

Subject index

Symbols

! 87
 $\Delta\beta$ 411–413
 $\Delta\chi^2$ 413
 β see regression, standardized coefficient
* see do-files, comments
+ see operators
== see operators
// see do-files, comments
>>> 86
(factor-variable operator) 310
(factor-variable operator) .. 310, 317, 322, 353, 355
& see operators
^ see operators
~ see operators
| see operators
|| 133
1:1 match 458–459
1:m match 464–467
30_year_averages.dta (dataset) .. 453

A

added-variable plot 334–336
additive index 98
adjusted count R^2 403–404
adjusted R^2 301–302
ado-files
 basics 497–499
 programming 499–515
adopath 499, 500
aggregate see collapse
Akaike’s information criterion 404
Aldrich–Nelson’s p^2 488
_all 53
alphanumeric variables see strings

analwe.dta (dataset) 77
analysis of variance see regression
angle() (axis-label suboption) 159
ANOVA see regression
Anscombe quartet ... 314–315, 324–325, 327
anscombe.dta (dataset) 324
anycount() (egen function) 110
ap.dta (dataset) 462
append 454, 468–470
arithmetic mean see average
arithmetical expressions see expressions
ASCII characters 476–478
association
 negative 282
 positive 282
 weak 282
ATE 263–265
augmented component-plus-residual
 plot 331
autocode() (function) 190
autocorrelation see regression, autocorrelation
average 9, 191, 194–195
average marginal effects .. 397–399, 420
average partial effects see average marginal effects
average treatment effect see ATE
avplots 335
aweight (weight type) .. 75, 77–78, 336
axis,
 labels 137, 157–159
 scales 149–151
 titles 137, 160–161
 transformations 150

B

_b[*name*] 85, 289
 backdoor criterion 319–320
 balanced panel data 462
 balanced repeated replication 235
bands() (**twoway mband** option)
 329–330
bar (graph type) 135–136, 139,
 184–186, 201
 bar charts 184–186, 201
 base category 305
 batch jobs see do-files
 Bayesian information criterion 404
BeautifulSoup (Python module) .. 453
berlin.dta (dataset) 223–226
 Bernoulli's distribution ... see binomial
 distribution
 beta see regression, standardized
 coefficient
 beta distribution 220–221
 between estimator 360
 bias 226
bin() (**histogram** option) 206–207
 binary code 476
 binary variables .. see variables, dummy
 binomial distribution 388–389
 bitmap graphs 170
bkp.dta (dataset) 459, 462
bkpequiv.dta (dataset) 459
 BLUE .. see Gauss–Markov assumptions
 book materials xxiii–xxiv
 bookstore 28
 boolean expressions 60, 63–64
 bootstrap 235
box (graph type) 132, 135–136,
 203–205
 box plots 15, 203–205
boxcox 343
 break see commands, break
browse 434
bs4 (Python module) 453
 by prefix .. 12, 69–71, 112–116, 125, 195
by() (**graph** option) 165–166
by() (**tabstat** option) 196–197
bysort 18, 70

byte (storage type) 128
 byte length 484

C

C 86
c. (factor-variable operator) 310,
 322, 353, 355
 calculator see pocket calculator
caption() (**graph** option) 162–163
capture 44, 496
 categories 176–177
 causality 261–264
cd 3
 center see variables, center
center 39, 300
 central limit theorem 228–230
 ChatGPT see LLM
 chi-squared
 likelihood-ratio 181
 Pearson 181
 classification tables 401–404
clear (option) 23
 clock position 153–154
cmdlog 37–40
 CMYK 142
 CNEF 468
 code-table 476
 coefficient of determination 296
coefplot 348–350
collapse 112, 466, 467
collect 201
column (option) 14, 180
 comma-separated values see
 spreadsheet format
 command + see commands, break
 command line see windows,
 Command
 commands
 abbreviations 7, 51–52
 access previous 7
 break 7
 e-class 81
 end of commands 42–43
 external 499

