



# 美國聯邦統計多元資料來源之應用與挑戰

美國聯邦統計機構近年受到調查回收率下降、成本攀升，以及使用者需求日益增加之影響，積極推動多元資料來源之串接應用，包括政府機關之行政紀錄、民間企業之資料庫、網際網路資訊等；本文旨在介紹美國對多元資料來源的創新應用與挑戰，供我國政府統計借鏡。

王珮倫、陳宜屏（行政院主計總處綜合統計處視察、研究員）

## 壹、前言

美國憲法明定每十年舉辦人口普查，作為重新分配眾議院席位之依據，另各項聯邦統計亦大量應用於與民衆權益息息相關的聯邦政府補助分配、稅率調整及社會福利津貼核算等，顯見美國政府決策對於聯邦統計之高度仰賴。由於高品質的統計有助於規劃、監測及評估公共政策，致相關需求日益殷切，惟受隱私意識抬頭影響，逐漸下降的調查回表率

已危及統計品質，單靠統計調查亦漸難以滿足各界對統計資料的及時性與細緻度要求。有鑑於此，國家科學、工程及醫學研究院所屬國家統計委員會 2017 年修訂之第 6 版「美國聯邦機構之原則與執行準則」（Principles and Practices for a Federal Statistical Agency），納入「運用多元資料來源編製統計，以滿足使用者需求」（practice 3），用以督促聯邦統計機構勇於創新，積極運用及整合跨機關間之資料。

## 貳、多元資料來源在政府統計所扮演之角色

美國聯邦統計機構在成本效益考量下，為減輕受查者負擔且兼顧必要資料之完整性，近年著手進行多元資料來源的整合運用，除傳統普查或抽樣調查外，另納入公部門的行政紀錄、民間企業的資料庫，甚至是來自感測器（如交通攝影）、地理空間及網際網路（如購物網報價）等各項資料。藉

由多元資料來源之應用並掌握最新整合串接之方法與技術，是目前美國統計機構有效回應使用者需求的重要關鍵策略。以下就目前美國聯邦統計主要參用之資料來源說明。

## 一、調查

統計調查為編製聯邦統計的主要方法，然而近年因隱私權意識高漲導致調查回收率下降，間接推升蒐集成本，加上預算有限，日漸難以兼顧統計調查的品質與數量，以滿足日趨多元細緻的使用者需求。因此，美國政府統計機構不斷思考如何使用其他資料來源代替部分調查問項，或補強調查統計，藉以減輕受訪者的負擔且能維持統計資料的品質。

## 二、行政紀錄

### (一) 現況

行政紀錄係政府機關為行政管理、監測或執法等目的而蒐集的資料，資料量相較統計調查為多；統計調查機關採用行政紀錄不僅可以節省資料蒐集的成本，

亦不致對受查者造成額外的負擔。行政紀錄除了可以作為調查母體名冊外，尚可用以簡化調查問項、插補或估算未回復之資料，甚至取代調查等。美國聯邦機構已有一些創新且成功的案例（表 1）。

### (二) 行政紀錄之挑戰

1. 跨機關資料取用困難重重：由於聯邦統計體系屬高度分散體制，加以法律及行政實務的限制，部會間資料整合取用極為不易，另各州及地方政府亦蒐集與持有許多重要之行政資料，惟因美國憲法保障其行政的獨立性，聯邦統計機構無權要求分享及運用這些資料，均阻礙了資料的管理與運用效能。
2. 資料不一致或不完整待克服：隨業務法規修訂，行政紀錄可能隨之變動，影

表 1 美國聯邦機構使用行政紀錄之創新實例

| 機關名稱  | 運作說明   |
|---|--|
| 商務部   | 1. 普查局使用國稅局的聯邦稅務資料創建普查企業登記冊，作為經濟普查及多數企業調查的母體資料。<br>2. 普查局經濟普查及自營作業者（non – employer business）調查之部分問項，以稅務行政紀錄取代。 |
| 勞動部   | 勞工統計局依據各州取得的失業保險計畫資料，規劃就業與薪資按季調查之方法與架構。  |
| 司法部   | 司法統計局依據聯邦調查局提供的全國案件報告系統中大量的犯罪報案資訊，完善全國犯罪受害調查資料分析。  |
| 農業部   | 經濟研究局及食品與營養局合辦全國家庭食品獲取與購買調查，透過使用補充營養援助計畫的行政資料，進行樣本分層並取代部分問項。   |
| 住宅與城市發展部  | 與商務部普查局合辦住宅調查試驗計畫，利用稅務資料取代部分調查項目，增補未回復之資料，及研究與評估調查結果的合理性。  |
| 衛生與公共服務部  | 國家衛生統計中心使用醫療保健和醫療補助服務中心之養老院、居家保健及臨終安養院等機構行政資料，取代全國護理之家調查和全國家庭和臨終關懷調查。  |
| 資料來源：Innovations in Federal Statistics — Combining Data Sources While Protecting Privacy 及 Principles and Practices for a Federal Statistical Agency : Sixth Edition. |  |

