**Nome Completo dos Alunos: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**Tarefa Aula 1 – Exploração de SNA em R**

**Questões Base**

* Explore as rotinas **Exemplo Rede.R** e **Exemplo Rede Two Mode.R** . Rode os códigos na plataforma R utilizando como base as tabelas Rede One Mode\_Tarefa Aula 1\_Paulista T4.xlsx e Rede Two Mode\_Tarefa Aula 1\_Paulista T4.xlsx. (atenção: não são as mesmas bases trabalhadas em sala).

**Análise das Medidas de Centralidade:**

Grau (nº de laços que um ator possui):

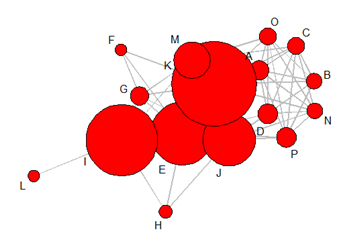
* Há 16 nós, sendo que os nós que mais possuem arestas são: K (11), J (10), E (9) e A (9).

Proximidade / Closeness (distância de um ator ao outro):

* Quanto maior o índice, maior a proximidade. Nesse caso, os nós K (0,79), E (0,71) e J (0,71)são os que estão mais próximos; enquanto o nó L (0,39) é o mais afastado.

Intermediação / Betweeness (interação entre atores não adjacentes):

* Índices iguais a zero indicam que o nó pode ser eliminado. Nesse caso, os nós L e F podem sair que não alteram a rede.
* Índices elevados indicam que o nó é ponte e, se forem eliminados, dividem a rede. Nesse caso, os nós ponte por ordem de importância são, K (18,70), I (15,31), E (13,18), J (10,48) e M (6,32).

> rede1 <- read.table("D:/0 Aula 28\_07\_18 Redes sociais/SNA e Text Mining Paulista T4/Trabalho 1/Rede One Mode\_Trab1.csv",header=TRUE,sep = ";", dec=",")

>

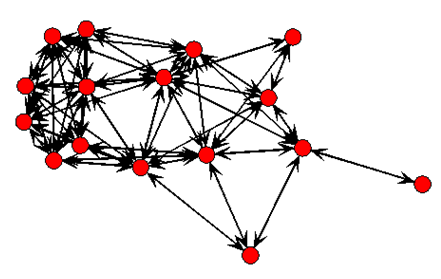
> # Construindo a rede a partir da matriz de relações (0 e 1)

> grede1 <- rede1[,2:17]

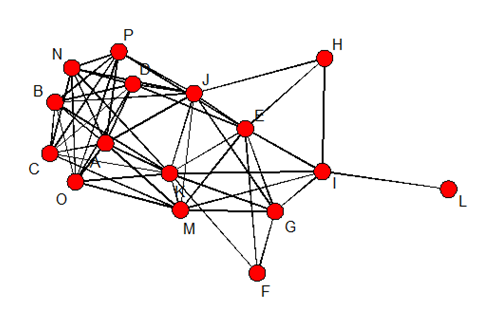
> rownames(grede1) <- rede1[,1]

> par(mfrow=c(1,1))

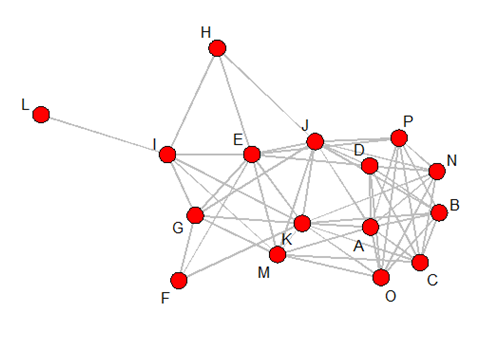
> gplot(grede1)



> gplot(grede1,gmode="graph",displaylabels = TRUE)



> gplot(grede1,gmode="graph",displaylabels = TRUE,edge.col="gray",usearrows=FALSE)



> # Explorando a rede

> degree(grede1,gmode="graph",cmode="indegree")

[1] 9 7 7 7 9 3 6 3 6 10 11 1 8 7 7 7

> closeness(grede1,gmode="graph")

[1] 0.6818182 0.6250000 0.6000000 0.6250000 0.7142857 0.5357143 0.6250000 0.5172414 0.6250000

[10] 0.7142857 0.7894737 0.3947368 0.6818182 0.6250000 0.6000000 0.6250000

> # qto maior o closeness, mais proximo

> betweenness(grede1,gmode="graph")

[1] 1.953846 1.010989 1.371429 2.079121 13.177289 0.000000 1.733333 0.400000 15.307692

[10] 10.481685 18.699634 0.000000 6.323443 1.010989 1.371429 2.079121

> # qto maior, mais ponte é (não pode ser eliminado, pois rompe a rede)

> # zero significa que pode ser eliminado que não impacta a rede

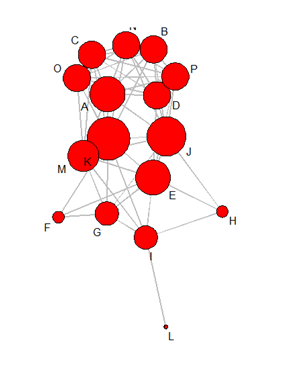
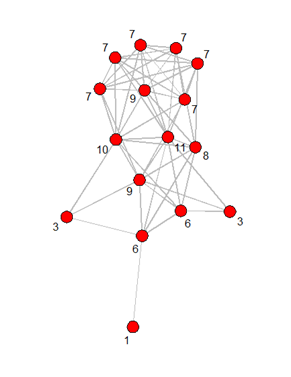
> # Aprimorando a representação da rede

> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.cex=degree(grede1,gmode="graph",cmode="indegree")/3)

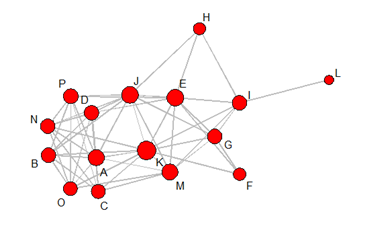
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,label=degree(grede1,gmode="graph",cmode="indegree"))

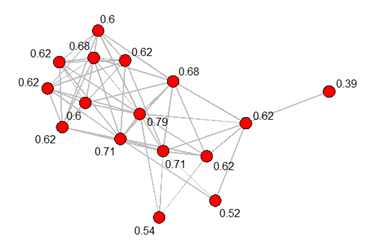
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.cex=closeness(grede1,gmode="graph")\*2)



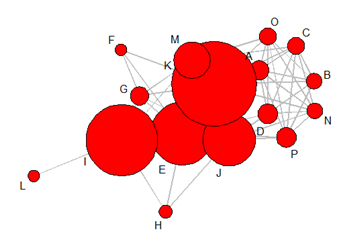
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,label=round(closeness(grede1,gmode="graph"),digits=2))



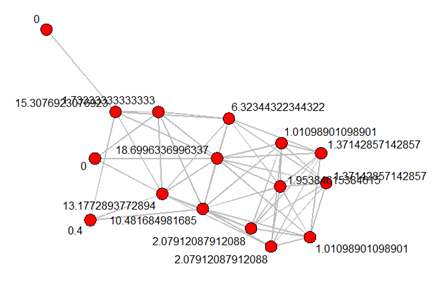
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.cex=betweenness(grede1,gmode="graph")/3+1)



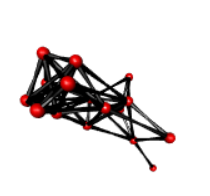
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,label=betweenness(grede1,gmode="graph"))



> # Gráfico 3D da rede

> gplot3d(grede1)



* Faça pequenas modificações na tabela e veja seus resultados.

**Análise do Impacto das Alterações (eliminação nós L e F):**

Grau (nº de laços que um ator possui):

* Antes de eliminar os nós L e F, tínhamos 16 nós no total. Após a alteração, temos 30 nós. Aparentemente, os nós da rede foram duplicados para espelhar as 2 situações (a anteiror e a atual).
* Se observaros a tabela resumo dos dados e o gráfico abaixo, vemos que o número de laços se manteve, independente da eliminação dos nós L e F *(obs. 1 na tabela abaixo).*

Proximidade / Closeness (distância de um ator ao outro):

* Quanto maior o índice, maior a proximidade. Nesse caso, os nós K (0,56) e J (0,54) são os mais próximos *(obs. 2 na tabela abaixo)*, diferente da rede anterior, quando os os maiores índices eram dos nós K (0,79), E (0,71) e J (0,71).
* Agora, o nó mais afastado é o H (0,40).

Intermediação / Betweeness (interação entre atores não adjacentes):

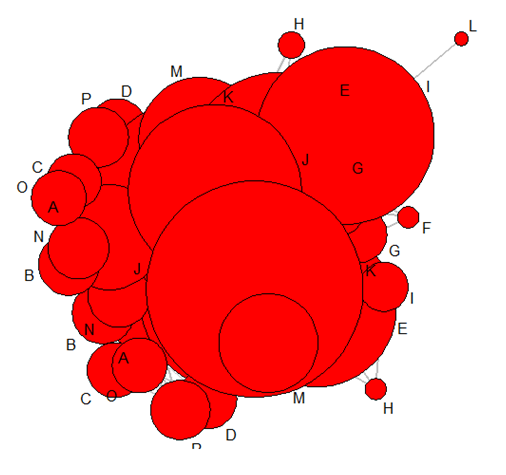
* Índices iguais a zero indicam que o nó pode ser eliminado. Nesse caso, os nós L e F saíram da rede e no momento, nenhum nó possui intermediação zero.
* Índices elevados indicam que o nó é ponte e, se forem eliminados, dividem a rede. Nesse caso, temos mais nós ponte e com índices mais elevados *(obs.3 na tabela abaixo)*. Por ordem de importância são, K’ (43,52), I’ (35,10), J’ (34,15), E’ (33,39), A’ (19,47), M’ (18,26), G’ (11,19) E N’ (10,13).



