The University of Texas at Dallas

Naveen Jindal School of Management

Store Product sales Analysis

A store is interested in determining the associations between items purchased from the Health and Beauty Aids department and the Stationery Department. The store chose to conduct a market basket analysis of specific items purchased from these two departments. TRANSACTIONS contains information about over 400,000 transactions made over the past three months. The following 17 products are represented in the data set: bar soap, bows, candy bars, deodorant, greeting cards, magazines, markers, pain relievers, pencils, pens, perfume, photo processing, prescription medications, shampoo, toothbrushes, toothpaste, and wrapping paper.

Name	Model Role	Data Type	Description
STORE	Ignore	Numeric	Identification number of the store
TRANSACTION	Ident	Numeric	Transaction identification number
PRODUCT	Target	Categorical	Product purchased
QUANTITY	Ignore	Numeric	Quantity of this product purchased

1.> a.), b.), c.)

```
> # Generate rules and sort by lift
> rules <- apriori(transactions_data1, parameter = list(supp = 0.03, conf = 0.20, minlen = 2, maxlen = 4))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target   ext
        0.2    0.1    1 none FALSE            TRUE       5    0.03      2      4  rules FALSE

Algorithmic control:
 filter tree heap memopt load sort verbose
    0.1 TRUE TRUE  FALSE TRUE    2    TRUE

Absolute minimum support count: 6000

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[17 item(s), 200000 transaction(s)] done [0.03s].
sorting and recoding items ... [13 item(s)] done [0.00s].
creating transaction tree ... done [0.05s].
checking subsets of size 1 2 3 done [0.00s].
writing ... [9 rule(s)] done [0.00s].
creating S4 object  ... done [0.01s].
> rules <- sort(rules, by="lift", decreasing=TRUE)
>
> # Limit output to 2 digits
> options(digits=2)
>
```

d.)

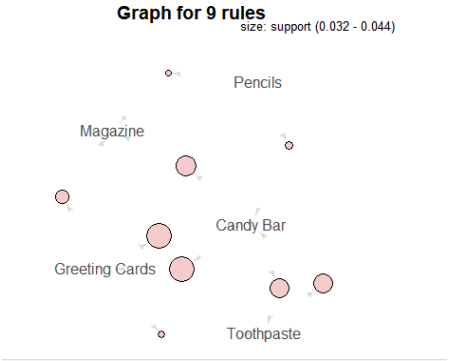
Only nine rules exist so inspect(rules [1:10]) will show out of bound error.

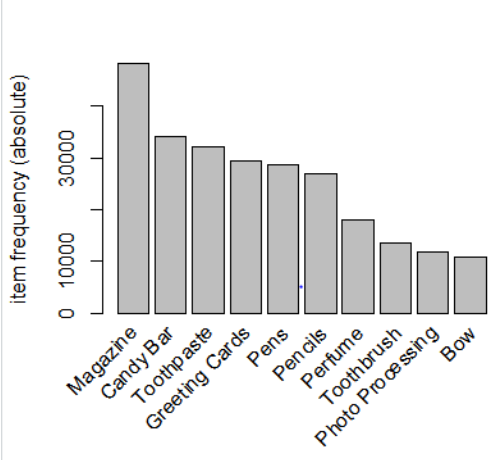
Therefore running inspect(rules [1:9]).

```
     lhs                    rhs                support confidence lift count
[1] {Greeting Cards}  => {Candy Bar}         0.044   0.30       1.74 8732
[2] {Candy Bar}       => {Greeting Cards}    0.044   0.26       1.74 8732
[3] {Toothpaste}      => {Candy Bar}         0.040   0.25       1.45 7956
[4] {Candy Bar}       => {Toothpaste}        0.040   0.23       1.45 7956
[5] {Pencils}         => {Candy Bar}         0.033   0.24       1.43 6603
[6] {Greeting Cards}  => {Toothpaste}        0.032   0.22       1.36 6416
[7] {Greeting Cards}  => {Magazine}          0.036   0.25       1.03 7267
[8] {Candy Bar}       => {Magazine}          0.041   0.24       0.98 8107
[9] {Pencils}         => {Magazine}          0.032   0.23       0.97 6326
```

e.) Remove duplicates

```
> redundant_index <- is.redundant(rules)
> redundant_index
[1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
> pruned_rules <- rules[!redundant_index]
```





2.> a.) Highest lift value is 1.74. The rules that have max lift.

{Greeting Cards} => {Candy Bar}

{Candy Bar} => {Greeting Cards}

b.)Lift is calculated using the following formula:

Lift=confidence of rule/support for the consequent=P(Consequent|Antecedent)/P(Consequent)

Where Consequent is RHS and Antecedent is LHS

c.)Interpreting first five rules.

[1] {Greeting Cards} => {Candy Bar} Suggest candy bars to people who bought greeting cards as most pe ople who buy greeting cards also tend to buy candy bars.
[2] {Candy Bar} => {Greeting Cards} Suggest greeting cards to people who bought candy bars as most people who buy candy bars also tend to buy greeting cards.
[3] {Toothpaste} => {Candy Bar} Suggest candy bars to people who bought toothpaste as most people who buy toothpaste also tend to buy candy bars.
[4] {Candy Bar} => {Toothpaste} Suggest toothpaste to people who bought candy bars as most people who buy candy bars also tend to buy toothpaste.
[5] {Pencils} => {Candy Bar} Suggest candy bars to people who bought pencils as most people who bu y pencils also tend to buy candy bars.

d.)There is no redundancy. We can access these rules as a decision maker. Like advertising Cand y bar, magazine and toothpaste to people who buy Greeting cards. Advertizing greeting card, to othpaste and magazine to people who buy candy bars etc.