**Groceries Dataset:**

**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**

**Ans:**

> library(arules)

> groceries<-read.transactions(file.choose(),format="basket")

> inspect(groceries[1:10])

items

[1] {bread,margarine,ready,

citrus,

fruit,semi-finished,

soups}

[2] {fruit,yogurt,coffee,

tropical}

[3] {milk,

whole}

[4] {,meat,

cheese,

fruit,yogurt,cream,

pip,

spreads}

[5] {bakery,

life,

milk,condensed,

milk,long,

other,

product,

vegetables,whole}

[6] {cleaner,

milk,butter,yogurt,rice,abrasive,

whole}

[7] {rolls/buns}

[8] {(appetizer),

beer,liquor,

other,

vegetables,UHT-milk,rolls/buns,bottled}

[9] {plants,

pot}

[10] {milk,cereals,

whole}

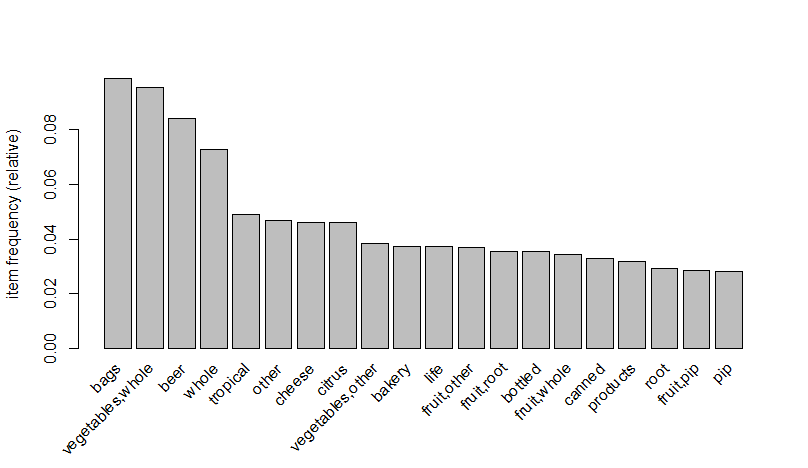
> class(groceries)

[1] "transactions"

attr(,"package")

[1] "arules"

> itemFrequencyPlot(groceries,topN=20)



1. **For support = 0.002,confidence = 0.05,minlen=3**

> groceries\_rules<-apriori(groceries,parameter = list(support = 0.002,confidence = 0.05,minlen=3))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen

0.05 0.1 1 none FALSE TRUE 5 0.002 3

maxlen target ext

10 rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 19

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

set transactions ...[6928 item(s), 9835 transaction(s)] done [0.11s].

sorting and recoding items ... [257 item(s)] done [0.00s].

creating transaction tree ... done [0.01s].

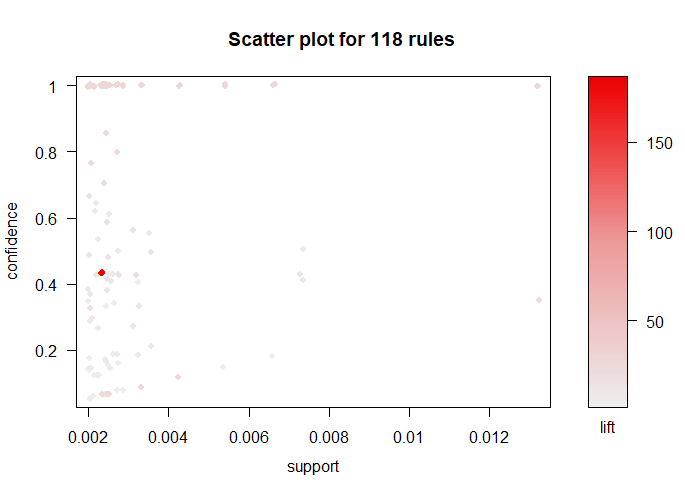
checking subsets of size 1 2 3 4 done [0.03s].

writing ... [118 rule(s)] done [0.00s].

