

Hierarchical Agglomerative Single Linkage Clustering Simulation

MA Mingyu 14110562D

4/5/2017

COMP4433 Assignment 2 Question 3d

derek.ma@connect.polyu.hk derek.ma

Set Up

Import data, delete first column, set initial cluster centers to first two records and set k is equal to 2.

```
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

#PREPARE DATA IN USE
data <- read.csv("~/Google Drive/_DM/2_Assignments/Ass2/data_q3.csv", stringsAsFactors = FALSE)
#delete first column for this specific case because it is not a data record
data <- data[,2:length(colnames(data))]
```

Preprocessing and Normalization

```
#Normalization
min1 <- min(data$B)
max1 <- max(data$B)
min2 <- min(data$C)
max2 <- max(data$C)
min3 <- min(data$D)
max3 <- max(data$D)
min4 <- min(data$E)
max4 <- max(data$E)
min5 <- min(data$F)
max5 <- max(data$F)
min6 <- min(data$G)
max6 <- max(data$G)

data_normalized <- data %>%
```

```

mutate(BN = (B - min1)/(max1-min1)) %>%
mutate(CN = (C - min2)/(max2-min2)) %>%
mutate(DN = (D - min3)/(max3-min3)) %>%
mutate(EN = (E - min4)/(max4-min4)) %>%
mutate(FN = (F - min5)/(max5-min5)) %>%
mutate(GN = (G - min6)/(max6-min6))

data_normalized <- data_normalized[,c(7,8,9,10,11,12)]
colnames(data_normalized) <- c("B", "C", "D", "E", "F", "G")
data <- data_normalized

```

Set Parameters

In this program, a clusters data frame will be used to save which nodes are in which clusters. Initial clusters distribution is saved in `clusters` data frame. After `n` iterations, the clusters distribution information will be saved in `clustersn` data frame.

```

# Create initial proximity matrix
proximityMatrix <- data.frame(matrix(NA, ncol = nrow(data), nrow = nrow(data)))
colnames(proximityMatrix) <- (1:nrow(data))
# Create a data frame to save all clusters
clusters <- data.frame(matrix(NA, ncol = nrow(data), nrow = nrow(data)))
colnames(clusters) <- (1:nrow(data))
colnames(clusters) <- paste("cluster", colnames(clusters), sep = "_")
clusters[1,] <- 1:nrow(data)

```

Basic Function

Distance Function

```

distance <- function(vector1, vector2){
  #Euclidean distance
  #Input: two vectors of data with same length
  #Input example: c(1,2,3); c(2,3,4)
  count <- 0
  for (i in 1:length(vector1)){
    count = count + (vector1[i] - vector2[i])^2
  }
  count^(1/2)
}

```

indexOfMin Function - Find row and column number for the minimum proximity in the proximity matrix

```

indexOfMin <- function(objectsData){
  #set all 0 values to NA
  valueMinInRows <- rep(NA,nrow(objectsData))
  colMinInRows <- rep(NA,nrow(objectsData))

```

```

objectsData[objectsData == 0] <- Inf
objectsData[is.na(objectsData)] <- Inf
for (i in 1:nrow(objectsData)){
  colMinInRows[i] <- which.min(as.numeric(objectsData[i,]))
  valueMinInRows[i] <- min(objectsData[i,])
}
row <- which.min(valueMinInRows)
col <- colMinInRows[row]
c(row,col)
}

```

minOfMatrix Function - return the minimum value except 0 and NA in the matrix

```

minOfMatrix <- function(objectsData){
  #set all 0 values to NA
  valueMinInRows <- rep(NA,nrow(objectsData))
  colMinInRows <- rep(NA,nrow(objectsData))
  objectsData[objectsData == 0] <- Inf
  objectsData[is.na(objectsData)] <- Inf
  for (i in 1:nrow(objectsData)){
    colMinInRows[i] <- which.min(as.numeric(objectsData[i,]))
    valueMinInRows[i] <- min(objectsData[i,])
  }
  min(valueMinInRows)
}

```

updateClusters Function - Based on the minimum proximity information, combine clusters

```

updateClusters <- function(index1,index2,clustersData){
  newClustersData <- clustersData
  baseIndex <- sum(!is.na(clustersData[,index2]))
  addIndex <- sum(!is.na(clustersData[,index1]))
  for (i in 1:addIndex){
    newClustersData[baseIndex+i,index2] <- clustersData[i,index1]
  }
  newClustersData[,-index1]
}

```

updateProximity Function - calculate proximity between each pair of clusters and generate proximity matrix

```

updateProximity <- function(clustersData, objectsData, proximityData){
  newProximityData <- data.frame(matrix(NA,nrow=ncol(clustersData),ncol=ncol(clustersData)))
  colnames(newProximityData) <- colnames(clustersData)
  for (irow in 1:ncol(clustersData)){
    for (icol in irow:ncol(clustersData)){

```

```

# For each pair of clusters
currentMinDist <- 0
for (pointClu1 in 1:sum(!is.na(clustersData[,irow]))){
  for (pointClu2 in 1:sum(!is.na(clustersData[,icol]))){
    # For each pair of points
    dist_value <- distance(
      as.numeric(objectsData[clustersData[pointClu1,irow],]),
      as.numeric(objectsData[clustersData[pointClu2,icol],])
    )
    if (dist_value < currentMinDist || currentMinDist == 0){
      currentMinDist <- dist_value
    }
  }
}
# save min distance between these two clusters
newProximityData[irow,icol] <- currentMinDist
}
}
newProximityData
}

```

Iteration

For each iteration, following steps will be implemented: 1 - Calculate proximity between each pair of clusters and update proximityMatrix 2 - Find out where is the minimum proximity in the proximity matrix 3 - Combine clusters and update the clusters table For each iteration, following tabs will be printed to answer the questions: d(1) - updated distance matrix: `proximityMatrixn`(n is the times of iterations) d(2) - partitions produced: `clustersn` can show the partitions for each iteration d(2) - final dendrogram: attached after this program

