Itemset Mining in R

We'll start with the example from the slides

```
d <- matrix(nrow=6, ncol=5, byrow=TRUE,
            data= c(1,1,0,1,1,
                    0,1,1,0,1,
                    1,1,0,1,1,
                    1,1,1,0,1,
                    1,1,1,1,1,
                    0,1,1,1,0))
colnames(d) <- c('A', 'B', 'C', 'D', 'E')
d
```

```
##      A B C D E
## [1,] 1 1 0 1 1
## [2,] 0 1 1 0 1
## [3,] 1 1 0 1 1
## [4,] 1 1 1 0 1
## [5,] 1 1 1 1 1
## [6,] 0 1 1 1 0
```

Load arules

```
library(arules)
```

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

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

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

Build a transaction object from the matrix

```
t <- as(d, "transactions")
t
```

```
## transactions in sparse format with
##  6 transactions (rows) and
##  5 items (columns)
```

```
inspect(t)
```

```
##     items
## [1] {A,B,D,E}
## [2] {B,C,E}
## [3] {A,B,D,E}
## [4] {A,B,C,E}
## [5] {A,B,C,D,E}
## [6] {B,C,D}
```

Note: When arules (or rCBA) talks about "support" they mean what the book calls "rsupport". supp / ntransactions

use apriori to extract itemsets with minsup = 3 (rsup=0.5)

```
frequent_itemsets <- apriori(d, parameter=list(support=0.5, target="frequent itemsets"))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE            TRUE       5     0.5      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: 3
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[5 item(s), 6 transaction(s)] done [0.00s].
## sorting and recoding items ... [5 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 4 done [0.00s].
## writing ... [19 set(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
inspect(frequent_itemsets)
```

```
##      items     support   count
## [1]  {C}       0.6666667 4
## [2]  {D}       0.6666667 4
## [3]  {A}       0.6666667 4
## [4]  {E}       0.8333333 5
## [5]  {B}       1.0000000 6
## [6]  {C,E}     0.5000000 3
## [7]  {B,C}     0.6666667 4
## [8]  {A,D}     0.5000000 3
## [9]  {D,E}     0.5000000 3
## [10] {B,D}     0.6666667 4
## [11] {A,E}     0.6666667 4
## [12] {A,B}     0.6666667 4
## [13] {B,E}     0.8333333 5
## [14] {B,C,E}   0.5000000 3
## [15] {A,D,E}   0.5000000 3
## [16] {A,B,D}   0.5000000 3
## [17] {B,D,E}   0.5000000 3
## [18] {A,B,E}   0.6666667 4
## [19] {A,B,D,E} 0.5000000 3
```

Sort our itemsets by support

```
sorted_itemsets <- sort(frequent_itemsets, by="support")
inspect(sorted_itemsets)
```

```
##      items     support   count
## [1]  {B}       1.0000000 6
## [2]  {E}       0.8333333 5
## [3]  {B,E}     0.8333333 5
## [4]  {C}       0.6666667 4
## [5]  {D}       0.6666667 4
## [6]  {A}       0.6666667 4
## [7]  {B,C}     0.6666667 4
## [8]  {B,D}     0.6666667 4
## [9]  {A,E}     0.6666667 4
## [10] {A,B}     0.6666667 4
## [11] {A,B,E}   0.6666667 4
## [12] {C,E}     0.5000000 3
## [13] {A,D}     0.5000000 3
## [14] {D,E}     0.5000000 3
## [15] {B,C,E}   0.5000000 3
## [16] {A,D,E}   0.5000000 3
## [17] {A,B,D}   0.5000000 3
## [18] {B,D,E}   0.5000000 3
## [19] {A,B,D,E} 0.5000000 3
```

