AssociationRules.R

Wow

Mon Nov 05 15:49:35 2018

```
#####################################################
#            Association Rule Mining
#####################################################
library(mlbench)
data(Zoo)
head(Zoo)
```

```
##           hair feathers  eggs  milk airborne aquatic predator toothed
## aardvark  TRUE    FALSE FALSE  TRUE    FALSE   FALSE     TRUE    TRUE
## antelope  TRUE    FALSE FALSE  TRUE    FALSE   FALSE    FALSE    TRUE
## bass     FALSE    FALSE  TRUE FALSE    FALSE    TRUE     TRUE    TRUE
## bear      TRUE    FALSE FALSE  TRUE    FALSE   FALSE     TRUE    TRUE
## boar      TRUE    FALSE FALSE  TRUE    FALSE   FALSE     TRUE    TRUE
## buffalo   TRUE    FALSE FALSE  TRUE    FALSE   FALSE    FALSE    TRUE
##          backbone breathes venomous  fins legs  tail domestic catsize
## aardvark     TRUE     TRUE    FALSE FALSE    4 FALSE    FALSE    TRUE
## antelope     TRUE     TRUE    FALSE FALSE    4  TRUE    FALSE    TRUE
## bass         TRUE    FALSE    FALSE  TRUE    0  TRUE    FALSE   FALSE
## bear         TRUE     TRUE    FALSE FALSE    4 FALSE    FALSE    TRUE
## boar         TRUE     TRUE    FALSE FALSE    4  TRUE    FALSE    TRUE
## buffalo      TRUE     TRUE    FALSE FALSE    4  TRUE    FALSE    TRUE
##            type
## aardvark mammal
## antelope mammal
## bass       fish
## bear     mammal
## boar     mammal
## buffalo  mammal
```

```
library(Matrix)
library(arules)
```

```
##
## Attaching package: 'arules'
```

```
## The following objects are masked from 'package:base':
##
##     abbreviate, write
```

```
summary(Zoo)
```

```
##     hair          feathers          eggs            milk
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
```

```
##  FALSE:58        FALSE:81        FALSE:42        FALSE:60
##  TRUE :43        TRUE :20        TRUE :59        TRUE :41
##
##
##
##
##   airborne        aquatic         predator        toothed
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:77        FALSE:65        FALSE:45        FALSE:40
##  TRUE :24        TRUE :36        TRUE :56        TRUE :61
##
##
##
##
##   backbone        breathes        venomous          fins
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:18        FALSE:21        FALSE:93        FALSE:84
##  TRUE :83        TRUE :80        TRUE :8         TRUE :17
##
##
##
##
##       legs           tail          domestic        catsize
##  Min.   :0.000   Mode :logical   Mode :logical   Mode :logical
##  1st Qu.:2.000   FALSE:26        FALSE:88        FALSE:57
##  Median :4.000   TRUE :75        TRUE :13        TRUE :44
##  Mean   :2.842
##  3rd Qu.:4.000
##  Max.   :8.000
##
##             type
##  mammal       :41
##  bird         :20
##  reptile      : 5
##  fish         :13
##  amphibian    : 4
##  insect       : 8
##  mollusc.et.al:10
```

```
# Create Items
#items <- as(Zoo, "transactions")

# We need to fix the error of column 13 first
class(Zoo[,13])
```

```
## [1] "integer"
```

```
table(Zoo[,13])
```

```
##
##  0  2  4  5  6  8
## 23 27 38  1 10  2
```

```
has_legs <- Zoo[,13]>0 #Convert legs attribute into binary class attribute
has_legs
```

```
##   [1]  TRUE  TRUE FALSE  TRUE  TRUE  TRUE  TRUE FALSE FALSE  TRUE  TRUE
##  [12]  TRUE FALSE FALSE  TRUE  TRUE  TRUE  TRUE FALSE FALSE  TRUE  TRUE
##  [23]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
##  [34]  TRUE FALSE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE  TRUE  TRUE  TRUE
##  [45]  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
##  [56]  TRUE  TRUE  TRUE  TRUE  TRUE FALSE FALSE FALSE  TRUE  TRUE  TRUE
##  [67] FALSE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE FALSE FALSE  TRUE FALSE
##  [78] FALSE  TRUE  TRUE FALSE FALSE FALSE  TRUE  TRUE  TRUE FALSE  TRUE
##  [89]  TRUE  TRUE  TRUE  TRUE FALSE  TRUE  TRUE  TRUE  TRUE  TRUE  TRUE
## [100] FALSE  TRUE
```

```
table(has_legs)
```

```
## has_legs
## FALSE  TRUE
##    23    78
```

```
Zoo[,13] <- has_legs
colnames(Zoo)[13] <- "has_legs" #rename column 13 from legs to has_legs
head(Zoo)
```

```
##           hair feathers  eggs  milk airborne aquatic predator toothed
## aardvark  TRUE    FALSE FALSE  TRUE    FALSE   FALSE     TRUE    TRUE
## antelope  TRUE    FALSE FALSE  TRUE    FALSE   FALSE    FALSE    TRUE
## bass     FALSE    FALSE  TRUE FALSE    FALSE    TRUE     TRUE    TRUE
## bear      TRUE    FALSE FALSE  TRUE    FALSE   FALSE     TRUE    TRUE
## boar      TRUE    FALSE FALSE  TRUE    FALSE   FALSE     TRUE    TRUE
## buffalo   TRUE    FALSE FALSE  TRUE    FALSE   FALSE    FALSE    TRUE
##          backbone breathes venomous  fins has_legs  tail domestic catsize
## aardvark     TRUE     TRUE    FALSE FALSE     TRUE FALSE    FALSE    TRUE
## antelope     TRUE     TRUE    FALSE FALSE     TRUE  TRUE    FALSE    TRUE
## bass         TRUE    FALSE    FALSE  TRUE    FALSE  TRUE    FALSE   FALSE
## bear         TRUE     TRUE    FALSE FALSE     TRUE FALSE    FALSE    TRUE
## boar         TRUE     TRUE    FALSE FALSE     TRUE  TRUE    FALSE    TRUE
## buffalo      TRUE     TRUE    FALSE FALSE     TRUE  TRUE    FALSE    TRUE
##            type
## aardvark mammal
## antelope mammal
## bass       fish
## bear     mammal
## boar     mammal
## buffalo  mammal
```

```r
# Alternatives for dealing with numeric factor:
# 1. Use each unique value as an item: Zoo[,13] <- as.factor([,13])
# 2. Use discretize for continuous data (see ?discretize ): Zoo[,13] <-
discretize(legs,
#    categories = 2, method="interval")

# Convert all logical attributes into factor attributes
for(i in 1:ncol(Zoo)){
  Zoo[,i] <- as.factor(Zoo[,i])
}

# Convert data into a set of items
items <- as(Zoo, "transactions") #All attributes need to be a factor!
