Q1) Write the apriori algorithm. Solve the following by Applying apriori algorithm

Consider the transaction below.

TID	Item
T_1	1,3,4
T_2	2,3,5
T_3	1,2,3,5
T_4	2,5
T_5	1,3,5

Create all possible association rule by generating Candidate set Considering minimum Support as 2 and the Confidence % for the rule should be >= 60%.

Ans) Candidate Set C_1

Item	Support Count
1	3
2	3
3	4
4	1
5	4

L_1

f_1	frequent Itemset
1	3
2	3
3	4
5	4

Candidate Set C_2

Item	Support count
1,2	1
1,3	3
1,5	2
2,3	2
2,5	3
3,5	3

L_2

f_1	frequent Itemset
1,3	3
1,5	2
2,3	2
2,5	3
3,5	3

Candidate Set C_3

Item	Support Count
1, 3, 5	2
2, 3, 5	2

L_2

f_i	frequent set
1, 3, 5	2
2, 3, 5	2

Candidate set C_4 can't be generated, then the Item set wil one Identified as two Final frequent Item set are

1, 3, 5   and   2, 3, 5

Rules for $\{1, 3, 5\} \rightarrow \{1,3\} \{1,5\} \{3,5\} \{1\} \{3\} \{5\}$

① $1 \wedge 3 \rightarrow 5$

$$\text{Support} = \frac{2/5}{3/5} = \frac{2}{3} = 66\%.$$

② $1 \wedge 5 \rightarrow 3$

$$\text{Support} = \frac{2}{2} = 100\%.$$

③ $3 \wedge 5 \rightarrow 1$

$$\text{Support} = \frac{2}{3} = 66\%.$$

④ $1 \rightarrow 3 \wedge 5$

$$\text{Support} = \frac{2}{3} = 66\%.$$

⑤ $3 \rightarrow 1 \wedge 5$

$$\text{Support} = \frac{2}{4} = 50\%.$$

⑥ $5 \rightarrow 1 \wedge 3$

$$\text{Support} = \frac{2}{4} = 50\%.$$

Rule for $\{2,3,5\} \to \{2,3\}, \{2,5\} \{3,5\} \{2\} \{3\} \{5\}$

① $2 \cap 3 \to 5$

Support $= \dfrac{2/5}{2/4} = 100\%$.

② $2 \cap 5 \to 3$

Support $= 2/3 = 66\%$.

③ $3 \cap 5 \to 2$

Support $= 2/2 = 66\%$

④ $2 \to 3 \cap 5$

Support $= 2/3 = 66\%$.

⑤ $3 \to 2 \cap 5$

Support $= 2/4 = 56\%$.

⑥ $5 \to 2 \cap 3$

Support $= 2/4 = 56\%$.

Threshold confidence $= \{1,3\} \{1,5\} \{3,5\} \{1\}$
$\{2,3\} \{2,5\} \{3,5\} \{2\}$

Pair (Juice, Soda): h(Juice, Soda) = $(3+4) \times 7 = 0$

Pair (Juice, water): h(Juice, water) = $(3+5) \times 7 = 1$

Pair (Soda, water): h(Soda, water) = $(4+5) \times 7 = 2$.

Hash Bucket	Count	pair
0	1	(Tea, water), (Juice, Soda)
1	1	(Juice, water)
2	1	(Soda, water)
3	1	(coffee, Tea)
4	1	(coffee, Juice)
5	2	(coffee, Soda), (Tea, Juice)
6	2	(coffee, water), (Tea, Soda).

Count ≥ 2

Pair	Count
Tea, water	3
Juice, Soda	3
Coffee, Soda	4
Tea, Juice	4
coffee, water	2
Tea, Soda	4

partition 1

TID	Items
T1	Coffee, Tea, Juice
T2	Coffee, Soda
T3	Tea, Soda
T4	Coffee, Juice
T5	Coffee, Tea .

Partition 2

TID	Items
T6	Soda, Juice
T7	Coffee, Soda, Juice
T8	Tea, Juice
T9	Coffee, Tea, Soda
T10	Soda, water

Partition 3

TID	Items
T11	Coffee, Tea, water
T12	Tea, water
T13	Juice, Soda, water
T14	Coffee, Juice, water
T15	Tea, Soda, Juice .

Partition 1

Itemset	Count
Coffee	4
Tea	4
Juice	3
Soda	2
(Coffee, Tea)	2
(Coffee, Juice)	2
(Tea, Soda)	2

Partition 2

Itemset	Count
Coffee	3
Soda	4
Juice	3
(Soda, Juice)	2
(Coffee, Soda)	2

Partition 3

Itemset	Count
Coffee	2
Tea	3
Juice	3
Soda	3
water	3
(Tea, water)	2
(Juice, Soda)	2

Item Set	Count
Coffee	4 + 3 + 2 = 9
Tea	4 + 0 + 3 = 7
Juice	3 + 3 + 3 = 9
Soda	2 + 4 + 3 = 9
water	0 + 0 + 3 = 3
(Coffee, Juice)	2 + 0 + 0 = 2
Coffee, Tea	2 + 0 + 0 = 2
Tea, Soda	2 + 0 + 0 = 2
Soda, Juice	0 + 2 + 2 = 4
Coffee, Soda	0 + 2 + 0 = 2
Tea, water	0 + 0 + 2 = 2
Juice, Soda	0 + 0 + 2 = 2

frequent itemset :- Coffee, Tea, Juice, Soda, water, (soda, Juice)