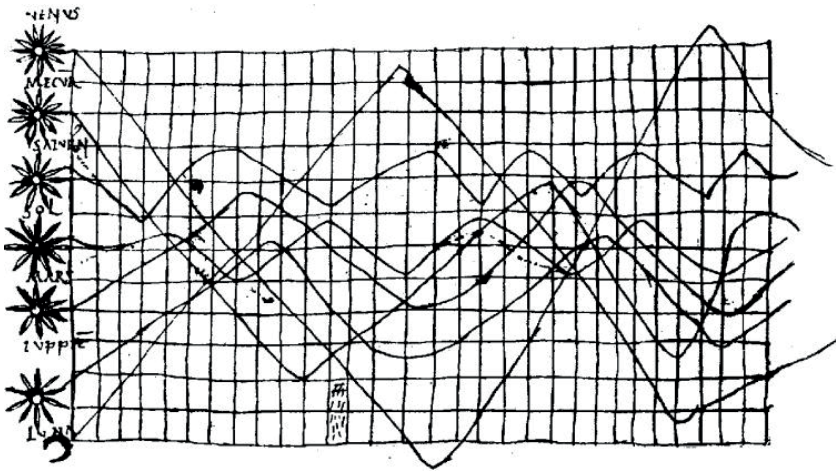


Het grafiekenboek

Het grafiekenboek

Jelke Bethlehem



Amsterdam University Press

Ontwerp omslag: Gijs Mathijs Ontwerpers, Amsterdam
Ontwerp binnenwerk: Crius Group, Hulshout

ISBN 978 94 6372 098 4
e-ISBN 978 90 4855 620 5 (e-pub)
DOI 10.5117/9789463720984
NUR 910

© Jelke Bethlehem / Amsterdam University Press B.V., Amsterdam 2022

Alle rechten voorbehouden. Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand, of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder voorafgaande schriftelijke toestemming van de uitgever.

Voor zover het maken van kopieën uit deze uitgave is toegestaan op grond van artikel 16B Auteurswet 1912 j^o het Besluit van 20 juni 1974, Stb. 351, zoals gewijzigd bij het Besluit van 23 augustus 1985, Stb. 471 en artikel 17 Auteurswet 1912, dient men de daarvoor wettelijk verschuldigde vergoedingen te voldoen aan de Stichting Reprorecht (Postbus 3051, 2130 KB Hoofddorp). Voor het overnemen van gedeelte(n) uit deze uitgave in bloemlezingen, readers en andere compilatiewerken (artikel 16 Auteurswet 1912) dient men zich tot de uitgever te wenden.

De uitgeverij heeft ernaar gestreefd alle copyrights van in deze uitgave opgenomen illustraties te achterhalen. Aan hen die desondanks menen alsnog rechten te kunnen doen gelden, wordt verzocht contact op te nemen met Amsterdam University Press.

Inhoud

1	Inleiding	7
2	Terug in de tijd	15
3	Ingrediënten van een grafiek	27
4	Richtlijnen voor grafieken	45
5	Grafieken voor de verdeling van een kwantitatieve variabele	83
6	Grafieken voor de verdeling van een kwalitatieve variabele	99
7	Grafieken voor de samenhang tussen kwantitatieve variabelen	129
8	Grafieken voor de samenhang tussen kwalitatieve variabelen	157
9	Grafieken voor de samenhang tussen gemengde variabelen	177
10	Grafieken voor tijdreeksen	187
11	Thematische kaarten	201
	Literatuur	215
	Register	219

1 Inleiding

1.1 Waarom een grafiek?

In 2012 deed het Centraal Bureau voor de Statistiek (CBS) onderzoek naar pesten op internet. Daarbij ging het om zaken als roddel, getreiter, stalken, chantage en bedreiging. Aan een steekproef van Nederlanders werd gevraagd of ze wel eens gepest werden en op welke manier ze werden gepest. Dit onderzoek leverde allerlei interessante uitkomsten op. De vraag is dan hoe je deze uitkomsten het beste kunt presenteren. Dat kan op drie verschillende manieren.

