### Statistical breadth and depth

You want a statistical package that does everything you need and also addresses 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 wrangling 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 can easily script your analysis or even create new Stata commands that work just like those shipped with Stata. Stata’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 community-contributed commands with just a few mouse clicks.

### Stay connected

Stata, Mata, and NetCourse are registered trademarks of StataCorp LLC, College Station, Texas, USA.

Numerics by Stata is a trademark of StataCorp LLC.
Tools for all of your data science needs

Exact statistics
- exact logistic and Poisson regression
- exact case-control statistics
- binomial tests
- Fisher's exact test for > 2 tables

Extended regression models (ERM)
- endogenous covariates
- sample selection
- nonrandom treatment
- panel data
- account for time-varying covariate alone or in combination
- continuous, interval-censored, binary, and ordinal outcomes

Finite mixture models (FMMs)
- prefix for 17 estimators
- mixtures of a single estimator
- mixtures combining multiple estimators or distributions
- continuous, binary, mixed, censored, interval-censored, truncated, and survival outcomes

Generalized linear models (GLMs)
- prefix for 46 estimation commands
- continuous, multilevel, nested, ordered, rank-ordered, and stereotype logistic
- generalized multistep probit
- zero-inflated and left-truncated count models
- selection models
- marginal effects

Choice models
- discrete choice
- rank-ordered alternatives
- conditional logit
- multinomial probit
- nested logit
- mixed logit
- panel data
- case-specific and alternative-specific predictors
- results—expected probabilities, covariate effects, comparisons across alternatives
- more

Cluster analysis
- hierarchical clustering
- kmeans and kmedian nonhierarchical clustering
- dendrograms
- stopping rules
- user-extensible analyses
- more

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

Data wrangling
- data transformations
- data frames
- match-merge
- import/export data
- GDPC
- SOL
- Unicoda
- by-group processing
- append files
- sort
- row-column operations
- labeling
- save results
- more

DSGE models
- specify model algebraically
- solve model
- estimate parameters
- identification diagnostics
- policy and transition matrices
- IRFs
- dynamic forecasts
- more

Epidemiology
- standardized rates
- case-control
- cohort
- matched case-control
- Mantel-Haenszel
- pharmacokinetics
- ROC analysis
- ICD-10
- more

Spatial autoregressive models
- spatial lags of dependent variable, independent variables, and autoregressive errors
- fixed and random effects in panel data
- endogenous covariates
- analyze spillover effects
- more

Survey methods
- multistage designs
- bootstrap, BRR, jackknife, linearized, and SDR variance estimation
- post-stratification
- calibration
- raking
- DEFF
- predictive margins
- means, proportions, ratios, totals
- summary tables
- almost all estimators supported
- more

Survival analysis
- Kaplan-Meier and Nelson-Aalen estimators
- Cox regression
- frailty
- parametric models (frailty, random effects)
- competing risks
- hazards
- time-varying covariates
- left-, right-, and interval-censoring
- Weibull, exponential, and Gompertz models
- more

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
- forecast models
- Hausman tests
- more

Time series
- ARIMA
- ARFIMA
- ARCH/GARCH
- VAR
- VECM
- multivariate GARCH
- unobserved-components models
- dynamic factors
- state-space models
- Markov-switching models
- business calendars
- tests for structural breaks
- threshold regressions
- forecasts
- impulse-response functions
- unit-root tests
- filters and smoothers
- rolling and recursive estimation
- more

Treatment effects/Causal inference
- inverse probability weighting
- doubly robust methods
- propensity-score matching
- regression adjustment
- covariate matching
- multilevel treatments
- endogenous treatments
- average causal effects
- ATEs on the treated (ITTs)
- potential-outcome means (POMs)
- continuous, binary, count, fractional, and survival outcomes
- panel data
- more

Community-contributed maximum-likelihood
- write likelihood using simple expressions
- script programs
- Mata
- matrix programs
- numeric or analytic derivatives
- automatic support for survey data
- perform linear and nonlinear tests
- or even marginal analyses
- automatic MLE, robust, bootstrap, and jackknife
- more

Other statistical methods
- kappa measure of interrater agreement
- Cronbach's alpha
- stepwise regression
- tests of normality
- more

Seasonal Adjustment Factor for 28-day Moving Average

Deriving Optimal Input Spins

Panel/daytime analysis
- series estimation
- kernel regression
- multiple covariates
- discrete and continuous covariates
- point estimates and CIs
- automatic optimization
- online bandwidth selection
- slices of response surface with CIs and associated plots
- more

