

Topic: DBF 資料檔匯入 R 之操作說明

Author: Ming-Chang Lee

Date: 2010/01/03

文章說明採用二種方法將 DBF 檔案匯入至 R。方法 1 是將 DBF 轉換為 TXT 再匯入 R，方法 2 是使用 foreign package 直接將 DBF 匯入 R。

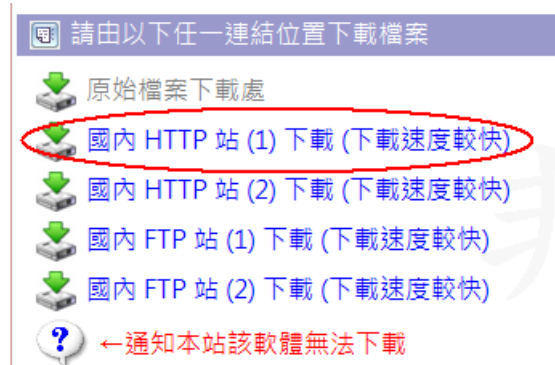
### 方法 1：DBF 轉換為 TXT 再匯入 R

#### 1. 安裝 DBFViewer

##### Step 1

首先下載可開啟 DBF 檔案 之免費軟體 「DBFViewer」，下載首頁。

<http://www.softking.com.tw/soft/download.asp?fid3=14729>，選取任何下載點。



##### Step 2

將下載後 DBFVIEW.ZIP 解壓縮至桌面，產生二個檔案：「INSTALL.EXE」與「FILE\_ID.DIZ」。

##### Step 3

執行 INSTALL.EXE 並安裝軟體。

## 2. 啟動 DBFViewer 並轉換檔案至 TXT

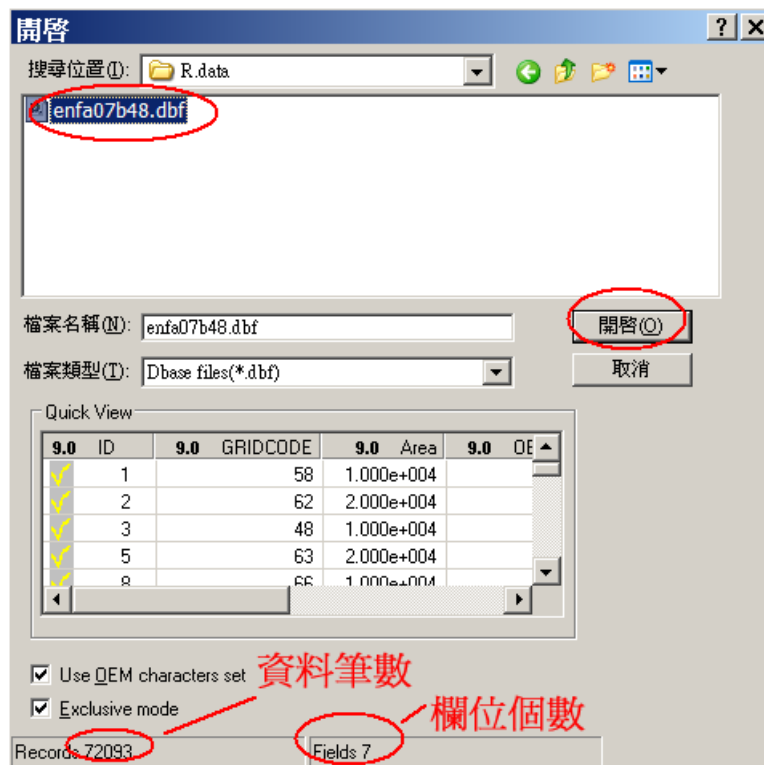
## Step 1

啟動 DBFViewer 2000。



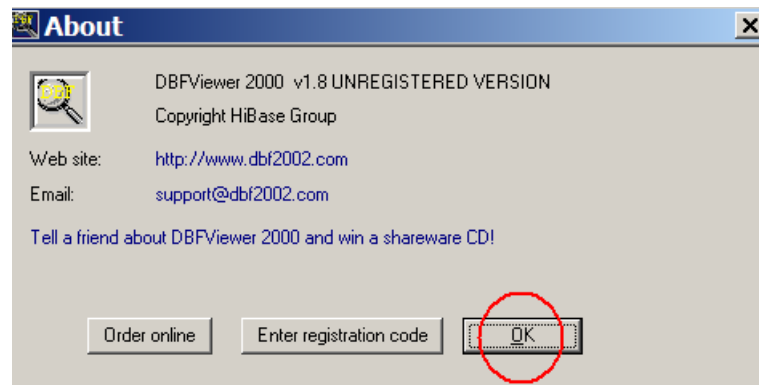
## Step 2

開啟 視窗選取 enfa07b48.dbf (本例將資料儲存於 C:\R.data 目錄) → 按 開啟



## Step 3

顯示提示視窗 → 按 「OK」。



