

Included with this lecture is a survey and the dataset that resulted from this survey. This survey was distributed in schools in Lebanon. The purpose of the survey was to see what sort of occupations are students at school level interested in. The survey includes questions about 22 occupations, in addition to asking the respondents about their age and gender. In Lebanon, there is a tendency for some types of occupations to have a high percent of men and a low percent of women. This survey wanted to test children's interest in the occupations when they were still in school. The variable *gender* was included to see whether there were gender differences, while the variable *age* was included to see whether children's occupational interests tend to change with time.

You should now be able to use what you have learned in this course to try to extract the story that is hidden in this dataset. Create a do-file that imports the dataset, labels the variables, and that labels the values for categorical variables. Note that the first row in the dataset contains the variable names. You can change these names if you wish. You can do whatever you want. This is your project.

Once you have a well-formatted Stata data file, save it on your computer and start looking at the summary statistics and the tables in order to see what is going on. Don't forget to use the **by** prefix in order to see whether there might be gender differences. Enjoy.