

Comprehensive Statistical Software

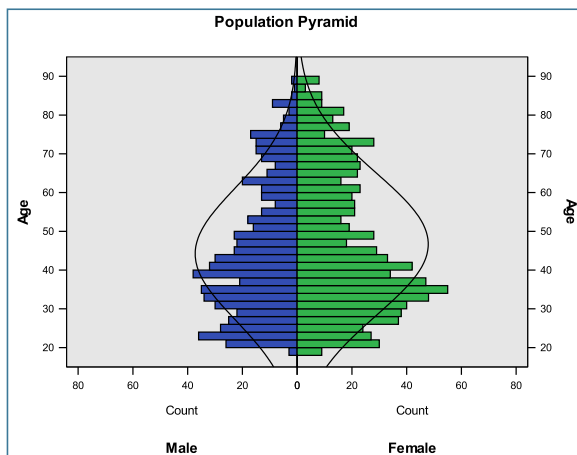
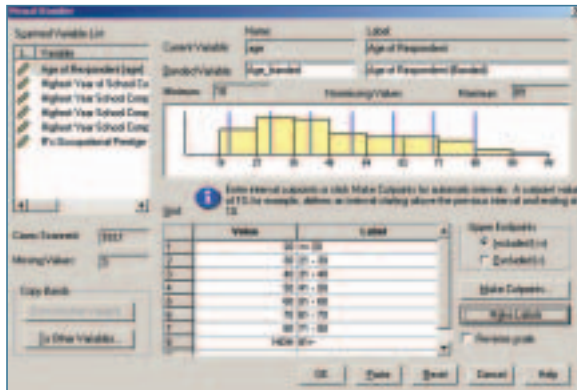
Solve business and research problems using SPSS for Windows, a statistical and data management package for analysts and researchers. Unlike other data analysis packages, SPSS is easier to use, has a correspondingly lower total cost of ownership, and comprehensively addresses the entire analytical process. SPSS Base is an integral part of this process, providing functionality for data access, data management and preparation, data analysis, and reporting. It enables you to work confidently with add-on modules and other products in the SPSS family, which provide capabilities for planning, data collection, and deployment, and additional functionality for areas SPSS Base addresses. The following sections describe SPSS Base highlights and list the statistics, data management procedures, and features that make SPSS Base the leading desktop data management and statistical software. Other products in the SPSS family are summarized on the last page of this document.

Access and analyze massive datasets quickly

SPSS for Windows makes it easy for you to quickly access, manage, and analyze any kind of dataset, including corporate databases or data downloaded from the Web. Make your analysis as fast as possible—let your server do the heavy computation work. Just combine SPSS for Windows with the optional SPSS Server.

Prepare your data for analysis quickly and easily

Before you can analyze your data, you need to prepare it for analysis. Numerous techniques and features built into SPSS Base enable easy data preparation. Unlike other statistical software, there's no need to purchase additional products to help you complete this important pre-analysis step. Following on the next page are summaries of just a few data management highlights of SPSS.



SPSS includes a number of data management techniques that help you prepare data for analysis. The top image shows the Visual Bander, which enables you to set cutpoints in your data. Here, ages have been broken into groups. The lower image shows a population pyramid, one of many chart types available in the presentation graphics system. Population pyramids display the distribution between two groups. This example compares counts of male and female respondents by age.

With SPSS Base, you can easily set up data dictionary information (for example, value labels and variable types) and prepare your data for analysis more quickly using the Define Variable Properties tool. SPSS presents a list of values and counts of those values so you can add the information. Once the data dictionary is set up, you can apply it using the Copy Data Properties tool. The data dictionary acts as a template, so you can apply it to other data files and to other variables within the same file.

SPSS makes it easy for you to identify duplicate cases, so you can eliminate them prior to your analysis. Use the Identify Duplicate Cases tool to set parameters and flag duplicates so that you can keep track of them for the record. Further, SPSS makes it easy to prepare continuous-level data for analysis. The Visual Bander enables you to easily create bands (for example, break income into income “bands” of 10,000 or break ages into groups). A data pass provides a histogram that enables you to specify cutpoints in an intelligent manner. You can then automatically create value labels from the specified cutpoints (for example, “21-30”).

You can also restructure your data files to prepare them for analysis. For example, take a data file that has multiple cases per subject and restructure the data to put all data for each subject into a single record. SPSS gives you the flexibility to complete the reverse action—you can take a data file that has a single case per subject and spread the data across multiple cases.

In SPSS 13.0, you can now more easily work with dates and times using the new Date and Time Wizard. Make calculations with dates and times, create time/date variables from strings containing date variables (such as “12/18/02”), and bring time/date data from a variety of sources into SPSS. You can also parse individual time/date units, such as year, from time/date variables to apply filters. For example, parse start dates to examine employees who started with your organization in 1995.

Analyze data with comprehensive techniques

Go beyond summary statistics and row-and-column math. SPSS gives you a wide range of statistical procedures for basic analysis, including counts, crosstabs, cluster, descriptives, factor analysis, regression, and cluster analysis. Once you complete your analysis, you can write data back to your database with ease—longer variable names are allowed in SPSS, so you can maintain variable names from your databases or spreadsheets. For even more analytical power, use SPSS Base with a variety of add-on modules, such as SPSS Regression Models™ and SPSS Advanced Models™, that focus on data analysis (details are on the last page).

See results more clearly with presentation graphics

Produce professional-looking graphs in SPSS using the presentation graphics system. The graphics system provides you with features that make it easy to create the graph you want and produce readable graph output. The presentation graphics system also gives you control at both creation and edit time, to help ease your workload in a production setting. Create a chart or graph once, and then use your specifications to create hundreds more just like it.

Present your best results with report OLAP

OLAP technology transforms the way you create and share information. Report OLAP in SPSS provides you with a fast, flexible way to create, distribute, and manipulate information for ad hoc decision making. Create tables, graphs, and report cubes that feature unique, award-winning pivoting technology and enable you to discover new insights into your data. Swap rows, columns, and layers of report cubes—or quickly change information and statistics in graphs—for new levels of understanding. You can even convert a table to a graph with just a few mouseclicks.

