Heatmap using R (introduction)

Permanent link to video: http://tinyurl.com/mv690/seminars/heatmap
Permanent link to this lab: http://tinyurl.com/mv690/seminars/heatmap/lab

Heat maps are useful for visualizing multivariate data but must be applied properly. It is a popular technique to display dense and intuitive information, especially for high-dimensional data. Heat maps use color to represent numbers. When used with suitable color scales, clustering can dramatically affect our ability to see structure in heat maps. After rows and columns are arranged according to similarity, previously undetectable patterns can become obvious (blocks of cells with similar colors).

Correlation matrix

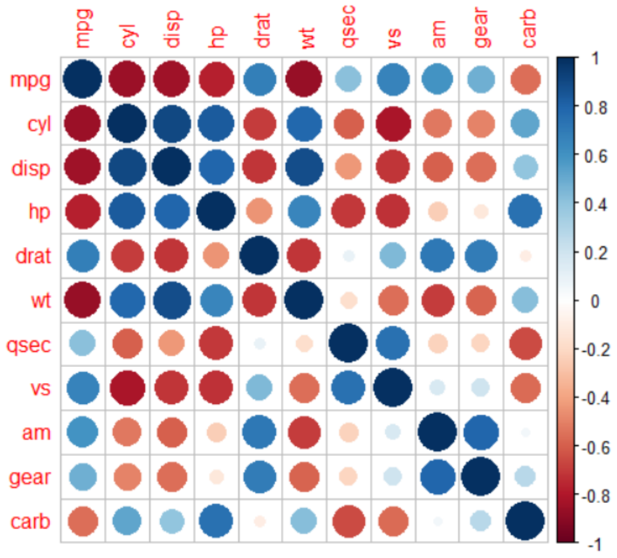
```
## if there is any error message, please install the package first
library(corrplot)
M = cor(mtcars)  ## firstly, compute the correlation
corrplot(M) ## the correlation values are represented by color and size by default
```

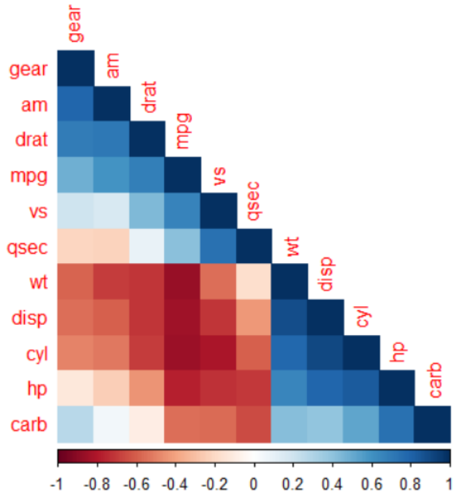
It’s good, but hard to identify the variable similarity. We can reorder the variables:

```
corrplot(M, method = 'color', order = 'AOE')
corrplot(M, method = 'color', order = 'AOE', type = 'lower')  ##only keep lower of correlation matrix
```

It’s also easy enough if you want to label the significance of p values.

```
testRes = cor.mtest(mtcars, conf.level = 0.95)
corrplot(M, p.mat = testRes$p, method = 'color', diag = FALSE, type = 'upper', sig.level = c(0.001, 0.01, 0.05), pch.cex = 0.9, insig = 'label_sig', pch.col = 'grey20', order = 'AOE',col = COL2('RdYlBu'))
```





With hierarchical cluster tree

First step is to install the package:

```
install.packages('BiocManager')
BiocManager::install("ComplexHeatmap")
library(ComplexHeatmap)
```

In this mini-tutorial, we will use 'iris' dataset as an example. In total 150 records from 3 species:

```
df <- scale(iris[, 1:4]) ## scale dataset besides the Species column
rownames(df) <- paste0('r', 1:150) ## give each record one rowname
head(df) ## just check the dataset
```

