

② Given a simple transactional data base X:

TID	T01	T02	T03	T04	T05	T06	T07	T08
Items	A,B,C,D	A,C,D,F	C,D,E,G,A	A,D,F,B	B,C,G	D,F,G	A,B,G	C,D,F,G

Using the threshold values Support = 25%, Confidence = 60%, find all large itemsets in database X.

Ans) Given,

TID	T01	T02	T03	T04	T05	T06	T07	T08
Items	{A,B,C,D}	{A,C,D,F}	{C,D,E,G,A}	{A,D,F,B}	{B,C,G}	{D,F,G}	{A,B,G}	{C,D,F,G}

Calculating Support

We can use Apriori algorithm to find all large itemsets.

Calculating Support threshold

$$\text{Given Support} = 25\% = 0.25$$

$$\text{Total transactions} = 8$$

$$\text{Support threshold} = 0.25 \times 8 = 2$$

Generating 1-item sets.

$$\{A\} \rightarrow \text{Support Count} = 5, \text{Support} = \frac{5}{8} \times 100 = 62.5\%$$

$$\{B\} \rightarrow \text{Support Count} = 4, \text{Support} = \frac{4}{8} \times 100 = 50\%$$

$$\{C\} \rightarrow \text{Support Count} = 5, \text{Support} = \frac{5}{8} \times 100 = 62.5\%$$

$$\{D\} \rightarrow \text{Support Count} = 6, \text{Support} = \frac{6}{8} \times 100 = 75\%$$

$$\{E\} \rightarrow \text{Support Count} = 1, \text{Support} = \frac{1}{8} \times 100 = 12.5\%$$

$$\{F\} \rightarrow \text{Support Count} = 4, \text{Support} = \frac{4}{8} \times 100 = 50\%$$

$$\{G\} \rightarrow \text{Support Count} = 5, \text{Support} = \frac{5}{8} \times 100 = 62.5\%$$

Large 1-item sets with Support threshold  $\geq 2$  are  $\{\{A\}, \{B\}, \{C\}, \{D\}, \{F\}, \{G\}\}$

Generating 2-item sets.

Item Set	Support Count
$\{A, B\}$	3
$\{A, C\}$	3
$\{A, D\}$	4
$\{A, F\}$	2
$\{A, G\}$	2
$\{B, C\}$	2
$\{B, D\}$	2
$\{B, F\}$	1
$\{B, G\}$	2
$\{C, D\}$	4
$\{C, F\}$	2
$\{C, G\}$	3

$$\{D, F\} \quad 4$$

$$\{D, G\} \quad 3$$

$$\{F, G\} \quad 2$$

Large 2-itemsets (Support Count  $\geq 2$ )

$\{\{A, B\}, \{A, C\}, \{A, D\}, \{A, F\}, \{A, G\}, \{B, C\}, \{B, D\}, \{B, G\}, \{C, D\},$   
 $\{C, F\}, \{C, G\}, \{D, F\}, \{D, G\}, \{F, G\}\}$

Generate 3 itemsets from 2 itemsets

Itemset	Support Count		
$\{A, B, C\}$	1	$\{A, D, G\}$	1
$\{A, B, D\}$	2	$\{A, F, G\}$	0
$\{A, B, F\}$	1	$\{B, C, D\}$	1
$\{A, B, G\}$	1	$\{B, C, F\}$	0
$\{A, C, D\}$	3	$\{B, C, G\}$	1
$\{A, C, F\}$	1	$\{B, D, F\}$	1
$\{A, C, G\}$	1	$\{B, F, G\}$	0
$\{A, D, F\}$	2	$\{C, D, F\}$	2
		$\{C, D, G\}$	2
		$\{C, F, G\}$	1
		$\{D, F, G\}$	2

Large 3-itemsets (Support Count  $\geq 2$ )

$\{\{A, B, D\}, \{A, C, D\}, \{A, D, F\}, \{C, D, F\}, \{C, D, G\}, \{D, F, G\}\}$