

Survey Research Data Archive Newsletter

本期目錄

■ 資料釋出消息.....	1
■ 資料工程初探—以開放資料清理轉換為例.....	32
■ 會員申請程序說明.....	50
■ 徵稿啓事	51

中央研究院 人文社會科學研究中心 調查研究專題中心

中華民國105年9月



國內
郵資已付

台北郵局許可證
台北字第2111 號

台灣郵政台北誌第395號執照登記為雜誌交寄



資料釋出消息

邱亦秀

本資料庫於 105 年 6 月至 8 月間改版更新及資料開放訊息如下：

【資料開放：長期主題型調查】

■ 華人家庭動態資料庫的建立：第十四年計畫(CVII2012)

計畫主持人：簡錦漢

計畫執行單位：中央研究院人文社會科學研究中心華人家庭研究計畫

經費補助單位：中央研究院人文社會科學研究中心

調查執行期間：2012/01/09-2012/04/30

華人家庭動態資料庫(Panel Study of Family Dynamics, 簡稱 PSFD)是以華人家庭成年樣本為對象的固定樣本追蹤調查，涵蓋的年齡層自青年至老年，並根據主樣本延伸出來的子女等親屬，再進行追蹤訪問。

CVII2012 調查對象為對象為主樣本群滿 16 歲子女(1987-1995 年出生)。實地的面訪調查於 2012 年 1 月 9 日至 4 月 30 日間執行。完訪的樣本數為 751 案。問卷內容包含以下主題：個人資料、教育經驗、工作經驗、婚姻與配偶資料、居住安排及家庭決策與支出等。

釋出項目計有：資料使用說明、問卷檔、ASCII 資料檔、SPSS 資料檔、STATA 資料檔、SAS 欄位定義程式、過錄編碼簿及附錄。

■ **華人家庭動態資料庫的建立：第十五年計畫(RR2014、CVIII2014、RCI2014)**

計畫主持人：章英華

計畫執行單位：中央研究院人文社會科學研究中心華人家庭研究計畫

經費補助單位：中央研究院人文社會科學研究中心

計畫執行期間：2013/08/01~2014/07/31

華人家庭動態資料庫(Panel Study of Family Dynamics, 簡稱 PSFD)是以華人家庭成年樣本為對象的固定樣本追蹤調查，涵蓋的年齡層自青年至老年，並根據主樣本延伸出來的子女等親屬，再進行追蹤訪問。

RR2014 調查對象為主樣本(含 1953-64、1935-54、1964-76、1977-1983 年出生的主樣本群)及年滿 25 歲的主樣本子女，完訪樣本數為 4,672 案。問卷內容包含：生活與工作狀況、居住安排、家庭決策與收支、子女生育與教養。

CVIII2014 調查對象為對象為主樣本群滿 16 歲子女(1990-1997 年出生)。實地的面訪調查於 2014 年 1 月 19 日至 6 月 30 日間執行。完訪的樣本數為 650 案。問卷內容包含以下主題：個人資料、教育經驗、工作經驗、婚姻與配偶資料、居住安排及家庭決策與支出等。

RCI2014 調查對象為主樣本之 1987-1989 年出生子女，完訪樣本數為 342 案。問卷內容包含：個人資料、教育經驗、工作經驗、婚姻與配偶資料、家庭價值與態度、親屬資料、居住安排、家庭決策與支出、子女教養、家庭關係與和諧。

釋出項目計有：資料使用說明、問卷檔、ASCII 資料檔、SPSS 資料檔、STATA 資料檔、SAS 欄位定義程式、過錄編碼簿及附錄。

2013 年第一次及第二次社會意向調查計畫

計畫主持人：楊文山

計畫執行單位：中央研究院社會學研究所

經費補助單位：中央研究院社會學研究所

調查執行期間：第一次 2013/07/09~2013/07/31

第二次 2013/11/19~2013/12/06

計畫執行期間：2013/01/01~2013/12/31

台灣社會意向調查主要研究目的在掌握台灣社會快速變化的脈動，以類似「社會總體檢」的精神，建構固定、有效的指標，定期衡量社會面。

第一次調查針對 2013 年前半年的新興社會議題與台灣重要的社會變遷，新增了四項主題：(1)世代、(2)核電與公投、(3)H7N9、(4)教育。

第二次調查，新增了四項主題：(1)民主品質(2)網路隱私 (3)中國對台灣影響 (4)同性戀與家庭。

調查母體為台灣、澎湖地區年滿十八歲以上、且家中有電話之民衆。抽樣方法採三階段機率抽樣，即先以分層二階段 PPS 抽樣法抽出電話號碼，其後，於中選電話號碼內再進行戶中抽樣。第一次調查抽樣電話數為 13,200 案，最後完成樣本數為 1,215 案；第二次調查抽樣電話數為 12,000 案，最後完成樣本數為 1,235 案。

釋出項目計有：資料使用說明、問卷檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及調查報告書。

【資料開放：政府調查資料】

■ 勞動部勞動力發展署「企業用人與就業市場趨勢專題調查」

勞動部勞動力發展署透過「台灣就業通」網站，辦理「企業用人與就業市場趨勢專題調查」系列，包括以網站求才會員隨機抽樣進行的「2015 新鮮人第一次工作就上手大調查」，和以國內 2 千大企業之企業會員為調查對象進行的「2015 企業人才進用與展望大調查」。

前者主要分析各產業企業聘用職場新鮮人的現況，以了解新鮮人在職場上有哪些適應不良或需要改進的地方。調查期間為 104 年 8 月 11 日至 9 月 11 日，調查方式採電話問卷方式，共計回收有效樣本數 1,101 份。後者主要瞭解企業對於人才進用的現況，以及企業對於未來展望為何。調查期間為 104 年 11 月 1 日至 12 月 2 日，採電話問卷方式進行調查，有效樣本數為 1,068 份。

釋出項目計有：資料使用說明、問卷檔、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿和報告書。

※本系列調查資料僅限「院內會員使用」。

■ 勞動部統計處「104 年職業災害失能勞工就業關懷調查」

勞動部統計處為了解職業災害失能勞工罹災前後之工作與就業狀況，及其需要政府協助項目，提供勞動部輔導職業災害失能勞工重返工作職場及規劃職業安全衛生政策之參考。

調查對象為臺灣地區 101 年 7 月至 103 年 6 月間曾經請領職業災害失能一次給付及年金給付之本國籍勞工。抽樣方法採分層比例隨機抽樣，依調查對象受傷前的勞工保險投保單位之行業區分，以勞工保險投保單位之所在地區及調查對象之年齡為分層變數。有效樣本依調查對象受傷前的勞工保險投保單位之行業、所在地區、調查對象之年齡及失能等級結構，作卡方適合度檢定，以確定其代表性。

調查項目包括受傷前後之工作狀況、復健情形、需要政府協助項目及工作安全衛生認知等。調查時期自 104 年 6 月 15 日起至 104 年 7 月 14 日，採電話訪問調查，共計回收有效樣本 3,013 份。

釋出項目計有：資料使用說明、問卷檔、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿、調查報告及摘要分析等。

■ 勞動部統計處「103 年、104 年育嬰留職停薪期滿復職關懷調查」

勞動部為了解申請育嬰留職停薪者期滿後回復原職及對復職協助措施之需求與建議，以促進性別工作平等相關政策之參考，辦理此項調查。本調查以台灣地區已請領育嬰留職停薪津貼者為抽樣母體，採郵寄問卷調查的方式，於 103 年完成有效樣本 5,525 份，104 年完成有效樣本 5,549。調查項目包括：

- 一、育嬰留職停薪狀況及重返職場情形
- 二、申請育嬰留職停薪前的就業狀況
- 三、育嬰留職停薪相關權益認知
- 四、基本資料

釋出項目計有：資料使用說明、問卷檔、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及報告書等。

■ 教育部「2011 年成人教育調查」

本調查旨在瞭解臺閩地區成人教育參與現況，包含參與正規學習、非正規學習與非正式學習活動之情形。調查樣本為全國年滿 18 歲以上的民衆，不含正規教育學制就讀之學生，但回流教育學生(如在職專班)則亦為調查對象。調查目的包括：(一)瞭解臺閩地區成人教育參與現況；(二)分析成人不同社會人口背景變項參與成人教育的分佈情形；(三)探究臺閩地區成人教育機構的經營現況；(四)提供研擬臺閩地區成人教育政策之參據。為達成上述各項目的，本調查分從個人層面與機構層面進行調查。個人層面部份以電腦輔助電話調查系統進行電話訪問，依臺閩地區 22 縣市 18 歲以上的人口各抽取所需樣本戶進行訪問，有效調查樣本為 11,000 人；機構層面部份則採郵寄問卷法，機構類型涵蓋家庭教育中心、大專院校、空中大學、高中職進修部、國中小補校、教育部立案之文教基金會、社區大學、長青學苑、樂齡學習中心、外籍配偶家庭服務中心、生活美學館、政府所屬教育訓練機構、博物館、美術館、科教館、公共圖書館、救國團等，母數為 2,101 所，共計回收 1,105 份，其中辦理成人教育者有 1,049 份。

釋出項目計有：資料使用說明、問卷檔、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及報告書。

■ 行政院主計總處「104 年事業人力僱用狀況調查」

主計總處為配合社會經濟發展情勢及人力決策之需要，定期蒐集具有前瞻性與政策性資料，並充實按月受雇員工薪資調查之不足，自民國 80 年 10 月創辦事業人力僱用狀況調查。本調查辦理初期(民國 80 年 10 月至 84 年 4 月)為每半年辦理一次，採工業及服務業二部門輪流辦理，自 85 年以後改為按年辦理(88 年因經費問題停辦)，且調查範圍擴大至工業及服務業全體。

本調查之基本問項大致可分為二部分：(一)各職類短缺員工概況及僱用條件(二)人員過剩及解決措施，同時並歷次辦理時，配合當時社經狀況及政策需要進行機動問項調整設計，歷次規劃項目大致包括外籍員工僱用情形、未來人員需求計畫、中高齡勞工僱用意願、部分時間僱用情形等。

調查對象為台灣地區工業及服務業之公民營企業場所單位，按各細行業採「截略分層隨機抽樣法」，其中公營事業、加工出口區與科學園區內之廠商為全查樣本，非全查層樣本採不等機率抽樣法，為減輕廠家每年連續填表之負擔，除全查樣本外，採每年輪換之原則。採郵寄通信調查為主，網際網路填報為輔並以電話聯繫催詢追蹤，共完成 9,948 案。

釋出項目計有：資料使用說明、問卷檔、ASCII 資料檔、SPSS 資料檔、STATA 資料檔、SAS 欄位定義程式、過錄編碼簿及報告書。

■ 行政院主計總處「103 年受僱員工動向調查」

行政院主計總處為配合社會經濟發展需要，蒐集受僱員工行業別、規模別、僱用型態別進退情形、受僱員工概況、非薪資報酬結構、調薪狀況、計薪方式及缺工情形等統計資料，並結合運用公、勞保及勞退新制檔、人口及住宅普查檔、戶籍檔之員工受僱及異動情形、教育程度別等資料，以供政府研訂經建計畫、規劃企業經營、釐訂產業、人力及輔導就業政策之參據，並供企業界調整員工薪資、改善經營管理、訂定人力計畫及提撥退休金之參考。

調查對象為 17 大行業之公民營事業單位(場所)及其進入(年底在職)與退出之受僱員工為對象，但不包括國防部所屬各軍事機關之工廠、免辦工商登記之機關團體福利社或消費合作社、學校之實習場所、救濟機關之習藝所、監獄人犯之作業場所等之受僱員工。採用最新工商及服務業普查母體資料檔，並依以下過程取得調查樣本：

- 1、應用最近一年「職類別薪資調查」之更正後樣本名冊。
- 2、按各細行業採「截略分層隨機抽樣法」，截略點以下之廠商採「分層隨機抽樣法」，共抽出約一萬家樣本。
- 3、採用哈奇斯達蘭尼斯之最適分層方法決定各層之分界限。

主要調查項目包含：1.一般概況(事業單位及填表人資訊)、2.年內全年受僱員工進退狀況、3.年內全年退休之受僱員工個人概況、4.年底在職受僱員工概況、5.年底員工之計薪方式、6.年內調整經常性薪資狀況及參考因素、7.全年非薪資報酬支付狀況、8.調查年 2 月底受僱員工人數、9.調查年 2 月底有無短缺員工。調查方法按各行業不同，分別採派員實地訪問調查、郵寄通信調查或兩者並行，另提供網路填報方式輔助調查辦理，103 年共完成 9,825 份。

釋出項目計有：資料使用說明、問卷檔、SPSS 資料檔、STATA 資料檔、ASCII 資料檔、SAS 欄位定義程式、過錄編碼簿及報告書。

■ 人力運用擬-追蹤調查資料庫 103 年-104 年

「人力運用擬追蹤調查：103 年 -104 年」利用主計總處進行調查時樣本輪換的特性，串連 103 年與 104 年人力運用調查重複的樣本，使相關分析元素具有兩個時點的分析資料，提供學術界有意義的分析資源。

釋出項目計有：資料使用說明、SPSS 資料檔、STATA 資料檔、SAS 資料檔、過錄編碼簿。

■ 交通部觀光局「97 年至 102 年國家風景區遊客調查」

交通部觀光局為瞭解國家風景區遊客旅遊決策、遊憩行為、旅遊活動偏好與滿意度及消費行為，以供相關機關規劃與改善觀光設施、提昇旅遊服務品質及訂定觀光發展策略之參考，特進行國家風景區遊客調查。

台灣 13 個國家風景區，包括「東北角暨宜蘭海岸」、「東部海岸」、「澎湖」、「大鵬灣」、「花東縱谷」、「馬祖」、「日月潭」、「參山」、「阿里山」、「茂林」、「北海岸及觀音山」、「雲嘉南濱海」及「西拉雅」等，本研究針對到訪 13 個國家風景區轄內遊憩據點之年滿 12 歲以上本國遊客為調查對象。