items

## transactions in sparse format with
##  101 transactions (rows) and
##  39 items (columns)

# Inspect transactions
colnames(items)

##  [1] "hair=FALSE"         "hair=TRUE"          "feathers=FALSE"
##  [4] "feathers=TRUE"      "eggs=FALSE"         "eggs=TRUE"
##  [7] "milk=FALSE"         "milk=TRUE"          "airborne=FALSE"
## [10] "airborne=TRUE"      "aquatic=FALSE"      "aquatic=TRUE"
## [13] "predator=FALSE"     "predator=TRUE"      "toothed=FALSE"
## [16] "toothed=TRUE"       "backbone=FALSE"     "backbone=TRUE"
## [19] "breathes=FALSE"     "breathes=TRUE"      "venomous=FALSE"
## [22] "venomous=TRUE"      "fins=FALSE"         "fins=TRUE"
## [25] "has_legs=FALSE"     "has_legs=TRUE"      "tail=FALSE"
## [28] "tail=TRUE"          "domestic=FALSE"     "domestic=TRUE"
## [31] "catsize=FALSE"      "catsize=TRUE"       "type=mammal"
## [34] "type=bird"          "type=reptile"       "type=fish"
## [37] "type=amphibian"     "type=insect"        "type=mollusc.et.al"

inspect(head(items))

##     items            transactionID
## [1] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
```

```
##      has_legs=TRUE,
##      tail=FALSE,
##      domestic=FALSE,
##      catsize=TRUE,
##      type=mammal}         aardvark
## [2] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=FALSE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=TRUE,
##      domestic=FALSE,
##      catsize=TRUE,
##      type=mammal}         antelope
## [3] {hair=FALSE,
##      feathers=FALSE,
##      eggs=TRUE,
##      milk=FALSE,
##      airborne=FALSE,
##      aquatic=TRUE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=FALSE,
##      venomous=FALSE,
##      fins=TRUE,
##      has_legs=FALSE,
##      tail=TRUE,
##      domestic=FALSE,
##      catsize=FALSE,
##      type=fish}           bass
## [4] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
```

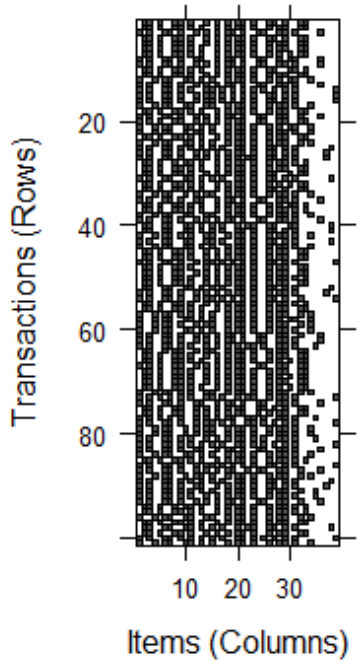
```
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=FALSE,
##      domestic=FALSE,
##      catsize=TRUE,
##      type=mammal}         bear
## [5] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=TRUE,
##      domestic=FALSE,
##      catsize=TRUE,
##      type=mammal}         boar
## [6] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=FALSE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=TRUE,
##      domestic=FALSE,
##      catsize=TRUE,
##      type=mammal}         buffalo

# Look at a first 3 transactions as a matrix.
as(items, "matrix")[1:3,]

##          hair=FALSE hair=TRUE feathers=FALSE feathers=TRUE eggs=FALSE
## aardvark      FALSE      TRUE           TRUE          FALSE       TRUE
## antelope      FALSE      TRUE           TRUE          FALSE       TRUE
## bass           TRUE     FALSE           TRUE          FALSE      FALSE
##          eggs=TRUE milk=FALSE milk=TRUE airborne=FALSE airborne=TRUE
## aardvark     FALSE      FALSE      TRUE           TRUE          FALSE
```

```
## antelope     FALSE      FALSE      TRUE           TRUE         FALSE
## bass          TRUE       TRUE     FALSE           TRUE         FALSE
##          aquatic=FALSE aquatic=TRUE predator=FALSE predator=TRUE
## aardvark          TRUE        FALSE          FALSE          TRUE
## antelope          TRUE        FALSE           TRUE         FALSE
## bass             FALSE         TRUE          FALSE          TRUE
##          toothed=FALSE toothed=TRUE backbone=FALSE backbone=TRUE
## aardvark         FALSE         TRUE          FALSE          TRUE
## antelope         FALSE         TRUE          FALSE          TRUE
## bass             FALSE         TRUE          FALSE          TRUE
##          breathes=FALSE breathes=TRUE venomous=FALSE venomous=TRUE
## aardvark          FALSE          TRUE           TRUE         FALSE
## antelope          FALSE          TRUE           TRUE         FALSE
## bass               TRUE         FALSE           TRUE         FALSE
##          fins=FALSE fins=TRUE has_legs=FALSE has_legs=TRUE tail=FALSE
## aardvark       TRUE     FALSE          FALSE          TRUE       TRUE
## antelope       TRUE     FALSE          FALSE          TRUE      FALSE
## bass          FALSE      TRUE           TRUE         FALSE      FALSE
##          tail=TRUE domestic=FALSE domestic=TRUE catsize=FALSE catsize=TRUE
## aardvark     FALSE           TRUE         FALSE         FALSE         TRUE
## antelope      TRUE           TRUE         FALSE         FALSE         TRUE
## bass          TRUE           TRUE         FALSE          TRUE        FALSE
##          type=mammal type=bird type=reptile type=fish type=amphibian
## aardvark        TRUE     FALSE        FALSE     FALSE          FALSE
## antelope        TRUE     FALSE        FALSE     FALSE          FALSE
## bass           FALSE     FALSE        FALSE      TRUE          FALSE
##          type=insect type=mollusc.et.al
## aardvark       FALSE              FALSE
## antelope       FALSE              FALSE
## bass           FALSE              FALSE
```

```r
# Look at the transaction 1,2, and 3 as sets of items
inspect(items[1:3])
```

```
##     items            transactionID
## [1] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=FALSE,
##      domestic=FALSE,
```

```
##      catsize=TRUE,
##      type=mammal}         aardvark
## [2] {hair=TRUE,
##      feathers=FALSE,
##      eggs=FALSE,
##      milk=TRUE,
##      airborne=FALSE,
##      aquatic=FALSE,
##      predator=FALSE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=TRUE,
##      domestic=FALSE,
##      catsize=TRUE,
##      type=mammal}         antelope
## [3] {hair=FALSE,
##      feathers=FALSE,
##      eggs=TRUE,
##      milk=FALSE,
##      airborne=FALSE,
##      aquatic=TRUE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=FALSE,
##      venomous=FALSE,
##      fins=TRUE,
##      has_legs=FALSE,
##      tail=TRUE,
##      domestic=FALSE,
##      catsize=FALSE,
##      type=fish}           bass

# Plot the binary matrix. Dark dots represent of a particular item in a
particular transaction
image(items)
```




```
# Look at the support (relative frequency) of items in the data set. Here we
look at the 10 most frequent items
itemFrequency(items,type="absolute") #support count
```

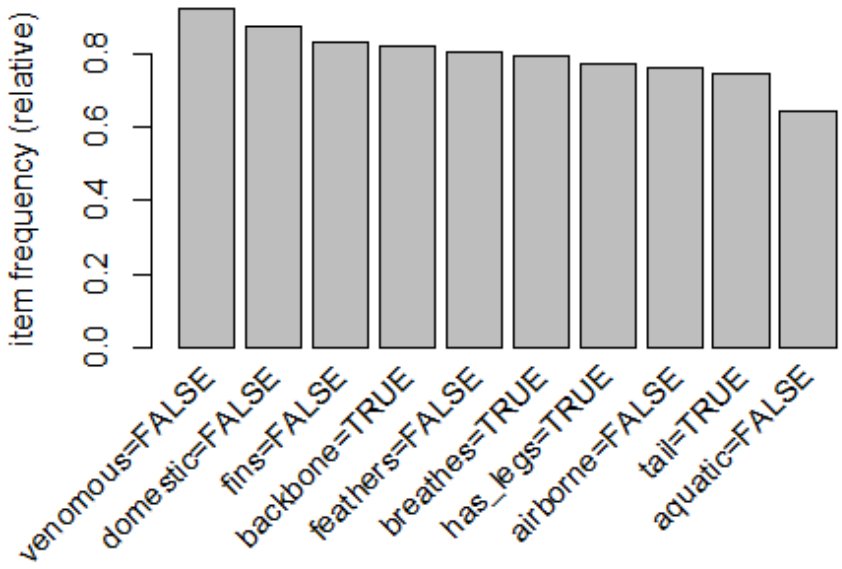
```
