Associative Rules

Durga Gaddam

October 21, 2016

Objective

The objective of this article is to explain how Apriori Algorithm is used in providin recommendations in market basket analysis and, customer purchasing behaviour.

Support(X)= Count(X)/N

confidence(X->Y)= Support(X,Y)/Support(X)

```r
###install.packages("arules")
###install.packages("arulesViz")
###install.packages("fpc")

###library(arules)
###library(arulesViz)
require(arules)
```

```
## Loading required package: arules
```

```
## Warning: package 'arules' was built under R version 3.3.1
```

```
## Loading required package: Matrix
```

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

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

```r
require(arulesViz)
```

```
## Loading required package: arulesViz
```

```
## Warning: package 'arulesViz' was built under R version 3.3.1
```

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

```r
data("Groceries")

summary(Groceries)
```

```
## transactions as itemMatrix in sparse format with
##  9835 rows (elements/itemsets/transactions) and
##  169 columns (items) and a density of 0.02609146
##
## most frequent items:
##       whole milk other vegetables       rolls/buns             soda
##             2513             1903             1809             1715
##           yogurt          (Other)
##             1372            34055
##
## element (itemset/transaction) length distribution:
## sizes
##    1    2    3    4    5    6    7    8    9   10   11   12   13   14   15
## 2159 1643 1299 1005  855  645  545  438  350  246  182  117   78   77   55
##   16   17   18   19   20   21   22   23   24   26   27   28   29   32
##   46   29   14   14    9   11    4    6    1    1    1    1    3    1
##
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   2.000   3.000   4.409   6.000  32.000
##
## includes extended item information - examples:
##        labels  level2           level1
## 1 frankfurter sausage meat and sausage
## 2     sausage sausage meat and sausage
## 3  liver loaf sausage meat and sausage
```

```
class(Groceries)
```

```
## [1] "transactions"
## attr(,"package")
## [1] "arules"
```

```
apply(Groceries@data[,10:20],2, function(r) paste(Groceries@itemInfo[r,"labels"], collapse=", "))
```

```
##  [1] "whole milk, cereals"
##  [2] "tropical fruit, other vegetables, white bread, bottled water, chocolate"
##  [3] "citrus fruit, tropical fruit, whole milk, butter, curd, yogurt, flour, bottled water, dishes"
##  [4] "beef"
##  [5] "frankfurter, rolls/buns, soda"
##  [6] "chicken, tropical fruit"
##  [7] "butter, sugar, fruit/vegetable juice, newspapers"
##  [8] "fruit/vegetable juice"
##  [9] "packaged fruit/vegetables"
## [10] "chocolate"
## [11] "specialty bar"
```

```
itemsets <- apriori(Groceries, parameter = list(minlen=1, maxlen=1, support=0.02, target="frequent item
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE            TRUE       5    0.02      1
##  maxlen            target   ext
##       1 frequent itemsets FALSE
## 
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
## 
## Absolute minimum support count: 196 
## 
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[169 item(s), 9835 transaction(s)] done [0.01s].
## sorting and recoding items ... [59 item(s)] done [0.00s].
## creating transaction tree ... done [0.01s].
## checking subsets of size 1
```

```
## Warning in apriori(Groceries, parameter = list(minlen = 1, maxlen = 1,
## support = 0.02, : Mining stopped (maxlen reached). Only patterns up to a
## length of 1 returned!
```

```
##  done [0.00s].
## writing ... [59 set(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
summary(itemsets)
```

```
## set of 59 itemsets
## 
## most frequent items:
## frankfurter     sausage         ham        meat     chicken     (Other) 
##           1           1           1           1           1          54 
## 
## element (itemset/transaction) length distribution:sizes
##  1 
## 59 
## 
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max. 
##       1       1       1       1       1       1 
## 
## summary of quality measures:
##     support       
##  Min.   :0.02105  
##  1st Qu.:0.03015  
##  Median :0.04809  
##  Mean   :0.06200  
##  3rd Qu.:0.07666  
##  Max.   :0.25552  
## 
## includes transaction ID lists: FALSE 
## 
## mining info:
##       data ntransactions support confidence
##  Groceries          9835    0.02          1
```