Iteration Round 1

```

proximityMatrix1 <- updateProximity(clusters,data,proximityMatrix)
combineIndex <- indexOfMin(proximityMatrix1)
minProximity <- minOfMatrix(proximityMatrix1)
clusters1 <- updateClusters(combineIndex[1],combineIndex[2],clusters)
proximityMatrix1

```

##	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7
## 1	0	0.3868302	2.153812	2.0461324	0.9914041	1.2657780	0.8305813
## 2	NA	0.0000000	2.192372	2.0597899	0.9438613	1.2533039	0.7958934
## 3	NA	NA	0.0000000	0.2505083	1.2741056	0.9927235	1.4373687
## 4	NA	NA	NA	0.0000000	1.1380268	0.9048853	1.3097704
## 5	NA	NA	NA	NA	0.0000000	0.3769650	0.2382569
## 6	NA	NA	NA	NA	NA	0.0000000	0.5551944
## 7	NA	NA	NA	NA	NA	NA	0.0000000
## 8	NA	NA	NA	NA	NA	NA	NA
## 9	NA	NA	NA	NA	NA	NA	NA
## 10	NA	NA	NA	NA	NA	NA	NA
## 11	NA	NA	NA	NA	NA	NA	NA
## 12	NA	NA	NA	NA	NA	NA	NA

```
## 13      NA      NA      NA      NA      NA      NA      NA
## 14      NA      NA      NA      NA      NA      NA      NA
## 15      NA      NA      NA      NA      NA      NA      NA
##      cluster_8 cluster_9 cluster_10 cluster_11 cluster_12 cluster_13
## 1  0.41542848 2.1350191 1.8250894 2.0532672 0.2806071 1.7252293
## 2  0.07972749 2.1092177 1.8143344 2.0206805 0.2294484 1.6518821
## 3  2.14388710 0.5944402 0.5235198 0.4521832 2.1258241 0.9033566
## 4  2.00939976 0.3790698 0.3344590 0.2513604 1.9970008 0.6704361
## 5  0.89958244 1.2346499 0.8789843 1.0970056 0.9282836 0.7968364
## 6  1.20693720 1.0798897 0.6959714 0.8851258 1.2300109 0.7818445
## 7  0.76268199 1.4100600 1.0538655 1.2805460 0.7782574 0.9605974
## 8  0.00000000 2.0534697 1.7677223 1.9684629 0.2192357 1.6022205
## 9      NA 0.0000000 0.5130612 0.3344254 2.0447576 0.5997525
## 10     NA      NA 0.0000000 0.2941505 1.7739539 0.4180999
## 11     NA      NA      NA 0.0000000 1.9803900 0.5350671
## 12     NA      NA      NA      NA 0.0000000 1.6248030
## 13     NA      NA      NA      NA      NA 0.0000000
## 14     NA      NA      NA      NA      NA      NA
## 15     NA      NA      NA      NA      NA      NA
##      cluster_14 cluster_15
## 1  1.2805097 2.1415080
## 2  1.2685334 2.1554439
## 3  1.0280121 0.3162543
## 4  0.8630906 0.1783285
## 5  0.4954452 1.2470129
## 6  0.5233692 1.0226351
## 7  0.6290820 1.4277620
## 8  1.2179163 2.1017690
## 9  0.8856337 0.3267749
## 10 0.6928935 0.4634837
## 11 0.8470162 0.3298998
## 12 1.1963091 2.0854497
## 13 0.6267244 0.7405878
## 14 0.0000000 0.9511731
## 15      NA 0.0000000
```

```
minProximity
```

```
## [1] 0.07972749
```

```
clusters1
```

```
##      cluster_1 cluster_3 cluster_4 cluster_5 cluster_6 cluster_7 cluster_8
## 1           1           3           4           5           6           7           8
## 2           NA           NA           NA           NA           NA           NA           2
## 3           NA           NA           NA           NA           NA           NA           NA
## 4           NA           NA           NA           NA           NA           NA           NA
## 5           NA           NA           NA           NA           NA           NA           NA
## 6           NA           NA           NA           NA           NA           NA           NA
## 7           NA           NA           NA           NA           NA           NA           NA
## 8           NA           NA           NA           NA           NA           NA           NA
## 9           NA           NA           NA           NA           NA           NA           NA
## 10          NA           NA           NA           NA           NA           NA           NA
## 11          NA           NA           NA           NA           NA           NA           NA
## 12          NA           NA           NA           NA           NA           NA           NA
## 13          NA           NA           NA           NA           NA           NA           NA
```

```
## 14      NA      NA      NA      NA      NA      NA      NA
## 15      NA      NA      NA      NA      NA      NA      NA
##      cluster_9 cluster_10 cluster_11 cluster_12 cluster_13 cluster_14
## 1         9         10         11         12         13         14
## 2         NA         NA         NA         NA         NA         NA
## 3         NA         NA         NA         NA         NA         NA
## 4         NA         NA         NA         NA         NA         NA
## 5         NA         NA         NA         NA         NA         NA
## 6         NA         NA         NA         NA         NA         NA
## 7         NA         NA         NA         NA         NA         NA
## 8         NA         NA         NA         NA         NA         NA
## 9         NA         NA         NA         NA         NA         NA
## 10        NA         NA         NA         NA         NA         NA
## 11        NA         NA         NA         NA         NA         NA
## 12        NA         NA         NA         NA         NA         NA
## 13        NA         NA         NA         NA         NA         NA
## 14        NA         NA         NA         NA         NA         NA
## 15        NA         NA         NA         NA         NA         NA
##      cluster_15
## 1         15
## 2         NA
## 3         NA
## 4         NA
## 5         NA
## 6         NA
## 7         NA
## 8         NA
## 9         NA
## 10        NA
## 11        NA
## 12        NA
## 13        NA
## 14        NA
## 15        NA
```

