

Walter Krämer

So lügt man mit Statistik

Piper München Zürich

Von Walter Krämer liegen in der Serie Piper außerdem vor:

Denkste! (2443)

Lexikon der populären Irrtümer (2446)

Lexikon der populären Listen (2591)

Das neue Lexikon der populären Irrtümer (2797)

Ungekürzte Taschenbuchausgabe

Piper Verlag GmbH, München

1.Auflage September 2000

2.Auflage Juli 2001

© 1991,1997 Campus Verlag GmbH, Frankfurt/Main

Umschlag: Büro Hamburg

Stefanie Oberbeck, Katrin Hoffmann

Umschlagabbildung: Paul Vismara / ZEFA

Foto Umschlagrückseite: Ingo Rappers

Satz: Friedrich Pustet, Regensburg

Druck und Bindung: Clausen & Bosse, Leck

Printed in Germany ISBN 3-492-23038-5

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.

Goethe, Gespräche mit Eckermann

Inhalt

Vorwort.....	9
1. Die Illusion der Präzision.....	15
2. Die betrügerische Basis.....	27
3. Vorsicht Kurve!	37
4. Protzige Prozent.....	51
5. Manipulierte Mittelwerte	61
6. Trügerische Trends	73
7. Der synthetische Superlativ	87
8. Die vorsortierte Stichprobe	97
9. Das frisierte Piktogramm	111
10. Wie es in den Wald hineinschallt.....	121
11. Statistisch oder wirklich arbeitslos?.....	137
12. Arm und reich - alles gleich?	145
13. Allerlei Kartentricks.....	155
14. Korrelation kontra Kausalität.....	165
15. Fluglärm erzeugt AIDS.....	177
16. Mythos Sozialprodukt.....	187
17. Welchen Zahlen kann man noch vertrauen?	197

Vorwort

*»Get your facts first, and then
you am 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 solange sie lebten. Die Arbeitslosen werden je nach Quelle weniger und mehr, die Krebsgefahr und das Ozonloch nehmen zu, ein Drittel aller Menschen leben unter dem Existenzminimum, ein Bundesbürger verzehrt im Jahr 5,8 Liter Speiseeis, Musiker leben länger als andere, Landluft ist gesund, Landluft ist ungesund, im Jahr 2020 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 60 Millionen Jahren aus, von Schweinefleisch bekommt man Herzinfarkt, Ausländer sind kriminell, Wähler der Grünen dagegen häufig impotent (kein Scherz - dies erklärt 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, so daß 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 gerne 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 zwei Dutzend 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 Wollens, 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 rechtfertigt 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 Parteigrenzen; es wird von Linken wie Rechten, Progressiven wie Konservativen gleichermaßen praktiziert. Allenfalls nimmt es mit der Gewißheit zu, mit der sich der Daten-Kosmetiker 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 Ozon-Experte 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 Manipulationsversuchen 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 verwerflich 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 Statistik-Studium. 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 Darrel 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, so daß 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 an den Universitäten Dortmund und Hannover. Andere haben meine Kollegen Götz Trenkler und Peter Pflaumer eingebracht. Herr Dr. Kroker vom Institut der Deutschen Wirtschaft in Köln hat mir freundlicherweise Material zur Samstagsarbeit zugeschickt. Bei den Graphiken und beim Redigieren des Manuskriptes haben mich Heide Abhoff und Elke Gebhardt unterstützt. Zusammen mit Clemens Tilke und Olaf Korn haben sie auch beim Verifizieren von Quellen und Zitaten mitgewirkt. Ralf Runde hat das Manuskript kontrollgelesen und mich vor diversen stilistischen Entgleisungen bewahrt. Tilmann Deutler, Monika und Michael Eberstaller, Hermann Göppl, Doris Krämer, Sonja Michels, Axel Scheffner, Johannes Schornstein, Claudia Schütze und Michael Schmidt haben zahlreiche eigene Ideen eingebracht. Benedikt Burkard vom Campus Verlag schließlich hat die Endredaktion besorgt und das Manuskript an vielen Stellen besser lesbar gemacht. Ich danke an dieser Stelle allen 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.

1. Die Illusion der Präzision

*»Der Mangel an mathematischer Bildung
gibt sich durch nichts so auffallend zu erkennen
wie durch maßlose Schärfe im Zahlenrechnen.«*

Carl Friedrich Gauß

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 zuviel 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 Mark«, was ähnlich klingt, wie »etwa 60 Mark, aber genau weiß ich das nicht mehr«, habe ich geschrieben »Ausgaben für Büromaterial: DM 58,24«. Das akzeptiert der Sachbearbeiter vermutlich eher als »etwa 60 Mark«; 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 Ungarisch als Muttersprache gibt. Auch diese Statistik ist sicher falsch, es sei denn, Gott der Allmächtige selber 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.

Wie alt war Methusalem

Diesen Vertrauensvorschuß, den wir krummen im Vergleich zu glatten Zahlen geben, erschleichen sie sich durch einen logischen Kurzschluß, ausgehend von der Alltagsweisheit »round numbers are always false« oder spiegelbildlich und auf Deutsch: eine korrekte Zahl ist immer krumm. Ich habe zum Beispiel in meinem ganzen Leben noch keine Telefonrechnung über exakt 100 Mark gehabt und kann mich an kein Essen mit meiner Frau oder mit Freunden erinnern, das genau 200 oder 300 Mark gekostet hätte - solche glatten Zahlen kommen praktisch niemals vor (siehe aber die Ausnahme auf der nächsten Seite).

Daraus ziehen wir dann unbewußt den Umkehrschluß, daß jede krumme Zahl korrekt sein muß, und das ist falsch. Darauf zählt der Supermarkt, der seine Marmelade für DM 2,97 statt drei Mark verkauft, genauso wie der Kfz-Experte, der vor Gericht den Bremsweg eines Autos mit 63,59 Meter angibt, oder wie der Angler, der uns seinen letzten Fang beschreibt. Nie von Dutzenden von Hechten prahlen - 23 waren es, und einer - gottverdammich - ist vom Haken wieder abgerutscht.

Dieser Zahlentrick ist so alt wie die Zahlen selbst. Adam, so schreibt die Bibel, wurde 930 Jahre alt, sein Sohn Set 912, und der berühmte Metuschelach alias Methusalem lebte sogar 969 Jahre lang. Was der Verfasser der Genesis

METZGEREI SACK
PFINZTALSTR. 13

A00 W04 5628
1.03.95 012:03 B16

K: DM/kg DM

FLEISCHWURST
24,7% FETTGEHALT
0,202 18,80 3,80

TH. ROTWURST
14% FETTGEHALT
0,122 23,80 4,33

ZIGEUNER-KASSELER
0,104 31,80 3,31

MORTADELLA BOLOGNA
0,091 22,90 2,08

HAUSMACHER LEBERW.
0,187 17,80 3,33

PFEFFERBRATEN
0,096 32,80 3,15

6 ** 20,00

N.181
ES BEDIENTE SIE:
FRAU LÖFFLER

*Solche Rechnungen kommen nur ganz selten vor:
aus lauter kleinen krummen Zahlen wird eine große runde*

damit wahrscheinlich meinte, war: alle diese Männer wurden ungewöhnlich alt. Jedoch hatte er ganz klar erkannt, daß ein seriöser Sachbuchautor so nicht schreibt - man könnte ja glauben, er hätte schlampig recherchiert. Die 969 Jahre für Methusalem dagegen zeigen jedem klar: der Autor hat auch die Details im Griff.

Daher war auch Goliath nicht »ziemlich groß« oder »riesengroß«, sondern exakt sechs Ellen und eine Handbreit groß, und das heilige Zelt, das die Israeliten bei ihrem Marsch durch die Wüste mitführten, war nicht nur erstaunlich lang und breit, sondern genau achtundzwanzig Ellen lang und achtzehn Ellen breit. Die berühmte Bundeslade war zweieinhalb Ellen lang, anderthalb Ellen breit und anderthalb Ellen hoch, das dabei verarbeitete Gold wog 29 Talente und 730 Schekel, das Silber 100 Talente und 1.775 Schekel, »also die Hälfte eines Schekels für jeden von zwanzig Jahren und darüber, der zur Veranlagung vortreten mußte, also für 603.550 Mann«.

Die Erschaffung der Erde datiert der deutsche Reformator Philipp Melanchthon auf das Jahr 3963 vor Christus, der englische Theologe John Lightfoot ist sogar noch präziser: »Himmel und Erde und alles was dazugehört wurden vom dreifaltigen Gott zusammen und zur gleichen Zeit erschaffen am Sonntag, dem 21. Oktober 4004 vor Christus, 9 Uhr morgens.« Damit ist jeder Zweifel ausgeräumt.

Das Statistische Bundesamt läßt wissen, Ende des Jahres X hätten 61 Millionen, 140 Tausend und 461 Personen in der Bundesrepublik gelebt (davon 2.228.788 arbeitslos und 4.240.500 ausländischer Staatsangehörigkeit), zusammen mit 24.502.300 Schweinen, 1.779.000 Schafen und 15.806.300 Rindern auf den Bauernhöfen unserer Republik. Wir lesen, daß es im Jahr Y insgesamt 667.259 unbeschränkt vermögenssteuerpflichtige Bundesbürger mit einem Gesamtvermögen von 479.437 Millionen Mark gegeben hat, daß im Jahr Z 12.828 Bundesbürger an Lungenentzündung gestorben sind, daß ein Angestellter in der Energiewirtschaft durchschnittlich 59.826 Mark, ein Arbei-

ter im Bergbau aber nur 44.744 Mark im Jahr verdient, daß es auf deutschen Campingplätzen insgesamt 435 474 Übernachtungen von Gästen aus Dänemark und 325519 Verkehrsunfälle mit Personenschaden gegeben hat, oder daß ein Bundesbürger pro Jahr 11,2 kg Geflügelfleisch verzehrt.

<u>CIVILIANS</u>				
(a)	<u>World War I</u>	-	Net known	
(b)	<u>World War II</u>			
	<u>Allied</u>			
	United Kingdom	60,595
	Belgium	90,000
	China	An enormous number
	Denmark	Unknown
	France	152,000
	Netherlands	242,000
	Norway	3,638
	U.S.S.R.	6,000,000
				6,548,233
				6,548,233
	<u>Enemy</u>			
	Germany	800,000
	Austria	125,000
	Italy	180,000
	Japan	600,000
	Poland	5,000,000
	Yugoslavia	Large number
				6,705,000
				6,705,000

Präzision durch Addition: Die wahre Zahl der Opfer kennt der Liebe Gott allein (aus Fighting with Figures, London 1995, einer Sammlung englischer Kriegsstatistiken des Zweiten Weltkriegs)

Woher will man das so exakt wissen? Die Wohnbevölkerung der Bundesrepublik ist selbst unmittelbar nach Volkszählungen nur ungenau bekannt, zwischen zwei Volkszählungen erst recht. Von den acht Ziffern der Bevölkerungszahl ist bestenfalls auf die ersten zwei Verlaß, und oft nicht einmal das. Noch schlimmer ist das Meßproblem bei Teilmengen der Bevölkerung wie bei Ausländern oder Arbeitslosen, worauf wir später in einem eigenen Kapitel nochmals zurückkommen - durch minimale Änderungen der Definition verschieben wir deren Zahl mühelos um mehrere Hunderttausend in jede Richtung, die uns paßt. Viehzählungen sind aus verschiedenen Gründen - darüber gibt es ganze Bücher - notorisch ungenau. Die Vermögenssteuerstatistik ist aus anderen Gründen eine der unzuverlässigsten Erhebungen überhaupt, und die Todesursachenstatistik ist auch nicht viel genauer (bei pathologischen Nachuntersuchungen stellt sich regelmäßig ein Drittel der offiziellen Todesursachen als falsch heraus). Bei den Einkommen aus abhängiger Beschäftigung ist unklar, was und was nicht eigentlich dazugehört (ist etwa das Kantinenessen des Arbeiters oder der Dienstwagen des Angestellten mitgezählt?). Nur Gott im Himmel weiß, wie viele Dänen ohne Meldeschein auf deutschen Campingplätzen übernachten und sich so der amtlichen Statistik entziehen, und die 11,2 kg Geflügelfleisch pro Bundesbürger sind ebenfalls dubios, weil z. B. völlig offenbleibt, wieviel davon weggeworfen bzw. von Hunden oder Katzen aufgeessen worden ist.

Trotzdem kommen diese Zahlen mit viel Autorität daher. Aber diese ist nur aufgemalt, beim leichtesten Kratzen am Lack der hinteren Ziffern kommt die ganze Hochstapellei ans Tageslicht.

Wenn wir also in der Zeitschrift *Nature* lesen, daß es in England 30.946 Prostituierte mit zusammen 4.641.900 Sexualkontakten jährlich gibt, so sollten wir das nicht zu wörtlich nehmen. Laut *Kicker Sportmagazin* haben die Olympischen Spiele in Barcelona 1992 alles in allem 409.196.600 Mark gekostet, aber jede andere Zahl zwischen zwei und vier Milliarden wäre genauso richtig. Und das gilt für die ganze Latte moderner Präzisionsziffern. Ein Keksfabrikant wirbt mit der Meldung, pro Monat würden in den USA 59.080.165 Portionen seiner Backwaren gegessen. Nach dem *World-Factbook* des amerikanischen CIA lebten am 21. März 1995 in der Volksrepublik China 1.127.519.327 Menschen. Laut *Bild-Zeitung* arbeitet eine typische Ehefrau pro Tag insgesamt eine Stunde, 50 Minuten und 13 Sekunden nur für ihren Mann (darunter 4 Minuten Hemden bügeln, 2 Minuten 30 Sekunden Bett machen, 1 Minute Barthaare aus dem Ausguß fischen und 15 Sekunden Klobrille schließen). Im September 1992 wurden nach Zählung des Statistischen Bundesamtes 2.384 Gartenschirme aus Polen nach Deutschland eingeführt, dafür 934 gebrauchte Klaviere von Deutschland nach England exportiert. Der Gesamtumsatz an Aktien und Renten an der Frankfurter Börse belief sich am 27. Oktober 1994 auf 17.903.906.077 Mark (und 89 Pfennig). Laut der Zeitschrift *Wachtturm* der Zeugen Jehovas vom 1. Januar 1996 haben im Jahr 1995 zusammen 244.591 Argentinier an Gedächtnismählern dieser Religionsgemeinschaft teilgenommen (verglichen mit 287.321 Deutschen, 21 Libyern, 2.262 Makedoniern und 3.109 Zyprioten). Die Gesamtkosten des Ersten Weltkriegs beliefen sich für alle kriegführenden Parteien zusammen auf 186.333.637.097 Dollar. Im Jahr 1992 wurden in Deutsch-

land 523.253 Fahrräder gestohlen (so der ADAC), und so weiter - alle diese Zahlen sind eigentlich nur grobe, mit Phantasieziffern garnierte Schätzungen.



59,080,165.

The number of packets of our Keebler Cookies they munched through every month in the States last year, thus helping us to achieve a 23% return to shareholders, on average, over the past 10 years.

*So viele Packungen Keebler-Cookies essen die Amerikaner
jeden Monat*

Teile und herrsche

Solche Hochstapelei muß nicht immer böse Absicht sein. Unsere amtlichen Statistiker in Wiesbaden z.B. sind gesetzlich angehalten, die Daten so zu publizieren wie sie sie erheben; sie dürfen gar nicht runden, so gern sie es vielleicht auch täten.

Andere krumme Zahlen werden erst durch das Addieren oder Dividieren krumm; sie stapeln quasi wider Willen hoch, wie jener Bibelschüler, der auf die Frage nach dem Alter der Welt mit »4.000 Jahre und eine Woche« geantwortet haben soll.

»Woher weißt du das denn so genau?«

»Na ja, letzte Woche haben wir das so gelernt, und da war die Welt 4.000 Jahre alt.«

Wenn wir also unser Vermögen zählen, und rechnen

Haus: 600.000,00 Mark

Bargeld: 3.934,58 Mark

zusammen: 603.934,58 Mark

so ist diese präzise Endsumme nur scheinbar präzise. Der Wert des Hauses ist nur bis auf, sagen wir, +/- 50.000 Mark bekannt, genauer kann auch die Endsumme nicht sein.

Die gleiche Scheinpräzision erzeugt die Division. Z.B. lese ich im *Handelsblatt*, daß ein bundesdeutscher Diplom-Physiker ein durchschnittliches Anfangsgehalt von 66.667 Mark im Jahr bezieht. Aber auch das ist zu präzise, um wahr zu sein, und in diesem Fall verrät die Zahl auch, wo sie hergekommen ist - ich will geteert und gefedert sein, wenn sie anders als durch Division von 200.000 durch 3 entstanden ist. Vermutlich haben drei der befragten Firmen einen Physiker eingestellt und sich vage erinnert, daß deren Startgehalt rund 70.000 (zweimal) bzw. 60.000 Mark gewesen ist.

Wenn wir also im Radio hören, daß ein Reisbauer in Bangladesch pro Jahr exakt 98 Mark und 13 Pfennige verdient, so suggeriert diese Zahl ganz ohne böse Absicht eine Recherche auf Heller und Pfennig, die nie stattgefunden hat. Vermutlich hat man nur das grob auf 2 Milliarden Taka (die Landeswährung von Bangladesch) geschätzte Volkseinkommen auf die rund 91 Millionen Einwohner umgelegt und dann mit dem aktuellen DM-Taka-Wechselkurs umgerechnet. Von diesen Zutaten ist nur der Wechselkurs exakt; das Sozialprodukt und die Bevölkerungszahl sind wilde Schätzungen. Wirft man aber alle Zutaten in einen Topf, kommt wieder eine (vermeintlich) exakte Zahl heraus.

So kommen auch die verblüffend exakten Angaben über unseren Konsum an Nahrungsmitteln zustande. Wenn man unseren Datensammlern in Wiesbaden glauben darf, verbraucht etwa ein 2-Personen-Rentnerhaushalt im Monat 8,3 Liter Milch, 8.540 Gramm Kartoffeln, 34 Gramm Tee, 1.027 Gramm Bohnenkaffee, 1.286 Gramm Rind- und 2.098 Gramm Schweinefleisch. Auch hier wurden die groben Globalzahlen erst durch Division mit der Zahl der Haushalte »exaktifiziert«.

Anderswo lese ich, daß von allen Völkern dieser Welt die Luxemburger mit 3.713 Kalorien pro Kopf und Tag die größten Vielfraße sind. Auch diese Statistik kann nur mittels der Grob-rein-fein-raus-Methode geboren worden sein. Nach ihr belegt etwa Germany (West) mit 3.381 Kalorien pro Kopf und Tag den 19. Platz, hinter Griechenland mit 3.400 und vor Kanada mit 3.374 Kalorien pro Kopf und Tag, aber diese Zahlen sind trotz aller vermeintlichen Exaktheit so ungenau, daß für die Bundesrepublik kein Rang von 1 bis 30 ausgeschlossen ist.

Bei Ländern wie Luxemburg kommt als weitere Fehlerquelle noch der Konsum durch Touristen dazu. Die Statistiker setzen hier nämlich oft Verkauf = Konsum. Von den Luxemburger 3.713 Kalorien pro Kopf und Tag landen aber sicher mehrere hundert in Magen und Darm von Ausländern. Daher führte Luxemburg auch zeitweilig die internationale Liga der Zigarettenraucher an, nicht, weil Luxemburger unmäßige Nikotinanhänger sind, sondern wegen der preiswerten Zigaretten und der Großeinkäufe der Ausländer. So könnten wir auch schließen, daß eine Familie im Schwarzwald durchschnittlich 20 Kuckucksuhren hat.

Professionelle Datensammler kennen natürlich solche Fehlermargen ganz genau. Es heißt dann, Partei X be-

käme, wäre nächsten Sonntag Wahl, 38 Prozent »plus/minus 1 Prozent«. Mit anderen Worten, der Wähleranteil liegt zwischen 37 und 39 Prozent, und das auch nicht sicher, sondern nur mit einer gewissen, wenn auch erheblichen Wahrscheinlichkeit. Da wir aber ungern Kleingedrucktes lesen, ziehen wir oft zu Unrecht den Hut vor Zahlen, die hinter einer dicken Schminke unnötiger Ziffern weit weniger exakt sind, als sie zunächst scheinen.

Literatur

Der Klassiker zu Schein und Sein im Wirtschaftsleben ist O. Morgenstern: *On the Accuracy of Economic Observations*, Princeton 1950 (2. Auflage Princeton 1963; deutsche Übersetzung: *Über die Genauigkeit wirtschaftlicher Beobachtungen*, Wien 1965). Siehe insbesondere Kapitel 1. III: »Scheingenauigkeit«. Auf den Polarforscher Peary wurde ich aufmerksam durch P. Wallich: »Polar Heat. The Argument Continues over an Explorer's Good Name«, *The Scientific American*, März 1990, S. 9-10. Die ganze Debatte wird ausführlich aufgerollt in einer von der Foundation for the Promotion of Navigation herausgegebenen Studie: *Robert E. Peary and the North Pole: A Report to the National Geographic Society*, Rockville 1989. Siehe auch die Besprechung dieses Buches durch D. Fordham in *Nature* 344, 26.4.1990, S. 902. Über Stolpersteine beim Schweinezählen informieren ausführlich H. Strecker, R. Wiegert und J. Peeters: *Messung der Antwortvariabilität auf Grund von Erhebungsmodellen mit Wiederholungszählungen*, Göttingen 1983. Der Kalorienverbrauch der Völker dieser Welt ist nachzulesen in C. L. Taylor und D. Jodice: *World Handbook of Political and Social Indicators*, Vol. 1 (3. Aufl.), New Haven 1983, S. 156-159, die Statistik der Kriegsoffer ist aus *Fighting with Figures*, London 1995, und zur Verlässlichkeit der Todesursachenstatistik siehe L. Jahn u. a.: *Studie zur Verbesserung der Validität und Reliabilität der amtlichen Todesursachenstatistik*, Baden-Baden 1995.

2. Die betrügerische Basis

»Dann hätte ich gerne noch ein Bier.«

Reaktion eines Alkoholikers auf die Information, daß 70 Prozent aller Verkehrsunfälle von nicht alkoholisierten Autofahrern verursacht werden

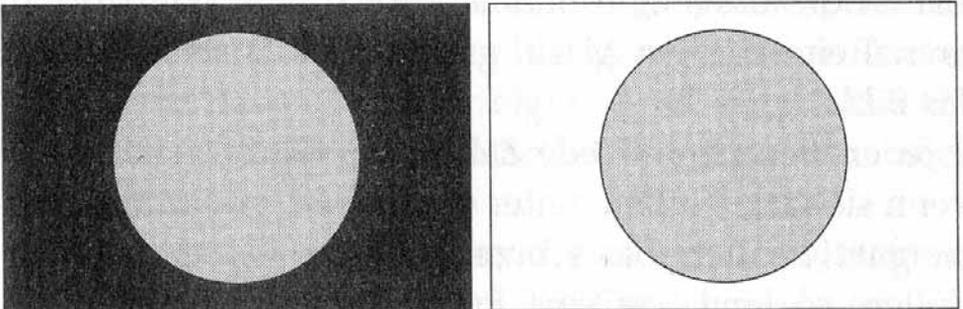
Ein Einzelhändler kauft eine Ware für 100 DM ein und schlägt sie für 200 DM wieder los. Wieviel Prozent macht seine Handelsspanne aus?

»Eine Unverschämtheit!« sagen wir als Kunde. »Ein Aufschlag von glatt 100 Prozent!«

»So schlimm ist das nun auch wieder nicht«, sagt der Händler. »50 Prozent Verdienst sind wirklich nicht zuviel.«

Offenbar haben hier beide recht. 100 DM sind 100 Prozent von 100 Mark und 50 Prozent von 200 Mark. Trotzdem ist der Eindruck je nach der Basis durchaus ein anderer.

Ein und dieselbe Sache sieht sehr verschieden aus, je nachdem, womit man sie vergleicht. Betrachten wir etwa im folgenden Diagramm die Farbe grau. Vor schwarzem Hintergrund erscheint sie eher weiß, vor weißem Hintergrund dagegen eher schwarz:



Genauso die Statistik: Wer kennt nicht die legendäre Erfolgsmeldung, abwechselnd der *Prawda*, der *Bild-Zeitung* und dem *Neuen Deutschland* zugeschrieben, zum Ausgang eines Sportwettkampfs mit zwei Mannschaften: »Unsere Jungens ehrenvoller Zweiter!« sagt die *Prawda*-Version. »Amerikaner Vorletzter.« Ein Rabatt für Barzahler lockt die Kunden an, eine Gebühr für Scheckbenutzer schreckt sie ab. Ein Freibetrag für Kinder freut die Steuerzahler sehr, eine Extrasteuer für kinderlose Arbeitnehmer regt die Steuerzahler auf. Eine Operation, bei der 90 Prozent aller Patienten überleben, läßt uns fröhlich in die Zukunft blicken, eine Operation, bei der 10 Prozent der Patienten sterben, treibt uns den Angstschweiß auf die Stirn.

Ein und dasselbe Glas ist sowohl halb voll wie auch halb leer, und ein und dieselbe Zahl ist sowohl groß wie klein, je nachdem, womit man sie vergleicht. Wenn etwa die Neuverschuldung des Bundes in einem Jahr 0,049 Billionen Mark beträgt, so sagen wir vielleicht: »Was für ein Pappentstiel!« Wenn wir dagegen hören, daß der Bund 49 Milliarden DM neue Schulden macht, fragen wir besorgt, wo das noch enden soll. Und blankes Entsetzen fällt uns angesichts einer Schuldenlast von 49.000 Millionen DM an.

In Wahrheit ist die Verschuldung aber in allen Fällen gleich. 49.000 Millionen DM ist das gleiche wie 49 Milliarden oder 0,049 Billionen DM, klingt jedoch nach mehr. Einmal ist die Basis eine Billion, einmal eine Milliarde und einmal eine Million Mark, und je nach Basis ändert sich das Bild.

Jeder Mensch und jede Zahl werden groß, so sagt man, wenn sie viele Nullen hinter sich haben, und mittels einer geeignet kleinen Basis erzeugen wir ein Gefolge von Nullen, so lang wie eine Fronleichnamsprozession. Die

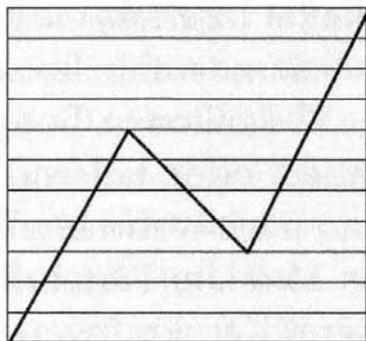
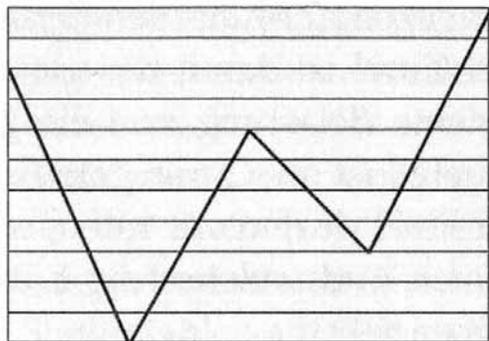
Radioaktivität eines Stoffes, d. h. die Zahl der Atomkernumwandlungen pro Sekunde, wird auf zwei Arten ausgedrückt, in Curie (Ci) und Becquerel (Bq). Ein Becquerel bedeutet 1 Zerfall, ein Curie dagegen 37 Milliarden Zerfälle pro Sekunde. Ein Tausendstel Curie ist damit das gleiche wie 37 Millionen Becquerel, eine Belastung von einigen hundert oder tausend Becquerel ist in Curie gemessen kaum noch wahrnehmbar. Dreimal dürfen wir raten, welches Maß die Partei der Grünen und welches die Lobby unserer Kernkraftwerksbetreiber liebt...

... o Täler weit, o Höhen ...

Ganz allgemein ist Vorsicht immer dann geboten, wenn eine Basis nicht von außen vorgegeben ist. Das gilt besonders beim Vergleich vom Gestern mit dem Heute. Je nachdem, was wir mit »gestern« meinen, ist der Vergleich mal mehr und mal weniger schmeichelhaft: Bei den Arbeitslosen oder bei der Inflation z. B. sieht die Jetzt-Zeit um so besser aus, je höher die Werte der Vergangenheit; deshalb läßt sich die aktuelle Regierung gerne an historischen Höchstständen messen. Bei Wirtschaftswachstum oder Handelsüberschüssen dagegen sieht die Jetzt-Zeit um so besser aus, je *kleiner* die Werte der Vergangenheit; deshalb nimmt die aktuelle Regierung hier gerne relative Flauten als Meßplatte des eigenen Erfolges. (Und umgekehrt die Opposition: sie mißt gute Wirtschaftsdaten gerne an noch besseren.)

Deshalb beginnen Kurven, die einen Aufwärtstrend beweisen sollen, gern in einem Tal. Ein Schaubild wie das

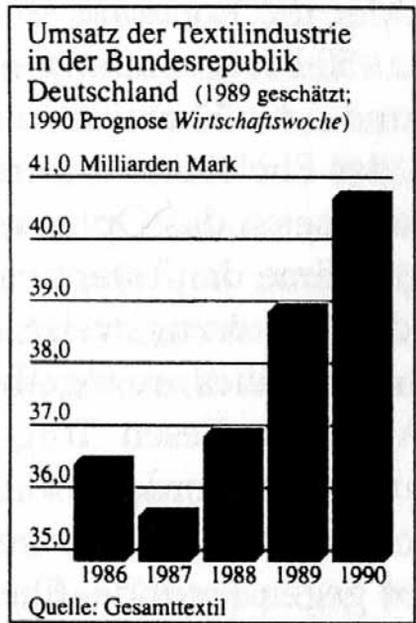
linke etwa kann man in einer positiv gemeinten Meldung lange suchen:



Wichtig für die Optik: Wo fängt die Kurve an?

Lassen wir dagegen die Kurve so wie rechts im Tal beginnen, sieht alles schon viel besser aus. Ein Beispiel: die deutsche Textilindustrie. Hier gelingt es der *Wirtschaftswoc*he, im Abstand von zehn Monaten ein negatives und ein positives Bild der selben Industrie zu zeichnen. Das linke Schaubild auf der gegenüberliegenden Seite illustriert den folgenden Text: »Schon seit Jahren sinkt in der Bundesrepublik der Anteil der Textilien am gesamten privaten Verbrauch. Das wird sich auch 1989 kaum ändern: Bei einem prognostizierten Anstieg des privaten Konsums um real 2,5 Prozent gegenüber dem Vorjahr dürfte sich der Textilverbrauch um nicht mehr als 1,5 Prozent erhöhen.« Diese eher pessimistische Einschätzung spiegelt sich auch in der Graphik wider: Sie beginnt auf einem Gipfel und endet in einem Tal.

Ganz anders dagegen die nur wenige Monate später entstandene Graphik zu genau dem gleichen Sachverhalt: sie beginnt in einem Tal und steigt dann stetig an; der Begleittext lautet: »Die bundesdeutsche Textilindustrie



Die im wesentlichen gleiche Kurve - je nach Lage der Talsohle einmal optimistisch, einmal pessimistisch

wird auch 1990 kräftig weiterwachsen. In diesem Jahr hat sie mit einem Umsatzplus von fünf bis sechs Prozent zu rechnen.«

Die statistischen Gefahren des Ehelebens

Diese Bedeutung der Basis bei Vergleichen zeigt sich am besten, wenn man dabei, ob aus Versehen oder in Betrugsabsicht, total danebengreift. Vor einigen Jahren philosophierte eine amerikanische Zeitung über Gewaltverbrechen in den USA. Die meisten geschehen zu Hause, in Küche, Wohn- und Schlafzimmer. Die Zeitung schloß daraus: man schläft nachts sicherer im Central Park.

Oder die Londoner *Times* spekuliert unter der Schlagzeile »50 Prozent aller weiblichen Mordopfer vorn eigenen Freund oder Ehemann umgebracht« über die Gefährlichkeit des Ehelebens. Nur rund 10 Prozent aller ermordeten Frauen seien das Opfer von völlig fremden Tätern, in der Regel käme der Täter aus dem Kreis der engeren Familie (und zwar vorzugsweise am Wochenende oder im Urlaub, »when families are together«).

Auch an diesen Trugschlüssen ist eine falsche Basis schuld. Daß Mordopfer in der Regel von Menschen umgebracht werden, die sie persönlich kennen, ist kein Argument gegen Freunde, Ehemänner und Bekannte, sondern eine Folge davon, daß persönliche Bekannte öfter als völlig Fremde ein Motiv und eine Gelegenheit zu einem Mord besitzen. Auch Männer werden vorzugsweise von persönlichen Bekannten ermordet, wenn sie denn ermordet werden, und daß diese Morde dann geschehen, wenn »families together« sind, liegt einfach daran, daß Mord per Telepathie nur selten möglich ist.

Deshalb geschehen auch in New Yorker Schlafzimmern und Küchen mehr Morde als im Central Park - nicht weil Wohnungen als solche ein gefährliches Pflaster wären, sondern weil wir dort weit häufiger als anderswo zugegen sind. Die Wahrscheinlichkeit, im Central Park erstochen, erschossen oder erwürgt zu werden (gegeben man übernachtet dort), ist ganz im Gegenteil weit größer, als wenn man nachts zu Hause bleibt.

Und wo wir schon einmal bei einem so unerfreulichen Thema wie Mord, Totschlag und Sterben sind: Auch die vermeintliche größere Selbstmordgefährdung Jugendlicher im Vergleich zu alten Menschen beruht auf einer falschen Vergleichsbasis. Unter dem Titel »Im Alter wirst

du glücklicher« argumentierte eine deutsche Zeitung einmal wie folgt: Der Anteil der Selbstmorde an allen Todesfällen ist bei Jugendlichen unter 20 mit 25 Prozent am größten, verglichen mit zehn Prozent bei 30- bis 40-jährigen und weniger als zwei Prozent bei den über 70-jährigen. »Es verringert sich also der Entschluß zum Selbstmord immer mehr, je weiter das Alter fortschreitet«, schreibt die Zeitung, wir würden also mit wachsendem Alter immer glücklicher.

In Wahrheit ist jedoch genau das Gegenteil der Fall. Die Selbstmorde steigen mit höherem Alter an, von weniger als fünf pro hunderttausend in der Gruppe der unter 20-jährigen bis auf fast fünfzig pro hunderttausend bei den über 70-jährigen. Je älter wir werden, desto eher scheiden wir aus freien Stücken aus dem Leben, und zwar zu allen Zeiten und in allen Ländern. Daß dennoch die Selbstmorde gerade bei Jugendlichen eine solch prominente Rolle spielen, liegt allein daran, daß Jugendliche eben generell nur selten sterben. Sie haben selten Krebs und Kreislaufleiden, sie haben keine Altersschwäche und keinen Alzheimer, und auch Schlaganfälle oder Leberschäden kommen unter Jugendlichen nicht sehr häufig vor. Mit anderen Worten, in diesem Alter sind Unfall, Mord und Selbstmord fast die einzigen Todesursachen, die noch übrig bleiben, so daß der hohe Anteil an Selbstmördern unter den Verstorbenen in dieser Altersklasse überhaupt nicht überrascht.

Ich erinnere mich an einen Artikel in einem Reisemagazin: Autofahrer führen kurz vor dem Ziel nicht mehr so konzentriert wie sonst, bei Urlaubsreisen sei deshalb vor allem auf den letzten Kilometern aufzupassen. Grundlage war eine Statistik, wonach die meisten Unfälle in einem

Radius von 30 km um den Wohnort des Verursachers geschehen. Der Gedanke, daß Autofahrer vor allem deswegen so oft in der Nähe ihres Wohnortes verunglücken, weil sie dort auch öfter als anderswo herumfahren, kam dem Artikelschreiber nicht.

Oder die »Gesellschaft zur Förderung der Freizeitwissenschaften« stellt eine Untersuchung mit dem Fazit vor, daß »Freizeit ein günstiger Nährboden für Kriminalität sein kann«. Zechprellerei, Ladendiebstahl, Schwarzfahren, aber auch Mord und Totschlag fänden vorzugsweise in des Täters Freizeit statt (gleicher Fehler wie die *Times*: wann denn sonst!). Dann wieder meint der ADAC - weil die meisten Verkehrsunfälle bei moderaten Geschwindigkeiten geschehen -, daß schnelles Autofahren nicht gefährlich sei, oder glauben andere, Autofahren wäre tagsüber gefährlicher als nachts (denn 70 Prozent aller Verkehrsunfälle geschehen tags).

Der Fehler ist immer der gleiche. Mit derselben Logik könnten wir auch schließen, daß Krankenhäuser lebensgefährlich sind (über die Hälfte aller Bundesbürger sterben dort), daß öffentliche Schulen Kriminelle produzieren (fast alle Insassen bundesdeutscher Strafanstalten sind auf einer öffentlichen Schule gewesen), oder daß man vom Zeitunglesen Haarausfall bekommt (welcher von diesem Malheur Betroffene hat nicht vorher Zeitungen gelesen?).

Bei Tempo 200 geschehen vor allem deshalb weniger Unfälle als bei Tempo 50, weil kaum jemand so schnell fährt. Bei Tempo 400 wurde meines Wissens bisher noch kein einziger Unfall auf deutschen Autobahnen registriert. Aus dem gleichen Grund geschehen tagsüber mehr Unfälle als nachts, obwohl Fahren bei Dunkelheit sicher

gefährlicher ist - mangels Verkehrsaufkommen gibt es nachts viel weniger Gelegenheit.

Wenn im *Stern* zu lesen ist: »Fußballspieler sind ... die reinsten Bruchpiloten. Sie verursachen fast die Hälfte der jährlich rund eine Million Sportunfälle«, so stimmt das einfach nicht. Fußballspieler verursachen vor allem deswegen mehr Unfälle als andere, weil es mehr Fußballspieler gibt. Die größten Bruchpiloten sind eher die alpinen Skifahrer, die mit 8 Prozent der Unfälle weit mehr Schaden anrichten, als es ihrem Anteil an den Sportlern der Nation entspricht.

Bei den Skiunfällen selbst regiert die gleiche Konfusion. »Über 50 Prozent der Verunglückten, die im Kantonsspital Chur behandelt werden, kommen aus der Bundesrepublik«, schreibt der *Spiegel*; er meint damit, die Deutschen rasten wilder auf den Schweizer Pisten herum als andere. Das mag stimmen oder auch nicht - ohne Anteil der Bundesbürger an den Skifahrern insgesamt sagt diese Statistik gar nichts aus. Bei 80 Prozent Urlaubsgästen aus der Bundesrepublik wäre ein Anteil von nur 50 Prozent an den Verletzten sogar ein Qualitätsmerkmal.

Genauso unbewiesen bleibt daher auch eine Theorie der *New York Times* zu den Ursachen des Alkoholismus. Demnach ist einer der Gründe, warum Männer oft zur Flasche greifen, niemand anders als die angetraute Ehefrau. Denn zwei Drittel aller Säufer sind verheiratet, und so kombinierte man, daß die Ehefrau den Drang zur Flasche fördern muß.

In Wahrheit besagt aber der Anteil der Ehemänner unter den Alkoholikern für sich allein genommen nicht das mindeste. Dazu müssen wir den Anteil der Ehemänner an der Risikogruppe (d.h. den Männern im heiratsfähigen

Alter) kennen, und sollte dieser zwei Drittel übersteigen, was ich fast vermute, deutet obige Statistik sogar ganz im Gegenteil auf eine ernüchternde Wirkung des Ehelebens hin.

Literatur

Zur Bedeutung der Verpackung bei der Wahrnehmung von Daten siehe A. Tverski und D. Kahneman: »Rational Choice and the Framing of Decisions«, *Journal of 'Business* 1986, 251-278. Auf unser kriminelles Freizeitverhalten wurde ich aufmerksam durch U. Stock: »Blau ohne Blaulicht«, *Die Zeit*, 27. 4.1990, S. 81 (eine satirische Besprechung von Hedwig Lerchenmüller-Hilse: *Freizeitwünsche und ihre illegale Verwirklichung*). Weitere Beispiele findet man in H. Swoboda, *Knaurs Buch der modernen Statistik*, München 1971 (eine wunderschöne, leider vergriffene Einführung in Sinn und Unsinn der Statistik); St. K. Campbell: *Flaws and Fallacies in Statistical Thinking*, Englewood Cliffs 1974 (besonders Kap. 9: Improper Comparisons), oder im 7. Kapitel meines Buches: *Denkste! Trugschlüsse aus der Welt des Zufalls und der Zahlen*, Campus Verlag 1995.

3. Vorsicht Kurve!

*»Wie leicht doch bildet man sich eine falsche Meinung
geblendet von dem Glanz der äußeren Erscheinung.«*

Molière

»Der Mensch ist ein Augentier«, schreibt Ludwig Reiners in seinem Lehrbuch der deutschen Stilkunst. »Das Auge ist noch lange aufnahmefähig, wenn der Verstand schon ermatet ist. Diese Aufnahmebereitschaft muß man ausnützen.«

Amen, können wir dazu nur sagen. Und Amen sagen auch die Manipulateure.

