

At this point, you have the necessary tools to allow you to jump into a dataset and extract some basic information from it. However, you will rarely find a complete dataset. What I mean is that it will almost always be the case that you need to perform some sort of modification to your dataset. For example, you might want to delete a variable because you do not need it and would like to simplify your dataset. Or perhaps you find yourself needing to create a variable. Why would you want to create a variable? I was once working on a dataset that contained a variable that measured the number of students in each major in a university. What I really wanted however was the percent of students in each major. This mean that I had to get the number of students in a major and divide it by the total number of students and then to multiply this ratio by 100.

In other instances you might want to rename a variable. We have already renamed variables in this course, but we used the graphical user interface. In this section we will learn how to rename variables using the command prompt. This is actually what the next lecture is about.