



Universidad Autónoma del Estado de Hidalgo

Doctorado en Ciencias Económico Administrativas

Temas Selectos I: Estadística para las CEA

Complejidad Económica | Lab 44-A

CUADERNO R.MARKDOWN

Académico:

Dra. Carla Carolina Pérez Hernández

Alumna:

Ana Grisel Sanjuan Merida

263501

Fecha de entrega: 10 de marzo de 2023.



LAB 44-A (MD)

AnaGSanjuanM

2023-03-10

LABORATORIO 44-A

ACADÉMICO: CARLA CAROLINA PÉREZ HERNÁNDEZ

ALUMNA: ANA GRISEL SANJUAN MERIDA

A tale of two heatmap functions

<https://divingintogeneticsandgenomics.rbind.io/post/a-tale-of-two-heatmap-functions/>

Se instala la paquetería

```
install.packages("gplots")
```

```
install.packages("heatmaps")
```

```
install.packages("pheatmap")
```

Se cargan las librerías

```
library(stats)
library(gplots)
```

```
##
```

```
## Attaching package: 'gplots'
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
## lowess
```

Medida de la expresión génica de 4 genes (h1, h2, l1 y l2) en 8 puntos de tiempo

```
h1 <- c(10,20,10,20,10,20,10,20)
```

```
h2 <- c(20,10,20,10,20,10,20,10)
```

```
l1 <- c(1,3,1,3,1,3,1,3)
```

```
l2 <- c(3,1,3,1,3,1,3,1)
```

Generamos matriz denominada mat para enlazar los genes

```
mat <- rbind(h1,h2,l1,l2)
```

Se corren los siguientes comandos para generar el plot

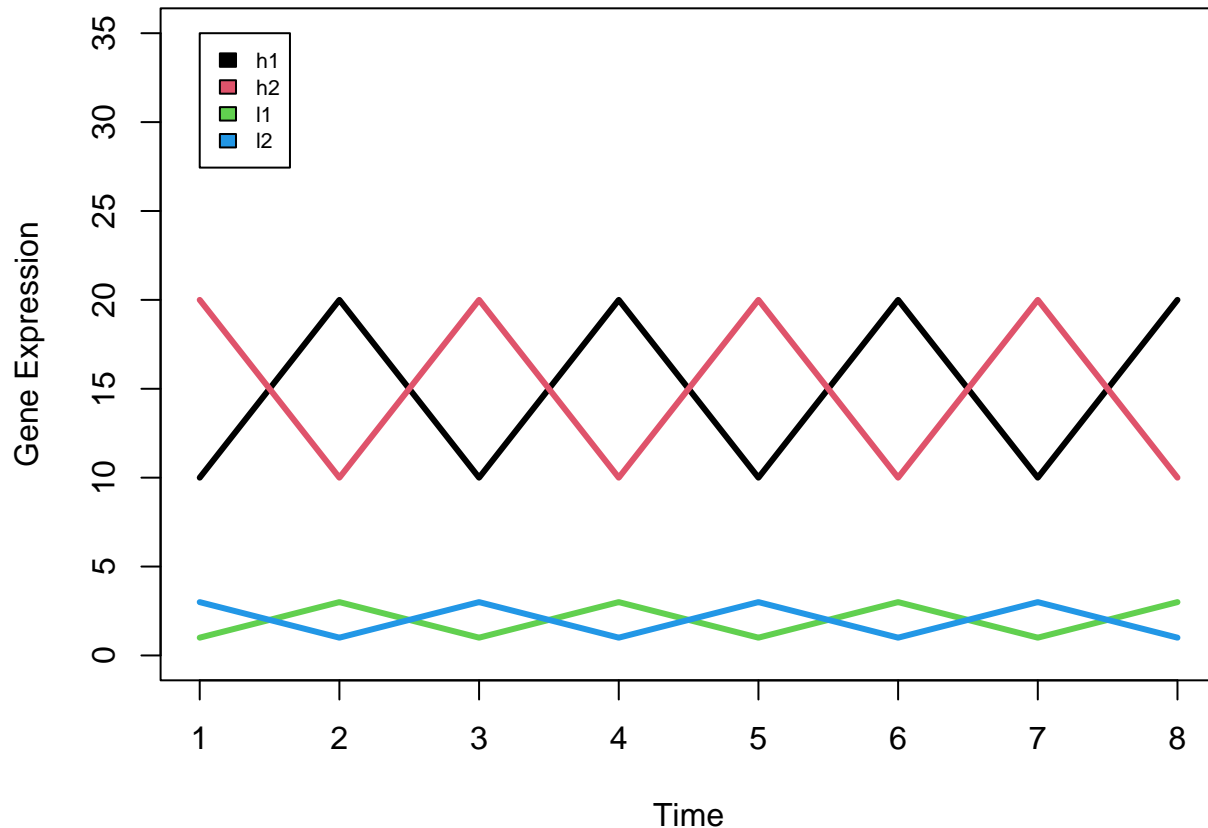
```

par(mfrow =c(1,1), mar=c(4,4,1,1))
plot(1:8,rep(0,8), ylim=c(0,35), pch="", xlab="Time", ylab="Gene Expression")

for (i in 1:nrow(mat)) {
  lines(1:8,mat[i,], lwd=3, col=i)
}

legend(1,35,rownames(mat), 1:4, cex=0.7)

