Die beiden weiteren Beiträge dieses Heftes von *Rolf Biehler* und *Heinz Steinbring* sowie von *Karin Sensenschmidt* und *Peter Weinberg* sind im Rahmen eines Pilotprojektes *(GRAPHDAS)* am Inst. f. Did. d. Mathematik (IDM) der Universität Bielefeld entstanden, in dem in Kooperation mit LehrerInnen mehrere Fortbildungs-Workshops sowie anschließende Unterrichtsversuche zur EDA durchgeführt wurden. Die Erfahrungsberichte von *Karin Sensenschmidt* und *Peter Weinberg* vermitteln einen anschaulichen Eindruck vom Aufbau, Arbeitsweise und Problemen in einer Unterrichtsreihe zur EDA. Die Erfahrungen aus diesen Unterrichtsreihen sind dann in weitere Lehrerfortbildungen und Unterrichtsversuche sowie in die Weiterentwicklung eines Unterrichtskonzeptes zur »Entdeckenden Statistik« eingeflossen. In dem Beitrag von *Rolf Biehler* und *Heinz Steinbring* werden Konzepte, Begründungen und Erfahrungen aus diesen Unterrichtsversuchen dargestellt. Dabei geht es darum, den Stellenwert der neuen Darstellungs- und Analysemittel der EDA auf der Basis praktischer Erfahrungen neu zu bestimmen und die Schwierigkeiten, Chancen und Bedingungen für die *Detektivarbeit mit Daten* im Mathematikunterricht herauszuarbeiten.

In den Projekten, die den Beiträgen dieses Heftes zugrundeliegen, ist im Unterricht selber nur selten mit Rechnern gearbeitet worden, vor allem, weil die notwendige Ausstattung vor Ort nicht zur Verfügung stand. Rechner mit flexibler Statistiksoftware, wie sie weiter oben erwähnt wurde, standen aber im Rahmen der Lehrerfortbildung und teilweise auch zur Unterrichtsvorbereitung zur Verfügung. Die Beiträge dieses Heftes sollen auch deutlich machen, daß man für den Unterricht in EDA nicht notwendigerweise Rechner zur Verfügung haben muß. Andererseits zeigen die praktischen Erfahrungen doch, daß eine Verstärkung *authentischer Detektivarbeit mit Daten* ohne flexible Software im Unterricht nur sehr schwer verwirklicht werden kann.

Biehler, R. (1982): Explorative Datenanalyse – Eine Untersuchung aus der Perspektive einer deskriptiv-empirischen Wissenschaftstheorie. IDM Materialien und Studien Bd. 24. Bielefeld: Universität Bielefeld

Biehler, R. /Rach, W. (1990): Softwaretools zur Statistik und Datenanalyse: Beispiele, Anwendungen und Konzepte aus didaktischer Sicht. Neue Medien im Unterricht. Soest: Soester Verlagskontor

Borovcnik, M./Ossimitz, G. (1987): Materialien zur Beschreibenden Statistik und Explorativen Datenanalyse. Schriftenreihe Didaktik der Mathematik der UBW Klagenfurt 11. Wien/Stuttgart; Hölder-Pichler-Tempsky/Teubner

Gnanadesikan/R./Kettenring, J. R./Siegel, A. F./Tukey, P. A. (1982): Themen aus der Datenanalyse: Begriffe, Methoden, Beispiele und Pädagogik. In: Der Mathematikunterricht, 28(1), 28–56

Heidbrink, H. (1991): Geheimlehre. MACup-Test: Statistikprogramme. In: MACup (3/91), 68–76

Landwehr, J. M./Watkins, A. E. (1986): Exploring Data. Palo Alto: Dale Seymour Publications

Winter, H. (1981): Zur Beschreibenden Statistik in der Sekundarstufe I – Rechtfertigungsgründe und Möglichkeiten zur Integration der Stochastik in den Mathematikunterricht. In: Stochastik im Schulunterricht. Schriftenreihe Didaktik der Mathematik Bd. 3 (S. 279-304). Wien/Stuttgart: Hölder-Pichler-Tempsky/Teubner.

Rolf Biehler

Aus: Der Mathematikunterricht, 37(6), 1991, 5-32

Entdeckende Statistik, Stengel-und-Blätter, Boxplots: Konzepte, Begründungen und Erfahrungen eines Unterrichtsversuches

von Rolf Biehler und Heinz Steinbring

1. Konzeptionelle Elemente der Explorativen Datenanalyse im Unterricht

1.1 Zum Kontext der Unterrichtsversuche

Inwieweit können Ideen, Techniken und Anwendungsbeispiele der Explorativen Datenanalyse (EDA) den Mathematikunterricht bereichern? Wie reagieren SchülerInnen und LehrerInnen auf Ideen der EDA? Wie muß die EDA ggf. umgestaltet werden, damit sie einen pädagogisch sinnvollen Platz im Mathematikcurriculum des allgemeinbildenden Mathematikunterrichts bekommen kann?

Wir untersuchen diese Fragen im GRAPHDAS-Projekt, in dessen Rahmen wir verschiedene Fortbildungsworkshops für LehrerInnen durchführen und zusammen mit ihnen Unterricht planen, beobachten und analysieren. Ein erster Unterrichtsversuch fand 1988 statt. Die Erfahrungberichte von Karin Sensenschmidt und Peter Weinberg in diesem Heft stammen aus dieser Phase. Im Anschluß daran haben wir umfangreicheres Material erarbeitet (vgl. Biehler und Steinbring [1990]), das dann LehrerInnen in einem zweiten Unterrichtsversuch im März 1990 zur Verfügung stand, in dem die EDA in etwa zwölf 8. Klassen unterrichtet wurde. Der folgende Aufsatz bezieht die Erfahrungen aus der zweiten Phase mit ein, die auch dadurch gekennzeichnet war, daß die unmittelbare Betreuung der LehrerInnen nicht so intensiv wie in der ersten Phase möglich war.

In unserem Projekt stand im Vordergrund, wie man den Unterricht in Beschreibender Statistik in der Sekundarstufe I verändern kann. Dies ist naheliegend, da in der EDA Methoden bereitgestellt werden, die den Wahrscheinlichkeitsbegriff nicht voraussetzen. Dabei ist uns bewußt, daß im Grunde der gesamte Aufbau eines Stochastikstranges in den Sekundarstufen noch einmal neu durchdacht werden muß, wenn man im Rahmen der Beschreibenden Statistik die Schwerpunkte in Richtung EDA verschiebt. Wir sind sogar der Auffassung, daß die EDA die Basis dafür verbreitert, Methoden der beurteilenden Statistik angemessener motivieren zu können. An anderer Stelle ist dazu das Konzept der beurteilenden Statistik als kritischer Statistik entwickelt worden; vgl. Biehler [1990].

1.2 Leitbild: Datendetektiv

Das Leitbild, das der amerikanische Statistiker J. W. Tukey [1977] für die EDA geprägt hat, ist das eines Detektivs, welcher in Daten interessante Strukturen und Besonderheiten aufdeckt, gefundenen Hinweisen nachgeht und Hypothesen entwickelt.

Bei der Analyse von Datensätzen sollen Erkenntnisse über *außermathematische Zusammenhänge* entwickelt werden. *Interpretation von Ergebnissen* im Sachkontext ist eine wichtige Komponente. Dabei wird davon ausgegangen, daß die EDA eine bestimmte Perspektive zur Lösung eines Problems beisteuert. Eine Lösung derart, daß dem »Substanzexperten« eine definitive Vorschrift erteilt wird, ist weder möglich noch wird sie angestrebt.

Diese Arbeitsweise macht die EDA interessant (und schwierig) für den Unterricht:

- Detektivarbeit ist spannend, macht Spaß.
- EDA erfordert und ermöglicht offene Arbeitsweisen im Unterricht; der Schüler kann sich als Experte (z. B. für inhaltliche Aspekte der Daten) einbringen.
- EDA ermöglicht es, Anwendungen mit offenem Ende zu behandeln, wo die Mathematik eine Perspektive zur Lösung beisteuert.
- EDA betont interpretative und begriffliche Aspekte der Schulmathematik: Mathe ist nicht nur Rechnen.
- EDA kann und sollte sinnvoll durch Software unterstützt werden, mit der Daten flexibel verarbeitet werden können: ein möglicher Beitrag zur informationstechnischen Bildung.

Die Arbeitsweisen der EDA zu charakterisieren, ist ein schwieriges Problem. Gegenüber üblichen Arbeitsweisen in der Statistik

- wird interaktiv und experimentell mit den Daten gearbeitet;
- werden i. d. R. mehrere Analysemethoden und Darstellungsweisen verwendet;
- werden häufig Daten vor einer Analyse erst noch geeignet transformiert und ausgewählt;
- spielen graphische Darstellungen für das Erkennen von Strukturen und Besonderheiten eine wesentliche Rolle.

In der Literatur sind verschiedene Beispiele zur EDA ausgearbeitet worden, die exemplarisch nicht nur einzelne Techniken, sondern explorative Arbeitsweisen an »Fallstudien« veranschaulichen wollen. In der deutschsprachigen Literatur findet man z. B.

- Selbstmordraten in verschiedenen Ländern und Altersgruppen, vgl. Biehler [1982];
- Niederschläge in verschiedenen Kontinenten. Seehöhe und Durchschnittstemperatur u. a., vgl. Borovcnik und Ossimitz [1987];
- Unfallstatistiken in der Bundesrepublik, vgl. Biehler und Rach [1990];
- Temperaturen in Bamberg und Norderney, Geburten in London, vgl. Biehler [1990];
- Frühlingswetter, vgl. Nordmeier [1989].

Eine detaillierte Herausarbeitung von Arbeitsweisen der EDA und ihre Entwicklung im Kontext der Statistik findet man bei Biehler [1982]. Borovcnik, et al. [1987] fassen Kernpunkte dieser Arbeitsweisen für die Lehrerbildung zusammen, allerdings mit teilweise unterschiedlicher Akzentuierung.

Nach unseren Erfahrungen in der Lehrerfortbildung und in der Schule ist es hilfreich, die Tätigkeiten eines Datendetektivs zunächst noch allgemeiner, als explorative Tätigkeiten, zu charakterisieren, um deutlich zu machen, daß diese Aspekte zum Gegenstand des Unterrichts selber gehören sollen und nicht einfach zum methodisch-kommunikativen Beiwerk für das Lernen von Verfahren. Zur Tätigkeit gehören die Komponenten:

- Fragestellungen entwickeln, präzisieren;
- Hypothesen aufwerfen, kritisieren;
- Auswahl geeigneter Methoden: Ausprobieren, Vergleichen mit anderen Methoden, Angemessenheit einschätzen;
- Modifizieren von Graphiken;
- Abwandlung und Neuerfindung von Mitteln;
- Interpretation von graphischen und numerischen Ergebnissen;
- Beschaffung von Hintergrundinformationen und weiteren Daten;
- Zusammenfassung und Kommunikation der Ergebnisse (schriftlich, mündlich).

Für den Mathematikunterricht enthält dies viele neue Elemente; wir werden in Abschnitt 3 genauer darauf zu sprechen kommen.

1.3 Das Handswerkszeug: die Mittel der Datenanalyse

Eine solche komplexe Tätigkeit auszuüben, setzt das flexible Beherrschen von Handwerkszeug voraus. John Tukey [1977] hat für die elementare Datenanalyse einige neue Darstellungsmittel entwickelt, die besser als die herkömmlichen einen flexiblen Umgang mit

Daten unterstützen sollen, vor allem dann, wenn man mit Papier und Bleistift arbeitet. Die umfassendste einführende Darstellung explorativer Techniken und Graphiken findet man in der deutschsprachigen Literatur bei Polasek [1988]. Zu dem Handwerkszeug gehören folgende Gruppen:

- ein Repertoire elementarer graphischer Darstellungen;
- ein Repertoire von numerischen statistischen Verfahren;
- Techniken und Begriffe zum Organisieren von Daten (Datentypen, Transformationen);
- (offene) Begriffe zur Analyse von Daten, z. B. Streuung, mittlerer Wert, Ausreißer, Abhängigkeiten.

Unser Ziel war zunächst, zu untersuchen, wieweit man mit einem *Minimalprogramm* an Begriffen und Darstellungen kommt, wobei vor allem die Möglichkeiten der neuen Darstellungen der EDA untersucht werden sollten. Es sollte auch geprüft werden, wie sich die neuen zu den bisherigen Mitteln verhalten. Entscheidend ist, daß der Gebrauch dieser Mittel so gelernt wird, daß sie dann flexibel und kritisch im Rahmen von Detektivarbeit mit Daten eingesetzt werden können. Wir sind von folgenden Elementen ausgegangen:

Tab. 1: Übersicht über elementare Darstellungen, Begriffe und Aufgabenstellungen der EDA

	Darstellungen	*Aufgabenstellungen* Begriffe
1	Tabelle Stengel-und-Blätter-Schaubild Säulendiagramm	*Darstellungen von Daten* Verteilung Streuung Symmetrie/Schiefe Ausreißer, Zerfallen in Gruppen
2	5-Zahlen-Zusammenfassung Boxplot	*Zusammenfassung und Entfalten von Daten* Median, Quartile, Extremwerte Streuungsmaße: Quartilabstand und Spannweite *Vergleich von Datensätzen*
3	Streudiagramm/Liniendiagramm – mehrere Abhängigkeiten in einem Diagramm – eingezeichnete Zusammenfassungen (Geraden, Streckenzüge)	*Untersuchung von Abhängigkeiten* Datentransformation Daten = Modell + Residuen Gerade nach Augenmaß 3-Gruppen-Gerade Medianlinie

Im Rahmen unserer Unterrichtsversuche ist der Block 3 dann nur ansatzweise ausgearbeitet und realisiert worden. Es war aber zumindestens wichtig, über einfache Darstellungen für zeitliche Abhängigkeiten zu verfügen.

