

Nutzung von SAS-Studio unter SAS OnDemand for Academics anhand eines Lernskriptes

Gisela Büchele
gisela.buechele@uni-ulm.de

Martin Rehm
martin.rehm@uni-ulm.de

Raphael S. Peter
raphael.peter@uni-ulm.de

Leonie Hezler
leonie.hezler@uni-ulm.de

Judith Vilsmeier
judith.vilsmeier@uni-ulm.de

Rainer Mucbe
rainer.mucbe@uni-ulm.de

Universität Ulm
Institut für Epidemiologie und
Medizinische Biometrie
Helmholtzstraße 22
89081 Ulm

Zusammenfassung

In vielen Studiengängen im Bereich Gesundheitswesen, Life Science oder anderen verwandten Fachrichtungen ergibt sich während des Studiums die Notwendigkeit, empirische Daten mit einem statistischen Auswertungsprogramm zu analysieren.

Jedoch zeigt die Erfahrung, dass das Erlernen von SAS auf der Programmierenebene zeitintensiv ist. Die SAS-Studio-Oberfläche bietet sich als sinnvolle Alternative an, da sie eine deutschsprachige, menü-basierte Auswertung von Daten mit einer großen Bandbreite an deskriptiven, graphischen und test-statistischen Methoden anbietet.

In SAS-Studio kann man nicht nur eigene SAS-Programme schreiben und ausführen, sondern auch die in SAS-Studio enthaltenen vordefinierten Tasks zum Analysieren der Daten verwenden. Die Tasks basieren auf SAS-Systemprozeduren und bieten die Möglichkeit, über eine Point-and-Click-Steuerung Dateien und Variablen auszuwählen sowie anschließend die entsprechende SAS-Syntax generieren zu lassen. Mit Zugriff über *SAS OnDemand for Academics* steht Studierenden nach Registrierung die Software kostenlos, überall und jederzeit in der Cloud zur Verfügung.

Literatur zur Anwendung von SAS-Studio ist bereits verfügbar, jedoch unseres Wissens nach nur auf Englisch und ohne detaillierte Anleitungen zu den Analysen vorausgehenden Datenaufbereitungsschritten. Um eine selbstständige Einarbeitung zu ermöglichen, fertigen wir gerade ein Skript zur biometrischen Auswertung mittels SAS-Studio auf *SAS OnDemand for Academics* an. Dieses Skript soll den Studierenden verschiedenster Studiengänge den Einstieg in SAS ermöglichen.

Schlüsselwörter: SAS OnDemand for Academics, SAS Studio, Task, Lehre

1 Hintergrund

In vielen Studiengängen im Bereich Gesundheitswesen, Life Science oder anderen verwandten Fachrichtungen ergibt sich während des Studiums oder spätestens bei der Bachelor- oder Masterarbeit die Notwendigkeit, empirische Daten mit einem statistischen Auswertungsprogramm zu analysieren. Die Programmieroberfläche von SAS ist dabei nicht immer die erste Wahl, da der Einarbeitungsaufwand in die Programmiersprache sehr groß ist und zusätzlich oftmals noch Lizenzkosten auf die Studierenden zukommen. [1]

Mit SAS-Studio auf dem online-Zugang von *SAS OnDemand for Academics* [2] werden zumindest diese beiden Argumente aufgehoben. Mit *SAS OnDemand for Academics* steht allen Studierenden ein Server mit SAS kostenfrei zu Verfügung, lediglich eine Registrierung ist notwendig. Mit einem Internet-Browser haben registrierte Benutzer jederzeit Zugang zu SAS-Analyse-Anwendungen wie Studio, Enterprise Guide, Enterprise Miner oder Forecast Studio. Die Datenauswertung setzt allerdings ein Hochladen aller Datensätze auf den SAS-Server voraus.

2 SAS OnDemand for Academics

SAS OnDemand for Academics ist eine online bereitgestellte Oberfläche für das Lehren und Lernen von Datenmanagement und Datenanalyse mit SAS.