```



Para calcular la distancia

```
dist(mat)
```

```

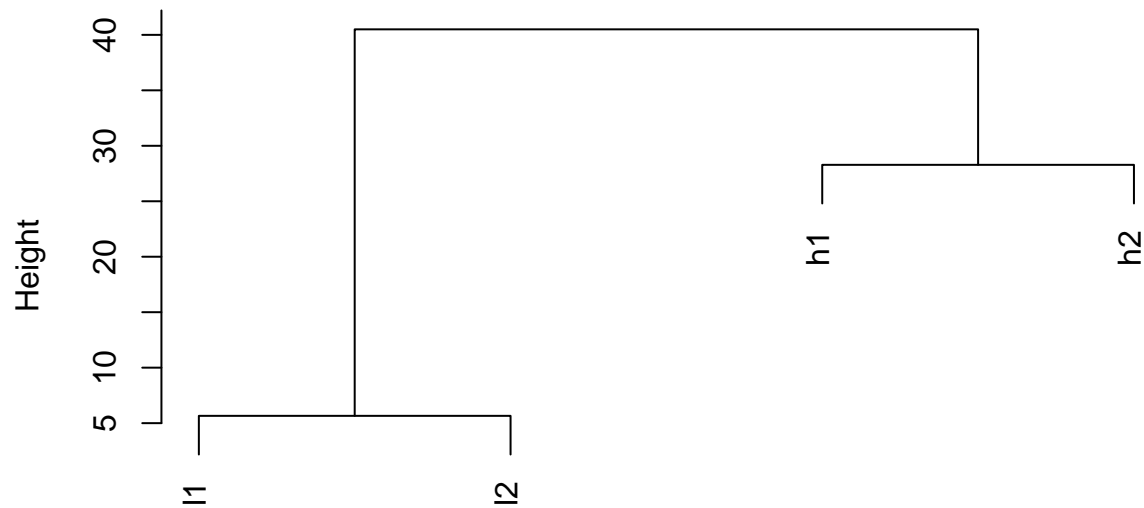
##          h1          h2          l1
## h2 28.284271
## l1 38.470768 40.496913
## l2 40.496913 38.470768  5.656854

```

Se usa el método predeterminado para el enlace: completo

```
plot(hclust(dist(mat)))
```

Cluster Dendrogram

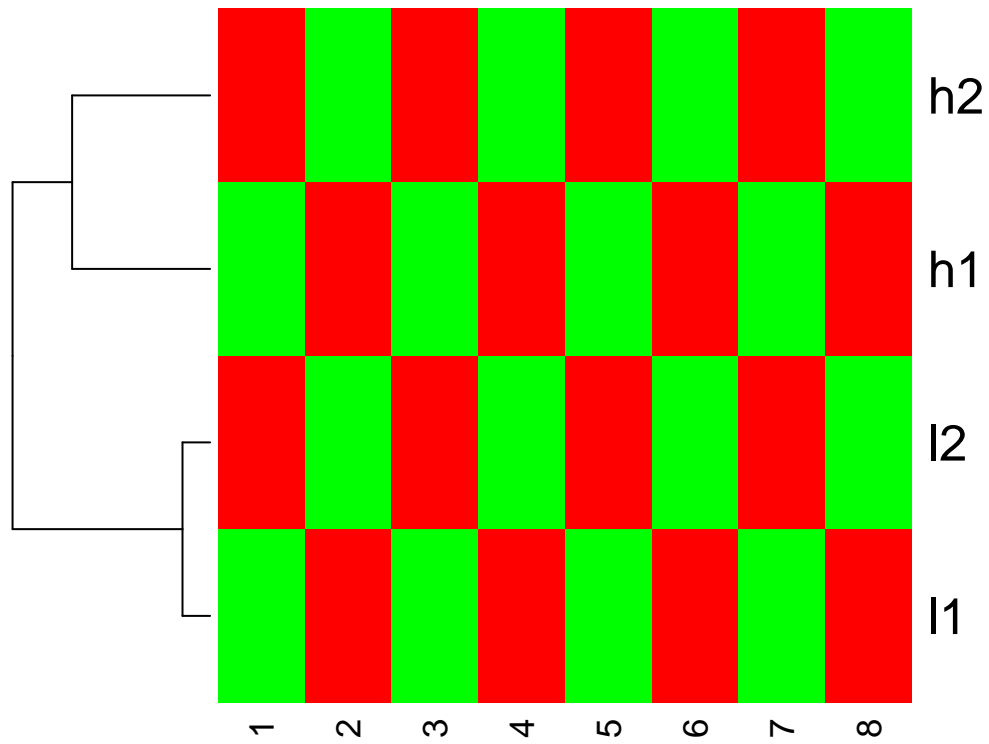


```
dist(mat)  
hclust (*, "complete")
```

Heatmap predeterminado, para obtener explícitamente parámetros

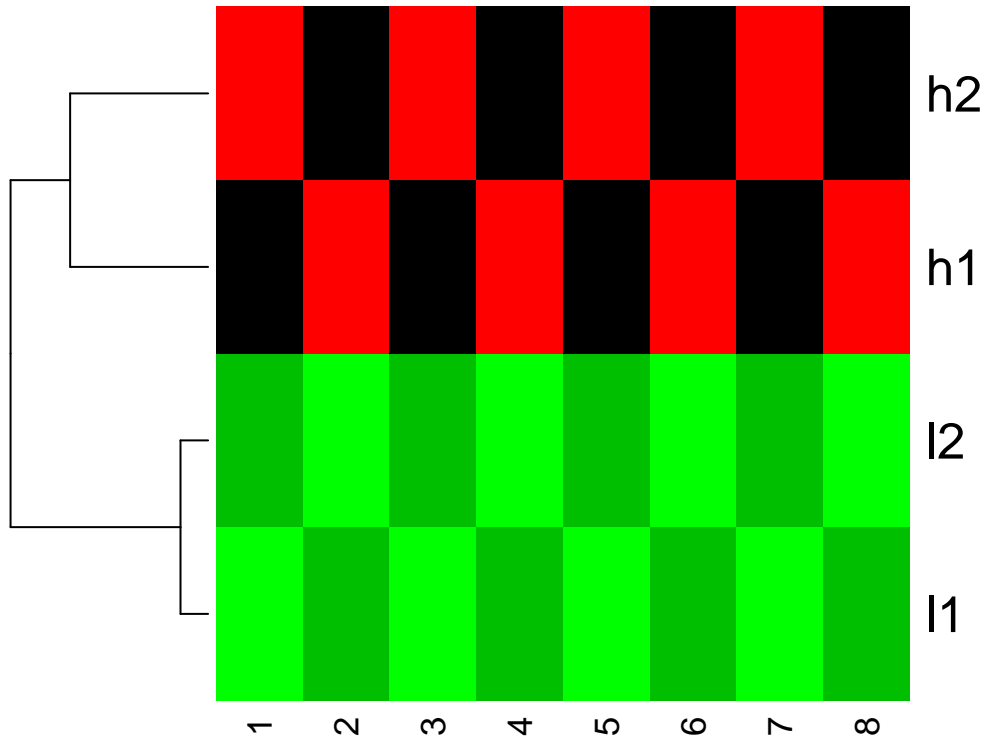
Escala en los renglones con escala de color verde y roja

```
heatmap(mat, Colv=NA, col=greenred(10), scale = "row")
```