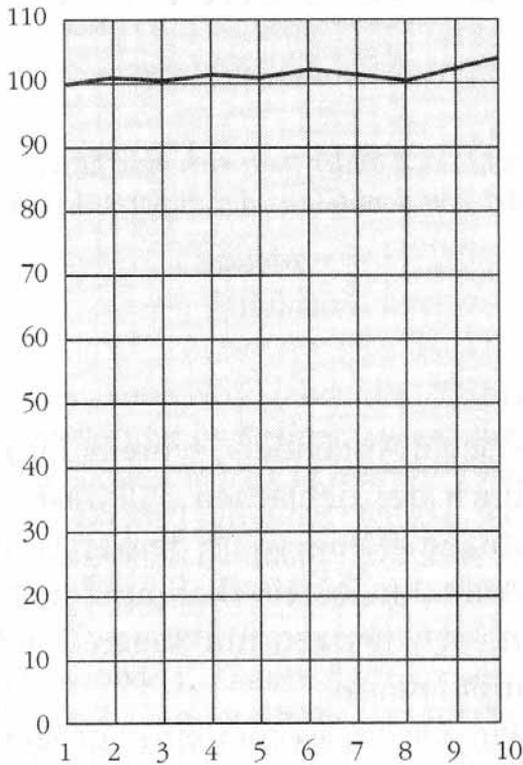
Nehmen wir den Pressechef einer größeren Firma, der eine Festschrift zum 10jährigen Firmenjubiläum herauszugeben hat. Leider war jedoch der Betriebserfolg alles andere als toll. Von Gewinn kann keine Rede sein, das Betriebsklima ist schlecht, jedes zweite Jahr schaut der Konkursrichter vorbei, mit einem Wort: der Pleitegeier setzt bereits zur Landung an. Trotzdem erwartet der Vorstand eine möglichst vorteilhafte Darstellung.

Beobachten wir also unseren Pressechef, wie er dieses Aschenputtel doch noch auf Prinzessin umfrisiert. Hier sind die Umsätze:

100 101 100,5 102 101,5 103 102,5 101,5 103 105

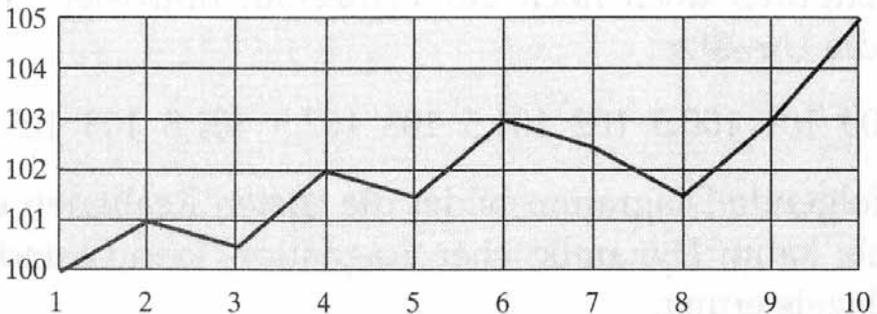
Das folgende Diagramm bildet die tristen Realitäten ehrlich ab: kaum Dynamik, eher Stagnation, kein Grund zu viel Begeisterung.

Umsatz

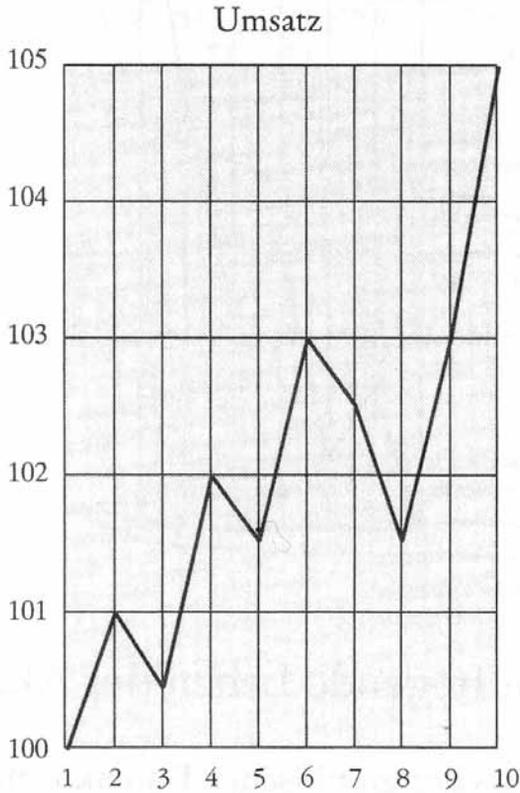


Um Platz zu sparen, schneiden viele einen Teil der senkrechten Achse ab, so wie im nächsten Diagramm. Das ist erlaubt, wenn jeder sieht, etwa an einer unterbrochenen senkrechten Achse, daß abgeschnitten worden ist. Falls dieser Hinweis fehlt: Manipulation.

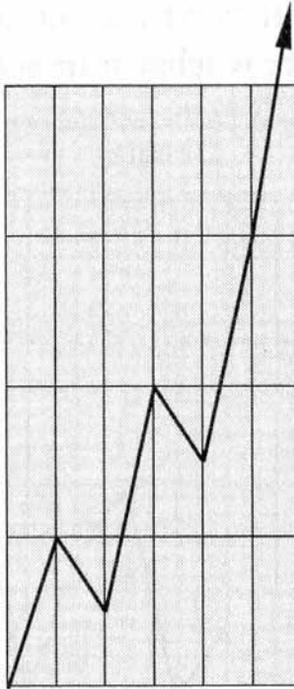
Umsatz



Im nächsten Schritt wird die senkrechte Achse wieder gestreckt; der Umsatz wächst nun schon dynamischer.

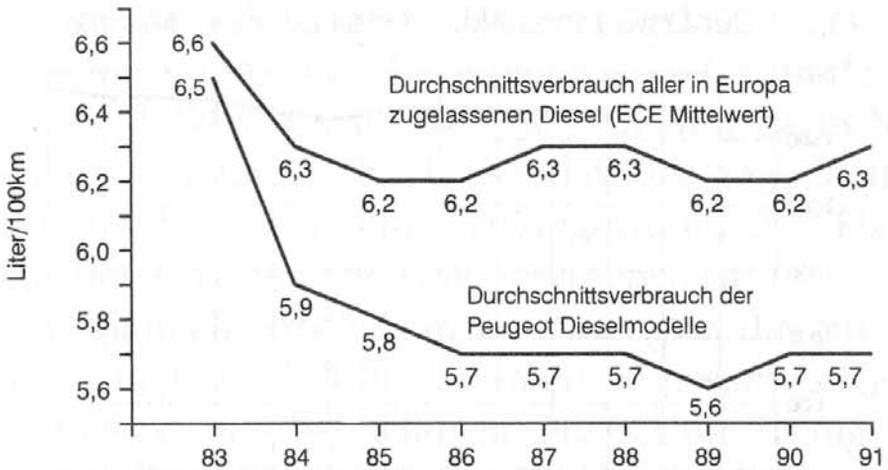


Ein bißchen stört noch der vergleichsweise flache Anstieg zwischen den Jahren 6 und 10. Aber auch hier ist Abhilfe nicht weit: Wir lassen einfach die Zwischenjahre weg! Eine Einheit auf der waagerechten Achse steht damit sowohl für ein Jahr wie für drei. Zugleich wird die Decke des Diagramms etwas gesenkt, so daß die Umsatzkurve diese am rechten Rand wie eine durch die Wolken schießende Rakete mit Macht durchbricht. Damit jeder Zweifel entfällt, wohin die Reise geht, erhält der rechte Rand der Kurve auch noch einen dicken Pfeil. Außerdem lassen wir die Zahlen an den Achsen weg: Voilà!



Das aufregende Leben der Aktien

Die Alltagspraxis der graphischen Datenkosmetik operiert meist in der Mitte zwischen dem letzten Schaubild und dem Ausgangsdiagramm. Selbst in der Werbung, wo man fast schon definitionsgemäß an Übertreibungen gewöhnt ist, findet man Graphiken wie die letzte nicht allzuoft. Die Regel sind hier eher Sünden wie die folgende, aus einer Annonce für Peugeot: Gezeigt werden soll, daß die Dieselmotoren dieses Autoproduzenten sparsamer sind als die der Konkurrenz. Und wenn man den reinen Zahlen glauben darf, sind sie das auch, wenngleich nur wenig. Deshalb, um diesen Abstand aufzublasen, schneidet die Werbung den größten Teil der senkrechten Achse ab (gegenüberliegende Seite):



Die phänomenale Sparsamkeit der Peugeot-Motoren: Schein ...

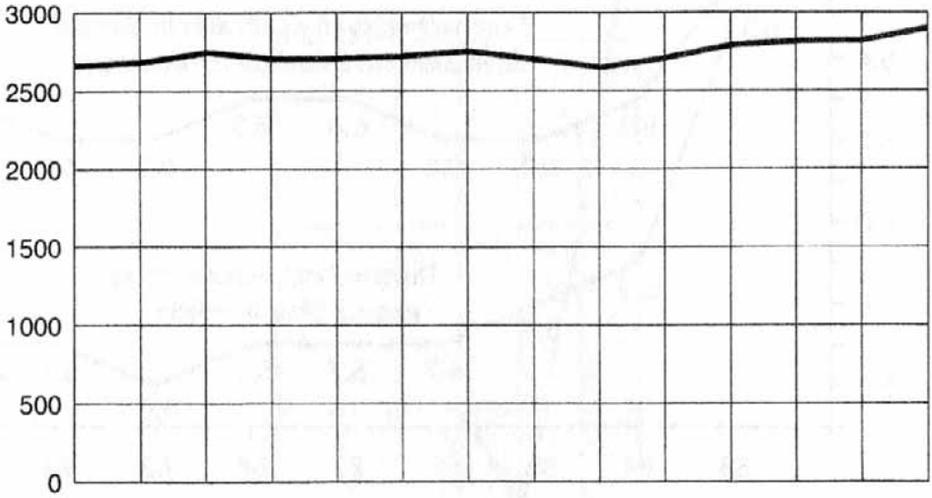


... und Sein

Solche Bilder findet man in Zeitungen und Zeitschriften sehr oft. Da wir aber in Annoncen ohnehin nicht nach der Wahrheit suchen, bleibt der Schaden meist begrenzt.

Ärgerlicher werden diese Faktenverdrehungen in redaktionellen Texten, wo man sie weniger erwartet. Die nächste Graphik, eine durch Abschneiden der senkrechten und Dehnen der waagerechten Achse dramatisierte

Der Dow-Jones-Aktienindex über vier Monate



So wird aus einer zahmen Ziege



SOURCE: DOW JONES & CO.

... ein Raging Bull

Kurve des amerikanischen Dow-Jones-Aktienindex - im Originaltext mit dem Titel »Raging Bull« oder auf deutsch »rasender Bulle« - dient dem Nachrichtenmagazin *Newsweek* zur Bestätigung der These, daß der amerikanische Aktienmarkt gerade eine außergewöhnliche Hausse durchlaufe, was aber gewaltig übertrieben ist.

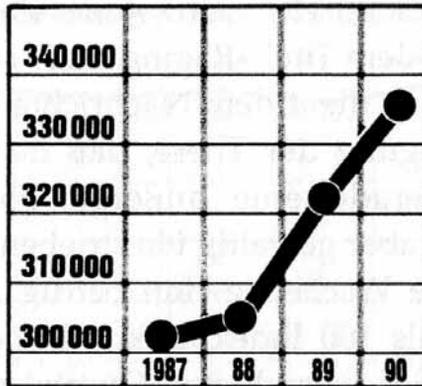
Der maximale Wochengewinn betrug in diesem Zeitraum weniger als 100 Indexpunkte; insgesamt stieg der Index in dem hier betrachteten Quartal um weniger als zehn Prozent - für die Aktionäre durchaus angenehm, aber bei weitem kein Rekord, denn soviel steigt (und fällt) ein Aktienindex ab und zu an einem Tag.

Diese Übertreibung wird sofort deutlich, wenn wir ein korrektes Diagramm ohne abgeschnittene Achse dagegenhalten - die ganze heiße Luft verpufft.

Auch die Graphik auf Seite 44, aus einem Zeitungsartikel zum Thema »Frauen und Kriminalität«, verzerrt die wahren Sachverhalte. In dem Bericht ist viel von »Frustration« und »Ausbeutung« die Rede (»Wird die Frau zur Täterin, weil die Seele trauert?«); deshalb neigten Frauen heute mehr als früher zu Gewalt.

Aber diese These wird durch die Graphik nicht bewiesen; die Aussage »geht steil nach oben« kommt nur durch das Abschneiden der senkrechten Achse zustande.

Würde diese bei 0 statt bei 300.000 beginnen, wäre sofort klar, daß der Anstieg der Delikte alles andere als ungewöhnlich ist, daß er sogar hinter dem allgemeinen Anstieg der Kriminalität zurückbleibt, daß also Frauen verglichen mit Männern eher friedlicher als rabiater werden.



Geht steil nach oben In vier Jahren stieg die Zahl der tatverdächtigen Frauen um fast 40 000 (Quelle: Bundeskriminalamt Wiesbaden)

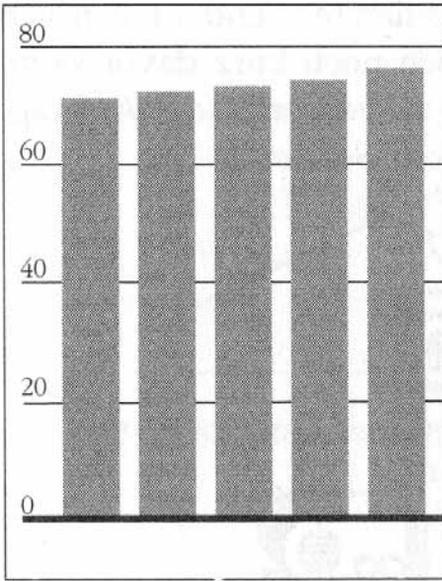
Ein Anstieg von 300.000 auf 335.000 entspricht einem Wachstum von 11,7% in vier Jahren oder einem durchschnittlichen Wachstum von weniger als 3% pro Jahr

... daß sich die Balken biegen

Statt Kurven findet man oft auch Säulen oder Balken. Die beiden Diagramme oben auf Seite 45 zum Beispiel zeigen die Kundenentwicklung einer bekannten deutschen Bank.

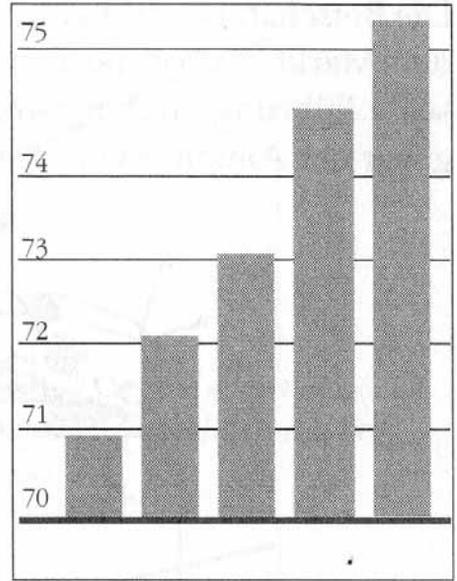
Genauso lügen die Säulendiagramme der Kredite und Kundeneinlagen im gleichen Heft. »Unsere Graphiken zeigen, daß die Einlagen von Kunden - dies sind vor allem Spareinlagen und Sparbriefe - und auch die Kredite an Kunden - vor allem Baufinanzierungen - stark angestiegen sind«, heißt es im Text, die Säulen auf Seite 45 unten sind »Beweis«.

Zuweilen werden Säulen auch beschnitten, um *Abstände* statt absolute Zahlen zu vergrößern, so wie in der folgenden Zeitungsannonce eines französischen Radiosenders.



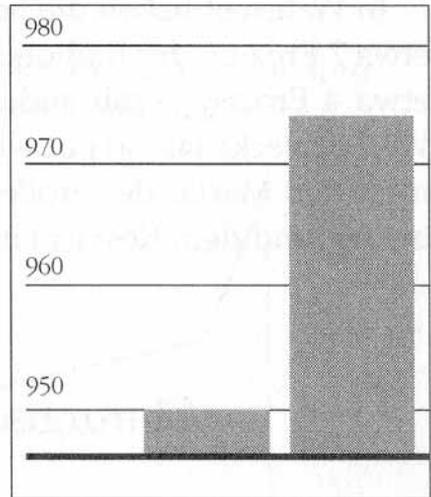
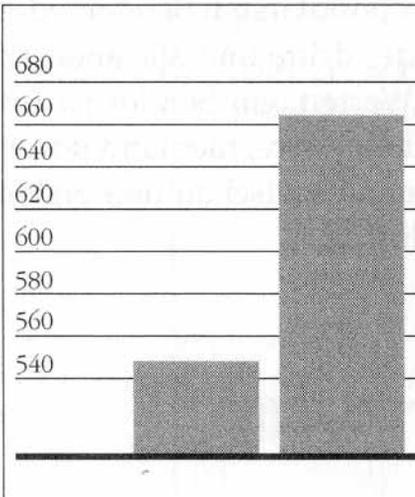
Die Kundenentwicklung der letzten fünf Jahre, wie sie wirklich ist (Anzahl in Tausend)

Kredite an Kunden in Mio. DM



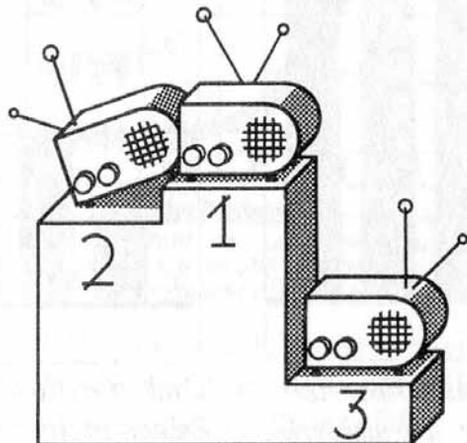
Und so stellte die Bank die Zahlen in ihrer Firmenzeitschrift dar

Kundeneinlagen in Mio. DM



Diese Graphiken lügen gleich zweifach: erstens sind die Säulen unten abgeschnitten, und zweitens zieht sich die Skala noch nach oben in die Länge

Die Botschaft ist: die beiden populärsten Sender teilen sich den Markt, wobei der erste auch noch kurz davor steht, seine Führung an den zweiten zu verlieren (der Auftraggeber der Annonce war der zweite Sender).

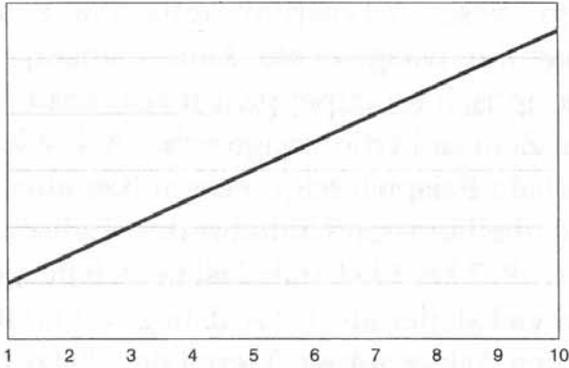


Die Graphik will den Abstand zwischen 2 und 3 vergrößern, den zwischen 1 und 2 verkleinern

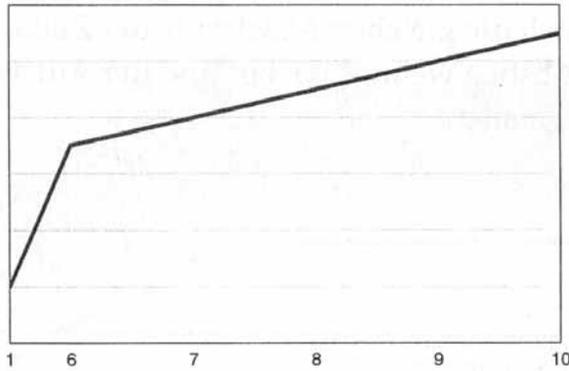
In Wahrheit haben die beiden populärsten Sender jeder etwa 7 Prozent der Radiohörer, der dritte und alle anderen etwa 4 Prozent - mit anderen Worten, ein Sender ist für Werbezwecke fast so gut wie jeder andere, niemand dominiert den Markt, der große Abstand zwischen den ersten beiden und dem Rest ist eine Illusion.

Chirurgische Eingriffe

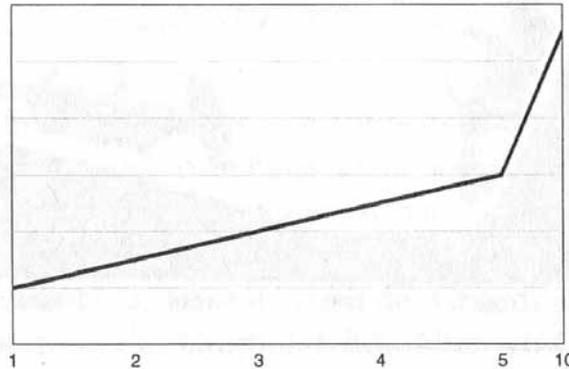
Statt der ganzen Achse lassen sich auch Teile von Achsen stauchen oder dehnen. In der nächsten Kurve bzw. besser gesagt Geraden ist der absolute Zuwachs in allen Perioden



Eine Gerade signalisiert einen konstanten absoluten Anstieg pro Periode



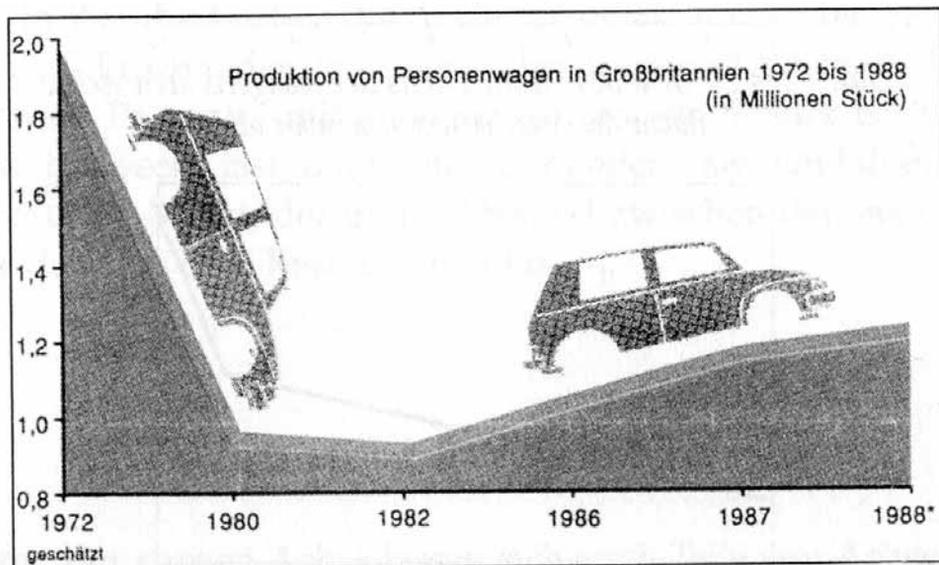
Durch das Dehnen des zweiten Teils der waagerechten Achse flacht das Wachstum scheinbar ab



Durch das Dehnen des ersten Teils der waagerechten Achse nimmt das Wachstum scheinbar zu

gleich. Um dieses Wachstum optisch abzuschwächen, macht man die waagerechte Achse länger; um dieses Wachstum optisch aufzupeppen, macht man ab dem gewünschten Zeitpunkt die waagerechte Achse kürzer.

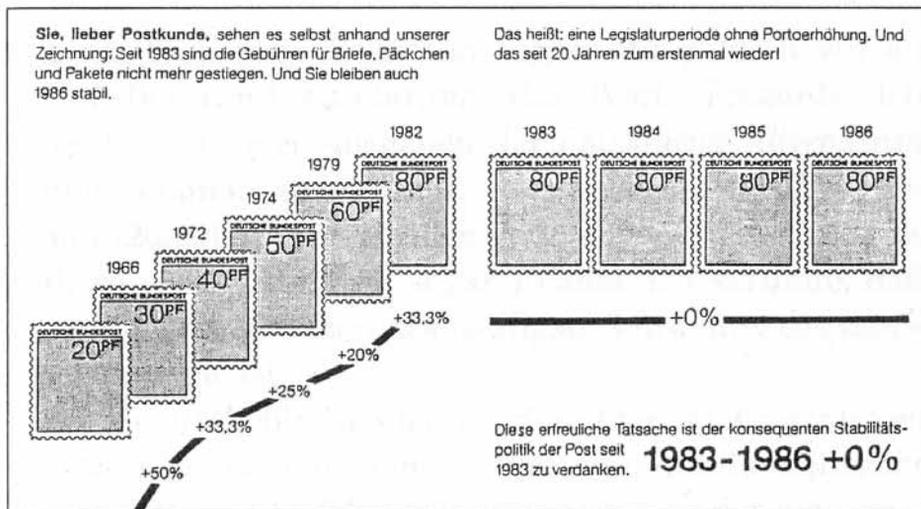
Das nächste Beispiel zeigt, vermutlich eher ungewollt verzerrend, die Berg- und Talfahrt der englischen Autoindustrie von 1972 bis 1988; dabei ist jedoch der Abstieg nur deshalb so viel steiler als der Aufstieg, weil dafür auf der waagerechten Achse nur ein Viertel des Platzes zur Verfügung steht wie für den zeitlich genauso langen Aufstieg, der dann folgt. Hätten die beiden gleich langen Zeitabschnitte auch die gleichen Abschnitte der Zeitachse belegt, wäre der Abstieg weitaus flacher und der Aufstieg weitaus steiler ausgefallen.



Die unterschiedlich steilen Straßenstücke sind ein Artefakt der waagerechten Achse

Das nächste und letzte Beispiel hat sich die gute Deutsche Bundespost zur Feier einer ungewöhnlichen Periode konstanter Preise ausgedacht. Hier ist die ungleiche Behandlung der Achsenabschnitte, die man bei der englischen Autoproduktion noch dem Ungeschick des Graphikers zu Buche schreiben könnte, ganz offenbar und mit leicht durchschaubarer Zielsetzung gewollt.

Seit 1983 stabile Gebühren



Die Preisstabilität wird durch das Dehnen der waagrechten Achse erzeugt

In Wahrheit war die Periode 1983-1986, deren Loblied hier gesungen wird, durchaus nicht extraordinär - die Postgebühren sind früher teilweise viel länger konstant geblieben - die 20-Pfennig-Briefmarke etwa regierte länger als ein Jahrzehnt (das ist wohl auch der Grund, warum das Anfangsdatum in der Graphik fehlt). Hier wird ganz einfach durch Dehnen am rechten Rand der waagerechten Achse eine Konstanz erzeugt, wo keine ist.

Literatur

Wie man Kurven- und Balkendiagramme korrekt und datenfreundlich zeichnet, findet man in vielen Büchern. Ich empfehle H. Riedwyl: *Graphische Gestaltung von Zahlenmaterial*, Bern 1975, G. Zelazny: *Wie aus Zahlen Bilder werden*, Wiesbaden 1988, und in aller Bescheidenheit auch mein eigenes Buch *So überzeugt man mit Statistik*, Campus Verlag 1994, besonders die Kapitel 3 und 4.

4. Protzige Prozent

»Ein Prozentsatz! Wahrhaftig!
Wahrhaftig, die Kerle haben prächtige Wörter
- die sind so beruhigend, so wissenschaftlich ...

Wenn man einmal sagt >Prozentsatz<,
braucht man sich wohl nicht mehr zu beunruhigen ...«

Fjodor M. Dostojewski: Schuld und Sühne

In einer typischen deutschen Tageszeitung lesen wir an einem Tag rund hundertmal das Wort »Prozent«. Ich habe das für zwei Ausgaben der *Frankfurter Allgemeinen Zeitung* einmal nachgezählt. Einmal erschien »Prozent« genau 126mal (ohne Anzeigen und Feuilleton), ein anderes Mal, an einem Samstag, sogar 135mal. Ich vermute, daß dieses Wort inzwischen das häufigste Substantiv der deutschen Sprache ist.

Wohl- und übelwollende Benutzer gleichermaßen schätzen es wegen seiner Aura von mathematischer Neutralität und Sachlichkeit. »Prozent« - oder, wie man im Norden Deutschlands vornehm sagt, »Percent« - riecht nach Kaufmannskontor und doppelter Buchführung; die Seriosität quillt nur so aus den Knopflöchern. Prozente stehen für Glaubwürdigkeit und Autorität, Prozente strahlen Gewißheit aus, Prozente zeigen, daß man rechnen kann, sie verleihen Autorität und Überlegenheit, um so mehr, und wahrscheinlich dadurch noch verstärkt, als so mancher Adressat einer modernen Prozentpredigt überhaupt nicht weiß, was eigentlich Prozente sind.

Auf die Frage »Wieviel sind 40 Prozent?« antworteten bei einer Umfrage des englischen »Office of Fair Trading«

rund die Hälfte aller befragten Bankkunden mit »einer von 40« oder »ein Viertel«. Oder um mit der *Norderneyer Badezeitung* zu sprechen: »Fuhr vor einigen Jahren noch jeder zehnte Autofahrer zu schnell, so ist es heute jeder fünfte. Doch auch fünf Prozent sind zu viele, und so wird weiterhin kontrolliert, und die Schnellfahrer haben zu zahlen.«

Mit Prozenten ist also trefflich Schaum zu schlagen. »Aufklärungsquote von 104 Prozent« meldet die Jenaer Polizei (gelesen in der *Welt*). »Mehr Mörder als die Jenaer Polizei fängt niemand: 104,8 Prozent aller Fälle von Mord, Mordversuch und Tötung klärten die Beamten 1993 auf.«

Die Lösung: die Polizisten hatten die in einem Jahr geklärten Fälle durch die im gleichen Jahr *gemeldeten* Fälle geteilt. Dann sind natürlich »Quoten« so wie oben keine Kunst, ja sogar beliebig zu verbessern: Geschieht in einem Jahr ein Mord, und drei »alte« Morde werden aufgeklärt, haben wir eine Aufklärungsquote von 300 Prozent...

Das phänomenale Wachstum der Frauenquote

Prozente sind natürlich nicht allein für Unfug da. Wenn ich weiß, 5.000 Menschen wählen Partei X, bin ich kaum klüger als zuvor. In einer Kommunalwahl mit 6.000 Wählern ist das viel, in einer Landtagswahl mit einer Million Wählern ist das gerade ein halbes Prozent, und in einer Bundestagswahl ist das so gut wie nichts. Daher interessiert bei Wahlen und Wahrscheinlichkeiten vor allem der

Prozentanteil. »Bei einer Lotterie gewinnt jedes dritte Los«, lese ich in einem Mathematikbuch für die 7. Klasse. »Bei einer anderen sind unter 800 Losen 270 Gewinne. Bei welcher Lotterie sind die Gewinnchancen größer?«

Auf der anderen Seite ist genauso wahr, daß eine Prozentzahl Informationen nicht nur liefert, sondern auch verschluckt. Links gehen zwei Zahlen ein, nämlich Zähler und Nenner, rechts kommt eine einzige heraus. Das kostet Information. Die Quotienten $\frac{1}{5}$, $\frac{7}{35}$ oder $\frac{117}{585}$ liefern alle den gleichen Prozentsatz, nämlich 20 Prozent, obwohl die Zähler (wie auch die Nenner) sehr verschieden sind. Wer einen davon zu verstecken Anlaß hat, redet daher gerne in Prozenten.

»Wir haben den Anteil unserer weiblichen Abgeordneten um 50 Prozent erhöht!« meldete nach einer Wahl der Vorsitzende der FDP. Begeistertes Klatschen der Fraktion, nach einer Meldung in der *taz*. »Na ja«, meinte einer der Abgeordneten leise, »zuerst waren es vier, jetzt sind es sechs ...«

Hier hatte die Prozent-Statistik wohl eher eine ironische Betonung weiblicher Unterrepräsentanz zum Zweck, wie vielleicht auch die notorische Info-Broschüre einer US-amerikanischen Universitätsfakultät, die einmal mit der Statistik »50% unserer Studentinnen heiraten einen Professor« um Zulauf warb. Vielleicht erhoffte man sich davon in der Tat zu einer Zeit, als Studieren für Frauen noch nicht selbstverständlich war, mehr Zuspruch seitens des weiblichen Geschlechts. Nötig war es jedenfalls. Wie eine Nachfrage ergab, waren am ganzen Fachbereich nur zwei Frauen immatrikuliert; eine davon war die Gattin des Dekans.

Dieses Verstecken kleiner Zahlen hinter imponierenden

Prozenten hat eine lange Tradition. Im alten Reichstag gab es einmal eine Anfrage zur wachsenden Unmoral in einem Dorf in Oberbayern; dort hatte die amtliche Statistik eine Zunahme der unehelichen Geburten um immerhin 200 Prozent notiert, »eine Sündenstatistik«, wie sich Altmeister Ernst Wagemann in seinem *Narrenspiegel der Statistik* erinnert, »die dadurch veranlaßt war, daß eine Frau, die dem Zivilstand der Ledigen angehörte, Drillinge geboren hatte.«

Allen Wallis und Harry Roberts berichten in *The Nature of Statistics* von den Nöten eines Weltkrieg-II-Radaroffiziers, der lange vergeblich um Verstärkung seiner Gruppe von sieben Technikern eingekommen war. Als einer davon nervlich zusammenbrach, meldete der Offizier nur lakonisch, »14 Prozent des Personals wegen Überlastung verrückt geworden«; darauf erhielt er mehr Verstärkung als er brauchen konnte.

Oder nehmen wir einen Landwirt, dessen Viehbesitz aus 57 Prozent Kühen, 14 Prozent Schweinen und 29 Prozent sonstigem Vieh besteht. Wer käme da auf die Idee, daß der stolze Agronom ganze 4 Kühe, 2 Schafe und ein Schwein sein eigen nennt? Angefangen bei den bekannten »80 Prozent der Testpersonen«, die sich für Kaffee X, Zahnpasta Y oder Rasenmäher Z entscheiden, bis hin zur hundertprozentigen Zustimmung der Belegschaft zur freiwilligen unbezahlten Sonntagsarbeit haben alle diese Prozente den gleichen Zweck: das Verstecken der Basis, auf die sich der Prozentanteil bezieht.

Besondere Vorsicht ist geboten bei 33,3 und 66,6 Prozent. Die Basis, die dadurch verborgen werden soll, hat bis zum Beweis des Gegenteils vermutlich gerade den Umfang drei.

Aber auch andere Dinge lassen sich hinter Prozenten leicht verstecken. Wenn wir zehn Personen fragen: »Welche Partei wählst du?«, und vier entscheiden sich für X, so sind das 40 Prozent. Aber niemand, egal ob er oder sie schon einmal etwas von »Stichprobenfehlern« gehört hat oder nicht, wird wohl viel Geld darauf verwetten, daß dies dem wahren Wahlergebnis nahekommt...

Fragen wir dagegen 1000 Wähler, und 400 wählen X, so glauben wir die Wahlprognose eher (weil mit der Größe der Stichprobe die Zuverlässigkeit der Daten steigt).

Jeder zweite lebt allein

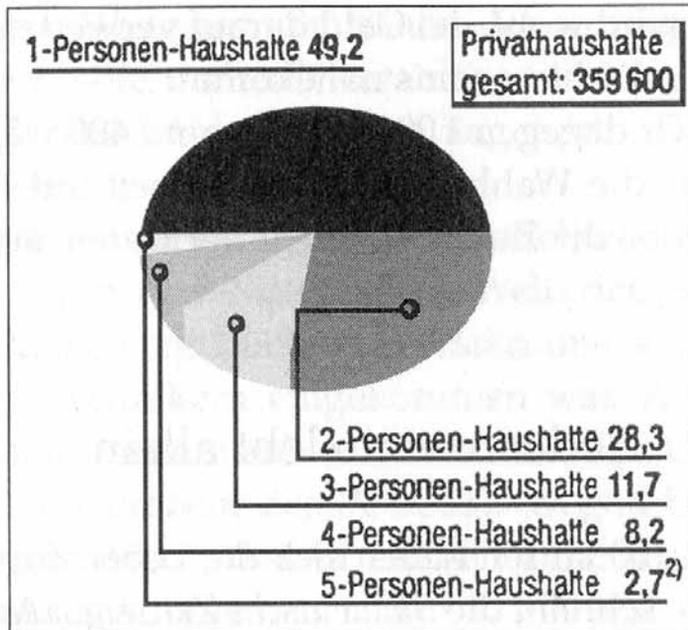
»In nahezu 100 Jahren haben sich die Lebensformen stark gewandelt«, schreibt die *Süddeutsche Zeitung*. »Anfang dieses Jahrhunderts war das Miteinander in der Großfamilie Normalität. Fast die Hälfte der Bevölkerung ... wohnte in Haushalten mit fünf und mehr Personen. Ganz anders heute: zwei Drittel der Bevölkerung leben allein oder zu zweit.« Beweis: 66 Prozent aller Haushalte bestehen aus weniger als drei Personen ...

In Frankfurt bestehen sogar 77,5 Prozent aller Haushalte aus weniger als drei Personen, 49,2 Prozent davon sind Single-Haushalte: Jeder zweite lebt allein ...

Wenn man einer Londoner Staatsanwältin glauben darf, ist dieses Verschwinden der Großfamilie nur zu begrüßen. Denn Großfamilien machen kriminell. Nur in wenigen Fälle von Jugendkriminalität, mit denen die Staatsanwältin befaßt war, kamen die Übeltäter aus Ein-Kind-Familien. Je größer die Familien, desto krimineller.

Jeder zweite lebt allein

Haushaltsgrößen in Frankfurt (Anteile in Prozent)¹⁾



1) Stand: 1994. 2) Aussagewert wegen geringer Basis eingeschränkt.
Quelle: Amt für Statistik, Wahlen und Einwohnerwesen, Frankfurt

Trotz 50% Single-Haushalten leben weit weniger als 50 Prozent der Frankfurter allein

Auch hier der gleiche Fehler beim Umrechnen von Haushalten auf Personen: Kleine Haushalte machen zwar einen großen Prozentsatz der Haushalte, aber einen weit kleineren Prozentsatz der Personen aus.

Der Wachstum-Turbo

Prozente täuschen oft auch bei Wachstumsraten; sie machen leicht aus einem lauen Lüftchen einen wilden Sturm. Nehmen wir eine Firma mit folgenden Umsätzen:

100 101 102,5

Der Umsatz wächst zunächst um 1 Prozent und dann um $1,5/101 = 1,49$ Prozent. Das ist nicht eben viel; die eine Wachstumsrate liegt so flach auf dem Bauch wie die andere, mit »Dynamik« meinen wir in der Regel etwas anderes.

Drücken wir dagegen die Ausgangsdaten statt in Wachstumsraten in *Wachstumsraten von Wachstumsraten* aus, so ändert sich das Bild. Wenn der Umsatz im ersten Jahr um 1 Prozent und im zweiten Jahr um 1,49 Prozent wächst, so wächst das *Umsatzwachstum* um stolze 49 Prozent!

Diese Rechnung ist zwar mathematisch einwandfrei, stellt aber die Realität auf den Kopf. Wenn wir von einer Steigerung des Umsatzwachstums von fast 50 Prozent erfahren, stellen wir uns dahinter kaum die obigen Ausgangsdaten vor. Viele werfen nämlich Wachstumsraten und Wachstumsraten von Wachstumsraten in einen Topf, selbst gestandene Wirtschaftsjournalisten, wie der folgende Ausschnitt aus der *Berliner Zeitung* zeigt. Unter der Überschrift »Wird das nächste Jahr weniger teuer?« heißt es da:

Preissteigerungsrate in neuen Ländern 1992 im Durchschnitt bei 12 Prozent. Die Teuerungsrate in den neuen Bundesländern könnte sich 1993 im Jahresdurchschnitt auf etwa sieben Prozent belaufen und damit gegenüber diesem Jahr nahezu halbieren.

Dies stellt das Hamburger HWWA Institut für Wirtschaftsforschung in seinem jüngsten Konjunkturbericht fest.

Mit anderen Worten, das nächste Jahr wird in Wahrheit überhaupt nicht »weniger teuer«; die Preise steigen weiterhin, nur nicht mehr so schnell.

Wachstumsraten von Wachstumsraten gehören also in den Giftschränk des Statistikers. Sie putschen schlappe Daten auf und töten positive Daten ab; mit der wahren Datenlage haben sie im allgemeinen nichts zu tun.

Außerdem leiden sie noch an einer weiteren Verwechslungsmöglichkeit, nämlich der zwischen *Prozentpunkten* des Wachstums und Wachstumsraten selbst. Wenn etwa unser Krankenkassen-Beitragsatz von 12 Prozent des Einkommens auf 13 Prozent steigt, so wächst er um einen Prozentpunkt oder um $\frac{1}{12} = 8,3$ Prozent. Dieser Unterschied ist vielen oft nicht klar.

Bei einem Ausgangs-Prozentsatz von weniger als 100 ist die Differenz in Punkten immer kleiner als in Prozent. Wer diese optisch kleiner machen will, nimmt hier gerne Punkte statt Prozent. Bei einem Ausgangssatz von mehr als 100 ist die Differenz in Punkten dagegen größer als in Prozent. Ein Anstieg des Preisindex von 150 auf 160 entspricht einem Plus von 10 Punkten oder $\frac{10}{150} = 6,6$ Prozent. Hier sind daher, falls die Änderung möglichst klein erscheinen soll, letztere sehr beliebt.

Prozente sind somit sowohl als Schlaf- wie auch als Aufputschmittel zu gebrauchen. Wie ein Fernglas, das man falsch herum vors Auge hält, machen sie kleine Dinge groß, aber auch große Dinge klein. »Ich hätte gerne 2 Milliarden Mark«, sagt der Landwirtschafts- zum Finanzminister, weil er Geld für seine Bauern braucht. Der Finanzminister erbleicht. »Jetzt stell dich nicht so an«,

drängt sein Kollege, »noch nicht einmal ein zehntel Prozent vom Sozialprodukt!«

So gesehen steht der Finanzminister wie ein Knauser da. Kein Wunder also, daß jeder, der Geld von uns haben will, dies gerne in Prozenten einer möglichst großen Summe tut.

Literatur

Die Statistik zu den mangelhaften Kenntnissen in Prozentrechnung unserer englischen Nachbarn stammt aus R. Young: »The Perils and Pitfalls of Percentages«, *The Times*, 2.5.1990. Eine Übersicht statistischer Methoden liefert J. L. Reiss: *Statistical Methods for Rates and Proportions*, New York 1973 (für Profis); für Wachstumsraten siehe auch mein Buch *Statistik verstehen*, Campus Verlag 1992.

5. Manipulierte Mittelwerte

*»Mit der ganzen Algebra ist man oftmals nur ein Narr,
wenn man nicht noch etwas anderes weiß.«*

Friedrich der Große

Jeder weiß, was ein Durchschnitt ist. Wenn Bauer A drei und Bauer B fünf Kühe hat, so hat im Durchschnitt jeder vier. Nichts ist einfacher als das.