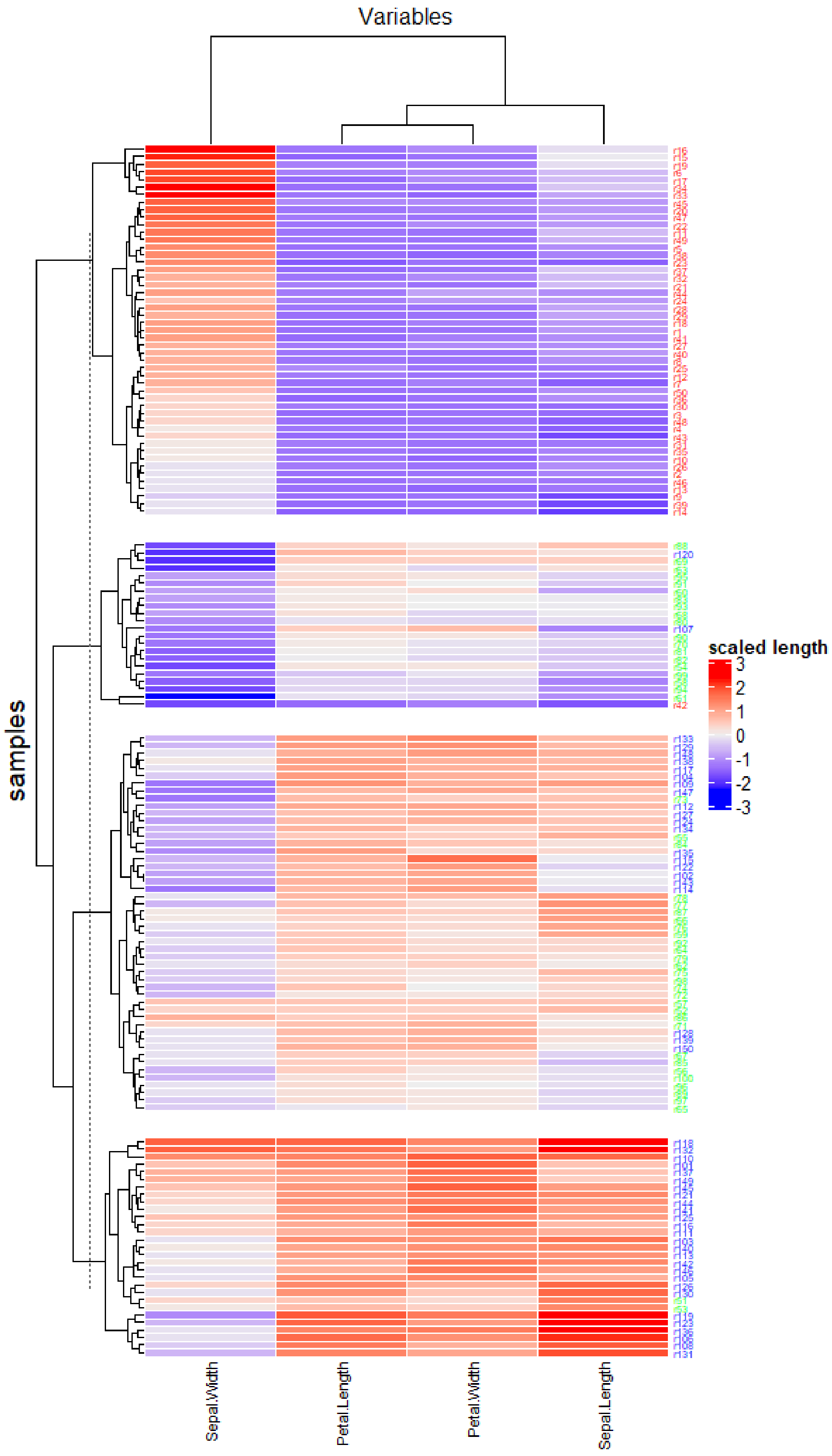
Here, we can plot the heatmap:

```
Heatmap(df,
        name = "scaled_length", #title of legend
        column_title = "Variables", row_title = "Samples",
        row_names_gp = gpar(fontsize = 7) # Text size for row names
)
```

