Q-STEP STATA 'HOW TO' GUIDES:
STATA COMMAND SHEET

Creator: Joshua Townsley

Command	Use	Options	Example
use	Open a .dta file	, clear (to clear existing data)	use "Dropbox/Example.dta"
help	Get help on a topic		help graph

VIEWING DATA

Command	Use	Options	Example
summarize	Show summary statistics for your variable(s)	, detail (show extra statistics)	summarize variable1
tabulate	Make a table with your variable(s)	, nolabel (show raw numbers not value labels) , column (show column percentages) , row (show row percentages) , cell (show cell percentages)	tabulate variable1 variable2, column
browse	Browse your data in a spreadsheet table		browse
sort	Sort your data according to a variable		sort variable1

EDITING DATA

Command	Use	Options	Example
generate	Generate a new variable		gen variable3=.
recode	Recode an existing variable		recode age (18/29=1 "Under 30") (30/44=2 "30-44") (45/59=3 "45-59") (60/100=4 "60+"), gen(agecategory)
drop	Get rid of a variable		drop variable1
sample	Take a sample of your data, discarding the rest	, count (sample with replacement)	sample 50

MAKING GRAPHS

Command	Use	Options	Example
histogram	Create a histogram	, discrete (treats data as a categorical variable) *Various graph options exist, see histogram help for details*	histogram variable1, width(5) color(navy)
kdensity	Create a density plot	, normal (draw a normal distribution for comparison)	kdensity variable1, normal
scatter	Create a scatterplot	, mlabel(variable1) (add labels using variable1)	scatter variable1
graph box	Create a boxplot	, over(variable1) (create separate plots for different values of variable1)	graph box variable1, over(variable2)

BIVARIATE ANALYSIS

Command	Use	Options	Example
correlate	Produce a table of correlations		correlate variable1 variable2
pwcorr	Produce a table of pairwise correlations	, sig (run a significance test for the correlation(s))	pwcorr variable1 variable2, sig
ttest	Run a t test on a variable	, by(variable2) (test whether the means are different by different values of another variable) , unequal (do not assume equal variances in both groups)	Ttest variable1, by(variable2)
tab	Produce a cross tabulation, with a chi-square test	, chi2 (to produce chi-square test)	tab variable1 variable2, chi2

REGRESSION

Command	Use	Options	Example
regress	Run an OLS regression	i.variable1 (creates dummy variables for each value of variable1)	regress variable1 variable2 i.variable3
logit	Run a logit regression for a binary dependent variable		logit variable1 variable2
predict	After a regression, calculate predicted values of the dependent variable	, res (create residuals)	predict variable1_predict, res
margins	After a logit regression, calculate marginal probabilities	, at(variable1=1 variable2-2) (set other variables to specific values) , atmeans (set other variables to their means)	margins

Sheffield Methods Institute.