



政府統計數據整合與治理

數據管理及應用攸關政府施政決策的方向與效能，本文爰就地方政府數據治理之思維加以分析探討，以增進各界對地方統計彙管應用實務之瞭解。

謝邦昌（輔仁大學資源與事業發展副校長）

壹、前言

管仲曾言：「不明於計數，而欲舉大事，猶無舟楫，而欲經於水險也。」「舉事必成，不知計數不可。」；「明法審數」則可知道方向以及瞭解輕重，可用以治理國家，也可以擺渡人生。古時智者就已經知道數據的重要性。然而，數據要有用，要能產生效能，須具備多項要素，包括：

- （一）權責及輔導單位之設立。
- （二）資料長期有系統的整理，產生有用的資訊及知識。
- （三）善用巨量資料，強化應

用數據內涵。

本文將以地方政府數據治理角度簡要說明。

貳、數據整合策略探討

一、權責及輔導單位之設立

隨著社經環境的演化，政府所面臨的數據種類也更複雜化，對政府單位而言，如何在面對這些資料當中，把它條理化、歸納化是一個很大的課題。因此，如能設立一個輔導單位協助推動，將可以讓這些資料更具條理、更具說服力。

當然，建構相關的對應制度以區分何種資料由哪個單位

去管理、產製，這是政府單位所面臨的第一個課題。另外，當已歸納分析由何者產製時，它的內涵如何去做相關的定義，以及制度的建立，又是另一個課題。因此，在數據整合的過程中，如能有相關權責及輔導單位成立，將可使數據的整合及應用，更容易推動（下頁圖 1）。

二、資料長期有系統的整理，產生有用的資訊及知識

政府的施政及作為是長期性的，要能面對各種外在的衝擊而做適度的調整，為達成

這個目標，對資料 (Data) 的蒐集一定要經過系統化、長期性的整理，再經由相關資料的分析，才能產生有用的資訊 (Information)，這些資訊在經過實務的探討，才能產生所謂的知識 (Knowledge)。

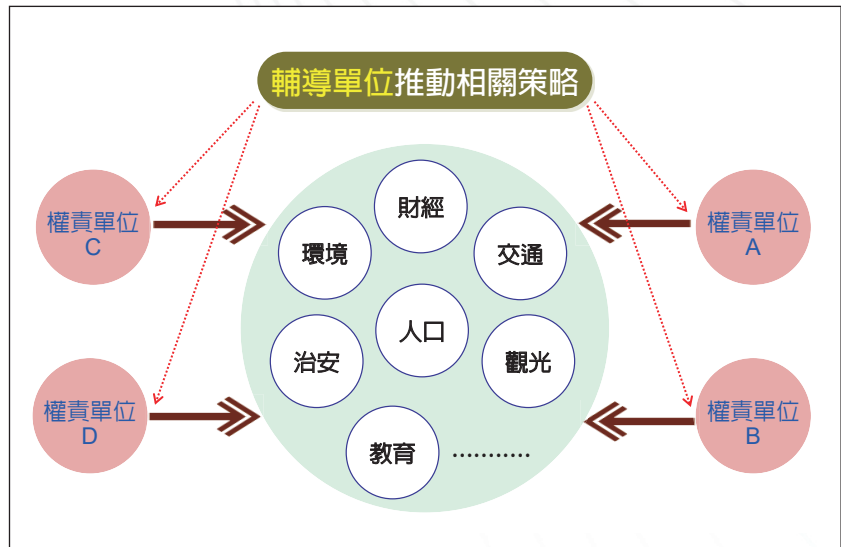
事實上，在政府長期施政的過程中，都會產製相關的資料，而這些資料目前都散布在報表、書刊或資訊系統等各項投入中，藉由資料的有效萃取、資料的交換或整合技術等轉換程序，才能產出有用的資訊及知識 (圖 2)。

三、善用巨量資料，強化應用數據內涵

由於每一類別的巨量資料收集不易，必然需要能將資料好好歸納並整理，以便將資料發揮出效益來。目前資料的運用，通常以採礦方式對各種巨量資料加以整理。

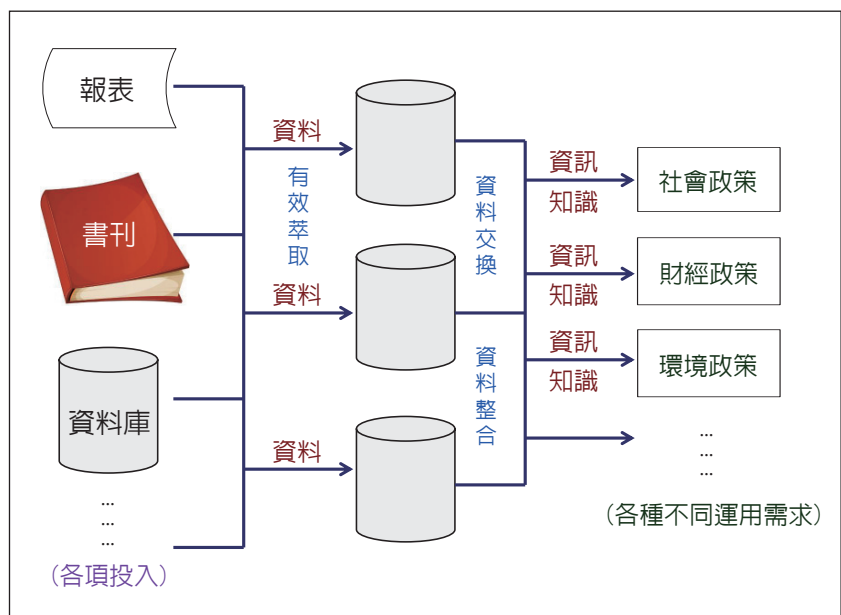
資料的收集要產生效能，所收集的數據定是要有用且完整的數據。此外，資料採礦是近年來資料庫應用領域中，相當熱門的議題，即為大家

