compgen2021: Week 2 exercises

Yuna Son

Exercises for Week2

For this set of exercises we will be using the expression data shown below:

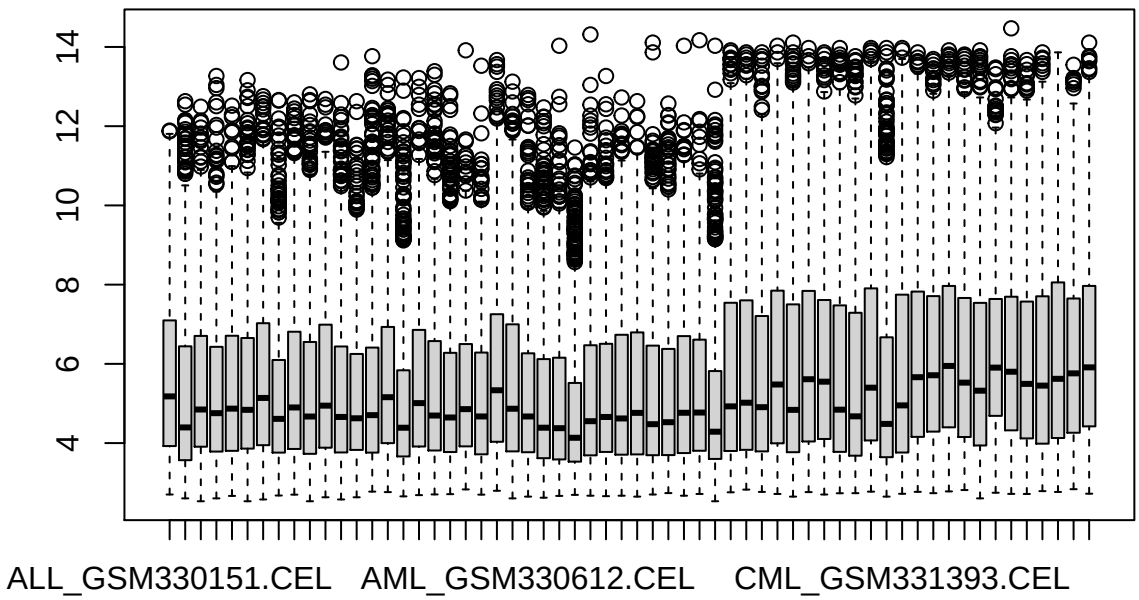
```
expFile=system.file("extdata",
                    "leukemiaExpressionSubset.rds",
                    package="compGenomRData")
mat=readRDS(expFile)
```

Clustering

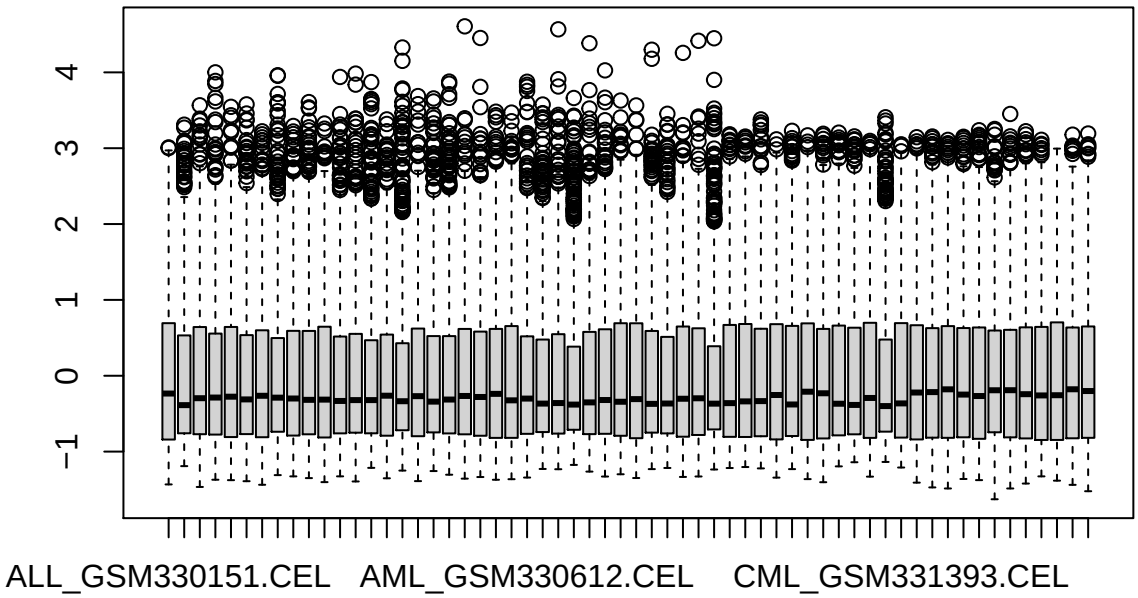
1. We want to observe the effect of data transformation in this exercise. Scale the expression matrix with the `scale()` function. In addition, try taking the logarithm of the data with the `log2()` function prior to scaling. Make box plots of the unscaled and scaled data sets using the `boxplot()` function. [Difficulty: **Beginner/Intermediate**]

solution:

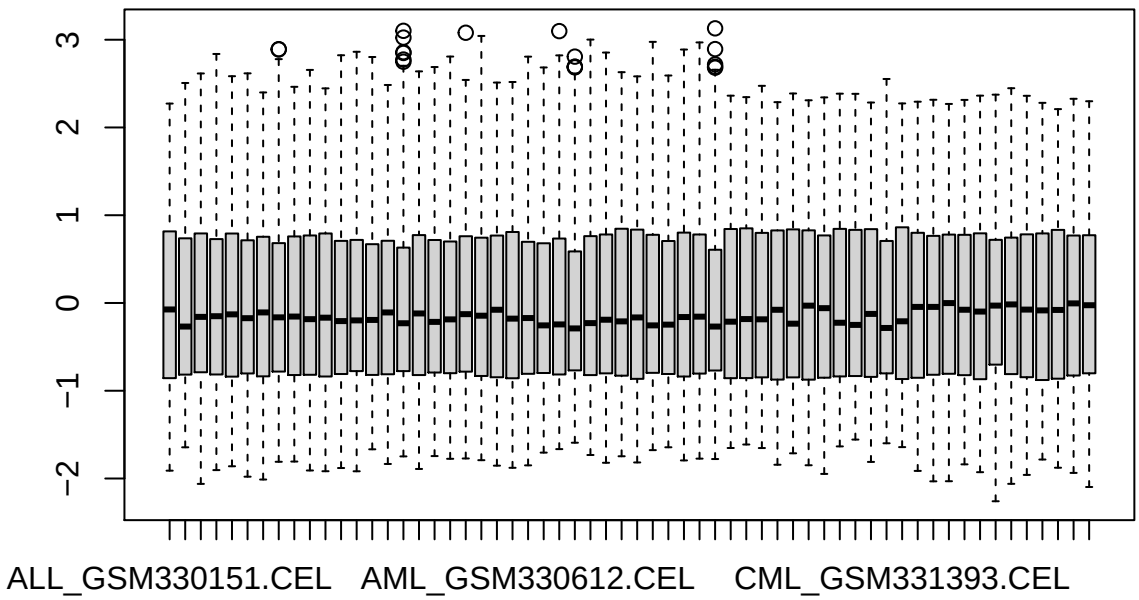
```
# Plot the unscaled data
p1 <- boxplot(mat, main = "Boxplot of the unscaled expression matrix")
```




```
# scale the expression matrix
scaled_exp = scale(mat)
p2 <- boxplot(scaled_exp, main = "Boxplot of the scaled expression matrix")
```




```
# log2 processing before scaling
log2_exp = log2(mat)
scaled_exp2 = scale(log2_exp)
p3 <- boxplot(scaled_exp2, main = "Boxplot of the log2 & scaled expression matrix")
```



2. For the same problem above using the unscaled data and different data transformation strategies, use the `ward.d` distance in hierarchical clustering and plot multiple heatmaps. You can try to use the `pheatmap` library or any other library that can plot a heatmap with a dendrogram. Which data-scaling strategy provides more homogeneous clusters with respect to disease types? [Difficulty: **Beginner/Intermediate**]

