

Berichte aus der Statistik

**Andreas Hilbert,
Ralf Minkenber**g (Hrsg.)

KSFE 2012

Proceedings der 16. Konferenz der SAS®-Anwender
in Forschung und Entwicklung (KSFE)



Shaker Verlag
Aachen 2012

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Copyright Shaker Verlag 2012

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

Printed in Germany.

ISBN 978-3-8440-1306-1

ISSN 1619-0963

Shaker Verlag GmbH • Postfach 101818 • 52018 Aachen

Telefon: 02407 / 95 96 - 0 • Telefax: 02407 / 95 96 - 9

Internet: www.shaker.de • E-Mail: info@shaker.de

Vorwort

Der vorliegende Proceedingsband enthält die 30 ausgearbeiteten Fassungen der Beiträge zur 16. Konferenz der SAS Anwender in Forschung und Entwicklung (KSFE). Die Konferenz selbst fand vom 07. – 09. März 2012 erstmals in Dresden statt und war so die östlichste KSFE überhaupt. Ausgerichtet wurde die Tagung vom Business Intelligence Competence Center der TU Dresden unter Leitung von Prof. Dr. Andreas Hilbert in Zusammenarbeit mit dem KSFE e.V. – unterstützt durch die Firma SAS Deutschland sowie zahlreiche weitere Sponsoren wie HMS Analytical Software GmbH, iCASUS GmbH, Systematika Information Systems GmbH, viadee Unternehmensberatung GmbH, JMP Software D-A-CH, dsquare.de, KYBEIDOS Gesellschaft für Systeme zur Unternehmenssteuerung mbH, Condat AG und Pharmanet I3.

Das Hörsaalzentrum der TU Dresden war an den drei Tagen der Konferenz fast vollständig in der Hand von rund 400 Datenanalysten, Statistikern, Wissenschaftlern und Unternehmensberatern aus Deutschland, Österreich, der Schweiz, den USA und Großbritannien, die ihre Erfahrungen und Kenntnisse im Umgang mit der SAS-Software teilten, diskutierten und vertieften. Für den Mittwoch meldete sich dabei knapp die Hälfte der Gäste sowohl für Tutorien zur praktischen Arbeit mit SAS-Produkten als auch für eine Zertifizierung im Umgang mit der SAS-Programmiersprache an. Diese international anerkannte Zertifizierung wurde in Dresden erstmalig im Rahmen einer KSFE-Tagung angeboten und fand bemerkenswert positiven Anklang bei den Teilnehmern. An den beiden übrigen Tagen wurden in den Hörsälen des Hörsaalzentrums parallel Beiträge zu aktuellen Themen in den klassischen Bereichen der Statistik, des Data Mining und des Text Mining, aber auch zu modernen Problemstellungen wie der Social Media Analysis, dem Risk Management oder der High-Performance-Analytics präsentiert. Daneben gab es noch zahlreiche kleine und große „Tipps & Tricks“, die die Grundidee der Konferenz – den Austausch von Erfahrungen – unterstrichen. Weitere Diskussionsmöglichkeiten gab es natürlich auch im Rahmen der Poster-Ausstellung, in den Pausen, bei den Ständen der zahlreichen Aussteller oder auch während der Abendveranstaltung.

Ein Highlight der Konferenz war dann noch die Key-Note von John Sall. Im Kontext der High-Performance-Analytics erläuterte der Amerikaner im voll besetzten Hörsaal, wie moderne Hochleistungs-PCs statistische Analysen in der heutigen Zeit verändern. John Sall, Mitinhaber des amerikanischen Unternehmens SAS Institute gründete im Jahr 1976 mit drei Partnern das Unternehmen SAS Institute und ist heute noch im Vorstand des Unternehmens aktiv. 1989 entwickelte Sall die Visualisierungserweiterung JMP, die inzwischen in vielen SAS-Produkten Einzug gefunden hat. Ebenfalls im Kontext von JMP ging dann auch der diesjährige KSFE Best Paper Award an Prof. David Meintrup von der Hochschule für Angewandte Wissenschaft aus Ingolstadt. In seinem Vortrag erläuterte er sein Vorgehen bei der statistischen Versuchsplanung mit JMP.

Dass die Tagung bei den Teilnehmern regen Anklang fand, zeigte sich auch in der Resonanz der Feedback-Bögen, die im Anschluss an die Tagung ausgewertet wurden. An dieser Stelle möchten wir Ihnen als Teilnehmer ein herzliches Dankschön für die konstruktiven Rückmeldungen ausdrücken.

Bedanken möchten wir uns auch für die sehr gute Zusammenarbeit und die wertvolle Unterstützung während der Vorbereitung und der Durchführung der Konferenz beim KSFE-Organisationskomitee. Die Weitergabe der Erfahrungen aus vorausgehenden Konferenzen war dabei besonders hilfreich.

Weiterhin möchten wir Frau Carina Ortseifen und Frau Henriette Höhle vom Universitätsrechenzentrum Heidelberg für die professionelle Erstellung dieses Konferenzbandes danken. Nicht zuletzt möchten wir auch allen beteiligten Mitarbeiterinnen und Mitarbeitern des Business Intelligence Competence Center der TU Dresden für ihren Einsatz vor, während und nach der Konferenz unseren Dank aussprechen. Besonderer Dank geht dabei vor allem an Alexander Müller, Andreas Schieber und Tobias Weiß. Durch ihr großes Engagement wurde der Erfolg dieser Konferenz erst möglich.

