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

##importing the data set

library(arules)

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

inspect(groceries[1:10])

class(groceries)

# itemFrequencyPlot can be applicable only for transaction data

# count of each item from all the transactions

itemFrequencyPlot(groceries,topN=20)

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

##118 rules are generated from the the data

library(arulesViz)

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

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

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

plot(groceries\_rules,method = "mosaic")

# Sorting rules by confidence

rules\_conf <- sort(groceries\_rules,by="confidence")

inspect(rules\_conf)

# Sorint rules by lift ratio

rules\_lift <- sort(groceries\_rules,by="lift")

inspect(rules\_lift)

##Changing the supprort and trying in the dataset

groceries\_rules2<-apriori(groceries,parameter = list(support = 0.005,confidence = 0.05,minlen=3))

summary(groceries\_rules2)

##if we increase the support value to 0.005 then it is only giving 12 rules

plot(groceries\_rules2,method = "grouped")

plot(groceries\_rules2,method = "graph")

##now we see by increasing the confidence

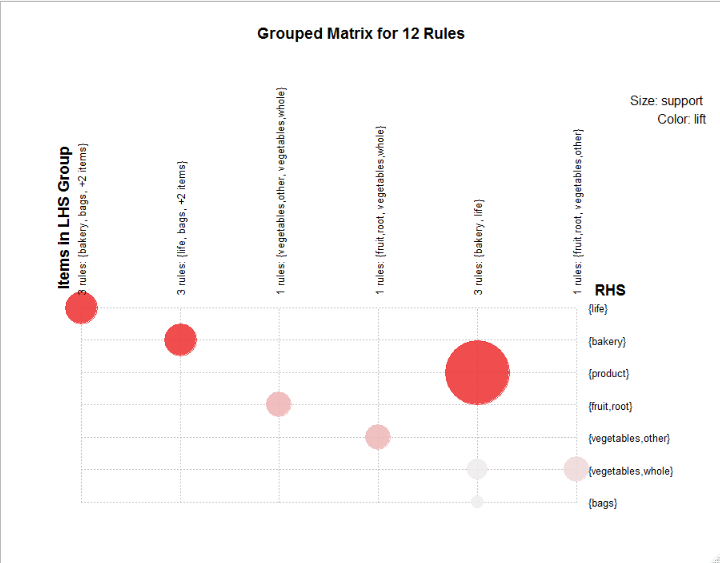
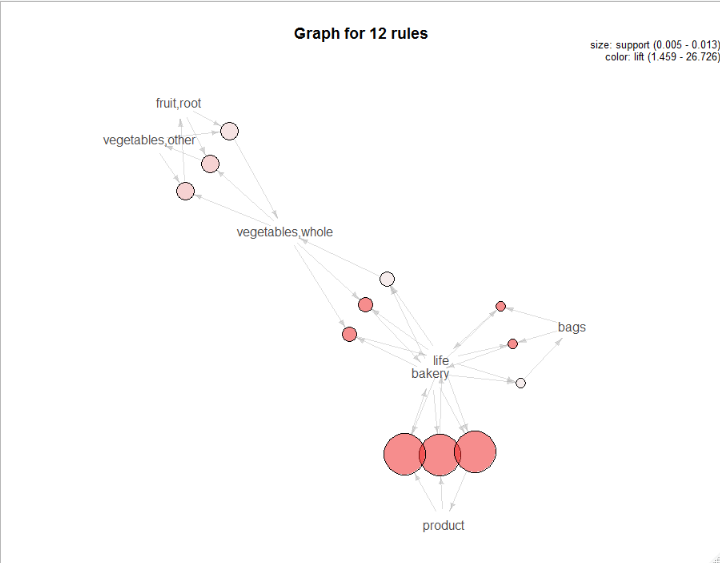
groceries\_rules1<-apriori(groceries,parameter = list(support = 0.002,confidence = 0.09,minlen=3))

