

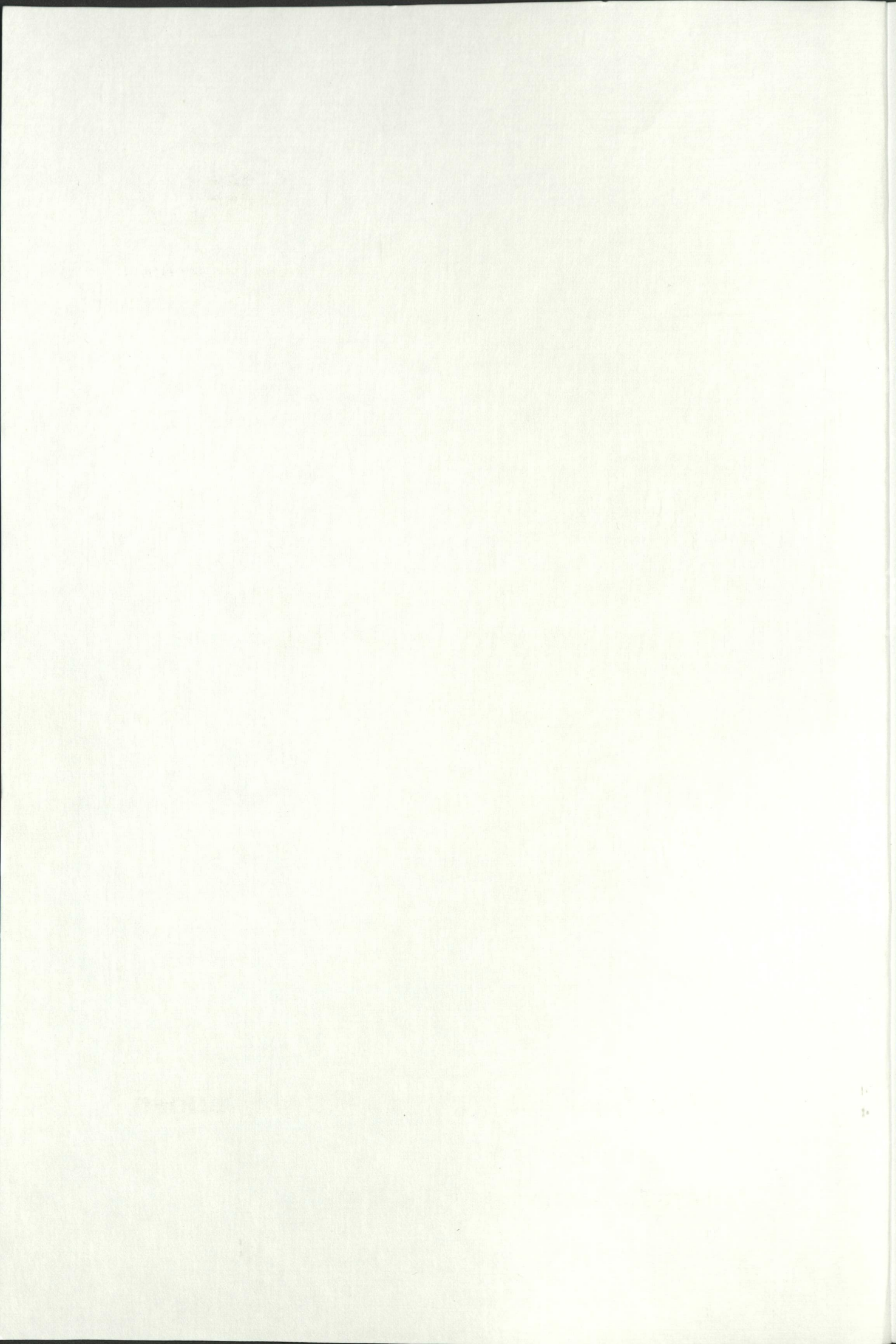
Statistisches Bundesamt

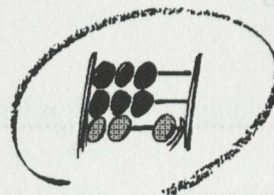
Bernd Schmidt

Fachliche Erprobung der Statistik-Software SPSS

Methoden - Verfahren - Entwicklungen

Materialien und Berichte





Statistisches Bundesamt

0	Vermerkungen	6
---	--------------	---

1	Einsatz von SPSS im Rahmen von § 7-Erhebungen	7
---	---	---

1.1	Erhebungen nach § 7 Abs. 1 BStatG	7
-----	-----------------------------------	---

1.2	Erhebungen nach § 7 Abs. 2 BStatG	11
-----	-----------------------------------	----

2	Das SPSS-Leistungsspektrum	15
---	----------------------------	----

2.1	Bernd Schmidt	15
-----	---------------	----

2.2	Der Dateiaufruf von SPSS 6.0.1	19
-----	--------------------------------	----

3	Fachliche Erprobung der Statistik-Software SPSS	25
---	---	----

3.1	Handhabung	26
-----	------------	----

3.2	Systemvoraussetzungen/Systemcharakteristika	29
-----	---	----

3.2.1	Stabilität des Systems und Laufzeiteffizienz/Schnelligkeit der	29
-------	--	----

3.2.1.1	Umsetzbarkeit statistischer Prozeduren	35
---------	--	----

3.2.1.2	Daten(werk)-Management	37
---------	------------------------	----

3.2.1.3	Ergebnisdarstellung	39
---------	---------------------	----

3.2.1.4	Produktsupport	40
---------	----------------	----

3.2.2	Neuerungen in den Versionen SPSS 7.0/7.5	44
-------	--	----

3.2.3	Empfehlungen für den zukünftigen Einsatz von SPSS	45
-------	---	----

3.2.4	Materialien und Berichte	46
-------	--------------------------	----

3.3		46
-----	--	----

3.4		46
-----	--	----

3.5		46
-----	--	----

3.6		46
-----	--	----

3.7		46
-----	--	----

3.8		46
-----	--	----

3.9		46
-----	--	----

3.10		46
------	--	----

3.11		46
------	--	----

3.12		46
------	--	----

3.13		46
------	--	----

3.14		46
------	--	----

3.15		46
------	--	----



Statistisches Bundesamt

Bernd Schmidt

Herausgeber und Vertriebsstelle:

Statistisches Bundesamt, *Gruppe IB - Institut für Forschung und Entwicklung in der Bundesstatistik*,
65180 Wiesbaden



Fachliche Informationen zu dieser Veröffentlichung können Sie direkt beim Statistischen Bundesamt erfragen:
Gruppe IB, Telefon: 06 11 / 75 26 32 oder Fax: 06 11 / 75 39 50

Erschienen im April 1998



Allgemeine Informationen über das Statistische Bundesamt und sein Datenangebot erhalten Sie:

- im Internet: <http://www.statistik-bund.de>

oder bei unseren Allgemeinen Auskunftsdiensten

65180 Wiesbaden

- Telefon: 06 11 / 75 24 05
- Telefax: 06 11 / 75 33 30
- E-Mail: auskunftsdienst@stba.bund400.de

Zweigstelle Berlin

Postfach 276

10124 Berlin

- Telefon: 030 / 23 24 68 66
- Telefax: 030 / 23 24 68 72
- E-Mail: stba-berlin.infodienst@t-online.de

Statist. Bundesamt - Bibliothek



24-00947

© Copyright: Statistisches Bundesamt, Wiesbaden 1998
Alle Rechte vorbehalten.

Sämtliche Daten sowie alle Teile der Dokumentation unterliegen dem Urheberrecht. Alle Rechte sind geschützt. Jegliche Vervielfältigung oder Verbreitung, ganz oder teilweise, bedarf der vorherigen Zustimmung.

Methoden - Verfahren - Entwurf - Erhebung - Materialien und Berichte

0	Vorbemerkungen	6
1	Einsatz von SPSS im Rahmen von § 7-Erhebungen	7
1.1	Erhebungen nach § 7 Abs. 1 BStatG	7
1.2	Erhebungen nach § 7 Abs. 2 BStatG	11
2	Das SPSS-Leistungsspektrum	15
2.1	Der Funktionsumfang von SPSS 6.0.1	15
2.2	Der Dateiaufbau von SPSS 6.0.1	19
3	Beurteilung von SPSS in der praktischen Anwendung	25
3.1	Beurteilungskriterien	26
3.2	Kriterienbezogene Bewertung	29
3.2.1	Handhabung	29
3.2.1.1	Systemvoraussetzungen/Systemcharakteristika	29
3.2.1.2	Stabilität des Systems und Laufzeiteffizienz/Schnelligkeit der Umsetzbarkeit statistischer Prozeduren	30
3.2.1.3	Daten(satz)-Management	35
3.2.1.4	Benutzerfreundlichkeit	37
3.2.2	Spektrum der statistischen Prozeduren	39
3.2.3	Ergebnisdarstellung	40
3.2.4	Produktsupport	49
3.3	Neuerungen in den Versionen SPSS 7.0/7.5	50
4	Empfehlungen für den zukünftigen Einsatz von SPSS im Statistischen Bundesamt	54

Abbildungen

	Seite
Abbildung 1: Anwendungsfenster und Dateneditor-Fenster von SPSS 6.0	18
Abbildung 2: SPSS-Daten-Datei angezeigt im SPSS-Dateneditor	20
Abbildung 3: Darstellung einer SPSS-Ausgabe im Ausgabe-Fenster	20
Abbildung 4: Darstellung einer SPSS-Grafikdatei	21
Abbildung 5: Zusammenspiel der vier verschiedenen Dateiebenen bei der Durchführung von statistischen Prozeduren	24
Abbildung 6: Darstellung univariater Statistiken im SPSS-Output-Fenster	41
Abbildung 7: Beispiel für eine Boxplot-Darstellung im SPSS-Grafikfenster	44
Abbildung 8: In SPSS erstellte und in Winword eingebundene statistische Grafiken	45
Abbildung 9: Zwei in SPSS erstellte einfache Balkendiagramme	47
Abbildung 10: In SPSS erstelltes gestapeltes bzw. gruppiertes Balkendiagramm	48
Abbildung 11: Einsatzgebiete von SPSS im Statistischen Bundesamt	54

Tabellen

	Seite
Tabelle 1: Wichtige Features des SPSS-Basismoduls	16
Tabelle 2: Praxisbezogene Beurteilungskriterien	28
Tabelle 3: Laufzeitverhalten von SPSS	32

Literaturverzeichnis.....	59
Anlage 1: Berechnung des Zusammenhangs zwischen Anlage- investitionen, Umsätzen und Mitarbeitern 1995 (Pearson'scher Korrelationskoeffizient) aus den Angaben der Investitionsschnellerhebung	61
Anlage 2: Tabellarischer Report zu den Anlageinvestitionen und Umsätzen 1995	62
Anlage 3: SPSS-Tabelle zu den Anlageinvestitionen, Ausrüstungs- investitionen und Bauinvestitionen 1995 im früheren Bundesgebiet und in den neuen Ländern	63
Anlage 4: Schulungsprogramm von SPSS im Zeitraum 9'97 bis 3'98	64
Anlage 5: Überblick über wichtige Features von SPSS 7.5	65

0 Vorbemerkungen

Statistische Softwarepakete für den PC haben in den letzten zehn Jahren als Datenaufbereitungs-, auswertungs- und analyseinstrumente in der Praxis zunehmend an Bedeutung gewonnen. Mit der rasanten Entwicklung im PC-Hardwarebereich – immer leistungsfähigere Prozessoren und Programmumgebungen, schnell ansteigende Speicherkapazitäten – stellen diese Programme ernstzunehmende Ergänzungen/Alternativen zu den Statistikprogrammen der Großrechnergeneration dar. Eines der bekanntesten Programme auf dem einschlägigen Softwaremarkt, das frühzeitig auch als PC-Version angeboten wurde und das, nachdem sich die Benutzeroberfläche MS Windows¹ als Marktstandard etabliert hatte, frühzeitig auch als Windows-Version erhältlich war, ist das integrierte Statistiksoftwarepaket SPSS (Statistical Package for the Social Sciences).

Mit der Intention, Erfahrungen mit entsprechender Software zu bekommen, wurde SPSS im Statistischen Bundesamt schon zu Beginn der 90er Jahre vom Institut für Forschung und Entwicklung² im Rahmen von Erhebungen für besondere Zwecke nach § 7 BStatG sowohl zur Datenaufbereitung als auch zur Datenauswertung und Ergebnispräsentation eingesetzt. In der Zwischenzeit sind allein in der Gruppe I B neun § 7-Erhebungen mit dem SPSS-Instrumentarium durchgeführt worden, insofern kann auf einen schon sehr vielfältigen Erfahrungsschatz im praktischen Umgang mit SPSS zurückgegriffen werden. Ziel dieser Arbeit ist es, aufbauend auf diese Erfahrungen eine praxisbezogene Bewertung durchzuführen und auf Möglichkeiten und Grenzen von SPSS im Hinblick auf eine breitergefächerte Nutzung im Rahmen der Statistikproduktion im Statistischen Bundesamt hinzuweisen.

Dieser Sachverhalt gewinnt vor dem Hintergrund der Planungen, bis zur Jahrtausendwende das "Statistische Informationssystem des Bundes" STATIS-BUND durch das "gemeinsame neue statistische Informationssystem" GENESIS zu ersetzen, zusätzlich an Bedeutung. Datenauswertungen sollen dann zunehmend von der Großrechnerebene auf die Ebene von Mikrocomputern verlagert werden.

¹ Nachfolgend wird bei der Erwähnung von Microsoft-Produkten auf das vorangestellte Kürzel "MS" verzichtet.

² Früher Gruppe IE; jetzt Gruppe IB.

1 Einsatz von SPSS im Rahmen von § 7-Erhebungen

Erhebungen für besondere Zwecke ergänzen das aus einer Vielzahl einzelstatistischer Rechtsgrundlagen bestehende System der Bundesstatistik um die Möglichkeit, Erhebungen *ohne* Einschaltung des Gesetzgebers mit dem Ziel der Erfüllung eines kurzfristig auftretenden Datenbedarfs oberster Bundesbehörden (§ 7 Abs. 1 BStatG) oder der Klärung wissenschaftlich-methodischer Fragestellungen auf dem Gebiet der Bundesstatistik (§ 7 Abs. 2 BStatG) durchzuführen. Im Rahmen von § 7-Erhebungen dürfen maximal 10 000 Erhebungseinheiten befragt werden; die Auskunftserteilung ist freiwillig. Nicht zuletzt aufgrund der großen inhaltlichen Bandbreite waren die in den letzten Jahren durchgeführten § 7-Erhebungen gut geeignet für einen Praxistest von SPSS. Nachfolgend wird ein Abriß über die einzelnen Erhebungen gegeben:

1.1 Erhebungen nach § 7 Abs. 1 BStatG

- Die erste § 7-Erhebung, bei der SPSS eingesetzt wurde, war **"Der in Plattenbauweise errichtete Wohnungsbestand in der ehemaligen DDR – Struktur, Erhaltungszustand, Umfeld, Akzeptanz durch die Mieter"**. Der Bearbeitungszeitraum der Erhebung erstreckte sich auf die Jahre 1991 und 1992. Das Erhebungsziel bestand darin, die Struktur des Gebäude- und Wohnungsbestandes herauszuarbeiten, das heißt Aussagen zu machen über die Größe und regionale Verteilung des Wohnungsbestandes, über seine Nutzung, über die Größe der Wohnungen sowohl hinsichtlich der Fläche als auch hinsichtlich der Zahl der Räume und über die Art der Beheizung sowie über die dabei verwendete Energie. Darüber hinaus sollte der Zustand der Gebäude und Wohnungen charakterisiert und eine abgestufte Dringlichkeitsliste über notwendige Sanierungsmaßnahmen aufgestellt werden. Ziel war es auch, das Umfeld der Gebäudekomplexe darzustellen und die Akzeptanz der Wohnungen bei den Mietern zu erfragen³. Schriftlich befragt wurden in einer ersten Erhebungsstufe 1 131 Wohnungsbaugenossenschaften und Wohnungsverwaltungen in den neuen Ländern; in einer zweiten Erhebungsstufe wurden aus den in der ersten Erhebungsstufe erfaßten

³ Vgl. Deckl, S.; Bechtold, S.: Wohngebäude in Plattenbauweise. Erste Ergebnisse einer Erhebung nach § 7 Abs. 1 BStatG bei Eigentümern und Verwaltern; in: *Wirtschaft und Statistik*, 2/1992, S. 88.

Wohngebäuden 2 003 Mieterhaushalte über persönliche Interviews zu Wohnumfeld und Wohnsituation befragt.

Das Statistikprogrammpaket SPSS (DOS-Version SPSS/PC+ 4.0) wurde überwiegend zur **Datenaufbereitung und -auswertung**, im Rahmen der Eigentümerbefragung aber auch zusätzlich zur **Datenerfassung**⁴ eingesetzt. Für die Zwecke der Datenerfassung wurde im Statistischen Bundesamt mit dem in SPSS zur Verfügung stehenden **Data-Entry-Modul** eine Eingabemaske entwickelt, die in ihrer Struktur den in der Erhebung eingesetzten Fragebogen widerspiegelte. In die Datenerfassungssapplikation wurden **erste grundlegende Plausibilitätskontrollen** – wie zum Beispiel Werteober- und untergrenzen – eingebaut, ferner wurde die Funktion des "**dynamic routing**", sprich die Eingabeführung über Sprungfolgen, ebenfalls in die Applikation integriert. Die dv-technische Datenaufbereitung (Plausibilitätsprüfungen) ebenso wie die anschließende Datenauswertung sowohl der Eigentümer- als auch der Mieterbefragung erfolgte ausschließlich mit der o. g. SPSS-Version am PC.

- In der im Jahr 1992 durchgeführten Erhebung zu den "**Einmaligen Leistungen der Hilfe zum Lebensunterhalt nach §21 BSHG**" wurde ebenfalls die o. g. DOS-Version von SPSS eingesetzt. Ziel der Erhebung war es, dem Bundesministerium der Finanzen differenzierte statistische Informationen im früheren Bundesgebiet zu den einmaligen Leistungen, die auf Antrag ergänzend zu den laufenden Sozialhilfeleistungen gewährt werden, zu geben. Als Erhebungseinheit dienten 650 amtliche Stellen, die mit der konkreten Bearbeitung und Gewährung von Sozialhilfe befaßt sind und die über 6 284 Bedarfsgemeinschaften Auskunft geben konnten.⁵ Was die organisatorisch-technische Durchführung der Erhebung insbesondere unter dem Gesichtspunkt der SPSS-Nutzung betrifft, so wurde im Prinzip wie in der Erhebung über den Wohnungsbestand in Plattenbauweise verfahren. Ein wichtiger Unterschied betraf jedoch die Datenerfassung. Für die Datenerfassung wurde nicht eine Applikation über das SPSS Data-Entry Modul entwickelt, sondern – ebenfalls um erste Erfahrungen zu sammeln – das vom Niederländischen Statistischen Zentralamt (CBS) entwickelte integrierte **Programmpaket Blaise** genutzt. Die über Blaise erfaßten Datensätze ließen

⁴ Im Rahmen der Mieterbefragung wurde ein privatwirtschaftliches Institut mit der Feldarbeit beauftragt, das auch die Datenerfassung durchführte.

⁵ Vgl. Bechtold, S., Bihler, W., Deininger, D.: Einmalige Leistungen der Hilfe zum Lebensunterhalt nach § 21 BSHG im Jahr 1991, in: *Wirtschaft und Statistik*, 2/1993, S. 113 ff.

sich problemlos in SPSS einlesen⁶, wobei auch die in Blaise erstellten variablenspezifischen **Metainformationen** – wie z. B. Variablenname, Variablenlabels, Variablentyp, Merkmalsausprägungen – **beim Einlesen der Datensätze in SPSS übernommen wurden.**

- Im Rahmen der Erhebung "**Einkommensverhältnisse 1992 von Familien mit Kindern in Deutschland**" wurde zum ersten Mal die Windows Version von SPSS (SPSS 6.0) eingesetzt. Das Ziel der Erhebung bestand u. a. darin, aktuelle Informationen zur Einkommensphäre und zur Einkommensentwicklung in bestimmten Lebensphasen von Familien mit wirtschaftlich abhängigen Kindern zu bekommen; darüber hinaus sollte auch die Wohnungssituation und die Frage der Erwerbstätigkeit von Frauen in entsprechenden Familien näher analysiert werden. Als Auswahl- und Erhebungseinheit dienten Familien im früheren Bundesgebiet und in den neuen Ländern. Befragt wurden über persönliche Interviews 4 000 Ehepaare mit Kindern, 1 250 Alleinerziehende und 750 Ehepaare ohne Kinder. Die Feldarbeit, die Datenerfassung und der überwiegende Teil des Mikro-Editings im Rahmen der Plausibilitätsprüfungen wurde durch die GfK Marktforschung durchgeführt. Umfangreiche Plausibilitätsprüfungen auf der Ebene des Datengesamtzusammenhangs (**Makro-Editing**) waren insbesondere beim Brutto- und Nettoeinkommen aufgrund der Sensibilität aber auch aufgrund der Komplexität des Merkmals notwendig. Diese Arbeiten wurden allesamt zentral im Statistischen Bundesamt über SPSS unter Nutzung des dafür vorgesehenen Instruments der **explorativen Datenanalyse** durchgeführt. Die sich anschließende Datenauswertung, v. a. aber auch die gesamte Ergebnispräsentation (Veröffentlichungstabellen, Ergebnisgrafiken) wurde ebenfalls über die in SPSS hierfür zur Verfügung stehenden Grafik- und Tabellen-Module vorgenommen. Intensiven Einsatz im Rahmen der **Analyse der Einkommenverteilung** erfuhren die in SPSS zur Verfügung stehenden **grafischen Aufbereitungs- und Präsentationsmöglichkeiten** in Form von Box-Plot- und Stem-and-Leaf-Darstellungen, darüber hinaus wurde aus dem umfangreichen statistischen Methodenspektrum für Analysezwecke v. a. die **Möglichkeit zur Varianzanalyse** mehrfach eingesetzt, die auf so einfache wie effiziente Art und Weise erlaubt, signifikante Unterschiede zwischen Gruppenmittelwerten – wie zum Beispiel potentielle statistisch signifikante Unterschiede im

⁶ Für diesen Zweck kann auf das Blaise-Tool "Cameleon" zurückgegriffen werden.

arithmetischen Mittel der Familien-Bruttojahreseinkommen der einzelnen Familientypen – nachweisen zu können, ohne komplizierte und langwierige Programmierungen anstellen zu müssen.⁷

- Die erste *Unternehmensbefragung*, in der WinSPSS im Rahmen von § 7-Erhebungen genutzt wurde, war die im Auftrag der EU-Kommission und im Rahmen des EG-Aktionsprogramms FORCE (Formation Continue en Europe) durchgeführte **"Erhebung zur beruflichen Weiterbildung in Unternehmen"**. Zugrundegelegt wurde ein zweistufiges Erhebungskonzept: Ende 1993 wurden in der Vorerhebung ca. 10 000 Unternehmen befragt, um grundlegende Informationen zum Angebot beruflicher Weiterbildung in Unternehmen zu bekommen. Anfang 1994 wurden diese Informationen vertieft und um weiterführende Inhalte (Kosten und Nutzung der Weiterbildung, Fragen der Weiterbildungskonzeption etc.) ergänzt.⁸ Was die dv-technische Abwicklung betrifft, wurde an die sich in den vorangegangenen § 7-Erhebungen bewährte dezentrale PC-Lösung angeknüpft, d. h. die Datenerfassung erfolgte über Blaise, die Datenaufbereitung und -auswertung wurde über WinSPSS durchgeführt. Im Rahmen der Datenaufbereitung und -auswertung konnten neue Erkenntnisse über die Eignung von SPSS zur **Beseitigung von Item-Nonresponse** und zur **Aggregation von Einzeldaten** gewonnen werden. Die Beseitigung von Item-Nonresponse wurde in Absprache mit Eurostat über ein kombiniertes Verfahren unter Einbeziehung von Referenzklasseninformationen und geeigneten Informationen aus dem entsprechenden unvollständigen Fragebogen vorgenommen. Fachliche Konzeption und dv-technische Umsetzung (Programmierung) wurde ausschließlich in der Fachabteilung unter Nutzung von SPSS realisiert, wobei die sich bietende Möglichkeit zu einer **engen personellen Verzahnung** fachlicher und dv-technischer Belange nicht zuletzt aufgrund des flexiblen und schnellen Informationsflusses sowohl unter Effektivitäts- als auch unter Effizienzgesichtspunkten sehr positiv beurteilt werden konnte. Ebenfalls in der Praxis bewährt hat sich die Aggregation von Einzeldaten unter Zuhilfenahme von SPSS. Aus Datenschutzgründen wurde an Eurostat nur ein aggregierter Datensatz bestehend aus 96 Schichten (Kombinationen aus Wirtschaftszweigen und Beschäftigtengrößenklassen) übermittelt. Die notwendige Datenaggregation konnte auf schnelle

⁷ Vgl. Wein, E.: Einkommensverhältnisse 1992 von Familien mit Kindern in Deutschland; Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1994, S. 38 ff.