##         hair=FALSE          hair=TRUE     feathers=FALSE
##                 58                 43                 81
##      feathers=TRUE         eggs=FALSE          eggs=TRUE
##                 20                 42                 59
##         milk=FALSE          milk=TRUE     airborne=FALSE
##                 60                 41                 77
##      airborne=TRUE      aquatic=FALSE       aquatic=TRUE
##                 24                 65                 36
##     predator=FALSE      predator=TRUE      toothed=FALSE
##                 45                 56                 40
##       toothed=TRUE     backbone=FALSE      backbone=TRUE
##                 61                 18                 83
##     breathes=FALSE      breathes=TRUE     venomous=FALSE
##                 21                 80                 93
##      venomous=TRUE         fins=FALSE          fins=TRUE
##                  8                 84                 17
##     has_legs=FALSE      has_legs=TRUE         tail=FALSE
##                 23                 78                 26
##          tail=TRUE     domestic=FALSE      domestic=TRUE
##                 75                 88                 13
##      catsize=FALSE       catsize=TRUE        type=mammal
##                 57                 44                 41
##          type=bird       type=reptile          type=fish
```

```
##                  20                   5                  13
##     type=amphibian        type=insect type=mollusc.et.al
##                   4                   8                  10
```

```
itemFrequency(items,type="relative") #support
```

```
##         hair=FALSE          hair=TRUE     feathers=FALSE
##         0.57425743         0.42574257         0.80198020
##      feathers=TRUE         eggs=FALSE          eggs=TRUE
##         0.19801980         0.41584158         0.58415842
##         milk=FALSE          milk=TRUE     airborne=FALSE
##         0.59405941         0.40594059         0.76237624
##      airborne=TRUE      aquatic=FALSE       aquatic=TRUE
##         0.23762376         0.64356436         0.35643564
##     predator=FALSE      predator=TRUE      toothed=FALSE
##         0.44554455         0.55445545         0.39603960
##       toothed=TRUE     backbone=FALSE      backbone=TRUE
##         0.60396040         0.17821782         0.82178218
##     breathes=FALSE      breathes=TRUE     venomous=FALSE
##         0.20792079         0.79207921         0.92079208
##      venomous=TRUE         fins=FALSE          fins=TRUE
##         0.07920792         0.83168317         0.16831683
##     has_legs=FALSE      has_legs=TRUE         tail=FALSE
##         0.22772277         0.77227723         0.25742574
##          tail=TRUE     domestic=FALSE      domestic=TRUE
##         0.74257426         0.87128713         0.12871287
##      catsize=FALSE       catsize=TRUE        type=mammal
##         0.56435644         0.43564356         0.40594059
##          type=bird       type=reptile          type=fish
##         0.19801980         0.04950495         0.12871287
##     type=amphibian        type=insect type=mollusc.et.al
##         0.03960396         0.07920792         0.09900990
```

```
itemFrequencyPlot(items,topN=10)
```




```
############### Alternative encoding ################
# No items for FALSE (use NAs instead)
Zoo2 <- Zoo
Zoo2[Zoo2==FALSE] <- NA
Zoo2[,13] <- Zoo2[,13]>0

## Warning in Ops.factor(Zoo2[, 13], 0): '>' not meaningful for factors

colnames(Zoo2)[13] <- "has_legs" #rename column 13 from legs to has_legs
for(i in 1:ncol(Zoo2)){
  Zoo2[,i] <- as.factor(Zoo2[,i])
}
summary(Zoo2)

##     hair     feathers     eggs       milk     airborne   aquatic
##  FALSE: 0   FALSE: 0   FALSE: 0   FALSE: 0   FALSE: 0   FALSE: 0
##  TRUE :43   TRUE :20   TRUE :59   TRUE :41   TRUE :24   TRUE :36
##  NA's :58   NA's :81   NA's :42   NA's :60   NA's :77   NA's :65
##
##
##
##
##   predator   toothed    backbone   breathes   venomous     fins
##  FALSE: 0   FALSE: 0   FALSE: 0   FALSE: 0   FALSE: 0   FALSE: 0
##  TRUE :56   TRUE :61   TRUE :83   TRUE :80   TRUE : 8   TRUE :17
##  NA's :45   NA's :40   NA's :18   NA's :21   NA's :93   NA's :84
##
```

```
##
##
##
##  has_legs      tail     domestic   catsize             type
##  NA's:101   FALSE: 0   FALSE: 0   FALSE: 0   mammal       :41
##             TRUE :75   TRUE :13   TRUE :44   bird         :20
##             NA's :26   NA's :88   NA's :57   reptile      : 5
##                                               fish         :13
##                                               amphibian    : 4
##                                               insect       : 8
##                                               mollusc.et.al:10
```

```
head(Zoo2)
```

```
##          hair feathers eggs milk airborne aquatic predator toothed
## aardvark TRUE     <NA> <NA> TRUE     <NA>    <NA>     TRUE    TRUE
## antelope TRUE     <NA> <NA> TRUE     <NA>    <NA>     <NA>    TRUE
## bass     <NA>     <NA> TRUE <NA>     <NA>    TRUE     TRUE    TRUE
## bear     TRUE     <NA> <NA> TRUE     <NA>    <NA>     TRUE    TRUE
## boar     TRUE     <NA> <NA> TRUE     <NA>    <NA>     TRUE    TRUE
## buffalo  TRUE     <NA> <NA> TRUE     <NA>    <NA>     <NA>    TRUE
##          backbone breathes venomous fins has_legs tail domestic catsize
## aardvark     TRUE     TRUE     <NA> <NA>     <NA> <NA>     <NA>    TRUE
## antelope     TRUE     TRUE     <NA> <NA>     <NA> TRUE     <NA>    TRUE
## bass         TRUE     <NA>     <NA> TRUE     <NA> TRUE     <NA>    <NA>
## bear         TRUE     TRUE     <NA> <NA>     <NA> <NA>     <NA>    TRUE
## boar         TRUE     TRUE     <NA> <NA>     <NA> TRUE     <NA>    TRUE
## buffalo      TRUE     TRUE     <NA> <NA>     <NA> TRUE     <NA>    TRUE
##            type
## aardvark mammal
## antelope mammal
## bass       fish
## bear     mammal
## boar     mammal
## buffalo  mammal
```

```
# Convert data into a set of items
items2 <- as(Zoo2, "transactions")
items2
```

```
## transactions in sparse format with
##  101 transactions (rows) and
##  37 items (columns)
```

```
# Inspect transactions
colnames(items2)
```

```
##  [1] "hair=FALSE"         "hair=TRUE"          "feathers=FALSE"
##  [4] "feathers=TRUE"      "eggs=FALSE"         "eggs=TRUE"
##  [7] "milk=FALSE"         "milk=TRUE"          "airborne=FALSE"
## [10] "airborne=TRUE"      "aquatic=FALSE"      "aquatic=TRUE"
```

```