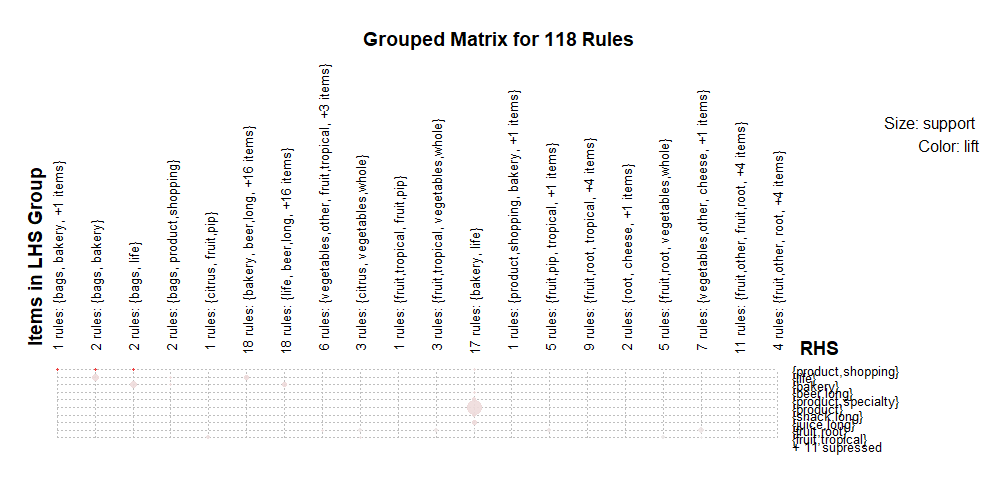
creating S4 object ... done [0.02s].

> library(arulesViz)

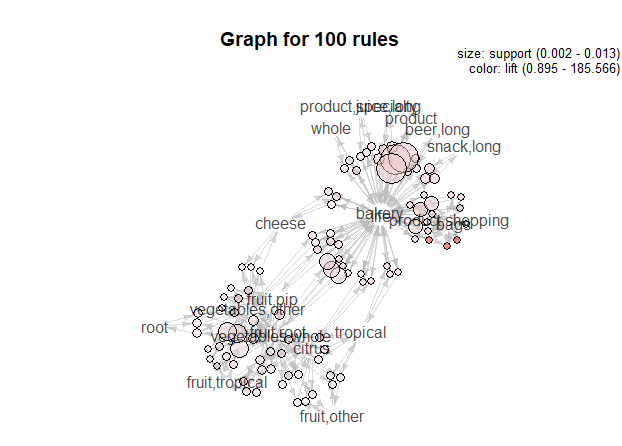
> plot(groceries\_rules,method = "scatterplot")



> plot(groceries\_rules,method = "grouped")



> plot(groceries\_rules,method = "graph")



1. **For support = 0.003,confidence = 0.05,minlen=3**

> groceries\_rules<-apriori(groceries,parameter = list(support = 0.003,confidence = 0.05,minlen=3))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen

0.05 0.1 1 none FALSE TRUE 5 0.003 3

maxlen target ext

10 rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 29

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

set transactions ...[6928 item(s), 9835 transaction(s)] done [0.13s].

sorting and recoding items ... [173 item(s)] done [0.00s].

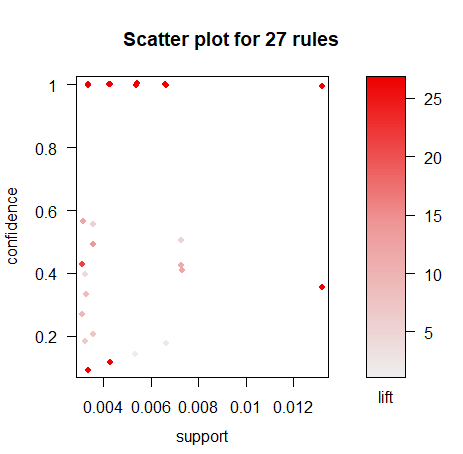
creating transaction tree ... done [0.00s].