⁸ Hinsichtlich weiterführender Informationen vgl. Schmidt, B.: Berufliche Weiterbildung in Unternehmen, Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1996, S. 17 ff.

und unkomplizierte Weise durchgeführt werden, da in WinSPSS hierfür ein entsprechendes Tool zur Verfügung steht, das umfangreiche Programmierarbeiten überflüssig machte.

- Im Jahr 1995/96 wurde mit der **"Investitionsschnellerhebung"** die gegenwärtig letzte Erhebung nach § 7 Abs. 1 BStatG durchgeführt, in der WinSPSS Anwendung gefunden hat. Hierbei handelte es sich um eine Befragung von ca. 10 000 Unternehmen des privaten Unternehmenssektors (ohne primären Sektor und Wohnungswirtschaft) zu ihren Investitionsaktivitäten in den Jahren 1994 und 1995.⁹ Über die gängigen Arbeitsprozeduren im Rahmen der Datenaufbereitung und -auswertung hinaus fand SPSS Anwendung im Rahmen der **Ergebnisevaluation**. Um die Qualität der Erhebungsergebnisse einschätzen zu können, wurde in den Bereichen Bergbau, Verarbeitendes Gewerbe, Baugewerbe über das SPSS-Instrumentarium ein differenzierter Ergebnisvergleich zwischen Investitionsschnellerhebungsergebnissen und Ergebnissen der entsprechenden Jahresherhebungen sowohl auf der Ebene der Einzeldatensätze (Mikroebene) als auch auf der Ebene von wirtschaftsbereichsspezifischen Gesamtergebnissen (Makroebene) durchgeführt. Erwähnenswert ist auch die kurze Bearbeitungszeit der Erhebung: Die Erhebung wurde in lediglich 8 Monaten konzipiert und realisiert, was aufgrund der schnellen und flexiblen dv-technischen Umsetzung – die Datenaufbereitung und Datenauswertung nahm nur ca. 1 Monat in Anspruch – ein wichtiger Verdienst von SPSS war.

1.2 Erhebungen nach § 7 Abs. 2 BStatG

- Die **Tourismusstichprobe** sollte als Methodenuntersuchung sowohl Erhebungsinhalte als auch mögliche Erhebungsverfahren auf ihre Eignung für eine Befragung zum Reiseverhalten testen.¹⁰ In der Erhebung sollten alle Reisen (Privat-, Urlaubs-, Erholungs-, Kur-, Besuchs-, Geschäfts-, Dienstreisen) mit mindestens einer Übernachtung, die die Bundesbürger im Jahr 1992 unternahmen, erfaßt werden. Die Erhebung wurde im früheren Bundesgebiet gänzlich, in den neuen Ländern zur Hälfte von Infratest

⁹ Hinsichtlich weiterführender Informationen zur Erhebungskonzeption und -realisation vgl. Schmidt, B.: Investitionsschnellerhebung, Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1997, S. 7 ff.

¹⁰ Weiterführende Informationen zur Erhebung vgl. Lüüs, H. P., Ehling, M., Eppmann, H., Götz, K.: Tourismusstichprobe 1992; Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1994.

Sozialforschung GmbH als computerunterstützte Telefonbefragung durchgeführt. Im Rahmen der Erhebung wurde von ca. 10 000 Befragten Auskunft über das Reiseverhalten von etwa 17 500 Personen gegeben.¹¹ SPSS wurde im Rahmen der Tourismusstichprobe zur Datenauswertung benutzt, Plausibilitätskontrollen dagegen wurden ausschließlich bei Infratest durchgeführt. Neben den klassischen Häufigkeitsauswertungen wurden im Rahmen der Datenanalyse eine Reihe verschiedener **Logit-Modelle** berechnet und ausgewertet, um festzustellen, welche persönlichen Merkmale für das Reiseverhalten bestimmend sind.¹² Die Analyse des funktionalen Zusammenhangs zwischen einer abhängigen Variablen (z. B. Reishäufigkeit) und (mehreren) erklärenden Variablen (z. B. Schulabschluß, Berufstätigkeit, Alter, Geschlecht), die alle *qualitativer* Natur sind, läßt sich in WinSPSS menügestützt über logarithmisch-lineare Modelle bestimmen, wobei diesbezüglich zwischen hierarchischen, allgemeinen und logit-linearen Modellen unterschieden wird.

- In den Jahren 1991/1992 wurde bei 6 400 Haushalten eine **"Zeitbudgeterhebung"** durchgeführt. Ziel der Erhebung war es, (1) eine geeignete Methode zur Erfassung der Zeitverwendung zu entwickeln und zu testen, (2) mit Hilfe der erhobenen Daten ein Satellitensystem Haushaltsproduktion zu den VGR aufzubauen und (3) Zeitverwendungsstrukturen von Bevölkerungsgruppen und Haushaltstypen zur Verfügung zu stellen.¹³ SPSS wurde im Rahmen der Auswertungsphase als **Ergänzung zu STATIS-BUND** v. a. für den letztgenannten Analysebereich eingesetzt. Um Unterschiede in den Zeitverwendungsstrukturen der einzelnen Bevölkerungs- und Haushaltsgruppen lokalisieren und quantifizieren zu können wurden aus dem umfangreichen Methodenspektrum von WinSPSS auch die Möglichkeiten zu **Cluster- und Faktorenanalyse** sowie **Diskriminanzanalyse** genutzt. Ebenfalls erwähnt werden muß der Einsatz von WinSPSS im Rahmen der Bearbeitung kurzfristiger Anfragen von Personen oder Institutionen zu spezifischen Zeitverwendungssachverhalten, die mit den bisherigen Veröffentlichungen nicht abgedeckt werden konnten. Eine ganze

¹¹ In den Ausführungen ist die vom Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen parallel durchgeführte Telefonbefragung auf der Grundlage des § 7 Abs. 2 BStatG, bei der u. a. auch der sog. Tagestourismus abgebildet wurde, nicht berücksichtigt.

¹² Vgl. Lüüs, H. P., Ehling, M., Eppmann, H., Götz, K.: Tourismusstichprobe 1992, a. a. O., S. 36.

¹³ Vgl. Statistisches Bundesamt (Hrsg.): Die Zeitverwendung der Bevölkerung. Methode und erste Ergebnisse der Zeitbudgeterhebung 1991/1992, Wiesbaden 1995, S. 2ff.

Reihe an Sonderauswertungen wurden auf diese Weise mit relativ wenig Aufwand in relativ kurzer Zeit durchgeführt.

- Die ersten Erfahrungen mit der **Aufbereitung und Auswertung von Längsschnittdaten** über WinSPSS können seit dem Jahr 1994 über das "**Europäische Haushaltspanel**" gewonnen werden. Das Europäische Haushaltspanel, das zunächst auf drei Befragungswellen (1994 bis 1996) angelegt¹⁴ und inzwischen um drei weitere Wellen (1997 bis 1999)¹⁵ aufgestockt wurde, soll international vergleichbare Informationen zur Entwicklung von Einkommen, Lebensstandard, Sozialversicherungssystemen, Armutsdynamik und zur sozialen Ausgrenzung geben. Dieselben 5 000 Haushalte, die sich schon 1994 an der Erhebung in Deutschland beteiligten, wurden auch in den Jahren 1995 und 1996 zu denselben Inhalten befragt. Gegenwärtig findet im Statistischen Bundesamt eine umfassende Validierung und Evaluation der ersten drei Wellen statt, aus denen differenzierte Erkenntnisse hinsichtlich der Eignung von SPSS für die Aufbereitung und Auswertung von Längsschnittdaten (hohe Komplexität des Datenmanagements) resultieren werden. Im Rahmen der Validierungsmaßnahmen ist gegenwärtig schon für die in der Erhebung abgebildeten Gesundheitsindikatoren eine **multiple Regressionsanalyse** unter Nutzung von SPSS durchgeführt worden, mit dem Ziel der empirischen Überprüfung, ob die abgebildeten Gesundheitsindikatoren ein valides Instrumentarium zur Erfassung von gesundheitlichem Befinden darstellen.¹⁶ Vergleichbare Analysen bieten sich auch in anderen Fragenbereichen an, insofern wird sich noch mehr Gelegenheit ergeben, dieses in SPSS menügestützt bereitgestellte Analysemodul breitgefächert zu testen.

- Im Jahr 1996 wurde mit einem Gesamtstichprobenumfang von gut 700 privaten Haushalten in ausgewählten Testgebieten der Länder Baden-Württemberg, Bayern, Brandenburg, Nordrhein-Westfalen und Sachsen eine **Testerhebung zur Neukonzeption der Laufenden Wirtschaftsrechnungen (LWR)** durchgeführt. Die primäre Aufgabe der Testerhebung bestand darin, zu überprüfen, ob das Erhebungsdesign und

¹⁴ Vgl. zur Darstellung des Erhebungskonzeptes Bechtold, S., Meyer, K.: Das Europäische Haushaltspanel. Darstellung des Erhebungskonzeptes und erster Ergebnisse; in: Wirtschaft und Statistik, 5/1996, S. 296 ff.

¹⁵ Für diese Wellen wird in Deutschland keine eigene Erhebung mehr durchgeführt, sondern das Europäische Haushaltspanel in das Sozio-oekonomische Panel (Soep) des DIW integriert.

¹⁶ Vgl. Stirner, H.: Soziale Indikatoren im Europäischen Haushaltspanel am Beispiel der Fragen zur Gesundheit; in: Wirtschaft und Statistik, 9/1997, S. 627 ff.

die Erhebungsinstrumente für die Erhebungspraxis geeignet sind, um die für die Neukonzeption vorgegebenen Ziele zu erreichen. Weiterhin sollten anhand der Ergebnisse und Erfahrungen der Testerhebung verbliebene Schwachstellen aufgedeckt und beseitigt werden.¹⁷ Ein zentrales Element der Neukonzeption war die Entwicklung eines DV-Systems, mit dessen Hilfe Datenerfassung und Datenplausibilisierung vollständig **PC-gestützt** im Dialogverfahren durchgeführt werden kann. Vorgesehen ist auch, für die Datenauswertung ein PC-Statistiksoftwarepaket einzusetzen. Um die Praxistauglichkeit der Vorgehensweise einschätzen zu können, war es nur konsequent, die **wissenschaftliche Begleituntersuchung** mit SPSS auszuwerten. Die Testerhebung hat gezeigt, daß das PC-gestützte integrierte Erhebungsmanagement für die Zukunft eine schnellere Ergebnisbereitstellung und wesentliche Effizienzgewinne bei der Statistikerstellung erwarten läßt¹⁸, zu dem auch SPSS einen gewichtigen Beitrag leisten wird.

¹⁷ Vgl. Gertkemper, F.: Ergebnisse der Testerhebung zum neuen Konzept der Laufenden Wirtschaftsrechnungen, in: *Wirtschaft und Statistik* 12/1997, S. 872 ff.

¹⁸ Vgl. Schwamb, H. J., Theis, T., Wein, E.: Integriertes Erhebungsmanagement, in: *Methoden, Verfahren, Entwicklungen*, Sonderausgabe 1/98, S. 9.

2 Das SPSS-Leistungsspektrum

SPSS ist das weltweit am häufigsten angewandte Programmsystem zur statistischen Datenanalyse.¹⁹ Den Grundstein legten im Jahr 1965 die beiden Studenten Norman Nie und Dale Bent vom Bereich Politikwissenschaft der Stanford Universität in San Francisco mit der Entwicklung der ersten Programmversion, die auf die Programmiersprache FORTRAN aufbaute und auf einem (Groß-)Rechner des Typs IBM 7090 lauffähig war. Die Programme der Pionierzeit, die damals noch als Lochkartenpäckchen vorlagen, wurden im Laufe der Jahre ständig technisch und fachlich-inhaltlich (Aufnahme neuer Statistikprozeduren) weiterentwickelt. Die erste PC Version von SPSS, lauffähig unter dem Betriebssystem MS-DOS, erschien 1983 unter dem Namen **SPSS/PC+** auf dem Markt und begründete die schnell voranschreitende weltweite Verbreitung des Programms. Anfang der 90er Jahre wurde dann die Windows-Version von SPSS eingeführt. Während die PC-Version des Programms im Prinzip eine abgespeckte Großrechner-Version war, so kommt **SPSS für Windows** – zumindest was das Handling des Programms betrifft – nahezu einer Neuentwicklung gleich. Einerseits deckt WinSPSS alle Möglichkeiten der Großrechner-Version ab, andererseits kann das Programm, von einigen wenigen Ausnahmen abgesehen, **benutzerfreundlich** ohne Kenntnis der Programmsyntax menügesteuert über entsprechende Dialogboxen bedient werden. Sarnow²⁰ spricht im Zusammenhang mit WinSPSS von einer "überzeugenden Portierungsarbeit", bei der man tatsächlich die "Leichtigkeit der Benutzung" vorfindet, die der Windows-Philosophie entspricht.

2.1 Der Funktionsumfang von SPSS 6.0.1

Das SPSS-Programmpaket ist **modular** aufgebaut. Es besteht aus einem sog. **Basis-Modul** und kann durch eine Vielzahl von **Zusatzmodulen** ergänzt werden. Grundlegende statistische Berechnungen und Analysen können schon mit dem Basismodul durchgeführt werden; darüber hinaus sind alle wichtigen Datenmanagement-Tools, Grafik-Tools sowie ein Reportgenerator enthalten. Um eine Vorstellung vom Funktionsumfang allein schon des Basismoduls zu geben, werden nachfolgend wichtige statistische

¹⁹ Hinsichtlich der nachfolgenden Ausführungen zur Entwicklung von SPSS vgl. Bühl, A., Zöfel, P.: *SPSS für Windows 6.1: Praxisorientierte Einführung in die moderne Datenanalyse*; Bonn, Paris, Reading, Mass. u. a. 1995, S. 17 ff.

²⁰ Sarnow, K.: Abstraktionsinstrumente. Statistiksysteme: von Shareware bis High-End; in: c't 1994, Heft 3, S. 186.

Datenaufbereitungs-, Auswertungs- und Präsentationsfunktionen tabellarisch aufgelistet²¹:

Tabelle 1: Wichtige Features des SPSS-Basismoduls

Statistische Kennziffern	Parametrische und nicht-parametrische Tests	Korrelations- und Regressionsanalyse	Varianzanalyse	Grafiken	Tabellen
arithmetisches Mittel	t-Test bei unabhängigen Stichproben	Pearson Korrelationskoeffizient	Einfaktorielle Varianzanalyse	Balken	Kreuztabellen
Median	t-Test bei abhängigen Stichproben	Spearman Rangkorrelationskoeffizient	Mehrfaktorielle Varianzanalyse	Linien	Reports (Tabellarische Berichte)
Modus	Chi-Quadrat-Anpassungs-Test	Kendall tau-b Koeffizient		Flächen	einfache Häufigkeitstabellen
Summe	Binominal-Test	Kendall tau-c Koeffizient		Kreis	
Varianz	Sequenzanalyse	Gamma		Hoch-Tief	
Standardabweichung	Kolmogorov-Smirnow Test	Somers d		Box-Plot	
Standardfehler	Mann-Whitney U	Kappa-Koeffizient		Fehlerdiagramm	
Minimum	Moses-Test	Kontingenzkoeffizient		Histogramm	
Maximum	Kolmogorov-Smirnow Z	Phi-Koeffizient		Stem-and-Leaf-Diagramm	
Spannweite	Wald-Wolfowitz	Cramers V		Scatterplot	
Schiefe der Verteilung	Kruskal-Wallis H	Lambda		Pareto-Diagramme	
Exzeß der Verteilung	Wilcoxon	Unsicherheitskoeffizient		Kontrollkarten	
Quartile	Vorzeichen-Test	Eta		Normalverteilungsplots	
benutzerdefinierte Percentile	Mc Nemar	Distanzmaße		Sequenz- und Zeitreihendiagramme	
5%-trim	Friedman	Partielle Korrelation			
Maximum-Likelihood-Schätzer	Kendall W	einfache lineare Regressionsanalyse			
Ausreißer	Cochran Q	Multiple lineare Regressionsanalyse			
Z-Score (standardisierte Werte)		Residualanalyse			

Ergänzt werden die dargestellten Funktionen um ein sehr breites Spektrum an Daten-/Dateien-Managementfunktionen. Hierzu zählen

²¹ Die Auflistung erhebt keinen Anspruch auf Vollständigkeit.

- zahlreiche optional wählbare **Import/Export-Formate** für die Datenmatrix (u. a. .dat (ASCII), .xls, .wk, .slk, .dbf Formate) und **Schnittstellen** zu anderen Programmen (u. a. zu Excel, Word für Windows), aufgrunddessen nicht nur Datensätze zur Weiterverarbeitung in andere Anwendungsprogramme *übertragen* werden können, sondern auch Informationen (v. a. Grafiken) zwischen Anwendungsprogrammen über **OLE** (Object Linking and Embedding) und **DDE** (Dynamic Data Exchange) *verknüpft* werden können.
- menügestütztes **Zusammenfügen von Datendateien**, wobei auf flexible Art und Weise sowohl Datendateien, die gleiche Variablen aber unterschiedliche Fälle enthalten, als auch Datendateien, die die gleichen Fälle aber unterschiedliche Variablen enthalten, zusammengefügt werden können.
- menügestütztes **Aggregieren von Daten** nach selbstdefinierten Break- (Gruppierungs-) Variablen. Als Aggregierungsfunktion stehen u. a. die Summe und der Mittelwert der Merkmalsausprägungen und die Anzahl der Fälle optional gewichtet oder ungewichtet zur Verfügung; darüber hinaus wird auf Wunsch auch die Anzahl der Fälle je Gruppierungsvariable (Schicht) mit ausgegeben.
- menügestützte **Datenselektion** (u. a. bedingte Auswahl, Zufallsstichprobe, Auflistung von Variablen und Fällen, Fallsortierung) und **Datenmodifikation** ((Bedingte) Berechnung neuer Variablen, Umcodierung von Werten, Rangtransformationen, Gewichtung von Fällen).
- menügestützte **Abrufbarkeit von Informationen** (Datensatzbeschreibung) über alle in der Datei enthaltenen Variablen, wie z. B. Variablenname, Variablenlabels, Merkmalsausprägungen, Variablenformat und Variablenposition.

Ergänzt werden diese Funktionen um – vom Aufbau und vom Layout her – typische Windows-Menüs ("Datei...öffnen, ...speichern, ...drucken etc". und "Bearbeiten ...Ausschneiden, ...kopieren, ...einfügen, ...löschen etc.") und durch eine kontextsensitive Hilfefunktion.

Nachfolgend ist das SPSS-Anwendungsfenster (Menüleiste) einschließlich ein für SPSS typisches Dateneditor-Fenster dargestellt:

Abbildung 1: Anwendungsfenster und Dateneditor-Fenster von SPSS 6.0

SPSS für Windows - [J:\msoffice\spss\lsc\auswert\daten5b...]

3:ma_94_w

ums_94

ums_95

1 336763 284990

2 318656 317716

3 524724 458557

4 429387 360331

5 352976 362722

6 3201145 3212079

7 997663 1010389

8 2788765 2473724

9 852422 1069986

10 13133004 12346435

11 397749 462309

12 940362 902740

13 7244 1054

14 581562 590212

15 2864494 2909480

SPSS-CHAID

8561 4209 0

0 0 0

28118 20342 0

1064 843 0

66360 65345 0

98915 0 98915 239253 113289

13900 0 13900 11557 0

214 0 214 100 0

9649 0 9649 16417 0

0 0 0 0 0

468 0 468 363 0

3638 22 3616 641 0

SPSS-Prozessor bereit

S 11 Ab 3 11/12 Bei 24,2 cm Ze 27 Sp 42 14:54

Neben dem SPSS-Basismodul gibt es noch mehrere **Zusatzmodule** sowohl zur Erweiterung des Spektrums der Datenauswertungs- und -analysemöglichkeiten als auch zur Erweiterung/Verbesserung der Ergebnispräsentationsmöglichkeiten²²:

1. **SPSS Professional Statistics:** Stellt u. a. Verfahren zur Messung von Ähnlichkeiten und Unterschieden in den Daten, zur Klassifikation der Daten und zur Identifikation zugrundeliegender Dimensionen zur Verfügung. Dieses Zusatzmodul ermöglicht die Clusteranalyse, die Faktorenanalyse, die Diskriminanzanalyse, die multidimensionale Skalierung, die Berechnung von Ähnlichkeits- und Unähnlichkeitsmaßen, die Zuverlässigkeitsanalyse sowie die gewichtete und zweistufige Kleinste-Quadrate-Regression.
2. **SPSS Advanced Statistics:** Stellt Verfahren für die logistische Regression, die log-lineare Analyse, die multivariate Varianzanalyse, die nichtlineare Regression mit Ne-

²² Hinsichtlich der nachfolgenden Ausführungen vgl. Norusis, M.: SPSS für Windows. Anwenderhandbuch für das Base System Version 6.0, 1994, S. iii ff. und Bühl, A., Zöfel, P.: SPSS für Windows 6.1: Praxisorientierte Einführung in die moderne Datenanalyse, a. a. O., S. 28.

benbedingungen, die Probit-Analyse, Cox-Regression, Überlebensanalyse nach der Kaplan-Meier- und der Sterbetafel-Methode zur Verfügung.

