

107000000AU130002

創辦內政大數據連結應用專案計畫

內政部統計處自行研究報告

中華民國 107 年 11 月

本報告內容及建議，純屬研究人員意見，不代表本機關意見

目 錄

第 1 章 緒論.....	1
第 1 節 計畫緣起與目的.....	1
第 2 節 計畫範圍與架構.....	3
第 3 節 計畫功能與特色.....	4
第 2 章 文獻回顧.....	6
第 1 節 國內.....	6
第 2 節 國外.....	7
第 3 章 推動方式.....	9
第 1 節 成立工作小組.....	9
第 2 節 辦理教育訓練.....	10
第 3 節 規劃大數據試辦委外案.....	12
第 4 章 主要成果.....	14
第 1 節 教育訓練成果.....	14
第 2 節 自行研究成果.....	24
第 3 節 規劃大數據試辦委外案.....	27
第 5 章 效益評估.....	32
第 1 節 創新性.....	32
第 2 節 應用性.....	33
第 3 節 效益性.....	33
第 6 章 結論與建議.....	35
第 1 節 結論.....	35
第 2 節 建議.....	36
參考文獻.....	38

圖目錄

圖 1、106 年內政統計研討會 花政務次長開幕致詞及座談會一隅	2
圖 2、業務議題診斷表 (A 表)、業務資料清洗表 (B 表)	5
圖 3、內政數據分析及決策應用工作小組架構圖	10
圖 4、內政大數據分析工作坊 (第 1 期) 由沈主任秘書淑妃 (前排右 6) 主持 開訓	11
圖 5、107 年內政大數據分析技術講座	12
圖 6、你覺得做數據分析最痛苦的是什麼? (複選, 最多 3 項, N=40)	13
圖 7、內政大數據分析工作坊成果—模擬城市 M.....	15
圖 8、內政大數據分析工作坊成果—道路交通事故預防	15
圖 9、內政大數據分析工作坊成果—新住民群聚	16
圖 10、107 年第 1 場內政大數據分析技術講座「敏捷式有效議題開發」上課情 形	22
圖 11、107 年第 2 場內政大數據分析技術講座上課情形	23
圖 12、107 年第 3 場內政大數據分析技術講座上課情形	24
圖 13、「道路交通安全肇事原因之地域與個人因素探討」分析成果	25
圖 14、「道路交通安全肇事原因之地域與個人因素探討」影片中統計圖	26
圖 15、「潛在風險住宅分布情形探討」分析成果	27
圖 16、內政大數據分析主題、議題與 KPI 彙整表.....	29
圖 17、內政大數據人房地資料連結簡表	30

摘 要

「統計檢驗政策、支援決策」是 80 年代政府統計界，喊得震天價響的施政名詞，惟隨主事者人去政息，慢慢的大家就不多提了，甚至有人覺得這只是八股口號、華而不實。惟隨大數據分析潮流興起，不僅國內外民間企業蓬勃運用，政府部門亦高度重視，美國國會甚至將循證決策（Evidence Based Policy Making）列為政府預算編列重要參考因素，使得統計分析於政府決策之功能再次被高舉強調，蔚為當今顯學。

為深化政府循證決策功能，落實數字管理精神，內政部由統計處（以下簡稱本處）主導成立工作小組，並推動本專案計畫。執行過程強調結合各單位業務專長，相互支援，激發團隊合作綜效。

推動辦理事項包括：1.大數據分析工作坊：採業務、統計、資訊 3 類人員混合分組方式，以客製化教材及視覺化分析軟體講授與實機操作，全程並有本處派員輔導進行。2.大數據技術講座：邀請產官學界專精人士，以專題講授搭配學員互動模式，進行大數據技術實例探究，俾利獲取最新資訊並啟發業務應用方向。3.大數據試辦委外案：「資料清洗」及「找出議題」為當前推動大數據分析最需解決之問題，占大數據分析工作量 6~8 成，耗時費工，需藉由外部力量來協助完成，爰針對較複雜費工之連結應用，辦理委外服務案，引進專業資訊服務，協助本部進行議題與資料盤點、資料庫清洗、資料運用分析（建立 SOP），並進行 2 項前導案例之實作。過程中本處積極調和委外廠商與業務單位之分工合作，以達專案計畫推動目標。

關鍵字：大數據、循證決策、資料清洗、資料盤點。

ABSTRACT

Keywords: Big Data, Evidence Based Policy Making, Database Cleansing, Data Inventory

“Statistics examines policies and supports decision-making”, an earth-shaking slogan of government statistics in the 1980s, was gradually not being mentioned with the change of ruling. However, with the rise of big data analysis, statistics has become widely used not only by domestic and foreign private enterprises, but also got great appreciation from the government departments. The US Congress even lists Evidence Based Policy Making as an important reference for government budgeting. Statistical analysis has been emphasized in the function of government policy-making once again.

In order to deepen the governmental function of evidence-based policy-making and implement the spirit of digital management, the Ministry of the Interior, led by the department of statistics, sets up a working group and promotes this project plan. The implementation process emphasizes combining business expertise of each department, mutual support, and inspiring team synergy.

