

Stata Class 3

Alexander C. Lembcke

Exercise 1

Start a new do-file and add the commands to reproduce the following exercises. Open the data set we created last week.

- a) Rename the management score variables to start with the same prefix, e.g. "m_peeps".
- b) Use the "order" command to bring all management variables together.
- c) Generate a variable that is equal to the total number of hours worked in a firm.
- d) List the first 10 observations to see whether your new variable is correctly defined.
- e) Drop that variable from the sample.
- f) Create a dummy variable, i.e. a variable that takes the values 0 and 1, that is 1 if a firm has at least 1 employee that is a union member and 0 otherwise. Hint use generate and replace in combination with "if".
- g) Create a variable that is the sum of all individual management scores. Compare it to the existing variable that combines the management scores ("management"). Why do they differ? Account for the reason for the difference and check again whether scores differ.

Exercise 2

Keep using the same do-file and add the commands to reproduce the following exercises.

- a) Create the conditional expectation function (CEF), i.e. the sample analog of $E(Y_i|X_i)$, where Y is log sales and X is the total servers using site file info.
- b) Create a graph that consists of a scatterplot of the above variables overlaid by the CEF.
- c) To save memory we want to replace the string variable 'analyst' with a byte variable. Create this variable without loss of information.
- d) Generate a variable that is the sum of 'totserver' for each country and year.
- e) Generate a dummy variable for each country.

Exercise 3

Keep using the same do-file and add the commands to reproduce the following exercises.

- a) Generate a variable that is equal to 1.1 and summarize that data for all observations where the variable is 1.1.
- b) Generate a variable that indicates the number of observations for each firm.
- c) Generate a variable that is 1 if a firms management score increased compared to the previous interview.
- d) Generate a ranking of firms by management score for each year and country.
- e) If a firm is observed three times or more, generate and indicator that tells you in which period the management score was second highest.

Exercise 4

Keep using the same do-file and add the commands to reproduce the following exercises.

- a) Create a scatter plot of management scores and sales using only 10% of the observations for each country and year (select the observations randomly).
- b) Use the appropriate options to suppress the legend, add axis descriptions and a title. Create a color version for presentations and a monochrome version for the paper. Save them.
- c) Highlight the observations from the US.