Problem 1 (Horgan book, 2.1.4, p22): It is required to estimate the number of message buffers in use in the main memory of the computer system at Roantree retailers Ltd. To do this, 20 programs were run and the number of message buffers in use was found to be

141, 146, 157, 151, 152, 140, 142, 156, 150, 140
139, 135, 143, 146, 146, 152, 140, 136, 149, 148

In R, calculate the average number of buffers used. What is the standard deviation? Would you say these data are skewed?

Problem 2 (Horgan book, 1.1 (pp11, 12), 2.1.3 (p22), 3.1.3 (pp36, 37)): In a class of 50 students of computing, 23 are female and 27 are male. The results of their first year Java examination are as follows:

Females: 57, 59, 78, 79, 60, 65, 68, 71, 75, 48, 51, 55, 56, 41, 43, 44, 75, 78, 80, 81, 83, 83, 85

Males: 48, 49, 49, 30, 30, 31, 32, 35, 37, 41, 86, 42, 51, 53, 56, 42, 44, 50, 51, 65, 67, 51, 56, 58, 64, 64, 75

(a) Read these data into R by storing them in the following ways:

- As two vectors, one for the females and one for the males.
- As one vector, with a separate factor vector designating the gender.

(b) If it was discovered that the 34th was entered incorrectly and should have obtained the mark 46 instead of 86, use an appropriate editing procedure to change this (single) mark.

(c) Obtain the summary statistics for each gender and for the entire class.

(d) Calculate the deciles for each gender and for the entire class.

(e) Form the stem and leaf display for each gender.

(f) Construct a boxplot for each gender.

Problem 3 (Horgan book, p8, 3.1.1, 2 (p36)): The examination results for a class of 119 students pursuing a computing degree are given on the website http://www.janehorgan.com as a text file called results.txt.

(a) Use these data to develop a boxplot of all the subjects on the same graph.

(b) Obtain a stem and leaf of each subject in results.txt. Does it appear that patterns are emerging?

Problem 4 (Based on the Excel file (CA2xx_survey_answers) of answers to the questionnaire (CA2xx_Questionnaire), both previously circulated):

(a) Read the survey answers into a suitable data frame in R. *It is first necessary to save the file, from within Excel, in comma separated (csv) format (other formats are also possible). Then, in R, use "read.csv" to read from the file in csv format.*

(b) Use R to find out how many males and females there are (check the questionnaire for the questions!), and to find the average height and weight of males and females.

***Note**: There will be additional problem(s) related to these survey data in later sessions.*