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

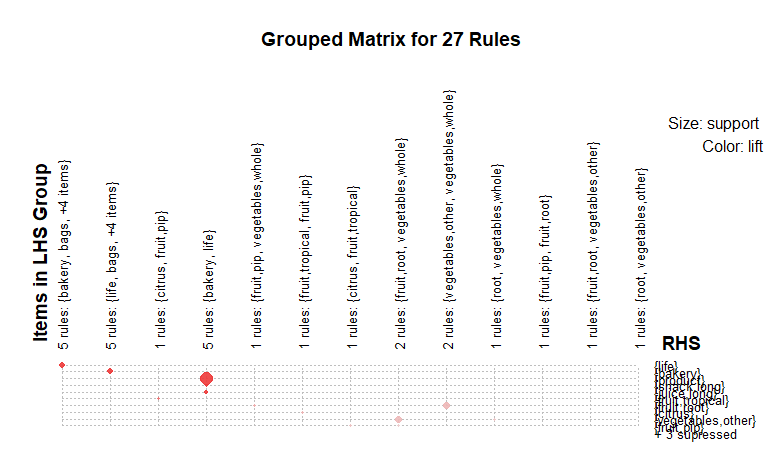
writing ... [27 rule(s)] done [0.00s].

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

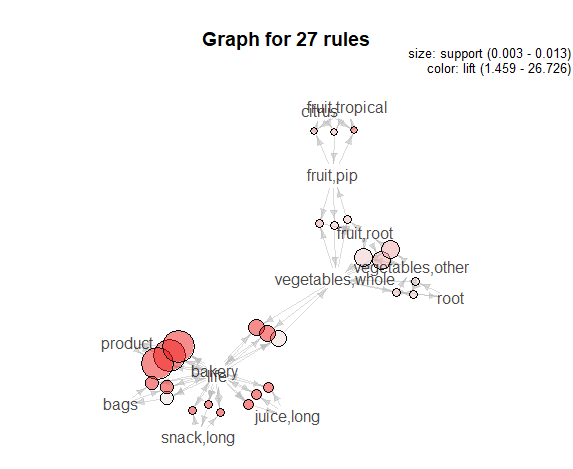
> plot(groceries\_rules,method = "scatterplot")



> plot(groceries\_rules,method = "grouped")



> plot(groceries\_rules,method = "graph")



1. **For support = 0.0025,confidence = 0.05,minlen=3**

> groceries\_rules<-apriori(groceries,parameter = list(support = 0.0025,confidence = 0.05,minlen=3))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen

0.05 0.1 1 none FALSE TRUE 5 0.0025 3

maxlen target ext

10 rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 24

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

set transactions ...[6928 item(s), 9835 transaction(s)] done [0.11s].

sorting and recoding items ... [209 item(s)] done [0.00s].

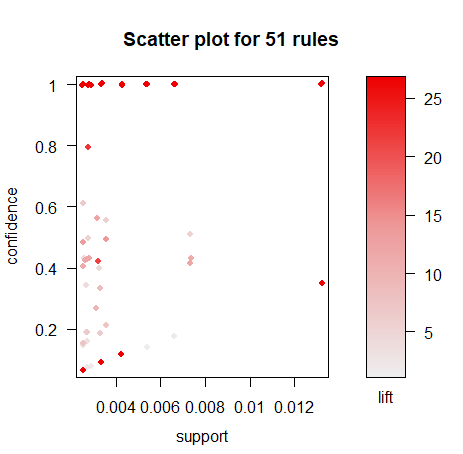
creating transaction tree ... done [0.01s].

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

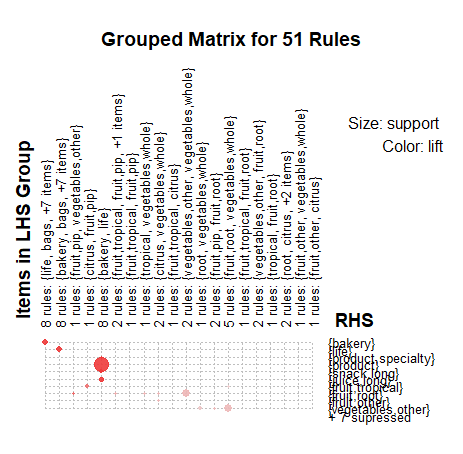
writing ... [51 rule(s)] done [0.00s].

creating S4 object ... done [0.02s].

