

Lof voor dit boek

“Dit speelse boekje over manipuleren met cijfers is compact – je zou het op een regenachtige zondagmiddag of in een paar treinritten kunnen uitlezen. Maar dat zegt niets over de kwaliteit – die is fenomenaal. In het ene voorbeeld na het andere onthult Huff alle trucs die ingezet worden om lezers met cijfers, grafieken en illustraties te overbluffen en in de val te lokken.”

– WALL STREET JOURNAL

“Op de redactie raadplegen we regelmatig dit fantastische boekje van Darrell Huff. Het is onmogelijk om een vorm van statistische misleiding te bedenken die hij niet heeft blootgelegd.”

– NEW YORK TIMES

“De auteur en illustrator bundelen hun enorme talent in dit vlot geschreven, fraai geïllustreerde boekje dat niet alleen erg informatief is, maar ook leuk om te lezen. Het neemt alle wind uit de opgeblazen statistische zeilen.”

– LIBRARY JOURNAL

“Precies op tijd – dit boek moest geschreven worden.
Het is overtuigend en leest als een trein.”

– MANAGEMENT REVIEW

“Dit verrassende boekje zal gegarandeerd
je vertrouwen in cijfers ondermijnen.”

– ATLANTIC

“Hoewel dit boek voor het eerst verscheen in 1954, is het nu relevanter dan ooit. Een fantastische inleiding in de toepassing van statistiek.”

– BILL GATES

Inhoud

Inleiding 7

1. De scheve steekproef 11

2. Het goedgekozen gemiddelde 27

3. De kleine getalletjes die ontbreken 37

4. Veel gedoe om bijna niks 53

5. De verbluffende grafiek 59

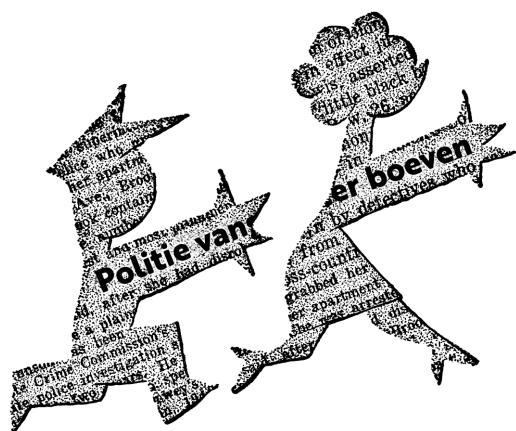
6. Het eendimensionale plaatje 65

7. Het semi-verwante cijfer 73

8. Oorzaak en gevolg 85

9. Meester in de statistipulatie 97

10. Cijfers op hun nummer zetten 117



Inleiding

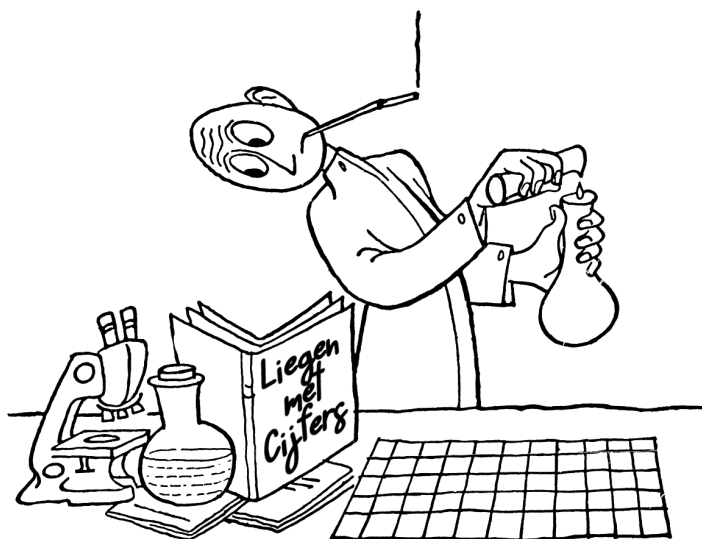
“De misdaad tiert hier welig,” zei mijn schoonvader, kort nadat hij was verhuisd van Iowa naar Californië. Dat wist hij dankzij de lokale krant die hij iedere dag las. Die krant stond erom bekend geen enkele misdaad in de regio over te slaan. De conclusie van mijn schoonvader getuigde van huis-tuin-en-keuken-statistiek. Zijn conclusie was gebaseerd op een steekproef, maar wel eentje die een vertekend beeld gaf. Hij baseert zich namelijk op een semi-verband en veronderstelt dat de ruimte en aandacht die een krant schenkt aan misdaadgevallen een maatstaf is voor het misdaadgehalte.

