

# OPHI

OXFORD POVERTY & HUMAN DEVELOPMENT INITIATIVE

[www.ophi.org.uk](http://www.ophi.org.uk)



UNIVERSITY OF  
OXFORD

## Stata Features

Mauricio Apablaza  
OPHI

*Tabita, Kenya*



*Rabiya, India*



*Stephanie, Madagascar*



*Agatha, Madagascar*



*Dalma, Kenya*



*Ann-Sophia, Kenya*



*Valérie, Madagascar*



# Why STATA?

- Transformation tool
- Common Language
- Adaptable, scalable
- 2 ways of interaction (windows/commands)
- Easy storage of processes or syntaxes (do files)
- Other options: R, SPSS, GRETLM (open source), even excel (Depending on the sample size)

# STATA

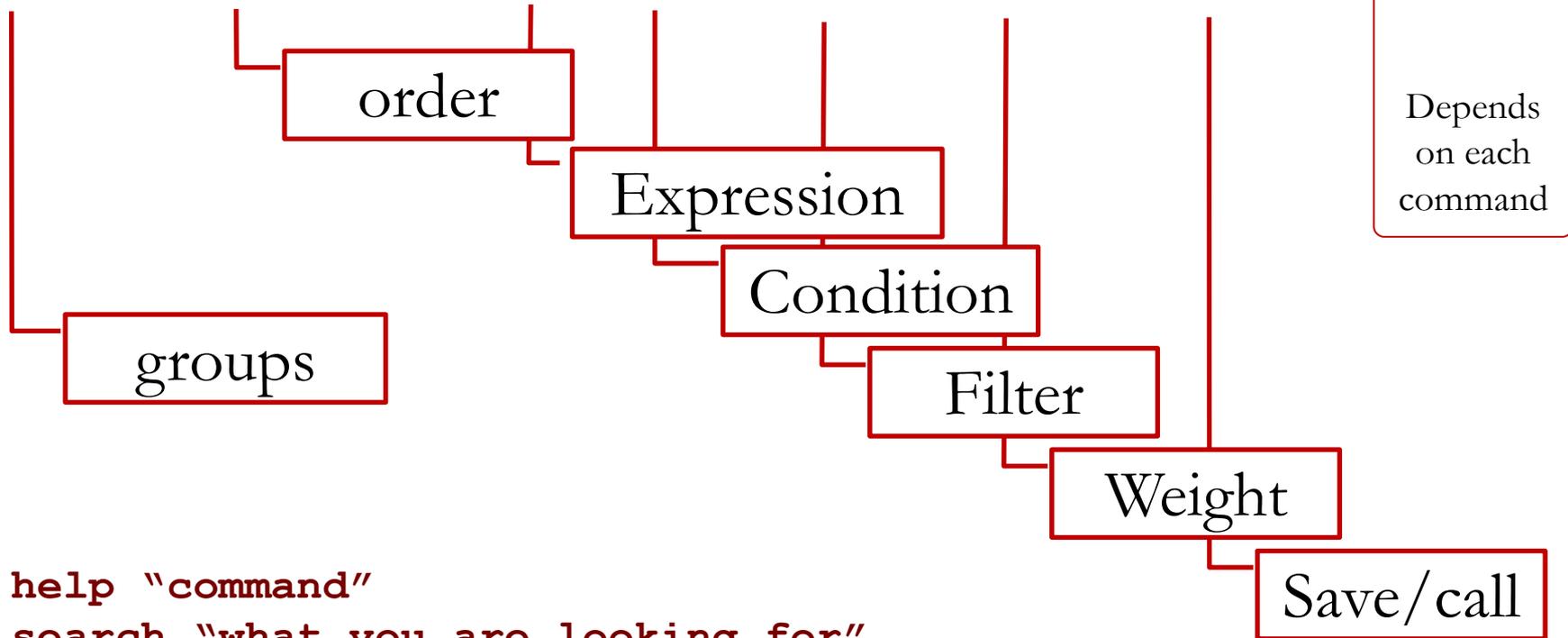
The screenshot shows the Stata/SE 11.1 interface. The window title is "Stata/SE 11.1 - [Results]". The menu bar includes File, Edit, Data, Graphics, Statistics, User, Window, and Help. The toolbar contains various icons for file operations and analysis. The main window is divided into several panels:

- Review**: A panel on the left side of the main window, currently showing the "Command" tab.
- Results**: The main central area displaying the Stata startup screen, including the STATA logo, version number (11.1), copyright information (2009 StataCorp LP), and contact details for StataCorp in Texas. It also shows "Notes" regarding memory and variable limits, and a message "running C:\Program Files\Stata11\profile.do ...".
- Variables**: A panel at the bottom left, currently empty, with columns for "Name" and "Label".
- Commands**: A panel at the bottom right, currently empty, for entering Stata commands.

