Market Basket Analysis

2024-04-11

```
library(arules)
```

```
## Warning: package 'arules' was built under R version 4.2.3

## Loading required package: Matrix

##
## Attaching package: 'arules'

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

```
data(Groceries)
str(Groceries)
```

```
## Formal class 'transactions' [package "arules"] with 3 slots
##   ..@ data       :Formal class 'ngCMatrix' [package "Matrix"] with 5 slots
##   .. .. ..@ i       : int [1:43367] 13 60 69 78 14 29 98 24 15 29 ...
##   .. .. ..@ p       : int [1:9836] 0 4 7 8 12 16 21 22 27 28 ...
##   .. .. ..@ Dim     : int [1:2] 169 9835
##   .. .. ..@ Dimnames:List of 2
##   .. .. .. ..$ : NULL
##   .. .. .. ..$ : NULL
##   .. .. ..@ factors : list()
##   ..@ itemInfo   :'data.frame':  169 obs. of  3 variables:
##   .. ..$ labels: chr [1:169] "frankfurter" "sausage" "liver loaf" "ham" ...
##   .. ..$ level2: Factor w/ 55 levels "baby food","bags",..: 44 44 44 44 44 44 44 42 42 41 ...
##   .. ..$ level1: Factor w/ 10 levels "canned food",..: 6 6 6 6 6 6 6 6 6 6 ...
##   ..@ itemsetInfo:'data.frame':  0 obs. of  0 variables
```

```
inspect(Groceries)
```

```
##        items
## [1]    {citrus fruit,
##         semi-finished bread,
##         margarine,
##         ready soups}
## [2]    {tropical fruit,
##         yogurt,
##         coffee}
## [3]    {whole milk}
## [4]    {pip fruit,
##         yogurt,
##         cream cheese ,
##         meat spreads}
## [5]    {other vegetables,
##         whole milk,
##         condensed milk,
##         long life bakery product}
## [6]    {whole milk,
##         butter,
##         yogurt,
##         rice,
##         abrasive cleaner}
## [7]    {rolls/buns}
## [8]    {other vegetables,
##         UHT-milk,
##         rolls/buns,
##         bottled beer,
##         liquor (appetizer)}
## [9]    {pot plants}
## [10]   {whole milk,
##         cereals}
## [11]   {tropical fruit,
##         other vegetables,
##         white bread,
##         bottled water,
##         chocolate}
## [12]   {citrus fruit,
##         tropical fruit,
##         whole milk,
##         butter,
##         curd,
##         yogurt,
##         flour,
##         bottled water,
##         dishes}
## [13]   {beef}
## [14]   {frankfurter,
##         rolls/buns,
##         soda}
## [15]   {chicken,
##         tropical fruit}
## [16]   {butter,
##         sugar,
##         fruit/vegetable juice,
##         newspapers}
## [17]   {fruit/vegetable juice}
## [18]   {packaged fruit/vegetables}
## [19]   {chocolate}
## [20]   {specialty bar}
## [21]   {other vegetables}
## [22]   {butter milk,
##         pastry}
```

```
library(arulesViz)
gr_rules=apriori(Groceries,parameter = list(support=0.001,conf=0.8))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.8    0.1    1 none FALSE            TRUE       5   0.001      1
##  maxlen target  ext
##      10  rules TRUE
##
## 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 [0.00s].
## sorting and recoding items ... [157 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 5 6 done [0.01s].
## writing ... [410 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
inspect(gr_rules[1:10])
```