釋出項目計有：資料使用說明、問卷檔、SPSS 資料檔、STATA 資料檔、ASCII 資料檔、SAS 欄位定義程式、過錄編碼簿及報告書。

※本資料為公共版，SRDA 會員登入後直接下載使用。

■ 交通部統計處「96、98、100 和 103 年機車使用狀況調查」

「機車使用狀況調查」的目的主要在蒐集臺灣地區機車之持有、使用、通(勤)學、支出及機車使用者對政府相關措施之意見等資料，以供相關施政參據。

本調查以臺閩地區為調查區域範圍，以機車車主(或使用者)為調查對象，並以機車車輛為調查單位。調查週期為每 2 年辦理 1 次，採郵寄問卷調查。調查項目包括：

- 1、使用者基本資料、持有情形、使用情形、通勤(學)情形、支出情形及相關措施之意見；
- 2、每月花費之停車費含路邊、路外、建築物附設停車位等之停車費；
- 3、機車平均每次使用時間、行駛公里等，來回算 1 次；
- 4、全年保養維修費用包含零件、配件、機油更換、車身及引擎維修，但不含燃料稅、汽油費及折舊費。

釋出項目計有：資料使用說明、問卷檔、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及報告書等。

※本資料為公共版，SRDA 會員登入後直接下載使用。

■ 交通部統計處「98 年-104 年民衆日常使用運具狀況調查」

「民衆日常使用運具狀況調查」創辦於民國 98 年，為估算各縣市之公共運輸市占率，了解民衆未搭乘公共運輸工具(以下簡稱公共運具)之原因，並探究民衆搭乘各項公共運具之滿意度，俾作為交通部推動「公路公共運輸提昇計畫」之施政參考依據，協助相關政策規劃並提供各界參考應用。

原調查範圍為臺灣地區，自民國 99 年起擴大調查範圍為臺閩地區(包括臺灣地區及金馬地區)。調查係對臺閩地區年滿 15 歲以上民衆進行電話訪問。調查結果已對受訪者性別、年齡、縣市及鄉鎮進行樣本代表性檢定後加權處理。

釋出項目計有：資料使用說明、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及摘要分析等。

※本資料為公共版，SRDA 會員登入後直接下載使用。

■ 交通部統計處「98 年自行車使用狀況調查」

為瞭解自行車的使用情形、騎自行車通勤(學)的潛在需求、對自行車行車安全的看法、對設置自行車專用道的看法及民衆持有及預計購買自行車的狀況，供作推動自行車相關政策之參據，交通部統計處特創辦「自行車使用狀況調查」。

此調查係於 98 年 11 月 3 日至 22 日委託全國意向顧問公司以電話訪問方式辦理，針對年滿 12 歲以上民衆，以臺灣地區住宅電話為母體隨機抽樣(電話號碼後 4 碼隨機)，調查結果並對受訪者性別、年齡及地區進行樣本代表性檢定。此次調查成功訪問 9,236 位 12 歲以上民衆，在 95%信心水準下，抽樣誤差為 $\pm 1\%$

釋出項目計有：資料使用說明、ASCII 資料檔、SPSS 資料檔、STATA 資料檔、SAS 欄位定義程式、過錄編碼簿及摘要分析。

※本資料為公共版，SRDA 會員登入後直接下載使用。

■ 國家發展委員會「93 至 98 年民衆對電子化政府相關議題的看法」

「行政院研究發展考核委員會 (現為國家發展委員會)」為瞭解民衆對電子化政府的看法，辦理「民衆對電子化政府相關議題的看法」民意調查。

本次釋出 93 到 98 年調查資料，調查是以臺灣地區住宅電話為母體，並以尾數 2 碼隨機進行，訪問對象是以台灣地區 20 歲以上之成年人。93 年 4 月調查有效樣本 1,099 案；93 年 10 月調查有效樣本 1,094 案；94 年調查有效樣本 1,090 案；95 年調查有效樣本 1,089 案；96 年調查有效樣本 1,112 案；97 年調查有效樣本 1,083 案；98 年調查有效樣本 1,106 案。

釋出項目計有：資料使用說明、問卷檔、ASCII 資料檔、SPSS 資料檔、STATA 資料檔、SAS 欄位定義程式、過錄編碼簿及摘要表。

■ 行政院原住民族委員會「97 年、102 年、103 年原住民就業狀況調查」

行政院原住民族委員會為蒐集並了解原住民族勞動力狀況、就業、失業情形、勞動意願等問題，以抽樣調查方式，進行原住民族就業狀況調查，調查結果提供政府相關單位制定原住民族各項勞動政策、就業服務及職業訓練之參考。

「原住民就業狀況調查」以設籍在臺閩地區之住戶內，十五歲以上之本國籍具有原住民身分者為調查對象，採多階段抽樣方法，以電訪訪問方式訪問原住民家戶成員之就業狀況。

本次釋出資料分為會員版及限制版兩種，限制版提供族別、縣市鄉鎮地區等變項，因具有敏感資訊，因此只提供現場使用(Onsite Service)及遠端連線使用(Remote Service)，亦不提供資料下載，僅能取回經分析後無機敏資訊的報表。

釋出項目計有：資料使用說明、問卷檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿、報告書(102 年)。

【資料開放：限制性資料】

■ 行政院原住民族委員會「103 年原住民族經濟狀況調查」

為能持續掌握我國原住民族家庭經濟現況，本研究以抽樣調查方法瞭解我國原住民族家庭收支及經濟狀況，包含家庭的收支狀況、經濟壓力狀況、貧窮狀況、住宅及財產擁有狀況、收入及負債狀況等。訪問對象為臺灣地區各縣市鄉鎮之原住民族家戶(指戶內至少有 1 人 15 歲以上且具有原住民族身分者)的主要或次經濟戶長。調查時間於 103 年 9 月 21 日開始面對家戶訪問，並於同年 10 月 20 日完成。有效樣本戶數為 5,214 戶。

限制版資料提供族別、縣市鄉鎮地區等變項，因具有敏感資訊，因此只提供現場使用(Onsite Service)及遠端連線使用(Remote Service)，亦不提供資料下載，僅能取回經分析後無機敏資訊的報表。

釋出項目計有：資料使用說明、問卷檔、SPSS 資料檔、STATA 資料檔。

【資料開放：個別型研究計畫調查資料】

以下個別型研究計畫，依學科類別及登錄號順序列出計畫名稱及計畫主持人。如欲查閱詳細的內容，請依照登錄號翻到後方頁面，除計畫名稱及計畫主持人之外，表格中會再提供計畫執行單位、計畫執行期間及摘要等資訊。

政治學

- E10102 隱藏黨派傾向選民之真實意向：調查與預測／劉正山
E10124 「台灣人/中國人認同」的再探究：兩岸交流下的台灣民衆身分認同／蒙志成
E10129 公共服務動機對政策擴散之影響：「閱讀起步走」計畫之個案分析／孫同文

社會學

- C00218 態度量表之選項數量、標示語及不平衡設計(I)／廖培珊
C00244 台灣地區基因體意向調查與資料庫建置之規劃(II)／于若蓉
E10112 從「巧語」到「現實」－伙伴關係與專業團隊運作模式之建構／黃源協
E10127 性剝削抑或性交易？兒童及少年性交易防制工作服務個案之需求與處遇策略探究／陳毓文
E10128 社會服務產業的制度化和網絡鑲嵌：以機構照護服務產業為例／紀金山
E10136 台灣與香港的社會企業之治理研究／官有垣

管理學

- E10204 陸客來臺觀光餐旅產業之藍海策略：以個體價值承諾、組織辨識、組織契合度之觀點探討／楊仁德
E10253 員工價值主張：探索進階人力資源管理根基／溫金豐
E10305 二元俱存是否可以提升供應商績效？聯盟二元俱存的跨層次交互作用之角色／黃銘章
E10313 逃離組織或發出建言以對抗不當督導？從情感事件理論探討不當督導產生的影響／李芸蘋
E10314 網路拍賣再購買意圖影響因素：詐欺以及價值觀點／何建達

教育學

- C00141 生產力、效率與品質面衡量：知識密集服務產業：大學生與資訊素養／傅祖壇

- E94023 學前教師的異動傾向：社會交換理論與個人－組織適配理論之驗證／陳怡靖
- E10050 性別、學習、與成功老化：高等教育下之老人正規學習經驗／陳麗光
- E10222 國小教師社會互動學習能力教學自我效能量表的發展與調查研究／黃永和
- E10231 高齡教育社會效果之調查及國際比較／林麗惠

經濟學

- C00143 知識經濟與電子商務之整合性研究：大學生在校所學與就業表現／傅祖壇
- C00144 知識經濟與電子商務之整合性研究：技職生在校所學與就業表現／傅祖壇

傳播學

- E10107 紙質媒體和網路媒體對低健康素養婦女之教育成效比較分析：以台北市新移民母親為例／洪雅玲
- E10125 社群網站對於台灣政治傳播及年輕選民的影響／林日璇

區域研究與地理

- C00231 我國國民全球暖化節能減碳核心素養之養成／朱瑞玲
- D00125 莫拉克颱風社會衝擊與復原調查(第三期)／國家災害防救科技中心
- E99026 影迷涉入、偶像崇拜與目的地意象關係之研究：影視觀光之觀點／顏昌華
- E99040 新瓶與老酒－從「休閒參與樣態」探討休閒參與、幸福感、知覺壓力的關係／吳宗瓊
- E10252 地方連結與活動涉入及環境衝擊調適之關係探討／鄭佳昆

C00141			
計畫名稱(中)	生產力、效率與品質面衡量：知識密集服務產業：大學生與資訊素養		
計畫主持人	傅祖壇		
計畫執行單位	中央研究院調查研究專題中心	計畫執行期間	
計畫摘要	<p>本研究主要目的為探究高等教育學校(公私立)在推動知識經濟上之作法，及該作法對學生生產力之影響，比較大學間之效率與生產力差異，並探究知識經濟或資訊科技對學生生產力的影響。</p> <p>本計畫是以台灣大學、政治大學、清華大學、中正大學、成功大學、中山大學、輔仁大學、淡江大學、銘傳大學、東海大學、逢甲大學、文化大學等十二所大專院校三年級學生為研究對象。除了追蹤 2002 年有接受過訪問的學生外，另還加上 2002 年應受訪而未受訪學號為雙號的學生。資料蒐集以自填方式執行之。於 2005 年 5 月開始執行到 6 月。最後，實際完訪樣本數為 2,628 案。</p>		

C00143			
計畫名稱(中)	知識經濟與電子商務之整合性研究：大學生在校所學與就業表現		
計畫主持人	傅祖壇		
計畫執行單位	中央研究院經濟研究所	計畫執行期間	2004/01/01~2015/12/31
計畫摘要	<p>知識經濟是提升國家競爭力之關鍵，高等教育更是知識經濟養成最重要的場所。有鑑於此，政府與學界均非常重視大學的教育內涵及成效。本研究之主要目的，在於探討大學生在校所學，對其未來在職場上之表現是否有所關連。</p> <p>此計畫調查對象區分為二種類，一類為國內 16 所大學經濟相關學系八十六至八十八學年度之男性畢業生，與八十八至九十學年度之女性畢業生；另一類為國內 24 所大學之商學院(非經濟學系)於八十六學年度之男性畢業生與八十八學年度之女性畢業生。抽樣方法乃是利用各大學所提供之學生名單，進行全面性之普查。在二種類調查中，經濟學系寄發 5,504 位受訪者，商學院寄發 7,940 位受訪者。資料蒐集以郵寄問卷方式執行，經濟學系調查時間為 2004 年 1 月 15 日至 2 月 29 日，商學院則為 2004 年 3 月 1 日至 4 月 9 日。最後，實際完訪樣本數，經濟學系 1,134 案、商學院 1,345 案。</p>		

C00144			
計畫名稱(中)	知識經濟與電子商務之整合性研究：技職生在校所學與就業表現		
計畫主持人	傅祖壇		
計畫執行單位	中央研究院經濟研究所	計畫執行期間	2004/01/01~2015/12/31
計畫摘要	<p>科技進步日新月異，而在科技進步的背後要有知識經濟來支撐，知識的創新和傳承仰仗高等教育的人才培育，我國高等教育近年來亦蓬勃發展。本研究之主要目的，在於探討商業類科技職畢業生在校所學，對其未來在職場上之表現是否有所關連。</p> <p>此計畫為國內 27 所大專技職院校日間部商業類科八十六學年度之男性畢業生，與八十八學年度之女性畢業生。利用各大專技職院校所提供之學生名單，進行全面性之普查，以郵寄問卷方式執行，總共寄發 13,120 位受訪者。於 2004 年 11 月開始執行至 2005 年 1 月。最後，實際完訪樣本數為 1,892 案。</p>		

C00218			
計畫名稱(中)	態度量表之選項數量、標示語及不平衡設計(I)		
計畫主持人	廖培珊		
計畫執行單位	中央研究院人文社會科學研究中心	計畫執行期間	2010/01/01~2010/12/31
計畫摘要	<p>本研究企圖以態度量表之選項設計作為測試討論之主體，以實驗設計的形式，對於一般民衆為組成之代表性樣本進行施測，以期所獲取之研究結果可廣泛應用於以調查訪問作為測量工具的實證研究中。在設計態度量表的選項類型時應同時考慮選項數量與相對應之標示語內容，因此本研究分兩階段進行，皆以 split-ballot 實驗設計的方式，運用電腦輔助電話訪問系統蒐集調查資料，來比較不同選項數量及標示語對於回答類型之影響。</p> <p>本計畫以電話訪問進行資料蒐集工作，調查母體為年滿十八歲以上且家中有電話之民衆而調查訪問地區則為台灣地區（不含澎湖及金馬地區）。正式訪問期間為民國 99 年 1 月 11 日至 2 月 5 日，完訪數為 3,339 案。</p>		

