

---

## Anwendungsmöglichkeiten von Software in der Archäologie

### - Datenbank, Datenanalyse, Grafik -

Uwe Schoenfelder (1)  
(unter Mitarbeit von Manfred Erdelen)

Computeranwendungen in der Archäologie sind in Deutschland schon relativ früh praktiziert worden. Impulse gingen wie in vielen anderen Bereichen von den USA aus, die schon in den frühen 50er Jahren den Großrechner für archäologische Fragestellungen nutzten.(2) In der Bundesrepublik publizierten bereits in den späten 60er Jahren Steuer und Goldmann erste konkrete EDV-Auswertungen, ab den 70er Jahren stieg die Zahl solcher Arbeiten ständig. (3) Trotzdem sind Computeranwender in der Archäologie noch relativ selten anzutreffen, so ist auch der Meinungs austausch auf wenige "Insider" beschränkt. Deshalb richtet sich dieser Aufsatz in erster Linie an Archäologen, die noch keine oder nur wenig Erfahrung mit der EDV-gestützten Aufnahme und Verarbeitung archäologischer Daten besitzen.(4)

Mit Sicherheit liegt einer der wesentlichen Gründe für die noch geringe EDV-Akzeptanz in der Problematik des Programmierens verborgen, wie im übrigen schon Gebühr u. Kampfmeyer (1980/81, 6) betonten. Das Gespür bzw. die Begabung, Lösungen durch Algorithmen zu erreichen, ist nicht jedermanns Sache. Außerdem sind viele Programme erst nach langen Probeversuchen lauffähig, abgesehen davon, daß man natürlich erst einmal die Syntax einer Programmiersprache erlernen muß. So ist die Entwicklung zu begrüßen, daß es seit einigen Jahren kommerzielle, auch für die Archäologie nutzbare Software-Pakete für Großrechner und PCs gibt, die infolge ihrer Menü-Steuerung ohne Programmierkenntnisse relativ einfach zu handhaben sind. Solche Programme sind auch für den kleinen bis mittleren Bereich archäologischer Datenanalyse interessant, so für Magister- und Promotionsarbeiten; sie sind aber auch beispielsweise für Ausgrabungsanalysen oder im Museum bei Inventarisationsarbeiten anwendbar.(5)

Einige dieser zumeist kommerziellen Programme, die z.T. fünfzig Mann-Jahre Entwicklungsarbeit bedeuten, möchten wir im folgenden kurz vorstellen. Dabei kann es sich, aufgrund der Vielzahl der mittlerweile auf den Markt gebrachten Programme, nur um eine subjektive Auswahl handeln. Es wurde allerdings auf die praktische Anwendbarkeit für den Archäologen, die Kompatibilität zu anderen Programmen sowie die weite Verbreitung geachtet. Im übrigen sind fast alle Programme, die wir nachfolgend behandeln werden, an den Rechenzentren (in unserem Fall an den Universitäten Bochum und Essen) vorhanden und jedem Studenten und Wissenschaftler mit einer Rechenerlaubnis zugänglich. Auf die für den Archäologen wichtigsten Bereiche, die mit solchen kommerziellen, leicht zugänglichen/weit verbreiteten Software-Programmen sinnvoll abgedeckt und verbunden werden können, wollen wir im Rahmen dieses Aufsatzes näher eingehen: (6)

- Datenbanken
- Statistische Auswertung
- Graphik

## 1. Datenbanken

Das systematische Erfassen von Daten ist ein alltäglicher Vorgang. Ob wir mit einem Zettelkasten oder einem Aktenordner arbeiten, immer werden Daten auf bestimmte Weise geordnet. Nicht anders arbeitet eine EDV-Datenbank. Allerdings müssen bei Computer-Datenbanken die Datentypen fest in Feldern organisiert werden. Wie dies geschieht, werden wir anhand konkreter Beispiele mit archäologischen Daten aufzeigen. Ausgangspunkt unserer Beispiele bilden die aus einer Dissertation stammenden Daten mehrerer Gräberfelder der Bronze- und beginnenden Eisenzeit am unteren Niederrhein (Schoenfelder 1989).

Zunächst wurden die Variablen bzw. Felder, wie z.B. Fundort, Meßtischblattnummer oder Keramiktypologie (abgekürzt FNDORT, M\_BLATT\_NR, TYPOLOG - siehe Abb. 1) mit einer dBASE-Maske (siehe Bemerkungen dazu w.u.) festgelegt, danach der Inhalt oder Wert dieser Felder kodiert eingegeben (Abb. 2).

### dBASE IIIplus, dBASE IV 1.1

Das von der Software-Firma Ashton-Tate entwickelte Datenbanksystem dBASE IIIplus und sein Nachfolger dBASE IV gehören samt ihren Vorläufern zu den am weitesten verbreiteten Datenbanksystemen vor allem im mittleren Bereich. Wir wollen im folgenden die ältere Version dBASE IIIplus behandeln - weil gerade diese Version besonders bekannt ist - und auch auf die aktuelle, stark verbesserte Version dBASE IV 1.1 näher eingehen, die sich in manchem von ihrem Vorgänger unterscheidet.

```
Structure for database: C:\WEDAUI.dbf
Number of data records: 663
Date of last update : 06/24/88
```

Field	Field Name	Type	Width	Dec	
1	FNDORT	Character	4		
2	TAFEL ABB	Character	14		
3	FND NR	Character	5		
4	HRKNFT S	Character	2		
5	FUNDSTELLE	Character	4		
6	LND BNDLND	Character	4		
7	REG BEZ	Character	4		
8	KREIS ALT	Character	2		
9	KREIS NEU	Character	4		
10	GEMEINDE A	Character	2		
11	GEMEINDE N	Character	7		
12	F ORT NR A	Character	3		
13	F ORT NR N	Character	7		
14	M_BLATT_NR	Character	4		
15	KOORD R_H	Character	10		
16	VERBLEIB	Character	3		
17	AUSGRAEBER	Character	3		
18	AUTOREN	Character	3		
19	LIT JAHR	Character	7		
20	GEOGR LAGE	Character	5		
21	PALEOGR_LA	Character	2		
22	GEOL L ALL	Character	2		
23	GEOL L SPZ	Character	2		
24	BODENART	Character	6		
25	BEFUND SIC	Character	2		
26	GRAEBERFEL	Character	1		
27	BEST ART	Character	2		
28	HUEGEL NR	Character	4		
29	HUEGEL HOE	Character	2		
30	HUEGEL DIA	Character	2		
31	GRAB NR	Character	4		
32	GRAB TYP	Character	2		
33	GR GRUB TY	Character	1		
34	GRÜBHOEHE	Character	2		
35	GRUB DIA	Character	2		
36	GRABENTYP	Character	2		
37	STRATA FND	Character	2		
38	STRATA BOD	Character	2		
39	SIED STYP	Character	2		
40	GRUB NR	Character	3		
41	GRUB TYP	Character	1		
42	GRUB HOEHE	Character	2		
43	GRUB DIAM	Character	2		
44	STRATIGR	Character	1		
45	HORTF ART	Character	2		
46	ENZLF ART	Character	2		
47	FUNDE SICH	Character	2		
48	KERAMIK	Character	1		
49	TYPOLOG	Character	5		
50	GEFAESS TY	Character	2		
51	RANDLIP TY	Character	2		
52	RAND TYP	Character	2		
53	HALSTYP	Character	2		
54	SCHULTR TY	Character	2		
55	BAUCH TY	Character	2		
56	UNTERTEIL	Character	2		
57	BODEN TYP	Character	2		
58	VERZ ART	Character	4		
59	VERZ STELL	Character	2		
60	RAND DIA	Character	3		
61	BODEN DIA	Character	3		
62	BAUCH DIA	Character	3		
63	WAND DIA	Character	2		
64	HOEHE	Character	3		
65	MAGER ART	Character	1		
66	MAGER KLAS	Character	2		
67	OF AUSSEN	Character	2		
68	OF INNEN	Character	2		
69	OB AUSSEN	Character	2		
70	OB INNEN	Character	2		
71	ERHALTUNG	Character	1		
72	INVENT_NR	Character	9		
73	METALL	Character	1		
74	METALL ART	Character	2		
75	SCHWERT	Character	2		
76	BEILE	Character	2		
77	LANZEN	Character	2		
78	NADELN	Character	2		
79	RINGE	Character	2		
80	DOLCHE	Character	2		
81	MESSER	Character	2		
82	ANHAENGER	Character	2		
83	PFEILSPITZ	Character	2		
84	FIBELN	Character	2		
85	GUERTELBL	Character	2		
86	SICHELN	Character	2		
87	GEFAESSE	Character	2		
88	SONSTIG	Character	2		
89	ERHALT A	Character	1		
90	INVENT_NRA	Character	9		
91	STEIN	Character	1		
92	STEINART	Character	2		
93	FUNKTION	Character	2		
94	ERHALT B	Character	1		
95	INVENT_NRB	Character	9		
96	TON	Character	1		
97	SPINWIRTEL	Character	1		
98	WEBGEWICHT	Character	1		
99	SONST C	Character	1		
100	ERHALT C	Character	1		
101	INVENT_NRC	Character	9		
102	HOLZ	Character	1		
103	HOLZART	Character	2		
104	FUNKT D	Character	2		
105	ERHALT D	Character	1		
106	INVENT_NRD	Character	9		
107	GLAS	Character	1		
108	GLASTYP	Character	2		
109	ERHALT E	Character	1		
110	INVENT_NRE	Character	9		
111	KNOCHEN	Character	1		
112	KNOCH TYP	Character	1		
113	FUNKT F	Character	1		
114	ERHALT F	Character	1		
115	INVENT_NRF	Character	9		
116	SKELETT	Character	1		
117	SKELET TYP	Character	1		
118	ALTER KLAS	Character	2		
119	GESCHLECHT	Character	1		
120	ANORMAL	Character	2		
121	INDIV ANZ	Character	2		
122	GEWICHT	Character	4		
123	ERHALT G	Character	1		
124	INVENT_NRG	Character	9		
125	BOTANIK	Character	1		
126	BOTANIKART	Character	2		
127	ERHALTUNGH	Character	1		
128	INVENT_NRH	Character	9		
				** Total **	367

