

BMI 882 Homework 1

Li Ge

2020-02-27

The homework requirement is [here](#).

Reproduce Figure 1b

First, I downloaded the [source materials](#), specifically, [Everything \(177 MB\)](#). Unzip it, go to the subfolder `/RNowebSource`, and then the `examiningDoxorubicinInDetail.Rnw` file for the details of section 2 - Case study 1: Doxorubicin. Then straight copy and paste the minimal code required to recreate figure 1b. The subtitles are just the R section names in their original `examiningDoxorubicinInDetail.Rnw` file. I do have to copy and paste the `/RawData/PottiNatMed` subfolder to my working directory to load the data to just run the original code.

loadGenes

```
doxorubicinGenes <-  
  read.table(file.path("RawData", "PottiNatMed",  
                        "GeneListsNov07", "Adria(final).txt"),  
             sep="\t", header=TRUE)  
dim(doxorubicinGenes)  
doxorubicinGenes[1:3,]  
doxorubicinGenes <- as.character(doxorubicinGenes[, "Probe.Set.ID"])  
doxorubicinGenes[1:3]
```

loadDoxo07

```
tempDoxorubicin07Header1 <-  
  read.table(file.path("RawData", "PottiNatMed", "Adria_ALL.txt"), sep="\t",  
             nrows=1, header=FALSE)  
tempDoxorubicin07Header1 <- as.vector(t(tempDoxorubicin07Header1))  
  
tempDoxorubicin07Header2 <-  
  read.table(file.path("RawData", "PottiNatMed", "Adria_ALL.txt"), sep="\t",  
             skip=1, nrows=1, header=FALSE)  
tempDoxorubicin07Header2 <- as.vector(t(tempDoxorubicin07Header2))  
  
tempDoxorubicin07Header1[1:25]  
tempDoxorubicin07Header2[1:25]  
  
table(tempDoxorubicin07Header1)  
table(tempDoxorubicin07Header2)
```

buildDoxo07Info

```
tempSampleNames <-  
  c(paste("Training", c(1:22), sep=""),  
    paste("Test", c(1:122), sep=""))  
  
tempGroup <- c(rep("Training", 22), rep("Test", 122))  
tempStatus <- tempDoxorubicin07Header2  
tempStatus[tempStatus == "Sens"] <- "Sensitive"  
tempStatus[tempStatus == "NR"] <- "Resistant"  
tempStatus[tempStatus == "Resp"] <- "Sensitive"  
  
doxorubicin07Info <-  
  data.frame(sampleGroup = tempGroup,  
             status      = tempStatus,  
             row.names   = tempSampleNames)  
  
doxorubicin07Info[c(1:2,22:25),]  
  
#rm(doxorubicin07Header1, doxorubicin07Header2, tempGroup, tempStatus)  
rm(list=ls(pattern="^temp"))
```

loadDoxo07Numbers

```
doxorubicin07Numbers <-  
  read.table(file.path("RawData", "PottiNatMed", "Adria_ALL.txt"), sep="\t",  
            skip=2, header=FALSE)  
  
colnames(doxorubicin07Numbers) <- rownames(doxorubicin07Info)  
  
doxorubicin07Numbers[1:4,c(1:2,22:25)]
```

getDoxoCors

```
doxorubicin07Cors <- cor(doxorubicin07Numbers)  
sum(doxorubicin07Cors > 0.9999)  
sum(diag(doxorubicin07Cors) > 0.9999)
```

plotDoxoHighCors

```
doxorubicin07HighCors <- (doxorubicin07Cors > 0.9999)  
  
same07Status <- matrix(rep(doxorubicin07Info[, "status"], 144), 144, 144)  
same07Status <- (same07Status == t(same07Status))  
  
temp07Ties <- which(doxorubicin07HighCors & same07Status, arr.ind=TRUE)  
temp07Ties <- temp07Ties[temp07Ties[,1] != temp07Ties[,2],]  
temp07BadTies <- which(doxorubicin07HighCors & (!same07Status), arr.ind=TRUE)
```

```

plot(1:144, 1:144, pch=".", cex=3,
     xlim=c(0.5,144.5), ylim=c(0.5,144.5),
     xaxs="i", yaxs="i",
     xlab="Column in Adria_ALL.txt",
     ylab="Column in Adria_ALL.txt",
     main="Identical Columns in Adria_ALL.txt",
     )
points(temp07Ties[,1],temp07Ties[,2],pch=21,bg="white",lwd=1)
points(temp07BadTies[,1],temp07BadTies[,2],pch=24,bg="red",lwd=2)
abline(h=22.5, v=22.5)

abline(v=32, lty="dashed")

```