Für die Vermittlung des Handwerkszeugs kann nun nicht einfach das Leitbild »Datendetektiv« zugrundegelegt werden. SchülerInnen und LehrerInnen muß bewußt werden, daß das Erlernen des Handwerkzeugs eine eigenständige Perspektive impliziert. Man kann neue Darstellungen nicht von *einzelnen* Anwendungen heraus begründen, sondern nur von der Gesamtperspektive her. Auch ist es nötig, neue Darstellungsmittel und Verfahren für sich zu untersuchen. Dabei dominiert die Perspektive »Was kann man mit dem Boxplot alles machen?« und nicht »Wie kann ich die vorliegenden Daten möglichst gut analysieren?«. Andererseits lernt man den Gebrauch des Handwerkszeugs nur durch den Gebrauch. Hierin liegt ein schwieriges Dilemma, das in ähnlicher Weise allen anwendungsbezogenen Mathematikunterricht charakterisiert und das in den unterrichtlichen Realisierungen immer wieder aufscheint.

1.4 Wechselbeziehung von Daten und Mitteln als Problem bei der Strukturierung von Unterrichtseinheiten

Zur Untersuchung eines bestimmten, vorgegebenen Datensatzes muß man die geeigneten Analysewerkzeuge benutzen, dazu paßt nicht jedes beliebige statistische Arbeitsmittel. Was die geeigneten Werkzeuge in einem solchen Falle sind, ist nicht von vornherein entschieden; erst im Prozeß der versuchsweisen und erprobenden Anwendung wird die Wirksamkeit der richtigen Analysemittel bei gleichzeitiger Transformation und Anpassung der vorliegenden Daten deutlich.

Dieses Problem bei der Auswahl geeigneter Analyse- und Darstellungsmittel, mit dem der statistische Experte bei seiner Arbeit mit Daten konfrontiert ist, erhält im Unterricht eine zusätzliche Schwierigkeit. Wenn das neue Wissen erst erworben werden muß, dann können die Schüler natürlich noch nicht auf eigene Erfahrungen mit statistischen Diagrammen zurückgreifen und zudem haben sie noch kein »Gefühl für die Daten«, d. h. sie können den Daten nicht »ansehen«, welche Mittel zu ihrer Analyse wohl am ehesten geeignet sein werden. Wie läßt sich dann ein Unterrichtsablauf gestalten, der selbst den explorativen Geist der EDA im Lernprozeß ansatzweise zur Geltung bringen möchte und in dem die elementaren Darstellungsmittel und Begriffe der EDA nicht einfach schematisch vom Lehrer vorgegeben und dann von den Schülern auf Datensätze angewendet werden?

Weder sollte der Einstieg in den Unterricht zur EDA eine bloße Vermittlung von Rezepten für die statistische Arbeit sein, noch kann der Beginn statistischen Arbeitens völlig offen gestaltet werden. Es muß so etwas wie eine »Mischung« zwischen beiden Vorgehensweisen angestrebt werden, und zwar »global« im Aufbau einer Unterrichtseinheit als auch »lokal« bei der Bearbeitung von Aufgaben.

Nach den Unterrichtserfahrungen scheint uns die folgende Struktur einer Unterrichtseinheit sinnvoll zu sein:

Orientierungsphase:	Vermittlung einer Idee davon, was EDA bedeutet;
Grundphase:	Einführung der elementaren Formen der Darstellungsmittel, Begriffe und Verfahren;
Aufbauphase:	Arbeit mit und an verschiedenen Datensätzen;
(kurze) Abschlußphase:	Vorbereitung des Tests der Unterrichtsreihe.

Als lokale Lösung haben wir in unserem Arbeitsmaterial Aufgaben entwickelt, in denen durch *nach Offenheit gestufte Teilaufgaben* verschiedene Aspekte abgedeckt werden:

- Anfertigung einer Darstellung, deren Typ vorgegeben wird;
- Entnehmen von einfachen Informationen daraus;
- Vorschlag weiterer Auswertungen – ggf. Vorgabe von Graphiken zur Interpretation;
- Aufwerfen von Interpretationsfragen zur gezielten Fokussierung auf die Daten;
- Konfrontation mit vorgegebenen Interpretationen;
- Öffnen für allgemeinere Fragestellungen, in denen man sich auch von der vorgeschlagenen Analyse lösen kann.

1.5 Kriterien bei der Auswahl von Datensätzen und Beispielen

Die Auswahl geeigneter Datensätze ist ein schwieriges Problem. Die folgenden Dimensionen sollten bei der Auswahl berücksichtigt werden (siehe auch Konold [1990]):

1. inhaltlich (von der Sache her) interessant und relevant (zu machen);
2. Bedeutung/Definition der Variablen ist verständlich zu machen;
3. Hintergrundinformationen für Interpretation verfügbar;
4. enthält mehrere Variablen (minimaler Reichtum);
5. mittlerer Umfang und Aggregationsgrad;
6. Analyse bringt interessante Muster und Besonderheiten zum Vorschein;
7. zeigt Nützlichkeit und Grenzen des EDA-Werkzeugkastens;
8. erlaubt, »Modellcharakter« von Daten zu thematisieren.

Unter Modellcharakter verstehen wir dabei, daß ein Dateneinsatz immer nur einen (quantifizierten) Ausschnitt aus der Wirklichkeit darstellt, selbst wenn er viele Variablen enthält. Wir haben im Unterricht mit Daten zu Unfallstatistiken gearbeitet, wo dies Problem besonders kraß ist. Schreckliche Einzelschicksale mit vielen Begleitumständen reduzieren sich auf Zahlen. Andererseits macht die Datenanalyse es möglich, eine Gesamtperspektive einzunehmen und globale Einflußfaktoren identifizieren zu helfen. Diese Aspekte müssen bewußt gemacht werden.

2. Zur Struktur von graphischen Darstellungen und Begriffen der EDA im Spiegel unterrichtspraktischer Erfahrungen

2.1 Umgang mit Graphiken

Arbeitsweisen und Funktionen von Graphiken haben wir schlagwortartig in der Fortbildung charakterisiert durch

exploratives und kommunikatives Mittel
Graphiken können genutzt werden, um anderen etwas mitzuteilen oder um selber etwas neu in Daten zu entdecken.

Wertespeicher und relationales Diagramm
Graphiken können zum Darstellen und Ablesen einzelner Werte genutzt werden. Wichtiger ist aber, sie zur Mitteilung oder Erkenntnis von Beziehungen/Strukturen zu nutzen. Hierbei wird die menschliche Wahrnehmungsfähigkeit ausgenutzt; es besteht aber auch die Gefahr »optischer Täuschungen«. Auf welche Relationen es in der Statistik ankommt, muß erst gelernt werden.

Modellcharakter[1]
Jede Graphik stellt eine Perspektive auf die Daten dar. Hiermit ist die Gefahr der Verfälschung, aber auch die Chance der Konzentration auf bestimmte Aspekte gegeben. Welche Graphiken man für einen Datensatz wählt, hängt davon ab, was man mitteilen möchte bzw. was man untersuchen möchte.

Variation und Vielfalt
Graphiken zu verändern, eine Vielfalt von Graphiken neben- oder hintereinander zu benutzen, ist eine wichtige Arbeitsweise der EDA, die wichtige Einsichten in die Daten bringen kann. Allerdings besteht die Gefahr, in einer Vielzahl von Darstellungen zu »ertrinken«. Die Koordination mehrerer Graphiken erhöht die kognitiven Anforderungen.

2.2 Graphische und semigraphische Darstellungen

Wir wollen zunächst eine Übersicht über die elementaren graphischen und semigraphischen Darstellungen geben, um zu verdeutlichen, wie sich Boxplot und Stengel-und-Blätter-Schaubild einordnen lassen. Die Darstellungsmöglichkeiten hängen davon ab, wie viele Variable man in einer Graphik darstellen will und von welchem Typ die Variablen sind. Wir wollen als Hintergrund einen Schülerdatensatz benutzen, bei dem verschiedene Variablen in einer Tabelle vorliegen, u. a.

Variable	*Typ*
Schülername	Namensvariable (identifizierend)
Schulwegzeit (Min.)	quantitativ (numerisch)
benutztes Verkehrsmittel	diskret (qualitativ, kategorial)

Tab. 2: Schulweg in Minuten, Klasse 8g — Datenurliste

Stefanie	45	Bus	Kirsten	20	Bus
Bettina	28	Bus	Svenja	40	Bus
Köray	30	Bahn	Cornelia	57	Bus/Bahn
Cemil	31	Bahn	Hendrik	16	Bus
Snezana	50	Bahn	Luciano	40	Bus
Daniel	5	Auto	Tamara	12	Rad
Viviane	8	Rad	Adrian	40	Skateboard
Marietta	10	Rad	Simone	—	
Tobias	5	zu Fuß	Reinhard	38	Bahn/Bus
Marvin	18	zu Fuß	Michael	5	zu Fuß
Ergin	25	Bus	Ullrich	38	Bus
Christian	6	zu Fuß	Almira	17	Bahn
Krezimir	15	Rad	Daniela	13	Bahn
Karin	19	Auto	Melanie	7	zu Fuß
Songül	25	Bus	Annika	45	Bus
Dagmar	—		Sabine	—	
Angelika	30	Bus			

(Quelle der Daten: Martin-Niemöller-Gesamtschule Bielefeld, Klasse 8g, Schuljahr 1988/89)

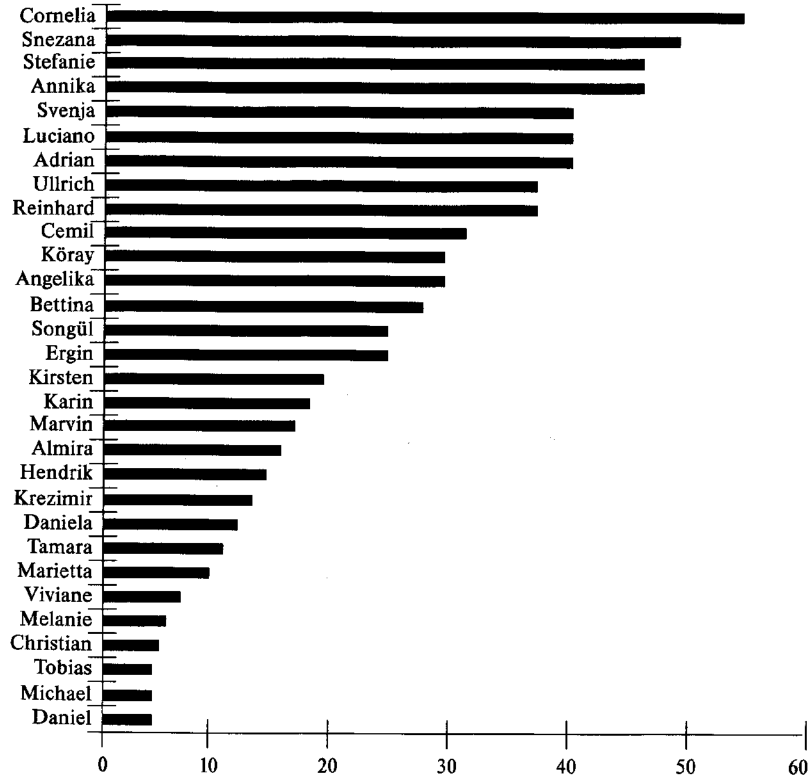
Tab. 3: Klassifikation elementarer Graphik

A. 1 numerische Variable	1. Strichliste 2. Stengel-und-Blätter 3. Punktdiagramm 4. Histogramm mit absoluter 5. — mit relativer Häufigkeit 6. Boxplot 7. Mittelwert-plus-Streuung-Diagramm
B. 1 kategoriale Variable	8. Strichliste 9. Säulendiagramm (Histogramm mit natürlichen natürlichen Klassen) — mit absoluter Häufigkeit 10. — mit relativer Häufigkeit 11. Tortendiagramm
C. 1 Name mit 1 numerischen Variablen (Schülername u. Schulwegzeit)	12. Balken-/Säulendiagramm 13. Stengel-und-Blätter
D. 2 numerische Variable: *Abhängigkeit* (Körpergewicht in Abhängigkeit vom Lebensalter; Körpergewicht in Abhängigkeit von Körpergröße)	14. Streudiagramm 15. Liniendiagramm 16. Balken-/Säulendiagramm 17. Streudiagramm mit Randverteilungen 18. zeitl. gestrecktes Stengel-und-Blätter
E. mehrere numerische Variable: *Verteilungsvergleich* (Körpergröße von Mädchen und Jungen)	19. *Multiple Diagramme:* Für jede Variable wird ein Diagramm aus A. erstellt. Diese werden dann nebeneinander oder übereinander gelegt.
F. 1 numerische und 1 kategoriale Variable (Schulwegzeit und Verkehrsmittel)	20. *Multiple Diagramme:* Für jede Merkmalsausprägung der kategorialen Variable wird ein Diagramm aus A. erstellt. Diese werden dann nebeneinander oder übereinander gelegt. 21. Stengel-und-Blätter
G. 2 kategoriale Variable: *Abhängigkeit* (Geschlecht und benutztes Verkehrsmittel)	22. Mehrfeldertafel
H. mehrere kategoriale Variable: *Verteilungsvergleich* (benutztes Verkehrsmittel in Stadt- und Landschulen)	23. *Multiple Diagramme:* Für jede Variable wird ein Diagramm aus B. erstellt. Diese werden dann nebeneinander oder übereinander gelagert.

Die Fälle E und F führen mehr oder weniger auf dieselben Graphiktypen, man kann sie auch als die entsprechenden *multiplen* Diagramme bezeichnen (multiple Strichliste, Stengel-und-Blätter usw.). Zum Teil gibt es Sonderformen, z. B. wenn zwei Balkendiagramme Rücken-an-Rücken dargestellt sind (z. B. bei Bevölkerungspyramiden). Zur Erläuterung sollen für den Schülerdatensatz einige Graphiken dargestellt werden.

In Abbildung 1 wird ein Balkendiagramm (C) dargestellt, dies ist eine auch in Massenmedien geläufige Darstellung. Die Kinder sind hier nicht in alphabetischer Reihenfolge, sondern nach der Dauer ihrer Schulwegzeit aufgeführt; dies erleichtert eine Reihe von Analysen.

Abb. 1: Balkendiagramm — Schulwegzeit in Minuten



Für SchülerInnen ist dies zunächst eine plausible naheliegende Veranschaulichung der Daten auf der Basis einer Tabelle, die alle Informationen enthält. Viele Daten in Massenmedien sind von dem in Abbildung 1 dargestellten Typ, auch wenn sie durch schmückendes Beiwerk und alphabetische Ordnung oft zu nicht mehr dienen können, als einzelne Werte abzulesen.

