Heatmaps

SYSKA

What are Heatmaps and Why Should You Care?

- Heatmaps are a useful way of visualising data tables by replacing numbers with colours.
- Used in conjunction with clustering methods in order to find patterns in the data.
- They are also very nice to look at.

What You Will Need

- R
- One of the many Heatmap packages available.

Heatmap.3

- https://gist.github.com/nachocab/3853004
- Heatmap.3 allows multiple axis labelling/annotations
- Heatmap.3 does not have its own R package but is merely raw code
- To use it easily, save the code into a new R script then anytime you want to use it, source its directory location. E.g. source ("/home/username/Desktop/RScripts/heatmap3code.R")

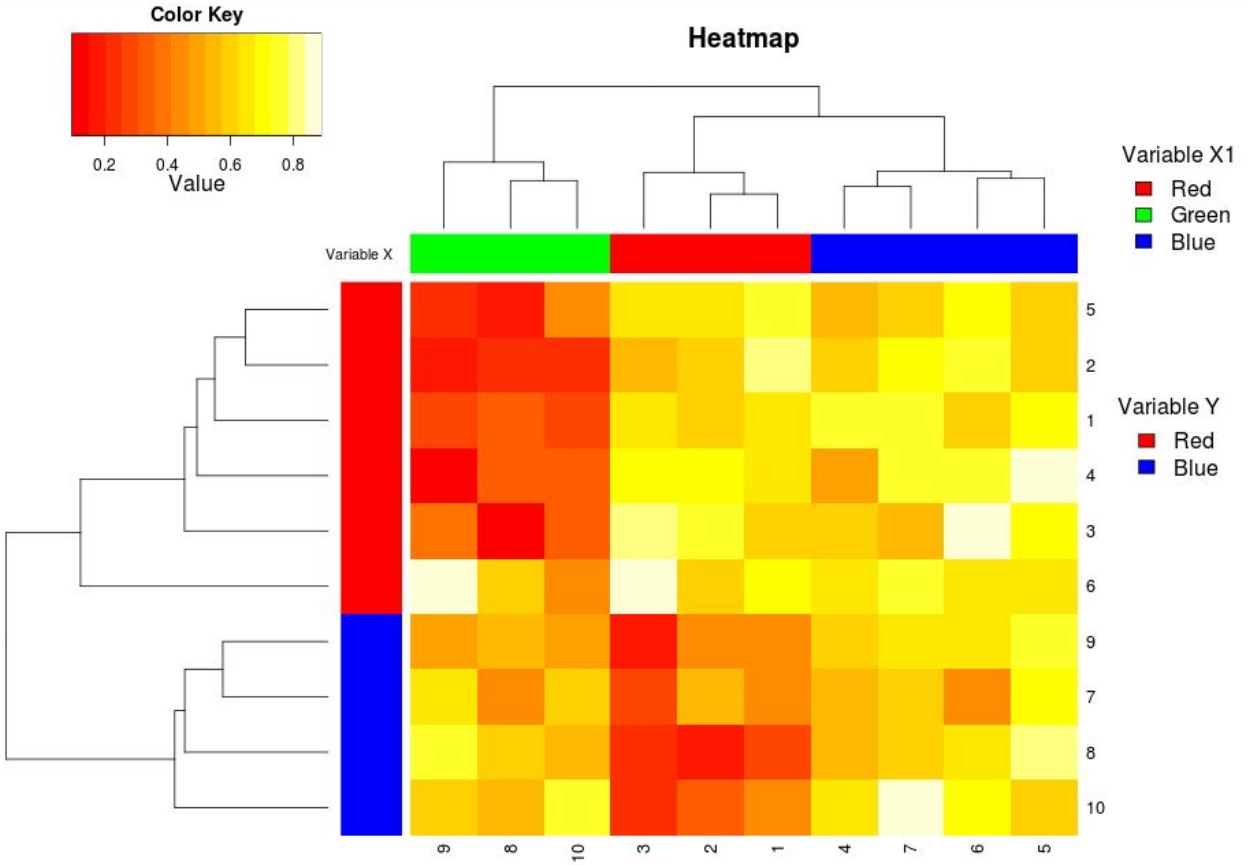
Let's Make Some Heatmaps

Simple First Example Usage

```r
#Creates a matrix
mat=matrix(c(rnorm(6,0.70,0.10),rnorm(4,0.30,0.10),rnorm(6,0.70,0.10),
             rnorm(4,0.30,0.10),rnorm(6,0.70,0.10),rnorm(4,0.30,0.10),
             rnorm(40,0.70,0.10),rnorm(5,0.30,0.10),rnorm(5,0.60,0.10),
             rnorm(5,0.30,0.10),rnorm(5,0.60,0.10),rnorm(5,0.30,0.10),
             rnorm(5,0.60,0.10)),nrow=10)

#Creates a matrix of colum groups
cgroup1 = c(rep("red",3),rep("blue",4),rep("green",3))
column_annotation = as.matrix(cgroup1)
colnames(column_annotation) = c("Variable X")

#Clusters rows into groups then creates a matrix of those groups
rowGroup = cutree(hclust(dist((mat))),k=2)
row_annotation = rbind(c("red","blue")[rowGroup])
rownames(row_annotation) = c("Variable Y")
```

