Updated July 2018

Tabulations and Summary Statistics for One Variable

This set of notes tells how to use Stata to produce one-way frequency tables and to calculate summary descriptive statistics for a single variable. It assumes that you have set Stata up on your computer (see the "Getting Started with Stata" handout), and that you have read in the set of data that you want to analyze (see the "Reading in Stata Format (.dta) Data Files" handout).

In Stata, most tasks can be performed either by issuing commands within the "Stata command" window, **or** by using the menus. These notes illustrate both approaches, using the data file "GSS2016.DTA" (this data file is posted here: https://canvas.harvard.edu/courses/53958).

Obtaining frequency tables

The "tabulate" command produces frequency tables for a single variable. Using the command window, you issue the command

```
tab <varname>
```

where you fill in the variable name for which you want a frequency distribution. For the variable "marital" (marital status) in GSS2016.DTA, you would type the command

```
tab marital
```

This produces the following report in the "Stata Results" window:

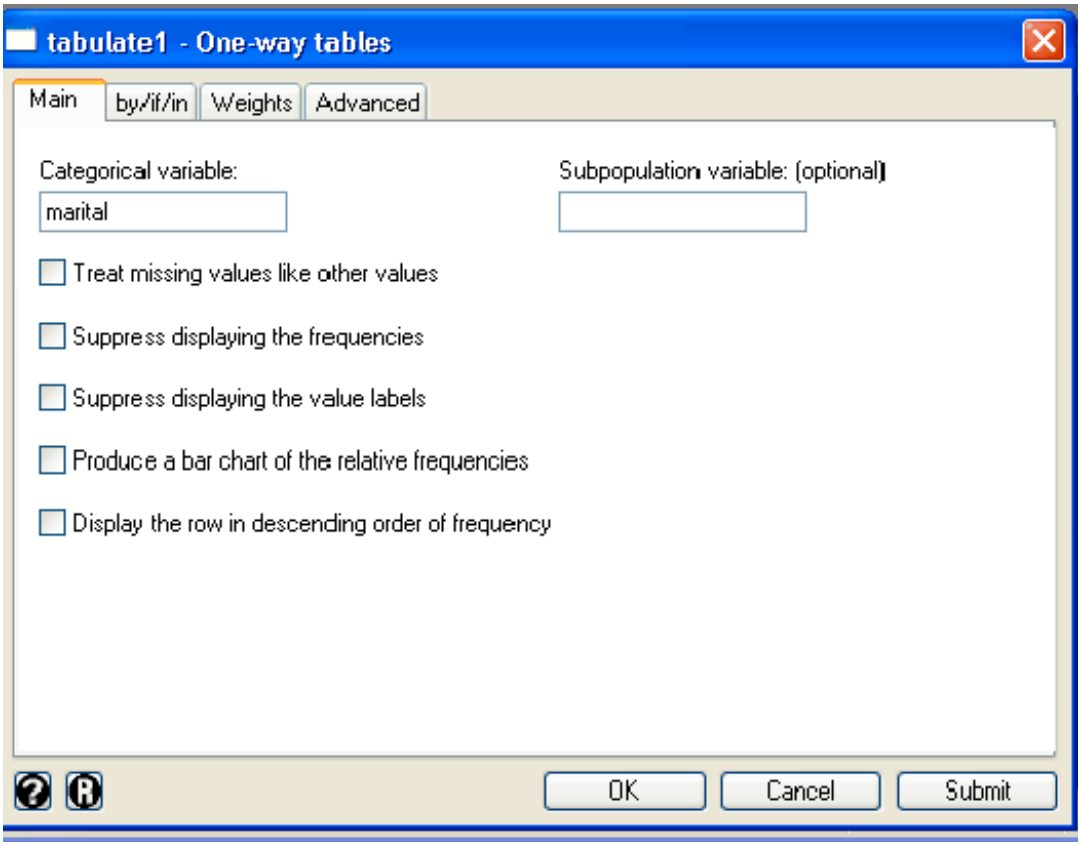
```
. tab marital

      marital |
       status |      Freq.     Percent        Cum.
--------------+-----------------------------------
      married |      1,212       42.29       42.29
      widowed |        251        8.76       51.05
     divorced |        495       17.27       68.32
    separated |        102        3.56       71.88
never married |        806       28.12      100.00
--------------+-----------------------------------
        Total |      2,866      100.00
```

Using the menus, you would proceed as follows

click on "Statistics"
click on "Summaries, tables, and tests"
click on "Frequency tables"
click on "One-way tables"

A window like this will then open up:



You fill in the variable you want frequencies for in the "Categorical variable:" box either by typing the variable name or selecting it from a list that appears in the drop-down menu, and click "OK". The frequency distribution shown above will then appear in the "Stata Results" window. You can check the boxes for the "options" you want; checking the fourth box gives you a plot in place of a table.

If you want to get frequencies for several variables at once, you can use the "tab1" command (note that the spelling is: "tab-one" NOT "tab-el"). For instance

```
tab1 marital sex childs
```

produces separate frequency distributions for marital status, sex, and number of children.

A similar menu-based approach to obtaining multiple sets of frequencies is available. To use it, click on "Multiple one-way tables" after "Frequency tables" in the above instructions.