Abb. 1 Datenbank-Struktur (aus: Schoenfelder 1989, Abb. 13).

CURSOR	<-- -->	UP	DOWN	DELETE	Insert Mode: Ins
Char:	Del	Exit/Save:	^End		
Word:	Home End	Page:	PgUp PgDn	Field: ^Y	Abort: Esc
		Help:	F1	Record: ^U	Memo: ^Home

```

FNDORT      0001
TAFEL_ABB  T1,1/87,1a
FND_NR      00001
HRKNFT_S    11
FUNDSTELLE 0001
LND_BNDLND 0101
REG_BEZ     5100
KREIS_ALT   01
KREIS_NEU   5112
GEMEINDE_A  00
GEMEINDE_N  0000000

```

```

EDIT          ||<C:>||WEDAUI          ||Rec: 1/663          ||

```

Abb. 2 Codierte Dateneingabe in dBASE IIIplus-Maske (Modus: Edit).

Erst die dritte erweiterte dBASE-Version (dBASE III bzw. IIIplus) erlaubt im Gegensatz zu den früheren Versionen I bis II – die Verwendung der Assist-Benutzeroberflächen, wobei Befehle aus einem vorgegebenen Menü nur noch ausgewählt zu werden brauchen. Diese Menü-Befehle sowie deren Unterbefehle (sog. "pull-down-menus") enthalten sozusagen "vorprogrammierte" Routinen (z.B. für die Felderdefinition einer Datenbank, das Herausfiltern von bestimmten Informationen etc.); somit sind Programmierkenntnisse nicht erforderlich. Man kann jedoch, und das ist ein großer Vorteil, dBASE IIIplus auch als Programmiersprache verwenden und so über die Menü-Steuerung zur systematischen Programmierung gelangen.(7) Im Menü (Abb. 3), oberste Zeile, finden sich bestimmte Oberbegriffe (es existiert auch eine deutsche Fassung des dBASE-Programms):

Set up - Create - Update - Position - Retrieve - Organize - Modify - Tools

- Während *Set up* schon bestehende Dateien lädt, kann man mit *Create* eine neue Datei mit einer Datenbankstruktur eröffnen. Insgesamt 128 Felder sind erlaubt, d.h. 128 Variablen wie Fundnummer, Keramiktyp etc. können numerisch oder alphanumerisch eingegeben werden (Abb. 1). Auf dem Bildschirm sind während des Eingabevorganges immer die wichtigsten Funktionen (z.B. Löschen, Einfügen) sowie bestimmte Informationen (Dateiname, Feldnummer, Laufwerk) sichtbar. Nachdem man die Datenbankstruktur so festgelegt hat, lassen sich sofort die ersten Daten eingeben. Weiter können mit *Create* andere dBASE-Dateitypen gebildet werden, so z.B. individuelle Daten- oder Druckformate.

- Mit *Update* ist man in der Lage, seine bereits eingegebenen Daten zu aktualisieren oder in bestimmter Weise anzeigen zu lassen (Abb. 2).

- Das Menüfeld *Position* erlaubt u.a. ein gezieltes Suchen nach Begriffen oder der Recordnummer, d.h. der laufenden Nummer der Eintragung.

- Mit *Retrieve* lassen sich kleinere Analysen durchführen, so können u.a. für einen gegebenen Variablenwert (z.B. Schichtlage) die Werte der anderen Variablen wie Keramiktypen etc. ermittelt und ausgedruckt werden (Abb. 4).

- Außerdem läßt sich mit dem Unterbefehl *Report*, der eine unter *Create* erstellte Report-Datei aufruft, eine Druckausgabe erzielen.

- Das Menüfeld *Organize* enthält Befehle, mit denen man seine Daten nach bestimmten Gesichtspunkten (etwa eine Literaturliste nach Namen oder Sachgebieten) ordnen oder kopieren und daraus eine neue Datenbank entstehen lassen kann.

Database file	DIERSFDT.DBF
Format for Screen Query	KALBECK.DBF
Catalog View	LS1.DBF
Quit dBASE III PLUS	RHEINBER.DBF
	SORT.DBF
	SORTGEO.DBF
	SORTNUM.DBF
	SORTZEIT.DBF
	ALLWED.DBF
	WEDAUII.DBF
	WEDAUIII.DBF
	WEDAUIV.DBF
	WEDNUMM.DBF
	WISSING.DBF
	WREST.DBF
	BARSORT.DBF
	SASKART.DBF

Command: USE C:  
 ASSIST ||<C:>||WEDAUI ||Rec: 1/663 ||  
 . Select - \_ .  
 Select a database file.

Abb. 3 dBASE IIIplus Hauptmenü.

List	Execute the command Specify scope Construct a field list Build a search condition Build a scope condition
Display	
Report	
Label	
Sum	No more conditions Combine with .AND. Combine with .OR.
Average	
Count	

Command: LIST STRATA\_BOD, TYPOLOG FOR STRATA\_BOD = '02'  
 ASSIST ||<C:>||WEDAUI ||Rec: 1/663 ||  
 . Select - \_ .  
 Select a logical operator for the FOR clause.

Record#	STRATA_BOD	TYPOLOG
8	02	02o0C
9	02	07N0D
10	02	00000
11	02	00000
12	02	00000
13	02	01o4A
14	02	07N1A
15	02	00000
16	02	00000
17	02	19000
18	02	19000
19	02	19000
20	02	00000
21	02	00000

ASSIST ||<C:>||WEDAUI ||Rec: 1/663 ||  
 Press any key to continue work in ASSIST.

Abb. 4 Retrieve-Menü dBASE IIIplus (Abfrage in Command-Zeile) mit Datenausgabe.

- *Modify* erlaubt u.a. Änderungen der ursprünglichen, mit *Create* erstellten Datenbankstruktur. So kann man Felder bzw. Variablen nachtragen, wenn man z.B. bei einer Grabungsanalyse das Feld Bodenart vergessen hat. Weiterhin lassen sich die übrigen dBASE-Dateiarten (*Report*, *Query*, *Format*) im nachhinein ändern.

- Das *Toolsfeld* enthält u.a. Funktionen, die auch im Betriebssystem enthalten sind, so daß man, um beispielsweise Dateien anzuzeigen oder zu löschen, das dBASE-Programm nicht zu verlassen braucht. Auch kann man sich dort die gesamte Datenbankstruktur anzeigen und ausdrucken lassen.

- Mit den Befehlen *Import* und *Export* lassen sich Daten sowohl aus anderen Software-Systemen nach dBASE, aber auch dBASE-Daten in andere Systeme übertragen.

Nachzutragen wären noch einige wichtige Einzel-Befehle, die im direkten Befehlsmodus gestartet werden können und nicht im Assist-Menü vorhanden sind.

- Die Befehle *append from* oder *join to* erlauben das Verknüpfen von Dateien.

- Der Befehl *copy to* mit seinen Erweiterungen (siehe auch Bemerkungen im Abschnitt "Statistik" weiter unten) wandelt dBASE-Dateien in diverse ASCII-Formate um, die wiederum von anderen Software-Systemen, z.B. Statistik- oder Graphikpaketen, gelesen werden können.

Besonders interessant sind Anwendungen, mit deren Hilfe sich mehrere Datenbanken miteinander verbinden oder hierarchische Strukturen schaffen lassen, beispielsweise kann man bei einer Grabung eine Hauptdatenbank im Museum einrichten, die mittels kleinerer Datenbanken draußen vor Ort immer wieder aktualisiert wird.

