

Statistiek, gegevens en statistisch denken

► Inhoud

- I.1 De wetenschap statistiek
- I.2 Soorten statistische toepassingen in het bedrijfsleven
- I.3 Basiselementen van statistiek
- I.4 Soorten gegevens
- I.5 Gegevens verzamelen
- I.6 Opgaven

► Statistiek in de praktijk

- I.1 Kwaliteitsverbetering: bedrijven in de VS reageren op de Japanse uitdaging

WAT WE GAAN BEHANDELEN

Wat is statistiek? Is het een studiegebied, een reeks getallen die een samenvatting geeft van onze economische situatie, het koersverloop van een aandeel of de marktvoorwaarden in een bepaald gebied? Of is het, zoals in een populair boek (Tanur *et al.*, 1989) wordt gesuggereerd, 'een gids voor het onbekende'? We zullen in hoofdstuk I zien dat elk van deze omschrijvingen gebruikt kan worden om te leren begrijpen wat statistiek is. We zullen zien dat er twee soorten statistiek bestaan: *beschrijvende statistiek*, die zich richt op het ontwikkelen van grafische en numerieke samenvattingen om een bepaald verschijnsel in het bedrijfsleven te kunnen beschrijven, en *verklarende statistiek*, die

gebruikmaakt van deze numerieke samenvattingen om te helpen bij het nemen van beslissingen in het bedrijfsleven. Het hoofdonderwerp van dit boek is de verklarende statistiek. Dat betekent dat we je voornamelijk zullen laten zien hoe je statistiek kunt gebruiken om gegevens te interpreteren en deze te gebruiken voor het nemen van beslissingen. In veel banen in de industrie, bij de overheid, in de medische wereld en op andere gebieden moeten besluiten genomen worden die zijn gebaseerd op gegevensverzamelingen, waardoor een goed begrip van deze statistische methoden voor jou belangrijk praktisch voordeel kan opleveren.

1.1 De wetenschap statistiek

Wat betekent statistiek voor jou? Associeer je het met NIPO-enquêtes, werkloosheidscijfers, faalkansen of met de verdraaiing van de feiten met behulp van getallen (statistiek als leugen!)? Of zie je het alleen maar als een verplicht onderdeel van je opleiding? We hopen je ervan te kunnen overtuigen dat statistiek een zinnige en nuttige wetenschap is, met een vrijwel eindeloos scala aan toepassingen in het bedrijfsleven, bij de overheid, en in de exacte en sociale wetenschappen. Ook zul je zien dat statistieken alleen kunnen liegen als ze verkeerd worden gebruikt. Ten slotte laten we de sleutelrol zien die statistiek speelt in het kritisch denken – in de klas, op het werk of in het dagelijks leven. Ons doel is je ervan te overtuigen dat de tijd die je aan dit onderwerp besteedt een goede investering zal blijken te zijn.

De *Van Dale* 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 de kans op een storing bij een bepaald type auto berekent of de resultaten van een NIPO-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 de vorm van **gegevens**, in het analyseren van deze informatie, en in het trekken van conclusies hieruit. Verder bepalen statistici welke informatie in een gegeven probleem belangrijk is, en of de conclusies die uit een onderzoek voortkomen betrouwbaar zijn.

DEFINITIE 1.1

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 van statistische toepassingen in het bedrijfsleven en bij de overheid uit de praktijk tegenkomen, die te maken hebben met het nemen van beslissingen en het trekken van conclusies.

