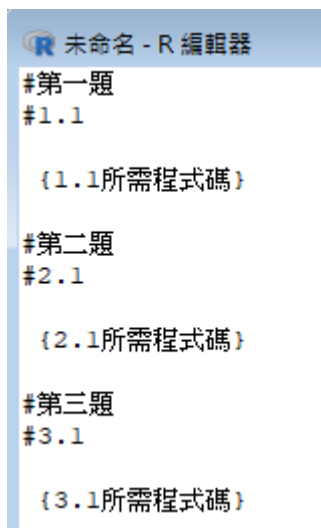


中級程式設計 R 語言應用 & R 語言新手村資料分析篇

期末作業

1. 請注意每一題的題目說明，完成所需要的程式碼。
2. 變數名稱以及函數名稱請使用題目中所設定的名稱命名。若沒規定請自行命名有意義的名稱，並加上註解說明。
3. 東華學生作業檔案請使用學號命名，例如：FIN_410721123。
非本校學生請使用姓名命名，例如：FIN_張小明。
4. 無法執行、執行程式結果不符合題目要求、設定資料有誤、程式無註解、變數名稱沒有使用規定的名稱，斟酌扣分。
5. 每題之間請使用註解表明清楚。例如第一題，請同學在 R 編輯器裡輸入
第一題，然後下一列開始是第一題所需要的程式碼，範例如下。



```
# 未命名 - R 編輯器
#第一題
#1.1
{1.1所需程式碼}

#第二題
#2.1
{2.1所需程式碼}

#第三題
#3.1
{3.1所需程式碼}
```

6. 選擇題答案部份，請將題號和答案寫在 word 檔上，並將 word 檔轉成 pdf 檔後，和程式題所要繳交的.R 檔，壓縮後一起上傳繳交。

第一部份:單選題，總共 10 題，每題 5 分，總分 50 分。請將題號和答案使用 word 作答完後，轉成 pdf 檔，並且和程式題的.R 檔一起壓縮後上傳。

1. R 語言宣告資料框時，需要使用下列那個函數。

- (A) c()
- (B) data.frame()
- (C) list()
- (D) for()

2. R 語言寫自訂函數時，需要使用下列那個函數。

- (A) sample()
- (B) names()
- (C) function()
- (D) list()

3. 選取資料框整個欄位資料時，使用下列那一個符號可以幫助我們更有效率。

- (A) %%
- (B) \$\$
- (C) \$
- (D) **

4. 現在有三筆學生資料向量分別是 StudentId, StudentSex, StudentName, 要合併成列表，並將列表中的每個項目命名為學生證號碼、學生性別、學生姓名。請問以下那一個程式碼，可以幫助我們獲得下列合併成列表的結果：

```
$學生證號碼  
[1] 1234 5678 4567  
  
$學生性別  
[1] "Male" "Female" "Male"  
  
$學生姓名  
[1] "John" "Mary" "Jack"
```

- (A) StudentList <- matrix(StudentId, StudentSex, StudentName)
c(StudentList) <- c("學生證號碼", "學生性別", "學生姓名")
StudentList
- (B) StudentList <- list(StudentId, StudentSex, StudentName)
c(StudentList) <- c("學生證號碼", "學生性別", "學生姓名")
StudentList
- (C) StudentList <- list(StudentId, StudentSex, StudentName)
names(StudentList) <- data.frame("學生證號碼", "學生性別", "學生姓名")
StudentList
- (D) StudentList <- list(StudentId, StudentSex, StudentName)
names(StudentList) <- c("學生證號碼", "學生性別", "學生姓名")
StudentList

5. 以下程式片段為使用 `sample()` 亂數函數模擬擲骰子的情況，請問如果只執行此程式片段一次，有可能會出現下列那一項結果。

`sample(1:6, 7)`

- (A) 4 6 3 2 1 5
(B) 6 4 6 1
(C) 3 2 5 4 1 6 7
(D) 錯誤，不能取用比總體還大的樣本
6. 在 R 語言中，如果需要使用亂數，我們會使用亂數函數，請問下列那一個函數可以幫助我們完成。
(A) `function()`
(B) `sample()`
(C) `names()`
(D) `geom_point()`
7. 以下程式碼為成績判斷程式，假設成績變數設為 `x`，如果 `x` 的值大於等於九十分就輸出 A。如果 `x` 的值小於九十分，並且大於等於七十分，就輸出 B。當 `x` 的值小於七十分就輸出 C，請問以下程式碼那一個正確。

