The Stata News

Statistics Graphics Data Management & Analysis

Statalist is now a forum!

Statalist is a discussion group for discussing all things Stata, and it has proven to be a great resource for learning Stata since 1994. Because the focus is on discussion, it has also proven useful for advanced users.

On March 31, Statalist was moved from being an old-fashioned mailing list to being a modern, web-based forum. It's better and, to use the modern jargon, it scales better, too. The old Statalist-the-mailing-list had 5,229 members as of last count and averaged some 47-plus posts per work day. It was Statalist's very success that limited Statalist's enrollment. Most people don't want that many emails dumped on them every day.

Statalist-the-modern-forum can go far beyond those numbers and only gets more useful.

You can learn a lot from Statalist-the-modern-forum just by lurking. You can get even more out of it by joining and participating in the discussions. Either way, visit

statalist.org

Statalist has always been and continues to be run by Stata users, although StataCorp set up the forum software and is running the servers for the new Statalist.

Dr. Marcello Pagano of Harvard University has been Statalist's moderator since 1996, and he continues in that capacity on the new Statalist. The original Statalist's searchable archives are still available and will continue to be available; visit **statalist.org/archives**. They are a valuable resource for Stata users.

But the most valuable resource is the new Statalist, and we recommend you try it out. You can join (for free, as usual) or you can lurk. We bet you'll lurk awhile and then join. Among so many other reasons, Statalist is a great place to catch the ear of Stata's development staff. All development



staff are required to at least lurk, and most join. As E.F. Hutton's marketing department was fond of saying, "When E.F. Hutton talks, people listen". At Statalist, it's the other way around. If you speak, we guarantee the development staff is listening, and they will respond.

But as important—and often more important—is that other Stata users are listening and responding, too. Statalist is not just about reaching Stata's developers. Most discussion is between Stata users—and highly competent ones at that. Statalist is a great place to learn how to get the most out of Stata.

Pay Statalist a visit. Go to **statalist.org**.

A message from Marcello Pagano

[Marcello Pagano, School of Public Health, Harvard University, has been Statalist's moderator since 1996. – Ed.]

Statalist has now moved from its comfortable mailing-list format that it has worn the last 20 years to a new format. This should improve things tremendously because will we not only be able to transfer all the good things about Statalist, such as fantastic response and help from terrific listers, but also be able to post graphs, data, equations, etc. Plus, it is all on the web, so our mailboxes will not be filled by Statalist mailings and can now be dedicated

In the spotlight:
Meet Stata's new xtmlogit command 2
Path diagram for multinomial logit
with random effects3
New from Stata Press 4
New from the Stata Bookstore 5
Stata Conference Boston 2014 6

2014 Stata Users Group meetings	8
Public training courses	9
NetCourses™1	0
ICPSR Summer Program1	1
Upcoming events1	1
They're here! Series 6 bookmarks 1	2

The Stata News

Executive Editor............ Karen Strope
Production Supervisor... Annette Fett



to being notified about prizes won and bequeathments made by long-lost uncles who went to the ends of the world to make their fortune.

One of the important reasons for the change was that Bill Mahoney—somebody most Statalisters don't even know and who has been working in the background keeping the software and hardware going all these years—is about to retire. That plus the Sun workstation that has housed Statalist is turning 20, and in dog years that is like 400 years, so it must be put out to pasture.

So goodbye to

Statalist@hsphsun2.harvard.edu,
which will shut down at the end of
May, and hello to statalist.org, which

is now up and running.

Go visit. Join, it is still free—no better bang for your buck—and you can continue to feast on Nick's nuggets of wisdom. I will remain the moderator.

A lot of work went on behind the scenes to make this happen. First, a committee was put together to argue design issues: Kit Baum, Maarten Buis, Nick Cox, Bill Gould, Phil Schumm, and myself. Then the hard work of implementation was carried out by Alan Riley and his crew: Kevin Crow, Pete Huckelba, and Annette Fett. Once you visit statalist.org, you will see what a terrific job they have done.