C00231			
計畫名稱(中)	我國國民全球暖化節能減碳核心素養之養成		
計畫主持人	朱瑞玲		
計畫執行單位	中央研究院民族學研究所	計畫執行期間	2010/06/01~2011/06/30
計畫摘要	<p>本研究主要目的希望透過訪問調查，了解一般人對於全球暖化與氣候變遷的態度和想法，希望藉此能幫助我們對於台灣地區的生活環境有更進一步的改善。所收集的資料，一方面用以檢證西方既有的相關理論在台灣社會是否適用，並進行氣候變遷相關環境能源議題的跨國比較。另一方面也由資料所呈現的事實中，提供教育部門及學校單位節能減碳正確知識建構及核心素養養成過程之參考改善。</p> <p>研究母體為台北市及台北縣板橋市、三重市、中和市、永和市、新莊市、新店市、汐止市及淡水鎮等八個河岸市鎮，18 歲以上 70 歲以下者(民國 28 年 1 月 1 日至民國 80 年 12 月 31 日以前出生者)，這二個母體均採分層二階段 PPS 抽樣法，以面對面訪問(face-to-face interviewing)方式進行資料蒐集，台北市共抽取 1,734 案，完成 555 案；台北縣共抽取 1,670 案，完成 627 案。</p>		

D00125			
計畫名稱(中)	莫拉克颱風社會衝擊與復原調查(第三期)		
計畫主持人	國家災害防救科技中心		
計畫執行單位	國家災害防救科技中心	計畫執行期間	2012/01/01~2012/12/31
計畫摘要	<p>98 年 8 月 6 日莫拉克颱風侵襲臺灣，對臺灣造成重大災情，為詳細瞭解莫拉克颱風受災家戶的災害衝擊以及復原情況，國家災害防救科技中心於 99 年至 101 年針對向各縣市政府社會處申請「莫拉克颱風安遷救助金」之同一家戶進行每年一次的追蹤調查。本調查的調查範圍有南投縣、嘉義縣、高雄市、臺南市、屏東縣以及臺東縣等 6 個縣市，至 101 年為止，本調查共累積 3 年的資料。101 年為「莫拉克颱風社會衝擊與復原調查」第三次調查，共成功訪問 1,542 戶，調查面向有社會情況、需求與協助、生活狀況、心理狀況以及永久住宅議題，調查結果將用以瞭解受災家戶 101 年的復原狀況與復原趨勢，並根據調查結果提出策略建議。</p>		

C00244			
計畫名稱(中)	台灣地區基因體意向調查與資料庫建置之規劃：2011 年面訪調查		
計畫主持人	于若蓉		
計畫執行單位	中央研究院人文社會科學研究中心調查研究專題中心	計畫執行期間	2011/01/01~2011/12/31
計畫摘要	<p>近年來，隨著基因科技的發展，醫學、農業生產技術等方面產生劇烈的變革。然而直至目前為止，我們仍無法確知這些科技可能帶來哪些後果，也不確知這些結果是否能被人類所控制。儘管各國都在積極發展基因科技，對於其發展可能涉及的倫理、法律、社會等層面的問題，不論政府或民間，瞭解仍相當有限。</p> <p>本計畫的目的，在探討影響民衆對基因科技相關看法與態度的因素，並試圖探究其中間機制。此次調查內容，有部份題組與前一次調查相同，以觀察認知、看法或態度的長期變遷趨勢。另外，也加入一些題組，來探究基因科技、醫學研究、基因改造食品等相關議題，藉以瞭解社會大眾對於基因科技的風險與利益、醫學研究的資金來源與利益歸屬、基因改造食品的購買意願等方面的看法。</p> <p>調查對象為台灣地區具有本國國籍，年齡在 18 歲至 70 歲的民衆(民國 29 年 1 月 1 日至 81 年 12 月 31 日以前出生)。採面對面訪問(face-to-face interviewing)的方式。正式調查原預定完訪 1,500 案，實際抽樣時依預定的膨脹係數抽出 3,134 案，最後有 1,503 案完訪。</p>		

E94023		
計畫名稱(中)	學前教師的異動傾向：社會交換理論與個人－組織適配理論之驗證	
計畫主持人	陳怡靖	
計畫執行單位	中華醫事學院幼兒保育系	計畫執行期間 2005/08/01~2006/07/31
計畫摘要	<p>由於社會變遷，雙薪家庭與雙生涯家庭逐漸成為現代家庭的主要型態，許多幼兒很早就被送到學校—上幼稚園或是托兒所，學前教師成為負責照顧幼兒、教育幼兒的重要角色。但根據過去研究顯示幼教老師高流失率與高異動率的現象一直是台灣學前教育的困擾問題，由於教師異動對於幼兒發展與生活適應，以及園所的整體發展都有不利的影響；因此，研究學前教師異動的影響因素，設法減低教師的異動實為幼教研究領域的重要課題。過去有關職業異動的研究為數不少，但教師異動的研究卻不多，以學前教師為研究對象的更是少數，因此有關教師異動這個主題，幼教領域有必要累積更多的研究結果以建立本土的、專門領域的理論解釋。本研究基於這個觀點，先以實際訪談的結果來分析可能影響學前教師異動的因素，訪談結果發現社會交換與個人-組織適配理論可能能夠解釋教師異動的心態與異動傾向，因此研究者希望能以全國性大樣本的調查來檢證這兩個理論對於台灣地區學前教師的異動傾向是否具有解釋性。</p> <p>研究結果發現，超過 50%的學前教師都表示曾經想異動，而且已經著手進行異動計畫、或是打算一年內就異動的大約佔 20%，顯示學前教師對於這份工作的穩定性普遍不足，教師異動心態確實值得關注。而影響教師異動傾向的關鍵性因素是月薪低、工作負荷量大、通勤時間、工作滿意度低，這些是預測教師異動傾向的重要因素，此一結果符應了「錢少、事多、離家遠」的工作心態。此外本研究也發現物質交換、分配公平是許多背景變項(例如教師的年資、教育程度、婚姻狀況、任教班別、學校班級數)對於異動傾向的中介影響因素。顯示公平理論以及社會交換理論對於學前教師的異動傾向具有解釋效果，而適配理論則否；因此本研究初步確立公平理論、交換理論係透過工作滿意度的中介而影響異動傾向的關係模式。</p>	

E99026			
計畫名稱(中)	影迷涉入、偶像崇拜與目的地意象關係之研究：影視觀光之觀點		
計畫主持人	顏昌華		
計畫執行單位	國立臺中技術學院休閒事業經營系	計畫執行期間	2010/08/01~2011/07/31
計畫摘要	<p>隨著影視娛樂產業及海外旅遊的成長，刺激全球影視旅遊的快速發展。影視旅遊又稱為電影引發之觀光旅遊，意指遊客因該旅遊地區或景點曾於電視、錄影帶或電影中出現而前往觀光旅遊。近年來影視旅遊之相關研究成長快速，然而，國內卻較少學者針對此議題進行探討。過去研究發現影迷涉入程度與旅遊目的地意象之間並未有顯著關係，然而，由於影迷涉入程度愈高，可能轉為偶像崇拜，進而形成旅遊目的地意象。因此，偶像崇拜可能是存在影迷涉入與目的地意象之間的中介變數。此外，地理上的距離也可能是影響其間關係的重要因素。因此，本研究計畫目的在探討影迷涉入、偶像崇拜及目的地意象之關係，其次，了解國內和國外旅客在影迷涉入和目的地意象關係之差異。本研究結果可以提供管理者未來進行旅遊目的地行銷之參考，最後，根據研究結果發現，提出行銷策略之實務意涵及未來研究方向。</p>		

E99040			
計畫名稱(中)	新瓶與老酒－從「休閒參與樣態」探討休閒參與、幸福感、知覺壓力的關係		
計畫主持人	吳宗瓊		
計畫執行單位	國立東華大學觀光暨休閒遊憩學系	計畫執行期間	2010/08/01~2014/01/31
計畫摘要	<p>本研究分為三個階段三個年度進行，第一階段(第一年)為建構「休閒參與樣態」指標，並檢視其與可能影響因子之關係。研究群體為台北縣市成年大眾(配合主計處的統計分類，採 15 歲(含)以上)，主要是因為希望所建構的休閒參與指標能反應一般大眾的「休閒參與樣態」，完成有效問卷為 2,011 份。</p> <p>第二個階段(第二年)，主要檢視與比較分析「休閒參與樣態」與「幸福感」之關係。問卷設定的研究群體，為退休早期者，其年齡範圍為 50 歲以上且退休時間五年內者稱之。但退休後另外從事全職工作，且領有全新者不包含在本研究對象之內。抽取樣本方式為立意抽樣，即在各縣市火車站或民衆聚集處尋找合適之研究對象。判定標準先以目</p>		

E99040	
	<p>測為主，攔截近似年齡的民衆，進一步詢問其年齡與條件是否符合本研究之對象後，再給予問卷填答，有效問卷為 258 份。</p> <p>第三個階段(第三年)，主要檢視與比較分析「休閒參與樣態」與「知覺壓力」之關係。問卷設定的研究群體，為職業婦女者，由於缺乏職業婦女的完整名單，因此本研究之抽取樣本方式為立意抽樣，以大台北地區為主，在台北火車站或職業婦女較常出入之地區尋找合適之研究對象，有效問卷為 305 份。</p>

E10050			
計畫名稱(中)	性別、學習、與成功老化：高等教育下之老人正規學習經驗		
計畫主持人	陳麗光		
計畫執行單位	國立成功大學老年學研究所	計畫執行期間	2011/08/01~2013/07/31
計畫摘要	<p>為探索性別、高齡學習、與成功老化的關連性，本計畫以註冊於各大學系所之高齡者為對象，擬訂兩年期程，藉兼用量性與質性資料的混合式研究設計，達成兩項研究目的：(1)描繪高等教育環境中的高齡學習者特徵；(2)探討高齡學習者於高等教育環境下之學習經驗及學習與其主觀所欲成功老化之可能關係。第一年以問卷方式進行資料之蒐集，並與一個常模做比較；第二考量第一年的量性資料決定參與者標準，以便立意採樣，資料之來源以深度訪談為主。本計畫的執行將更豐富政策制訂者、實務工作者、學術研究者、與一般民衆對高齡者異質性及其多元學習觀點與需求的瞭解，亦有助於對未來世代共學做最好的思考與預備。有效問卷 287 份。</p>		

E10102			
計畫名稱(中)	隱藏黨派傾向選民之真實意向：調查與預測		
計畫主持人	劉正山		
計畫執行單位	國立中山大學政治學研究所	計畫執行期間	2012/08/01~2015/07/31
計畫摘要	<p>本研究計畫為三年期的應用/技術型計畫，目標在於結合理論及調查的實務，探求隱藏黨派傾向選民之真實意向，增加學術界對於隱藏黨派傾向選民之真實意向的洞察能力。我國有相當高比例的選民會在調查中隱藏其真實的政黨傾向。這樣的選民不見得只是項目無反應(item-nonresponse)而已，他們往往宣稱自己是中間或中立選民。從理論上來看，學者有足夠的理由相信，多數選民有其真實的政黨傾向，只是現有的測量方式不易洞察。也因為這個長久以來理論與實務的落差，造成目前要瞭解政黨支持分佈，得依靠有實際開票結果作對照的選舉調查；非選舉期間所進行的調查之預測力十分有限。本年度的主題文獻仍在整理中，但資料收集的工作已經開始。本年度除了文獻之外，亦啟動了第二年的部份工作，先精進研究方法，將 multiple imputation 應用在今年度收集的電話調查資料，來增強對中間選民的政黨傾向的預測能力。</p>		

E10107			
計畫名稱(中)	紙質媒體和網路媒體對低健康素養婦女之教育成效比較分析：以台北市新移民母親為例		
計畫主持人	洪雅玲		
計畫執行單位	輔仁大學影像傳播學系	計畫執行期間	2005/08/01~2006/07/31
計畫摘要	<p>根據內政部資料顯示，2010年，台灣有47萬新移民，其子女數目亦超過20萬人，這個研究關心這個特定族群。所謂的新移民母親，大多來自越南、印尼、泰國、中國、菲律賓等地的東南亞籍配偶，其中大部分教育水平較低，且大多嫁入台灣的貧困家庭，因此他們的孩子經常因缺乏適當的照顧而生病。目前紙質媒體和網路媒體，可能是最流行的健康宣傳媒體，人們依賴它們來搜尋醫療信息。本研究目的在於比較紙質媒體和網路媒體所提供的兒童過敏保健資訊，對於低健康素養的新移民母親的教育成效，從而建立一套低識字病人適用的健康傳播媒體的設計準則。這項研究將分五個階段實施，首先回顧既存的文獻，以確定可以用來評價低識字病人適用的健康宣傳媒體的策</p>		

E10107	
	略，方法和工具；其次，邀請新移民母親進行焦點訪談，以了解他們對於兒童過敏保健的基本概念，與使用何種標準來評價健康宣傳媒體。第三設計製作具體可用的測試媒體，包括靜態宣傳單和動態的網站，介紹常見的兒童過敏症的發生原因、症狀、預防；第四，預先測量問卷的可靠性和有效性，以確保所有問卷中的問題皆是良好的衡量指標；第五，進行媒體觀前觀後的理解能力測試，共募集 50 位低健康素養的新移民母親，隨機分為宣傳單與網站兩組，分別探索其教育成效。

