

Rezensionen



CHRISTOF WOLF & HENNING BEST (Hg.), 2010: Handbuch der sozialwissenschaftlichen Datenanalyse. Wiesbaden: VS Verlag für Sozialwissenschaften. ISBN: 978-3-531-16339-0, 1.098 Seiten, 79,95 EUR.

Christof Wolf (GESIS und Universität Mannheim) und Henning Best (GESIS, Mannheim) haben sich mit dem Handbuch der sozialwissenschaftlichen Datenanalyse das Ziel gesetzt, der quantitativen empirischen Sozialforschung ein deutschsprachiges Einführungs-, Überblicks- und Nachschlagewerk zu elaborierten Analyseverfahren zur Verfügung zu stellen. Primäre Adressaten des Herausgeberbandes sind alle Anwender fortgeschrittener quantitativ-empirischer Methoden und damit auch Doktoranden und Studierende der Sozialwissenschaften in höheren Fachsemestern (S. 4). Auf insgesamt rund 1.100 Seiten (sic!) wird – unter Mitzählung der einleitenden Bemerkungen von Henning Best und Christof Wolf – in 40 Einzelbeiträgen ein methodologischer und methodischer Bogen gespannt, der von A wie *Analyse kategorialer Daten* (Kapitel 18 von Hans-Jürgen Andreß) bis Z wie *Zeitreihenanalyse* (Kapitel 40 von Rainer Metz) reicht. Sofern die spezifische Zielsetzung eines Kapitels nicht ein abweichendes Vorgehen erfordert, folgen alle Beiträge dabei einem formal gleichen Aufbau. Zunächst wird das behandelte Analyseverfahren in allgemeinen Worten vorgestellt, bevor seine mathematischen Grundlagen dargestellt und diskutiert werden. Im Anschluss werden

exemplarische Anwendungen aufgezeigt, wobei in den meisten Beiträgen möglichst einheitlich ALLBUS- oder SOEP-Daten als Datenbasis herangezogen werden. Den Abschluss eines jeden Kapitels bilden Hinweise zu häufig vorkommenden Anwendungsfehlern, die insbesondere für diejenigen von Interesse sind, die sich in eine bestimmte Analyseverfahren neu einarbeiten wollen, und kommentierte Hinweise auf weiterführende Literatur zum behandelten Verfahren. Jedem Kapitel ist zudem ein viertel- bis einseitiger Abstract vorangestellt, der einen kurzen Überblick über die behandelten Inhalte, den konkreten Aufbau und die Kernaussagen des Beitrages gibt.

Thematisch ist das Handbuch der sozialwissenschaftlichen Datenanalyse in sechs Sinnabschnitte gegliedert. Den ersten Teil *Einführung* bilden die einleitenden Worte von Best und Wolf sowie ein Beitrag von Karl-Dieter Opp zur *Kausalität als Gegenstand der Sozialwissenschaften und der multivariaten Statistik* (Kapitel 2). In der für ihn charakteristischen, prägnanten und präzisen Art behandelt Opp das Kausalitätsproblem, das auf Grund der Struktur der nutzbaren Daten vielen Analysen der empirischen Sozialforschung immanent ist. Obwohl es sich bei diesem Kapitel im engeren Sinne nicht um einen Beitrag zu Datenanalyseverfahren handelt, schärfen Opps Ausführungen noch einmal den Blick dafür, was statistische Analyseverfahren in den Sozialwissenschaften zu leisten vermögen und was nicht. Insofern ist der Beitrag wichtig und mit seiner Einordnung am Anfang des Handbuchs richtig platziert.