```
##      lhs                           rhs                    support confidence    coverage      lift cou
## [1]  {liquor,
##       red/blush wine}          => {bottled beer}     0.001931876  0.9047619 0.002135231 11.235269
## [2]  {curd,
##       cereals}                 => {whole milk}       0.001016777  0.9090909 0.001118454  3.557863
## [3]  {yogurt,
##       cereals}                 => {whole milk}       0.001728521  0.8095238 0.002135231  3.168192
## [4]  {butter,
##       jam}                     => {whole milk}       0.001016777  0.8333333 0.001220132  3.261374
## [5]  {soups,
##       bottled beer}            => {whole milk}       0.001118454  0.9166667 0.001220132  3.587512
## [6]  {napkins,
##       house keeping products} => {whole milk}       0.001321810  0.8125000 0.001626843  3.179840
## [7]  {whipped/sour cream,
##       house keeping products} => {whole milk}       0.001220132  0.9230769 0.001321810  3.612599
## [8]  {pastry,
##       sweet spreads}           => {whole milk}       0.001016777  0.9090909 0.001118454  3.557863
## [9]  {turkey,
##       curd}                    => {other vegetables} 0.001220132  0.8000000 0.001525165  4.134524
## [10] {rice,
##       sugar}                   => {whole milk}       0.001220132  1.0000000 0.001220132  3.913649
```

```
gr_rules=sort(gr_rules,by='support',decreasing=T)
inspect(gr_rules[1:10])
```

```
##      lhs                          rhs                    support confidence    coverage     lift count
## [1]  {citrus fruit,
##       tropical fruit,
##       root vegetables,
##       whole milk}              => {other vegetables} 0.003152008  0.8857143 0.003558719 4.577509    31
## [2]  {other vegetables,
##       curd,
##       domestic eggs}           => {whole milk}       0.002846975  0.8235294 0.003457041 3.223005    28
## [3]  {hamburger meat,
##       curd}                    => {whole milk}       0.002541942  0.8064516 0.003152008 3.156169    25
## [4]  {herbs,
##       rolls/buns}              => {whole milk}       0.002440264  0.8000000 0.003050330 3.130919    24
## [5]  {tropical fruit,
##       herbs}                   => {whole milk}       0.002338587  0.8214286 0.002846975 3.214783    23
## [6]  {citrus fruit,
##       root vegetables,
##       other vegetables,
##       yogurt}                  => {whole milk}       0.002338587  0.8214286 0.002846975 3.214783    23
## [7]  {pork,
##       other vegetables,
##       butter}                  => {whole milk}       0.002236909  0.8461538 0.002643620 3.311549    22
## [8]  {tropical fruit,
##       root vegetables,
##       yogurt,
##       rolls/buns}              => {whole milk}       0.002236909  0.8148148 0.002745297 3.188899    22
## [9]  {tropical fruit,
##       grapes,
##       whole milk}              => {other vegetables} 0.002033554  0.8000000 0.002541942 4.134524    20
## [10] {root vegetables,
##       other vegetables,
##       yogurt,
##       fruit/vegetable juice} => {whole milk}       0.002033554  0.8333333 0.002440264 3.261374    20
```

```
gr_rules
```

```
## set of 410 rules
```

```
redundant_rules=is.redundant(gr_rules)
redundant_rules
```

```
##   [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [37] FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [85] FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##  [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [121] FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [205] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [217] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [229] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [241] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [253] FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE  TRUE  TRUE FALSE FALSE
## [265] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [277] FALSE FALSE FALSE FALSE FALSE  TRUE  TRUE FALSE FALSE FALSE FALSE FALSE
## [289] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [301] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [313] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [325] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [337] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [349] FALSE FALSE FALSE FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE  TRUE
## [361] FALSE FALSE FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE
## [373] FALSE FALSE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## [385] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  TRUE FALSE  TRUE
## [397] FALSE FALSE FALSE FALSE FALSE FALSE FALSE  TRUE  TRUE FALSE  TRUE  TRUE
## [409] FALSE FALSE
```

