Using S-PLUS

Splus # To start program

sink("filename") # To save S-PLUS commands and output in a file
options(echo=T)
# Suppose X is the GPA of students in a certain class
X = c(2.1, 3.5, 2.2, 2.4, 3.3, 1.8, 4.0, 3.2, 2.1, 2.8, 2.5, 2.5, 2.6, 3.0,
1.6, 1.9, 3.2, 3.6, 2.8, 2.9, 2.5, 2.9, 2.4, 3.7, 3.2) # Entering vectors

mean.x = mean(X); var.x = var(X); sd.x = sqrt(var.x); median.x = median(X)
mean.x
[1] 2.748

cat("mean.x = ", mean.x, "Median = ", median.x, "sd = ", sd.x, fill=T)
mean.x = 2.748 Median = 2.8 sd = 0.623244200400881

stem(X); # Stem and leaf diagram

N = 25 Median = 2.8
Quartiles = 2.4, 3.2

Decimal point is at the colon

  1 : 689
  2 : 11244
  2 : 55568899
  3 : 02223
  3 : 567
  4 : 0

postscript(file="eda.ps", print.it=F) # To save graphic output in a file
par(mfrow=c(1,2))
hist(X,xlab="Histogram"); boxplot(X, xlab="Boxplot")
Y = X^2; plot(x,y)

eda function(X); # Writing functions in S-PLUS
{
  mean.x = mean(X); sd.x = sqrt(var(X));
  hist(X); stem(X); summary(X)
}
eda(X)

# Entering matrices
mat1 - matrix(c(3,1,4,5,6,7,3,4,9,7,4,22), ncol=4, byrow=T)

mat1[,1]; mat1[,2:3]; mat1[2,3]

q() # Quit program

Miscellaneous
==========

1) S BATCH infile outfile  # To send a batch file
2) source("source.file")  # To use a source file
3) Entering lists:
   list1 - scan(what=list(city="",x=0,y=0))
         Denver 12 23
         "New York" 12 34
4) Reading ASCII Text Files
   a) Vectors: vec1 - scan("vec.dat")
   b) Matrices: mat1 - matrix(scan("mat.dat",byrow=T,ncol=3)
   c) Lists: list1 - scan("list.dat",what=list(id="",var1=0,var2=0)
   d) Tables: dat1 - read.table("data",header=F, col.names=c("V1","V2"))
5) Output S-PLUS data objects as ASCII
     dump("x1","x1.file")

6) Useful functions
   rep(); is.na() # To identify missing values; X[is.na(X)] - 9999
   tapply(); sapply(); cbind(); rbind()
   par(mfrow=c(2,3)); data.frame(); table(); length(); sort()
   attach(); detach(); abline(); pairs();
   jackknife(); bootstrap()

Homework 1

A)

Consider the data auto.stats in S-PLUS. To use the data, create a data frame as follows

auto.data - data.frame(auto.stats)
attach(auto.data)
auto.data[1,] gives the first record for AMC Concord:

4099  22               3          2          2.5
Rear.Seat Trunk Weight Length Turning.Circle Displacement Gear.Ratio
27.5  11  2930  486  40  121  3.5

1. Perform EDA for the following variables: Weight, Miles.per.gallon

2. Compute jackknife estimates of bias and variance for the sample trimmed mean (with 1% trimming) of Length and Displacement.

B)