

Customer ID	Transaction ID	Items Bought
1	0001	$\{a, d, e\}$
1	0024	$\{a, b, c, e\}$
2	0012	$\{a, b, d, e\}$
2	0031	$\{a, c, d, e\}$
3	0015	$\{b, c, e\}$
3	0022	$\{b, d, e\}$
4	0029	$\{c, d\}$
4	0040	$\{a, b, c\}$
5	0033	$\{a, d, e\}$
5	0038	$\{a, b, e\}$

***A. treating each transaction ID as a market basket***

1. Compute the support for itemset  $\{b, d, e\}$ :
2. Compute the confidence for the association rules  $\{b, d\} \rightarrow \{e\}$  and  $\{e\} \rightarrow \{b, d\}$ .

***B. Treating each custom ID as a market basket.*** Each item should be treated as a binary variable (1 if an item appears in at least one transaction bought by the customer, and 0 otherwise.)

1. Compute the support for itemsets  $\{b, d, e\}$ :
2. Compute the confidence for the association rules  $\{b, d\} \rightarrow \{e\}$  and  $\{e\} \rightarrow \{b, d\}$

Consider the following set of frequent 3-itemsets:

$\{1, 2, 3\}$ ,  $\{1, 2, 4\}$ ,  $\{1, 2, 5\}$ ,  $\{1, 3, 4\}$ ,  $\{1, 3, 5\}$ ,  $\{2, 3, 4\}$ ,  $\{2, 3, 5\}$ ,  $\{3, 4, 5\}$ . Assume that there are only five items in the data set.

Q1: List all candidate 4-itemsets obtained by the candidate generation procedure in Apriori.

Q2: List all candidate 4-itemsets that survive the candidate pruning step of the Apriori algorithm.