Matters to be promoted and carried out include 3 parts: 1. Big Data Analysis Workshop: Mixing people specializing in business, statistics and information technology into groups, using customized teaching materials, visual analysis software by computer practice. 2. Big Data Technology Lecture: Inviting professionals from industry and academia to teach and interact with students, conducting case studies on big data technologies, to obtain the latest information and to inspire business applications. 3. Big Data Trail Outsourcing Case: “Data Cleansing” and “Issue Discovering” are two of the most important issues to be solved in the current promotion of big data analysis, with the workload of 60~80%, which is time-consuming, labor-intensive and needed to be assisted by external forces. Therefore, for more complex data connections and applications, introducing professional information services to assist us in making an inventory of issue and data, as well as conducting database cleansing and analysis of data utilization (establishing SOP), with implementation of two aforementioned leading cases by outsourcing services. During the process, we actively manage cooperation between the subcontractors and the business units to achieve the goal of the project.

第1章 緒論

政府大數據分析已成為施政趨勢，而要做好政府統計大數據分析，需要三方面的配合，資訊技術、統計方法與專業知識，因此充分結合資訊、統計與業務單位是推動大數據不可或缺的條件，本處遂於去（106）年底在機關長官全力支持下，創設內政部大數據工作小組及專案計畫，成為政府機關中第一個由統計單位主導成立部會級大數據工作小組及專案計畫。隨後立即展開相關大數據連結應用，支援決策之系列工作，如工作坊、技術講座、案例實作、委員會、研討會等，密集推動，並獲初步成果。

第1節 計畫緣起與目的

「統計檢驗政策、支援決策」是 80 年代政府統計界，喊得震天價響的施政名詞，惟隨主事者人去政息，慢慢的大家就不多提了，甚至有人覺得這只是八股口號、華而不實。惟隨大數據分析潮流興起，不僅國內外民間企業蓬勃運用，政府部門亦高度重視，美國國會甚至將循證決策（Evidence Based Policy Making）列為政府預算編列重要參考因素，使得統計分析於政府決策之功能再次被高舉強調，蔚為當今顯學。

內政部統計處有感於大數據分析對於支援政府決策之重要，為

提升同仁對大數據應用的瞭解，帶動同仁思考如何創造內政資料支援決策的價值，於去（106）年 8 月 1 日舉辦「內政統計研討會」，主題為「內政大數據之串聯應用」，邀請花政務次長敬群開幕致詞，並擔任綜合座談主持人，由戶政司張琬宜司長、地政司王靚琇司長、資訊中心施明德主任，及統計處饒志堅處長共同參與與談（圖 1）。



圖 1、106 年內政統計研討會 花政務次長開幕致詞及座談會一隅
（右下圖左起依次為地政司王靚琇司長、戶政司張琬宜司長、花敬群政務次長、統計處饒志堅處長、資訊中心施明德主任）

研討會最後獲致 3 點決議：

1. 進行「內政業務議題與統計資訊診斷」：各單位依業務內容及各項議題檢視統計資訊之需求，以強化各單位對於議題之瞭解及數據分析過程之掌握。

2. 進行「內政統計資料清洗」：各單位檢視現行系統內之資料或結果表單，進行必要之清理與檢討。
3. 成立「內政數據分析與決策應用工作小組」：以機制建立為起點，培養各單位數據分析應用能力，協助重大核心政策或法案之制定。

為落實會議決議，統計處旋與相關單位討論，擬訂「內政大數據連結應用專案計畫」及「內政數據分析及決策應用工作小組設置要點」，由統計處擔任幕僚單位，於 106 年 11 月 16 日提報本部第 20 次部務會報，並於 12 月 14 日奉核定，以台內統字第 1062001031 號函送相關單位施行。此為政府機關中，第一個由統計單位主導成立之部會級大數據工作小組及專案計畫，全力推動相關大數據連結應用，提升數據分析應用價值。

第2節 計畫範圍與架構

本專案計畫提升各單位及所屬機關資料品質與自主分析能量，強化數據支援決策功能為充分發揮數據支援決策功能，提升各業務單位核心資料庫品質與自主分析能量，及拓展跨單位大數據連結應用，達到業務、資訊、統計三合一統合綜效，以強化內政部政策議定、法案推動及議題辯護能力，本專案計畫主要推動工作包括業務議題診斷、業務資料清洗與分析能量提升，及建置內政大數據資料庫，規劃「智

慧型數據分析平臺」，提供內政決策分析。

第3節 計畫功能與特色

本專案計畫依執行先後順序，可概分為議題、資料、分析三大主軸，並力求議題（目標）清楚，資料（手段）明確、分析（結論）具體可行，以發揮數字管理功能。執行過程強調融合各單位專長，相互支援，激發團隊合作綜效。整體內容具下列特點：

1. **建立標準化作業規範**：無論議題發想、與資料連結、清洗、分析皆力求建立標準作業機制，制訂各種標準化文件格式（SOP），留下專業知識（know-how）軌跡，除可提升工作效率，更有利經驗傳承。例如本計畫制訂「業務議題診斷表（A表）」及「業務資料清洗表（B表）」2種主要表（圖2），內容包括議題名稱、背景說明、相關法規、資料需求項目、業務統計、應用統計、公務統計、關鍵指標、決策應用、現存問題、提案單位、主協辦單位、改進措施及預期效益等，藉這些系統化、格式化表格，盤點出重要業務議題及需配合的統計資料，各種資訊一目瞭然。

