DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
CU
CHANDIGARH
UNIVERSITY
Discover. Learn. Empower.

NAAC
GRADE A+
ACCREDITED UNIVERSITY

EXPERIMENT 1.3

Student Name: Jatin
Branch: CSE
UID: 20BCS5951
Section:NTPP- 605 B
Subject – Data Mining Lab
Semester – 6

Aim:

Demonstration of association rule mining using Apriory algorithm on supermarket data.

Objective:

To find interesting associations and relationships among large sets of data items with the help of Association rule mining

CODE-

```
install.packages("arules")
install.packages("arulesViz")
library(arules)
library(arulesViz)

install.packages("RColorBrewer")
library(RColorBrewer)

data("swiss")
print(swiss)
summary(swiss)
```

DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
Discover. Learn. Empower.

NAAC
GRADE A+
ACCREDITED UNIVERSITY

RStudio

File   Edit   Code   View   Plots   Session   Build   Debug   Profile   Tools   Help

Go to file/function   Addins ▾

Source

Console   Terminal ×   Jobs ×

~/ →

```
> library(arules)
> library(arulesviz)
> library(RColorBrewer)
> data("swiss")
> print(swiss)
               Fertility Agriculture Examination Education Catholic Infant.Mortality
Courtelary          80.2        17.0          15        12     9.96             22.2
Delemont            83.1        45.1           6         9    84.84             22.2
Franches-Mnt        92.5        39.7           5         5    93.40             20.2
Moutier             85.8        36.5          12         7    33.77             20.3
Neuveville          76.9        43.5          17        15     5.16             20.6
Porrentruy          76.1        35.3           9         7    90.57             26.6
Broye               83.8        70.2          16         7    92.85             23.6
Glane               92.4        67.8          14         8    97.16             24.9
Gruyere             82.4        53.3          12         7    97.67             21.0
Sarine              82.9        45.2          16        13    91.38             24.4
Veveyse             87.1        64.5          14         6    98.61             24.5
Aigle               64.1        62.0          21        12     8.52             16.5
Aubonne             66.9        67.5          14         7     2.27             19.1
Avenches            68.9        60.7          19        12     4.43             22.7
Cossonay            61.7        69.3          22         5     2.82             18.7
Echallens           68.3        72.6          18         2    24.20             21.2
Grandson            71.7        34.0          17         8     3.30             20.0
Lausanne            55.7        19.4          26        28    12.11             20.2
La Vallee           54.3        15.2          31        20     2.15             10.8
Lavaux              65.1        73.0          19         9     2.84             20.0
Morges              65.5        59.8          22        10     5.23             18.0
Moudon              65.0        55.1          14         3     4.52             22.4
Nyone               56.6        50.9          22        12    15.14             16.7
Orbe                57.4        54.1          20         6     4.20             15.3
Oron                72.5        71.2          12         1     2.40             21.0
Payerne             74.2        58.1          14         8     5.23             23.8
Paysd'enhaut        72.0        63.5           6         3     2.56             18.0
Rolle               60.5        60.8          16        10     7.72             16.3
Vevey               58.3        26.8          25        19    18.46             20.9
Yverdon             65.4        49.5          15         8     6.10             22.5
Conthey             75.5        85.9           3         2    99.71             15.1
Entremont           69.3        84.9           7         6    99.68             19.8
Herens              77.3        89.7           5         2   100.00             18.3
Martigwy            70.5        78.2          12         6    98.96             19.4
Monthey             79.4        64.9           7         3    98.22             20.2
St Maurice          65.0        75.9           9         9    99.06             17.8
Sierre              92.2        84.6           3         3    99.46             16.3
Sion                79.3        63.1          13        13    96.83             18.1
Boudry              70.4        38.4          26        12     5.62             20.3
La Chauxdfnd        65.7         7.7          29        11    13.79             20.5
Le Locle            72.7        16.7          22        13    11.22             18.9
Neuchatel           64.4        17.6          35        32    16.92             23.0
Val de Ruz          77.6        37.6          15         7     4.97             20.0
ValdeTravers        67.6        18.7          25         7     8.65             19.5
V. De Geneve        35.0         1.2          37        53    42.34             18.0
Rive Droite         44.7        46.6          16        29    50.43             18.2
Rive Gauche         42.8        27.7          22        29    58.33             19.3
> summary(swiss)
```

```
> summary(swiss)
   Fertility      Agriculture     Examination      Education        Catholic       Infant.Mortality
 Min.   :35.00   Min.   : 1.20   Min.   : 3.00   Min.   : 1.00   Min.   :  2.150   Min.   :10.80
 1st Qu.:64.70   1st Qu.:35.90   1st Qu.:12.00   1st Qu.: 6.00   1st Qu.:  5.195   1st Qu.:18.15
 Median :70.40   Median :54.10   Median :16.00   Median : 8.00   Median : 15.140   Median :20.00
 Mean   :70.14   Mean   :50.66   Mean   :16.49   Mean   :10.98   Mean   : 41.144   Mean   :19.94
 3rd Qu.:78.45   3rd Qu.:67.65   3rd Qu.:22.00   3rd Qu.:12.00   3rd Qu.: 93.125   3rd Qu.:21.70
 Max.   :92.50   Max.   :89.70   Max.   :37.00   Max.   :53.00   Max.   :100.000   Max.   :26.60
```