solution:

```
library(pheatmap)
```

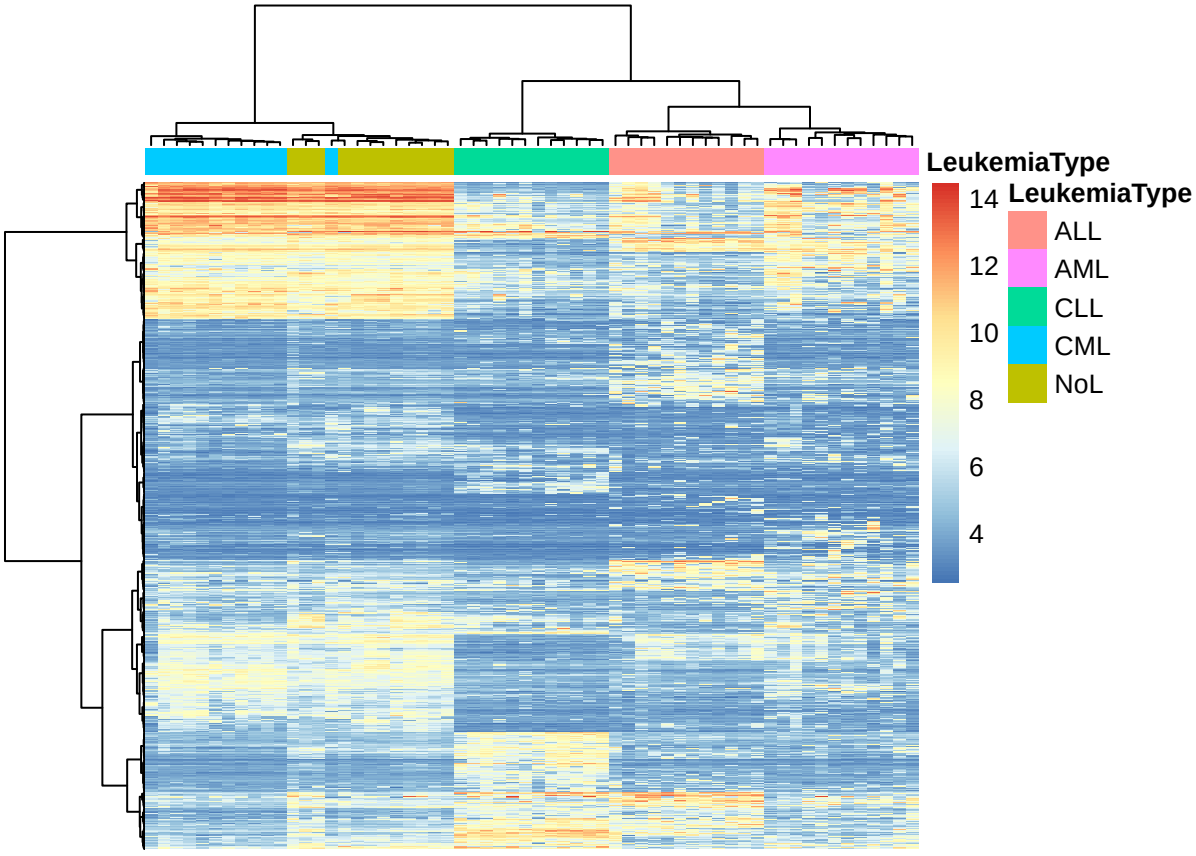
```
## Warning: package 'pheatmap' was built under R version 4.0.5
```

```
colnames(mat)
```

```
##  [1] "ALL_GSM330151.CEL" "ALL_GSM330153.CEL" "ALL_GSM330154.CEL"
##  [4] "ALL_GSM330157.CEL" "ALL_GSM330171.CEL" "ALL_GSM330174.CEL"
##  [7] "ALL_GSM330178.CEL" "ALL_GSM330182.CEL" "ALL_GSM330185.CEL"
## [10] "ALL_GSM330186.CEL" "ALL_GSM330195.CEL" "ALL_GSM330201.CEL"
## [13] "AML_GSM330532.CEL" "AML_GSM330546.CEL" "AML_GSM330559.CEL"
## [16] "AML_GSM330566.CEL" "AML_GSM330571.CEL" "AML_GSM330574.CEL"
## [19] "AML_GSM330580.CEL" "AML_GSM330584.CEL" "AML_GSM330593.CEL"
## [22] "AML_GSM330603.CEL" "AML_GSM330611.CEL" "AML_GSM330612.CEL"
## [25] "CLL_GSM330933.CEL" "CLL_GSM330934.CEL" "CLL_GSM330969.CEL"
## [28] "CLL_GSM330979.CEL" "CLL_GSM330980.CEL" "CLL_GSM330982.CEL"
```

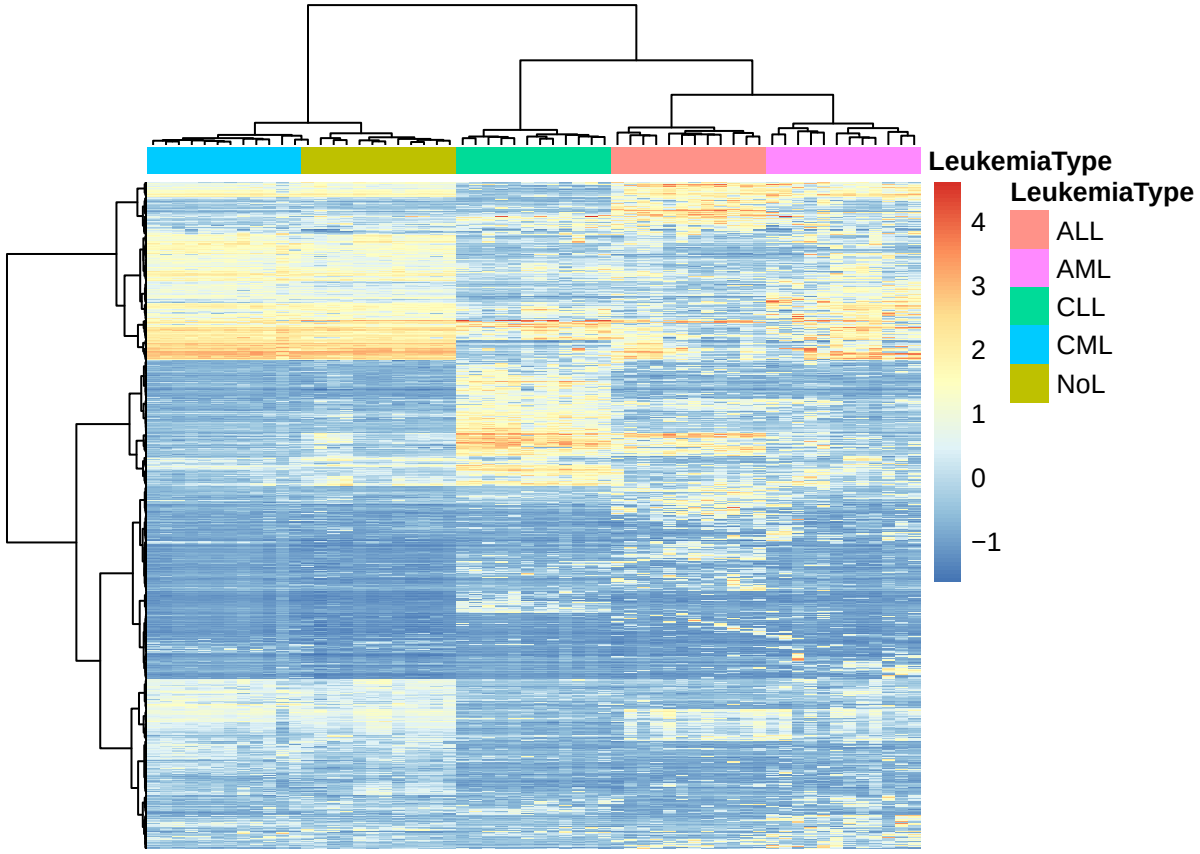
```
## [31] "CLL_GSM330987.CEL" "CLL_GSM330999.CEL" "CLL_GSM331004.CEL"
## [34] "CLL_GSM331009.CEL" "CLL_GSM331037.CEL" "CLL_GSM331048.CEL"
## [37] "CML_GSM331377.CEL" "CML_GSM331378.CEL" "CML_GSM331381.CEL"
## [40] "CML_GSM331382.CEL" "CML_GSM331383.CEL" "CML_GSM331386.CEL"
## [43] "CML_GSM331387.CEL" "CML_GSM331388.CEL" "CML_GSM331389.CEL"
## [46] "CML_GSM331390.CEL" "CML_GSM331392.CEL" "CML_GSM331393.CEL"
## [49] "NoL_GSM331660.CEL" "NoL_GSM331661.CEL" "NoL_GSM331663.CEL"
## [52] "NoL_GSM331666.CEL" "NoL_GSM331668.CEL" "NoL_GSM331670.CEL"
## [55] "NoL_GSM331671.CEL" "NoL_GSM331672.CEL" "NoL_GSM331673.CEL"
## [58] "NoL_GSM331674.CEL" "NoL_GSM331675.CEL" "NoL_GSM331677.CEL"
```

```
# set the leukemia type annotation for each sample
annotation_col = data.frame(
                    LeukemiaType =substr(colnames(mat),1,3))
rownames(annotation_col)=colnames(mat)

# Without any scaling
pheatmap(mat,show_rownames=FALSE,show_colnames=FALSE,
         annotation_col=annotation_col,
         scale = "none",clustering_method="ward.D",
         clustering_distance_cols="euclidean")
```