Een hoop winters geleden rapporteerden verschillende onderzoekers, onafhankelijk van elkaar, over de heilzame werking van antihistaminepillen. Die onderzoeken toonden aan dat in een substantieel aantal gevallen verkoudheid na de behandeling verdween. Dat zorgde voor veel bombarie, vooral in advertenties, en de verkoudheidsmedicijnen overspoelden de markt. Maar de hele kwestie was eigenlijk vooral gebaseerd op hoop en de merkwaardige weigering om achter de statistieken te kijken

naar een feitje dat al heel lang bekend was. Henry G. Felsen, een komiek, maar geen medische autoriteit, wist veel eerder al te vertellen dat bij een adequate behandeling de verkoudheid na zeven dagen is verdwenen, maar als je verkoudheid helemaal niet behandelt, duurt genezing ongeveer een week.

En zo gaat het met veel dingen die je leest of hoort. Gemiddelden, causale verbanden en trends zijn niet altijd wat ze lijken. Als je verder kijkt dan je neus lang is, dan zie je veel meer ... of juist veel minder.

De geheime taal van cijfers en statistieken, die in onze feitenminnende cultuur grote aantrekkingskracht hebben, wordt gebruikt om dingen op te blazen, verwarring te stichten en zaken mooier voor te spiegelen. Statistische technieken en termen zijn noodzakelijk om te kunnen berichten over sociale en economische trends, bedrijfsresultaten en opiniepeilingen, maar zonder schrijvers die die termen eerlijk en zorgvuldig gebruiken en lezers die precies weten wat ze betekenen, leiden ze al snel tot semantische onzin.



Bij het aanhalen van wetenschappelijk onderzoek in de media hebben misleidende statistische cijfers het beeld verdrongen van de held in de witte jas die tot 's avonds laat in een slecht verlicht laboratorium staat te zwoegen. Als een 'dun laagje bladgoud' kan een statistisch cijfer een belangrijk inzicht vermommen tot iets wat het niet is. Een sluw verpakt statistisch cijfer is beter dan een leugen; het is misleidend, maar ze kunnen je niet op onwaarheden betrappen.

Dit boek is een soort inleiding in het bedriegen met statistiek. Daarmee lijkt het misschien op een handleiding voor oplichters, maar ik kan mijn bedoelingen wellicht het beste vergelijken met de gepensioneerde inbreker wiens memoires de basis vormen voor een cursus slotenkraken. Niet voor dieven – die kennen al deze trucs al – maar voor de eerlijke burger, die met deze kennis inbrekers kan weren.



Hoofdstuk 1

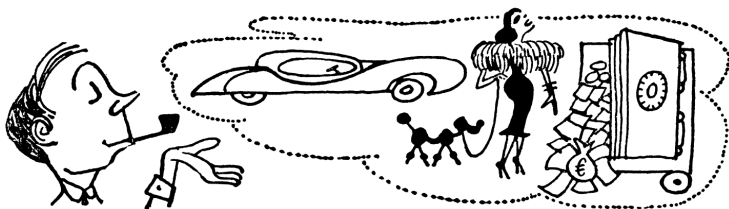
De scheve steekproef

“De gemiddelde Yale-student uit de lichter van 1924, verdient 25.111 dollar per jaar,” stond ooit in *Time Magazine*, dat daarmee reageerde op een artikel uit de *New York Sun*. Voor die tijd een aanzienlijk bedrag.

