

THE STATA NEWS

Volume 18 Number 4

October/November/December 2003

www.stata.com

Stata 8 for Linux64

StataCorp will begin shipping Stata 8 for Linux64 in early December. Linux64 is a 64-bit operating system that runs on AMD's line of 64-bit chips, including the Opteron and the Athlon 64.

We are excited about this new platform because our initial tests have shown that the price-to-performance ratio is excellent for this series of processors, and we expect a wide range of machines to be available. Linux64 joins the other 64-bit platforms currently supported by StataCorp: Sun Solaris and SGI IRIX.

Stata for Linux64 breaks the 2 GB memory limitation shared by most 32-bit operating systems, allowing users with large datasets to access 4 GB, 8 GB, and more. In theory, Stata for Linux64 can use up to 61 billion GB, but we suspect most users will have slightly less memory than that installed.

There is one caveat. If you are considering one of these computers and want access to more than 2 GB of memory, pay special attention to the motherboard's specifications. Several of the motherboards currently shipping for AMD's 64-bit chips support only 2 or 4 GB of RAM even though the chip could address more memory. If you need more memory, be sure that you purchase one of the computers with a motherboard that will support that memory.

Stata reads Haver Analytics databases

Stata 8 for Windows can now read databases compiled and distributed by Haver Analytics (www.haver.com), which maintains over 100 databases that are widely used in economic analysis and business decision making. These databases form one of the most comprehensive collections of financial, macroeconomic, commodity-price, and industry-level data for both developed and developing countries. Haver Analytics also maintains many specialized databases for the U.S. and European countries, including 14 databases of regional data for the U.S. and 6 databases for EUROSTAT and OECD data.

Two new commands provide access to the Haver databases: `haver describe` describes the variables in a Haver database, and `haver use` reads variables from a Haver database into memory. For example,

```
haver describe filename
```

provides a one-line description of each variable in the specified database. Specifying the `detail` option produces a table of output for each variable. After you use `haver describe` to identify the variables of interest, typing

```
haver use varlist filename
```

reads the requested variables into memory. To put all the variables from the specified dataset into memory, you would simply type

```
haver use filename
```

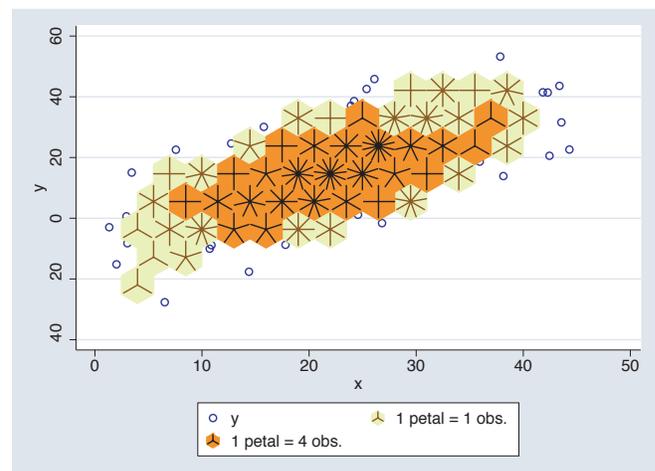
These new commands are included in the October 28, 2003, free update to Stata 8 for Windows. You can obtain this update by typing `update` and following the instructions. After updating, you can access online documentation by typing `help haver`.

Density-distribution sunflower plots

Stata 8 has a new `sunflower` command that draws density-distribution sunflower plots as described in

Dupont, W. D. and W. D. Plummer, Jr. 2003. Density distribution sunflower plots. *Journal of Statistical Software* 8(3): 1–11.

Density-distribution sunflower plots are useful for displaying bivariate data that are too dense to give a visually accurate impression of the joint distribution.

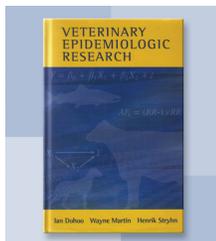


These sunflower plots are available in the current free update to Stata 8. Type `update` and follow the instructions. You can access online help by typing `help sunflower`.

