

## Minimum Frequent Item set without Candidate Generation

### (FP Growth) FREQUENT PATTERN GROWTH.

- First find out the support count of items and arrange the items in frequently descending order.

T ID	List of itemids	
T <sub>100</sub>	<u>I<sub>1</sub></u> , <u>I<sub>2</sub></u> , <u>I<sub>5</sub></u>	I <sub>1</sub> - 6
T <sub>200</sub>	<u>I<sub>2</sub></u> , <u>I<sub>4</sub></u>	I <sub>2</sub> - 7
T <sub>300</sub>	<u>I<sub>2</sub></u> , <u>I<sub>3</sub></u>	I <sub>3</sub> - 6
T <sub>400</sub>	<u>I<sub>1</sub></u> , <u>I<sub>2</sub></u> , <u>I<sub>4</sub></u>	I <sub>4</sub> - 2
T <sub>500</sub>	<u>I<sub>1</sub></u> , <u>I<sub>3</sub></u>	I <sub>5</sub> - 2
T <sub>600</sub>	<u>I<sub>2</sub></u> , <u>I<sub>3</sub></u>	
T <sub>700</sub>	<u>I<sub>1</sub></u> , <u>I<sub>3</sub></u>	
T <sub>800</sub>	<u>I<sub>1</sub></u> , <u>I<sub>2</sub></u> , <u>I<sub>3</sub></u> , <u>I<sub>5</sub></u>	
T <sub>900</sub>	<u>I<sub>1</sub></u> , <u>I<sub>2</sub></u> , <u>I<sub>3</sub></u>	

→ First we have to find out min-support count table and prune it.

$$100\% = 9 \text{ transactions}$$

$$22\% = x$$

$$x = \frac{22}{100} \times 9 = 1.98 \approx 2$$

The minimum support count is 2

Itemset	support count
{I <sub>2</sub> }	7
{I <sub>1</sub> }	6
{I <sub>3</sub> }	6
{I <sub>4</sub> }	2
{I <sub>5</sub> }	2

T_ID	Items	
100	I <sub>1</sub> , I <sub>2</sub> , I <sub>5</sub>	I <sub>2</sub> , I <sub>1</sub> , I <sub>5</sub>
200	I <sub>2</sub> , I <sub>4</sub>	I <sub>2</sub> , I <sub>4</sub>
300	I <sub>2</sub> , I <sub>3</sub>	I <sub>2</sub> , I <sub>3</sub>
400	I <sub>1</sub> , I <sub>2</sub> , I <sub>4</sub>	I <sub>2</sub> , I <sub>1</sub> , I <sub>4</sub>
500	I <sub>1</sub> , I <sub>3</sub>	I <sub>1</sub> , I <sub>3</sub>
600	I <sub>2</sub> , I <sub>3</sub>	I <sub>2</sub> , I <sub>3</sub>
700	I <sub>1</sub> , I <sub>3</sub>	I <sub>1</sub> , I <sub>3</sub>
800	I <sub>1</sub> , I <sub>2</sub> , I <sub>3</sub> , I <sub>5</sub>	I <sub>2</sub> , I <sub>1</sub> , I <sub>3</sub> , I <sub>5</sub>
900	I <sub>1</sub> , I <sub>2</sub> , I <sub>3</sub>	I <sub>2</sub> , I <sub>1</sub> , I <sub>3</sub>

step 1 - Insert 1<sup>st</sup> transaction {I<sub>2</sub>, I<sub>1</sub>, I<sub>5</sub>}

Item ID	Support count	Node Link	
I <sub>2</sub>	7	→	I <sub>2</sub> :1
I <sub>1</sub>	6	→	I <sub>1</sub> :1
I <sub>3</sub>	6	→	I <sub>3</sub> :1
I <sub>4</sub>	2	→	I <sub>4</sub> :1
I <sub>5</sub>	2	→	I <sub>5</sub> :1

step 2 - Insert 2<sup>nd</sup> transaction {I<sub>2</sub>, I<sub>4</sub>}

Item ID	support count	Node Link	
I <sub>2</sub>	7	→	I <sub>2</sub> :2
I <sub>1</sub>	6	→	I <sub>1</sub> :1
I <sub>3</sub>	6	→	I <sub>3</sub> :1
I <sub>4</sub>	2	→	I <sub>4</sub> :1
I <sub>5</sub>	2	→	I <sub>5</sub> :1

Step 3 - Insert 3<sup>rd</sup> transaction {I<sub>2</sub>I<sub>3</sub>}

Item	support	Node
ID	count	Link
I <sub>2</sub>	7	•-----> I <sub>2</sub> :3
I <sub>1</sub>	6	•-----> I <sub>1</sub> :1
I <sub>3</sub>	6	•-----> I <sub>3</sub> :1
I <sub>4</sub>	2	•-----> I <sub>4</sub> :1
I <sub>5</sub>	2	•-----> I <sub>5</sub> :1

transaction

step 4 - Insert 4<sup>th</sup> transaction {I<sub>2</sub>I<sub>1</sub>I<sub>4</sub>}