Panel or longitudinal data
- random and fixed effects with robust standard errors
- linear mixed models
- endogenous effects, random effects
- fixed effects
- Polya–Poisson
- penalized estimation
- estimation commands
- fixed effects
- mixed effects
- fixed effects
- random effects
- clustered and unclustered data
- clustered observations
- fixed and random effects
- more

Generalized linear models (GLMs)
- prefix for 46 estimation commands
- binomial, Poisson, count, logit, and probit
- negative binomial
- quasi-family
- quasi-likelihood
- gamma
- inverse-gamma
- binomial
- Poisson
- gamma
- inverse-gamma
- logit
- probit
- tobit
- negative binomial
- more

Nonlinear models
- prefix for 17 estimators
- mixtures of a single estimator
- mixtures combining multiple estimators or distributions
- continuous, binary, mixed, censored, interval-censored, truncated, and survival outcomes
- more

Bayesian analysis
- thousands of built-in models
- univariate and multivariate models
- linear and nonlinear models
- multilevel models
- continuous, binary, ordinal, and count outcomes
- bayes: prefix for 46 estimation commands
- continuous univariate and discrete priors
- add your own models
- multiple chains
- convergence diagnostics
- posterior summaries
- hypothesis testing
- model fit
- model comparison
- more

Tools for all of your data science needs

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

Basic statistics
- summaries
- cross-tabulations
- correlations
- t and F tests
- equality-of-variance tests
- tests of proportions
- confidence intervals
- factor variables
- more

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

Multiple imputation
- nine univariate imputation methods
- multivariate normal imputation
- chained equations
- explore pattern of misssingness
- manage imputed datasets
- fit model and pool results
- transform parameterizations
- joint tests of parameter estimates
- predictors
- more

Multivariate effects
- factor analysis
- principal components
- discriminant analysis
- rotation
- multivariate scaling
- Procrustean analysis
- correspondence analysis
- biplots
- dendrograms
- user-extensible analyses
- more

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

Nonparametric regression
- series estimation
- kernel regression
- multiple covariates
- discrete and continuous covariates
- point estimates and CIs
- automatic optimization
- online bandwidth selection
- slices of response surface with CIs and associated plots
- more

Power, precision, and sample size
- power
- sample size
- effect size
- minimum detectable effect
- CIs
- means
- proportions
- variances
- correlations
- ANOVA
- regression
- cluster randomized designs
- case-control
- cohort studies
- contingency tables
- survival analysis
- balanced or unbalanced designs
- results in tables or graphs
- more

Programming features
- add new commands
- scripting
- object-oriented programming
- menu and dialog-box programming
- standardized rates
- case-control
- cohort
- matched case-control
- Mantel-Haenszel
- pharmacokinetics
- ROC analysis
- ICD-10
- more

Spatial autoregressive models
- spatial lags of dependent variable, independent variables, and autoregressive errors
- fixed and random effects in panel data
- endogenous covariates
- analyze spillover effects
- more

Survey methods
- multistage designs
- bootstrap, BRR, jackknife, linearized, and SDR variance estimation
- post-stratification
- calibration
- raking
- DEFF
- predictive margins
- means, proportions, ratios, totals
- summary tables
- almost all estimators supported
- more

Survival analysis
- Kaplan-Meier and Nelson-Aalen estimators
- Cox regression
- frailty
- parametric models (frailty, random effects)
- competing risks
- hazards
- time-varying covariates
- left-, right-, and interval-censoring
- Weibull, exponential, and Gompertz models
- more

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
- forecast models
- Hausman tests
- more

Time series
- ARIMA
- ARFIMA
- ARCH/GARCH
- VAR
- VECM
- multivariate GARCH
- unobserved-components models
- dynamic factors
- state-space models
- Markov-switching models
- business calendars
- tests for structural breaks
- threshold regressions
- forecasts
- impulse-response functions
- unit-root tests
- filters and smoothers
- rolling and recursive estimation
- more

Treatment effects/Causal inference
- inverse probability weighting
- doubly robust methods
- propensity-score matching
- regression adjustment
- covariate matching
- multilevel treatments
- endogenous treatments
- average causal effects
- ATEs on the treated (ITTs)
- potential-outcome means (POMs)
- continuous, binary, count, fractional, and survival outcomes
- panel data
- more

Community-contributed maximum-likelihood
- write likelihood using simple expressions
- script programs
- Mata
- matrix programs
- numeric or analytic derivatives
- automatic support for survey data
- perform linear and nonlinear tests
- or even marginal analyses
- automatic MLE, robust, bootstrap, and jackknife
- more

Other statistical methods
- kappas
- measures of interrater agreement
- Cronbach's alpha
- stepwise regression
- tests of normality
- more