```
# 2020/11/20(五), 109 學年第一學期 資料科學應用 R 小考(1)
#
# 學號: A107260028 姓名: 張雅筑
#
# 本檔案為各題之程式碼檔，無執行結果


library(readxl)
# 1(a)
my.data <- read.csv("Calculus-score-A.csv", header = TRUE, skip = 2)
xlsx_file <- "Calculus-score-B.xls"
excel_sheets(xlsx_file)
my.data1 <- read_excel(xlsx_file, sheet = "工作表 1", na = "NA", skip = 2)
my.data[c(1:5, 36:40), ]
as.data.frame(head(my.data1, 5))
as.data.frame(tail(my.data1, 5))

# 1(b)
my.data2 <- as.data.frame(my.data1)
names(my.data)[1:12] <- c("座號", "學號", "姓名", "性別", "quiz.1.", "quiz.2.",
"quiz.3.", "quiz.4.", "TA", "MidtermExam", "FinalExam", "Attendance") #change
variable name
names(my.data2)[1:12] <- c("座號", "學號", "姓名", "性別", "quiz.1.", "quiz.2.",
"quiz.3.", "quiz.4.", "TA", "MidtermExam", "FinalExam", "Attendance") #change
variable name
my.dataA <- transform(my.data,class = "A") # 增加列
my.dataB <- transform(my.data2,class = "B") # 增加列
names(my.data2) == names(my.data) #ensure names are the same
score <- rbind(my.dataA, my.dataB) #rbind two data frames.
score[38:43,]

# 1(c)
score[is.na(score)] <- 0 # 使用 is.na（）將 NA 替換為 0
Q <- score[5]*0.07 + score[6]*0.07 + score[7]*0.08 + score[8]*0.08 + score[9]*0.15 +
score[10]*0.25 + score[11]*0.30 + score[12]
x <- c(Q[1:95,])
y <- ifelse(x >= 100, 100, x)
y1 <- as.data.frame(y)
```

```
names(y1)[1] <- c("學期成績")
y1

# 1(d)
w <- ifelse(60 > y &    y >= 50, x, (sep="0"))
w1 <- as.data.frame(w)
L <- which(w1 > 0) #找某元素在向量中的下標，可以用函數 which 實現
score[L,]

# 1(e)
A <- which(score[,13] == "A")
B <- which(score[,13] == "B")
# A 班總成績平均各為多少
sum(y1[A,]) / length(A)
# B 班總成績平均各為多少
sum(y1[B,]) / length(B)
A1 <- which(score[,4] == "女")
B1 <- which(score[,4] == "男")
#  女生總成績平均各為多少
sum(y1[A1,]) / length(A1)
#  男生總成績平均各為多少
sum(y1[B1,]) / length(B1)

# 1(f)
A2 <- ifelse(60 > y &    score[,13] == "A", x, (sep="0"))
A3 <- as.data.frame(A2)
A4 <- which(A3 > 0)
# A  班學期成績不及格比例為多少?
length(A4) / length(A)
B2 <- ifelse(60 > y & score[,13] == "B" & score[,4] == "男", x, (sep="0"))
B3 <- as.data.frame(B2)
B4 <- which(B3 > 0)
# B  班男同學學期成績不及格比例為多少?
length(B4) / length(B)

# 1(g)
score1 <- transform(score,score = y1)
names(score1)[14] <- c("score")
```

```
SG <- score1[A1,]
SB <- score1[B1,]
SG1 <- order(SG$score, decreasing = TRUE)
SB1 <- order(SB$score, decreasing = TRUE)
SG2 <- SG[SG1,]
SB2 <- SB[SB1,]
head(SG2, 5)
head(SB2, 5)

# 2(a)
set.seed <- c(123456)
Letters.code <- c(sample(LETTERS[1:5], 20, replace=T))
i <- c(1:length(Letters.code))
Numbers.code <- ifelse(Letters.code[i] == "A" , "1",
                       ifelse(Letters.code[i] == "B" , "3",
                              ifelse(Letters.code[i] == "C" , "2",
                                     ifelse(Letters.code[i] == "D" , "3",
                                            ifelse(Letters.code[i] == "E" , "1",
NA)))))

# 2(b)
survey <- data.frame(Letters.code = Letters.code, Numbers.code = Numbers.code)
survey
```