# Collapse

Say that you have a dataset with observations of people, gender, and salary. Say that you want to know the average salary by gender.

What you are trying to do is collapse the dataset. Instead of individual observations, you want an observation per gender.

The collapse command allows you to do that:

collapse (mean) salary, by(gender)

The resulting dataset with have just one observation per the “by” variable combinations.