

1 Statistiek, gegevens en een kritische houding

WAT WE GAAN BEHANDELEN

- Waar gaat statistiek over?
- Hoe wordt statistiek in de praktijk toegepast?
- We bespreken het verband tussen statistiek en gegevens.
- We maken verschil tussen populaties en steekproeven.
- We introduceren de belangrijkste elementen van een proces.
- We maken verschil tussen beschrijvende en verklarende statistiek.
- We benadrukken het belang van een kritische houding.

Statistiek in de praktijk

Enkele jaren had het Amerikaans tv-programma *20/20* een aflevering met de titel *Feit of Fictie? Ontmaskering van zogenaamde enquêtes*. De focus lag op wat we tegenwoordig ‘nepnieuws’ of ‘alternatieve feiten’ zouden noemen: misinformatie die vaak breed wordt uitgemeten in de media (kranten, tijdschriften, tv-programma’s, sociale media zoals Twitter en Instagram enzovoort). In het programma werden verschillende misleidende (en mogelijk onethische) enquêtes gepresenteerd, die waren uitgevoerd door bedrijven of belangengroepen die een specifiek doel voor ogen hadden. Enkele van deze onderzoeken worden genoemd in tabel 1.1, samen met meer recente misleidende studies die in productadvertenties zijn gebruikt.

Voor deze *Statistiek in de praktijk* richten we ons op een onderzoek dat gesponsord is door Mars Corp. en gepubliceerd is in het tijdschrift *Nature Neuroscience* (december 2014). Onderzoekers van de Columbia University onderzochten of cacaosupplementen een hersengebied zouden kunnen verbeteren dat met de leeftijd achteruitgaat, en dat geassocieerd wordt met leeftijdsgebonden geheugenverlies. Ze concludeerden dat cacaosupplementen inderdaad het cognitieve vermogen van ouderen kunnen stimuleren. Vergelijkbare studies beweren dat chocolade het risico op hart- en vaatziekten vermindert en helpt bij

gewichtsverlies. Media zoals *The New York Times* brachten deze resultaten met koppen als: 'Wil je je geheugen verbeteren? Overweeg om chocolade te eten!', 'Goed nieuws voor chocoladeliefhebbers: hoe meer je eet, hoe lager het risico op hartziekten,' of simpelweg 'Chocolade is goed voor je.' Deze berichten hebben de vraag naar chocolade waarschijnlijk vergroot. Terwijl de totale snoepverkoop daalde, steeg de verkoop van chocolade in de vs van \$14,2 miljard in 2007 naar \$18,9 miljard in 2017.

Bij de gezondheidsclaims rond chocolade kunnen vraagtekens geplaatst worden. Zo ontdekte *Vox* (7 november 2017)¹ dat chocolade-bedrijven zoals Mars miljoenen hebben geïnvesteerd in wetenschappelijke studies en onderzoeksbeurzen om de voordelen van cacao te benadrukken. Van de 100 recente door Mars gesponsorde onderzoeken naar cacao, chocolade en gezondheid, waren er maar liefst 98 gunstig voor de snoepfabrikant. Dit hoge percentage doet volgens *Vox* twijfels rijzen over de kwaliteit van de studies, aangezien Mars en andere fabrikanten de positieve resultaten kunnen gebruiken om hun producten te promoten.

Analyse van een cacaostudie

De kritiek van *Vox* richtte zich onder meer op de chocoladestudie in *Nature Neuroscience*. De onderzoekers verdeelden 37 proefpersonen willekeurig over vier groepen:

1. Hoge dosis cacao + beweging (8 proefpersonen)
2. Hoge dosis cacao + geen beweging (11 proefpersonen)
3. Lage dosis cacao + beweging (9 proefpersonen)
4. Lage dosis cacao + geen beweging (9 proefpersonen)

Deelnemers in de hoge-dosisgroepen kregen dagelijks 900 mg cacaoflavonoïden, terwijl deelnemers in de lage-dosisgroepen slechts 10 mg kregen. De deelnemers in de bewegingsgroepen deden vier keer per week een uur aan aerobics. Na drie maanden werd onderzocht of cacaoflavonoïden cognitieve achteruitgang in een hersengebied dat geassocieerd wordt met geheugenverlies konden tegengaan. Dit gebeurde met MRI-scans en een objectherkenningstaak. Ook werd bekeken of beweging invloed had op het geheugen.