Der zweite Teil des Handbuchs *Grundlagen der Datenanalyse* versammelt – wie der Titel bereits nahelegt – Arbeiten, die sich mit den Basics der sozialwissenschaftlichen Datenanalyse befassen. Manuela Pötschke widmet sich in ihrem Beitrag *Datengewinnung und Datenaufbereitung* (Kapitel 3)

den Problemen, die durch die Auswahl des Erhebungsverfahrens und im Prozess der Datengewinnung und -aufbereitung entstehen können. Im Zentrum der Betrachtung stehen dabei die verschiedenen Formen der Befragung. Cornelia Weins gibt in ihrem Beitrag *Uni- und bivariate deskriptive Statistiken* (Kapitel 4) einen Überblick über die gängigen Lage- und Streuungsmaße sowie grundlegende Verfahren der bivariaten Datenauswertung. Horst Degen widmet sich der *Graphische(n) Datenexploration* (Kapitel 5), wobei er den Schwerpunkt auf die adäquate Visualisierung univariater Verteilungen legt. *Der Umgang mit fehlenden Werten* (Kapitel 6) wird anschließend von Martin Spieß thematisiert. Er beleuchtet dabei vor allem Möglichkeiten und Grenzen von Imputationsverfahren, greift aber auch Aspekte der Datengewichtung auf. Explizit um *Gewichtung* (Kapitel 7) geht es in dem Beitrag von Siegfried Gabler und Matthias Ganninger. Die Autoren erklären Prozesse und Ziele von Gewichtungsverfahren anschaulich und illustrativ am Beispiel der dritten Erhebungswelle des ESS. Der Beitrag *Grundlagen des Statistischen Schließens* (Kapitel 8) von Steffen M. Kühnel und Dagmar Krebs schafft dann einen Übergang zur Inferenzstatistik. Den Schwerpunkt setzen Kühnel und Krebs bei den gängigen Schätz- und Testverfahren. Einen entgegengesetzten Weg beschreitet Susumu Shikano in seiner *Einführung in die Inferenz durch den nichtparametrischen Bootstrap* (Kapitel 9) in dem er das weniger genutzte (und wohl auch weniger bekannte) Bootstrapping in seiner Logik und seiner Anwendung anschaulich erläutert. Den Abschluss des zweiten Teils des Handbuchs bilden Thomas Gautschis Ausführungen zur *Maximum-Likelihood Schätztheorie* (Kapitel 10). Sein Beitrag setzt zwar ein mathematisches Verständnis auf etwas höherem Niveau voraus, bietet aber gleichzeitig eine wirklich gute Gelegenheit, die Logik von ML-Verfahren kompakt nachzuvollziehen.

Im dritten Teil des Handbuchs geht es um *Messen und Skalieren*. Beatrice Rammstedt widmet sich hier zunächst ganz grundsätzlich den Themen *Reliabilität, Validität, Objektivität* (Kapitel 11), bevor sich Joachim Gerich mit der *Thurstone- und Likertskalierung* (Kapitel 12) sowie der *Guttman- und Mokkenskalisierung* (Kapitel 13) befasst. Zusammengenommen bieten beide Kapitel nicht nur einen guten Einblick in die Logik der jeweiligen Skalierungsverfahren, sondern lassen auch die Unterschiede zwischen ihnen pointiert hervortreten. Im 14. Kapitel thematisieren Christian Geiser und Michael Eid dann die *Item-Response-Theorie*. Im Zentrum stehen dabei Variationen des Rasch-Modells, an dem sie exemplarisch und anschaulich die Logik und Anwendung von IRT-Modellen verdeutlichen. *Hauptkomponentenanalyse und explorative Faktorenanalyse* (Kapitel 15) sind das Thema des Beitrages von Hans-Georg Wolff und Johann Bacher. Die beiden Autoren bieten einen kompakten und zugleich doch umfassenden ersten Einblick in diese Form der Datenanalyse bzw. Datenreduktion. Den Abschluss des dritten Teils bilden Beiträge von Jörg Blasius zur *Korrespondenzanalyse* (Kapitel 16) und von Ingwer Borg zur *Multidimensionalen Skalierung* (Kapitel 17). Beide Autoren gehören auf ihrem jeweiligen Gebiet ohne Zweifel zu den führenden Experten, was den Beiträgen unschwer anzumerken ist.