So go try it out and learn to love it. As we said on the list, Statalist has a long history of helpful, constructive, and polite discussion. I expect to see that continue in the Statalist family. If you have never joined, go and take a peek.



In the spotlight: Meet Stata's new xtmlogit command

Okay, so that title is a bit unfair. We added **xtologit** and **xtoprobit**, but the truth is, there is no **xtmlogit** command. The good news is that you can still fit multinomial logit models to panel data.

If you are like me, you love Stata's intuitive panel commands. I'm new to StataCorp, and before I got here, I thought panel-data analysis meant I needed a command that started with xt. It turns out that Stata is far more flexible than I knew, and that flexibility comes from a command I never really paid much attention to—gsem. Don't let the name fool you. It's still intuitive. Really.

Suppose that I want to analyze the probability of three work-related outcomes for women: employed, unemployed, and not participating in the labor force. I have data on 4,685 women along with some covariates that I think might influence their outcomes.

varname	Description				
idcode	ndividual identifier				
workstat	status (1=employed, 2=unemployed, 3=out of labor force)				
age	age in years				
race	race (1=white, 2=black, 3=other)				
student	indicator for being a student in the current period				
msp	indicator for married, spouse present				

Knowing what I know about the **xt** syntax, I imagine this should work:

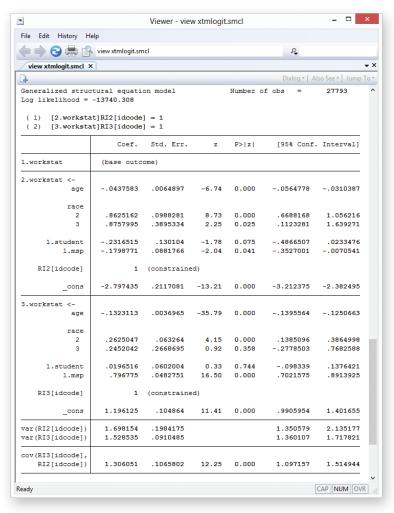
xtmlogit workstat age i.race i.student i.msp, re

But what would **re** mean here? Should Stata assume I want one random effect for each level of **workstat** or one random effect overall? If I have multiple random effects, should they be independent or should they be correlated?

Behind the scenes, Stata is doing a lot of work for you.

The intuition for setting up the **gsem** syntax is that I get to tell Stata exactly what I want. Because I'm fitting a multinomial model, I will omit employed as the base category and use factor-variable notation to specify a separate equation for the other outcomes (for example, **2.workstat** means unemployed). I tell Stata that the outcome is a function of the covariates with <-*varlist*. I want one latent variable that varies at the individual level, which I call **RI#[idcode]**, to capture the random effect for outcome #.

gsem (2.workstat <- age i.race i.student i.msp
 RI2[idcode])
 (3.workstat <- age i.race i.student i.msp
 RI3[idcode]), mlogit</pre>



Suppose I am interested in how marital status affects work status. The section for **2.workstat** shows a significant negative coefficient for **msp**, so I conclude that married women have a lower probability of being unemployed relative to being employed. Similarly, married women appear to have a higher probability of being out of the labor force. But that's just **mlogit** stuff. If you're reading this, you probably are more interested in what is different.

The var(RI2[idcode]) and var(RI3[idcode]) terms give the variances of the random effects. In practical terms, these are the variances of the idiosyncratic individual effects. I'm used to reporting random effects in terms of standard deviations. Each of these variance estimates is stored in e(b), so I can see my usual standard deviations this way:

. display sqrt(_b[var(RI2[idcode]):_cons]) 1.3031322

Just beneath that, **gsem** reports the covariance of my two random terms. The covariance is significant, so I conclude there is an underlying correlation between the two random intercepts.

