General Sir John Kotelawala Defence University

Faculty of Management, Social Sciences and Humanities

Department of Languages

BSc in Applied Data Science Communication

D L S NADAVI - D/ADC/23/0028

Y M S S B SENAVIRATHNA - D/ADC/23/0030

A G A U S GUNASEKARA - D/ADC/23/0034

M L B T S PERERA - D/ADC/23/0047

Fundamentals of Data Mining / LB 2114

Year 2: Semester 1

Assignment number 2

30.04.2024

Table of contents

Page no

Association Rule Mining with Student Dataset

Association Rule Mining with Student Dataset

1) Introduction

Association rule mining is a technique used to identify hidden links between variables in huge datasets. The goal of association rule mining is to find patterns or correlations between distinct items, which can then be used to predict whether specific goods would be purchased or used together. Association rule mining has a wide range of applications, including market basket research, consumer segmentation, and fraud detection.
This report describes about students' academic performance and behavior with respect to familial and educational background, lifestyle choices and socio-economic factors. The aim of creating this report is to explore the association rules and patterns that exist and their potential impact on students' academic performance and well-being. This paper outlines all of the procedures involved in creating association rules from a data set using R in a straightforward and logical manner.

2) Data Set

The data set was taken from:
https://github.com/Emmanuel96/apriori_association_rule_mining/tree/master/Dataset

This dataset includes information about various attributes of students, with a focus on factors that may influence their academic performance and behavior. These attributes encompass a broad spectrum ranging from demographic details to familial and educational background, as well as lifestyle choices and socio-economic indicators. Each entry in the dataset corresponds to a student enrolled in a particular school, providing a rich repository of data for analysis.

3) Explanation and Preparation of the Data Set

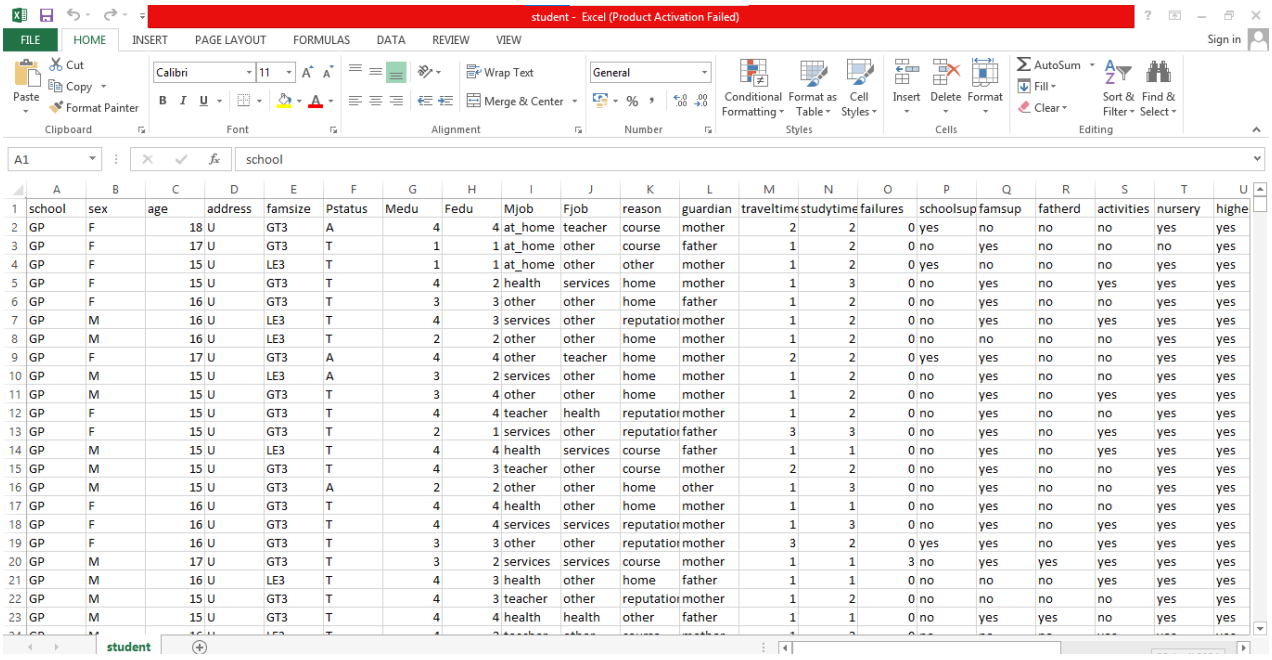
a. Explanation of the Data Set

Student data set has been used for the association rule mining task. There are 33 columns and 1046 rows in the data set.

Attributes of the data set are,

1. School - The school the student attends
2. Sex - Gender of the student (Male or Female)
3. Age - Age of the student
4. Address - Type of address of the student (urban or rural)
5. Famsize - Family size (small or large)
6. Pstatus - Parent's cohabitation status ('T' - living together, 'A' - living apart)
7. Medu - Mother's education level (1 - none, 2 - primary education (4th grade), 3 - 5th to 9th grade, 4 - secondary or higher education)
8. Fedu - Father's education level (same scale as Medu)
9. Mjob - Mother's job
10. Fjob - Father's job

11. Reason - Reason for choosing the current school
12. Guardian - Student's guardian
13. Traveltime - Home to school travel time (1 - <15 min., 2 - 15 to 30 min., 3 - 30 min. to 1 hour, or 4 - >1 hour)
14. Studytime - Weekly study time (1 - <2 hours, 2 - 2 to 5 hours, 3 - 5 to 10 hours, or 4 - >10 hours)
15. Failures - Number of past class failures
16. Schoolsup - Whether the student receives educational support from the school (yes or no)
17. Famsup - Whether the student receives educational support from the family (yes or no)
18. Fatherd - Father's educational support level (1 - low, 2 - medium, or 3 - high)
19. Activities - Extra-curricular activities participation (yes or no)
20. Nursery - Whether the student attended nursery school (yes or no)
21. Higher - Desire to pursue higher education (yes or no)
22. Internet - Internet access at home (yes or no)
23. Romantic - In a romantic relationship (yes or no)
24. Famrel - Quality of family relationships (from 1 - very bad to 5 - excellent)
25. Freetime - Free time after school (from 1 - very low to 5 - very high)
26. Goout - Going out with friends frequency (from 1 - very low to 5 - very high)
27. Dalc - Workday alcohol consumption (from 1 - very low to 5 - very high)
28. Walc - Weekend alcohol consumption (from 1 - very low to 5 - very high)
29. Health - Current health status (from 1 - very bad to 5 - very good)
30. Absences - Number of school absences
31. G1 - First period grade (from 0 to 20)
32. G2 - Second period grade (from 0 to 20)
33. G3 - Final grade (from 0 to 20)



school	sex	age	address	famsize	Pstatus	Medu	Fedu	Mjob	Fjob	reason	guardian	traveltime	studytime	failures	schoolsup	famsup	fatherd	activities	nursery	highe
GP	F	18	U	GT3	A	4	4	at_home	teacher	course	mother	2	2	0	yes	no	no	no	yes	yes
GP	F	17	U	GT3	T	1	1	at_home	other	course	father	1	2	0	no	yes	no	no	no	yes
GP	F	15	U	LE3	T	1	1	at_home	other	other	mother	1	2	0	yes	no	no	no	yes	yes
GP	F	15	U	GT3	T	4	2	health	services	home	mother	1	3	0	no	yes	no	yes	yes	yes
GP	F	16	U	GT3	T	3	3	other	other	home	father	1	2	0	no	yes	no	no	yes	yes
GP	M	16	U	LE3	T	4	3	services	other	reputatio	mother	1	2	0	no	yes	no	yes	yes	yes
GP	M	16	U	LE3	T	2	2	other	other	home	mother	1	2	0	no	no	no	no	yes	yes
GP	F	17	U	GT3	A	4	4	other	teacher	home	mother	2	2	0	yes	yes	no	no	yes	yes
GP	M	15	U	LE3	A	3	2	services	other	home	mother	1	2	0	no	yes	no	no	yes	yes
GP	M	15	U	GT3	T	3	4	other	other	home	mother	1	2	0	no	yes	no	yes	yes	yes
GP	F	15	U	GT3	T	4	4	teacher	health	reputatio	mother	1	2	0	no	yes	no	no	yes	yes
GP	F	15	U	GT3	T	2	1	services	other	reputatio	father	3	3	0	no	yes	no	yes	yes	yes
GP	M	15	U	LE3	T	4	4	health	services	course	father	1	1	0	no	yes	no	yes	yes	yes
GP	M	15	U	GT3	T	4	3	teacher	other	course	mother	2	2	0	no	yes	no	no	yes	yes
GP	M	15	U	GT3	A	2	2	other	other	home	other	1	3	0	no	yes	no	no	yes	yes
GP	F	16	U	GT3	T	4	4	health	other	home	mother	1	1	0	no	yes	no	no	yes	yes
GP	F	16	U	GT3	T	4	4	services	services	reputatio	mother	1	3	0	no	yes	no	yes	yes	yes
GP	F	16	U	GT3	T	3	3	other	other	reputatio	mother	3	2	0	yes	yes	no	yes	yes	yes
GP	M	17	U	GT3	T	3	2	services	services	course	mother	1	1	3	no	yes	yes	yes	yes	yes
GP	M	16	U	LE3	T	4	3	health	other	home	father	1	1	0	no	no	no	yes	yes	yes
GP	M	15	U	GT3	T	4	3	teacher	other	reputatio	mother	1	2	0	no	no	no	no	yes	yes
GP	M	15	U	GT3	T	4	4	health	health	other	father	1	1	0	no	yes	yes	no	yes	yes

b. Preparation of the Data Set

As the dataset is completely suitable for do association rule mining and has no NULL values in the dataset, we didn't had much work to do to prepare the dataset. Therefore, first we read and understood the dataset and applied the association rule mining into the dataset using R software.