Der vierte Teil des Handbuchs *Analyse von Häufigkeiten, Gruppen und Beziehungen* wird mit einem Beitrag von Hans-Jürgen Andreß zur *Analyse kategorialer Daten* (Kapitel 18) eröffnet, der sich darauf konzentriert, Besonderheiten und Unterschiede von Logit-Modellen, log-linearen Modellen und der gewichteten Regressionsanalyse nach Grizzle, Starmer und Koch herauszuarbeiten. Im Anschluss geben Manuel C. Völkle und Edgar Erdfelder einen gelungenen Einblick in die *Varianz- und Kovarianzanalyse* (Kapitel 19), bevor Reinhold Decker, Silvia Rašković und Kathrin Brunsiek kompetent

über das Verfahren der *Diskriminanzanalyse* (Kapitel 20) informieren. Die *Clusteranalyse* (Kapitel 21) ist Thema des anschließenden Beitrages von Michael Wiedenbeck und Cornelia Züll. Sie konzentrieren sich auf die Behandlung der Clusterzentrenanalyse und agglomerativer Verfahren der Clusterbildung. Johann Bacher und Jeroen K. Vermunt geben einen instruktiven ersten Einblick in die *Analyse latenter Klassen* (Kapitel 22), wobei sie u.a. die Probleme der Validitätsprüfung ins Zentrum ihrer Ausführungen rücken. Im letzten Kapitel des vierten Teils behandeln Hans J. Hummell und Wolfgang Sodeur – zwei Wegbereiter der sozialwissenschaftlichen Netzwerkforschung in Deutschland – mit großer Expertise Logik und Verfahren der *Netzwerkanalyse* (Kapitel 23).

Mit dem fünften Teil *Regressionsverfahren für Querschnittsdaten* wird ein Einstieg in die nach wie vor wachsende Familie der regressionsanalytischen Methoden gegeben. Nachdem Christof Wolf und Henning Best hier zunächst kompakt und kompetent die *Lineare Regressionsanalyse* (Kapitel 24) vorstellen, erörtert Dieter Ohr vertiefend und versiert *Modellannahmen und Regressionsdiagnostik (der linearen Regression)* (Kapitel 25), bevor Henning Lohmann spezifisch und instruktiv *Nicht-Linearität und Nicht-Additivität in der multiplen Regression* (Kapitel 26) diskutiert. Zusammengekommen bieten die drei Kapitel einen wirklich guten Einstieg für die fortgeschrittene Handhabung linearer Regressionsmodelle. Ben Jann widmet sich in seinem anschließenden Beitrag zur *Robuste(n) Regression* (Kapitel 27) den spezifischen Problemen, die die so genannten Ausreißer in OLS-Regressionen hervorrufen können, und gibt einen Einblick in regressionsdiagnostische Verfahren jenseits der üblichen Residuenanalyse. Wolfgang Langer offeriert im Anschluss in gewohnter Kompetenz einen Überblick über die *Mehrebenenanalyse mit Querschnittsdaten* (Kapitel 28), bevor Jost Reinecke und Andreas Pöge sich versiert der Logik und der Handhabung von

Strukturgleichungsmodelle(n) (Kapitel 29) annehmen und Petra Stein die besonderen Probleme der *Regression mit unbekanntem Subpopulationen* (Kapitel 30) diskutiert. Das 31. Kapitel *Logistische Regression* von Henning Best und Christof Wolf stellt – analog dem 24. Kapitel – eine kompetent geschriebene Einführung in den Umgang mit (binären) logistischen Regressionsanalysen dar. Ergänzt und vertieft wird die Darstellung durch die nachfolgende versierte Abhandlung von Steffen M. Kühnel und Dagmar Krebs zur *Multinomiale(n) und ordinale(n) Regression* (Kapitel 32), bevor Gerhard Tutz instruktiv die Besonderheiten der *Regression für Zählvariablen* (Kapitel 33) darlegt und Gerrit Bauer zum Abschluss des fünften Teils Hinweise für die *Graphische Darstellung regressionsanalytischer Ergebnisse* (Kapitel 34) gibt.