Now, before you rush out and try this, there are a few caveats. Behind the scenes, Stata is doing a lot of work

for you. Every random effect has to be numerically integrated out, and that is time intensive. All the usual rules about time increasing with respect to the number of random effects apply. However, now the number of random effects increases with the number of equations.

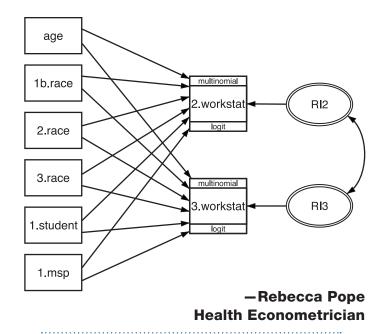
This little model took about five minutes to fit, equivalent to the time spent for a trip to the breakroom for coffee. But, if I had added several more outcomes or random coefficients ... Let's just say it's always nice to take a vacation.

You can see more examples of what **gsem** is capable of in the [SEM] manual. A couple of my other personal favorites are the Heckman selection model and endogenous treatment effects.

- Rebecca Pope Health Econometrician

Path diagram for multinomial logit with random effects

If you are already familiar with **gsem**, you probably wonder why there is no path diagram in the previous article. I was afraid that if I included it, the panel-data folks would think it was just another article about **gsem** and skip the whole thing. As a treat for sticking with me through all that preceding text, here is the missing path diagram.



New from Stata Press

A Gentle Introduction to Stata, Fourth Edition

Author: Alan C. Acock Publisher: Stata Press Copyright: 2014

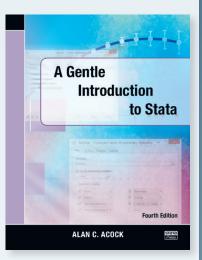
ISBN-13: 978-1-59718-142-6

Price: \$54.00

Updated for Stata 13, this book delves into new features, such as power and sample-size analysis using the **power** suite of features (specifically two-sample tests of means, as well as one-way, two-way, and repeated measures ANOVA) and effect-size computations using **esize** and **estat esize**. This edition also boasts a new chapter that gets you up and running with Stata's SEM Builder and **sem** and **gsem**. These are just a few examples of the features covered in Acock's book.

The Stata technical group has the following to say about the book:

Acock is quite careful to teach the reader all aspects of using Stata. He covers data management, good work habits (including the use of basic dofiles), basic exploratory statistics (including graphical displays), and analyses using the standard array of basic statistical tools (correlation, linear and



logistic regression, and parametric and nonparametric tests of location and dispersion). He also successfully introduces some more advanced topics such as multiple imputation and structural equation modeling in a very approachable manner.

Acock teaches Stata by using the menus and dialog boxes while still stressing the value of do-files. In this way, he ensures that all types of users can build good work habits. Each chapter has exercises that the motivated reader can use to reinforce the material.

Read the entire review and order online at stata-press.com/books/gentle-introduction-to-stata.

An Introduction to Stata for Health Researchers, Fourth Edition

Authors: Svend Juul and Morten Frydenberg

Publisher: Stata Press Copyright: 2014

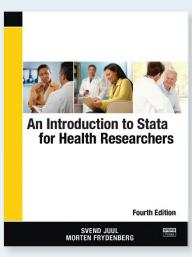
ISBN-13: 978-1-59718-135-8

Price: \$52.00

The fourth edition has been substantially revised based on new features from Stata 12 and Stata 13. The updated material has been streamlined while including new features in Stata.

The Stata technical group has the following to say about the book:

Svend Juul and Morten Frydenberg's An Introduction to Stata for Health Researchers, Fourth Edition is distinguished in its careful attention to detail. The reader will learn not only the skills for statistical analysis but also the skills to make the analysis reproducible. The authors use a friendly,



down-to-earth tone and include tips gained from a lifetime of collaboration and consulting.