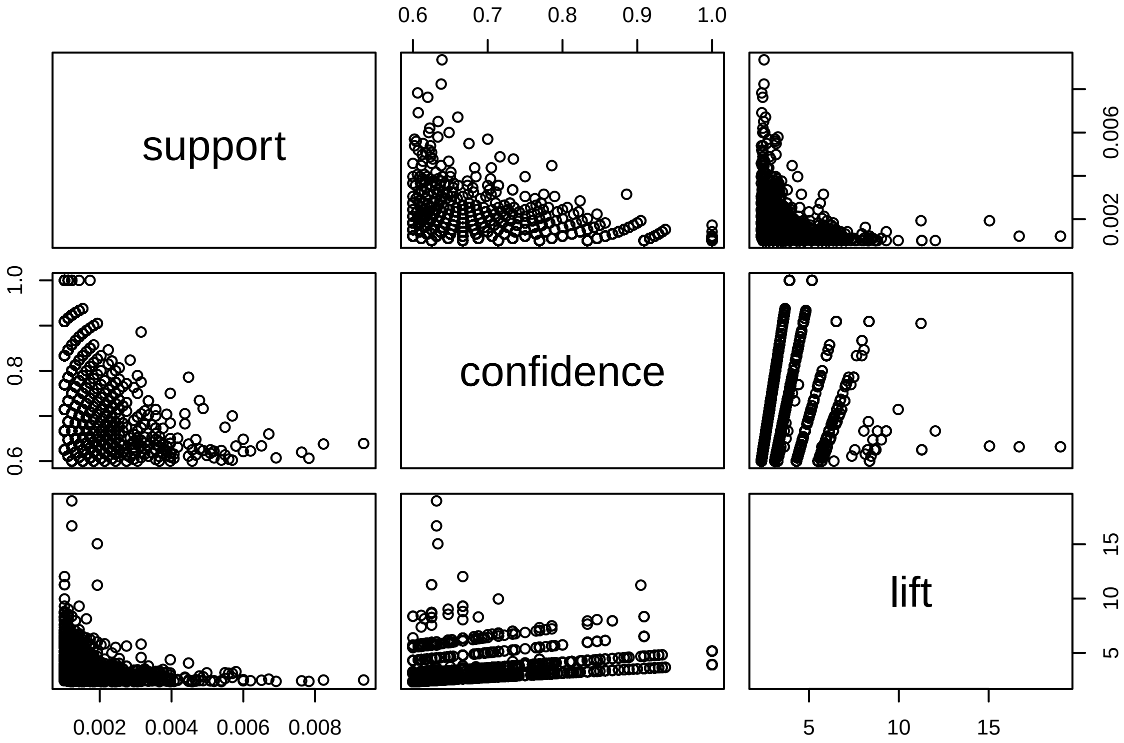
```
inspect(head(sort(itemsets, by = "support"),10))
```

```
##      items               support
## [1]  {whole milk}        0.25551601
## [2]  {other vegetables} 0.19349263
## [3]  {rolls/buns}        0.18393493
## [4]  {soda}              0.17437722
## [5]  {yogurt}            0.13950178
## [6]  {bottled water}     0.11052364
## [7]  {root vegetables}  0.10899847
## [8]  {tropical fruit}   0.10493137
## [9]  {shopping bags}     0.09852567
## [10] {sausage}           0.09395018
```

