

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:A106260020 姓名:楊錄綺

```
> #ex1(a)
> score.a <- read.csv("data/Calculus-score-A.csv", na = "NA", fileEncoding = 'BIG5', skip=2)
> score.b <- read_excel("data/Calculus-score-B.xls", 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`
> head(score.a)
  座號    學號  姓名  性別  X7. X7..1 X8. X8..1 X15. X25. X30.
1    1 1401405008 希瑄彥  男   10     0    5    20  0.0   55   50
2    2 2401550880 張泓丞  男   25    40   70    87 80.0   46   68
3    3 3404550061 張安婕  女   18    15  48    33 86.7   54   79
4    4 4404550042 柯政學  男   10    10  NA    NA 13.3    2    0
5    5 5404550023 謝文躍  女   35    45  52    97 86.7   55   67
6    6 6404550000 張樺玫  男   30    35  90    67 86.7   54   48

Times
1     2
2     9
3     9
4     7
5     9
6     9
> head(score.b)
# A tibble: 6 x 12
  座號    學號  姓名  性別 `0.070000000000000000...` `0.070000000000000000...`
  <dbl> <dbl> <chr> <chr>          <dbl>          <dbl>
1     1 14.05e8 史文羽...  男              60              81
2     2 24.05e8 鄭樺好...  男              80             100
3     3 34.05e8 張敬安...  男              10              40
4     4 44.05e8 何筑亦...  女              15              25
5     5 54.05e8 張 儀...  女              30              45
6     6 64.02e8 雷西萱...  男              75              78
# ... with 6 more variables: `0.080000000000000002...7` <dbl>,
#   `0.080000000000000002...8` <dbl>,
#   `0.14999999999999999` <dbl>, `0.25` <dbl>,
```

```

# `0.2999999999999999` <dbl>, Times <dbl>
> tail(score.a)
  座號    學號  姓名  性別 X7. X7..1 X8. X8..1  X15. X25.
35   35 404550328 李梅祐  男   20    25   55    32   86.7   41
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

  X30. Times
35    48     7
36    81     9
37     0     7
38    42     9
39    62     9
40   100     3
> tail(score.b)
# A tibble: 6 x 12
  座號    學號  姓名  性別 `0.070000000000000000...` `0.070000000000000000...`
  <dbl> <dbl> <chr> <chr>                <dbl>                <dbl>
1    50 4.05e8 許 何...  男                      88                      73
2    51 4.05e8 鄭鈺尤...  女                      80                      85
3    52 4.05e8 楊宜路...  男                      48                      35
4    53 4.05e8 張渝好...  男                       0                      38
5    54 4.05e8 廖暄安...  男                      50                      70
6    55 5.00e8 楊毅亦...  女                       5                      35
# ... with 6 more variables: `0.0800000000000000002...7` <dbl>,
#   `0.0800000000000000002...8` <dbl>,
#   `0.149999999999999999` <dbl>, `0.25` <dbl>,
#   `0.299999999999999999` <dbl>, Times <dbl>
>
> #ex11(b)
> class.a <- rep(c("A"), times = 40)
> class.b <- rep(c("B"), times = 55)
> score.a$class <- data.frame(class.a)
> score.b$class <- data.frame(class.b)
>
> names(score.a) <- c("座號", "學號", "姓名", "性別", "quiz(1)", "quiz(2)", "quiz(3)", "quiz(4)", "TA",
"MidtermExam", "FinalExam", "Attendance", "class")
> names(score.b) <- c("座號", "學號", "姓名", "性別", "quiz(1)", "quiz(2)", "quiz(3)", "quiz(4)", "TA",

```

```
"MidtermExam", "FinalExam", "Attendance", "class")
```

```
> score <- rbind(score.a, score.b)
```

```
> score[38:43,]
```

	座號	學號	姓名	性別	quiz(1)	quiz(2)	quiz(3)	quiz(4)
38	38	404550431	沈泓霏	女	15	25	53	67
39	39	404550442	許安霏	女	53	60	80	72
40	40	404550453	李政宜	男	80	100	85	100
41	1	404550465	史文羽	男	60	81	100	97
42	2	404685071	鄭樺妤	男	80	100	100	92
43	3	404685084	張敬安	男	10	40	62	93

```
TA MidtermExam FinalExam Attendance class.a
```

38	93.3	29	42	9	A
39	100.0	61	62	9	A
40	100.0	95	100	3	A
41	100.0	90	83	6	B
42	100.0	92	97	2	B
43	100.0	65	84	9	B

```
> #ex1(c)
```

```
> score[is.na(score)] <- 0
```

```
> scoretotal <- c(as.matrix(score[, 5:11]))%% as.vector(c(0.07, 0.07, 0.08, 0.08, 0.15, 0.25, 0.3))+score[,12])
```

```
> scoretotal[scoretotal >= 100] <- c(100)
```

```
> scoretotal
```

