

次級資料分析與 R 軟體之應用

Secondary Data Analysis and the Application of R

2018 (107) 第二學期

授課教師：陳語婕 (email: chenyl@yahoo.com)

上課時間地點：星期四 234 [ST019]

Office Hours: By appointment

課程簡介

次級資料分析 (Secondary data analysis) 是使用既有的統計資料來對研究主題進行實證分析的社會研究方法，而既有的統計資料包含人口普查及各種大型學術資料庫，例如台灣社會變遷基本調查計畫。相對於蒐集初級資料 (Primary data) 而進行的研究，使用次級資料的優勢為：(1) 研究者為取得大型樣本的資料所需花費的時間及金錢成本較低；(2) 資料庫的蒐集通常已採用隨機抽樣原則，分析該類資料的研究成果將能推論至整體社會；(3) 資料庫的蒐集通常為跨時間及跨區域，可用來檢驗跨時變化或跨域比較的研究主題。隨著網路介面的發達及電腦運算速度的提升，對於可公開下載的次級資料庫的運用更是日益方便，次級資料分析便成為當前常見的社會學研究方法之一。

目前社會科學常使用的統計軟體相當多元，但通常以使用者付費的為主。R 是一個程式語言且為少數免付費、開放原始碼 (open-source) 及跨平台的統計運算及繪圖軟體，它可在 Microsoft Windows、Mac OS 及多數 UNIX 平台上運作。它的操作介面雖然完全以程式語言控制，但學習門檻低，可以作為學習程式語言邏輯的入門，且其繪製的圖表常能見於學術期刊上。R 的擴充性強，有各種套件 (packages) 及說明檔案可供下載使用來進行統計分析；資料讀取也與其他統計軟體的格式相容，可讀取的格式包括純文字檔、Microsoft Excel、SAS、SPSS、Stata 等。

這門針對大學部全年級的選修課程有下列學習目標：

- (1) 增強學生的實務分析經驗，延伸社會統計及社會研究方法兩門必修課之所學；
- (2) 培養學生搜尋、整合及分析次級資料的能力；

- (3) 通過學習 R 的程式語言及其應用，了解電腦程式語言的邏輯，為日後進修或就業奠基。

課程教材

- (1) 吳明隆，2015，《R 統計軟體應用分析實務》。台北：五南。
- (2) Jared P. Lander (鍾振蔚譯)，2015，《R 軟體資料分析基礎與應用》。台北：旗標。
- (3) Liahna E. Gordon, 2016, "Real research; research methods sociology students can use". Thousand Oaks: Sage.

參考書目及網路資源

- (1) Pierre Lafaye de Micheaux, Remy Drouilhet, Benoit Liquet, 2013, "The R software; fundamentals of programming and statistical analysis", New York, NY: Springer New York.
- (2) David W. Stewart、Michael A. Kamins (董旭英、黃儀娟譯)，2000，《次級資料研究法》。台北：弘智文化。
- (3) R 軟體 (<https://cran.rstudio.com/>)。
- (4) R 軟體及程式編輯介面 (<http://www.rstudio.com/products/rstudio/download/>)。
- (5) 互動式 R 學習套件 swirl (<https://swirlstats.com/>)。

課程評量

本課程評量主要依據作業報告、考試成績及參與表現：

- (1) 平日小組作業 (15%)：包含隨堂電腦軟體操作及輸出。
- (2) 期中期末小組口頭與書面報告 (30%、35%)：包含關於 R 軟體的操作與分析內容的詮釋；請針對次級資料庫中所選擇之研究議題所做之分析與詮釋，引用及書寫格式請比照《臺灣社會學刊》撰稿體例，若有抄襲之情事以零分計。
- (3) 分組個人表現 (10%)：由同學互相評鑑。
- (4) 課堂參與出席 (10%)：隨堂點名，正式請假需出示紙本證明。

課程進度與閱讀教材

週次	日期		課程進度	閱讀教材
1	二月	21 日	課程簡介；學生分組	
2		28 日	和平紀念日放假一天	
3	三月	7 日	次級資料分析簡介與認識台灣常用次級資料庫；學生分組確定	課程教材(3) Chapters 4, 5
4		14 日	R 軟體的基本操作及程式語言邏輯簡介	課程教材(1) 第一、二章 課程教材(2) 第一至四章
5		21 日	次級資料庫的下載與匯入；學生選定研究分析主題	課程教材(1) 第三、四章 課程教材(2) 第六、十二章
6		28 日	研究議題資料基本描述統計及資料視覺化的呈現：使用 R 檢視資料及繪圖	課程教材(1) 第六章 課程教材(2) 第七章
7	四月	4 日	兒童節放假一天	
8		11 日	研究議題資料之摘要統計，包含變異量分析	課程教材(1) 第八、十二、十三章 課程教材(2) 第十五章
9		18 日	期中考週	
10		25 日	次級資料常見之態度相關資料蒐集模式及適用之分析法：信度分析與因素分析	課程教材(1) 第十七章
11	五月	2 日	相關度建立後之模型建立：線性模型分析	課程教材(1) 第十五章 課程教材(2) 第十六章
12		9 日	相關度建立後之模型建立：其他線性模型分析	課程教材(1) 第十九章 課程教材(2) 第十七章
13		16 日	進行推論之前：模型診斷	課程教材(2) 第十八章
14		23 日	畢業考週；態度資料之模型建立：結構方程模式	lavaan package, http://lavaan.ugent.be/
15		30 日	從資料庫裡尋找變項：資料分群	課程教材(2) 第二十二章
16	六月	6 日	期末報告與講解	

17		13 日	期末報告與講解	
18		20 日	期末考週	