```
## Ex1

dataset <- read.table('wine.txt', header = T)   # se serve aggiungi il nome
delle colonne col.names=c('',..)
dataset

load("mcshapiro.test.RData")

group1   <- factor(dataset$color)    # qui species è il primo grouping
group2   <- factor(dataset$region)     # qui gender è il secondo grouping

x  <- dataset[,1]    # salvo la variabile numerica in x

g <- length(levels(group1))
b <- length(levels(group2))
n <- length(x)/(g*b)

# Check hyp
group_12 <- group1
levels(group_12) <- c('R_P','R_T','R_V','W_P','W_T','W_V')  #variable which
divides the g*b groups (only for plot)
group_12[group1=='red' & group2=='Piemonte'] <- 'R_P'
group_12[group1=='red' & group2=='Toscana'] <- 'R_T'
group_12[group1=='red' & group2=='Veneto'] <- 'R_V'
group_12[group1=='white' & group2=='Piemonte'] <- 'W_P'
group_12[group1=='white' & group2=='Toscana'] <- 'W_T'
group_12[group1=='white' & group2=='Veneto'] <- 'W_V'


# 1) normality (multivariate) in each group (4 test)
load("mcshapiro.test.RData")
Ps <- c(shapiro.test(x[ group_12==levels(group_12)[1]])$p,
         shapiro.test(x[ group_12==levels(group_12)[2]])$p,
         shapiro.test(x[ group_12==levels(group_12)[3]])$p,
         shapiro.test(x[ group_12==levels(group_12)[4]])$p,
         shapiro.test(x[ group_12==levels(group_12)[5]])$p,
         shapiro.test(x[ group_12==levels(group_12)[6]])$p)
Ps
#   0.2369077 0.4612391 0.9702243 0.6525198 0.5279147 0.6612603
# ok

# 2) homogeneity of the covariance (qualitatively)
S1 <-  var(x[ group_12==levels(group_12)[1]])
S2 <-  var(x[ group_12==levels(group_12)[2]])
S3 <-  var(x[ group_12==levels(group_12)[3]])
S4 <-  var(x[ group_12==levels(group_12)[4]])
S5 <-  var(x[ group_12==levels(group_12)[5]])
S6 <-  var(x[ group_12==levels(group_12)[6]])

# test of homogeneity of variances (1D)
# H0: sigma.1 = .. = sigma.g
# H1: there exist i,j s.t. sigma.i!=sigma.j

bartlett.test(x, group_12)
```

```
# ok


### Model with interaction (complete model):
### x.ijk = mu + tau.i + beta.j + gamma.ij + eps.ijk; eps.ijk~N(0,sigma^2),
###    i=1,2 (effect Fact1), j=1,2 (effect Fact2)

fit.aov2.int <- aov(x ~ group1 + group2 + group1:group2)  # aov(y ~ factor1 +
factor2 + interaction)
summary.aov(fit.aov2.int)


# b)
# reduce
fit.aov2.ad <- aov(x ~ group1 + group2)
summary.aov(fit.aov2.ad)


fit.aov1 <- aov(x ~ group1)
summary.aov(fit.aov1)

# Estimate the great mean mu:
m <- mean(x)

# Estimate tau.i, beta.j:
tau1  <- mean(x[group1=='red']) - m
tau2 <- mean(x[group1=='white']) - m  # devo avere g righe


# variance:
DF <- fit.aov1$df # fit è il modello finale

Spooled <- sum(fit.aov1$res^2)/DF
Spooled  # estimate for the variance


# BonfCI
DF <- fit.aov1$df # fit è il modello finale


Spooled <- sum(fit.aov1$res^2)/DF   # SS_res= Spooled * DF= sum(eps_hat_i) =
sum(fit$res^2) ~ sigma^2 Chi(n-g)
means <- tapply(x, group1, mean)
names(means) <- levels(group1)
means

n_group = length(x)/2   # qui sto assumendo che i livelli di ogni gruppo
abbiano stessa numerosità
alpha <- 0.01
k     <- 2+1 # g Conf Int for the means and 1 for the variance
# se voglio solo Bonf per le medie o per var ricordati di cambiare il k !

# l'intervallo per le g medie lo trovo dividendo una normale per la radice
della Chi/DF -> t
```

```
BF_means    <- rbind(cbind(means - sqrt(Spooled / n_group) * qt(1 - alpha /
(2*k), DF),
                          means + sqrt(Spooled / n_group) * qt(1 - alpha /
(2*k), DF)))

# l'intervallo per sigma^2 lo trovo dalla distribuzione di SS_res ~ sigma^2
Chi(n-g)
# => chi^2_alpha/2 < SS_res / sigma^2 < chi^2_(1-alpha/2) con SS_res= Spooled
* DF
BF_var = c(Spooled * DF / qchisq(1 - alpha / (2*k), DF),
           Spooled * DF / qchisq(alpha / (2*k), DF))

BF = c(BF_means, BF_var)
BF

BF_means
BF_var
```