1.2 Soorten statistische toepassingen in het bedrijfsleven

Statistiek betekent voor de meeste mensen ‘beschrijven met getallen’. Maandelijkse werkloosheidscijfers, het percentage nieuwe bedrijven dat failliet gaat en het percentage motorstoringen dat in een bepaalde periode voor een bepaald type auto optreedt: dit zijn allemaal voorbeelden van statistische beschrijvingen van grote hoeveelheden gegevens die over een bepaald onderwerp zijn verzameld. Vaak worden de gegevens geselecteerd uit een grotere verzameling gegevens waarvan we de kenmerken willen schatten. We noemen dit selectieproces het *nemen van een steekproef*. Zo zou je bijvoorbeeld de leeftijden van een groep videotheekbezoekers kunnen verzamelen om een schatting te kunnen maken van de gemiddelde leeftijd van *alle* bezoekers van die videotheek. Dan zou je je schatting kunnen gebruiken om ervoor te zorgen dat de reclame van de videotheek op de juiste leeftijdsgroep is gericht. Je ziet dat statistiek twee verschillende processen omvat: (1) het beschrijven van gegevensverzamelingen en (2) het trekken van conclusies (schattingen, beslissingen, voorspellingen enzovoort) ten aanzien van deze gegevensverzamelingen, op basis van een steekproef. Deze

twee brede gebieden waarin de toepassingen van de statistiek kunnen worden onderverdeeld, noemen we: *beschrijvende statistiek* en *verklarende statistiek*.

DEFINITIE 1.2

Beschrijvende statistiek gebruikt numerieke en grafische methoden om patronen in een gegevensverzameling te ontdekken, om de informatie die uit een gegevensverzameling kan worden gewonnen samen te vatten, en om deze informatie op een overzichtelijke manier te presenteren.

DEFINITIE 1.3

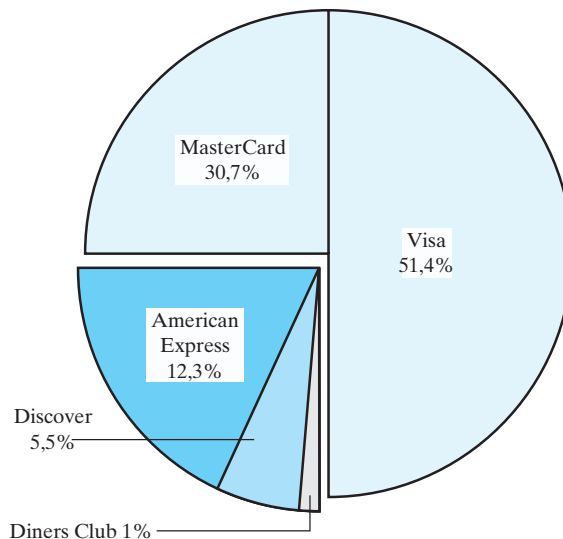
Verklarende statistiek gebruikt steekproefgegevens voor het maken van schattingen, het nemen van beslissingen en het doen van voorspellingen of voor andere generalisaties met betrekking tot een grotere verzameling gegevens.

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 onderzoeken van een aantal studies uit het bedrijfsleven, die het toepassen van statistische methoden illustreren.

► Studie I 'Marktaandeel VS van creditcards' (*The Nilson Report*, 8 okt. 1998)

The Nilson Report verzamelde gegevens over alle aankopen met creditcards en debitcards in de Verenigde Staten gedurende de eerste zes maanden van 1998. Het bedrag van elke aankoop werd geregistreerd en geassocieerd naar het type creditcard dat werd gebruikt. De resultaten zijn te zien in de grafiek van Associated Press, figuur 1.1. Uit de grafiek is duidelijk te zien dat de helft van de aankopen werd gedaan met een VISA-card en een kwart met een MasterCard. Omdat figuur 1.1 het type creditcard toont dat in alle aankopen met creditcard in de eerste helft van 1998 werd gebruikt, is deze grafiek een voorbeeld van beschrijvende statistiek.