Obs. 1

Obs. 2

Obs. 3



> library(sna)

> library(network)

> library(rgl)

> library(ggplot2)

> rede1 <- read.table("D:/0 Aula 28\_07\_18 Redes sociais/SNA e Text Mining Paulista T4/Trabalho 1/Rede One Mode\_Trab1\_alt.csv",header=TRUE,sep = ";", dec=",")

> rede1

Label A B C D E G H I J K M N O P

1 A 0 1 1 1 0 0 0 0 1 1 1 1 1 1

2 B 1 0 1 1 0 0 0 0 1 1 0 0 1 1

3 C 1 1 0 1 0 0 0 0 0 1 1 1 0 1

4 D 1 1 1 0 1 0 0 0 1 0 0 1 1 0

5 E 0 0 0 1 0 1 1 1 1 1 1 0 0 1

6 F 0 0 0 0 1 1 0 0 0 1 0 0 0 0

7 G 0 0 0 0 1 0 0 1 1 1 1 0 0 0

8 H 0 0 0 0 1 0 0 1 1 0 0 0 0 0

9 I 0 0 0 0 1 1 1 0 0 1 1 0 0 0

10 J 1 1 0 1 1 1 1 0 0 1 1 1 0 1

11 K 1 1 1 0 1 1 0 1 1 0 1 1 1 0

12 L 0 0 0 0 0 0 0 1 0 0 0 0 0 0

13 M 1 0 1 0 1 1 0 1 1 1 0 0 1 0

14 N 1 0 1 1 0 0 0 0 1 1 0 0 1 1

15 O 1 1 0 1 0 0 0 0 0 1 1 1 0 1

16 P 1 1 1 0 1 0 0 0 1 0 0 1 1 0

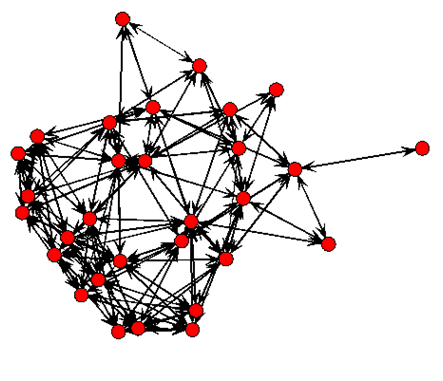
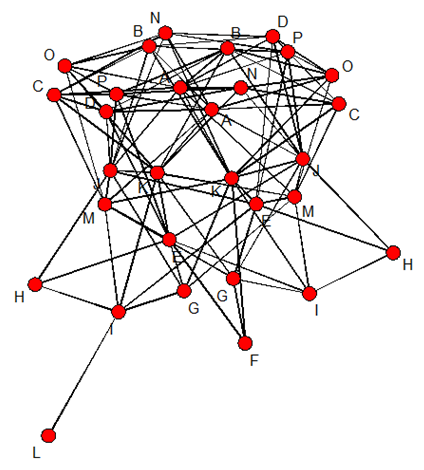
> grede1 <- rede1[,2:15]

> rownames(grede1) <- rede1[,1]

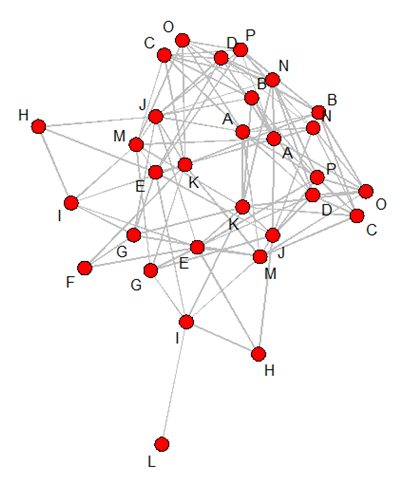
> par(mfrow=c(1,1))

> gplot(grede1)

> gplot(grede1,gmode="graph",displaylabels = TRUE)

> gplot(grede1,gmode="graph",displaylabels = TRUE,edge.col="gray",usearrows=FALSE)



> degree(grede1,gmode="graph",cmode="indegree")

[1] 9 7 7 7 8 3 5 3 5 10 10 1 8 7 7 7 9 7 7 7 9 6 3 6 10 11 8 7 7 7

> closeness(grede1,gmode="graph")

[1] 0.5178571 0.4833333 0.4677419 0.4833333 0.5178571 0.4264706 0.4677419 0.4142857 0.4531250

[10] 0.5370370 0.5576923 0.3222222 0.5178571 0.4833333 0.4677419 0.4833333 0.5178571 0.4833333

[19] 0.4677419 0.4833333 0.5178571 0.4677419 0.4027778 0.4677419 0.5370370 0.5576923 0.5000000

[28] 0.4833333 0.4677419 0.4833333

> betweenness(grede1,gmode="graph")

[1] 19.365920 10.320878 8.802899 9.900651 30.739995 1.608203 8.951653 2.654506 7.518634

[10] 34.502691 38.222642 0.000000 23.386898 10.320878 8.802899 9.900651 19.467364 10.128902

[19] 8.810357 9.751273 33.393228 11.193444 1.528938 35.100030 34.152563 43.519501 18.263868

[28] 10.128902 8.810357 9.751273

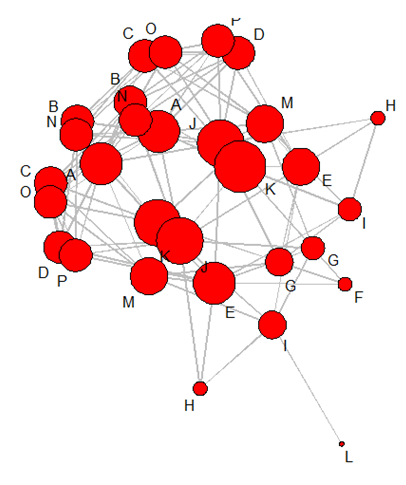
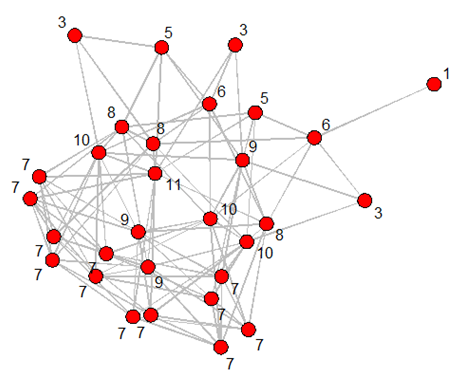
>

> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.cex=degree(grede1,gmode="graph",cmode="indegree")/3)

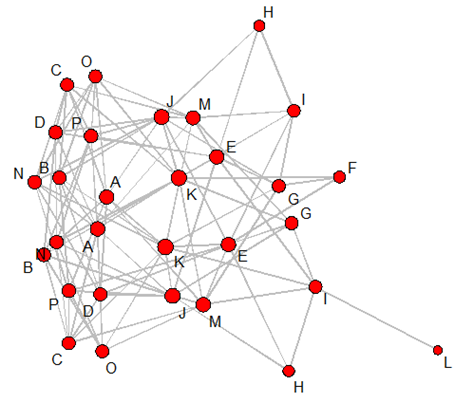
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,label=degree(grede1,gmode="graph",cmode="indegree"))

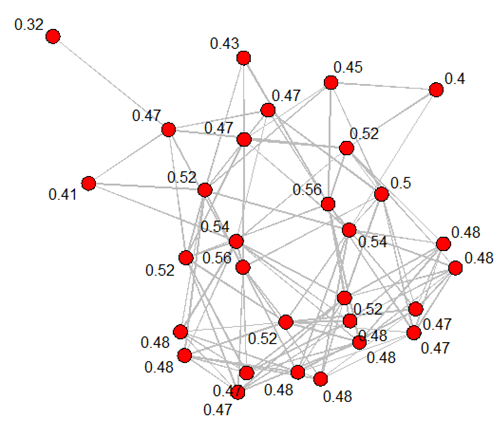
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.cex=closeness(grede1,gmode="graph")\*2)



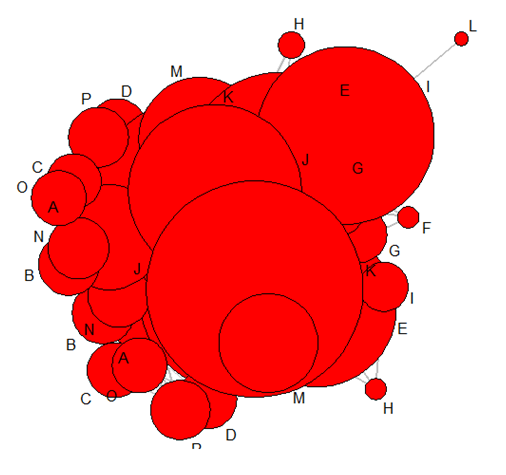
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,label=round(closeness(grede1,gmode="graph"),digits=2))



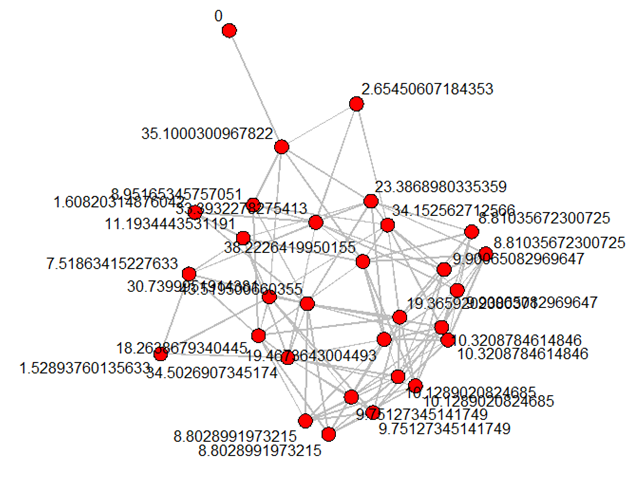
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.cex=betweenness(grede1,gmode="graph")/3+1)



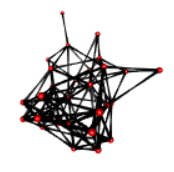
> gplot(grede1,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,label=betweenness(grede1,gmode="graph"))



> # Gráfico 3D da rede1

> gplot3d(grede1)



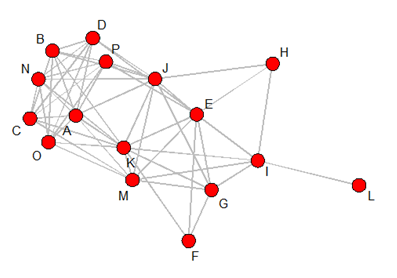
* Inclua outras análises em seu código (usando as extensões **sna**, **network** ou **igraph**) e comente os resultados (seja criativo!).
* Os gráficos são iguais, independentemente do tipo da extensão !!!!!

