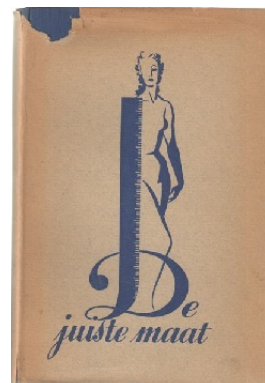


1.2 Statistische variabelen

Inleiding

In 1947 deden de wiskundigen H. Freudenthal en J. Sittig in opdracht van warenhuis De Bijenkorf onderzoek naar de lichaamsafmetingen van 5001 vrouwen. De Bijenkorf wilde op grond daarvan een maatsysteem voor kleding opzetten. Beide wiskundigen pasten op de grote verzameling kwantitatieve gegevens statistische methoden toe. Ze maakten overzichtelijke verdelingen van de lengtes, de gewichten, de voetslengte, de kniehoogte, en dergelijke van hun steekproef. De resultaten werden gepubliceerd in een boek met titel 'De juiste maat'.



Figuur 1

Je leert in dit onderwerp

- onderscheid maken tussen kwalitatieve en kwantitatieve variabelen;
- het verschil tussen relatieve en absolute frequenties.

Voorkennis

- het begrip statistische variabele;
- de basisvaardigheden van het werken met MS-Excel.

Verkennen

Opgave V1

Hier staan de gegevens over de gezinsomvang van leerlingen uit twee 4-havoklassen:

H4A: 3, 5, 2, 6, 5, 5, 2, 3, 4, 6, 4, 3, 5, 2, 5, 4, 4, 3, 4, 4, 5, 3, 7, 4, 4, 5, 2, 4

H4B: 4, 2, 5, 3, 4, 2, 4, 3, 5, 2, 3, 4, 6, 3, 3, 4, 3, 5, 4, 2, 4, 2, 5, 3

- a Van deze data is een tabel gemaakt. Verklaar de getallen in deze tabel.

gezinsomvang (aantal personen)	frequenties H4A	frequenties H4B
2	4	
3	5	
4	9	
5	7	
6	2	
7	1	

Tabel 1

- b Vul de tabel ook in voor H4B.

Je moet nog verder met de tabel rekenen, dus het is handig om met Excel te werken.

In de frequentietabel staan de absolute frequenties. Door ze allemaal te delen door het totale aantal leerlingen krijg je relatieve frequenties die je ook kunt geven als percentages. Je kunt dan gemakkelijker beide tabellen vergelijken.

- c Maak een nieuwe tabel met relatieve frequenties.
- d Beschrijf de verschillen tussen H4A en H4B.

Uitleg

Er zijn grofweg twee soorten statistische variabelen:

- variabelen die in een getal zijn uit te drukken, bijvoorbeeld de lengte van een mens, de hoogte van het inkomen, de omvang van het gezin; dit zijn kwantitatieve variabelen;
- variabelen die een kenmerk zijn en niet in een getal zijn uit te drukken, bijvoorbeeld het geslacht, de kleur ogen, de godsdienst, de bloedgroep, de naam; dit zijn kwalitatieve variabelen.

Het is vaak handig om je gegevens te ordenen in frequentietabellen. Daarin hebben niet alleen de data een logische volgorde, maar is ook meteen te zien hoe vaak een bepaalde waarde van de statistische variabele voorkomt, de absolute frequentie. Om twee (verschillende) datasets goed te kunnen vergelijken, is het handiger om met relatieve frequenties (in procenten) te werken, dat wil zeggen de absolute frequenties gedeeld door het totale aantal, eventueel in procenten.

Opgave 1

In deze opgave wordt telkens een variabele beschreven die betrekking heeft op de werknemers van een bepaald bedrijf. Geef aan of het een kwalitatieve of een kwantitatieve variabele betreft. Geef ook aan of het maken van een frequentietabel zinvol is.

- De schoenmaat.
- De lengte.
- Het inkomen in euro.
- De burgerlijke staat.
- Het aantal dienstjaren bij het bedrijf.
- Het aantal personen waaruit het huishouden van de betreffende werknemer bestaat.

Opgave 2

In deze tabel zie je de gegevens over de gezinsomvang van leerlingen uit twee 4-havoklassen.

