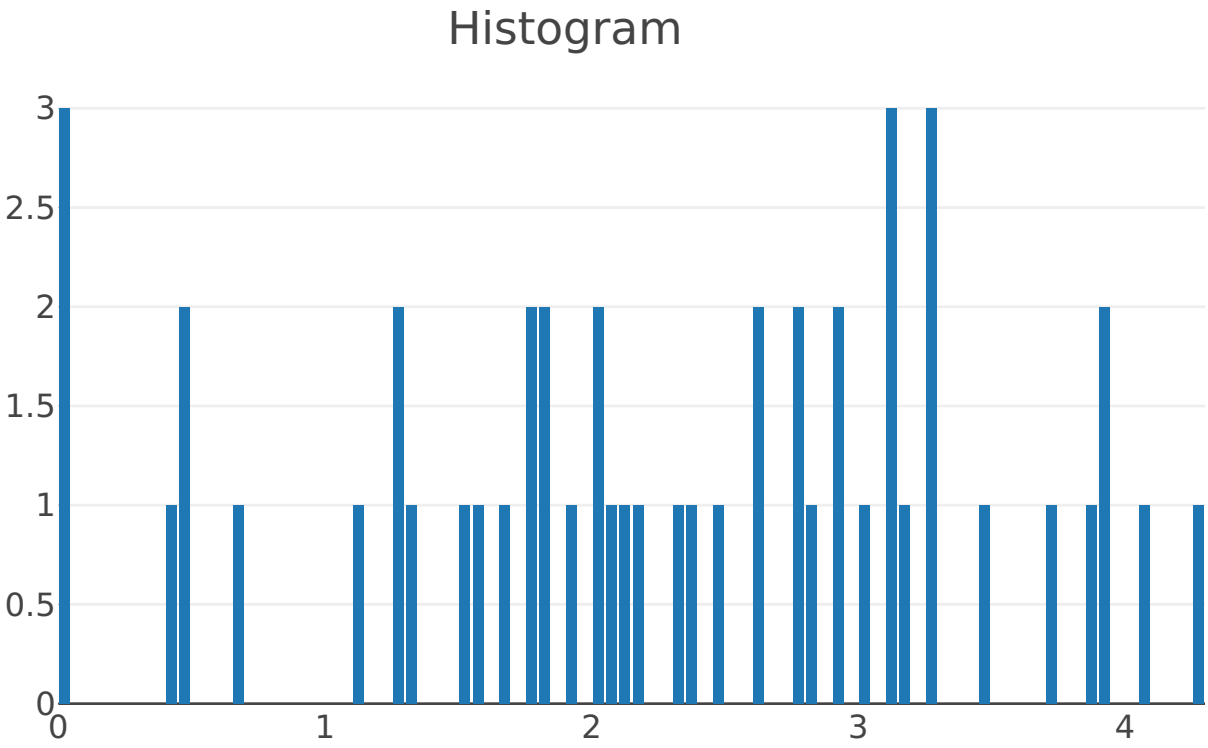
Count data DE analysis

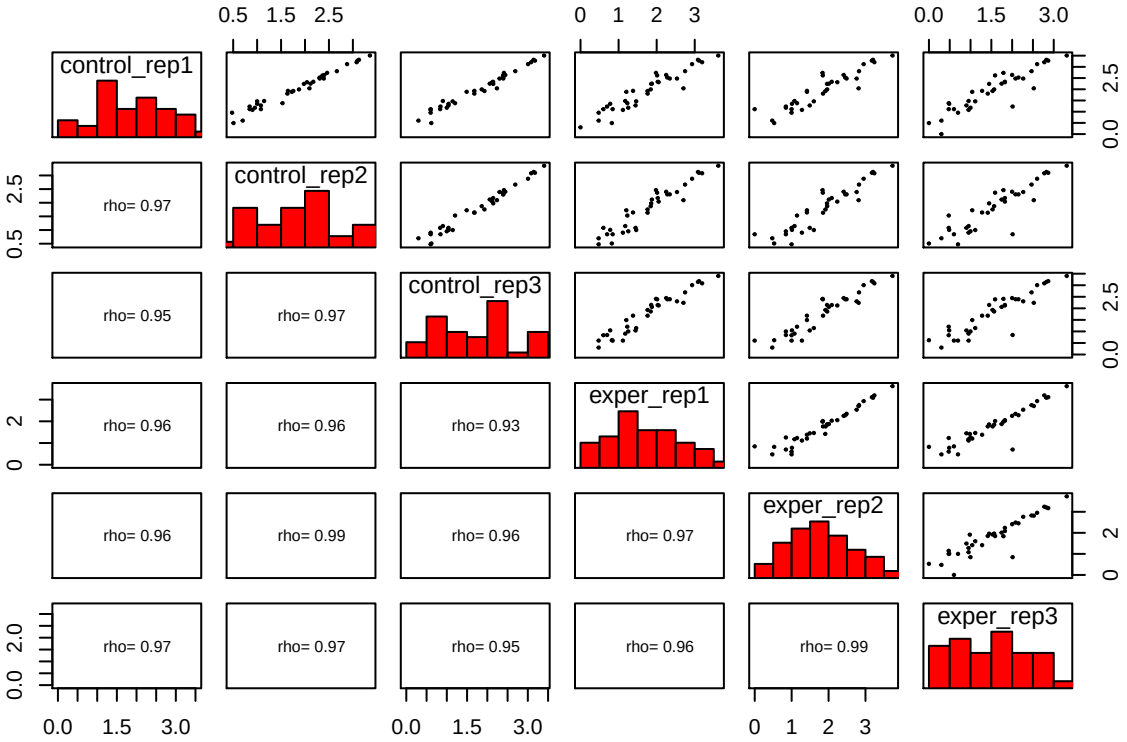
[1]. Reading the data.

The merged the count dara table will be read from a web URL to be able to run this rmarkdown anywhere. In Eukaryotes only a subset of all genes are expressed in a given cell. Expression is therefore a bimodal distribution, with non-expressed genes having counts that result from experimental and biological noise. It is important to filter out the genes that are not expressed before doing differential gene expression. You can decide which cutoff separates expressed vs non-expressed genes by looking your histogram we created.



[2]. All2all scatter plots

To check the reproducibility of biological replicates, we use all2all plots.