4) Association Rule Mining

Association rule mining is a type of unsupervised machine learning technique that discovers connections between two or more items in large datasets. It was proposed by Agrawal et al in 1993. It's a popular system in data mining which has a wide range of operations in various fields, such as request market basket analysis, customer segmentation, and fraud discovery. The two most important measures used in association rule mining are support and confidence.

- Support: This measures how frequently the particulars in the rule appear together in the dataset. A high support value indicates that the rule is constantly being.
- Confidence: This measures how likely it's that the consequent item will do if the precedent item occurs. Strong rules are indicated by a high confidence value.

A third metric called lift, can be used to compare confidence with anticipated confidence, or how numerous times an if- also statement is anticipated to be set up true.

5) Implementation in R

Packages used

1) **arules:** A complete R package for mining association rules and frequent item sets from transaction data is called `arules`. The association rules that describe the relationships between items in transactional datasets can be generated and evaluated by this package. Recommendation systems, market basket analysis, and other applications involving transactional data analysis frequently use this package.
2) **arulesviz:** Specifically created for the purpose of visualizing association rules and item sets, the `arulesviz` package is an extension of the `arules` package. To assist users in exploring and interpreting the outcomes of association rule mining, it provides a range of visualization techniques. Scatter plots, matrix plots, and graph-based representations of item sets and rules are some examples of these visualizations.

Explanation of the experimental procedure and Visualization of the results

Step 01

Import the dataset.

```
> #import data set
> data=read.csv("student.csv",header=T, colClasses="factor")
> data
   school sex age address famsize Pstatus Medu Fedu     Mjob     Fjob
1      GP   F  18       U     GT3       A    4    4  at_home  teacher
2      GP   F  17       U     GT3       T    1    1  at_home    other
3      GP   F  15       U     LE3       T    1    1  at_home    other
4      GP   F  15       U     GT3       T    4    2   health services
5      GP   F  16       U     GT3       T    3    3    other    other
6      GP   M  16       U     LE3       T    4    3 services    other
7      GP   M  16       U     LE3       T    2    2    other    other
8      GP   F  17       U     GT3       A    4    4    other  teacher
9      GP   M  15       U     LE3       A    3    2 services    other
10     GP   M  15       U     GT3       T    3    4    other    other
11     GP   F  15       U     GT3       T    4    4  teacher   health
12     GP   F  15       U     GT3       T    2    1 services    other
13     GP   M  15       U     LE3       T    4    4   health services
```

Step 02

Use the ‘name ()’ function to get the column names of the dataset.

```
> names(data)
 [1] "school"     "sex"        "age"        "address"    "famsize"
 [6] "Pstatus"    "Medu"       "Fedu"       "Mjob"       "Fjob"
[11] "reason"     "guardian"   "traveltime" "studytime"  "failures"
[16] "schoolsup"  "famsup"     "fatherd"    "activities" "nursery"
[21] "higher"     "internet"   "romantic"   "famrel"     "freetime"
[26] "goout"      "Dalc"       "Walc"       "health"     "absences"
[31] "G1"         "G2"         "G3"
>
```

Step 03

Use ‘head ()’ and ‘tail ()’ functions to get first and last 6 rows in the dataset.

```
> head(data)
  school sex age address famsize Pstatus Medu Fedu     Mjob     Fjob
1     GP   F  18       U     GT3       A    4    4  at_home  teacher
2     GP   F  17       U     GT3       T    1    1  at_home    other
3     GP   F  15       U     LE3       T    1    1  at_home    other
4     GP   F  15       U     GT3       T    4    2   health services
5     GP   F  16       U     GT3       T    3    3    other    other
6     GP   M  16       U     LE3       T    4    3 services    other
      reason guardian traveltime studytime failures schoolsup famsup
1     course   mother          2         2        0       yes     no
2     course   father          1         2        0        no    yes
3      other   mother          1         2        0       yes     no
4       home   mother          1         3        0        no    yes
5       home   father          1         2        0        no    yes
6 reputation   mother          1         2        0        no    yes
  fatherd activities nursery higher internet romantic famrel freetime
1      no         no     yes    yes       no       no      4        3
2      no         no      no    yes      yes       no      5        3
3      no         no     yes    yes      yes       no      4        3
4      no        yes     yes    yes      yes      yes      3        2
5      no         no     yes    yes       no       no      4        3
6      no        yes     yes    yes      yes       no      5        4
  goout Dalc Walc health absences G1 G2 G3
1     4    1    1      3        4  0 11 11
2     3    1    1      3        2  9 11 11
3     2    2    3      3        6 12 13 12
4     2    1    1      5        0 14 14 14
5     2    1    2      5        0 11 13 13
6     2    1    2      5        6 12 12 13
>
```

```
> tail(data)
     school sex age address famsize Pstatus Medu Fedu     Mjob     Fjob reason guardian
1040     MS   F  18       U     GT3       T    1    1    other    other course   mother
1041     MS   M  20       U     LE3       A    2    2 services services course    other
1042     MS   M  17       U     LE3       T    3    1 services services course   mother
1043     MS   M  21       R     GT3       T    1    1    other    other course    other
1044     MS   M  18       R     LE3       T    3    2 services    other course   mother
1045     MS   M  19       U     LE3       T    1    1    other  at_home course   father
     traveltime studytime failures schoolsup famsup fatherd activities nursery higher
1040          2         2        1        no     no      no        yes     yes    yes
1041          1         2        2        no    yes     yes         no     yes    yes
1042          2         1        0        no     no      no         no      no    yes
1043          1         1        3        no     no      no         no      no    yes
1044          3         1        0        no     no      no         no      no    yes
1045          1         1        0        no     no      no         no     yes    yes
     internet romantic famrel freetime goout Dalc Walc health absences G1 G2 G3
1040       no       no      1        1     1    1    1      5        0  6  5  0
1041       no       no      5        5     4    4    5      4       11  9  9  9
1042      yes       no      2        4     5    3    4      2        3 14 16 16
1043       no       no      5        5     3    3    3      3        3 10  8  7
1044      yes       no      4        4     1    3    4      5        0 11 12 10
1045      yes       no      3        2     3    3    3      5        5  8  9  9
>
```

Step 04

Use the 'summary ()' function to get the summary of the dataset.

```
> summary(data)
    school     sex          age          address       famsize       Pstatus    Medu
 GP    :772   F  :591   16     :281   address:  1   famsize:  1   A      :121    :  1
 MS    :272   M  :453   17     :277   R      :285   GT3    :738   Pstatus:  1   0:  9
 school:  1   sex:  1   18     :222   U      :759   LE3    :306   T      :923   1:202
                        15     :194                                             2:289
                        19     : 56                                             3:238
                        20     :  9                                             4:306
                        (Other):  6
 Fedu          Mjob           Fjob           reason        guardian   traveltime
  :  1   at_home :194   at_home : 62   course    :430   father  :243    :  1
 0:  9   health  : 82   Fjob    :  1   home      :258   guardian:  1   1:623
 1:256   Mjob    :  1   health  : 41   other     :108   mother  :728   2:320
 2:324   other   :399   other   :584   reason    :  1   other   : 73   3: 77
 3:231   services:239   services:292   reputation:248                  4: 24
 4:224   teacher :130   teacher : 65

 studytime failures     schoolsup      famsup    fatherd         activities     nursery
  :  1      :  1    no       :925   famsup:  1   no  :824   activities:  1   no     :209
 1:317     0:861    schoolsup:  1   no    :404   paid:  1   no        :528   nursery:  1
 2:503     1:120    yes      :119   yes   :640   yes :220   yes       :516   yes    :835
 3:162     2: 33
 4: 62     3: 30
```

Step 05

Use the 'str ()' function to get the structure of the dataset.

```
> str(data)
'data.frame':   1045 obs. of  33 variables:
 $ school    : Factor w/ 3 levels "GP","MS","school": 1 1 1 1 1 1 1 1 1 1 ...
 $ sex       : Factor w/ 3 levels "F","M","sex": 1 1 1 1 1 2 2 1 2 2 ...
 $ age       : Factor w/ 9 levels "","15","16","17",..: 5 4 2 2 3 3 3 4 2 2 ...
 $ address   : Factor w/ 3 levels "address","R",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ famsize   : Factor w/ 3 levels "famsize","GT3",..: 2 2 3 2 2 3 3 2 3 2 ...
 $ Pstatus   : Factor w/ 3 levels "A","Pstatus",..: 1 3 3 3 3 3 3 1 1 3 ...
 $ Medu      : Factor w/ 6 levels "","0","1","2",..: 6 3 3 6 5 6 4 6 5 5 ...
 $ Fedu      : Factor w/ 6 levels "","0","1","2",..: 6 3 3 4 5 5 4 6 4 6 ...
 $ Mjob      : Factor w/ 6 levels "at_home","health",..: 1 1 1 2 4 5 4 4 5 4 ...
 $ Fjob      : Factor w/ 6 levels "at_home","Fjob",..: 6 4 4 5 4 4 4 6 4 4 ...
 $ reason    : Factor w/ 5 levels "course","home",..: 1 1 3 2 2 5 2 2 2 2 ...
 $ guardian  : Factor w/ 4 levels "father","guardian",..: 3 1 3 3 1 3 3 3 3 3 ...
 $ traveltime: Factor w/ 5 levels "","1","2","3",..: 3 2 2 2 2 2 2 3 2 2 ...
 $ studytime : Factor w/ 5 levels "","1","2","3",..: 3 3 3 4 3 3 3 3 3 3 ...
 $ failures  : Factor w/ 5 levels "","0","1","2",..: 2 2 2 2 2 2 2 2 2 2 ...
 $ schoolsup : Factor w/ 3 levels "no","schoolsup",..: 3 1 3 1 1 1 1 3 1 1 ...
 $ famsup    : Factor w/ 3 levels "famsup","no",..: 2 3 2 3 3 3 2 3 3 3 ...
 $ fatherd   : Factor w/ 3 levels "no","paid","yes": 1 1 1 1 1 1 1 1 1 1 ...
 $ activities: Factor w/ 3 levels "activities","no",..: 2 2 2 3 2 3 2 2 2 3 ...
 $ nursery   : Factor w/ 3 levels "no","nursery",..: 3 1 3 3 3 3 3 3 3 3 ...
 $ higher    : Factor w/ 3 levels "higher","no",..: 3 3 3 3 3 3 3 3 3 3 ...
 $ internet  : Factor w/ 3 levels "internet","no",..: 2 3 3 3 2 3 3 2 3 3 ...
 $ romantic  : Factor w/ 3 levels "no","romantic",..: 1 1 1 3 1 1 1 1 1 1 ...
 $ famrel    : Factor w/ 6 levels "","1","2","3",..: 5 6 5 4 5 6 5 5 5 6 ...
 $ freetime  : Factor w/ 6 levels "","1","2","3",..: 4 4 4 3 4 5 5 2 3 6 ...
```

