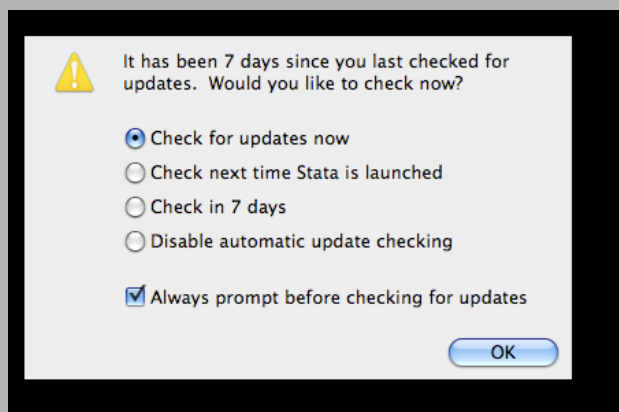


Stata v 12 Illustration First Session

Launch Stata

PC Users	Mac Users
<p>___ START > ALL PROGRAMS > Stata; or ___ Double click on the Stata icon on your desktop</p>	<p>___ APPLICATIONS > STATA folder > Stata; or ___ Double click on the Stata icon on your dock</p>

You might see something like the following:



Stata will ask you if you would like to check for updates. **Click OK.**

Tip! Checking for updates is a good idea. And it doesn't take long.

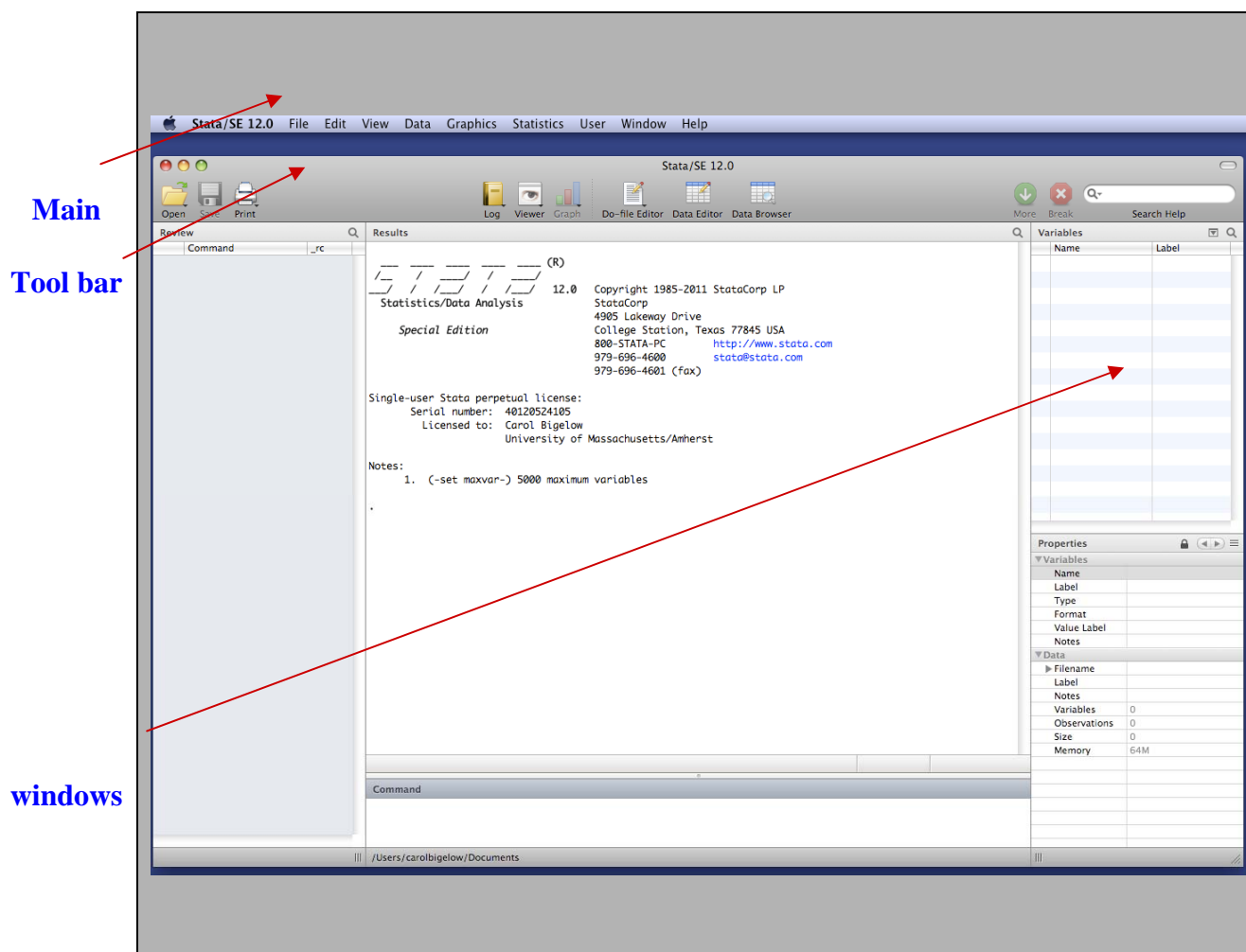
If this check is not done for you automatically, you can do it manually as follows from the main toolbar at top:

