

Handig met getallen 5

Verbanden - 1.12 Misleidende grafieken

Rekenboek gecijferdheid voor de pabo



Zonder verband laat alles los (Erasmusbrug Rotterdam)

Auteur: Ruud Houweling

Adviezen: Arjen de Vetten

ISBN: 978 94 90681 289

Voldoet aan de Kennisbasis wiskunde voor de pabo



1.12 Misleidende grafieken

Er bestaat een veelzeggende uitdrukking over statistiek: je hebt leugens, grote leugens en statistiek. Blijkbaar is statistiek de overtreffende trap van liegen. Toch is het niet zo dat statistieken liegen, maar wel dat mensen statistieken bewust of onbewust misbruiken.

Sommige journalisten gebruiken statistische gegevens soms onbedoeld onjuist als ze zelf onvoldoende verstand van statistiek hebben. Ze trekken dan conclusies die ze niet kunnen trekken. Het komt ook voor dat instanties of bedrijven bewust een deel van beschikbare statistische gegevens gebruiken en zo een vertekend beeld geven. Dat doen ze met een bepaald doel.

In beide gevallen wordt gemanipuleerd met cijfers en wordt statistiek gebruikt om een gelijk aan te tonen of een gunstige conclusie te trekken. Dat kan bewust misleidend zijn, maar ook onbewust. Het gevaar bestaat dat conclusies en meningen gebaseerd zijn op onjuiste, of zelfs gemanipuleerde, informatie. Het doorzien van onjuist gebruik van statistische gegevens is lastig. Daarom eerst een voorbeeld.

Voorbeeld 1 Politieke voorkeur

Marketingbureau 'U vraagt, wij draaien' doet onderzoek naar de politieke voorkeur van de Nederlander van 18 jaar en ouder. Daarvoor ondervraagt het bureau een grote groep mensen in Wassenaar en een groep van vergelijkbare grootte in Oost-Groningen.

De belangrijkste vraag bij dit onderzoek is natuurlijk: Geeft het onderzoek een betrouwbaar beeld van de politieke voorkeur van de Nederlander boven de 18 jaar?

In dat verband geven de resultaten van een recente verkiezing in die beide delen van Nederland interessante informatie.

Uitslag (%)	PvdA	VVD	CDA	PVV	SP	Overige
Oost-Groningen	40	11	6	14	18	11
Wassenaar	11	53	8	9	4	15

De verschillen in stemgedrag tussen de beide regio's bij deze verkiezingen zijn groot. Dat is mogelijk een afspiegeling van de verschillen in politieke voorkeur tussen Wassenaar en Oost-Groningen.

Terug naar de vraag of het onderzoek van marketingbureau 'U vraagt, wij draaien' iets zegt over de politieke voorkeur van de Nederlander boven de 18 jaar. Die kans is klein, gezien het verschil in stemgedrag in de regio's waar het onderzoek plaats vindt. Hoogstens zegt het onderzoek iets over de politieke voorkeur van de inwoners van Oost-Groningen en Wassenaar.

Reflectie:

Aan welke eisen moet een onderzoek voldoen om iets te kunnen zeggen over de politieke voorkeur van de Nederlander van 18 jaar of ouder?

Het gaat erg ver om alle factoren die de kwaliteit van een onderzoek bepalen uitgebreid te bespreken. Bovendien rekent de 'Kennisbasis Rekenen-Wiskunde voor de Pabo' dat niet tot de bagage van de leerkracht basisschool. Toch is het belangrijk om bij het lezen en interpreteren van grafieken en vooral bij het trekken van conclusies te weten of het om goed onderzoek gaat. Goed onderzoek is niet alleen representatief, maar ook betrouwbaar en valide.

Representativiteit

Een steekproef is representatief als de steekproef van het onderzoek goed vergelijkbaar is met de totale populatie waar de steekproef onderdeel van is en uit is geselecteerd. Een representatief onderzoek wil zeggen dat het onderzoek het mogelijk maakt om op basis van de resultaten van de steekproef betrouwbare uitspraken te doen over de totale populatie. In Voorbeeld 1 is de totale populatie 'alle Nederlanders van 18 jaar of ouder'.

