

Laboratorio44A-MD

Leislle R. Manjarrez O.

2023-03-11

Hecho con gusto por Leislle R. Manjarrez O.

Laboratorio 44A- Heatmaps con gplots

Instalamos la paqueteria `install.packages("gplots")` `install.packages("ggplot2")`

Se llaman librerias

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

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

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

```
library(ggplot2)
```

Creamos un objeto (mat) suponiendo que contiene los valores de 4 genes (l1, l2, h1, h2) los cuales son medidos en 8 puntos de tiempo

Graficamos para observar que h1 y h2 son altos e l1 y l2 son bajos

```
h1 <- c(10,20,10,20,10,20,10,20)
h2 <- c(20,10,20,10,20,10,20,10)

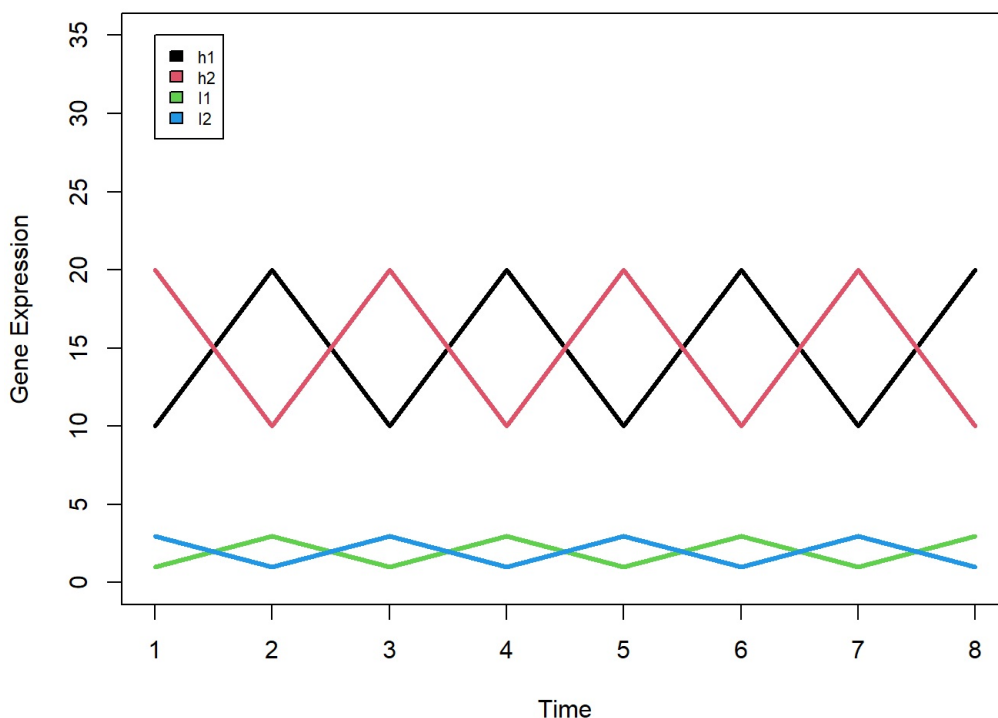
l1 <- c(1,3,1,3,1,3,1,3)
l2 <- c(3,1,3,1,3,1,3,1)

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

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)
```



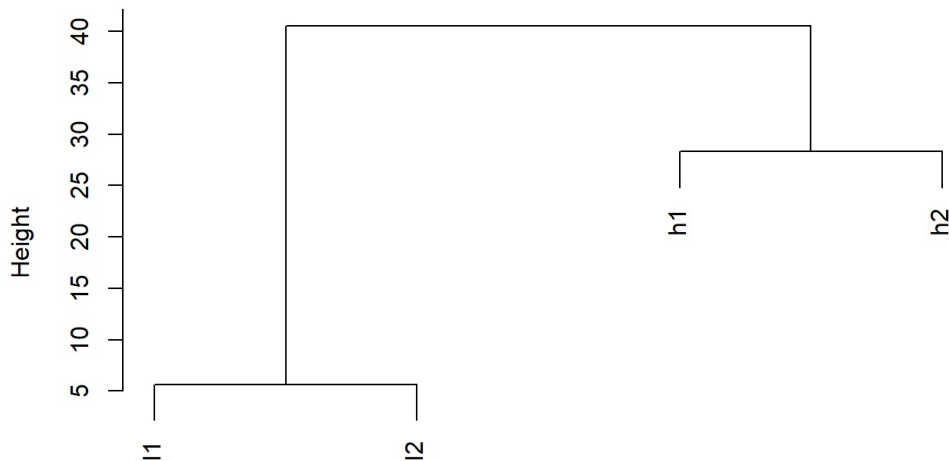
Calculamos la distancia entre estos objetos/genes y graficamos en un dendrograma

```
dist(mat)
```

```
##           h1           h2           I1
## h2 28.284271
## I1 38.470768 40.496913
## I2 40.496913 38.470768  5.656854
```

