



## S-PLUS® 8 ENTERPRISE DEVELOPER: COMPREHENSIVE FEATURE LIST

### S PROGRAMMING LANGUAGE

The award-winning S programming language is at the core of S-PLUS. The only language created specifically for exploratory data analysis and statistical modeling, the S programming language allows you to create statistical applications up to five times faster than with other languages.

- Object-oriented, interpreted 4GL language
- Interactive exploration and fast prototyping
- Rich data structures: vector, matrix, array, data frame, list and many more
- User-defined functions, objects, classes, methods and libraries
- Library of over 4000 functions for data manipulation, graphics, statistical modeling, and integration
- CSAN library of available packages
- Dialogs for data preparation, charting and statistical modeling
- Interactive command-line with history recall
- Manage objects with Object Explorer<sup>1</sup>
- Script file editor<sup>1</sup>
- Multiple data and graphics windows
- Cut-and-paste to Word, PowerPoint and Excel<sup>1</sup>
- Integrated Excel spreadsheets<sup>1</sup>
- PowerPoint Wizard: quickly create slides from charts<sup>1</sup>
- Create custom toolbars, menus and dialogs<sup>1</sup>
- On-line help and manuals
- Eclipse based development environment

### S-PLUS WORKBENCH DEVELOPMENT ENVIRONMENT

Rapidly create reliable statistical applications with this integrated development environment for S programmers.

- Based on industry-standard Eclipse framework
- Check-in and check-out files with source code control system integration
- Intelligent editor for S programs with line numbering, automatic indentation, and syntax highlighting
- Project, file and task management
- Automatic syntax error detection
- Code outline browser
- Command-line console with history recall
- Object and search path views
- Analytic step-by-step debugger
- Analytic profiling
- Package system for improved porting and deployment

### GRAPHICAL USER INTERFACE

A convenient window-based GUI puts common tasks at your fingertips with easy-to-use menus and dialogs

- File import and export dialogs
- Database import and export dialogs<sup>1</sup>

### SCALABLE PIPELINE ARCHITECTURE

Scale statistical applications to gigabytes of data without the need for additional RAM or 64-bit architectures with this library of data types and functions for programming with large data sets.

- Data types for out-of-memory vectors, data frames, and time series
- Use familiar S functions, operators and programming style
- Scalable algorithms for data manipulation, charting and modeling
- High-performance data preparation tools: aggregate, merge, sort, partition, filter and more
- Data manipulation using built-in SQL processor
- Hexagonal binning plots to explore structure of large data sets
- Scalable model estimation: univariate statistics, linear regression, analysis of variance, logistic regression, poisson regression, quasi-likelihood, K-means clustering, principal components
- Scalable model scoring for more than 20 model types

### GRAPHICAL FUNCTIONS

Explore data and create custom charts with this library of graphical functions in the S language

- Scatterplots, histograms, pie charts, box plots, bar charts, dot charts, time series charts, 3-D wireframe charts, image plots and many more.

## S-PLUS® 8 ENTERPRISE DEVELOPER: COMPREHENSIVE FEATURE LIST

- Brush and spin dynamic visualization
- Programmatic control over colors, lines, axes, annotations and layout
- Unique Trellis™ graphics – create multiple charts conditioned by levels of one or more variables
- Create interactive, embedded web-based charts with S-PLUS Graphlets™
- Element-Specific Graph arguments for plots and command-line graphics

### INTEGRATION

S-PLUS is an open system, designed to integrate with the systems you already have.

### Data and graphics formats

- ASCII: fixed format, comma-separated, and tab-delimited
- Spreadsheets: Excel, Lotus 1-2-3, Quattro Pro
- Application data: SAS 7/8/9, SPSS, Matlab, Minitab, Sigma Plot, Systat, STATA, Gauss, Epi Info and more
- Database files: Paradox, dBase, Access, FoxPro
- Financial data sources: LIM, Bloomberg, FAME
- Native database clients: SQL Server<sup>1</sup>, Oracle, Sybase, IBM DB2
- ODBC interface to compliant databases
- Export graphics as PDF, PostScript, GIF, PNG, JPG, WMF, bitmap, TIFF and more

### APIs and system interfaces

- APIs for C, C++, Java and Fortran
- Language support for pipes, sockets, and files
- DDE, COM and OLE interfaces<sup>1</sup>
- XML import and export
- Reporting in XML, PDF, HTML and RTF

### STATISTICAL & NUMERICAL TECHNIQUES

S-PLUS is the most comprehensive statistical analysis package available, and includes all of the following capabilities:

#### Basic Statistics

- Summary statistics

- Crosstabulations
- Correlation and covariance
- Probabilities, quantiles, densities and random number generation from many distributions
- Durbin-Watson statistic

### Hypothesis Tests and Confidence Intervals

- One-sample and two-sample t-test and Wilcoxon
- Paired t-test
- Correlation: Pearson, Kendall's tau, Spearman's rho
- Goodness-of-Fit: Chi-square, Kolmogorov-Smirnov, Shapiro-Wilk
- Rank tests: Kruskal-Wallis, Friedman
- Proportions: exact Binomial test, Normal approximation
- Contingency tables and tests for independence: Chi-square, Fisher, Mantel-Haenszel, McNemar