3. **SPSS Tables:** Zusatzmodul zur Erstellung von Tabellen. Darüber hinaus ermöglicht das Modul die Auswertung von Mehrfachantworten.
4. **SPSS Trends:** Ermöglicht Analysen zur Bedarfsvorhersage und von Zeitreihen. Eingesetzt werden u. a. Kurvenanpassungsmodelle, glättende Modelle, Methoden für die Schätzung autoregressiver Funktionen. In dieses Modul sind darüber hinaus auch verschiedene Verfahren zur Beseitigung von Missing-Werten integriert.
5. **SPSS Categories:** Beinhaltet Verfahren zur Conjoint-Analyse und zur optimalen Skalierung (z. B. die Korrespondenzanalyse).
6. **SPSS CHAID:** Unterstützt die tabellarische Analyse kategorialer Daten; erstellt Vorhersagemodelle und erzeugt leicht lesbare Baumstrukturdiagramme.²³
7. **SPSS LISREL 7:** Enthält Verfahren zur Analyse linearer Strukturgleichungen.

Darüber hinaus weist der Hersteller darauf hin, daß das System vom Benutzer erweitert werden kann, indem C- oder FORTRAN-Programme eingefügt werden können.²⁴

2.2 Der Dateiaufbau von SPSS 6.0.1

Der SPSS-Dateiaufbau besteht aus **vier verschiedenen Ebenen:**

1. **Daten-Dateien:** Daten-Dateien sind **tabellenförmig aufgebaut**, werden im **Daten-editor-Fenster angezeigt** und ähneln im Layout einem Arbeitsblatt in Tabellenkalkulationsprogrammen. Sämtliche (statistische) Prozeduren in SPSS beziehen sich gewöhnlich auf die Daten in der aktuell geöffneten Daten-Datei. In SPSS ist es nicht möglich, mehrere Daten-Dateien simultan zu öffnen und zu bearbeiten. Nachfolgend ist eine typische Daten-Datei dargestellt, wobei die **Zeilen die Erhebungsfälle** und die **Spalten die Erhebungsvariablen** repräsentieren.

²³ CHAID stellt im Gegensatz zu den anderen Modulen eine Stand-Alone-Software dar und benötigt daher nicht das Base-System.

²⁴ Vgl. diesbezüglich Norusis, M.: SPSS für Windows Anwenderhandbuch, a. a. O., S. iv.

Abbildung 2: SPSS-Daten-Datei angezeigt im SPSS-Dateneditor

	ordnr_wk	bran	ma_94	ma_94_o	ma_94_w	ma_95	ma_95_o	ma_95_w	umi
1	1001007	1	0	0	0	0	0	0	
2	1001008	1	0	0	0	0	0	0	
3	501006	1	3416	0	3416	2922	0	2922	
4	301006	2	98	0	98	90	0	90	
5	301005	3	696	0	696	675	0	675	
6	501003	1	14012	.	.	13204	.	.	
7	201004	3	515	0	515	489	0	489	
8	1001001	1	17396	0	17396	16478	0	16478	
9	301003	3	2021	0	2021	1919	0	1919	
10	501001	1	71709	0	71709	66751	0	66751	
11	501007	1	91	0	91	93	0	93	
12	501004	1	5138	0	5138	5008	0	5008	
13	1201003	1	46	46	0	3	3	0	
14	501005	1	3373	0	3373	3051	0	3051	
15	501002	1	2933	4	2929	2863	6	2857	
16	1201001	1	12876	12876	0	11044	11044	0	

2. **Ausgabe-Dateien:** Ausgabe- (Output-) Dateien sind Textdateien. In diesen Dateien werden die **jeweiligen Systeminformationen** und die **Ergebnisse statistischer Prozeduren** – soweit sie textlicher Art sind (wie z. B. deskriptive Statistiken, Kreuztabellen, Reports, Korrelationsmatrizen) – **hineingeschrieben**. Der Inhalt dieser Dateien kann bearbeitet, d. h. gespeichert, gelöscht, kopiert, ergänzt, gedruckt etc. werden. Nachfolgende Abbildung 3 zeigt ein SPSS-Output-Fenster, wobei exemplarisch am (gewichteten) Datensatz der FORCE-Erhebung eine Häufigkeitsauszählung über das Angebot von Lehrveranstaltungen in Unternehmen dargestellt ist.

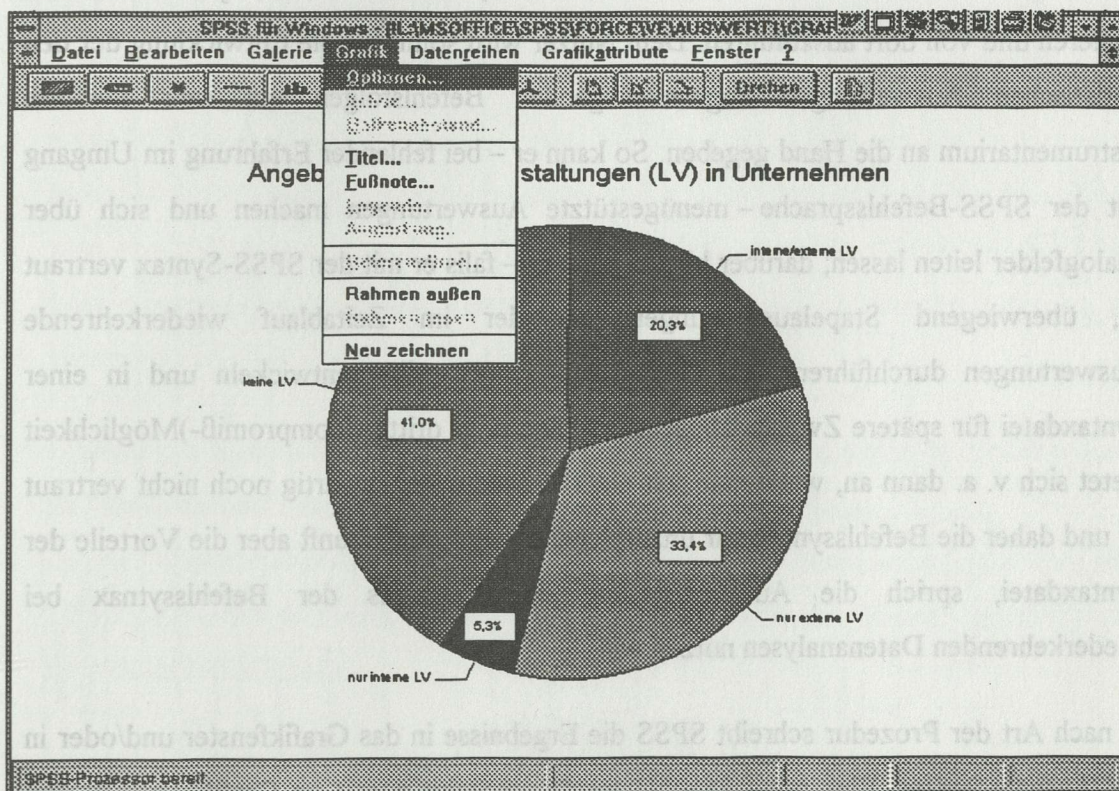
Abbildung 3: Darstellung einer SPSS-Ausgabe im Ausgabe-Fenster

SPSS für Windows - [Ausgabe1]					
Datei Bearbeiten Daten Transformieren Statistik Grafik Extras Fenster ?					
Pause Rollen Runden Glossar					
-> WEIGHT -> BY hr_fakto . -> FREQUENCIES -> VARIABLES=f1 .					
F1 Angebot von traditionellen Lehrveranstaltungen in Unternehmen					
Value Label	Value	Frequency	Percent	Valid Percent	Cum. Percent
Ja	1	114853	59,1	59,1	59,1
Nein	2	79378	40,9	40,9	100,0
	Total	194231	100,0	100,0	
Valid cases		194231	Missing cases 0		

Das Beispiel zeigt, daß im SPSS-Ausgabefenster neben den Ergebnissen der Häufigkeitsauszählung auch Systeminformationen (hier: Hinweis auf den gewichteten Datensatz "weight by hr_fakto") und die der Häufigkeitsauszählung zugrundeliegende Befehlssyntax (hier: "Frequencies Variables = f1") ausgewiesen wird.

3. **Syntax-Dateien:** Syntax-Dateien sind gleichzusetzen mit dem **Befehlseditor**. Hier können unter Nutzung der SPSS-Steuersprache eigene Befehlsfolgen unmittelbar erstellt (Programmierung) oder aber über Dialogfelder ausgewählte Befehle hineinkopiert, von dort ausgeführt, bearbeitet und auch dauerhaft gespeichert werden.
4. **Grafik-Dateien:** In diesen Dateien können **Grafiken und Plots bearbeitet und gespeichert werden**, die auf der Grundlage von Daten aus einer Daten-Datei in SPSS erstellt wurden.

Abbildung 4: Darstellung einer SPSS-Grafikdatei



Eine über das sog. SPSS-Grafik-Karussell erstellte Grafik kann nicht nur als Grafik-Datei gespeichert und bearbeitet werden, sondern auch in andere Anwendungen (Winword, Powerpoint etc.) eingebunden werden.

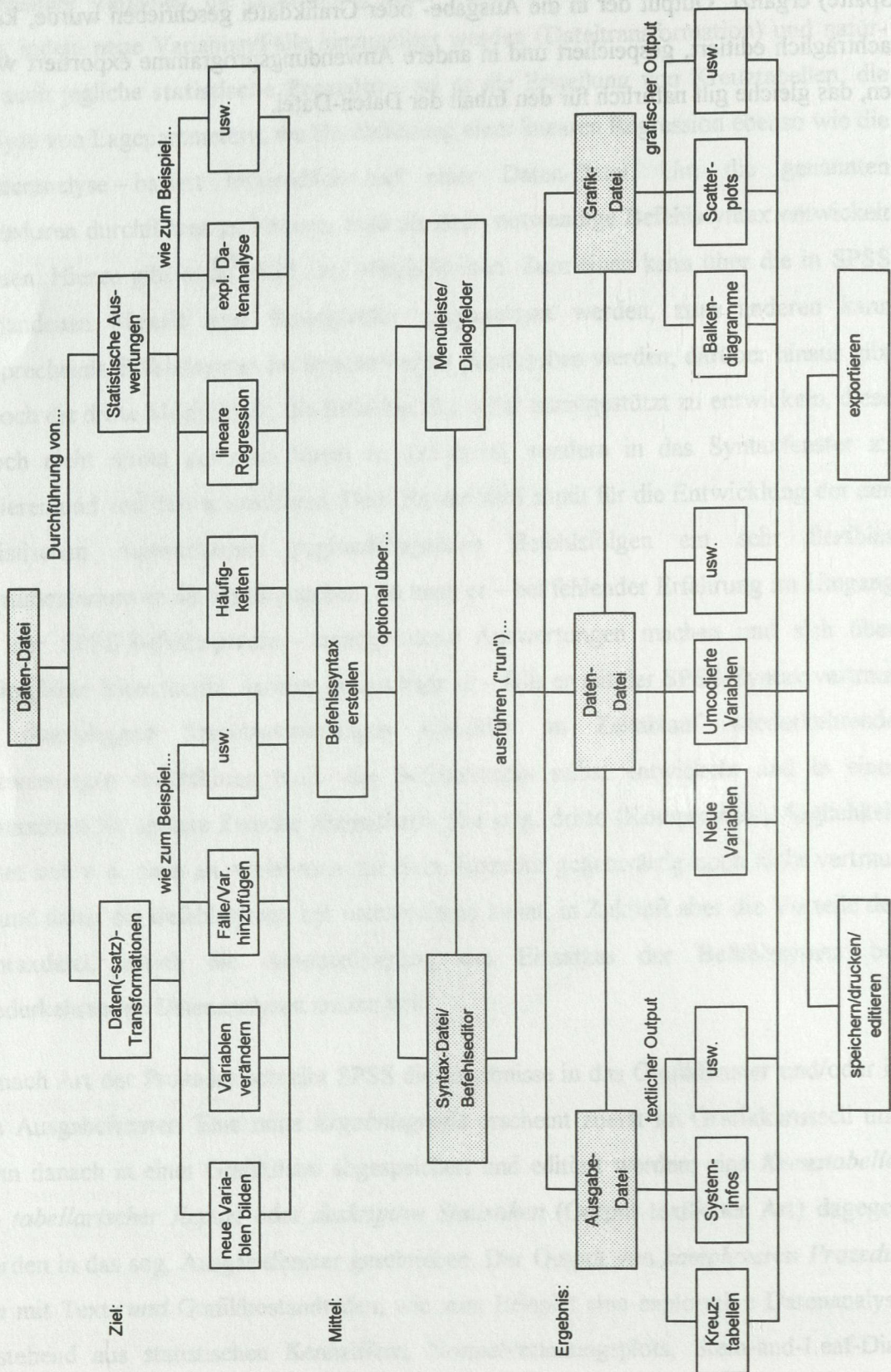
Abbildung 5 auf der Seite 24 soll Einsatz und Zusammenspiel der vier Datei-Ebenen nochmals deutlich machen:

Grundlage jeglicher **Variablen- und Dateitransformation** – wie z. B. die Veränderung bestehender Variablen, die Berechnung neuer Variablen, das Verschmelzen zweier Dateien, indem neue Variablen/Fälle hinzugefügt werden (Dateitransformation) und natürlich auch jegliche **statistische Prozedur** – sei es die Erstellung von Kreuztabellen, die Analyse von Lageparametern, die Durchführung einer linearen Regression ebenso wie die Clusteranalyse – basiert letztendlich auf einer Daten-Datei. Um die genannten Prozeduren durchführen zu können, muß die dazu notwendige Befehlssyntax entwickelt werden. Hierzu gibt es in SPSS drei Möglichkeiten: Zum einen kann über die in SPSS vorhandenen **Menüs und Dialogfelder** vorgegangen werden; zum anderen kann entsprechende Befehlssyntax im **Syntaxfenster** geschrieben werden; darüber hinaus gibt es noch die dritte Möglichkeit, die Befehlssyntax zwar menügestützt zu entwickeln, diese jedoch nicht direkt aus dem Menü zu aktivieren, sondern in das Syntaxfenster zu kopieren und von dort auszuführen. Dem Nutzer wird somit für die Entwicklung der den statistischen Auswertungen zugrundeliegenden Befehlsfolgen ein sehr flexibles Instrumentarium an die Hand gegeben. So kann er – bei fehlender Erfahrung im Umgang mit der SPSS-Befehlssprache – menügestützte Auswertungen machen und sich über Dialogfelder leiten lassen; darüber hinaus kann er – falls er mit der SPSS-Syntax vertraut ist, überwiegend Stapelauswertungen und/oder im Zeitablauf wiederkehrende Auswertungen durchführen muß – die Befehlssyntax selbst entwickeln und in einer Syntaxdatei für spätere Zwecke abspeichern. Die o. g. dritte (Kompromiß-)Möglichkeit bietet sich v. a. dann an, wenn man mit einer Prozedur gegenwärtig noch nicht vertraut ist und daher die Befehlssyntax nur unzureichend kennt, in Zukunft aber die Vorteile der Syntaxdatei, sprich die Automatisierung des Einsatzes der Befehlssyntax bei wiederkehrenden Datenanalysen nutzen will.

Je nach Art der Prozedur schreibt SPSS die Ergebnisse in das Grafikfenster und/oder in das Ausgabefenster. Eine reine *Ergebnisgrafik* erscheint zuerst im Grafikkarussell und kann danach in einer Grafikdatei abgespeichert und editiert werden; eine *Kreuztabelle*, ein *tabellarischer Report* oder *deskriptive Statistiken* (Output textlicher Art) dagegen werden in das sog. Ausgabefenster geschrieben. Der Output von *komplexeren Prozeduren* mit Text- und Grafikbestandteilen, wie zum Beispiel eine explorative Datenanalyse bestehend aus statistischen Kennziffern, Normalverteilungsplots, Stem-and-Leaf-Dia-

grammen, Histogrammen erscheint je nach Art der Ergebnisse im Ausgabe- wie auch im Grafikfenster. Datentransformationen beziehen sich unmittelbar auf die Daten-Datei. Eine neu gebildete Variable zum Beispiel wird direkt ins Datenfenster übernommen, d. h. die Datenmatrix – bestehend aus Fällen und Variablen – wird um die neue Variable (Spalte) ergänzt. Output der in die Ausgabe- oder Grafikdatei geschrieben wurde, kann nachträglich editiert, gespeichert und in andere Anwendungsprogramme exportiert werden, das gleiche gilt natürlich für den Inhalt der Daten-Datei.

Abbildung 5: Zusammenspiel der vier verschiedenen Dateiebenen bei der Durchführung von statistischen Prozeduren



3 Beurteilung von SPSS in der praktischen Anwendung

Um die Eignung von SPSS im praktischen Einsatz im Statistischen Bundesamt strukturiert bewerten zu können, ist es notwendig, zunächst Kriterien zu entwickeln, an denen die Beurteilung ausgerichtet werden kann. Dies wird im Gliederungspunkt 3.1 vorgenommen. Im Anschluß daran wird im Gliederungspunkt 3.2 die eigentliche Beurteilung von SPSS anhand dieser Kriterien durchgeführt.

Im Zusammenhang mit der Beurteilung von SPSS muß berücksichtigt werden, daß

- sich die **Leistungsfähigkeit der Hardware** im Zeitablauf enorm verbessert hat, was eine Bewertung von SPSS im praktischen Einsatz erschwert, da manche Problemfelder *hardwarespezifisch* sind.
- **verschiedene SPSS Versionen** im Einsatz waren (DOS-Version, Windows Version) und die – sich auch gegenwärtig noch im Einsatz befindliche – Version 6.0.1 schon seit geraumer Zeit durch die Versionen 7.0 und 7.5 abgelöst wurde, die unter Windows 95 bzw als Netzwerk-Version unter Windows NT laufen und die in wichtigen Modulen grundlegende Neuerungen aufweisen. Da im Statistischen Bundesamt bis zum Abschluß der Vernetzung kein Umstieg auf eines der beiden Betriebssysteme erfolgt, konnte mit den neueren Versionen von SPSS keinerlei praktische Erfahrungen gesammelt werden.
- Erhebungen nach § 7 BStatG sich in gewissen Punkten **von anderen Erhebungen der statistischen Ämter unterscheiden**. Hervorgehoben werden muß die Freiwilligkeit der Auskunfterteilung und die Beschränkung des Erhebungsumfangs auf maximal 10 000 Erhebungseinheiten ("kleine Stichproben"), was das Bewertungsspektrum letztendlich einschränkt.
- hier nur die im praktischen Umgang mit SPSS gemachten Erfahrungen wiedergegeben werden können; der Anspruch auf eine umfassende wissenschaftliche Bewertung der Software kann daher nicht erhoben werden, da nur ein Teil des SPSS-Spektrums an technischen und methodischen Möglichkeiten in der Praxis eingesetzt wurde.²⁵

²⁵ So wurden zum Beispiel die Analysemodule Chaid, Lisrel und Zeitreihen aufgrund fehlender praktischer Einsatzfelder nicht genutzt.

Darüber hinaus wären zu diesem Zweck auch umfassende Vergleichsrechnungen für die Beurteilung der numerischen Präzision und der Angemessenheit der statistischen Prozeduren notwendig.

- die Bewertung zwangsläufig **subjektive Bestandteile** enthält, da nicht zuletzt ein so wichtiges Kriterium wie die "Benutzerfreundlichkeit" eines Programms nur eingeschränkt an objektiven Gesichtspunkten festgemacht werden kann.

3.1 Beurteilungskriterien

Um eine möglichst strukturierte Beurteilung vornehmen zu können, ist es sinnvoll, die Prüfkriterien in geeignete Kriteriengruppen zusammenzufassen. Orientiert an Haag/Kieser²⁶, Hilbert et al.²⁷ und der Universität von Manchester²⁸ wurden folgende Beurteilungskriteriengruppen gebildet:

1. Handhabung

Das Kriterium "Handhabung" wird hierbei im weiteren Sinne verstanden, d. h. nicht nur die "**Benutzerfreundlichkeit**" – wie z. B. die Verständlichkeit und Übersichtlichkeit der Programmenüs und Dialogfelder oder der für eine angemessene Nutzung notwendige Grad an Programm-(Vor-)Kenntnissen – wird hierunter subsumiert, sondern auch Kriterien wie die "**Systemvoraussetzungen/Systemcharakteristika**" (u. a. Hardware-Voraussetzungen, Betriebssystemvoraussetzungen, Schnittstellen mit externen Programmen, Programmstruktur), die "**Schnelligkeit/Laufzeiteffizienz**", die "**Stabilität des Systems**" (Systemabsturzanfälligkeit) und die Möglichkeiten des "**Datenmanagements**" (u. a. Möglichkeiten des Datenimports und -exports und der Datentransformation, Vorhandensein einer Variablen- bzw. Fallzahlenbegrenzung) und das "**File-Handling**" (Angemessenheit der Datensatzorganisation).

²⁶ Haag, U.; Kieser, M.: Statistische Auswertungssysteme: eine Einführung in ihre Anwendung, Konstruktion und Bewertung; Stuttgart, New York, Jena 1992, S. 137 ff.

²⁷ Hilbert, A., Bankhofer, U., Elsholz, B., Niederöcker, U.: Statistiksoftware unter Windows – Eine Marktanalyse, Arbeitspapiere zur Mathematischen Wirtschaftsforschung, Heft 124, Universität Augsburg 1995, S. 3 ff.

²⁸ MIDAS, Manchester Computing University of Manchester (Campbell, M.): Comparative Analysis of Statistical Packages and Data Bases for Statistics, Material zum TES-Course SUP-104-E, Course Slides, Day 2, S. 5-12.

2. Spektrum der statistischen Prozeduren

Kernelement jeder statistischen Software sind die bereitgestellten statistischen Prozeduren. Unter Beurteilungsgesichtspunkten wurde hier der Schwerpunkt nicht auf die generelle Methodenvielfalt gelegt, sondern darauf geachtet, inwieweit die in der herkömmlichen Statistikproduktion im Rahmen der Datenaufbereitungs- und -auswertungsphase benötigten Methoden und Instrumente **vorhanden und anwendbar** sind. Was die **numerische Zuverlässigkeit** – sprich die Genauigkeit/Angemessenheit der Ergebnisse der statistischen Prozeduren – betrifft, werden aufgrund fehlender praxisbezogener Vergleichsmöglichkeiten nur Prüfergebnisse von externen Studien, die zu diesem Zweck durchgeführt wurden, wiedergegeben.