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

Cluster Dendrogram



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

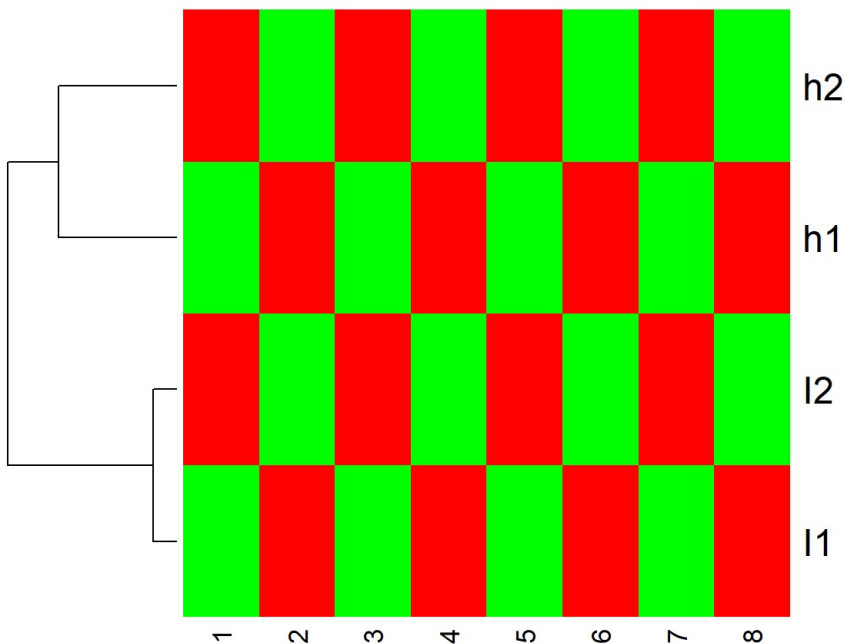
Observamos que la medida de distancia predeterminada es la euclidiana por lo que h1 y h2 estan mas cerca, asi como I1 e I2 Busquemos la ayuda sobre heatmap

```
?heatmap
```

```
## starting httpd help server ... done
```

El valor predeterminado es fila si el simbolo es falso y ninguno en caso contrario. Vamos a marcar parametros explicitos

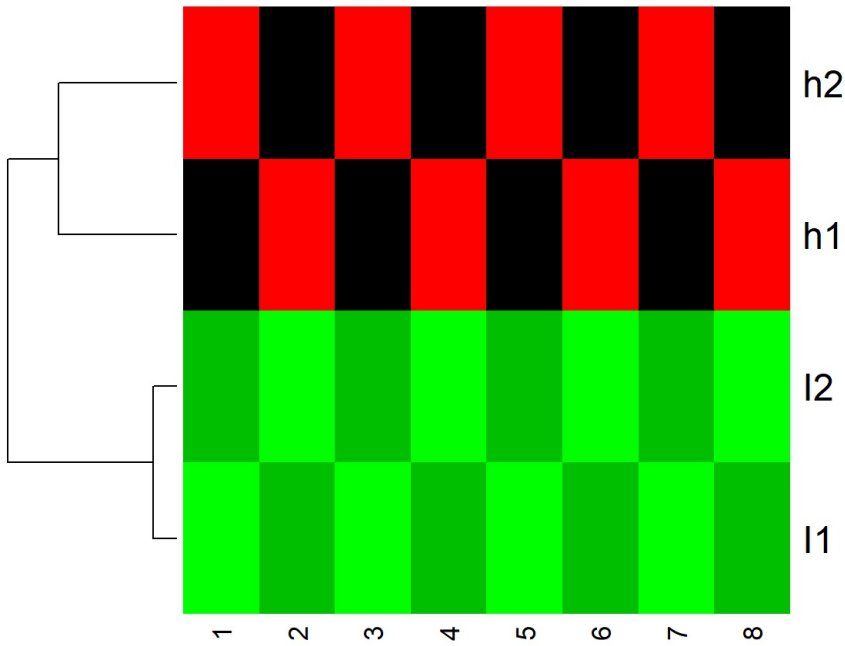
```
heatmap(mat, Colv = NA, col=greenred(10), scale = "row")
```