- commands (*continued*)
 - internal 499
 - long see do-files, line breaks
 - r-class 81–82
 - search 16
- comments see do-files, comments
- communities.dta** (dataset) ... 233, 436
- component-plus-residual plot .. 330–331
- compound quotes 510
- compress** 444, 475
- compute see **generate**
- cond()** (function) 100–101, 510
- conditional-effects plot 396–397
- conditions see **if** qualifier
- confidence intervals 248–249, 351
- confounder 264–265, 267, 319, 341, 359, 360, 369, 372, 373, 375, 376
- connect()** (**scatter** option) .. 144–147
- connected** (plottype) 144–147
- constraint** 428
- contingency table see frequency tables, two-way
- contract** 77
- control variables 262–264
- Cook's *D* 336–340
- cooksd** (**predict** option) 339
- correlation coefficient 282–283
- count R^2 402–404
- counterfactual concept of causality .. see causality
- covariate see variables, independent
- covariate pattern 404
- cprplot** 330–331
- Cramér's *V* 181
- .csv** see spreadsheet format
- Ctrl+Break** see commands, break
- Ctrl+C** see commands, break
- cumulated probability function see probit model
- D**
- D1 (time-series operator) 367
- DAG ... 260–261, 264, 319–320, 372, 373
- Data Editor 448–449
- data matrix 433–434
- data region 137
- data types see storage types
- data-generating process 259–261
- data1.dta** (dataset) 4–5, 25, 32, 45, 46, 48, 81, 93, 102, 115, 131, 230, 231, 234, 236, 237, 240, 249, 268, 328, 391, 433, 451, 454, 475, 488, 501
- data2w.dta** (dataset) 362
- datasets 435–436
 - combine 454, 457–470
 - describe 5–6
 - export 475
 - hierarchical 115, 464–467
 - import 436–440
 - load 4–5
 - nonmachine-readable 448–452
 - panel data 361–362
 - preserve 144
 - rectangular 434
 - reshape 361–364
 - restore 144
 - save 27, 475
 - sort 8–9
 - text files 440–448
- date()** (function) 123–124
- dates
 - combining datasets by 452
 - display format 122–123
 - elapsed 122–123
 - from strings 123–124
 - variables 121–125
- .dct** (extension) 68
- decode** 351
- DEFF see design effect
- DEFT see design effect
- degrees of freedom see **df**
- delete see **erase**
- #delimit** 42–43
- density 205, 207–209
- describe** 2–3, 5–6
- design effect 238
- destring** 117–118
- detail** (option) 13, 81

- df..... 295
- dfbeta..... 333
- diary..... 125–126
- diary.dta (dataset)..... 125
- dictionary..... 445–448
- DID estimator..... 360, 365–371
- difference-in-difference..... see DID estimator
- dir..... 4, 24
- directed acyclic graph..... see DAG
- directory
 - change..... 3
 - contents..... 4, 24
 - working directory.... xxiv, 3–4, 68
- discard..... 499
- discrepancy..... 338, 409
- discrete (histogram option).....
 - 182–183
- display..... 62, 123, 219, 289, 488, 490–491
- distributions
 - describe..... 175–215
 - grouped..... 187–190
- .do (extension)..... 68
- do..... 23
- do-files
 - analyzing..... 45
 - basics..... 22–24
 - comments..... 41–42
 - create..... 45–46
 - editors..... 22–23
 - error messages..... 23–24
 - execute..... 23
 - exit..... 44–45
 - from interactive work..... 35–40
 - line breaks..... 42–43
 - master..... 46–49
 - organization..... 45–49
 - version control..... 43
- doedit..... 22, 36
- dot (graph type).... 135–136, 186–187, 201–203
- dot charts..... 186–187, 201–203
- double (storage type)..... 126, 128
- drop..... 6, 53
- .dta (extension)..... 68
- dummy variables..... see variables, dummy
- Dynamic documents..... 24–26
- dyndoc..... 24–26
- E**
 - e() (stored results)..... 81–86
 - e(b)..... 85
 - e(V)..... 85
 - e-class..... see commands, e-class
 - edit..... 448–449
 - efficiency..... 226–227
 - egen..... 109–110, 369
 - egenmore..... 110
 - elapsed dates..... see dates, elapsed
 - elapsed time..... see time, elapsed
 - EMF..... 170
 - Encapsulated PostScript..... see EPS
 - encode..... 117–118, 466
 - end..... 494
 - end (Python)..... 86
 - endogenous variable..... see variables, dependent
 - Enhanced Metafile..... see EMF
 - Epanechnikov kernel..... 209–210
 - EPS..... 170
 - erase..... 48
 - ereturn list..... 84
 - error messages
 - ignore..... 44
 - invalid syntax..... 11
 - estat classification..... 402
 - estat dwatson..... 344
 - estat gof..... 405
 - estat ic..... 404, 415
 - estimand..... 226–227, 230, 250, 252
 - estimate..... 226
 - estimates restore..... 85
 - estimates store..... 85, 346–347
 - estimator... 79, 226–227, 234, 322, 325, 356, 358, 360, 365, 376
 - eststo..... 347–348
 - esttab..... 346–348
 - etable..... 346–347

