

REZENSION

M. Borovcnik – G. Ossimitz:

Materialien zur Beschreibenden Statistik und Explorativen Datenanalyse

Schriftenreihe Didaktik der Mathematik, Universität für Bildungswissenschaften in Klagenfurt; Band 11

Wien: Verlag Hölder-Pichler-Tempsky, 1987

Stuttgart: Verlag B.G. Teubner, 1987 – 226 S.

Heinz Trauerstein, Bielefeld

1. Vorbemerkungen

Die Explorative Datenanalyse (EDA) ist ein relativ neuer Bereich der Statistik. Es geht hier um das Explorieren von Daten, um die Suche nach Mustern und Strukturen in Datensätzen. Tukey vergleicht in seinem grundlegenden Werk (1977) die Explorative Datenanalyse mit der Tätigkeit eines Detektivs auf der Spurensuche. Er braucht *Hilfsmittel* wie z.B. feines Pulver zum Erkennen von Fingerabdrücken und *Sachverstand*, um an den richtigen Stellen nach Fingerabdrücken zu suchen. Die Hilfsmittel der EDA sind vor allem graphische Darstellungen, durch die Besonderheiten und Regelmäßigkeiten innerhalb der Daten offengelegt werden sollen. Der Sachverstand bezieht sich auf die Vertrautheit mit der Sachsituation, aus der die Daten stammen. Der Beurteilenden Statistik entspricht nach Tukey die Tätigkeit des Richters. Er prüft, welche Gesetze herangezogen werden können und ob die vorliegenden Fakten zur Urteilsfindung ausreichen. Ebenso werden in der Beurteilenden Statistik aufgrund des Skalenniveaus und der verteilungsmäßigen Voraussetzungen die statistischen Verfahren ausgewählt. Mit ihrer Hilfe entscheidet man, ob die vorliegenden Daten ausreichen, um eine Hypothese anzunehmen bzw. zu verwerfen.

Die Beschreibende Statistik, deren vornehmliche Aufgabe im Zusammenfassen und Darstellen von Daten besteht, hat durch die Vorgehensweise und die Arbeitstechniken der EDA eine bedeutsame Ausweitung erfahren. Als großer Vorteil der EDA für den Unterricht wird die relative Elementarität der mathematischen Mittel im Vergleich zur Beurteilenden Statistik gesehen. Der Lehrer kann in der Sekundarstufe I schon "interessante" Fragestellungen in Angriff nehmen, die er wegen der wahrscheinlichkeitstheoretischen Voraussetzungen der Beurteilenden Statistik bisher allenfalls in der Sekundarstufe II ansprechen konnte. Biehler (1982) weist im Vorwort seines Buches darauf hin, daß man von der EDA neue Impulse für die Gestaltung eines anwendungsbezogenen Stochastikunterrichts erwartet, da ihre Techniken den Schülern die Möglichkeit für einen selbsttätigen experimentellen Umgang mit Daten bieten. Auch Engel (1982, S. 58) verspricht sich von der EDA eine Aufwertung der Beschreibenden Statistik, die bisher im Unterricht ein "ideenarmes Thema von geringem Umfang" war.

Vor diesem Hintergrund durfte man auf das vorliegende Buch von Manfred Borovcnik und Günther Ossimitz gespannt sein und hoffen, daß die Autoren dem Leser etwas von der graphischen und numerischen Detektivarbeit, von dem Suchen nach Spuren und Auffälligkeiten in Daten vermitteln können. Um es vorweg zu sagen: Ich bin nicht enttäuscht worden.

2. Zum Inhalt

Das Buch gliedert sich in 4 Teile, die unabhängig voneinander gelesen werden können:

Teil A: Explorative Arbeitstechniken – Fallbeispiele

Teil B: Statistische Grundtätigkeiten

Teil C: Explorative Datenanalyse und politische Bildung im Unterricht

Teil D: Zum Hintergrund der explorativen Datenanalyse.

Im Teil A werden auf 72 Seiten 7 Fallbeispiele ausführlich besprochen. Die Autoren haben sie so ausgewählt, daß grundlegende Arbeitstechniken der EDA in konkreten Sachsituationen mit realen Daten erarbeitet werden. Dabei geht es um folgende Sachzusammenhänge:

1. Beherbergungsbetriebe in Klagenfurt,
2. Tore im europäischen Fußball,
3. die Keimfähigkeit verschiedener Gemüsesorten,
4. Niederschläge in Afrika, Südamerika und Australien,
5. Seehöhe und Durchschnittstemperatur,
6. die Handelsflotten der Welt,
7. Länge und Einzugsgebiet von Flüssen.