## [13] "predator=FALSE"     "predator=TRUE"      "toothed=FALSE"
## [16] "toothed=TRUE"       "backbone=FALSE"     "backbone=TRUE"
## [19] "breathes=FALSE"     "breathes=TRUE"      "venomous=FALSE"
## [22] "venomous=TRUE"      "fins=FALSE"         "fins=TRUE"
## [25] "tail=FALSE"         "tail=TRUE"          "domestic=FALSE"
## [28] "domestic=TRUE"      "catsize=FALSE"      "catsize=TRUE"
## [31] "type=mammal"        "type=bird"          "type=reptile"
## [34] "type=fish"          "type=amphibian"     "type=insect"
## [37] "type=mollusc.et.al"
```

```
inspect(head(items2))
```

```
##     items           transactionID
## [1] {hair=TRUE,
##      milk=TRUE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      catsize=TRUE,
##      type=mammal}        aardvark
## [2] {hair=TRUE,
##      milk=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      tail=TRUE,
##      catsize=TRUE,
##      type=mammal}        antelope
## [3] {eggs=TRUE,
##      aquatic=TRUE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      fins=TRUE,
##      tail=TRUE,
##      type=fish}          bass
## [4] {hair=TRUE,
##      milk=TRUE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      catsize=TRUE,
##      type=mammal}        bear
## [5] {hair=TRUE,
##      milk=TRUE,
##      predator=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
```

```
##      breathes=TRUE,
##      tail=TRUE,
##      catsize=TRUE,
##      type=mammal}        boar
## [6] {hair=TRUE,
##      milk=TRUE,
##      toothed=TRUE,
##      backbone=TRUE,
##      breathes=TRUE,
##      tail=TRUE,
##      catsize=TRUE,
##      type=mammal}        buffalo
```

```
# Look at the support (relative frequency) of items in the data set. Here we
look at the 10 most frequent items
itemFrequency(items2,type="absolute") #support count
```

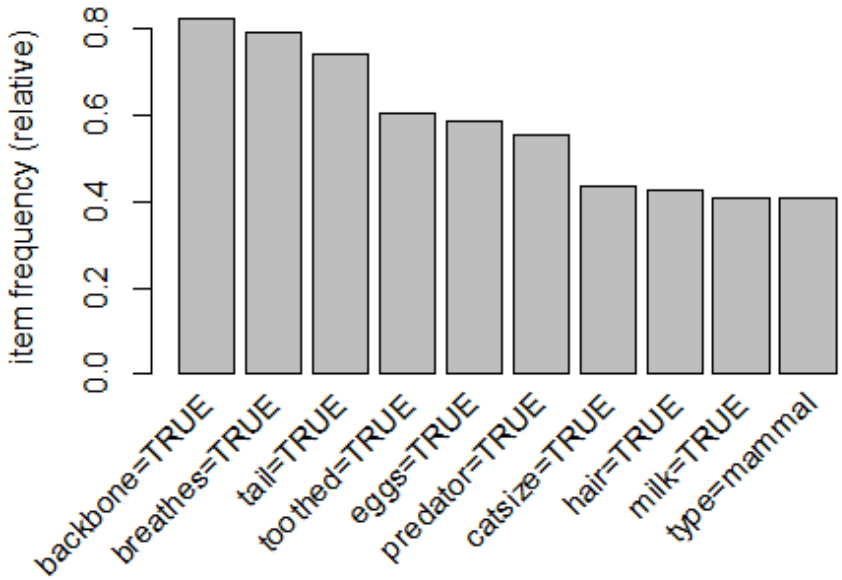
```
##         hair=FALSE          hair=TRUE     feathers=FALSE
##                  0                 43                  0
##      feathers=TRUE         eggs=FALSE          eggs=TRUE
##                 20                  0                 59
##         milk=FALSE          milk=TRUE     airborne=FALSE
##                  0                 41                  0
##      airborne=TRUE      aquatic=FALSE       aquatic=TRUE
##                 24                  0                 36
##     predator=FALSE      predator=TRUE      toothed=FALSE
##                  0                 56                  0
##       toothed=TRUE     backbone=FALSE      backbone=TRUE
##                 61                  0                 83
##     breathes=FALSE      breathes=TRUE     venomous=FALSE
##                  0                 80                  0
##      venomous=TRUE         fins=FALSE          fins=TRUE
##                  8                  0                 17
##         tail=FALSE          tail=TRUE     domestic=FALSE
##                  0                 75                  0
##      domestic=TRUE      catsize=FALSE       catsize=TRUE
##                 13                  0                 44
##        type=mammal          type=bird       type=reptile
##                 41                 20                  5
##          type=fish     type=amphibian        type=insect
##                 13                  4                  8
## type=mollusc.et.al
##                 10
```

```
itemFrequency(items2,type="relative") #support
```

```
##         hair=FALSE          hair=TRUE     feathers=FALSE
##         0.00000000         0.42574257         0.00000000
##      feathers=TRUE         eggs=FALSE          eggs=TRUE
##         0.19801980         0.00000000         0.58415842
##         milk=FALSE          milk=TRUE     airborne=FALSE
```

```
##          0.00000000          0.40594059          0.00000000
##       airborne=TRUE       aquatic=FALSE        aquatic=TRUE
##          0.23762376          0.00000000          0.35643564
##      predator=FALSE       predator=TRUE       toothed=FALSE
##          0.00000000          0.55445545          0.00000000
##        toothed=TRUE      backbone=FALSE       backbone=TRUE
##          0.60396040          0.00000000          0.82178218
##      breathes=FALSE       breathes=TRUE      venomous=FALSE
##          0.00000000          0.79207921          0.00000000
##       venomous=TRUE          fins=FALSE           fins=TRUE
##          0.07920792          0.00000000          0.16831683
##          tail=FALSE           tail=TRUE      domestic=FALSE
##          0.00000000          0.74257426          0.00000000
##       domestic=TRUE       catsize=FALSE        catsize=TRUE
##          0.12871287          0.00000000          0.43564356
##         type=mammal          type=bird        type=reptile
##          0.40594059          0.19801980          0.04950495
##           type=fish      type=amphibian         type=insect
##          0.12871287          0.03960396          0.07920792
## type=mollusc.et.al
##          0.09900990
```

```r
itemFrequencyPlot(items2,topN=10)
```



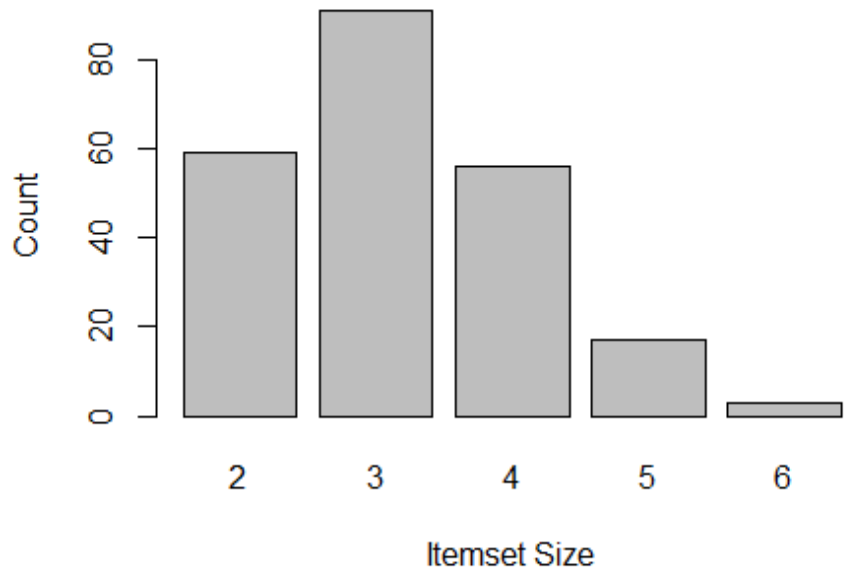

```r
# Select transactions that contain a certain item
items_insects <- items2[items2 %in% "type=insect"]
inspect(items_insects)
```

```
##     items           transactionID
## [1] {eggs=TRUE,
##      breathes=TRUE,
##      type=insect}        flea
## [2] {eggs=TRUE,
##      airborne=TRUE,
##      breathes=TRUE,
##      type=insect}        gnat
## [3] {hair=TRUE,
##      eggs=TRUE,
##      airborne=TRUE,
##      breathes=TRUE,
##      venomous=TRUE,
##      domestic=TRUE,
##      type=insect}        honeybee
## [4] {hair=TRUE,
##      eggs=TRUE,
##      airborne=TRUE,
##      breathes=TRUE,
##      type=insect}        housefly
## [5] {eggs=TRUE,
##      airborne=TRUE,
##      predator=TRUE,
##      breathes=TRUE,
##      type=insect}        ladybird
## [6] {hair=TRUE,
##      eggs=TRUE,
##      airborne=TRUE,
##      breathes=TRUE,
##      type=insect}        moth
## [7] {eggs=TRUE,
##      breathes=TRUE,
##      type=insect}        termite
## [8] {hair=TRUE,
##      eggs=TRUE,
##      airborne=TRUE,
##      breathes=TRUE,
##      venomous=TRUE,
##      type=insect}        wasp
```

```
############## Find frequent itemsets ###############
# Default of apriori function -> support = 0.1, minimum length = 1, maximum
length = 10
freq_itemset <- apriori(items, parameter=list(target="frequent"))

## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE            TRUE       5     0.1      1
```

```
##  maxlen            target   ext
##      10 frequent itemsets FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 10
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[39 item(s), 101 transaction(s)] done [0.00s].