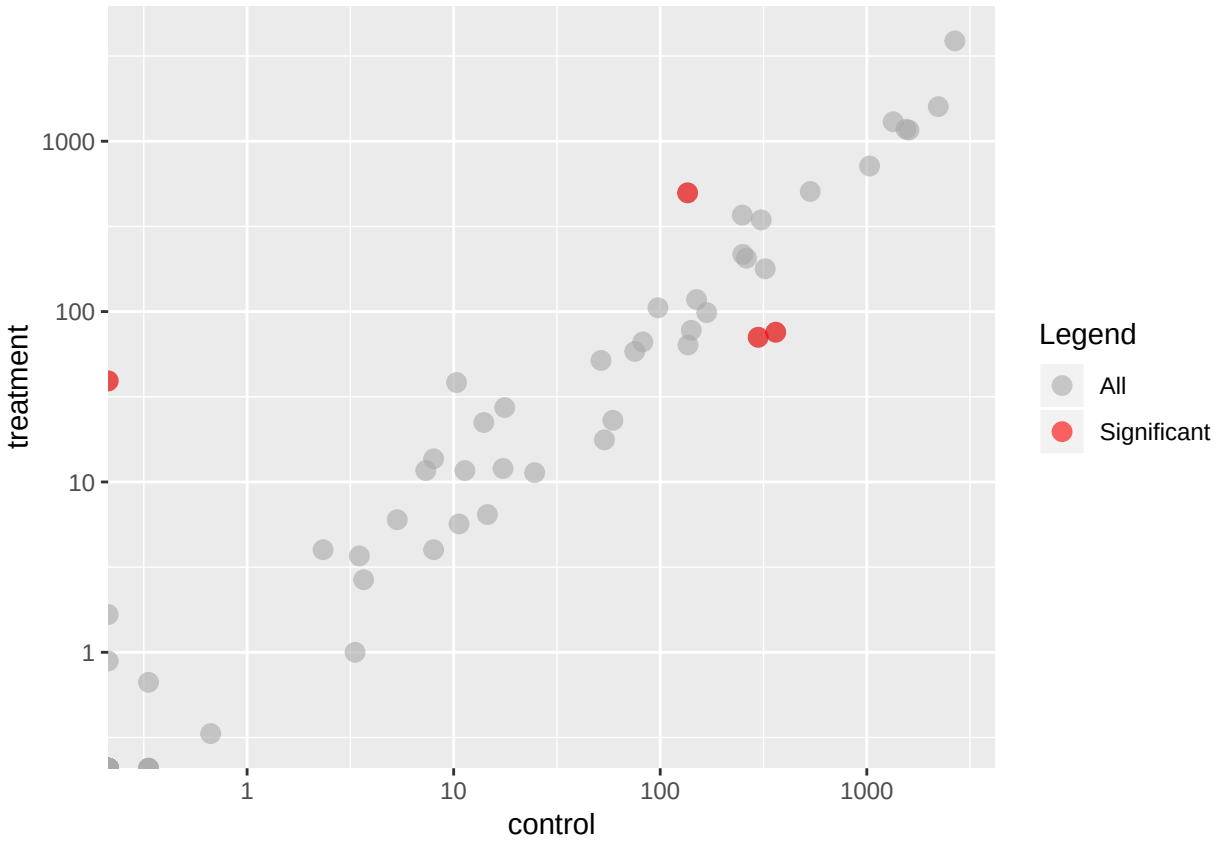


[3]. DESeq ANALYSIS

The goal of Differential gene expression analysis is to find genes or transcripts whose difference in expression, when accounting for the variance within condition, is higher than expected by chance.

