

```

install.packages("NbClust")
projectdir <- "C:/rezawork/study/r-studio/edx-marketing analytics/week2-marketsegmentation"
setwd(projectdir)
getwd()
seg_data <- read.csv(file = "SegmentationData.csv",row.names=1)
head(seg_data)
# standardization
std_seg_data <- scale(seg_data[,c("Trendy", "Styling", "Reliability", "Sportiness", "Performance", "Comfort")])
dist <- dist(std_seg_data, method = "euclidean")
as.matrix(dist)[1:5,1:5]
# hierarchical clustering
set.seed(1990)
clust <- hclust(dist, method = "ward.D2")
plot(clust)
# cutree 4
h_cluster <- cutree(clust, 4)
rect.hclust(clust, k=4, border="red")
table(h_cluster)

#the clustering variables means by cluster
hclust_summary <- aggregate(std_seg_data[,c("Trendy", "Styling", "Reliability", "Sportiness",
                                             "Performance", "Comfort")],by=list(h_cluster),FUN=mean)

hclust_summary

# cutree 3
plot(clust)
h_cluster <- cutree(clust, 3)
rect.hclust(clust, k=3, border="red")
table(h_cluster)

#the clustering variables means by cluster
hclust_summary <- aggregate(std_seg_data[,c("Trendy", "Styling", "Reliability", "Sportiness",
                                             "Performance", "Comfort")],by=list(h_cluster),FUN=mean)

hclust_summary

# rename those clusters according to their characteristics
h_cluster <- factor(h_cluster,levels = c(1,2,3),labels = c("Perf.", "Comfort", "Appearance"))

# focus on a given cluster
plot(cut(as.dendrogram(clust), h=9)$lower[[3]])

# Number of Clusters, we must use only 5 variables out of 6 to avoid collinearity issues.
library(NbClust)
set.seed(1990)
NbClust(data=std_seg_data[,1:5], min.nc=3, max.nc=15, index="all", method="ward.D2")

# Targeting the Clusters/segments Demographics
library(gmodels)
CrossTable(seg_data$MBA,h_cluster,prop.chisq = FALSE, prop.r = T, prop.c = T,prop.t = F,chisq = T)

# Targeting the Clusters/segments Choice
CrossTable(h_cluster, seg_data$Choice, prop.chisq = FALSE, prop.r = T, prop.c = T,prop.t = F,chisq = T)

# which isnot statistically proved

```