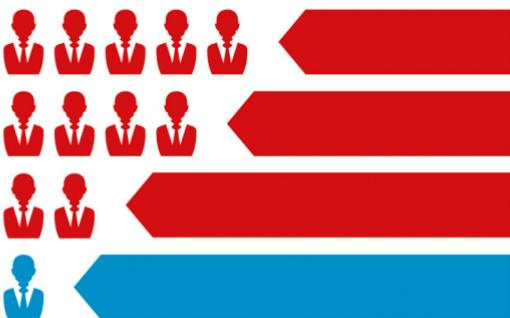
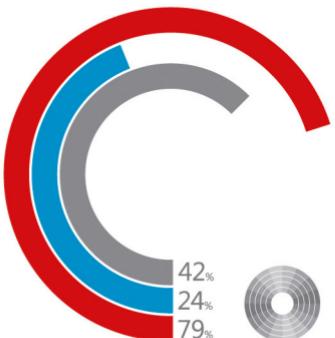


SO LÜGT MAN MIT STATISTIK



So lügt man mit Statistik

Walter Krämer ist Professor für Wirtschafts- und Sozialstatistik an der TU Dortmund und Autor verschiedener Bestseller.

Walter Krämer

SO LÜGT MAN MIT STATISTIK

Campus Verlag
Frankfurt/New York

ISBN 978-3-593-50459-9

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlags unzulässig. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Copyright © 2015 Campus Verlag GmbH, Frankfurt am Main
Umschlaggestaltung: total italic, Thierry Wijnberg,
Amsterdam/Berlin

Umschlagmotiv: Thinkstock

Innengestaltung und Satz: Oliver Schmitt

Gesetzt aus: Univers und Yoga Pro

Druck und Bindung: Beltz Bad Langensalza

Printed in Germany

Dieses Buch ist auch als E-Book erschienen.

www.campus.de

Inhalt

Vorwort zur Neuausgabe 7

Vorwort 11

1 Die Illusion der Präzision 17

2 Vorsicht, Kurve! 29

3 Hunde beißen lieber Männer 47

4 Die Konfusion des Staatsanwalts 57

5 Manipulierte Mittelwerte 63

6 Protzige Prozente 75

7 Trügerische Trends 85

8 Der synthetische Superlativ 93

9 Die vorsortierte Stichprobe 101

10 Das frisierte Piktogramm 117

11 Wie es in den Wald hineinschallt ... 129

12 Statistisch oder wirklich arbeitslos? 145

13 Arm und reich - alles gleich? 153

14 Allerlei Kartentricks 163

15 Korrelation und Kausalität 171

16 Fluglärm erzeugt Aids 185

17 Welchen Zahlen kann man noch vertrauen? 195

Register 201

»Man hat behauptet, die Welt werde durch Zahlen regiert:
das aber weiß ich, dass die Zahlen uns belehren,
ob sie gut oder schlecht regiert werde.«

Johann Wolfgang von Goethe, Gespräche mit Eckermann

Vorwort zur Neuausgabe

Kann man einem Autor eine größere Freude machen, als ihn 25 Jahre nach Erscheinen seines Buches zu einem neuen Vorwort aufzufordern? Schließlich heißt das ja: Das Buch wird immer noch gekauft. Und das hätte ich damals, im Sommer 1990, als ich es auf einer abgelegenen alten Farm im englischen Devonshire geschrieben habe, tatsächlich nicht gedacht.

Inzwischen ist es zu einem Klassiker mit mehreren Dutzend Auflagen und Übersetzungen ins Italienische, Koreanische und Chinesische geworden. Das ist für mich als Autor angenehm, für mich als Statistiker sowohl gut wie schlecht. Gut, weil das bedeutet: Die Menschen sind an dem Thema interessiert. Schlecht, weil das Thema immer noch ein Thema ist. Verglichen mit 1990 scheinen zwar gewisse Arten von Missbrauch und Schlampereien in den letzten Jahren an Verbreitung eingebüßt zu haben - Datengrafiken mit abgeschnittenen Füßen und selektiv gestauchten Achsen wie in dem Kapitel »Vorsicht, Kurve!« etwa sieht man heute weniger, hier haben die Journalisten, die Hauptlieferanten meiner Beispiele, einiges dazugelernt. Dafür schleichen sich andere Arten des Missbrauchs ein. Von der brutalen Fälschung ganzer Staatsbilanzen à la Griechenland will ich dabei gar nicht reden (das hat übrigens die alte DDR mindestens genauso gut gekonnt, siehe weiter unten in Kapitel 17). Auf einer Ebene tiefer, aber mit langfristig genauso fatalen

