

Contents of User's Guide

BASICS

1	Read this—it will help	3
1.1	Getting Started with Stata	4
1.2	The User's Guide and the Reference manuals	4
1.2.1	Sample datasets	5
1.2.2	Cross-referencing	6
1.2.3	The index	7
1.2.4	The subject table of contents	7
1.2.5	Typography	7
1.2.6	Vignettes	8
1.3	What's new	8
1.3.1	New matrix language	9
1.3.2	Survey statistics	9
1.3.3	Longitudinal/panel data	12
1.3.4	Time-series statistics	13
1.3.5	Multivariate statistics	14
1.3.6	Survival analysis	19
1.3.7	General-purpose statistics	19
1.3.8	New ML features	25
1.3.9	Functions and expressions	26
1.3.10	Data management	28
1.3.11	Graphics	29
1.3.12	User interface	32
1.3.13	Programming	36
1.3.14	Documentation	38
1.3.15	What's more	38
1.4	References	39
2	A brief description of Stata	41
3	Resources for learning and using Stata	45
3.1	Overview	45
3.2	The http://www.stata.com web site	46
3.3	The http://www.stata-press.com web site	46
3.4	The Stata listserv	47
3.5	The Stata Journal and the Stata Technical Bulletin	47
3.6	Updating and adding features from the web	48
3.6.1	Official updates	48
3.6.2	Unofficial updates	48
3.7	NetCourses	49
3.8	Books and other support materials	50
3.9	Technical support	50
3.9.1	Register your software	50
3.9.2	Before contacting technical support	50
3.9.3	Technical support by email	51
3.9.4	Technical support by phone or fax	51
3.9.5	Comments and suggestions for our technical staff	51

4	Stata's online help and search facilities	53
4.1	Introduction	53
4.2	help: Stata's online manual pages	53
4.3	search: Stata's online index	55
4.4	Accessing help and search from the Help menu	56
4.5	More on search	56
4.6	More on help	57
4.7	help contents: Stata's online table of contents	58
4.8	search: All the details	59
4.8.1	How search works	59
4.8.2	Author searches	60
4.8.3	Entry ID searches	60
4.8.4	FAQ searches	61
4.8.5	Return codes	61
4.9	net search: Searching net resources	61
5	Flavors of Stata	63
5.1	Platforms	63
5.2	Stata/SE, Intercooled Stata, and Small Stata	63
5.2.1	Determining which version you own	64
5.2.2	Determining which version is installed	64
5.3	Size limits comparison of Stata/SE, Intercooled Stata, and Small Stata	64
5.4	Speed comparison of Stata/SE, Intercooled Stata, and Small Stata	65
5.5	Features of Stata/SE	65
6	Setting the size of memory	67
6.1	Memory size considerations	67
6.2	Setting memory on the fly: Stata/SE	67
6.2.1	Advice on setting maxvar	68
6.2.2	Advice on setting matsize	69
6.2.3	Advice on setting memory	70
6.3	Setting memory size on the fly: Intercooled Stata	71
6.4	The memory command	71
6.5	Virtual memory and speed considerations	73
6.6	An issue when returning memory to Unix	73
7	–more– conditions	75
7.1	Description	75
7.2	set more off	75
7.3	The more programming command	75
8	Error messages and return codes	77
8.1	Making mistakes	77
8.1.1	Mistakes are forgiven	77
8.1.2	Mistakes stop user-written programs and do-files	77
8.1.3	Advanced programming to tolerate errors	78
8.2	The return message for obtaining command timings	78