# 論述》統計 · 調查



響資料之一致性與可比較性；此外，行政紀錄有精確度不一、資料儲存格式不利使用、未完整建檔等缺點，均有賴統計機構投入資源研究並加以克服，方可提升資料應用價值。

### 三、其他資料來源

#### (一) 現況

近年隨數據 (Data) 爆炸性的增長，引起各國政府統計機構對大數據應用之高度關注，惟因涵蓋範圍與資料定義釐清費時，法規面與技術面尚存困難，致目前仍鮮少據以編製政府統計。美國正探討如何使用調查與行政紀錄以外，包括來自網際網路、感測器，或私人企業所提供之數據等非傳統資料來源，產製聯邦統計。聯邦機構已有部分使用非傳統資料來源之實例，如商務部經濟分析局利用調查、健康保險理賠紀錄及民間資料庫等編製衛生保健衛星帳，並與普查局合作利用廠商銷售與交易資料來強化消費者支出

統計，另勞動部勞工統計局利用網路耙梳方法蒐集特定商品價格 (表 2)。

#### (二) 非傳統資料 (如網路資料) 來源之挑戰

1. 資料代表性不足：網路資料通常無法涵蓋所需資料的全部內涵與範圍，例如蒐集網際網路商品售價資料編製消費者物價指數，雖具發布之及時性並可節省調查成本，惟因未包括非線上消費者支出之資訊，致無法準確地反映全體消費者實際購買之價格

變動現況。

2. 資料缺乏可比較性：統計機構無法管控這類資料的一致性，當民間資料庫 (或網站) 因故停止營運或資料內容因定義或範圍改變而發生顯著變化時，都將危及時間數列資料的比較分析。

## 參、多元資料來源之整合

### 一、暢通多元資料來源之取用

表 2 美國聯邦機構使用非傳統資料來源之創新實例

| 資料來源  | 運作說明  |
|---|---|
| 健康保險理賠紀錄 (Health Insurance Claims Records)              | 商務部經濟分析局整合調查資料、健保理賠紀錄及 Truven Health Market Scan 公司的商業理賠和就診資料庫，於 2015 年首次發布 2000 年至 2010 年的衛生保健衛星帳。 |
| 銷售與交易資料之彙整資料庫 (Sales and Transactions Data Aggregators) | 商務部經濟分析局與普查局使用商店銷售掃描與信用卡交易的第三方 (third - party) 彙整資料庫，用以改善按月零售業調查的銷售估計值，未來將使用交易資料，以加強對消費者支出的初步估計。    |
| 網路耙梳 (Web Scraping)                                     | 勞動部勞工統計局於消費者物價統計計畫中，針對選定商品 (如電視機、相機、洗衣機等) 特性，進行網絡耙梳蒐集資料，並於更換商品花色時以特徵迴歸法估算價格變動。                      |

資料來源：Principles and Practices for a Federal Statistical Agency : Sixth Edition.

(一) 聯邦政府推動措施

聯邦管理暨預算局 (Office of Management and Budget, OMB) 所屬聯邦統計方法委員會於 2008 年成立「行政、替代及混合資料小組委員會」(the Subcommittee on Administrative, Alternative, and Blended Data)，已就行政紀錄之取得、使用及評估訂定參考案例與協議；OMB 再於 2014 年發布的工作指南指出，聯邦統計納入行政紀錄應作為執行準則之慣例。運用行政紀錄及其他非傳統