Folgen werden auch die Sünden im Umgang mit Unsicherheiten und Wahrscheinlichkeiten immer schlimmer, und auch der korrekte Umgang mit Prozenten scheint unserer Spezies so fremd wie eh und je.

Ich glaube inzwischen, dass viele dieser Fehler eine Folge - nicht fehlender Intelligenz, sondern der internen Verdrahtung unseres Gehirns sind, quasi das genetische Erbe, das wir von den Affen Afrikas in das dritte Jahrtausend hineingetragen haben. Wie der Psychologe und Wirtschaftsnobelpreisträger Daniel Kahneman in seinem Weltbestseller *Schnelles Denken langsames Denken* versucht hat nachzuweisen, ist dieser von den Affen ererbte und von ihm System 1 genannte Teil unseres Gehirns ein notorisch schlechter Statistiker. Es kann sehr gut addieren und Summen vergleichen, aber es kann keine Durchschnitte bilden, nicht multiplizieren und nicht dividieren (und erst recht natürlich keine Wurzeln ziehen oder Logarithmen ausrechnen). Und da ein Großteil unseres Denkens - das schnelle Denken - nahezu automatisch und ohne unser Zutun mittels System 1 geschieht, machen wir ebenso automatisch immer Fehler, wenn wir statt zu addieren hätten multiplizieren müssen. Und viele statistische Trugschlüsse sind genau von dieser Art.

Diese Fehlerquelle ist allenfalls genetisch, d.h. in 100 000 Jahren, und nicht mit Büchern wie diesem auszumerzen. Was dieses Buch aber versucht, ist ein Appell an das von Kahneman so genannte System 2, das ist der Teil unseres Gehirns (das langsame Denken), der nur mit willentlicher Anstrengung funktioniert und Mühe macht, der eigentliche Denker sozusagen. Deswegen schalten wir diesen Teil nur ein, wenn wir glauben, dass es sich lohnt. Und dieses Buch will zeigen: Ja es lohnt sich, es will seine Leser

animieren, beim Umgang mit Statistik öfter als vielleicht bisher den eigentlichen Denker zu benutzen, in der Hoffnung, dass ganz im Sinn des bekannten deutschen Statistikers Ernst Wagemann über das Falsche das Richtige dann doch ans Tageslicht gelangt: »Es hat sich die alte Erfahrung bewährt«, schreibt Wagemann in seinem *Narrenspiegel der Statistik* (1935), »dass das Wesen der Dinge unserem Verständnis am besten durch die Betrachtung ihrer Kehrseite erschlossen wird.« Von dieser Kehrseite sehen Sie auf den folgenden Seiten genug. Vielleicht macht das auch ein wenig neugierig auf die Vorderseite?

*Walter Krämer
Dortmund, im Sommer 2015*

Vorwort

»Get your facts first, and then
you can distort 'em as you please.«

Mark Twain

Wir schlagen morgens unsere Zeitung auf - und ehe das letzte Blatt gewendet ist, haben wir mehr Statistiken gesehen als Goethe oder Schiller während ihres ganzen Lebens. Die Arbeitslosen werden je nach Quelle weniger oder mehr, die Krebsgefahr und das Ozonloch nehmen zu, ein Drittel aller Menschen lebt unter dem Existenzminimum, ein Bundesbürger verzehrt im Jahr 8 Liter Speiseeis, Musiker leben länger als andere, Landluft ist gesund, Landluft ist ungesund, im Jahr 2050 gibt es 10 Milliarden Menschen auf der Welt, der Dow-Jones-Index hat 30 Punkte zugelegt, Tennisspieler B hat noch nie gegen einen brilletragenden Linkshänder verloren, der jünger war als er selbst, 25 Prozent mehr Frauen als Männer biegen falsch in Einbahnstraßen ein, die Dinosaurier starben vor 65 Millionen Jahren aus, von Schweinefleisch bekommt man Herzinfarkt, Ausländer sind kriminell, Wähler der Grünen dagegen häufig impotent (kein Scherz - dies erklärte ein Hamburger Sexualforscher tatsächlich vor dem Deutschen Bundestag), in New York schläft man am sichersten im Central Park, noch 30 Jahre trennen uns vom nächsten