#Using Apriori function
#**rule1**
rule1 <- apriori(swiss, parameter = list(supp = 0.01, conf = 0.2))

inspect(rule1[1:10])
inspect(head(rule1,7))
inspect(head(sort(rule1,by = "lift"),5))

plot(rule1)
plot(rule1,method="grouped")

```
> #Using Apriori function
> rule1 <- apriori(swiss, parameter = list(supp = 0.01, conf = 0.2))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target   ext
        0.2    0.1    1 none FALSE            TRUE       5    0.01      1     10  rules TRUE

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

Absolute minimum support count: 0

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[18 item(s), 47 transaction(s)] done [0.00s].
sorting and recoding items ... [18 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 done [0.00s].
writing ... [4316 rule(s)] done [0.00s].
creating S4 object  ... done [0.00s].
```

```
> inspect(rule1[1:10])
     lhs    rhs                       support   confidence coverage lift count
[1]  {}  => {Education=[1,7)}         0.2978723 0.2978723  1        1    14
[2]  {}  => {Examination=[14,19)}     0.3191489 0.3191489  1        1    15
[3]  {}  => {Catholic=[6.64,76)}      0.3191489 0.3191489  1        1    15
[4]  {}  => {Agriculture=[41,63.4)}   0.3191489 0.3191489  1        1    15
[5]  {}  => {Examination=[3,14)}      0.3191489 0.3191489  1        1    15
[6]  {}  => {Infant.Mortality=[19,20.8)} 0.3191489 0.3191489 1     1    15
[7]  {}  => {Fertility=[65.4,75.9)}   0.3191489 0.3191489  1        1    15
[8]  {}  => {Infant.Mortality=[10.8,19)} 0.3404255 0.3404255 1     1    16
[9]  {}  => {Fertility=[35,65.4)}     0.3404255 0.3404255  1        1    16
[10] {}  => {Agriculture=[63.4,89.7)} 0.3404255 0.3404255  1        1    16