Iteration Round 2

```
proximityMatrix2 <- updateProximity(clusters1,data,proximityMatrix1)
combineIndex <- indexOfMin(proximityMatrix2)
minProximity <- minOfMatrix(proximityMatrix2)
clusters2 <- updateClusters(combineIndex[1],combineIndex[2],clusters1)
proximityMatrix2
```

```
##      cluster_1 cluster_3 cluster_4 cluster_5 cluster_6 cluster_7 cluster_8
## 1         0  2.153812  2.0461324  0.9914041  1.2657780  0.8305813  0.3868302
## 2         NA  0.000000  0.2505083  1.2741056  0.9927235  1.4373687  2.1438871
## 3         NA      NA  0.0000000  1.1380268  0.9048853  1.3097704  2.0093998
## 4         NA      NA      NA  0.0000000  0.3769650  0.2382569  0.8995824
## 5         NA      NA      NA      NA  0.0000000  0.5551944  1.2069372
## 6         NA      NA      NA      NA      NA  0.0000000  0.7626820
## 7         NA      NA      NA      NA      NA      NA  0.0000000
## 8         NA      NA      NA      NA      NA      NA      NA
## 9         NA      NA      NA      NA      NA      NA      NA
```

```
## 10      NA      NA      NA      NA      NA      NA      NA
## 11      NA      NA      NA      NA      NA      NA      NA
## 12      NA      NA      NA      NA      NA      NA      NA
## 13      NA      NA      NA      NA      NA      NA      NA
## 14      NA      NA      NA      NA      NA      NA      NA
##      cluster_9 cluster_10 cluster_11 cluster_12 cluster_13 cluster_14
## 1  2.1350191  1.8250894  2.0532672  0.2806071  1.7252293  1.2805097
## 2  0.5944402  0.5235198  0.4521832  2.1258241  0.9033566  1.0280121
## 3  0.3790698  0.3344590  0.2513604  1.9970008  0.6704361  0.8630906
## 4  1.2346499  0.8789843  1.0970056  0.9282836  0.7968364  0.4954452
## 5  1.0798897  0.6959714  0.8851258  1.2300109  0.7818445  0.5233692
## 6  1.4100600  1.0538655  1.2805460  0.7782574  0.9605974  0.6290820
## 7  2.0534697  1.7677223  1.9684629  0.2192357  1.6022205  1.2179163
## 8  0.0000000  0.5130612  0.3344254  2.0447576  0.5997525  0.8856337
## 9      NA      0.0000000  0.2941505  1.7739539  0.4180999  0.6928935
## 10     NA      NA      0.0000000  1.9803900  0.5350671  0.8470162
## 11     NA      NA      NA      0.0000000  1.6248030  1.1963091
## 12     NA      NA      NA      NA      0.0000000  0.6267244
## 13     NA      NA      NA      NA      NA      0.0000000
## 14     NA      NA      NA      NA      NA      NA
##      cluster_15
## 1  2.1415080
## 2  0.3162543
## 3  0.1783285
## 4  1.2470129
## 5  1.0226351
## 6  1.4277620
## 7  2.1017690
## 8  0.3267749
## 9  0.4634837
## 10 0.3298998
## 11 2.0854497
## 12 0.7405878
## 13 0.9511731
## 14 0.0000000
```

```
minProximity
```

```
## [1] 0.1783285
```

```
clusters2
```

```
##      cluster_1 cluster_3 cluster_5 cluster_6 cluster_7 cluster_8 cluster_9
## 1           1           3           5           6           7           8           9
## 2           NA           NA           NA           NA           NA           2           NA
## 3           NA           NA           NA           NA           NA           NA           NA
## 4           NA           NA           NA           NA           NA           NA           NA
## 5           NA           NA           NA           NA           NA           NA           NA
## 6           NA           NA           NA           NA           NA           NA           NA
## 7           NA           NA           NA           NA           NA           NA           NA
## 8           NA           NA           NA           NA           NA           NA           NA
## 9           NA           NA           NA           NA           NA           NA           NA
## 10          NA           NA           NA           NA           NA           NA           NA
## 11          NA           NA           NA           NA           NA           NA           NA
## 12          NA           NA           NA           NA           NA           NA           NA
## 13          NA           NA           NA           NA           NA           NA           NA
```

## 14	NA	NA	NA	NA	NA	NA	NA
## 15	NA	NA	NA	NA	NA	NA	NA
##	cluster_10	cluster_11	cluster_12	cluster_13	cluster_14	cluster_15	
## 1	10	11	12	13	14	15	
## 2	NA	NA	NA	NA	NA	4	
## 3	NA	NA	NA	NA	NA	NA	
## 4	NA	NA	NA	NA	NA	NA	
## 5	NA	NA	NA	NA	NA	NA	
## 6	NA	NA	NA	NA	NA	NA	
## 7	NA	NA	NA	NA	NA	NA	
## 8	NA	NA	NA	NA	NA	NA	
## 9	NA	NA	NA	NA	NA	NA	
## 10	NA	NA	NA	NA	NA	NA	
## 11	NA	NA	NA	NA	NA	NA	
## 12	NA	NA	NA	NA	NA	NA	
## 13	NA	NA	NA	NA	NA	NA	
## 14	NA	NA	NA	NA	NA	NA	
## 15	NA	NA	NA	NA	NA	NA	

Iteration Round 3

```

proximityMatrix3 <- updateProximity(clusters2,data,proximityMatrix2)
combineIndex <- indexOfMin(proximityMatrix3)
minProximity <- minOfMatrix(proximityMatrix3)
clusters3 <- updateClusters(combineIndex[1],combineIndex[2],clusters2)
proximityMatrix3