- Welk probleem doet zich voor als je de frequenties van beide groepen vergelijkt?
- Wat zou je met deze frequenties nog moeten doen?
- Maak een nieuwe tabel met relatieve frequenties.
- Hoeveel procent gezinnen met 4 of meer kinderen hoort bij H4A? En bij H4B?

gezinsomvang (aantal personen)	frequenties	
	H4A	H4B
2	4	5
3	5	7
4	9	7
5	7	4
6	2	1
7	1	0

Tabel 2

Theorie en voorbeelden

Om te onthouden

Er bestaan verschillende soorten variabelen en dus verschillende soorten data. Een belangrijk onderscheidend kenmerk is:

- Een **kwantitatieve variabele** kan wel in een getal worden uitgedrukt (zoals lengte, hoogte van het inkomen, omvang van het gezin en dergelijke).
- Een **kwalitatieve variabele** beschrijft een bepaald kenmerk dat niet noodzakelijk in een getal wordt uitgedrukt (bijvoorbeeld geslacht, kleur ogen, godsdienst, bloedgroep, naam en dergelijke).

Bij veel statistische variabelen kun je een **frequentietabel** maken. Door te tellen hoe vaak een bepaalde waarde van die variabele voorkomt, krijg je de **absolute frequentie** van die waarde. Deel je die absolute frequentie door het totale aantal waarden, dan krijg je de **relatieve frequentie** van die waarde. Relatieve frequenties (vaak in procenten uitgedrukt) maken het vergelijken van twee datasets gemakkelijker.

Veel van de tabellen in deze paragraaf zijn beschikbaar als bestand. In het Practicum **Statistiek: Data presenteren en vergelijken** leer je hoe je met statistische gegevens kunt werken.

Voorbeeld 1

In 1947 hielden de wiskundigen Freudenthal en Sittig een **statistisch onderzoek ten behoeve van een nieuw maatsysteem voor vrouwenkleding** in opdracht van het warenhuis De Bijenkorf. Onder andere maten zij de mouwlengte van 5001 vrouwen in cm nauwkeurig. Je ziet een frequentietabel met de data.

mouwlengte	frequentie
49	3
50	11
51	22
52	53
53	89
54	163
55	250
56	405
57	519
58	660
59	578
60	653
61	560
62	421
63	260
64	159
65	106
66	52
67	18
68	15
69	3
70	0
71	1

- Welke absolute en welke relatieve frequentie horen er bij een mouwlengte van 60 cm?
- Hoeveel procent van deze vrouwen had een mouwlengte vanaf 64 cm?

Antwoord

- Er zijn 653 vrouwen met een mouwlengte van 60 cm, de absolute frequentie is dus 653. De relatieve frequentie is nu $\frac{653}{5001} \approx 0,131$. Dat is ongeveer 13%.
- Een mouwlengte vanaf 64 cm betekent dat je met de mouwlengtes 64, 65, ..., 71 te maken hebt. Het gaat daarbij om $159 + 106 + 52 + 18 + 15 + 3 + 0 + 1 = 354$ vrouwen. En dat is: $\frac{354}{5001} \approx 0,071$. Dat is ongeveer 7%.

Opgave 3

Bekijk de tabel uit **Voorbeeld 1**.

Tabel 3

- Is mouwlengte een kwalitatieve of kwantitatieve statistische variabele?
 - kwalitatieve
 - kwantitatieve
- Maak zelf een tabel met relatieve frequenties voor de mouwlengtes.
- Welk nut heeft de tabel voor De Bijenkorf, denk je?

Opgave 4

Bekijk de kruistabel die Freudenthal en Sittig maakten van de mouwlengte en de kniehoogte in cm van 5001 vrouwen.

