

「四化」形塑下的教育統計

教育部統計處催生教育大數據計畫，擴增逾半人力及經費，力行組織及業務變革，並以四化－巨量化、國際化、資訊化及技術化形塑新團隊，業務質量及人員知能大幅躍進。

鄭靜芬、馮漢昌、劉嘉蕙、郭溫慈、李書惠（教育部統計處專員、專員、專員、專員、工程師）

壹、前言

考量本部業管之各項個體資料已趨完備，本處爰推動「跨機關大數據統計資料庫建置計畫」，期透過跨機關巨量資料連結，達成學校填報減量及統計資訊延伸之雙重目標。經部內各單位就資料提供、處理、資訊需求、資安等議題進行協商及個資盤點後，計畫終獲部長核定，並於實體隔離資料室建置完成後，正式於 110 年 7 月啓動，本處亦獲致 11 名恆定人力及每年 900 萬元之計畫執行經費，旋即導入國際政府統

計新思維，援引組織變革理念，以巨量化、國際化、資訊化及技術化為基調，實施單位組織及業務重整；轉型迄今，成效斐然，屢獲部長及業務單位稱許，無論組織績效及人員素質，皆有脫胎換骨之蛻變與成長。

貳、新時代教育統計——「四化」

一、教育統計巨量化

統計業務巨量化有助於提升資訊品質、擴增資訊層面，並降低行政查報或統計調查作業負擔及成本；而處理及分析

大數據，更是新時代統計人員不可或缺之技術及知能。相關執行成果如下：

（一）創建教育跨域大數據統計資料庫

透過資料盤點及協商，按年跨域整合連結本部、內政部、勞動部等 30 餘項大數據資料集，並運用資料處理及資料庫管理等相關技術，於清整資料後建置橫斷及縱貫（cross-sectional and longitudinal）教育跨域大數據統計資料庫（下頁圖 1），作為本部循證決策（EBDM）之重要基礎。

(二) 援引循證決策概念，產製大數據統計結果

運用教育跨域大數據統計資料庫，產製「性別平等教育法」規範之學生懷孕統計結果，以彌補學校難以掌握相關情事之缺憾，並大幅降低統計黑數及行政查報負擔；復構思學生體能、身心疾病、打工、跨學區就讀、獎學金與公費出國等觀測主題，產製相關統計結果，並運用大數據技術深入剖析，有效延伸資料應用之層面及價值。

(三) 創建實體隔離環境，落實資安管控機制

為配合大數據計畫推

展，確保執行過程資安無虞，爰於計畫啟動前完成實體隔離環境建置，期以斷網、監控及門禁等措施，嚴密維護個資；另同步研訂本部「跨機關大數據統計資料庫機敏資料暨實體隔離環境作業規範」，並奉部長核定後實施，以為個資索取、處理、應用、稽核及實體隔離環境使用相關事項之準則。

二、教育統計國際化

身為新時代統計人員，自應逐時掌握國際政府統計之規範、技術、做法及新知，並通盤審酌國內情勢，擇適者效法

之，係統計業務創新升級之重要來源。近期教育統計領域之國際化執行成果如次：

(一) 創編聯合國永續發展目標—教育統計指標

聯合國永續發展目標—教育統計指標 (Sustainable Development Goal 4 Indicators, SDG 4 Indicators) 係全球教育領域重要議題之統計指標，爰依聯合國教科文組織

(UNESCO) 定義，廣泛蒐集相關政府機關資料，於 109 年及 110 年陸續創編並發布「中等以下教育完成率」、「國小學齡前一年組織學習參與率」、「青年與成人正規及非正規教育訓練參與率」、「教師流失率」、「教師在職訓練率」及「提供基礎設施服務之學校比率」等指標，以及各指標之性別平衡指數 (Gender Parity Indices, 「教師在職訓練率性別平衡指數」如下頁圖 2)，以呈現我國教育永續發展輪廓，建立當前教育質量之比較基礎，供各界參用。

圖 1 「教育部跨機關大數據統計資料庫建置計畫」基本作業流程



資料來源：教育部統計處。

專題

(二) 試編重要國際組織之教育統計指標

為評估高級中等教育資源投入效率及成果，爰依 OECD 定義，依據高級中等學校查報相關資料，試編高級中等教育完成率（Upper Secondary Completion Rate）統計指標，以具體衡量學生入學後完成學業能

