

Survey Research Data Archive Newsletter

本期目錄

■ 資料釋出消息.....	1
■ R「SURVEY」套件簡介--以 104 年網路沉迷研究為例.....	11
■ 會員申請程序說明.....	25
■ 資料捐贈說明	26
■ 徵稿啓事	27

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12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 9 月至 11 月間改版更新及資料開放訊息如下：

【資料改版】

■ 台灣社會變遷基本調查計畫 2014 第六期第五次

本次改版內容包括：

1. 抽樣分層變項 stratum2，其第六、七層合併為第六層。
2. 花東地區為同一抽樣分層，第一階段(PSU)抽出花蓮與台東，第二階段直接抽出受訪者，故無(SSU)抽出鄉里之編號。更正花蓮縣與台東縣這兩個縣市的 PSU 與 SSU 數值。

■ 「台灣教育長期追蹤資料庫」後續調查：教育與勞力市場的連結

將原為「會員版」的資料改以「公共版」方式開放使用，而「公共版」資料不限下載資格，加入 SRDA 會員完成電子郵件確認，取得網路會員身分後即可下載。

本次改版為「公共版」的調查資料如下：

C00303_1 「『台灣教育長期追蹤資料庫』後續調查：教育與勞力市場的連結」
-2005/2007 高中職五專學生核心樣本 2009 年調查

C00303_2 「『台灣教育長期追蹤資料庫』後續調查：教育與勞力市場的連結」
-2001/2003 年高中職五專學生樣本 2009 年調查

C00303_3 「『台灣教育長期追蹤資料庫』後續調查：教育與勞力市場的連結」 -
2001/2003 年高中職五專學生樣本 2010 年調查

E10259_1 「『台灣教育長期追蹤資料庫』後續調查：教育與勞力市場的連結」
-2005/2007 高中職五專學生核心樣本 2013 年調查

【資料開放：長期主題型調查】

■ 台灣社會變遷基本調查計畫 2015 第七期第一次

計畫主持人：傅仰止、章英華、杜素豪、廖培珊

計畫執行單位：中央研究院社會學研究所

經費補助單位：科技部

計畫執行期間：2015/01/01-2015/12/31

台灣社會變遷基本調查(以下簡稱「變遷調查」)由科技部前身行政院國家科學委員會之人文及社會科學發展處在 1983 年推動，1985 年成功完成第一次全台灣代表性樣本的調查，並從 1990 年第二期計畫開始，每年都進行兩份獨立的問卷調查。研究規劃的主要原則，是以其中一份問卷每間隔五年重覆同樣的調查主題(一共有五個固定主題)，從事貫時性之調查，收集含有兩個時間點以上的資料供比較分析，以突顯社會變遷的趨勢。另外一份問卷則有彈性地採用不同的研究主題。

第七期第一次調查(簡稱七期一次)依例分為兩個主題進行。主題一「綜合問卷」是第一個進行第七波調查的問卷。七期一次的問卷設計中，有些題組進行部分調整，有些則進行較大的調整，對保留的題目，則盡量維持原來的題幹與答項，以利長期比較，共計成功完訪問卷為 2,034 份。主題二「工作與生活」是本計畫執行第二次的主题，首次調查為十年前之五期一次工作與生活組問卷。本年度的工作與生活組問卷參考 2005 年的問卷設計原則，並加入 ISSP 2015 年「工作取向」以及 EASS 2014 年「工作生活」的核心題組，共計成功完訪 2,031 份問卷。

本次同時釋出會員版及限制版資料，限制版資料較會員版增加調查地區村里編號，與受訪者行職業開放題(公司名稱除外)，提供有相關需求的使用者進一步分析使用。

釋出項目計有：資料使用說明、問卷、SPSS 資料檔、STATA 資料檔、調查報告及計畫說明手冊。

■ 台灣教育長期追蹤資料庫後續調查：教育與勞力市場的連結 -2001/2003 年國中學生之非核心樣本 2011-2013 年調查

計畫主持人：關秉寅教授

計畫執行單位：國立政治大學社會學系

計畫執行期間：2009/01/01-2012/12/31；2013/06/01-2017/05/31

調查執行期間：2011/08/09-2011/09/02；2012/03/08-2012/3/23；

2013/10/29-2013/11/12

「『台灣教育長期追蹤資料庫』後續調查：教育與勞力市場的連結」(Taiwan Education Panel Survey and Beyond，簡稱 TEPS-B)是「台灣教育長期追蹤資料庫」(Taiwan Education Panel Survey，簡稱 TEPS)的後續追蹤調查，自 2009 年起繼續追蹤 TEPS 當年接受調查的學生樣本。這些學生目前正值人生進入勞力市場或繼續升學的階段，故 TEPS-B 研究計畫的研究問題與理論分析架構聚焦在高等教育擴張後，大學品質和科系選擇對進入勞力市場後的職業表現，及性別在科系選擇和職場上的表現差異。

TEPS 於 2001 年以台灣地區 90 學年度(2001 年)就讀國中一年級學生為抽樣母群，共抽取約 20,000 名學生為調查對象。此國中學生樣本於 2003 年再度接受調查(TEPS-B 稱之為 Panel 1)，其中又有四千多名於 2005 年及 2007 年由 TEPS 持續兩波追蹤其高中職五專就讀情況(TEPS-B 稱此樣本及其同時接受調查之高中職五專同學為 Panel 2)。曾接受 TEPS 四波調查、橫跨 Panel 1 及 Panel 2 之受訪學生稱之為核心樣本(Core Panel)。此次接受調查之國中學生樣本為曾接受過 TEPS 四波調查的學生的國中同學，即國中學生中非核心樣本，亦即為核心樣本的國中同學(Non-Core Panel)。TEPS-B 自中央研究院調查研究專題中心取得 TEPS 此波追蹤學生樣本名單共 15,794 案(簡稱為 Panel-1 NCP)，完訪樣本數為 6,966 案。此次最終資料釋出樣本數含未完訪者，仍為 15,794 案。電訪問卷內容包括受訪者個人資料、教育經歷及工作經歷等題項。

釋出項目包括：資料使用說明、問卷、SPSS 資料檔、STATA 資料檔、過錄編碼簿與報告書等。

■ 台灣青少年成長歷程研究：國三樣本(J3)第八波，青少年問卷

計畫主持人：伊慶春

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

計畫執行期間：2008/01/01-2008/12/31

此計畫擬由生命歷程的研究架構，探討臺灣青少年進入成年初期的成長過程與生活經驗。其中，家庭、學校與社區為青少年與成年初期之主要生活場域，在此研究中，這三制度間的交互作用被假定為影響青少年成長歷程的主要社會機制。該計畫經由長期性、貫時性的固定樣本追蹤研究設計，完整描繪出青少年至成年初期的轉折機制與成長軌跡。並試圖由教育期望、教育成就、職業發展的機會結構，及社會資本之組成如何運用於各不同成長歷程中等重要議題的討論，提出影響不同成長模式之重要社會機制及其運作方式。