Item	support	Node
ID	count	Link
I <sub>2</sub>	7	•-----> I <sub>2</sub> :4
I <sub>1</sub>	6	•-----> I <sub>1</sub> :2
I <sub>3</sub>	6	•-----> I <sub>3</sub> :1
I <sub>4</sub>	2	•-----> I <sub>4</sub> :1
I <sub>5</sub>	2	•-----> I <sub>5</sub> :1

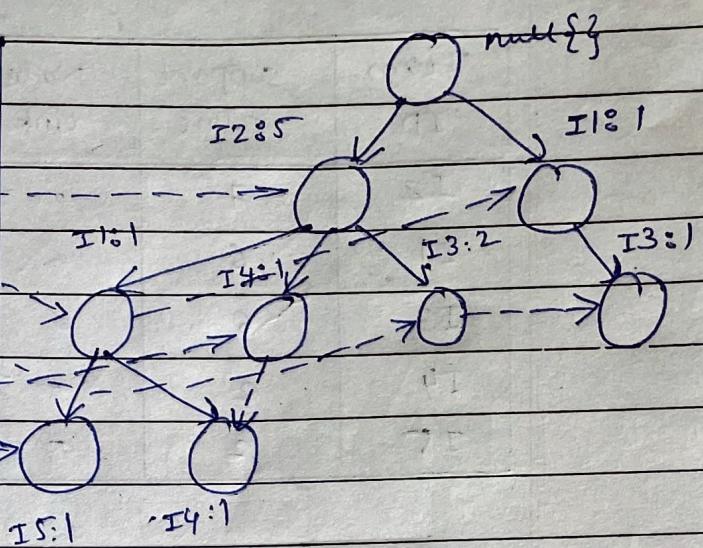
step

Item 5: Insert 5<sup>th</sup> transaction {I<sub>1</sub>, I<sub>3</sub>}

Item	support	Node
ID	count	Link
I <sub>2</sub>	7	•-----> I <sub>2</sub> :4
I <sub>1</sub>	6	•-----> I <sub>1</sub> :2
I <sub>3</sub>	6	•-----> I <sub>3</sub> :1
I <sub>4</sub>	2	•-----> I <sub>4</sub> :1
I <sub>5</sub>	2	•-----> I <sub>5</sub> :1

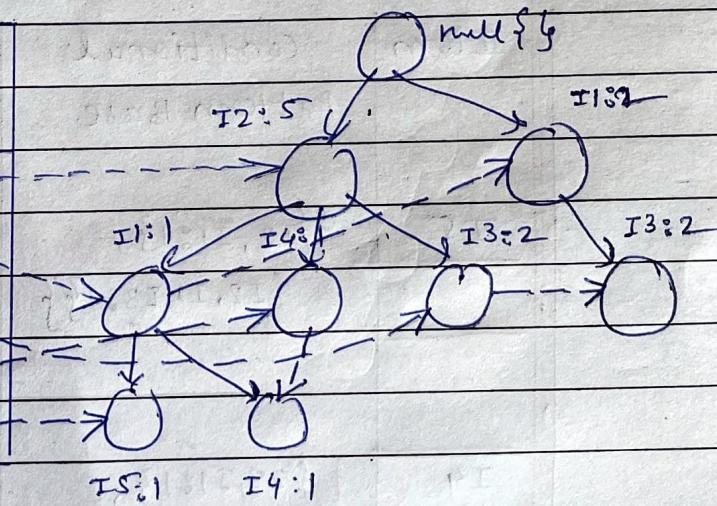
step 6 : Insert 6th transaction { I<sub>2</sub> I<sub>3</sub> }

Item	support	Node
ID	count	Link
I <sub>2</sub>	7	• - - - ->
I <sub>1</sub>	6	• - - - ->
I <sub>3</sub>	6	• - - - ->
I <sub>4</sub>	2	• - - - ->
I <sub>5</sub>	2	• - - - ->



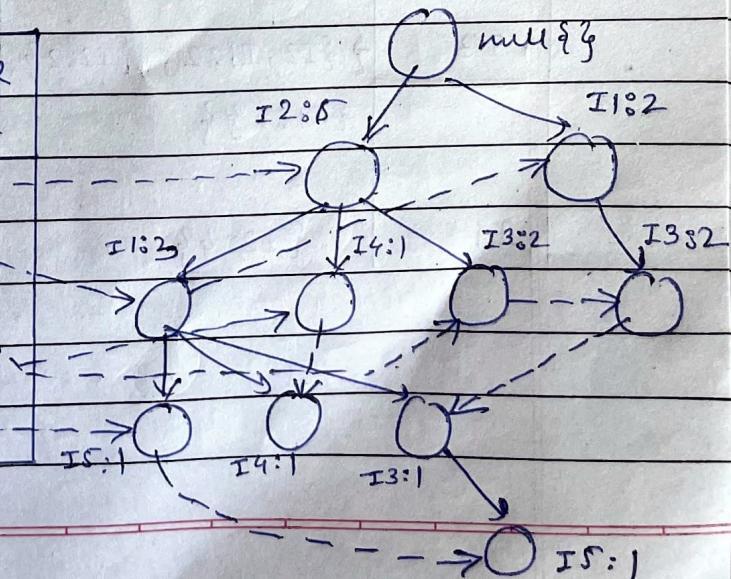
step 7 : Insert 7th transaction { I<sub>1</sub> I<sub>3</sub> }

