

美國『社會概況調查資料瀏覽平台 (GSS Data Explorer)』介紹

林雅琪

一、背景與目的

問卷調查為瞭解社會概況、人民意見的方法之一，而長期性的調查不但能作為觀察社會長期變遷趨勢的依據，對於研究者而言更是相對寶貴的研究資源。其中，在美國由芝加哥大學的獨立研究機構 NORC (National Opinion Research Center) 負責資料蒐集、NSF (National Science Foundation) 資助的長期調查計畫－「社會概況調查」 (General Social Survey, 以下簡稱為 GSS¹)，從 1972 年開始即透過問卷調查的方式，蒐集美國當代社會資料以瞭解相關態度、行為及屬性的變化趨勢及原因，並將調查資料公開，作為社會科學等領域研究者的免費研究分析資源。

和我們台灣的「社會變遷基本調查」(Taiwan Social Change Survey, 以下簡稱 TSCS) 相同的是，TSCS 和 GSS 兩者均加入最大的國際社會調查組織－ISSP² (International Social Survey Program)，該組織是由各國進行社會調查的單位組成，至今已有美國、法國、日本、台灣等 50 多個國家加入，每年各國會針對共同關注的議題建立題組、並進行調查訪問，例如：公民權、健康、環境、宗教、社會網絡等跨國調查題組³。

而隨著調查年度增加，累積的問卷題目、變項數、成功樣本數等研究資料量愈來愈可觀，至 2014 年止 GSS 資料庫有高達 5,629 個變項及問卷題目，使得使用者在搜尋資料時可能花費較多的時間。另外，當資料量日益增加，如沒有人將資料進一步分析，轉化成有價值的知識或研究著作，如此儘管有再大量的調查資料仍舊是一些數據。有鑑於此，NORC 持續開發、提升其功能，並於 2015 年 12 月 18 日起推出了社會概況調查資料瀏覽平台 (GSS Data Explorer)⁴，讓使用者容易搜尋及取得資料、出版品與相關研究資源、並能直接在平台上線上分析、下載資料、將資料共享給其他使用者等功能，讓使用者更容易使用 GSS 所蒐集的調查資料。本文將介紹 GSS Data Explorer 的功能及特性、利用圖解說明方式，示範各項功能的使用流程與操作結果，提供給可能需要利用 GSS 資料的使用者、或對 GSS Data Explorer 有興趣者作為參考。

1 有關 GSS 介紹，可參考 SRDA 學術調查研究資料庫通訊第 27 期「長期系列調查資料的加值應用—以 1972-2006 年美國社會概況調查資料 (GSS) 串連為例」一文，下載網址：<https://srda.sinica.edu.tw/publish/emagzinecont/37>。

2 ISSP，網址為 <http://www.issp.org/>。

3 ISSP 各題組調查結果，<http://zcat.gesis.org/webview/index.jsp>，請點選 ISSP 再點選 by Module Topic。

4 GSS Data Explorer，網址為 <https://gssdataexplorer.norc.org/>。

二、功能與特色

GSS Data Explore 主要有四大特色：變項搜尋與儲存、資料分析、資料擷取、專案工作區，而資料視覺化 (Visualize Data) 功能目前尚在建置中。各項特色概述如下：

- 變項搜尋與儲存 (Search and Save Variables)：使用者可以透過變項搜尋，尋找所需的調查題目、變項等相關調查資料，並能進而將選取的變項加入變項購物車 (Variable Cart)。
- 資料分析 (Analyze Data)：在簡易操作的介面上，對已儲存在變項購物車中的變項或自行新增選取的變項，進行交叉、多層次、相關、迴歸等資料分析。
- 資料擷取 (Extract Data)：提供不同輸出格式及資料檔 (SAS、SPSS、Stata、R)，讓使用者自行選擇欲下載的擷取資料格式。
- 專案工作區 (Project Workspaces)：在專案工作區中，使用者可以儲存歷次搜尋結果、儲存的變項及資料分析、擷取的紀錄。另外，經由專案設定，可以將專案中既存資料與其他擁有 MyGSS 帳號的使用者共享。

雖然任何使用者均能瀏覽變項和題目、儲存變項，然而如需分析資料、建立專案，甚至將專案分享給其他人，則必須要有 MyGSS 帳號，因此建議有上述需求的使用者，可以先申請 MyGSS 帳號，帳號申請流程為：至註冊頁面⁵上填寫資料，再至電子信箱收取確認信完成註冊步驟，隨後即可輸入帳號及密碼進行登入。以下為 GSS Data Explorer 各項功能的使用流程及操作結果：

(一) 變項搜尋與儲存 (Search and Save Variables)

透過「主題」、「調查年度」或「關鍵字」，逐步限縮查詢的範圍，進而搜尋合適的題目、變項以及出版品。其中，當使用者所輸入的關鍵字，和 GSS 資料庫的變項名稱或是調查題目相同時，即符合搜尋條件。

「假設我們想要瞭解過去十年「收入 (income)」跟「幸福 (happiness)」之間的關係，該透過哪些步驟來搜尋到需要的資料？」本文將以此假設來示範：首先，在關鍵字中輸入「income」，再篩選 2004-2014 調查年度，結果顯示有 51 個變項符合搜尋條件。使用者可以藉由瀏覽變項說明 (Description) 或調查題目 (Survey Questions)，從中尋找真正需要的變項。假設我們需要的是受訪者個人收入資料，瀏覽變項說明得知「rincome」這個變項符合需求，點選變項左邊購物車符號，即可把「rincome」儲存至變項購物車 (如圖一)。