受訪對象的選定，以受訪者國三時(2000 年)學校坐落在臺北市、臺北縣及宜蘭縣的學生為母體，採分層比例隨機抽樣，三個地區的抽樣彼此獨立。第一階段先抽

選學校，依據各地區的都市發展特性區分出不同的層級，依據每個層級的學生人數比例決定該層級應抽取的學校數。第二階段在抽出的每所學校中一年級和三年級各隨機抽取兩個班級，所抽之班級即對全班之學生進行測試。

此波調查為國三學生樣本的第八波調查，相當於碩二階段的學生，於 2008 年 11 月採電訪及郵寄問卷進行追蹤訪問，最後，完成 1,798 筆有效樣本。問卷主題包括就業或就學經驗與家庭、社區生活與家人關係。

本次同時釋出會員版及限制版資料，限制版資料較會員版多提供國中所在地的鄉鎮市區、郵遞區號、大學學校科系名稱等，提供有相關需求的使用者進一步分析使用。

釋出項目計有：資料使用說明、問卷、SPSS 資料檔、STATA 資料檔、過錄編碼簿及訪員手冊。

【資料開放：政府調查資料】

■ 行政院主計總處「104 年家庭收支調查」

家庭收支調查始於民國五十三年，每兩年調查一次，但從民國五十九年起，改為每年調查一次，目前由行政院主計總處及新北市、臺北市、臺中市、臺南市、高雄市等直轄市政府主計處分別按年辦理。

104 年臺灣地區家庭收支調查係以各直轄市及縣市為副母體，在每個副母體內採分層二段隨機抽樣方法抽出受訪戶，樣本數計 16,528 戶。本調查項目包括：一、戶口組成，二、家庭設備，三、住宅概況，四、記載事項，五、經常性收入，六、經常性支出。

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

■ 行政院主計總處「104 年受僱員工薪資調查」

受僱員工薪資調查旨在就工商業場所人力需求面，按月蒐集台灣地區各行業事業單位受僱員工人數、薪資、工時及進退狀況等資料，以明瞭整體勞動市場人力需求以及工時與薪資變動趨勢。

本調查範圍，包含工業及服務業等各大行業，採派員實地訪查、通信調查及網際網路填報等分別進行，其抽樣方法採用「截略分層隨機抽樣法」。並應用統計結果及工業生產、國民所得統計資料編製產量生產力統計、產值生產力統計及多因素

生產力統計。另按年附帶辦理受僱員工動向調查及事業人力僱用狀況等專案調查，俾深入了解各行業廠商人力需求及受僱人員異動狀況，以補按月調查之不足。

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

■ 交通部統計處「遊覽車營運狀況調查」

「遊覽車營運狀況調查」其調查目的主要在蒐集臺閩地區遊覽車之基本資料、行駛里程及能源使用情況、車輛保養及維修費用、車輛營運狀況、配屬人員基本資料、公司基本資料、營運狀況與未來展望及對政府採取之輔導管理措施等之意見，以提供釐訂遊覽車輔導管理政策及發展觀光事業之參考。

本調查以臺閩地區為調查區域範圍，每 2 年調查一次，以遊覽車駕駛人及業者為調查對象，甲表調查單位為遊覽車車輛，中型車 (10~25 人座) 採全查，大型車 (26 人座以上) 以隨機抽樣法抽取；乙表調查單位則為遊覽車業者(公司)，採全查。調查方式採郵寄問卷調查，另輔以電話催收。

本調查共釋出 96、98、100、102 等四個年度的資料，釋出項目計有：資料使用說明、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及報告書(98 年僅有摘要分析)。

■ 交通部統計處「自用小客車使用狀況調查」

「自用小客車使用狀況調查」其調查目的在於蒐集臺閩地區自用小客車車輛之使用、通勤(學)、行駛、支出情形、不同的油價對車輛使用頻率及油費支出，以及駕駛人對各項措施之看法，以提供政府釐訂自用小客車管理措施之依據。調查區域範圍為臺閩地區，包括臺北市、高雄市、臺灣省及福建省。調查對象以自用小客車車主(或使用者)為主，調查單位為每一自用小客車車輛。調查方式採郵寄問卷方式辦理。

本調查共釋出 97、99、101、103 等四個年度的資料，釋出項目計有：資料使用說明、問卷、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及報告書(97 年僅有提要分析與調查方法概述)。

■ 交通部統計處「小客貨車租賃業營運狀況調查」

「小客貨車租賃業營運狀況調查」的目的在於蒐集臺閩地區(包括臺北市、高雄市、臺灣省及福建省)小客貨車租賃業基本資料、經營短租、長租營運狀況、出租車輛行駛、保險、違規及肇事狀況、對相關措施之意見及未來展望等資料。

調查對象為臺閩地區小客貨車租賃業者，採郵寄問卷調查。調查項目包括：
1.公司基本資料；2.公司經營短租業務狀況(租期未滿 1 年)；3.公司經營長租業務狀況(租期 1 年及以上)；4.保險、違規及肇事狀況；5.對管理措施之意見；6.未來展望。

本調查共釋出 97、101 等兩個年度的資料，釋出項目計有：資料使用說明、問卷、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿及報告書等。

■ 勞動部統計處「103 年外籍勞工工作及生活關懷調查」

為了解外籍勞工在臺工作與生活概況、負擔仲介費用情形，以及對法令及申訴管道的認知與運用情形，提供政府擬訂外籍勞工政策及管理制度的參據，於 103 年 6 月以「事業類外籍勞工(製造業、營造業)」及「家庭類外籍看護工」為調查對象，辦理「103 年度外籍勞工工作及生活關懷調查」。

事業類母體資料來源為經核准僱用之製造業、營造業之外籍勞工資料檔，採分層比例抽樣，以行業別、國籍別為分層變數，至少完成 4,200 份有效樣本，採用通信調查方式，另輔以多元管道催收調查表；家庭類母體資料來源為經核准僱用之外籍家庭看護工資料檔，採分層比例抽樣，以縣市別、國籍別為分層變數，至少完成 3,300 份有效樣本，調查方式以中、外語人員進行電話調查，另輔以通信調查。

調查項目包括：基本資料、仲介公司收取費用及服務概況、工作概況、薪資、工時及相關福利，以及對政府提供法令及申訴管道的瞭解及使用狀況等。調查時期自民國 103 年 6 月 1 日起至 7 月 31 日止，計回收有效樣本 7,841 人(事業類外籍勞工 4,451 人，家庭看護工 3,390 人)。