Fijn voor die student natuurlijk! Maar wacht even... Wat betekent dat indrukwekkende gemiddelde eigenlijk? In eerste instantie lijkt hiermee bewezen, dat als je je zoon of dochter naar Yale stuurt, jouw pensioen is zekergestellt en hij of zij ook vroeg kan gaan rentenieren, toch? Twee dingen aan dit cijfer zullen iemand met een kritische blik meteen opvallen: het is verrassend nauwkeurig en het is onwaarschijnlijk hoog.

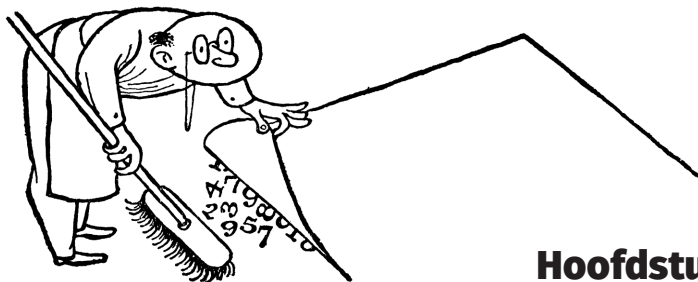
De kans dat het gemiddelde inkomen van zo'n wijdverspreide groep tot op de dollar nauwkeurig bekend is, is heel klein. Het is niet erg waarschijnlijk dat iemand precies weet hoe hoog zijn of haar eigen inkomen afgelopen jaar was, tenzij het inkomen alleen bestond uit salaris. Maar hele hoge inkomens bestaan zelden alleen uit salaris; mensen in deze categorie hebben vaak ook inkomsten uit bezit of beleggingen.

Bovendien is dit fraaie gemiddelde vrijwel zeker berekend aan de hand van de inkomens die de voormalige Yale-studenten zélf hebben opgegeven. Zelfs als in 1924 op de Universiteit van Yale een erencode gold, dan weten we nog niet zeker dat iedereen ook 25 jaar later, ten tijde van het onderzoek, nog steeds eerlijk antwoord heeft gegeven. Sommige mensen overdrijven hun inkomen, uit schaamte, ijdelheid of optimisme. Anderen geven juist een lager inkomen op, zelfs in de belastingaangifte vrees ik, en houden dit hardnekkig vol, ook als een journalist ernaar vraagt. De belastinginspecteur leest immers ook de krant! Het is mogelijk dat deze twee effecten, aandikken en afzwakken, elkaar opheffen, maar dat is niet waarschijnlijk. Het ene effect is mogelijk veel sterker dan het andere, alleen weten we niet welke.



We beginnen dus te twijfelen aan een cijfer waarvan ons gezond verstand ons al zegt dat het eigenlijk niet kan kloppen. Laten we eens proberen de vinger op de zere plek te leggen en de mogelijke oorzaak aan te wijzen. De denkfout die 25.111 dollar oplevert als het ‘gemiddelde inkomen’ van een groep studenten, terwijl het daadwerkelijke gemiddelde best eens de helft van dat bedrag zou kunnen zijn.

Maak kennis met de steekproef, die in veel statistische berekeningen wordt gebruikt. De basisbeginselen van een steekproef zijn eenvoudig, hoewel in de praktijk subtiele aanpassingen heb-



Hoofdstuk 3

De kleine getalletjes die ontbreken

“Uit onderzoek blijkt dat mensen die poetsen met tandpasta van Doakes 23% minder gaatjes hebben” – het staat in grote letters geschreven. 23% minder vaak naar de tandarts klinkt aanlokkelijk, dus lees je verder. Het onderzoek is, zo blijkt, verricht door een geruststellend ‘onafhankelijk’ laboratorium en de cijfers zijn gecontroleerd door een betrouwbare accountant. Wat wil je nog meer?

Maar als je niet goedgelovig bent, dan weet je uit ervaring dat de ene tandpasta niet veel beter is dan de andere. Hoe kan het dan dat Doakes zulke gunstige resultaten kan presenteren? Liegen ze alsof het gedrukt staat, in koeieletters nog wel? Nee. En ze hoeven ook niet te liegen, want er zijn makkelijkere en effectievere manieren om aan overtuigende cijfers te komen.