> # Extensão sna

> gplot(grede,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.sna=degree(grede,gmode="graph",cmode="indegree")/3)

There were 15 warnings (use warnings() to see them)

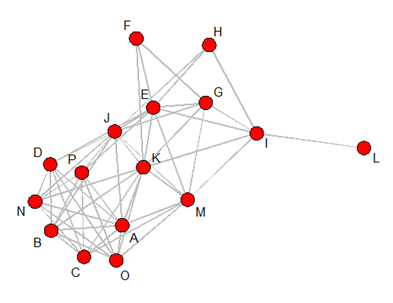


> # Extensão network

> gplot(grede,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.network=degree(grede,gmode="graph",cmode="indegree")/3)

There were 15 warnings (use warnings() to see them)

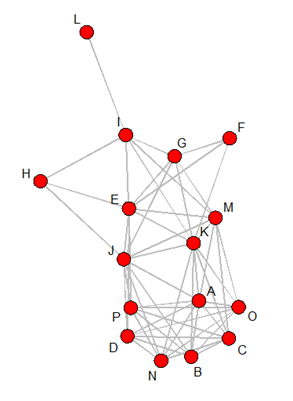


> # Extensão igrapgh

> gplot(grede,gmode="grede",displaylabels = TRUE,

+ edge.col="gray",usearrows=FALSE,vertex.igraph=degree(grede,gmode="graph",cmode="indegree")/3)

There were 15 warnings (use warnings() to see them)



* Desafio: Baseado na tabela da Rede Two Mode desta tarefa, faça uma análise de agrupamento (cluster analysis) do tipo hierárquico aglomerativo (dendrograma) das pessoas ou dos produtos adquiridos por elas, levando em consideração apenas a estrutura de relações entre elas. Comente como implementou e discuta os resultados, comparando com a rede construída. Utilize a plataforma R e o script de exemplo de uso de *Cluster Analysis* em R.

*Dica: após a seleção dos grupos, desenhe a rede e represente os nós das pessoas (ou produtos) com cores de acordo com o grupo correspondente.*

> rede2 <- read.table("D:/0 Aula 28\_07\_18 Redes sociais/SNA e Text Mining Paulista T4/Trabalho 1/Rede Two Mode\_Trab1.csv",header=TRUE,sep = ";", dec=",")

>

> head(rede2)

X iPhone iPad Livro.Harry.Potter jogo.MineCraft Camisa.do.Corinthians Bola.de.Futebol Flauta.Transversal

1 João 1 1 0 1 1 0 1

2 Maria 1 1 1 1 1 1 1

3 José 0 0 0 0 1 0 1

4 Paulo 1 0 0 0 1 0 0

5 Pedro 0 0 1 0 1 1 0

6 Luisa 0 0 1 0 1 1 0

Lista.Telefônica Caixa.de.Fósforos Calculadora Detergente

1 1 1 1 0

2 1 1 1 1

3 0 0 1 0

4 0 0 1 0

5 1 1 1 0

6 1 1 1 0

> row.names(rede2) <- rede2$X

>

> # Implementa o algoritmo hierárquico e apresenta o dendrograma

> hc <- hclust(dist(rede2), "average") # explorar com outros métodos de distância

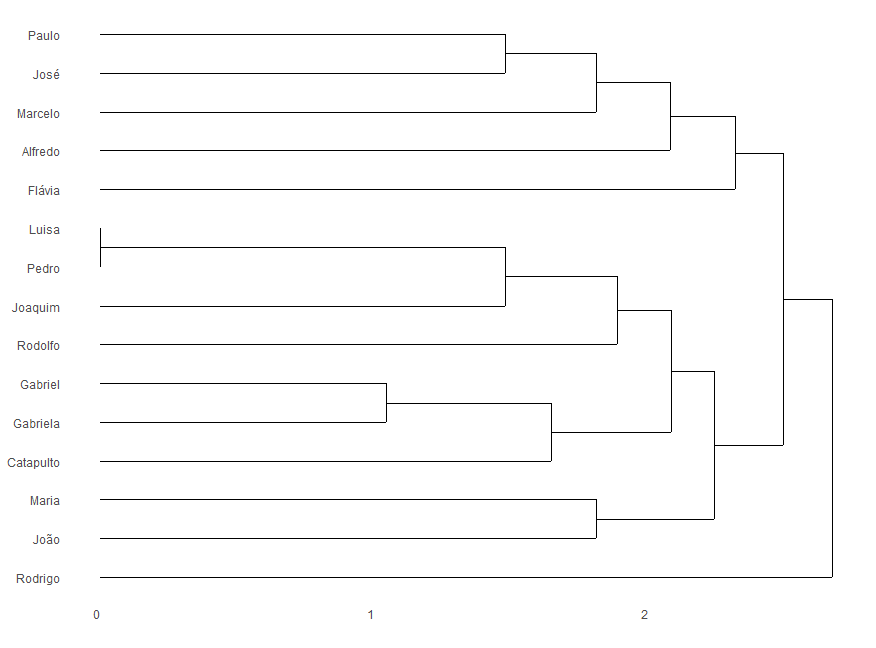
Warning message:

In dist(rede2) : NAs introduced by coercion

> p <- ggdendrogram(hc, rotate=FALSE)

> print(p)

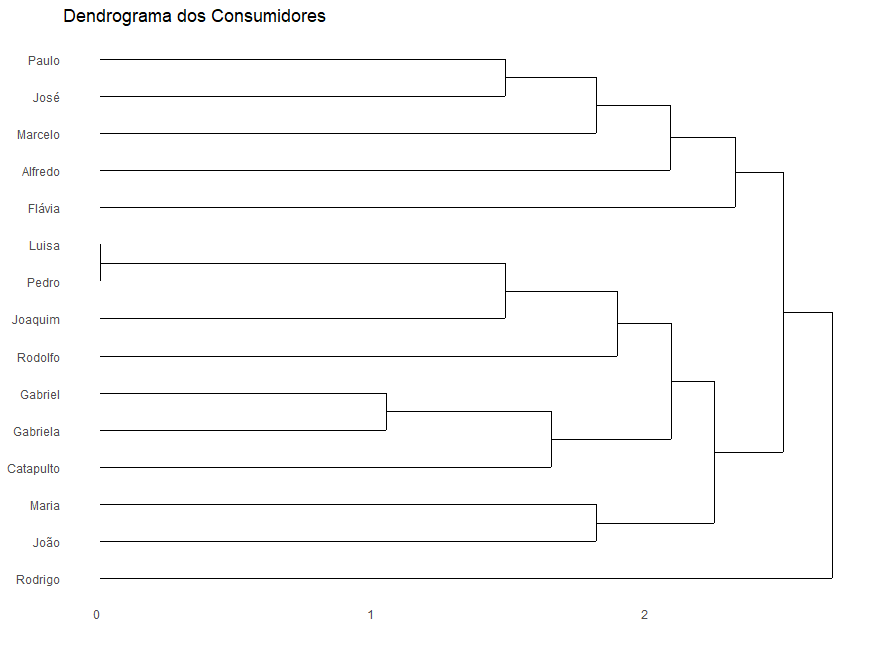
> ggdendrogram(hc, rotate=TRUE)



> hcdata <- dendro\_data(hc)

> ggdendrogram(hcdata, rotate=TRUE, size=2) + labs(title="Dendrograma dos Consumidores")

>



> # "Cortando" a árvore em 4 grupos

> grupos <- cutree(hc,k=4)

> grupos

João Maria José Paulo Pedro Luisa Marcelo Alfredo Joaquim Gabriela Flávia Catapulto

1 1 2 2 1 1 2 2 1 1 3 1

Rodrigo Gabriel Rodolfo

4 1 1

>

> # Analisando as principais variáveis a partir dos grupos

> boxplot(rede2$iPhone ~ grupos, col = "blue", main = 'Box Plot do iPhone')

> boxplot(rede2$iPad ~ grupos, col = "blue", main = 'Box Plot do iPad')

> boxplot(rede2$Livro.Harry.Potter ~ grupos, col = "blue", main = 'Box Plot do Livro Harry Potter')

> boxplot(rede2$jogo.MineCraft ~ grupos, col = "blue", main = 'Box Plot do Jogo MineCraft')

> boxplot(rede2$Camisa.do.Corinthians ~ grupos, col = "blue", main = 'Box Plot da Camisa do Corinthians')

> boxplot(rede2$Bola.de.Futebol ~ grupos, col = "blue", main = 'Box Plot da Bola de Futebol')