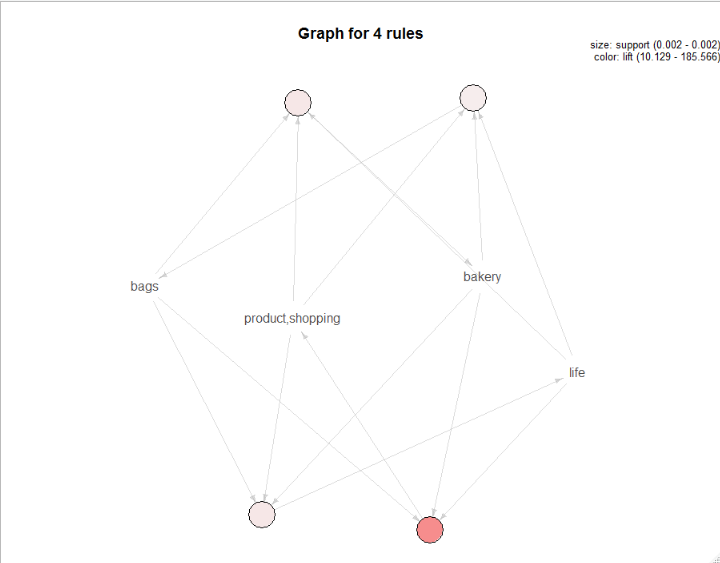
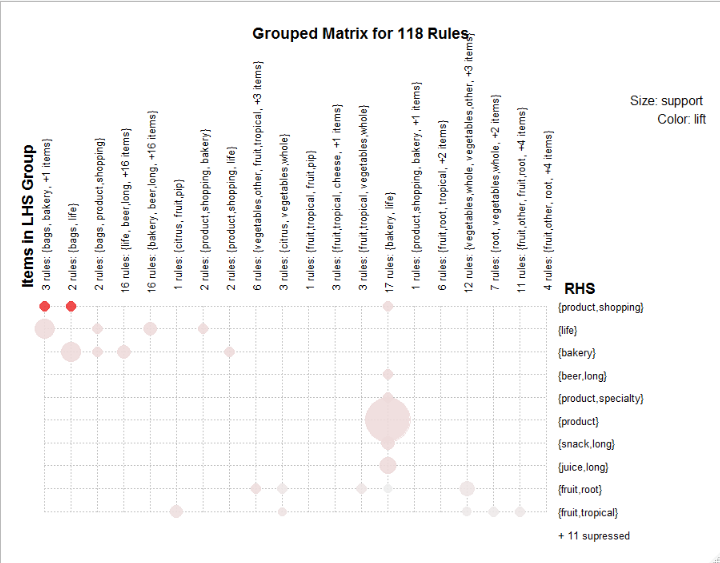
##105rules is obtained by increasing confidence to 4

##now we see by increasing the min len

groceries\_rules3<-apriori(groceries,parameter = list(support = 0.002,confidence = 0.05,minlen=4))

##we got only 4 rules





Book dataset

##in book data set I have a query regarding the min len in the basket the model is not performing with min len greater than the 1 but any how I have done it by keeping the model length as the 1

With this association rules formed please clear the doubt

##importing the data set

library(arules)

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

inspect(book[1:10])

class(book)

# itemFrequencyPlot can be applicable only for transaction data

# count of each item from all the transactions

itemFrequencyPlot(book,topN=15)

book\_rules<-apriori(book,parameter = list(support = 0.0009,confidence = 0.005,minlen=1))

summary(book\_rules)

##28 rules are generated from the the data

library(arulesViz)

plot(book\_rules,method = "scatterplot")

plot(book\_rules,method = "graph")

# Sorting rules by confidence

rules\_conf <- sort(book\_rules,by="confidence")

inspect(rules\_conf)

# Sorint rules by lift ratio

rules\_lift <- sort(book\_rules,by="lift")

inspect(rules\_lift)

##Changing the supprort and trying in the dataset

book\_rules1<-apriori(book,parameter = list(support = 0.009,confidence = 0.005,minlen=1))

summary(book\_rules1)

##14 rules generated by increasing the support value

plot(book\_rules1,method = "scatterplot")

plot(book\_rules1,method = "graph")

##now we see by decresing the confidence

book\_rules2<-apriori(book,parameter = list(support = 0.0009,confidence = 0.001,minlen=1))

summary(book\_rules2)

##122 rules is obtained by decresing confidence

plot(book\_rules2,method = "scatterplot")

plot(book\_rules2,method = "graph")