In einem 8. Abschnitt werden Vorschläge zu 5 kleinen Unterrichtsprojekten mit folgenden Themen gemacht: Unser Alter, Dauer von Musiktiteln, der Schulweg, ein Reaktionstest und Ergebnisse von Schularbeiten.

Im Teil B wird ein Sachthema auf 50 Seiten exemplarisch vertieft. Ausgehend von einem Zeitungsartikel zum Thema Scheidungen analysieren die Autoren zunächst das Scheidungsverhalten in Kärnten und konstruieren verschiedene mathematische Modelle zur Erklärung der in den Daten beobachteten Phänomene. Es wird untersucht, wie gut die Modelldaten jeweils die reale Entwicklung wiedergeben und ob die Modelle auf andere Datensätze – hier die Scheidungszahlen für ganz Österreich – anwendbar sind.

Auch Zusammenhänge zwischen Scheidungen und der Ehedauer bzw. Scheidungen und der Kinderzahl sprechen die Autoren an. Es wird deutlich, wie vorsichtig man bei der Interpretation von erkannten Mustern oder Abweichungen sein muß, und daß weitere Daten, die man zur Beantwortung einer Frage heranzieht, in der Regel wieder neue Fragen aufwerfen, die eigentlich noch geklärt werden sollten. Angeregt durch die Kritik an einem Zeitungsartikel stellen die Verfasser die Frage nach der Messung von "Scheidungsintensität" und diskutieren dazu verschiedene mögliche Beziehungszahlen. Hier zeigt sich – wie die Autoren auch betonen (S. 17) –, daß die Schwierigkeiten bei der EDA weniger in der Mathematik als vielmehr in der (außermathematischen) Sache begründet sind und daß "der Verlauf der Analyse von der Kenntnis der Sachzusammenhänge geleitet wird" (S. 201).

Im letzten Kapitel von Teil B werden einige Anregungen für die Behandlung weiterer Projektthemen im Unterricht gegeben. Die Vorschläge gruppieren sich um die Themenbereiche Arbeitswelt, Wachstum und Verkehr und beinhalten auch Hinweise, wie man sich entsprechende Daten beschaffen kann.

Der Teil C besteht aus zwei ausführlich dargestellten Unterrichtsreihen.

Bernhard Kröpfl hat in seiner Unterrichtseinheit zum Thema "Wir lesen Zeitung" Darstellungsmittel der beschreibenden Statistik wie: Tabellen, Zackenkurven (Polygonzüge), Stabdiagramme und Kreisdiagramme in einer 3. Klasse AHS (das entspricht bei uns einem 7. Schuljahr am Gymnasium oder der Realschule) behandelt. Besonders interessiert waren die Schüler an den "legalen" Manipulationsmöglichkeiten von Graphiken. Für Stabdiagramme und Polygonzüge werden eine Vielzahl solcher Möglichkeiten anhand von Schaubildern aus Zeitungen vorgestellt. Am Ende drängt sich dem Betrachter die Erkenntnis auf, daß es die optimale Darstellung für einen Sachverhalt gar nicht gibt, daß Manipulation mithin unumgänglich ist. Der Autor meint daher, "daß es nie früh genug sein kann, unsere Schüler in Manipulationstechniken zu unterweisen", und zwar so, daß die Schüler diese Techniken selber anwenden und deren Wirkung bewußt erleben (S. 156). Den an weiteren motivierenden Beispielen zur Beschreibenden Statistik interessierten Leser möchte ich in diesem Zusammenhang auf das Heft von G. Fillbrunn (1981) hinweisen, das auch für den Unterricht in Klasse 7 konzipiert ist.

In der zweiten Unterrichtsreihe von Teil C stellt Helga Jungwirth ein Projekt vor, bei dem eine 6. Klasse AHS (10. Schuljahr) den eigenen Mathematikunterricht mit Methoden der EDA untersucht. Die Schüler und Schülerinnen haben in Gruppenarbeit Fragebögen zum Mathematikunterricht mit Items zur Notengebung, zur Einstellung und zum Interesse am Fach Mathematik, sowie zur Beurteilung verschiedener Unterrichtsformen entwickelt. Die Fragebögen und ihre Auswertungen sind im Anhang beigelegt.

