2020/11/06(五), 109 學年第一學期 資料科學應用 R 作業(2)

學號:A107260007　　姓名:蕭亞青

```
> #2020/11/06 回家作業
>
> #ex1.13
> lm.obj <- lm(airquality$Wind ~ airquality$Temp)
> lm.anova <- anova(lm.obj)
> lm.summary <- summary(lm.obj)
> x <- c(class(lm.anova), str(lm.anova))
Classes ‘anova’ and 'data.frame':	2 obs. of  5 variables:
 $ Df     : int  1 151
 $ Sum Sq : num  396 1491
 $ Mean Sq: num  395.71 9.87
 $ F value: num  40.1 NA
 $ Pr(>F) : num  2.64e-09 NA
```

```
 - attr(*, "heading")= chr [1:2] "Analysis of Variance Table₩n" "Response:
airquality$Wind"
> attributes(lm.summary)
$names
 [1] "call"          "terms"         "residuals"
 [4] "coefficients"  "aliased"       "sigma"
 [7] "df"            "r.squared"     "adj.r.squared"
[10] "fstatistic"    "cov.unscaled"

$class
[1] "summary.lm"

> lm.summary$r.squared
[1] 0.2097529
>
> #ex1.20
> my.data <- read.delim("data/statlog_vehicle_846x18.txt")
> data.frame(my.data)
    no class compactness circularity distance
```

1	1	0	96	55	103
2	2	0	101	56	100
3	3	0	93	35	66
4	4	0	101	48	107
5	5	0	87	38	85
6	6	0	95	48	104
7	7	0	98	55	101
8	8	0	107	53	103
9	9	0	103	50	98
10	10	0	77	38	63
11	11	0	89	41	75
12	12	0	98	55	101
13	13	0	96	55	98
14	14	0	97	59	108
15	15	0	92	39	91
16	16	0	73	37	53
17	17	0	101	53	103
18	18	0	79	40	80
19	19	0	80	37	57

20	20	0	94	38	84
21	21	0	97	50	108
22	22	0	95	46	105
23	23	0	99	46	105
24	24	0	85	39	77
25	25	0	77	38	75
26	26	0	88	35	50
27	27	0	100	45	100
28	28	0	102	54	100
29	29	0	106	49	107
30	30	0	95	45	80
31	31	0	103	54	107
32	32	0	93	35	72
33	33	0	85	36	78
34	34	0	91	45	75
35	35	0	82	38	53
36	36	0	107	52	101
37	37	0	98	54	104
38	38	0	103	54	91

39	39	0	108	51	103
40	40	0	84	39	90
41	41	0	78	36	60
42	42	0	98	45	76
43	43	0	101	51	105
44	44	0	90	36	78
45	45	0	97	48	94
46	46	0	111	54	103
47	47	0	103	55	100
48	48	0	92	46	79
49	49	0	101	56	100
50	50	0	94	39	89

	radiusratio	pr.axis	max.length	scatterratio
1	201	65	9	204
2	215	69	10	208
3	154	59	6	142
4	222	68	10	208
5	177	61	8	164
6	214	67	9	205

7	228	70	9	210
8	221	66	11	209
9	212	63	9	193
10	135	59	5	130
11	143	56	7	146
12	219	69	11	225
13	161	54	10	215
14	227	70	11	224
15	191	62	8	176
16	111	54	6	126
17	203	63	9	195
18	133	55	7	147
19	116	55	6	125
20	158	55	9	169
21	211	65	10	214
22	219	68	9	201
23	209	64	11	197
24	151	59	8	150
25	144	59	6	147

26	121	58	5	114
27	209	65	8	201
28	163	53	10	213
29	194	57	11	214
30	186	62	7	164
31	218	64	12	222
32	172	62	7	149
33	149	55	7	147
34	154	57	6	150
35	125	59	5	133
36	218	64	11	202
37	186	59	10	213
38	179	57	11	220
39	197	60	11	211
40	180	60	7	177
41	116	56	6	123
42	166	60	7	157
43	212	68	10	209
44	179	64	8	157