Mit »Durchschnitt« ist hier, wie meistens in der Alltagssprache, das sogenannte »arithmetische Mittel« gemeint, die Summe der Werte geteilt durch ihre Anzahl, in unserem Falle drei plus fünf geteilt durch zwei. Es ist handlich, leicht zu berechnen, es verdichtet große Datenmassen elegant zu einer einzigen kompakten Zahl.

Leider tut es der guten Dinge dabei oft zuviel, wie folgende Version eines uralten Statistiker-Witzes zeigt (Copyright Franz Josef Strauß): »Zwei Männer sitzen im Wirtshaus, der eine verdrückt eine Kalbshaxe, der andere trinkt zwei Maß Bier. Statistisch gesehen ist das für jeden eine Maß Bier und eine halbe Haxe, aber der eine hat sich überfressen und der andere ist besoffen.«

Das arithmetische Mittel verkleistert oft eine große Ungleichheit - es schweigt sich zur Streuung um das Mittel völlig aus. Wenn es in unserem Dorf zehn Bauern gibt, von denen einer 40 Kühe hat, und alle anderen haben nichts, so hat auch im Mittel jeder vier. Für die neun Habenichtse ist das aber nur ein schwacher Trost. Offenbar macht es einen

Unterschied, *ob* sich die Werte dicht um das Mittel sammeln, oder ob sie in alle Winde streuen, aber diesen Unterschied sieht man dem Mittelwert nicht an.

Die mittlere Lufttemperatur in Plymouth/England beträgt über das Jahr gerechnet 13 Grad (Tageswerte), damit fast das gleiche wie in Minneapolis in den USA. Trotzdem ist das Klima an diesen Orten keinesfalls gleich moderat und angenehm. Das Jahresmittel von 13 Grad in Plymouth verschweigt z.B. eine bemerkenswert kleine Streuung der Temperaturen über das Jahr hinweg; selbst im Februar, dem kältesten Monat, zeigt das Thermometer immerhin noch 8, im Juli dagegen nicht mehr als 21 Grad. Frost und große Hitze gibt es nicht. »Während die relativ kühlen Sommer manche Früchte wie Aprikosen und Pfirsiche nur unter dem Schutz einer Hecke oder einer Mauer reifen lassen«, teilt uns ein Reiseführer mit, »machen die milden Winter es möglich, zahlreiche subtropische Gewächse zu akklimatisieren. Myrten und immergrüne südeuropäische Sträucher findet man fast überall im Freien.«

Davon können die Farmer um Minneapolis nur träumen, ihnen frieren im Winter die Ohren ab. Im Januar zeigt das Thermometer durchschnittlich minus 15, dafür im Sommer über 30, zuweilen sogar an die 40 Grad. Trotzdem gibt es, über das Jahr gemittelt, zu Plymouth keinen Unterschied.

Ein Mittelwert ohne Streuung ist also nur die Hälfte wert. Zu einem seriösen Durchschnitt gehört in aller Regel auch ein Maß für die Abweichung davon. Wie dieses aussieht, sei uns hier egal - bei nackten Mittelwerten ist immer Vorsicht angezeigt.

Welcher Durchschnitt darf's denn sein?

Das arithmetische Mittel, obwohl für die meisten Zwecke völlig ausreichend, ist nicht das einzige. Sein größter Konkurrent ist der *Zentralwert* oder *Median*. Damit ist, grob gesagt, die Zahl gemeint, die genauso viele linke wie rechte Nachbarn hat. Wenn die Daten wie die Orgelpfeifen vor uns stehen, ist der Median der Wert, der in der Mitte



steht. Er ist sogar noch einfacher zu finden als das arithmetische Mittel und hat auch andere Vorteile. Erstens erhält man so immer einen tatsächlich realisierten Wert; Durchschnittsfamilien mit 1,7 Kindern oder die $3\frac{1}{2}$ Sexualpartner pro Bundesbürger, mit denen Statistiker gerne das Publikum erheitern, gibt es dabei nicht. Zweitens ist immer die Hälfte aller Werte kleiner und die andere Hälfte größer als der Median, wie wir das eigentlich von jedem vernünftigen Mittelwert erwarten. (Das arithmetische Mittel kann sich zwar an diese Regel halten, muß es aber nicht. In unserem Beispieldorf etwa haben neun Bauern

weniger als das arithmetische Mittel von vier Kühen, und nur ein einziger hat mehr.) Und drittens und letztens versieht der Median auch dann noch seinen Dienst, wenn Merkmalswerte nicht mehr addier- und dividierbar sind (wie etwa Schulnoten oder Hitparadenplätze), das arithmetische Mittel also machtlos ist, denn ein einfaches Sortieren der Daten der Größe nach genügt.

In unserem Dorf mit den 10 Bauern und 40 Kühen (die aber alle einem einzigen Krösus gehören) ist der Median des Kuhbesitzes null. Neun Nullen und eine Vierzig nebeneinander aufgereiht, ergeben in der Mitte immer eine Null. Das arithmetische Mittel ist dagegen größer, nämlich vier.

Das ist kein Zufall. In diesem Beispiel sind die Merkmalswerte, wie die Statistiker sagen, »schief verteilt« (rechts-schief, ganz genau gesagt): Kleine Werte treten häufig auf, große immer seltener. Solche Verteilungen sind typisch für Merkmale wie Einkommen, Vermögen oder Grundbesitz; je weiter man nach oben steigt, desto dünner wird die Luft. Bei solchen Merkmalen liegt das arithmetische Mittel *immer* oberhalb des Medians, manchmal sogar drastisch oberhalb. Wenn wir also lesen, das Einkommen im Sultanat Brunei betrage durchschnittlich 54.000 Mark im Jahr, verglichen mit 46.000 in der Bundesrepublik, so heißt das nicht unbedingt, daß ein typischer Bundesbürger verglichen mit einem Untertan des Sultans ärmer ist. Eher ist das Gegenteil der Fall. In beiden Fällen ist nämlich mit »Durchschnitt« das arithmetische Mittel, d. h. Volkseinkommen geteilt durch Wohnbevölkerung, gemeint. Jedoch ist die Einkommensverteilung in Brunei erheblich schief als bei uns. Läßt man den Sultan, der als reichster Mensch der Erde gilt, und seine Familie weg,

sieht das Bild in Brunei schon ganz anders aus. Solche »Ausreißer«, wie die Statistiker sagen, ziehen das arithmetische Mittel an sich heran wie ein Magnet.

Dem Median dagegen sind Ausreißer egal. Wenn unser Oberbauer 400 statt 40 Kühe hat, bleibt der Median bei null. Das arithmetische Mittel dagegen gerät völlig aus dem Gleichgewicht, steigt von 4 auf 40 - ein einziger Merkmalswert hat das arithmetische Mittel völlig in der Hand.

Wenn also der Präsident der Bundesärztekammer vom mittleren Einkommen der deutschen Ärzte spricht, meint er in der Regel nicht das arithmetische Mittel, sondern den



»Sollen wir das arithmetische Mittel als durchschnittliche Körpergröße nehmen und den Gegner erschrecken, oder wollen wir ihn einlullen und nehmen den Median?«

Median. Wann immer im *Deutschen Ärzteblatt* von Geld die Rede ist, erinnert man sich gern daran, daß es außer dem arithmetischen auch noch andere Mittelwerte gibt.

Auf der anderen Seite rufen Kritiker, die meinen, deutsche Ärzte verdienten viel zu viel, gern das arithmetische Mittel in den Zeugenstand; dieses ist wie bei allen rechtschiefen Verteilungen immer größer als der von den Ärzten selbst bevorzugte Median. Ein niedergelassener Laborarzt etwa verdient nach Abzug der Praxiskosten, wenn man das arithmetische Mittel nimmt, durchschnittlich 700.000 Mark im Jahr, beim Median dagegen 200.000 DM weniger (immer noch zuviel, aber das ist ein Buch für sich).

Gewichtete Mittelwerte

Zuweilen wird auch das gewöhnliche arithmetische Mittel angewandt, wenn eigentlich das »gewogene« arithmetische Mittel zuständig wäre. Ein Imbißbuden-Pächter im Wilden Westen wurde, wenn ich einem amerikanischen Kollegen glauben darf, einmal gefragt, woraus sein Wildragout bestehe. »Naja, es ist auch Pferdefleisch dabei«, gab er ehrlich zu. Und wieviel Pferdefleisch? »Halb und halb«, sagt daraufhin der Gastronom. »Ein Kaninchen und ein Pferd.«

Das gewöhnliche arithmetische Mittel rechnet ebenso; es behandelt alle Werte gleich. Meistens ist das auch korrekt, manchmal aber nicht. Wenn in einer Firma die Frauen 20 Mark pro Stunde und die Männer 30 Mark pro Stunde verdienen, so ist der Durchschnittslohn für alle nicht notwendig 25 Mark; je nach Zahl der Männer und

der Frauen ist er einmal mehr und einmal weniger: Gibt es mehr Männer, ist der Durchschnitt größer, gibt es mehr Frauen, ist der Durchschnitt kleiner. Nur bei gleich vielen Frauen wie Männern entspricht der wahre Durchschnittslohn dem arithmetischen Mittel 25, sonst aber nicht.

Der wahre Durchschnitt ist das sogenannte »gewogene« arithmetische Mittel von 20 und 30, bei sieben Frauen und drei Männern etwa

$$0,7 \cdot 20 + 0,3 \cdot 30 = 23.$$

Die Gewichte - hier 0,7 und 0,3 - beim gewichteten arithmetischen Mittel sind meistens aus dem Kontext klar. Manchmal aber auch nicht. Angenommen etwa, ein Land erhebt an seiner Grenze Zoll. Nehmen wir weiter an, es importiere nur Autos und Lebensmittel, etwa Dosenwurst. Der Zoll auf Automobile betrage 50 und der auf Lebensmittel 10 Prozent. Wie hoch ist der durchschnittliche Zoll?

Das gewöhnliche arithmetische Mittel nützt uns hier wenig. Autos sind für den internationalen Warenaustausch wichtiger als Dosenwurst; die einzelnen Zollsätze müssen unterschiedlich gewichtet werden. Die Frage ist nur, wie?

Ein auf den ersten Blick vernünftiges Gewicht ist offenbar der wertmäßige Anteil des Produkts an der Gesamteinfuhr - so gehen intensiv gehandelte Wirtschaftsgüter stärker in den Durchschnitt ein. Angenommen, in unserem Beispiel tragen Automobile 80 und Lebensmittel 20 Prozent zum Importwert bei. Nach diesem Gewichtungsschema beträgt der durchschnittliche Zoll damit

$$0,8 \cdot 50\% + 0,2 \cdot 10\% = 42\%.$$

Dieses System hat aber einen Pferdefuß. Angenommen, die einheimische Automobilindustrie überredet die Regierung zu einer Zollerhöhung auf 200 Prozent für die ausländische Konkurrenz. Damit steigt natürlich auch der durchschnittliche Zoll. Zumindest sollte man das erwarten. Der erste Zoll bleibt gleich, der zweite steigt, also nimmt der Durchschnitt zu. Jeder vernünftige Mittelwert hat diese Eigenschaft.

Bei konstanten Gewichten stimmt das auch. Nicht notwendig aber mit unserem System. Als Konsequenz der Zollerhöhung geht nämlich wie gewünscht der Importanteil ausländischer PKWs zurück, etwa auf 10 Prozent. Damit ändern sich aber auch die Gewichte unseres Mittelwerts. Vor der Zollerhöhung hatten PKWs ein Gewicht von 0,8, nach der Zollerhöhung haben PKWs ein Gewicht von 0,1 und Lebensmittel ein Gewicht von 0,9, und damit *sinkt* der durchschnittliche Zoll; er beträgt jetzt nur noch

$$0,1 \cdot 200\% + 0,9 \cdot 10\% = 29\%!$$

In abgeschwächter Form finden wir dieses »Paradox des Zolls« auch bei unserem Preisindex für die Lebenshaltung. Auch diese berühmte Statistik ist ein gewogener arithmetischer Mittelwert, konkret: ein arithmetisches Mittel der Preisverhältnisse der Güter, aus denen sich der Warenkorb zusammensetzt: Lebensmittel, Urlaubsreisen, Postgebühren, Miete, Auto etc. Deren aktuelle Preise werden geteilt durch die Preise in der sogenannten Basisperiode, und diese Preisverhältnisse werden dann mittels eines gewichteten arithmetischen Mittels zu einem Durchschnitt kondensiert - dem Preisindex für die Lebenshaltung.

Die amtlichen Gewichte orientieren sich hierzulande

an den relativen Ausgaben für das jeweilige Produkt; sie werden periodisch an Verbrauchsveränderungen angepaßt. Wenn aber ein bestimmtes Gut, etwa Urlaubsreisen nach Acapulco, besonders stark im Preis gestiegen und daraufhin im Umsatz abgefallen ist, geht dessen Gewicht zurück, so daß regelmäßig die Inflation mit neuen Gewichten kleiner ist als mit alten. Die häufigen Klagen aus Arbeitnehmerkreisen, das lange Festhalten am gleichen Warenkorb sei ungerecht, da den Konsumenten nicht mehr angepaßt, sind also kontraproduktiv. Die Praxis des Statistischen Bundesamtes ist im Gegenteil recht arbeitnehmerfreundlich; würde in Wiesbaden mit jährlich neuen Gewichten gerechnet, wären die Preissteigerungen nicht größer, wie einige Gewerkschaftsfunktionäre glauben, sondern kleiner.

Angst vorm Fliegen

Eine weitere Konfusion betrifft die Zahl, durch die man beim arithmetischen Mittel die Merkmalsumme teilt. Beispiel Verkehrssicherheit: Womit ist man sicherer unterwegs, mit dem Flugzeug oder mit der Bahn? (Das Auto als Killer Nr. 1 lassen wir von vornherein außen vor.)

Hier gibt es zwei Antworten, eine aus dem Kopf und eine aus dem Bauch. Der Kopf sagt uns, Fliegen ist sicherer. Haben wir nicht tausendmal gelesen, in den gleichen Zeitungen und Illustrierten, die keine Katastrophe auslassen, um mit Flugzeugwracks und Leichenteilen ihre Auflage zu steigern, daß Fliegen trotzdem sicherer ist? Im Durchschnitt, so das Standardargument, kommen dadurch weniger Menschen um als durch die Eisenbahn.

Die Frage ist jedoch, wie rechnen wir den Durchschnitt aus? Wodurch teilen wir beim Bilden unseres Mittelwerts *die* Zahl der Opfer, was kommt in den Nenner unseres Bruches?

Der Standard-Nenner sind die insgesamt zurückgelegten Passagier-Kilometer. Damit erhalten wir:

Bahn:	9 Verkehrstote pro 10 Milliarden Passagierkilometer
Flugzeug:	3 Verkehrstote pro 10 Milliarden Passagier kilometer

Diese oder eine ähnliche Statistik kennen wir; sie bestätigt die bekannte Beruhigungsformel, denn so gesehen ist Bahnreisen in der Tat gefährlicher. Wenn diese Statistik stimmt, kommen in der Eisenbahn dreimal so viele Menschen ums Leben wie im Luftverkehr.

Wieso bricht uns dann beim Besteigen eines Flugzeugs der Angstschweiß aus, beim Besteigen eines Zuges aber nicht? Wie die nächste Statistik zeigt, ist das gar nicht so irrational, wie viele denken. Mit dem gleichen Recht, mit dem wir die Zahl der Verkehrstoten auf die zurückgelegten Kilometer beziehen, können wir sie nämlich auch auf die Stunden beziehen, die wir in Gefahr verbringen. Für mich selbst jedenfalls ist diese Zahl weit wichtiger - schließlich habe ich auch beim Besteigen eines Bettes keine Angst, obwohl die Wahrscheinlichkeit, darin zu sterben, fast 99 Prozent beträgt.

Um daher zu entscheiden, ob ich rationalerweise Angst haben darf oder nicht, ist die Wahrscheinlichkeit, auf den nächsten tausend Kilometern umzukommen, vielleicht gar nicht so wichtig. Viel mehr interessiert mich die Wahrscheinlichkeit, in den nächsten 10 *Minuten* (oder in der

nächsten Stunde) umzukommen. Mit anderen Worten, nicht die Passagier-Kilometer, sondern die Passagier-*Stunden* gehören in den Nenner unseres Bruchs. Die Durchschnitte betragen jetzt:

Bahn:	7 Verkehrstote pro 100 Millionen Passagier-Stunden
Flugzeug:	24 Verkehrstote pro 100 Millionen Passagier-Stunden

Der Vorteil des Fliegens hat sich also umgekehrt - pro Stunde produziert der Flugverkehr mehr als dreimal so viel tödliche Unfälle wie die Eisenbahn.

Literatur

Mehr über Mittelwerte und Preisindizes findet man in jedem Standard-Lehrbuch der Statistik. Mein Favorit ist Bamberg/Baur: *Statistik*, 10. Auflage, München 1996. Für eine alternative Einführung ohne viele Formeln und Symbole siehe auch mein eigenes Buch *Statistik verstehen*, Campus Verlag 1992.

Das Paradox des Zolls ist aus S. P. Chambers: »Statistics and Intellectual Integrity«, *Journal of the Royal Statistical Society*, 1965, 1-16. Die Statistik des sicheren Reisens stammt aus dem gleichnamigen Aufsatz von Lopez-Real in *Stochastik in der Schule*, 1/1989, S. 28-31. Die Karikatur auf Seite 50 ist aus Stephen K. Campbell: *Flaws and Fallacies in Statistical Thinking*, Englewood Cliffs 1974 (mit freundlicher Genehmigung von Prentice Hall, Inc.).

6. Trügerische Trends

*»Nach drüben ist die Aussicht uns verrannt.«
Faust, in Auerbachs Keller*

»Binnen 170 Jahren«, errechnete Mark Twain, »hat sich der untere Mississippi um 240 Meilen verkürzt. Das macht im Durchschnitt $1\frac{1}{3}$ Meile pro Jahr. Daher sieht jeder Mensch, es sei denn, er ist blind oder ein Idiot, daß vor einer Million Jahren der untere Mississippi mehr als eine Million dreihunderttausend Meilen lang gewesen ist und in den Golf von Mexiko hinausragte wie ein Angelstock. Genauso sieht man sofort, daß heute in 742 Jahren der untere Mississippi nur noch eine Meile und dreiviertel messen wird. ... Das ist das Faszinierende an der Wissenschaft: man erhält die tollsten Ergebnisse aus so gut wie nichts ...«

Hier hat Mark Twain etwas getan, was alle Menschen seit Adam und Eva gerne tun: er hat einen Trend, wie die Statistiker sagen, »extrapoliert«. Wenn der Mississippi mit dem Schrumpfen fortfährt wie gehabt, ist er tatsächlich in einigen hundert Jahren nicht länger als ein Bach. Und wenn er schon immer so geschrumpft ist wie zu Zeiten von Mark Twain, dann muß er vor einer Million Jahren mehr als eine Million Meilen lang gewesen sein. Selbst »Blinde und Idioten« sehen das sofort.

Wenn das Argument trotzdem ganz offensichtlich

hinkt, so nicht wegen seiner Logik - die ist nämlich einwandfrei -, sondern wegen seiner Annahmen. Diese sind, wie die Statistiker zuweilen sagen, »heroisch«, oder auf gut deutsch ganz offensichtlich falsch. Der Mississippi ist durchaus nicht immer nur geschrumpft. Mit dieser Unterstellung nimmt Mark Twain nur eine alte menschliche Schwäche aufs Korn, nämlich gerne anzunehmen, alles wäre schon immer so gewesen und würde immer weitergehen wie gehabt.

Diese mentale Trägheit ist uns angeboren wie das Atmen und das Aufrechtgehen. »Warum hast du deinen Regenmantel vergessen?« hat mich mehr als einmal in einem verregneten Urlaub meine Frau gefragt, und ich mußte kleinlaut zugestehen: »Weil zu Hause schönes Wetter war.«

Genauso kleben wir auch anderswo gern an einem aktuellen Trend. Fliegen im Frühling die Maikäfer einmal seltener, fürchten wir um deren Fortbestand. Nehmen sie dagegen wieder Überhand, zählen wir die Jahre, bis uns die Insektenplage das letzte Haar vom Kopf gefressen hat. Bleibt der Winter dreimal kalt, haben wir vor der nächsten Eiszeit Angst. Bleibt die Kälte aber aus, fürchten wir, daß der Nordpol schmilzt. Als einige Wochen der Wind in unseren Landen etwas fester blies, stand »Deutschland wird Orkanland« in der *Bild-Zeitung*. Als die Bundesrepublik für einige Jahre den Weltrekord der niedrigsten Geburtenrate hielt, sahen die Medien eine Entvölkerung der Republik voraus. Im Augenblick dagegen stöhnt man eher über Übervölkerung und fragt sich, wann der Platz nicht mehr für alle reicht. Genau wie vorher das Geburtendefizit erscheint jetzt der positive Wanderungssaldo (die Differenz zwischen Zu- und Abwanderern) für alle Zeiten

zementiert. Sobald wir in die Zukunft sehen, ist Phantasie nicht mehr gefragt. Wie kurvenreich auch immer der bisherige Weg, wir denken gern: ab jetzt geht's nur noch geradeaus.

Neben mentaler Trägheit ist daran vielleicht auch ein Bedürfnis nach Gewißheit schuld - nichts quält mehr als die völlige Ungewißheit, was die Zukunft bringt. Verglichen damit ist das Wissen, daß die Welt am 17. November 2017 endgültig untergeht, regelrecht beruhigend.

Weiter so!

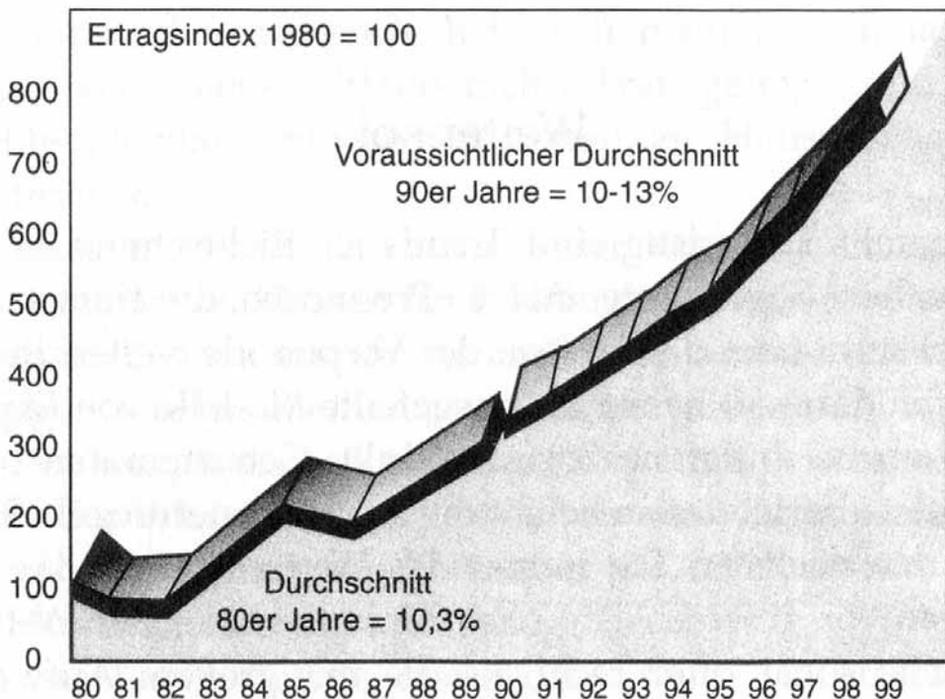
Allenfalls kurzfristig sind Trends als Richtschnur zu gebrauchen. Sogenannte »naive« Prognosen, die einfach das Wachstum oder das Niveau der Vorperiode weiterführen, treffen dann oft besser als ausgefeilte Modelle von Experten, ob bei Inflation, Sozialprodukt, Geburtenraten oder Wetter - bei letzterem behält die simple Bauernregel »Sage für den nächsten Tag immer das Wetter voraus, das am Tag zuvor gewesen ist« nach Meinung mancher Wetterfrösche sogar öfter recht als die mit großem Aufwand produzierte Wettervorhersage aus Offenbach.

Je länger jedoch der Prognosehorizont, desto unzuverlässiger der vergangene Trend. Wenn, wie das Statistische Landesamt Schleswig-Holstein uns wissen läßt, eine durchschnittliche Schleswig-holsteinische Milchkuh im Jahr 1970 rund 4.200 Liter Milch gegeben hat, im Jahr 1980 dagegen 4.700 Liter, und im Jahr 1990 sogar 5.200 Liter, so wäre ich trotz dieses schönen Trends doch sehr vorsichtig

mit der Prognose, daß sie dann im Jahr 2000 5.700 Liter und im Jahr 2010 6.200 Liter geben wird ...

Daher kann ich auch der folgenden Graphik aus der Werbung eines Vermögensberaters nicht vertrauen, die einen aktuellen Börsentrend gleich um zehn Jahre verlängert; angesichts der bekannten Unsicherheiten auf den Kapitalmärkten erscheint mir das doch sehr gewagt:

Gewinnwachstum der deutschen Aktiengesellschaften von 1980 bis zum Jahr 2000

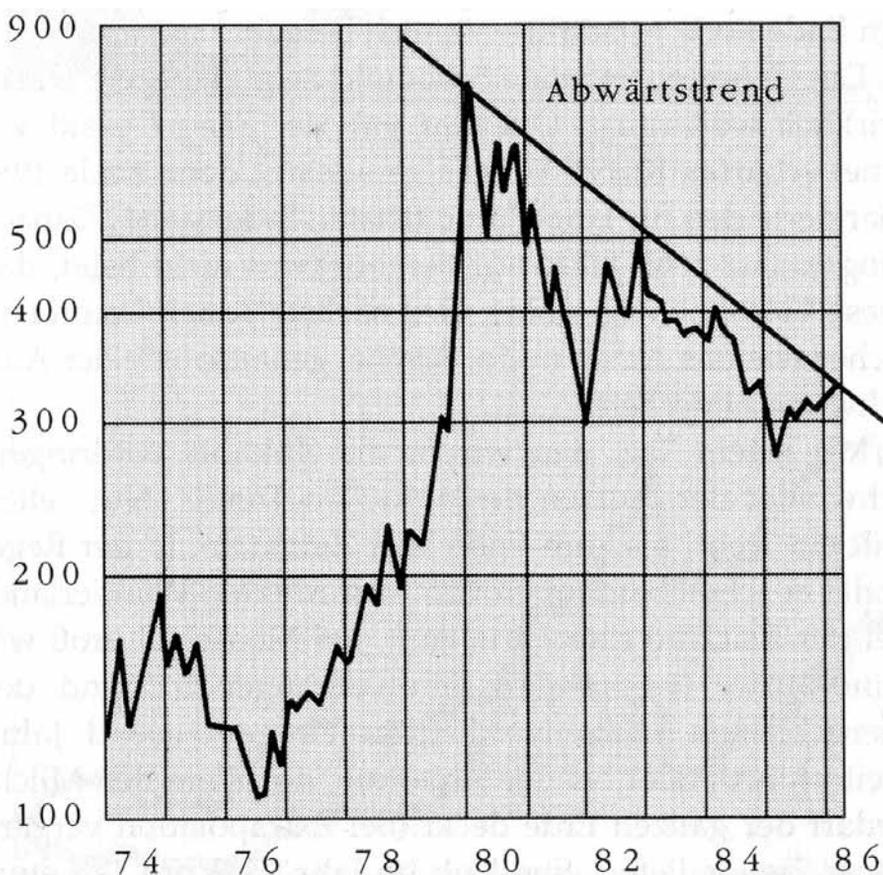


Ein vergangener Trend, zehn Jahre in die Zukunft fortgeschrieben

Noch weiter in die Zukunft hat ein Sportreporter einmal den Meilenweltrekord vorhergesagt. Nachdem der erste offizielle Rekord (eines Engländers namens Webster aus dem Jahr 1865) binnen 30 Jahren von 4 Minuten 44 Sekunden um rund 30 Sekunden auf 4 Minuten 17 Sekunden

verbessert worden war, argumentierte er so: »Pro Jahr eine Sekunde weniger. Die vier Minuten fallen damit spätestens im Jahr 1912.« - »Ist ja toll«, meinte darauf der Chefredakteur. »Damit sind wir 1972 bei drei Minuten und Ostern 2132 schneller als der Schall.«

Aus den gleichen Gründen ist auch der folgende »Abwärtstrend« des Goldpreises aus der *Wirtschaftswoche* zu baldiger Umkehr verurteilt, denn sonst gäbe es Gold bald umsonst, bzw. man bekäme sogar noch Geld dazu (und wie wir heute wissen, hat sich dieser Trend auch sehr bald umgekehrt - zehn Jahre später steht Gold wieder bei 400).



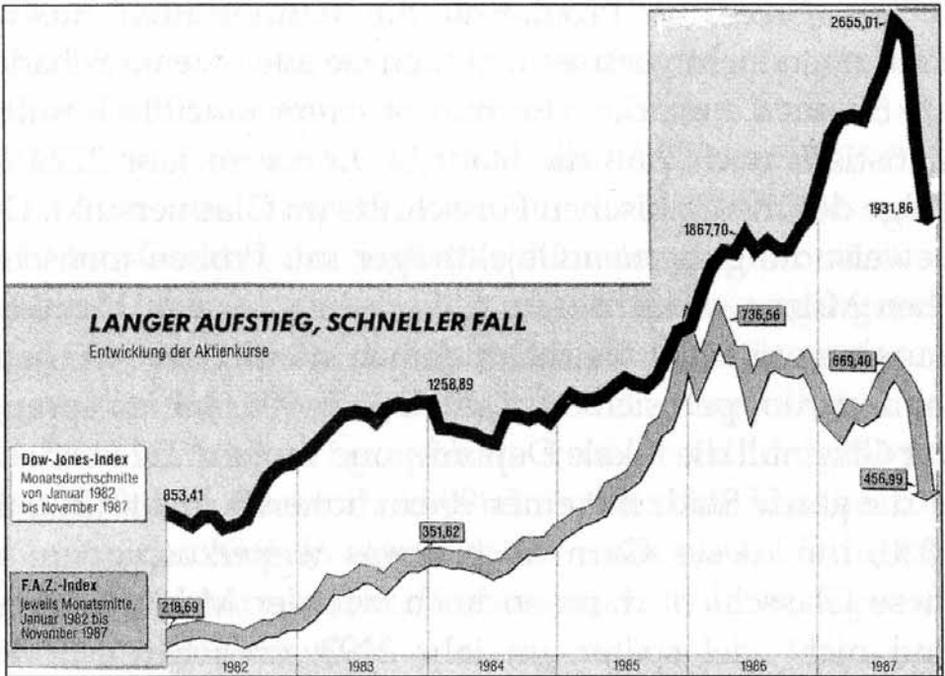
Auch dieser Trend kann unmöglich ewig weitergehen

Trendextrapolierer sind Autofahrer, die nachts ohne Licht auf einer geraden Straße fahren - sie haben nur solange Glück, wie keine Kurve kommt. Spätestens dann zeigt sich der große Nachteil jeder Trendverlängerung: daß sie nur solange die wahre Richtung zeigt, wie der Trend nicht bricht.

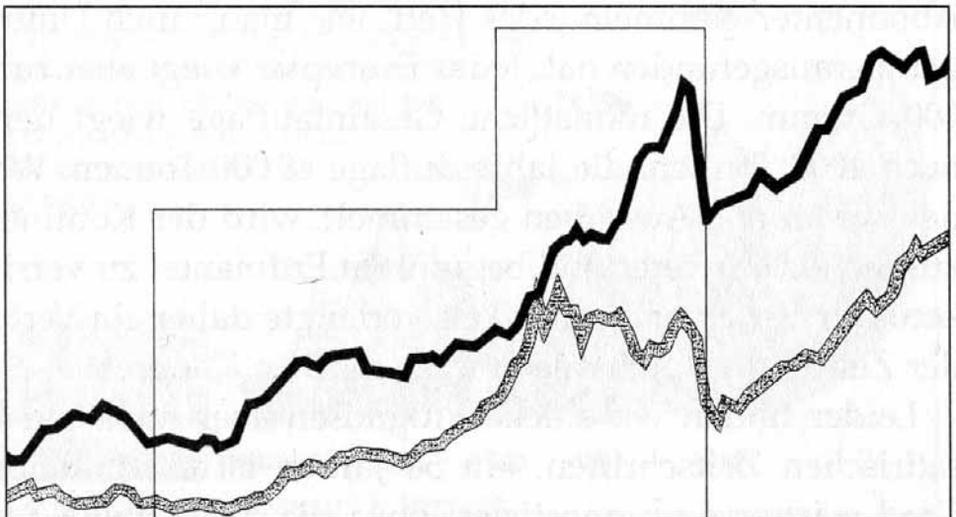
Der auf der gegenüberliegenden Seite abgebildete Trend aus dem *Spiegel* 1987 brach schon kurz nachdem die Druckerschwärze trocken war. Die Graphik zeigt zwei Aktienindizes, den Dow-Jones und den der *FAZ*; sie ist kurz nach dem Börsenkrach vom Oktober 1987 entstanden und suggeriert nur einen Weg (durch einen dicken Pfeil am Ende noch verstärkt), nämlich bergab.

Das darunter stehende Schaubild zeigt dagegen, wie es wirklich weiterging. Offenbar gab der *Spiegel* exakt vor einer scharfen Kurve Vollgas geradeaus, denn Ende 1987 markierte den für lange Jahre tiefsten Indexstand. Danach ging es fast zehn Jahre nur bergauf (was nicht heißt, daß diese Hausse ewig währt. Der nächste Krach kommt so sicher wie das Amen in der Kirche, genauso wie der Aufschwung hinterher).

Mit jedem Tag, den wir in die Zukunft vordringen, schwindet der Nutzen des aktuellen Trends. Nur selten hält ein Trend bis zum Ende aller Zeiten an. In der Regel endet er schon bald im großen Absurditäten-Wunderland, bei der Zeit, die rückwärts läuft, bei Menschen groß wie Dinosaurier (wenn man den vergangenen Trend des menschlichen Längenwachstums einige tausend Jahre weiterführt), oder bei der Superkuh, die allein den Milchbedarf der ganzen Erde deckt (bei Extrapolation vergangener Trends liefert eine Kuh im Jahr 2500 pro Tag etwa 40.000 Liter Milch).



Dieser Trend war bald zu Ende



So ging die Reise wirklich weiter

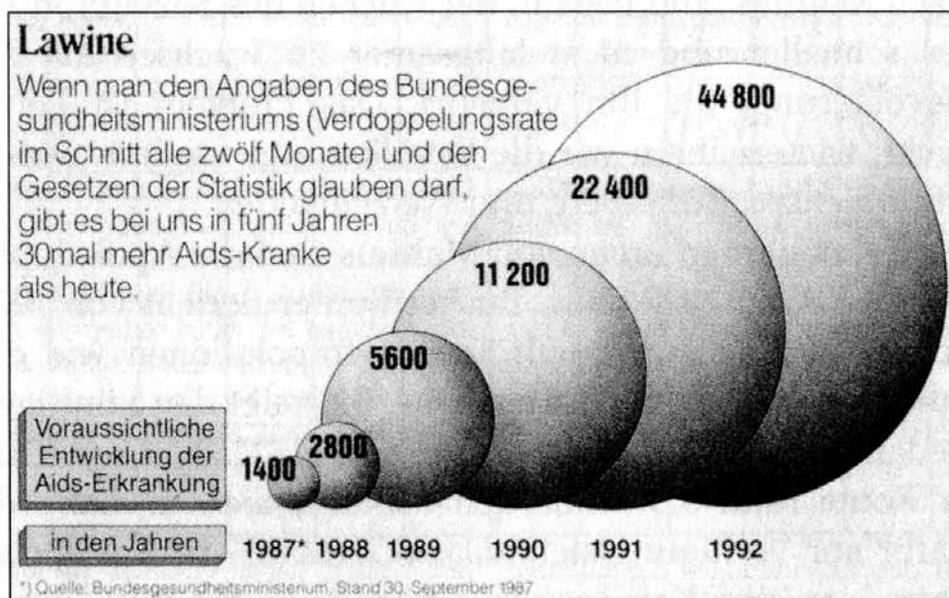
Solange solche Prognosen die Humorseiten unserer Zeitungen nicht verlassen, richten sie auch wenig Schaden an. So wies etwa das »Journal of Irreproducible Results« statistisch nach, daß die Stadt St. Louis im Jahr 2224 als Folge des medizinischen Fortschritts im Glas versinkt. Der Beweis: die gläsernen Objektträger mit Proben menschlichen Milzgewebes, die im Labor eines lokalen Krankenhauses untersucht werden, nehmen ständig zu. Alle neun Jahre verdoppelt sich die Zahl. In den 90er Jahren sprengt der Glasmüll die lokale Deponie, und im Jahr 2224 bedeckt er die ganze Stadt mit einer 90 cm hohen Schicht. (Im Jahr 2350, um dieses Garn noch etwas weiterzuspinnen, ist diese Glasschicht dann so hoch wie der Mount Everest, und nicht viel später, im Jahr 2493, erreichen wir den Mond. Quasi im Vorbeigehen haben wir damit auch das ganze amerikanische Raumfahrtprogramm als überflüssig nachgewiesen.)

Ein anderer Artikel dieser Zeitschrift zeigt, daß der nordamerikanische Kontinent an der Zeitschrift *National Geographic* zugrundegehen wird: Fast alle 10 Millionen Abonnenten sammeln jedes Heft, wie man durch Umfragen herausgefunden hat. Jedes Exemplar wiegt aber rund 400 Gramm. Die monatliche Gesamtauflage wiegt demnach 4.000 Tonnen, die Jahresauflage 48.000 Tonnen. Werden weiter alle Ausgaben gesammelt, wird der Kontinent zusehends schwerer und beginnt im Erdmantel zu versinken. Der Verfasser des Artikels verlangte daher ein Verbot der Zeitschrift durch den amerikanischen Kongreß.

Leider finden wir solche Prognosen aber nicht nur in satirischen Zeitschriften. »In 50 Jahren ist es rund fünf Grad wärmer«, prognostiziert etwa die *Bild-Zeitung*, und der *Spiegel* sieht auf einem Titelblatt voraus, wie der Kölner

Dom im Meer versinkt. Andere sagen die Übernahme der Republik durch Ausländer, drei Autos pro Haushalt in Westdeutschland, einen ständig wachsenden Energieverbrauch oder unseren Untergang durch AIDS voraus.

Die folgende Graphik aus dem Jahre 1987, die auch noch die »Gesetze der Statistik« zum Zeugen nimmt, manipuliert dabei sogar auf zweifache Art: Ein vergangener Trend, der sich schon wenig später als brüchig erwiesen hat, wird erstens völlig unbegründet fortgeführt, und zweitens auch noch graphisch irreführend dargestellt: Die Kranken nehmen in fünf Jahren um das 32fache, die Volumina der Kugeln aber um das mehr als 600fache zu. Auf solche Manipulationen kommen wir in einem eigenen Kapitel nochmals zurück.



Eine Kunstlawine aus im wesentlichen heißer Luft

Das Ende ist nahe

Auch die Wissenschaft ist gegen trügerische Trend-Fortschreibung nicht immun. Den unrühmlichen Anfang machte ein berühmter Ökonom des 18. Jahrhunderts, der ehrwürdige Reverend Robert Thomas Malthus, mit seinen bekannten Prognosen von Nahrungsmitteln und Bevölkerung. Malthus hatte ausgerechnet, daß menschliche Populationen dazu tendieren, sofern niemand sie bremst, sich wie das Glas über St. Louis in konstanten Zeitabständen zu verdoppeln. Dieser exponentiell wachsenden Bevölkerung stellte Malthus ein, wie er glaubte, auf ewig nur linear wachsendes Sozialprodukt gegenüber, das die Bevölkerung nie ernähren könne, mit Hunger und Elend für die Masse der Menschen als unabwendbarer Konsequenz.

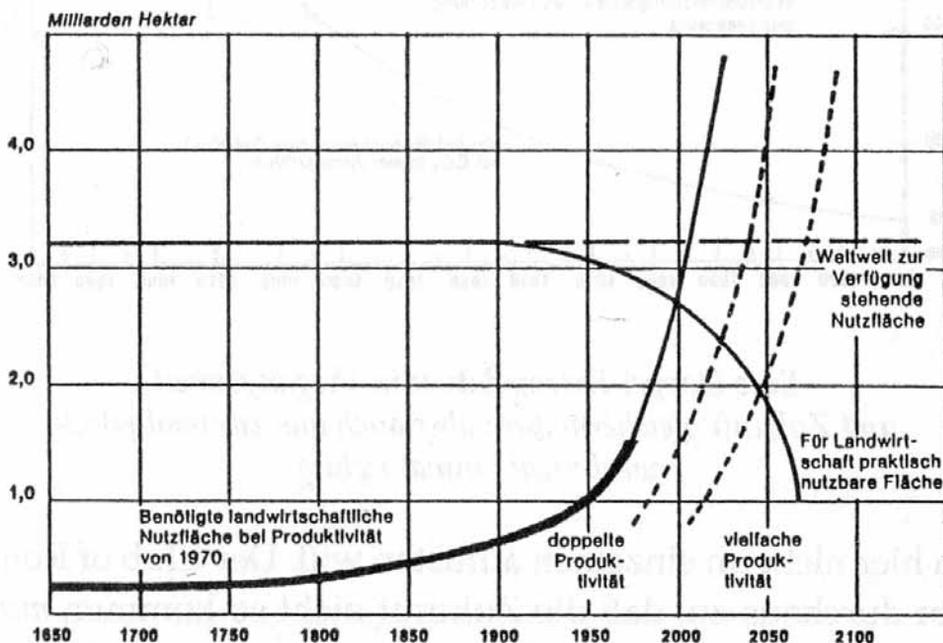
Damalige Trends gaben Malthus durchaus recht, aber wie wir heute wissen, kehrten sie sich zumindest in den westlichen Industrienationen sehr bald um. Schon kurz nach Malthus' Tod begann der Kuchen des Sozialproduktes schneller und nicht langsamer zu wachsen als die Bevölkerung, die ihn verzehrt. Das Problem ist heute nicht, wie ernähren wir die Bevölkerung, sondern wohin mit Milchsee, Butterberg und Weizenhalden.