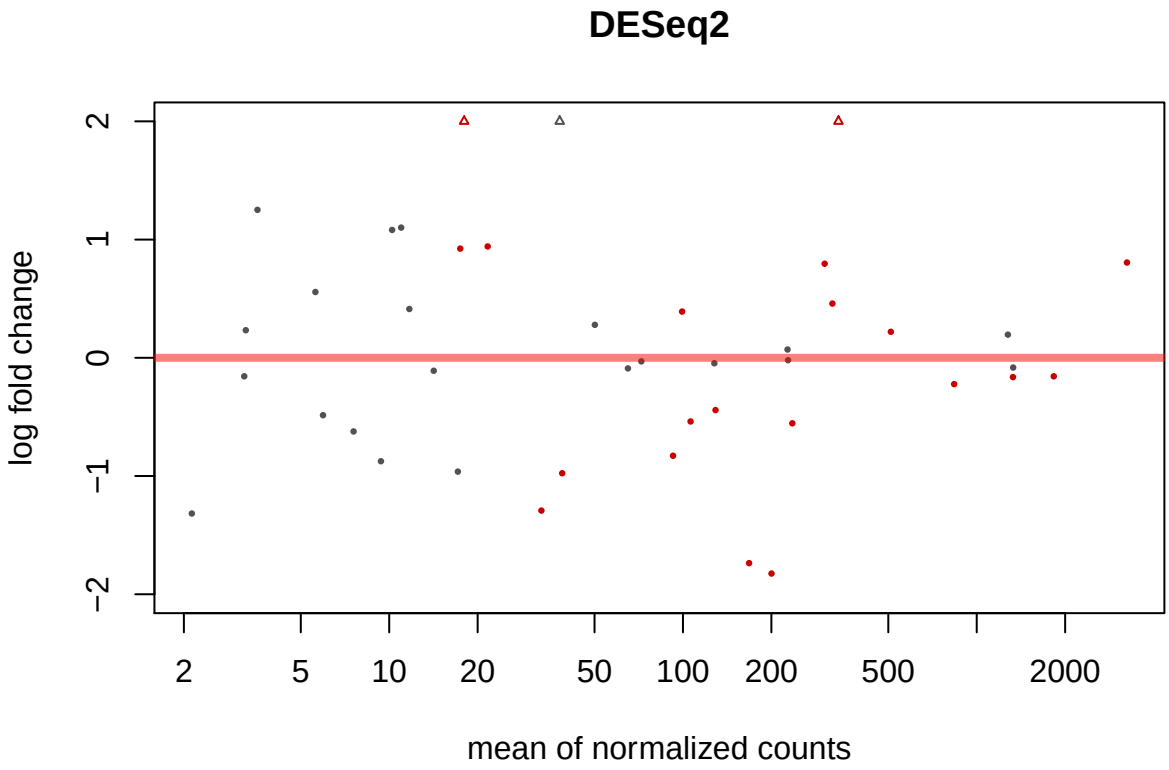
The first step is to indicate the condition that each column (experiment) in the table represent. Here we define the correspondence between columns and conditions. Make sure the order of the columns matches to your table.