Om te bepalen of een onderzoek al of niet representatief is, kun je vragen stellen als:

1. Behoort de groep deelnemers (de steekproef) tot de totale groep (de populatie) waarover het onderzoek uitspraken wil doen en is de steekproef vergelijkbaar met de populatie?
2. Is de steekproef aselekt (niet selectief), dat wil zeggen: hebben alle leden van de populatie dezelfde kans om aan het onderzoek deel te nemen?
3. Hoe groot is de steekproef, met andere woorden: hoeveel mensen hebben er meegedaan aan het onderzoek? Zijn dat genoeg mensen om conclusies te kunnen trekken?
4. Hebben er mensen geweigerd mee te werken aan het onderzoek? Hoe groot is dat aantal in verhouding tot het totaal aantal deelnemers?

Het onderzoek uit Voorbeeld 1 'scoort' niet goed op deze vragen. Het antwoord op vraag 1 is 'nee' en de antwoorden op de andere drie vragen zijn niet duidelijk.

Betrouwbaarheid

Representativiteit is een onderdeel van betrouwbaarheid. Betrouwbaarheid betekent eigenlijk vooral 'herhaalbaarheid', dat wil zeggen: het onderzoek levert bij herhaling van het onderzoek met een vergelijkbare doelgroep ook vergelijkbare resultaten en vergelijkbare conclusies op. Is dat niet zo, dan moet je grote vraagtekens zetten bij de betrouwbaarheid van dat onderzoek.

Of een onderzoek betrouwbaar is, hangt van nog veel meer factoren af. De representativiteit van de steekproef is belangrijk, maar er zijn veel complexere factoren van invloed. De verwachtingen van de onderzoekers zelf blijken invloed te hebben op de resultaten van het onderzoek. In goed onderzoek worden maatregelen genomen om dat te voorkomen, bijvoorbeeld door (delen van) het onderzoek uit te laten voeren door iemand die niet weet wat het doel van het onderzoek is.

Je kunt je vast ook wel iets voorstellen bij:

- De kwaliteit van de vragen van een vragenlijst.
Zijn ze eenduidig, op meer manieren te interpreteren of zelfs suggestief?
- De beschikbaarheid van de populatie.
Kun je een representatieve steekproef samenstellen die groot genoeg is?
- De vorm van het onderzoek.
Gebruikt het onderzoek vragenlijsten, gestructureerde interviews of observaties? Onjuiste of onwaarschijnlijke antwoorden komen bijvoorbeeld vaker voor bij een vragenlijst dan bij een gestructureerd interview dat wordt opgenomen en achteraf wordt bekeken en gescoord.
- Het beschikbare budget.
Is er genoeg geld om noodzakelijke maatregelen te nemen die de betrouwbaarheid zo veel mogelijk garanderen?

Validiteit

Een goed onderzoek is, behalve representatief en betrouwbaar, ook valide. Validiteit is de mate waarin datgene wat je observeert of meet daadwerkelijk hetzelfde is als wat je wilt observeren of meten. Met andere woorden: je moet als onderzoeker wat je wilt onderzoeken op de juiste manier onderzoeken en met daarvoor geschikte middelen meten. Als voorbeeld:

- Het gemiddeld aantal ooievaars in Nederland is geen belangrijke factor in de toename of afname van het aantal kinderen dat wordt geboren en daarmee geen goede indicatie.
- Temperatuur meet je met een thermometer en niet met een rolmaat.

Significantie

Voor de interpretatie en conclusies van een (representatief, betrouwbaar en valide) onderzoek is de term 'significantie' van belang. Gebruik je de term 'significant' in een gesprek, dan bedoel je waarschijnlijk 'belangrijk' of 'betekenisvol'. Maar, gebruik je 'significant' in de context van een causaal (oorzakelijk) verband, dan bedoel je daar iets anders mee. Je wilt dan duidelijk maken dat zo'n verband - een effect of een relatie - niet berust op toeval.

