```
> #2020/11/27(五), 109 學年第一學期 資料科學應用 HW(4)
> #學號: a107260008 姓名: 林玥澔
> 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)
[1] "工作表 1"
> b <- read_excel(xlsx_file, 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`
> a[c(1:5, 36:40), ]
    座號          學號      姓名 性別 X7. X7..1 X8.
1       1 401405008   希瑄彥    男   10     0   5
2       2 401550880   張泓丞    男   25    40  70
3       3 404550061   張安婕    女   18    15  48
4       4 404550042   柯政學    男   10    10  NA
5       5 404550023   謝文躍    女   35    45  52
36     36 404550369   陳王霖    女   55    73  92
37     37 404550420   何瑄穎    男   28    10  35
38     38 404550431   沈泓霏    女   15    25  53
39     39 404550442   許安霏    女   53    60  80
40     40 404550453   李政宜    男   80   100  85
    X8..1   X15. X25. X30. Times
1      20    0.0   55   50     2
2      87   80.0   46   68     9
3      33   86.7   54   79     9
4      NA   13.3    2    0     7
5      97   86.7   55   67     9
36     73  100.0   72   81     9
37      3   66.7   30    0     7
38     67   93.3   29   42     9
39     72  100.0   61   62     9
40    100  100.0   95  100     3
> as.data.frame(head(b, 5))
```

	座號	學號	姓名	性別
1	1	404550465	史文羽	男
2	2	404685071	鄭樺妤	男
3	3	404685084	張敬安	男
4	4	404685099	何筑亦	女
5	5	404685100	張　儀	女

	0.070000000000000007...5
1	60
2	80
3	10
4	15
5	30

	0.070000000000000007...6
1	81
2	100
3	40
4	25
5	45

	0.080000000000000002...7
1	100
2	100
3	62
4	40
5	70

	0.080000000000000002...8	0.14999999999999999
1	97	100.0
2	92	100.0
3	93	100.0
4	13	93.3
5	61	93.3

	0.25	0.29999999999999999	Times
1	90	83	6
2	92	97	2
3	65	84	9
4	36	5	9
5	29	48	4

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

	座號	學號	姓名	性別

```
1      51 404685407 鄭鈺尤    女
2      52 404685905 楊宜路    男
3      53 404685013 張渝妤    男
4      54 404685119 廖暄安    男
5      55 499555916 楊毅亦    女
  0.070000000000000007...5
1                      80
2                      48
3                       0
4                      50
5                       5
  0.070000000000000007...6
1                      85
2                      35
3                      38
4                      70
5                      35
  0.080000000000000002...7
1                     100
2                      48
3                      60
4                      20
5                      45
  0.080000000000000002...8 0.14999999999999999
1                      85                  100
2                      98                  100
3                      40                   87
4                      85                  100
5                      55                   87
  0.25 0.29999999999999999 Times
1   89                  95  9
2   50                  62  9
3   49                  25  1
4   54                  69  4
5   58                  60  3
> # 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
> 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
> 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
 [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
 [9] TRUE TRUE TRUE TRUE
> hh <- rbind(my.dataA, my.dataB) #rbind two data frames.
> hh[38:43,]
```

	座號	學號	姓名	性別	quiz.1.	quiz.2.
38	38	404550431	沈泓霏	女	15	25
39	39	404550442	許安霏	女	53	60
40	40	404550453	李政宜	男	80	100
41	1	404550465	史文羽	男	60	81
42	2	404685071	鄭樺妤	男	80	100
43	3	404685084	張敬安	男	10	40

	quiz.3.	quiz.4.	TA	MidtermExam	FinalExam
38	53	67	93.3	29	42
39	80	72	100.0	61	62
40	85	100	100.0	95	100
41	100	97	100.0	90	83
42	100	92	100.0	92	97
43	62	93	100.0	65	84

	Attendance	class
38	9	A
39	9	A
40	3	A
41	6	B
42	2	B
43	9	B

```
> # 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]
```

	FinalExam
1	17.0
2	29.4
3	32.7
4	7.0
5	29.1
6	23.4
7	14.1
8	32.1
9	16.9
10	1.0
11	31.9
12	24.9
13	2.0
14	6.0
15	30.6
16	29.5
17	0.0
18	3.0
19	27.6
20	28.1
21	20.1
22	25.2
23	3.0
24	26.5
25	35.4
26	34.5
27	31.8
28	36.9
29	30.3
30	28.5
31	30.0
32	23.1
33	33.4
34	35.4
35	21.4
36	33.3
37	7.0

38	21.6
39	27.6
40	33.0
41	30.9
42	31.1
43	34.2
44	10.5
45	18.4
46	18.9
47	19.1
48	36.0
49	9.0
50	15.4
51	13.5
52	27.9
53	24.6
54	18.6
55	20.5
56	12.6
57	20.5
58	33.6
59	26.0
60	9.8
61	16.2
62	26.5
63	19.2
64	3.0
65	15.0
66	9.0
67	26.1
68	27.9
69	12.0
70	20.2
71	9.0
72	25.5
73	22.8
74	24.0
75	20.4

76	30.0
77	22.8
78	29.4
79	33.1
80	9.0
81	27.4
82	31.5
83	26.5
84	13.5
85	24.9
86	13.5
87	25.8
88	30.5
89	37.5
90	33.9
91	37.5
92	27.6
93	8.5
94	24.7
95	21.0

```
> 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
```

