**MARKETING AND RETAIL ANALYTICS**

**GROUP ASSIGNMENT**

**ON**

**“MARKET BASKET ANALYSIS OF A GROCERY STORE”**

**PGP-BABI-HYD**

**SUBMISSION BY: GROUP 3**

**MEMBERS:**

**Asha Murmu**

**Devlina Chai**

**Nabasish Bhattacharjee**

**Naveen Kumar Sambangi**

**Ramya Boodidha**

**TABLE OF CONTENTS**

1. INTRODUCTION …………………………………………………………………………………………………….………….…….…3
2. INSTALL & LOAD REQUIRED PACKAGES ………………………………………………………………………………….....3
3. IMPORT DATA INTO R ………………………………………………………………………………………………………….….…3
4. SUMMARY OF GROCERY TRANSANCTIONS …………………………………………………………………………….….3
5. TOP 20 “FREQUENTLY BOUGHT” ITEMS ……………………………………………………………………..…….……….4
6. CREATE ASSOCIATION RULES ………………………………………………………………………………………….………….5

* SCENARIO 1: SUPPORT = 0.001 AND CONFIDENCE = 0.8
* SCENARIO 2: RESTRICTING NUMBER OF ITEMS ON LHS TO 2 WITH SUPPORT = 0.001 AND CONFIDENCE = 0.8
* SCENARIO 3: CHECK THE RULES WITH DIFFERENT VALUES OF SUPPORT = 0.0015, 0.0018, 0.0025

1. BUSINESS RECOMMENDATIONS FROM MARKET BASKET ANALYSIS ………………………………………..10
2. APPENDIX: R-CODE ………………………………………………………………………………………………………………….12

**1. INTRODUCTION**

**OBJECTIVE:**

On a grocery receipts data, the objective is to perform Market Basket analysis and look at association rules, to understand what goods can be bundled and what types of promotions or discounts can be offered on the products either online or at the retail store.

**DATA SET:**

10000 collection of receipts with each line representing 1 receipt and the items purchased by a customer. Each line is called a transaction and each column in a row represents an item.

**GLIMPSE OF THE DATA SET:**



**2. INSTALL AND LOAD THE REQUIRED PACKAGES**

Following packages are installed and loaded:

arules, arulesViz, datasets, dplyr, Rcpp, magrittr, readxl, ggplot2, cluster.

**3. IMPORT DATA INTO R**

Set the working directory and then the data from the given groceries.csv is loaded into “trans” using a read.transactions() function.

**4. SUMMARY OF GROCERY TRANSACTIONS**

summary() function is used to look at the summary of the transactions of the grocery data, following is the output and important insights from summary output:

**SUMMARY OUTPUT:**

**> summary(trans)**

transactions as itemMatrix in sparse format with

**9835** rows (elements/itemsets/transactions) and

**169** columns (items) and a density of 0.0261

most frequent items:

whole milk other vegetables rolls/buns soda yogurt Other

2513 1903 1809 1715 1372 34055

element (itemset/transaction) length distribution:

sizes

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21

2159 1643 1299 1005 855 645 545 438 350 246 182 117 78 77 55 46 29 14 14 9 11

22 23 24 26 27 28 29 **32**

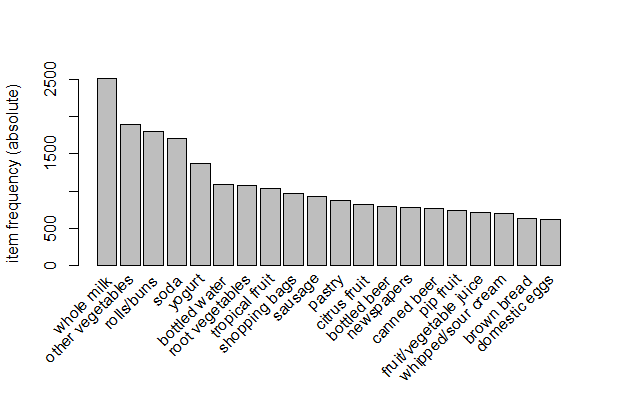
4 6 1 1 1 1 3 **1**

**INSIGHTS FROM SUMMARY:**

* There are 9835 transaction records in the given data set
* Total number of **unique items** is 169
* At-most 32 items were purchased in one of the transactions (that is, number of items purchased in a transaction range from 1 to a max of 32)

**5. TOP 20 “FREQUENTLY BOUGHT” ITEMS**

Following is the plot depicting top 20 frequently bought items along with the number of times the item appeared in 9835 transactions:



**INSIGHTS FROM THE FREQUENCY PLOT:**

* Among all the items, **“WHOLE MILK”** is the **most frequently bought item** as it appeared in almost 2500 transactions.
* After the whole milk, next most frequently bought items are “OTHER VEGETABLES”, “ROLLS/BUNS”, “SODA”, which appeared in almost 1800 transactions.

