**Assignment 1 – Books dataset**

Using Apriori algorithm: Below are the 3 steps to be accomplished for the assignment and necessary solution is being provided in the subsequent sections

**Problem Statement:**

Prepare rules for the all the data sets

1) Try different values of support and confidence. Observe the change in number of rules for different support, confidence values

2) Change the minimum length in apriori algorithm

3) Visulize the obtained rules using different plots

**Solution:**

Step 1: Try different values of support and confidence. Observe the change in number of rules for different support, confidence values

Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.05,confidence = 0.4,minlen=5))

> Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.05,confidence = 0.4,minlen=5))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen maxlen

0.4 0.1 1 none FALSE TRUE 5 0.05 5 10

target ext

rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 100

set item appearances ...[0 item(s)] done [0.00s].

set transactions ...[11 item(s), 2000 transaction(s)] done [0.00s].

sorting and recoding items ... [9 item(s)] done [0.00s].

creating transaction tree ... done [0.00s].

checking subsets of size 1 2 3 4 5 done [0.00s].

writing ... [10 rule(s)] done [0.00s].

creating S4 object ... done [0.00s].

> inspect(Book\_rules)

lhs rhs support confidence lift

[1] {ChildBks,YouthBks,DoItYBks,GeogBks} => {CookBks} 0.0510 0.9026549 2.094327

[2] {YouthBks,CookBks,DoItYBks,GeogBks} => {ChildBks} 0.0510 0.8500000 2.009456

[3] {ChildBks,YouthBks,CookBks,DoItYBks} => {GeogBks} 0.0510 0.6219512 2.253446

[4] {ChildBks,YouthBks,CookBks,GeogBks} => {DoItYBks} 0.0510 0.6144578 2.178928

[5] {ChildBks,CookBks,DoItYBks,GeogBks} => {YouthBks} 0.0510 0.5730337 2.315288

[6] {ChildBks,DoItYBks,ArtBks,GeogBks} => {CookBks} 0.0535 0.8991597 2.086217

[7] {CookBks,DoItYBks,ArtBks,GeogBks} => {ChildBks} 0.0535 0.8230769 1.945808

[8] {ChildBks,CookBks,DoItYBks,ArtBks} => {GeogBks} 0.0535 0.6524390 2.363910

[9] {ChildBks,CookBks,ArtBks,GeogBks} => {DoItYBks} 0.0535 0.6407186 2.272052

[10] {ChildBks,CookBks,DoItYBks,GeogBks} => {ArtBks} 0.0535 0.6011236 2.494289

count

[1] 102

[2] 102

[3] 102

[4] 102

[5] 102

[6] 107

[7] 107

[8] 107

[9] 107

[10] 107

> #List out those books which are with higher lift ratio.

> inspect(head(sort(Book\_rules, by = "lift")))

