

Subject index

A

abbreviations 42
adjust 211, 274
ado-files 291–297
ado-path 5
adoupdate 23
analysis of variance 120
aorder 93
append 93–94
archiving data 323–324
arithmetic operators 80
ASCII data 66–68
audit trail 307–308
axes (graphs) 140–146
axis labels (graphs) 140–143
axis ticks (graphs) 140–143

B

backing up data 322–323
baplot 265
bar graphs 159–163
binomial regression 193–194
binreg 193–194
bkup 322
Bland–Altman plot 265–266
box-and-whisker plot 157–158
browse 106
by-graphs 175–176
by: (prefix) 40

C

c() (returned settings and constants) .. 286
capture 302
case–control data 194–204
 matched 202–204, 224–226
categorical predictors 212–215
cc 194–197

cci 126
cd 62–63
chi-squared test 108
ci 123–124
cii 124
classification of diseases 281
clogit 224–226
clustered observations 226–228
cmdlog 18
codebook 308–310
codebook (command) 103
cohort data without censorings ... 191–194
collapse 98
colors (graphs) 152
combining files 93–96
combining graphs 176–177
command log 18
Command window 8
comments in do-files and ado-files 41
comparing measurements 263–267
competing events (Cox regression) .. 257–
 258
complex strings 56–57
compress 58
conditional logistic regression ... 224–226
confidence intervals 123–124
confidentiality 325–326
continue 302
contract 96–98
copying graphs to documents 190
copying tables to documents 17
correcting errors 316
correlate 261
correlation coefficient 260–261
Cox regression 238–244
creturn list 286
cs 191–194

csi 126
 cut () (egen function) 83–84

D

Data Browser 12
 Data Editor 12
 data entry 64–66, 312–314
 data protection 325–326
 dataset label 71
 date formats 50–51
 date variables 50–52
 db 13
 debugging programs 296–297
 decimal periods and commas 18, 66
 delayed entry (Cox regression) ... 241–244
 #delimit 41
 describe 102
 destring 55
 diagnostic tests 267–275
 combining information 272–275
 diagnostics, linear regression 207–209
 diagt 269–270
 dialogs 13, 177–185
 direct standardization 253–256
 disease classification 281
 display 127
 do 14–15
 do-file 13–14
 Do-file Editor 12
 dotplot (graph command) 158–159
 double quotes 40–41
 drop 91–92
 dropped-line plots 171–172
 dstdize 254–256
 dummy variables 212–215
 duplicates 96

E

e () (returned results) 284–285
 effect modification 215–217
 egen 82–84
 egenmore (egen functions) 84
 encode 55
 encryption 326
 entering data 64–66, 312–314

EpiData (program for entering data) ... 66
 epitab (family of commands) 191
 ereturn list 284–285
 error correction 316
 error finding 315–316
 error messages 25–27
 estat gof 222–223
 estat phtest 240
 Excel (Microsoft) 68–69
 exercises 331–342
 exit 6
 exlogistic 223, 225
 expand 97–98
 explicit subscripting 87–89
 expoisson 247
 extended functions 81–84

F

FAQs 24
 fdasave 70
 fdause 70
 file path 62–63
 file types 29–30
 findit 23–24
 Fisher's exact test 108
 folder structure (Windows) 343–345
 foreach 298–299
 format 45–46
 formats
 date 50–51
 numeric 45–46
 string 54
 forvalues 299–302
 function plots 173–174
 functions
 extended 81–84
 mathematical 80–81
 statistical 81

G

generalized linear models 224
 generate 77–79
 gladder 115–116
 gold standard 267

- goodness-of-fit test (logistic regression) . . .
 222–223
- graph area 130
- graph bar (graph command) . . . 159–163
- graph box (graph command) . . . 157–158
- graph combine (graph command) . . 176–
 177
- graph command syntax 131–134
- graph display (graph command) . . . 190
- Graph Editor 185–189
- graph export (graph command) . . . 190
- graph matrix (graph command) . . . 174
- graph options 132–133, 140–151
- Graph Recorder 188
- graph save (graph command) 189
- graph schemes 137–140
- graph size 134–137
- graph use (graph command) 189
- graphs 129–190
- grid lines (graphs) 140–141
- group() (egen function) 83
- groups 112
- gsort 93
- H**
- help 22
- histogram (graph command) . . . 155–157
- I**
- icd9 281
- if (command) 297
- if (qualifier) 36–38
- immediate commands 125–127
- in (qualifier) 38
- incidence rates 229–231
- incidence-rate ratio 231–233, 246–247
- indicator variables 212–215
- indirect standardization 248–253
- infile 67–68
- infix 68
- input 65
- insheet 67
- installing Stata 1
- interaction 215–217
- interrater agreement 259–260
- intraclass correlation coefficient 263
- ir 232
- iri 125–126
- K**
- kap 259–260
- Kaplan–Meier survivor function 235–
 237
- kappa statistic 259–260
- keep 91–92
- kernel density curve 156–157
- L**
- label data 71
- label define 72
- label list 73
- label values 72
- label variable 71
- labels 71–76
- ladder 115–116
- legend (graph option) 147–149
- level() (option) 123
- life table survival analysis 233–234
- lincom 220–223
- line patterns (graphs) 153
- line plots 165–171
- linear regression 205–218
 diagnostics 207–209
- list 104–106
- log 15
- log book 310–312
- log file 15–16
- logical expressions 78–79
- logical operators 36–38
- logistic (command) 218–223
- logistic regression 218–223
- logit 218–220
- log-rank test 237–238
- log-scaled axes (graphs) 143–145
- loneway 263
- long command lines 41–42
- lroc 273–274
- ltable 233–234