Vemos que h1 y h2 se agrupan igual que I1 e I2. Tambien observamos que h1,h2 tienen valores grandes pero tienen el mismo color rojo que I1 e I2, lo que confirma que el mapa de calor se agrupa primero y luego escala la fila para representar el color.

Ahora quitaremos la bascula dentro de heatmap

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



No cambia la forma de agrupar pero si el color. Vamos a escalar los genes antes de introducirlos al mapa de calor

```
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
## I1 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143
## I2  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143 -0.9354143  0.9354143
##      [,8]
## h1  0.9354143
## h2 -0.9354143
## I1  0.9354143
## I2 -0.9354143
## attr("scaled:center")
## h1 h2 I1 I2
## 15 15  2  2
## attr("scaled:scale")
##      h1      h2      I1      I2
## 5.345225 5.345225 1.069045 1.069045
```

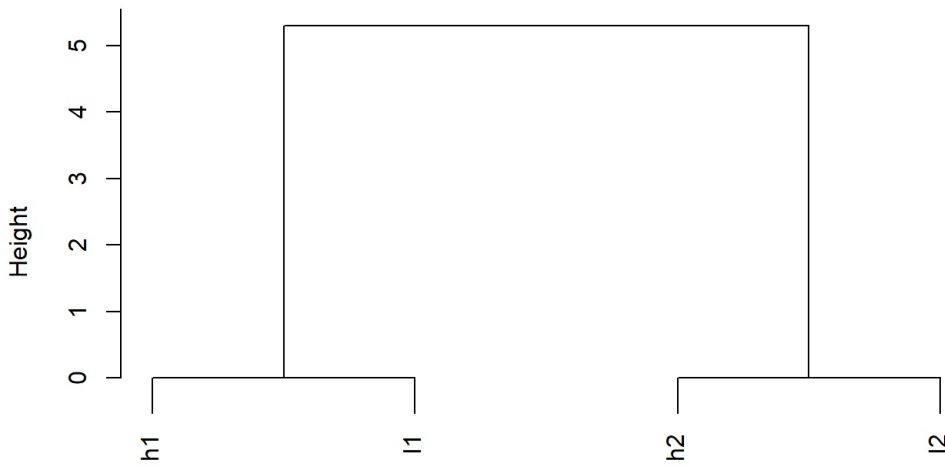
Volvemos a calcular la distancia entre genes visualizando los cambios mediante el dendrograma

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

```
##      h1      h2      I1
## h2 5.291503
## I1 0.000000 5.291503
## I2 5.291503 0.000000 5.291503
```

