SESSION-3: STATA DATA MANAGEMENT

Course detail: http://julhas.com/jsedutech/stata-level-one.html
Mentor: Julhas Sujan

Session Outline:

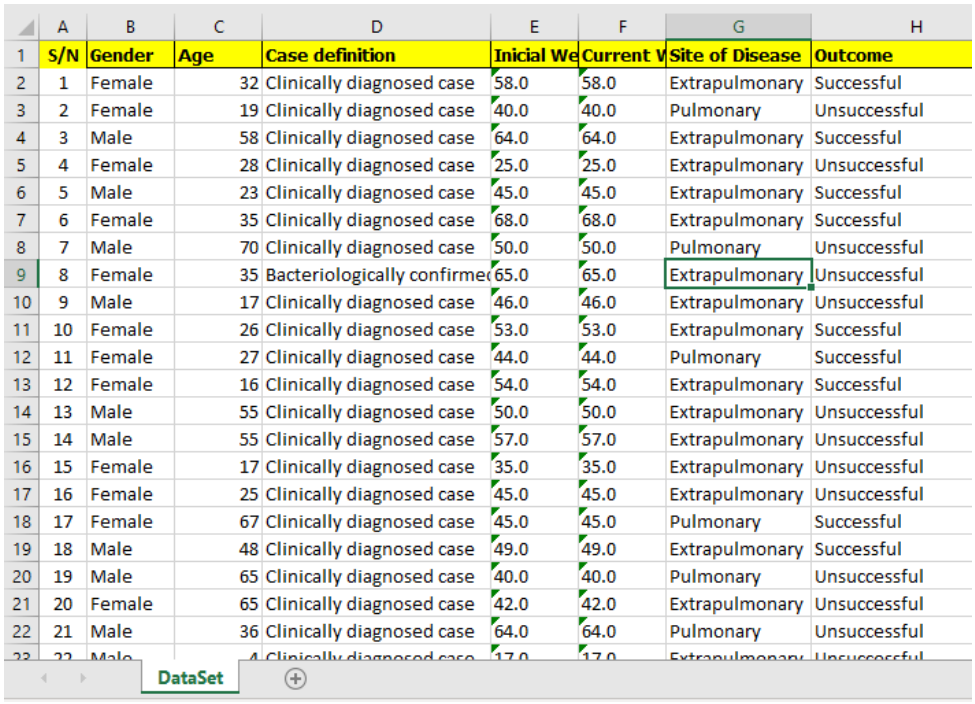
- Guest speech
- Recap Session-2 (Commands: describe, codebook, list, and summarize)
- Excel Dataset preparation and Import
- Variables (Generate, Rename, Replace, Drop by using editor and commands)
- Operators (addition, subtraction, multiplication, division, power)
- Data types (Byte, Integer, Long, Float, Double)
- Log file (Opening, Log storage and Closing)
- DO file (Opening, Creating, Saving commands, and Clear Widows)
- Self Practice
- Assignments
- Open discussion

Lession-1: Recap Session-2

- Opening Stata
- Commads - describe, codebook, list, and summarize

Lession-2: Excel DataSet preperation

- For this tutorial we have prepared an excel dataset for demosntration. It contains:
  - Serial number
  - Age
  - Gender
  - Case Definition
  - Initial Weight
  - Current Weight
  - Outcome



