

## **Einführung in ein Statistikprogramm – Vergleich der SAS/Enterprise Guide 4.1 Software mit SPSS für Windows**

Christof Kögler  
Universitätsrechenzentrum Heidelberg  
Im Neuenheimer Feld 293  
69120 Heidelberg  
Christof.Koegler@googlemail.com

Carina Ortseifen  
Universitätsrechenzentrum Heidelberg  
Im Neuenheimer Feld 293  
69120 Heidelberg  
Carina.Ortseifen@urz.uni-heidelberg.de

### **Zusammenfassung**

Angesichts der Weiterentwicklungen der beiden großen Statistiksoftwarehersteller SAS und SPSS werden mit diesem Beitrag die jüngsten Produkte aus beiden Häusern – SAS/Enterprise Guide Software 4.1 und SPSS 15.0 für Windows - aus der Perspektive des Ausbilders verglichen. Im Vordergrund steht dabei die Frage nach der Anwenderfreundlichkeit, besonders für den Anfänger. Umfragen in den Kursen im Wintersemester 2006/2007 am Rechenzentrum der Universität Heidelberg (URZ) sowie eigene Erfahrungen zeigen, dass sich beide Programme in vielen Bereichen aneinander annähern und sich die Argumente für den Einsatz des einen oder des anderen Programms nicht mehr ausschließlich mit inhaltlichen und funktionalen Aspekten begründen lassen, sondern monetäre Aspekte hinzugezogen werden.

**Schlüsselwörter:** Vergleich, SPSS, SAS/Enterprise Guide, Benutzerfreundlichkeit, Ausbildung

## **1 Einleitung**

Am Rechenzentrum der Universität Heidelberg (URZ) werden seit mehr als 20 Jahren Kurse in SAS und SPSS für Studierende und Mitarbeiter durchgeführt. Galt bis vor wenigen Jahren noch die Empfehlung – sofern vom Betreuer der Arbeit, vom Studien- oder Fachbereich keine Vorgaben gemacht wurden –, SPSS für einzelne, weniger umfangreiche Auswertungen einzusetzen und SAS immer dann, wenn viele wiederkehrende Aktionen notwendig sind und Programmiererfahrung vorhanden ist, so muss diese Empfehlung in Hinblick auf die SAS/Enterprise Guide Software (SAS/EG) und die neuen Möglichkeiten der Programmintegration (Python, Visual Basic) bei SPSS neu überdacht werden.

Die Einführung in SPSS erfolgt am URZ seit vielen Jahren innerhalb eines zweitägigen Blockkurses. Für das Arbeiten mit der SAS Software wird seit ca. 3 Jahren mit einem ebenfalls zweitägigen Kurs zur SAS/EG begonnen und darauf aufbauend das Programmieren mit SAS vorgestellt.

Im aktuellen Wintersemester 2006/2007 wurde in den beiden Einführungskursen SPSS und SAS/EG eine Erhebung durchgeführt, um mehr über die Benutzerfreundlichkeit der Programme bei Anfängern zu erfahren. Ergänzend haben wir unsere eigenen Erfahrungen zusammengetragen, um vergleichend die Vor- und Nachteile der beiden Statistikprogramme herauszuarbeiten.

Im Folgenden wird zunächst der Aufbau der am URZ durchgeführten Einführungskurse zu Statistiksoftware beschrieben (Abschnitt 2), anschließend der Fragebogen und die Ergebnisse der Kursumfragen vorgestellt (Abschnitt 3), unsere eigenen Erfahrungen dargestellt (Abschnitt 4) sowie abschließend die gesamten Erkenntnisse zusammengefasst (Abschnitt 5).

## 2 Kursaufbau

Die beiden Einführungskurse in SAS/EG und SPSS sind vom Zeitumfang und Inhalt gleichartig aufgebaut: Der Unterricht im Schulungsraum (1-2 Teilnehmer pro PC, 16 Geräte) wird an zwei aufeinanderfolgenden Tagen von jeweils 9-16 Uhr durchgeführt und besteht aus Vorlesungs- und zugehörigen Übungsteilen [1].

