

THE STATA NEWS

January/February/March 2010

Vol 25 No 1

Stata makes a difference

Learn how the National Data Bank for Rheumatic Diseases uses Stata to tackle even the most challenging data-management tasks.

p. 1

In the spotlight

Get an introduction to Stata's new suite of commands for multiple imputation, and see how the Variables Manager introduced in Stata 11 simplifies dataset management.

p. 2

Stata Conference Boston 2010

Mark your calendars and make travel arrangements to attend the annual Stata Conference, being held this year in Boston.

p. 9

Also in this issue

New from Stata Press	4
New from the Stata Bookstore	5
Seminars on Stata.....	7
Public training courses	8
Users Group meetings	9
Upcoming NetCourses	10

The Stata News

Executive Editor: Brian Poi
Production Supervisor: Annette Fett

Stata makes a difference at the National Data Bank for Rheumatic Diseases

Editor's note: This column profiles organizations and individuals who use Stata in innovative ways to solve an array of challenges, from complex data-management tasks, to advanced statistical analysis, to graphical and textual report generation. We hope these profiles help you think creatively and learn how Stata can make you more productive.

Fred Wolfe, MD, leads the National Data Bank for Rheumatic Diseases (NDB).

The NDB collects self-reported information directly from patients using 28-

page questionnaires mailed at six-month intervals, gathering information on use of services, medical costs, financial status, functional ability, quality of life, psychological status, treatments received and their side effects, and long-term outcomes pertaining to illness, work, and death. Patients are typically referred to the NDB by their rheumatologist, and critical medical events are validated by obtaining medical records.

The NDB

Unlike other data banks that collect data primarily from administrative sources like Medicare or insurance companies or from physicians, hospitals, and laboratories, the NDB data sourced directly from patients allows researchers to answer questions that are most germane to patients but that cannot be answered based on other databases' information. These questions include things like treatment efficacy in

the community rather than efficacy in randomized clinical trials, whether patients use the treatments, whether patients report less pain, and how the disease affects patients' daily lives.

NDB staff input the nearly 10,000 variables of data into a Microsoft SQL database. Wolfe and his team then use Stata programs to build and update the dataset, a step that is done nightly and takes about 6 hours to run. Stata programs then check data consistency. "Immortalized" datasets are created upon the completion of a six-month survey phase. The current immortalized dataset contains nearly 600,000 observations on 89,000 patients. Auxiliary

"Stata allowed us to do data management in a flexible, useful, cost-effective way that we couldn't do otherwise." — Fred Wolfe

programs allow database managers to apply value labels and account for missing values and allow users to extract, manipulate, and process variables of interest. All told, the NDB uses over 1,000 programs and do-files, mostly written by Wolfe or

his colleague, Kaleb Michaud, a senior analyst at NDB and assistant professor of medicine at the University of Nebraska Medical Center.

The database is used to publish research about rheumatic diseases in peer-reviewed journals. Roughly 125 papers that rely on this data have been published. Because the data bank is complex and accessed by Stata commands, researchers using the NDB data typically work with a member of Wolfe's staff. Michaud adds, "When research collaborators want to work with the data, we highly recommend that they use Stata for the analysis; serious medical students, residents, and fellows who take on research projects with me all use it."

The NDB also maintains safety registries, long-term observational studies that monitor adverse events among patients receiving new drugs.

Continued on p. 2

Getting started with Stata

Learning that many colleagues were switching to Stata, Wolfe was tempted when he discovered that data management in Stata would free him from the arduous task of writing SAS loops. Finding SAS to be expensive and bloated, and S-Plus and (later) R to be hard to learn and unwieldy for data management, he quickly fell for Stata in 1995. Wolfe writes, “Stata programs were one of the things that made Stata great for us.” Stata’s flexible programming language allowed for all sorts of contingencies in the data and facilitated reporting. Wolfe continues, “In a data bank that was always changing, Stata allowed us to do data management in a flexible, useful, cost-effective way that we couldn’t do otherwise.”

Wolfe also credits the Stata community, including the Statalist email group, the *Stata Journal* and *Stata Technical Bulletin*, and the Statistical Software Components (SSC) archive. He says that postings written by Nick Cox, a long-time member of the Stata community, taught him more about programming than he could have learned anywhere else. Wolfe summarizes the Stata community: “Today, between the manuals, the archives, and Googling Stata issues, Stata is a continuing teacher. I certainly learned what not to do.”

Stata’s role—data management, analysis, and reporting

The NDB relies heavily on Stata’s data-management facilities, including its support for ODBC connectivity, extensive macro manipulation features, and commands like `egen` and `merge`. Statistical techniques such as linear regression, logit modeling, fixed- and random-effects estimation, mixed modeling, and survival analysis are all carried out using Stata. Stata’s comprehensive graphics capabilities are a vital component of the NDB’s data-management and reporting tasks. Wolfe asks, “How could I live without Stata and the Statalist?”