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

■ 勞動部統計處「人力需求調查」

勞動部為了瞭解事業單位短期之人力需求，各細職類人力需求增減及原因，與其他勞工相關議題等，每年進行四次的『人力需求調查』。

調查對象為台灣地區參加勞工保險，且僱用人數為 30 人以上之事業單位，採「分業分層比例隨機抽樣法」，以員工規模為分層變數，以電腦輔助電話訪問(CATI)的方式進行問卷調查，每季大約回收 3,000 份有效樣本。

本調查共釋出 103、104 等兩個年度的資料，釋出項目計有：資料使用說明、問卷、CSV 資料檔、SPSS 資料檔、STATA 資料檔、過錄編碼簿。

【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。

社會學

E10043 職場女性懷孕經驗及相關政策之研究／陳芬苓

E10226 權能激發觀點：受照顧者自我效能之量表發展與驗證／劉素芬

管理學

E10227 外派人員組織承諾對其偏差行為的影響：正式與社會控制機制的干擾效果
／邱雅萍

E10405 以資訊系統成功模式探討醫師使用行動醫療之重要因素／溫信財

E10408 溫泉旅館服務場景對顧客行為意圖影響之多層次分析：個人層級與商店層級變數之中介與干擾角色檢驗／張國謙

區域研究與地理

E99024 咖啡文化在台灣發展脈絡與市場之研究／陳文蓉

E99024			
計畫名稱(中)	咖啡文化在台灣發展脈絡與市場之研究		
計畫主持人	陳文蓉		
計畫執行單位	德霖技術學院餐旅管理系	計畫執行期間	2010/08/01~2011/07/31
計畫摘要	<p>在國人消費能力提升的當下，咖啡飲品的經濟潛力值得期待，然而，台灣發展咖啡文化的經驗不若歐美日等國豐富，仿效他國並不利於自身產業的發展與競爭力，因此，本研究從文化與消費行為觀點，探討咖啡文化在台灣的形成與發展、融入台灣文化後的特色，以及台灣咖啡族的消費行為。本研究採用專家訪談與文獻探討的研究方法，瞭解咖啡文化在台灣發展的脈絡、並與歐美日等國家之咖啡文化進行差異比較分析；採用問卷調查的研究方法，探討咖啡文化對台灣咖啡族消費行為之影響以及區域次文化、消費者特性(包括生活型態、人格特質等)對咖啡文化與消費者行為之間的影響關係。本研究建構之咖啡文化與消費行為與體驗之關係模式，可供後續餐飲觀光休閒領域的學者探討與研究咖啡文化與經營相關議題作參考、提供台灣咖啡業者在遴選、培育與發展等人力資源規劃以及區隔市場策略作參考，以及建議政府幫助業者改善經營績效，培植與發展獨特的台灣咖啡文化，進而提高國家之競爭力與打響台灣的國際知名度。</p>		

E10043			
計畫名稱(中)	職場女性懷孕經驗及相關政策之研究		
計畫主持人	陳芬苓		
計畫執行單位	國立臺北大學社會工作學系	計畫執行期間	2011/08/01~2013/10/31
計畫摘要	<p>本計畫的目的在於探討職場女性的懷孕經驗及所面臨的風險，並分析建構性別友善的政策制度。希望能夠探索女性工作者在懷孕事件之後，個體與社會制度之間的關係轉變，而懷孕婦女如何體認此轉變，如何體認到社會文化中對懷孕婦女工作者的規訓，如何體認到職場中可能的生育風險，如何因應身體改變自己原有的勞動慣習，以及觀察到懷孕對職場女性「美／專業」的威脅；本研究並進一步探討懷孕的負面可能影響。</p> <p>本計畫預計實施二年，第一年研究目的在於瞭解女性學術研究者及醫護產業女性工作者在懷孕過程中的經驗與所面臨的懷孕風險，並且整體性比較國內有工作／無工作之懷孕婦女之流產及職業懷孕婦女工作中斷的因素。第二年的工作是持續瞭解在不同職場中之懷孕女性經驗差異，針對製造業及服務業的懷孕女性，探索其懷孕過程中的經驗及風險差異，並且研究國內外制度如何因應及規範職場婦女懷孕之風險，以及如何為懷孕女性勞工提供友善措施。</p>		

E10226			
計畫名稱(中)	權能激發觀點：受照顧者自我效能之量表發展與驗證		
計畫主持人	劉素芬		
計畫執行單位	國立台北大學社會工作學系	計畫執行期間	2013/10/01~2014/09/30
計畫摘要	<p>此研究目的為，以權能激發為理論觀點，發展一份適用於國內之「受照顧者自我效能量表」。</p> <p>在研究方法方面，此研究採用表面效度、內在一致性效度、因素分析和建構效度等分析進行量表建構，以專家意見和結構性問卷進行調查，有效樣本共 307 人。</p> <p>在研究結果方面，此研究發現「受照顧者自我效能量表」共包括 4 個次向度，共 26 個題項。分析結果顯示，本量表的內在一致性信度達 .866，四個次量表的信度皆良好。本量表的聚合效度和區辨效度均良好。本量表可用於未來受照顧者接受社會生物心理介入服務之評估工具。</p>		

E10227			
計畫名稱(中)	外派人員組織承諾對其偏差行為的影響：正式與社會控制機制的干擾效果		
計畫主持人	邱雅萍		
計畫執行單位	中原大學企業管理學系	計畫執行期間	2013/08/01~2014/09/30
計畫摘要	<p>企業全球化趨勢下，多國企業愈來愈仰賴在不同文化能有效運作的外派人員，藉以管理海外運作、提供技術支援，並成為與母公司聯繫的窗口。外派人員對於企業國際化的成功與否扮演著重要的角色，然而身處於高度不確定的時代，員工偏差行為的例子屢見不鮮。在學術研究上，探討職場偏差行為的研究仍少，有關外派人員偏差行為的研究更是少見，因此國際人力資源管理領域之外派人員偏差行為議題需要更多研究投入。</p> <p>此研究從代理理論觀點，探討外派人員組織承諾與偏差行為關係，並了解母公司控制機制的干擾效果。此研究以台灣企業之外派人員為研究對象，共回收 158 個有效樣本，採多元迴歸分析驗證此研究的觀念架構。實證結果顯示：情感承諾對職場偏差行為具有顯著的負向影響；持續承諾對職場偏差行為具有顯著的正向影響。在母公司的控制機制上，正式控制機制會顯著影響外派人員的組織承諾(情感承諾、持續承諾)與職場偏差行為間之關係。換言之，採行高程度的正式控制機制，會對情感承諾與職場偏差行為(人際偏差、組織偏差)間的負向關係有強化的效果；對持續承諾與職場偏差行為的正向關係有弱化的效果。</p>		

