CSCI 6350
Reading Assignment 1

Read paper "Fast Algorithms for Mining Association Rules" by R. Agrawal and R. Srikant. Think about answers to the following questions as you read the paper. You will be asked to discuss your answer to these questions in class.

Questions:
1. What is the Apriori property?
2. How is the Apriori property used in the frequent itemset generation process?
3. How is the Apriori property used in the rule generation step of the Apriori algorithm?
4. What is the main purpose of the "subset" function (section 2.1.2)? How does the chosen data structure support an *efficient* "subset" function?
5. What is the most computationally expensive step/process in the Apriori algorithm?
6. What is the main advantage of AprioriTid?
7. When does AprioriTid perform better than Apriori? When does Apriori perform better than AprioriTid?
8. What is the main advantage of the AprioriHybrid algorithm?
9. When evaluating the performance of various association rule mining algorithms, *synthetic data* is constructed. What are the set of factors one needs to take into account when constructing this data?
10. Why is it that the experiments only compare the speed of various algorithms, and not comparing the association rules generated from these systems?