

Assignment 4

LPO 9951 / Fall 2015

For this assignment, I'd like you to make use of the datafile and do files that you generated as part of assignment 3. Using these, complete the following steps:

1. Subset the data on gender. Create one data file for males and one for females (hint: use **keep if** or **drop if**).
2. Merge the two datasets together, and tabulate the **_merge** variable. What are the results?
3. Now split the dataset by variables. To do this you will need to either use the **drop** command or use the **save** command with a variable list. Make sure that in each dataset, you include the student id.
4. Add a new line to each dataset. Alter the id in the new observation so that the two files do not match.
5. Repeat the **merge** command again, but this time create a result where the two additional (fake) observations are dropped. Do this using the **merge** command itself, not **drop if** or **keep if**.
6. Repeat the merge, but this time only keep the variables in the *master* dataset.
7. Repeat the merge, but this time only keep the variables in the *using* dataset.

Submit your results in a do file per normal procedure. Remember that since I don't have access to your small dataset, I need a do file that will cleanly create the dataset from the original data and then complete the operations above.