RNA-seq analysis in R

Initial exploration of RNA-seq data - solutions

Contents

Data

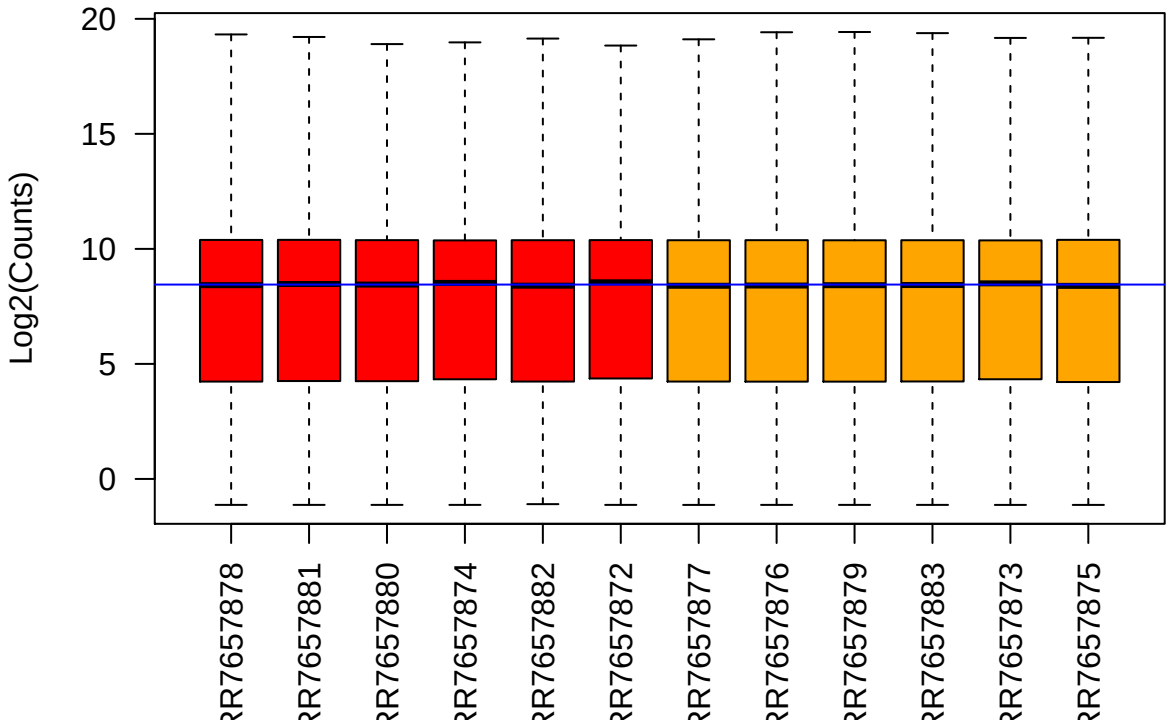
```
library(tximport)
library(DESeq2)
library(tidyverse)
```

Challenge 1

1. Use the `DESeq2` function `rlog` to transform the count data. This function also normalises for library size.
2. Plot the count distribution boxplots with this data. How has this effected the count distributions?

```
rlogcounts <- rlog(filtCounts)

# Check distributions of samples using boxplots
boxplot(rlogcounts,
        xlab="",
        ylab="Log2(Counts)",
        las=2,
        col=statusCols)
# Let's add a blue horizontal line that corresponds to the median logCPM
abline(h=median(as.matrix(rlogcounts)), col="blue")
```


Log2(Counts)
20
15
10
5
0
RR7657878
RR7657881
RR7657880
RR7657874
RR7657882
RR7657872
RR7657877
RR7657876
RR7657879
RR7657883
RR7657873
RR7657875