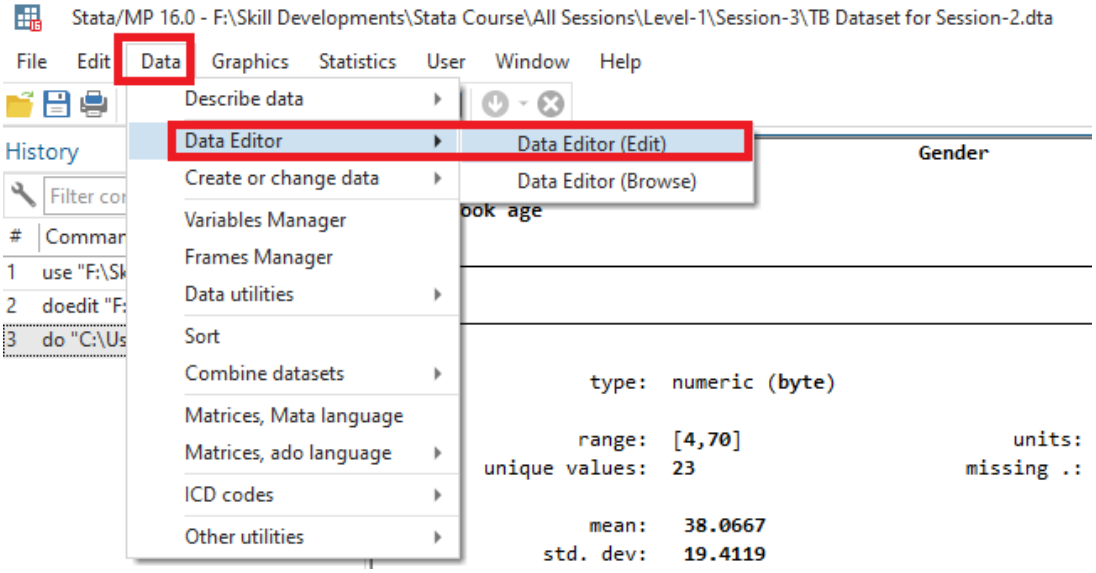
	A	B	C	D	E	F	G	H
1	S/N	Gender	Age	Case definition	Inicial We	Current V	Site of Disease	Outcome
2	1	Female	32	Clinically diagnosed case	58.0	58.0	Extrapulmonary	Successful
3	2	Female	19	Clinically diagnosed case	40.0	40.0	Pulmonary	Unsuccessful
4	3	Male	58	Clinically diagnosed case	64.0	64.0	Extrapulmonary	Successful
5	4	Female	28	Clinically diagnosed case	25.0	25.0	Extrapulmonary	Unsuccessful
6	5	Male	23	Clinically diagnosed case	45.0	45.0	Extrapulmonary	Successful
7	6	Female	35	Clinically diagnosed case	68.0	68.0	Extrapulmonary	Successful
8	7	Male	70	Clinically diagnosed case	50.0	50.0	Pulmonary	Unsuccessful
9	8	Female	35	Bacteriologically confirme	65.0	65.0	Extrapulmonary	Unsuccessful
10	9	Male	17	Clinically diagnosed case	46.0	46.0	Extrapulmonary	Unsuccessful
11	10	Female	26	Clinically diagnosed case	53.0	53.0	Extrapulmonary	Successful
12	11	Female	27	Clinically diagnosed case	44.0	44.0	Pulmonary	Successful
13	12	Female	16	Clinically diagnosed case	54.0	54.0	Extrapulmonary	Successful
14	13	Male	55	Clinically diagnosed case	50.0	50.0	Extrapulmonary	Unsuccessful
15	14	Male	55	Clinically diagnosed case	57.0	57.0	Extrapulmonary	Unsuccessful
16	15	Female	17	Clinically diagnosed case	35.0	35.0	Extrapulmonary	Unsuccessful
17	16	Female	25	Clinically diagnosed case	45.0	45.0	Extrapulmonary	Unsuccessful
18	17	Female	67	Clinically diagnosed case	45.0	45.0	Pulmonary	Successful
19	18	Male	48	Clinically diagnosed case	49.0	49.0	Extrapulmonary	Successful
20	19	Male	65	Clinically diagnosed case	40.0	40.0	Pulmonary	Unsuccessful
21	20	Female	65	Clinically diagnosed case	42.0	42.0	Extrapulmonary	Unsuccessful
22	21	Male	36	Clinically diagnosed case	64.0	64.0	Pulmonary	Unsuccessful
23	22	Male	4	Clinically diagnosed case	17.0	17.0	Extrapulmonary	Unsuccessful

DataSet

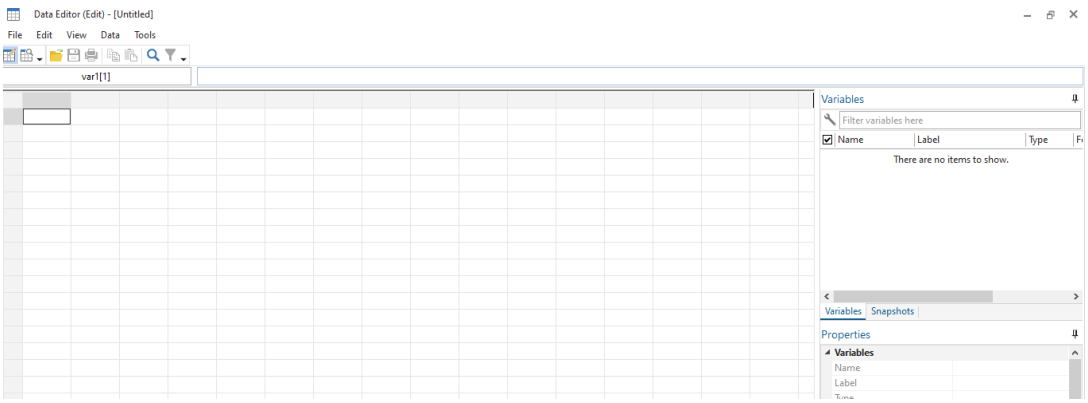
- Now we want to import this dataset. Select all the columns and copy this dataset.

