

Subject index

Symbols

β weights 272
 η^2 222
 ϕ 131
* comment 83
/* and */ comment 83

A

acquiring datasets 449–450
add, label define option 2
agreement, intraclass correlation ... 255
alpha reliability 369
ameans command 95
analysis of covariance see ANCOVA
analysis of variance see ANOVA
ANCOVA 231–242
ANOVA,
 degrees of freedom 221
 equal-variance test 221
 one-way 215–216
 one-way power analysis ... 257–260
 power analysis 257–264
 repeated-measures 249
 repeated-measures power
 analysis 262–264
 two-way 243–249
 two-way power analysis ... 260–262
ANOVA assumptions 216
ATS at UCLA 202

B

bar chart 105, 344
bar graph of means 227
Bartlett test of equal variances ... 221
beta weights 209, 272
 limitation 308
binary variables 292–293

binscatter command ... 196–200, 313
block regression 291–299
blog see Stata Blog
Bonferroni multiple-comparison test ...
 206, 219
bookstore, Stata 446
bootstrap estimation of standard errors
 283
bootstrap regression 278
Boston College Stata site 444
box plot 115, 230–231

C

casewise deletion 203
categorical covariates 233, 237
categorical predictors, regression ... 293
categorical variable 299–301
cause and effect 194
center command 305
centering 317–318
chi-squared
 table 127
 test 125, 346
chitable command 127–129
clear command 29
cls command 16
codebook command 43–45, 53–54,
 136
codebook example 25–28
coding system 24–28
coefficient of variation 114
Cohen's d 169–170, 174–182
Cohen's f 258
collinearity 287
command structure 76
Command window 8, 10

- confidence interval
 regression line..... 209
 slope..... 209
constant 209
continuous covariates..... 233–235
continuous variable 313
conventions used in the book..... 1
copying, HTML format..... 18, 87
copying results to word processor... 18,
 86
correlate command..... 203–204
correlation,
 interpreting 200
 limitation..... 308
 multiple comparison..... 206
correlation ratio..... 222, 241–242
count outcome variables 337
Cramér’s *V*, measure of association ...
 131
creating value labels..... 55–58
criterion-related validity 376
cross-tabulation 122
curve fitting..... 311–317
- D**
data,
 long format 160, 217
 wide format 158, 217
Data Editor..... 29–33, 40–41
dataset contents 43
dataset,
 acquisition 449–450
 c10interaction..... 301, 326
 c11barchart..... 344
 cancer 10, 11
 census..... 325
 censusfv 19
 chapter6_aspirin 134
 chapter13_missing..... 399
 chores..... 171
 create 21–23
 depression..... 213
 descriptive_gss..... 98, 117, 118
 divorce 331
 download xxix, 449–450
 dataset, *continued*
 environ 334
 firstsurvey..... 43
 firstsurvey_chapter4 78, 83,
 90
 flourishing_bmi 414, 415
 gss2002 and 2006_chapter12 ...
 392
 gss2002_chapter10..... 440
 gss2002_chapter6 146, 147
 gss2002_chapter7.. 155, 161, 162,
 164, 172, 187
 gss2002_chapter8..... 213
 gss2002_chapter9..... 265
 gss2002_chapter10..... 325, 326
 gss2002_chapter11..... 359
 gss2006_chapter6.. 122, 135, 140,
 143, 146
 gss2006_chapter6_10percent....
 143
 gss2006_chapter8 .. 190, 212, 213
 gss2006_chapter8_selected.. 206
 gss2006_chapter9..... 225, 232
 gss2006_chapter9_2way 243
 gss2006_chapter12..... 362, 369
 gss2006_chapter12_selected....
 383
 intraclass..... 256
 kappa1..... 373
 kuder-richardson..... 371
 long..... 160
 nlsw88..... 327
 nlsy97_chapter7 183, 187
 nlsy97_chapter11..... 337, 346
 nlsy97_selected_variables.....
 266, 291
 ops2004..... 268, 412
 partyid..... 218, 228, 266
 positive 73
 regsmpl..... 311, 399, 440
 relate xxix, 50, 73
 relate_small..... xxix
 retest..... 367
 severity 359
 spearman 211, 213