```

##	cluster_1	cluster_3	cluster_5	cluster_6	cluster_7	cluster_8	cluster_9
## 1	0	2.153812	0.9914041	1.2657780	0.8305813	0.3868302	2.1350191
## 2	NA	0.000000	1.2741056	0.9927235	1.4373687	2.1438871	0.5944402
## 3	NA	NA	0.000000	0.3769650	0.2382569	0.8995824	1.2346499
## 4	NA	NA	NA	0.000000	0.5551944	1.2069372	1.0798897
## 5	NA	NA	NA	NA	0.000000	0.7626820	1.4100600
## 6	NA	NA	NA	NA	NA	0.000000	2.0534697
## 7	NA	NA	NA	NA	NA	NA	0.000000
## 8	NA	NA	NA	NA	NA	NA	NA
## 9	NA	NA	NA	NA	NA	NA	NA
## 10	NA	NA	NA	NA	NA	NA	NA
## 11	NA	NA	NA	NA	NA	NA	NA
## 12	NA	NA	NA	NA	NA	NA	NA
## 13	NA	NA	NA	NA	NA	NA	NA
##	cluster_10	cluster_11	cluster_12	cluster_13	cluster_14	cluster_15	
## 1	1.8250894	2.0532672	0.2806071	1.7252293	1.2805097	2.0461324	
## 2	0.5235198	0.4521832	2.1258241	0.9033566	1.0280121	0.2505083	
## 3	0.8789843	1.0970056	0.9282836	0.7968364	0.4954452	1.1380268	
## 4	0.6959714	0.8851258	1.2300109	0.7818445	0.5233692	0.9048853	
## 5	1.0538655	1.2805460	0.7782574	0.9605974	0.6290820	1.3097704	
## 6	1.7677223	1.9684629	0.2192357	1.6022205	1.2179163	2.0093998	
## 7	0.5130612	0.3344254	2.0447576	0.5997525	0.8856337	0.3267749	
## 8	0.000000	0.2941505	1.7739539	0.4180999	0.6928935	0.3344590	
## 9	NA	0.000000	1.9803900	0.5350671	0.8470162	0.2513604	
## 10	NA	NA	0.000000	1.6248030	1.1963091	1.9970008	
## 11	NA	NA	NA	0.000000	0.6267244	0.6704361	


```
## 12      NA      NA      NA      NA 0.0000000 0.8630906
## 13      NA      NA      NA      NA      NA      NA 0.0000000
```

```
minProximity
```

```
## [1] 0.2192357
```

```
clusters3
```

```
##      cluster_1 cluster_3 cluster_5 cluster_6 cluster_7 cluster_9 cluster_10
## 1           1           3           5           6           7           9          10
## 2           NA          NA          NA          NA          NA          NA          NA
## 3           NA          NA          NA          NA          NA          NA          NA
## 4           NA          NA          NA          NA          NA          NA          NA
## 5           NA          NA          NA          NA          NA          NA          NA
## 6           NA          NA          NA          NA          NA          NA          NA
## 7           NA          NA          NA          NA          NA          NA          NA
## 8           NA          NA          NA          NA          NA          NA          NA
## 9           NA          NA          NA          NA          NA          NA          NA
## 10          NA          NA          NA          NA          NA          NA          NA
## 11          NA          NA          NA          NA          NA          NA          NA
## 12          NA          NA          NA          NA          NA          NA          NA
## 13          NA          NA          NA          NA          NA          NA          NA
## 14          NA          NA          NA          NA          NA          NA          NA
## 15          NA          NA          NA          NA          NA          NA          NA
##      cluster_11 cluster_12 cluster_13 cluster_14 cluster_15
## 1           11          12          13          14          15
## 2           NA           8          NA          NA           4
## 3           NA           2          NA          NA          NA
## 4           NA          NA          NA          NA          NA
## 5           NA          NA          NA          NA          NA
## 6           NA          NA          NA          NA          NA
## 7           NA          NA          NA          NA          NA
## 8           NA          NA          NA          NA          NA
## 9           NA          NA          NA          NA          NA
## 10          NA          NA          NA          NA          NA
## 11          NA          NA          NA          NA          NA
## 12          NA          NA          NA          NA          NA
## 13          NA          NA          NA          NA          NA
## 14          NA          NA          NA          NA          NA
## 15          NA          NA          NA          NA          NA
```