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

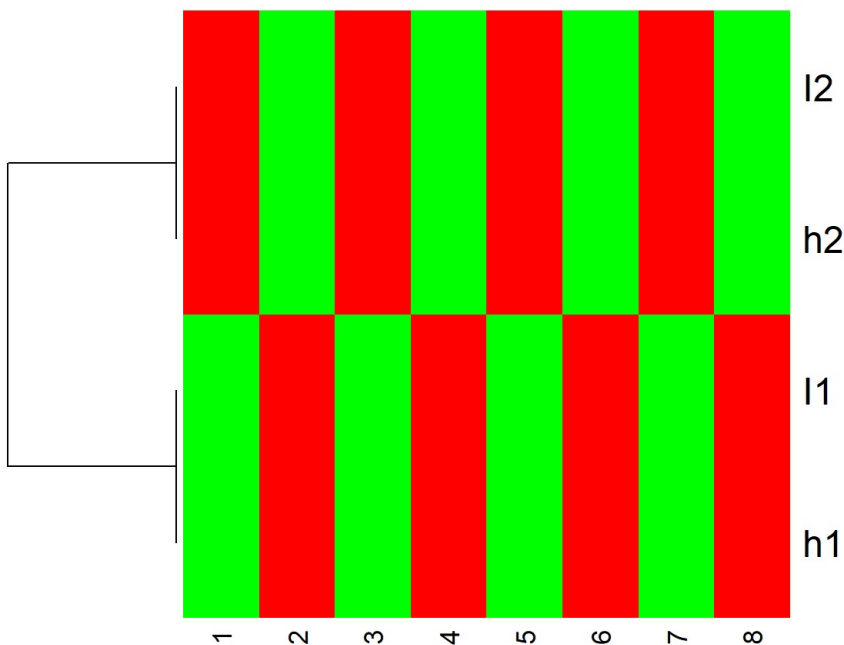
Cluster Dendrogram



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

Vemos que ahora h1 w l1 están agrupados igual que h2 e l2. Graficamos este resultados en un mapa de calor

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



Es importante pensar si es necesario escalar o no los datos antes porque como vemos esto afecta el aspecto de los mapas de calor. Suponiendo que no se escalaron los datos de antemano y queremos que l1 y h1 se agrupen, así como l2 y h2, es posible utilizar una medida de distancia diferente. Hagamos un ejemplo y grafiquemos para observar los cambios

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

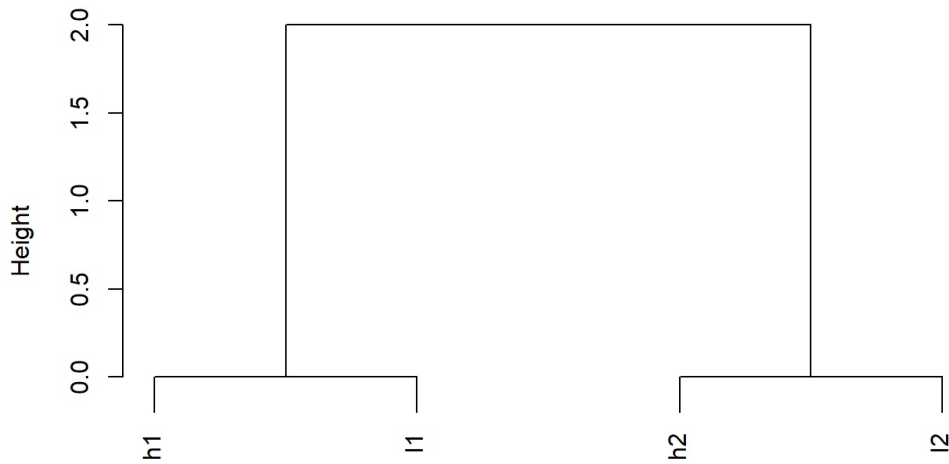
```
##      h1 h2 l1 l2
## h1   1 -1  1 -1
## h2  -1  1 -1  1
## l1   1 -1  1 -1
## l2  -1  1 -1  1
```

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

```
##      h1 h2 I1 I2
## h1   0  2  0  2
## h2   2  0  2  0
## I1   0  2  0  2
## I2   2  0  2  0
```