GAU, eine Klimakatastrophe droht, und wäre am nächsten Sonntag Wahl, käme Partei X nicht in den Bundestag.

Viele dieser Statistiken sind falsch. Einige sind bewusst manipuliert, andere nur unpassend ausgesucht. In einigen sind schon die reinen Zahlen falsch, in anderen sind Zahlen nur irreführend dargestellt. Dann wieder werden Äpfel und Birnen zusammengeworfen, Fragen suggestiv gestellt, Trends fahrlässig fortgeschrieben, Raten, Quoten oder Mittelwerte kunstwidrig berechnet, Wahrscheinlichkeiten vergewaltigt oder Stichproben verzerrt, sodass Lüge und Statistik für viele zusammengehören wie Pat und Patachon. Seit Benjamin Disraelis »There are three kinds of lies: lies, damned lies, and statistics« hageln die einschlägigen Bonmots auf die armen Statistiker nur so herab.

Diese Verunglimpfungen sind zwar billig, aber in gewissem Sinn wahr. Billig, weil sich jedes Werkzeug sowohl ge- als auch missbrauchen lässt. Hier steht die Statistik sicher nicht allein. Wahr, weil wir alle nur zu gern die Welt so sehen, wie wir sie gern hätten, statt so, wie sie wirklich ist. Wir benutzen die Statistik nur zu oft »wie ein Betrunkener einen Laternenpfahl: vor allem zur Stütze unseres Standpunkts und weniger zum Beleuchten eines Sachverhalts« (Andrew Lang).

Dieses Buch fasst meine einschlägigen Erfahrungen aus vielen Berufsjahren als Statistiker zusammen. Dabei predige ich nicht (jedenfalls nicht mit Absicht) vom hohen Ross des professionellen Datensammlers. Die ehrliche Präsentation von Fakten ist keine Sache des Könnens, sondern des Willens, und vor allem daran scheint es vielen Datensündern offenbar zu mangeln.

Es gibt ein Gentleman's Agreement unserer Medien, dass im Dienste einer guten Sache die Wahrheit nicht so wichtig

ist. So rechtfertigte das *Deutsche Ärzteblatt* einen Fehler in der Aids-Statistik - nämlich durch sogenanntes »Kumulieren« die aktuellen Krankenstände höher darzustellen, als sie wirklich sind - mit den Forschungsgeldern, die so leichter einzuwerben seien. »Wenn die Kumulierung zu diesem Effekt beiträgt«, lesen wir dort schwarz auf weiß, »dann sollten wir es noch eine Weile dabei belassen.«

Dieses Reklamieren einer »License to Lie« im Dienste eines subjektiven oder objektiven guten Zwecks kennt keine Parteidgrenzen; es wird von Linken wie Rechten, Progressiven wie Konservativen gleichermaßen praktiziert. Allenfalls nimmt es mit der Gewissheit zu, mit der sich der Datenkosmetiker im Besitz der absoluten Wahrheit wähnt. Wer sicher weiß, dass die Welt in zwanzig Jahren untergeht, wenn nicht dieses oder jenes geschieht, fühlt sich durch Konventionen wie Faktentreue und Sachlichkeit in seinem Rettungswerk oft sehr gehemmt. »Je dramatischer wir die Sache sehen, desto besser für die Menschheit«, führt etwa ein amerikanischer Klimaforscher als Entschuldigung für reichlich gewagte Trendextrapolationen unseres Wetters an; als Wissenschaftler müsse man »manchmal auch ein bisschen Panik verursachen, damit man gehört wird«, stößt ein Ozonexperte aus Deutschland in das gleiche Horn.