HELP > OFFICIAL UPDATES

Introduction to the Windows in Stata

When Stata starts up, your screen will display the following (unless you have changed your windows preferences):

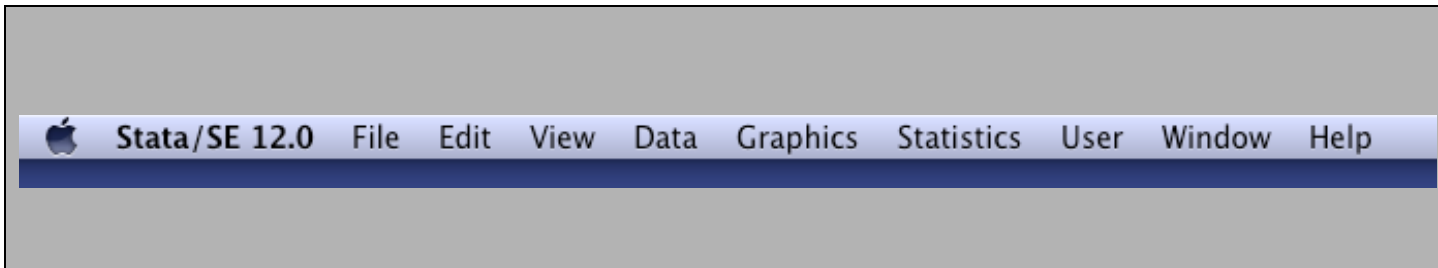
- **main menu:** This is the narrow toolbar located at the top of your screen,
- **tool bar:** This is located just below main menu, plus
- **windows:** In this picture, you see the variables window.



Main Menu

Tip - To obtain a description of each icon, simply place your mouse over the icon. A description should appear.

The main menu is similar in layout to the main menu in many software programs. Each selection (“File”, “Edit”, etc) produces a drop down menu from which you can do additional, related, selections

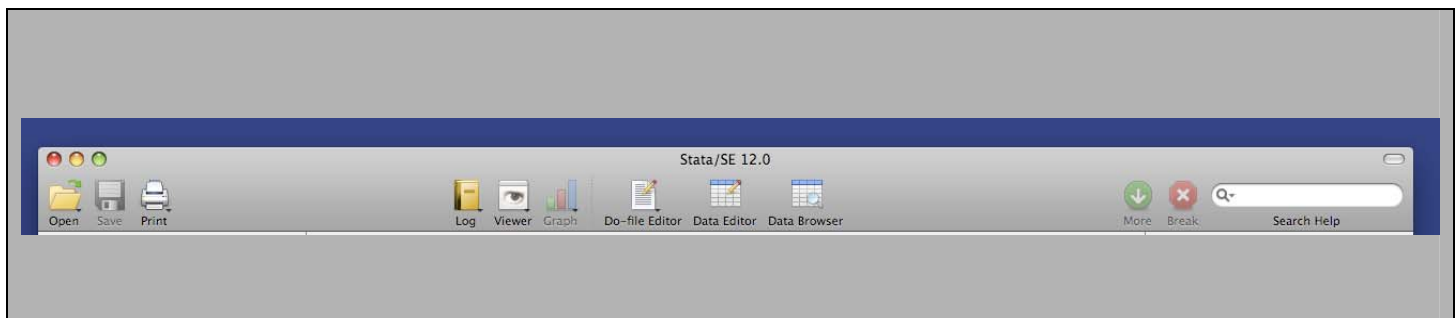


Key to Main Menu

	Drop Down Menus
File	<ul style="list-style-type: none"> • Open: open/use a stata data set • Save/Save as: save the current stata data set • Do: Execute a stata do-file • Filename: Copy a file to the command line • Print: Print log or graph • Exit: Quit Stata
Edit	<ul style="list-style-type: none"> • Copy/Paste: use to copy /paste of text among command, results and log windows • Copy/Table: use to copy a table in the results window to another file
View	Click here to view and manage data, add variable labels, change orientation of windows, view and modify graphs, etc.
Data	Click here to browse or edit data. From here you can use menus to obtain variable labels and discrete value code labels.
Graphics	Click here to obtain menu driven graphs (eg – histogram, XY scatter, etc)






Statistics	Click here to obtain menu driven statistical analyses (eg – normal theory linear regression, analysis of variance, logistic regression, etc)
User	Click here for menus of user-supplied stata commands
Window	Click here to navigate among the windows
Help	Click here for help.

