

Katharina Schüller



Statistik und Intuition

Alltagsbeispiele kritisch hinterfragt

SACHBUCH



Springer Spektrum

Statistik und Intuition

Über den Autor

Katharina Schüller ist Diplom-Statistikerin, Statistik-Expertin bei DRadio Wissen, Lehrbeauftragte an verschiedenen Hochschulen und ausgezeichnet als „Statistikerin der Woche“ durch die American Statistical Association. In zahlreichen Vorträgen und Publikationen klärt sie auf über den richtigen (und falschen) Gebrauch von Statistik.

Katharina Schüller

Statistik und Intuition

Alltagsbeispiele kritisch hinterfragt



Springer Spektrum

Katharina Schüller
STAT-UP
München
Deutschland

ISBN 978-3-662-47847-9 ISBN 978-3-662-47848-6 (eBook)
DOI 10.1007/978-3-662-47848-6

Die Deutsche Nationalbibliothek verzeichnetet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Spektrum
© Springer-Verlag Berlin Heidelberg 2015

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen.

Planung: Iris Ruhmann

Gedruckt auf säurefreiem und chlorfrei gebleichtem Papier

Springer Berlin Heidelberg ist Teil der Fachverlagsgruppe Springer Science+Business Media (www.springer.com)

Geleitwort

Dieses Buch sollte man allen Nutzern und Nutzerinnen sowie Studierenden der Statistik, nicht zuletzt Journalisten, als Pflichtlektüre auf den Nachttisch legen. Man kann es nämlich sehr gut auch häppchenweise lesen, außerdem ist es unterhaltsam und an zahlreichen flott geschriebenen Fallbeispielen aufgehängt, sodass die Einführung in die Denk- und Vorgehensweise der Statistik quasi schmerzlos nebenbei geschieht.

Diese Denk- und Vorgehensweise der Statistik ist den meisten Menschen heutzutage immer noch genauso fremd wie unseren Vorfahren vor einer Million Jahren in den Savannen Afrikas. Wenn wir dem großen Psychologen, Statistiker und Wirtschafts-Nobelpreisträger Daniel Kahneman glauben dürfen, ist ein großer Teil unseres Denkapparates, der von den Affen ererbte Teil, ein miserabler Statistiker. Wir schließen hemmungslos von Einzelfällen auf das Kollektiv, vernachlässigen dabei Variabilität und Stichprobenfehler, wir verwechseln systematisch Korrelation und Kausalität, konstruieren gerne künstliche Muster in zufälliges Chaos hinein, und wir addieren, wo wir eigentlich multiplizieren müssten. Gegen diese Gendefekte anzuschreiben, ist nicht leicht und wurde bereits oft versucht, unter anderem auch

von mir selbst. Umso größer ist meine Freude, mit Katharina Schüller eine derart kompetente Mitstreiterin gefunden zu haben. Man kann ihrem Buch nur eine maximale Verbreitung wünschen.

Walter Krämer

Dank

Mein ganz besonderer Dank geht an meine Kollegen Stefan Fritsch, Adrian Hambeck, David Hillmann, Carina Hindinger, Daniel Schalk und Claudia Stuckart, die große Teile der Recherchen und insbesondere der grafischen Darstellungen übernommen haben.

Ausdrücklich gilt mein Dank darüber hinaus Walter Krämer für unzählige beflügelnde Anregungen und stets bedingungslose Unterstützung, Alain Thierstein für die Idee der Gliederung dieses Buches und für wertvolle Hinweise zur Visualisierung, außerdem Frank Schneider sowie Bernhard, Valentina, Viviana und Maximilian Schüller für grenzenlose Geduld und ebenso grenzenlos kritische Rückmeldung, wie sie nur liebende Familienmitglieder geben können.

Inhalt

1	Statistik: selbstverständlich?	1
1.1	Einleitung.....	1
1.2	Über Denkmuster.....	7
1.3	Heuristik und Statistik.....	13
1.4	Wie Statistiker arbeiten.....	25
2	Politik und Weltgeschehen	31
2.1	Warum Steuerschätzungen danebenliegen.....	31
2.2	Wann man arbeitslos ist	36
2.3	Wie 17-Jährige wählen	42
2.4	Wie man Wahlen fälscht	47
2.5	Wie man Korruption misst	53
2.6	Wie man ein gutes Land wird.....	59
2.7	Warum Sparsamkeit nicht tötet	64
3	Wirtschaft und Unternehmen	71
3.1	Wer wirklich arm ist.....	71
3.2	Wie sich Ungleichheit entwickelt	76
3.3	Wie die Mittelschicht schrumpft.....	82
3.4	Was ein Mensch zum Leben braucht	88
3.5	Warum Schwermetall nicht reich macht	94
3.6	Was Absolventen wollen.....	98
3.7	Warum Frauen Vorstände verlassen	103