E10112			
計畫名稱(中)	從「巧語」到「現實」－伙伴關係與專業團隊運作模式之建構		
計畫主持人	黃源協		
計畫執行單位	國立暨南國際大學	計畫執行期間	2012/08/01~2013/07/31
計畫摘要	<p>為檢視臺灣社會服務之夥伴關係與專業團隊的運作現況，本研究以研究者於 2012 年所建構的「有效夥伴關係」與「有效專業團隊」的兩份量表，選取長期照顧服務領域進行調查研究。本研究採集群分析法進行抽樣，分別抽到臺北市、臺中市、宜蘭縣、嘉義縣及屏東縣等五縣市，並總共以五縣市 292 個組織做為本研究之研究對象。在資料收集方面，是由研究者或訪員到各機關(構)進行面對面的訪問，最終完成 138 個機構，計 256 份有效問卷。透過本研究，期能據以從政策、管理與實務等面向，提出有助於建構長期照顧之夥伴關係與專業團隊的相關建議，進而讓夥伴關係與專業團隊在長期照顧領域的實踐，能夠從「理想」朝向「現實」邁進，進而提昇我國長期照顧服務輸送的品質。</p>		

E10124			
計畫名稱(中)	「台灣人/中國人認同」的再探究：兩岸交流下的台灣民衆身分認同		
計畫主持人	蒙志成		
計畫執行單位	國立成功大學政治學系	計畫執行期間	2012/08/01~2014/07/31
計畫摘要	<p>此研究擬針對「台灣人/中國人認同」此一重大主題，探究為何台灣民衆在身份認同上，從民主轉型以來，存在著穩定增加的「台灣人」認同趨勢。此外，不論在那一時期的台灣政治發展歷程，具有「台灣人與中國人」的混合認同，為何長期扮演著台灣民衆身份認同中的主要趨勢。最後，從 1980 年代末期以來的兩岸人民交流，如何在實務上影響、並在概念上型塑出台灣民衆的認同思路。</p> <p>此計畫擬在「當前民族主義思路下的身份認同」之理論架構下、輔以族群認同研究所提出之「認同循環論」，從台灣民衆因兩岸間不同民族主義思路的接觸出發，到因不同交流層次的衝突摩擦而後沈澱所形成的認同內涵，結合整合性民意調查所呈現之「台灣人/中國人」認同的長期走向，完成具詮釋性且能連結經驗數據的理論論述，並依其所導引出來的命題建構可驗證的測量變數與統計假設。期望此研究主題，一方面在理論詮釋上，能夠妥善地提供台灣民衆身份認同的動態解釋，另一方面，在方法上，亦希望立基於目前所大型民調資料累積的資料成果，以多元的統計模型與估計方法，突破既有的資料建構限制，來進行相關的台灣民衆身份認同的動態分析研究。</p>		

E10125			
計畫名稱(中)	社群網站對於台灣政治傳播及年輕選民的影響		
計畫主持人	林日璇		
計畫執行單位	國立交通大學傳播與科技系	計畫執行期間	2012/02/01~2014/07/31
計畫摘要	<p>有鑑於現有的研究及文獻多為橫向研究探討社群網站使用與線上線下政治參與的相關性，缺乏長期縱向研究，因此此系列研究針對 2012 年總統大選，於大選前一個月(第一次)及大選後兩週(第二次)，透過問卷蒐集受測者社群網站的政治活動使用及投票、態度等方面進行縱向研究(longitudinal study)調查。</p> <p>招募方法除了透過向全校學生發校園公告鼓勵填寫問卷之外，研究團隊也透過臉書上的新生社團張貼招募文宣。此研究的研究對象</p>		

E10125	
	<p>不限於大學生，因為研究問題提出了大學生及較年長的使用者在臉書政治使用上的差別，因此亦針對其他年齡層的選民收集資料，於 PPT 上各大看板張貼招募訊息，第一份問卷於選前一個月前開始收資料，而第二份問卷於選後隔天立即向第一波參與者寄出問卷填寫邀請信。研究團隊提供各 600 份\$100 元的全家禮券給第一波及第二波的參與者，並於兩份問卷皆收完後的一個月寄出 1,200 份禮券給全台參與者(約二月下旬)。</p> <p>此研究共收集到 1,036 份有效問卷(第一波，有資格投票者)，而第二波參與填寫問卷人數為 523 人(510 人有資格投票)，兩波合計為 50.48%的保留率。</p> <p>此系列研究論述臉書上看到別人的波文等，也算是一種被動的線上政治參與。由於臉書上，容易無意間看到不同論點的政治資訊，對於政治沒有興趣或沒有相關知識的使用者，透過臉書上連結的波文，也能看到相關資訊。亦即，人際連結成為政治相關資訊穿透的工具。因此，傳統的政治參與定義在社群網站上，有不同的形式與定義。此研究將 differential gains model 延伸到社群網站上被動的政治參與，亦即，看到波文也算是政治討論的一種形式，因此在社群網站上，不一定要真正波文討論，但接收到資訊也能達成 differential gains，並進一步對選舉有更高的興趣。</p>

E10127			
計畫名稱(中)	性剝削抑或性交易？兒童及少年性交易防制工作服務個案之需求與處遇策略探究		
計畫主持人	陳毓文		
計畫執行單位	國立臺灣大學社會工作學系	計畫執行期間	2012/08/01~2014/07/31
計畫摘要	<p>現今被保護安置之性交易個案多非被押賣的少年，安置為主的處遇模式常遭受一些質疑，然而在實務運作上卻未有改變。本研究企圖能夠了解以這些孩子及服務他們的社工人員為關注對象，檢視性交易兒少之創傷經驗、身心狀況及對性交易行為的態度，並探究兒少及社工對現行處遇模式的看法。本研究分兩年進行，第一年採量化研究設計，針對國內目前因違反「兒童及少年性交易防制條例」而接受緊急、短期或長期安置處遇之未滿 18 歲，且安置至少兩週以上之兒童少年進行問卷調查，以了解其對性交易行為的看法，成長經驗以及身心健康</p>		

E10127	
	<p>狀況，共完成 335 份有效問卷。調查結果顯示，有極大比例的受訪者曾經使用毒品、喝酒，其過去成長經驗中也多有遭肢體暴力與性暴力傷害的經驗，而其自殘與自殺的比例也遠高於一般青少年，或其他因家庭失功能而接受安置的兒童少年。有相當高比例的受訪者來自失功能的家庭，她們長期受到各種形式的家庭暴力對待，有將近四分之一的少女受到他人不當的性對待，經歷各種風險與傷害健康的行為，而且有相當高比例的少女曾經以自殘的方式對待自己，有自殺意念或行為的比較也不低。</p>

E10128			
計畫名稱(中)	社會服務產業的制度化和網絡鑲嵌：以機構照護服務產業為例		
計畫主持人	紀金山		
計畫執行單位	靜宜大學社會工作與兒童少年福利學系	計畫執行期間	2012/08/01~2014/07/31
計畫摘要	<p>在當代社會服務需求日增，各國因面臨財政危機卻得削減社會福利的窘境下，福利多元主義開始廣受青睞，民間部門資源對補足福利服務缺口的重要性也隨之大增。然而，由於福利資源緊縮的趨勢，福利多元化議題若僅侷限福利觀點和非營利面向，勢將擴大志願失靈的問題，並低估民間部門動員的潛力。因此本計畫採取廣義的產業觀點和組織場域的實證取向，再聚焦於臺灣早期殘補式社會福利和近期社會福利政策發展所形成的機構照護服務產業，以深入探討此一與西方福利國家發展過程有別，但結果卻又相似之多元服務體系特質和意義。本計畫以兩年為期程，結合政府統計和二手資料分析、問卷調查、深入訪談、焦點團體等研究方法，進行整合性研究著重探討臺灣機構照護服務產業體系的多元特質、產業的組織趨同現象與變遷，及探討其產業趨同的制度化和網絡鑲嵌機制，再據以探討此一產業發展社會企業的可行性和條件。</p>		

E10129			
計畫名稱(中)	公共服務動機對政策擴散之影響：「閱讀起步走」計畫之個案分析		
計畫主持人	孫同文		
計畫執行單位	國立暨南國際大學公共行政與政策學系	計畫執行期間	2012/08/01~2014/07/31
計畫摘要	<p>台中縣沙鹿鎮立深波圖書館，是台灣第一個推動「Bookstart」的發源地。當時該圖書館陳錫冬館長，曾在天下雜誌中閱讀到英國 Book Trust 基金會推動 Bookstart 的相關報導；在此同時，為配合文建會所推動的「公共圖書館空間及營運改善計劃」，致使陳館長將 Bookstart 視為改善圖書館營運項目的契機。正因陳館長的動心起念，使得深波圖書館成為台灣 Bookstart 運動的濫觴。</p> <p>Bookstart 計畫從英國擴散到台灣，目前國內多是從幼兒教育的外部觀點出發，並論及其對於嬰幼兒的學習效果；相對地，鮮見由內部因素分析 Bookstart 政策的擴散。再者，政策擴散之後的成效，亦取決於行政人員對於該項政策的態度與動機。事實上，台灣 Bookstart 政策方案的擴散，圖書館館長的態度至關重要。然而，若要超越市場誘因機制的設計、傳統激勵理論的論述，就必須思索為何行政人員基於何種原因，誘發一個館長不但關心 Bookstart 的發展，更納為自己所屬組織公共服務的重點項目。故而，本研究的主要核心問題，即在於「圖書館館長個人的何種公共服務動機，會成為影響 Bookstart 政策擴散的決定因素」。本研究是以全國圖書館參與 Bookstart 政策的發展過程為主要範疇，因此藉由全國推行「Bookstart」政策，探討政策擴散的演變過程。除知悉圖書館願意參與「Bookstart」政策原因之外，更希冀進一步瞭解圖書館館長的公共服務動機，是否為政策擴散中之重要因素。一方面能提供爾後「Bookstart」政策至其他圖書館擴散之可行性策略，另一方面亦可知悉政策執行者之公共服務動機，是否為影響政策擴散之重要因素，進而於第二年建構台灣政策擴散的模式。</p>		

E10136			
計畫名稱(中)	台灣與香港的社會企業之治理研究		
計畫主持人	官有垣		
計畫執行單位	國立中正大學社會福利學系	計畫執行期間	2012/08/01~2014/07/31
計畫摘要	<p>本實證研究計畫探討台港兩地社會企業的治理結構與功能的動態演變情形，對於政府的政策制訂者、社會企業的領導者、實務工作者，以及探索此議題的學術研究者而言，皆有其重要性與價值性。瞭解社會企業的治理情況與功能發揮良窳，可以指引這類 NPO 有效地實踐其社會價值(social values)並發揮社會影響力(social impacts)，畢竟社會企業存在的目的，除了工具性的經濟目的外，最重要的即是達成社會目的(social purpose)。樣本共計發放 430 家社會企業(包含工作整合型、社區發展型、社會合作社，服務提供與產品銷售、公益創投與其他類型等)，計回收 110 家，回收率為 25.6%。</p>		

E10204			
計畫名稱(中)	陸客來臺觀光餐旅產業之藍海策略：以個體價值承諾、組織辨識、組織契合度之觀點探討		
計畫主持人	楊仁德		
計畫執行單位	國立高雄餐旅大學旅館管理系	計畫執行期間	2013/08/01~2014/07/31
計畫摘要	<p>本研究旨在探討員工－組織契合度、員工－工作契合度、組織內競爭、組織聲望、變革領導以及個人專業能力，對於旅館業員工在顧客服務表現之理論驗證。用於本比較研究之樣本數總計為 619 份，分別蒐集自臺灣國內之國際與本土品牌連鎖旅館，運用 LISREAL 軟體進行分析後，二種不同型態之品牌連鎖旅館呈現以下結果：一、組織外部聲望對於員工－組織契合度和員工－工作契合度具有顯著效果，促使員工顧客服務表現有正向影響。二、個人專業能力與組織內競爭，對員工－組織契合度於統計數據並無達到顯著水準。本土品牌連鎖旅館資料經分析發現，變革領導對於員工－組織契合度及員工－工作契合度之知覺作用具有顯著影響。針對國際品牌連鎖旅館，各變數間關係並不顯著，僅於組織內競爭與員工－組織契合度之層面為顯著相關，但本土品牌連鎖旅館卻不顯著。</p>		

E10222			
計畫名稱(中)	國小教師社會互動學習能力教學自我效能量表的发展與調查研究		
計畫主持人	黃永和		
計畫執行單位	國立臺北教育大學教育學系	計畫執行期間	2013/08/01~2014/07/31
計畫摘要	<p>此研究旨在發展一具有信效度的國小教師社會互動教學自我效能量表，並針對臺北地區教師進行抽樣調查研究，以瞭解教師社會互動教學自我效能的現況，期能有助於國內社會互動教學與研究的發展。此研究根據相關文獻探討結果，進行量表編製、預試與信效度考驗，然後再以分層隨機叢集方式抽取臺北市教師與新北市教師進行正式問卷施測。研究結果指出：(1)本研究量表各題鑑別度良好，所有題目在測量同一特質上具有高度一致性；因素分析結果可以得出「合作互動」、「討論互動」與「上台互動」三個因素，因素負荷量均達.75 以上，可解釋變異達 84.64%；Cronbach α 信度係數介於.90~.97 之間；本研究量表與中文版教師效能感量表的相關係數為.74，效標關聯效度良好；(2)臺北地區國小教師對於自己是否具備能力指導學生互動學習的效能評估，其平均數約只有中上程度把握，而感到非常有把握的教師只有 5~8%左右；(3)不同地區、性別、學歷、年資與班級學生數的臺北地區國小教師，其社會互動教學自我效能並無顯著不同。最後，此研究針對社會互動教學與自我效能研究提出數點建議。</p>		