Do the same experiment using eclat

```
frequent_itemsets2 <- eclat(d, parameter=list(support=0.5))
```

```
## Eclat
##
## parameter specification:
##  tidLists support minlen maxlen            target   ext
##     FALSE     0.5      1     10 frequent itemsets FALSE
##
## algorithmic control:
##  sparse sort verbose
##       7   -2    TRUE
##
## Absolute minimum support count: 3
##
## create itemset ...
## set transactions ...[5 item(s), 6 transaction(s)] done [0.00s].
## sorting and recoding items ... [5 item(s)] done [0.00s].
## creating bit matrix ... [5 row(s), 6 column(s)] done [0.00s].
## writing  ... [19 set(s)] done [0.00s].
## Creating S4 object  ... done [0.00s].
```

```
sorted_itemsets2 <- sort(frequent_itemsets2, by="support")
inspect(sorted_itemsets2)
```

```
##      items     support   count
## [1]  {B}       1.0000000 6
## [2]  {B,E}     0.8333333 5
## [3]  {E}       0.8333333 5
## [4]  {B,C}     0.6666667 4
## [5]  {B,D}     0.6666667 4
## [6]  {A,B,E}   0.6666667 4
## [7]  {A,B}     0.6666667 4
## [8]  {A,E}     0.6666667 4
```

```
## [9]  {A}       0.6666667 4
## [10] {D}       0.6666667 4
## [11] {C}       0.6666667 4
## [12] {B,C,E}   0.5000000 3
## [13] {C,E}     0.5000000 3
## [14] {A,B,D,E} 0.5000000 3
## [15] {A,B,D}   0.5000000 3
## [16] {A,D,E}   0.5000000 3
## [17] {B,D,E}   0.5000000 3
## [18] {D,E}     0.5000000 3
## [19] {A,D}     0.5000000 3
```

Figure out he association rules for the frequent itemset

```
rules <- ruleInduction(frequent_itemsets)
inspect(rules)
```

```
##      lhs        rhs support   confidence lift
## [1]  {C}     => {B} 0.6666667 1.0000000  1.0
## [2]  {D}     => {B} 0.6666667 1.0000000  1.0
## [3]  {E}     => {A} 0.6666667 0.8000000  1.2
## [4]  {A}     => {E} 0.6666667 1.0000000  1.2
## [5]  {A}     => {B} 0.6666667 1.0000000  1.0
## [6]  {E}     => {B} 0.8333333 1.0000000  1.0
## [7]  {B}     => {E} 0.8333333 0.8333333  1.0
## [8]  {C,E}   => {B} 0.5000000 1.0000000  1.0
## [9]  {D,E}   => {A} 0.5000000 1.0000000  1.5
## [10] {A,D}   => {E} 0.5000000 1.0000000  1.2
## [11] {A,D}   => {B} 0.5000000 1.0000000  1.0
## [12] {D,E}   => {B} 0.5000000 1.0000000  1.0
## [13] {B,E}   => {A} 0.6666667 0.8000000  1.2
## [14] {A,E}   => {B} 0.6666667 1.0000000  1.0
## [15] {A,B}   => {E} 0.6666667 1.0000000  1.2
## [16] {B,D,E} => {A} 0.5000000 1.0000000  1.5
## [17] {A,D,E} => {B} 0.5000000 1.0000000  1.0
## [18] {A,B,D} => {E} 0.5000000 1.0000000  1.2
```

Rerun the above experiment using fpgrowth (which appears to be broken)

```
#library(rCBA)
#rules <- fpgrowth(train=t, support = 0.5, #confidence = 0.7)
#inspect(rules)
```

Let's at the grocery store example

```
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   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  level2           level1
## 1 frankfurter sausage meat and sausage
## 2     sausage sausage meat and sausage
## 3  liver loaf sausage meat and sausage
```