Auf Darstellungen der *Verteilung* der Daten unter zeitweiliger Abstraktion von den zugehörigen Personen überzugehen, d. h. auf Darstellungen der Kategorie A, ist ein Standpunktwechsel, der bewußt begründet und vollzogen werden muß. Darstellungen der Kategorie A sind SchülerInnen i. a. nicht vertraut, kommen in Massenmedien auch selten vor.

Abbildung 2 zeigt ein Stengel-und-Blätter-Schaubild für die Schülerdaten, das aus der Tabelle entsteht, indem man die Ziffern der Zahlen auf Stengel (Zehner) und Blätter (Einer) aufteilt.

Hier stellt die Zeile 1|0 2 3 die Zahlen 10, 12 und 13 dar. Durch die Möglichkeit, schnell die Personen wieder den im Stengel-und-Blätter-Schaubild dargestellten Zahlen zuordnen zu können oder einige von ihnen zusätzlich in dem Diagramm zu notieren (hier: Cornelia), stellt das Stengel-und-Blätter-Schaubild gewissermaßen einen Übergang zum Histogramm dar, in dem radikal die Verbindung zu den zugehörigen Personen unterbrochen ist. Aus diesem Grund taucht das Stengel-und-Blätter-Schaubild in Tabelle 1 auch in A *und* C auf.

Wenn wir das benutzte Verkehrsmittel als kategoriale Variable hinzunehmen, so ergeben sich z. B. die folgenden Darstellungsmöglichkeiten der Kategorie F:

Abb. 2: Schulwegzeiten
Geordnetes Stengel-und-Blätter-Schaubild
2-fache Unterteilung
Einheit: Minuten

0		
*	5 5 5 6 7 8	
1	0 2 3	
*	5 6 7 8 9	
2	0	
*	5 5 8	
3	0 0 1	
*	8 8	
4	0 0 0	
*	5 5	
5	0	
*	7	Cornelia
6		

Abb. 3: Schulwegzeiten und Stengel-und-Blätter-Schaubild, unterschieden nach Schülern, die motorisiert (A) oder nicht motorisiert (N) zur Schule kommen.
Einheit: Minuten

0	
*	N N N N N A
1	N N A
*	N N A A A
2	A
*	A A A
3	A A A
*	A A
4	N A A
*	A A
5	A
*	A
6	

Abb. 4: Schulwegzeiten nach Verkehrsmitteln
Multiples Stengel-und-Blätter-Schaubild
Einheit: Minuten

	Bahn/ Bus	zu Fuß	Rad	Auto	Skate-board
0					
*		5 5 6 7	8	5	
1	3		0 2		
*	6 7	8	5	9	
2	0				
*	5 5 8				
3	0 0 1				
*	8 8				
4	0 0				0
*	5 5				
5	0				
*	7				

Als weitere Varianten kommen noch das Rücken-an-Rücken-Stengel-und-Blätter-Schaubild und das zeitlich gestreckte Stengel-und-Blätter-Schaubild hinzu. Auch kann man durch feinere und gröbere Unterteilung die Darstellung in Abbildung 2 strecken und stauchen (vgl. den Artikel von Sensenschmidt und Weinberg in diesem Heft und bei Biehler und Steinbring [1990]). Dieser Variantenreichtum in Verbindung damit, daß auch mit Papier und Bleistift zwischen diesen Varianten relativ leicht gewechselt werden kann, ist einzigartig. Das Balkendiagramm aus Abbildung 1, aber auch ein Histogramm ist schwerer zu erstellen und ohne Computer kaum in zeitlich angemessener Weise zu variieren.

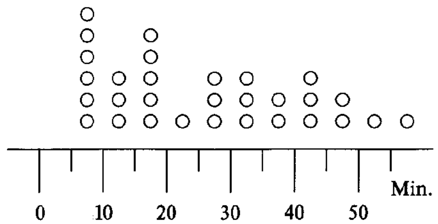
Das Stengel-und-Blätter-Schaubild als semigraphische Darstellung kommt einem Histogramm am nächsten, enthält aber noch mehr Informationen über die einzelnen Daten, so daß die Beziehung zwischen dem Zahlenwert und dem zugehörigen Merkmalsträger leichter erkennbar ist. Der entscheidende Vorteil des Stengel-und-Blätter-Schaubilds liegt darin, daß seine Herstellung und Veränderung mit Papier-und-Bleistift wesentlich effektiver ist als bei den Histogrammen:

- Daten können direkt in ein Stengel-und-Blätter-Schaubild hinein erhoben/notiert werden;
- die Klasseneinteilung ergibt sich pragmatisch und braucht zunächst nicht eigens thematisiert werden, Weiterentwicklung zum Histogramm möglich;
- mit dem Stengel-und-Blätter-Schaubild ist eine Sortierung der Daten erreicht;
- es kann leicht gestreckt und gestaucht werden;
- ein Stengel-und-Blätter-Schaubild kann leicht nach Kategorien aufgespalten werden;
- ein multiples oder Rücken-an-Rücken-Diagramm kann erstellt werden;
- Median und Quartile können durch Abzählen im Stengel-und-Blätter-Schaubild bestimmt und veranschaulicht werden

Diese Eigenschaften machen das Stengel-und-Blätter-Schaubild zu einem idealen Kandidaten für einen Einstieg in die Datenanalyse, wo zunächst das Vertrautwerden mit Daten und einfachen flexiblen Analysetechniken mit Papier-und-Bleistift im Vordergrund stehen soll.

Ist diese Darstellung nun auch noch dann hilfreich, wenn man – wie in der statistischen Praxis – über flexible Computersysteme verfügt, mit denen man auch andere Darstellungen schnell variieren kann? Übliche Histogramme haben den prinzipiellen Nachteil, daß durch die Säulendarstellung die einzelnen Daten nicht mehr sichtbar sind. Eine Mischform wäre ein »Histogramm mit Einzeldatendarstellung«.

Abb. 5: Histogramm mit Einzeldatendarstellung



Anm.: Die Klasseneinteilung wurde analog zu der aus Abb. 2 gewählt.

Steht dies so zur Verfügung, daß man einzelne Punkte per Mausklick identifizieren und dauerhaft beschriften kann und daß man die Klasseneinteilung variieren kann, so treffen praktisch alle Unterscheidungen, die beispielsweise Borovcnik [1990, S. 78] herausstellt, kaum mehr zu. Das Stengel-und-Blätter-Schaubild hat allerdings auch für den Profi spezielle diagnostische Fähigkeiten:

1. Numerische Muster, z. B. die Bevorzugung bestimmter Ziffern oder Zahlenwerte sind erkennbar.
2. Die Verteilung innerhalb eines Blattes ist zugänglich, z. B. ist prinzipiell feststellbar, ob dort Lücken vorliegen. Da sich solche Lücken innerhalb eines Blattes aber nicht optisch aufdrängen, wird diese Möglichkeit häufig nicht wirksam.
3. Wenn es nicht nur um Häufigkeiten geht, sondern auch um die numerischen Werte einzelner Zahlen, dann ist ein Stengel-und-Blätter-Schaubild eine gute Mischung zwischen graphischer und numerischer Information. Dies gilt besonders für kleinere Datensätze.

Grenzen des Stengel-und-Blätter-Schaubilds:
Es versagt dort, wo nicht Rohdaten vorliegen, sondern nur noch Häufigkeitsdaten oder umfangreiche Datensätze. Dies trifft auf viele Sozialdaten zu, wo für eine Bevölkerung bestimmte Häufigkeitsverteilungen erhoben wurden, z. B. Einkommensverteilungen. Sobald es auf die *Form* der Verteilung und ihre Interpretation ankommt, d. h. auf die Verteilung der relativen Häufigkeiten unter Abstraktion von den absoluten Anzahlen, ist das übliche Histogramm die einzig brauchbare Darstellung (vgl. hierzu Biehler [1982, S. 35 ff.]). Das Curriculum von Rosebery [1988] enthält beispielsweise zur Entwicklung des Verteilungsbegriffs

Aufgaben (»distribution teasers«), wo Schüler aufgefordert werden, die qualitative Form der Verteilung von Zufallsgrößen zu skizzieren, z. B.

- die Verteilung der Entfernung, die Leute täglich zum Arbeitsplatz fahren;
- die Altersverteilung der Bevölkerung in einem Staat;
- die Verteilung der Küstenlänge bei den amerikanischen Bundesstaaten.

Über solche Aktivitäten kann exemplarisch klar werden, wie eine Verteilungs*form* in Verbindung zu inhaltlichen Charakteristika steht.

Abb. 6: Schulwegzeiten in Minuten
Ermittlung der Viertelwerte,
5-Zahlenzusammenfassung

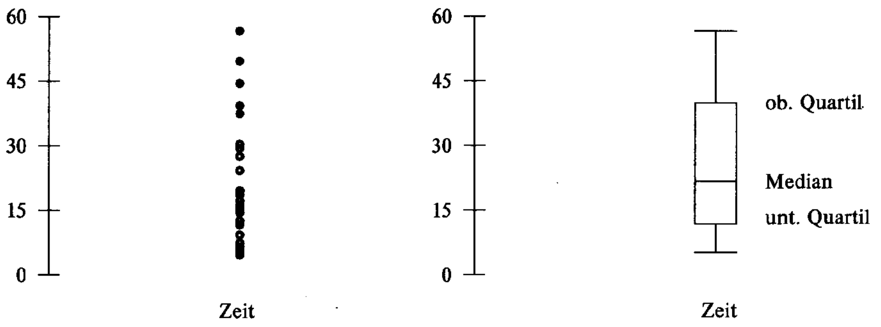
Anhand des Stengel-und-Blätter-Schaubildes lassen sich auch auf einer handlungsmäßig anschaulichen Grundlage die Kennzahlen Median sowie die Quartile einführen. In einer 5-Zahlen-Zusammenfassung werden die Ergebnisse notiert:

```
0 |
* | 5 5 5 6 7 8
1 | 0 [2] 3      <------  unteres Quartil/
                          Viertelwert
* | 5 6 7 8 9
2 | 0 1          <------  Median
* | 5 5 8
3 | 0 0 1
* | 8 [8]        <------  oberes Quartil/
                          Viertelwert
4 | 0 0 0
* | 5 5
5 | 0
* | 7
6 |
```

Median	22,5	
Viertelwerte	12	38
Extremwerte	5	57

Die Werte der 5-Zahlenzusammenfassung können dann in Form eines Boxplots dargestellt werden. Für das Verständnis dieser Darstellungsform ist es nützlich, zum Vergleich das Punktdiagramm heranzuziehen: Beide Darstellungsweisen plotten die Punkte gewissermaßen entlang einer Zahlenachse, wobei der Boxplot eine Auswahl und Zusammenfassung der Werte darstellt.

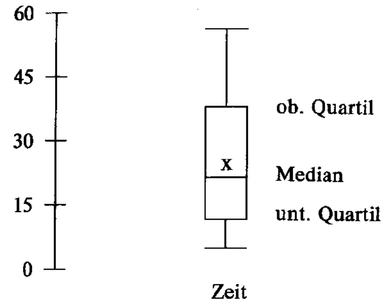
Abb. 7: Schulwegzeiten in Minuten
Punktdiagramm und Boxplot



Die entscheidende Grundidee beim Boxplot besteht darin, eine numerische Zusammenfassung graphisch darzustellen, um den Vergleich mehrerer Verteilungen zu erleichtern. Zur Begründung dieser Darstellungsform, die einer manchmal in der Praxis zu beobachtenden Ablehnung oder aber dogmatischen Übernahme vorbeugen könnte, ist es hilfreich, diese Darstellung im Kontext anderer »Mittelwert-plus-Streuung«-Diagramme zu sehen und ihre spezifischen Vorteile zu erkennen (vgl. hierzu Biehler [1990]).

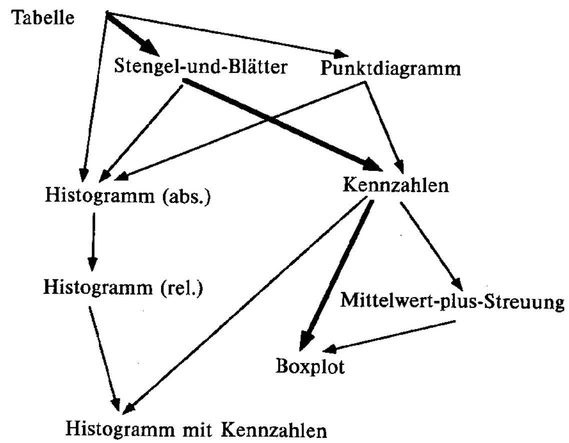
Abb. 8: Schulwegzeiten in Minuten
Boxplot mit arithmetischem Mittel

In der Schule wurde auch häufig das arithmetische Mittel parallel zu den anderen Kennzahlen mit entwickelt. Dieser Mittelwert kann in die Darstellung mit aufgenommen werden, z. B. durch ein Kreuzchen im Boxplot. Diese Darstellung erlaubt es dann auch, Beziehungen zwischen arithmetischem Mittel und Median experimentell zu untersuchen.



Im folgenden Diagramm wird eine netzförmige Übersicht über die elementare Graphik gegeben.

Abb. 9: Netz elementarer Graphik für eine (numerische) Variable



Die Graphiken in Abbildung 9 beziehen sich alle auf die Kategorie (A) aus Tabelle 3. Sie bleiben insofern im Rahmen der traditionellen Inhalte der beschreibenden Statistik, als hierbei die *Verteilung einer Variablen* im Vordergrund steht. Geht man von einem Datensatz mit mehreren Variablen aus, so ergibt sich die folgende Netzstruktur (Abb. 10, s. S. 16).

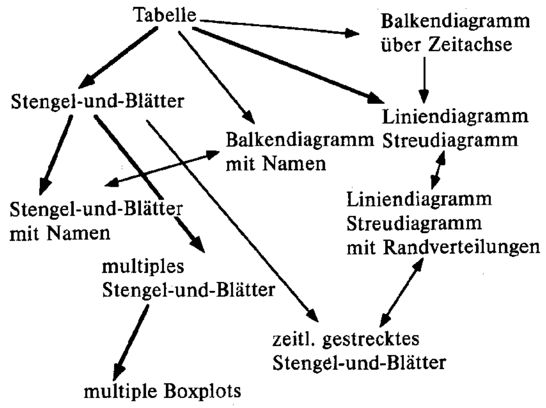
Beim Streu- und Liniendiagramm mit Randverteilungen können die Randverteilungen durch Punktdiagramme, Boxplots o. ä. dargestellt werden. Zusammen mit dem zeitlich gestreckten Stengel-und-Blätter-Schaubild stellen diese Diagramme Übergänge zwischen ein- und zweidimensionalen Darstellungen dar.