Excel files 435–439
 exclusion restriction 372
exit 27
exit (in do-files) 44–45
 exit Stata 27
 exogenous variable see variables,
 independent
exp() (function) 66
expand 77
 expected value 325
 export see datasets, export
 exposure .. 261, 319, 320, 322, 328, 340,
 341, 358, 360, 369, 372, 373,
 376, 416
 expressions 61–67
 extensions 68

F

F test 296–297
 factor variables .. 39, 305–306, 309–311,
 416, 418
 FAQ 16, 28
 fence 203–204
filefilter 443
 filenames 67–69
 first difference model see regression,
 first difference
 Fisher’s exact test 181
fitstat 404
 five-number summary 195, 203
 fixed format 445–448
 fixed-effects model 368–370
float() (function) 129
float (storage type) 128
foreach 71–74, 463, 508–509
forvalues 74–75, 225
 frames 470
fre 178
 free format 443–445
 frequencies
 absolute 177
 conditional 180–181
 relative 177
 frequency tables 14–15
 one-way 177–178

 two-way 178–182
 frequency weights see weights
 frequently asked questions see FAQ
freqwe.dta (dataset) 76
function (plottype) 135–136, 248,
 312
 functions 65–67, 99–102
fweight (weight type) 75–77
fxsize() (graph option) 167–168
fysize() (graph option) 167–168

G

gamma coefficient 181
 Gauss curve ... see normal distribution
 Gauss distribution see normal
 distribution
 Gauss–Markov assumptions 324
generate 19–20, 27, 93–116, 128
generate() (**tabulate** option) 305
 GIF 170
gladder (statistical graph) 137
.gph (extension) 68
graph 131–170
 graph region 137, 148
 graphical causal model see DAG
 Graphics Interchange Format .. see GIF
 graphs
 3-D 137
 combining 166–168
 connecting points 146
 editor .. 138–139, 155–156, 168–169
 elements 137–139
 export 169–170
 multiple 163–168
 overlay 163–165
 print 168–169
 titles 162–163
 types 135–137
 weights 336
 grid lines 151, 159
 grouping
 by quantiles 189
 intervals with arbitrary width .. 190
 intervals with same width
 189–190

grstyle 133
gsample 234
GSOEP xxiii, 5, 103, 105, 236–237,
 240, 359, 455–456, 468

H

help 15–17
help files 516–517
histogram 182–184, 205–207
histogram (plottype) 135–136
 History window .. see windows, History
 homoskedasticity see regression,
 homoskedasticity
 Hosmer–Lemeshow test 405
 Huber/White/sandwich estimator .. 343

I

i. (factor-variable operator) .. 305–306,
 353, 355, 418
if qualifier 10–11, 58–61
imargin() (graph combine option) ...
 167
import see datasets, import
import delimited 441–443
in qualifier 8–9, 57–58
 inference
 causal 217, 259–267, 319–323,
 356–376
 descriptive 217, 229–258
 prediction 356–358
 predictive .. 217–218, 267–277, 323
 statistical 217–267
infile 440, 443–448
infix 440, 443, 448
 influential data points 331–340,
 409–413
inlist() (function) 66–67, 100
input 448–451
 inputting data 448–452
inrange() (function) 67, 100
inspect 176
 instrumental variable regression
 .. see regression, instrumental
 variables
 interaction terms 307–311, 417–420

invnormal() (function) 221
IPython (Python module) 286
ipywidgets (Python module) 286
irecode() (function) 101
iscale() (graph combine option)
 167
 ISO 8859 characters 477–478
iso8859.dta (dataset) 479
 item nonresponse see nonresponse
 iteration block 399–400
ivregress 375–376
weight (weight type) 75

