**STATA WORK OUTLINE**

Part I – Accessing Stata

* Create folder structure for project
* Access Stata on machines
* Walk through to Stata GUI

Part II – Importing data and basic commands

* Identify & change working directory (**cd**)
* Import data
  + Start with **help** import to see help file and various import methods
  + Use **insheet** for csv file
* **Save** data
* See how to look at the data through **browse**
* **Destring** data **[start with fiscalyears, quarters and spending/amount as string that easy to destring; can go deeper into instring in week 2]**
* Get a sense of the data (variables, numeric v string) with **describe**
* Sort data (via **sort** and **gsort**) and **list** the top expenditures in dataset
* Get summary statistics (**summarize**) and look that them in tabular form (**table**)
* Use the **if** to subset data (explain Relational Operators (>, <, >=, <=, ==, !=) and Logical Operators (&, |), and **inlist**)

Part II – Group Exercise I

* Import exercise file into Stata
* Determine how many string and numeric variables there are
* Report the average spending on Infrastructure for Afghanistan
* Which category received the most funding in Kenya in 2013? How much?

Part III – Working with variables

* **Generate** a new variable that reports spending in millions of dollars
* Add a data label (**label variable**) to the new variable to describe it and **format** it to have comma
* Create value labels (**label define**) for quarters, apply them (**label value**), and view them through the **codebook**
* **Encode** categories and sectors variables
* Change the **order** of the variables
* **Drop unnecessary string variables (category, sector) and spent2 as well as** observations where the operating unit is “Worldwide”

Part IV **–** Group Exercise II

* Create a binary variable for any USAID observations
* Label the variable, and create variable labels where “0” is “No” and “1” is “Yes”
* Create a table summarizing total USAID vs other spending by year

Part V – More work with data

* **Collapse** the dataset to aggregate spending country level rather than by activity
* Use **egen** to create a ranking and fiscal year mean

Part VI – Stata Q&A and Distribute Homework Assignment