

This section has hopefully made it clear that having a well-formatted dataset is extremely important. As you have seen, it is very easy to format a dataset using Stata. Of course, the larger the number of variables, the longer it will take, but I guarantee that it is well worth it. Since formatting the data is extremely important, this section will end with an activity. Attached to this lecture you will find an Excel file that contains data. You will notice that the first row contains the names of the variables. Also attached to this lecture is the survey that was used in order to collect the data. You should use this survey in order to:

- Label the dataset
- Label the variables
- Create value labels and assign them to the proper variables

It is important that you complete this task because in the next few sections we will be using this dataset in order to learn how we can take a closer look at the data.