Jan/Feb/Mar
Vol 29 No 1

The Stata® News

Statistics Graphics Data Management & Analysis

In the spotlight: Export tables to Excel®

A new feature in Stata 13, **putexcel**, allows you to easily export matrices, expressions, and stored results to an Excel file. Combining **putexcel** with a Stata command's stored results allows you to put the table displayed in your Stata Results window in an Excel file. Let me show you.

A stored result is simply a scalar, macro, or matrix stored in memory after you run a Stata command. The two main types of stored results are e-class (for estimation commands) and r-class (for general commands). You can list a command's stored results after it has been run by typing **ereturn list** (for estimation commands) or **return list** (for general commands). Let's try a simple example by loading the auto dataset and running **correlate** on the variables **foreign** and **mpg**:

```
. sysuse auto
(1978 Automobile Data)

. correlate foreign mpg
(obs=74)

             |  foreign      mpg
-------------+------------------
     foreign |   1.0000
         mpg |   0.3934   1.0000
```

Because **correlate** is not an estimation command, we use **return list** to see its stored results.

```
. return list

scalars:
                  r(N) =  74
                r(rho) =  .3933974152205484

matrices:
                  r(C) :  2 x 2
```

Now we can use **putexcel** to export these results to Excel.

The basic syntax of **putexcel** is

```
putexcel excel_cell=(expression) …
             using filename [, options]
```

If you are working with matrices, the syntax is

```
putexcel excel_cell=matrix(expression) …
             using filename [, options]
```

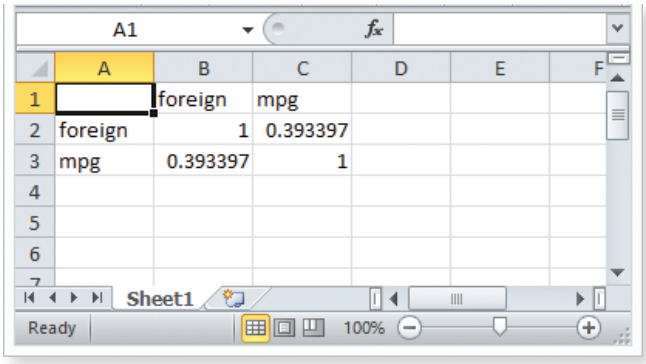
It is easy to build the above syntax in the **putexcel** dialog. We have a helpful video on our YouTube channel about the dialog (**stata.com/videos13/saving-estimation-results-to-excel**). Let's list the matrix **r(C)** to see what it contains.

```
. matrix list r(C)

symmetric r(C)[2,2]
             foreign         mpg
foreign            1
    mpg   .39339742           1
```

To re-create the table in Excel, we need to export the matrix **r(C)** with the matrix row and column names. In your Stata Command window, type

```
. putexcel A1=matrix(r(C), names) using corr
```

To export the matrix row and column names, we used the **names** option after we specified the matrix **r(C)**. When we open the file **corr.xlsx** in Excel, the table below is displayed.



In the Spotlight: Double-robust treatment effects (two wrongs don't make a right, but one does) 3
New from the Stata Gift Shop 5
Public training courses 6
NetCourses 7
Upcoming events 8
Short courses 8
New from the Stata Bookstore 9
Stata Conference 10
2014 Stata Users Group meetings 11
Don't forget: You can go green! 12

The Stata News
Executive Editor ..... Karen Strope
Production Supervisor ... Annette Fett

Next let's try a more involved example. Reload the auto dataset, and run a tabulation on the variable **foreign**. Because **tabulate** is not an estimation command, we use **return list** to see its stored results.

```
. sysuse auto, clear
(1978 Automobile Data)

. tabulate foreign

   Car type |      Freq.     Percent        Cum.
------------+-----------------------------------
   Domestic |         52       70.27       70.27
    Foreign |         22       29.73      100.00
------------+-----------------------------------
      Total |         74      100.00

. return list

scalars:
                  r(N) =  74
                  r(r) =  2
```