De onderzoekers concludeerden dat beweging geen effect had, maar cacaoflavonoïden wel. Deelnemers met een hoge dosis cacao lieten grotere verbeteringen in cognitieve prestaties zien dan die met een lage dosis. Maar volgens *Vox* trokken de onderzoekers conclusies die verder gingen dan hun data toelieten. Zo claimden ze dat cacaoflavonoïden leeftijdsgebonden geheugenverlies met 30 jaar konden terugdraaien. Dr. Henry Drysdale van het *Center for Evidence-Based Medicine* in Oxford gaf aanvullende kritiek. Hij wees erop dat het verbeteren van geheugen in drie maanden geen relevante maatstaf is voor werkelijke

leeftijdsgebonden achteruitgang. Ook benadrukte hij de noodzaak van een veel grotere steekproef en een langere onderzoeksperiode. Tot slot betoogde Drysdale dat het meten van hersenactiviteit in een MRI-scanner en het uitvoeren van een objectherkenningstaak weinig zeggen over het alledaagse functioneren van het geheugen.

De impact van misleidende claims

Uiteindelijk concludeerde Vox dat deze studie slechts aantoonde dat cacaosupplementen hersenactiviteit tijdelijk kunnen verbeteren op basis van specifieke en beperkte tests. Dit is verre van een geldige bevestiging dat cacao het geheugen verbetert.

Terug naar de aflevering van 20/20. Deze eindigde met een interview met Cynthia Crossen, auteur van *Tainted Truth: The Manipulation of Fact in America*. Ruim 20 jaar voordat de term ‘nepnieuws’ werd bedacht, waarschuwde Crossen:

‘Als iedereen cijfers misbruikt om ons bang te maken en ons ergens toe te bewegen, hoe goed dat doel ook is, verliezen we de kracht van cijfers. Onderzoek kan ons helpen betere beslissingen te nemen, maar dan moet het onderzoek wel betrouwbaar zijn. Daar maak ik me zorgen over.’

TABEL 1.1 Voorbeelden van ‘nepnieuws’

Nepnieuws (bron)	Werkelijke onderzoeksinformatie/ onjuistheid
Het eten van haverzemelen is een makkelijke en goedkope manier om je cholesterol te verlagen. (<i>Quaker Oats</i>)	Om je cholesterol te verlagen, zou je dieet uitsluitend uit haverzemelen moeten bestaan.
Een op de vier Amerikaanse kinderen onder de 12 heeft honger of loopt risico op honger. (<i>Food Research and Action Center</i>)	Gebaseerd op vragen als: ‘Eet je soms minder dan je eigenlijk zou willen?’, ‘Maak je de porties van een maaltijd weleens kleiner?’ en ‘Is het ooit voorgekomen dat je maar een paar verschillende soorten voedsel voor je kinderen op tafel zette omdat je geen geld meer had om eten te kopen?’
Er is een sterke link tussen de golfhandicap van een CEO en de prestaties van het bedrijf: hoe lager de handicap van de CEO (dus hoe beter de golfer), hoe beter het bedrijf presteert. (<i>New York Times</i>)	De enquête werd gestuurd naar de CEO’s van de 300 grootste Amerikaanse bedrijven, maar slechts 51 CEO’s antwoordden. De gegevens van enkele top-CEO’s werden weggelaten.
Voor de invoering van federale gezondheids-wetgeving werd verwacht dat 30% van de werkgevers ‘zeker’ of ‘waarschijnlijk’ zou stoppen met het aanbieden van zorgverzekeringen. (<i>McKinsey & Company Survey</i>)	De enquête werd online ingevuld door 1329 bedrijven uit de particuliere sector in de VS. De vragen waren zo gesteld dat het logisch leek om te stoppen met het aanbieden van zorgverzekeringen.