E10405			
計畫名稱(中)	以資訊系統成功模式探討醫師使用行動醫療之重要因素		
計畫主持人	溫信財		
計畫執行單位	臺北醫學大學醫務管理學系	計畫執行期間	2015/08/01~2016/07/31
計畫摘要	<p>隨著行動裝置的普及，愈來愈多醫院導入 iPad 行動醫療資訊系統，當臨床醫師巡視住院病患時，可以不再攜帶厚重的病歷，透過無線網路就能將病歷資料傳到 iPhone 或 iPad 上，即時接收病人最新檢查報告與影像報告，向病患解說病情，縮短醫病距離。</p> <p>由於醫療資訊系統行動化，是未來資訊發展之趨勢，故本研究目的為：了解臨床醫師使用本院行動醫療之系統穩定度、資訊傳輸、介面操作性、服務品質及對行動醫療的使用頻率、成效、接受度及滿意度。本研究透過問卷調查，以北部某區域教學醫院中使用行動醫療資訊系統之醫師為研究對象，共獲得 126 份有效樣本。</p>		

E10408			
計畫名稱(中)	溫泉旅館服務場景對顧客行為意圖影響之多層次分析: 個人層級與商店層級變數之中介與干擾角色檢驗		
計畫主持人	張國謙		
計畫執行單位	致理科技大學休閒遊憩管理系	計畫執行期間	2015/08/01~2016/07/31
計畫摘要	<p>由於顧客體驗與員工行為對於企業營運之成功是密不可分的，但在餐旅研究領域中卻少以服務行銷金三角模型結合 S-O-R(刺激—有機體—反應)模型之觀點來探討顧客體驗與員工行為間之關係，並予以實證性地檢驗。因此，結合個人與公司層次的研究議題，本研究則是以服務行銷金三角模型之觀點為基礎，實證性地檢驗公司層級變數(服務氣候與員工投入)對於個人層級變數關係間(刺激：服務場景—有機體：顧客情感—反應：行為意圖)之跨層次調節影響。本研究以溫泉旅館中之顧客、管理階層及員工作為資料蒐集對象(可避免同源偏差)，且為避免統計分析時受資料巢套的影響，因此以階層線性模式探究變數中的假設關係。</p> <p>研究結果顯示員工行為影響了服務的傳遞，並且進一步澄清為何(公司層級的服務氣候與員工投入之跨層次調節角色)與如何(個人層級的知覺服務場景與顧客情感)間的交互效果顯著地影響了顧客於服務企業中的消費體驗。</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。



R 「survey」 套件簡介-- 以 104 年網路沉迷研究為例

王俞才

一、R 是什麼？

R 軟體是基於 1976 年由 John Chambers 在 AT&T 的貝爾實驗室開發的 S 語言所建構而成的。在 1992 年，現職紐西蘭奧克蘭大學的兩位統計系教授 Ross Ihaka 和 Robert Gentleman(兩位教授的名字字首都是 R，因此 R 也成為了這個軟體的名稱)為了教授初階的統計課程，開始發展 R 軟體，而現階段則是由 R Development Core Team 負責開發。此外，S 語言的發明者 John Chambers 也是 R Development Core Team 的成員之一，因此不難了解 S 語言與 R 軟體之間有著密不可分的關係。

目前的 R 軟體為開放原始碼的自由軟體(Free and Open-Source Software)，提供了包含資料處理、統計計算及繪圖等完整且強大的功能，所有人都可以免費取得 R 軟體的原始碼並加以修改或擴充其功能。此外，R 軟體也提供了多種執行檔(包含 Windows、MacOS 和 Linux 等)方便使用者在不同的平台上執行。

二、如何學習 R？

對於 R 軟體的初學者而言，最大的問題不外乎就是「我該使用哪個函數或套件？」。其實 R 是相當容易學習且使用上也非常簡單的。如圖 1 所示，R 程式撰寫流程基本上依序為確認目標、資料準備、資料前處理，接著尋找合適的套件、函數或自行撰寫程式，最後再將結果彙整。使用者可以藉由 R 軟體網站中 CRAN(Comprehensive R Archive Network)Task Views 尋找合適的套件，也可以使用網站中的 Google 站內搜尋尋找需要的套件或函數。如同本篇文章要介紹的「survey」套件，就可以在 CRAN Task Views 底下的 Social Sciences 分類中找到。此外，若在 R 軟體中使用「??關鍵字」也可以幫助使用者找到相關的函數，而輸入「?函數名稱」或「help(函數名稱)」則可搜尋該函數的定義或範例。

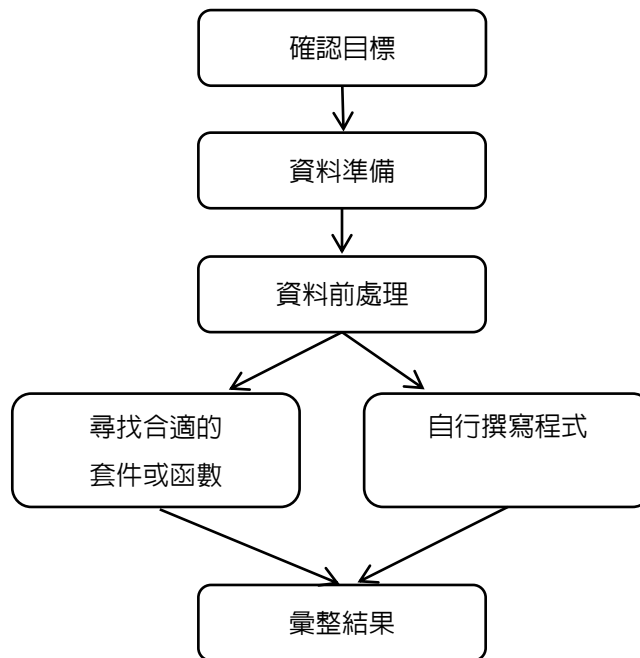


圖 1. 程式撰寫流程圖

三、 範例資料

有鑑於國人上網越來越容易、上網時間越來越長，網路使用風險似乎也有日益嚴重趨勢，國家發展委員會特於「104 年數位機會調查」委外服務案中規畫「網路沉迷研究」子案，主要是為了評估民衆因上網所造成的人際、健康與時間管理等等問題，期望透過深度討論國內網路沉迷的現象，作為政府日後相關政策的參考向。該調查是以臺灣及金馬地區的 22 個縣市為調查範圍，並以年滿 12 歲且有上網經驗的民衆為研究對象進行電話訪問。該調查已於 105 年 2 月 15 日釋出，有興趣的讀者可至 SRDA 學術調查研究資料庫網站申請下載。