5 MyGSS 帳號註冊頁面：https://gssdataexplorer.norc.org/d/users/sign_up

The screenshot shows the GSS Data Explorer interface. At the top, there are navigation links like 'Features', 'Media Room', 'Help and Resources', 'About', 'Contact', 'Register', and 'Sign In'. Below this is a search bar with the text 'MY GSS > Search Data'. The main search area is titled 'Search Data' and contains several filters: a 'Keyword' field with 'income' entered, a 'Filter by Years' slider set to '2004 - 2014', and a 'Filter by Subject' dropdown menu. A 'SEARCH' button is visible. Below the search filters, there are tabs for 'Variables' and 'Survey Questions'. The search results section shows '51 Results Matching Criteria'. A table lists variables with columns for 'Variable Name', 'Description', and 'Years Available'. The table includes rows for 'income', 'income98', and 'income06'. Annotations 1 through 6 are placed on the interface to highlight specific features: 1 points to the keyword input, 2 to the year filter, 3 to the variable tabs, 4 to the variable selection icon, 5 to the variable name, and 6 to the year availability grid.

圖一 變項搜尋與儲存功能⁶

- ① 輸入關鍵字
- ② 篩選調查年度
- ③ 符合搜尋條件筆數，依變項/問卷題目，瀏覽詳細資料
- ④ 儲存變項
- ⑤ 點擊變項瀏覽詳細資訊
- ⑥ 綠色方格表示篩選的調查年度中，哪些年有調查資料

以上範例是我們很明確的知道需要哪些變項，透過關鍵字搜尋即可快速找到所

6 本文圖片擷取自 GSS Data Explorer 官網：<https://gssdataexplorer.norc.org/>。

需的變項，然而，當不確定需要哪些變項時，則可以從主題來搜尋，勾選有興趣的主題。假設對於老年主題有興趣，則勾選「Aged」主題，再從中尋找需要的變項（如圖二）。瀏覽與儲存變項方式請參考圖一。

The screenshot shows the GSS Data Explorer interface. At the top, there are navigation links: Features, Media Room, Help and Resources, About, Contact, Register, and Sign In. Below this is a search bar with the text "MY GSS > Search Data" and a "VARIABLES" dropdown showing "2". The main section is titled "Search Data" and contains a search bar with the placeholder "Enter a Keyword", a "Filter by Years: 1972 - 2014" range selector, and a "Filter by Subject:" dropdown menu. The "Filter by Subject:" dropdown is circled in red and labeled with a "1". The dropdown menu is open, showing a list of subjects. The "Aged" subject is selected and highlighted in red, and a red arrow points from the "Aged" subject in the dropdown to a red box labeled "2" that highlights the "Search Subjects" modal window. The modal window shows a list of subjects with checkboxes next to them. The "Aged" subject is checked. Below the modal window, there is a table with 6 results matching the criteria. The table has columns for "Variable Name" and "Description".

| Variable Name | Description |
|---------------|--------------------------------------|
| natsoc | Social security |
| spretire | Govt spending on retirement benefits |
| aidold | Govts resp: provide for the elderly |
| eldhelp | Who should provide help for elderly |
| eldcost | Who should pay for help for elderly |
| aged | Should aged live with their children |

圖二 主題搜尋功能

- 1 點選「加號」
- 2 勾選主題類別

然而，如果從變項說明或調查題目尚無法確定「rincome」是否符合我們需求時，也可以點擊「rincome」這個變項，進一步瀏覽該變項的原始問卷題目和選項、不同調查年度的次數分配結果等詳細資料，再決定是否需要儲存此變項。此外在搜尋結果頁面，亦提供該變項歸類於哪一個主題、有哪些其他相關的變項等資訊（如圖三）。相同地，可以參考上述說明來搜尋與儲存「happiness」這個變項。

GSS Data Explorer BETA Features Media Room Help and Resources About Contact Register Sign In

MY GSS > Search Data > Variable Detail

← Back To Variable Results

Respondents income ①

Variable: **rincome**
Subject: **Income** ②

ADDED TO CART ⑥

Survey Questions ③

Questions associated with this variable:
Did you earn any income from (OCCUPATION DESCRIBED IN OCC-INDUSTRY) in [the previous year]?
Yes (ASK RINCOME)
No [See REMARKS]
A. IF YES: In which of these groups did your earnings from (OCCUPATION IN OCC) for last year-[the previous year]-fall? That is, before taxes or other deductions. Just tell me the letter.

Year Availability

Available in 28 of 30 years total.