> inspect(head(rule1,7))
     lhs    rhs                       support   confidence coverage lift count
[1]  {}  => {Education=[1,7)}         0.2978723 0.2978723  1        1    14
[2]  {}  => {Examination=[14,19)}     0.3191489 0.3191489  1        1    15
[3]  {}  => {Catholic=[6.64,76)}      0.3191489 0.3191489  1        1    15
[4]  {}  => {Agriculture=[41,63.4)}   0.3191489 0.3191489  1        1    15
[5]  {}  => {Examination=[3,14)}      0.3191489 0.3191489  1        1    15
[6]  {}  => {Infant.Mortality=[19,20.8)} 0.3191489 0.3191489 1     1    15
[7]  {}  => {Fertility=[65.4,75.9)}   0.3191489 0.3191489  1        1    15
> inspect(head(sort(rule1,by = "lift"),5))
     lhs                                              rhs                 support    confidence
[1]  {Agriculture=[63.4,89.7), Catholic=[6.64,76)}  => {Education=[1,7)} 0.02127660 1
[2]  {Fertility=[65.4,75.9), Examination=[3,14)}    => {Education=[1,7)} 0.10638298 1
[3]  {Examination=[3,14), Catholic=[2.15,6.64)}     => {Education=[1,7)} 0.04255319 1
[4]  {Catholic=[76,100], Infant.Mortality=[19,20.8)} => {Education=[1,7)} 0.08510638 1
[5]  {Fertility=[65.4,75.9), Catholic=[76,100]}     => {Education=[1,7)} 0.06382979 1
     coverage   lift     count
[1]  0.02127660 3.357143 1
[2]  0.10638298 3.357143 5
[3]  0.04255319 3.357143 2
[4]  0.08510638 3.357143 4
[5]  0.06382979 3.357143 3
> plot(rule1)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
>
> plot(rule1)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
> plot(rule1,method="grouped")
```
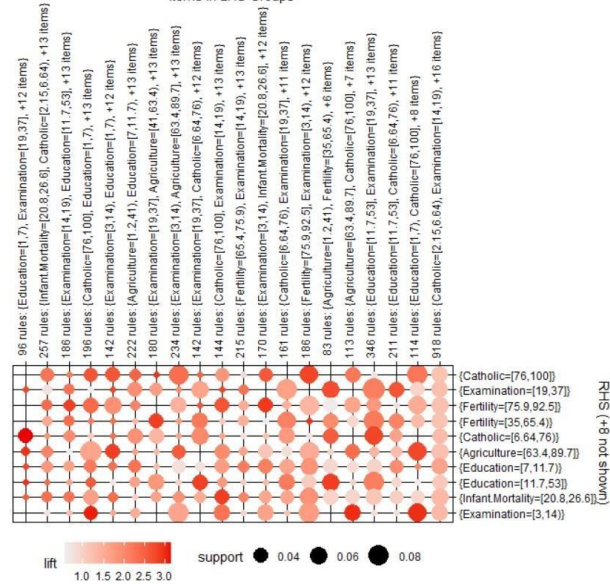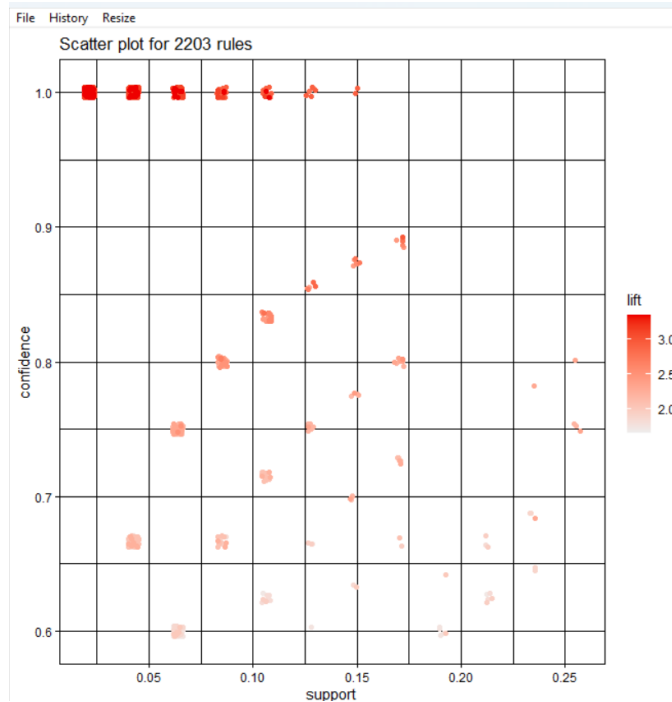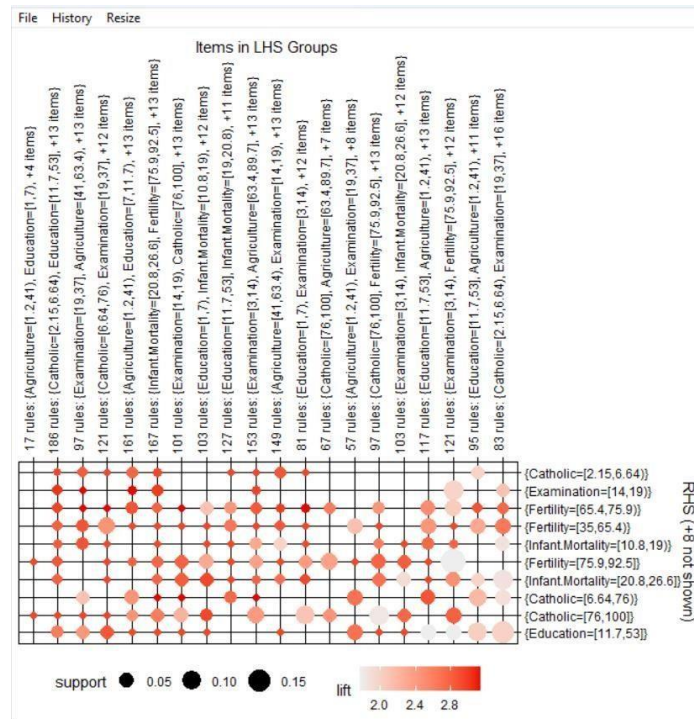
DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
Discover. Learn. Empower.