- dataset, *continued*
- wide 157
 - wide9 250
- degrees of freedom, 127
- ANOVA 221
 - one-sample *t* test 164
- dependent *t* test 171
- dependent variable 124
- describe command 10, 45–46, 51
- dfbeta command 286
- dialog box,
- alpha 369
 - anova 243
 - codebook 43–45
 - correlate 203
 - describe 45–46
 - egen 68–70
 - generate 63, 65–66
 - graph bar 141–142, 227
 - graph box 230
 - graph hbox 115
 - graph pie 102
 - histogram 14–15, 105–107
 - kwallis 228
 - lfit 210
 - logistic 338–339
 - margins 236, 238–239, 245
 - nestreg 319
 - oneway 218–219
 - open 64
 - pcorr 273–274
 - power 177–178
 - prtest 155–161
 - recode 59–60, 291
 - regress 207, 269
 - rename 55
 - rvplot 280
 - scatter 191–195
 - sdtest 170
 - sktest 110
 - Submit vs. OK 17
 - summarize 11
 - tab1 98
 - tabi 138–139
 - table 140
- dialog box, *continued*
- tabstat 114
 - tabulate 66–67, 122–126, 136
 - ttest 162–163, 166–169, 172
- dictionary file 50
- difference of means test 164
- difference of proportions test 157
- discontinuity graph 198–200
- display command 169–170, 342, 431–432
- do-file,
- continuation line 165
 - introduction 6
- Do-file Editor 81–85
- download datasets 449–450
- drop command 71
- dummy variables 292–293
- E**
- effect size 169–170, 222, 241, 273
- η^2 222, 241–242
- egen command 63, 68–70, 363
- egen count command 363
- egen rowmean command 70, 363
- egen rowmiss command 69
- egenmore command 69
- entering data 29–33
- equal variance, Bartlett test 221
- esize command 169–170
- estat esize command 241–242
- estat vif postestimation command ..
..... 287–288
- estimates store command 346
- Excel
- exporting data 47
 - importing data 47
- exit Stata 18, 43
- exponentiation 342
- external validity 203
- F**
- F* ratio 221
- F* test of unequal variances 170
- Facebook see Stata on Facebook

factor analysis, 378
 commonality 381
 eigenvalue 381, 386
 exploratory factor analysis 379
 extraction 381
 factor score 381, 390
 loading 381
 oblique rotation 381, 388
 orthogonal rotation 381, 387
 PCF 380
 PF 379
 postestimation 382
 principal component analysis. 380
 principal-component factor analysis 380
 promax 388
 rotation 381
 scree plot 381, 386
 simple structure 381
 varimax 387

factor variable 299–301

fonts, fixed 86

format,
 numeric 31
 string 32

fre command 100–102

frequency distributions 99

ftable command 221

full mediation 435

G

gamma, measure of association 136

generalized linear model 427–428

generalized structural equation modeling 413–441

generate command 63–68

geometric mean 95

glm command 427–428

Goodman and Kruskal’s gamma, measure of association 136

GradPlan 446

graph,
 alternative scattergram ... 196–200
 bar chart 105, 142, 344
 box plot 115

graph, *continued*
 collinearity 287
 discontinuity 198–200
 hanging rootogram 275
 heteroskedasticity 279
 histogram 108, 275
 medians 231
 overlay two-way showing
 interaction effects 303
 pie chart 102
 residual versus fitted 280
 scattergram 190–196

