

Contents

Figures
Preface

Getting Started

Dataset CD
Notes

Chapter 1 Introduction to Stata

Commands Covered

describe; codebook *varname*; set more off; File→Log→Begin;
search *keyword*; help *command_name*; which *package_name*;
ssc install *package_name*

Obtaining Information about a Dataset
Obtaining Information about Variables
Creating and Viewing a Log File
Printing Results and Copying Output
Getting Help
Installing Statistical Modules
Exercises
Notes

Chapter 2 Descriptive Statistics

Commands Covered

tabulate *varname*; summarize *varname*, detail; sktest *varname*;
histogram *varname*, d percent; histogram *varname*, percent;
sort *varname*; list *varname*; The Graph Editor Recorder

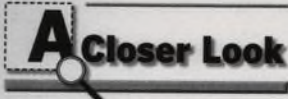
Interpreting Measures of Central Tendency and Variation
Describing Nominal Variables
Describing Ordinal Variables
Describing Interval Variables

x
xiii

1
2
3
5

5
7
8
9
10
12
13
13

15
145
146
155
157
157
15

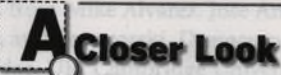
Obtaining Bar Charts and Histograms	22
<i>Creating Bar Charts</i>	23
 A Closer Look Graph Editor Recorder	28
<i>Creating Histograms</i>	29
Obtaining Case-level Information with sort and list	32
Exercises	33
Notes	42

Chapter 3 Transforming Variables **43**

Commands Covered

recode, generate (); generate, recode (); xtile, nquantiles (); generate; tabulate, generate (); label variable; label define; label values; drop; aorder

A Workbook Convention: Weighting the nes2008 Dataset	44
Transforming Categorical Variables	44
Transforming Interval Variables	46

 **A Closer Look**

The xtile Command 48

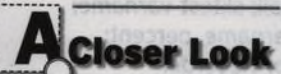
The label define and label values Commands	50
Creating an Additive Index	51
Creating Indicator Variables	53
Exercises	54
Notes	58

Chapter 4 Making Comparisons **59**

Commands Covered

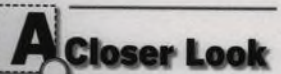
tabulate *dep_var indep_var*, column; tabulate *indep_var*, summarize (*dep_var*); format; _gwtmean, *dep_var_mean=dep_var* [if], by (*indep_var*); if; twoway (line *dep_var indep_var*, sort); replace; graph bar (mean) *dep_var*, over (*indep_var*)

Cross-tabulation Analysis	59
Mean Comparison Analysis	61
Visualizing Relationships with Line Charts and Bar Charts	62

 **A Closer Look**

The format Command 63

<i>Graphing an Interval-level Dependent Variable</i>	63
--	----

 **A Closer Look**

The if Qualifier 65

<i>Graphing a Categorical Dependent Variable</i>	70
--	----

A Closer Look

The replace Command

71

Exercises

73

Notes

90

Chapter 5 Making Controlled Comparisons

93

Commands Covered

```
bysort cntrl_var: tabulate dep_var indep_var, col; tabulate cntrl_var
indep_var, summarize (dep_var); graph bar dep_var, over (cntrl_var)
over (indep_var); twoway (line dep_var indep_var if cntrl_var==value1,
sort) (line dep_var indep_var if cntrl_var==value2, sort)
```

Cross-tabulation Analysis with a Control Variable

94

Bar Charts for Controlled Comparisons with a Categorical Dependent Variable

97

Mean Comparison Analysis with a Control Variable

100

Line Charts for Controlled Comparisons with an Interval-level Dependent Variable

103

Exercises

106

Notes

119

Chapter 6 Making Inferences about Sample Means

121

Commands Covered

```
ttest varname = testvalue; ttest varname, by(group_var); robvar
varname, by(group_var)
```

Describing a Sample Mean

122

Testing the Difference between Two Sample Means

127

Exercises

132

Notes

135

Chapter 7 Chi-square and Measures of Association

137

Commands Covered

```
(tabulate option) chi2; (tabulate option) V; somersd indep_var dep_var;
lambda dep_var indep_var
```