Der sechste und letzte Teil des Handbuchs schlägt schließlich den Bogen zur *Analyse von zeitbezogenen Daten*. Markus Gangl thematisiert hier zunächst die *Nichtparametrische Schätzung kausaler Effekte mittels Matchingverfahren* (Kapitel 35). Josef Brüderl geht im Anschluss auf die *Kausalanalyse mit Paneldaten* (Kapitel 36) ein, bevor Hans-Peter Blossfeld die *Survival- und Ereignisanalyse* (Kapitel 37) vorstellt. Das Thema des Beitrages von Florian Schmiedek und Julia K. Wolff sind *Latente Wachstumskurvenmodelle* (Kapitel 38) und Stefani Scherer gibt gemeinsam mit Josef Brüderl einen Einblick in die *Sequenzdatenanalyse* (Kapitel 39), bevor Rainer Metz last but not least die *Zeitreihenanalyse* (Kapitel 40) eingehend erörtert. Summarisch lässt sich festhalten, dass alle sechs Beiträge dieses Teils die jeweiligen Autorinnen und Autoren als versierte Kenner der behandelten Materie ausweisen und in der Regel zugleich auch ihr didaktisches Können dokumentieren. Gleichwohl bringt die longitudinale Perspektive der Analyseverfahren eine gewisse Komplexitätssteigerung mit sich, die die Auseinandersetzung – möglicherweise – erschwert. Interessierten Neueinsteigern sei daher empfohlen, sich auf jeden Fall

zunächst noch einmal ihrer Grundkenntnisse der sozialwissenschaftlichen Datenanalyse zu versichern, bevor sie sich mit dem spezifischen Feld der Analyse zeitbezogener Daten auseinandersetzen.

Ein Mammutwerk wie das Handbuch der sozialwissenschaftlichen Datenanalyse bietet zwangsläufig auch Anhaltspunkte für Detailkritik und es wäre billig, hier exemplarisch einzelne Aspekte herauszugreifen. Dies würde jedoch in keiner Weise der Leistung des Herausgeberbandes gerecht und daher soll im Folgenden eher eine Gesamtwürdigung versucht werden. An allererster Stelle ist dabei herauszustellen, dass die Umsetzung des Konzepts insgesamt als sehr gelungen bezeichnet werden kann und dass das Handbuch der sozialwissenschaftlichen Datenanalyse in keiner einschlägigen Fachbibliothek fehlen sollte. Ungeachtet aller qualitativen Schwankungen, die in einem solchen Werk zwischen einzelnen Beiträgen stets bestehen, ist der Herausgeberband insgesamt sowohl mit Blick auf die versammelte fachliche Expertise wie auch hinsichtlich der didaktischen Aufbereitung der Inhalte auf einem sehr hohen Niveau anzusiedeln. Dazu dürfte nicht zuletzt beigetragen haben, dass alle Einzelabhandlungen einem eingehenden Review-Verfahren unterzogen wurden (S. 5). Als weiterer wichtiger Pluspunkt ist in diesem Zusammenhang die Begleithomepage zum Handbuch zu nennen (<http://www.handbuch-datenanalyse.de/>). Auf ihr können nicht nur Zusammenfassung und Gliederung der einzelnen Kapitel sowie die Kurzportraits der Autorinnen und Autoren eingesehen werden. Zudem sind dort die – wohl unausweichlichen – Errata dokumentiert und es besteht die Möglichkeit, die Analysesyntax zu den in den einzelnen Kapiteln behandelten Auswertungen abzurufen oder über ein dafür eingerichtetes Kontaktformular Anfragen und Kommentare an die Herausgeber zu senden.