Folgende Vorteile bei der Verwendung von *SAS OnDemand for Academics* liegen vor:

- Kostenloser Online-Zugriff auf SAS-Software
- Kein Download, keine Installation und Konfiguration von Softwareanwendungen
- Studierende können SAS jederzeit und überall nutzen
- Lehrende können ohne Wartungs- oder IT-Aufwand für die Softwareanwendungen unterrichten
- Die Software ist immer in der aktuellsten Version in der Cloud verfügbar

Voraussetzung für die Verwendung des Angebotes ist eine nicht-kommerzielle Nutzung und eine Registrierung für *SAS OnDemand for Academics*. Bei der Registrierung wird ein SAS-Profil mit einem Benutzernamen und einem Passwort angelegt.

3 SAS-Studio

SAS-Studio ist eine webbrowsers-basierte Programmierumgebung für die SAS-Software. Durch die Bereitstellung in einem unterstützenden Rahmen (mit z.B. Syntax-Autovervollständigungsfunktion und Popup-Tipps zu Schlüsselwörtern und Prozeduren) eignet sich SAS-Studio auch gut für den Einstieg in die Programmierung mit SAS.

SAS-Studio ermöglicht die deutschsprachige, menü-basierte Auswertung von Daten mit einer großen Bandbreite an datentechnischen, deskriptiven, graphischen und test-

statistischen Methoden sowie Regressionsanalysen bis hin zu multivariaten Analysen, Clusteranalysen oder Data Mining.

Vorteile bei der Verwendung von SAS-Studio:

- SAS-Code kann auf einem beliebigen Gerät über einen Webbrowser geschrieben und ausgeführt werden
- Unterstützung bei der Erstellung von SAS-Code durch die Syntax-Autovervollständigungsfunktion und Popup-Tipps zu Schlüsselwörtern und Prozeduren
- Zusätzlich besteht die Möglichkeit, über eine Point-and-Click-Steuerung (Tasks) SAS-Code generieren zu lassen

Im Folgenden wird ein Beispiel für den Einsatz von Tasks gezeigt:

In einer Box-Plot Grafik soll die Verteilung von LDL-Cholesterin-Werten getrennt für die Patientengruppen mit und ohne koronare Herzkrankheit dargestellt werden.



Abbildung 1: Auswahl von Dateien und Variablen in Tasks

Sobald ein Datensatz ausgewählt wurde, können, der Methode entsprechend, Variablen verschiedene Rollen (z.B. abhängige Variable oder Klassifizierungsvariable) zugewiesen werden (s. Abbildung 1).

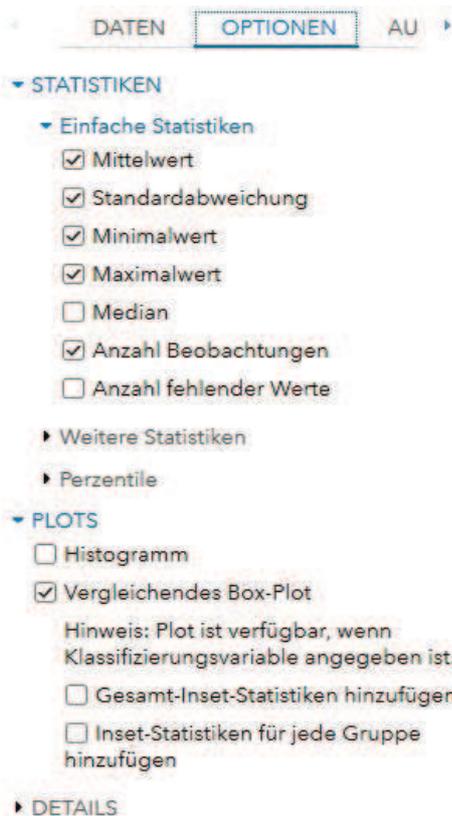


Abbildung 2: Auswahl von Optionen in Tasks

Anhand der angegebenen Informationen und ausgewählten Optionen (s. Abbildung 2) wird automatisch SAS-Syntaxcode generiert, der abgespeichert oder auch verändert werden kann (Abbildung 3).