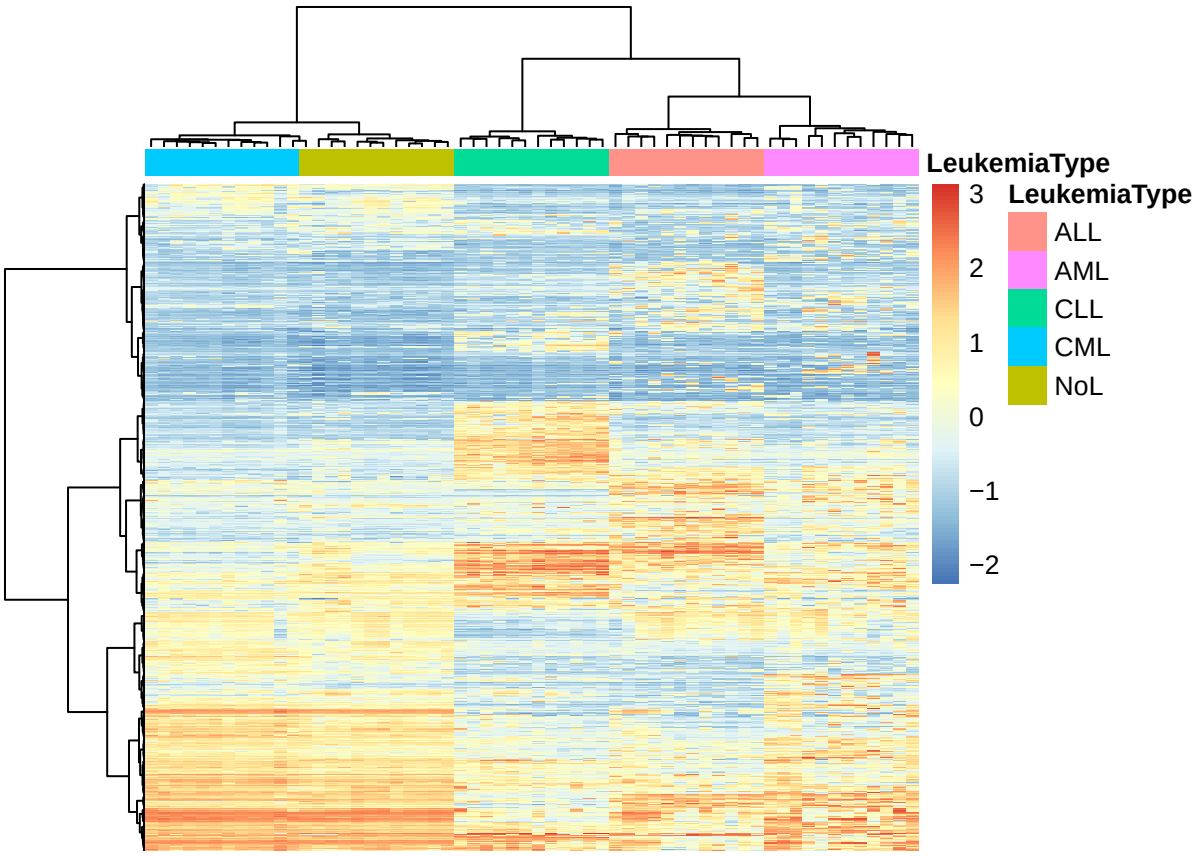



```
# Heatmap after Scaling
pheatmap(scaled_exp,show_rownames=FALSE,show_colnames=FALSE,
         annotation_col=annotation_col,
```

```
          scale = "none",clustering_method="ward.D",
          clustering_distance_cols="euclidean")
```




```
# Heatmap after log2 and scaling
pheatmap(scaled_exp2,show_rownames=FALSE,show_colnames=FALSE,
          annotation_col=annotation_col,
          scale = "none",clustering_method="ward.D",
          clustering_distance_cols="euclidean")
```



Non-scaled heatmap showed the poor clustering compared to others since we can see CML and NoL are mixed. Also, log2 and scaled heatmap showed more better differential gene expression profiles specifically among ALL, AML, CLL groups (more distinct red colors of different profiles in each group).

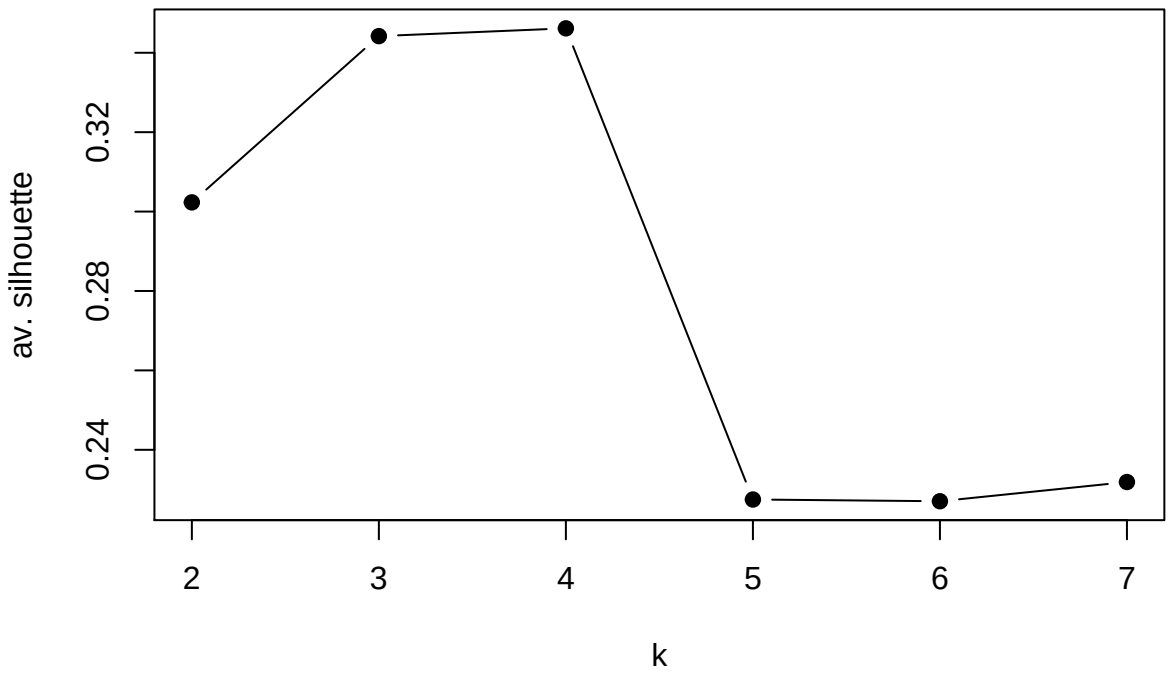
3. For the transformed and untransformed data sets used in the exercise above, use the silhouette for deciding number of clusters using hierarchical clustering. [Difficulty: **Intermediate/Advanced**]

solution:

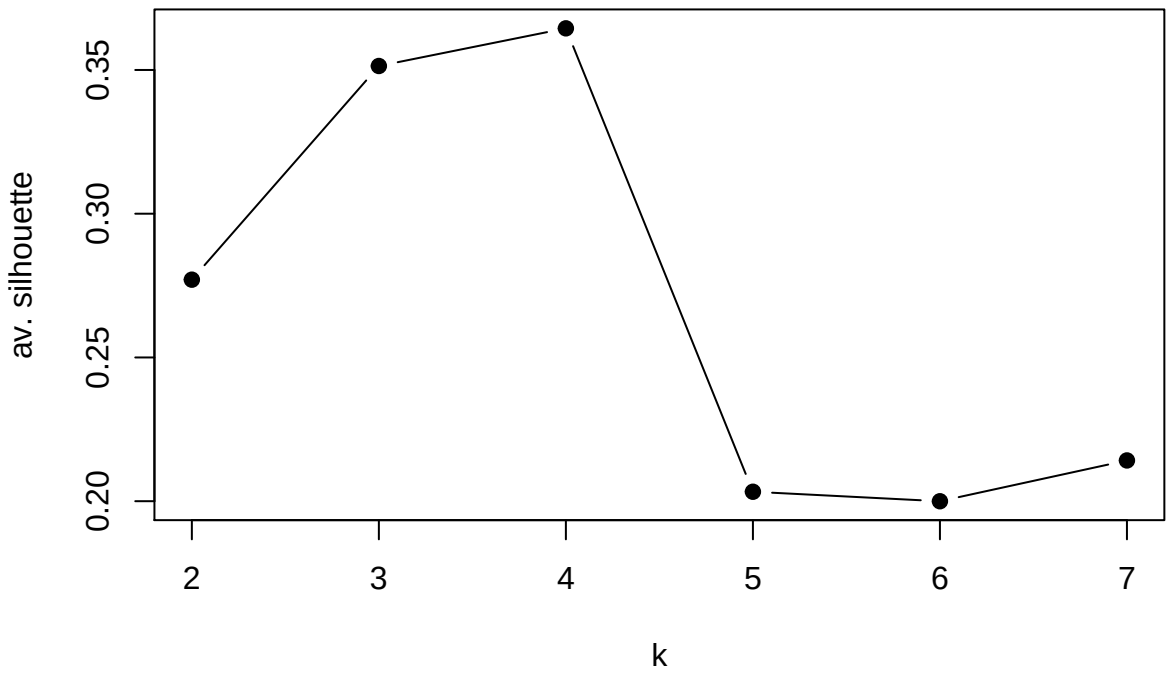
```
library(cluster)
```

```
## Warning: package 'cluster' was built under R version 4.0.5
```