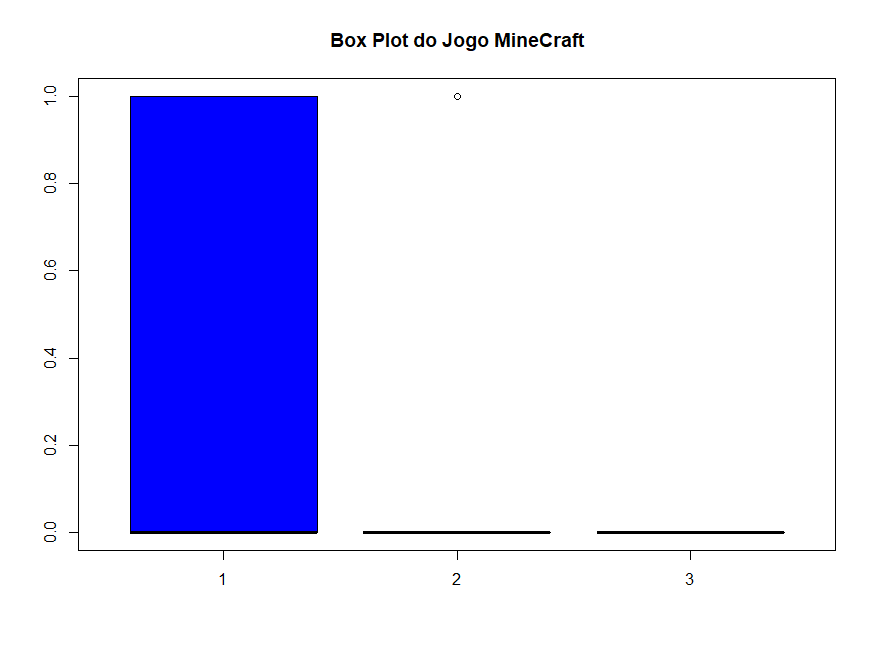
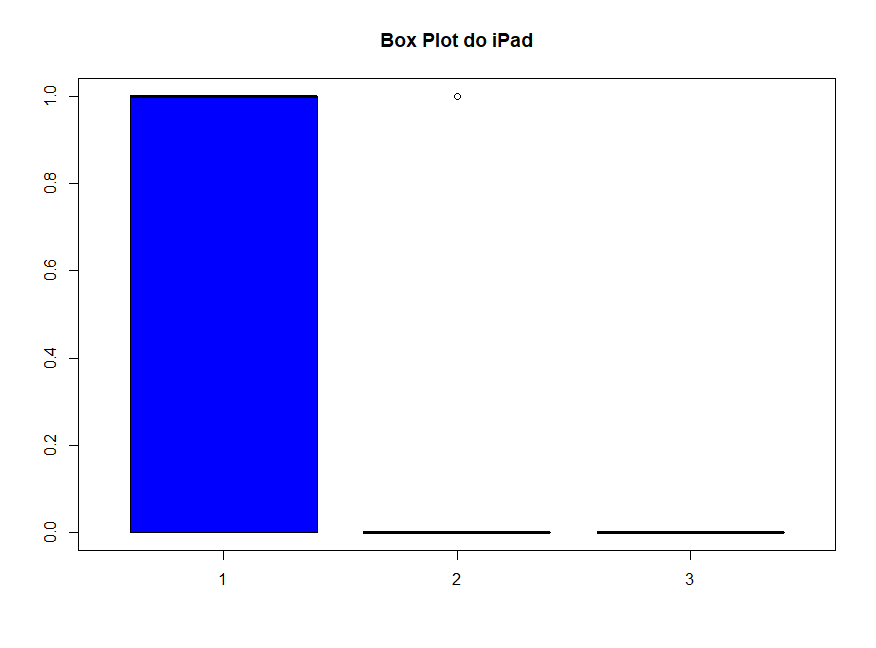
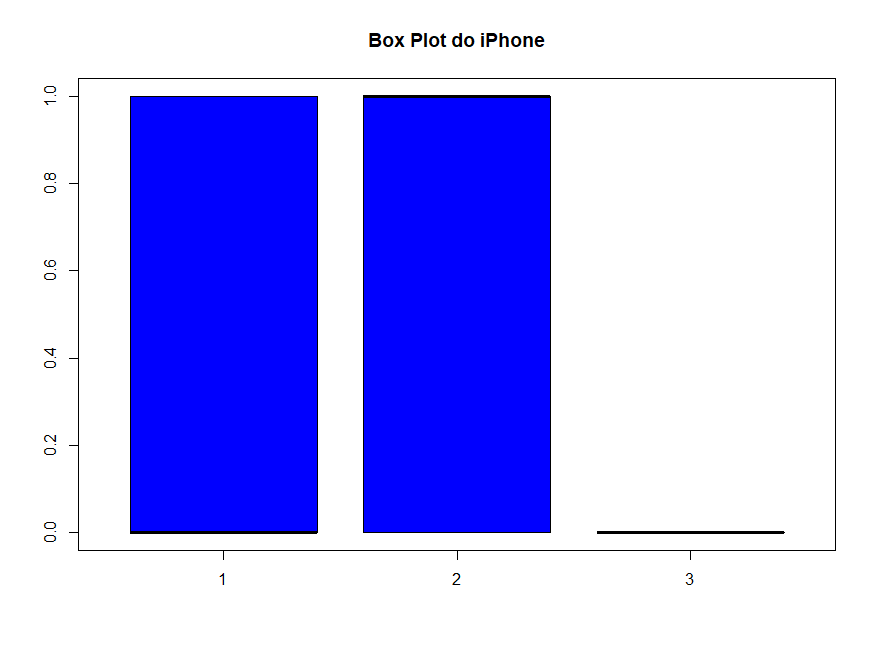
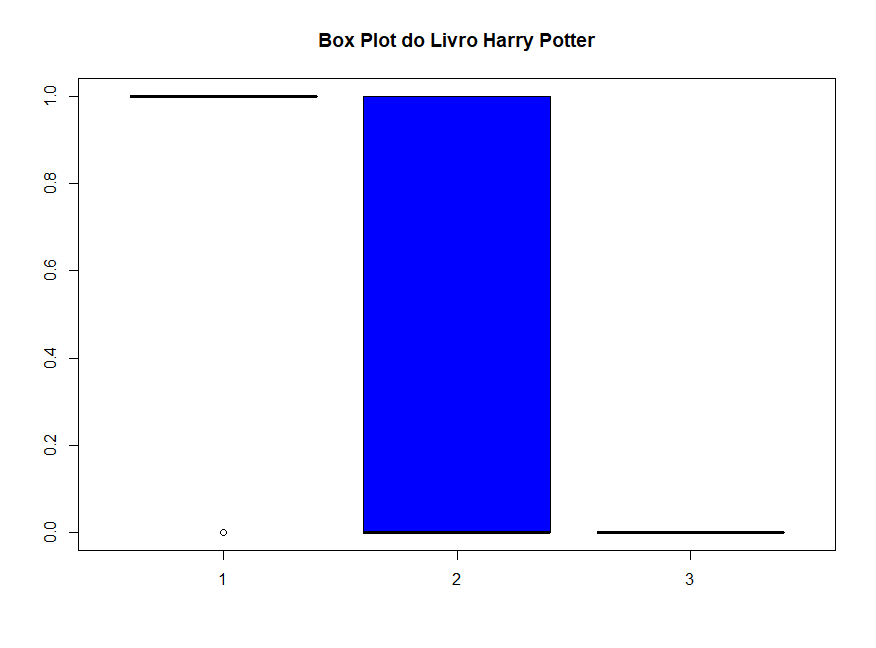
> boxplot(rede2$Flauta.Transversal ~ grupos, col = "blue", main = 'Box Plot da Flauta Transversal')

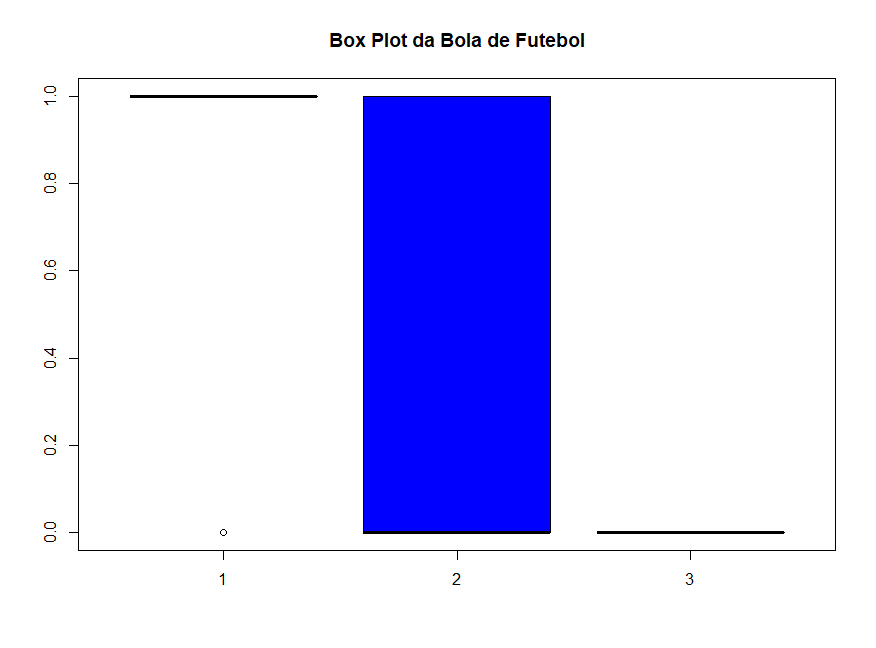
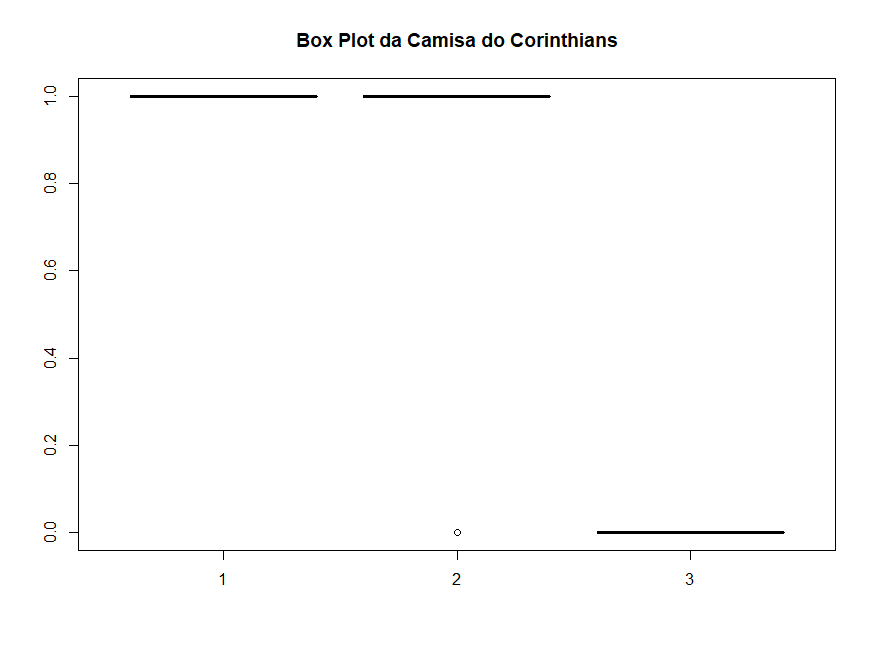
> boxplot(rede2$Lista.Telefônica ~ grupos, col = "blue", main = 'Box Plot da Lista Telefônica')

