

軟體的選擇與使用

表一 統計軟體使用調查 (得票數/百分比)

年度 (投票數)	2014 (3285)	2013 (1880)	2012 (798)
軟體			
RapidMiner	1453/44.2%	737/39.2%	213/26.7%
R	1264/38.5%	704/37.4%	245/30.7%
Excel	847/25.8%	527/28.0%	238/29.8%
KNIME	492/15.0%	110/5.9%	174/21.8%
Weka / Pentaho	558/17.0%	269/14.3%	118/14.8%
StatSoft Statistica	56/1.7%	170/9.0%	112/14.0%
SAS base	357/10.9%	202/10.7%	101/12.7%
MATLAB	277/8.4%	186/9.9%	80/10.0%
IBM SPSS Statistics	253/7.7%	164/8.7%	62/7.8%
SAS Enterprise Miner	235/7.2%	110/5.9%	46/5.8%
Python	639/19.5%	250/13.3%	119/14.9%

註：整理自 KDnuggets 2014、2013 及 2012 的調查結果 (綠底是自由軟體)。

表一 統計軟體使用調查

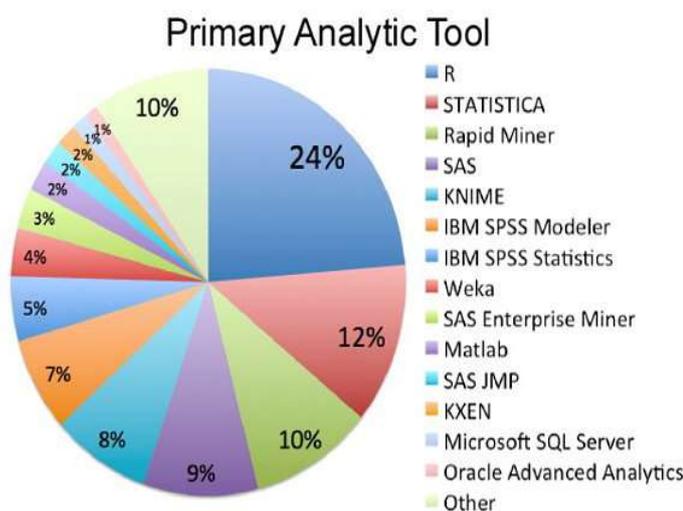


圖 1 主要的分析工具軟體

取自 2013 Data Miner Survey Summary Report

圖 1 主要的分析工具軟體

類似功能的軟體通常不只一種，例如：統計分析軟體可以為作為巨量資料（big data）的理解與分析的工具。在臺灣既有的文獻中，不難發現 SPSS（Statistical Product and Service Solutions）一直是教育學門最常見的統計分析軟體。坊間出版的教育研究法書籍中，SPSS 的使用幾乎伴隨著統計範例一起出現於書中。可是除了 IBM 公司的 SPSS 外，國際上仍有其他類似功能的套裝軟體。如：美國史丹索特公司（statsoft）的 STATISTICA、統計系統分析公司的 SAS（Statistics Analysis System）及 The MathWorks 公司的 MATLAB 都是數值分析及統計應用的商業軟體。而在自由軟體方面，R 語言（<http://www.r-project.org/>）、RapidMiner（<http://sourceforge.net/projects/rapidminer/>）及 GNU Octave（www.gnu.org/software/octave/）都是長年更新具有類似功能的軟體。

在全球使用的統計上，致力於大數據分析領域的 KDnuggets（www.kdnuggets.com）藉由問卷搜集各界使用何種軟體來完成資料探索、巨量資料分析於真實的專案中（What Analytics, Data mining, Big Data software you used in the past 12 months for a real project?），發現最多人使用的是免費下載的自由軟體。就整體趨勢而言，也是越來越多人選擇自由軟體。在問卷中 RapidMiner 是 2014 年及 2013 年最多人使用的軟體，而 2012 年最多人使用的是 R 軟體，這兩個軟體都是自由軟體。另外一個統計軟體調查，美國 Rexer Analytics 公司（<http://www.rexeranalytics.com/>）於 2013 年 1259 個受訪者中，R 軟體是最多人使用的資料探勘軟體。

從以上的調查中可以發現，自由軟體在統計分析上佔有重要的角色，而 SPSS 並不是最廣泛使用的軟體。另一方面，若細看 Rexer Analytics 公司的年度報告（2013 Data Miner Survey Summary Report），許多研究者不會只用一種軟體來計算統計數據，利用不同軟體作為交叉比對是比較可靠的作法。

大部份人對軟體的選擇是被動的，在求學階段老師教授何種軟體，學生畢業後也使用相同的軟體；進入職場所使用的軟體也是來自當時的環境，大家習慣依附既有的資訊環境。可是軟體的進步實在非常快速，短短的時間，手機上的 APP 軟體廣受大家好評，並融入生活與學習之中，但 APP 軟體並不存在於數年前的資訊環境。許多網站都有軟體使用的排行榜，也許可以提供不同的選擇參考。如：每天都有百萬人次下載的 SourceForge，提供各種自由軟體的下載排行，網址為 <http://sourceforge.net/top>。臺灣 PCHOME 的分類下載排行，網址：<http://toget.pchome.com.tw/rank/>，也提供各式中文軟體的下載排行。如果不知道任何軟體排行的網址，只要在搜尋網站輸入“most used software”或“軟體排行”就可找到各式各樣的軟體排行。

如何選擇軟體、善用軟體變成現代人生活不得不面對的問題。也許我們並不知道什麼軟體是最好的、最適合自己的，但參考並試用國際上大多數人的選擇，給自己多一些彈性，不要侷限於特定軟體，才有機會得到更有效率的服務。