

SAS-Kurse am HRZ der Universität Dortmund - Ein Erfahrungsbericht -

Dr. Andreas Christmann

Universität Dortmund, Hochschulrechenzentrum

Abstract

Am Hochschulrechenzentrum der Universität Dortmund finden regelmäßig Kurse zu SAS statt. Es wird ein Überblick über die behandelten Themen gegeben. Es hat sich herausgestellt, daß Personen, die noch nicht mit SAS gearbeitet haben, besonders die leicht bedienbaren Module wie zum Beispiel SAS/INSIGHT erlernen möchten. Personen, die bereits Erfahrung mit SAS gesammelt haben, äußern häufig den Wunsch, mehr über die SAS-Syntax oder über SAS/IML zu erfahren.

Einleitung

Am Hochschulrechenzentrum der Universität Dortmund finden regelmäßig Kurse zu SAS statt. Anfängerkurse werden zu Beginn des Wintersemesters und Fortgeschrittenenkurse zu Beginn des Sommersemesters abgehalten. Die Kurse richten sich an Studierende und Mitarbeiter/innen aller Fachbereiche sowie in der Regel auch an Institutionen der Universität, wie z.B. Angestellte der Bibliotheken oder der Verwaltung. Die Kurse dauern 3 bis 5 Tage. Die Kursdauer beträgt 3 Stunden pro Tag. Die Teilnehmerzahl variiert stark, etwa zwischen 5 und 30 Personen. Die Vorkenntnisse zu SAS und zu Statistik variieren ebenfalls stark, da Personen aus verschiedenen Fachbereichen und aus verschiedenen Semestern an den Kursen teilnehmen. Somit kommen die Teilnehmer i.a. mit sehr verschiedenen Erwartungshaltungen.

Art der Kurse

Die Kurse sind stark anwendungsorientiert. Es hat sich als sinnvoll erwiesen, zwei bis drei kleinere Beispieldatensätze zu verwenden, um die Verwendung von SAS zu demonstrieren. Trotz der knappen Zeit wird auf Wunsch der Teilnehmer zum Teil versucht, die Überlegungen darzustellen, die den statistischen Verfahren zugrunde liegen. Manchmal ist es sinnvoll, allgemeine Hinweise zur statistischen Datenverarbeitung zu geben und auf „Fallen“ hinzuweisen. Meist ist der letzte Tag reserviert für spezielle Fragen der Kursteilnehmer.

Themen in Anfänger- und Fortgeschrittenen-Kursen

Zunächst werden einige allgemeine Dinge zu SAS erläutert, wie die Hochschullandeslizenz, Preise, Versionen, SAS-Module und allgemeine Informationsquellen zum SAS-System. Es wird in aller Kürze versucht darzulegen, wie die Programme SAS, SPSS und S-PLUS untereinander einzuordnen sind. Vor- und Nachteile der Verwendung von Syntax und Mouse werden erläutert.

Import- und Export-Möglichkeiten von Daten, Tabellen oder Plots werden gezeigt am Beispiel von Microsoft Word und Excel, LaTeX sowie SPSS. Es wird gezeigt, wie deutsche Umlaute und das „ß“ im SAS-System verwendet werden können.

Themen für Anfängerkurse

Da viele Kursteilnehmer, die noch nicht mit SAS gearbeitet haben, zunächst nur ungern die Syntax erlernen möchten, hat sich die Verwendung von SAS/INSIGHT als besonders geeignet erwiesen. An einem praktischen Beispiel wird demonstriert, wie von der eigentlichen Fragestellung ausgehend über die Erstellung eines Kodierplans unter Beachtung der

Skalenniveaus der Variablen die Daten in das SAS-System eingegeben werden können. Die numerischen und graphischen Möglichkeiten von SAS/INSIGHT werden demonstriert.

Von den vielfältigen Möglichkeiten des DATA-Steps werden nur die grundlegendsten Punkte wie temporäre vs. permanente Dateien, I/O-Funktionen, Möglichkeiten der Datenmodifikation und die Verwendung von Funktionen beispielhaft behandelt. Nur in der Praxis besonders oft verwendete Prozeduren werden in sehr einfachen Beispielen behandelt, wie zum PROC FORMAT, PRINT, CONTENTS, MEANS, CAPABILITY, FREQ, CORR, GPLOT und REG.

Gegen Ende des Kurses werden kurz weitere Module wie SAS/ASSIST oder SAS/LAB gezeigt.

Themen für Fortgeschrittenenkurse

Personen, die bereits Erfahrung mit SAS gesammelt haben, äußern häufig den Wunsch, mehr über die SAS-Syntax zu erfahren. Deshalb werden in den Fortgeschrittenenkursen komplizierte Arten des Datenimports, DO-Schleifen, ARRAYS, sowie Datums- und Zeit-Formate behandelt. Aus dem Modul SAS/STAT werden in der Regel die Prozeduren REG, GLM, NPAR1WAY und GPLOT behandelt. Zum Teil möchten Kursteilnehmer Grundzüge von SAS/IML oder die explorativen Methoden von SAS/INSIGHT kennenlernen, etwa zur Erkennung von Ausreißern oder Untergruppen. Bei Bedarf kann die Erstellung einer Eingabemaske mit PROC FSEDIT oder die Verwendung von einfachen SAS/Makros demonstriert werden.

Aus Zeitgründen können in der Regel leider nicht alle genannten Themen behandelt werden. Dies trifft auch auf andere SAS-Module zu, die im Rahmen der Hochschullandeslizenz verfügbar sind.

SAS-Kurse des Fachbereichs Statistik

Der Fachbereich Statistik führt für seine Studierenden eigene SAS-Kurse durch. Der Hörerkreis ist entsprechend homogener und verfügt über ein breiteres Statistik-Wissen als derjenige bei den Kursen des Hochschulrechenzentrums. Die Kursdauer beträgt in der Regel 1 bis 2 Wochen und behandelt auch SAS-Prozeduren für komplexere Fragestellungen.

Fazit für das HRZ-Dortmund

Anfänger bevorzugen oft die SAS-Module, die mit der Mouse bedient werden können. Insbesondere SAS/INSIGHT hat sich für den Einstieg in das SAS-System als geeignet erwiesen. Personen, die schon Erfahrungen mit SAS gesammelt haben, möchten oft ihre Kenntnisse über die SAS-Syntax vertiefen. Es sei betont, daß die Erfahrungen am HRZ der Universität Dortmund sich selbstverständlich von denen an anderen Universitäten unterscheiden können.

Literatur

Höschel, H.-P., Maucher, C., Schacht, S., Strüby, R. (1996). Datenanalyse mit dem SAS-System. SAS Institute Inc.

Dufner, J., Jensen, U., Schumacher, E. (1992). Statistik mit SAS. Teubner, Stuttgart.

Graf, A., Ortseifen, C. (1995). Statistische und grafische Datenanalyse mit SAS. Spektrum, Heidelberg.