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

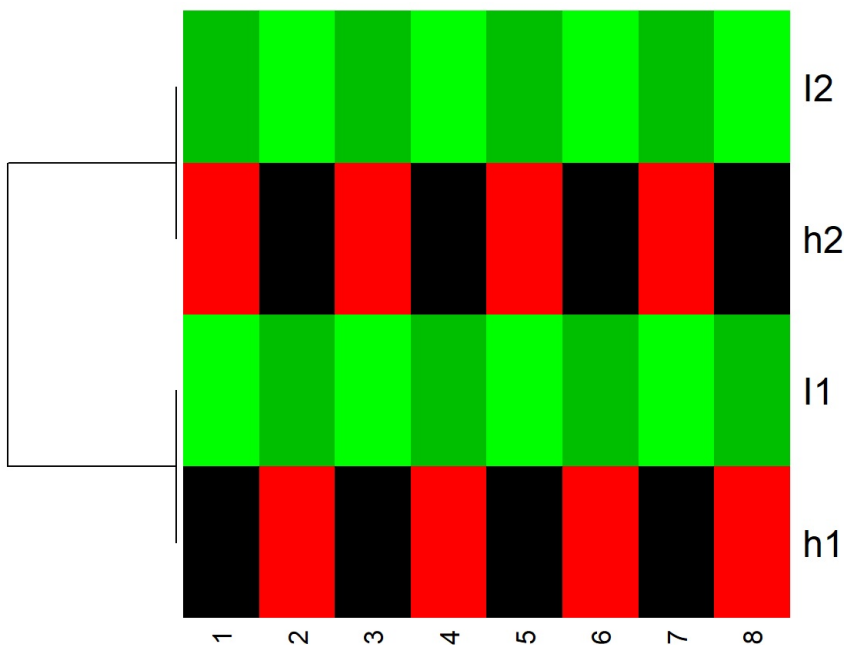
Cluster Dendrogram



```
as.dist(1 - cor(t(mat)))
hclust (*, "complete")
```

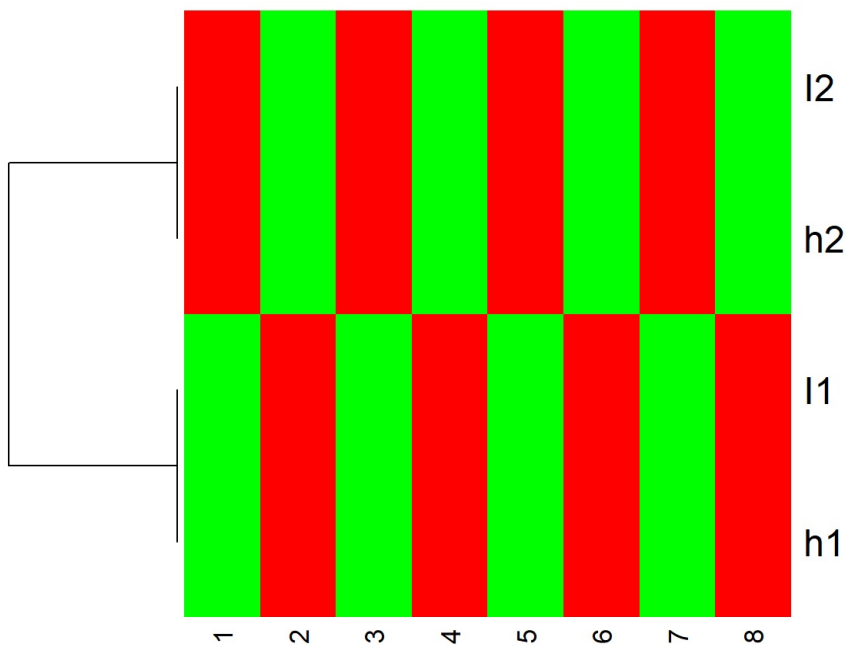
Ahora graficamos el mapa de calor configurando la escala en ninguno