graph bar command 141–142, 227, 344

graph box command 230–231

Graph Editor 104–105

graphics book 447

gsem command 413–441
 logistic regression 425–428

GUI interface, **Edit** ▷ **Preferences** 8

H

harmonic mean 95

help 6
 video 19
 web-based 19

help, listcoef option 342–343

help label command 43

heteroskedasticity 279

hierarchical regression 291–299, 319–321, 353–355

histogram 108

histogram command 13–17, 275

HTML format 87

I

ice command 399

imputation see multiple imputation

increment in R^2 273

independent variable 124

indicator variables 292–293

interaction term 301–304

interactive table 138

intercept 209

interquartile range 114

interval-level variables 94
 intraclass correlation 255, 256

J

`jitter()`, `scatter` option 193

K

`kappa` 372
 `kappa, weighted` 374
 `kappa with three raters` 374
 `keep` command 71
 Kendall's tau, measure of association ..
 136
 Kruskal–Wallis test, ANOVA alternative
 228
 Kuder–Richardson coefficient of
 reliability 371
 kurtosis 97, 110, 277–278

L

`label variable` command 57
 labeling values 33
 labeling variables 23
 likelihood-ratio chi-squared test... 346–
 347
 limitations of Stata 450
 `list` command 79–80, 218
 `list` option, `nolabel` 218
 `listcoef` command 342–344
 listwise deletion 203
 log, `.smcl` extension 87
 log files 87
 log files and graphs 88
 `logistic` command 338–341
 logistic regression 329–360, 433
 bar chart 344
 exponentiation 342
 hypothesis tests 346
 interpreting odds ratio 341
 likelihood-ratio chi-squared test...
 346
 logits 336
 McFadden pseudo- R^2 340
 nested 353–355
 nonlinear 332

logistic regression, *continued*

 odds ratio 334
 percentage change 341
 pseudo- R^2 340
 S-curve 332
 vs. OLS regression 333
 Wald chi-squared test 346

`logit` command 332, 338–341,
 426–428

`logits` 336
 long format 160, 167, 217, 251–252
 `lrdrop1` command 346–347
 `lrtest` command 346

M

MAR 395–397
 `margins` command .. 236, 238–240, 306,
 350–352
 `marginsplot` command 246–249,
 306–307
 maximum number of variables 450
 MCAR 395–396
 McFadden pseudo- R^2 340
 mean squares 221
 measure of association,
 η^2 222, 241–242
 ϕ 131
 odds ratio 133
 V 131
 `median` command 185
 median, graph box plot 231
 mediation 434–438
 correlated residuals 439
 full 435
 partial 435
 menu, open 64
 `mi estimate` command .. 399, 405–406
 `mi impute chained` command 399
 `mi impute` command 403–404
 `mi impute mvn` command 399, 404
 `mi register` command 403
 `mi set` command 403–404
 `mibeta` command 406–408
 `mim` command 399
 missing, count 363

- missing values .. 27, 54, 65, 77, 393–412
 - types 53, 54
- misstable** command 400–401
- more** command 4, 44
- multicollinearity 287
- multiple comparison,
 - Bonferroni 219
 - Scheffé 219
 - Šidák 219
- multiple comparison and correlation...
 - 206
- multiple correlation 270
- multiple imputation 393–412
- multiple regression command .. 269–273
- multiple regression diagnostics 283
- multiple regression with interaction
 - term 301–307
- multiple regression,
 - block 291–299
 - categorical predictors 293
 - dummy variables 293
 - hierarchical 291–299
 - indicator variables 293
 - influential case 283
 - nested 291–299
 - outlier 283
 - residual 279
 - weighted data 289
- mvdecode** command 54, 65, 165

- N**
- naming variables 25
- nested regression 291–299, 319–321,
 - 353–355
- nestreg** command .. 297–299, 319–321,
 - 353–355
- NetCourses 449
- nolabel**, **list** option 218
- nominal-level variables 94
- nonlinear 332
- nonlinear regression 308–321
- nonparametric ANOVA alternatives
 - 228
- nonparametric tests 183
 - Mann–Whitney 183

- nonparametric tests, *continued*
 - median 184
 - rank sum 183
- normally distributed residuals 278
- numlabel** command 80, 102