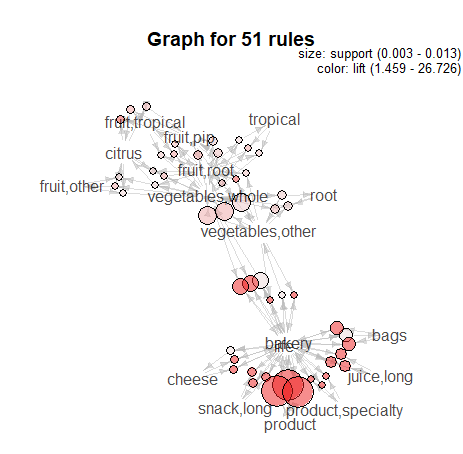
> plot(groceries\_rules,method = "scatterplot")



> plot(groceries\_rules,method = "grouped")



> plot(groceries\_rules,method = "graph")



1. **For support = 0.0035,confidence = 0.65,minlen=3**

> groceries\_rules<-apriori(groceries,parameter = list(support = 0.0035,confidence = 0.65,minlen=3))

Apriori

Parameter specification:

confidence minval smax arem aval originalSupport maxtime support minlen maxlen target ext

0.65 0.1 1 none FALSE TRUE 5 0.0035 3 10 rules FALSE

Algorithmic control:

filter tree heap memopt load sort verbose

0.1 TRUE TRUE FALSE TRUE 2 TRUE

Absolute minimum support count: 34

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

set transactions ...[6928 item(s), 9835 transaction(s)] done [0.24s].

sorting and recoding items ... [151 item(s)] done [0.00s].

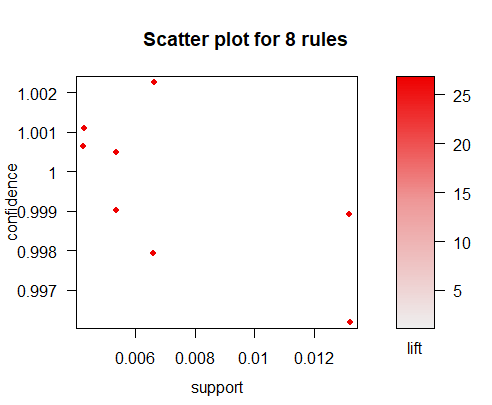
creating transaction tree ... done [0.01s].

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

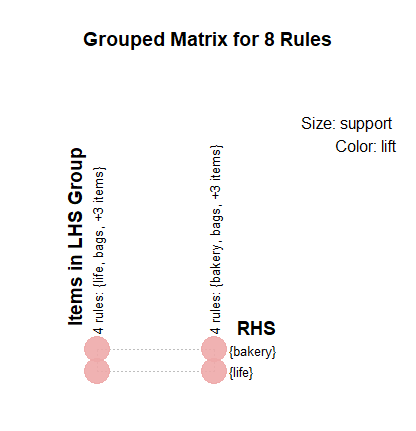
writing ... [8 rule(s)] done [0.00s].

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

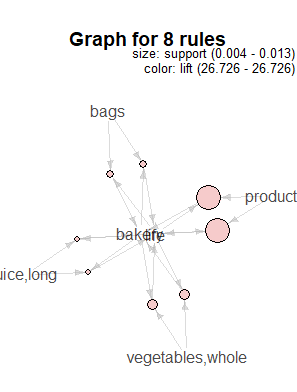
> plot(groceries\_rules,method = "scatterplot")



> plot(groceries\_rules,method = "grouped")



> plot(groceries\_rules,method = "graph")



**So From Above 4 cases we observed that,**

1. **Lower the Confidence level Higher the no. of rules.**
2. **Lower the minlen(minimum length), Higher the no. of Rules are getting generated.**
3. **Higher the Support, lower the no. of rules.**
4. **So, we can conclude that these 3 parameters different combinations can generate different rules.**