3. Ergebnisdarstellung

Hierunter wird die Beurteilung des **Spektrums an Möglichkeiten zur Präsentation von Ergebnissen** verstanden; die **Bedienbarkeit** der hierfür zur Verfügung stehenden Menüs; die **Qualität des Outputs** und die Frage, inwieweit die Ergebnisse auch in andere Programme **exportiert** werden können. Das Hauptaugenmerk wird dabei auf die Möglichkeiten der **Tabellenerstellung** (Tables-Modul) und der **Grafikerstellung** (Grafik-Modul) gelegt.

4. Produktsupport und Serviceleistung

Neben der Beurteilung der spezifischen Programmeigenschaften ist auch von Interesse, welche **Schulungen** von SPSS (Frequenz, Inhalte, etc.) angeboten werden und inwieweit eine **Hot-Line zur Problembewältigung** existiert.

In der nachfolgenden Tabelle 2 sind die wichtigsten Beurteilungskriterien überblicksartig aufgelistet:

²⁹ Vgl. Norups, M.: SPSS für Windows Anwenderhandbuch, v. O. S. 1.

³⁰ So zum Beispiel die Touristenzusatzgewinne 1992.

Tabelle 2: Praxisbezogene Beurteilungskriterien

Beurteilungskriterium	Erläuterungen
I. Handhabung <ul style="list-style-type: none"> • Systemvoraussetzungen/Systemcharakteristika • Schnelligkeit/Laufzeiteffizienz/Stabilität des Systems • Daten(satz)-Management • Benutzerfreundlichkeit 	<ul style="list-style-type: none"> - Hardware- und Software- (Betriebssystem-) Voraussetzungen - Kompatibilität/ Schnittstellen mit anderen Programmen - Programmstruktur - der statistischen Prozeduren/Ergebnisdarstellung - der programmiertechnischen Umsetzung - Speichermöglichkeit - System-Absturzanfälligkeit - Maximalzahl von Fällen/Variablen - Datenimport und -export - Datenformate - Datentransformation - Datenselektion - Datensatzorganisation - Notwendige (Vor-)Kenntnisse für die Nutzung - Verständlichkeit/Übersichtlichkeit der Programmenüs/Dialogfelder usw. - Wahlmöglichkeiten bei der Befehlseingabe - Handbuch, Hilfefunktion
II. Spektrum der statistischen Prozeduren	<ul style="list-style-type: none"> - Methoden zur Datenaufbereitung, Datenauswertung und Datenanalyse - numerische Zuverlässigkeit
III. Ergebnisdarstellung	<ul style="list-style-type: none"> - Tabellen-Modul - Grafik-Modul - Integrierbarkeit von SPSS-Tabellen und Grafiken in externe Programme
IV. Produktsupport	<ul style="list-style-type: none"> - Qualität der Schulungen - Hot-Line für die Problembewältigung

Um eine einigermaßen konsistente Beurteilung vornehmen zu können, beschränkt sich die Bewertung – obwohl anfang der 90er Jahre noch einige § 7-Erhebungen mit der SPSS DOS-Version durchgeführt wurden – auf die seit 1993 im Statistischen Bundesamt im Einsatz befindliche **Windows Version von SPSS** (SPSS für Windows Version 6.0.x). Was die **Hardware-Ausstattung** betrifft, wird der Beurteilung als **Minimalstandard** ein **PC mit 486 DX Prozessor (50 Mhz) mit 8 MB Arbeitsspeicher (RAM)** zugrundegelegt.

3.2 Kriterienbezogene Bewertung

3.2.1 Handhabung

3.2.1.1 Systemvoraussetzungen/Systemcharakteristika

Entsprechend dem Stand der (Hardware-)Entwicklung empfahl der Hersteller im Jahr 1994 in seinem zur Version 6.0 erschienenen Anwenderhandbuch als Hardware-Ausstattung (1) einen 80486/DX Prozessor mit 33 Mhz (mindestens jedoch einen 80386 Prozessor) und (2) 8 MB RAM (mind. jedoch 4 MB). Darüber hinaus sollte Windows 3.1 im erweiterten Modus mit einer permanenten Auslagerungsdatei von 10 MB laufen und ein SMARTDrive mit einer Windows-Cache-Größe von mindestens 512 KB installiert werden.²⁹

Diese Systemangaben müssen jedoch im Zusammenhang mit dem Stand der Entwicklung der Hardware zum Zeitpunkt der Markteinführung der Version 6.0 gesehen werden. Generell gilt natürlich, je besser die Hardwareausstattung (Arbeitsspeicher, Prozessor-Leistung, Festplattenkapazität) desto benutzerfreundlicher und effizienter die Bedienung und die Laufeigenschaften und desto besser auch die Bewertung des Programms. Der Praxis-einsatz hat gezeigt, daß es – um das gesamte Spektrum des Programms auch noch bei größeren Datensätzen in effizienter Weise nutzen zu können und nicht ständig am Ressourcenlimit zu arbeiten und sich der Gefahr von Systemabstürzen auszusetzen – empfehlenswert ist, den Arbeitsspeicher auf zumindest 16 MB (besser 32 MB) aufzustocken und einen Pentium-Prozessor einzusetzen. Nicht vernachlässigt werden sollte jedoch der Sachverhalt, daß einige der im 1. Gliederungspunkt genannten § 7-Erhebungen³⁰ mit – gemessen am heutigen Standard – leistungsmäßig sehr eingeschränkten PCs durchgeführt wurden, wobei aus der Sicht der technischen Abläufe trotzdem die **positiven Eindrücke** bei weitem überwogen.

Was die **Kompatibilität bzw. die Schnittstellen** mit anderen Anwender-Programmen betrifft, so bietet SPSS vielfältige Möglichkeiten. Unproblematisch gestaltet sich der Import/Export von Dateien, die im ASCII-Format abgespeichert sind. Darüber hinaus hat SPSS aber auch Schnittstellen für den Datenaustausch mit marktführenden Tabellen-

²⁹ Vgl. Norusis, M.: SPSS für Windows Anwenderhandbuch, a. a. O., S.1.

³⁰ So zum Beispiel die Tourismusstichprobe 1992.

kalkulationsprogrammen und Datenbanken wie zum Beispiel MS Excel, Lotus 123 und dBase. SPSS wurde daher auch in der einschlägigen Literatur bezüglich des Kriteriums "Systemoffenheit" sehr positiv beurteilt. So vergaben zum Beispiel *Bamberg/Baur*³¹ bei ihrem Vergleich von 16 verschiedenen Statistikpaketen mit allgemeinem Methodenspektrum SPSS die bestmögliche Bewertung.

Die **flexible Programmstruktur** von SPSS bringt v. a. unter Kostengesichtspunkten Vorteile mit sich. Durch den **modularen Aufbau** des Programms muß nicht ein Statistik**gesam**tpaket gekauft werden, es besteht vielmehr die Möglichkeit, neben dem Basismodul, das schon die grundlegenden Funktionen zur Datentransformation, statistischen Auswertung und Grafikerstellung enthält, nur die (Zusatz-) Module³² anzuschaffen oder nachträglich zu ergänzen, die entsprechend dem jeweiligen Arbeitsgebiet auch wirklich benötigt werden.

Insgesamt gesehen kann SPSS daher unter dem Gesichtspunkt "Systemvoraussetzungen/Systemcharakteristika" sehr positiv beurteilt werden. Zu empfehlen ist jedoch – um das gesamte Leistungsspektrum von SPSS (Auswertungs-, Datentransformations-, Ergebnisdarstellungsmöglichkeiten) möglichst effizient nützen zu können – einen, gemessen am heutigen Stand der technischen Entwicklung möglichst leistungsfähigen PC mit einem reichlich ausgestatteten Arbeitsspeicher einzusetzen.

3.2.1.2 Stabilität des Systems und Laufzeiteffizienz/Schnelligkeit der Umsetzbarkeit statistischer Prozeduren

Die Beurteilung der Stabilität des Systems – sprich die **Speichereffizienz** und das Auftreten von unvorhersehbaren **Systemabstürzen** – hängt unmittelbar von der entsprechenden Hardware-Ausstattung ab. Generell kann festgehalten werden, daß SPSS sehr **hohe Ansprüche an die Arbeitsspeicher- und Festplattenkapazität** des PCs stellt. Die Gefahr von System- und Prozedurabbrüchen kann jedoch deutlich eingeschränkt werden, wenn – wie oben schon betont wurde – mit einem möglichst **großen Arbeitsspeicher** gearbeitet und SPSS eine möglichst **große permanente Auslagerungsdatei** zur Verfügung gestellt wird. Dies setzt letztendlich auch ausreichend **freie Festplattenkapazitäten** voraus.

³¹ Bamberg, G., Baur, F.: Statistik, 8. Auflage, München 1993, S. 263.

³² Vgl. die auf den Seiten 18 und 19 beschriebenen Zusatzmodule.

Konkrete Größenangaben zur Hardwareausstattung sind nicht sinnvoll, da diese unmittelbar vom Auswertungsgegenstand – sprich von der Größe des zu analysierenden Datensatzes (Anzahl der Fälle *und* Variablen) – determiniert wird. Für die Auswertungszwecke der § 7-Erhebungen waren nicht selten PCs mit lediglich 8 MB RAM und 66 Mhz Prozessorleistung im Einsatz, wobei SPSS unter Kosten-/ Nutzenaspekten fast durchgängig positiv bewertet werden konnte. Diese Bewertung sollte auch nicht mit dem Argument der für § 7-Erhebungen typischen kleinen Datensätzen in Frage gestellt werden, da dies zwar für die *Fallzahl* nicht jedoch für die Anzahl der *Variablen* gilt. Die Erfahrungen haben nämlich gezeigt, daß die Speicher- und Laufzeiteffizienz des Programms mindestens ebenso stark von der Anzahl der Fälle wie von der Anzahl der Variablen (Auswertungssachverhalte) im Datensatz determiniert wird, wobei zumindest einige Datensätze von § 7-Erhebungen (vgl. das EU-Haushaltspanel oder die Zeitbudgeterhebung) eine auch im Hausvergleich eher **überproportional große Variablenzahl** beinhalteten.

Sind bestimmte statistische Prozeduren mit den SPSS zur Verfügung gestellten Ressourcen nicht durchführbar, informiert das Programm i. d. R. über das Problem und bricht die Prozedur ab. Gelegentlich besteht die Möglichkeit über eine (temporäre) Anhebung des **virtuellen Arbeitsspeichers** (vorab eingestellt sind 512 KB) entsprechende Prozeduren doch noch erfolgreich durchführen zu können.

Im Zusammenhang mit den Prozedur-/Systemabbrüchen muß betont werden, daß diese Probleme über alle im 1. Gliederungspunkt genannten § 7-Erhebungen gewöhnlich nur in **Ausnahmefällen bei sehr speicherintensiven Prozeduren** (z. B. bei umfangreichen Variablentransformationen) aufgetreten sind. Im übrigen weist der Hersteller darauf hin, daß es nahezu keine programmtechnischen Beschränkungen hinsichtlich der Größe des Datensatzes gibt³³; Probleme mit der Programmstabilität aber auch im Hinblick auf das Kriterium "Laufzeiteffizienz" haben ihre Ursache folglich unmittelbar in einer unzureichenden Hardwareausstattung, wobei nicht vergessen werden darf, daß aufgrund der rasanten Weiterentwicklung in diesem Bereich sich auch die hardwarebedingten Grenzen der Einsatzmöglichkeiten des Programms kontinuierlich nach oben verschieben werden.

³³ Vgl. auch die Ausführungen im Gliederungspunkt 3.2.1.3

Aussagen zum **Laufzeitverhalten des Programms** können ebenfalls nur sehr schwer gemacht werden, da hierfür mehrere Faktoren verantwortlich sind, wobei das **Zusammenwirken** der Faktoren und nicht deren isolierte Existenz den Zeitbedarf für eine Arbeitsprozedur determinieren. Oben wurde schon auf die in der Praxis gewonnenen Erfahrungen hingewiesen, daß bei gegebener Hardwareausstattung und Systemumgebung nicht nur die **Anzahl der Fälle** die Laufzeit von statistischen Prozeduren entscheidend beeinflußt, sondern auch die **Anzahl der im Datensatz befindlichen Variablen**, und zwar unabhängig davon, ob diese Variablen in die entsprechende Prozedur direkt einbezogen sind oder nicht. Nachfolgend wird anhand ausgewählter Prozeduren ein kleiner Einblick in das Laufzeitverhalten von SPSS gegeben (vgl. Tabelle 3).³⁴

Tabelle 3: Laufzeitverhalten von SPSS

§ 7-Datensatz	Anzahl der Fälle im Datensatz	Anzahl der Variablen im Datensatz	Statistikprozedur	Ausführungszeit
FORCE-Vorerhebung	9 178	52	einfache Häufigkeitsauszählung ¹	3 sec.
FORCE-Haupterhebung	4 081	386	einfache Häufigkeitsauszählung ¹	10 sec.
Investitionsschnellerhebung (Originaldatensatz)	6 166	79	tabellarischer Bericht ²	9 sec.
Investitionsschnellerhebung (aufgestockter Datensatz)	86 324	79	tabellarischer Bericht ²	180 sec.
Europäisches Haushaltspanel	9 918	830	explorative Datenanalyse ³	61 sec.
Zeitbudgeterhebung	32 326	4 213	Datensatzsortierung ⁴	2 700 sec.
Zeitbudgeterhebung	32 326	810	Datensatzsortierung ⁴	391 sec.
Zeitbudgeterhebung	32 326	4 213	Kreuztabellierung ⁵	533 sec.
Zeitbudgeterhebung	32 326	810	Kreuztabellierung ⁵	76 sec.
Zeitbudgeterhebung	32 326	4 213	einfache Häufigkeitsauszählung ⁶	443 sec.
Zeitbudgeterhebung	32 326	810	einfache Häufigkeitsauszählung ⁶	78 sec.

¹ **Häufigkeitsauszählung** der dichotomen Variablen "Teilnahme von Mitarbeiter(innen) des Unternehmens an Weiterbildung im weiteren Sinne" in Form von (1) Austauschprogrammen, (2) Coaching, (3) Einarbeitung neuer Mitarbeiter, (4) Job-Rotation, (5) Lernstatt, (6) Fachmessen, (7) Qualitätszirkel. Variablenausprägungen: 1 = "Ja"; 2 = "Nein".

² **Auszählung und (Tabellen-)Darstellung** der Anlage-, Ausrüstungs- und Bauinvestitionen 1995 (**Summe, Mittelwert, Fallzahl**) im früheren Bundesgebiet und in den neuen Ländern nach Beschäftigtengrößenklassen.

³ Berechnung und Darstellung von wichtigen **Lageparametern** (arithmetisches Mittel, Median, 5%-trim, 95%-Vertrauensintervall, Minimum, Maximum, Varianz, M-Schätzer, Percentile, Extremwert, Schiefe und Kurtosis der Verteilung); Darstellung der Häufigkeitsverteilung in Form eines **Stem-and-Leaf-Diagramms** und grafische Darstellung des **Normalverteilungsplots** (auch trendbereinigt), des **Box-Plots** und des **Histogramms** der Variablen "persönliches Einkommen".

⁴ **Sortiervariable**: Höhe des Haushaltseinkommens (klassifiziert).

⁵ **Kreuztabelle** der Variablen: "Haushaltseinkommen (klassifiziert)" nach "8 Haushaltstypen".

⁶ **Einfache Häufigkeitsauszählung** der Variablen "Anzahl der Erwerbstätigen im Haushalt (klassifiziert)".

³⁴ Dargestellt sind die reinen Laufzeiten, d. h. nachträglich notwendige (Um-)Formatierungen für Darstellungs- und Präsentationszwecke sind nicht berücksichtigt.

Zugrundegelegt wurde ein PC mit einem Pentium 133 Mhz Prozessor und 16 MB RAM, wobei 3 MB RAM exklusiv für SPSS reserviert wurden. Darüber hinaus wurde SPSS eine permanente Auslagerungsdatei in Höhe von 51 MB auf der Festplatte zur Verfügung gestellt.

Die Beispiele machen deutlich, daß Standardprozeduren wie z. B. einfache Häufigkeitsauszählungen, Kreuztabellen aber auch aufwendigere tabellarische Berichte unter Einbeziehung einer Gruppierungsvariable oder umfangreichere explorative Analysen bei der neben vielen (univariaten) Statistiken auch eine ganze Reihe an Grafiken (Box-Plots, Stem-and-Leaf-Diagramm, Histogramm, Normalverteilungsplots) aufgebaut werden, bei Datensätzen mit relativ geringer Fallzahl (unter 10 000) und Variablenzahl häufig eine Sache von wenigen Sekunden ist, aber auch bei Datensätzen mit einer deutlich größeren Fallzahl (vgl. z. B. den auf 86 000 Einheiten aufgestockten Datensatz der Investitions-schnellerhebung) noch immer im unteren Minutenbereich vonstatten gehen kann.

Aus den Beispielen geht jedoch auch hervor, daß das Laufzeitverhalten des Programms sehr stark durch die **Variablenanzahl** im Datensatz determiniert wird, wie – zum einen – der Vergleich der Laufzeiten derselben Häufigkeitsauszählung im Datensatz der FORCE-Vorerhebung (52 Variablen) und im Datensatz der Haupterhebung (386 Variablen) und – zum anderen – die Beispielsprozeduren mit unterschiedlich breiten Datensätzen der Zeitbudgeterhebung³⁵ zeigen. Das Zeitbudget-Beispiel macht deutlich, daß eine statistische Prozedur wie die Bildung einer einfachen Häufigkeitstabelle oder einer zweidimensionalen Kreuztabelle bei einer Fallzahl von 32 000 lediglich etwas mehr als eine Minute aber auch sieben bis acht Minuten in Anspruch nehmen kann, je nachdem, ob die Prozedur mit dem kompletten Datensatz (4 213 Variablen) oder aber mit einem auf 810 Variablen reduzierten Datensatz durchgeführt wird. Allein schon aus diesem Grund empfiehlt es sich, im Vorfeld statistischer Auswertungen den Datengrundfile an den **analysespezifischen Bedarf anzupassen**, d. h. diejenigen Variablen und Fälle zu selektieren, die für die jeweiligen Auswertungszwecke notwendig sind. SPSS stellt für diese Zwecke **komfortable Datenmanagement-Tools** zur Verfügung.

Vor dem Hintergrund der Abhängigkeit des SPSS-Laufzeitverhaltens von der Variablenzahl im Datensatz muß auch auf den Sachverhalt hingewiesen werden, daß eine Vielzahl

³⁵ Der Grundfile der Zeitbudgeterhebung besteht aus 32 326 Fällen und 4 213 Variablen und umfaßt 240 MB!

der von den statistischen Ämtern durchzuführenden Erhebungen insbesondere im **Unternehmensbereich** zwar hohe Fallzahlen aufweisen, im Vergleich zum EU-Panel oder zur Zeitbudgeterhebung jedoch relativ **wenige Merkmale** umfassen. Insofern kann davon ausgegangen werden, daß SPSS nicht nur ein adäquates Auswertungs- und Analyseinstrument für ("kleine") § 7-Erhebungen ist, sondern durchaus auch für eine ganze Reihe anderer Erhebungen der statistischen Ämter **effiziente Einsatzmöglichkeiten** bieten kann.

Im Zusammenhang mit der Laufzeiteffizienz darf darüber hinaus nicht vergessen werden, daß die **Schnelligkeit der programmiertechnischen Umsetzung** gewünschter statistischer Prozeduren mindestens ebenso bedeutend ist wie die reinen PC-Laufzeiten. Hier hat SPSS den Vorteil, daß aufgrund der Menüunterstützung die gängigen statistischen Auswertungen ohne zeitraubende und auf grundlegende (Vor-) Kenntnisse aufbauende Programmierarbeiten realisiert werden können. Aus diesem Grund kann die Konzeption einschließlich der technischen Umsetzung gewünschter Auswertungen (Ergebnistabellen, Grafiken etc.) von *einer* Person durchgeführt werden und macht keine Arbeitsteilung zwischen DV-Spezialisten und Fachstatistiker erforderlich, was allein schon aufgrund der **geringeren Koordinations- und Abstimmungserfordernisse** und der kürzeren Informations-, Entscheidungs- und Handlungswege weniger zeitintensiv ist.

Zusammenfassend kann festgehalten werden, daß die im Rahmen von § 7-Erhebungen gemachten Erfahrungen hinsichtlich der Schnelligkeit der Umsetzbarkeit gewünschter statistischer Prozeduren in SPSS insgesamt gesehen sehr positiv ausgefallen sind. Dies umso mehr, wenn man die Beurteilung der Laufzeiteffizienz nicht nur auf die reinen PC-Laufzeiten beschränkt, sondern auch den **Zeitaufwand für die konzeptionelle und programmiertechnische Umsetzung** gewünschter statistischer Prozeduren in die Bewertung mit einbezieht. Hier hat SPSS den Vorzug, daß viele statistische Prozeduren in Sekundenschnelle menügestützt über einige wenige Mausklicke abgerufen werden können, alternativ jedoch auch der Weg über die Befehlssyntax gewählt werden kann, der mehr Flexibilität bietet und auch bei Stapelverarbeitungen und Auswertungsroutinen Vorteile zeitigt.

Größere (hardwarebedingte) Probleme traten bei sehr speicherintensiven Auswertungsprozeduren und Daten(satz)modifikationen der Zeitbudgeterhebung auf, und dort v. a. verursacht durch die ungewöhnlich breite Datenmatrix (ca. 4 200 Variablen!). Hier zeigte

sich, daß die meisten Prozeduren extrem beschleunigt werden konnten, wenn der entsprechende Grundfile auf die analysespezifisch erforderlichen Variablen reduziert wurde, was aufgrund des sehr effizienten Daten(satz-)Managements in SPSS gewöhnlich sehr schnell bewerkstelligt werden konnte.

Darüber hinaus muß auf das **enorme Zeiteinsparungspotential** hingewiesen werden, daß mit **Innovationen im Bereich der PC-Hardwareausstattung** einschließlich einer möglichst optimal an den anwenderspezifischen Bedarfen ausgerichteten Programmumgebung einhergeht: Die Bildung der neuen personenbezogenen Variablen "Zeitverwendung für Studium/Schule" in Kombination mit dem Merkmal (1) "mit wem ausgeübt" bzw. (2) "wo ausgeübt", bei der für jede Person des Haushalts 288 Zeittakte in Verbindung mit sechs "mit-wem"- bzw. zwei "wo"-Kategorien durchgecheckt werden mußten, beanspruchte mit dem bisher für Auswertungszwecke der Zeitbudgeterhebung verfügbaren PC (133 Mhz Prozessor, 32 MB Arbeitsspeicher) im Netzbetrieb **1 1/4 Stunden** reine Laufzeit; dieselbe Prozedur konnte dagegen mit dem seit kurzem im Einsatz befindlichen deutlich leistungsstärkeren PC (233 Mhz Prozessor, 32 MB Arbeitsspeicher), der zudem als Stand-Alone-PC konfiguriert ist, in lediglich **7 Minuten** durchgeführt werden. Die o. g. Datentransformation aus der Datenaufbereitungs- und Auswertungspraxis der Zeitbudgeterhebung macht eindrucksvoll deutlich, welche enormen *kurzfristigen* Entwicklungsmöglichkeiten im Bereich der (PC-) Programmlaufzeiten aufgrund von Leistungssteigerungen im Hardwarebereich auch für die Zukunft zu erwarten sind.

