```
> #2020/11/06(五), 109 學年第一學期 資料科學應用 R 作業(2)
>
> #學號:A107260008 姓名:林玥澔
>
>    ## 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"
>    R <- lm.summary["r.squared"]
> class(R)
[1] "list"
```

```
> R2 <- as.numeric(R)
> class(R2)
[1] "numeric"
> R2^2
[1] 0.04399628
>    # ex1.20
>    my.data <- read.table("statlog_vehicle_846x18 (1).txt", header = T, sep= "\t")
>    str(my.data)
'data.frame':    846 obs. of    20 variables:
  $ no                  : int    1 2 3 4 5 6 7 8 9 10 ...
  $ class               : int    0 0 0 0 0 0 0 0 0 0 ...
  $ compactness     : int    96 101 93 101 87 95 98 107 103 77 ...
  $ circularity     : int    55 56 35 48 38 48 55 53 50 38 ...
  $ distance          : int    103 100 66 107 85 104 101 103 98 63 ...
  $ radiusratio     : int    201 215 154 222 177 214 228 221 212 135 ...
  $ pr.axis           : int    65 69 59 68 61 67 70 66 63 59 ...
  $ max.length      : int    9 10 6 10 8 9 9 11 9 5 ...
  $ scatterratio : int    204 208 142 208 164 205 210 209 193 130 ...
  $ elongatedness: int    32 32 46 32 40 32 31 32 34 52 ...
  $ pr.axis.1        : int    23 24 18 24 20 23 24 24 22 18 ...
  $ max.length.1 : int    166 169 128 154 129 151 168 163 161 130 ...
  $ scaledvmi        : int    227 227 162 232 186 227 236 222 214 145 ...
  $ scaledvma        : int    624 651 304 641 402 628 661 653 567 247 ...
  $ scaledradius : int    246 223 120 204 130 202 245 212 185 139 ...
  $ skewness          : int    74 74 64 70 63 74 72 66 64 79 ...
  $ skewness.1      : int    6 6 5 5 1 5 1 0 5 13 ...
  $ kurtosis          : int    2 5 13 38 25 9 6 1 5 21 ...
  $ kurtosis.1      : int    186 186 197 190 198 186 188 191 198 183 ...
  $ hollows           : int    194 193 202 202 205 193 197 201 204 187 ...
> dim(my.data)
[1] 846    20
> head(my.data, 5)
   no class compactness circularity distance radiusratio pr.axis
1   1       0          96          55      103         201      65
2   2       0         101          56      100         215      69
3   3       0          93          35       66         154      59
4   4       0         101          48      107         222      68
5   5       0          87          38       85         177      61
```

```
    max.length scatterratio elongatedness pr.axis.1 max.length.1
1            9          204            32         23           166
2           10          208            32         24           169
3            6          142            46         18           128
4           10          208            32         24           154
5            8          164            40         20           129
    scaledvmi scaledvma scaledradius skewness skewness.1 kurtosis
1         227       624          246       74          6        2
2         227       651          223       74          6        5
3         162       304          120       64          5       13
4         232       641          204       70          5       38
5         186       402          130       63          1       25
    kurtosis.1 hollows
1          186      194
2          186      193
3          197      202
4          190      202
5          198      205
> tail(my.data, 5)
        no class compactness circularity distance radiusratio pr.axis
842 842     3          87          45       66          139       58
843 843     3          95          43       76          142       57
844 844     3          90          44       72          157       64
845 845     3          89          46       84          163       66
846 846     3          85          36       66          123       55
        max.length scatterratio elongatedness pr.axis.1 max.length.1
842            8          140            47         18           148
843           10          151            44         19           149
844            8          137            48         18           144
845           11          159            43         20           159
846            5          120            56         17           128
        scaledvmi scaledvma scaledradius skewness skewness.1 kurtosis
842         168       294          175       73          3       12
843         173       339          159       71          2       23
844         159       283          171       65          9        4
845         173       368          176       72          1       20
846         140       212          131       73          1       18
        kurtosis.1 hollows
```

```
842             188         196
843             187         200
844             196         203
845             186         197
846             186         190
> object.size(my.data)
70816 bytes
> print(object.size(my.data), units = "Kb")
69.2 Kb
>
> ##ex1.28
> z <- read.table("stock-data.txt", skip = 1, header = T, sep = "\t" )
> dim(z)
[1] 60 10
> head(z, 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(z, 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          成交金
額
56          旺宏   100       8     14.50     10.25          11.84   152,177 8,137,500,167
57          旺宏   100       9     12.65     10.40          11.55   108,879 5,542,998,380
58          旺宏   100      10     12.00     10.25          11.31    68,571 3,041,525,834
59          旺宏   100      11     13.65     10.85          12.54   167,018 9,538,526,797
60          旺宏   100      12     12.85     11.15          12.17   115,192 5,070,210,532
        成交股數  週轉率百分比
56 687,167,610             20.31
57 479,779,350             14.18
```

```
58 268,710,697                7.94
59 760,264,306               22.47
60 416,455,073               12.31
>
> ##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")
> DateTime1 <- paste("2018", Dates, Time)
> DateTime <- strptime(DateTime1, format="%Y %m%d %H:%M",    tz = "UTC")
> Items <- as.factor(Items1)
> Volume <- as.numeric(Volume1)
> 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
>
> ## ex1.33(b)
> id <- 1:length(Dates)
> Q <- id [Dates >=    "0701"]
> mysale[Q, ]
                    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
9    2018-08-13 03:00:00    shoes      6208
10 2018-10-29 14:00:00    shirt      1425
>
```