圖 1 權責及輔導單位之業務關聯



資料來源：作者自行繪製。

圖 2 資料、資訊及知識之投入產出轉換過程



資料來源：作者自行繪製。

專題



口語中的 Data Mining，但它其實不是新東西，因為 Data Mining 使用的分析方法，如預測模型（迴歸、時間數列）、資料庫分割（Database Segmentation）、連接分析（Link Analysis）、偏差偵測（Deviation Detection）等，其實過去美國政府從第二次世界大戰前，就在人口普查以及軍事方面使用這些技術。但是資訊科技的進展超乎想像，新工具的出現，例如關聯式資料庫、物件導向資料庫、柔性計算理論（包括 Neural Network、Fuzzy Theory、Genetic Algorithms、Rough Set 等）、人工智慧的應用（如知識工程、專家系統），以及網路通訊技術的發展，使得從資料堆中挖掘之寶藏常常能超越歸納範圍的關係，Data Mining 便成為企業智慧的一部分。

在政府統計數據整理與治理方面，是否能與 AI 做結合，一直是大家關心的課題，目前 Data Mining 已被許多研究人員視為結合資料庫系統與機器學習技術的重要領域，許多產業

界人士也認為此領域是一項增加各企業潛能的重要指標。此領域蓬勃發展的原因：現代的企業體經常蒐集了大量資料，包括生產、財務、市場、客戶、供應商、競爭對手以及未來趨勢等重要資訊，但是資訊超載與無結構化，使得企業決策單位無法有效利用現存的資訊，甚至使決策行為產生混亂與誤用。如果能透過資料採礦技術，從巨量的資料庫中，發掘出不同的資訊與知識出來，作為決策支援之用，必能產生企業的競爭優勢。

基本上，Data Mining 是用來將你的資料中隱藏的資訊挖掘出來，所以 Data Mining 其實是所謂的 Knowledge Discovery 的一部分，Data Mining 使用了許多統計分析與 Modeling 的方法，到資料中尋找有用的特徵以及關聯性。Knowledge Discovery 的過程對 Data Mining 的應用成功與否有重要的影響，只有它才能確保 Data Mining 能獲得有意義的結果。

綜合以上，要應用 Data Mining，必須要有完善格式

化的電腦資料庫。各地方政府統計資料庫目前僅供資料查詢用，似可利用其做 Data Mining，找出一些相關問題，作為政策擬訂的參考。目前政府基本統計資料已漸趨完備，如何利用 Data Mining 工具來挖掘出具有建設性的資訊，需要瞭解 Data Mining 的應用，常用的有分類、推估、預測、關聯分組及同質分組，尚有其他功能可針對不同的需求及決策系統，發展出更多的應用方向，此不細說。

若以另一實際案例「查稅選案」來做說明，它是應用判別分析法的基本原理之觀念，首先找出最能判斷群體差異的變數，亦即識別能力較高之變數，再建立判別函數，用實證資料中的個體做判別歸類，最後比較各種分析方法的優劣。最後的結果是依不誠實度排序選案，來進行選案查核工作，因為此法能以查核最少的家數，而得到較高之補稅額，用最少的人力有效提升稅額，相當智慧化，也可以將此法運用於健保的核價及抓弊。

參、政府數據整合應用實例

經觀察行政院主計總處及地方政府相關應用平台，已陸續建構相關數據整合及應用機制，若能有效引導運用，將能夠提升施政效能。所建構之項目，茲列舉如下：

一、縣市重要統計指標

行政院主計總處為展現各縣市重要施政成果，彙編有縣市重要統計指標（圖3），且國內報章媒體也常引用作為縣市施政評比之依據。

二、視覺化圖表展現

資料視覺化（Data Visu-

alization）就是把資訊轉化成以不同型式圖表或圖像化來呈現，由眼睛接收和靠著大腦綜合辨識能力，將隱藏未知的訊息，轉換成有用的資訊。

視覺化圖表展現皆使用各項可視化工具（如 Power BI、Tableau 等視覺化軟體），將資料結果和設計兩者結合，從多樣龐雜（Variety）的資料中萃取出有用的資訊，製作成簡單易懂的資訊圖表（Infographics）（下頁圖4）。