Heatmap con la escala desactivada

```
heatmap(mat, Colv = NA, col=greenred(10), scale = "none")
```



Escalando los genes antes de introducirlos al heatmap

```
mat.scaled<- t(scale(t(mat), center=TRUE, scale = TRUE))
mat.scaled
```

```
##          [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
## h1 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143
## h2  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143
## l1 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143
## l2  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143
##          [,8]
## h1  0.9354143
## h2 -0.9354143
## l1  0.9354143
## l2 -0.9354143
## attr("scaled:center")
## h1 h2 l1 l2
## 15 15  2  2
## attr("scaled:scale")
##      h1      h2      l1      l2
## 5.345225 5.345225 1.069045 1.069045
```

Cambio e la distancia entre genes

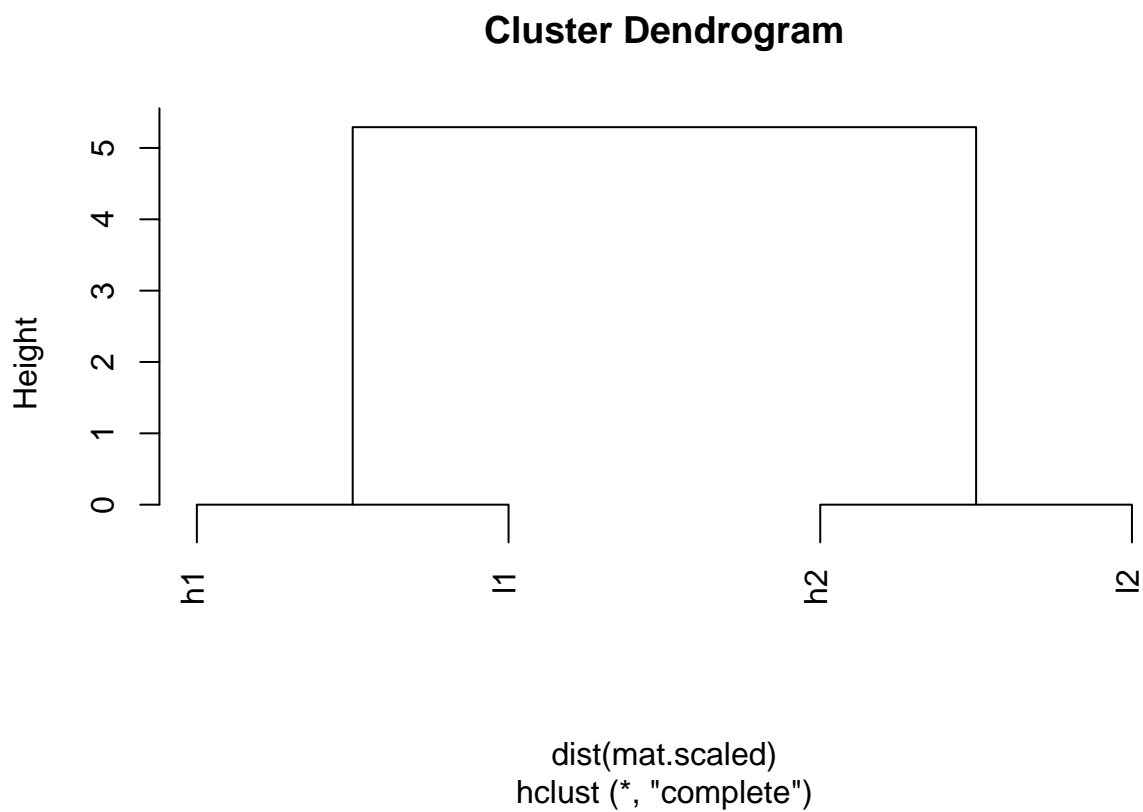
```
dist(mat.scaled)
```

```
##           h1           h2           l1
## h2 5.291503
## l1 0.000000 5.291503
## l2 5.291503 0.000000 5.291503
```

Obtener el plot

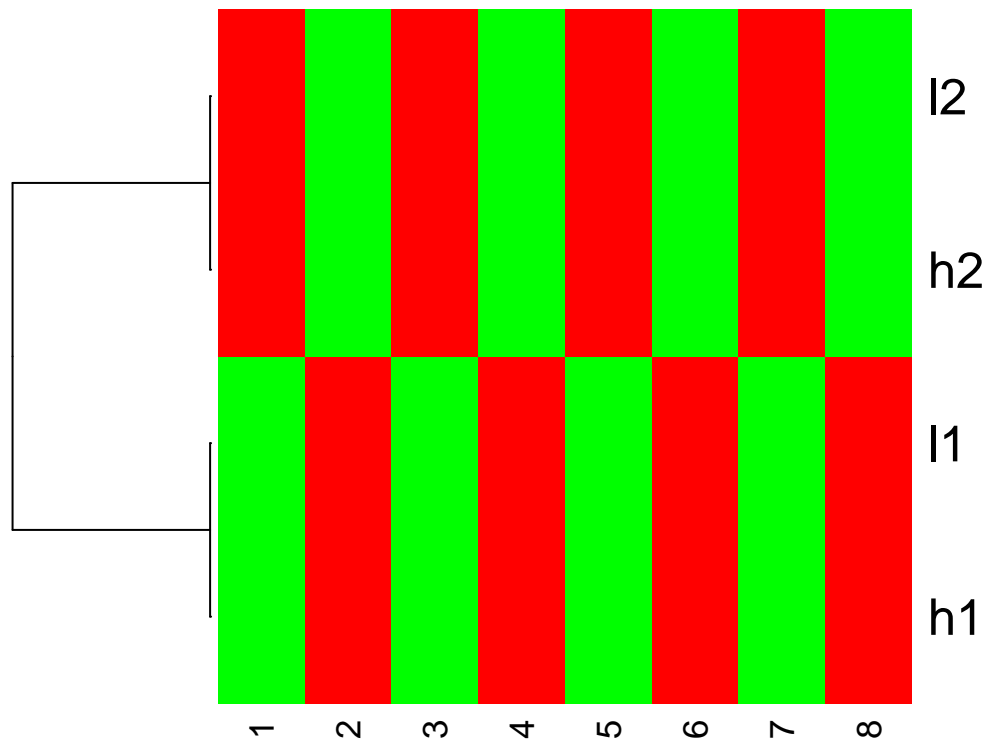
Ahora h1 y l1 están agrupados juntos; l2 y h2 están agrupados juntos

```
plot(hclust(dist(mat.scaled)))
```