E10231			
計畫名稱(中)	高齡教育社會效果之調查及國際比較		
計畫主持人	林麗惠		
計畫執行單位	國立中正大學成人及繼續教育學系	計畫執行期間	2013/08/01~2014/07/31
計畫摘要	<p>此研究建基在 101 年度「高齡教育的社會效果評估：健康及公民參與指標之建構與應用」的執行基礎上，主要以活躍老化觀點、生態系統理論，以及影響社會效果的基本概念模型作為理論基礎，並結合高齡教育、社會效果的相關研究，旨在發展「高齡教育的社會效果評估：學習者觀點」量表，並進行高齡學習者的調查。為達研究目的，此研究採問卷調查法進行研究，針對樂齡學習中心、樂齡大學、長青學苑的學習者發出問卷，共計發出 10,690 份，回收有效問卷 9,033 份，有效回收率為 84.50%，並針對調查所得資料進行統計分析，共獲致八項結論，包括：第一、高齡教育社會效果包含健康、公民參與兩個層面；第二、健康層面包含健康知識、健康狀態兩個向度；第三、公民</p>		

E10231	
	參與包含個人、家庭、社區、社會四個向度；第四、高齡教育的社會效果不會因為年齡的不同而有所差異；第五、女性參與高齡教育的社會效果大於男性；第六、低教育程度的參與者產生較高的社會效果；第七、各類高齡教育機構的課程類型有不同程度的社會效果，長青學苑的人文藝術課程社會效果最佳，樂齡學習中心的貢獻課程社會效果最佳，樂齡大學的學校特色課程社會效果最佳；第八、相較於長青學苑、樂齡大學而言，樂齡學習中心產生較大幅度的社會效果。

E10252			
計畫名稱(中)	地方連結與活動涉入及環境衝擊調適之關係探討		
計畫主持人	鄭佳昆		
計畫執行單位	國立臺灣大學園藝暨景觀學系暨研究所	計畫執行期間	2013/08/01~2014/07/31
計畫摘要	<p>地方連結近幾年來已成為景觀休憩領域研究的重點之一，然而對於地方連結如何形成、如何變動、以及如何影響環境態度的機制仍然不清楚。相關文獻指出人們會因為活動涉入而與地點產生情感連結，並且在面對環境改變時，不同程度的地方連結會在不同時間點有不同的變化；而不同目的性之使用者在地點改變活動無法進行之後反應也可能不相同。這些理論需要再進一步的比較與確認。有鑑於此，本研究以固定樣本追蹤法分別針對全園關閉及部份功能整建關閉之地點，調查使用者在中斷前、後之情感依附變化與活動狀況，並追蹤其調適行為。期望能對地方連結在地點中斷後的改變上提出一個整合的架構，以供環境規劃及休憩管理上之參考。</p>		

E10253			
計畫名稱(中)	員工價值主張：探索進階人力資源管理根基		
計畫主持人	溫金豐		
計畫執行單位	國立中山大學人力資源管理研究所	計畫執行期間	2013/08/01~2015/07/31
計畫摘要	<p>數十年來，人力資源管理功能致力於採取不同的措施以提高組織與員工的績效時，大多站在組織觀點而非從員工角度。隨著人才的競爭越來越激烈，「員工價值」的議題在組織研究逐漸受到重視，但是尚未發展出穩固的理論及研究方法。本研究從行銷領域的顧客價值主張觀點出發，嘗試討論以員工價值創造為基礎的人力資源管理實務；我們透過焦點團體訪談，先探索描述員工價值主張的構面，並進一步發展出量表，以量化資料實際分析績效良好的企業可能採取的員工價值主張，我們發現了三種員工價值主張的型態，並說明採取這些員工價值主張的企業特性，以提供進一步的實務與研究建議。</p>		

E10305			
計畫名稱(中)	二元俱存是否可以提升供應商績效？聯盟二元俱存的跨層次交互作用之角色		
計畫主持人	黃銘章		
計畫執行單位	靜宜大學企業管理學系(所)	計畫執行期間	2014/08/01~2015/07/31
計畫摘要	<p>具備「二元俱存」的組織能力被視為組織的動態能力，因此，二元俱存能力成為提升廠商績效的重要因素。雖然部分研究的結果支持二元俱存能力與廠商績效之間的正向關係但近來也有部分的研究及實證結果並未支持此一主張。因此，二元俱存能力對於廠商績效的影響如何？仍是具有爭議的議題。本研究認為高度整合可能可以創造一個二元俱存的供應鏈網絡系絡，克服探索策略與利用策略之間的張力，使得供應鏈中的個別廠商得以受惠，因而釐清廠商二元俱存能力對績效影響的關係。本研究以中衛體系的廠商進行實證，共計回收 30 家中心廠與 103 家協力廠的資料，配對後進行 HLM 分析。研究結果發現：廠商的二元俱存能力對其績效有正向的影響；供應鏈整合的程度對供應鏈的二元俱存能力有正向影響；供應鏈二元俱存能力對於供應商績效有正向影響，同時對於供應商二元俱存能力與供應商績效之間的正向關係有正向的調節效果。</p>		

E10313			
計畫名稱(中)	逃離組織或發出建言以對抗不當督導？從情感事件理論探討不當督導產生的影響		
計畫主持人	李芸蘋		
計畫執行單位	國立臺南大學行政管理學系	計畫執行期間	2014/08/01~2015/07/31
計畫摘要	<p>不當督導是一種職場上常見的情形，並且往往為員工帶來負面的影響。本研究以情感事件理論為基礎，提出一個理論架構來討論不當督導對員工的影響，試圖了解不當督導是如何影響到員工在情感上與行為上的反應，同時，考量到潛在的組織因素可能也會對員工有所影響，因此也將組織公平納入模型中探討。本研究目的在於了解員工對於不當督導的情感與行為反應機制，另外也試圖了解組織公平是否為重要的調節因素，影響員工對於不當督導的所產生的反應。研究資料於 2014 年 1 月至 6 月收集，最後完成之有效樣本數為 233 件。</p>		


E10314			
計畫名稱(中)	網路拍賣再購買意圖影響因素：詐欺以及價值觀點		
計畫主持人	何建達		
計畫執行單位	國立中興大學科技管理研究所	計畫執行期間	2014/08/01~2015/07/31
計畫摘要	<p>近年來網路普及與網路拍賣低營運成本、低進入障礙的特性，使得有越來越多的人使用網路拍賣來進行交易，其中僅有少數線上拍賣賣家獲利。因此，瞭解線上買家的購買與再購行為，有助於賣家找出潛在的回流顧客，進而降低成本、增加營收，缺乏信任則是線上交易的另一項阻礙因素。因此，了解在線上拍賣中如何建立信任是一個重要議題，在這個研究中，我們主要討論線上拍賣的購買意圖，過去的研究裡有不同的觀點來討論購買意圖，但卻很少討論到詐欺以及社會影響，同時在這個研究中，我們藉由科技接受模式(TAM)、社會影響、詐欺控制、來討論這些變數是否會影響線上拍賣的再購買意圖。</p> <p>本研究以奇摩拍賣使用者為主要研究對象，結果也意味著，在網路上的拍賣者需採取差異化策略來取得網路拍賣市場的購買意圖。</p>		

欲更進一步了解上述學術調查資料及個別型研究計畫，歡迎利用 SRDA 網站上的【[資料查詢](#)】→【[簡易查詢](#)】功能，以計畫名稱或計畫主持人尋得該筆資料。或可在簡易查詢中設定以登錄號為查詢條件，參照表格右上方的英文數字共 6 碼(例如 C00125)輸入查詢。

「政府調查資料」除公共版之外，需登入 SRDA 網站的「政府調查資料」專區，依資料提供單位找到您需要的資料，勾選欲申請的年度後提出申請，經審核通過後下載。

「限制性資料」需要另外填具書面申請資料，可選擇現場使用(On-site)或遠距使用(Remote Service)，經審核通過後開通帳號。限制性資料只能攜回分析完成後的報表，無法取得原始資料。詳情請參考 SRDA 網站「限制性資料」專區，網址：<https://srda.sinica.edu.tw/srda/restrict/index.html>。

如需取得英文化的資料，請至 SRDA 的英文網頁下載或提出申請。上述資料的使用權限，皆需具備 SRDA 一般會員或院內會員的資格。



資料工程初探 以開放資料清理轉換為例

王學治

【前言】

近年在開放資料趨勢盛行下，普羅大眾能獲得的資料已走向多元，雖然資源豐富，但從資料使用者的角度來看，取得到利用之間，實務上仍存在不少挑戰，例如多樣化檔案格式的處理、檔案內文編碼的轉換、與解讀資料結構的能力與熟悉度等等，對於開放資料的使用者而言，形成一個技術性門檻。進一步來說，這些門檻是資料工程(Data Engineering)¹領域嘗試解決的問題之一，為了使資料能進一步被統合使用，包含截取(extract)、轉換(transform)、讀取(load)等傳統資料倉儲整理資料的手段，在資料工程領域中持續地被精進與探討著。

由於針對處理不同檔案格式的需求，衍生出的相關工具五花八門，加上使用不同程式語言處理資料時，還會面臨如何選擇合適的套件(package)或函式庫(library)等問題。有鑑於此，本文淺析利用開放資料時，在處理資料過程中會遇到的狀況，透過一些背景概念的介紹，以及如何藉由資料工程的程序與工具來處理應對，並以一筆開放資料集為範例，示範將檔案格式從「JSON」轉換成「CSV」，以便後續使用統計軟體進行分析的過程，提供讀者作為使用開放資料的參考。

一、開放資料的定義

在介紹資料工程概念之前，首先對開放資料的定義作簡單介紹。根據開放知識基金會（Open Knowledge Foundation，簡稱 OKFN）開放資料手冊²中敘述，開放資料 (Open Data)指的是「資料能被任何人自由地使用，重新使用與散佈－我們至多只能要求來源標示，與以相同方式分享」。

「開放資料的精神」主要是破除取得對象與再使用目的之限制，讓任何使用者都能不受限制的再次加以利用。此議題原本跟資料儲存的媒介(檔案格式)無關，但由於期望資料儲存的方式，能盡可能地降低資料減損的可能性，並擴大資料的獲取性，於是開放資料本身會被要求必須以「機器可讀取」的格式來儲存；此外，為了能夠普及與長久保存，通常會優先考量以文字形式的電子檔案格式來儲存。有興趣了解此背景的讀者，可進一步閱讀「Digital Ice Age」文章³ (2009)。

1 資料工程 (Data Engineering)參考閱讀

<https://gist.github.com/dataengineer/aaefc8465dcbf6885b38c82931a9fe>。

2 什麼是開放資料？http://opendatahandbook.org/guide/zh_TW/what-is-open-data/。

3 The Digital Ice Age：<http://www.popularmechanics.com/technology/gadgets/a1028/4201645/>。

在 OKFN 手冊所引用的開放定義⁴中，以開放格式散布被認為是作品具備「開放性」的一個必備條件。於是，兼具「機器可讀取」的格式，又能夠被長久利用、保存等條件下，原本在網際網路中常見用來傳輸分享的電子檔案格式，如 XML、JSON、CSV 等，很自然地就變成「開放資料」傳遞格式的首選。

一、開放資料應用層面的挑戰

前言已提及，根據資料提供者的選擇，開放資料除了提供不同檔案格式給使用者運用之外，由於儲存資料的內文編碼(encoding)的不同，也會造成使用者讀取開放資料上的困擾。加上文件內儲存的資訊結構，常是以各資料領域專家(Domain Expert)常用的慣例格式或該領域特殊的規範語法。例如地理資訊系統(GIS)領域中，使用 GeoJSON 檔案格式來傳遞地理資訊，使用者解讀時必須先對該領域或語法規範有所理解，方能正確解析出所傳遞的資訊。

綜上所述，使用者可能面臨資料處理的狀況如下：

- 提供者以多樣化的檔案格式傳遞開放資料
- 資料內容所使用的編碼方式(encoding)儲存可能不同
- 內容的資料結構需要判讀，結構化程度也不同
- 不同檔案格式可選用的工具支援豐富程度不一
- 處理資料的程式語言，有不同的函式與套件可供處理特定檔案格式

二、資料工程是什麼？又如何協助回應上述挑戰？

如果要用一句話來解釋資料工程領域，筆者會推薦參考大數據分析顧問公司 Insight Data 所提出的概念⁵：「資料工程師讓資料科學家更有效率的完成工作(Data engineers enable data scientists to do their jobs more effectively!)」；也就是說，如果要讓資料科學家更專注在尋求問題解答的過程，則會需要有資料工程背景的人員協助處理「資料蒐集、資料倉儲、資料清理、資料格式轉換、資料查詢」等工作⁶，以減少資料科學家的負擔，於是這些工作項目組合在一起，便形成了資料工程領域的內涵。

4 OKFN 開放定義：Open Definition - Defining Open in Open Data, Open Content and Open Knowledge。

5 資料工程 (Data Engineering)參考閱讀 #4 Data Science vs Data Engineering
<https://gist.github.com/dataengeer/aaefc8465dcbbf6885b38c82931a9fe>。

6 資料工程 (Data Engineering)參考閱讀 #5 那些大數據書不會教的資料工程
<https://gist.github.com/dataengeer/aaefc8465dcbbf6885b38c82931a9fe>

有興趣深入了解的讀者，可以參考 Quora 問答⁷ 來進一步理解資料工程領域的相關工作。下表擷取一部分資料領域中的工作，來協助回應上述挑戰。其中，處理使用開放資料「檔案格式轉換」與「資料結構變化」的問題，屬於本文重點，以下將一一介紹。

