Berichte aus der Statistik

Dietmar Beyer,
Carina Ortseifen (Hrsg.)

SAS® in Hochschule und Wirtschaft

Proceedings der 8. Konferenz der SAS®-Anwender
in Forschung und Entwicklung (KSFE)

KSFE
KONFERENZ DER SAS®-ANWENDER IN FORSCHUNG UND ENTWICKLUNG

Shaker Verlag
Aachen 2004

Bibliografische Information der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Copyright Shaker Verlag 2004
Alle Rechte, auch das des auszugsweisen Nachdruckes, der auszugsweisen oder vollständigen Wiedergabe, der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, vorbehalten.

Printed in Germany.

ISBN 3-8322-3280-X
ISSN 1619-0963

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen
Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9
Internet: www.shaker.de • eMail: info@shaker.de

Vorwort

Die 8. Konferenz der SAS Anwender in Forschung und Entwicklung (KSFE) fand am 26. und 27. Februar 2004 an der Fachhochschule Schmalkalden statt. Die Veranstaltung wurde gemeinsam vom Fachbereich Informatik der Fachhochschule Schmalkalden und dem Verein Kooperation der SAS-Anwender in Forschung und Entwicklung e.V. ausgerichtet.

Die Statistikausbildung und die Zusammenarbeit zwischen Hochschule und Wirtschaft hat am Standort Schmalkalden eine lange Tradition. Im Jahr 1902 als „Königliche Fachschule für Kleineisen- und Stahlwaren-Industrie“ gegründet, erwarb der Standort sich schnell einen besonderen Ruf durch die hervorragende technische Ausrüstung und die für die damalige Zeit modernen Lehrmethoden. In den 50er Jahren des letzten Jahrhunderts erfolgte die Umwandlung zur „Ingenieurschule für Maschinenbau“. 1991 wurde die Fachhochschule Schmalkalden gegründet. Heute hat sich die FH Schmalkalden mit über 2800 Studenten in acht Studiengängen einen festen Platz innerhalb der Thüringer Hochschulen erobert.

Mit ca. 225 Teilnehmern aus dem Hochschulbereich und der Industrie fand auch diese unter dem Motto ‚SAS in Hochschule und Wirtschaft’ stehende Tagung wieder ein großes Interesse. Wie die vorangehenden Konferenzen hatte auch die 8. KSFE das Ziel, ein deutschsprachiges Forum für SAS-Anwender aus dem Hochschul- und Forschungsbereich zu schaffen, deren Hauptanwendungsgebiete in der Datenverarbeitung und -analyse liegen. Diesem Ziel entsprachen die Vorträge, Tutorien und Demonstrationen zu den Schwerpunktthemen Betriebswirtschaftliche Anwendungen insbesondere Controlling, Daten-Management, Data Mining und Web Mining, SAS Anwendungen in Medizin, Naturwissenschaft, Statistik und Informatik, Tipps und Tricks zur SAS Anwendung sowie Ausbildung in und mit SAS. Die Pausen und die Abendveranstaltungen boten eine gute Gelegenheit zum Gedanken- und Informationsaustausch unter den Konferenzteilnehmern.

Dieser Tagungsband enthält neben den Langfassungen der meisten der Konferenzbeiträge auch eine Referentenliste. Die CD enthält zusätzlich das Programm, die Kurzfassungen und Bilder von der Veranstaltung.

Wie in den vergangenen Jahren wurde auch dieses Mal für den besten Beitrag ein Preis ausgeschrieben. Der Preis, Gutscheine für den Besuch einer Schulung bei SAS Institute im Wert von € 1500,-, wurde von der Firma

Systematika GmbH in Heidelberg gestiftet. Die Vorträge wurden von den Konferenzteilnehmern anhand eines kleinen Fragebogens, der jeweils zu Beginn der Vortragssitzung unter den Zuhörern verteilt wurde, bewertet. Die Gewinnerin dieser Ausschreibung war Frau Stephanie Roll vom Charité-Klinikum Berlin.

Wir möchten den folgenden Firmen danken, die durch ihr Sponsoring diese Konferenz maßgeblich unterstützten: dem Hauptsponsor SAS (Heidelberg) sowie iCASUS (Heidelberg), HMS (Heidelberg) und Condat AG (Berlin).