**6. CREATE ASSOCIATION RULES**

**FOREWORDS:** Since the given data set is small with wide variety of items, we are starting with a very low value of support = 0.001 and very good confidence level of 0.8. Later on, we will increase the support value a bit and check the association rules.

**SCENARIO 1: SUPPORT = 0.001 AND CONFIDENCE = 0.8**

**STEPS FOLLOWED IN DERIVING THE ASSOCIATION RULES:**

* Obtain the rules by building a “Apriori Algorithm” with support= 0.001 and confidence = 0.8
* Sort the rules in descending order of confidence
* Check for any redundant rules and it is found that there are **18 redundant rules**
* Analysis is done on remaining non-redundant rules,which is as follows:

**SUMMARY OF ASSOCIATION RULES:**

**> summary(rules)**

set of **392** rules

rule length distribution (lhs + rhs):sizes

3 4 5 6

29 **227** 130 6

Min. 1st Qu. Median Mean 3rd Qu. Max.

3.0 4.0 4.0 4.3 5.0 6.0

summary of quality measures:

support confidence lift count

Min. :**0.00102** Min. :**0.80** Min. : **3.1** Min. :10.0

1st Qu.:0.00102 1st Qu.:0.83 1st Qu.: 3.3 1st Qu.:10.0

Median :0.00122 Median :0.85 Median : 3.6 Median :12.0

Mean :0.00125 Mean :0.87 Mean : 4.0 Mean :12.3

3rd Qu.:0.00132 3rd Qu.:0.91 3rd Qu.: 4.4 3rd Qu.:13.0

Max. :**0.00315** Max. :**1.00** Max. :**11.2** Max. :31.0

mining info:

data ntransactions support confidence

trans 9835 0.001 0.8

**INSIGHTS FROM SUMMARY OF ASSOCIATION RULES:**

* There are 392 non-redundant rules
* Out of these 392 rules, majority (227 rules) of the rules are 4 items Long
* All the rules have very good CONFIDENCE in the range of 80% to 100%
* All the rules have very good LIFT in the range of 3.1 to 11.2 (MAX)
* Maximum support that can be used for this data is 0.0031 (that is 0.3%)

**INSPECT FIRST 20 ASSOCIATION RULES:**

> inspect(rules[1:20])

lhs rhs support confidence lift count

[1] {rice,sugar} => {whole milk} 0.0012 1 3.9 12

[2] {canned fish,hygiene articles} => {whole milk} 0.0011 1 3.9 11

[3] {butter,rice,root vegetables} => {whole milk} 0.0010 1 3.9 10

[4] {flour,root vegetables,whipped/sour cream} => {whole milk} 0.0017 1 3.9 17

[5] {butter,domestic eggs,soft cheese} => {whole milk} 0.0010 1 3.9 10

[6] {citrus fruit,root vegetables,soft cheese} => {other vegetables} 0.0010 1 5.2 10

[7] {butter,hygiene articles,pip fruit} => {whole milk} 0.0010 1 3.9 10

[8] {hygiene articles,root vegetables,whipped/sour cream} => {whole milk} 0.0010 1 3.9 10

[9] {hygiene articles,pip fruit,root vegetables} => {whole milk} 0.0010 1 3.9 10

[10] {cream cheese,domestic eggs,sugar} => {whole milk} 0.0011 1 3.9 11

[11] {curd,domestic eggs,sugar} => {whole milk} 0.0010 1 3.9 10

[12] {cream cheese,domestic eggs,napkins} => {whole milk} 0.0011 1 3.9 11

[13] {brown bread,pip fruit,whipped/sour cream} => {other vegetables} 0.0011 1 5.2 11

[14] {grapes,tropical fruit,whole milk,yogurt} => {other vegetables} 0.0010 1 5.2 10

[15] {ham,pip fruit,tropical fruit,yogurt} => {other vegetables} 0.0010 1 5.2 10

[16] {ham,pip fruit,tropical fruit,whole milk} => {other vegetables} 0.0011 1 5.2 11

