

#2020/11/27(五), 109 學年第一學期 資料科學應用 HW(4)

#學號: a107260008 姓名: 林玥浩

```
library(readxl)
```

1(a)

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

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

```
excel_sheets(xlsx_file)
```

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

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

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

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

1(b)

```
g <- as.data.frame(g)
```

```
names(a)[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(g)[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(a, class = "A") # 增加列
```

```
my.dataB <- transform(g, class = "B") # 增加列
```

```
names(g) == names(a) #ensure names are the same
```

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

```
hh[38:43,]
```

1(c)

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

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

```
xx <- c(qq[1:95,])
```

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

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

```
names(yue)[1] <- c("學期成績")
```

```
yue
```

1(d)

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

```
wg <- as.data.frame(weiy)
L <- which(wg > 0) #找某元素在向量中的下標，可以用函數 which 實現
hh[L,]
```

```
# 1(e)
A <- which(hh[,13] == "A")
B <- which(hh[,13] == "B")
sum(y1[A,]) / length(A)
sum(y1[B,]) / length(B)
A1 <- which(hh[,4] == "女")
B1 <- which(hh[,4] == "男")
sum(y1[A1,]) / length(A1)
sum(y1[B1,]) / length(B1)
```

```
# 1(f)
Ag <- ifelse(60 > y & hh[,13] == "A", x, (sep="0"))
Aw <- as.data.frame(Ag)
Aq <- which(Aw > 0)
length(Aw) / length(A)
nw <- ifelse(60 > y & hh[,13] == "B" & hh[,4] == "男", x, (sep="0"))
ny <- as.data.frame(nw)
nm <- which(ny > 0)
length(nm) / length(B)
```

```
# 1(g)
run <- transform(hh, hh = y1)
names(run)[14] <- c("hh")
oc <- run[A1,]
ow <- run[B1,]
oc1 <- order(oc$hh, decreasing = TRUE)
ow1 <- order(ow$hh, decreasing = TRUE)
oc2 <- oc[oc1,]
ow2 <- ow[ow1,]
head(oc2, 5)
head(ow2, 5)
```

```
# 2(a)
```

```
set.seed <- c(123456)
dwp <- c(sample(LETTERS[1:5], 20, replace=T))
e <-c()
for(i in 1:20){
  if(dwp[i] == "A")
    e[i] <- 1
  else if(dwp[i] == "E")
    e[i] <- 1
  else if(dwp[i] == "C")
    e[i] <- 2
  else
    e[i] <- 3
}
cat(e)

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