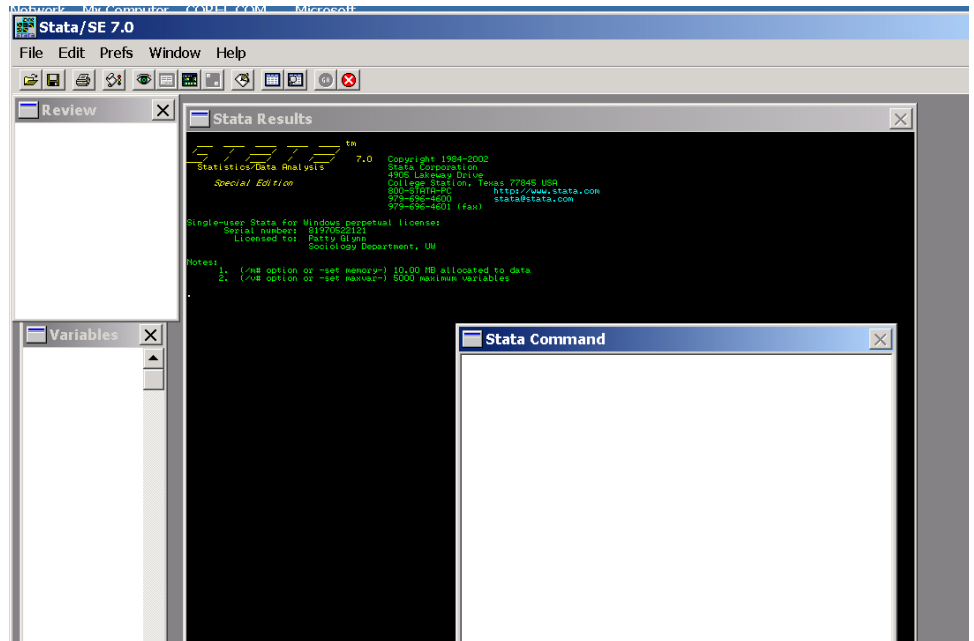


A Little Stata Session¹

Following is a very basic introduction to Stata. I highly recommend the tutorial available at: <http://www.ats.ucla.edu/stat/stata/default.htm>

When you bring up Stata, you will see several windows. You issue your commands from the **Command** window by typing a command, and then pressing Enter. You will see your results in the **Results** window. The **Review** window stores commands that you have already issued. If you want to submit the command again, you can click on it, and this will send it over to the Command window. You may alter it if you want, and then press Enter to submit it. The Variables window lists all of the variables in your current data set. You may click on any variable to send the variable name over to the Command window.



I usually start my Stata session by opening a log file. This file will have a record of all of the commands that I issue, and the results from Stata. You may open a log file with point and click, or by issuing commands. There are two types of logs - a formatted ".smcl" file, or an ASCII ".log" file. My preference is for the ASCII .log file because I can easily bring it into a word processor. The point and click method of opening a log file can be found by clicking on the icon that looks like a scroll. You can also open a log file with the command:



log using "full path for log" For example: `log using "D:\learnStata\first.log"`

For many data sets and procedures, you will need to increase the amount of memory that Stata allocates. The amount of memory that you can allocate will depend on the memory in the machine you are using.

Getting help in Stata: You may click on Help on the menu bar. And, you may issue the command "whelp" or "help". The difference is that "whelp" brings up a separate window with the information, while "help" will put the information in your results window. For example issuing the command: `whelp log` will provide you with information about the "log" command.