Toolbar



The default arrangement of windows and toolbars in Stata includes a tool bar located below the main menu. It provides short cuts to some frequently used facilities. Many of the icons, like the icons in the main menu bar, produce drop down menus from which you can make additional, related, selections,.

Key to Toolbar

Tool	Description
	Open Use to open a stata data set for use in Stata
	Save Use to save a stata data set. See also the drop down menu for the Save As utility
	Print Use to print the contents of the currently active window. If you want to print the contents of a different window, press and hold the “print” tool to select the window you wish to print.
	Log Use to create, pause, resume or exit a log of your session. This is an alternative to the FILE > LOG instruction from the main menu.
	Viewer The viewer tool is used, primarily, as a resource for obtaining help.

Key to Toolbar – *continued*

Tool	Description
	<p>Bring graph window to front Click here to bring the graph window to the front. See also the drop down menu for specific graph selections.</p>
	<p>Do File Editor Click here to launch the do file editor. This is the window you activate to write your own Stata-do files</p>
	<p>Data Editor Click here to create a new data set if there is no current data set or, to edit the current data set if it exists. Tip – Do not use the data editor to make changes to your data set. Instead, make changes using commands in a do-file or log so that you save a record.</p>
	<p>Data Browser Click here to view the data , without risk of inadvertently making changes.</p>
	<p>More This tells stata to continue processing commands when it has paused.</p>
	<p>Break This tells stata to stop the current command(s) it is executing</p>

Five Main Windows

The five main windows are: (1) “command”, (2) “results”, (3) “review”, (4) “variables” and (5) “properties”. Stata provides other windows as well: Viewer, Data Editor, Data Browser, Do-file Editor, Graph, and Graph Editor. More on these later.

Key to Five Main Windows

	Window	Description
1	Command	This is where you type your command
2	Results	This window shows the output from commands, including error messages
3	Review	This window is a record of your previous commands. You can scroll through this window and highlight previous commands that you want to execute again.
4	Variables	This is a handy listing of your variables and labels.
5	Properties	<i>Tip!</i> This is new to version 12. Use this to manage your variables – names, labels, notes, formats, and storage types

Tips for the Use of Windows

To navigate between windows and to select the window you wish to make active	From the main menu at top: WINDOW From drop down menu: choose window that you want to be active
To resize a window	Position cursor at the edge of the window. Click and drag.
Stata version 12 lets you choose the arrangement and order of the windows on your screen.	From the main menu at top: VIEW> LAYOUT From the drop down menu: Choose the orientation and order of the windows on your screen.

How to use this illustration

In the pages that follow, you will see **green**, **black**, and **blue**. These colorations are things that I did using MS Word.

Green: These are comments that I typed into the STATA command line. Notice that they all begin with an asterisk. I encourage you to use comments liberally. It will save you a lot of grief later when trying to recall what you did and why.

Tip! Each command is introduced to you twice, once as a “generic” and presented in green, and once in illustration of its actual use in black.

Black: Bold black are actual Stata commands that I typed and executed. Note – You do *not* type the leading period.

Blue: I’ve colored in BLUE the output that Stata produces so that you can compare it with the output you get

.***** Open log

```
. log using /users/carolbigelow/Desktop/junk.log, replace
      name: <unnamed>
      log: /users/carolbigelow/Desktop/junk.log
      log type: text
opened on: 3 Feb 2012, 10:56:12

. ***** Set working directory, access data from website, save copy to working directory
. cd/Users/carolbigelow/Desktop
/Users/carolbigelow/Desktop

. use "http://www.pauldickman.com/survival/ivf"
(In Vitro Fertilization data)

. save ivf, replace
file ivf.dta saved
```

```
.
. *****
. ***** View structure of data, compact codebook for all data, view labelbook
. describe
```

Contains data from ivf.dta

```
obs:          641                In Vitro Fertilization data
vars:          6                3 Feb 2012 10:56
size:          8,333
```

```
-----
      storage  display  value
variable name  type    format  label    variable label
-----
id             float   %9.0g           identity number
matage        byte    %8.0g           maternal age (years)
hyp           byte    %8.0g           hypertension (1=yes, 0=no)
gestwks       float   %9.0g           gestational age (weeks)
sex           byte    %8.0g           sex of the baby
bweight       int     %8.0g           birthweight (g)
-----
```