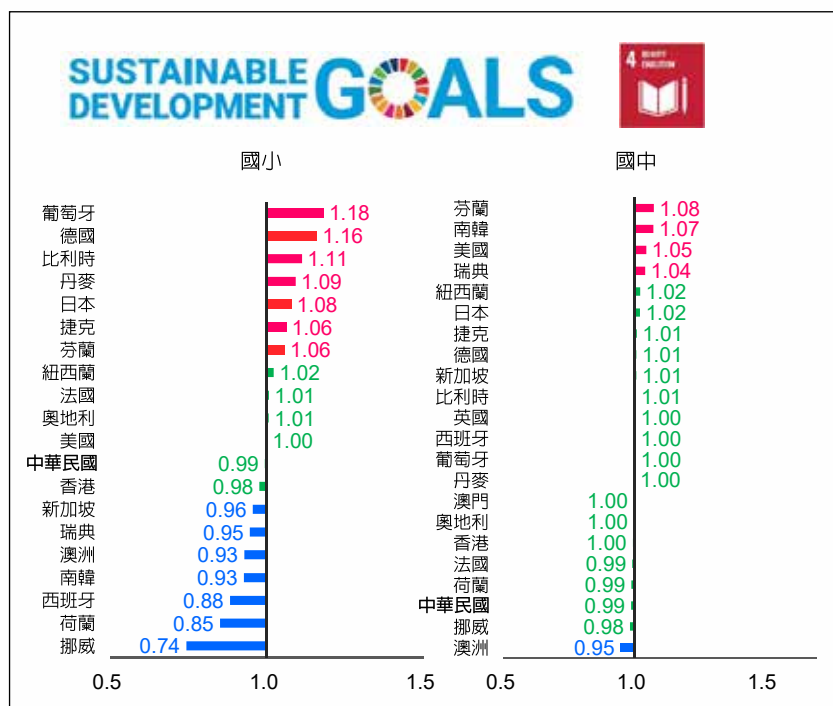
力；並依聯合國教科文組織（UNESCO）定義，依各級學校填報資訊，試編粗畢業率（Gross Graduation Ratio）統計指標，以量化評估各級學校畢業年齡人口之學業完成程度。此二項指標適可提供邁入下一教育階段或勞動市場之人力資源量化衡量標準，實係相關教育政策之制

定參據，亦係教育統計國際接軌之重要元素。

(三) 定期查閱國際統計新知及規範

為永續發展本處國際化重點工作，爰設置種子人員定期查閱國際統計技術、新知及規範，並藉學習型組織概念，建立知識庫分享全體同仁，期發揮擴散效果，提升人員專業素養及國際觀。

圖 2 教師在職訓練率性別平衡指數



說明：1. 澳門、英國、德國、西班牙、比利時、荷蘭、奧地利、捷克、丹麥、冰島等國之國中部分為 2018 年資料，餘均為 2019 年資料。
 2. 本指數係以教師在職訓練率之女性、男性數據比例計算而得，以綠色表示呈性別平衡國家數值，女性優於男性者以紅色表示，男性優於女性則以藍色表示。
 資料來源：教育部教師在職訓練資料、聯合國教科文組織（UNESCO）各國資料。

三、教育統計資訊化

隨資訊科技發展，統計工作亦應就資料處理質效及資訊陳示技術求新求變，以產製優質統計資訊，提升統計服務質量。近期本處人員積極開發各項資訊系統及資訊視覺化功能，成效斐然，執行成果如次：

(一) 自力創建「教育統計動態視覺化平台」

為順應國際間政府統計之視覺化趨勢，爰自力創建「教育統計動態視覺化平臺」

(圖 3)，除納入原建置之「主題式互動統計圖表專區」，復應用尖端技術增設下列專區：

1. 各級學校地理資訊及地區別統計查詢專區：以行政

區域、各級學校教育統計圖表資料結合 2D 地理位置圖徵為基底，再應用 ArcGIS 開發工具，導入 3D 建物模型技術(圖 4)，將地理統計資訊之呈現躍

升至立體場域；另增設「縣市統計視覺化儀表板」，期以智慧、直觀及互動方式，提供全方位視角之教育統計決策分析。

2. 體驗式互動統計專區：以情境體驗 (Situational Experiencing) 概念，設計開發教育統計相關之互動遊戲，增進使用者對於統計資訊之趣味性、親和性及理解程度。
3. 動態資訊統計圖表專區：應用 Flourish 及 D3.js 最新製圖技術，融入動態資訊圖表 (Dynamic Infographic) 概念元素，期藉高度親和之競跑圖、3D 寰宇圖及熱點圖等動態元素，生動呈現統計結果之歷年變化趨勢，精準傳遞資訊意涵。

(二) 開發「教育統計作業系統」

各級學校教育為基本教育統計主軸，為有效整合、應用跨教育階段統計作業，爰開發「教育統計作業系統」

圖 3 教育統計動態視覺化平臺首頁



資料來源：教育部統計處。

圖 4 各校 3D 立體圖



資料來源：教育部統計處。

專題

(圖 5)，將跨教育階段資料填報、製表、分析及列印等功能系統化，以提升統計

資訊品質、服務質量及作業效率。

(三) 自力創建「教育性別統

計動態資訊圖表」

教育領域之性別議題向受各界關注，爰運用 Infogram 動態資訊圖像工具，創建「教育性別統計動態資訊圖表」(圖 6)，期以行動優先 (mobile first) 之響應式 (RWD) 網頁為基底，藉縮放、翻轉及淡入淡出等可視化動態效果，賦予資訊圖表嶄新面貌，提升統計易讀性及服務質量。

(四) 規劃建置「教育統計決策支援系統」

為降低本部各單位之統計專業落差，提升本處統計服務效率及多元性，爰以「資訊範圍極大化」、「資訊產製便捷化」、「資訊應用權限化」為核心概念，自力規劃開發「教育統計決策支援系統」(下頁圖 7)，其核心概念係將本部各類量化資料悉數納入系統，並建置資料倉儲及多維度表格分析模型，將各類資料歸納整合於多層次選單，由使用者自