Die neueste Version dBASE IV 1.1 - voll kompatibel zu dBASE IIIplus - besitzt, neben vielen Verbesserungen (z.B. 256 Felder pro Datensatz), auf die wir im folgenden noch eingehen werden, einen im Gegensatz zur älteren Version wesentlich veränderten Menüaufbau sowie geänderte Oberbegriffe.(8) Zu diesen im sog. "Kontrollzentrum" zu findenden Oberbegriffen, die wiederum - ähnlich dBASE IIIplus - Untermenüs bzw. Unterbefehle besitzen, kann man jeweils Dateien bilden oder dort bestehende Dateien aufrufen (Abb. 5).

```

Catalog  Tools  Exit                               5:59:29
                                     dBASE IV CONTROL CENTER
                                     CATALOG: C:\DBASE\UNTITLED.CAT

```

Data	Queries	Forms	Reports	Labels	Application
<create>	<create>	<create>	<create>	<create>	<create>
ADREMA ARCHKOR3 BARSORT4 DFUNDST DUISBURG KERAMIK STATIST1			LIT_PRT	ADREMA	

File:           New file  
Description: Press ENTER on <create> to create a new file

Help:F1 Use:— Data:F2 Design:Shift-F2 Quick Report:Shift-F9 Menus:F10

Abb. 5 dBASE IV Hauptmenü (Control Center).

Im Data-Menü können Dateien neu erstellt und/oder andere Dateien aufgerufen sowie neue Daten einer schon bestehenden Datei hinzugefügt werden. Diese Dateien werden in einem Katalog zusammengestellt. In dem Untermenü Tools befinden sich u.a. Befehle, um Daten zu im- oder exportieren, Hilfsprogramme, eine Datenschutzroutine (*Protect*). Außerdem können die Daten mitsamt den Variablenmasken in verschiedener Form (*Edit*, *Browse*) dargestellt oder verändert werden (z.B. *Append*).

Das Query-Menü beinhaltet die von IBM entwickelte Datenbankabfrageroutine QBE (Query By Example) und kann auf recht einfache Weise Daten unter den verschiedensten Aspekten suchen, aber auch verändern. Es wirkt ähnlich wie die dBASE IIIplus-Oberbefehle Retrieve und Organize. DBASE IV QBE besitzt zwei wichtige Eigenschaften. Zum einen können sog. "virtual files" bzw. "views" (Ad-hoc-Abfragen, die die originale Datenbankdatei nicht verändern, nur die Abfragedatei mit der Extension QRY wird gespeichert) gebildet, zum anderen Datenbankdateien verändert werden, z.B. indem man sie sortiert oder miteinander verknüpft.

Von den Ausgabe-orientierten Menüs *Forms*, *Reports* und *Labels* sei nur das *Reports*-Menü besonders erwähnt, da es für das Ausdrucken am wichtigsten ist und außerdem einige wesentliche Verbesserungen zu der Vorgänger-Version aufweist. Sieben Haupt-Befehle (*Layout*, *Fields*, *Bands*, *Words*, *Go to*, *Exit*) können im Reports-Menü angewählt werden. Mit ihnen kann man z.B. Felder aus der gegenwärtigen Datenbank holen, aber auch spezielle Felder eigens für den Druck bilden, längere Texte schreiben oder bestimmte graphische Hervorhebungen (Box, Linien etc.) vornehmen. Fünf sog. "report bands" (Bereiche, auf denen geschrieben oder Felder positioniert werden können) erleichtern anschaulich das Positionieren von Text oder bestimmten Feldern direkt auf dem Bildschirm (Abb. 6). Eine wesentliche Verbesserung zu dem Report-Befehl von dBASE IIIplus besteht in der Möglichkeit, auch einzelne Blätter eines "reports" auszudrucken sowie bestimmte Seitenformate und Fonts (Schrifttypen) selbst zu bestimmen.

Während der Menü-Befehl *Forms* für individuelle Daten-Eingabemasken zuständig ist, können mit dem Menü-Befehl *Labels* Etikettenformate erstellt werden.

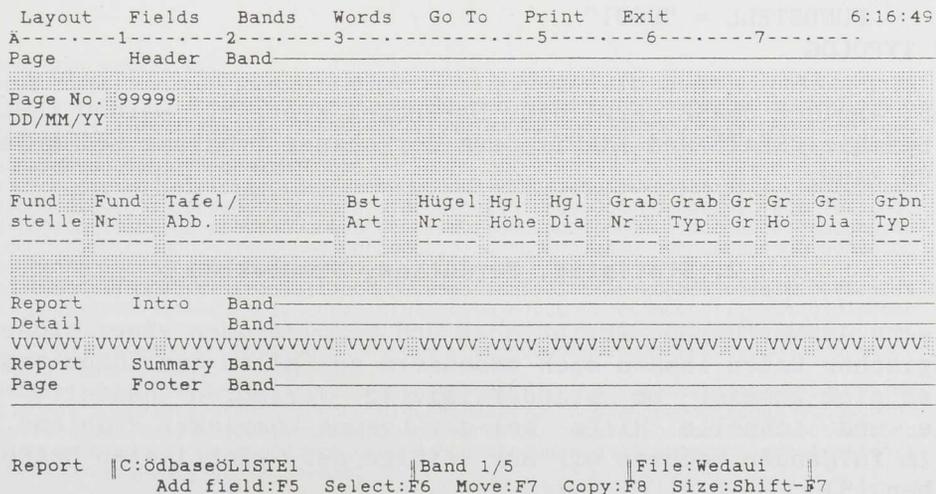


Abb. 6 dBASE IV Report-Menü.

---

Mit dem Menü-Befehl *Applications* ist man einerseits in der Lage, eine Abfolge von Menü-Befehlen sowie von eigenen Programmen festzulegen, die es ermöglicht, bestimmte, immer wiederkehrende Aufgaben - wie z.B. Datenbanken zu aktualisieren oder Ausdrucke oder Etiketten zu generieren - zusammenzufassen. Andererseits kann man, ähnlich den Vorgängerversionen zu dBASE, eigene Programme mit der Programmiersprache dBASE schreiben, die sich wiederum mit der Datenbank-Abfragesprache SQL (embedded SQL, s.u.) kombinieren lassen.

### Der dBASE-Compiler Clipper

Der allerdings nur für den Anwender mit Programmierkenntnissen interessante dBASE-Compiler "Clipper" (Nantucket GmbH, Köln) optimiert den Gebrauch von dBASE IIIplus in mehreren Bereichen. So wandelt er z.B. Programmdateien (Extension .PRG) in eine Form (Extension .EXE) um, in der sie direkt vom Betriebssystem, ohne die Verwendung der dBASE-Software, geladen werden können und durch deren Quellcode besser geschützt sind. Außerdem besitzt Clipper - neben vielen anderen Vorteilen - die Möglichkeit, 1024 Felder bzw. 2048 Speichervariablen zu eröffnen. Allerdings ist Clipper nur bis zur dBASE Version IIIplus aufwärts kompatibel, der erweiterte Befehlssatz von dBASE IV 1.1 ist noch nicht eingegliedert worden.

### SQL (Structured Query Language)

SQL ist eine von IBM in San Jose/Kalifornien seit dem Jahr 1974 immer weiterentwickelte Abfragesprache, die für entsprechende Datenbanken im Großrechnerbereich entworfen wurde (van der Lans 1987; Weber 1990). Viele Hersteller haben die Sprache SQL implementiert, somit existieren mehrere "Dialekte" (z.B. SQL/DS, Oracle, INGRES/SQL etc.), die sich z.T. erheblich voneinander unterscheiden können.

dBASE IV ist in der Lage, SQL-kompatibel zu arbeiten. Ein kleines SQL-Programm, das (unser dBASE-Datenbankbeispiel benutzend) die Keramiktypen und deren Verzierung für den Fundort Duisburg-Wedau bzw. die Fundstelle Grabung Bonnet 1895/96, Neudorf Hügel 1-7 (= "0001") heraussucht:

```
SELECT      TYPOLOG, VERZ
FROM        FUNDORT = "0001"
WHERE       FUNDSTELL = "0001"
ORDER BY    TYPOLOG
```

Zuvor muß, ähnlich dBASE, erst eine Datenbank erstellt werden (SQL verwendet die Befehle CREATE und INPUT), aus der obiges Programm seine Informationen holen kann.

## 2. Statistik, Seriation, Simulation

Statistische Auswertungen, Seriationen und Simulationen einer großen Anzahl archäologischer Daten lassen sich besonders gut Hilfe des Computers berechnen, da es sich zumeist um standardisierte Verfahren handelt, die eine wertvolle und schnelle Hilfe bei der Lösung komplexer Probleme bedeuten können. Im folgenden möchten wir nur einige der wichtigsten Methoden kurz beschreiben.(9)

Bei der statistischen Auswertung können einfache Methoden, aber auch komplexere mathematische Verfahren angewandt werden:



---

## Datenmodifikation

Sowohl SPSS als auch SAS enthalten eine Reihe von Befehlen zur Änderung und Erweiterung der eingelesenen Daten. Dafür einige Beispiele (SPSS):

- *compute* (Berechnen neuer Variablen)
- *recode* (Umcodierung von Werten in eine neue Variable)
- *missing value* (Definition fehlender Werte)
- *select if* (Einschränkung von bestimmten Fällen)
- *aggregate* (Zusammenfassen von Variablenwerten in neue Variablen)
- *weight by* (Festlegung einer Variablen, die für jede Beobachtung den entsprechenden Gewichtungsfaktor enthält)

Des Weiteren sind iterative Anweisungen ("Makros", d.h. kleine Programme), temporäre Datenmodifikationen, die Anwendung mathematischer Funktionen usw., insgesamt also das Schreiben regelrechter Auswertungsprogramme (Prozeduren) möglich.