Step 06

Use the 'dim ()' function to get the dimension of the data set which includes the number of rows and columns in the data set.

```
> dim(data)
[1] 1045   33
> |
```

Step 07

Get columns for association rule mining.

```
> #get columns for Association Rule Mining
> data_new=data[,16:23]
> View(data_new)
> |
```

ARM.R* × data_new ×

Filter

	schoolsup	famsup	fatherd	activities	nursery	higher	internet	romantic
1	yes	no	no	no	yes	yes	no	no
2	no	yes	no	no	no	yes	yes	no
3	yes	no	no	no	yes	yes	yes	no
4	no	yes	no	yes	yes	yes	yes	yes
5	no	yes	no	no	yes	yes	no	no
6	no	yes	no	yes	yes	yes	yes	no
7	no	no	no	no	yes	yes	yes	no
8	yes	yes	no	no	yes	yes	no	no
9	no	yes	no	no	yes	yes	yes	no
10	no	yes	no	yes	yes	yes	yes	no

Showing 1 to 11 of 1,045 entries, 8 total columns

Step 08

Use colSums () function to compute the sums of columns.

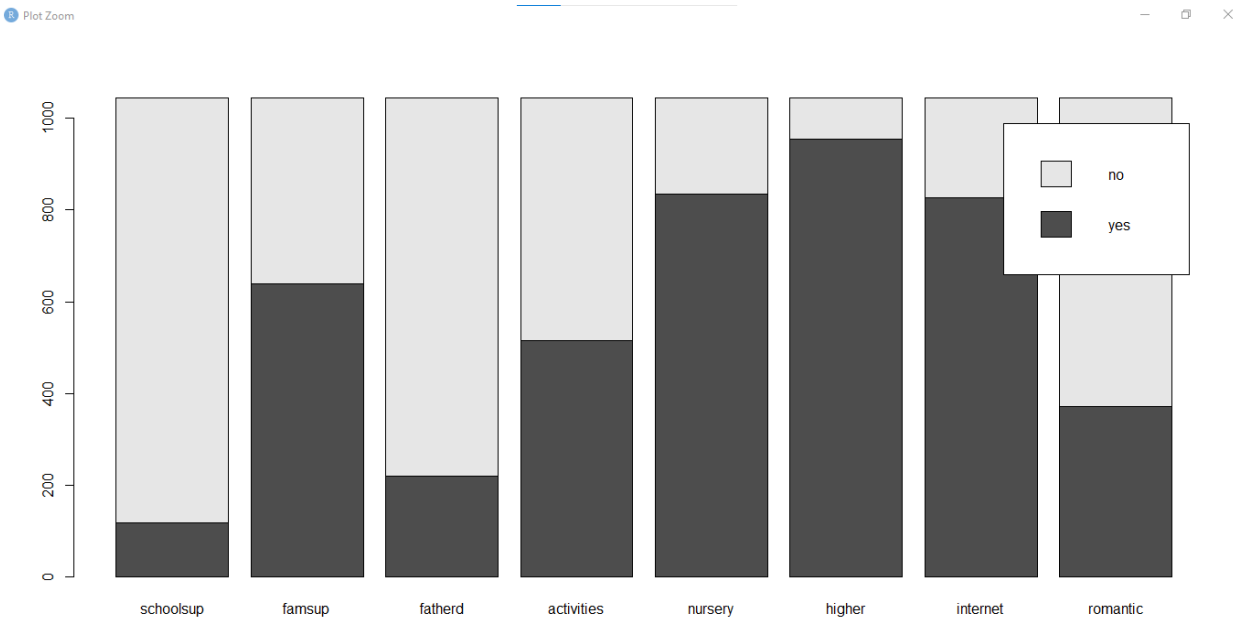
```
> #  Only YES columns
> yes=colSums(data_new=="yes")
> yes
 schoolsup     famsup    fatherd activities    nursery     higher   internet   romantic
       119        640        220        516        835        955        827        371
> |
```

```
> # Only NO columns
> no=colSums(data_new=="no")
> no
 schoolsup     famsup    fatherd activities    nursery     higher   internet   romantic
       925        404        824        528        209         89        217        673
> 

> #Get both YES & NO columns
> sub=rbind(yes,no)
> sub
    schoolsup famsup fatherd activities nursery higher internet romantic
yes       119    640      220        516     835    955      827      371
no        925    404      824        528     209     89      217      673
> 
```

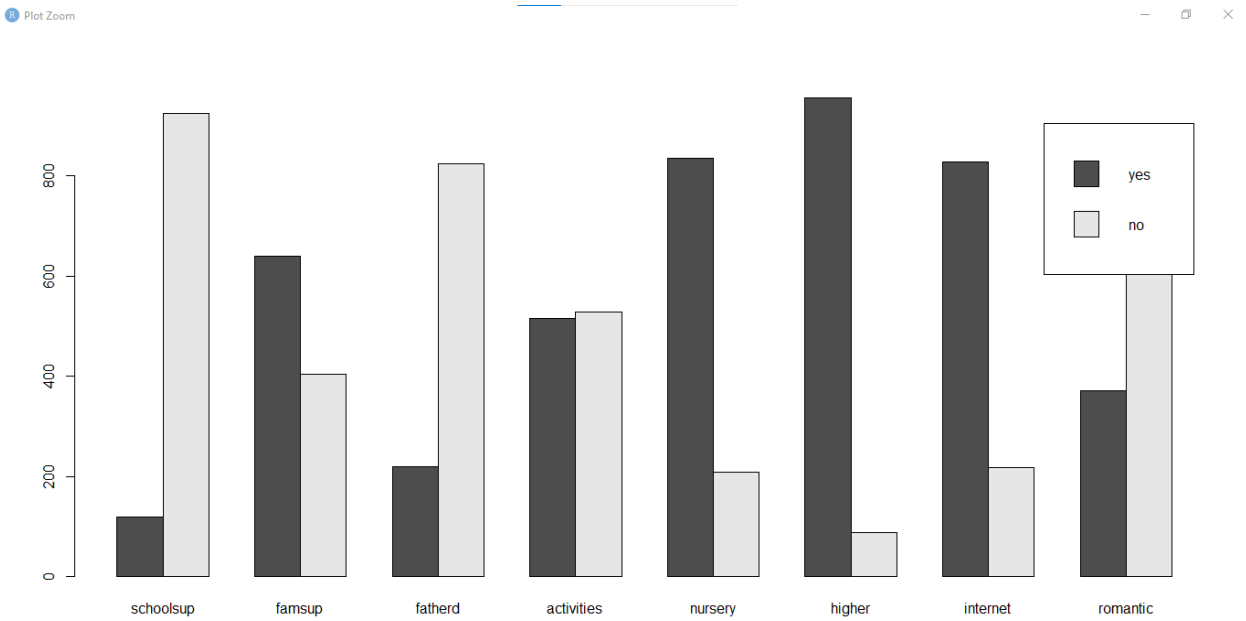
Step 09

Plot and explore the "student" dataset with barplot () function.

```
> barplot(sub,legend=rownames(sub))
> 
```



```
> barplot(sub,beside = T,legend=rownames(sub))
>
```



Step 10

Install and activate “arules” package.

```
> library(arules)
Loading required package: Matrix

Attaching package: 'arules'

The following objects are masked from 'package:base':

    abbreviate, write

Warning messages:
1: package 'arules' was built under R version 4.2.3
2: package 'Matrix' was built under R version 4.2.3
>
```

Step 11

Create Association rules.

```
> #Get the rules
> rules=apriori(data_new)
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen
        0.8    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: 104

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[24 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [15 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [1036 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

Get the summary of these rules.

In here, we got 1036 rules associated with the student dataset.

```
> summary(rules)
set of 1036 rules

rule length distribution (lhs + rhs):sizes
  1   2   3   4   5   6   7
  2  42 179 333 311 143  26

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.000   4.000   4.000   4.392   5.000   7.000

summary of quality measures:
    support         confidence        coverage            lift            count
 Min.   :0.1005   Min.   :0.8000   Min.   :0.1005   Min.   :0.8802   Min.   :105.0
 1st Qu.:0.1273   1st Qu.:0.8427   1st Qu.:0.1455   1st Qu.:0.9836   1st Qu.:133.0
 Median :0.1684   Median :0.8742   Median :0.1919   Median :1.0158   Median :176.0
 Mean   :0.2119   Mean   :0.8777   Mean   :0.2422   Mean   :1.0178   Mean   :221.4
 3rd Qu.:0.2519   3rd Qu.:0.9091   3rd Qu.:0.2900   3rd Qu.:1.0463   3rd Qu.:263.2
 Max.   :0.9139   Max.   :1.0000   Max.   :1.0000   Max.   :1.1749   Max.   :955.0

mining info:
     data ntransactions support confidence                    call
 data_new          1045     0.1        0.8 apriori(data = data_new)
>
```

Inspect the rules.

