

SAS-Analyst – Kurse für Mediziner: erste Erfahrungen

Rainer Muche, Andreas Habel

Abteilung Biometrie und Medizinische Dokumentation, Universität Ulm

Telefon: 0731 / 50-26903

eMail: rainer.muche@medizin.uni-ulm.de

Abstract

Von der Abteilung Biometrie und Medizinische Dokumentation werden seit einigen Jahren Einführungskurse in die SAS-spezifische Syntaxsprache für Mediziner an der Universität Ulm angeboten. Dabei hat sich gezeigt, daß es einen Bedarf für eine einfach gehaltene, menügesteuerte Oberfläche für diejenigen Nutzerkreise gibt, die nur einmal eine Auswertung, z.B. für eine medizinische Dissertation, benötigen bzw. nur selten kleine Standardauswertungen auszuführen haben. Für diese Nutzer halten wir das Erlernen des komplexen SAS-Systems, speziell der SAS-Syntaxsprache, für sehr aufwendig.

Die Oberfläche SAS-Analyst ist ein Zusatzpaket zum Softwarepaket SAS ab der Version 6.12. Sie beinhaltet eine Dateneingabemöglichkeit, die sich an Tabellenkalkulationsprogramme anlehnt. Viele Standardverfahren der statistischen Auswertung sind von dieser Oberfläche maus- und menügesteuert erreichbar. Zusätzlich zum Lieferumfang von SAS bietet die Analyst-Oberfläche einige Möglichkeiten z.B. zur Fallzahlberechnung und Poweranalyse an, so daß sie auch für Nutzer der "normalen" SAS-Oberfläche interessant sein kann.

Einleitung

Letztes Jahr haben wir auf der KSFE von den Erfahrungen bei der Durchführung verschiedenster SAS-Kurse für Mediziner und Medizinische Dokumentare berichtet [1]. Dabei hat es sich gezeigt, daß es Nutzergruppen gibt, für die die bisher angebotenen Einführungskurse in die SAS-Syntax nicht besonders geeignet sind. Deshalb haben wir damals einen SAS-Kurs auf Basis der Analyst-Oberfläche angekündigt.

Problemstellung

Das Ausbildungsumfeld, in dem wir tätig sind, ist die Medizin. Dies ist ein Bereich, in dem in der Ausbildung und Forschung die Biometrie immer wichtiger wird und dementsprechend die Software eine wesentliche Rolle spielt. Unser Ziel der Bemühungen mit dem Angebot von verschiedenen SAS-Kursen in den letzten Jahren ist es nun, den Nutzern den Einstieg zu erleichtern.

Bei unserem Klientel haben wir häufig beobachtet, daß ein Einstiegskurs in die SAS-Syntax nicht das geeignete Mittel ist, sie in die Lage zu versetzen, eigene kleine Auswertungen durchzuführen. Ein wichtiger Punkt war und ist es, daß das Erlernen einer Syntaxsprache am Anfang sehr zeitintensiv ist. Dazu kommt, daß die Teilnehmer oft kein Fingerspitzengefühl für Syntax mitbringen. Formalia, wie z.B. das Semikolon haben die Teilnehmer nach einigen Tagen immer noch nicht verinnerlicht. Das Angebot von Konkurrenzprodukten, die mit Maus und Menü bedient werden können, steigert auch nicht unbedingt die Lernmotivation. Häufig wird ja sogar Excel als Statistikpaket benutzt!

Weiterhin haben wir aus gutem Grund SAS-GRAPH nicht in den Einführungskurs eingebunden. Da in der Medizin Grafiken in Veröffentlichungen und Dissertationen aber immer benötigt werden, war auch dieser Punkt ein Anlaß, über Alternativen nachzudenken.

Unsere Folgerung war dementsprechend, für unsere Statistiksoftwarekurse ein maus- und menügesteuertes System einzusetzen. Da an unserer Universität über das Rechenzentrum nur SAS über eine Landeslizenz angeboten wird, haben wir uns natürlich gefreut, das mit SAS-ANALYST eine solche Oberfläche eingeführt wurde.

Zielgruppendifkussion

Die Zielgruppen, die wir versuchen zu erreichen, sind in der Medizin folgendermaßen zu umschreiben: erstens einmal Medizinstudenten, die während des Studiums eine Dissertation anfertigen. Dazu müssen sie in der Regel einen wohldefinierten Datensatz einmal auswerten. Zweitens Kliniker, die während ihrer Assistenzzeit einige kleinere Veröffentlichungen herausbringen, um in ihrer Karriere weiterzukommen. In der Regel reichen dafür eher „einfache“ statistische Standardmethoden.