FIGUUR 1.1: Creditcard marktaandeel VS
Bron: *The Nilson Report*, 8 okt. 1998

► Studie 2 'Storingen bij een autotype'

Een bepaald type automeerk wil een garantietermijn van 3 jaar of 100.000 kilometer gaan invoeren. Uit proeven met versnelde verouderingsprocessen bleek dat dit zonder al te veel kosten ingevoerd zou kunnen worden. Het bedrijf wilde echter in de praktijk controleren of het versnelde ouderdomsproces met de werkelijkheid overeenkwam. Hiertoe werd in het jaar 2000 van 1000 in dat jaar net op de weg gebrachte auto's genoteerd, wanneer er voor het eerst een belangrijke storing optrad. Voor een tweetal groepen storingen, bij de ophanging en bij de besturing, is de kilometerstand van de eerste storing binnen 3 jaar weergegeven in tabel 1.1.

TABEL 1.1: Aantal storingen bij een bepaalde kilometerstand

Km-stand bij eerste storing	Ophanging	Besturing
<40.000	0	1
40.000 - <50.000	1	2
50.000 - <60.000	3	1
60.000 - <70.000	5	6
70.000 - <80.000	4	2
80.000 - <90.000	2	0
90.000 - <100.000	6	9
100.000 - <110.000	29	41
>110.000	950	938

Uit tabel 1.1 kunnen we de conclusie trekken dat in hooguit 11% van de auto's in deze tabel binnen de voorgenomen garantietermijn een storing optreedt. Uit deze tabel is niet te zien of bij een en dezelfde auto beide storingen voorgekomen zijn. In dat geval is het percentage auto's nog minder. Als we daarnaast ook nog een tabel hebben waarin de kosten van deze reparaties zijn opgenomen, kunnen we een schatting maken van de kosten van deze garantietermijn. Indien de fabrikant in dat jaar de eerste is die met zo'n garantie komt, kan hij bepalen of het commercieel aantrekkelijk is om deze garantietermijn in te voeren.

► Studie 3 'Onderzoek naar benodigd aantal ambulances aan stationering'

In veel regio's in Nederland is de ambulancedienst per dorp of stad geregeld. Dit leidt vaak tot onnodig hoge kosten in plaats van het niet efficiënt inzetten van ambulances. Ook kan de rijtijd naar de plaats van het ongeval en vervolgens naar het dichtstbijzijnde ziekenhuis onnodig lang zijn. Om dit te onderzoeken heeft men in de kop van Noord-Holland gegevens verzameld van de gemiddelde tijd die een ambulance kwijt is om naar een ongeval te rijden en om vervolgens het slachtoffer naar het dichtstbijzijnde ziekenhuis te brengen, waar hij of zij adequaat kan worden behandeld. Aan de hand van deze gegevens kan men onder andere bepalen hoeveel ambulances er nodig zijn om dit vervoer te regelen. Verder kunnen we ook een afstandstabel ontwikkelen, zodat we niet alleen een tabel in benodigde tijd, maar ook in afgelegde kilometers krijgen. Door middel van een rekenprogramma kunnen we vervolgens analyseren wat de gevolgen in rijtijd en benodigd aantal ambulances zal zijn als we een aantal verschillende stationeringsplaatsen samenvoegen tot één stationeringsplaats.

Net als studie 2 is deze studie een voorbeeld van *verklarende statistiek*. De gereden kilometers van de ambulances die geselecteerd zijn uit een aantal stationeringsplaatsen worden gebruikt om conclusies te trekken met betrekking tot de gereden kilometers vanuit een centrale stationeringsplaats.

Deze studies vormen drie praktijkvoorbeelden van het gebruik van statistiek in het bedrijfsleven, de economie en bij de overheid. Merk op dat elke studie een analyse van de gegevens bevat, hetzij om de gegevens te beschrijven (studie 1), hetzij om conclusies te trekken over een verzameling gegevens (studies 2 en 3).

1.3 Basiselementen van statistiek

Statistische methoden zijn in het bijzonder nuttig voor het bestuderen, analyseren en kennis verkrijgen van *populaties*.

