

如何使用「學術調查研究資料庫」的調查資料

王文心

本文的目的是希望讓有興趣使用二手資料的人士，瞭解如何從「學術調查研究資料庫」中，找到符合其所需要的資料，並在取得資料後，能夠順利的進行分析。

一、瞭解申請及使用資料規定

「學術調查研究資料庫」蒐集國內以面訪、郵寄問卷、或電話訪問等方式得到的調查資料，這些調查資料的研究領域包括社會、政治、經濟、教育、心理、公共衛生等。在經過我們對資料的分類整理與檢誤後對外開放，供學術界人士參考使用。基於維護資料提供者與學術調查研究資料庫的權益，本室訂定有資料釋出辦法，規定申請資格、申請手續、及使用規則（詳細辦法請見本期通訊第 66 頁）。在您申請資料前請先查看您是否符合申請資格，以及在您得到資料後所必須遵守的使用規則。

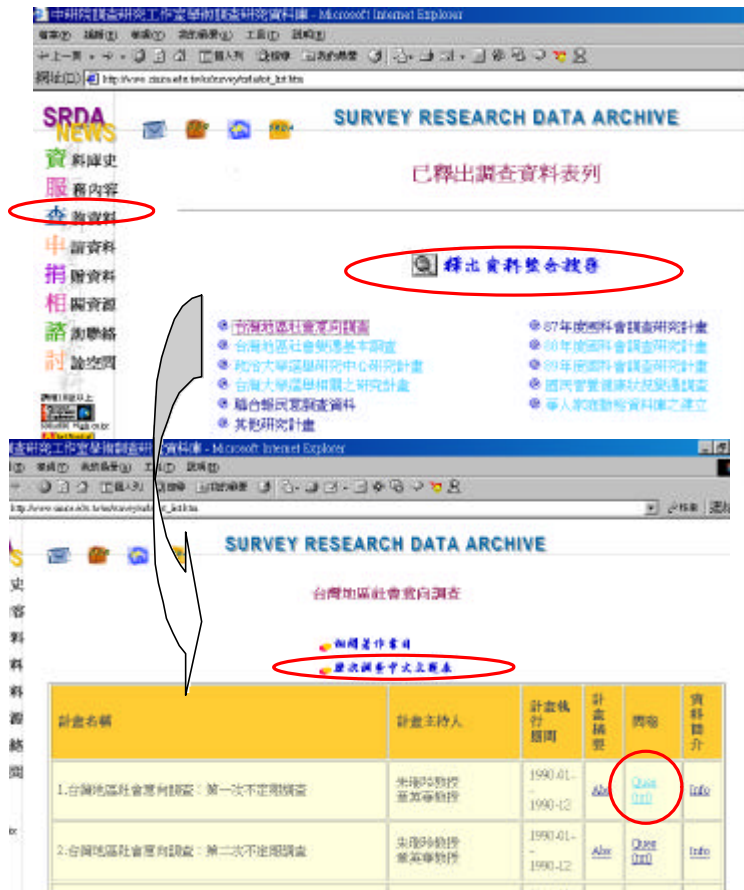
二、如何找尋適合您的調查資料

除親自前來本室查閱書面資料外，我們架設網站，透過網路提供豐富的資訊與便捷的檢索方式，協助有需要使用二手調查資料的人士，找到適合使用的資料。

學術調查研究資料庫的網址為：
<http://www.sinica.edu.tw/as/survey/srda/>，在網站上可查詢的項目如下（圖一）：

1. 整合搜尋：目前提供計畫名稱、關鍵字與計畫主持人的整合搜尋，近期內將提供全文檢索，可以任意字查詢問卷、及摘要的內容。
2. 主題表：大型資料如社會變遷、社會意向、華人家庭動態資料庫等，本資料庫網站上提供已彙整好的歷次調查主題表，您可以利用主題表快速找到相關資料。
3. 空白問卷：我們在網路上放置各筆調查計畫使用的空白問卷以供瀏覽，其中有純文字（txt）和 PDF 兩種格式，如欲瀏覽 PDF 問卷時，申請者端的電腦須先裝 Acrobat Reader（此為免費的共享軟體，我們網站上有提供下載軟體的連結點），利用滑鼠單點問卷位置，即可開啟檔案閱讀問卷。
4. 資料簡介（Info）：在此項目中可查詢調查方式、母體地區、抽樣方式、樣本數、使用的統計軟體、可提供的資料項目等資訊。