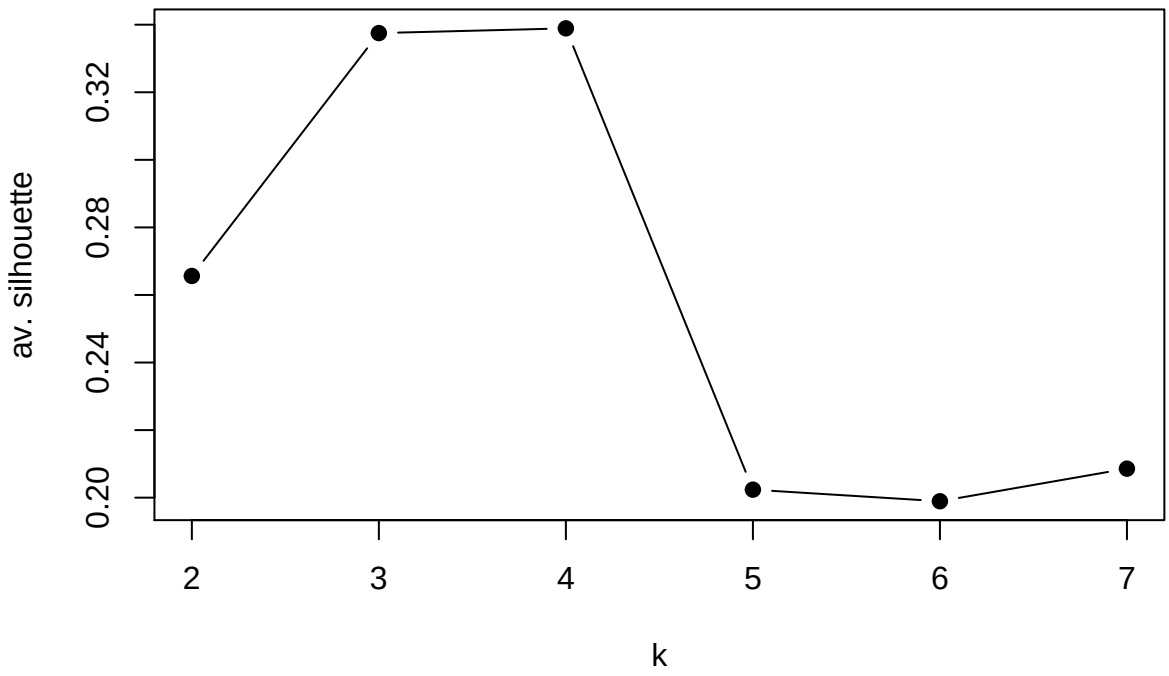
```
set.seed(101)

# calculate the average silhouette value for different k-values in untransformed data.
Ks=sapply(2:7,
    function(i)
      summary(silhouette(pam(t(mat),k=i)))$avg.width)
plot(2:7,Ks,xlab="k",ylab="av. silhouette",type="b",
     pch=19)
```




```
# calculate the average silhouette value for different k-values in scaled data.
Ks=sapply(2:7,
    function(i)
      summary(silhouette(pam(t(scaled_exp),k=i)))$avg.width)
plot(2:7,Ks,xlab="k",ylab="av. silhouette",type="b",
     pch=19)
```



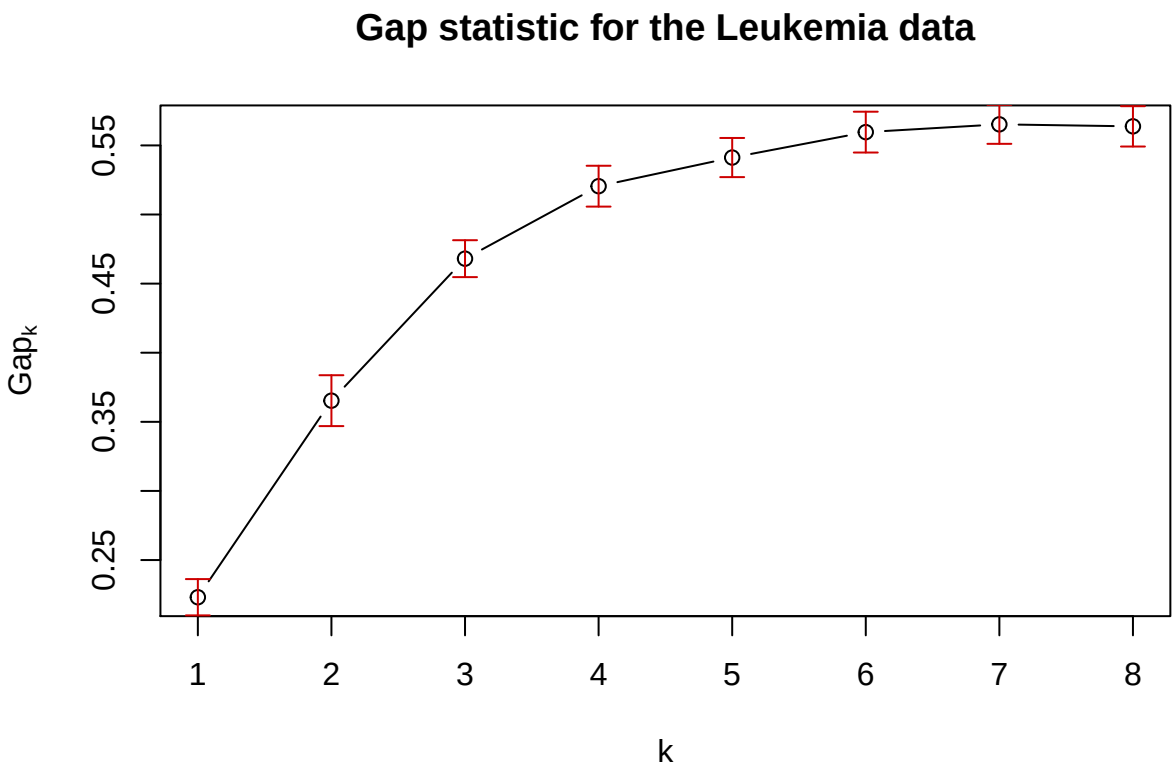

```
# calculate the average silhouette value for different k-values in log2/scaled transformed data.
Ks=sapply(2:7,
    function(i)
      summary(silhouette(pam(t(scaled_exp2),k=i)))$avg.width)
plot(2:7,Ks,xlab="k",ylab="av. silhouette",type="b",
     pch=19)
```



It seems that k = 4 is the best value for clustering.

4. Now, use the Gap Statistic for deciding the number of clusters in hierarchical clustering. Is the same number of clusters identified by two methods? Is it similar to the number of clusters obtained using the k-means algorithm in the unsupervised learning chapter. [Difficulty: **Intermediate/Advanced**]

```
library(cluster)
set.seed(101)
# define the clustering function
pam1 <- function(x,k)
  list(cluster = pam(x,k, cluster.only=TRUE))

# calculate the gap statistic
pam.gap= clusGap(t(mat), FUN = pam1, K.max = 8,B=50)

# plot the gap statistic across k values
plot(pam.gap, main = "Gap statistic for the Leukemia data")
```



Gap statistic method shows that k = 7 is the best but if we consider the error bars in the figure, k = 6 will be the lowest optimal number of clusters. Previous Silhouette method gave us k = 4 as an optimal number so they are not the same numbers. But our biological group number is 5 so we can say our k = 4 or k = 6 is close to the biological value. Also, it may show the possibility of "not yet found" biological subgroups.

Dimension reduction

We will be using the leukemia expression data set again. You can use it as shown in the clustering exercises.

1. Do PCA on the expression matrix using the `princomp()` function and then use the `screeplot()` function to visualize the explained variation by eigenvectors. How many top components explain 95% of the variation? [Difficulty: **Beginner**]

solution:

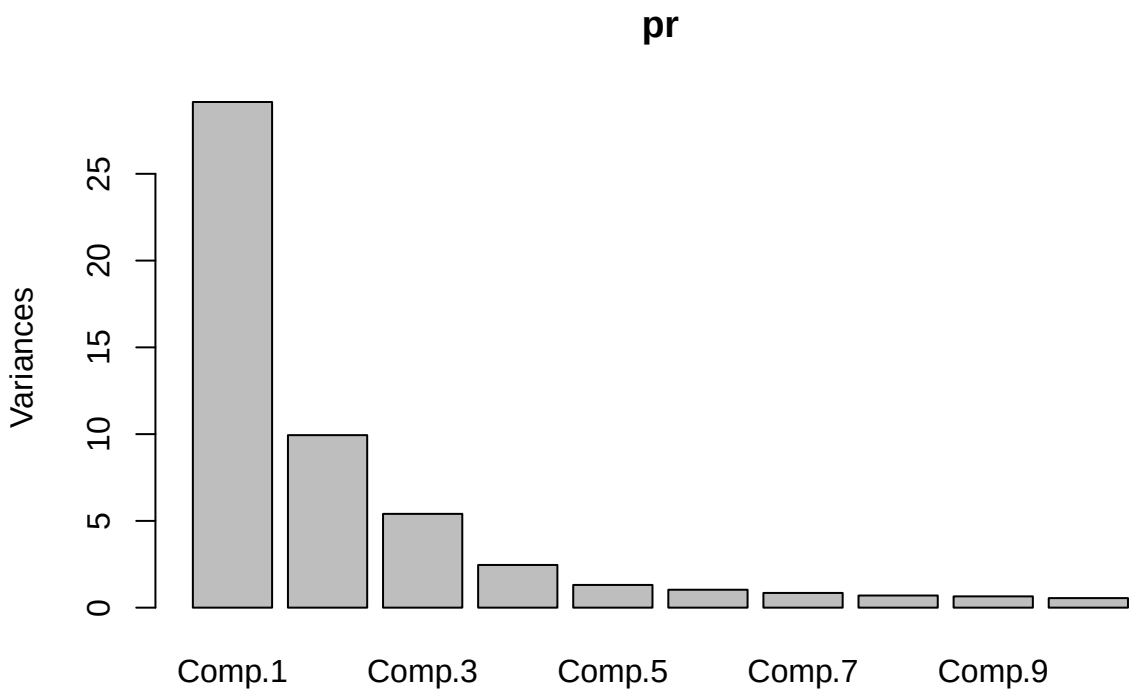
```
# Calculate the PCA with scaled data
pr= princomp(scale(mat))
summary(pr)
```

```
## Importance of components:
##                            Comp.1    Comp.2     Comp.3     Comp.4     Comp.5
## Standard deviation     5.3980464 3.1529029 2.32472774 1.56870823 1.14416985
## Proportion of Variance 0.4861346 0.1658458 0.09016281 0.04105515 0.02184058
## Cumulative Proportion  0.4861346 0.6519803 0.74214316 0.78319830 0.80503889
##                            Comp.6     Comp.7     Comp.8     Comp.9     Comp.10
```

```
## Standard deviation     1.0163041 0.92085568 0.83618790 0.80637087 0.741512765
## Proportion of Variance 0.0172318 0.01414707 0.01166517 0.01084808 0.009173193
## Cumulative Proportion  0.8222707 0.83641776 0.84808292 0.85893101 0.868104199
##                            Comp.11     Comp.12    Comp.13     Comp.14
## Standard deviation     0.734978612 0.684352550 0.65828459 0.636403547
## Proportion of Variance 0.009012238 0.007813454 0.00722954 0.006756915
## Cumulative Proportion  0.877116437 0.884929891 0.89215943 0.898916345
##                            Comp.15     Comp.16     Comp.17    Comp.18
## Standard deviation     0.626302672 0.615372546 0.600370341 0.56906486
## Proportion of Variance 0.006544128 0.006317707 0.006013423 0.00540265
## Cumulative Proportion  0.905460473 0.911778180 0.917791603 0.92319425
##                            Comp.19     Comp.20     Comp.21     Comp.22
## Standard deviation     0.549322382 0.524753964 0.509814376 0.489873632
## Proportion of Variance 0.005034286 0.004594039 0.004336181 0.004003607
## Cumulative Proportion  0.928228538 0.932822578 0.937158759 0.941162365
##                            Comp.23     Comp.24     Comp.25    Comp.26
## Standard deviation     0.482433327 0.473364327 0.462804890 0.44960365
## Proportion of Variance 0.003882915 0.003738301 0.003573379 0.00337243
## Cumulative Proportion  0.945045280 0.948783581 0.952356961 0.95572939
##                            Comp.27     Comp.28     Comp.29    Comp.30
## Standard deviation     0.439205787 0.424763283 0.421822014 0.39135109
## Proportion of Variance 0.003218247 0.003010074 0.002968532 0.00255515
## Cumulative Proportion  0.958947638 0.961957712 0.964926244 0.96748139
##                            Comp.31     Comp.32     Comp.33     Comp.34
## Standard deviation     0.382994366 0.373072389 0.364965083 0.355318345
## Proportion of Variance 0.002447192 0.002322039 0.002222214 0.002106292
## Cumulative Proportion  0.969928586 0.972250625 0.974472839 0.976579130
##                            Comp.35     Comp.36     Comp.37     Comp.38
## Standard deviation     0.338241345 0.326528505 0.322846728 0.316324586
## Proportion of Variance 0.001908695 0.001778793 0.001738906 0.001669357
## Cumulative Proportion  0.978487826 0.980266619 0.982005525 0.983674882
##                            Comp.39     Comp.40     Comp.41     Comp.42
## Standard deviation     0.308522560 0.296659654 0.287414886 0.279974690
## Proportion of Variance 0.001588024 0.001468251 0.001378167 0.001307738
## Cumulative Proportion  0.985262906 0.986731156 0.988109323 0.989417061
##                            Comp.43     Comp.44      Comp.45      Comp.46
## Standard deviation     0.250015853 0.245318258 0.2348965297 0.2282002770
## Proportion of Variance 0.001042842 0.001004021 0.0009205269 0.0008687916
## Cumulative Proportion  0.990459903 0.991463925 0.9923844514 0.9932532430
##                             Comp.47      Comp.48      Comp.49      Comp.50
## Standard deviation     0.2246585841 0.2127630877 0.2052917511 0.1851652001
## Proportion of Variance 0.0008420334 0.0007552241 0.0007031148 0.0005720079
## Cumulative Proportion  0.9940952763 0.9948505004 0.9955536152 0.9961256231
##                            Comp.51      Comp.52      Comp.53      Comp.54
## Standard deviation     0.1791512892 0.1739143228 0.1657289040 0.1652361222
## Proportion of Variance 0.0005354552 0.0005046078 0.0004582261 0.0004555051
## Cumulative Proportion  0.9966610783 0.9971656861 0.9976239121 0.9980794173
##                            Comp.55      Comp.56      Comp.57      Comp.58
## Standard deviation     0.1570079504 0.1543267078 0.1381601383 0.1305610759
## Proportion of Variance 0.0004112695 0.0003973429 0.0003184555 0.0002843876
## Cumulative Proportion  0.9984906868 0.9988880297 0.9992064852 0.9994908728
##                            Comp.59      Comp.60
## Standard deviation     0.1283559255 0.1184982639
## Proportion of Variance 0.0002748623 0.0002342649
```

```
## Cumulative Proportion  0.9997657351 1.0000000000
```

```
# Visualize the explained variation by eigenvectors
screeplot(pr)
```



Comp25 has 0.952356961 of cumulative proportion. So, top 25 components explain 95% of the variation.

2. Our next tasks are removing the eigenvectors and reconstructing the matrix using SVD, then we need to calculate the reconstruction error as the difference between the original and the reconstructed matrix. HINT: You have to use the `svd()` function and equalize eigenvalue to 0 for the component you want to remove. [Difficulty: **Intermediate/Advanced**]