Ben je nieuwsgierig naar opvattingen over de 'status van het vak leraar basisonderwijs' en houd je daarom een straatinterview voor het Centraal station Utrecht tijdens de NOT (de Nederlandse Onderwijs Tentoonstelling), dan zijn je resultaten waarschijnlijk sterk beïnvloedt door de factoren 'leraar' en 'NOT-bezoek' en niet representatief. M.a.w.: ze zijn niet 'significant' voor je vraagstelling en zeggen weinig over je oorspronkelijke vraag omdat ze sterk door twee 'toevalsfactoren' zijn beïnvloedt.

Significantie kun je bepalen m.b.v. statistische testen. Een significantie van 100% wil zeggen: er is geen enkele kans dat je resultaten zijn veroorzaakt door toeval. Dat komt natuurlijk zelden of nooit voor. De meest gebruikte regel m.b.t. significantie is de 95% regel: je kunt met 95% zekerheid zeggen dat je resultaten (een effect of relatie) niet ontstaan zijn door toeval.

Met andere woorden: je bevindingen zijn zeker 95 van de 100 keer echt en maximaal 5 van de 100 keer niet. Je accepteert dus ook een kans van 5% dat wat je vindt niet waar is.

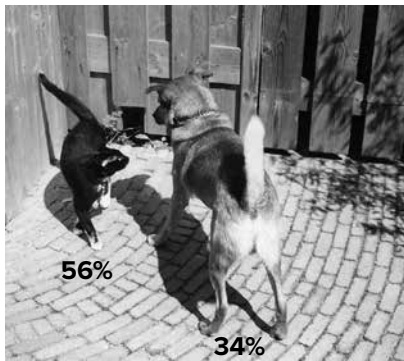
Zoals gezegd: deze onderwerpen zijn van belang bij het verrichten van zorgvuldig onderzoek en als informatie bij het lezen en interpreteren van grafieken. Misschien vind je het interessant om je hier verder in te verdiepen, bijvoorbeeld in het licht van je onderzoeksstage. Maar, in de toets over de Kennisbasis kom je geen vragen tegen over representativiteit, betrouwbaarheid en validiteit.

Conclusie:

Kenmerkend voor statistiek is dat het lijkt of je er van alles mee aan kunt tonen. Bij het interpreteren van statistische gegevens en het werken met grafieken moet je goed opletten of het onderzoek zelf op een juiste manier is gedaan. Wees kritisch op de vraag of de indruk die grafieken wekken juist is, of dat er misschien sprake is van manipulatie. Met andere woorden: op grond van welke gegevens zijn conclusies getrokken? Soms is het overduidelijk wat de grafiek zogenaamd moet aantonen en soms zijn grafieken bewust suggestief gemaakt om een bepaalde boodschap over te brengen.

Voorbeeld 2 Als kat en hond

De procentuele verdeling van honden en katten in Nederland is weergegeven in dit pictogram.



Opdracht:

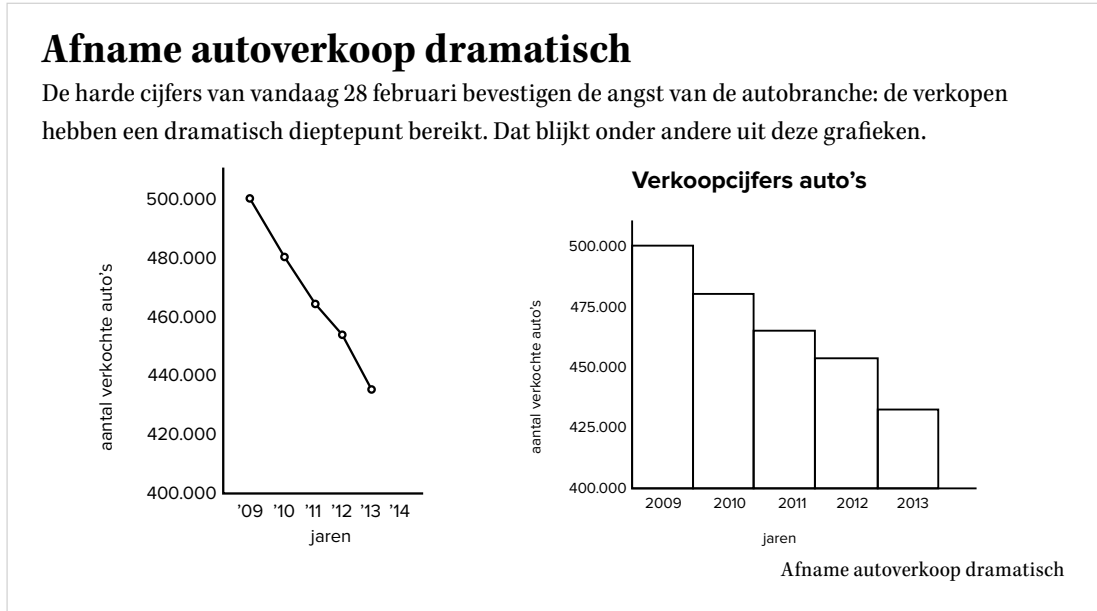
Noteer wat je opvalt aan deze foto en de percentages.

Oplossing:

Hond Joepie is als grootste afgebeeld en kat Miepie als kleinste. Dat suggereert dat er meer honden zijn dan katten. Maar, de percentages onder de dieren geven aan dat er (bijna) tweemaal zoveel katten zijn als honden. De grootte van de figuurtjes in de pictogram zet je dus op het verkeerde been. Een staaf- of cirkeldiagram is hier beter op z'n plaats. Een beelddiagram met de juiste verhoudingen kan ook.

Voorbeeld 3 Afname autoverkoop dramatisch

De titel van een krantenartikel suggereert dat de autoverkoop dramatisch is gedaald. Een grafiek in het artikel onderstreept de spectaculaire daling.



Opdracht:

Wat is er mis met deze grafieken?

Oplossing:

De gegevens in deze tabel geven de werkelijke situatie weer.

Verkoopjaren	2009	2010	2011	2012	2013
Aantal verkochte auto's	500.000	480.000	465.000	455.000	435.000

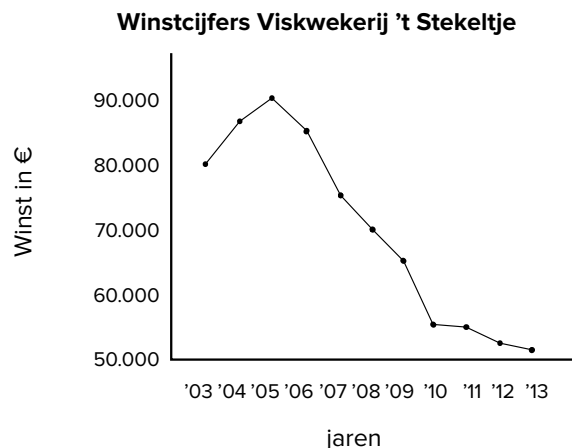
Toelichting:

Beide grafieken beginnen bij 400.000 auto's en niet bij 0. Omdat de verticale schaal van de grafieken bij 400.000 begint, lijkt het alsof de verkoopcijfers tot bijna 0 zijn teruggelopen. Als je ze tekent met een verticale schaal die wel bij 0 begint, zie je ook een daling, maar die is niet zo spectaculair. De totale daling is 65.000 over 4 jaar, oftewel 13%. De tabel vertelt dus het ware verhaal.

► Opgave 1 Dramatische winstdaling 't Stekeltje

De crisis slaat ook toe bij viskwekerij 't Stekeltje. Deze grafiek laat dat zien.

