Association rule mining

Revisit the notes on association rule mining and the R example on music playlists: **playlists.R** and **playlists.csv**. Then use the data on grocery purchases in **groceries.txt** and find some interesting association rules for these shopping baskets. The data file is a list of shopping baskets: one person's basket for each row, with multiple items per row separated by commas. Pick your own thresholds for lift and confidence; just be clear what these thresholds are and say why you picked them. Do your discovered item sets make sense? Present your discoveries in an interesting and visually appealing way.

Notes:

This is an exercise in visual and numerical story-telling. Do be clear in your description of what you've done, but keep the focus on the data, the figures, and the insights your analysis has drawn from the data, rather than technical details.

The data file is a list of baskets: one row per basket, with multiple items per row separated by commas. You'll have to cobble together your own code for processing this into the format expected by the "arules" package. This is not intrinsically all that hard, but it is the kind of data-wrangling wrinkle you'll encounter frequently on real problems, where your software package expects data in one format and the data comes in a different format. Figuring out how to bridge that gap is part of the assignment, and so we won't be giving tips on this front.

```r
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```r
library(arules)
```

```
## Loading required package: Matrix
##
## Attaching package: 'Matrix'
##
## The following objects are masked from 'package:tidyr':
##
##     expand, pack, unpack
##
##
## Attaching package: 'arules'
##
## The following object is masked from 'package:dplyr':
```

```
##
##      recode
##
## The following objects are masked from 'package:base':
##
##      abbreviate, write
```
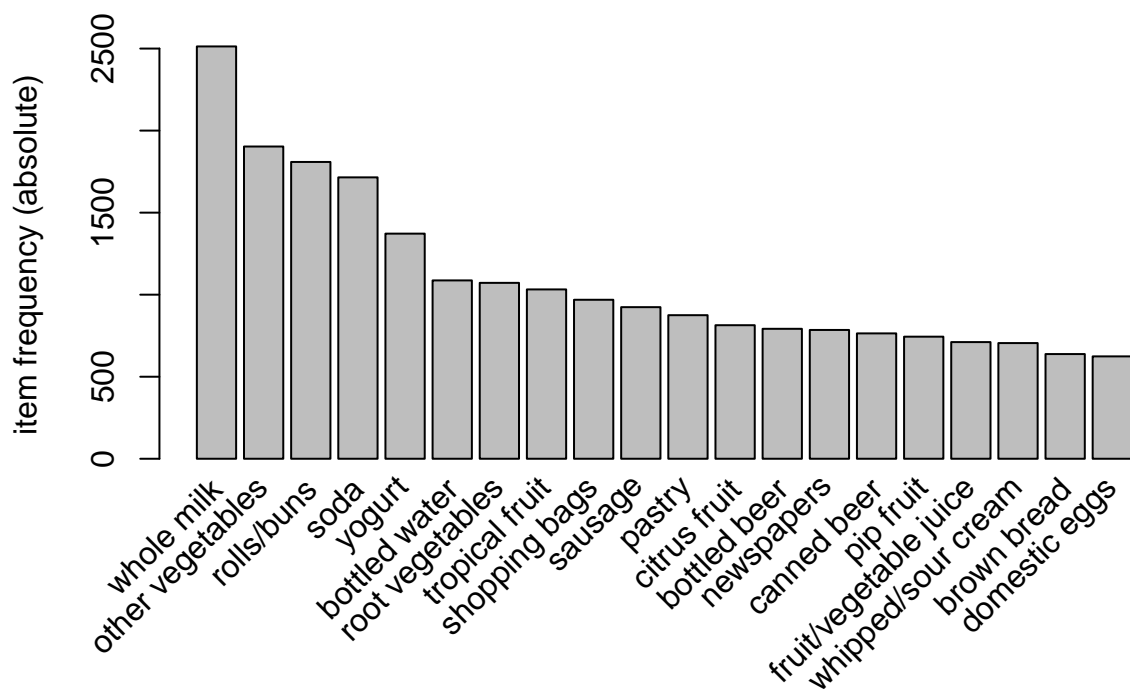
```r
library(arulesViz)

groceries_raw <- read.transactions("/Users/aradelfin/Desktop/Intro to Machine Learning/groceries.txt",

summary(groceries_raw)
```

```
## 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   16
## 2159 1643 1299 1005  855  645  545  438  350  246  182  117   78   77   55   46
##   17   18   19   20   21   22   23   24   26   27   28   29   32
##   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
## 1 abrasive cleaner
## 2 artif. sweetener
## 3   baby cosmetics
```

```r
itemFrequencyPlot(groceries_raw, topN=20, type="absolute", main="Top 20 Most Frequent Items")
```

Top 20 Most Frequent Items


```r
rules <- apriori(groceries_raw, parameter=list(support=0.01, confidence=0.3, maxlen=4))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.3    0.1    1 none FALSE            TRUE       5    0.01      1
##  maxlen target  ext
##       4  rules TRUE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 98
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[169 item(s), 9835 transaction(s)] done [0.00s].
## sorting and recoding items ... [88 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4

## Warning in apriori(groceries_raw, parameter = list(support = 0.01, confidence =
## 0.3, : Mining stopped (maxlen reached). Only patterns up to a length of 4
## returned!
```

```
## done [0.00s].
## writing ... [125 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

summary(rules)

```
## set of 125 rules
##
## rule length distribution (lhs + rhs):sizes
##  2  3
## 69 56
##
##    Min. 1st Qu.  Median   Mean 3rd Qu.    Max.
##   2.000   2.000   2.000  2.448   3.000   3.000
##
## summary of quality measures:
##     support          confidence        coverage           lift
##  Min.   :0.01007   Min.   :0.3079   Min.   :0.01729   Min.   :1.205
##  1st Qu.:0.01149   1st Qu.:0.3454   1st Qu.:0.02888   1st Qu.:1.608
##  Median :0.01454   Median :0.3978   Median :0.03711   Median :1.789
##  Mean   :0.01859   Mean   :0.4058   Mean   :0.04783   Mean   :1.906
##  3rd Qu.:0.02217   3rd Qu.:0.4496   3rd Qu.:0.05663   3rd Qu.:2.155
##  Max.   :0.07483   Max.   :0.5862   Max.   :0.19349   Max.   :3.295
##      count
##  Min.   : 99.0
##  1st Qu.:113.0
##  Median :143.0
##  Mean   :182.8
##  3rd Qu.:218.0
##  Max.   :736.0
##
## mining info:
##            data ntransactions support confidence
##   groceries_raw         9835    0.01        0.3
##                                                                   call
##   apriori(data = groceries_raw, parameter = list(support = 0.01, confidence = 0.3, maxlen = 4))
```

inspect(rules)