Generando el heatmap

```
heatmap(mat.scaled, Colv = NA, col=greenred(10), scale = "none")
```



Si aún no se escalan los datos pero se desea que l1 y h1 se agrupen juntos, al igual que l2 y h2, se puede usar la medida de distancia diferentes

Correlación entre genes para asignar valores de 1 y -1

```
cor(t(mat))
```

```
##      h1 h2 l1 l2
## h1   1 -1  1 -1
## h2  -1  1 -1  1
## l1   1 -1  1 -1
## l2  -1  1 -1  1
```

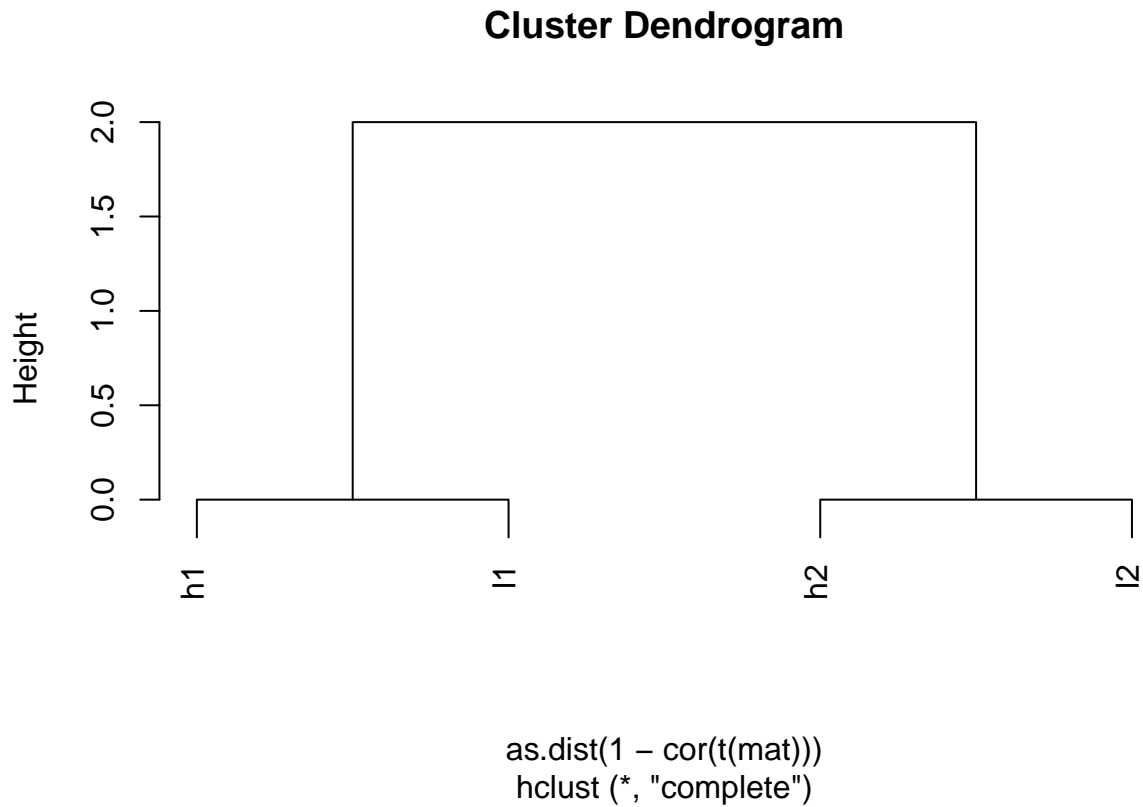
Correlación para definir la distancia

```
1- cor(t(mat))
```

```
##      h1 h2 l1 l2
## h1   0  2  0  2
## h2   2  0  2  0
## l1   0  2  0  2
## l2   2  0  2  0
```

