20210611
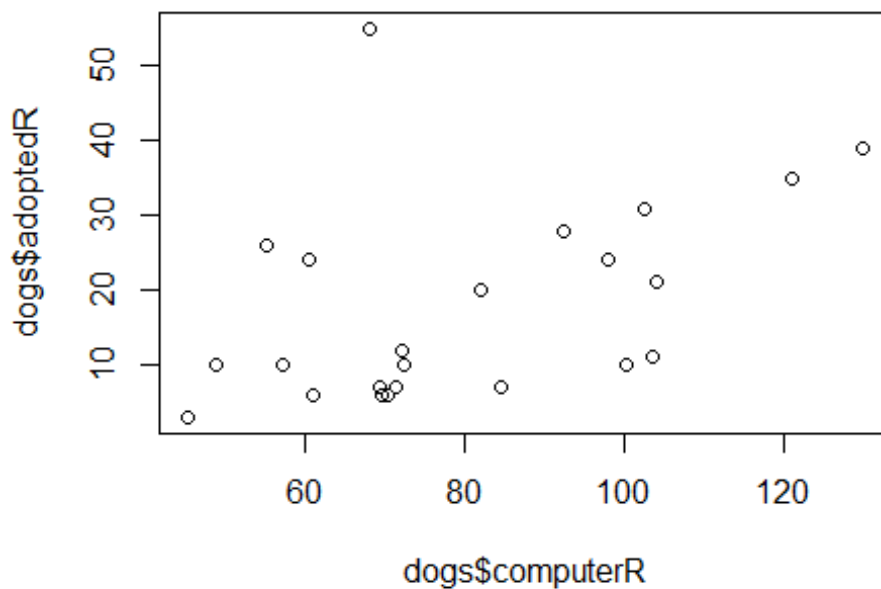
```
#####109 學年度第二學期 R 語言#####
rm(list=ls())
setwd("C:/Users/User/Desktop/R/R-project/R Statistics/data")

#### 2021/06/11 #### 迴歸分析
# import data #
dogs <- read.table("dogs.txt")
dogs <- read.table("dogs.txt", header=T)
plot(dogs$computerR,dogs$adoptedR)
```



```
#若出現亂碼
#colnames(dogs) <- c("a", "b", "c", "d", "e", "f", "g", "h", "i", "j",
"k", "l", "m")

### b0, b1 ### way1
x <- dogs$computerR
Y <- dogs$adoptedR
# b1
beta1 <- sum((x-mean(x))*(Y-mean(Y)))/sum((x-mean(x))^2)
beta1
```

```
## [1] 0.2575605
```

```r
# b0
beta0 <- mean(Y)-beta1*mean(x)
beta0
```

```
## [1] -2.860113
```

```r
### b0, b1 ### way2
slm.model <- lm(adoptedR ~ computerR, data=dogs)
summary(slm.model)
```

```
##
## Call:
## lm(formula = adoptedR ~ computerR, data = dogs)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -12.950  -8.263  -2.939   6.845  40.338
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -2.8601     9.4447  -0.303   0.7650
## computerR     0.2576     0.1137   2.264   0.0343 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 12.18 on 21 degrees of freedom
## Multiple R-squared:  0.1963, Adjusted R-squared:  0.158
## F-statistic: 5.128 on 1 and 21 DF,  p-value: 0.03425
```

```r
####################
### 平方和 ### way1
SSTotal <- sum((dogs$adoptedR - mean(dogs$adoptedR))^2)
SSR <- 0.1963*SSTotal #R-squared 有經過四捨五入
SSE <- SSTotal - SSR
# 以下方法較精確
SSE <- deviance(slm.model)
SSR <- SSTotal - SSE

### 平方和 ### way2
anova(slm.model)
```

```
## Analysis of Variance Table
##
## Response: adoptedR
##           Df  Sum Sq Mean Sq F value  Pr(>F)
## computerR  1  760.77  760.77  5.1277 0.03425 *
## Residuals 21 3115.67  148.37
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#####################
### 區間估計 ### way 1
qt(0.975,21)

## [1] 2.079614

c(-2.8601 - 2.079614*9.4447, -2.8601 + 2.079614*9.4447)

## [1] -22.50143  16.78123

c(0.2576 - 2.079614*0.1137,0.2576 + 2.079614*0.1137)

## [1] 0.02114789 0.49405211

## 區間估計 預設 a=0.05 ## way 2
confint(slm.model)

##                    2.5 %     97.5 %
## (Intercept) -22.50152824 16.7813027
## computerR     0.02102233  0.4940987

confint(slm.model, level=0.95)

##                    2.5 %     97.5 %
## (Intercept) -22.50152824 16.7813027
## computerR     0.02102233  0.4940987

####################
## 其他 ##
names(slm.model)

##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "xlevels"       "call"          "terms"         "model"

View(slm.model$coef)

# 取代 tidy #
tab1 <- tidy(slm.model)
