

maar vanuit methodologisch oogpunt om selectiebias en verwante vormen van misrepresentatie van een populatie of domein te verminderen.

## Epiloog

Beperken we ons even tot de statistiek, dan is genoegzaam bekend dat fundamenteel statistisch onderzoek dikwijls moeizaam of niet met belangrijke prijzen gehonoreerd wordt, niet alleen binnen de wiskunde, maar ook wanneer het gaat om een ruimer, interdisciplinair domein. Daarnaast is het een feit dat statistische bijdragen sedert de opkomst van het vak in de tweede helft van de 19e eeuw er lange tijd bekaaid vanaf kwamen in de traditionele wetenschapshistorische en -filosofische literatuur. In de canonvorming van de ideeëngeschiedenis bleven zij daarmee stelselmatig onderbelicht. Daarin kwam ongeveer veertig jaar geleden enige verandering, maar dan toch vooral door het werk van gespecialiseerde historici van de statistiek, onder wie voornoemde Stephen Stigler.

Tot slot nog even dit. In 1922, precies 100 jaar geleden ontving Albert Einstein de Nobelprijs voor Natuurkunde, zoals bekend niet voor de relativiteitstheorie en merkwaardigerwijs ook nog eens met terugwerkende kracht voor het jaar 1921. Wie nog in de veronderstelling mocht verkeren dat de toekenning van 's werelds meest prestigieuze prijs als vanzelfsprekend transparant, rationeel en eerlijk verloopt, wordt ruw wakker geschud in een tweetal recente, historische artikelen in het *Nederlands Tijdschrift*

voor *Natuurkunde* (Van Lunteren, 2021; Kox, 2022). Niets menselijks bleek de commissieleden vreemd; persoonlijke voorkeuren en antipathieën, oude rekeningen die nog moesten worden vereffend, verborgen agenda's en soms ronduit bizarre compromissen speelden een niet te onderschatten rol. Zelfs Einstein kreeg ermee te maken. Hoe dan ook, bij een al te idyllische voorstelling van de sacrale status en betekenis van wetenschappelijke prijzen kunnen gedetailleerde historische case-studies soms een welkome nuancering bieden.

### LITERATUUR

- Kox, A. (2022). Terug in de tijd. *Nederlands Tijdschrift voor Natuurkunde*, 1, 31–32.
- Lunteren, F. van. (2021). De wrok van Arrhenius; De beginjaren van de Nobelprijzen voor de Natuur- en Scheikunde, *Nederlands Tijdschrift voor Natuurkunde*, 1, 14–17.
- Starmans, R. J. C. M. (2012). Betwiste paradigma's en vermeende revoluties; De moeizame ironie van Stiglers Wet. *Filosofie-Tijdschrift*, 22(2).
- Starmans, R. J. C. M. (2022a). De reikwijdte van de counterfactual; Over causaliteit, potential outcomes en grafische modellen. *STAtOR*, 23(1).
- Starmans, R. J. C. M. (2022b). De Sokal-affaire in retrospectief; Over parodieën, crises en wantrouwen in wetenschap. *Filosofie-Tijdschrift*, 32(1).

RICHARD STARMANS is verbonden aan de Faculteit Bèta-wetenschappen (Department of Information and Computing Sciences) van de Universiteit Utrecht en aan Tilburg University. Hij doet onderzoek op het snijvlak van filosofie, statistiek en informatica.

E-mail: starmans@cs.uu.nl

## Good news from Young Statisticians



After another long lockdown-winter, the Young Statisticians section was finally able to organize its first event of 2022, which took place on April 7th: a Statistics Café on causality. Oisín Ryan, Bart Eggen and Jesse Krijthe each gave interesting talks on different aspects of this central topic within statistics.

We aim to soon plan a company visit to an organization that uses statistics and data science in daily practice. (If you have any suggestions for this, don't hesitate to contact us!) To keep informed on our future events, or to subscribe to our newsletter, make sure to visit our website: [vvsor.nl/young-statisticians](https://vvsor.nl/young-statisticians).

## PEILINGPRAKTIJKEN

### Rommel niet met de assen van een grafiek!



Om de omvang van allerlei verschijnselen uit een grafiek te kunnen aflezen, heb je assen nodig. Die assen horen voorzien te zijn van maatstreepjes en bij die maatstreepjes horen waarden te staan. Bij het maken van een grafiek moet je die assen op een correcte manier tekenen. Doe je dat niet dan leidt de grafiek al gauw tot verkeerde interpretatie.