NAAC
GRADE A+
ACCREDITED UNIVERSITY

Scatter plot for 4316 rules
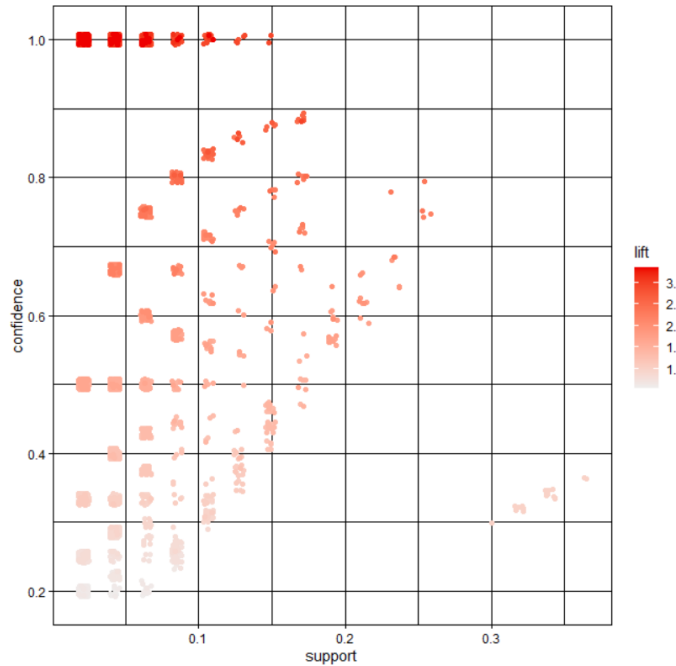
Items in LHS Groups

#rule2

```
rule2 <- apriori(swiss ,parameter = list(supp = 0.007, conf = 0.6))
inspect(head(rule2,5))
plot(rule2)
```

plot(rule2, method = "grouped")

```
> plot(rule2)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
> plot(rule2, method = "grouped")
> rule3 <- apriori(swiss,parameter = list(supp = 0.003, conf = 0.8))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target  ext
        0.8    0.1    1 none FALSE              TRUE       5   0.003      1     10  rules TRUE

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

Absolute minimum support count: 0

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[18 item(s), 47 transaction(s)] done [0.00s].
sorting and recoding items ... [18 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 done [0.00s].
writing ... [1717 rule(s)] done [0.01s].
creating S4 object  ... done [0.00s].
```
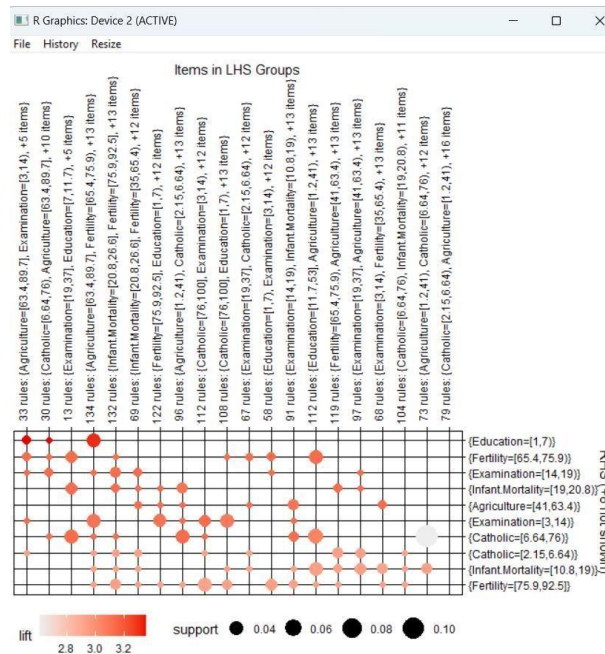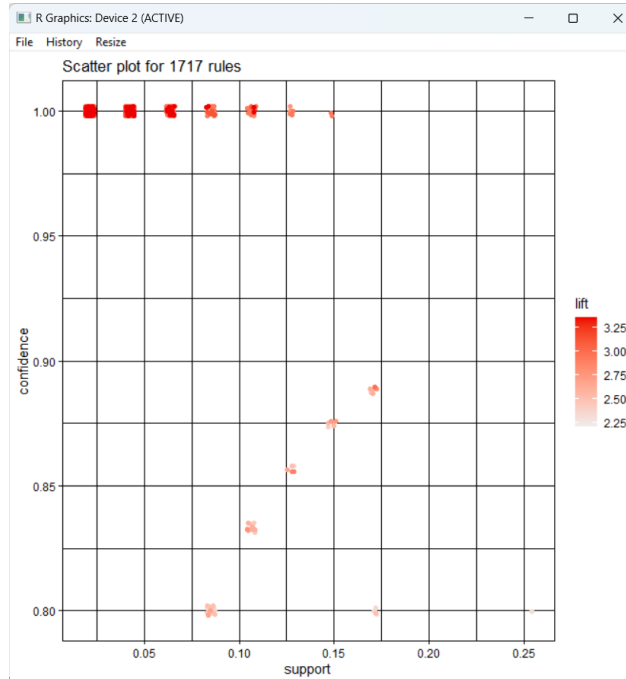
```
> inspect(head(rule3,5))
    lhs                                      rhs                              support    confidence coverage   lift     count
[1] {Examination=[3,14)}                  => {Catholic=[76,100]}             0.25531915 0.8        0.31914894 2.350000 12
[2] {Education=[1,7), Catholic=[6.64,76)} => {Examination=[14,19)}           0.02127660 1.0        0.02127660 3.133333  1
[3] {Examination=[14,19), Education=[1,7)} => {Infant.Mortality=[20.8,26.6]} 0.06382979 1.0        0.06382979 2.937500  3
[4] {Education=[1,7), Catholic=[6.64,76)} => {Fertility=[65.4,75.9)}         0.02127660 1.0        0.02127660 3.133333  1
[5] {Education=[1,7), Catholic=[6.64,76)} => {Agriculture=[63.4,89.7]}       0.02127660 1.0        0.02127660 2.937500  1
> plot(rule3)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
> plot(rule3, method = "grouped")
```

DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
Discover. Learn. Empower.

NAAC
GRADE A+
ACCREDITED UNIVERSITY

#rule3

rule3 <- apriori(swiss,parameter = list(supp = 0.003, conf = 0.8))
inspect(head(rule3,5))
plot(rule3)
plot(rule3, method = "grouped")

```
> rule3 <- apriori(swiss,parameter = list(supp = 0.003, conf = 0.8))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target  ext
        0.8    0.1    1 none FALSE            TRUE       5   0.003      1     10  rules TRUE

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

Absolute minimum support count: 0

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[18 item(s), 47 transaction(s)] done [0.00s].
sorting and recoding items ... [18 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 done [0.00s].
writing ... [1717 rule(s)] done [0.01s].
creating S4 object  ... done [0.00s].
```
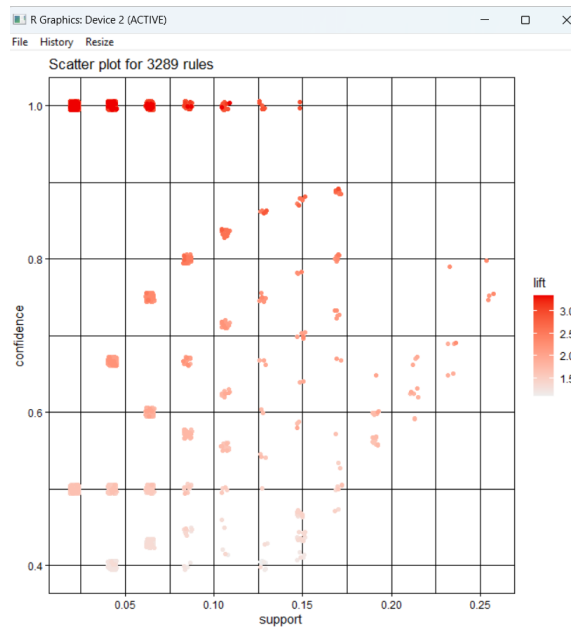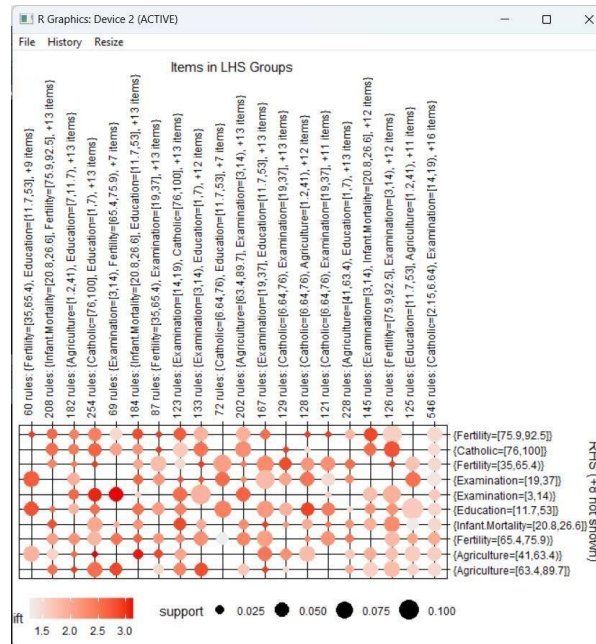
```
> inspect(head(rule3,5))
    lhs                                      rhs                          support    confidence coverage   lift     count
[1] {Examination=[3,14]}                  => {Catholic=[76,100]}          0.25531915 0.8        0.31914894 2.350000 12
[2] {Education=[1,7], Catholic=[6.64,76]} => {Examination=[14,19]}        0.02127660 1.0        0.02127660 3.133333 1
[3] {Examination=[14,19], Education=[1,7]} => {Infant.Mortality=[20.8,26.6]} 0.06382979 1.0     0.06382979 2.937500 3
[4] {Education=[1,7], Catholic=[6.64,76]} => {Fertility=[65.4,75.9]}       0.02127660 1.0        0.02127660 3.133333 1
[5] {Education=[1,7], Catholic=[6.64,76]} => {Agriculture=[63.4,89.7]}     0.02127660 1.0        0.02127660 2.937500 1
> plot(rule3)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
> plot(rule3, method = "grouped")
```