Sorted by: id

```
. codebook, compact
```

```
Variable  Obs Unique      Mean      Min      Max  Label
```

```
-----
id          641    641      321      1    641  identity number
matage     641     21  33.97192   23    43  maternal age (years)
hyp        639     2   .1392801    0     1  hypertension (1=yes, 0=no)
gestwks   641    177  38.68725  24.69  42.35  gestational age (weeks)
sex        641     2   1.49142     1     2  sex of the baby
bweight   641    295  3129.137   630   4650  birthweight (g)
-----
```

```
. labelbook
```

```
value label sex
```

```
-----
      values                                labels
      range: [1,2]                          string length: [4,6]
      N: 2                                    unique at full length: yes
      gaps: no                               unique at length 12: yes
      missing .*: 0                          null string: no
                                           leading/trailing blanks: no
                                           numeric -> numeric: no

definition
      1  male
      2  female

variables:  sex
```

```
. ***** Detailed review of one continuous variable (matage) and one discrete variable (hyp)
```

```
. codebook matage
```

```
-----
matage

                                maternal age (years)
-----
      type:  numeric (byte)

      range:  [23,43]                units:  1
      unique values:  21              missing .:  0/641

      mean:    33.9719
      std. dev: 3.87046

      percentiles:    10%    25%    50%    75%    90%
                       29     31     34     37     39
```

```
. codebook hyp
```

```

hyp
      hypertension (1=yes, 0=no)
-----
      type: numeric (byte)
      range: [0,1]
      unique values: 2
      units: 1
      missing .: 2/641
      tabulation: Freq. Value
                   550  0
                   89  1
                   2   .

```

```

.
. *****
. ***** Listing of first 10 observations with separator after 5
. list in 1/10, separator(5)

```

```

+-----+
| id  matage  hyp  gestwks  sex  bweight |
+-----+
1. | 1      33    0    37.74  female  2410 |
2. | 2      34    0    39.15  female  2977 |
3. | 3      34    0    35.72  female  2100 |
4. | 4      30    0    39.29   male   3270 |
5. | 5      35    0    38.38  female  2620 |
+-----+
6. | 6      37    0    37.86   male   3260 |
7. | 7      31    0    40.06   male   3750 |
8. | 8      31    1    34.81  female  1450 |
9. | 9      33    1    38.81   male   3200 |
10. | 10     33    0    40.35  female  3675 |
+-----+

```

```

. ***** Listing of first 50 observations - Nice layout. Each block shows study id at left and only 2
decimal places shown

```

```

. slist in 1/50, id(id) noobs decimal(2)

```

```

id  matage  hyp  gestwks  sex  bweight
1   33     0   37.74    2   2410
2   34     0   39.15    2   2977
3   34     0   35.72    2   2100

```

Some output omitted

```

48  33     0   40.15    1   3570
49  32     0   41.29    1   3660
50  31     0   33.60    2   1920

```

```

.
. ***** SINGLE VARIABLE DESCRIPTIVES
. ***** SINGLE DISCRETE variable. Options after the comma are optional
. tab1 hyp, missing sort plot
-> tabulation of hyp

