

# Association rules

Complete example

# Input

TID	List of item IDs
1	1,2,5
2	2,4
3	2,3
4	1,2,4
5	1,3
6	2,3
7	1,3
8	1,2,3,5
9	1,2,3

Min support threshold =  $2/9=22\%$

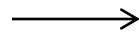
Min confidence threshold = 75%

# Generate all frequent itemsets

Frequent 1-itemsets. Candidates: {1,2,3,4,5}

Count candidates

C1	Counts
1	6
2	7
3	6
4	2
5	2



F1	Counts
1	6
2	7
3	6
4	2
5	2

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

# Candidate generation $C2 = F1 \times F1$

F1	1	2	3	4	5
F1					
1	{1,2}				
2	{1,3}				
3	{1,4}				
4	{1,5}				
5	{2,3}				
	{2,4}				
	{2,5}				
	{3,4}				
	{3,5}				
	{4,5}				

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

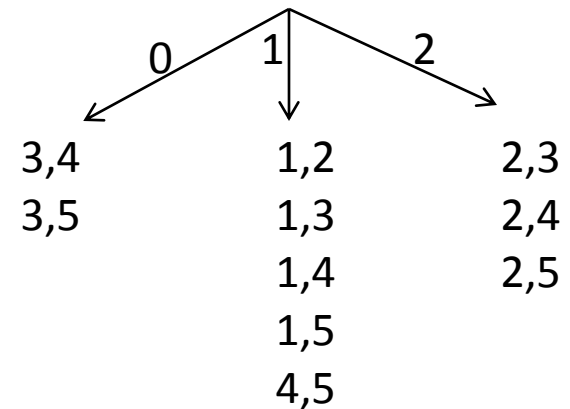
# Candidate C2 counting

C2	
{1,2}	
{1,3}	
{1,4}	
{1,5}	
{2,3}	
{2,4}	
{2,5}	
{3,4}	
{3,5}	
{4,5}	

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

Hash tree of candidates. Hash function: mod 3. Max leaf size = 3

$0 \bmod 3 = 3 \bmod 3 = 0$   
 $1 \bmod 3 = 4 \bmod 3 = 1$   
 $2 \bmod 3 = 5 \bmod 3 = 2$



Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

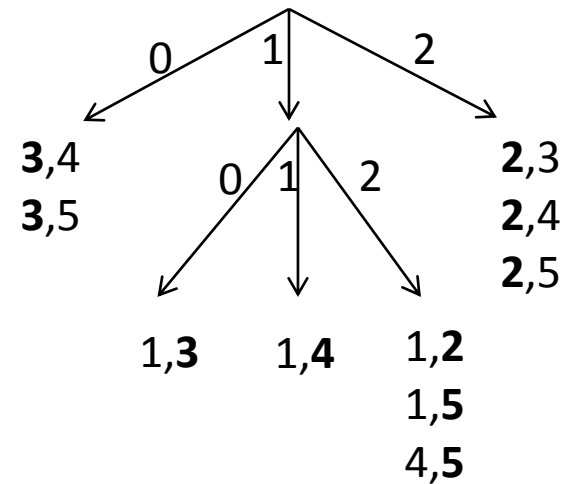
# Candidate C2 counting

C2	
{1,2}	
{1,3}	
{1,4}	
{1,5}	
{2,3}	
{2,4}	
{2,5}	
{3,4}	
{3,5}	
{4,5}	

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

Hash tree of candidates. Hash function: mod 3. Max leaf size = 3

$0 \bmod 3 = 3 \bmod 3 = 0$   
 $1 \bmod 3 = 4 \bmod 3 = 1$   
 $2 \bmod 3 = 5 \bmod 3 = 2$