Inhaltlich werden folgende Themenbereiche behandelt:

1. **Arbeitsumgebung** (Oberfläche), **Aufruf und Beendigung des Programms**, **Hilfe-System** Online-Hilfe, kontext-spezifische Hilfe, Literatur und Web-Links
2. **SAS- bzw. SPSS-Tabellen** neu anlegen und auf vorhandene zurückgreifen, aus Fremdformaten wie Excel importieren und auch exportieren, bearbeiten – neue Variablen berechnen, Datumsvariablen, Fälle/Variablen löschen, Filtern / Untergruppen bilden, Tabellen verknüpfen und zusammenfassen/aggregieren
3. **Analysen und Grafiken:** Am Beispiel der Kennwerte für metrische und nominale Merkmale werden die Dialogfenster vorgestellt; am Beispiel des t-Tests und des Chi-Quadrat-Tests wird die Darstellung von Testergebnissen gezeigt; Überblick über die implementierten statistischen Verfahren; am Beispiel eines Balkendiagramms wird der Aufbau einer Grafik erläutert; Export der Ergebnisse in Word.
4. **Weiterführende Themen:** Einstellung der Arbeitsumgebung, Optionen, Veränderung des Ergebnis-Layout, Arbeiten mit der Syntax

## 3 Umfrage und Ergebnisse

Im laufenden Wintersemester 2006/2007 wurden die Teilnehmer beider Einführungskurse anhand eines Fragebogens über ihre Eindrücke befragt. Der Schwerpunkt lag hierbei auf der intuitiven Bedienbarkeit der Programme für Einsteiger. Die Befragung fand direkt nach Abschluss des Kurses statt.

### 3.1 Der Fragebogen

Die Kursteilnehmer wurden zum einen zu ihrem Eindruck vom jeweiligen Programm befragt, zum anderen zu ihren Vorkenntnissen im Bereich Statistik und Statistiksoftware, sowie der Motivation zum Kursbesuch. Dabei lag der Fokus unseres Interesses auf dem ersten Teil, während der zweite lediglich der Kontrolle auf andere Einflussfaktoren diente.

Die Teilnehmer wurden in einer Freitextfrage nach ihren ersten Eindrücken vom jeweiligen Programm gefragt. Außerdem sollten sie angeben, ob sie das Programm weiterempfehlen würden und ob sie vorhaben, es selbst weiter zu verwenden. Schließlich konnten sie auf einer vierstufigen Skala die intuitive Bedienbarkeit der Software sowie ihr eigenes Vermögen, nach dem Kurs mit dem Programm alleine zurechtzukommen, bewerten.

Im „Kontrollteil“ wurde das Fachgebiet der Teilnehmer und ihre Motivation zum Kursbesuch erfragt. Hier interessierte uns, ob die Teilnehmer an einer konkreten Auswertung arbeiteten, allgemeines Interesse zeigten oder vorwiegend am Scheinerwerb interessiert waren. Schließlich wurden die Teilnehmer nach ihren Vorkenntnissen zu Statistik, Windowssoftware im Allgemeinen, Statistiksoftware, SAS und SPSS befragt.

### 3.2 Ergebnisse

Etwa zwei Drittel der Kursteilnehmer gaben den ausgefüllten Fragebogen zurück. Wegen der geringen Fallzahl ist eine statistische Auswertung nur bedingt sinnvoll. Bezüglich der Benutzerfreundlichkeit der Programme – dem Hauptfokus unseres Interesses – wurden beide Programme überwiegend positiv bewertet.

Von 16 Befragten zu SPSS gaben 15 an sie würden das Programm weiterempfehlen und weiterhin benutzen. Auf die Frage nach der intuitiven Bedienbarkeit des Programms antworteten 13 mit „eher ja“ und drei mit „eher nein“.