Results

Research produced with NDB has led to important insights and real changes in recommended treatments. Among their many findings, NDB researchers were the first to show that methotrexate reduced mortality in rheumatoid arthritis (RA) patients. They demonstrated or confirmed the association between RA and heart attacks, stroke, skin cancer, and lymphomas, and they showed there is no increase in cancer and cardiovascular risk from biologic therapy of RA. The NDB documented rates of work disability among RA patients and identified predictors thereof. They published the first longitudinal study on joint replacement in RA patients, and they published the classic and definitive papers on the erythrocyte sedimentation rate in rheumatic diseases. Recently, they determined the rate of retinal toxicity of the common arthritis and lupus drug hydroxychloroquine, and their findings are now being turned into recommendations for treatment monitoring.

The NDB also used Stata to develop clinical assessments, including the HAQ-II functional questionnaire and the fibromyalgia diagnostic questionnaire. They showed that patient questionnaires could be used in clinical settings and were important in predicting outcomes.

Their database has also allowed them to learn lessons that are more broadly applicable. For example, the NDB’s extensive questionnaires have led them to understand rates of non-response and what makes a survey question good or bad. As Wolfe describes the process, “The data bank is an epidemiologic textbook on data collection and errors, missing data, biases, causality, and on and on. Stata made it easy to learn these things.”

Conclusion

Stata plays a crucial role at the NDB under the direction of Fred Wolfe. From managing raw survey data to integrity checking, to advanced statistical analyses and report generation, Stata provides the tools the NDB needs to get the job done.

— Brian Poi, Executive Editor
and Senior Economist

In the spotlight: Multiple imputation

Editor’s note: In this column, Stata developers tell you what they find useful and interesting about the new features in Stata 11.

Many real-world datasets are incomplete—some observations may not have data for all the variables you wish to use in your analysis. A popular way to handle these situations is to use Rubin’s multiple imputation (MI) procedure. Like listwise deletion, another common way to handle missing data, MI is applicable to a wide variety of analyses. Unlike listwise deletion, which results in a loss of data, MI preserves all available information in the dataset and can therefore lead to more efficient estimation.

MI is a simulation-based method consisting of three steps:

- 1) *imputation* creates multiple imputed (completed) datasets according to a chosen imputation model
- 2) *complete-data analysis* performs primary analysis of interest on each of the datasets
- 3) *pooling* consolidates results from step 2 into one MI inference using Rubin’s combination rules

In Stata, you can use the `mi` command to perform multiple imputation. `mi impute` performs imputation assuming the data are “missing-at-random” (MAR). Stata’s `mi estimate` combines the second and third steps into one easy-to-use command.

Let’s look at `mi` in action. We have data from a case-control study examining the relationship between smoking and heart attacks. As shown in the first Results window, one of the covariates, body mass index, is missing in 22 of 154 observations. We fill in missing values of the continuous variable `bmi` by using `mi impute regress` to obtain multiply imputed data, and then we analyze those data using logistic regression and pool the results. The MI Control Panel is available to guide you through all the steps.

First, we create 20 imputations of `bmi` using `mi impute regress`:

The image shows three Stata windows. On the left is the 'MI - Multiple-Imputation Control Panel' with the 'Impute' tab selected. In the center is a smaller window titled 'mi impute regress - Impute using linear regression' where 'bmi' is the imputed variable and 'attack smokes age hsggrad female' are the independent variables. On the right is a 'Results' window showing the output of the command `. misstable summarize bmi`, which displays the distribution of `bmi` for observed and imputed data.

Variable	Obs=.	Obs>.	Obs<.	Unique values	Min	Max
bmi	22		132	132	17.22643	38.24214

Another 'Results' window shows the output of `. mi impute regress bmi = attack smokes age hsggrad female, add(20)`, including a table of observations per `m`:

Variable	observations per <i>m</i>			total
	complete	incomplete	imputed	
bmi	132	22	22	154

Next we fit the logistic regression to the multiply imputed data using the `mi estimate:` prefix. Here we show the command syntax instead of using the MI Control Panel:

The screenshot shows the output of the command `. mi estimate, dots: logit attack smokes age bmi hsggrad female`. It includes summary statistics for the imputations and the logistic regression coefficients.