內政大數據連結應用專案計畫

業務議題診斷(樣張)				
議題編號	A 統 001	提案單位 統計處	主辦單位 統計處	協辦單位 營建署、戶政司
日期	106 年 9 月 14 日			
議題名稱	老屋分布及其影響人數			
背景說明	1. 臺灣目前屋齡 30 年以上老屋有 398 萬戶，部分老舊建築物抗震防災能力不足，可見老屋改建具迫切性，其坐落分布等資訊具政策應用價值。 2. 為應抗震防災救援需求，找出老屋及住於其內之幼年、老年人口正確位址，以備不時之需。			
相關法規	《都市危險及老舊建築物加速重建條例》106 年 3 月 16 日立法院一讀			
重要記事	無			
診斷結果				
①	資料名稱：房屋稅籍檔	資料編號：B 統 001		
	資料項目：房屋屋齡、樓層、結構、地址、[統計區代碼]			
	統計單位：宅、棟			
	業務統計：房屋稅籍統計	權責單位：財政部		
	應用統計：內政部不動產資訊平台-房屋稅籍住宅類數量 http://pip.moi.gov.tw/V2/E/SCRE0401.aspx <住宅供給>	應用單位：營建署		
	公務統計：財政統計年報-房屋稅稅源 https://www.mof.gov.tw/Pages/Detail.aspx?nodeid=185&pid=740728	公統單位：財政部統計處		
資料名稱：戶籍資料	資料編號：無			

內政大數據連結應用專案計畫

業務資料清洗(樣張)				
資料編號	B 統 002	提案單位 統計處	主辦單位 營建署	協辦單位 無
日期	106 年 8 月 14 日			
資料名稱	全國建管系統			
資料項目	使用執照、建照執照、建築執照、基地面積、建築面積、總樓地板面積、建築物高度、地下避難面積、法定空地面積、建造類別、地址、[統計區代碼]			
統計單位	件、面積、工程造價			
業務統計	營建統計資訊 http://www.cpami.gov.tw/chinese/index.php?option=com_categorytable&view=categorytable&categoryid=155&Itemid=102			權責單位：營建署
應用統計	1. 景氣指標及對策信箋: https://www.ndc.gov.tw/News_Content.aspx?n=4BAFF85D273162D4&sn=A93837405548FF94&=F8E6EFP9E8181156 2. 工業生產統計: http://www.moea.gov.tw/mrs/dos/bulletin/Bulletin.aspx?kind=6&rtm=1&menu_id=6725			應用單位： 1. 國家發展委員會 2. 經濟部統計處
公務統計	公務統計報表程式 http://www.cpami.gov.tw/chinese/index.php?option=com_content&view=article&id=160&Itemid=102			公統單位：營建署主計室
盤點結果				
項次	資料項目	起始時點	品質	說明
現存問題	1. 使用執照、建照執照、建築執照...	95 年至 105 年	高	建置為 80 年間，之前資料有缺漏，80~95 年資料未涵括所有縣市資料。
	2. 缺[統計區代碼]	95 年至 105 年	中高	增加[統計區代碼]可產製空間統計資料，結合其他空間產品，提升應用價值。

圖 2、業務議題診斷表 (A 表)、業務資料清洗表 (B 表)

2. **建置集中化取用環境**：截至 106 年底，內政部及所屬機關計有 798 個資訊系統，875 個資料集，項目內容繁多複雜。本計畫配合本部資訊向上集中（軟硬體集中管理）政策，建置資料集中化取用環境，將相關資料集作一統整。此項作為具有節省取用成本（避免重複購置軟硬體）、齊一統計定義標準（一致性資料定義及背景資料說明）、提升資料品質（可相互勾稽除錯）、增進管理效能（有效分層權限系統管理）及強化資訊安全（提升資安等級）等優點，串成一個大型資料集，大幅增加交叉應用項目，提升資料應用範圍，發揮大數據之價值。

3. **建立全面化團隊合作**：為落實本計畫之執行，特別成立前述「內政數據分析決策應用工作小組」，由花政務次長親任召集人，相關部內單位主管及所屬機關副首長擔任委員，提升執行層次、強化執行力道，統計處處長則為執行秘書，負責實際業務推動。另設計跨單位提案機制，將大數據分析工作推展至部內各單位及所屬，強化單位間橫向合作。例如：可由統計處提出議題需求，請營建署主辦，統計處、地政司協辦（當然也可以由業務單位提案，統計處主辦），只要工作小組討論通過即據以執行，此項設計重點在強調靈活彈性，並兼具務實合理。