本文做為範例的資料檔取自「104 年網路沉迷研究」，經過簡單的資料前處理(如：刪除遺失值、重新編碼...等)後，並假設其為真實「母體」資料。隨後再分別進行三種抽樣設計：簡單隨機抽樣、分層隨機抽樣及群集抽樣作為本篇文章的三種範例資料¹。首先我們就利用一些基本的函數來對母體資料集(pop)做簡單的探索性分析。

¹ 若讀者需要該範例資料檔或程式碼，歡迎來信索取：wuysi@gate.sinica.edu.tw

```

> head(pop) #顯示前 6 筆資料
  gender    age unlimited cluster minute  area
1   女 40-49 歲      是      12   120  南部地區
2   女 40-49 歲     不是       1   300  北部地區
3   女 12-19 歲     不是      12   120  南部地區
4   女 40-49 歲     不是       1    50  北部地區
5   女 40-49 歲      是       8    60  北部地區
6   女 40-49 歲     不是       8   480  北部地區

> tail(pop) #顯示尾 6 筆資料
  gender    age unlimited cluster minute  area
1390  男 60 歲以上      是       1    30  北部地區
1391  男 60 歲以上      是       3   180  北部地區
1392  男 60 歲以上      是      12   120  東部地區
1393  男 60 歲以上     不是       1   180  中部地區
1394  男 60 歲以上     不是       3    90  中部地區
1395  男 60 歲以上     不是       3    30  北部地區

> dim(pop) #取得資料集維度
[1] 1395    6

> names(pop) #取得變數名稱
[1] "gender"  "age"    "unlimited"  "cluster"  "minute"  "area"
    
```

從執行結果可以發現，母體資料集 pop 共有 1395 個觀測值及 6 個變數，且透過 names() 函數可以得知 6 個變數的變數名稱：

gender：請問您的性別是？

age：請問您大約幾歲？

unlimited：請問您採用的是吃到飽方案嗎？

cluster：群集

minute：請問您平均每天上網多久(分鐘)？

area：請問您目前住在哪一個地區？

接著我們可以利用 `summary()` 函數來查看各個變數分佈的情形。對於資料集有初步的了解後，我們就可以接續進行三種不同的抽樣設計。

> summary(pop)			#查看各變數分佈的情形			
gender	age	unlimited	cluster	minute	area	
男:678	12-19 歲 :219	是 :744	Min. :1.000	Min. :2.0	北部地區:647	
女:717	20-29 歲 :275	不是:651	1st Qu.: 1.000	1st Qu. : 90.0	中部地區:350	
	30-39 歲 :337		Median :5.000	Median : 180.0	南部地區:351	
	40-49 歲 :349		Mean :5.351	Mean : 209.7	東部地區: 35	
	50-59 歲 :174		3rd Qu.: 9.000	3rd Qu. : 300.0	離島地區: 12	
	60 歲以上: 41		Max. :14.000	Max. :1200.0		

1. 簡單隨機抽樣：

這部分我們將採用 `sample()` 函數來進行簡單隨機抽樣，該函數也是在 R 軟體中最基本的抽樣函數，其基本格式為：

`sample(x, size, replace=FALSE, prob=NULL)`

其中 `x` 為一正整數或含有一個或多個元素的向量；`size` 為一正整數，代表想要被抽出的樣本數；`replace` 詢問是否為抽出可放回的抽樣；`prob` 用於設定各樣本被抽出的機率。

接著我們就利用 `sample` 函數來執行簡單隨機抽樣，並設定抽出的樣本數為 200。同時為了讓讀者可以跟我得到相同的抽樣結果，在執行任何抽樣過程之前，我先利用了 `set.seed()` 函數來建立隨機起始點。

> set.seed(1)			#建立隨機起始點			
> srs<-pop[sample(1:nrow(pop), 200),]			#執行簡單隨機抽樣			
> head(srs)			#顯示前 6 筆資料			
gender	age	unlimited	cluster	minute	area	
371	男 40-49 歲	是	9	390	北部地區	
519	男 12-19 歲	不是	5	240	北部地區	
798	女 20-29 歲	是	12	300	北部地區	
1265	男 20-29 歲	是	9	60	南部地區	
281	男 40-49 歲	是	1	120	北部地區	
1249	男 20-29 歲	不是	1	240	南部地區	

2. 分層隨機抽樣：

分層隨機抽樣是將母體依照某衡量標準，區分成若干個不重複的子母體，我們稱之為『層』，且層與層之間有很大的變異性，而層內的變異性較小。在區分不同層後，再從每一層中利用簡單隨機抽樣抽出所須比例的樣本數，最後將所得各層樣本合起來即為樣本。利用分層隨機抽樣可保持樣本資料與母體分佈的一致性，在分析資料時也可以減少資料不平衡的問題。

分層抽樣我們可以透過 `sampling` 套件中的 `strata()` 函數來實現，其基本格式為：

```
strata( data , stratanames=NULL , size , method=c("srswor" , "srswr" ,  
"poisson" , "systematic") , pik , description=FALSE )
```

其中 `data` 為待被抽樣的資料集；`stratanames` 為將被作為分層依據的變數名稱；`size` 用於設定各分層中將要被抽出的樣本數，該值的順序必須與該變數中各水準出現的順序一致，且必須將資料集按照該變數的水準進行升冪排列；`method` 用於選擇 4 種抽樣方法，分別為隨機抽出不放回(`srswor`)、隨機抽出放回(`srswr`)、卜松(`poisson`)及系統抽樣(`systematic`)，預設為隨機抽出不放回；`pik` 用於設定各分層被抽出的機率；`description` 參數則是詢問是否輸出包含各分層基本資訊的抽樣結果。

這邊我們將利用資料集中的變數 `area` 作為分層的依據，並設定樣本數為 200，而各層應抽出的人數則依「比例配置法」(Proportional allocation)決定。其執行過程如下：

```
> round(prop.table(table(pop$area))*200)                                #查看各層應抽出的人數
北部地區  中部地區  南部地區  東部地區  離島地區
      93      50      50       5       2

> poporder<-pop[order(pop$area) , ]                                    #將資料集依變數 area 的水準排序
> install.packages("sampling")                                         #安裝 sampling 套件
> library("sampling")                                                  #載入 sampling 套件
> set.seed(1)                                                           #建立隨機起始點
> stratID<-strata(poporder ,                                          #執行分層隨機抽樣
                 stratanames="area" ,
                 size=c(93 , 50 , 50 , 5 , 2) ,
                 method="srswor" ,
                 description= TRUE)
```



```

Stratum 1
Population total and number of selected units: 647 93
Stratum 2
Population total and number of selected units: 350 50
Stratum 3
Population total and number of selected units: 351 50
Stratum 4
Population total and number of selected units: 35 5
Stratum 5
Population total and number of selected units: 12 2
Number of strata 5
Total number of selected units 200
> head(stratID) #顯示前 6 筆資料
      area  ID_unit  Prob  Stratum
9  北部地區    9  0.1437403    1
15 北部地區   15  0.1437403    1
33 北部地區   33  0.1437403    1
40 北部地區   40  0.1437403    1
42 北部地區   42  0.1437403    1
49 北部地區   49  0.1437403    1
> strat<-getdata(poporder , stratID$ID_unit) #取得分層抽樣所得資料
> head(strat) #顯示前 6 筆資料
      gender  age  unlimited  cluster  minute  area
11  女 40-49 歲    是         3    180 北部地區
17  女 40-49 歲   不是        14    480 北部地區
35  女 40-49 歲    是         3    840 北部地區
42  女 40-49 歲    是         4    120 北部地區
44  女 40-49 歲    是         8     60 北部地區
119 女 40-49 歲    是        12    240 北部地區