- O**
- odds ratio 133, 334
 - interpretation 341
 - percentage change 341
- OLS regression vs. logistic regression...
 - 333
- omega2** command 241
- one-sample *t* test 162
- one-way ANOVA 215–216
- oneway** option
 - bonferroni** 219
 - scheffe** 219
 - sidak** 219
- open
 - existing dataset 9–11
 - Stata-installed dataset 10
- optifact** command 444
- option,
 - add, label define** 2
 - help, listcoef** 342–343
 - jitter(), scatter** 193
 - percent, listcoef** 343
- ordinal-level variables 94
- ordinary least-squares regression .. 413–
 - 415
- outlier 283

- P**
- paired *t* test 171
- part correlation 273
- partial mediation 435
- pasting, 87
 - reformatting 86
- pasting results to word processor, ... 86
 - formatting 86
- path analysis 434–438
- pcorr** command 273, 295
- Pearson’s chi-squared 136
- percent, listcoef** option 343

- pie chart 102
- plot a confidence interval 210
- Poisson regression 337
- postestimation command,
 estat vif 287–288
 margins 306
 marginsplot 306–307
 predict 284–286, 302
 test 296–297
- power, unequal standard deviations ...
 178
- power analysis 173–182
 one-way ANOVA 257–260
 repeated-measures ANOVA ... 262–
 264
 two-way ANOVA 260–262
- power command 177–182
- power oneway command 257–260
- power repeated command ... 262–264
- power twomeans command ... 179–182
- power twoway command 260–262
- powerreg command 321–324
- predict postestimation command
 284–286, 302
- predict postregression 284–286
- predictive validity 376
- probability tables 127
- product term 304
- project outline 50
- prophecy formula 368
- proportions,
 one-sample test 155
 two-sample test 157
- prtest command 155–161
- pseudo- R^2 340
- pwcorr command 204–206
- pweights 290
- pwmean command 223–225
- Q**
- quadratic model,
 centering 317–318
 examining 319–321
 fitting 311–317
- qualifier,
 if with missing values 77
 in 78
- R**
- R^2 169–170, 270
 change 273
- random sample 191
 how to draw 153
- random sampling 151
- randomization 151
 alternative to 232
 how to perform 153
- ranksum command 183
- recode command 59–61, 291
- recoding 165
 ranges 59
- regress command .. 207–210, 269–270,
 293–296, 301–304
- regression,
 block 291–299
 bootstrap 278
 categorical predictors 293
 dummy variables 293
 hierarchical 291–299
 indicator variables 293
 influential case 283
 nested 291–299
 outlier 283
 residual 279
 robust 278
 weighted data 289
- regression diagnostics 283
- regression line, plotting 195
- regression with interaction term .. 301–
 307
- reliability,
 alpha 369
 equivalent forms 368
 kappa 372
 kappa with three raters 374
 Kuder–Richardson coefficient of re-
 liability 371
 prophecy formula 368
 split-half 368