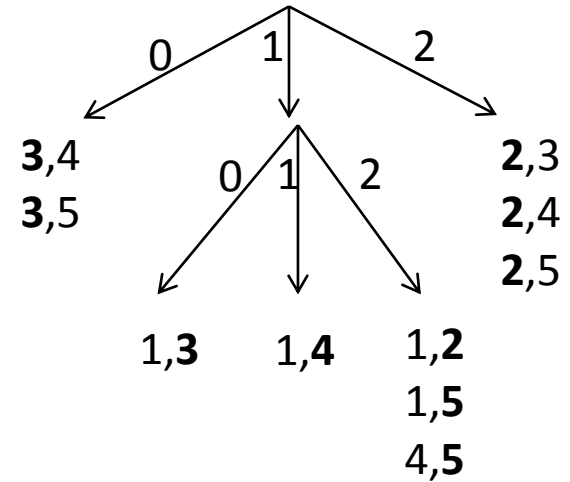
Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

# Candidate C2 counting

C2	
{1,2}	
{1,3}	
{1,4}	
{1,5}	
{2,3}	
{2,4}	
{2,5}	
{3,4}	
{3,5}	
{4,5}	

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

$0 \bmod 3 = 3 \bmod 3 = 0$   
 $1 \bmod 3 = 4 \bmod 3 = 1$   
 $2 \bmod 3 = 5 \bmod 3 = 2$



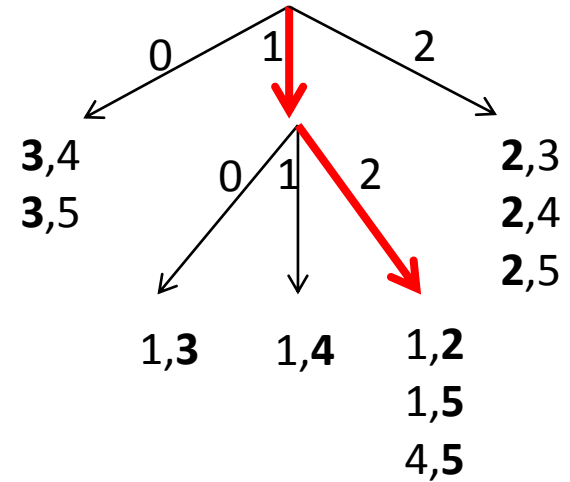
Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

# Candidate C2 counting

C2	
{1,2}	++
{1,3}	
{1,4}	
{1,5}	++
{2,3}	
{2,4}	
{2,5}	
{3,4}	
{3,5}	
{4,5}	

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

$0 \bmod 3 = 3 \bmod 3 = 0$   
 $1 \bmod 3 = 4 \bmod 3 = 1$   
 $2 \bmod 3 = 5 \bmod 3 = 2$



Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2



# Candidate C2 counting

C2	
{1,2}	++
{1,3}	
{1,4}	
{1,5}	++
{2,3}	
{2,4}	
{2,5}	++
{3,4}	
{3,5}	
{4,5}	

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

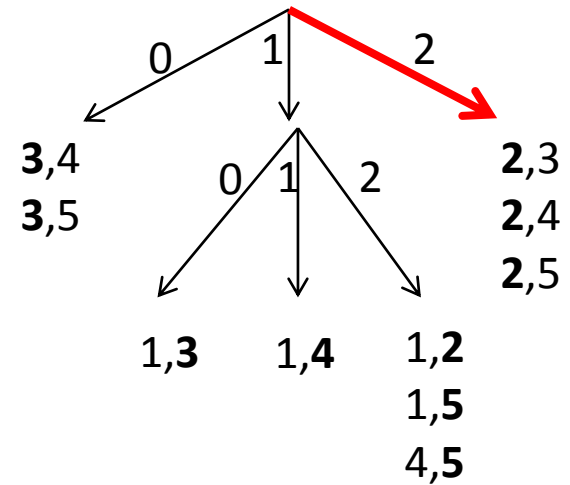
$0 \bmod 3 = 3 \bmod 3 = 0$

$1 \bmod 3 = 4 \bmod 3 = 1$

$2 \bmod 3 = 5 \bmod 3 = 2$



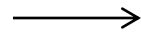
6 comparisons instead 10



Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

# Candidate C2 counting ...

C2	Count
{1,2}	4
{1,3}	4
{1,4}	1
{1,5}	2
{2,3}	4
{2,4}	2
{2,5}	2
{3,4}	0
{3,5}	1
{4,5}	0



