

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

學號:A107260016

姓名:凌倫敏

(請依照規定)貼上執行程式碼及執行結果。

詳見: R 程式作業繳交方式

<http://www.hmwu.idv.tw/web/teaching/doc/R-how-homework.pdf>

```
> ## ex1.13(a)
> lm.obj <- lm(airquality$Wind ~ airquality$Temp)
> lm.anova <- anova(lm.obj)
> lm.summary <- summary(lm.obj)
> class(lm.anova)
[1] "anova"          "data.frame"
> 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"
> ## ex1.13(b)
> attributes(lm.summary)
$names
 [1] "call"          "terms"          "residuals"      "coefficients"
 [5] "aliased"        "sigma"          "df"              "r.squared"
 [9] "adj.r.squared" "fstatistic"     "cov.unscaled"

$class
[1] "summary.lm"

> attr(lm.summary, "names")
 [1] "call"          "terms"          "residuals"      "coefficients"
 [5] "aliased"        "sigma"          "df"              "r.squared"
```

```

[9] "adj.r.squared" "fstatistic"    "cov.unscaled"
> names(lm.summary)
[1] "call"          "terms"          "residuals"      "coefficients"
[5] "aliased"        "sigma"          "df"             "r.squared"
[9] "adj.r.squared" "fstatistic"    "cov.unscaled"
> R <- lm.summary["r.squared"]
> class(R)
[1] "list"
> R2 <- as.numeric(R) # 轉成數字
> class(R2)
[1] "numeric"
> R2^2
[1] 0.04399628
> ## ex1.20

```

```

> my.data <- read.table("statlog_vehicle_846x18.txt", row.names=1)
> my.data

```

	V2	V3	V4	V5	V6	V7
no class	compactness	circularity	distance	radiusratio	pr.axis	
1	0	96	55	103	201	65
2	0	101	56	100	215	69
3	0	93	35	66	154	59
4	0	101	48	107	222	68
5	0	87	38	85	177	61
6	0	95	48	104	214	67
7	0	98	55	101	228	70
8	0	107	53	103	221	66
9	0	103	50	98	212	63
10	0	77	38	63	135	59
11	0	89	41	75	143	56
12	0	98	55	101	219	69
13	0	96	55	98	161	54
14	0	97	59	108	227	70
15	0	92	39	91	191	62
16	0	73	37	53	111	54
17	0	101	53	103	203	63
18	0	79	40	80	133	55
19	0	80	37	57	116	55
20	0	94	38	84	158	55

21	0	97	50	108	211	65
22	0	95	46	105	219	68
23	0	99	46	105	209	64
24	0	85	39	77	151	59
25	0	77	38	75	144	59
26	0	88	35	50	121	58
27	0	100	45	100	209	65
28	0	102	54	100	163	53
29	0	106	49	107	194	57
30	0	95	45	80	186	62
31	0	103	54	107	218	64
32	0	93	35	72	172	62
33	0	85	36	78	149	55
34	0	91	45	75	154	57
35	0	82	38	53	125	59
36	0	107	52	101	218	64
37	0	98	54	104	186	59
38	0	103	54	91	179	57
39	0	108	51	103	197	60
40	0	84	39	90	180	60
41	0	78	36	60	116	56
42	0	98	45	76	166	60
43	0	101	51	105	212	68
44	0	90	36	78	179	64
45	0	97	48	94	198	63
46	0	111	54	103	171	50
47	0	103	55	100	194	62
48	0	92	46	79	176	64
49	0	101	56	100	168	55
50	0	94	39	89	194	62
51	0	89	36	72	141	56

V8

V9

V10

V11

V12

V13

no max.length scatterratio elongatedness pr.axis max.length scaledvmi

1	9	204	32	23	166	227
2	10	208	32	24	169	
227						
3	6	142	46	18	128	162

4	10	208	32	24	154	
232						
5	8	164	40	20	129	186
6	9	205	32	23	151	227
7	9	210	31	24	168	236
8	11	209	32	24	163	
222						
9	9	193	34	22	161	214
10	5	130	52	18	130	
145						
11	7	146	46	19	137	
170						
12	11	225	30	25	178	
231						
13	10	215	31	24	175	
226						
14	11	224	30	25	186	
225						
15	8	176	37	21	137	
196						
16	6	126	55	18	128	
135						
17	9	195	34	22	162	
210						
18	7	147	47	19	135	
172						
19	6	125	54	18	125	
142						
20	9	169	39	20	130	
196						
21	10	214	31	24	156	
232						
22	9	201	33	23	148	
223						
23	11	197	34	23	152	
212						
24	8	150	45	19	134	
176						

