Introduction to Stata 13 Graphics

Dawn Koffman
Office of Population Research
Princeton University
September 2013
Stata 13 Graphics

Pros:

Many graph types and plot types provided
Multiple plot types may be overlaid
Can easily change overall look of graphs
Same options available for most types of graphs
Very flexible

Cons:

Sometimes slow
Large syntax: 680 page graphics manual!
  Stata 13 Graphics Manual is only available on-line:

Stata Graphics References:

http://data.princeton.edu/stata/Graphics.html, by German Rodriguez
Stata 13 Graphics Manual (may want to start with “graph intro”)
Stata Graphics Syntax

```bash
graph <graphtype>
  graph bar

graph twoway <plottype>
  graph twoway scatter
  graph twoway line
  graph twoway lfit
  graph twoway lfitci

graphs commands may have options
  some options have suboptions or a list of options
    graph twoway scatter var1 var2, xlabel(30(10)100, labsize(small))

appearance of graph defined by graph elements:
  data - marker symbols, lines
  elements within plot region – text, marker labels, line labels
  elements outside plot region – titles, legend, notes, axis labels, tick marks, axis titles
  size and shape of plot region and entire graph
```
Stata Graphics Syntax: A Simple Example

sysuse uslifeexp.dta, clear

graph twoway line le year

    /* OR */

twoway line le year

    /* OR */

line le year
Using Schemes

line le year, scheme(slmono)

line le year, scheme(economist)

/* to see list of scheme names:
graph query, schemes

to change default scheme:
set scheme schemename
*/
Multiple Dependent Variables

```
line le_wmale le_wfemale le_bmale le_bfemale year
```

- **Life expectancy, white males**
- **Life expectancy, white females**
- **Life expectancy, black males**
- **Life expectancy, black females**

Year range from 1900 to 2000.
Adding Text

```
line le_wmale le_wfemale le_bmale le_bfemale year ///
, text(32 1920 "{bf:1918} {it:Influenza} Pandemic", place(3))
```

![Graph showing life expectancy trends from 1900 to 2000. The graph highlights a significant drop in life expectancy around 1918, labeled as the "1918 Influenza Pandemic." Legend indicates lines for white males, white females, black males, and black females.](image-url)
Overlaying Two-Way Plot Types

scatter le year if year >= 1950 || lfit le year if year >= 1950
/* OR */

scatter ///
le year if year >= 1950 ///
|| lfit le year if year >= 1950
/* OR */

twoway ///
(scatter le year if year >= 1950) ///
(lfit le year if year >= 1950)
/* OR */

#delimit ;
twoway
(scatter le year if year >= 1950)
(lfit le year if year >= 1950);
#delimit cr
Overlaying Two-Way Plot Types

scatter le year if year >= 1925 ///
|| lfit le year if year >= 1925 & ///
    year < 1950 ///
|| lfit le year if year >= 1950

/* OR */

twoway ///
(scatter le year if year >= 1925) ///
(lfit le year if year >= 1925 & ///
    year < 1950) ///
(lfit le year if year >= 1950)

/* OR */

#delimit ;
scatter le year if year >= 1925
|| lfit le year if year >= 1925 & year < 1950
|| lfit le year if year >= 1950;
#delimit cr
Overlaying Two-Way Plot Types

#delimit ;
scatter le_male le_female year if year >= 1950
|| lfit le_male year if year >= 1950
|| lfit le_female year if year >= 1950;
#delimit cr
Adding a Title and Removing the Legend

```stata
#delimit ;
scatter le_male le_female year if year >= 1950 || lfit le_male year if year >= 1950 || lfit le_female year if year >= 1950 ,
    title("US Male and Female Life Expectancy, 1950–2000")
text(75 1978 "Female", place(3))
text(68 1978 "Male", place(3))
    legend(off);
#delimit cr
```

US Male and Female Life Expectancy, 1950-2000

```
```
Showing Confidence Intervals, Labelling Axes, Modifying Legend

sysuse lifeexp.dta, clear
#delimit ;
twoway
(1fitci lexp safewater if region == 2) /* North America */
(scatter lexp safewater if region == 2)
, title("Life expectancy at birth by access to safe water, 1998")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(ring(0) pos(5) order(2 "Linear fit" 1 "95% CI"));
#delimit cr
Markers Labels and Subtitles

twoway
(lfitci lexp safegwater if region == 2) /* North America */
(scatter lexp safegwater if region == 2, mlabel(country))
,title("Life expectancy at birth by access to safe water, 1998")
subtitle("North America")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(ring(0) pos(5) order(2 "Linear fit" 1 "95% CI"));

Life expectancy at birth by access to safe water, 1998
North America

Percent of population with access to safe water

Life expectancy at birth

Canada
Jamaica
Puerto Rico
Canada
Dominican Republic
Haiti
Nicaragua
El Salvador
Honduras
Guatemala
Mexico
Panama
Trinidad and Tobago

Linear fit 95% CI
generate pos = 12 if country == "Panama"
replace pos = 12 if country == "Honduras"
replace pos = 10 if country == "Cuba"
replace pos = 9 if country == "Jamaica"
replace pos = 9 if country == "El Salvador"
replace pos = 9 if country == "Trinidad and Tobago"
replace pos = 9 if country == "Dominican Republic"
#delimit ;
twoway
(lfitci lexp safewater if region == 2) /* North America */
(scatter lexp safewater if region == 2
    , mlabel(country) mlabvposition(pos))
, title("Life expectancy at birth by access to safe water, 1998")
subtitle("North America")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(ring(0) pos(5) order(2 "Linear fit" 1 "95% CI"))
plotregion(margin(r+10));
#delimit cr
Position of Marker Labels