De belangrijkste troef in dit verhaal is de ondeugdelijke steekproef – dat wil zeggen, statistisch ondeugdelijk. Maar precies goed voor het doel dat Doakes voor ogen had. Als je de kleine lettertjes had gelezen, had je kunnen weten dat de testgroep slechts bestond uit twaalf personen. Eigenlijk nog best sportief

van Doakes dat ze dat in de kleine lettertjes vermeldden, want de meeste adverteerders laten die informatie achterwege zodat zelfs statistici slechts kunnen raden welk foefje was toegepast. Een steekproef van een dozijn is nog niet eens zo heel slecht – het kan nog erger.

Enige tijd geleden kwam er een product op de markt met de naam Tandpoeder van Dr. Cornish – de fabrikant claimde dat was aangetoond dat het poeder “erg goed hielp tegen gaatjes”. Het poeder bevatte ureum en laboratoriumonderzoek zou hebben uitgewezen dat ureum tandcariës kon bestrijden. Maar die resultaten waren eigenlijk nutteloos, want het ging om een vooronderzoek waarbij het poeder getest was door precies ... zes proefpersonen.

Maar laten we nog eens kijken naar het gemak waarmee Doakes zo’n mooie claim kon maken zonder iets onwaars te verkondigen en die resultaten zelfs gecertificeerd kon krijgen. Laat een willekeurig groepje mensen zes maanden lang bijhouden hoeveel gaatjes ze hebben en vraag ze dan over te stappen op de tandpasta van Doakes. Er zijn dan drie mogelijke uitkomsten: de testgroep krijgt daarna beduidend meer gaatjes, beduidend minder gaatjes of ongeveer hetzelfde aantal gaatjes. Als de eerste of laatste uitkomst zich voordoet, dan verbergt Doakes de resultaten ergens onderin een bureaula (waar niemand ze ziet) en begint opnieuw. Vroeg of laat, kwestie van toeval, rapporteert de testgroep een sterke verbetering van het aantal gaatjes, goed genoeg voor een mooie advertentiecampagne. Het maakt eigenlijk niet uit of ze met Doakes, hun oude vertrouwde tandpasta of zuiveringszout hadden gepoetst. Het belang van een kleine testgroep is evident: bij een grote testgroep is het verschil dat veroorzaakt wordt door toeval doorgaans erg klein en dus niet advertentiewaardig – met een verbetering van 2 procent ga je niet meer tandpasta verkopen.

Hoe je louter op basis van toeval met een klein aantal gevallen een nikszeggende uitkomst kunt krijgen, kun je ook zelf gemakkelijk testen. Gooi gewoon een muntje op. Hoe vaak gooi je kop? In 50% van de gevallen natuurlijk, dat weet iedereen. Toch? Laten we dat gewoon eens checken...

Ik heb tien keer een munt opgegooid en het was acht keer kop, dus dat bewijst dat het opgooien van een munt in 80 procent van de gevallen kop oplevert. Ahum.

ECHT GETEST (één test)



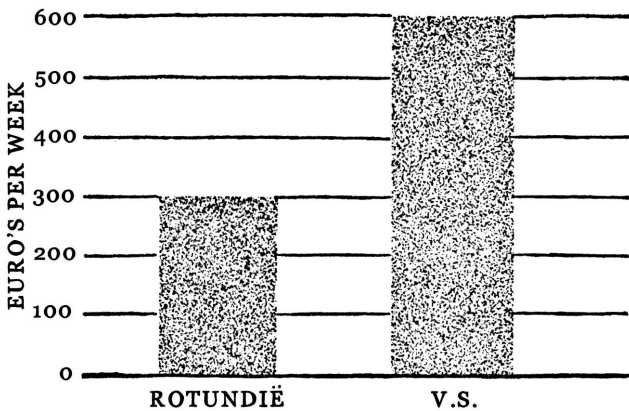
Wetenschappelijk bewijs dat het opgooien van een munt in 80 procent van de gevallen kop oplevert.