## sorting and recoding items ... [34 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 6 7 8 9 10

## Warning in apriori(items, parameter = list(target = "frequent")): Mining
## stopped (maxlen reached). Only patterns up to a length of 10 returned!

##  done [0.08s].
## writing ... [238445 set(s)] done [0.05s].
## creating S4 object  ... done [0.11s].

freq_itemset

## set of 238445 itemsets

freq_itemset <- apriori(items,
parameter=list(support=0.5,minlen=2,maxlen=8,target="frequent"))

## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE            TRUE       5     0.5      2
##  maxlen            target   ext
##       8 frequent itemsets FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 50
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[39 item(s), 101 transaction(s)] done [0.00s].
## sorting and recoding items ... [16 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 6 done [0.02s].
## writing ... [226 set(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
freq_itemset

## set of 226 itemsets

# Sort by support
freq_itemset <- sort(freq_itemset, by="support")
inspect(head(freq_itemset, n=10))

##      items                              support   count
## [1]  {venomous=FALSE,domestic=FALSE} 0.8019802 81
## [2]  {backbone=TRUE,venomous=FALSE}  0.7821782 79
## [3]  {fins=FALSE,has_legs=TRUE}      0.7623762 77
## [4]  {venomous=FALSE,fins=FALSE}     0.7623762 77
## [5]  {breathes=TRUE,fins=FALSE}      0.7524752 76
## [6]  {breathes=TRUE,venomous=FALSE}  0.7425743 75
## [7]  {backbone=TRUE,tail=TRUE}       0.7326733 74
## [8]  {venomous=FALSE,has_legs=TRUE}  0.7326733 74
## [9]  {feathers=FALSE,airborne=FALSE} 0.7227723 73
## [10] {breathes=TRUE,has_legs=TRUE}   0.7227723 73

# Look at frequent itemsets with many items
barplot(table(size(freq_itemset)), xlab="Itemset Size", ylab="Count")
```




```
# Look at itemsets with size > 4 {5,6,7,8}
inspect(freq_itemset[size(freq_itemset)>4])

##      items                support count
## [1]  {backbone=TRUE,
```

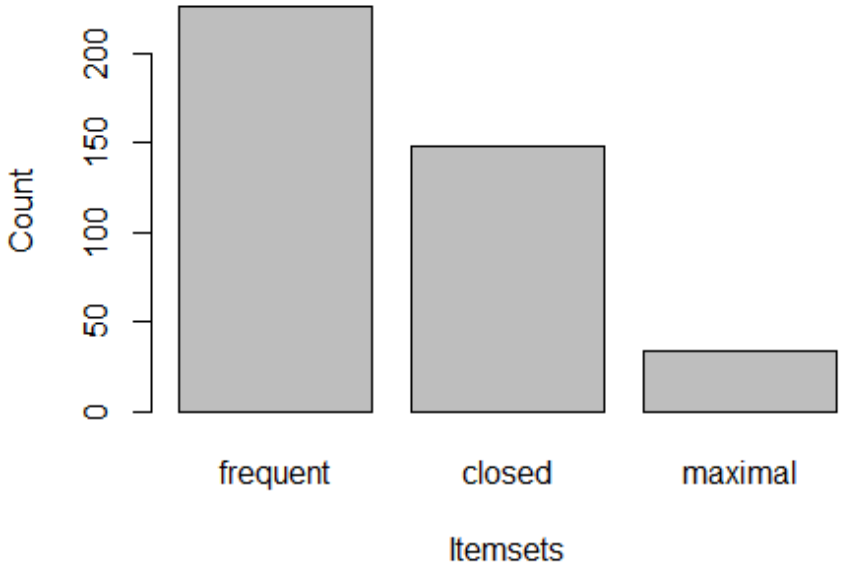
```
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE}  0.6138614    62
## [2]  {aquatic=FALSE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE}  0.5643564    57
## [3]  {breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE,
##       domestic=FALSE} 0.5643564    57
## [4]  {backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       has_legs=TRUE,
##       tail=TRUE}      0.5544554    56
## [5]  {backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       tail=TRUE}      0.5544554    56
## [6]  {feathers=FALSE,
##       airborne=FALSE,
##       toothed=TRUE,
##       backbone=TRUE,
##       venomous=FALSE} 0.5445545    55
## [7]  {backbone=TRUE,
##       breathes=TRUE,
##       fins=FALSE,
##       has_legs=TRUE,
##       tail=TRUE}      0.5445545    55
## [8]  {breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE,
##       tail=TRUE}      0.5445545    55
## [9]  {backbone=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE,
##       tail=TRUE}      0.5445545    55
## [10] {backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE,
##       tail=TRUE}      0.5445545    55
```

```
## [11] {aquatic=FALSE,
##       backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE}     0.5148515    52
## [12] {backbone=TRUE,
##       breathes=TRUE,
##       fins=FALSE,
##       has_legs=TRUE,
##       domestic=FALSE} 0.5148515    52
## [13] {backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       has_legs=TRUE,
##       domestic=FALSE} 0.5148515    52
## [14] {backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       domestic=FALSE} 0.5148515    52
## [15] {aquatic=FALSE,
##       backbone=TRUE,
##       breathes=TRUE,
##       fins=FALSE,
##       has_legs=TRUE}  0.5049505    51
## [16] {aquatic=FALSE,
##       backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       has_legs=TRUE}  0.5049505    51
## [17] {aquatic=FALSE,
##       backbone=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE}  0.5049505    51
## [18] {backbone=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE,
##       domestic=FALSE} 0.5049505    51
## [19] {aquatic=FALSE,
##       backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
##       has_legs=TRUE}  0.5049505    51
## [20] {backbone=TRUE,
##       breathes=TRUE,
##       venomous=FALSE,
##       fins=FALSE,
```

```
##       has_legs=TRUE,
##       domestic=FALSE} 0.5049505    51

####### Concise representation of itemsets ########
# 1. Maximal frequent itemsets
max_freq_itemset <- freq_itemset[is.maximal(freq_itemset)]
max_freq_itemset #instead of storing 226 itemsets, we can store only 34
itemsets

## set of 34 itemsets

inspect(head(sort(max_freq_itemset, by="support")))

