

# Stata Workshop

At MINAGRI

---

Roshni Khincha and Sakina Shibuya

DIME, World Bank

August, 2018



# Section 1

---

# Why learn stata?

## **Excel vs Stata**

Can I use Excel?

## The main reasons to use Stata

- In Excel you make changes directly to the data and save new versions of the data set
- In Stata you make changes to the instructions on how to get from the raw data to the final analysis and save new versions of the instructions
- Since Stata is a more statistics oriented software, processing the data to create analytical products can be a lot easier.

## The main reasons to use Stata

- Powerful tool with many capabilities:
  - Descriptive statistics
  - Inference statistics
  - Complex data analysis
- But it's also good for beginner programmers:
  - User friendly interface
  - Relatively easy programming language that can be learned while you're using the software

## What's the fuss about do-files?

- Its through the do-file you communicate your work to other members in your team, both current and future
- Think of the do-files as instructions on how to get from raw data to final report
- For a simple task you can enter commands manually. But for more complex tasks you need to write a recipe, or a list of instructions

# How to open up a data set in Stata?

**How to open up a data set in Stata?**

## Three ways to tell Stata what to do

- Drop-down menus
  - An easy place to start but quickly becomes inefficient
- Command window
  - Faster than menus but require that you are familiar with the command
- Do-file
  - The only feasible way to run long instructions
  - Use menus and command window to figure out what you need to write, then copy to a do file



**Open a dataset - menus**

### Open a dataset command window

- When you use the menus, Stata produces the code for that action (except for Data Browse)
  - Highlight, right-click and copy the code
  - Paste the code in the command window
  - Hit enter

## Task 1

1. Open Stata and then open the EICV household data set **cs\_s0\_s5\_household.dta** using the menu: File → Open. Navigate to where you saved the material for this lab. Select the data set and click *Open*
2. Browse to check that you have data: Data → Data Editor → Data Editor Browse
3. Describe to get additional information on the data: Data → Describe data → Describe data in memory or in a file.
  - A new window will open
  - Select In memory and press OK

## Task 1

- You can see that one the second command printed information on your screen.
  - The first part is the command used
  - The second part are the results

## Task 1

- You can perform both tasks by typing the in your command prompt.  
This will yield the same results
- Type *browse* in the command window and press enter
- Type *describe* and press enter

## **An introduction to Stata** Stata interface

•

•



•

## Section 2

---

**How can we delete irrelevant variables?**

## Delete variables

- blah blah

```
. use "$data\cs_s0_s5_household.dta", clear  
  
. keep hhid province district ur2012 s5cq2 s5cq4 s5cq8 s5cq15 s5c  
> q23 s5bq2 s5cq22 s5cq13 s5cq17
```

```
. use "$data\cs_s0_s5_household.dta", clear  
  
. keep hhid province district ur2012 s5cq2 s5cq4 s5cq8 s5cq15 s5c  
> q23 s5bq2 s5cq22 s5cq13 s5cq17
```

## Section 3

---