Inside this issue:

Stata 8 for Linux64	1
Stata reads Haver Analytics databases	1
Density-distribution sunflower plots	1
From the Stata Bookstore	2
Latest NetCourse schedule	3
New ODBC data exporting support	4
Update to Stata 8.2	4

From the Stata Bookstore



Title: *Veterinary Epidemiologic Research*
 Authors: I. Dohoo, W. Martin, H. Stryhn
 Publisher: Atlantic Veterinary College
 Copyright: 2003
 Pages: 706; hardcover
 ISBN: 0-919013-41-4
 Price: \$98.00

Veterinary Epidemiologic Research

Veterinary Epidemiologic Research by Ian Dohoo, Wayne Martin, and Henrik Stryhn is a graduate-level text on the principles and methods of veterinary epidemiologic research. Although many of the examples relate to veterinary epidemiology, the principles apply equally to human epidemiology; only some of the diseases may not be familiar.

All terms and epidemiological measures are clearly defined, and all notations and formulas are illustrated 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: www.ucei.ca/ver.

A complete table of contents and online ordering information can be found at www.stata.com/bookstore/ver.html. You can also order using the enclosed bookstore order form.



Title: *Statistical Methods for Rates and Proportions, 3rd ed*
 Authors: J. Fleiss, B. Levin, M. Cho Paik
 Publisher: John Wiley & Sons
 Copyright: 2003
 Pages: 760; hardcover
 ISBN: 0-471-52629-0
 Price: \$94.75

Statistical Methods for Rates and Proportions, 3d ed

Statistical Methods for Rates and Proportions by Joseph Fleiss, Bruce Levin, and Myunghee Cho Paik covers a wide variety of methods for analyzing categorical data, with the emphasis and examples drawn from clinical medicine, epidemiology, psychiatry, psychopathology, and public health. The book is aimed at researchers and students who have at least a year's course in applied statistics, including a thorough grounding in chi-squared tests and correlation. The mathematical background is restricted to high school algebra.

This book has the look and feel of a textbook, complete with examples used to demonstrate methods and motivate questions and exercises (with selected answers) at the end of each chapter. A detailed

bibliography at the end of each chapter provides a valuable desk reference for those interested in probing a topic in more depth.

Covered topics include statistical inference for a single proportion, significance in a fourfold table, sample sizes, randomization, comparative studies, randomized controlled trials, logistic and Poisson regression, regression models for matched samples, correlated binary data, missing data, measurement of interrater agreement, and standardization of rates.

A complete table of contents and online ordering information can be found at www.stata.com/bookstore/smrp.html. You can also order using the enclosed bookstore order form.



Title: *Design of Experiments, 2d ed*
 Authors: R. Kuehl
 Publisher: Brooks/Cole
 Copyright: 2000
 Pages: 666; hardcover
 ISBN: 0-534-36834-4
 Price: \$98.75

Design of Experiments: Statistical Principles of Research Design and Analysis, 2d ed

The second edition of *Design of Experiments: Statistical Principles of Research Design and Analysis* by Robert O. Kuehl is an excellent introduction to the methods of research design. The book emphasizes application but provides sufficient theory for a basic understanding of the statistical principles involved. College algebra and basic probability and statistics are sufficient mathematical background for readers of this book.

Chapter 1 emphasizes performing good science through proper planning and research design. The importance of starting with your research hypothesis, understanding sources of variation, understanding the difference between experimental units and observational units, and the benefits from using replication, randomization, blocking controls, and covariates is explained.

Chapter 2 goes through all the nuts and bolts of ANOVA in a clear and precise way. Chapter 3 covers multiple treatment comparisons. Chapter 4 discusses diagnostic methods for testing model assumptions. Chapters 5, 6, and 7 explain factorial and nested treatments designs for fixed, random, and mixed models. The remaining chapters (8 through 17) are devoted to more complicated experimental designs: complete block, incomplete block, fractional factorial, split-plot, repeated measures, cross-over, and analysis of covariance.

Throughout the text, examples and exercises come from research in life sciences, agriculture, engineering, industry, and chemistry. This book is ideal for research professionals learning research design and analysis of variance (ANOVA). Unlike many texts, the underlying goal of providing answers to real research hypotheses does not get lost while presenting the mathematical details.

A complete table of contents and online ordering information can be found at www.stata.com/bookstore/doe.html. You can also order using the enclosed bookstore order form.