##     items              support count
## [1] {backbone=TRUE,
##      venomous=FALSE,
##      tail=TRUE,
##      domestic=FALSE} 0.6039604    61
## [2] {feathers=FALSE,
##      airborne=FALSE,
##      venomous=FALSE,
##      domestic=FALSE} 0.5742574    58
## [3] {feathers=FALSE,
##      fins=FALSE,
##      domestic=FALSE} 0.5445545    55
## [4] {feathers=FALSE,
##      airborne=FALSE,
##      toothed=TRUE,
##      backbone=TRUE,
##      venomous=FALSE} 0.5445545    55
## [5] {backbone=TRUE,
##      breathes=TRUE,
##      venomous=FALSE,
##      fins=FALSE,
##      has_legs=TRUE,
##      tail=TRUE}      0.5445545    55
## [6] {predator=TRUE,
##      domestic=FALSE} 0.5346535    54

# 2. Closed frequent itemsets
closed_freq_itemset <- freq_itemset[is.closed(freq_itemset)]
closed_freq_itemset #instead of storing 226 itemsets, we can store only 148
itemsets

## set of 148 itemsets

inspect(head(sort(closed_freq_itemset, by="support")))

##     items                           support   count
## [1] {venomous=FALSE,domestic=FALSE} 0.8019802 81
## [2] {backbone=TRUE,venomous=FALSE}  0.7821782 79
## [3] {fins=FALSE,has_legs=TRUE}      0.7623762 77
```

```
## [4] {venomous=FALSE,fins=FALSE}     0.7623762 77
## [5] {breathes=TRUE,fins=FALSE}      0.7524752 76
## [6] {breathes=TRUE,venomous=FALSE}  0.7425743 75
```

```
# Maximal frequent itemsets is a subset of Closed frequent itemsets
# Closed frequent itemsets is a subset of Frequent itemsets
barplot(c(frequent=length(freq_itemset),closed=length(closed_freq_itemset),ma
ximal=length(max_freq_itemset)),
  ylab="Count", xlab="Itemsets")
```




```
############## Mine association rules ###############
rules <- apriori(items) #This extracts lots of rules; default minimum length,
support and confidence = 1, 0.1 and 0.8

## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.8    0.1    1 none FALSE            TRUE       5     0.1      1
##  maxlen target   ext
##      10  rules FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 10
```

```
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[39 item(s), 101 transaction(s)] done [0.00s].
## sorting and recoding items ... [34 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 6 7 8 9 10

## Warning in apriori(items): Mining stopped (maxlen reached). Only patterns
## up to a length of 10 returned!

##  done [0.09s].
## writing ... [1517191 rule(s)] done [0.34s].
## creating S4 object  ... done [0.92s].

rules

## set of 1517191 rules

print(object.size(rules), unit="Mb")

## 104.4 Mb

inspect(head(rules))

##     lhs                rhs              support   confidence lift
## [1] {}              => {feathers=FALSE} 0.8019802 0.8019802  1.000000
## [2] {}              => {backbone=TRUE}  0.8217822 0.8217822  1.000000
## [3] {}              => {fins=FALSE}     0.8316832 0.8316832  1.000000
## [4] {}              => {domestic=FALSE} 0.8712871 0.8712871  1.000000
## [5] {}              => {venomous=FALSE} 0.9207921 0.9207921  1.000000
## [6] {domestic=TRUE} => {predator=FALSE} 0.1089109 0.8461538  1.899145
##     count
## [1] 81
## [2] 83
## [3] 84
## [4] 88
## [5] 93
## [6] 11

# Create less rules (increase minimum length, minimum support, and minimum
confidence)
rules <- apriori(items, parameter=list(support=0.5, confidence=0.9))

## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.9    0.1    1 none FALSE            TRUE       5     0.5      1
##  maxlen target   ext
##      10  rules FALSE
##
## Algorithmic control:
```

```
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 50
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[39 item(s), 101 transaction(s)] done [0.00s].
## sorting and recoding items ... [16 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 6 done [0.00s].
## writing ... [381 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].

rules

## set of 381 rules

inspect(head(rules))

##     lhs                rhs              support   confidence lift
## [1] {}              => {venomous=FALSE} 0.9207921 0.9207921  1.000000
## [2] {predator=TRUE} => {domestic=FALSE} 0.5346535 0.9642857  1.106737
## [3] {hair=FALSE}    => {eggs=TRUE}      0.5346535 0.9310345  1.593805
## [4] {eggs=TRUE}     => {hair=FALSE}     0.5346535 0.9152542  1.593805
## [5] {hair=FALSE}    => {milk=FALSE}     0.5544554 0.9655172  1.625287
## [6] {milk=FALSE}    => {hair=FALSE}     0.5544554 0.9333333  1.625287
##     count
## [1] 93
## [2] 54
## [3] 54
## [4] 54
## [5] 56
## [6] 56

# Minimum length of itemsets must be greater than or equal to 2
rules <- apriori(items, parameter=list(minlen=2, support=0.5,
confidence=0.9))

## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.9    0.1    1 none FALSE            TRUE       5     0.5      2
##  maxlen target   ext
##      10  rules FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 50
```

```
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[39 item(s), 101 transaction(s)] done [0.00s].