The book is based on the assumption that the reader has some basic knowledge of statistics but no knowledge of Stata. The authors build the reader's abilities as a builder would build a house: laying a firm foundation in Stata, framing a general structure in which good work can be accomplished, adding the details that are particular to various types of statistical analyses, and, finally, trimming with a thorough treatment of graphics and special topics such as power and sample-size computations.

Read the entire review and order online at stata-press.com/books/introduction-stata-health-researchers.

Speaking Stata Graphics

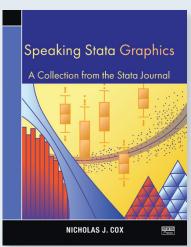
Author: Nicholas J. Cox

Publisher: Stata Press Copyright: 2014

ISBN-13: 978-1-59718-144-0

Price: \$45.00

Speaking Stata Graphics is ideal for researchers who want to produce effective publication-quality graphs. A compilation of articles from the popular "Speaking Stata" column by Nicholas J. Cox, this book provides valuable insights about Stata's built-in and user-written statistical graphics commands.



Read the table of contents and order online at stata-press.com/books/speaking-stata-graphics.

One Hundred Nineteen Stata Tips, Third Edition

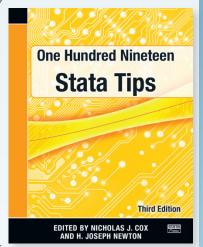
Editors: Nicholas J. Cox and H. Joseph Newton

Publisher: Stata Press Copyright: 2014

ISBN-13: 978-1-59718-143-3

Price: \$35.00

One Hundred Nineteen Stata Tips provides concise and insightful notes about commands, features, and tricks that will give you a deeper understanding of Stata. The book is a result of contributions from the Stata community that have appeared in the Stata Journal since 2003.



Read the table of contents and order online at

stata-press.com/books/stata-tips.

New from the Stata Bookstore

Statistical Analysis of Human Growth and Development

Author: Yin Bun Cheung

Publisher: Chapman & Hall/CRC

Copyright: 2014

ISBN-13: 978-1-43-987154-6

Price: \$82.50

This book covers a broad array of statistical techniques useful to researchers in the fields of human growth and development.

Topics include an overview of basic statistical concepts, causality, linear regression, quantile regression, logistic regression, survival analysis, mixed models, fractional polynomials, item—response theory,



validity, reliability, and multiple imputation.

Read the entire review or order online at stata.com/bookstore/statistical-analysis-of-human-growth-and-development.

Quasi-Least Squares Regression

Authors: Justine Shults and Joseph M. Hilbe

Publisher: Chapman & Hall/CRC

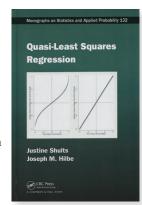
Copyright: 2014

ISBN-13: 978-1-4200-9993-5

Price: \$79.50

Quasi-Least Squares Regression is a great resource for graduate students and researchers interested in estimating population-averaged effects from longitudinal and clustered data.

The book provides an introduction to quasi-least-squares (QLS) estimators and provides many examples using author-written



Stata commands. QLS estimators extend GEE estimators to incorporate a wider set of correlation structures. The QLS estimators also allow researchers to determine the correlation structure that best fits their data.

Read the entire review and order online at stata.com/bookstore/quasi-least-squares-regression.



Come join us in historic Boston, home to Fenway Park and the Harvard Museum of Natural History, for two days of networking and Stata exploration. Don't miss this opportunity to connect with colleagues and fellow researchers as well as Stata developers.