第2章 文獻回顧

第1節 國內

1.交通部高速公路局 ETC 資料分析

以 ETC 大數據資料預測高速公路的車流量走勢，以適當匝道儀控紓解塞車，進一步運用建立交通預報與預警機制，以及提升服務品質。

2.勞動部「大專畢業生就業追蹤系統」

以大專院校畢業生基本資料，勾稽勞保等投保資料檔，及替代役、在學及入出境等資料，俾了解大專畢業生就業流向及起薪，以規劃就業政策及輔導。

3.衛生福利部食藥署「食藥戰情中心」

透過大數據管理、巨量資料分析技術導入，協助建立我國食品安全風險管理與危機處理機制模式，落實源頭管理並保障全體國民健康。

4.警政署「全國毒品情資資料庫」

運用大數據分析毒品流通、交易網絡，勾勒完整毒品犯罪圖像，以向上溯源等策略，強力掃蕩毒品。

第2節 國外

1.美國_證據基礎決策委員會 Evidence-Based Policymaking Commission

93年參眾兩院預算委員會主席商議立法成立「證據基礎決策委員會 (Evidence-Based Policymaking Commission)」，研究如何充分運用資料評估聯邦政府計畫及稅式支出之效益，以作為預算編立之參據，以及未來修法加強保護個資並排除資料運用障礙之參考，105年3月國會通過法案，由總統歐巴馬簽署實施。歐巴馬曾說：透過數據之掌握與創新可以使我們所做的變得更好，他主張設立“大數據創新生態系統 (Big Data innovation ecosystem)”，讓更多聯邦機構在此平台提取並分析資訊，培養制定決策的能力。並提出「為統計目的提供並使用公務登記資料指引」備忘錄，要求各機關首長促進計畫與統計單位間的密切合作，發展堅實的數據管理策略。而美國為加速落實與實踐，眾議院更於106年11月通過「證據基礎決策基本法 (Foundations for Evidence-Based Policymaking Act)」，規範聯邦機關應提出證據建立計畫、設置評估長，以統籌機關之政策評估與分析作業。

2.荷蘭_城市資料中心 Urban Data Center

93年荷蘭新統計法開始施行，賦予「荷蘭中央統計局」獨立機

構的地位，並賦予其擁有所有單位資料的權限，所有政府機關全部行政資料都要交給統計局。105 年統計局成立了城市數據中心，讓 180 個全國性的公務登記資料可以連結，成為各大都市市政府的「資訊幕僚」，因為一離開中央，荷蘭有 95% 的市政府面臨資料能力不足窘境。城市數據中心就是要作為中央與地方資料的橋樑，把中央與地方的資料都整理在一起，進而幫助地方政府提升施政服務品質。

第3章 推動方式

第1節 成立工作小組

以機制建立為起點，培養各單位數據分析應用能力，協助重大核心政策或法案之制定。

為落實會議決議，統計處旋與相關單位討論，擬訂「內政大數據連結應用專案計畫」及「內政數據分析及決策應用工作小組設置要點」，由統計處擔任幕僚單位，於106年11月16日提報本部第20次部務會報，並於12月14日奉核定，以台內統字第1062001031號函送相關單位施行（附錄1）。內政數據分析及決策應用工作小組置委員十人，其中一人為召集人，由本部政務次長兼任；一人為副召集人，由本部主任秘書兼任；其餘委員由下列人員派兼之（圖3）。此為政府機關中，第一個由統計單位主導成立之部會級大數據工作小組及專案計畫，全力推動相關大數據連結應用，提升數據分析應用價值。

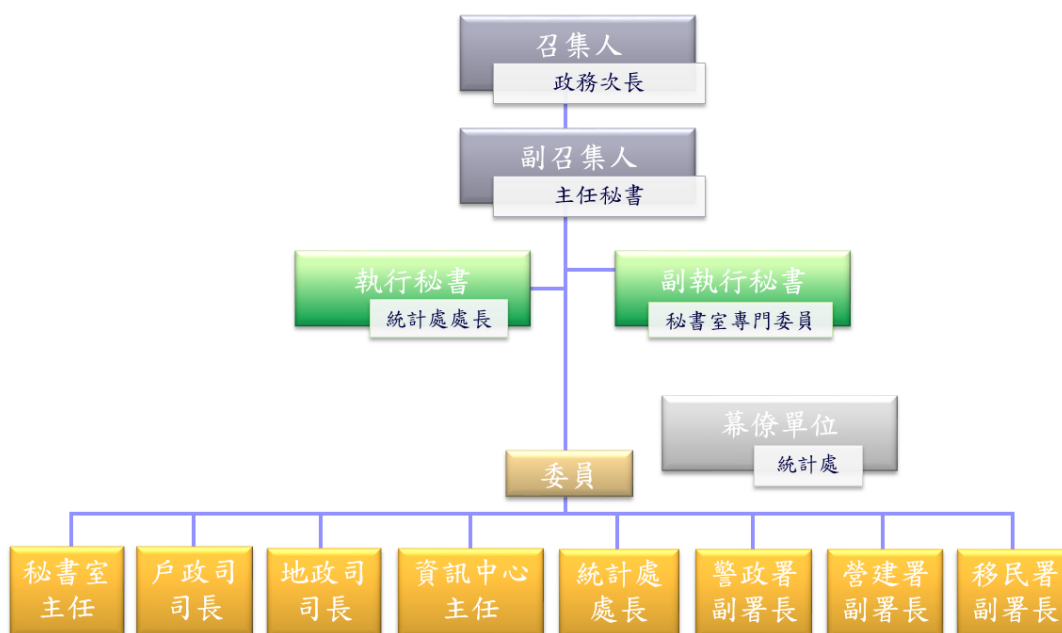


圖 3、內政數據分析及決策應用工作小組架構圖

第2節 辦理教育訓練

為使本部暨所屬機關同仁了解本計畫之重要性、策略架構、執行方法，並思考經辦業務與本部重要政策議題之關聯性，進而透過數據分析精進施政作為與政策論述，以強化本部政策制定、法案推動及議題辯護能力，爰推動辦理下列教育訓練：