```

3. 群集抽樣：

群集抽樣的方法就是將母體分成幾個群集(或部落、區域)，再從這幾個群集中抽出數個群集進行抽樣或普查。有時群集抽樣又稱部落抽樣或叢聚抽樣。在考慮使用群集抽樣時，一般會要求各群集對資料整體有較好的代表性，即群集間的變異小，而群集內的變異大。因此當群與群之間差距較大時，群集抽樣常常會出現分佈不廣或樣本代表性較差等缺點。

我們可以透過 `sampling` 套件中的 `cluster()` 函數來執行群集抽樣，其基本格式為：

```
cluster( data, clustername, size, method=c("srswor", "srswr", "poisson",  
"systematic"), pik, description=FALSE )
```

該函數的參數除了 `clustername` 與 `size` 略有差異外，其餘參數的涵義都跟 `strata()` 函數相同。`clustername`，顧名思義，指用來劃分群組的變數名稱。而 `size` 為一正整數，代表欲被抽出的群集數。

以下我們將利用資料集中的變數 `cluster2` 作為分群的依據，隨機抽出 3 個群集。其抽樣過程如下：

```
> set.seed(1) #建立隨機起始點
> cluID<-cluster(pop, #執行群集抽樣
  clustername="cluster",
  size=3,
  method="srswor",
  description=TRUE)
Number of selected clusters: 3
Population total and number of selected units 1395 230
> head(cluID) #顯示前 6 筆資料
```

	Cluster	ID_unit	Prob
1	4	529	0.2142857
2	4	690	0.2142857
3	4	78	0.2142857
4	4	936	0.2142857
5	4	596	0.2142857
6	4	799	0.2142857

² 變數 `cluster` 為作者以隨機亂數生成，僅供範例說明使用，並未考量其分佈是否恰當或樣本代表的好壞。

```

> cluster<-getdata(pop,cluID$ID_unit) #取得群集抽樣所得資料
> head(cluster) #顯示前 6 筆資料

```

	gender	age	unlimited	cluster	minute	area
529	女	30-39 歲	是	4	480	北部地區
690	女	30-39 歲	是	4	360	南部地區
78	女	40-49 歲	不是	4	60	南部地區
936	女	30-39 歲	不是	4	120	南部地區
596	女	30-39 歲	不是	4	420	中部地區
799	女	20-29 歲	不是	4	120	北部地區

四、「survey」套件

survey 套件是由現職紐西蘭奧克蘭大學生物統計學教授 Thomas Lumley 所撰寫及負責維護更新，主要的功能在處理複雜的調查資料分析。survey 套件的功能相當強大，包含了最基本的求取平均數、總和、比值到迴歸模型的建立及主成分分析，甚至是資料視覺化...等功能，對於從事有關調查資料分析工作的研究人員或學生而言有相當大的幫助。此外，survey 套件有一個專屬的網站(圖 2)，其網址為 <http://r-survey.r-forge.r-project.org/survey/>，該網站不只包含了套件內容的簡介，還有許多有關 R 軟體的基礎教學。

Survey analysis in R

This is the homepage for the `survey` package, which provides facilities in R for analyzing data from complex surveys. The current version is 3.29. A much earlier version (2.2) was published in *Journal of Statistical Software*.

An experimental package for very large surveys such as the American Community Survey can be found [here](#).

A port of a much older version of the survey package (version 3.6-8) to S-PLUS 8.0 is available from CSAN (thanks to Patrick Aboiyom at laushtefit).

Features:

- Means, totals, ratios, quantiles, contingency tables, regression models, loglinear models, survival curves, snk tests, for the whole sample and for domains.
- Variances by Taylor linearization or by replicate weights (BRR, jackknife, bootstrap, multistage bootstrap, or user-supplied).
- Multistage sampling with or without replacement.
- PPS sampling with or without replacement. Horvitz-Thompson and Yates-Grundy estimates and a range of approximations.
- Post-stratification, generalized raking, raking, GREG estimators, trimming of weights.
- Two-phase designs. Estimated weights for augmented IPW estimators.
- Graphics.
- Support for using multiply imputed data.
- Database-backed design objects for large data sets (now with replicate weights, too).
- Some support for parallel processing on multicore computers.
- Multivariate analysis: principal components, factor analysis (experimental).
- Likelihood ratio (Rao-Scott) tests for glm, Cox models, loglinear models.

The `NEWS` file gives a history of features and bug fixes.

Comparison shopping
Alex Zaslavsky keeps a comprehensive list of [survey analysis software](#) for the ASA Section on Survey Research Methods.

User-generated ratings and reviews of this package (and others) at [cran.r-project.org](#).

Using the survey package:

- Specifying a survey design
- Creating replicate weights
- Sample summary statistics
- Using supplied replicate weights
- Domain (subpopulation) estimation
- Tables of summary statistics
- Post-stratification and calibration
- Linearly PSUs
- Regression models
- Tests of association
- Stratification within PSUs
- Graphics
- Multiple comparison and ordinal logistic regression
- Database-backed survey objects
- Programming with survey objects

Technical notes and comparisons with other software
Some examples (in PDF) translated from Stata and SUDAAN examples at [UCLA Academic Technology Services](#).