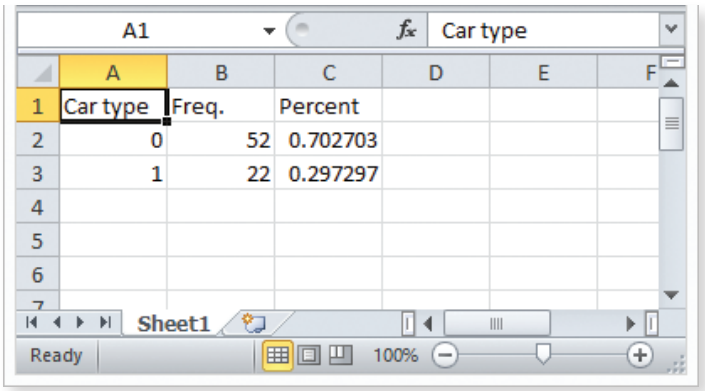
tabulate is different from most commands in Stata: it does not automatically save all the results we need in the stored results. We need to use the **matcell()** and **matrow()** options of **tabulate** to save its results into two Stata matrices.

```
. tabulate foreign, matcell(freq) matrow(names)

   Car type |      Freq.     Percent        Cum.
------------+-----------------------------------
   Domestic |         52       70.27       70.27
    Foreign |         22       29.73      100.00
------------+-----------------------------------
      Total |         74      100.00

. matrix list freq

freq[2,1]
    c1
r1  52
r2  22

. matrix list names

names[2,1]
    c1
r1   0
r2   1
```

The **putexcel** commands below create a basic tabulation table in Excel.

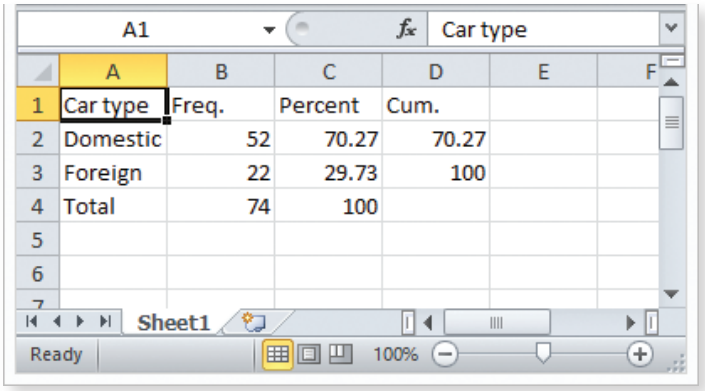
```
. putexcel A1=("Car type") B1=("Freq.")
      C1=("Percent") using results, replace
. putexcel A2=matrix(names) B2=matrix(freq)
      C2=matrix(freq/r(N)) using results, modify
```

Here is the resulting Excel table:



You probably noticed that this table does not include cumulative percentages or the total number of cars. Moreover, our "Car type" column contains the numeric values of the **foreign** variable rather than the value labels **Domestic** and **Foreign**.

With a bit of programming, you can overcome these limitations. On the Stata Blog, I have posted a short do-file that exports the table by **tabulate** exactly as it appears in the Stata Results window. Go to **blog.stata.com/2013/09/25/export-tables-to-excel** to get it. With that program, we get this Excel spreadsheet:



In the blog post, I also provide a simple command that combines **tabulate** and **putexcel** into one handy command. I explain how to use **putexcel** to format the exported Excel tables in another blog post—**blog.stata.com/retain-format-with-putexcel**. You can also learn how to quickly export estimation results at **stata.com/stata13/create-word-and-excel-files**.

—Kevin Crow
Senior Software Developer

In the spotlight: Double-robust treatment effects (two wrongs don't make a right, but one does)

If you ever wanted an extra shot at getting your treatment-effects model right, **teffects** can help you.

teffects allows you to write a model for the treatment and a model for the outcome. We will show how—even if you misspecify one of the models—you can still get correct estimates using doubly robust estimators.

In experimental data, the treatment is randomized so that a difference between the average treated outcomes and the average nontreated outcomes estimates the average treatment effect (ATE).

"The shocking fact is that only one of the two models must be correct to estimate the ATE ..."

Suppose you want to estimate the ATE of a mother's smoking on her baby's birthweight. The ethical impossibility of asking a random selection of pregnant women to smoke mandates that these data be observational. Which women choose to smoke while pregnant almost certainly depends on observable covariates, such as the mother's age.

We use a conditional model to make the treatment as good as random. More formally, we assume that conditioning on observable covariates makes the outcome conditionally independent of the treatment. Conditional independence allows us to use differences in model-adjusted averages to estimate the ATE.