Aber nicht nur Retter der Menschheit möchten Dinge anders erscheinen lassen, als sie wirklich sind. Diese Versuchung teilt ein Greenpeace-Demonstrant mit einem erfolglosen Konzernchef vor der Hauptversammlung oder mit einem Jäger nach der Jagd. Sie spürt ein treuloser Ehepartner genauso wie ein Kurdirektor, der um Gäste werben soll, oder wie der Präsident des ADAC bei pro und kontra Tempolimit. Seit es Daten gibt, gibt es auch die Datenmanipulation.

Zwischen diesen mehr oder weniger dreisten Manipu-

lationsversuchen mache ich im Folgenden keinen Unterschied - hier landen alle statistischen Falschmünzer im gleichen Topf. Damit meine ich natürlich nicht, dass eine irreführende Werbung für ein Waschmittel genauso verwerlich ist wie etwa die Manipulation der Krebsgefahr; von einer höheren Warte gibt es dabei durchaus einen Unterschied. Hier kommt es mir jedoch allein auf die Methoden und nicht auf Anlass oder Absicht an.

Um diese Methoden zu durchschauen, braucht es kein Statistikstudium. Wer die Erinnerung an den einschlägigen Unterricht an Schule oder Hochschule noch schwer im Magen spürt, kann ganz beruhigt sein - mehr als die vier Grundrechenarten und den Willen, den kleinen und großen Datensünden des Alltagslebens auf die Spur zu kommen, braucht man zum Verständnis dieses Buches nicht. Schließlich muss man auch nicht selbst Caruso heißen, um zu hören, ob ein anderer richtig singt.

Die Grundmuster vieler Manipulationen mit Statistik haben schon viele andere aufgedeckt und vorgestellt. Eine konstante Quelle der Inspiration war mir etwa Darrell Huff mit seinem immer noch taufrischen Klassiker *How to Lie With Statistics* (New York 1954), dessen deutsche Ausgabe *Wie lügt man mit Statistik* (Zürich 1956) längst vergriffen ist. Seit damals haben jedoch Statistiker wie Datensünder viel dazugelernt, sodass es außer beim Titel kaum Überschneidungen mit den folgenden Seiten gibt. Eine sehr einfühlsame Betrachtung in deutscher Sprache bieten auch Thomas Lachs und Eva Maria Nesvadba in *Statistik - Lügen oder Wahrheit* (Wien 1986). Weitere Literatur- und Quellenhinweise werden jeweils am Ende eines Kapitels nachgereicht.

Viele der folgenden Exponate verdanke ich meinen Studenten, andere geschätzten Kollegen. Bei der Umsetzung der

Buchidee habe ich in vielfältiger Weise Unterstützung erfahren. Ich danke an dieser Stelle allen Beteiligten herzlich und halte der guten Form halber fest, dass verbleibende Fehler und insbesondere kontroverse Wert- und Vorurteile, die aus den folgenden Seiten herausgelesen werden können, allein dem Autor anzulasten sind.

»Heaven and earth were created together, in the same instant, by the Trinity on Sunday, twentyfirst of October, 4004 B.C., at nine o'clock in the morning.«

James Ussher (1581-1656), Erzbischof von Armagh, in seinem Buch Annals of the West

Kürzlich saß ich an meiner Steuererklärung. Man kennt das ja. So und so viele Ausgaben für das und das, aber wo ist der Beleg? Vergessen und verschlampt. Ärgerlich, wenn man nur deshalb zu viel Steuern zahlen muss.

Aber halt! Akzeptiert nicht das Finanzamt bestimmte Kosten auch ohne Beleg? Nach einem Urteil des Bundesfinanzhofs reicht es aus, wenn »nicht nur ein geringes Maß an Wahrscheinlichkeit« für deren Echtheit spricht, und diese Wahrscheinlichkeit, obwohl am einfachsten durch Belege zu untermauern, ist grundsätzlich auch anders herstellbar. Die Frage ist nur, wie.