1. **大數據分析工作坊**：課程強調團隊合作，將業務、統計、資訊 3 類人員混合分組，參訓學員 40 名，分別來自本部及所屬機關計 11 個單位（圖 4），透過討論、實作、成果報告方式，發揮學中做、做中學之精神，將訓練課程與業務需求結合，以客製化教材進行 Tableau 視覺化分析軟體講授與實機操作，激發分析興趣與熱情，全程並有本處派員輔導進行，此課程業於本（107）年問卷調查，

學員就「內容對目前或未來辦理的業務有幫助」及「對提升個人專業知識有所助益」之滿意度分別達 89.2%及 94.6%。



圖 4、內政大數據分析工作坊（第 1 期）由沈主任秘書淑妃（前排右 6）主持開訓

2.大數據技術講座：邀請產官學界專精人士，以專題講授方式進行大數據技術實例探究，俾利獲得最新資訊並啟發業務應用方向。本年 4~6 月間先分別由邱泊寰博士、楊雅惠總經理及梁德馨教授完成「敏捷式有效議題開發」、「輿情探勘之統計原理」及「如何用對統計方法作對數據分析」3 場次講座（圖 5），本部及所屬機關參加者超過 200 人次，深獲好評，滿意度皆超過 9 成，對本部未來如何開發有效議題、如何將輿情探勘技術與民意調查進一步結合運用，以及如何作出正確的數據分析成果咸具助益。下半年將持續根據本計畫進程，適時安排相關講座，積極協助同仁解決大數據分析遭遇之困難。



圖 5、107 年內政大數據分析技術講座

第3節 規劃大數據試辦委外案

根據針對本部大數據工作坊 40 位學員所作之「你覺得做數據分析最痛苦的是甚麼？（複選，最多 3 項）」調查結果（圖 6），認為「3.【錯誤百出】資料亂七八糟的」所占比率最高（65%），其次為「4.【訂出議題】不知道要做什麼才是真有用的」亦占 50%；顯示「資料清洗」及「找出議題」為當前推動大數據分析最需解決之問題。由於資料清洗（data cleansing）占大數據分析工作量 6~8 成，耗時耗工，往往非本部現有人力所能承擔，需藉由外部力量來協助完成，因此本計畫特別針對較複雜費工之連結應用，規劃辦理委外服務案，期引進專業資訊服務，協助本部進行議題與資料盤點、資料庫清洗、資料運用分析，並進行 2 項前導案例之實作，以達上揭專案計畫推動目標。

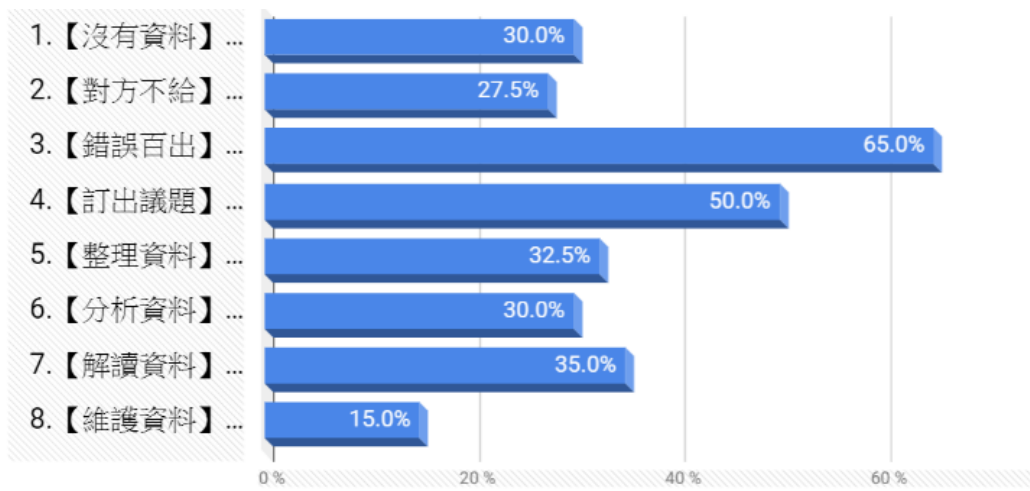


圖 6、你覺得做數據分析最痛苦的是什麼？（複選，最多 3 項，N=40）

第4章 主要成果

第1節 教育訓練成果

雖然本專案自去年底正式施行以來尚未滿 1 年，但透過詳細規劃、密集推動、各單位全力配合，及長官的指導支持，已召開 2 次工作小組委員會議、1 次工作會議、3 場技術講座、1 項委外服務案、1 期工作坊及成果發表會。成果發表會具體案例包括人口（戶政）、交通事故（警政）、新住民之群聚（移民）、地籍清理（地政）、人民團體管理（合團）、昇降設備管理、中低收入戶住宅需求（營建）、火災高風險對象（消防）及危老建築改建（綜合）等 9 篇報告，每篇皆詳述議題名稱、動機背景、資料來原、主要視覺化 Tableau 分析結果及未來發展，並由該單位（機關）長官指導評述，咸表成果豐碩並具業務參考價值。