De eerste manier is het schrijven van een tekst over de uitkomsten. Voor het onderzoek naar pesten op internet zou die tekst er bijvoorbeeld zo uit kunnen zien:

In 2012 gaf 3 procent van de Nederlandse bevolking van 15 jaar of ouder aan weleens gepest te zijn via internet. Dit zijn naar schatting ruim 400 duizend personen. Jongeren worden veel meer met cyberpesten geconfronteerd dan ouderen. Van de 15- tot 18-jarigen had 10 procent in 2012 te maken met pesten via internet, tegen minder dan 1 procent van de 65-plussers. Natuurlijk maken jongeren ook meer gebruik van het internet en zijn ze actiever op sociale media, die vaak als platform voor pestgedrag dienen.

Deze tekst bevat de belangrijkste uitkomsten van het onderzoek. Veel details staan erin. Toch is het lastig om er een globaal beeld uit te halen.

Een tweede manier om de uitkomsten samen te vatten is in de vorm van een tabel. Een tabel is een numeriek overzicht waarin vooral cijfers staan en haast geen tekst. Tabel 1.1.1 bevat de uitkomsten van het onderzoek naar cyberpesten.

Tabel 1.1.1 Slachtoffers van cyberpesten, naar leeftijd, in 2012

Leeftijd	Percentage slachtoffers
15 tot 18 jaar	10,3%
18 tot 25 jaar	7,1%
25 tot 45 jaar	3,1%
45 tot 65 jaar	2,1%
65 jaar en ouder	0,9%

Deze kleine tabel is nog wel redelijk overzichtelijk. Zo is eenvoudig te zien in welke leeftijdscategorie cyberpesten het meest voorkomt (15 tot 18 jaar) en in welke groep het minst (65 jaar en ouder). Bij grotere en ingewikkelder tabellen is de interpretatie meestal een stuk lastiger.

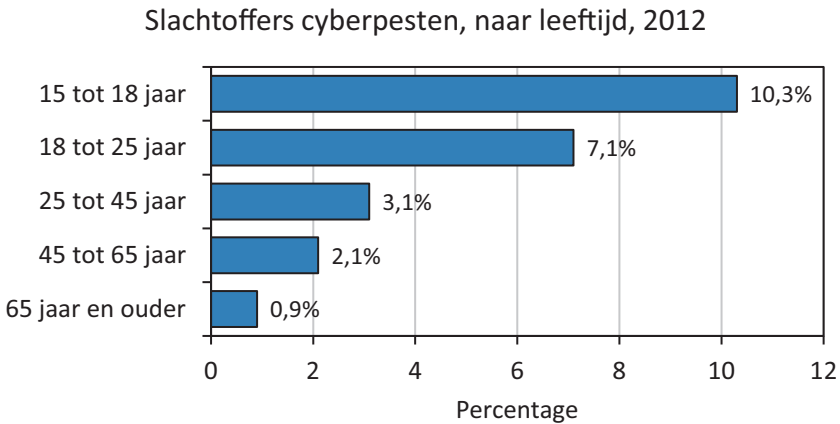
Een derde manier om de uitkomsten van een onderzoek te presenteren, is met een grafiek. In een grafiek is vaak eenvoudiger te zien wat er aan de hand is. Dat maakt een grafiek geschikt voor een groot en breed publiek. Maar pas op: er zijn ook gevaren. Je kunt gegevens op allerlei manieren grafisch weergeven, maar niet elke manier is goed. Het is vrij eenvoudig om mensen op het verkeerde been te zetten. Daardoor kunnen ze de verkeerde conclusies trekken. Het is daarom belangrijk dat je zorgvuldig bent bij het maken van een grafiek. En dat betekent ook dat je moet oppassen als je een grafiek van een onderzoeker of in de krant onder ogen krijgt.