## Statistik-Prozeduren

Für jede einzelne Statistikprozedur in SPSS oder SAS müssen kleine Programmdateien erstellt werden. Einfache Analysen wie die SPSS-Prozeduren Condescriptive (Mittelwerte, Standardabweichung, Minimum/Maximum), Frequencies (Häufigkeitsverteilungen, Histogramme), Crosstabs (Kombinationstabellen bzw. Kontingenztafeln) sowie Plots (Diagramme) lassen sich leicht durchführen.

Die in der Demo-Diskette befindlichen kleinen SPSS-Beispielprogramme (CONDES.TXT, CROSS.TXT, PLOT.TXT, FREQ.TXT) enthalten jeweils zwei Blöcke, die aus den Befehlsschlüsselwörtern und deren Parametern bestehen.

Die Befehle gliedern sich in drei Abschnitte:

- **Datendefinition:** Der Befehl TITLE stellt die Überschrift dar, FILE HANDLE bezieht sich auf die zu untersuchende Datendatei (bei Überlänge muß die Record-Länge, in unserem Fall 351 Zeichen, angegeben werden), DATA LIST beschreibt die einzulesenden Variablen, die analysiert werden sollen und gibt deren Position an, VARIABLE LABELS und VALUE LABELS beschreibt kodierte oder abgekürzte Begriffe ausführlicher, MISSING VALUES legt diejenigen Werte fest, die nicht in der Untersuchung berücksichtigt werden sollen.
- **Datenmodifikation:** SELECT IF beschränkt die Auswahl der Fälle einer Variablen.
- **Datenanalyse:** CONDESCRIPTIVE, CROSSTABS, PLOT, FREQUENCIES.

Komplexere Analysen wie die Varianz-, Regressions-, Faktoren-, Cluster- oder Diskriminanz-Analyse erfordern ein höheres mathematisches Niveau und somit eine gründliche Durchdringung der verschiedenen Methoden und Ansätze. Im folgenden seien nur einige allgemeine grundlegende Informationen über diese Analyseverfahren angeführt. (10)

Die Varianzanalyse (Demo-Diskette: VARIANZ.TXT) prüft die Wirkung einer oder mehrerer unabhängiger Variablen (mindestens mit Nominalskalierung) auf eine oder mehrere abhängige Variablen (nur metrisches Skalenniveau erlaubt). In unserem einfachen Beispiel wird der Einfluß bestimmter Gefäßty-

---

pen (d.h. mehrere unabhängige Variablen) auf die Gefäßhöhe (d.h. eine abhängige Variable) untersucht. Die dementsprechende SPSS-Prozedur heißt ONEWAY. Rechnerisch wird dabei die Gesamtvarianz (bestehend aus der Varianz zwischen den Gefäßtypen und innerhalb der Gefäßtypen) zugrunde gelegt. Die weiteren Rechenschritte bestehen aus der Summe der Abweichungsquadrate, den Freiheitsgraden, der mittleren Quadratsumme sowie dem F-Test. Diverse Tests (z.B. Duncan, Scheffe, Student-Newman-Keuls) sind möglich.

Wichtig ist der Vergleich der Werte der empirischen und theoretischen F-Verteilung. Keine Wirkung der unabhängigen Variablen (Gefäßtypen) ist dann zu vermuten, wenn der empirische F-Wert kleiner als der tabellierte Testwert ist. In unserem Beispiel ist der empirische Wert sehr hoch und bestätigt, ebenso wie die Test-Prozeduren (siehe Graphiken mit Asterisk), die Bedeutung der Höhe in bezug auf die Gefäßtypen. Die Testgraphik bestätigt die Typenscheidung bezüglich der Höhe in Hochformen (Typen 1-4) und Niedrigformen (Typen 6-8).

Will man den Einfluß mehrerer unabhängiger Variablen auf mehrere abhängige Variablen testen, dann wendet man am besten die SPSS-Prozeduren ANOVA oder MANOVA an. Erstere ist imstande, auch Einweg-Varianzanalysen wie ONEWAY zu leisten (allerdings können mit ONEWAY zusätzliche Tests durchgeführt werden).

Die Regressionssanalyse (Demo-Diskette: REGRESS.TXT) prüft, ähnlich der Varianzanalyse, die Beziehungen zwischen einer abhängigen und einer bzw. mehreren unabhängigen Variablen. Im Gegensatz zur Varianzanalyse gilt sie allerdings nur für metrische Variablen. So kann man beispielsweise den Einfluß bestimmter Werte (Randdurchmesser etc. als unabhängige Variablen) auf die Höhe eines Gefäßtyps (abhängige Variable) untersuchen.

Während die einfache Regression die Wirkung einer unabhängigen auf eine abhängige Variable untersucht, prüft die sog. "multiple" Regression (SPSS-Prozedur REGRESSION) die Wirkung mehrerer unabhängiger Variablen auf eine abhängige Variable. Die lineare Regressionsanalyse der SPSS-Prozedur REGRESSION unterstellt dabei eine lineare Beziehung zwischen Regressand (abhängige Variable) und Regressor (unabhängige Variable). Graphisch stellt sich dieser Bezug als eine Gerade dar, die sich der empirischen Punkteverteilung möglichst gut anpaßt. Anhand der Abweichungen von dieser Geraden lassen sich nun Voraussagen über die Beziehungen der einzelnen Variablen treffen.

Das Bestimmtheitsmaß  $r^2$ , das sich aus der erklärten Abweichung und der Gesamtabweichung ergibt, ist um so größer, je höher der Anteil der erklärten Abweichung (hier der unabhängigen Variablen "Gefäßhöhe") an der Gesamtabweichung ist. So bedeutet in unserem Beispiel der  $r^2$ -Wert .60117 (zu Bodendurchmesser), daß etwa 60 % Einfluß auf die Variable "Bod" zurückzuführen sind.

Korrelation/Faktorenanalyse (Demo-Diskette: FAKTOR.TXT): Allgemeines Ziel der Faktorenanalyse ist die Reduktion der in den Daten enthaltenen Information auf möglichst wenige Komponenten, ohne daß damit ein zu starker Informationsverlust einhergeht. Konkret versucht man, die zahlreichen im ursprünglichen Datenmaterial erhobenen Variablen zu wenigen neuen "Faktoren" rechnerisch zu kombinieren, wobei diese Kombination die ursprüngliche Datenstruktur möglichst getreu widerspiegelt. Sofern sich solche Faktoren bestimmen lassen, kann man versuchen, sie anschließend zu interpretieren, wenn man ihre "Ladungen" betrachtet, d.h. die Beiträge der einzelnen Ausgangsvariablen zum betreffenden Faktor.

---

Die Programme zur praktischen Durchführung von Faktorenanalysen finden sich ebenfalls in den einschlägigen Statistikpaketen (z.B. SPSS: FACTOR; SAS: FACTOR, PRINCOMP). In unserer Beispieldatei FAKTOR.TXT werden die metrischen Werte der Gefäßtypen in bezug auf die Fundorte miteinander verglichen.

Ziel der Clusteranalyse (Demo-Diskette: ROHDAT.TXT, CLUST1.TXT, CLUST2.TXT) ist es, "Objekte" mittels numerischer Verfahren zu klassifizieren. Diejenigen Objekte (z.B. Gefäßtypen, Fundstellen, Kulturen), die bezüglich ihrer "Merkmale" (Vorkommen/Fehlen bzw. quantitative Ausprägung z.B. von Größen- und Formmaßen, Inventargegenständen, Begräbnisformen) die größte Ähnlichkeit aufweisen, werden zu Gruppen zusammengefaßt. Die einander ähnlichsten Gruppen werden wiederum in "Übergruppen" vereint usw., so daß eine Hierarchie nach Art eines Stammbaumes entsteht. (Die numerischen Klassifikationsverfahren wurden tatsächlich vorwiegend im Bereich der biologischen Statistik entwickelt.)

Alternativ zu dem gerade umrissenen "agglomerativen" und "hierarchischen" Verfahren gibt es aber auch divisive Methoden (ausgehend von der Gesamtzahl aller Objekte werden Gruppen durch Unterteilung gebildet) und nicht-hierarchische Cluster-Bildung (bei der z.B. ein Objekt mehreren Gruppen zugleich angehören kann).