Die modernen Erben von Malthus sind die Apokalyptiker des »Club of Rome«. Ihr berühmter Bericht von 1972 schockte vor allem durch Trendextrapolationen wie die auf der gegenüberliegenden Seite. Sie zeigt den künftigen Bedarf an Ackerland sowie die maximale Fläche (laut Club of Rome rund 3,3 Milliarden Hektar), die auf der Erde dafür zur Verfügung steht. Das Schaubild »beweist«, daß ganz klar eine Katastrophe droht, denn die noch ungenutzten Reserven sind irgendwann im nächsten Jahrhun-

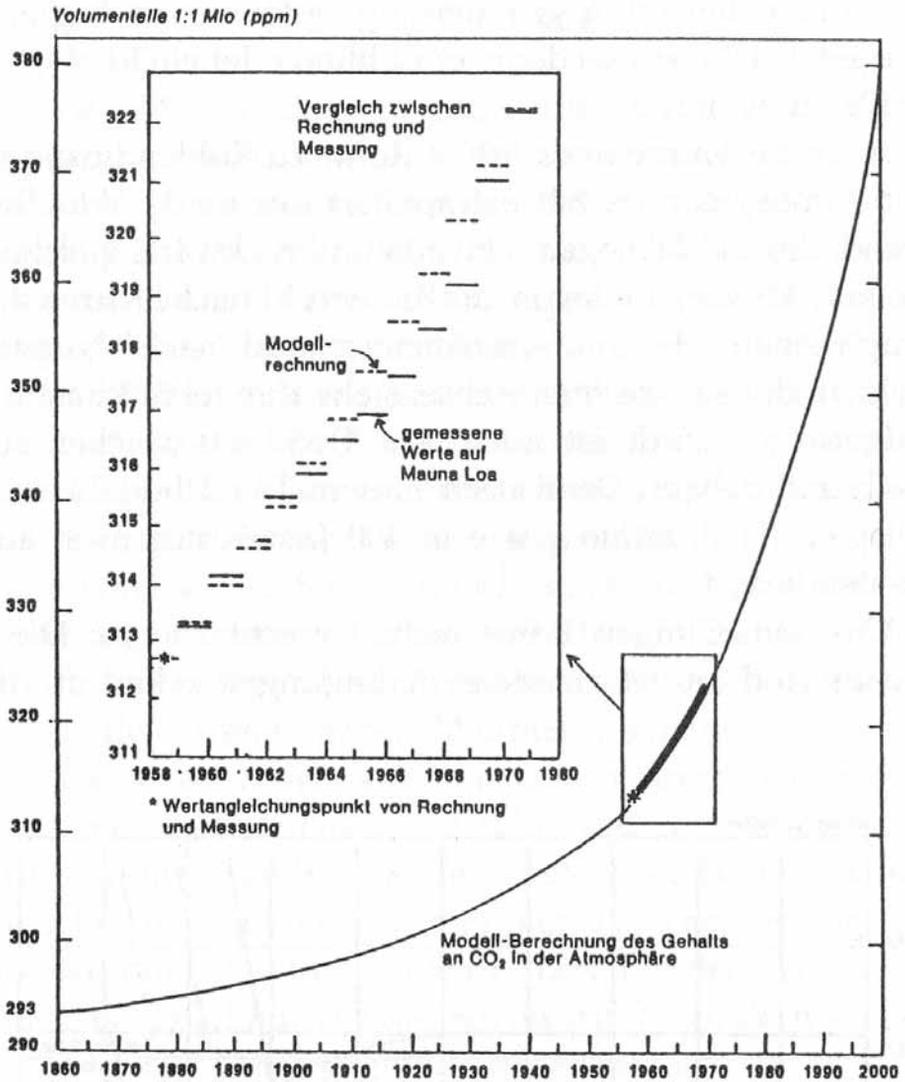
dert unter den Pflug gebracht. Spätestens dann beginnt, wie jeder sieht, »es sei denn, er ist blind oder ein Idiot«, die große Hungersnot.

Auch die Kurve des Club of Rome zu Kohlendioxyd in der Atmosphäre (S. 84) extrapoliert nur einen aktuellen Trend, dieses Mal sogar vorwärts und rückwärts gleichermaßen. Abgesehen davon, daß eine recht flache Kurve mit graphischen Mitteln (Abschneiden und nachfolgendes Dehnen der senkrechten Achse; siehe Kapitel 3) künstlich aufgerichtet wird, ist auch der Trend als solcher auf Treibsand gebaut. Denn nach diesem Trend besteht zum Beispiel die Erdatmosphäre in 400 Jahren nur noch aus Kohlendioxyd.

Um dem Club of Rome nicht Unrecht zu tun: Diese Trends sind an verschiedene Bedingungen geknüpft, die



*Wenn guter Wille Wissenschaft ersetzt:
unhaltbare Trendextrapolationen des Club of Rome*



Eine Doppel-Extrapolation in Vergangenheit und Zukunft gleichermaßen: aber auch aus zweimal falsch wird nicht einmal richtig

ich hier nicht im einzelnen auflisten will. Der Club of Rome gibt durchaus zu, daß die Zukunft nicht so kommen *muß*, sondern nur so kommen *kann*. Aber diese Einschränkung gerät bei vielen seiner Anhänger leicht in Vergessenheit.

Literatur

Die eher pessimistischen Trendextrapolationen des Club of Rome sind nachzulesen in D. Meadows et al.: *Die Grenzen des Wachstums*, Stuttgart 1972. Die Gegenposition dazu beziehen Herman Kahn in *Die Zukunft der Welt*, München 1981, Julian Simon und Herman Kahn in *The Resourceful Earth*, London 1984, und Julian Simon in *Scarcety and abundance*, London 1994. Über das Zusammenspiel von Wissenschaft und Medien bei der Produktion apokalyptischer Trends informiert der Artikel »Warner, Windmacher, Wissenschaftler« in der *Zeit* vom 23.3.1990. Speziell zum Thema Trend und AIDS siehe auch K. M. Wittkowski: »How to lie with AIDS-Statistics - Bemerkungen zur Beeinflussung politischer Entscheidungen durch irreführende Darstellung von Daten«, *Arbeit und Sozialpolitik* 10/1988, S. 306-312. Einen streng wissenschaftlichen Beweis, daß die Erde im Jahr 2026 untergeht, liefern Heinz von Foerster u. a. in »Doomsday: Friday 13. November A.D. 2026«, *Science*, Nov. 1960, 1291-1295. Die graphische Darstellung der Goldpreisentwicklung fand ich in der *Wirtschaftswoche* vom 31.1.1986, Seite 100.

7. Der synthetische Superlativ

»Unter den Blinden ist der Einäugige König.«

Volksmund

»Besuchen Sie das Exmoor Vogel-Paradies!«, lese ich in einem bunten Werbeheft. »Genießen Sie die einzigartige Landschaft, die Heimat der größten Sammlung tropischer Vögel im Norden von Devonshire.«

Leider weiß ich nicht, wie viele Vogelparks es im Norden Devonshires, einer wirklich wunderschönen Landschaft im Süden Englands, im ganzen gibt. Vermutlich nur einen einzigen. Klar, daß dieser dann der größte, schönste, gewaltigste, phänomenalste etc. von allen ist.

Dergleichen Superlative sind nicht echt, sie sind synthetisch hergestellt. Sie schmücken den Besitzer so wie das Kompliment »Einer der gefeiertsten Münchner Landschaftsmaler der ersten Hälfte des 19. Jahrhunderts«, das die Fernsehsendung »1.000 Meisterwerke« einmal für den Pinselkünstler Rottmann fand.

Besonders häufig kommen solche synthetischen Superlative in Prospekten für Touristen-Attraktionen vor: »der einzig erhaltene Adelssitz des Umlandes von Bremen« (Schloß Schönebeck), »die einzige Großkabinenbergbahn Norddeutschlands« (Bad Harzburg), »die größte innerstädtische Parkanlage des Ruhrgebiets« (der Hardt-Park in Wuppertal), »der schönste Profanbau der Renaissance

nördlich der Alpen« (Augsburger Rathaus), »der letzte schwimmende hölzerne Rahsegler Europas« (die Barke »Seute Deern« in Bremerhaven) und so weiter. Während eines Urlaubes in England habe ich das einzige Buntglasfenster von Chagall, das älteste noch aktive Theater, die kleinste Kirche und den ältesten propellergetriebenen eisernen Ozeandampfer dieses Landes aufgesucht, nicht gerechnet all die Attraktionen, deren Singularität ich übersehen habe. Wohin uns auch die Reise führt, ein Dutzend Superlative sind in jedem Fall dabei...

Wir können blind in eine Kiste greifen - wir finden einen Superlativ. Mein Sohn ist Abonnent der »auflagenstärksten Fachzeitschrift für Münzsammler in Westdeutschland«. Dann wird ROSAT, der »größte jemals in Westeuropa gebaute Wissenschaftssatellit«, im Weltraum ausgesetzt. Er hat das »leistungsstärkste Röntgenteleskop der Welt«. Anderswo wird der »größte Kokain-Schmugglerring, der sich je in einer amerikanischen Fluggesellschaft breitgemacht hat«, von findigen Kriminalisten aufgedeckt, das teuerste (besterhaltene, am längsten verschollene, am wenigsten bekannte) Gemälde eines norddeutschen Heidemalers versteigert, oder der wärmste Sonntag im Juni seit zehn Jahren registriert.

Manchmal sieht man diesen Superlativen sofort an, daß sie synthetisch sind, manchmal nicht. Wie kann man sich, um die Meinung zu einer Person befragt, von der man nicht viel hält, elegant und ohne zu lügen aus der Affäre ziehen? Statt zu sagen: Herr X (oder Frau Y) brilliert durch Mittelmäßigkeit, beschreibt man sie als besten Fachvertreter unter 35 an der ganzen Fakultät.

Einen Oscar für jeden

Wie mittelmäßig auch eine Leistung sein mag, für einen Superlativ ist immer Platz. Wer heute in Kunst und Unterhaltung arbeitet, kommt früher oder später um einen Preis nicht mehr herum. Ein Impresario, der auch einem schwach begabten Schützling endlich einen Preis verschaffen will, schreibt (um das Prinzip einmal auf die Spitze zu treiben) einfach einen für den meistgehörten Nachwuchssänger unter 23 Jahren aus Rumänien aus. Den begehrten Oscar gibt es inzwischen nicht nur für den besten Film oder für die beste Regie, sondern auch für die besten Kostüme, die besten Spezialeffekte, die beste Hintergrundmusik und bald auch für den besten unverheirateten mexikanischen Maskenbildner ohne Abitur.

So bekommen am Ende alle einen Preis - in gewisser Weise auch verdient, denn fast jeder kann irgend etwas besser als alle anderen Menschen auf der Welt, ob Vogelstimmen-Imitieren unter Wasser, Dauerflötenspielen, Apfelsinenschalen oder Gurkenschneiden (240 Scheiben aus einer Gurke ist hier der Weltrekord), Kirschkerne-spucken (aktueller Rekord bei 19 Metern und 86 Zentimetern), Baumstammrollen, Küssen, Drachen-fliegen-Lassen oder Memorieren der Lottozahlen von 1955 bis zur Gegenwart. Der eine erkennt mit verbundenen Augen eine Automarke nur am Geräusch der zuschlagenden Tür, der andere Musikstücke auch dann, wenn man sie rückwärts spielt, und wiederum ein anderer hat das komplette Kursbuch der Bundesbahn im Kopf.

Der Betrug beginnt in dem Moment, wo die Kriterien oder die Vergleichsbasis, oder beide, genau für diesen Einzelfall zurechtgeschustert worden sind. Um relativ der

oder die beste zu sein, reicht es offensichtlich aus, daß alle zugelassenen Konkurrenten schlechter sind. Die entscheidende Frage ist dabei, wer kontrolliert die Zulassung? Ist diese offen für alle, oder wird dabei mit Blick auf den Ausgang vorsortiert?

Hier unterscheiden sich synthetische von echten Superlativen durch die Vergleichsbasis. Bei einem echten Superlativ liegt diese unverrückbar fest; sie wird nicht dem Zweck der Übung angepaßt. Bei einem synthetischen Superlativ dagegen werden die Konkurrenten, wie die des seinerzeitigen Box-Schwergewichts-Weltmeisters Floyd Patterson, der diesen Titel vor allem durch Siege über zweitklassige Konkurrenten gewonnen hatte, handverlesen ausgesucht.

Es gibt einen einfachen Test, ob ein Superlativ synthetisch ist: Wir erweitern die Vergleichsbasis. Bleibt der Superlativ bestehen, ist er echt. Geht er dagegen wie Floyd Patterson beim ersten Kontakt mit einem nicht handverlesenen Konkurrenten k.o. (Boxfreunde werden sich erinnern, daß dies Sonny Liston war), so ist er falsch. Wenn also mein Super-Vogelpark aus Nord-Devonshire im Süden Devons einen schöneren Konkurrenten hat, so stinkt das Lob im Reiseführer klar nach Manipulation.

Von diesem Eingrenzen der Kriterien lebt etwa das *Guinness Buch der Rekorde*. Anstelle eines in jeder Hinsicht makellosen und einzigartigen »Spitzenproduktes der menschlichen Evolution« (Josef Stalin im Zeugnis zeitgenössischer Intellektueller), eines Helden sowohl der Tat wie auch der Wissenschaft, suchen wir jemanden, der in zehn Minuten auf einem Arm die meisten Liegestütze schafft. Unter dem Tiel »I'm bigger and better than you« hat der *Economist* gezeigt, daß jede der 24 OECD-Nationen

die anderen in einer Konkurrenz mit Abstand schlägt: In Australien wohnen mehr Menschen als anderswo im eigenen Haus, die Schweden zahlen die meisten Steuern, die Deutschen trinken das meiste Bier, die Schweizer haben das meiste Geld, die Amerikaner das größte Sozialprodukt, kein Land nimmt einen größeren Teil seines Sozialprodukts durch den Tourismus ein als Österreich, und nirgends ist dieses im letzten Jahrzehnt so stark gestiegen wie in der Türkei. Selbst Island hält zumindest einen Rekord, nämlich in der Verbreitung des Spieles »Trivial Pursuit«.

Immer, wenn Buncol ein Heimtor schießt

Ihre höchste Vollendung erfährt die Kunst des synthetischen Superlativs im Sport. Je nach Kriterium und Vergleichsbasis liefert ein und dieselbe Leistung sowohl positive wie negative Superlative, soviel das Papier verträgt. Angenommen etwa, Tennisspieler B erzielt in den ersten 10 Spielen der Saison folgende Bilanz:

Spiel	Turnier	Erdteil	Platz	Ausgang
1	1	Australien	Rasen	verloren
2	2	Amerika	Hartplatz	verloren
3	3	Europa	Rasen	gewonnen
4	3	Europa	Rasen	gewonnen
5	3	Europa	Rasen	gewonnen
6	4	Europa	Hartplatz	verloren
7	5	Europa	Rasen	gewonnen
8	5	Europa	Rasen	gewonnen
9	5	Europa	Rasen	gewonnen
10	5	Europa	Rasen	gewonnen

Daraus lassen sich etwa die folgenden Schlagzeilen ableiten, die alle unbestreitbar richtig sind:

»B im Aufwärtstrend! Mehr als 75 Prozent der letzten acht Spiele gewonnen!«

»B im Abwärtstrend! In mehr als der Hälfte seiner Turniere schon in der ersten Runde ausgeschieden.«

»Phänomenal! B in Europa auf Rasen ohne Niederlage!«

»B's traurige Bilanz: Bisher auf Hartplatz ohne Sieg.«

Und dergleichen Posaunenstöße mehr. Für jede vorgefaßte Meinung finden wir auch den passenden Superlativ.

Alle diese Statistiken sind wahr und trotzdem falsch. Es stimmt, B *hat* mehr als 75 Prozent seiner letzten acht Spiele gewonnen, noch nie in Europa auf Rasen verloren oder auf Hartplatz gewonnen und ist tatsächlich in drei von fünf Turnieren in der ersten Runde ausgeschieden. Die Zahlen lügen nicht. Nur der Zahlenverwerter lügt.

Der dazu benutzte Trick, nämlich die Optimierung der Bezugsmenge, wird am deutlichsten als solcher entlarvt, wenn man das Prinzip auf die Spitze treibt, d.h. die Bezugsmenge ganz verschwinden läßt. Angenommen, B hat noch nie gegen einen Linkshänder gespielt. Dann sind die folgenden Aussagen beide wahr:

»Armer B! Noch nie gegen einen Linkshänder gewonnen!«

»Bravo, B! Gegen Linkshänder noch ohne Satzverlust!«

So produzieren wir auch mit minimalen Daten noch einen schönen Knalleffekt.

Solche synthetischen Superlative werden aber in der Regel schon bald von den Ereignissen überrollt. »Spezialist für Verlängerungen: Becker verlor noch nie ein Abbruch-Match« las ich etwa in der *Hannoverschen Allge-*

meinen Zeitung, kurz bevor Becker sein wichtigstes Abbruch-Match von allen, das Endspiel in Wimbledon, gegen Stefan Edberg glatt verlor.

Auch die Superlative in den folgenden Steckbriefen von vier Fußball-Bundesliga-Clubs (mittlerweile z.T. schon wieder abgestiegen) aus dem *Kicker-Sportmagazin* überlebten größtenteils den nächsten Spieltag nicht. Zunächst ist für jede Mannschaft (FC Homburg, Werder Bremen, VFL Bochum und Fortuna Düsseldorf) noch vollkommen korrekt der bisherige Saisonverlauf zusammengefaßt. Die Bezugsmenge ist dabei klar, nämlich die bisherige Saison. Wir erfahren, daß der FC Homburg drei von vier Elfm Metern verwandelt und noch keine rote Karte eingefangen hat. Werder Bremen hat sogar sechs von sechs Elfm Metern ins gegnerische Netz gesetzt, aber auch schon eine rote Karte einkassiert. Der Spitzenreiter im Verursachen von Elfme-

FC Homburg

Immer, wenn der
Bunco! trifft . . .

Torschützen: Bunco! (5), Wojcicki (5), Schäfer (2), Freiler (2), Beck.

Eigentor: —

Elfmeter für: 4/3.

Elfmeter gegen: 4/2.

Eingesetzte Spieler: 23.

Rote Karte: —

Gelbe Karten: Freiler (4), Frenken (4), Schäfer (3), Mörsdorf (3), Dooley (3), Wojcicki (3), Geschlecht (3), Beck (2), Bunco! (2), Fuchs, Knoll, Stickroth, Ehrmantraut, Kelsch, Müller, Jambo.

Letzte Spiele: UNNUUNNSS.

Besonderes: Zweiter Heimsieg in Folge. Wiederum 2:1, wiederum mit gleicher Torfolge (wie gegen Blau-Weiß): 1:0 durch Bunco!, 2:1 durch Wojcicki-Elfmeter. Immer wenn Bunco! ein Heimtor schoß, verlor Homburg nicht!

Werder Bremen

Daheim ist Völler
meist voll da

Torschützen: Völler (12), Kutzop (7), Burgsmüller (5), Ordenewitz (3), Votava (3), Schaaf (3), Meier, Möhlmann.

Eigentor: —

Elfmeter für: 6/6.

Elfmeter gegen: 6/6.

Eingesetzte Spieler: 21.

Rote Karte: Ruländer.

Gelbe Karten: Kutzop (5), Meier (4), Sauer (4), Möhlmann (3), Votava (3), Neubarth, Hermann, Schaaf, Völler, Burgsmüller, Ruländer, Ordenewitz, Otten, Pezzey.

Letzte Spiele: SSSNUNSSNS.

Besonderes: Unheimliche Heimserie bleibt bestehen: Im Weser-Stadion ohne Niederlage seit dem 19. Mai 1984! Rudi Völler erzielte zum siebenten Male in dieser Saison das 1:0 für Werder.

Vfl Bochum

Das Toreschießen
total verlernt

Torschützen: Schulz (5), Leifeld(4),
Woelk (4), Kree (2), Lameck (2), Oswald,
Wegmann, Benatelli.

Eigentor: —.

Elfmeter für: 4/2.

Elfmeter gegen: 4/4.

Eingesetzte Spieler: 20.

Rote Karte: Reekers.

Gelbe Karten: Woelk (4), Kempe (4), Os-
wald (4), Reekers (3), Leifeld (3), Kree (2),
Benatelli (2), Schulz, Lameck,
Nehl, Heinemann.

Letzte Spiele: SUUNNUUSNN.

Besonderes: Siebenmal in den letzten
neun Spielen ohne Tor. Letzte vier Heim-
spiele: 0:0, 0:1, 0:0, 0:1. Daheim auch seit
dem 11. Oktober ohne Sieg! Erste Bun-
desliga-Heimniederlage gegen Nürnberg!

Fortuna Düsseldorf

Was ist mit dieser
Abwehr los?

Torschützen: Jensen (6), Preetz (4), Bok-
kenfeld (2), Dusend (2), Weikl (2), De-
mandt (2), Blättel (2), Keim, Fach, Thomas,
Krümpelmann.

Eigentor: —.

Elfmeter für: 3/3.

Elfmeter gegen: 9/7.

Eingesetzte Spieler: 21.

Rote Karte: Bockenfeld, Blättel.

Gelbe Karten: Wojtowicz (5), Weikl (3),
Blättel (3), Bockenfeld (2), Grabotin (2),
Jakobs (2), Kremers (2), Thomas (2),
Keim, Schmadtke, Krümpelmann, Kaiser.

Letzte Spiele: SNNNSUNNUN.

Besonderes: Ohne Sieg in den letzten
sechs Spielen (2:10 Punkte). Zum vier-
zehnten Male in dieser Saison mehr als
zwei Gegentore!

tern ist dagegen mit neun Stück Fortuna Düsseldorf, und so weiter. Alle diese Statistiken betreffen die gesamte bis dato absolvierte Saison; sie erlauben damit einen fairen Vergleich.

Die Optimierung der Bezugsmenge beginnt bei »Besonderes«. »Immer, wenn Buncol ein Heimtor schoß, verlor Homburg nicht« ist z. B. eine besondere Erwähnung wert, wie auch, daß der VFL Bochum siebenmal in den letzten neun Spielen ohne Tor geblieben ist. Warum aber ausgerechnet neun? Nun, weiter oben sehen wir, daß der VFL das erste Spiel gewonnen hat. Diese kleine Erweiterung der Bezugsmenge ruiniert daher den ganzen Rekord. Genauso stellen wir bei allen anderen Besonderheiten fest, daß sie bei Licht betrachtet gar nicht so besonders sind.

Literatur

Die Rekorde im Kirschkerne-spucken und Gurkenschnneiden sind aus N. McWhirter (Hrsg.): *The Guinness Book of Records*, Edition 28, London 1981. Die Hackordnung der Nationen ist aus »I'm bigger and better than you«, *The Economist*, 21. 4. 1990, S. 108.

8. Die vorsortierte Stichprobe

*»Es ist mir noch heute rätselhaft, daß man herausbringt,
was sechzig Millionen Menschen denken,
wenn man zweitausend Menschen befragt.
Erklären kann ich das nicht. Es ist eben so.«*

Elisabeth Noelle-Neumann, Meinungsforscherin

Ein Psychiater schrieb einmal, die ganze Menschheit sei verrückt. Gefragt, wie er zu dieser Meinung käme, sagte er: Sehen sie sich doch die Leute an, die in meiner Praxis sind ...

Eine deutsche Frauenzeitschrift behauptet, Mädchen seien intelligenter als Jungen. Statistischer »Beweis«: An den Gymnasien Baden-Württembergs hätten das Jahr zuvor 5.00 Jungen, aber nur 3.00 Mädchen das Klassenziel verfehlt. Fazit: »Was meßbare Intelligenz angeht, kommen Frauen besser weg.«

So einfach ist die Sache aber nicht. Erstens fehlen die Gesamtzahlen der Jungen und Mädchen auf den baden-württembergischen Gymnasien. Die sind aber, wie wir im Kapitel über Prozente gesehen haben, ganz zentral. Wenn insgesamt nur 4.000 Mädchen ein Gymnasium besuchten, von denen 3.800 durchgefallen sind, würde die obige Statistik dem weiblichen Geschlecht nur wenig schmeicheln.

Lassen wir aber diesen Punkt einmal beiseite und nehmen an, nicht nur die absolute, sondern auch die relative Zahl der Wiederholer sei bei den Mädchen kleiner. Selbst dann sagt die Statistik immer noch nichts aus. Zweitens sind nämlich die Gymnasiasten keine Zufallsstichprobe

aller schulpflichtigen Jugendlichen. Zumindest dürfen wir hoffen, daß sie im Durchschnitt etwas klüger sind als andere. Keiner käme daher auf die Idee, der Intelligenzquotient von Oberschülern wäre repräsentativ; diese Stichprobe ist ganz klar, wie die Statistiker sagen, »verzerrt«.

Verzerrt ist das Gegenteil von repräsentativ: Gewisse Teilmengen der Grundgesamtheit - wie hier die besonders begabten unter allen Jugendlichen - sind in der Stichprobe systematisch überrepräsentiert, andere sind systematisch unterrepräsentiert. Nach dem Zweiten Weltkrieg ließ die amerikanische Militärregierung per Stichprobe den Ernährungszustand der deutschen Stadtbevölkerung untersuchen; sie stellte dazu an Bahnhöfen und öffentlichen Plätzen Waagen auf, mit denen wurden die vorbeikommenden Personen gewogen. Wer kam in diese Stichprobe herein? Vor allem gut ernährte Menschen, die viel reisten, während ihre vom Hunger geschwächten und bettlägerigen Landsleute systematisch übersehen wurden.

Diese Verzerrung bewirkt, daß in unserem Ausgangsbeispiel die Intelligenz der untersuchten Jugendlichen vermutlich oberhalb des Durchschnitts liegt. Sie wäre für die Frage »Sind Jungen dümmer als Mädchen?« nicht ganz so schlimm, denn unserer Illustrierten kommt es ja nicht auf Intelligenz an sich, sondern vor allem auf die *Unterschiede* an. Aber auch dazu hat die Stichprobe ohne zusätzliches Wissen keinen Wert, denn die Verzerrung ist nicht notwendig bei Jungen und bei Mädchen gleich. Die höhere Wiederholerquote bei Jungen könnte etwa auch dadurch zustande kommen, daß mehr Jungen als Mädchen von ehrgeizigen Eltern auf ein Gymnasium geschickt werden, die auf einer anderen Schule besser aufgehoben wären.

Noch vor nicht allzulanger Zeit etwa waren Mädchen auf Gymnasien in der Minderheit, und man darf vermuten, zugleich auch eine Auswahl der intelligentesten ihres Geschlechts. Daß dann auch prozentual weniger sitzenbleiben, sollte niemanden verwundern.

Während einer Gastprofessur in Kanada hatte ich einmal zu meiner großen Verblüffung festgestellt, daß ausländische Studenten, vor allem aus China und Hongkong, oft besser waren als die kanadischen; in meinen Kursen fiel kein einziger von ihnen durch, verglichen mit einem Drittel der Einheimischen. Eine Weile spielte ich daher mit der Theorie, daß Asiaten vielleicht begabter für Statistik sind, dann fiel mir auf, daß hier der gleiche Mechanismus wirkt: Eine Familie in Hongkong schickt ihren Sohn oder ihre Tochter doch nur bei einem sicheren Erfolg zum Studium nach Kanada (d.h., nur die wirklich Cleveren werden nach Übersee geschickt), verglichen mit weit bescheideneren Standards in Kanada. Vermutlich gibt es dort genauso viele Dummköpfe und Genies wie überall, nur war die Auswahl, die mir auf der Universität begegnete, eine andere als bei den Studenten aus Hongkong.

Mord und Totschlag

So schließen wir in einem fort zu Unrecht aus verzerrten Stichproben. »Ich glaube, hier werden mehr Leute umgebracht als bei uns«, bemerkte meine Frau nach drei Wochen Aufenthalt in England. Ihr war aufgefallen, daß kein Medien-Tag ohne Mordgeschichte möglich war.

In Wahrheit werden in England aber weniger Menschen umgebracht als in der Bundesrepublik - die Zeitungen schreiben nur mehr dazu. Irgendwie sind Engländer auf eine morbid-abstrakte Weise mehr als andere an Mord und Totschlag interessiert. Nicht umsonst ist hier die Heimat vieler berühmter Kriminalschriftsteller, von Agatha Christie bis Arthur Conan Doyle, und sind in Madame Tussauds Wachsfigurenkabinett neben Königen und Premierministern auch berühmte Schwerverbrecher ausgestellt. Der Eindruck eines Landes voller Massenmörder ist also zum Teil auch ein »Verdienst« der Medien.

In Ländern mit totalitärer Regierung erleben wir das ganze umgekehrt: Durch eine spiegelverkehrte Verzerrung der Nachrichten in die andere Richtung glaubt man dort, daß Kriminalität nicht existiert. Chinesen etwa halten sich für vergleichsweise unbedroht von Raub und Mord, zum Teil auch deshalb, weil die Zensur Berichte dazu systematisch unterdrückt.

Die Nachrichten, die wir im Radio hören, im Fernsehen sehen und in der Zeitung lesen, sind immer nur eine Stichprobe aller potentiellen Nachrichten, und zwar eine je nach Vorurteil und Weltbild der Medien selektierte Stichprobe, deshalb sehen wir die Welt nicht, wie sie ist, sondern wie die Redaktion der *Tagesschau* sie sieht.

Das große Studentensterben

Nicht immer ist bei solchen Verzerrungen Absicht oder Vorurteil im Spiel; oft kommen die Stichproben auch schon verzerrt in den Nachrichten-Redaktionen an. Wenn

ich etwa in der Presse lese, Entbindungen zu Hause seien sicherer als im Krankenhaus, oder daß sogenannte »Retortenbabys« dreimal häufiger als andere Babys bei der Geburt sterben, so sind auch hier zumindest teilweise verzerrte Stichproben im Spiel. Einmal bleibt eine schwangere Frau bei drohenden Komplikationen nicht gern daheim, so daß fast alle »kritischen« Fälle im Krankenhaus versammelt sind, und zum anderen sind Mütter künstlich gezeugter Babys bei deren Geburt im Durchschnitt fünf Jahre älter als andere. Kein Wunder, daß diese Geburten dann weniger glatt verlaufen als der Durchschnitt der übrigen Entbindungen. In beiden Fällen betrachten wir die Realität durch ein gefärbtes Glas, so wie der Portier im Ritz, der ungläubig den Kopf schüttelt, als man ihm erzählte, in Indien herrsche Hungersnot.

»Kann ich nicht bestätigen. Ich habe schon viele Inder gesehen, aber keiner sah so aus, als ob er nichts zu essen hätte.«

Oder nehmen wir die folgende Hiobsbotschaft aus der Londoner *Times*: »60 Prozent aller Piloten in der zivilen Luftfahrt sterben vor dem 65. Lebensjahr!« (30. März 1990, Seite 2). Die Vorsitzenden der Pilotengewerkschaft seien aufs höchste alarmiert, eine sicher nicht billige Studie der »International Federation of Airline Pilots Associations« mit Befragung von 70.000 Piloten zu Trink- und Rauchgewohnheiten, Streß und sexuellen Spannungen sei geplant, um dieses mysteriöse Pilotensterben aufzuklären ...

Der folgende Rat ist dagegen ganz umsonst: Nachsehen, wie die Stichprobe zustandekam. Die *Times* schreibt von Pensionskassen und Lebensversicherungen in Großbritannien, Südamerika und Kanada. Auch wenn wir einmal unterstellen, deren Kunden wären für die Piloten

weltweit repräsentativ, bleibt die Frage, worauf sich diese 60 Prozent Todesfälle beziehen.

Gemeint ist offenbar, wie es in der Meldung ja auch wörtlich heißt: zwei Drittel aller Piloten, die es derzeit gibt, sterben vermutlich vor ihrem 65. Lebensjahr. So genau kann die *Times* das aber gar nicht wissen, es sei denn, sie steht mit dem lieben Gott in Telefonkontakt. Schließlich leben diese Piloten ja noch. Vermutlich sind also die 60 Prozent so zu verstehen: Von den aktiven und ehemaligen Piloten, die im letzten Jahr verstorben sind, waren 60 Prozent weniger als 65 Jahre alt.

Diese Stichprobe der im Jahr zuvor verstorbenen Piloten ist aber verzerrt. Das ist leicht zu sehen, wenn wir das Problem einmal auf Fußballspieler übertragen. »Alarm!«, so könnte eine Schlagzeile der *Bild-Zeitung* verkünden, »Bundesliga-Fußballspieler werden keine 65 Jahre alt!«

Die Fußball-Bundesliga existiert seit 1963. Die ältesten Spieler waren damals Mitte 30 und damit 1990, zur Zeit des Times-Artikels, keine 65. Mit anderen Worten, alle bis dato aus welchen Gründen auch immer verstorbenen Spieler - Alkohol, Herzinfarkt, Verkehrsunfall, eben alles, was uns auch schon jung bedroht - können bei ihrem Tod noch keine 65 Jahre alt gewesen sein; sie starben aber nicht an den Spätfolgen des Fußballspiels, sondern weil es in dieser Risikogruppe keine älteren Personen gab. Und wenn es in welcher Gruppe von Menschen auch immer keine über 65jährigen gibt, kann auch keiner im Alter von über 65 sterben (oder wie es so schön in der *Hessischen Allgemeinen* einmal hieß: »Statistisch erwiesen: Aus parkenden Autos, in denen keine Fotoapparate rumliegen, werden auch keine Fotoapparate geklaut«).

So gesehen ist das mysteriöse Pilotensterben also gar

nicht mehr so mysteriös: Die zivile Luftfahrt hat zwischen 1960 und 1990 gewaltig expandiert, so daß die meisten aktiven und ehemaligen Piloten zur Zeit des *Times-Artikels* weniger als 65 Jahre zählten. Zwanzig Jahre später sieht das ganze schon ganz anders aus; damals waren die Veteranen eine kleine Minderheit - nicht notwendig, weil sie wie die Fliegen vor dem Rentenalter starben, sondern vielleicht auch deshalb, weil es nicht viele ältere Piloten gab. (Übrigens: noch gefährlicher als Piloten leben die Studenten; diese sterben schon mit unter 30...)

Die Statistik und der Sensenmann

»Im Sommersemester X legten die ersten Absolventen des Studiengangs Betriebswirtschaftslehre an der Universität Y in Z ihre Diplomprüfung ab«, läßt eine große deutsche Universität verlauten. »Die Absolventen benötigten für ihr Studium maximal neun Semester und erzielten, so berichtet der Fachbereich, überdurchschnittlich gute Noten. Sie zeigten damit, daß die Bemühungen des Fachbereichs so- und-so um eine Verkürzung der Studiendauer durch sinnvolle Gestaltung der Studienordnung erfolgreich sind...«

Vielleicht waren die »Bemühungen um eine Verkürzung der Studiendauer« wirklich erfolgreich, vielleicht aber auch nicht. Die zitierte Statistik jedenfalls ist dazu ohne Wert. Denn wie viele Semester konnten die ersten Absolventen neun Semester nach Eröffnung eines Studiengangs maximal studieren...

Auch die gemeldeten »überdurchschnittlich guten Noten« überraschen keinesfalls. Die kleine Minderheit der

Studierenden, die innerhalb der Regelzeit ihr Studium beenden, sind natürlich auch begabter als die Bummelanten; daraus auf die Studierenden insgesamt zu schließen, ist reichlich naiv. Hätte die für diese Meldung verantwortliche Pressestelle gewartet, bis auch der letzte Anfänger des ersten Studienjahres in den Hafen des Examens eingelaufen wäre, hätte sich ein ganz anderes Bild ergeben. Schließlich würden wir auch niemals auf den Einfall kommen, die Zeiten der zehn Schnellsten eines Marathonlaufes auf alle Starter insgesamt zu übertragen ...

In abgeschwächter Form kopieren wir diesen Trugschluß auch bei der Berechnung unserer sogenannten »Lebenserwartung« - die mittlere Zeitspanne, die wir von Geburt bis Tod erleben, derzeit in Deutschland für Männer 74 und für Frauen 78 Jahre. Mit anderen Worten, ein Baby, das heute geboren wird, verweilt durchschnittlich noch weitere 74 bzw. 78 Jahre auf dieser schönen Welt. So interpretieren nicht nur Journalisten und Menschen auf der Straße, sondern auch Lebensversicherungen, Rentenfonds und Sozialpolitiker diese Zahl.

Diese Interpretation ist aber falsch, ein heute geborener Junge wird im Durchschnitt nicht 74, sondern 77-78 Jahre alt, und ein heute geborenes Mädchen wird im Durchschnitt nicht 78, sondern vermutlich über 80 Jahre alt.

Die Zahlen 74 und 78 beruhen auf der aktuellen sogenannten »Sterbetafel«; die gibt grob gesagt für jede Altersklasse an, wieviel Prozent in einem Jahr gestorben sind. Zur Zeit zum Beispiel sterben 23 von 10.000 40jährigen Männern und 13 von 10.000 40jährigen Frauen in ihrem 41. Lebensjahr, und genauso wird auch für alle anderen Altersklassen ausgerechnet, wie viele den nächsten Geburtstag nicht erleben.

Bei der Berechnung der Lebenserwartung wird nun unterstellt, daß dies auch zukünftig so bleibt, und das ist alles andere als klar. Wer sagt denn, daß in 40 Jahren, wenn die heute Geborenen 40 Jahre sind, diese Sterbe-Wahrscheinlichkeiten gleichgeblieben sind! Sie werden im Gegenteil vermutlich kleiner sein, wenn man den Prognosen der Demographen glauben darf. Wie lange die heute Geborenen im Durchschnitt wirklich leben, wissen wir erst, wenn der oder die letzte gestorben ist, d. h. vor Ende des nächsten Jahrhunderts sicher nicht. Sollte dann jemand diese Zeilen lesen und nachrechnen, kommt vermutlich ein größeres Mittel als 74 bzw. 78 heraus.

Verschiedenen Wirtschaftsunternehmen wie Versicherungen, die mit anderer Leute Tod und Leben Geld verdienen, ist diese Verzerrung durchaus angenehm. Prämien für Lebensversicherungen werden nach der aktuellen Sterbetafel kalkuliert, sie sind damit zu hoch. Auf der anderen Seite rechnen unsere Rentenstrategen nur mit durchschnittlich 9 bzw. 13 Jahren Rentenzahlungen, und das greift zu tief. Pro durchschnittlichem Rentner und pro durchschnittlicher Rentnerin werden weit längere Zahlungen nötig sein als diese 9 bzw. 13 Jahre (immer ein Rentenalter von 65 unterstellt), die Anforderungen an unsere Rentenkasse werden weit stärker steigen als wir heute glauben.

Was Bilder alter Meister und amerikanische Bomber gemeinsam haben

Trotz aller Verzerrungen sind Stichproben eines der nützlichsten Werkzeuge, das die Statistik kennt. Bei einem Alkoholtest reicht eine Blutprobe - eine Stichprobe im wahrsten Sinn des Wortes - zur Bestimmung unserer Promille völlig aus; dazu müssen uns die freundlichen Verkehrspolizisten nicht das ganze Blut ablassen. Eine unverzerrte Stichprobe darf also durchaus für das Ganze stehen. Genauso reichen, wenn wir nur die Bevölkerung vorher gut durchmischen, 2.000 Personen völlig aus (auch wenn das Frau Noelle-Neumann mysteriös erscheint), um recht genau zu wissen, wie viele Bürger am nächsten Sonntag Partei X wählen oder nach Spanien in Urlaub fahren wollen.

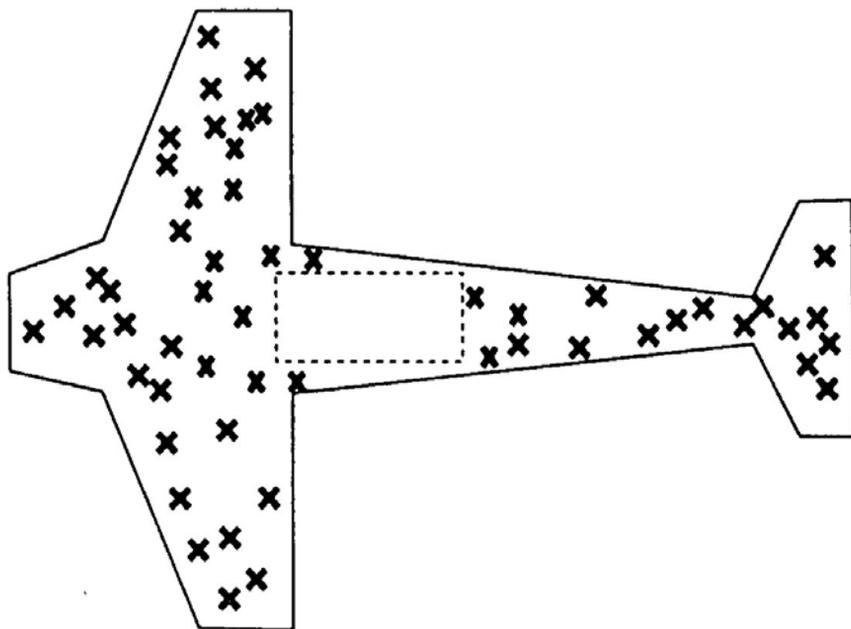
Die Statistiker simulieren das »Durchmischen«, das unser Körper bei einer Blutprobe von allein besorgt, indem sie eine sogenannte »Zufallsstichprobe« ziehen. Grob gesagt bedeutet das: Jeder hat die gleiche Chance, in die Stichprobe hineinzukommen. In diesem Fall spricht man auch von einer »reinen Zufallsstichprobe«. Daneben wenden unsere Meinungsforscher auch noch andere Verfahren wie etwa »repräsentative« Stichproben an, die den Zufall etwas an die Leine nehmen; diese Einzelheiten interessieren uns hier nicht. Wichtig ist allein: Die Chance, in der Stichprobe zu enden, hängt nicht vom untersuchten Merkmal ab.