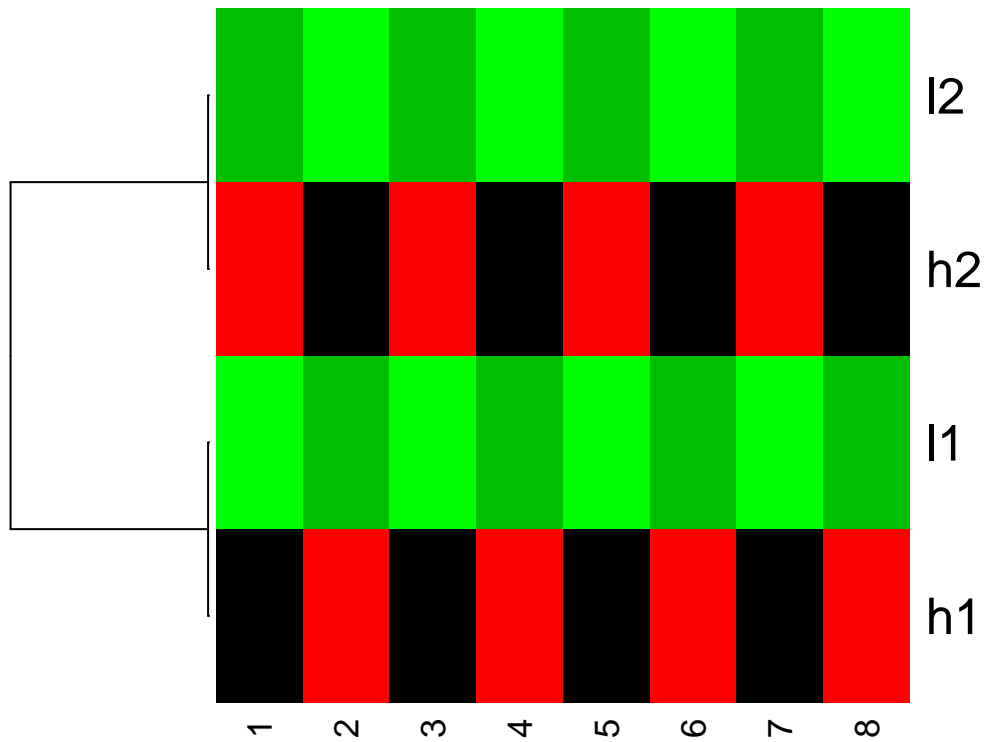

Generamos el dendrograma

```
hc <- hclust(as.dist(1-cor(t(mat))))  
plot(hc)
```



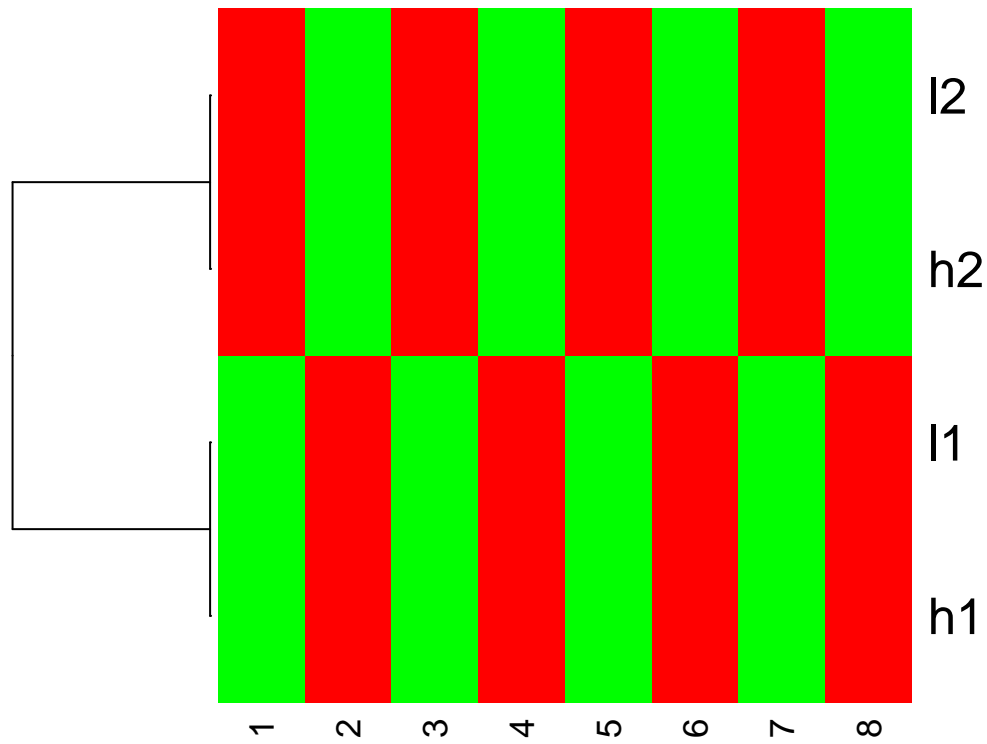
Heatmap sin escala

```
heatmap(mat, Colv = NA, Rowv=as.dendrogram(hc), col=greenred(10), scale = "none")
```



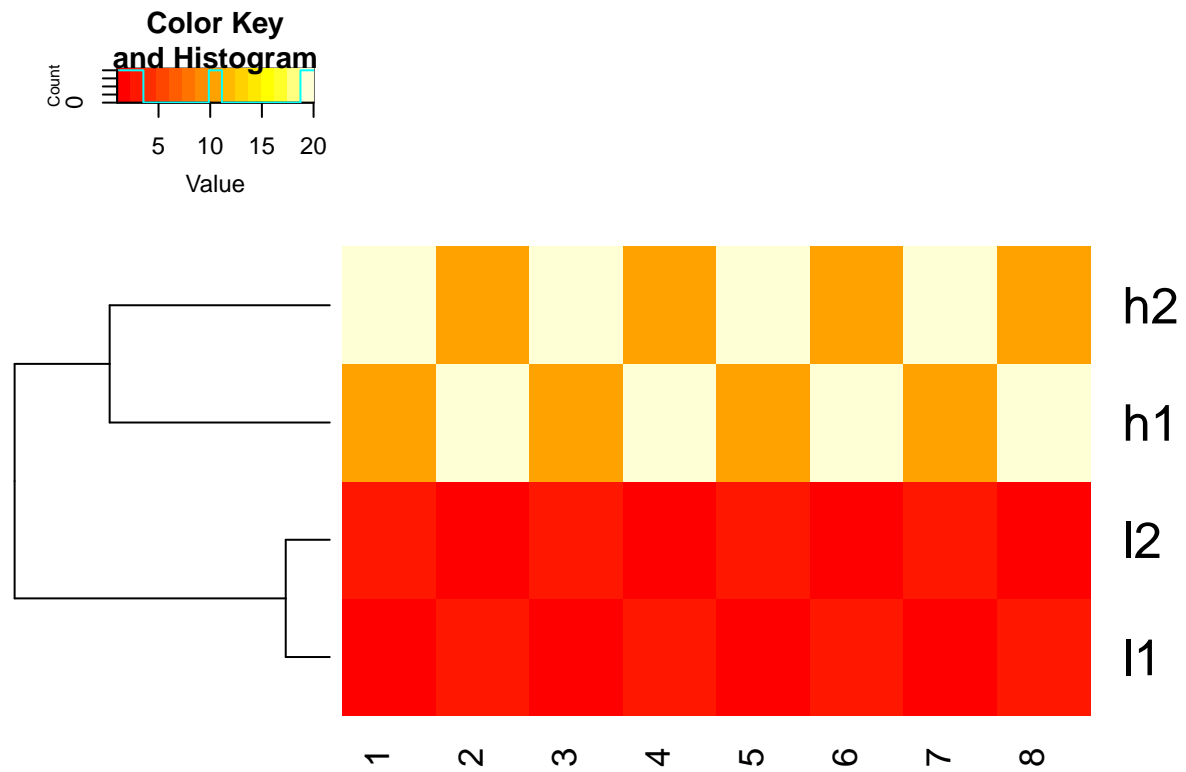
Heatmap con la escala en los renglones

```
heatmap(mat, Colv = NA, Rowv=as.dendrogram(hc), col=greenred(10), scale = "row")
```