A preliminary selection of user-written talks:

Profile analysis

Phil Ender UCLA Statistical Consulting Group

Optimal interval design for phase I oncology clinical trials

Bryan Fellman MD Anderson Cancer Center

Using Stata for educational accountability and compliance reporting

Billy Buchanan Mississippi Department of Education

Computer simulation of patient flow through an operating suite

David Clark Maine Medical Center, Portland, Maine

Classification using random forests in Stata and R

Linden McBride Cornell University

Making interactive online visualizations with stata2leaflet and stata2d3

Robert Grant St George's Medical School, University of London

A midas retouch regarding diagnostic meta-analysis

Ben Dwamena University of Michigan

Mata routines for solution of nonlinear systems using interval methods

Matthew Baker Hunter College and the Graduate Center, CUNY

bipolate: A Stata command for bivariate interpolation with particular application to 3D graphing

Joseph Canner Johns Hopkins University School of Medicine

Stata hybrids: Updates and ideas

James Fiedler Universities Space Research Association

ctgov: A suite of Stata commands for reporting trial results to ClinicalTrials.gov

Phil Schumm Department of Health Studies, University of Chicago



StataCorp will also be represented:

Transformation survival models

Yulia Marchenko StataCorp LP

Estimating average treatment effects from observational data using teffects

David Drukker StataCorp LP

The popular Wishes and Grumbles and Report to users will also be given.

Come early, stay late

Whether you stay for the JSM or just to relax, be sure to enjoy what Boston has to offer. Take a cruise in Boston Harbor, walk the Freedom Trail, visit Fenway Park, and have a bowl of "chowdah". Boston is a great city with plenty to do and see.

Scientific committee

- Stephen Soldz (Chair)
 Boston Graduate School of Psychoanalysis
- Kit Baum Boston College
- Marcello Pagano Harvard University

When	July 31–August 1, 2014				
Where	Omni Parker House 60 School Street Boston, Massachusetts				
Who	Stata 13 developers You and Stata users from around the world				

Accommodations

The Omni Parker House is offering a special rate of \$229 per night for Stata Conference attendees staying between July 29 and August 2, 2014. Book by June 30 to receive the special rate.

Reserve your room now.

Omni Parker House 60 School Street Boston, Massachusetts 02108 (617) 227-8600

Visit **stata.com/boston14** for web reservation information.

Register online:

stata.com/boston14



Stata Users Group meetings are international conferences where the use of Stata is discussed across a wide-ranging breadth of fields and environments. The meetings are open to everyone. In past years, participants have traveled from around the world to attend the events. Representatives from StataCorp will also be in attendance.

Find out more at stata.com/meeting.

Germany			
Date	June 13 (workshop June 12)		
Venue	University of Hamburg Hamburg, Germany		
Cost	TBA		
Details	stata.com/meeting/germany14		

Portugal					
Date	September 19				
Venue	Instituto Superior de Economia e Gestão Lisbon, Portugal				
Cost	ТВА				
Details	stata.com/meeting/portugal14				

Nordic and Baltic			
Date	September 5		
Venue	Aarhus University Aarhus, Denmark		
Cost	Free		
Details	stata.com/meeting/nordic-and-baltic14		

Italy					
Date	Date November 13-14 (workshop November 1				
Venue	Hotel NH Touring Milano Milan, Italy				
Cost	€95 meeting €400 meeting + workshop; students: 35% off				
Details	stata.com/meeting/italy14				

UK			
Date	September 11–12		
Venue	Cass Business School London, UK		
Cost	Both days: £96 regular; £66 students		
	Single day: £66 regular; £48 students		
Details	stata.com/meeting/uk14		

Public training courses

Learn Stata from StataCorp's experts. These courses are ideal for researchers and individuals who want to learn more or gain a deeper understanding of Stata.

■ Regression Modeling Using Stata

June 2, 2014, Washington, DC

Regression modeling is a fundamental tool for researchers who want to establish causal quantitative relationships from observational data. Learn the theoretical concepts necessary to understand regression models and how to implement them using Stata, and learn to reinforce those concepts with exercises and examples you will solve with the assistance of the instructor.

■ Panel-Data Analysis Using Stata

June 3-4, 2014, Washington, DC

Learn both the theory and the practice of panel-data analysis. After introducing the fixed-effects and random-effects approaches to unobserved individual-level heterogeneity, the course covers linear models with exogenous covariates, linear models with endogenous variables, dynamic linear models, and nonlinear models. Exercises and Stata examples supplement the lessons.