Die fett gesetzten Pfeile markieren einen Minimalweg, der von uns ins Auge gefaßt wurde für eine Einführung, in der weitgehend mit Papier und Bleistift gearbeitet wird.

Nach unserer Erfahrung ist es für die Lehrerbildung essentiell wichtig, aber auch für den Unterricht lohnend, möglichst über den Minimalweg hinauszugehen, die obigen Netze vollständiger zu durchlaufen und ggf. auch noch Darstellungen für diskrete Variable hinzuzunehmen (Tabelle 3, B, G, H). Dies wird i. d. R. nicht dadurch geschehen können, daß SchülerInnen in all diesen Darstellungen Kompetenz erwerben. Ein Überblickswissen über die Vor- und Nachteile der verschiedenen Graphiken ist aber nützlich. Dies könnte z. B. an einem Datensatz durch vorgegebene Graphiken erarbeitet werden.

Hiermit kann man zwei Problematiken begegnen, die wir im Unterricht beobachtet haben. Zum einen sind den SchülerInnen die Beziehungen zwischen verschiedenen Graphiken keineswegs unmittelbar klar, sondern müssen eigens zum Thema gemacht werden. Zum anderen bringen SchülerInnen und LehrerInnen in einen offen gestalteten Unterricht eigene Graphikformate ein, und die neuen Darstellungen müssen mit ihren spezifischen Leistungen und Schwächen im Kontext der alten deutlich werden.

Abb. 10: Netz elementarer Graphik für ein und zwei Variablen



2.3 Begriffe und Aufgabenstellungen der Datenanalyse

Im Unterricht hat sich gezeigt, daß eine begriffliche Orientierung besonders wichtig ist. Es kommt nicht nur darauf an, bestimmte Kennzahlen ausrechnen zu können, sondern das Anwenden von Begriffen zur Beschreibung und Interpretation einer Darstellung im Rahmen der unterrichtlichen Kommunikation wird ebenso wichtig. Ferner ist es wichtig, neben quantitativen Begriffen parallel mit offenen qualitativen Begriffen zu arbeiten, anstatt diese nach »genetischer« Einführung zu vergessen. Die Tabelle 4 gibt eine Übersicht.

Tab. 4: Qualitative und quantitative Begriffe der Datenanalyse

qualitativ und offen	*quantitativ*
mittlerer Wert	Median, arithmetisches Mittel
Streuung	Spannweite, Quartilabstand
Kennzahlen	Quartile, Minimum, Maximum (Rangplatz, Prozentwert)
Verteilung	
Dichte	
Form der Verteilung	
Ausreißer	
Zerfall in Gruppen	

Im Sinne eines begrifflichen Minimalprogramms wurden bei den Unterrichtsvorschlägen und -realisierungen zur EDA die Begriffe Median, Quartile, Minimum, Maximum sowie – darauf aufbauend – die Streuungsmaße Spannweite und Quartilabstand verwendet, da diese relativ einfach (durch Zählen) zu bestimmen sind, und vor allem der Quartilabstand ein einfacher zu verstehendes quantitatives Maß für die Streuung ist, als z. B. die Standardabweichung.

Für eine flexible Praxis der EDA ist es aber wichtig, daß die quantitativ gefaßten Begriffe als Präzisierungen von allgemeineren, qualitativen Begriffen zu verstehen sind, vgl. Mosteller und Tukey [1977, S. 17 ff.]. Bei der Anwendung muß ggf. geprüft werden, ob die quantitativen Begriffe vernünftig sind oder ob alternative Präzisierungen gewählt oder neu definiert werden müssen. In der Präzisierung steckt eine gewisse Willkür, derer man sich bewußt sein muß. Die Wahl der 25 %- und 75 %-Werte für eine Standardzusammenfassung im Boxplot (statt z. B. 40 %, 60 % oder 10 %, 90 %) ist pragmatisch gerechtfertigt, kann aber zunächst nicht weiter begründet werden. In der Anwendung kann es sogar sinnvoll sein, davon abzuweichen.

Neben der »Einfachheit« spricht für die »neuen« Begriffe Median und Quartilabstand, das sie nicht so stark von Ausreißern beeinflußt sind, wie die traditionellen Maße (Resistenz, Robustheit). Die traditionellen Streuungsmaße basieren aber auf dem an sich schlüssigen Konzept der Messung einer »mittleren Abweichung vom Mittelwert«, das vor allem bei symmetrischen Verteilungen sehr nützlich ist und weiterhin in vielen Situationen seine Anwendung und Bedeutung findet. Das arithmetische Mittel $\bar{x}$ hat eine Reihe von Vorteilen, die durch die Beziehung

$$n \cdot \bar{x} = \sum_{i=1}^{n} x_i$$

vermittelt wird (Ausgleichseigenschaft, Schwerpunkteigenschaft, Gesamtsumme ist aus $\bar{x}$ rekonstruierbar). Aus diesen Gründen wäre ein Verzicht auf die traditionellen Begriffe der Statistik nicht unproblematisch.

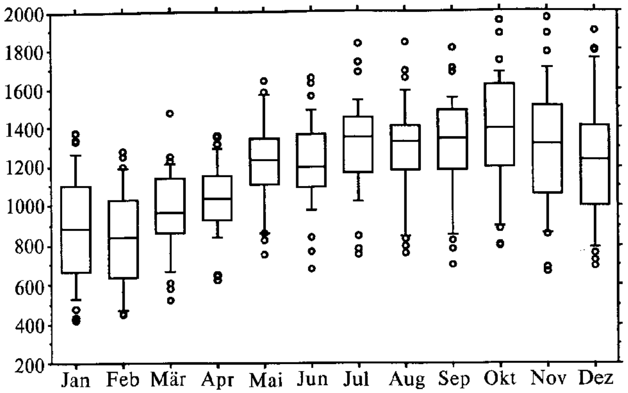
In der obigen Tabelle 4 sind eine Reihe von weiteren qualitativen Begriffen aufgeführt, die für die Analyse wichtig sind. Der Begriff des *Ausreißers* und des *Zerfallen in Gruppen* ist bei der Analyse von Graphiken wichtig: Hier liegen oft zentrale Hinweise auf inhaltliche Aspekte der Daten vor. Es war intendiert, daß diese Konzepte im Rahmen einer Einführung nicht weiter präzisiert werden. Dies wäre auch im Rahmen der Beschreibenden Statistik gar nicht sinnvoll möglich. Vielmehr geht es darum, diese Begriffe bezogen auf die verwendeten Darstellungsmittel zur Analyse qualitativ zu verwenden. In dem Boxplot von Abbildung 6 etwa stellt man fest, daß die Daten zwischen unterem Quartil und Median »dichter« liegen als zwischen Median und oberem Quartil. Eine unterschiedliche »Dichte« der Daten läßt sich anschaulich auch im Punktdiagramm und im Stengel-und-Blätter-Schaubild feststellen.

Für LehrerInnen und SchülerInnen war es schwierig, auch die qualitativen Begriffe als Fachbegriffe in sorgfältiger, disziplinierter Weise einzuführen und zu benutzen. Ebenso schwierig war es, die in Tabelle 1 als *Aufgabenstellungen der Datenanalyse* gekennzeichneten Tätigkeiten *Zusammenfassen, Entfalten, Vergleichen, Abhängigkeit untersuchen* nicht *vollständig* mit der Anfertigung bestimmter Graphiken oder Verfahren zu identifizieren. Der *Vergleich von Verteilungen* war ein Ziel- und Ausgangspunkt der Grundphase des Unterrichts. Vergleiche können und sollen schon anhand von (multiplen) Stengel-und-Blätter-Schaubildern vorgenommen werden. Im Laufe des Unterrichts verfeinern und erweitern sich die Mittel dazu.

Aus den Unterrichtsversuchen haben wir gelernt, daß es wichtig ist, diesen allgemeineren Standpunkt stärker zur Geltung zu bringen. Die Auswahl von Vergleichskriterien und von graphischen Darstellungen zu Vergleichszwecken wird sonst leicht durch die schematische Anfertigung von Boxplots ersetzt oder durch die Anwendung von Kriterien, die SchülerInnen spontan einbringen, wenn Datensätze verglichen werden sollen, nämlich Minimalwert und Maximalwert. Aber auch die Beschreibung und Interpretation von mehreren Boxplots ist nicht trivial.

Als Beispiel zeigen wir zwei Diagramme zu dem Datensatz mit Verkehrstoten, auf den wir auch in Abschnitt 3 noch weiter eingehen. Beide Diagramme wurden im Unterricht eines Lehrers von Schülern erstellt bzw. zur Interpretation vorgelegt, nachdem der Vergleich von Verteilungen anhand von Boxplots an einfachen Beispielen geübt worden war.

Abb. 11: Verkehrstote in der Bundesrepublik Deutschland
1953 - 1987, Boxplots für monatsspezifische Verteilungen

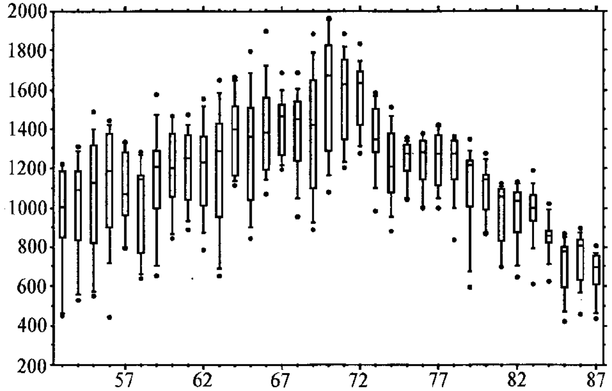


Anm.: 1) In dem Januar-Boxplot sind 35 Daten enthalten, die Anzahl der Verkehrstoten im Januar 1953, Januar 1954, ..., Januar 1987.

2) Die Antennen der Boxplots reichen in Abbildung 11 und 12 bis zum 10 % (90 %)-Wert. Diese Variante wird von Cleveland (1985) als Vereinfachung der ursprünglichen Definition von Tukey vorgeschlagen.

Die Veränderung des Medians ist eines der zu beobachtenden Phänomene. Nicht jedem fällt aber auf, daß die »Wintermonate« Okt., Nov., Dez., Jan., Febr. eine höhere Streuung als die anderen Monate aufweisen. Aber selbst wenn dies registriert wird: Welche allgemeinen Vorstellungen können sich SchülerInnen zu Situationen machen, bei denen sich die Streuung verändert, unabhängig von den Mittelwerten? Erfahrungsgemäß muß hierauf beim Thema Verteilungsvergleich wesentlich mehr Gewicht gelegt werden.

Abb. 12: Verkehrstote in der Bundesrepublik Deutschland
1953 - 1987, Boxplots für die Verteilung innerhalb der Jahre



Anm.: In jedem dieser Boxplots sind die jeweiligen 12 Monate eines Jahres enthalten.

Abbildung 12 zeigt eine Fülle von merkwürdigen Strukturen und Besonderheiten, denen man weiter nachgehen könnte, z. B. die geringer werdende Streuung nach 1971, die Ausreißer 1979 das Auf-und-Ab des Medians usw.

Mit der Anfertigung von Boxplots erreicht man, daß man zunächst sowohl mittlere Werte (Median), Streuung, Form der Verteilung und Ausreißer in einen Vergleich einbeziehen kann. Häufig ist es aber nicht sinnvoll, dabei stehen zu bleiben. Im vorliegenden Fall z. B. könnte eine schrittweise Steigerung der Komplexität sinnvoll sein, z. B. könnten erst nur Mittelwerte (Mediane oder arithmetische Mittel) verglichen werden. Letzterer hätte im vorliegenden Fall sogar den Vorteil, daß die Gesamtzahl rekonstruierbar wird. Danach könnte man die Entwicklung der Streuung in einem eigenen Diagramm untersuchen.

3. Aufbauphase: Graphiken, Begriffe und offenes Material im Unterricht

3.1 Verschiedene Arbeitsweisen im Unterricht zur Entdeckenden Statistik

Ein wichtiges Ziel des Unterrichtsversuches, sowohl im Rahmen der Fortbildung als auch während der beratenden Unterrichtsbeobachtungen, bestand darin, die besondere Denk- und Arbeitsweise der Entdeckenden Statistik in den Mathematikunterricht einzubringen. Wie läßt sich die explorative, offene, statistische Detektivarbeit im Unterricht verwirklichen? Nach einer unterrichtlichen Phase der Einführung und Einübung des elementaren statistischen Handwerkszeugs lassen sich mehr und mehr Möglichkeiten für offenes Arbeiten realisieren. Wie müßte die statistische Detektivarbeit vorbereitet werden? Welche wichtigen Erfahrungen ergeben sich aus diesen ersten Unterrichtsversuchen? Welche konstruktiven Vorschläge zu Verbesserungen und Alternativen lassen sich aus der Reflexion über die curricularen Intentionen und deren unterrichtlicher Umsetzung herleiten?

In den verschiedenen Unterrichtsreihen zur Entdeckenden Statistik ließen sich (idealtypisch) drei Varianten der Aufbauphase identifizieren:

- 1. Typ: Vertiefende Übung der in der Grundphase erarbeiteten Begriffe und Verfahren;
- 2. Typ: Orientierung auf inhaltlich interessante Sachsituationen und Daten: Bearbeitung kleiner und mittlerer Datensätze;
- 3. Typ: Weitgehend offene Problemsituationen mit zunehmender Komplexität: in den Datensätzen, bei Graphiken und Darstellungsmitteln, und in begrifflichen Beziehungen und Strukturen.

Teilweise wurden die im Unterricht beobachtbaren Vorgehensweisen durch unser Arbeitsmaterial nahegelegt; die dort vorgeschlagenen Datensätze und Beispiele konnten sowohl als bloßes Übungsmaterial eingesetzt werden (Typ 1), sie boten aber auch Möglichkeit zu öffnenden und explorativen Arbeitsweisen (Typ 2). Offen gestaltete Projektarbeit sollte nach Vorschlag des Materials zusätzlich und aufbauend auf der Unterrichtsreihe organisiert werden, konnte jedoch ansatzweise im Unterricht selbst beobachtet werden (Typ 3). Wesentliche Intention der Aufbauphase ist es, nach der Einführung in den Gebrauch der elementaren Handwerkszeuge der EDA nun der Detektivarbeit mehr Raum zu gewähren.