The regression-adjustment (RA) estimator uses a model for the outcome. The RA estimator uses a difference in the average predictions for the treated and the average predictions for the nontreated to estimate the ATE. Below we use **teffects ra** to estimate the ATE when conditioning on the mother's marital status, her education level, whether she had a prenatal visit in the first trimester, and whether it was her first baby.

```
. webuse cattaneo2
(Excerpt from Cattaneo (2010) Journal of Econometrics 155: 138-154)

. teffects ra (bweight mmarried prenatal1 fbaby medu) (mbsmoke)

Iteration 0:   EE criterion =  4.582e-24
Iteration 1:   EE criterion =  5.097e-26

Treatment-effects estimation                    Number of obs      =      4642
Estimator      : regression adjustment
Outcome model  : linear
Treatment model: none
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-230.9541	24.34012	-9.49	0.000	-278.6599	-183.2484
POmean						
mbsmoke nonsmoker	3402.548	9.546721	356.41	0.000	3383.836	3421.259

Mothers' smoking lowers the average birthweight by 231 grams.

The inverse-probability-weighted (IPW) estimator uses a model for the treatment instead of a model for the outcome; it uses the predicted treatment probabilities to weight the observed outcomes. The difference between the weighted treated outcomes and the weighted nontreated outcomes estimates the ATE. Conditioning on the same variables as

above, we now use **teffects ipw** to estimate the ATE:

```
. teffects ipw (bweight) (mbsmoke mmarried prenatal1 fbaby medu)

Iteration 0:   EE criterion =  1.701e-23
Iteration 1:   EE criterion =  4.947e-27

Treatment-effects estimation                    Number of obs      =      4642
Estimator      : inverse-probability weights
Outcome model  : weighted mean
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-231.1516	24.03183	-9.62	0.000	-278.2531	-184.0501
POmean						
mbsmoke nonsmoker	3402.219	9.589812	354.77	0.000	3383.423	3421.015

Mothers' smoking again lowers the average birthweight by 231 grams.

We could use both models instead of one. The shocking fact is that only one of the two models must be correct to estimate the ATE, whether we use the augmented-IPW (AIPW) combination proposed by Robins and Rotnitzky (1995) or the IPW-regression-adjustment (IPWRA) combination proposed by Wooldridge (2010).

The AIPW estimator augments the IPW estimator with a correction term. The term removes the bias if the treatment model is wrong and the outcome model is correct, and the term goes to 0 if the treatment model is correct and the outcome model is wrong.

The IPWRA estimator uses IPW probability weights when performing RA. The weights do not affect the accuracy of the RA estimator if the treatment model is wrong and the outcome model is correct. The weights correct the RA estimator if the treatment model is correct and the outcome model is wrong.

We now use **teffects aipw** to estimate the ATE:

```
. teffects aipw (bweight mmarried prenatal1 fbaby medu)   ///
>         (mbsmoke mmarried prenatal1 fbaby medu)

Iteration 0:   EE criterion =  2.153e-23
Iteration 1:   EE criterion =  1.802e-26

Treatment-effects estimation                    Number of obs      =      4642
Estimator      : augmented IPW
Outcome model  : linear by ML
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-229.7809	24.96839	-9.20	0.000	-278.718	-180.8437
POmean						
mbsmoke nonsmoker	3403.122	9.564165	355.82	0.000	3384.376	3421.867

Mothers' smoking lowers the average birthweight by 230 grams.

Finally, we use **teffects ipwra** to estimate the ATE:

```
. teffects ipwra (bweight mmarried prenatal1 fbaby medu)  ///
>         (mbsmoke mmarried prenatal1 fbaby medu)

Iteration 0:   EE criterion =  3.901e-22
Iteration 1:   EE criterion =  1.373e-25