At the bottom of the window, the status bar shows "d:\Economics\Data" on the left and "CAP NUM OVR" on the right.

# STATA Syntax

[by varlist:]command [varlist][=exp][if exp][in range][weight][using filename][,options]

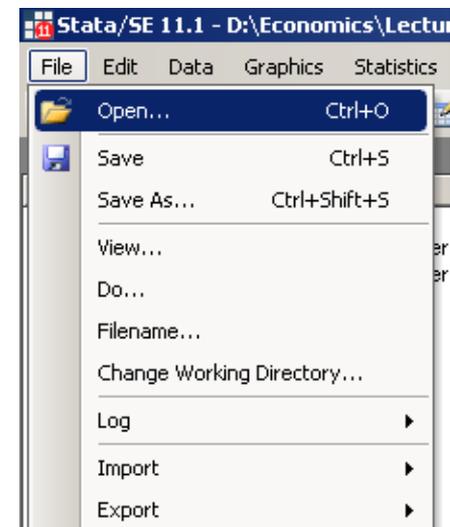
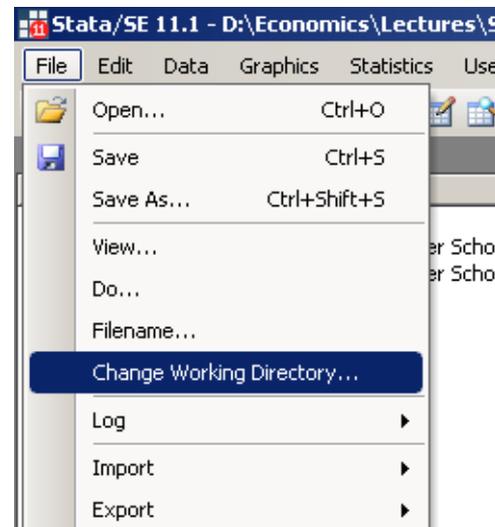
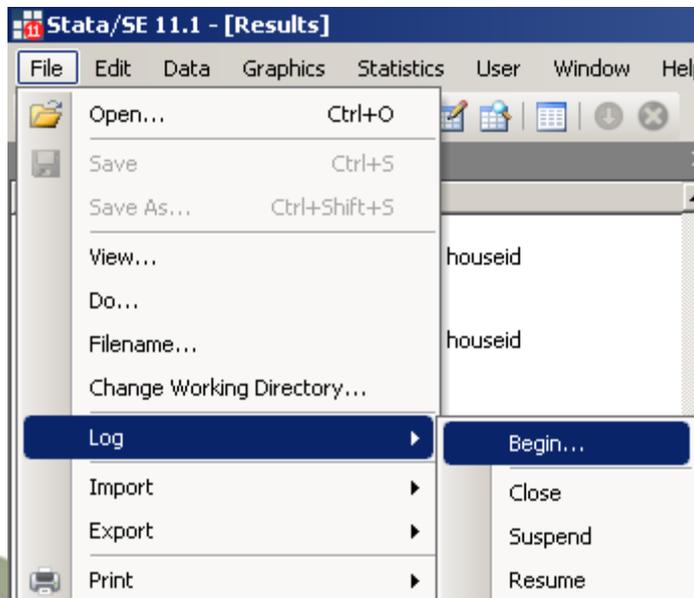


- `help "command"`
- `search "what you are looking for"`

# STATA (starting)

[by varlist:]command [varlist][=exp][if exp][in range][weight][**using filename**][,options]

```
log using "my_first_do_file.log", replace  
cd "d:\your directory"  
use "Sample_Bhutan.dta"
```



# STATA (few examples)

[by varlist:]command [varlist][=exp][if exp][in range][weight][**using filename**][,options]

