Due June 5TH

1. Run NETWORK>SUBGROUP>LOUVAIN METHOD of partitioning (with UCINET)

using the [**aNobii** bookshelf (cosine similarity)](https://www.space.ntu.edu.tw/navigate/s/4DE84DAC075647E7A97D6C843C2CE78FQQY). Save the results Louv attribute file so it can be opened in NetDraw, along with the book pair similarity file.

Again, run NEWORK>SUBROUP>NEWMAN COMMUNITY DETECTION using the same file. Note that you will need to first dichotomize the data by setting the similarity thresholds (For example, aNobii at 0.7).

You can use Transform>dichotomize or Transform>dichotomize interactive function to determine the proper threshold and visualize the data in NetDraw.

Determine the proper level of clustering by examining the Q value. Do include the visualization of the results in your report.

1. Community-detection with Gephi

Use the [**aNobii** bookshelf (cosine similarity)](https://www.space.ntu.edu.tw/navigate/s/4DE84DAC075647E7A97D6C843C2CE78FQQY) data

Run both Modularity (i.e. Louvain method) and

Girvan-Newman [aNobii (dichotomized at 0.7)](https://www.space.ntu.edu.tw/navigate/s/06A5759147E1470194D433C69683B84BQQY)

Based on results from Q1 and Q2, discuss with clustering method produces the best result.

1. Filtering and modularity analysis (with Gephi)

Try to apply modularity grouping method to Public health [2008-2017](https://www.space.ntu.edu.tw/navigate/s/89A53D911ADD4159BFB66EC890175F5EQQY) dataset

Explore different filters (e.g. edge weight, degree range, giant component etc. to make your graphs more manageable and interpretable. State what nodes and edges should be filtered out and give your rationales. (criteria for filtering: the modularity analysis should produce high modularity value and “manageable” number of modularity classes)

Combine filtering and modularity analysis to make sense of the clustering. You might want to consult graduate students in Public health to help you determine the best grouping of research interests.

Perform the same analysis with [Marvel social network](https://www.space.ntu.edu.tw/navigate/s/28F3DAAC32F1451FB4BD52DFBB23E82CQQY) , see if you can come up with satisfactory clustering based your understanding of the stories.

1. Structural equivalence and position analysis

For this assignment, you will run various similarity and structural equivalent methods on “TRADE” data in the UCIENT data folder. Notice that after running similarity/dissimilarity command in “TOOLS”, an “output” file will be saved. Make sure you use different names for the output files and record the folder in which they are saved as you will need these output files for clustering and MDS analyses.

1. Similarity; Run Tools>Similarities

on TRADE and compare the results using different similarity measures (i.e. “Correlation”, “Matches”, or “Jaccard” ).

Next use the same command to analyze the TRADE data.

Dissimilarity; Run Tools>Dissimilarities and Distances

Notice that the dissimilarity data can be saved into new data file for later analysis.

1. Run Hierarchical clustering on the results of the similarity (e.g. correlation and matches) and dissimilarity analyses;
   1. Run Network>Roles & Positions > Structural>Profile   
      Generate and compare **clustering diagrams** (dendrograms) for the same school thesis committee data and TRADE data and choose the algorithm that best capture the profile similarity of the faculty/nations;
   2. Run Tools>Scaling and Decomposition> Metric (for valued network) / Non-metric (for binary network)

Compare the visual displays of either the thesis or TRADE data generated by Cluster analysis and MDS, see if the visual configurations produced by these two visualization tools agree with each other. You can experiment different dimension values (i.e. 2 or 3) and determine the appropriateness of the MDS results by checking the Stress number. (MDS can also be performed by NETDREW, first you need to make sure whether the data is numerical or non-numerical).

\*This data has been selected by Wasserman and Faust (1994) from a list of 63 countries given by Smith and White (1988). The selection was intended to be a representative sample of countries which spanned the globe physically, economically and politically and was used by them in their network analysis book. The data records interaction of the countries with respect to TRADE of four goods, namely**: manufactured goods, food and live animals, crude materials (not food) and minerals and fuels**. The final matrix records exchange of **diplomats** between the countries. All TRADE (including the diplomats) is from the row to the column. The TRADE Attribute data lists average population growth between 1970 and 1981, average GNP growth (per capita) over the same period, secondary school enrollment ratio in 1981, and energy consumption in 1981 (in kilo coal equivalents per capita).

Questions (<http://www.analytictech.com/ucinet/help/webhelp.html>):

1. Tools > Similarity and Distance
   1. Mode: columns or rows > similarity col by col
   2. Mode: Matricies > similarity between matrices (e.g. TRADE has 5 matrices, hence the output similarity matrix is 5x5)
2. Tools > Scaling > MDS & Tools > Clustering both takes a similarity (or distance) matrix as input.
3. Network > Roles & Positions > Structural> Profile  
   Comparisons of **rows and columns** of data matrices  
   Output:

**SE**: Similarity or distance matrix (based on Measure of profile similarity/distance)