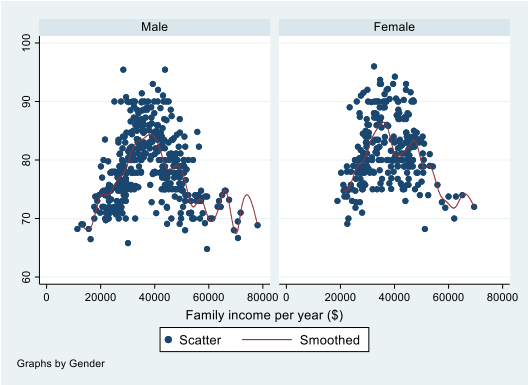
Visualizing data in Stata – Section 7.4
Najib Mozahem

At this point, it would be instructive to discuss how the user can manipulate the look of graphs produced with the **twoway** command. We have already seen how to change the look of bar graphs as well as box and dot plots. Some of the options used in those cases, apply as well to the **twoway** command, but there are also many differences.

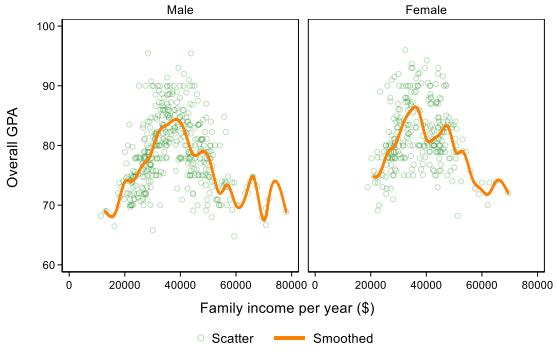
In the last section, we used the following command:

twoway (scatter gpa income) (mspline gpa income), by(gender) legend(label(1 "Scatter") label(2 "Smoothed"))



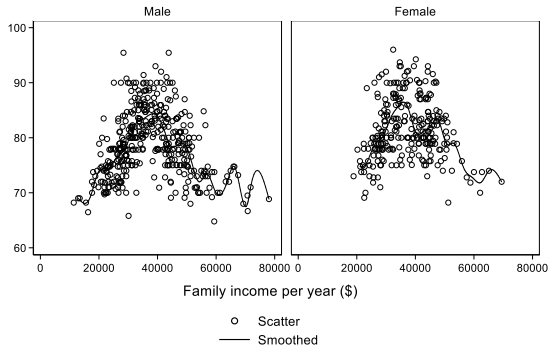
The following command produces the exact same graph but in a more attractive way:

twoway (scatter gpa income, mcolor(green%25)) (mspline gpa income, lcolor(orange) lwidth(thick)), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed") cols(2)) ytitle("Overall GPA") scheme(lean1)



As I said before, what I usually do is that I start with a scheme and modify it the way I like. In this case, I specified the *lean1* scheme, which is one of my favorites because it produces a clean graph. This is the graph using only the scheme as an option:

twoway (scatter gpa income) (mspline gpa income), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed")) scheme(lean1)



I next decided to add some color to it:

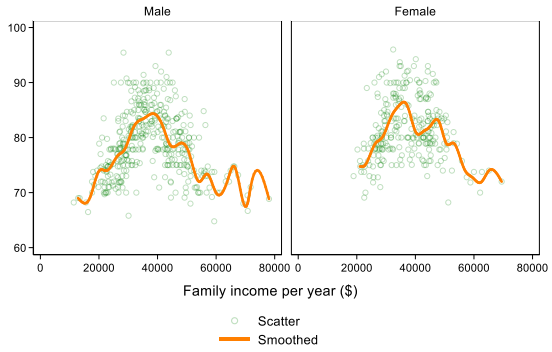
twoway (scatter gpa income, mcolor(green)) (mspline gpa income, lcolor(orange)), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed")) scheme(lean1)

The first graph is a scatter plot, and to change the colors of the markers I use the **mcolor()** option. The second graph is a spline, and to change the color of the line used to draw the spline I use the **lcolor()** option. Next I decided to make the scatter points transparent because what I want is for the spline to take center stage. This is accomplished by specifying the **%** character after the color, and following this with the percent opacity that I want:

twoway (scatter gpa income, mcolor(green%25)) (mspline gpa income, lcolor(orange)), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed")) scheme(lean1)

Now I want to make the spline more visible, so I make the line thicker:

twoway (scatter gpa income, mcolor(green%25)) (mspline gpa income, lcolor(orange) lwidth(thick)), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed")) scheme(lean1)

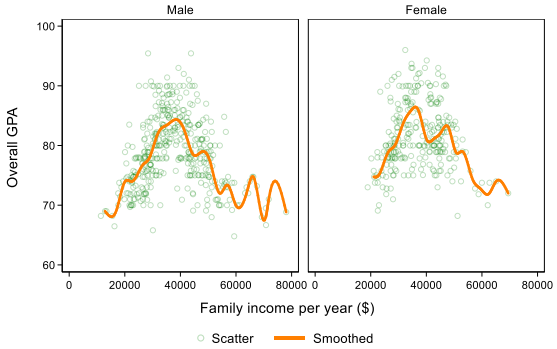


Next I decide to add a title to the y-axis in order to make it clear what the graph is plotting:

twoway (scatter gpa income, mcolor(green%25)) (mspline gpa income, lcolor(orange) lwidth(thick)), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed")) ytitle("Overall GPA") scheme(lean1)

Finally, I want both groups in the legend to appear on the same line nest to each other since there is a lot of empty space at the bottom. This is accomplished by specifying the **cols()**option where I tell Stata that I want the legend to have two columns:

twoway (scatter gpa income, mcolor(green%25)) (mspline gpa income, lcolor(orange) lwidth(thick)), by(gender, note("")) legend(label(1 "Scatter") label(2 "Smoothed") cols(2)) ytitle("Overall GPA") scheme(lean1)



As you can see, customizing a graph makes a huge difference and it personalizes your work to a considerable extent. There are more options that are available in Stata, and the best way to get to know them is to play around with the graphs.