C2	Count
{1,2}	4
{1,3}	4
{1,5}	2
{2,3}	4
{2,4}	2
{2,5}	2

Frequent 2-itemsets:

{1,2}:4, {1,3}:4, {1,5}:2, {2,3}:4, {2,4}:2, {2,5}:2

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

# Candidate generation $C3 = F2 \times F2$

C2	{1,2}	{1,3}	{1,5}	{2,3}	{2,4}	{2,5}

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

C2
{1,2}
{1,3}
{1,5}
{2,3}
{2,4}
{2,5}

C3	
{1,2,3}	
{1,2,5}	
{1,3,5}	
{2,3,4}	
{2,3,5}	
{2,4,5}	

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {1,5}:2, {2,3}:4, {2,4}:2, {2,5}:2

# Candidate pruning C3

C2
{1,2}
{1,3}
{1,5}
{2,3}
{2,4}
{2,5}

C3	
{1,2,3}	
{1,2,5}	
{1,3,5}	
{2,3,4}	
{2,3,5}	
{2,4,5}	

C3	
{1,2,3}	
{1,2,5}	

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {1,5}:2, {2,3}:4, {2,4}:2, {2,5}:2

# Candidate C3 counting

C3	Count
{1,2,3}	2
{1,2,5}	2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {1,5}:2, {2,3}:4, {2,4}:2, {2,5}:2

# All frequent itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {1,5}:2, {2,3}:4, {2,4}:2, {2,5}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

## Maximal frequent itemsets

(cannot be extended into frequent super-itemset):

1-itemsets: ~~{1}:6~~, ~~{2}:7~~, ~~{3}:6~~, {4}:2, ~~{5}:2~~

2-itemsets: ~~{1,2}:4~~, ~~{1,3}:4~~, ~~{1,5}:2~~, ~~{2,3}:4~~, {2,4}:2, ~~{2,5}:2~~

3-itemsets: {1,2,3}:2, {1,2,5}: 2

# All frequent itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6, {4}:2, {5}:2

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {1,5}:2, {2,3}:4, {2,4}:2, {2,5}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

## Maximal frequent itemsets

(cannot be extended into frequent super-itemset):

2-itemsets: {2,4}:2

3-itemsets: {1,2,3}:2, {1,2,5}: 2

## Closed frequent itemsets

(cannot be extended into frequent super-itemset with the same support count):

1-itemsets: {1}:6, {2}:7, {3}:6, ~~{4}:2~~, ~~{5}:2~~

2-itemsets: {1,2}:4, {1,3}:4, ~~{1,5}:2~~, {2,3}:4, {2,4}:2, ~~{2,5}:2~~

3-itemsets: {1,2,3}:2, {1,2,5}: 2

# Output: closed frequent itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {2,3}:4, {2,4}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