lhs rhs support confidence lift

[1] {ChildBks,CookBks,DoItYBks,GeogBks} => {ArtBks} 0.0535 0.6011236 2.494289

[2] {ChildBks,CookBks,DoItYBks,ArtBks} => {GeogBks} 0.0535 0.6524390 2.363910

[3] {ChildBks,CookBks,DoItYBks,GeogBks} => {YouthBks} 0.0510 0.5730337 2.315288

[4] {ChildBks,CookBks,ArtBks,GeogBks} => {DoItYBks} 0.0535 0.6407186 2.272052

[5] {ChildBks,YouthBks,CookBks,DoItYBks} => {GeogBks} 0.0510 0.6219512 2.253446

[6] {ChildBks,YouthBks,CookBks,GeogBks} => {DoItYBks} 0.0510 0.6144578 2.178928

count

[1] 107

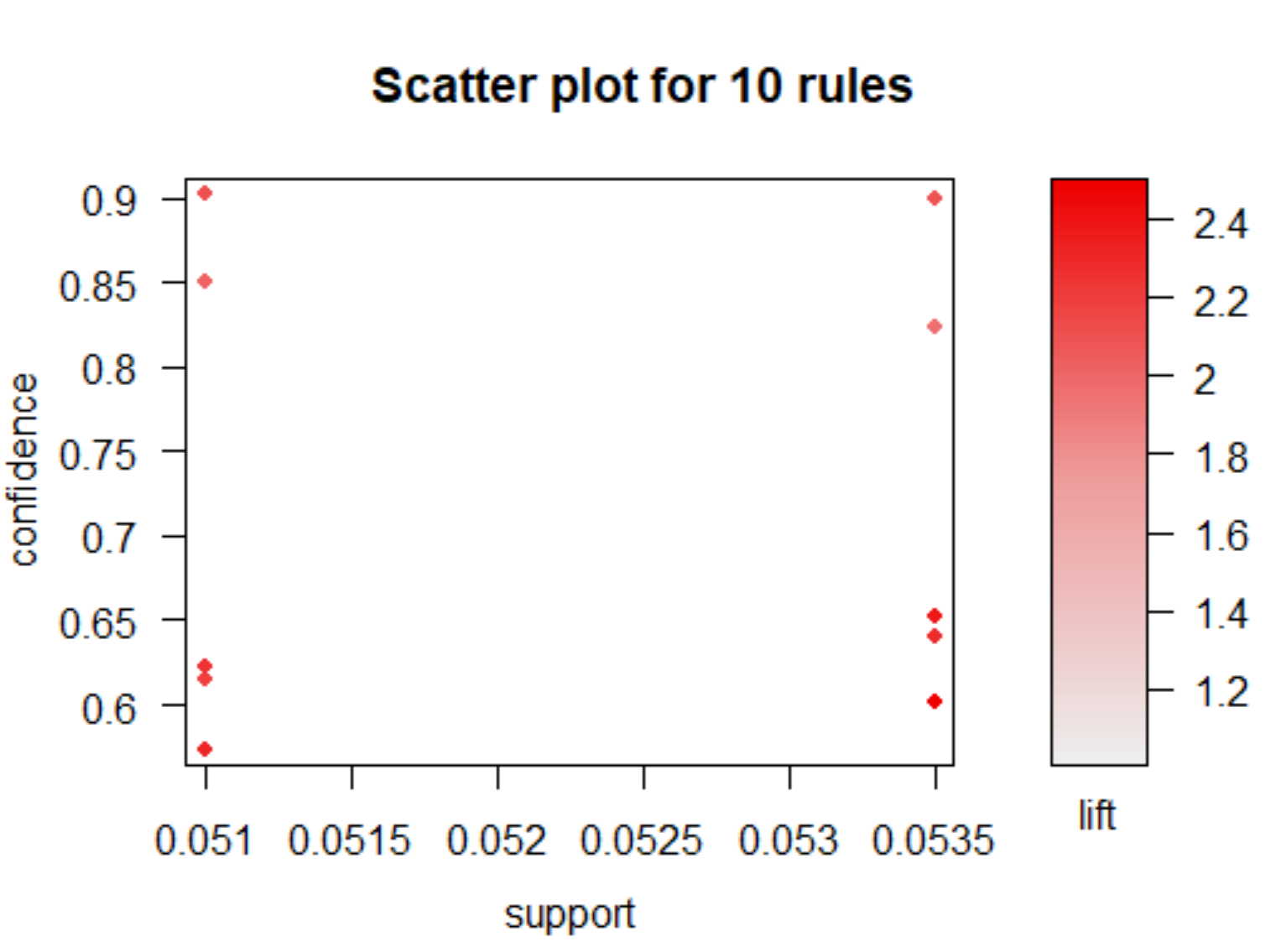
[2] 107

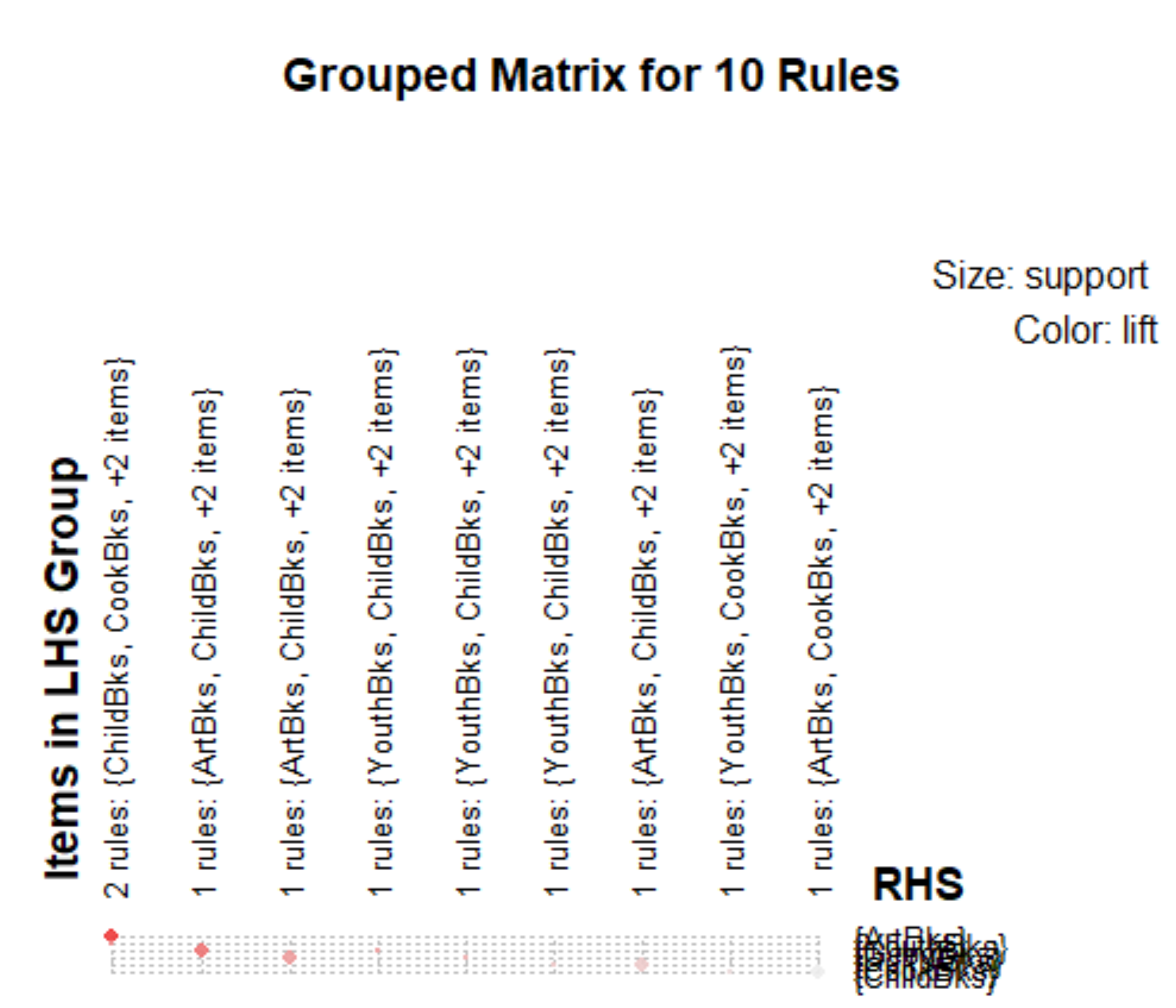
[3] 102

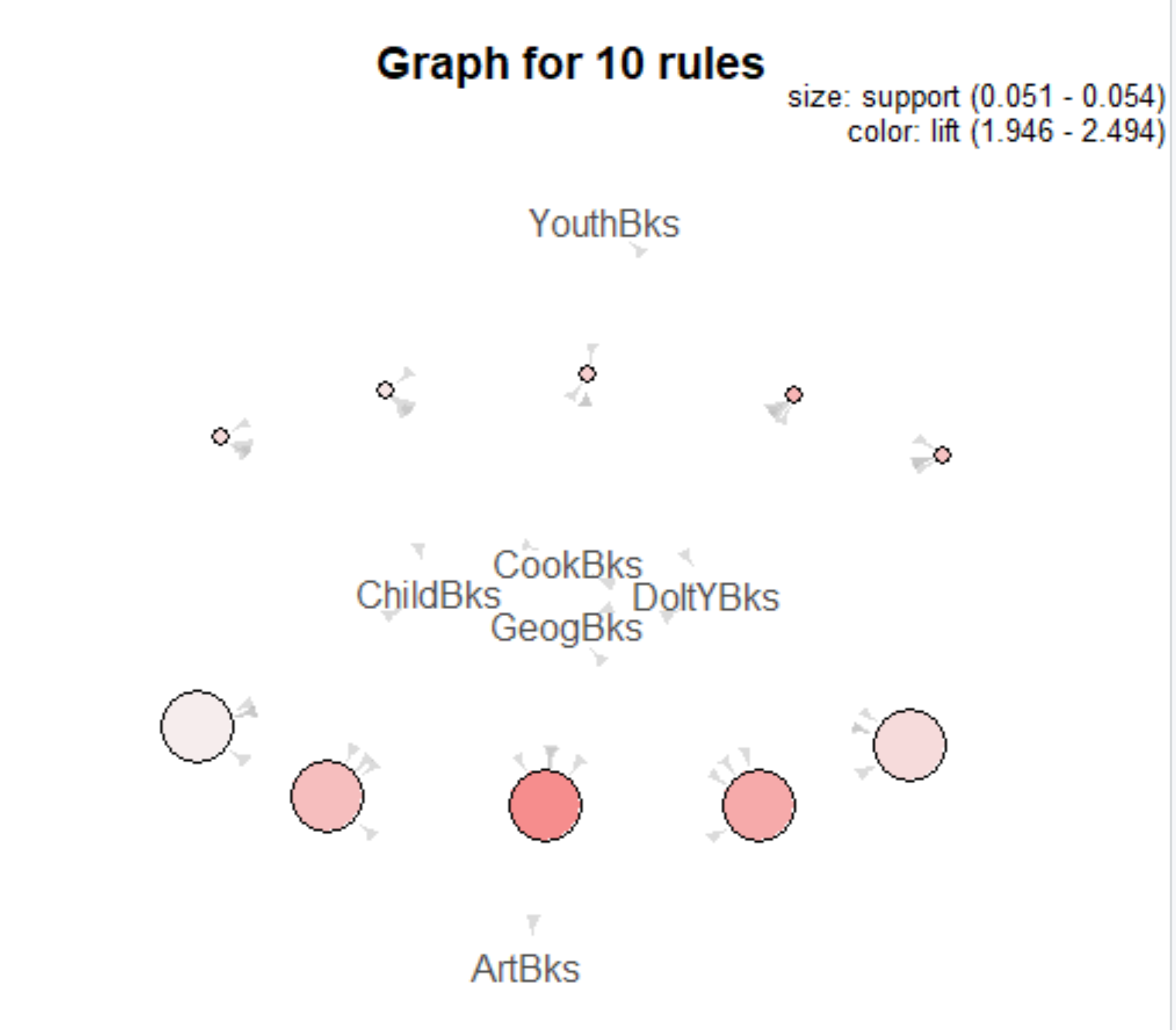
[4] 107

[5] 102

[6] 102







**2nd rule:** Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.02,confidence = 0.5,minlen=3))

inspect(head(sort(Book\_rules, by = "lift")))