Latest NetCourse™ schedule

The Stata NetCourses™ have been well received by participants, as is reflected in the following remarks:

- “I thought the material was excellent. I was also impressed by the opportunity the students had to ask questions and the efforts made by the Course Leaders to answer them regardless of how complicated the question.”
- “Stata has put together a great course, and I want to thank you for making it so accessible. It has been years since I dealt with statistics and programming language, but I have a project that seems to be just what Stata was developed to support.”
- “The NetCourses are simply great.”

A NetCourse is a “lecture” posted to the NetCourse web site on Friday. After reading the lecture over the weekend or on Monday, participants can post questions and comments to the course discussion area. Course Leaders respond to the questions and comments on Tuesday and Thursday. The other participants are encouraged to amplify or otherwise respond to the questions or comments as well. The next lecture is then posted on Friday, and the process repeats. After the last lecture, discussion continues for a few additional weeks until the course concludes.

A brief summary of upcoming NetCourses is listed below. For more details on how NetCourses work and for course syllabi, visit www.stata.com/info/products/netcourse.

An enrollment form for upcoming NetCourses has been enclosed with the *Stata News*. You can also enroll online at www.stata.com/info/products/netcourse/enrollment.html.

NC-101. Introduction to Stata

NC-101 is designed to take smart, knowledgeable people and turn them into proficient interactive Stata users. The course covers not just the obvious, such as getting data into Stata, but also covers detailed techniques and tricks to make you a powerful Stata user. Many of Stata’s key concepts are explored from web update features and match-merging to using `by` groups and explicit subscribing.

Introduction to Stata
Prerequisites
Stata 8
Dates offered
January 23 – March 5, 2004
Course Leaders
May Boggess and Derek Wagner
Enrollment Deadlines
January 19
Price
\$95

NC-151. Introduction to Stata programming

NC-151 is intended for all Stata users. Through a combination of lectures, example applications, and carefully chosen problems, the course addresses the full range of methods and techniques necessary for you to be most productive in the Stata environment. Beginning with effective ways to organize both simple and complicated analyses in Stata, NetCourse 151 then moves into programming elements that can be used to work more efficiently. Key programming topics include macro processing, program flow of control, using do-files, programming ado-files, Monte-Carlo simulations, and bootstrapped standard errors.

Introduction to Stata programming
Prerequisites
Stata 8; basic knowledge of using Stata interactively
Dates offered
January 23 – March 5, 2004
Course Leaders
Kevin Crow and Brian Poi
Enrollment Deadlines
January 19
Price
\$125

NC-152. Advanced Stata programming

The goal of NC-152 is to turn knowledgeable Stata programmers into Stata programming masters. The course assumes that you know what you need to program and, to some extent, how to do it. Now, you want to add your own commands to Stata and learn to produce professional-level tools that you and others can use.

NC-152 teaches you how to create and debug new commands that work the same way as the commands in Stata. You will learn how to parse both standard and nonstandard Stata syntax using the intuitive `syntax` command, how to manage and process saved results, how to process `by` groups, and more.

Advanced Stata programming
Prerequisites
Stata 8; NC-151 or equivalent knowledge
Date offered
January 23 – March 12, 2004
Course Leaders
May Boggess and Brian Poi
Enrollment Deadline
January 19
Price
\$150

New ODBC data exporting support

With the addition of new data exporting commands, Stata 8 can now fully interact with any ODBC data source. You can take complete control of your ODBC data source from within Stata. New features include

- the ability to insert Stata data into an ODBC database table with the `odbc insert` command
- scripting SQL commands or interactively managing a data source with SQL using the `odbc exec` command
- batch processing of multiple predefined SQL commands via `odbc sqlfile`

Stata 8 for Windows, Macintosh, and Linux can act as ODBC clients. Any registered ODBC source, regardless of platform, can be accessed as a server.