圖 2. SURVEY 套件網站

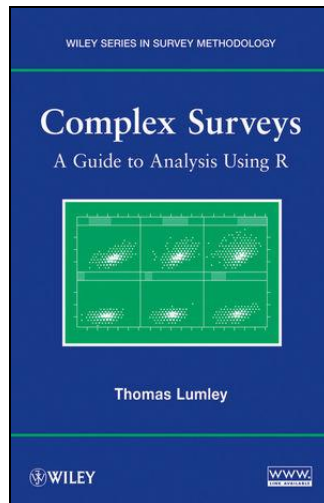


圖 3. Complex Surveys: A Guide to Analysis Using R

另外，Thomas Lumley 也撰寫了一本書叫「Complex Surveys: A Guide to Analysis Using R」(圖 3)，這本書也是以 survey 套件為基礎，讀者都可以透過網站或書籍對 survey 套件有更進一步的認識。

在使用 survey 套件時最主要有兩個步驟要執行：第一個是我們必須提供資料集名稱及抽樣設計(含抽樣方法、母體大小、權數……)的資訊以建立 survey.design 物件；第二個步驟則是以先前建立的 survey.design 物件，再利用指定的函數來執行統計分析、建立模型或繪圖。而這些指定的函數名稱都是以 svy*為開頭，例如：svymean()、svytotal()、svyratio()、svychisq()、svyglm()等。

在 survey 套件中 svydesign()即是用來建立 survey.design 物件的函數，它包含了許多可以設定的參數，在此我們僅說明最重要的幾個參數：

ids：為資料集中的一個變數名稱。當採用群集抽樣時，需給定每個樣本所屬的群集。~1 或~0 表示沒有分群。

strata：作為分層依據的變數名稱。NULL 表示沒有分層。

fpc：為一向量，給定每個樣本所屬層別的樣本數大小。

data：為一包含原始數據的資料集。

下面我們將以先前利用簡單隨機抽樣、分層隨機抽樣及群集抽樣等三種抽樣方式所抽出的樣本對 survey 套件及其函數做簡單的介紹，推估國人每天平均的上網時間。

1. 簡單隨機抽樣：

在說明 survey 套件的使用方法之前，我們先以簡單隨機抽樣樣本中的變數 minute，並用土法煉鋼的方式來計算國人每天上網時間的平均數及其變異數的估計量，其公式如下：

$$\hat{\mu}_x = \frac{1}{n} \sum_{i=1}^n X_i = \frac{1}{200} (390 + 240 + 300 + \dots + 180) = 199.75 \quad \dots\dots\dots(a)$$

$$\widehat{Var}[\hat{\mu}_x] = \frac{N-n}{N} \times \frac{\widehat{Var}[x]}{n} = \frac{1395-200}{1395} \times \frac{30980.59}{200} = 132.6946 \quad \dots\dots\dots(b)$$

其中，

$$\widehat{Var}[x] = \sum_{i=1}^n \frac{(X_i - \bar{x})^2}{n-1} = \frac{[(390 - 199.75)^2 + \dots + (180 - 199.75)^2]}{200 - 1} = 30980.59$$

由上面的計算結果可以得知國人每天上網的平均時間為 199.75 分鐘，且變異數的估計量為 132.69 分鐘。接著我們就使用 survey 套件來實現此一過程，第一個步驟為建立 survey.design 物件：

```
> install.packages("survey") #安裝 survey 套件
> library("survey") #載入 survey 套件
> des.srs<-svydesign( ids=~1 , #建立 survey.design 物件
                    strsta=NULL ,
                    fpc=rep(1395 , 200) ,
                    data=srs)
> summary(des.srs) #查看 des.srs 的摘要
Independent Sampling design
svydesign(ids = ~1 , strsta = NULL , fpc = rep(1395 , 200) , data = srs)
Probabilities:
  Min. 1st Qu. Median Mean 3rd Qu. Max.
0.1434 0.1434 0.1434 0.1434 0.1434 0.1434
Population size (PSUs): 1395
Data variables:
[1] "gender" "age" "unlimited" "cluster" "minute" "area"
```

透過 `summary()` 函數我們可以得知 `survey.design` 物件(`des.srs`)為 Independent Sampling design，代表這個抽樣設計為簡單隨機抽樣，而且當我們透過 `fpc` 參數指定母體大小時，這個簡單隨機抽樣會被認定是抽出不放回的。一旦我們完成建立 `survey.design` 物件之後，我們就可以透過 `survey` 套件中許多 `svy*` 的函數來做資料的統計分析。

```
> svymean(~minute, design=des.srs) #求取上網時間的平均數
```

	mean	SE
minute	199.75	11.519

透過 `svymean()` 函數，我們可以得到國人每天上網時間的平均數、以及平均數的標準差分別為 199.75 分鐘和 11.519 分鐘。這個結果與先前公式(a)與公式(b) ($\sqrt{132.694} = 11.519$) 所得到結果一樣。`survey` 套件中大部分的 `svy*` 函數都跟 `svymean` 一樣，以方程式(`~variable name`)的寫法去選擇需要被分析的變數。他跟你熟悉的 `lm()` 函數的寫法有點類似，但他的變數通常只放在「~」的一邊。接著我們可以以中央極限定理(CLT)為基礎，用手或是利用 `confint()` 函數來計算平均數的信賴區間。

```
> confint(svymean(~minute, design=des.srs)) #求取平均數 95%的信賴區間
```

	2.5 %	97.5 %
minute	177.1726	222.3274

假設現在我們想了解在不同性別下國人每天的平均上網時間，則我們可以利用 `svyby()` 函數來幫我們實現：

```
> svyby(~minute, by=~gender, design = des.srs, FUN=svymean) #求取不同性別下的平均上網時間
```

	gender	minute	se
男	男	181.4216	13.54838
女	女	218.8265	18.64260

當我們只想知道男生(或特定族群)的特性時，那麼我們可以利用 `subset()` 函數來實現：

```
> svymean(~minute, design=subset(des.srs, gender=="男")) #求取男生的平均上網時間
```

	mean	SE
minute	181.42	13.548

2. 分層隨機抽樣：

相較於簡單隨機抽樣，建立分層隨機抽樣的 survey.design 物件是需要多一些程序的。首先我們必須從母體資料集(pop)找出每個分層的母體樣本數，並將它建立成一個新的物件：

```
> class.table<-as.data.frame(table(pop$area))           #找出每個分層的樣本數，並建立資料集
> class.table
      Var1  Freq
1  北部地區  647
2  中部地區  350
3  南部地區  351
4  東部地區   35
5  離島地區   12
```

接著我們必須按照分層隨機抽樣樣本(strat)中每個樣本所屬的分層將 class.table 中的各分層人數匯入，匯入的方法可以利用 merge()函數來執行。merge()函數可以簡單且快速地将兩個擁有相同變數名稱的資料作合併。但需注意的是，由於變數 area 已套用中文的選項說明，因此資料合併後的排列順序會依照各水準字首的筆劃多寡，按遞增方式呈現(「中」部地區<「北」部地區<「東」部地區<「南」部地區<「離」島地區)。

```
> names(class.table)[1]<-"area"                       #將 class.table 的第一個變數名稱更改為 area
> strat<-merge(strat, class.table, by="area")         #將 strat 與 class.table 資料集做合併
> head(strat)                                         #顯示前 6 筆資料
      area gender  age  unlimited  cluster  minute  Freq
1  中部地區  女  40-49 歲    不是      5     240  350
2  中部地區  女  40-49 歲    不是      9      60  350
3  中部地區  女  40-49 歲     是       1      60  350
4  中部地區  女  40-49 歲    不是      1     180  350
5  中部地區  女  40-49 歲    不是      1     240  350
6  中部地區  女  12-19 歲     是       1      60  350
```