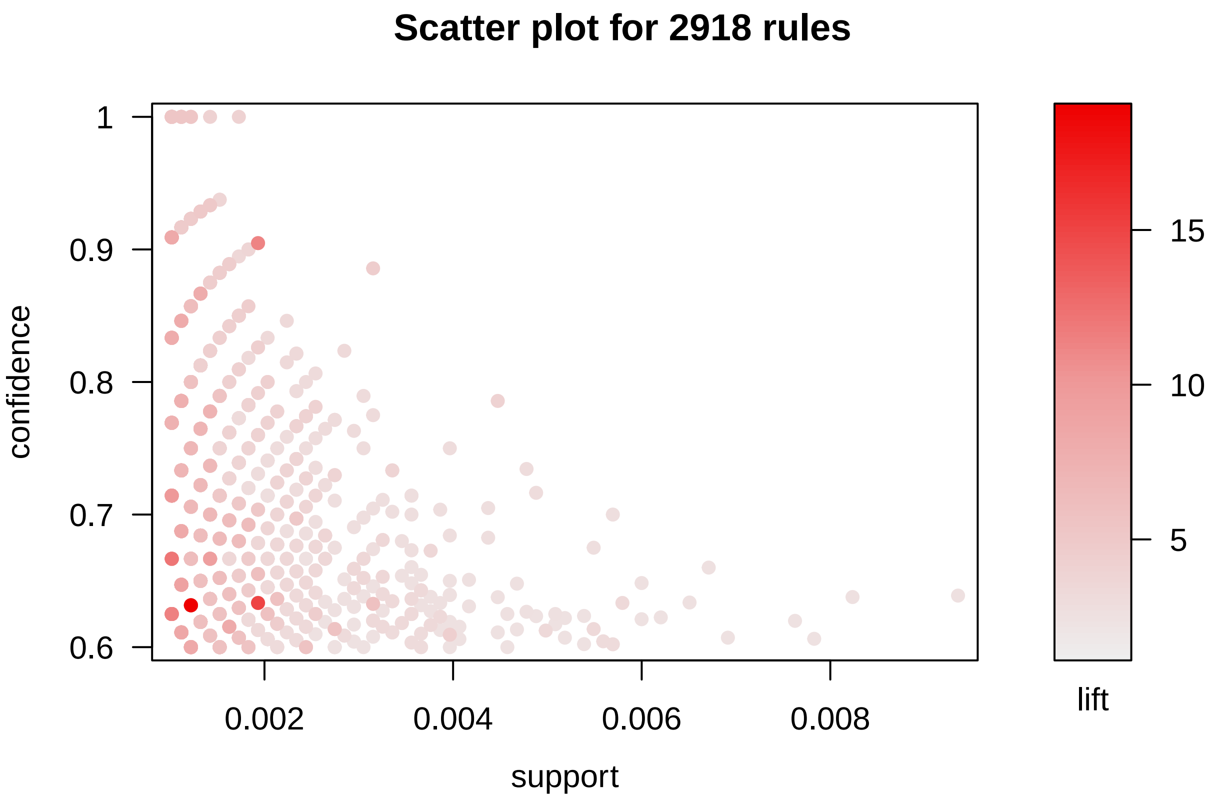
```
rules <- apriori(Groceries, parameter = list(support=0.001, confidence=0.6, target="rules"))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.6    0.1    1 none FALSE            TRUE       5   0.001      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: 9
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[169 item(s), 9835 transaction(s)] done [0.01s].
## sorting and recoding items ... [157 item(s)] done [0.00s].
## creating transaction tree ... done [0.01s].
## checking subsets of size 1 2 3 4 5 6 done [0.04s].
## writing ... [2918 rule(s)] done [0.00s].
## creating S4 object  ... done [0.01s].
```

```
plot(rules@quality)
```


support
confidence
lift
0.6
0.7
0.8
0.9
1.0
0.002
0.006
1.0
0.8
0.6
5
10
15
0.002
0.004
0.006
0.008
5
10
15

```
plot(rules)
```