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



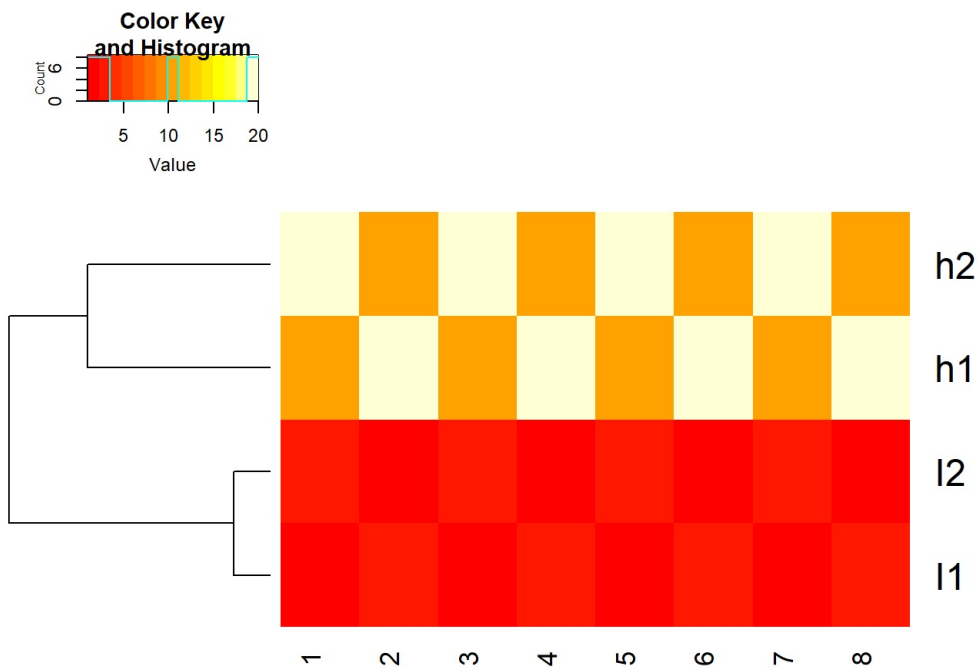
Vemos como siguen agrupados I2 y h2, así como I1 y h1 pero por el rango de valores es diferente ya que I1 e I2 son verdes (valores pequeños) y h1 y h2 son rojos (valores grandes). Si en la escala (scale) consideramos los renglones (row) entonces adecuaremos el gráfico

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



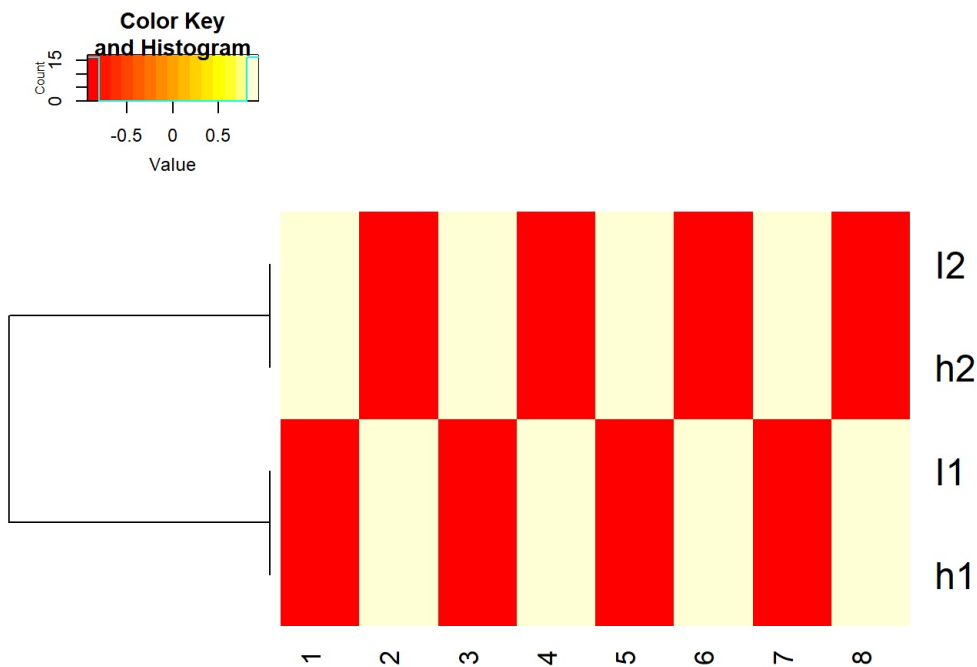
Es importante poner atención a la escala de color pues en heatmap.2 el valor predeterminado es ninguno