M

macros 288–290
 manuals 21, 327–328
 marker symbols (graphs) 154
 master do-file 312
 matched case–control data 202–204,
 224–226
 matching datasets 94–96
 mathematical functions 80–81
 matrix graphs 174
 mcc 202–204
 mean 117
 measurement comparison 263–267
 measurement, reproducibility 259–263
 measurement variation 260–263
 memory considerations 57–59
 merge 94–96
 merging datasets 94–96
 meta-analysis 281
 mhodds 199–201
 missing (option) 40
 missing values 47–48, 317
 mkdir 62–63
 more 9
 mortality rates 229–231
 multinomial logistic regression 224
 multiple axes (graphs) 145–146
 multiple-failure data (Cox regression)
 257
 mvdecode 47
 mvencode 47

N

_N 87
 _n 87
 nestreg 218
 NetCourses 24
 noLabel (option) 40
 nonparametric tests 122–123
 normal curve 156–157
 notes 74
 NoteTab Light 17
 numeric
 formats 45–46
 lists 35–36

numeric, *continued*

 ranges 36
 variables 45–53
 numLabel 72

O

ODBC 70
 odbccommand 70
 oneway 120
 online help 21–24
 operators
 arithmetic 80
 logical 36–38
 relational 36–38
 options 39–40
 graph 132–133
 order 92–93
 ordered logit estimation 224
 outfile 69
 output log 15–16
 outsheet 69

P

panel data 228
 parallel lists 300–301
 pharmacokinetic data 281
 pk (family of commands) 281
 plot area 130
 poisson (command) 246–247
 Poisson regression 246–247
 polytomous logistic regression 224
 postestimation commands 283
 power estimation 279–281
 precision 48–50
 predict
 after linear regression 207–208
 after logistic regression 221–222
 predictive values 268–270
 preferences, window 7–12
 preserve 91
 profile.do 16, 302–303
 programs 291–297
 proportion 110
 proportional hazards regression .. 238–241
 protecting data 325–326

Q

qnorm 114
 Q–Q plot 114
 qualifiers 36–38
 questionnaires 308
 quotes
 double 40–41
 single 288–289

R

r() (returned results) 283–284
 random-effects models 228
 random numbers 277
 random sampling 92, 277–278
 range plots 171
 ranksum 123
 receiver operating characteristic 270–272
 recode 84–86
 regress 205–218
 regression analysis 205–228
 regression diagnostics 207–209
 relational operators 36–38
 rename 92
 reordering variables 92–93
 replace 77–79
 reproducibility of measurements 259–263
 reshape 98–99
 reshaping data 96–100
 residual-versus-fitted plot 208–209
 restore 91
 Results window 9–10
 return list 283–284
 Review window 10
 risktable (option) 237
 robust variance estimates 226–228
 ROC analysis 270–272
 roccomp 272
 roctab 270–271
 rvfplot 208–209

S

sample 92, 277
 sample-size estimation 279–281
 sampsi 279–281
 SAS 70

save 61–64
 saved results 283–288
 saving graphs 189
 scalars 290
 scatterplots 163–165
 schemes (graphs) 137–140
 sdpair 266–267
 sdtest 121–122
 search 22–23
 selecting observations 92
 selecting variables 91
 sensitivity 268–270
 separate 100
 set trace 296
 shell 302
 simulations 278–279
 single quotes 288–289
 slist 106
 SMCL format 16–17
 SMR 248–253
 sort 93
 specificity 268–270
 spreadsheets 68–70
 SSC archive 23
 standard populations 253–256
 standardization 247–256
 direct 253–256
 indirect 248–253
 standardized mortality ratio 248–253
 Stata manuals 21, 327–328
 Statalist 25
 statistical functions 81
 statsby: (prefix) 286–288
 Stat/Transfer 68
 stcox 238–244
 stcurve 240–241
 stir 232
 storage types 48–50
 stphplot 239–240
 stptime 230–231
 strate 250
 stratified analysis 191–204, 231–233
 case–control studies 194–204
 cohort studies 191–194
 incidence rates 231–233

string formats 54
string variables 54–57
sts graph 236
sts list 235
sts test 237
stset 234–235
stsplot 244–246
subscripting 87–89
summarize 104
survey data 227–228, 282
survival analysis 233–238
svy: (prefix) 227–228
svyset 227–228
swilk 114
syntax 291
syntax diagrams 31–35
sysdir 5
sysuse 64

T

tab1 107
tab2 108–109
tabi 125
table 111, 119–120
tabm 111–112
tabodds 197–199
tabstat 117–119
tabulate 107
tabulate, summarize() 119
technical support 25
test 214–215
testparm 214–215
text editors 17
text (graphs) 147–151
time-dependent covariates (Cox regression)
..... 256
time-series data 282
time variables 52–53
title (graphs) 147
tostring 55
transformations 115–116
translation programs 68
ttest 120–121
ttesti 127

twoway connected (graph command) ...
..... 167–169
twoway dropline (graph command) ...
..... 172
twoway function (graph command) ...
..... 173–174
twoway line (graph command) ... 165–
167
twoway rcap (graph command) 169
twoway scatter (graph command)
..... 163–165

U

update 2–5
updating Stata 2–5
use 61–64

V

value labels 72
variable labels 71
variable lists 35
variables 45–59
 dates and times 50–53
 numeric 45–53
 string 54–57
Variables window 11
variation of measurements 260–263
version control 43
view 17
Viewer window 11–12

W

webuse 64
weights 38–39
which 295
window preferences 7–12
Windows (Microsoft) 343–347
WinZip 322

X

xi: (prefix) 214–216
XML format 68–69
xmlsave 68
xmluse 69
xpose 100

<code>xt</code> (family of commands)	228
<code>xtlogit</code>	228