732A61/TDDD41 Data Mining - Clustering and Association Analysis
Lecture 9: Exercises

Jose M. Peña
IDA, Linköping University, Sweden

Exercises

- Run the Apriori algorithm on the following transactional database with minimum support equal to one transaction. Explain step by step the execution.

Transaction id	Items bought
1	A, B, C
2	X, Y, Z
3	A, Y, C
4	X, B, Z

- Repeat the exercise above with the following additional constraint: Find the frequent itemsets that contain the item A. Make clear when and how the constraint is used. Incorporate the constraint into the algorithm, i.e. do not simply run the algorithm and afterwards consider the constraint.

- Let the items A, B, C, X, Y and Z have a price of respectively -3, -2, -1, 1, 2 and 3 units. Repeat the exercise above with the following additional constraint: Find the frequent itemsets whose range is smaller than 3 (recall that the range is the price of the most expensive item minus the price of the cheapest item). Make clear when and how the constraint is used. Incorporate the constraint into the algorithm, i.e. do not simply run the algorithm and afterwards consider the constraint.

- Repeat the exercises above with the FP grow algorithm.

- Apply the Simple algorithm to the frequent itemset XBZ on the database above in order to find association rules with confidence greater than 0.5.