

MA4413 2009 Assignment 1

Instructions

The submission deadline is 2pm Monday 26th October 2009. Answers should be typed, 12pt font with 1.5 line spacing, and submitted as a single pdf file with the following naming convention **Assign1-0123456.pdf** where the trailing digits are your UL Student Identification Number. Keep your answers brief and concise. Verbosity will be penalised. You should also provide an accompanying **.txt** file or **.R** file containing your R code. Use the same naming convention. It is in your own interests to keep your R code as well annotated as possible. This R code file should be clearly organised and use a comment line before each batch of commands to indicate which question is being considered. Email both files as attachments (not zipped) to one single email to **kevin.hayes@ul.ie**.

Pulse Data

Measure your pulse rate (beats per minute). Then toss a coin. If it comes up heads, run on the spot for one minute and then take your pulse again. If it comes up tails, sit quietly for one minute and then take your pulse again.

1. #Pulse-Q1: Load and Code

Load the data from the file **pulse.txt** into R into your own data frame called **pulse**. The following variables should be entered as **factors**, where the factor levels are given in brackets: **Ran** (1 = yes if subject

ran on the spot, 2 = no); **Smokes** (1 = yes if subject smokes regularly, 2 = no); **Sex** (1 = male, 2 = female). The variable **Activity** should be entered as an ordered factor (1 = slight, 2 = moderate, 3 = high).

2. #Pulse-Q2: Imperial to Metric

Since these data were collected in America, the heights were measured in inches and the weights in pounds. Convert these to metres and kilograms.

3. Pulse-Q3: Simple Tables

Obtain a table for the variable **Sex** – how many males and females are there in the sample? Repeat the process for **Ran**, **Smokers** and **Activity**. Do you notice any data problems?

4. Pulse-Q4: More Tables

Obtain a table cross-tabulating the variables **Sex** and **Smokers**. What percentage of males and females smoke? What percentage of people who smoke are female? What percentage of males have a high level of activity? What about females?

5. Pulse-Q5: Exploratory Data Analysis

Examine the heights using a Stem&Leaf diagram. Use this graphical display to identify the heights of the tallest and shortest persons in the class. Obtain a histogram and boxplot displays of the variable height. Describe the differences between the displays. Which do you think is most useful? Justify your answer.

6. **Pulse-Q6: Descriptive Statistics**

Calculate basic summary statistics for the initial pulse rates, `Pulse1`. Repeat the previous calculates to obtain summary information for males and females separately. Repeat this analysis for `Pulse2` – remember, the pulse rates after one minute will differ depending on whether the person ran or not.

7. **Pulse-Q7: Body Mass Index**

A person's body mass index (BMI) is defined as their weight (kg) divided by their height-squared (m^2). calculate BMI for all subjects. Save this derived variable in the data frame. Obtain a display of BMI values for the group as a whole, and then for males and females separately. Consider other groups (and crossing of groups) as appropriate.

8. **Pulse-Q8: After Minus Before Analysis**

Calculate a new variable `Change` as `Pulse2 - Pulse1` and store this new variable the data frame `pulse`. Analyse the data using the variable `change` as the primary response.

9. **Pulse-Q9: Personal Data**

Append your own pulse rate to the data frame `pulse`. Complete the additional information on sex, smoking and usual level of activity.

10. Complete the examination of the the pulse values using the commands you have learnt. Write a few paragraphs as a summary of your investigation. What effect do sex, smoking, usual level of activity and exercise have on a person's pulse rate.