Item	support	Node
set	count	Link
I <sub>2</sub>	7	• - - - ->
I <sub>1</sub>	6	• - - - ->
I <sub>3</sub>	6	• - - - ->
I <sub>4</sub>	2	• - - - ->
I <sub>5</sub>	2	• - - - ->



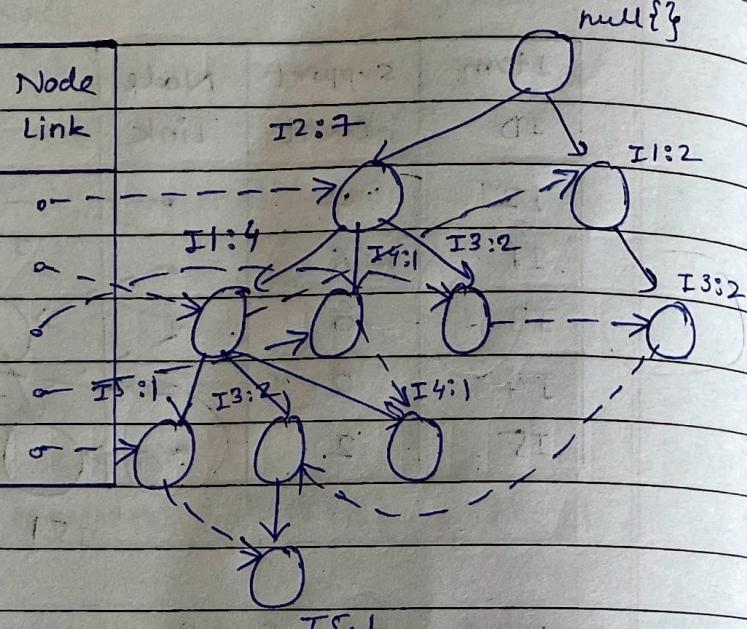
step 8 : Insert 8th transaction { I<sub>2</sub> I<sub>1</sub>, I<sub>3</sub> I<sub>5</sub> }

Item	support	Node
Set	count	Link
I <sub>2</sub>	7	• - - - ->
I <sub>1</sub>	6	• - - - ->
I <sub>3</sub>	6	• - - - ->
I <sub>4</sub>	2	• - - - ->
I <sub>5</sub>	2	• - - - ->



step 9: Insert 9<sup>th</sup> transaction {I2 I1 I3}

Item	support	Node
ID	count	Link
I2	7	o -
I1	6	a -
I3	6	a -
I4	2	o -
I5	2	o -



Item	Conditional Pattern Base	conditional FP tree	Freq pattern generation
I5	$\{\{I_2, I_1 : 1\}, \{I_2, I_1, I_3 : 1\}\}$	$\langle I_2 : 2, I_1 : 2 \rangle$	$\{I_2, I_5 : 2\}, \{I_1, I_5 : 2\}, \{I_2, I_1, I_5 : 2\}$
I4	$\{\{I_2, I_1 : 1\}, \{I_2 : 1\}\}$	$\langle I_2 : 2 \rangle$	$\{I_2, I_4 : 2\}$
I3	$\{\{I_2, I_1 : 2\}, \{I_2 : 2\}, \{I_1 : 2\}\}$	$\langle I_2 : 4, I_1 : 2 \rangle$	$\{I_2, I_3 : 4\}, \{I_1, I_3 : 4\}, \{I_2, I_1, I_3 : 2\}$
I1	$\{I_2 : 4\}$	$\langle I_2 : 4 \rangle$	$\{I_2, I_1 : 4\}$

The FP tree is mined as follows.

Start from each frequent length-1 pattern. construct its conditional pattern base, then construct its FP tree and perform mining recursively on such a tree. Mining the FP tree by creating conditional & sub-pattern bases. Start from the bottom of the header table.

- Item I5 path is I2, I1, I5 and I2, I1, I3, I5  
(I2 I1 : 1) (I2 I1 I3 : 1)
- Item I4 path is I2, I4 and I2 I1 I4  
(I2 : 1) and (I2 I1 : 1)
- Item I3 path is I2, I3 , I1, I3 and I2, I1, I3  
(I2 : 2) (I1 : 2) (I2 I1 : 2)
- Item I1 path is I2, I1  
(I2 : 4)
- Item I2 path is empty.

- 2) A database has 5 transactions. Let minimum support be 60% and min confidence be 80%. Find all frequent itemsets using Apriori and FP growth.

Tid	Items bought
T <sub>100</sub>	{ M, O, N, K, E, Y }
T <sub>200</sub>	{ D, O, N, K, E, Y }
T <sub>300</sub>	{ M, A, K, E }
T <sub>400</sub>	{ M, U, C, K, Y }
T <sub>500</sub>	{ C, O, O, K, I, E }