Der *erste Typ* der Aufbauphase bestand vor allem aus einer schematischen Übung der erlernten Inhalte der EDA mit teilweiser Einführung von Varianten. In der Art der Unterrichtsführung und den Arbeitsweisen der Schüler unterschied er sich kaum von der Grundphase; es war eine »vorgezogene« Vorbereitung auf den Test. Die behandelten Datensätze waren bloße »Anwendungsfälle« für die Techniken, interessante Phänomene in den Daten kamen kaum zur Sprache. Die Sachhintergründe der Daten waren voneinander isoliert, es gab keine rote Linie. Dies ist eine »extreme« Form, wie sie durch unser Material und die vorbereitende Fortbildung nicht intendiert war; sie zeigte sich vor allem im Unterricht der LehrerInnen, die nur mit schriftlichem Material in ihrer Vorbereitung und Durchführung des Unterrichts gearbeitet haben. Ein erstes Verständnis von der explorativen Vorgehensweise der EDA ist wohl nur in einem sozialen Rahmen des gemeinsamen Lernens, so wie es in den Fortbildungen intendiert wurde, möglich.

Die Aufbauphasen des *zweiten Typs* strebten eine Ausweitung an, indem hier – teils bezogen auf übersichtliche neue Datensätze, teils durch Erweiterung von schon aus der Grundphase her bekannten Datensätzen – nun eine offenere, explorative Arbeitsweise verbunden mit der Bearbeitung interessanter inhaltlicher Aspekte in Angriff genommen wurde. Es wurden beispielsweise über mehrere Unterrichtsstunden hinweg unterschiedliche Datensätze aus gleichen Inhaltsbereichen (z. B. dem Straßenverkehr in der BRD in verschiedenen Jahrgängen) behandelt; es wurden mit vorbereiteten Arbeitsblättern zu Daten in Schülergruppen verschiedene graphische Darstellungen erstellt und hinterher in der Klasse im Hinblick auf ihren unterschiedlichen Gebrauchswert diskutiert, oder es wurden zu einzelnen noch erläuterungsbedürftigen inhaltlichen Fragen zusätzliche Daten und Informationen eingeholt. Ansatzweise nahm die statistische Arbeit hier solche Formen an, wie sie im Rahmen kleinerer Projekte auftreten werden. Dieser zweite Typ einer *Orientierung auf inhaltlich interessante Situationen* eröffnete neue und sachlich vielfältige Analysemöglichkeiten und hatte damit im großen und ganzen eine eigenständige Funktion in der gesamten Unterrichtsreihe. Gute Beispiele für diesen Typ der Aufbauphase werden in dem Bericht von Sensenschmidt und Weinberg (in diesem Heft) beschrieben.

Der *dritte Typ* mit *weitgehend offenen Problemsituationen mit zunehmender Komplexität* läßt sich bis zu einem gewissen Grade als »Erweiterung« des zweiten Typs charakterisieren. Zum einen wurden den SchülerInnen in einzelnen Unterrichtsstunden z. T. sehr umfangreiche Datentabellen zur Analyse vorgelegt, zum anderen war dies häufig mit sehr offenen Fragestellungen und Arbeitsaufträgen verbunden. Die fehlende, direkte Anleitung durch den Lehrer sollte eine offene, explorative Arbeitsweise stimulieren, sie setzt jedoch Kompetenzen bei den SchülerInnen voraus, die sie z. T. erst noch erwerben müssen. Die Durchführung eines solchen offenen Unterrichts erfordert strategische Vorüberlegungen und Planungen. In den Beobachtungen hat sich gezeigt, daß die Schüler zwar mit interessanten Fragen an die Daten herangehen, sich dabei aber auf einzelne Fakten konzentrieren. Häufig ist es nicht möglich, direkt auf solche Fragen eine einfache, mathematische Antwort zu geben – im Unterricht wird dies jedoch oft durch Anwendung von rechnerischen Verfahren gemacht. Damit ergeben sich für einen offenen Unterricht zur Entdeckenden Statistik vor allem zwei Probleme:

- Diskussion und weitere Klärung der von SchülerInnen formulierten Analyseprobleme, bei Vermeidung von voreiligen, rechnerischen Lösungsversuchen.
- Gemeinsam überprüfen, welche Arbeitsmittel der EDA geeignet sind und welche eventuell noch benötigt werden, um sich dem Problem zu nähern; hier ist insbesondere der Einsatz von graphischen Diagrammen gegenüber den rechnerischen Verfahren für ein begriffliches Verstehen der Problemfrage sehr wichtig.

Im folgenden wird ein Beispiel aus einer Unterrichtsstunde vorgestellt, an dem viele Probleme eines offenen Unterrichts zur Entdeckenden Statistik exemplarisch sichtbar werden. Nach der Grundphase hat der Lehrer in dieser Klasse zunächst einen möglichst offenen Arbeitsstil angestrebt, um exploratives Arbeiten zu entwickeln und die Fragen und Ideen der SchülerInnen zur Geltung kommen zu lassen. Diese Aufbauphase entspricht dem 3. Typ einer Bearbeitung von offenen Problemsituationen mit zunehmender Komplexität.

Statistische Detektivarbeit zeichnet sich dadurch aus, gute Fragen an das Problem zu formulieren, Hypothesen zu entwickeln, nach weiteren Indizien zu suchen, die geeigneten Analysemittel einzusetzen bzw. für das Problem zu verändern und dann die »Lösung des Falles« mitzuteilen (vgl. 1.2). Eine so verstandene Detektivarbeit ist sehr voraussetzungsvoll und muß geübt werden. Die Datensätze in unserem Material sollten als Beispiele auf diesem Weg zum EDA-Detektiv dienen; weitgehend offenes und wenig strukturiertes statistisches Arbeiten in Form eines Projektes scheint erst dann möglich. Dementsprechend sind

umfangreiche Datensätze zu komplexen Sachverhalten auch nicht einfach »Anwendungsfelder« des elementaren Handwerkszeugs der EDA, wie z. B. von Stengel-und-Blätter-Schaubild, Boxplot und einfachen numerischen Verfahren, sie erfordern vielmehr die Neuinterpretation und die Variation dieses elementaren Handwerkszeugs.

3.2 Mit welchen Schwierigkeiten ist ein offen strukturierter Unterricht zur Entdeckenden Statistik konfrontiert? – Beispiele aus dem Unterricht

Im Verlaufe einer Unterrichtsreihe befassen sich die SchülerInnen und der Lehrer ca. 6 Stunden mit dem folgenden Beispiel. Die Bearbeitung erfolgt »spontan«, und Planungen für die weitere unterrichtliche Vorgehensweise macht der Lehrer im wesentlichen in Reaktion auf spezifische Vorschläge der SchülerInnen für die statistische Untersuchung; dabei benutzt der Lehrer selbst den Computer, um tabellarische Rechnungen durchzuführen und später auch, um graphische Darstellungen, z. B. Boxplots vorzubereiten. Die numerischen Rechnungen werden von den SchülerInnen im Unterricht jedoch vollständig »per Hand« durchgearbeitet, die vom Computer erzeugten Diagramme werden nur ergänzend und als fertige Blätter bzw. OHP-Folien den SchülerInnen zugänglich gemacht.

Zu Beginn der Bearbeitung des Beispiels legt der Lehrer eine OHP-Folie auf mit einem sehr umfangreichen Datensatz zu den *Verkehrstoten in der Bundesrepublik Deutschland in den Jahren 1953 bis 1987, nach Jahren und Monaten* auf (siehe die vollständige Tabelle[2] bei Sensenschmidt und Weinberg in diesem Heft, Auszüge dieser Tabelle der jährlichen und der monatlichen Summen siehe Tabelle 5 und 6). Die vom Lehrer allgemein formulierte Aufgabe besteht für die SchülerInnen darin, nach Besonderheiten und Strukturen zu suchen. Es werden größte und kleinste Werte benannt; auch wird ein Auf- und ein Absteigen der Werte über die Jahre hinweg angemerkt. Anschließend ergeben sich aus einfachen, teils zufällig und spontan getroffenen Schüleräußerungen schwierige begriffliche Probleme und komplizierte Vorstellungen.

Tab. 5: Summen der Anzahlen der Verkehrstoten in der Bundesrepublik Deutschland für alle Monate der Jahre 1953 bis 1987

Monat	Jan.	Febr.	März	April	Mai	Juni	Juli	Aug.	Sept.	Okt.	Nov.	Dez.
Summe	31284	29031	33799	36215	42969	42716	45894	45224	45895	48143	45342	43403

Tab. 6: Die jährliche Anzahl der Verkehrstoten in der Bundesrepublik Deutschland für die Jahre 1953 bis 1987

Jahr	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962
Summe	11449	12071	12791	13427	13004	12169	13822	14406	14543	14445
Jahr	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972
Summe	14513	16494	15753	16868	17084	16636	16646	19193	18753	18811
Jahr	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982
Summe	16302	14614	14870	14820	14978	14662	13222	13041	11674	11608
Jahr	1983	1984	1985	1986	1987					
Summe	11732	10199	8400	8948	7967					

In der Diskussion kommt zur Sprache, daß es im Februar fast immer die wenigsten Verkehrstoten gibt. Darauf merkt ein Schüler an, der Februar habe ja auch nur 28 Tage, eine für Schüler typische Äußerung, die eine Einzelbeobachtung hervorhebt. Es wird kein Vorschlag gemacht oder vom Lehrer eingebracht, mit Hilfe einer graphischen Darstellung, in Form eines Streudiagramms z. B., die Monatswerte direkt miteinander zu vergleichen. Solche graphisch-funktionalen Darstellungen erlauben es, die Einzelbeobachtungen in einen strukturellen Zusammenhang einzuordnen. Demgegenüber wird später überlegt, wie man mit rechnerischen Methoden die Monate miteinander vergleichen könne. Dazu soll die (durchschnittliche) Anzahl der *täglichen* Verkehrstoten eines Monats (für alle Jahre zwischen 1953 und 1987) ermittelt werden. Wie *rechnet* man dies aus? Und was *bedeutet* dieser Wert? Immerhin hat man eine komplizierte begriffliche Konstruktion für den intendierten Vergleich der Monatswerte eingeführt: *die durchschnittlichen täglichen Januar- (bzw. Monats-)toten für alle Januarmonate zwischen '53 und '87.* Wie kann man diesen Begriff verstehen, welche statistische Interpretation benötigt man hier?

Versucht man die dahinterstehende »Formel« aufzuschreiben, so erhält man:

sei i = 1953, 1954, ..., 1987,
sei j = Januar, Februar bis Dezember ($j = 1, 2, \ldots 12$)
sei n_{ij} = Anzahl der Verkehrstoten im i-ten Jahr und j-ten Monat, und
sei m_j = Anzahl der Tage im j-ten Monat,

$$\text{dann ist } d_j = \frac{1}{m_j}\left(\frac{1}{35}\sum_{i=1953}^{1987} n_{ij}\right)$$

die durchschnittliche Anzahl der Verkehrstoten an einem Tag im Monat j für alle Jahre. Diese komplizierte Formel – die im Unterricht selbst nicht ausgearbeitet wurde – verdeutlicht die schwierige Begriffsbildung.

Die SchülerInnen berechnen die durchschnittliche Anzahl der Verkehrstoten für den Januar zwischen 1953 und 1987. Die in der Tabelle enthaltene Summe von 31284 aller Verkehrstoten aller Januarmonate zwischen '53 und '87 soll durch 35 dividiert werden, weil es insgesamt 35 Monate gibt. In der folgenden Unterrichtssituation zeigt sich dann, wie schwierig für die SchülerInnen die Vorstellung der täglichen Durchschnittsanzahl für einen Monat aller 35 Jahre ist; nur durch enge Führung des Lehrers gelangt man dahin, nun den erhaltenen Wert durch 31, die Zahl der Januartage zu dividieren.

Jetzt wird zum Februar übergegangen. Die Feststellung, daß der Februar 28 Tage hat, muß in eine Rechnung umgesetzt werden. Die Zahl der Februartoten darf nicht durch 31 dividiert werden. Was ist hier die korrekte Zahl? Nacheinander erfolgen von den SchülerInnen Vorschläge, wie »Durch 28!«, »Durch 28,5!« und »Durch 28,3!«, bis die richtige Einsicht in die Abfolge der Schaltjahre zum Divisor 28,25 führt. An diesen beiden Monaten sowie dem März wird dann beispielhaft deutlich, wie *die durchschnittlichen täglichen Januar- (bzw. Monats-)toten für alle Januarmonate zwischen '53 und '87* nacheinander berechnet werden sollen. Die Rechnungen ergeben dann z. B. folgende (gerundete) Werte:

Januar:	28,8 Tote	
Februar:	29,4 Tote	
März:	31,2 Tote	
April:	34,5 Tote	usw.

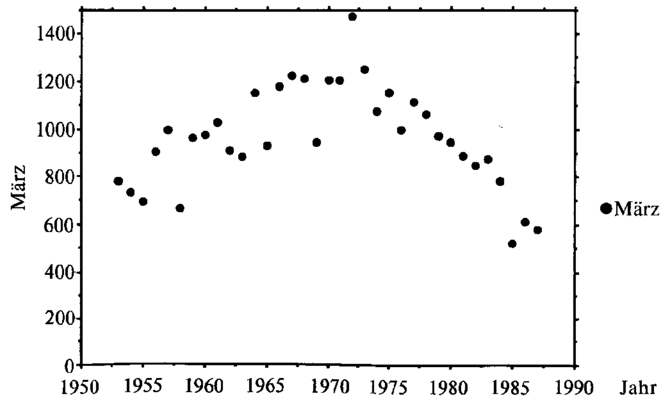
Es ist für den Lehrer an dieser Stelle offenbar schwierig, im komplexen Unterrichtsgeschehen die Grundorientierung der EDA, Daten graphisch darzustellen und zu interpretieren, zur Geltung zu bringen; es wird eine rein rechnerische Vorgehensweise weiterverfolgt. Während der Rechnungen zu dieser kleinen Tabelle, mehren sich die kritischen Rufe und ein Schüler fragt laut: »Wie können denn 0,5 Menschen sterben?« Diese Frage macht ein Verstehensproblem bei statistischen Verhältnissen prägnant sichtbar: Durchschnittswerte, wie z. B. 34,5 Verkehrstote, werden ganz konkret und sachbezogen interpretiert, repräsentieren jedoch einen »idealisierten« Wert, der eigentlich eine komplizierte Beziehung zwischen verschiedenen Anzahlen, hier den Tagen, Monaten und einer Anzahl von Jahren wiedergibt.