Iteration Round 4

```
proximityMatrix4 <- updateProximity(clusters3,data,proximityMatrix3)
combineIndex <- indexOfMin(proximityMatrix4)
minProximity <- minOfMatrix(proximityMatrix4)
clusters4 <- updateClusters(combineIndex[1],combineIndex[2],clusters3)
proximityMatrix4
```

```
##      cluster_1 cluster_3 cluster_5 cluster_6 cluster_7 cluster_9 cluster_10
## 1           0  2.153812 0.9914041 1.2657780 0.8305813 2.1350191 1.8250894
## 2           NA 0.0000000 1.2741056 0.9927235 1.4373687 0.5944402 0.5235198
## 3           NA      NA 0.0000000 0.3769650 0.2382569 1.2346499 0.8789843
## 4           NA      NA      NA 0.0000000 0.5551944 1.0798897 0.6959714
```

```
## 5      NA      NA      NA      NA 0.0000000 1.4100600 1.0538655
## 6      NA      NA      NA      NA      NA 0.0000000 0.5130612
## 7      NA      NA      NA      NA      NA      NA 0.0000000
## 8      NA      NA      NA      NA      NA      NA      NA
## 9      NA      NA      NA      NA      NA      NA      NA
## 10     NA      NA      NA      NA      NA      NA      NA
## 11     NA      NA      NA      NA      NA      NA      NA
## 12     NA      NA      NA      NA      NA      NA      NA
##      cluster_11 cluster_12 cluster_13 cluster_14 cluster_15
## 1      2.0532672 0.2806071 1.7252293 1.2805097 2.0461324
## 2      0.4521832 2.1258241 0.9033566 1.0280121 0.2505083
## 3      1.0970056 0.8995824 0.7968364 0.4954452 1.1380268
## 4      0.8851258 1.2069372 0.7818445 0.5233692 0.9048853
## 5      1.2805460 0.7626820 0.9605974 0.6290820 1.3097704
## 6      0.3344254 2.0447576 0.5997525 0.8856337 0.3267749
## 7      0.2941505 1.7677223 0.4180999 0.6928935 0.3344590
## 8      0.0000000 1.9684629 0.5350671 0.8470162 0.2513604
## 9              NA 0.0000000 1.6022205 1.1963091 1.9970008
## 10             NA      NA 0.0000000 0.6267244 0.6704361
## 11             NA      NA      NA 0.0000000 0.8630906
## 12             NA      NA      NA      NA 0.0000000
```

```
minProximity
```

```
## [1] 0.2382569
```

```
clusters4
```

```
##      cluster_1 cluster_3 cluster_6 cluster_7 cluster_9 cluster_10 cluster_11
## 1           1           3           6           7           9          10          11
## 2           NA          NA          NA           5          NA          NA          NA
## 3           NA          NA          NA          NA          NA          NA          NA
## 4           NA          NA          NA          NA          NA          NA          NA
## 5           NA          NA          NA          NA          NA          NA          NA
## 6           NA          NA          NA          NA          NA          NA          NA
## 7           NA          NA          NA          NA          NA          NA          NA
## 8           NA          NA          NA          NA          NA          NA          NA
## 9           NA          NA          NA          NA          NA          NA          NA
## 10          NA          NA          NA          NA          NA          NA          NA
## 11          NA          NA          NA          NA          NA          NA          NA
## 12          NA          NA          NA          NA          NA          NA          NA
## 13          NA          NA          NA          NA          NA          NA          NA
## 14          NA          NA          NA          NA          NA          NA          NA
## 15          NA          NA          NA          NA          NA          NA          NA
##      cluster_12 cluster_13 cluster_14 cluster_15
## 1           12          13          14          15
## 2           8           NA          NA           4
## 3           2           NA          NA          NA
## 4           NA          NA          NA          NA
## 5           NA          NA          NA          NA
## 6           NA          NA          NA          NA
## 7           NA          NA          NA          NA
## 8           NA          NA          NA          NA
## 9           NA          NA          NA          NA
## 10          NA          NA          NA          NA
## 11          NA          NA          NA          NA
```

```
## 12      NA      NA      NA      NA
## 13      NA      NA      NA      NA
## 14      NA      NA      NA      NA
## 15      NA      NA      NA      NA
```

Iteration Round 5

```
proximityMatrix5 <- updateProximity(clusters4,data,proximityMatrix4)
combineIndex <- indexOfMin(proximityMatrix5)
minProximity <- minOfMatrix(proximityMatrix5)
clusters5 <- updateClusters(combineIndex[1],combineIndex[2],clusters4)
proximityMatrix5
```

```
##      cluster_1 cluster_3 cluster_6 cluster_7 cluster_9 cluster_10 cluster_11
## 1         0  2.153812 1.2657780 0.8305813 2.1350191  1.8250894  2.0532672
## 2        NA  0.000000 0.9927235 1.2741056 0.5944402  0.5235198  0.4521832
## 3        NA      NA 0.0000000 0.3769650 1.0798897  0.6959714  0.8851258
## 4        NA      NA      NA 0.0000000 1.2346499  0.8789843  1.0970056
## 5        NA      NA      NA      NA 0.0000000  0.5130612  0.3344254
## 6        NA      NA      NA      NA      NA 0.0000000  0.2941505
## 7        NA      NA      NA      NA      NA      NA 0.0000000
## 8        NA      NA      NA      NA      NA      NA      NA
## 9        NA      NA      NA      NA      NA      NA      NA
## 10       NA      NA      NA      NA      NA      NA      NA
## 11       NA      NA      NA      NA      NA      NA      NA
##      cluster_12 cluster_13 cluster_14 cluster_15
## 1  0.2806071  1.7252293  1.2805097  2.0461324
## 2  2.1258241  0.9033566  1.0280121  0.2505083
## 3  1.2069372  0.7818445  0.5233692  0.9048853
## 4  0.7626820  0.7968364  0.4954452  1.1380268
## 5  2.0447576  0.5997525  0.8856337  0.3267749
## 6  1.7677223  0.4180999  0.6928935  0.3344590
## 7  1.9684629  0.5350671  0.8470162  0.2513604
## 8  0.0000000  1.6022205  1.1963091  1.9970008
## 9      NA  0.0000000  0.6267244  0.6704361
## 10     NA      NA  0.0000000  0.8630906
## 11     NA      NA      NA  0.0000000
```

```
minProximity
```

```
## [1] 0.2505083
```

```
clusters5
```

```
##      cluster_1 cluster_6 cluster_7 cluster_9 cluster_10 cluster_11
## 1         1         6         7         9         10         11
## 2        NA        NA         5        NA        NA        NA
## 3        NA        NA        NA        NA        NA        NA
## 4        NA        NA        NA        NA        NA        NA
## 5        NA        NA        NA        NA        NA        NA
## 6        NA        NA        NA        NA        NA        NA
## 7        NA        NA        NA        NA        NA        NA
## 8        NA        NA        NA        NA        NA        NA
## 9        NA        NA        NA        NA        NA        NA
## 10       NA        NA        NA        NA        NA        NA
```

```
## 11      NA      NA      NA      NA      NA      NA
## 12      NA      NA      NA      NA      NA      NA
## 13      NA      NA      NA      NA      NA      NA
## 14      NA      NA      NA      NA      NA      NA
## 15      NA      NA      NA      NA      NA      NA
##      cluster_12 cluster_13 cluster_14 cluster_15
## 1          12          13          14          15
## 2           8          NA          NA           4
## 3           2          NA          NA           3
## 4          NA          NA          NA          NA
## 5          NA          NA          NA          NA
## 6          NA          NA          NA          NA
## 7          NA          NA          NA          NA
## 8          NA          NA          NA          NA
## 9          NA          NA          NA          NA
## 10         NA          NA          NA          NA
## 11         NA          NA          NA          NA
## 12         NA          NA          NA          NA
## 13         NA          NA          NA          NA
## 14         NA          NA          NA          NA
## 15         NA          NA          NA          NA
```

