In [18]:
```
#Libraries
library(arules)
library(arulesViz)
library(tidyverse)
```

In [22]:
```
# Reading transactions
df <- read.transactions("../Dataset/Groceries.csv", sep=",")
```

In [23]:
```
# View Transaction
inspect(df[1:10])
```

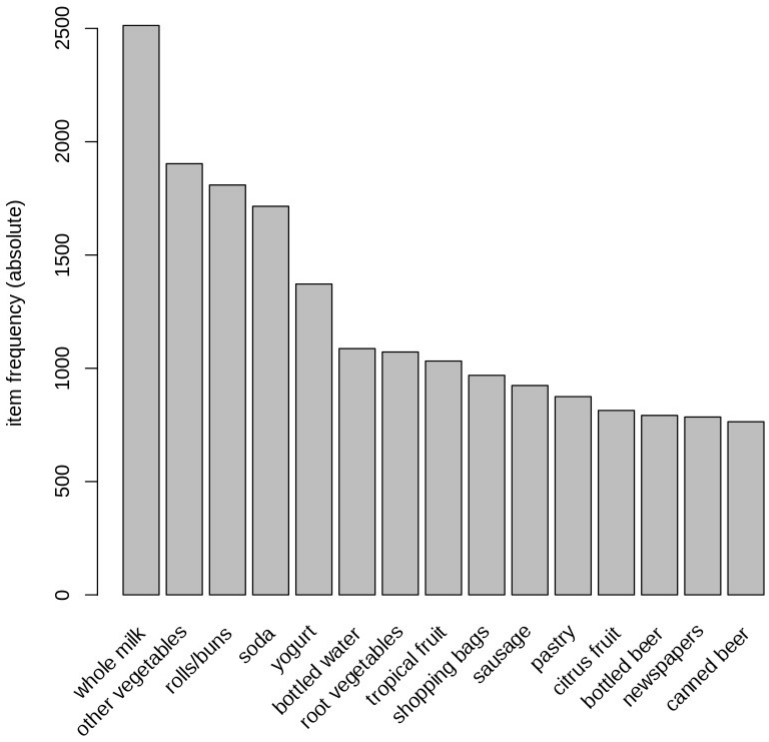
```
     items
[1]  {citrus fruit,
      margarine,
      ready soups,
      semi-finished bread}
[2]  {coffee,
      tropical fruit,
      yogurt}
[3]  {whole milk}
[4]  {cream cheese,
      meat spreads,
      pip fruit,
      yogurt}
[5]  {condensed milk,
      long life bakery product,
      other vegetables,
      whole milk}
[6]  {abrasive cleaner,
      butter,
      rice,
      whole milk,
      yogurt}
[7]  {rolls/buns}
[8]  {bottled beer,
      liquor (appetizer),
      other vegetables,
      rolls/buns,
      UHT-milk}
[9]  {pot plants}
[10] {cereals,
      whole milk}
```

In [24]:
```
# summarize whole data set
summary(df)
```

```
transactions as itemMatrix in sparse format with
 9835 rows (elements/itemsets/transactions) and
 169 columns (items) and a density of 0.026

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   1
6
2159 1643 1299 1005  855  645  545  438  350  246  182  117   78   77   55   4
6
```

```
  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       2       3       4       6      32

includes extended item information - examples:
            labels
1 abrasive cleaner
2 artif. sweetener
3   baby cosmetics
```

In [25]:
```
# Frequency plot of top 15 items
itemFrequencyPlot(df, topN=15, type="absolute")
```



In [40]:
```
# Get the rules
rules = apriori(df, parameter = list(supp = 0.003, conf = 0.3))
```

```
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen
        0.3    0.1    1 none FALSE            TRUE       5   0.003      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: 29

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[169 item(s), 9835 transaction(s)] done [0.00s].
sorting and recoding items ... [136 item(s)] done [0.00s].
```

```
creating transaction tree ... done [0.01s].
checking subsets of size 1 2 3 4 5 done [0.01s].
writing ... [1361 rule(s)] done [0.00s].
```