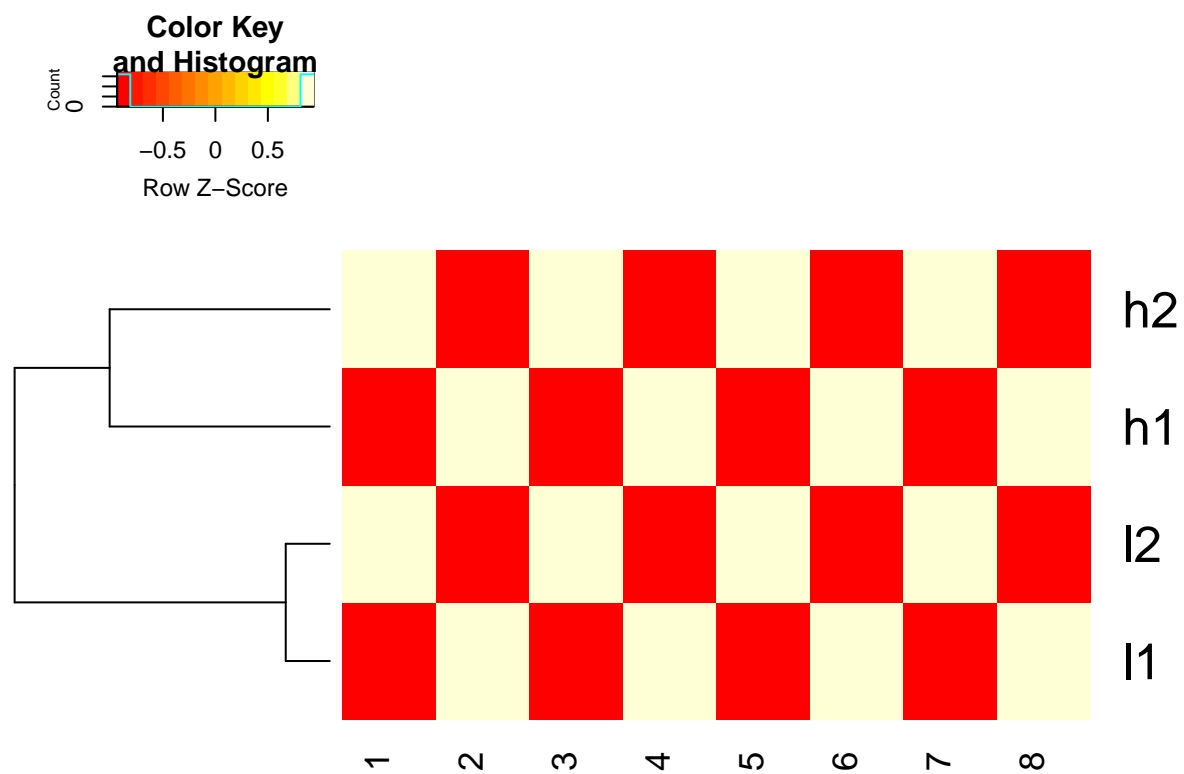
Valores predeterminados de heatmap.2 con ninguna escala

```
heatmap.2(mat, trace = "none", Colv= NA, dendrogram = "row", scale = "none")
```