lhs rhs support confidence lift count

[1] {DoItYBks,ArtBks,ItalCook} => {ItalArt} 0.0250 0.6849315 14.12230 50

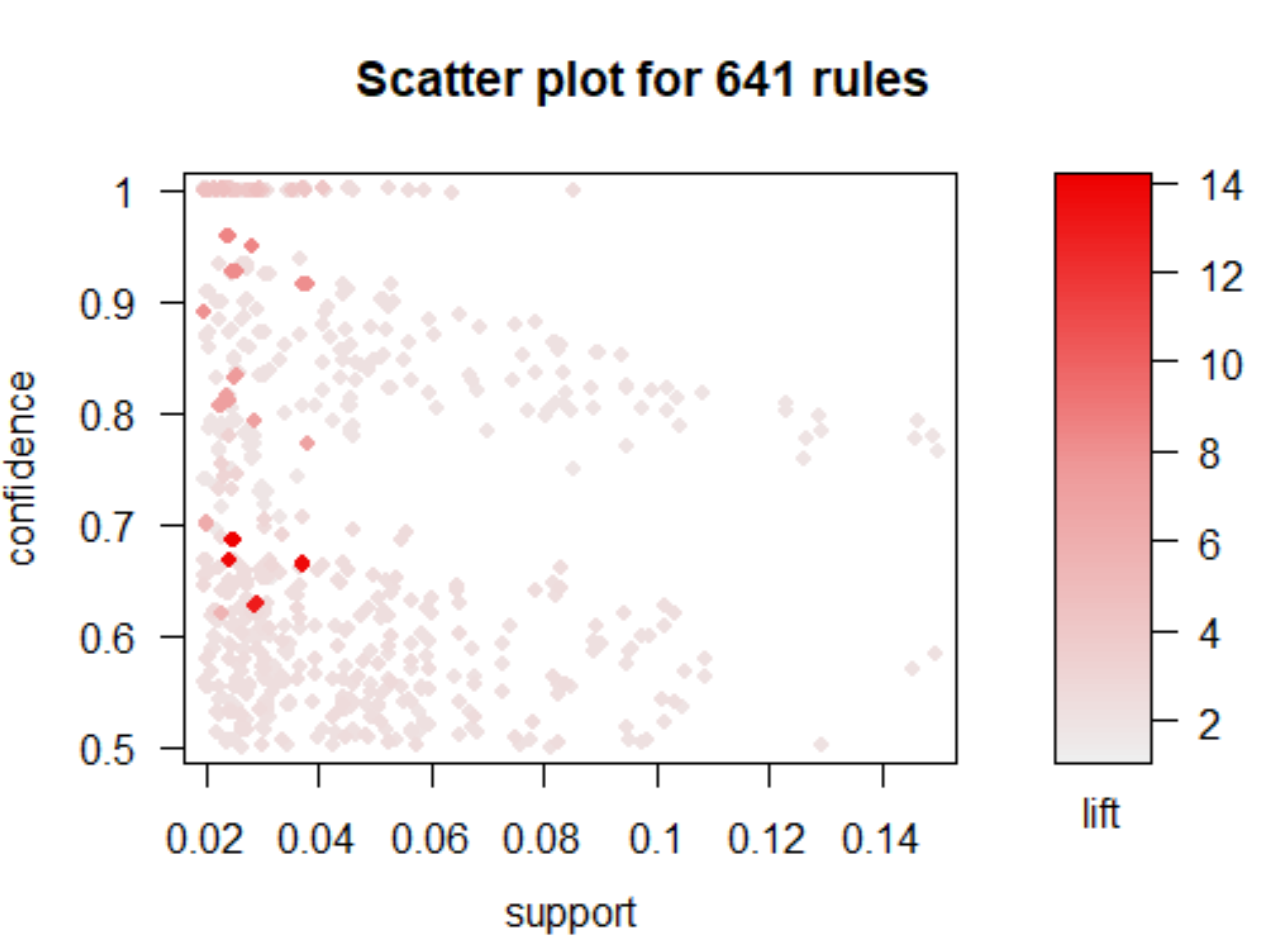
[2] {CookBks,DoItYBks,ArtBks,ItalCook} => {ItalArt} 0.0250 0.6849315 14.12230 50