Deshalb fragen wir zum Thema »Wie viele Deutsche sind katholisch?« natürlich nicht die Teilnehmer der Christmette im Kölner Dom; wir nehmen auch nicht, wenn

wir das Durchschnittseinkommen aller Bundesbürger wissen wollen, als Stichprobe das *Who is Who in Germany* - in beiden Fällen hängt die Wahrscheinlichkeit, für die Stichprobe gezogen zu werden, deutlich von dem untersuchten Merkmal ab.

Man muß kein Statistiker sein, um zu sehen, wohin das systematische Unter- bzw. Überrepräsentieren gewisser Merkmalsträger führt. Im Zweiten Weltkrieg haben die Amerikaner einmal ihre Bomber auf Treffer durch deutsche Flugabwehrgeschütze untersucht: Welche Teile des Bombers werden am häufigsten getroffen, wo sind zusätzliche Panzerungen nötig?

Das nächste Schaubild zeigt, wo diese Treffer lagen:



Die Kreuze markieren die Stellen, wo an heimkehrenden Bombern Flugabwehrgeschosse eingeschlagen hatten; die innerhalb des gestrichelten Rechtecks getroffenen Bomber kamen nicht zurück.

über den ganzen Bomber gleichmäßig verteilt, mit einer Ausnahme: in einem bestimmten Viereck am Rumpf gab es keine Treffer. Warum? Diese Stelle lag nicht im Windschatten der deutschen Flak - hier war der Benzintank, die hier getroffenen Bomber kamen nicht nach Hause ...

Oder nehmen wir ein weniger martialisches, modernes Beispiel - Geldanlage in alter oder neuer Kunst. Sollen wir unsere Spargroschen darin anlegen oder nicht?

»Na klar«, sagte unser Freund, der Galerist. »Etwas Besseres gibt es nicht. Sieh dir nur die Auktionen an! Die Preise steigen wie verrückt!«

Und das tun sie auf den ersten Blick auch in der Tat. Verfolgen wir etwa ein bestimmtes Gemälde, das heute seinen Besitzer wechselt, rückwärts durch die Zeit, kommen wir aus dem Staunen über die Preissprünge von einem Besitzer zum nächsten nicht heraus. Wer vor 100 Jahren als Dekoration fürs Treppenhaus einige Bilder eines obskuren Pinselkünstlers namens van Gogh erstand, wäre heute vermutlich reicher als die Nachfolger von Rockefeller - in einem einzigen Jahr stiegen laut Londoner *Economist* die Preise von auktionierten Gemälden um 44,5 Prozent. Dagegen fallen Aktien und andere Wertpapiere, erst recht das Sparbuch, deutlich ab.

Also auf zu Sotheby's?

Einen Moment, sagt der Statistiker. Sind die heute verkauften Gemälde auch eine Zufallsstichprobe aus allen Gemälden, die es gibt? Sind wirklich *alle* Bilder, die heute die Wände von Museen, Büros und Wohnzimmern auf dieser Erde zieren, auch vertreten unter denen, die auf Auktionen oder sonstwie aktenkundig den Besitzer wechseln? Nur dann stehen doch die Preise der gekauften und verkauften Bilder für den Wert der Bilder insgesamt.

Mit etwas Nachdenken sehen wir aber sofort: Die heute ersteigerten oder in Galerien gekauften Bilder sind alles andere als zufällig ausgewählt. Hier werden ganz im Gegenteil nur die Rosinen aus einem ziemlich trockenen Kuchen herausgepickt. Ein Bild, das keiner haben will, erscheint in der Statistik nicht - die Stichprobe ist deutlich zugunsten hoher Preissprünge verzerrt; genauso können wir auch zeigen (wir fragen nur die Gewinner), daß Lottoscheine gute Kapitalanlagen sind.

Wer also schöne Bilder liebt, sollte sie kaufen und sich daran freuen; reich wird er dadurch nicht.

Das gleiche Auswahlprinzip ist vermutlich auch für die vermeintlich höhere Lebenserwartung von Musikern verantwortlich, die periodisch in den Medien für Gesprächsstoff sorgt. So hatte man etwa in den USA einmal eine Lebenserwartung für Orchesterdirigenten von 73,4 Jahren ermittelt, 4 Jahre mehr als für den Rest der männlichen Bevölkerung, mit dem impliziten Rat der *New York Times* (in der Ausgabe vom 5. Dezember 1978): Lebe länger! werde Dirigent!

Pustekuchen, sagt wieder der Statistiker. Ein männlicher Säugling mit angeborener Begabung zum Dirigenten lebt im Mittel keine Minute länger als ein Baby ohne Musiktalent. Zumindest gibt es keinen Anlaß zur Annahme des Gegenteils; beide werden mit gleicher Wahrscheinlichkeit mehr als 30, 40 oder 50 Jahre alt. Die vermeintlich höhere Lebenserwartung der Musiktalente kommt nur dadurch zustande, daß alle diejenigen aus der Stichprobe herausfallen, die schon jung gestorben sind: nur wer die ersten 20, 30 Jahre seines Lebens übersteht, wird potentiell Dirigent. Würden auch die früh verstorbenen Musiktalente in die Rechnung eingebracht, wäre die

Lebenserwartung von Musikern vermutlich keinen Tag größer als für andere Leute auch.

Auch hier hat man das Pferd von hinten aufgezäumt, man hat nur die erfolgreichen (in diesem Fall die Überlebenden) gezählt. Mit der gleichen Methode könnten wir »beweisen«, daß Bankdirektoren, Universitätsprofessoren oder katholische Bischöfe länger leben als andere. Das ist zwar wahr, aber keine Folge des Berufs, sondern der Tatsache, daß man dergleichen Posten in der Regel erst in fortgeschrittenen Jahren übernimmt. Die Merkmale »fromm«, »geschäftstüchtig« oder »intelligent« (sofern man Professoren dieses Merkmal zugesteht) an sich dagegen haben mit langem Leben nichts zu tun.

Literatur

Zu den Fallstricken einer Investition in Kunst siehe L. Cook: »Buying works of art for Investment«, *The Times*, 7. 7.1990, S. 48, oder »Death of a salesman«, *The Daily Telegraph*, 9. 7.1990, S. 16; zu den irreführenden mittleren Studienzeiten siehe L. Davies: »Verzerrte Zahlen vom Wissenschaftsrat«, *Deutsche Universitäts-Zeitung 17A992*, S. 25-28.

Weitere Beispiele für Fehlschlüsse aus Stichproben gibt auch Kapitel 8 in meinem Buch *Denkste! Trugschlüsse aus der Welt des Zufalls und der Zahlen*, Campus Verlag 1995; daraus ist auch die Graphik des Bombers mit den Flaktreffern entnommen. Wie man diese Fehlschlüsse vermeidet, lernt man in vielen Statistik-Büchern wie etwa D. Freedman, R. Pisani und R. Purves: *Statistics*, New York 1978 (die meiner Meinung nach immer noch beste, leider nur auf Englisch vorliegende Einführung in die Statistik überhaupt). Das Beispiel zur Gesundheit der deutschen Stadtbevölkerung ist aus J. Härtung: *Statistik*, 10. Auflage, München 1995; auch hier wird ausführlich und gründlich vorgeführt, wie sich die oben vorgeführten Fehlschlüsse vermeiden lassen.

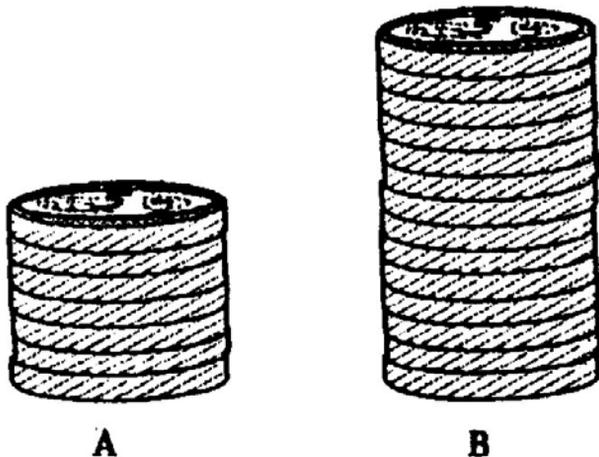
9. Das frisierte Piktogramm

»Oh, wie der Falschheit Außenseite glänzt«

Shakespeare, Der Kaufmann von Venedig

In Land A verdient ein Arbeiter pro Stunde 10 Mark, in B dagegen 20 Mark. Wie stellen wir diesen Sachverhalt am besten graphisch dar?

Aus dem dritten Kapitel kennen wir schon das sogenannte Säulendiagramm: zwei Säulen, eine doppelt so hoch wie die andere. Wenn wir zur besseren Veranschaulichung statt dessen zwei Stapel von Münzen nehmen, erhalten wir ein sogenanntes »Piktogramm«:



Der rechte Stapel symbolisiert ein doppelt so hohes Einkommen

Piktogramme wie dieses bereiten Zahlen auf seriöse Weise graphisch auf; sie helfen uns, die Daten schneller und zugleich auch dauerhafter abzuspeichern. Ein gutes

Piktogramm übersetzt abstrakte Daten quasi in normales Deutsch. Auch wer nicht lesen kann, versteht die Botschaft ohne Schwierigkeiten.

Aber Piktogramme können eine Botschaft auch verzerren. Angenommen, wir sind an einer Betonung von B's Vorsprung interessiert. Das erste Piktogramm nützt dazu nichts, es stellt die Dinge so dar, wie sie wirklich sind: in B verdient man das Doppelte wie in A, und damit Schluß.

Ganz anders dagegen die folgende Version, mit Banknoten statt Münzen: statt zweier Stapel von Münzen, einer doppelt so hoch wie der andere, zwei Geldscheine, einer doppelt so lang (und so breit) wie der andere.



A

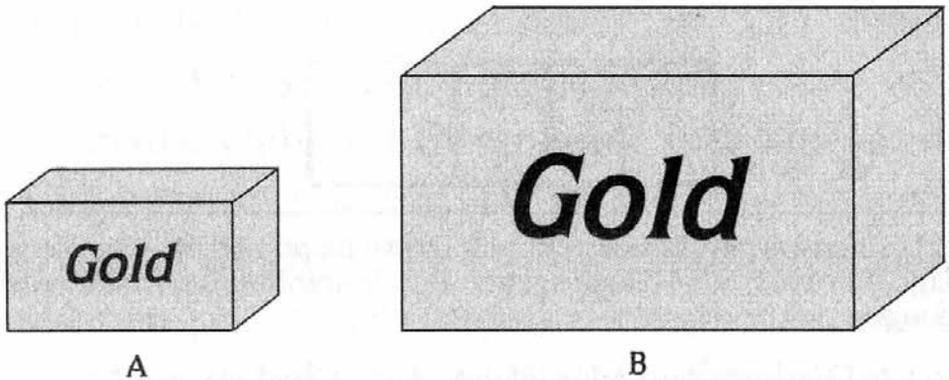


B

Irreführend: Die Fläche des rechten Scheins ist viermal so groß wie die des linken

Die Ränder der Scheine verhalten sich zwar korrekt wie zwei zu eins, aber darauf kommt es hier doch gar nicht an. Bei zweidimensionalen Figuren setzen wir Größe doch mit *Fläche*, nicht mit Seitenlänge gleich. Die Fläche von B's Banknote ist aber *viermal* so groß wie die von A, und mit genau diesem Eindruck verlassen wir das Piktogramm.

Diese Verzerrung läßt sich durch räumliche Darstellung nochmals steigern. Dazu ersetzen wir B's Einkommen durch einen Barren aus Gold, mit doppelt so langen Kanten wie in A:

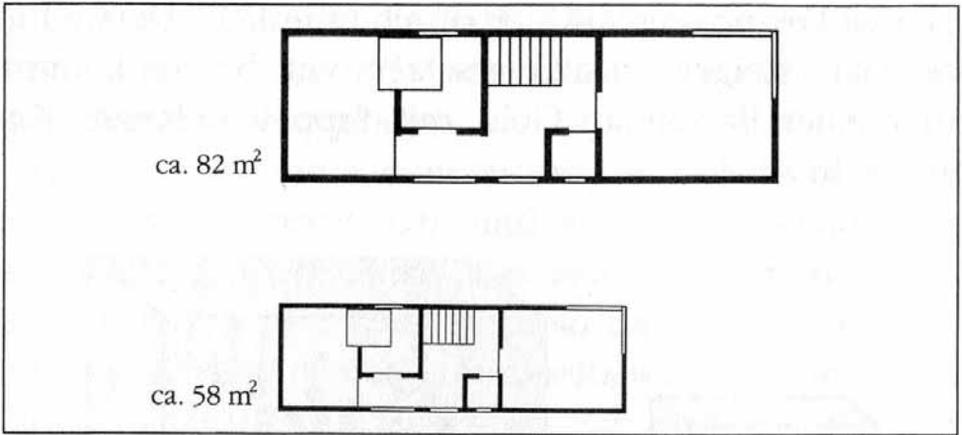


Irreführend: Das Volumen des rechten Goldbarrens ist achtmal so groß wie das des linken

Hier stellt unser Gehirn aber keine Korrespondenz der Kanten, sondern der Volumen her, und diese verhalten sich jetzt sogar wie 8:1 zuungunsten von A.

Die wundersame VW-Vermehrung

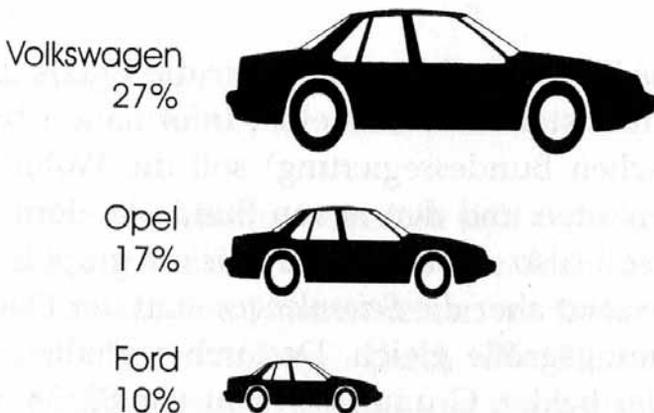
Soweit die Theorie - betrachten wir die Praxis. Die Graphik der nächsten Seite (aus einer Informationsbroschüre der deutschen Bundesregierung) soll die Wohnungsgrößen in den alten und den neuen Bundesländern deutlich machen, soll also explizit zwei *Flächen* graphisch übersetzen; sie setzt aber die *Seitenlängen* statt der Flächen mit der Wohnungsgröße gleich. Dadurch verhalten sich die Flächen der beiden Grundrisse nicht wie 82:58, was den wahren Größenverhältnissen entspräche, sondern wie



Die westdeutsche Wohnung ist viel zu groß

116:58 - die bundesdeutsche Wohnung erscheint nicht 41 Prozent, sondern 100 Prozent größer als die der alten DDR.

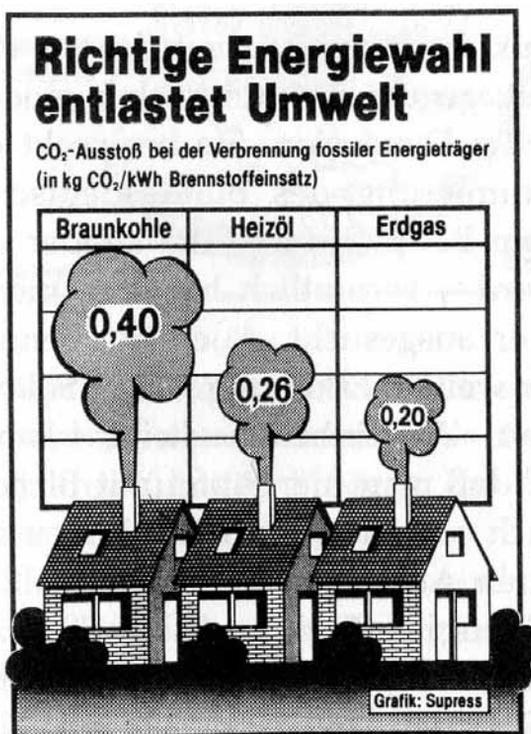
Auch die nächste Graphik zur deutschen Autoproduktion verzerrt die wahren Verhältnisse: der Unterschied zwischen VW und Ford etwa ist viel zu groß; während VW in Wahrheit nur rund dreimal soviel Autos absetzt, ist das VW-Auto in der Graphik achtmal größer (ganz abgesehen



Das VW-Auto übertreibt den Vorsprung

davon, daß die reine Zahl der abgesetzten Autos wenig sagt - viel wichtiger ist hier der Wert bzw. Umsatz, also das, was die Produzenten für ihre Autos einnehmen).

Genauso drastisch überzeichnen die beiden nächsten Graphiken die wahren Fakten. Die erste will zeigen, daß die Braunkohle beim Verbrennen mehr CO₂ erzeugt als andere fossile Energieträger, schießt dabei aber weit über ihr Ziel hinaus: während Braunkohle in Wahrheit doppelt soviel CO₂ erzeugt wie Erdgas, ist die zugehörige Rauchwolke viermal so groß. Die zweite Graphik vergleicht die Subventionen pro Beschäftigten für verschiedene Länder der EG und übertreibt genau so stark: Verglichen mit dem Subventionsriesen Luxemburg erscheint das deutsche Männchen wie ein Zwerg, aber in Wahrheit liegen die Pro-



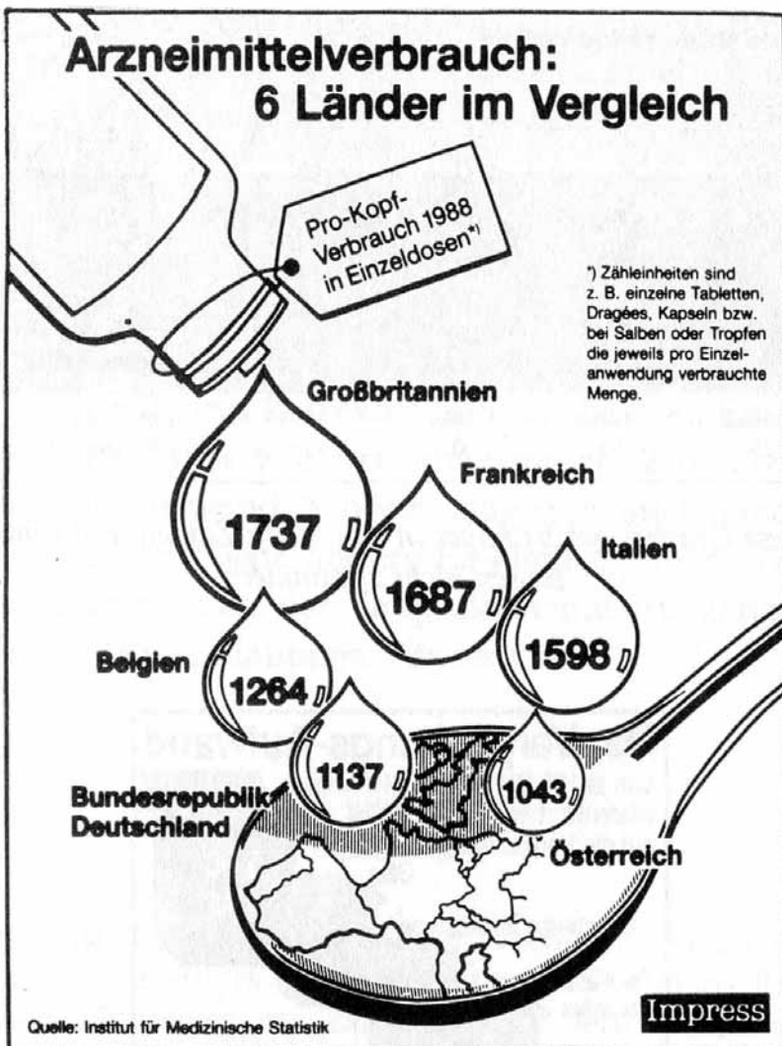
Diese Wolken verhüllen die Wahrheit



Der Luxemburger Subventionsriese ist in Wahrheit gar nicht so viel größer als der deutsche

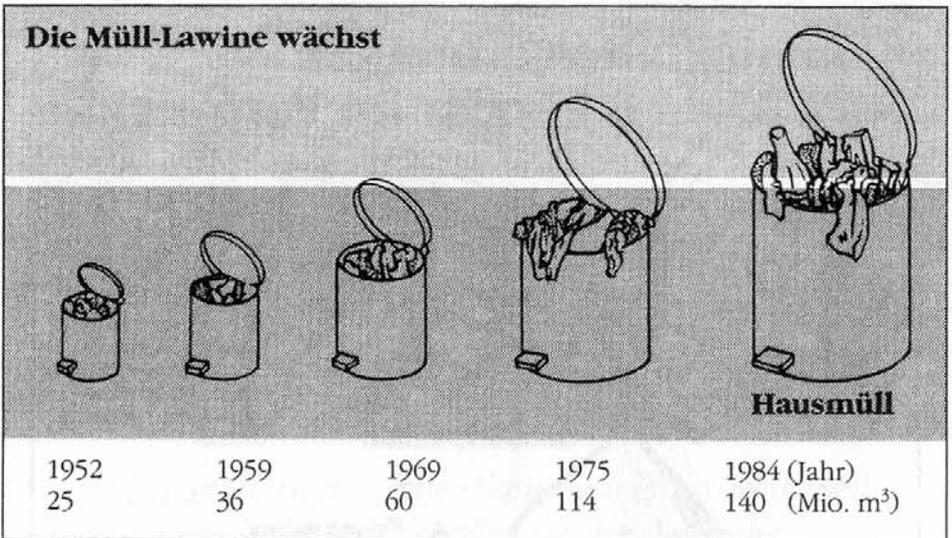
Kopf-Subventionen hierzulande gerade einmal zwanzig Prozent hinter denen in Luxemburg zurück (und auch das ist noch geschönt, weil nur die alten Bundesländer einbezogen sind).

Die Graphik des Arzneiverbrauchs aus einer Pharmafachzeitschrift (gegenüber) führt schon einen Schritt weiter in die dritte Dimension. Sie bezweckt offenbar eine optische Minimierung des bundesdeutschen Pharmakonsums, zum Beispiel durch die Länder, mit denen er verglichen wird - vermutlich hat man hier die größten Pillenschlucker ausgesucht. Auch die benutzte Maßeinheit sagt ohne weitere Erklärung nichts. Solange nicht klar ist, was genau »Zähleinheiten« heißt, können wir nicht ausschließen, daß man hier Äpfel mit Birnen vergleicht. Und schließlich verzerrt auch das Volumen der Blutstropfen den wahren Arzneikonsum. So verhält sich der britische zum deutschen Verbrauch wie 1737:1137, verglichen mit einem Volumenverhältnis der Blutstropfen von 5464:1137. Der englische Verbrauch wird also über- und der bundesdeutsche unterschätzt.

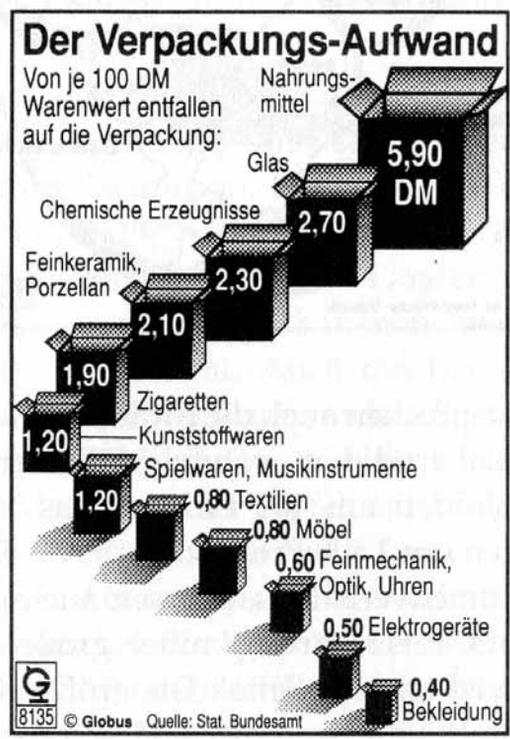


Genauso entsprechen auch die Mülltonnen auf Seite 118 auch nicht annähernd dem wahren Müll. Der Müll selbst wächst in 32 Jahren um 460 Prozent, das Volumen der Tonnen dagegen um 1.400 Prozent.

Falsche Volumenverhältnisse lassen auch den Verpackungsaufwand für Nahrungsmittel größer erscheinen, als er wirklich ist (gleiche Seite). Die größte Kiste symbolisiert einen rund 15fachen Aufwand, verglichen mit der kleinsten, ist aber 50mal so groß.



Diese Graphik gehört selber in den Müll: Zahlen und Eimer passen nicht zusammen



Die Kisten täuschen: in Wahrheit sind die Unterschiede kleiner

Vorsicht also bei Datengraphik in höheren Dimensionen - die beiden passen schlecht zusammen. Allenfalls, wenn die darzustellenden Zahlen selber schon Flächen oder Volumen sind, lassen sich zwei- oder dreidimensionale Figuren oder Gegenstände sinnvoll für Datengraphiken verwenden: Hätte man etwa in der Graphik der Wohnungsgröße die Flächen und in der Graphik der Müllawine das Volumen den jeweiligen Zahlen angepaßt, wäre ein annähernd korrektes Bild entstanden. Eindimensionale Daten aber wie Arzneiverbrauch, Zinssätze oder Verpackungsaufwand gehören dagegen auch eindimensional dargestellt; jede andere Einkleidung erschwert nur das Rückübersetzen der Graphik in die Ausgangsdaten und öffnet Manipulationen Tür und Tor.

Literatur

Wie man Zahlen korrekt in Bilder übersetzt, zeigen Herbert Koberstein: *Statistik in Bildern*, Stuttgart 1973; Hans Riedwyl: *Graphische Gestaltung von Zahlenmaterial*, Bern 1975; Gene Zelazny: *Wie aus Zahlen Bilder werden*, Wiesbaden 1988; oder Jürgen Geßler: *Statistische Graphik*, Basel 1993. Wer sich nicht an der englischen Sprache stört, kann auch einen Blick in die beiden Klassiker von Edward Tufte: *The Visual Display of Quantitative Information* (Cheshire 1982) und *Envisioning Information* (Cheshire 1990) werfen. Weitere Beispiele für Manipulationen mit Datengraphiken inklusive ausführlicher Hinweise, wie man es besser macht, finden Sie ferner in meinem eigenen Buch *So überzeugt man mit Statistik* (Campus Verlag 1994).

10. Wie es in den Wald hineinschallt...

*»Wenn du eine weise Antwort verlangst,
mußt du vernünftig fragen«*

Goethe

Nach einer Umfrage der IG Metall lehnen 95 Prozent aller bundesdeutschen Arbeiter das Arbeiten am Samstag ab. Nach einer zeitgleichen Umfrage des Offenbacher Marplan-Instituts dagegen sind 72 Prozent aller Beschäftigten auch zum Arbeiten am Wochenende bereit.

Das paßt nicht ganz zusammen; hier ist ganz offensichtlich etwas faul. Und es ist auch leicht zu sehen, was da faul ist: der Fragebogen selbst.

»Votum für das freie Wochenende« steht bei der IG Metall in Balkenlettern obenan. Damit ist das Ergebnis programmiert. Ich frage mich, wann die IG Metall diese teure Aktion mit mehr als 2 Millionen Fragebögen überhaupt noch ausgewertet hat.

Auch die Fragen selbst sind suggestiv; sie stellen sicher, daß jeder auch ohne dieses Deckblatt weiß, wie abzustimmen ist. Zum Stichwort »Wochenende« etwa beginnt die Frage so: »Die Gewerkschaften haben die 5-Tage-Woche von montags bis freitags in den fünfziger/sechziger Jahren durchgesetzt... Dadurch sind für alle zusätzliche Möglichkeiten gemeinsamer Freizeitgestaltung entstanden, an die wir uns gewöhnt haben.« Dann folgen nach »Was entspricht Deiner/Ihrer Meinung?« folgende Alternativen:

- Nach meiner Ansicht wäre die Abschaffung des freien Wochenendes ein schwerer Schlag für Familie, Freundschaften, Partnerschaften, für Geselligkeit, Vereine, den Sport und das Kulturleben (1)
- Ich halte den gemeinsamen Freizeitraum des Wochenendes für nicht so wichtig. Seine Abschaffung würde zur besseren Auslastung der Freizeit- und Verkehrseinrichtungen führen (2)
- Weiß nicht/keine Angabe (3)

Diese Worte lassen keinen Zweifel: Wer bei (1) nicht zustimmt, ist ein Klassenfeind.

Genauso ist auch zum Stichwort »Der Samstag« die gewünschte Antwort gleich mit eingebaut. »Die Arbeitgeber und manche Politiker wollen vor allem den Samstag wieder zum normalen Arbeitstag machen«, heißt es hier. »Wie wäre das, wenn Du/Sie regelmäßig am Samstag arbeiten müßtest/müßten?« mit den Alternativen »Würde mir nichts ausmachen« und »Wäre Verlust an Lebensqualität«. Die 95 Prozent aller Stimmen für Alternative 2 überraschen hier niemanden. Viel eher sollte schon bedenklich stimmen, daß trotz dieses ideologischen Trommelfeuers immerhin noch fünf Prozent aller Befragten ihr Kreuz dort markierten, wo es für einen Gewerkschafter eigentlich nicht hingehört.

Genauso suggestiv, wenn auch mit umgekehrter Absicht, fragt auch das Marplan-Institut. Auf die Frage »Inwieweit wären Sie bereit, samstags zu arbeiten, wenn es für die wirtschaftliche Situation Ihres Unternehmens gut wäre?« bietet es folgende Antworten an:

- Gelegentlich, wenn dafür an einem anderen Tag arbeitsfrei ist

- Häufiger, an mehreren Samstagen (etwa 8-12mal jährlich, wenn dafür ein Zusatzurlaub von mehreren zusammenhängenden Tagen herauskommt)
- Abwechselnd, einmal die Woche 6 Tage lang, also einschließlich Samstag, und in der nächsten Woche 4 Tage, so daß Sie in dieser Woche ein »Drei-Tage-Wochenende« zur Verfügung haben; dies etwa an 20 Samstagen im Jahr
- Nein, nicht bereit

Genauso wie die IG Metall nur die negativen Seiten sieht, werden hier die Vorteile herausgestellt, als gälte es allein, aus einem Kuchen die größte Rosine herauszupicken. Außerdem wird durch den Zusatz »wenn es für die wirtschaftliche Situation Ihres Unternehmens gut wäre« die letzte Alternative quasi vom Platz gestellt - wer sägt schon gern am Ast, auf dem er selber sitzt, bzw. entlarvt sich gern als Egoist. Die 72 Prozent Zustimmung zu »gelegentlich samstags arbeiten« überraschen daher keinesfalls.

Zuweilen gelingt es sogar dem gleichen Institut, die öffentliche Meinung binnen Tagen umzukehren. Im September 1983 fragte Emnid die Bundesbürger im Auftrag des Nachrichtenmagazins *Panorama*: Sind Sie für oder gegen die Nachrüstung? Nur 14 Prozent der Befragten waren dafür. Sechs Tage später fragte Emnid das gleiche im Auftrag des Bundesverteidigungsministeriums - jetzt waren 54 Prozent dafür.

Die konkrete Frage bei der *Panorama-Umfrage* lautete: »Wenn die Verhandlungen zwischen den Vereinigten Staaten und der Sowjetunion erfolglos bleiben, sollen demnächst auch bei uns in der Bundesrepublik neue Raketen aufgestellt werden? Sind Sie für oder gegen die Aufstellung neuer Raketen?«

Die konkrete Frage bei der im Auftrag des Verteidigungsministers durchgeführten Umfrage lautete:

»Der Westen muß gegenüber der Sowjetunion stark genug bleiben. Deshalb ist es nötig, in Westeuropa moderne Atomwaffen aufzustellen, wenn die Sowjetunion ihre neuen Mittelstreckenwaffen nicht abbaut.«

Dann sollten die Befragten sagen, ob sie dieser Aussage eher zustimmten oder eher nicht zustimmten, und 58 Prozent stimmten zu.

»Die Zerstörung der schützenden Ozonschicht führt zu Hautkrebs und vielen anderen Gesundheits- und Umweltproblemen. Unterstützen Sie die Forderung von Greenpeace, daß die Firma DuPont, der weltweit größte Erzeuger von Ozon-vernichtenden Chemikalien, deren Produktion sofort beenden soll?« fragt Greenpeace in den USA.

Natürlich unterstützen wir.

»Befürworten Sie eine weitere Kontrollbehörde zusätzlich zu den vielen, die wir schon haben, oder sollen die bestehenden effizienter arbeiten?« fragt eine Vereinigung von Unternehmern, die sich gegen eine Konsumentenschutzbehörde wehrt.

Natürlich meinen hier die meisten, die bestehenden Behörden sollten erstmal effizienter arbeiten, ehe man an eine weitere Behörde denkt. Auch in diesen Fällen ist die Antwort in der Frage bereits eingebaut, man kann sich so wie bei der Nachrüstungs-Umfrage oder wie bei der IG-Metall-Aktion und bei der Marplan-Umfrage die Frage sparen.

Mit der folgenden Fragesequenz könnten wir vermutlich »beweisen«, daß 90 Prozent aller Bundesbürger *für* ein Tempolimit auf unseren Autobahnen sind:

1. Wie Sie vielleicht wissen, gehen schon seit Jahren die Unfallzahlen auf unseren Straßen zurück. Finden Sie das gut? (Na klar!)
2. An vielen Unfällen sind auch Raser schuld. Meinen Sie, man sollte etwas gegen diese Raser tun? (Unbedingt!)
3. Wären Sie bereit, zur Erhöhung der Verkehrssicherheit auch selbst langsamer zu fahren? (Hmm, na ja ...)
4. Meinen Sie, wenn Sie selbst langsamer fahren, dann sollten andere das auch? (Wenn schon, denn schon!) etc. ...

Genauso erzeugen wir aber auch eine Mehrheit für freie Fahrt auf der Autobahn:

1. Wie Sie vielleicht wissen, gelten deutsche Autos als die sichersten der Welt. Finden Sie das gut?
2. 40 Prozent (oder ein anderer imponierender Prozentsatz) unserer Autos werden exportiert, nicht zuletzt wegen ihrer hervorragenden Fahreigenschaften auch bei hoher Geschwindigkeit; sie sichern damit hunderttausende von Arbeitsplätzen. War Ihnen dieser Zusammenhang bewußt?
3. Die meisten Unfälle gibt es bei mäßigen Geschwindigkeiten innerhalb geschlossener Ortschaften. Meinen Sie, daß wir diese vorrangig bekämpfen sollten? etc. ...

Am Schluß dieses Fragebogens macht wohl selbst Joschka Fischer dann sein Kreuz bei »freier Fahrt«.

Umfragen wie diese wollen die öffentliche Meinung nicht messen, sondern machen. Ob die Grünen zur Atomkraft, die Pharma-Industrie zu Tierversuchen, ob die Privaten Fernsehsender zu Rundfunkgebühren oder der

Hartmann-Bund zur freien Ärzteswahl - man weiß vorher schon, was rauskommt: Solche Zahlen sind am besten im Papierkorb aufgehoben.

Darf man beim Beten rauchen?

Diese zentrale Bedeutung der konkreten Frage ist professionellen Meinungsforschern lange bekannt. Im März 1955 etwa fragte Elisabeth Noelle-Neumann eine repräsentative Stichprobe von Arbeitern: »Finden Sie, daß in einem Betrieb alle Arbeiter in der Gewerkschaft sein sollten?«, mit folgendem Resultat:

Dafür	44%
Dagegen	20%
Unentschieden	36%

Dann legte sie einer anderen, gleichgroßen und ebenfalls repräsentativen Stichprobe die gleiche Frage vor, nur mit dem Zusatz »... oder muß man es jedem einzelnen überlassen, ob er in der Gewerkschaft sein will oder nicht?« Ergebnis:

Dafür	24%
Dagegen	70%
Unentschieden	6%

Dieser scheinbar unschuldige Zusatz halbiert damit die Anhängerschaft der Gewerkschaften von vorher 44 auf hinterher nur noch 24 Prozent; zugleich läßt er die Gegner von vorher 20 auf hinterher 70 Prozent anwachsen - eine

mehr als dreifach größere Opposition nur wegen eines kleinen Nebensatzes.

Einen großen Unterschied macht auch, ob man etwas verbieten oder »nicht erlauben soll«. 54 Prozent der Befragten in einer amerikanischen Umfrage glaubten, daß die USA öffentliche Angriffe auf die Demokratie verbieten sollten. 75 Prozent dagegen glaubten, daß die USA öffentliche Angriffe auf die Demokratie nicht erlauben sollten. Wenn wir fragen: »Begrüßen Sie eine Legalisierung der Abtreibung?«, erhalten wir andere Ergebnisse, als wenn wir fragen: »Begrüßen Sie eine Legalisierung des Schwangerschaftsabbruchs?«, und wenn wir fragen: »Soll die Bundesrepublik Deutschland den Zustrom von Asylanten begrenzen?« sieht die Antwort anders aus, als wenn wir fragen: »Soll die Bundesrepublik Deutschland den Zustrom von Asylbewerbern begrenzen?«

Ein Beispiel (wer den Witz schon kennt, bitte unten weiter lesen): Zwei Mönche diskutieren, ob man beim Beten rauchen darf. »Fragen wir doch unsere Vorgesetzten!«

»Fehlanzeige«, kommt der erste zurück. »Ich habe gefragt: »Darf ich beim Beten rauchen?««

»Und?«

»Der Abt hat nein gesagt.«

»Komisch«, sagt der zweite Mönch, »mein Abt hat ja gesagt.«

»Was hast du denn gefragt?«

»Darf ich beim Rauchen beten?«

Solche Zufälligkeiten der Fragestellung können sogar den Lauf der Weltgeschichte ändern. Zu Beginn der Watergate-Affäre fragte das Meinungsforschungsinstitut Gallup die Amerikaner zu ihrer Meinung über ein »Im-

peachment« des Präsidenten. Sind Sie dafür oder dagegen? Die meisten waren dagegen. Und da amerikanische Senatoren und Kongreßabgeordnete auf die öffentliche Meinung mehr hören als auf den Lieben Gott, war das ein Freibrief für den Präsidenten Nixon (er schwenkte diese Umfragen den Journalisten monatelang wie eine Fahne vor der Nase).

Dann fand man aber heraus, daß die meisten Amerikaner überhaupt nicht wissen, was »Impeachment« heißt. Viele glaubten »Verurteilung«, und das ging selbst denen, die das Verhalten des Präsidenten prinzipiell mißbilligten, doch etwas zu weit. Deshalb änderte Gallup leicht die Frage, statt »Befürworten Sie ein Impeachment« fragte man: »Soll der Präsident sich einer Untersuchungskommission des Senates stellen?« (was genau das gleiche heißt) und die Umfragen schlugen geradezu dramatisch um. Eine große Mehrheit stimmte diesem Vorschlag zu, die Politiker paßten sich pflichtgemäß der neuen Mehrheit an, und die Impeachment-Prozedur begann - beziehungsweise kurz vor dem Ernstfall trat Nixon freiwillig zurück.

Sag doch einfach ja

Eine weitere Quelle irreführender Umfrageergebnisse sind unterschiedlich viele Möglichkeiten bei den Antworten. Eine Zeitungsmeldung »59 Prozent aller Deutschen sind glücklich« ist je nach Fragebogen sehr verschieden zu bewerten. Die Frage:

Sind Sie	glücklich	<input type="radio"/>
	unglücklich	<input type="radio"/>
	weder/noch	<input type="radio"/>

wird einen anderen Prozentsatz liefern als die Frage:

Sind sie	glücklich	<input type="radio"/>
	zufrieden	<input type="radio"/>
	eher zufrieden	<input type="radio"/>
	eher unzufrieden	<input type="radio"/>
	unzufrieden	<input type="radio"/>
	unglücklich	<input type="radio"/>

Bei Ja-Nein-Entscheidungen ist es ferner wichtig, welche Alternative in der Ja-Form steht. Denn die meisten Menschen sagen lieber ja als nein. So stellten amerikanische Meinungsforscher einmal die Frage: »Stimmen Sie der Behauptung zu: Für die zunehmende Kriminalität in unserem Land sind in erster Linie die Menschen mit ihrem individuellen Fehlverhalten und nicht die gesellschaftlichen Verhältnisse verantwortlich?« 60 Prozent der Befragten sagten ja. Dann fragten sie die gleichen Leute im gleichen Fragebogen etwas weiter unten: »Stimmen Sie der Behauptung zu: Für die zunehmende Kriminalität in unserem Land verantwortlich sind in erster Linie die gesellschaftlichen Verhältnisse und nicht die Menschen mit ihrem individuellen Fehlverhalten?«, und wieder sagten 60 Prozent der Befragten ja.

Wenn wir also wollen, daß die deutschen Fußballfreunde eine Abschaffung der Bundesliga-Winterpause fordern, müssen wir fragen: »Soll die Winterpause in der Bundesliga abgeschafft werden?« Wenn wir lieber einen Fortbestand der Winterpause hätten, müssen wir fragen: »Soll die Winterpause bleiben?«

Kontaktprobleme

Der nächste Fallstrick ist die Art des Interviews. Bei schriftlichen Befragungen z. B. antworten viele einfach nicht, und zwar in aller Regel diejenigen, denen die ganze Sache nicht so wichtig ist. Wenn wir einer Zeitung die folgende Karte beilegen: »Glauben Sie, daß die Erde demnächst untergeht? Bitte antworten Sie mit Ja oder Nein und schicken Sie die Karte an die Redaktion«, wer wird hier wohl mitmachen? Vor allem Untergangspropheten, die wegen Ozonloch oder Rinderwahnsinn nicht mehr schlafen können, und die Zeitung hat die gewünschte Schlagzeile: »90 Prozent aller Deutschen glauben, daß die Erde demnächst untergeht!«

Wenn ich also in *Readers Digest* lese, mehr als 80 Prozent der in den USA verlorenen Brieftaschen würden samt Inhalt dem Besitzer retourniert, so glaube ich das einfach nicht. Diese Statistik basiert auf der Umfrage eines Journalisten, der in einem Leitartikel seine Leser fragte: »Haben Sie schon einmal Ihre Geldbörse verloren? Dann schreiben Sie mir bitte, ob Sie sie zurückerhalten haben oder nicht.« Ungefähr 600 Menschen haben geantwortet, mehr als 80 Prozent davon hatten ihre Geldbörse zurückerhalten.