DEFINITIE 1.4

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

Voorbeelden van populaties zijn: (1) *alle* werknemers in Nederland, (2) *alle personen* die een bepaald merk mobiele telefoon hebben gekocht, (3) *alle* geregistreerde storings van een bepaald type cv in de regio Noord-Holland. (4) auto's die afgelopen jaar van een bepaalde lopende band zijn gekomen, (5) de *totale* voorraad reserveonderdelen van de onderhoudsdienst van KLM, (6) de *totale* verkoop van de 'drive-through'-afdeling van een restaurant van McDonald's in een bepaald jaar, en (7) de verzameling van *alle* ongelukken op een bepaald stuk snelweg gedurende een vakantieperiode. De eerste twee voorbeelden van populaties (1-2) zijn verzamelingen van personen; de volgende drie (3-5) zijn verzamelingen objecten, de volgende (6) is een verzameling transacties, en de laatste (7) is een verzameling gebeurtenissen. Merk ook op 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 Nederland.

DEFINITIE 1.5

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

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

Als we een bepaalde variabele bestuderen, is het handig om hiervoor een numerieke representatie te vinden. Zo'n numerieke representatie is echter vaak niet zonder meer voorhanden, daarom speelt het meten een belangrijke rol in statistisch onderzoek. **Metten** is het proces waarbij we getallen toekennen aan variabelen van individuele eenheden van een populatie. We zouden bijvoorbeeld de voorkeur voor een voedselproduct kunnen meten door een consument te vragen de smaak van het product een cijfer toe te kennen op een schaal van 1 tot 10. Of we zouden de leeftijd van de beroepsbevolking kunnen meten door eenvoudigweg aan elke werknemer te vragen hoe oud hij of zij is. In andere gevallen worden voor het meten instrumenten gebruikt, zoals stopwatches, weegschalen en schuifmaten.

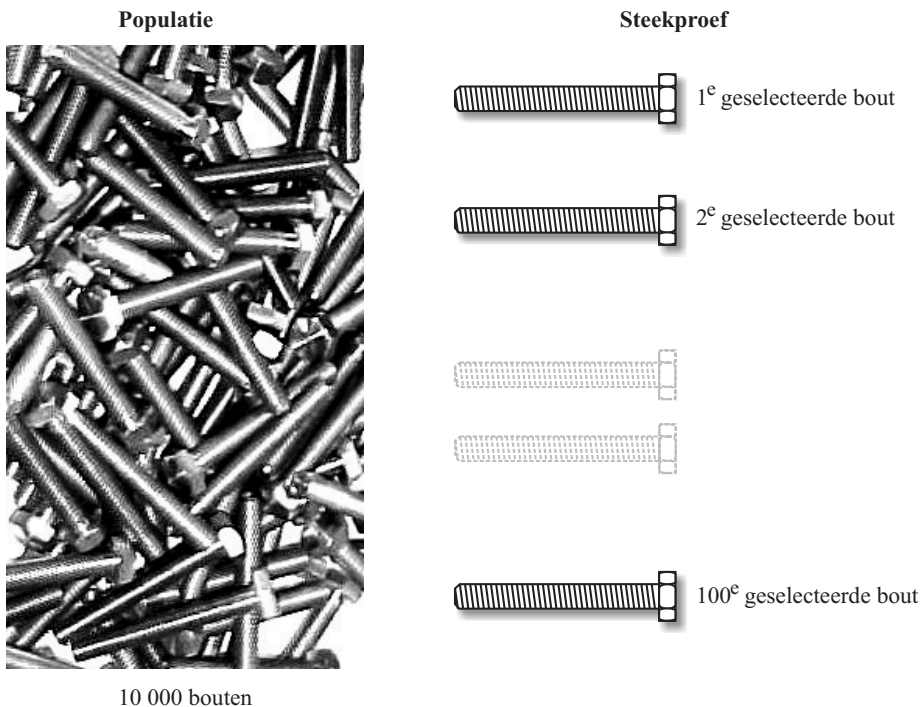
Als de populatie die we willen bestuderen klein is, is het mogelijk om een variabele te meten voor elke eenheid in de populatie. Als we bijvoorbeeld het beginsalaris willen meten van iedereen die het afgelopen jaar zijn ingenieursdiploma heeft gehaald op de TU Delft, is het in ieder geval doenlijk om elk salaris te bepalen. Als we een variabele meten voor