Die Erwartungshaltung der Personen in diesen Zielgruppen möchten wir folgendermaßen zusammenfassen:

1. sie möchten eine kurze Einarbeitungszeit, da die Daten meist schon vorliegen,
2. sie favorisieren eine Oberfläche, die sich an Bekanntem orientiert,
3. sie brauchen eine Oberfläche, die wenige spezifische Statistikkenntnisse voraussetzt,
4. sie wünschen ein System, in dem Dateneingabe und Grafikbearbeitung integriert sind, um nicht mit mehreren Systemen arbeiten zu müssen.

Aus dem Vorhergesagten soll aber auch folgern, daß eine solche Oberfläche nicht ein Allheilmittel sein kann. Folgenden Zielgruppen werden wir weiterhin empfehlen, die SAS-Syntax zu erlernen, da sie ein viel mächtigeres Werkzeug als eine maus- und menügesteuerte Oberfläche ist:

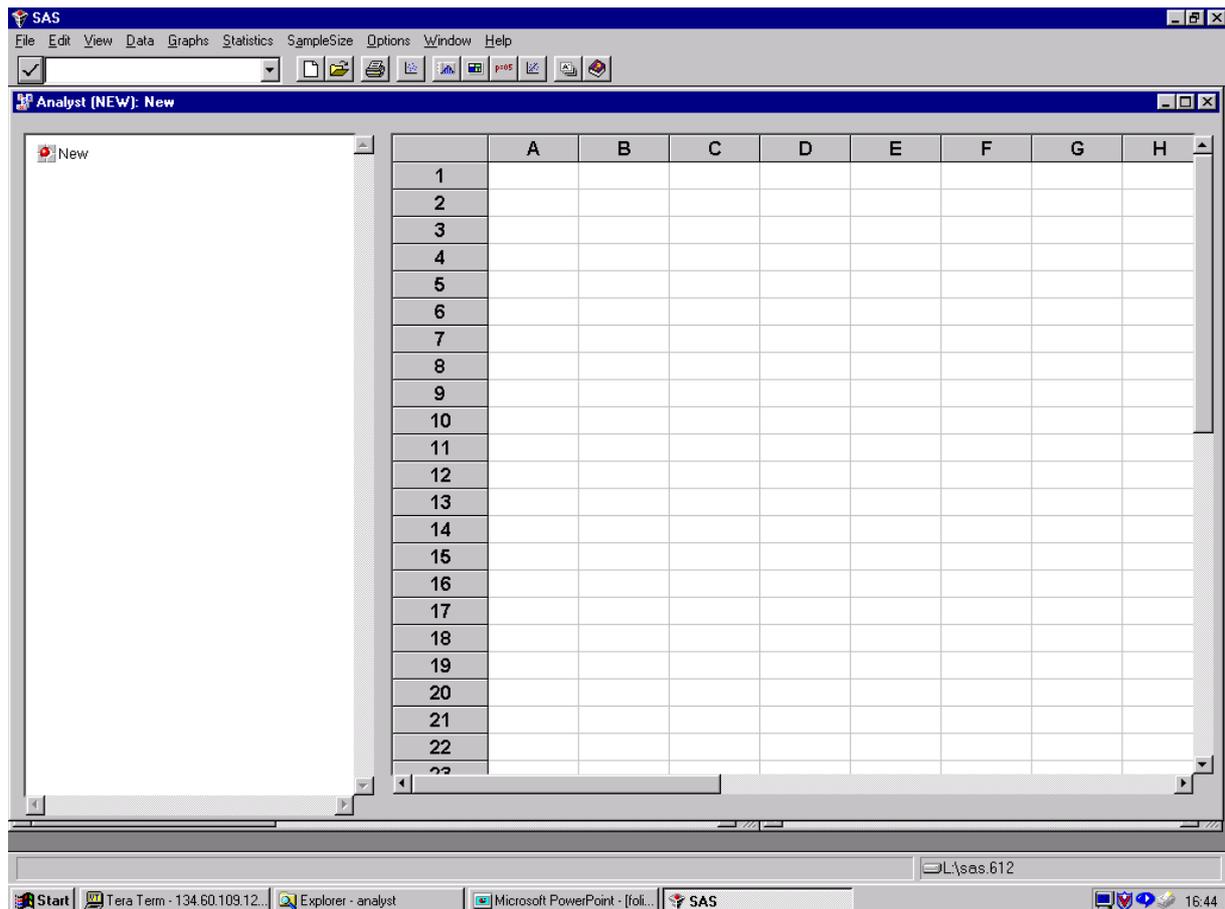
1. Kliniker, die häufig Daten auszuwerten haben,
2. Kliniker, die an einem großen Projekt arbeiten und dort regelmäßig Auswertungen zu erstellen haben,
3. Anwender multivariater Methoden, da diese nicht alle umgesetzt sind,
4. Personen aus Berufsgruppen, die professionell mit Daten und deren Handhabung zu tun haben (z.B. Medizinische Dokumentare).