Ich habe es so gemacht: Statt »60 Euro«, was ähnlich klingt wie »etwa 60 Euro, aber genau weiß ich das nicht mehr«, habe ich geschrieben »Ausgaben für Büromaterial: Euro 58,24«. Das akzeptiert der Sachbearbeiter vermutlich eher als »etwa 60 Euro«; die Gefahr, dass er weitere Beweise sehen will, ist hier nur halb so groß. Mit anderen Worten, ich habe mit Daten manipuliert.

Am 6. April des Jahres 1909 hatte der Amerikaner Robert E. Peary als erster Mensch der Welt den Nordpol erreicht. Zumindest glaubte er das. Beziehungsweise er wollte, dass die Welt ihm diese Leistung glaubte; deshalb gab er seine Position mit 89 Grad, 57 Minuten und 11 Sekunden nördlicher Breite an, rund 5 Kilometer vom Pol entfernt – für alle praktischen Zwecke ein Volltreffer, und genau dieser Effekt war auch geplant.

In Wahrheit konnte Peary seine Position unmöglich bis auf 30 Meter genau bestimmen (denn auf nichts anderes laufen die 11 Bogensekunden hinaus). Selbst heute, mit den modernen Hilfsmitteln der Satellitennavigation, wäre das ein kleines Meisterstück. Damals war es ein Ding der Unmöglichkeit. Selbst Pearys Freunde geben zu, dass er mit seinen beschränkten Mitteln die Position bestenfalls bis auf 6 Bogenminuten oder 10 Kilometer genau hätte angeben können. Die restlichen Ziffern hatte er sich einfach ausgedacht.

Irgendwo habe ich einmal gelesen, dass es auf der Welt 8 Millionen 523 Tausend und exakt 12 Menschen mit Ungarisches Muttersprache gibt. Auch diese Statistik ist sicher falsch, es sei denn, Gott der Allmächtige selbst hätte sie erstellt. Wahrscheinlich hat ein Hobby-Datensünder alle Ziffern außer der ersten à la Peary frei erfunden.

Trotzdem glauben wir sie ihm. Wenn wir in einer fremden Ortschaft fragen: »Wie weit ist es zur nächsten Post?«, und einer sagt: »Na ja, drei Kilometer oder so, dann links«, ein anderer aber: »2,4 Kilometer geradeaus, dann rechts«, wem glauben wir? Gehen wir an der kritischen Kreuzung links oder rechts?

Ich glaube, die meisten gehen rechts. Wir gehen rechts aus dem gleichen Grund, warum wir einem Nordpolforscher

glauben, der seine Position bis auf die Bogensekunde kennt, oder einem Linguisten, der uns bis auf die einzelne Seele sagt, wer diese oder jene Sprache spricht. Kleider machen Leute und Ziffern machen Zahlen, und je mehr Ziffern eine Zahl umhängen hat, desto mehr vertrauen wir ihr auch.

Auch in der Literatur ist dieses Stilmittel weit verbreitet. »Mit fröhlichem Lächeln tat Gargantua seinen Hosenlatz auf«, schreibt Jonathan Swift in *Gullivers Reisen*, »zog den Spritzschlauch hervor und ließ einen solchen Erguss auf die Umstehenden los, dass 260 418 Pariser eines bitteren und feuchten Todes starben, Weiber und Kinder gar nicht eingerechnet.«

Oder was ist von folgender Passage aus Büchners *Hessischen Landboten* zu halten:

Im Großherzogtum Hessen sind 718,373 Einwohner, die geben an den Staat jährlich 6,363,364 Gulden, als

1) Direkte Steuern	2,128,131 fl.
2) Indirekte Steuern	2,478,264 fl.
3) Domänen	1,547,394 fl.
4) Regalien	46,938 fl.
5) Geldstrafen	98,511 fl.
6) Verschiedene Quellen	64,198 fl.
<hr/>	
	6,363,363 fl.

Natürlich konnte Büchner diese Zahlen so genau nicht kennen. Wo er sie her hat, wissen wir nicht. Vermutlich hat er nur die ungefähre Größenordnung gewusst und den Rest als literarisches Stilmittel frei dazuerfunden. Denn wie lahm klingt doch der Satz: »Die Bürger zahlen jährlich rund 2 Millionen Taler direkte Steuern«? Bei den 2 128 131 Talern dage-