

First step was to make an age by health dataset, so I did:

```
table age, contents (mean health p25 health p75 health) replace
```

I take the new dataset and give the new variables appropriate names (command **rename**). Once we have the health by age dataset, we can make a simple graph with the following command. Note that with the graphing commands, it is nearly always easier to use menus, and STATA produces the graph and the correct command line syntax for you.

```
twoway connected health_mean age, title(Self Reported Health Status)  
subtitle("1 is best, 5 is worst") xtitle(Age) ytitle(Health)
```



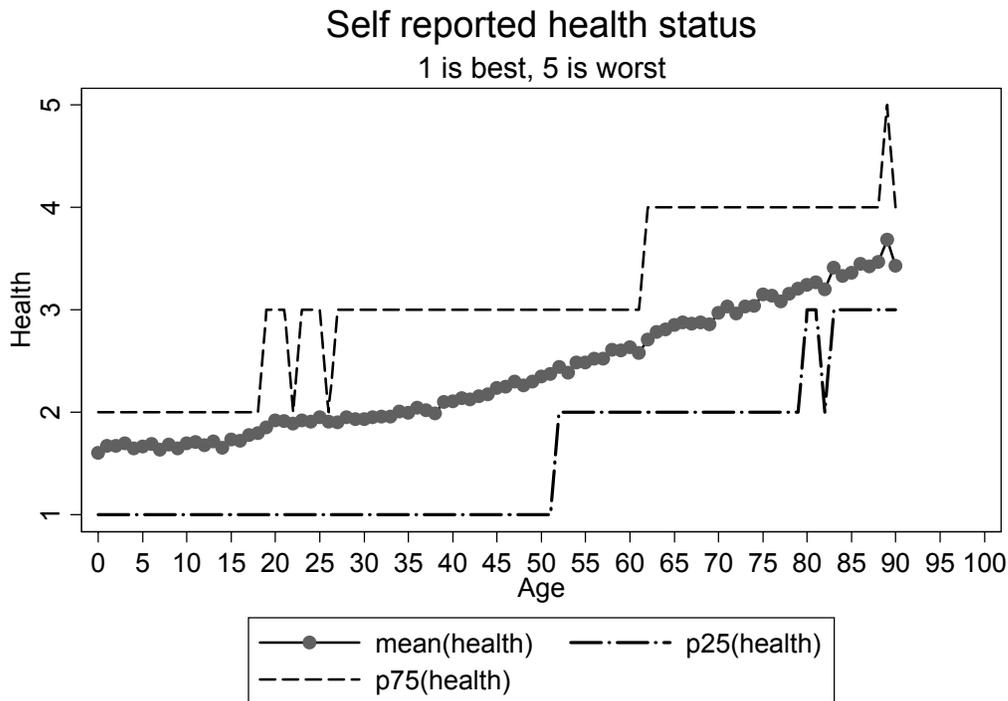
Obviously health status gets worse with age. How about the distribution of health statuses?

We may want to set the X axis scale to have a tick mark every 5 years, and we would like to add the 25 percentile and the 75 percentile of the health distribution, statistics which we included in our table command. To get STATA to give us an X axis scale that goes from 0 to 100 in steps of 5, we input the range 0 (5) 100. Similarly if we wanted to produce this Y axis scale rather than relying on the default, we would ask for 1 (1) 5.

```

twoway (connected health_mean age, clcolor(black) clpat(solid))
      (connected health_p25 age, msymbol(none) clcolor(black)
      clpat(longdash_dot) clwidth(medthick)) (connected health_p75 age,
      msymbol(none) clcolor(black) clpat(dash)), ytitle(Health) xtitle(Age)
      xlabel(0 (5) 100) title(Self reported health status) subtitle("1 is
      best, 5 is worst")

```



Here it seems that the dispersion of health status doesn't change that much with time. The distance between p25 and p75 is 1 category in the early years and in the end, but is often 2 categories in the middle years.

