

# Interaktive Datenverarbeitung und -analyse im Statistischen Landesamt

## SAS auf dem Großrechner

Als zentrale Aufgabe führt das Statistische Landesamt Baden-Württemberg im Bund-Länder-Verband der amtlichen Statistik die Erhebung, Aufbereitung und Auswertung von Bundes- und Landesstatistiken sowie Basisanalysen, zum Beispiel im Bereich Bevölkerung oder Umwelt, sowie statistisch-methodische Untersuchungen durch.

Im Rechenzentrum des Statistischen Landesamts war bisher (historisch gewachsen) eine Vielzahl von unterschiedlichen Programmen bzw. Programmsystemen für den mathematisch-statistischen Bereich im Einsatz, zu nennen sind hier unter anderem die BMD-Programme (Biomedical Computer Programm) von der GMD oder das SSP-Paket (Statistical Science Package) von der IBM oder Programme für Zeitreihenanalyse.

Bei diesen Programmen handelte es sich um sogenannte Batch-Anwendungen, die nach Angabe der entsprechenden Steuerinformationen und Eingabedaten durch EDV-Fachkräfte auf der Großrechneranlage ablaufen. Seitens der statistischen Fachbereiche ist seit längerem gewünscht worden, mathematisch-statistische Verfahren benutzerfreundlich in eigener Verantwortung abwickeln zu können.

Im Jahr 1988 wurde daher nach eingehender Prüfung im Rechenzentrum, vor allem zur Ablösung der bisherigen mathematisch-statistischen Programmpakete, aber auch zur Verbesserung der graphischen Datenverarbeitung, das Programmsystem SAS (Statistical Analysis System) beschafft.

SAS ist ein umfassendes und vielseitiges Programmpaket zur Datenanalyse, Datenverwaltung und Präsentation. Ursprünglich für statistische Aufgaben entwickelt, ist das Basisprodukt nach und nach durch verschiedene Zusatzprodukte, zum Beispiel für Graphik, Tabellenkalkulation, Textverarbeitung, Berichterstellung, Computerleistungsmessung usw. ergänzt worden, so daß inzwischen mit diesem integrierten Gesamtsystem nahezu alle Anforderungen von Statistikern bewältigt werden können:

- mathematisch-statistische Verfahren
- graphische Verfahren
- Programmiersprache
- ausbaufähiges menügesteuertes Endbenutzersystem
- Schnittstelle zu Dateien und Datenbanken
- Tabellen- und Textverarbeitung.

Die Software wird von der Firma SAS Institute, mit Sitz in Heidelberg, seit 1982 in Deutschland vertrieben und hat hier mittlerweile nach eigenen Angaben ca. 1 200 Installationen. Die gesamte Programmpalette steht für Großrechner und PCs bei gleicher Syntax bzw. gleichem Leistungsumfang zur Verfügung. Im Statistischen Landesamt ist SAS als interaktives System unter dem tp-Monitor TSO im Einsatz.

Bisher werden die Komponenten SAS/BASE als Grundsoftware für mathematisch-statistische Methoden und zur Datenverwaltung, SAS/GRAPH für die Erstellung von Prä-

sentationsgraphiken und thematischen Karten, SAS/FSP zur Tabellenkalkulation und Textverarbeitung, sowie SAS/AF zur Generierung einfach zu handhabender Menüsysteme direkt für den Endbenutzer eingesetzt. Außerdem sind Erweiterungen in Richtung Qualitätskontrolle, Zeitreihenanalyse, Operations Research und Schnittstellen zu den Datenbanksystemen IMS oder ADABAS möglich. Gerade diese Direktzugriffsmöglichkeit auf viele andere externe Datenhaltungen ist einer der großen Pluspunkte von SAS im Vergleich zu ähnlichen Systemen.