	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
attack					
smokes	1.210533	.360707	3.36	0.001	.5035523 1.917514
age	.0362795	.0155306	2.34	0.019	.0058399 .0667191
bmi	.1110053	.0485159	2.29	0.022	-.01575 .2062606
hsggrad	.1769547	.4078711	0.43	0.664	-.6224732 .9763826
female	-.1012866	.4183743	-0.24	0.809	-.9212886 .7187154
_cons	-5.680722	1.717448	-3.31	0.001	-9.049706 -2.311738

Using MI to account for the missing values of `bmi`, we find that `age` is a statistically significant contributing factor toward heart attacks. Had we used listwise deletion, we would not have been able to identify `age` as statistically significant.

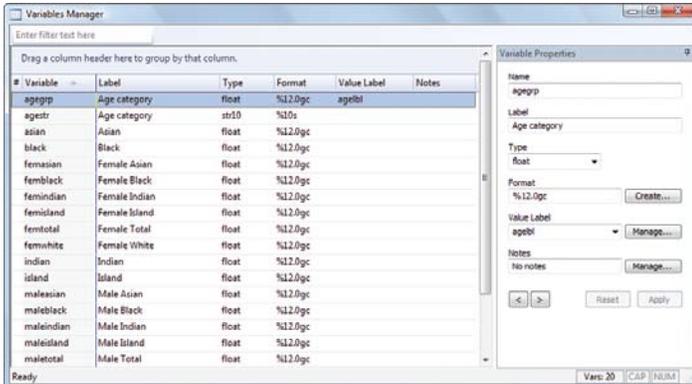
Other methods are available in `mi impute` for imputing single or multiple variables of different types. `mi estimate` supports most Stata estimation procedures, including those for survival and survey data analysis.

Postestimation features include `mi test` and `mi testtransform` for testing linear and nonlinear hypotheses within the MI framework.

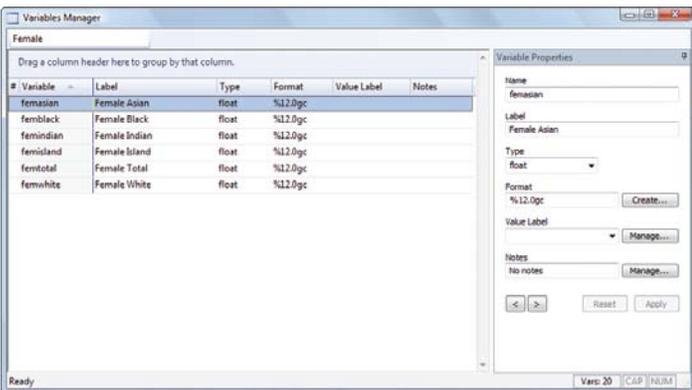
`mi` provides an extensive set of data manipulation commands for complete management of multiply imputed data, including merging, appending MI datasets, and more. In addition, `mi` provides validation routines to verify consistency of multiply imputed data.

In the spotlight: The Variables Manager

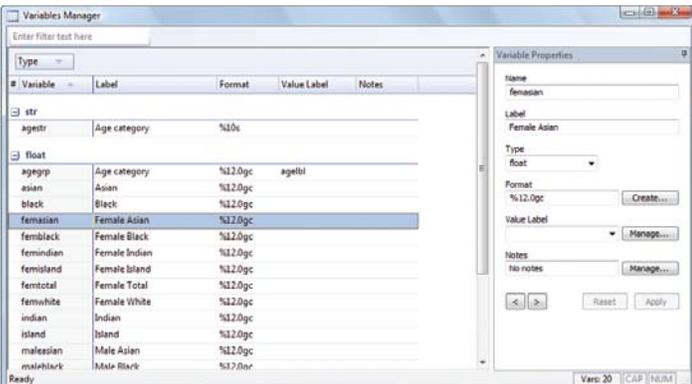
Stata 11's new Variables Manager provides one-stop shopping for changing and managing your variables. Click on a variable to change its name, label, type, format, value label, and notes. Click on a column heading to sort it in ascending order; click on it again for descending order.



You can filter variables to view just a subset of a potentially large dataset. When you type a string in the filter box, the Variables Manager will show only those variables whose characteristics match your search criterion. Here we filter variables for which "Female" appears in any of the columns:

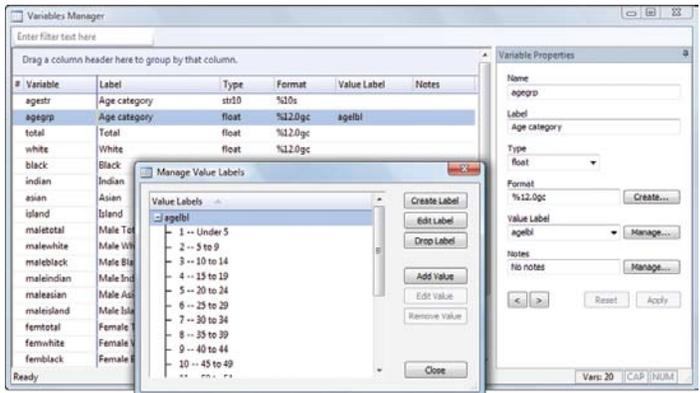


You can also group variables by any of the characteristics shown in the column headings. Drag a column header to the top of the dialog window, and variables will be grouped by that characteristic. Here we group variables by storage type.