Iteration Round 6

```
proximityMatrix6 <- updateProximity(clusters5,data,proximityMatrix5)
combineIndex <- indexOfMin(proximityMatrix6)
minProximity <- minOfMatrix(proximityMatrix6)
clusters6 <- updateClusters(combineIndex[1],combineIndex[2],clusters5)
proximityMatrix6
```

```
##      cluster_1 cluster_6 cluster_7 cluster_9 cluster_10 cluster_11
## 1           0  1.265778  0.8305813  2.135019  1.8250894  2.0532672
## 2          NA  0.000000  0.3769650  1.079890  0.6959714  0.8851258
## 3          NA          NA  0.0000000  1.234650  0.8789843  1.0970056
## 4          NA          NA          NA  0.000000  0.5130612  0.3344254
## 5          NA          NA          NA          NA  0.0000000  0.2941505
## 6          NA          NA          NA          NA          NA  0.0000000
## 7          NA          NA          NA          NA          NA          NA
## 8          NA          NA          NA          NA          NA          NA
## 9          NA          NA          NA          NA          NA          NA
## 10         NA          NA          NA          NA          NA          NA
##      cluster_12 cluster_13 cluster_14 cluster_15
## 1  0.2806071  1.7252293  1.2805097  2.0461324
## 2  1.2069372  0.7818445  0.5233692  0.9048853
## 3  0.7626820  0.7968364  0.4954452  1.1380268
## 4  2.0447576  0.5997525  0.8856337  0.3267749
## 5  1.7677223  0.4180999  0.6928935  0.3344590
## 6  1.9684629  0.5350671  0.8470162  0.2513604
## 7  0.0000000  1.6022205  1.1963091  1.9970008
## 8          NA  0.0000000  0.6267244  0.6704361
## 9          NA          NA  0.0000000  0.8630906
## 10         NA          NA          NA  0.0000000
```

```
minProximity
```

```
## [1] 0.2513604
```

```
clusters6
```

```
##      cluster_1 cluster_6 cluster_7 cluster_9 cluster_10 cluster_12
## 1           1           6           7           9          10          12
## 2           NA          NA           5          NA          NA           8
## 3           NA          NA          NA          NA          NA           2
## 4           NA          NA          NA          NA          NA          NA
## 5           NA          NA          NA          NA          NA          NA
## 6           NA          NA          NA          NA          NA          NA
## 7           NA          NA          NA          NA          NA          NA
## 8           NA          NA          NA          NA          NA          NA
## 9           NA          NA          NA          NA          NA          NA
## 10          NA          NA          NA          NA          NA          NA
## 11          NA          NA          NA          NA          NA          NA
## 12          NA          NA          NA          NA          NA          NA
## 13          NA          NA          NA          NA          NA          NA
## 14          NA          NA          NA          NA          NA          NA
## 15          NA          NA          NA          NA          NA          NA
##      cluster_13 cluster_14 cluster_15
## 1           13          14          15
## 2           NA          NA           4
## 3           NA          NA           3
## 4           NA          NA          11
## 5           NA          NA          NA
## 6           NA          NA          NA
## 7           NA          NA          NA
## 8           NA          NA          NA
## 9           NA          NA          NA
## 10          NA          NA          NA
## 11          NA          NA          NA
## 12          NA          NA          NA
## 13          NA          NA          NA
## 14          NA          NA          NA
## 15          NA          NA          NA
```

Iteration Round 7

```
proximityMatrix7 <- updateProximity(clusters6,data,proximityMatrix6)
combineIndex <- indexOfMin(proximityMatrix7)
minProximity <- minOfMatrix(proximityMatrix7)
clusters7 <- updateClusters(combineIndex[1],combineIndex[2],clusters6)
proximityMatrix7
```

```
##      cluster_1 cluster_6 cluster_7 cluster_9 cluster_10 cluster_12 cluster_13
## 1           0  1.265778 0.8305813  2.135019  1.8250894  0.2806071  1.7252293
## 2           NA  0.000000 0.3769650  1.079890  0.6959714  1.2069372  0.7818445
## 3           NA          NA 0.0000000  1.234650  0.8789843  0.7626820  0.7968364
## 4           NA          NA          NA  0.000000  0.5130612  2.0447576  0.5997525
## 5           NA          NA          NA          NA  0.0000000  1.7677223  0.4180999
## 6           NA          NA          NA          NA          NA  0.0000000  1.6022205
```

```
## 7      NA      NA      NA      NA      NA      NA      NA 0.0000000
## 8      NA      NA      NA      NA      NA      NA      NA      NA
## 9      NA      NA      NA      NA      NA      NA      NA      NA
## cluster_14 cluster_15
## 1  1.2805097 2.0461324
## 2  0.5233692 0.8851258
## 3  0.4954452 1.0970056
## 4  0.8856337 0.3267749
## 5  0.6928935 0.2941505
## 6  1.1963091 1.9684629
## 7  0.6267244 0.5350671
## 8  0.0000000 0.8470162
## 9      NA 0.0000000
```

```
minProximity
```

```
## [1] 0.2806071
```

```
clusters7
```

```
## cluster_6 cluster_7 cluster_9 cluster_10 cluster_12 cluster_13
## 1      6      7      9      10      12      13
## 2      NA      5      NA      NA      8      NA
## 3      NA      NA      NA      NA      2      NA
## 4      NA      NA      NA      NA      1      NA
## 5      NA      NA      NA      NA      NA      NA
## 6      NA      NA      NA      NA      NA      NA
## 7      NA      NA      NA      NA      NA      NA
## 8      NA      NA      NA      NA      NA      NA
## 9      NA      NA      NA      NA      NA      NA
## 10     NA      NA      NA      NA      NA      NA
## 11     NA      NA      NA      NA      NA      NA
## 12     NA      NA      NA      NA      NA      NA
## 13     NA      NA      NA      NA      NA      NA
## 14     NA      NA      NA      NA      NA      NA
## 15     NA      NA      NA      NA      NA      NA
## cluster_14 cluster_15
## 1      14      15
## 2      NA      4
## 3      NA      3
## 4      NA      11
## 5      NA      NA
## 6      NA      NA
## 7      NA      NA
## 8      NA      NA
## 9      NA      NA
## 10     NA      NA
## 11     NA      NA
## 12     NA      NA
## 13     NA      NA
## 14     NA      NA
## 15     NA      NA
```