In het Nederlands hebben we het gewoonlijk over *grafieken*, *diagrammen* of *plots*. In het Engels zijn er ook drie verschillende termen gangbaar: *graphs*, *charts* en *plots*. Het is niet duidelijk of het bij deze termen om verschillende zaken gaat of dat het verschillende namen voor hetzelfde zijn. Vaak zie je dat deze termen door elkaar worden gebruikt. Dat is ook in dit boek het geval. Zie Robbins (2013) voor een discussie over de Engelse terminologie.

Niet elk type grafiek is geschikt om de 'boodschap' in de gegevens op een goede manier over te brengen. De sterke en zwakke punten van veel grafieken zullen in dit boek aan de orde komen. Zo zullen we het hebben over cirkeldiagrammen, staafdiagrammen, histogrammen, puntenwolken, lijndiagrammen en thematische kaarten.

Voor de gegevens over cyberpesten komt het *staafdiagram* in aanmerking. Het resultaat staat in figuur 1.1.1. In de grafiek is een duidelijke trend te zien: cyberpesten neemt af naarmate je ouder wordt. Cyberpesten komt het vaakst voor bij jongeren van 15 tot 18 jaar. Cyberpesten komt het minst voor bij ouderen vanaf 65 jaar. Voor deze groep ligt het percentage onder de 1%. Dat is niet zo verwonderlijk, want ouderen zijn veel minder actief op sociale media.

De grafiek in figuur 1.1.1 bevat maar vijf getallen (de vijf percentages). Een grafiek kan vooral haar kracht tonen als zij veel gegevens in beeld samenvat. Een grafiek kan onthullend zijn. Een grafiek kan het inzicht in de gegevens vergroten door het spontaan tonen van allerlei aspecten die in tekst of tabellen verborgen zouden zijn gebleven. Een grafiek vormt dus een effectief instrument voor het overbrengen van de boodschap die in een grote hoeveelheid gegevens opgesloten zit. De voordelen van een grafiek zijn al door Schmid (1983) op een rij gezet:

Figuur 1.1.1 Een staafdiagram

- Een grafiek brengt de boodschap in de gegevens snel en direct over. Je bent daardoor in staat om de belangrijkste aspecten van een grote hoeveelheid gegevens in één oogopslag te zien.
- Een grafiek is een krachtig instrument om een veelomvattend en samenhangend beeld te geven van de gegevens.
- Een grafiek heeft meer overtuigingskracht dan een stuk tekst of een tabel. Een grafiek doet meer dan het alleen maar signaleren van een bepaald verschijnsel. Een grafiek laat ook zien dat dit verschijnsel zich daadwerkelijk voordoet.
- Een grafiek kan onthullend zijn. Een grafiek vergroot het inzicht in de gegevens door het spontaan zichtbaar maken van tot dan toe verborgen gebleven aspecten. Daarmee stimuleert een grafiek analytisch denken en verder onderzoek.
- Een grafiek trekt door haar aantrekkelijke uiterlijk (vergeleken met tekst en tabellen) eerder de aandacht van de lezer. Daardoor daagt een grafiek lezers uit zich te verdiepen in het onderwerp.

Grafieken kun je voor twee verschillende doeleinden gebruiken: *presentatie* en *exploratie*. Bij *presentatie* van grafieken gaat het om het overbrengen van een boodschap door een onderzoeker aan anderen. De over te brengen boodschap staat centraal. Alles moet erop gericht zijn die boodschap zo eenvoudig en correct mogelijk over te brengen, ook bij een niet-statistisch geschoolde lezer. Dit soort grafieken vinden we bijvoorbeeld terug op de tv (in het journaal), in de krant, in de sociale media (Facebook, Twitter, enz.) en in het Statistisch Zakboek van het

CBS. Mooie voorbeelden van grafieken voor publicatie zijn te vinden in het boek van Robbins (2013).

