



# Contents

---

---

<b>Preface</b> .....	<b>ix</b>
Notes on the Eighth Edition .....	x
Acknowledgments .....	xi
<b>1 Stata and Stata Resources</b> .....	<b>1</b>
A Typographical Note .....	1
An Example Stata Session .....	2
Stata's Documentation and Help Files .....	7
Searching for Information .....	8
StataCorp .....	9
The <i>Stata Journal</i> .....	10
Books Using Stata .....	11
<b>2 Data Management</b> .....	<b>13</b>
Example Commands .....	14
Creating a New Dataset by Typing in Data .....	16
Creating a New Dataset by Copy and Paste .....	21
Specifying Subsets of the Data: <b>in</b> and <b>if</b> Qualifiers .....	23
Generating and Replacing Variables .....	26
Missing Value Codes .....	29
Using Functions .....	32
Converting Between Numeric and String Formats .....	36
Creating New Categorical and Ordinal Variables .....	39
Using Explicit Subscripts with Variables .....	41
Importing Data from Other Programs .....	42
Combining Two or More Stata Files .....	46
Collapsing Data .....	49
Reshaping Data .....	52
Using Weights .....	55
Creating Random Data and Random Samples .....	57
Writing Programs for Data Management .....	61
<b>3 Graphs</b> .....	<b>65</b>
Example Commands .....	65
Histograms .....	68
Box Plots .....	71
Scatterplots and Overlays .....	74
Line Plots and Connected-Line Plots .....	80
Other Twoway Plot Types .....	85

Bar Charts and Pie Charts	87
Symmetry and Quantile Plots	90
Adding Text to Graphs	93
Graphing with Do-Files	95
Retrieving and Combining Graphs	96
Graph Editor	97
Creative Graphing	100
<b>4 Survey Data</b>	<b>107</b>
Example Commands	108
Declare Survey Data	108
Design Weights	110
Poststratification Weights	113
Survey-Weighted Tables and Graphs	115
Bar Charts for Multiple Comparisons	119
<b>5 Summary Statistics and Tables</b>	<b>123</b>
Example Commands	123
Summary Statistics for Measurement Variables	125
Exploratory Data Analysis	127
Normality Tests and Transformations	129
Frequency Tables and Two-Way Cross-Tabulations	133
Multiple Tables and Multi-Way Cross-Tabulations	136
Tables of Means, Medians and Other Summary Statistics	139
Using Frequency Weights	140
<b>6 ANOVA and Other Comparison Methods</b>	<b>143</b>
Example Commands	144
One-Sample Tests	145
Two-Sample Tests	148
One-Way Analysis of Variance (ANOVA)	151
Two- and <i>N</i> -Way Analysis of Variance	154
Factor Variables and Analysis of Covariance (ANCOVA)	155
Predicted Values and Error-Bar Charts	158
<b>7 Linear Regression Analysis</b>	<b>163</b>
Example Commands	163
Simple Regression	167
Correlation	170
Multiple Regression	174
Hypothesis Tests	179
Dummy Variables	181
Interaction Effects	185
Robust Estimates of Variance	190
Predicted Values and Residuals	192
Other Case Statistics	197
Diagnosing Multicollinearity and Heteroskedasticity	202

Confidence Bands in Simple Regression .....	205
Diagnostic Graphs .....	209
<b>8 Advanced Regression Methods .....</b>	<b>215</b>
Example Commands .....	215
Lowess Smoothing .....	217
Robust Regression .....	221
Further <b>rreg</b> and <b>qreg</b> Applications .....	227
Nonlinear Regression — 1 .....	230
Nonlinear Regression — 2 .....	233
Box–Cox Regression .....	238
Multiple Imputation of Missing Values .....	241
Structural Equation Modeling .....	244
<b>9 Logistic Regression .....</b>	<b>251</b>
Example Commands .....	252
Space Shuttle Data .....	254
Using Logistic Regression .....	258
Marginal or Conditional Effects Plots .....	262
Diagnostic Statistics and Plots .....	264
Logistic Regression with Ordered-Category $y$ .....	268
Multinomial Logistic Regression .....	270
Multiple Imputation of Missing Values — Logit Regression Example .....	278
<b>10 Survival and Event-Count Models .....</b>	<b>283</b>
Example Commands .....	284
Survival-Time Data .....	286
Count-Time Data .....	288
Kaplan–Meier Survivor Functions .....	290
Cox Proportional Hazard Models .....	293
Exponential and Weibull Regression .....	299
Poisson Regression .....	303
Generalized Linear Models .....	307
<b>11 Principal Component, Factor and Cluster Analysis .....</b>	<b>313</b>
Example Commands .....	314
Principal Component Analysis and Principal Component Factoring .....	315
Rotation .....	318
Factor Scores .....	321
Principal Factoring .....	323
Maximum-Likelihood Factoring .....	325
Cluster Analysis — 1 .....	327
Cluster Analysis — 2 .....	331
Using Factor Scores in Regression .....	336
Measurement and Structural Equation Models .....	344

<b>12 Time Series Analysis</b>	<b>351</b>
Example Commands	351
Smoothing	353
Further Time Plot Examples	359
Recent Climate Change	363
Lags, Lead and Differences	366
Correlograms	371
ARIMA Models	374
ARMAX Models	382
<b>13 Multilevel and Mixed-Effects Modeling</b>	<b>387</b>
Example Commands	388
Regression with Random Intercepts	390
Random Intercepts and Slopes	395
Multiple Random Slopes	400
Nested Levels	404
Repeated Measurements	406
Cross-Sectional Time Series	410
Mixed-Effects Logit Regression	415
<b>14 Introduction to Programming</b>	<b>423</b>
Basic Concepts and Tools	423
Do-files	423
Ado-files	424
Programs	425
Local macros	426
Global macros	427
Scalars	427
Version	427
Comments	428
Looping	429
If ... else	430
Arguments	431
Syntax	432
Example Program: <b>multicat</b> (Plot Many Categorical Variables)	434
Using <b>multicat</b>	437
Help File	441
Monte Carlo Simulation	445
Matrix Programming with Mata	452
<b>Dataset Sources</b>	<b>457</b>
<b>References</b>	<b>461</b>
<b>Index</b>	<b>469</b>