Nepnieuws (bron)	Werkelijke onderzoeksinformatie/ onjuistheid
Meer dan 80% van de tandartsen beveelt Colgate-tandpasta aan hun patiënten aan, aldus een reclame. (<i>Colgate-Palmolive Company</i>)	Bij het invullen van de enquête konden er meerdere tandpasta's aanbevolen worden. De Advertising Standards Authority oordeelde dat de reclame misleidend was, omdat deze impliceerde dat 80% van de tandartsen Colgate boven alle andere merken aanbeveelt. De reclame werd verboden en Colgate kreeg een boete.
Een reclame voor Kellogg's Frosted Mini-Wheats beweerde dat klinisch bewezen was dat de ontbijtgranen het concentratievermogen van kinderen met bijna 20% verbeteren. (<i>Kellogg Company</i>)	Slechts de helft van de kinderen liet enige verbetering zien in concentratievermogen. Slechts 1 op de 7 verbeterde met 18% of meer, en slechts 1 op de 9 verbeterde met 20% of meer. Kinderen die Frosted Mini-Wheats aten, werden vergeleken met kinderen die alleen water als ontbijt kregen. (<i>Kellogg betaalde \$4 miljoen om een rechtszaak over deze misleidende claim te schikken.</i>)
Walmart beweerde in een reclame dat het 'verantwoordelijk was voor een prijsdaling van 3,1%' en dat het bedrijf 'klanten meer dan \$700 per jaar bespaart.' (<i>Global Insight</i>)	Het Economic Policy Institute gaf aan dat de studie gebaseerd was op Walmarts invloed op de consumentenprijsindex (CPI), maar 60% van het CPI wordt bepaald door diensten, niet door producten die bij Walmart te koop zijn. (<i>De reclame werd ingetrokken wegens misleiding.</i>)
Comcast vroeg in een enquête aan consumenten welke aanbieder, Comcast of DIRECTV, meer HD-kanalen had. Ze kregen kanaallijsten te zien van DIRECTV (lijstnummer 387) en Comcast (lijstnummer 429). (<i>NAD Case Report No. 5208</i>)	De National Advertising Division (NAD) verwierp de enquête nadat ontdekt werd dat het hogere lijstnummer (#429) een subtiele, maar effectieve suggestie gaf dat Comcast meer kanalen aanbood.
NPR publiceerde een onderzoek waaruit bleek dat tieners die hun smartphone vijf uur of langer per dag gebruiken 71% meer kans hebben op een risicofactor voor zelfdoding of depressie. (<i>Journal of Abnormal Psychology</i>)	De data waren gebaseerd op door tieners zelf gerapporteerd gebruik van smartphones, wat vaak niet overeenkomt met daadwerkelijk gebruik dat met een app wordt gemeten. Daarnaast waren de metingen van verslaving en zelfdodingsrisico gebaseerd op onbetrouwbare methoden.

In de volgende afleveringen van 'Statistiek in de praktijk' in dit hoofdstuk bespreken we, in verband met misleidende enquêtes en onderzoeken, een aantal kernbegrippen uit de statistiek, namelijk:

- het vaststellen van de populatie, de steekproef en de het trekken van conclusies;
- het vaststellen van de gegevensverzamelmethode en het gegevenstype;
- het kritisch beoordelen van de ethiek van een statistisch onderzoek.

1.1 De wetenschap statistiek

Waar denk je aan bij het woord statistiek? Aan enquêtes, werkloosheidscijfers van het Centraal Bureau voor de Statistiek of misschien aan het zodanig presenteren van cijfers dat ze een vals beeld schetsen (hierover bestaat een beroemd boekje: *How to lie with statistics*)? Of zie je statistiek alleen maar als een verplicht onderdeel van je opleiding? We hopen dat je na het bestuderen van dit boek het met ons eens bent dat statistiek een nuttige wetenschap is, met een vrijwel eindeloos scala aan toepassingen in het bedrijfsleven, bij de overheid, in de natuurwetenschappen en in de sociale wetenschappen. Ten slotte laten we de sleutelrol zien die statistiek speelt in het kritisch denken – tijdens college, op het werk of in het dagelijks leven. Daarbij zul je ook zien dat je met een kritische houding valse presentaties van resultaten kunt doorprikken.

De *Random House College Dictionary* definieert statistiek als ‘de wetenschap die zich bezighoudt met het verzamelen, classificeren, analyseren en interpreteren van informatie of gegevens’. Een statisticus is dus niet alleen maar iemand die werkloosheidscijfers berekent of de resultaten van een enquête in een tabel zet. Professionele statistici hebben een opleiding in *de wetenschap statistiek*. Dat wil zeggen dat ze opgeleid zijn in het verzamelen van numerieke informatie, in het analyseren van deze informatie, en in het trekken van conclusies hieruit. Verder bepalen statistici welke informatie bij een gegeven probleem van belang is, en hoe betrouwbaar de conclusies zijn die uit een onderzoek voortkomen.

Statistiek is de wetenschap van **gegevens**. Zij omvat het verzamelen, classificeren, samenvatten, organiseren, analyseren en interpreteren van numerieke informatie.

In de volgende paragraaf zul je een aantal voorbeelden tegenkomen van statistische toepassingen waarbij het gaat om het nemen van beslissingen en het trekken van conclusies.