```
##         lhs                    rhs                  support confidence   coverage     lift cou
## [1]  {hard cheese}        => {whole milk}       0.01006609  0.4107884 0.02450432 1.607682    9
## [2]  {butter milk}        => {other vegetables} 0.01037112  0.3709091 0.02796136 1.916916   10
## [3]  {butter milk}        => {whole milk}       0.01159126  0.4145455 0.02796136 1.622385   1
## [4]  {ham}                => {whole milk}       0.01148958  0.4414062 0.02602949 1.727509   1
## [5]  {sliced cheese}      => {whole milk}       0.01077783  0.4398340 0.02450432 1.721356   10
## [6]  {oil}                => {whole milk}       0.01128622  0.4021739 0.02806304 1.573968   1
## [7]  {onions}             => {other vegetables} 0.01423488  0.4590164 0.03101169 2.372268   14
## [8]  {onions}             => {whole milk}       0.01209964  0.3901639 0.03101169 1.526965   1
## [9]  {berries}            => {yogurt}           0.01057448  0.3180428 0.03324860 2.279848   10
## [10] {berries}            => {other vegetables} 0.01026945  0.3088685 0.03324860 1.596280   10
## [11] {berries}            => {whole milk}       0.01179461  0.3547401 0.03324860 1.388328   1
## [12] {hamburger meat}     => {other vegetables} 0.01382816  0.4159021 0.03324860 2.149447   13
## [13] {hamburger meat}     => {whole milk}       0.01474326  0.4434251 0.03324860 1.735410   14
```

4

```
## [14]  {hygiene articles}        => {whole milk}      0.01281139  0.3888889 0.03294357 1.521975  1
## [15]  {sugar}                   => {other vegetables} 0.01077783  0.3183183 0.03385867 1.645119  1
## [16]  {sugar}                   => {whole milk}      0.01504830  0.4444444 0.03385867 1.739400  1
## [17]  {waffles}                 => {whole milk}      0.01270971  0.3306878 0.03843416 1.294196  1
## [18]  {long life bakery product} => {whole milk}     0.01352313  0.3614130 0.03741739 1.414444  1
## [19]  {dessert}                 => {other vegetables} 0.01159126  0.3123288 0.03711235 1.614164  1
## [20]  {dessert}                 => {whole milk}      0.01372649  0.3698630 0.03711235 1.447514  1
## [21]  {cream cheese}            => {yogurt}          0.01240468  0.3128205 0.03965430 2.242412  1
## [22]  {cream cheese}            => {other vegetables} 0.01372649  0.3461538 0.03965430 1.788977  1
## [23]  {cream cheese}            => {whole milk}      0.01647178  0.4153846 0.03965430 1.625670  1
## [24]  {chicken}                 => {other vegetables} 0.01789527  0.4170616 0.04290798 2.155439  1
## [25]  {chicken}                 => {whole milk}      0.01759024  0.4099526 0.04290798 1.604411  1
## [26]  {white bread}             => {other vegetables} 0.01372649  0.3260870 0.04209456 1.685268  1
## [27]  {white bread}             => {whole milk}      0.01708185  0.4057971 0.04209456 1.588147  1
## [28]  {chocolate}               => {whole milk}      0.01667514  0.3360656 0.04961871 1.315243  1
## [29]  {coffee}                  => {whole milk}      0.01870869  0.3222417 0.05805796 1.261141  1
## [30]  {frozen vegetables}       => {other vegetables} 0.01779359  0.3699789 0.04809354 1.912108  1
## [31]  {frozen vegetables}       => {whole milk}      0.02043721  0.4249471 0.04809354 1.663094  2
## [32]  {beef}                    => {root vegetables} 0.01738688  0.3313953 0.05246568 3.040367  1
## [33]  {beef}                    => {other vegetables} 0.01972547  0.3759690 0.05246568 1.943066  1
## [34]  {beef}                    => {whole milk}      0.02125064  0.4050388 0.05246568 1.585180  2
## [35]  {curd}                    => {yogurt}          0.01728521  0.3244275 0.05327911 2.325615  1
## [36]  {curd}                    => {other vegetables} 0.01718353  0.3225191 0.05327911 1.666829  1
## [37]  {curd}                    => {whole milk}      0.02613116  0.4904580 0.05327911 1.919481  2
## [38]  {napkins}                 => {whole milk}      0.01972547  0.3766990 0.05236401 1.474268  1
## [39]  {pork}                    => {other vegetables} 0.02165735  0.3756614 0.05765125 1.941476  2
## [40]  {pork}                    => {whole milk}      0.02216573  0.3844797 0.05765125 1.504719  2
## [41]  {frankfurter}             => {rolls/buns}      0.01921708  0.3258621 0.05897306 1.771616  1
## [42]  {frankfurter}             => {whole milk}      0.02053889  0.3482759 0.05897306 1.363029  2
## [43]  {brown bread}             => {whole milk}      0.02521607  0.3887147 0.06487036 1.521293  2
## [44]  {margarine}               => {other vegetables} 0.01972547  0.3368056 0.05856634 1.740663  1
## [45]  {margarine}               => {whole milk}      0.02419929  0.4131944 0.05856634 1.617098  2
## [46]  {butter}                  => {other vegetables} 0.02003050  0.3614679 0.05541434 1.868122  1
## [47]  {butter}                  => {whole milk}      0.02755465  0.4972477 0.05541434 1.946053  2
## [48]  {newspapers}              => {whole milk}      0.02735130  0.3426752 0.07981698 1.341110  2
## [49]  {domestic eggs}           => {other vegetables} 0.02226741  0.3509615 0.06344687 1.813824  2
## [50]  {domestic eggs}           => {whole milk}      0.02999492  0.4727564 0.06344687 1.850203  2
## [51]  {fruit/vegetable juice}   => {whole milk}      0.02663955  0.3684951 0.07229283 1.442160  2
## [52]  {whipped/sour cream}      => {other vegetables} 0.02887646  0.4028369 0.07168277 2.081924  2
## [53]  {whipped/sour cream}      => {whole milk}      0.03223183  0.4496454 0.07168277 1.759754  3
## [54]  {pip fruit}               => {other vegetables} 0.02613116  0.3454301 0.07564820 1.785237  2
## [55]  {pip fruit}               => {whole milk}      0.03009659  0.3978495 0.07564820 1.557043  2
## [56]  {pastry}                  => {whole milk}      0.03324860  0.3737143 0.08896797 1.462587  3
## [57]  {citrus fruit}            => {other vegetables} 0.02887646  0.3488943 0.08276563 1.803140  2
## [58]  {citrus fruit}            => {whole milk}      0.03050330  0.3685504 0.08276563 1.442377  3
## [59]  {sausage}                 => {rolls/buns}      0.03060498  0.3257576 0.09395018 1.771048  3
## [60]  {sausage}                 => {whole milk}      0.02989324  0.3181818 0.09395018 1.245252  2
## [61]  {bottled water}           => {whole milk}      0.03436706  0.3109476 0.11052364 1.216940  3
## [62]  {tropical fruit}          => {other vegetables} 0.03589222  0.3420543 0.10493137 1.767790  3
## [63]  {tropical fruit}          => {whole milk}      0.04229792  0.4031008 0.10493137 1.577595  4
## [64]  {root vegetables}         => {other vegetables} 0.04738180  0.4347015 0.10899847 2.246605  4
## [65]  {root vegetables}         => {whole milk}      0.04890696  0.4486940 0.10899847 1.756031  4
## [66]  {yogurt}                  => {other vegetables} 0.04341637  0.3112245 0.13950178 1.608457  4
## [67]  {yogurt}                  => {whole milk}      0.05602440  0.4016035 0.13950178 1.571735  5
```

