```
#2020/11/06(五), 109學年第一學期 資料科學應用 R作業
(2) > #學號: A107260024      姓名: 游閔超
> #1.13
> lm.obj <- lm(airquality$Wind ~ airquality$Temp)
> lm.anova <- anova(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"
> lm.summary <- summary(lm.obj)
> 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
> #1.20
> x1 <- read.table("data/statlog_vehicle_846x18.txt", header = TRUE, sep = "\t")
> dim(x1)
[1] 846  20
> head(x1,5)
  no class compactness circularity distance radiusratio
1 1    0       96           55         103      201
2 2    0       101          56         100      215
3 3    0       93           35         66       154
4 4    0       101          48         107      222
5 5    0       87           38         85       177
  pr.axis max.length scatterratio elongatedness pr.axis.1
1   65        9          204           32        23
2   69        10         208           32        24
3   59        6          142           46        18
4   68        10         208           32        24
5   61        8          164           40        20
  max.length.1 scaledvmi scaledvma scaledradius skewness
1    166         227       624         246        74
2    169         227       651         223        74
```

	no	class	compactness	circularity	distance	radiusratio
3		128	162	304	120	64
4		154	232	641	204	70
5		129	186	402	130	63

	skewness.1	kurtosis	kurtosis.1	hollows
1	6	2	186	194
2	6	5	186	193
3	5	13	197	202
4	5	38	190	202
5	1	25	198	205

```
> tail(x1,5)
```

	no	class	compactness	circularity	distance	radiusratio
842	842	3	87	45	66	139
843	843	3	95	43	76	142
844	844	3	90	44	72	157
845	845	3	89	46	84	163
846	846	3	85	36	66	123

	pr.axis	max.length	scatterratio	elongatedness	pr.axis.1
842	58	8	140	47	18
843	57	10	151	44	19
844	64	8	137	48	18
845	66	11	159	43	20
846	55	5	120	56	17

	max.length.1	scaledvmi	scaledvma	scaledradius	skewness
842	148	168	294	175	73
843	149	173	339	159	71
844	144	159	283	171	65
845	159	173	368	176	72
846	128	140	212	131	73

	skewness.1	kurtosis	kurtosis.1	hollows
842	3	12	188	196
843	2	23	187	200
844	9	4	196	203
845	1	20	186	197
846	1	18	186	190

```
> print(object.size(x1), units = "KB")
69.2 Kb
> #1.28
> x2 <- read.table("data/stock-data.txt", header  = TRUE, skip=1)
> head(x2,5)
```

	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	78.3	69.6	74.30	263,999
2	台積電	100	2	77.0	69.9	72.54	235,159
3	台積電	100	3	72.2	65.7	69.74	276,434
4	台積電	100	4	73.9	68.0	71.37	211,611
5	台積電	100	5	76.9	73.0	74.96	213,185

	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
> tail(x2,5)
   半導體公司 年度 月份 最高價 最低價 加權平均價 成交筆數
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
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
> #1.33
> Dates <- c("180924", "181112", "181231", "181105", "180604", "180219", "180416",
"180611", "180813", "181029")
> Time <- c("01:00", "04:00", "16:00", "23:00", "08:00", "09:00", "07:00", "17:00", "03:00",
"14:00")
> s <- paste(Dates, Time)
> DateTime <- as.POSIXlt(strptime(s, format = "%y%m%d %H:%M", tz = "UTC" ))
> class(DateTime)
[1] "POSIXlt" "POSIXt"
> Items <- as.factor( c("shirt", "shirt", "pants", "jacket", "jacket", "shirt", "jacket", "jacket",
"shoes", "shirt"))
> class(Items)
[1] "factor"
> Volume <- c(7951, 159,1958, 6848, 3762, 3678, 8696, 9045, 6208, 1425)
> class(Volume)
[1] "numeric"
> mySale <- data.frame(DateTime, Items, Volume)
> print(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
> #ex1.33(b)
> Items[Dates >= "0700"]
 [1] shirt  shirt  pants  jacket jacket shirt  jacket jacket
```

```
 [9] shoes  shirt
Levels: jacket pants shirt shoes
> sum(Volume[Dates >= "0700"], na.rm=T)
[1] 49730
```