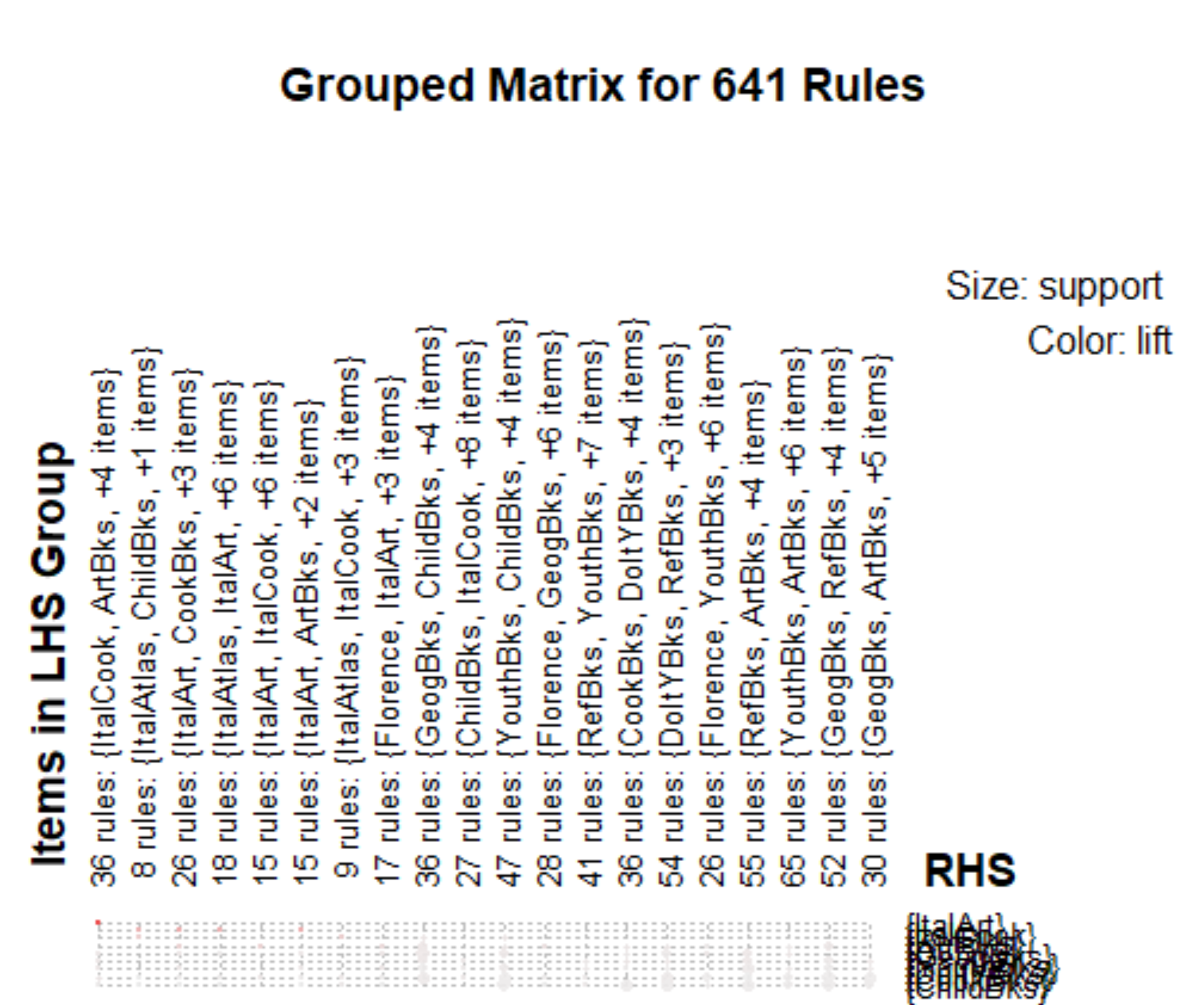
[3] {ArtBks,GeogBks,ItalCook} => {ItalArt} 0.0240 0.6666667 13.74570 48

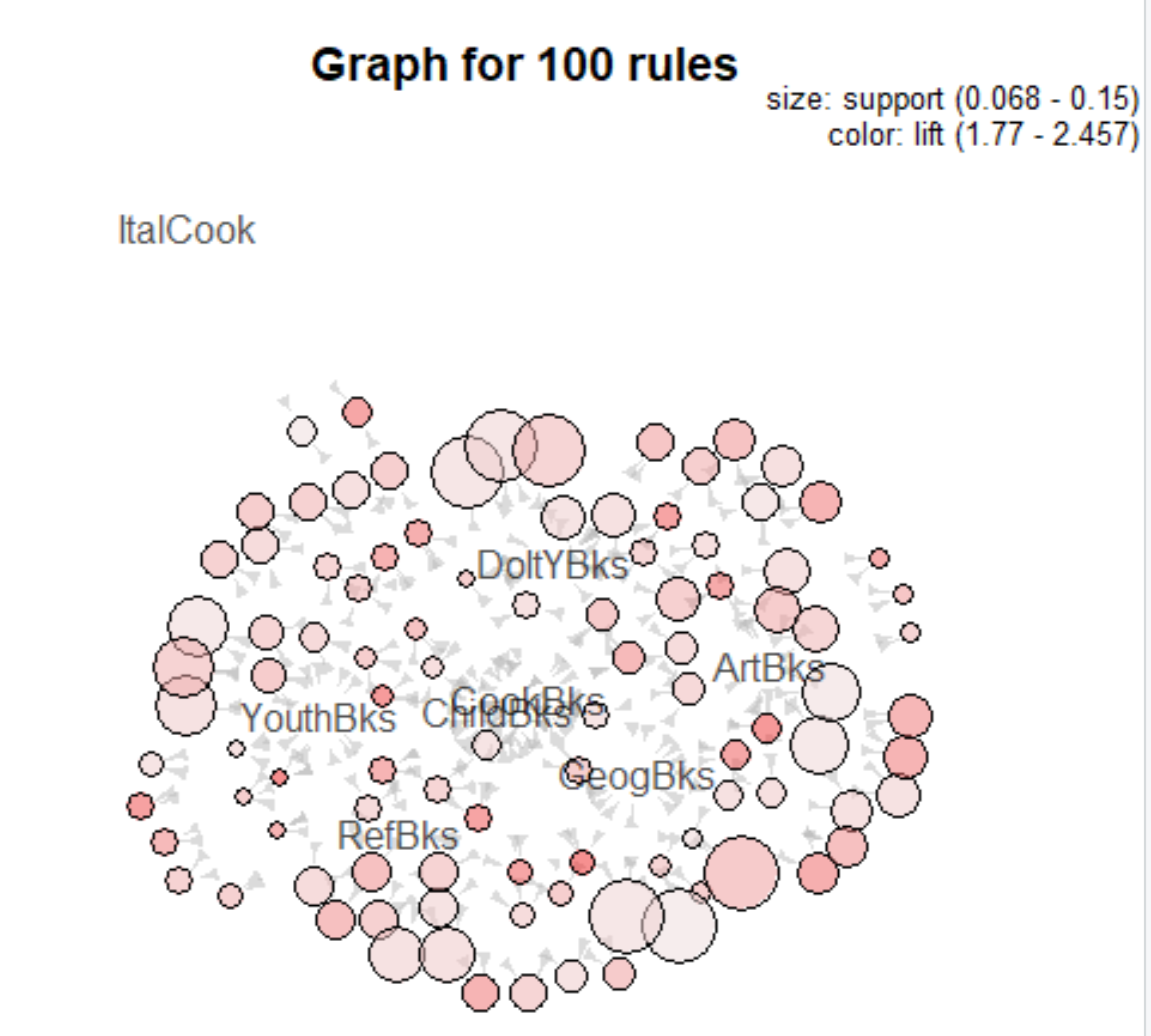
[4] {CookBks,ArtBks,GeogBks,ItalCook} => {ItalArt} 0.0240 0.6666667 13.74570 48