```
> inspect(rules)
```

```
        romantic=no}     => {schoolsup=no} 0.1224880  0.8767123 0.1397129 0.9904480   128
[1032] {schoolsup=no,
        fatherd=no,
        activities=no,
        higher=yes,
        internet=yes,
        romantic=no}     => {nursery=yes}  0.1224880  0.8000000 0.1531100 1.0011976   128
[1033] {schoolsup=no,
        famsup=yes,
        fatherd=no,
        nursery=yes,
        internet=yes,
        romantic=no}     => {higher=yes}   0.1550239  0.9818182 0.1578947 1.0743455   162
[1034] {famsup=yes,
        fatherd=no,
        nursery=yes,
        higher=yes,
        internet=yes,
        romantic=no}     => {schoolsup=no} 0.1550239  0.8709677 0.1779904 0.9839582   162
[1035] {schoolsup=no,
        famsup=yes,
        fatherd=no,
        higher=yes,
        internet=yes,
        romantic=no}     => {nursery=yes}  0.1550239  0.8393782 0.1846890 1.0504794   162
[1036] {schoolsup=no,
        famsup=yes,
        fatherd=no,
        nursery=yes,
        higher=yes,
        romantic=no}     => {internet=yes} 0.1550239  0.8019802 0.1933014 1.0133849   162
>
```

Since there are too many rules we need to reduce them into smaller number of rules. Therefore, we specify the parameters as follows to get the rules under the confidence of 0.9

```
> #Get the rules under the confidence of 0.9
> rules_1=apriori(data_new,parameter = list(conf=0.9))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen
        0.9    0.1    1 none FALSE            TRUE       5     0.1      1     10
 target  ext
  rules TRUE

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

Absolute minimum support count: 104

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[24 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [15 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [313 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

Get the summary of the rules.

In here, we got 313 rules associated with the student dataset.

```
> summary(rules_1)
set of 313 rules

rule length distribution (lhs + rhs):sizes
  1   2   3   4   5   6   7
  1  11  53 101  99  40   8

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.000   4.000   4.000   4.399   5.000   7.000

summary of quality measures:
    support         confidence        coverage           lift            count
 Min.   :0.1005   Min.   :0.9000   Min.   :0.1005   Min.   :0.9848   Min.   :105.0
 1st Qu.:0.1301   1st Qu.:0.9150   1st Qu.:0.1378   1st Qu.:1.0168   1st Qu.:136.0
 Median :0.1742   Median :0.9298   Median :0.1847   Median :1.0323   Median :182.0
 Mean   :0.2164   Mean   :0.9358   Mean   :0.2319   Mean   :1.0358   Mean   :226.1
 3rd Qu.:0.2622   3rd Qu.:0.9548   3rd Qu.:0.2756   3rd Qu.:1.0554   3rd Qu.:274.0
 Max.   :0.9139   Max.   :1.0000   Max.   :1.0000   Max.   :1.1749   Max.   :955.0

mining info:
     data ntransactions support confidence
 data_new          1045     0.1        0.9
                                                call
 apriori(data = data_new, parameter = list(conf = 0.9))
>
```

Inspect the rules.

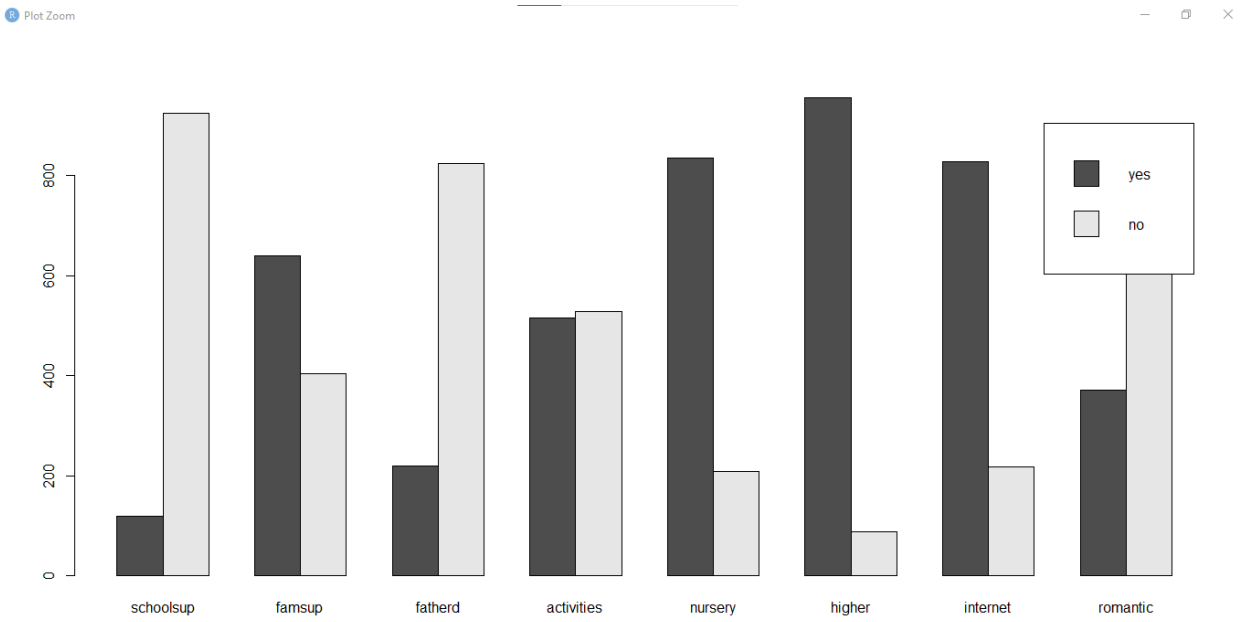
```
[310] {schoolsup=no,
       famsup=yes,
       activities=no,
       nursery=yes,
       internet=yes,
       romantic=no}     => {higher=yes}   0.1023923  0.9907407 0.1033493 1.0841090
107
[311] {schoolsup=no,
       famsup=yes,
       fatherd=no,
       activities=no,
       nursery=yes,
       internet=yes}    => {higher=yes}   0.1110048  0.9133858 0.1215311 0.9994641
116
[312] {schoolsup=no,
       fatherd=no,
       activities=no,
       nursery=yes,
       internet=yes,
       romantic=no}     => {higher=yes}   0.1224880  0.9014085 0.1358852 0.9863579
128
[313] {schoolsup=no,
       famsup=yes,
       fatherd=no,
       nursery=yes,
       internet=yes,
       romantic=no}     => {higher=yes}   0.1550239  0.9818182 0.1578947 1.0743455
162
> |
```

As we are focused on the factors that may influence the academic performance and behavior of the students, we need rules with “Yes”. For that we need to get the summary of the dataset to find the highest number of yes results.

```
> summary(data_new)
     schoolsup      famsup    fatherd         activities     nursery
 no       :925   famsup:  1   no  :824   activities:  1   no     :209
 schoolsup:  1   no    :404   paid:  1   no        :528   nursery:  1
 yes      :119   yes   :640   yes :220   yes       :516   yes    :835
    higher        internet       romantic
 higher:  1   internet:  1   no      :673
 no    : 89   no      :217   romantic:  1
 yes   :955   yes     :827   yes     :371
> |
```

Creating a plot is the easiest way to find the highest number of yes results.

```
> barplot(sub,beside = T,legend=rownames(sub))
> |
```



According to the plot "higher" (desire to pursue higher education) contains the highest number of yes results. Since we want to see rules where students are desiring to pursue higher education we use the following code to get those rules under the confidence of 0.7

```
> #Get the rules under the confidence of 0.7
> rules_1_new=apriori(data_new,parameter = list(conf=0.7),
+                     appearance = list(rhs=c("higher=yes"),default="lhs"))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target
        0.7    0.1    1 none FALSE            TRUE       5     0.1      1     10  rules
  ext
 TRUE

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

Absolute minimum support count: 104

set item appearances ...[1 item(s)] done [0.00s].
set transactions ...[24 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [15 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [345 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

Get the summary of the above rules.

In here, we got 345 rules.

```
> summary(rules_1_new)
set of 345 rules

rule length distribution (lhs + rhs):sizes
  1   2   3   4   5   6   7
  1  14  59 111 103  49   8

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.000   4.000   4.000   4.391   5.000   7.000

summary of quality measures:
    support         confidence        coverage           lift            count
 Min.   :0.1005   Min.   :0.7955   Min.   :0.1005   Min.   :0.8704   Min.   :105
 1st Qu.:0.1292   1st Qu.:0.8832   1st Qu.:0.1416   1st Qu.:0.9664   1st Qu.:135
 Median :0.1751   Median :0.9186   Median :0.1914   Median :1.0052   Median :183
 Mean   :0.2201   Mean   :0.9161   Mean   :0.2402   Mean   :1.0024   Mean   :230
 3rd Qu.:0.2622   3rd Qu.:0.9475   3rd Qu.:0.2880   3rd Qu.:1.0368   3rd Qu.:274
 Max.   :0.9139   Max.   :1.0000   Max.   :1.0000   Max.   :1.0942   Max.   :955

mining info:
     data ntransactions support confidence
 data_new          1045     0.1        0.7

call
 apriori(data = data_new, parameter = list(conf = 0.7), appearance = list(rhs = c("higher=ye
s"), default = "lhs"))
>
```

Inspect the rules.