```

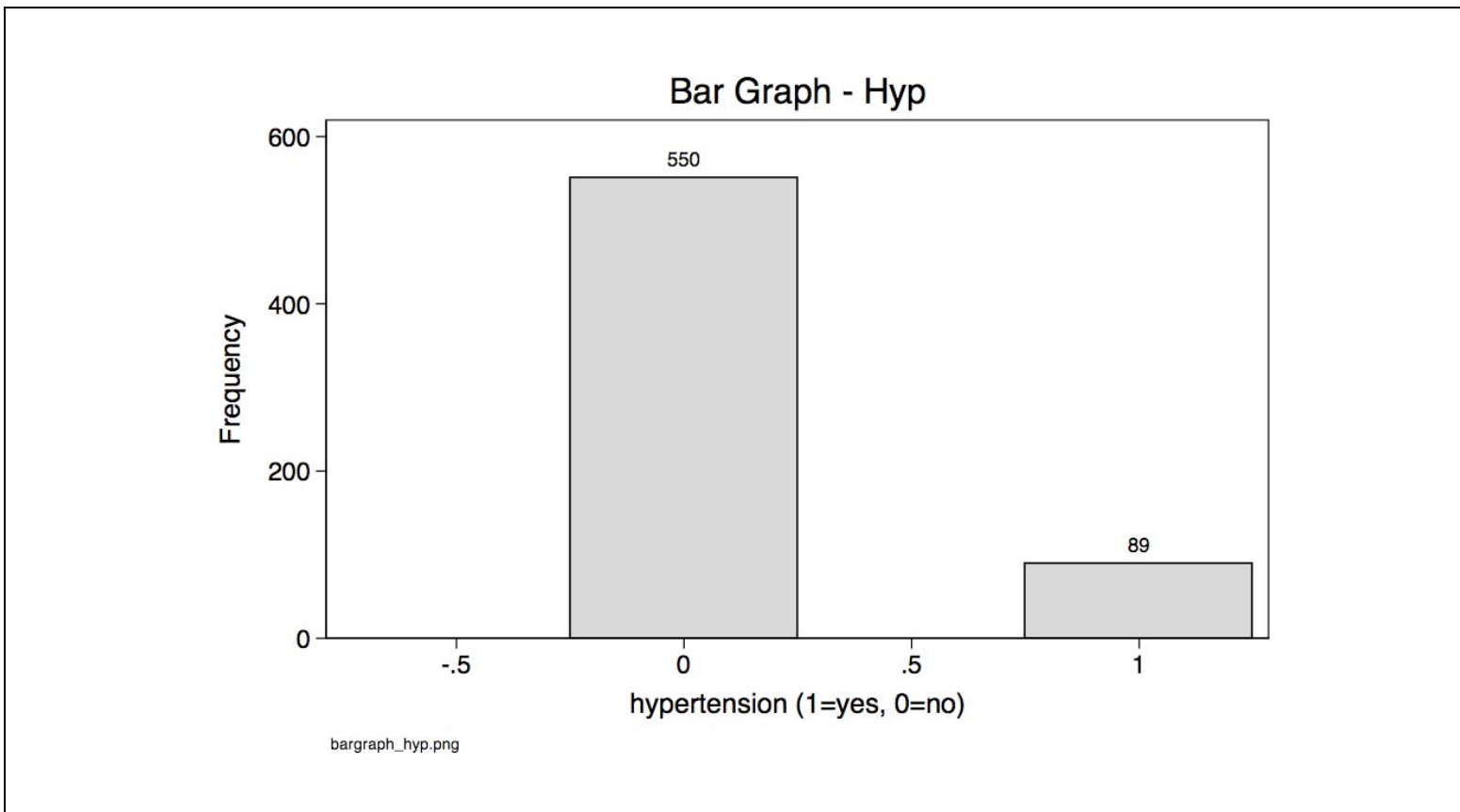
hypertensio		
n (1=yes,		
0=no)	Freq.	
0	550	*****
1	89	*****
.	2	

Total	641	

```

. histogram hyp, discrete frequency addlabels gap(50) title("Bar Graph - Hyp")
caption("bargraph_hyp.png", size(vsmall))
(start=0, width=1)
. graph export "bargraph_hyp.png", replace

```



```

.
.
. ***** SINGLE VARIABLE DESCRIPTIVES

```

```
. ***** SINGLE CONTINUOUS Variable. Lots of choices. Options after comma are optional
. summarize matage, detail
```

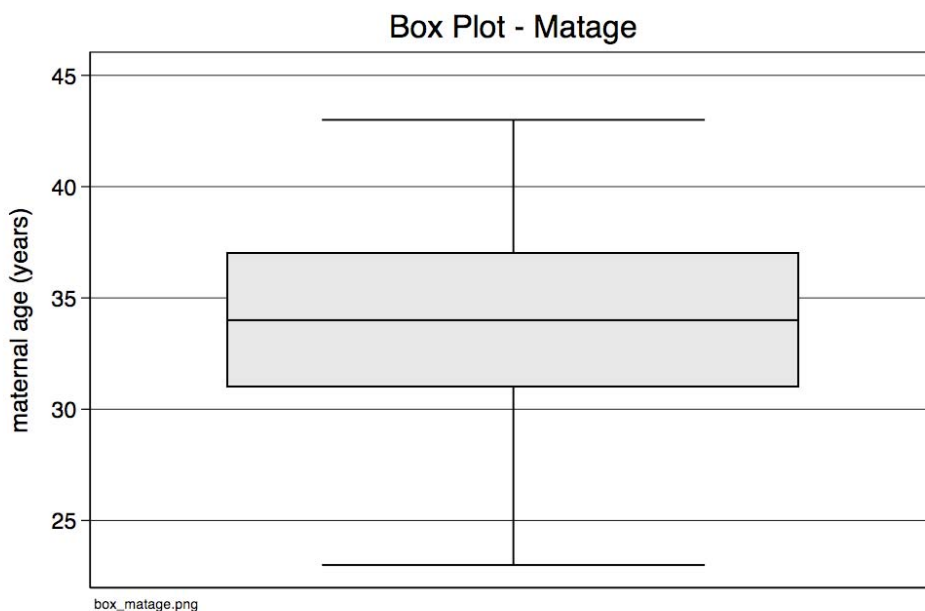
maternal age (years)

