RPubs *brought to you by RStudio*

Example 5

Comments (–)   Share   Hide Toolbars

Finally, let's go fishing for substitute products. So far we were concerned with complements, i.e. items and itemsets that showed high `lift`. In other words, they were deliberately bought together much more often than it were warranted by sheer chance.

Let's consider case "Bottled beer Vs. Canned beer" and prove that people tend to buy either one or the other, and rarely do they buy both, qualifying these two as substitute products.

```
rules <- apriori(Groceries,
                 parameter = list(support=.001,
                                  conf = .01,
                                  minlen=2,
                                  maxlen=2,
                                  target='rules'
                                  ))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport support minlen maxlen target    ext
##        0.01    0.1    1 none FALSE            TRUE   0.001      2      2  rules FALSE
##
## Algorithmic control:
##  filter tree heap memopt load sort verbose
##     0.1 TRUE TRUE  FALSE TRUE    2    TRUE
##
## Absolute minimum support count: 9
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[169 item(s), 9835 transaction(s)] done [0.01s].
## sorting and recoding items ... [157 item(s)] done [0.00s].
## creating transaction tree ... done [0.01s].
## checking subsets of size 1 2 done [0.00s].
## writing ... [5818 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

Let's only look at the rules where "beer" is present at both left- and right-hand-side of the rule and add `chiSquared` p-value to prove statistical significance of our findings:

```
quality(rules)$chi  <- interestMeasure(rules, measure='chi', significance=T, Groceries)
inspect(subset(rules, lhs %pin% 'beer' & rhs %pin% 'beer'))
```

```
##      lhs              rhs             support confidence lift   chi
## 4785 {canned beer} => {bottled beer} 0.002644 0.03403    0.4226 8.743e-07
## 4786 {bottled beer} => {canned beer}  0.002644 0.03283    0.4226 8.743e-07
```