```
> inspect(rules_1_new)
```

```
32
[342] {schoolsup=no,
       famsup=yes,
       activities=no,
       nursery=yes,
       internet=yes,
       romantic=no}     => {higher=yes} 0.1023923  0.9907407 0.1033493 1.0841090   1
07
[343] {schoolsup=no,
       famsup=yes,
       fatherd=no,
       activities=no,
       nursery=yes,
       internet=yes}    => {higher=yes} 0.1110048  0.9133858 0.1215311 0.9994641   1
16
[344] {schoolsup=no,
       fatherd=no,
       activities=no,
       nursery=yes,
       internet=yes,
       romantic=no}     => {higher=yes} 0.1224880  0.9014085 0.1358852 0.9863579   1
28
[345] {schoolsup=no,
       famsup=yes,
       fatherd=no,
       nursery=yes,
       internet=yes,
       romantic=no}     => {higher=yes} 0.1550239  0.9818182 0.1578947 1.0743455   1
62
```

Get the rules under the confidence of 0.8

```
> #Get the rules under the confidence of 0.8
> rules_2=apriori(data_new,parameter = list(conf=0.8),
+                 appearance = list(rhs=c("higher=yes"),default="lhs"))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen
        0.8    0.1    1 none FALSE            TRUE       5     0.1      1     10
 target  ext
  rules TRUE

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

Absolute minimum support count: 104

set item appearances ...[1 item(s)] done [0.00s].
set transactions ...[24 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [15 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [344 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

Get the summary of the rules.

In here, we got 344 rules.

```
> summary(rules_2)
set of 344 rules

rule length distribution (lhs + rhs):sizes
  1   2   3   4   5   6   7
  1  14  59 111 102  49   8

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   1.00    4.00    4.00    4.39    5.00    7.00

summary of quality measures:
    support         confidence        coverage           lift
 Min.   :0.1005   Min.   :0.8043   Min.   :0.1005   Min.   :0.8802
 1st Qu.:0.1292   1st Qu.:0.8836   1st Qu.:0.1423   1st Qu.:0.9668
 Median :0.1756   Median :0.9187   Median :0.1919   Median :1.0053
 Mean   :0.2205   Mean   :0.9164   Mean   :0.2405   Mean   :1.0028
 3rd Qu.:0.2622   3rd Qu.:0.9477   3rd Qu.:0.2883   3rd Qu.:1.0370
 Max.   :0.9139   Max.   :1.0000   Max.   :1.0000   Max.   :1.0942
     count
 Min.   :105.0
 1st Qu.:135.0
 Median :183.5
 Mean   :230.4
 3rd Qu.:274.0
 Max.   :955.0

mining info:
     data ntransactions support confidence
 data_new          1045     0.1        0.8

call
 apriori(data = data_new, parameter = list(conf = 0.8), appearance = list(rhs = c("h
igher=yes"), default = "lhs"))
>
```

Inspect the rules.

```
> inspect(rules_2)
```

```
07
[342] {schoolsup=no,
       famsup=yes,
       fatherd=no,
       activities=no,
       nursery=yes,
       internet=yes}     => {higher=yes} 0.1110048  0.9133858 0.1215311 0.9994641   1
16
[343] {schoolsup=no,
       fatherd=no,
       activities=no,
       nursery=yes,
       internet=yes,
       romantic=no}      => {higher=yes} 0.1224880  0.9014085 0.1358852 0.9863579   1
28
[344] {schoolsup=no,
       famsup=yes,
       fatherd=no,
       nursery=yes,
       internet=yes,
       romantic=no}      => {higher=yes} 0.1550239  0.9818182 0.1578947 1.0743455   1
62
>
```

Get the rules under the confidence of 0.9

```
> #Get the rules under the confidence of 0.9
> rules_3=apriori(data_new,parameter = list(conf=0.9),
+                 appearance = list(rhs=c("higher=yes"),default="lhs"))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target
        0.9    0.1    1 none FALSE            TRUE       5     0.1      1     10  rules
  ext
 TRUE

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

Absolute minimum support count: 104

set item appearances ...[1 item(s)] done [0.00s].
set transactions ...[24 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [15 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [215 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

Get the summary of the rules.

In here, we got 215 rules.

```
> summary(rules_3)
set of 215 rules

rule length distribution (lhs + rhs):sizes
 1  2  3  4  5  6  7
 1  8 36 62 64 36  8

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.000   4.000   5.000   4.488   5.000   7.000

summary of quality measures:
    support         confidence        coverage            lift            count
 Min.   :0.1005   Min.   :0.9000   Min.   :0.1005   Min.   :0.9848   Min.   :105.0
 1st Qu.:0.1388   1st Qu.:0.9221   1st Qu.:0.1464   1st Qu.:1.0090   1st Qu.:145.0
 Median :0.1923   Median :0.9407   Median :0.2019   Median :1.0294   Median :201.0
 Mean   :0.2402   Mean   :0.9436   Mean   :0.2562   Mean   :1.0325   Mean   :251.0
 3rd Qu.:0.2933   3rd Qu.:0.9663   3rd Qu.:0.3120   3rd Qu.:1.0574   3rd Qu.:306.5
 Max.   :0.9139   Max.   :1.0000   Max.   :1.0000   Max.   :1.0942   Max.   :955.0

mining info:
     data ntransactions support confidence
 data_new          1045     0.1        0.9

call
 apriori(data = data_new, parameter = list(conf = 0.9), appearance = list(rhs = c("higher=ye
s"), default = "lhs"))
>
```

Inspect the rules

```
> inspect(rules_3)
      lhs                    rhs            support confidence  coverage      lift count
[1]   {}                  => {higher=yes} 0.9138756  0.9138756 1.0000000 1.0000000   955
[2]   {schoolsup=yes}     => {higher=yes} 0.1110048  0.9747899 0.1138756 1.0666549   116
[3]   {fatherd=yes}       => {higher=yes} 0.2066986  0.9818182 0.2105263 1.0743455   216
[4]   {activities=yes}    => {higher=yes} 0.4602871  0.9321705 0.4937799 1.0200191   481
[5]   {famsup=yes}        => {higher=yes} 0.5722488  0.9343750 0.6124402 1.0224313   598
[6]   {romantic=no}       => {higher=yes} 0.6028708  0.9361070 0.6440191 1.0243265   630
[7]   {internet=yes}      => {higher=yes} 0.7311005  0.9238210 0.7913876 1.0108827   764
[8]   {nursery=yes}       => {higher=yes} 0.7358852  0.9209581 0.7990431 1.0077499   769
[9]   {schoolsup=no}      => {higher=yes} 0.8028708  0.9070270 0.8851675 0.9925060   839
[10]  {famsup=yes,
       nursery=no}        => {higher=yes} 0.1062201  0.9250000 0.1148325 1.0121728   111
[11]  {nursery=no,
       internet=yes}      => {higher=yes} 0.1454545  0.9156627 0.1588517 1.0019555   152
[12]  {internet=no,
       romantic=no}       => {higher=yes} 0.1291866  0.9000000 0.1435407 0.9848168   135
[13]  {nursery=yes,
       internet=no}       => {higher=yes} 0.1502392  0.9022989 0.1665072 0.9873323   157
[14]  {fatherd=yes,
       activities=yes}    => {higher=yes} 0.1062201  0.9736842 0.1090909 1.0654450   111
[15]  {fatherd=yes,
       activities=no}     => {higher=yes} 0.1004785  0.9905660 0.1014354 1.0839178   105
[16]  {famsup=yes,
       fatherd=yes}       => {higher=yes} 0.1598086  0.9823529 0.1626794 1.0749307   167
```

Step 12

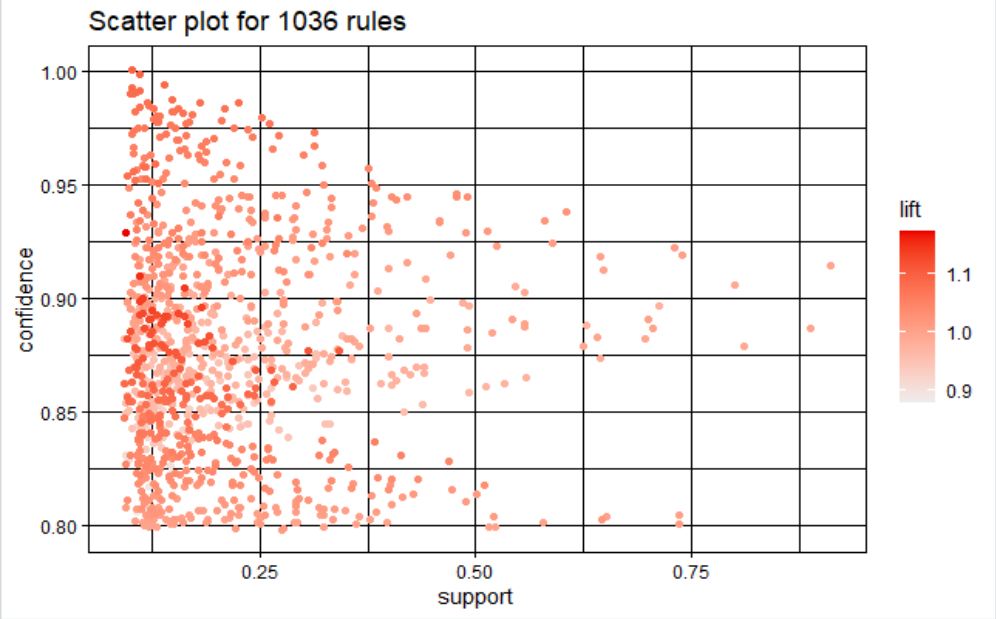
Visualize these rules.