4	Wissen und Technik	109
4.1	Wie man Forschung fälscht	109
4.2	Wann Waschmaschinen kaputt gehen	115
4.3	Was Pünktlichkeit bedeutet	122
4.4	Wo Kinder verunglücken	128
4.5	Wie gefährlich Windräder sind	134
4.6	Was weibliche Wirbelstürme anrichten	139
4.7	Was der Klimaschutz kostet	144
5	Gesundheit und Ernährung	149
5.1	Was uns umbringt	149
5.2	Wie Essen unsere Gesundheit bedroht	155
5.3	Wer etwas zu lachen hat	161
5.4	Wer welches Fleisch isst	166
5.5	Warum Alkohol nicht gesund ist	172
5.6	Wie häufig psychische Erkrankungen sind	176
5.7	Warum Dummheit nicht krank macht	180
6	Gesellschaft und Leben	189
6.1	Wie der Fußball statistisch wurde	189
6.2	Was Namen und Daten aussagen	194
6.3	Wer attraktiv und glücklich ist	202
6.4	Warum Monogamie im Trend liegt	205
6.5	Wie oft Partner einander schlagen	211
6.6	Wie man Schuld nicht herleitet	217
6.7	Wer liest, schreibt und kriminell wird	227
7	Das Handwerkszeug	235
7.1	Daten und Skalen	235
7.2	Mittelwerte und Streuung	236
7.3	Prozente und Risiken	242
7.4	Der Zentrale Grenzwertsatz	247
7.5	Testen und Schätzen	251
7.6	Konfidenzintervalle	262
7.7	Zusammenhangsanalysen	264
7.8	Regressions- und komplexere Modelle	268
7.9	Zeitreihen und Prognosen	272

Statistisch denken: Elf Regeln für den Alltag	281
Sachverzeichnis	285

Abbildungsnachweis

Bei sämtlichen Abbildungen mit Ausnahme von Abbildung 6 handelt es sich um eigene Darstellungen, mit Ausnahme von Abbildung 7 auch auf Basis eigener Berechnungen. Im Folgenden sind die Quellen für das verwendete Zahlenmaterial genannt. Nicht verzeichnete Abbildungen beruhen auf simulierten Daten bzw. eigenen Erhebungen.

- Abb. 1 Transparency International: Corruption Perceptions Index 2014. www.transparency.org, 2015.
- Abb. 2 UNAIDS: AIDSinfo Online Database. www.unaids.org, o.J.
- Abb. 3 Oxfam International: Oxfam Issue Briefing: Wealth: Having It All And Wanting More. London, Oxfam House, 2015.
- Abb. 5 Böcking, D.: Soziale Ungleichheit: Deutschland wird amerikanischer. Spiegel Online, 05.12.2011.
- Abb. 6 Wikipedia: https://commons.wikimedia.org/wiki/File:Metal_bands_per_country.png (eigene Einfärbung und Legende), lizenziert unter der CC BY-SA 3.0 gemäß <https://creativecommons.org/licenses/by-sa/3.0/deed.en>

- Abb. 7 Kienbaum: Absolventenstudie 2007/08. Gummersbach, 2008 und Kienbaum: Absolventenstudie 2009/10. Gummersbach, 2010.
- Abb. 10 Opitz-Neumann, N. et al.: Kinderunfallatlas. Bundesanstalt für Straßenwesen, Bergisch Gladbach, Mensch und Sicherheit Heft M 232, 2012 und Stadt Rosenheim: Amt für Verkehrswesen, persönliche Mitteilung vom 16.01.2013.
- Abb. 11 Deutscher Fleischer-Verband: Geschäftsbericht 2014. S. 36–39, Frankfurt a.M., 2015 und Deutscher Fleischer-Verband: Geschäftsbericht 2012. S. 38–41, Frankfurt a.M., 2013.
- Abb. 12 Europäischer Rat: Soziale Indikatoren: Zusammengefasste Scheidungsziffer. www.gesis.org, o.J. und Statistisches Bundesamt: Eheschließungen, Ehescheidungen. www.destatis.de, o.J.
- Abb. 13 Schlack, R. et al.: Körperliche und psychische Gewalterfahrungen in der deutschen Erwachsenenbevölkerung. Ergebnisse zur Studie zur Gesundheit Erwachsener in Deutschland (DEGS1). Bundesgesundheitsblatt 56(5/6), S. 755–764, 2013.
- Abb. 17 Grotlüschen, A. und Riekmann, W.: Funktionaler Analphabetismus in Deutschland. Münster, Waxmann, 2012 und Statistisches Bundesamt: Statistik der Allgemeinbildenden Schulen, Absolventen/ Abgänger: Bundesländer, Schuljahr, Geschlecht, Schulabschlüsse, Tabelle 21111-0014. www.gesis.destatis.de, o.J.
- Abb. 30 Vigen, T.: Spurious Correlations. www.tylervigen.com, 2015.

Tabellennachweis

Im Folgenden sind die Quellen für das in den Tabellen verwendete (Zahlen-)Material genannt, soweit externe Quellen verwendet wurden. Tabelle 6 beruht auf eigenen Erhebungen.