		Statistisch onderzoek 1947																						
		De Bijenkorf - Amsterdam																						
		5001 vrouwen met leeftijd > 18 jr.																						
		kniehoogte in cm																						
		34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54		
mouwlengte in cm	71																			1			1	
	70																							
	69														1			1				1	3	
	68										2	2	1		3	3	2	2					15	
	67											3	1	2	4	2	2	3					18	
	66											4	5	5	4	6	10	7	7	3	1		1	52
	65							1		2	7	3	15	13	26	16	11	6	3	1	1	1	106	
	64					1		2	2	5	5	13	22	26	31	22	18	4	4	4			159	
	63						3	1	5	5	20	36	44	48	40	29	16	6	4	2	1		260	
	62						2	6	15	21	61	61	84	62	59	27	14	6	1	2			421	
	61						3	2	6	29	43	65	99	111	91	61	33	9	7	1			560	
	60						3	5	17	40	68	112	117	121	84	49	31	2	2	2			653	
	59			1	1	2	11	20	60	110	108	88	66	72	22	12	3	1		1			578	
	58				3	8	21	50	72	110	141	103	90	34	22	5	1						660	
	57			1	2	9	28	60	80	92	87	72	53	20	12	2	1						519	
	56		1	2	4	13	30	50	79	82	60	40	32	8	3		1						405	
55			3	6	11	27	38	46	46	39	14	16	1	3								250		
54			4	7	11	17	32	37	23	14	8	4	1	1	3	1						163		
53			3	6	6	16	12	17	15	8	6											89		
52		1	2	3	5	9	13	10	5	4	1											53		
51		1	2		4	4	5	2	2	1	1											22		
50	1					3		1	3		3											11		
49			1	2																		3		
		1	3	19	34	76	178	313	495	632	738	675	665	466	343	195	88	45	18	12	2	3	5001	

Figuur 2

- Hoeveel procent van de onderzochte vrouwen heeft een mouwlengte van 60 cm?
- Welke combinatie van mouwlengte en kniehoogte komt het meeste voor?

- c Hoeveel procent van de vrouwen met een mouwlengte van 60 cm heeft een kniehoogte van 50 cm?
- d Hoeveel procent van de onderzochte vrouwen heeft een mouwlengte van meer dan 68 cm en een kniehoogte van meer dan 50 cm?
- e Hoeveel bedroeg de gemiddelde mouwlengte van de 5001 vrouwen?
- f Hoeveel bedroeg de gemiddelde kniehoogte?

Opgave 5

Stel dat je het onderzoek van Freudenthal en Sittig dit jaar opnieuw zou uitvoeren. Je kunt nu (met de moderne hulpmiddelen) veel meer vrouwen in De Bijenkorf laten opmeten. Als je de resultaten wilt vergelijken met die van Freudenthal en Sittig, is het verstandig om met relatieve frequenties te werken. Laat met een rekenvoorbeeld zien waarom dat zo is.

Voorbeeld 2

In de tabel in het bestand **schoenmaten** vind je deze turftabel van de schoenmaten van mensen in de schoenenwinkel van v.O. in H. Maak zelf een frequentieverdeling (in Excel of een soortgelijk programma) door de bijbehorende kolommen met absolute frequenties en die met relatieve frequenties te maken.

Schoenmaten op zaterdag 14 juli 2007 in H bij schoenenzaak v.O.	
schoenmaat	aantal
36	III
37	IIIII
38	IIIII IIII
39	IIIII IIIII IIIII III
40	IIIII IIIII II
41	IIIII III
42	IIIII IIIII
43	IIII
44	I

Figuur 3

Antwoord

Het is verstandig om het bestand eerst op je eigen pc op te slaan en er dan pas in te gaan werken! In het bestand **schoenmaten antwoord** vind je de absolute frequenties en de relatieve frequenties.

Ga na hoe je het programma gebruikt om het rekenwerk voor je te doen. Als je op een cel klikt, kun je zien welke formule er wordt gebruikt om de inhoud van die cel te berekenen.

Opgave 6

Bekijk **Voorbeeld 2**. Sla het Excelbestand eerst op je eigen pc op. Als je het vervolgens opent kun je er in werken.

- a Maak zelf in Excel een frequentietabel en een relatieve frequentietabel. Doe eventueel eerst het **Practicum**
- b Voeg bij de absolute frequenties één waarneming 36 toe. Hoeveel getallen veranderen er in de frequentietabel?
- c Hoeveel getallen veranderen er in de relatieve frequentietabel?