3.2.1.3 Daten(satz)-Management

Wie in den bisherigen Ausführungen schon mehrfach angedeutet wurde, bietet SPSS sehr umfassende und effiziente Möglichkeiten des Datenmanagements. Der **Datenaustausch** von und zu den gängigen Tabellenkalkulationsprogrammen und Datenbanken verläuft aufgrund vorhandener Schnittstellen meist problemlos. Beim Öffnen von Excel-Tabellen (.xls-Format) zum Beispiel wird nicht nur die eigentliche *Datenmatrix* in den Dateneditor von SPSS übernommen, es steht zusätzlich die Option der Übernahme vorhandener *Variablenamen* zur Verfügung. In SPSS erstellte *Tabellen* und *Grafiken* können nicht nur in andere Windowsanwendungen (z. B. Winword) eingefügt werden, sondern auch über **OLE** (Object Linking and Embedding) und **DDE** (Dynamic Data Exchange) verknüpft werden. In der praktischen Arbeit hat sich jedoch gezeigt, daß OLE-Verbindungen

zwischen SPSS und Winword sehr speicherintensiv sind, wodurch sich der praktische Nutzen derartiger Verknüpfungen relativiert.

Was die **Datenformate** betrifft, so unterstützt SPSS ausgehend von Text-, numerischen, Dezimal- und Datumsformaten alle gängigen Formate; darüber hinaus können fehlenden Werten zwei Arten von Missing-Formaten – System-Missings und Benutzer-Missings zugewiesen werden, die in den jeweiligen statistischen Prozeduren entsprechend den Notwendigkeiten behandelt werden.

Gemäß Herstellerangaben gibt es hinsichtlich der **Größe des bearbeitbaren Datensatzes** weder bei den Fallzahlen noch bei der Anzahl der Variablen eine definierte Obergrenze. Im Zusammenhang mit der verarbeitbaren Variablenzahl muß angemerkt werden, daß das System vom Hersteller selbst mit bis zu **32 000 Variablen** in einer Datendatei getestet wurde.³⁶ Allerdings mußten hierbei die gewünschten Prozeduren über den Befehlseditor ausgeführt werden, da Windows die tatsächliche Anzahl an Variablen, die über Dialogboxen zugänglich sind, auf etwa **4 500 Variablen** beschränkt. In der praktischen Arbeit im Rahmen der § 7-Erhebungen wurde diese Grenze jedoch bisher noch nicht erreicht.

Für die Zwecke der **Datenmanipulation** bietet SPSS – alternativ menü- oder syntaxgestützt – ein umfassendes Instrumentarium zur Bearbeitung und **Transformation von Dateien** (Sortierung, Transponierung, Verschmelzung/Teilung, Datenaggregation, Datengewichtung, (bedingte) Auswahl von Fällen (auch Ziehung einer Zufallsstichprobe) usw.) und zur **Transformation von Daten** (z. B. neue Variablen berechnen, Variablen umcodieren, löschen von Daten, ändern von Datenformaten, Verfahren zur Ersetzung von Missing-Werten usw.). Sehr effizient sind in diesem Zusammenhang die Möglichkeiten zum **internen Verbinden von Dateien** (hinzufügen von Variablen bzw. Fällen), aber auch zur **Erzeugung einer Teilmenge von Variablen** (KEEP-and-DROP-Befehl). Die vielfältigsten, überwiegend positiven Erfahrungen mit der Anpassung der Datensätze an die analysebedingten Notwendigkeiten konnten diesbezüglich mit dem EU-Haushaltspanel (hier resultierten eine Vielzahl von Datensätzen allein schon aus den unterschiedlichen Befragungswellen) und der Zeitbudgeterhebung (hier erforderte v. a. der überpro-

³⁶ Vgl. Norusis, M.: SPSS für Windows Anwenderhandbuch, a. a. O., S. 499.

portional breite Datengrundfile die Bildung analysespezifischer Sub-Files) gemacht werden.

In SPSS ist es nicht möglich, mehrere Datendateien gleichzeitig zu öffnen und zu bearbeiten. Das heißt jedoch nicht, daß **Multi-Tasking** generell unmöglich ist. Werden Auswertungsprozeduren mit längeren Laufzeiten gestartet, besteht die Möglichkeit, im Rahmen der noch verfügbaren Systemressourcen (Arbeitsspeicher!) weitere Arbeiten sowohl in SPSS (z. B. Outputbearbeitungen, Entwurf einer Programmsyntax für eine zukünftige Auswertung) als auch in anderen Anwenderprogrammen (z. B. Textverarbeitung) durchzuführen.

Abschließend muß noch darauf hingewiesen werden, daß die im Rahmen der § 7-Erhebungen erworbenen **positiven Eindrücke mit den SPSS-Datenmanagementfunktionen** auch in der einschlägigen **Fachliteratur bestätigt werden**.³⁷

3.2.1.4 Benutzerfreundlichkeit

Die **Systembedienbarkeit**, sprich die Art der zur Verfügung stehenden Benutzeroberfläche, die Verständlichkeit/Übersichtlichkeit der Programmenüs, Dialogboxen usw., die Wahlmöglichkeiten zur Befehlseingabe aber auch die Qualität der Hilfefunktion im System einschließlich der erhältlichen Handbücher stellt ein Kriterium dar, dessen Bewertung sehr stark vom subjektiven Geschmack des jeweiligen Benutzers abhängt. Generell kann festgehalten werden, daß bezogen auf die § 7-Erhebungen, in denen SPSS angewandt wurde, auch Personen, die keinerlei Programmvorkenntnisse hatten, nach **Besuch eines Intensivkurses** (Dauer 3 bis 5 Tage) eine ausreichende Wissensbasis vermittelt bekamen, um anschließend die wichtigsten **Datenaufbereitungs- und Datenauswertungsarbeiten** im Rahmen der Statistikerstellung mit SPSS durchführen zu können.

Die Windows-Philosophie, sprich das Programm so einfach, übersichtlich und "selbsterklärend" wie möglich zu gestalten, wurde mit WinSPSS sehr überzeugend realisiert. Dies gilt für die **Benutzeroberfläche** und die Ausgestaltung der **Programmenüs und Dialogboxen**, aber auch generell für die in SPSS **realisierte Fenstertechnik und**

³⁷ Vgl. zum Beispiel die Ausführungen in Hilbert, A., Bankhofer, U., Elsholz, B., Niederöcker, U.: Statistiksoftware für den PC, a. a. O.; S. 33 und 53. oder in Sarnow, K.: Abstraktionsinstrumente. Statistiksysteme: von Shareware bis High-End, a. a. O.; S. 186.

Benutzerführung. Durch die Unterscheidung zwischen Editier-, Output- und Grafikfenster mit jeweils eigenen Menü- und Symbolleisten³⁸ wird ein müheloser Wechsel zwischen Input und Output und ein ebenso **intuitives wie einfaches Arbeiten** ermöglicht. Bei der Durchführung der meisten Prozeduren kann gewählt werden zwischen der Benutzung der **Windows-Bedienelemente und des Syntaxeditors**, was dem Nutzer – je nach Vorliebe, (Syntax-)Kenntnissen und/oder Art der Auswertung – ein adäquates Maß an **Freiheitsgraden** bietet.

Bei eventuell auftretenden Unklarheiten kann die ins System integrierte **Hilfefunktion** abgerufen oder aber die umfangreiche **Programmdokumentation** in Anspruch genommen werden. Das Hilfesystem ist **kontextbezogen und menügestützt** ausgebaut. Für jemanden, der über den Befehlseditor arbeitet, ist ferner die Möglichkeit zum Abrufen **kontextbezogener Syntaxinformationen** sehr nützlich. Darüber hinaus bietet SPSS umfangreiche und recht **übersichtliche Handbücher** zu den einzelnen Programm-Modulen (Base System, Advanced Statistics, Professional Statistics, Tables, LISREL/PRELIS, Trends, CHAID) und zur SPSS-Programmiersprache (SPSS Syntax Guide) an. Aufgrund der nahezu weltweiten Verbreitung von SPSS gibt es in der Zwischenzeit eine Vielzahl weiterer auf dem Markt erhältlicher SPSS-Anwenderhandbücher, die bei zusätzlichem Informationsbedarf ebenfalls zu Rate gezogen werden können.

Im Zusammenhang mit den in SPSS durchführbaren Statistikprozeduren muß angemerkt werden, daß SPSS i. d. R. weder die Angemessenheit eingegebener Prozeduren kommentiert noch bei der Anwendung statistischer Verfahren darüber informiert, ob bzw. gegebenenfalls welche dem jeweiligen Verfahren zugrundeliegenden Voraussetzungen erfüllt sein müssen und welche Implikationen sich hieraus ergeben. Insbesondere bei der Anwendung komplexer Analysen (Regressionsrechnungen, Clusteranalysen, usw.) ist daher eine umfassende fachliche Methodenkenntnis und das Wissen um die Voraussetzungen, Möglichkeiten und Grenzen der statistischen Verfahren Grundvoraussetzung.

Zusammenfassend kann festgehalten werden, daß SPSS was die Einfachheit und Übersichtlichkeit der Programmbedienbarkeit betrifft, nahezu keine Wünsche offenläßt. Die meisten der im Rahmen der Statistikerstellung anfallenden Prozeduren lassen sich in **SPSS in sehr kurzer Zeit erlernen und anwenden**. Die positive Bewertung gilt hierbei

³⁸ Vgl. auch die Ausführungen im Gliederungspunkt 2.2 zum SPSS-Dateiaufbau.

sowohl für die Durchführung menügestützter als auch – aufgrund der **relativ leicht erlernbaren Steuersprache** – für die syntaxgestützte Auswertung.³⁹

3.2.2 Spektrum der statistischen Prozeduren

Im 2. Gliederungspunkt wurde der modular aufgebaute Funktionsumfang von SPSS schon umfassend beschrieben, insofern muß auf das Spektrum der in SPSS zur Verfügung stehenden Auswertungsmodule nicht mehr im einzelnen eingegangen werden.

Bezogen auf die § 7-Erhebungen, in denen SPSS eingesetzt wurde, kann festgehalten werden, daß die in SPSS zur Verfügung stehenden Tools zum Datenmanagement, zur Datenmodifikation und -transformation wie auch das Spektrum der vorhandenen statistischen Prozeduren **für die benötigten Auswertungs- und Analysezwecke ein adäquates Instrumentarium** darstellten. Dies gilt nicht nur für die unter Auswertungsgesichtspunkten relativ einfach gestrickten Erhebungen vergleichbar der Investitionsschnellerhebung, sondern auch für das EU-Haushaltspanel, für die Erhebung zur beruflichen Weiterbildung und für die Erhebung über die Einkommensverhältnisse von Familien mit Kindern, die sowohl unter Aufbereitungs- als auch unter Auswertungsgesichtspunkten sehr hohe Ansprüche an o. g. statistisches Instrumentarium stell(t)en. Bezüglich der genannten Erhebungen muß hervorgehoben werden, daß sämtliche Datenaufbereitungs- und -auswertungsarbeiten im StBA unter Nutzung von SPSS durchgeführt wurden.

Hinsichtlich der **numerischen Zuverlässigkeit** der in SPSS abrufbaren statistischen Prozeduren liegen keine eigenen Erfahrungen vor. Um dies einschätzen zu können, wären Vergleichsrechnungen – zum Beispiel mit STATIS-BUND – notwendig, die jedoch im Rahmen der vorliegenden Arbeit aufgrund des damit verbundenen Zeitaufwands nicht geleistet werden konnten. Zieht man jedoch die einschlägige Fachliteratur zu Rate, so kann auf die Studie von *Hilbert, Bankhofer et al.*⁴⁰ verwiesen werden, in der die numerische Zuverlässigkeit/Laufzeiteffizienz von neun (PC-) Statistiksoftware-Programmen für

³⁹ Die Universität Manchester kommt bei ihrem Test verschiedener Statistikprogramme für SPSS zum gleichen Ergebnis.

Vgl. MIDAS, Manchester Computing University of Manchester (Campbell, M.): *Comparative Analysis of Statistical Packages and Data Bases for Statistics*, a. a. O., S. 45.

⁴⁰ Vgl. Hilbert, A., Bankhofer, U., Elsholz, B., Niederöcker, U.: *Statistiksoftware für den PC*, a. a. O.; S. 18 ff und Anlage B.1 "Aggregierte Datenmatrix mit 67 Merkmalen" auf den Seiten 52 und 53.

MS Windows⁴¹ getestet und an SPSS **acht von zehn möglichen Punkten** vergeben wurde. Was die **numerische Genauigkeit** einzelner Verfahren betrifft, so erreichte SPSS bei den *deskriptiven Statistiken* 9,7 (Mittelwert über alle getesteten Programme hinweg: 8,6), bei den *Korrelationsrechnungen* 9,9 (8,8), bei den *Regressionsrechnungen* 8,5 (8,9), beim *Datensatz von Longley* 9,9 (9,6), bei den *Einstichprobentests* 8,6 (7,1), bei den *Tests bei zwei unabhängigen Stichproben* 10 (7,6), bei der Berechnung *mathematischer Terme* 8,1 (6,3), beim *Statistics Quiz* 8,9 (7,8) und bei der *Varianzanalyse* 9,7 (8,5) der 10 maximal erreichbaren Punkte. Die größten Probleme bezüglich der numerischen Genauigkeit zeigte SPSS – wie viele andere Softwarepakete auch – bei *zwei verbundenen* (4,8 Punkte, Mittelwert: 6,0) und *mehreren Stichproben* (5,1 Punkte, Mittelwert: 6,0). Neben Tests auf numerische Genauigkeit wurden auch noch methodisch-orientierte **Experimente** hinsichtlich der statistischen Adäquatheit einzelner Prozeduren durchgeführt. SPSS erhielt hier bei der Beschreibung bi- und multivariater Verteilungen, bei den Regressions- und Korrelationsrechnungen, bei parametrischen Tests, beim Test zweier bzw. mehrerer unabhängiger Stichproben, beim Test zweier oder mehrerer verbundener Stichproben, bei der Diskriminanzanalyse und bei der Zeitreihenanalyse die Maximalpunktzahl 10. Zusammenfassend kommen die Autoren bezüglich ihrer Tests zur numerischen Zuverlässigkeit und Laufzeiteffizienz zu dem Ergebnis, "daß die **Genauigkeit der Ergebnisse** (der getesteten Programme; der Verf.) **im Durchschnitt sehr gut ist**". Bemängelt wurde die in fast allen getesteten Programmen fehlenden Problemhinweise und Lösungs- und Alternativvorschläge bei der Durchführung von statistischen Prozeduren.⁴²

3.2.3 Ergebnisdarstellung

In diesem Gliederungspunkt wird zunächst auf die sich in SPSS bietenden Möglichkeiten zur Darstellung von **Ergebnissen im Textmodus** eingegangen, die zum Beispiel in Form einer Tabelle, einer Auflistung von statistischen Kennziffern oder als textliche Wiedergabe komplexer statistischer Prozeduren (Regressionsrechnung, Faktorenanalyse etc.) im sog. **SPSS-Outputfenster** abgebildet werden. Daneben stellt SPSS mit seinem Grafik-Modul ein Instrument zur **grafischen Ergebnispräsentation** zur Verfügung.

⁴¹ In den Test einbezogen wurden die Programme Almo 32, Sigmastat 1.0/Sigmaplot 2.0, S-Plus 3.2, SPSS 6.1, Statgraphics Plus 1.1, Statistica 4.5, Systat 5.03, Unistat 3.0 und Winstat 3.0.

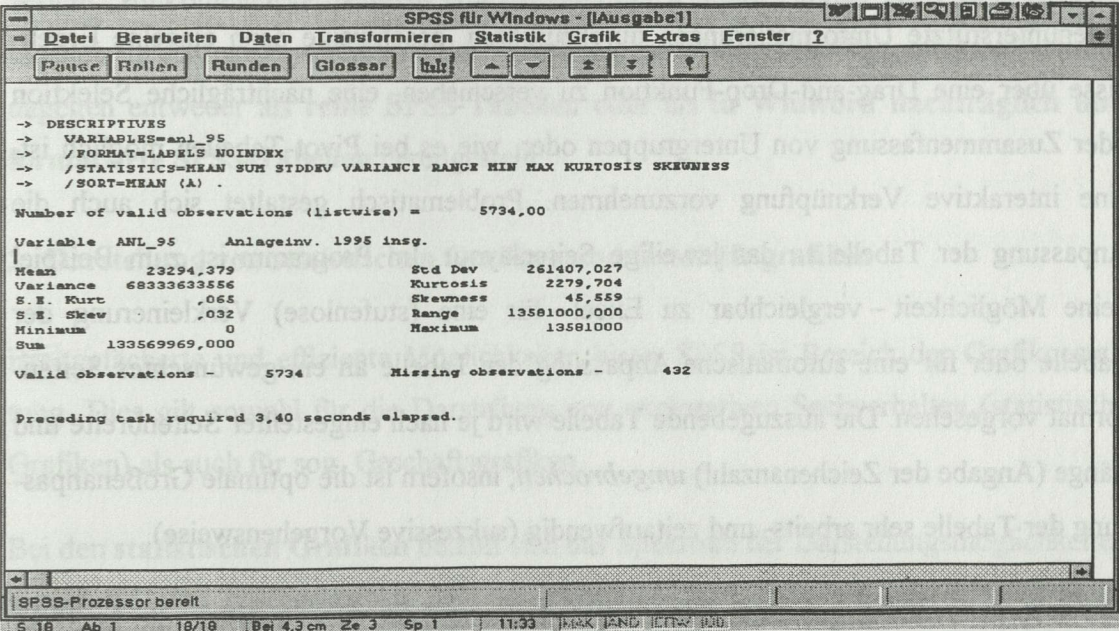
⁴² Vgl. Hilbert, A., Bankhofer, U., Elsholz, B., Niederöcker, U.: Statistiksoftware für den PC, a. a. O.; S. 30.

1. Ergebnisdarstellung im Textmodus

Wie oben schon angesprochen wurde, schreibt SPSS die Ergebnisse von statistischen Prozeduren **textlicher Art** in das sog. **Ausgabefenster**.⁴³ Neben den eigentlichen Ergebnissen der jeweiligen Statistikprozedur können darüber hinaus auch weitere Informationen, wie zum Beispiel die der jeweiligen Auswertung zugrundeliegende **Befehlssyntax**, **Fehler** und **Warnungen** und die zur Verfügung stehenden **Systemressourcen** ausgegeben werden. Der Aufbau der Darstellung einer SPSS-Ausgabe ist im Prinzip für alle Prozeduren gleich: SPSS gibt zunächst die Systeminformationen einschließlich der jeweiligen Befehlssyntax der Statistikprozedur (beides optional) im Ausgabefenster aus. Im Anschluß daran folgt die Darstellung der Statistikergebnisse.

Nachfolgend ist ein typisches SPSS-Outputfenster mit Ergebnissen deskriptiver univariater Statistiken zu den Anlageinvestitionen von Unternehmen aus der Investitionsschnellerhebung (im Beispiel bestehend aus arithmetischem Mittel, Varianz, Standardfehler des Exzesses und der Schiefe, Minimum, Summe, Standardabweichung, Exzeß, Schiefe, Spannweite und Maximum) abgebildet.

Abbildung 6: Darstellung univariater Statistiken im SPSS-Output-Fenster



⁴³ Einen optischen Eindruck von der SPSS-Ausgabe wurde schon im Rahmen der Beschreibung des SPSS-Dateiaufbaus mit der Darstellung einer einfachen univariaten Häufigkeitstabelle vermittelt (vgl. Abbildung 3 auf der Seite 20).

Darüber hinaus sind in den Anlagen 1 und 2 mit der Darstellung des statistischen Zusammenhangs zwischen Umsätzen und Anlageinvestitionen in Form des Pearson'schen Korrelationskoeffizienten und der Darstellung eines tabellarischen Reports zwei weitere Beispiele aufgeführt. In diesem Zusammenhang muß darauf hingewiesen werden, daß alle drei Beispiele nicht über den Befehlseditor sondern über das **Statistik-Menü** ausgeführt wurden, wozu lediglich einige wenige Mausklicke notwendig waren.

Der Output kann nachträglich editiert, ausgedruckt und gespeichert werden. Hierzu muß angemerkt werden, daß es zwar relativ einfach ist Textbeiträge nachträglich hinzuzufügen, problematisch wird es jedoch, wenn ein komplexerer Output, wie zum Beispiel der tabellarische Bericht in Anlage 2, nachträglich umformatiert werden soll. Für diesen Zweck sind die Möglichkeiten in SPSS eher **eingeschränkt und nicht sehr komfortabel**. Dies gilt nicht nur für oben aufgeführte Statistiken sondern insbesondere auch für die **Gestaltung von Tabellen**, wofür in SPSS ein separates **Tabellen-Modul** zur Verfügung steht. In Anlage 3 ist ein Beispiel für eine einfache, in SPSS erstellte Tabelle abgebildet. SPSS bietet zwar eine Vielzahl Formatierungsmöglichkeiten an, die menügestützt abgerufen werden können, im Vergleich zu einer Spezialsoftware (Excel, Winword etc.) sind die Möglichkeiten zur *Tabellennachbearbeitung* jedoch sehr begrenzt und das **Tabellenlayout eher bescheiden**. In SPSS ist es im Prinzip nicht möglich spalten- und zeilenunterstützte Umformatierungen durchzuführen, geschweige denn einzelne Ergebnisse über eine Drag-and-Drop-Funktion zu verschieben, eine nachträgliche Selektion oder Zusammenfassung von Untergruppen oder, wie es bei Pivot-Tabellen möglich ist, eine interaktive Verknüpfung vorzunehmen. Problematisch gestaltet sich auch die Anpassung der Tabelle an das jeweilige Seitenlayout. Im Programm ist zum Beispiel keine Möglichkeit – vergleichbar zu Excel – für eine (stufenlose) Verkleinerung der Tabelle oder für eine automatische Anpassung der Tabelle an ein gewünschtes Seitenformat vorgesehen. Die auszugebende Tabelle wird je nach eingestellter Seitenbreite und -länge (Angabe der Zeichenanzahl) *umgebrochen*; insofern ist die optimale Größenanpassung der Tabelle sehr arbeits- und zeitaufwendig (sukzessive Vorgehensweise).