Deploy results to maximize their benefits

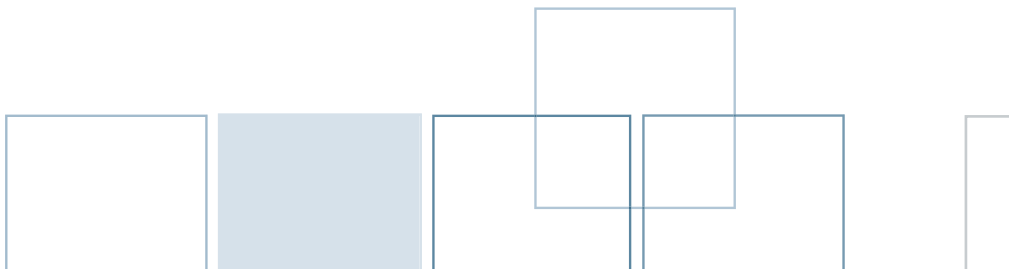
Ensure that people who don't have SPSS installed on their machines can interact with the results. Other software from SPSS Inc. makes it easy. You can publish results from SPSS using SmartViewer® Web Server™ or score results using SmartScore® for deployment. (Visit www.spss.com for more information on these deployment products.)

Features

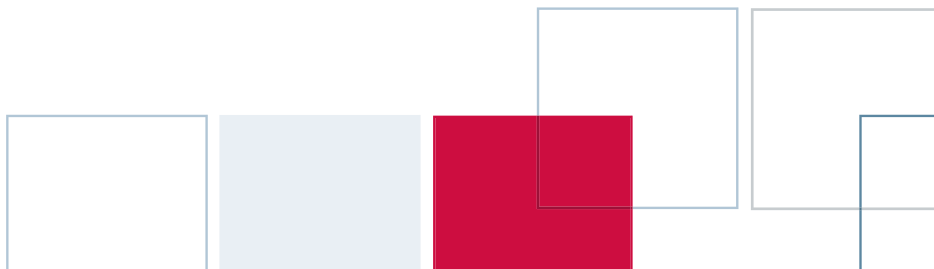
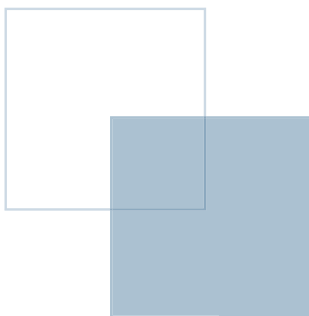
General operations

- Use the scripting facility to:
 - Create, edit, and save scripts
 - Build customized form interfaces
 - Assign scripts to toolbar icons or menus
 - Automatically execute scripts whenever certain events occur
 - Select from a repository of popular scripts from the SPSS Script Library
 - Support SAX Basic 6 to make scripting easier and more reliable
- Use automation to:
 - Integrate SPSS with other desktop applications
 - Build custom applications using Visual Basic®, Power Builder, and C++
 - Integrate SPSS into larger custom applications (such as Microsoft® Word or Excel)
 - Enable more than 400 automation methods
 - Access extensive online documentation
- Use the HOST command to take advantage of the operating system functionality in SPSS. This command enables applications to “escape” to the operating system and execute other programs in sync with the SPSS session.
- Prevent syntax jobs from breaking when you create a common or main project directory that enables you to include transformations for multiple projects
 - Better manage multiple projects, syntax files, and datasets
- Specify interactive syntax rules using the INSERT command
- Apply splitters through the Data Editor to more quickly and easily understand wide and long datasets
- Select the customizable toolbar feature to:
 - Assign procedures, scripts, or other software products
 - Select from standard toolbar icons or create your own
- Work with multidimensional pivot tables/report cubes to:
 - Rearrange columns, rows, and layers by dragging icons for easier ad hoc analyses
 - Toggle between layers by clicking on an icon for easier comparison between subgroups
 - Enable online statistical help for choosing statistical procedures or chart types and interpreting results; realistic application examples are included
- Change text attributes such as fonts, colors, bolding, italics, and others
- Change table attributes such as number formats, line styles, line width, column alignments, background/foreground shading, enable or disable lines, and more
- Selectively display or hide rows, columns, or labels to highlight important findings
- Enable task-oriented help with step-by-step instructions:
 - View case studies that show you how to use selected statistics and interpret results
 - Select the Statistics Coach™, which helps you choose the best statistical procedure or graph
 - Receive examples to help interpret your results using Results Coach™
 - Work through tutorials
 - View chart examples in the Chart Gallery
 - Select “Show Me” buttons, which link to the tutorial for more in-depth help when you need it
 - Use “What’s This?” help, which provides pop-up definitions of statistical terms and rules of thumb
- Use formatting capabilities for output to:
 - Transform a table into a graph for more visually compelling communication
 - Create cleaner-looking tables using box characters in Draft Viewer
 - Show correlation coefficients together with their significance level (as well as n) in correlations using the default output display
- Work the way you want when you control whether, upon activation, a table is opened in place or in its own window
- Stamp date and time into the journal file for easy reference
- Right-click on an SPSS syntax file icon to run a command file without needing to go through production mode
- See bottom-line information fast with OLAP cubes
- Use drop-down lists for easier access to different layers
- Set default output format to text/ASCII
- Set permanent page settings
- Set a column width for all pivot tables and define text wrapping
- See an improved default appearance of pivot table output. For example, set table column width so it automatically adjusts to fit content.
- Choose whether to use scientific notation to display small numbers
- Interact with reports your colleagues publish on the intranet/Internet with the optional SmartViewer Web Server
- Control number of digits of precision in presentations
- Add footnotes and annotations
- Display values or labels
- Reorder categories within a table to display results most effectively
- Group or ungroup multiple categories in rows or columns under a single heading that spans the rows or columns
- Use one of 16 pre-formatted TableLooks™ for quick and consistent formatting of results
- Create and save customized formats as TableLooks for your own personalized style
- Rotate table labels
- Work with the Viewer to organize, view, and move through results, and:
 - Keep a record of your work using the “append” default in journal files
 - Establish more intelligent placement of page breaks in Draft Viewer

Features subject to change based on final product release. □ Symbol indicates a new feature.



