Why use Stata?

Integrated menus and dialogs • Intuitive command syntax • Fully programmable
Fast, accurate, and easy to use • One package—no separate modules
Available for Windows, Mac, and Unix

Resources
Find Stata tutorials, training, technical support, publications—including 12,000 pages of documentation, filled with fully worked examples—and more at stata.com/support. Or access the shared knowledge of over 6,500 Stata users on the Stata forum at stata.com/statalist. Beyond stata.com, the Stata resource links offer even more tutorials, examples, books, and course materials. Visit stata.com/links. StataCorp’s publishing arm, Stata Press, publishes books by Stata users. Browse the list of available titles at stata-press.com. Also check out the academically indexed Stata Journal, available at stata-journal.com.

Stay Connected

stata.com/why-use-stata

Stata, Mata, and NetCourse are registered trademarks of StataCorp LP, College Station, Texas, USA.
Numerics by Stata is a trademark of StataCorp LP.
Hundreds of statistical tools at your fingertips

ANOVA/MANOVA
balanced and unbalanced designs • contrasts, designs, factorial, nested, and mixed designs • marginal means • repeated measures

Basic statistics
summaries • cross-tabulations • correlations • t and t tests • equality-of-variances tests • tests of proportions • confidence intervals • factor variables

Bayesian analysis
thousands of built-in models • univariate, multivariate, linear, and nonlinear models • continuous, binary, ordinal, and count outcomes • 18 priors • add your own models • adaptive Metropolis-Hastings sampling • Gibbs sampling • convergence diagnostics • posterior summaries • hypothesis testing • model comparison

Binary, count, and discrete outcomes
conditional, multinomial, nested, ordered, rank-ordered, and stereotype logistic • logistic, probit, tobit • marginal effects • multinomial probit • Poisson and negative binomial • selection models • zero-inflated, truncated, and censored outcomes

Cluster analysis
hierarchical clustering • k-means and k-median • nonhierarchical clustering • dendrograms • stopping rules • user-extensible analyses

Contrasts, pairwise comparisons, and margins
compare means, intercepts, or slopes • compare with reference category, adjacent category, grand mean, etc. • orthogonal polynomials • multiple-comparison adjustments • graph estimated means and contrasts • interaction plots

Data management
data transformations • match-merge • import/export • ODBC • SQL • XML • by-group processing • append files • sort • row/column transposition • labeling • saved results • Unicode

Documentation
23 manuals • 12,000 pages • seamless navigation • thousands of worked examples • quick starts • methods and formulas • references

Embedded statistical computations
Numerics by Stata

Epidemiology
standardization of rates • case-control • cohort • matched case-control • Mantel-Haenszel • pharmacokinetics • ROC analysis • ICD-10

Exact statistics
exact logistic and Poisson regression • exact case-control statistics • binomial tests • Fisher’s exact test for r x c tables

Functions
statistical • random-number • mathematical • string • date and time

Generalized linear models (GLMs)
ten link functions • user-defined links • save distributions • ML and IRLS estimation • nine variance estimators • seven residuals

GMM and nonlinear regression
generalized method of moments (GMM) • nonlinear regression

Graphics
lines • bars • areas • ranges • contours • confidence intervals • interactions plots • survival plots • publication quality • customize anything • Graph Editor

Installation Qualification
IQ report for regulatory agencies such as the FDA • installation verification accessibility for persons with disabilities

IRT (Item response theory)
binary response models: 1PL, 2PL, 3PL • Ordinal response models: graded response, partial credit, rating scale • normal response model • item characteristic curves • test characteristic curves • information function graphs

Linear models
regression • censored outcomes • endogenous regressors • bootstrap, jackknife, robust, and cluster-robust variance • contrasts • quantile regression • OLS

Longitudinal data/panel data
random effects • robust standard errors • continuous, binary, count, ordered, censored, and survival outcomes • GEE • dynamic panel-data models • instrumental variables • panel unit-root tests

Matrix programming—Mata
interactive sessions • large-scale development projects • optimization • matrix inversions • decompositions • eigenvalues and eigenvectors • LAPACK engine • real and complex numbers • string matrices • interface to Stata datasets and matrices • numerical derivatives • object-oriented programming

Multilevel mixed-effects models
continuous, binary, count, and survival outcomes • two-, three-, and higher-level models • generalized linear models • random intercepts • random slopes • crossed random effects • BLUPs of effects and fitted values • hierarchical models • residu al error structures • support for survey data