J

jackknife 235
Java 86
jitter() (scatter option) 381
joinby 454
 JPG ... see Joint Photographic Experts
 Group
 JPG 170

K

ka_temp.dta (dataset) 144, 453
kdensity 209–210, 221, 227
kdensity (plottype) 210–211
 Kendall's tau-b 181
 kernel density estimator 207–211
 key variable 459–460

L

label 20, 105, 106
label data 475
 labels,
 and values 107–108
 datasets 475
 display 107–108
 values 20–21, 106
 variables 20, 105–106
labmask 351
 large language model see LLM
legend 137, 161–162
length() (function) 484
levelsof 270, 508–509
 leverage 338, 409

- `lfit` (graph type) 133
- `lfit` (plotype) 163–164, 291
- likelihood 389–390
- likelihood-ratio χ^2 401
- likelihood-ratio test 414–415
- `line` (plotype) 139, 144–147
- linear probability model see regression, LPM
- linear regression see regression
- linearity assumption 328–331, 405–409
- `list` 6–8
- LLM 17–19, 90–91, 173–174
- `ln()` (function) 65
- `local` 83, 365, 489–493
- local macro see macro
- local mean regression 406
- localization 482–483
- `loess` see LOWESS
- `log()` (function) 95–96
- `log` (scale suboption) 150
- log files
 - finish recording 44
 - interrupt recording 39
 - log commands 37–40
 - SMCL 43–44
 - start recording 43–44
- logarithm 95–96
- logical expressions see expressions
- `logistic` 395
- logistic regression,
 - coefficients 393–397
 - command 392–393
 - dependent variable 384–388
 - diagnostic 405–413
 - estimation 388–391
 - fit 400–405
- `logit` 392–393
- logit model see logistic regression
- logits 386–388
- loops
 - `foreach` 71–74
 - `forvalues` 74–75
- `lower()` (function) 118
- LOWESS 331, 406–407
- `lowess` 406–407
- `lowess` (plotype) 406–407
- LPM see regression, LPM
- `lrtest` 414
- M**
 - `m:1 match` see `1:m match`
 - machine learning 269–277, 323
 - macro
 - extended macro functions 511–512
 - local 83–84, 489–493
 - manuals xxv
 - `margin()` (graph option) 148–149
 - marginal effects 398
 - `margins` 291–293, 300–301, 317, 354–355, 395–396, 408–409, 416–417, 420, 426
 - `marginsplot` 293, 300–301, 317, 354–355, 396–397, 408–409, 416–417, 426
 - marker
 - colors 142
 - labels 137, 152–155
 - opacity 142
 - options 140
 - sizes 144
 - symbols 137, 141
 - transparency 142
 - `marksample` 506, 513
 - master data 459
 - `match` see datasets, combine
 - match type 458–459
 - `matplotlib` (Python module) 171, 286
 - `matrix` (graph type) 135–136, 329, 332
 - maximum 9, 193–195
 - maximum likelihood
 - principle 388–391
 - search domain 400
 - `mband` (plotype) 329–330
 - `mdb.dta` (dataset) 116, 121
 - `mdy()` (function) 122
 - `mdyhms()` (function) 126–127

mean see average
 mean 230–231, 237, 242, 256–257,
 357–358
 mean comparison test 255–258
 median 193–195
 median regression 340
 median-trace 329–330
 memory see RAM
 merge 454, 457–467
 _merge (variable) 459–460, 462
 metadata 463
 mi 245–247
 mi() see missing()
 minimum 9, 193–195
 minutes() (function) 127–128
 missing values see missings
 missing() (function) 66, 307–308
 missings
 coding 451–452
 definition 6
 encode 104
 in expressions 60–61, 64–65
 set 11–12, 102–103
 ML see maximum likelihood
 ml 391
 mlabel() (scatter option) ... 152–155
 mlabposition() (scatter option)....
 153–154
 mlabsize() (scatter option)
 153–154
 mlabvposition() (scatter option) ...
 154
 MLE see maximum likelihood
 mlogit 424–425
 more off 225
 MSS 295–296
 multicollinearity 305
 multinomial logistic regression
 423–426
 multiple imputation 240–247
 mvdecode 11, 12, 103
 mvencode 104