SAS-Analyst-Oberfläche

Im Folgenden sollen jetzt einige Spezifika von SAS-ANALYST beschrieben werden, die wir auf unsere Anwendung untersucht haben. Zuerst einige Hinweise für diejenigen von Ihnen, die das Zusatzpaket noch nicht haben, es aber mal gerne angucken möchten. Analyst.exe kann kostenlos über die Homepage von SAS (WWW.SAS.COM) unter den Stichworten Download und Version 6 heruntergeladen werden. Sie benötigen dazu die Site-Number ihrer SAS-Installation.

Wenn Sie das File in Ihr SAS-Hauptverzeichnis kopieren (SAS.EXE steht da) und durch doppelklick öffnen, installiert sich das Paket selbständig in ihr SAS-System. Aufgerufen wird SAS-Analyst über Globals --> Analyze --> Statistical Analysis. Dieser letzte Menüpunkt wird bei der Installation angelegt.

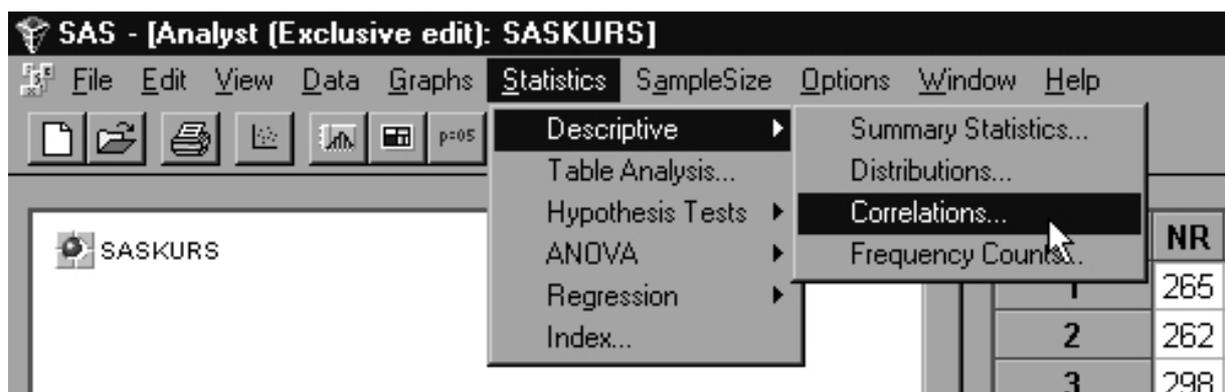
Die sich dann öffnende Oberfläche sieht folgendermaßen aus: oben gibt es eine Windows-angepaßte Menüleiste, links ein Fenster, in dem die durchgeführten Auswertungen in einer Baumstruktur angezeigt und wieder ausgewählt werden können. Im rechten Teil gibt es ein Tabellenblatt, in das Daten eingegeben oder eingelesen werden können.



Biometrische Methoden auf der Analyst-Oberfläche

Jetzt noch einige Anmerkungen zu den Auswertungsmöglichkeiten auf der Analyst-Oberfläche. Auf der Oberfläche finden sich nicht alle Auswertungsmöglichkeiten von SAS (wie z.B. in SPSS). Es wurden einige umgesetzt, allerdings finden sich nicht immer alle Optionen, die man in der SAS-Syntax zur Verfügung hat.