Use apriori to extract frequent itemsets with minsup 25%

```
grocery_itemsets <- apriori(Groceries, parameter=list(support=0.01, target="frequent itemsets", minlen=
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE            TRUE       5    0.01      2
##  maxlen            target   ext
##      20 frequent itemsets FALSE
##
## 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 [0.01s].
## sorting and recoding items ... [88 item(s)] done [0.00s].
## creating transaction tree ... done [0.01s].
## checking subsets of size 1 2 3 4 done [0.00s].
## writing ... [245 set(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
inspect(sort(grocery_itemsets, by="support"))
```

```
##       items                                              support    count
## [1]   {other vegetables,whole milk}                      0.07483477 736
## [2]   {whole milk,rolls/buns}                            0.05663447 557
## [3]   {whole milk,yogurt}                                0.05602440 551
## [4]   {root vegetables,whole milk}                       0.04890696 481
## [5]   {root vegetables,other vegetables}                 0.04738180 466
## [6]   {other vegetables,yogurt}                          0.04341637 427
## [7]   {other vegetables,rolls/buns}                      0.04260295 419
## [8]   {tropical fruit,whole milk}                        0.04229792 416
## [9]   {whole milk,soda}                                  0.04006101 394
## [10]  {rolls/buns,soda}                                  0.03833249 377
```

```
## [11]  {tropical fruit,other vegetables}                  0.03589222 353
## [12]  {whole milk,bottled water}                         0.03436706 338
## [13]  {yogurt,rolls/buns}                                0.03436706 338
## [14]  {whole milk,pastry}                                0.03324860 327
## [15]  {other vegetables,soda}                            0.03274021 322
## [16]  {whole milk,whipped/sour cream}                    0.03223183 317
## [17]  {sausage,rolls/buns}                               0.03060498 301
## [18]  {citrus fruit,whole milk}                          0.03050330 300
## [19]  {pip fruit,whole milk}                             0.03009659 296
## [20]  {whole milk,domestic eggs}                         0.02999492 295
## [21]  {sausage,whole milk}                               0.02989324 294
## [22]  {tropical fruit,yogurt}                            0.02928317 288
## [23]  {bottled water,soda}                               0.02897814 285
## [24]  {other vegetables,whipped/sour cream}              0.02887646 284
## [25]  {citrus fruit,other vegetables}                    0.02887646 284
## [26]  {whole milk,butter}                                0.02755465 271
## [27]  {whole milk,newspapers}                            0.02735130 269
## [28]  {yogurt,soda}                                      0.02735130 269
## [29]  {sausage,other vegetables}                         0.02694459 265
## [30]  {whole milk,fruit/vegetable juice}                 0.02663955 262
## [31]  {whole milk,curd}                                  0.02613116 257
## [32]  {pip fruit,other vegetables}                       0.02613116 257
## [33]  {root vegetables,yogurt}                           0.02582613 254
## [34]  {whole milk,brown bread}                           0.02521607 248
## [35]  {other vegetables,bottled water}                   0.02480935 244
## [36]  {soda,shopping bags}                               0.02460600 242
## [37]  {tropical fruit,rolls/buns}                        0.02460600 242
## [38]  {whole milk,shopping bags}                         0.02450432 241
## [39]  {sausage,soda}                                     0.02430097 239
## [40]  {root vegetables,rolls/buns}                       0.02430097 239
## [41]  {whole milk,margarine}                             0.02419929 238
## [42]  {rolls/buns,bottled water}                         0.02419929 238
## [43]  {other vegetables,shopping bags}                   0.02318251 228
## [44]  {root vegetables,other vegetables,whole milk}      0.02318251 228
## [45]  {yogurt,bottled water}                             0.02297916 226
## [46]  {other vegetables,pastry}                          0.02257245 222
## [47]  {other vegetables,domestic eggs}                   0.02226741 219
## [48]  {other vegetables,whole milk,yogurt}               0.02226741 219
## [49]  {pork,whole milk}                                  0.02216573 218
## [50]  {pork,other vegetables}                            0.02165735 213
## [51]  {citrus fruit,yogurt}                              0.02165735 213
## [52]  {beef,whole milk}                                  0.02125064 209
## [53]  {other vegetables,fruit/vegetable juice}           0.02104728 207
## [54]  {pastry,soda}                                      0.02104728 207
## [55]  {tropical fruit,root vegetables}                   0.02104728 207
## [56]  {rolls/buns,pastry}                                0.02094560 206
## [57]  {tropical fruit,soda}                              0.02084392 205
## [58]  {yogurt,whipped/sour cream}                        0.02074225 204
## [59]  {frankfurter,whole milk}                           0.02053889 202
## [60]  {whole milk,frozen vegetables}                     0.02043721 201
## [61]  {whole milk,bottled beer}                          0.02043721 201
## [62]  {tropical fruit,pip fruit}                         0.02043721 201
## [63]  {other vegetables,butter}                          0.02003050 197
## [64]  {citrus fruit,tropical fruit}                      0.01992883 196
```

```
## [65]  {beef,other vegetables}                                 0.01972547 194
## [66]  {whole milk,napkins}                                    0.01972547 194
## [67]  {other vegetables,margarine}                            0.01972547 194
## [68]  {rolls/buns,newspapers}                                 0.01972547 194
## [69]  {sausage,yogurt}                                        0.01962379 193
## [70]  {rolls/buns,shopping bags}                              0.01952211 192
## [71]  {other vegetables,newspapers}                           0.01931876 190
## [72]  {frankfurter,rolls/buns}                                0.01921708 189
## [73]  {whole milk,coffee}                                     0.01870869 184
## [74]  {other vegetables,brown bread}                          0.01870869 184
## [75]  {yogurt,fruit/vegetable juice}                          0.01870869 184
## [76]  {root vegetables,soda}                                  0.01860702 183
## [77]  {tropical fruit,bottled water}                          