- Tab. 1 Nach Haller, M.: Recherchieren. Ein Handbuch für Journalisten. Ölschläger, München, 3. Auflage, 1989, S. 36, eigene Kürzungen und Ergänzungen.
- Tab. 2 Statistisches Landesamt Bremen: Statistisches Jahrbuch 2011. Bremen, 2012, eigene Berechnungen.
- Tab. 3 Enste, D. H. et al.: Mythen über die Mittelschicht. Wie schlecht steht es wirklich um die gesellschaftliche Mitte? Roman-Herzog-Institut, München, 2011, eigene Berechnungen.
- Tab. 4 IPCC: Climate Change: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 2014 und World Bank: Gross domestic product 1960–2012. World Development Indicators Database, 17.04.2015, eigene Berechnungen.

- Tab. 5 Appel, B. et al.: EHEC Outbreak 2011. Investigation of the Outbreak Along the Food Chain. Bundesinstitut für Risikobewertung, BfR-Wissenschaft 03, Berlin, 2012, Deutsche Sepsis-Hilfe e.V.: Informationen zur Sepsis. www.sepsis-hilfe.org, o.J. (Stand: 10.05.2015), Helmholtz Zentrum München: Lungenentzündung – Verbreitung. www.lungeninformationsdienst.de (Stand: 10.05.2015), o.J., Robert-Koch-Institut: Infektionsgeschehen von besonderer Bedeutung. Supplement zum Epidemiologischen Bulletin Nr. 21, 30.05.2011, Robert-Koch-Institut: Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2011. Berlin, 2012; Robert-Koch-Institut: Infektionsepidemiologisches Jahrbuch meldepflichtiger Krankheiten für 2013. Berlin, 2014, Statistisches Bundesamt: Die 10 häufigsten Todesursachen. www.destatis.de, o.J., Sülberg, D.: Die Altersappendizitis – Der CRP-Wert als Entscheidungshilfe. Dissertation an der Ruhr-Universität Bochum, 2008, eigene Berechnungen.
- Tab. 7 Nach Seyler, H.: Mammographie-Screening. Kontroverse über Nutzen und Risiken. Pro Familia Medizin Nr. 1, Februar 2014, S. 7, eigene Kürzungen und Berechnungen.

1

Statistik: selbstverständlich?

1.1 Einleitung

Statistik spekuliert nicht. Statistik beruht auf harten Fakten. „Es ist ein Buch über Demografie, das ist in Ordnung, das sind einfach Zahlen“, sagte Frank Schirrmacher über Thilo Sarrazins Buch „Deutschland schafft sich ab“.

Ich bin Statistikerin. Mein Arbeitsalltag besteht darin, Zahlen und Fakten zusammenzutragen und zu analysieren, um daraus Informationen zu gewinnen. Informationen, die mir und anderen helfen zu begreifen, wie unsere Welt funktioniert – unsere Gesellschaft, unsere Umwelt, unsere Wirtschaft und natürlich wir Menschen selbst. Statistiker betreiben, vereinfacht gesagt, Recherche – nur eben mit ganz bestimmten, formalisierten Mitteln. Recherche ist, so definiert es der Journalist Hans Leyendecker, „ein professionelles Verfahren, mit dem Aussagen über Vorgänge beschafft, geprüft und beurteilt werden“. Recherche hat zunächst einmal etwas mit Handwerk zu tun. Statistisches Handwerkszeug wird angesichts der Verfügbarkeit großer Datenmengen auch für Journalisten immer wichtiger. Denn Datenjournalismus heißt, Daten ihre Geschichte erzählen zu lassen, Erklärungen zu finden und Entwicklungen vorherzusagen.

Daten, das heißt heute oft: Big Data. Gemeint sind Datenmengen, die zu groß, zu komplex und oft zu heterogen sind, um sie noch von Hand zu verarbeiten. Das bedeutet zugleich, dass die klassischen Analysewerkzeuge, die viele von uns in der Ausbildung oder im Studium erlernt haben, für Big Data nicht mehr funktionieren. Mit dem Taschenrechner und Excel-Tabellenblättern können wir die Informationen nicht mehr analysieren, die in Tweets stecken oder in den Datenströmen, die Smartphones ununterbrochen in die Cloud schicken. Dort, in den hoch frequenten, feinräumigen, vielfach verknüpften Datenströmen liegt aber das Wissen verborgen, das Politik und Wirtschaft heute benötigen: Wo genau bei der Hochwasserkatastrophe 2013 in Dresden der nächste Damm zu brechen drohte, posteten Tausende freiwilliger Helfer auf Twitter und Facebook. Wie sich Menschen in einer Einkaufspassage bewegen und welche Werbung sie dabei womöglich zu Gesicht bekommen, verraten ihre Smartphones.

