

```

library(car)
library(MASS)
library(class)
library(sp)
library(lattice)
library(geoR)
library(gstat)
rm(list = ls())
setwd('C:/Users/gianm/Desktop/TDEApplied/Exam 15 Giugno')
load('mcshapiro.test.RData')
df <- read.table("stoneflakes.txt")
head(df)

plot(df[,1],df[,2], pch = 16, xlab = 'Length', ylab = 'Width')

df.e<- dist(df, method='euclidean') # distance matrix

## single linkage

df.es <- hclust(df.e, method='ward.D') # clustering gerarchico con Ward
linkage

coph.es <- cophenetic(df.es)
coph.es

cor(coph.es, df.e)
#0.8898677
x11()
par(mfrow= c(1,2))
image(as.matrix(df.e), main='Euclidean', asp=1 )
image(as.matrix(coph.es), main='Single', asp=1 )

x11() # dendrogramma
plot(df.es, main='euclidean-Ward', hang=-0.1, xlab='', labels=F, cex=0.6,
sub='')
rect.hclust(df.es, k=3)

help(cutree)
cluster.es <- cutree(df.es, k=3)
cluster.es

dim1 <- length(cluster.es[which(cluster.es == 1)]) #
dim2 <- length(cluster.es[which(cluster.es == 2)]) #
dim3 <- length(cluster.es[which(cluster.es == 3)])
dim1
dim2
dim3

centroids <- apply (df, 2, function (x) tapply (x, cluster.es, mean))
centroids

x11()
plot(df[,1],df[,2], pch = 16, xlab = 'Length', ylab = 'Width', col =
cluster.es)

df$clu <- cluster.es

```

```

mcshapiro.test(df[df$clu == 1,1:2])
mcshapiro.test(df[df$clu == 2,1:2])
mcshapiro.test(df[cluster.es == 3,1:2])

mod <- manova(cbind(Length, Width)~ clu, data = df)
summary.manova(mod, test = 'Wilks')

n <- dim(df)[1]
n1 <- dim1
n2 <- dim2
n3 <- dim3
m1 <- sapply(df[df$clu == 1,1:2],mean)
m2 <- sapply(df[df$clu == 2,1:2],mean)
m3 <- sapply(df[df$clu == 3,1:2],mean)
p <-2
g<-3
alpha <- 0.1
k <- p*g*(g-1)/2
qT <- qt(1-alpha/(2*k), n-g)

W <- summary.manova(mod)$SS$Residuals
m <- sapply(df,mean) # estimates mu

inf12 <- m1-m2 - qT * sqrt( diag(W)/(n-g) * (1/n1+1/n2) )
sup12 <- m1-m2 + qT * sqrt( diag(W)/(n-g) * (1/n1+1/n2) )
inf13 <- m1-m3 - qT * sqrt( diag(W)/(n-g) * (1/n1+1/n3) )
sup13 <- m1-m3 + qT * sqrt( diag(W)/(n-g) * (1/n1+1/n3) )
inf23 <- m2-m3 - qT * sqrt( diag(W)/(n-g) * (1/n2+1/n3) )
sup23 <- m2-m3 + qT * sqrt( diag(W)/(n-g) * (1/n2+1/n3) )

CI <- list('cluster1-cluster2'=cbind(inf12, m1-m2, sup12), 'cluster1-
cluster3'=cbind(inf13, m1-m3,sup13), 'cluster2-cluster3'=cbind(inf23,m2-m3,
sup23))
CI

```