

**ASSIGNMENT C3**  
**CORE METHODS IN EDUCATIONAL DATA MINING**  
**PROFESSOR RYAN BAKER**  
**KNOWLEDGE STRUCTURE**  
**DUE NOON, WEDNESDAY APRIL 12**

The goal of this assignment is to build a knowledge structure mapping for the simulated data set in 8items.csv. You can assume that this data set involves data from an 8-item test given to students.

You may use any method to discover the knowledge structure in this data set. It is acceptable to use Barnes's Q-matrix method, Learning Factors Analysis, Learning Factors Transfer Analysis, Partial Order Knowledge Spaces, Factor Analysis, or any other method.

You can use any existing software package, or can implement your own code in any programming language. It is also acceptable to do this assignment by hand in Excel (show your work).

Your hand-in must include a q-matrix which indicates the mapping between each item and one or more skills. You should also give the evidence you used to decide that this was the best mapping. You must also turn in any code or Excel files (Matlab files, Python code, SPSS code, etc. etc.) used in your computations. If you used an existing package, it is acceptable to simply state what package you used. You should also turn in a document explaining how you completed the assignment (e.g. the method and implementation of it that you used, and how you assessed model goodness). You will be graded on completeness and comprehensibility of your hand-in, whether you correctly and validly apply the method you choose to this data, and whether the methods you chose fit the requirements of this assignment.

**BONUS:** The student who succeeds in producing the knowledge structure closest to the original knowledge structure used to simulate the data gets the bonus. "Best" is defined as the percent agreement between the original knowledge structure, and your knowledge structure. Cross-validation does not need to be used for this assignment.