- Use outline representation for quick output location
 - Select an icon in the outline and see corresponding results displayed in the content pane
 - Reorder charts, tables, and other objects by dragging icons in the outline
 - Selectively collapse or expand the outline to view or print selected results
 - Contain tables, charts, and objects in a single content pane for easy review and access
 - Right-justify, left-justify, or center output
 - Work seamlessly between applications using OLE 2.0 client and server
 - Create presentation-quality reports
 - Use notes table for complete documentation of each analysis
 - Create and save analysis specifications for repetitive tasks or unattended processing
 - Use the enhanced production mode facility with dialog interface and macros for easier periodic reporting
 - Have full control over table splitting with improved pagination and printing
 - Select the print preview option
 - Enter your own commands, if you wish, via a command line input window
 - Refer to explanations of statistical terms through the on-screen statistical glossary
 - Receive support for the Microsoft MAPI standard for electronically mailing text and graphics files
 - Export output to Word and:
 - Convert pivot tables to Word tables with all formatting saved
 - Convert graphics into static pictures
 - Export output to Microsoft PowerPoint® and:
 - Convert pivot tables to tables in PowerPoint with all formatting saved
 - Convert graphics into static pictures
 - Export output to Excel and:
 - Put tables on the same sheet or on separate sheets within one Excel workbook file
 - Export only the current view or all layers of an SPSS pivot table
 - Place each pivot table layer on the same sheet or on separate sheets within one Excel workbook
 - Receive wheel-mouse support for Output Viewer scroll
 - Switch output languages (for example, switch between Japanese and English)
- Graphic capabilities**
- Categorical charts
 - ☐ 3-D Bar: Simple, cluster, and stacked
 - Bar: Simple, cluster, stacked, drop-shadow, and 3-D
 - Line: Simple, multiple, and drop-line
 - Area: Simple and stacked
 - Pie: Simple, exploding, and 3-D effect
 - High-low: High-low-close, difference area, and range bar
 - Boxplot: Simple and clustered
 - Error bar: Simple and clustered
 - ☐ Error bars: Add to bar, line and area charts, confidence level, S.D., and S.E.
 - Scatterplots
 - Simple, grouped, scatterplot matrix, and 3-D
 - Fit lines: Linear, quadratic or cubic regression, and Lowess smoother; confidence interval control for total or subgroups; and display spikes to line
 - Bin points by color or marker size to prevent overlap
 - Density charts
 - ☐ Population pyramids: Mirrored axis to compare distributions; with or without normal curve
 - ☐ Dot charts: Stacked dots show distribution; symmetric, stacked, and linear
 - Histograms: With or without normal curve; custom binning options
 - Quality control charts
 - Pareto
 - X-Bar
 - Range
 - Sigma
 - Individual chart
 - Moving range chart
 - Diagnostic and exploratory charts
 - Caseplots and time-series plots
 - Probability plots
 - Autocorrelation and partial autocorrelation function plots
 - Cross-correlation function plots
 - Receiver-Operating Characteristics (ROC)
 - Editing options
 - ☐ Automatically reorder categories in differing order (descending or ascending) or by different sort methods (value, label, or summary statistic)
 - Data value labels
 - ☐ Drag to any position on your chart, add connecting lines, and match font color to subgroup
 - Select and edit specific elements directly within a chart: Colors, text, and styles
 - Choose from a wide range of line styles and weights
 - Display gridlines, reference lines, legends, titles, footnotes, and annotations
 - ☐ Y=X reference line
 - Layout options
 - ☐ Paneled charts: Create a table of subcharts, one panel per level or condition, showing multiple rows and columns
 - 3-D effects: Rotate, modify depth, and display backplanes
 - Chart templates
 - Save selected characteristics of a chart and apply them to others automatically. You can apply the following attributes at creation or editing time:
 - Layout
 - Titles
 - Footnotes and annotations
 - Chart element styles
 - Data element styles
 - Axis scale range
 - Axis scale settings
 - Fit and reference lines
 - Scatterplot point binning
 - ☐ Tree-view layout and finer control of template bundles
 - Graph export
 - BMP
 - EMF
 - EPS
 - JPEG
 - PCT
 - PNG
 - TIFF
 - WMF



Analyze

Descriptive statistics

Reports

- OLAP cubes enable you to:
 - Quickly estimate changes in the mean or sum between any two related variables using percent change. For example, easily see how sales increase from quarter to quarter.
 - Create case summaries
 - Create report summaries
 - Generate presentation-quality reports using numerous formatting options
 - Generate case listing and case summary reports with statistics on break groups

Frequencies

- Frequency tables: Frequency counts, percent, valid percent, and cumulative percent
- Long string variables
- Option to order your output by analysis or by table
- More compact output tables by eliminating extra lines of text where they're not needed
- Central tendency: Mean, median, mode, and sum
- Dispersion: Maximum, minimum, range, standard deviation, standard error, and variance
- Distribution: Kurtosis, kurtosis standard error, skewness, and skewness standard error
- Percentile values: Percentiles (based on actual or grouped data), quartiles, and equal groups
- Formats: Condensed or standard, sorted by frequency or values, or generate index of tables
- Charts: Bar, histogram, or pie chart

Descriptives

- Central tendency: Mean and sum
- Dispersion: Maximum, minimum, range, standard deviation, standard error, and variance
- Distribution: Kurtosis and skewness
- Z scores: Compute and save as new variables
- Display order: Ascending or descending order on means and variable name

Explore

- Confidence intervals for mean
- Descriptives: Interquartile range, kurtosis, kurtosis standard error, median, mean, maximum, minimum, range, skewness, skewness standard error, standard deviation, standard error, variance, five percent trimmed mean, and percentages
- M-estimators: Andrew's wave estimator, Hampel's M-estimator, Huber's M-estimator, and Tukey's biweight estimator
- Extreme values and outliers identified
- Grouped frequency tables: Bin center, frequency, percent, valid, and cumulative percent
- Plots: Construct plots with uniform scale or dependent on data values
 - Boxplots: Dependent variables and factor levels together
 - Descriptive: Histograms and stem-and-leaf plots
 - Normality: Normal probability plots and detrended probability plots with Kolmogorov-Smirnov and Shapiro-Wilks statistics
 - Spread versus level plots using Levene's test: Power estimation, transformed, or untransformed
 - Shapiro-Wilk's test of normality in EXAMINE allows for 5,000 cases when weights are not specified

Crosstabs

- Three-way relationships in categorical data with Cochran's and Mantel-Haenszel statistics allow you to go beyond the limits of a two-way crosstab
- Counts: Observed and expected frequencies
- Percentages: Column, row, and total
- Long string variables
- Residuals: Raw, standardized, and adjusted standardized
- Marginals: Observed frequencies and total percentages
- Tests of independence: Pearson and Yates corrected Chi-square, likelihood ratio Chi-square, and Fisher's exact test
- Test of linear association: Mantel-Haenszel Chi-square
- Measure of linear association: Pearson r

- Nominal data measures: Contingency coefficient; Cramer's V; Phi; Goodman and Kruskal's Lambda (asymmetric and symmetric); Tau (column or row dependent); and uncertainty coefficient (asymmetric and symmetric)
- Ordinal data measures: Goodman and Kruskal's Gamma, Kendall's Tau-b and Tau-c, Somers' D (asymmetric and symmetric), and Spearman's Rho
- Nominal by interval measure: Eta
- Measure of agreement: Cohen's Kappa
- Relative risk estimates for case control and cohort studies
- Display tables in ascending or descending order
- Frequency counts written to file
- McNemar's test
- Option to use integer or non-integer weights
- Tables converted to charts

Descriptive ratio statistics

- Help for understanding your data using:
 - Coefficient of dispersion
 - Coefficient of variation
 - Price-related differential (PRD)
 - Average absolute deviance