問題	資料工程領域相關工作	
	擷取	轉換
檔案格式轉換	○	
內文編碼	○	
資料結構變化		○
處理程式語言	○	○
轉換工具	○	○

三、資料結構化種類與轉換須留意的細節

在進行轉換前，讀者必須先了解「資料結構化」的基礎概念，並留意轉換上的細節。按照資料內容的結構化程度不同，大致可分為下列三種類別：

1. 結構化資料
2. 半結構化資料
3. 非結構化資料

「結構化資料」或者稱為表格化資料(**tabular data**)⁸，是一種「每列(**rows**)有相同數量的欄(**columns**)，且每列中的個別欄所紀錄的資料值(**value**)都為相同欄位屬性值」的資料種類。結構化資料常見於網頁超文本(**HTML**)的表格資料，或是關聯式資料庫(**relational-database**)的表格資料中；若轉化成純文字檔案，則以 **csv**(**comma separated value**)格式最為常見。**csv** 是以逗號區隔個別欄位值，也稱作分隔符號文字 (**delimited text**)的一種，其儲存內容規範請參照 **W3C 標準 4180 附錄 A**⁹。由於 **csv** 格式內容簡單，容易產生與編輯，是常用的資料交換格式之一。

「半結構化資料」相對於「結構化資料」，是將「欄位定義」與「資料值」混

7 資料工程 (Data Engineering)參考閱讀 #6 What is data engineering? – Quora
<https://gist.github.com/dataengineer/aaaefc8465dcbbf6885b38c82931a9fe>

8 結構化資料：<https://www.w3.org/TR/2015/PR-tabular-data-model-20151117/#dfn-tabular-data>

9 <https://www.w3.org/TR/2015/PR-tabular-data-model-20151117/#standards>。

合儲存的一種檔案格式；不同於結構化資料事先嚴格規範特定欄位的值域，半結構化資料能較彈性的紀錄欄位與值，所以也常用於資料交換，特別是在組織之間交換資訊時使用。例如常見的 XML 與 JSON 檔案格式，均屬於半結構化資料的一種。

半結構化資料的另外一個特色是常以「巢狀結構」方式來紀錄資料；由父子層級欄位標籤(tag)與值，複合成巢狀結構片段來組合想傳達的資料。由於巢狀結構彈性極大，對需要驗證 JSON/XML 內容者，可額外指定 schema¹⁰或 DTD¹¹綱要，作為資料收取方自行驗證之用。

「非結構化資料」為可變長度、內文格式沒有明確定義的文本，常見檔案格式為 txt。表一整理三種結構化資料，並附上範例與檔案格式，提供讀者對照比較。

表一、結構化資料對照表

資料分類	表格tabular	巢狀結構資料	文字資料、圖、二進位檔
屬性	結構化資料	半結構化	非結構化
範例	age, name, place 25, marx, Taipei	<people> <age>25</age> <name>marx</name> <place>Taipei</place> </people>	My name is marx, 25 years old, lived in Taipei City
檔案格式	csv,xls	JSON/XML/HTML	txt、PNG、PDF

值得注意的是，轉換不同結構化屬性檔案時，由於目標格式結構化屬性的不同，而有資料表現能力的落差。例如由「半結構化資料」轉換為「結構化資料」時，由於來源格式可能有多層巢狀(nested)資訊，當目標格式要求「同一欄位」的值必須描述一致的內容時，就會導致資訊轉換對應上的落差。此時，必須以目標檔案格式所能表達資料的能力，做出適當的選擇，例如將一個 JSON 檔存為多個 csv 檔案，或是改由 Excel 分頁的方式儲存，才能保留完整的資訊。

此外，若僅在語法上符合檔案格式要求，而未滿足「結構化資料定義值」的需求(一個欄位內容屬性要一致)，轉換後的結果也將難以再利用，實務上應避免這種情形。例如，將以下內容(半結構化的資料值)視為文字，直接套入 csv 欄位之中：

"<age>25</age>"

雖然符合此結構化欄位的定義(文字欄位)，內容卻出現半結構化標籤文字

10 W3C XML Schema : <https://www.w3.org/XML/Schema> ; JSON Schema : <http://json-schema.org/> 。

11 DTD (document type definition) : <https://www.w3.org/TR/REC-xml/#dt-doctype> 。

(`<age>25</age>`)，若要利用此欄位的内容(文字 25)，必須經過消除`<age></age>`標籤的步驟；也就是說，欄位内容均符合定義，但卻無法直接使用。對此狀況，可透過改變欄位值定義(改為數值)，或請提供者在提供檔案前先刪除 tag，方便後續利用。理解内容結構化的屬性差異與須留意的細節，可避免轉換後資料使用上的困擾。

四、JSON 格式轉換 CSV 格式範例

為使讀者能夠體會轉換過程，本文以政府資料開放平臺中「各縣市統計區 15 歲以上人口五歲年齡組教育程度統計」資料集¹²作為示範。

【資料介紹與說明】

- 此資料集內按照 22 個縣市行政區劃分，各行政區分別提供 JSON 與 XML 檔案兩種版本供下載使用，本文將挑選其中一行政區資料，並採用 JSON 格式來進行示範，此僅為範例操作，由於行政區界線可能變化，讀者於資料處理利用上仍須留意。
- 本例挑選「台中市統計區資料」
- 本資料收集 15 歲以上人口的五歲年齡組之「教育程度」的人口數交叉統計
- 本資料依「二級發布區」¹³加以統計發布

【資料轉換三步驟】

由於原始 JSON 資料檔內文結構屬於半結構化資料，考慮到使用 csv 格式(結構化資料)的通用性，將資料格式轉換為 csv 檔。

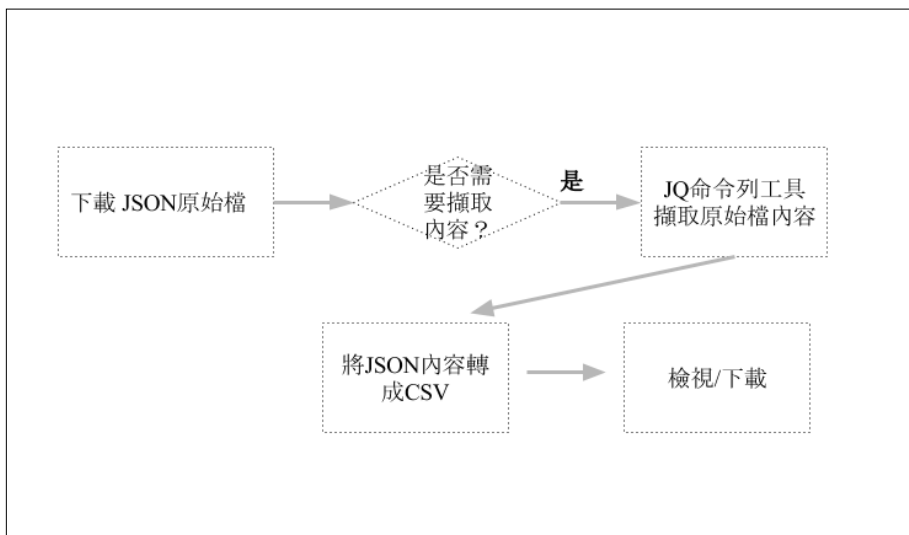
轉換目標：資料格式轉換 (由 JSON → csv)。

轉換步驟如圖一所示，說明如下：

1. 需要從多層半結構化的 JSON 檔中，定位、擷取並輸出所需統計資料部份。
2. 挑選適合的轉換工具，包含命令列或線上工具；這些工具必須支援原始格式的内文編碼。
3. 挑選適當工具來檢視轉換出的 csv 内容，方便驗證轉換内容的正確性。

12 資料集網址：<http://data.gov.tw/node/18629>

13 「二級發布區」的意義可參考「社會經濟資料庫」分組入口網 (http://210.65.89.57/stat/Web/Portal/STAT_PortalHome.aspx)常見問題，或至內政部統計處網站，參閱專題分析(<http://sowf.moi.gov.tw/stat/topic/list.htm>)中有關統計區分類系統文章。



圖一·資料格式轉換步驟

建議讀者在開始轉換之前，可先透過文字編輯器開啓下載的檔案，或利用一般程式語言(R、Python、PHP 等)以相關函式進行讀取動作，檢視原始資料格式，以了解資料原貌。

【資料轉換過程中使用工具】

轉換過程中將用到以下五種工具，取得方式詳見轉換工具¹⁴(①②⑤)。

1. jq (用途 json2json 過濾、挑選轉換)
2. json2csv 命令列工具 (轉換格式)
3. 線上 json2csv 工具(轉換格式工具)
4. csvkit 套裝內的 csvlook 命令列工具 (檢視工具)
5. 文字編輯器 sublime text(檢視工具)

本例透過 linux 作業系統環境執行，相同工作亦可於 windows 環境上完成。表二整理各個轉換步驟所使用的工具及用途。

14 <https://gist.github.com/dataengeer/4e6f8ff0abb7ee787b49645311e9b78>

表二·資料轉換步驟與使用工具說明

轉換步驟	使用工具	工具說明	工具用途
1 從多層半結構化的JSON檔中，定位、擷取並輸出所需統計資料部份。	sublime text	GUI檢視工具	檢視json
	jq	命令列工具	擷取json
2 挑選適合的轉換工具，包含命令列或線上工具；這些工具必須支援原始下載格式的内文編碼	json2csv	命令列工具	轉換工具
	http://www.convertcsv.com/json-to-csv.htm https://konklone.io/json/	線上json2csv網站	轉換工具
3 挑選適當工具來檢視轉換出的csv內容，方便驗證轉換內容的正確性	csvlook	命令列工具	檢視csv

【步驟一：擷取所需 JSON 內容】

1.1 取得原始資料


進入「政府資料開放平臺」資料集頁面(圖二)，下載「台中市統計區」(方框標示方框處)的 JSON 資料檔，下載並儲存成「taichung.json」檔。

The screenshot shows a data collection page with the following details:

- 資料集描述:** 各縣市統計區15歲以上人口五歲年齡組教育程度統計(二)統發布區
- 主要欄位說明:** 15歲以上五歲年齡組(15-19歲、20-24歲、25-29歲、30-34歲、35-39歲、40-44歲、45-49歲、50-54歲、55-59歲、60-64歲、65歲以上)與教育程度(研究以上、大學專科、高中職、國中初職、小學及自修、不識字、未詳(只有統計區有))的人口數交叉統計 ※內政部統計處於政府資料開放平臺所提供各縣市最新之統計區人口資料格式為json及xml，使用瀏覽器瀏覽時為一連串之文字數字，一般是可讓程式設計師開發程式進行資料應用而並非是亂碼，若您想要下載的格式為csv檔(可用excel瀏覽)，可參見國土資訊系統社會經濟資料庫分組網頁(segis.moi.gov.tw)的社會經濟資料庫共通平台進行下載。
- 資料資源:** A list of data resources for different counties/cities, each with a 'JSON' icon and a '檢視資料' (View Data) button. The entry for '台中市統計區' is highlighted with a red box and an arrow.

圖二·資料集下載頁面

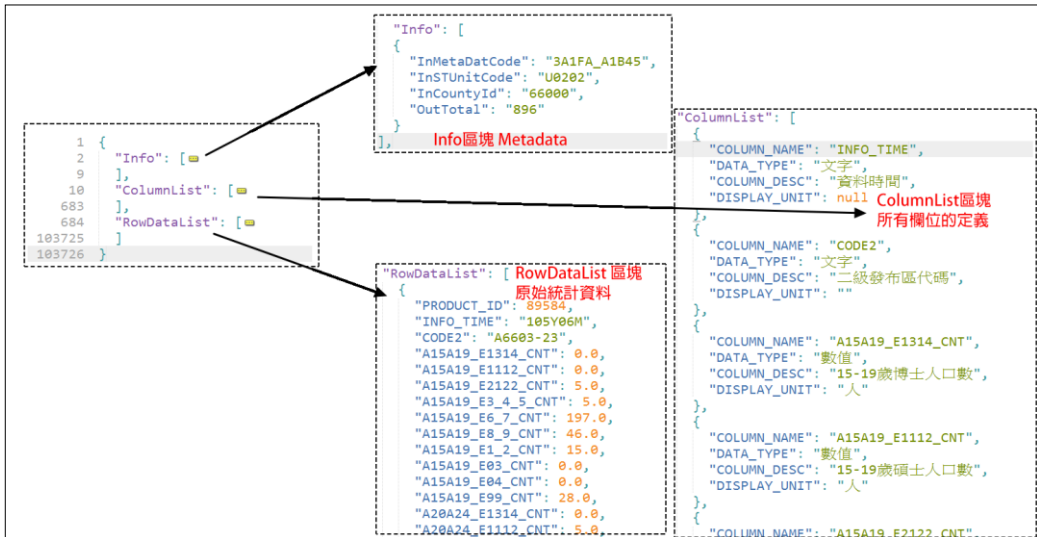
1.2 檢視下載 JSON 檔

打開下載後的原始檔，可發現總行數共有 103,611 行，內容可分為三個區塊，分別為紀錄總筆數、資料序號的「Info」區，紀錄資料型態、欄位英文名稱的「ColumnList」區，以及紀錄統計數值的「RowDataList」區(參考表三及圖三)。圖三節錄各區塊部份內容，點擊可展開區塊內容。以 Info 區塊為例，從「OutTotal」欄位可知總筆數為 896 筆(圖四)。

表三·JSON 區塊說明

JSON 區塊	區塊意義	內容說明
Info	Metadata	開放平台序號、縣市代碼、總筆數等
CplumnList	欄位資料	欄位名稱、資料型態、欄位說明等
RowDataList	統計資料	對應統計數值

此外，從圖三的 ColumnList 區塊可知「Code2」欄位的文字是代表某一個二級發布區、「A15A19_E1314_CNT」欄位的數值是紀錄 15-19 歲博士人口數……依此類推，因此對應到圖五的 RowDataList 區塊後，便可從原始檔案內容看出，編號 AA6603-23 這個二級統計區，其區域內 15-19 歲 (A15A19_E1314_CNT)、20-24 歲 (A20A24_E1314_CNT)、25-29 歲 (A25A29_E1314_CNT)、30-34 歲 (A30A34_E1314_CNT) 年齡層的博士人口數分別為 0、0、2、4 人。