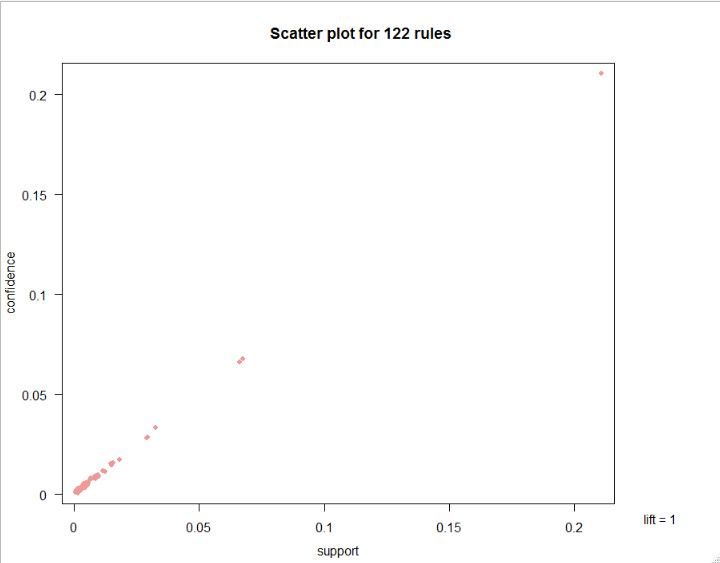
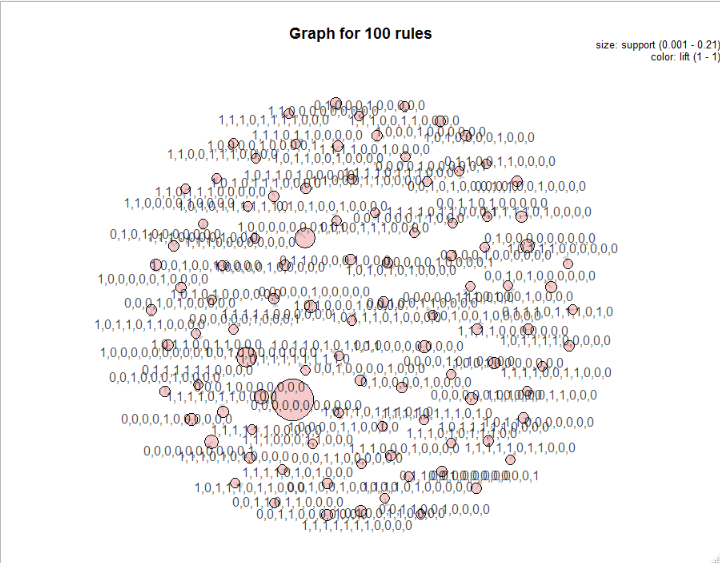
##now we see by decreasing the min len

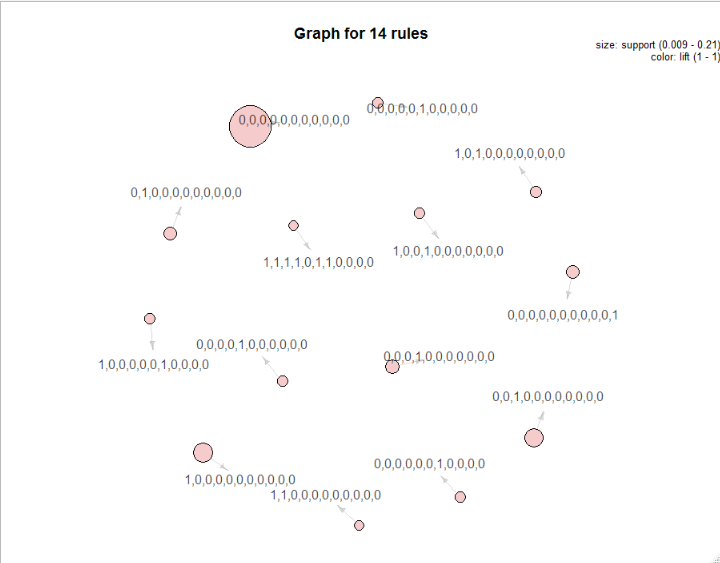
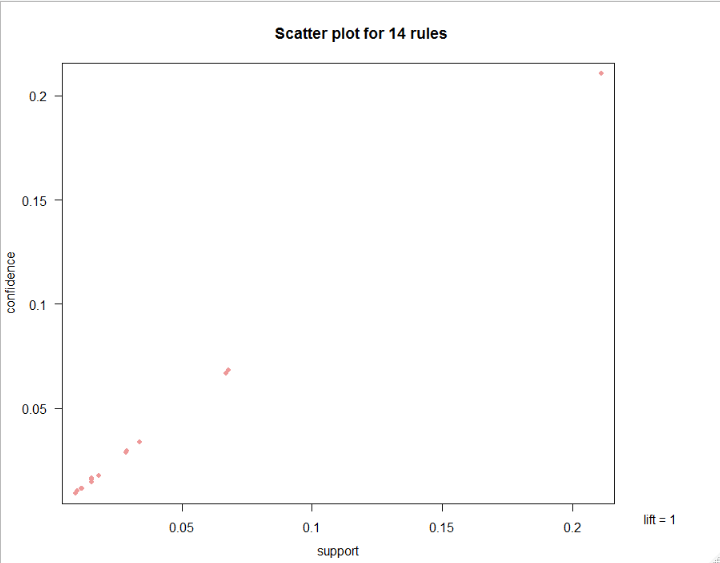
book\_rules3<-apriori(book,parameter = list(support = 0.001,confidence = 0.0001,minlen=1))