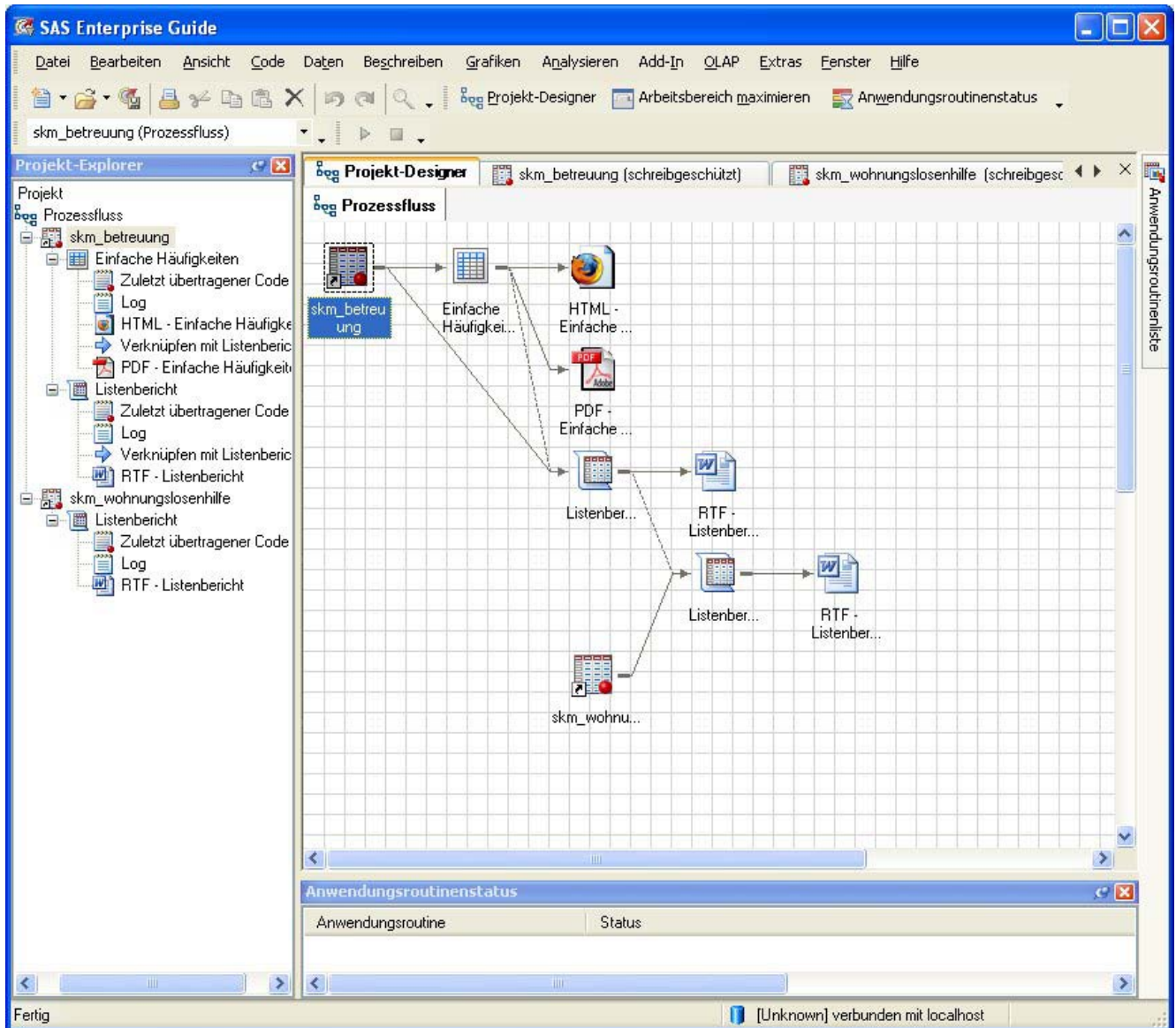
Ein ähnliches Bild ergibt sich für die SAS/Enterprise Guide Software. Von den zehn Teilnehmern des Einführungskurses, die an der Befragung teilnahmen, würden neun das Programm weiterempfehlen, acht haben vor es weiter zu verwenden. Zwei Teilnehmer halten SAS/EG für „absolut“ und fünf für „eher“ intuitiv bedienbar, während zwei dem Programm diese Eigenschaft „eher nicht“ und einer „gar nicht“ zuschreiben würden.

Die Einschätzungen der Eingängigkeit der SAS/EG sind also etwas weiter gestreut als die von SPSS, in beiden Fällen fällt die Bewertung jedoch überwiegend positiv aus. Die Reaktionen der Kursteilnehmer in den Freitextantworten spiegeln teilweise die Einschätzungen wider, die die Kursleiterin während des Unterrichts abgegeben hatte.

## 4 Beobachtungen und Erfahrungen

SPSS und dem SAS/Enterprise Guide liegen zwei unterschiedliche Philosophien zugrunde. SPSS orientiert sich stark an anderen Windowsanwendungen. Dateien werden geöffnet, angezeigt, bearbeitet und schließlich gespeichert oder die Änderungen verworfen. Seit der Version 14 können (endlich) auch mehrere Tabellen gleichzeitig geöffnet werden.

SAS funktioniert, bedingt durch die Logik des Systems, anders: Änderungen an einer Tabelle werden sofort festgeschrieben und können nicht rückgängig gemacht werden. Mit dem Enterprise Guide hat SAS ein weiteres Konzept eingeführt: das Projekt. Ein Projekt enthält Verweise zu Daten, Abfragen, Notizen, Syntax und Ergebnisse. Die Abbildung 1 zeigt dieses Konzept in der praktischen Anwendung.



**Abbildung 1:** Das Projekt in der Prozessflussansicht des SAS/EG 4.1

Allgemein sind die einzelnen Elemente der Softwarepakete im SAS/EG vergleichsweise stark integriert. Alle Elemente sind im Projekt zusammengefasst und können zueinander in Beziehung gesetzt werden. Bei SPSS ist der Integrationsgrad deutlich geringer – jede Tabelle, Syntax- oder Ergebnisdatei wird in einem eigenen Fenster geöffnet. SPSS kennt keine Beziehungen zwischen einzelnen Dateien.