Een veel voorkomende fout is de as niet laten beginnen bij 0. Als de as een omvang, hoeveelheid, waarde, duur, aantal of percentage moet aanduiden, dan is er een natuurlijk nulpunt. De as moet dan bij dit nulpunt beginnen. Als dat niet het geval is, dan lijken allerlei verschijnselen groter dan ze in werkelijkheid zijn. Er is dus sprake van overdrijving. De grafiek is suggestief. Aan de hand van een paar voorbeelden laten we zien hoe een grafiek je op het verkeerde been kan zetten.

### De nieuwe Amerikaanse zorgverzekering Obamacare

Een eerste voorbeeld van gerommel met de as komt van Fox News. Dat is een van de grote nieuwszenders in de

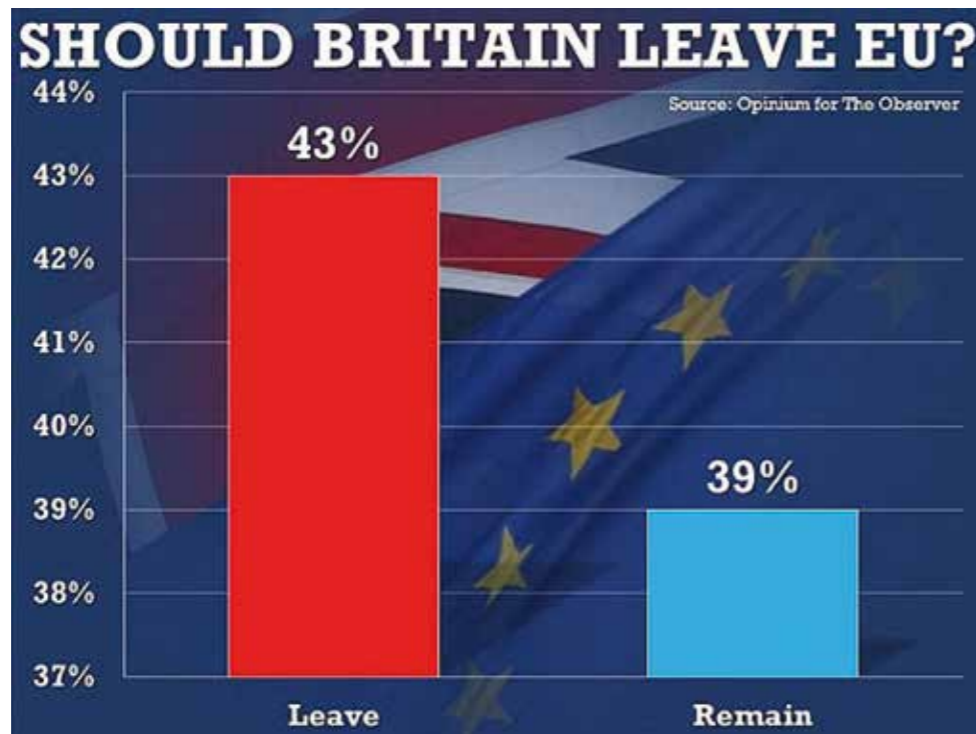
Verenigde Staten. De zender is de stem van de rechtse, conservatieve Amerikanen. In de berichtgeving liet Fox geen mogelijkheid onbenut om de regering van president Barack Obama in een kwaad daglicht te zetten.

Het was een van de belangrijke doelen van de Amerikaanse president Barack Obama om het systeem van gezondheidsverzekeringen te hervormen. Daarvoor voerde hij een nieuwe wet in: de *Patient Protection and Affordable Care Act*. De wet verplichtte iedere Amerikaan een zorgverzekering af te sluiten. En de zorgverzekeraars mochten niemand weigeren. In de volksmond werd deze wet *Obamacare* genoemd. De Republikeinen waren fel tegen de wet. Daarom probeerde Fox News de invoering van Obamacare af te schilderen als een volledige flop. De regering Obama mikte erop dat op 31 maart 2014 zo'n 7 miljoen Amerikanen zich moesten hebben aangemeld voor de nieuwe zorgverzekering. Dus probeerde Fox News op 27 maart 2014 met de linker grafiek in figuur 1 aan te geven dat dit doel bij lange na niet was gehaald.