elke eenheid van een populatie, is het resultaat een **telling** van de populatie. In de meeste gevallen zullen de populaties waarin we zijn geïnteresseerd echter veel groter zijn, met wellicht vele duizenden of zelfs een oneindig aantal eenheden. Voorbeelden van grote populaties zijn o.a. de populaties die genoemd zijn na definitie 1.4, en verder: alle facturen die in het afgelopen jaar zijn opgesteld door een *AEX-genoteerde* onderneming en alle personen die een vliegticket geboekt hebben via KLM. Het zou veel te veel tijd en/of geld kosten om een telling te houden voor zulke populaties. Een mogelijk geschikt alternatief is het selecteren en bestuderen van een *deelverzameling* van de eenheden in die populatie.

DEFINITIE 1.6

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

Veronderstel bijvoorbeeld dat een onderneming de breeksterkte van een door hen gefabriceerd type bout wil controleren. Deze bouten moeten bij een bepaalde uitgeoefende kracht afbreken ter beveiliging van een bepaalde constructie. Er zijn twee redenen om niet alle gefabriceerde bouten te testen. Ten eerste kunnen we de bouten alleen maar testen met een vernietigend onderzoek. We kunnen de bout namelijk alleen maar testen door te onderzoeken bij welke sterkte deze breekt. Indien de bout is gebroken, is deze verder onbruikbaar. Ten tweede is het aantal gefabriceerde bouten (ongeveer 10.000 per dag) te groot om te testen. De onderneming test daarom slechts 100 bouten uit de gefabriceerde 10.000 van die dag (zie figuur 1.2). De fabrikant kan nu van deze bouten de breeksterkte en de gemiddelde waarde bepalen.



FIGUUR 1.2: Een steekproef van 10.000 gefabriceerde bouten

Nadat de relevante variabele(n) voor elke eenheid in de steekproef (of populatie) is/zijn gemeten, worden de gegevens geanalyseerd, hetzij met beschrijvende, hetzij met verklarende statistische methoden. De onderneming zou bijvoorbeeld alleen maar geïnteresseerd kunnen zijn in het bepalen van de breeksterkte van de 100 bouten in de steekproef. Het is echter waarschijnlijker dat het bedrijf de informatie van de steekproef zal willen gebruiken om gevolgtrekkingen te maken over de populatie van 10.000 bouten.

DEFINITIE 1.7

Een **statistische gevolgtrekking** is een schatting, een voorspelling of een andere generalisatie met betrekking tot een populatie die gebaseerd is op informatie uit een steekproef.

Dat wil zeggen dat we de informatie uit de steekproef gebruiken om iets te weten te komen over de hele populatie. Zo kan de onderneming uit de steekproef van 100 bouten een schatting maken van de breeksterkte in de populatie van 10.000 bouten. De gevolgtrekking van de onderneming met betrekking tot de breeksterkte van de bouten kan worden gebruikt om te beslissen of het fabricagesysteem nog in orde is.

De voorafgaande definities en voorbeelden laten vier van de vijf elementen van een vraagstuk in de verklarende statistiek zien: een populatie, een of meer relevante variabelen, een steekproef en een gevolgtrekking. Maar met een gevolgtrekking alleen zijn we er nog niet. We moeten ook de **betrouwbaarheid** weten – met andere woorden: we moeten weten hoe goed de gevolgtrekking is. De enige manier waarop we er zeker van kunnen zijn dat een gevolgtrekking voor een populatie correct is, is door de gehele populatie in onze steekproef te betrekken. We kunnen echter vanwege *bependingen in de middelen* (dat wil zeggen onvoldoende tijd en/of geld) meestal niet met hele populaties werken, zodat we onze gevolgtrekkingen moeten baseren op slechts een deel van de populatie (een steekproef). Het is daarom belangrijk om, wanneer dit ook maar mogelijk is, de betrouwbaarheid van elke gevolgtrekking die wordt gemaakt, te bepalen en te vermelden. Betrouwbaarheid is dus het vijfde element van vraagstukken in de verklarende statistiek.