9	The Break key	81
9.1	Making Stata stop what it is doing	81
9.2	Side-effects of pressing Break	82
9.3	Programming considerations	82
10	Keyboard use	83
10.1	Description	83
10.2	F-keys	83
10.3	Editing keys in Stata for Windows, Macintosh, and Unix(GUI)	85
10.4	Editing keys in Stata for Unix(console)	86
10.5	Editing previous lines in Stata for all operating systems	87
10.6	Tab expansion of variable names	88
ELEMENTS		
11	Language syntax	91
11.1	Overview	91
11.1.1	varlist	91
11.1.2	by varlist:	92
11.1.3	if exp	94
11.1.4	in range	95
11.1.5	=exp	97
11.1.6	weight	97
11.1.7	options	99
11.1.8	numlist	101
11.1.9	datelist	101
11.1.10	Prefix commands	102
11.2	Abbreviation rules	103
11.2.1	Command abbreviation	104
11.2.2	Option abbreviation	104
11.2.3	Variable-name abbreviation	105
11.3	Naming conventions	105
11.4	varlists	106
11.4.1	Lists of existing variables	106
11.4.2	Lists of new variables	107
11.4.3	Time-series varlists	108
11.5	by varlist: construct	110
11.6	File-naming conventions	114
11.6.1	A special note for Macintosh users	115
11.6.2	A special note for Unix users	115
12	Data	117
12.1	Data and datasets	117
12.2	Numbers	118
12.2.1	Missing values	118
12.2.2	Numeric storage types	122
12.3	Dates	122

12.4	Strings	123
12.4.1	Strings containing identifying data	123
12.4.2	Strings containing categorical data	123
12.4.3	Strings containing numeric data	124
12.4.4	String storage types	124
12.5	Formats: controlling how data are displayed	124
12.5.1	Numeric formats	124
12.5.2	European numeric formats	127
12.5.3	Date formats	128
12.5.4	Time-series formats	129
12.5.5	String formats	130
12.6	Dataset, variable, and value labels	131
12.6.1	Dataset labels	131
12.6.2	Variable labels	132
12.6.3	Value labels	133
12.6.4	Labels in other languages	138
12.7	Notes attached to data	139
12.8	Characteristics	140
13	Functions and expressions	143
13.1	Overview	143
13.2	Operators	144
13.2.1	Arithmetic operators	144
13.2.2	String operators	144
13.2.3	Relational operators	145
13.2.4	Logical operators	146
13.2.5	Order of evaluation, all operators	147
13.3	Functions	147
13.4	System variables (<code>_variables</code>)	148
13.5	Accessing coefficients and standard errors	148
13.5.1	Simple models	149
13.5.2	ANOVA and MANOVA models	149
13.5.3	Multiple-equation models	149
13.6	Accessing results from Stata commands	150
13.7	Explicit subscripting	151
13.7.1	Generating lags and leads	151
13.7.2	Subscripting within groups	152
13.8	Time-series operators	155
13.8.1	Generating lags and leads	155
13.8.2	Operators within groups	155
13.9	Label values	156
13.10	Precision and problems therein	157
14	Matrix expressions	159
14.1	Overview	159
14.1.1	Definition of a matrix	159
14.1.2	<code>matsize</code>	160
14.2	Row and column names	160
14.2.1	The purpose of row and column names	161
14.2.2	Three-part names	163
14.2.3	Setting row and column names	164

14.2.4	Obtaining row and column names	165
14.3	Vectors and scalars	165
14.4	Inputting matrices by hand	165
14.5	Accessing matrices created by Stata commands	166
14.6	Creating matrices by accumulating data	167
14.7	Matrix operators	167
14.8	Matrix functions	168
14.9	Subscripting	169
14.10	Using matrices in scalar expressions	170
15	Printing and preserving output	171
15.1	Overview	171
15.1.1	Starting and closing logs	172
15.1.2	Appending to an existing log	174
15.1.3	Temporarily suspending and resuming logging	174
15.2	Placing comments in logs	175
15.3	Logging only what you type	175
15.4	The log-button alternative	176
15.5	Printing logs	176
16	Do-files	177
16.1	Description	177
16.1.1	Version	178
16.1.2	Comments and blank lines in do-files	179
16.1.3	Long lines in do-files	180
16.1.4	Error handling in do-files	181
16.1.5	Logging the output of do-files	183
16.1.6	Preventing <code>--more--</code> conditions	184
16.2	Calling other do-files	184
16.3	Running a do-file (Stata for Windows)	184
16.4	Running a do-file (Stata for Macintosh)	185
16.5	Running a do-file (Stata for Unix)	186
16.6	Programming with do-files	187
16.6.1	Argument passing	187
16.6.2	Suppressing output	188
17	Ado-files	189
17.1	Description	189
17.2	What is an ado-file?	189
17.3	How can I tell if a command is built in or an ado-file?	190
17.4	Can I look at an ado-file?	190
17.5	Where does Stata look for ado-files?	191
17.5.1	Where are the official ado directories?	192
17.5.2	Where is my personal ado directory?	192
17.6	How do I install an addition?	192
17.7	How do I add my own ado-files?	193
17.8	How do I install official updates?	193

