**Installing R:**

You can install R using the command line version (https://cran.cnr.berkeley.edu) or R studio (https://www.rstudio.com/), an IDE for R.

**R- Processing Data**

#Read in the fpkm data

> G<-read.table("all\_genes.fpkm",header=T)

> M<-as.matrix(G[,c(3:21)])

#create boxplot

> boxplot(log2(M),use.cols=TRUE,main="log2FPKM")

#Let's normalize the datasets

######Quantile Normalization of the fpkm data

#load the preprocessCore library from Bioconductor, which contains the normalize.quantiles function

> source("http://bioconductor.org/biocLite.R")

> biocLite("preprocessCore")

> library(preprocessCore)

#now create the normalized quantile data

> N<- normalize.quantiles(M)

> boxplot(log2(N),use.cols=TRUE,main="log2FPKM-QNorm")

#Now lets see how they cluster using PCA

#See end of document if you have installation problems with rgl on ubuntu

> install.packages("rgl")

> library(rgl)

> P<-prcomp(t(G[,c(3:21)]))

> plot3d(P$x,xlab="PC1",ylab="PC2",zlab="PC3",axes=FALSE)

> spheres3d(P$x[1:6,1:3], radius=10,col="red")

> spheres3d(P$x[1:6,1:3], radius=100,col="red")

> spheres3d(P$x[7:11,1:3], radius=100,col="yellow")

> spheres3d(P$x[12:16,1:3], radius=100,col="blue")

> spheres3d(P$x[17:19,1:3], radius=100,col="green")

> box3d(col="#989898")

**HEATMAP**

> install.packages("gplots")

> library(gplots)

> data = read.table("all\_quantile.txt",header=T)

> data = log2(data)

####!!!WARNING!!!This heatmap plots every gene for every experiment, so do it only if you have time to waste and want to see how it looks like. Generally, this kind of global heatmap is not very useful

> heatmap.2(as.matrix(data), col=redgreen(75), scale="row", key=T, keysize=1.5,density.info="none", trace="none",cexCol=0.9, labRow=NA,margins=c(10,10))

#lets just do top 100 fold changed genes

> data = read.table("heatmap\_genes\_top100.txt",header=T)

> data = log2(data)

> data2 = normalize.quantiles(as.matrix(data))

> colnames(data2) = colnames(data)

> rownames(data2) = rownames(data)

> heatmap.2(as.matrix(data2), col=redgreen(100), scale="row", key=T, keysize=1.5,density.info="none", trace="none",cexCol=0.9, labRow=NA,margins=c(10,10))

#create file

> jpeg(file="heatmap\_top100.jpg")

**CORRELATIONS**

#this data is already normalized, otherwise, you should normalize first!

> data = read.table("all\_quantile.txt",header=T)

> cor(data,method="pearson")

> cor(data,method="spearman")

#Installation issues with rgl library and X11 unmet dependency on Ubuntu 14.04

> sudo apt-get install libglew-dev libcheese7 libcheese-gtk23 libclutter-gst-2.0-0 libcogl15 libclutter-gtk-1.0-0 libclutter-1.0-0

> sudo apt-get build-dep r-cran-rgl

Then in R

> install.packages(“rgl”)