- a. De winst over 2013 lijkt bijna € 0,-. Klopt dat? Licht je antwoord toe.
- b. Hoeveel procent is de winst in de periode 2008 tot en met 2013 gedaald?
- c. Teken een eerlijkere winstgrafiek. Gebruik ruitjespapier van 0,5 cm x 0,5 cm.



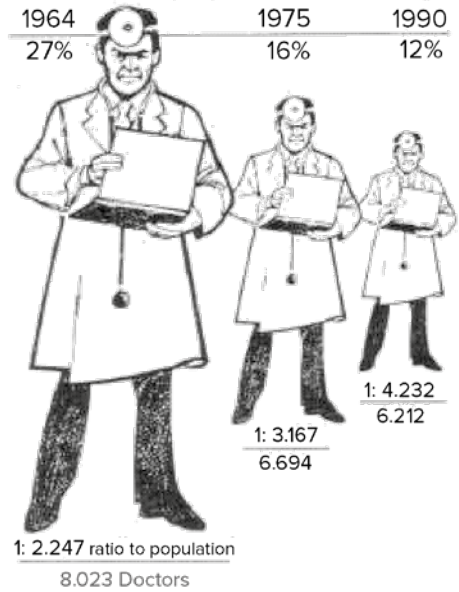
► **Opgave 2** Huisartsen en inwoners

In deze beeldgrafiek over het aantal huisartsen staan veel gegevens: jaartallen, percentages, aantallen artsen en ook de verhouding 'inwoners per arts'.

In 1964 is 27% van alle artsen alleen huisarts. Er zijn 8.023 huisartsen, gemiddeld is er 1 huisarts op 2.247 inwoners. De gegevens van 1975 en 1990 staan ook in de beeldgrafiek. Beantwoord de vragen.

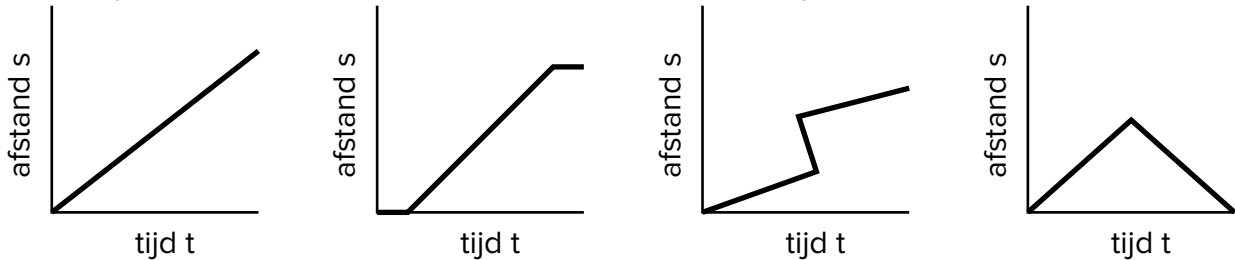
- Hoe zijn deze gegevens grafisch weergegeven?
- Is de beeldgrafiek misleidend? Licht je antwoord toe.

Percentage of Doctors Devoted Solely to Family Practice		
1964	1975	1990
27%	16%	12%



► **Opgave 3** De fietstocht van Ruud

Ruud maakt een fietstocht van vier dagen en zet die in vier grafieken. Op de assen staan geen getallen. Toch kun je uitspraken over de fietstocht doen aan de hand van de grafieken.



Over grafiek 1 kun je bijvoorbeeld zeggen: Ruud vertrekt en fietst steeds met dezelfde snelheid. Beantwoord de vragen.

- Beschrijf de 2e dag van de fietstocht aan de hand van de 2e grafiek.
- Beschrijf de derde dag. Wat gaat er volgens de grafiek mis, denk je?
- De grafiek van dag 4 ziet er wat vreemd uit. Geef een beschrijving van die dag.