```
## [68]  {rolls/buns}              => {whole milk}       0.05663447  0.3079049 0.18393493 1.205032   5!
## [69]  {other vegetables}        => {whole milk}       0.07483477  0.3867578 0.19349263 1.513634   7:
## [70]  {curd,
##        yogurt}                   => {whole milk}       0.01006609  0.5823529 0.01728521 2.279125     !
## [71]  {curd,
##        whole milk}               => {yogurt}           0.01006609  0.3852140 0.02613116 2.761356     !
## [72]  {other vegetables,
##        pork}                     => {whole milk}       0.01016777  0.4694836 0.02165735 1.837394   1(
## [73]  {pork,
##        whole milk}               => {other vegetables} 0.01016777  0.4587156 0.02216573 2.370714   1(
## [74]  {butter,
##        other vegetables}         => {whole milk}       0.01148958  0.5736041 0.02003050 2.244885   1:
## [75]  {butter,
##        whole milk}               => {other vegetables} 0.01148958  0.4169742 0.02755465 2.154987   1:
## [76]  {domestic eggs,
##        other vegetables}         => {whole milk}       0.01230300  0.5525114 0.02226741 2.162336   1:
## [77]  {domestic eggs,
##        whole milk}               => {other vegetables} 0.01230300  0.4101695 0.02999492 2.119820   1:
## [78]  {fruit/vegetable juice,
##        other vegetables}         => {whole milk}       0.01047280  0.4975845 0.02104728 1.947371   1(
## [79]  {fruit/vegetable juice,
##        whole milk}               => {other vegetables} 0.01047280  0.3931298 0.02663955 2.031756   1(
## [80]  {whipped/sour cream,
##        yogurt}                   => {other vegetables} 0.01016777  0.4901961 0.02074225 2.533410   1(
## [81]  {other vegetables,
##        whipped/sour cream}       => {yogurt}           0.01016777  0.3521127 0.02887646 2.524073   1(
## [82]  {whipped/sour cream,
##        yogurt}                   => {whole milk}       0.01087951  0.5245098 0.02074225 2.052747   1(
## [83]  {whipped/sour cream,
##        whole milk}               => {yogurt}           0.01087951  0.3375394 0.03223183 2.419607   1(
## [84]  {other vegetables,
##        whipped/sour cream}       => {whole milk}       0.01464159  0.5070423 0.02887646 1.984385   1
## [85]  {whipped/sour cream,
##        whole milk}               => {other vegetables} 0.01464159  0.4542587 0.03223183 2.347679   1
## [86]  {other vegetables,
##        pip fruit}                => {whole milk}       0.01352313  0.5175097 0.02613116 2.025351   1:
## [87]  {pip fruit,
##        whole milk}               => {other vegetables} 0.01352313  0.4493243 0.03009659 2.322178   1:
## [88]  {other vegetables,
##        pastry}                   => {whole milk}       0.01057448  0.4684685 0.02257245 1.833421   1(
## [89]  {pastry,
##        whole milk}               => {other vegetables} 0.01057448  0.3180428 0.03324860 1.643695   1(
## [90]  {citrus fruit,
##        root vegetables}          => {other vegetables} 0.01037112  0.5862069 0.01769192 3.029608   1(
## [91]  {citrus fruit,
##        other vegetables}         => {root vegetables}  0.01037112  0.3591549 0.02887646 3.295045   1(
## [92]  {citrus fruit,
##        yogurt}                   => {whole milk}       0.01026945  0.4741784 0.02165735 1.855768   1(
## [93]  {citrus fruit,
##        whole milk}               => {yogurt}           0.01026945  0.3366667 0.03050330 2.413350   1(
## [94]  {citrus fruit,
##        other vegetables}         => {whole milk}       0.01301474  0.4507042 0.02887646 1.763898   1:
## [95]  {citrus fruit,
##        whole milk}               => {other vegetables} 0.01301474  0.4266667 0.03050330 2.205080   1:
```