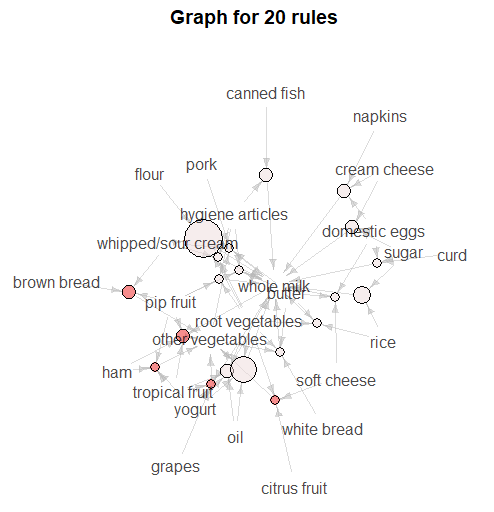
[17] {oil,root vegetables,tropical fruit,yogurt} => {whole milk} 0.0011 1 3.9 11

[18] {oil,other vegetables,root vegetables,yogurt} => {whole milk} 0.0014 1 3.9 14

[19] {butter,other vegetables,root vegetables,white bread} => {whole milk} 0.0010 1 3.9 10

[20] {butter,other vegetables,pork,whipped/sour cream} => {whole milk} 0.0010 1 3.9 10

**GRAPHICAL DISPLAY OF FIRST 20 ASSOCIATION RULES:**

****



**IMPORTANT INSIGHTS FROM FIRST 20 ASSOCIATION RULES:**

* All the first 20 rules have 100% confidence and lift > 1, indicating that these rules will have very good operational usefulness
* RHS for all the first 20 rules is either “WHOLE MILK” OR “OTHER VEGETABLES”.
* Rice and sugar

canned fish and hygiene articles

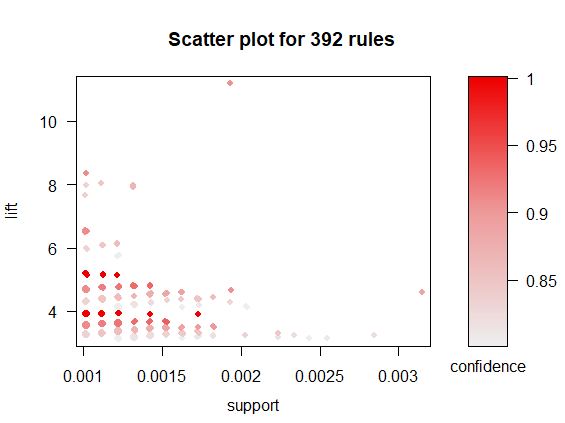
butter and rice and root vegetables

flour, root vegetables and whipped cream

butter, domestic eggs and soft cheese

**All the above 5 purchase patterns led to the purchase of “WHOLE MILK”,** **that means, for example: with 100% confidence we can say that customers who bought rice and sugar also bought “whole milk”.**

**SCATTER PLOT OF ALL ASSOCIATION RULES:**



**INSIGHTS FROM SCATTER PLOT:**

* Majority of the rules have a support value of 0.001, and many of the rules are concentrated in the support range of 0.001 to 0.002.
* As support values increases, number of rules is decreasing, with support = 0.0031, there is only one rule.
* Lift for the rules is in the range of 3 to a maximum of 11
* Darker red dots indicate very good confidence level of 90% to 100%, rules which have support values from 0.001 to 0.002 have more concentration of darker red spots.

**FURTHER ANALYSIS ON TRANSACTIONS RELATED TO “WHOLE-MILK” & “OTHER VEGETABLES”**

**foreword:** from the association rules, “whole milk” and “other vegetables” seem to very important ones and hence we are trying to analyse them further:

**To find out what customers had purchased before buying ‘Whole Milk’ (This will help understand the patterns that led to the purchase of ‘whole milk’:**