N

_N 113–114, 125

_n 112–113, 125
 net 30
 net install 30–31, 347
 NetCourses 28
 newlist 73
 nofreq (option) 180
 nonlinear relationships ... 312, 415–417
 nonresponse 239
 normal distribution 221–222, 228
 density 421
 density function 421–422
 normal() (function) 252–253,
 256–257
 note() (graph option) 162–163
 notes 105
 null hypothesis 251, 256
 null model 399–400
 numlabel 107, 108
 numlist 67, 73
 numpy (Python module) 286

O

observations
 definition 5
 list 6–7
 odds 384–386
 odds ratio 385–386, 394–395
 odds-ratio interpretation 394–395
 ologit 429
 OLS 286–288
 operators 62–63
 options 13, 55–57
 order 475
 ordered logistic regression see
 proportional-odds model
 ordinal logit model see
 proportional-odds model
 ordinary least squares see OLS
 outcome 55, 218, 259, 261–265, 267,
 268, 270, 319, 320, 323, 325,
 328, 358, 360, 368, 369, 372,
 382, 389, 394

P

pandas (Python module) 87, 453

- panel data see datasets, panel data
 - partial correlation.....see regression,
 standardized coefficient
 - partial regression plot.....see
 added-variable plot
 - partial residual plot.....see
 component-plus-residual plot
 - PATE.....263, 321, 322, 356
 - pctile.....108
 - PDF.....170
 - Pearson χ^2404–405
 - Pearson residual.....404–405
 - Pearson's r .. see correlation coefficient
 - pequiv07kkh.dta (dataset).....468
 - percentiles.....see quantiles
 - pie (graph type) 135–136, 186, 201
 - pie charts.....186, 201
 - pip install.....87
 - plot region.....137, 148–149
 - plotregion() (graph option).....
 148–149
 - PNG.....170
 - pocket calculator.....62
 - popst1.dta (dataset).....465, 466
 - population average treatment effect ...
 see PATE
 - Portable Document Format ... see PDF
 - Portable Network Graphic see PNG
 - post-stratification weights..see weights
 - postest.....291
 - postestimation command.....291
 - PostScript.....see PS
 - potential outcome.....261–262
 - ppfad.dta (dataset).....463
 - PPS.....see sample, PPS
 - PRE.....268
 - predict.....291, 357–358, 395
 - predicted values.....290–293
 - Pregibon's $\Delta\beta$see $\Delta\beta$
 - preserve.....144
 - primary sampling unit.....232, 237
 - probability density function.....248
 - probability interpretation.....395, 397
 - probability proportional to size.....see
 sample, PPS
 - probit.....422–423
 - probit model.....421–423
 - profile plot.....300–301, 416–417
 - progex.dta (dataset).....487
 - program.....493–497
 - programs
 and do-files.....494
 - debugging.....495, 506
 - defining.....493–494
 - in do-files.....495–497
 - naming.....495
 - redefining.....494
 - syntax.....503–506, 509–511
 - proportion.....266
 - proportional reduction in error.....see
 PRE
 - proportional-odds model.....429–430
 - prtest.....195
 - PS.....170
 - pseudo R^2400
 - PSID.....455, 468
 - PSU.....see primary sampling unit
 - pwd.....24, 68
 - pweight (weight type).....75, 78–79
 - pyplot (Python class).....171
 - Python.....86–91, 453–454
 classes.....88
 - graphs.....170–172
 - methods.....88–89
 - modules.....87–88
 - Package Index.....87
 - Stata Function Interface 88–90
 - version.....87
 - python.....86, 171
 - python query.....87
 - python search.....87
 - python which.....88, 453
- Q**
- quantile.....213
 - quantile plot.....211–214
 - quantiles.....192–195
 - quartiles.....193–195
 - quietly.....508
 - Q–Q plots.....214–215