0.01850534 182
## [78]  {soda,fruit/vegetable juice}                            0.01840366 181
## [79]  {pip fruit,yogurt}                                      0.01799695 177
## [80]  {chicken,other vegetables}                              0.01789527 176
## [81]  {other vegetables,whole milk,rolls/buns}                0.01789527 176
## [82]  {other vegetables,frozen vegetables}                    0.01779359 175
## [83]  {yogurt,pastry}                                         0.01769192 174
## [84]  {citrus fruit,root vegetables}                          0.01769192 174
## [85]  {chicken,whole milk}                                    0.01759024 173
## [86]  {beef,root vegetables}                                  0.01738688 171
## [87]  {curd,yogurt}                                           0.01728521 170
## [88]  {other vegetables,curd}                                 0.01718353 169
## [89]  {whole milk,white bread}                                0.01708185 168
## [90]  {root vegetables,whipped/sour cream}                    0.01708185 168
## [91]  {tropical fruit,other vegetables,whole milk}            0.01708185 168
## [92]  {soda,bottled beer}                                     0.01698017 167
## [93]  {citrus fruit,rolls/buns}                               0.01677682 165
## [94]  {whole milk,chocolate}                                  0.01667514 164
## [95]  {whole milk,cream cheese }                              0.01647178 162
## [96]  {frankfurter,other vegetables}                          0.01647178 162
## [97]  {other vegetables,bottled beer}                         0.01616675 159
## [98]  {bottled water,bottled beer}                            0.01576004 155
## [99]  {domestic eggs,rolls/buns}                              0.01565836 154
## [100] {sausage,shopping bags}                                 0.01565836 154
## [101] {root vegetables,bottled water}                         0.01565836 154
## [102] {pip fruit,root vegetables}                             0.01555669 153
## [103] {whole milk,yogurt,rolls/buns}                          0.01555669 153
## [104] {yogurt,newspapers}                                     0.01535333 151
## [105] {yogurt,shopping bags}                                  0.01525165 150
## [106] {tropical fruit,whole milk,yogurt}                      0.01514997 149
## [107] {whole milk,sugar}                                      0.01504830 148
## [108] {sausage,root vegetables}                               0.01494662 147
## [109] {hamburger meat,whole milk}                             0.01474326 145
## [110] {rolls/buns,margarine}                                  0.01474326 145
## [111] {butter,yogurt}                                         0.01464159 144
## [112] {soda,newspapers}                                       0.01464159 144
## [113] {whipped/sour cream,rolls/buns}                         0.01464159 144
## [114] {other vegetables,whole milk,whipped/sour cream}        0.01464159 144
## [115] {yogurt,brown bread}                                    0.01453991 143
## [116] {rolls/buns,fruit/vegetable juice}                      0.01453991 143
## [117] {root vegetables,whole milk,yogurt}                     0.01453991 143
## [118] {other vegetables,napkins}                              0.01443823 142
```

```
## [119] {root vegetables,domestic eggs}                       0.01433655 141
## [120] {yogurt,domestic eggs}                                0.01433655 141
## [121] {onions,other vegetables}                             0.01423488 140
## [122] {yogurt,margarine}                                    0.01423488 140
## [123] {bottled water,fruit/vegetable juice}                 0.01423488 140
## [124] {pip fruit,rolls/buns}                                0.01392984 137
## [125] {sausage,tropical fruit}                              0.01392984 137
## [126] {other vegetables,whole milk,soda}                    0.01392984 137
## [127] {hamburger meat,other vegetables}                     0.01382816 136
## [128] {soda,canned beer}                                    0.01382816 136
## [129] {tropical fruit,whipped/sour cream}                   0.01382816 136
## [130] {citrus fruit,pip fruit}                              0.01382816 136
## [131] {whole milk,dessert}                                  0.01372649 135
## [132] {other vegetables,cream cheese }                      0.01372649 135
## [133] {other vegetables,white bread}                        0.01372649 135
## [134] {tropical fruit,fruit/vegetable juice}                0.01372649 135
## [135] {beef,rolls/buns}                                     0.01362481 134
## [136] {pork,root vegetables}                                0.01362481 134
## [137] {rolls/buns,bottled beer}                             0.01362481 134
## [138] {whole milk,long life bakery product}                 0.01352313 133
## [139] {soda,chocolate}                                      0.01352313 133
## [140] {citrus fruit,bottled water}                          0.01352313 133
## [141] {tropical fruit,shopping bags}                        0.01352313 133
## [142] {pip fruit,other vegetables,whole milk}               0.01352313 133
## [143] {other vegetables,coffee}                             0.01342145 132
## [144] {butter,rolls/buns}                                   0.01342145 132
## [145] {pip fruit,soda}                                      0.01331978 131
## [146] {tropical fruit,pastry}                               0.01321810 130
## [147] {citrus fruit,other vegetables,whole milk}            0.01301474 128
## [148] {root vegetables,butter}                              0.01291307 127
## [149] {root vegetables,other vegetables,yogurt}             0.01291307 127
## [150] {whole milk,hygiene articles}                         0.01281139 126
## [151] {citrus fruit,soda}                                   0.01281139 126
## [152] {root vegetables,shopping bags}                       0.01281139 126
## [153] {whole milk,waffles}                                  0.01270971 125
## [154] {other vegetables,chocolate}                          0.01270971 125
## [155] {root vegetables,whole milk,rolls/buns}               0.01270971 125
## [156] {brown bread,soda}                                    0.01260803 124
## [157] {rolls/buns,brown bread}                              0.01260803 124
## [158] {sausage,pastry}                                      0.01250635 123
## [159] {yogurt,cream cheese }                                0.01240468 122
## [160] {yogurt,frozen vegetables}                            0.01240468 122
## [161] {domestic eggs,soda}                                  0.01240468 122
## [162] {yogurt,napkins}                                      0.01230300 121
## [163] {other vegetables,whole milk,domestic eggs}           0.01230300 121
## [164] {tropical fruit,root vegetables,other vegetables}     0.01230300 121
## [165] {tropical fruit,other vegetables,yogurt}              0.01230300 121
## [166] {root vegetables,other vegetables,rolls/buns}         0.01220132 120
## [167] {onions,whole milk}                                   0.01209964 119
## [168] {soda,napkins}                                        0.01199797 118
## [169] {root vegetables,fruit/vegetable juice}               0.01199797 118
## [170] {sausage,bottled water}                               0.01199797 118
## [171] {tropical fruit,root vegetables,whole milk}           0.01199797 118
## [172] {pork,soda}                                           0.01189629 117
```