## Step 4

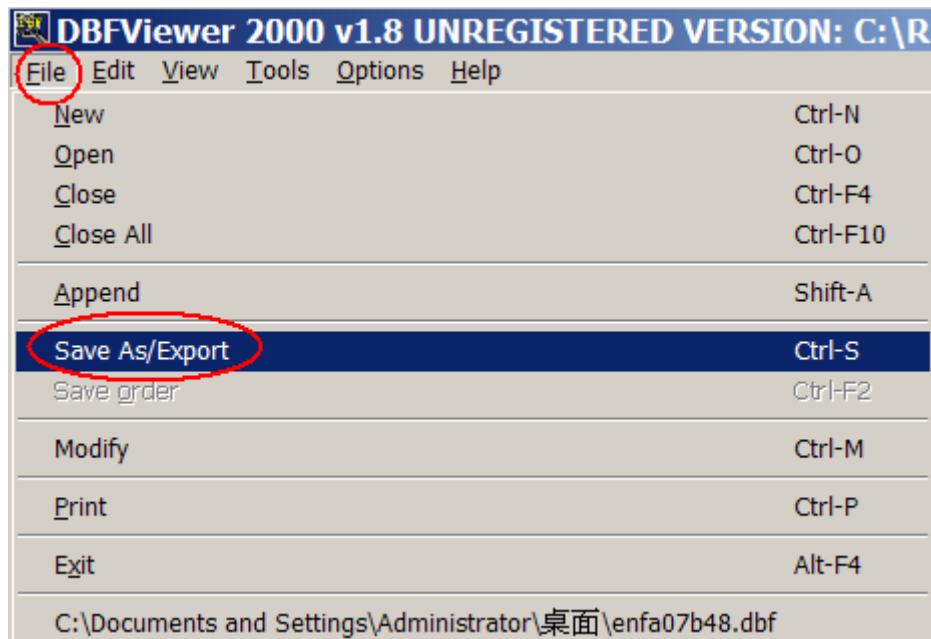
視窗顯示 72093 筆資料與欄位名稱。

	9.0 GRILL OD	9.0 Arr	9.0 OBS (VE)	9.0 ENF	9.0 areal	ab	SPEC ES
1	58	1.000e+004	0	0.580000	0.010000	grus	
2	62	2.000e+004	0	0.620000	0.020000	grus	
3	48	1.000e+004	0	0.480000	0.010000	grus	
5	63	2.000e+004	0	0.630000	0.020000	grus	
8	66	1.000e+004	0	0.660000	0.010000	grus	
9	54	1.000e+004	0	0.540000	0.010000	grus	
10	60	1.000e+004	0	0.600000	0.010000	grus	
11	55	1.000e+004	0	0.550000	0.010000	grus	
13	58	1.000e+004	0	0.580000	0.010000	grus	
14	53	2.000e+004	0	0.530000	0.020000	grus	
16	56	1.000e+004	0	0.560000	0.010000	grus	
17	57	3.000e+004	0	0.570000	0.030000	grus	
19	61	1.000e+004	0	0.610000	0.010000	grus	
20	57	1.000e+004	0	0.570000	0.010000	grus	
21	58	1.000e+004	0	0.580000	0.010000	grus	