Treatment-effects estimation                    Number of obs      =      4642
Estimator      : IPW regression adjustment
Outcome model  : linear
Treatment model: logit
```

bweight	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
ATE						
mbsmoke (smoker vs nonsmoker)	-227.4408	25.62591	-8.88	0.000	-277.6667	-177.215
POmean						
mbsmoke nonsmoker	3403.027	9.56025	355.96	0.000	3384.289	3421.765

Mothers' smoking lowers the average birthweight by 227 grams.

All of these results tell a similar story, so we assume that both the outcome and the treatment models are correct. When both models are correct, the AIPW estimator is more efficient than either the RA or the IPW estimator. We started off in search of robustness and ended up with extra efficiency.

References

Robins, J. M., and A. Rotnitzky. 1995. Semiparametric efficiency in multivariate regression models with missing data. *Journal of the American Statistical Association* 90: 122–129.

Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

—David M. Drukker
Director of Econometrics

Public training courses

Learn Stata from StataCorp's experts. These two-day courses are ideal for researchers and individuals that want to learn or have a deeper understanding of Stata.

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

March 4–5, 2014, Washington, DC
April 21–22, 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 supplement the lessons and Stata examples.

Estimating Average Treatment Effects Using Stata

March 6–7, 2014, Washington, DC

Learn how and when to use Stata's treatment-effects estimators to analyze treatment effects in observational data. Use regression adjustment, inverse probability weights, doubly robust methods, propensity-score matching, and covariate matching to estimate average treatment effects (ATEs) and ATEs on the treated. We will cover the conceptual and theoretical underpinnings of treatment effects as well as many examples using Stata.

Structural Equation Modeling Using Stata

March 24–25, 2014, Washington, DC

Learn how to illustrate, specify, and estimate structural equation models in Stata using Stata's SEM Builder and the **sem** command. The course introduces several types of models, including path analysis, confirmatory factor analysis, full structural equation models, and latent growth curves. Exercises supplement the lessons and Stata examples.

Multilevel/Mixed Models Using Stata

April 23–24, 2014, Washington, DC

Measure and account for clustering and grouping at multiple levels. Whether linear or nonlinear, multilevel modeling allows for random intercepts and slopes at multiple levels, reducing the problems of too much or too little data aggregation. The course is interactive, uses real data, offers ample opportunity for specific research questions, and provides exercises to reinforce what you learn.

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

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

MARCH							APRIL						
SU	MO	TU	WE	TH	FR	SA	SU	MO	TU	WE	TH	FR	SA
						1			1	2	3	4	5
2	3	4	5	6	7	8	6	7	8	9	10	11	12
9	10	11	12	13	14	15	13	14	15	16	17	18	19
16	27	18	19	20	21	22	20	21	22	23	24	25	26
23	24	25	26	27	28	29	27	28	29	30			
30	31												

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

NEW

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.

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: March 7–April 18, 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: March 7–April 18, 2014
Cost: $125

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: $150

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**.

Upcoming events

MPSA 2014 Annual Conference

Midwest Political Science Association
April 3–6 in Chicago, Illinois
Attending: Kristin MacDonald, Senior Statistician

AERA 2014 Annual Meeting

American Educational Research Association
April 3–7 in Philadelphia, Pennsylvania
Attending: Chuck Huber, Senior Statistician

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; 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: Chuck Huber, Senior Statistician, and Bill Rising, Director of Educational Services

stata.com/news/conferences

Short courses

Looking for Stata courses in your area? Check our list of third-party short courses, held throughout the world.

Short courses are held by a variety of institutions to help people learn more about statistics and Stata. Short courses are offered by institutions other than StataCorp and may be of interest to Stata users.

Currently, short courses are scheduled in

- USA
- Belgium
- Brazil
- Kenya
- Pakistan
- Singapore
- United Arab Emirates
- UK

However, more courses are continually added.

Go to **stata.com/meeting/short-courses** to see what courses are available in your area.

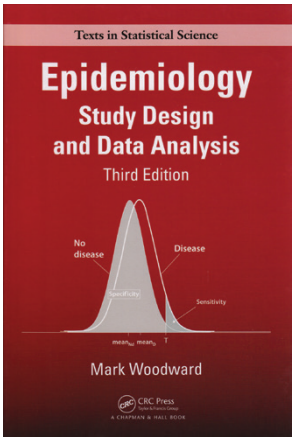
If you are teaching a short course using Stata, please notify us at **service@stata.com** of the course name, instructor, details, and dates, and we will add it to our webpage.

stata.com/meeting/short-courses

New from the Stata Bookstore

Epidemiology: Study Design and Data Analysis, Third Edition

Author:	Mark Woodward
Publisher:	Chapman & Hall/CRC
Copyright:	2013
ISBN-13:	978-1-439-83970-6
Price:	$84.75