Percentiles		Smallest		
1%	25	23		
5%	27	23		
10%	29	24	Obs	641
25%	31	24	Sum of Wgt.	641
50%	34		Mean	33.97192
		Largest	Std. Dev.	3.87046
75%	37	42		
90%	39	43	Variance	14.98046
95%	40	43	Skewness	-.2659265
99%	41	43	Kurtosis	2.523825

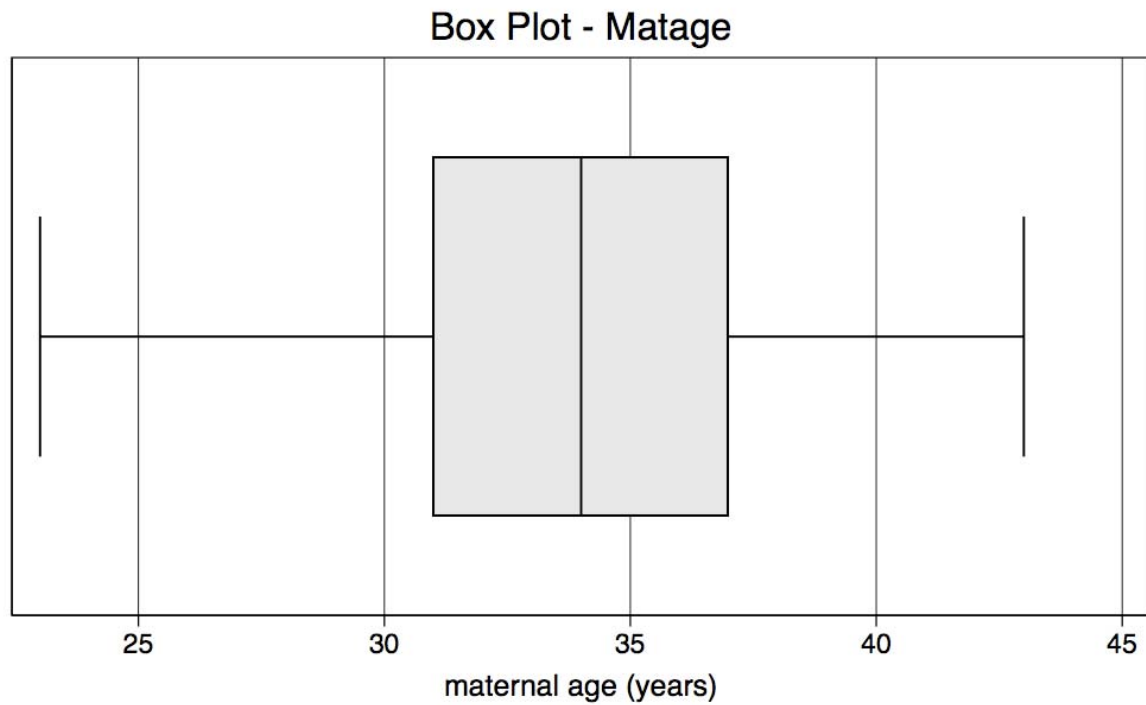
```
. tabstat matage, stat(n mean sd min p25 p50 p75 max semean skewness kurtosis) col(stat) format(%8.2f)
```

variable	N	mean	sd	min	p25	p50	p75	max	se(mean)	skewness	kurtosis
matage	641.00	33.97	3.87	23.00	31.00	34.00	37.00	43.00	0.15	-0.27	2.52

```
. graph box matage, title("Box Plot - Matage") caption("box_matage.png", size(vsmall))
. graph export "box_matage.png", replace
```

