

Revision 7

Revision

Question 1.

```
# We create the points in R
x1 <- c(6, 5)
x2 <- c(4, 6)
x3 <- c(1, 1)
x4 <- c(5, 4)
x5 <- c(2, 1)
x6 <- c(3, 5)
x7 <- c(7, 4)
x8 <- c(1, 3)

X <- rbind(x1, x2, x3, x4, x5, x6, x7, x8) # a, b and c are combined per row
colnames(X) <- c("x", "y") # rename columns

X # display the points

##      x y
## x1 6 5
## x2 4 6
## x3 1 1
## x4 5 4
## x5 2 1
## x6 3 5
## x7 7 4
## x8 1 3

# The distance is found using the dist() function:
distance <- dist(X, method = "euclidean")
distance # display the distance matrix

##      x1      x2      x3      x4      x5      x6      x7
## x2 2.236068
## x3 6.403124 5.830952
## x4 1.414214 2.236068 5.000000
## x5 5.656854 5.385165 1.000000 4.242641
## x6 3.000000 1.414214 4.472136 2.236068 4.123106
## x7 1.414214 3.605551 6.708204 2.000000 5.830952 4.123106
## x8 5.385165 4.242641 2.000000 4.123106 2.236068 2.828427 6.082763

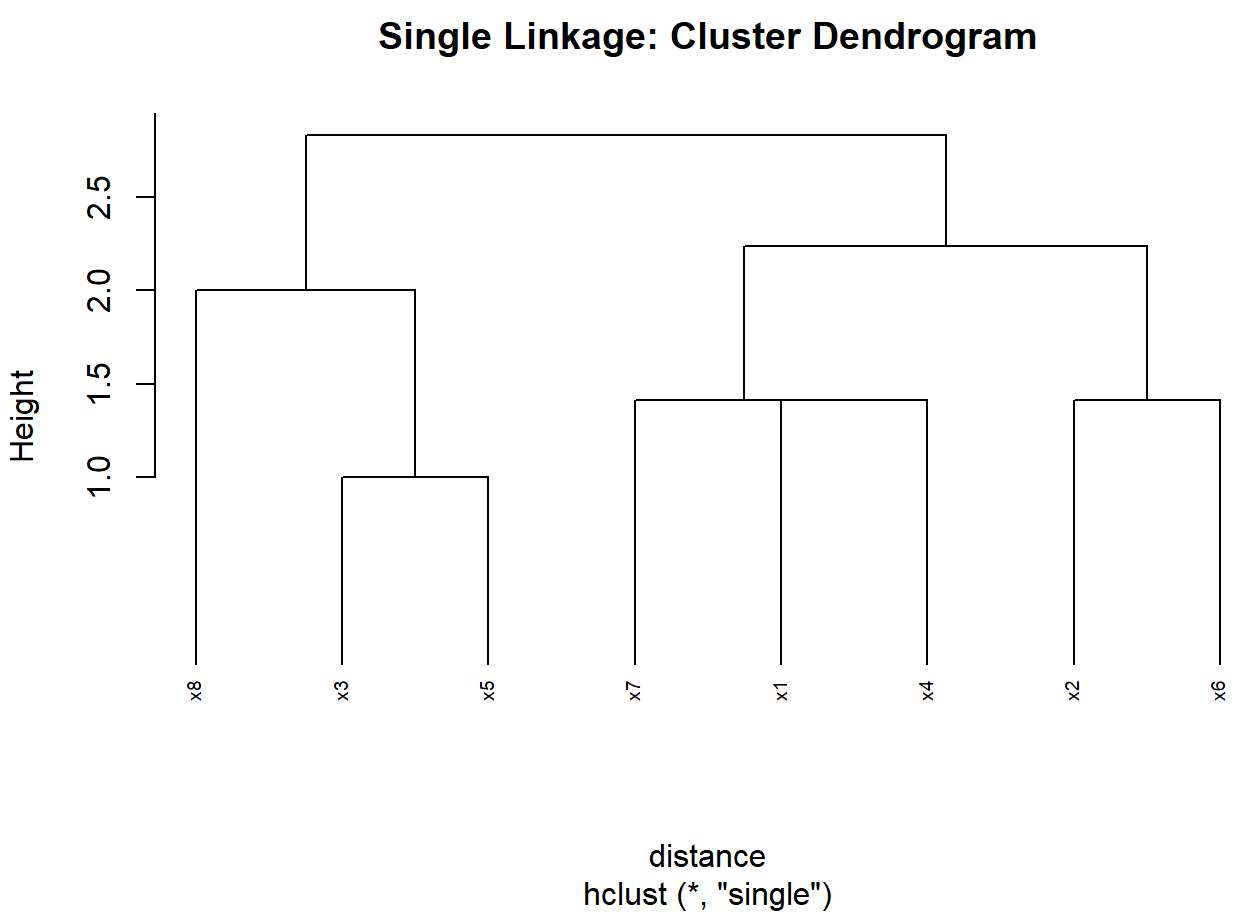
# Agglomerative Hierarchical Clustering for single linkage
start.time <- Sys.time()
ahc.s <- hclust(distance, method="single")
ahc.s

##
## Call:
## hclust(d = distance, method = "single")
##
## Cluster method   : single
## Distance         : euclidean
## Number of objects: 8

end.time <- Sys.time()
print(end.time - start.time)

## Time difference of 0.00798893 secs

# Plotting a dendrogram for single linkage
plot(ahc.s, hang = -1, cex = 0.6, main="Single Linkage: Cluster Dendrogram")
```



Question 2.

```
# We create the points in R
x1 <- c(6, 5, 4)
x2 <- c(0, 0, 0)
x3 <- c(1, 1, -2)
x4 <- c(7, 3, -1)
x5 <- c(2, -1, -1)
x6 <- c(4, 4, 4)

d <- rbind(x1, x2, x3, x4, x5, x6) # a, b and c are combined per row
colnames(d) <- c("x", "y", "z") # rename columns

d # display the points

##      x  y  z
## x1 6  5  4
## x2 0  0  0
## x3 1  1 -2
## x4 7  3 -1
## x5 2 -1 -1
## x6 4  4  4

# The distance is found using the dist() function:
d.dist <- dist(d, method = "euclidean")
d.dist # display the distance matrix

##      x1      x2      x3      x4      x5
## x2 8.774964
## x3 8.774964 2.449490
## x4 5.477226 7.681146 6.403124
## x5 8.774964 2.449490 2.449490 6.403124
## x6 2.236068 6.928203 7.348469 5.916080 7.348469

# Agglomerative Hierarchical Clustering for single linkage
start.time <- Sys.time()
ahc.c <- hclust(d.dist, method="complete")
ahc.c

##
## Call:
## hclust(d = d.dist, method = "complete")
##
## Cluster method   : complete
## Distance         : euclidean
## Number of objects: 6

end.time <- Sys.time()
print(end.time - start.time)

## Time difference of 0.004143953 secs

# Plotting a dendrogram for single linkage
plot(ahc.c, hang = -1, cex = 0.6, main="Complete Linkage: Cluster Dendrogram")
```