72 73 74 75 76 77 78 80 82 83 84 85 86 87 88 89 90 91 93 94 96 98 00 02 04 06 08 10 12 14

Summary by Year Showing All Years

| Code | Label | 1974 | 1975 | 1976 | 1977 | 1978 | 1980 | 1982 | 1983 | 1984 | 1985 | 1986 | 1987 | 1988 | 1 |
|------|-----------------|------|------|------|------|------|------|------|------|------|------|------|------|------|---|
| 1 | Lt \$1000 | 69 | 92 | 56 | 60 | 70 | 45 | 60 | 49 | 44 | 42 | 41 | 52 | 34 | |
| 2 | \$1000 to 2999 | 5 | 109 | 104 | 86 | 113 | 76 | 78 | 68 | 67 | 56 | 50 | 72 | 53 | |
| 3 | \$3000 to 3999 | 49 | 70 | 59 | 67 | 48 | 37 | 61 | 39 | 33 | 47 | 39 | 40 | 44 | |
| 4 | \$4000 to 4999 | 67 | 65 | 60 | 45 | 53 | 29 | 49 | 31 | 39 | 24 | 34 | 43 | 21 | |
| 5 | \$5000 to 5999 | 64 | 50 | 45 | 64 | 52 | 45 | 46 | 38 | 39 | 32 | 28 | 37 | 26 | |
| 6 | \$6000 to 6999 | 48 | 47 | 55 | 59 | 47 | 31 | 49 | 36 | 29 | 46 | 26 | 35 | 28 | |
| 7 | \$7000 to 7999 | 57 | 59 | 44 | 57 | 44 | 36 | 42 | 39 | 41 | 34 | 32 | 48 | 35 | |
| 8 | \$8000 to 9999 | 89 | 101 | 78 | 103 | 80 | 88 | 78 | 75 | 52 | 42 | 46 | 64 | 40 | |
| 9 | \$10000 - 14999 | 155 | 175 | 166 | 215 | 195 | 179 | 235 | 191 | 176 | 164 | 155 | 194 | 147 | |
| 10 | \$15000 - 19999 | 60 | 59 | 85 | 110 | 123 | 136 | 152 | 132 | 115 | 123 | 124 | 144 | 117 | |
| 11 | \$20000 - 24999 | 30 | 30 | 34 | 40 | 60 | 80 | 132 | 115 | 122 | 113 | 105 | 145 | 121 | |
| 12 | \$25000 or more | 35 | 23 | 42 | 49 | 57 | 100 | 143 | 181 | 176 | 235 | 226 | 317 | 279 | |
| 13 | Refused | 37 | 21 | 23 | 20 | 24 | 23 | 64 | 46 | 23 | 38 | 26 | 29 | 44 | |
| 98 | Don't know | 15 | 6 | 10 | 8 | 12 | 8 | 8 | 12 | 15 | 8 | 6 | 11 | 5 | |
| 99 | No answer | 0 | 16 | 13 | 54 | 8 | 59 | 22 | 15 | 9 | 0 | 0 | 12 | 14 | |
| 0 | Not applicable | 593 | 567 | 625 | 493 | 551 | 496 | 651 | 532 | 493 | 530 | 532 | 576 | 473 | |

Summary statistics

| | |
|---------------|-------|
| Valid cases | 34892 |
| Missing cases | 24707 |

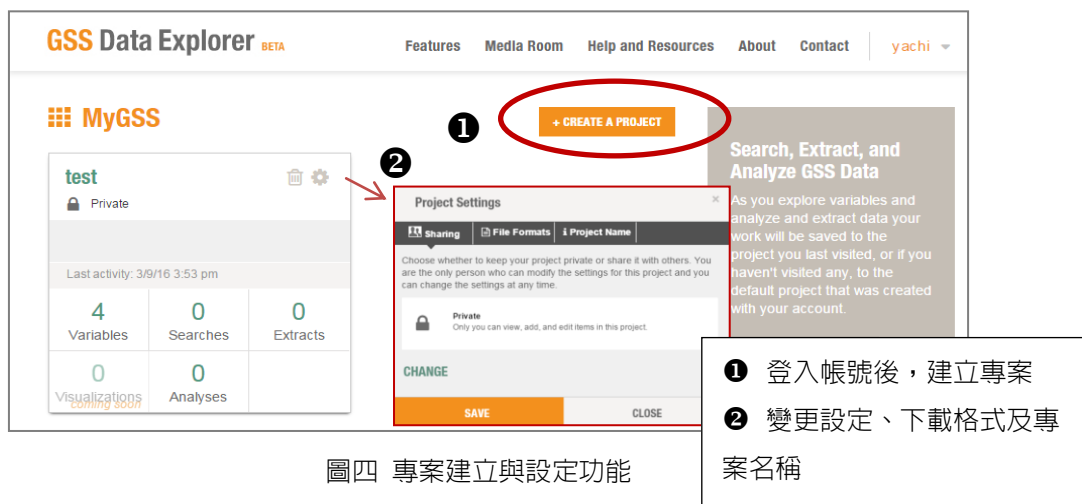
① 變項說明
② 變項名稱、所屬主題分類
③ 原始問卷題目及選項
④ 其他相關變項
⑤ 各調查年度次數分配
⑥ 儲存變項

圖三 瀏覽詳細資料頁面

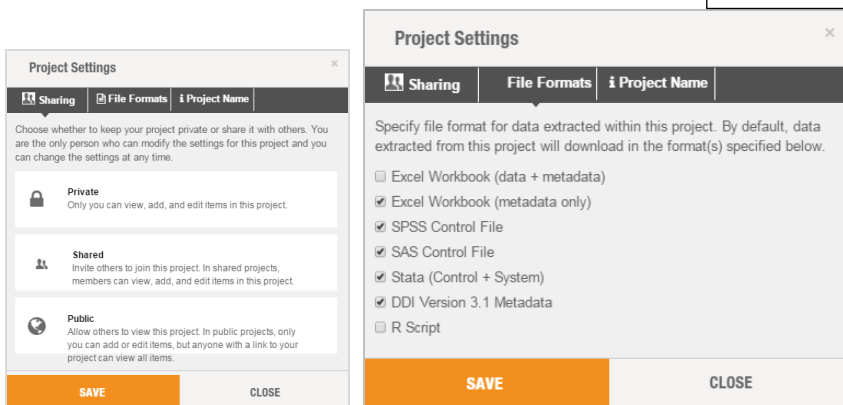
(二) 專案工作區 (Project Workspaces)

