Market Basket Analysis

Author: Nandini Ethirajulu

```
# Data Description
# The dataset (marketbasket.xlsx) provided by RM's Information Technology
group contains purchases by 1000 customers over a one-week period. Each
customer's purchase behavior is recorded with variables indicating the type
of bread (wheat, white, or none), jelly (grape, strawberry, or none), and
peanut butter (creamy, natural, or none).

#A) Creating a Binary Transaction Matrix The R script reads the dataset,
preprocesses it by converting variables to factors, and creates a binary
transaction matrix



#Read the data
#install.packages("readxl")
library(readxl)

## Warning: package 'readxl' was built under R version 4.3.2

setwd("C:/Users/nandi/Documents/Personal/Academics/Projects/Market Basket
Analysis")


marketbasket_dats <- read_excel("marketbasket.xlsx", sheet = "Data")



###########Preprocessing########
#Renaming the column name from "Peanut Butter" to "Peanut_Butter"
colnames(marketbasket_dats)[3]<- "Peanut_Butter"

#converting the column values to factors
marketbasket_dats$Bread <- as.factor(marketbasket_dats$Bread)
marketbasket_dats$Jelly <- as.factor(marketbasket_dats$Jelly)
marketbasket_dats$Peanut_Butter <- as.factor(marketbasket_dats$Peanut_Butter)



######### Creating a Binary Transactional Matrix ##########

marketbasket_transaction_matrix <- as.matrix(marketbasket_dats)

marketbasket_transaction_matrix[marketbasket_transaction_matrix != "none"] <-
1
marketbasket_transaction_matrix[marketbasket_transaction_matrix == "none"] <-
0
```

```r
##Binary Transaction Matrix created
head(marketbasket_transaction_matrix)
```

```
##      Bread Jelly Peanut_Butter
## [1,] "1"   "1"   "0"
## [2,] "1"   "1"   "0"
## [3,] "0"   "0"   "0"
## [4,] "1"   "0"   "0"
## [5,] "1"   "0"   "0"
## [6,] "0"   "0"   "0"
```

```r
######### Writing the matrix completely to a tab  delimited text file
#########

write.table(marketbasket_transaction_matrix, sep = "\t", file =
"marketbasket_binary_transaction_matrix.txt",
            row.names = FALSE,
            col.names = colnames(marketbasket_transaction_matrix))



# B) Initial Study and Recommendations
#initial study of the data to get a better understanding of RM customer
behavior about these three products


library("arules")
```

```
## Warning: package 'arules' was built under R version 4.3.2
```

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

```
## Warning: package 'Matrix' was built under R version 4.3.3
```

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

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

```r
#converting data into transactions
marketbasket_transactions <- as(marketbasket_dats, "transactions")

##Applying apriori algorithm to extract the association rules

marketbasket_rules  <- apriori(marketbasket_transactions, parameter =
list(support = 0.1, confidence = 0.6))
```

```
## Apriori
##
## Parameter specification:
##  confidence minval smax arem  aval originalSupport maxtime support minlen
##         0.6    0.1    1 none FALSE            TRUE       5     0.1      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: 100
##
## set item appearances ...[0 item(s)] done [0.00s].
## set transactions ...[9 item(s), 1000 transaction(s)] done [0.00s].
## sorting and recoding items ... [8 item(s)] done [0.00s].
## creating transaction tree ... done [0.00s].
## checking subsets of size 1 2 3 done [0.00s].
## writing ... [19 rule(s)] done [0.00s].
## creating S4 object  ... done [0.00s].
```

```
#Apriori rules sorted by lift percentage and filtered top 10 ranking
associations


inspect(head(sort(marketbasket_rules, by = "lift"), n = 10))
```

```
##      lhs                                      rhs           support
confidence
## [1]  {Jelly=grape, Peanut_Butter=creamy} => {Bread=white} 0.124
0.9763780
## [2]  {Bread=white, Peanut_Butter=creamy} => {Jelly=grape} 0.124
0.6666667
## [3]  {Jelly=grape}                       => {Bread=white} 0.306
0.8793103
## [4]  {Bread=white}                       => {Jelly=grape} 0.306
0.6637744
## [5]  {Bread=white, Peanut_Butter=none}   => {Jelly=grape} 0.144
0.6515837
## [6]  {Peanut_Butter=creamy}              => {Bread=white} 0.186
0.8157895
## [7]  {Jelly=grape, Peanut_Butter=none}   => {Bread=white} 0.144
0.8044693
## [8]  {Jelly=none, Peanut_Butter=none}    => {Bread=none}  0.278
0.6289593
## [9]  {Bread=none, Peanut_Butter=none}    => {Jelly=none}  0.278
0.8714734
## [10] {Bread=none}                        => {Jelly=none}  0.325
0.8666667
```

```
##      coverage lift     count
## [1]  0.127    2.117957 124
## [2]  0.186    1.915709 124
## [3]  0.348    1.907398 306
## [4]  0.461    1.907398 306
## [5]  0.221    1.872367 144
## [6]  0.228    1.769608 186
## [7]  0.179    1.745053 144
## [8]  0.442    1.677225 278
## [9]  0.319    1.518246 278
## [10] 0.375    1.509872 325
```

##All the top 10 rankings have lift value greater than 1 indicating that the products in the lhs increases the probability of item purchase in the rhs.

######## Analysis and advise #########

 # From the top 10 rankings of association rules of the three product combinations, it is recommended to promote the item combinations of Jelly - grape, Peanut_Butter - creamy and white bread, as it is expected to more prominently increase the sales and frequently bought by the customers.

#Combinations of white bread and creamy peanut butter are most likely to increase the purchase of the grape jelly. hence it is suggested to align together in the store to increase the sales and purchase frequency.

#In addition to the above, some customers are also interested only in buying jelly-grape along with white breads, without purchasing any type of peanut butter. So it is recommended to apply a combo or bundling options for these two products to elevate the customer's interests and choices.

#One other important transactions are creamy peanut butter purchases, which are increasing the customer's likeliness to purchase white bread most often.

##In scenarios where customer's did not purchase any jelly or peanut butter, they also avoided the pourchase of white bread. Similarly, even when they were not buying breads or jellys, they have ignored the purchase of peanut butter.

##Finally based on our initial analysis we recommend David to implement thoughtful promotions, offers or bundling options for the above combination items in the product list to increase the sales.