SPSS macht es dem Einsteiger jedoch in anderen Bereichen leichter, da das Programm sich stärker Windows-typisch verhält. So wird der Einsteiger leicht vom SAS-typischen

Speicherverhalten überrumpelt, da SAS bekanntlich alle Änderungen an den Daten sofort festschreibt. Ein weiteres SAS-Konzept, das bei Einsteigern zu Verwirrung führt, ist die Bibliothek (Library). Der SAS/EG-Anwender ist an einigen Stellen, beispielsweise beim Anlegen neuer Tabellen oder bei der Definition von Formaten, ohne Vorwarnung mit dem Bibliothekensystem konfrontiert.

Die Menüorganisation ist beim SAS/EG derzeit besser gelöst als bei SPSS, die Analyse-Menüs *Beschreiben*, *Grafiken* und *Analysieren* sind klarer strukturiert als beim Konkurrenten. SPSS plant, in den kommenden Versionen im Grafikmenü aufzuräumen, sodass der Benutzer nicht mehr mit mehreren verschiedenen Möglichkeiten zur gleichen Funktionalität konfrontiert sein wird.

Bei der Etikettierung von Werten (Formate) bietet SPSS mehr Übersichtlichkeit. Die Wertelabels lassen sich, einmal definiert, nach Belieben ein- oder ausblenden. Die SAS-Formate sind umständlicher zu bedienen, haben jedoch den Vorteil, dass Tabellen-übergreifend mit ihnen gearbeitet werden kann. SPSS bietet zudem die Möglichkeit fehlende Werte benutzerdefiniert zu kodieren, während dies bei SAS nur über Umwege möglich ist.

Die SAS/EG integriert die erzeugte Syntax besser als der Konkurrent. Die Syntax wird direkt im „Log“ ausgegeben, was den Einstieg in die Programmierung erleichtert. Andererseits ist der erzeugte Code unübersichtlich, aufgrund der Verwendung von SQL, ODS und Makros. In SPSS muss die Syntax getrennt „eingefügt“ werden. Der Anfänger hat so keinen Kontakt zum Programmcode. Automatisierung ist in SPSS auch nur über die Syntax möglich, wohingegen der SAS/EG die Möglichkeit bietet, über den Prozessfluss und die Verwendung von Parametern Abläufe zu reproduzieren.

Beide Programme haben ihre Stärken und Schwächen. Vor der Einführung des SAS/EG galt die Faustregel „Für einmalige Benutzung nimm das intuitive SPSS, für ‚richtige Statistik‘ verwende SAS.“ Heute können die beiden Programme aufgrund ihrer Eigenschaften nicht mehr klar einem bestimmten Nutzungsprofil zugeordnet werden. Die konzeptuellen Stärken des SAS/EG liegen in der Visualisierung der Abläufe im Projekt, die von SPSS in der besseren Kompatibilität mit anderen Windowsanwendungen bezüglich Speicherverhalten und Dateimanagement.

## 5 Fazit

Mit dem Enterprise Guide hat SAS, was die Einsteigerfreundlichkeit betrifft, zum Konkurrenten aufgeschlossen. SPSS hat sich hingegen in anderen Bereichen verbessert, wie beispielsweise im Bereich Grafik – hier soll die Umstellung zur einheitlichen Grafikerstellung mit der Version 16 endgültig vollzogen werden. Auch die Teilnehmer der Einführungskurse am URZ Heidelberg haben beide Produkte als intuitiv bedienbar bewertet. Eine Empfehlung für eines der Programme auszusprechen gestaltet sich zunehmend schwierig. Oft ist der entscheidende Faktor die Vorliebe des Professors oder die Verfügbarkeit am Institut.

Die beiden Hersteller arbeiten also an den Defiziten und nähern sich gegenseitig an. In akademischen Kreisen wird angesichts dessen in Zukunft der Preis sicherlich eine im-

mer größere Rolle bei der Entscheidung für das eine oder das andere Programm spielen. Hinweise dafür finden sich in der Entwicklung der Lizenzzahlen an der Universität Heidelberg aus der jüngsten Vergangenheit.

## **Literatur**

- [1] Unterlagen zum SAS/EG- und SPSS-Kurs: <http://www.urz.uni-heidelberg.de> und dann die Stichworte *Statistik > Kurse > Unterlagen zu SAS-Kursen* bzw. *Unterlagen zu SPSS-Kursen* anklicken.