Records: 72093    4858.08 KB    Last mod: 31.12.2009 00:50:22    OEM    20:13:05

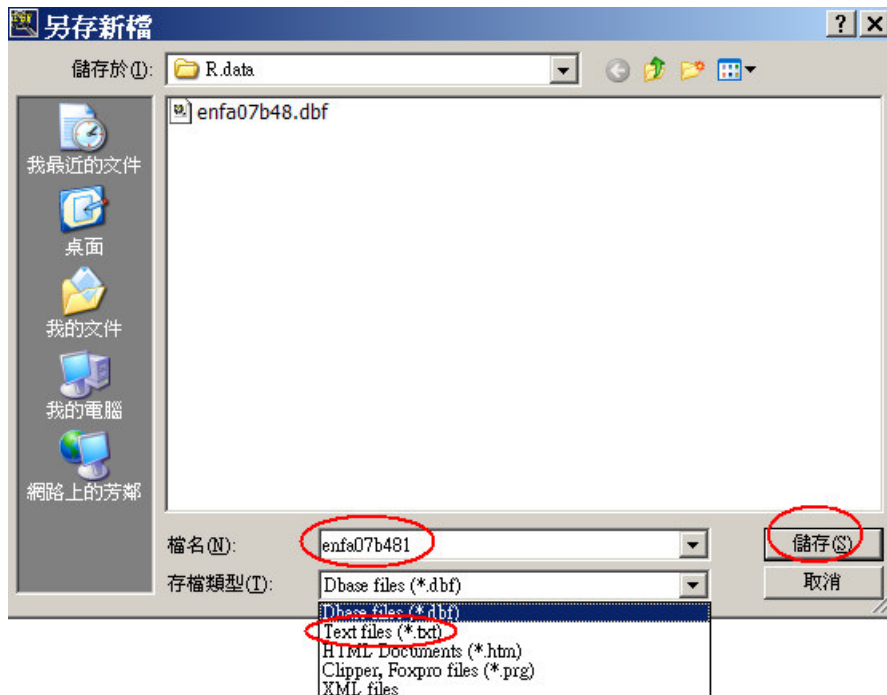
## Step 5

選取 File → Save As/Export 準備匯出成 TXT 文字檔。



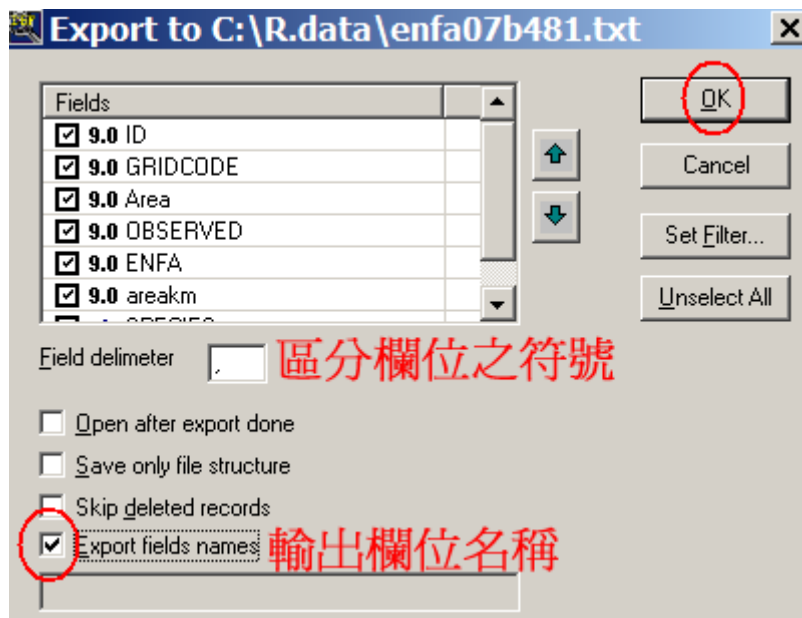
## Step 6

另存新檔 視窗 → 輸入 檔名 → 選取 「Text files (\*.txt)」 存檔類型 → 儲存。



## Step 7

設定 Field delimiter 為「,」 → 將 Export fields names 打勾 → 按 OK。

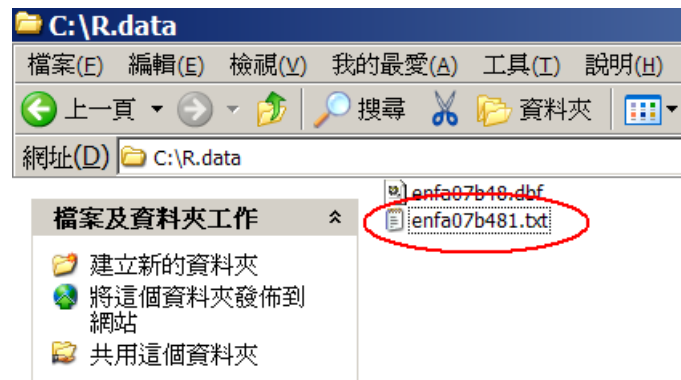


## Step 8

關閉 DBFViewer 2000。

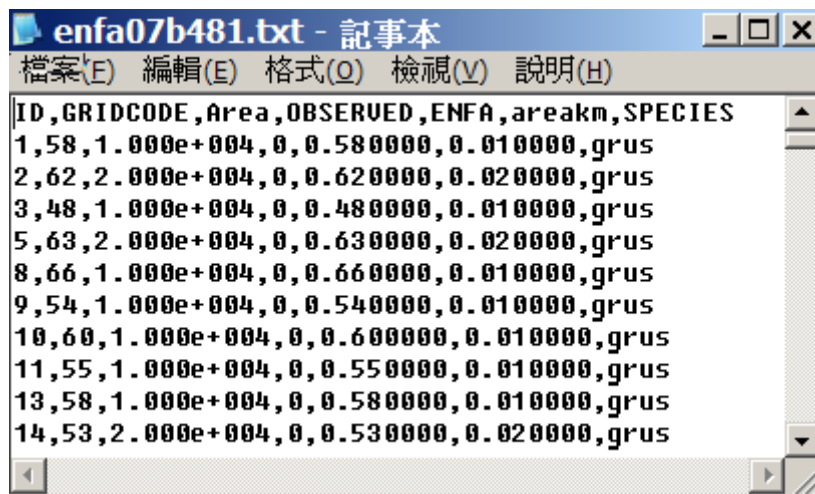
## Step 9

開啟 C:\R.data\ enfa07b481.txt 。



## Step 10

DBF 檔案已成功轉換為 TXT 。



### 3. 匯入 TXT 檔案至 R

參考 R 執行結果，其中「enfa07b48.dbf」已轉換為「enfa07b481.txt」。

```
> # Topic: AUC{PresenceAbsence}
> # Date: 2010/01/03
> # Author: ALAN LEE
> # Web: http://web.ydu.edu.tw/~alan9956/
> # Set the working directory
> workpath <- "C:/R.data"
> setwd(workpath)
> getwd() # Get working directory
[1] "C:/R.data"
> # Load the library
> library(PresenceAbsence)
> # Import txt file into R
> # 7 個欄位名稱: ID, GRIDCODE, Area, OBSERVED, ENFA, areakm, SPECIES
> speciology.tmp <- read.table(file="enfa07b481.txt", header = TRUE,
sep = ",")
> speciology.tmp[1:3,] # View the top 3 records
  ID GRIDCODE  Area OBSERVED ENFA areakm SPECIES
1  1      58 10000         0 0.58  0.01   grus
2  2      62 20000         0 0.62  0.02   grus
3  3      48 10000         0 0.48  0.01   grus
> # Note: 依 auc {PresenceAbsence} 線上說明, 使用 PresenceAbsence 資料格式
> # DATA[,1] plot ID text
> # DATA[,2] observed values zero-one values
> # DATA[,3] predicted probabilities from first model numeric (between
0 and 1)
> # DATA[,4] predicted probabilities from second model, etc...
> # 使用三欄為 SPECIES, PRESENCE, 物種機率為 ENFA.
> # 選取指定的欄位: 物件名稱$欄位名稱
> id <- speciology.tmp$ID
> observed <- speciology.tmp$OBSERVED
> enfa <- speciology.tmp$ENFA
> speciology <- data.frame(id, observed, enfa)
> speciology[1:3,]
```

```
id observed enfa
1 1          0 0.58
2 2          0 0.62
