Case 1
R Scripting

Part 1. R Scripting

1. Calculate the following sum:

$$x = \sum_{i=1}^{25} \left(\frac{2^i}{i} + \frac{3^i}{i^2} \right)$$

   a. What is the value of x?_____

2. Execute the following lines which create two vectors of random integers which are chosen with replacement from the integers 0, 1,… , 999. Both vectors have length 250.

```
>set.seed(50)
>xVec <- sample(0:999, 250, replace=T)
>yVec <- sample(0:999, 250, replace=T)
```

   a. Pick out the values in yVec which are > 600.
   b. What are the index positions in yVec of the values which are > 600?
   c. What are the values in xVec which correspond to the values in yVec which are > 600? (By correspond, we mean at the same index positions.)
   d. Create the vector $\left(|x_1 - \bar{x}|^{\frac{1}{2}}, |x_2 - \bar{x}|^{\frac{1}{2}}, ..., |x_{250} - \bar{x}|^{\frac{1}{2}} \right)$ where $\bar{x}$ denotes the mean of the vector xVec
   e. How many values in yVec are within 200 of the maximum value of the terms in yVec?
   f. How many numbers in xVec are divisible by 2? (Note that the modulo operator is denoted %%.)
   g. Sort the numbers in the vector xVec in the order of increasing values in yVec.
   h. Pick out the elements in yVec at index positions 1; 4; 7; 10; 13; : : : :

Part 2. Dataset Analysis

3. One of R's built-in dataset within the MASS library is the Survey Dataset. This dataset comprises of measurements and answers taken from 237 students of statistics at the University of Adelaide. The following variables are available:

- Sex: gender of student
- Wr.Hnd: span width in cm (from thumb to pinky) of the writing hand
- NW.Hnd: span width in cm (from thumb to pinky) of the non-writing hand
- W.Hnd: writing hand
- Fold: When folding your arms - which one is on top?
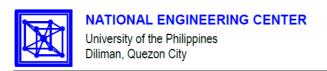- Pulse: beats per minute
- Clap: When clapping your hands - which on is on top?
- Exer: How often do you exercise?
- Smoke: How often do you smoke?
- Height: body length in cm
- M.I: Preference of either metric (cm/m) or imperial (feet/inches) units?
- Age: age in years

Run the following lines of code to load the survey data into the R environment:

```
>library(MASS)
>data(survey)
>attach(survey)
```

Answer the following questions using various R scripts:

1. What is the average pulse rate per minute of all students surveyed?
2. Do the two oldest students smoke?
3. Which factors might have an influence on the student's pulse?
4. It is generally believed that the pulse of an individual decreases with increasing age?

References: Nastasiya F. Grinberg & Robin J. Reed

Prof. Nicolai Meinshausen