Compare means

Means

- Better model creation with harmonic and geometric means
- Cells: Count, mean, standard deviation, sum, and variance
- All-ways totals
- Measure of analysis with Eta and Eta²
- Test of linearity with r and r²
- Results displayed in report, crosstabular, or tree format
- Statistics computed for total sample

Features subject to change based on final product release. □ Symbol indicates a new feature.



t tests

- One sample *t* test to compare sample mean to a reference mean of your choice
- Independent sample statistics: Compare sample means of two groups for both pooled and separate-variance estimates with Levene's test for equal variances
- Paired sample statistics: Correlation between pairs, difference between means, and two-tailed probability for test of no difference and for test of zero correlation between pairs
- Statistics: Confidence intervals, counts, degrees of freedom, mean, two-tailed probability, standard deviation, standard errors, and *t* statistic

Oneway ANOVA

- Contrasts: Linear, quadratic, cubic, higher-order, and user-defined
- Post hoc tests: Duncan, LSD, Bonferroni, Student-Newman-Keuls, Scheffe, Tukey's alternate test, and Tukey's HSD
- Twelve post hoc tests: Sidak, Gabriel, Hochberg, Tamhane, Dunnett T3, Games-Howell, Dunnett C, Dunnett (2 sided), Dunnett (control), Ryan-Einot-Gabriel-Welsch F (REGW), REGW Range, and Waller-Duncan
- ANOVA statistics: Between- and within-groups sums of squares, degrees of freedom, mean squares, F ratio, and probability of F
- Fixed-effects measures: Standard deviation, standard error, and 95 percent confidence intervals
- Random effects measures: Estimate of variance components, standard error, and 95 percent confidence intervals
- Group descriptive statistics: Maximum, mean, minimum, number of cases, standard deviation, standard error, and 95 percent confidence interval
- Homogeneity of variance test: Levene's test
- Read and write matrix materials
- Equality of means: Reach accurate results when variances and sample sizes vary across different groups
 - Brown-Forsythe test
 - Welch test

ANOVA models—simple factorial

- Create custom models without limits on maximum order of interaction
- Work faster because you don't have to specify ranges of factor levels
- Choose the right model using four types of sum of squares
- Increase certainty with better data handling in empty cells
- Perform lack-of-fit tests to select your best model
- Choose from one of two designs: Balanced or unbalanced
- Use analysis of covariance with up to 10 covariates methods: Classic experimental, hierarchical, and regression
- Enter covariates control: Before, with, or after main effects
- Set interaction to: None, 2-, 3-, 4-, or 5-way
- Select from the following statistics: ANOVA, means and counts table, multiple classification analysis, unstandardized regression coefficients, and *n*-way cell means
- Choose up to 10 independent variables
- Reach predicted values and deviations from the mean in MCA table

Correlate

Bivariate

- Pearson *r*, Kendall's Tau-*b*, and Spearman
- One- and two-tailed probabilities
- Means, number of non-missing cases, and standard deviations
- Cross-product deviations and covariances
- Coefficients displayed in matrix or serial format

Partial

- One- and two-tailed probabilities
- Mean, number of non-missing cases, and standard deviation
- Zero-order correlations
- Up to 100 control variables
- Up to five order values
- Correlations displayed in matrix or serial string format, lower triangular, or rectangular correlation matrix

Distances

- Compute proximities between cases or variables
- Dissimilarity measures
 - Interval measure: Euclidean and squared Euclidean distance, Chebychev distance metric, city-block or Manhattan distance, distance in an absolute Minkowski power metric, and customized
 - Counts measures: Chi-square and Phi-square
 - Binary measures: Euclidean and squared Euclidean distance; size, pattern, and shape difference; variance dissimilarity measure; and Lance and Williams nonmetric
- Similarity measures
 - Interval measures: Pearson correlation and cosine
 - Binary measures: Russell and Rao; simple matching; Jaccard; dice (or Czekanowski or Sorenson); Rodgers and Tanimoto; Sokal and Sneath 1 through 5; Kulczynski 1 and 2; Hamann, Goodman and Krusal Lambda; Anderberg's D; Yule's Y coefficient of colligation; Yule's Q; Ochiai; dispersion similarity measure; and fourfold point correlation
- Standardize data values: Z scores, range of -1 to 1, range of 0 to 1, maximum magnitude of 1, mean of 1, and standard deviation of 1
- Transform measures: Absolute values, dissimilarities into similarities, similarities into dissimilarities, and rescale proximity values to a range of 0 to 1
- Identification variable specification
- Printed matrix of proximities between items
- Improved performance: Proximities are five to 50 times faster than versions earlier than SPSS 11.0
- Improved scalability for proximities between variable matrices

Regression—linear regression

- Methods: Backward elimination, forced entry, forced removal, forward entry, forward stepwise selection, and R^2 change/test of significance
- Equation statistics: Akaike information criterion (AIC), Ameniya's prediction criterion, ANOVA tables (F, mean square, probability of F, regression, and residual sum of squares), change in R^2 , F at step, Mallows's C_p , multiple R, probability of F, R^2 , adjusted R^2 , Schwarz Bayesian criterion (SBC), standard error of estimate, sweep matrix, and variance-covariance matrix
- Descriptive statistics: Correlation matrix, covariance matrix, cross-product deviations from the mean, means, number of cases used to compute correlation coefficients, one-tailed probabilities of correlation coefficients, standard deviations, and variances
- Independent variable statistics: Regression coefficients (B, standard errors of coefficients, standardized regression coefficients, approximate standard error of standardized regression coefficients, and t), tolerances, zero-order, part, and partial correlations, and 95 percent confidence interval for unstandardized regression coefficient
- Variables not in equation: Beta or minimum tolerance
- Durbin-Watson
- Colinearity diagnostics: Condition indexes, eigenvalues, variance inflation factors, variance proportions, and tolerances
- Plots: Casewise, histogram, normal probability, detrended normal, partial, outlier, and scatterplots
- Create and save variables:
 - Prediction intervals: Mean and individual
 - Predicted values: Unstandardized, standardized, adjusted, and standard error of mean

- Distances: Cook's distances, Mahalanobis' distance, and leverage values
- Residuals: Unstandardized, standardized, Studentized, deleted, and Studentized deleted
- Influence statistics: $dfbetas$, standardized $dfbetas$, $dffits$, standardized $dffits$, and covariance ratios
- Option controls: F-to-enter, F-to-remove, probability of F-to-enter, probability of F-to-remove, suppress the constant, regression weights for weighted least-squares model, confidence intervals, maximum number of steps, replace missing values with variable mean, and tolerance
- Regression coefficients displayed in user-defined order
- System files can contain parameter estimates and their covariance and correlation matrices through the Outfile command
- Solutions can be applied to new cases or used in further analysis
- Decision making can be further improved throughout your organization when you export your models via XML