Een tweede gebruik van grafieken is *exploratie*. Daarbij maakt een onderzoeker de grafiek niet voor een ander, maar voor zichzelf. Het is een hulpmiddel bij het onderzoeken van statistische gegevens. De grafiek moet helpen om onverwachte patronen en structuren aan het licht te brengen. Is er een bepaald soort samenhang? Zijn er objecten (personen, bedrijven) die zich afwijkend gedragen? Veel voorbeelden van het gebruik van grafieken voor exploratie zijn te vinden in Tukey (1977). Dit soort grafieken zullen niet vaak in publicaties terechtkomen. Een onderzoeker zal er vele maken en ze misschien bijna allemaal weer weggooien.

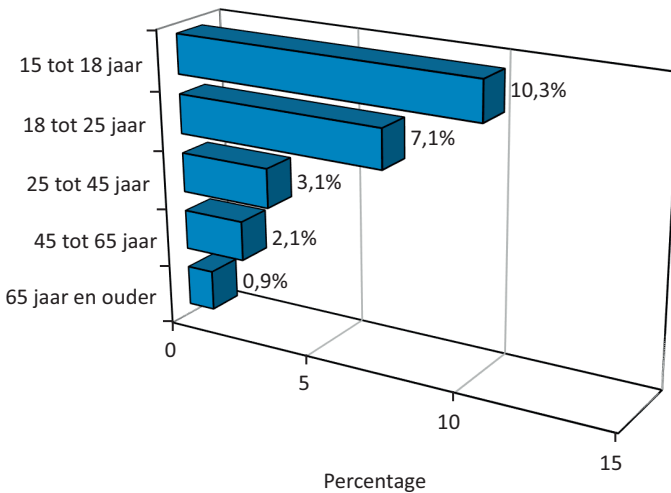
1.2 Het maken van grafieken

Een grafiek kan dus een krachtig hulpmiddel zijn om de (statistische) boodschap over te brengen die in de verzamelde gegevens verborgen zit. En dat geldt in het bijzonder voor lezers zonder kennis van statistiek. Grafieken kunnen zinvoller en aantrekkelijker zijn dan tabellen met getallen. Daarom is het niet verbazingwekkend dat je grafieken vaak terugziet in onderzoeksrapporten en in populaire media (zoals websites, kranten, tv en sociale media). Het gebruik van grafieken is echter niet zonder risico's. Slecht ontworpen grafieken kunnen een verkeerde boodschap overbrengen. Publicaties staan er helaas vol mee. Problemen worden vaak veroorzaakt door ontwerpers van grafieken die onvoldoende statistische kennis hebben. Ze hebben meer aandacht voor de aantrekkelijkheid van het grafisch ontwerp dan voor de statistische inhoud.

Niet elke grafiek is dus een goede grafiek. En het helpt niet dat er veel computerprogrammatuur beschikbaar is voor het maken van grafieken. Voorbeelden zijn statistische programma's zoals SPSS, SAS, Stata en R. Er zijn ook meer algemene programma's waarmee je goede grafieken kunt maken, maar helaas ook slechte grafieken. Soms leiden de standaard-instellingen van de programma's al tot verkeerde grafieken.

Figuur 1.2.1 toont een staafdiagram dat is gemaakt met *Microsoft Excel*. Er zijn maar een paar klikken nodig om dit staafdiagram met een driedimensionaal perspectief te maken. Veel deskundigen op het terrein van grafieken raden het gebruik van driedimensionaal perspectief af omdat het de interpretatie van een grafiek bemoeilijkt. In paragraaf 4.5 van hoofdstuk 4 (Richtlijnen voor grafieken) gaan we dieper in op het gebruik van dit soort driedimensionale grafieken.

Figuur 1.2.1 Een voorbeeld van een slechte grafiek die gemaakt is met Microsoft Excel



Bron: Centraal Bureau voor de Statistiek

‘Eén plaatje zegt meer dan duizend woorden.’ Deze spreuk wordt vaak aangehaald om aan te geven hoe waardevol grafieken kunnen zijn. Er wordt dan ook nog bij gezegd dat dit een oud Chinees spreekwoord is. Dat is echter niet het geval. De spreuk is de vertaling van ‘One picture is worth a thousand words.’ Die spreuk is in de jaren twintig van de vorige eeuw bedacht om aan te geven dat het gebruik van plaatjes in advertentieteksten belangrijk is. Ook al is de spreuk geen oud Chinees spreekwoord, toch illustreert hij het principe dat één plaatje vaak beter in staat is om een boodschap over te brengen dan een groot stuk tekst.