Managing value labels is easy with the Variables Manager as well. Select a variable with the value label you want to change, and then click on the Manage... button next to the value-label drop-down box. Among other

tasks, you can add values to an existing label or drop the label altogether. By selecting a variable without a value label, you can attach an existing label or create a new one.



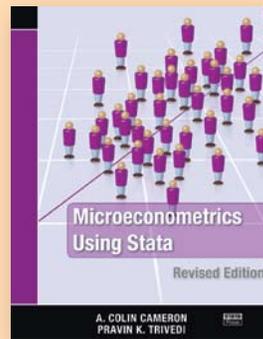
Another powerful feature of the Variables Manager is that you can select multiple variables. Change the data type, display format, and the value label on multiple variables at once.

Reproducibility is a hallmark of Stata's philosophy, and the Variables Manager lives up to that standard while making managing your variables a snap. All the changes you make in the Variables Manager issue Stata commands to achieve their results, so you can log your changes for future reference.

— James Hassell, Senior Software Engineer

New from Stata Press

Microeconometrics Using Stata, Revised Edition



Authors: A. Colin Cameron and Pravin K. Trivedi
 Publisher: Stata Press
 Copyright: 2010
 Pages: 706; paperback
 ISBN-10: 1-59718-073-4
 ISBN-13: 978-1-59718-073-3
 Price: \$65.00

Microeconometrics Using Stata, Revised Edition, by A. Colin Cameron and Pravin K. Trivedi, is an outstanding introduction to microeconometrics and how to do microeconomic research using Stata. Aimed at students and researchers, this book covers topics left out of microeconometrics textbooks and omitted from basic introductions to Stata. Cameron and Trivedi provide the most complete and up-to-date survey of microeconomic methods available in Stata.

The revised edition has been updated to reflect the new features available in Stata 11 that are germane to microeconomists. Instead of using `mfx` and the user-written `margin` commands, the revised edition uses the new `argins` command, emphasizing both marginal effects at the means and average marginal effects. Factor variables, which allow you to

specify indicator variables and interaction effects, replace the `xi` command. The new `gmm` command for generalized method of moments and nonlinear instrumental-variables estimation is presented, along with several examples. Finally, the chapter on maximum likelihood estimation incorporates the enhancements made to `m1` in Stata 11.

Early in the book, Cameron and Trivedi introduce simulation methods and then use them to illustrate features of the estimators and tests described in the rest of the book. While simulation methods are important tools for econometricians, they are not covered in standard textbooks. By introducing simulation methods, the authors arm students and researchers with techniques they can use in future work. Cameron and Trivedi address each topic with an in-depth Stata example, and they reference their 2005 textbook, *Microeconometrics: Methods and Applications*, where appropriate.

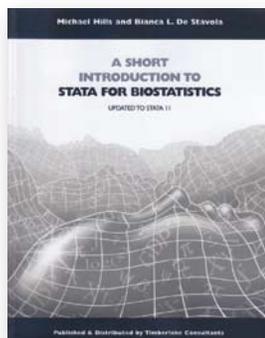
The authors also show how to use Stata's programming features to implement methods for which Stata does not have a specific command. Although the book is not specifically about Stata programming, it does show how to solve many programming problems. These techniques are essential in applied microeconometrics because there will always be new, specialized methods beyond what has already been incorporated into a software package.

The unique combination of topics, intuitive introductions to methods, and detailed illustrations of Stata examples make *Microeconometrics Using Stata* an invaluable, hands-on addition to the library of anyone who uses microeconomic methods.

You can find the table of contents and online ordering information at www.stata-press.com/books/musr.html.

New from the Stata Bookstore

A Short Introduction to Stata for Biostatistics



Authors: Michael Hills and Bianca L. De Stavola
 Publisher: Timberlake Consultants
 Copyright: 2009
 Pages: 188; paperback
 ISBN-10: 0-9557076-4-1
 ISBN-13: 978-0-9557076-4-3
 Price: \$52.00

A Short Introduction to Stata for Biostatistics bridges the information in the *Getting Started* manual and the *Reference* manuals by providing a more detailed introduction to the most often used analytic methods in biomedical research. Although it is written specifically for biostatisticians, epidemiologists, and health professionals new to Stata, the book is useful for more experienced users wanting more in-depth knowledge of both Stata commands and biostatistical issues. The book is hands on, intended to be used

while working with Stata, and includes a CD-ROM containing the datasets and several author-written programs.