In [41]:
```
# Summarize the rules
summary(rules)
```

```
set of 1361 rules

rule length distribution (lhs + rhs):sizes
  2   3   4   5
148 912 291  10

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
    2.0     3.0     3.0     3.1     3.0     5.0

summary of quality measures:
    support         confidence      coverage          lift           count
 Min.   :0.003   Min.   :0.30   Min.   :0.004   Min.   : 1.2   Min.   : 30
 1st Qu.:0.003   1st Qu.:0.36   1st Qu.:0.008   1st Qu.: 1.9   1st Qu.: 34
 Median :0.004   Median :0.43   Median :0.010   Median : 2.2   Median : 41
 Mean   :0.006   Mean   :0.45   Mean   :0.014   Mean   : 2.4   Mean   : 58
 3rd Qu.:0.006   3rd Qu.:0.52   3rd Qu.:0.015   3rd Qu.: 2.7   3rd Qu.: 59
 Max.   :0.075   Max.   :0.89   Max.   :0.193   Max.   :11.4   Max.   :736

mining info:
 data ntransactions support confidence
   df          9835   0.003        0.3
```

In [42]:
```
# Print the rules
inspect(rules[1:10])
```

```
     lhs                      rhs                support confidence coverage
[1]  {liquor}                 => {bottled beer}     0.0047  0.42       0.0111
[2]  {cereals}                => {whole milk}       0.0037  0.64       0.0057
[3]  {candles}                => {whole milk}       0.0031  0.34       0.0089
[4]  {soups}                  => {other vegetables} 0.0032  0.46       0.0068
[5]  {Instant food products}  => {hamburger meat}   0.0031  0.38       0.0080
[6]  {Instant food products}  => {whole milk}       0.0031  0.38       0.0080
[7]  {specialty cheese}       => {other vegetables} 0.0043  0.50       0.0085
[8]  {specialty cheese}       => {whole milk}       0.0038  0.44       0.0085
[9]  {chocolate marshmallow}  => {whole milk}       0.0032  0.35       0.0090
[10] {flower (seeds)}         => {other vegetables} 0.0038  0.36       0.0104
     lift count
[1]   5.2 46
[2]   2.5 36
[3]   1.3 30
[4]   2.4 31
[5]  11.4 30
[6]   1.5 30
[7]   2.6 42
[8]   1.7 37
[9]   1.4 31
[10]  1.9 37
```

In [43]:
```
# Sort the rules by confidence
rules = sort(rules, by = "confidence")
options(digits = 2)
```

In [44]:
```
inspect(rules[1:10])
```

```
     lhs                    rhs                support confidence coverage li
ft count
[1]  {citrus fruit,
```

```
      root vegetables,
      tropical fruit,
      whole milk}          => {other vegetables}  0.0032       0.89   0.0036
4.6    31
[2]  {butter,
      root vegetables,
      yogurt}              => {whole milk}        0.0031       0.79   0.0039
3.1    30
[3]  {citrus fruit,
      root vegetables,
      tropical fruit}      => {other vegetables}  0.0045       0.79   0.0057
4.1    44
[4]  {brown bread,
      other vegetables,
      root vegetables}     => {whole milk}        0.0032       0.78   0.0041
3.0    31
[5]  {butter,
      onions}              => {whole milk}        0.0031       0.75   0.0041
2.9    30
[6]  {curd,
      tropical fruit,
      yogurt}              => {whole milk}        0.0040       0.75   0.0053
2.9    39
[7]  {curd,
      domestic eggs}       => {whole milk}        0.0048       0.73   0.0065
2.9    47
[8]  {butter,
      tropical fruit,
      yogurt}              => {whole milk}        0.0034       0.73   0.0046
2.9    33
[9]  {root vegetables,
      tropical fruit,
      whipped/sour cream} => {other vegetables}  0.0034       0.73   0.0046
3.8    33
[10] {butter,
      curd}                => {whole milk}        0.0049       0.72   0.0068
2.8    48
```