- reliability, *continued*
 - test–retest 367
 - weighted kappa 374
 - rename** command 55, 251
 - rename variable 251
 - repeated-measures *t* test 171
 - repeated-measures ANOVA 249
 - repeated-measures design 262–264
 - replace** command 63, 165
 - reshape** command 251–252
 - reshape wide to long format... 251–252
 - residual 278, 279
 - Results window 8
 - clear 16
 - more** 4
 - scroll size 4
 - reverse coding 363
 - reverse-code variables 58
 - Review window 8
 - robust estimation of standard errors...
 - 283
 - robust regression 278
 - root mean squared error 209
- S**
- sample, draw 191
 - sample** command 143, 191
 - save** command 41–42
 - saveold** command 42
 - scale construction 362
 - reverse coding 363
 - scale creation 68
 - scatter** command 191–197
 - scattergram 190–200
 - scattergram with confidence interval ..
 - 210
 - scattergram with **jitter()** option.. 193
 - scattergram, alternative 196–200
 - Scheffé, multiple comparison 219
 - schemes 15
 - scientific notation 272
 - scree plot 385–386
 - S*-curve 332
 - sdtest** command 170–171
 - search** command 6, 444
 - seed for starting a random sample.. 191
 - sem** advantages 413
 - SEM Builder 415–422
 - restore default settings 423
 - sem** command 413–441
 - sembuilder** command 415
 - semipartial correlation 273, 295
 - set more off** command 4
 - set seed** command 143, 152, 190
 - Šidák, multiple comparison 219
 - significance,
 - statistical 202
 - substantive 202
 - skewness 97, 110, 277–278
 - sktest** command 277–278
 - slope 209
 - SMCL 87
 - spearman** command 211
 - Spearman’s rho 211
 - split-half reliability 368
 - ssc** command 444
 - standardized beta coefficient 209
 - standardized beta weights 272
 - standardized regression coefficients ...
 - 272
 - Stat/Transfer 450
 - Stata Blog 445
 - Stata, limitations 450
 - Stata on Facebook 445
 - Stata on Twitter 445
 - Stata bookstore 446
 - Stata code for textbook examples.. 444
 - Stata/IC limitations 450
 - Stata Journal* 446
 - Stata listserver 445
 - Stata Markup and Control Language..
 - see SMCL
 - Stata NetCourses 449
 - Stata Portal, UCLA 202, 444, 447
 - Stata screen 7
 - Stata tutorial 449
 - Statalist 445
 - statistical significance versus
 - substantive significance ... 202
 - structural equation modeling.. 413–441

sum of squares.....221
summarize command..11–13, 110, 207,
 276–277
summarize(), **tabulate** option...243–
 249
 summary of data.....27
sunflower command.....193
sysuse command.....10

T

t test,
 one-sample.....162
 power analysis.....177–182
 two-sample.....164
tab command.....143
tab1 command.....99, 203
tab2 command.....292
tabdisp command.....252
tabi command.....138–139
 table.....122
 probability.....127
 summary statistics.....229
 table calculator.....138–139
table command.....140–141
tabstat command.....114, 229
tabulate command.....62, 66–
 68, 99, 122–126, 134, 136–138,
 243–249
 tau, measure of association.....136
 test of significance,
 kurtosis.....110
 skewness.....110
test postestimation command....296–
 297
 test–retest reliability.....367
 tests,
 Bonferroni.....206
 chi-squared.....125
 dependent *t* test.....171
 likelihood-ratio chi-squared...346–
 347
 likelihood-ratio chi-squared with
 logistic regression.....346
 logistic regression.....346
 long format for portions.....160

tests, *continued*

 Mann–Whitney.....183
 median.....184
 multiple comparison with correla-
 tions.....206
 nonparametric.....183
 one-sample *t* test.....162
 paired *t* test.....171
 proportions.....155, 157
 rank sum.....183
 repeated-measures *t* test.....171
 skewness and kurtosis....277–278
 two-sample *t* test.....164
 unequal variances.....170
 Wald chi-squared test.....346
 wide format for proportions...158
 z test for proportions.....161
 z test with logistic regression..346
 textbook examples using Stata
 commands.....444
 tolerance.....287–288
 toolbar,
 Stata for Mac.....9
 Stata for Windows.....9
ttest command.....162–169, 172
 tutorial.....449
 Twitter.....see Stata on Twitter
 two-by-two table.....122
 two-way ANOVA.....243–249
twoway command.....195–196, 303

U

UCLA Stata Portal.....19, 202, 444
 reshaping data wide to long...251
 user-written commands.....444

V

validity,
 criterion related.....376
 external.....203
 predictive.....376
 value labels.....23, 33–40, 55–58, 80
 variable labels.....23
 variable name.....22, 25
 reserved.....22

Variables Manager 33–40, 55–57, 71
Variables window 8
variance inflation factor 287–288
VIF see variance inflation factor

W

Wald chi-squared test 346
web search 444
weighted data 289–291
weighted kappa 374
weights, `pweights` 290
wide format 158, 168, 217, 251–252
working directory 9

X

`xtreg` command 256
`xtset` command 253

Z

z test,
 one-sample proportion 155
 two-sample proportion 157
zero-inflated models 337