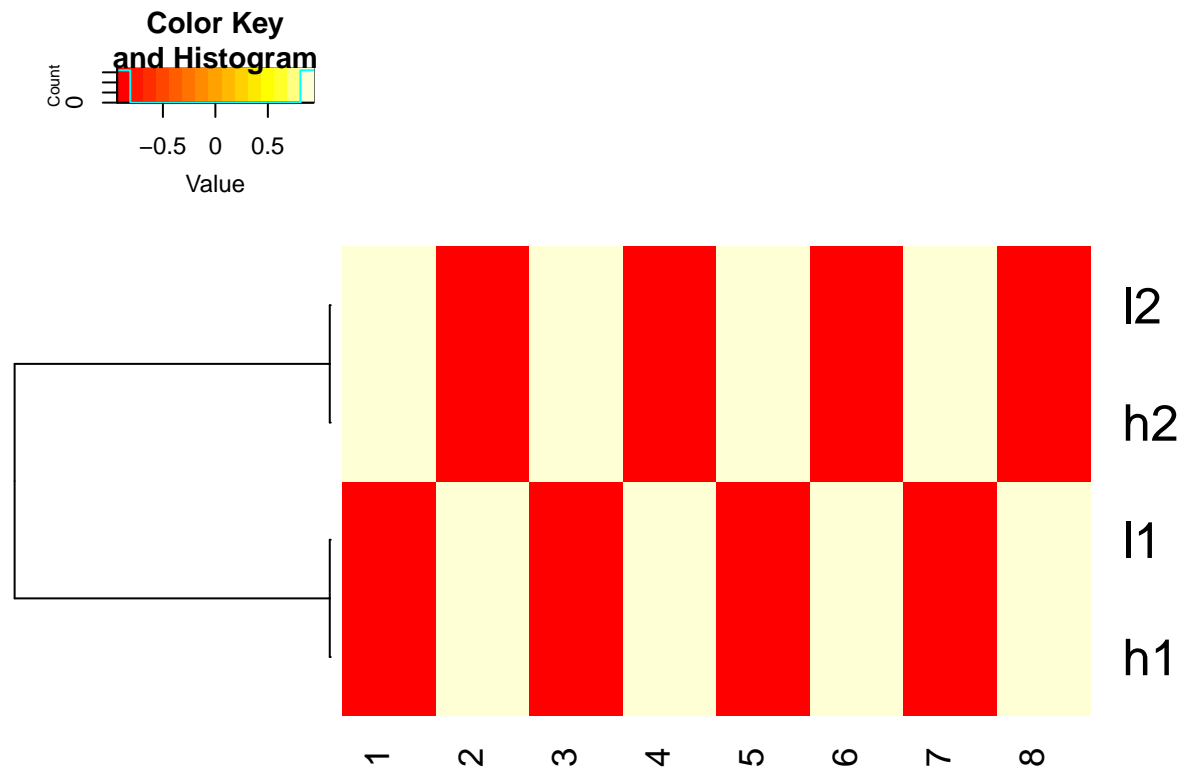
Funciones de heatmap en R: primero se agrupa y luego usa el argumento de escala (si está configurado) para representar los datos

```
heatmap.2(mat, trace = "none", Colv= NA, dendrogram = "row", scale = "row")
```



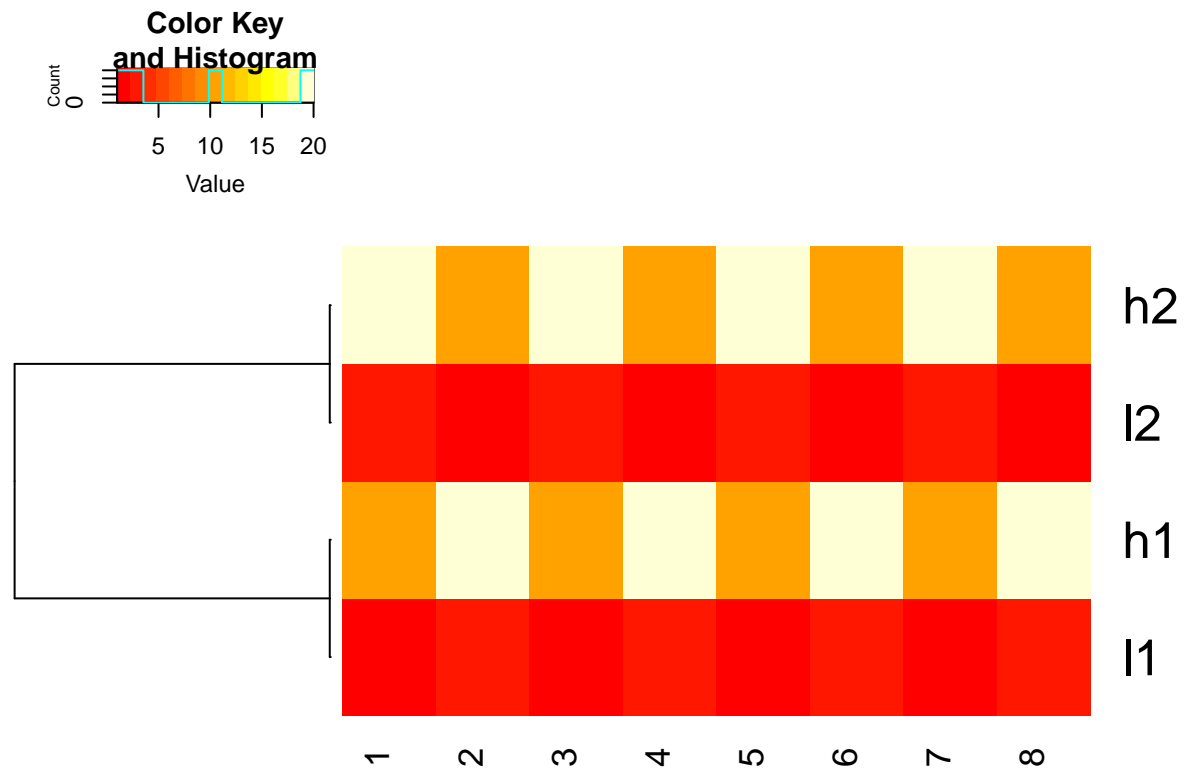
Datos escalados explícitamente primero y uso de la distancia euclidiana

```
heatmap.2(t(scale(t(mat), center=TRUE, scale=TRUE)), trace = "none", Colv= NA, dendrogram = "row", scal
```