45	198	63	9	181
46	171	50	11	221
47	194	62	11	212
48	176	64	8	162
49	168	55	11	214
50	194	62	9	172

	elongatedness	pr.axis.1	max.length.1	scaledvmi
1	32	23	166	227
2	32	24	169	227
3	46	18	128	162
4	32	24	154	232
5	40	20	129	186
6	32	23	151	227
7	31	24	168	236
8	32	24	163	222
9	34	22	161	214
10	52	18	130	145
11	46	19	137	170
12	30	25	178	231

13	31	24	175	226
14	30	25	186	225
15	37	21	137	196
16	55	18	128	135
17	34	22	162	210
18	47	19	135	172
19	54	18	125	142
20	39	20	130	196
21	31	24	156	232
22	33	23	148	223
23	34	23	152	212
24	45	19	134	176
25	46	19	132	167
26	59	17	122	132
27	32	23	147	231
28	31	24	173	219
29	31	24	161	224
30	40	20	145	188
31	30	25	174	221

32	44	19	124	169
33	45	19	128	168
34	44	19	146	170
35	51	18	128	152
36	33	23	164	219
37	32	24	172	223
38	31	25	170	220
39	31	24	160	222
40	37	21	131	209
41	55	17	124	141
42	42	20	148	184
43	32	24	162	222
44	42	19	126	182
45	36	21	155	200
46	30	25	172	227
47	31	24	175	217
48	41	20	149	183
49	31	24	175	219
50	38	21	135	191

	scaledvma	scaledradius	skewness	skewness.1
1	624	246	74	6
2	651	223	74	6
3	304	120	64	5
4	641	204	70	5
5	402	130	63	1
6	628	202	74	5
7	661	245	72	1
8	653	212	66	0
9	567	185	64	5
10	247	139	79	13
11	317	156	76	18
12	748	216	74	6
13	683	221	76	3
14	732	218	70	10
15	466	151	67	3
16	227	147	82	1
17	571	210	68	5
18	311	144	76	8

19	229	132	81	8
20	430	155	69	9
21	683	218	72	7
22	602	201	69	5
23	575	159	65	0
24	331	133	73	0
25	315	136	80	16
26	192	138	74	21
27	611	189	72	5
28	669	201	76	12
29	670	172	67	0
30	406	178	65	11
31	728	199	67	0
32	334	125	62	5
33	321	134	64	10
34	335	180	66	16
35	259	146	87	0
36	610	192	65	17
37	665	217	73	1

38	707	198	72	1
39	661	187	67	7
40	469	145	71	4
41	221	121	78	3
42	371	186	69	13
43	653	224	73	5
44	367	142	66	1
45	494	189	64	20
46	727	201	69	15
47	666	219	73	10
48	396	178	67	2
49	681	224	74	2
50	444	121	63	4

	kurtosis	kurtosis.1	hollows	
1	2	186	194	
2	5	186	193	
3	13	197	202	
4	38	190	202	
5	25	198	205	

6	9	186	193
7	6	188	197
8	1	191	201
9	5	198	204
10	21	183	187
11	5	184	188
12	14	187	195
13	6	185	193
14	25	186	198
15	23	192	200
16	15	176	184
17	5	191	198
18	30	181	193
19	5	178	184
20	15	190	195
21	29	188	197
22	38	191	202
23	33	194	205
24	16	184	193

25	20	181	187
26	4	182	187
27	5	189	195
28	27	187	195
29	39	192	206
30	18	199	204
31	18	189	200
32	30	203	210
33	24	197	203
34	2	193	198
35	0	177	183
36	2	197	206
37	26	186	195
38	32	186	198
39	3	190	200
40	38	190	198
41	16	178	185
42	10	190	196
43	23	186	195

```
44            20            192        198

45            11            199        203

46             6            190        198

47            14            187        194

48            10            191        198

49             3            185        192

50            23            201        209

 [ reached 'max' / getOption("max.print") -- omitted 796 rows ]

> head(my.data, 5)

   no class compactness circularity distance

1  1     0            96            55         103

2  2     0            101           56         100

3  3     0             93            35          66

4  4     0            101           48         107

5  5     0             87            38          85

   radiusratio pr.axis max.length scatterratio

1           201       65           9           204

2           215       69          10           208

3           154       59           6           142
```

