#preambolo

setwd("~/Desktop/Università/Analisi dei dati/esercitazioni R/Progetto Data")

library(readxl)

library(FactoMineR)

#-------------------------------

#PCA

macchine <- read\_xlsx("matrice\_city\_cars.xlsx" , sheet = "Foglio1")

macchine2 <- data.frame(macchine, row.names = "elenco")

macchine2.pca <- PCA(macchine2)

macchine2.pca$eig

macchine2.pca$var$cor

macchine2.pca$ind$coord

macchine2.pca$ind$cos2

#--------------------------------

#Cluster Analysis

#Gerarchico

cor(macchine2)

boxplot.matrix(as.matrix(macchine2))

dist.euclidea <- dist(macchine2, method = "euclidean")

cluster1 <- hclust(dist.euclidea, method = "complete")

cluster1$merge

cluster1$height

plot(cluster1)

abline(h=300, col="red")

cluster2 <- hclust(dist.euclidea, method = "average")

plot(cluster2)

abline(h=250, col="blue")

cluster3 <- hclust(dist.euclidea, method = "single")

plot(cluster3)

abline(h=125, col="green")

distanza.manhat<- dist(macchine2, method = "manhattan")

cluster4 <- hclust(distanza.manhat, method = "complete")

plot(cluster4)

abline(h=700, col="purple")

#-----------------------------------

#Cluster Analyisis

#non gerarchica

set.seed(234)

clusterNG1 <- kmeans(macchine2, centers = 2, algorithm = "MacQueen")

clusterNG1$cluster

clusterNG1$centers

clusterNG1$totss

clusterNG1$withinss

clusterNG1$tot.withinss

clusterNG1$betweenss

clusterNG1$size

summary(macchine2[which(clusterNG1$cluster==1),])

summary(macchine2[which(clusterNG1$cluster==2),])