Het is de betrouwbaarheidsmaat die aan een gevolgtrekking wordt toegevoegd, die zorgt voor een waterscheiding tussen de wetenschap statistiek enerzijds en de kunst van het voorspellen van de toekomst anderzijds. Een handler kan, net als een statisticus, een steekproef onderzoeken (je hand) en daaruit gevolgtrekkingen maken over de populatie (je levensloop). In tegenstelling tot statistische gevolgtrekkingen, bestaat er voor de gevolgtrekkingen van de handler echter geen betrouwbaarheidsmaat.

Veronderstel dat we het gemiddelde gewicht van een populatie verblikken willen schatten uit het gemiddelde gewicht van een steekproef van blikken. Door gebruik te maken van statistische methoden kunnen we de *schattingsfout begrenzen*. Deze begrenzing is niets anders dan een getalwaarde die door onze schattingsfout (het verschil tussen het gemiddelde gewicht van de steekproef en het gemiddelde gewicht van de populatie verblikken) waarschijnlijk niet wordt overschreden. We zullen in latere hoofdstukken zien dat deze grens een maat is voor de onzekerheid van onze gevolgtrekking. De betrouwbaarheid van statistische gevolgtrekkingen zal verder in dit boek nog geregeld ter sprake komen. Op dit moment willen we alleen maar dat je beseft dat een gevolgtrekking niet compleet is zonder een betrouwbaarheidsmaat.

DEFINITIE 1.8

Een **betrouwbaarheidsmaat** is een uitspraak (meestal kwantitatief) over de mate van onzekerheid die gepaard gaat met een statistische gevolgtrekking.

We besluiten dit onderdeel met een samenvatting van de elementen van zowel beschrijvende als verklarende statistische problemen, en een voorbeeld ter illustratie van de betrouwbaarheidsmaat.

Vier elementen van beschrijvende statistiek

1. de relevante populatie;
2. één of meer variabelen (kenmerken van de populatie- of steekprofeenheden) die onderzocht moeten worden;
3. tabellen, grafieken, of numerieke hulpmiddelen om een samenvatting te geven;
4. conclusies met betrekking tot de gegevens, gebaseerd op de patronen die naar voren zijn gekomen.

Vijf elementen van verklarende statistiek

1. de relevante populatie;
2. één of meer variabelen (kenmerken van de populatie- of steekprofeenheden) die onderzocht moeten worden;
3. de steekproef van populatie-eenheden;
4. de gevolgtrekking over de populatie, gebaseerd op informatie in de steekproef;
5. een betrouwbaarheidsmaat voor de gevolgtrekking.

1.4 Soorten gegevens

Je hebt nu geleerd dat statistiek de wetenschap van gegevens is, en dat gegevens worden verkregen door de waarden van een of meer variabelen te meten van de eenheden in de steekproef (of populatie). Alle gegevens (en daarmee de variabelen die we meten) kunnen in twee algemene categorieën worden ingedeeld: *kwantitatieve* gegevens en *kwalitatieve* gegevens.

Kwantitatieve gegevens zijn gegevens die voorkomen op een van nature voorkomende numerieke schaal. Voorbeelden van kwantitatieve gegevens zijn:

1. De temperatuur (in graden Celsius) waarop elke eenheid in een steekproef van 20 stukjes hittebestendig plastic begint te smelten.
2. Het huidige werkloosheidspercentage voor elk van de 12 provincies in Nederland.
3. De scores van een steekproef van 50 studenten voor het vak statistiek aan de TU Delft.
4. Het aantal vrouwelijke leidinggevendenden in elke onderneming in een steekproef van 75 productiebedrijven.