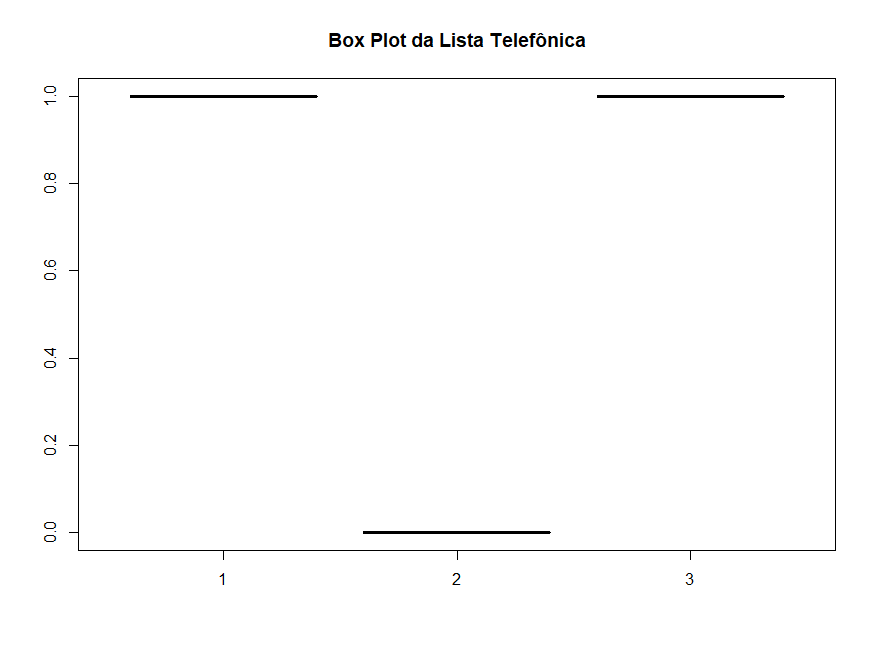
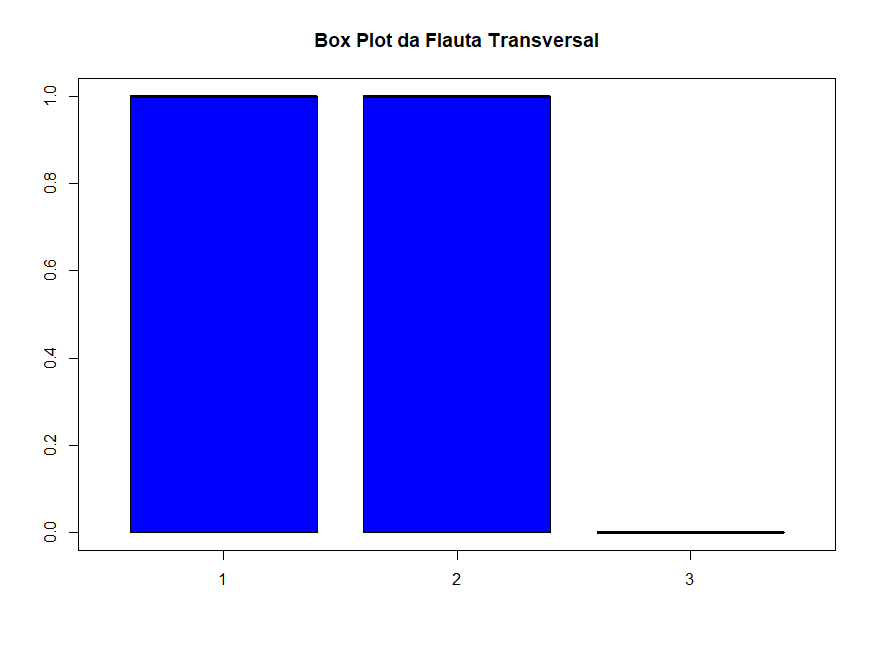
> boxplot(rede2$Caixa.de.Fósforos ~ grupos, col = "blue", main = 'Box Plot da Caixa de Fósforos')

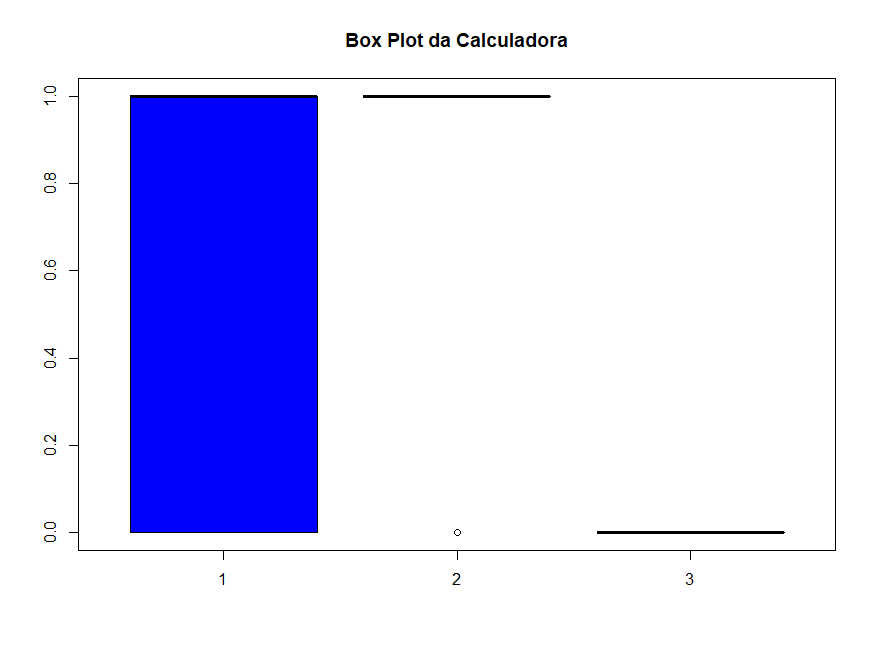
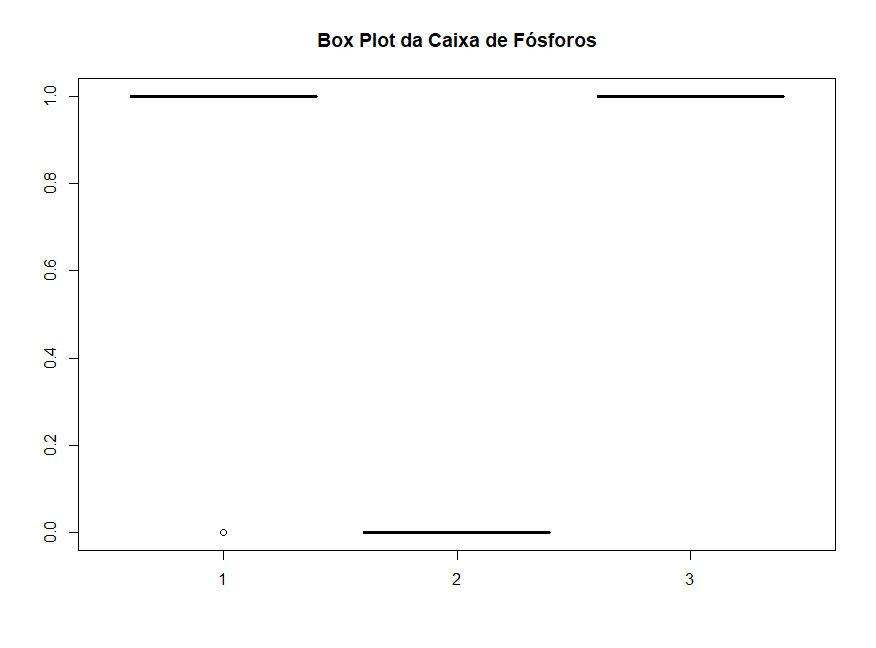
> boxplot(rede2$Calculadora ~ grupos, col = "blue", main = 'Box Plot da Calculadora')

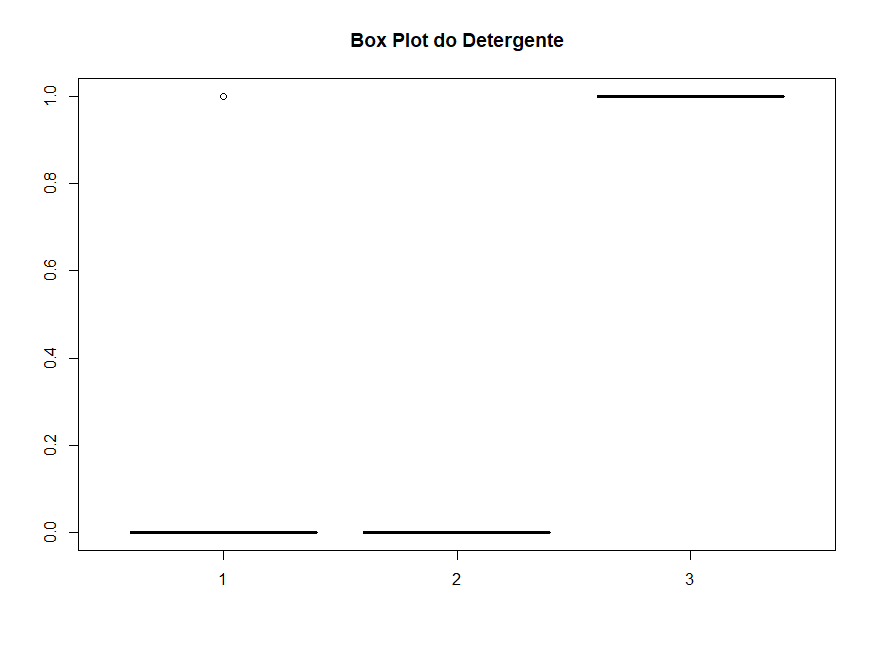
> boxplot(rede2$Detergente ~ grupos, col = "blue", main = 'Box Plot do Detergente')











> # Refazendo os clusters, agora com as variáveis padronizadas

> rede2\_padr <- rede2

> for (i in 2:12) rede2\_padr[,i] <- scale(rede2\_padr[,i])

>

> # Implementa o algoritmo hierárquico e apresenta o dendrograma

> hc2 <- hclust(dist(rede2\_padr), "average") # explorar com outros métodos de distância

Warning message:

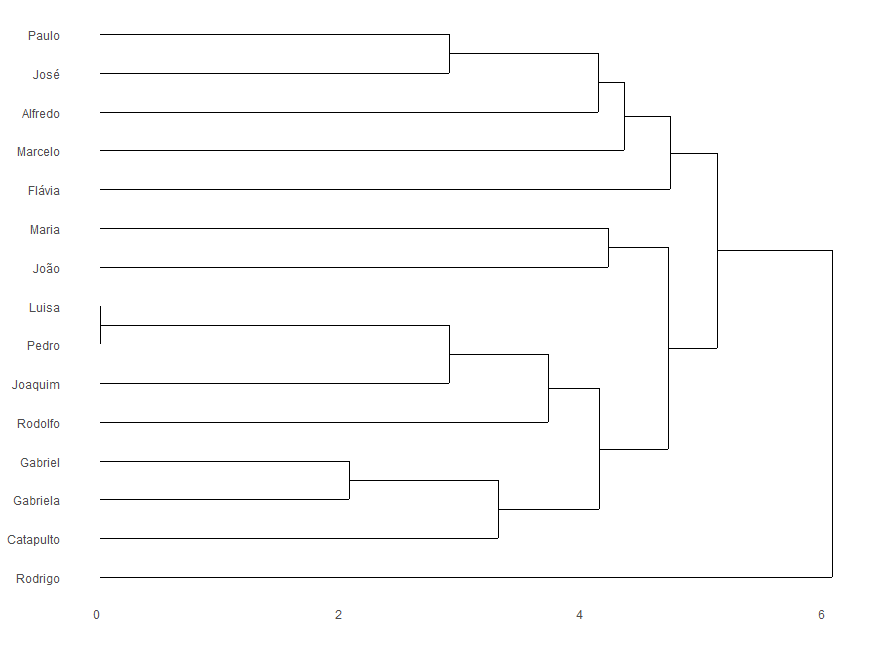
In dist(rede2\_padr) : NAs introduced by coercion

> p <- ggdendrogram(hc2, rotate=FALSE)

> print(p)

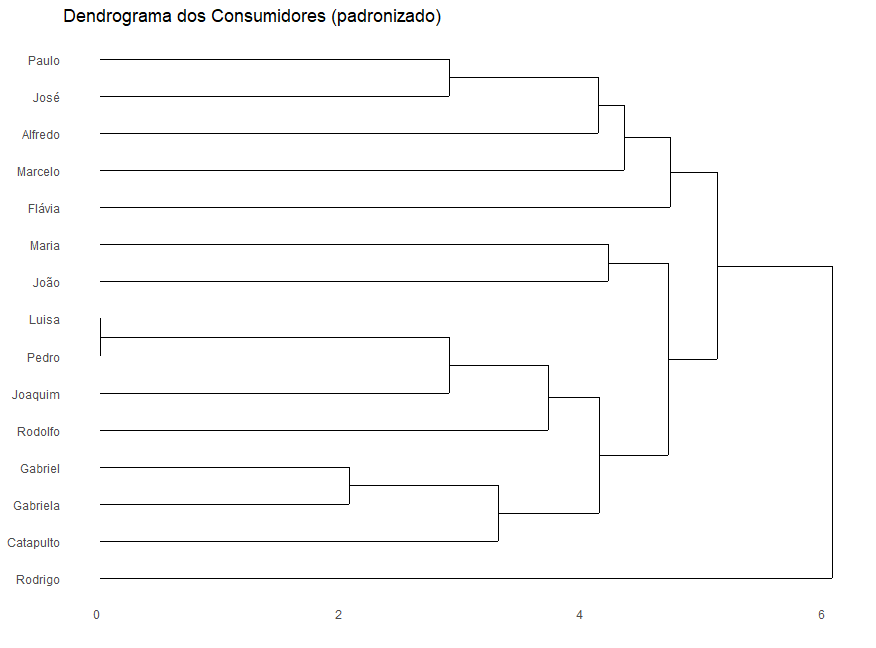
> ggdendrogram(hc2, rotate=TRUE)

>



> hcdata2 <- dendro\_data(hc2)

> ggdendrogram(hcdata2, rotate=TRUE, size=2) + labs(title="Dendrograma dos Consumidores (padronizado)")



> # "Cortando" a árvore em 4 grupos

> grupos2 <- cutree(hc2,k=4)

> grupos2

João Maria José Paulo Pedro Luisa Marcelo Alfredo Joaquim Gabriela Flávia Catapulto

1 1 2 2 1 1 2 2 1 1 3 1

Rodrigo Gabriel Rodolfo

4 1 1

> # Analisando as principais variáveis a partir dos grupos2

> # Analisando as principais variáveis a partir dos grupos2

> boxplot(rede2$iPhone ~ grupos2, col = "red", main = 'Box Plot do iPhone')

> boxplot(rede2$iPad ~ grupos2, col = "red", main = 'Box Plot do iPad')

> boxplot(rede2$Livro.Harry.Potter ~ grupos2, col = "red", main = 'Box Plot do Livro Harry Potter')

> grupos2 <- cutree(hc2,k=4)

> boxplot(rede2$jogo.MineCraft ~ grupos2, col = "red", main = 'Box Plot do Jogo MineCraft')

> boxplot(rede2$Camisa.do.Corinthians ~ grupos2, col = "red", main = 'Box Plot da Camisa do Corinthians')

> boxplot(rede2$Bola.de.Futebol ~ grupos2, col = "red", main = 'Box Plot da Bola de Futebol')

> boxplot(rede2$Flauta.Transversal ~ grupos2, col = "red", main = 'Box Plot da Flauta Transversal')

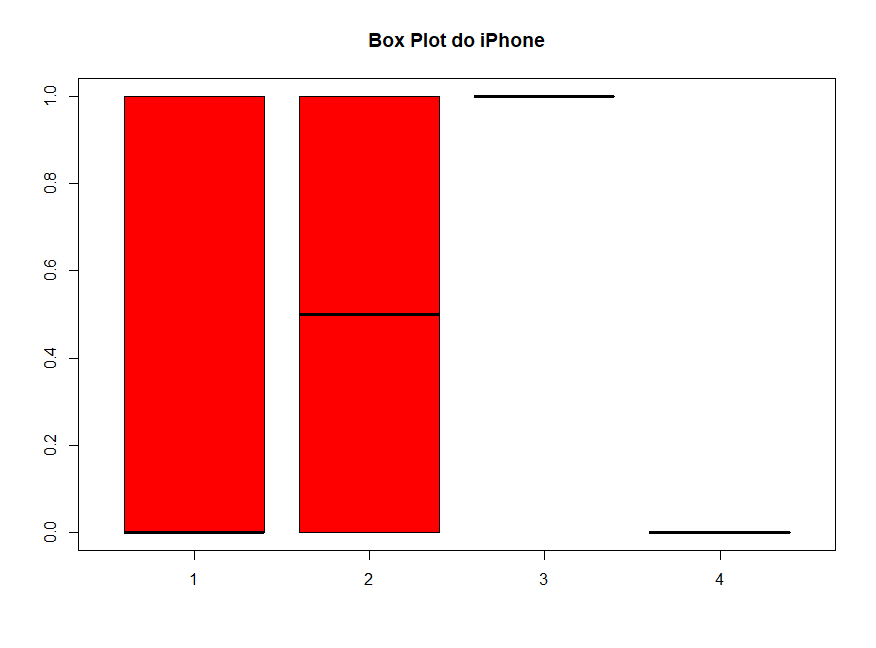
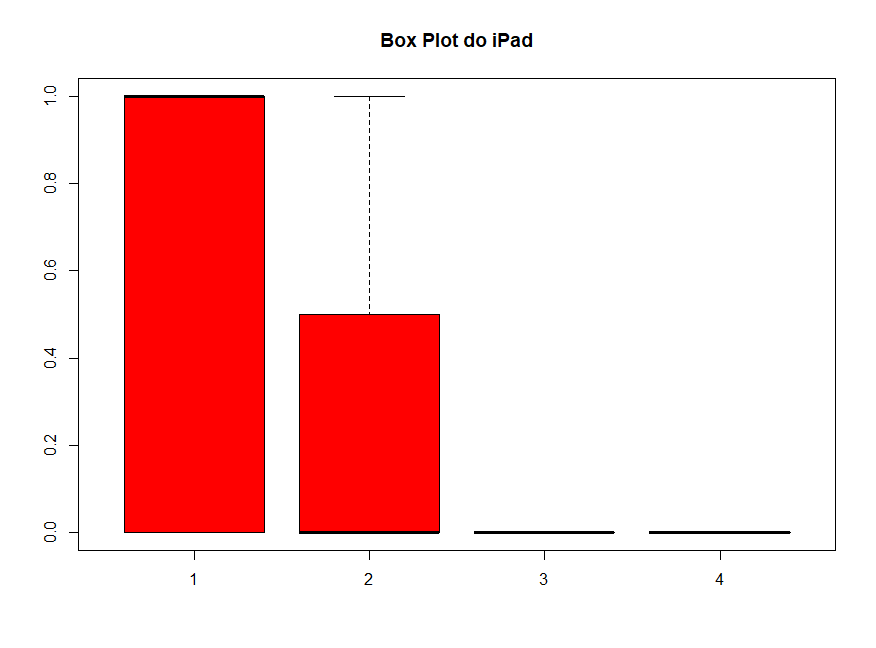
> boxplot(rede2$Lista.Telefônica ~ grupos2, col = "red", main = 'Box Plot da Lista Telefônica')