	A	B	C	D	E	F	G	H
1	S/N	Gender	Age	Case definition	Inicial We	Current W	Site of Disease	Outcome
2	1	Female	32	Clinically diagnosed case	58.0	58.0	Extrapulmonary	Successful
3	2	Female	19	Clinically diagnosed case	40.0	40.0	Pulmonary	Unsuccessful
4	3	Male	58	Clinically diagnosed case	64.0	64.0	Extrapulmonary	Successful
5	4	Female	28	Clinically diagnosed case	25.0	25.0	Extrapulmonary	Unsuccessful
6	5	Male	23	Clinically diagnosed case	45.0	45.0	Extrapulmonary	Successful
7	6	Female	35	Clinically diagnosed case	68.0	68.0	Extrapulmonary	Successful
8	7	Male	70	Clinically diagnosed case	50.0	50.0	Pulmonary	Unsuccessful
9	8	Female	35	Bacteriologically confirme	65.0	65.0	Extrapulmonary	Unsuccessful
10	9	Male	17	Clinically diagnosed case	46.0	46.0	Extrapulmonary	Unsuccessful
11	10	Female	26	Clinically diagnosed case	53.0	53.0	Extrapulmonary	Successful
12	11	Female	27	Clinically diagnosed case	44.0	44.0	Pulmonary	Successful
13	12	Female	16	Clinically diagnosed case	54.0	54.0	Extrapulmonary	Successful
14	13	Male	55	Clinically diagnosed case	50.0	50.0	Extrapulmonary	Unsuccessful
15	14	Male	55	Clinically diagnosed case	57.0	57.0	Extrapulmonary	Unsuccessful
16	15	Female	17	Clinically diagnosed case	35.0	35.0	Extrapulmonary	Unsuccessful
17	16	Female	25	Clinically diagnosed case	45.0	45.0	Extrapulmonary	Unsuccessful
18	17	Female	67	Clinically diagnosed case	45.0	45.0	Pulmonary	Successful
19	18	Male	48	Clinically diagnosed case	49.0	49.0	Extrapulmonary	Successful
20	19	Male	65	Clinically diagnosed case	40.0	40.0	Pulmonary	Unsuccessful
21	20	Female	65	Clinically diagnosed case	42.0	42.0	Extrapulmonary	Unsuccessful
22	21	Male	36	Clinically diagnosed case	64.0	64.0	Pulmonary	Unsuccessful

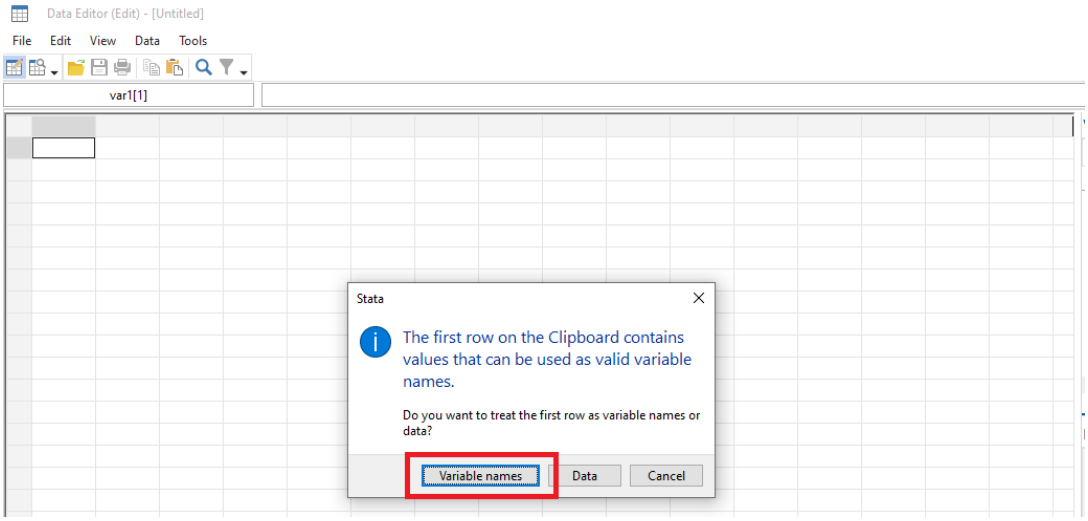
- Open Stata >> Data >> Data Editor >> Data Editor (Edit)



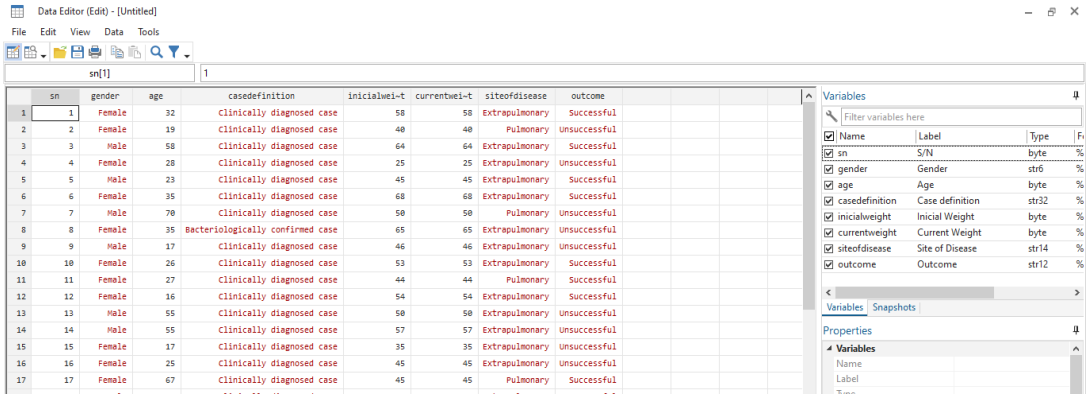
- Click here and you can see the following Blank window