Verwerken

Opgave 7

Stel dat je een statistisch onderzoek houdt onder de leerlingen in je eigen jaargroep. Je verzamelt de volgende gegevens:

- het geslacht (mannelijk of vrouwelijk)
- de lichaamslengte
- het lichaamsgewicht
- de omvang van het gezin
- de afstand tot school
- de bloedgroep
- de kleur van de ogen
- de beoefende sport(en)

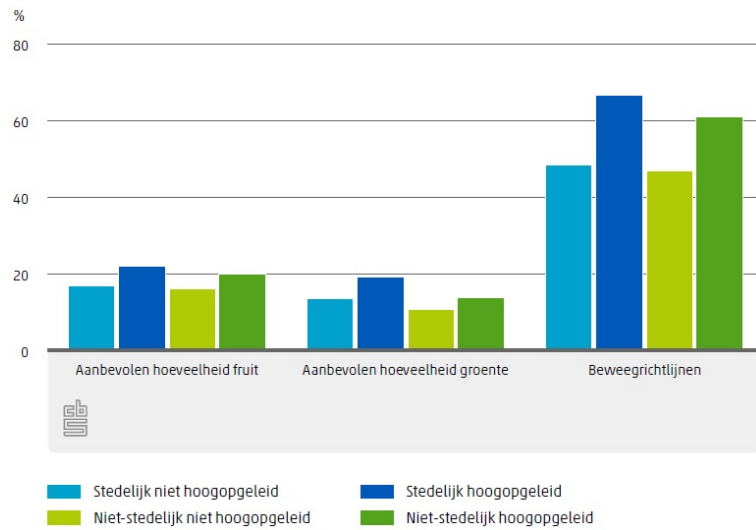
Sommige van deze variabelen geven een hoeveelheid of een grootte aan, andere geven een kenmerk aan.

- a Welke statistische variabelen in dit onderzoek zijn kwalitatief?
- b Welke statistische variabelen in dit onderzoek zijn kwantitatief?

Opgave 8

Bekijk het diagram van het CBS uit de Jongerenmonitor 2019. Het geeft de hoeveelheid jongeren weer die zich houden aan een drietal gezondheidsrichtlijnen.

7.2.5 Volgen richtlijnen, 18-24 jaar, 2015/2018



Figuur 4

- a Is de frequentie hier absoluut of relatief?
 - A. relatief
 - B. absoluut
- b Welke zes vragen zijn er minimaal gesteld aan de jongeren om dit diagram te kunnen maken?
- c Welke statistische variabelen in dit onderzoek zijn kwantitatief?
- d Welke statistische variabelen in dit onderzoek zijn kwalitatief?

Opgave 9

Voor een biologiepracticum moet het aantal slakken op een stuk grond worden geteld. Het stuk grond wordt daartoe in stukken van 1 m² verdeeld. Iedere leerling telt het aantal slakken op vier van die stukken. Hier zie je de resultaten.

aantal slakken per m ²	2	3	4	5	6	7	8	9
frequentie	16	14	7	4	2	3	1	1

Tabel 4

- a Om welke statistische variabele gaat het hier?
- b Hoeveel m² is de oppervlakte van het stuk grond?
- c Hoeveel leerlingen hebben er geteld?
- d Hoeveel slakken zijn er in totaal geteld?
- e Hoeveel slakken zijn er gemiddeld per m² gevonden?

Opgave 10

De tabel laat het aantal geslaagden zien op havo en vwo gedurende drie schooljaren.