Um anspruchsvolle Ergebnistabellen zu erstellen, empfiehlt es sich daher, die in SPSS erstellten Tabellen in eine **Spezialsoftware zu transferieren** und dort nachzubearbeiten. Hierzu muß die Tabelle markiert, in die Zwischenablage kopiert und zum Beispiel in

Excel oder Winword über das Menü "Bearbeiten ...einfügen" integriert werden.⁴⁴ Dies geht jedoch i. d. R. nicht ohne größere nachträgliche Umformatierungen zumindest des *Tabellenkopfes* und der *Tabellenvorspalte* einher.

Zusammenfassend kann festgehalten werden, daß das in SPSS zur Verfügung gestellte Tabellen-Modul zwar **vielfältige Möglichkeiten** zur menügestützten Tabellengestaltung bietet, problematisch wird es jedoch, wenn die Tabellen im SPSS-Ausgabefenster layoutmäßig nachbearbeitet werden müssen. Für die Ausgabe von Ergebnistabellen, bei denen dem Layout geringere Bedeutung zukommt, dürfte das SPSS-Tabellenmodul im allgemeinen zufriedenstellende Qualität und Komfort bieten. Sobald jedoch komplexe und/oder vom Layout her anspruchsvolle Tabellen erstellt werden müssen, ist der Einsatz einer entsprechenden Spezialsoftware zu empfehlen, da diese Programme mehr Flexibilität und auch vielfältigere (Formatierungs-) Möglichkeiten bieten und daher zwangsläufig für die Entwicklung von Präsentationstabellen besser geeignet sind.

Was die Praxis der Tabellenerstellung über SPSS im Rahmen von **§ 7-Erhebungen** betrifft, so wurden die o. g. Möglichkeiten ausgeschöpft: In den meisten Projektberichten zu den bisher durchgeführten § 7-Erhebungen wurden in den Text integrierte Ergebnistabellen überwiegend mit **Winword oder Excel** erstellt. Der *Tabellenanhang* im Projektbericht "Einkommensverhältnisse von Familien mit Kindern", ferner auch viele Arbeitstabellen für Eurostat im Rahmen der "Erhebung zur beruflichen Weiterbildung" wurden dagegen entweder als **reine SPSS-Tabellen** oder als **in Winword nachträglich umformatierte SPSS-Tabellen** bereitgestellt.

2. Darstellung von statistischen Grafiken und Geschäftsgrafiken

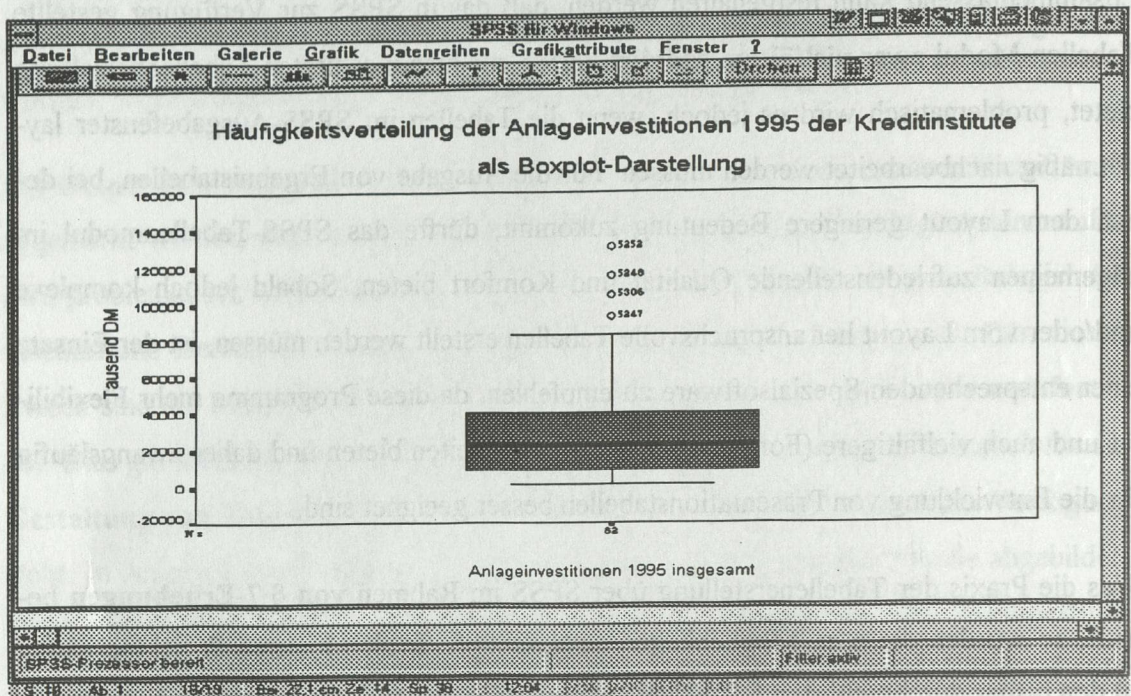
Breitgefächerte und effiziente Möglichkeiten bietet SPSS im Bereich der Grafikgestaltung. Dies gilt sowohl für die Darstellung von explorativen Sachverhalten (statistische Grafiken) als auch für sog. Geschäftsgrafiken.

Bei den **statistischen Grafiken** beläuft sich das Spektrum der Darstellungsmöglichkeiten ausgehend von Histogrammen, Normalverteilungs- und Boxplots über Fehlerbalken, Streudiagrammen bis zu Sequenz- und Zeitreihendiagrammen. Um einen Eindruck vom Grafikmenü vermitteln zu können, ist in Abbildung 7 ein **SPSS-Grafikfenster** aufge-

⁴⁴ In Winword muß im Anschluß daran noch der Befehl "Text in Tabelle" aus dem Menü "Tabelle" ausgeführt werden.

führt, wobei am Datensatz der Investitionsschnellerhebung die Häufigkeitsverteilung der Anlageinvestitionen 1995 der Kreditinstitute als **Boxplot** dargestellt ist.

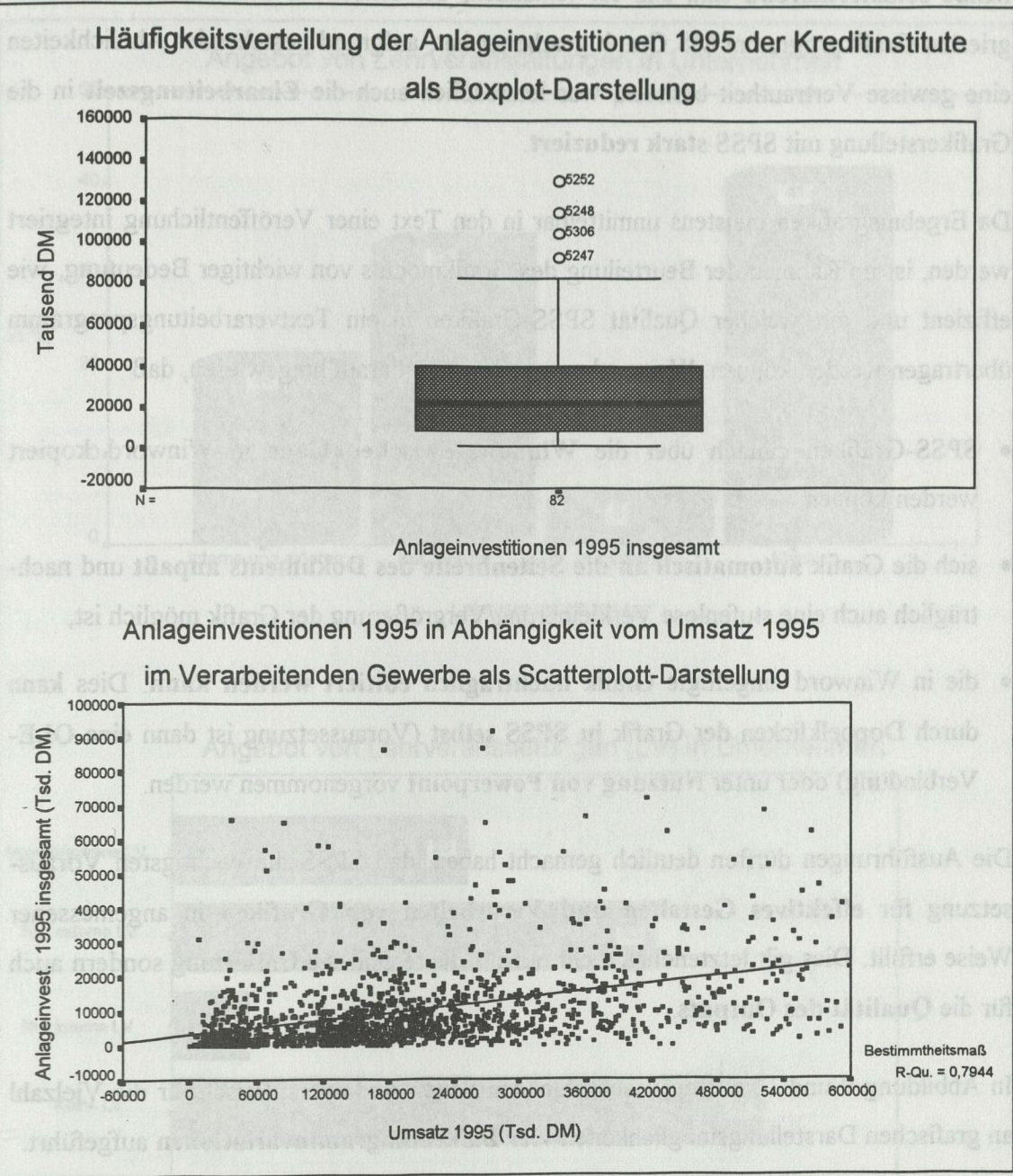
Abbildung 7: Beispiel für eine Boxplot-Darstellung im SPSS-Grafikfenster



Ein über das SPSS-Grafikmenü erstelltes Boxplot erscheint zunächst im **SPSS-Grafikkarussell**; anschließend kann der Inhalt des Grafikkarussells als eigenständige Grafik abgespeichert, editiert, gedruckt oder in eine externe Software übertragen werden. Die Übertragung nach Winword zum Beispiel kann problemlos über die **Zwischenablage** durchgeführt werden, wobei optional auch eine **OLE-Verbindung** hergestellt werden kann, d. h. in SPSS vorgenommene Grafikaktualisierungen werden automatisch auch im Winword-Dokument wirksam. Auf der nachfolgenden Seite ist o. g. Abbildung 7 und zusätzlich ein in SPSS erstelltes und mit wenigen Mausklicken über die Zwischenablage in das Winword-Dokument integriertes **Streudiagramm mit Regressionsgerade** abgebildet.

Eine über die Windows-Zwischenablage in Winword übertragene Grafik paßt sich automatisch an das nutzerspezifisch festgelegte Seitenlayout an, wodurch **arbeitsintensive Größenanpassungen der Grafik entfallen**.

Abbildung 8: In SPSS erstellte und in Winword eingebundene statistische Grafiken



Vergleichbar einfach und vielfältig sind auch die Möglichkeiten zur Erstellung von Geschäftsgrafiken, d. h. von für Veröffentlichungen vorgesehene **Ergebnisgrafiken**. Wie im Rahmen der Ausführungen zum SPSS-Leistungsspektrum schon dargestellt wurde, können unter Verwendung des Grafik-Moduls von Balkendiagrammen und Liniendiagrammen angefangen bis zu Flächendiagrammen, Kreisdiagrammen und Hoch-Tief-Diagrammen eine Vielzahl von Grafikarten menügeleitet erstellt werden. Was die Entwicklung einschließlich einer eventuell notwendigen Nachbearbeitung der Grafiken an-

belangt, kann festgehalten werden, daß die einzelnen Menüs und Dialogfelder in **hohem Maße selbsterklärend** sind und für jemanden, der schon mit dem in Winword integrierten Grafikprogramm MS Graph gearbeitet hat, aufgrund bestehender Ähnlichkeiten eine gewisse Vertrautheit besitzen, was letztendlich auch die **Einarbeitungszeit** in die Grafikerstellung mit SPSS **stark reduziert**.

Da Ergebnisgrafiken meistens unmittelbar in den Text einer Veröffentlichung integriert werden, ist im Rahmen der Beurteilung des Grafikmoduls von wichtiger Bedeutung, wie effizient und mit welcher Qualität SPSS-Grafiken in ein Textverarbeitungsprogramm übertragen werden können. Weiter oben wurde schon darauf hingewiesen, daß

- SPSS-Grafiken einfach über die **Windows-Zwischenablage** in Winword kopiert werden können,
- sich die Grafik **automatisch an die Seitenbreite des Dokuments anpaßt** und nachträglich auch eine stufenlose Verkleinerung/Vergrößerung der Grafik möglich ist,
- die in Winword eingefügte Grafik **nachträglich editiert werden kann**. Dies kann durch Doppelklicken der Grafik in **SPSS selbst** (Voraussetzung ist dann eine OLE-Verbindung) oder unter **Nutzung von Powerpoint** vorgenommen werden.

Die Ausführungen dürften deutlich gemacht haben, daß SPSS die wichtigsten Voraussetzung für **effektives Gestalten und Verarbeiten von Grafiken** in angemessener Weise erfüllt. Dies gilt letztendlich nicht nur für die technische Umsetzung sondern auch für die **Qualität des Outputs**.

In Abbildung 9 und 10 auf den nachfolgenden Seiten sind exemplarisch für die Vielzahl an grafischen Darstellungsmöglichkeiten **vier Balkendiagrammvariationen** aufgeführt.

Abbildung 9: Zwei in SPSS erstellte einfache Balkendiagramme

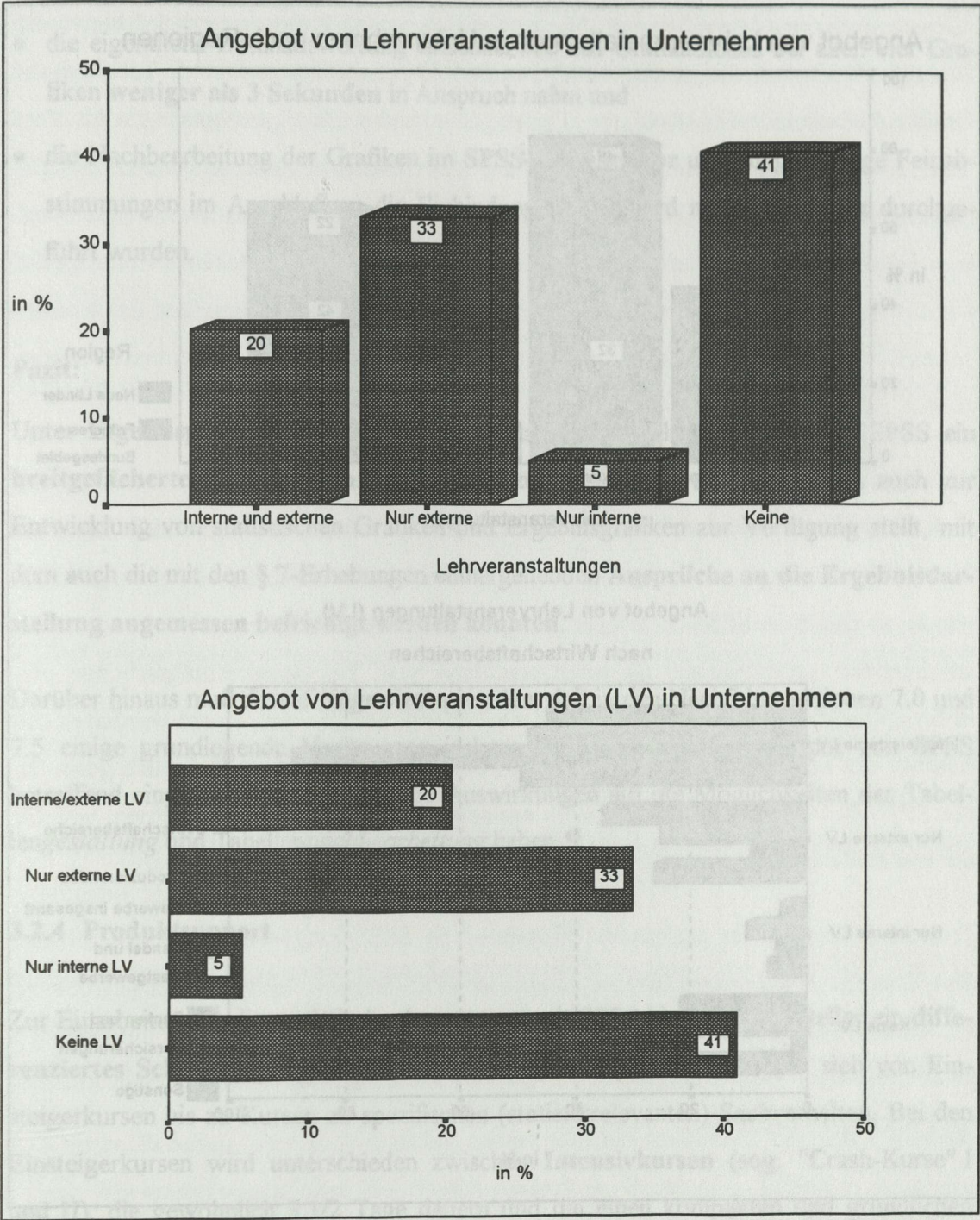
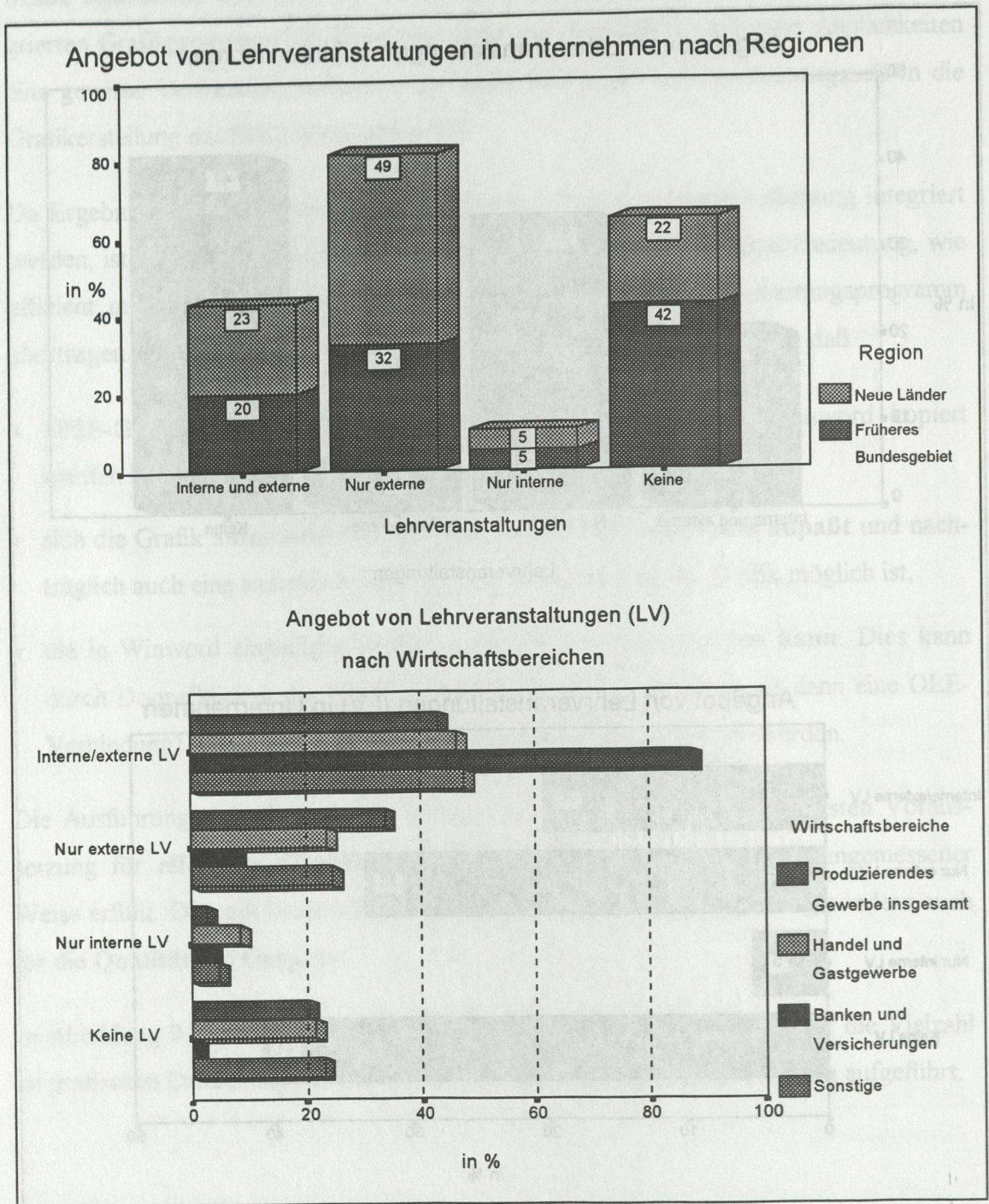


Abbildung 10: In SPSS erstelltes gestapeltes bzw. gruppiertes Balkendiagramm



Die Datengrundlage für alle vier Grafiken stellt der Datensatz der Vorerhebung zur beruflichen Weiterbildung bestehend aus 9 178 Fällen dar. Angemerkt werden muß dabei, daß

- die Vorarbeiten zur Grafikerstellung **menügestützt** mit wenigen Mausklicken durchgeführt wurden,
- die eigentliche Datenauswertung einschließlich des Grafikaufbaus bei allen vier Grafiken **weniger als 3 Sekunden** in Anspruch nahm und
- die Nachbearbeitung der Grafiken im **SPSS-Grafikeditor** und einige wenige Feinabstimmungen im Anschluß an die Einbindung in Winword mit **Powerpoint** durchgeführt wurden.

Fazit:

Unter Ergebnispräsentationsgesichtspunkten kann festgehalten werden, daß SPSS ein **breitgefächertes Instrumentarium** sowohl zur Erstellung von Tabellen als auch zur Entwicklung von statistischen Grafiken und Ergebnisgrafiken zur Verfügung stellt, mit dem auch die mit den § 7-Erhebungen einhergehenden **Ansprüche an die Ergebnisdarstellung angemessen befriedigt werden konnten**.

Darüber hinaus muß darauf hingewiesen werden, daß in den Nachfolgeversionen 7.0 und 7.5 einige grundlegende Neuerungen insbesondere den textlichen Output von SPSS betreffend eingeführt worden sind, die Auswirkungen auf die Möglichkeiten der **Tabelengestaltung** und **Tabellennachbearbeitung** haben.⁴⁵

3.2.4 Produktsupport

Zur Einarbeitung in die statistische Auswertung mit SPSS bietet der Hersteller ein **differenziertes Schulungsprogramm** an.⁴⁶ Das Angebotsspektrum erstreckt sich von Einsteigerkursen bis zu Kursen zu spezifischen (statistikrelevanten) Sachverhalten. Bei den Einsteigerkursen wird unterschieden zwischen **Intensivkursen** (sog. "Crash-Kurse" I und II), die gewöhnlich 3 1/2 Tage dauern und die einen kompakten und gründlichen Einstieg in die vielfältigen Datenmanagement- und Auswertungsmöglichkeiten von SPSS (Crash-Kurs I) bzw. weiterführende Kenntnisse für die Aufbereitung von Daten und deren Auswertung (Crash-Kurs II) vermitteln sollen, und sog. **Grundkursen**, die lediglich

⁴⁵ Vgl. die Ausführungen im Gliederungspunkt 3.3.