[5] {ArtBks,ItalCook} => {ItalArt} 0.0375 0.6637168 13.68488 75

[6] {CookBks,ArtBks,ItalCook} => {ItalArt} 0.0375 0.6637168 13.68488 75







3rd Rule: Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.03,confidence = 0.5,minlen=4))

> Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.03,confidence = 0.5,minlen=4))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen maxlen

0.5 0.1 1 none FALSE TRUE 5 0.03 4 10

target ext

rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 60

set item appearances ...[0 item(s)] done [0.00s].

set transactions ...[11 item(s), 2000 transaction(s)] done [0.00s].

sorting and recoding items ... [11 item(s)] done [0.00s].

creating transaction tree ... done [0.00s].

checking subsets of size 1 2 3 4 5 6 done [0.01s].

writing ... [249 rule(s)] done [0.00s].

creating S4 object ... done [0.00s].

> inspect(head(sort(Book\_rules, by = "lift")))

lhs rhs support confidence lift count

[1] {CookBks,ArtBks,ItalCook} => {ItalArt} 0.0375 0.6637168 13.684883 75

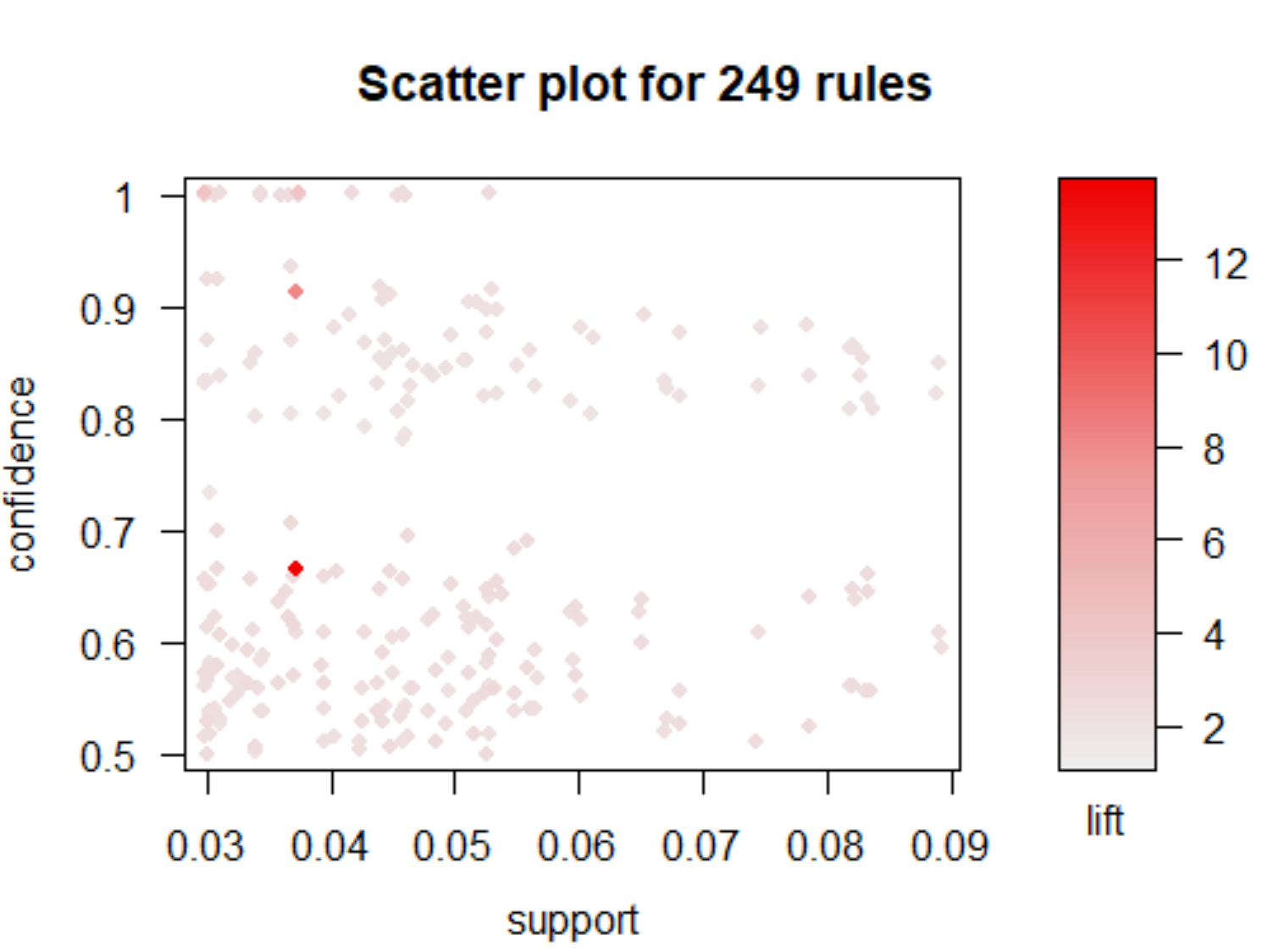
[2] {CookBks,ArtBks,ItalArt} => {ItalCook} 0.0375 0.9146341 8.058451 75

