

> #2020/11/27(五), 109 學年第一學期 資料科學應用 HW(4)

> #學號: a107260050 姓名: 袁宇昕

> library(readxl)

> # 1(a)

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

> 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`

> head(x, 5)

	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

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

> tail(x, 5)

	座號	學號	姓名	性別	X7.	X7..1	X8.
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

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

> head(y, 5)

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

```
  座號  學號 姓名 性別 `0.070000000000000~`  
  <dbl> <dbl> <chr> <chr>          <dbl>  
1     1 1 4.05e8 史文羽~ 男           60  
2     2 4.05e8 鄭樺好~ 男           80  
3     3 4.05e8 張敬安~ 男           10  
4     4 4.05e8 何筑亦~ 女           15  
5     5 4.05e8 張 儀~ 女           30
```

```
# ... with 7 more variables:
```

```
# `0.0700000000000000007...6` <dbl>,  
# `0.0800000000000000002...7` <dbl>,  
# `0.0800000000000000002...8` <dbl>,  
# `0.14999999999999999` <dbl>, `0.25` <dbl>,  
# `0.29999999999999999` <dbl>, Times <dbl>
```

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

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

```
  座號  學號 姓名 性別 `0.070000000000000~`  
  <dbl> <dbl> <chr> <chr>          <dbl>  
1    51 4.05e8 鄭鈺尤~ 女           80  
2    52 4.05e8 楊宜路~ 男           48  
3    53 4.05e8 張渝好~ 男            0  
4    54 4.05e8 廖暄安~ 男           50  
5    55 5.00e8 楊毅亦~ 女            5
```

```
# ... with 7 more variables:
```

```
# `0.0700000000000000007...6` <dbl>,  
# `0.0800000000000000002...7` <dbl>,  
# `0.0800000000000000002...8` <dbl>,  
# `0.14999999999999999` <dbl>, `0.25` <dbl>,  
# `0.29999999999999999` <dbl>, Times <dbl>
```

```
> # 1(b)
```

```
> x1 <- as.data.frame(x)
```

```
> names(x1)[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(y)[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
```

```
> x2 <- transform(x1, class = "A")
```

```
> y1 <- transform(y, class = "B")
```

```
> names(x2) == names(y1)
```

```
[1] TRUE TRUE TRUE TRUE TRUE TRUE TRUE TRUE
```

```
[9] TRUE TRUE TRUE TRUE TRUE
```

```
> score <- rbind(x2, y1)
```

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

```
> score
```

	座號	學號	姓名	性別	quiz.1.	quiz.2.
1	1	401405008	希瑄彥	男	10	0
2	2	401550880	張泓丞	男	25	40
3	3	404550061	張安婕	女	18	15
4	4	404550042	柯政學	男	10	10
5	5	404550023	謝文躍	女	35	45
6	6	404550000	張樺玫	男	30	35
7	7	404550057	徐 敬	女	12	10
8	8	404550075	王亞詠	女	30	45
9	9	404550013	王承庭	男	30	15
10	10	404550001	高羽黛	女	25	10
11	11	404550189	丁易偉	女	80	100
12	12	404550177	彥氏丞	男	40	20
13	13	404550155	婕 學	女	0	10
14	14	404550143	躍庭祐	女	9	0
15	15	404550161	玫鈞霖	男	25	30
16	16	404550149	家聿穎	女	47	60
17	17	404550126	詠吾筑	男	3	5
18	18	404550294	庭馨儀	女	5	0
19	19	404550272	黛婕西	女	20	13
20	20	404550250	偉茹麥	男	2	20
21	21	404550238	氏存水	女	0	23
22	22	404550216	家芳榮	女	50	28
23	23	404550294	嫣玲杰	女	0	0
24	24	404550272	美慈蒙	女	13	48
25	25	404550250	鈺玲豪	男	55	25
26	26	404550215	宜 慈	男	40	58
27	27	404550287	柯陳渝	男	0	45
28	28	404550361	紀鍾暄	男	50	50
29	29	404550340	蔡鍾毅	男	35	15
30	30	404550328	梅林茹	男	25	65
31	31	404550386	王柯辰	女	45	53
32	32	404550360	曾紀歲	女	45	55