[1]	33.450	70.010	67.995	10.895	73.375	67.015	38.920
[8]	78.345	48.525	20.455	96.435	67.295	16.150	22.030
[15]	73.990	79.400	8.560	24.245	61.905	61.340	49.915
[22]	68.570	10.995	67.055	68.000	69.210	65.635	84.040
[29]	66.100	78.540	75.330	69.860	72.240	82.260	54.765
[36]	88.460	30.205	55.245	77.920	99.150	94.030	97.060
[43]	81.350	40.535	55.375	62.355	61.310	50.450	21.600
[50]	36.700	32.150	76.810	48.200	52.550	69.700	43.360
[57]	60.910	94.070	77.990	23.950	39.100	80.600	72.850
[64]	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
[78]	67.800	85.140	21.300	72.200	78.410	82.300	51.510
[85]	74.660	45.200	65.300	87.220	100.000	95.720	100.000
[92]	72.590	44.460	70.000	59.350			

> #exl1(d)

> score[55 <= scoretotal & scoretotal > 60,]

	座號	學號	姓名	性別	quiz(1)	quiz(2)	quiz(3)	quiz(4)
2	2	401550880	張泓丞	男	25	40	70	87
3	3	404550061	張安婕	女	18	15	48	33
5	5	404550023	謝文躍	女	35	45	52	97
6	6	404550000	張樺玫	男	30	35	90	67
8	8	404550075	王亞詠	女	30	45	85	65
11	11	404550189	丁易偉	女	80	100	100	93
12	12	404550177	彥氏丞	男	40	20	55	38
15	15	404550161	玫鈞霖	男	25	30	85	53
16	16	404550149	家聿穎	女	47	60	70	82
19	19	404550272	黛婕西	女	20	13	60	43
20	20	404550250	偉茹麥	男	2	20	50	65
22	22	404550216	家芳榮	女	50	28	80	97
24	24	404550272	美慈蒙	女	13	48	73	65
25	25	404550250	鈺玲豪	男	55	25	65	60
26	26	404550215	宜 慈	男	40	58	65	30
27	27	404550287	柯陳渝	男	0	45	53	40
28	28	404550361	紀鍾暄	男	50	50	90	68
29	29	404550340	蔡鍾毅	男	35	15	82	28
30	30	404550328	梅林茹	男	25	65	62	41
31	31	404550386	王柯辰	女	45	53	77	57
32	32	404550360	曾紀歲	女	45	55	83	64
33	33	404550344	呂蔡嘉	男	30	40	90	78
34	34	404550322	廖繆傑	男	30	35	100	82
36	36	404550369	陳王霖	女	55	73	92	73
39	39	404550442	許安霏	女	53	60	80	72
40	40	404550453	李政宜	男	80	100	85	100
41	1	404550465	史文羽	男	60	81	100	97
42	2	404685071	鄭樺妤	男	80	100	100	92
43	3	404685084	張敬安	男	10	40	62	93
46	6	401550816	雷西萱	男	75	78	67	58
47	7	401555528	張麥笙	女	60	33	15	65
52	12	404720005	王恭莞	女	53	25	80	85
55	15	404720038	丁乃愛	男	55	70	85	80
57	17	404720150	曾銘清	男	65	63	15	50
58	18	404720161	劉莞韋	男	95	86	85	75
59	19	404720172	曾易佳	男	80	65	98	75
62	22	404720106	詹書飛	男	65	80	80	85

63	23	404720117	劉徐卉	男	65	90	70	65
67	27	404720624	張羽廖	男	45	0	30	65
72	32	404720723	茹呂怡	男	35	50	30	90
74	34	404720929	芳李米	男	35	15	40	65
76	36	404720135	萱何兆	男	35	25	85	55
77	37	404720231	笙沈品	男	40	58	45	70
78	38	404720338	倩許欣	男	50	60	60	30
79	39	404720436	曼李儷	女	60	40	73	90
81	41	404720638	恭倩鄭	男	20	15	50	70
82	42	404720736	琇曼張	女	58	40	80	55
83	43	404720834	張雅益	女	60	75	85	85
85	45	404720040	劉杰乃	男	40	43	85	80
87	47	404720245	黃品銘	男	70	15	60	50
88	48	404720343	廖林韋	男	96	55	85	70
89	49	404720541	詹傑仙	女	98	80	98	98
90	50	404685109	許 何	男	88	73	85	100
91	51	404685407	鄭鈺尤	女	80	85	100	85
92	52	404685905	楊宜路	男	48	35	48	98
94	54	404685119	廖暄安	男	50	70	20	85