In this case a total sum of 10 counts separates well expressed from non-expressed genes. You can change this value and padj value and log2FoldChange cutoffs according to your data



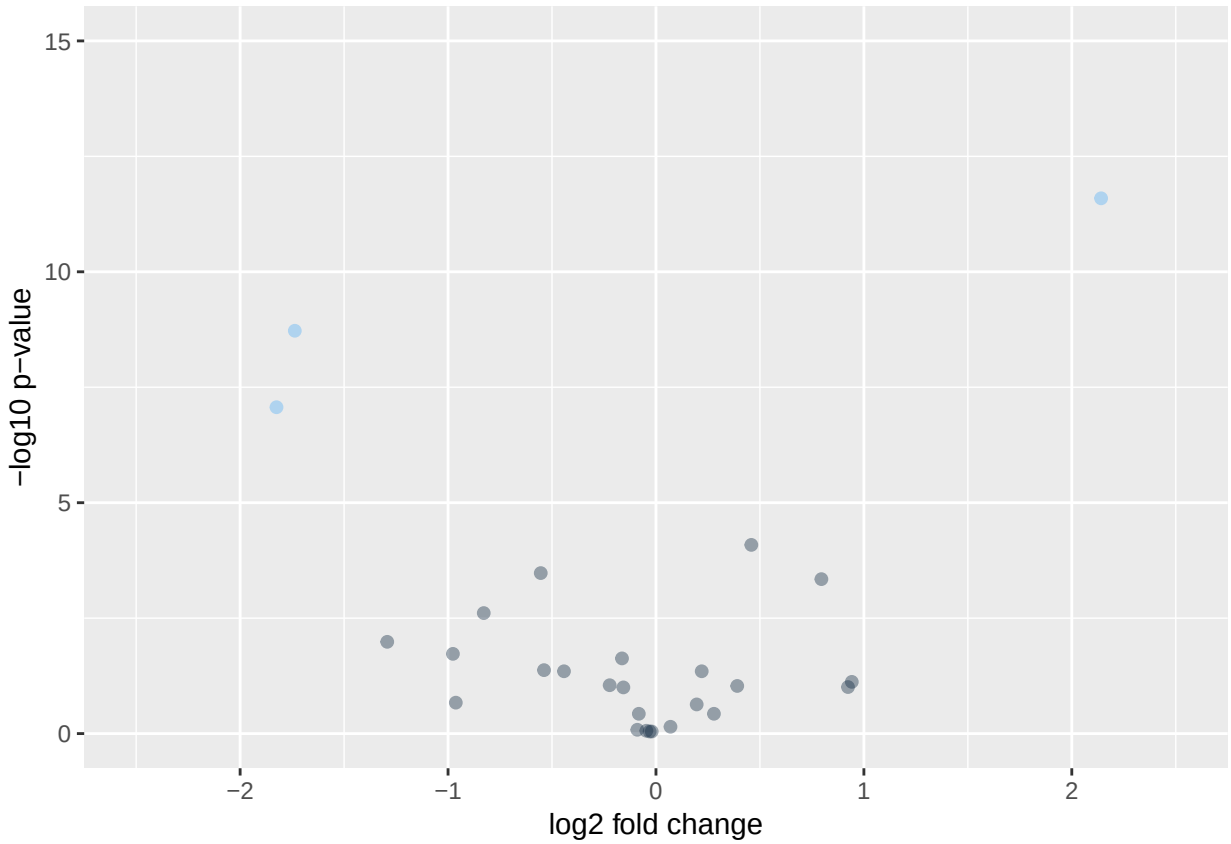
[4]. MA Plot

The Second way to visualize it, we use MA plots. For MA Plot there is another builtin function that you can use



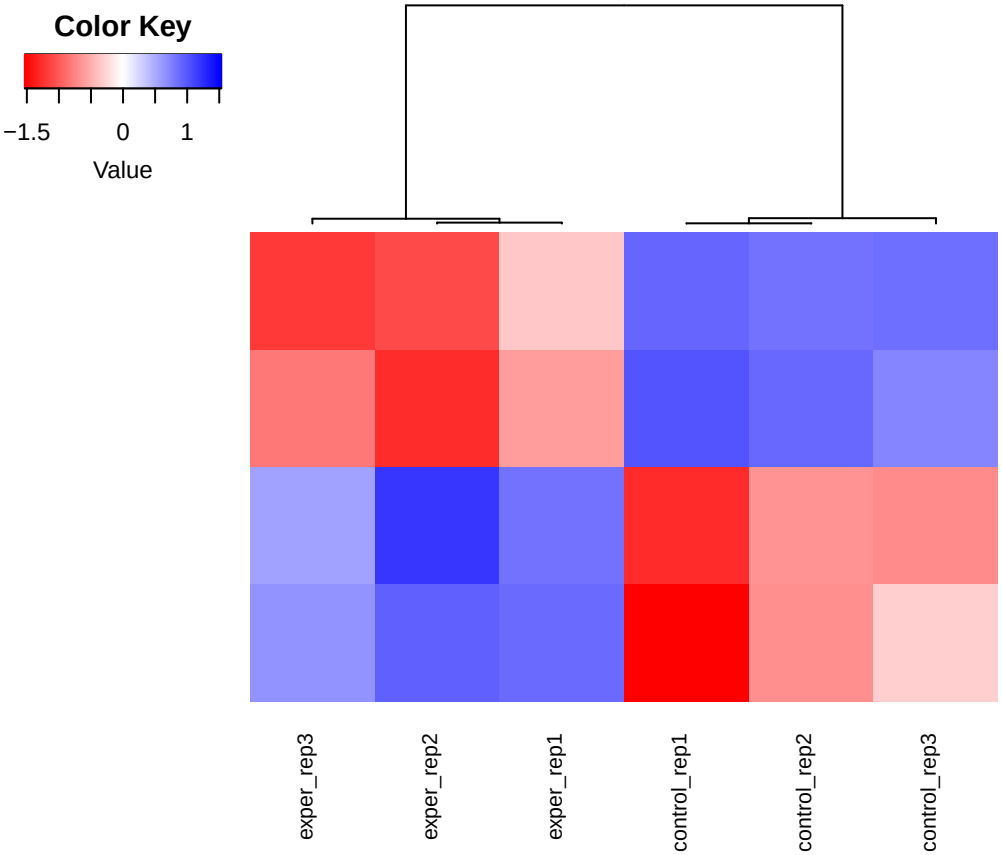
[5]. Volcano Plot

The third way of visualizing the data is making a Volcano Plot. Here on the x axis you have log2foldChange values and y axis you have your -log10 padj values. To see how significant genes are distributed. Highlight genes that have an absolute fold change > 2 and a padj < 0.01



[6] Heatmap

The forth way of visualizing the data that is widely used in this type of analysis is clustering and Heatmaps.


Color Key
−1.5
0
1
Value
exper_rep3
exper_rep2
exper_rep1
control_rep1
control_rep2
control_rep3