3 3          0 0.48
> auc(speciology)
      AUC      AUC.sd
1 0.8682436 0.01553680
> # end
```

## 方法 2：直接使用 foreign package 匯入 R

R - foreign package 可匯入多種檔案格式：ARFF, DBF, Minitab Portable Worksheet, SPSS, SAS Permanent Dataset 等資料格式。主要修改以下內容：

```
> library(foreign)
> speciology.tmp <- read.dbf(file="enfa07b48.dbf")
```

參考 R 執行結果，其中直接使用匯入「enfa07b48.dbf」方法。

```
> # Topic: AUC{PresenceAbsence}
> # Date: 2010/01/03
> # Author: ALAN LEE
> # Web: http://web.ydu.edu.tw/~alan9956/
> # Set the working directory
> workpath <- "C:/R.data"
> setwd(workpath)
> getwd() # Get working directory
[1] "C:/R.data"
> # Load the library
> library(foreign)
Warning message:
package 'foreign' was built under R version 2.10.1
> library(PresenceAbsence)
Warning message:
package 'PresenceAbsence' was built under R version 2.10.1
> # Import txt file into R
> # 7 個欄位名稱: ID, GRIDCODE, Area, OBSERVED, ENFA, areakm, SPECIES
> speciology.tmp <- read.dbf(file="enfa07b48.dbf")
> speciology.tmp[1:3,] # View the top 3 records
  ID GRIDCODE  Area OBSERVED ENFA areakm SPECIES
1  1      58 10000         0 0.58  0.01   grus
2  2      62 20000         0 0.62  0.02   grus
3  3      48 10000         0 0.48  0.01   grus
> # Note: 依 auc {PresenceAbsence} 線上說明, 使用 PresenceAbsence 資料格式
> # DATA[,1] plot ID text
> # DATA[,2] observed values zero-one values
> # DATA[,3] predicted probabilities from first model numeric (between
```



```
0 and 1)
> # DATA[,4] predicted probabilities from second model, etc...
> # 使用三欄為 SPECIES, PRESENCE ,物種機率為 ENFA.
> # 選取指定的欄位：物件名稱$欄位名稱
> id <- speciology.tmp$ID
> observed <- speciology.tmp$OBSERVED
> enfa <- speciology.tmp$ENFA
> speciology <- data.frame(id, observed, enfa)
> speciology[1:3,]
  id observed enfa
1  1         0 0.58
2  2         0 0.62
3  3         0 0.48
> auc(speciology)
      AUC      AUC.sd
1 0.8682436 0.01553680
> # end
```

--- END ---