Grundsätzlich kritisch anzumerken sind lediglich vier Punkte: Zunächst muss festgehalten werden, dass das Handbuch über kein

Stichwortregister verfügt, was für ein Werk mit dem oben skizzierten Anspruch einen deutlichen Mangel darstellt. Davon abgesehen fällt in der Gesamtschau auf, dass die Aufgabe, Hinweise zu häufigen Fehlern zu geben, von den Autorinnen und Autoren mit sehr unterschiedlichem Einsatz gelöst wurde. Während viele Kapitel sich dadurch auszeichnen, dass die Autoren mit viel Engagement bemüht sind, Novizen des diskutierten Verfahrens vor den typischen Anfängerfehlern zu bewahren, vermitteln einzelne Kapitel hier eher den Eindruck einer unliebsamen formalen Anforderung nachzukommen. Dabei dürfte gerade die Diskussion häufig auftretender Fehler für diejenigen, die sich an einem für sie neuen Analyseverfahren erproben wollen, eine Schlüsselpassage des jeweiligen Kapitels sein – ganz zu schweigen davon, dass sie auch für fortgeschrittene Anwender eine Art Vergewisserung darstellen kann, methodisch-analytisch sauber zu arbeiten. Kurzum: Hier hätten die Herausgeber strenger auf die durchgängige Einhaltung der aufgestellten Maßstäbe pochen sollen.

Des Weiteren ist anzumerken, dass insbesondere einige der im zweiten und dritten Teil des Handbuchs versammelten Beiträge nach Dafürhalten des Rezensenten einen anderen Personenkreis, als die von den Herausgebern anvisierten Gruppen, adressieren. Ohne dadurch die inhaltliche und didaktische Qualität der Beiträge schmälern zu wollen, behandeln sie dennoch Gegenstände, die eher der grundständigen universitären Ausbildung in empirischer Sozialforschung zuzurechnen sind und die spätestens bei fortgeschrittenen Studierenden als selbstverständliches Basiswissen voraussetzbar sein sollten. (Wobei natürlich auch denkbar ist, dass der Rezensent sich an dieser Stelle irrt.) Selbstverständlich spricht nichts dagegen, in einem Handbuch der sozialwissenschaftlichen Datenanalyse auch grundlegende Inhalte zu behandeln und den Adressatenkreis des Werkes entsprechend auf Studierende in der Anfangsphase ihrer

Ausbildung auszuweiten. Nur wäre es dann vielleicht empfehlenswert gewesen, den basalen Themenbereichen mehr Raum zu gewähren und sie in kleineren Sinneinheiten detaillierter zu behandeln.

Schließlich ist festzuhalten, dass für ein Handbuch der sozialwissenschaftlichen Datenanalyse noch einige weitere Kapitel zu einer Reihe anderer Themenbereiche denkbar gewesen wären. Die Herausgeber weisen zwar in ihren einleitenden Bemerkungen zu Recht darauf hin, dass ein vollständiger Überblick über den Stand des Faches in *einem* Buch schlechterdings nicht möglich ist (S. 5), gleichwohl hat man bereits bei der Durchsicht des Inhaltsverzeichnisses das Gefühl, dass bestimmte Themenkomplexe noch hätten Berücksichtigung finden können. Beispielsweise hätte sich der Rezensent Kapitel zu *Analysestrategien*, zur Conjoint-Analyse und zu vielversprechenden Innovationen in der sozialwissenschaftlichen Datenanalyse gewünscht. Anderen Leserinnen und Lesern des Handbuchs wird es mit anderen Themen wahrscheinlich ähnlich ergehen. So paradox es daher für ein Buch von über eintausend Seiten klingen mag: Mehr wäre mehr gewesen. Allerdings muss es ja bei Neuauflagen – zu denen es aufgrund des Mehrwerts des Sammelbandes und der daraus resultierenden Nachfrage ohne Zweifel kommen wird – nicht zwangsläufig bei einem einbändigen Werk bleiben.

ULRICH ROSAR, UNIVERSITÄT DÜSSELDORF

* * * * *