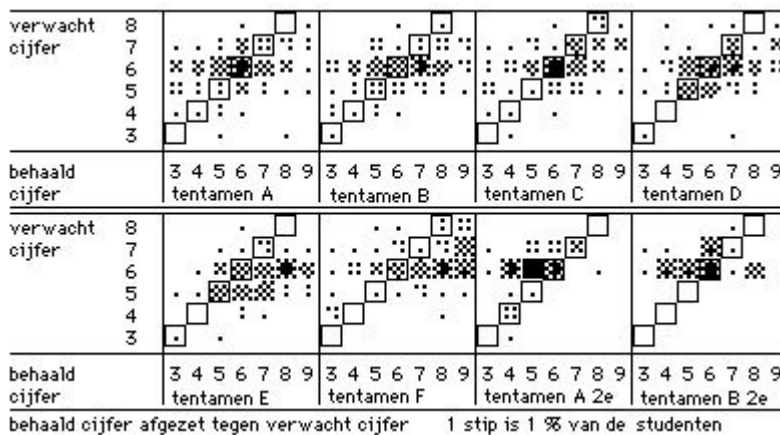
Voortgezet onderwijs; geslaagden naar onderwijssoort									
Onderwerpen	Aantal examenkandidaten			Aantal geslaagden			Percentage geslaagden		
	2005/06	2006/07	2007/08 ^a	2005/06	2006/07	2007/08 ^a	2005/06	2006/07	2007/08 ^a
Periodes	aantal	aantal	aantal	aantal	aantal	aantal	%	%	%
Vwo	31633	32961	35150	29287	30331	32398	93	92	92
Vwo natuurprofiel totaal	14967	15940	17233	13750	14570	15723	92	91	91
Vwo natuur en techniek	3765	3510	3970	3476	3213	3662	92	92	92
Vwo natuur en gezondheid	9641	10268	11259	8771	9275	10097	91	90	90
Vwo natuur ongedeeld	1561	2162	2004	1503	2082	1964	96	96	98
Vwo maatschappijprofiel totaal	16657	17007	17908	15528	15747	16666	93	93	93
Vwo economie en maatschappij	10087	10142	10801	9413	9392	10112	93	93	94
Vwo cultuur en maatschappij	6359	6623	6858	5911	6126	6309	93	92	92
Vwo maatschappij ongedeeld	211	242	249	204	229	245	97	95	98
Havo	42983	44560	46313	38105	39755	41371	89	89	89
Havo natuurprofiel totaal	12008	12641	13634	10743	11181	12306	89	88	90
Havo natuur en techniek	3699	3618	3871	3319	3203	3492	90	89	90
Havo natuur en gezondheid	7360	7853	8704	6546	6906	7815	89	88	90
Havo natuur ongedeeld	949	1170	1059	878	1072	999	93	92	94
Havo maatschappijprofiel totaal	30974	31911	32679	27361	28571	29065	88	90	89
Havo economie en maatschappij	15705	16203	16959	13607	14321	14993	87	88	88
Havo cultuur en maatschappij	14751	15182	15264	13294	13778	13649	90	91	89
Havo maatschappij ongedeeld	518	526	456	460	472	423	89	90	93

Figuur 5

- Welke variabelen worden er onderzocht? Welke soort variabele betreft het?
- Waar vind je absolute frequenties en waar relatieve frequenties?
- Op de havo is in 2007/2008 het geslaagdenpercentage 89. Toon dit aan door een berekening.
- Het aantal examenkandidaten op de havo neemt absoluut gezien toe. Neemt het aantal examenkandidaten relatief bekeken ook toe?

Opgave 11

Je ziet kruistabellen van zes tentamens en twee hertentamens.

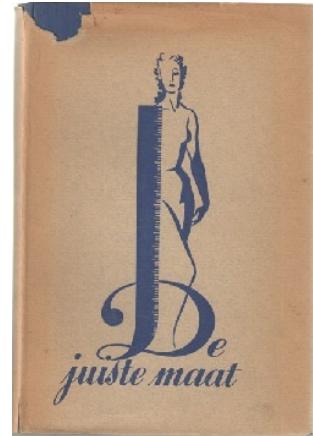


Figuur 6

- Welke statistische variabelen worden hier vergeleken?
- Hoe zie je of een tentamen slechter gemaakt is dan verwacht?
- Welke tentamens werden duidelijk slechter gemaakt dan de meeste studenten hadden verwacht?
- Kun je zeggen dat de twee hertentamens moeilijker waren dan de bijbehorende eerste versies?

Toepassen

In 1951 verscheen bij uitgeverij Stafleu in Leiden het boek 'De Juiste Maat', met als ondertitel 'Lichaamsafmetingen van Nederlandse vrouwen als basis voor een nieuw maatsysteem voor damesconfectiekleding'. Auteurs van dit boek waren J. Sittig, Adviesbureau voor Toegepaste Statistiek, en prof. dr. H. Freudenthal, Rijksuniversiteit Utrecht. Het onderzoek was gehouden in opdracht van N.V. Magazijn De Bijenkorf, Amsterdam. In het kader van dit onderzoek zijn bij 5001 vrouwelijke klanten van de Bijenkorf vijftien lichaamsmaten opgemeten. Vervolgens is gekeken welke van deze maten het meest bruikbaar zijn om een **maatsysteem voor kleding** op te baseren.



