➢ Association Rules

Tid	Items
10	A, B, C
20	A, C, D, F
30	B, C, D, E
40	B, C, E
50	C, E, F
60	C, E, F
70	A, B, C, F
80	A, B
90	B, C, E
100	B, F

We will use apriori algorithm to generate candidate itemsets.

a) Minimum support=0.25 and confidence=0.75

1st scan:

Itemset	Sup
A	0.4
B	0.7
C	0.8
D	0.2
E	0.5
F	0.5

Itemset	sup
A	0.4
B	0.7
C	0.8
E	0.5
F	0.5

2nd scan:

Itemset	Sup
A,B	0.3
A,C	0.3
A,E	0
A,F	0.2
B,C	0.5
B,E	0.3
B,F	0.2
C,E	0.5
C,F	0.4
E,F	0.2

Itemset	Sup
A,B	0.3
A,C	0.3
B,C	0.5
B,E	0.3
C,E	0.5
C,F	0.4

3rd Scan:

Itemset	Sup
A,B,C	0.2
A,C,F	0.1
B,C,E	0.3

→

Itemset	Sup
B,C,E	0.3

So, the association rules with minimum support=0.25 and confidence=0.75 are:

A->B

A->C

E->C

F->C

B,E->C

b)

We want to find association rules with minimum support=0.25 and confidence=0.5

The candidate itemsets are the same as previously as we want minimum support=0.25

So, the association rules with minimum support=0.25 and confidence=0.5 are:

A->B

A->C

B->C

C->B

E->B
C->E

E->C

C->F

F->C

E->B,C
B,C->E

C,E->B

B,E->C

c) Minimum support=0.35 and confidence=0.5

1st Scan:

Itemset	Sup
A	0.4
B	0.7
C	0.8
D	0.2
E	0.5
F	0.5

Itemset	Sup
A	0.4
B	0.7
C	0.8
E	0.5
F	0.5

2nd Scan:

Itemset	Sup
A,B	0.3
A,C	0.3
A,E	0
A,F	0.2
B,C	0.5
B,E	0.3
B,F	0.2
C,E	0.5
C,F	0.4
E,F	0.2

Itemset	Sup
B,C	0.5
C,E	0.5
C,F	0.4

If we do 3rd scan, we will observe that we will generate itemsets with length=3 that will contain subsets that are not frequent(support<0.35). So, we will not generate them.

So, the association rules with minimum support=0.35 and confidence=0.5 are:

B->C

C->B

C->E

E->C

C->F

F->C

Discussion-Results:

We used Apriori algorithm to generate association rules for a transactional database with different support and confidence thresholds. In the first case (a) with support=25% and confidence=75% we found strong association rules. By reducing the confidence threshold in the second case (b) to 50% we found additional association rules, forming a superset of the rules in part (a). In the third case (c) by increasing the support threshold to 35% we result in a subset of the association rules from part (b). Overall, with a lower confidence threshold more rules were generated, while increasing the support threshold fewer rules were generated. Additionally, it is noteworthy that two association rules (E->C and F->C) consistently appeared in all three parts (a, b, and c) regardless of the change in support and confidence thresholds. That indicates that these two rules (E->C and F->C) are robust and can be deemed as reliable rules within the context of the transactional database.