Het verschil tussen de twee staven in het linker staafdiagram is inderdaad bijzonder groot. De grafiek suggereert dat het aantal aanmeldingen zelfs minder is dan de helft van de doelstelling. Maar een nadere blik



Figuur 1. Fox News rommelt met een grafiek over Obamacare. Bron: Fox News, 2014



Figuur 2. De Daily Mail rommelt met een grafiek over een Brexit-peiling. Bron: Daily Mail, 2016

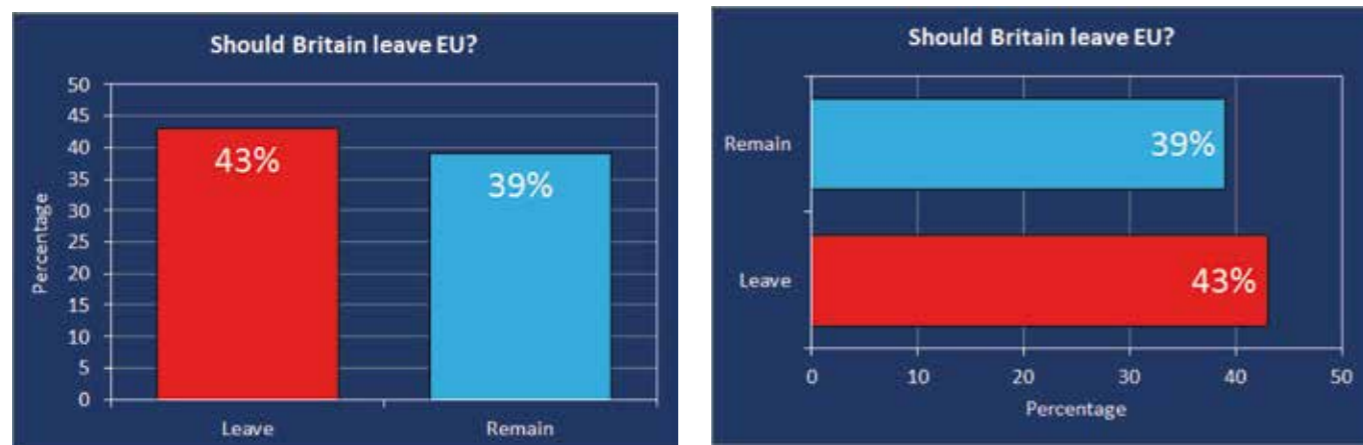
op de cijfers leert dat in feite het verschil helemaal niet zo groot is: er zijn immers al 6 miljoen aanmeldingen en dat is maar 1 miljoen minder dan de doelstelling van ruim 7 miljoen. De staven weerspiegelen dus niet de werkelijke cijfers. Dat komt omdat de verticale as niet bij 0 begint. Dat is niet eenvoudig vast te stellen omdat die verticale as helemaal geen schaalverdeling heeft gekregen. Daarmee is dit een bijzonder slechte grafiek. Als je niet goed oplet, wordt je al snel op het verkeerde been gezet.

Er was veel kritiek van deskundigen op de grafiek van Fox News. Dat leidde ertoe dat Fox News de grafiek een paar dagen later (op 1 april 2014) corrigeerde. De rechter

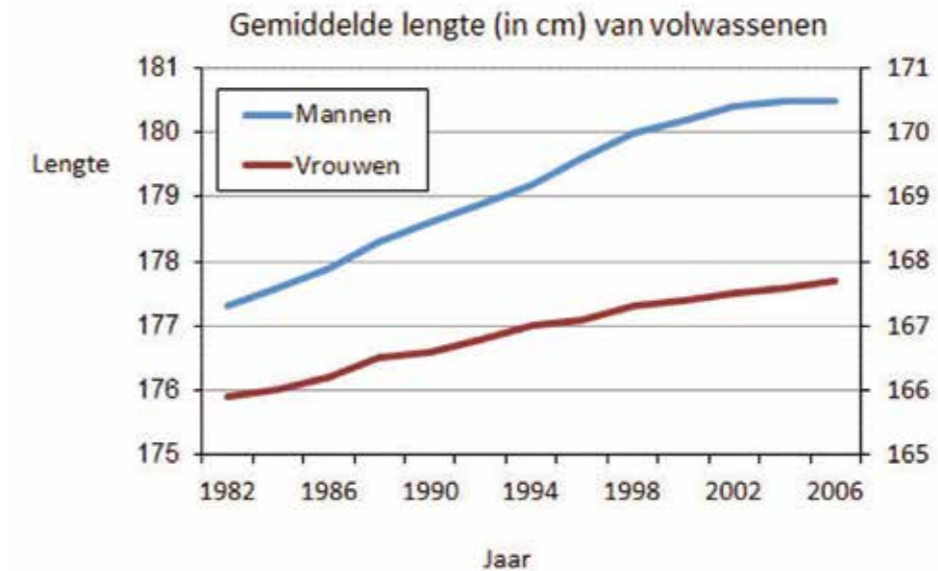
grafiek in figuur 1 is de gecorrigeerde versie van de grafiek. De verticale as heeft nu een keurige schaalverdeling. En die schaalverdeling begint, zoals het hoort, bij 0. Het verschil tussen beide staven is nu veel minder groot. De grafiek geeft een veel realistischer beeld van de situatie.