Strenggenommen gehören Clustertechniken nicht zur (schließenden) Statistik im engeren Sinne: es geht nicht um den Schluß von Stichproben auf Parameter einer Grundgesamtheit, sondern darum, die Objekte nach einem festgelegten (und reproduzierbaren) Verfahren zu klassifizieren, im Grunde also nichts anderes als eine - wenn auch streng geordnete - Art der Datendarstellung. Die entsprechenden Techniken wurden ursprünglich größtenteils empirisch entwickelt und erprobt. In den letzten Jahren bauten zahlreiche Studien aber auch die theoretischen Grundlagen der Clusteranalyse stark aus, wobei nunmehr auch echt statistische Aspekte einbezogen wurden (z.B. Tests auf optimale Zahl von Gruppen).

Die vielfältigen Methoden der Bestimmung von Clustern unterscheiden sich zum einen darin, welches Maß sie für die "Ähnlichkeit" (bzw. Unterschiedlichkeit) zweier Objekte anhand ihrer Merkmale verwenden. Es können sowohl binäre (vorhanden/fehlend) als auch quantitative Merkmale herangezogen werden. Jedes Merkmal (z.B. Durchmesser, Größe, Vorhandensein von Henkeln) bildet eine Achse im vieldimensionalen Merkmalsraum; jedem Objekt kann dann, je nach konkreter Ausprägung der Merkmale, eine Position in diesem Raum zugewiesen werden. Ähnlichkeit läßt sich dann als Entfernung zwischen den Objekten auffassen. Außerdem muß festgelegt werden, wie die Ähnlichkeit zwischen Gruppen von Objekten zu bestimmen ist, z.B. als Abstand zwischen den beiden am weitesten voneinander entfernten oder aber den am dichtesten benachbarten Objekten der beiden Gruppen, oder als Durchschnitt aller paarweisen Entfernungen.

Die typische Schwierigkeit bei der Anwendung der Clusteranalyse besteht in der Wahl geeigneter Ähnlichkeitsmaße und Clusterbildungsverfahren. Es gibt - wegen des empirischen Charakters vieler Techniken - hierfür weniger A-priori-Leitlinien als etwa bei der Wahl statistischer Tests, die weitgehend durch die Fragestellung festgelegt ist. Dazu kommt, daß die Resultate einer Clusteranalyse oft stark von den gewählten Verfahren abhängen und verschiedene Techniken also zu durchaus differierenden Resultaten führen können. Ein gründliches Verständnis der vielen angebotenen Methoden, ihrer Möglichkeiten und vor allem ihrer Grenzen, sollte also unbedingt vorhanden sein, bevor man die Clusteranalyse als Forschungsinstrument einsetzt.

---

Die benötigten Programme oder Prozeduren werden im Rahmen mehrerer Statistikpakete angeboten, so z.B. als Prozedur CLUSTER in SPSS und FASTCLUS in SAS. Das wohl umfangreichste, nur der Clusteranalyse gewidmete Paket dürfte CLUSTAN sein (Autor D. Wishart, University of Edinburgh). Das Bonner Seriationspaket (s.u.) enthält ebenfalls Cluster-Algorithmen.

Da unsere Funddatei pro Fund einen Record aufweist, muß für die Clusteranalyse, die die Funde einer Einheit (Grab) untersucht, die Datenstruktur so umgeschrieben werden, daß die Funde dieser Einheit in einem Record vereint sind (ROHDAT.TXT). In unseren beiden Beispieldateien (CLUST1.TXT = SAS, CLUST2.TXT = SPSS-X) werden die Fundorte nach Anzahl der Gefäßtypen "geclustert".

Die Diskriminanzanalyse (Demo-Diskette: DISCRIM.TXT) dient der Zuordnung von Objekten zu möglichst scharf voneinander abgegrenzten Gruppen. Auch hier wird versucht, die ursprünglich erfaßten Variablen zu (wenigen) neuen Funktionen zu kombinieren. Anhand dieser Funktionen kann für jedes Objekt ein Wert (score) berechnet werden, der es erlaubt, das Objekt einer von mehreren Gruppen zuzuordnen. Die Funktionen sollten dabei so beschaffen sein, daß die gefundenen Objekt-Gruppen möglichst deutlich kontrastieren.

Eine typische archäologische Anwendung wäre z.B. die folgende: Es liege eine Anzahl von Keramikgefäßen vor, die einer früheren Kultur A oder einer späteren Kultur B angehören oder aber unklare Zwischenformen aufweisen. An den Gefäßen wird eine Anzahl von Maßen erhoben. Mittels der Diskriminanzanalyse könnte nun aus diesen Maßen eine Funktion berechnet werden, die es erlaubt, die "Übergangsformen" mit größtmöglicher (und angebarerer!) statistischer Sicherheit entweder der Kultur A oder der Kultur B zuzuordnen. Zu beachten ist, daß immer ein Teil der vorhandenen Daten (bekannter Gruppenzugehörigkeit!) zur Ermittlung der Diskriminanzfunktion benötigt wird; diese können dann auf die verbliebenen Daten angewandt werden.

Auch die Diskriminanzanalyse gehört heute zum Standardrepertoire der meisten Statistikpakete (SPSS: DISCRIMINANT; SAS: DISCRIM, NEIGHBOR, CANDISC, STEPDISC). In unserer Beispieldatei (DISCRIM.TXT) werden die Fundorte und die Gefäßtypen mit ihren Meßwerten jeweils auf Gruppenzugehörigkeit überprüft, wobei Testdaten (Gruppe unbekannt) und Analysedaten (Gruppe bekannt) untersucht werden (Beispiele 1-4).

### Vergleich der Pakete SAS und SPSS

Hinsichtlich des Angebotes an statistischen Prozeduren gibt es für den "Normalanwender" kaum Unterschiede zwischen SAS und SPSS: beide Pakete enthalten alle elementaren sowie eine große Vielzahl komplexer und spezieller Statistikverfahren. Unterschiede bestehen eher in der Art der Benutzung. In beiden Fällen ist natürlich ein einfacher Aufruf von Prozeduren und deren Steuerung durch zusätzliche Parameter möglich. Beide Programme bieten auch die Möglichkeit zur Modifikation von Daten (Berechnung neuer Variablen aus den vorhandenen, Umcodierung von Datenwerten usw.) und zur Auswahl bestimmter Fälle bzw. Beobachtungen. Bei SAS besteht hierfür jedoch eine wesentlich größere Flexibilität, da regelrechte Programme in einer PL/1-ähnlichen Sprache geschrieben werden können. Dem programmiererfahrenen Anwender wird damit die Möglichkeit gegeben, seine statistische Auswertung in praktisch beliebig komplexer, aber dennoch übersichtlicher Weise auf seine speziellen Bedürfnisse zuzuschneiden. Darüber hinaus bietet SAS ein gegenüber SPSS wesentlich umfangreicheres und leistungsfähigeres Graphik-"Unterpaket" (SAS/GRAPH); die Qualität ist so gut, daß manche Anwender SAS als reine Graphiksoftware einsetzen, ohne seine statistischen Prozeduren zu nutzen.(11)

---

Auch für einige andere, speziellere Bereiche der Statistik, z.B. Zeitreihenanalyse (SAS/ETS), Matrixmanipulation (SAS/IML), Operations Research (SAS/OR) usw., bietet SAS Prozedursammlungen an. Diese umfassende Leistungsfähigkeit bedingt andererseits (neben hohen Lizenzkosten) einen erhöhten Lernaufwand seitens des Anwenders, der sich erst seinen Weg durch das große Angebot bahnen muß.

Für Standard-Auswertungen ohne spezielle und komplizierte "Sonderwünsche" wird daher nach wie vor auch der Einsatz des weniger flexiblen, aber dafür leichter zu bedienenden SPSS zu erwägen sein; auch dessen "Performance" ist vom "Normalbenutzer" kaum je zu überschreiten. Die relativ starre Syntax früherer SPSS-Versionen wurde mit der Version SPSS-X durch eine anschaulichere und flexiblere Kommandosprache ersetzt. Auch bei SPSS ist jetzt ein Dialogbetrieb möglich, der in seinen Grundzügen der sehr komfortablen interaktiven SAS-Benutzeroberfläche entspricht. Beide Pakete weisen also Vorzüge und Nachteile auf; die Entscheidung darüber, welches man einsetzen sollte, hängt - wie immer im EDV-Bereich - von den spezifischen Bedürfnissen des Anwenders ab.

