

**BASIC ASSIGNMENT 5**  
**CORE METHODS IN EDUCATIONAL DATA MINING**  
**PROFESSOR RYAN S. BAKER**

**DUE VIA ASSISTMENTS NOVEMBER 9, 2020, 5PM, EASTERN STANDARD TIME**

In this assignment, you will need to build a Bayesian Knowledge Tracing model for this week's data set. This data set is a subset of the data set used in:

Baker, R.S.J.d., Corbett, A.T., Roll, I., Koedinger, K.R. (2008) Developing a Generalizable Detector of When Students Game the System. *User Modeling and User-Adapted Interaction*, 18, 3, 287-314.

This variables in the data set are:

- ID – a unique ID for every student action in the Cognitive Tutor used
- Lesson – the tutor lesson the action comes from
- Student – a deidentified ID for the student
- KC – the knowledge component (skill) involved
- Item – the problem step in the learning system
- Right – is the student action right (1) or not right (0)
- Firstattempt – is this the student's first attempt at the problem step (1)?
- Time – how long did the student attempt take?

You should complete questions 1-9 of this assignment in Microsoft Excel, or a similar spreadsheet program.

**You will receive full credit if you complete questions 1-9.**

**Question 10 is not graded.** If you decide to try question 10, you should complete it using Java and BKT-BF, at <https://www.upenn.edu/learninganalytics/ryanbaker/BKT-BruteForce.zip> . Again, question 10 is optional and not graded.

Final note: If ASSISTments displays check marks and X's on the left panel of the assignment, or a performance bar on the bottom, don't worry. **It is OK to ask for hints and make errors. If you complete the assignment, you will receive full credit.**