⁴⁶ Vgl. Anlage 4 mit Kursen zu unterschiedlichen Themengebieten der Monate September 1997 bis März 1998.

ein bis zwei Tage dauern und erstes Basiswissen vermitteln. Die Kurse zu spezifischen Sachverhalten umfassen ausgehend von der "Regressions- und Varianzanalyse", "Segmentieren und Klassifizieren", "Zeitreihenanalyse und Prognosemodelle" bis zu der "Datenanalyse mit neuronalen Netzen" und der "Ergebnispräsentation mit Tabellen und Grafiken" ein sehr breitgefächertes Themengebiet und dauern gewöhnlich ein bis zwei Tage. Die Schulungen finden in der Regel am Stammsitz der deutschen Niederlassung in München statt, werden auf Wunsch jedoch auch vor Ort durchgeführt, was die gleichzeitige Schulung einer größeren Zahl an Mitarbeiterinnen und Mitarbeitern ermöglicht. Darüber hinaus bietet der Hersteller auch auf die speziellen Anforderungen des Kunden zugeschnittene Kurse an, in denen mit Daten der jeweiligen Institution gearbeitet werden kann.

Die eigenen Erfahrungen mit dem Schulungsprogramm von SPSS konzentrieren sich v. a. auf die **Intensivkurse**. Hierzu kann festgehalten werden, daß in den 3 1/2 Tagen ein **breitgefächertes, praxisbezogenes und gut strukturiertes Programm** ausgerichtet wurde, so daß auch SPSS-"Anfänger" **ausreichende Kenntnisse** erwerben konnten, um anschließend grundlegende Datenaufbereitungs- und Datenauswertungsarbeiten mit SPSS im Rahmen der § 7-Erhebungen durchzuführen. Einige dieser Kurse wurden in den Räumen des StBA abgehalten. Hierbei wurde auch von dem Angebot Gebrauch gemacht, das SPSS-Instrumentarium an kleineren **amtseigenen Datensätzen** zu erlernen, was für die anschließende praktische Umsetzung erlernter Kenntnisse von Vorteil war.

Treten im Rahmen des praktischen Einsatzes von SPSS Fragen auf, bietet SPSS eine **Hot-Line für Problembewältigungen** an. Der SPSS-Support reicht von der gewöhnlich kostenlosen (telefonischen) Beantwortung kleinerer Anfragen bis zu umfassenden kostenpflichtigen Beratertätigkeiten vor Ort im Rahmen der Durchführung eines konkreten Projektes (Coaching).

3.3 Neuerungen in den Versionen SPSS 7.0/7.5

Im Vergleich zu der oben beschriebenen und bewerteten Version 6.0.x sind in den Nachfolgeversionen 7.0 und 7.5 einige **grundlegende Neuerungen** integriert worden, die positive Auswirkungen auf die Bewertung von SPSS haben. Da die beiden Versionen nur unter Windows'95 bzw. Windows NT 3.51 oder 4.0 lauffähig sind, im Amt gegenwärtig

jedoch noch Windows 3.1 benützt wird, konnte mit diesen neuen Versionen leider noch kein differenziertes praktisches Know-how erworben werden. Die nachfolgenden Ausführungen stützen sich daher neben den Erfahrungen, die aus dem Testen einer **Demo-Version** von SPSS 7.5 resultieren, lediglich auf **Informationsmaterial** des Herstellers. Im Zusammenhang mit der Demo-Version muß jedoch hervorgehoben werden, daß es sich hierbei nicht um eine Applikation zu reinen Demonstrationszwecken handelte, sondern um das für einen bestimmten Zeitraum freigeschaltete und somit flexibel einsetzbare "normale" SPSS-Programm mit sämtlichen Features. Insofern konnten zumindest einige Neuerungen im Programm ansatzweise getestet werden.

Einen Großteil der Neuerungen betrifft den **SPSS-Output**. Ausgabe wie Ausgabe-Fenster wurden grundsätzlich überarbeitet, was die Möglichkeiten der Gestaltung und die Qualität der Darstellung von **Ergebnissen im Textmodus** deutlich verbessert. Die o. g. negativen Gesichtspunkte insbesondere in Bezug auf die Nachbearbeitung von Tabellen wurden deutlich entschärft. In der Version 7.5

- sind **umfangreiche Formatierungsmöglichkeiten** zur *nachträglichen* Veränderung des Aussehens von Tabellen möglich. So können jetzt nachträglich angefangen von Textattributen (Schriftart, Farben, etc.), Tabellenattributen (Zahlenformate, Spaltenausrichtung, etc.), Hervorhebung wichtiger Ergebnisse bis zu Umgruppierungen von Tabellenkategorien eine Vielzahl an Formatierungen durchgeführt werden, die in der Version 6.0 und 6.1 noch nicht möglich waren.
- können **multidimensionale Pivot-Tabellen** erstellt werden. Dies vereinfacht die Tabellennachgestaltung, da durch die Unterstützung der Drag-and-Drop-Funktion Zeilen- und Spalteninformationen einfach durch Mausziehen vertauscht werden können.
- können **Tabellen**, aber auch jeglicher andere **Output textlicher Art, ohne Verlust von Formatierungen** über die Zwischenablage in Excel oder Winword integriert werden. Hierbei kann auch die Möglichkeit zur **Herstellung einer OLE-Verbindung** genutzt werden, was die Aktualisierung von Ergebnissen in Tabellen erleichtert.
- können **Tabellenfeatures** über eine neue Skriptsprache **standardisiert und automatisiert** werden, was die wiederkehrende Produktion und Veröffentlichung von Ergebnissen in Form von Tabellen(bänden) wesentlich erleichtert.

Die Einführung von **"Visual Basic"** als Skriptsprache neben der eigentlichen Befehlssprache soll mehr Flexibilität im Hinblick auf die Ausrichtung des Programms an den nutzerspezifischen Bedarfen ermöglichen. So hat der Nutzer angefangen von der **frei konfigurierbaren Menüleiste** bis zur **Erstellung eigener Skripte** vielfältige Möglichkeiten zur **individuellen Gestaltung seiner Arbeitsumgebung**. Darüber hinaus können mit der Skriptsprache in Zukunft eigenständige Applikationen zur **Automatisierung komplexer Arbeiten** und des **Ergebnis-Outputs** entworfen werden.

Auch wenn die meisten der genannten Features bisher noch nicht explizit auf Praxistauglichkeit getestet werden konnten, kann davon ausgegangen werden, daß durch die **Steigerung der Flexibilität und der Programmoffenheit** – gemäß Herstellerangaben läßt sich SPSS 7.5 problemlos in andere Programme und Anwendungen integrieren – die schon in der Version 6.0 und 6.1 vorhandenen effizienten Möglichkeiten im Bereich des Programm-Handlings und im Bereich der Arbeitsorganisation und -realisation einschließlich der Ergebnisdarstellung nochmals verbessert wurden, wobei die noch bestehenden Schwachpunkte v. a. im Bereich der Tabellengestaltung (auch der tabellarischen Reports und Kreuztabellen) verringert wurden.

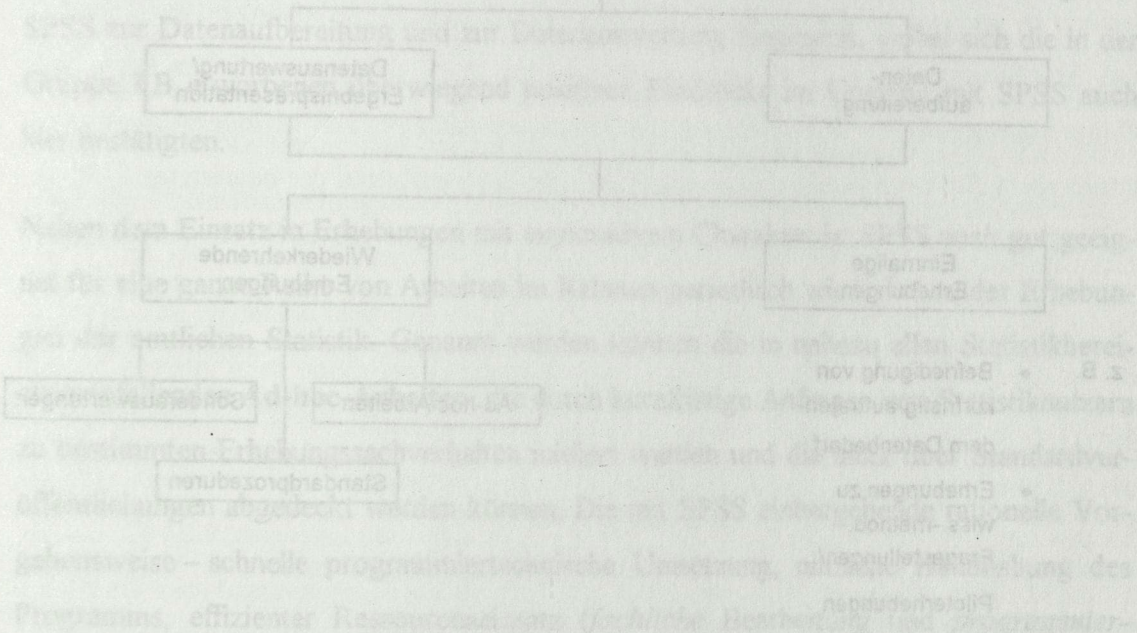
Durch die Unterstützung des HTML-Formats ist es gemäß Herstellerangaben auch möglich, **SPSS-Tabellen ins Internet oder Intranet zu stellen; dasselbe gilt für Grafiken**, sofern sie vorher im JPEG-Format abgespeichert werden. Darüber hinaus unterstützt SPSS im Rahmen des Exports von Grafiken *außerhalb* der Windows-Umgebung neuerdings neben dem genannten Exportformat JPEG auch die Formate WMF, CGM, TIFF, PICT, EPS und BMP, was unter dem Gesichtspunkt der Systemoffenheit ebenfalls positiv zu bewerten ist.

Auch im Bereich der **statistischen Prozeduren** gibt es eine Reihe von zusätzlichen Features. Genannt werden können u. a. im Bereich der Häufigkeitstabellen die Ausgabe des Mc Nemar's Tests, im Bereich der Faktorenanalyse die Rotationsmethoden (Varimax-, Equamax-, Quartimax-, Promax- und Oblimin-Rotation), im Bereich der Diskriminanzanalyse die Jackknife-Schätzung und im Bereich der Clusteranalyse die K-Clusterzentrenanalyse.

Was den strukturellen Aufbau von SPSS betrifft, so wurde seit der Version 7.0 eine ganze Reihe **neuer eigenständiger Module** entwickelt, die eine Vielzahl zusätzlicher

Analysemöglichkeiten bieten: U. a. **Exact Test** zur Analyse und Auswertung von Studien mit kleinen Fallzahlen; **AMOS** zur Modellierung von Strukturgleichungen und zur Durchführung konfirmatorischer Faktorenanalyse; **Diamond** als Ergänzung zum Tables-Modul zur Darstellung komplexer Informationen; **Neural Connection** (Neuronales Netzwerk) zur Vorhersage, Klassifikation, Zeitreihenanalyse und Datensegmentation und **Missing Values Analysis** zur Erkennung der Verteilung und Muster fehlender Werte und der Substitution fehlender Werte anhand von Schätzwerten aus Regressionsanalysen. Darüber hinaus wird für die Gestaltung von Eingabemasken seit kurzem auch ein neues **Data-Entry-Modul** zur Verfügung gestellt.⁴⁷ Getestet werden konnte keines der genannten Module, Aussagen über die Angemessenheit/den praktischen Nutzen der Prozeduren können daher nicht gemacht werden.

Einen abschließenden Überblick über die Features der neuen Version SPSS 7.5 gibt die Anlage 5.⁴⁸



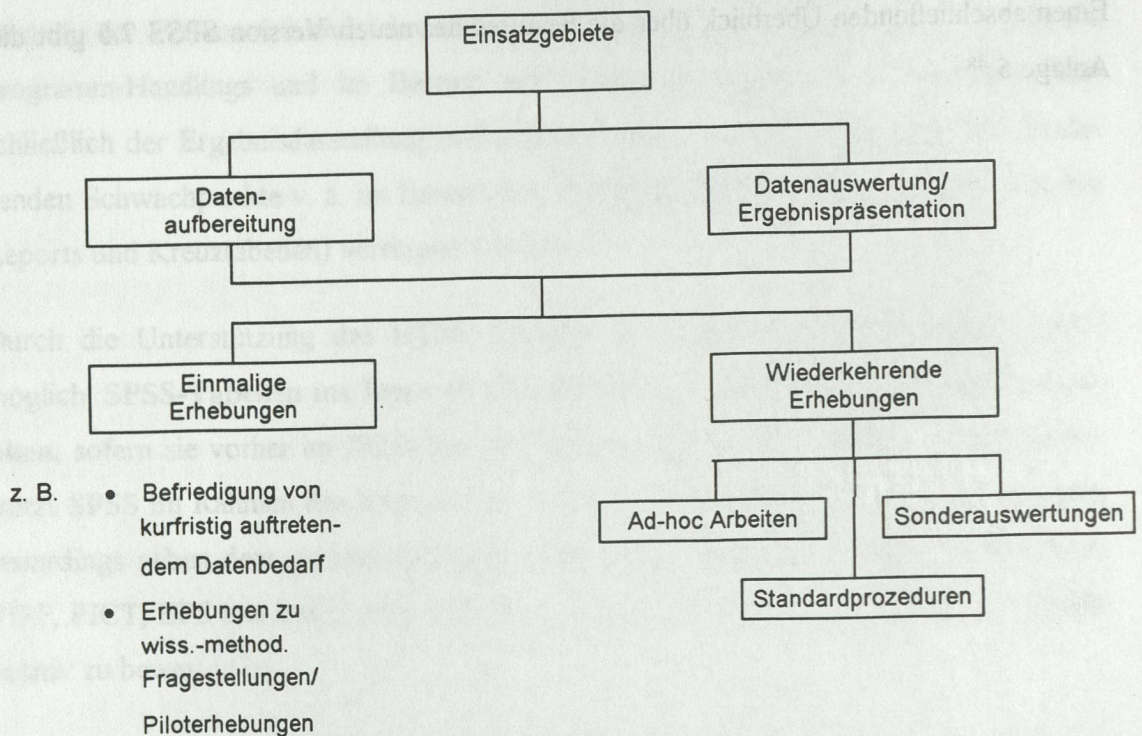
⁴⁷ Ein Data-Entry-Modul war bisher nur in die früheren SPSS-DOS-Versionen integrierbar.

⁴⁸ Angelehnt an die Checkliste in Janetzko, D.: Trendforscher. Universelle Statistikprogramme unter Windows; in: c't 1997, Heft 13, S. 228.

4 Empfehlungen für den zukünftigen Einsatz von SPSS im Statistischen Bundesamt

Die bisherigen Ausführungen dürften deutlich gemacht haben, daß SPSS gut geeignet ist für eine **breitgefächerte Anwendung** im Statistischen Bundesamt. Als Einsatzgebiete kommen nicht nur (kleine) Erhebungen mit **einmaligem Charakter** – vergleichbar den Erhebungen nach § 7 BStatG – in Betracht, sondern auch turnusmäßig **wiederkehrende Erhebungen** (Monatserhebungen, Jahresehebungen etc.) aus dem umfangreichen Erhebungsprogramm der amtlichen Statistik.

Abbildung 11: Einsatzgebiete von SPSS im Statistischen Bundesamt



Aufgrund der schnellen programmiertechnischen Umsetzung gewünschter (Statistik-) Prozeduren ist SPSS letztendlich prädestiniert für den Einsatz in Erhebungen mit **explorativem Charakter**, d. h. in Erhebungen, die neu initiiert oder inhaltlich umgestaltet werden und die daher hohe Ansprüche an die DV-Flexibilität generell und an die Effizienz der Aufbereitungs- und Auswertungstools der eingesetzten Statistiksoftware im spezifischen stellen. Neben dem Einsatz in Erhebungen nach § 7 Abs. 1 BStatG, die dem vorrangigen Ziel dienen, kurzfristig auftretenden Datenbedarf oberster Bundesbehörden

zu befriedigen, ist SPSS auch gut geeignet für die Klärung wissenschaftlich-methodischer Fragestellungen und für den Gebrauch in Test- bzw. Piloterhebungen zur Gewinnung und Verarbeitung von statistischen Informationen für den Zweck der Überprüfung und Evaluation von Erhebungsprozessen und -papieren im Vorfeld der Einführung einer neuen Statistik oder im Rahmen der Umgestaltung/Modifikation (bestimmter Sachverhalte) einer bereits existierenden Statistik. Ein gutes Beispiel für letztgenannten Erhebungszweck stellt die Testerhebung zur "Prüfung eines alternativen Erhebungsbogens für die mündliche Befragung in Haushaltserhebungen" dar, die gegenwärtig von der für den Mikrozensus (MZ) zuständigen Gruppe durchgeführt wird und v. a. dazu dient, zu testen, inwieweit Modifikationen an den Erhebungspapieren des MZ (v. a. am Interviewerbogen) aus fragebogenmethodischer Sicht zu Effizienz-/Qualitätssteigerungen führen können (Fragenaufbau, -führung, Filterung, etc.). In diesem Zusammenhang wurde – vergleichbar zur Vorgehensweise in der von Gruppe I B durchgeführten § 7-Erhebungen – Blaise zur Datenerfassung und (teilweisen) Datenplausibilisierung und SPSS zur Datenaufbereitung und zur Datenauswertung eingesetzt, wobei sich die in der Gruppe I B erworbenen überwiegend positiven Eindrücke im Umgang mit SPSS auch hier bestätigten.

Neben dem Einsatz in Erhebungen mit explorativem Charakter ist SPSS auch gut geeignet für eine ganze Reihe von Arbeiten im Rahmen periodisch wiederkehrender Erhebungen der amtlichen Statistik. Genannt werden können die in nahezu allen Statistikbereichen anfallenden **Ad-hoc-Arbeiten**, die durch kurzfristige Anfragen von Statistiknutzern zu bestimmten Erhebungssachverhalten initiiert werden und die nicht über Standardveröffentlichungen abgedeckt werden können. Die mit SPSS einhergehende rationelle Vorgehensweise – schnelle programmiertechnische Umsetzung, einfache Handhabung des Programms, effizienter Ressourceneinsatz (*fachliche* Bearbeitung und *programmiertechnische* Umsetzung der Anfrage kann in *Personalunion* erfolgen und macht *keine* kosten- und zeitintensive Arbeitsteilung notwendig) – ermöglicht eine flexible und schnelle dv-technische Abwicklung der häufig nicht sehr komplexen Anfragen. Sehr gute Erfahrungen diesbezüglich können/konnten v. a. im Rahmen der Erhebung zur beruflichen Weiterbildung in Unternehmen und der Zeitbudgeterhebung gemacht werden. Statistische Informationen zu wichtigen erhebungsspezifischen Kennziffern, die von einigen Nutzern zum Teil noch am selben Tag benötigt wurden (z. B. für wichtige Sitzungen), konnten mit SPSS in den meisten Fällen termingerecht zur Verfügung gestellt werden.

Äußerst fraglich ist es jedoch, ob für die Bearbeitung entsprechend kurzfristiger Anfragen, die Ad-hoc-Auswertungen und Tabellierungen erforderlich machen, die weniger flexiblen amtseigenen DV-Ressourcen (STATIS-BUND) gleichermaßen geeignet sind. Um hier kundenfreundlicher werden zu können, empfiehlt es sich, SPSS als **ergänzendes Instrumentarium** in die konventionelle DV-Umgebung (jetzt STATIS-BUND; später Genesis) zu integrieren. Im übrigen wäre diese Vorgehensweise für die Statistischen Ämter nicht neu: Das **Statistische Landesamt Rheinland-Pfalz** hat schon im Jahr 1987 eine **SPSS-Batch-Version** als **ergänzendes Instrumentarium** im Rahmen des **Landesinformationssystems (LIS)** für Aufbereitungs- und Auswertungszwecke eingesetzt, um den spezifischen, kurzfristigen Auswertungswünschen von LIS-Benutzern angemessen nachkommen zu können.⁴⁹

Darüber hinaus erschließt sich für SPSS auch ein Betätigungsfeld auf dem Gebiet der **Sonderauswertungen oder Ergänzungserhebungen**. Im Mikrozensus zum Beispiel werden in unterschiedlichen zeitlichen Intervallen Ergänzungserhebungen zu Urlaubs- und Erholungsreisen, zu Gesundheitssachverhalten, zur Altersvorsorge, zu Aus- und Weiterbildung und zur Wohnsituation der Haushalte mit zum Teil geringeren Auswahl-sätzen als in der Grunderhebung durchgeführt, die auch aufgrund der begrenzten Zahl an Variablen durchaus geeignet sein könnten für einen Einsatz von SPSS im Rahmen der Datenaufbereitung und -auswertung. Die schrittweise Einführung von SPSS für Auswertungszwecke im Rahmen des Mikrozensus wäre letztendlich die konsequente Fortführung des mit der Umstellung der Datenerfassung und der maschinellen Signier- und Plausibilitätskontrolle von der Großrechnerebene auf ein dialogorientiertes APC-gestütztes Verfahren (Blaise) eingeschlagenen Weges hin zu einem **integrierten, vollständig PC-gestützten Erhebungsmanagement** entsprechend dem Vorbild der neukonzeptierten laufenden Wirtschaftsrechnungen.⁵⁰

Potentielle Einsatzgebiete für SPSS bieten sich nicht nur auf dem Gebiet der Sonder-, Ergänzungs- und Ad-hoc-Arbeiten sondern auch im Rahmen des **"normalen" Erhebungsgeschäfts** des Statistischen Bundesamtes und zwar v. a. im Bereich der Stichpro-

⁴⁹ Vgl. Lehmann, R. E.: Das Landesinformationssystem. SPSS-Batch-Version als Ergänzung des Abrufinstrumentariums, in: Statistische Monatshefte Rheinland-Pfalz 3/87, S. 71 ff.