### The Bonn Seriation and Archaeological Statistics Package

Das Bonner Programmpaket ist nur für den PC-Gebrauch konzipiert worden.(12) Es enthält in seiner aktuellen Version 4.1 folgende Programme:

- Seriation
- Simulation
- Clusteranalysis
- Social Status Analysis
- Harris Matrix Analysis

Die einzelnen Programmdateien lassen sich übersichtlich menügesteuert starten. Konkrete Anwendungsbeispiele für jede Prozedur sowie die Help-Funktion (über Taste F1) machen das gesamte Paket sehr benutzerfreundlich. Die Form der Dateneingabe ist festgelegt, Fremdformate (dBASE, SPSS) können eingelesen werden, wenn sie dem Aufbau der Dateneingabe entsprechen. Die mathematischen Grundlagen der wichtigsten Algorithmen aller Hauptprogramme werden in einem gesonderten Kapitel des Handbuchs ausführlich behandelt. Hilfsprogramme, z.B. zur vergleichenden Betrachtung zweier Listen (browse) oder der Einrichtung einer "RAM-Disk" (Pseudo-Festplatte im Hauptspeicher mit schnellem Zugriff), werden bereitgestellt. Wertvolle Literaturhinweise vervollständigen das Programmpaket.

Seriation (13) (Abb. 8): Nach Korbel (1981) verstehen wir unter dem Begriff Seriation "... alle Verfahren, mit denen verschiedene Objekte so einander zugeordnet werden können, daß Serien entstehen, die ihr relatives Alter widerspiegeln." Allerdings sind auch andere Interpretationen als chronologische (z.B. Geschlechtsgebundenheit von Typen) möglich.

Seriationstechniken haben den Zweck, Individuen oder Gruppen von Individuen in einer Dimension oder entlang einer Achse anzuordnen. Je nach Genauigkeit der Typendefinition bzw. der Langlebigkeit von Typen fällt die Qualität der Seriation aus. Auch sind bei der Aufbereitung der Seriationsdaten Gewichtungen der chronologischen Signifikanz von Typen möglich, die allerdings umstritten sind (Scollar et al. 1990, 180). Ausgangspunkt ist, wie bei der Clusteranalyse, eine rechteckige Datenmatrix. Meist stehen die größeren Einheiten (z.B. Fundstellen, Gräber) in den Zeilen, die Variablen (z.B. Typen, Merkmale, Eigenschaften) in den Spalten.



---

Anhand statistischer Verfahren (z.B. Kammerer-Goldmann-Verfahren) werden Objekte aus diversen Einheiten (Gräber, Siedlungsgruben etc.), die als geschlossene Funde zu behandeln sind, nach Ähnlichkeit der Zusammensetzung und räumlich möglichst nahe an einer Diagonalen geordnet. Seriiert werden kann aufgrund des Vorhandenseins/Nicht-Vorhandenseins (Presence/Absence Seriation; P-Matrix, benannt nach dem "Erfinder" dieser Technik, Sir Flinders Petrie; nach Wilkinson wird die P-Matrix "ideal" an der Diagonalen angeordnet) oder der prozentualen Verteilung von Typen (Abundance Seriation; die unimodale Q-/R-Matrix, wobei "Q" gewählt wurde, weil dieser der nächste Buchstabe zu "P" ist, "R" steht für den Amerikaner Robinson). Man kann die Merkmale in bezug auf die größeren Einheiten (Q-Technik) oder umgekehrt die größeren Einheiten in bezug auf die Merkmale (R-Technik) untersuchen.

Das Bonner Paket benutzt als Ordnungsverfahren das sog. "Reciprocal Averaging" (RA, alias "analysis factorielle des correspondences", alias "correspondence analysis"). Es handelt sich dabei um eine Methode zum Ordnen von Tabellen, in denen Fälle (z.B. Fundorte) als Zeilen und Merkmale (z.B. die Anzahlen von Objekten verschiedener Keramiktypen) als Spalten aufgeführt sind. Das Ziel des Verfahrens (in der Archäologie auch "Seriation" genannt) besteht unter anderem darin, die Zeilen und Spalten der Tabelle so umzuordnen, daß die stark besetzten Zellen möglichst nahe der Hauptdiagonalen stehen; Fundorte (Zeilen) mit gleichem Fundinventar und Merkmale (Spalten) mit ähnlichem Fundortspektrum sollen zusammengruppiert werden. Der benutzte Algorithmus läßt sich im Prinzip leicht verstehen und nachvollziehen:

Zunächst wird allen Merkmalen (Fundtypen, Spalten) ein beliebiger Gewichtungsfaktor (z.B. laufende Nummer) zugeordnet. Für jeden Fall wird dann ein Punktwert (Score) berechnet, nämlich der gewichtete Mittelwert (über alle Spalten) aus den an diesem Fundort vorkommenden Objekttypen, unter Benutzung des oben genannten Gewichtungsfaktors. Diese Prozedur wird dann für die Merkmale (Spalten) wiederholt: man berechnet den gewichteten Mittelwert über alle Fundorte (Zeilen). Die Gewichte sind jetzt die zuvor erhaltenen Fundort-Scores; so erhält man Scores für alle Objekte. Anhand dieser werden nun wieder neue Fundort-Scores ermittelt usw. Die jeweils neuen Scores werden bei jedem Schritt skaliert (z.B. auf den Bereich 0 bis 100).

Es läßt sich zeigen (Hill 1973), daß nach einer endlichen (und meist recht geringen) Zahl von Iterationen die Fundort- wie auch die Objekt-Scores auf einen stabilen Wert konvergieren, der unabhängig von den willkürlich bestimmten Anfangswerten ist. Ordnet man die Zeilen und Spalten der Tabelle nun in der Reihenfolge der entsprechenden Scores, so ergibt sich die gewünschte oben beschriebene Anordnung.

In der Praxis verfährt man nicht immer nach dem beschriebenen Iterationsverfahren (Scollar et al. 1990), sondern macht sich zunutze, daß RA sich auch als Eigenwertproblem auffassen läßt, mit naher Verwandtschaft etwa zur Hauptkomponentenanalyse (so lassen sich auch mehrdimensionale Seriationen durchführen, wobei allerdings das Auftreten bestimmter Artefakte zu beachten ist, vgl. Hill u. Gauch 1980). In entsprechenden Rechnerprogrammen werden daher effiziente matrizenalgebraische Techniken eingesetzt, die schnell zum Ergebnis führen.

Bezüglich der konkreten Benutzung des Bonner Paketes läßt sich folgendes sagen: Das Format der Programmausgabe der Seriation (Prozeduren Getprtf, Prep, Print, Compare), die im übrigen auch umgekehrt werden kann (Prozedur Reverse), ist unter Angabe gewisser Parameter steuerbar. Vergleiche bezüglich der Unterschiede verschiedener Seriationsläufe (Cmprint, Compact, Compare) sind möglich. Die Eingabedaten können mit den Befehlen Split (unterteilt Typen) und Merge (faßt Typen zusammen) nachträglich bearbeitet werden.

---

Simulation: Das Programm Sersim simuliert ein "idealtypisches" Gräberfeld oder einen archäologischen Komplex, deren Analysen mit denen eines realen Gräberfeldes bzw. Komplexes verglichen werden können. Parameter des Vergleiches sind die Anzahl der Typen und Gräber, die Benutzungsdauer des Gräberfeldes, die durchschnittliche Verwendungszeit der Typen etc.

Clusteranalysen (Abb. 9) versuchen, wie oben bereits erwähnt, anhand von Ähnlichkeitskoeffizienten (Chi-Quadrat, euklidische Distanz u.v.m.) bestimmte Merkmale, Typen usw. in vorher unbekannte Kategorien zu fassen. Ausgangspunkt ist eine rechteckige Datenmatrix. Das Bonner Paket kann Clusteranalysen (Prozedur CA) für absolute oder relative Häufigkeitsdaten (abundance data) analysieren. Möglichkeiten zu einer Untersuchung binärer Datenmatrizen (presence/absence data) sind in der neuen Version 4.1 ebenfalls gegeben. Mit Hilfe der neuen Prozedur "MAP CLUSTERS" kann man einen Plan mit der räumlichen Verteilung der Cluster erstellen.

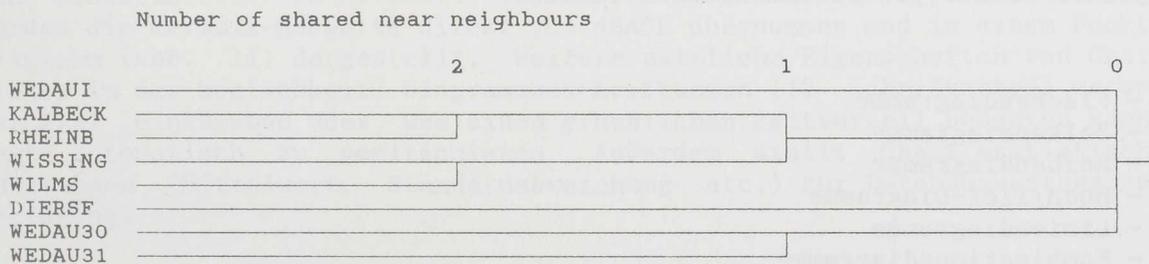
Incidence Statistic: Mit diesem Programm werden Histogramme erzeugt, die z.B. die Verteilung der Anzahl von Grabfunden oder die Anzahl von Gräbern, die bestimmte Typen enthalten, anzeigen. Spezielle Typen können bestimmt und im Histogramm angezeigt werden.