使用專案工作區功能，需要先登入 MyGSS 帳號，專案工作區不但可以幫助使用者儲存搜尋及資料分析等歷程紀錄，並能經由專案設定來選擇不公開 (Private)、共享 (Shared) 或是公開 (Public) 的狀態。

共享是藉由 E-mail 方式邀請其他 MyGSS 帳號的使用者加入，共享的使用者可以瀏覽、新增以及編輯專案內容，此功能對於需要共同研究的研究者們相對便利，亦能促進使用 GSS 資料作為課堂教學；和共享不同的是，只要知道公開專案連結者均可以瀏覽公開的專案，但對於專案內容不能編輯或新增。然而，不論是共享或公開的專案，都只有最初建立專案的使用者才有權限去變更專案設定的狀態。此外，如對於下載資料的格式有特別偏好，也可以在此先進行設定，上述的設定功能在專案建立後亦能隨時變更。以下是建立新專案、專案設定的流程 (如圖四、圖五)。

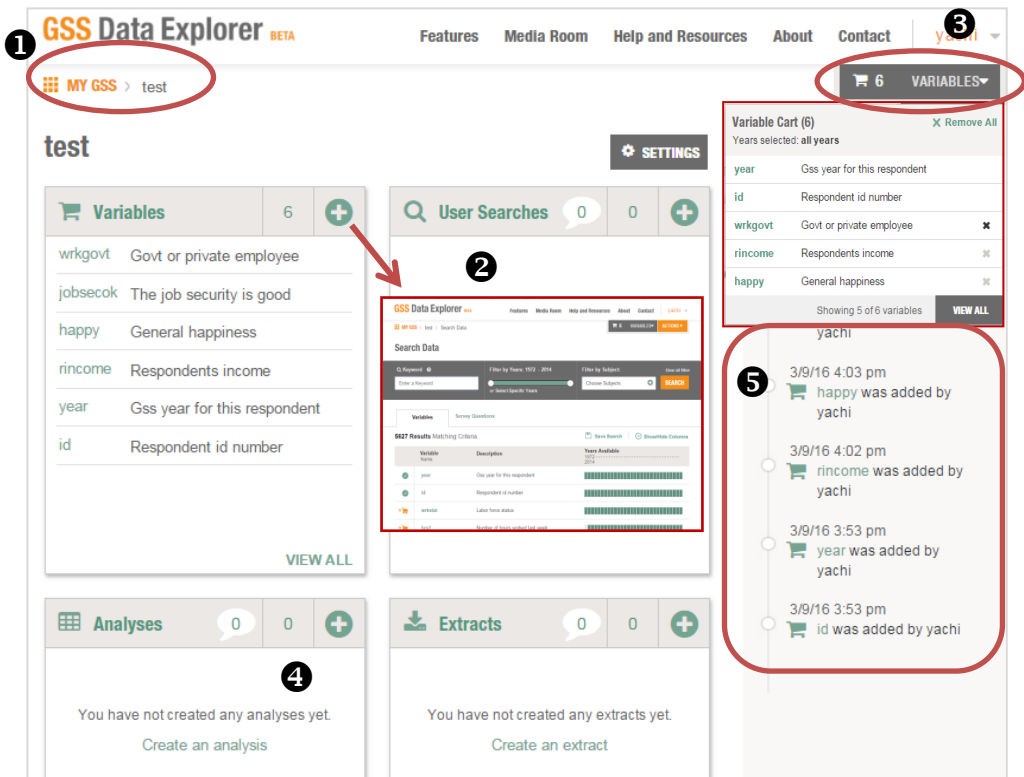


圖四 專案建立與設定功能



圖五 專案建立後設定變更功能

專案建立完成後，即可新增搜尋、儲存變項等資料流程，如需移除已儲存的變項，點選右上角變項購物車則可以移除，而從右列可以瀏覽專案工作的時間紀錄。此外，因使用者可以建立兩個以上的專案，因此在進行儲存變項等工作前，需事先選定特定專案（如圖六）。