一、工作坊教育訓練 9 篇成果：

1. 模擬城市 M：

選定一模擬城市，利用地理位置及人口分布，搭配人口特性，包含：男女比、婚姻狀況、年齡、教育程度等，歸納分析該模擬城市各區所需公共設施及政府政策，以利提供政府相關機關施政參考，並可結合民間資源推動相關措施（圖 7）。

模擬城市M

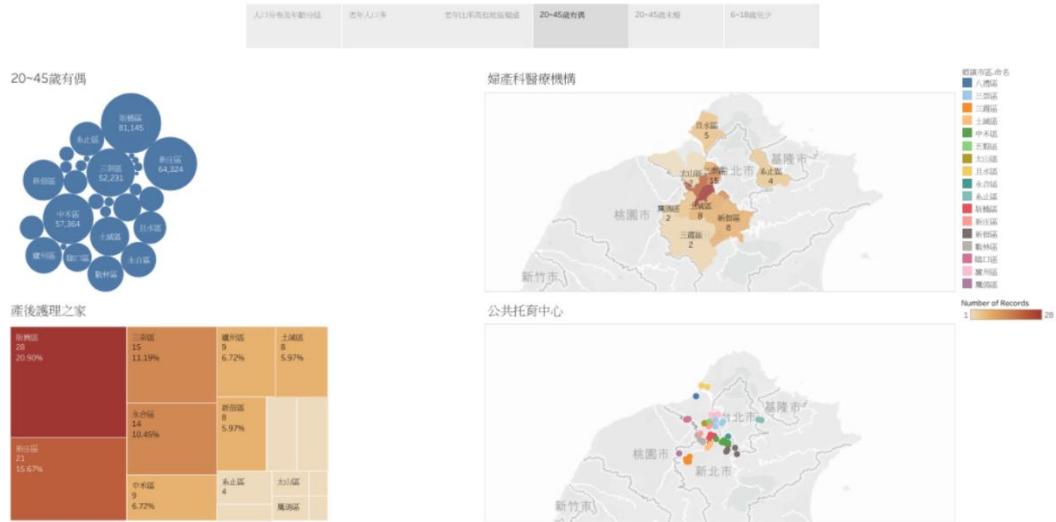


圖 7、內政大數據分析工作坊成果—模擬城市 M

2. 道路交通安全-以 A2 類交通事故概況為例：

以 105 年發生 A2 類交通事故 30 萬筆資料，找出易肇事路段、時段與車種，作為勤務規劃依據（圖 8）。



圖 8、內政大數據分析工作坊成果—道路交通事故預防

3. 新住民是否有群聚現象？新住民居住之屋齡及房價情形為何？

新住民曾為外籍勞工之人數有多少？

以新住民人數與房屋稅籍檔及實價登錄交易資料進行比

對，分析新住民人數分布及變動情形，並瞭解居住屋齡及房價等資訊。另就外籍勞工申請案資料及新住民資料進行比對，瞭解新住民曾為外籍勞工之人數及比率（圖 9）。

新住民是否有群聚現象？其居住之屋齡及房價情形為何？曾為外籍勞工之人數有多少？

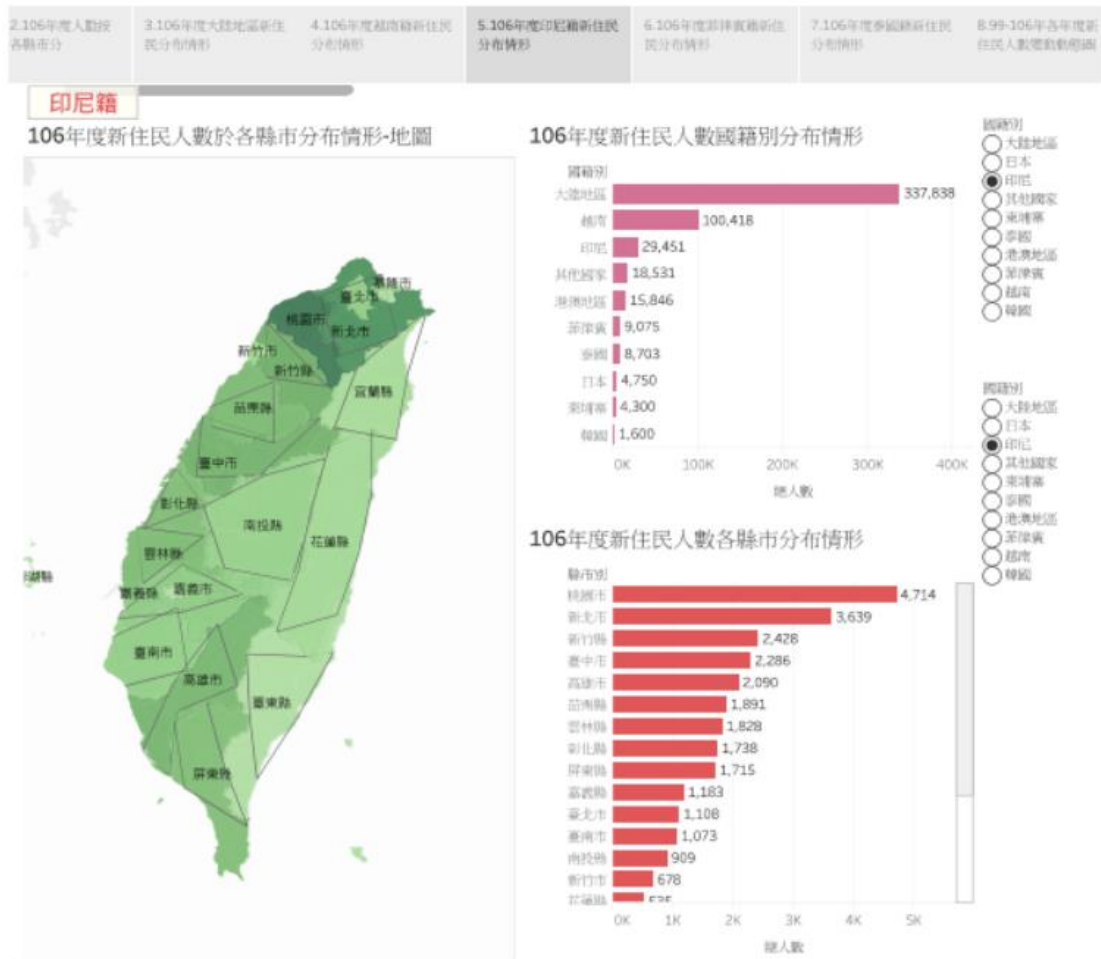


圖 9、內政大數據分析工作坊成果—新住民群聚

4.辦理地籍清理執行成效-以新北市為例

臺灣光復之初期有地籍登記不完整或與現行法令規定不符之情形，影響土地之有效利用，亦妨礙民眾財產權利之行使，以新北市近年地籍清理成果了解標售及囑託國有成果，並分析

本制度執行成效。

5.政府天秤的管制重量：哪些人民團體應公開財務狀況？