- Paste (Ctrl+V) and click on the Variables name button

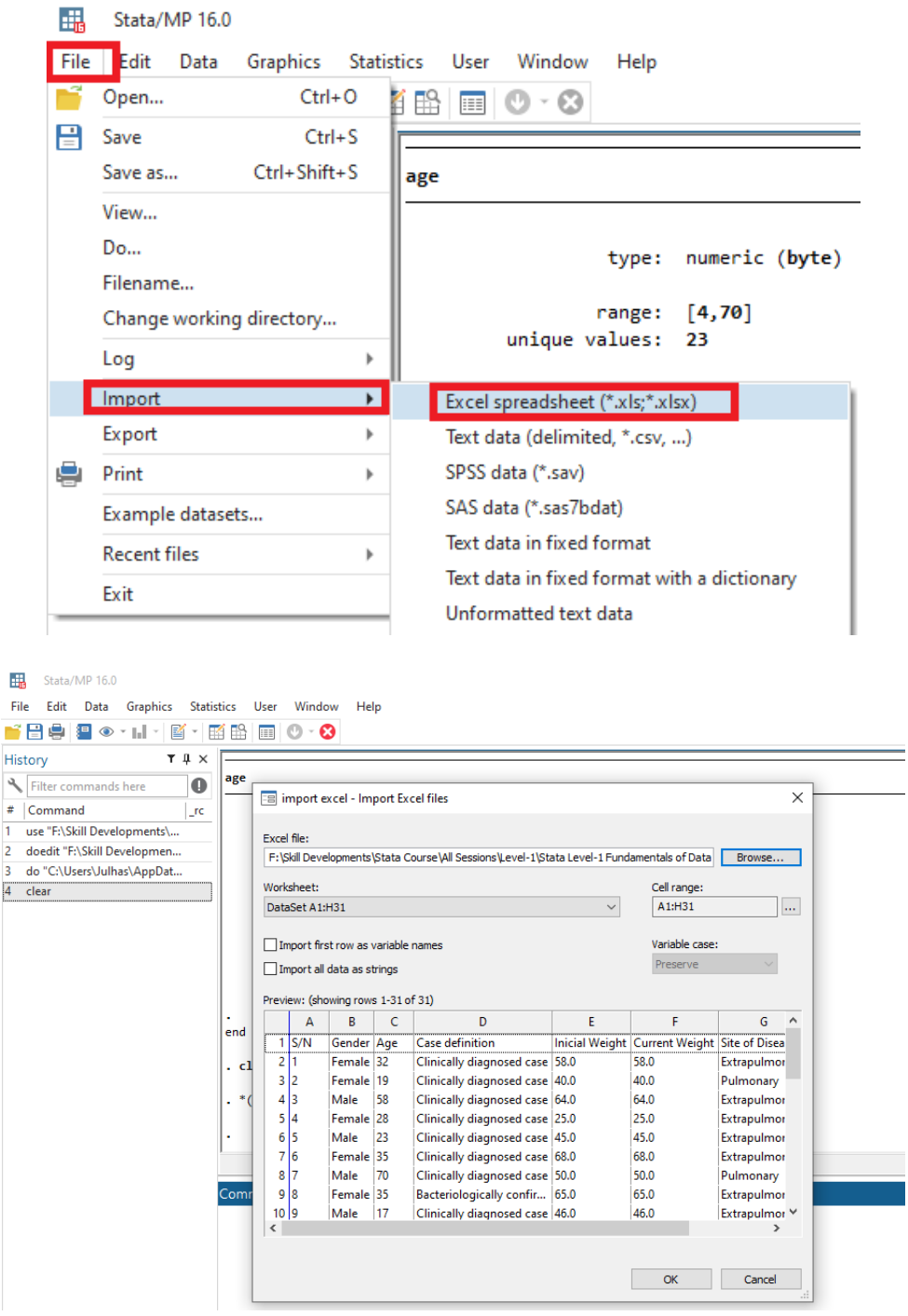


- Dataset is ready



- Save: Don't forget to save. You can from the Menu >> File >> Save as or **Crtl + Shift + S**.