Usando $1 - \text{cor}(x)$ como distancia y no escalar

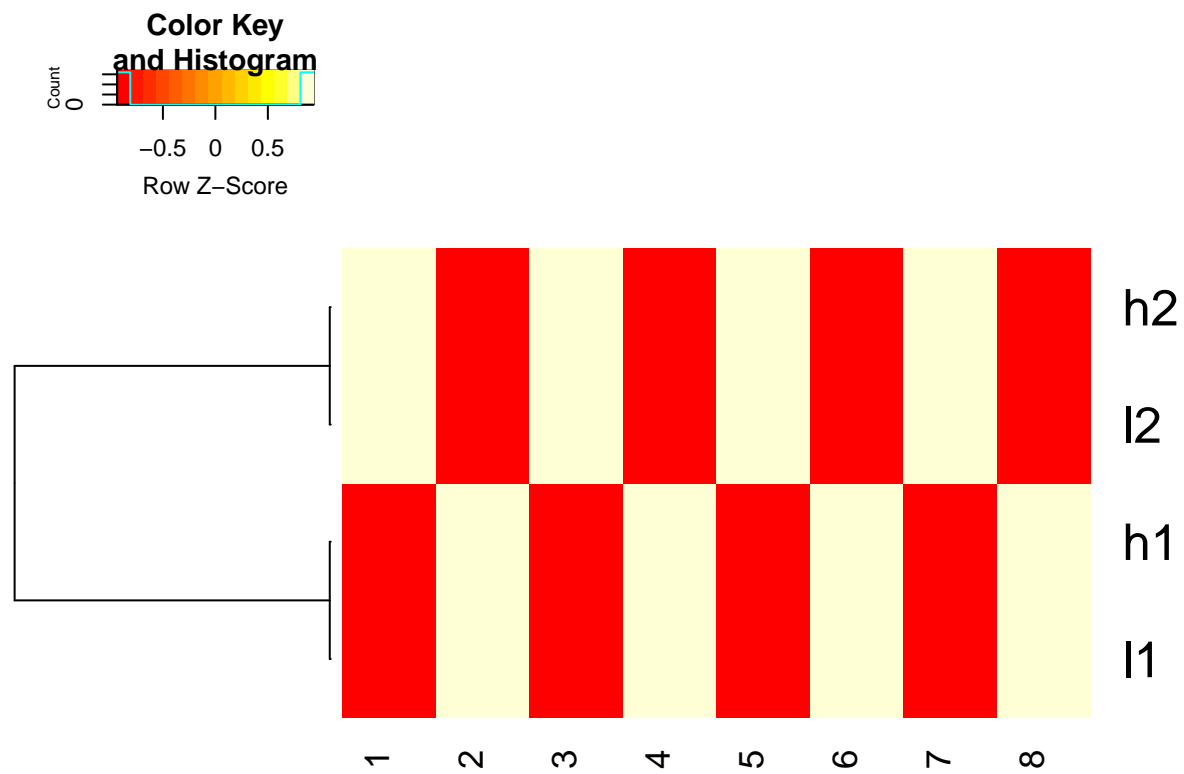
```
heatmap.2(mat, trace = "none",
  Colv= NA, dendrogram = "row",
  scale = "none",
  hclust=function(x) hclust(x, method='complete'), distfun=function(x) as.dist(1-cor(t(x))))
```



Usando $1 - \text{cor}(x)$ como distancia y no escalar

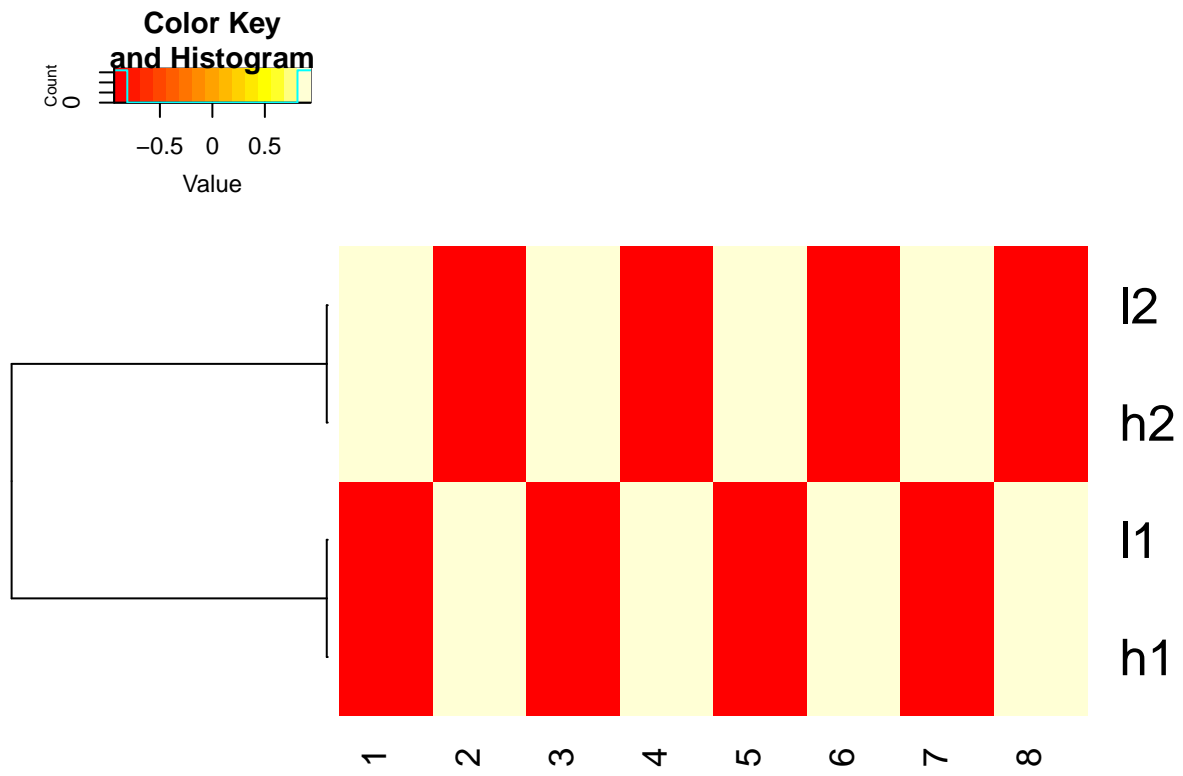
Usar la escala en la función `heatmap.2` para representar los colores

```
heatmap.2(mat, trace = "none",
          Colv= NA, dendrogram = "row",
          scale = "row",
          hclust=function(x) hclust(x, method='complete'), distfun=function(x) as.dist(1-cor(t(x))))
```



Escala y uso de $1 - \text{cor}(x)$ como distancia

```
heatmap.2(t(scale(t(mat), center=TRUE, scale=TRUE)), trace = "none",
          Colv= NA, dendrogram = "row",
          hclust=function(x) hclust(x, method='complete'), distfun=function(x) as.dist(1-cor(t(x))))
```

FIN DE LABORATORIO 44-A