Big Data betrifft aber nicht nur Unternehmen und Politiker, sondern auch Journalisten. Weil Big Data neue Erwartungen an die Berichterstattung schafft, können Journalisten etwas von Statistikern lernen. Die Medien leisten einen wesentlichen Beitrag dazu, welches Bild wir alle uns von der Welt und den alltäglichen Ereignissen machen, wie wir diese Ereignisse in einen Kontext bringen und welche Muster wir zu erkennen glauben. Und weil dieses Bild immer stärker durch Zahlen und Statistiken geformt und untermauert wird, braucht jeder, der Zeitungen liest, Nachrichten im Fernsehen ansieht oder im Radio hört und auf der Suche nach Neuigkeiten im Internet surft, ein grundlegendes Verständnis von Statistik.

Der Wunsch, die Welt begreifbar zu machen, ist offenbar so stark, dass selbst völlig sinnlose Zahlen in seriösen Medien präsentiert werden, und das sogar, wenn im selben Artikel auf den statistischen Unfug hingewiesen wird. So schrieb die „Frankfurter Allgemeine Zeitung“ über eine Befragung von Pegida-Demonstranten, die Befunde seien keinesfalls repräsentativ und man könne nichts über den typischen Pegida-Demonstranten sagen. Im übernächsten Abschnitt wurden dann aber exakte Prozentsätze zur politischen Orientierung, zum Wahlverhalten, zum Vertrauen in politische und öffentliche Institutionen und zur Haltung von Pegida-Demonstranten gegenüber der Demokratie genannt, als ob es sich um Fakten handeln würde.

Frank Schirrmachers Kommentar ist aus statistischer Sicht, vorsichtig gesagt, etwas unglücklich formuliert. Vielleicht wollte er ausdrücken, dass man Zahlen aus der Demografie nennen darf, dass das in Ordnung ist und kein Tabu sein sollte. Wir können ihn selbst nicht mehr fragen. Beim Wort genommen lässt sein Satz jedoch gleich zwei problematische Interpretationen zu.

Demografie, oder allgemeiner Statistik, besteht einfach nur aus Zahlen – das ist die erste mögliche Lesart –, und zwar aus Zahlen, die oft wenig bis gar nichts darüber aussagen, wie unsere Welt wirklich ist. Die Welt ist eben komplex, weil sie nur durch Beziehungen und Kontexte zu verstehen ist. Selbst der Städtestatistiker Ansgar Schmitz-Veltin schreibt, der Wunsch der Städtestatistik, die gesellschaftliche Realität mit „objektiven“ Zahlen zu beschreiben, habe vermutlich noch nie ernsthaft funktioniert. Dem will ich widersprechen. Statistik ist ein Mittel, Informationen zusammenzufassen und zu komprimieren.

Dabei geht naturgemäß etwas verloren. Aber genau deswegen ist Statistik ein Mittel, Informationen überhaupt erst transportierbar zu machen, so wie Sprache. Denn dasselbe geschieht, wenn ein Journalist eine Reportage schreibt oder ein Musiker sein Album in Form von MP3-Dateien veröffentlicht. Ein grundlegendes Verständnis von Statistik, von Demografie, Wirtschafts- und Sozialstatistik, aber auch von Wahrscheinlichkeitsrechnung steht deshalb auf einer Stufe mit der Fähigkeit zum Lesen und Schreiben. Statistik zu verstehen ist eine notwendige Fähigkeit, um die Welt, in der wir leben, einordnen und bewerten zu können und um Entscheidungen unter Unsicherheit zu treffen. Allein die Beantwortung der Frage, ob und wie viel bestimmte Statistiken aussagen, sei es für sich allein genommen oder im Kontext weiterer Informationen, setzt schon ein grundlegendes Verständnis für Statistik voraus.

Die zweite mögliche Lesart – es ist Demografie und deshalb in Ordnung – stellt Sarrazins Buch auf eine Stufe mit den Publikationen der statistischen Ämter. Auch das ist falsch. Statistische Ämter wählen nicht aus, um bestimmte (politische) Thesen zu stützen. Man kann sich zwar darüber streiten, ob die amtliche Statistik alles Relevante abbildet und ob umgekehrt alles in der amtlichen Statistik Abgebildete relevant ist, aber das Auswahlkriterium ist nicht eine bestimmte politische Haltung. (Gegen Ausnahmen kämpft beispielsweise die „Radical Statistics Group“ in Großbritannien: „Statistics should inform, not drive politics“.) Statistische Ämter wählen allerdings im Sinne einer Qualitätssicherung aus. Das heißt, ihre Quellen, Erhebungsmethoden und Grundgesamtheiten sind bekannt und anerkannt. Das heißt nicht, dass die amtliche Statistik perfekt und unan-

greifbar wäre. Auch bei ihr treten Fehler auf, aber diese werden, sobald sie bekannt sind, immerhin korrigiert. Bei wissenschaftlichen Studien geschieht das meist nicht. Doch nicht nur die Auswahl, sondern auch die Darstellung von Daten kann eine Manipulation bewirken. Absolutwerte, Prozentwerte, die Berechnung von Quoten – jede dieser Darstellungen betont bestimmte Aspekte der Daten und lenkt unsere Aufmerksamkeit weg von anderen. So zeigten Psychologen in Hongkong erst kürzlich in einem Experiment, dass Probanden sich relativ zu anderen beurteilten und nicht absolut. Konkret bewerteten sie ihre eigene Leistung ziemlich schlecht, wenn man ihnen sagte, sie hätten fünf Prozent Fehler gemacht und jemand anderes nur zwei Prozent. Ziemlich zufrieden waren sie aber, wenn sie erfuhren, dass sie 95 Prozent richtig hätten und die andere Person 98 Prozent ...