R

r() (stored results).....81–83
r(max) (stored result).....81–82
r(mean) (stored result)...73–74, 81–82
r(min) (stored result).....81
r(N) (stored result).....81–82
r(sd) (stored result).....81–82
r(sum) (stored result).....81
r(sum_w) (stored result).....81
r(Var) (stored result).....81–82
r-class.....see commands, r-class
 R^2268, 296, 323
 random numbers.....73, 96, 218–219
range() (scale suboption)....149–150
ranksum.....195
.raw (extension).....68, 440–441
 raw data.....see **.raw** (extension)
rbeta() (function) ..220–221, 259–260
recode() (function).....101, 189–190
recode.....108
 reference lines.....137, 151
regress.....16, 21, 53, 288
 regression
 adjusting interpretation.....318
 ANOVA table.....293–295
 autocorrelation.....343–344
 coefficient.....288–293, 299–301
 command.....288, 298–299
 diagnostics.....324–344
 estimating causal effects..319–323
 first difference.....366–368
 fit.....296–297
 for description.....316–317
 homoskedasticity.....341–343
 instrumental variables....372–376
 linear.....21–22, 281–377
 LPM.....380–384
 multiple.....297–298
 nonlinear relationships....311–315
 omitted variables.....341
 panel data.....359–371
 plots of coefs.....348–352
 prediction.....323, 356–358
 profile plots.....352–355
 residuals.....293

 simple.....284–288
 standardized coefficient...302–304
 tables.....345–348
 usages.....315–323
replace.....19–20, 27, 73, 93–116
requests (Python module).....453
reshape.....363–364
residual (**predict** option).....293
 residual
 definition.....285
 sum.....287, 295
 residual sum of squares.....see RSS
 residual-versus-fitted plot....326–327,
 342–343, 383–384
 response variable.....see variables,
 dependent
restore.....144
 Results window..see windows, Results
resultsset.....350
return list.....82–83
reverse (scale suboption).....150
 RGB.....142
rnormal() (function).....96, 222
 root MSE.....296
round() (function).....99
row (option).....14, 180
rowmiss() (**egen** function).....110
rplot.do.....283
 RSS.....268, 275, 287, 295
rstudent (**predict** option).....342
 Rubin causal model.....see causality
run.....272, 498
runiform() (function).....73, 96,
 218–219, 224
runiformint() (function).....219
 running counter.....112–113
 running sum.....114–116
rvfplot.....326–327

S

sample
 PPS.....232–233
 SRS.....225, 230–231
 cluster.....231–232
 multistage.....232

- sample (*continued*)
 - stratified 233–234
 - two-stage 232
- sample** 223
- sampling distribution 225–229, 234–235
- sampling probability 225, 232–233
- sampling weights *see* weights
- SAS files 435, 440
- save** 27, 475
- Scalable Vector Graphic *see* SVG
- scatter** 132
- scatter** (plottype) 135–136, 139
- scatterplot 281–282
- scatterplot matrix 328–329, 332
- scatterplot smoother 329–330
- scikit-learn** (Python module) ... 269
- _se**[*name*] 85
- search** 16, 30, 31
- secondary sampling unit 232
- selection probability *see* sampling probability
- sensitivity 402–403
- separate** 215
- set** 4
- set obs** 219
- set seed** 218–219
- sfi** *see* Python (Stata Function Interface)
- sign interpretation 394
- significance test 250–258
- simple random sample *see* sample, SRS
- SJ 16, 28
- SJ-ados 30–31
- SMCL 43–44, 516–517
- .smcl** (extension) 68
- SOEP *see* GSOEP
- soepuse** 31, 454
- sort** 8–9
- sort** (**scatter** option) 147
- specificity 402–403
- spreadsheet format 440–443
- SPSS files 435, 440
- sqrt()** (function) 65–66
- SRS *see* sample, SRS
- ssc install** 31, 39, 133, 178, 349, 351
- SSC-ados 31
- SSU *see* secondary sampling unit
- standard deviation 9, 191
- standard error .. 226–227, 230–231, 256
- standard normal distribution *see* normal distribution
- Stata Journal* *see* SJ
- Stata Press 28
- Statalist 28
- statsby** 350–351
- stereotype model 427–429
- storage types 117, 128–129
- stored results ... 73–74, 81–84, 289–290
- strings 444–445, 452
 - in expressions 118
 - replace substrings 119–121
 - storage type 117
 - to dates 123–124
 - to numeric 117–118
 - variables 116–121
- strong first stage 372
- strpos()** (function) 118–119
- strtrim()** (function) 120
- Student's *t* distribution 257
- subinstr()** (function) 120–121
- subscripts 115–116
- substr()** (function) 119–120
- subtitle()** (**graph** option) ... 162–163
- sum()** (function) 114–116
- summarize** 9–10, 13, 51, 55, 81, 193–194
- summarize()** (**tabulate** option) 195–196
- summary graphs 201–203
- summary tables 195–201
- superposition 202–203
- SVG 170
- svy** (prefix) ... 236–238, 257–258, 288, 344, 393
- svyset** 236–237, 288, 393
- symmetry plot 341–342
- sympplot** (statistical graph) 342