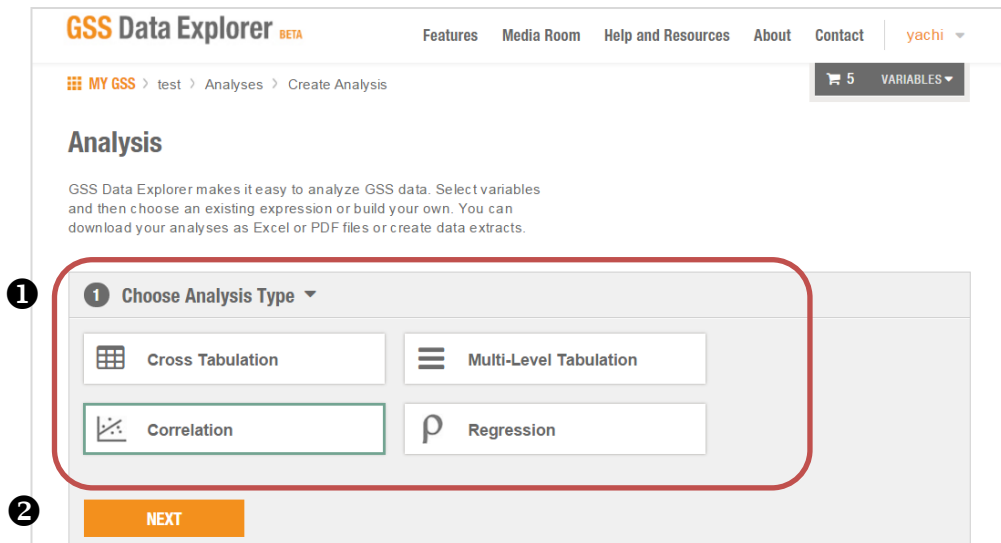
圖六 專案建立與設定功能

- ❶ 選擇在此專案中進行工作
- ❷ 點選加號進行變項搜尋與儲存
- ❸ 點選後彈出購物車清單，移除不需要的變項
- ❹ 點選新增資料分析
- ❺ 專案工作時間紀錄

(三) 資料分析 (Analyze Data)

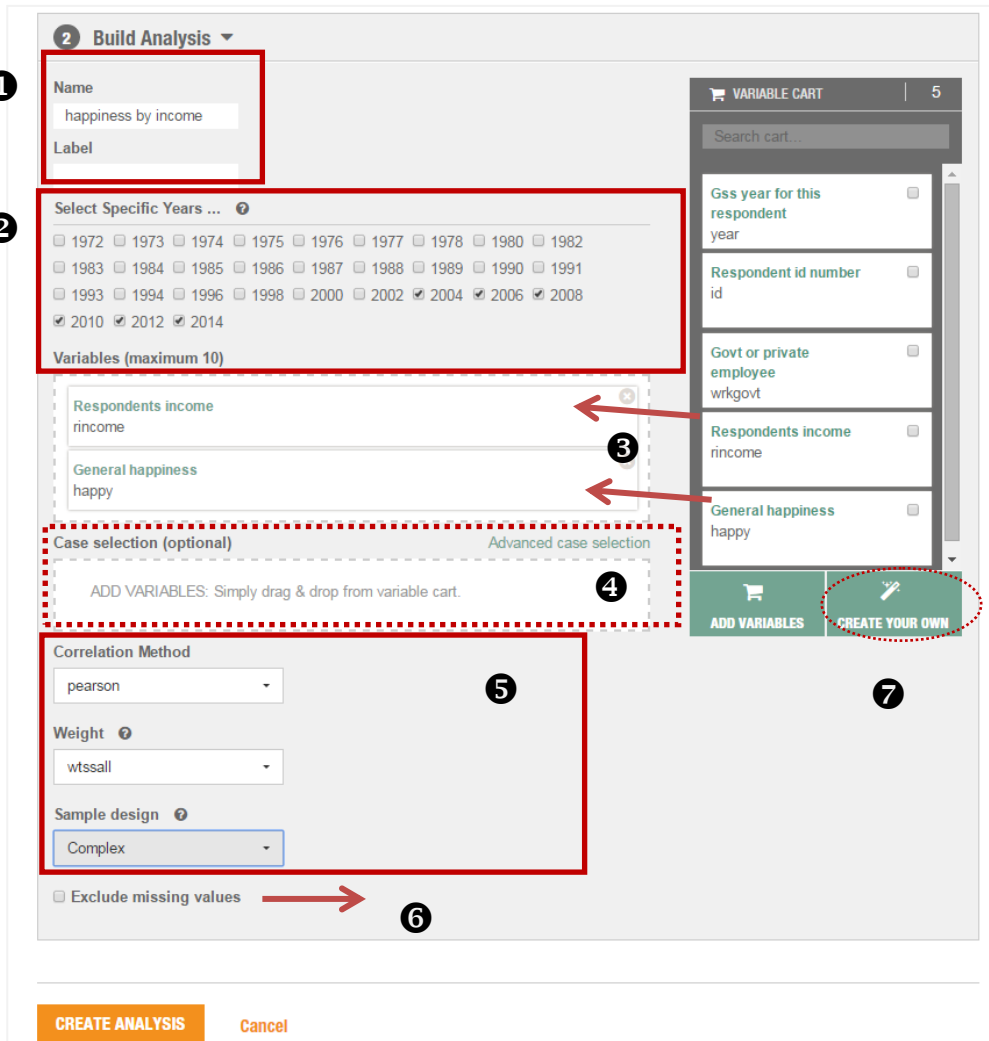
使用者可以從變項購物車中已儲存的變項或自行新增變項 (在線上對變項重新編碼)，直接在簡易操作的介面上進行交叉、多層次、相關、迴歸之資料分析。另外，系統會提供預設的相關分析方法、設定權值變項及抽樣設計的類型，然而使用者也可以依研究需求自行選擇合適的選項。