圖三·內文區塊結構

```

"Info": [
{
  "InMetaDatCode": "3A1FA_A1B45",
  "InSTUnitCode": "U0202",
  "InCountyId": "66000",
  "OutTotal": "896"
}
],
    
```

圖四·Info 區塊

圖五「原始統計資料(RowDataList)」的部份，為年齡分組與教育程度分類統計值(二級發布區區內統計)，也是本文後續擷取之目標。

```

{
  "COLUMN_NAME": "A65UP_E04_CNT",
  "DATA_TYPE": "數值",
  "COLUMN_DESC": "65歲以上不識字人口數",
  "DISPLAY_UNIT": "人"
},
{
  "COLUMN_NAME": "A65UP_E99_CNT",
  "DATA_TYPE": "數值",
  "COLUMN_DESC": "65歲以上未詳人口數",
  "DISPLAY_UNIT": "人"
}
],
"RowDataList": [
{
  "PRODUCT_ID": 89584,
  "INFO_TIME": "105Y06M",
  "CODE2": "A6603-23",
  "A15A19_E1314_CNT": 0.0,
  "A15A19_E1112_CNT": 0.0,
  "A15A19_E2122_CNT": 5.0,
  "A15A19_E3_4_5_CNT": 5.0,
  "A15A19_E6_7_CNT": 197.0,
  "A15A19_E8_9_CNT": 46.0,
  "A15A19_E1_2_CNT": 15.0,
  "A15A19_E03_CNT": 0.0,
  "A15A19_E04_CNT": 0.0,
  "A15A19_E99_CNT": 28.0,
  "A20A24_E1314_CNT": 0.0,
  "A20A24_E1112_CNT": 5.0,
  "A20A24_E2122_CNT": 227.0,
  "A20A24_E3_4_5_CNT": 12.0,
  "A20A24_E6_7_CNT": 51.0,
  "A20A24_E8_9_CNT": 2.0,
  "A20A24_E1_2_CNT": 0.0,
  "A20A24_E03_CNT": 0.0,
  "A20A24_E04_CNT": 0.0,
  "A20A24_E99_CNT": 0.0,
  "A25A29_E1314_CNT": 2.0,
  "A25A29_E1112_CNT": 63.0,
  "A25A29_E2122_CNT": 187.0,
  "A25A29_E3_4_5_CNT": 10.0,
  "A25A29_E6_7_CNT": 31.0,
  "A25A29_E8_9_CNT": 10.0,
  "A25A29_E1_2_CNT": 0.0,
  "A25A29_E03_CNT": 0.0,
  "A25A29_E04_CNT": 0.0,
  "A25A29_E99_CNT": 0.0,
  "A30A34_E1314_CNT": 4.0,
  "A30A34_E1112_CNT": 48.0,
  "A30A34_E2122_CNT": 170.0,

```

圖五·年齡分組對照教育程度統計原始資料(部分)截圖

1.3 利用命令列工具 jq 來擷取出所需要區塊的 JSON 文字

先簡介 jq 命令列工具：

[jq 命令列工具說明]

jq 是一種處理 JSON 文字檔的命令列轉換工具。jq 能讀取 JSON 文字檔內容或文字串流(stream)作為輸入，並且能指定過濾(filter)語法，從輸入(input)中過濾出結果，並產生 JSON 文字檔輸出(output)。

接著，下指令處理網路上下載的原始檔 (檔名：taichung.json)，並指定擷取 RowDataList 區塊(用 jq 的過濾器 filter 語法來指定)，然後將之輸出為 JSON 文字檔，檔名為 tai-rows.json。

完整的轉換指令與分解說明如下：

<code>jq</code>	<code>-c</code>	<code>'.</code>	<code>RowDataList</code>	<code> </code>	<code>.'</code>	<code>[</code>	<code>]</code>	<code>'</code>	<code>taichung.json</code>	<code>></code>	<code>tai-rows.json</code>
①			②						③		④

- ① jq 命令與 -c 參數(產生壓縮格式的輸出)，不加-c 參數則不壓縮。
- ② jq filter 語法(完整說明請參照¹⁵) 取出 RowDataList[]的內容。
- ③ 輸入檔 taichung.json。
- ④ 輸出檔 tai-rows.json。

[關於 jq 過濾器更多說明]

jq 過濾器(filter)是透過語法指定需要過濾的內容，單次執行 jq 命令可指定一個以上的過濾器。過濾器如同篩子，篩選輸入，並將符合指定過濾語法的內容輸出；多個過濾器之間用「|」符號區隔，並且前一個 filter 語法的輸出可以為下一個 filter 語法的輸入。更多過濾器的用法請參考 jq 官網文件。

【步驟二：執行 json2csv 轉換】

本步驟將 tai-rows.json 轉換成 csv。csv 格式為統計軟體常用輸入格式，透過這樣的轉換，便利讀者使用統計軟體分析資料。

¹⁵ <https://stedolan.github.io/jq/manual/>。

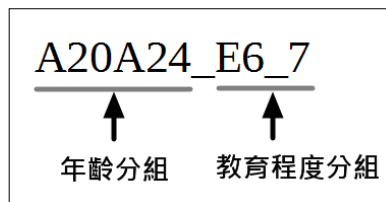
2.1 透過命令列工具 json2csv 來轉換

首先，示範以命令列工具 `json2csv`，將步驟 1.3 產生的 JSON 文字檔內容，輸出成 `csv` 文字格式的方式。由於 `json2csv` 指令必須指定輸出之 JSON 檔案的欄位名稱，表四整理了 JSON 檔案中，用以組成欄位名稱的教育程度和年齡的分組方式及其代碼。

表四·教育程度與年齡分組代碼對照表

組別	代碼	說明	組別	代碼	說明
1	A15A9	15-19 歲	1	E1314	博士
2	A20A24	20-24 歲	2	E1112	碩士
3	A25A29	25-29 歲	3	E2122	大學院校
4	A30A34	30-34 歲	4	E3_4_5	專科
5	A35A39	35-39 歲	5	E6_7	高中職
6	A40A44	40-44 歲	6	E8_9	國中初職
7	A45A49	45-49 歲	7	E1_2	小學
8	A50A54	50-54 歲	8	E03	自修
9	A55A59	55-59 歲	9	E4	不識字
10	A60A64	60-64 歲	10	E99	未詳
11	A65UP	65 歲以上			

完整的欄位名稱規則為「年齡分組_教育程度分組_CNT」，如果了解「20-24 歲年齡層具備高中職教育程度」分區統計值，應該挑選「A20A24_E6_7」欄位，如圖六所示。同理，要找出 25-29 歲大學與專科層級的分區統計值，則可分別選擇「A25A29_E2122_CNT」、「A25A29_E3_4_5_CNT」欄位。



圖六·年齡分組對照教育程度分組

如須一次轉出多個 `csv` 欄位值，在指令中必須使用小寫逗號「,」串接輸出的欄位，並用雙引號(")把全部的欄位包裹在一起。本例中，我們將指定「A20A24_E6_7_CNT、A25A29_E2122_CNT、A25A29_E3_4_5_CNT」三欄位，完整指令與細部說明如下：

```
< tai-rows.json json2csv -p -k "A20A24_E6_7_CNT,A25A29_E2122_CNT,A25A29_E3_4_5_CNT" > result.csv
```

- ① tai-rows.json 為輸入檔
- ② json2csv 指令 (-p 參數不列出 header 列)
- ③ -k 參數列出指定輸出的欄位。輸出欄位用逗號串接，前後加括號(逗號與欄位名稱中間不可加入空白)
- ④ 輸出結果儲存為 result.csv 檔

【輸出欄位補充說明】

如果後續要用地理區域呈現各統計值，您需要指定輸出二級發布區代碼 CODE2。另外，根據資料集網頁上的說明，本資料集更新頻率為不定時，請讀者留意統計時間的欄位 INFO_TIME 以避免統計時間上可能的落差。

以上是透過下指令的方式將 JSON 檔轉換成 csv 文字。但如果不想安裝工具，亦可以採用以網站方式進行線上轉換。

2.2 透過線上網站轉換成 csv

本文介紹兩個提供「線上轉換 JSON→csv 檔」的網站並進行比較，分別是：

- <http://www.convertcsv.com/json-to-csv.htm>
- <https://konklone.io/json/>

由於本範例筆數(896 筆)不多，線上轉換的速度不算太慢，如果原始來源檔案大小超過 1MB 以上，或讀者操作時發現轉換速度過慢，仍建議讀者使用命令列工具來轉換。

2.2.1 透過網站 convertcsv.com 線上轉換成 csv

轉換工具網址：<http://www.convertcsv.com/json-to-csv.htm>

【轉換步驟說明】

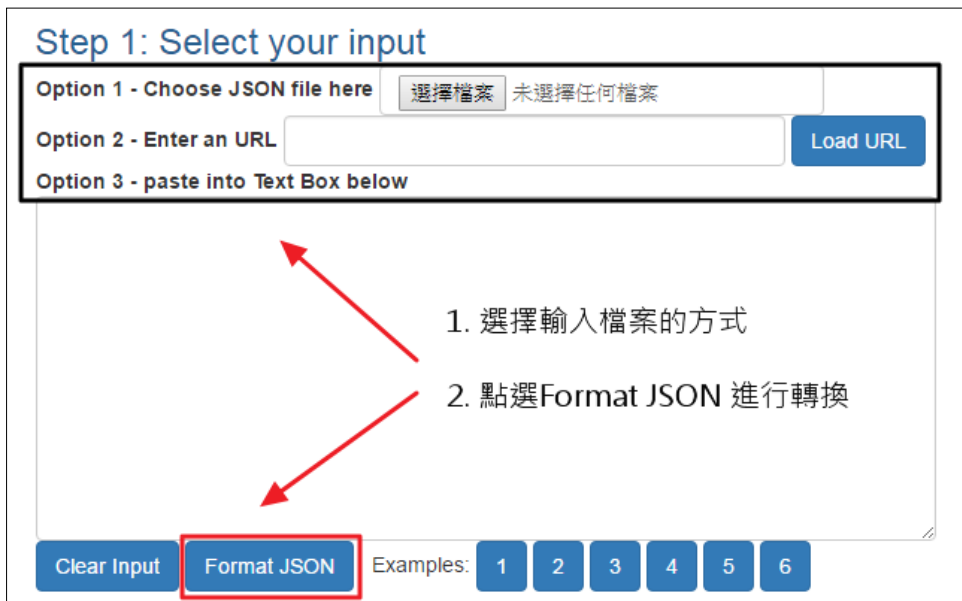
Step 1：首先選擇輸入的方式(Select your input)，一共有三種選擇。讀者可選擇其中一種進行轉換：

1. 上傳檔案(Choose JSON file here)
2. 輸入網址(Enter an URL)
3. 貼上 JSON 文字(Paste into Text Box below)

1.與 3.的選項比較直觀，不多作說明。而選項 2.「輸入網址」指的是輸入一串網址(可容許轉換網站線上直接存取 JSON 文件)，至於能不能轉換出結果，端視原始檔是否能被網站可解析而定，通常是 JSON 文字檔內文的結構越簡單，解析成功的可能性越大。此外，若只是想體驗轉換的操作，網站上提供六個範例資料(Examples 1~6)可直接試用。

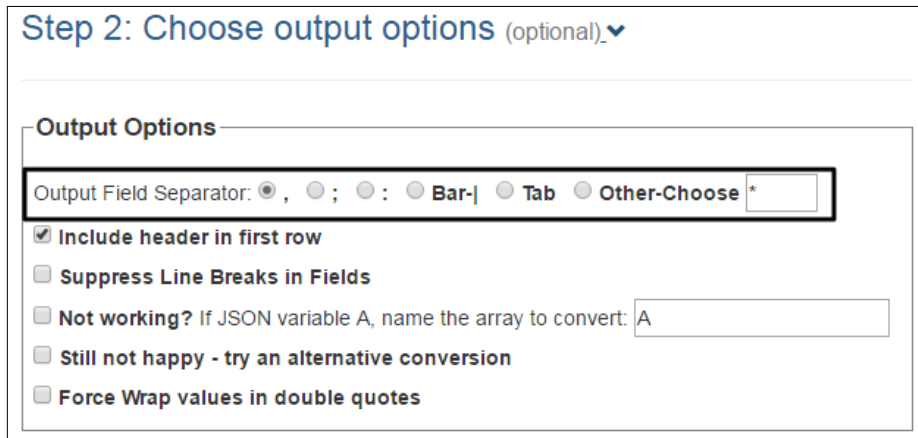
經過實際測試，貼上下載本文範例檔案的網址並按 **Load URL**，雖然可以將 JSON 原始檔案的文數字內容載入到下方的文字框中，但無法被該網站解析。因此改用「上傳檔案」的方式，選擇步驟一處理後的 JSON 檔案後上傳。

圖七·網站 convertcsv.com 輸入頁面



Step 2：調整輸出選項

本步驟為可選項目，可忽略不調整，預設為以逗號分隔、並於第一列列出欄位名稱。傳統上，csv 檔案採用逗號為欄位分隔符號；如果需要挑選其他分隔符號，則可於「Output Field Separator」設定分隔符號為分號(;)、冒號(:)、直線(|)、Tab、或者其他(如星號*)，選項畫面請參考圖八。