DEFINITIE 1.9

Kwantitatieve gegevens zijn meetwaarden die worden geregistreerd op een van nature voorkomende numerieke schaal.

In contrast hiermee kunnen kwalitatieve gegevens niet op een numerieke schaal worden gemeten; ze kunnen alleen in categorieën worden ingedeeld. Voorbeelden van kwalitatieve gegevens zijn:

1. de politieke-partijconnectie (Democratisch, Republikeins, of Onafhankelijk) in een steekproef van 50 CEO's;
2. de kwaliteitsstatus (al of niet defect) van elke computerchip uit een serie van 100, door Intel gefabriceerd;

3. de afmetingen van een auto (klein, compact, middelgroot of groot) die wordt gehuurd door elke zakenreiziger uit een steekproef van 30 zakenreizigers;
4. de rangorde die een smaaktester (beste, slechtste enzovoort) toekent aan vier merken barbecuesaus, in een panel van 10 testers.

Vaak kennen we willekeurige numerieke waarden toe aan kwalitatieve gegevens om de invoer in de computer en de analyse makkelijker te maken. Maar deze toegekende waarden zijn slechts codes: optellen, aftrekken, vermenigvuldigen of delen van deze codes heeft geen betekenis. Zo zouden we bijvoorbeeld de volgende codering kunnen gebruiken: Democraat = 1, Republikein = 2 en Onafhankelijk = 3. Evenzo kan een smaaktester de barbecuesausen een code toekennen van 1 (beste) tot 4 (slechtste). Dit zijn slechts willekeurig gekozen numerieke codes voor de kenmerken, die buiten de test geen toepassing hebben.

DEFINITIE 1.10

Kwalitatieve gegevens zijn meetwaarden die niet op een natuurlijk voorkomende numerieke schaal kunnen worden gemeten; ze kunnen alleen worden geclassificeerd in één categorie uit een groep categorieën.

Zoals te verwachten valt, hangen de statistische methoden voor het beschrijven, vastleggen en analyseren van gegevens af van het soort gegevens (kwantitatief of kwalitatief) dat wordt gemeten. In de volgende hoofdstukken zullen we vele bruikbare methoden beschrijven. Maar eerst bespreken we een aantal belangrijke ideeën met betrekking tot het verzamelen van gegevens.

1.5 Gegevens verzamelen

Als je eenmaal het soort gegevens hebt gekozen – kwantitatieve of kwalitatieve – dat het meest geschikt is voor het onderhavige probleem, zul je de gegevens moeten gaan verzamelen. In het algemeen zijn er vier manieren om gegevens te verkrijgen:

1. Gegevens uit een *gepubliceerde bron*
2. Gegevens uit een *experiment*
3. Gegevens uit een *enquête*
4. Gegevens verkregen uit *observatie*

Soms zijn de relevante gegevens al verzameld en beschikbaar in een **gepubliceerde bron**, zoals een boek, tijdschrift of een krant. Stel dat je de werkloosheidspercentages in de 12 provincies van Nederland wilt bestuderen en analyseren. Deze gegevensverzameling (naast vele andere) is te vinden in de publicaties van het Centraal Bureau voor de Statistiek (CBS) die jaarlijks door het ministerie van Sociale Zaken aan het CBS worden verstrekt. Evenzo kan iemand die geïnteresseerd is in maandelijks hypotheekaanvragen voor nieuwe huizen, deze gegevens vinden in een andere publicatie van het Kadaster.¹

¹ Bij gepubliceerde gegevens maken we vaak een onderscheid tussen de *primaire bron* en de *secundaire bron*. Als degene die publiceert zelf de gegevens heeft verzameld, is er sprake van een primaire bron, anders van een secundaire bron.