* **3765 rules** are obtained by building a “Apriori Algorithm” with **RHS= “whole milk”,** support= 0.001 and confidence = 0.8, out of which let us inspect first 10 rules:

lhs rhs support confidence lift count

[1] {rice,sugar} => {whole milk} 0.0012 1 3.9 12

[2] {canned fish,hygiene articles} => {whole milk} 0.0011 1 3.9 11

[3] {butter,rice,root vegetables} => {whole milk} 0.0010 1 3.9 10

[4] {flour,root vegetables,whipped/sour cream} => {whole milk} 0.0017 1 3.9 17

[5] {butter,domestic eggs,soft cheese} => {whole milk} 0.0010 1 3.9 10

[6] {butter,hygiene articles,pip fruit} => {whole milk} 0.0010 1 3.9 10

[7] {hygiene articles,root vegetables,whipped/sour cream} => {whole milk} 0.0010 1 3.9 10

[8] {hygiene articles,pip fruit,root vegetables} => {whole milk} 0.0010 1 3.9 10

[9] {cream cheese,domestic eggs,sugar} => {whole milk} 0.0011 1 3.9 11

[10] {curd,domestic eggs,sugar} => {whole milk} 0.0010 1 3.9 10

**Customers who bought ‘Whole Milk’ also bought what products (KIND OF COLLABORATIVE FILTERING):**

* **6 rules** are obtained by building a “Apriori Algorithm” with **LHS= “whole milk”,** support= 0.001 and confidence = 0.8, out of which let us inspect 6 rules:

lhs rhs support confidence lift count

[1] {whole milk} => {other vegetables} 0.075 0.29 1.5 736

[2] {whole milk} => {rolls/buns} 0.057 0.22 1.2 557

[3] {whole milk} => {yogurt} 0.056 0.22 1.6 551

[4] {whole milk} => {root vegetables} 0.049 0.19 1.8 481

[5] {whole milk} => {tropical fruit} 0.042 0.17 1.6 416

[6] {whole milk} => {soda} 0.040 0.16 0.9 394

**To find out what customers had purchased before buying ‘other vegetables’ (This will help understand the patterns that led to the purchase of ‘other vegetables’:**

* **3342 rules** are obtained by building a “Apriori Algorithm” with **RHS= “other vegetables”,** support= 0.001 and confidence = 0.8, out of which let us inspect first 10 rules:

lhs rhs support confidence lift count

[1] {citrus fruit,root vegetables,soft cheese} => {other vegetables} 0.0010 1.00 5.2 10

[2] {brown bread,pip fruit,whipped/sour cream} => {other vegetables} 0.0011 1.00 5.2 11

[3] {grapes,tropical fruit,whole milk,yogurt} => {other vegetables} 0.0010 1.00 5.2 10

[4] {ham,pip fruit,tropical fruit,yogurt} => {other vegetables} 0.0010 1.00 5.2 10

[5] {ham,pip fruit,tropical fruit,whole milk} => {other vegetables} 0.0011 1.00 5.2 11

[6] {butter,fruit/vegetable juice,tropical fruit,whipped/sour cream} => {other vegetables} 0.0010 1.00 5.2 10

[7] {newspapers,rolls/buns,soda,whole milk} => {other vegetables} 0.0010 1.00 5.2 10

[8] {citrus fruit,root vegetables,tropical fruit,whipped/sour cream} => {other vegetables} 0.0012 1.00 5.2 12

[9] {oil,root vegetables,whole milk,yogurt} => {other vegetables} 0.0014 0.93 4.8 14

[10] {citrus fruit,root vegetables,tropical fruit,whole milk,yogurt} => {other vegetables} 0.0014 0.93 4.8 14

**Customers who bought ‘other vegetables’ also bought what products (KIND OF COLLABORATIVE FILTERING):**

* **6 rules** are obtained by building a “Apriori Algorithm” with **LHS= “other vegetables”,** support= 0.001 and confidence = 0.8, out of which let us inspect 6 rules:

lhs rhs support confidence lift count

[1] {other vegetables} => {whole milk} 0.075 0.39 1.51 736

[2] {other vegetables} => {root vegetables} 0.047 0.24 2.25 466

[3] {other vegetables} => {yogurt} 0.043 0.22 1.61 427

[4] {other vegetables} => {rolls/buns} 0.043 0.22 1.20 419

[5] {other vegetables} => {tropical fruit} 0.036 0.19 1.77 353

[6] {other vegetables} => {soda} 0.033 0.17 0.97 322

**SCENARIO 2: RESTRICTING NUMBER OF ITEMS ON LHS TO 2 WITH SUPPORT = 0.001 AND CONFIDENCE = 0.8**

* Obtain the rules (with a restriction of number of items on LHS to be 2) by building a “Apriori Algorithm” with support= 0.001 and confidence = 0.8