```
summary(redundant_rules)
```

```
##    Mode   FALSE    TRUE
## logical     392      18
```

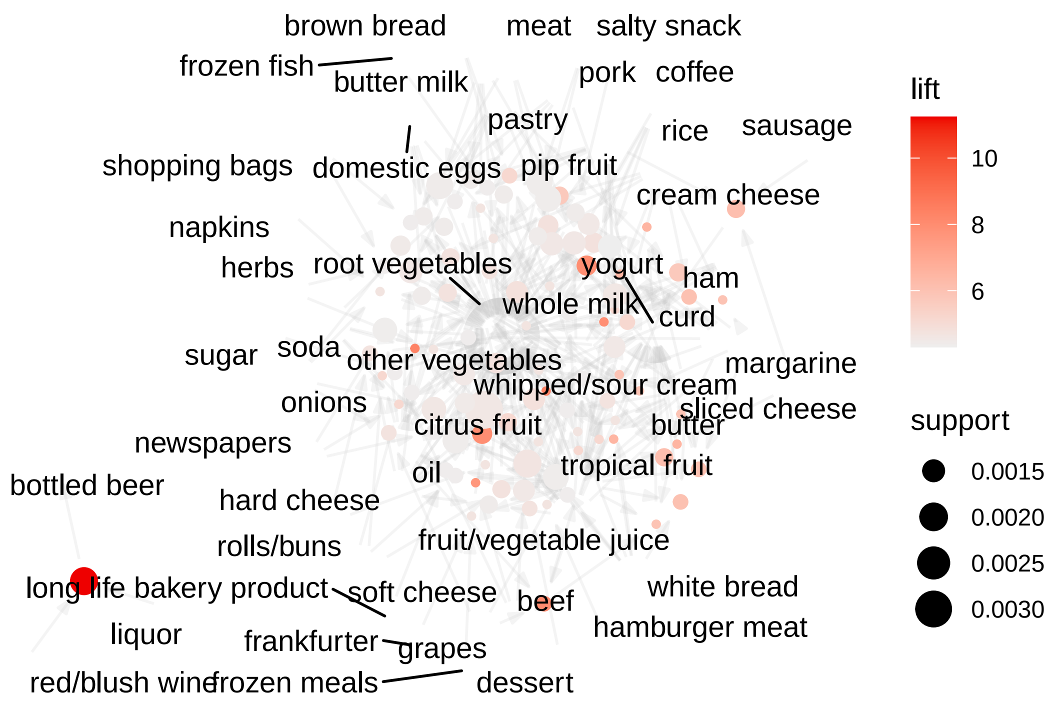
```
gr_rules=gr_rules[!redundant_rules]
inspect(gr_rules[1:5])
```

```
##     lhs                      rhs                     support confidence    coverage     lift count
## [1] {citrus fruit,
##      tropical fruit,
##      root vegetables,
##      whole milk}          => {other vegetables} 0.003152008  0.8857143 0.003558719 4.577509    31
## [2] {other vegetables,
##      curd,
##      domestic eggs}       => {whole milk}       0.002846975  0.8235294 0.003457041 3.223005    28
## [3] {hamburger meat,
##      curd}                => {whole milk}       0.002541942  0.8064516 0.003152008 3.156169    25
## [4] {herbs,
##      rolls/buns}          => {whole milk}       0.002440264  0.8000000 0.003050330 3.130919    24
## [5] {tropical fruit,
##      herbs}               => {whole milk}       0.002338587  0.8214286 0.002846975 3.214783    23
```

```
plot(gr_rules,method='graph')
```

```
## Warning: Too many rules supplied. Only plotting the best 100 using 'lift'
## (change control parameter max if needed).

## Warning: ggrepel: 1 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```



```r
#gr_rules=sort(gr_rules,by='support',decreasing=T,lhs)
plot(gr_rules,method='graph',interactive = T)
```

```
## Warning in plot.rules(gr_rules, method = "graph", interactive = T): The
## parameter interactive is deprecated. Use engine='interactive' instead.
```

```
## Warning: Too many rules supplied. Only plotting the best 100 using 'lift'
## (change control parameter max if needed).
```