```
> #2020/11/27(五), 109 學年第一學期 資料科學應用 HW4
> #學號: a107260086          姓名: 張允銓
> library(readxl)
> # 1(a)
> Data <- read.csv("Calculus-score-A.csv", header = TRUE, skip = 2)
> xlsx <- "Calculus-score-B.xls"
> excel_sheets(xlsx)
[1] "工作表 1"
> Data1 <- read_excel(xlsx, sheet = "工作表 1", na = "NA", skip = 2)
New names:
* `0.070000000000000007` -> `0.070000000000000007...5`
* `0.070000000000000007` -> `0.070000000000000007...6`
* `0.080000000000000002` -> `0.080000000000000002...7`
* `0.080000000000000002` -> `0.080000000000000002...8`
> Data[c(1:5, 36:40), ]
    座號         學號     姓名 性別 X7. X7..1 X8. X8..1   X15. X25. X30. Times
1       1 401405008 希瑄彥     男   10       0    5      20     0.0     55     50
2
2       2 401550880 張泓丞     男   25      40   70      87    80.0     46     68
9
3       3 404550061 張安婕     女   18      15   48      33    86.7     54     79
9
4       4 404550042 柯政學     男   10      10   NA      NA    13.3      2      0
7
5       5 404550023 謝文躍     女   35      45   52      97    86.7     55     67
9
36     36 404550369 陳王霖     女   55      73   92      73   100.0     72     81
9
37     37 404550420 何瑄穎     男   28      10   35       3    66.7     30      0
7
38     38 404550431 沈泓霏     女   15      25   53      67    93.3     29     42
9
39     39 404550442 許安霏     女   53      60   80      72   100.0     61     62
9
40     40 404550453 李政宜     男   80     100   85     100   100.0     95    100
3
> as.data.frame(head(Data1, 5))
  座號         學號     姓名 性別 0.070000000000000007...5
```

```
1       1 404550465  史文羽     男                              60
2       2 404685071  鄭樺妤     男                              80
3       3 404685084  張敬安     男                              10
4       4 404685099  何筑亦     女                              15
5       5 404685100  張　儀     女                              30
   0.070000000000000007...6 0.080000000000000002...7
0.080000000000000002...8
1                              81                            100
97
2                             100                            100
92
3                              40                             62
93
4                              25                             40
13
5                              45                             70
61
   0.14999999999999999 0.25 0.29999999999999999 Times
1                    100.0     90                    83        6
2                    100.0     92                    97        2
3                    100.0     65                    84        9
4                     93.3     36                     5        9
5                     93.3     29                    48        4
> as.data.frame(tail(Data1, 5))
   座號          學號     姓名  性別  0.070000000000000007...5
1     51 404685407  鄭鈺尤     女                              80
2     52 404685905  楊宜路     男                              48
3     53 404685013  張渝妤     男                               0
4     54 404685119  廖暄安     男                              50
5     55 499555916  楊毅亦     女                               5
   0.070000000000000007...6 0.080000000000000002...7
0.080000000000000002...8
1                              85                            100
85
2                              35                             48
98
3                              38                             60
40
```

```
4                                70                                20
85
5                                35                                45
55
   0.14999999999999999 0.25 0.29999999999999999 Times
1                   100     89                  95        9
2                   100     50                  62        9
3                    87     49                  25        1
4                   100     54                  69        4
5                    87     58                  60        3
>
> # 1(b)
> Data2 <- as.data.frame(Data1)
> 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(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
> DataA <- transform(Data,class = "A") #  增加列
> DataB <- transform(Data2,class = "B") #  增加列
> names(Data2) == names(Data) #ensure names are the same
 [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
> score <- rbind(DataA, DataB) #rbind two data frames.
> score[38:43,]
    座號       學號   姓名 性別 quiz.1. quiz.2. quiz.3. quiz.4.      TA
38    38 404550431 沈泓霏   女       15       25       53       67   93.3
39    39 404550442 許安霏   女       53       60       80       72  100.0
40    40 404550453 李政宜   男       80      100       85      100  100.0
41     1 404550465 史文羽   男       60       81      100       97  100.0
42     2 404685071 鄭樺妤   男       80      100      100       92  100.0
43     3 404685084 張敬安   男       10       40       62       93  100.0
    MidtermExam FinalExam Attendance class
38           29        42          9     A
39           61        62          9     A
40           95       100          3     A
41           90        83          6     B
42           92        97          2     B
```

```
43                65              84                9         B
>
> # 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]
> E <- c(Q[1:95,])
> F <- ifelse(E >= 100, 100, E)
> F1 <- as.data.frame(F)
> names(F1)[1] <- c("學期成績")
> F1
    學期成績
1      33.450
2      70.010
3      67.995
4      10.895
5      73.375
6      67.015
7      38.920
8      78.345
9      48.525
10     20.455
11     96.435
12     67.295
13     16.150
14     22.030
15     73.990
16     79.400
17      8.560
18     24.245
19     61.905
20     61.340
21     49.915
22     68.570
23     10.995
24     67.055
25     68.000
26     69.210
```