In [46]:
```
# Inspect the redundant rules
print("Showing first 10 rules")
rules[is.redundant(rules)]
inspect(rules[is.redundant(rules)][1:10])
rules = rules[!is.redundant(rules)]
```

```
[1] "Showing first 10 rules"
set of 85 rules
     lhs                     rhs                support confidence coverage li
ft count
[1]  {butter,
      other vegetables,
      root vegetables}     => {whole milk}        0.0042       0.63   0.0066
2.5    41
[2]  {butter,
      other vegetables,
      tropical fruit}      => {whole milk}        0.0034       0.61   0.0055
2.4    33
[3]  {curd,
      other vegetables,
      tropical fruit}      => {whole milk}        0.0032       0.60   0.0053
2.3    31
[4]  {other vegetables,
      tropical fruit,
      whipped/sour cream} => {whole milk}        0.0045       0.57   0.0078
2.2    44
[5]  {tropical fruit,
```

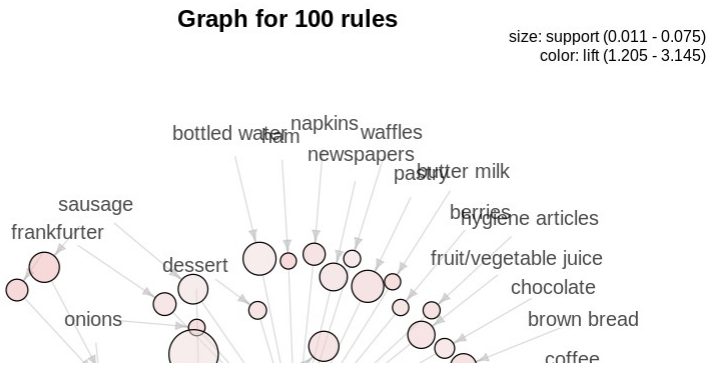
```
      whipped/sour cream,
      whole milk}          => {other vegetables}  0.0045       0.56   0.0079
2.9    44
[6]  {butter,
      tropical fruit,
      whole milk}          => {other vegetables}  0.0034       0.54   0.0062
2.8    33
[7]  {pastry,
      root vegetables,
      whole milk}          => {other vegetables}  0.0031       0.54   0.0057
2.8    30
[8]  {other vegetables,
      pastry,
      root vegetables}     => {whole milk}        0.0031       0.52   0.0059
2.0    30
[9]  {bottled water,
      tropical fruit,
      yogurt}              => {whole milk}        0.0037       0.51   0.0071
2.0    36
[10] {other vegetables,
      sausage,
      tropical fruit}      => {whole milk}        0.0031       0.51   0.0060
2.0    30
```