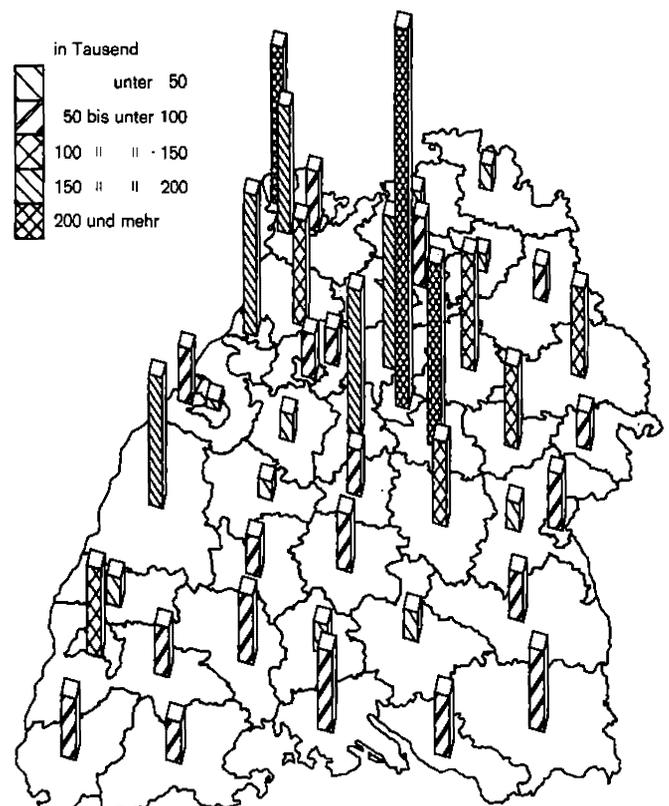
Die einzelnen Module beinhalten folgende wichtige Prozeduren und Funktionen:

### SAS/BASE

BASE ist die Grundsoftware, die auch für den Einsatz aller anderen Optionen benötigt wird. Sie umfaßt das gesamte Datenmanagement wie Eingabe, Datenmanipulation und Speicherung. Es kann praktisch jede Datenstruktur gelesen und dann in speziellen SAS-Bibliotheken abgelegt werden. Die Daten können dort jederzeit

Abbildung 1

### Beschäftigte in den Stadt- und Landkreisen Baden-Württembergs am 25. Mai 1987



Statistisches Landesamt Baden-Württemberg

264 90

aktualisiert, sortiert, verkettet oder dokumentiert werden. Das Grundmodul enthält außerdem Prozeduren für statistische Auswertungen. Das Repertoire reicht von einfachen beschreibenden Statistiken und Kreuztabellen bis hin zu komplexen multivariaten Verfahren wie Cluster- oder Diskriminanzanalyse.

### SAS/GRAPH

Der Leistungsumfang dieser Teilkomponente ist imponierender als bei manchem Spezialprodukt. Erstellt werden können die üblichen Linien-, Balken-, Kreis- oder Streudiagramme, auch dreidimensional, wobei der Anwender bei der Gestaltung der Graphik unter einer Vielzahl von Schriftarten, Schraffurmustern und Farbschattierungen auswählen kann. Ein interaktiver Katalog dient zur Speicherung und Verwaltung der Schaubilder. Hiermit ist auch eine gezielte Überlagerung mehrerer bereits erstellter Graphiken in einem Schaubild möglich. Am beeindruckendsten sind aber die Programmroutinen zur thematischen Kartographie, angefangen bei der einfachen Flächenkarte, bis hin zur dreidimensionalen, beliebig drehbaren Prismenkarte.

Aufgrund eigener Programmierung ist inzwischen auch der Einbau von Kreisdiagrammen oder Texten in Karten möglich. Darstellbar sind alle regionalen Einheiten, für die Koordinaten verfügbar sind. Im Statistischen Landesamt können Karten von Baden-Württemberg bis auf Gemeindeebene direkt am Graphikterminal erzeugt werden.