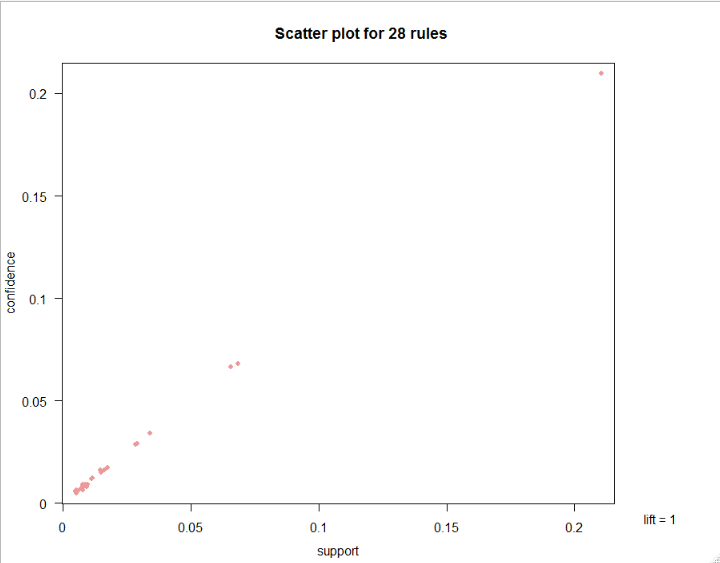
summary(book\_rules3)

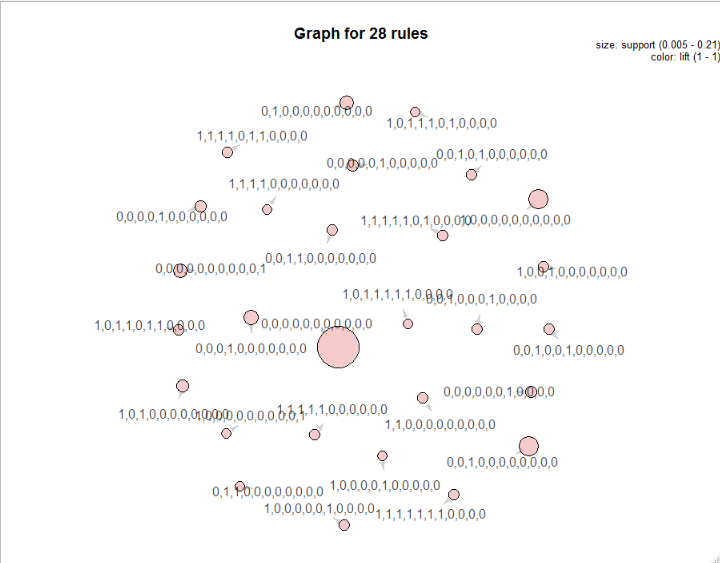
plot(book\_rules3,method = "graph")

##here the change in minlen does not gives any rules so the data set does not contain the rules with the increase in min len









My movies data set

##importing the data set

library(arules)