„Statistik ist für mich das Informationsmittel der Mündigen. Wer mit ihr umgehen kann, kann weniger leicht manipuliert werden. Der Satz ‚Mit Statistik kann man alles beweisen‘ gilt nur für die Bequemen, die keine Lust haben, genauer hinzusehen.“ Es sei daher notwendig, genau hinzusehen, fügte die Meinungsforscherin Elisabeth Noelle-Neumann hinzu. Genaues Hinsehen lässt sich im Wesentlichen auf zwei Grundaspekte zurückführen. Zuerst muss die Datenbasis stimmen, also das Rohmaterial, mit dem gearbeitet wird. Gleches gilt für das Handwerkszeug, mit dem diese Daten dann weiterverarbeitet werden. Schlechte Daten werden nicht besser durch ein komplexes statistisches Modell, und aus guten Daten werden falsche Aussagen, wenn die Methodik der Weiterverarbeitung nicht stimmt.

Bis dahin beschränkt sich statistisches Verständnis auf die routinierte Anwendung handlicher Fertigkeiten. Liest man gängige Bestseller über das „Lügen mit Statistik“, dann scheint die Angelegenheit auch recht einfach. Solche Bücher über den richtigen Umgang mit Statistiken gibt es zuhauf. Meistens handeln sie von sehr grundlegenden Problemen wie der Unterscheidung zwischen Median und Mittelwert oder dem richtigen Gebrauch von Grafiken. Sie verbleiben auf der Ebene der Beschreibung. Zudem sind die dargestellten Beispiele oft stark vereinfacht und stellen damit Journalisten, die scheinbar derart dumme Fehler machen, mehr oder weniger direkt als inkompetente Schreiberlinge dar.

Das Kernproblem ist ein gänzlich anderes. Statistisches Denken widerspricht unsererer Intuition, die von statistisch aufbereiteten „harten“ Zahlen und „objektiven“ Fakten in erster Linie Struktur und Objektivität erwartet. Doch Statistik ist nicht nur objektiv, Statistik ist nicht nur Handwerk. Statistik ist immer auch mit Subjektivität und Unsicherheit verbunden. Statistik arbeitet mit Wahrscheinlichkeiten, nicht mit Sicherheiten. Daraus erwachsen Fallstricke und Missverständnisse sowie oftmals überzogene Erwartungen, was Widersprüche zu unseren Denkgewohnheiten ergibt, die ständig nach „Wahrheiten“ und Mustern in unseren Beobachtungen suchen.

Die Nutzer von Statistik haben „häufig die Erwartung, nur mit genauen Zahlen auch genau planen zu können“, schreibt Ansgar Schmitz-Veltin und fährt fort: Dies „zu überwinden erfordert Mut“. Um den Mut, genauer hinzusehen und die Möglichkeiten wie auch die Grenzen von Statistik auszuloten, darum geht es in diesem Buch.

Zum Nachlesen:

Elhami, N.: Interview mit Frank Schirrmacher in DRadio Wissen, 02.09.2010.

Kwong, J. Y. Y. und Wong, K. F. E.: Fair or Not Fair? The Effects of Numerical Framing on the Perceived Justice of Outcomes. *Journal of Management* 40(6), S. 1558–1582, September 2014.

Leyendecker, H.: Was zur Hölle ist Recherche. Eröffnungsrede der nr-Fachkonferenz „Recherche reloaded“. Hamburg, 28.05.2011.

N. N.: Anti-Islam-Bewegung Pegida: Forscher zweifeln an Teilnehmerzahl der Demonstrationen. *Frankfurter Allgemeine Zeitung*, 19.01.2015.

Schmitz-Veltin, A.: Szenarien in der Stadtforschung – eine sinnvolle Ergänzung zu klassischen Vorausberechnungen? In: Szenarien zur demografischen, sozialen und wirtschaftlichen Entwicklung in Städten und Regionen, Themenbuch Stadtforschung und Statistik 1, S. 137–145, Köln, 2013.

1.2 Über Denkmuster

Statistik weckt bei Nicht-Statistikern zunehmend höhere Erwartungen. Nicht zuletzt liegt das an ihren neuen Kleidern. Heute sagt man „Big Data“, „Data Science“ oder „Analytics“ und meint damit eine modernere, „bessere“, oftmals auch nur eine praxisorientiertere Statistik. (Wer einen Blick auf die gängigen Statistik-Skripten an den Uni-

versitäten wirft, dem erschließt sich Letzteres ganz unmittelbar.)

