

# 網頁資料擷取與分析 R 3.5

## 範例試卷

### 【認證說明與注意事項】

- 一、本項考試為操作題，所需總時間為 90 分鐘，時間結束前需完成所有考試動作。成績計算滿分為 100 分，合格分數為 70 分。
- 二、操作題為四大題，第一大題 15 分，第二大題 20 分，第三大題 35 分，第四大題 30 分，總計 100 分。
- 三、操作題所需的檔案皆於 C:\ANS.CSF\各指定資料夾內讀取。題目存檔方式，請依題目指示儲存於 C:\ANS.CSF\各指定資料夾，測驗結束前必須自行存檔，並關閉開發環境，檔案名稱錯誤或未自行存檔者，均不予計分。
- 四、本項考試禁止使用外部套件，否則該題不予計分。
- 五、輸入與輸出的格式必須完全相同，每一行字、空白都要一樣，特別注意輸出後有無空白與換行。
- 六、注意全型、半型字元、英文字母大小寫、小數點的位數是否與題目的要求相同。
- 七、每一題於測驗時間內可進行重複送評。請提交程式碼檔案，而非執行檔或執行的結果。請注意提交的檔案是否適用該題目（請檢查有無交錯題目）。
- 八、每一題至少有 1 組評分測試資料顯示於題目中，且至少有 1 組隱藏的評分測試資料。
- 九、題目如有需要進行檔案讀寫，在本機撰寫程式碼自行測試時，程式開啟檔案或寫入檔案的路徑，是依據您電腦中程式的啟動位置。在提交評分時，程式所開啟或寫入的檔案，必須與程式碼檔在同一層。

壹、操作題 100%(第一題 15 分、第二題 20 分、第三題 35 分、第四題 30 分)

## 一、新北市公共自行車即時資訊

### 1. 作答須知：

- (1) 請至 C:\ANS.CSF\RW01 資料夾開啟 **RWD01.r** 編寫，完成結果請儲存於 C:\ANS.CSF\RW01 資料夾，檔案名稱請另存新檔為 **RWA01.r**。
- (2) 題目所需之檔案皆位於 C:\ANS.CSF\RW01 資料夾。
- (3) 程式所產出的檔案，須輸出與程式同一層資料夾。

### 2. 設計說明：

- (1) 請撰寫一程式，讀取新北市公共自行車即時資訊 **read.xml**，請將其中 sno (站點代號)、sna (中文場站名稱)、tot (場站總停車格) 等三個欄位轉存為 **write.csv** (需為 UTF-8 編碼格式)，各欄位內容之間以一個半形逗號隔開。

