```
> #2020/11/27(五), 109 學年第一學期 資料科學應用 R 作業(4)
> #學號:A107260102        姓名: 熊家瀋
> 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)
[1] "工作表 1"
> O <- 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`
> 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(O, 5))
   座號      學號      姓名  性別  0.070000000000000007...5
```

0.070000000000000007...6

1	1	404550465	史文羽	男	60	81
2	2	404685071	鄭樺妤	男	80	100
3	3	404685084	張敬安	男	10	40
4	4	404685099	何筑亦	女	15	25
5	5	404685100	張　儀	女	30	45

0.080000000000000002...7 0.080000000000000002...8 0.14999999999999999 0.25

1	100	97	100.0	90
2	100	92	100.0	92
3	62	93	100.0	65
4	40	13	93.3	36
5	70	61	93.3	29

0.29999999999999999 Times

1	83	6
2	97	2
3	84	9
4	5	9
5	48	4

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

座號　　學號　姓名　性別　0.070000000000000007...5 0.070000000000000007...6

1	51	404685407	鄭鈺尤	女	80	85
2	52	404685905	楊宜路	男	48	35
3	53	404685013	張渝妤	男	0	38

| 4 | 54 | 404685119 | 廖暄安 | 男 | 50 |
| 70 |
| 5 | 55 | 499555916 | 楊毅亦 | 女 | 5 |
| 35 |

| | 0.080000000000000002...7 | 0.080000000000000002...8 | 0.14999999999999999 |
| 0.25 |
1	100	85
100	89
2	48	98
100	50
3	60	40
87	49
4	20	85
100	54
5	45	55
87	58

	0.29999999999999999	Times
1	95	9
2	62	9
3	25	1
4	69	4
5	60	3

```
> # 1(b)
> j <- as.data.frame(O)
> 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
> o <- transform(data,class = "A") # 增加列
> ob <- transform(j,class = "B") # 增加列
> names(j) == names(data) #ensure names are the same
  [1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
> gg <- rbind(o, ob) #rbind two data frames.
> gg[38:43,]
    座號       學號     姓名 性別  quiz.1. quiz.2. quiz.3. quiz.4.       TA
MidtermExam
38    38 404550431  沈泓霏     女        15      25      53      67  93.3
```

29

39	39 404550442	許安霏	女	53	60	80	72	100.0

61

40	40 404550453	李政宜	男	80	100	85	100	100.0

95

41	1 404550465	史文羽	男	60	81	100	97	100.0

90

42	2 404685071	鄭樺妤	男	80	100	100	92	100.0

92

43	3 404685084	張敬安	男	10	40	62	93	100.0

65

	FinalExam	Attendance	class
38	42	9	A
39	62	9	A
40	100	3	A
41	83	6	B
42	97	2	B
43	84	9	B

```
> # 1(c)
> gg[is.na(gg)] <- 0 # 使用 is.na（）將 NA 替換為 0
> QQ <- gg[5]*0.07 + gg[6]*0.07 + gg[7]*0.08 + gg[8]*0.08 + gg[9]*0.15 +
gg[10]*0.25 + gg[11]*0.30 + gg[12]
> kK <- c(QQ[1:95,])
> y <- ifelse(kK >= 100, 100, kK)
> l <- as.data.frame(y)
> names(l)[1] <- c("學期成績")
> l
```

	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)
> TT <- ifelse(60 > y &    y >= 50, k, (sep="0"))
> t <- as.data.frame(TT)
> L <- which(t > 0) #找某元素在向量中的下標，可以用函數 which 實現
> gg[L,]
    座號      學號      姓名 性別 quiz.1. quiz.2. quiz.3. quiz.4.    TA
MidtermExam
```

	座號	學號	姓名	性別	quiz.1.	quiz.2.	quiz.3.	quiz.4.	TA	MidtermExam
35	35 404550328	李梅祐	男	20	25	55	32	86.7	41
38	38 404550431	沈泓霏	女	15	25	53	67	93.3	29
45	5 404685100	張　儀	女	30	45	70	61	93.3	29
48	8 403555042	張水兆	男	0	30	50	80	13.0	0
54	14 404720027	高凱瓊	男	15	40	35	60	80.0	42
68	28 404720722	楊佳聿	女	30	35	20	50	60.0	45
70	30 404720527	馨飛羽	男	15	0	45	65	93.0	44
75	35 404720037	勳　陳	男	20	30	22	60	80.0	50
84	44 404720932	曾林凱	女	55	18	30	50	80.0	58
95	55 499555916	楊毅亦	女	5	35	45	55	87.0	58

```
    FinalExam Attendance class
35         48            7       A
```

38	42	9	A
45	48	4	B
48	90	9	B
54	32	9	B
68	63	9	B
70	44	7	B
75	38	9	B
84	15	9	B
95	60	3	B

```r
> # 1(e)
> A <- which(gg[,13] == "A")
> B <- which(gg[,13] == "B")
> # A 班總成績平均各為多少
> sum(l[A,]) / length(A)
[1] 58.84575
> # B 班總成績平均各為多少
> sum(l[B,]) / length(B)
[1] 61.123
> PP <- which(h[,4] == "女")
> EK <- which(h[,4] == "男")
> # 女生總成績平均各為多少
> sum(l[PP,]) / length(PP)
[1] 58.95292
> # 男生總成績平均各為多少
> sum(l[EK,]) / length(EK)
[1] 60.90322
> # 1(f)
> AC <- ifelse(60 > y &    h[,13] == "A", k, (sep="0"))
> AG <- as.data.frame(AC)
> AH <- which(AG > 0)
> # A 班學期成績不及格比例為多少?
> length(AH) / length(A)
[1] 0.35
> BN <- ifelse(60 > y & h[,13] == "B" & h[,4] == "男", k, (sep="0"))
> BH <- as.data.frame(BN)
> BG <- which(BH > 0)
> # B 班男同學學期成績不及格比例為多少?
> length(BG) / length(B)
```

```
[1] 0.2909091
> # 1(g)
> J1 <- transform(score,score = y1)
> names(J1)[14] <- c("score")
> SJ <- J1[A1,]
> DB <- J1[B1,]
> S1 <- order(SJ$score, decreasing = TRUE)
> SD1 <- order(DB$score, decreasing = TRUE)
> S2 <- SJ[S1,]
> SC2 <- DB[SD1,]
> head(S2, 5)
      座號      學號      姓名 性別 quiz.1. quiz.2. quiz.3. quiz.4.      TA
MidtermExam
89    49 404720541 詹傑仙    女      98      80      98      98 100.0
96
91    51 404685407 鄭鈺尤    女      80      85     100      85 100.0
89
11    11 404550189 丁易偉    女      80     100     100      93   93.3
90
36    36 404550369 陳王霖    女      55      73      92      73 100.0
72
79    39 404720436 曼李儷    女      60      40      73      90 100.0
68
      FinalExam Attendance class      score
89         95         9     B 100.000
91         95         9     B 100.000
11         93         4     A   96.435
36         81         9     A   88.460
79         87         7     B   85.140
> head(SC2, 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 劉莞葦    男      95      86      85      75 100
```

80

41   1 404550465 史文羽   男   60   81   100   97 100

90

FinalExam Attendance class score

	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)
> Yy <- c(sample(LETTERS[1:5], 20, replace=T))
> Qx <-c()
> for(i in 1:20){
+     if(Yy[i] == "A")
+         Qx[i] <- 1
+
```