Statistische Datenanalyse mit der Analyst Application

Ulrich Reincke

SAS Institute

The **Analyst Application** is a task-oriented interface that provides access to basic statistical analyses.

Designed for the desktop environment of PCs and workstations, it is a useful tool for students and researchers as well as for experienced SAS users who desire a convenient interface to widely used statistical techniques.
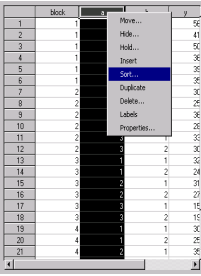

The Analyst Application provides three facilities for accessing data: a native dialog for opening SAS data sets and importing other data files, a dialog for accessing SAS library members, and a link to the query window for creating SQL queries. Using the File > Open dialog, you can open SAS data sets, (SPSS portable files,) Excel spreadsheets, dBase files, and many other PC file formats. The application also provides a task for creating new SAS libraries, and you can start from scratch by typing in new variables and values directly in the Data Table.

The **Data Table** provides a spreadsheet-like view of the open data set. There you can perform elementary data management tasks. You can point-and-click to directly modify your data table to perform operations to sort the table, insert columns, add rows, change column properties, and display variable labels as column headings. In addition, you can type data directly into the data table to create a new data set or modify existing observations.

The application provides a wide range of analytical and graphical tasks. You can compute descriptive statistics, analyze crosstabulations, and fit models with analysis of variance and regression analysis. The application also provides some tasks not covered by SAS procedures, such as a variety of hypothesis tests and basic sample size and power computations.
In Version 8, the statistical capabilities of the Analyst Application have been strengthened to include survival analysis (there is a life tables task and a proportional hazards task (performs Cox regression)), repeated measurements, multivariate statistics, and mixed models.

The Analyst Application provides many graphs that are independent (first column of graphs) of a particular statistical task. You can produce several types of graphs, including histograms, box-and-whisker plots, and scatter plots.

In addition, most statistical tasks include integrated graphs (second column of graphs) appropriate for that specific type of analysis. For example, when you perform a distribution analysis, you can request a histogram of your variable with a superimposed Normal distribution. When you perform a regression analysis, you can request a plot of the residuals versus predicted values. t-distribution, means and interaction plots are also possible.
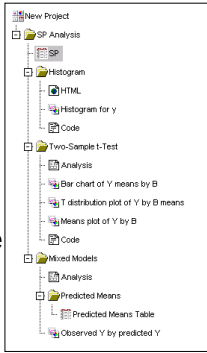
The Analyst Application organizes results at the project level in a Windows Explorer-like **Project Tree**. When you run an analytical or data management task, the application will add a folder to the Project Tree that contains your results, which can include analysis results, output data sets, graphics, SAS code, and HTML output. A single project can include analysis on multiple data sets, and you can click on any data node in the Project Tree to open that data in the Data Table and then proceed to perform further analysis. In addition, you can set preferences for whether nodes for HTML output and SAS code appear in the Project Tree.

The data used for this demonstration is saved in an Excel spreadsheet named BPMULT. This data set includes results from the Treatment of Mild Hypertension Trial (TOMHS) conducted at the University of Minnesota. TOHMS was a clinical trial comparing several antihypertensive medications for optimal treatment of mild hypertension. The data has been extended to include a hypothetical study that evaluates the effects of medication on severe hypertension.

The response variable is systolic blood pressure (SBP) measured at 3, 6, and 9 months on a total of 128 subjects in the treatment group and 230 subjects in the placebo (or control) group. There is also a baseline blood pressure measurement stored as a separate variable. Other variables include clinic, treatment, and a binary variable indicating whether the patient complied with the entire treatment regimen.

Zusammenfassung

Die Analyst Application ist

eine anwendungsorientierte

point-and-click Oberfläche

zur Statistischen Datenanalyse

Release Information

Required Products
- Base SAS
- SAS/STAT
- SAS/GRAPH
- SAS/ASSIST

Added Functionality
- SAS/IML
- SAS/ACCESS

Platforms
Windows 95/98/NT
UNIX environments (HP-UX, Solaris, Digital UNIX, AIX)
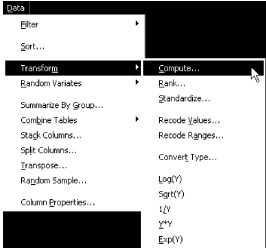Alpha VMS
VAX/VMS
OS/2
...

The Analyst Application is an interface that uses capabilities from several SAS products. In order to run the application, you are required to license **base SAS**, **SAS/STAT**, **SAS/GRAPH**, and **SAS/ASSIST**. The additional functionality provided by **SAS/IML** is limited to scatter plot ellipses in the Correlations task and partial regression plots in the Linear Regression task. **SAS/ACCESS** is required to access a wide range of file formats such as an Excel spreadsheet or Oracle database.

In Version 8, the Analyst Application will be available for Windows 95/98/NT, HP, Solaris, Digital UNIX, RS/6000, Alpha VMS, VAX/VMS, and OS/2. Recent news is that Silicon Graphics and Intel ABI will also be included for V8.

The Analyst Application provides more extensive facilities for **data management** and data manipulation. You can combine data sets by columns and rows, split and stack columns, and use the redesigned Compute Window to create new variables. New Recode dialogs make it easy to recode values and ranges of an existing variable. A Stack Columns task creates a new data set by stacking specified columns into a single column while preserving the structure of other columns. Menus offer quick access to commonly used transformations such as log and square root, and you can easily produce random variates from a wide range of distributions.