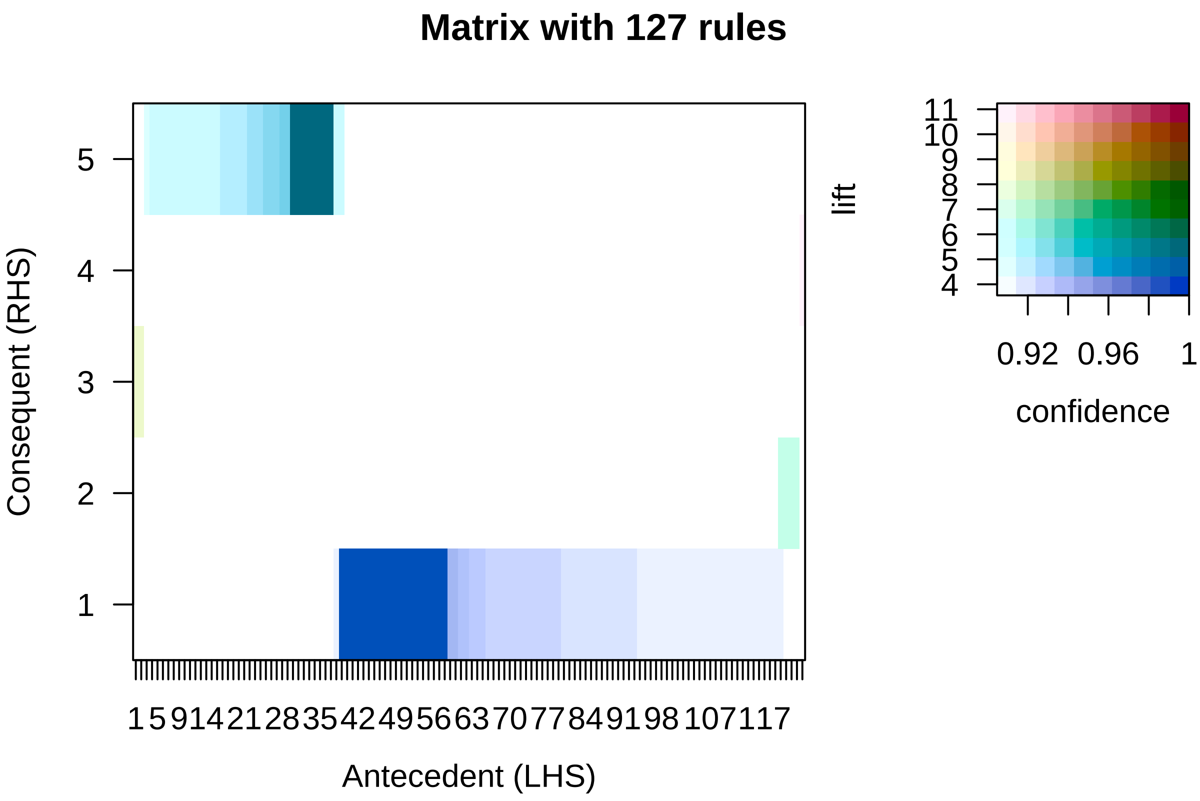


```
confidentRules <- rules[quality(rules)$confidence>0.9]
confidentRules
```

```
## set of 127 rules
```

```
plot(confidentRules, method="matrix", measure=c("lift", "confidence"), control= list(reorder=TRUE))
```

