Q.1) A database has four transactions. Let min_sup = 60% min_conf = 80%. Apply Apriori algorithm to find the frequent itemsets and the strong association rules.

Transaction	Date	List of items
T100	10/15/99	{K, A, D, B}
T200	10/15/99	{D, A, C, E, B}
T300	10/19/99	{C, A, B, E}
T400	10/22/99	{B, A, D}

Soln:- Support threshold = 60%.

∴ min_sup = 0.6 × no: of transaction
= 0.6 × 4 = 2.4

∴ min_sup = 2.4

① Count of each Itemset (C_1) by scanning the database

Itemset	Count
{A}	4
{B}	4
{C}	2
{D}	3
{E}	1
{K}	1

~~② Prune step (L_1) :- Remove min_sup below value~~

(2) Prune step (L_1):- C_1 show that Itemset {C}, {E} and {K} doesnot meet the min_sup.

Itemset	Count
{A}	4
{B}	4
{D}	3

(3) Join step (C_2):- Form C_2 from $L_1 \bowtie L_1$ and find out their occurrences.

Itemset	Count
{A, B}	4
{A, D}	3
{B, D}	3

(4) Prune step (L_2):- C_2 shows that all Itemset does meet the min_sup.

Itemset	Count
{A, B}	4
{A, D}	3
{B, D}	3

(5) Join step (C_3):- Form C_3 from $L_2 \bowtie L_2$ and find out their occurrences.

Itemset	Count
{A, B, D}	3

Thus, {A, B, D} it is frequent.

⑥ Generate Association Rule:-
From frequent itemset discovered above, the association could be:-

- {A, B} → {D}
confidence = support {A, B, D} / support {A, B}
= (3/4) × 100
= 75 %.

- {A, D} → {B}
confidence = support {A, B, D} / support {A, D}
= (3/3) × 100
= 100 %.

- {B, D} → {A}
confidence = Support {A, B, D} / support {B, D}
= (3/3) × 100
= 100 %.

- {A} → {B, D}
confidence = support {A, B, D} / support {A}
= ~~75~~ (3/4) × 100
= 75 %.

{B} → {A, D}

Confidence = support {A, B, D} / support {B}

= (3/4) × 100

= 75 %.

{D} → {A, B}

Confidence = support {A, B, D} / support {D}

= (3/3) × 100

= 100 %.

So if minimum confidence is 80 %, then {A, D} → {B}, {B, D} → {A}, and {D} → {A, B} can be considered as strong association rules.

Q.2) Generate frequent pattern tree for the following transaction with 30% minimum support.

Transaction ID	Items
T1	E, A, D, B
T2	D, A, C, E, B
T3	C, A, B, E
T4	B, A, D
T5	D
T6	D, B
T7	A, D, E
T8	B, C

Soln:- Sup_count = 8 × 30% = 8 × 0.3 = 2.4

Sup_count = 2.4

Step 1 :- Scan the database for count the each dataset

Item set	Sup count
A	5
B	6
C	3
D	6
E	4

Step 2:- Sort the set of frequent Itemset in the order of descending support order.

Itemset	Sup count
B	6
D	6
A	5
E	4
C	3

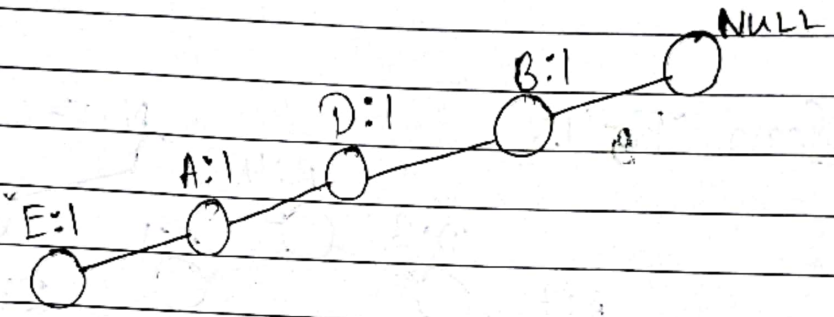
Step 3:- Scan the database for 2nd time sort item in each transaction according to descending support count

