

Subject index

A

abbreviations 41
ado-files 296–302
ado-path 6
adoupdate 23
analysis of variance 136
append 91
archiving data 116–117
arithmetic operators 81
ASCII data 68–69
audit trail 99–100
axes (graphs) 241–245
axis labels (graphs) 241–243
axis ticks (graphs) 241–243

B

backing up data 114–115
bar graphs 257–261
binreg 175
bkup 114
Bland–Altman plot 209–211
box-and-whisker plot 255–256
browse 126
by-graphs 273
by: (prefix) 39

C

c() (system settings and constants) ... 292
capture 305
case–control data 149–154
 matched 152–154, 176
categorical predictors 160–170
cc 149–151
cci 142
cd 64–65
centile 131
chi-squared test 127

ci 139
cii 140
classification of diseases 227
clogit 176
clustered observations 176–178
codebook 102–103
codebook (command) 123
cohort data without censorings ... 145–147
collapse 96
colors (graphs) 251
combining files 91–94
combining graphs 274–275
Command window 9
comments in do-files and ado-files 40
comparing measurements 209–212
complex strings 58–59
compress 60
conditional logistic regression 176
confidence intervals 139–140
confidentiality 115–116
contract 94–95
copying graphs to documents 285
copying tables to documents 18
correcting errors 108–109
count 129
Cox regression 186–197
 checking proportionality ... 190–191
 stratification 189–190
 time-varying coefficients ... 193–197
 time-varying covariates 193–197
creturn list 292
cs 145–147
csi 142
cumul 133
cut() (egen function) 84–85

D

Data Editor 12–13
 data entry 66–68, 103–104
 data protection 115–116
 dataset label 73
 date formats 53
 date variables 53–54
 db 14
 debugging programs 305–307
 decimal periods and commas 46
 #delimit 40
 describe 122
 destring 57
 diagnostic tests 216–222
 diagt 218–219
 dialogs 14, 275–280
 direct standardization 205–207
 disease classification 227
 display 143
 do 15
 do-file 14–15
 Do-file Editor 13
 dotplot 256
 drop 89
 dropped-line plots 269–270
 dstdize 207

E

e() (returned results) 290–291
 egen 83–85
 encode 57
 encryption 116
 entering data 66–68, 103–104
 EpiData (program for entering data) 68
 epitab (family of commands) 145
 ereturn list 290–291
 error correction 108–109
 error finding 107–108
 error messages 26–27
 estat gof 175
 estat phtest 191
 Excel (Microsoft) 70–71
 exercises 313–324
 exit 7
 exlogistic 176

expand 95
 explicit subscripting 88
 extended functions 82–85

F

factor variables 160–171
 FAQs 25
 fdasave 71
 fdause 71
 file path 64–65
 file types 29–30
 filenames 100–101
 findit 24
 Fisher's exact test 127
 folder structure (Windows) 325–327
 foreach 302–304
 format 45–46
 formats
 date 53
 numeric 45–46
 string 56
 forvalues 304
 function plots 270–271
 functions
 extended 82–85
 mathematical 81–82
 statistical 82
 fvset 161

G

generate 79–80
 gold standard 216
 graph area 232
 graph bar 257–261
 graph box 255–256
 graph combine 274–275
 graph command syntax 233–236
 graph display 285
 Graph Editor 280–284
 graph export 285
 graph matrix 271–272
 graph options 234, 241–250
 Graph Recorder 284
 graph save 284
 graph schemes 239–241

graph size 236–238
 graph use 284
 graphs 231–286
 grid lines (graphs) 241–242
 group() (egen function) 84
 groups 130
 gsort 90

H

help 22
 histogram 254–255
 Hosmer–Lemeshow test 175

I

icd9 227
 if (command) 302
 if (qualifier) 36–37
 immediate commands 140–143
 in (qualifier) 37
 incidence-rate ratio 148–149
 indirect standardization 202–204
 infile 69
 infix 69
 input 67
 insheet 68–69
 installing Stata 3
 interactions 164–170
 intraclass correlation coefficient 215
 ir 148
 iri 141–142

K

Kaplan–Meier curve 183–186
 keep 89
 kernel density curve 254–255

L

label data 73
 label define 74
 label list 75
 label values 74
 label variable 73
 labels 73–78
 lag and lead functions 87
 legend (graph option) 246–248

level() (option) 139
 lincom 159–160
 line patterns (graphs) 252
 line plots 264–269
 linear regression 155–157
 list 124–126
 local 294–295
 log 16–18
 log book 105
 log file 15–16
 logical expressions 80
 logical operators 36–37
 logistic 171–175
 logistic regression 171–175
 logit 171–175
 log-scaled axes (graphs) 243–244
 long command lines 40–41
 longitudinal data 227
 lrtest 175, 189
 lvr2plot 158

M

macros 294–295
 global 295
 local 294–295
 manuals 21, 311–312
 marker symbols (graphs) 253
 master do-file 106
 matched case–control data .. 152–154, 176
 matching datasets 92–94
 mathematical functions 81–82
 matrix graphs 271–272
 mcc 152–154
 mdy() (date function) 54
 measurement comparison 209–212
 measurement, reproducibility ... 213–216
 measurement variation 213–214
 memory considerations 59–61
 merge 92–94
 merging datasets 92–94
 meta-analysis 227
 missing (option) 39
 missing values 48–50, 109
 mkdir 64–65
 —more— 10