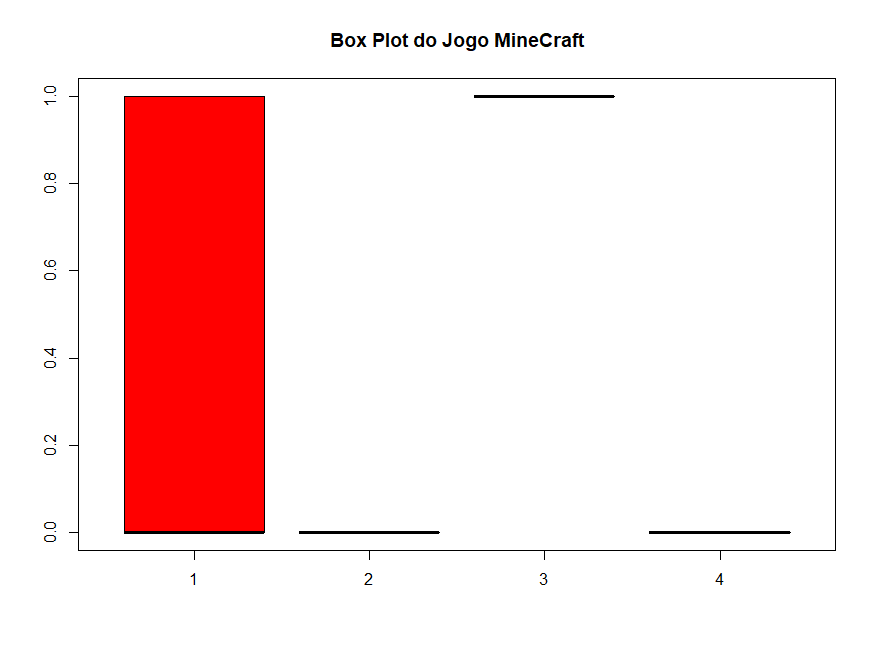
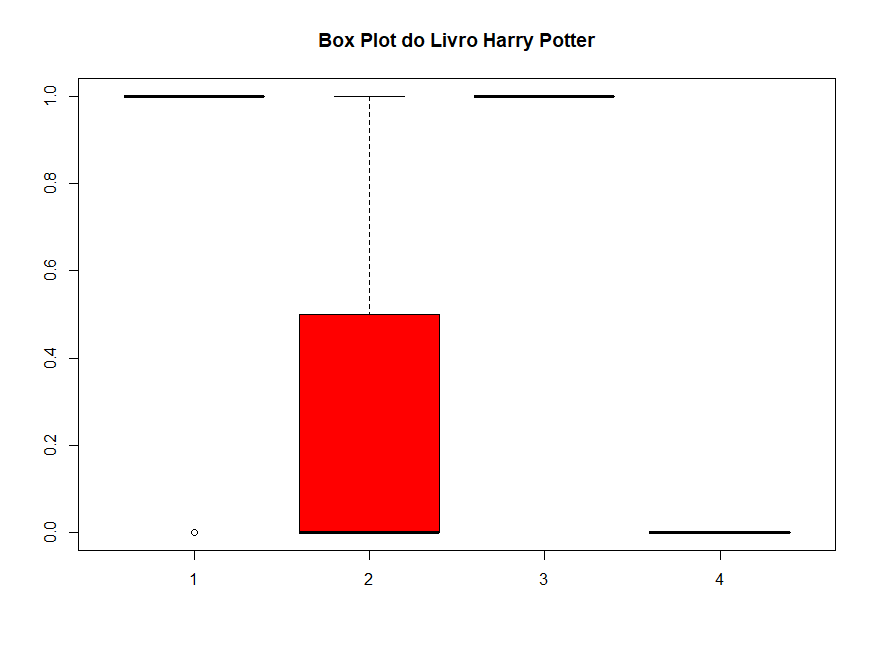
> boxplot(rede2$Caixa.de.Fósforos ~ grupos2, col = "red", main = 'Box Plot da Caixa de Fósforos')

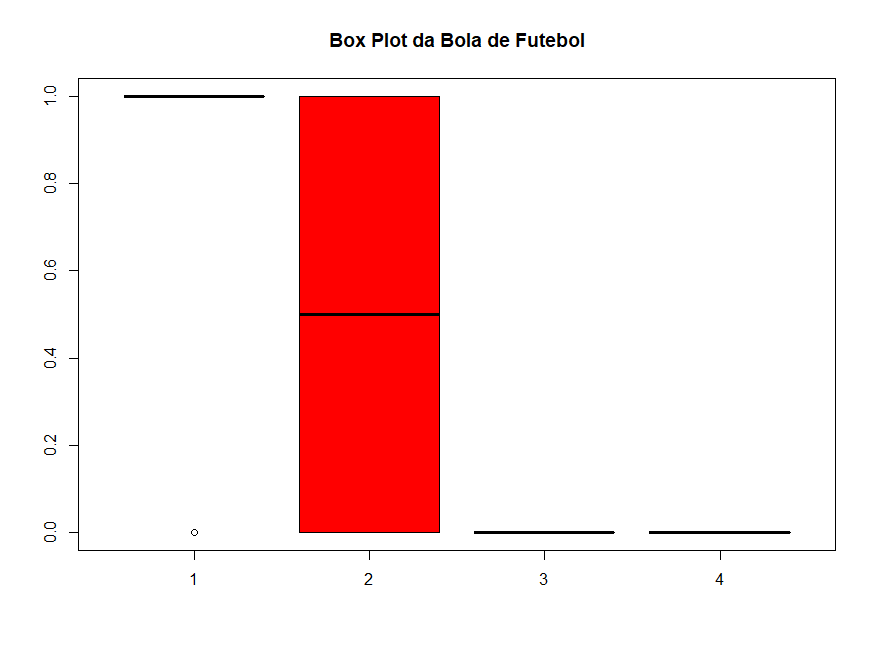
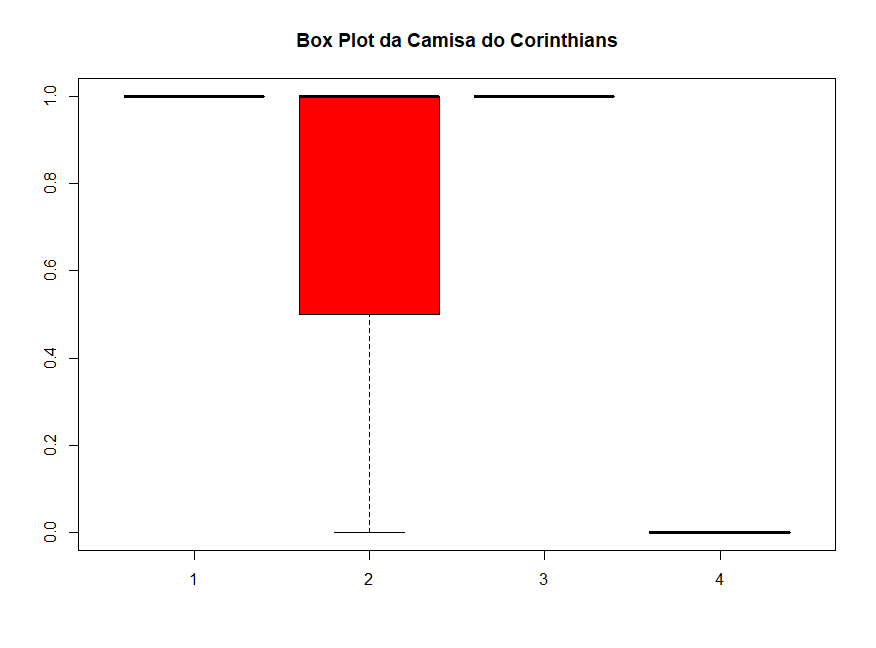
> boxplot(rede2$Calculadora ~ grupos2, col = "red", main = 'Box Plot da Calculadora')

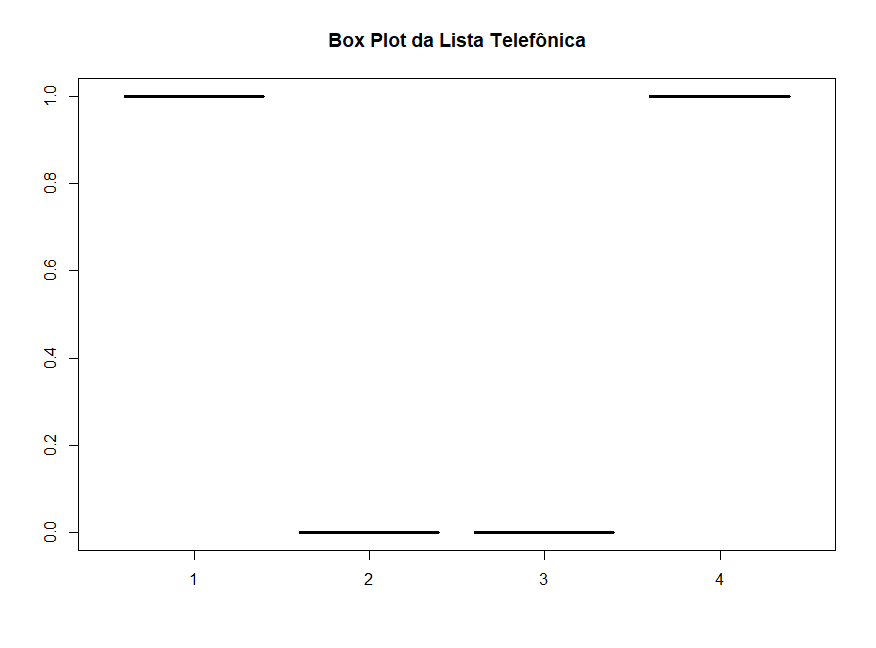
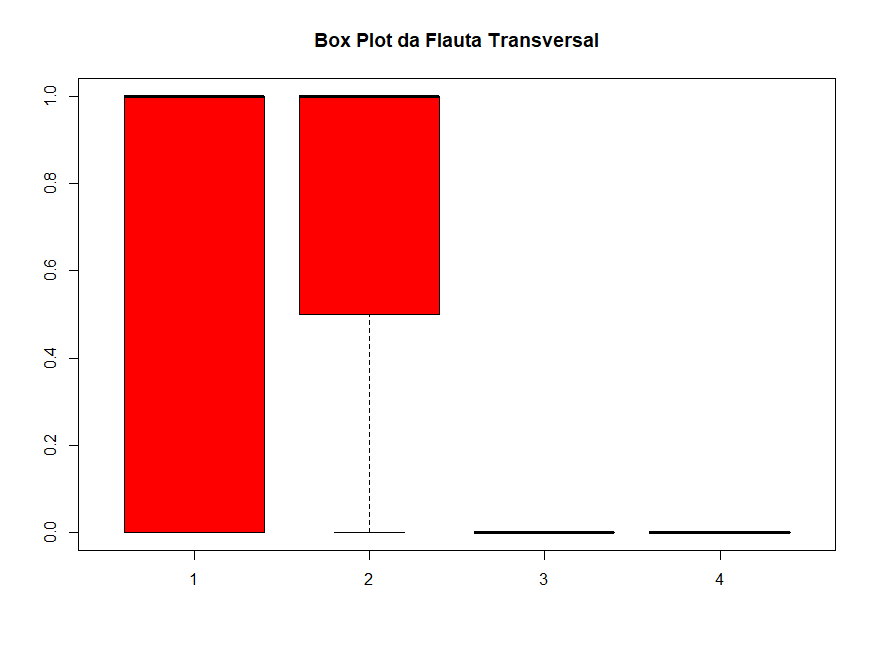
> boxplot(rede2$Detergente ~ grupos2, col = "red", main = 'Box Plot do Detergente')