\* 提示：只需要輸出資料，不需要輸出欄位名稱。

### 3. 輸入輸出：

#### (1) 輸入說明

讀取 **read.xml**

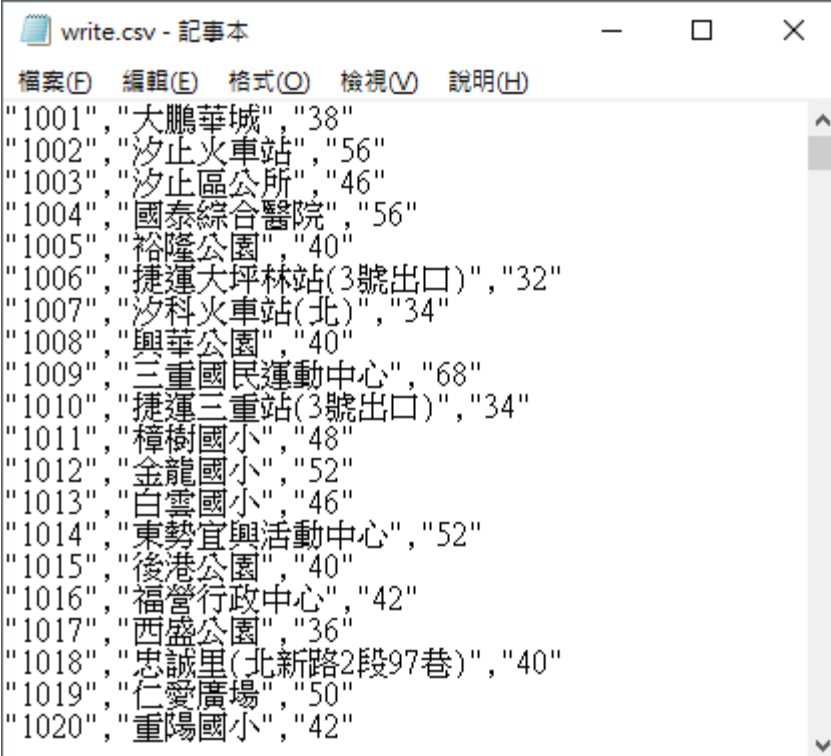
#### (2) 輸出說明

將三個欄位的內容：sno、sna、tot，輸出至 **write.csv** 檔案，各欄位內容之間以一個半形逗號隔開

#### (3) 範例輸入

無

範例輸出



```
"1001","大鵬華城","38"  
"1002","汐止火車站","56"  
"1003","汐止區公所","46"  
"1004","國泰綜合醫院","56"  
"1005","裕隆公園","40"  
"1006","捷運大坪林站(3號出口)","32"  
"1007","汐科火車站(北)","34"  
"1008","興華公園","40"  
"1009","三重國民運動中心","68"  
"1010","捷運三重站(3號出口)","34"  
"1011","樟樹國小","48"  
"1012","金龍國小","52"  
"1013","白雲國小","46"  
"1014","東勢宜興活動中心","52"  
"1015","後港公園","40"  
"1016","福營行政中心","42"  
"1017","西盛公園","36"  
"1018","忠誠里(北新路2段97巷)","40"  
"1019","仁愛廣場","50"  
"1020","重陽國小","42"
```

4. 評分項目：

項	目	配分	得分
(1)	符合設計說明輸出正確格式	15	
總	分	15	

## 二、美元收盤匯率

### 1. 作答須知：

- (1) 請至 C:\ANS.CSF\RW02 資料夾開啟 **RWD02.r** 編寫，完成結果請儲存於 C:\ANS.CSF\RW02 資料夾，檔案名稱請另存新檔為 **RWA02.r**。
- (2) 題目所需之檔案皆位於 C:\ANS.CSF\RW02 資料夾。
- (3) 程式所產出的檔案，須輸出與程式同一層資料夾。

### 2. 設計說明：

- (1) 請撰寫一程式，爬取 **read.html**，取得「新臺幣對美元銀行間成交之收盤匯率」資料，並將其中日期、NTD/USD 兩個欄位的名稱與資料轉存為 **write.csv**（需為 UTF-8 編碼格式）。

### 3. 輸入輸出：

#### (1) 輸入說明

爬取 **read.html** 內的資料

#### (2) 輸出說明

日期、NTD/USD 兩個欄位的名稱與資料，輸出至 **write.csv**

#### (3) 範例輸入

無

範例輸出

```
write.csv - 記事本
檔案(F) 編輯(E) 格式(O) 檢視(V) 說明(H)
"date","NTD_USD"
"2018-06-26",30.412
"2018-06-25",30.403
"2018-06-22",30.307
"2018-06-21",30.302
"2018-06-20",30.17
"2018-06-19",30.186
"2018-06-15",30.002
"2018-06-14",29.938
"2018-06-13",29.89
"2018-06-12",29.86
"2018-06-11",29.821
"2018-06-08",29.816
"2018-06-07",29.74
```

### 4. 評分項目：

項	目	配分	得分
(1)	符合設計說明輸出正確格式	20	
總	分	20	

### 三、矩陣

#### 1. 作答須知：

- (1) 請至 C:\ANS.CSF\RW03 資料夾開啟 **RWD03.r** 編寫，完成結果請儲存於 C:\ANS.CSF\RW03 資料夾，檔案名稱請另存新檔為 **RWA03.r**。

#### 2. 設計說明：

- (1) 利用程式內提供的數據，依下列要求進行輸出：
  - (a) 請隨機產生 5~15 之間，15 個正整數並輸出
  - (b) 請將 (a) 轉成 3x5 的 X 矩陣並輸出
  - (c) 請輸出 X 矩陣的最大值
  - (d) 請輸出 X 矩陣的最小值
  - (e) 請輸出 X 矩陣的總和
  - (f) 請輸出 X 矩陣四個角落的元素內容

#### 3. 輸入輸出：

##### (1) 輸入說明

無

##### (2) 輸出說明

- (a) 請隨機產生 5~15 之間，15 個正整數並輸出
- (b) 請將 (a) 轉成 3x5 的 X 矩陣並輸出
- (c) 請輸出 X 矩陣的最大值
- (d) 請輸出 X 矩陣的最小值
- (e) 請輸出 X 矩陣的總和
- (f) 請輸出 X 矩陣四個角落的元素內容