33	33	404550344	呂蔡嘉	男	30	40
34	34	404550322	廖繆傑	男	30	35
35	35	404550328	李梅祐	男	20	25
36	36	404550369	陳王霖	女	55	73
37	37	404550420	何瑄穎	男	28	10
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
44	4	404685099	何筑亦	女	15	25
45	5	404685100	張 儀	女	30	45
46	6	401550816	雷西萱	男	75	78
47	7	401555528	張麥笙	女	60	33
48	8	403555042	張水兆	男	0	30
49	9	404720052	柯榮品	男	0	0
50	10	404720063	謝卡欣	男	30	25
51	11	404720074	張莉儷	男	25	10
52	12	404720005	王恭莞	女	53	25
53	13	404720016	王琇易	男	15	5
54	14	404720027	高凱瓊	男	15	40
55	15	404720038	丁乃愛	男	55	70
56	16	404720049	張瑞書	女	20	28
57	17	404720150	曾銘清	男	65	63
58	18	404720161	劉莞韋	男	95	86
59	19	404720172	曾易佳	男	80	65
60	20	404720184	黃瓊品	男	15	0
61	21	404720195	廖愛靜	男	30	30
62	22	404720106	詹書飛	男	65	80
63	23	404720117	劉徐卉	男	65	90
64	24	404720211	許怡怡	男	30	10
65	25	404720427	鄭 芊	男	30	5
66	26	404720526	楊亞易	女	25	NA
67	27	404720624	張羽廖	男	45	0
68	28	404720722	楊佳聿	女	30	35
69	29	404720429	吾靜宜	男	55	20
70	30	404720527	馨飛羽	男	15	0

71	31	404720625	婕曾卉	男	10	0
72	32	404720723	茹呂怡	男	35	50
73	33	404720821	存廖芊	男	20	15
74	34	404720929	芳李米	男	35	15
75	35	404720037	勳 陳	男	20	30
76	36	404720135	萱何兆	男	35	25
77	37	404720231	笙沈品	男	40	58
78	38	404720338	倩許欣	男	50	60
79	39	404720436	曼李儷	女	60	40
80	40	404720534	雅清史	女	0	15
81	41	404720638	恭倩鄭	男	20	15
82	42	404720736	琇曼張	女	58	40
83	43	404720834	張雅益	女	60	75
84	44	404720932	曾林凱	女	55	18
85	45	404720040	劉杰乃	男	40	43
86	46	404720147	曾鈞瑞	女	45	40
87	47	404720245	黃品銘	男	70	15
88	48	404720343	廖林韋	男	96	55
89	49	404720541	詹傑仙	女	98	80
90	50	404685109	許 何	男	88	73
91	51	404685407	鄭鈺尤	女	80	85
92	52	404685905	楊宜路	男	48	35
93	53	404685013	張渝妤	男	0	38
94	54	404685119	廖暄安	男	50	70
95	55	499555916	楊毅亦	女	5	35

quiz.3. quiz.4. TA MidtermExam FinalExam

1	5	20	0.0	55	50
2	70	87	80.0	46	68
3	48	33	86.7	54	79
4	NA	NA	13.3	2	0
5	52	97	86.7	55	67
6	90	67	86.7	54	48
7	38	3	80.0	32	17
8	85	65	93.3	60	77
9	72	37	86.7	27	33
10	50	NA	46.7	24	0
11	100	93	93.3	90	93
12	55	38	86.7	71	53