Zum Schluß ihres Beitrages skizziert die Autorin "eine Philosophie des Unterrichts von Explorativer Datenanalyse" (S. 167 ff.), die andeutungsweise durch Begriffe wie "offene Mathematik", "problemorientierter Unterricht", "Platz für Eigeninitiative" und "soziales Lernen" charakterisiert werden kann.

Der Teil D beginnt mit einer überblickartigen Zusammenfassung und Charakterisierung der im Buch benutzten explorativen Arbeitstechniken. Bei der Analyse eindimensionaler Daten sind das: Stamm- und Blatt-Diagramm, Zahlenzusammenfassung und Kastenschaubilder. Bei der Analyse zweidimensionaler Daten beginnt man in der Regel mit der Darstellung der Daten durch Punktwolken und kann je nach vorliegender Situation in verschiedene Richtungen weiterarbeiten: "fit-and-residual" Analyse, Datentransformation, Zerlegung von Punktwolken oder Quotientenbildung aus beiden Variablen. Charakteristisch für diese Arbeitstechniken ist nach Darstellung der Autoren ihre Flexibilität und Robustheit, der geringe Aufwand, die praktische Ausrichtung, das visuelle Arbeiten und vor allem die Möglichkeit, die im Datenmaterial versteckten Besonderheiten und Regelmäßigkeiten deutlich hervortreten zu lassen.

Am Ende des Buches stellen die Autoren die Leitideen der EDA zusammen und vergleichen sie mit der klassischen Statistik. "Die EDA versteht sich demnach nicht als Konkurrenz ... zur Beurteilenden Statistik sondern als eine Ergänzung, die den Weg zu neuen Hypothesen aus vorliegenden Daten erleichtern soll" (s. 199). "Wir sehen Explorative Datenanalyse ... als eine konsequent eigenständig entwickelte Beschreibende Statistik, die nicht bloß eine Propädeutik zur Beurteilenden 'Mathematischen' Statistik bleibt, sondern bereits für sich ein mächtiges Werkzeug zur Gewinnung neuer Einsichten bietet" (S. 199).

Abschließend gehen die Autoren auf ihr Verständnis von Mathematik im Sinne einer "Offenen Mathematik" und auf die sich daraus ableitenden Perspektiven für den Unterricht

ein. Dabei beziehen sie sich auf das Buch von Fischer/Malle (1985) (S. 10).

3. Stellungnahme zu einzelnen Punkten

Zunächst muß darauf hingewiesen werden, daß das Buch für die Bedürfnisse von Schule und Unterricht konzipiert wurde und nicht für die Anwendung der EDA in der Forschung und der Praxis. Denn dort werden auch mathematisch anspruchsvollere Verfahren herangezogen und vor allem kommen Computerprogramme verstärkt zum Einsatz. Meiner Meinung nach bieten sich auch im Unterricht Möglichkeiten, die hier vorgestellten "Papier- und Bleistift-Techniken" durch den Einsatz des Computers sinnvoll zu ergänzen. Im Buch wird dieses Thema leider nur in einem Satz erwähnt (S. 17). Der Lehrer sollte auch etwas davon erfahren, daß die EDA durchaus kontrovers diskutiert wird und einige Statistiker ihr ablehnend gegenüberstehen.

Nun zum zentralen Anliegen der Autoren: der Spurensuche in den Daten und ihrer inhaltlichen Deutung im jeweiligen Sachzusammenhang. In den Fallbeispielen von Teil A kann man meiner Meinung nach etwas von dieser "Detektivarbeit" spüren. Da wird bei den "Toren im europäischen Fußball" nicht nur mit Toren gerechnet, um die Schüler zu motivieren, sondern es werden hinter den Daten liegende Sachverhalte aufgedeckt, die den interessierten Leser zu neuen Fragestellungen anregen können wie z.B.: Wenn solche offensiven bzw. defensiven Spielauffassungen für ganze Regionen nachgewiesen werden können, vielleicht läßt sich ähnliches auch bei einem Vergleich einzelner Vereine über einen längeren Zeitraum zeigen – oder ist etwa die Person des Trainers der entscheidende Faktor?