The first four chapters provide an overview of data entry and management commands, including those used to create, label, and drop variables and used to sort observations. After two chapters on graphics, the bulk of the book details methods used in data description and analysis. Beginning with commands used to create frequency tables and summary statistics, the book proceeds to describe commands used for univariate and multivariate analyses, including linear regression, Poisson regression, logistic regression, and survival data analysis. Included among the final chapters is a useful tutorial for writing your own Stata programs.

New additions for Stata 11 include a section on competing-risks analysis, new chapters on report generation and on meta analysis, and a description of the factor-variable notation.

You can find the table of contents and online ordering information at www.stata.com/bookstore/sisb.html.

Veterinary Epidemiologic Research, 2nd Edition



Authors: Ian Dohoo, Wayne Martin, and Henrik Stryhn
 Publisher: VER Inc.
 Copyright: 2010
 Pages: 865; hardcover
 ISBN-10: 0-919013-60-0
 ISBN-13: 978-0-919013-60-5
 Price: \$119.00

Veterinary Epidemiologic Research, Second Edition, is a new edition of a popular graduate-level text on veterinary epidemiology. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology (except some of the diseases may not be familiar).

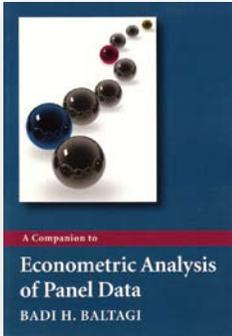
New are chapters on Bayesian methods, spatial data, and the epidemiology of infectious diseases. There are also expanded discussions of meta-analysis, diagnostic tests, survival analysis, controlled studies, clustering, and repeated measures.

All terms and epidemiological measures are clearly defined, and all notations and formulas are identified with examples. Designs discussed in this text include cohort studies, case-control studies, two-stage sampling designs, and controlled trials. Several statistical models are also discussed: linear regression, logistic regression, multinomial logistic regression, the Poisson model, survival analysis, and mixed-effects models.

The datasets used in the examples, as well as examples using Stata, are thoroughly described in the text and are available on the book's web site.

You can find the table of contents and online ordering information at www.stata.com/bookstore/ver.html.

A Companion to Econometric Analysis of Panel Data



Author: Badi H. Baltagi
 Publisher: Wiley
 Copyright: 2009
 Pages: 312; paperback
 ISBN-10: 0-470-74403-0
 ISBN-13: 978-0-470-74403-1
 Price: \$39.75

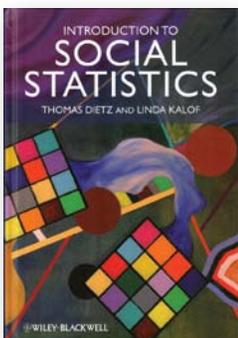
A Companion to Econometric Analysis of Panel Data will appeal to students and researchers learning about panel data with Baltagi's *Econometric Analysis of Panel Data, Fourth Edition* and other graduate-level econometrics texts. *Companion* includes derivations of results shown without proof in Baltagi's textbook and provides more numerical examples, most of which were completed using Stata.

Working through *Companion* is like having a personal tutor help you work through a graduate text like Baltagi's. The organization of material in *Companion* parallels that of Baltagi's panel-data text but presents the material in a pedagogical manner consisting of some background material followed by many sample exercise questions and complete solutions. While one may at first frown at the idea of a book including solutions to exercises, this text delivers them in a way that truly motivates learning rather than mere copying: the questions are sufficiently involved that by seeing how to solve them, students develop a method to attack mathematical proofs and think like an econometrician.

Baltagi's *A Companion to Econometric Analysis of Panel Data*, while focused on panel data, is more generally useful to all students of econometrics, because it teaches and instills tools that are applicable to an array of problems.

You can find the table of contents and online ordering information at www.stata.com/bookstore/ceapd.html.

Introduction to Social Statistics: The Logic of Statistical Reasoning



Authors: Thomas Dietz and Linda Kalof
 Publisher: Wiley-Blackwell
 Copyright: 2009
 Pages: 569; hardcover
 ISBN-10: 1-4051-6902-8
 ISBN-13: 978-1-4051-6902-8
 Price: \$69.00

In *Introduction to Social Statistics: The Logic of Statistical Reasoning*, Thomas Dietz and Linda Kalof present and expand on standard introductory

statistics. They detail how statistical concepts have or could relate to everyday events, such as studies of crime versus poverty or treaty participation. The authors take their time exposing each new topic and are not shy about delving into ancillary topics, such as profiling famous statisticians or exploring the role of women in the field of statistics. Such discussions are boxed in gray to readily set them apart from the main discussion.