13	40	NA	60.0	5	0
14	50	30	60.0	0	0
15	85	53	100.0	54	72
16	70	82	100.0	61	75
17	NA	NA	40.0	8	0
18	30	NA	93.3	18	0
19	60	43	86.7	43	62
20	50	65	80.0	42	77
21	30	7	93.3	45	37
22	80	97	100.0	35	54
23	NA	NA	53.3	0	0
24	73	65	93.3	45	65
25	65	60	100.0	8	88
26	65	30	100.0	21	85
27	53	40	93.3	37	76
28	90	68	100.0	50	93
29	82	28	100.0	34	71
30	62	41	100.0	82	65
31	77	57	100.0	51	70
32	83	64	100.0	52	47
33	90	78	100.0	22	88
34	100	82	100.0	51	88
35	55	32	86.7	41	48
36	92	73	100.0	72	81
37	35	3	66.7	30	0
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
44	40	13	93.3	36	5
45	70	61	93.3	29	48
46	67	58	93.3	35	33
47	15	65	87.0	65	47
48	50	80	13.0	NA	90
49	5	NA	73.0	5	NA
50	30	10	60.0	21	38

51	10	15	73.0	13	15
52	80	85	80.0	73	63
53	15	90	87.0	3	52
54	35	60	80.0	42	32
55	85	80	100.0	49	45
56	10	70	80.0	36	22
57	15	50	80.0	57	55
58	85	75	100.0	80	82
59	98	75	80.0	64	80
60	5	0	73.0	7	16
61	20	20	80.0	14	24
62	80	85	100.0	63	65
63	70	65	100.0	68	64
64	20	20	67.0	12	NA
65	10	65	100.0	35	20
66	NA	NA	67.0	0	0
67	30	65	73.0	55	57
68	20	50	60.0	45	63
69	5	15	73.0	44	30
70	45	65	93.0	44	44
71	0	66	67.0	5	0
72	30	90	100.0	56	55
73	10	55	80.0	28	46
74	40	65	100.0	48	50
75	22	60	80.0	50	38
76	85	55	100.0	58	70
77	45	70	100.0	70	46
78	60	30	100.0	34	68
79	73	90	100.0	68	87
80	NA	NA	60.0	9	NA
81	50	70	100.0	71	68
82	80	55	100.0	57	75
83	85	85	100.0	71	65
84	30	50	80.0	58	15
85	85	80	100.0	63	53
86	10	50	93.0	28	35
87	60	50	100.0	39	66
88	85	70	100.0	75	95

89	98	98	100.0	96	95
90	85	100	100.0	83	83
91	100	85	100.0	89	95
92	48	98	100.0	50	62
93	60	40	87.0	49	25
94	20	85	100.0	54	69
95	45	55	87.0	58	60

Attendance class

1	2	A
2	9	A
3	9	A
4	7	A
5	9	A
6	9	A
7	9	A
8	9	A
9	7	A
10	1	A
11	4	A
12	9	A
13	2	A
14	6	A
15	9	A
16	7	A
17	0	A
18	3	A
19	9	A
20	5	A
21	9	A
22	9	A
23	3	A
24	7	A
25	9	A
26	9	A
27	9	A
28	9	A
29	9	A
30	9	A

31	9	A
32	9	A
33	7	A
34	9	A
35	7	A
36	9	A
37	7	A
38	9	A
39	9	A
40	3	A
41	6	B
42	2	B
43	9	B
44	9	B
45	4	B
46	9	B
47	5	B
48	9	B
49	9	B
50	4	B
51	9	B
52	9	B
53	9	B
54	9	B
55	7	B
56	6	B
57	4	B
58	9	B
59	2	B
60	5	B
61	9	B
62	7	B
63	0	B
64	3	B
65	9	B
66	9	B
67	9	B
68	9	B

69	3	B
70	7	B
71	9	B
72	9	B
73	9	B
74	9	B
75	9	B
76	9	B
77	9	B
78	9	B
79	7	B
80	9	B
81	7	B
82	9	B
83	7	B
84	9	B
85	9	B
86	3	B
87	6	B
88	2	B
89	9	B
90	9	B
91	9	B
92	9	B
93	1	B
94	4	B
95	3	B

```
> # 1(c)
```