27	65.635
28	84.040
29	66.100
30	78.540
31	75.330
32	69.860
33	72.240
34	82.260
35	54.765
36	88.460
37	30.205
38	55.245
39	77.920
40	99.150
41	94.030
42	97.060
43	81.350
44	40.535
45	55.375
46	62.355
47	61.310
48	50.450
49	21.600
50	36.700
51	32.150
52	76.810
53	48.200
54	52.550
55	69.700
56	43.360
57	60.910
58	94.070
59	77.990
60	23.950
61	39.100
62	80.600
63	72.850
64	22.050

```
65     47.200
66     20.800
67     61.550
68     58.300
69     40.800
70     55.000
71     26.280
72     70.050
73     49.450
74     62.900
75     54.960
76     74.900
77     71.360
78     67.800
79     85.140
80     21.300
81     72.200
82     78.410
83     82.300
84     51.510
85     74.660
86     45.200
87     65.300
88     87.220
89    100.000
90     95.720
91    100.000
92     72.590
93     44.460
94     70.000
95     59.350
>
> # 1(d)
> w <- ifelse(60 > F &    F >= 50, E, (sep="0"))
> w1 <- as.data.frame(w)
> P <- which(w1 > 0) #找某元素在向量中的下標，可以用函數 which 實現
> score[P,]
    座號          學號     姓名 性別  quiz.1. quiz.2. quiz.3. quiz.4.      TA
```

```
35     35 404550328 李梅祐   男        20        25        55        32 86.7
38     38 404550431 沈泓霏   女        15        25        53        67 93.3
45      5 404685100 張　儀   女        30        45        70        61 93.3
48      8 403555042 張水兆   男         0        30        50        80 13.0
54     14 404720027 高凱瓊   男        15        40        35        60 80.0
68     28 404720722 楊佳聿   女        30        35        20        50 60.0
70     30 404720527 馨飛羽   男        15         0        45        65 93.0
75     35 404720037 勳　陳   男        20        30        22        60 80.0
84     44 404720932 曾林凱   女        55        18        30        50 80.0
95     55 499555916 楊毅亦   女         5        35        45        55 87.0
     MidtermExam FinalExam Attendance class
35            41        48          7     A
38            29        42          9     A
45            29        48          4     B
48             0        90          9     B
54            42        32          9     B
68            45        63          9     B
70            44        44          7     B
75            50        38          9     B
84            58        15          9     B
95            58        60          3     B
>
> # 1(e)
> A <- which(score[,13] == "A")
> B <- which(score[,13] == "B")
> # A 班總成績平均各為多少
> sum(F1[A,]) / length(A)
[1] 58.84575
> # B 班總成績平均各為多少
> sum(F1[B,]) / length(B)
[1] 61.123
> A1 <- which(score[,4] == "女")
> B1 <- which(score[,4] == "男")
> #  女生總成績平均各為多少
> sum(F1[A1,]) / length(A1)
[1] 58.95292
> #  男生總成績平均各為多少
> sum(F1[B1,]) / length(B1)
```

```
[1] 60.90322
>
> # 1(f)
> A2 <- ifelse(60 > F &    score[,13] == "A", E, (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)
[1] 0.35
> B2 <- ifelse(60 > F & score[,13] == "B" & score[,4] == "男", E, (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] 0.2909091
>
> # 1(g)
> score1 <- transform(score,score = F1)
> names(score1)[14] <- c("score")
> SK <- score1[A1,]
> ST <- score1[B1,]
> SK1 <- order(SK$score, decreasing = TRUE)
> ST1 <- order(ST$score, decreasing = TRUE)
> SK2 <- SK[SK1,]
> ST2 <- ST[ST1,]
> head(SK2, 5)
    座號       學號     姓名 性別 quiz.1. quiz.2. quiz.3. quiz.4.       TA
89    49 404720541  詹傑仙    女       98       80       98       98 100.0
91    51 404685407  鄭鈺尤    女       80       85      100       85 100.0
11    11 404550189  丁易偉    女       80      100      100       93    93.3
36    36 404550369  陳王霖    女       55       73       92       73 100.0
79    39 404720436  曼李儷    女       60       40       73       90 100.0
    MidtermExam FinalExam Attendance class     score
89           96        95          9     B 100.000
91           89        95          9     B 100.000
11           90        93          4     A    96.435
36           72        81          9     A    88.460
79           68        87          7     B    85.140
```

```
> head(ST2, 5)
    座號       學號    姓名 性別 quiz.1. quiz.2. quiz.3. quiz.4.   TA MidtermExam
40    40 404550453 李政宜   男       80      100       85      100 100
95
42     2 404685071 鄭樺妤   男       80      100      100       92 100
92
90    50 404685109 許　何   男       88       73       85      100 100
83
58    18 404720161 劉莞韋   男       95       86       85       75 100
80
41     1 404550465 史文羽   男       60       81      100       97 100
90
    FinalExam Attendance class score
40        100          3     A 99.15
42         97          2     B 97.06
90         83          9     B 95.72
58         82          9     B 94.07
41         83          6     B 94.03
>
> # 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 2 3 1 3 1 1 2 3 3 3 2 1 3 2 2 1 1 1 3>
> # 2(b)
> LN <- data.frame(Letters.code = y, Numbers.code = x)
> LN
```

	Letters.code	Numbers.code
1	C	2
2	C	2
3	B	3
4	E	1
5	B	3
6	A	1
7	A	1
8	C	2
9	D	3
10	D	3
11	D	3
12	C	2
13	E	1
14	D	3
15	C	2
16	C	2
17	E	1
18	E	1
19	A	1
20	B	3

>