### SAS/FSP

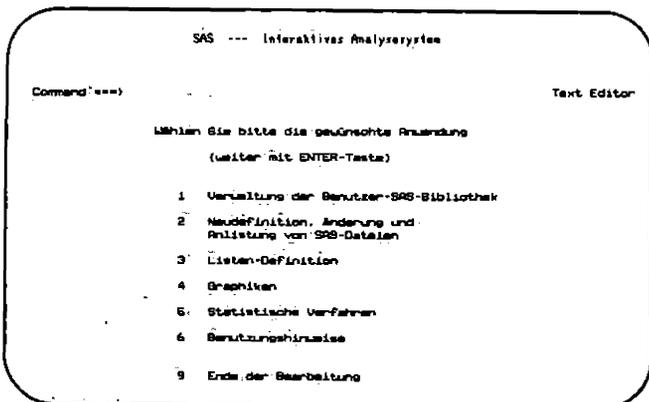
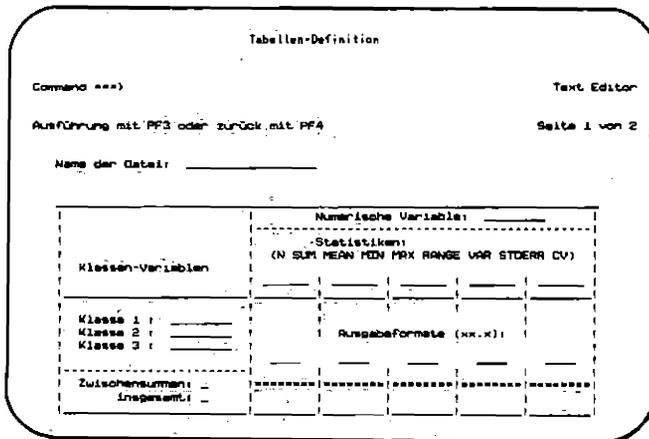
Zu FSP gehören die interaktiven Prozeduren zur mehrdimensionalen elektronischen Spreadsheet-Verarbeitung, Tabellen- bzw. Berichterstellung und Textverarbeitung. Innerhalb der Tabellenkalkulation besteht eine direkte Verbindung zur SAS-Programmiersprache. FSP dient aber auch zur Datenerfassung, sei es manuell oder durch Lesen externer Dateien. Hilfreich sind hierbei vom Anwender konstruierbare Plausibilitätskontrollen mit Festlegung von Wertebereichen, Initialisierungen oder Sperrung von Feldern.

### SAS/AF

Bei AF handelt es sich um eine bildschirmorientierte Software zur Generierung von menügesteuerten und benutzerfreundlichen Anwendungen, die dann eine Schnittstelle zum Endbenutzer bilden, der nur über wenig detaillierte EDV-Kenntnisse verfügt. SAS besteht jedoch nicht nur aus einer großen Menge von fertigen Prozeduren, die aufgerufen und mit bestimmten Parametern belegt werden müssen, sondern ist darüber hinaus zugleich eine Programmiersprache der 4. Generation mit eigener Syntax, mit der im Statistischen Landesamt schon etliche Programme entwickelt worden sind.

Abbildung 2

### Beispiele für Auswahlmensüs



### Interaktives Analyse- und Auswertungssystem

Als bisher wichtigste und umfassendste *Eigenentwicklung* des Statistischen Landesamtes Baden-Württemberg ist hier das interaktive Analysesystem „SAS62“ zur dezentralen Datenverarbeitung durch die statistischen Fachbereiche des Amtes zu nennen. Dieses Programmsystem wird über einfache, hierarchisch aufgebaute Auswahlmensüs gesteuert und deckt die wichtigsten und meistbenutzten statistischen Auswertungen ab. Es können damit sowohl externe Dateien gelesen als auch manuell Daten eingegeben werden. Die Variablenwerte können nach verschiedenen Merkmalen aggregiert, selektiert oder in Intervalle eingeteilt und dann in einer eigenen Bibliothek gespeichert und verwaltet werden. Das Angebot an mathematisch-statistischen Funktionen reicht von Summen- und Mittelwertbildungen über Kreuztabellen und Korrelationen bis zu Regressions- und Varianzanalysen. Zur Veranschaulichung und Präsentation der Ergebnisse dienen eine Reihe von verschiedenen Standardgraphiken sowie die Möglichkeit der Druckausgabe in Form von mehrfach untergliederten Tabellen.

