STA 315 2.0 Programming and Data Analysis with R

Tutorial – 2

1. Enter the following data into a vector.
   244, 191, 160, 187, 180, 176, 174, 205, 211, 183, 211, 180, 194, 200
   Name the elements as 1, 2, 3, A, B, C, D, E, F, G, H, I, J, K

2. What is the class of the object defined by the expression x <- c(4, "a", TRUE)?

3. Suppose I have a vector x <- c(3, 5, 1, 10, 12, 6) and I want to set all elements of this vector that are less than 6 to be equal to zero. What R code achieves this?

4. Write your own code to calculate standard deviation for the following data set.

Patient ID	1	2	3	4	5	6	7
Weight	60	72	57	90	95	72	70

Help: SD = $\sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n-1}}$

5. Write a code to create following matrix.

```
  a1 b1 c1
A  1  5 -2
B  1  2 -1
C  3  6 -3
```

6. If y<-matrix(c(1,3,0,9,5,-1),nrow=3,byrow=T)
   x<-matrix(c(3,4,-2,6),nrow=2,byrow=T)
   Decide what each of the following is and use R to check your answers:
   >y%*%x
   >t(x)
   >solve(x)

With x and y as above, decide what the following subscript operations are and use R to check your answers.

(a) x[1,]

(b) x[2,]

(c) x[,2]

(d) y[1,2]

(e) y[,2:3]

7. Following data represent the annual sales annual advertising expenditutes for seven companies selling mobile phones. Develop a least square simple linear regression model to predic annual sales by annual advertising expenditure.

Sales($million/year)	Advertising($million/year)
2580	1.2
11942	2.6
9845	2.2
27800	3.2
18926	2.9
4800	1.5
14550	2.7

Help: inverse of a matrix – solve()

8. Create a vector x of the elements -1, 0, 1, . . . , 6. Write code of the form x['something'], to extract

(a) the elements of x that are < 0,

(b) the elements of x that are > 0,

(c) the elements of x that are <= 3,

(d) the elements of x that are strictly less than zero or strictly greater than 4

hint: recall that | denotes 'or'

(e) the first element,

(f) the elements with indexes 2 and 4,

(g) the elements of x that are not zero

hint: '!=' denotes 'not equal to'.

9. Suppose you have the following data:

height	**58**	**59**	**60**	**61**	**62**	**63**	**64**	**65**	**66**	**67**	**68**	**69**	**70**	**71**	**72**
weight	**115**	**117**	**120**	**123**	**126**	**129**	**132**	**135**	**139**	**142**	**146**	**150**	**154**	**159**	**164**

(a) create a matrix of data from the above given measurement.

```
      [,1] [,2]
 [1,]   58  115
 [2,]   59  117
 [3,]   60  120
 [4,]   61  123
 [5,]   62  126
 [6,]   63  129
 [7,]   64  132
 [8,]   65  135
 [9,]   66  139
[10,]   67  142
[11,]   68  146
[12,]   69  150
[13,]   70  154
[14,]   71  159
[15,]   72  164
```

Your output should look like this.

(b) How would you create a matrix that has height and weight as the two rows instead of columns? You may look up the help on “matrix” if necessary.

(c) Create a data frame to store the above data set and name the columns as height and weight.

(d) How would you get R to give you the height and weight of the 8th student in the data set?

10. Built-in-dataframe

We use built-in data frames in R for our tutorials. For example, here is a built-in data frame in R, called mtcars.

>data()

>mtcars

The top line of the table, called the header, contains the column names. Each horizontal line afterward denotes a data row, which begins with the name of the row, and then followed by the actual data. Each data member of a row is called a cell.

To retrieve data in a cell, we would enter its row and column coordinates in the single square bracket "[]" operator. The two coordinates are separated by a comma. In other

words, the coordinates begins with row position, then followed by a comma, and ends with the column position. The order is important.

Here is the cell value from the first row, second column of mtcars
> mtcars[1, 2]

we can use the row and column names instead of the numeric coordinates
> mtcars["Mazda RX4", "cyl"]

the number of data rows in the data frame is given by the nrow function
> nrow(mtcars)

number of columns of a data frame is given by the ncol function
> ncol(mtcars)

details of the mtcars data set is available in the R documentation
> help(mtcars)

Instead of printing out the entire data frame, it is often desirable to preview it with the head function beforehand.
> head(mtcars)
>head(mtcars,2)
>tail(mtcars)

Challenge
Try to get descriptive statistics for “cyl”.