#delimit ;
twoway
(scatter lexp safewater if region == 2 | region == 3
, mlabel(country))
,title("Life expectancy at birth by access to safe water, 1998")
subtitle("North and South America")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
plotregion(margin(r+10));
#delimit cr
Position of Marker Labels and Legend Display

```
replace pos = 9 if country == "Argentina"
replace pos = 9 if country == "Canada"
replace pos = 9 if country == "Cuba"
replace pos = 9 if country == "Panama"
replace pos = 9 if country == "Venezuela"
replace pos = 9 if country == "Jamaica"
replace pos = 9 if country == "Dominican Republic"
replace pos = 9 if country == "Ecuador"
replace pos = 9 if country == "El Salvador"
replace pos = 12 if country == "Puerto Rico"
#delimit ;
twoway
(scatter lexp safewater if region == 2
,mlabel(country) mlabvposition(pos))
(scatter lexp safewater if region == 3
,mlabel(country) mlabvposition(pos))
,title("Life expectancy at birth by access to safe water, 1998")
subtitle("North and South America")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(ring(0) pos(5) order(1 "North America" 2 "South America") cols(1));
#delimit cr
```
Marker Size and Symbol, Line Color

twoway
(scatter lexp safewater if region == 2
   ,mlabel(country) mlabvposition(pos) msize(small))
(scatter lexp safewater if region == 3
   ,mlabel(country) mlabvposition(pos) msize(small) msymbol(circle_hollow))
(lfit lexp safewater if region == 2, clcolor(navy))
(lfit lexp safewater if region == 3, clcolor(maroon))
,title("Life expectancy at birth by access to safe water, 1998")
subtitle("North and South America")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(ring(0) pos(5) cols(1) order(1 "North America" 2 "South America"
   3 "North America linear fit" 4 "South America linear fit"));
twoway
(scatter lexp safewater if region == 2
,mlabel(country) mlabvposition(pos) msize(small) mcolor(black) mlabcolor(black))
(scatter lexp safewater if region == 3
,mlabel(country) mlabvposition(pos) msize(small) mcolor(black) mlabcolor(black)
msymbol(circle_hollow))
(lfit lexp safewater if region == 2, clcolor(black))
(lfit lexp safewater if region == 3, clcolor(black) clpattern(dash))
,title("Life expectancy at birth by access to safe water, 1998", color(black))
subtitle("North and South America")
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(ring(0) pos(5) cols(1) order(1 "North America" 2 "South America"
3 "North America linear fit" 4 "South America linear fit"));
By-Graph: Separate Graphs for Each Subset of Data

#delimit ;
twoway scatter lexp safewater, by(region, total) ,ytitle("Life expectancy at birth")
   xtitle("Percent of population with access to safe water"); #delimit cr

Graphs by Region
By-Graph Options

twoway scatter lexp safewater
,by(region,total style(compact)
    title("Life expectancy by access to safe water") note(""))
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water");

Life expectancy by access to safe water

<table>
<thead>
<tr>
<th></th>
<th>Eur &amp; C.Asia</th>
<th>N.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S.A.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Axis Scale, Ticks and Labels

twoway scatter lexp safewater
, by(region,total style(compact)
    title("Life expectancy by access to safe water") note(""))
xscale(range(20 100))
xtick(20(10)100)
xlabel(30(10)100, labsize(small))
xtitle("Percent of population with access to safe water")
ytitle("Life expectancy at birth")
ylabel(55(5)80, angle(0));
Storing Graphs in Memory

twoway
(scatter lexp safegwater if region == 2,
mcolor(black) msize(small)
mlabel(country) mlabvposition(pos) mlabcolor(black))
(lfit lexp safegwater if region == 2, clcolor(black))
nname(north_america, replace)
subtitle("North America", color(black))
ylabel(,angle(0))
ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water")
legend(off);
Storing Graphs in Memory

twoway
(scatter lexp_sa safewater if region == 3, mcolor(black) msize(small) mlabel(country) mlabvposition(pos) mlabcolor(black))
(lfit lexp safewater if region == 3, clcolor(black))
,name(south_america, replace) subtitle("South America", color(black)) ylabel(, angle(0)) ytitle("Life expectancy at birth")
xtitle("Percent of population with access to safe water") legend(off);
Combining Graphs

graph combine north_america south_america,
  title("Life expectancy by access to safe water", color(black)) col(1);

Life expectancy by access to safe water

Life expectancy at birth vs Percent of population with access to safe water.
Combining Graphs

graph combine north_america south_america, title("Life expectancy by access to safe water", color(black)) xcommon ycommon xsize(7) ysize(10.5) col(1);
Saving and Including Stata Graphs

save graph in portable format (format determined by filename extension)

vector formats contain drawing instructions (.wmf .emf .ps .eps .pdf)
resolution independent
work well if graph my be resized

```
graph export north_amERICA.wmf
```

raster formats save graph pixel-by-pixel (.png)
use current resolution
work well if including graph on web pages

```
graph export north_amERICA.png
```

include "portable-format-graph" in Windows application (Word, Powerpoint):

Insert -> Picture -> From File
Using Mata Functions to Add Graphs to Word Document*

create Stata graphs and and use `graph export` to save graphs in portable format

```stata
sysuse uslifeexp
line le year
graph export us_lifeexp_overall.emf, replace

line le_wmale le_wfemale le_bmale le_bfemale year
graph export us_lifeexp_race_gender.emf, replace
```

use Mata functions to:
create Word document
add Stata graphs
save Word document

```mata
mata:
    dh = _docx_new()
    _docx_image_add(dh, "us_lifeexp_overall.emf")
    _docx_image_add(dh, "us_lifeexp_race_gender.emf")
    rc = _docx_save(dh, "us_lifeexp_graphs.docx")
end
```

* New functions available in Stata 13