### Graphische Erweiterungen

Das SAS-System ist auch für Erweiterungen im graphischen Bereich nutzbar. Hier ist ein Online-Menüsystem zu erwähnen mit dem spezielle Schaubilder zu den Bereichen Bodennutzung, Viehwirtschaft und Umwelt dargestellt werden können. Die Präsentationsgraphiken und themati-

schen Karten können wahlweise mehr- oder einfarbig für jede Gemeinde erzeugt werden.

## Datenschutz

Das beschriebene interaktive Datenaufbereitungs- und -analysesystem, das es den Fachbereichen ermöglicht, Erfassungs-, Aufbereitungs- und Analysearbeiten mit statistischen Daten ohne Programmierunterstützung durchzuführen, ist seit Mitte 1989 nach entsprechenden Schulungen im Praxistest. Zur Gewährleistung der Belange des Datenschutzes werden den Fachbereichen nur Summendaten bzw. anonymisierte Einzeldaten aus dem Bereich des Großrechners zur Verfügung gestellt.

## Ausblick

Die Erstellung derartiger Menüsysteme für Anwendungen ist wohl die adäquate Nutzung der maschinellen Datenverarbeitung für Fachbereiche: Da die Mächtigkeit und Kom-

plexität von SAS als Gesamtsystem eine mehrwöchige Ausbildung und vor allem regelmäßige Nutzung voraussetzen, ist SAS als Programmiersprache nur für EDV-Spezialisten geeignet. Zur Zeit wird das interaktive System im Statistischen Landesamt von ca. 25 wissenschaftlichen Mitarbeitern genutzt.

Parallel zum Einsatz des Menues wurden und werden natürlich ständig SAS-Programme für gezielte Anforderungen in den Sachgebieten Mathematische Statistik und Graphik geschrieben.

Im Großrechnerbereich ist damit ein ganz wesentlicher Schritt in Richtung dezentrale Datenverarbeitung gemacht worden. Parallel laufen zur Zeit im Statistischen Landesamt die Überlegungen bzw. Vorbereitungen, durch den Einsatz von Personalcomputer weitere Verbesserungen in Richtung EDV-gestützter Erfassung, Aufbereitung und Auswertung von (kleinen) Statistiken, Verwaltungsunterstützung sowie Textverarbeitung bzw. Text- und Datenkommunikation zu erreichen.

Dipl.-Math. Karl-Albert Heilmann

## Drei Charaktere - eine Trilogie

Kassette



Das noch junge Bundesland Baden-Württemberg — es entstand erst nach dem Zweiten Weltkrieg — nimmt einen vorderen Platz in der Bundesrepublik Deutschland ein, nicht zuletzt auf wirtschaftlich-technischem Gebiet. Auch die Vielfalt des landschaftlichen und kulturellen Angebots trägt zur Attraktivität des Landes bei.

Die Bändchen dieser Trilogie weisen auf wichtige Gründe des Aufschwungs hin: Fleiß, Erfindungsreichtum, Sparsamkeit und kritisches Nachdenken — alles dies zeichnet die Menschen dieser Regionen seit vielen Generationen aus.

Die Darstellungen beschreiben den jeweiligen Volkscharakter und informieren über topographisch-statistische Gegebenheiten im vergangenen Jahrhundert. Die Wiedergabe von Sitten, Sagen, Schwänken, Anekdoten und Gebräuchen rundet diese lebendigen Berichte ab. Sie bieten die Gelegenheit, Baden-Württemberg und seine Bewohner früher und heute in entspannender Lektüre kennenzulernen.

Dr. Guntram Palm  
Finanzminister des Landes Baden-Württemberg

3 Broschüren  
(mit Kassette)  
im Format  
DIN A 5;  
Gesamtumfang  
111 Seiten;  
Preis: DM 14,-;  
ISBN:  
3-923292-16-3

Herausgeber: Statistisches Landesamt Baden-Württemberg  
Postfach 10 60 33, 7000 Stuttgart 10