#
ss <- summary(slm.model)
View(ss$coefficients)

### 複迴歸 ###
model2 <- lm(adoptedR ~ computerR + crimeR + captured + killedR,
             data=dogs)
summary(model2)
```

```
##
## Call:
## lm(formula = adoptedR ~ computerR + crimeR + captured + killedR,
##     data = dogs)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -13.108  -3.548  -1.165   3.440  14.741
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.277e+01  1.253e+01   3.414  0.00309 **
## computerR    5.544e-02  9.493e-02   0.584  0.56645
## crimeR       9.753e-03  4.573e-03   2.133  0.04695 *
## captured    -5.213e-05  5.348e-05  -0.975  0.34261
## killedR     -6.488e-01  1.246e-01  -5.209 5.92e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 7.254 on 18 degrees of freedom
## Multiple R-squared:  0.7556, Adjusted R-squared:  0.7013
## F-statistic: 13.92 on 4 and 18 DF,  p-value: 2.423e-05

### 練習題：講義 p.19 (5) ###
x <- c(3.7, 3.6, 2.8, 5.6, 3.3, 2.2, 3.3,
       3.1, 3.2, 3.5, 5.2, 4.6, 5.8, 3)
y <- c(5.7, 5.9, 6.7, 9.5, 5.4, 3.5, 6.2,
       4.7, 6.1, 4.9, 10.7, 7.6, 11.8, 4.1)

d1 <- data.frame(x, y)

xm <- mean(d1$x)
ym <- mean(d1$y)

d1$x_m <- d1$x - xm
d1$y_m <- d1$y - ym
d1$xy_m <- d1$x_m*d1$y_m
d1$x_m2 <- (d1$x_m)^2

b1 <- sum(d1$xy_m)/sum(d1$x_m2)
b0 <- ym - b1*xm

##########
r <- sum(d1$xy_m)/(13*sd(d1$x)*sd(d1$y)) # way1
r1 <- cor(x, y) # way2
r2 <- cov(x, y)/(sd(x)*sd(y)) # way3

b1 <- r*(sd(d1$y)/sd(d1$x))
b0 <- mean(d1$y) - b1*mean(d1$x)
```

```
###########################
x <- c(3.7, 3.6, 2.8, 5.6, 3.3, 2.2, 3.3,
       3.1, 3.2, 3.5, 5.2, 4.6, 5.8, 3)
y <- c(5.7, 5.9, 6.7, 9.5, 5.4, 3.5, 6.2,
       4.7, 6.1, 4.9, 10.7, 7.6, 11.8, 4.1)

d1 <- data.frame(x, y)
#
m1 <- lm(y~x, data = d1)
summary(m1)

##
## Call:
## lm(formula = y ~ x, data = d1)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.1508 -0.7573 -0.2971  0.6446  2.1012
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.2088     0.9949  -1.215    0.248
## x             2.0742     0.2536   8.178    3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.9993 on 12 degrees of freedom
## Multiple R-squared:  0.8479, Adjusted R-squared:  0.8352
## F-statistic: 66.88 on 1 and 12 DF,  p-value: 2.999e-06

#
anova(m1)

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## x          1 66.785  66.785  66.879 2.999e-06 ***
## Residuals 12 11.983   0.999
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

#
confint(m1)

##                  2.5 %    97.5 %
## (Intercept) -3.376484 0.9588061
## x            1.521562 2.6267836
```