An Oberthemen werden hier die Deskriptive Analyse (FREQ, MEANS, UNIVARIATE, CORR) und Mehrfeldertafeln (FREQ) angeboten. An Hypothesentests finden sich Ein- und Zweistichproben tests für Raten und Mittelwerte. Unter ANOVA finden sich ein- und mehrfaktorielle Varianzanalysen. Hier findet sich z.B. auch NPAR1WAY als nichtparametrische, einfaktorielle Varianzanalyse. Unter REGRESSION findet man lineare Regressionsmodelle (REG) und logistische Regression (LOGISTIC). Über den INDEX läßt sich auch durch Ausschauen der statistischen Methodik das entsprechende Menü finden.



Im Wesentlichen stecken unter den Menüs die altbekannten Prozeduren. Aber manchmal wurden Auswertungen neu programmiert. Zum Beispiel wurde der Zweistichproben-t-Test nicht über PROC TTEST eingebunden, sondern komplett neu programmiert.

Wir haben die Analyst-Oberfläche darauf hin untersucht, ob die statistischen Methoden, die unsere Zielgruppe aus der Biomathematik-Vorlesung kennen, alle vorhanden sind. Die gute Nachricht: bis auf Überlebenszeitanalysen (Kaplan-Meier und Log-rank-Test) ist alles da. Diese werden in der Version 8 folgen. Die nicht so gute Nachricht: manches ist versteckt und muß mühsam im Kurs und den Begleitunterlagen erläutert werden. Auf die Frage von Teilnehmern nach dem „Warum das so ist“, muß man dann meist passen.

Vor- und Nachteile der Analyst-Oberfläche

Wir sehen also folgende Vorteile in der Nutzung dieser Oberfläche:

- die Oberfläche ist maus- und menügesteuert, so daß die vielen Syntaxfehler wegfallen
- die Oberfläche ist Windows-ähnlich, so daß keine Hemmschwelle bei der Nutzung existiert
- Daten aus Standardpaketen wie Excel können eingelesen werden
- durch die Beschränkung auf ausgewählte Analysemethoden wird allerschlimmstes Datenquälen eher verhindert
- es gibt einen Zusammenhang zur SAS-Syntax, so das die Analyst-Oberfläche als Einsteiger-Möglichkeit in SAS genutzt werden kann.

Wo es Vorteile gibt, gibt es auch fast immer Nachteile, so auch hier:

- der wichtigste ist das Fehlen eines Handbuchs bzw. einer Online-Hilfe. Für die nächste Version gibt es ein Handbuch [2]
- das Tabellenblatt verführt dazu, es ohne nachzudenken anzuwenden. Da man fast sofort auf Fehler und Probleme stößt, schreckt dies die Anwender sofort ab und dieses Urteil wird auf das ganze Produkt übertragen.
- SAS ist sehr mächtig im Datenmanagement. Wenn in der Analyst-Oberfläche daran gespart wird, verliert die Software stark an Attraktivität. Set und Merge sollen in Version 8 laufen.
- Die Klassierung von Variablen ist in dieser Version so umgesetzt, daß man kaum selber darauf kommt. Wir haben es wochenlang versucht und dann bei SAS nachgefragt.
- eine Beschränkung in den Sortiermöglichkeiten des Datensatzes.
- es sind nicht alle Optionen der hier eingebundenen Prozeduren abrufbar. Das ist unverständlich, da doch nur ein Feld auf der Menümaske eingebaut werden muß. Wenn man einige Optionen schon umsetzt, warum dann nicht alle?
- Die Inkonsequenz bei der Erzeugung des Source-Codes ist ärgerlich, wenn man auf dem Analyst-Kurs dann mit einem Syntax-Kurs aufbauen möchte.
- In der Analyst-Oberfläche ist auch der Grafikeditor des SAS-Systems integriert, der auf jedenfall viel zu unzulänglich ist.

Wir haben uns trotzdem entschlossen, einen solchen Analyst-Kurs anzubieten aus folgenden Gründen:

1. an der Uni Ulm ist nur SAS lizenziert,
2. die angebotenen Methoden sind für Anwender fast immer ausreichend,
3. die Anwender gewöhnen sich an SAS, so daß diejenigen, die doch häufig auswerten, leichter auf SAS-Syntax zurückgreifen können.