The screenshot shows the 'CODE' tab of the SAS Task Wizard. It displays the SAS code generated from the selected options. The code is as follows:

```
1 /* Taskcode generiert von SAS Studio 3.7 */
2
3 ods noproctitle;
4 ods graphics / imagemap=on;
5
6 proc means data=WORK.KHKGESAMT chartype mean std min max n vardef=df;
7     var LDL_Cholesterin_mg_dl;
8     class vorhandene_KHK;
9 run;
10
11 proc sort data=WORK.KHKGESAMT out=WORK.TempSorted2236;
12     by vorhandene_KHK;
13 run;
14
15 proc boxplot data=WORK.TempSorted2236;
16     plot (LDL_Cholesterin_mg_dl)*vorhandene_KHK / boxstyle=schematic;
17 run;
18
19 proc datasets library=WORK noprint;
20     delete TempSorted2236;
21 run;
22
```

Abbildung 3: Task-generierter SAS-Syntaxcode

Mit „Ausführen“ wird die Auswertung ausgeführt und im Ergebnisfenster (s. Abbildung 4) dargestellt. Log-Informationen und Ergebnisse werden, wie gewohnt, in separaten Reitern dargestellt.

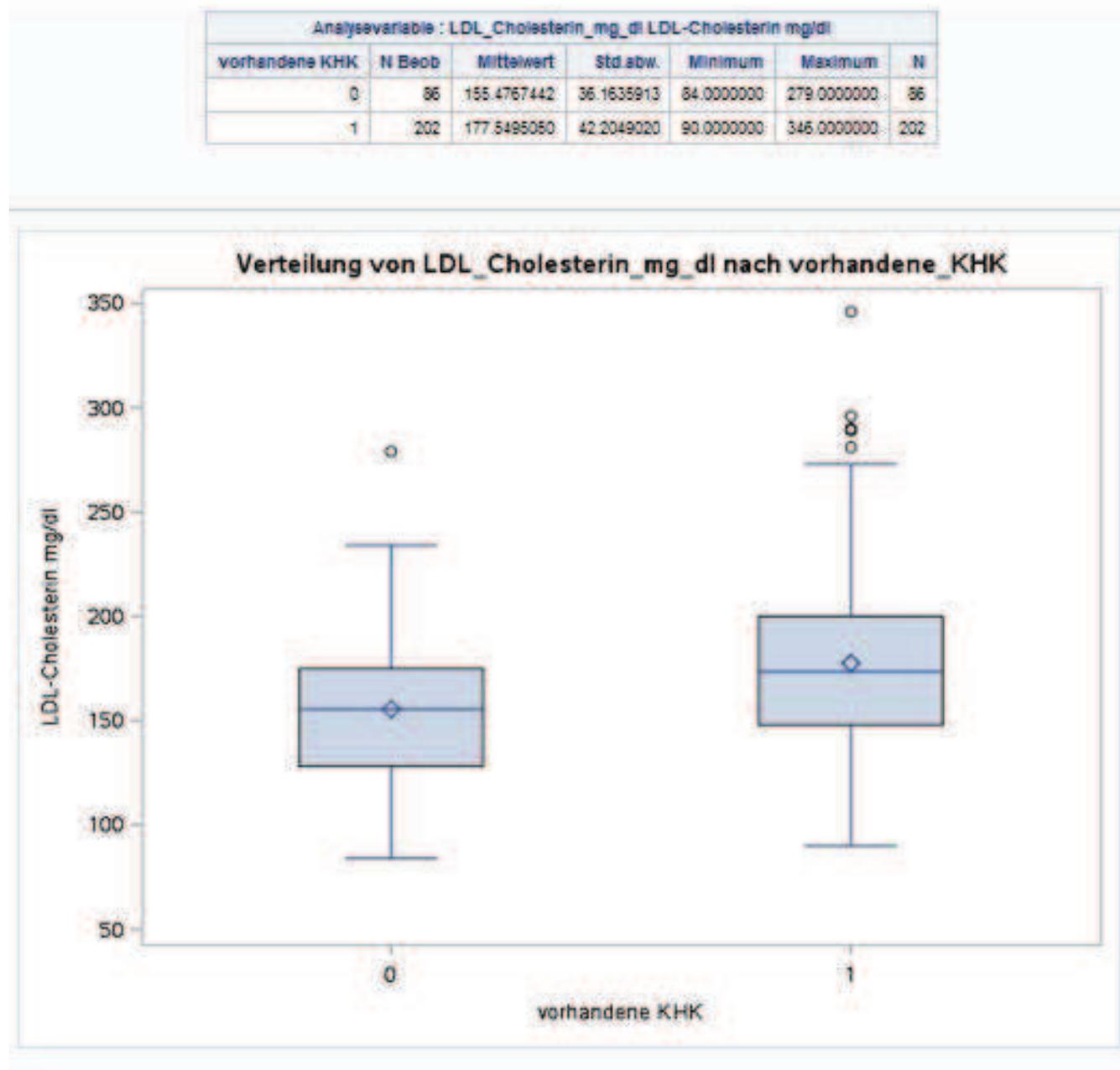


Abbildung 4: Ergebnis der Task-Auswertung

4 Anwendung

In den kombinierten Praktika „Statistik“ und „Epidemiologie“ im Bachelorstudiengang *Mathematische Biometrie* an der Universität Ulm wird seit dem Wintersemester 2017/2018 ein SAS-Programmierungskurs auf der *SAS OnDemand for Academics*-Oberfläche angeboten. Jedoch zeigt die Erfahrung aus diesem und anderen Kursen, dass

das Erlernen von SAS auf der Programmierenebene zeitintensiv ist. Für andere Studiengänge, bei denen einerseits im Curriculum keine umfangreichen Übungen eingeplant sind und andererseits die Kenntnis eines Statistikprogrammes für die spätere Berufswahl nicht zwingend erforderlich ist, ist gerade die SAS-Studio-Oberfläche eine sinnvolle Alternative.