Enough information to generate rules

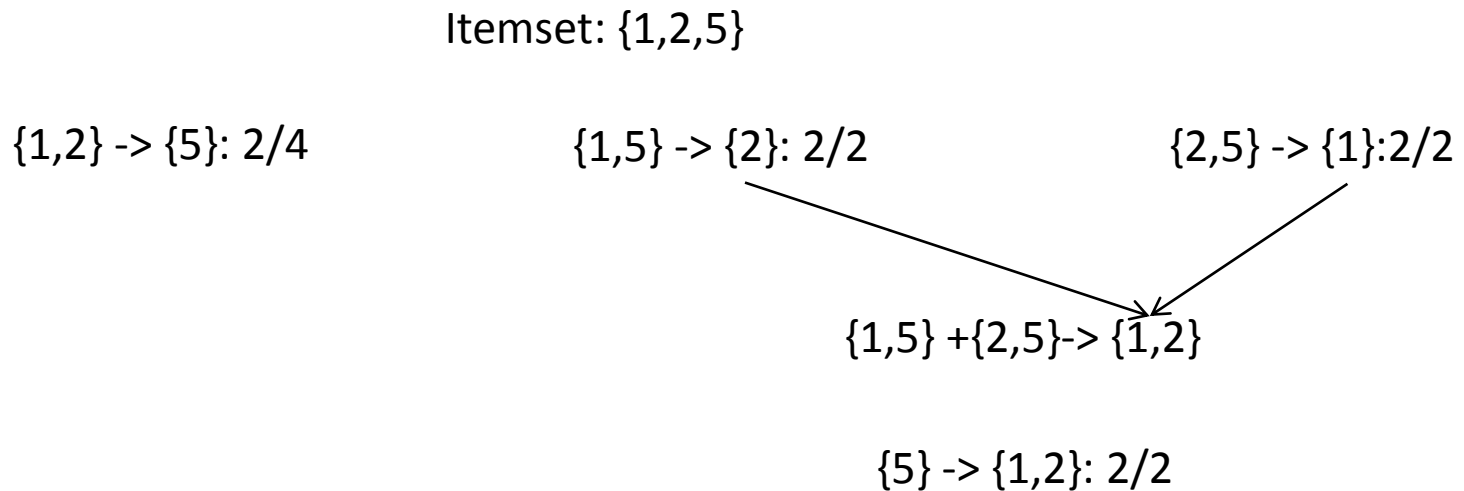


# Rule generation: 3-itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {2,3}:4, {2,4}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2



Rules:  $\{1,5\} \rightarrow \{2\}: 2/9, 1$ ;  $\{2,5\} \rightarrow \{1\}: 2/9, 1$ ;  $\{5\} \rightarrow \{1,2\}: 2/9, 1$

# Rule generation: 3-itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {2,3}:4, {2,4}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

Itemset: {1,2,3}

{1,2} -> {3}: 2/4

{1,3} -> {2}: 2/4

{2,3} -> {1}:2/4

All do not pass confidence threshold

Rules: {1,5} -> {2}: 2/9, 1; {2,5} -> {1}: 2/9, 1; {5} -> {1,2}: 2/9, 1

# Rule generation: 2-itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {2,3}:4, {2,4}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

{1} -> {2}: 4/6

{1} -> {3}: 4/6

{2} -> {3}: 4/7

{2} -> {4}: 2/7

{2} -> {1}: 4/7

{3} -> {1}: 4/6

{3} -> {2}: 4/6

{4} -> {2}: 2/2

{1} -> {5}: 2/6

{2} -> {5}: 2/7

{5} -> {1}: 2/2

{5} -> {2}: 2/2

Rules: {1,5} -> {2}: 2/9, 1; {2,5} -> {1}: 2/9, 1; {5} -> {1,2}: 2/9, 1

# Rule generation: 2-itemsets

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {2,3}:4, {2,4}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

~~{1} → {2}: 4/6~~  
~~{2} → {1}: 4/7~~

~~{1} → {3}: 4/6~~  
~~{3} → {1}: 4/6~~

~~{2} → {3}: 4/7~~  
~~{3} → {2}: 4/6~~

~~{2} → {4}: 2/7~~  
~~{4} → {2}: 2/2~~

~~{1} → {5}: 2/6~~  
~~{5} → {1}: 2/2~~

~~{2} → {5}: 2/7~~  
~~{5} → {2}: 2/2~~

Rules: {1,5} → {2}: 2/9, 1; {2,5} → {1}: 2/9, 1; {5} → {1,2}: 2/9, 1

# All association rules

Frequent 1-itemsets: {1}:6, {2}:7, {3}:6

Frequent 2-itemsets: {1,2}:4, {1,3}:4, {2,3}:4, {2,4}:2

Frequent 3-itemsets: {1,2,3}:2, {1,2,5}: 2

Rules:

{1,5} -> {2}: 2/9, 1; {2,5} -> {1}: 2/9, 1; {5} -> {1,2}: 2/9, 1

{4} -> {2}: 2/9, 1; {5} -> {1}: 2/9, 1; {5} -> {2}: 2/9, 1

List of item IDs
1,2,5
2,4
2,3
1,2,4
1,3
2,3
1,3
1,2,3,5
1,2,3