Abschließend möchten wir Sie noch auf die 17. Konferenz der SAS Anwender in Forschung und Entwicklung vom 28.02. – 01.03.2013 an der Universität Ulm hinweisen. Veranstalter im kommenden Jahr ist das Institut für Epidemiologie und Medizinische Biometrie unter der Leitung von Prof. Dr. Dietrich Rothenbacher und der KSFE e.V. Vielleicht sind ja auch Sie bei der 17. KSFE 2013 in Ulm mit dabei, um Ihr Fachwissen an andere Teilnehmer weiter zugeben!

Dresden, Juni 2012

Prof. Dr. Andreas Hilbert
Technische Universität Dresden
Chairman

Ralf Minkenberg
KSFE e.V.
Vorsitzender

Inhalt

<i>A. Bachert</i> SAS in der Praxis–Tipps und Tricks für SAS Web Report Studio 4.3	1
<i>P. Bewerunge</i> Das Beste aus zwei Welten Aufruf von R-Funktionen mit PROC IML	13
<i>T. Bruckner, A. Deckert</i> Standardisierte Auswertung von Safety-Daten mit SAS	27
<i>A. Deckert</i> Erstellen flexibler Makros am Beispiel einer SQL-Routine	35
<i>D. Dolic</i> Verwendung von PROC METADATA zur Steuerung eines zentralen Metadatenmodells	51
<i>S. Erbslöh</i> Tipps, Tricks und Stolperfallen zum Umwandeln von Variablen	57
<i>C. Erler</i> Laufzeitoptimierung bei der Verknüpfung großer Datenmengen Ein Vergleich zwischen MERGE und JOIN	77
<i>T. Euler, T. Otte</i> Anwendung eines SAS/STAT-Modells ohne SAS/STAT in einem CRM-Projekt	83
<i>S. Frenzel, S. Baumert, R. Koob, R. Neumüller</i> Performante Erzeugung von Berichten in (vor)formatierten Excel-Dateien	91
<i>B. Gigic, A. Deckert</i> SAS Backstage	103
<i>M. Haffner, A. Mangold, B. Braun</i> PROC SOAP, PROC HTTP und der ganze REST Webservices und SAS	113
<i>B. Heinen</i> JMP10: Sneak Preview für Entwicklung und Modellierung	127
<i>K. Heinrich</i> Social Media Analysis	139

<i>H.-J. Helms, N. Benda, T. Friede</i>	149
Auswertung und Fallzahlberechnung in Dosis-Findungs-Studien mit aktiver Kontrolle mit SAS/STAT und SAS/IML	
<i>B. P. Jäger, C. Malsch, P. E. Rudolph, K.-E. Biebler</i>	159
Randomisierungsverfahren mit SAS	
<i>J. Lang</i>	177
Einstieg in Text Analytics für SAS Enterprise Guide Anwender: Von der Datenquelle zum Bericht - und darüber hinaus...	
<i>M. Malzer, A. Pfahlberg, W. Uter, J. Hegewald, O. Gefeller</i>	193
SAS Makro zur CART-Analyse am Beispiel eines dermatologisch-epidemiologischen Datensatzes	
<i>B. Mayer</i>	207
Der Umgang mit fehlenden Werten in Verlaufsstudien im Falle von Drop-outs	
<i>D. Meintrup</i>	219
Statistische Versuchsplanung mit JMP – von der Klassik zur Moderne	
<i>R. Muche, B. Mayer</i>	235
Medizinische Statistik mit ??? Ein Plädoyer für eine Windows-angenäherte SAS-Oberfläche	
<i>K. Mühlenbruch</i>	249
Ein Makro zur Berechnung von Diskriminanz- und Reklassifizierungsstatistiken für die Verbesserung eines Prädiktionsmodells bei Anwendung der Cox-Regression	
<i>H. Ramroth</i>	263
Proc Transpose oder Do-it-yourself	
<i>S. Reimann</i>	273
SAS DataStep Component Interface – Neue Objekte im DataStep	
<i>S. Reimann</i>	283
ODS Report Writing Interface – Neue Möglichkeiten der Berichterstellung	
<i>A. Schieber, A. Hilbert</i>	297
SAS Enterprise Miner in der Lehre – Von der Einführung zur Zertifizierung	

<i>H. Schmidt, C. Tilke</i>	307
Ein modulares Statistikmakro-Paket für direkte und indirekte Meta-Analysen	
<i>J. Schmidtke, W. Mönkemeyer, K. Schmidt</i>	325
Methoden der Risikobewertung	
<i>O. Schoffer, A. Niedostatek, S. J. Klug</i>	335
Berechnung des relativen Überlebens auf Basis von Krebsregisterdaten	
<i>H.-F. Ulbrich</i>	349
Modellierung rechts-zensierter Zählraten ein SAS-Makro unter Nutzung der Prozedur NLMIXED	
<i>A. Zenk, V. Michel, J. Schmidtke</i>	357
Tipps zur SAS-Programmierung in der PIAFStat-Umgebung	