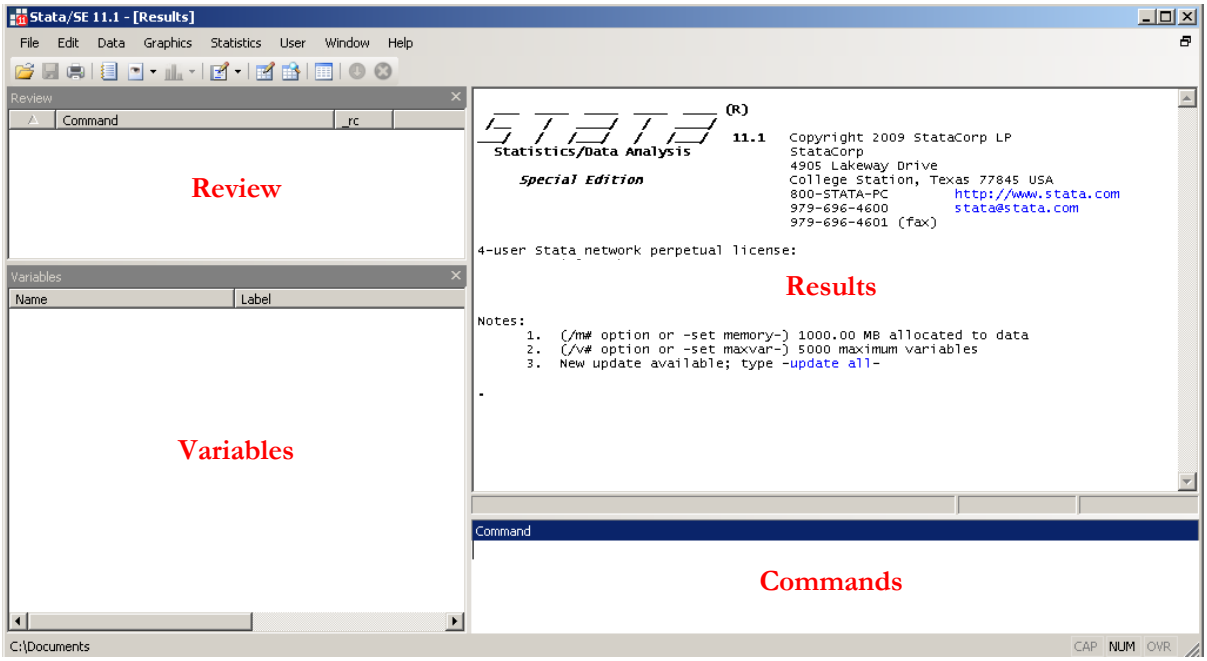
Oxford Poverty and Human Development Initiative (OPHI)
http://ophi.qeh.ox.ac.uk
Oxford Dept of International Development,
Queen Elizabeth House, University of Oxford

Basic STATA commands

Typical STATA window



Exploring your data

• Create a do file	doedit
• Change your directory	cd "**c:\your directory**"
• Open your database	use **database**, clear
• Import from Excel (csv file)	insheet using "**filename**.csv"
Filter (after the following commands)	if var1==3 or if var1=="male"
Equivalence symbols:	== equal; ~= not equal; != not equal; > greater than; >= greater than or equal; < less than; <= less than or equal; & and; | or.
Weight	[iw=weight] or [aw=weight]
• Browse your database	browse
• Look for a variables	lookfor "**any word/topic**"
• Summarize a variable (mean, standard deviation, min. and max.)	su **variable1 variable2 variable3**
• Tabulate a variable (per category)	tab **variable1** (add a second variables for cross tabs)
• Statistics for variables by subgroups	tabstat **variable1 variable2**, s(n mean) by(group)
• Information of a variable (coding)	codebook **variable1**, tab(99)
• Keep certain variables (use drop for the opposite)	keep var1 var2 var3
• Save a dataset	save **filename**, [replace]

Creating Variables

• Generate a new variable (a number or a combinations of other variables)	gen **new_variable** = 1 gen **new_variable** = **variable1**+ **variable2**
• Generate a new variable conditional (with restriction of range)	gen **new_variable** = 1 if **variable1**==0 gen **new_variable** = **variable1** if **variable2**==0
• Replace data in an existent variable	replace **variable1**=1 replace **variable1**= **variable1**+ **variable2**
• Replace data in an existent variable conditional	replace **variable1** = 1 if **variable2**==3 replace **variable1** =1/**variable2** if **variable1**==0
• Create a dummy variable (1 if the condition is satisfied, 0 otherwise)	gen **male**=(**sex**==2) gen **london**=(**region**==2)
• Create a variable 2 (specific actions)	egen **maxvar**=max(**var1**) egen **minvar**=min(**var1**) egen **rowsum**=rsum(**var1 var2 var3**)
• Label a variable	Label variable variable1 "Name of the variable 1"

Generating a loop

Assign one variable to the household	bys hh_id: egen hhvar=max(var1) bys hh_id: egen hhvar=min(var1)

Loop with strings	foreach var in string1 string2 string3 {
	gen new_`var' = 0.3
	}

Loop with numbers	forvalues x=1(1)13 {
	gen newvar`x'=`x'
	}

Explanation	forvalues x=minimum(interval)maximum {
	any action that you want but instead of the number you will write `x'
	}

Other

• Creating a matrix	Automatically	matrix a=J(n° of row, n° of columns, content)
	Directly	matrix a=[1,2,3\4,5,6] (, separates columns and \ rows)
• See the matrix a		matrix list a
• Create a log file to record your work		log using **filename**.log/smcl, [append replace]
• Close the log (at the end of the file)		log close

More information at http://www.stata.com/links/resources1.html