```
## [173] {pastry,shopping bags}                                 0.01189629 117
## [174] {berries,whole milk}                                   0.01179461 116
## [175] {rolls/buns,chocolate}                                 0.01179461 116
## [176] {tropical fruit,newspapers}                            0.01179461 116
## [177] {beef,yogurt}                                          0.01169293 115
## [178] {rolls/buns,napkins}                                   0.01169293 115
## [179] {whole milk,butter milk}                               0.01159126 114
## [180] {other vegetables,dessert}                             0.01159126 114
## [181] {root vegetables,frozen vegetables}                    0.01159126 114
## [182] {whipped/sour cream,soda}                              0.01159126 114
## [183] {ham,whole milk}                                       0.01148958 113
## [184] {root vegetables,newspapers}                           0.01148958 113
## [185] {other vegetables,whole milk,butter}                   0.01148958 113
## [186] {other vegetables,yogurt,rolls/buns}                   0.01148958 113
## [187] {canned beer,shopping bags}                            0.01138790 112
## [188] {tropical fruit,domestic eggs}                         0.01138790 112
## [189] {whole milk,oil}                                       0.01128622 111
## [190] {rolls/buns,canned beer}                               0.01128622 111
## [191] {pork,rolls/buns}                                      0.01128622 111
## [192] {frankfurter,soda}                                     0.01128622 111
## [193] {bottled water,newspapers}                             0.01128622 111
## [194] {sausage,citrus fruit}                                 0.01128622 111
## [195] {whole milk,salty snack}                               0.01118454 110
## [196] {frankfurter,yogurt}                                   0.01118454 110
## [197] {root vegetables,margarine}                            0.01108287 109
## [198] {rolls/buns,coffee}                                    0.01098119 108
## [199] {root vegetables,pastry}                               0.01098119 108
## [200] {bottled water,shopping bags}                          0.01098119 108
## [201] {tropical fruit,whole milk,rolls/buns}                 0.01098119 108
## [202] {chicken,root vegetables}                              0.01087951 107
## [203] {root vegetables,curd}                                 0.01087951 107
## [204] {citrus fruit,whipped/sour cream}                      0.01087951 107
## [205] {whole milk,yogurt,whipped/sour cream}                 0.01087951 107
## [206] {whole milk,sliced cheese}                             0.01077783 106
## [207] {other vegetables,salty snack}                         0.01077783 106
## [208] {other vegetables,sugar}                               0.01077783 106
## [209] {sausage,pip fruit}                                    0.01077783 106
## [210] {other vegetables,whole milk,bottled water}            0.01077783 106
## [211] {other vegetables,long life bakery product}            0.01067616 105
## [212] {sausage,brown bread}                                  0.01067616 105
## [213] {tropical fruit,brown bread}                           0.01067616 105
## [214] {fruit/vegetable juice,shopping bags}                  0.01067616 105
## [215] {pip fruit,pastry}                                     0.01067616 105
## [216] {berries,yogurt}                                       0.01057448 104
## [217] {pip fruit,bottled water}                              0.01057448 104
## [218] {other vegetables,whole milk,pastry}                   0.01057448 104
## [219] {curd,whipped/sour cream}                              0.01047280 103
## [220] {other vegetables,whole milk,fruit/vegetable juice}    0.01047280 103
## [221] {whole milk,yogurt,soda}                               0.01047280 103
## [222] {other vegetables,butter milk}                         0.01037112 102
## [223] {citrus fruit,domestic eggs}                           0.01037112 102
## [224] {citrus fruit,fruit/vegetable juice}                   0.01037112 102
## [225] {citrus fruit,root vegetables,other vegetables}        0.01037112 102
## [226] {berries,other vegetables}                             0.01026945 101
```

```
## [227] {white bread,soda}                                 0.01026945 101
## [228] {tropical fruit,curd}                              0.01026945 101
## [229] {margarine,bottled water}                          0.01026945 101
## [230] {citrus fruit,whole milk,yogurt}                   0.01026945 101
## [231] {frozen vegetables,rolls/buns}                     0.01016777 100
## [232] {frankfurter,root vegetables}                      0.01016777 100
## [233] {root vegetables,brown bread}                      0.01016777 100
## [234] {margarine,soda}                                   0.01016777 100
## [235] {butter,whipped/sour cream}                        0.01016777 100
## [236] {pork,other vegetables,whole milk}                 0.01016777 100
## [237] {other vegetables,yogurt,whipped/sour cream}       0.01016777 100
## [238] {sausage,other vegetables,whole milk}              0.01016777 100
## [239] {whole milk,hard cheese}                           0.01006609  99
## [240] {other vegetables,waffles}                         0.01006609  99
## [241] {curd,rolls/buns}                                  0.01006609  99
## [242] {tropical fruit,napkins}                           0.01006609  99
## [243] {frankfurter,sausage}                              0.01006609  99
## [244] {sausage,fruit/vegetable juice}                    0.01006609  99
## [245] {whole milk,curd,yogurt}                           0.01006609  99
```