Literatur zur Anwendung von SAS-Studio ist bereits verfügbar, jedoch unseres Wissens nach nur auf Englisch und ohne detaillierte Anleitungen zu den Analysen vorausgehenden Datenaufbereitungsschritten. [3,4] Um eine selbstständige Einarbeitung zu ermöglichen, fertigen wir gerade ein Skript zur biometrischen Auswertung mittels SAS-Studio auf *SAS OnDemand for Academics* an. Es beinhaltet die Datenaufbereitung und verschiedene statistische Analyseverfahren, vom Dateneinlesen über Klassierung von Variablen, deskriptive Statistik und statistische Tests bis hin zur Überlebenszeitanalyse und Fallzahlplanung. Dieses Skript soll den Studierenden verschiedenster Studiengängen den Einstieg in SAS ermöglichen [5].

Die erste Anwendung des Skriptes soll im Seminar „Q1/Biometrie“ im Studiengang *Humanmedizin* an der Universität Ulm im Rahmen eines Lehrprojekts stattfinden. In einem Beitrag von Muche *et al.* [6] werden der Ansatz des „Flipped Classroom“ für einen Statistiksoftwarekurs mit SAS-Studio auf *SAS OnDemand for Academics* sowie der Einsatz des Skripts in einem Lehrprojekt in diesem Tagungsband vorgestellt.

Literatur

- [1] R. Muche, G. Büchele, A. Imhof, A. Habel: Erfahrungen mit SAS-Kursen für unterschiedliche Nutzergruppen. Proceedings der 3. KSFE-Konferenz, Shaker-Verlag, Aachen (1999), 177-182
- [2] SAS OnDemand for Academics. https://www.sas.com/en_us/software/on-demand-for-academics.html (aufgerufen am 26.07.2018)
- [3] R. Cody: Biostatistics by Example using SAS Studio. SAS Institute, Cary NC, 2016
- [4] G. Der, B. S. Everitt. Essential Statistics Using SAS University Edition. SAS Institute, Cary NC, 2015
- [5] G. Büchele, R. Muche: Problem-basiertes Lernen im Rahmen einer SAS-9- Einführungsvorlesung. Proceedings der 10. KSFE-Konferenz, Shaker Verlag, Aachen (2006), 35-44
- [6] R. Muche, G. Büchele, M. Rehm, B. Mayer: Flipped Classroom mit SAS onDemand – SAS-Studio in der Biometrieausbildung im Studiengang Humanmedizin. Proceedings der 23. KSFE-Konferenz, Shaker-Verlag, Aachen (2019), im Druck