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



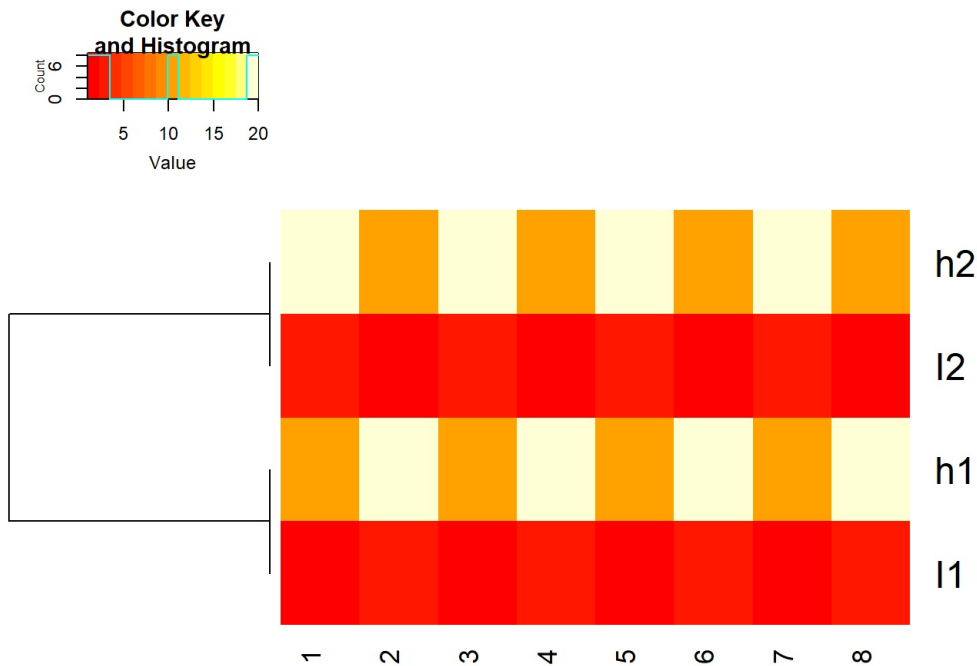
Es un hecho que en R primero se agrupa y luego usa el argumento de escala (si está configurado) para representar datos. Nuevamente vamos a escalar primero los datos explícitamente con la distancia euclidiana

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



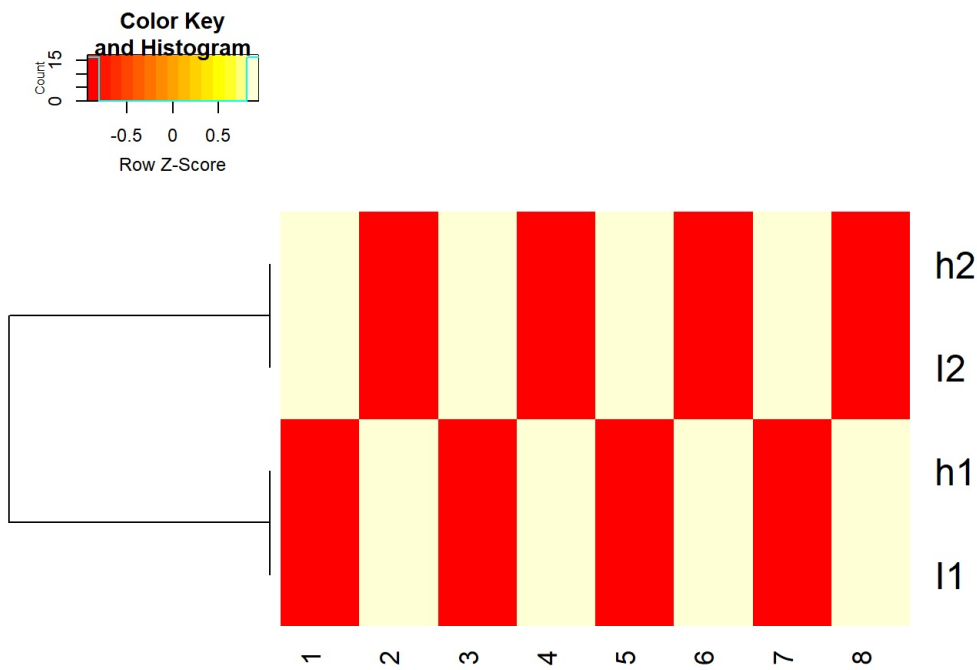
Usaremos $1-\text{cor}(x)$ como distancia y no escalaremos de antemano

```
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))))
```



Nuevamente usaremos $1-\text{cor}(x)$ como distancia y no escalaremos de antemano, pero usaremos la escala de la funcion heatmap.2 para representar 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))))
```



Por ultimo escalamos de antemano y usamos 1-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))))
```