Analytics, so definiert es das Institute for Operations Research and the Management Science (INFORMS), sei „the scientific process of transforming data into insight for making better decisions“. Das klingt neu und aufregend. Doch schon vor rund 25 Jahren beschrieb der indische Statistiker C. R. Rao die Statistik als „Methode, Informationen aus Daten zwecks Entscheidungsfindung zu extrahieren“. Beides macht glauben, dass sich mit modernen statistischen Methoden Unsicherheit besiegen lässt, dass sich Entscheidungen berechnen lassen. Rao klärt uns sofort auf: Sicheres Wissen entstehe in einer neuen Art des Denkens aus der Kombination von unsicherem Wissen und dem Wissen über das Ausmaß der Unsicherheit.

Aber für unsere vertrauten Denkmuster klingt das arg konstruiert. Wir verstehen unter sicherem Wissen das Gegenteil von unsicherem Wissen. Viele Nicht-Statistiker erhoffen sich von Statistik die Grundlage zu Entscheidungen auf Basis der Rationalität nackter Zahlen. Schließlich ist Mathematik neben der Philosophie die einzige Wissenschaft, die sich mit Wahrheiten beschäftigt. Schon in der Schule stand am Ende jeder Mathe-Aufgabe ein Wert, der entweder richtig oder falsch war. Der Kopf weiß, dass es bei Statistik um Wahrscheinlichkeiten geht, aber der Bauch erwartet klare „mathematische“ Aussagen. Als solche werden Statistiken dann aufgefasst.

Wenn wir als Statistiker uns dagegen wehren, führt dies bei Laien oft zu einem Umschwung ins Gegenteil. Wer zuvor noch an die Unbestechlichkeit der Zahlen glaubte, der kommt nun zur Überzeugung, dass Statistiker zu keinen

konkreten Aussagen fähig seien und der Instinkt dem ganzen Zahlenhokusokus sowieso überlegen sei. Eine derartige Überzeugung reift besonders dann, wenn die Statistik partout nicht zum gewünschten Ergebnis kommen will. Und solche Missverständnisse richten leider oft nachhaltigen Schaden an. Denn Statistik hilft der produzierenden Industrie, Ressourcen besser auszuschöpfen, Abfälle und Emissionen zu reduzieren und damit effizienter zu wirtschaften. Sie dient weitsichtigen Politikern dazu, Steuermittel verantwortungsvoll einzusetzen, indem sie beispielsweise demografische Entwicklungen bei ihrer Schulbedarfsplanung berücksichtigen. Kurz gesagt: Wer nachhaltige Entscheidungen treffen will, kommt an der Statistik als Mittel zur systematischen Datenanalyse nicht vorbei.

Eine wesentliche Fähigkeit, die von Führungskräften verlangt wird, ist Entscheidungskompetenz. Aber worin besteht diese? Sie ist sicher nicht die weit verbreitete Überzeugung, dass es genau eine richtige, gute und Erfolg versprechende Lösung gebe und mindestens eine falsche, schlechte, zum Scheitern verurteilte Alternative. Ein ähnlicher Denkfehler findet sich bei Journalisten: Es gebe etwas „Wahres“, und die Wahrheit sei dauerhaft. Es gebe etwas „Falsches“, „Schlechtes“, das ein Journalist aufdecken muss. Wie verlockend ist es, wenn Unternehmen in ihren Pressemitteilungen, Wissenschaftler in ihren Studien und Politiker in ihren Reden diese „Wahrheit“ auf dem Silbertablett präsentieren. Das Problem ist die Realität mit ihrem „Sowohl-als-auch“ und ihrer Unsicherheit.

Der grundlegende Denkfehler liegt in der Erwartung, dass Entscheidungen mit Hilfe von Datenanalyse berechnet werden können. Doch ohne Unsicherheit braucht man

keine Entscheidung. Und wenn es nichts zu entscheiden gibt, braucht man keine Entscheider.

Fehler beim Verständnis von Statistik entstehen häufig dann, wenn Entscheidungen und Meinungen eigentlich schon feststehen und die Statistik einer Art „post-dezisionistischen Argumentation“ dienen soll – sie rechtfertigt das, was man schon längst einfach so beschlossen hat. Das Traurige daran ist, dass viele inzwischen schon glauben, dass es „sowieso immer so läuft“. Und wenn die Statistik zur getroffenen Entscheidung passt, unterstellen deren Gegner häufig, dass diese Statistik nur passend ausgesucht oder passend gemacht worden sei. Bestätigt ein Gutachten die politische Linie der herrschenden Partei, dann waren die Gutachter eben gekauft. Belegt eine Studie im Auftrag eines Windelherstellers den ökologischen Vorteil von Wegwerfwindeln, dann war die Studie eben manipuliert.