Althans, in de wereld van de tandpastastatistiek. Probeer het zelf maar eens. Heel misschien is het resultaat 50-50, maar de kans is groot van niet. Net als in mijn geval, is het heel aannemelijk dat je niet eens in de buurt van de 50-50 uitkomt. Alleen als je geduld genoeg hebt om de munt duizend keer op te gooien, dan kom je in de buurt (maar niet helemaal) van de echte waarschijnlijkheid. Dus alleen wanneer het aantal pogingen groot genoeg is, is de wet van het gemiddelde een bruikbare beschrijver of voorspeller.

vergelijken. Een afgekapte staafgrafiek is net zo misleidend als de afgekapte lijngrafiek waarmee je eerder hebt kennism gemaakt. Je vindt de staafgrafiek vooral in studieboeken, bedrijfsdocumentatie en opiniebladen. Dat geldt ook voor diens opgeleukte nakomelingen, de beeldgrafieken.

Stel, ik wil het gemiddelde weekloon van een timmerman in de VS en Rotundië met elkaar vergelijken. In de VS is dat weekloon € 600 en in Rotundië € 300. Het moet in het oog springen, dus alleen die cijfers noemen is niet voldoende. Ik maak daarom een staafgrafiek. (Overigens, als die € 600 per week niet overeenkomt met het enorme bedrag dat jij moest betalen voor je nieuwe schutting, bedenk dan dat jouw timmerman mogelijk niet iedere week zo hard heeft gewerkt als aan jouw schutting. Bovendien heb ik niet verteld wélk gemiddelde wordt bedoeld en hoe dat is berekend, dus het heeft geen zin te gaan muggeziften. Zie je hoe gemakkelijk het is om je te verschuilen achter een statistische façade als je verdere informatie achterwege laat? Je vermoedde vast al dat ik dit voorbeeld heb verzonnen, maar je was geheid minder achterdochtig geweest als ik in plaats van € 600 het weekloon € 598,30 had gebruikt.



Dat is 'm dan – de euro's per week staan aangegeven langs de y-as. Dit is een duidelijk en eerlijk plaatje – twee keer zoveel geld is twee keer zo hoog in de grafiek en daar lijkt het ook op.

De grafiek is alleen niet heel erg aantrekkelijk, toch? Dat is snel te verbeteren door een plaatje te gebruiken dat meer lijkt op geld dan een staaf: geldzakken. Eén geldzak voor de arme sloebers uit Rotundië en twee voor het loon in de VS. Of drie voor de Rotundiërs en zes voor de Amerikanen. Hoe dan ook, het blijft ook met geldzakken een eerlijke grafiek. Geen vuiltje bedrog aan de lucht.



Ik zou het hierbij kunnen laten als ik alleen informatie wilde overbrengen. Maar ik wil meer. Ik wil duidelijk maken dat de Amerikaanse timmerman véél beter af is dan zijn Rotundiaanse collega. Dus hoe groter ik het verschil tussen 300 en 600 kan laten lijken, hoe beter het mijn argumenten kracht bijzet.

Om heel eerlijk te zijn (wat ik natuurlijk niet nastreef) wil ik het verschil flink overdrijven, zonder betrappt te kunnen worden op valsheid in geschrifte. Dat kan – op een manier waarmee je bijna elke dag voor de gek gehouden wordt.

Eerst teken ik een geldzak die staat voor de driehonderd euro van de Rotundiërs en vervolgens teken ik een geldzak die twee keer zo hoog is voor de zeshonderd euro van de Amerikanen. Dat is in verhouding, toch?



Kijk, dat is de suggestie die ik wil wekken. Het weekloon van de Rotundiërs valt in het niet bij het weekloon van de Amerikanen. De truc is natuurlijk dat een geldzak die twee keer zo hoog is ook twee keer zo breed is. Daardoor neemt de geldzak niet twee maar vier keer zoveel ruimte in beslag in de grafiek.

De bedragen zeggen twee keer zo groot, maar de visuele indruk (en die is veel dominanter) zegt vier keer zo groot. Of nog erger... Omdat een geldzak een object is dat in werkelijkheid drie dimensies heeft is de tweede geldzak ook nog eens twee keer zo diep als de eerste. Van ruimtemeetkunde weet je wellicht dat als de afmetingen van een kubus verdubbelen, de inhoud verachtvoudigd. Als de kleine geldzak € 300 bevat, dan bevat de grote geldzak – met acht keer het volume – maar liefst € 2400 en geen € 600!