1.2 Soorten statistische toepassingen

Statistiek betekent voor de meeste mensen ‘beschrijven met getallen’. Maandelijkse werkloosheidscijfers, het percentage mislukte levertransplantaties en het percentage leidinggevende vrouwen in een bepaalde bedrijfssector: dit zijn allemaal voorbeelden van een beschrijving van grote hoeveelheden gegevens. Vaak worden de gegevens geselecteerd uit een grotere verzameling waarvan we de kenmerken willen schatten. We noemen dit proces **steekproef trekken**. Zo zou je bijvoorbeeld de leeftijden kunnen

registreren van bezoekers van webwinkels die geïnteresseerd zijn in een bepaald product. Hiermee kun je dan een schatting maken van de leeftijdsverdeling van *alle* potentiële bezoekers die in dat product zijn geïnteresseerd. Die schatting zou je kunnen gebruiken om de reclame van een webshop af te stemmen op de juiste leeftijdsgroep. Je ziet dat het bij statistiek om twee verschillende processen gaat: (1) het beschrijven van gegevensverzamelingen en (2) het trekken van conclusies (schattingen, beslissingen, voorspellingen enzovoort) op basis van een steekproef. Deze twee belangrijkste gebieden waarop de statistiek wordt toegepast, noemen we de beschrijvende statistiek en de verklarende statistiek.

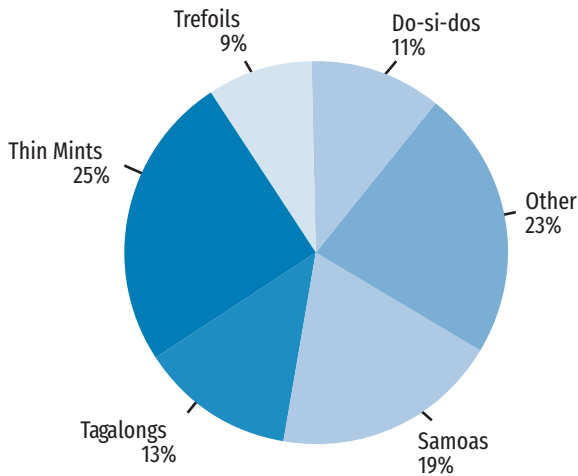
De **beschrijvende statistiek** gebruikt numerieke en grafische methoden om patronen in een gegevensverzameling te ontdekken, om de informatie in een gegevensverzameling samen te vatten en om deze informatie op een overzichtelijke manier te presenteren.

De **verklarende statistiek** gebruikt steekproefgegevens voor het schatten, het nemen van beslissingen en het voorspellen. De verklarende statistiek wordt ook wel **inductieve statistiek** of **inferentiële statistiek** genoemd.

Hoewel we in de volgende hoofdstukken zowel beschrijvende als verklarende statistiek zullen behandelen, zal het accent in dit boek liggen op de verklarende statistiek. Laten we beginnen met het bespreken van een aantal onderzoeken waarin statistische methoden werden toegepast.

Onderzoek 1. Bestverkopende koekjes van scoutingmeisjes²

Sinds 1917 hebben Amerikaanse scoutingmeisjes dozen met koekjes verkocht. Momenteel zijn er acht soorten te koop: Thin Mints, Samoas, Tagalongs, Peanut Butter Patties, Do-si-dos, Caramel DeLites, Peanut Butter Sandwiches en Trefoils. Elk van de naar schatting 150 miljoen dozen van de door de scoutingmeisjes verkochte koekjes wordt geclassificeerd naar soort. De resultaten zijn samengevat in figuur 1.1. Uit de grafiek kun je duidelijk aflezen dat Thin Mints het beste wordt verkocht (25%), gevolgd door Samoas (19%) en Tagalongs (13%). Omdat de figuur de omzet van de verschillende soorten koekjes beschrijft, is de grafiek een voorbeeld van beschrijvende statistiek.