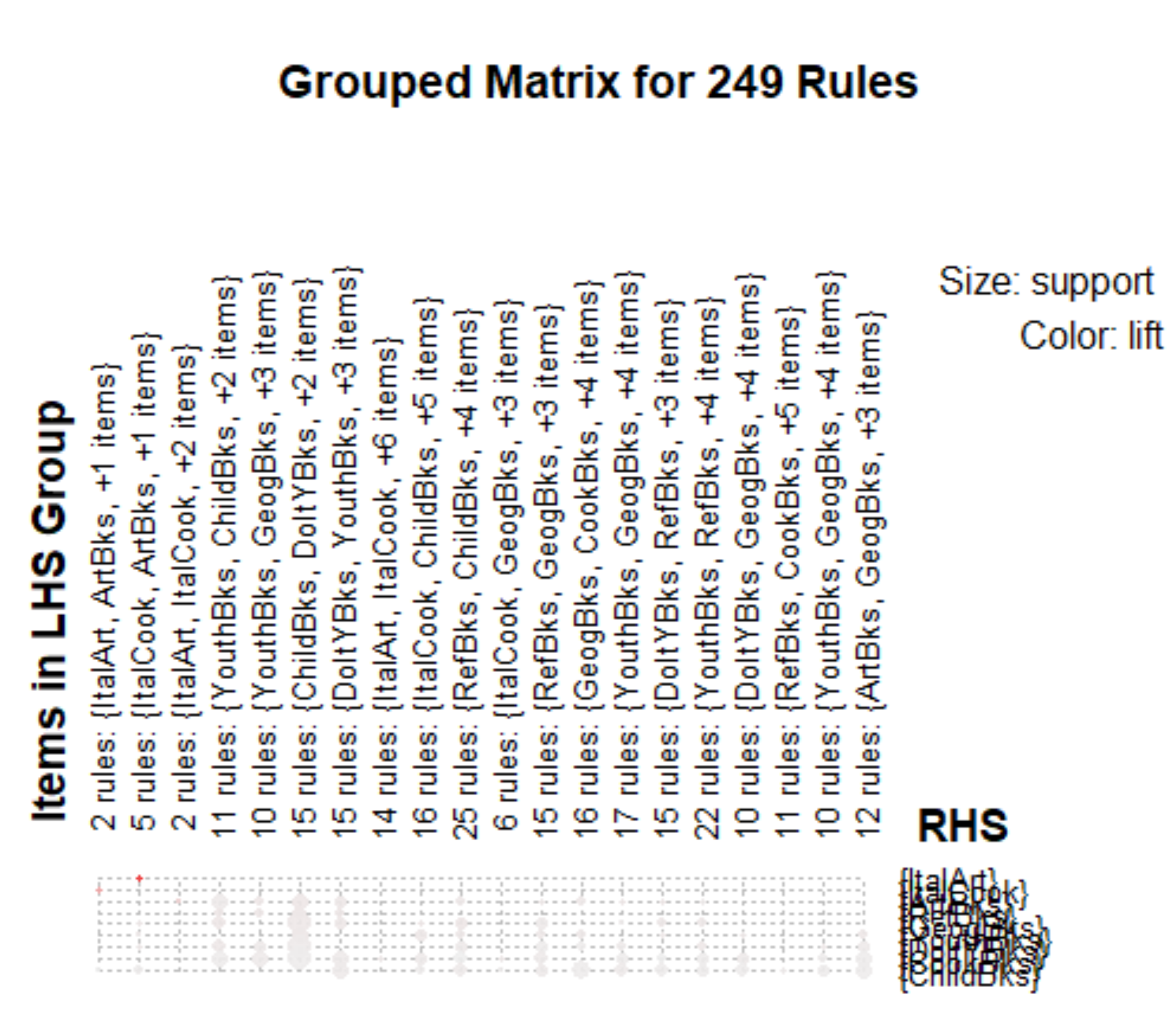
[3] {CookBks,ItalCook,ItalArt} => {ArtBks} 0.0375 1.0000000 4.149378 75