此處的分層隨機抽樣樣本多了一個變數 Freq，於是我們就可以開始建立屬於分層隨機抽樣的 survey.design 物件，他跟簡單隨機抽樣的 survey.design 物件只有兩個地方略有差異：一個是在參數 strata 需要指定分層依據的變數名稱，也就是變數 area；另一個則是參數 fpc 需利用變數 Freq 來給定每個樣本所屬層別的樣本數大小。

```

> des.strat<-svydesign(ids=~1, strata=~area, fpc=~Freq, data=strat) #建立 survey.design
物件
> summary(des.strat) #查看 des.strat 的摘要
Stratified Independent Sampling design
svydesign(ids = ~1, strata = ~area, fpc = ~Freq, data = strat)
Probabilities:
  Min. 1st Qu. Median Mean 3rd Qu. Max.
0.1425 0.1428 0.1429 0.1434 0.1437 0.1667
Stratum Sizes:
      北部地區 中部地區 南部地區 東部地區 離島地區
obs          93      50      50      5      2
design.PSU    93      50      50      5      2
actual.PSU   93      50      50      5      2
Population stratum sizes (PSUs):
中部地區 北部地區 東部地區 南部地區 離島地區
  350      647      35      351      12
Data variables:
[1] "area" "gender" "age" "unlimited" "cluster" "minute" "Freq"

```

接下來就可以跟簡單隨機抽樣一樣利用 `svy*` 函數做一連串的統計分析，這裡我們就不再贅述。

```

svymean(~minute, design=des.strat) #求取上網時間的平均數
      mean      SE
minute 225    11.634

```

3. 群集抽樣：

群集抽樣樣本(cluster)是從 14 個群集當中採用一階段群集抽樣(One-Stage Cluster)方式隨機抽取 3 個群集而組成的樣本。群集抽樣的 `survey.design` 物件跟簡單隨機抽樣不同之處在於變數 `ids`，此處我們必須給定每個樣本所屬的群集，也就是變數 `cluster`。另外參數 `fpc` 需要給定母體的群集個數。


```

> des.clu<-svydesign(ids=~cluster, strata=NULL, fpc=rep(14, nrow(cluster)), data=cluster)
#建立 survey.design 物件
> summary(des.clu)
#查看 des.clu 的摘要
1 - level Cluster Sampling design
With (3) clusters.
svydesign(ids = ~cluster, strata = NULL, fpc = rep(14, nrow(cluster)),
  data = cluster)
Probabilities:
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 0.2143  0.2143  0.2143  0.2143  0.2143  0.2143
Population size (PSUs): 14
Data variables:
[1] "gender"  "age"     "unlimited" "cluster" "minute"  "area"
> svymeans(~minute, design=des.clu)
#求取上網時間的平均數
      mean      SE
minute 251.65  5.1541

```

以上是針對 survey 套件所作粗淺的介紹，當然讀者也可以透過 CRAN Task Views 選擇其他合適的套件加以運用。但如同先前所言，survey 套件所提供的函數是相當完善的，且擁有專屬的網站提供解說及教學。不僅如此，假設你對於套件有任何疑問，可以隨時寫 E-mail 給套件的作者，若套件的作者發現有任何缺失，也會立即修正錯誤或擴充套件的功能。當然這就是其他統計軟體所無法取代的優點之一。

五、資料引用

國家發展委員會資訊管理處(2016)。104 年網路沈迷研究 (AE120001)【原始數據】。取自中央研究院人文社會科學研究中心調查研究專題中心學術調查研究資料庫 doi:10.6141/TW-SRDA-AE120001-1。

六、參考書籍

Thomas Lumley(2010). Complex Surveys A guide to Analysis Using R. WILEY(U.S.A)

陳景祥(2010)，”R 軟體：應用統計方法”，台北市：東華書局。

黃文、王正林(2016)，”利用 R 語言打通大數據的經脈(第二版)”，台北市：佳魁資訊。



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、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。

三、資格認證

(一)一般會員：依照申請人的在職或在學身分，期限分為一年及兩年。會員效期到期之前，本資料庫將主動與會員聯絡資料校正及資格延長事宜。會員須於校正期間內確認有無資料異動情形，否則暫停其一般會員權限。

(二)院內會員：申請人須於資料庫網站填具申請表，提供中央研究院電子郵件信箱後並經確認，再提供中央研究院服務證影本向本資料庫提出申請。本資料庫於每年 11 月以電子郵件校正資料，離職後會員權限失效。

(三)網路會員：申請人填寫申請表並完成電子郵件確認程序後即為網路會員；一般會員及院內會員若屆期末進行身分校正，亦轉為網路會員。

四、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	○	—	—	—	—	○	—

五、會員義務

SRDA 在努力提供更好服務的同時，亦需要了解會員們的意見。因此有一些需履行之義務項目，請會員配合實行。

- (一)資料使用：敬請遵守「會員約定條款」的各項規定。
- (二)書目提供：會員在下載資料使用後，發表相關篇章論著等，敬請提供書目資料，嘉惠他人作為研究方向的輔助參考之用。若能提供作品全文亦非常歡迎。
- (三)意見調查：針對會員一年來的使用情形，本資料庫將進行例行性的年度整體滿意度調查，以作為資料庫發展及改善的參考。

資料捐贈說明

SRDA 亦歡迎學者及各界捐贈資料！為使這些得來不易的珍貴資料發揮最大效用，本資料庫工作人員會謹慎處理每筆捐贈資料，並負起妥善保管之責。我們蒐集的資料主要是以調查方法進行的量化資料，公開釋出前，我們會先瞭解該筆資料之智慧財產權歸屬，並取得所有權人/機構同意授權本資料庫公開學術使用。

本資料庫對資料捐贈者提供以下服務項目：

1. 定期／不定期寄贈「學術調查研究資料庫」出版品；
2. 資料整理與檢誤相關諮詢服務；
3. 提供其捐贈資料之申請情況；
4. 提供資料使用者利用其捐贈資料發表之衍生著作。

若您願意捐贈資料給「SRDA 學術調查研究資料庫」蒐藏並對外開放，或有任何疑問，歡迎洽詢「中央研究院人社中心調查研究專題中心」資料組，我們將竭誠為您服務。電話：(02)27871825，資料組蘇脩惠小姐。



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
中華民國一〇五年十二月第五十九期

GPN : 2009101543