Dietz and Kalof cover various topics, including describing and displaying univariate data, plotting relationships between variables, probability, sampling distributions, estimation, confidence intervals, hypothesis tests, ANOVA, contingency tables, and regression.

You can find the table of contents and online ordering information at www.stata.com/bookstore/itss.html.

Additional featured titles

Econometric Analysis of Panel Data, 4th Edition

Author: Badi H. Baltagi
 Price: \$52.00

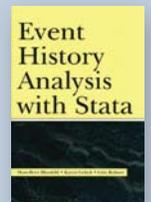
If you need to know how to perform estimation and inference on panel data from an econometric standpoint, then this is the book to read.



Event History Analysis with Stata

Authors: Hans-Peter Blossfeld, Katrin Golsch, and Götz Rohwer
 Price: \$33.00

This book presents survival analysis from a social-science perspective. It is aimed at the professional social scientist but could also serve as a graduate-level text.



Fixed Effects Regression Models

Author: Paul D. Allison
 Price: \$16.75

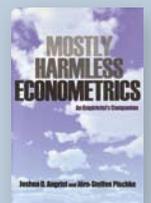
This is a useful handbook that concentrates on the application of fixed-effects methods for a variety of data situations, from linear regression to survival analysis.



Mostly Harmless Econometrics: An Empiricist's Companion

Authors: Joshua D. Angrist and Jörn-Steffen Pischke
 Price: \$28.00

All graduate students and researchers should read this book. This instructive and irreverent romp through micro-econometrics is as much of a page-turner as we are likely to see in a book about statistical methods.



For more information on these and other titles, visit www.stata.com/bookstore/.

Seminars on Stata



Date: April 8, 2010

Venue: Waterview Conference Center
1919 North Lynn Street
Arlington, VA 22209

Price: \$195

Register: www.stata.com/meeting/dcsem10/

Join us in Washington, D.C., at the Waterview Conference Center on Thursday, April 8, 2010, for Seminars on Stata—a series of high-level seminars on Stata and statistics. Learn how to work more efficiently and take advantage of Stata's unique features for various types of data, including panel and complex survey, as well as data that contain missing values.

The meeting will run from 9:30 AM to 3:30 PM, and lunch will be provided. Registration begins at 8:45 AM with a continental breakfast.

The speakers are

Roberto G. Gutierrez, *Director of Statistics*
Bill Rising, *Director of Educational Services*

Program

Easy automation and reproducible analysis

Learn how to use both script files and the Stata GUI (menus, dialog boxes, Variables Manager, Data Editor, and Do-file Editor) to perform reproducible analyses with both result and command logging.

Panel/longitudinal data and multilevel mixed-effects modeling

We will briefly cover the wide range of commands in Stata for estimating models of continuous, count, and binary outcomes with fixed effects and random effects. We will then extend random-effects estimation to intercepts and coefficients at multiple levels. These multilevel models are estimated by `xtmixed` for continuous outcomes, `xtmelogit` for binary outcomes, and `xtmepoisson` for count outcomes. All three commands share a similar syntax for model specification and for postestimation analysis.

Survey data

Most of Stata's estimation commands are equipped to automatically handle data from complex surveys. So long as we declare the survey aspects of our data, the estimates and their standard errors are adjusted for pre- and poststratification, multilevel sampling (clustering), and weighted sampling. We will cover declaring survey data and estimation, as well as the three primary survey variance estimators—linearization, balanced repeated replication, and jackknife.

Multiple imputation for missing data

Multiple imputation provides a unified framework for handling missing data that is missing at random (MAR) or missing completely at random (MCAR). We will introduce Stata's suite of `mi` commands for imputation, estimation, and data management.

Special topics

We will cover a number of special topics: 1) how the division of estimation and postestimation (estimates, tests and confidence intervals of linear and nonlinear combinations, marginal effects, linear and nonlinear predictions, etc.) provides a common and powerful framework for performing analyses, 2) Stata's extensibility and its relation to the active Stata user community, and 3) graphics, graphics editing, and creating custom graph profiles. We will also briefly discuss how our newly learned knowledge applies to other estimation areas, such as survival analysis, univariate and multivariate time series, and multivariate methods.

Logistics organizers

Sarah Marrs (smarrs@stata.com)
Karen Strope (kstrope@stata.com)

Public training courses

Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals

Date: May 19–20, 2010

Location: CompuWorks
99 Bedford Street
Boston, MA 02111

Instructor: Bill Rising, StataCorp

Cost: \$950

Register: www.stata.com/training/enroll.html
Enrollment is limited to 20 participants.