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

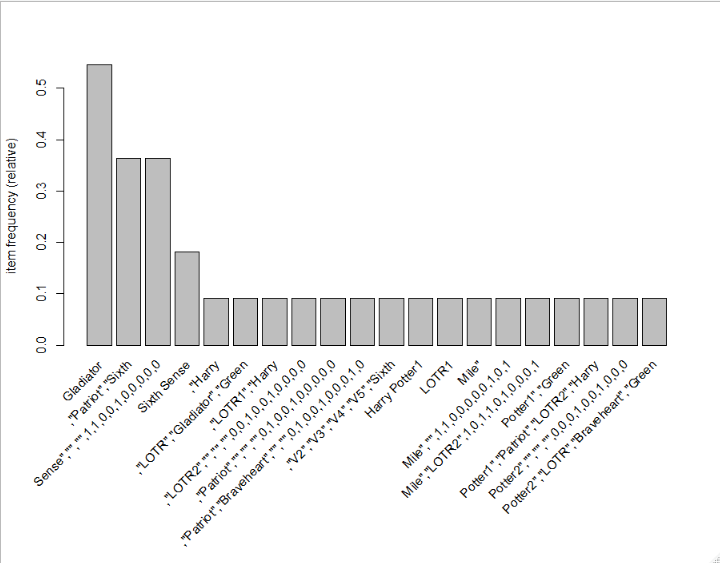
inspect(mov[1:10])

class(mov)

# itemFrequencyPlot can be applicable only for transaction data

# count of each item from all the transactions

itemFrequencyPlot(mov,topN=20)



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

summary(mov\_rules)

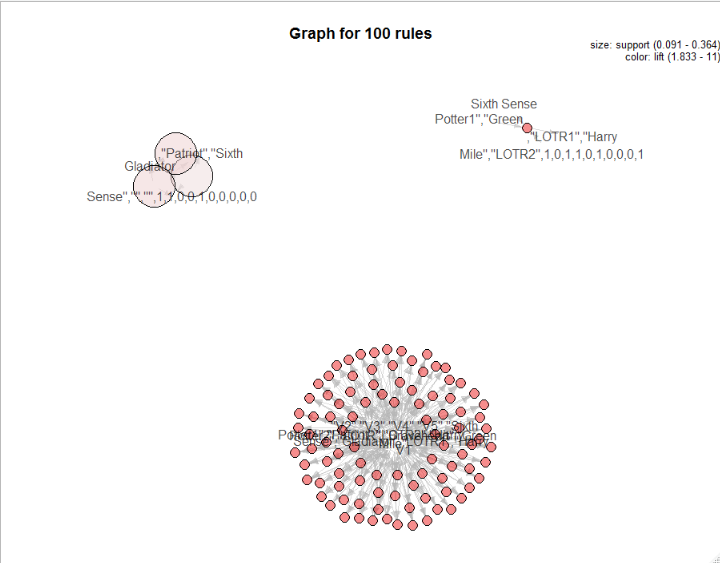
##181 rules are generated from the the data

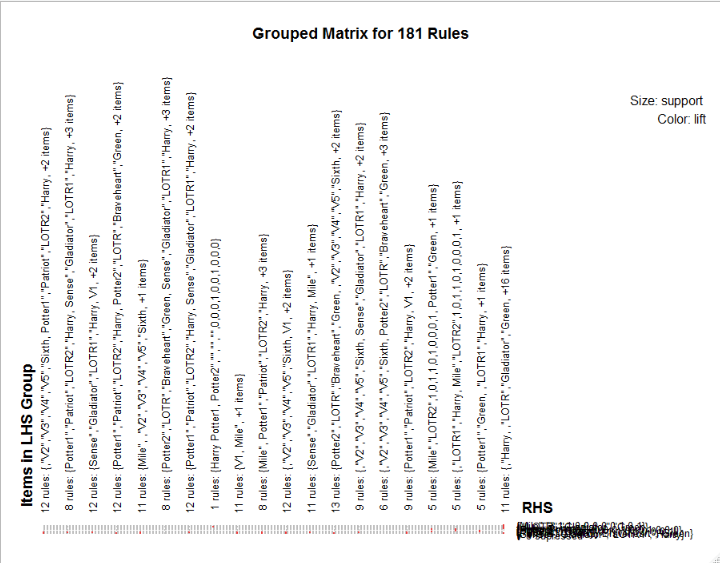
library(arulesViz)