⁵⁰ Hinsichtlich detaillierter Ausführungen zum integrierten Erhebungsmanagement am Beispiel der laufenden Wirtschaftsrechnungen vergleiche Schwamb, H. J., Theis, T., Wein, E.: Integriertes Erhebungsmanagement, a. a. O.

benerhebungen der Wirtschaftsstatistiken, die sehr häufig einen nicht allzu großen Erhebungsumfang besitzen und auch von den Erhebungsinhalten her betrachtet im Vergleich zu Haushalts- und Personenbefragungen nicht allzu umfangreich sind. Genannt werden können z. B. die Kostenstrukturerhebungen, im Bereich des Handels die Großhandelsstatistiken (Befragung von ca. 13 000 Unternehmen), die Einzelhandelsstatistiken (ca. 33 000 Unternehmen) oder die Gastgewerbestatistiken (ca. 11 000 Unternehmen), darüber hinaus aber auch im Bereich des Produzierenden Gewerbes Erhebungen im Verarbeitenden Gewerbe (jährliche Investitionserhebung/Monatsbericht: (Betriebe von) max. 68 000 Unternehmen) und im Baugewerbe (Jahreserhebung einschließlich Investitionserhebung: max. 35 000 Unternehmen; Monatsbericht im Bauhaupt- bzw. Ausbaugewerbe: Betriebe von max. 20 000 Unternehmen (Bauhauptgewerbe) bzw. max. 9 000 ausbaugewerbliche Betriebe), die sich allesamt noch in einer fallmäßigen Größenordnung bewegen, die für einen (breiten) Einsatz von SPSS durchaus geeignet sein dürften. Vor dem Hintergrund der ständig steigenden Leistungsfähigkeit der Hard- und Software sollte man darüber hinaus für die **Unternehmensstatistiken** untersuchen, ob nicht auch in diesen Bereichen ein vollständig PC-gestütztes integriertes Erhebungsmanagement, bei dem die Datenerfassung und -aufbereitung (v. a. die Plausibilitätsprüfungen) über Blaise und die Datenauswertung und Präsentation über eine PC-Auswertungssoftware (z. B. SPSS) abgewickelt wird, zu **Effizienzgewinnen** (Zeit- und Kosteneinsparungen) führen könnte. Die Annahme, daß eine derartige Vorgehensweise mit Effizienzgewinnen einhergehen könnte, wird durch die Erfahrungen des Landesamtes für Datenverarbeitung und Statistik Nordrhein-Westfalen gestützt, das allein durch den Umstieg auf eine PC-gestützte Aufbereitung der Gehalts- und Lohnstrukturerhebung 1995 unter Nutzung von Blaise III 50 Personenmonate einsparen konnte.⁵¹ Es kann davon ausgegangen werden, daß über eine konsequente Dezentralisierung *auch* der Auswertungsprozesse dieses Ergebnis noch verbesserungsfähig ist.

Die zunehmende PC-Orientierung im Rahmen des Erhebungsmanagements hätte natürlich auch Konsequenzen für den **Prozeß der Leistungserstellung**: Die verstärkte Nutzung von PC-Aufbereitungs- (Blaise) und -auswertungssoftware (SPSS) hätte eine Verlagerung von (DV-) Aufgaben auf die fachliche Organisationseinheit zur Folge. Die

⁵¹ Pricking, T., et. al.: New approaches to data processing integration in North Rhine Westphalian Statistics based on Blaise III, in: Actes de la 4^e Conférence Internationale des Utilisateurs de Blaise, Institut National De La Statistique Et Des Etudes Economiques (INSEE), Paris 1997, S. 202.

zunehmenden Anforderungen aus dv-technischer Sicht in den Fachbereichen kann letztendlich nur über eine Bündelung des einschlägigen Fachwissens unmittelbar in den für die Durchführung der Statistik verantwortlichen Facheinheiten erreicht werden; d. h. die gegenwärtig überwiegend zentral koordinierten DV-Ressourcen müßten enger an die Fachabteilungen gekoppelt werden. Einhergehen dürfte dies mit Effizienzgewinnen auf

- **inhaltlichem Gebiet:** Die stärkere Anbindung des DV-Bereichs an die Fachabteilungen verbessert dessen Einblick in und das Verständnis für fachlich-inhaltliche Belange. Bestmögliche Kenntnis der Erhebungszusammenhänge ist letztendlich Grundvoraussetzung für eine effiziente und effektive - an den fachlichen Erhebungsnotwendigkeiten - ausgerichtete DV-technische Erhebungsrealisation.
- **zeitlichem Gebiet:** Die engere Verzahnung zwischen fach- und dv-technischen Belangen führt allein schon aufgrund der geringeren Koordinations- und Abstimmungserfordernissen und der kürzeren Informations-, Entscheidungs- und Handlungswege zu Zeiteinsparungen im Erhebungszyklus.
- **finanziellem Gebiet:** Die mit einer PC-gestützten Datenaufbereitung und Datenauswertung einhergehenden Zeiteinsparungen haben unmittelbaren Einfluß auf den mit der Erhebungsrealisation einhergehenden Kostenaufwand.

Zusammenfassend kann daher festgehalten werden, daß die verstärkte PC-Orientierung sowohl unter **Effizienz-** (rationeller Mitteleinsatz: frühzeitigere Datenbereitstellung; Kosteneinsparungen) als auch unter **Effektivitätsgesichtspunkten** (bessere Zielerreichung: Qualitätssteigerungen aufgrund geringerer Reibungsverluste zwischen fachlicher Erhebungskonzeption und DV-technischer Umsetzung) befürwortet werden sollte.

Literaturverzeichnis

- Bamberg, G., Baur, F.: Statistik, 8. Auflage, München 1993
- Bechtold, S., Bihler, W., Deininger, D.: Einmalige Leistungen der Hilfe zum Lebensunterhalt nach § 21 BSHG im Jahr 1991, in: *Wirtschaft und Statistik* 2/1993, S. 113 ff.
- Bechtold, S., Meyer, K.: Das Europäische Haushaltspanel. Darstellung des Erhebungskonzeptes und erster Ergebnisse, in: *Wirtschaft und Statistik* 5/1996, S. 296 ff.
- Bühl, A., Zöfel, P.: SPSS für Windows 6.1: Praxisorientierte Einführung in die moderne Datenanalyse, Bonn, Paris, Reading, Mass. u. a. 1995
- Deckl, S., Bechtold, S.: Wohngebäude in Plattenbauweise. Erste Ergebnisse einer Erhebung nach § 7 Abs.1 BStatG bei Eigentümern und Verwaltern, in: *Wirtschaft und Statistik* 2/1992, S. 88 ff.
- Gertkemper, F.: Ergebnisse der Testerhebung zum neuen Konzept der Laufenden Wirtschaftsrechnungen, in: *Wirtschaft und Statistik* 12/1997, S. 872 ff.
- Haag, U., Kieser, M.: Statistische Auswertungssysteme: eine Einführung in ihre Anwendung, Konstruktion und Bewertung, Stuttgart, New York, Jena 1992
- Hilbert, A., Bankhofer, U., Elsholz, B., Niederöcker, U.: Statistiksoftware unter Windows – Eine Marktanalyse, Arbeitspapiere zur Mathematischen Wirtschaftsforschung, Heft 124, Universität Augsburg 1995
- Janetzko, D.: Trendforscher. Universelle Statistikprogramme unter Windows, in: *c't* 1997, Heft 13, S. 224 ff.
- Lehmann, R. E.: Das Landesinformationssystem. SPSS-Batch-Version als Ergänzung des Abrufinstrumentariums, in: *Statistische Monatshefte Rheinland-Pfalz* 3/87, S. 71 ff.

- Lüüs, H. P., Ehling, M., Eppmann, H., Götz, K.: Tourismusstichprobe 1992, Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1994
- MIDAS, Manchester Computing University of Manchester (Campbell, M.): Comparative Analysis of Statistical Packages and Data Bases for Statistics, Material für den TES-Course SUP-104-E, Course Slides, Day 2
- Norusis, M. J.: SPSS für Windows. Anwenderhandbuch für das Base System Version 6.0, 1994
- Pricking, T., et. al.: New approaches to data processing integration in North Rhine Westphalian Statistics based on Blaise III, in: Actes de la 4^e Conférence Internationale des Utilisateurs de Blaise, Institut National De La Statistique Et Des Etudes Economiques (INSEE), Paris 1997, S. 49 ff.
- Sarnow, K.: Abstraktionsinstrumente. Statistiksysteme: von Shareware bis High-End, in: c't 1994, Heft 3, S. 174 ff.
- Schmidt, B.: Berufliche Weiterbildung in Unternehmen, Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1996
- Schmidt, B.: Investitionsschnellerhebung, Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1997
- Schwamb, H. J., Theis, T., Wein, E.: Integriertes Erhebungsmanagement, in: Methoden, Verfahren, Entwicklungen, Sonderausgabe 1/98
- Statistisches Bundesamt (Hrsg.): Die Zeitverwendung der Bevölkerung. Methode und erste Ergebnisse der Zeitbudgeterhebung 1991/1992, Wiesbaden 1995
- Stirner, H.: Soziale Indikatoren im Europäischen Haushaltspanel am Beispiel der Fragen zur Gesundheit, in: Wirtschaft und Statistik 9/1997, S. 627 ff.
- Wein, E.: Einkommensverhältnisse 1992 von Familien mit Kindern in Deutschland, Projektbericht zur Erhebung nach § 7 BStatG, Wiesbaden 1996

Anlage 1: Berechnung des Zusammenhangs zwischen Anlageinvestitionen, Umsätzen und Mitarbeitern 1995
(Pearson'scher Korrelationskoeffizient) aus den Angaben der Investitionsschnellerhebung

File l:\msoffice\spss\ise\auswert\daten5b.sav
Created: 17 Oct 97 11:45:21 - 79 variables and 6.166 cases

```
-> CORRELATIONS
-> /VARIABLES=anl_95 ums_95 ma_95
-> /PRINT=TWOTAIL SIG
-> /STATISTICS DESCRIPTIVES
-> /MISSING=PAIRWISE .
```

PEARSON CORR problem requires 192 bytes of workspace.

Variable	Cases	Mean	Std Dev
ANL_95	5734	23294,3790	261407,0266
UMS_95	6006	339545,9328	1724986,2604
MA_95	6152	976,6938	7005,2425

- - Correlation Coefficients - -

	ANL_95	UMS_95	MA_95	
ANL_95	1,0000 (5734) P= ,	,6383 (5585) P= ,000	,7961 (5725) P= ,000	= Korrelationskoeffizient nach Pearson = Fallzahl (n) = Irrtumswahrscheinlichkeit (p)
UMS_95	,6383 (5585) P= ,000	1,0000 (6006) P= ,	,7473 (5996) P= ,000	
MA_95	,7961 (5725) P= ,000	,7473 (5996) P= ,000	1,0000 (6152) P= ,	

(Coefficient / (Cases) / 2-tailed Significance)

" . " is printed if a coefficient cannot be computed

Preceding task required 9,78 seconds elapsed.

Anlage 2: Tabellarischer Report zu den Anlageinvestitionen und Umsätzen 1995

```

-> Report
-> /FORMAT= CHWRAP(ON) PREVIEW(OFF) CHALIGN(BOTTOM) UNDERSCORE(ON)
-> ONEBREAKCOLUMN(OFF) CHDSPACE(1) SUMSPACE(0) AUTOMATIC NOLIST
-> BRKSPACE(0)
-> PAGE(1) MISSING'. ' LENGTH(1, 99999) ALIGN(LEFT) TSPACE(1) FTSPACE(1)
-> MARGINS(1,129)
-> /TITLE=
-> LEFT 'Anlageinvestitionen und Umsätze 1994 und 1995 in Deutschland'+
-> ' insgesamt, früheres Bundesgebiet und neue Länder'
-> RIGHT 'Seite )PAGE'
-> /VARIABLES
-> anl_95 (VALUES) (RIGHT) (OFFSET(0)) (19)
-> anl_95_o (VALUES) (RIGHT) (OFFSET(0)) (19)
-> anl_95_w (VALUES) (RIGHT) (OFFSET(0)) (19)
-> ums_95 (VALUES) (RIGHT) (OFFSET(0)) (17)
-> /MISSING=LIST
-> /BREAK (TOTAL) 'Gesamtwerte' (SKIP(1)) /SUMMARY
-> SUM( anl_95 ) SKIP(1) SUM( anl_95_o ) SUM( anl_95_w ) SUM( ums_95 ) 'Summe'
-> /SUMMARY MEAN( anl_95 ) MEAN( anl_95_o ) MEAN( anl_95_w ) MEAN( ums_95 )
-> 'Mittelwert'
-> /BREAK wib5 (LABELS) (LEFT) (OFFSET(0)) (47)
-> /SUMMARY SUM( anl_95 ) SKIP(0) SUM( anl_95_o ) SUM( anl_95_w ) SUM( ums_95 )
-> 'Summe'
-> /SUMMARY MEAN( anl_95 ) MEAN( anl_95_o ) MEAN( anl_95_w ) MEAN( ums_95 )
-> 'Mittelwert' .

```

REPORT problem requires 6060 bytes of memory to store specifications for this task.

Anlageinvestitionen 1995 in Deutschland, in den neuen Ländern und im früheren Bundesgebiet und Umsätze 1995 (Tsd. DM)

4 Wirtschaftsbereiche	Anlageinvestitionen 1995 (Deutschland insgesamt)	Anlageinvestitionen 1995 (neue Länder)	Anlageinvestitionen 1995 (früheres Bundesgebiet)	Umsatz 1995 (Deutschland insgesamt)
Produzierendes Gewerbe				
Summe	63337477	15681411	47656066	1286069489
Mittelwert	21063	5215	15848	427692
Handel, Gastgewerbe				
Summe	6737672	1781511	4956161	318911383
Mittelwert	5318	1406	3912	251706
Verkehr, Nachrichten- übermittlung				
Summe	34692898	15083118	19609780	164390219
Mittelwert	87608	38089	49520	415127
Sonstige Dienstleistungen (incl. Wohnungswesen)				
Summe	12914491	4352156	8562335	94128957
Mittelwert	18215	6138	12077	132763
Gesamtwerte				
Summe	117682538	36898196	80784342	1863500048
Mittelwert	21878	6860	15018	346440

REPORT problem required an additional 1400 bytes of memory.

Preceding task required 10,82 seconds elapsed.

Anlage 3: SPSS-Tabelle zu den Anlageinvestitionen, Ausrüstungsinvestitionen und Bauinvestitionen 1995 im früheren Bundesgebiet und in den neuen Ländern

```
-> * Basic Tables.
-> TEMPORARY.
-> VARIABLE LABELS fbnr_2 ' '.
-> TABLES
-> /BOXCHARS = SYSTEM
-> /FORMAT LIGHT VBOX FRAME SPACE BLANK MARGINS(1,150) LENGTH(1,9999)
-> CWIDTH(24,19,19) INDENT(2) MISSING(' ') WRAPCHARS(' /-') LLAYER
-> /OBSERVATION anl_95_o anl_95_w aus_95_o aus_95_w bau_95_o bau_95_w
-> /FTOTAL $t 'Zusammen'
-> /TABLES (fbnr_2 + $t )
-> BY (anl_95_o + anl_95_w + aus_95_o + aus_95_w + bau_95_o + bau_95_w)
-> /STATISTICS
-> sum( 'Summe')
-> /TTITLE=
-> CENTER 'Anlageinvestitionen, Ausrüstungsinvestitionen und Bauinvestitionen'+
-> ' 1995 ' .
```

Anlageinvestitionen, Ausrüstungsinvestitionen und Bauinvestitionen 1995 im früheren Bundesgebiet und in den neuen Ländern

Wirtschaftsbereiche	Anlageinvestitionen 1995 (neue Länder)	Anlageinvestitionen 1995 (früheres Bundesgebiet)	Ausrüstungs- investitionen (neue Länder)	Ausrüstungs- investitionen 1995 (früheres Bundesgebiet)	Bauinvestitionen 1995 neue Länder)	Bauinvestitionen 1995 (früheres Bundesgebiet)
	in Tsd. DM	in Tsd. DM	in Tsd. DM	in Tsd. DM	in Tsd. DM	in Tsd. DM
Bergbau, Gewinnung v. Steinen und Erden	935724	1812915	682379	1562774	253345	250141
Verarbeitendes Gewerbe	7794989	37023827	5330401	33800690	2534204	4981247
Energie- und Wasserversorgung	5851943	7905318	3506873	6242725	2345070	1954698
Baugerbe	1121197	1431466	718588	997057	439886	486344
Großhandel	746988	1856503	378143	1262324	383857	632696
Einzelhandel	982027	2008192	498931	1362219	507081	701179
Gastgewerbe	53626	1118275	20775	929918	34968	192229
Verkehr	8757690	10518057	2543268	7501137	6502488	5754388
Nachrichtenübermittlung	6325428	9092253	5434090	7581733	891338	1510520
Kreditgewerbe	1268952	2526101	364114	1786859	926338	739242
Versicherungsgewerbe	1047029	1968681	45477	596681	1024567	1459549
Sonst. Dienstleistungen						
insgesamt (ohne Gast- gewerbe, Wohnungswesen)	1653598	3943147	2012482	7055161	1004065	1243436
Wohnungswesen	2815062	4620188	58613	913136	2762474	3786549
Zusammen	39354253	85824923	21594134	71592414	19609681	23692218

Preceding task required 6,87 seconds elapsed.

Anlage 4: Schulungsprogramm von SPSS im Zeitraum 9'97 bis 3'98

SEPTEMBER	2.-3.	München	Einführung in die Auswertung von Befragungen
	9.-10.	München	Datenanalyse mit Neuronalen Netzen
	11.-12.	München	SPSS-Auswertungen in der Medizin
	16.-19.	München	SPSS für Windows - Crash-Kurs
OKTOBER	7.-10.	München	SPSS für Windows - Crash-Kurs 2
	14.-15.	München	Zeitreihenanalyse und Prognosemodelle
	21.-22.	München	Teleform Workshop
	28.-30.	München	Multivariate Auswertungsverfahren in der Marktforschung
	28.-31.	Düsseldorf	SPSS für Windows - Crash-Kurs
	15.-16.	Berlin	SPSS für Windows - Grundkurs
NOVEMBER	11.-12.	München	Einführung in die Auswertung von Befragungen
	13.-14.	München	Ergebnispräsentation mit Tabellen und Grafiken
	18.-21.	München	SPSS für Windows - Crash-Kurs
	25.	München	Die SPSS 7.5 Skript-Sprache
	26.-27.	Frankfurt	Einführung in die statistische Datenanalyse und SPSS
	27.-28.	München	Regressions- und Varianzanalyse
DEZEMBER	2.-3.	München	Database Marketing mit SPSS und CHAID
	9.-12.	München	SPSS für Windows - Crash-Kurs
	9.-12.	Düsseldorf	SPSS für Windows - Crash-Kurs 2
JANUAR	14.-15.	München	Segmentieren und Klassifizieren
	20.-23.	München	SPSS für Windows Crash-Kurs
	27.-28.	München	SPSS-Auswertungen in der Medizin
	28.-29.	Frankfurt	SPSS für Windows - Grundkurs
	28.-29.	Düsseldorf	Datenanalyse mit Neuronalen Netzen
FEBRUAR	3.-4.	München	Einführung in die statistische Datenanalyse und SPSS
	5.	München	Die SPSS 7.5 Skript-Sprache
	10.-13.	Düsseldorf	SPSS für Windows Crash-Kurs
	17.-20.	München	SPSS für Windows Crash-Kurs 2
MÄRZ	3.-5.	München	Multivariate Auswertungsverfahren in der Marktforschung
	10.-13.	München	SPSS für Windows Crash-Kurs
	18.-19.	München	Database Marketing mit SPSS und CHAID
	24.-25.	München	Einführung in die Auswertung von Befragungen

Kurse zu SPSS Syntax oder Neuerungen in SPSS 7.5 richten wir auf Anfrage gerne ein. Bitte rufen Sie uns an.

Anlage 5: Überblick über wichtige Features von SPSS 7.5*

Allgemeines	Befehlssprache Basic-Dialekt Datenblatt On-Line-Hilfe max. Anzahl Variablen: 32 000 max. Anzahl Fälle: keine Beschränkung
Windows-Integration	DDE-Funktion OLE-Funktion ODBC
Daten-Import-/Export	ASCII BMDP dBase II/III/IV Excel 4 Excel 5/95/97 HTML-Export Lotus 1-2-3 Ver. 4/5 SYLK
Grafik-Export	BMP CGM EPS JPEG PICT TIFF WMF Windows-Zwischenablage
Lage- und Streumaße	
Grafiken	Tortendiagramm Balkendiagramm Histogramm (2D, 3D) Streudiagramm (2D, 3D) Rotation multiple Streudiagramme Box & Whisker Plots Stem-and-Leaf-Plots Wahrscheinlichkeitsplots Normalverteilungsplots Brushing
Auswertungsplanung	Normalverteilungstests Varianzhomogenitätsplots Power-Analyse
Nicht-parametrische Tests	Binominal Fisher Exact Chi-Quadrat Mantel-Haenszel Wilcoxon

Anlage 5: Überblick über wichtige Features von SPSS 7.5*

	MC-Nemar Mann-Whitney/U Moses Kolmogorov-Smirnov Wald-Wolfowitz Kruskal-Wallis/H Cochrans Q Friedman
Korrelationsmaße	Phi Cramers V Wilks Lambda Spearmans Rang Kendalls Rangkorrelation Kendalls Tau Kendalls W/ concordance Pearson Korrelationskoeffizient Eta kanonische Korrelation (S)
Regression	einfache multiple Regression polynomiale Regression schrittweise Regression logistische Regression robuste Regression Probit-Analyse gewichtete Regression Durbin-Watson Test
t-Test und Varianzanalyse	t-Test (abhängige/unabhängige Stichproben) n-faktorielle Varianzanalyse Kovarianzanalyse multivariate Varianzanalyse multivariate Kovarianzanalyse Varianzkomponentenschätzung
Post-hoc-Mittelwertvergleiche	Tukey HSD-Test (Tukeys a) Tukeys b Bonferronis t-Test Dunnetts Test Scheffe-Test
Logit, Loglineare Modelle	
Clusteranalyse	hierarchische Verfahren K-Means-Verfahren Cluster Methoden within groups/between groups avarage linkage single complete linkage Zentroid Clustering Median Clustering Wards-Methode

Anlage 5: Überblick über wichtige Features von SPSS 7.5*

	Cluster Graph Dendogramm
Proximitätsmaße	City-Block-Metrik euklidische Distanz quadratische euklidische Distanz Tschebyscheff Power-Distanzen Minkowski Kosinus
Faktorenanalyse	ortogonale Rotation schiefwinklige Rotation Promax-Rotation
Diskriminanzanalyse	einstufige schrittweise
multidimensionale Skalierung	metrische ordinale
Conjoint-Analyse	
Reliabilitätsanalyse	
Survival-Analyse	Kaplan-Meier Sterbetafeln Regressionsmodelle

* Berücksichtigt sind die SPSS-Module Base System, Professional Statistics und Advanced Statistics.

Quelle: Janetzko, D.: Trendforscher. Universelle Statistikprogramme unter Windows, in: c't 1997, Heft 13, Seite 228.