syntax 503–506, 509–511
 syntax checks 503
 syntax diagram 51

T

tab-separated values ... see spreadsheet
 format
tab1 178
tab2 181
table 181, 197–201
tabstat 194–195
tabulate 14, 56, 107, 177–182, 305
 tagged-image file format see TIFF
 tau-b see Kendall's tau-b
 Taylor linearization 235
teffects 358
tempvar 513–515
test 257–258, 266
 test dataset 275–277
 test statistic 253
text() (**twoway** option) 154–155
 text files 436, 440–448
 textbox options 161
 tick lines 137, 159
 TIFF 170
 time
 display-format 127
 elapsed 125–128
 from strings 127
 time-series varlist 367
titanic.dta (dataset) 259
title() (**graph** option) 162–163
 total sum of squares see TSS
 total variance see TSS
trace 495
 training dataset 275–277
translate 43
 TSS 268, 274, 294–296
ttest 195, 257
 two-stage least squares 373
 two-way fixed-effects model 371
 two-way table see frequency tables,
 two-way
twostep 352

twoway (graph type) 132, 135–136,
 291
.txt (extension) 68

U

U-shaped relationship 314–315, 330
 Unicode 94–95, 436, 475–485
unicode 479
 uniform distribution 218, 220
 unit nonresponse see nonresponse
update 29
 updating Stata 28–29
upper() (function) 118, 483
use 4–5, 23, 433
using 67–69
 using data 459
udstrlength() (function) 484
ustrlength() (function) 484
ustrsortkey() (function) 482–483
ustrupper() (function) 483

V

V see Cramér's *V*
 value labels see labels, values
valuelabel (axis-label suboption)
 159
 variables
 _all 53
 allowed names 94–95
 categorical 304, 379
 center 39, 73–74, 82, 300–301,
 307–308
 definition 5–6
 delete 6, 53
 dependent 281
 dummy 96–97, 299, 304–306
 generate 19–20, 93–105
 group 187–190
 identifier 451
 independent 281
 list 7, 52–55
 multiple codings 452
 names 105
 ordinal 427
 replace 19–20, 93–105

variables (*continued*)
 temporary 513–515
 transformations 137, 311–315,
 331, 340, 343
 Variables window see windows,
 Variables
 variance see standard deviation
 variation 294
 varlist see variables, list
 vce(robust) 343
 vector graphs 170
 version 43
 view 43–44

W

webscraping 453–454
 webuse xxiv
 weight (weight type) 75
 weights 75–79, 237–240
 whisker 203–204
 who.dta (dataset) 281, 313
 wildcards 54
 windows,
 changing 23
 Command 1–2
 font sizes for 2
 History 1–2, 7
 preferences for 2
 Results 1–2
 scroll back in 4
 Variables 1–2
 within estimator 360, 365–371
 working directory see directory,
 working directory

X

xlabel() (twoway option) 157–159
 xline() (twoway option) 151
 .xls (extension) see Excel files
 xtick() (twoway option) 159
 xpequiv.dta (dataset) 469
 xscale() (twoway option) 149–151
 xsize() (graph option) 148
 xstata 1
 xstata-mp 1

xstata-se 1
 xt commands 360–361
 xtick() (twoway option) 159
 xtile 108, 189, 343
 xtitle() (twoway option) 160–161
 xtreg 370
 xtset 367

Y

ylabel() (twoway option) 157–159
 yline() (twoway option) 151
 ytick() (twoway option) 159
 yscale() (twoway option) 149–151
 ysize() (graph option) 148
 ytick() (twoway option) 159
 ytitle() (twoway option) 160–161

Z

z transformation 252
 zip archive xxiv