Figuur 7

Bekijk een deel van de uitkomst van het onderzoek in het bestand **Statistiek Bijenkorf 1947**.

Opgave 12

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

- Om welk type statistische variabelen gaat het hier?
- Welke klassenindeling is er gebruikt voor de variabele *lengte*? En voor de variabele *gewicht*?
- Maak een tabel van de relatieve frequenties van de lengtes van de vrouwen.
- Hoeveel procent van de vrouwen heeft een lengte vanaf (afgerond) 1,56 m tot en met 1,68 m?

Opgave 13

Je ziet een kruistabel voor lengte en gewicht die de wiskundigen Freudenthal en Sittig maakten van 5001 vrouwen in het kader van hun onderzoek naar een beter systeem voor kledingmaten van vrouwen in opdracht van warenhuis De Bijenkorf uit 1947. Bekijk de tabel.

		Statistisch onderzoek 1947																									
		De Bijenkorf - Amsterdam																									
		5001 vrouwen met leeftijd > 18 jr.																									
		gewicht in kg																									
		42	45	48	51	54	57	60	63	66	69	72	75	78	81	84	87	90	93	96	99	102	105	108	111	114	
lengte in cm	186								1					1	2	2					1						1
	183								1	1	5	4	2	3	2				3								6
	180																			3							21
	177				1	2	2	4	13	19	15	10	9	4	4	4	1	1			4		1				94
	174				1	1	6	15	17	22	22	20	13	12	5	3	7	4	2	2	2	1	1	1			156
	171				2	7	25	41	50	68	34	32	22	25	18	8	9	8	9	1			1		1	1	362
	168			1	13	24	39	59	74	73	55	43	35	22	40	20	20	20	5	8	4	3	2	1	1	1	563
	165		1	3	20	52	109	108	92	85	85	64	55	48	44	28	19	10	8	6	6	2			1		846
	162		2	10	42	92	103	127	100	85	90	59	56	47	28	38	13	11	9	4	3	3	1		1		924
	159	1	3	22	58	79	80	83	96	89	82	59	53	37	32	23	12	10	6	4	3	2					834
	156	1	7	16	35	45	62	57	70	58	51	59	38	28	25	9	10	6	6		1	2					586
	153	1	2	18	24	22	39	42	44	33	32	20	21	9	10	7	9	2	2	1	2		1	1			342
	150		7	11	15	12	19	19	20	22	19	8	13	4	2	1	3	1									176
	147	3	5	3	4	10	8	6	9	4	4	6	2		1	1				1							67
	144			2	1	1		2	3	1	1	1	1	1													14
141	1				3	1					1	1														8	
138							1																			1	
		7	27	86	216	350	493	564	591	560	495	386	322	242	213	142	103	76	48	27	24	13	7	3	3	3	5001