En dat is precies de illusie die mijn sluwe grafiekje wekt. Hoewel het Amerikaanse weekloon twee keer zo hoog is, blijf je achter met de indruk dat de Amerikaanse timmerman wel acht keer zo veel verdient.

En wat ik doe is niet strafbaar. Ik doe slechts wat zoveel mensen doen. Het blad *Newsweek* deed het met geldzakken, de brancheorganisatie American Iron and Steel Institute deed het



Hoofdstuk 7

Het semi-verwante cijfer

Als je iets wilt bewijzen dat je niet kunt bewijzen, toon dan iets anders aan en doe net alsof dat hetzelfde is. Na de confrontatie met statistiek duizelt het de meeste mensen zo, dat bijna niemand het verschil zal opmerken. Het semi-verwante cijfer is een hulpmiddel dat je geregeld van pas zal komen.

Je kunt niet bewijzen dat jouw wondermiddel verkoudheid geneest, maar je kunt wel de resultaten van een legitiem onderzoek publiceren (in kapitalen) waaruit blijkt dat 15 gram van het spul binnen 11 seconden 31.10^8 ziektekiemen heeft gedood in een reageerbuisje. En als je toch bezig bent, zorg ervoor dat het onderzoek is verricht door een gerenommeerd laboratorium.

Publiceer het hele rapport. Fotografeer een model in een witte doktersjas en plaats die foto naast de onderzoeksresultaten. Maar zeg niks over de slimmigheidjes die je in je verhaal hebt gestopt. Het is niet aan jou – toch? – om de lezer erop te wijzen dat een antiseptisch middel dat goed werkt in een reageerbuisje misschien niet werkt bij mensen, vooral niet nadat het sterk is verdund om te voorkomen dat mensen hun keel branden. Maak

Als we 1% zakken herinner je je misschien nog de zaak van de sapcentrifuge. In de vele advertenties kon je lezen dat het apparaat “26% meer sap perste” hetgeen was “bewezen door laboratoriumonderzoek” en “onderschreven door het Good Housekeeping Institute.” Dat klinkt goed. Als je een sapcentrifuge kunt kopen die 26% meer sap perst, waarom zou je dan nog een ander apparaat kiezen?

Zonder verder in te gaan op het feit dat laboratoriumtests (vooral ‘onafhankelijke’ laboratoriumtests) de meest vreemde dingen hebben aangetoond, wat betekent dat cijfer nou eigenlijk? Zesentwintig procent meer dan wat?

Met veel moeite kon uiteindelijk achterhaald worden dat deze sapcentrifuge 26% meer sap uitperste dan een ouderwetse handpers. Stel je voor dat je op basis van deze informatie je koopbeslissing neemt – misschien was het wel de slechtste sapcentrifuge op de markt. Het percentage van 26% was niet alleen verdacht nauwkeurig, het was ook compleet irrelevant.

Adverteerders zijn niet de enigen die je met cijfers voor de gek houden als je niet oplet. Een artikel over rijveiligheid in het krantenmagazine *This Week* vertelde, ongetwijfeld met de beste bedoelingen, wat er kon gebeuren als je “met 110 kilometer per uur over de snelweg raasde, zwalkend van links naar rechts”. Volgens het artikel was de kans dat je die dolle rit zou overleven ’s ochtends om zes uur vier keer zo groot als ’s avonds om zes uur. Het bewijs: “’s Avonds om zes uur gebeuren er vier keer zoveel dodelijke ongelukken als ’s ochtends om zes uur.” Dat klopt wel ongeveer, maar de conclusie over je overlevingskansen volgt daar niet uit.

Er komen ’s avonds meer mensen om in het verkeer om de doodeenvoudige reden dat er dan ook meer mensen op de weg zijn. Jij, als eenzame bestuurder, loopt misschien ’s avonds meer gevaar, maar de cijfers bewijzen het een noch het ander.