In [12]:
```
inspect(rules[1:10])
```

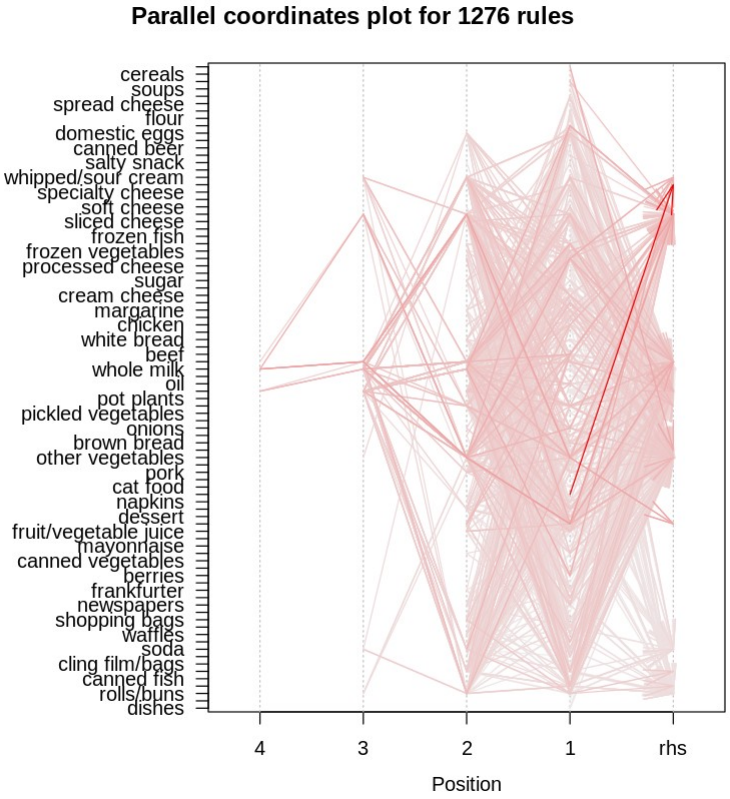
```
     lhs                     rhs                support confidence coverage li
ft count
[1]  {citrus fruit,
      root vegetables,
      tropical fruit,
      whole milk}         => {other vegetables}  0.0032       0.89   0.0036
4.6    31
[2]  {butter,
      root vegetables,
      yogurt}             => {whole milk}        0.0031       0.79   0.0039
3.1    30
[3]  {citrus fruit,
      root vegetables,
      tropical fruit}     => {other vegetables}  0.0045       0.79   0.0057
4.1    44
[4]  {brown bread,
      other vegetables,
      root vegetables}    => {whole milk}        0.0032       0.78   0.0041
3.0    31
[5]  {butter,
      onions}             => {whole milk}        0.0031       0.75   0.0041
2.9    30
[6]  {curd,
      tropical fruit,
      yogurt}             => {whole milk}        0.0040       0.75   0.0053
2.9    39
[7]  {curd,
      domestic eggs}      => {whole milk}        0.0048       0.73   0.0065
2.9    47
[8]  {butter,
      tropical fruit,
      yogurt}             => {whole milk}        0.0034       0.73   0.0046
2.9    33
[9]  {root vegetables,
      tropical fruit,
      whipped/sour cream} => {other vegetables}  0.0034       0.73   0.0046
3.8    33
[10] {butter,
      curd}               => {whole milk}        0.0049       0.72   0.0068
2.8    48
```

In [14]:
```
# Plot the graphs for the rules
plot(rules, method = "graph")
```

```
Warning message:
“plot: Too many rules supplied. Only plotting the best 100 rules using ‘suppor
t’ (change control parameter max if needed)”
```



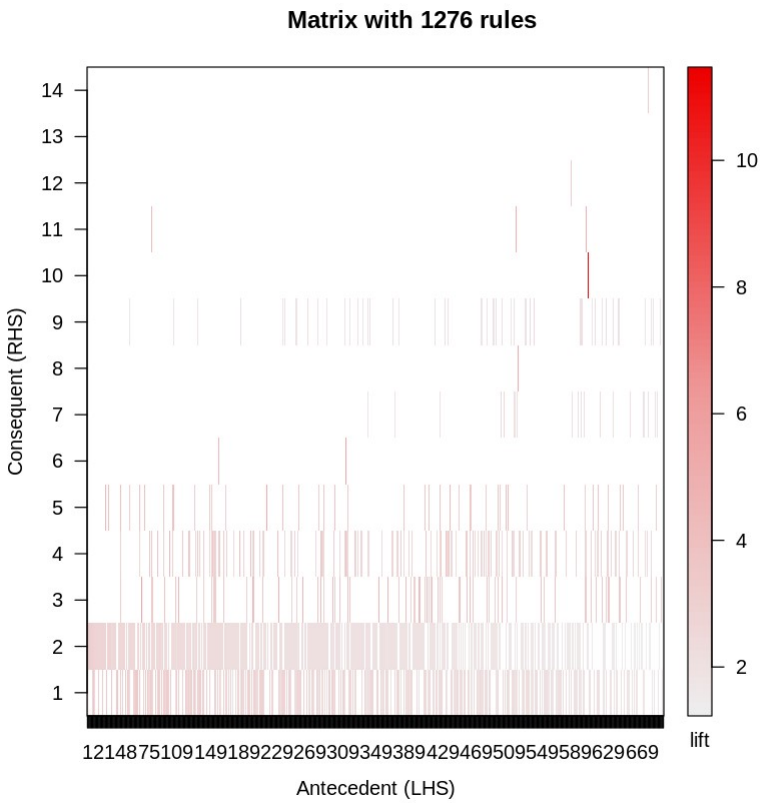
In [15]:
```
plot(rules, method = "paracoord")
```



In [53]:
```
plot(rules, method = "matrix", control = list(reorder = "none"))
```