25	6	147	46	19	132
167					
26	5	114	59	17	122
132					
27	8	201	32	23	147
231					
28	10	213	31	24	173
219					
29	11	214	31	24	161
224					
30	7	164	40	20	145
188					
31	12	222	30	25	174
221					
32	7	149	44	19	124
169					
33	7	147	45	19	128
168					
34	6	150	44	19	146
170					
35	5	133	51	18	128
152					
36	11	202	33	23	164
219					
37	10	213	32	24	172
223					
38	11	220	31	25	170
220					
39	11	211	31	24	160
222					
40	7	177	37	21	131
209					
41	6	123	55	17	124
141					
42	7	157	42	20	148
184					
43	10	209	32	24	162
222					

44	8	157		42	19	126
182						
45	9	181		36	21	155
200						
46	11	221		30	25	172
227						
47	11	212		31	24	175
217						
48	8	162		41	20	149
183						
49	11	214		31	24	175
219						
50	9	172		38	21	135
191						
51	7	138		48	18	126
163						
	V14	V15	V16	V17	V18	V19
V20						
no scaledvma scaledradius skewness skewness kurtosis kurtosis hollows						
1	624	246	74	6	2	186
194						
2	651	223	74	6	5	186
193						
3	304	120	64	5	13	197
202						
4	641	204	70	5	38	190
202						
5	402	130	63	1	25	198
205						
6	628	202	74	5	9	186
193						
7	661	245	72	1	6	188
197						
8	653	212	66	0	1	191
201						
9	567	185	64	5	5	198
204						
10	247	139	79	13	21	183

187						
11	317	156	76	18	5	184
188						
12	748	216	74	6	14	187
195						
13	683	221	76	3	6	185
193						
14	732	218	70	10	25	186
198						
15	466	151	67	3	23	192
200						
16	227	147	82	1	15	176
184						
17	571	210	68	5	5	191
198						
18	311	144	76	8	30	181
193						
19	229	132	81	8	5	178
184						
20	430	155	69	9	15	190
195						
21	683	218	72	7	29	188
197						
22	602	201	69	5	38	191
202						
23	575	159	65	0	33	194
205						
24	331	133	73	0	16	184
193						
25	315	136	80	16	20	181
187						
26	192	138	74	21	4	182
187						
27	611	189	72	5	5	189
195						
28	669	201	76	12	27	187
195						
29	670	172	67	0	39	192

206						
30	406	178	65	11	18	199
204						
31	728	199	67	0	18	189
200						
32	334	125	62	5	30	203
210						
33	321	134	64	10	24	197
203						
34	335	180	66	16	2	193
198						
35	259	146	87	0	0	177
183						
36	610	192	65	17	2	197
206						
37	665	217	73	1	26	186
195						
38	707	198	72	1	32	186
198						
39	661	187	67	7	3	190
200						
40	469	145	71	4	38	190
198						
41	221	121	78	3	16	178
185						
42	371	186	69	13	10	190
196						
43	653	224	73	5	23	186
195						
44	367	142	66	1	20	192
198						
45	494	189	64	20	11	199
203						
46	727	201	69	15	6	190
198						
47	666	219	73	10	14	187
194						
48	396	178	67	2	10	191


```

198
49      681      224      74      2      3      185
192
50      444      121      63      4      23      201
209
51      286      130      72      1      1      187
192
[ reached 'max' / getOption("max.print") -- omitted 795 rows ]
> str(my.data)
'data.frame':  847 obs. of  19 variables:
 $ V2 : chr  "class" "0" "0" "0" ...
 $ V3 : chr  "compactness" "96" "101" "93" ...
 $ V4 : chr  "circularity" "55" "56" "35" ...
 $ V5 : chr  "distance" "103" "100" "66" ...
 $ V6 : chr  "radiusratio" "201" "215" "154" ...
 $ V7 : chr  "pr.axis" "65" "69" "59" ...
 $ V8 : chr  "max.length" "9" "10" "6" ...
 $ V9 : chr  "scatterratio" "204" "208" "142" ...
 $ V10: chr  "elongatedness" "32" "32" "46" ...
 $ V11: chr  "pr.axis" "23" "24" "18" ...
 $ V12: chr  "max.length" "166" "169" "128" ...
 $ V13: chr  "scaledvmi" "227" "227" "162" ...
 $ V14: chr  "scaledvma" "624" "651" "304" ...
 $ V15: chr  "scaledradius" "246" "223" "120" ...
 $ V16: chr  "skewness" "74" "74" "64" ...
 $ V17: chr  "skewness" "6" "6" "5" ...
 $ V18: chr  "kurtosis" "2" "5" "13" ...
 $ V19: chr  "kurtosis" "186" "186" "197" ...
 $ V20: chr  "hollows" "194" "193" "202" ...
> dim(my.data)
[1] 847  19
> my.data[c(1:6, 843:847), ]
      V2      V3      V4      V5      V6      V7
no  class compactness circularity distance radiusratio pr.axis
1     0         96         55        103        201        65
2     0        101         56        100        215        69
3     0         93         35         66        154        59
4     0        101         48        107        222        68

```

5	0	87	38	85	177	61
842	3	87	45	66	139	58
843	3	95	43	76	142	57
844	3	90	44	72	157	64
845	3	89	46	84	163	66
846	3	85	36	66	123	55