Curve estimation

- Eleven types of curves are available for specification
- Regression summary displays: Curve type, R^2 coefficient, degrees of freedom, overall F test and significance level, and regression coefficients
- Trend-regression models available: Linear, logarithmic, inverse, quadratic, cubic, compound, power, S, growth, exponential, and logistic

Nonparametric tests

- Chi-square: Specify expected range (from data or user-specified) and frequencies (all categories equal or user-specified)
- Binomial: Define dichotomy (from data or cut point) and specify test proportion
- Runs: Specify cut points (median, mode, mean, or specified)
- One sample: Kolmogorov-Smirnov, uniform, normal, and Poisson
- Two independent samples: Mann-Whitney U, Kolmogorov-Smirnov Z, Moses extreme, and Wald-Wolfowitz runs
- k-independent samples: Kruskal-Wallis H and median
- 2-related samples: Wilcoxon, sign, and McNemar
- k-related samples: Friedman, Kendall's W, and Cochran's Q
- Descriptives: Maximum, mean, minimum, number of cases, and standard deviation

Multiple response

- Crosstabulation tables: Cell counts, cell percentages based on cases or responses, column and row, and two-way table percentages
- Frequency tables: Counts, percentage of cases, or responses
- Both multiple-dichotomy and multiple-response groups can be handled

Data reduction

Factor

- Number of cases and variable labels for analysis can be displayed
- Input from correlation matrix, factor, loading matrix, covariance matrix, or raw data case file
- Output of correlation matrix or factor matrix

Features subject to change based on final product release. □ Symbol indicates a new feature.



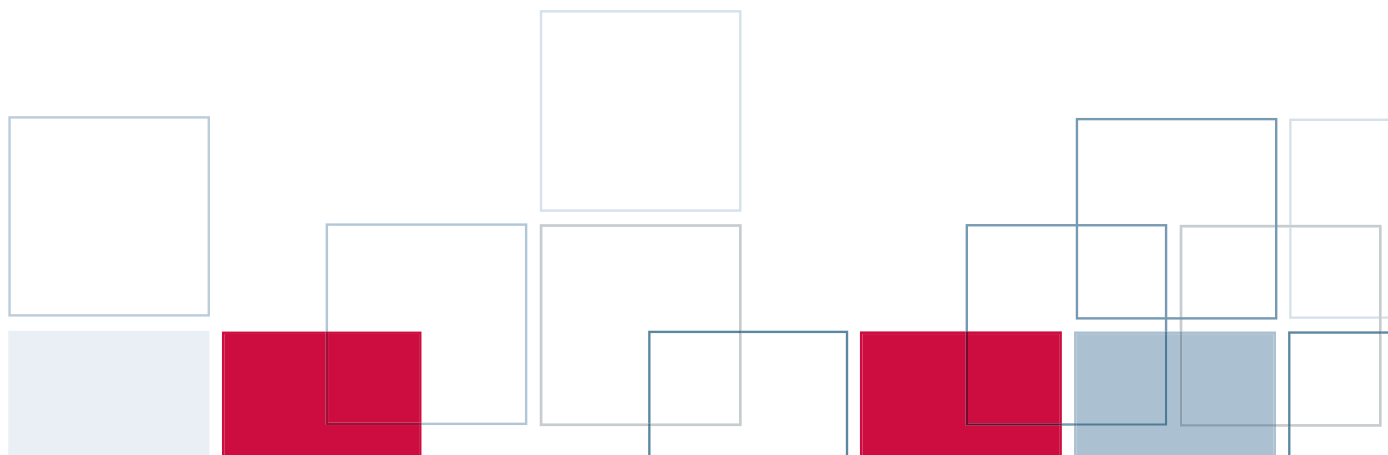
- Seven extraction methods available for when analysis is performed on correlation matrices or raw data files: Principal component, principal axis, Alpha factoring, image factoring, maximum likelihood, unweighted least squares, and generalized least squares
- Rotation methods: Varimax, equamax, quartimax, promax, and oblimin
- Display: Initial and final communalities, eigenvalues, percent variance, unrotated factor loadings, rotated factor pattern matrix, factor transformation matrix, factor structure, and correlation matrix (oblique rotations only)
- Covariance matrices can be analyzed using three extraction methods: Principal component, principal axis, and image
- Factor scores: Regression, Bartlett, and Anderson-Rubin
- Factor scores saved as active variables
- Statistics available: Univariate correlation matrix, determinant and inverse of correlation matrix, anti-image correlation and covariance matrices, Kaiser-Meyer-Olkin measure of sampling adequacy, Bartlett's test of sphericity, factor pattern matrix, revised communalities, eigenvalues and percent variance by eigenvalue, reproduced and residual correlations, and factor score coefficient matrix
- Plots: Scree plot and plot of variables in factor space
- Matrix input and output
- Post-rotational calculated through sum-of-squares loadings
- Solutions applied to new cases or to use in further analysis with the select subcommand
- Factor score coefficient matrix exported to score new data (syntax only)

Classify

TwoStep cluster analysis

- Group observations into clusters based on a nearness criterion. This procedure uses a hierarchical agglomerative clustering procedure in which individual cases are successively combined to form clusters whose centers are far apart. This algorithm is designed to cluster large numbers of cases. It passes the data once to find cluster centers and again to assign cluster memberships. Cluster observations by building a data structure called the CF Tree, which contains the cluster centers. The CF Tree is grown during the first stage of clustering and values are added to its leaves if they are close to the cluster center of a particular leaf.
 - Categorical-level and continuous-level data can be used
 - Distance measures: Euclidean distance and the likelihood distance
 - Criteria command tunes the algorithm so that:
 - The initial threshold can be specified to grow a CF Tree
 - The maximum number of child nodes a leaf node may have can be set
 - The maximum number of levels a CF Tree may have can be set
 - HANDLENOISE subcommand so you can treat outliers in a special manner during clustering. The default value of noise percent is zero, equivalent to no noise handling. The value can range between zero and 100.
 - INFILE subcommand allows the algorithm to update a cluster model in which a CF Tree is saved as an XML file using the OUTFILE subcommand

