dit boekje.

Hoofdstuk 1

In dit hoofdstuk vergelijken we twee - ogenschijnlijk tegengestelde - uitspraken over de Amerikaanse defensieuitgaven. Voor een groot deel is deze tegenspraak te herleiden tot het verschil tussen *relatief* en *absoluut*. Maar er blijken ook nog heel andere zaken mee te kunnen spelen. Zo wordt er naar defensieuitgaven per hoofd van de bevolking gekeken, terwijl het in het geheel niet duidelijk is of dat wel zo logisch is. Ook speelt de inflatie een verwarrende rol.

Enkele begrippen: relatief, absoluut, percentages, inflatie.

Hoofdstuk 2

De walvis wordt met uitsterven bedreigd. Weer een uitspraak die met grafieken wordt onderbouwd. Overigens blijkt niet *de* walvis met uitsterven bedreigd te worden, maar vele soorten. Andere soorten weer niet. Bovendien worden de 1972-getallen vergeleken met de 'oorspronkelijke' getallen, terwijl niemand echt duidelijk maakt hoe je daaraan komt. Enkele grafische mogelijkheden om het uitsterven in beeld te brengen worden bestudeerd, alle met voor- en nadelen.

Enkele begrippen: staafdiagram, frequentie, relatieve frequentie, cirkeldiagram.

Hoofdstuk 3

Bevolkingsgegevens van de Orkney-eilanden worden bestudeerd. Daarbij wordt gebruik gemaakt van een bijzonder staafdiagram, het *histogram* en van een *polygoon*. Verder kijken we naar *bevolkingspiramides* en naar de rol van *klassebreedtes* daarbij. Een bevolkingspiramide met *frequentiedichtheden* geeft vaak een eerlijker beeld dan een gewone bevolkingspiramide met ongelijke klassebreedtes.

Enkele begrippen: histogram, polygoon, bevolkingspiramides, klassen, klassebreedtes, frequentiedichtheid.

Hoofdstuk 4

Dit is een verzameling van verschillende 'grafische voorstellingen'. Deze worden bekeken naar hun eerlijkheid en duidelijkheid.

Geen nieuwe begrippen.

Hoofdstuk 5

In hoofdstuk 5 worden verschillende centrummaten ingevoerd. Naast het (rekenkundig) *gemiddelde* zijn dat de *mediaan* en *modus*.

De centrummaten kunnen worden gebruikt om een verzameling waarnemingen met één getal te karakteriseren.

Enkele begrippen: centrummaat, gemiddelde, mediaan, modus.

Hoofdstuk 6

Aangezien het geven van de centrummaat vaak onvoldoende informatie geeft, kan ook nog een getal gegeven worden dat als maat voor de spreiding dient. Er worden in dit hoofdstukje twee spreidingsmaten behandeld.

Enkele begrippen: spreidingsmaat, gemiddelde absolute afwijking, standaarddeviatie.

Hoofdstuk 7

Terug naar de Orkneys. Er wordt dieper ingegaan op het maken van histogrammen. We beginnen met een frequentietabel, kiezen een klassebreedte en maken een histogram of polygoon. Te kleine of te grote klassebreedtes leiden tot ongewenste verschijnselen.

Soms wordt gewerkt met een somfrequentietabel. Deze heeft b.v. als voordeel dat in één oogopslag te zien is hoeveel mensen jonger dan 30 jaar zijn, in het geval van Papa Westray.

Enkele begrippen: frequentietabel, somfrequentietabel, frequentiepolygoon, somfrequentiepolygoon.

Hoofdstuk 8

Een somfrequentiepolygoon is een cumulatieve grafiek, d.w.z. als je bij 30 op de x -as zit is de waarde die daarbij hoort, de som van alle 'eerdere' waarden, bij Papa Westray dus alle mensen jonger dan 30.

Een bijzondere cumulatieve grafiek is de Lorenz-kromme, die gebruikt kan worden om de mate waarin twee zaken gelijkmatig verdeeld zijn te illustreren.

Voorbeelden: de spreiding van de bevolking over een land;
de spreiding van inkomens over de bevolking.

Enkele begrippen: cumulatief, Lorenz-kromme, Gini-coëfficiënt.

Hoofdstuk 9

Een tamelijk bijzondere grafiek is de *driehoeksgrafiek*. Deze kan geschikt zijn als je te maken hebt met drie zaken die tezamen constant zijn. Zeker niet altijd een erg duidelijke grafiek.

Enkele begrippen: driehoeksgrafiek.

Hoofdstuk 10

Hierin vinden we wat driedimensionale grafieken. De turftabellen die daarbij horen zijn tweedimensionaal. Zowel histogrammen als echte grafieken (computer) worden bekeken. Een puntendiagram - dat heel veel lijkt op een tweedimensionale turftabel - kan duidelijkheid verschaffen omtrent het verband of correlatie tussen twee zaken.

Enkele begrippen: 2-dimensionale turftabel, puntendiagram, correlatie, 3-dimensionaal histogram, 3-dimensionale grafiek.

Hoofdstuk 11

Dit is een voorbeeld van een wetenschappelijk artikel waarbij rijkelijk gebruik wordt gemaakt van grafieken.

Geen nieuwe begrippen.

De bijlage gaat over de bevolkingsspreiding op Indonesië en sluit aan bij het slot van hoofdstuk 6.