R computing for Business Data Analytics

Homework 1

Due date: 23:59PM October 23(Sunday), 2016

Please e-mail your homework (.pdf) and the associated *R* code (.*R*) to hchuang.om@gmail.com.

The email title must be R_HW1_GroupName. Each group submits ONLY ONE copy.

NO late homework will be accepted.

Q1. (10%) Finish the following tasks using *R*.

(a) Create a vector called **downtime**. The vector should contain the following numbers: 0, 1, 2, 12, 12, 14, 18, 21, 21, 23, 24, 25, 28, 29, 30, 30, 30, 33, 36, 44, 45, 47, and 51.

(b) Calculate the mean, median, min, max, and range of **downtime**.

(c) Calculate the standard deviation, 5 percentile, and 95 percentile of downtime (Hint: Use the *quantile* function).

(d) What is the most frequent number? What is the frequency? (Hint: Use table( ))

(e) Use *which*( ) to take the most frequent number from the **downtime** vector.

P.S.: Show all the *R* functions that you use.

Q2. (10%) Use *rep*( ) and *seq*( ) as needed to create the two vectors.

(a) 0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4

(b) 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5

Q3. (10%)

(a) Create a 4x3 matrix that stores the values below. Also name each column correctly (x, y, z)

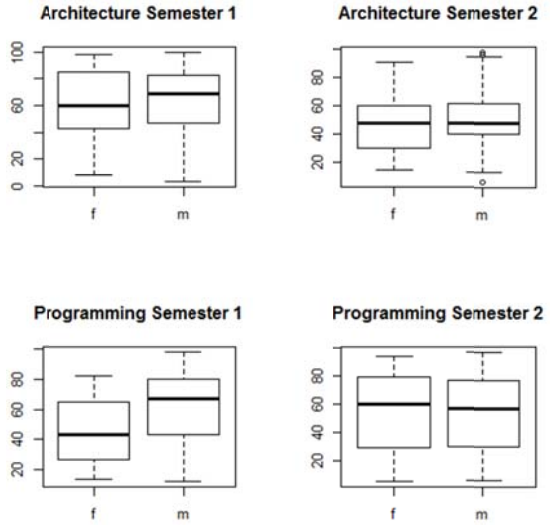
x	y	*z*
61	13	4
175	21	18
111	24	14
124	23	18

(b) Display the row 1, column 3 element of the matrix.

Q4. (10%) Calculate $\sum_{i=1}^{N} 1/i$, and compare with log(*N*)+0.6 for *N* = 500, 2000, 8000.

Q5. (10%) The equation $x^7 + 10000x^6 + 1.06x^5 + 10600x^4 + 0.0605x^3 + 605x^2 + 0.0005x + 5$ has exactly one real root. Write an *R* program to find the root. What is the root? How many iterations of Newton's method are required to find this root if the initial guess is x=0?

Q6. (10%) Based on the *results.txt* file (in Lecture 1), write *R* code to reproduce the graph below.



Q7. (10%)

(a) Compute 4!, 50!, and 5000! (Hint: Use the *factorial* function)

(b) Compute $\binom{4}{2}, \binom{50}{20}, \text{and} \binom{5000}{2000}$

(c) The *factorial* function tends to return Infinity when its argument is large. To tackle this, apply *log*( ) and *sum*( ) to compute 5000! and $\binom{5000}{2000}$. Express your answers in terms of $e^?$.

Q8. (10%)

(a) Use *R* to create a *vector* that contains all integers (整數) from 1 to 100 that are NOT divisible by 2, 3, or 7. Do NOT use loops (Hint: *which*( ) will help).

(b) Create a 10by10 *identity matrix*. That is, all diagonal (對角) elements are 1 and all remaining elements are 0 (Hint: *diag*( ) will help).

Then use two different ways to make all the non-zero elements 5 (Hint: the first can be *diag*( ) and the second can be *which*( ) + *logical comparisons*).

Q9. (10%) Use *while*( ) to write an *R* function that prints out all prime numbers <= *n* (where *n* is an integer). After writing the function, set *n*=100 and show me the results.

Q10. (10%) Consider the function *y*=*f*(*x*) defined by $y = f(x) = \begin{cases} -x^3, \forall x \le 0 \\ x^2, \forall x \in (0,1] \\ \sqrt{x}, \forall x > 1 \end{cases}$.

Write an *R* function to calculate *y* using *if* statements.

Generate the following plot for *x*=seq(-2, 2, 0.1).