- multiple axes (graphs) 244–245
- mvdecode 48
- mvencode 49
- N**
- _N 87
- _n 87
- NetCourses 25
- nolabel (option) 39
- nonparametric tests 138
- normal curve 254–255
- notes 76
- Numbering observations 87
- numeric
 - formats 45–46
 - lists 35
 - ranges 36
 - variables 45–56
- numlabel 74
- O**
- ODBC 71
- odbc (command) 71
- oneway 136
- online help 22–24
- operators
 - arithmetic 81
 - logical 36–37
 - relational 36–37
- options 38–39
- order 90
- outfile 70
- output log 15–16
- outsheet 70
- P**
- panel data 227
- pharmacokinetic data 227
- pk (family of commands) 227
- plot area 232
- poisson 199–201
- Poisson regression 199–201
- postestimation, regression 157–160
- power estimation 225–227
- precision 50–52
- predict 158
- predictive values 217–219
- preferences, window 8–13
- prefixes 39
- preserve 305
- programs 296–302
- proportional hazards regression .. 186–197
- protecting data 115–116
- Q**
- qnorm 132
- Q–Q plot 132
- qualifiers 36–37
- quotes
 - double 40
 - single 294
- R**
- r () (returned results) 289–290
- random-effects models 176
- random numbers 223
- random sampling 90, 223–224
- range plots 270
- ranksum 138
- rates 197–201
 - Poisson regression 199–201
 - tabulating 197
- receiver operating characteristic .. 219–222
- recode 85–86
- regress 156–157
- regression
 - postestimation 157–160
- regression analysis 155–178
- relational operators 36–37
- rename 90
- reordering variables 90
- replace 79–80
- reproducibility of measurements 213–216
- reshape 96–97
- reshaping data 94–98
- restore 305
- Results window 9–10
- return list 289–290
- Review window 10

risktable (option).....184
 rnormal() (function).....223
 robust variance estimates.....177–178
 ROC analysis.....219–222
 roctab.....221–222
 runiform() (function).....223
 rvfplot.....158
 rvpplot.....158

S

sample.....90, 223
 sample-size estimation.....225–227
 sampsi.....225–227
 SAS.....71
 save.....63–66
 saved results.....289–294
 saving graphs.....284
 scalar.....295–296
 scalars.....295–296
 scatterplots.....261–263
 schemes (graphs).....239–241
 sdtest.....137–138
 search.....23
 selecting observations.....89
 selecting variables.....89
 sensitivity.....217–219
 separate.....97
 set trace.....306
 simulations.....224
 single quotes.....294
 slist.....125
 SMR.....202–204
 sort.....90
 specificity.....217–219
 spreadsheets.....70–71
 SSC archive.....23
 standard populations.....205–207
 standardization.....202–207
 direct.....205–207
 indirect.....202–204
 standardized mortality ratio.....202–204
 Stata manuals.....21, 311–312
 Statalist.....25
 statistical functions.....82
 statsby: (prefix).....292–294

Stat/Transfer.....70
 stci.....185
 stcox.....186–197
 stcurve.....188
 stir.....149
 storage types.....50–52
 stphplot.....190
 stpower.....226
 stptime.....197
 strate.....204
 stratified analysis.....145–154
 case–control studies.....149–154
 cohort studies.....145–147
 incidence rates.....148–149
 string() (function).....58
 string formats.....56
 string variables.....56–59
 sts graph.....183
 sts list.....185
 sts test.....186
 stset.....181–183, 192
 stsplot.....193
 subscripting.....88
 substr() (string function).....58
 summarize.....123
 survey data.....227
 survival analysis.....179–201
 syntax.....296–297
 syntax diagrams.....31–34
 sysdir.....6
 sysuse.....66

T

tab1.....126
 tab2.....127–128
 tabi.....141
 table.....129, 135
 tabm.....130
 tabodds.....151–152
 tabstat.....134–135
 tabulate.....126
 technical support.....25
 testparm.....163–164
 texp() (option).....195–197
 text editors.....18

text (graphs) 246–250
 time-series data 227
 time-to-event data 179–201
 time variables 54–56
 title (graphs) 246
 tostring 58
 transformations 133
 translation programs 70
 ttest 136–137
 ttesti 142
 tvf() (option) 195–197
 twoway connected 265–267
 twoway dropline 269
 twoway function 270–271
 twoway line 264–265
 twoway rcap 267
 twoway scatter 261–263

U

update 4–5
 updating Stata 4–5
 use 63–66

V

value labels 74
 variable labels 73
 variable lists 35
 variable names 100–101
 variables 45–61
 date and time 53–56
 numeric 45–56
 string 56–59
 Variables Manager 11
 Variables window 10
 variation of measurements 213–214
 varlist 35
 vce() (option) 177–178
 version control 42
 view 18
 Viewer window 11–12

W

webuse 66
 weights 38

window preferences 8–13
 Windows (Microsoft) 325–329

X

xi 170–171
 XML format 70–71
 xmlsave 70
 xmluse 71
 xpose 98