```
## [96]  {other vegetables,
##        sausage}              => {whole milk}       0.01016777 0.3773585 0.02694459 1.476849  1
## [97]  {sausage,
##        whole milk}           => {other vegetables} 0.01016777 0.3401361 0.02989324 1.757876  1
## [98]  {bottled water,
##        other vegetables}     => {whole milk}       0.01077783 0.4344262 0.02480935 1.700192  1
## [99]  {bottled water,
##        whole milk}           => {other vegetables} 0.01077783 0.3136095 0.03436706 1.620783  1
## [100] {root vegetables,
##        tropical fruit}       => {other vegetables} 0.01230300 0.5845411 0.02104728 3.020999  1
## [101] {other vegetables,
##        tropical fruit}       => {root vegetables}  0.01230300 0.3427762 0.03589222 3.144780  1
## [102] {root vegetables,
##        tropical fruit}       => {whole milk}       0.01199797 0.5700483 0.02104728 2.230969  1
## [103] {tropical fruit,
##        yogurt}               => {other vegetables} 0.01230300 0.4201389 0.02928317 2.171343  1
## [104] {other vegetables,
##        tropical fruit}       => {yogurt}           0.01230300 0.3427762 0.03589222 2.457146  1
## [105] {tropical fruit,
##        yogurt}               => {whole milk}       0.01514997 0.5173611 0.02928317 2.024770  1
## [106] {tropical fruit,
##        whole milk}           => {yogurt}           0.01514997 0.3581731 0.04229792 2.567516  1
## [107] {rolls/buns,
##        tropical fruit}       => {whole milk}       0.01098119 0.4462810 0.02460600 1.746587  1
## [108] {other vegetables,
##        tropical fruit}       => {whole milk}       0.01708185 0.4759207 0.03589222 1.862587  1
## [109] {tropical fruit,
##        whole milk}           => {other vegetables} 0.01708185 0.4038462 0.04229792 2.087140  1
## [110] {root vegetables,
##        yogurt}               => {other vegetables} 0.01291307 0.5000000 0.02582613 2.584078  1
## [111] {root vegetables,
##        yogurt}               => {whole milk}       0.01453991 0.5629921 0.02582613 2.203354  1
## [112] {rolls/buns,
##        root vegetables}      => {other vegetables} 0.01220132 0.5020921 0.02430097 2.594890  1
## [113] {rolls/buns,
##        root vegetables}      => {whole milk}       0.01270971 0.5230126 0.02430097 2.046888  1
## [114] {other vegetables,
##        root vegetables}      => {whole milk}       0.02318251 0.4892704 0.04738180 1.914833  2
## [115] {root vegetables,
##        whole milk}           => {other vegetables} 0.02318251 0.4740125 0.04890696 2.449770  2
## [116] {other vegetables,
##        whole milk}           => {root vegetables}  0.02318251 0.3097826 0.07483477 2.842082  2
## [117] {soda,
##        yogurt}               => {whole milk}       0.01047280 0.3828996 0.02735130 1.498535  1
## [118] {other vegetables,
##        soda}                 => {whole milk}       0.01392984 0.4254658 0.03274021 1.665124  1
## [119] {soda,
##        whole milk}           => {other vegetables} 0.01392984 0.3477157 0.04006101 1.797049  1
## [120] {rolls/buns,
##        yogurt}               => {other vegetables} 0.01148958 0.3343195 0.03436706 1.727815  1
## [121] {rolls/buns,
##        yogurt}               => {whole milk}       0.01555669 0.4526627 0.03436706 1.771563  1
## [122] {other vegetables,
##        yogurt}               => {whole milk}       0.02226741 0.5128806 0.04341637 2.007235  2
```

```
## [123] {whole milk,
##        yogurt}               => {other vegetables} 0.02226741   0.3974592 0.05602440 2.054131    2
## [124] {other vegetables,
##        rolls/buns}           => {whole milk}        0.01789527   0.4200477 0.04260295 1.643919    1
## [125] {rolls/buns,
##        whole milk}           => {other vegetables} 0.01789527   0.3159785 0.05663447 1.633026    1
```

KKF

```
high_lift_rules <- subset(rules, lift > 2)
high_confidence_rules <- subset(rules, confidence > 0.4)

inspect(high_lift_rules)
```

```
##       lhs                        rhs                        support confidence   coverage      lift count
## [1]   {onions}                => {other vegetables} 0.01423488  0.4590164 0.03101169 2.372268   140
## [2]   {berries}               => {yogurt}           0.01057448  0.3180428 0.03324860 2.279848   104
## [3]   {hamburger meat}        => {other vegetables} 0.01382816  0.4159021 0.03324860 2.149447   136
## [4]   {cream cheese}          => {yogurt}           0.01240468  0.3128205 0.03965430 2.242412   122
## [5]   {chicken}               => {other vegetables} 0.01789527  0.4170616 0.04290798 2.155439   176
## [6]   {beef}                  => {root vegetables}  0.01738688  0.3313953 0.05246568 3.040367   171
## [7]   {curd}                  => {yogurt}           0.01728521  0.3244275 0.05327911 2.325615   170
## [8]   {whipped/sour cream}    => {other vegetables} 0.02887646  0.4028369 0.07168277 2.081924   284
## [9]   {root vegetables}       => {other vegetables} 0.04738180  0.4347015 0.10899847 2.246605   466
## [10]  {curd,
##        yogurt}               => {whole milk}        0.01006609  0.5823529 0.01728521 2.279125    99
## [11]  {curd,
##        whole milk}           => {yogurt}            0.01006609  0.3852140 0.02613116 2.761356    99
## [12]  {pork,
##        whole milk}           => {other vegetables} 0.01016777  0.4587156 0.02216573 2.370714   100
## [13]  {butter,
##        other vegetables}     => {whole milk}        0.01148958  0.5736041 0.02003050 2.244885   113
## [14]  {butter,
##        whole milk}           => {other vegetables} 0.01148958  0.4169742 0.02755465 2.154987   113
## [15]  {domestic eggs,
##        other vegetables}     => {whole milk}        0.01230300  0.5525114 0.02226741 2.162336   121
## [16]  {domestic eggs,
##        whole milk}           => {other vegetables} 0.01230300  0.4101695 0.02999492 2.119820   121
## [17]  {fruit/vegetable juice,
##        whole milk}           => {other vegetables} 0.01047280  0.3931298 0.02663955 2.031756   103
## [18]  {whipped/sour cream,
##        yogurt}               => {other vegetables} 0.01016777  0.4901961 0.02074225 2.533410   100
## [19]  {other vegetables,
##        whipped/sour cream}   => {yogurt}            0.01016777  0.3521127 0.02887646 2.524073   100
## [20]  {whipped/sour cream,
##        yogurt}               => {whole milk}        0.01087951  0.5245098 0.02074225 2.052747   107
## [21]  {whipped/sour cream,
##        whole milk}           => {yogurt}            0.01087951  0.3375394 0.03223183 2.419607   107
## [22]  {whipped/sour cream,
##        whole milk}           => {other vegetables} 0.01464159  0.4542587 0.03223183 2.347679   144
## [23]  {other vegetables,
##        pip fruit}            => {whole milk}        0.01352313  0.5175097 0.02613116 2.025351   133
## [24]  {pip fruit,
```

8