##### (3) 範例輸入

無

範例輸出

[1] "隨機正整數："

[1] 8 13 9 14 15 5 10 14 11 10 15 9 12 11 6

[1] "X 矩陣內容："

[,1] [,2] [,3] [,4] [,5]

[1,] 8 14 10 10 12

[2,] 13 15 14 15 11

[3,] 9 5 11 9 6

[1] "最大：15"

[1] "最小：5"

[1] "總和：162"

[1] "四個角落元素："

[,1] [,2]

[1,] 8 12

## 4. 評分項目：

項	目	配分	得分
(1)	符合設計說明輸出正確格式	35	
總	分	35	

#### 四、市場成交行情：折線圖

##### 1. 作答須知：

- (1) 請至 C:\ANS.CSF\RW04 資料夾開啟 **RWD04.r** 編寫，完成結果請儲存於 C:\ANS.CSF\RW04 資料夾，檔案名稱請另存新檔為 **RWA04.r**。
- (2) 題目所需之檔案皆位於 C:\ANS.CSF\RW04 資料夾。
- (3) 程式所產出的檔案，須輸出與程式同一層資料夾。

##### 2. 設計說明：

- (1) 請讀取果菜市場香蕉成交行情 **read.csv** 資料，主要有兩個欄位：成交日期、成交平均價。再以 **plot** 函數輸出折線圖 **chart.png**，輸出圖表的參數如下：
  - 圖表標題：Market Average Price
  - 以成交日期為 X 軸，X 軸名稱：date
  - 以成交平均價為 Y 軸，Y 軸名稱：NT\$
  - Y 軸下限 15、上限 25

##### 3. 輸入輸出：

###### (1) 輸入說明

讀取 **read.csv** 的內容

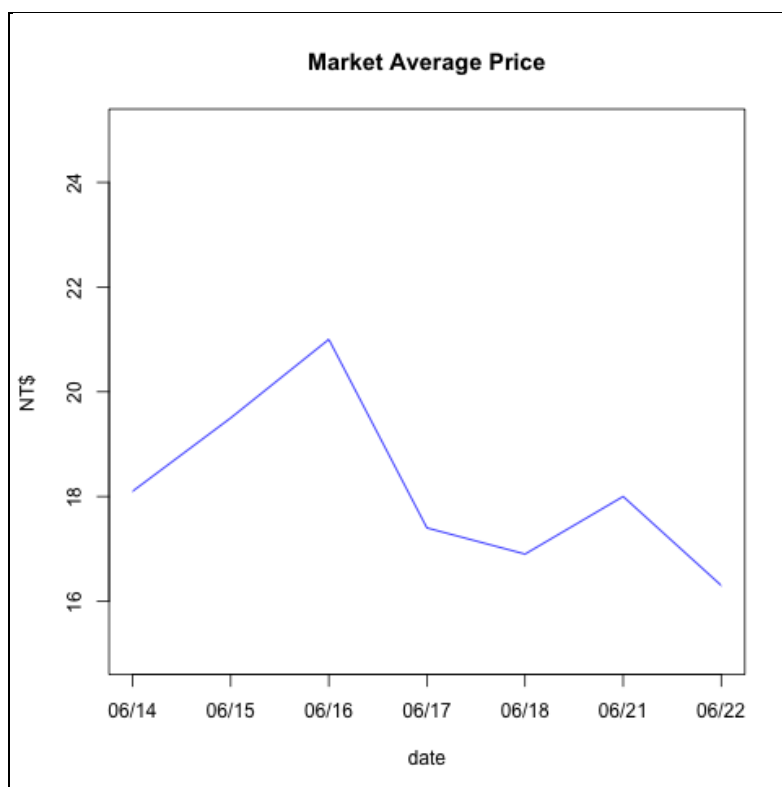
###### (2) 輸出說明

輸出 **chart.png** 圖檔

###### (3) 範例輸入

無

範例輸出



\* 注意：base plotting system 的版本，會造成輸出的圖檔有差異，但不影響評分的準確性。評分系統會將您的程式，於伺服器中運行輸出圖檔進行評分。

4. 評分項目：

項	目	配分	得分
(2)	符合設計說明輸出正確格式	30	
總	分	30	