Woodward's third edition of *Epidemiology: Study Design and Data Analysis* has two target audiences: researchers who need statistical solutions to epidemiology problems and statisticians who wish to learn how their science applies to epidemiology. This book successfully presents statistical principles in epidemiology in a manner that is neither too theoretical nor too replete with medical jargon. It provides complete treatment of the topic, from simple contingency tables to meta-analysis. The book uses real data throughout—more than 20 large datasets are cataloged for download—and the end of each chapter has exercises. Woodward makes Stata code for working many of the examples available for download.

Topics include basic terminology, causality, descriptive statistics, testing of means, relative risks versus odds ratios, exact tests based on tables, tests for linear and nonlinear trends, confounding and interaction, direct and indirect standardization, cohort designs, case–control studies, intervention studies, power and sample size, linear models (including analysis of variance), logistic and other models for binary responses, survival analysis (including Cox regression), and meta-analysis. The third edition has been expanded to include risk scores and clinical decision rules, bootstrapping, multiple imputation, binomial regression models, competing risks, propensity scoring, and splines.

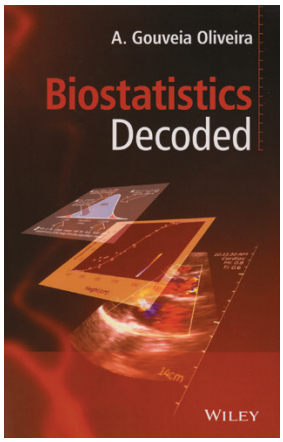
Read more or order online at
stata.com/bookstore/epidemiology-sdda.

Biostatistics Decoded

Author:	A. Gouveia Oliveira
Publisher:	Wiley
Copyright:	2013
ISBN-13:	978-1-119-95337-1
Price:	$54.50

Biostatistics Decoded is an introduction to biostatistics for medical professionals and clinical researchers. Oliveira emphasizes concepts and basic calculations that will provide the reader with a foundation for understanding the study designs and statistical methods reported in the scientific literature.

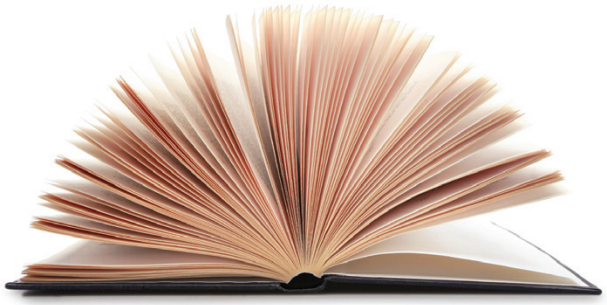
The book is comprehensive and includes basic descriptive and inferential statistics as well as advanced topics such as the analysis of longitudinal studies, survival analysis, factor analysis, and meta-analysis. A variety of study designs are also covered, including stratified and multistage sampling designs and modern experimental designs such as adaptive clinical trials and noninferiority trials.



Oliveira avoids mathematical proofs, instead using diagrams, graphs, and simulations to illustrate ideas. Familiarity with basic arithmetic, square roots, and logarithms is sufficient, and no knowledge of calculus is necessary. All examples are worked using Stata.

Read more or order online at
stata.com/bookstore/biostatistics-decoded.

The Stata Bookstore offers a wide variety of titles on statistics and Stata that our development staff feels may be of interest to Stata users.



Many more titles online!

stata.com/bookstore

STATA® CONFERENCE

BOSTON 2014

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.

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

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

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.

BOSTON 2014

Register online:

stata.com/boston14

2014 Stata Users Group meetings

Save the date

The 12th German Stata Users Group meeting will be held at the University of Hamburg on Friday, June 13, 2014. Everyone who is interested in using Stata is invited.

Registration is now open. Visit **stata.com/meeting/germany14** for details.

The final program, cost, and venue will be circulated in April 2014.

Scientific committee

- Dirk Enzmann
  University of Hamburg
- Johannes Giesecke
  University of Bamberg
- Ulrich Kohler
  University of Potsdam
- Kai-Uwe Schnapp
  University of Hamburg

Find more details online at **stata.com/meeting/germany14**.

Date	June 13, 2014
Venue	University of Hamburg Hamburg, Germany
Cost	TBA
Details	stata.com/meeting/germany14

Save the date

The 2014 UK Stata Users Group meeting is a two-day international conference where the use of Stata is discussed across a wide-ranging breadth of fields and environments. The meeting is open to everyone.

If you are interested in giving a presentation, feel free to contact the scientific organizers now.

Scientific committee

- Nicholas J. Cox
  Durham University
- Patrick Royston
  MRC Clinical Trials Unit at UCL

Find more details online at **stata.com/meeting/uk14**.

Date	September 11–12, 2014
Venue	Cass Business School London, UK
Cost	TBA
Details	stata.com/meeting/uk14

More dates and locations available soon

stata.com/meeting

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

Return service requested.

Contact us

979-696-4600 979-696-4601 (fax)
service@stata.com stata.com
Please include your Stata serial number with all correspondence.

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

/StataCorp /Stata /StataCorp /+Stata blog.stata.com

Don't forget:
You can go green!

Just a reminder: the *Stata News* is now available electronically.

Be eco-friendly and go paperless. Have the *Stata News* delivered straight to your inbox.

Sign up to receive future issues of the *Stata News* in electronic format before it arrives in the mail. We will send you an email with the *News* content as soon as it is available. The electronic format of the *Stata News* has all the same information as the printed version—the only difference is the convenience.

Go to

stata.com/news-delivery

to choose your delivery method. Select email, printed, or both.

Get your *News* your way.