```
##        whole milk}          => {other vegetables} 0.01352313  0.4493243 0.03009659 2.322178   133
## [25] {citrus fruit,
##        root vegetables}     => {other vegetables} 0.01037112  0.5862069 0.01769192 3.029608   102
## [26] {citrus fruit,
##        other vegetables}    => {root vegetables}  0.01037112  0.3591549 0.02887646 3.295045   102
## [27] {citrus fruit,
##        whole milk}          => {yogurt}           0.01026945  0.3366667 0.03050330 2.413350   101
## [28] {citrus fruit,
##        whole milk}          => {other vegetables} 0.01301474  0.4266667 0.03050330 2.205080   128
## [29] {root vegetables,
##        tropical fruit}      => {other vegetables} 0.01230300  0.5845411 0.02104728 3.020999   121
## [30] {other vegetables,
##        tropical fruit}      => {root vegetables}  0.01230300  0.3427762 0.03589222 3.144780   121
## [31] {root vegetables,
##        tropical fruit}      => {whole milk}       0.01199797  0.5700483 0.02104728 2.230969   118
## [32] {tropical fruit,
##        yogurt}              => {other vegetables} 0.01230300  0.4201389 0.02928317 2.171343   121
## [33] {other vegetables,
##        tropical fruit}      => {yogurt}           0.01230300  0.3427762 0.03589222 2.457146   121
## [34] {tropical fruit,
##        yogurt}              => {whole milk}       0.01514997  0.5173611 0.02928317 2.024770   149
## [35] {tropical fruit,
##        whole milk}          => {yogurt}           0.01514997  0.3581731 0.04229792 2.567516   149
## [36] {tropical fruit,
##        whole milk}          => {other vegetables} 0.01708185  0.4038462 0.04229792 2.087140   168
## [37] {root vegetables,
##        yogurt}              => {other vegetables} 0.01291307  0.5000000 0.02582613 2.584078   127
## [38] {root vegetables,
##        yogurt}              => {whole milk}       0.01453991  0.5629921 0.02582613 2.203354   143
## [39] {rolls/buns,
##        root vegetables}     => {other vegetables} 0.01220132  0.5020921 0.02430097 2.594890   120
## [40] {rolls/buns,
##        root vegetables}     => {whole milk}       0.01270971  0.5230126 0.02430097 2.046888   125
## [41] {root vegetables,
##        whole milk}          => {other vegetables} 0.02318251  0.4740125 0.04890696 2.449770   228
## [42] {other vegetables,
##        whole milk}          => {root vegetables}  0.02318251  0.3097826 0.07483477 2.842082   228
## [43] {other vegetables,
##        yogurt}              => {whole milk}       0.02226741  0.5128806 0.04341637 2.007235   219
## [44] {whole milk,
##        yogurt}              => {other vegetables} 0.02226741  0.3974592 0.05602440 2.054131   219
```

```r
inspect(high_confidence_rules)
```

```
##       lhs                 rhs                    support confidence  coverage     lift count
## [1]  {hard cheese}     => {whole milk}        0.01006609  0.4107884 0.02450432 1.607682    99
## [2]  {butter milk}     => {whole milk}        0.01159126  0.4145455 0.02796136 1.622385   114
## [3]  {ham}             => {whole milk}        0.01148958  0.4414062 0.02602949 1.727509   113
## [4]  {sliced cheese}   => {whole milk}        0.01077783  0.4398340 0.02450432 1.721356   106
## [5]  {oil}             => {whole milk}        0.01128622  0.4021739 0.02806304 1.573968   111
## [6]  {onions}          => {other vegetables}  0.01423488  0.4590164 0.03101169 2.372268   140
## [7]  {hamburger meat}  => {other vegetables}  0.01382816  0.4159021 0.03324860 2.149447   136
## [8]  {hamburger meat}  => {whole milk}        0.01474326  0.4434251 0.03324860 1.735410   145
## [9]  {sugar}           => {whole milk}        0.01504830  0.4444444 0.03385867 1.739400   148
```
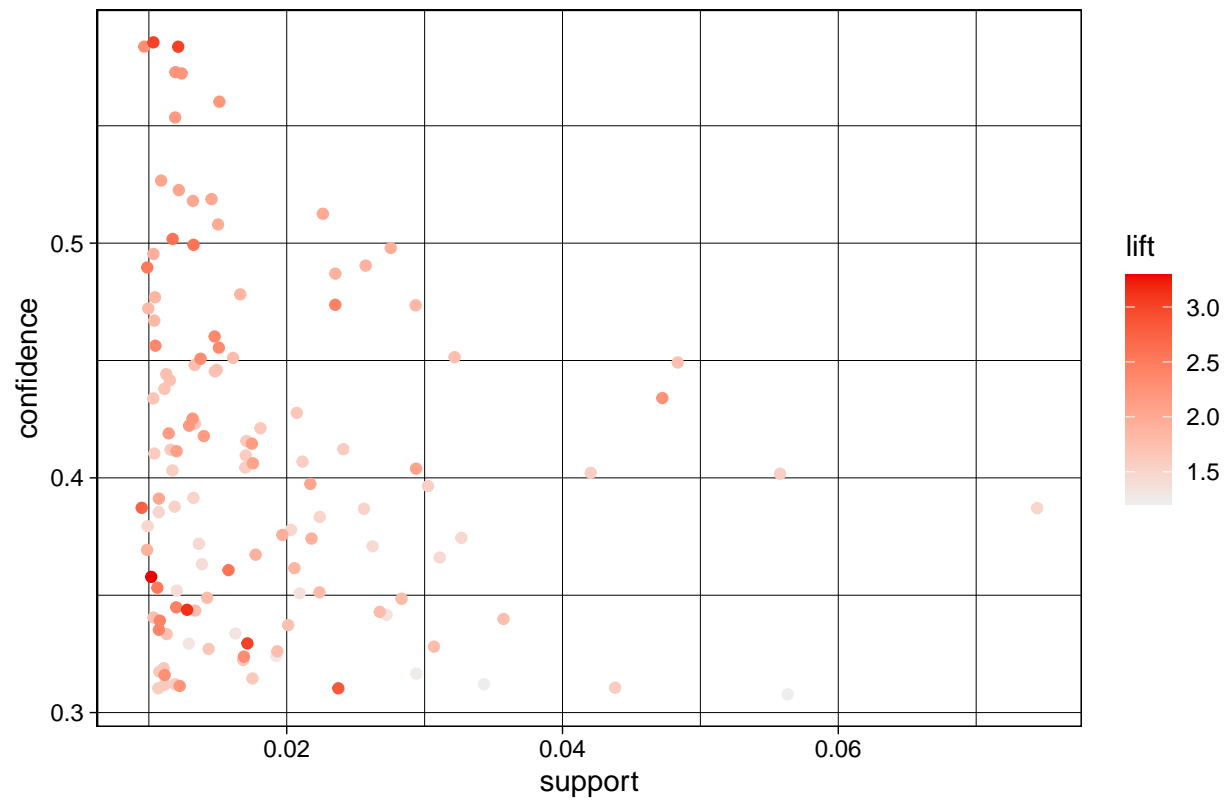
9

