Module 7: Recommendation Engine

Assignment Solution

Module 7: Assignment Solution

Perform the following activities in R

Task 1: Association Rule mining

→ Load Adult data into R.

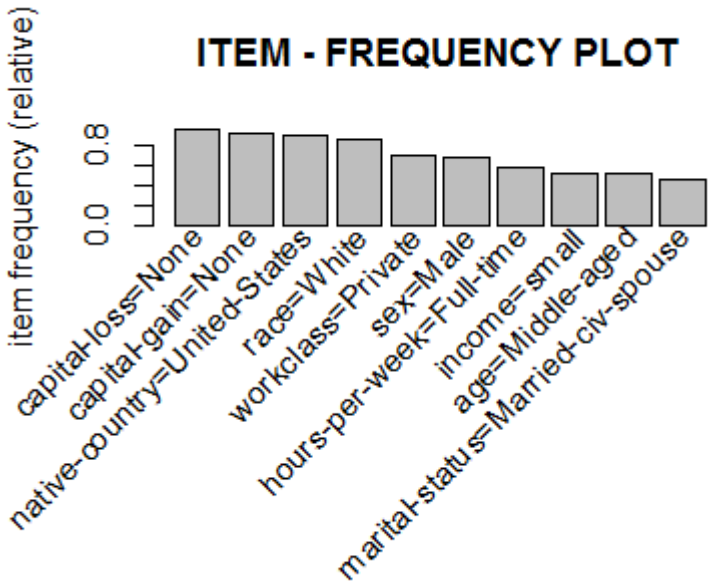
Solution :

```
#Load the data into R
install.packages("arules")
library(arules)

data(package = "arules")$results
data("Adult")

inspect(Adult[1:30])
```

→ Plot a graph of ten items with highest support.

Solution :

```
#Plotting items with highest support values
itemFrequencyPlot(Adult,topN = 10,main = "ITEM - FREQUENCY PLOT")
```



→ Generate association rules with lhs values as "capital-loss=None" and "native-country=United-States"

Solution :

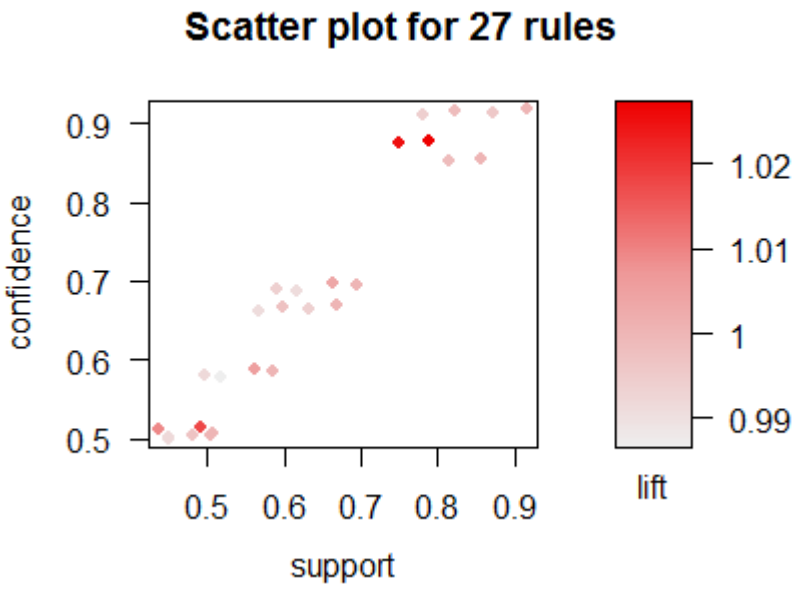
```
#Generating rules
colnames(Adult)
model<-apriori(Adult,parameter = list(support = 0.2,confidence  = 0.5),
               appearance = list(default = "rhs",
                                 lhs = c("capital-loss=None","native-country=United-States")))
```

```
> summary(model)
set of 27 rules

rule length distribution (lhs + rhs):sizes
 1  2  3
 7 14  6

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.000   1.500   2.000   1.963   2.000   3.000

summary of quality measures:
    support         confidence          lift
 Min.   :0.4366   Min.   :0.5002   Min.   :0.9864
 1st Qu.:0.5056   1st Qu.:0.5460   1st Qu.:0.9936
 Median :0.5983   Median :0.6667   Median :0.9984
 Mean   :0.6335   Mean   :0.6824   Mean   :0.9998
 3rd Qu.:0.7642   3rd Qu.:0.8543   3rd Qu.:1.0000
 Max.   :0.9174   Max.   :0.9174   Max.   :1.0271

mining info:
  data ntransactions support confidence
 Adult         48842     0.2        0.5
```

→ Use arulesViz package to generate some visualizations from the rules

Solution :

```
#Plotting stats
library(arulesViz)
plot(model)
```



Task 2: Recommendation Engine

→ Load the movie.csv file into R and convert the file into realRatingMatrix

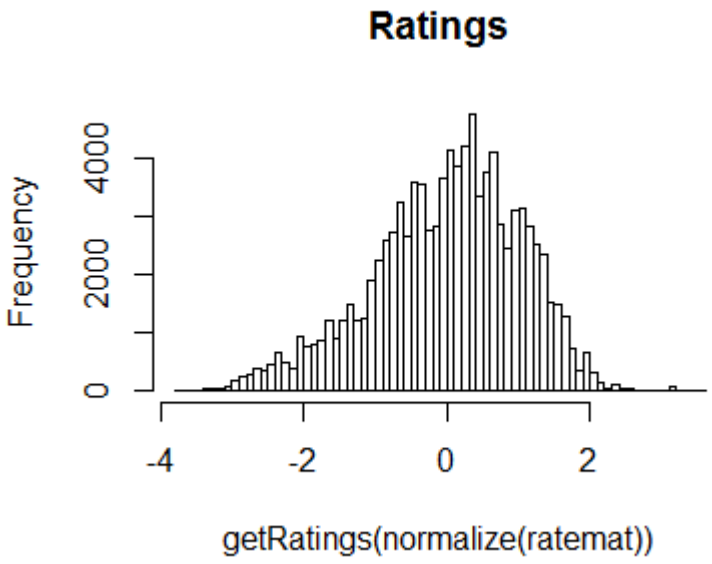
Solution :

```
#Load the file
ratings<-read.csv("movie.csv")

#Converting into realRatingMatrix
library(recommenderlab)
ratemat<-as(ratingmat,"realRatingMatrix")
head(as(ratemat,"data.frame"))
```

```
> head(as(ratemat,"data.frame"))
     user   item rating
1   User1 Movie1      5
453 User1 Movie2      3
584 User1 Movie3      4
674 User1 Movie4      3
883 User1 Movie5      3
969 User1 Movie6      5
```

→ Explore the data and plot a histogram of normalized ratings

```
#Exploring the model
dim(ratemat)
image(ratemat)
#User Ratings
hist(getRatings(normalize(ratemat)),breaks = 100, main = "Ratings")
```



→ Build a User bases collaborative filter which recommends 5 items to user100 and user101

```
#Building model
itemrec<-Recommender(ratemat,method = "UBCF")
getModel(itemrec)


#Predicting values
predrec<-predict(itemrec,ratemat[100:101],n=5)
predrec

head(as(predrec,"list"))
```

```
> head(as(predrec,"list"))
$User100
[1] "Movie332" "Movie307" "Movie311" "Movie902" "Movie984"

$User101
[1] "Movie174" "Movie172" "Movie127" "Movie56"  "Movie258"
```