### Regression

- Basic linear regression
- Polynomial regression
- Model diagnostics
- Prediction and confidence intervals
- Stepwise selection of models
- Parametric spline models
- Constrained regression
- Logistic regression
- Generalized linear models

### Analysis of Variance

- Univariate and multivariate ANOVA
- Flexible specification of variables, covariables, interactions, nesting, transformations
- Automatic generation of dummy variables
- Choice of contrasts
- Type III sums of squares
- Designed experiments: one-way, two-way, factorial, split-plot, unbalanced, fractional factorial designs, response surface methods, robust designs, taguchi methods and more
- Variance component estimation

## S-PLUS® 8 ENTERPRISE DEVELOPER: COMPREHENSIVE FEATURE LIST

- Multiple comparisons: Fisher, Tukey, Dunnett, Sidak, Bonferroni, Scheffé, simulation-based

### Nonlinear Regression and Maximum Likelihood

- Nonlinear regression
- Nonlinear maximum likelihood
- Quasi-likelihood
- Constrained nonlinear regression

### Nonparametric Regression

- Generalized additive models (GAMs)
- Smoothers: loess, super, kernel, spline
- Projection Pursuit, ACE, and AVAS

### Tree Models

- Classification trees
- Regression trees
- Pruning, shrinking, and splitting
- Scoring

### Correlated Data Analysis

- Longitudinal data and repeated measures analysis
- Linear (LME), nonlinear (NLME), and generalized mixed effects (GLMM) models
- Generalized Estimating Equations (GEE)
- Biexponential, first-order compartment, four-parameter logistic models
- User-defined correlation structures

### Resampling

- Bootstrap
- Jackknife

### Multivariate Analysis

- Canonical correlation
- Discriminant analysis
- Factor analysis
- Multidimensional scaling
- Principal components
- Biplots

### Cluster Analysis

- K-means
- Hierarchical clustering
- Monothetic clustering
- Model-based clustering
- Crisp and fuzzy clustering
- Divisive and agglomerative methods

### Quality Control

- Shewhart chart
- Cusum chart
- Charts based on  $\bar{x}$ , s, np, p, c, u

### Power and Sample Size

- Normal mean
- Binomial proportion

### Survival Analysis

- Kaplan-Meier curves
- Cox proportional hazards models with mixed effects
- Left, right, and interval censoring
- Time-dependent covariates and strata
- Multiple event models
- Competing risk models
- Frailty models
- Parametric survival
- Expected survival
- Person years analysis
- Aalen's Additive Regression Model

### Time Series Analysis

- Autocovariance, autocorrelation and partial autocorrelation
- Smoothed periodograms
- Box-Jenkins ARIMA models
- Classical and robust AR
- Long-memory models
- Seasonal decompositions
- Fourier transformations
- Classical and robust smoothers and filters

### Robust Statistics

- Robust estimation and inferences



## S-PLUS® 8 ENTERPRISE DEVELOPER: COMPREHENSIVE FEATURE LIST

- Robust MM regression
- Robust GLM, ANOVA, covariance, principal components, and discriminant analysis
- Least trimmed squares regression
- Minimum absolute residual regression
- Visually compare robust and traditional methods

### Missing Data

- Multiple imputation
- Gaussian, logistic, and conditional Gaussian models

### Date, Time, and Calendar Data

- Univariate and multivariate time series
- Aggregation, alignment, merging, and interpolation
- Times and dates from milliseconds to millennia
- Time zones with international daylight savings rules
- Holidays and financial market closures
- Custom time and date formats
- Relative time, time sequence, and event objects
- Powerful time-series charting

### Mathematical Computations

- Vector and matrix algebra
- Matrix decompositions
- Systems of linear equations
- Locate roots
- Nonlinear optimization
- Constrained optimization

- Ordinary differential equations
- Numerical integration

### ADDITIONAL LIBRARIES

Libraries from Insightful Research and the S-PLUS user community offer additional capabilities

- MASS: Modern and Applied Statistics libraries (Venables, Ripley) included
- Hmisc and Design libraries for biostatistical and epidemiologic modeling (Harrell) included
- Insightful Research libraries available for download

### ADD-ON MODULES

Optional modules add additional capabilities to S-PLUS:

- S+ArrayAnalyzer: microarray analysis<sup>1</sup>
- S+EnvironmentalStats: environmental statistics<sup>1</sup>
- S+FinMetrics: financial econometrics
- S+NuoOPT: large-scale constrained optimization
- S+SeqTrial: Clinical trial design and analysis<sup>1</sup>
- S+SpatialStats: analysis of spatial data
- S+Wavelets: wavelet and signal series analysis

### SUPPORTED PLATFORMS

- Windows 2000, Windows XP, Windows Vista
- Sun Solaris (SPARC)
- Red Hat and SUSE Linux (Intel)