Social Status Analysis: Dieses Programm sortiert die soziale Stellung von Typen oder Gräbern nach quantitativen und/oder qualitativen Gesichtspunkten und stellt die Ergebnisse in drei Histogrammen (type index distribution, highest status type index, sum of status indices of types) dar. Bei der qualitativen Analyse können die Typen gewichtet werden.

Harris Matrix Analysis: Ebenfalls neu ist die Harris Matrix Analysis, die stratigrafische Schichtzuweisungen korrelieren, modifizieren, untersuchen und in einem Diagramm darstellen kann.

Near Neighbour Clustering of GRABCLUSTER

Distance Measure: Euclidian Distance  
 Number of Neighbours considered: 2



Level	#clusters	mean # el.	#el. in residue
1	1	4.0	4
2	0	0.0	8

Abb. 9 Cluster-Dendrogramm, Gräberfelder Niederrhein (The Bonner Seriation and Archaeological Statistics Package - Scollar et al. 1990).

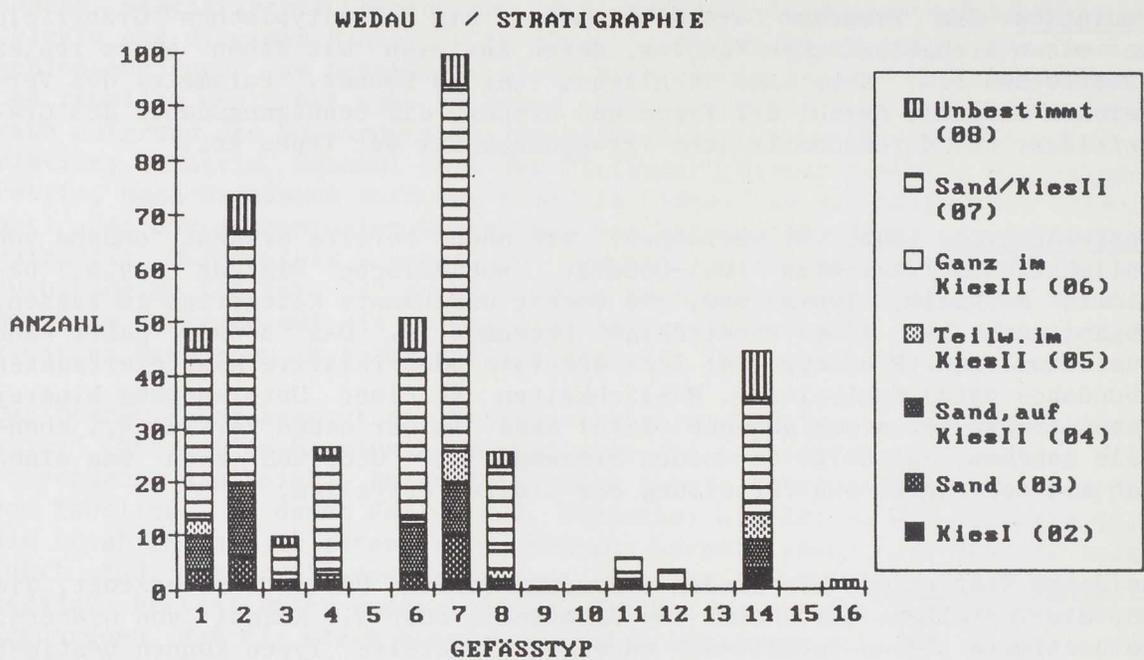


Abb. 10 Chart Säulendiagramm (aus: Schoenfelder 1989, Abb. 98).

### 3. Graphik

Sowohl SPSS, SAS als auch das Bonner Seriations- und Statistikpaket enthalten zahlreiche Möglichkeiten der graphischen Darstellung. Es sind dies vor allem Seriationstabellen, Kreuztabellen, Histogramme, Dendrogramme, Kreuzdiagramme. Will man jedoch seine Daten ausschließlich in Diagrammform darstellen, sind sog. Business-Graphik-Programme empfehlenswert. Eines der am meisten verbreiteten Programme wollen wir - quasi stellvertretend für viele ähnliche Systeme (z.B. Harvard Graphics) - hier kurz vorstellen.

Das Graphikpaket Chart wurde von der Software-Firma Microsoft entwickelt. Es enthält 49 mögliche Darstellungsweisen von Diagrammen, die sich zu folgenden Grundtypen zusammenfassen lassen:

- Flächendiagramme
- Balkendiagramme
- Reihendiagramme
- Hoch/Tief-Diagramme
- Liniendiagramme
- Kombinationsdiagramme
- Kreisdiagramme
- Punktdiagramme
- Spannweitendiagramme
- 3-D-Diagramme

Chart besitzt eine übersichtliche Menüsteuerung (ähnlich dem Textverarbeitungsprogramm Word), die vom Anwender recht einfach zu handhaben ist. (14). Die Daten (maximal 128 Werte pro Graphik) gibt man entweder selbst ein oder übernimmt sie von einer Datenbank, in unserem Fall von dBASE IIIplus, das von Chart ohne weitere Umwandlung akzeptiert wird.

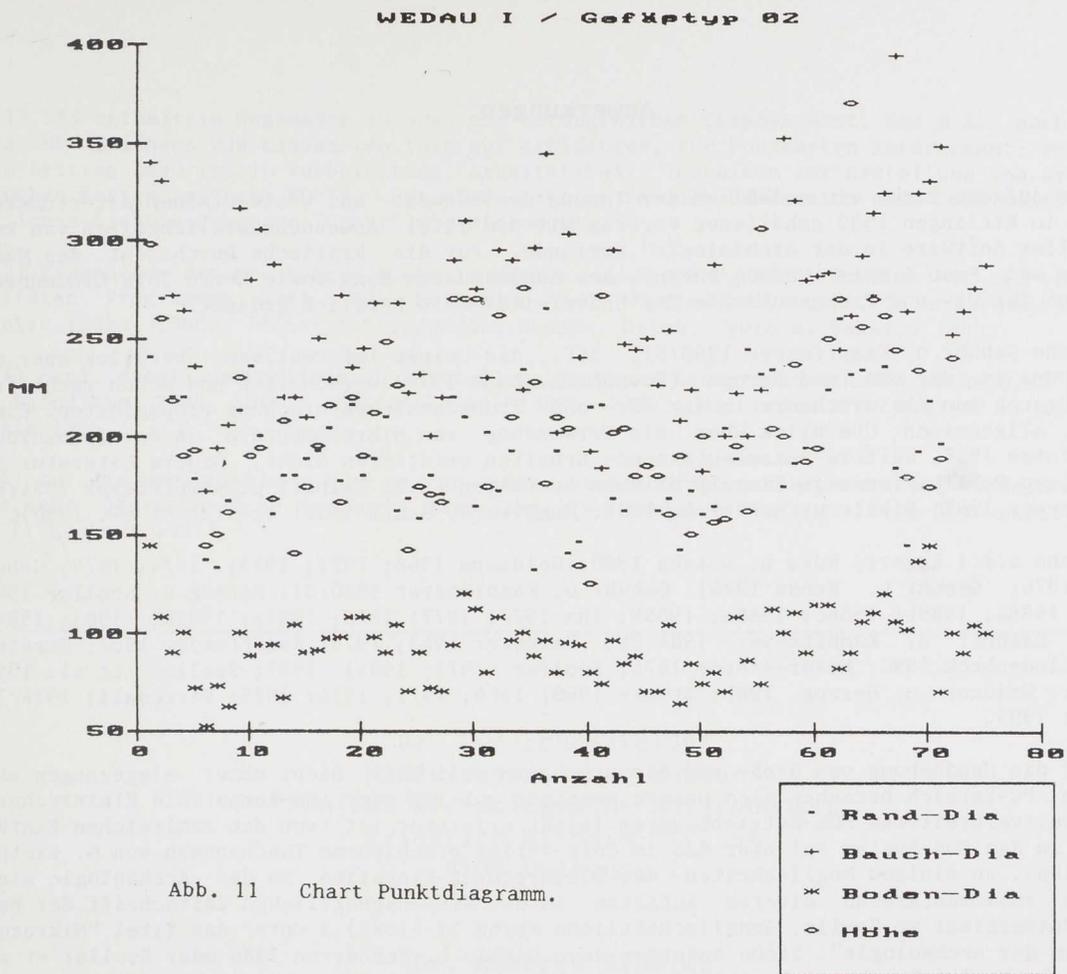


Abb. 11 Chart Punktdiagramm.