針對上述已經儲存到專案工作區的「rincome」跟「happiness」這兩個變項，如何進行相關分析？首先，從圖六專案工作區的分析區塊新增分析，選擇欲分析之類型後按下一步，再針對分析的內容，例如：建立分析資料之名稱、選擇調查年度、篩選特定的選項、是否需排除遺漏值等設定，待設定完成後點選建立分析，即開始進行線上相關分析 (如圖七、八)。



圖七 資料分析設定功能 (一)


- 1 選取資料分析類型後
- 2 點選下一步開始進行設定

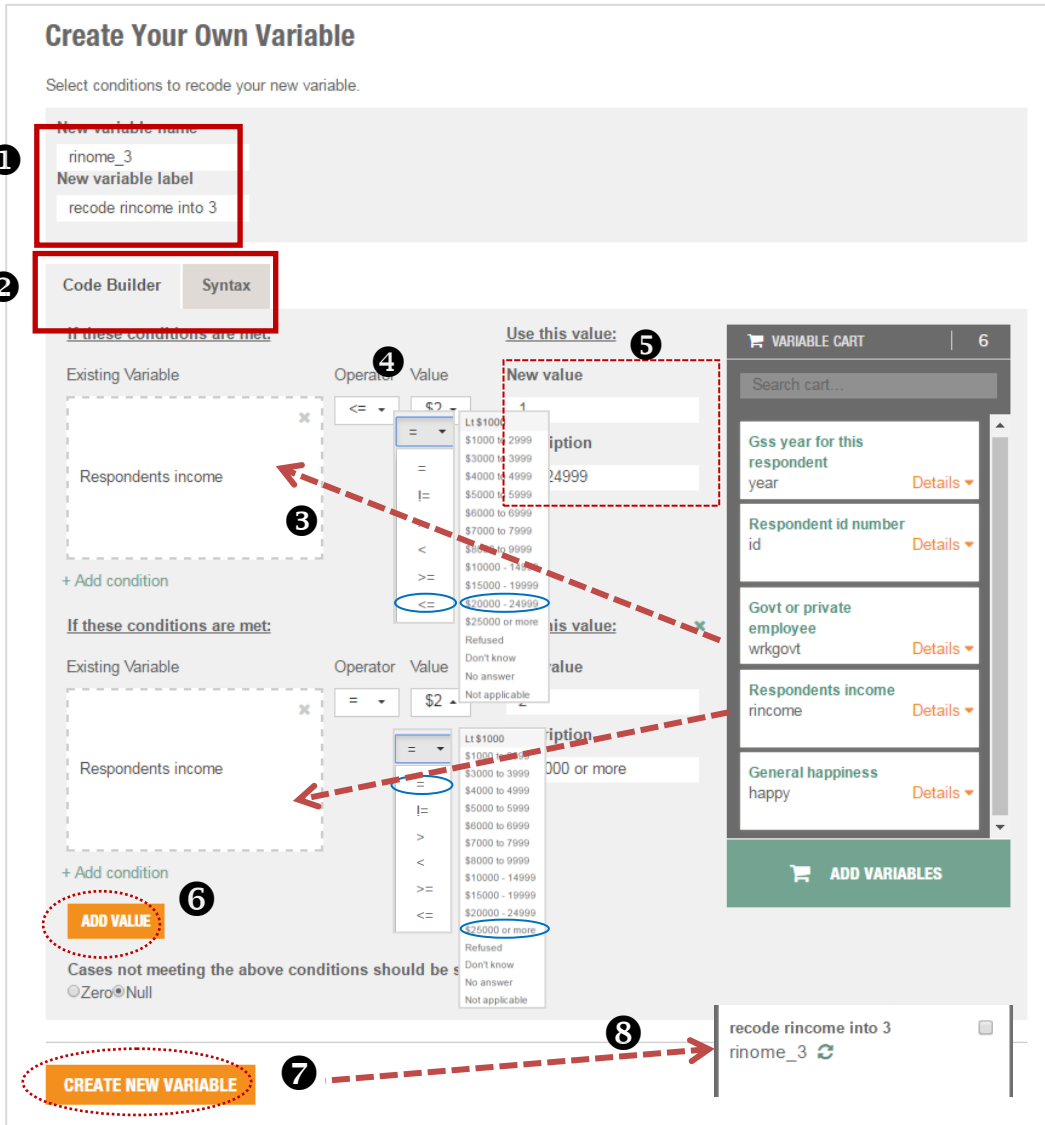


圖八 資料分析設定功能（二）

- ① 為要分析的資料命名及說明
- ② 勾選要分析的調查年度
- ③ 將要分析的變項，從右方拖曳至此
- ④ 勾選需要分析變項的哪幾個選項
- ⑤ 依研究需求自行選擇：相關分析的方法、設定權值變項、抽樣設計
- ⑥ 勾選是否需排除遺漏值
- ⑦ 自行建立新變項（重新編碼）如圖九

從上述設定中，如仍無法滿足研究所需的選項分類，則可以從「CREATE YOUR

OWN」功能，透過點選或輸入語法的方式對變項進行重新編碼。以下假設我們需要將「rincome」從原本 12 個分類 (Lt \$1000、\$1000 to 2999、\$3000 to 3999、...、\$25000 or more)，重新編碼成 2 個分類 (小於等於\$24999、大於\$25000)，該如何用點選方式來設定。而當成功新增變項後，在變項購物車中即會出現有「」符號的變項，表示該變項是被重新編碼過的。(如圖九)。



Create Your Own Variable

Select conditions to recode your new variable.