Alternatively you can import the same dataset from the following menu: File >> Import >> Excel Spreadsheet



Command: import excel "F:\Skill Developments\Stata Course\All Sessions\Level-1\Stata Level-1 Fundamentals of Data Analysis-Data
> set.xlsx", sheet("DataSet") clear

Lession-3: Variables (Generate, Rename, Replace, Drop by using editor and commands)

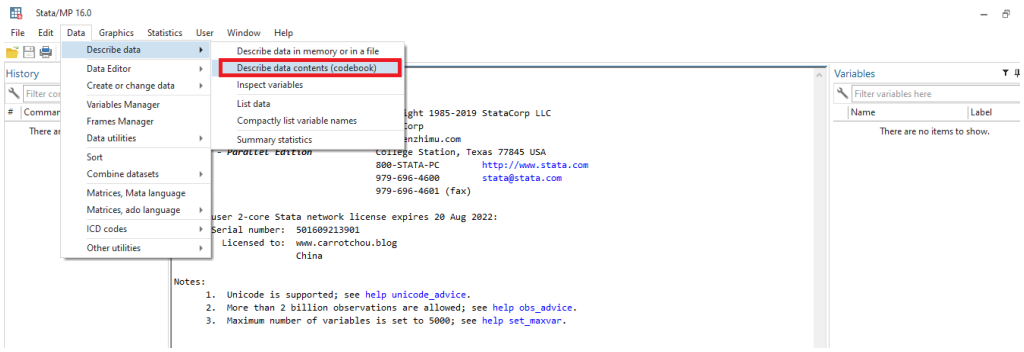
- **Generate:** To create a new variable in Stata using the generate command, usually abbreviated gen:
    - gen variable=something
    - generate location = .
    - gen location = "Dhaka"
- **Replace:** You can change the value of an existing variable using replace. Since replace can destroy data, it has no abbreviation. The basic syntax: replace variable=something
    - replace location = 'Rajshahi
- **Recoding with generate and replace:**
    - gen education = 12 if age >= 18
      replace gender="1" if gender=="Male"
- **Rename**: You can change the name of a variable with the rename command
    - rename location location_type
- **Drop varibale:** You can delete any variable by using 'drop' command
    - drop variable_name
    - drop location

Lession-4: Operators

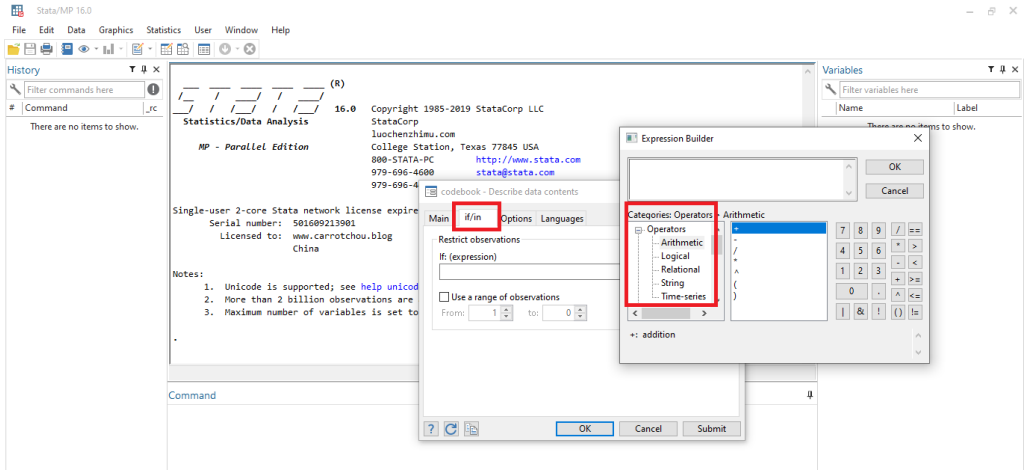
- The operators defined in Stata are given in the table below:

Arithmetic		Logical		Relational (numeric and string)	
+	addition	!	not	>	greater than
-	subtraction	|	or	<	less than
*	multiplication	&	and	>=	> or equal
/	division			<=	< or equal
^	power			==	equal
				!=	not equal
+	string concatenation				

- **Example to see the operators:**
    - Open Stata >> Clik on Data Menu >> Describe data >> Describe data contents (Coodbook)