Obtaining summary statistics

Summary statistics include the mean, standard deviation, and so on. They are obtained using the "summarize" command in Stata. To obtain summary statistics for education using the command window, one would issue the following command:

```
summ educ
```

which produces this output:

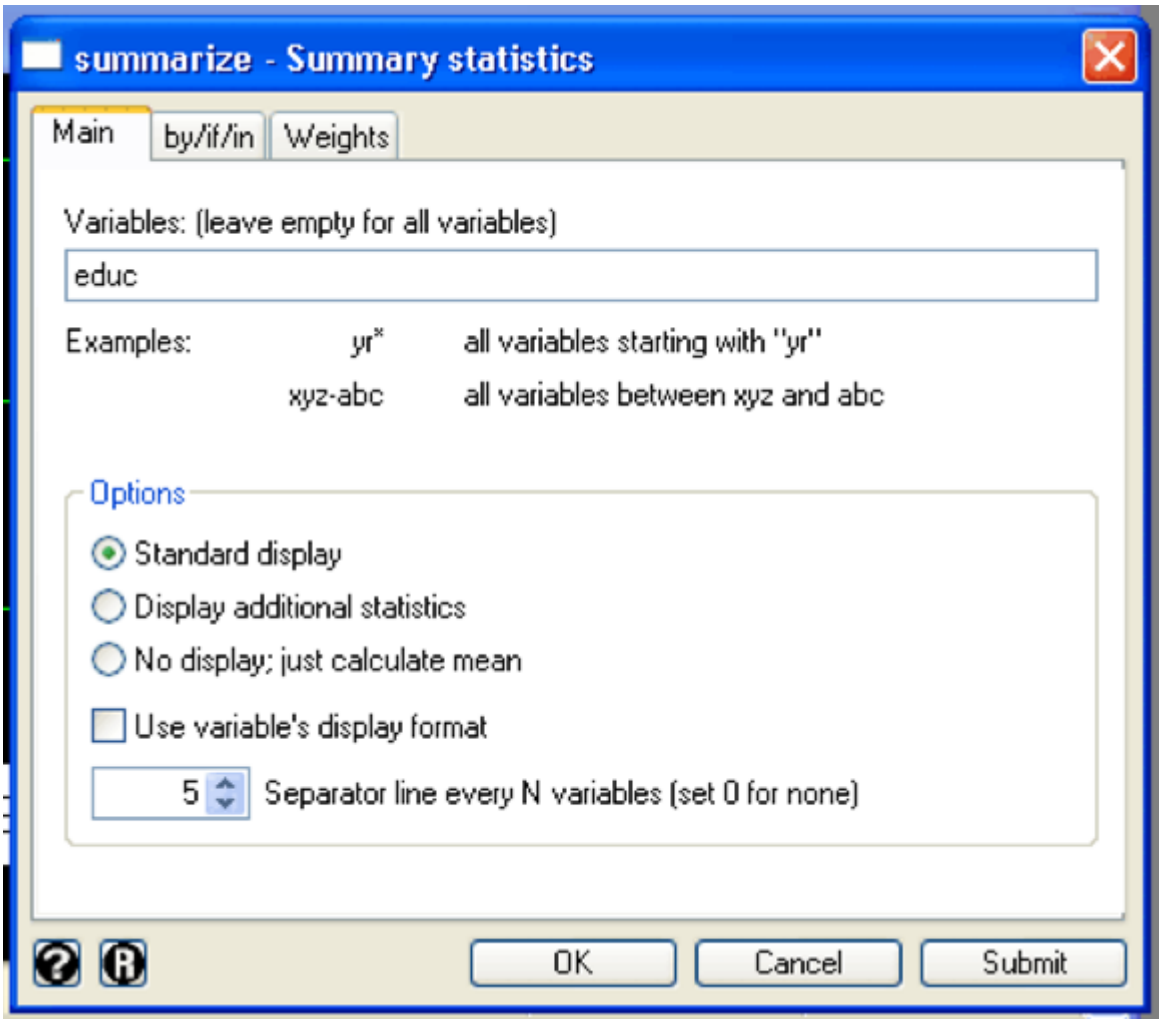
```
. summ educ

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
        educ |      2,858    13.73723    2.963886          0         20
```

With the menu-based approach, you proceed as follows:

click on "Statistics"
click on "Summaries, Tables, and Tests"
click on "Summary and Descriptive Statistics"
click on "Summary Statistics"

A window like the one on the following page opens:



Fill in the variable(s) of interest to you in the "Variables:" window and click "OK"; the report shown above will then appear.

The statistics reported for the "Standard display" are rather few, including only the mean, standard deviation, minimum value, and maximum value. If you want a fuller report with more statistics (including several percentiles), add the "detail" option to the "summarize" command, or click the "Display additional statistics" button in the above window. For example,

```
summ educ, detail
```

yields an elaborated report for education shown on the next page:

```
. summ educ, detail

              highest year of school completed
-------------------------------------------------------------
      Percentiles      Smallest
 1%            6              0
 5%            9              0
10%           11              1       Obs                2,858
25%           12              1       Sum of Wgt.        2,858

50%           13                      Mean           13.73723
                      Largest         Std. Dev.      2.963886
75%           16             20
90%           18             20       Variance        8.78462
95%           19             20       Skewness       -.1855541
99%           20             20       Kurtosis        3.820384
```