Die Abbildungen 10 und 11 geben zwei Beispiele der übersichtlichen graphischen Darstellung archäologischer Daten. Gerade statistische Tabellen, die aufgrund ihres Zahlengewirrs nicht sehr übersichtlich sind, lassen sich mit Hilfe z.B. eines Balkendiagramms (Abb. 10) sehr anschaulich darstellen. In unserem Fall mußten die Tabellenwerte per Hand eingetippt werden, weil sie von der SPSS-Kreuztabelle der Großrechnerversion SPSS-X nicht automatisch übernommen werden konnten (allerdings gibt es, wie oben bereits erwähnt, eine Schnittstelle zu MS-Chart für die Version SPSS-PC+). Im zweiten Fall wurden die Keramik-Meßwerte direkt von dBASE übernommen und in einem Punktdiagramm (Abb. 11) dargestellt. Weitere nützliche Eigenschaften von Chart liegen in der Möglichkeit, Diagrammbeschriftungen (12 Schriftarten) entweder frei einzugeben oder, was einen erheblichen Zeitvorteil bedeuten kann, Texte automatisch zu positionieren. Außerdem stellt Chart statistische Funktionen (Mittelwert, Standardabweichung etc.) zur Datenauswertung zur Verfügung.

Als Ausgabemedium möchten wir einen grafikfähigen Laserdrucker empfehlen, da bei komplizierter Graphik Einzelheiten mit den üblichen Nadeldruckern schwer darstellbar sind.

\* \* \*

Anm. d. Red.: Die im Text erwähnte Demo-Diskette mit Anwendungsbeispielen kann bei der Redaktion angefordert werden. Auf ihr befindet sich auch unter dem Dateinamen SOFTLIT.TXT die hier nur als Kurzzitat genannte Literatur.

---

## Anmerkungen

(1) Dem Aufsatz liegt ein anläßlich der Tagung des südwest- und westdeutschen Altertumsverbandes in Ettlingen 1989 gehaltener Vortrag mit dem Titel "Anwendungsmöglichkeiten von kommerzieller Software in der Archäologie" zugrunde. Für die kritische Durchsicht des Manuskripts sei Frau Irmela Herzog/Rheinisches Landesmuseum Bonn sowie Herrn Jörg Lindenbeck/Institut für Ur- und Frühgeschichte der Universität Köln herzlich gedankt.

(2) Siehe Gebühr u. Kampffmeyer 1980/81, 3ff., die einen informativen Überblick über die Entwicklung in den USA und Europa (Forschungsstand 1979) verschaffen und schon recht früh den Gebrauch von Kleinrechnern in der Ur- und Frühgeschichtsforschung propagierten. Einen kurzen, allgemeinen Überblick über die Verwendung von Mikrocomputern in der Archäologie gibt Hintze 1988. Weitere zusammenfassende Arbeiten existieren nicht, neuere Literatur ist in einigen Publikationen zu Einzelproblemen enthalten (z.B. Kalhoff u. Kampffmeyer 1984/85; Kampffmeyer 1987; Dibble u. McPherron 1988; Ruggles u. Rahtz 1988; Scollar et al. 1990).

(3) Siehe z.B.: Eggert, Kurz u. Wotzka 1980; Goldmann 1968; 1971; 1972; 1974; 1979; Gebühr 1975; 1976; Gebühr u. Kunow 1976; Gebühr u. Kampffmeyer 1980/81; Herzog u. Scollar 1987; Hintze 1986a; 1986b; 1986c; 1988a; 1988b; Ihm 1970; 1977; 1978; 1982a; 1982b; 1983; 1984a; 1984b; Kalhoff u. Kampffmeyer 1984/85; Kammerer 1968; 1979; Kampffmeyer 1987; Karstens 1978; Lindenbeck 1990; Meier-Arendt 1975; Scollar 1974; 1985; 1987; Scollar et al. 1990; Scollar, Weidner u. Herzog 1985; Steuer 1968; 1969; 1971; 1974; 1979; Vértessalij 1976/77; Ziegert 1983.

(4) Auf die Handhabung von Groß- und Kleinrechnern soll hier nicht näher eingegangen werden. Im PC-Bereich beziehen sich unsere Aussagen auf IBM oder IBM-kompatible Kleinrechner, deren weitverbreitetes DOS-Betriebssystem leicht erlernbar ist (von den zahlreichen Einführungen in das DOS-System sei hier das im Chip-Verlag erschienene Taschenbuch von G. Biethan empfohlen). Zu einigen Möglichkeiten des Kleinrechner-Einsatzes in der Archäologie siehe zuletzt zusammenfassend diverse Aufsätze in der Wissenschaftlichen Zeitschrift der Humboldt-Universität zu Berlin, Gesellschaftliche Reihe 37 (1988) 3 unter dem Titel "Mikrocomputer in der Archäologie". Siehe außerdem dazu Dibble u. McPherron 1988 oder Scollar et al. 1990. - Im Großrechnerbereich gibt es, je nach Rechnerart, mehrere Betriebssysteme. Wir arbeiteten mit einer IBM 4381 R23 (Betriebssystem VM/SP Rel.4 mit CMS Rel.4 / Universität-GHS Essen) und einer Cyber 170 Modell 855 (Betriebssystem NOS und NOS/VE / Universität Bochum). Beschreibungen und/oder Einführungen zu diesen oder anderen Systemen sind an den jeweiligen Rechenzentren der Universitäten oder anderer Institutionen (z.B. Behörde etc.) erhältlich. - Mittels eines Modems oder Akustikkopplers ist eine Kommunikation zwischen Großrechner und PC möglich.

(5) Die wenigen, eigens für die Archäologie entwickelten Programme sind entweder überhaupt nicht, nur schwer zugänglich oder in umständlichen Programmlistings publiziert. Das Bonner Seriations- und Statistikpaket (Scollar et al. 1990) oder das Kartierungsprogramm Sitefit (Lindenbeck 1990) bilden eine der wenigen Ausnahmen.

(6) Folgende Bereiche werden nicht behandelt: Bildverarbeitung, CAD-Programme (Erstellen von Zeichnungen), Textverarbeitung - Layout, DFÜ (Datenfernübertragung), Netzwerke, Utilities (Datensicherung, Datenrettung etc.), Expertensysteme. - Hinsichtlich der digitalisierten Bildverarbeitung sowie der Anwendungen von Kartierungs- oder CAD-Programmen (computer-gestütztes Zeichnen) in der Archäologie siehe Beiträge in diesem Heft.

(7) Weiterführende Literatur (außer den Handbüchern) zu dBASE IIIplus ist reichlich vorhanden. Empfehlenswert sind die Bücher von A. Simpson (1987a; 1987b). Die wichtigsten Daten von dBASE IIIplus sind dem Datenblatt Ashton-Tate 1986 zu entnehmen.

(8) Eine ausführliche Beschreibung der neuen dBASE IV-Versionen siehe in Herbert 1989; Hursh u. Hursh 1988; Simpson 1991.

(9) Ausführliche Besprechungen/Diskussionen und weiterführende Literatur zu den höheren statistischen Methoden, der Seriation und Simulation siehe beispielsweise in: Hintze 1986a; 1986b; 1986c; 1988a; 1988b; Leredde u. Djindjian 1980; Rozoy 1987; Scollar et al. 1990.

(10) Weiterführende Literatur (Auswahl): Carter u. Frurip 1985; Harting 1986; Hintze 1986; 1988; Ihm, Lüning u. Zimmermann 1978; Leredde u. Djindjian 1980; Ringrose 1988; Schubö u. Uehlinger 1986; Schuchard-Ficher et al. 1980; Scollar et al. 1990; Weber 1988.

---

---

(11) SAS enthält im Gegensatz zu SPSS ein umfangreiches Graphikpaket, das u.a. auch Routinen zum Zeichnen von Landkarten (bis auf Kreisebene, für Fundkarten interessant) vorsieht. Ein Beitrag dazu ist in Vorbereitung (Arbeitstitel: "Techniken zur Erstellung von archäologischen Karten mittels EDV"). Bei SPSS gibt es für die PC-Version eine Schnittstelle zum Business-Graphik-Programm "Chart" von Microsoft.

(12) Eingehende Beschreibungen siehe in Herzog u. Scollar 1987; Scollar et al. 1990. Zu weiteren Programmen bzgl. der Seriation und Clusteranalyse, P-, Q- und R-Matrix siehe Hintze 1986a; 1986b; 1986c; 1988; Duncan, Hodson, Orton, Tyers u. Vekaria 1988.

(13) Bzgl. einer ausführlicheren Darstellung diverser Seriationsmethoden siehe z.B. Leredde u. Djindjian 1980, 60ff.; Hintze 1986a; 1986b; 1986c; Karstens 1978; Scollar et al. 1990; Wenig 1986.

(14) Auf die Menübefehle wollen wir an dieser Stelle nicht weiter eingehen, sondern nur auf das sehr übersichtliche Handbuch verweisen. Im Handel gibt es weitere Benutzerliteratur (z.B. Lambert 1987).

---

Dr. Uwe Schoenfelder  
Niederrheinisches Museum  
"Projekt Archäologische Ausgrabungen"  
Friedrich-Wilhelm-Str. 64  
4100 Duisburg 1

Dr. Manfred Erdelen  
Hochschulrechenzentrum  
der Universität - GHS Essen  
Schützenbahn 79  
4300 Essen 1