```
> library(arulesViz)
Warning message:
package 'arulesViz' was built under R version 4.2.3
>
```

Step 13

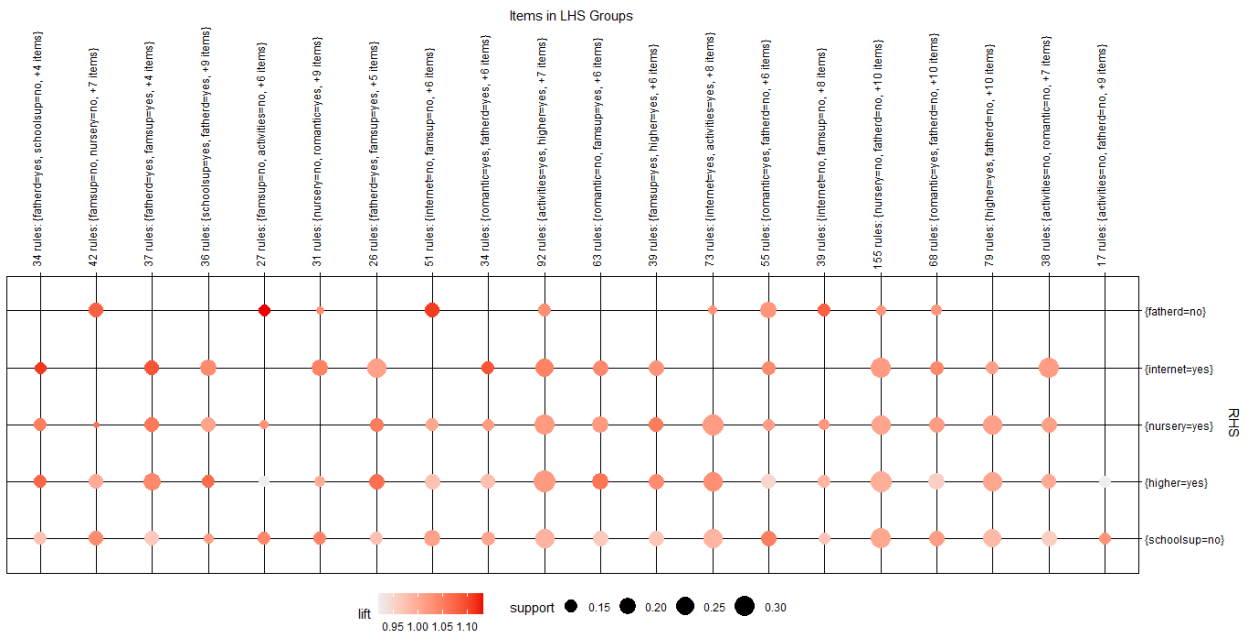
Plot the rules.

```
> plot(rules)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
>
```



Plot the rules in groups

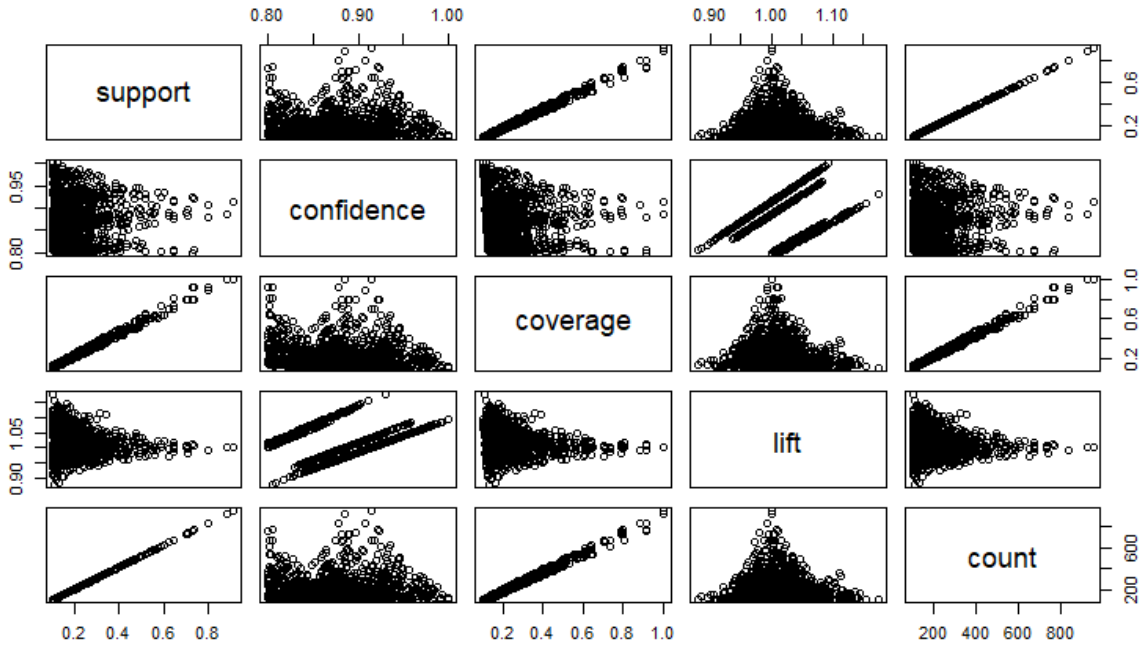
```
> plot(rules,method = "grouped")
>
```



Step 14

Display a scatterplot matrix to compare the support, confidence, and lift.

```
> plot(rules@quality)
>
```



To determine the factors influencing students' desire for higher education with a confidence level of 0.7, we'll analyze the following variables: educational support from the school, father's educational support level, participation in extracurricular activities, attendance at nursery school, internet access at home, and being in a romantic relationship.

```
> rules_new=apriori(data_new,parameter=list(conf=0.7),
+                   appearance=list(rhs=c("higher=yes"),
+                                   lhs=c("schoolsup=no","fatherd=no","activities=n
o","nursery=yes","internet=yes","romantic=no"),
+                                   default="none"))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target
        0.7    0.1    1 none FALSE            TRUE       5     0.1      1     10  rules
  ext
 TRUE

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

Absolute minimum support count: 104

set item appearances ...[7 item(s)] done [0.00s].
set transactions ...[7 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [7 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [64 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

Get the summary of the above rules.

In here we got 64 rules that implies the above mentioned conditions in the students' data set.

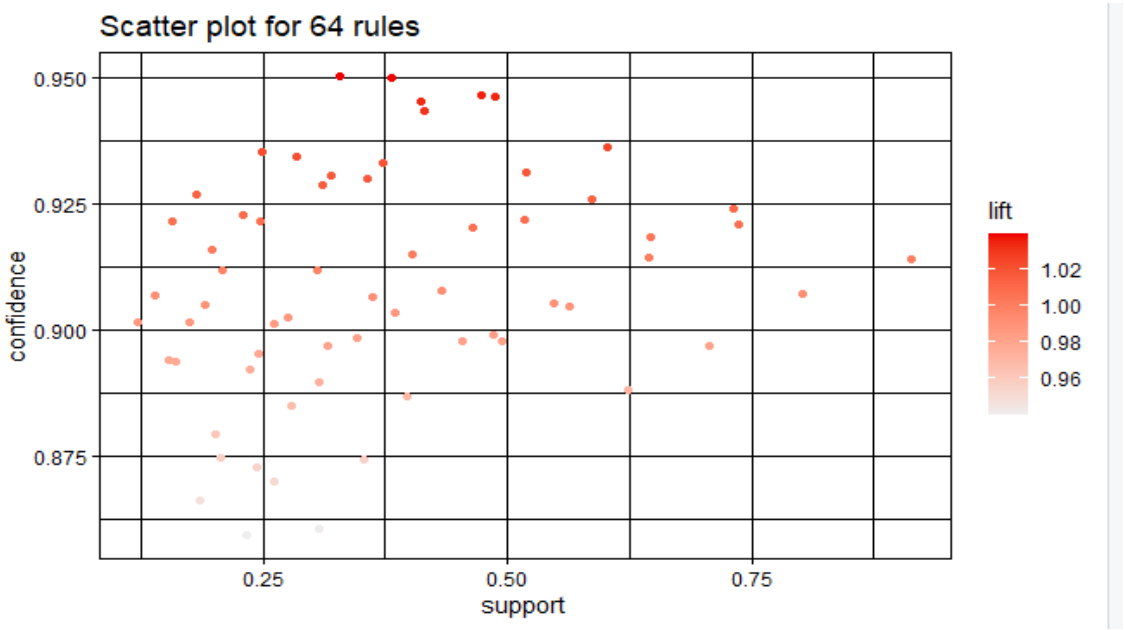
```
> summary(rules_new)
set of 64 rules

rule length distribution (lhs + rhs):sizes
 1  2  3  4  5  6  7
 1  6 15 20 15  6  1

   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
      1       3       4       4       5       7

summary of quality measures:
    support         confidence        coverage          lift
 Min.   :0.1225   Min.   :0.8592   Min.   :0.1359   Min.   :0.9401
 1st Qu.:0.2416   1st Qu.:0.8963   1st Qu.:0.2675   1st Qu.:0.9808
 Median :0.3378   Median :0.9069   Median :0.3699   Median :0.9924
 Mean   :0.3772   Mean   :0.9088   Mean   :0.4140   Mean   :0.9945
 3rd Qu.:0.4859   3rd Qu.:0.9243   3rd Qu.:0.5218   3rd Qu.:1.0114
 Max.   :0.9139   Max.   :0.9503   Max.   :1.0000   Max.   :1.0398
     count
 Min.   :128.0
 1st Qu.:252.5
 Median :353.0
 Mean   :394.1
 3rd Qu.:507.8
 Max.   :955.0

mining info:
     data ntransactions support confidence
 data_new          1045     0.1        0.7
```

Inspect the rules.

```
> inspect(rules_new)
     lhs                    rhs             support   confidence  coverage   lift      count
[1]  {}                  => {higher=yes} 0.9138756 0.9138756 1.0000000 1.0000000   955
[2]  {activities=no}     => {higher=yes} 0.4535885 0.8977273 0.5052632 0.9823298   474
[3]  {romantic=no}       => {higher=yes} 0.6028708 0.9361070 0.6440191 1.0243265   630
[4]  {fatherd=no}        => {higher=yes} 0.7071770 0.8968447 0.7885167 0.9813641   739
[5]  {internet=yes}      => {higher=yes} 0.7311005 0.9238210 0.7913876 1.0108827   764
[6]  {nursery=yes}       => {higher=yes} 0.7358852 0.9209581 0.7990431 1.0077499   769
[7]  {schoolsup=no}      => {higher=yes} 0.8028708 0.9070270 0.8851675 0.9925060   839
[8]  {activities=no,
      romantic=no}       => {higher=yes} 0.3062201 0.9116809 0.3358852 0.9975985   320
[9]  {fatherd=no,
      activities=no}     => {higher=yes} 0.3531100 0.8744076 0.4038278 0.9568125   369
[10] {activities=no,
      internet=yes}      => {higher=yes} 0.3464115 0.8982630 0.3856459 0.9829161   362
[11] {activities=no,
      nursery=yes}       => {higher=yes} 0.3617225 0.9064748 0.3990431 0.9919018   378
[12] {schoolsup=no,
      activities=no}     => {higher=yes} 0.3971292 0.8867521 0.4478469 0.9703204   415
[13] {fatherd=no,
      romantic=no}       => {higher=yes} 0.4641148 0.9203036 0.5043062 1.0070338   485
[14] {internet=yes,
      romantic=no}       => {higher=yes} 0.4736842 0.9464627 0.5004785 1.0356582   495
[15] {nursery=yes,
      romantic=no}       => {higher=yes} 0.4880383 0.9461967 0.5157895 1.0353670   510
[16] {schoolsup=no,
```

Plot the result.

```
> plot(rules_new)
>
```



Step 15

Explore Association rules using interactive manipulations and viewing using shiny.