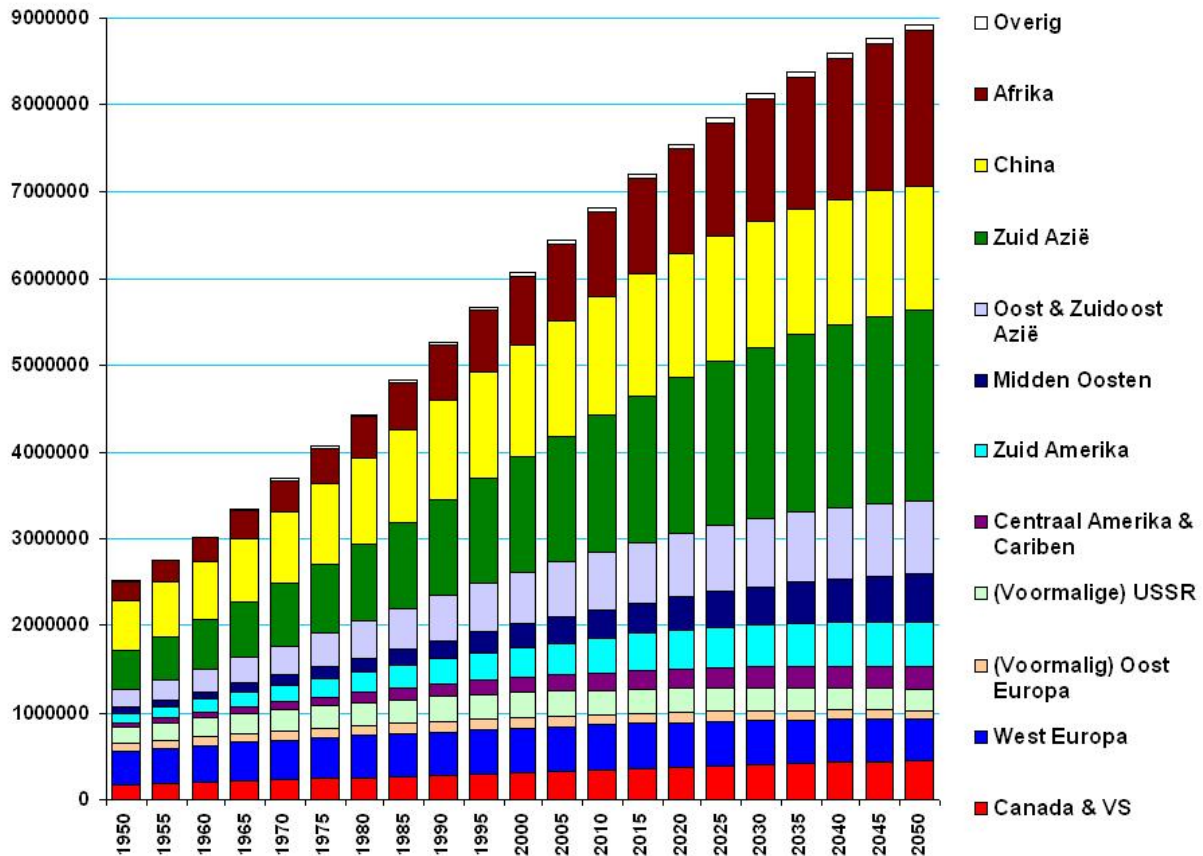
Figuur 8

- Welke statistische variabelen zie je?
- De variabelen worden afgerond. Welke frequentie staat er in de cel waar een vrouw van 170 cm en 70 kg is ingedeeld? Is dit getal absoluut of relatief?
- Hoeveel procent van de vrouwen is kleiner dan 151,5 cm?
- Hoeveel procent is naar schatting kleiner dan 153 cm?

Testen

Opgave 14

Bekijk de figuur met betrekking tot de geografische spreiding van de wereldbevolking. Het stond op 22 september 1999 in NRC/Handelsblad.



Figuur 9

- Welke statistische variabelen zie je hier? Zijn ze kwantitatief of kwalitatief?
- Hoeveel miljard mensen zijn er volgens dit diagram in 100 jaar bijgekomen?
- In welk deel van de wereld komt er absoluut gezien het grootste aantal mensen bij in 100 jaar?
- In welk deel van de wereld komt er relatief gezien het grootste aantal mensen bij in 100 jaar?
- In 1950 woonde ongeveer de halve wereldbevolking in Azië. Hoe zit dat volgens deze voorspelling in 2050?

Practicum

Met **Excel** (een spreadsheetprogramma, een rekenblad) werken is bij statistiek eigenlijk onontbeerlijk. Je kunt er grote hoeveelheden gegevens in kwijt.

Als je nog nooit of heel weinig met Excel hebt gewerkt, doe dan eerst de practica:

- [Basistechnieken Excel](#)
- [Diagrammen met Excel](#)

Speciaal voor het werken met Excel bij het presenteren van tabellen en diagrammen met statistische data is het practicum:

- [Data presenteren](#)



© 2021

Deze paragraaf is een onderdeel van het Math4All wiskundemateriaal.

Math4All stelt het op prijs als onvolkomenheden in het materiaal worden gemeld en ideeën voor verbeteringen in de content of dienstverlening kenbaar worden gemaakt.

Email: f.spijkers@math4all.nl

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