Yes or no. Motivate you answer (try to be concise and complete):

1. Performing hierarchical clustering with a correlation-based similarity metric applied on non-rescaled data (mean centered, variance rescaled) gives exactly the same results as when you would use hierarchical clustering with a Euclidian distance applied on the non-rescaled data.
2. Increasing the number of clusters as a parameter in K Means clustering reduces the cluster size
3. The is the output of a PCA analysis on the following dataset

Explain in words what PCA is doing.

What is the meaning/interpretation of loadings and scores (see graph below) .





3)

a) What would this code be doing? Golub is a gene-patient dataset. The dataset contains x patients and 3051 genes. Explain the code line per line. Explain why you need to perform the different steps. Explain what you expect in the output.

golub\_m = golub – rowMeans(golub) [code on the whole dataset]

SD = apply(golub, 1, sd, na.rm=TRUE)

golub\_r = golub\_m/SD

d.euclidian <- dist(golub\_r, method="euclidian")

m <- as.matrix(dist(golub))

dim(m)

#(3051,3051)

kmeangolub100= kmeans(d.euclidian\_r,100)

b) what will happen if you change the following parameter setting

kmeangolub10= kmeans(d.euclidian\_r,10)

Details: A matrix containing 52 timepoints (columns) and 154 metabolites (rows). (dim 154 X 52)

PCAres<-prcomp(t(mDC), scale = TRUE, center=TRUE)

**plot(predict(PCAres)[,1],predict(PCAres)[,2])**

**abline(v=0, col="gray")**

**abline(h=0, col="gray")**

**text(predict(PCAres)[,1],predict(PCAres)[,2], labels=sub("X(.+h)(\\..)?","\\1",rownames(t(mDC))),cex=1, adj = c(0,0))**

ylim=c(-8,10)**)**

PCAres$rotation

> PCAres$rotation[,1:2]

PC1 PC2

X0h 0.021863836 0.023906003

X0h.1 0.039028678 0.017691394

X0h.2 0.005461632 -0.017034873

X0h.3 -0.041738981 -0.120975500

X0h.4 -0.013439588 0.040876208

X0h.5 -0.004382840 0.098624472

X0h.6 -0.005265209 -0.027082855

X1h 0.100321907 -0.170260192

X1h.1 0.124223220 -0.158749737

X1h.2 0.113246106 -0.259660111

X1h.3 0.085097561 -0.262912989

X1h.4 0.107609797 -0.217270096

X1h.5 0.121947367 -0.200702477

X1h.6 0.106006583 -0.248034658

X1h.7 0.079960771 -0.283068089

X4h 0.143140565 -0.137883847

X4h.1 0.147126391 -0.153135963

X4h.2 0.136592457 -0.192953349

X4h.3 0.124793438 -0.236023856

X4h.4 0.145724443 -0.163820371

X4h.5 0.143321845 -0.153699220

X12h 0.163343705 -0.001921062

X12h.1 0.168373363 -0.011169471

X12h.2 0.166339755 -0.057639004

X12h.3 0.156473767 -0.036013092

X12h.4 0.163372589 -0.063960887

X12h.5 0.165558375 -0.073457096

X12h.6 0.161875346 -0.072861979

X12h.7 0.146561592 -0.124896560

X24h 0.161965803 0.075651813

X24h.1 0.168887749 0.023097883

X24h.2 0.164064620 0.054576321

X24h.3 0.161470931 0.024223839

X24h.4 0.167193641 0.062521375

X24h.5 0.164769605 0.034092888

X24h.6 0.165659697 0.020066778

X48h 0.153519030 0.131926136

X48h.1 0.158996413 0.128349314

X48h.2 0.161156899 0.127114467

X48h.3 0.158791128 0.113065151

X48h.4 0.158122105 0.135715208

X48h.5 0.161889076 0.125211253

X48h.6 0.155522618 0.108017688

X48h.7 0.161425788 0.099800912

X96h 0.150080075 0.168187010

X96h.1 0.154343044 0.162643355

X96h.2 0.155816215 0.151972897

X96h.3 0.154543525 0.152397362

X96h.4 0.148660901 0.159353433

X96h.5 0.152688568 0.159208078

X96h.6 0.154228756 0.161343858

X96h.7 0.154120584 0.139673120