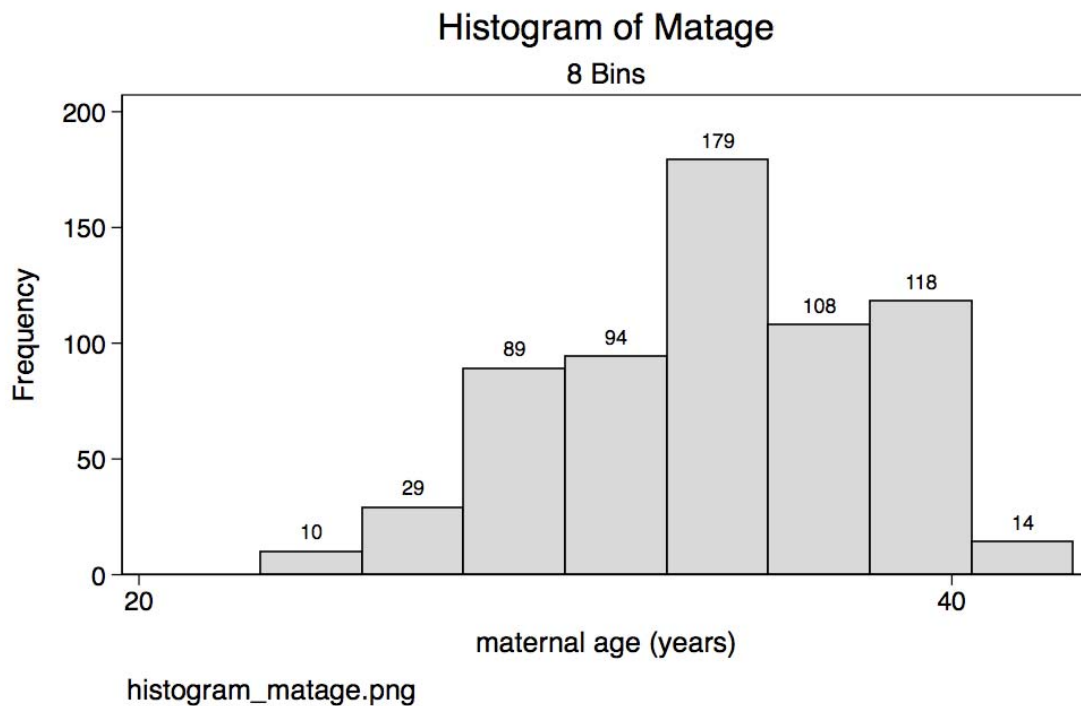


```
. graph hbox matage, title("Box Plot - Matage") caption("hbox_matage.png", size(vsmall))  
. graph export "hbox_matage.png", replace
```



hbox_matage.png

```
. histogram matage, bin(8) frequency addlabels title("Histogram of Matage") subtitle("8 Bins")  
caption("histogram_matage.png", size(vmsall))  
(bin=8, start=23, width=2.5)  
(note: named style vmsall not found in class gsize, default attributes used)  
  
. graph export "histogram_matage.png", replace
```



```
. ***** TWO VARIABLE DESCRIPTIVES
```

```
. ***** TWO DISCRETE variables cross tab, with display of missing
. tab2 sex hyp, missing row column cell
```

-> tabulation of sex by hyp

```
+-----+
| Key      |
|-----+
| frequency|
| row percentage|
| column percentage|
| cell percentage|
+-----+
```

sex of the baby	hypertension (1=yes, 0=no)			Total
	0	1	.	
male	273	52	1	326
	83.74	15.95	0.31	100.00
	49.64	58.43	50.00	50.86
	42.59	8.11	0.16	50.86
female	277	37	1	315
	87.94	11.75	0.32	100.00
	50.36	41.57	50.00	49.14
	43.21	5.77	0.16	49.14
Total	550	89	2	641
	85.80	13.88	0.31	100.00
	100.00	100.00	100.00	100.00
	85.80	13.88	0.31	100.00

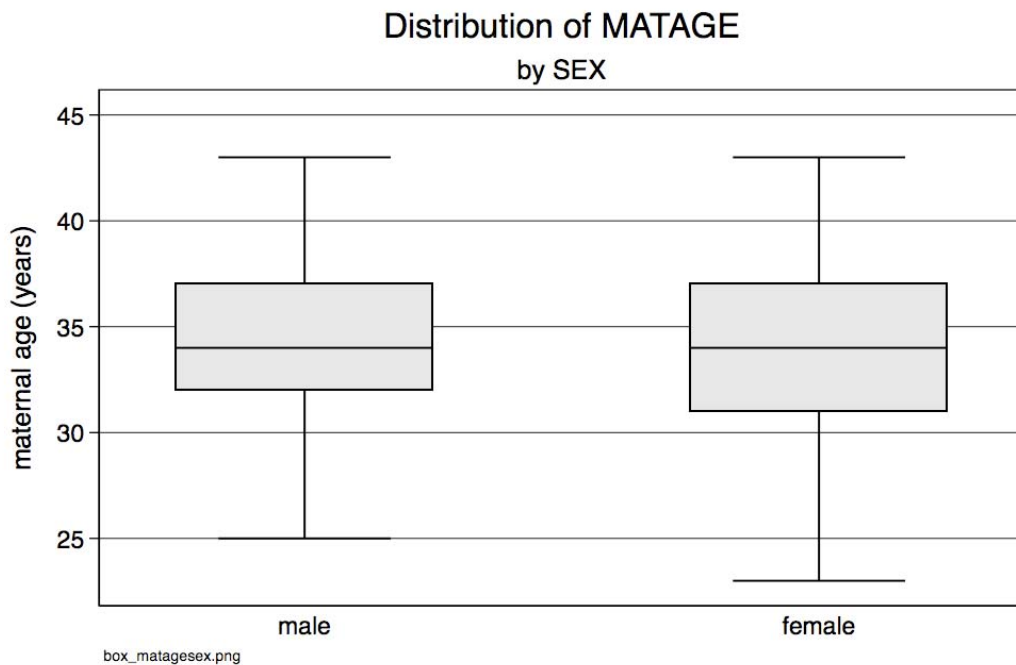
```
. ***** One CONTINUOUS, by levels of one DISCRETE. Must sort by DISCRETE variable first
. sort sex
. tabstat matage, by(sex) stat(n mean sd semean) col(stat) format(%8.2f)
```

