In preparation of the interview, please work through the following short exercise:

- Load the Lalonde data set (see attachment) into either Stata or R. The data belongs to a study that looked at the effectiveness of a job training program (the treatment) on the real earnings of an individual, a couple years after completion of the program. It consists of a number of demographic variables as well as a treatment indicator and the real earnings in the year 1978 (the response).
- Produce a table with summary statistics (incl. mean, standard deviation) for the variables in the dataset.
- Produce *two* histograms showing the distribution of real earnings 1978 (re78) for individuals in the treatment and in the control group *separately*.
- Test whether the difference between treatment and control group is statistically significant at conventional levels. What do you find?

Please send back the code *and* the results by Tue 27 February (midnight).

There is no need to extensively format results, but you should comment every step of your code and the results you obtain. Also, don't worry if you don't manage to address all the points. You will be able to explain how you went about solving the exercise and where you might have faced difficulties during the interview.