Insbesondere sei allen beteiligten Mitarbeitern des örtlichen Organisationsteams des Fachbereichs Informatik der Fachhochschule Schmalkalden für ihren Einsatz vor, während und nach der Konferenz gedankt. Durch ihr Engagement wurde der Erfolg dieser Konferenz erst möglich.

Unser Dank gilt auch dem KSFE Organisationsteam für die gute Zusammenarbeit bei der Vorbereitung und Durchführung der Konferenz. Deren jahrelange Erfahrung erleichterte die Arbeit vor Ort erheblich.

Und schließlich danken wir auch Frau Henriette Höhle vom Universitätsrechenzentrum Heidelberg für die Erstellung des Tagungsbandes und den Mitarbeitern des Instituts für Medizinische Informatik der Universität Gießen für die Erstellung der CD.

Zum Schluss möchten wir noch darauf hinweisen, dass die 9. KSFE vom 3. – 4. März 2005 an der Humboldt Universität zu Berlin stattfinden wird. Nähere Informationen dazu finden Sie unter www.ksfe.de.

Schmalkalden/Heidelberg, im August 2004

Prof. Dietmar Beyer
FH Schmalkalden, FB Informatik
Chairman

Dr. Carina Ortseifen
Universität Heidelberg
Vorsitzende KSFE e.V.

Inhalt

H.-P. Altenburg **1**
Über die multivariate Aufdeckung von Ausreißern in Kohortenstudien

S. Beimel **9**
Erzeugen von PDF-Dateien ohne ODS: Das Makro Lst2Pdf

A. Christmann **17**
Über statistische Probleme bei der Analyse von Daten aus dem Bereich der Kraftfahrzeugversicherung

Ch. Gottermeier **25**
Datenqualität im Data Mining

S. Hentsch, A. Aulhorn **61**
Data Mining zur Prognose der In-Vitro-Fertilisation

W. Himmel, F. Krolll **69**
Text Mining zur Analyse der Anfragen an ein medizinisches Expertenforum

B. Jäger, M. Wodny, P.E. Rudolph **81**
Simulationen mit SAS, dargestellt für die Erzeugung der Prüfgröße des David-Tests

M. Kappler **97**
Vorstellung eines SAS-Makros zur Dokumentation von Programmen in Multi-User Umgebungen

O. Kruse **105**
Verborgene Schätze heben: Die Offset-Option in PROC GENMOD

A. Mangold, M. Pilz **119**
SAS Enterprise Guide erweitern mit Add-Ins
Anleitung zum Einstieg

S. Menzler, G. Knapp **151**
Powervergabe bei Klinischen Studien mit flexiblen Designs

R. Minkenberg **165**
Möglichkeiten und Grenzen der Steuerung von MS Excel durch SAS mittels DDE – Theorie und Praxis

R. Muche, Ch. Ring, Ch. Ziegler **173**
Ein SAS-Makro Paket für die Entwicklung und Validierung eines logistischen Regressionsmodells

D. Neubert, R. Kaluscha, M. Urban **187**
Integration von SAS V8 mit einer Oracle-Datenbank in einen wissenschaftlichen Arbeitsplatz unter Linux

A. Oelerich, Th. Poddig **195**
Evaluierung quantitativer Ratingverfahren

C. Ortseifen **213**
Schritt für Schritt-Einführung in die Prozedur Tabulate

C. Ortseifen, G. Pfister, O. Schoffer, L. Tschiersch **237**
Tipps und Tricks für einen leichteren Umgang mit der SAS Software

A. Reuß, H.J. Zwiesler **265**
Stornoanalyse in einem Unfallversicherungsbestand

B. Richter, Th. Dittmann **277**
Sächsische Statistiken mit SAS

S. Roll **287**
Multiple Imputation in SAS

P.E. Rudolph, A. Tuchscherer, B. Jäger, M. Tuchscherer **295**
Bootstrap in der Diskriminanzanalyse

Th. Rüdiger **311**
SAS Dictionary Tables zur Nutzung von Selektionen am Beispiel des AXA-CRM-Explorers

J. Spilke, H.P. Piepho, N. Mielenz, X. Hu **319**
Einhaltung des statistischen Risikos 1. Art in unbalancierten gemischten linearen Modellen bei unterschiedlichen Restriktionen der REML-Schätzer und Berechnung deren Schätzfehler

B. Tawakkoli **333**
Wechselwirkung zwischen internetbasierten Diskussionsforen und Kapitalmarkt – Eine empirische Analyse unter Anwendung des Leistungsspektrums von SAS