圖一 「學術調查研究資料庫」網站



三、取得資料的方式

1. 電子資料：

基於目前網路的方便性，我們建議申請者以電子郵件附檔的方式來收取所申請的資料，不但可縮短取得資料的時間，亦可以省下儲存媒介（磁片、光碟片）和郵寄的費用。

2. 書面資料：

「學術調查研究資料庫」盡量將全部蒐集到的資料電子化，除以下兩種狀況外，原則上我們建議使用者利用申請的電子檔，自行列印書面資料。

- (1) 原計畫有提供裝訂好的紙本：目前可供索取的紙本資料，包括社會變遷及社會意向各次調查的問卷、執行報告書，以及這兩項調查的部份督導手冊或訪員手冊。
- (2) 只有書面資料：本室限於人力物力或資料本身的限制，有些資料仍無法做到全部電子化，例如台灣地區生育力調查

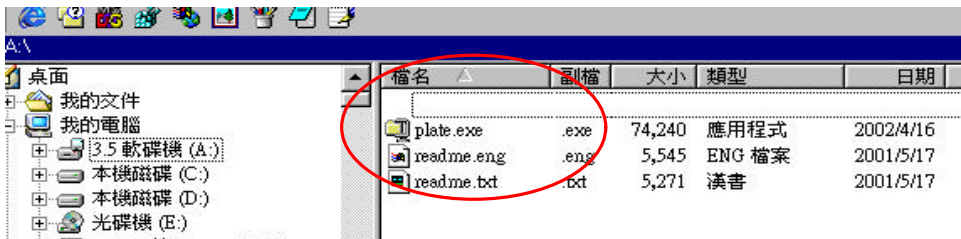
(KAP)，該調查目前只有資料檔是電子檔，其餘部份只有書面資料。

四、資料內容：

為讓資料使用者在作二次分析時，能正確地依循原計畫主持人的研究設計及專業考量，我們盡力提供表一的資料項目供申請者參考使用。為了避免檔案過大，使得申請者的信箱無法接

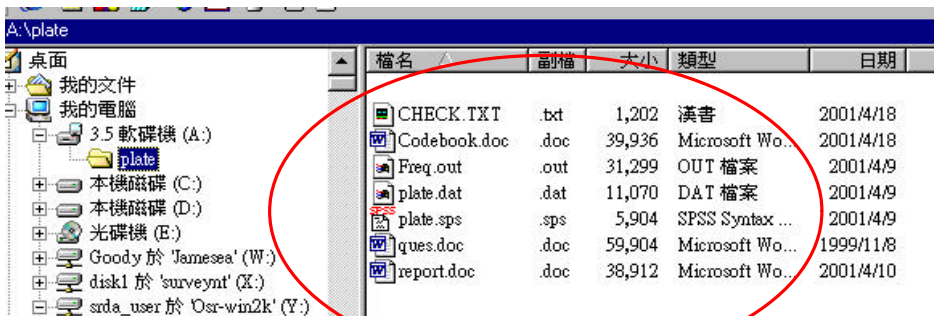
收，我們在釋出資料時，會將讀我檔案 (readme.txt) 之外的其他資料項目，以 WinZip 格式壓縮成自解壓縮檔 (.exe)。申請者取得資料後，請先閱讀讀我檔案，瞭解是否是您申請的資料。以圖二為例，使用者申請後所收到的有三個檔案，分別為 plate.exe (相關調查資料) 的壓縮檔、英文讀我檔案 (readme.eng) 與中文讀我檔案 (readme.txt)。