Nach dieser Phase der Berechnung von 12 Durchschnittswerten, öffnet der Lehrer wiederum die Arbeitsweise und fragt, was die SchülerInnen nun allgemein interessiert. »Ich möchte wissen, wieviel Tote es heute gibt!« wirft ein Schüler in die Diskussion ein. Er meint damit ausrechnen zu wollen, wie groß die Anzahl der Verkehrstoten für den aktuellen Tag, den 8. März 1990 voraussichtlich sein wird. Insgesamt gesehen hat dieser Schüler eine für statistische Untersuchungen ganz legitime Frage gestellt: Es soll aus vorhandenem Datenmaterial eine statistische Prognose abgegeben werden. Für den einführenden Unterricht in die Entdeckende Statistik wird eine solche Fragestellung jedoch schnell sehr voraussetzungsvoll: Was bedeutet eine solche Prognose? Welche statistischen Mittel sind erforderlich? Welche Verfahren müssen eingesetzt werden?

Der Lehrer läßt sich auf diese Frage ein, und ein Schüler formuliert einen ungenauen Ansatz, der jedoch eine zentrale Idee für eine Prognose zu reflektieren versucht. »Wir müssen die *abfallende Tendenz* in den Märzmonaten irgendwie ausrechnen!« Im weiteren zeigt sich, daß die Idee der *abfallenden Tendenz* noch viel zu allgemein und zu wenig operativ ist. Mehrere Schüler äußern zudem, daß es im März für die letzten Jahre im Grunde keine abfallende, sondern seit 1985 eine teils steigende Tendenz zu beobachten gibt. Die abfallende Tendenz gilt für das ganze Jahr (im Durchschnitt), aber im großen gilt sie auch für die Märzmonate, wenn man ihren Verlauf seit 1970 und nicht nur die letzten vier Jahre betrachtet. Darauf kommt der Vorschlag, zum Monat April überzuwechseln, denn hier sei wirklich eine abfallende Tendenz gegeben. Die Äußerungen der SchülerInnen verdeutlichen, wie strikt deterministisch und kaum statistisch sie die *abfallende Tendenz* interpretieren: nur wenn eine strikt monoton abfallende Zahlenfolge gegeben ist, wollen einige SchülerInnen von einer abfallenden Tendenz sprechen.

Der Lehrer versucht das Problem der »fallenden Tendenz« durch rechnerisches Herangehen zu bearbeiten. Wie anders hätte eventuell die weitere Bearbeitung verlaufen können, wenn im Unterricht ein Computer mit statistischer Software zur Verfügung gestanden hätte? Die hiermit unmittelbar gegebene Möglichkeit, schnell ein Streudiagramm zu erstellen, hätte anders als der bloße rechnerische Umgang mit den Zahlen den Blick auf die »funktionalen« Beziehungen in den Daten gelenkt. Zum Beispiel bekommt man für die Werte der Märzmonate (1953–1987) mit dem Streudiagramm zusätzliche Orientierungen und Vorstellungen zum Trendbegriff, dazu was bei den SchülerInnen »fallende und steigende Tendenz« heißt (vgl. Abb. 13).

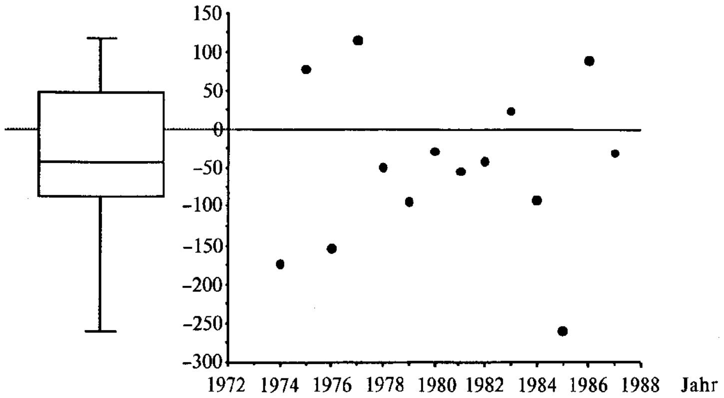
Abb. 13: Streudiagramm
Anzahl der täglichen Verkehrstoten in der Bundesrepublik Deutschland in den Märzmonaten 1953 bis 1987



Im Unterricht wird die Auswertung auf die letzten 14 Jahre von 1974 bis 1987 eingeschränkt; für die einzelnen Monatswerte müssen die SchülerInnen nun die (absoluten) Zuwächse oder Abnahmen der Anzahlen der Verkehrstoten von Jahr zu Jahr bestimmen. Diese einzelnen Werte ergeben zunächst die lokalen Schwankungen, man kann anschaulich keine globale Tendenz von Zu- oder Abnahme ersehen. Insbesondere da dieser Ansatz nicht in ein zeitliches, graphisches Diagramm übertragen wird, bleiben positive Möglichkeiten ungenutzt und die SchülerInnen warten nur auf Anweisungen, wie konkret weiter zu rechnen ist.

Anstatt die Daten in einem zeitlichen Diagramm zu visualisieren, werden im Unterricht nun für die berechneten Differenzwerte die Boxplots für alle 12 Monate von den SchülerInnen erstellt. Deren vergleichende Interpretation stellt sich vor allem deshalb als schwierig heraus, da ganz unerwartet negative Werte auftreten und im Boxplot verarbeitet werden müssen. Diese Anwendung des Werkzeugs »Boxplot« auf den vorliegenden Datensatz ist mit neuartigen Rechenproblemen konfrontiert, die von einer vergleichenden Interpretation wegführen. Wie sähe dies beispielsweise für die Differenzwerte des Monats März (1974 bis 1987) aus? In der Abbildung 14 ist der entsprechende Boxplot mit einem Streudiagramm dargestellt.

Abb. 14: Jährliche Veränderungsraten in der Anzahl der Verkehrstoten in der Bundesrepublik Deutschland
Streudiagramm mit Randverteilung (als Boxplot) für die Jahre 1974 bis 1987



Zunahme bzw. Abnahme der Anzahlen der Verkehrstoten im Vergleich zum Vorjahr., 1974 bis 1987

Dem Boxplot kann man nun im Vergleich mit dem Streudiagramm entnehmen, daß »im Mittel« die Anzahl der Verkehrstoten im Monat März von Jahr zu Jahr um 50 abnimmt. In diesem Problemkontext erhält somit der Boxplot eine neuartige begriffliche Nutzung, die sich auf die Schätzung der »mittleren« Veränderungsrate konzentriert, wobei zugleich die Unsicherheit dieser Schätzung durch die Streuung im Boxplot veranschaulicht wird. Den Boxplot als graphische Zusammenfassung der Verteilung der Differenzwerte aufzufassen und zudem die Boxplots aller zwölf Monate interpretierend zu vergleichen, ist sehr problematisch und stellt hohe begriffliche Anforderungen, die von den SchülerInnen nicht unmittelbar zu erwarten sind.

Für die Entwicklung eines besseren Verständnisses der *abfallenden Tendenz,* wenn man die Behandlung dieses Begriffs an dieser Stelle im Unterricht denn überhaupt für möglich erachtet, ist die Einführung und Nutzung graphischer Darstellungsmittel unbedingt erforderlich. So hätte die Darstellung der 35 jeweiligen monatlichen Werte in einem Streudia-

gramm – wie wir es hier in Abbildung 13 gemacht haben – das Augenmerk stärker auf den globalen Gesamtverlauf lenken können. Das rechnerische Vorgehen engt den offenen Unterricht stark in seinen Kommunikationsformen ein, graphische Darstellungen ermöglichen weiterhin vielfältige, entdeckende Perspektiven auf das Problem und gleichzeitig die Entwicklung interpretativen und begrifflichen Verständnisses.

3.3 Die vielfältigen Bedingungen in Unterrichtsversuchen – Möglichkeiten für Entdeckenden Statistikunterricht auf der Grundlage praktischer Erfahrungen

Die beobachteten Probleme zum Einsatz von Daten und zum Gebrauch und zur Interpretation von graphischen Diagrammen im Entdeckenden Statistikunterricht zeigen, wie schwierig und begrifflich anspruchsvoll ein solches Vorgehen im traditionellen Mathematikunterricht ist. Die hier diskutierten Schwierigkeiten dürfen jedoch nicht zu dem schnellen Urteil verleiten, diese explorative, statistische Arbeit mit jungen SchülerInnen sei so kompliziert, so daß man einen solchen Unterricht eigentlich gar nicht realisieren könnte. Es muß bedacht werden, daß diese Unterrichtsversuche selbst einen explorativen Charakter hatten, und daß auf einen Schlag gleich mehrere Probleme in Angriff genommen werden mußten:

- die Einführung neuer curricularer Inhalte;
- eine neuartige, explorative Arbeitsweise für die SchülerInnen;
- die neue, relationale Auffassung von graphischen Diagrammen (vgl. Biehler [1985]) (für LehrerInnen und SchülerInnen);
- ein »offenes«, mathematisches Arbeiten ohne die üblichen, eindeutigen rechnerischen Ergebnisse und festen Verfahren;
- ungewohnter Umfang an sprachlichen Darstellungen und Diskussionen bei der Bewertung und Interpretation von Daten und Diagrammen.

Nach unseren Beobachtungen und Erfahrungen in Unterricht und Fortbildung liegt das größte Problem bei der Einführung der Entdeckenden Statistik darin, daß sie in der von uns vertretenen explorativen Vorgehensweise sehr stark mit den Arbeitsweisen des traditionellen Mathematikunterrichts in Konflikt gerät. Demgegenüber darf man bei diesen Unterrichtsversuchen die positiven Ansätze und Beispiele nicht zu gering bewerten. Die kritische Reflexion der Unterrichtsbeobachtungen im Rahmen des kooperativen Unterrichtsversuches dient nicht einfach zur Analyse von Fehlern und Mißständen, wie sie sich im praktischen Unterrichtsablauf ereignen, sie sollten uns dazu Hinweise geben, an welchen Stellen in der geplanten Unterrichtsreihe strukturelle Änderungen vorgenommen werden müssen, in welchen Unterrichtsabschnitten paradigmatische Beispiele (beispielhafte »Szenarien«) für Planung und Unterrichtsdurchführung vonnöten sind, und in welcher Weise und welchem Umfang die Vorbereitung der LehrerInnen durch Fortbildung und Material verbessert werden muß.

So sind die von uns im Unterricht beobachteten Schwierigkeiten bei der explorativen Analyse von komplexen Datensätzen nicht bloß als Verstehensprobleme der SchülerInnen oder als fehlende Kenntnisse der LehrerInnen zu erklären; im Unterricht ist es schwierig, eine Balance zwischen der Offenheit der explorativen Arbeitsweise und der Einführung (»Definition«) von angemessenen Begriffen, rechnerischen Mitteln und graphischen Diagrammen herzustellen. Zudem können die Begriffe der EDA nicht so »präzise« definiert werden, wie dies sonst im Matheunterricht geschieht. Und so ergibt sich aufgrund der explorativen Vorgehensweise, daß durch spontane Problem- und Fragestellungen plötzlich neue Begriffe von relativ großer Tragweite benötigt werden, die man nicht einfach im Rahmen scheinbar logisch organisierter Fachstrukturen vorab entwickeln kann.

Die Bearbeitung der komplexen Datentabelle zu den Verkehrstoten in der BRD zeigt diese Probleme exemplarisch auf. Zunächst muß festgestellt werden, daß die SchülerInnen eigenständig gute Anmerkungen und interessante Fragen formulieren, wie man sie für einen explorativen Unterricht unbedingt braucht. Die Bemerkung des einen Schülers, daß der Februar ja nur 28 Tage hat, wird vom Lehrer nicht aufgegriffen, geht anfangs bei dem Vergleich der Monatszahlen fast vollständig unter und wird erst später behandelt. Diese Bemerkung zielt auf die erforderliche Einführung eines »passenden« Durchschnittes zum Vergleich der Monate untereinander. Noch weitergehender und statistisch voraussetzungsvoller ist die von dem zweiten Schüler formulierte Problemstellung: »Ich möchte wissen, wieviel Tote es heute gibt!« (am 8. März 1990). Hiermit wird eine statistische Prognose erfragt. Diese legitime und für einen offen angelegten Unterricht wünschenswerte Frage macht wichtige statistische Analysebegriffe erforderlich.

In diesen Problemen sieht der Lehrer nicht so sehr die Notwendigkeit, Begriffe und Mittel der elementaren EDA zu variieren und auf die vorliegenden Besonderheiten anzupassen, er versucht vielmehr wie im traditionellen Matheunterricht, das vorab in der Grundphase erlernte Handwerkszeug von Stengel-und-Blätter-Schaubild, Median und Boxplot auf diesen Datensatz zu anzuwenden. Allein mit diesen Mitteln kann man die mit der Prognose verbundene zeitliche Entwicklung nicht analysieren und verstehen. Hier benötigt man eine elementare Form des Liniendiagramms.

Im folgenden sollen mögliche Rahmenbedingungen für eine elementare Bearbeitung der Schülerfrage »Ich möchte wissen, wieviel Tote es heute gibt!« dargestellt werden. Eine solche Problemstellung müßte eigentlich zuerst »entschlüsselt«, sie müßte in einen statistischen Rahmen gestellt werden. Kann man überhaupt Vorhersagen machen? Unterliegt der »abfallenden Tendenz« eine Gesetzmäßigkeit, die man mathematisch fassen kann? Es passieren doch täglich viele, unvorhersagbare Dinge, die eine genaue Prognose für einen einzigen Tag unmöglich machen? Diese und ähnliche Einstiegsfragen helfen, die Unterschiede zwischen deterministischem, gesetzmäßigen Verhalten und einer statistischen Aussage bzw. Prognose besser zu verstehen.