V8 V9 V10 V11 V12

V13

no max.length scatterratio elongatedness pr.axis max.length scaledvmi

1	9	204	32	23	166
227					
2	10	208	32	24	169
227					
3	6	142	46	18	128
162					
4	10	208	32	24	154
232					
5	8	164	40	20	129
186					
842	8	140	47	18	148
168					
843	10	151	44	19	149
173					
844	8	137	48	18	144
159					
845	11	159	43	20	159
173					
846	5	120	56	17	128
140					

V14 V15 V16 V17 V18 V19

V20

no scaledvma scaledradius skewness skewness kurtosis kurtosis hollows

1	624	246	74	6	2	186
194						
2	651	223	74	6	5	186
193						
3	304	120	64	5	13	197
202						

```

4      641      204      70      5      38      190
202
5      402      130      63      1      25      198
205
842    294      175      73      3      12      188
196
843    339      159      71      2      23      187
200
844    283      171      65      9      4       196
203
845    368      176      72      1      20      186
197
846    212      131      73      1      18      186
190

```

```

> n <- nrow(my.data)
> p <- ncol(my.data)
> myData <- matrix(rnorm(n*p), ncol = p, nrow=n)
> print(object.size(myData), units = "Mb")
0.1 Mb
> ## ex1.28
> my.data2 <- read.table("stock-data.txt", header = TRUE, skip = 1, sep="\t")
> dim(my.data2)
[1] 60 10
> my.data2[c(1:5, 56:60), ]

```

	半導體公司	年度	月份	最高價	最低價	加權平均價	成交筆數
1	台積電	100	1	78.30	69.60	74.30	263,999
2	台積電	100	2	77.00	69.90	72.54	235,159
3	台積電	100	3	72.20	65.70	69.74	276,434
4	台積電	100	4	73.90	68.00	71.37	211,611
5	台積電	100	5	76.90	73.00	74.96	213,185
56	旺宏	100	8	14.50	10.25	11.84	152,177
57	旺宏	100	9	12.65	10.40	11.55	108,879
58	旺宏	100	10	12.00	10.25	11.31	68,571
59	旺宏	100	11	13.65	10.85	12.54	167,018
60	旺宏	100	12	12.85	11.15	12.17	115,192

	成交金額	成交股數	週轉率百分比
1	100,578,274,926	1,353,616,348	5.22
2	74,985,055,548	1,033,654,452	3.98

3	88,459,924,495	1,268,289,393	4.89
4	70,177,023,098	983,177,475	3.79
5	74,005,599,560	987,256,484	3.80
56	8,137,500,167	687,167,610	20.31
57	5,542,998,380	479,779,350	14.18
58	3,041,525,834	268,710,697	7.94
59	9,538,526,797	760,264,306	22.47
60	5,070,210,532	416,455,073	12.31

```
> attributes(my.data2)
```

```
$names
```

```
[1] "半導體公司" "年度" "月份" "最高價"
[5] "最低價" "加權平均價" "成交筆數" "成交金額"
[9] "成交股數" "週轉率百分比"
```

```
$class
```

```
[1] "data.frame"
```

```
$row.names
```

```
[1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
[23] 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44
[45] 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60
```

```
> class("成交筆數")
```

```
[1] "character"
```

```
> class("成交金額")
```

```
[1] "character"
```

```
> class("成交股數")
```

```
[1] "character"
```

```
> ## ex1.33(a)
```

```
> 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")
```

```
> Items1 <-c ( "shirt", "shirt", "pants", "jacket", "jacket", "shirt", "jacket", "jacket",
"shoes", "shirt")
```

```
> Volume1 <-c ("7951", "159", "1958", "6848", "3762", "3678", "8696", "9045",
"6208", "1425")
```

```
> class(Dates)
```

```

[1] "character"
> class(Time)
[1] "character"
> class(Items1)
[1] "character"
> class(Volume1)
[1] "character"
> DateTime1 <- paste("2018", Dates, Time)
> DateTime <- strptime(DateTime1, format="%Y %m%d %H:%M", tz = "UTC")
> Items <- as.factor(Items1)
> Volume <- as.numeric(Volume1)
> class(DateTime)
[1] "POSIXlt" "POSIXt"
> class(Items)
[1] "factor"
> class(Volume)
[1] "numeric"
> mysale <- data.frame (DateTime, Items, Volume)
> mysale

```

	DateTime	Items	Volume
1	2018-09-24 01:00:00	shirt	7951
2	2018-11-12 04:00:00	shirt	159
3	2018-12-31 16:00:00	pants	1958
4	2018-11-05 23:00:00	jacket	6848
5	2018-06-04 08:00:00	jacket	3762
6	2018-02-19 09:00:00	shirt	3678
7	2018-04-16 07:00:00	jacket	8696
8	2018-06-11 17:00:00	jacket	9045
9	2018-08-13 03:00:00	shoes	6208
10	2018-10-29 14:00:00	shirt	1425