FIGUUR 1.1 Grafiek van de omzet van koekjes van scoutingmeisjes (Minitab)

Onderzoek 2. Grof gedrag: maakt het echt uit op de werkvloer?³

Uit eerder onderzoek is gebleken dat grofheid op de werkvloer kan leiden tot vergeldings- en contraproductief gedrag. Er is echter weinig onderzoek gedaan naar de manier waarop grof gedrag de taakuitvoering van een slachtoffer beïnvloedt. In een onderzoek werden universiteitsstudenten die deelnamen aan een managementcursus willekeurig toegewezen aan een van twee experimentele condities: grofheidsconditie (45 studenten) en controlegroep (53 studenten). Elke leerling werd gevraagd om binnen vijf minuten zoveel mogelijk toepassingen voor een steen op te schrijven. Het totaal aantal toepassingen diende als prestatie maatstaf voor elke leerling. Bij studenten in de grofheidsconditie vertoonde de begeleider grof gedrag door de studenten als groep uit te schelden omdat ze onverantwoordelijk en onprofessioneel waren (vanwege een te laat binnengekomen klasgenoot). In de controlegroep werden er geen opmerkingen over de te laat arriverende klasgenoot gemaakt. Zoals je zou verwachten ontdekten de onderzoekers dat de prestatieniveaus van leerlingen in de grofheidsconditie over het algemeen lager waren dan de prestatieniveaus van leerlingen in de controlegroep. Daarom concludeerden ze dat grofheid op de werkplek de werkprestaties negatief beïnvloedt. Dit onderzoek is een voorbeeld van het gebruik van inferentiële statistiek. De onderzoekers gebruikten gegevens die verzameld waren over 98 studenten in een gesimuleerde werkomgeving om een conclusie te trekken over de prestatieniveaus van alle werknemers die werden blootgesteld aan grofheid op het werk.

Deze onderzoeken bieden praktijkvoorbeelden van toepassingen van de statistiek. Je ziet dat de gegevensverzameling wordt beschreven (Onderzoek 1), of dat er conclusies over een gegevensverzameling worden getrokken (Onderzoek 2).

HET GEBRUIK VAN TECHNOLOGIE, STATISTISCHE SOFTWARE

Er bestaan verschillende computerprogramma's voor het analyseren van gegevens, het uitvoeren van statistische technieken en het weer-geven in grafieken met behulp van grafische tools. De keuze voor een bepaald softwarepakket hangt af van verschillende factoren, waaronder de complexiteit van de analyse, het beschikbare budget en de voorkeuren van de gebruiker.

Excel is een veelgebruikt programma voor gegevensanalyse en biedt basisstatistieken, zoals gemiddelden, standaardafwijkingen en correlaties. Voor meer geavanceerde analyses zijn echter vaak add-ins of externe software nodig. Er zijn specifieke softwareprogramma's ontworpen voor het verwerken van grote datasets en voor het omgaan met complexe statistische modellen voor meer geavanceerde analysebehoefte en statistische analyses. Deze bieden een breed scala aan ingebouwde tools en functies die specifiek zijn afgestemd op statistici en data-analisten in verschillende werkvelden, zoals kwaliteitsmanagement, productie, gezondheidszorg en onderzoek. Enkele populaire opties voor Windows zijn Minitab, SPSS en SAS. Deze bieden gratis proefversies of academische licenties voor studenten en onderzoekers. JASP is een opkomend vergelijkbaar statistisch softwarepakket dat bekendstaat om zijn gebruiksvriendelijke interface en open-sourcebenadering. De statistische methodes die in dit boek aan bod komen, worden gedemonstreerd met behulp van output van deze softwarepakketten.

Voornoemde softwarepakketten zijn ontworpen voor statistische analyse en kwaliteitsverbetering in industriële en zakelijke omgevingen. In aanvulling daarop zijn er programmeertalen voor statistische toepassingen, waaronder gegevensanalyse, *machine learning*, kunstmatige intelligentie (*artificial intelligence*), webontwikkeling, automatisering, visualisatie en meer. R en Python zijn voorbeelden van krachtige programmeertalen die veel worden gebruikt voor statistische analyse en *data science*. Hierbij moet code worden geschreven voor het uitvoeren van analyses en het genereren van visualisaties. Dit valt buiten de reikwijdte van dit boek.

1.3 Basiselementen van de statistiek

Statistische methoden zijn in het bijzonder nuttig voor het bestuderen en analyseren van populaties bestaande uit experimentele eenheden.

Een **experimentele eenheid** is een object (bijvoorbeeld een persoon, een ding, een transactie of een gebeurtenis) waarvan we gegevens vastleggen.

Een **populatie** is een verzameling eenheden (bijvoorbeeld personen, objecten, transacties of gebeurtenissen) die we willen bestuderen.