Scatter plot for 1717 rules

Items in LHS Groups

#rule4

rule4 <- apriori(swiss ,parameter = list(supp = 0.02, conf = 0.4))
inspect(head(rule4,5))

plot(rule4)
plot(rule4, method = "grouped")

```
> rule4 <- apriori( swiss ,parameter = list(supp = 0.02, conf = 0.4))
Apriori

Parameter specification:
 confidence minval smax arem  aval originalSupport maxtime support minlen maxlen target  ext
        0.4    0.1    1 none FALSE            TRUE       5    0.02      1     10  rules TRUE

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

Absolute minimum support count: 0

set item appearances ...[0 item(s)] done [0.00s].
set transactions ...[18 item(s), 47 transaction(s)] done [0.00s].
sorting and recoding items ... [18 item(s)] done [0.00s].
creating transaction tree ... done [0.00s].
checking subsets of size 1 2 3 4 5 6 done [0.00s].
writing ... [3289 rule(s)] done [0.01s].
creating S4 object  ... done [0.00s].
```

```
> inspect(head(rule4,5))
    lhs                       rhs                           support   confidence coverage  lift     count
[1] {Education=[1,7]}      => {Examination=[3,14]}          0.1914894 0.6428571  0.2978723 2.014286 9
[2] {Examination=[3,14]}   => {Education=[1,7]}             0.1914894 0.6000000  0.3191489 2.014286 9
[3] {Education=[1,7]}      => {Fertility=[65.4,75.9)}       0.1276596 0.4285714  0.2978723 1.342857 6
[4] {Fertility=[65.4,75.9)} => {Education=[1,7]}            0.1276596 0.4000000  0.3191489 1.342857 6
[5] {Education=[1,7]}      => {Infant.Mortality=[10.8,19)}  0.1276596 0.4285714  0.2978723 1.258929 6
> plot(rule4)
To reduce overplotting, jitter is added! Use jitter = 0 to prevent jitter.
> plot(rule4, method = "grouped")
```

DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
CU CHANDIGARH UNIVERSITY Discover. Learn. Empower.

NAAC GRADE A+
ACCREDITED UNIVERSITY

Items in LHS Groups

#rule5

```
rule5 <- apriori( swiss ,parameter = list(supp = 0.02, conf = 0.4))
inspect(head(rule5,5))
plot(rule5)
plot(rule5, method = "grouped")
```

DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
Discover. Learn. Empower.

NAAC GRADE A+
ACCREDITED UNIVERSITY