## sorting and recoding items ... [16 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 6 done [0.00s].
## writing ... [380 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].

rules

## set of 380 rules

inspect(head(rules)) #default: highest support

##     lhs                 rhs               support   confidence lift
## [1] {predator=TRUE} => {domestic=FALSE} 0.5346535 0.9642857  1.106737
## [2] {hair=FALSE}    => {eggs=TRUE}      0.5346535 0.9310345  1.593805
## [3] {eggs=TRUE}     => {hair=FALSE}     0.5346535 0.9152542  1.593805
## [4] {hair=FALSE}    => {milk=FALSE}     0.5544554 0.9655172  1.625287
## [5] {milk=FALSE}    => {hair=FALSE}     0.5544554 0.9333333  1.625287
## [6] {hair=FALSE}    => {domestic=FALSE} 0.5346535 0.9310345  1.068574
##     count
## [1] 54
## [2] 54
## [3] 54
## [4] 56
## [5] 56
## [6] 54

summary(rules)

## set of 380 rules
##
## rule length distribution (lhs + rhs):sizes
##   2   3   4   5   6
##  35 125 145  62  13
##
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.000   3.000   4.000   3.718   4.000   6.000
##
## summary of quality measures:
##     support         confidence          lift            count
##  Min.   :0.5050   Min.   :0.9012   Min.   :0.9788   Min.   :51.00
##  1st Qu.:0.5248   1st Qu.:0.9375   1st Qu.:1.0860   1st Qu.:53.00
##  Median :0.5545   Median :0.9672   Median :1.1954   Median :56.00
##  Mean   :0.5724   Mean   :0.9642   Mean   :1.1905   Mean   :57.81
##  3rd Qu.:0.6040   3rd Qu.:1.0000   3rd Qu.:1.2469   3rd Qu.:61.00
##  Max.   :0.8020   Max.   :1.0000   Max.   :1.7119   Max.   :81.00
##
## mining info:
```

```
##   data ntransactions support confidence
##  items           101     0.5        0.9

quality(head(rules))

##     support confidence     lift count
## 1 0.5346535  0.9642857 1.106737    54
## 2 0.5346535  0.9310345 1.593805    54
## 3 0.5346535  0.9152542 1.593805    54
## 4 0.5544554  0.9655172 1.625287    56
## 5 0.5544554  0.9333333 1.625287    56
## 6 0.5346535  0.9310345 1.068574    54

# Look at rules with highest lift
rules <- sort(rules, by="lift")
inspect(head(rules, n=10))

##      lhs                              rhs          support   confidence
## [1]  {milk=FALSE,venomous=FALSE} => {eggs=TRUE}  0.5148515 1.0000000
## [2]  {hair=FALSE,eggs=TRUE}      => {milk=FALSE} 0.5346535 1.0000000
## [3]  {eggs=TRUE}                 => {milk=FALSE} 0.5742574 0.9830508
## [4]  {milk=FALSE}                => {eggs=TRUE}  0.5742574 0.9666667
## [5]  {eggs=TRUE,domestic=FALSE}  => {milk=FALSE} 0.5247525 0.9814815
## [6]  {eggs=TRUE,venomous=FALSE}  => {milk=FALSE} 0.5148515 0.9811321
## [7]  {hair=FALSE,milk=FALSE}     => {eggs=TRUE}  0.5346535 0.9642857
## [8]  {milk=FALSE,domestic=FALSE} => {eggs=TRUE}  0.5247525 0.9636364
## [9]  {milk=FALSE,domestic=FALSE} => {hair=FALSE} 0.5148515 0.9454545
## [10] {hair=FALSE}                => {milk=FALSE} 0.5544554 0.9655172
##      lift     count
## [1]  1.711864 52
## [2]  1.683333 54
## [3]  1.654802 58
## [4]  1.654802 58
## [5]  1.652160 53
## [6]  1.651572 52
## [7]  1.650726 54
## [8]  1.649615 53
## [9]  1.646395 52
## [10] 1.625287 56

# Create rules using the alternative encoding (No FALSE item)
rules2 <- apriori(items2, parameter=list(minlen=2, support=0.5,
confidence=0.9))

## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.9    0.1    1 none FALSE            TRUE       5     0.5      2
##  maxlen target   ext
##      10  rules FALSE
```

```
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 50
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[22 item(s), 101 transaction(s)] done [0.00s].
## sorting and recoding items ... [6 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 done [0.00s].
## writing ... [4 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].

rules2

## set of 4 rules

inspect(sort(rules2, by="lift"))

##     lhs                           rhs             support   confidence
## [1] {toothed=TRUE}             => {backbone=TRUE} 0.6039604 1.0000000
## [2] {toothed=TRUE,tail=TRUE}   => {backbone=TRUE} 0.5148515 1.0000000
## [3] {tail=TRUE}                => {backbone=TRUE} 0.7326733 0.9866667
## [4] {breathes=TRUE,tail=TRUE}  => {backbone=TRUE} 0.5940594 0.9836066
##     lift     count
## [1] 1.216867 61
## [2] 1.216867 52
## [3] 1.200643 74
## [4] 1.196919 60

summary(rules2)

## set of 4 rules
##
## rule length distribution (lhs + rhs):sizes
## 2 3
## 2 2
##
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##     2.0     2.0     2.5     2.5     3.0     3.0