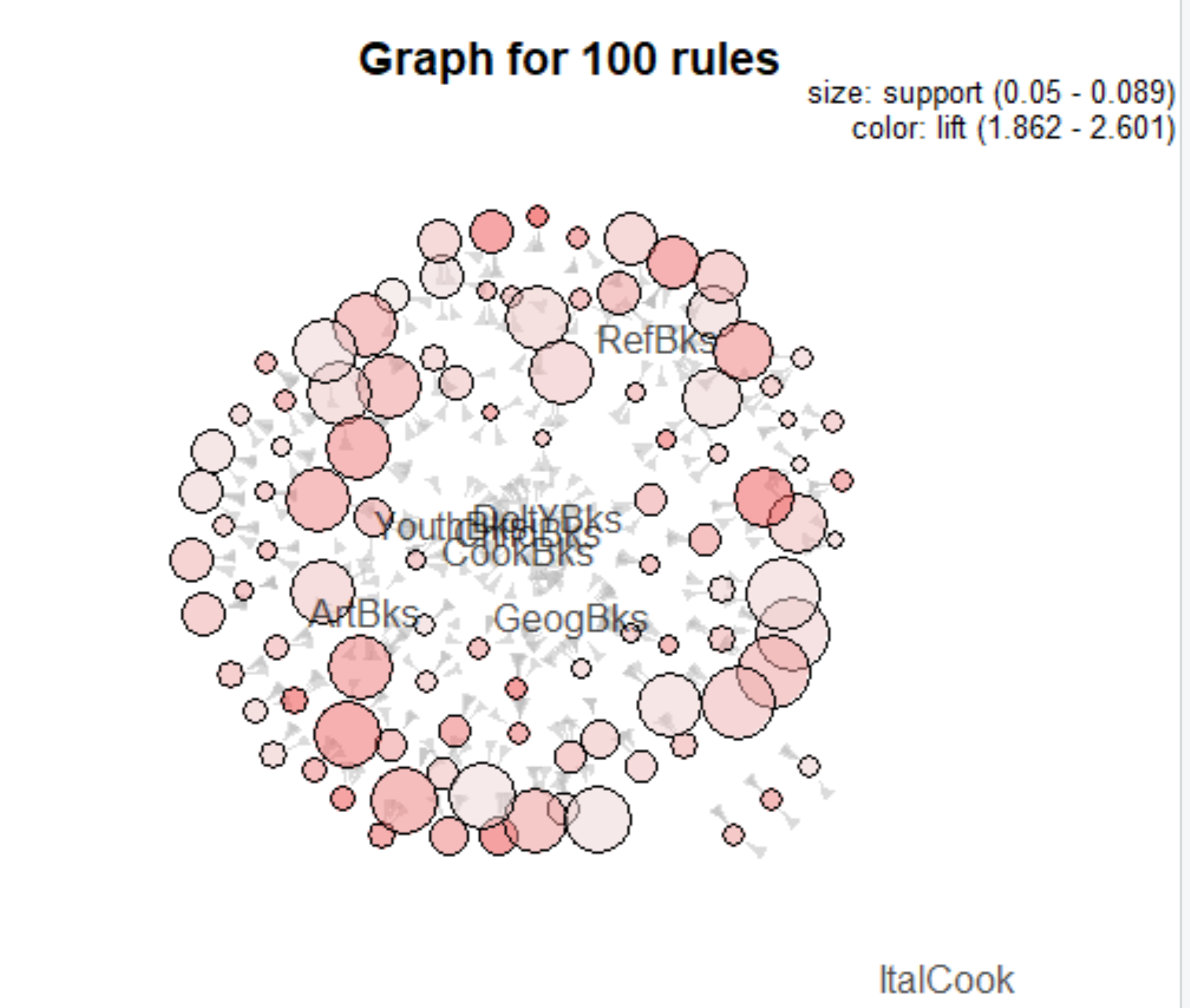
[4] {ChildBks,CookBks,ItalArt} => {ArtBks} 0.0300 1.0000000 4.149378 60

[5] {ChildBks,YouthBks,DoItYBks} => {RefBks} 0.0530 0.5578947 2.600908 106

[6] {CookBks,DoItYBks,ItalCook} => {ArtBks} 0.0365 0.6239316 2.588928 73







4th Rule:Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.02,confidence = 0.5,minlen=6))

|  |
| --- |
| Book\_rules <- apriori(as.matrix(Book\_data),parameter = list(support = 0.02,confidence = 0.5,minlen=6))  Apriori  Parameter specification:  confidence minval smax arem aval originalSupport maxtime support minlen maxlen  0.5 0.1 1 none FALSE TRUE 5 0.02 6 10  target ext  rules FALSE  Algorithmic control:  filter tree heap memopt load sort verbose  0.1 TRUE TRUE FALSE TRUE 2 TRUE  Absolute minimum support count: 40  set item appearances ...[0 item(s)] done [0.00s].  set transactions ...[11 item(s), 2000 transaction(s)] done [0.00s].  sorting and recoding items ... [11 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 ... [26 rule(s)] done [0.00s].  creating S4 object ... done [0.00s]. |
|  |
| |  | | --- | | > |   > inspect(head(sort(Book\_rules, by = "lift")))  lhs rhs support confidence  [1] {ChildBks,YouthBks,CookBks,DoItYBks,ArtBks} => {GeogBks} 0.0310 0.6966292  [2] {ChildBks,YouthBks,CookBks,DoItYBks,GeogBks} => {ArtBks} 0.0310 0.6078431  [3] {ChildBks,YouthBks,CookBks,DoItYBks,GeogBks} => {RefBks} 0.0270 0.5294118  [4] {ChildBks,CookBks,DoItYBks,RefBks,GeogBks} => {YouthBks} 0.0270 0.6000000  [5] {ChildBks,YouthBks,CookBks,RefBks,ArtBks} => {GeogBks} 0.0225 0.6617647  [6] {ChildBks,YouthBks,CookBks,ArtBks,GeogBks} => {DoItYBks} 0.0310 0.6666667  lift count  [1] 2.524019 62  [2] 2.522171 62  [3] 2.468120 54  [4] 2.424242 54  [5] 2.397698 45  [6] 2.364066 62 |