Transaction ID	Item_bought
T1	{B, D, A, E}
T2	{B, D, A, E, C}
T3	{B, A, E, C}
T4	{B, D, A}
T5	{D}
T6	{B, D}
T7	{D, A, E}
T8	{B, C}

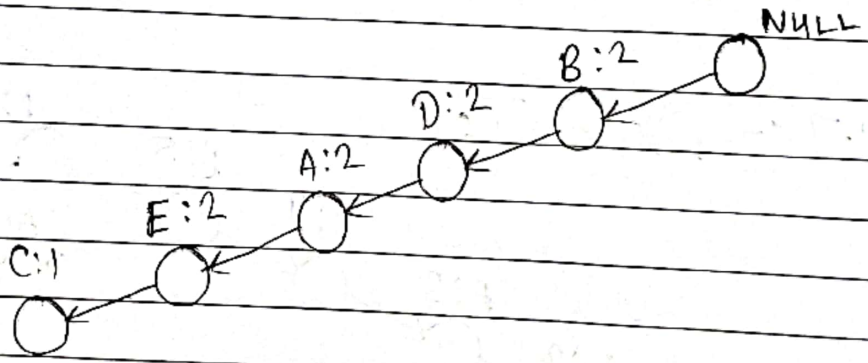
Step 4:- Construct of Frequent Pattern Tree.

4.1:- Create a root node with label null (O^{NULL})