### Willen de Britten uit de EU?

Een tweede voorbeeld van het manipuleren van de as van een grafiek komt van een Britse peiling over *Brexit*, het verlaten van de Europese Unie (EU) door het Verenigd Koninkrijk (VK). De grafiek in figuur 2 is afkomstig uit de



Figuur 3. Twee versies van de gerepareerde grafiek over de Brexit-peiling



Figuur 4. Het CBS rommelt met de assen van een grafiek met de lengte van mannen en vrouwen. Bron: CBS-Webmagazine, 17 januari 2008

Britse krant *Daily Mail* van 3 april 2016. De boodschap lijkt duidelijk: een groot deel van de bevolking (de rode staaf) wil dat het VK de EU verlaat (*Brexit*). Slechts een klein deel van de bevolking (de blauwe staaf) wil dat het VK in de EU blijft.

Helaas, er is wat mis met de grafiek: de verticale as begint niet met 0, maar bij 37%. Door deze truc lijkt het verschil tussen 'Leave' and 'Remain' veel groter dan het in werkelijkheid is. De krant overdrijft dus het verschil.

Figuur 3 bevat twee gerepareerde versies van het staafdiagram in figuur 2. De belangrijkste wijziging is dat de as nu bij 0 begint. Daardoor is het verschil tussen 'Leave' en 'Remain' veel kleiner. Dat klopt wel, want het verschil is maar vier procentpunten.

Een tweede wijziging is dat de rechter grafiek in figuur 3 gedraaid is zodat de staven nu horizontaal zijn getekend in plaats van verticaal. Een grafiek als deze biedt meer ruimte voor (horizontale) teksten en bevordert daarmee de leesbaarheid. De rechter grafiek heeft mijn voorkeur.

### Zijn mannen veel langer dan vrouwen?

Ten slotte nog een derde voorbeeld van gerommel met de as. Het gaat om een grafiek die in 2008 door het Centraal Bureau voor de Statistiek (CBS) is gemaakt. In figuur 4 is deze grafiek gereproduceerd.

Een eerste blik op deze grafiek roept de indruk op van een sterke toename van de gemiddelde lengte. Dat geldt vooral voor mannen (de blauwe lijn) en in iets mindere mate voor vrouwen (de rode lijn). Verder wekt de grafiek

de suggestie dat vrouwen gemiddeld maar half zo lang zijn als mannen.

Er zijn twee problemen met de grafiek. In de eerste plaats zijn er twee verticale assen. De linker as is voor de mannen en de rechter as moet je gebruiken voor de vrouwen. De twee assen lopen niet synchroon. Daardoor kun je mannen en vrouwen niet met elkaar vergelijken. In de tweede plaats beginnen beide assen niet bij het natuurlijke nulpunt. De linker verticale as begint bij 175 en de rechter bij 165.

Als je dezelfde as zou gebruiken voor mannen en vrouwen en bovendien die as bij 0 zou laten beginnen, dan krijg je een veel rustiger beeld van de ontwikkeling van de gemiddelde lengte van mannen en vrouwen. De groei gaat minder hard en de verschillen tussen mannen en vrouwen zijn minder groot.

### Conclusie

Je ziet regelmatig grafieken langskomen waarin gemiddeld is met de verticale as. Zulke grafieken roepen meestal een verkeerd beeld op van de situatie. Daarom is het goed om altijd eerst even naar de as van een grafiek te kijken voordat je hem gaat interpreteren.

JELKE BETHLEHEM werkte bij het CBS en is emeritus hoogleraar aan de Universiteit Leiden. Hij is een expert op het gebied van steekproeven, vragenlijsten en weergave van onderzoeksresultaten. Deze onderwerpen behandelt hij regelmatig in zijn blog. E-mail: mail@jelkebethlehem.nl