

Supplementary Notes #1

COMP 578 Data Mining and Data Warehousing

Exercise on Association Rule Mining

1. You are given the data set below:

Trans ID	Items Purchased
100	Bread, butter, noodles
200	Butter, chips
300	Butter, coke
400	Bread, butter, chips
500	Bread, coke
600	Butter, coke
700	Bread, coke
800	Bread, butter, coke, noodles
900	Bread, butter, coke

Fill in the blanks below:

- a. People who buy ___ are 2 times more likely to also buy ___.
 - b. People who buy butter are ___ times more likely to also buy noodles.
2. Mine the above data set for interesting associations.
 - a. Use $\text{min_sup} = 20\%$ and $\text{min_conf} = 50\%$.
 - b. How many times do you have to go through the data set to discover all interesting association?
 - c. Repeat a to c using the Apriori Algorithm.