Statistische Analysen schaffen für den, der statistisches Denken nicht gelernt hat, eine scheinbare Sicherheit – oder das genaue Gegenteil, ein abgrundtiefer Misstrauen gegen diese „größtmögliche Art des Lügens“. Dieses Problem wird im Zeitalter von Big Data massiv zunehmen. Einerseits gibt es Fälle wie den von Kenneth Rogoff, dem Harvard-Professor und früheren Chefökonom des Internationalen Währungsfonds. Rogoff und seine Kollegin Carmen Reinhart zogen aus ihren Daten recht radikale Schlüsse über die ökonomischen Folgen einer hohen öffentlichen Verschuldung. Prompt zitierten Politiker in den USA und Europa diese Ergebnisse eifrig als Beleg für die Notwendigkeit einer strikten Sparpolitik. Das war allerdings etwas voreilig, denn drei Jahre später stellte sich heraus: Die Wissenschaftler hatten sich verrechnet.

Auf der anderen Seite trauen Menschen offenbar eher ihrem eigenen Bauchgefühl als einem Algorithmus. Das untersuchten Forscher der Universität Pennsylvania in einem Experiment. Ihre Probanden sollten verschiedene Prognosen abgeben; zugleich bekamen sie Vorschläge für diese Prognosen von einem Computer. Sobald man ihnen zeigte, dass der Algorithmus nicht perfekt war, verließen sie sich fast ausschließlich auf ihre Intuition. Obwohl in Summe der Algorithmus nachweislich eine deutlich höhere Trefferquote besaß als jeder Mensch, verziehen ihm die wenigen Teilnehmer seine seltenen Fehler. Wie man Menschen trotzdem davon überzeugen kann, der Statistik zu trauen und nicht (nur) ihrer Intuition, darauf kann die Wissenschaft bis heute keine Antwort geben.

Statistische Analysen nutzen Algorithmen, mathematische Formeln und deren Umsetzung in Programmcode. Statistik ist dennoch weit mehr als eine Sammlung von Rezepten zum „Kneten“ von Daten. Statistik ist eine spezielle Art des Denkens. Ein guter Statistiker denkt nicht in „richtig“ oder „falsch“, in „null“ oder „eins“, in „sicher“ oder „unsicher“. Statistik, egal ob auf kleine oder unvorstellbar große Datenmengen angewandt, liefert vor allem eine Information: Sie beantwortet die Frage, wie groß die restliche Unsicherheit ist, die man selbst mit der kompliziertesten Mathematik, den leistungsstärksten Computern und den cleversten Statistikern auf diesem Planeten nicht beseitigen kann.

Warum fällt den meisten Menschen statistisches Denken so schwer? Eine Antwort auf diese Frage gibt der Nobelpreisträger Daniel Kahneman in seinem Buch „Schnelles Denken, langsames Denken“. Um den Alltag in einer kom-

plexen Welt, die ständig rasche Entscheidungen erfordert, zu bewältigen, müssen wir in der Lage sein, Ähnlichkeiten zwischen Situationen zu erkennen und Erfahrungen zu verallgemeinern. Wir entwickeln Heuristiken, überschlägige Denkregeln und Urteilsweisen, weil es einfach unmöglich ist, ständig alle Eventualitäten, Wechselwirkungen, Restriktionen und Folgen unseres Tuns abzuwägen. Entscheidungen „aus dem Bauch heraus“ sind solche, die auf unbewussten empirischen Analysen unserer Lebenserfahrung beruhen und daraus ableiten, was vermutlich richtig ist, weil es schon einmal ganz gut funktioniert hat.

Die gute Nachricht: Jeder von uns arbeitet ganz intuitiv mit Statistik. Wir haben eine ungefähre Vorstellung davon, wie viel Geld wir für den Wochenendeinkauf einstecken müssen, weil wir abschätzen, was unser Warenkorb im Durchschnitt kostet. Wir planen Pufferzeiten ein, wenn wir zu einem wichtigen Termin fahren, wobei wir Extremwertstatistik betreiben. Wir jonglieren mit Trends („diese Woche habe ich ein halbes Kilo abgenommen, da passt mir bis zum Sommerurlaub mein neuer Bikini“) und Korrelationen („um die Uhrzeit kam da noch nie ein Auto“). Wir analysieren also Daten in Form von Beobachtungen und Erfahrungen mit Hilfe empirischer Methoden und statistischer Verfahren, etwa der Berechnung von Mittelwerten. Und das tun wir jeden Tag.

Zum Nachlesen:

Dietvorst, B. J. et al.: Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General 144(1), S. 114–126, Februar 2015.

Eichengreen, B.: Der eigentliche Skandal. Die ZEIT, 02.05.2013.

Kahneman, D.: Schnelles Denken, langsames Denken. München, Siedler Verlag, 2012.

Rao, C. R.: Was ist Zufall? Statistik und Wahrheit. München, Prentice Hall, 1995.

