An Introduction to Stata – Section 3.3
Najib Mozahem

In the previous lecture, we gave our variables meaningful names. In this lecture, we will give them a meaningful description. These descriptions are called labels in Stata. This is done using the label command. Once again, we ask Stata for help:

help label

We see from the output that to label a variable we need to use the **label** command followed by the word **variable**, followed by the variable name, and finally followed by the label itself. Take for example the first variable. We know that the variable *id* stores the students' id numbers. Therefore, I would like to attach to it the label "Student identification number". This is done like this:

label variable id "Student identification number"

Looking at the right-hand side, we can clearly see that the variable now has a label attached to it. Now we need to do the same for the other variables. This is accomplished by executing the following commands:

label variable gender "Student is male or female?"

label variable gpa "Student overall GPA"

label variable course_title "Name of the course"

label variable credits "Credits assigned to course"

label variable grade "Numeric grade on course"

label variable semester "Semester course was taken"

Our dataset is starting to look good now. There is one more thing that we need to label at this point, and this is the data itself. If you recall, when we asked for help about the **label** command, there was an option to label the data itself. The original dataset that we looked at in the first section in this course was actually labelled. Labelling the data allows the researcher to know what sort of information does the data contain. In our case, the dataset contains students' grades, so we choose to label it as such:

label data "This dataset contains the grades achieved by some students on certain courses"

If you now run the **describe** command you will see that the dataset and all the variables have been labelled.