內政部為活化公民社會，將朝廢人民團體法，另立政黨法、社會團體法與職業團體法等三法取代，其中，「社會團體法草案」已草擬完成，除鬆綁社會團體成立門檻外，管理手段也由「行政介入」轉為資訊公開的「全民監督」。由於社會團體可享有賦稅優惠、參與公共事務及可向社會募款，為防止發生逃漏稅或產生其他弊端，依照草案內容，未來社會團體除會務、涉及個人隱私等相關資訊外，每年達一定預算規模以上社會團體每年度工作計畫、預算、財務報表，以及所接受補助、捐贈，與支付獎助及捐助等資訊，都要公開刊載，上網公告供民眾線上查詢閱覽，強化公共監督功能，本次利用各級人民團體活動概況調查所抽樣團體，分析社會團體及職業團體經費資源分配，找出財務透明應鎖定何種團體。

6.為防範升降設備維護保養紀錄不實情事，及確保升降設備確實定期保養，以維護公共安全

分析目前各縣市各類型升降設備比例，瞭解各縣市政府升降設備到期月份件數，推估升降設備檢查人所需量，並透過各縣市政府維護廠商所占比例及件數，從各廠商維護人員數推斷

是否足夠，以加強昇降設備維護保養，提升管理機制及法制作業，掌握昇降設備維護情形。

7.居住於土壤液化潛勢區老舊住宅低收入戶，地方政府應優先輔導入住公營社會住宅

現階段各地方政府刻正陸續興辦社會住宅，初步提供之戶數尚難於短時間內大量釋出，對於提供服務之政府，應評估潛在之需求對象(客戶)，爰統計低收入戶，居住於土壤液化潛勢區之老舊住宅情形，由地方政府社福單位優先宣導、關懷及輔導入住，利用財政部房屋稅籍檔、衛生福利部中低收入戶檔及國家災害防救科技中心(NCDR)土壤液化潛勢區資料找出 106 年底新北市 29 行政區之潛勢區屋齡 30-80 年之低收入戶。

8.從火災統計檢視火災高風險對象

分析統計各縣市火災發生頻繁之種類(森林田野及建築物火警)及易致人命傷亡(建築物)火警之原因，以建築物火警案件最多的大台北地區探究火災原因相關的「人為」及「外在環境」因素，得知火災高風險對象(學齡前兒童及高齡)。