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

```
> Q <- score[5]*0.07 + score[6]*0.07 + score[7]*0.08 + score[8]*0.08 + score[9]*0.15
+ score[10]*0.25 + score[11]*0.30 + score[12]
```

```
> Q1 <- c(Q[1:95,])
```

```
> y <- ifelse(Q1 >= 100, 100, Q1)
```

```
> y1 <- as.data.frame(y)
```

```
> names(y1)[1] <- c("學期成績")
```

```
> y1
```

```
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)
```

```
> y2 <- ifelse(60 > y & y >= 50, Q1, (sep="0"))
```

```
> y3 <- as.data.frame(y2)
```

```
> y4 <- which(y3 > 0)
```

```
> score[y4, ]
```

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

48	50	80	13.0	0	90
54	35	60	80.0	42	32
68	20	50	60.0	45	63
70	45	65	93.0	44	44
75	22	60	80.0	50	38
84	30	50	80.0	58	15
95	45	55	87.0	58	60

Attendance class

35	7	A
38	9	A
45	4	B
48	9	B
54	9	B
68	9	B
70	7	B
75	9	B
84	9	B
95	3	B

```
> # 1(e)
```

```
> Ascore <- which(score[,13] == "A")
```

```
> Bscore <- which(score[,13] == "B")
```

```
> sum(y1[Ascore,]) / length(Ascore)
```

```
[1] 58.84575
```

```
> sum(y1[Bscore,]) / length(Bscore)
```

```
[1] 61.123
```

```
> A1 <- which(score[,4] == "女")
```

```
> B1 <- which(score[,4] == "男")
```

```
> sum(y1[A1,]) / length(A1)
```

```
[1] 58.95292
```

```
> sum(y1[B1,]) / length(B1)
```

```
[1] 60.90322
```

```
> # 1(f)
```

```
> A2 <- ifelse(60 > y & score[,13] == "A", x, (sep="0"))
```

```
> A3 <- as.data.frame(A2)
```

```
> A4 <- which(A3 > 0)
```

```
> length(A4) / length(Ascore)
```

```
[1] 13.225
```

```
> B2 <- ifelse(60 > y & score[,13] == "B" & score[,4] == "男", x, (sep="0"))
```

```

> B3 <- as.data.frame(B2)
> B4 <- which(B3 > 0)
> length(B4) / length(Bscore)
[1] 11.16364
> # 1(g)
> score1 <- transform(score,score = y1)
> names(score1)[14] <- c("score")
> SG <- score1[A1,]
> SB <- score1[B1,]
> SG1 <- order(SG$score, decreasing = TRUE)
> SB1 <- order(SB$score, decreasing = TRUE)
> SG2 <- SG[SG1,]
> SB2 <- SB[SB1,]
> head(SG2, 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    score
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(SB2, 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 score

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)

> y <- c(sample(LETTERS[1:5], 20, replace=T))

> x <-c()

> for(i in 1:20){

+ if(y[i] == "A")

+ x[i] <- 1

+ else if(y[i] == "E")

+ x[i] <- 1

+ else if(y[i] == "C")

+ x[i] <- 2

+ else

+ x[i] <- 3

+ }

> cat(x)

2 3 3 1 1 1 3 2 1 1 1 1 2 2 1 1 3 1 2 1

> # 2(b)

> LN <- data.frame(Letters.code = y, Numbers.code = x)

> LN

Letters.code Numbers.code

1 C 2

2 D 3

3 B 3

4 A 1

5 A 1

6	E	1
7	B	3
8	C	2
9	A	1
10	A	1
11	A	1
12	E	1
13	C	2
14	C	2
15	A	1
16	A	1
17	D	3
18	E	1
19	C	2
20	E	1
>		
>		