- MEMALLOCATE subcommand specifies the maximum amount of memory in megabytes (MB) that the cluster algorithm should use
- Missing data: Exclude both user-missing and system-missing values, or let user-missing values be treated as valid
- Option to standardize continuous-level variables or leave them at the original scale
- Ability to specify the number of clusters, specify the maximum number of clusters, or let the number of clusters automatically be chosen
 - Algorithms available for determining the number of clusters: Bayesian Information criterion (BIC) or Akaike Information criterion (AIC)
- Output written to a specified filename as XML
- Final model output saved or use an option that updates the model later with more data
- Plots:
 - Bar chart of frequencies for each cluster
 - Pie chart showing observation percentages and counts within each cluster
 - Importance of each variable within each cluster: The output is sorted by the importance rank of each variable
- Plot options:
 - Comparisons (one plot per cluster or one plot per variable)
 - Measure of variable importance (parametric or non-parametric)
 - Ability to specify alpha level when considering importance



- Print options:
 - Information criterion (AIC or BIC) for different numbers of clusters
 - Two tables describing the variables in each cluster. In one table, means and standard deviations are reported for continuous variables. The other table reports frequencies of categorical variables. All values are separated by cluster.
 - List of clusters and number of observations in each cluster
- Cluster number saved for each case to the working data file

Cluster

- Use one of six linkage methods to determine clusters: Single linkage (nearest neighbor), average linkage between groups, centroid (average linkage within groups), complete linkage (farthest neighbor), median, and Ward
- Provide the same set of similarity and dissimilarity measures as in proximity
- Save cluster memberships as new variables
- Save distance matrices for use in other procedures
- Display: Agglomeration schedules, cluster membership, and distance matrices
- Use proximities between variable matrices for improved scalability
- Choose from the following plots: Horizontal and vertical icicle plots and dendrogram plots of cluster solutions
- Specify case identifiers for tables and plots
- Have the ability to accept matrix input and produce matrix output

Quick cluster

- Squared Euclidean distance
- Centers selected by widely spaced cases, first K cases, or direct specification
- Cluster membership saved as a variable
- Two methods provided for updating cluster centers
- K-means clustering algorithms

Discriminant

- Variable selection methods: Direct entry, Wilks' Lambda minimization, Mahalanobis' distance, smallest F ratio, minimization of sum of unexplained variation for all pairs, and largest increase in Rao's V
- Statistics:
 - Summary: Eigenvalues, percent and cumulative percent of variance, canonical correlations, Wilks' Lambda, and Chi-square tests
 - At each step: Wilks' Lambda, equivalent F, degrees of freedom, and significance of F for each step; F-to-remove; tolerance; minimum tolerance; F-to-enter; and value of statistic for each variable not in equation
 - Final: Standardized canonical discriminant function coefficients, structure matrix of discriminant functions, and functions evaluated at group means
 - Optional: Means, standard deviations, univariate F ratios, pooled within-groups covariance and correlation matrices, matrix of pairwise F ratios, Box's M test, group and total covariance matrices, unstandardized canonical discriminant functions, classification results table, and classification function coefficients
- Rotation of coefficient (pattern) and structure matrices
- Output displayed step by step and/or in summary form
- In classification stage: Prior probabilities, equal, proportion of cases, or user-specified
- All groups, cases, territorial maps, and separate groups plotted
- Casewise results saved to system file for further analysis
- Matrix files read/written, including additional statistics: Counts, means, standard deviations, and Pearson correlation coefficients
- Solutions applied to new cases or for use in further analysis
- Jackknife estimates provided for misclassified error rate
- Decision making further improved by exporting your models throughout your organization via XML

Scaling

- Reduce your data and improve measurement with reliability
- Find the hidden structure in your similarity data using ALSCAL multidimensional scaling

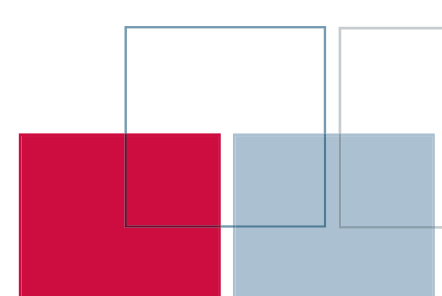
Matrix operations

- Write your own statistical routines in the compact language of matrix algebra

Data management

- Prepare continuous-level data for analysis with the Visual Bander, which allows you to easily create bands (for example, breaking income into income "bands" of \$10,000 or breaking ages into demographic groups)
 - Specify cutpoints in an intelligent manner using a histogram created through a data pass
 - Automatically create value labels based on your cutpoints
 - Copy bands to other variables
- Create your own custom programs with the Output Management System (OMS). Turn output from SPSS procedures into data (SPSS data files, XML, or HTML) and create your own programs for:
 - Bootstrapping
 - Jackknifing and Leaving One Out methods
 - Monte Carlo simulations
- ▣ Create custom programs in SPSS, even if you have little or no experience with SPSS syntax, using the OMS Control Panel
- Easily clean your data when you identify duplicate records through the user interface with the Identify Duplicate Cases tool
- Make sense and keep track of your data files by adding notes to them with the Data File Comments command
- Prevent the accidental destruction of data by making the dataset read-only
- Easily set up all of your value labels to prepare your data for analysis using the Define Variable Properties tool
 - Set up data dictionary information, including value labels and variable types
 - Intelligently add labels because a data pass made first enables SPSS to present a list of values and counts of those values
 - Save time by being able to enter data and value labels directly onto the grid rather than having to use nested dialogs

Features subject to change based on final product release. Symbol indicates a new feature.