Update to Stata 8.2

That's right, the current version of Stata is 8.2, though for convenience, we continue to call it Stata 8. A glance at the front page of this *News* will show three all-new, major extensions that have been made to Stata in the last few months—and that is really just a small part. Stata 8 was released in January of this year, and since that time, we have made more enhancements, including

- commands to support the FDA file format for NDAs
- ODBC support for Macintosh and Linux
- support for user-written DLLs in C
- special time-series and panel-data graphs
- indexes for all *Stata Journals* published in 2003
- BHHH, DFP, and BFGS optimization in maximum likelihood `m1`
- multiple labels supporting multiple languages
- stem filtering of files
- speed improvements, particularly in graphics
- improved default formatting of dates on graphs
- searching in the Viewer
- improved support for encapsulated PostScript (EPS)
- improved support for enhanced metafiles (EMF)
- support for portable network graphics (PNG) and TIFF
- ODBC insertion statements and batch SQL processing
- a host of new window features for the Macintosh
- better default convergence criterion for maximum likelihood
- a host of small fixes to commands
- access to the manual datasets from the **Help | Contents** menu
- new “easy” graph dialog boxes
- `gllamm` internal support leading to ~200% speed improvement
- support for Haver Analytics databases
- sunflower plots

These enhancements are all free to Stata 8 users and easy to install. The updates are 100% compatible with any programs, analyses, or do-files you are currently using. Type `update` in Stata 8 and follow the instructions. Stata gets better every day. Use `update` to stay current.

THE STATA NEWS is published four times a year and is free to all registered users of Stata.



INTERNATIONAL RESELLERS

CreActive snc

Serving Italy
tel: +39 0575 333297
email: staff@creactive.net

IEM

Serving Botswana, Lesotho, Mozambique, Namibia, South Africa, Swaziland, Zimbabwe
tel: +27-11-8286169
email: instruments@mweb.co.za

Informatique Inc

Serving Japan
tel: +81-3-3505-1250
email: sales@informatiq.co.jp

SOFTWARE shop Inc

Serving Bolivia, Chile, Colombia, Ecuador, Peru, Venezuela
tel: 425-651-4090
email: ventas@software-shop.com

Timberlake Consultores Brasil

Serving Brazil
tel: +55-11-3263-1287
email: info@timberlake.com.br

Timberlake Consultants Polska

Serving Poland
tel: +48 600 370 435
email: info@timberlake.pl

INTERNATIONAL DISTRIBUTORS

Chips Electronics

Serving Brunei, Indonesia, Malaysia, Singapore
tel: 62 - 21 - 452 17 61
email: puyuh23@indo.net.id

Cosinus Computing BV

Serving The Netherlands
tel: +31 416-378 125
email: info@cosinus.nl

Dittrich & Partner Consulting

Serving Austria, Czech Republic, Germany, Hungary, Poland
tel: +49 2 12 / 26 066 - 0
email: sales@dpc.de

Ixon Technology Company Ltd

Serving Taiwan
tel: +886-(0)2-27045535
email: hank@ixon.com.tw

Mercostat Consultores

Serving Argentina, Brazil, Paraguay, Uruguay
tel: 598-2-613-7905
email: mercostat@adinet.com.uy

Metrika Consulting

Serving the Baltic States, Denmark, Finland, Iceland, Norway, Sweden
tel: +46-708-163128
email: sales@metrika.se

MultiON Consulting S.A. de C.V.

Serving Belize, Costa Rica, El Salvador, Guatemala, Honduras, Mexico, Nicaragua, Panama
tel: 52 (55) 55 59 40 50
email: info@multion.com.mx

Ritme Informatique

Serving Belgium, France, Luxembourg
tel: +33 (0)1 42 46 00 42
email: info@ritme.com

Scientific Solutions S.A.

Serving Switzerland
tel: 41 (0)21 711 15 20
email: info@scientific-solutions.ch

Survey Design & Analysis Services

Serving Australia, New Zealand
tel: +61 (0)3 9878 7373
email: sales@survey-design.com.au

Timberlake Consultants

Serving Eire, U.K.
tel: +44 (0)208 697 3377
email: info@timberlake.co.uk

Timberlake Consulting S.L.

Serving Spain
tel: +34 (9) 5 560 14 30
email: timberlake@timberlakeconsulting.com

Timberlake Consultores, Lda.

Serving Portugal
tel: +351 214 702 869
email: timberlake.co@mail.telepac.pt

Tstat S.r.l.

Serving Italy
tel: +39 0864 210101
email: tstat@tstat.it

Vishvas Marketing-Mix Services

Serving India
tel: 91-22-25892639
email: bandya@ymmsmeet.com

HOW TO REACH US

StataCorp

4905 Lakeway Drive
College Station TX 77845
USA

PHONE 979-696-4600

FAX 979-696-4601

EMAIL stata@stata.com

WEB www.stata.com

Please include your Stata serial number with all correspondence.