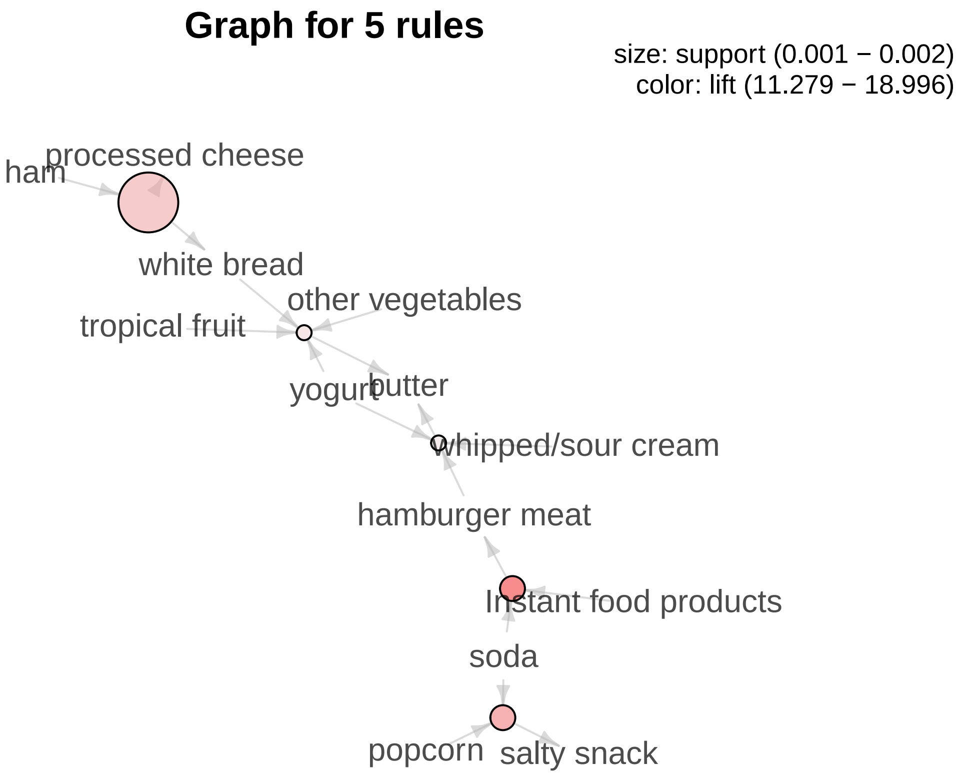
```
## Itemsets in Antecedent (LHS)
##   [1] "{tropical fruit,other vegetables,whole milk,yogurt,oil}"
##   [2] "{citrus fruit,other vegetables,soda,fruit/vegetable juice}"
##   [3] "{tropical fruit,whipped/sour cream,fruit/vegetable juice}"
##   [4] "{whole milk,butter,whipped/sour cream,soda}"
##   [5] "{tropical fruit,root vegetables,whole milk,yogurt,oil}"
##   [6] "{tropical fruit,whipped/sour cream,hard cheese}"
##   [7] "{yogurt,oil,coffee}"
##   [8] "{hamburger meat,tropical fruit,whipped/sour cream}"
##   [9] "{herbs,whole milk,fruit/vegetable juice}"
##  [10] "{root vegetables,whole milk,butter,white bread}"
##  [11] "{tropical fruit,whole milk,whipped/sour cream,fruit/vegetable juice}"
##  [12] "{tropical fruit,butter,yogurt,white bread}"
##  [13] "{pork,whole milk,butter milk}"
##  [14] "{root vegetables,onions,napkins}"
##  [15] "{pip fruit,butter milk,fruit/vegetable juice}"
##  [16] "{citrus fruit,root vegetables,whole milk,yogurt,whipped/sour cream}"
##  [17] "{citrus fruit,whole milk,whipped/sour cream,cream cheese }"
## [18] "{hard cheese,oil}"
## [19] "{tropical fruit,dessert,whipped/sour cream}"
## [20] "{grapes,onions}"
## [21] "{tropical fruit,yogurt,whipped/sour cream,fruit/vegetable juice}"
## [22] "{citrus fruit,whole milk,whipped/sour cream,domestic eggs}"
## [23] "{tropical fruit,whipped/sour cream,soft cheese}"
## [24] "{citrus fruit,root vegetables,cream cheese }"
## [25] "{pip fruit,butter,pastry}"
## [26] "{butter,whipped/sour cream,soda}"
## [27] "{root vegetables,whole milk,yogurt,rice}"
## [28] "{root vegetables,whole milk,yogurt,oil}"
## [29] "{citrus fruit,tropical fruit,root vegetables,whole milk,yogurt}"
## [30] "{citrus fruit,root vegetables,soft cheese}"
## [31] "{pip fruit,whipped/sour cream,brown bread}"
## [32] "{ham,tropical fruit,pip fruit,whole milk}"
## [33] "{tropical fruit,grapes,whole milk,yogurt}"
## [34] "{whole milk,rolls/buns,soda,newspapers}"
## [35] "{citrus fruit,tropical fruit,root vegetables,whipped/sour cream}"
## [36] "{tropical fruit,butter,whipped/sour cream,fruit/vegetable juice}"
## [37] "{ham,tropical fruit,pip fruit,yogurt}"
## [38] "{frankfurter,tropical fruit,frozen meals}"
## [39] "{tropical fruit,root vegetables,yogurt,oil}"
## [40] "{butter,soft cheese,domestic eggs}"
## [41] "{root vegetables,other vegetables,yogurt,oil}"
## [42] "{root vegetables,whipped/sour cream,hygiene articles}"
## [43] "{pork,other vegetables,butter,whipped/sour cream}"
## [44] "{sausage,tropical fruit,root vegetables,rolls/buns}"
## [45] "{curd,domestic eggs,sugar}"
## [46] "{canned fish,hygiene articles}"
## [47] "{citrus fruit,whipped/sour cream,rolls/buns,pastry}"
## [48] "{root