- Save work by easily copying dictionary information from one variable to another and from one dataset to another using the Copy Data Properties tool
 - Copy dictionary information (such as variable and value labels) between variables and datasets using the template facility
 - Receive a ready means of cloning dictionaries
- Analyze more data, more efficiently—file size considerations are practically eliminated (especially when used in conjunction with the optional SPSS Server 13.0)
- Assign like variable attributes to multiple variables simultaneously
- Easily select rows and columns to paste information elsewhere
- Easily reorder your variables
- Save time by sorting data directly in the Data Editor
- Avoid reformatting column widths for each new session
- Gain speed by creating customized keyboard options
- Restructure data files that have multiple cases per subject and restructure data to put all data for each subject into a single record (restructure data files from a univariate form to multivariate form)
- Restructure data files that have a single case per subject and spread data across multiple cases (restructure data files from a multivariate form to a univariate form)
- Keep variables when saving a data file with an easy-to-use graphical user interface
- Identify and select variables using your own organization scheme as you sort variables according to variable labels in a list box
- Display variable labels in a dialog; use up to 256 characters
- Display variable labels as a tool tip in the Data Editor
- Assign aliases to fields with the Database Administrator Assistant
- Save SQL queries for later use
- Create prompted queries
- Select data more easily using the “where” clause
- Set any character or combination of characters as the delimiter between fields in an ASCII text file
- Use included ODBC software drivers
- Write SPSS data files from within other applications, such as Excel, using the SPSS ODBC driver
- Use virtually unlimited numbers of variables and cases
- Specify and work with subsets of variables
- Enter, edit, and browse data in the Data Editor’s spreadsheet format
- ▣ Easily work with dates and times using the Date and Time Wizard
 - Create a date/time variable from a string containing a date/time variable
 - Create a date/time variable from variables that include individual date units, such as month or year
 - Parse individual date/time units from date/time variables
 - Calculate with dates and times
- Display values or value labels in Data Editor cells
- With a right mouseclick, receive direct access to variable information within dialog boxes
- Rename variables
- Sort cases
- Choose from several data formats: Numeric, comma, dot, scientific notation, date, dollar, custom currency, and string
- Set an option to show currency as comma- or decimal-delimited
- Choose system missing and up to three user-defined missing values per variable
- Create value labels up to 60 characters
- Create variable labels up to 256 characters
- Insert and delete variables and cases
- Reorder variables
- Search for values of a selected variable
- Transpose working files
- Aggregate data using an extensive set of summary functions
 - ▣ Save aggregated values directly to your active file
- Split files to apply analyses and operations to subgroups
- Select cases either permanently or temporarily
- Process first n cases
- Select random samples of cases for analysis
- Select subsets of cases for analysis
- Weigh cases by values of a selected variable
- Specify random number seeds
- Rank data
- Use neighboring observations for smoothing, averaging, and differencing fast Fourier transformations and their inverse
- More accurately describe your data using longer variable names (up to 64 bytes)
 - Work more easily with data from databases and spreadsheets that include longer variable names than allowed in versions earlier than SPSS 12.0
- ▣ Ensure data containing longer text strings (up to 32,767 bytes) is not truncated or lost when working with open-ended question responses, data from other software that allows long text strings, or other types of long text strings

File management

- Truly minimize data handling with conversion-free/copy-free data access in SQL databases. Save time by not needing to convert data into SPSS format (especially when used in conjunction with the optional SPSS Server 13.0)
- Import data (including compound documents) from current versions of Excel without needing the Database Wizard
 - Read columns that contain mixed data types without any loss of data
 - Automatically read columns with mixed data types as string variables and read all values as valid string variables

- Work more efficiently as you run multiple sessions on one desktop. For example, on jobs that take a long time, you can use SPSS in another session as long as the licenses are available
- Easily read in and define ASCII data using a Text Wizard similar to the one provided in Excel
 - Use text qualifiers to make reading in data even easier
- Increase the accuracy and repeatability of your syntax files with search and replace enhancements
- Read database tables using the Database Wizard
 - Drag-and-drop join support
- Export tables and text as ASCII output
- Save tables as HTML and charts as JPG formats to post SPSS results on the Internet or your intranet
- Directly access the SPSS home page on the Internet
- Translate files to and from Excel, Lotus® 1-2-3®, and dBASE®
- Read raw data from fixed, free-field, or tab-delimited ASCII files
- Write data to fixed-format or tab-delimited ASCII files
- Read complex file structures: Hierarchical files, mixed record types, repeating data, and non-standard file structures
- Read and write SPSS/PC+™ system files
- Merge files
- Display and apply data definitions from an SPSS data file to a working file
- Update master files using transaction files

- Read and write data matrices
- Save many intermediate results for further analysis
- Read recent versions of SAS® files
- Export data files to SAS
- Export data files to current versions of Excel

Transformations

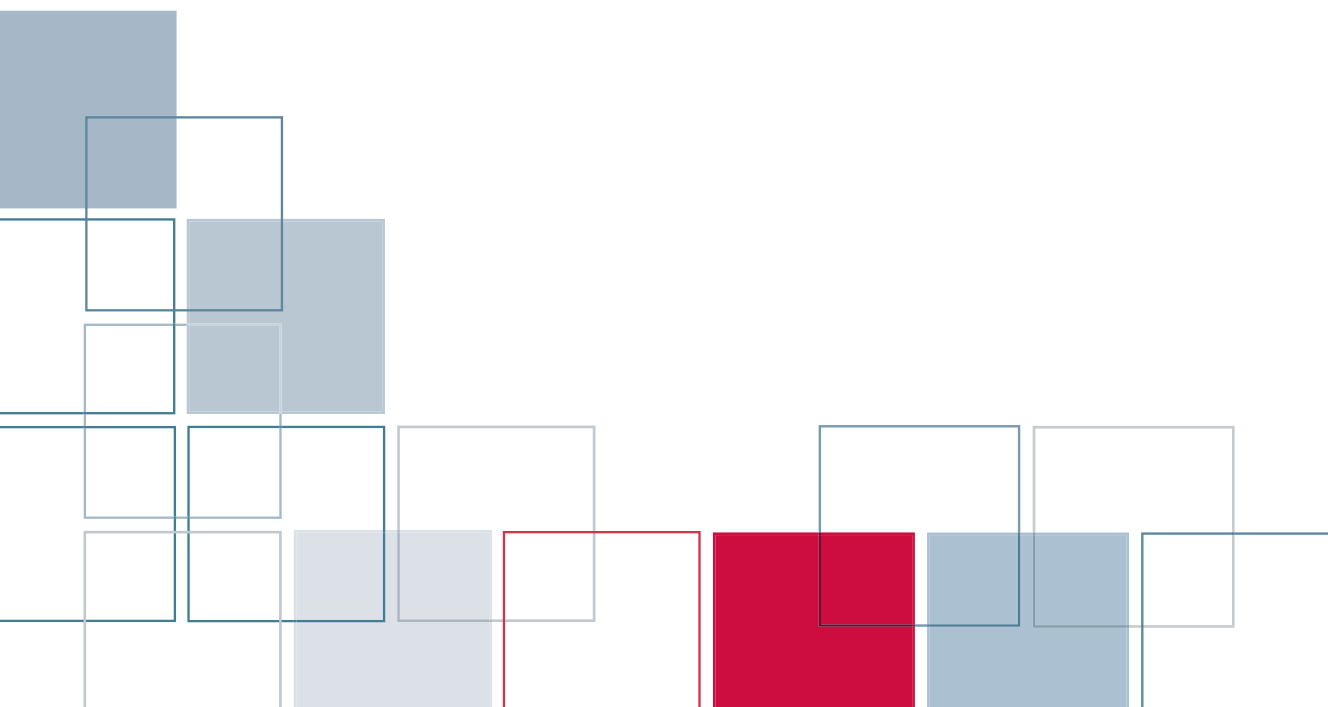
- Compute new variables using arithmetic, cross-case, date and time, logical, missing-value, random-number, and statistical or string functions
- Count occurrences of values across variables
- Recode string or numeric values
- Automatically convert string variables to numeric variables using the autorecode command
 - ☐ Use an autorecode template to append existing recode schemes
 - ☐ Recode multiple variables simultaneously
 - ☐ Autorecode blank strings so that they are defined as “user-missing”
- Create conditional transformations using do if, else if, else, and end if structures
- Use programming structures such as do repeat-end repeat, loop-end loop, and vectors
- Make transformations permanent or temporary
- Execute transformations immediately, in batch mode, or on demand
- Use cumulative distribution, inverse cumulative distributions, and random number generator functions: Beta, Cauchy, Chi-square, Exponential, F, Gamma, Laplace, logistic, lognormal, Normal, Pareto, Student t, uniform, and Weibull