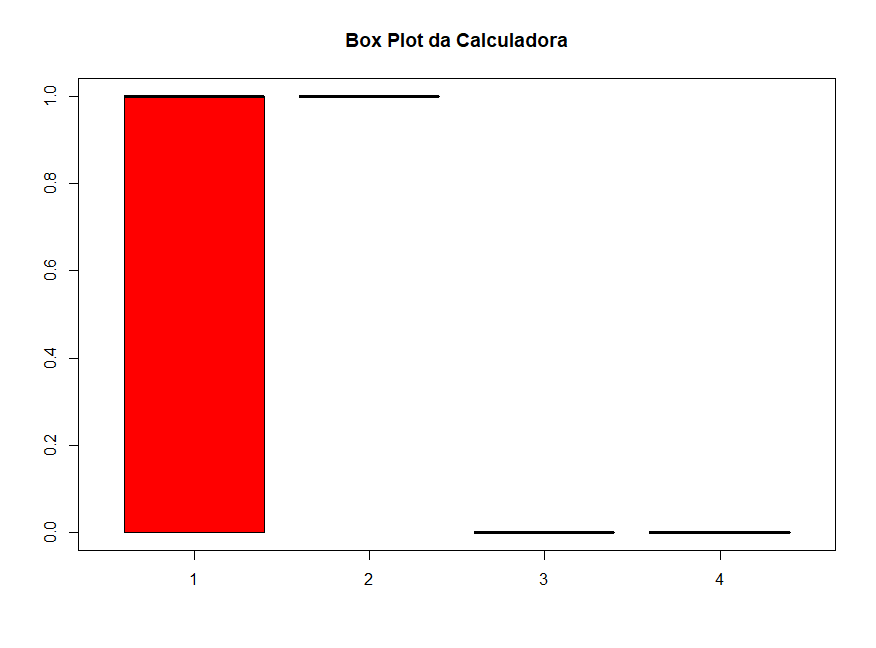
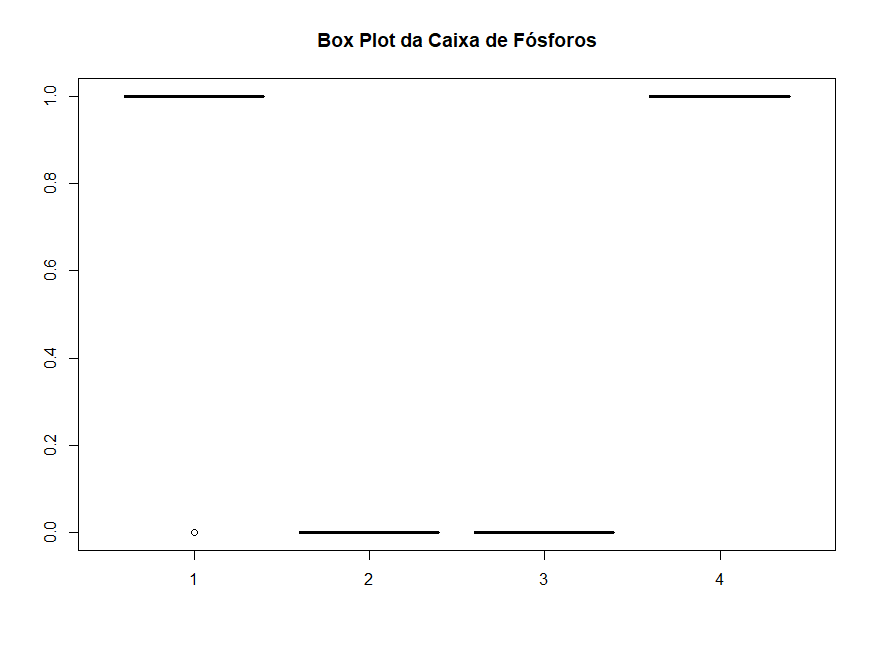
>

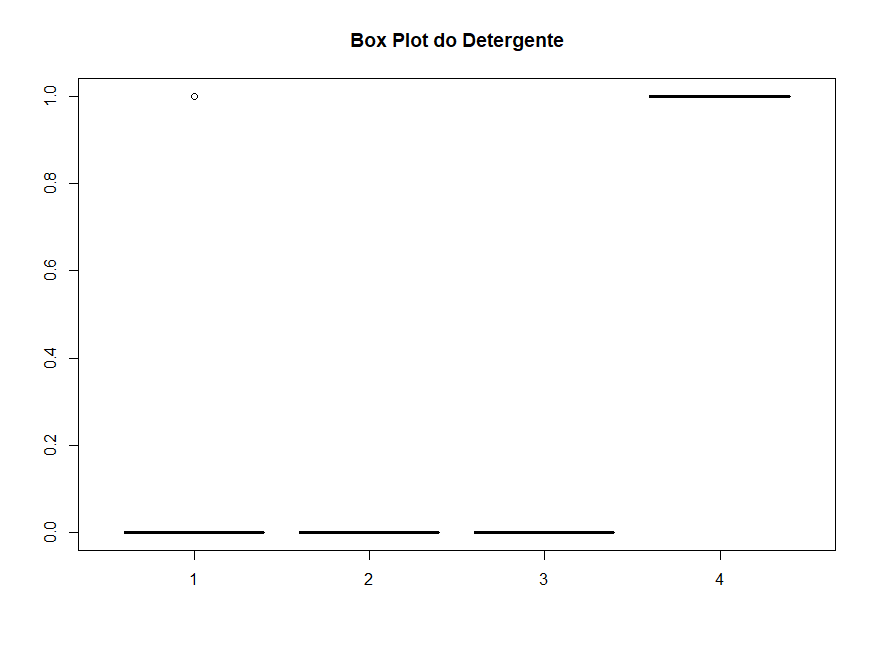












>

> # Desenhar os gráficos de rede

> # professor quer produto em branco e pessoa em qquer cor

> compras1 <- read.table("D:/0 Aula 28\_07\_18 Redes sociais/SNA e Text Mining Paulista T4/Trabalho 1/Rede Two Mode\_Trab1.csv",header=TRUE,sep = ";", dec=",")

> compras1

X iPhone iPad Livro.Harry.Potter jogo.MineCraft Camisa.do.Corinthians Bola.de.Futebol Flauta.Transversal

1 João 1 1 0 1 1 0 1

2 Maria 1 1 1 1 1 1 1

3 José 0 0 0 0 1 0 1

4 Paulo 1 0 0 0 1 0 0

5 Pedro 0 0 1 0 1 1 0

6 Luisa 0 0 1 0 1 1 0

7 Marcelo 1 0 0 0 0 1 1

8 Alfredo 0 1 1 0 1 1 1

9 Joaquim 1 0 1 0 1 1 1

10 Gabriela 0 1 1 1 1 1 0

11 Flávia 1 0 1 1 1 0 0

12 Catapulto 0 1 0 0 1 1 1

13 Rodrigo 0 0 0 0 0 0 0

14 Gabriel 0 1 1 1 1 1 1

15 Rodolfo 1 0 1 0 1 1 1

Lista.Telefônica Caixa.de.Fósforos Calculadora Detergente

1 1 1 1 0

2 1 1 1 1

3 0 0 1 0

4 0 0 1 0

5 1 1 1 0

6 1 1 1 0

7 0 0 1 0

8 0 0 1 0

9 1 1 1 0

10 1 1 0 0

11 0 0 0 0

12 1 1 0 0

13 1 1 0 1

14 1 1 0 0

15 1 0 0 0

>

> gcompras1<- compras1[, 2:12]

> rownames(gcompras1) <- compras1[,1]

> gcompras1

iPhone iPad Livro.Harry.Potter jogo.MineCraft Camisa.do.Corinthians Bola.de.Futebol Flauta.Transversal

João 1 1 0 1 1 0 1

Maria 1 1 1 1 1 1 1

José 0 0 0 0 1 0 1

Paulo 1 0 0 0 1 0 0

Pedro 0 0 1 0 1 1 0

Luisa 0 0 1 0 1 1 0

Marcelo 1 0 0 0 0 1 1

Alfredo 0 1 1 0 1 1 1

Joaquim 1 0 1 0 1 1 1

Gabriela 0 1 1 1 1 1 0

Flávia 1 0 1 1 1 0 0

Catapulto 0 1 0 0 1 1 1

Rodrigo 0 0 0 0 0 0 0

Gabriel 0 1 1 1 1 1 1

Rodolfo 1 0 1 0 1 1 1

Lista.Telefônica Caixa.de.Fósforos Calculadora Detergente

João 1 1 1 0

Maria 1 1 1 1

José 0 0 1 0

Paulo 0 0 1 0

Pedro 1 1 1 0

Luisa 1 1 1 0

Marcelo 0 0 1 0

Alfredo 0 0 1 0

Joaquim 1 1 1 0

Gabriela 1 1 0 0

Flávia 0 0 0 0

Catapulto 1 1 0 0

Rodrigo 1 1 0 1

Gabriel 1 1 0 0

Rodolfo 1 0 0 0

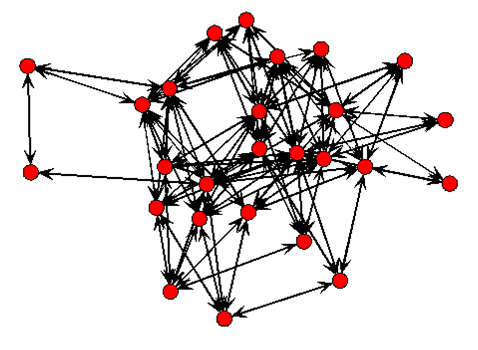
>

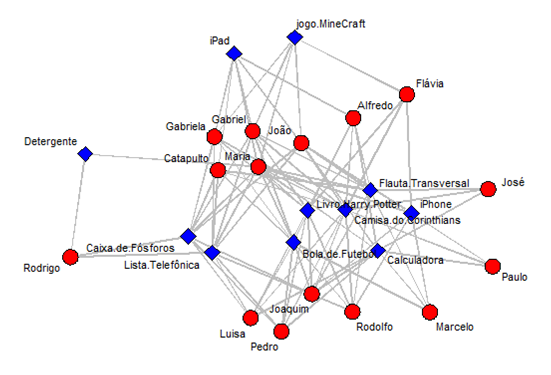
> # Gráfico rede Two Mode

> gplot(gcompras1)

> gplot(gcompras1, gmode="twomode", displaylabels = TRUE, edge.col="gray", label.cex = 0.7, usearrows=FALSE)

>





**###### DESAFIO AQUI É COLOCAR PRODUTO EM BRANCO E PESSOAS EM OUTRAS COR !!!!!**