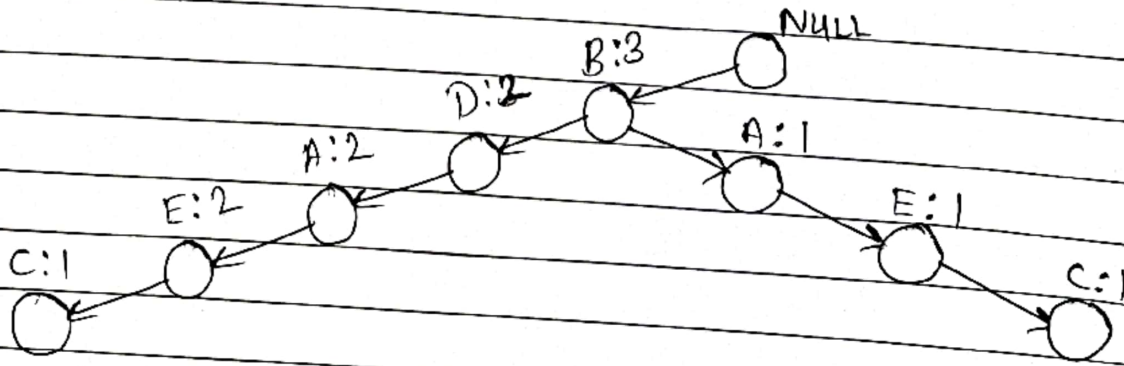
4.2:- Scan T_1



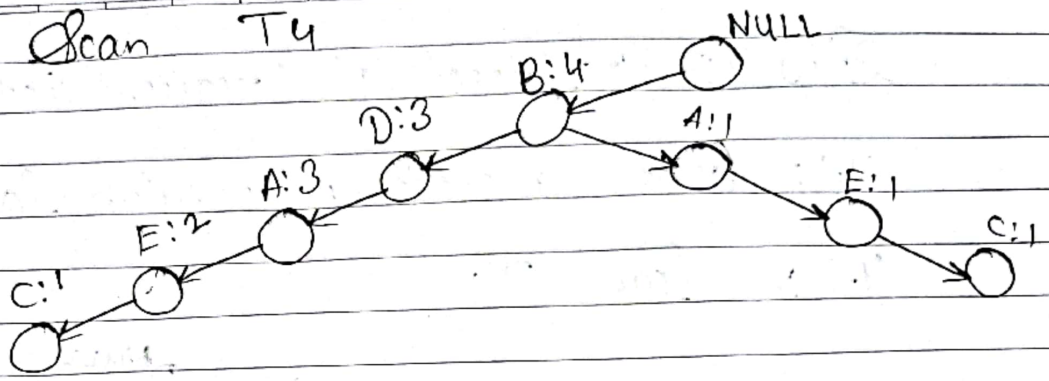
4.3:- Scan T_2



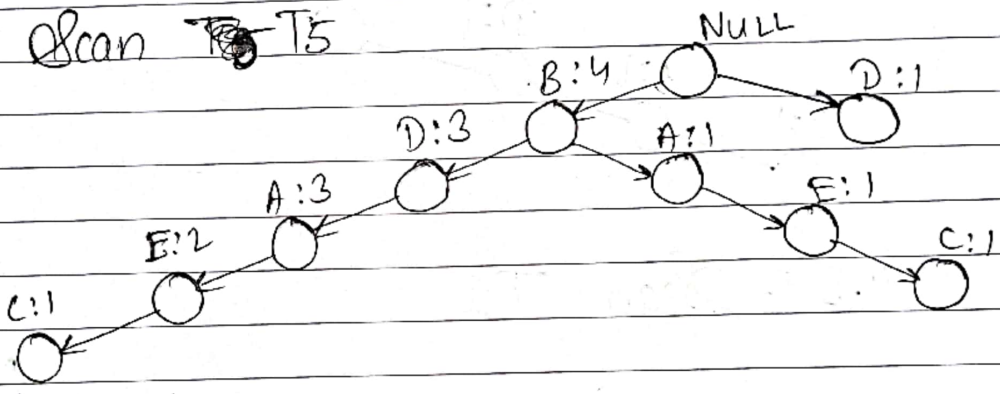
4.4:- Scan T_3



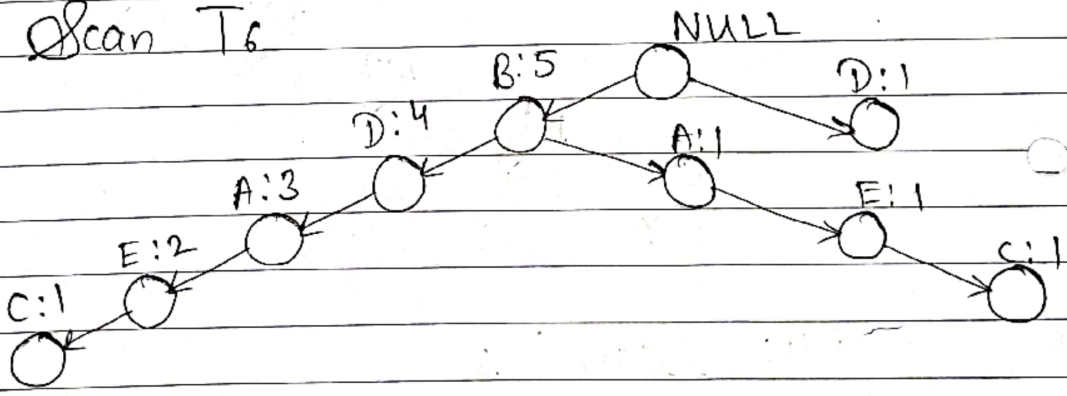
4.5 :- Scan T4



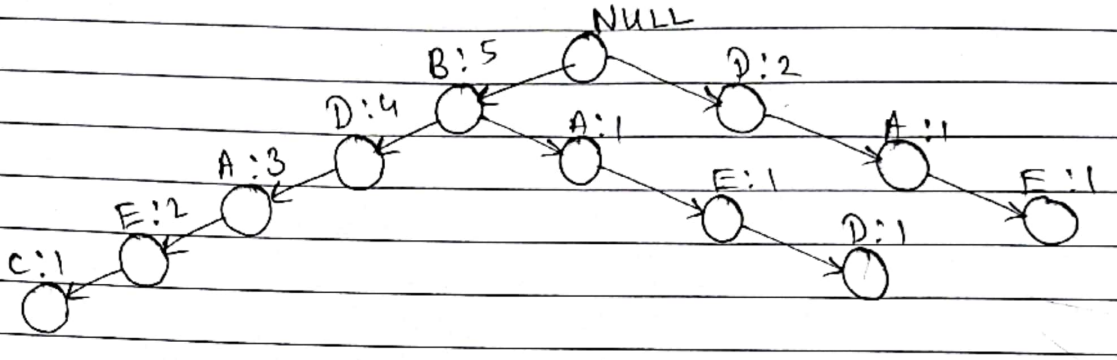
4.6 :- Scan T5



4.7 :- Scan T6



4.7:- Scan T7



4.8 :- Scan T8