4	222	68	10	208
5	177	61	8	164

	elongatedness	pr.axis.1	max.length.1	scaledvmi
1	32	23	166	227
2	32	24	169	227
3	46	18	128	162
4	32	24	154	232
5	40	20	129	186

	scaledvma	scaledradius	skewness	skewness.1
1	624	246	74	6
2	651	223	74	6
3	304	120	64	5
4	641	204	70	5
5	402	130	63	1

	kurtosis	kurtosis.1	hollows
1	2	186	194
2	5	186	193
3	13	197	202
4	38	190	202

5 25 198 205

> tail(my.data, 5)

no class compactness circularity distance

842 842 3 87 45 66

843 843 3 95 43 76

844 844 3 90 44 72

845 845 3 89 46 84

846 846 3 85 36 66

radiusratio pr.axis max.length scatterratio

842 139 58 8 140

843 142 57 10 151

844 157 64 8 137

845 163 66 11 159

846 123 55 5 120

elongatedness pr.axis.1 max.length.1 scaledvmi

842 47 18 148 168

843 44 19 149 173

844 48 18 144 159

845 43 20 159 173

```
846                  56           17                128          140

      scaledvma scaledradius skewness skewness.1

842           294                175          73                3

843           339                159          71                2

844           283                171          65                9

845           368                176          72                1

846           212                131          73                1

      kurtosis kurtosis.1 hollows

842           12            188         196

843           23            187         200

844            4            196         203

845           20            186         197

846           18            186         190

>

> #ex1.28

> my.data1 <- read.table("data/stock-data.txt", header=TRUE, sep="₩t", fill=T)

> head(my.data1, 5)

   民國 100 年 5 家半導體公司股票月成交資訊.元.股.       X

1                                          半導體公司  年度
```

2 台積電 100

3 台積電 100

4 台積電 100

5 台積電 100

X.1 X.2 X.3 X.4 X.5

1 **月份 最高價 最低價 加權平均價 成交筆數**

2 1 78.3 69.6 74.3 263,999

3 2 77 69.9 72.54 235,159

4 3 72.2 65.7 69.74 276,434

5 4 73.9 68 71.37 211,611

X.6 X.7 X.8

1 **成交金額 成交股數 週轉率百分比**

2 100,578,274,926 1,353,616,348 5.22

3 74,985,055,548 1,033,654,452 3.98

4 88,459,924,495 1,268,289,393 4.89

5 70,177,023,098 983,177,475 3.79

> tail(my.data1, 5)

民國 100 **年** 5 **家半導體公司股票月成交資訊.元.股.** X

57 **旺宏** 100

58 旺宏 100

59 旺宏 100

60 旺宏 100

61 旺宏 100

```
    X.1    X.2    X.3    X.4        X.5                  X.6
57    8   14.5 10.25 11.84 152,177 8,137,500,167
58    9 12.65   10.4 11.55 108,879 5,542,998,380
59   10      12 10.25 11.31   68,571 3,041,525,834
60   11 13.65 10.85 12.54 167,018 9,538,526,797
61   12 12.85 11.15 12.17 115,192 5,070,210,532
                  X.7    X.8
57 687,167,610 20.31
58 479,779,350 14.18
59 268,710,697   7.94
60 760,264,306 22.47
61 416,455,073 12.31
>
> #ex1.33
```

```
> Dates <- c("0924", "1112", "1231", "1105", "0604", "0219", "0416", "0611", "0813",
"1029")
> Time <- c("01:00", "04:00", "16:00", "23:00", "08:00", "09:00", "07:00", "17:00",
"03:00", "14:00")
> Items <- factor(c("shirt", "shirt", "pants", "jacket", "jacket", "shirt", "jacket",
"jacket", "shoes", "shirt"))
> Volume <- as.numeric(c(7951,159,1958, 6848, 3762, 3678, 8696, 9045, 6208,
1425))
> DatesTime <- paste(Dates, Time)
> strptime(DatesTime, "%m%d %H%M")
 [1] NA NA NA NA NA NA NA NA NA NA
> as.POSIXct(DatesTime, format="%m%d %H%M", tz="UTC")
 [1] NA NA NA NA NA NA NA NA NA NA
> my.Sale <- data.frame(Items, Volume)
> my.Sale
      Items Volume
1     shirt     7951
2     shirt      159
3     pants     1958
```

4 jacket 6848

5 jacket 3762

6 shirt 3678

7 jacket 8696

8 jacket 9045

9 shoes 6208

10 shirt 1425

>