```
> # 2020/11/20(五), 109學年第一學期 資料科學應用 R小考(1)-加分
> #
> # 學號:A106260105        姓名:黃念慈
> #
>
> # ex1(a)
> calculusA<- read.table("data/Calculus-score-A.csv",
+                        sep = ",", skip = 3)
> names(calculusA) <- c("No", "ID", "Name", "Gender",
+                       "Quiz1", "Quiz2", "Quiz3", "Quiz4",
+                       "TA", "Midterm", "Final", "ATT")
> head(calculusA, 5)
  No        ID   Name Gender Quiz1 Quiz2 Quiz3 Quiz4   TA Midterm
Final ATT
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
> tail(calculusA, 5)
   No        ID   Name Gender Quiz1 Quiz2 Quiz3 Quiz4    TA Midterm
Final ATT
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
> library(readxl)
```

```
> calculusB<- read_excel("data/Calculus-score-B.xls", skip = 2 )
New names:
* `0.070000000000000007` -> `0.070000000000000007...5`
* `0.070000000000000007` -> `0.070000000000000007...6`
* `0.080000000000000002` -> `0.080000000000000002...7`
* `0.080000000000000002` -> `0.080000000000000002...8`
> names(calculusB) <- c("No", "ID", "Name", "Gender", "Quiz1",
+                       "Quiz2", "Quiz3", "Quiz4", "TA", "Midterm",
"Final", "ATT")
> head(calculusB, 5)
# A tibble: 5 x 12
     No        ID Name   Gender Quiz1 Quiz2 Quiz3 Quiz4    TA Midterm
Final   ATT
  <dbl>     <dbl> <chr>  <chr>  <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
<dbl> <dbl>
1     1 404550465 史文羽 男        60    81   100    97 100         90
83     6
2     2 404685071 鄭樺妤 男        80   100   100    92 100         92
97     2
3     3 404685084 張敬安 男        10    40    62    93 100         65
84     9
4     4 404685099 何筑亦 女        15    25    40    13  93.3       36
5     9
5     5 404685100 張　儀 女        30    45    70    61  93.3       29
48     4
> tail(calculusB, 5)
# A tibble: 5 x 12
     No        ID Name   Gender Quiz1 Quiz2 Quiz3 Quiz4    TA Midterm
Final   ATT
  <dbl>     <dbl> <chr>  <chr>  <dbl> <dbl> <dbl> <dbl> <dbl>   <dbl>
<dbl> <dbl>
1    51 404685407 鄭鈺尤 女        80    85   100    85   100       89
95     9
2    52 404685905 楊宜路 男        48    35    48    98   100       50
62     9
3    53 404685013 張渝妤 男         0    38    60    40    87       49
25     1
4    54 404685119 廖暄安 男        50    70    20    85   100       54
```

```
69     4
5    55 499555916 楊毅亦 女          5    35    45    55    87      58
60     3
>
> # ex1(b)
> calculusA$class <- "A"
> calculusB$class <- "B"
> Score.all <- rbind(calculusA, calculusB)
> score.all <- as.data.frame(Score.all)
> score.all[38:43,]
   No        ID   Name Gender Quiz1 Quiz2 Quiz3 Quiz4    TA Midterm
Final ATT class
38 38 404550431 沈泓霏     女    15    25    53    67  93.3      29
42   9     A
39 39 404550442 許安霏     女    53    60    80    72 100.0      61
62   9     A
40 40 404550453 李政宜     男    80   100    85   100 100.0      95
100   3     A
41  1 404550465 史文羽     男    60    81   100    97 100.0      90
83   6     B
42  2 404685071 鄭樺妤     男    80   100   100    92 100.0      92
97   2     B
43  3 404685084 張敬安     男    10    40    62    93 100.0      65
84   9     B
>
> # ex1(c)
> score.all[is.na(score.all)] <- 0
> score.all$total <-
+   (score.all$Quiz1*0.07+score.all$Quiz2*0.07
+   +score.all$Quiz3*0.08 +score.all$Quiz4*0.08
+   +score.all$TA*0.15+score.all$Midterm*0.25
+   +score.all$Final*0.30+score.all$ATT)
> id.scroe100 <- id <- 1:length(score.all$total)
> id.scroe100[score.all$total > 100]
[1] 89 91
> score.all$total[89] <- 100
> score.all$total[91] <- 100
> score.all$total
```

```
 [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
>
> # ex1(d)
> subset(score.all, (score.all$total >= 55) & (score.all$total < 60))
   No        ID   Name Gender Quiz1 Quiz2 Quiz3 Quiz4   TA Midterm
Final ATT class
38 38 404550431 沈泓霏     女    15    25    53    67 93.3      29
42   9     A
45  5 404685100 張  儀     女    30    45    70    61 93.3      29
48   4     B
68 28 404720722 楊佳聿     女    30    35    20    50 60.0      45
63   9     B
70 30 404720527 馨飛羽     男    15     0    45    65 93.0      44
44   7     B
95 55 499555916 楊毅亦     女     5    35    45    55 87.0      58
60   3     B
    total
38 55.245
45 55.375
68 58.300
```

```
70 55.000
95 59.350
>
> # ex1(e)
> classA.total <- subset(score.all, class == "A", total)
> colMeans(classA.total)
   total
58.84575
> classB.total <- subset(score.all, class == "B", total)
> colMeans(classB.total)
 total
61.123
> male.total <- subset(score.all, Gender == "男", total)
> colMeans(male.total)
   total
60.90322
> female.total <- subset(score.all, Gender == "女", total)
> colMeans(female.total)
   total
58.95292
>
> # ex1(f)
> A.fail <- nrow(subset(score.all, (total < 60) & (class == "A")))
> A.all <- nrow(subset(score.all, class == "A"))
> A.fail/A.all
[1] 0.35
> maleB.fail <- nrow(subset(score.all, (total < 60) & (class == "B")
& (Gender == "男")))
> maleB.all <- nrow(subset(score.all, (class == "B") & (Gender == "男
")))
> maleB.fail/maleB.all
[1] 0.4102564
>
> # ex1(g)
> Score.female <- subset(score.all, Gender == "女")
> score.female<- Score.female[order(-Score.female$total),c(13, 2, 3,
14)]
> score.female$rank <- c(1:36)
```

```
> head(score.female, 5)
   class        ID   Name   total rank
89     B 404720541 詹傑仙 100.000    1
91     B 404685407 鄭鈺尤 100.000    2
11     A 404550189 丁易偉  96.435    3
36     A 404550369 陳王霖  88.460    4
79     B 404720436 曼李儷  85.140    5
>
> Score.male <- subset(score.all, Gender == "男")
> score.male<- Score.male[order(-Score.male$total),c(13, 2, 3, 14)]
> score.male$rank <- c(1:59)
> head(score.male, 5)
   class        ID   Name total rank
40     A 404550453 李政宜 99.15    1
42     B 404685071 鄭樺妤 97.06    2
90     B 404685109 許  何 95.72    3
58     B 404720161 劉莞韋 94.07    4
41     B 404550465 史文羽 94.03    5
>
> # ex2(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
>
> # ex2(b)
> code.df <- as.data.frame(Letters.code)
> code.df$Numbers.code <- Numbers.code
> code.df
   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