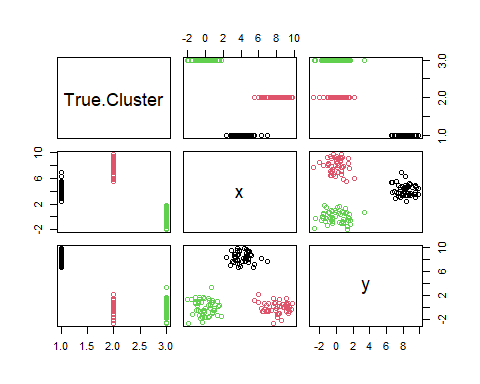
In Class Assignment Module 7

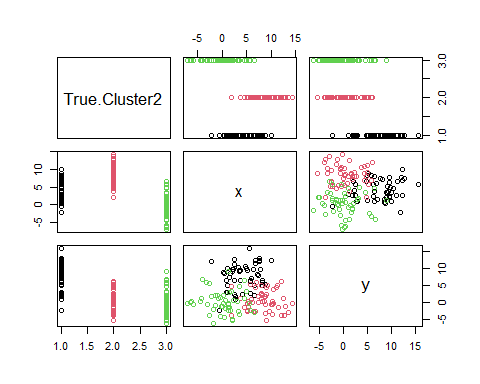
Arturo Ortiz

2023-11-29

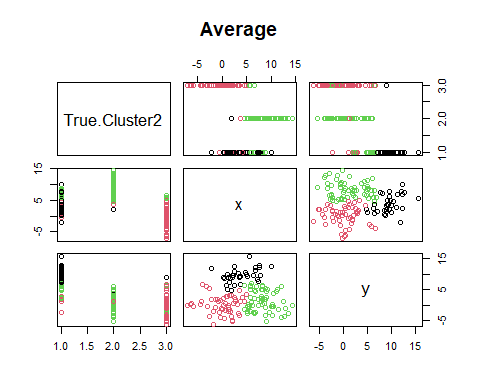
d1 <- cbind(rep(1, 50), rnorm(50, 4, 1), rnorm(50, 8, 1))  
d2 <- cbind(rep(2, 50), rnorm(50, 8, 1), rnorm(50, 0, 1))  
d3 <- cbind(rep(3, 50), rnorm(50, 0, 1), rnorm(50, 0, 1))  
  
mydata1 <- rbind(d1, d2, d3)  
mydata1 <- data.frame(mydata1)  
colnames(mydata1) <- c("True.Cluster", "x", "y")  
  
mydata1\_dist <- dist(mydata1)  
mydata1\_hclust <- hclust(mydata1\_dist, method = "single")  
mydata1\_cutree <- cutree(mydata1\_hclust, 3)  
plot(mydata1, col = mydata1\_cutree)



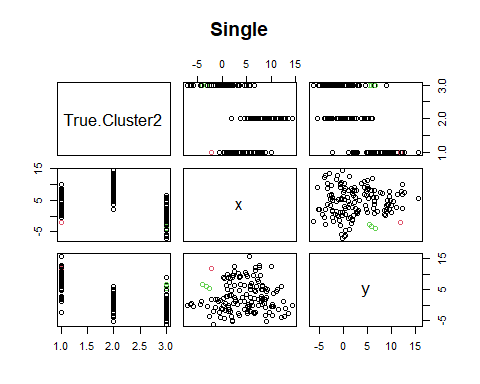
d1 <- cbind(rep(1, 50), rnorm(50, 4, 3), rnorm(50, 8, 3))  
d2 <- cbind(rep(2, 50), rnorm(50, 8, 3), rnorm(50, 0, 3))  
d3 <- cbind(rep(3, 50), rnorm(50, 0, 3), rnorm(50, 0, 3))  
  
mydata2 <- rbind(d1, d2, d3)  
mydata2 <- data.frame(mydata2)  
colnames(mydata2) <- c("True.Cluster2", "x", "y")  
plot(mydata2, col = mydata2$True.Cluster2)



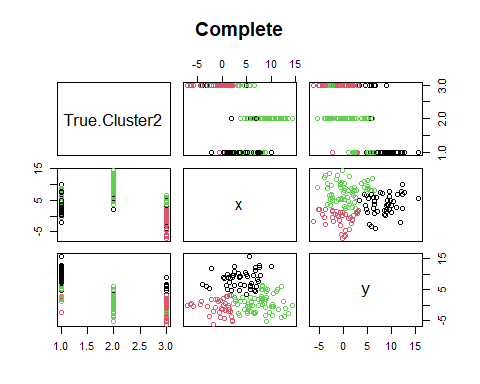
mydata2\_dist <- dist(mydata2)  
mydata2\_hclust <- hclust(mydata2\_dist, method = "average")  
mydata2\_cutree <- cutree(mydata2\_hclust, 3)  
plot(mydata2, col = mydata2\_cutree, main = "Average")



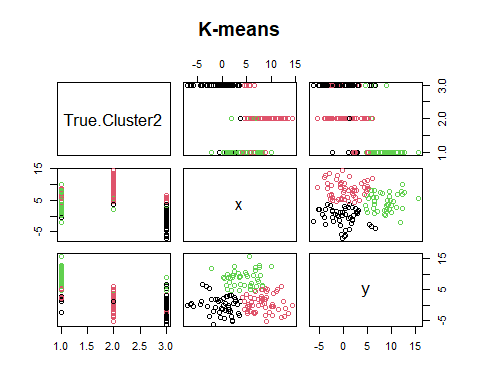
mydata2\_hclust1 <- hclust(mydata2\_dist, method = "single")  
mydata2\_cutree1 <- cutree(mydata2\_hclust1, 3)  
plot(mydata2, col = mydata2\_cutree1, main = "Single")



mydata2\_hclust3 <- hclust(mydata2\_dist, method = "complete")  
mydata2\_cutree3 <- cutree(mydata2\_hclust3, 3)  
plot(mydata2, col = mydata2\_cutree3, main = "Complete")



mydata2\_kmeans <- kmeans(mydata2, centers = 3)  
plot(mydata2, col = mydata2\_kmeans$cluster, main = "K-means")



Because hierarchical clustering shows a higher degree of agreement between the clusters it produces, it performs better compared to K-means. Because hierarchical clustering successfully clusters a greater proportion of data points, it performs better in this situation, as indicated by the higher diagonal values.”