Leider sagt das nicht, daß 80 Prozent aller verlorenen Geldbörsen zurückgegeben werden; es sagt allein, daß 80 Prozent der Verlierer, die sich auf den Aufruf meldeten, ihre Geldbörse zurückerhalten hatten; die Verlierer, deren Börsen sich *nicht* wiederfanden - vermutlich die große Mehrheit -, sahen keinen großen Anlaß, deshalb einen Brief zu schreiben. Nur wem ein seltenes Ereignis widerfährt, verspürt den Drang, dies seiner Umwelt mitzuteilen,

und eine verlorene und nicht retournierte Geldbörse ist nach meiner eigenen Erfahrung alles andere als selten.

»Werden im Deutschen zu viele englische Wörter benutzt?« fragten die Dortmunder *Ruhr-Nachrichten* im Dezember 1996. »Sagen Sie uns telefonisch heute von 11 bis 18 Uhr ihre Meinung. Wenn Sie meinen, daß zu viele englische Wörter benutzt werden, sprechen Sie ein >Ja< auf das Band, sonst >nein<.«

Leider habe ich von dieser Umfrage zu spät erfahren - ich hätte gerne mehrmals angerufen und laut »ja« gesagt (mir geht das Pidgin-Englisch in unserer schönen deutschen Sprache ganz gewaltig auf die Nerven). Aber hätte ich auch angerufen, wenn mir diese Verhunzung gleichgültig wäre? So gesehen sind also die 97,4 Prozent Ja-Stimmen keine große Überraschung.

Aber auch Interviews von Mensch zu Mensch sind nicht frei von Tücken. Ein Statistiker-Kollege vom Meinungsforschungsinstitut Emnid erzählte mir einmal, wie eine Umfrage zum Bekanntheitsgrad von PC-Herstellern - ganz ohne Absicht - in die Hosen ging: Welche PC-Hersteller kennen Sie? fragt der Interviewer, und klappt, um die Antwort einzutippen, sein Notebook mit dem großen Firmenzeichen »Compaq« auf dem Deckel auf.

Auch wer die Fragen stellt, ist nicht egal. Wenn weiße Frauen weiße Frauen fragen: »Sind Sie für die Todesstrafe?« kommt etwas anderes heraus, als wenn schwarze Männer weißen Frauen dieselbe Frage stellen. Bei einer Umfrage zur Rassendiskriminierung unter schwarzen amerikanischen Soldaten fühlten sich 11 Prozent der von einem weißen Interviewer Befragten diskriminiert; war der Interviewer ebenfalls ein Schwarzer, stieg der Prozentsatz auf 35 Prozent.

Oder die Befragten reden den Befragern nach dem Mund. So achten die Deutschen beim Autokauf in erster Linie auf Sicherheit und Benzinverbrauch (57 bzw. 53 Prozent der Befragten), kaum dagegen auf PS (22 Prozent) oder Geschwindigkeit (16 Prozent) - es gehört sich eben so. 73 Prozent der Deutschen wollen ihre Organe spenden, 76 Prozent sehen im Fernsehen am liebsten die Nachrichten, 91 Prozent lehnen Gewalt gegen Asylbewerber ab - alles politisch sehr korrekt, aber ob diese Antworten wirklich des Volkes wahre Meinung wiedergeben, weiß der Liebe Gott allein.

Diese Tendenz der Befragten, die Antworten dem Zeitgeist bzw. der Meinung des Interviewers anzupassen, ist auch für das größte Desaster der jüngeren Meinungsforschung verantwortlich, die Prognose für die Wahl zur DDR-Volkskammer am 18. März 1990, denn die Vorliebe der Medien und Meinungsmacher für die SPD war für keinen Wahlbürger zu übersehen. Viele sagten daher, nach der Wahlabsicht gefragt, nur um des lieben Friedens willen: SPD. Schließlich will man ja kein Spielverderber sein. Und so wurde die SPD noch am Wahltag als absolute Mehrheitspartei gehandelt, während sie in Wahrheit nur auf rund ein Fünftel aller Stimmen kam.

Warum es mehr verheiratete Frauen als verheiratete Männer gibt

Als ob die armen Meinungsforscher damit noch nicht Probleme genug hätten, kommt als nächste Komplikation auch noch die Unzuverlässigkeit der Befragten selbst

hinzu, einmal wegen der Tendenz, wie wir eben gesehen haben, zum Interviewer nett zu sein (»The Acquiescence Quagmire«), aber auch wegen persönlicher Interessen, die ebenfalls eine Umfrage verderben können.

So brachte etwa eine Umfrage zur Haushaltsgröße im griechischen Athen zutage, daß dort mehr Menschen wohnen als in München, Hamburg und Berlin zusammen - rund sechs Millionen, wenn wir der Stadtverwaltung glauben dürfen.

Des Rätsels Lösung: Die Athener Stadtwerke planten eine Rationierung des Trinkwassers; jeder Athener sollte künftig höchstens 90 Liter Wasser täglich verbrauchen dürfen. Um diese Rationierung umzusetzen, wurden die Haushalte gefragt: »Wie viele Personen wohnen hier?« Die *Hannoversche Allgemeine Zeitung* gibt uns die Antwort: »Vertraut man den ausgefüllten Wasserwerks-Formularen, dann gibt es Zweizimmerwohnungen, in denen acht oder gar zehn Personen hausen. Selbst Haushalte mit 15 Mitgliedern sind keine Seltenheit. Ungeniert werden jetzt sogar entfernte Verwandte vom Lande nach Athen >eingemeindet<. Und Angehörige, die schon vor Jahren nach Australien oder in die USA ausgewandert sind, tauchen in den Fragebögen als heimgekehrte Mitbewohner auf.«

Wenn wir dem amerikanischen Statistiker Jerome Cohen glauben dürfen, so hat es zu Beginn des Jahrhunderts in ein und derselben chinesischen Provinz einmal 28 Millionen Menschen und einmal 105 Millionen Menschen gegeben - die erste Zählung diente der Musterung und Steuererhebung, die zweite der Verteilung von Hilfsgütern während einer Hungersnot.

Wegen dieser Neigung der Befragten, zu ihren eigenen Gunsten, wo es geht, etwas zu schummeln, bezweifle ich

auch sehr, ob wir jemals per Umfrage die wahre Häufigkeit des Sexualverkehrs erfahren werden. Eine jüngere Studie aus Großbritannien etwa ermittelte für englische Frauen durchschnittlich 2,9 verschiedene Sexualpartner über das Leben verteilt, für englische Männer dagegen 11, und das kommt mir im Licht der alten Weisheit »It takes two to tango« doch sehr spanisch vor, genauso wie ein anderes Resultat einer Umfrage aus Großbritannien, daß es dort weit mehr verheiratete Frauen als verheiratete Männer gibt - in beiden Fällen wurden offenbar die Antworten den jeweiligen Standards von Schicklichkeit bzw. Männlichkeitsbedürfnis angepaßt.

Selbst bei harmlosen Fragen ist auf die Antworten nicht unbedingt Verlaß, wie eine Umfrage einer Studentenzeitung unter Erstsemestern zeigt: Neben Alter, Hobbys, Herkunft und Beruf der Eltern wurde auch nach dem Ehestand gefragt, mit folgendem Ergebnis:

Ledig	1561
Verheiratet	16
Weiß nicht	11

Genauso hat es wenig Sinn, zu fragen, wer seine Kinder schlägt oder alkoholabhängig ist. Auch Fragen wie »Glauben Sie an Gott?«, »Haben Sie Angst vorm Fliegen?«, »Nehmen Sie jeden Tag ein Bad?« oder »Wären Sie gerne Bundeskanzler?« provozieren aus verschiedenen Gründen systematisch die Unwahrheit. Wenn ich also in der Presse lese, 76 Prozent aller Manager von morgen erstrebten in erster Linie »ein Leben mit einem idealen Gleichgewicht zwischen Familie, Arbeit und Freizeitaktivität«, so denke ich, was soll das überhaupt? Wer wirklich ohne Rücksicht auf Verluste nur den Vorstandsvorsitz von

Daimler-Benz anstrebt, sagt dies ganz sicher nicht einem Meinungsforschungsinstitut. Wenn uns also Umfragen berichten, die Deutschen seien besonders kinderlieb, hielten Ausländer für eine kulturelle Bereicherung und läsen gern ein gutes Buch, so wissen wir, was wir davon zu halten haben.

Literatur

Was ich über Umfragen weiß, kenne ich aus E. Noelle-Neumann: *Umfragen in der Massengesellschaft (Einführung in die Methoden der Demoskopie)*, Reinbek 1963, oder H. Schuman und St. Presser: *Questions and Answers in Attitude Surveys*, New York 1981. Aus diesen Lehrbüchern sind auch verschiedene der obigen Beispiele entlehnt. Weitere Beispiele sind aus W. Schneider: *Unsere tägliche Desinformation*, Hamburg 1984, und F. Barnes: »Can you trust those polls?« *Readers Digest*, Juli 1995. Ein nützliches Kompendium der Umfrageproblematik insgesamt, mit verschiedenen Aufsätzen zu Interviewer-Effekt, Frageform und Frageanordnung liefern R. Lehtonen und E. Pahkinen: *Practical Methods for Design and Analyses of Complex Surveys*, New York 1995. Über ein Teilgebiet informieren L. Suchman und B. Jordan: »International Trouble in Face to Face Interviews«, *Journal of the American Statistical Association* 85, 1990, S. 232-253 (mit Diskussion). Einen guten Überblick über die bundesdeutsche Umfragepraxis vermittelt F. Bauske: *Umfragen in der Empirischen Sozialforschung 1945-1982: Datenbestandskatalog des Zentralarchivs für Empirische Sozialforschung*, Frankfurt 1983 (Campus). Zur Statistik des Sexualverkehrs siehe S. J. Gurman: »Six of one ...«, *Nature* 342, 2.11.1989, S. 12, und A. M. Johnson et al.: »Surveying Sexual Attitudes«, *Nature* 343, 11.1.1990, S. 109. Auf die wundersame Vermehrung der Athener wurde ich aufmerksam durch einen gleichnamigen Artikel in der *Hannoverschen Allgemeinen Zeitung* vom 20.3.1993, S. 4, und für eine kritische Analyse

deutscher Umfragepraktiken siehe »Ohrfein an der Haustür« in *Der Spiegel* 26/1994. Und für einen flotten, wenn auch unfairen Verriß der modernen Meinungsforschung überhaupt siehe auch Hans Weigel: *Der exakte Schwindel oder: Der Untergang des Abendlandes durch Zahlen und Ziffern*, Graz 1977. Selten habe ich mich bei der Lektüre eines Buches so geärgert wie zugleich auch amüsiert.

11. Statistisch oder wirklich arbeitslos?

»Die Statistik soll wieder eine Größe werden,
mit der man rechnen kann.«

*Kommentar des >Neuen Deutschland< zur Praxis der amtlichen
DDR-Statistik, zu Zeiten von Gemüseknappheit die schweren
und reichlich vorhandenen Melonen statt dem Obst
dem Gemüse zuzurechnen*

Zur Zeit sind in der Bundesrepublik rund vier Millionen Menschen arbeitslos. Die wahre Zahl weiß niemand so genau. Zwar meldet die Bundesanstalt für Arbeit in Nürnberg monatlich eine exakte Zahl, aber diese Zahl ist nur scheinbar exakt - wie wir schon mehrfach gesehen haben, täuschen solche präzisen Statistiken oft nur über eine fundamentale Unsicherheit hinweg.

Das gilt besonders für die Arbeitslosigkeit. Manche glauben, daß die amtliche Statistik hier systematisch unter-, andere, daß sie übertreibt. Periodische Schlagzeilen wie »Rückgang der Arbeitslosigkeit nur auf Papier« oder »Anstieg der Pseudo-Arbeitslosigkeit« drücken diese Zweifel deutlich aus, hier denkt so mancher: »Wußte ich es doch!« Wie kaum eine andere weckt die Statistik der Arbeitslosigkeit das alte Vorurteil, daß mit Statistik alles zu beweisen sei.

Die folgende Meldung der *Hannoverschen Allgemeinen Zeitung* etwa, unter der Schlagzeile »Arbeitslosenquote verschönt«, steht für den verbreiteten Glauben, die amtlichen Zahlen seien viel zu klein:

»Der vom Bundespresseamt gemeldete Rückgang der Arbeitslosigkeit hat praktisch nicht stattgefunden. Die von der Regie-

rungsbehörde errechnete Verringerung der Zahl der Vollzeitarbeitslosen um 95.313 innerhalb eines Jahres beruht auf einer Umstellung der Statistik, sozialen Hilfen und Arbeitsbeschaffungsmaßnahmen. Dies geht aus einem Vergleich mit Angaben der Bundesanstalt für Arbeit in Nürnberg hervor.«

Die folgende Meldung aus einer anderen Zeitung dagegen, betitelt »Arbeitslos auch durch eine falsche Statistik«, steht für die genauso weit verbreitete Meinung, die amtlichen Zahlen seien viel zu groß:

»Die Verbesserung der Arbeitsmarktstatistik ist eine wesentliche Voraussetzung für die gezielte Bekämpfung der Arbeitslosigkeit. So heißt es in einem Positionspapier der CDU/CSU-Bundestagsfraktion über »Grundzüge der Wirtschafts- und Arbeitsmarktpolitik«. In der Arbeitslosenstatistik sollen nach dem Willen der Unionsfraktion nur noch solche Arbeitslosen erfaßt werden, »die dem Arbeitsmarkt wirklich zur Verfügung stehen und eine Arbeitsstelle nachfragen, die für sie die einzige oder eine wesentliche Erwerbsquelle darstellt«. Für die »wohlstandsbedingte Arbeitslosigkeit« bei der es sich in der Diktion des Unionspapiers um Bürger handele, die nur bedingt an der Aufnahme einer Tätigkeit interessiert sind, beispielsweise weil sie Einkünfte aus Vermögen oder durch einen gutverdienenden Ehepartner haben, könne gegebenenfalls eine gesonderte Statistik eingerichtet werden. Weiterhin solle die Arbeitsmarktstatistik dadurch verbessert werden, daß die tatsächliche Verfügbarkeit eines Arbeitslosen für den Arbeitsmarkt besser kontrolliert und eine getrennte Statistik für Voll- und Teilzeitarbeitsuchende geführt wird.«

Was also geht hier vor? Setzt uns die Bundesanstalt für Arbeit bzw. die Regierung à la DDR-Statistik einmal Gemüse mit und einmal ohne Melonen vor? Existieren viele »offizielle« Arbeitslose wirklich nur in der Phantasie eines Oppositions-Statistikers, oder haben wir im Gegenteil statistisch viel zu viele davon?

Stille Reserven

Die Antwort ist: alle Seiten haben recht. Arbeitslosigkeit ist einer der schwammigsten Begriffe überhaupt, mit Spielraum für Dutzende von Definitionen, von denen eine das gleiche Existenzrecht hat wie die andere, und je nach Definition kommt unter dem Strich einmal die eine und einmal die andere Zahl heraus.

Die Bundesanstalt für Arbeit in Nürnberg macht sich die Sache dabei besonders einfach (nicht aus Bequemlichkeit, sondern weil die Zahlen so am schnellsten verfügbar sind); sie zählt als arbeitslos, wen ein Arbeitsamt als arbeitssuchend registriert, und wer außerdem

- mehr als 18 Stunden in der Woche arbeiten will,
- nicht nur vorübergehend Arbeit sucht,
- älter als 15 und jünger als 65 Jahre ist,
- dem Arbeitsmarkt sofort zur Verfügung steht.

Es ist also gar nicht so einfach, hierzulande amtlich arbeitslos zu sein. Wer nur eine Teilzeitarbeit von weniger als 18 Stunden oder eine Ferienstelle sucht, wer wegen Krankheit vorübergehend nicht zur Verfügung steht oder die Suche per Arbeitsamt ganz einfach aufgegeben hat, der ist damit auch nicht arbeitslos. So betrachtet sind die amtlichen Zahlen also viel zu klein.

Andere Länder sind hier sogar noch restriktiver. In Frankreich oder England zählen Teilzeitarbeitslose überhaupt nicht mit, auch bei mehr als 18 gewünschten Wochenstunden nicht. Anderswo muß man vor Arbeitslosigkeit beschäftigt gewesen sein, d. h. Schulabgänger oder frischgebackene Akademiker ohne Anstellung sind per definitionem niemals arbeitslos.

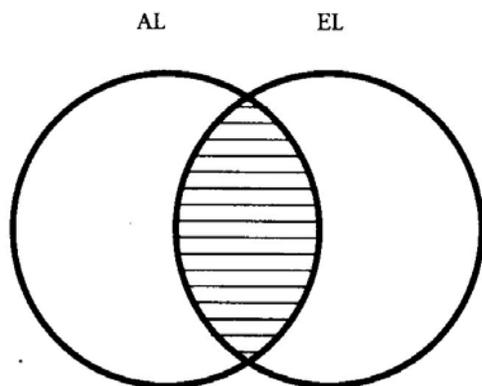
Aber auch liberalere Definitionen kommen vor. In den USA etwa ist auch ein 90jähriger Rentner, der vergeblich nach einer Stelle sucht, im Sinne der Statistik arbeitslos. Dort ist die Statistik also viel umfassender.

Ein Begriff, der unserer intuitiven Vorstellung von »arbeitslos« vermutlich noch am nächsten kommt, ist die sogenannte »Erwerbslosigkeit«. Erwerbslos ist jeder Mann (und jede Frau), ganz gleich wie alt, der gegen Entgelt arbeiten will, ganz gleich wie lange, und bei diesem Bemühen nicht erfolgreich ist. Diese Definition umfaßt (sollte man meinen) die »offiziellen« Arbeitslosen, aber auch viele andere - die entmutigten und nicht mehr registrierten Arbeitswilligen, aber auch die Millionärs-Gattin, die sich emanzipieren will, den joblosen Werkstudenten oder den Rentner, der einen Hausmeisterposten sucht - und erhöht damit auf den ersten Blick die Arbeitslosigkeit,

Dabei vergessen aber viele, daß die amtliche Arbeitslosenstatistik auch »Übergangsarbeitslose«, Schwarzarbeiter und sogenannte »Kindergeld-Arbeitslose« mitzählt, die vor allem zum Abkassieren der Unterstützung als arbeitslos gemeldet sind. Da hier das Interesse an einer Erwerbsarbeit nur vorgeschoben ist, sind diese Personen zwar amtlich arbeitslos, aber nicht erwerbslos in obigem Sinn. So gesehen sind die amtlichen Zahlen also zu groß.

Das folgende Schaubild beseitigt diesen Widerspruch. Es stellt im linken Kreis die amtlichen Arbeitslosen und im rechten Kreis die Erwerbslosen dar. Keine dieser Mengen ist eine Teilmenge der anderen; die Kreise überschneiden sich, wobei die Schnittmenge alle nach *beiden* Definitionen Arbeitslosen umfaßt. Daneben gibt es aber auch solche, die zwar amtlich arbeitslos, aber mangels Arbeitsabsicht nicht erwerbslos sind (der linke Halbmond), wie auch andere

(der rechte Halbmond), die trotz bester Absichten nicht in den Karteien der Arbeitsämter aufzufinden und der sogenannten »stillen Reserve« zuzuordnen sind.



Welcher Halbmond größer ist, sei hier dahingestellt. Üblicherweise bestreitet die jeweilige Regierung, daß der rechte, und die Opposition, daß der linke überhaupt existiert, aber in Wahrheit enthalten beide wohl mehrere hunderttausend Menschen.

Quotenzauber

Damit sind die Probleme der Arbeitslosenstatistik aber nicht erschöpft. Diese weist ja neben den absoluten Zahlen auch noch die Arbeitslosenhilfe aus, den Quotienten aus tatsächlichen und potentiellen Arbeitslosen, und als hätten wir mit dem Zähler nicht schon Probleme genug, kommt hier noch eine verbreitete Unkenntnis dazu, was bei der Arbeitslosenquote eigentlich im Nenner steht. »Schock! Jeder Zehnte arbeitslos!« verkündete etwa die *Bild-Zeitung* vor einigen Wintern in großen Lettern und dokumentierte

damit, daß sie es auch nicht weiß. Jeder zehnte von 80 Millionen Bundesbürgern ergibt 8 Millionen Arbeitslose, und so viele hatten wir noch nie.

In Wahrheit werden die Arbeitslosen durch die *Erwerbspersonen* geteilt. Sämtliche sogenannte »Nichterwerbspersonen« fallen aus dem Nenner heraus. Aber auch gewisse Erwerbspersonen wie Selbständige, Beamte und Soldaten, zusammen immerhin mehr als drei Millionen Menschen, fallen aus dem bundesdeutschen Nenner heraus, der damit kleiner ist als anderswo und so die Arbeitslosenquote größer macht (wenn der Nenner wächst, wird ein Bruch kleiner, und wenn der Nenner schrumpft, wird der Bruch größer). Mit anderen Worten, durch eine Erweiterung des Nenners, die weit weniger auffällt als ein Herumdoktern an den Arbeitslosenzahlen selbst, würde die deutsche Arbeitslosenquote schlagartig *reduziert* (wie etwa in Großbritannien 1986 vorgeführt, wo ohne einen einzigen Arbeitslosen weniger die Quote über Nacht von 13,6 auf 12,2 Prozent gefallen ist).

Wenn man internationale Arbeitslosenquoten auf den gleichen Nenner bringt, löst sich daher mancher nationale Erfolg in statistisches Wohlgefallen auf. In Belgien und den Niederlanden etwa steigt die Arbeitslosenquote bei dieser Rechnung nicht unerheblich an; die bundesdeutsche Quote wie auch die französische dagegen sinken.

Wir sehen also, wie bei Wein und Käse ist auch bei Arbeitslosenquoten nicht eine Sorte wie die andere, und damit haben wir auch schon das Dilemma auf den Punkt gebracht. Nicht die Vielfalt der möglichen Definitionen ist das eigentliche Problem, die alle ihre Vor- und Nachteile und keinen offensichtlich besten Kandidaten haben, sondern die Unkenntnis der Konsumenten über die Spann-

breite des Angebots. Solange wir uns auf eine Sorte verständigen, schadet das auch nichts. Greift aber der Datenhändler unter die Theke und offeriert uns eine Zahl, deren Zusammensetzung nur er selber kennt, dann heißt es aufgepaßt.

Literatur

Das grundsätzliche Dilemma, daß Sachverhalte und Begriffe sich nicht immer decken, heißt in der Statistik auch »Adäquationsproblem«; es wird in vielen Lehrbüchern der Wirtschafts- und Sozialstatistik ausführlich gewürdigt; siehe etwa Peter von der Lippe: *Wirtschaftsstatistik*, Stuttgart 1990. Für einen Überblick der Arbeitslosen-Problematik siehe Ulrich Gramer: »Konzeptionelle Probleme der Arbeitsmarktstatistik aus der Sicht der Arbeitsmarktforschung«, *Allgemeines Statistisches Archiv* 74, 1990, S. 1-18.

12. Arm und reich - alles gleich?

*»Sozialämter müssen Schultüten bezahlen -
Die Schultüte gehört zum notwendigen Lebensunterhalt
von Kindern. Denn ohne Schultüte sind Jungen und
Mädchen bei der Einschulung sozial ausgegrenzt und
werden in ihrem Selbstwertgefühl beeinträchtigt.«*

*Pressenotiz zu einem Urteil
des Bundesverwaltungsgerichts*

Ich habe in meinem ganzen Leben keine Schultüte bekommen. Als ich noch in die Schule ging, konnten meine Eltern sich so etwas nicht leisten, und als sie es sich leisten konnten, ging ich nicht mehr in die Schule. Aber auch vorher wäre ich nie auf die Idee gekommen, unsere Familie wäre arm gewesen.

Heute wären wir es aber. »Die Armut greift um sich«, »Armut - eine bedrückende Realität«, »Jeder achte lebt in Armut«, diese und ähnliche Schlagzeilen muß ich heute ständig lesen. Und diese Schlagzeilen meinen nicht Somalia oder Bangladesch, sondern die reiche Bundesrepublik.

Das ist das große Kreuz der Wirtschafts- und Sozialstatistik: Wie grenzen wir Dinge, Personen, Tatbestände sinnvoll voneinander ab? Nicht nur Armut oder Arbeitslosigkeit, auch vieles andere, was wir zählen, messen und vergleichen, sieht je nach Brille, durch die wir es betrachten, einmal so und einmal anders aus.

Laut Statistischem Bundesamt ist fast jeder zehnte Deutsche heute schwerbehindert - nicht weil wir wirklich immer kränker würden, sondern weil die Meinung, was »schwerbehindert« eigentlich bedeutet, heute sehr viel weiter faßt als noch vor 50 Jahren, vom Rest des Krank-

heitsspektrums ganz zu schweigen. Würden wir die deutsche Meßlatte für Krankheit auf andere Länder übertragen, so wären mehr als eine Milliarde Chinesen, die kein »r« aussprechen können, nach unserer Sicht der Dinge krank, und sie hätten Anspruch auf eine Sprachtherapie.

Oder wie viele Menschen sterben jedes Jahr durch Unfälle auf der Straße? Das Statistische Jahrbuch sagt: im Jahr 1995 waren es in Deutschland 9.435 Menschen. Aber wenn jemand erst eine Woche nach dem Unfall stirbt? Oder erst nach drei Monaten? In der Bundesrepublik wird hier die Grenze bei 31 Tagen angelegt, in der alten DDR lag sie bei 72 Stunden . . .

Oder wie viele Menschen können nicht lesen oder schreiben? Nach dem »Bericht zur Bekämpfung des Analphabetismus in der Bundesrepublik« gibt es heute rund 500.000 Analphabeten in Deutschland, verglichen mit weniger als 10.000 zu Kaiser Wilhelms Zeiten. Wieso? Damals war ein Analphabet, wer seinen Namen nicht schreiben konnte. Heute ist ein Analphabet »eine Person, die sich nicht beteiligen kann an all den zielgerichteten Aktivitäten ihrer Gruppe und ihrer Gemeinschaft, bei denen Lesen, Schreiben und Rechnen erforderlich ist und an der weiteren Nutzung dieser Kulturtechniken für ihre weitere Entwicklung und die der Gesellschaft«.

Oder nehmen wir den verwandten Problemkreis der Schulbildung: Hier habe ich Deutschland (West) in einer Internationalen Statistik von 22 Industrienationen auf dem drittletzten Platz gesehen, weit hinter England und den USA, obwohl die Schulen dort nicht besser und nach meiner eigenen Erfahrung in aller Regel weitaus schlechter sind. Diese Statistik hatte zunächst einmal nur die *Zeiten* in der Schule angerechnet, also reine Bewahranstal-

ten wie eine US-amerikanische Highschool mit deutschen Gymnasien in einen Topf geworfen, und dann noch unsere Lehrlingsausbildung, um die uns die ganze Welt beneidet, nicht als Schule mitgezählt.

Oder nehmen wir die Kriminalität. Bringt ein Amokläufer in Großbritannien zehn Menschen um, so geht das in London als zehn Morde (einfach), im schottischen Glasgow als ein Mord (zehnfach) in die Kriminalitätsstatistik ein. Wer nach elf Bier noch ein zwölftes trinken kann, gilt in München als ein ganzer Kerl und in Teheran als Krimineller; Kriegsdienstverweigerer, Ehebrecher, Umweltverschmutzer, Gotteslästerer, Tierquäler, Vergewaltiger sind einmal kriminell und einmal nicht, und für gewisse Delikte wie die Folterung von Gefangenen kann man hier ins Gefängnis und da auf die Ordensliste kommen.

Oder nehmen wir die Säuglingssterblichkeit. In einer internationalen Statistik war einmal zu lesen, daß in Deutschland (West) von 1.000 Babys 19 die Geburt nicht überleben, mehr als z.B. in Hongkong (15) oder Singapur (14), also in Ländern, von denen wir anzunehmen geneigt sind, daß sie der Bundesrepublik hinsichtlich Hygiene und Gesundheitswesen eher unterlegen sind. Der Grund: die Definition von »Säuglingssterblichkeit«. In Deutschland meint man damit alle lebend Geborenen, die im ersten Lebensjahr versterben, ohne die Totgeborenen, und das ist der Trick: Würden wir wie vielfach anderswo die noch am ersten Tag oder bis zur Taufe verstorbenen Babys als »tot geboren« zählen, hätten wir mit einem Federstrich die Säuglingssterblichkeit halbiert.

Unter anderem auf diese Weise kam auch der vermeintliche Vorsprung des Ostens vor dem Westen Deutschlands bei der Säuglingssterblichkeit zustande: Im Westen zählt

ein Kind als lebend geboren, wenn es atmet *oder* wenn sein Herz schlägt (ist beides nicht der Fall, und wiegt es außerdem noch mehr als 1.000 Gramm, gilt es als totgeboren; wiegt es weniger als 1.000 Gramm, gilt es als Fehlgeburt und geht überhaupt nicht in die Statistik der Todesfälle ein). Stirbt dann ein solches Kind, erhöht das die Säuglingssterblichkeitsstatistik. In der DDR dagegen galt ein Kind als lebend geboren, wenn es atmete *und* wenn sein Herz schlug. War eines von beiden nicht der Fall, galt es von Anfang an als tot geboren; wenn das Kind dann wirklich starb, ging das die Säuglingssterblichkeit nichts an. Denn es hatte ja nie offiziell gelebt...

Die wundersame Vermehrung des Eskimos

Manchmal sind diese je nach Definition einmal so und einmal so erscheinenden Statistiken nicht zu vermeiden; sie sind kein Produkt eines absichtlichen Manipulationsversuchs, eher eine Konsequenz der natürlichen Unbestimmtheit, die vielen Begriffen unausweichlich innewohnt. Solange der liebe Gott nicht von oben dekretiert, wie wir weltweit Streiks, Schulbildung, Kriminalität und Säuglingssterblichkeit zu messen haben, müssen wir mit diesen Unterschieden in den Zahlen leben.

Die Manipulation und der Betrug beginnen da, wo wir die Begriffe so bestimmen, daß das Ergebnis uns ins Weltbild paßt. »Je nach Berechnungsbasis und >erkenntnisleitendem Interesse< läßt sich der Anteil von Arbeiterkindern an Schule und Hochschule verdoppeln oder halbieren«, schreibt Konrad Adam in der *FAZ*. »Um das Gewicht der

noch bevorstehenden Aufgabe zu betonen, wird der Wert anklägerisch auf 7 Prozent gedrückt, soll dagegen die Größe des Geleisteten ins Licht gerückt werden, dann steigt die Quote auf stolze 26 Prozent. Wer glaubt, der Chancengleichheit nahe zu sein, erinnert gern daran, daß zwei von fünf Studenten Väter haben, die nur den Hauptschulabschluß besitzen, und wer gar meint, das Ziel sei schon erreicht, sieht sich durch die Beobachtung bestätigt, daß jeder zweite Student der erste seiner Familie ist, der eine Hochschule besucht.«

Wenn also die Finanzstatistik vieler islamischer Länder keine Zinsen kennt, so nicht, weil es dort, wie der Koran verlangt, keine Zinsen gibt - sie heißen nur »Verwaltungskosten«. Wenn es in manchen katholischen Ländern kaum Abtreibungen und Ehescheidungen gibt, so nicht nur deshalb, weil dort so wenige Abtreibungen und Ehescheidungen stattfinden - sie heißen nur »Annullierungen« bzw. »Fehlgeburten«. Und wenn in weltbekannten Zentren des Sex-Tourismus wie in Thailand kaum Prostitution zu existieren scheint, so nicht, weil es dort keine Prostituierte gibt - sie heißen einfach »Service workers«.

Diese Schwammigkeit der Begriffe ist eine große Versuchung. In Kinderfilme dürfen Fernsehsender nur sehr wenig Werbung packen. Kein Problem - die Filme heißen dann »Familienfilme«. Pro Spielfilm darf man nur soundso viele Minuten werben? Auch nicht schlimm, wir kaufen drei oder vier Filme zu dem gleichen Thema und machen daraus eine »Serie«. Denn in Serien darf man öfter werben ...

Oder was ist ein Tourist? Die deutsche Touristikbranche, die gerne jährlich neue Reiser rekorde meldet, zählt hier auch Geschäfts- und Studienreisende oder Wochen-

endbesucher mit, so daß wir, je nachdem wen wir befragen, auf 50, 60, 70 oder 80 Millionen Urlaubsreisen jährlich kommen.

Oder was ist eine Wohnung? Die amtliche Statistik der alten DDR feierte Jahr für Jahr neue Wohnungsbaurekorde, aber viele dieser neuen »Wohnungen« waren nur Plätze in Altenheimen oder modernisierte Altbauwohnungen; insgesamt war die amtliche DDR-Wohnungsbau-statistik um mehr als eine Million Einheiten zu hoch.

Daß selbst Privatpersonen bestimmte Begriffe einmal so und einmal so betrachten, zeigt die letzte Volkszählung in Kanada: die »Ureinwohner« werden immer zahlreicher; sie haben sich seit der letzten Zählung nur fünf Jahre vorher fast verdoppelt. Aber nicht, weil sie wirklich mehr geworden wären, sondern weil immer mehr Kanadier sich als Ureinwohner sehen. Ein gestiegenes Selbstbewußtsein plus verschiedene soziale Vergünstigungen machen diesen Status immer attraktiver, »mehr und mehr Leute finden es akzeptabel, zuzugeben, wer sie sind«, und so stuft sich ein Kanadier, dessen Ur-Ur-Urgroßvater Indianer oder Eskimo gewesen ist, heute gern als Ureinwohner ein.

Der Maharadscha-Trick

Aber zurück zur Armut. Werden wir wirklich, wie die Medien uns glauben machen wollen, immer ärmer, oder geht es auch hier nicht mit rechten Dingen zu?

Nach einer amerikanischen Studie leidet eines von acht Kindern in den reichen USA an Hunger. Ich habe mir die

Studie einmal angesehen; sie basierte auf einer Umfrage unter Haushalten, in der etwa gefragt wurde:

- Hatte Ihr Haushalt jemals zuwenig Geld, um Lebensmittel einzukaufen?
- Essen Ihre Kinder jemals weniger als sie nach Ihrer Meinung müßten, weil das Geld für Lebensmittel nicht reicht?
- Haben Ihre Kinder jemals geklagt, sie wären hungrig?

und so weiter. Ab einer gewissen Anzahl »ja« wurden dann die Kinder dieses Haushalts als hungrig eingestuft. Mit anderen Worten, zwischen Kindern, die einmal in fünf Jahren, und solchen, die täglich hungern, macht diese Studie keinen Unterschied, genausowenig wie zum Grad des Hungers, ob lebensbedrohlich oder nicht sofort gestillter Appetit. Würde man als »Hunger« das bezeichnen, was man in Indien oder Bangladesch mit diesem Wort bezeichnet, wäre der »Hunger« in den USA praktisch verschwunden.

Noch einfacher machte es sich die Weltbank in einer Studie von 1980: Sie definierte alle Amerikaner als hungrig, die weniger als die durchschnittliche Zahl von Kalorien zu sich nehmen, und konstatierte voll moralischer Entrüstung, daß fast die Hälfte aller Amerikaner Hunger leiden. (Zur Erinnerung: Es gehört zum Wesen eines Durchschnitts, daß rund die Hälfte aller Werte darunter liegen; beim Median ist es sogar genau die Hälfte ...)

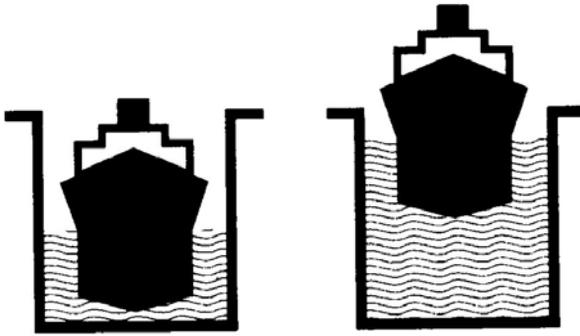
Mit dieser Methode, nämlich alle Personen unterhalb eines Durchschnitts hungrig oder arm zu nennen, zieht sich auch die bundesdeutsche Armutslobby wie Münchenhausen am eigenen Zopf aus ihrem Begründungsnotstand heraus. Hier stellvertretend für Dutzende von ähnlichen

Berichten ein Ausschnitt aus der Bielefelder *Neuen Westfälischen*; sie berichtet unter der Schlagzeile »Immer mehr Armut in reicher Republik«:

»Die Armut wird in der reichen Bundesrepublik ein immer größeres Problem. Nach einer gestern vom Deutschen Gewerkschaftsbund (DGB) und dem paritätischen Wohlfahrtsverband veröffentlichten Studie lebt jeder zehnte Westdeutsche an oder unter der Armutsschwelle... >Noch nie lebten in der reichen Bundesrepublik so viele Arme wie zur Zeit<, so faßt die stellvertretende DGB-Vorsitzende ... gestern in Düsseldorf das Ergebnis einer Studie >Armut im Wohlstand< zusammen ... Besonders häufig von Armut betroffen seien Arbeitslose, alleinerziehende Frauen und kinderreiche Familien.«

Aber wer ist eigentlich arm? »Als arm gilt in der Studie, wer weniger als die Hälfte des durchschnittlichen Einkommens zur Verfügung hat«, meldet die Zeitung weiter, und damit wird klar, warum die Armut nicht verschwinden kann: selbst wenn alle Bundesbürger, reiche wie arme, doppelt oder dreimal soviel Geld verdienen oder sonstwie ausgezahlt bekämen als zuvor - der Anteil der Menschen unter der Hälfte des Durchschnitts bliebe der gleiche, der er vorher war; selbst ein Verzehnfachen oder Verhundertfachen ändert daran nichts - der Anteil der Personen unter der Hälfte des Durchschnitts rührt sich keinen Millimeter von der Stelle. So wie der Tiefgang eines Schiffes in der Schleuse bei jedem Wasserstand der gleiche bleibt, genauso bleibt die Armut, ganz egal, wie reich wir werden, per Begriffsbestimmung immer gleich.

Ein Vergleich von »Armut« über Raum und Zeit hinweg ist sinnlos, wenn dabei der Wasserspiegel in der Schleuse alias die Armutsgrenze schwankt. In der obigen Meldung war die Armutsgrenze bei 2.000 Mark im Monat angesetzt



*Ganz gleich wie hoch das Wasser in der Schleuse steht:
der Teil des Schiffes unter der Wasserlinie bleibt immer gleich*

- darunter ist eine Familie arm. Hätte die *Neue Westfälische* statt dessen die Meßlatte von 1970 angelegt - damals weniger als 1.200 Mark im Monat - hätte die Armut drastisch *abgenommen*. Und hätte sie die heute noch unbekannte Grenze des Jahres 2020 als Kriterium verwendet, wären selbst viele heute Reiche plötzlich arm.

Ob durch die Presse oder unsere eigenen Ansprüche verleitet, auch wir selber tendieren dazu, diese Latte ständig anzuheben. Auf die Frage »Können Sie sich mehr als das Allernötigste leisten?« sagten 1972 zwei Drittel aller Bundesbürger ja, zwölf Jahre später aber nur noch ein Drittel, obwohl auch für die unteren Einkommen der Lebensstandard ständig angestiegen war. Aber nach unserer eigenen Einschätzung sind wir in dieser Zeit verarmt ...

Die gleichen sinnwidrigen Resultate produziert dieser Gummistandard auch bei internationalen Vergleichen. Wenn wir etwa in Deutschland die gleiche Armutsgrenze nehmen wie in Indien, sind alle Bundesbürger reich. Wenn wir in Indien die gleiche Grenze nehmen wie in Deutschland, sind fast alle Inder arm (obwohl sehr viele dieser

»Armen« sich selbst als durchaus reich bezeichnen würden).

Was solche »Armutquoten« wirklich messen, ist also weniger die Armut, wie wir sie üblicherweise verstehen, als vielmehr die Ungleichheit, die Diskrepanz zwischen denen, die relativ viel, und denen, die relativ wenig haben. Auch das kann man mit Recht als Ärgernis betrachten - es gibt durchaus gute Gründe, moralische wie ökonomische, das Volkseinkommen gleichmäßiger aufzuteilen - aber mit wahrer Armut, mit der Sorge um das nackte Überleben hat das alles nichts zu tun, über diese »DGB-Armut« können die wirklich Armen dieser Welt nur lachen. Und anders als diese wahre läßt sich die vom DGB beklagte Wohlstandsarmut spielend leicht beheben - wir nehmen den Reichen ihre Mehrverdienste weg, dann haben alle das gleiche, und die Armut ist verschwunden.

Literatur

Zu den Zahlenspielereien der Touristik-Branche siehe »Zahlenspiele«, *Frankfurter Allgemeine Zeitung*, 14.3.1991. Zu den Ureinwohnern Kanadas siehe Leo Wieland: »Immer mehr Kanadier wollen als Ureinwohner gelten«, *Frankfurter Allgemeine Zeitung*, 29.5.1993.

Eine Generalabrechnung mit der amtlichen DDR-Statistik und ihren Helfern im Westen liefert P. von der Lippe: »Materialien zum Bericht zur Lage der Nation im geteilten Deutschland 1987 als statistische Grundlage für die Staatsverträge mit der ehemaligen DDR« in: *Ökonomische Erfolge und Mißerfolge der deutschen Vereinigung - Eine Zwischenbilanz*, Stuttgart 1994, S. 3-36. Und daß Datenmanipulationen auch in anderen Ländern blühen und gedeihen, zeigt Cynthia Crossen in *Tainted Truth: The Manipulation of Fact in America*, New York 1994.