Install and load the arulesviz() package.

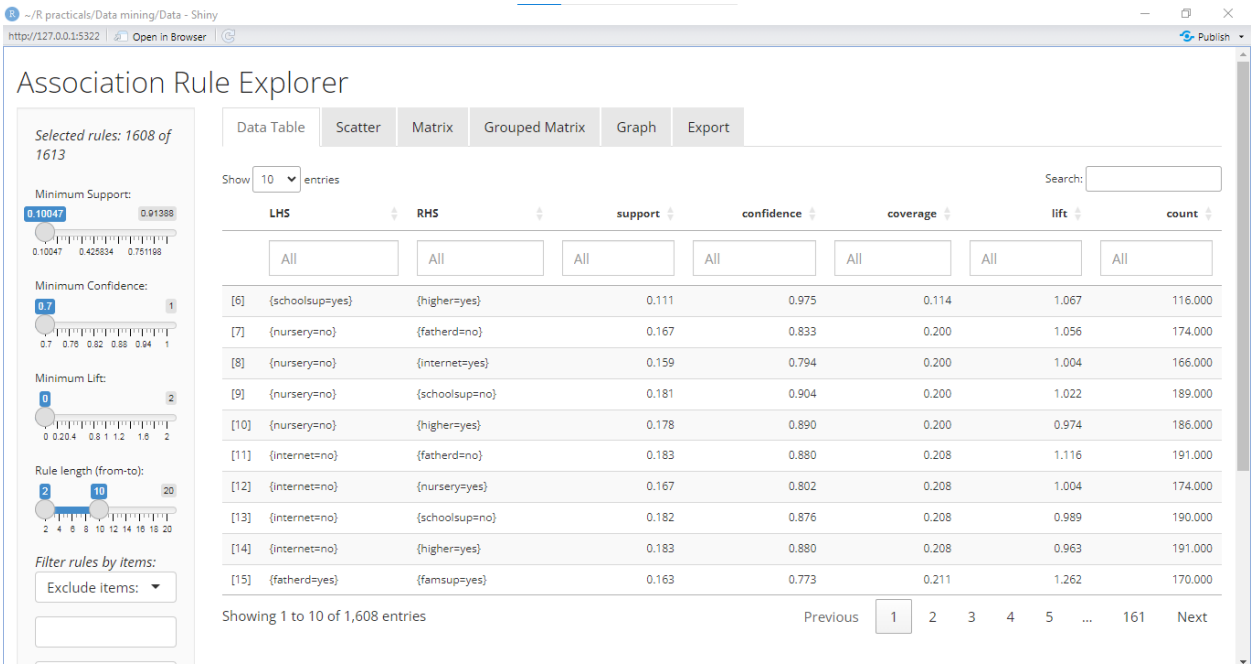
```
> library(arulesViz)
>
```

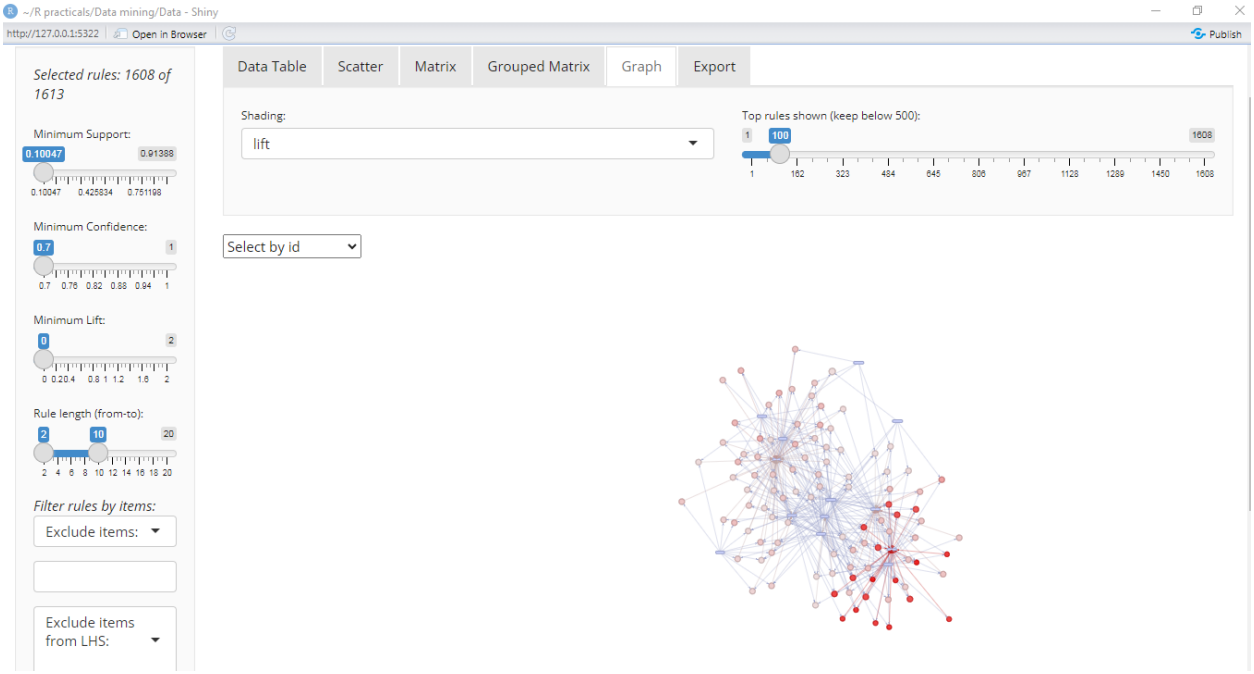
Get the rules under the confidence = 0.7

```
> rules_ex=apriori(data_new,parameter = list(conf=0.7))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen
        0.7    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: 104

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[24 item(s), 1045 transaction(s)] done [0.00s].
sorting and recoding items ... [15 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 7 done [0.00s].
writing ... [1613 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
>
```

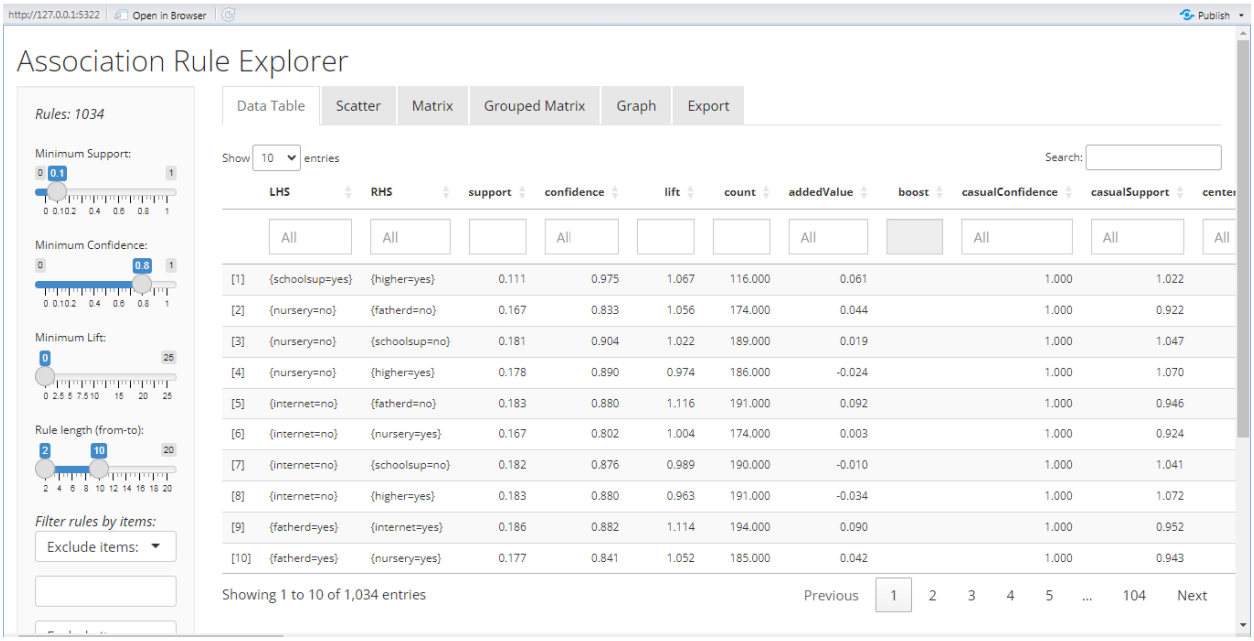
Explore association rules using ruleExplorer() function.