- Standard bivariate normal variates with correlation r, Half Normal, inverse Gaussian, Studentized range, and Studentized maximum modulus
- Work with cumulative distribution and random number generator for discrete distribution functions: Bernoulli, Binomial, Geometric, Hypergeometric, Negative Binomial, and Poisson
- Use cumulative distribution for Non-central distribution: Non-central Beta, non-central Chi-square, non-central F, and non-central T
- Use density/probability functions for:
 - Continuous distributions: Beta, standard bivariate normal with correlation r, Cauchy, Chi-square, exponential, F, Gamma, Half Normal random, inverse Gaussian, Laplace, logistic, lognormal, normal, Pareto, Student t, uniform, and Weibull
 - Discrete distributions: Bernoulli, binomial, geometric, hypergeometric, negative binomial, and Poisson
- Use non-central density/probability functions for: Non-central Beta, non-central Chi-square, non-central F distribution, and non-central t distribution
- Select two-tail probabilities: Chi-square and F
- Use auxiliary function: Logarithm of the complete Gamma function

System requirements

- Operating system: Windows® XP, 2000, and Me are the preferred platforms. Windows 98 is also supported.
- Hardware: Intel® Pentium®-compatible processor
- Memory: 128MB RAM minimum
- Minimum free drive space: 220MB
- SVGA monitor

Features subject to change based on final product release. ☐ Symbol indicates a new feature.



SPSS Server products

SPSS Server 13.0

SPSS Server 13.0 enables SPSS users in your organization to work with large data files for better decision making. The client/server version combines SPSS 13.0 for Windows with SPSS Server 13.0 and a wide range of add-on modules to deliver enterprise-strength scalability and enhanced performance.

SPSS family

Not only is SPSS powerful, but it is also a flexible system that adapts to your evolving needs. Add more analytical power, as you need it, with optional add-on modules and stand-alone software from the SPSS family. Unless otherwise noted, the products described below are add-on modules that require you to use the corresponding version of SPSS Base to operate.

SPSS Regression Models

Predict behavior or events when your data go beyond the assumptions of linear regression techniques. Perform multinomial or binary logistic regression and nonlinear regression, weighted least squares, two-stage least squares, and probit analysis.

SPSS Advanced Models

SPSS Advanced Models' powerful multivariate techniques include mixed level models, general linear modeling (GLM), variance component estimation, MANOVA, Kaplan-Meier estimation, Cox regression, hiloglinear, loglinear, Survival Analysis, and polytomous universal models (PLUM; for ordinal outcomes).

SPSS Classification Trees™

With this new add-on module, you can create highly visual classification and decision trees directly within SPSS for segmentation, stratification, prediction, data reduction and variable screening, interaction identification, and category merging and discretizing continuous variables. These highly visual trees enable you to present results in an intuitive manner—so you can more clearly explain categorical results to non-technical audiences.

SPSS Tables™

Use the recently enhanced SPSS Tables to present survey, customer satisfaction, polling, and compliance reporting in a tabular format. Features, such as a table builder preview, included inferential statistics, and improved data management capabilities make it easy to clearly communicate your results with clients, decision makers, and other report readers.

SPSS Exact Tests™

SPSS Exact Tests always provides you with correct p-values, regardless of your data structure, even if you have a small number of cases, have subset your data into fine breakdowns, or have variables where 80 percent or more of the responses are in one category.

SPSS Categories™

Unleash the full potential of your categorical data through perceptual maps with optimal scaling and dimension reduction techniques. This add-on module provides you with everything you need to analyze and interpret multivariate data and their relationships more completely. A new procedure greatly enhances SPSS Categories 13.0's functionality for multiple correspondence analysis.

SPSS Trends™

Improve forecasting with complete time-series analyses, including multiple curve-fitting and smoothing models and methods for estimating autoregressive functions.

SPSS Conjoint™

SPSS Conjoint helps market researchers develop successful products. By performing conjoint analysis, you learn what product attributes are important in the consumer's mind and what the most preferred attribute levels are, and can perform pricing studies and brand equity studies.

SPSS Missing Value Analysis™

If values are missing from your data, this procedure may find some relationships between the missing values and other variables. In addition, the missing values procedure can estimate what the value would be if data weren't missing. It can estimate the mean, covariance matrix, and correlation matrix through regression or the expectation maximization (EM) algorithm.

SPSS Maps™

Add powerful mapping analysis procedures that are integrated seamlessly with SPSS. Create high-quality maps and perform demographic analysis directly from SPSS data.

SPSS Complex Samples™

Incorporate complex sample designs into data analysis for more accurate analysis of complex sample data. SPSS Complex Samples, with specialized planning tools and statistics, reduces the risk of reaching incorrect or misleading inferences for stratified, clustered, or multistage sampling.

Amos™

Support your research and theories by extending standard multivariate analysis methods when using this stand-alone software package for structural equation modeling (SEM). Build attitudinal and behavioral models that more realistically reflect complex relationships, because any numeric variable, whether observed or latent, can be used to predict any other numeric variable.

SPSS Text Analysis for Surveys™

SPSS Text Analysis for Surveys is a stand-alone software package that offers a combination of linguistic technologies and manual techniques to categorize responses to open-ended questions. To enhance your quantitative analysis, you can export the results as categories or dichotomies for analysis and graphing in SPSS Base or Excel.

SPSS Data Entry™ and Dimensions™ products

SPSS Inc. offers a variety of stand-alone products that help you enter and capture data for survey research. SPSS Data Entry products provide you options for desktop- or Web-based data entry, useful when networking multiple stations. Dimensions give you the ability to automatically capture data online, by telephone, through handheld devices, or when using paper forms that you scan. All of these products work with SPSS 13.0 for Windows, enabling you to seamlessly analyze your survey data.

To learn more, please call Technologies4Targeting Ltd +44 (0)1733 890790 or visit www.tech4t.co.uk/spss

SPSS is a registered trademark and the other SPSS products named are trademarks of SPSS Inc. All other names are trademarks of their respective owners.
© Copyright 2004 SPSS Inc. S135PC-0804

The SPSS logo is displayed in white, bold, sans-serif capital letters on a solid red rectangular background.