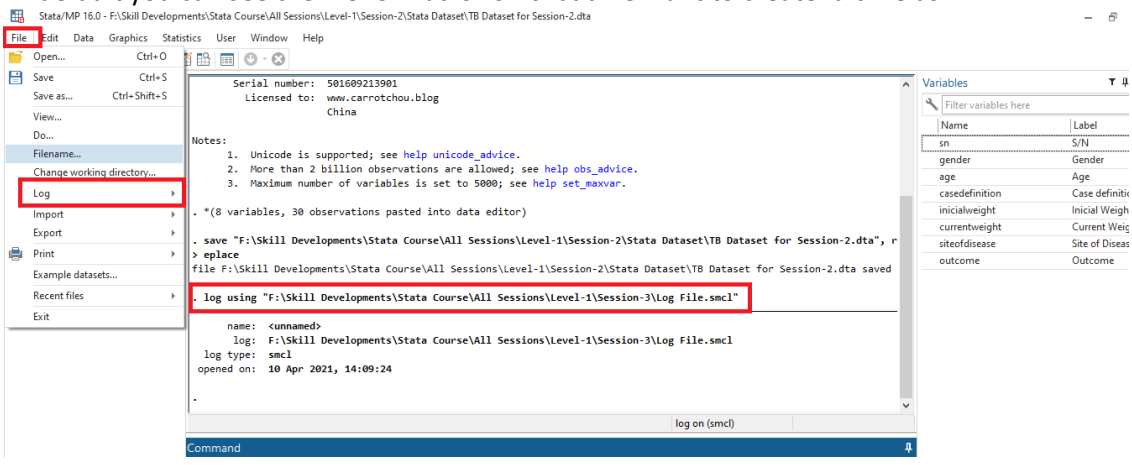


Next:



Lession-5: Log File

- **Create log file using Editor**: File >> Log >> Begin and Save this log file in your work directory. In default you can see the file format is . smcl but we want to create .txt file as:

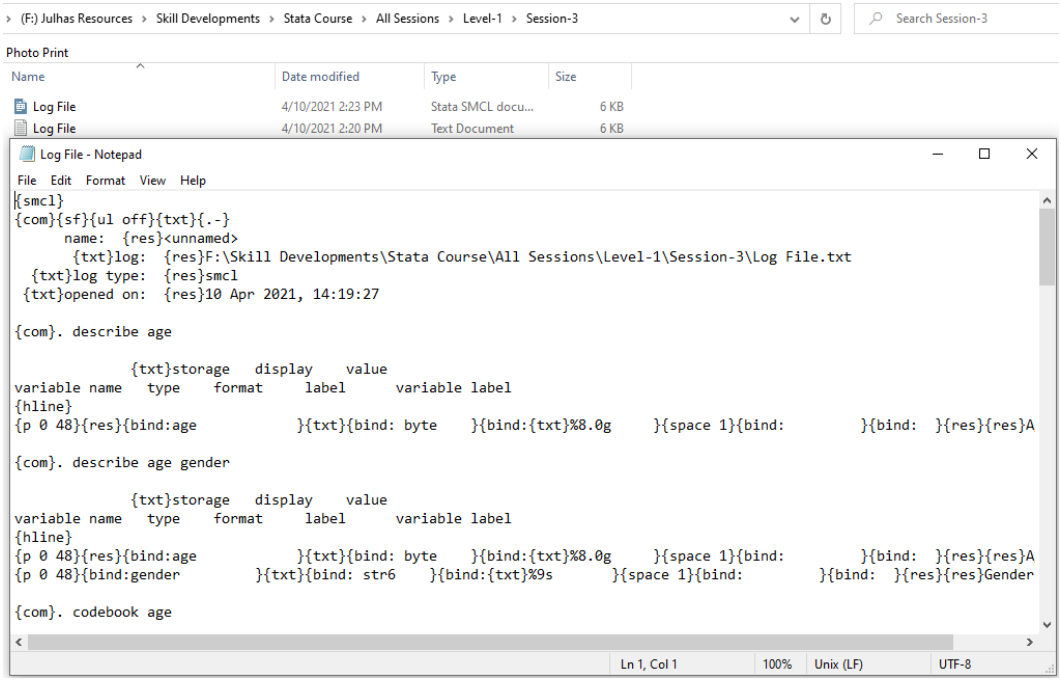


- **Commad:**
  - Syntax: log using “working directory”
  - Example:

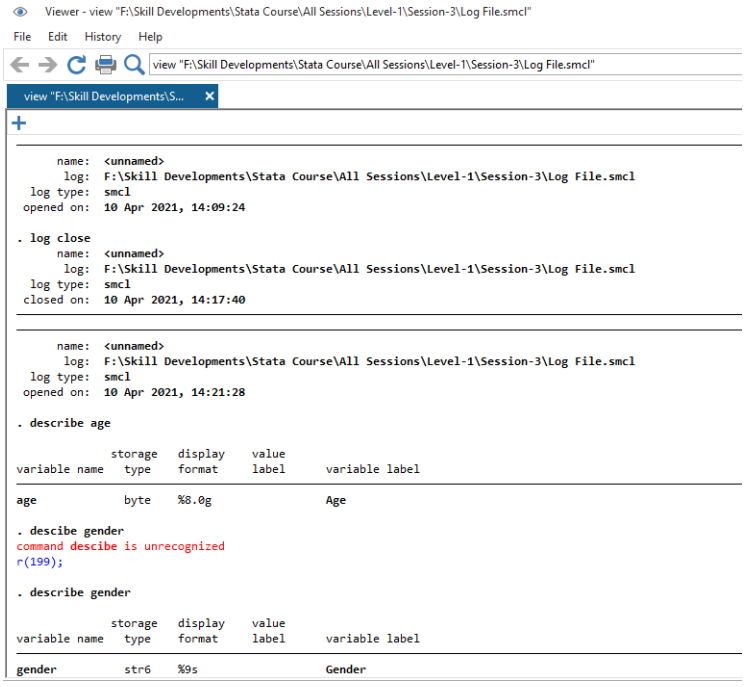
1. log using "F:\Skill Developments\Stata Course\All Sessions\Level-1\Session-3\Log File.txt"
2. log using "F:\Skill Developments\Stata Course\All Sessions\Level-1\Session-3\Log File. smcl"