**INSPECT FIRST 10 ASSOCIATION RULES:**

lhs rhs support confidence lift count

**[1] {liquor,red/blush wine} => {bottled beer} 0.0019 0.90 11.2 19**

[2] {cereals,curd} => {whole milk} 0.0010 0.91 3.6 10

[3] {cereals,yogurt} => {whole milk} 0.0017 0.81 3.2 17

[4] {butter,jam} => {whole milk} 0.0010 0.83 3.3 10

[5] {bottled beer,soups} => {whole milk} 0.0011 0.92 3.6 11

[6] {house keeping products,napkins} => {whole milk} 0.0013 0.81 3.2 13

[7] {house keeping products,whipped/sour cream} => {whole milk} 0.0012 0.92 3.6 12

[8] {pastry,sweet spreads} => {whole milk} 0.0010 0.91 3.6 10

[9] {curd,turkey} => {other vegetables} 0.0012 0.80 4.1 12

[10] {rice,sugar} => {whole milk} 0.0012 1.00 3.9 12

**INSIGHT FROM FIRST 10 RULES:**

* First rule has a MAXIMUM LIFT of 11.2 and a VERY GOOD CONFIDENCE of 90%, indicating that **most of the people** **who bought Liquor and red wine also bought bottled beer.** Because of high lift and high confidence, this rule has **high revenue generating value** for the business.

**SCENARIO 3: CHECK THE RULES WITH DIFFERENT VALUES OF SUPPORT = 0.0015, 0.0018, 0.0025**

**CASE A: SUPPORT = 0.0015 AND CONFIDENCE = 0.8**

* **60 rules** were obtained by building a “Apriori Algorithm” with support= 0.0015 and confidence = 0.8, out of which let us inspect **first 10 rules:**

lhs rhs support confidence lift count

[1] {flour,root vegetables,whipped/sour cream} => {whole milk} 0.0017 1.00 3.9 17

[2] {cream cheese,other vegetables,sugar} => {whole milk} 0.0015 0.94 3.7 15

[3] {root vegetables,sausage,tropical fruit,yogurt} => {whole milk} 0.0015 0.94 3.7 15

[4] {liquor,red/blush wine} => {bottled beer} 0.0019 0.90 11.2 19

[5] {fruit/vegetable juice,tropical fruit,whipped/sour cream} => {other vegetables} 0.0019 0.90 4.7 19

[6] {butter,pip fruit,whipped/sour cream} => {whole milk} 0.0018 0.90 3.5 18

[7] {domestic eggs,tropical fruit,whipped/sour cream} => {whole milk} 0.0018 0.90 3.5 18

[8] {butter,root vegetables,tropical fruit,yogurt} => {whole milk} 0.0017 0.89 3.5 17

[9] {ham,pip fruit,tropical fruit} => {other vegetables} 0.0016 0.89 4.6 16

[10] {onions,root vegetables,tropical fruit} => {other vegetables} 0.0016 0.89 4.6 16

**CASE B: SUPPORT = 0.0018 AND CONFIDENCE = 0.8**

* **20 rules** were obtained by building a “Apriori Algorithm” with support= 0.0018 and confidence = 0.8, out of which let us inspect **first 10 rules:**

lhs rhs support confidence lift count

[1] {liquor,red/blush wine} => {bottled beer} 0.0019 0.90 11.2 19

[2] {fruit/vegetable juice,tropical fruit,whipped/sour cream} => {other vegetables} 0.0019 0.90 4.7 19

[3] {butter,pip fruit,whipped/sour cream} => {whole milk} 0.0018 0.90 3.5 18

[4] {domestic eggs,tropical fruit,whipped/sour cream} => {whole milk} 0.0018 0.90 3.5 18

[5] {citrus fruit,root vegetables,tropical fruit,whole milk} => {other vegetables} 0.0032 0.89 4.6 31

[6] {butter milk,pork} => {other vegetables} 0.0018 0.86 4.4 18

[7] {butter,other vegetables,pork} => {whole milk} 0.0022 0.85 3.3 22

[8] {fruit/vegetable juice,other vegetables,root vegetables,yogurt} => {whole milk} 0.0020 0.83 3.3 20