■ Time-Series Analysis Using Stata

June 5-6, 2014, Washington, DC

Learn how to efficiently use Stata for time-series analysis. Topics include data management, estimation of univariate and

multivariate stationary and non-stationary time series, model selection, hypothesis testing, and interpretation. Exercises and Stata examples supplement the lessons.

■ Using Stata Effectively: Data Management, Analysis, and Graphics Fundamentals June 23–24, 2014, Washington, DC

Aimed at both new Stata users and those who wish to learn techniques for efficient day-to-day use of Stata, this course enables you to use Stata in a reproducible manner, making collaborative changes and follow-up analyses much simpler. Exercises and Stata examples supplement the lessons.

Computers with Stata installed are provided at all public training sessions. Enrollment is limited.

Find out more and enroll at stata.com/public-training.

June						
SU	MO	TU	WE	TH	FR	SA
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

Save \$195 when you combine

Regression Modeling Using Stata and Panel-Data Analysis Using Stata.

Find out how at stata.com/public-training.

NetCourses[™]

NetCourses are convenient web-based courses that teach you how to exploit the full power of Stata. Learn Stata from the comfort of your own home or office!

Introduction to Survival Analysis **Using Stata**



Learn how to effectively analyze survival data using Stata. We cover censoring, truncation, hazard rates, and survival functions. Topics include data preparation, descriptive statistics, life tables, Kaplan-Meier curves, and semiparametric (Cox) regression and parametric regression. Discover how to set the survival-time characteristics of your dataset just once and apply any of Stata's many estimators and statistics to those data.

This course is written for everyone who uses Stata, whether health researchers or social scientists.

> Dates: June 13-August 1, 2014

Cost: \$295

Don't forget our other courses!

Introduction to Stata

Learn how to use all of Stata's tools and become a sophisticated Stata user. You will understand the Stata environment, how to import and export data from different formats, how Stata's intuitive syntax works, data management in Stata, matching and merging, how to analyze subgroups of data, how to reproduce your work and document it for publication and review, how to interact with the Stata community online, and more.

Videos, lectures, examples, and exercises combine to develop your Stata skills.

> Dates: July 11-August 22, 2014

Cost: \$95

Introduction to Stata Programming

Become an expert in organizing your work in Stata. Make the most of Stata's scripting language to improve your workflow and create concretely reproducible analyses. Learn how branching, looping, flow of control, and accessing saved estimation results can speed up your work and lead to more complete analyses. Learn bootstrapping and Monte Carlo simulations, too.

> Dates: July 11-August 22, 2014

Cost: \$150



Advanced Stata Programming

Learn how to create and debug your own commands that are indistinguishable from the commands in Stata. You will be able to parse both standard and nonstandard Stata syntax using the intuitive syntax command, to manage and process stored results, to post your own stored results, to process by-groups, to create data management commands, to program your own maximum likelihood estimator, and more. In short, learn to create commands that act just like the commands that ship with Stata.

> Dates: October 10-November 28, 2014

Cost: \$175

Introduction to Univariate Time Series with Stata

Learn about univariate time-series analysis with an emphasis on the practical aspects most needed by practitioners and applied researchers. Written for a broad array of users, including economists, forecasters, financial analysts, managers, and anyone who wants to analyze time-series data. Become expert in handling date and date-time data; time-series operators; time-series graphics; basic forecasting methods; and ARIMA, ARMAX, and seasonal models.

We provide lecture material, detailed answers to the questions posted at the end of each lecture, and access to a discussion board on which you can post questions for other students and the course leader to answer.

> Dates: October 10-November 28, 2014

Cost: \$295

stata.com/netcourse

The dates above don't work for you? No problem! NetCourseNow allows you to set the time and work at your own pace as well. It also gives you a personal NetCourse instructor to guide you through the course. Visit stata.com/netcourse/ncnow.