- **Check log history:** Visit your working directory and open the Log File >> F:\Skill Developments\Stata Course\All Sessions\Level-1\Session-3\Log File.txt

1. Text file



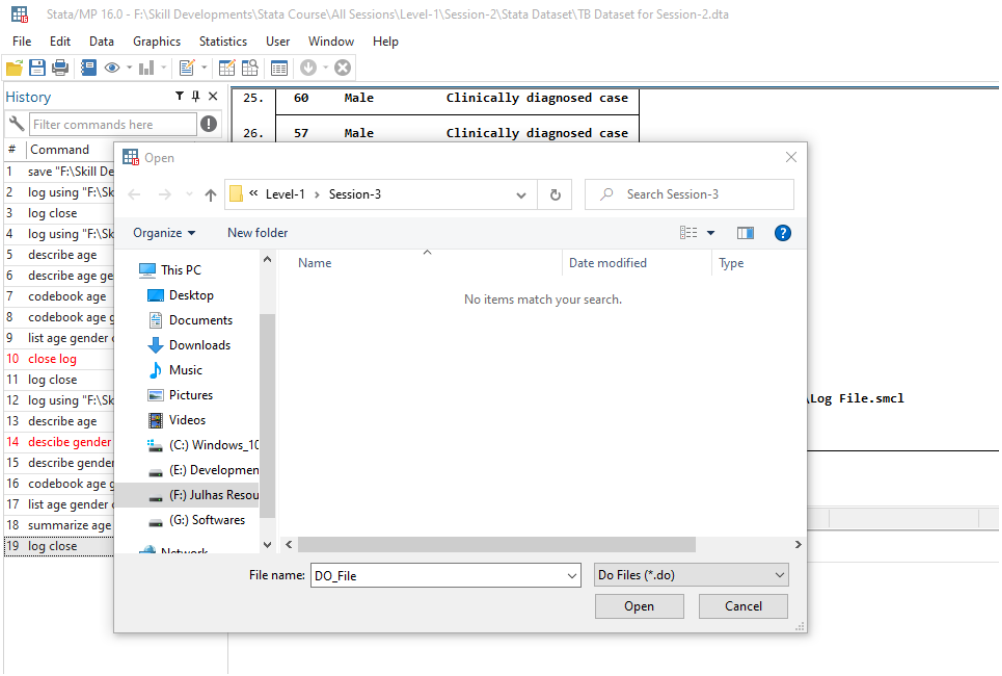
2. Stata file format



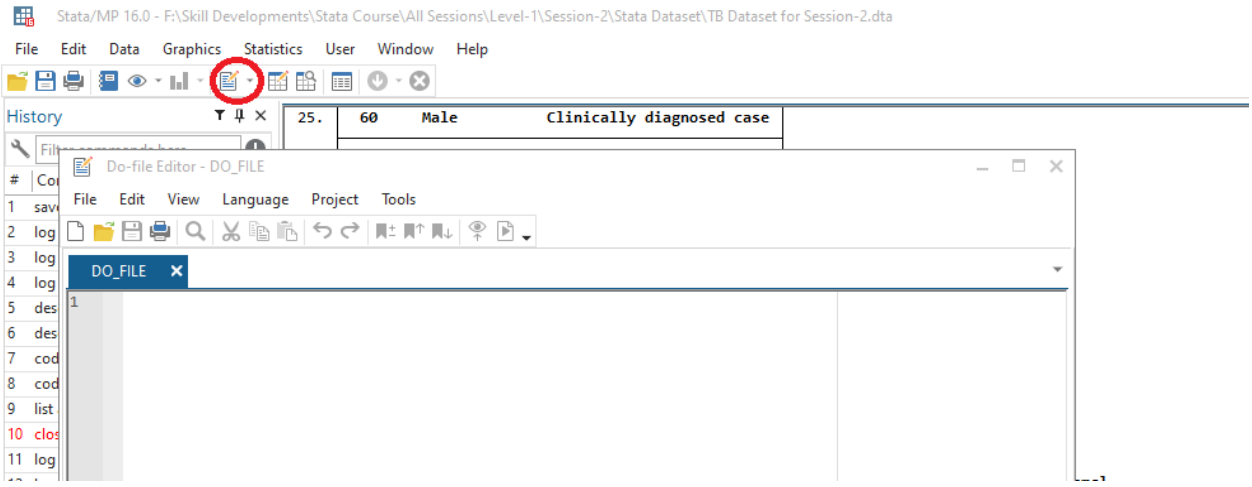
- **Close log: You must have to close the log:**
  - log close
- **Log append**: If you want to store all of your daily activity log then you can use append to store all results in a single file as:
  - log using "F:\Skill Developments\Stata Course\All Sessions\Level-1\Session-3\Log File.smcl", **append**