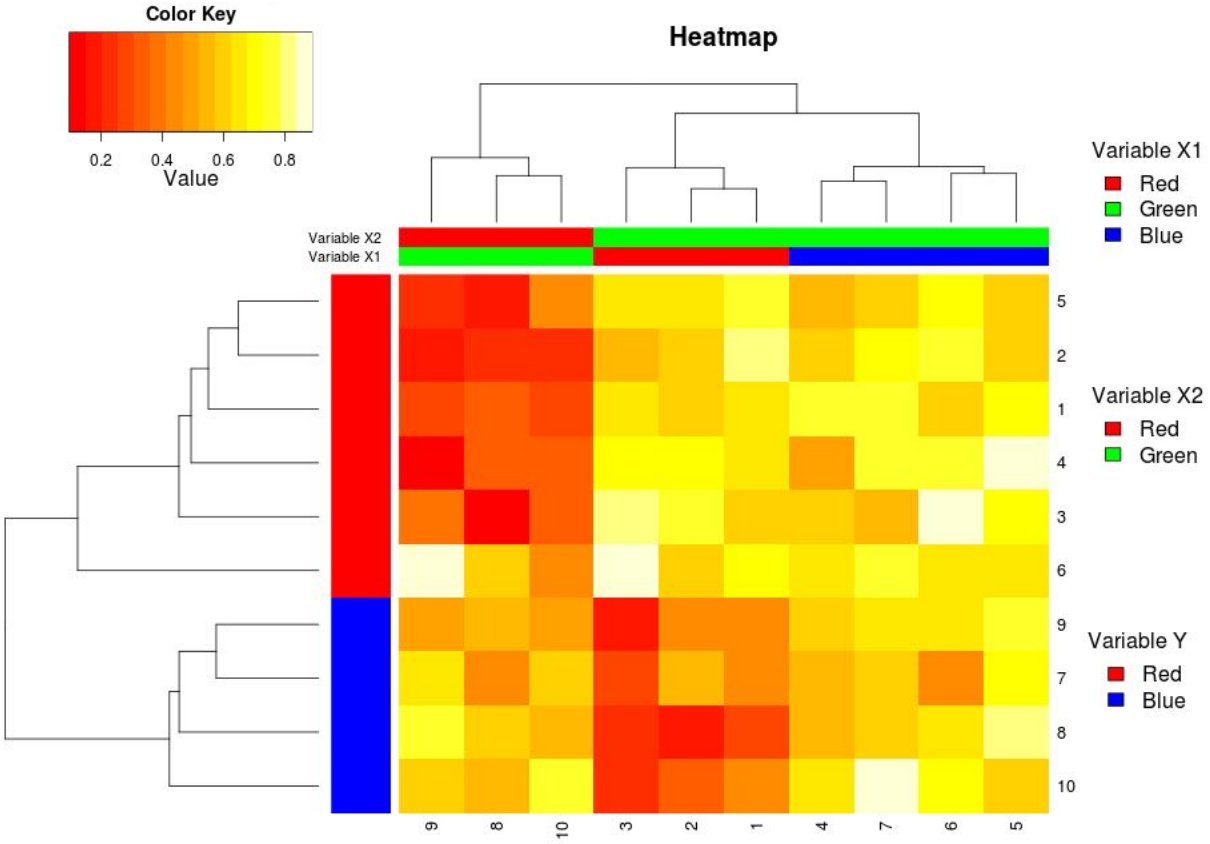
Simple First Example Usage

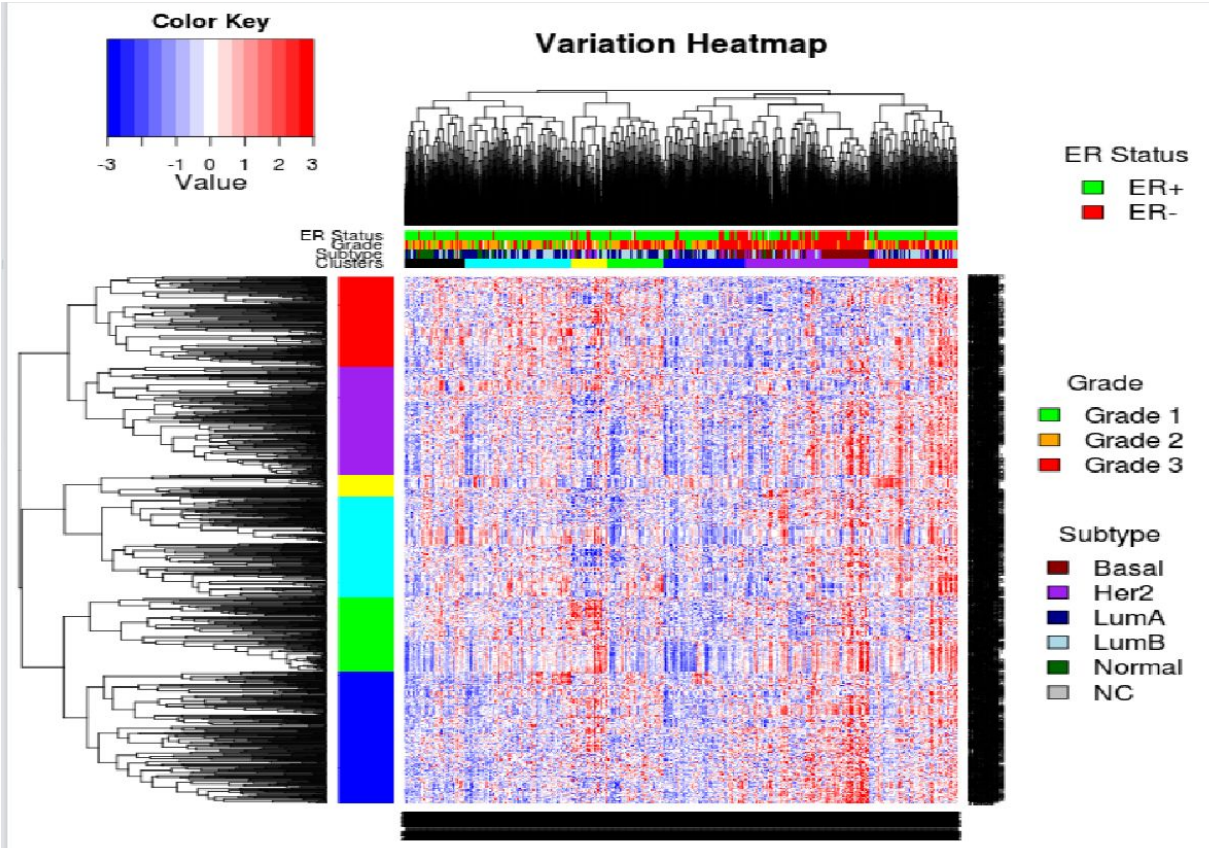
```
18  #Produces Heatamp
19  heatmap.3(mat, RowSideColors=row_annotation,
20            distfun=function(x) dist(x,method = "euclidean"),
21            ColSideColors=column_annotation,
22            margins = c(5,15),
23            main =  "Heatmap")
24
25  legend("topright", legend=c("Red", "Green", "Blue"),
26         fill=c("Red", "Green", "Blue"), bty="n", lty=0,
27         box.lty=0,title="Variable X1")
28  legend("right", legend=c("Red", "Blue"), fill=c("Red", "Blue"),
29         bty="n", lty=0, box.lty=0,title=" Variable Y   ")
```


Color Key
0.2
0.4
0.6
0.8
Value
Heatmap
Variable X
Variable X1
Red
Green
Blue
Variable Y
Red
Blue
5
2
1
4
3
6
9
7
8
10
9
8
10
3
2
1
4
7
6
5

Simple Second Example Usage

```
# multiple columns

#Creates a second matrix of colum groups
cgroup2 = c(rep("green",7),rep("red",3))
#Combines the two colum groups into a single matrix (note: Use rbind if producing row annotation)
column_annotation = cbind(cgroup1,cgroup2)
#Sets names
colnames(column_annotation) = c("Variable X1", "Variable X2")

#Clusters rows into groups then creates a matrix of those groups