18	Programming Stata	195
18.1	Description	196
18.2	Relationship between a program and a do-file	196
18.3	Macros	199
18.3.1	Local macros	200
18.3.2	Global macros	200
18.3.3	The difference between local and global macros	201
18.3.4	Macros and expressions	202
18.3.5	Double quotes	203
18.3.6	Extended macro functions	205
18.3.7	Macro increment and decrement functions	206
18.3.8	Macro expressions	207
18.3.9	Advanced local macro manipulation	208
18.3.10	Advanced global macro manipulation	208
18.3.11	Constructing Windows filenames using macros	209
18.3.12	Accessing system values	210
18.3.13	Referencing characteristics	210
18.4	Program arguments	211
18.4.1	Named positional arguments	213
18.4.2	Incrementing through positional arguments	214
18.4.3	Using macro shift	216
18.4.4	Parsing standard Stata syntax	217
18.4.5	Parsing immediate commands	218
18.4.6	Parsing nonstandard syntax	218
18.5	Scalars and matrices	220
18.6	Temporarily destroying the data in memory	220
18.7	Temporary objects	220
18.7.1	Temporary variables	221
18.7.2	Temporary scalars and matrices	221
18.7.3	Temporary files	222
18.8	Accessing results calculated by other programs	222
18.9	Accessing results calculated by estimation commands	225
18.10	Saving results	226
18.10.1	Saving results in <code>r()</code>	227
18.10.2	Saving results in <code>e()</code>	228
18.10.3	Saving results in <code>s()</code>	230
18.11	Ado-files	231
18.11.1	Version	233
18.11.2	Comments and long lines in ado-files	233
18.11.3	Debugging ado-files	233
18.11.4	Local subroutines	234
18.11.5	Development of a sample ado-command	235
18.11.6	Writing online help	240
18.12	A compendium of useful commands for programmers	246
18.13	References	246
19	Immediate commands	247
19.1	Overview	247
19.1.1	Examples	248
19.1.2	A list of the immediate commands	249
19.2	The display command	250

20	Estimation and postestimation commands	251
20.1	All estimation commands work the same way	251
20.2	Standard syntax	254
20.3	Replaying prior results	255
20.4	Cataloging estimation results	256
20.5	Specifying the estimation subsample	258
20.6	Specifying the width of confidence intervals	259
20.7	Obtaining the variance-covariance matrix	260
20.8	Obtaining predicted values	261
20.8.1	Using predict	262
20.8.2	Making in-sample predictions	263
20.8.3	Making out-of-sample predictions	263
20.8.4	Obtaining standard errors, tests, and confidence intervals for predictions	264
20.9	Accessing estimated coefficients	265
20.10	Performing hypothesis tests on the coefficients	267
20.10.1	Linear tests	267
20.10.2	Using test	269
20.10.3	Likelihood-ratio tests	270
20.10.4	Nonlinear Wald tests	271
20.11	Obtaining linear combinations of coefficients	272
20.12	Obtaining nonlinear combinations of coefficients	273
20.13	Obtaining marginal effects	274
20.14	Obtaining robust variance estimates	275
20.15	Obtaining scores	280
20.16	Weighted estimation	283
20.16.1	Frequency weights	283
20.16.2	Analytic weights	284
20.16.3	Sampling weights	285
20.16.4	Importance weights	287
20.17	A list of postestimation commands	287
20.18	References	288