13. Allerlei Kartentricks

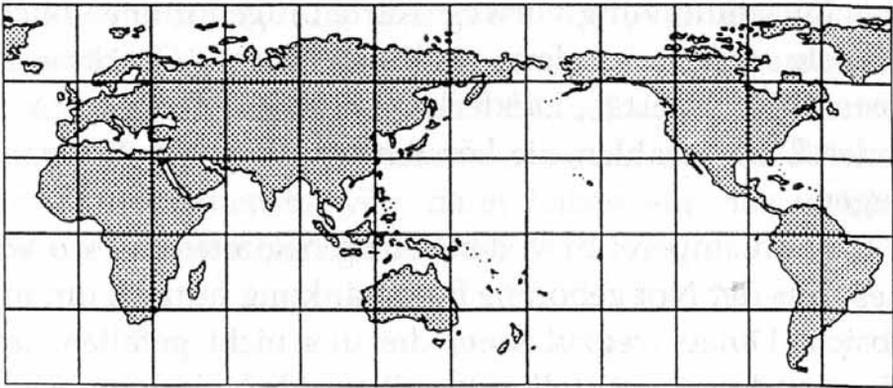
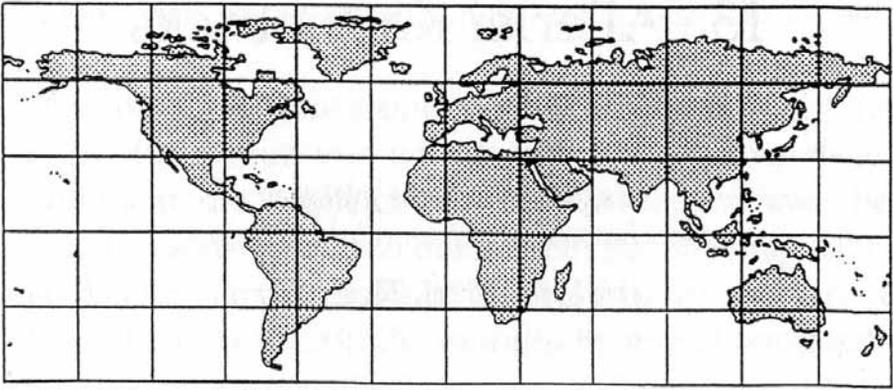
*»Es gibt keine reine Wahrheit,
aber ebensowenig einen reinen Irrtum.«*

Friedrich Hebbel, Tagebücher

Eine Entschuldigung vorweg: Karten lügen immer. Nicht weil sie wollen, sondern weil sie müssen. Ob Straßenkarten, Wetterkarten, Länderkarten - sie müssen aus Millionen Daten wählen, sie können nie die ganze Wahrheit zeigen.

Die Schlamperei bzw. der Betrug beginnen da, wo wir diese aus der Not geborene Beschränkung nutzen, um mit Absicht Dinge wegzulassen, die uns nicht gefallen, um Dinge anders darzustellen, als sie wirklich sind, um Sachverhalte zu beschönigen oder zu dramatisieren. Das fängt schon mit den Rändern einer Karte an. Denn bekanntlich ist die Erde eine Kugel, mit einer Oberfläche ohne natürliches Zentrum, so daß jede zweidimensionale Weltkarte auch schon ein kleines Werturteil enthält. Die erste der beiden Weltkarten auf der nächsten Seite etwa zeigt die Welt, wie wir Europäer sie gerne sehen: Europa wie die Spinne im Netz ziemlich in der Mitte, der Rest der Erde drumherum. An dieses Bild sind wir gewöhnt, den meisten fällt dabei nicht auf, wie wir Europäer uns so auch optisch zu den Herren dieses Globus machen.

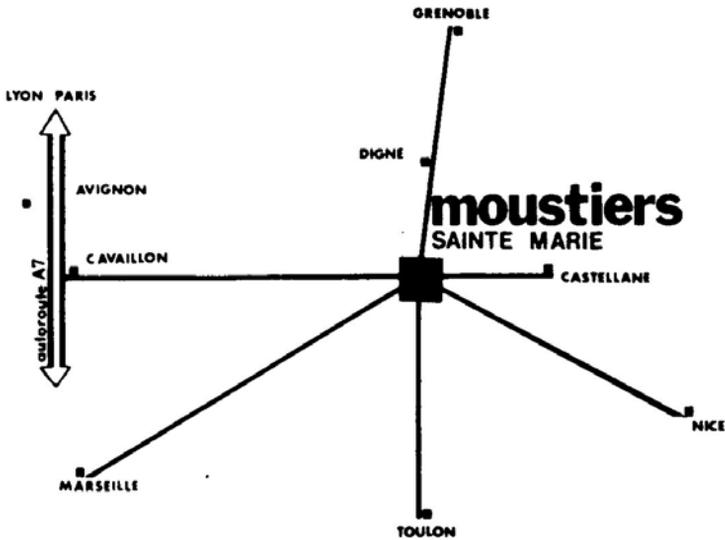
Die zweite Karte zeigt die Welt, wie die Japaner, Chinesen und Australier sie sehen. Das Reich der Mitte in der



Wo ist das Zentrum? Je nach Rahmen ändert sich das Bild

Mitte, und die »Gweilos«, die Barbaren, an den Rändern (alternativ könnte man auch daran denken, die Karten auf den Kopf zu stellen; dann wären die armen Australier »Down under« plötzlich Spitzenklasse. Nirgendwo steht geschrieben, daß der Norden in Karten immer oben liegen muß, und früher war der Norden auch in vielen Karten unten).

Dergleichen Werturteile sind bei Weltkarten leider nicht zu umgehen - irgendwohin muß man nun mal die Mitte setzen. Und daß jeder natürlich gern die Mitte da sieht, wo er oder sie selber sich befindet, kann ich gut verstehen.

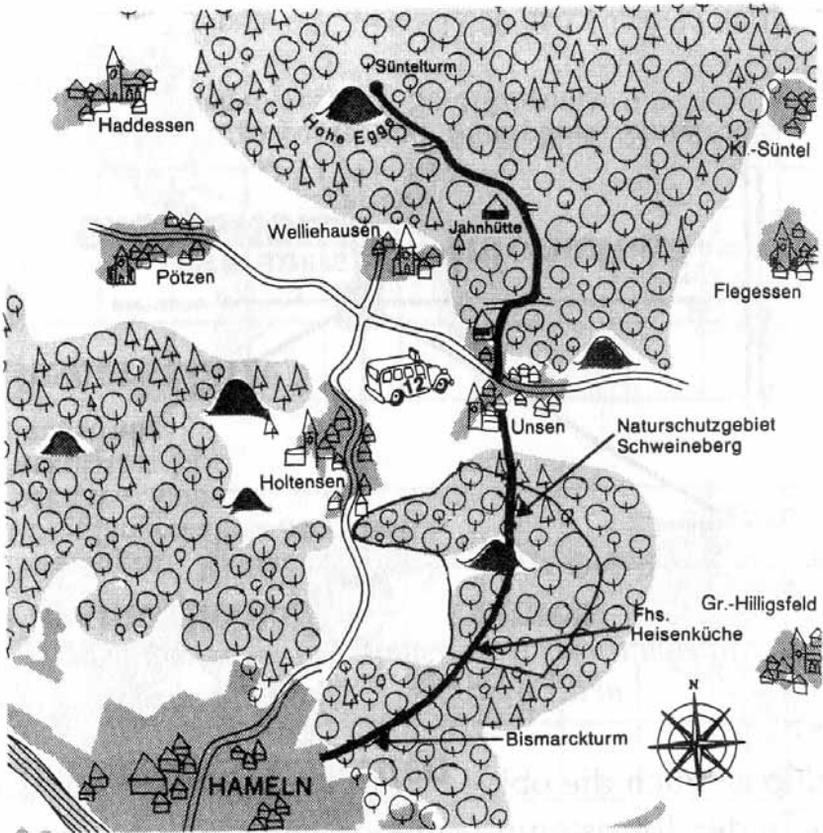


*Der Mittelpunkt des Universums: Moustiers-Sainte Marie
in der französischen Provence*

Deshalb ist auch die obige Karte, die mir eine freundliche Dame in der Touristeninformation eines südfranzösischen Bergdorfes einmal in die Hände drückte, als eher läßliche Sünde anzusehen.

Auch andere Aspekte ihrer Umwelt zeigen Touristenkarten oft durch eine rosarote Brille. Versuchen Sie doch einmal, auf den Karten, wie sie Urlaubsorte an Besucher verteilen, das lokale Heizwerk oder die Ruine der abgebrannten Ziegelei zu finden ...

Da wir aber in Touristenprospekten nur selten nach der Wahrheit suchen, sind solche Sünden zu verzeihen. Eindeutiger ist die Sache, wenn die Leser solche Übertreibungen nicht erwarten, wenn der Rahmen vorgegeben ist, und wir markieren wie der KGB auf alten Rußlandkarten gewisse Orte gar nicht oder falsch. Und das macht nicht allein der KGB: Versuchen Sie doch einmal, auf einer ame-



Wo bitte geht's zur Mülldeponie?

rikanischen Straßenkarte den Zugang zum »Camp David« des amerikanischen Präsidenten ausfindig zu machen.

Das größte Betrugspotential entfalten Karten aber durch das Gleichsetzen von *Fläche* mit Größe, Macht und militärischer Bedrohung. Wir alle kennen die Karten aus der Zeit des Kalten Krieges, die den Westen der riesigen, durch die sogenannte »Merkatorprojektion« nochmals weiter aufgeblasenen Landmasse der ehemaligen Sowjetunion gegenüberstellen, die die armen kleinen Nato-Staaten fast erdrückt, oder Karten des Nahost-Konfliktes, in dem das kleine Israel in einem Meer von Feinden geradezu ertrinkt. Mit diesen Mitteln haben schon die Nazis

im Zweiten Weltkrieg die Engländer als Aggressoren porträtieren wollen - denn wer könnte sich erdreisten, ein Riesengebilde wie das britische Empire anzugreifen auch nur zu versuchen.



*Aus der in Amerika verkaufte Nazi-Propagandaschrift
Facts in Review vom 5. Februar 1940*

Heimtückischer, weil nicht so leicht durchschaubar, sind solche Karten, wenn sie die Flächen je nach bestimmten Daten unterschiedlich färben. Angenommen etwa, wir wollen wissen, wie dicht die Bundesbürger und -bürgerinnen zusammenwohnen. Die Karte auf der nächsten Seite stellt dies graphisch für die sechzehn Bundesländer dar: Bei mehr als 200 Einwohnern pro Quadratkilometer ist das Land schwarz eingefärbt, sonst weiß.



Eine typische Dichtekarte: die Bevölkerungsdichte der Bundesrepublik Deutschland. Alle Länder mit mehr als 200 Menschen pro Quadratkilometer sind schwarz eingefärbt, die anderen weiß

Wie wir sehen, geht fast die ganze Republik in weiß: in den meisten Bundesländern wohnen weniger als 200 Menschen auf einem Quadratkilometer, weit weniger als durchschnittlich etwa in Belgien (320), Holland (350) oder England (360) - wir scheinen in Deutschland sehr viel Platz zu haben.

In Wahrheit leben in Deutschland rund 220 Menschen auf einem Quadratkilometer, genauso viele wie in England (wenn wir Schottland, Wales und Nordirland dazu nehmen). Daß die Karte etwas anderes zu sagen scheint, liegt daran, daß hierzulande wie auch anderswo so viele Menschen in den Ballungszentren leben, im Ruhrgebiet, um Stuttgart, Hamburg, Frankfurt, Leipzig, Dresden. Die

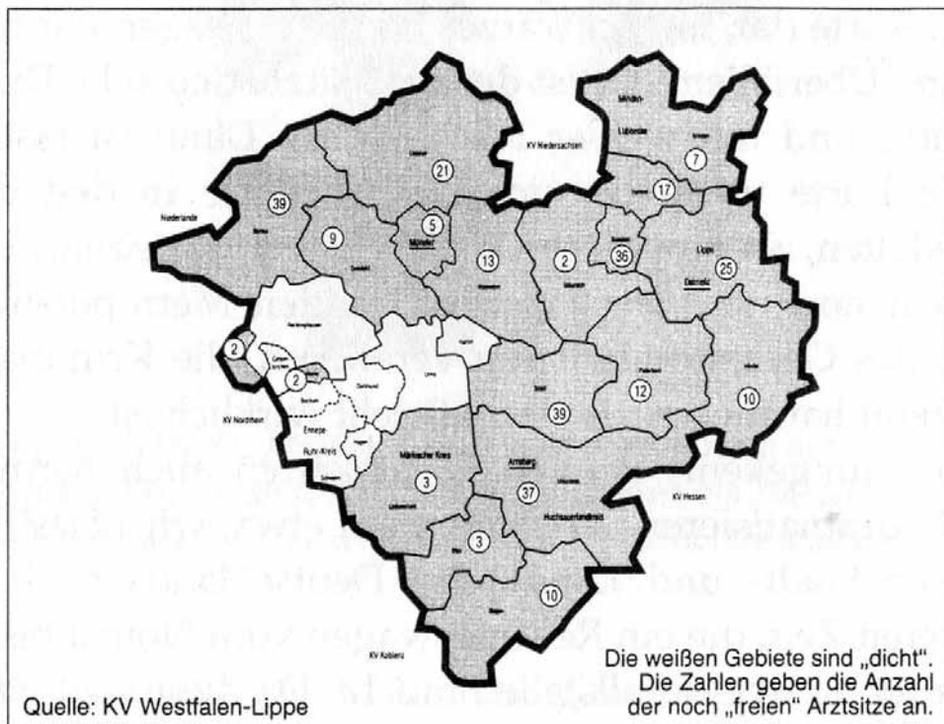
Bundesbürger drängen sich also in wenigen dicht besiedelten Regionen; würden wir uns gleichmäßiger auf die Republik verteilen, wäre die ganze Karte schwarz.

Solche »Dichtekarten« tauchen viele Daten in ein falsches Licht. Angenommen, wir stellen die räumliche Verteilung der Kriminalität, gemessen durch die Banküberfälle pro Jahr und Land- bzw. Stadtkreis in einer solchen Dichtekarte dar, mit schwarzer Fläche bei sagen wir mehr als drei Überfällen (das ist der geschätzte deutsche Durchschnitt), und mit weißer Fläche sonst. Dann ist fast die ganze Karte weiß! Da sich die Überfälle in den dicht besiedelten, aber an Fläche kleinen Städten konzentrieren, können noch so viele Überfälle in den Metropolen das Weiß des Gesamtbilds kaum verändern, die Kriminalität erscheint harmloser, als sie vielleicht wirklich ist.

Und umgekehrt können Dichtekarten auch harmlose Daten dramatisieren. Angenommen etwa, wir klassifizieren die Stadt- und Landkreise Deutschlands nach der mittleren Zeit, die ein Rettungswagen vom Notruf bis zur Ankunft an der Unfallstelle braucht. Da diese Zeit naturgemäß in dünn besiedelten Gebieten länger ist, kommt so leicht die Illusion zustande, die ganze Republik wäre mit Rettungsdiensten unterversorgt (während wir in Wahrheit eher zu viele als zu wenige dieser bei näherem Hinsehen gar nicht so selbstlosen Samariter haben).

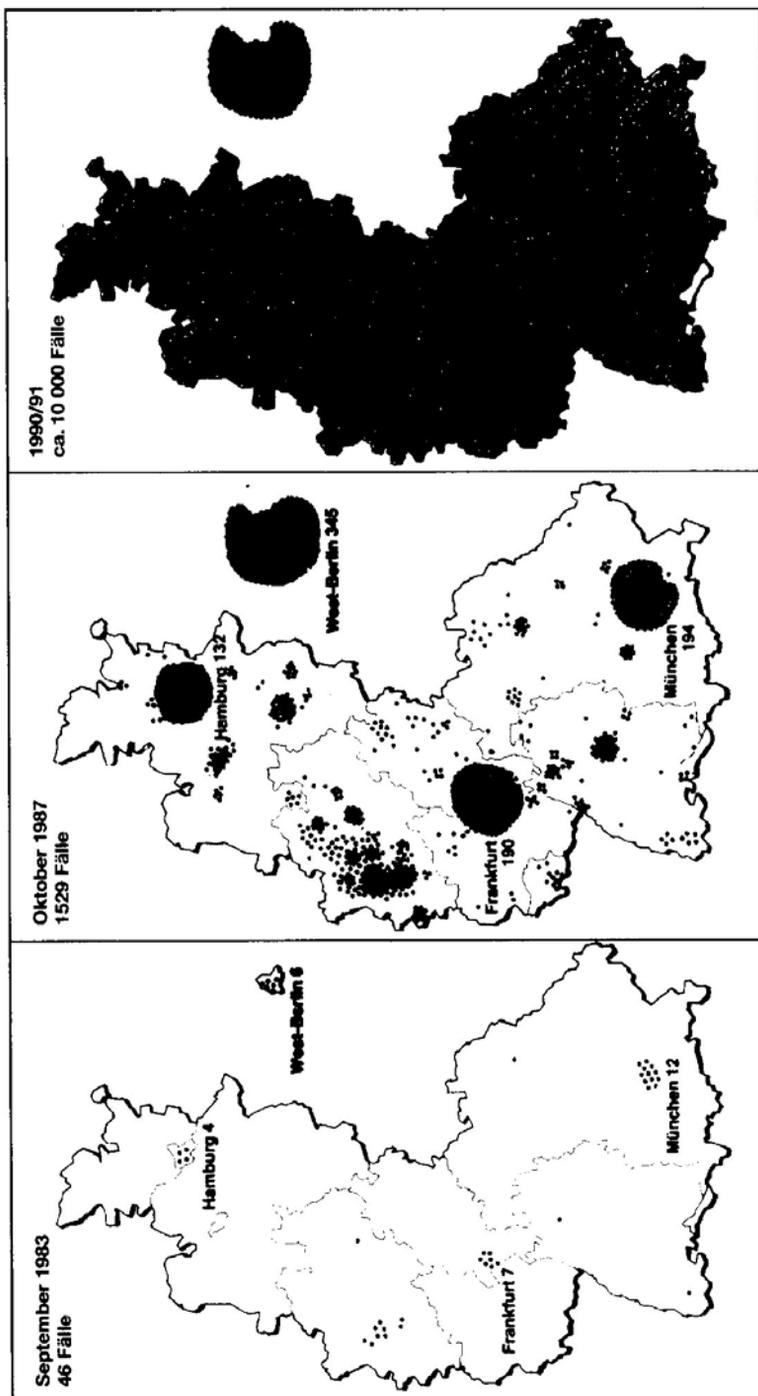
Die nächste Dichtekarte ist eher zu den Verharmlosern zu rechnen. Sie gibt an, in welchem Landkreis der Kassenärztlichen Vereinigung Westfalen-Lippe sich noch Kassenärzte niederlassen können. Sie zeigt fast die gesamte Fläche offen = grau. Aber für frisch approbierte Ärzte ist das nur ein schwacher Trost: Zwar zeigt die Karte nur für wenige Kreise dicht = weiß, aber davon hat der hoff-

nungsvolle Nachwuchsmediziner nichts. Denn Ärzte behandeln keine Quadratkilometer, sondern Menschen, und die wohnen in den weißen Kreisen Dortmund, Hagen, Bochum, Unna, Recklinghausen - die Dominanz der grauen alias noch offenen Niederlassungsgebiete führt völlig in die Irre.



Eine Dichtekarte für die Versorgung mit niedergelassenen Ärzten

Die letzte Graphik zeigt die Ausbreitung von AIDS (sehr anschaulich, so wie den »Schwarzen Tod« des Mittelalters, schwarz). Kein Mensch käme bei solchen Bildern auf den Gedanken, daß pro Jahr in Deutschland weit mehr Menschen an der Grippe sterben als an AIDS.



Der Siegeszug des Schwarzen Todes findet vor allem in den Redaktionsstuben der deutschen Medien statt

Literatur

Eine systematische Übersicht der Lüge mit Karten finden Sie bei Mark Monmonier: *Eins zu einer Million. Die Tricks und Lügen der Kartographen*, Basel 1996 (Birkhäuser). Für weitere Beispiele und eine Anleitung, wie man es richtig macht, siehe auch mein eigenes Buch *So überzeugt man mit Statistik*, Campus Verlag 1994, besonders Kapitel 9: Kartogramme. Die AIDS-Graphik stammt aus dem *Stern*, Heft 48 vom 19.11.1987.

14. Korrelation kontra Kausalität

*»Während einer nur Zahlen und Zeichen im Kopf hat,
kann er nicht dem Kausalzusammenhang auf die Spur kommen.«*

Schopenhauer

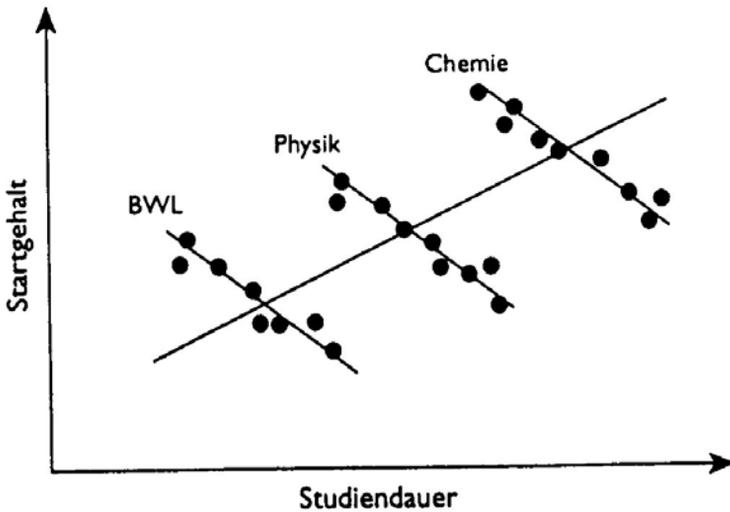
Unter dem Stichwort »Methusalems machen Kasse« lese ich im *Handelsblatt*: »Ein langes Studium zahlt sich in barer Münze aus. Zu diesem überraschenden Ergebnis kommt eine Studie über die Einstiegsgehälter von Berufsanfängern, für die die Deutsche Gesellschaft für Personalführung 44 Firmen befragt hat.«

Das überrascht mich in der Tat. Seit Jahren bete ich meinen Studenten vor: Laßt euch nicht so hängen, zieht zügig euer Studium durch, jedes Semester über der Regelstudienzeit drückt aufs Gehalt. Und nun das!

Wer mir bis hierher gefolgt ist, ahnt natürlich: das *Handelsblatt* ist einem Trugschluß aufgesessen (den es, das sei hier zur Ehrenrettung angemerkt, auch selber korrigiert). Dieser Trugschluß, einer der häufigsten in der Statistik überhaupt, ist der falsche Schluß von Korrelation auf Kausalität, die Unterstellung, daß, wenn zwei Variablen parallel verlaufen, die eine die andere am Zügel führen muß.

Dieser Schluß ist manchmal richtig, manchmal falsch. Bei der Studiendauer ist er falsch.

Es ist eine Illusion, daß langes Studieren ein hohes Starteinkommen fördert. In Wahrheit ist das Gegenteil der Fall. In den akademischen Fächern, in denen ich selbst



Hält man das Studienfach konstant, gehen die Startgehälter mit steigender Dauer des Studiums zurück. Wirft man alle Fächer in einen Topf, scheint das Startgehalt dagegen mit einem langen Studium zu steigen

aktiv oder passiv an der Ausbildung beteiligt war oder bin, haben »Bummelstudenten« eher bescheidene Aussichten auf eine gute Anstellung; in aller Regel geht das Startgehalt mit steigender Semesterzahl zurück.

Daran ändern auch das *Handelsblatt* und seine Studien-Statistik nichts. Die dort gemeldete positive Korrelation von Studiendauer und Gehalt liegt einfach daran, daß alle Fächer in einen Topf geworfen sind. Und in diesem großen Topf haben bzw. hatten diejenigen Hochschulabsolventen und -absolventinnen mit den langwierigsten Fächern, wie Chemie und Medizin, die höchsten Startgehälter. Aber diese hohen Startgehälter hatten sie nicht wegen der Länge, sondern wegen der Schwere ihres Studiums; die Länge als solche trägt dazu überhaupt nichts bei. Im Gegenteil, auch in der Chemie ist ein schneller Abschluß einen Bonus wert, nur ist dieser schnelle Abschluß hier viel schwieriger.

Ganz anders die Betriebswirtschaft. Die hier geforderten Fähigkeiten sind eher praktischer als akademischer Natur und in kürzerer Zeit vermittelbar (falls überhaupt: das Zeug zum Vorstandssprecher der Deutschen Bank hat man oder hat man nicht, und wenn nicht, nützt auch das beste Studium nichts). Daher schließen auch Diplom-Betriebswirte (FH) mit ihren sechs Semestern Studium die Liste der Startgehälter nach unten ab, aber nicht, weil sie zu schnell studieren, sondern weil hier ein Überangebot besteht. Auch hier wird, wie in der Chemie, ein schnelles Studium belohnt, wenn auch von einer kleineren Basis aus.

Hält man aber die dritte Variable, nämlich das Studienfach, konstant, ist der Zusammenhang zwischen Studiedauer und Gehalt in allen Fächern negativ, so wie im nächsten Schaubild dargestellt.

In abgeschwächter Form tritt dieses Phänomen auch bei der bekannten Korrelation zwischen Rauchen und frühem Sterben auf. Hier wird in der Regel das Rauchen als Ursache für frühes Sterben unterstellt. Wenn man der Bundeszentrale für gesundheitliche Aufklärung glauben darf, stirbt z. B. ein 30jähriger Raucher mit einem Konsum von ein bis zwei Päckchen Zigaretten am Tag sechs Jahre früher als ein Nichtraucher, laut Wissenschaftlichem Institut der Ortskrankenkassen sogar 12 Jahre früher.

Daraus folgt jedoch noch nicht, daß alle diese Jahre dem Tabak anzulasten sind. Denn Raucher werden auch öfter als Nichtraucher ermordet oder von Bussen überfahren, und das kann mit dem Rauchen als solchem wohl kaum direkt zusammenhängen. Vielmehr ist hier eine Hintergrund-Variable »Charakter« bzw. »Raucherpersönlichkeit« im Spiel, die viele gefährliche Gewohnheiten glei-

chermaßen fördert, so daß unsere modernen Raucher auch dann früher, wenn auch mit einem kleineren Vorsprung, sterben würden, hätte Kolumbus Amerika und den Tabak nie entdeckt.

Genauso stützt auch die hohe Korrelation von Schulbildung und Einkommen für sich allein noch keineswegs die These, daß Bildung bare Münze bringt. Hier wird eine dritte Variable »Ehrgeiz« und »Talent« vernachlässigt, die Schul- und Berufserfolg in gleicher Weise positiv berührt. Mit anderen Worten, viele erfolgreiche Menschen hätten auch ohne Abitur Erfolg gehabt (genauso wie viele Raucher auch ohne Tabak früher sterben würden). Berufserfolg und Bildung gehen zum Teil auch deshalb Hand in Hand, weil beide sich bei Menschen mit bestimmten Eigenschaften häufen.

Die Sache mit den Klapperstörchen

Durch das geschickte Auslassen solcher Hintergrundvariablen verdreht man jede Wahrheit leicht ins Gegenteil. Mitte der 70er Jahre etwa fand sich eine große amerikanische Universität der Diskriminierung gegen Frauen angeklagt. Sie lasse prozentual mehr männliche als weibliche Bewerber zu und habe gefälligst etwas für die Emanzipation zu tun.

»Falsch!« konterte die beschuldigte Universität. »Wir lassen allem Männerüberschuß zum Trotz in allen Fällen mehr Frauen zu als Männer!«

Und so war es auch. Der Frauenanteil insgesamt war klein nicht wegen der Diskriminierung des weiblichen

Geschlechts, sondern weil sich Frauen auf Fächer kaprizierten, in denen ein großer Bewerberüberhang bestand. Allein deshalb blieben relativ mehr Frauen als Männer vor der Tür. In jedem einzelnen Fach waren sie bei der Zulassung aber *erfolgreicher* als Männer.

Wie schon bei den Bummelstudenten entsteht auch hier ein falscher Eindruck durch Übersehen der Hintergrundvariablen »Studienfach«. Hätten sich Frauen ebenso zahlreich wie Männer für Fächer beworben, in denen der Zugang leichter war, wären sie auch stärker an der Universität vertreten.

Aus dem gleichen Grund können auch die Klagen unserer Ärzte über sinkende Einkommen nicht ganz überzeugen. Wenn man etwa der Kassenärztlichen Bundesvereinigung glauben darf, ging das durchschnittliche Einkommen der niedergelassenen Ärzte in der Bundesrepublik in den letzten Jahren um mehrere Prozent zurück. In Wahrheit nahm jedoch das Einkommen der Ärzte, von Ausnahmen abgesehen, weiter zu. Daß trotzdem das Durchschnittseinkommen sank, liegt einfach an der wachsenden Zahl junger Ärzte mit noch kleiner Praxis, die erst am Anfang ihrer Großverdiener-Karriere stehen. Mit anderen Worten, auch wenn in jeder Altersklasse das Einkommen steigt, kann das Durchschnittseinkommen trotzdem sinken, nämlich wenn die Belegung der relativ einkommensschwachen Jahrgänge steigt.

Oder nehmen wir die Krebsgefahr. Hier müssen wir fast täglich in den Medien hören oder lesen, wie uns diese Menschheitsgeißel von Jahr zu Jahr vermehrt bedroht. In Wahrheit ist jedoch das Gegenteil der Fall: Die Wahrscheinlichkeit, an Krebs zu sterben, hat quer durch alle Altersklassen und für Männer wie für Frauen gleicher-

maßen in den letzten dreißig Jahren *abgenommen*. Daß dennoch immer mehr von uns an dieser Krankheit sterben (aktuell rund 23 Prozent verglichen mit weniger als 10 Prozent zu Anfang des Jahrhunderts), liegt allein daran, daß wir im Durchschnitt immer älter werden. Auch hier wird eine Hintergrundvariable, nämlich Lebensalter, übersehen. Daß die Zeitgenossen Kaiser Wilhelms so selten Krebs bekamen, lag nicht an ihrem gesunden Lebenswandel, sondern daran, daß sie im Durchschnitt schon mit 45 Jahren an anderen Krankheiten gestorben sind.

Solche übersehenen Hintergrundvariablen produzieren Nonsenskorrelationen zuhauf. Angefangen bei den Klapperstörchen, deren Zahl hoch positiv mit den bundesdeutschen Geburten korreliert, über die Zahl der unverheirateten Tanten eines Menschen und den Kalziumgehalt seines Skeletts (negative Korrelation), Heuschnupfen und Weizenpreis (negative Korrelation), Schuhgröße und Lesbarkeit der Handschrift (positive Korrelation), Schulbildung und Einkommen (positive Korrelation) bis zu Ausländeranteil und Kriminalität (positive Korrelation) spannt sich ein weiter Bogen eines falsch verstandenen bzw. absichtlich mißbrauchten Korrelationsbegriffs.

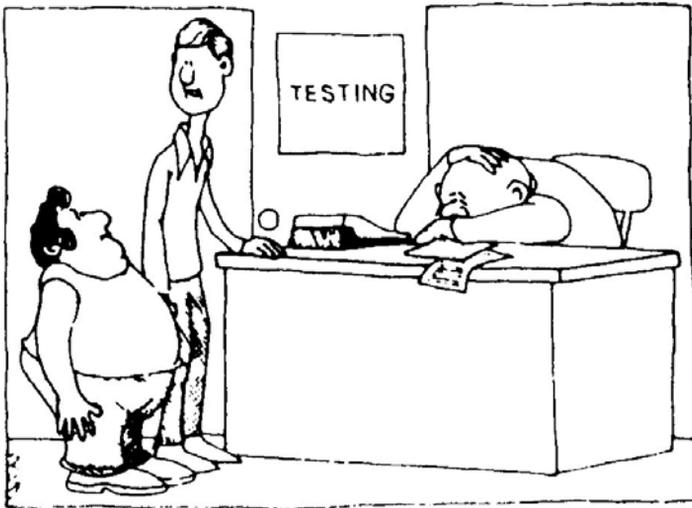
Bei den Geburten und Klapperstörchen macht das weiter nichts. Zwar sind diese tatsächlich in manchen Gegenden eng korreliert, aber trotzdem glaubt deswegen niemand, daß der Storch die Kinder bringt. Die positive Korrelation von Ausländeranteil und Kriminalität in den Gemeinden der Bundesrepublik ist schon gefährlicher; hier unterstützt die Statistik unter Umständen nur ein ebenso populäres wie falsches Vorurteil, denn große Gemeinden ziehen sowohl Ausländer wie Kriminelle an.

Die negative Korrelation von Weizenpollenallergien

und Weizenpreis, über die aus den Staaten des mittleren Westens der USA berichtet wird, entsteht dagegen durch das Wetter: Wenn der Weizen wegen guten Wetters gut gedeiht und heftig blüht, sinkt, wie jeder Ökonomiestudent im ersten Semester lernt, aufgrund des hohen Angebots der Preis. Bei unverheirateten Tanten und Kalziumgehalt wie auch bei Schuhgröße und Handschrift ist dagegen das Alter der jeweiligen Person der Bösewicht: Junge Menschen haben mehr unverheiratete Tanten als ältere, dafür in den Knochen weniger Kalzium. Ältere Schüler haben größere Füße und eine schönere Handschrift (und ältere Männer weniger Haare, aber mehr Geld ...).

Die Spitze aller vernachlässigten Hintergrund-Variablen hält jedoch die Zeit. Zeitreihendaten wie Volkseinkommen, Staatsverschuldung, Aktien- oder Preisindex, Studentenzahlen, Auslandsurlauber, Konsum von Südfrüchten oder Mitgliedschaft im Deutschen Fußballbund zeigen aus verschiedenen Gründen oft einen monotonen Trend. Weist dieser Trend bei beiden Variablen in die gleiche Richtung (ob nach oben oder unten, ist dabei egal), so korrelieren die Variablen positiv, und weisen die Trends in verschiedene Richtungen, so korrelieren die Variablen negativ, unabhängig davon, ob ein kausaler Zusammenhang besteht oder nicht. Dieser gemeinsame Trend ist etwa für die Korrelation von Klapperstörchen und Geburten in der Bundesrepublik verantwortlich, denn beide Variablen nahmen lange Zeit im Gleichschritt ab. Hier sprießen die Korrelationen nur so aus den statistischen Jahrbüchern hervor. Wer zum Beispiel glaubt, an allem Übel dieser Welt sei nur der Reichtum schuld (eher geht ein Kamel durchs Nadelöhr ...), findet hier reichlich Bestätigung. Da in allen westlichen Industrienationen seit

mehreren Jahrzehnten das Volkseinkommen steigt, ist dieses automatisch mit allen Variablen, die in dieser Zeit ebenfalls gestiegen sind, wie Mord und Totschlag, Krebs, Verkehrsunfällen, Alkoholkonsum und Ehescheidung, bestens korreliert.



»Er sagt, wir ruinieren seine ganze schöne Korrelation zwischen Größe und Gewicht.«

Der Punkt ist: Eine positive oder negative Korrelation, das heißt ein wie auch immer definierter Gleich- oder Widerklang von zwei Variablen, kann auch andere Ursachen als eine Abhängigkeit in der einen oder anderen Richtung haben. Bei zwei blonden Geschwistern wird auch niemand sagen, weil das erste blond ist, ist auch das zweite blond; beide haben diese Eigenschaft von den Eltern geerbt. Genauso kann eine Korrelation auch anderswo statt durch wechselseitige Abhängigkeit durch einen verborgenen Regisseur im Hintergrund, alias durch eine vergessene dritte Einflußgröße entstehen, welche die beiden Ausgangsvariablen in die gleiche oder auch in entgegengesetzte Richtungen zwingt.

Viele Korrelationen sind auch ein reines Kind des Zufalls und haben mit keiner Kausalität, weder zwischen den jeweiligen Variablen noch zwischen diesen und irgendeiner anderen, etwas zu tun. Einige Jahrzehnte verliefen zum Beispiel Aktienkurse und Rucksäcke erstaunlich parallel, und es hat auch nicht an Erklärungen dafür gefehlt. Genauso erwarteten amerikanische Börsengurus lange Zeit gespannt die »Superbowl«, das Finale der amerikanischen Football-Saison, denn bei einem Sieg der »National Football Conference« schloß das Börsenjahr mit einem Gewinn, und bei einem Sieg der »American Football Conference« mit einem Verlust; auch hier besteht nicht der geringste Zusammenhang.

Ehemänner leben länger

Selbst wenn der Zufall oder eine dritte Variable als Verursacher einer Korrelation ausscheiden, ist damit immer noch nichts über die *Richtung* einer möglichen Kausalität gesagt. Die Bewohner der Neuen Hebriden glaubten bis vor kurzem, daß Läuse gut für die Gesundheit seien. Die Korrelation zwischen diesen Variablen war in der Tat beeindruckend, wenn man Darrell Huff glauben darf, der diese Statistik ausgegraben hat. Läuse und Gesundheit traten gern im Tandem auf, gesunde Insulaner hatten Läuse, kranke nicht. Läuse wurden regelrecht als Haustiere behandelt, jeder hätte gerne möglichst viel davon gehabt.

In Wahrheit verlief die Ursache-Wirkung-Beziehung gerade umgekehrt. Nicht die Läuse vertreiben die Krank-

heit, sondern die Krankheit vertreibt die Läuse. Kranke Menschen haben häufig Fieber, und das wird Läusen, die die Hitze hassen, schnell zu heiß.

Für solche Trugschlüsse müssen wir aber nicht auf die Neuen Hebriden gehen.

Unter der Schlagzeile »Gesundes Meditieren« meldet die *Bild-Zeitung*, daß Meditieren den Herzinfarkt bekämpfen soll. Nach einer amerikanischen Studie hätten 87 Prozent der Meditierenden »eine sehr niedrige Herzinfarkt-Rate« gehabt. Auch hier kann die Ursache dort liegen, wo die Meldung sie vermutet, vielleicht aber auch anderswo: Menschen, für die Streß und Herzinfarkt Fremdwörter sind, meditieren gern.

Daher irren vermutlich auch die Horrorgeschichten unserer Illustrierten über Unterschiede in der regionalen Sterblichkeit. So sterben etwa in vielen Kurorten mit besonders guter Luft mehr Menschen als in Industriegebieten. Jedoch bringt nicht das gute Klima die Menschen um, sondern umgekehrt: Kranke und Menschen, die bald sterben, ziehen gerne in ruhige Orte mit gesunder Luft. Oder um mit der *Ditmarscher Landeszeitung* zu sprechen: »Wenn sich die Sterbefälle von 54 auf 65 erhöhten, so war diese Tatsache nicht etwa auf eine höhere Sterblichkeit zurückzuführen, sondern auf die Erweiterung des Brunsbüttelkooger Krankenhausbetriebs.«

Oder was sollen wir von der folgenden Pressemeldung halten (aus der *Bild-Zeitung*, aber ähnlich fast jeden Sommer auch in anderen Zeitungen zu finden):

»Ehemänner leben länger - Eingefleischte Junggesellen zwischen 45 und 54 Jahren sollten doch den Sprung ins Abenteuer Ehe wagen. Laut Statistik der Epidemiologischen Fakultät an der Universität Kalifornien werden nämlich 23 Prozent der

ledigen Männer dieser Altersgruppe in den nächsten zehn Jahren sterben. Das Todesrisiko verheirateter Männer liegt dagegen nur bei 11 Prozent.«

Die reinen Zahlen hinter dieser Meldung treffen durchaus zu: Ehemänner leben im Durchschnitt wirklich länger als Junggesellen, fünf Jahre etwa in der Bundesrepublik und fünfzehn Jahre in Japan - um soviel leben, wenn wir den Demographen glauben dürfen, Männer mit angetrauten Ehefrauen wirklich länger.

Nicht zutreffend ist aber die Erklärung, die wir dafür in den Medien oft lesen: daß Ehemänner länger leben, *weil* sie Ehemänner sind. In Wirklichkeit verläuft die Kausalrichtung genau umgekehrt.

Ehemänner leben nicht deshalb länger, weil sie Ehemänner sind, sondern sie werden Ehemänner, weil sie länger leben. Faktoren wie Reichtum und Gesundheit, die innerhalb wie außerhalb der Ehe das lange Leben fördern, fördern gleichzeitig auch die Ehe, und Faktoren wie Armut oder Krankheit, die innerhalb wie außerhalb der Ehe das Leben eher verkürzen, sind der Ehe eher abträglich. Oder wie einer meiner Studenten einmal formulierte: »Die phänomenale Schlagzeile >Ehemänner leben länger<, welche jedem verheirateten Mann eine gewisse Form der Genugtuung und Sicherheit geben wird, erweist sich nach der Lektüre des Artikels als haltlos.«

Damit will ich nicht bestreiten, daß auch die Ehe selbst durchaus das Überleben fördern kann. So mögen etwa verheiratete Männer gesünder essen als unverheiratete, ganz einfach weil getreu dem hergebrachten Rollenmuster Frauen immer noch besser und vor allem gesünder kochen. Und wie jeder Psychologe weiß, kann auch die emotionale Stabilität einer glücklichen Ehe die Gesundheit und die

Lebenserwartung beider Partner positiv berühren. Diese widerstreitenden Kausalbeziehungen - in der Demographie als »marriage selection« und »marriage protection« bekannt - sind nicht leicht zu trennen und deswegen noch Gegenstand von wissenschaftlichen Debatten. Worauf es mir hier ankommt, ist allein: Betrachten wir die Wahrscheinlichkeit, binnen zehn Jahren zu sterben, für zwei Gruppen von Männern, verheiratete und nicht verheiratete, dann sagen die Zahlen für sich allein genommen über Wohl und Wehe des Ehelebens überhaupt nichts aus.

