

Exercise session 1

Course: Introductory Econometrics

Lecturer: Dmytro Vikhrov

Date: September 16, 2014.

Problem 1

Open the dataset *401ksubs* located in (File → Open data → Sample file...).

1. Display the dataset for visual inspection. For this select all variable by pressing Ctrl+A, then go to Data → Display values. Explore the following commands in Data tab: Edit values, Add observations and Dataset structure.
2. Plot the histogram of annual income (right click on *inc*, then choose Frequency distribution).
3. Compute descriptive statistics (right click on *inc*, then Summary statistics).
4. Compute Gini coefficient for the income data (Variable → Gini coefficient). Interpret your findings.
5. Plot the histogram of $\ln income$. Explain how and why it is different from the histogram of income.

Problem 2

Now split the sample into several sub-samples.

1. Compute descriptive statistics and plot histograms for males and females (Sample → Restrict, based on criterion...). Interpret your findings.
2. Re-do the above exercise for married and unmarried individuals. Explain your intuition.

Problem 3 Let's now move to bivariate graphs. Make three scatter plots: income against family size, income against age and financial assets against income. Further, compute correlation coefficients for each pair of variables. Explain your intuition.

Problem 4 Now let's add time dimension to our analysis. Open dataset *greene5_1* and make time series plots of variables that you might find interesting.

Problem 5 Open dataset *greene14_1* and explore the panel structure of the dataset. Make panel plots of variable *Q* by company (*unit*).