ADVICE

21	Inputting data	291
21.1	Overview	291
21.2	Determining which input method to use	292
21.2.1	Entering data interactively	292
21.2.2	If the dataset is in binary format	293
21.2.3	If the data are simple	293
21.2.4	If the dataset is formatted and the formatting is significant	294
21.2.5	If there are no string variables	295
21.2.6	If all the string variables are enclosed in quotes	296
21.2.7	If the undelimited strings have no blanks	297
21.2.8	If you make it to here	297
21.3	If you run out of memory	297
21.4	Transfer programs	298
21.5	ODBC sources	298
21.6	Reference	299
22	Combining datasets	301

23	Dealing with strings	303
23.1	Description	303
23.2	Categorical string variables	303
23.3	Mistaken string variables	304
23.4	Complex strings	305
23.5	Reference	306
24	Dealing with dates	307
24.1	Overview	307
24.2	Dates	308
24.2.1	Inputting dates	308
24.2.2	Conversion into elapsed dates	309
24.2.2.1	The mdy() function	309
24.2.2.2	The date() function	310
24.2.3	Displaying dates	312
24.2.4	Other date functions	315
24.2.5	Specifying particular dates (date literals)	315
24.3	Time-series dates	316
24.3.1	Inputting time variables	317
24.3.2	Specifying particular dates (date literals)	319
24.3.3	Time-series formats	319
24.3.4	Translating between time units	320
24.3.5	Extracting components of time	320
24.3.6	Creating time variables	320
24.3.7	Setting the time variable	321
24.3.8	Selecting periods of time	321
24.3.9	The %tg format	322
25	Dealing with categorical variables	323
25.1	Continuous, categorical, and indicator variables	323
25.1.1	Converting continuous variables to indicator variables	323
25.1.2	Converting continuous variables to categorical variables	324
25.1.3	Converting categorical variables to indicator variables	326
25.2	Using indicator variables in estimation	328
25.2.1	Testing the significance of indicator variables	329
25.2.2	Importance of omitting one of the indicators	330
26	Overview of Stata estimation commands	333
26.1	Introduction	333
26.2	Linear regression with simple error structures	334
26.3	ANOVA, ANCOVA, MANOVA, and MANCOVA	335
26.4	Generalized linear models	336
26.5	Binary-outcome qualitative dependent variable models	336
26.6	Conditional logistic regression	338
26.7	Multiple-outcome qualitative dependent variable models	338
26.8	Simple count dependent variable models	339
26.9	Linear regression with heteroskedastic errors	339
26.10	Stochastic frontier models	340
26.11	Linear regression with systems of equations (correlated errors)	340
26.12	Models with endogenous sample selection	341

26.13	Models with time-series data	341
26.14	Panel-data models	342
26.14.1	Linear regression with panel data	342
26.14.2	Censored linear regression with panel data	343
26.14.3	Generalized linear models with panel data	343
26.14.4	Qualitative dependent-variable models with panel data	344
26.14.5	Count dependent-variable models with panel data	344
26.14.6	Random-coefficient models with panel data	344
26.15	Survival-time (failure-time) models	344
26.16	Survey data	345
26.17	Multivariate and cluster analysis	345
26.18	Pharmacokinetic data	346
26.19	Specification search tools	346
26.20	Obtaining new estimation commands	347
26.21	Reference	347
27	Commands everyone should know	349
27.1	Forty-one commands	349
27.2	The by construct	350
28	Using the Internet to keep up to date	351
28.1	Overview	351
28.2	Sharing datasets (and other files)	352
28.3	Official updates	352
28.3.1	Example	353
28.3.2	Updating ado-files	354
28.3.3	Frequently asked questions about updating the ado-files	355
28.3.4	Updating the executable	355
28.3.5	Frequently asked questions about updating the executable	356
28.3.6	Updating both ado-files and the executable	356
28.4	Downloading and managing additions by users	357
28.4.1	Downloading files	357
28.4.2	Managing files	359
28.4.3	Finding files to download	359
28.5	Making your own download site	360