Analyzing Ordinal-level Relationships

138

Analyzing Nominal-level Relationships

143

A Problem with Lambda

145

Exercises

146

Notes

155

Chapter 8 Correlation and Linear Regression

157

Commands Covered

```
correlate varlist; regress dep_var indep_var(s); twoway
(scatter dep_var indep_var) (lfit dep_var indep_var)
```

The correlate Command and the regress Command

157

A Closer Look

R-squared and Adjusted R-squared: What's the Difference? 22159

- Creating a Scatterplot with a Linear Prediction Line 160
- Exploring Multivariate Relationships with Regression 163
- Exercises 165
- Notes 173

Chapter 9 Dummy Variables and Interaction Effects 175

Commands Covered

```
xi: regression dep_var i.indep_var; char varname [omit] #; test varname1 = varname2; predict newvar
```

Regression with Dummy Variables 175

A Closer Look

The test Command 178

- Interaction Effects in Multiple Regression 180
- Graphing Linear Prediction Lines for Interaction Relationships 183
- Exercises 186
- Notes 192

Chapter 10 Logistic Regression 193

Commands Covered

```
logit dep_var indep_var(s); logistic dep_var indep_var(s), [coef]; estimates store name; lrtest name, force; adjust indep_var1, by(indep_var2) pr gen(newvar); quietly; tabstat dep_var1 dep_var2, by(indep_var)
```

- The logit Command and the logistic Command 194
- Logistic Regression with Multiple Independent Variables 197

A Closer Look

The estimates Command and the lrtest Command 199

- Working with Predicted Probabilities: Models with One Independent Variable 200
- Working with Predicted Probabilities: Models with Multiple Independent Variables 203
- Exercises 208
- Notes 213

Chapter 11 Doing Your Own Political Analysis 215

- Five Doable Ideas 215
 - Political Knowledge 216
 - Economic Performance and Election Outcomes 216
 - State Courts and Criminal Procedure 216
 - Electoral Turnout in Comparative Perspective 217
 - Congress 217

4-7	Inputting Data	217
4-8	Stata-formatted Datasets	217
4-9	Microsoft Excel Datasets	218
4-10	PDF Format or Hand-coded Data	221
	Writing It Up	223
5-1	The Research Question	225
5-2	Previous Research	225
5-3	Data, Hypotheses, and Analysis	225
5-4	Conclusions and Implications	226
	Notes	226
Appendix		
5-8	Table A-1 Descriptions of Constructed Variables in gss2006	227
	Table A-2 Descriptions of Variables in the states Dataset	230
6-1	Table A-3 Descriptions of Variables in the world Dataset	235
8-1	Scatterplot with Linear Prediction Line	161
8-2	Scatterplot with Linear Prediction Line (unedited)	161
8-3	Editing a Scatterplot with a Linear Prediction Line	161
9-1	Linear Prediction Lines for Interaction Relationships	161
9-2	Linear Prediction Lines for Interaction Relationships (unedited)	161
9-3	Editing a Linear Prediction Overlay	161
10-1	A Logistic Regression Curve	161
10-2	A Logistic Regression Curve (edited)	161
10-3	Two Logistic Regression Curves	161
10-4	Two Logistic Regression Curves (edited)	161
11-1	Opening an Excel Dataset and Evaluating a Frequency Distribution (nominal variable)	161
11-2	Removing Commas from Data Values Using Excel	161
11-3	Copy/Pasting from Excel into Stata's Data Editor	161
11-4	Data in HTML Format	161
11-5	Copying HTML Data from Excel to Stata	161
11-6	Data in PDF Format	161
11-7	Entering Data Directly into the Data Editor	161
12-1	Obtaining a Histogram of an Interval Variable	161
12-2	Obtaining a Histogram of an Interval Variable (unedited)	161
12-3	Changing the Axis Rule for a Histogram	161
12-4	Changing the Bar Color in Bar Properties	161
12-5	Edited Histogram	161
13-1	An Unweighted and a Weighted Frequency Distribution	161
13-2	Creating Indicator Variables	161
14-1	Line Chart of the Relationship Between an Interval-level Dependent Variable and a Categorical Independent Variable	161
14-2	Two-way Graph	161
14-3	Changing Axis Properties	161
14-4	Changing the Y-axis Title	161
14-5	Adding a Chart Title	161
14-6	Adding a Chart Subtitle	161