Build rules from the itemsets

```
grocery_rules <- ruleInduction(grocery_itemsets, transactions=Groceries, confidence = 0.50)
inspect(sort(grocery_rules, by="confidence"))
```

```
##      lhs                                    rhs                support
## [1]  {citrus fruit,root vegetables}      => {other vegetables} 0.01037112
## [2]  {tropical fruit,root vegetables}    => {other vegetables} 0.01230300
## [3]  {curd,yogurt}                       => {whole milk}       0.01006609
## [4]  {other vegetables,butter}           => {whole milk}       0.01148958
## [5]  {tropical fruit,root vegetables}    => {whole milk}       0.01199797
## [6]  {root vegetables,yogurt}            => {whole milk}       0.01453991
## [7]  {other vegetables,domestic eggs}    => {whole milk}       0.01230300
## [8]  {yogurt,whipped/sour cream}         => {whole milk}       0.01087951
## [9]  {root vegetables,rolls/buns}        => {whole milk}       0.01270971
## [10] {pip fruit,other vegetables}        => {whole milk}       0.01352313
## [11] {tropical fruit,yogurt}             => {whole milk}       0.01514997
## [12] {other vegetables,yogurt}           => {whole milk}       0.02226741
## [13] {other vegetables,whipped/sour cream} => {whole milk}     0.01464159
## [14] {root vegetables,rolls/buns}        => {other vegetables} 0.01220132
## [15] {root vegetables,yogurt}            => {other vegetables} 0.01291307
##      confidence lift     itemset
## [1]  0.5862069  3.029608 224
## [2]  0.5845411  3.020999 229
## [3]  0.5823529  2.279125 214
## [4]  0.5736041  2.244885 216
## [5]  0.5700483  2.230969 230
## [6]  0.5629921  2.203354 236
## [7]  0.5525114  2.162336 217
## [8]  0.5245098  2.052747 220
## [9]  0.5230126  2.046888 238
## [10] 0.5175097  2.025351 222
## [11] 0.5173611  2.024770 232
## [12] 0.5128806  2.007235 244
## [13] 0.5070423  1.984385 221
```

```
## [14] 0.5020921  2.594890 237
## [15] 0.5000000  2.584078 235
```