Bovengenoemde spreuk is ook van toepassing op grafieken. Maar niet zonder meer, want er zijn goede en slechte grafieken. Een grafiek doet alleen haar werk wanneer zij accuraat, eenvoudig te interpreteren, overtuigend, onthullend en aantrekkelijk is. Helaas is dit niet altijd het geval.

1.3 Een goede grafiek

Niet elke grafiek is een goede grafiek. Als je niet bent getraind in het doorzien van de grafische voetangels en klemmen, kun je een heel verkeerd beeld krijgen opgedrongen. Dat is vooral vervelend omdat grafieken vaak worden gebruikt voor het overbrengen van statistische informatie aan het grote

publiek. Onderzoekers, grafisch ontwerpers en computerprogrammeurs doen er daarom goed aan zich te bezinnen op wat nu eigenlijk goede grafieken zijn. Schmid (1983) geeft een aantal eisen waaraan een goede grafiek zou moeten voldoen.

Nauwkeurigheid en aanzien

Je moet een grafiek nauwkeurig maken. Het mag niet zo zijn dat door onnauwkeurigheid of onzorgvuldigheid van de maker (mens of computer) een misleidend beeld ontstaat. Het aanzien van een grafiek bepaalt of deze belangstelling zal opwekken. Een goede grafiek doet dit vanwege haar nette, waardige en professionele uiterlijk.

Eenvoud

Hoofddoel van een grafiek moet zijn het overbrengen van de statistische boodschap in de gegevens. Dat moet zo eenvoudig mogelijk gebeuren. Daarbij moet je niet ter zake doende tekst en versiering vermijden. Dat leidt de aandacht af van de boodschap. Tufte (1983) duidt al deze overbodige poespas aan met *chart junk*.

Helderheid

De grafiek moet een objectieve statistische boodschap overbrengen. De figuur moet duidelijk zijn en niet tot dubbelzinnige of verkeerde interpretatie leiden. Je moet zonder al te veel moeite de juiste boodschap eruit kunnen halen. Een grafiek is bedoeld om tijd en inspanning te besparen. Zij mag dus geen puzzel zijn.

Vormgeving

Er zijn veel verschillende typen grafieken. Het type grafiek dat je kiest, moet in overeenstemming zijn met de aard van de over te brengen boodschap. Daarbij mag je functioneel gebruikmaken van kleuren of grijstinten. Je moet optische illusies vermijden. Verschillen in waargenomen contrast moeten indicatief zijn voor verschillen in waarde. Je moet je bij het maken van de grafiek bewust zijn van de fysiologische en psychologische processen die zich in de hersenen afspelen bij het waarnemen en interpreteren van grafieken.

De hierboven genoemde eisen hangen met elkaar samen. Zo bepaalt de vormgeving het aanzien van de grafiek, en de nauwkeurigheid de helderheid.

Bij het maken van een grafiek moet je deze eisen daarom gezamenlijk in acht nemen.

We gaan niet in op de cognitieve aspecten van de perceptie van grafieken. Meer daarover is bijvoorbeeld te vinden in Few (2004, 2009). Ook Snijkers (1985) beschrijft de cognitieve aspecten van grafieken.

Er zijn heel wat boeken geschreven over het maken van grafieken en wat daarbij mis kan gaan. Voorbeelden zijn Schmid (1983), Tufte (1983), Wainer (1997), Wallgren et al. (1996) en Robbins (2013). Een nuttige internationaal georiënteerde publicatie over het maken van grafieken is de gids van de Verenigde Naties. Zie hiervoor United Nations (2009).