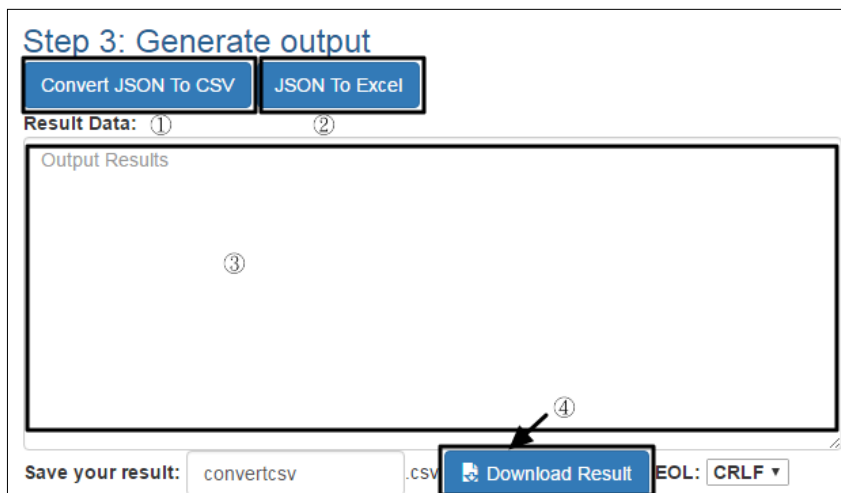
圖八·網站 convertcsv.com 輸出選項頁面

Step 3：將輸出結果另存為 csv 或 Excel 檔案

根據 Step 1 的輸入、Step 2 的輸出參數指定，Step 3 就是輸出結果。

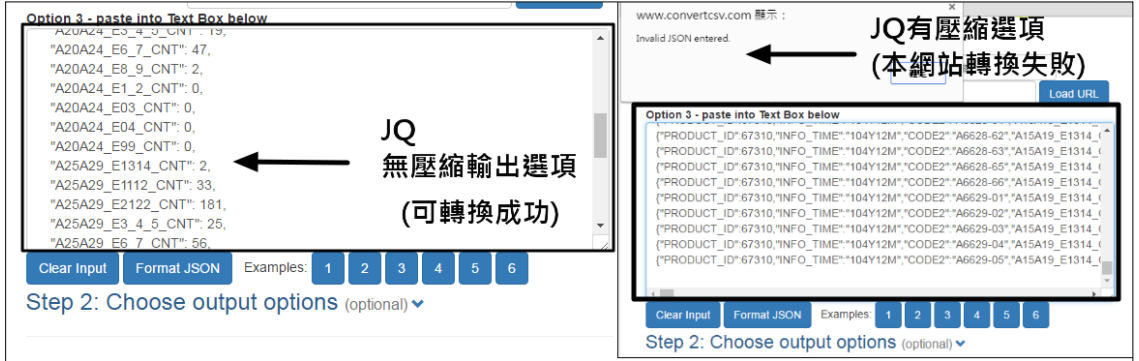
Step 3 的輸出方式有兩種，分別是「Convert JSON to CSV」與「JSON to Excel」，對應到界面(圖九)的① ② 兩個按鈕。如果按下①按鈕，結果會直接出現在下方「Output Results」(下圖③ 區域)。按下② 的結果會是跳出另存新檔的視窗。

如果要將區域③ 的結果輸出，需要按下「Download Result」按鈕(下圖④)另存新檔。可在「Save your result」中指定「另存的檔名」和「換行符號」(預設是 windows 格式 CR+LF)。



圖九·網站 convertcsv.com 輸出頁面

須提醒讀者注意的是，經過測試，該網站只接受無壓縮格式的 JSON 檔案或文字，如圖十所示，若是上傳經過壓縮的 JSON 檔案，會出現「Invalid JSON entered」的訊息，讀者欲利用該網站的話，務必調整步驟一 1.3 的指令參數(取消 -C 參數)。



圖十·網站 convertcsv.com 壓縮格式測試結果

該網站選項雖然比較多元，但整體的操作動線流程設計還有待改善，無法直覺的操作。雖然乍看之下並不覺得複雜，網站設計者也著實花了不少心力將「輸入、輸出、選項、教學範例」整合成單一界面，但對初次使用網站的人來說，須要一些時間測試與理解界面的正確使用方式。

2.2.2 透過網站 konklone.io 線上轉換成 csv

線上工具網址：<https://konklone.io/json/>

接下來介紹的 JSON to csv 資料轉換工具網站「Konklone.io」，收錄於美國白宮所建置名為「Project Open Data」的專案中¹⁶。操作上十分簡單直覺，只有「貼上文字」與「輸出下載」兩個步驟，不容易搞錯。另外，本網站可接受輸入壓縮格式與非壓縮格式 JSON 文字。

16 Project Open Data，網址 <https://project-open-data.cio.gov/>，點選 4.Tools 之 4.13 JSON-to-CSV Converter。

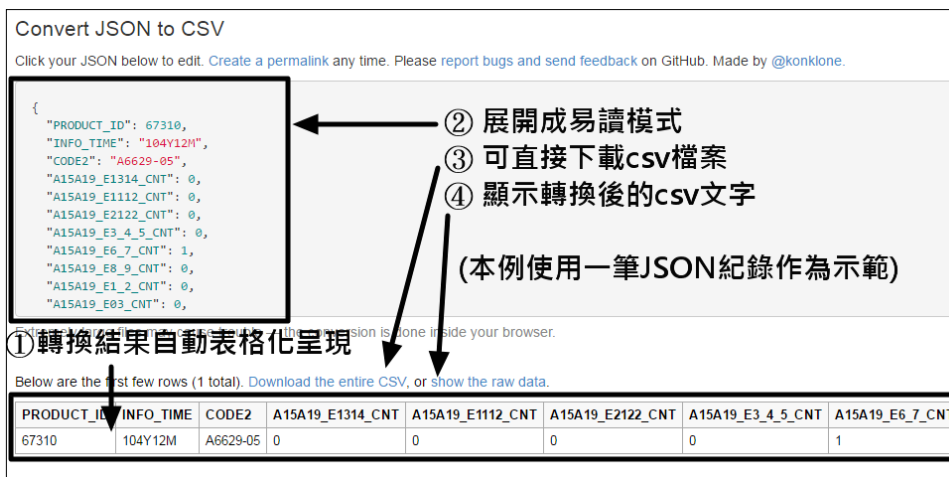
【2.2.2.1 輸入】



圖十一 · 網站 Konklone.io 輸入頁面

於上圖貼上文字後，網頁會自動排版展開成可讀樣式，並將結果輸出在下方，列出產出的筆數，以及提供下載 csv 檔的連結(圖十一為輸入，圖十二為輸出)。相比之下，是不是比前一個線上工具，更容易操作呢？

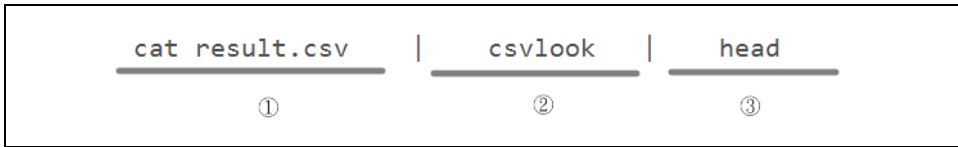
【2.2.2.2 結果與輸出】



圖十二 · 網站 Konklone.io 輸出頁面

【步驟三：透過命令列工具檢視轉換後的 csv】

在步驟三中，筆者使用 csvkit 套件為工具，示範利用命令列下 csvlook 指令的方式來檢視步驟二中轉換好格式的檔案內容。csvkit 還有其他諸如列出欄位名稱(csvcut)、排序(csvsort)、合併資料(csvjoin)等指令可供讀者應用來檢視或處理 csv 格式的檔案。

圖十三 · 用 `csvlook` 命令檢視步驟二輸出結果

指令說明：

- ① 讀取步驟二輸出結果 `result.csv` 內容
- ② 將內容用 `csvlook` 命令呈現，`csv` 檔案內容第一行顯示為欄位名稱
- ③ 用 `head` 命令輸出結果前 10 行，如下圖所示

A20A24_E6_7_CNT	A25A29_E2122_CNT	A25A29_E3_4_5_CNT
47	181	25
46	140	15
23	99	12
39	103	17
39	151	18
46	109	19
30	96	11

圖十四 · `csvlook` 命令輸出結果(前十行)

如此，輸出的欄位如果不對或有缺漏，都可透過步驟二、步驟三等提到的工具重新轉換或確認資料的正確性。此三步驟完結之後，轉出的 `csv` 檔案便可用統計軟體進行後續的分析、處理、運用。

【更多說明】

原始 csv 內容中，分隔符號(逗號)的數量是跟欄位數目有成正比，透過 csvlook 指令直接做編排，容易視覺上看出遺漏的欄位。但螢幕大小是有限的，本例資料轉出的 csv 無法透過畫面一目了然。讀者可將其視為資料正確性檢查的一種輔助。

【結語】

資料使用者在「取用到利用」資料過程時，往往面臨不少的考驗。本文介紹資料工程領域的一些概念與工具，透過觀念的分享與實例演練相關工具，來協助讀者初步體會資料整理、轉換處理過程中會遭遇的問題與克服的方法。

電腦科學領域有一句經典名言「Garbage In, Garbage out」(垃圾進，垃圾出)，說明如果輸入電腦的是一堆無用的資訊，即使電腦仍可按使用者指令運算處理，產出的結果也不具任何意義；相同地，包含政府與民間提供的開放資料、巨量資料等資料的處理利用亦是如此。若要避免因資料的謬誤，造成後續分析、運算推導出之結論可能出現「失之毫釐、差之千里」的遺憾，吾人執行資料處理的各個步驟皆應謹慎為之。

展望未來，資料工程領域仍在持續的進化，有兩個趨勢正在影響它的發展：

- (1)資料科學團隊分工精細化
- (2)聚焦整合資料能力，建構多元化的資料組合供利用

此進階議題的探討，留給有興趣的讀者，自行追蹤最新的發展與變化。

【參考書目】

Janssens, J. (2014). *Data Science at the Command Line*: O'Reilly Media.

Squire, M. (2015). *Clean Data* (pp. 272). Retrieved from <http://shop.oreilly.com/product/9781785284014.do>



會員申請程序說明

一、會員資格

(一) 一般會員(二年期)

1. 國內外公私立研究機構之專任研究人員；
2. 國內外公私立大專院校之專兼任教師；
3. 政府機構之專任研究人員；
4. 捐贈或授權資料予「學術調查研究資料庫」之個人或單位代表。

(二) 一般會員(一年期)

1. 國內外公私立大專院校大學部學生、碩博士班研究生；
2. 國內外公私立研究機構研究助理人員；
3. 國內外公私立大專院校研究助理人員；
4. 政府機構研究助理人員。

若不符合前述一般會員資格，但已取得博士學位者，可以個案申請方式處理。

(三) 院內會員

中央研究院專任研究人員

(四) 網路會員

不限資格之個人

二、資料使用權限

	學術調查資料		政府調查資料		限制性資料	SRDA:Nesstar 線上分析	
	公共版	會員版	抽查	普查		描述性統計	交叉表迴歸分析
一般會員	○	○	○	—	○	○	○
院內會員	○	○	○	○	○	○	○
網路會員	○	—	—	—	—	○	—

三、加入方式

請至 SRDA 網頁(<https://srda.sinica.edu.tw>) 點選「加入會員」，依步驟點選同意會員約定條款同意書中各項規定，填寫申請表提出申請，至個人電子郵件信箱中收取確認信並完成電子郵件確認程序。如需加入成為一般會員或院內會員，則請將在職(在學)文件影本以郵寄、傳真、手機拍照或掃描成電子檔附加於 Email 等方式擲回本中心。

- 審核時間：收到文件後一至三個工作天。
- 諮詢專線：(02)2787-1829；傳真：(02)2787-1802；Skype：csr_209。



徵稿啓事

「『學術調查研究資料庫』通訊」為中央研究院人文社會科學研究中心調查研究專題中心出版以服務學術界之刊物，每年三月、六月、九月及十二月以季刊形式發行。本中心為增加本通訊內容之多樣性及促進與讀者的互動交流，歡迎學界人士能夠提供以下相關議題之文章。

- 一、 二手資料之利用與分析
- 二、 調查資料之整理與檢誤
- 三、 國內外資料庫使用之經驗
- 四、 資料編碼或編碼簿之製作
- 五、 統計軟體之運用

寫作重點請著重於經驗交流與心得建議。請以 e-mail 的方式傳寄給我們，字數以 1,000-1,500 字為限，並以單篇完結之形式寫作。來稿由編輯委員會審閱決定刊登與否，一經採用將稿費酬謝，每千字 750 元。對接受刊登稿件，本刊編輯群有權刪改並置放於「『學術調查研究資料庫』通訊」電子版網站，如不同意者，請於稿件上聲明。文稿內容不代表本刊意見，如涉及著作權或其他文責問題，由作者自負法律責任。文章以真實姓名發表，投稿請附上您的聯絡方法(通訊地址與電話)。

『學術調查研究資料庫』通訊 Survey Research Data Archive Newsletter

出版單位 / 中央研究院人文社會科學研究中心 調查研究專題中心
地 址 / 台北市南港區研究院路二段 128 號
電 話 / (02)27871829 傳 真 / (02)27871802
網 址 / <https://srda.sinica.edu.tw>
電子郵件 / srda@gate.sinica.edu.tw
發 行 人 / 于若蓉
編輯顧問 / 王文心、杜素豪、廖培珊、楊孟麗、謝淑惠
主 編 / 羅婉云
文 編 / 王俞才、王學治、李孟諳、邱亦秀、
林雅琪、陳家玉、蘇婉雯、蘇脩惠
美 編 / 邱亦秀
印 刷 / 永億印刷實業股份有限公司
地 址 / 台北市南港區南港路三段 130 巷 3 弄 1 號 1 樓
中華民國九十一年四月創刊
中華民國一〇五年九月第五十八期

GPN : 2009101543