

```

setwd("E:/workingdirectory/phc6053")

#####
# USE THE SAS7BDAT PACKAGE TO READ IN THE DATA
#####
library(sas7bdat)
library(foreign)
df <- read.ssd("E:/workingdirectory/phc6053", "fghm81",
              sascmd="C:/Program Files/SASHome93/SASFoundation/9.3/sas.exe")
# df <- read.sas7bdat("wcgs.sas7bdat") # assignment 2 file?

#####
# MAKE SURE WE'RE LOOKING AT THE RIGHT STUFF
#####
is.data.frame(df)
colnames(df)
summary(df)

#####
# CREATE THE LOG OF SYSBP
#####
df$LNSBP <- log(df$SYSBP)

#####
# CREATE THE CATEGORIZED BMI VARIABLE
#####
df$BMIGROUPS <- factor(ifelse(df$BMI < 18.5,
                              1,
                              ifelse(df$BMI >= 18.5 &
                                      df$BMI < 25,
                                      2,
                                      ifelse(df$BMI >= 25 &
                                              df$BMI < 30,
                                              3,
                                              ifelse(df$BMI >= 30,
                                                      4,
                                                      NA))),
                              labels=c("Underweight",
                                      "Normal",
                                      "Overweight",
                                      "Obese")))

summary(df$BMIGROUPS)

#####
# CONDUCT THE NUMERIC SUMMARIES
#####
summary(df$SYSBP)
sd(df$SYSBP)
summary(df$LNSBP)
sd(df$LNSBP)
summary(df$BMI)
sd(df$BMI)
summary(df$AGE)
sd(df$AGE)
#####
# FREQ TABLE OF SEX AND BMIGROUPS
#####
df$SEX <- factor(df$SEX, labels=c("Male", "Female"))

```

```

table(df$SEX, df$BMIGROUPS)
prop.table(table(df$SEX, df$BMIGROUPS))

table(df$SEX)
prop.table(table(df$SEX))

table(df$BMIGROUPS)
prop.table(table(df$BMIGROUPS))
#####
# CORRELATION COEFFICIENT FOR Y=SYSBP, X=AGE
#####
cor(df$AGE, df$SYSBP)
cor.test(df$AGE, df$SYSBP)

#####
# CORRELATION COEFFICIENT FOR Y=SYSBP, X=BMI
#####
cor(df$BMI, df$SYSBP)
cor.test(df$BMI, df$SYSBP)
#####
# HIST, BOXPLOT, QQPLOT for SYSBP, LNSBP, BMI, AGE
#####
#SYSBP
hist(df$SYSBP)
boxplot(df$SYSBP)
qqnorm(df$SYSBP)
qqline(df$SYSBP, col="red")

#LNSBP
hist(df$LNSBP)
boxplot(df$LNSBP)
qqnorm(df$LNSBP)
qqline(df$LNSBP, col="red")

#BMI
hist(df$BMI)
boxplot(df$BMI)
qqnorm(df$BMI)
qqline(df$BMI, col="red")

#AGE
hist(df$AGE)
boxplot(df$AGE)
qqnorm(df$AGE)
qqline(df$AGE, col="red")

#####
# SCATTERPLOT Y=SYSBP X=AGE, with LOESS CURVE
#####
#plot(df$AGE, df$SYSBP)
lw1 <- loess(df$SYSBP ~ df$AGE)
plot(df$SYSBP ~ df$AGE, pch=16, col=adjustcolor("darkblue", alpha.f=0.5))
ord <- order(df$AGE)
lines(df$AGE[ord], lw1$fitted[ord], col=adjustcolor("red", alpha.f=0.6),
      lwd=3)

#####
# SCATTERPLOT Y=SYSBP X=BMI, with LOESS CURVE
#####
#plot(df$BMI, df$SYSBP)
lw1 <- loess(df$SYSBP ~ df$BMI)

```

```
plot(df$SYSBP ~ df$BMI, pch=16, col=adjustcolor("darkgreen", alpha.f=0.5))
ord <- order(df$BMI)
lines(df$BMI[ord], lw1$fitted[ord], col=adjustcolor("red", alpha.f=0.6),
      lwd=3)
#####
#
#####

save.image("E:/workingdirectory/phc6053/ass1.RData")
```