

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

學號: A106260105 姓名: 黃念慈

(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>

```
> # 1.13
> lm.obj <- lm(airquality$Wind ~ airquality$Temp)
> lm.anova <- anova(lm.obj)
> lm.summary <- summary(lm.obj)
>
> # ex1.13(a)
> class(lm(lm.anova))
[1] "lm"
> 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, "r.squared")
NULL
```

```

>
> # ex1.20
> statlogvehicle846x18 <-
read.delim("data/statlog_vehicle_846x18.txt")
> dim(statlogvehicle846x18)
[1] 846 20
> head(statlogvehicle846x18,4)
  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
  max.length scatterratio elongatedness pr.axis.1 max.length.1
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
  scaledvma scaledradius skewness skewness.1 kurtosis kurtosis.1
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
> tail(statlogvehicle846x18,4)
  no class compactness circularity distance radiusratio pr.axis
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
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
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

```

843      187      200
844      196      203
845      186      197
846      186      190

```

```
> print(object.size(statlogvehicle846x18), units = "Mb")
```

```
0.1 Mb
```

```
>
```

```
> # ex1.28
```

```
> varNames <- c("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", "成交股數", "週轉率百分比")
```

```
> stockdata <- read.table("data/stock-data.txt", header = F, sep =
"\t", col.names = varNames)
```

```
> class(stockdata)
```

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

```
> head(stockdata, 5)
```

半導體公司 年度 月份 最高價 最低價

1 民國 100 年 5 家半導體公司股票月成交資訊 (元, 股)

2 半導體公司 年度 月份 最高價 最低價

3 台積電 100 1 78.3 69.6

4 台積電 100 2 77 69.9

5 台積電 100 3 72.2 65.7

加權平均價 成交筆數 成交金額 成交股數 週轉率百分比

1

2 加權平均價 成交筆數 成交金額 成交股數 週轉率百分比

3 74.3 263,999 100,578,274,926 1,353,616,348 5.22

4 72.54 235,159 74,985,055,548 1,033,654,452 3.98

5 69.74 276,434 88,459,924,495 1,268,289,393 4.89

```
> tail(stockdata, 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

58 旺宏 100 8 14.5 10.25 11.84 152,177 8,137,500,167

59 旺宏 100 9 12.65 10.4 11.55 108,879 5,542,998,380

60 旺宏 100 10 12 10.25 11.31 68,571 3,041,525,834

61 旺宏 100 11 13.65 10.85 12.54 167,018 9,538,526,797

62 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

```

58 687,167,610      20.31
59 479,779,350      14.18
60 268,710,697       7.94
61 760,264,306      22.47
62 416,455,073      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")
> Items <- as.factor(c("shirt", "shirt", "pants", "jacket",
"jacket", "shirt", "jacket", 'jacket', "shoes", "shirt"))
> Volume <- c(7951, 159, 1958, 6848, 3762, 3678, 8696, 9045, 6208,
1425)
> Year <- 2018
> datetime <- paste(Year, Dates, Time)
> DateTime <- as.POSIXct(strptime(datetime, format="%Y %m%d %H:%M",
tz = "UTC"))
> 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)
> mySaleorder <- mySale[order(mySale$DateTime,decreasing = F),]
> mySaleorder7 <- tail(mySaleorder,6)
> Items
[1] shirt shirt pants jacket jacket shirt jacket jacket shoes

```

```
[10] shirt
Levels: jacket pants shirt shoes
> sum(subset(mySaleorder7,Items=="shoes",select="Volume"))
[1] 6208
> sum(subset(mySaleorder7,Items=="shirt",select="Volume"))
[1] 9535
> sum(subset(mySaleorder7,Items=="jacket",select="Volume"))
[1] 6848
> sum(subset(mySaleorder7,Items=="pants",select="Volume"))
[1] 1958
```