[9] {rice,yogurt} => {other vegetables} 0.0019 0.83 4.3 19

[10] {herbs,shopping bags} => {other vegetables} 0.0019 0.83 4.3 19

**CASE C: SUPPORT = 0.0025 AND CONFIDENCE = 0.8**

* **3 rules** were obtained by building a “Apriori Algorithm” with support= 0.0025 and confidence = 0.8, out of which let us inspect all **3 rules:**

lhs rhs support confidence lift count

[1] {citrus fruit,root vegetables,tropical fruit,whole milk} => {other vegetables} 0.0032 0.89 4.6 31

[2] {curd,domestic eggs,other vegetables} => {whole milk} 0.0028 0.82 3.2 28

[3] {curd,hamburger meat} => {whole milk} 0.0025 0.81 3.2 25

**7. BUSINESS RECOMMENDATIONS FROM MARKET BASKET ANALYSIS**

**A) GOODS THAT CAN BE BUNDLED TOGETHER/SHELF SPACE:**

**All the below bundles can be sold with 100% confidence since rules show that they have Lift > 1 and confidence is 100%**

* **(Liquor, Red Wine) can be bundled with Bottled Beer:** this bundle can increase the revenue to a great extent as lift is 11.2(maximum of all rules) with a confidence of 90%
* (Rice and sugar) can be bundled with whole-milk
* (canned fish and hygiene articles) can be bundled with whole-milk
* (butter, rice and root vegetables) can be bundled with whole-milk
* (flour, root vegetables and whipped cream) can be bundled with whole-milk
* (butter, domestic eggs and soft cheese) can be bundled with whole-milk
* (citrus fruit, root vegetables, soft cheese) can be bundled with other vegetables
* (brown bread, pip fruit, whipped cream) can be bundled with other vegetables
* (grapes, tropical fruit, whole milk, yogurt) can be bundled with other vegetables
* (ham, pip fruit, tropical fruit, yogurt) can be bundled with other vegetables
* (ham, pip fruit, tropical fruit, whole milk) can be bundled with other vegetables

**B) DISCOUNTS CAN BE OFFERED ON THE FOLLOWING:**

**LOSS LEADER CONCEPT: Discounts should be offered on frequently bought items, so that it will help the business revenue to increase, and therefore we suggest that discounts can be offered on following products:**

Whole milk

Other vegetables

Rolls/buns

Soda

Yogurt

Bottled water

Root veggies

Tropical fruits

Sausage

Pastry

Citrus Fruit

Bottled beer

Pip fruit

Whipped cream

Brown bread

Eggs

**C) CROSS SELLING:**

Since Whole milk, other vegetables, rolls/buns, soda, yogurt, bottled water, root veggies, sausage, pastry, bottled beer, pip fruit, citrus fruit, brown bread, whipped cream, eggs are the top 20 frequently bought items- these items can be **bundled with some unpopular products**, and this kind of cross selling helps in increase in sales and the revenue.

**D) STORE LAYOUT CAN BE DESIGNED BASED ON THE FOLLOWING:**

**1) BREAK THE PURCHASE INERTIA:**

Products with low price and which are bought frequently like **Whole milk, brown bread, eggs, citrus fruit, buns, yogurt** **can be placed at the entrance** to break the purchase inertia of the customers

**2) MAKE CUSTOMERS SPEND MORE TIME IN THE STORE:**

Products like bottled beer, other vegetables which are **popular** can be **placed at the farthest end of the store**, so that the customers look at a greater number of products while walking till the end and there are **chances of buying more and increasing the basket size**

**E) FOR ONLINE STORES:**

* Products listed in “goods that can be bundled together” section has to be placed in their **same category/theme** on the website/app
* Products with high lift and high confidence can be shown in the form of **“DISPLAY ADS”** and personalized emails also can be sent for **targeted marketing**
* Popular products like vegetables, fruits, bottle beer can be displayed on the home page of the website, so that the **customer spends more time on the website**, which might increase the revenue.
* **Discounts on combo Offers**: Providing discounts on products that can be bundled together can also increase the revenue
* **Buy one get one offers/ Upselling:** unpopular and medium cost items can be put in this category to increase the sales
* **Other ways to increase the revenue:**

Offers on digital transactions thru money wallets

Referral discounts/offers

Advertise on social media

**8. APPENDIX: R-CODE**

################################

# Market Basket Analysis

################################

# install and load the required packages

install.packages("readxl") # For reading Excel file

install.packages("Rcpp")

install.packages("dplyr") # For glimpse, etc.