1 New variable name
rinome_3
New variable label
recode rinome into 3

2 Code Builder Syntax

3 Existing Variable
Respondents income

4 Operator
=

5 Value
\$24999

6 New value
\$25000 or more

7 Cases not meeting the above conditions should be
Zero/Null

8 CREATE NEW VARIABLE

6 VARIABLE CART
Search cart...
Gss year for this respondent year
Respondent id number id
Govt or private employee wrkgovt
Respondents income rinome
General happiness happy
ADD VARIABLES

圖九 新增變項功能

- ❶ 輸入新增的變項名稱與變項說明
- ❷ 選擇用點選或語法方式做設定
- ❸ 將要重新編碼的變項，從右方拖曳至此
- ❹ 點選重新編碼條件
- ❺ 為重新編碼的選項輸入數值說明
- ❻ 點選後，繼續重新編碼
- ❼ 以上設定完成請點選新增變項
- ❽ 完成後，在變項購物車即會出現被重新編碼的新增變項「rincome_3」

分析完成，除了提供分析結果外，並可以檢視分析前所做的設定，而分析結果提供表格、圖表等不同格式供使用者選擇下載。以下為相關分析結果頁面、可以選擇要匯出 PDF 檔或是 R 程式碼。如分析完成後，需變更設定或調整分析的變項，只要點選右上方，即可重新編修設定、並重新分析，例如「happiness」在之前設定中沒有排除遺漏值，即可透過重新編修設定、再次取得相關分析結果（如圖十）。另外，亦可以針對每一筆分析資料作評論、註解的功能（如圖十一）。

MY GSS > test > Analyses > happiness by income

Analysis Summary

happiness by income

Selected years: 2004, 2006, 2008, 2010, 2012, 2014
 Variables: Respondents income, happy_recode
 Case selection: (Combined.year IN (2004, 2006, 2008, 2010, 2012, 2014)) AND (Combined.rincome IN (1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12))
 Correlation method: pearson
 Exclude missing values: Yes
 Weight var: WTSSALL
 Sample design: Simple Random Sample

EXPORT CREATE VISUALIZATION

PDF R Code

```

$correlation
      rincome happy_recode
rincome  1.00000000  0.03688923
happy_recode 0.03688923  1.00000000

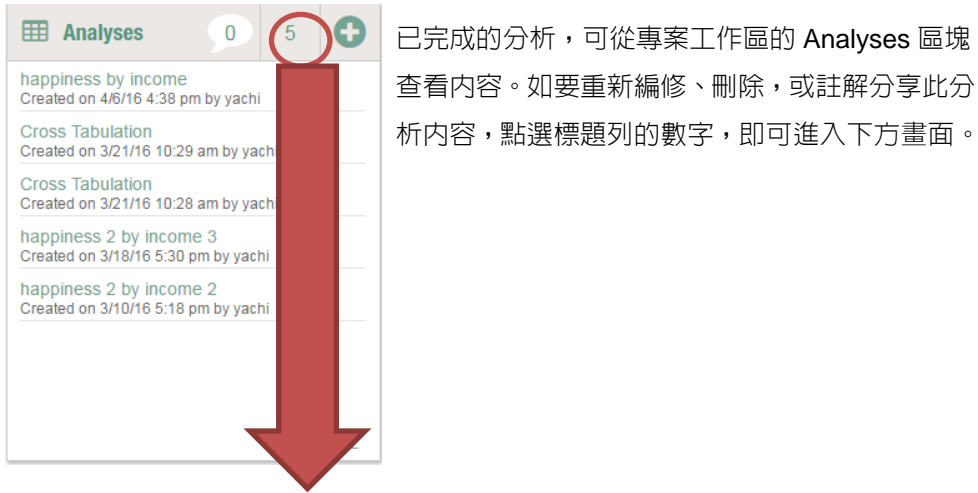
$std.err
      rincome happy_recode
rincome  1.612755e-16 1.029898e-02
happy_recode 1.029898e-02 1.334981e-16

$.value
      rincome happy_recode
rincome  6.200571e+15 3.581834e+00
happy_recode 3.581834e+00 7.490745e+15

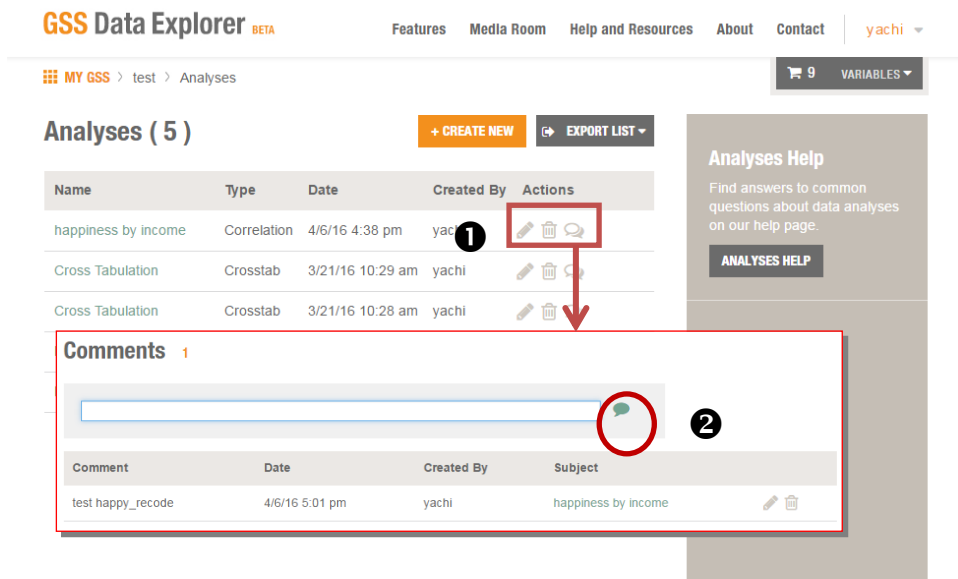
$.p.value
      rincome happy_recode
rincome  0.0000000000 0.0003429121
happy_recode 0.0003429121 0.0000000000
  
