URBAN EDGE

FP Growth Algorithm

① Consider the following Example. Apply FP-Growth Algorithm and Construct FP-tree. Given MinSup = 3

TID	items
1	a, b, c
2	b, c, d
3	d, e
4	a, b, d
5	a, b, c, e
6	a, b, c, d

NOTE: If you're not given "MinSup" please skip the "Step-1" and Draw the FP-tree for all transaction.

Soln: STEP 1: Count all items in whole transactions to get Support Count.

Items	Support
a	4
b	5
c	4
d	4
e	2

STEP 2: Apply the Threshold (MinSup) and remove item which passes Support Count less than MinSup.
In the above table, item 'e' Support count is less than MinSup, So it should be removed.

STEP 3: Arrange items in Descending Oder of there Support Count. and alter the table accordingly.

Item	Support
b	5
a	4
c	4
d	4

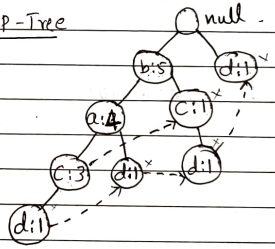
→

TID	items
1	b, a, c
2	b, c, d
3	d
4	b, a, d
5	b, a, c
6	b, a, c, d

∴ b, a, c, d will be the FP-tree construction pattern (as per descending Oder).

STEP 4: FP-Tree Construction, we go through Each of transaction with fall two Condition
* items whose Support Count is less than Min-Sup Should be removed from Transaction
* Descending-Oder item list (pattern) for FP-tree Construction should be maintained

FP-Tree



STEP 5: By using Constructed FP-Tree Compare MinSupport to get Association. ∴ Associations are {b, a, c}

→ Mining of FP-tree is Summarized below

Count of Each Item in Conditional Pattern Base

items	Conditional Pattern Base	Conditional FP-Tree	FP Generated
d	{b,a,c:1}, {b,c:1}	{b:2, c:2}	{b,d:2} {c,d:2} {b,c,d:2}
c	{b,a:3} {b:1}	{b:4, a:3}	{b,c:4} {a,c:3} {b,a,c:3}
a	{b:4}	{b:4}	{b,a:4}
b	{}		

minimum count should be considered

Frequent itemsets

① Min Sup = 3

TID	Items.
1	{M, O, N, K, E, Y}
2	{D, O, N, K, E, Y}
3	{A, E, K, M}
4	{C, K, M, U, Y}
5	{C, E, I, K, O}

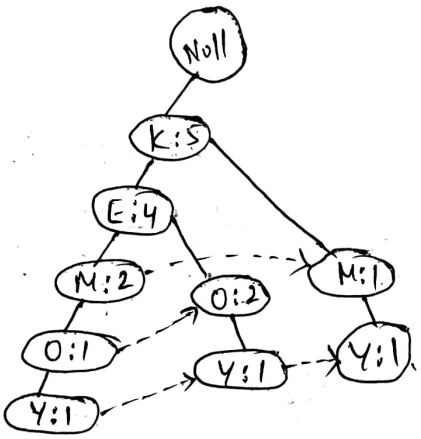
Sol^n:

A	1
C	2
D	1
E	4
I	1
K	5
M	3
N	2
O	3
U	1
Y	3

E:4, K:5, M:3, O:3, Y:3

K	5
E	4
M	3
O	3
Y	3

TID	Items.
1	K, E, M, O, Y
2	K, E, O, Y
3	K, E, M
4	K, M, Y
5	K, E, O

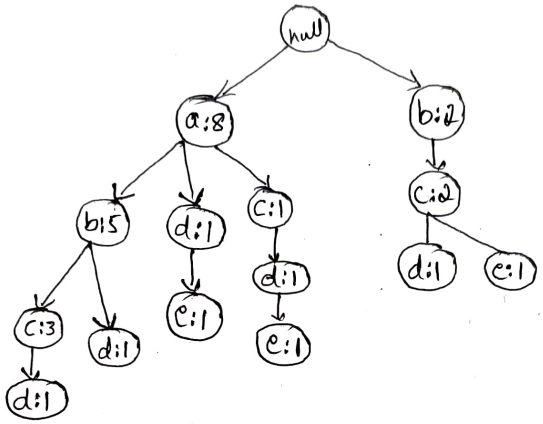


→ list of frequent itemset considering Suffix.

items	paths	Count of Each Item in path.	Candidate itemset with count of Each w.r.t. Updated transaction table
Y	{K, E, M, O: 1} {K, E, O: 1} {K, M: 1}	{K: 3} → others < Min Sup	{K, Y: 3}
O	{K, E, M: 1} {K, E: 2}	{K: 3} {E: 3}	{K, O: 3} {E, O: 3} {K, E, O: 3}
M	{K, E: 2} {K: 1}	{K: 3}	{K, M: 3}
E	{K: 4}	{K: 4}	{K, E: 4}
K	{5}		

FP- Growth Algorithm

TID	items
1	{a,b}
2	{b,c,d}
3	{a,c,d,e}
4	{a,d,e}
5	{a,b,c}
6	{a,b,c,d}
7	{a}
8	{a,b,c}
9	{a,b,d}
10	{b,c,e}



MinSup=2

Conditional FP Tree

items	Paths / Conditional Pattern Base	Count of Each item in path /	Candidate itemset with Count of Each w.r.t transaction table (check the Occurrence in Predicted transaction table)
e	{a,c,d:1}{a,d:1}{b,c:1}	{a:2}{b:1}{c:2}{d:2}	{a,e:2}{c,e:2}{d,e:2}{a,c,e:1}{a,d,e:2}{c,d,e:1}{a,c,d,e:1}
d	{a,b,c:1}{a,b:1}{a,c:1}{a:1}{b,c:1}	{a:4}{b:3}{c:3}	{a,d:4}{b,d:3}{c,d:3}{a,b,d:2}{a,c,d:2}{b,c,d:2}{a,b,c,d:1}
c	{a,b:3}{a:1}{b:2}	{a:4}{b:5}	{a,c:4}{b,c:3}{a,b,c:2}{c:6}
b	{a:5}{2}	{a:5}	{a,b:5}{b:7}
a	{8}		{a:8}

∴ The list of frequent itemsets ordered by their corresponding suffixes

e - {e, ae, ce, de, ade}
d - {d, ad, bd, cd, abd, acd, bcd}
c - {c, ac, bc, abc}
b - {b, ab}
a - {a}