rowGroup = cutree(hclust(dist((mat))),k=2)
row_annotation = rbind(c("red","blue")[rowGroup])
rownames(row_annotation) = c("Variable Y")

#Produces heatmap
heatmap.3(mat, RowSideColors=row_annotation,
          distfun=function(x) dist(x,method = "euclidean"),
          ColSideColors=column_annotation,
          margins = c(5,15),
          main = "Heatmap")

#Adds Legends
legend("topright", legend=c("Red", "Green", "Blue"),
       fill=c("Red", "Green", "Blue"), bty="n",
       lty=0, box.lty=0,title="Variable X1")
legend("right", legend=c("Red", "Green"),
       fill=c("Red", "Green"), bty="n",
       lty=0, box.lty=0,title="Variable X2")
legend("bottomright", legend=c("Red", "Blue"),
       fill=c("Red", "Blue"), bty="n", lty=0,
       box.lty=0,title=" Variable Y   ")
```


Color Key
0.2
0.4
0.6
0.8
Value
Heatmap
Variable X2
Variable X1
Variable X1
Red
Green
Blue
Variable X2
Red
Green
Variable Y
Red
Blue
5
2
1
4
3
6
9
7
8
10
9
8
10
3
2
1
4
7
6
5


Color Key
-3
-1
0
1
2
3
Value
Variation Heatmap
ER Status
Grade
Subtype
Clusters
ER Status
ER+
ER-
Grade
Grade 1
Grade 2
Grade 3
Subtype
Basal
Her2
LumA
LumB
Normal
NC

Code

- The Full code can be found in the comments