Multiple imputation
nine univariate imputation methods • multivariate normal imputation • chained equations • explore pattern of missingness • manage imputed datasets • fit model and pool results • transform parameters • joint tests of parameter estimates • predictions

Multivariate methods
factor analysis • principal components • discriminant analysis • rotation • multidimensional scaling • Procrustean analysis • correspondence analysis • biplots • dendrograms • user-extensible analyses

Nonparametric methods
Wilcoxon-Mann-Whitney signed ranks, and Kruskal-Wallis tests • Spearman and Kendall correlations • Kolmogorov-Smirnov tests • exact binomial CIs • survival data • ROC analysis • smoothing • bootstrapping

Power and sample size
doctoral • sample size • effect size • minimum detectable effect • means • proportions • variances • correlations • ANOVA • case-control studies • cohort studies • contingency tables • survival analysis • balanced or unbalanced designs • results in tables or graphs

Programmable maximum likelihood
user-specified functions • NR, DFF, BFGS, BHHH • OIM, OPG, robust, bootstrap, and jackknife SEs • Wald tests • survey data • numeric or analytic derivatives

Resampling and simulation methods
bootstrap • jackknife • Monte Carlo simulation • permutation tests

SEM (Structural equation modeling)
graphical-path diagram builder • standardized and unstandardized estimates • modification indices • direct and indirect effects • continuous, binary, count, ordinal, and survival outcomes • multilevel models • random slopes and intercepts • factor scores • groups and tests of invariance • goodness of fit • handles MAF data by FIML • correlated data • survey data

Simple maximum likelihood
specify likelihood using simple expressions • no programming required • survey data • standard, robust, bootstrap, and jackknife SEs • matrix estimators

Survey methods
multistage and bootstrap, BRR, jackknife, linearized, and SDR variance estimation • poststratification • DEFF • predicted margins • proportions, ratios, and totals • summary tables • virtually all estimators supported

Survival analysis
Kaplan-Meier and Nelson-Aalen graphs • Cox regression (frailty) • parametric models (frailty, random effects) • competing risks • hazards • time-varying covariates • left- and right-censoring • multilevel models

Tests, predictions, and effects
Wald tests • LR tests • linear and nonlinear combinations • predictions and generalized predictions • marginal means • least-squares means • adjusted means • marginal and partial effects • forest plots • Hausman tests

Time series
ARMA • ARFIMA • ARCH/GARCH • VAR • VECD • multivariate GARCH • ARCH • unobserved-components model • dynamic factors • state-space models • Markov-switching model • forecasts • impulse-response functions • unit-root tests • CUSUMs and smoothers

Treatment effects
inverse-probability weighting (IPW) • doubly robust methods • propensity-score matching • regression adjustment • covariate matching • multilevel treatments • endogenous treatments • average treatment effects (ATEs) • ATET on the treated • potential-outcome means (POMs) • continuous, fractional, binary, count, and survival outcomes • balance and overlap diagnostics

Other statistical methods
Kapla measure of inter-rater agreement • Cronbach’s alpha • stepwise regression • tests of normality

Statistical breadth and depth
You need a statistical package that cannot only handle your specific requirements but also address the expanding range of statistical methods. Research professionals from all disciplines use Stata to manage and analyze their data—and so can you. Whether you are a student or a seasoned research professional, Stata gives you all the tools you need.

Easy to learn...
With Stata’s menus and dialogs, you get the best of both worlds. You can easily point-and-click or drag and drop your way to all of Stata’s statistical, graphical, and data-management features. Use Stata’s intuitive command syntax to quickly execute commands. You can even access the dialog boxes for each command directly from the online help system, which is a great way to explore all of Stata’s capabilities.

Fully programmable...
Stata has complete command-line scripting and programming facilities, including a full matrix programming language. You have access to everything you need to script your analyses or even to create new Stata commands—commands that work just like those shipped with Stata. Shinn’s Do-file Editor allows you to create files of Stata commands so analyses can be reproduced with the click of a button. With Stata’s version control, your scripts will continue to work seamlessly when you upgrade to a new version.

Whether you enter commands directly or use the menus and dialogs, you can create a log of all actions and their results to ensure the reproducibility and integrity of your analyses.

And extensible
Easily search, add, and use any of the thousands of user-written commands with just a few mouse clicks.