ICPSR Summer Program in Quantitative Methods of Social Research

June-August 2014

The ICPSR Summer Program in Quantitative Methods of Social Research is recognized throughout the world as the preeminent forum for basic and advanced training in the methodologies and technologies of social science research. ICPSR's general instructional philosophy emphasizes the integration of methodological strategies with the theoretical and practical concerns that arise in research on substantive social issues. The ICPSR Summer Program also creates a unique and supportive social environment that facilitates professional networking and encourages the exchange of ideas about the theory and practice of social science research.

The Summer Program takes place from early June to late August. In addition to two four-week sets of courses in Ann Arbor, Michigan, ICPSR offers intensive one-week workshops in Ann Arbor and at several additional sites.

In 2014, ICPSR is offering several workshops that will be of particular interest to the Stata user community:

Models for Categorical Outcomes Using Stata: Specification, Estimation, and Interpretation

Scott Long, PhD, Indiana University, Bloomington June 16–20

Structural Equation Modeling with Stata

Kristin MacDonald, PhD, StataCorp June 18–20

Panel-Data Analysis Using Stata

David Drukker, PhD, StataCorp July 7–11

Managing Statistical Research: The Workflow of Data Analysis

Scott Long, PhD, Indiana University, Bloomington July 7–11

Analyzing Multilevel and Mixed Models Using Stata

Bill Rising, PhD, StataCorp August 11–13

For further details about courses, visit www.icpsr.umich.edu/icpsrweb/sumprog.

Upcoming events

Visit with Stata developers at the following upcoming meetings:

APS 2014 Annual Convention

Association for Psychological Science

May 22–25 in San Francisco, California Attending: Chuck Huber, Senior Statistician

JSM 2014 Annual Meeting

Joint Statistical Meetings

August 2–7 in Boston, Massachusetts Attending: Kristin MacDonald, Senior Statistician; Xiao Yang, Senior Statistician; Charles Lindsey, Senior Statistician; Yulia Marchenko, Director of Biostatistics; and Bill Rising, Director of Educational Services

APA 2014 Annual Convention

American Psychological Association

August 7–10 in Washington, DC Attending: Chuck Huber and Kristin MacDonald, Senior Statisticians

ASA 2014 Annual Meeting

American Sociological Association

August 16–19 in San Francisco, California Attending: Rose Medeiros, Senior Statistician

APSA 2014 Annual Meeting

American Political Science Association

August 28–31 in Washington, DC Attending: Kristin MacDonald, Senior Statistician

APHA 2014 Annual Meeting

American Public Health Association

November 15–19 in New Orleans, Louisiana Attending: Bill Rising, Director of Educational Services

CAIR 2014 Annual Meeting

California Association for Institutional Research

November 19–21 in San Diego, California Attending: Chuck Huber, Senior Statistician

stata.com/news/conferences



StataCorp 4905 Lakeway Drive College Station, TX 77845-4512 **USA**

Return service requested.

Contact us

979-696-4600 service@stata.com 979-696-4601 (fax)

stata.com

Please include your Stata serial number with all correspondence.

Find a Stata distributor near you: stata.com/worldwide











State Copyright 2014 by State Corp LP. State is a registered trademark of State Corp LP.

They're here!

The next series of commemorative statistician bookmarks is now available.

Series 6 celebrates five prominent women in statistics and related fields: Stella Cunliffe, Grace Murray Hopper, Isobel Loutit, Ada Lovelace, and Margaret E. Martin.

You can order yours today for only \$2.25 in North America (\$3.50 elsewhere). The price includes shipping.

Collect them all!

घवत. Sala

stata.com/giftshop/bookmarks



Stella Cunliffe



Grace Hopper



Isobel Loutit



Ada Lovelace



Margaret Martin