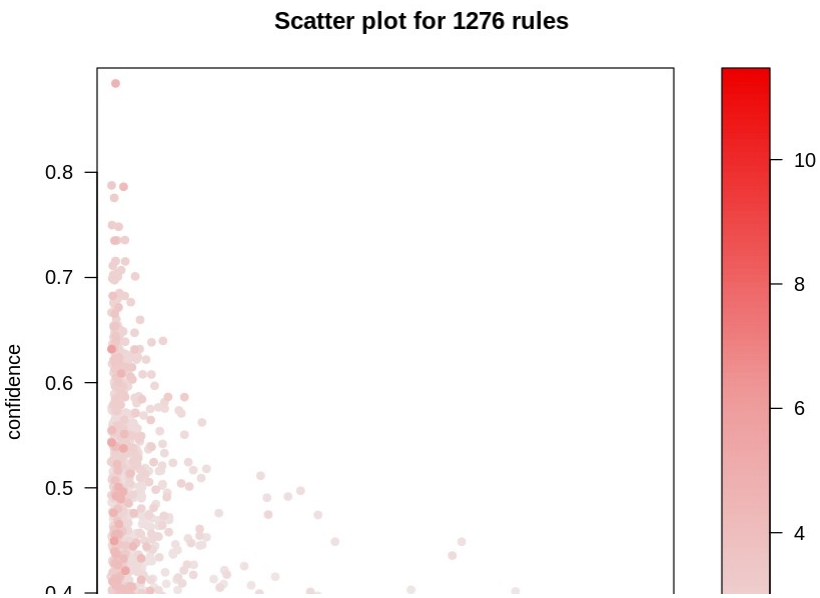
```
Itemsets in Antecedent (LHS)
  [1] "{citrus fruit,root vegetables,tropical fruit,whole milk}"
  [2] "{butter,root vegetables,yogurt}"
  [3] "{citrus fruit,root vegetables,tropical fruit}"
  [4] "{brown bread,other vegetables,root vegetables}"
  [5] "{butter,onions}"
  [6] "{curd,tropical fruit,yogurt}"
  [7] "{curd,domestic eggs}"
  [8] "{butter,tropical fruit,yogurt}"
  [9] "{root vegetables,tropical fruit,whipped/sour cream}"
 [10] "{butter,curd}"
 [11] "{domestic eggs,sugar}"
 [12] "{other vegetables,root vegetables,tropical fruit,yogurt}"
 [13] "{baking powder,yogurt}"
 [14] "{tropical fruit,whipped/sour cream,yogurt}"
 [15] "{citrus fruit,other vegetables,root vegetables,tropical fruit}"
 [16] "{butter,pork}"
```

```
 [17] "{butter,coffee}"
 [18] "{domestic eggs,other vegetables,whipped/sour cream}"
 [19] "{root vegetables,tropical fruit,yogurt}"
 [20] "{butter,hamburger meat}"
 [21] "{butter,hygiene articles}"
 [22] "{butter,other vegetables,whipped/sour cream}"
 [23] "{butter,other vegetables,yogurt}"
 [24] "{onions,root vegetables,whole milk}"
 [25] "{cream cheese,domestic eggs}"
 [26] "{other vegetables,pip fruit,root vegetables}"
 [27] "{sausage,tropical fruit,yogurt}"
 [28] "{frozen vegetables,other vegetables,yogurt}"
 [29] "{pip fruit,root vegetables,yogurt}"
 [30] "{root vegetables,sliced cheese}"
 .
 .
 ...Upto 695 Antecedents

Itemsets in Consequent (RHS)
 [1] "{other vegetables}"   "{whole milk}"         "{root vegetables}"
 [4] "{yogurt}"             "{tropical fruit}"     "{citrus fruit}"
 [7] "{soda}"               "{bottled beer}"       "{rolls/buns}"
[10] "{hamburger meat}"     "{whipped/sour cream}" "{bottled water}"
[13] "{pip fruit}"          "{sausage}"
```



In [17]:
```
plot(rules)
```

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


Scatter plot for 1276 rules
confidence
0.8
0.7
0.6
0.5
10
8
6
4