vegetables,other vegetables,butter,white bread}"
## [49] "{tropical fruit,root vegetables,other vegetables,yogurt,oil}"
## [50] "{pip fruit,butter,hygiene articles}"
## [51] "{cream cheese ,domestic eggs,napkins}"
## [52] "{pip fruit,root vegetables,hygiene articles}"
## [53] "{other vegetables,butter,whipped/sour cream,domestic eggs}"
## [54] "{rice,sugar}"
## [55] "{cream cheese ,domestic eggs,sugar}"
## [56] "{pip fruit,root vegetables,other vegetables,bottled water}"
## [57] "{root vegetables,whipped/sour cream,flour}"
## [58] "{root vegetables,butter,rice}"
## [59] "{sausage,tropical fruit,root vegetables,yogurt}"
## [60] "{other vegetables,cream cheese ,sugar}"
## [61] "{yogurt,domestic eggs,sugar}"
## [62] "{citrus fruit,domestic eggs,sugar}"
## [63] "{root vegetables,other vegetables,yogurt,rice}"
## [64] "{beef,tropical fruit,yogurt,rolls/buns}"
## [65] "{pip fruit,whipped/sour cream,cream cheese }"
## [66] "{root vegetables,whipped/sour cream,sugar}"
## [67] "{tropical fruit,domestic eggs,hygiene articles}"
## [68] "{pip fruit,other vegetables,whipped/sour cream,domestic eggs}"
## [69] "{pip fruit,root vegetables,other vegetables,brown bread}"
## [70] "{frankfurter,tropical fruit,root vegetables,yogurt}"
## [71] "{root vegetables,whipped/sour cream,soft cheese}"
##  [72] "{tropical fruit,butter,hygiene articles}"
##  [73] "{root vegetables,other vegetables,yogurt,hard cheese}"
##  [74] "{tropical fruit,long life bakery product,napkins}"
##  [75] "{rice,bottled water}"
##  [76] "{butter,whipped/sour cream,sliced cheese}"
##  [77] "{butter,whipped/sour cream,coffee}"
##  [78] "{tropical fruit,butter,yogurt,domestic eggs}"
##  [79] "{whipped/sour cream,house keeping products}"
##  [80] "{root vegetables,butter,white bread}"
##  [81] "{citrus fruit,tropical fruit,herbs}"
##  [82] "{tropical fruit,yogurt,whipped/sour cream,domestic eggs}"
##  [83] "{citrus fruit,butter,curd}"
##  [84] "{domestic eggs,margarine,fruit/vegetable juice}"
##  [85] "{tropical fruit,other vegetables,whipped/sour cream,domestic eggs}"
##  [86] "{pip fruit,other vegetables,yogurt,cream cheese }"
##  [87] "{tropical fruit,curd,yogurt,domestic eggs}"
##  [88] "{tropical fruit,root vegetables,rolls/buns,bottled water}"
##  [89] "{butter,curd,domestic eggs}"
##  [90] "{root vegetables,butter,yogurt,domestic eggs}"
##  [91] "{pip fruit,root vegetables,yogurt,fruit/vegetable juice}"
##  [92] "{soups,bottled beer}"
##  [93] "{tropical fruit,domestic eggs,sugar}"
##  [94] "{tropical fruit,root vegetables,yogurt,pastry}"
##  [95] "{root vegetables,other vegetables,rolls/buns,brown bread}"
##  [96] "{citrus fruit,other vegetables,yogurt,frozen vegetables}"
##  [97] "{pork,rolls/buns,waffles}"
