# table 5.1, Transactional data, Han and Kamber (2006) p.236
> # items : I1 I2 I3 I4 I5
> # dataset: total data = 9
> # Transaction ID   Items
> # T100            {I1,I2,I5},
> # T200            {I2,I4},
> # T300            {I2,I3},
> # T400            {I1,I2,I4},
> # T500            {I1,I3},
> # T600            {I2,I3},
> # T700            {I1,I3},
> # T800            {I1,I2,I3,I5},
> # T900            {I1,I2,I3}
>
> # step 1.
> # load "arules" package
> library(arules)
Loading required package: Matrix
Loading required package: lattice

Attaching package: 'Matrix'

The following object(s) are masked from package:stats :
xtabs

The following object(s) are masked from package:base :
colMeans,  
colSums,  
rcond,  
rowMeans,  
rowSums

Attaching package: 'arules'

The following object(s) are masked from package:base :

%in%

> # step 2.
> # prepare data
> a_list <- list(  
+   c("I1","I2","I5"),  
+   c("I2","I4"),  
+   c("I2","I3"),  
+   c("I1","I2","I4"),  
+   c("I1","I3"),  
+   c("I2","I3"),  
+   c("I1","I3"),  
+   c("I1","I2","I3","I5"),  
+   c("I1","I2","I3")  
+ )
>
> # set transaction names
> names(a_list) <- paste("T",1:9, "00", sep = "")
> a_list

$T100
[1] "I1" "I2" "I5"

$T200
[1] "I2" "I4"
$T300
(1) "I2" "I3"

$T400
(1) "I1" "I2" "I4"

$T500
(1) "I1" "I3"

$T600
(1) "I2" "I3"

$T700
(1) "I1" "I3"

$T800
(1) "I1" "I2" "I3" "I5"

$T900
(1) "I1" "I2" "I3"

> # force data into transactions
> table5_1 <- as(a_list, "transactions") # Force an Object to Belong
> to a Class >as (Object, Class)
> table5_1
transactions in sparse format with
9 transactions (rows) and
5 items (columns)
>
> # step 3.
> # analyze data
> # generate level plots to visually inspect binary incidence matrices
> image(table5_1) # result- Figure 1 Level plot
> summary(table5_1)
transactions as itemMatrix in sparse format with
9 rows (elements/itemsets/transactions) and
5 columns (items) and a density of 0.5111111
most frequent items:

<table>
<thead>
<tr>
<th>I2</th>
<th>I1</th>
<th>I3</th>
<th>I4</th>
<th>I5 (Other)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

element (itemset/transaction) length distribution:

sizes
- 2 3 4
- 5 3 1

Min. 1st Qu. Median Mean 3rd Qu. Max.
2.000 2.000 2.000 2.556 3.000 4.000

includes extended item information - examples:

labels
- 1  I1
- 2  I2
- 3  I3

includes extended transaction information - examples:

transactionID
- 1  T100
- 2  T200
- 3  T300

> # step 4.
> # find 1-items (L1)

> # provides the generic function itemFrequency and the frequency/support for all single items in an objects based on itemMatrix.

> itemFrequency(table5_1, type = "relative") # default: "relative"

<table>
<thead>
<tr>
<th>I1</th>
<th>I2</th>
<th>I3</th>
<th>I4</th>
<th>I5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6666667</td>
<td>0.7777778</td>
<td>0.6666667</td>
<td>0.2222222</td>
<td>0.2222222</td>
</tr>
</tbody>
</table>

> itemFrequency(table5_1, type = "absolute") # same as the textbook

I1 I2 I3 I4 I5
6 7 6 2 2

> # step 5.
> # create an item frequency bar plot for inspecting the item frequency
> distribution for objects based on itemMatrix
> itemFrequencyPlot(table5_1) # result-Figure 2 Item frequency bar plot
> # step 6.
> # mine association rules
> # rules <- apriori(table5_1, parameter = list(supp = 0.5, conf = 0.9, target = "rules"))
> rules<- apriori(table5_1) # Mine frequent itemsets, association rules or association hyperedges using the Apriori algorithm

parameter specification:
  confidence minval smax arem aval originalSupport support minlen maxlen 0.8 0.1 1 none FALSE TRUE 0.1 1 5
  target ext
  rules FALSE

algorithmic control:
  filter tree heap memopt load sort verbose
  0.1 TRUE TRUE FALSE TRUE 2 TRUE

apriori - find association rules with the apriori algorithm
version 4.21 (2004.05.09) (c) 1996-2004 Christian Borgelt
set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[5 item(s), 9 transaction(s)] done [0.00s].
sorting and recoding items ... [5 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 done [0.00s].
writing ... [10 rule(s)] done [0.00s].
creating S4 object ... done [0.00s].
>
> # step7.
> # display results
> inspect(table5_1) # display transactions
  items transactionID
  1 {I1, I2, I5} T100
  2 {I2,}
3 \{I2, \\
  I3\} \rightarrow \{I2\} \\
4 \{I1, \\
  I2, \\
  I4\} \rightarrow \{I2\} \\
5 \{I1, \\
  I3\} \rightarrow \{I2\} \\
6 \{I2, \\
  I3\} \rightarrow \{I2\} \\
7 \{I1, \\
  I3\} \rightarrow \{I2\} \\
8 \{I1, \\
  I2, \\
  I3, \\
  I5\} \rightarrow \{I2\} \\
9 \{I1, \\
  I2, \\
  I3\} \rightarrow \{I2\} \\
> \text{inspect(rules) \# display association}

<table>
<thead>
<tr>
<th>lhs</th>
<th>rhs</th>
<th>support</th>
<th>confidence</th>
<th>lift</th>
</tr>
</thead>
<tbody>
<tr>
<td>{I4}</td>
<td>{I2}</td>
<td>0.222222</td>
<td>1.285714</td>
<td>1.285714</td>
</tr>
<tr>
<td>{I5}</td>
<td>{I1}</td>
<td>0.222222</td>
<td>1.500000</td>
<td>1.500000</td>
</tr>
<tr>
<td>{I5}</td>
<td>{I2}</td>
<td>0.222222</td>
<td>1.285714</td>
<td>1.285714</td>
</tr>
</tbody>
</table>
| \{I1, \\
  I4\} | \{I2\} | 0.111111 | 1.285714 | 1.285714 |
| \{I3, \\
  I5\} | \{I1\} | 0.111111 | 1.500000 | 1.500000 |
| \{I3, \\
  I5\} | \{I2\} | 0.111111 | 1.285714 | 1.285714 |
| \{I1, \\
  I5\} | \{I2\} | 0.222222 | 1.285714 | 1.285714 |
| \{I2, \\
  I5\} | \{I1\} | 0.222222 | 1.500000 | 1.500000 |
| \{I1, \\
  I3, \\
  I5\} | \{I2\} | 0.111111 | 1.285714 | 1.285714 |
| \{I2, \\
  I5\} | \{I1\} | 0.111111 | 1.285714 | 1.285714 |
I3,
I5) => (I1) 0.1111111 1 1.500000

> # reference:
> # http://r-forge.r-project.org/projects/arules
> # end

Figure 1 Level plot
Reference