

Contents

	List of tables	xv
	List of figures	xvii
	Preface	xxi
	Support materials for the book	xxiii
1	Getting started	1
1.1	Conventions	1
1.2	Introduction	4
1.3	The Stata screen	7
1.4	Using an existing dataset	9
1.5	An example of a short Stata session	11
1.6	Summary	18
1.7	Exercises	18
2	Entering data	21
2.1	Creating a dataset	21
2.2	An example questionnaire	23
2.3	Developing a coding system	24
2.4	Entering data using the Data Editor	29
2.4.1	Value labels	32
2.5	The Variables Manager	33
2.6	The Data Editor (Browse) view	38
2.7	Saving your dataset	39
2.8	Checking the data	41
2.9	Summary	44
2.10	Exercises	44

3	Preparing data for analysis	47
3.1	Introduction	47
3.2	Planning your work	47
3.3	Creating value labels	53
3.4	Reverse-code variables	56
3.5	Creating and modifying variables	60
3.6	Creating scales	65
3.7	Saving some of your data	68
3.8	Summary	69
3.9	Exercises	70
4	Working with commands, do-files, and results	71
4.1	Introduction	71
4.2	How Stata commands are constructed	72
4.3	Creating a do-file	76
4.4	Copying your results to a word processor	82
4.5	Logging your command file	83
4.6	Summary	85
4.7	Exercises	86
5	Descriptive statistics and graphs for one variable	87
5.1	Descriptive statistics and graphs	87
5.2	Where is the center of a distribution?	88
5.3	How dispersed is the distribution?	92
5.4	Statistics and graphs—unordered categories	94
5.5	Statistics and graphs—ordered categories and variables	103
5.6	Statistics and graphs—quantitative variables	105
5.7	Summary	112
5.8	Exercises	113
6	Statistics and graphs for two categorical variables	117
6.1	Relationship between categorical variables	117
6.2	Cross-tabulation	118

6.3	Chi-squared test	121
6.3.1	Degrees of freedom	123
6.3.2	Probability tables	123
6.4	Percentages and measures of association	125
6.5	Odds ratios when dependent variable has two categories	128
6.6	Ordered categorical variables	130
6.7	Interactive tables	133
6.8	Tables—linking categorical and quantitative variables	135
6.9	Power analysis when using a chi-squared test of significance	138
6.10	Summary	140
6.11	Exercises	141
7	Tests for one or two means	143
7.1	Introduction to tests for one or two means	143
7.2	Randomization	146
7.3	Random sampling	148
7.4	Hypotheses	148
7.5	One-sample test of a proportion	149
7.6	Two-sample test of a proportion	151
7.7	One-sample test of means	155
7.8	Two-sample test of group means	157
7.8.1	Testing for unequal variances	163
7.9	Repeated-measures t test	164
7.10	Power analysis	166
7.11	Nonparametric alternatives	173
7.11.1	Mann–Whitney two-sample rank-sum test	174
7.11.2	Nonparametric alternative: Median test	175
7.12	Summary	175
7.13	Exercises	176
8	Bivariate correlation and regression	179
8.1	Introduction to bivariate correlation and regression	179

8.2	Scattergrams	180
8.3	Plotting the regression line	185
8.4	Correlation	186
8.5	Regression	192
8.6	Spearman's rho: Rank-order correlation for ordinal data	197
8.7	Summary	198
8.8	Exercises	198
9	Analysis of variance	201
9.1	The logic of one-way analysis of variance	201
9.2	ANOVA example	202
9.3	ANOVA example using survey data	209
9.4	A nonparametric alternative to ANOVA	212
9.5	Analysis of covariance	215
9.6	Two-way ANOVA	226
9.7	Repeated-measures design	231
9.8	Intraclass correlation—measuring agreement	237
9.9	Summary	239
9.10	Exercises	240
10	Multiple regression	243
10.1	Introduction to multiple regression	243
10.2	What is multiple regression?	244
10.3	The basic multiple regression command	245
10.4	Increment in R-squared: Semipartial correlations	249
10.5	Is the dependent variable normally distributed?	251
10.6	Are the residuals normally distributed?	254
10.7	Regression diagnostic statistics	258
	10.7.1 Outliers and influential cases	258
	10.7.2 Influential observations: DF_{beta}	261
	10.7.3 Combinations of variables may cause problems	262
10.8	Weighted data	264

12.3.5	Rater agreement—kappa (κ)	328
12.4	Validity	331
12.4.1	Expert judgment	331
12.4.2	Criterion-related validity	332
12.4.3	Construct validity	333
12.5	Factor analysis	334
12.6	PCF analysis	339
12.6.1	Orthogonal rotation: Varimax	342
12.6.2	Oblique rotation: Promax	344
12.7	But we wanted one scale, not four scales	345
12.7.1	Scoring our variable	346
12.8	Summary	347
12.9	Exercises	348
13	Working with missing values—multiple imputation	349
13.1	The nature of the problem	349
13.2	Multiple imputation and its assumptions about the mechanism for missingness	351
13.3	What variables do we include when doing imputations?	353
13.4	Multiple imputation	354
13.5	A detailed example	355
13.5.1	Preliminary analysis	355
13.5.2	Setup and multiple-imputation stage	360
13.5.3	The analysis stage	362
13.5.4	For those who want an R^2 and standardized β s	363
13.5.5	When impossible values are imputed	366
13.6	Summary	367
13.7	Exercises	369
A	What's next?	371
A.1	Introduction to the appendix	371
A.2	Resources	371

A.2.1	Web resources	372
A.2.2	Books about Stata	373
A.2.3	Short courses	375
A.2.4	Acquiring data	376
A.3	Summary	377
References		379
Author index		383
Subject index		385