圖 5 教育統計作業系統功能架構圖



資料來源：教育部統計處。

圖 6 教育性別統計動態資訊圖表



資料來源：教育部統計處。

訂表單格式及篩選資料，透過線上分析處理技術即時運算，簡易快捷呈現統計資訊；另運用權限管控方式，限制內參資訊之使用及公開，以建構多元、精確、迅捷及安全之統計資訊查詢服務。

四、教育統計技術化

新時代政府統計業務愈趨複雜多元，以大數據為核心之循證決策概念允為當前主流，亟須提升業務及人員之技術層次，方可以堅實可靠之統計資訊，作為決策強而有力之後盾。近期本處面對疫情及少子女化

等衝擊，運用妥適統計技術完成多項技術化精進業務如次：

(一) 精進高中以上學生人數預測模型，強化理論基礎及確度

學生人數規模及未來變化趨勢攸關教育資源配置等重要政策規劃與執行，近年因少子女化、生源短缺等問題，各界愈趨關注各教育階段學生人數預測結果，爰檢視原預測模式作業，運用分段式迴歸 (segmented regression)、移動平均 (moving average)、指數平滑 (exponential smoothing) 及整合移動平均

自我迴歸 (ARIMA) 等技術，優化大專 1 年級及高級中等教育 (包含五專及七年一貫制前 3 年) 學生預測作業，以強化理論基礎，有效提升統計確度及品質。

(二) 準確預測每日確診人數，發揮統計超前佈署功能為即時掌握校園疫情變化，爰每日接收本部校園安全暨災害防救通報處理中心近 1 日確診數資料，依資料走勢、模型配適度、預測結果合理性及交叉驗證結果，建立「資料、模型同步滾動 (roll planning) 修正」機制，以「移動窗格最佳化」方法決定各等級學生及教職員次日 7 日確診人數之最適預測模型，並於 2 小時內將預測結果密陳部長。預測作業期間，翌日預測值平均絕對百分比誤差 (MAPE) 僅 9.4%，並曾連續 7 日誤差低於 5%，誤差人數低於百人者計 9 日 (下頁圖 8)，並準確預測校園疫情高峰，於疫情期間

圖 7 教育統計決策支援系統規劃圖



資料來源：教育部統計處。

專題

成爲多項防疫措施施行及物資發放之準據，允爲校園疫情超前部署利器。

(三) 應用資料探勘及大數據技術，剖析成因預知未來爲協助學生探索性向，發展跨域整合能力及競爭力，爰以大專校院畢業生個資及勞動、財稅等大數據資料，運用決策樹 (decision tree) 及隨機森林 (random forest) 機器學習 (ML) 方法，建置大學畢業生跨域升讀預測模型，提供學生跨域學習相關資訊及調整學校跨域課程之參據；另爲維護懷孕學生之受教權，並促進學

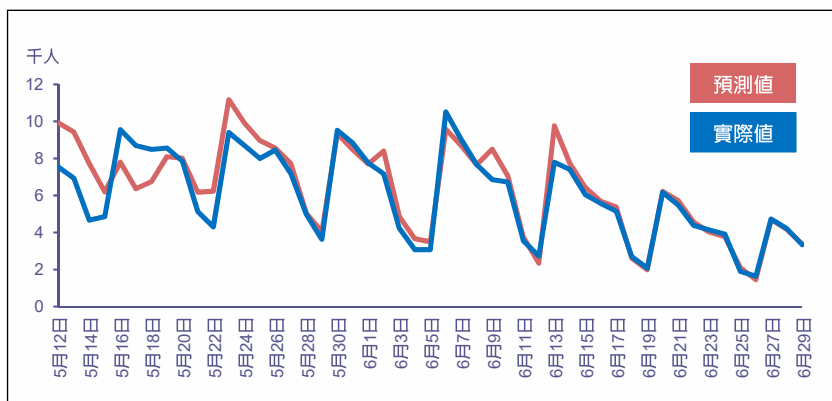
生身心健康發展，遂以本部各等級學生學籍與體適能、衛福部及內政部相關資料爲基礎，依懷孕、心理疾病、學生身體質量指數 (BMI) 及體適能等面向，分別建立羅吉斯迴歸及零膨脹負二項迴歸模型，以發掘各主題之潛在族群，作爲開設衛教、體能等課程及心理諮商輔導之參考。

參、結語

「循證決策」及「巨量資料」概念之興起，實係政府統計業務之發展方向及重點，作爲新時代統計人員，自應緊隨

其發展脈動；本處大數據計畫之施行，厥爲教育統計組織、業務及人員升級轉型之關鍵；而「四化」的推動，亦爲具體實踐循證決策概念之最佳例證。展望未來，教育統計仍將持續依循此一理念，引領業務及組織持續創新成長，並持續扮演數字智庫要角，成爲教育政策之指引明燈。❖

圖 8 預測翌日確診學生人數與實際值比較



資料來源：教育部統計處。