solution:

```
d=svd(scale(mat)) # apply SVD

# Reconstructing the matrix using SVD
new_comp <- append(d$d[1:25], c(rep(0, 35)))
diag(new_comp)
```

```
##           [,1]     [,2]     [,3]     [,4]     [,5]     [,6]     [,7]     [,8]
##  [1,] 170.7012  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [2,]   0.0000 99.70354  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [3,]   0.0000  0.00000 73.51435  0.00000  0.00000  0.00000  0.00000  0.00000
##  [4,]   0.0000  0.00000  0.00000 49.60691  0.00000  0.00000  0.00000  0.00000
```

```
##  [5,]   0.0000  0.00000  0.00000  0.00000 36.18183  0.00000  0.00000  0.00000
##  [6,]   0.0000  0.00000  0.00000  0.00000  0.00000 32.13836  0.00000  0.00000
##  [7,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000 29.12001  0.00000
##  [8,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000 26.44258
##  [9,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [10,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [11,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [12,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [13,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [14,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [15,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [16,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [17,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [18,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [19,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [20,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [21,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [22,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [23,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [24,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [25,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [26,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [27,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [28,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [29,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [30,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [31,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [32,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [33,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [34,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [35,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [36,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [37,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [38,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [39,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [40,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [41,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [42,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [43,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [44,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [45,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [46,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [47,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [48,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [49,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [50,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [51,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [52,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [53,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [54,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [55,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [56,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [57,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [58,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
```

```
## [59,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [60,]   0.0000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##           [,9]    [,10]    [,11]    [,12]    [,13]    [,14]    [,15]    [,16]
##  [1,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [2,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [3,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [4,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [5,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [6,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [7,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [8,] 0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [9,] 25.49969  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [10,]  0.00000 23.44869  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [11,]  0.00000  0.00000 23.24206  0.00000  0.00000  0.00000  0.00000  0.00000
## [12,]  0.00000  0.00000  0.00000 21.64113  0.00000  0.00000  0.00000  0.00000
## [13,]  0.00000  0.00000  0.00000  0.00000 20.81679  0.00000  0.00000  0.00000
## [14,]  0.00000  0.00000  0.00000  0.00000  0.00000 20.12485  0.00000  0.00000
## [15,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000 19.80543  0.00000
## [16,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000 19.45979
## [17,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [18,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [19,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [20,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [21,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [22,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [23,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [24,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [25,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [26,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [27,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [28,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [29,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [30,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [31,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [32,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [33,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [34,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [35,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [36,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [37,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [38,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [39,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [40,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [41,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [42,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [43,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [44,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [45,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [46,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [47,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [48,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [49,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [50,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [51,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
```

```
## [52,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [53,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [54,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [55,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [56,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [57,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [58,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [59,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
## [60,]  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000  0.00000
##          [,17]    [,18]   [,19]    [,20]    [,21]    [,22]    [,23]    [,24]
##  [1,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [2,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [3,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [4,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [5,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [6,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [7,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [8,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##  [9,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [10,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [11,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [12,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [13,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [14,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [15,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [16,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [17,] 18.98538  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [18,]  0.00000 17.99541  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [19,]  0.00000  0.00000 17.3711  0.00000  0.00000  0.00000  0.00000  0.00000
## [20,]  0.00000  0.00000  0.0000 16.59418  0.00000  0.00000  0.00000  0.00000
## [21,]  0.00000  0.00000  0.0000  0.00000 16.12175  0.00000  0.00000  0.00000
## [22,]  0.00000  0.00000  0.0000  0.00000  0.00000 15.49116  0.00000  0.00000
## [23,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000 15.25588  0.00000
## [24,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000 14.96909
## [25,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [26,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [27,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [28,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [29,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [30,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [31,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [32,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [33,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [34,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [35,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [36,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [37,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [38,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [39,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [40,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [41,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [42,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [43,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [44,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
```

```
## [45,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [46,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [47,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [48,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [49,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [50,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [51,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [52,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [53,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [54,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [55,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [56,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [57,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [58,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [59,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
## [60,]  0.00000  0.00000  0.0000  0.00000  0.00000  0.00000  0.00000  0.00000
##          [,25] [,26] [,27] [,28] [,29] [,30] [,31] [,32] [,33] [,34] [,35]
##  [1,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [2,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [3,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [4,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [5,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [6,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [7,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [8,]  0.00000     0     0     0     0     0     0     0     0     0     0
##  [9,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [10,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [11,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [12,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [13,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [14,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [15,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [16,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [17,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [18,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [19,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [20,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [21,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [22,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [23,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [24,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [25,] 14.63518     0     0     0     0     0     0     0     0     0     0
## [26,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [27,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [28,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [29,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [30,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [31,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [32,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [33,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [34,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [35,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [36,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [37,]  0.00000     0     0     0     0     0     0     0     0     0     0
```

```
## [38,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [39,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [40,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [41,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [42,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [43,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [44,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [45,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [46,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [47,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [48,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [49,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [50,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [51,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [52,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [53,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [54,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [55,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [56,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [57,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [58,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [59,]  0.00000     0     0     0     0     0     0     0     0     0     0
## [60,]  0.00000     0     0     0     0     0     0     0     0     0     0
##       [,36] [,37] [,38] [,39] [,40] [,41] [,42] [,43] [,44] [,45] [,46] [,47]
##  [1,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [2,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [3,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [4,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [5,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [6,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [7,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [8,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [9,]     0     0     0     0     0     0     0     0     0     0     0     0
## [10,]     0     0     0     0     0     0     0     0     0     0     0     0
## [11,]     0     0     0     0     0     0     0     0     0     0     0     0
## [12,]     0     0     0     0     0     0     0     0     0     0     0     0
## [13,]     0     0     0     0     0     0     0     0     0     0     0     0
## [14,]     0     0     0     0     0     0     0     0     0     0     0     0
## [15,]     0     0     0     0     0     0     0     0     0     0     0     0
## [16,]     0     0     0     0     0     0     0     0     0     0     0     0
## [17,]     0     0     0     0     0     0     0     0     0     0     0     0
## [18,]     0     0     0     0     0     0     0     0     0     0     0     0
## [19,]     0     0     0     0     0     0     0     0     0     0     0     0
## [20,]     0     0     0     0     0     0     0     0     0     0     0     0
## [21,]     0     0     0     0     0     0     0     0     0     0     0     0
## [22,]     0     0     0     0     0     0     0     0     0     0     0     0
## [23,]     0     0     0     0     0     0     0     0     0     0     0     0
## [24,]     0     0     0     0     0     0     0     0     0     0     0     0
## [25,]     0     0     0     0     0     0     0     0     0     0     0     0
## [26,]     0     0     0     0     0     0     0     0     0     0     0     0
## [27,]     0     0     0     0     0     0     0     0     0     0     0     0
## [28,]     0     0     0     0     0     0     0     0     0     0     0     0
## [29,]     0     0     0     0     0     0     0     0     0     0     0     0
## [30,]     0     0     0     0     0     0     0     0     0     0     0     0
```

```
## [31,]     0     0     0     0     0     0     0     0     0     0     0     0
## [32,]     0     0     0     0     0     0     0     0     0     0     0     0
## [33,]     0     0     0     0     0     0     0     0     0     0     0     0
## [34,]     0     0     0     0     0     0     0     0     0     0     0     0
## [35,]     0     0     0     0     0     0     0     0     0     0     0     0
## [36,]     0     0     0     0     0     0     0     0     0     0     0     0
## [37,]     0     0     0     0     0     0     0     0     0     0     0     0
## [38,]     0     0     0     0     0     0     0     0     0     0     0     0
## [39,]     0     0     0     0     0     0     0     0     0     0     0     0
## [40,]     0     0     0     0     0     0     0     0     0     0     0     0
## [41,]     0     0     0     0     0     0     0     0     0     0     0     0
## [42,]     0     0     0     0     0     0     0     0     0     0     0     0
## [43,]     0     0     0     0     0     0     0     0     0     0     0     0
## [44,]     0     0     0     0     0     0     0     0     0     0     0     0
## [45,]     0     0     0     0     0     0     0     0     0     0     0     0
## [46,]     0     0     0     0     0     0     0     0     0     0     0     0
## [47,]     0     0     0     0     0     0     0     0     0     0     0     0
## [48,]     0     0     0     0     0     0     0     0     0     0     0     0
## [49,]     0     0     0     0     0     0     0     0     0     0     0     0
## [50,]     0     0     0     0     0     0     0     0     0     0     0     0
## [51,]     0     0     0     0     0     0     0     0     0     0     0     0
## [52,]     0     0     0     0     0     0     0     0     0     0     0     0
## [53,]     0     0     0     0     0     0     0     0     0     0     0     0
## [54,]     0     0     0     0     0     0     0     0     0     0     0     0
## [55,]     0     0     0     0     0     0     0     0     0     0     0     0
## [56,]     0     0     0     0     0     0     0     0     0     0     0     0
## [57,]     0     0     0     0     0     0     0     0     0     0     0     0
## [58,]     0     0     0     0     0     0     0     0     0     0     0     0
## [59,]     0     0     0     0     0     0     0     0     0     0     0     0
## [60,]     0     0     0     0     0     0     0     0     0     0     0     0
##       [,48] [,49] [,50] [,51] [,52] [,53] [,54] [,55] [,56] [,57] [,58] [,59]
##  [1,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [2,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [3,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [4,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [5,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [6,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [7,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [8,]     0     0     0     0     0     0     0     0     0     0     0     0
##  [9,]     0     0     0     0     0     0     0     0     0     0     0     0
## [10,]     0     0     0     0     0     0     0     0     0     0     0     0
## [11,]     0     0     0     0     0     0     0     0     0     0     0     0
## [12,]     0     0     0     0     0     0     0     0     0     0     0     0
## [13,]     0     0     0     0     0     0     0     0     0     0     0     0
## [14,]     0     0     0     0     0     0     0     0     0     0     0     0
## [15,]     0     0     0     0     0     0     0     0     0     0     0     0
## [16,]     0     0     0     0     0     0     0     0     0     0     0     0
## [17,]     0     0     0     0     0     0     0     0     0     0     0     0
## [18,]     0     0     0     0     0     0     0     0     0     0     0     0
## [19,]     0     0     0     0     0     0     0     0     0     0     0     0
## [20,]     0     0     0     0     0     0     0     0     0     0     0     0
## [21,]     0     0     0     0     0     0     0     0     0     0     0     0
## [22,]     0     0     0     0     0     0     0     0     0     0     0     0
## [23,]     0     0     0     0     0     0     0     0     0     0     0     0
```

```
## [24,]     0     0     0     0     0     0     0     0     0     0     0     0
## [25,]     0     0     0     0     0     0     0     0     0     0     0     0
## [26,]     0     0     0     0     0     0     0     0     0     0     0     0
## [27,]     0     0     0     0     0     0     0     0     0     0     0     0
## [28,]     0     0     0     0     0     0     0     0     0     0     0     0
## [29,]     0     0     0     0     0     0     0     0     0     0     0     0
## [30,]     0     0     0     0     0     0     0     0     0     0     0     0
## [31,]     0     0     0     0     0     0     0     0     0     0     0     0
## [32,]     0     0     0     0     0     0     0     0     0     0     0     0
## [33,]     0     0     0     0     0     0     0     0     0     0     0     0
## [34,]     0     0     0     0     0     0     0     0     0     0     0     0
## [35,]     0     0     0     0     0     0     0     0     0     0     0     0
## [36,]     0     0     0     0     0     0     0     0     0     0     0     0
## [37,]     0     0     0     0     0     0     0     0     0     0     0     0
## [38,]     0     0     0     0     0     0     0     0     0     0     0     0
## [39,]     0     0     0     0     0     0     0     0     0     0     0     0
## [40,]     0     0     0     0     0     0     0     0     0     0     0     0
## [41,]     0     0     0     0     0     0     0     0     0     0     0     0
## [42,]     0     0     0     0     0     0     0     0     0     0     0     0
## [43,]     0     0     0     0     0     0     0     0     0     0     0     0
## [44,]     0     0     0     0     0     0     0     0     0     0     0     0
## [45,]     0     0     0     0     0     0     0     0     0     0     0     0
## [46,]     0     0     0     0     0     0     0     0     0     0     0     0
## [47,]     0     0     0     0     0     0     0     0     0     0     0     0
## [48,]     0     0     0     0     0     0     0     0     0     0     0     0
## [49,]     0     0     0     0     0     0     0     0     0     0     0     0
## [50,]     0     0     0     0     0     0     0     0     0     0     0     0
## [51,]     0     0     0     0     0     0     0     0     0     0     0     0
## [52,]     0     0     0     0     0     0     0     0     0     0     0     0
## [53,]     0     0     0     0     0     0     0     0     0     0     0     0
## [54,]     0     0     0     0     0     0     0     0     0     0     0     0
## [55,]     0     0     0     0     0     0     0     0     0     0     0     0
## [56,]     0     0     0     0     0     0     0     0     0     0     0     0
## [57,]     0     0     0     0     0     0     0     0     0     0     0     0
## [58,]     0     0     0     0     0     0     0     0     0     0     0     0
## [59,]     0     0     0     0     0     0     0     0     0     0     0     0
## [60,]     0     0     0     0     0     0     0     0     0     0     0     0
##       [,60]
##  [1,]     0
##  [2,]     0
##  [3,]     0
##  [4,]     0
##  [5,]     0
##  [6,]     0
##  [7,]     0
##  [8,]     0
##  [9,]     0
## [10,]     0
## [11,]     0
## [12,]     0
## [13,]     0
## [14,]     0
## [15,]     0
## [16,]     0
```

```
## [17,]     0
## [18,]     0
## [19,]     0
## [20,]     0
## [21,]     0
## [22,]     0
## [23,]     0
## [24,]     0
## [25,]     0
## [26,]     0
## [27,]     0
## [28,]     0
## [29,]     0
## [30,]     0
## [31,]     0
## [32,]     0
## [33,]     0
## [34,]     0
## [35,]     0
## [36,]     0
## [37,]     0
## [38,]     0
## [39,]     0
## [40,]     0
## [41,]     0
## [42,]     0
## [43,]     0
## [44,]     0
## [45,]     0
## [46,]     0
## [47,]     0
## [48,]     0
## [49,]     0
## [50,]     0
## [51,]     0
## [52,]     0
## [53,]     0
## [54,]     0
## [55,]     0
## [56,]     0
## [57,]     0
## [58,]     0
## [59,]     0
## [60,]     0
```

```
mat.re <- d$u %*% diag(new_comp) %*% t(d$v)
colnames(mat.re) <- colnames(mat)
# Calculate the reconstruction error
# We can calculate the root mean squared error calculation

rmse = sqrt(mean((mat - mat.re) ** 2))
rmse
```

```
## [1] 5.793702
```

```r
# Or use the package from Metrics
library(Metrics)
```

```
## Warning: package 'Metrics' was built under R version 4.0.5
```

```r
rmse(mat, mat.re)
```

```
## [1] 5.793702
```