Vor diesem Hintergrund müßte der spezifische Tag, der 8. März 1990, »statistisch« interpretiert werden: Geht es konkret um genau diesen einzelnen Tag, oder steht er für eine bestimmte Gruppe von Tagen? Es gäbe verschiedene Interpretationsweisen:

- Man möchte wissen, wie viele Verkehrstote es an einem beliebigen Tag des Jahres 1990 geben wird. Hierbei wäre der 8. März 1990 »ein Tag wie jeder andere« im Jahre 1990.
- Es gibt saisonale Unterschiede in den Anzahlen der Verkehrstoten über die Monate jeden Jahres. Der März gehört zu den jahreszeitlich frühen Monaten mit deutlich geringeren Anzahlen. Daher erscheint es sinnvoll, sich ausschließlich auf die Zahlen der Märzmonate zu beziehen. Dementsprechend repräsentiert der 8. März 1990 einen beliebigen Märztag.
- Ist es richtig, jedem Tag im März die gleiche Zahl (den Durchschnitt) zuzuordnen? Gibt es nicht immer auch Tage im Monat mit wenigen und mit vielen Verkehrstoten? Zur weiteren Bearbeitung dieser Frage müssen im statistischen Jahrbuch die täglichen Zahlen der Verkehrstoten im Monat März für mehrere Jahre zusätzlich herangezogen werden. Es könnte nun die monatliche Streuung der Daten mit Stengel-und-Blätter-Schaubild sowie Boxplot analysiert werden.
- Gibt es inhaltliche Bedingungen für das Schwanken der Anzahlen in den einzelnen Monaten? Die Diskussion dieser Frage kann zu der Einsicht führen, daß es abhängig von den Wochentagen unterschiedlich viele Verkehrstote gibt. An den Wochenendtagen (Freitag, Samstag, Sonntag) ist die Zahl merklich höher, an den anderen Tagen niedriger (vgl. Biehler und Rach [1990, S. 66 ff.]). Der 8. März 1990 war ein Donnerstag, damit ist zu erwarten, daß die Anzahl entsprechend geringer sein wird. Nun repräsentiert der 8. März 1990 die Donnerstage im Monat März 1990.
- Ist es sinnvoll, sich die Anzahlen der Verkehrstoten des 8. März von allen Jahren 1953 bis 1987 zu beschaffen, und dann auf dieser Grundlage eine Prognose zu machen? Warum wird es wohl nicht ausreichen, sich konkret auf die Tage 8. März 1953, 8. März 1954,

8. März 1955, 8. März 1956 usw. zu beschränken? Wenn eine Prognose für den 1. Weihnachtstag, den 25. Dezember 1990 oder für den Silvestertag, den 31. Dezember 1990 abgegeben werden sollte, wäre es in diesen Fällen sinnvoll, sich genau auf diese Daten für die einzelnen Jahre zu beschränken?

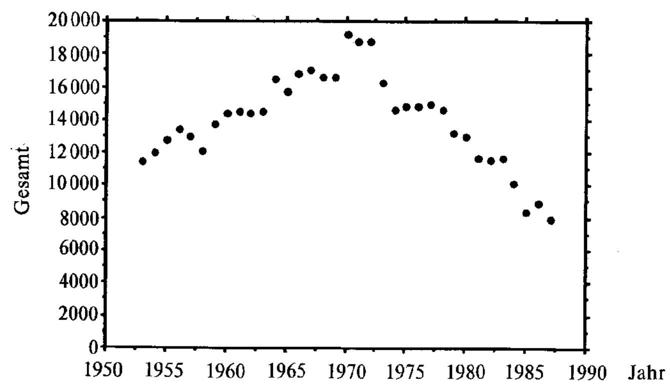
Mit diesen Fragen könnte ein erstes »statistisches Verständnis« davon entwickelt werden, was mit einer Prognose für den 8. März 1990 gemeint ist. Die verschiedenen Interpretationen des betrachteten Datums bieten Möglichkeiten, mit elementaren graphischen Diagrammen und numerischen Betrachtungen eine einfache Prognose für 1990 zu formulieren.

Letztlich erfordert eine intensive statistische Analyse einer solch' komplexen Datentabelle die Benutzung des Computers zusammen mit entsprechender interaktiver Software, die sowohl den Rechenaufwand reduziert als auch schnell graphische Diagramme zu erstellen gestattet. Man hat zu viele Daten, um einfach nur mit Papier und Bleistift an die Arbeit gehen zu können. Andererseits kann man im Unterricht auch nicht ohne gewisse Vorkenntnisse, ohne Fragen zu möglichen Analysen usw. mit dem Computer und der Statistik-Software direkt an die Daten herangehen, der Computer kann die Analyse nicht automatisch vornehmen oder durchführen.

Eine computergestützte statistische Analyse bedarf gewisser Vorkenntnisse der Schüler, Vertrautheit mit dem Sachkontext der Daten, der Möglichkeiten, an die Daten interessante Fragen zu stellen, und der Fähigkeit, Diagramme zu lesen, zu interpretieren und Beziehungen in verschiedenen Diagrammen miteinander zu vergleichen. Natürlich sollen diese Fähigkeiten im Umgang mit der Software auch weiterentwickelt werden. Die folgende beispielhafte Analyse versucht in diesem Sinne statistische Software zu nutzen.

Wird der 8. März als ein beliebiger Tag im Jahr angesehen, so könnte man die Entwicklung der jährlichen Zahlen untersuchen (vgl. Biehler und Steinbring [1990]). Beispielhaft könnten folgende Diagramme diskutiert werden:

Abb. 15: Streudiagramm
Anzahl der Verkehrstoten in der Bundesrepublik 1953-1987

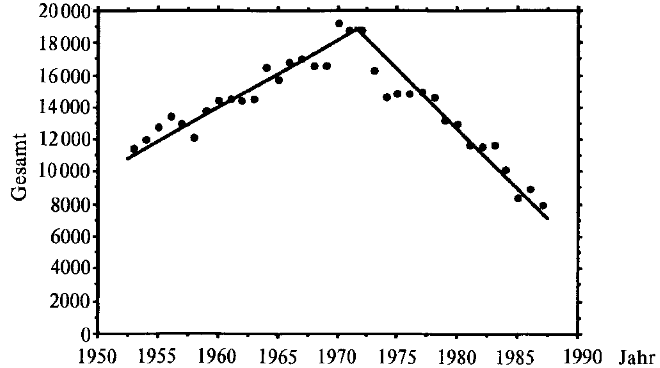


Folgende Fragen könnten Gegenstand der Diskussion sein:

- Was ist in diesem Diagramm dargestellt?
- Welch eine Form hat der zeitliche Verlauf der Anzahl der Verkehrstoten?
- Wie kann man mit Hilfe des Diagramms neue Beschreibungen zur Erklärung von »steigender und fallender Tendenz« entwickeln?

In dem folgenden Diagramm ist eine »Dachkurve« zur Darstellung einer möglichen linearen Entwicklung eingezeichnet worden.

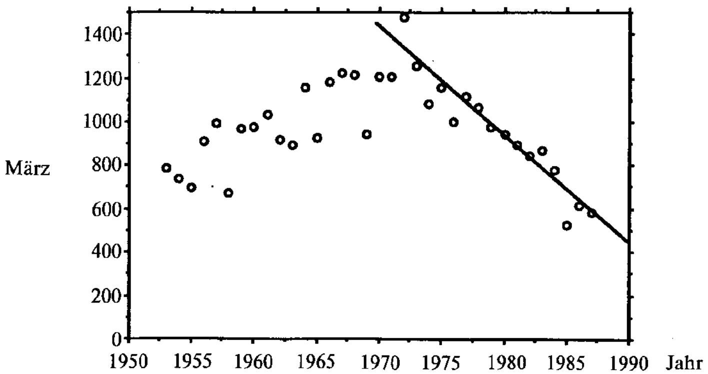
Abb. 16: Streudiagramm mit »Dachkurve«
Anzahl der Verkehrstoten in der Bundesrepublik Deutschland 1953-1987



Mit Hilfe dieses Diagramms läßt sich die Idee »abfallende Tendenz« vor dem Hintergrund des Verhältnisses von lokalen Schwankungen und globalen Verläufen besprechen. Auch ist es nun möglich, Vorstellungen für die Prognose zu entwickeln. Reicht es aus, einfach die Linie zu verlängern? Wie können die zu erwartenden lokalen Schwankungen in der Prognose berücksichtigt werden? Muß man nicht obere und untere Werte angeben? Desweiteren: Kann es mit der im Diagramm dargestellten linearen Entwicklung immer so weitergehen? Ist es wirklich möglich, die Anzahl der Verkehrstoten bis ca. 2000 auf Null zu drücken? Eine mögliche einfache statistische Antwort könnte lauten: Die Gesamtzahl wird voraussichtlich zwischen 7000 und 9000 Verkehrstoten liegen. Das bedeutet im Durchschnitt eine Zahl zwischen von ca. 19 bis 25 Verkehrstoten für einen Tag im Jahre 1990[3].

In vergleichbarer Weise könnte man vorgehen, wenn man sich auf die Märzmonate beschränkt. In dem folgenden Diagramm (Abb. 17) sind die Daten der Märzmonate graphisch repräsentiert. Entsprechend der Schätzung der »mittleren« Veränderungsrate für die Märzmonate mit Hilfe des Boxplots ist auch die Gerade mit Steigung -50 »per Hand« eingepaßt worden, um auf diese Weise eine einfache Abschätzung für das Jahr 1990 zu erhalten.

Abb. 17: Streudiagramm
Anzahl der Verkehrstoten in der Bundesrepublik Deutschland 1953 bis 1987
Monat März, mit einer Geraden zur Abschätzung der Veränderungen



Jetzt könnte die unterrichtliche Arbeit zu einer ersten Prognose kommen, bei der nach dem als linear angenommenen Verlauf der Entwicklung ein Wert von ca. 450 Verkehrstoten zu erwarten ist. Führt man in die Prognose noch in mehr oder weniger plausibler Weise Schwankungsbreiten ein, so könnte man als niedrigsten Wert ca. 350 und als höchsten ca. 550 Verkehrstote annehmen. Der tägliche durchschnittliche Wert läge dann zwischen ca. 11 und ca. 18 Verkehrstoten. Diese Prognose sollte mit der ersten verglichen und die Unterschiede sollten erläutert werden. Im März 1990 gab es in der (alten) BRD 568 Verkehrstote. Man muß also wohl feststellen, daß sich der abnehmende Trend seit 1985 nicht mehr fortsetzt! Am 8. März 1990 gab es 18 Verkehrstote.

Um die Schwankungen der Zahlen in den Märzmonaten zu studieren, benötigt man die täglichen Zahlen. Diese 31 Werte lassen sich dann zunächst einmal mit dem Stengel-und-Blätter-Schaubild darstellen. Das Schwanken wird in einem Streudiagramm deutlich sichtbar (vgl. Abb. 18 und 19).

Abb. 18: Streudiagramm
Anzahl der täglichen Verkehrstoten in der Bundesrepublik Deutschland im März 1986

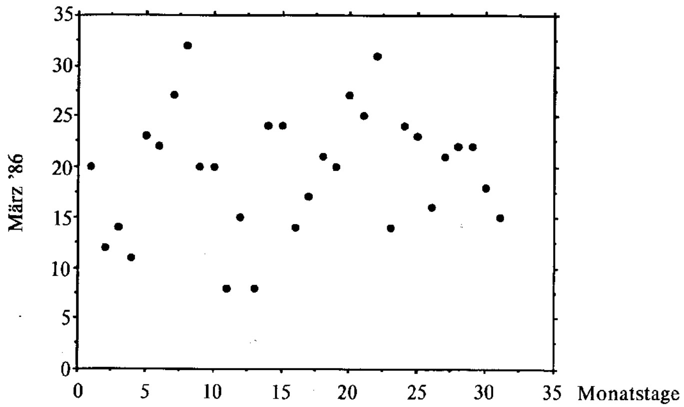
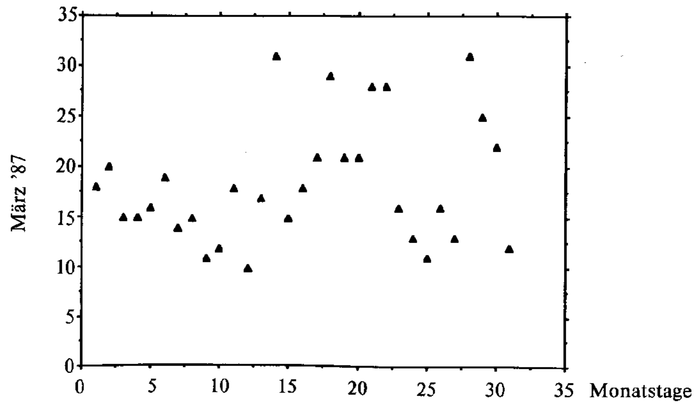
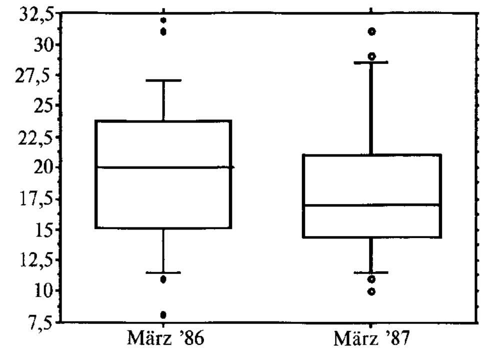


Abb. 19: Streudiagramm
Anzahl der täglichen Verkehrstoten in der Bundesrepublik Deutschland im März 1987



Diese streuenden Werte lassen sich mit Hilfe von Boxplots vergleichen (Abb. 20).