```
## [10] {cream cheese}           => {whole milk}       0.01647178 0.4153846 0.03965430 1.625670 162
## [11] {chicken}                => {other vegetables} 0.01789527 0.4170616 0.04290798 2.155439 176
## [12] {chicken}                => {whole milk}       0.01759024 0.4099526 0.04290798 1.604411 173
## [13] {white bread}            => {whole milk}       0.01708185 0.4057971 0.04209456 1.588147 168
## [14] {frozen vegetables}      => {whole milk}       0.02043721 0.4249471 0.04809354 1.663094 201
## [15] {beef}                   => {whole milk}       0.02125064 0.4050388 0.05246568 1.585180 209
## [16] {curd}                   => {whole milk}       0.02613116 0.4904580 0.05327911 1.919481 257
## [17] {margarine}              => {whole milk}       0.02419929 0.4131944 0.05856634 1.617098 238
## [18] {butter}                 => {whole milk}       0.02755465 0.4972477 0.05541434 1.946053 271
## [19] {domestic eggs}          => {whole milk}       0.02999492 0.4727564 0.06344687 1.850203 295
## [20] {whipped/sour cream}     => {other vegetables} 0.02887646 0.4028369 0.07168277 2.081924 284
## [21] {whipped/sour cream}     => {whole milk}       0.03223183 0.4496454 0.07168277 1.759754 317
## [22] {tropical fruit}         => {whole milk}       0.04229792 0.4031008 0.10493137 1.577595 416
## [23] {root vegetables}        => {other vegetables} 0.04738180 0.4347015 0.10899847 2.246605 466
## [24] {root vegetables}        => {whole milk}       0.04890696 0.4486940 0.10899847 1.756031 481
## [25] {yogurt}                 => {whole milk}       0.05602440 0.4016035 0.13950178 1.571735 551
## [26] {curd,
##       yogurt}                 => {whole milk}       0.01006609 0.5823529 0.01728521 2.279125  99
## [27] {other vegetables,
##       pork}                   => {whole milk}       0.01016777 0.4694836 0.02165735 1.837394 100
## [28] {pork,
##       whole milk}             => {other vegetables} 0.01016777 0.4587156 0.02216573 2.370714 100
## [29] {butter,
##       other vegetables}       => {whole milk}       0.01148958 0.5736041 0.02003050 2.244885 113
## [30] {butter,
##       whole milk}             => {other vegetables} 0.01148958 0.4169742 0.02755465 2.154987 113
## [31] {domestic eggs,
##       other vegetables}       => {whole milk}       0.01230300 0.5525114 0.02226741 2.162336 121
## [32] {domestic eggs,
##       whole milk}             => {other vegetables} 0.01230300 0.4101695 0.02999492 2.119820 121
## [33] {fruit/vegetable juice,
##       other vegetables}       => {whole milk}       0.01047280 0.4975845 0.02104728 1.947371 103
## [34] {whipped/sour cream,
##       yogurt}                 => {other vegetables} 0.01016777 0.4901961 0.02074225 2.533410 100
## [35] {whipped/sour cream,
##       yogurt}                 => {whole milk}       0.01087951 0.5245098 0.02074225 2.052747 107
## [36] {other vegetables,
##       whipped/sour cream}     => {whole milk}       0.01464159 0.5070423 0.02887646 1.984385 144
## [37] {whipped/sour cream,
##       whole milk}             => {other vegetables} 0.01464159 0.4542587 0.03223183 2.347679 144
## [38] {other vegetables,
##       pip fruit}              => {whole milk}       0.01352313 0.5175097 0.02613116 2.025351 133
## [39] {pip fruit,
##       whole milk}             => {other vegetables} 0.01352313 0.4493243 0.03009659 2.322178 133
## [40] {other vegetables,
##       pastry}                 => {whole milk}       0.01057448 0.4684685 0.02257245 1.833421 104
## [41] {citrus fruit,
##       root vegetables}        => {other vegetables} 0.01037112 0.5862069 0.01769192 3.029608 102
## [42] {citrus fruit,
##       yogurt}                 => {whole milk}       0.01026945 0.4741784 0.02165735 1.855768 101
## [43] {citrus fruit,
##       other vegetables}       => {whole milk}       0.01301474 0.4507042 0.02887646 1.763898 128
## [44] {citrus fruit,
##       whole milk}             => {other vegetables} 0.01301474 0.4266667 0.03050330 2.205080 128
```

