  






setwd("/Users/Lauren/Desktop/")

ALD <- read.csv("Ordered Team4-ALD\_4cohort\_dataset.csv")

ALD

control <- ALD[1:25, 3:152]

control

options(max.print = 99999)

control

control\_AC <- ALD[26:45, 3:152]

options(max.print = 99999)

control\_AC

mild\_AH <- ALD[46:55, 3:152]

options(max.print = 99999)

mild\_AH

severe\_AH <- ALD[56:78, 3:152]

options(max.print = 99999)

severe\_AH

ALD\_matrix <- as.matrix(read.csv("Ordered Team4-ALD\_4cohort\_dataset.csv"))

ALD\_matrix

cov\_ALD\_matrix <- cov(ALD\_matrix)

cohort\_vs\_species <- read.csv("cohorts\_species.csv")

dist.cohort\_vs\_species <- dist(cohort\_vs\_species, method="euclidean")

dist.cohort\_vs\_species

dist.control\_AC <- dist(control\_AC, method="euclidean")

dist.control\_AC

dist.mild\_AH <- dist(mild\_AH, method="euclidean")

dist.mild\_AH

dist.severe\_AH <- dist(severe\_AH, method="euclidean")

dist.severe\_AH

dist.control <- dist(control, method="euclidean")

dist.control

clustcontrol\_AC <- hclust(dist.control\_AC, method="single")

plot(as.dendrogram(clustcontrol\_AC), horiz = TRUE, xlab="Distance between Species for control\_AC", main="Dendrogram based on control\_AC cohort of 152 bacteria samples")

clustcontrol <- hclust(dist.control, method="single")

plot(as.dendrogram(clustcontrol), horiz = TRUE, xlab="Distance between Species for control", main="Dendrogram based on control cohort of 152 bacteria samples")

clustmild\_AH <- hclust(dist.mild\_AH, method="single")

plot(as.dendrogram(clustmild\_AH), horiz = TRUE, xlab="Distance between Species for mild\_AH", main="Dendrogram based on mild\_AH cohort of 152 bacteria samples")

dist.severe\_AH <- dist(severe\_AH, method="euclidean")

dist.severe\_AH

clustsevere\_AH <- hclust(dist.severe\_AH, method="single")

plot(as.dendrogram(clustsevere\_AH), horiz = TRUE, xlab="Distance between Species for severe\_AH", main="Dendrogram based on severe\_AH cohort of 152 bacteria samples")