## Short Examples

```
count
bys area: count
sort houseid slnp
browse slnp b11q1 b11q2 b11q3 b11q4 houseid
order Literate
```

# STATA (variables)

- Numerical
- String
- Categorical (coded)

```
. codebook b11q4, tab (20)
```

---

```
b11q4
```

---

```
type: numeric (byte)  
label: b11q4
```

```
range: [1,6] units: 1  
unique values: 6 missing .: 1/24786
```

```
tabulation: Freq. Numeric Label  
10177 1 married  
13042 2 never married  
439 3 divorced  
124 4 separated  
994 5 widow  
9 6 living together  
1 .
```

# STATA (weights)

- First, check the information about the construction of the survey.
- The most popular are:
  - Analytical weight (inversely proportional to the variance of an observation) `[aw=weight]` or
  - Frequency weight: number of duplicated observations `[fw=weight]`

# STATA: Exploring the dataset

Exploring variables:            age (numerical/individual)  
   toilet (categorical/household)

```
lookfor age  
summarize b11q3 [aw=weight]  
tabstat b11q3 [aw=weight], by(dcode) stats(mean min max)
```

```
lookfor toilet  
tab b2q16 [aw=weight]  
codebook b2q16  
tab b2q16 b2q17 [aw=weight]
```

And if we want to  
know the number of  
households?

If you have problems with your variables use `tostring` (transform numeric variables to string), `destring` (string to numeric), `encode` (Encode string into numeric), `decode` (Numeric variable to string variable)

# STATA: Creating new variables -gen

Creating a dummy for overcrowding  
(more than 3 people per room)

```
generate proom=(male+female)/b2q6
```

```
generate over1=(proom>3)
```

```
generate over2=0
```

```
replace over2=1 if proom>3
```

\* Check results

```
browse female male b2q6 proom over* if b11q2==1
```

# STATA: Creating new variables - egen

Creating variables with the mean, minimum, maximum age per household

```
bys houseid: egen eldest=max(b11q3)
bys houseid: egen youngest=min(b11q3)
bys houseid: egen meanage=mean(b11q3)
```

Other important (columns vs rows)

```
bys houseid: egen sum=total(b11q3)
egen total=rowtotal(deprivation*)
```

# STATA: Creating new variables

## Short Exercises (do file)

Create a dummy variable if the source of drinking water is: unprotected well, spring, river/lake/pond or other

**Create a dummy variable (per household) if the head of the household never went to school.**

# STATA: Creating new variables

In the stata memory

```
local w=0.33  
display `w'  
generate w_hnever=hnever*`w'  
  
summarize pce_real [aw=weight]  
return list  
local mean=r(mean)  
display `mean'  
local range=r(max)-r(min)  
display `range'
```

# STATA: Loops

## Repeated process

For instance (it is loop)

```
by area: count
```

Advanced Loops (using strings and numbers)

```
foreach var in string1 string2 {  
  local new_`var'=0.3  
}
```

```
forvalues j=1(1)5 {  
  matrix a[1,`j']=`j'  
}
```

# STATA

## Exercise

The survey include question of the about the household assets under the name b3q1XX. To see all of them write: `describe b3q1*`

In groups, define and calculate an assets index using at least 4 of the variables described.

Hints: To construct this index, first you transform the relevant variables to dummy variables. Then you should define an aggregation method (ej. If one assets is not achieved the household is deprived or an weighted average of the assets)

# STATA

## Merge 2 datasets

### **Merge 2 datasets with individual data**

```
merge houseid using floor, uniquising  
tab _merge  
drop _merge
```

### **Merge dataset with individual dataset with dataset with household dataset:**

```
merge houseid slnp using nationality, sort  
tab _merge
```

# OPHI

OXFORD POVERTY & HUMAN DEVELOPMENT INITIATIVE

[www.ophi.org.uk](http://www.ophi.org.uk)



UNIVERSITY OF  
OXFORD

## Stata Features

Mauricio Apablaza  
OPHI

*Tabita, Kenya*



*Rabiya, India*



*Stephanie, Madagascar*



*Agatha, Madagascar*



*Dalma, Kenya*



*Ann-Sophia, Kenya*



*Valérie, Madagascar*

