

```
> #2020/11/20(五), 109 學年第一學期 資料科學應用 R 小考(1)
```

```
> #學號:A107260036 姓名:顏郁芹
```

```
> #ex.1(a)
```

```
> x <- read.csv("Calculus-score-A.csv",skip=2)
```

```
> names(x) <- c("座號", "學號", "姓名", "性別", "quiz(1)", "quiz(2)", "quiz(3)",  
"quiz(4)", "TA", "MidtermExam", "FinalExam", "Attendance")
```

```
> head(x, 5)
```

	座號	學號	姓名	性別	quiz(1)	quiz(2)	quiz(3)	quiz(4)	TA
MidtermExam									
1	1	401405008	希瑄彥	男	10	0	5	20	0.0
55									
2	2	401550880	張泓丞	男	25	40	70	87	80.0
46									
3	3	404550061	張安婕	女	18	15	48	33	86.7
54									
4	4	404550042	柯政學	男	10	10	NA	NA	13.3
2									
5	5	404550023	謝文躍	女	35	45	52	97	86.7
55									

```
FinalExam Attendance
```

1	50	2
2	68	9
3	79	9
4	0	7
5	67	9

```
> tail(x, 5)
```

	座號	學號	姓名	性別	quiz(1)	quiz(2)	quiz(3)	quiz(4)	TA
MidtermExam									
36	36	404550369	陳王霖	女	55	73	92	73	100.0
72									
37	37	404550420	何瑄穎	男	28	10	35	3	66.7
30									
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
```

FinalExam Attendance

```
36      81      9
37       0      7
38     42      9
39     62      9
40    100      3
```

```
> library(readxl)
```

```
> y <- read_excel("Calculus-score-B.xls", skip=2)
```

New names:

```
* `0.0700000000000000007` -> `0.070000000000000007...5`
* `0.0700000000000000007` -> `0.070000000000000007...6`
* `0.0800000000000000002` -> `0.080000000000000002...7`
* `0.0800000000000000002` -> `0.080000000000000002...8`
```

```
> names(y) <- c("座號", "學號", "姓名", "性別", "quiz(1)", "quiz(2)", "quiz(3)",
"quiz(4)", "TA", "MidtermExam", "FinalExam", "Attendance")
```

```
> head(y, 5)
```

A tibble: 5 x 12

```
座號 學號 姓名 性別 `quiz(1)` `quiz(2)` `quiz(3)` `quiz(4)` TA
```

MidtermExam

```
<dbl> <dbl> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>
<dbl>
1 1 4.05e8 史文羽... 男 60 81 100 97
100 90
2 2 4.05e8 鄭樺妤... 男 80 100 100 92
100 92
3 3 4.05e8 張敬安... 男 10 40 62 93
100 65
4 4 4.05e8 何筑亦... 女 15 25 40 13
93.3 36
5 5 4.05e8 張儀... 女 30 45 70 61
93.3 29
```

```
# ... with 2 more variables: FinalExam <dbl>, Attendance <dbl>
```

```
> tail(y, 5)
```

```
# A tibble: 5 x 12
```

```
  座號  學號 姓名 性別 `quiz(1)` `quiz(2)` `quiz(3)` `quiz(4)`    TA
MidtermExam
  <dbl> <dbl> <chr> <chr>    <dbl>    <dbl>    <dbl>    <dbl> <dbl>
<dbl>
1    51 4.05e8 鄭鈺尤… 女      80      85     100      85
100      89
2    52 4.05e8 楊宜路… 男      48      35      48      98
100      50
3    53 4.05e8 張渝好… 男       0      38      60      40
87      49
4    54 4.05e8 廖暄安… 男      50      70      20      85
100      54
5    55 5.00e8 楊毅亦… 女       5      35      45      55
87      58
```

```
# ... with 2 more variables: FinalExam <dbl>, Attendance <dbl>
```

```
> #ex.1(b)
```

```
> options("max.print" = 10000)
```

```
> x$"class" <- "A"
```

```
> y$"class" <- "B"
```

```
> score <- rbind(x, y)
```

```
> score[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 鄭樺好 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
```

```
> #ex.1(c)
> score[is.na(score)] <- 0
> score$"學期成績" <- (score$"quiz(1)"*0.07 + score$"quiz(2)"*0.07 +
score$"quiz(3)"*0.08 + score$"quiz(4)"*0.08 + score$"TA"*0.15 +
score$"MidtermExam"*0.25 + score$"FinalExam"*0.30) + score$"Attendance"
> score$"學期成績" <- ifelse(score$"學期成績" >= 100, 100, score$"學期成績" )
> score$"學期成績"

 [1] 33.450 70.010 67.995 10.895 73.375 67.015 38.920 78.345
48.525 20.455
[11] 96.435 67.295 16.150 22.030 73.990 79.400 8.560 24.245
61.905 61.340
[21] 49.915 68.570 10.995 67.055 68.000 69.210 65.635 84.040
66.100 78.540
[31] 75.330 69.860 72.240 82.260 54.765 88.460 30.205 55.245
77.920 99.150
[41] 94.030 97.060 81.350 40.535 55.375 62.355 61.310 50.450
21.600 36.700
[51] 32.150 76.810 48.200 52.550 69.700 43.360 60.910 94.070
77.990 23.950
[61] 39.100 80.600 72.850 22.050 47.200 20.800 61.550 58.300
40.800 55.000
[71] 26.280 70.050 49.450 62.900 54.960 74.900 71.360 67.800
85.140 21.300
[81] 72.200 78.410 82.300 51.510 74.660 45.200 65.300 87.220
100.000 95.720
[91] 100.000 72.590 44.460 70.000 59.350
```

```
> #ex.1(d)
```

```
> subset(score, score$"學期成績" >= 55 & score$"學期成績"<60)
```

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								
45	5 404685100	張儀	女	30	45	70	61	93.3
29								
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								
95	55 499555916	楊毅亦	女	5	35	45	55	87.0
58								

FinalExam	Attendance	class	學期成績
-----------	------------	-------	------

38	42	9	A	55.245
45	48	4	B	55.375
68	63	9	B	58.300
70	44	7	B	55.000
95	60	3	B	59.350

```
> #ex.1(e)
```

```
> mean(score$"學期成績"[1:40])
```

```
[1] 58.84575
```

```
> mean(score$"學期成績"[41:95])
```

```
[1] 61.123
```

```
> bscore <- which(score[,4] == "男")
```

```
> mean(bscore)
```

```
[1] 51.23729
```

```
> gscore <- which(score[,4] == "女")
```

```
> mean(gscore)
```

```
[1] 42.69444
```

```
> #ex.1(f)
```

```
> q <- subset(score, 60 > score$"學期成績" & score[,13] == "A")
```

```
> length(q)/40
```

```
[1] 0.35
```

```

> W <- subset(score, 60 > score$"學期成績" & score[,13] == "B")
> length(W)/55
[1] 0.2545455
>
>
> #ex.2(a)
> set.seed(123456)
> Letters.code <- sample(LETTERS[1:5], 20, replace=T)
> Numbers.code <- ifelse(Letters.code %in% c("A", "E"), 1, ifelse(Letters.code == "C", 2, 3))
> Numbers.code
[1] 3 3 3 1 1 3 3 2 2 1 2 3 3 1 1 3 1 2 3 2
> #ex.2(b)
> data <- data.frame(Letters.code, Numbers.code)
> data

```

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

>