install.packages("magrittr") # For %>% Piping function

install.packages("ggplot2") # For graphical representation

install.packages("cluster")

install.packages("arules")

install.packages("arulesViz", dependencies = TRUE)

install.packages("datasets")

library(dplyr)

library(Rcpp)

library(magrittr)

library(readxl)

library(ggplot2)

library(cluster)

library(arules)

library(arulesViz)

library(datasets)

# Set the working directory

setwd("C:\\mydata\\files")

# Load the data into R

trans <- read.transactions('groceries.csv', format = 'basket', sep=',')

trans

inspect(trans[1:5])

# Summary of grocery transaction data

summary(trans)

# Plot of top 20 frequently bought products

itemFrequencyPlot(trans, topN=20, type='absolute')

#-------------------- Create Association Rules -----------------------------

# case 1 : support=0.001 , confidence = 0.8

# build Apriori algorithhm to find association rules

rules<- apriori(trans, parameter = list(supp=0.001, conf=0.8))

# sorting the rules in the descending order of confidence

rules<- sort(rules, by='confidence', decreasing = TRUE)

# check for any redundant rules

rules\_redundant <- rules[is.redundant(rules)]

rules\_redundant

### 18 redundent rules ###

inspect(rules\_redundant[1:18])

# consider non-redundant rules for further analysis

rules <- rules[!is.redundant(rules)]

### summary of 392 non-redundant rules ####

summary(rules)

# Display top "20" rules

inspect(rules[1:20])

# Graphically display top 20 rules

top20 <- rules[1:20]

plot(top20, method="graph")

# scatter plot

plot (rules, measure=c("support", "lift"), shading="confidence")

#------------------------------------------------------------------------------

# To find out what customers had purchased before buying ‘Whole Milk’.

# This will help understand the patterns that led to the purchase of ‘whole milk’.

rules\_milk1 <- apriori (trans, parameter=list (supp=0.001,conf = 0.08),

appearance = list (default="lhs",rhs="whole milk"), control = list (verbose=F))

rules\_milk1 <- sort(rules\_milk1, by='confidence', decreasing = TRUE)

inspect(rules\_milk1[1:10])

rules\_milk1

# To know : the Customers who bought ‘Whole Milk’ also bought what products ?

# In the equation, ‘whole milk’ is in LHS (left hand side).

rules\_milk2 <- apriori (trans, parameter=list (supp=0.001,conf = 0.15,minlen=2),

appearance = list (default="rhs",lhs="whole milk"), control = list (verbose=F))

rules\_milk2 <- sort(rules\_milk2, by='confidence', decreasing = TRUE)

rules\_milk2

# 1:7 --out of bounds

inspect(rules\_milk2[1:6])

#----------------------------------------

rules\_veg1 <- apriori (trans, parameter=list (supp=0.001,conf = 0.08),

appearance = list (default="lhs",rhs="other vegetables"), control = list (verbose=F))

rules\_veg1 <- sort(rules\_veg1, by='confidence', decreasing = TRUE)

rules\_veg1

inspect(rules\_veg1[1:10])

# To know : the Customers who bought ‘other vegetables’ also bought what products ?

rules\_veg2 <- apriori (trans, parameter=list (supp=0.001,conf = 0.15,minlen=2),

appearance = list (default="rhs",lhs="other vegetables"), control = list (verbose=F))

rules\_veg2 <- sort(rules\_veg2, by='confidence', decreasing = TRUE)

rules\_veg2

inspect(rules\_veg2[1:6])

# ----------------------------------------------------------------

# case 2: Restricting number of items on LHS to 2 and checking first 20 rules

rules\_two\_items <- apriori(trans, parameter = list(supp=0.001, conf=0.8))

options(digits=2)

inspect(rules\_two\_items[1:10])

# ----------------------------------------------------------------

# case 3: checking with other values of support:

rules\_a <- apriori(trans, parameter = list(supp=0.0015, conf=0.8)) # -- 60 rules

rules\_a<- sort(rules\_a, by='confidence', decreasing = TRUE)

inspect(rules\_a[1:10])

rules\_b <- apriori(trans, parameter = list(supp=0.0018, conf=0.8)) # -- 20 rules

rules\_b<- sort(rules\_b, by='confidence', decreasing = TRUE)

inspect(rules\_b[1:10])

rules\_c<- apriori(trans, parameter = list(supp=0.0025, conf=0.8)) # -- 3 rules

rules\_c<- sort(rules\_c, by='confidence', decreasing = TRUE)

inspect(rules\_c[1:3])