The root mean square error is 5.793702.

3. Produce a 10-component ICA from the expression data set. Remove each component and measure the reconstruction error without that component. Rank the components by decreasing reconstruction-error. [Difficulty: **Advanced**]

solution:

```r
library(fastICA)
```

```
## Warning: package 'fastICA' was built under R version 4.0.5
```

```r
ica.res=fastICA(t(mat),n.comp=10) # apply ICA
re.mat <- ica.res$S %*% ica.res$A
#re.mat
rmse(t(mat), re.mat)
```

```
## [1] 5.880021
```

```r
# Remove component i and calculate the reconstruction error without that component.
Error <- c()
for (i in 1:10){
  re.mat <- ica.res$S[, -i] %*% ica.res$A[-i, ]
  Error[i] <- rmse(t(mat), re.mat)

}
df <- data.frame(comp = c(1:10),
                 err = Error)

order(df$err)
```

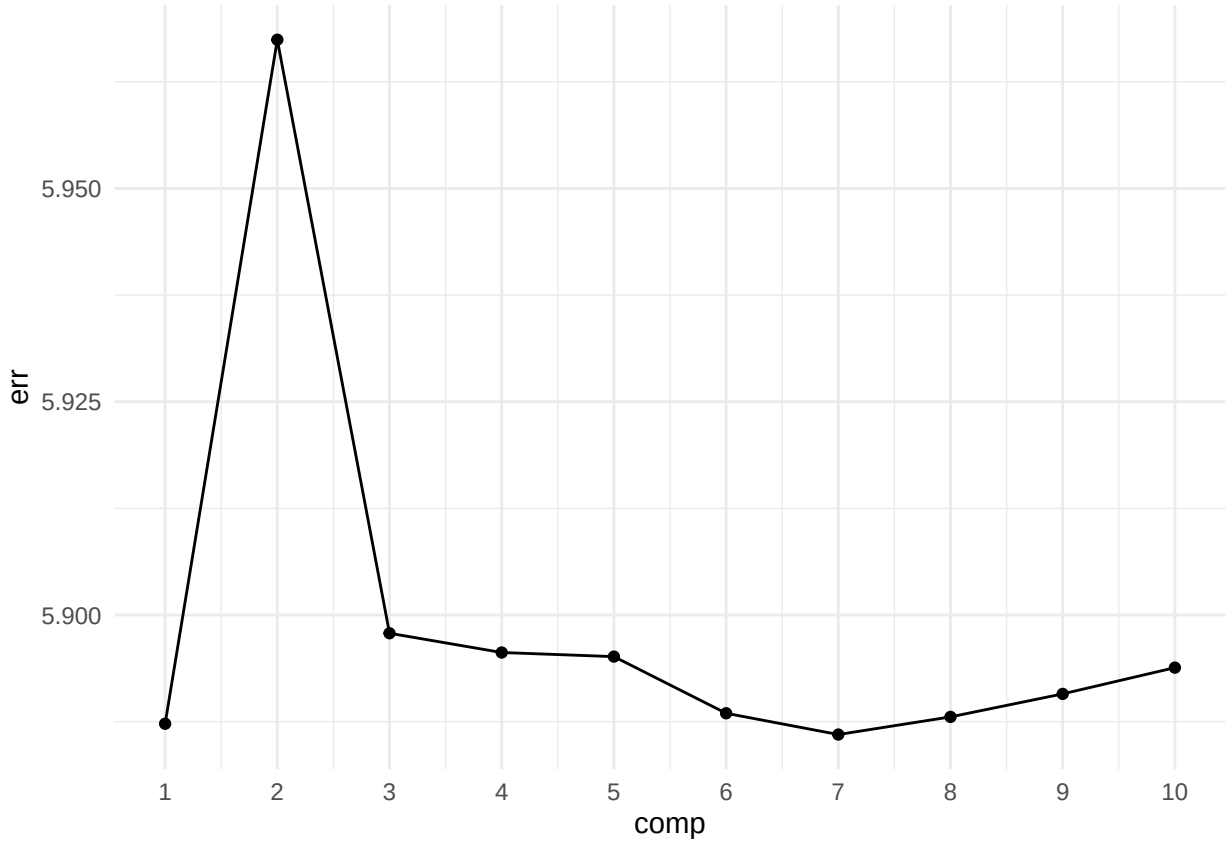
```
##  [1]  7  1  8  6  9 10  5  4  3  2
```

When we order the errors by decreasing reconstruction-error, we get "removing" component from the order is the best to reduce the error. Also, we can easily visualize the results by this simple line graph.

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.0.5
```

```
# graph for the Errors after removing each component
g <- ggplot(data = df, aes(x = comp, y = err)) +
  geom_line() +
  geom_point() +
  theme_minimal()

g + scale_x_continuous(breaks = c(1:10))
```



4. In this exercise we use the `Rtsne()` function on the leukemia expression data set. Try to increase and decrease perplexity t-sne, and describe the observed changes in 2D plots. [Difficulty: **Beginner**]

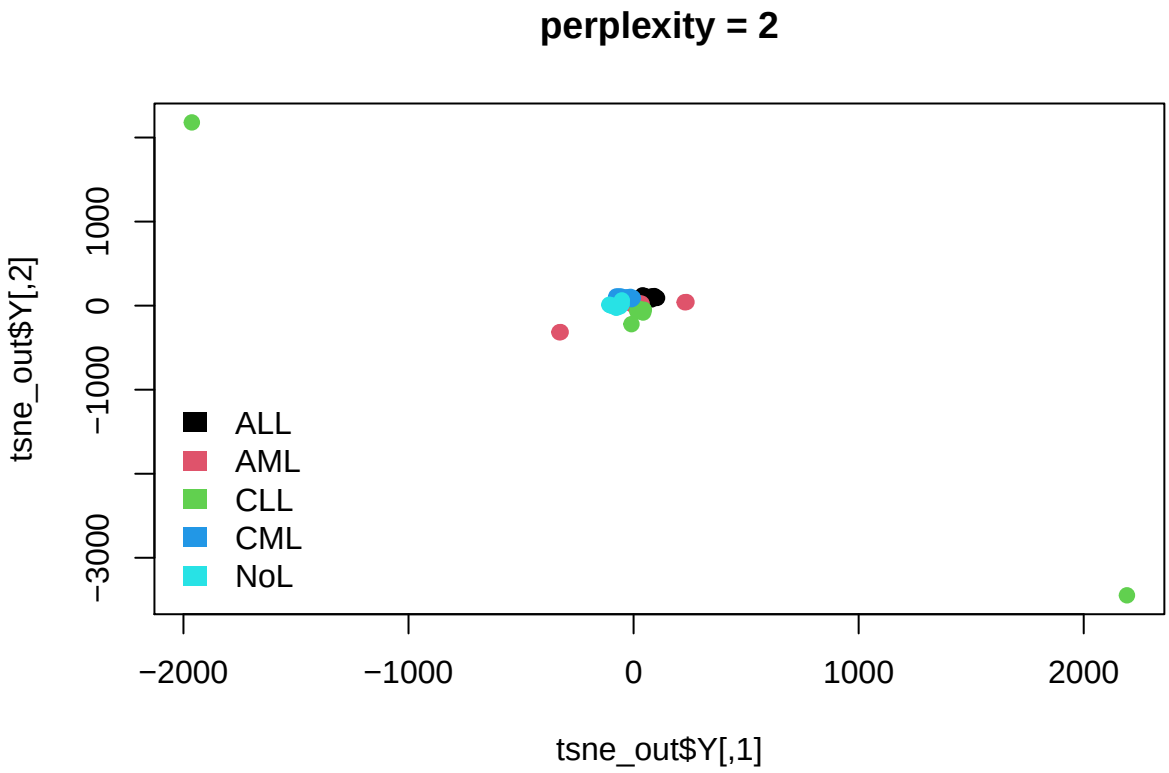
solution: put your text here