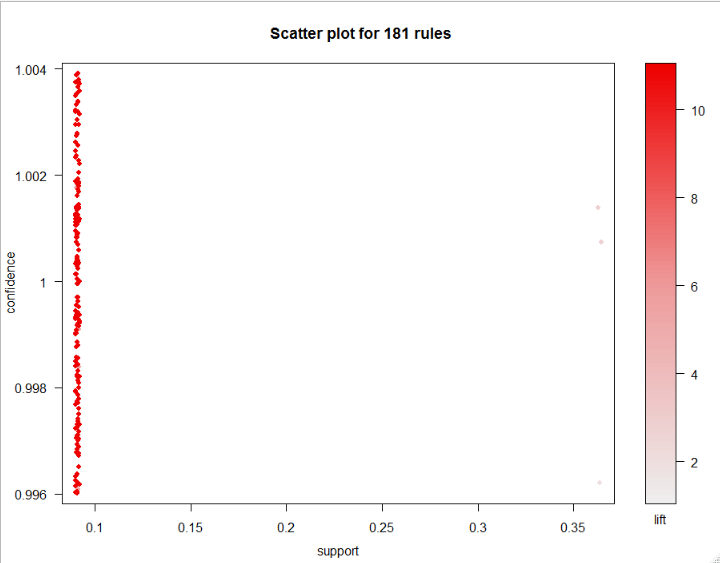
plot(mov\_rules,method = "scatterplot")

plot(mov\_rules,method = "grouped")

plot(mov\_rules,method = "graph")







Sorting rules by confidence

rules\_conf <- sort(mov\_rules,by="confidence")

inspect(rules\_conf)

# Sorint rules by lift ratio

rules\_lift <- sort(mov\_rules,by="lift")

inspect(rules\_lift)

##Changing the supprort and trying in the dataset

mov\_rules1 <- apriori(mov,parameter = list(support = 0.0156,confidence = 0.05,minlen=3))

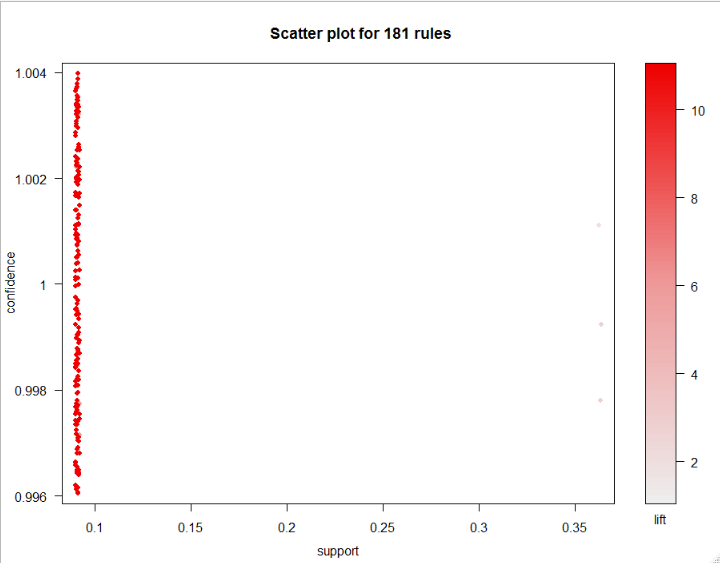
summary(mov\_rules1)

##if we increase the support value to 0.0156 but the st of rules remain same there are no addition of new rules are decrease in old rules

plot(mov\_rules1,method = "grouped")

plot(mov\_rules1,method = "graph")

plot(mov\_rules1,method = "scatterplot")



The above graphs are same as the before graph there is no change

##now we see by increasing the confidence

mov\_rules2<-apriori(mov,parameter = list(support = 0.05,confidence = 1,minlen=3))

summary(mov\_rules2)

plot(mov\_rules2,method = "grouped")

plot(mov\_rules2,method = "graph")

plot(mov\_rules2,method = "scatterplot")

##181rules there is no change in the rules after changing the confidence to 1

##now we see by incresing the min len

mov\_rules3<-apriori(mov,parameter = list(support = 0.05,confidence = 0.05,minlen=4))

summary(mov\_rules3)

##100 rules the rules have been decreased indicating the basket lenght of 4 contains less transactions satisfying our con,support

plot(mov\_rules3,method = "graph")

plot(mov\_rules3,method = "grouped")

plot(mov\_rules3,method = "scatterplot")