Here I go through a few gyrations to compare men's and women's health by age. Starting with the usual CPS dataset again, and remember that the graph command was entered using the menus because the graph command has too many options to try to remember:

```
table age, contents(mean health sd health) by(sex) replace

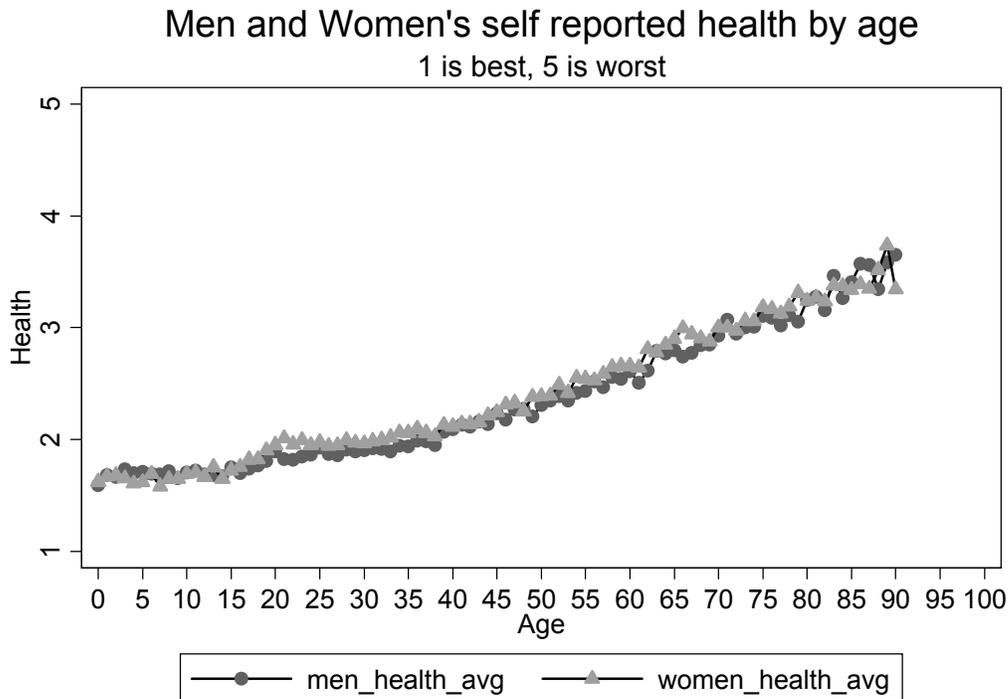
rename table1 health_avg

. rename table2 health_sd

. gen men_health_avg= health_avg if sex==1
(91 missing values generated)

. gen women_health_avg=health_avg if sex==2
(91 missing values generated)

. twoway (connected men_health_avg age, msymbol(circle) clpat(solid))
(connecting women_health_avg age, msymbol(triangle) clpat(solid)),
ytitle(Health) ylabel(1 (1) 5) xtitle(Age) xlabel(0 (5) 100) title(Men
and Women's self reported health by age) subtitle("1 is best, 5 is
worst")
```



It doesn't look from this graph as if there is much difference between men and women in their lifetime health trajectory.

The first thing I did here was create a new poverty level variable with 3 rather than 14 levels. Then I made the summary table. Then instead of creating new variables for health by poverty level, I used the 'if' syntax in the graph making dialog to make the selection.

```

gen new_pov=1 if povl1<6
(104051 missing values generated)

. replace new_pov=2 if povl1>5 & povl1<10
(37152 real changes made)

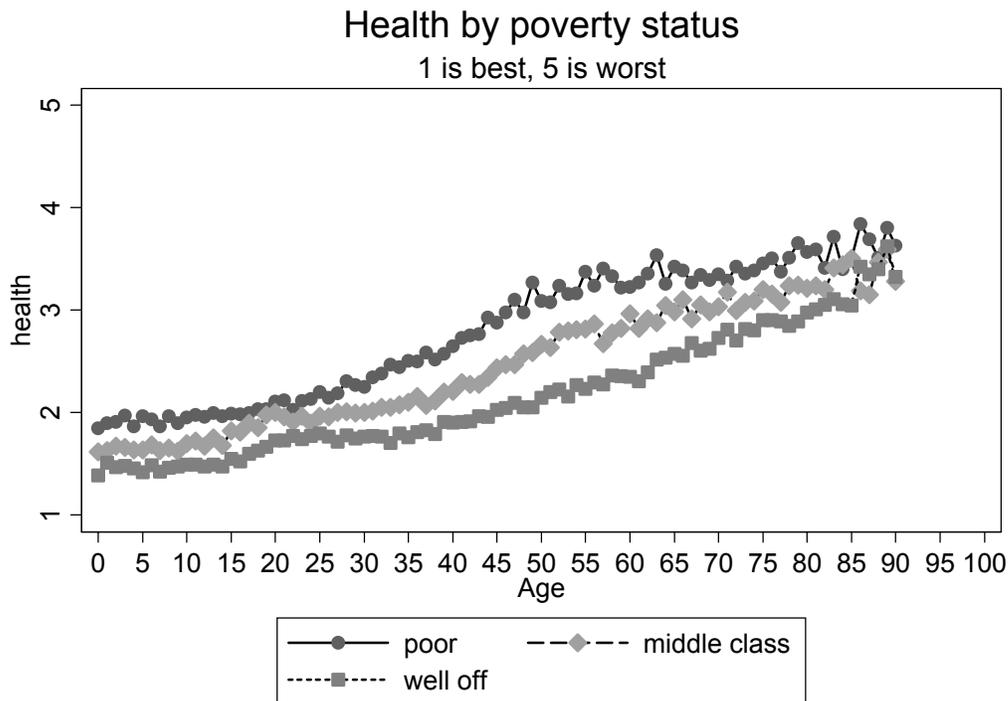
. replace new_pov=3 if povl1>9
(66899 real changes made)

. label define new_pov 1 "<1.5 x poverty level" 2 "1.5-3 x poverty
level" 3 ">3 x poverty level"

table age, contents(mean health) by(new_pov) replace

twoway (connected mean_health age if new_pov==1) (connected mean_health
age if new_pov==2) (connected mean_health age if new_pov==3),
ytitle(health) ylabel(1 (1) 5) xtitle(Age) xlabel(0 (5) 100)
title(Health by poverty status) subtitle("1 is best, 5 is worst")
legend(order(1 "poor" 2 "middle class" 3 "well off"))

```



It's perhaps not surprising that social class really does impact health status. At what ages does health disparity between poor and well off seem the greatest? Why do you think that is?