Lession-6: DO file (Opening, Creating and Saving commands)

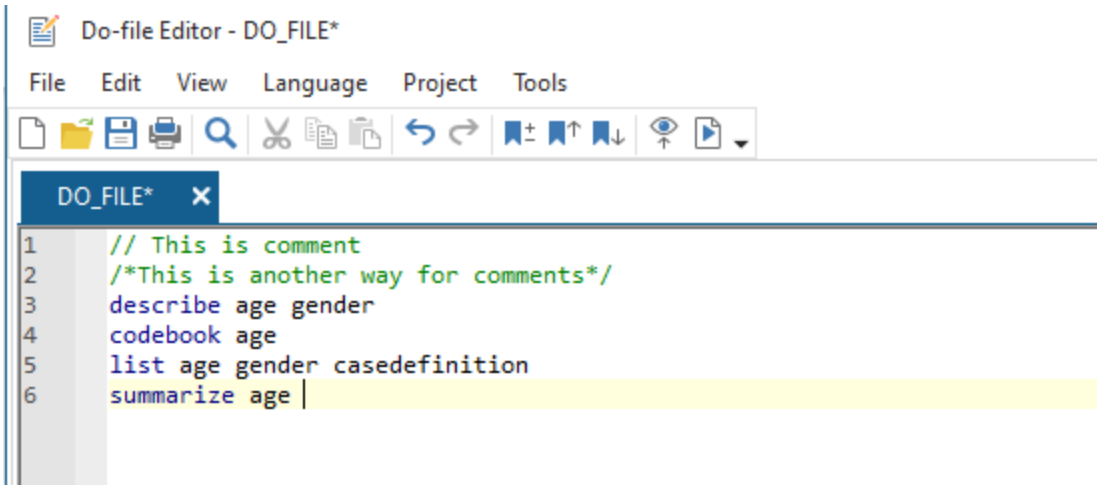
- Create from Editor: Open Stata >> File >> DO: Click here and save as DO_File



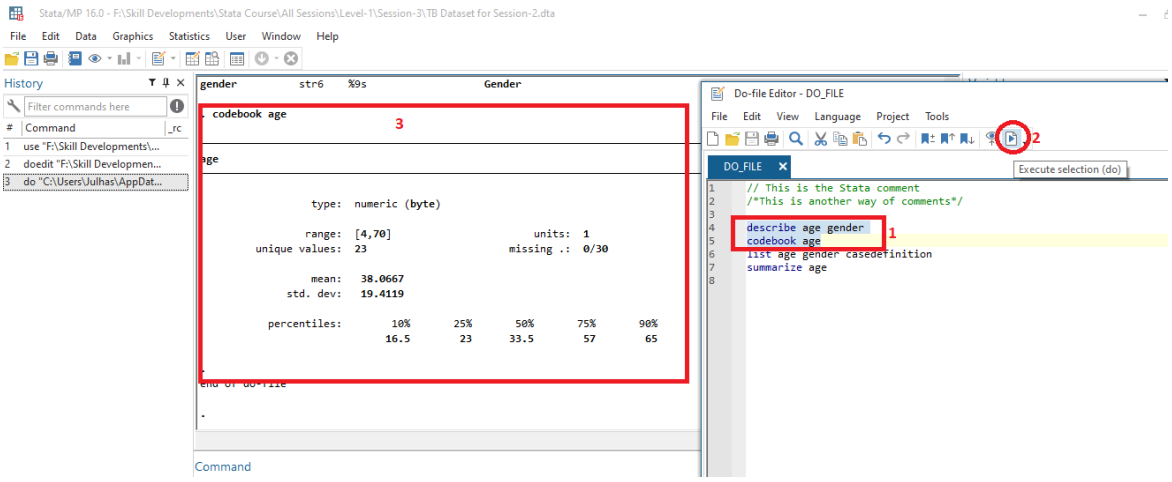
- You can create alternatively from the following menu: Open Stat >> Sub Menu >> Click here and save:



- **Writing and saving code:** You can write and save Stata commads in the DO as:



- **Execute DO file:** Select the commad or commands and click on the following icon and it will execute the command or commands:



1. Select the commands
2. Click here to execute
3. See the results

Next Session:

- Article writing style and Stata result input to your article
- Import DO file
- Edit DO file
- Tabulate
- Append
- Merge
- Concat
- Graphics: Editor and Command

Good Luck!

If you need any support, don't heasitate to let me know.