STATISTIEK IN DE PRAKTIJK

1.1 Kwaliteitsverbetering: Bedrijven in de VS reageren op de Japanse uitdaging

Gedurende de laatste drie decennia zijn bedrijven in de VS sterk onder druk komen te staan door de invoer van producten van superieure kwaliteit uit het buitenland. Zo steeg bijvoorbeeld van 1984 tot 1991 het marktaandeel van geïmporteerde auto's en lichte vrachtwagens van 22% naar 30%. Een ander voorbeeld is de markt voor televisie en video. Beide producten werden in de VS uitgevonden, maar sinds 1995 is er geen enkele Amerikaanse firma meer die ze maakt. Ze worden uitsluitend gemaakt door landen in het Verre Oosten, voornamelijk Japan.

Om een antwoord te vinden op deze uitdaging van de concurrentie, zijn bedrijven in de VS – zowel productiebedrijven als dienstverlenende bedrijven – begonnen met het nemen van initiatieven voor kwaliteitsverbetering. Veel van deze bedrijven leggen nu de nadruk op kwaliteitsmanagement in alle fasen en aspecten van hun bedrijf, van het productontwerp tot de productie, distributie, verkoop en service.

In het algemeen houden programma's voor kwaliteitsverbetering zich bezig met (1) achterhalen wat de klant wil, (2) deze behoeften te vertalen naar een productontwerp, en (3) het produceren en leveren van producten of diensten die voldoen aan, of beter zijn dan de ontwerpspecificaties. Bij al deze aspecten, maar in het bijzonder bij het derde aspect, is voor kwaliteitsverbetering procesverbetering nodig – voor zowel productieprocessen, distributieprocessen, serviceprocessen als supportprocessen.

Maar wat betekent het nu als je zegt dat een proces is verbeterd? In het algemeen betekent dit dat de klant van het proces (dat wil zeggen de gebruiker van de output) een grotere tevredenheid met de output aangeeft. Vaak zijn voor deze toename in tevredenheid verminderingen in de variatie van één of meer

procesvariabelen nodig. Dat wil zeggen dat een vermindering in de variatie van de outputstroom van het proces nodig is.

Maar hoe kan procesvariatie geregistreerd en verminderd worden? In de jaren twintig zorgde Walter Shewhart van Bell Telephone Laboratories voor misschien wel de belangrijkste doorbraak in de eeuw met betrekking tot de verbetering van processen. Hij realiseerde zich dat variatie in procesoutput onvermijdelijk is. Geen twee onderdelen die door een bepaalde machine worden geproduceerd zijn hetzelfde; geen twee transacties door een bankemployé zijn hetzelfde. Hij zag ook in dat variatie kan worden begrepen, geregistreerd en gecontroleerd met behulp van statistische methoden. Hij ontwikkelde een eenvoudige grafische techniek – de *controlekaart* – om te bepalen of productvariatie binnen aanvaardbare grenzen ligt. Deze methode geeft een richtlijn voor wanneer een productieproces moet worden bijgesteld of veranderd, en wanneer het ongemoeid moet worden gelaten. De methode kan aan het eind van het productieproces worden toegepast, of nog belangrijker, op verschillende plaatsen binnen het proces. We behandelen controlekaarten en andere hulpmiddelen voor het verbeteren van processen in hoofdstuk 10.

Gedurende het laatste decennium zijn, voornamelijk als gevolg van de Japanse uitdaging aan de kwaliteit van de Amerikaanse producten, controlekaarten en andere statistische hulpmiddelen algemeen in gebruik geraakt in de Verenigde Staten. Er zijn aanwijzingen dat de VS de uitdaging van de Japanse concurrentie kunnen aangaan: de meest prestigieuze kwaliteitsverbeteringsprijs ter wereld was ooit de Japanse Demingprijs. Tegenwoordig is het de Amerikaanse Malcolm Baldrige National Quality Award. Een andere indicatie dat de concurrentiepositie van de VS weer verbeterd is, is de verandering in het marktaandeel van de autofabrikanten in de VS: het aandeel van geïmporteerde auto's daalde van 30% in 1991 tot 26% in 1999.