The EPubs dataset

```
data("Epub")
summary(Epub)
```

```
## transactions as itemMatrix in sparse format with
##  15729 rows (elements/itemsets/transactions) and
##  936 columns (items) and a density of 0.001758755
##
## most frequent items:
## doc_11d doc_813 doc_4c6 doc_955 doc_698 (Other)
##     356     329     288     282     245   24393
##
## element (itemset/transaction) length distribution:
## sizes
##     1     2     3     4     5     6     7     8     9    10    11    12    13
## 11615  2189   854   409   198   121    93    50    42    34    26    12    10
##    14    15    16    17    18    19    20    21    22    23    24    25    26
##    10     6     8     6     5     8     2     2     3     2     3     4     5
##    27    28    30    34    36    38    41    43    52    58
##     1     1     1     2     1     2     1     1     1     1
##
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   1.000   1.000   1.000   1.646   2.000  58.000
##
## includes extended item information - examples:
##    labels
## 1 doc_11d
## 2 doc_13d
## 3 doc_14c
##
## includes extended transaction information - examples:
##       transactionID           TimeStamp
## 10792  session_4795 2003-01-01 20:59:00
## 10793  session_4797 2003-01-02 07:46:01
## 10794  session_479a 2003-01-02 10:50:38
```

Use apriori to find the rules with confidence 0.75 or better

```
epub_rules <- apriori(Epub, parameter=list(confidence=0.01, support=0.01,minlen=2))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##        0.01    0.1    1 none FALSE            TRUE       5    0.01      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: 157
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[936 item(s), 15729 transaction(s)] done [0.01s].
## sorting and recoding items ... [19 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 done [0.00s].
## writing ... [0 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
inspect(sort(epub_rules, by="confidence"))
```

Extract frequent itemsets

```
epub_itemsets <- apriori(Epub, parameter=list(target="frequent itemsets", support=0.01,minlen=2))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##          NA    0.1    1 none FALSE            TRUE       5    0.01      2
##  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: 157
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[936 item(s), 15729 transaction(s)] done [0.01s].
## sorting and recoding items ... [19 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 done [0.00s].
## writing ... [0 set(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
inspect(sort(epub_itemsets, by="support"))
```