The rest of this document will be a few Stata commands (in bold), and abbreviated output.

```
log using "D:\learnStata\first.log"
log: D:\learnStata\first.log
log type: text
opened on: 11 Sep 2002, 10:55:16
```

¹Prepared by Patty Glynn, University of Washington, C:\all\help\helpnew\stata_session.wpd. 9/11/02, modified 4/10/03, 5/13/03, 5/28/03.

```
. set memory 60000k
```

Current memory allocation

| settable | current value | description | memory usage (1M = 1024k) |
|-------------|---------------|-------------------------|------------------------------|
| set maxvar | 5000 | max. variables allowed | 1.733M |
| set memory | 5000k | max. data space | 4.883M |
| set matsize | 400 | max. RHS vars in models | 1.254M |
| | | | ----- |
| | | | 7.870M |

```
. use "D:\learnStata\1991 u.s. general social survey.dta", clear
```

```
. sum
```

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|----------|------|----------|-----------|-----|-----|
| sex | 1517 | 1.580751 | .4935988 | 1 | 2 |
| race | 1517 | 1.199077 | .4734917 | 1 | 3 |
| region | 1517 | 1.831246 | .835713 | 1 | 3 |
| happy | 1504 | 1.799202 | .6165596 | 1 | 3 |
| life | 980 | 1.59898 | .5693921 | 1 | 3 |
| sibs | 1505 | 3.931561 | 3.046618 | 0 | 26 |
| childs | 1509 | 1.901259 | 1.764724 | 0 | 8 |

```
. inspect
```

```
sex: respondent's sex
```

Number of Observations

| | | | Total | Integers | Non-Integers |
|--------|---|----------|-------|----------|--------------|
| | # | Negative | - | - | - |
| | # | Zero | - | - | - |
| | # | Positive | 1517 | 1517 | - |
| | # | | ----- | ----- | ----- |
| | # | Total | 1517 | 1517 | - |
| | # | Missing | - | - | - |
| +----- | | | ----- | | |
| 1 | 2 | | 1517 | | |

(2 unique values)

sex is labeled and all values are documented in the label.

```
race: race of respondent
```

Number of Observations

| | | | Total | Integers | Non-Integers |
|--------|---|----------|-------|----------|--------------|
| | # | Negative | - | - | - |
| | # | Zero | - | - | - |
| | # | Positive | 1517 | 1517 | - |
| | # | | ----- | ----- | ----- |
| | # | Total | 1517 | 1517 | - |
| | # | Missing | - | - | - |
| +----- | | | ----- | | |
| 1 | 3 | | 1517 | | |

(3 unique values)

OUTPUT ABBREVIATED

. tabulate sex

| respondent's sex | Freq. | Percent | Cum. |
|------------------|-------|---------|--------|
| male | 636 | 41.92 | 41.92 |
| female | 881 | 58.08 | 100.00 |
| Total | 1517 | 100.00 | |

. tabulate race

| race of respondent | Freq. | Percent | Cum. |
|--------------------|-------|---------|--------|
| white | 1264 | 83.32 | 83.32 |
| black | 204 | 13.45 | 96.77 |
| other | 49 | 3.23 | 100.00 |
| Total | 1517 | 100.00 | |

. tabulate race if (sex ==1)

| race of respondent | Freq. | Percent | Cum. |
|--------------------|-------|---------|--------|
| white | 545 | 85.69 | 85.69 |
| black | 71 | 11.16 | 96.86 |
| other | 20 | 3.14 | 100.00 |
| Total | 636 | 100.00 | |

. tabulate race if (sex ==2)

| race of respondent | Freq. | Percent | Cum. |
|--------------------|-------|---------|--------|
| white | 719 | 81.61 | 81.61 |
| black | 133 | 15.10 | 96.71 |
| other | 29 | 3.29 | 100.00 |
| Total | 881 | 100.00 | |

. tabulate region

| region of the united states | Freq. | Percent | Cum. |
|-----------------------------|-------|---------|--------|
| north east | 679 | 44.76 | 44.76 |
| south east | 415 | 27.36 | 72.12 |
| west | 423 | 27.88 | 100.00 |
| Total | 1517 | 100.00 | |

