// Name: Prajakta Rajendra Bhosale// Roll No: 4012

// Batch: P10

// Problem Statement: Create association rules for market basket analysis for

the given threshold using R.

//-------------------------------------------------------------------------------------------------

# Load the libraries

install.packages("arules")

library(arules)

install.packages("arulesViz")

library(arulesViz)

install.packages("datasets")

library(datasets)

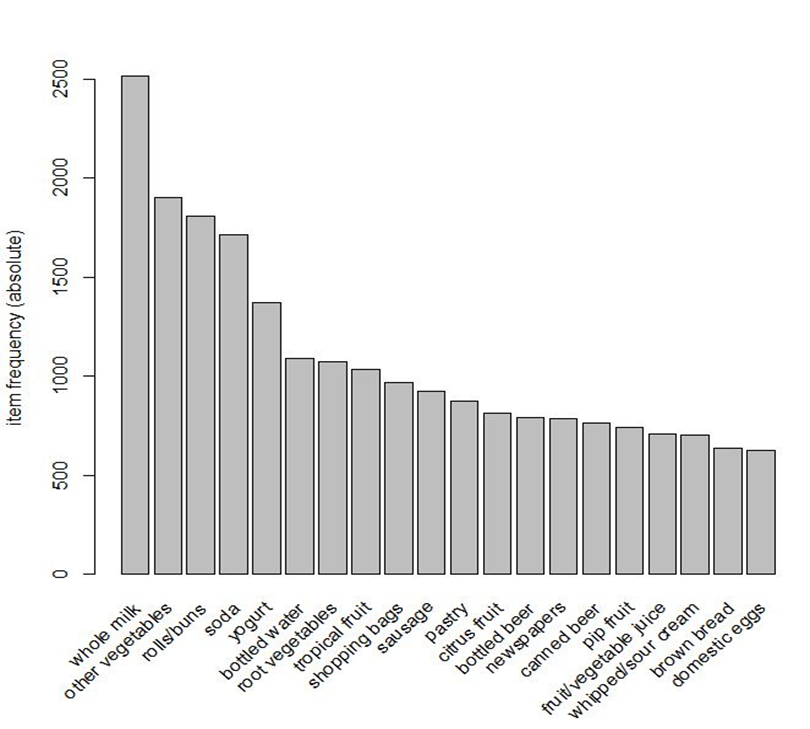
# Load the dataset

data("Groceries")

# Lets explore the data before we make any rule:

# Create an item frequency plot for the top 20 items

itemFrequencyPlot(Groceries,topN=20,type="absolute")



# You will always have to pass the minimum required support and confidence

# We set minimum support to 0.001

# We set minimum confidence to 0.8

# We then show the top 5 rules

# Get the rules

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

# Show the top 5 rules but only 2 digits

options(digits=2)

inspect(rules[1:5])

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

rules <- apriori(Groceries,parameter = list(supp=0.001,conf=0.8,maxlen=3))

subset.matrix <- is.subset(rules,rules)

subset.matrix[lower.tri(subset.matrix,diag = T)] <-NA

redundant <- colSums(subset.matrix,na.rm=T) >=1

rules.pruned <- rules[!redundant]

rules <- rules.pruned

rules <- apriori(data = Groceries,parameter = list(supp=0.001,conf=0.08),appearance = list(default="lhs",rhs="whole milk"),control = list(verbose=F))

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

inspect(rules[1:5])

# 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] {root vegetables,

butter,rice} => {whole milk} 0.0010 1 3.9 10

# [4] {root vegetables,

whipped/sour cream,

flour} => {whole milk} 0.0017 1 3.9 17

# [5] {butter,soft cheese,

domestic eggs} => {whole milk} 0.0010 1 3.9 10

rules <- apriori(data = Groceries,parameter = list(supp=0.001,conf=0.15,minlen=2),appearance = list(default="rhs",lhs="whole milk"),control = list(verbose=F))

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

inspect(rules[1:5])

# lhs rhs support confidence

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

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

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

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

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

# lift count

# [1] 1.5 736

# [2] 1.2 557

# [3] 1.6 551

# [4] 1.8 481

# [5] 1.6 416

library(arulesViz)

plot(rules,method = "graph",interactive = TRUE,shading = NA)