Unser Analyst-Kurs

Kursprinzipien

Bei der Umsetzung des Kurses verfolgen wir folgende Prinzipien:

Wir möchten nicht eine allgemeine Einführung geben, sondern von den biometrischen Auswertungsproblemen auf die Softwareumsetzung kommen und die Lösung auf der Analyst-Oberfläche zeigen. Es soll keine Vorlesung sein, sondern der Stoff soll direkt am Rechner nachvollzogen werden. Da die Nutzer später meist einen Datensatz auszuwerten haben, werden wir den Stoff auch an einem durchgehenden Beispiel mit realen Daten präsentieren. Damit die Teilnehmer vertraut werden im Umgang mit der Software, bieten wir Selbstlernphasen an, die betreut werden.

Skript

Da wir annehmen, dass das Gelernte nicht hängenbleibt, wenn man keine weitere Hilfe oder Nachschlagewerke anbietet, haben wir ein Begleitskript für den Kurs geschrieben. Da es hier um Erfahrungen geht, kann ich hier sagen, daß das ein sehr mühsames Geschäft ist, so ein Skript zu schreiben.

Kursplanung

Im Folgenden möchte ich unseren so gedachten Kursablauf kurz darstellen (Kapitel-Anmerkungen beziehen sich auf unser Skript):

SAS-Analyst - Kursplanung 1. Tag:

- | | |
|-------------|---|
| 8.30-9.15 | Begrüßung
Ziel und Übersicht
Skript erklären (Kap. 1)
An- und Abmelden (Kap. 2)
Manuelle Dateneingabe (Kap. 3.1) und Exclusive Edit (Anhang 3)
Import von Excel (Kap. 3.2 und Kap. 3.3)
Beispieldatensatz vorstellen und importieren (Anhang 2) |
| 9.15-10.00 | SAS-Dateien und Datensätze modifizieren (Kap. 4)
SAS-Formate (Anhang 6) |
| 10.00-10.30 | Pause |
| 10.30-11.00 | Dekriptive Beispielauswertung
Technische Hilfen bei Auswertungen (Kap. 5)
Arbeiten mit SAS-Dateien (Anhang 5) - nicht LIBNAME |
| 11.00-12.00 | Skalenniveau (Kap. 6.1)
Deskriptive Statistik (Kap. 7) |
| 12.00-13.00 | Pause |
| 13.00-15.00 | Selbständiges Üben |

SAS-Analyst - Kursplanung 2. Tag:

8.30-9.00	Wiederholung
9.00-10.00	Einführung in Zusammenhangsauswertungen (Kap. 6.3) Korrelation und Regression (Kap. 8)
10.00-10.30	Pause
10.30-11.30	Einführung in stat. Tests und Konfidenzintervalle (Kap. 6.4-6.6) Tests für unabhängige Beobachtungen (Kap. 9) Tests für abhängige Beobachtungen (Kap. 10) Einstichproben tests und Konfidenzintervalle (Kap. 11)
11.30-12.00	Grafiken bearbeiten (Anhang 7) Ausblick: Hinweise auf weitere Möglichkeiten z.B. Fallzahlberechnungen (Kap. 12) Literaturhinweise (Anhang 8)
13.00-15.00	Selbständiges Üben

Erfahrungen

Jetzt noch etwas zu unseren "Erfahrungen":

Am 29. November 1999 brannte es an der Universität. Das Rechenzentrum ist zwar nicht direkt abgebrannt, aber Labore direkt unter dem Rechenzentrum. Unter anderem betroffen ist die Nutzung des Schulungs-PC-Pools. Seit November 1999 können deshalb keine Kurse durchgeführt werden.

Wir haben einen Kurs kurz vor dem Brand durchgeführt mit freiwilligen Teilnehmern aus dem Bekannten- und Kollegenkreis. Diese sind sicher nicht repräsentativ für unsere Zielgruppe. Es hat sich aber gezeigt, daß der Plan durchzuführen ist. Einige kleine Korrekturen werden zu machen sein. Uns hat am Ende Zeit gefehlt für eine Kursreflektion und Hinweise, was die Teilnehmer nach dem Kurs an Hilfen erwarten können. Das Skript sollte vor dem Kurs zugestellt werden, da das Nachlesen und Blättern während des Kurses sehr viel Zeit verbraucht. Insgesamt war dies noch kein echter Test.

Literatur

- [1] R. Muche, G. Büchele, A. Imhof, A. Habel:
Erfahrungen mit SAS-Kursen für unterschiedliche Nutzergruppen
Proceedings der 3. KSFE, Heidelberg, 1999, S. 177-182
- [2] The Analyst Application, Users Guide, 56281