```
library("Rtsne")
```

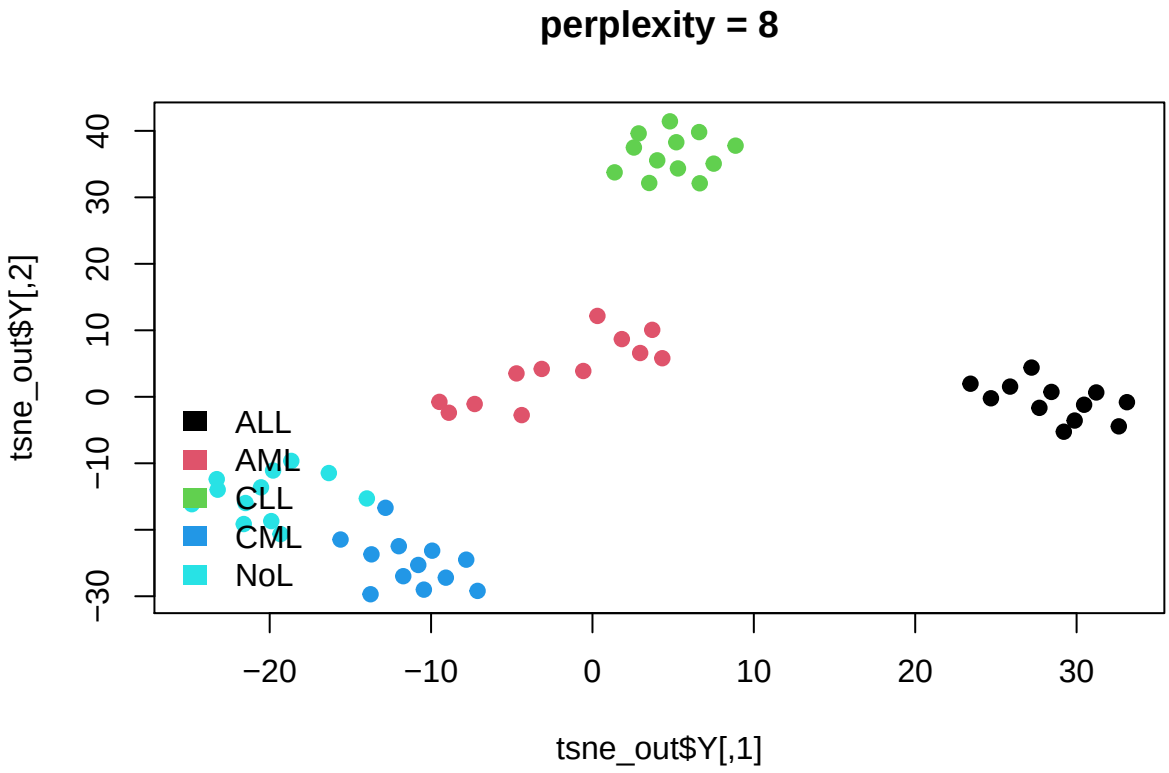
```
## Warning: package 'Rtsne' was built under R version 4.0.5
```

```
# set the leukemia type annotation for each sample
annotation_col = data.frame(
                    LeukemiaType =substr(colnames(mat),1,3))
rownames(annotation_col)=colnames(mat)

tsne_out <- Rtsne(t(mat),perplexity = 2) # Run TSNE
```

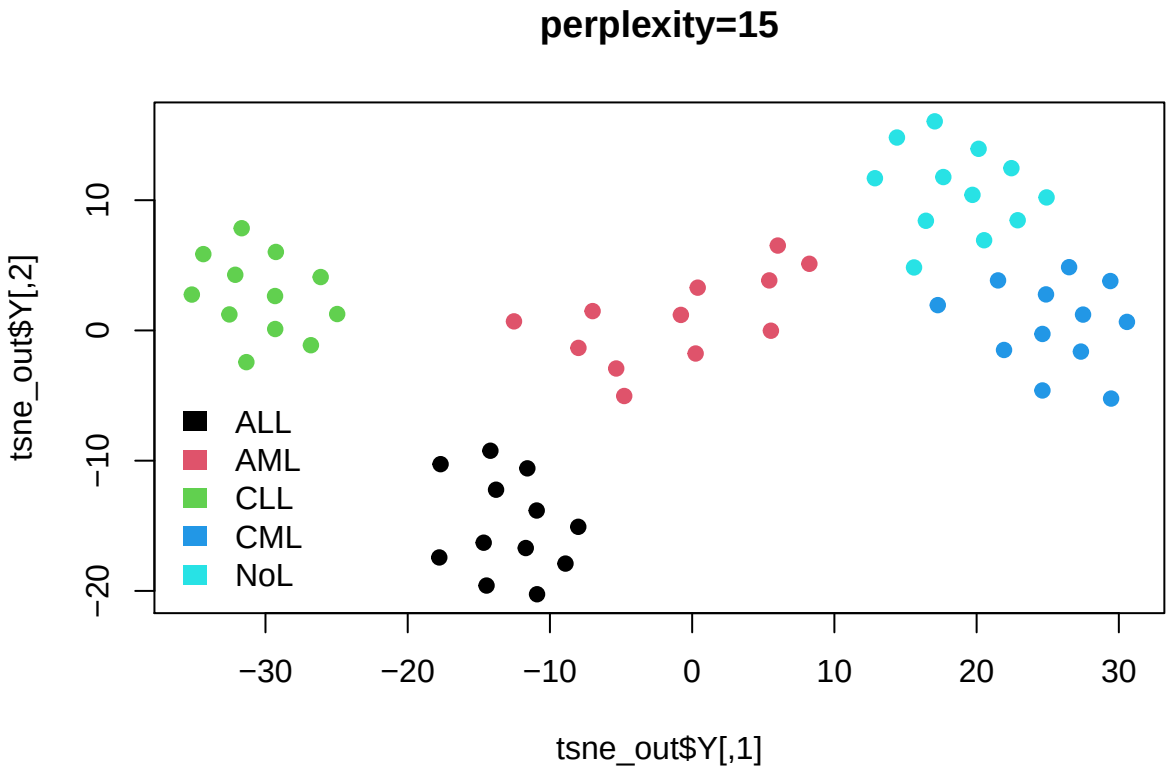
```
# Show the objects in the 2D tsne representation
plot(tsne_out$Y,col=as.factor(annotation_col$LeukemiaType),
     pch=19,  main="perplexity = 2")

# create the legend for the Leukemia types
legend("bottomleft",
       legend=unique(annotation_col$LeukemiaType),
       fill =palette("default"),
       border=NA,box.col=NA)
```




```
# Produce perplexity 8
tsne_out <- Rtsne(t(mat),perplexity = 8) # Run TSNE

# Show the objects in the 2D tsne representation
plot(tsne_out$Y,col=as.factor(annotation_col$LeukemiaType),
     pch=19, main="perplexity = 8")

# create the legend for the Leukemia types
legend("bottomleft",
       legend=unique(annotation_col$LeukemiaType),
       fill =palette("default"),
       border=NA,box.col=NA)
```




```
# Produce perplexity 15 tsne
tsne_out <- Rtsne(t(mat),perplexity = 15) # Run TSNE

# Show the objects in the 2D tsne representation
plot(tsne_out$Y,col=as.factor(annotation_col$LeukemiaType),
     pch=19, main="perplexity=15")

# create the legend for the Leukemia types
legend("bottomleft",
       legend=unique(annotation_col$LeukemiaType),
       fill =palette("default"),
       border=NA,box.col=NA)
```



When we reduced perplexity to 2, the t-SNE plot showed very poor separations among different samples. Although we can see the better separations when we increased perplexity, we don't see that much of improvement after certain numbers of perplexity. (Perplexity=15 was not significantly better than perplexity = 8.)