Voorbeelden van populaties zijn: (1) *alle* werknemers in Vlaanderen, (2) *alle* kiesgerechtigden in Nederland, (3) *alle* personen die een bepaald merk mobiele telefoon hebben gekocht, (4) *alle* auto's die het afgelopen jaar van een bepaalde lopende band zijn gerold, (5) de *totale* voorraad reserveonderdelen van de onderhoudsdienst van de spoorwegen, (6) *alle* dagomzetten van de 'drive-through'-afdeling van een McDonald's restaurant in een bepaald jaar, en (7) de verzameling van *alle* ongelukken op een bepaald stuk snelweg gedurende een vakantieperiode. De eerste drie voorbeelden van populaties zijn verzamelingen van personen; de volgende twee (4, 5) zijn verzamelingen van objecten, de volgende (6) is een verzameling van omzetten, en de laatste (7) is een verzameling gebeurtenissen. Je ziet ook dat elke verzameling alle eenheden van de desbetreffende populatie bevat.

Als we een populatie bestuderen, concentreren we ons op een of meer kenmerken of eigenschappen van de eenheden van die populatie. We noemen zulke kenmerken variabelen. Zo kunnen we bijvoorbeeld geïnteresseerd zijn in de variabelen leeftijd, geslacht, inkomen en/of het aantal jaren opleiding van de mensen die op dit moment werkloos zijn in de Europese Unie.

Een **variabele** is een kenmerk of eigenschap van een eenheid uit een populatie.

De term 'variabele' is afgeleid van het feit dat het kenmerk kan variëren over de verschillende eenheden in een populatie.

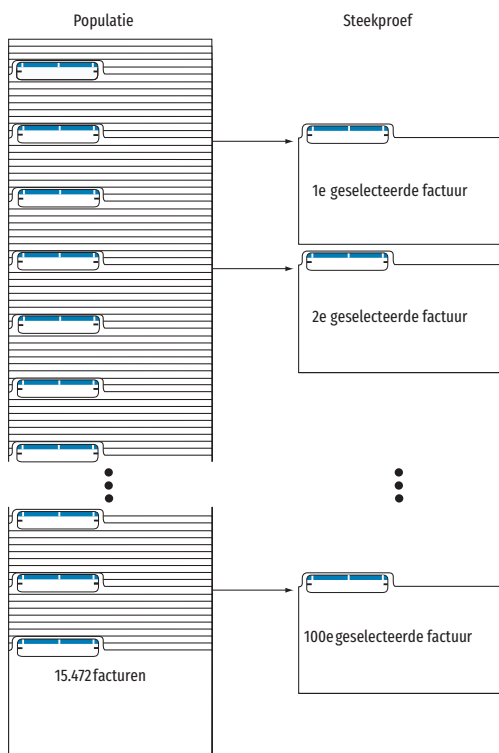
Vaak willen we de waarde van een variabele weergeven door middel van een getal. **Metten** is het proces waarbij we getallen toekennen aan variabelen. We kunnen bijvoorbeeld de voorkeur voor een voedselproduct meten door een consument te vragen een cijfer toe te kennen op een schaal van 1 tot 10 voor de smaak van het product. Of we zouden de leeftijd van de beroepsbevolking kunnen meten door aan elke werknemer te vragen hoe oud hij of zij is. In andere gevallen worden voor het meten instrumenten gebruikt, zoals stopwatches, weegschalen of schuifmaten.

Als de populatie die we willen bestuderen klein is, is het mogelijk om een variabele te meten voor elke eenheid in de populatie. Als we bijvoorbeeld het beginsalaris willen meten van iedereen die het afgelopen jaar een rechtenstudie

heeft afgerond aan de Katholieke Universiteit Leuven, is het mogelijk om elk salaris te bepalen. Als we een variabele meten voor iedere eenheid van een populatie, is het resultaat een zogenaamde **census** van de populatie. In de meeste gevallen zullen de populaties waarin we geïnteresseerd zijn echter veel groter zijn, met wellicht vele duizenden of zelfs een oneindig aantal eenheden. Voorbeelden van grote populaties zijn de zeven populaties die we zojuist hebben genoemd. Het zou veel te veel tijd en/of geld kosten om een census te houden voor zulke populaties. Een alternatief is het selecteren en bestuderen van een deelverzameling van de eenheden van die populatie.

Een **steekproef** is een deelverzameling van de eenheden van een populatie.

Veronderstel dat een boekhouding wordt gecontroleerd op fouten in de facturen. In plaats van het controleren van alle 15.472 facturen die in een bepaald jaar zijn verstuurd, kan een accountant een steekproef nemen van slechts 100 facturen (zie figuur 1.2). Hij kan dan de status (fout of geen fout) van elk van de facturen van de steekproef vastleggen (meten).



FIGUUR 1.2 Een steekproef van alle facturen van de onderneming