##  [98] "{sausage,butter,long life bakery product}"
##  [99] "{pip fruit,root vegetables,other vegetables,cream cheese }"
## [100] "{citrus fruit,other vegetables,butter,bottled water}"
## [101] "{root vegetables,domestic eggs,coffee}"
## [102] "{tropical fruit,pip fruit,yogurt,frozen meals}"
## [103] "{sausage,pip fruit,cream cheese }"
## [104] "{whipped/sour cream,long life bakery product,napkins}"
## [105] "{tropical fruit,butter,yogurt,sliced cheese}"
## [106] "{sausage,berries,butter}"
## [107] "{root vegetables,other vegetables,yogurt,waffles}"
## [108] "{tropical fruit,other vegetables,butter,yogurt,domestic eggs}"
## [109] "{butter,hygiene articles,napkins}"
## [110] "{tropical fruit,butter,frozen meals}"
## [111] "{citrus fruit,tropical fruit,other vegetables,domestic eggs}"
## [112] "{domestic eggs,margarine,bottled beer}"
## [113] "{other vegetables,butter,whipped/sour cream,napkins}"
## [114] "{tropical fruit,root vegetables,herbs,other vegetables}"
## [115] "{citrus fruit,tropical fruit,curd,yogurt}"
## [116] "{frankfurter,root vegetables,sliced cheese}"
## [117] "{curd,cereals}"
## [118] "{pork,root vegetables,other vegetables,butter}"
## [119] "{pastry,sweet spreads}"
## [120] "{root vegetables,butter,cream cheese }"
## [121] "{other vegetables,curd,whipped/sour cream,cream cheese }"
## [122] "{tropical fruit,other vegetables,butter,white bread}"
## [123] "{tropical fruit,whole milk,butter,sliced cheese}"
## [124] "{liquor,red/blush wine}"
## Itemsets in Consequent (RHS)
## [1] "{whole milk}"       "{yogurt}"           "{root vegetables}"
## [4] "{bottled beer}"     "{other vegetables}"
```



```r
highLiftRules <- head(sort(rules, by="lift"),5)
plot(highLiftRules, method = "graph", control = list(types="items"))
```

```
## Warning: Unknown control parameters: types

## Available control parameters (with default values):
## main  =  Graph for 5 rules
## nodeColors    =  c("#66CC6680", "#9999CC80")
## nodeCol   =  c("#EE0000FF", "#EE0303FF", "#EE0606FF", "#EE0909FF", "#EE0C0CFF", "#EE0F0FFF", "#EE121
## edgeCol   =  c("#474747FF", "#494949FF", "#4B4B4BFF", "#4D4D4DFF", "#4F4F4FFF", "#515151FF", "#53535
## alpha     =  0.5
## cex   =  1
## itemLabels    =  TRUE
## labelCol  =  #000000B3
## measureLabels     =  FALSE
## precision     =  3
## type  =  items
## layout    =  NULL
## layoutParams  =  list()
## arrowSize     =  0.5
## interactive   =  FALSE
## engine    =  igraph
## plot  =  TRUE
## verbose   =  FALSE
```


Graph for 5 rules
size: support (0.001 − 0.002)
color: lift (11.279 − 18.996)
ham
processed cheese
white bread
other vegetables
tropical fruit
yogurt
butter
whipped/sour cream
hamburger meat
Instant food products
soda
popcorn
salty snack