Iteration Round 8

```
proximityMatrix8 <- updateProximity(clusters7,data,proximityMatrix7)
combineIndex <- indexOfMin(proximityMatrix8)
minProximity <- minOfMatrix(proximityMatrix8)
clusters8 <- updateClusters(combineIndex[1],combineIndex[2],clusters7)
proximityMatrix8
```

```
##   cluster_6 cluster_7 cluster_9 cluster_10 cluster_12 cluster_13
## 1         0  0.376965  1.07989  0.6959714  1.206937  0.7818445
## 2        NA  0.000000  1.23465  0.8789843  0.762682  0.7968364
## 3        NA        NA  0.00000  0.5130612  2.044758  0.5997525
## 4        NA        NA        NA  0.0000000  1.767722  0.4180999
## 5        NA        NA        NA        NA  0.000000  1.6022205
## 6        NA        NA        NA        NA        NA  0.0000000
## 7        NA        NA        NA        NA        NA        NA
## 8        NA        NA        NA        NA        NA        NA
##   cluster_14 cluster_15
## 1  0.5233692  0.8851258
## 2  0.4954452  1.0970056
## 3  0.8856337  0.3267749
## 4  0.6928935  0.2941505
## 5  1.1963091  1.9684629
## 6  0.6267244  0.5350671
## 7  0.0000000  0.8470162
## 8        NA  0.0000000
```

```
minProximity
```

```
## [1] 0.2941505
```

```
clusters8
```

```
##   cluster_6 cluster_7 cluster_9 cluster_12 cluster_13 cluster_14
## 1         6         7         9         12         13         14
## 2        NA         5        NA         8        NA        NA
## 3        NA        NA        NA         2        NA        NA
## 4        NA        NA        NA         1        NA        NA
## 5        NA        NA        NA        NA        NA        NA
## 6        NA        NA        NA        NA        NA        NA
## 7        NA        NA        NA        NA        NA        NA
## 8        NA        NA        NA        NA        NA        NA
## 9        NA        NA        NA        NA        NA        NA
## 10       NA        NA        NA        NA        NA        NA
## 11       NA        NA        NA        NA        NA        NA
## 12       NA        NA        NA        NA        NA        NA
## 13       NA        NA        NA        NA        NA        NA
## 14       NA        NA        NA        NA        NA        NA
## 15       NA        NA        NA        NA        NA        NA
##   cluster_15
## 1         15
## 2          4
## 3          3
## 4         11
## 5         10
```

```
## 6      NA
## 7      NA
## 8      NA
## 9      NA
## 10     NA
## 11     NA
## 12     NA
## 13     NA
## 14     NA
## 15     NA
```

Iteration Round 9

```
proximityMatrix9 <- updateProximity(clusters8,data,proximityMatrix8)
combineIndex <- indexOfMin(proximityMatrix9)
minProximity <- minOfMatrix(proximityMatrix9)
clusters9 <- updateClusters(combineIndex[1],combineIndex[2],clusters8)
proximityMatrix9
```

```
##   cluster_6 cluster_7 cluster_9 cluster_12 cluster_13 cluster_14
## 1         0  0.376965  1.07989   1.206937  0.7818445  0.5233692
## 2        NA  0.000000  1.23465   0.762682  0.7968364  0.4954452
## 3        NA        NA  0.00000   2.044758  0.5997525  0.8856337
## 4        NA        NA        NA  0.000000  1.6022205  1.1963091
## 5        NA        NA        NA        NA  0.0000000  0.6267244
## 6        NA        NA        NA        NA        NA  0.0000000
## 7        NA        NA        NA        NA        NA        NA
##   cluster_15
## 1  0.6959714
## 2  0.8789843
## 3  0.3267749
## 4  1.7677223
## 5  0.4180999
## 6  0.6928935
## 7  0.0000000
```

```
minProximity
```

```
## [1] 0.3267749
```

```
clusters9
```

```
##   cluster_6 cluster_7 cluster_12 cluster_13 cluster_14 cluster_15
## 1         6         7         12         13         14         15
## 2        NA         5          8        NA        NA          4
## 3        NA        NA          2        NA        NA          3
## 4        NA        NA          1        NA        NA         11
## 5        NA        NA        NA        NA        NA         10
## 6        NA        NA        NA        NA        NA          9
## 7        NA        NA        NA        NA        NA        NA
## 8        NA        NA        NA        NA        NA        NA
## 9        NA        NA        NA        NA        NA        NA
## 10       NA        NA        NA        NA        NA        NA
## 11       NA        NA        NA        NA        NA        NA
## 12       NA        NA        NA        NA        NA        NA
```



```
## 13      NA      NA      NA      NA      NA      NA
## 14      NA      NA      NA      NA      NA      NA
## 15      NA      NA      NA      NA      NA      NA
```

