

Grafische Datenauswertung in Orlando II mit SAS/Graph[©] und SAS/Insight[©] Software

Hans Schwarz
SAS Institute GmbH

Durch die Integration von beiden Modulen in das SAS System ist es gewährleistet, daß sämtliche zur Verfügung stehenden Datenquellen grafisch ausgewertet werden können.

SAS/Graph Software

Sämtliche gängigen Grafiktypen wie **Liniendiagramme, Histogramme, 3-D-Grafiken, Kreisdiagramme, Konturdiagramme und räumliche Darstellungen auf Grund von Landkarten** werden unterstützt und zwar als Darstellung von Daten beliebiger Herkunft und mit beliebig großen Datenmengen.

Zur Illustration bzw. der Beschriftung der Grafiken stehen neben vielen Software Fonts auch sämtliche Hardware Fonts zur Verfügung.

Mit Hilfe der Prozedur GREPLAY lassen sich **mehrere Grafiken auf einer Ausgabeseite** darstellen.

Darüber hinaus bestehen mehrere Möglichkeiten an einer bestehenden Grafik Verbesserungen oder Veränderungen durchzuführen: ad hoc durch Benutzung des **grafischen Editors** und als wiederverwendbare Lösung für gleiche oder gleichartige Grafiken in Abhängigkeit sich verändernder SAS Dateien über das **Data Step Graphic Interface (DSGI)** und durch den Einsatz spezieller SAS Dateien über die **Annotate Facility**.

Besondere Stärke der SAS/Graph Software sind die vielen **frei definierbaren Schnittstellen zu Ausgabegeräten** mit der Prozedur GDEVICE.

Die weitere Öffnung nach außen bei dem SAS System macht sich auch bei SAS/Graph durch die Einführung von Ex- und Importfunktionen bemerkbar, so daß jetzt Bilddateien im **Bitmap, Tiff JPG Format u.a.** geschrieben werden können.

Mit Orlando II 6.12 werden die aus der SAS/AF Software stammenden grafischen **3-D Objekte** in einer kleinen AF Anwendung (GRAPH-N-GO) nutzbar gemacht. Die Anwendung läßt sich starten mit dem Kommando: **AF C=SASHELP.GRAPH.GOBJECT.FRAME** (ohne SAS/AF Lizenz),

SAS/Insight Software

Heutige Rechnergenerationen lassen es zu, daß statistische Zusammenhänge nicht mehr nur durch eine Unmenge von abstrakten Zahlen also z. B. Koeffizienten, Regressionen etc. dargestellt werden müssen, sondern erlauben eine interaktive und anschauliche Analyse durch grafische Aufbereitung.

SAS/Insight betreibt also Statistik nicht nur zahlenorientiert, sondern vor allem mit grafischen Darstellungen. Dank der grafischen Benutzeroberfläche bietet die Software einfachen Zugriff auf statistische Verfahren und stellt innerhalb des SAS Systems das Werkzeug für die interaktive Datenanalyse dar. SAS/Insight richtet sich sowohl an den grafischen Experten aber auch vor allem an diejenigen, die SAS lediglich als Auswertewerkzeug für ihre wissenschaftlich-technischen Aufgaben benutzen wollen, ohne sich mit statistischen Grundlagen auseinander setzen zu wollen oder zu können.

Für die Datenexploration stehen alle bekannten Techniken zur Visualisierung von Daten und ihrer Veranschaulichung zur Verfügung. Die Inhalte der einzelnen Fenster des Menusystems können interaktiv und dynamisch bearbeitet werden.

Mit der Anzeige von Daten in **Linien- und Balkendiagrammen** sowie den **Boxplots** verfügt SAS/Insight über die klassischen Grafiktypen der Datenvisualisierung

Neben diesen Klassikern kennt SAS/Insight auch **Streuungsdiagramme(Scatter Plot)** mit Regressionslinien, mit denen die 'Beziehung' von Variablen visuell dargestellt wird. Gehen mehr als zwei Variablen in die Analyse mit ein, erzeugt das SAS System eine **Scatter Plot Matrix**.

Räumliche Streuungsdiagramme mit Rotation zeigen eine 3 dimensionale Darstellung. Die Grafik läßt sich frei im Raum bewegen und die dargestellten Beobachtungen können mit Hilfe von Farben und Symbolen gekennzeichnet werden.

Bei der **Verteilungsanalyse** sind zwei Typen von Dichteschätzungen möglich:

Kernschätzung und parametrische Schätzung.

Zahlenorientierte Analysen:

Klassiker:

- Verteilungs- und Korrelationsanalyse
- einfache Vorhersagemodelle
- Hauptkomponentenanalyse

basierend auf **General-Linear-Models:**

- Regressionsanalysen
- Varianzanalysen
- Kovarianzanalysen

basierend auf **Generalized-Linear-Models**

- logistische Regressionsanalysen
- Poisson Regression

Eine Auswahl der wichtigsten Merkmale:

1. Alle grafischen und textuellen Objekte kommunizieren untereinander
2. Visuelle Identifikation von Ausreißern und auffälligen Beobachtungen durch Farben, Symbole und Anklicken
3. freie Achsenskalierung
4. Kennzeichnen von Gruppen von Beobachtungen mit ihren Ausprägungen .
5. Anzeige von Werten, Labels etc.
6. Unterdrückung der Kommunikation durch **Freeze Option**, um den Inhalt von Fenstern besser vergleichen zu können, die aufgrund veränderter Datensätze erstellt worden sind.
7. Brushing

Quellen:

1. SAS/Graph Software
Version 6, 1st Edition, Vol I+II, #56020
2. grafischer Editor
Version 6, 1st Edition, #56023
3. SAS/Insight User's Guide
Version 6, 3rd Edition #55582
Anwendungs- und Nachschlagewerk
4. DISK '93 Proceedings
Heinz-Bernd Spieker: Neuerungen in SAS/Insight Software
5. SEUGI Proceedings '94
Gerhard Held: Dynamic Data Analysis with SAS/Insight Software p 784
6. SEUGI Proceedings '93
Gerhard Held: Data Visualization with SAS/Insight Software
p 394

7. SUGI 17 Proceedings 1992
Heman Robinson: Analytic Modeling in SAS/Insight Software
p 1298
8. SUGI 18 Proceedings 1992
 - i)
p1100
 - ii) Greg Weier, Randall T. Thomas
Exploratory Data Analysis Using SAS Software
p1265