```
## [45] {bottled water,
##       other vegetables}      => {whole milk}       0.01077783  0.4344262 0.02480935 1.700192   106
## [46] {root vegetables,
##       tropical fruit}        => {other vegetables} 0.01230300  0.5845411 0.02104728 3.020999   121
## [47] {root vegetables,
##       tropical fruit}        => {whole milk}       0.01199797  0.5700483 0.02104728 2.230969   118
## [48] {tropical fruit,
##       yogurt}                => {other vegetables} 0.01230300  0.4201389 0.02928317 2.171343   121
## [49] {tropical fruit,
##       yogurt}                => {whole milk}       0.01514997  0.5173611 0.02928317 2.024770   149
## [50] {rolls/buns,
##       tropical fruit}        => {whole milk}       0.01098119  0.4462810 0.02460600 1.746587   108
## [51] {other vegetables,
##       tropical fruit}        => {whole milk}       0.01708185  0.4759207 0.03589222 1.862587   168
## [52] {tropical fruit,
##       whole milk}            => {other vegetables} 0.01708185  0.4038462 0.04229792 2.087140   168
## [53] {root vegetables,
##       yogurt}                => {other vegetables} 0.01291307  0.5000000 0.02582613 2.584078   127
## [54] {root vegetables,
##       yogurt}                => {whole milk}       0.01453991  0.5629921 0.02582613 2.203354   143
## [55] {rolls/buns,
##       root vegetables}       => {other vegetables} 0.01220132  0.5020921 0.02430097 2.594890   120
## [56] {rolls/buns,
##       root vegetables}       => {whole milk}       0.01270971  0.5230126 0.02430097 2.046888   125
## [57] {other vegetables,
##       root vegetables}       => {whole milk}       0.02318251  0.4892704 0.04738180 1.914833   228
## [58] {root vegetables,
##       whole milk}            => {other vegetables} 0.02318251  0.4740125 0.04890696 2.449770   228
## [59] {other vegetables,
##       soda}                  => {whole milk}       0.01392984  0.4254658 0.03274021 1.665124   137
## [60] {rolls/buns,
##       yogurt}                => {whole milk}       0.01555669  0.4526627 0.03436706 1.771563   153
## [61] {other vegetables,
##       yogurt}                => {whole milk}       0.02226741  0.5128806 0.04341637 2.007235   219
## [62] {other vegetables,
##       rolls/buns}            => {whole milk}       0.01789527  0.4200477 0.04260295 1.643919   176
```
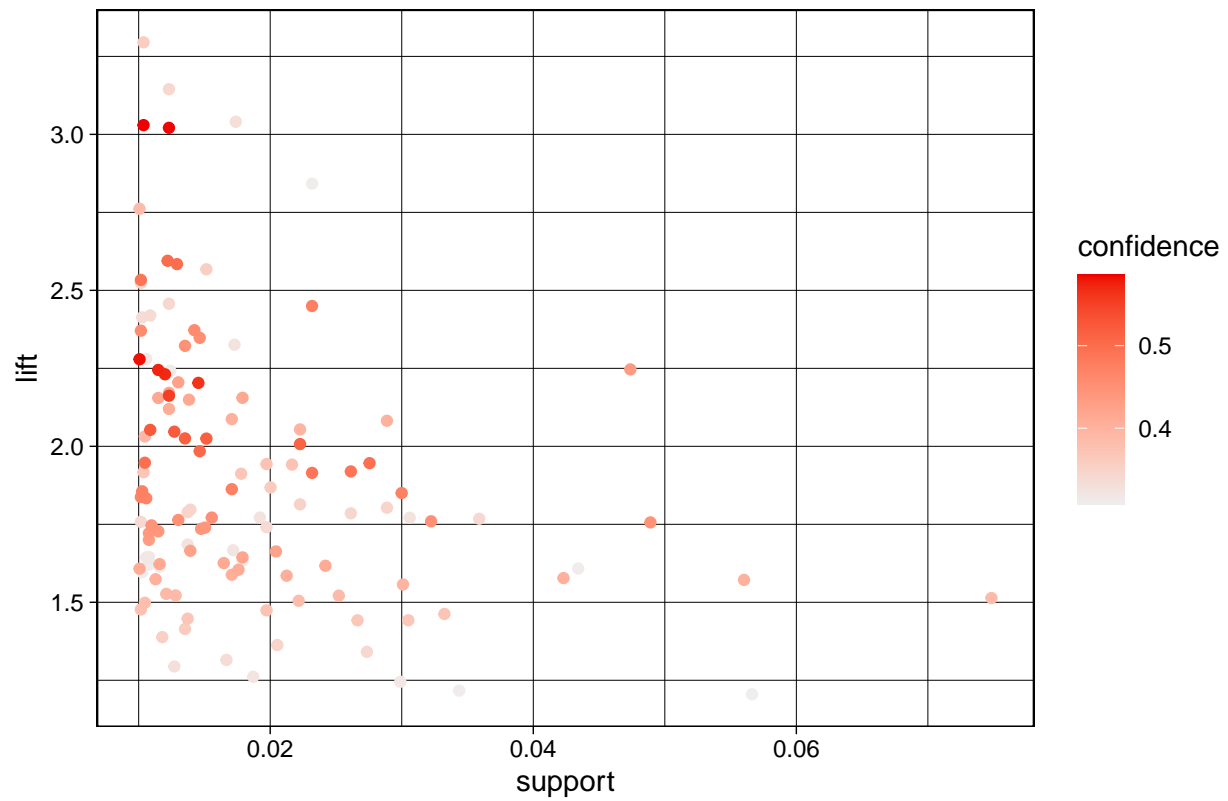
```
plot(rules)
```

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

Scatter plot for 125 rules

```
plot(rules, measure = c("support", "lift"), shading = "confidence")
```
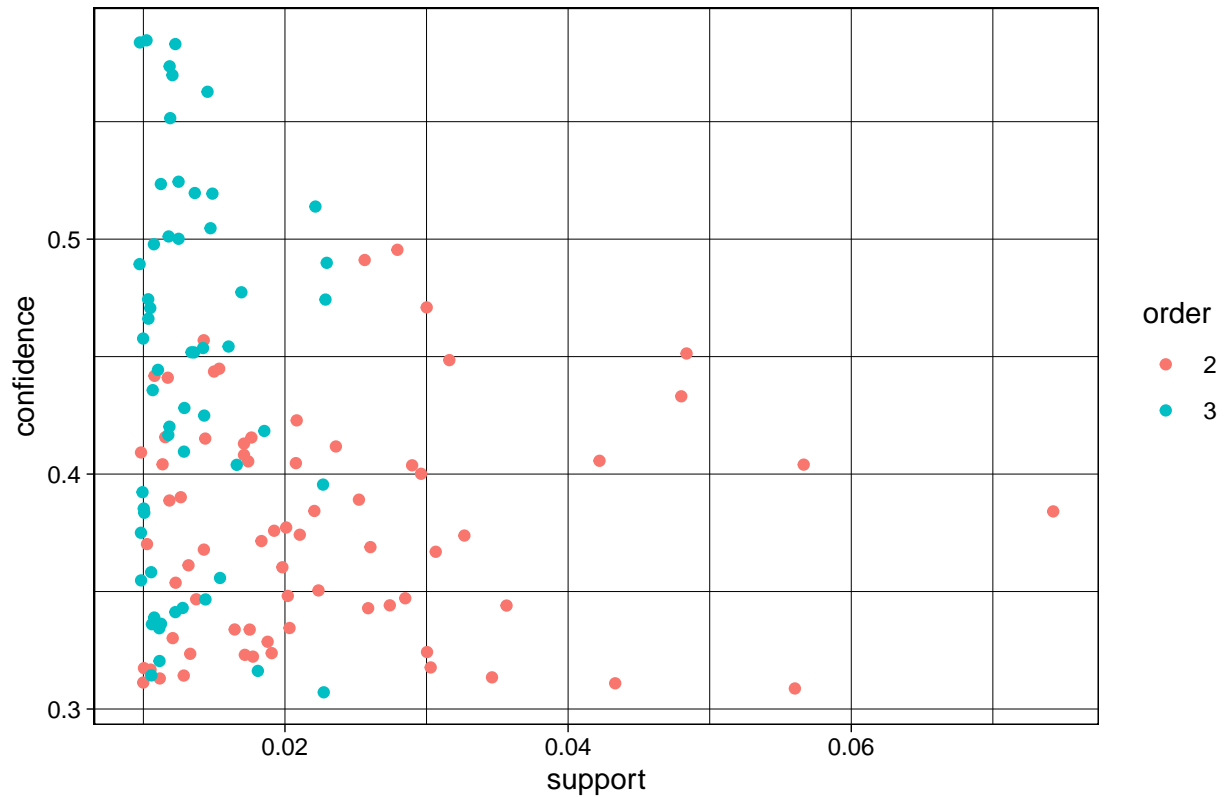
Scatter plot for 125 rules