圖二 在電腦檔案總管中看到「學術調查研究資料庫」提供申請者的檔案名稱畫面



資料確定無誤後，請依據讀我檔案中的解壓縮說明，將資料解壓縮到您指定的資料夾下，請記住您指定的資料夾路徑，這將來在利用統計軟體讀取資料時會使用到。在此例中，我們將 plate.exe 這個檔案解壓縮到 A 磁碟片的 Plate 資料夾下。

圖三 解壓縮後的在檔案總管中看到的檔案名稱畫面



完成解壓縮程序後，便可以在資料夾中，見到我們所提供的資料項目（圖三）。「學術調查研究資料庫」提供的資料項目與檔案的命名原則如表一所示，對照圖三右邊解壓縮後的檔名，可以知道我們提供的資料項目，由上而下依序是：資料整理報告、過錄編碼簿、次數分配結果、ASCII 格式的資料檔、SPSS 語法的欄位定義程式、問卷、以及研究報告書。

表一 資料項目與檔案名稱對照表

資料項目	檔案命名原則	說明
摘要	Abs*.txt	摘要為純文字格式，中文的摘要檔名為 Abs.txt，英文摘要檔名為 Abs_en.txt。
過錄編碼簿	Codebook.doc	記載該筆調查資料問卷每個答案對應的變項名稱、欄位、數值意涵、原始題目內容等，若該資料有兩份以上的過錄編碼簿時，我們會用 Code1.doc 及 Code2.doc ..來表示第一份問卷、第二份問卷...的過錄編碼簿。
資料檔	*.dat	ASCII 格式的數據資料，建議以漢書等純文字文書編輯軟體開啟，若使用 word、notepad 開啟時，超過頁寬的欄位會折到下一行。此類資料檔使用時請配合資料庫提供的欄位定義程式，然而使用者亦可利用過錄編碼簿，自行撰寫熟悉的統計程式來讀取資料。
	*.sav	儲存成 SPSS 的系統檔，必須有 SPSS 統計軟體才能開啟。
	*.sd2	儲存成 SAS 的系統檔，必須有 SAS 統計軟體才能讀取資料
	*.xls	以 MS EXCEL 的資料檔。
欄位定義程式	*.sas	以 SAS 統計語法撰寫的欄位定義程式。
	*.sps	以 SPSS 統計語法撰寫的欄位定義程式。
問卷檔	Ques.doc	該筆調查計畫使用的調查問卷，一般為 word 檔案，若為網路調查問卷或本室所蒐藏較早期的調查問卷，則副檔名就會是 *.html 或 *.txt。
次數分配結果	Freq.out	這是資料庫利用資料檔執行後得到的次數分配結果，為純文字格式，同樣的，若以 word、notepad 等開啟時，請將字體改成標楷體、10 號字，可得到較美觀且易閱讀的表格。
資料整理報告	Check.txt	在此文件中，我們說明了資料庫曾對資料做了哪些檢查，哪些資料是我們參考原始資料進行修改或製作的，以及提醒使用者注意的事項。
研究報告	Report.doc	計畫主持人提供的研究報告，若是有數個檔案時，會用一個資料夾來存放。

*為任意字或該研究計畫名稱的關鍵字。

五、閱讀資料

我們建議資料在解壓縮後，依下列的順序閱讀資料：

1. 資料整理報告：

瞭解資料庫曾經過資料做過哪些修改，以及資料檔中需請申請者特別注意的變項及數值。

2. 摘要與研究報告：

瞭解原計畫主持人的研究設計、研究限制、以及已有的研究成果。

3. 問卷：

瞭解該計畫調查問卷內的題目

4. 過錄編碼簿與欄位定義程式（或系統檔）：

瞭解變項名稱、資料屬性（數字或文字），準備進行分析。

六、使用者常見的問題

(Q & A):