Become intimately familiar with all three components of Stata: data management, analysis, and graphics. This two-day course is aimed at both new Stata users and those who would like to optimize their workflow and pick up tips for efficient day-to-day usage of Stata. Upon completion of the course, you will be able to use Stata efficiently for basic analyses and graphics. You will be able to do this in a reproducible manner, making collaborative changes and follow-up analyses much simpler. Finally, you will be able to make your datasets self-explanatory to your co-workers and your future self.

Whether you currently own Stata 11 or you are considering an upgrade or new purchase, this course will unquestionably make you more proficient with Stata's wide-ranging capabilities.

Course topics

- Stata basics
- Data management
- Workflow
- Analysis
- Graphics

For more information, see www.stata.com/training/eff_stata.html.

Want to be kept informed of new public training courses and other events?

Sign up for email alerts at

www.stata.com/alerts/

Multilevel/Mixed Models Using Stata

Date: May 6–7, 2010

Location: MicroTek
1101 Vermont Ave. NW
Suite 300
(at "L" Street)
Washington, DC 20005

Instructor: Roberto G. Gutierrez, StataCorp

Cost: \$1295

Register: www.stata.com/training/enroll.html
Enrollment is limited to 24 participants.

This two-day course is an introduction to using Stata to fit multilevel/mixed models. Mixed models contain both fixed effects analogous to the coefficients in standard regression models and random effects not directly estimated but instead summarized through the unique elements of their variance–covariance matrix. Mixed models may contain more than one level of nested random effects, and hence these models are also referred to as multilevel or hierarchical models, particularly in the social sciences. Stata's approach to linear mixed models is to assign random effects to independent panels where a hierarchy of nested panels can be defined for handling nested random effects.

Course topics

- Introduction to linear mixed models
- Random coefficients and hierarchical models
- Advanced topics
- Postestimation analysis
- Nonlinear models

For more information, go to www.stata.com/training/mixed.html.

Enrollment in public training courses is limited. Computers with Stata 11 installed are provided at all public training sessions. A continental breakfast, lunch, and an afternoon snack will also be provided. All training courses run from 8:30 AM to 4:30 PM each day. Participants are encouraged to bring a USB flash drive to all public training sessions; this is the safest and simplest way to save your work from the session.

For a complete schedule of upcoming training courses, visit www.stata.com/training/public.html.

Stata Conference Boston 2010

Dates: July 15–16, 2010

Venue: Omni Parker House, Boston
60 School Street
Boston, MA 02108



Cost: Single day \$125, students \$50
Both days \$195, students \$75

Submissions: March 31, 2010

Details: www.stata.com/meeting/boston10/

The Stata Conference Boston 2010 will be held on July 15 and 16 at the Omni Parker House hotel, located in downtown Boston near the Boston Common and the Park Street T.

Call for presentations

All users are encouraged to submit abstracts for possible presentations. Presentations on any Stata-related topic will be considered, including (but not limited to) the following topics:

- new user-written commands, including commands for modeling and estimation, graphical analysis, data management, or reporting
- use or evaluation of existing Stata commands
- methods for using Stata or for teaching statistics with Stata
- case studies of Stata use in novel areas or applications
- surveys or critiques of Stata facilities in specific fields
- comparisons of Stata to other software or use of Stata together with other software

The deadline for submissions is March 31, 2010.

Presentations by StataCorp team members include

Competing-risks survival regression in Stata

Roberto G. Gutierrez, *Director of Statistics*

Multiple imputation in Stata

Yulia U. Marchenko, *Senior Statistician*

Scientific organizers

Liz Allred, *Harvard School of Public Health*

Kit Baum, *Boston College*

Amresh Hanchate, *Boston University*

Marcello Pagano, *Harvard School of Public Health*

Logistics organizer

Sarah Marrs, *StataCorp LP*

Mexican Stata Users Group meeting

Date: April 29, 2010

Venue: Universidad Iberoamericana, Mexico City campus
Prolongación Paseo de la Reforma 880, Lomas de Santa Fe, México C.P. 01219, Distrito Federal, México

Cost: 1,160 pesos professionals; 580 pesos students

Details: www.stata.com/meeting/mexico10/

Organized by MultiON Consulting S.A. de C.V., distributor of Stata in Mexico and Central America, the second Stata Users Group meeting in Mexico will be held on Thursday, April 29, 2010, at the Universidad Iberoamericana in Mexico City. Make plans to attend. Presentations include:

Keynote lecture: Estimation of count-data panel models

Pravin K. Trivedi (*Indiana University*)