Summary for variables: matage
by categories of: sex (sex of the baby)

sex	N	mean	sd	se(mean)
male	326.00	34.21	3.79	0.21
female	315.00	33.72	3.94	0.22
Total	641.00	33.97	3.87	0.15

```
. graph box matage, over(sex) title("Distribution of MATAGE") subtitle("by SEX")
caption("box_matagesex.png", size(vsmall))

. graph export "box_matagesex.png", replace
```



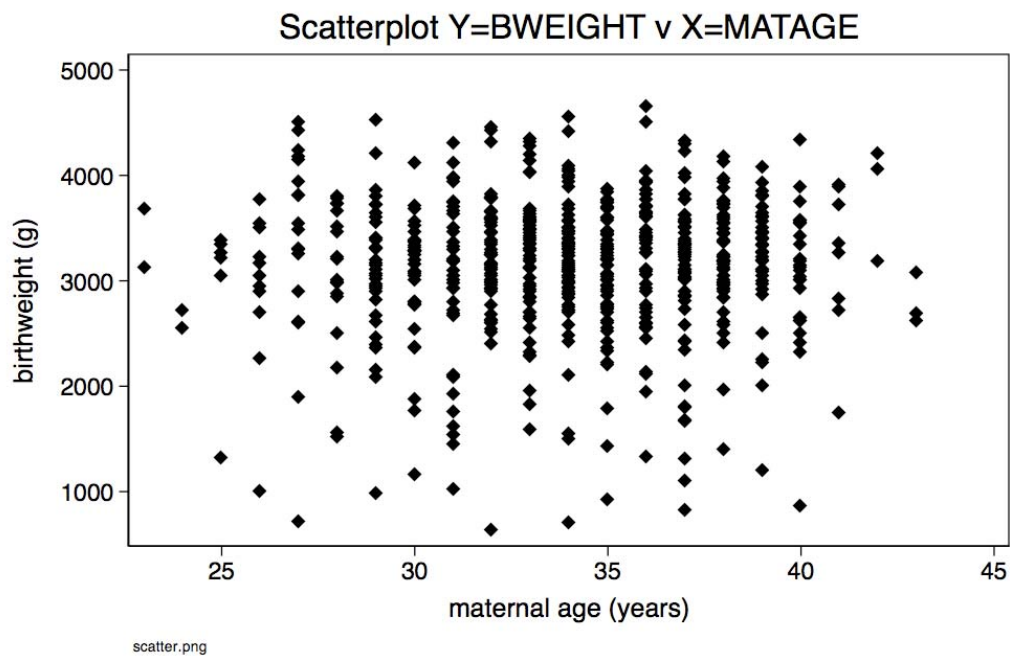
box_matagesex.png

```
.
.
. ***** TWO CONTINUOUS variables with a nice layout
. tabstat matage bweight, stat(n mean sd semean) col(stat) format(%8.2f)
```

variable	N	mean	sd	se(mean)
matage	641.00	33.97	3.87	0.15
bweight	641.00	3129.14	652.78	25.78

```
. graph twoway (scatter bweight matage, symbol(d)), title("Scatterplot Y=BWEIGHT v X=MATAGE")  
caption("scatter.png", size(vsmall))
```

```
. graph export "scatter.png", replace
```



```
. log close  
name: <unnamed>  
log: /users/carolbigelow/Desktop/junk.log  
log type: text  
closed on: 3 Feb 2012, 10:56:25
```