Eindrucksvoll finde ich auch das 5. Fallbeispiel. Es wird nicht nur gezeigt, wie man den Zusammenhang zwischen der durchschnittlichen Jahrestemperatur verschiedener Orte und ihrer Höhe über dem Meeresspiegel gewinnt, sondern auch – unter Zuhilfenahme der höhenbereinigten Temperaturresiduen –, welchen Einfluß die geographische Breite auf die durchschnittliche Jahrestemperatur eines Ortes in der Regel hat. Die Analyse der Ausreißerwerte liefert Hinweise auf klimatische Besonderheiten in Europa, wie den Einfluß des Golfstromes, das Festlandklima und das Mittelmeerklima.

Die Fallbeispiele in Teil A sind so leicht verständlich und ausführlich dargestellt, daß auch der mathematisch nicht besonders vorgebildete Leser sie auf andere Sachsituationen übertragen kann.

Etwas schwierig zu lesen sind die Kapitel 5 und 6 von Teil B. Zum einen fehlen jegliche visuellen Veranschaulichungen der aufgelisteten Daten und zum anderen machen solch "trockene" Erläuterungen, wie z.B. zu Tabelle 23 auf Seite 112, die Information zumindest für Schüler ungenießbar.

Sicherlich wollen die Autoren mit ihrem Buch auch Überzeugungsarbeit für die EDA leisten. Das wird unter anderem auf S. 35 f. deutlich, wo der Median und das arithmetische Mittel einander gegenübergestellt werden. Eine einseitige Bevorzugung des Medians, wie sie hier propagiert wird, ist sicher ebenso bedenklich wie die generelle und unkritische Benutzung des arithmetischen Mittels. Das letztere erscheint mir jedoch als die größere Gefahr, da selbst in diesem Buch (S. 103, Tab. 15) das arithmetische Mittel der Ehedauer bis zur Scheidung berechnet wird, wo der Median treffender wäre. Zu der hier angesprochenen Berechnung des arithmetischen Mittels von Schulnoten nur soviel: "Obwohl das

Mitteln von Schulnoten ... eigentlich ein Unfug" ist (S. 36), gibt es bei der Berechnung der Zeugnisnoten keine brauchbare Alternative. Hierzu möchte ich auf die amüsanten Ausführungen von H. Winter (1985) verweisen.

4. Schlußbemerkung

Das hier vorliegende Buch bietet vor allem wegen der dargestellten Fallbeispiele eine gut lesbare und leicht verständliche Einführung in die Explorative Datenanalyse. Es bietet neben zwei erprobten Unterrichtseinheiten auch für schulische Zwecke aufbereitetes Datenmaterial und eine Fülle von unterrichtspraktischen Anregungen. Ich möchte es daher nicht nur den Mathematiklehrern, sondern auch Kollegen mit anderen Fächern, wie z.B. Geographie, Biologie, Sozialkunde usw. empfehlen.

Literatur:

Biehler, R. (1982): Explorative Datenanalyse – Eine Untersuchung aus der Perspektive einer deskriptiv – empirischen Wissenschaftstheorie. Materialien und Studien Bd. 24, Institut für Didaktik der Mathematik, Bielefeld

Engel, A. (1982) (Hrsg.): Neue Ideen in der Statistik. In: Der Mathematikunterricht 28, Heft 1

Fillbrunn, G. (1981): Beschreibende Statistik (Zeitungsbeiträge als motivierendes und die Kritikfähigkeit förderndes Element in der beschreibenden Statistik, dargestellt an der Häufigkeitsverteilung und ihrer graphischen Darstellung sowie am Mittelwert und Median.) In: Lehrerfortbildung Mathematik, M 9. Landesstelle für Erziehung und Unterricht, Rotebühlstraße 133, Stuttgart 1

Fischer, R./ Malle, G. (1985): Mensch und Mathematik – Eine Einführung in didaktisches Denken und Handeln. Bibliographisches Institut, Mannheim

Tukey, J.W. (1977): Exploratory Data Analysis. Addison – Wesley, Reading

Winter, H. (1985): Die Zensurkonferenz. In: Mathematik lehren, Heft 8, S. 27–32