```

❶ 分析前的設定
 ❷ 相關分析結果
 ❸ 選擇要匯出的格式
 ❹ 重新編修設定

圖十 資料分析結果與匯出功能



已完成的分析，可從專案工作區的 Analyses 區塊查看內容。如要重新編修、刪除，或註解分享此分析內容，點選標題列的數字，即可進入下方畫面。



圖十一 資料分析評論

- 1 由左至右分別為：編修、刪除、註解
- 2 註解填寫完畢按此儲存

(四) 資料擷取 (Extract Data)

GSS 資料庫有超過 5000 個變項、40 年以上的調查資料。雖然網頁上仍提供快速下載⁷ (Quick Downloads) 全部資料檔的功能，但因下載全部的資料需要耗費很多時間，建議使用者事先利用擷取工具(Extract tool)，篩選出自己需要的變項後再進行資料下載。資料擷取同樣可以從專案工作區中進入來建立，依照設定步驟進行 (如圖十二)。

圖十二 資料擷取與設定功能

① 為擷取資料命名

7 快速下載全部資料檔之連結：

https://gssdataexplorer.norc.org/pages/show?page=gss%2Fgss_data。

- ② 從已加入變項購物車的變項選取或回到搜尋頁面搜尋，選擇要下載的變項
- ③ 篩選特定選項
- ④ 篩選需要的年份、欲下載的格式
- ⑤ 按此確認下載，下載完成後會顯示可下載的箭頭圖示
- ⑥ 下載壓縮檔

然而，由於從資料擷取流程下載的資料，並不是系統檔（如 SPSS 的 SAV 系統檔、STATA 的 DTA 系統檔），因此使用資料前需自行利用下載的欄位定義程式來讀取原始資料。例如，需要 SPSS 的 SAV 資料檔，則需利用下載的 GSS.sps 修改檔案存取的路徑後再讀取 GSS.dat。

(五) 其他功能

除上述特性與功能之外，在 GSS Data Explorer 上亦可透過「主題」、「調查年度」或「作者」，逐步限縮查尋的範圍，進而搜尋合適的出版品⁸，包含：過錄編碼簿、問卷、報告書、衍生著作等不同類型的相關資源。使用者可以選擇在線上瀏覽，或是直接下載整份文件（如圖十三）。

The screenshot shows the GSS Reports search interface. At the top, there are search filters: a 'Keyword' field containing 'happiness' (circled 1), a 'Years, by Range' slider from 2004 to 2014 (circled 2), and an 'Authors' field containing 'Showing All Authors' (circled 3). Below these is a 'VARIABLES' dropdown showing '2'. A horizontal menu (circled 4) includes 'All', 'Codebook', 'Questionnaires', 'Cross National Reports', 'Methodological Reports', 'Project Reports', 'Social Change Reports', 'Topical Reports', and 'Trend Reports'. Below the menu, a box (circled 5) indicates '24 Reports Matching Criteria'. A table lists search results with columns for Name, Description, Authors, and Applicable Years. The 'GSS Codebook Main Body' entry (circled 6) is highlighted. A 'Show/Hide Columns' dropdown menu is open on the right, showing checked options for Name, Description, Authors, and Years Studied.

| Name | Description | Authors | Applicable Years |
|-------------------------|--|---------|------------------|
| GSS Codebook Appendix A | Sampling Design and Weighting | | ■■■■■ |
| GSS Codebook Appendix N | Changes in Question Wording, Response Categories, and Format | | ■■■■■ |
| GSS Codebook Appendix T | General Social Survey Papers | | ■■■■■ |
| GSS Codebook Appendix V | Subject Index to Questions | | ■■■■■ |
| GSS Codebook Index | Index | | ■■■■■ |
| GSS Codebook Main Body | Main Body | | ■■■■■ |

圖十三 出版品及相關資源搜尋功能

8 相關出版品搜尋之連結：<https://gssdataexplorer.norc.org/documents>。

- ❶ 輸入關鍵字
- ❷ 篩選年度
- ❸ 輸入作者
- ❹ 選擇出版品類型
- ❺ 符合搜尋條件筆數
- ❻ 點擊瀏覽文件全文

三、結語

GSS Data Explorer 設計了簡易操作的介面、並整合了資料搜尋、分析、專案工作區等功能，協助使用者在大量的資料中更有效率的搜尋、取得所需的變項與相關資源。而持續開發中的視覺化功能亦值得期待，我們將持續關注 GSS Data Explorer 的現況，如有最新的發展，將為大家作介紹。