##
## summary of quality measures:
##     support         confidence          lift           count
##  Min.   :0.5149   Min.   :0.9836   Min.   :1.197   Min.   :52.00
##  1st Qu.:0.5743   1st Qu.:0.9859   1st Qu.:1.200   1st Qu.:58.00
##  Median :0.5990   Median :0.9933   Median :1.209   Median :60.50
##  Mean   :0.6114   Mean   :0.9926   Mean   :1.208   Mean   :61.75
##  3rd Qu.:0.6361   3rd Qu.:1.0000   3rd Qu.:1.217   3rd Qu.:64.25
##  Max.   :0.7327   Max.   :1.0000   Max.   :1.217   Max.   :74.00
```

```
##
## mining info:
##    data ntransactions support confidence
##  items2           101     0.5        0.9
```

```
quality(rules2)
```

```
##     support confidence     lift count
## 1 0.6039604  1.0000000 1.216867    61
## 2 0.7326733  0.9866667 1.200643    74
## 3 0.5148515  1.0000000 1.216867    52
## 4 0.5940594  0.9836066 1.196919    60
```

```
##### Calculate additional interestingness measures #####
interestMeasure(rules2, c("phi","gini","kappa"), transactions = items2)
```

```
##         phi       gini     kappa
## 1 0.5750851 0.09687286 0.4970575
## 2 0.7317620 0.15684747 0.7120867
## 3 0.4797344 0.06741236 0.3741755
## 4 0.5221857 0.07987075 0.4513354
```

```
# Add measures to the rules
quality(rules2) <- cbind(quality(rules2), interestMeasure(rules2,
c("phi","gini","kappa"), transactions = items2))
inspect(sort(rules2, by="gini"))
```

```
##     lhs                           rhs             support   confidence
## [1] {tail=TRUE}                => {backbone=TRUE} 0.7326733 0.9866667
## [2] {toothed=TRUE}             => {backbone=TRUE} 0.6039604 1.0000000
## [3] {breathes=TRUE,tail=TRUE} => {backbone=TRUE} 0.5940594 0.9836066
## [4] {toothed=TRUE,tail=TRUE}  => {backbone=TRUE} 0.5148515 1.0000000
##     lift     count phi       gini       kappa
## [1] 1.200643 74    0.7317620 0.15684747 0.7120867
## [2] 1.216867 61    0.5750851 0.09687286 0.4970575
## [3] 1.196919 60    0.5221857 0.07987075 0.4513354
## [4] 1.216867 52    0.4797344 0.06741236 0.3741755
```

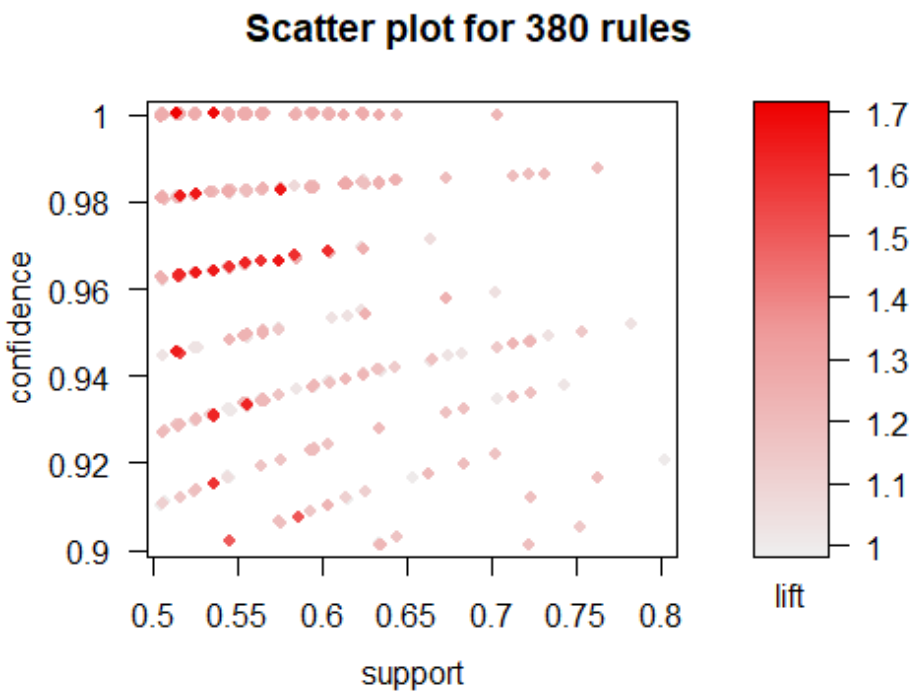
```
################### Export rules ###################
# Exporting rules as a CSV-file to be opened in MS Excel or other tools
setwd("C:/Users/Wow/Desktop") #Set file path to "C:\Users\Wow\Desktop"
write(rules, file="rules.csv", quote=TRUE)
write(rules2, file="rules2.csv", quote=TRUE)
```

```
######## Association rule visualization ############
library(arulesViz)
```

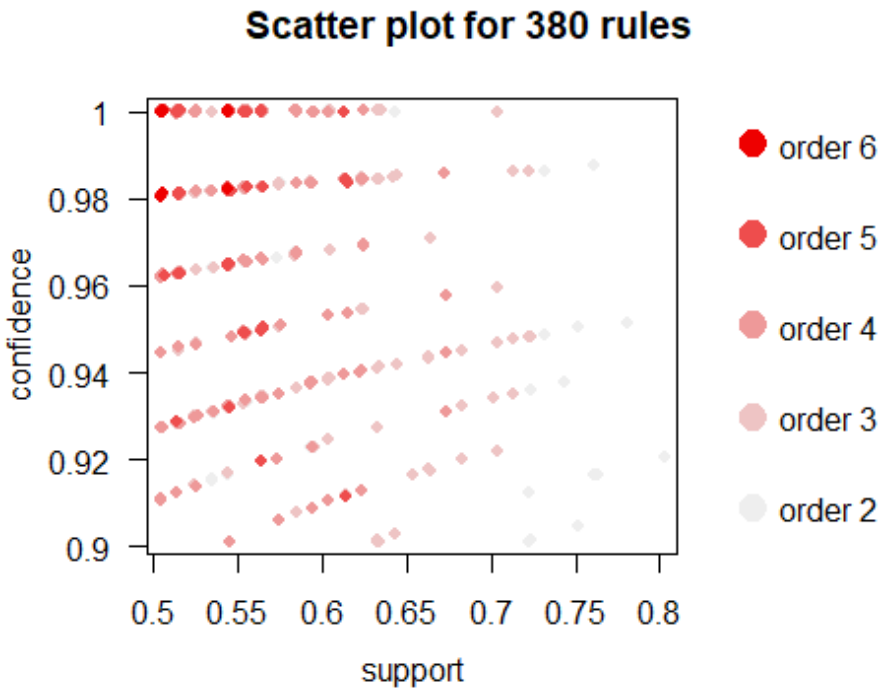
```
## Loading required package: grid
```

```
# Scatter Plot
plot(rules)
```

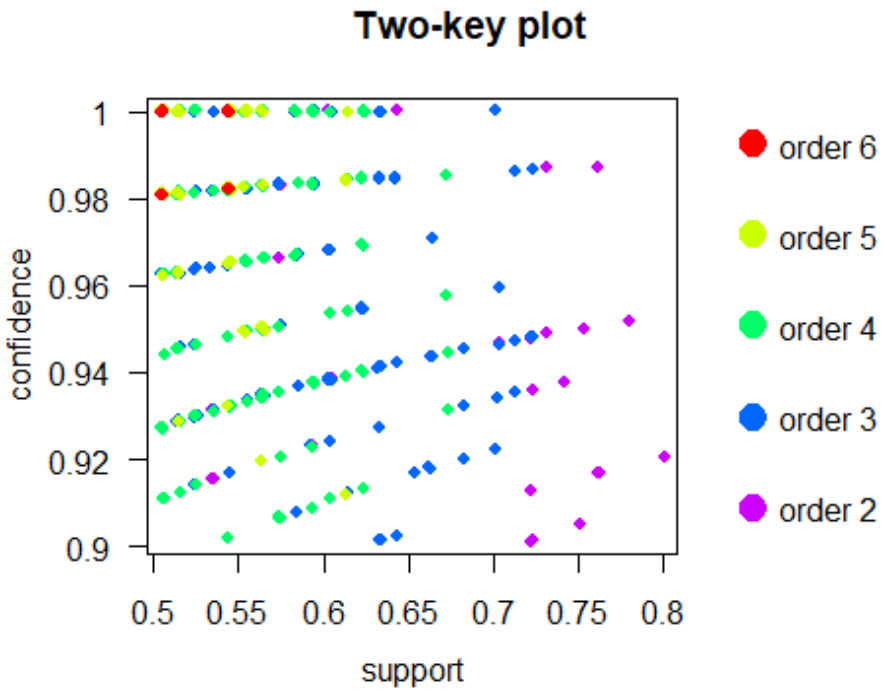
```
## To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
```




```
plot(rules, shading="order") #"order" means number of items

## To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
```




```
# Interactive Plot
#plot(rules, interactive=TRUE) #show interactive plot

# # Two-key plot: It is a scatterplot with shading = "order"
plot(rules, method = "two-key plot")

## To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
```




```
# Grouped matrix plot:
# Antecedents (columns) in the matrix are grouped using clustering technique.
# Groups are represented by the most interesting item (highest ratio of
support in the group to support in all rules) in the group.
# Balloons' location in the matrix indicates what consequent the antecedent
are connected.
plot(rules, method="grouped matrix")
```

Grouped Matrix for 380 Rules
ns in LHS Group
s: {milk=FALSE, hair=FALSE, +2 items}
s: {eggs=TRUE, milk=FALSE}
s: {hair=FALSE, eggs=TRUE, +2 items}
s: {feathers=FALSE, airborne=FALSE, +4 items}
s: {toothed=TRUE, venomous=FALSE, +1 items}
s: {tail=TRUE, fins=FALSE, +1 items}
s: {aquatic=FALSE, venomous=FALSE, +1 items}
s: {tail=TRUE, has_legs=TRUE, +2 items}
s: {toothed=TRUE, feathers=FALSE, +1 items}
s: {fins=FALSE, aquatic=FALSE, +4 items}
s: {fins=FALSE, breathes=TRUE, +7 items}
s: {has_legs=TRUE, aquatic=FALSE, +5 items}
s: {breathes=TRUE, aquatic=FALSE, +4 items}
s: {fins=FALSE, has_legs=TRUE, +6 items}
s: {tail=TRUE, feathers=FALSE, +7 items}
s: {airborne=FALSE, toothed=TRUE, +4 items}
s: {predator=TRUE, domestic=FALSE, +10 items}
s: {breathes=TRUE, has_legs=TRUE, +7 items}
s: {feathers=FALSE, toothed=TRUE, +3 items}
s: {airborne=FALSE, toothed=TRUE, +5 items}
+ 3 supressed
backbone=TRUE}
feathers=FALSE}
tail=TRUE}
Size: support
Color: lift