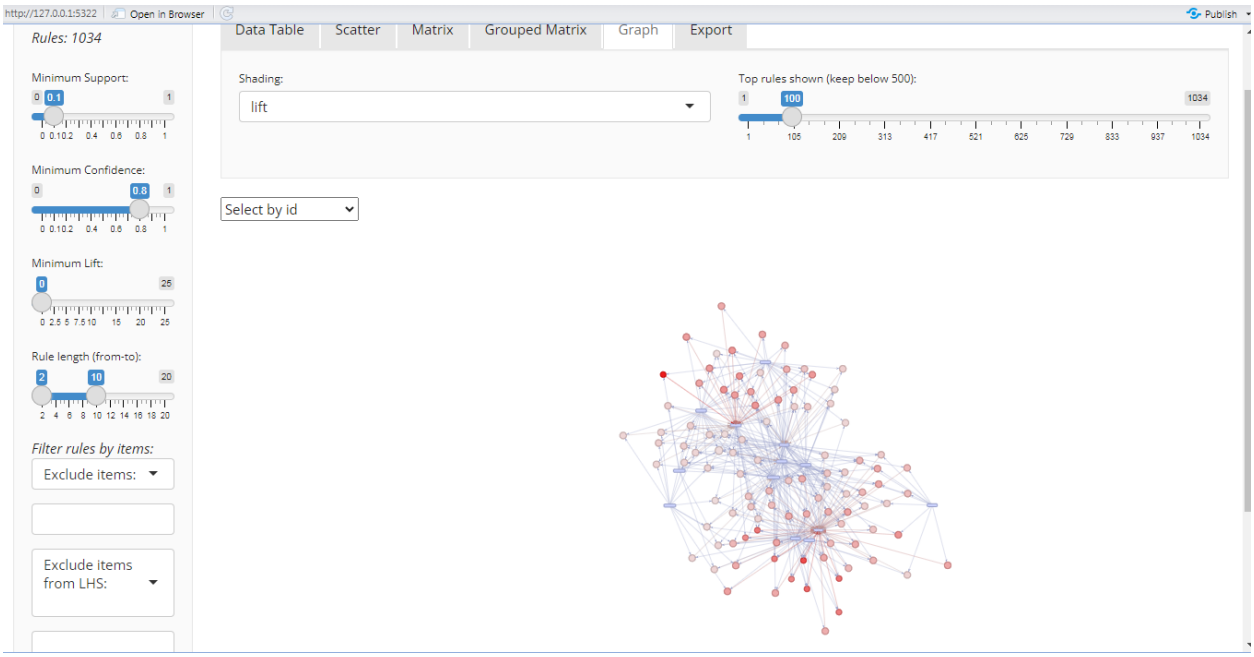
```
> ruleExplorer(rules_ex)
ruleExplorer started.
Loading required package: shiny
Warning: package 'shiny' was built under R version 4.2.3

Listening on http://127.0.0.1:5322
>
```


~/R practicals/Data mining/Data - Shiny
http://127.0.0.1:5322
Open in Browser
Publish
Association Rule Explorer
Selected rules: 1608 of 1613
Minimum Support:
0.10047
0.91388
Minimum Confidence:
0.7
1
Minimum Lift:
0
2
Rule length (from-to):
2
10
20
Filter rules by items:
Exclude items:
Data Table
Scatter
Matrix
Grouped Matrix
Graph
Export
Show 10 entries
Search:
LHS
RHS
support
confidence
coverage
lift
count
All
[6] {schoolsup=yes} {higher=yes} 0.111 0.975 0.114 1.067 116.000
[7] {nursery=no} {fatherd=no} 0.167 0.833 0.200 1.056 174.000
[8] {nursery=no} {internet=yes} 0.159 0.794 0.200 1.004 166.000
[9] {nursery=no} {schoolsup=no} 0.181 0.904 0.200 1.022 189.000
[10] {nursery=no} {higher=yes} 0.178 0.890 0.200 0.974 186.000
[11] {internet=no} {fatherd=no} 0.183 0.880 0.208 1.116 191.000
[12] {internet=no} {nursery=yes} 0.167 0.802 0.208 1.004 174.000
[13] {internet=no} {schoolsup=no} 0.182 0.876 0.208 0.989 190.000
[14] {internet=no} {higher=yes} 0.183 0.880 0.208 0.963 191.000
[15] {fatherd=yes} {famsup=yes} 0.163 0.773 0.211 1.262 170.000
Showing 1 to 10 of 1,608 entries
Previous 1 2 3 4 5 ... 161 Next

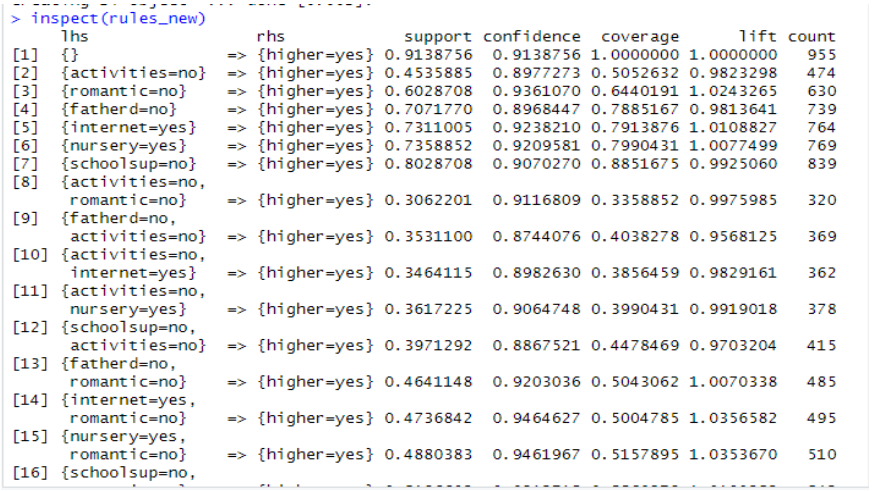

~/R practicals/Data mining/Data - Shiny
http://127.0.0.1:5322
Open in Browser
Publish
Selected rules: 1608 of 1613
Minimum Support:
0.10047
0.91388
Minimum Confidence:
0.7
1
Minimum Lift:
0
2
Rule length (from-to):
2
10
20
Filter rules by items:
Exclude items:
Exclude items from LHS:
Data Table
Scatter
Matrix
Grouped Matrix
Graph
Export
Shading:
lift
Top rules shown (keep below 500):
1
100
1608
Select by id

```
> ruleExplorer(data_new)
ruleExplorer started.
Converting dataset into transactions.

Listening on http://127.0.0.1:5322
```





6) Result, Analysis, Discussions



```
> inspect(rules_new)
     lhs                 rhs             support confidence  coverage      lift count
[1]  {}               => {higher=yes} 0.9138756  0.9138756 1.0000000 1.0000000   955
[2]  {activities=no}  => {higher=yes} 0.4535885  0.8977273 0.5052632 0.9823298   474
[3]  {romantic=no}    => {higher=yes} 0.6028708  0.9361070 0.6440191 1.0243265   630
[4]  {fatherd=no}     => {higher=yes} 0.7071770  0.8968447 0.7885167 0.9813641   739
[5]  {internet=yes}   => {higher=yes} 0.7311005  0.9238210 0.7913876 1.0108827   764
[6]  {nursery=yes}    => {higher=yes} 0.7358852  0.9209581 0.7990431 1.0077499   769
[7]  {schoolsup=no}   => {higher=yes} 0.8028708  0.9070270 0.8851675 0.9925060   839
[8]  {activities=no,
      romantic=no}    => {higher=yes} 0.3062201  0.9116809 0.3358852 0.9975985   320
[9]  {fatherd=no,
      activities=no}  => {higher=yes} 0.3531100  0.8744076 0.4038278 0.9568125   369
[10] {activities=no,
      internet=yes}   => {higher=yes} 0.3464115  0.8982630 0.3856459 0.9829161   362
[11] {activities=no,
      nursery=yes}    => {higher=yes} 0.3617225  0.9064748 0.3990431 0.9919018   378
[12] {schoolsup=no,
      activities=no}  => {higher=yes} 0.3971292  0.8867521 0.4478469 0.9703204   415
[13] {fatherd=no,
      romantic=no}    => {higher=yes} 0.4641148  0.9203036 0.5043062 1.0070338   485
[14] {internet=yes,
      romantic=no}    => {higher=yes} 0.4736842  0.9464627 0.5004785 1.0356582   495
[15] {nursery=yes,
      romantic=no}    => {higher=yes} 0.4880383  0.9461967 0.5157895 1.0353670   510
[16] {schoolsup=no,
```

The snippet of code output in the image appears to be related to data mining or machine learning.

Specifically, it shows the inspection of rules derived from some dataset.

Each rule consists of a left-hand side and a right-hand side, along with various metrics.

The goal is likely to discover interesting associations or patterns between conditions and outcomes.

Key Observations:

The command inspect (rules_new) has been executed.

The rules are associated with an outcome labeled as “higher=yes”.

Metrics provided for each rule include:

Support: Indicates the proportion of instances that satisfy the rule.

Confidence: Measures how often the outcome occurs when the conditions are met.

Coverage: Represents the proportion of instances covered by the rule.

Lift: Indicates the strength of association between conditions and the outcome.

Count: The number of instances that match the rule.

Example Rules (abbreviated):

Rule 1: {lhs} => {higher=yes} with high support and confidence.

Rule 2: {activities=no} => {higher=yes} with moderate support and confidence.

Rule 3: {romantic=no} => {higher=yes} with decent support and confidence.

Contextual Interpretation:

These rules could be part of a recommendation system, risk assessment, or decision-making process.

The conditions (lhs) may represent features or attributes, while the outcome (rhs) could be a desirable result.