Abb. 20: Boxplots
Anzahl der täglichen Verkehrstoten in der Bundesrepublik Deutschland im März 1986 und März 1987



Wie sind die hier beobachtbaren Extremwerte von ca. 8 und ca. 32 Verkehrstoten an Tagen im März mit den in der Prognose gegebenen Werten zu vergleichen? Ist es sinnvoll, sich auf die Donnerstage im März zu beschränken, um eine Prognose für Donnerstag, den 8. März 1990 abzugeben? Betrachten wir beispielhaft die Anzahl der Verkehrstoten an den März-Donnerstagen von 1986 und 1987:

Donnerstag	1. 3. '86	8. 3. '86	15. 3. '86	22. 3. '86	29. 3. '86	7. 3. '87	14. 3. '87	21. 3. '87	28. 3. '87
Verkehrstote	20	32	24	31	22	15	20	11	20

An den fünf März-Donnerstagen 1986 gab es 129 Verkehrstote, im Durchschnitt 25,8 Tote, das ist mehr als der tägliche Märzdurchschnitt von 19,7 Toten; anders ist es im März 1987, mit durchschnittlich 16,5 Verkehrstoten an einem Donnerstag, der geringer als der tägliche Märzdurchschnitt von 18,7 Toten ist. Es gibt zu wenig März-Donnerstage, es können starke lokale Schwankungen auftreten, die den globalen Effekt verdecken, daß über das ganze Jahr betrachtet die Donnerstage zu den Wochentagen zählen, an denen es etwas weniger Verkehrstote als durchschnittlich an einem Wochentag gibt.

4. Resümee

Wichtige Ausgangspunkte unserer Überlegungen waren eine kritische Haltung gegenüber dem aktuellen Stochastikunterricht, insbesondere der (beschreibenden) Statistik. Im alltäglichen Unterrichtsgeschehen verkümmert die Statistik (in der Jahrgangsstufe 7/8) doch sehr häufig zu einem Repertoire von numerischen Rezepten und wenigen, zumeist technisch komplexen Standardgraphiken.

Gegenüber dieser Situation scheint die elementare EDA nun positive Alternativen und wichtige Ergänzungen für den Stochastikunterricht insgesamt anzubieten. Die Entdeckende Statistik beginnt mit einem Set an Handwerkszeug, das in seinen Grundformen einfach aufgebaut und von den SchülerInnen leicht zu verstehen ist. So war es in der Tat in allen beobachteten Klassen der Fall, daß innerhalb von ca. 8 Unterrichtsstunden ein erstes Verständnis von der Beschaffenheit der elementaren Handwerksmittel der EDA (Stengel-und-Blätter-Schaubilder, numerische Zählverfahren für den Median usw., Boxplot und auch Li-

niendiagramm) bei allen SchülerInnen vorhanden war. Diese Mittel sind technisch einfach und ihre »Herstellung« bei der Analyse von kleinen und mittleren Datensätzen geht nach etwas Übung schnell von der Hand – anders als z. B. die Erstellung eines genau anzufertigenden Histogramms. In dieser Hinsicht wurde somit bestätigt, daß im Rahmen der Grundphase die Einführung in das Grundrepertoire des EDA-Handwerkszeugs in den Statistikunterricht ohne große Probleme möglich ist. Neben der Einfachheit dieser Mittel stellt sich natürlich die Frage, ob die Unterrichtsführung während der Grundphase nicht stärker dem traditionellen Vorgehen entsprach und somit hier das Gelingen positiv gefördert hat.

Die Ausweitung der Arbeitsmittel sowohl im Hinblick auf größere und damit »inhaltlich offenere« Datensätze als auch im Hinblick auf begriffliche Variationen, Neuinterpretationen und Neuentwicklungen von Darstellungs- und Analysemitteln zeigten dann in unterschiedlicher Weise neue und für den traditionellen Matheunterricht ungewohnte Probleme, auf Seiten der LehrerInnen und auf Seiten der SchülerInnen. Unsere nachträgliche kritische Analyse von Unterrichtsbeobachtungen hat uns doch recht plastisch deutlich gemacht, daß wir die Probleme mit der »begrifflichen« Seite der EDA unterschätzt haben. Von ihrem Selbstverständnis her betrachtet, scheint die EDA, vereinfacht gesagt, nur aus Darstellungsmitteln, aus numerischen Verfahren und aus geschickten, beispielhaften Anwendungen und dabei Variationen dieser Mittel zu bestehen. Die begrifflichen Aspekte werden sozusagen nebenher mitgeliefert, ergeben sich automatisch.

Die Dominanz der Mittel (Diagramme und Verfahren) gegenüber der begrifflichen Seite kommt dem traditionellen Mathematikunterricht insofern entgegen, als auch hier wenig explizite Reflexion auf begriffsbezogene Interpretationen gelegt wird. Eine zusätzliche Schwierigkeit besteht für die EDA nun darin, daß ihre Begriffe nicht im Rahmen einer theoretischen Struktur eindeutig festgelegt und formal definiert sind; es gibt somit auch keine schulgemäßen methodischen Definitionen dieser Begriffe. Das besondere Erfordernis des Verstehens dieser Begriffe liegt darin, daß ihre Bedeutung erst im Vollzuge von verschiedenen Anwendungen und Analysen von Daten sichtbar und erfahrbar wird. Für die Kommunikation zwischen LehrerInnen und SchülerInnen, und auch zwischen LehrerInnen und Forschern ist es aber notwendig, daß diese Begriffe terminologisch und von ihren begrifflichen Strukturen beschreibbar sind, auch wenn es keine einfache, übliche schulmathematische Definition gibt.

Vor allem haben die Probleme des offen strukturierten Unterrichts zur Entdeckenden Statistik während der Aufbauphase gezeigt, daß der begrifflichen Seite der EDA im Unterricht mehr explizite Aufmerksamkeit geschenkt weden muß, daß hier eine ausgleichende Balance zu den »Methoden« hergestellt werden muß. Unsere kritische Reflexion hat dementsprechend u. a. zu einer neuen Charakterisierung der in der elementaren EDA wichtigen Grundbegriffe geführt, in Form einer Ausarbeitung und Weiterführung der vor dem Versuch beschriebenen Struktur (vgl. 1.3, Tabelle 1 mit 2.3, Tabelle 4).

Unterrichtsversuche intendieren häufig die Erprobung eines neuen Fach*inhaltes*. Dazu werden vorab Materialien entwickelt, deren Brauchbarkeit in (teilweise mehreren) Unterrichtsdurchläufen erprobt wird und woran sich dann die Verbesserung der Materialien anschließt. Unser Augenmerk war darüber hinaus gerichtet auf: Bedingungen der Kooperation zwischen didaktischer Theorie und Unterrichtspraxis, Einflüssen der Interaktion im Unterrichtsverlauf zwichen LehrerIn und SchülerInnen, den dort auftretenden Verstehensproblemen, der jeweiligen Unterrichtsführung des Lehrers, seine spezifische Weise der Gestaltung der Aufbauphase usw. Dieses breite Spektrum an einwirkenden Aspekten hat uns ein komplexes Bild der Bedingungen einer Unterrichtsrealisierung vermittelt. Die neuen, kritischen und konstruktiven Einsichten, die wir dabei gewonnen haben und die nun in diesem Papier zu weiterführenden didaktischen Konzeptionen der elementaren EDA beigetragen haben, sind nur vor diesem Hintergrund praktischer Erfahrungen möglich geworden.

Die LeserInnen unseres Aufsatzes sollte bei seiner rückblickenden Bewertung über die hier referierten Unterrichtsbeobachtungen bedenken, daß die von uns hier entwickelten Konzepte und alternativen Vorgehensweisen den LehrerInnen während des Versuches noch nicht zur Verfügung gestanden haben, ja daß diese erst durch und nach dem Unterrichtsversuch ausgearbeitet werden konnten. An dieser Stelle bleibt es uns allen beteiligten LehrerInnen für ihre Bereitschaft, im Versuch mitzuwirken, zu danken; ohne ihr Einverständnis, an den alltäglichen Unterrichtsstunden teilzunehmen – was ja immer noch nicht selbstverständlich ist – und so einen umfassenden Einblick in die vielfältigen professionellen und persönlichen Bedingungen zu gewinnen, hätten wir diese konstruktive Weiterentwicklung nicht vornehmen können; die von uns formulierte Kritik, so muß gesagt werden, richtet sich nicht einfach gegen den beobachteten Unterricht, sondern ist vor allem eine Kritik aus der Sicht der Unterrichtspraxis, die zur Weiterentwicklung der didaktischen Konzeption der Entdeckenden Statistik beitragen soll.

Anmerkungen

1) Modellcharakter: Modelle haben einen Abbildungsaspekt, einen Auswahlaspekt und einen intentionalen Aspekt.
2) Die im Material enthaltene Tabelle umfaßte die Werte für die Jahre 1953–1987, während des Unterrichtsversuches sind den LehrerInnen dann auch die Zahlen für 1988 mitgeteilt worden. Jeweils im August des darauffolgenden Jahres werden die vollständigen Zahlen zum Straßenverkehr eines Jahres vom Statistischen Bundesamt veröffentlicht.
3) Tatsächlich gab es im Jahre 1990 in der (alten) BRD 7906 Verkehrstote.

Literatur

[1] *Biehler, R.* (1982): Explorative Datenanalyse – Eine Untersuchung aus der Perspektive einer deskriptiv-empirischen Wissenschaftstheorie. IDM Materialien und Studien Bd. 24. Bielefeld: Universität Bielefeld
[2] *Biehler, R.* (1985): Graphische Darstellungen. In: mateematica didactica, 8, S. 57–81
[3] *Biehler, R.* (1990): Daten analysieren mit dem Computer: Unterstützung von Begriffsbildung und Anwendungsorientierung in der Stochastik. In: Der Mathematikunterricht, 36 (6), S. 50–71
[4] *Biehler, R./Rach, W.* (1990): Softwaretools zur Statistik und Datenanalyse: Beispiele, Anwendungen und Konzepte aus didaktischer Sicht. Neue Medien im Unterricht. Soest: Soester Verlagskontor
[5] *Biehler, R./Steinbring, H.* (1990): Entdeckende Statistik im Mathematikunterricht. Materialien aus dem GRAPHDAS-Projekt. Bielefeld: Universität Bielefeld
[6] *Borovcnik, M.* (1990): Explorative Datenanalyse – Techniken und Leitideen. In: Didaktik der Mathematik, 18 (1), S. 61–80
[7] *Borovcnik, M./Ossimitz, G.* (1987): Materialien zur Beschreibenden Statistik und Explorativen Datenanalyse. Schriftenreihe Didaktik der Mathematik der UBW Klagenfurt 11. Wien/Stuttgart: Hölder-Pichler-Tempsky/Teubner
[8] *Cleveland, W. S.* (1985): The Elements of Graphing Data. Monterey: Wadsworth
[9] *Konold, C.* (1990): ChancePlus: A Computer Based Curriculum for Probability and Statistics. Annual Review: Year 1. Amherst: Scientific Reasoning Research Institute, University of Massachusetts
[10] *Mosteller, F./Tukey, J. W.* (1977): Data Analysis and Regression. A Second Course in Statistics. Reading: Addison-Wesley
[11] *Nordmeier, G.* (1989): »Erstfrühling« und »Aprilwetter«-Projekte in der Explorativen Datenanalyse. In: Stochastik in der Schule, 9 (3), S. 21–42
[12] *Polasek, W.* (1988): Explorative Daten-Analyse. Berlin: Springer
[13] *Rosebery, A. S.* (1989): Reasoning under uncertainty: Developing statistical reasoning. In: Journal of Mathematical Behavior, 8, S. 205–219
[14] *Tukey, J. W.* (1977): Exploratory Data Analysis. Reading: Addison-Wesley

Unterrichtsversuche zur Explorativen Datenanalyse in Klasse 8 – zwei Erfahrungsberichte[1]

von Karin Sensenschmidt und Peter Weinberg

Einleitung

Seit einigen Jahren gehört der Stochastikunterricht gemäß Richtlinien zum verbindlichen Bestandteil des Mathematikunterrichts an Gesamtschulen in Nordrhein-Westfalen. Vorgesehen ist für die Jahrgangsstufe 5/6 eine grundlegende Einführung in die Wahrscheinlichkeitsrechnung. Hier wird den Schülern der Wahrscheinlichkeitsbegriff näher gebracht und der mathematische Wahrscheinlichkeitsbegriff gegenüber dem umgangssprachlichen abgesetzt. Elementare Probleme werden mit Hilfe von Baumdiagrammen und der dazugehörigen Regeln gelöst. In der Jahrgangsstufe 7/8 folgt eine Unterrichtseinheit zur Statistik. Anhand von statistischen Untersuchungen, die die Schüler oft selbst vorbereiten und durchführen, werden gängige statistische Kenngrößen wie Mittelwert, Varianz und Standardabweichungen eingeführt. Außerdem werden zur Auswertung geeignete Diagrammformen wie Balken-, Stab-, Tortendiagramme und Zeitreihen herangezogen. Thema der Jahrgangsstufe 9/10 ist dann wieder die Wahrscheinlichkeitsrechnung. Mit Hilfe realer oder erdachter Experimente, die auf Baumdiagramme führen, werden die Grundlagen der Binomialverteilung erarbeitet. Der hier folgende Artikel befaßt sich nun mit einer Modifikation der Unterrichtseinheit zur Statistik der Jahrgangsstufe 7/8. Die Grobvorbereitung, die Planung und die Begleitung des Unterrichtsversuchs wurden gemeinsam von einer Projektgruppe getragen, in der außer uns Rolf Biehler und Heinz Steinbring vom Institut für Didaktik der Mathematik der Universität Bielefeld mitarbeiteten.

Die folgenden zwei Erfahrungsberichte konzentrieren sich auf unterschiedliche Aspekte: Im ersten Bericht steht mehr die Einführung und der Test am Ende im Vordergrund; im zweiten Bericht wird stärker dargestellt, wie mit einem komplexen Anwendungsbeispiel im Unterricht umgegangen wurde. Ausgehend von gemeinsamen Ideen nahmen unsere beiden parallel durchgeführten Unterrichtseinheiten teilweise einen recht unterschiedlichen Verlauf, wie es bei allem lebendigen Unterricht sein sollte, in den die Vorlieben und Kenntnisse aller Beteiligten eingehen können.

A. Interpretation von Daten beim Einstieg und beim Abschlußtest

Erfahrungsbericht von Peter Weinberg

1. Vorüberlegungen zur Unterrichtseinheit

In der Jahrgangsstufe 7/8 werden normalerweise statistische Kenngrößen wie arithmetisches Mittel und Standardabweichung mit den Schülern erarbeitet. In der Regel wird im Unterricht jedoch nicht thematisiert, daß diese Kenngrößen ihre Bedeutung erst voll entfalten, wenn die betrachteten Daten z. B. normalverteilt sind. Bei anderen Verteilungen (z. B. Gleichverteilung) können sie aber durchaus von untergeordneter Bedeutung sein.