TA MidtermExam FinalExam Attendance class.a

2	80.0	46	68	9	A
3	86.7	54	79	9	A
5	86.7	55	67	9	A
6	86.7	54	48	9	A
8	93.3	60	77	9	A
11	93.3	90	93	4	A
12	86.7	71	53	9	A
15	100.0	54	72	9	A
16	100.0	61	75	7	A
19	86.7	43	62	9	A
20	80.0	42	77	5	A
22	100.0	35	54	9	A
24	93.3	45	65	7	A
25	100.0	8	88	9	A
26	100.0	21	85	9	A
27	93.3	37	76	9	A
28	100.0	50	93	9	A
29	100.0	34	71	9	A
30	100.0	82	65	9	A
31	100.0	51	70	9	A

32	100.0	52	47	9	A
33	100.0	22	88	7	A
34	100.0	51	88	9	A
36	100.0	72	81	9	A
39	100.0	61	62	9	A
40	100.0	95	100	3	A
41	100.0	90	83	6	B
42	100.0	92	97	2	B
43	100.0	65	84	9	B
46	93.3	35	33	9	B
47	87.0	65	47	5	B
52	80.0	73	63	9	B
55	100.0	49	45	7	B
57	80.0	57	55	4	B
58	100.0	80	82	9	B
59	80.0	64	80	2	B
62	100.0	63	65	7	B
63	100.0	68	64	0	B
67	73.0	55	57	9	B
72	100.0	56	55	9	B
74	100.0	48	50	9	B
76	100.0	58	70	9	B
77	100.0	70	46	9	B
78	100.0	34	68	9	B
79	100.0	68	87	7	B
81	100.0	71	68	7	B
82	100.0	57	75	9	B
83	100.0	71	65	7	B
85	100.0	63	53	9	B
87	100.0	39	66	6	B
88	100.0	75	95	2	B
89	100.0	96	95	9	B
90	100.0	83	83	9	B
91	100.0	89	95	9	B
92	100.0	50	62	9	B
94	100.0	54	69	4	B

```
> #exl1(e)
```

```
> mean(scoretotal[which(score$class == "A")])
```

```
[1] 58.84575
```

```

> mean(scoretotal[which(score$class == "B")])
[1] 61.123
> boy <- scoretotal[which(score$性別 == "男")]
> mean(boy)
[1] 60.90322
> girl <- scoretotal[which(score$性別 == "女")]
> mean(girl)
[1] 58.95292

> #ex1(f)
> newscore <- data.frame(score, scoretotal)
> a.score <- newscore[1:40,]
> b.score <- newscore[41:95,]
> length(which(a.score$scoretotal < 60)) / length(a.score$座號)
[1] 0.35
> length(which(b.score$scoretotal < 60 & b.score$性別 == "男")) / length(b.score$座號)
[1] 0.2909091

> #ex1(g)
> boy.score <- newscore[score$性別 == "男",][order(boy, decreasing = T),]
> 名次 <- c(1, 2, 3, 4, 5)
> data.frame(名次, boy.score[1:5,])
  名次 座號      學號  姓名 性別 quiz.1. quiz.2. quiz.3.
40    1   40 404550453 李政宜  男     80     100     85
42    2    2 404685071 鄭樺好  男     80     100    100
90    3   50 404685109 許  何  男     88      73     85
58    4   18 404720161 劉莞韋  男     95      86     85
41    5    1 404550465 史文羽  男     60      81    100

  quiz.4.  TA MidtermExam FinalExam Attendance class.a
40     100 100           95       100           3       A
42      92 100           92       97           2       B
90     100 100           83       83           9       B
58      75 100           80       82           9       B
41      97 100           90       83           6       B

  scoretotal
40      99.15
42      97.06
90      95.72
58      94.07
41      94.03

```

>

```
> girl.score <- newscore[score$性別 == "女",][order(girl, decreasing = T),]
```

```
> data.frame(名次, girl.score[1:5,])
```

	名次	座號	學號	姓名	性別	quiz.1.	quiz.2.	quiz.3.
89	1	49	404720541	詹傑仙	女	98	80	98
91	2	51	404685407	鄭鈺尤	女	80	85	100
11	3	11	404550189	丁易偉	女	80	100	100
36	4	36	404550369	陳王霖	女	55	73	92
79	5	39	404720436	曼李儷	女	60	40	73

	quiz.4.	TA	MidtermExam	FinalExam	Attendance	class.a
89	98	100.0	96	95	9	B
91	85	100.0	89	95	9	B
11	93	93.3	90	93	4	A
36	73	100.0	72	81	9	A
79	90	100.0	68	87	7	B

	scoretotal
89	100.000
91	100.000
11	96.435
36	88.460
79	85.140

```
> #exl2(a)
```

```
> set.seed(123456)
```

```
> Letters.code <- sample(LETTERS[1:5], 20, replace=T)
```

```
> x <- c("A", "E")
```

```
> y <- c("B", "D")
```

```
> Numbers.code <- ifelse(Letters.code %in% x, 1, ifelse(Letters.code %in% y, 3, 2))
```

```
> #exl2(b)
```

```
> data.frame(Letters.code, Numbers.code)
```

	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