```
plot(rules, method = "two-key plot")
```

```
## To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
```
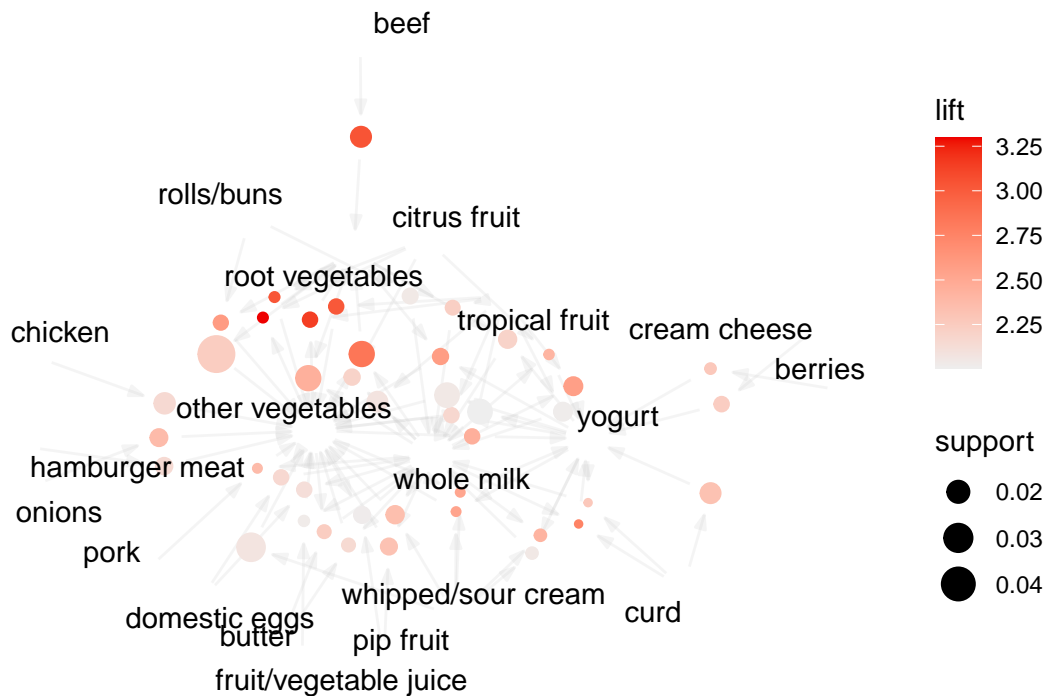
Scatter plot for 125 rules


```r
plot(high_lift_rules, method="graph", control=list(type="items"))
```

```
## Warning: Unknown control parameters: type

## Available control parameters (with default values):
## layout     =  stress
## circular   =  FALSE
## ggraphdots    =  NULL
## edges      =  <environment>
## nodes      =  <environment>
## nodetext  =  <environment>
## colors     =  c("#EE0000FF", "#EEEEEEFF")
## engine     =  ggplot2
## max    =  100
## verbose    =  FALSE
```

```
#graph_rules <- associations2igraph(high_lift_rules, associationsAsNodes = FALSE)
#igraph::write_graph(graph_rules, file='groceries_rules0.graphml', format = "graphml")
```

This analysis aims to uncover patterns in grocery shopping data that can inform business strategies like cross-selling, inventory management, and targeted marketing.

To ensure the most relevant and impactful rules were identified, thresholds were set for support, confidence, and lift. Specifically, a support threshold of 0.01 was used, meaning that the rule applies to at least 1% of all transactions. Confidence was set at 0.3, indicating that in 30% of cases where the left-hand side (LHS) items are bought, the right-hand side (RHS) items are also purchased. Additionally, a lift threshold of greater than 2.0 was applied, ensuring that the RHS items are at least twice as likely to be bought when the LHS items are purchased compared to when they are not.

Key Findings

1. Beef and Root Vegetables
Rule: {beef} => {root vegetables}
Lift: 3.04
Confidence: 33.13%

This rule suggests that customers who buy beef are three times more likely to purchase root vegetables compared to a random customer. The strong complementarity between these items makes this pairing a potential target for promotions or product bundling strategies.

2. Whipped/Sour Cream and Yogurt
Rule: {whipped/sour cream} => {yogurt}
Lift: 2.53
Confidence: 49.0%

This rule suggests that customers who purchase whipped or sour cream are more than twice as likely to buy yogurt as well. The confidence level indicates that nearly half of the transactions involving whipped or sour cream also include yogurt. This insight could be leveraged by offering promotions that combine these dairy products. For example, a "buy whipped cream, get a discount on yogurt" offer could increase sales of both items. Additionally, these items could be placed together in the dairy section to encourage customers to purchase them together.

3. Citrus Fruit, Root Vegetables, and Other Vegetables
Rule: {citrus fruit, root vegetables} => {other vegetables}
Lift: 3.03
Confidence: 58.6%

The association between citrus fruit, root vegetables, and other vegetables indicates a natural grouping in customer shopping habits. With customers being three times more likely to buy other vegetables when they already have citrus fruit and root vegetables in their basket.

Conclusion
In analyzing the grocery shopping data, we identified several significant association rules that reveal important insights into customer purchasing behavior. By focusing on patterns with high lift and confidence, we discovered strong correlations between products like beef and root vegetables, citrus fruit and other vegetables, and dairy combinations such as curd, yogurt, and whole milk. These insights offer actionable opportunities for enhancing business strategies, including cross-selling, targeted promotions, and optimized store layouts. For instance, leveraging these associations through strategic product placement or bundled offers can boost sales and improve customer satisfaction.