數據來源時應將下列因素納入考量，以評估其利弊與成本 (附圖)。

(二) 創建政府資料專責交換中心

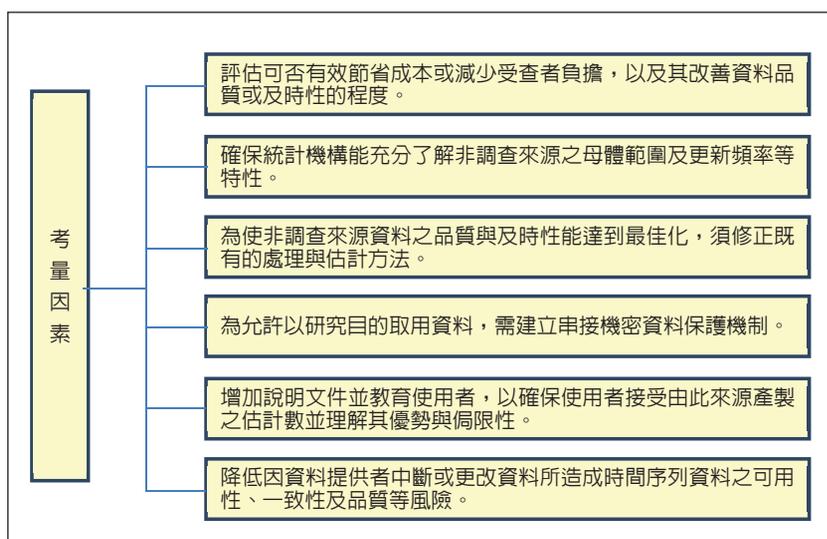
證據基礎決策委員會於 2017 年總結報告建議，應整合行政與調查資料用於統計計畫之研究與評估，並兼顧機敏資料之保護，為達此目的，有必要加強資料基礎設施、資料庫安全性及相關統計協議，並創建政府資料之專責交換中心，確保資料連結與使用之安全。

二、鬆綁資料使用限制

美國已於「機密資訊保護與統計效率法」(Confidential Information Protection and Statistical Efficiency Act, 簡稱 CIPSEA) 中授權普查局、經濟分析局及勞工統計局等 3 個統計機構可相互分享企業個別資料，其他機構之行政資料取用則仍受到法律與行政障礙的限制而困難重重，目前唯有透過修訂相關法規，才能解決資料取用的合法性。

另「美國聯邦機構之原則與執行準則」第 6 次修訂版指出，現行不同機構間的資料取用，係由個別統計機構與其他聯邦機構、州立機構及私人企業分別進行協商，形成人力與時間的浪費，因此建議加強 OMB 所屬跨機構統計政策委員會 (the Interagency Council on Statistical Policy) 之協調與評估權力，依資料串接的型態，發展適當的估算方法，制定估算之品質指標，並應用先進的資料隱私保護技術，以利資料取用。

附圖 使用非調查資料來源之考量因素



資料來源：Principles and Practices for a Federal Statistical Agency : Sixth Edition.



### 三、加強資料隱私保護

雖聯邦統計機構取用多元資料來源須依保密承諾將個別資料去識別化或匿名化，以防止洩露個人資料；惟相較於單一來源的單純性，當結合多元來源資料時，可能藉由其他特徵值的串接而辨識出個人或組織身分，進而增加個別資料外洩或身分遭盜用等風險。因此，統計機構與其部門業務單位應合作加強資料串接的隱私保護，建立有效且安全的資料處理方法，並保證相關資料蒐集之處理與儲存程序皆以統計為目的，不用於其他行政管理或管制用途。

### 肆、結論

我國政府統計業務，向來公務統計與調查統計並重，隨資訊技術進展迅速，統計基礎資料之蒐集與處理方法日益多元細緻，品質、數量及時效亦隨之精進；惟調查統計之效率與品質受到隱私權意識高漲、詐騙案件頻仍等經社環境變遷影響，與美國同樣面臨回收率

下降、成本增加之嚴峻挑戰。為能有所因應，並配合資料串接連結整合加值應用之國際統計發展潮流，本（107）年 6 月公布之最新版統計法，已增訂主計機構基於統計目的，或為減輕受查者負擔，除法律另有規定外，得向各該政府所屬機關要求提供公務或調查相關資料（第 18 條），各機關建立及修改行政資料處理系統應納入統計需求（第 17 條），以及為強化個別資料保密，保障受查者權益，明定各機關辦理統計業務人員如蒐集未經授權之資料與違反調查資料保密等不當行為之罰則（第 25 條）等，期能奠定資料供需雙方衡平友善之統計基礎。

為發揮綜效，統計機構應積極汲取相關實務經驗與新知，引進最新發展技術，持續強化人才培訓、資料基礎設施與資料庫安全性，並完備數據提供與取用等管制規範，以健全統計環境。至於網際網路、感測器及私人企業等非傳統資料來源之數據應用，屬政府統計的創新嘗試，各國尚處評估

與研究階段，未來應持續關注其發展，適時提供我國政府統計借鏡。

### 參考文獻

1. The National Academy of Sciences · Engineering · Medicine (2017), *Innovations in Federal Statistics — Combining Data Sources While Protecting Privacy*, Washington, DC.
2. The National Academy of Sciences · Engineering · Medicine (2017), *Principles and Practices for a Federal Statistical Agency : Sixth Edition*, Washington, DC.
3. The Commission on Evidence — Based Policymaking (2017), *The Promise of Evidence — Based Policymaking*, Washington, DC. ❖