

#2020/11/27(五), 109 學年第一學期 資料科學應用 HW4

#學號: a107260012 姓名: 江鴻麟

```
library(readxl)
```

# 1(a)

```
data <- read.csv("Calculus-score-A.csv", header = TRUE, skip = 2)
```

```
xlsx_file <- "Calculus-score-B.xls"
```

```
excel_sheets(xlsx_file)
```

```
"工作表 1"
```

```
c <- read_excel(xlsx_file, sheet = "工作表 1", na = "NA", skip = 2)
```

```
data[c(1:5, 36:40), ]
```

```
as.data.frame(head(c, 5))
```

```
as.data.frame(tail(c, 5))
```

# 1(b)

```
j <- as.data.frame(c)
```

```
names(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(j)[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
```

```
dataA <- transform(data, class = "A") # 增加列
```

```
dataB <- transform(j, class = "B") # 增加列
```

```
names(j) == names(data) #ensure names are the same
```

```
h <- rbind(dataA, dataB) #rbind two data frames.
```

```
h[38:43,]
```

# 1(c)

```
h[is.na(h)] <- 0 # 使用 is.na ( ) 將 NA 替換為 0
```

```
Q <- h[5]*0.07 + h[6]*0.07 + h[7]*0.08 + h[8]*0.08 + h[9]*0.15 + h[10]*0.25 +  
h[11]*0.30 + h[12]
```

```
k <- c(Q[1:95,])
```

```
y <- ifelse(k >= 100, 100, k)
```

```
l <- as.data.frame(y)
```

```
names(l)[1] <- c("學期成績")
```

|

# 1(d)

```
r <- ifelse(60 > y & y >= 50, k, (sep="0"))
```

```
t <- as.data.frame(r)
```

```
L <- which(t > 0) #找某元素在向量中的下標，可以用函數 which 實現  
h[L,]
```

# 1(e)

```
A <- which(h[,13] == "A")
```

```
B <- which(h[,13] == "B")
```

# A 班總成績平均各為多少

```
sum(l[A,]) / length(A)
```

# B 班總成績平均各為多少

```
sum(l[B,]) / length(B)
```

```
P <- which(h[,4] == "女")
```

```
E <- which(h[,4] == "男")
```

# 女生總成績平均各為多少

```
sum(l[P,]) / length(P)
```

# 男生總成績平均各為多少

```
sum(l[E,]) / length(E)
```

# 1(f)

```
A6 <- ifelse(60 > y & h[,13] == "A", k, (sep="0"))
```

```
A7 <- as.data.frame(A6)
```

```
A8 <- which(A7 > 0)
```

# A 班學期成績不及格比例為多少?

```
length(A8) / length(A)
```

```
B3 <- ifelse(60 > y & h[,13] == "B" & h[,4] == "男", k, (sep="0"))
```

```
B8 <- as.data.frame(B3)
```

```
B5 <- which(B8 > 0)
```

# B 班男同學學期成績不及格比例為多少?

```
length(B5) / length(B)
```

```

# 1(g)
L1 <- transform(score,score = y1)
names(L1)[14] <- c("score")
SJ <- L1[A1,]
DB <- L1[B1,]
SG1 <- order(SJ$score, decreasing = TRUE)
SK1 <- order(DB$score, decreasing = TRUE)
SG2 <- SJ[SG1,]
SB2 <- DB[SK1,]
head(SG2, 5)
0
head(SB2, 5)
3
# 2(a)
set.seed <- c(123456)
y <- c(sample(LETTERS[1:5], 20, replace=T))
x <-c()
for(i in 1:20){
  if(y[i] == "A")
    x[i] <- 1
  else if(y[i] == "E")
    x[i] <- 1
  else if(y[i] == "C")
    x[i] <- 2
  else
    x[i] <- 3
}
cat(x)

# 2(b)
MM <- data.frame(Letters.code = y, Numbers.code = x)
MM

```