Q5: Source R code

```r
# ---------> complete and run the following code for this assignment  <-------
#
#
# R code for STA302 or STA1001H1F assignment 3
# copyright by Dogan Akad
# date: Dec. 5, 2016
#

## Load in the data set
a3 = read.table("/Users/doganakad/Desktop/uoft/first semester/STA302/Assignments/A3/a3data.txt",sep="",header=T)
str(a3)
is.factor(a3$gender)
is.factor(a3$smoke)
a3$smoke = as.factor(a3$smoke)
is.factor(a3$smoke)
a3$logPlasma = log(a3$plasma)
## ==>Q1
## subtract gender,smoke,vitamin,plasma
a3_subset = a3[, -c(2,3,5,12)]
panel.pearson <- function(x, y,...) {
  horizontal <- (par("usr")[1] + par("usr")[2]) / 2;
  vertical <- (par("usr")[3] + par("usr")[4]) / 2.8;
  text(horizontal, vertical, format(abs(cor.test(x,y)$p.value), digits=2))
}
pairs(a3_subset, main = "A3 data", pch = 21, bg = c("red", "blue"), lower.panel = panel.pearson)
## ==> Q2 fit regression lines
## Regression line with calories only
m0 = lm(a3$logPlasma ~ a3$Calories)
## Regression line with calories and fat
m1 = lm(a3$logPlasma ~ a3$Calories + a3$fat)
## Regression line with calories and QI
m2 = lm(a3$logPlasma ~ a3$Calories + a3$QI)
summary(m0)
summary(m1)
summary(m2)
##==> Q3
# Multiple regression line with all possible predictor variables
m3 = lm(a3$logPlasma ~ a3$age + a3$gender + a3$smoke + a3$QI + a3$Vitamin + a3$Calories + a3$fat
        + a3$fiber + a3$alcohol + a3$CHL + a3$DBC)
summary(m3)


## Q4
# Regression with no predictors
nullmod = lm (logPlasma~1, data =a3)
# Regression with all the predictors
fullmod = m3
# Stepwise regression applying both ways
bothways = step ( nullmod , scope = list ( lower = formula ( nullmod ),upper = formula ( fullmod )),
direction ="both")
formula (bothways)
```