. tab1 sex race region

-> tabulation of sex

| respondent's sex | Freq. | Percent | Cum. |
|------------------|-------|---------|--------|
| male | 636 | 41.92 | 41.92 |
| female | 881 | 58.08 | 100.00 |
| Total | 1517 | 100.00 | |

-> tabulation of race

| race of respondent | Freq. | Percent | Cum. |
|--------------------|-------|---------|--------|
| white | 1264 | 83.32 | 83.32 |
| black | 204 | 13.45 | 96.77 |
| other | 49 | 3.23 | 100.00 |
| Total | 1517 | 100.00 | |

-> tabulation of region

| region of the united states | Freq. | Percent | Cum. |
|-----------------------------|-------|---------|--------|
| north east | 679 | 44.76 | 44.76 |
| south east | 415 | 27.36 | 72.12 |
| west | 423 | 27.88 | 100.00 |
| Total | 1517 | 100.00 | |

. tabulate region, plot

| region of the united states | Freq. | |
|-----------------------------|-------|-------|
| north east | 679 | ***** |
| south east | 415 | ***** |
| west | 423 | ***** |
| Total | 1517 | |

. tabulate sex race

| respondent's sex | race of respondent | | | Total |
|------------------|--------------------|-------|-------|-------|
| | white | black | other | |
| male | 545 | 71 | 20 | 636 |
| female | 719 | 133 | 29 | 881 |
| Total | 1264 | 204 | 49 | 1517 |

. codebook

```
sex ----- respondent's sex
      type: numeric (float)
      label: sex

      range: [1,2]                units: 1
unique values: 2                  coded missing: 0 / 1517

      tabulation:  Freq.  Numeric  Label
                   636    1    male
                   881    2    female
```

```
race ----- race of respondent
      type: numeric (float)
      label: race

      range: [1,3]                units: 1
unique values: 3                  coded missing: 0 / 1517

      tabulation:  Freq.  Numeric  Label
                   1264   1    white
                   204   2    black
                   49    3    other
```

```
region ----- region of the united states
      type: numeric (float)
      label: region

      range: [1,3]                units: 1
unique values: 3                  coded missing: 0 / 1517

      tabulation:  Freq.  Numeric  Label
                   679    1    north east
                   415    2    south east
                   423    3    west
```

. tabulate sex race

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 545 | 71 | 20 | 636 |
| female | 719 | 133 | 29 | 881 |
| Total | 1264 | 204 | 49 | 1517 |

. tabulate sex race if (region==1)

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 248 | 28 | 5 | 281 |
| female | 334 | 54 | 10 | 398 |
| Total | 582 | 82 | 15 | 679 |

```
. tabulate sex race if (region==2)
```

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 138 | 33 | 6 | 177 |
| female | 169 | 61 | 8 | 238 |
| Total | 307 | 94 | 14 | 415 |

```
. tabulate sex race if (region==3)
```

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 159 | 10 | 9 | 178 |
| female | 216 | 18 | 11 | 245 |
| Total | 375 | 28 | 20 | 423 |

```
. sort region
```

```
. by region: tabulate sex race
```

```
-> region = north eas
```

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 248 | 28 | 5 | 281 |
| female | 334 | 54 | 10 | 398 |
| Total | 582 | 82 | 15 | 679 |

```
-> region = south eas
```

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 138 | 33 | 6 | 177 |
| female | 169 | 61 | 8 | 238 |
| Total | 307 | 94 | 14 | 415 |

```
-> region = west
```

| respondent | race of respondent | | | Total |
|------------|--------------------|-------|-------|-------|
| 's sex | white | black | other | |
| male | 159 | 10 | 9 | 178 |
| female | 216 | 18 | 11 | 245 |
| Total | 375 | 28 | 20 | 423 |

```
# Create variable called hlthavg1 = mean average the variables within the parentheses.
```

```
. egen hlthavg1 = rmean( hlth1 hlth2 hlth3 hlth4)  
(503 missing values generated)
```

```
. exit
```