Q1：我是大三學生，因為老師的作業指定要用調查資料來分析資料，我可以申請您們的資料嗎？

Ans：可以的，但是要麻煩您在使用同意書上請課程指導老師簽名。

Q2：我急著要交作業，可是遇到週末，我怕寄出的使用同意書和學生證可能要到下禮拜才會寄到，可以先給我資料嗎？

Ans：可以的，但是須請您先用傳真的方式，將填好的同意書和證件

影本傳真給我們，我們收到您的傳真後，會先將資料依您指定的方式傳送給您，但請您務必盡快將應繳交之文件寄給我們。

Q3：我在電腦上看到的檔名為什麼和你們在讀我檔案中列的不一樣？

Ans：

這可能是因為電腦設定的問題，我們的檔名是由主檔名和副檔名兩部份組成的，有些電腦會設定成隱藏已知的副檔名，像是*.doc、*.txt 等，因此像是問卷檔可能只顯示 ques，而不會顯示 ques.doc。

這對您讀取檔案不會造成影響，但是您會看到幾個相同檔名的檔案，但實際上是不一樣的檔案，例如：srda.sps（欄位定義程式）、srda.dat（資料檔），在隱藏副檔名的狀況下，您在電腦上看到的檔名都只有 srda，而導致您不知該執行哪一個檔案，或是當您需要修改欄位定義程式中的資料檔路徑時，讓您誤以為資料檔名只有 srda，而造成統計軟體無法正確讀取資料。

此時，請您從檔案總管，點選工具列的「檢視」，在「資料夾選項」中，取消「隱藏已知檔案類型的副檔名」並在「隱藏檔」項目下，選擇「顯示所有檔案」，您便可看到和我們在讀我檔案中一樣的檔案。

Q4：為什麼我執行欄位定義程式後，出現找不到資料檔的訊息？

Ans：

可能是因為您存放資料檔的路徑與欄位定義程式中的資料存放路徑不一致，在我們所提供的欄位定義程式中的資料路徑，是我們在檢誤資料時讀取資料的路徑，與您放置資料檔的路徑不一定會相同。以您解壓縮後是將資料檔是存放在磁碟片(A:)的plate資料夾下為例(參考圖三)，請在欄位定義程式中找到下列指令，依範例的方式修改路徑。

SPSS 程式

```
Data list file="c:\plate.dat" rec=1.
```

修改為

```
Data list file="a:\plate\plate.dat"  
rec=1.
```

SAS 程式

```
Data srda;
```

```
Infile "c:\srda.dat" lrecl=80 pad;
```

修改為

```
Data srda;  
Infile "a:\plate\plate.dat" lrecl=80  
pad;
```

Q5：為什麼我從 spss 開啟「學術調查研究資料庫」提供的欄位定義程式，看到的是一堆亂碼？

Ans：因為我們在欄位定義程式中有變項說明及選項數值說明，這部份用的是中文，所以您必須設定 SPSS 使用中文字型，才能正常顯示。修改的方式請您從 SPSS 工具列的「View」下，進入「Fonts」後，選擇中文字集與字型。

Q6：提供的欄位定義程式是 SPSS，可是我習慣使用的是 SAS，你們可以幫我嗎？

Ans：如果您無法自行處理時，我們可以幫您利用 SPSS 讀取後，將資料轉存成 SAS 的系統檔，但如此會失去選項數值 (value label) 的資訊，屆時再麻煩您參考過錄編碼簿進行分析。

Q7：我覺得現在要取得資料的手續好麻煩，每次申請資料都得提供那些證件嗎？

Ans：目前我們正在評估實施會員制的可行性，也就是只要成為資料庫的會員之後，便可直接在網路上查詢或下載資料，而不受時間和空間的限制。但在新制度實施之前，還是得請您先按照目前的釋出辦法相關規定提出申請，不便之處，請多包涵。