It's fine, but not very pretty. We can customize the heatmap:

```
Heatmap(df, name = 'scaled length', column_title = 'Variables',
        row_title = 'samples',
        column_dend_height = unit(2, "cm"), ## change the height of
dendrogram of variables
        row_dend_width = unit(2, "cm"), ## same but for dendrogram of
samples
        row_dend_reorder = T,
        column_title_gp = gpar(fontsize = 12),
        column_names_gp = gpar(fontsize = 8),
        row_names_gp = gpar(col = rep(c('red', 'green', 'blue'), each
= 50), fontsize = 4), ## use color to represent species
        rect_gp = gpar(col = "white", lwd = .2), ## add white border
        row_km = 4, row_gap = unit(5, "mm"))
## split the sample dendrogram into 4 subgroups (dashed line), and add
spaces to separate them in the figure
```

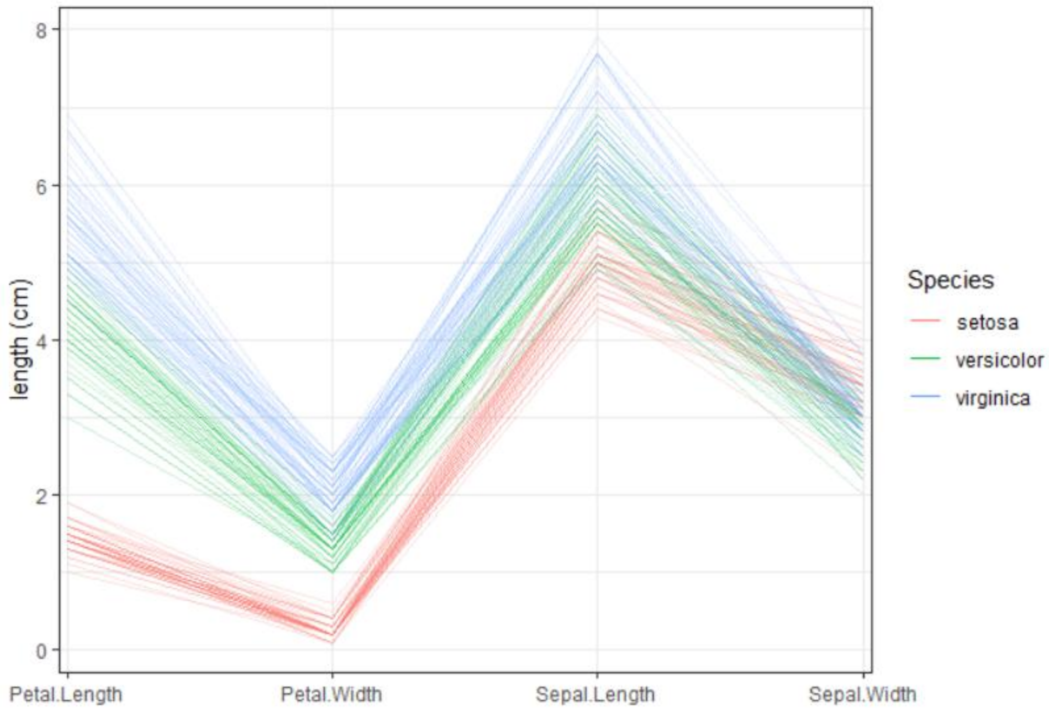
More details about this package, please check this website:
https://jokergoo.github.io/ComplexHeatmap-reference/book/


Variables
samples
scaled length
3
2
1
0
-1
-2
-3
Sepal.Width
Petal.Length
Petal.Width
Sepal.Length

Besides heatmap

Heatmap is suitable for visualizing intense information. However, it is not the only option available. If your datasets are in reasonable size with relatively low complexity, simple line plots should be recommended.

```
library(ggplot2) ## if there any error, please install this package
first
library(dplyr); library(tidyr) ## same here
head(iris) ## check the dataset
iris2 <- iris %>% mutate(id = 1:150) %>%
  pivot_longer(col = Sepal.Length:Petal.Width, names_to = 'var',
values_to = 'length') ## change the dataset from wide to long
ggplot(data = iris2,
       aes(x = var, y = length, group = id, col = Species)) +
geom_line(alpha = .2) +
  theme_bw() +
  scale_x_discrete(expand = c(0.01,0)) + labs(x = '', y = 'length
(cm)') +
  guides(col=guide_legend(override.aes=list(alpha=c(1, 1, 1))))
```


length (cm)
Petal.Length
Petal.Width
Sepal.Length
Sepal.Width
Species
setosa
versicolor
virginica