	學期成績
1	16.450
2	40.610
3	35.295
4	3.895
5	44.275
6	43.615
7	24.820
8	46.245
9	31.625
10	19.455
11	64.535
12	42.395

13	14.150
14	16.030
15	43.390
16	49.900
17	8.560
18	21.245
19	34.305
20	33.240
21	29.815
22	43.370
23	7.995
24	40.555
25	32.600
26	34.710
27	33.835
28	47.140
29	35.800
30	50.040
31	45.330
32	46.760
33	38.840
34	46.860
35	33.365
36	55.160
37	23.205
38	33.645
39	50.320
40	66.150
41	63.130
42	65.960
43	47.150
44	30.035
45	36.975
46	43.455
47	42.210
48	14.450
49	12.600
50	21.300

51	18.650
52	48.910
53	23.600
54	33.950
55	49.200
56	30.760
57	40.410
58	60.470
59	51.990
60	14.150
61	22.900
62	54.100
63	53.650
64	19.050
65	32.200
66	11.800
67	35.450
68	30.400
69	28.800
70	34.800
71	17.280
72	44.550
73	26.650
74	38.900
75	34.560
76	44.900
77	48.560
78	38.400
79	52.040
80	12.300
81	44.800
82	46.910
83	55.800
84	38.010
85	49.760
86	31.700
87	39.500
88	56.720

89	67.140
90	61.820
91	63.600
92	44.990
93	35.960
94	45.300
95	38.350

```
> # 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,]
```

	座號	學號	姓名	性別	quiz.1.	quiz.2.
30	30	404550328	梅林茹	男	25	65
36	36	404550369	陳王霖	女	55	73
39	39	404550442	許安霏	女	53	60
59	19	404720172	曾易佳	男	80	65
62	22	404720106	詹書飛	男	65	80
63	23	404720117	劉徐卉	男	65	90
79	39	404720436	曼李儷	女	60	40
83	43	404720834	張雅益	女	60	75
88	48	404720343	廖林韋	男	96	55

	quiz.3.	quiz.4.	TA	MidtermExam	FinalExam
30	62	41	100	82	65
36	92	73	100	72	81
39	80	72	100	61	62
59	98	75	80	64	80
62	80	85	100	63	65
63	70	65	100	68	64
79	73	90	100	68	87
83	85	85	100	71	65
88	85	70	100	75	95

	Attendance	class
30	9	A
36	9	A
39	9	A
59	2	B
62	7	B

63	0	B
79	7	B
83	7	B
88	2	B

```r
> # 1(e)
> A <- which(hh[,13] == "A")
> B <- which(hh[,13] == "B")
> sum(y1[A,]) / length(A)
[1] 58.84575
> sum(y1[B,]) / length(B)
[1] 61.123
> A1 <- which(hh[,4] == "女")
> B1 <- which(hh[,4] == "男")
> sum(y1[A1,]) / length(A1)
[1] 58.95292
> sum(y1[B1,]) / length(B1)
[1] 60.90322
> # 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)
[1] 0.025
> 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] 0.6363636
> # 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)
```

	座號	學號	姓名	性別	quiz.1.	quiz.2.
89	49	404720541	詹傑仙	女	98	80
91	51	404685407	鄭鈺尤	女	80	85
11	11	404550189	丁易偉	女	80	100
36	36	404550369	陳王霖	女	55	73
79	39	404720436	曼李儷	女	60	40

	quiz.3.	quiz.4.	TA	MidtermExam	FinalExam
89	98	98	100.0	96	95
91	100	85	100.0	89	95
11	100	93	93.3	90	93
36	92	73	100.0	72	81
79	73	90	100.0	68	87

	Attendance	class	hh
89	9	B	100.000
91	9	B	100.000
11	4	A	96.435
36	9	A	88.460
79	7	B	85.140

```
> head(ow2, 5)
```

	座號	學號	姓名	性別	quiz.1.	quiz.2.
40	40	404550453	李政宜	男	80	100
42	2	404685071	鄭樺妤	男	80	100
90	50	404685109	許 何	男	88	73
58	18	404720161	劉莞韋	男	95	86
41	1	404550465	史文羽	男	60	81

	quiz.3.	quiz.4.	TA	MidtermExam	FinalExam
40	85	100	100	95	100
42	100	92	100	92	97
90	85	100	100	83	83
58	85	75	100	80	82
41	100	97	100	90	83

	Attendance	class	hh
40	3	A	99.15
42	2	B	97.06
90	9	B	95.72
58	9	B	94.07
41	6	B	94.03

```
> # 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)
1 3 1 1 1 1 3 2 3 2 3 1 1 1 3 1 3 2 1 3> # 2(b)
> dw <- data.frame(Letters.code = dwp, Numbers.code = e)
> dw
     Letters.code Numbers.code
1              E            1
2              D            3
3              E            1
4              A            1
5              E            1
6              E            1
7              B            3
8              C            2
9              D            3
10             C            2
11             B            3
12             E            1
13             A            1
14             E            1
15             D            3
16             A            1
17             D            3
18             C            2
19             A            1
20             D            3
```

\>