

Consider the following data set. Each line contains the items in a transaction.

p1 1 3 5 7 11 Items ; 25 Total ; 6 Transactions
 p2 2 5 6 20
 p3 3 5 7 11 13
 p4 5 11 15 25
 p5 1 3 7 11
 p6 1 5 7 20

Q1[4]. Find all frequent itemsets (size 1,2,3,4, and so on), where the minimum support is 2/6. Show the steps of the Apriori algorithm.

$$F_1 = \left\{ \{1\} = 3, \{3\} = 3, \{5\} = 5, \{7\} = 4, \{11\} = 3, \{20\} = 2 \right\}$$

1: 3/6	5: 5/6	11: 3/6	20: 2/6
2: 1/6	6: 1/6	13: 1/6	25: 1/6
3: 3/6	7: 4/6	15: 1/6	

$$F_2 = \left\{ \begin{aligned} &\{1,3\} = 2, \{1,5\} = 2, \{1,7\} = 3, \{3,5\} = 2, \{3,7\} = 3, \\ &\{3,11\} = 2, \{5,7\} = 3, \{5,11\} = 2, \{5,20\} = 2, \\ &\{7,11\} = 2 \end{aligned} \right\}$$

1/3: 2/6	3/5: 2/6	5/7: 3/6	7/11: 2/6	1/20: 0/6
1/5: 2/6	3/7: 3/6	5/11: 2/6	7/20: 1/6	
1/7: 3/6	3/11: 2/6	5/20: 2/6		
1/11: 1/6	3/20: 0/6			
1/20: 1/6				

$$F_3 = \left\{ \{1,3,7\} = 2, \{1,5,7\} = 2, \{3,5,7\} = 2, \{3,7,11\} = 2 \right\}$$

1/3/5: 1/6	3/5/11: 1/6
1/3/7: 2/6	3/7/11: 2/6
1/5/7: 2/6	5/7/11: 1/6
3/5/7: 2/6	5/7/20: 1/6
1/3/11: 1/6	5/11/20: 0/6
1/7/11: 1/6	3/5/20: 0/6
	1/5/20: 1/6

F₄: None Meet Minimum Support

1/3/5/7: 1/6
1/3/7/11: 1/6
3/5/7/11: 1/6

Q2[4]. Examine the frequent item sets that you generated in the previous question. Use them to generate association rules, where the minimum confidence is 80%.

$1 \Rightarrow 3 : 2/3$	$3 \Rightarrow 1 : 2/3$	} F2 Association Rules: $1 \Rightarrow 7, 3 \Rightarrow 7, 20 \Rightarrow 5$
$1 \Rightarrow 5 : 2/3$	$5 \Rightarrow 1 : 2/5$	
$1 \Rightarrow 7 : 3/3$	$7 \Rightarrow 1 : 3/4$	
$3 \Rightarrow 5 : 2/3$	$5 \Rightarrow 3 : 2/5$	
$3 \Rightarrow 7 : 3/3$	$7 \Rightarrow 3 : 3/4$	
$3 \Rightarrow 11 : 2/3$	$11 \Rightarrow 3 : 2/3$	
$5 \Rightarrow 7 : 3/5$	$7 \Rightarrow 5 : 3/4$	
$5 \Rightarrow 11 : 2/5$	$11 \Rightarrow 5 : 2/3$	
$5 \Rightarrow 20 : 2/5$	$20 \Rightarrow 5 : 2/2$	
$7 \Rightarrow 11 : 2/4$	$11 \Rightarrow 7 : 2/3$	

$1, 3 \Rightarrow 7 : 2/2$	$3, 5 \Rightarrow 7 : 2/2$	} F3 Association Rules: $(1, 3 \Rightarrow 7), (1, 5 \Rightarrow 7), (3, 5 \Rightarrow 7),$ $(3, 11 \Rightarrow 7), (7, 11 \Rightarrow 3)$
$1, 7 \Rightarrow 3 : 2/3$	$3, 7 \Rightarrow 5 : 2/3$	
$3, 7 \Rightarrow 1 : 2/3$	$5, 7 \Rightarrow 3 : 2/3$	
$1, 5 \Rightarrow 7 : 2/2$	$3, 7 \Rightarrow 11 : 2/3$	
$1, 7 \Rightarrow 5 : 2/3$	$3, 11 \Rightarrow 7 : 2/2$	
$7, 5 \Rightarrow 1 : 2/3$	$7, 11 \Rightarrow 3 : 2/2$	

Rules w/ confidence: $1 \Rightarrow 7, 3 \Rightarrow 7, 20 \Rightarrow 5, (1, 3 \Rightarrow 7), (1, 5 \Rightarrow 7), (3, 5 \Rightarrow 7), (3, 11 \Rightarrow 7), (7, 11 \Rightarrow 3)$