Iteration Round 10

```
proximityMatrix10 <- updateProximity(clusters9,data,proximityMatrix9)
combineIndex <- indexOfMin(proximityMatrix10)
minProximity <- minOfMatrix(proximityMatrix10)
clusters10 <- updateClusters(combineIndex[1],combineIndex[2],clusters9)
proximityMatrix10
```

```
##   cluster_6 cluster_7 cluster_12 cluster_13 cluster_14 cluster_15
## 1         0  0.376965  1.206937  0.7818445  0.5233692  0.6959714
## 2        NA  0.000000  0.762682  0.7968364  0.4954452  0.8789843
## 3        NA        NA  0.000000  1.6022205  1.1963091  1.7677223
## 4        NA        NA        NA  0.0000000  0.6267244  0.4180999
## 5        NA        NA        NA        NA  0.0000000  0.6928935
## 6        NA        NA        NA        NA        NA  0.0000000
```

```
minProximity
```

```
## [1] 0.376965
```

```
clusters10
```

```
##   cluster_7 cluster_12 cluster_13 cluster_14 cluster_15
## 1         7         12         13         14         15
## 2         5          8        NA         NA          4
## 3         6          2        NA         NA          3
## 4        NA          1        NA         NA         11
## 5        NA        NA        NA         NA         10
## 6        NA        NA        NA         NA          9
## 7        NA        NA        NA         NA         NA
## 8        NA        NA        NA         NA         NA
## 9        NA        NA        NA         NA         NA
## 10       NA        NA        NA         NA         NA
## 11       NA        NA        NA         NA         NA
## 12       NA        NA        NA         NA         NA
## 13       NA        NA        NA         NA         NA
## 14       NA        NA        NA         NA         NA
## 15       NA        NA        NA         NA         NA
```

Iteration Round 11

```
proximityMatrix11 <- updateProximity(clusters10,data,proximityMatrix10)
combineIndex <- indexOfMin(proximityMatrix11)
minProximity <- minOfMatrix(proximityMatrix11)
clusters11 <- updateClusters(combineIndex[1],combineIndex[2],clusters10)
proximityMatrix11
```

```
##   cluster_7 cluster_12 cluster_13 cluster_14 cluster_15
## 1         0  0.762682  0.7818445  0.4954452  0.6959714
```

```
## 2      NA      0.000000  1.6022205  1.1963091  1.7677223
## 3      NA      NA      0.0000000  0.6267244  0.4180999
## 4      NA      NA      NA      0.0000000  0.6928935
## 5      NA      NA      NA      NA      0.0000000
```

```
minProximity
```

```
## [1] 0.4180999
```

```
clusters11
```

```
##      cluster_7 cluster_12 cluster_14 cluster_15
## 1           7          12          14          15
## 2           5           8          NA           4
## 3           6           2          NA           3
## 4          NA           1          NA          11
## 5          NA          NA          NA          10
## 6          NA          NA          NA           9
## 7          NA          NA          NA          13
## 8          NA          NA          NA          NA
## 9          NA          NA          NA          NA
## 10         NA          NA          NA          NA
## 11         NA          NA          NA          NA
## 12         NA          NA          NA          NA
## 13         NA          NA          NA          NA
## 14         NA          NA          NA          NA
## 15         NA          NA          NA          NA
```

Iteration Round 12

```
proximityMatrix12 <- updateProximity(clusters11,data,proximityMatrix11)
combineIndex <- indexOfMin(proximityMatrix12)
minProximity <- minOfMatrix(proximityMatrix12)
clusters12 <- updateClusters(combineIndex[1],combineIndex[2],clusters11)
proximityMatrix12
```

```
##      cluster_7 cluster_12 cluster_14 cluster_15
## 1           0  0.762682  0.4954452  0.6959714
## 2          NA  0.000000  1.1963091  1.6022205
## 3          NA      NA  0.0000000  0.6267244
## 4          NA      NA      NA  0.0000000
```

```
minProximity
```

```
## [1] 0.4954452
```

```
clusters12
```

```
##      cluster_12 cluster_14 cluster_15
## 1           12          14          15
## 2           8           7           4
## 3           2           5           3
## 4           1           6          11
## 5          NA          NA          10
## 6          NA          NA           9
## 7          NA          NA          13
```

```
## 8      NA      NA      NA
## 9      NA      NA      NA
## 10     NA      NA      NA
## 11     NA      NA      NA
## 12     NA      NA      NA
## 13     NA      NA      NA
## 14     NA      NA      NA
## 15     NA      NA      NA
```

Iteration Round 13

```
proximityMatrix13 <- updateProximity(clusters12,data,proximityMatrix12)
combineIndex <- indexOfMin(proximityMatrix13)
minProximity <- minOfMatrix(proximityMatrix13)
clusters13 <- updateClusters(combineIndex[1],combineIndex[2],clusters12)
proximityMatrix13
```

```
##   cluster_12 cluster_14 cluster_15
## 1          0   0.762682  1.6022205
## 2         NA   0.000000  0.6267244
## 3         NA          NA  0.0000000
```

```
minProximity
```

```
## [1] 0.6267244
```

```
clusters13
```

```
##   cluster_12 cluster_15
## 1          12          15
## 2           8           4
## 3           2           3
## 4           1          11
## 5          NA          10
## 6          NA           9
## 7          NA          13
## 8          NA          14
## 9          NA           7
## 10         NA           5
## 11         NA           6
## 12         NA          NA
## 13         NA          NA
## 14         NA          NA
## 15         NA          NA
```

Iteration Round 14

```
proximityMatrix14 <- updateProximity(clusters13,data,proximityMatrix13)
combineIndex <- indexOfMin(proximityMatrix14)
minProximity <- minOfMatrix(proximityMatrix14)
clusters14 <- updateClusters(combineIndex[1],combineIndex[2],clusters13)
proximityMatrix14
```

```
## cluster_12 cluster_15
## 1          0  0.762682
## 2          NA  0.000000
```

```
minProximity
```

```
## [1] 0.762682
```

```
clusters14
```

```
## [1] 15  4  3 11 10  9 13 14  7  5  6 12  8  2  1
```

Now all nodes are in one cluster. The final dendrogram has been formed.