Literatur

Ein immer noch lesenswerter Klassiker zu Scheinkorrelationen ist G. U. Yule: »Why Do We Sometimes Get Nonsense Correlations Between Time Series? - A Study in Sampling and the Nature of Time Series«, *Journal of the Royal Statistical Society* 89, 1926, S. 1-64; siehe auch mein eigenes Buch *Statistik verstehen* (Campus Verlag 1992). Die Debatte zur Geschlechter-Diskriminierung bei der Universitätszulassung wird ausführlich aufgerollt in P. J. Bickel, E. A. Hammel und J. W. O'Connell: »Sex Bias in Graduate Admissions: Data from Berkeley«, *Science* 187, 1975, S. 398-404. Zu »Simpson's Paradox« siehe auch das Kapitel über bedingte Wahrscheinlichkeiten in meinem Buch *Denkste! Trugschlüsse aus der Welt des Zufalls und der Zahlen* (Campus Verlag 1995). Zur Lebenserwartung von Rauchern und zum Zusammenhang von Rauchen, Selbstmord und Unfällen verweise ich auf mein Buch *Wir kurieren uns zu Tode* (Ullstein Verlag 1997), und zum Thema Ehe und Lebenserwartung siehe auch N. Goldmann: »Marriage Selection and Mortality Patterns«, *Demography* 30, 1993, 189-208.

15. Fluglärm erzeugt AIDS

»Von der Wahrscheinlichkeitsrechnung her
wird das eine oder andere Tor fallen.«

Christof Daum, Fußballtrainer

Ein Mann kommt zum Arzt.

»Also die Lage ist ziemlich ernst. Sie sind sehr krank. Statistisch gesehen überleben neun von zehn Patienten diese Krankheit nicht.«

Der Mann erbleicht.

»Sie haben aber Glück«, beruhigt ihn der Arzt. »Ich hatte schon neun Patienten mit den gleichen Symptomen, und die sind alle tot.«

»Verrechnet hat sich in der Nacht zum Donnerstag in Frankfurt ein 44jähriger Systemanalytiker und Mathematiker, der von Beamten einer Polizeistreife gebeten worden war, wegen starken Alkoholgenusses sein Fahrzeug stehen zu lassen«, teilt der Berliner *Tagesspiegel* mit. »Der Wissenschaftler versicherte, er werde seine Frau anrufen und sie bitten, ihn abzuholen.« Dann schloß er sein Auto ab und ging weg. Als die Beamten aber kurz darauf nochmals an der gleichen Stelle vorbeikamen, sahen sie unseren Freund am Steuer seines Wagens davonfahren. Anhalten, Blutprobe, Führerscheinentzug. »Mit einer zweiten Kontrolle hatte ich nicht gerechnet«, entschuldigte sich der Delinquent. »Vorhin wurde ich zum allerersten Mal überhaupt kontrolliert, und nach der Wahrscheinlich-

keitsrechnung findet die nächste Kontrolle erst in hundert Jahren statt...«

Vermutlich hatte unser Freund wie folgt gedacht: »Mit einer Wahrscheinlichkeit von - sagen wir - 1 zu 200 werde ich heute kontrolliert. Demnach beträgt die Wahrscheinlichkeit für *zwei* Kontrollen in derselben Nacht $(1/200)^2 = 1: 40.000$. Das kommt im Mittel alle 100 Jahre vor, also kann ich ganz beruhigt sein. Das müßte schon ein großer Zufall sein, wenn die Polizei mich heute nochmals kontrolliert.«

So funktionieren Wahrscheinlichkeiten aber nicht - unser Mathematiker hatte die sogenannte »bedingte« Wahrscheinlichkeit mit der unbedingten alias »normalen« Wahrscheinlichkeit eines Ereignisses verwechselt. Zwei Kontrollen in derselben Nacht sind in der Tat sehr unwahrscheinlich, aber die bedingte Wahrscheinlichkeit für zwei Kontrollen, gegeben, wir sind schon einmal aufgefallen, ist weit größer; bei unabhängigen Kontrollen ist sie genauso groß wie die Wahrscheinlichkeit für eine einzige Kontrolle.

Beispiele für unabhängige Ereignisse sind: »Im Jahr X wird Deutschland Fußball-Weltmeister« und »im Jahr X+1 gibt es ein Erdbeben auf den Fidschi-Inseln«. Die Wahrscheinlichkeit für das erste ist - sagen wir - 1:10, die Wahrscheinlichkeit für das zweite 1: 100, die Wahrscheinlichkeit für beides zusammen also 1: 1.000 (bei unabhängigen Ereignissen ist die Wahrscheinlichkeit für beide zusammen gerade das Produkt der einzelnen Wahrscheinlichkeiten). Kein Mensch käme hier auf den Gedanken: »Aha, die Deutschen sind Fußballweltmeister geworden. Also sinkt die Wahrscheinlichkeit für ein Erdbeben auf den Fidschi-Inseln von 1: 100 auf 1: 1.000.« Diese Wahrscheinlichkeit behält den gleichen Wert 1:100, den sie

auch vorher hatte, ganz gleich, wie die Deutschen Fußball spielen.

Oft wird diese Konsequenz der Unabhängigkeit aber - aus Absicht oder Unkenntnis - übersehen. Wenn etwa die *Frankfurter Allgemeine Zeitung* über einen Lottospieler schreibt: »Tippt er jede Woche im Samstagslotto zwei Reihen, dann ist der Volltreffer für ihn unausweichlich, vorausgesetzt, er wird 135.000 Jahre alt«, so stimmt das einfach nicht. Wenn die *FAZ* Pech hat, tippt ihr Lottospieler bis zum Ende aller Tage ohne eine Mark Gewinn. Zwar ist die Wahrscheinlichkeit eines Hauptgewinns, wie die *FAZ* ganz richtig ausgerechnet hat, 1:13.983.816, aber das bedeutet keinesfalls, daß man für einen sicheren Volltreffer 13.983.816 mal oder (bei zwei Versuchen pro Woche) 135.000 Jahre spielen muß. Bei unabhängigen Ziehungen - und die verschiedenen Ziehungen der Lottozahlen sind ganz klar unabhängig voneinander - ist es sowohl möglich, wenn auch bisher nur einmal vorgekommen, daß eine Kombination ein zweites Mal erscheint, als auch, daß sie nie gezogen wird.

Genauso ist es falsch, wenn die *Frankfurter Rundschau* in einem mit »Kernschmelzunfall auch bei uns wahrscheinlich« betitelten Artikel schreibt: »Ein GAU in 10.000 Reaktorjahren. Bisher 3.500 Jahre erreicht« und damit meint, daß bei einer Wahrscheinlichkeit von 1:10.000 für einen Unfall pro Jahr und Kernkraftwerk in 10.000 Jahren ein Unfall unausweichlich sei. Wenn die Wahrscheinlichkeit von 1:10.000 für einen GAU pro Jahr tatsächlich stimmt, folgt daraus zwar eine mittlere Wartezeit bis zum nächsten Unfall von 10.000 Jahren, wie man mit mathematischen Methoden berechnen kann, aber diese mittlere Wartezeit bleibt unabhängig von der Betriebsdauer immer gleich.

Ein GAU ist zwar nicht auszuschließen, aber die Wahrscheinlichkeit dafür nimmt mit der Betriebsdauer *nicht* zu (wegen immer besserer Sicherheitstechnik sogar eher ab). Bei unabhängigen Ereignissen hängt die Wahrscheinlichkeit *nicht* davon ab, was vorher war, sie treten ein oder auch nicht, aber ob sie eintreten, wissen wir erst hinterher.

Diese simple Wahrheit vergessen wir aber oft. »Großmütterchen, zero ist eben herausgekommen«, läßt Dostojewski seinen »Spieler« sagen, »also wird es lange Zeit nicht herauskommen. Sie werden viel verlieren, wenn Sie bis dahin immer auf zero setzen wollen. Warten Sie lieber noch ein Weilchen.« Und wie alle Dostojewski-Kenner wissen, kam kurz später noch mal zero, und das unbelehrbare Großmütterchen verließ die Wiesbadener Spielbank mit weit mehr Geld, als sie hereingetragen hatte.

Viele Menschen denken wie Dostojewskis Spieler. Tausende von Lottospielern suchen weiter solche Zahlen, die lange nicht gezogen worden sind, in der Hoffnung, ihre Chancen nähmen dadurch zu. Die Mutter dreier Jungen glaubt, daß ihr viertes Kind ein Mädchen werden wird (denn vier Jungen kommen nur einmal in sechzehn Fällen vor). Von Veteranen des Ersten Weltkrieges berichtet man, sie hätten gern in frischen Granattrichtern Deckung gesucht, weil es äußerst unwahrscheinlich sei, daß ein und derselbe Punkt zweimal getroffen wird. Und jeder kennt die Anekdote von dem Flugzeugpassagier, der sicherheits halber immer eine Bombe mit sich führt, »denn die Wahrscheinlichkeit für zwei Bomben in ein und demselben Flugzeug ist ja praktisch null«.

In all diesen Fällen wird der Fehler unseres Mathematikers und der Frankfurter Zeitungen wiederholt: Die Wahrscheinlichkeit etwa für vier Knabengeburt in Folge ist

in der Tat sehr klein, nämlich $(1/2)^4 = 1:16$. Die *bedingte* Wahrscheinlichkeit für einen Jungen, gegeben drei sind schon da, ist aber nicht 1:16, sondern genauso groß wie die unbedingte Wahrscheinlichkeit, nämlich 1:2. Den Spermien, die um eine Eizelle kämpfen, ist der Ausgang des letzten Rennens völlig gleich. Genauso denkt der Terrorist, der eine Bombe in ein Flugzeug schmuggelt, nicht lange nach, wer sonst noch eine Bombe mit sich führt, und die Kugeln, die bei der Ziehung der Lottozahlen in die bekannten Röhrchen fallen, scheren sich keinen Deut darum, was acht Tage vorher stattgefunden hat.

Die falsche Signifikanz der Signifikanz

Falsche Wahrscheinlichkeiten führen uns auch bei der Frage in die Irre, ob bestimmte Ereignisse durch Zufall erklärt werden können oder ob sie eine systematische Ursache haben. Unter der Schlagzeile »Oft Blutkrebs neben Atommeiler vom Typ Schneller Brüter« teilt uns die *Hannoversche Allgemeine Zeitung* mit: »Bei Menschen unter 25 Jahren, die in einem Umkreis von zwölf Kilometern um das nordschottische Atomkraftwerk Dounreay leben, ist die Zahl der Leukämiefälle zehnmal so hoch wie im statistischen Durchschnitt Großbritanniens.« In dem betreffenden Gebiet sei es in vier Jahren zu fünf Fällen von Blutkrebs gekommen, aber »rein statistisch gesehen« hätte es nur 0,5 Fälle geben dürfen. Damit sei der Zufall als Erklärung ausgeschlossen und das Kraftwerk als Verursacher entlarvt.

Wer weiß - vielleicht ist das Kraftwerk wirklich schuld.

Bei einem anderen Kraftwerk (Sellafield) hat man mit statistischen Methoden die Strahlenbelastung der dort Beschäftigten als Verursacher von Leukämie bei deren Kindern entlarvt. Mein Punkt ist nur, die Statistik der *Hannoverschen Allgemeinen* sagt zu diesem Thema nichts, mit der gleichen Methode »beweisen« wir jetzt, daß Fluglärm AIDS erzeugt.

Dazu müssen wir wissen, wann ein Ereignis durch Zufall erklärbar ist und wann nicht. Wenn wir zum Beispiel 10mal würfeln und würfeln jedesmal die Sechs, so ist das ein klares Zeichen: der Würfel ist gefälscht. Wäre er korrekt, käme 10mal »Sechs« äußerst selten vor (genau: mit Wahrscheinlichkeit $(1/6)^{10} = 0,00000002$). Wir können zwar nicht ausschließen, daß wir einen unschuldigen Würfel zu Unrecht verdächtigen, aber die Wahrscheinlichkeit dafür ist sehr gering.

Nach diesem Prinzip funktionieren alle statistischen Tests, auch der, auf den die *Hannoversche Allgemeine Zeitung* ihre Meldung stützt: Wir lehnen den Zufall als Alleinursache ab (und unterstellen damit automatisch, daß es eine systematische Kausalbegründung gibt), wenn das fragliche Ereignis nur bei Zufall äußerst unwahrscheinlich wäre (wobei »äußerst unwahrscheinlich« traditionellerweise bei fünf Prozent beginnt).

Genauso entscheiden wir jetzt, ob Fluglärm AIDS erzeugt: Wir unterstellen hypothetisch, daß nur der Zufall wirkt. Treten dann in der Nähe von Flughäfen mehr AIDS-Fälle als durch Zufall erklärbar auf, lehnen wir den Zufall als Verursacher ab.

Angenommen also, jeder zehnte Bundesbürger ist vom AIDS-Virus befallen (der Einfachheit halber - der wahre Anteil ist weit kleiner). Um zu testen, ob dieser Anteil in

der Nähe von Flugplätzen systematisch höher ist als anderswo, ziehen wir eine Zufallsstichprobe von je zehn Anliegern der sieben größten deutschen Flughäfen und prüfen nach, ob unter diesen häufiger AIDS auftritt als sonst. Die folgende Tabelle zeigt das Ergebnis (nach Flughäfen getrennt):

Flughafen	AIDS-Fälle
Berlin	0
Frankfurt	2
Hannover	0
Leipzig	1
München	0
Hamburg	4
Düsseldorf	0

Wie nicht anders zu erwarten, variiert die Häufigkeit von AIDS von Ort zu Ort. An einigen Orten zählen wir weniger Kranke, an anderen mehr. Insgesamt sind 7 der 70 untersuchten Personen infiziert, genauso viele wie im Durchschnitt der Bevölkerung; die Statistik liefert keinerlei Indiz, daß Fluglärm AIDS erzeugt. Diese Hypothese ist als Hirngespinnst entlarvt (bzw. statistisch korrekt: bis zum Eintreffen weiterer Indizien auf Eis gelegt).

Trotzdem »beweisen« wir jetzt statistisch, daß Fluglärm AIDS erzeugt. Wir lassen alle Plätze außer Hamburg weg. Mit anderen Worten, wir ziehen aus der Stichprobe eine Teilmenge heraus, aber diese Teilmenge ist *nicht* zufällig.

Der Betrug beginnt, wenn wir berechnen, ob vier oder mehr AIDS-Fälle bei zehn Personen mit dem Zufall noch vereinbar sind, und dabei so tun, als wären die zehn Personen zufällig ausgewählt. Dann ist diese Wahrschein-

lichkeit nämlich sehr klein. Bei zehn Prozent Infizierten in der Gesamtbevölkerung beträgt sie, wie man durch einige simple Regeln der Wahrscheinlichkeitsrechnung schnell herausfindet, nur rund ein Prozent, und das ist klein genug. Der Zufall als einziger Verursacher dieses Ergebnisses ist damit so unwahrscheinlich, daß er nach den Regeln der Kunst als ausgeschlossen gelten darf. Mit anderen Worten, etwas anderes als Zufall muß hinter den Daten stehen - die Daten sind, wie die Statistiker sagen, »signifikant«.

In Wahrheit sind sie aber alles andere als signifikant. Die Signifikanz wurde nur künstlich hergestellt. Der Trick ist immer der gleiche, nämlich eine für die Ausgangshypothese möglichst ungünstige Teilstichprobe auszusondern, bei der Berechnung von Wahrscheinlichkeiten aber zu verfahren, als wäre die Stichprobe zufällig ausgewählt.

Bei korrekter Berechnung wären vier oder mehr Infizierte in der ungünstigsten von sieben Teilstichproben durchaus nicht außergewöhnlich. Das kommt in mehr als fünf Prozent aller Fälle durch reinen Zufall vor und ist damit nach üblichen Maßstäben nicht signifikant.

Damit ist auch die Statistik der *Hannoverschen Allgemeinen Zeitung* ohne weitere Indizien als Beweisstück unbrauchbar. Bei damals weltweit rund 300 Kernkraftwerken ist die Wahrscheinlichkeit, daß in der Nähe *irgendeines* davon mehr als fünf Fälle von Blutkrebs auftreten, auf jeden Fall größer als fünf Prozent, auch dann, wenn Atomkraft in Wahrheit nichts mit Krebs zu schaffen haben sollte.

Problematisch ist die Stichprobenauswahl auch hinsichtlich Todesursache, Raum und Zeit. Vom Extremfall absoluter Gleichverteilung abgesehen, werden zu jedem

Zeitpunkt und an jedem Ort mehr Menschen an einer bestimmten Krankheit leiden oder sterben als anderswo, so daß wir durch geeignete Wahl der Stichprobe genauso zeigen können, daß Joghurt Magenkrebs und Bergsteigen Karies erzeugt, daß Flötenspieler Hämorrhoiden und Taxifahrer Pickel haben, daß in großen Städten der Alkoholismus und auf dem Land die Malaria grassiert, daß in der Nähe von Keksfabriken Magengeschwüre, in der Nähe von Bergwerken Klaustrophobien entstehen, daß man von Schellfisch Durchfall und von Haferflocken Haarausfall bekommt - alles was zum Überbrücken eines Sommerloches in den Medien hilfreich ist: Computer schuld an grauem Star, Kabelfernsehzuschauer häufiger als andere von Herzinfarkt bedroht, Skifahrer schizophran, bei Vollmond kommt der Sensenmann, mysteriöse Häufung siamesischer Zwillinge nach Mousse au Chocolat und was der Zufall alles sonst noch so zusammenwirft. Das einzige, was wir dazu brauchen, ist ein festes Vorurteil, eine hinreichend detaillierte Morbiditätsstatistik und eine selbstgestrickte Stichprobe.

Literatur

Zitiert habe ich »Kernschmelzunfall auch bei uns wahrscheinlich«, *Frankfurter Rundschau*, 10.5.1986, und »Oft Blutkrebs nahe Atommeiler vom Typ Schneller Brüter«, *Hannoversche Allgemeine Zeitung*, 11.6.1986. Mehr zur »Illusion der Induktion« bei A. K. Dewdney: *200% of Nothing*, New York 1993 (pikanterweise wurde in der deutschen Übersetzung der Abschnitt über »signifikante« Gefährdung durch Atomkraftwerke weggelassen) und in meinem eigenen Buch: *Denkste! Trugschlüsse aus der Welt des Zufalls und der Zahlen* (Campus Verlag 1995).

16. Mythos Sozialprodukt

*»... daß die Statistik eine ganz besondere, eine ganz
kuriose Wissenschaft ist«*

Wilhelm Raabe, Horacker

Das Bruttosozialprodukt der Bundesrepublik Deutschland beträgt zur Zeit rund 4 Billionen DM jährlich (eine Vier mit zwölf Nullen). Damit ist Deutschland hinter Rußland, Japan und den USA das viert-»produktivste« Land der Welt. (Pro Kopf gerechnet sieht die Sache anders aus, aber auch hier liegt Deutschland in der Spitzengruppe.)

Solche »Hitlisten« kehren jedes Jahr so sicher wieder wie das Weihnachtsfest. Sie provozieren heiße Diskussionen in den Medien und rote Köpfe in den Parlamenten, sie haben schon manchen Minister seinen Stuhl gekostet. Sie werden zu Zeugen aufgerufen, daß ein durchschnittlicher Schweizer 270mal produktiver (reicher, ökonomisch erfolgreicher, intelligenter, klüger ...) ist als ein Bürger Mosambiks (denn das Pro-Kopf-Sozialprodukt des Spitzenreiters Schweiz ist 270 mal größer als das des Schlußlichts Mosambik), daß der Stern der USA allmählich niedergeht (weil das vormals »reichste« Land pro Kopf gerechnet nun hinter der Schweiz und Japan nur noch an dritter Stelle steht), daß Italien England ökonomisch überholt, oder daß Reichtum und Wohlstand nur per Marktwirtschaft erreichbar sind. Wegen solcher Listen beneiden uns Österreicher und Franzosen und sehen wir selbst voller Ehrfurcht zu

Schweizern und Japanern auf. Kein anderes Kriterium stellt derart deutlich eine Hackordnung der Völker her wie das Sozialprodukt.

In diesem Kapitel sehen wir uns die Eingeweide dieses goldenen Kalbes einmal näher an.

Was mißt eigentlich das Bruttosozialprodukt?

Wenn wir unser Lexikon im Schrank nach dem Sozialprodukt befragen, werden wir oft mit Erklärungen abgespeist wie »Wert der Bruttonproduktion aller Wirtschaftseinheiten eines Landes«, »gesamtwirtschaftliche Bruttowertschöpfung« oder ähnlichem. Solche Definitionen erinnern mich immer an »Links ist da wo der Daumen rechts ist«: Hinterher ist man kaum klüger als zuvor. Denn was ist die »Bruttonproduktion aller Wirtschaftseinheiten eines Landes« überhaupt?

Ein Schlaumeier würde jetzt sagen: das Sozialprodukt, und damit hat sich die Katze exakt wieder in den Schwanz gebissen. Darüber hinaus ist diese Erklärung auch noch sachlich falsch, denn das Sozialprodukt mißt eindeutig *nicht* den Wert aller in einem Jahr produzierten Güter und Dienstleistungen einer Volkswirtschaft. Es mißt nur eine Teilmenge davon, und dazu noch eine Teilmenge, die von Jahr zu Jahr und von Volkswirtschaft zu Volkswirtschaft erheblich schwankt.

Nach aktueller Praxis erfaßt das Bruttosozialprodukt vor allem Güter und Dienstleistungen, die gegen Geld gehandelt werden: Autos, Häuser, Waschmaschinen, Lebensmittel, Heizöl, Taxifahrten, Dienstleistungen von Ärzten,

Krankenhäusern, Banken, Post. Das hat den großen Vorteil, daß man so im wahrsten Sinn des Wortes Äpfel und Birnen aufaddieren kann: 1.000 kg Äpfel à 4 Mark und 2.000 kg Birnen à 3 Mark ergibt ein Sozialprodukt von $1000 * 4 + 2000 * 3 = 100.00$ Mark.

Dieses System hat aber auch große Nachteile. So steigt etwa das Sozialprodukt auch dann, wenn sich real überhaupt nichts ändert und nur die Preise steigen. Das westdeutsche Bruttosozialprodukt ist etwa von 1960 bis 1990 um mehr als 1.000 Prozent gestiegen (von 300 Milliarden Mark auf rund 4.000 Milliarden Mark), aber von diesen über tausend Prozent Wachstum gehen mehrere hundert Punkte auf die Inflation zurück: Zu konstanten Preisen ist das Sozialprodukt nur um rund 200 Prozent gewachsen.

Das Problem der Schattenwirtschaft

Noch schwerer als dieses problematische Umrechnen von realen Gütern und Dienstleistungen in Geldeinheiten wiegt der Ausschluß von Produkten, die durchaus zu Wohlergehen und Zufriedenheit der Menschen beitragen, aber nicht gegen Geld gehandelt und daher auch nicht im »offiziellen« Sozialprodukt gemeldet werden: Die selbst statt in der Werkstatt aufgezogenen Winterreifen, das selbst tapezierte Wohnzimmer, das selbst getippte Manuskript, die selbst reparierte Armbanduhr: Was man selber macht statt von anderen gegen Entgelt machen läßt, fällt systematisch durch den Rost. Wenn wir vom Sohn des Nachbarn für 20 Mark den Rasen mähen lassen, steigt das Sozialprodukt um 20 Mark. Mähen wir den Rasen selbst,

bleibt das Sozialprodukt konstant. Das produzierte Gut, nämlich ein frisch gemähter Rasen, ist in beiden Fällen gleich, aber einmal zählt es zum Sozialprodukt und einmal nicht.

Der mit Abstand größte Brocken der so unterschlagenen Güter und Dienstleistungen ist die in der Regel unbezahlte Arbeit unserer Hausfrauen und Hausmänner: spülen, waschen, kochen, Treppen putzen, Kinder erziehen, Kranke pflegen, alles trägt zu unserem Wohlergehen bei, wird aber im Sozialprodukt nicht mitgezählt. Wenn also ein reicher Junggeselle seine Haushälterin heiratet, die vorher einen Monatslohn von 4.000 Mark bekam und nun, nach der Hochzeit, nichts anderes tut als vorher auch, nur ohne Honorar, so hat er mit einem Schlag das deutsche Sozialprodukt um $12 \times 4.000 = 48.000$ Mark pro Jahr reduziert. Inklusive aller dieser in privaten Haushalten »gratis« erbrachten Dienstleistungen (etwa durch Bewertung nach dem Bundesangestellten-Tarifvertrag) wäre unser Sozialprodukt im Handumdrehen um fast die Hälfte größer.

Ein weiterer großer Brocken, der im Gegensatz zur Haushaltsproduktion ganz offiziell zum Sozialprodukt gehört, der aber gleichfalls auf dem Weg in das Statistische Jahrbuch spurlos verschwindet, ist die Schwarzarbeit beziehungsweise »Schattenwirtschaft«, wie die Ökonomen sagen. Wenn in Land A ein Maurer für tausend Mark eine Grube gräbt und diesen Lohn versteuert, steigt das Sozialprodukt um tausend Mark. Wenn in Land B ein anderer Maurer die gleiche Arbeit schwarz erledigt, so steigt das »eigentliche« Sozialprodukt dort ebenfalls um tausend Mark. Jedoch bleibt dieser Beitrag unberücksichtigt, und das amtlich ausgewiesene Sozialprodukt fällt um tausend Mark zu niedrig aus.

Im Gegensatz zur Haushaltsproduktion wird in der Schattenwirtschaft also durchaus Ware gegen Geld gehandelt. Nur erfährt die amtliche Statistik nichts davon.

Das Ausmaß dieser Schattenwirtschaft wird heute in westlichen Industrienationen auf rund 10 Prozent des amtlichen Sozialprodukts geschätzt. Die höchste Quote (13 Prozent) vermutet man für Schweden und Italien, die kleinsten Quoten von 4,1 und 4,3 Prozent für Japan und die Schweiz. Die Bundesrepublik Deutschland (West) belegt mit geschätzten 8,6 Prozent einen Mittelplatz - bei einem Sozialprodukt von 4 Billionen Mark immerhin mehr als 300 Milliarden Mark. Um diese Summe ist also die offizielle deutsche Zahl zu klein.

Das Problem des Staatssektors

Der weitaus größte Teil des Sozialprodukts wird in Unternehmen produziert. Aber auch der Staat trägt dazu bei, und zwar absolut wie relativ gesehen immer mehr: vor allem die Dienstleistungen von Polizei, Justiz und Feuerwehr, von Schulen, Hochschulen und öffentlicher Verwaltung, ohne die ein Sozialsystem nicht funktioniert.

Das Problem ist nur: Wie wollen wir diese Güter bewerten? Die Beiträge der Unternehmen werden zu den Preisen bewertet, die sie am Markt erzielen, d. h. heiß begehrte und daher teure Güter tragen mehr zum Sozialprodukt bei als Ladenhüter, die niemand haben will.

Bei den Wohltaten des Staates funktioniert dieses System leider schlecht - sie werden uns aufgedrückt, ob wir wollen oder nicht. Zwar müssen wir durchaus dafür

zahlen, aber nur selten direkt, meistens indirekt über Steuern und Sozialabgaben. Wenn wir die Dienste eines Standesbeamten oder Polizisten beanspruchen, kostet uns das zunächst nichts oder nur minimale Gebühren, die aber die Kosten keineswegs decken. Erst in den jährlichen Etat-Debatten unserer Parlamente wird uns dafür eine Rechnung präsentiert.

Aushilfsweise gehen daher die vom Staat produzierten Güter und Dienstleistungen mit ihren Gestehungskosten in die Sozialprodukt-Berechnung ein, und das hat einen perversen Effekt: Je unwirtschaftlicher, schlampiger und ineffizienter unsere Staatsbürokratie, desto höher das Sozialprodukt! Wenn etwa das Katasteramt XY völlig unnötig 20 Mitarbeiter beschäftigt, obwohl die Arbeit auch von 10 getan werden könnte, hat es sich um das Sozialprodukt verdient gemacht. Würde es, ohne seine Dienste einzuschränken, die 10 überflüssigen Mitarbeiter in die freie Wirtschaft entlassen, wäre sein Produkt auf einmal nur die Hälfte wert.

Der Staatsbeitrag bläht das Sozialprodukt also künstlich auf und macht so die Untererfassung der Schattenwirtschaft teilweise wieder wett. Zum Glück, wenn man so will, laufen beide Fehler zudem oftmals parallel. Länder wie Italien, mit großem und notorisch ineffizientem Staatssektor, machen so zwar einerseits viel Lärm um nichts, leisten sich aber zum Ausgleich eine blühende Schattenwirtschaft wie sonst nirgends in Europa, in der ein Beamter vielleicht das Doppelte seines offiziellen, im Sozialprodukt vermerkten Gehaltes, für das er aber nicht viel leistet, mit Nebengeschäften verdient, für die er wirklich arbeitet, die aber im Sozialprodukt nicht erscheinen. Dann wieder gibt es Länder mit kleinem und effizientem Staatssektor,

aber auch marginaler Schattenwirtschaft wie die Schweiz, in denen *beide* Effekte, das statistische Aufblähen durch den Staatssektor und das statistische Komprimieren durch die Schattenwirtschaft, kleiner sind als anderswo. Ob diese Fehler sich aber wirklich international im gleichen Umfang neutralisieren, weiß niemand so genau.

Das Problem der Vorleistungen

Hinter der Erfassung und Bewertung der binnen eines Jahres produzierten Güter und Dienstleistungen lauern als nächster Stolperstein die sogenannten Vorleistungen. Angenommen, Robinson Crusoe erntet auf seiner Insel 3 Zentner Getreide, für das er aber einen Zentner Saatgut braucht. Offenbar beträgt sein Sozialprodukt damit nicht drei Zentner, sondern zwei: Von der Bruttoproduktion sind die Vorleistungen abzuziehen.

Die Frage ist nur: Was ist eine Vorleistung? Dieses Problem ist das kniffligste bei der Interpretation und Berechnung des Sozialproduktes überhaupt. Es wirft sogar einen philosophischen Schatten auf diese Materie, denn nach der reinen Lehre von Karl Marx sind etwa Kleidung und Ernährung der Lohnabhängigen nichts als Input im gesamtwirtschaftlichen Produktionsprozeß und damit von der Wertschöpfung (dem Mehrwert) abzuziehen.

Dieser Ansicht von Karl Marx konnten sich unsere Amtsstatistiker aber nicht anschließen. Sie ziehen vom Produktionswert eines Unternehmens nur die selbsterstellten bzw. von anderen *Unternehmen* gekauften, im Endprodukt verschwundenen Güter und Dienstleistungen ab.

Die Dienste der Produktionsfaktoren Arbeit und Kapital zählen nach gängiger Praxis nicht dazu.

Vor allem bei der Staatsproduktion führt das wieder zu perversen Konsequenzen. So schätzen wir etwa die Dienste von Polizei, Justiz und Feuerwehr wohl kaum um ihrer selbst. Vielmehr wären die meisten von uns sicher mehr als froh, wenn alle Menschen Engel wären, wenn wir Polizei und Feuerwehr überhaupt nicht brauchten. Mit anderen Worten, dieser Teil der Staatsproduktion ist eher als Vorleistung für das Funktionieren der Sozialgemeinschaft denn als eigenständiger Beitrag zu unserem Wohlstand anzusehen. Die Leistungen von Polizei, Justiz, Feuer- und Bundeswehr an sich will keiner haben - sie sind nur Inputs für das eigentlich gewünschte, Friede, Ordnung, Sicherheit, und daher streng genommen als Vorleistung vom Produktionswert abzuziehen.

Zur Zeit geschieht das jedoch nirgends auf der Welt - mit der Folge, daß zum Beispiel eine Tankerkatastrophe mit Riesenumweltschäden oder ein Erdbeben, ein Tornado, ein kleiner Krieg oder eine Flutkatastrophe das Sozialprodukt erhöhen: Die Schäden werden, falls überhaupt, nur unzureichend subtrahiert, die Arbeit der Retter und Helfer dagegen voll dem Produktionswert zugeschlagen.

Literatur

Dieses Kapitel ist meinem Buch *Statistik verstehen* entlehnt (Campus 1992). Die Technik der Sozialproduktberechnung wird in fast allen Lehrbüchern zur Wirtschaftsstatistik und der sogenannten »Volkswirtschaftlichen Gesamtrechnung« im Detail er-

klärt. Meine Favoriten sind hier der inzwischen mehrfach neu aufgelegte Klassiker von Alfred Stobbe: *Volkswirtschaftslehre 1: Volkswirtschaftliches Rechnungswesen* (Berlin 1966), und Peter von der Lippe: *Wirtschaftsstatistik* (4. Aufl. Stuttgart 1990). Im dritten Kapitel dieses letzten Buches sind sehr schön all die Tricks und Kniffe aufgeführt, die unsere Amtsstatistiker in Wiesbaden bei der konkreten Berechnung des Sozialprodukts verwenden.

Mit dem im amtlichen Sozialprodukt so wenig gewürdigten Beitrag der privaten Haushalte zur Wirtschaftsproduktion setzt sich der Aufsatz von M. Hilzenbecher »Die schattenwirtschaftliche Wertschöpfung der Hausarbeit« in den *Jahrbüchern für Nationalökonomie und Statistik* 1986 auseinander. Dort findet man auch Hinweise auf weitere Arbeiten zu diesem Problem. Der sonstigen Schattenwirtschaft nehmen sich ebenfalls eine ganze Reihe von Autoren an. Als erster Einstieg kann hier das Buch von Hannelore Weck und anderen: *Schattenwirtschaft* dienen (München 1984). Auch die sonstigen Mängel des Sozialprodukts als Wohlstandsmaß werden seit langem heiß diskutiert und etwa in dem Buch von Christian Leipert: *Die heimlichen Kosten des Fortschritts* (Frankfurt 1989) ins Rampenlicht gerückt.

17. Welchen Zahlen kann man noch vertrauen?

*»Ein falscher Zeuge bleibt nicht ungestraft;
und wer frech Lügen redet, wird umkommen.«*

Die Sprüche Salomos, 19,9

Zum Glück haben auch statistische Lügen kurze Beine (oder wie ein altes Sprichwort aus Angola sagt: Auch ein Furz, den du unter Wasser losläßt, kommt an die Oberfläche). »Und dann mußten für alle Unterteilungen die Ergebnisse nach den fünf möglichen Kriterien aufgeschlüsselt werden: Heilung, Besserung, keine Veränderung, Verschlechterung und Tod«, schreibt Alexander Solschenizyn in *Krebsstation*. »Die Ergebnisse notierte Sojas Helfer besonders aufmerksam. Ihm war sofort aufgefallen, daß fast keine endgültigen Heilungen, aber auch wenig Todesfälle darunter waren.«

»Aha, hier darf also niemand sterben, sie entlassen die Kranken rechtzeitig«, sagte Kostoglotkow.

Und genauso kommen wir auch selber den meisten statistischen Falschmünzern schnell auf die Schliche. Wenn etwa der griechische Geschichtschreiber Herodot uns wissen läßt, das Perserheer des Xerxes sei mit einer Übermacht von 5.283.220 Mann über den Bosphorus gezogen, so ist das leicht als Unfug zu entlarven. Erstens ist die Zahl viel zu genau; sie soll vermutlich nur das Unwissen des Autors übertünchen (siehe Kapitel 1), zweitens hätte ein solches Heer gar nicht auf das Schlachtfeld gepaßt, auf



Es nützt nichts, irgendwann kommt alles an den Tag

dem die Perser dann den Griechen gegenüberstanden. »Ich habe berechnet«, schreibt der Historiker Hans Delbrück, »daß, wenn wir uns dieses Landheer hintereinanderweg auf einer Straße marschierend denken, der Zug 420 Meilen lang gewesen wäre, d. h. daß, als die ersten vor Thermopylä ankamen, die letzten gerade aus Susa ausmarschieren konnten...«

In Wahrheit zählten die Perser rund 15.000 Krieger, wie Delbrück durch einfache Überlegungen zu Nachschub und Verpflegung ausrechnete; sie waren in beiden Perserkriegen den Griechen zahlenmäßig unterlegen.

Genauso lassen sich auch viele andere falsche Zahlen durch einfache Kontrollen leicht enttarnen. Von dem Normannenheer in der Schlacht bei Hastings 1066 wird gesagt, es sei so groß gewesen, daß es die Flüsse austrank, zwölf mal hunderttausend Mann; überlegt man aber, wie viele Schiffe die Normannen hatten und wie viele Krieger ein Schiff transportieren konnte, kommt man auf maximal viertausend. Die Stärke der Polen in der

Schlacht zu Tannenberg wird in einer Lübecker Chronik mit 5.100.000 Mann, die Zahl der Toten mit 630.000 angegeben. In Wahrheit zählten die Polen mitsamt Verbündeten rund 17.000 Mann, wie man wieder anhand der Bedürfnisse zu Essen, Trinken, Fortbewegen, Schlafen leicht berechnen kann - solche offensichtlichen Übertreibungen in historischen Berichten können niemanden lange täuschen. Wenn ein feindliches Heer mehr Krieger zählt als eine ganze Volkswirtschaft ernähren könnte, wenn Saddam Husseins Soldaten im Golfkrieg mehr amerikanische Flugzeuge abgeschossen haben wollen als die Amerikaner überhaupt besaßen, wenn Seuchen oder Erdbeben mehr Menschen weggraffen als in einer Gegend vorher lebten, dann haben statistische Falschmünzer schnell ausgespielt.

Das Wort hat der Genosse Mittag

Aber auch von dieser ewigen Bedrohung durch die Fakten abgesehen haben Datenfälscher kein leichtes Leben. Sie müssen nämlich nicht nur unsere Skepsis, sondern oft auch noch die Logik überlisten, und das wird mit jeder Lüge immer schwerer. Die letzte Abbildung dieses Buches zeigt eine Anweisung von Günter Mittag, Chef der Wirtschaft des Zentralkomitees der SED, an die Staatliche Zentralverwaltung für Statistik in der alten DDR; Mittag hätte gerne die Außenhandelszahlen etwas »aufgebessert«: statt eines sowieso schon falschen Überschusses von 521 Millionen Valutamark einen Überschuß von 910 Millionen. (In Wahrheit gab es in der fraglichen Periode, den

ersten drei Quartalen 1987, ein Defizit von 579 Millionen Mark; diese und andere unmittelbare Eingriffe der Partei in das statistische Außenbild der DDR sind in dem Bericht der Enquete-Kommission »Aufarbeitung von Geschichte und Folgen der SED-Diktatur in Deutschland« des Deutschen Bundestages nachzulesen, speziell in dem Gutachten meines Statistiker-Kollegen Peter von der Lippe.)

Aber diese Phantomzahlen brachten die armen DDR-Statistiker in des Teufels Küche. Denn wenn man in einem Jahr den Außenhandelsüberschuß vergrößert, wird dessen *Wachstum* ein Jahr später kleiner. Es sei denn, man fälscht nächstes Jahr von neuem ...

Und außerdem: wenn man *eine* Zahl fälscht, muß man auch andere Zahlen fälschen. Es reicht nicht, den Exportüberschuß um einige Millionen Phantasiemark zu vergrößern - mindestens eine der Zahlen zu Im- und Export ist dann ebenfalls zu fälschen (und wie wir in der Abbildung sehen, wurden die Importe in der Tat um die benötigten Millionen abgesenkt). Und wenn ich die Importe fälsche, muß ich auch die Zahlungsbilanz fälschen, und wenn ich die Zahlungsbilanz fälsche, muß ich die Statistik der Zentralbank fälschen, und wenn ich die Statistik der Zentralbank fälsche ... Kurz: es ist weit schwerer, eine Phantomwirtschaft statistisch konsistent zu simulieren, als konsequent die Wahrheit abzubilden; wäre die DDR politisch nicht ohnehin verschwunden - das ganze Kartenhaus der ausgedachten Zahlen wäre auch von selber unter dieser Last der Widersprüche bald zerfallen.

Solche groß angelegten amtlichen Lügen sind wie ein Schneeball: Je länger man sie wälzt, desto schwerer werden sie - auf Dauer kann man seine Umwelt mit Statistik nicht betrogen.

3. Damit auch per 30.9. ein Exportüberschuß gemeldet werden kann wird vorgeschlagen, Veränderungen in gleicher Höhe wie für das 1. Halbjahr vorzunehmen..

~~Daraus ergibt sich ein Exportüberschuß von 527 Mio VM.~~
910

Gleichzeitig wird damit erreicht, daß kein so starker Rückgang beim Export und Import im Vergleich zu den im Vorjahr gemeldeten Angaben eintritt.

Daraus ergibt sich ein zu meldender

zum gemeldeten Ist 30.9.1986

1817883
36696

Export	von 18 803 Mio VM	=	93,6 %
Import	von 148 282 Mio VM	=	107,9 %
Umsatz	von 137 085 Mio VM	=	105,7 %
Export- überschuß	von 527 Mio VM		

910

4. Der Exportüberschuß im Handel mit dem sozialistischen Wirtschaftsgebiet beträgt ~~2 044~~ Mio VM, so daß für den Außenhandel insgesamt ein Saldo von ~~142 309~~ Mio VM per 30.9.1987 an den RGW und UNO-Organen gemeldet ~~wird.~~

3054

Mit dem Abzug von 30,5
1150 helfen abgedeckt

Mit dem für Feind von 30,5
1315 helfen abgedeckt

Ein Originaldokument aus den Archiven der amtlichen DDR-Statistik die Zahl 910 in der Mitte ist von Günter Mittag höchstpersönlich eingetragen, die handschriftlichen Anmerkungen unten dokumentieren die Reaktion der DDR-Statistiker

Oder um mit Abraham Lincoln zu sprechen: »You can fool some of the people some of the time, but you cannot fool all of the people all of the time.«

Ein Trost daher zum Schluß: Falsche Zahlen sind meist leichter zu erkennen als falsche Worte oder falsche Bilder, und ohne unser Zutun kann uns niemand auf lange Sicht mit solchen Blendwerken für dumm verkaufen.

Amen.