- (A)

```
y <- 75
if (x > 90) {
  "A"
} else if (x < 90 & x >= 70) {
  "B"
} else {
  "C"
}
```
- (B)

```
x <- 75
if (x >= 90) {
  "A"
} else if (70 <= x < 90) {
  "B"
} else {
  "C"
}
```
- (C)

```
x <- 75
if (x >= 90) {
  "A"
} else if (x < 90 & x >= 70) {
  "B"
} else {
  "C"
}
```
- (D)

```
x <- 75
if (x >= 90) {
  "A"
} else if (x < 90) {
  "B"
} else {
  "C"
}
```

8. 以下程式碼為一種排序程式，可以幫助我們做數字的排序，假設輸入一個有五個數字的數值向量到此排序程式中，請問當迴圈 `i = 1`，`j = 3` 執行完後，此數值向量內的數值排列順序是下列那一個。

```
numbers <- c(5, 2, 4, 3, 1)
Sort <- function(numbers){
  temp <- 0
  for(i in 1:5){
    swap <- 0
    for(j in 1:4){
      if(numbers[j] > numbers[j+1]){
        temp <- numbers[j]
```

```

        numbers[j] <- numbers[j+1]
        numbers[j+1] <- temp
        swap <- swap + 1
    }
}
if(swap == 0) break
}
return(numbers)
}
Sort(numbers)

```

- (A) 2, 4, 3, 5, 1
- (B) 1, 2, 5, 4, 3
- (C) 5, 4, 1, 2, 3
- (D) 4, 3, 2, 1, 5

9. 假設 $x = 5$ ， $y = 10$ ，請問以下那一個程式可以幫我們完成**交換** x 和 y 的數值，使得 $x = 10$ ， $y = 5$ 。

- | | | | |
|------------------------|--------------------|------------------------|--------------------|
| (A) $\text{temp} <- x$ | (B) $x <- y$ | (C) $\text{temp} <- x$ | (D) $y <- x$ |
| $y <- x$ | $\text{temp} <- x$ | $x <- y$ | $\text{temp} <- x$ |
| $x <- \text{temp}$ | $y <- \text{temp}$ | $y <- \text{temp}$ | $x <- y$ |

10. 當我們需要**快速有效率**的宣告 100000 個內容為 0 的向量時，下列那一個程式片段可以達到我們的需求。

- (A) `rep(NA, times = 100000)`
- (B) `rep(0, times = 100000)`
- (C) `rep(NA, times = 100000)`
- (D) `rep(0, times = 100000)`

第二部份:程式題，總共 2 題，每題 25 分，總分 50 分。請儲存成.R 檔後，和單選題的 pdf 檔一起壓縮上傳。

第一題(25%): 請為以下的核准貸款資料表，建立一個決策樹。目的是預測以下情況是否核准貸款。

核准貸款資料表:

年紀	工作	房子	信用	核可貸款
年輕	無	無	一般	否
年輕	無	無	好	否
年輕	有	無	好	是
年輕	有	有	一般	是
年輕	無	無	一般	否
中年	無	無	一般	否
中年	無	無	好	否
中年	有	有	好	是
中年	無	有	極好	是
中年	無	有	極好	是
老年	無	有	極佳	是
老年	無	有	好	是
老年	有	無	好	是
老年	有	無	極佳	是
老年	無	無	一般	否

請問此情況是否核准貸款?

年紀	工作	房子	信用	核可貸款
年輕	無	無	好	???

第二題(25%): 請分析特拉斯執行長，伊隆·馬斯克(Elon Musk)訪談內容中，各詞彙出現的次數，將出現次數前六多的詞彙，使用長條圖表示。輸出答案要有前六多的詞彙資料框以及相對應的長條圖。

訪談內容檔案: Elon_Musk_interview.txt