The homework is due on **Monday, February 5th, 14:00.** Solving all questions will give you 100 points. You should submit a printout of your code, as well as e-mail the code to Evimaria (trezi@cs.helsinki.fi). There should be clear instructions on how to run the experiments. If you have pre-processed the data, you should either submit the code for preprocessing, or submit also the pre-processed dataset.

Question (100 points)   Download one of the AS Internet graphs, or the co-authorship graph from the home page of the course, and perform the following measurements.

1. Estimating the exponent of the degree distribution.

   - Make a log-log plot of the degree distribution and estimate the exponent by fitting a line.

   - Perform logarithmic binning, plot the resulting line, and estimate the exponent of the distribution by fitting a line to the curve.

   - Make a log-log plot of the cummulative distribution and estimate the exponent **of the degree distribution** by fitting a line to it.

   - Estimate the exponent by using the maximum likelihood estimate.

   Discuss your experimental findings.

   (Note: You may need to prune the low-end of the distribution to obtain a power-law distribution).

2. Computing the clustering coefficient.

   - Compute the clustering coefficient $C^{(2)}$.
   - Plot the $C(k)$ distribution.

   Discuss your findings.

3. Compute the average path length of any two reachable nodes in the graph.

4. Compute the degree correlation using the method of Newman. Discuss your results.

Bonus Question (20 points)   Randomly rewire the graph by swapping edges as explained in class, for a sufficiently large number of swaps. Create 100 such graphs, and for each one compute the clustering coefficient, the average path length, and the degree correlations. Create a histogram of the values, and observe where the actual value falls in this histogram. Discuss your observations.