近來地方政府也都不約而同地加強資料視覺化展現，開始陸續運用施政成果資料建置視覺化查詢專區。

三、數據整合應用資訊化

科技進步帶動可測量和取得數據的管道越來越多元，同樣地資料也越來越細緻，可是隨之而來將面臨資料量和種類的爆炸性成長。

行政院主計總處為自動化儲存資料，從眾多來源彙整並進行標準化，建立資料倉儲（Data Warehouse），也建置統計資訊化系統，供各地方政



資料來源：中華民國統計資訊網。

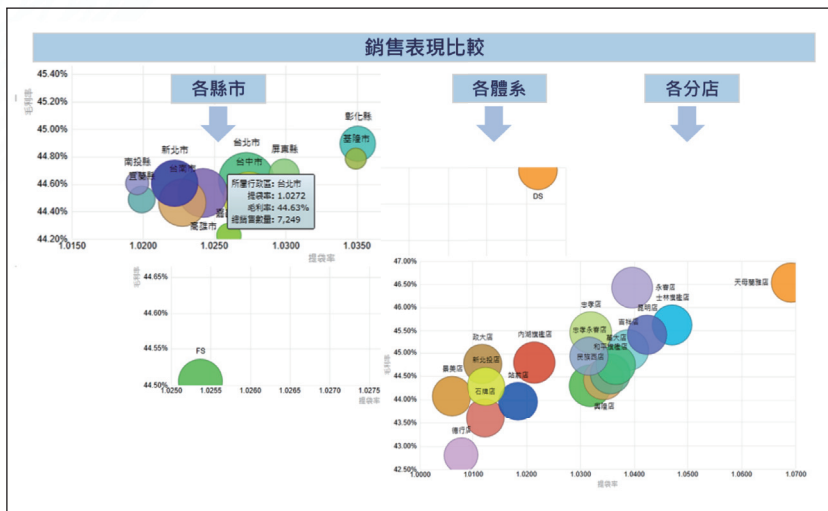
專題

府應用。

各地方政府的資料倉儲如能陸續建構，並將描述資料倉儲之資料字典 (Data Dictionary)

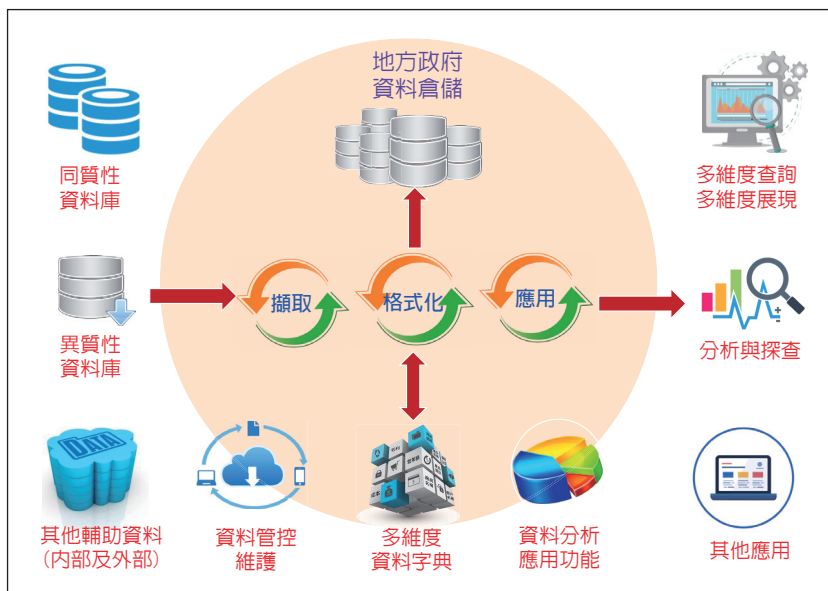
予以一致化，同時導入系統分析工具，對於資料長期有系統的整理，以及決策應用上，都將有極大助益 (圖 5)。

圖 4 視覺化圖表功能應用



資料來源：大數據分析 Excel Power BI 全方位應用第 3-42 頁。

圖 5 資料倉儲之建構及應用



資料來源：作者自行繪製。

肆、結語

基於地方政府對數據應用需求之迫切性，須以數據整合和治理兼顧的原則，設定分期、分級及階段性目標，有系統的優化資料彙整、管理及應用內涵。中央對地方數據的整合特別需要制定通則並兼顧其特色。數據應用效能的提升則有待設立權責或輔導單位，並建置職缺、投入空間及經費；並予以有系統的長期整理，包括數據的收集、規劃、及格式化。此外，在治理方面亦須制定法令，保護人民個資，使其運用能符合普世倫理規範。如此，政府統計數據整合與治理將會出現嶄新的契機。

參考文獻

1. 中華資料採礦協會，資料採礦 (Data Mining) 技術在主計業務上應用之研究計畫書。
2. 謝邦昌、鄭宇庭、宋龍華、陳妙華，大數據分析 Excel Power BI 全方位應用，基峰資訊。
3. 中華民國統計資訊網，縣市重要統計指標，<https://winsta.dgbas.gov.tw/StatWeb/vision/vision.html>。