1.3 Heuristik und Statistik

Der Unterschied zwischen Heuristik und Statistik liegt darin, dass Statistik erstens immer die Möglichkeit in Betracht zieht, dass die „Regelmäßigkeit“ unserer Daten auf Zufall beruht, und dass sie zweitens strikt zwischen Daten und Annahmen über diese Daten unterscheidet. Heuristik hinterfragt die Muster nicht. Dann werden wir überrascht wie in C. R. Raos Anekdote der Königsohn, der eines Tages unter den Bewohnern seines Reiches auf einen Mann traf, welcher ihm zum Verwechseln ähnlich sah. „War deine Mutter jemals in meinem Palast angestellt“, fragte er diesen. „Nein“, entgegnete der Mann, „aber mein Vater.“

„Fooled by randomness“ nennt Nassim Taleb, Börsenhändler und Bestseller-Autor, die menschliche Neigung, Zufallsprodukte als regelmäßige und systematische Muster zu deuten. Lässt man eine Million Affen auf Schreibmaschinen herumhämmern, könnte vielleicht irgendwann einer von ihnen Shakespeares Werke produzieren. Bloß macht das den Affen nicht zum Schriftsteller. Wollen wir wirklich darauf wetten, dass dieser Affe als nächstes Goethes „Faust“ eintippt? Wohl kaum. Doch wenn eine Million Händler an

der Börse zocken und einer von ihnen irgendwann durch Zufall astronomisch hohe Renditen erzielt, dann vertrauen ihm viele blindlings ihr Vermögen an.

Lassen wir eine Software beliebig viele Modelle an einem Datensatz berechnen, dann wird irgendeines davon rein durch Zufall perfekt passen. So funktioniert Data Mining. Es ist fast sicher, dass ein anderes Modell besser passt, sobald neue Daten hinzukommen. Wer glaubt, dass sich die Zukunft genau so fortsetzt wie die Vergangenheit, der verkennt, dass irgendein Ereignis der Vergangenheit von Zufall beeinflusst ist. Statistik bedeutet, den Zufall von der Wahrheit zu trennen. Data Mining ohne Nachdenken bedeutet, dass wir den Zufall zur Wahrheit erklären.

Wir sollten nicht gleich ein System vermuten, nur weil wir ein mathematisches Modell finden, das die Daten der Vergangenheit exakt nachbilden kann. Denn so ein Modell gibt es immer. Beispielsweise lassen sich zwei zufällig ausgewählte Datenpunkte genau durch eine Gerade beschreiben. Ob weitere, unbeobachtete Datenpunkte zwischen, vor oder hinter diesen beiden Punkten ebenfalls auf der Geraden liegen, ist keineswegs sicher. Deswegen sind lineare Interpolationen oder Trends mit höchster Vorsicht zu betrachten. Wir Menschen neigen dazu, derartige Muster in beobachteten Daten zu suchen und diese Muster in dem Moment, da wir sie entdecken, als „typisch“ für die Daten zu identifizieren. Daraus leiten wir ab, dass starke Abweichungen von diesen Mustern etwas zu bedeuten hätten, obgleich sie einfach zufällige Streuungen sein können.

Das Phänomen der „Regression zum Mittelwert“ beschreibt das statistische „Gesetz“, nach dem in einer Folge zufälliger Messwerte auf einen extrem ausgefallenen Wert

fast sicher einer folgt, der deutlich näher am Mittelwert der Daten liegt. Sich das vorzustellen ist nicht einfach, und so tendieren wir dazu, auf Ausreißer alarmiert zu reagieren. Die Statistik kennt deshalb das Konzept der *Signifikanz*, das uns beurteilen lässt, ob Strukturen oder Trends systematisch sind und ab wann Abweichungen von diesen Strukturen nicht mehr durch den Zufall zu erklären sind. Dass eine Beobachtung höchstwahrscheinlich keine zufällige Schwankung darstellt, heißt noch lange nicht, dass sie irgendeine praktische Bedeutung hätte, und so ist *Relevanz* etwas ganz anderes als Signifikanz.

Bei sehr großen Stichproben lassen sich sehr schwache Zusammenhänge oder sehr kleine Unterschiede als „signifikant“ nachweisen, aber praktisch gesehen mögen diese völlig vernachlässigbar sein. Bei sehr kleinen Stichproben werden hingegen womöglich selbst große Effekte nicht signifikant, sind nach statistischer Bewertung also Zufallsprodukte, auch wenn sie real eine hohe Bedeutung haben und man eigentlich darauf reagieren müsste.

Von Stichproben ist es nicht weit zum Begriff der *Repräsentativität*. Repräsentativität ist eines der wichtigsten Konzepte der angewandten Statistik, zugleich aber sehr schwer sicherzustellen. Anekdotische Einzelbeobachtungen sprechen Leser emotional an, Häufigkeitsdaten hingegen wirken sachlich und nüchtern: „Der Tod eines Menschen: Das ist eine Katastrophe. Hunderttausend Tote: Das ist eine Statistik!“, formulierte es Kurt Tucholsky. Dennoch stehen Daten in Bezug zu den realen Gegebenheiten, die dahinterstecken. Daten werden über Personen, Orte oder Zeitpunkte erhoben, die man statistische Subjekte nennt. Sie können mittels Fragebogen, Maßband oder komplexen Messinstrumenten gewonnen worden sein.