An impact evaluation of PROCAMPO

Omar Stabridis, Janet Zamudio, and Mario Paulín (*Consejo Nacional de Evaluación de la Política de Desarrollo Social*)

Stata as a tool for transparency and statistics dissemination: Measuring multidimensional poverty in México

Víctor H Pérez, Dulce Cano, and Rocío Espinosa (*Consejo Nacional de Evaluación de la Política de Desarrollo Social*)

Measuring poverty at state level using Stata

Carlos Guerrero de Lizardi and Manuel Lara Caballero (*Consejo Nacional de Evaluación de la Política de Desarrollo Social*)

Hierarchical linear models using Stata

Delno Vargas Chanes (*Colegio de México*); and María Merino (*ITAM*)

Generating descriptive statistics from the MXFLS

Alicia Santana Cartas (*Universidad Iberoamericana*)

Bivariate dynamic probit models for panel data

Alfonso Miranda (*Institute of Education, University of London*)

Selection bias correction based on the multinomial logit: An application to the Mexican labor market

Luis Huesca (*Centro de Investigación en Alimentación y Desarrollo*)

Generalized method of moments estimators in Stata

David Drukker (*Director of Econometrics, StataCorp LP*)

Using Stata to analyze size frequency of the life cycle of a Mexican desert spider

Irma Gisela Nieto-Castañeda (*Centro de Investigaciones Biológicas del Noroeste*); María Luisa Jiménez-Jiménez (*S.C.*); and Isaías H. Salgado-Ugarte (*FES Zaragoza UNAM*)

ML capabilities: Stata vs Gauss

Armando Sánchez Vargas (*Institute for Economic Research, UNAM*)

Analyzing data from complex survey designs

Isabel Cañette (*Senior Statistician, StataCorp LP*)

Report to users and “Wishes and grumbles”

Save the date

German Stata Users Group meeting

Sophia Rabe-Hesketh will be the keynote speaker.

Date: June 25, 2010

Venue: Berlin Graduate School of Social Sciences
Luisenstraße 56, House 1
10117 Berlin-Mitte, Germany

Cost: €35 professionals; €15 students

Submissions: March 30, 2010

Details: www.stata.com/meeting/germany10/

UK Stata Users Group meeting

Dates: September 9–10, 2010

Venue: Cass Business School
106 Bunhill Row, London EC1Y 8TZ, UK

Cost: Both days: £94.00 professionals; £64.63 students
Single day: £64.63 professionals; £47.00 students

Submissions: May 31, 2010

Details: www.stata.com/meeting/uk10/

Italian Stata Users Group meeting

Date: November 11–12, 2010

Venue: Grand Hotel Baglioni
Via Indipendenza, 8
40121 Bologna, Italy

Cost: €90 day 1 only; €375 day 1 + a training course

Submissions: August 30, 2010

Details: www.stata.com/meeting/italy10/

NetCourse™ schedule

Enroll by visiting www.stata.com/netcourse/.

NC101: Introduction to Stata

An introduction to using Stata interactively.

Dates: March 19–April 30, 2010

Enrollment deadline: March 18, 2010

Price: \$95

Details: www.stata.com/netcourse/nc101.html

NC151: Introduction to Stata Programming

An introduction to Stata programming dealing with what most statistical software users mean by programming, namely, the careful performance of reproducible analyses.

Dates: March 19–April 30, 2010

Enrollment deadline: March 18, 2010

Price: \$125

Details: www.stata.com/netcourse/nc151.html

NC152: Advanced Stata Programming

This course teaches you how to create and debug new commands that are indistinguishable from those of official Stata. It is assumed that you know why and when to program and, to some extent, how. You will learn how to parse both standard and nonstandard Stata syntax by using the intuitive **syntax** command, how to manage and process saved results, how to process bygroups, and more.

Dates: October 8–November 26, 2010

Enrollment deadline: October 7, 2010

Price: \$150

Details: www.stata.com/netcourse/nc152.html

NC461: Introduction to Univariate Time Series with Stata

This course introduces univariate time-series analysis, emphasizing the practical aspects most needed by practitioners and applied researchers. The course is written to appeal to a broad array of users, including economists, forecasters, financial analysts, managers, and anyone who encounters time-series data.

Dates: October 8–November 26, 2010

Enrollment deadline: October 7, 2010

Price: \$295

Details: www.stata.com/netcourse/nc461.html

Contact us

StataCorp
4905 Lakeway Dr.
College Station, TX 77845
USA

Phone 979-696-4600
Fax 979-696-4601
Email service@stata.com
Web www.stata.com

Please include your Stata serial number with all correspondence.



Copyright 2010 by StataCorp LP.

To locate a Stata international distributor near you, visit www.stata.com/worldwide/.