9.臺北市土壤液化潛勢區內建物及人口分布

利用國家災害防救科技中心(NCDR)提供之房屋稅稅籍檔加值資料計算方式，以地址欄位與建築物資料進行勾稽串連，

取得設籍於民國 89 年以前和 90 年以後完成之建築物內的人口數、戶數，完成本系統，期助益保障民眾生命及財產安全，提升整體居住品質。

9 篇報告雖然內容豐富，惟由於 3 週工作坊時間有限，無法完全連結分析各項資料，殊為可惜，因此請業務單位依評述意見再予精進，提報後續工作會議。

二、萃取精華延伸分析：

由各篇評述內容發現，其中 4 篇發展潛力雄厚，可列為重點議題項目，深入研析，提升應用價值，包括：

1. 無殼青年家庭與家庭組織型態：連結戶籍與房屋稅籍資料，探討資料議題包括 20-45 歲青年家庭人口數、無殼家庭戶數、房屋所有人年齡、性別、教育程度、地域別、樓層別、屋齡別分布(全國房屋所有人中年紀最小的是幾歲？20 歲以下擁有房屋者人數？有無性別差異？地域別差異？獨居老人居住於 5 樓公寓之人數？鄰近有無醫院或 119) 及家庭組織型態地域別分布(單人家庭、單親家庭、老人家庭、獨居老人家庭之變化？居住環境)。可供社會住宅及人口生育政策參用。
2. 新住民群落與居住生育狀況：連結新住民資料、房屋稅籍檔與戶籍檔，探討資料議題包括第 1、2、3 代新住民人口數、地域

分布（有無群聚特性？特性為何？）；新住民無殼家庭戶數、房屋所有人年齡、性別、教育程度、地域別、樓層別、屋齡別分布（住宅面積與非新住民有無顯著差異？）及新住民家庭組織型態（是否與非新住民有顯著差異？是否單親家庭較多？）。可供移民署管理及照顧政策參用。

3. **交通事故天地人事特性預防**：連結事故檔（事）、戶籍檔（人）、氣象檔（天）、GIS 路網圖（地），探討資料議題包括交通事故件數、死亡人數、受傷人數按事故別、第一當事人年齡、性別、教育程度、婚姻狀況、車齡、駕駛年資、降雨量、日照方向、事故時間、地域、地點分（酒駕與非酒駕地點是否有差異？男性與女性事故地點是否有差異？好發事故地段之特性？不同時段事故地點是否有差異？天候、時段與地點之關聯性？國道與一般道路事故差異性？），供為研擬更適切道安改進措施及因應對策。

4. **潛在風險住宅分布情形**：連結戶籍檔、房屋稅籍檔、斷層圖、土壤液化潛勢區圖、順向坡圖、違建檔、實價登錄檔、中低收入戶檔、找出潛在風險較高建築（屋齡大於 50 年、5 層樓公寓、非鋼筋結構、位於斷層帶、土壤液化高潛勢區、順向坡、頂樓加蓋違建、內住老年人口、附近房屋交易頻繁、中低收入

戶)之分布，供為改建政策參用。

本專案計畫業將上述 4 篇議題列為重點項目，其中第 1、2 項涉及資料清洗，將於後續「大數據試辦委外案」優先執行，第 3、4 項之後續研析則分別由警政署（統計室）及統計處主責辦理，預計本年底可提出完整報告，達成統計支援決策之目的。

三、技術講座成果：

1.敏捷式有效議題開發

為使同仁了解內政大數據連結應用過程中，如何開發好議題，於本（107）年 4 月 11 日舉辦本年度第一場內政大數據分析技術講座，邀請卡米爾股份有限公司邱泊寰博士及其團隊進行案例說明及互動討論，各單位參加同仁共 69 位。

課程由國外政府單位如何以大數據資料結合人工智慧處理相關議題的案例導入，並釐析開發有效議題的四大關鍵包括「尋找目標群眾」、「使用在什麼樣的情境」、「需使用什麼樣的資料」、「可以達成什麼效益」，接著分組討論實作，依學員業務屬性分為「消防營建」、「地政秘書」、「警政移民」、「戶政民政」4 組，並將統計、資訊人員混編在各組，由該公司成員帶領並擔任各組小組長，實際藉由腦力激盪共同討論方式，尋找有效益的業務議題。

之後進行分組報告，包括「使用人工智慧產製各項業務應答集」、「預測溺水者流向」及「解決殯葬業詐騙問題」等，上述議題皆是利用大數據來協助第一線工作人員減輕工作負荷，並讓民眾對施政更有感，創造政績亮點。本次技術講座著重實際操作練習，祛除心裡疑懼，激發學員開發議題及工具運用的潛能，逾 8 成學員認為對業務有所助益，並表達參加後續技術講座的高度意願，最後在彼此回饋的心得分享後圓滿結束（圖 10）。



圖 10、107 年第 1 場內政大數據分析技術講座「敏捷式有效議題開發」上課情形

2.文字探勘技術在公部門的應用及陷阱

為使同仁瞭解大數據分析實務上相關經驗及案例，於本（107）年 5 月 22 日舉辦本年第 2 場內政大數據分析技術講座，包括參（競）賽經驗及文字探勘應用分享，各單位共 65 位同仁參加。

本次活動先由戶政司分享本年參加總統盃社會創新黑客松參賽的經驗，參賽題目為「人口少子化及高齡化地圖」及「尋

找消失的邊緣人」，思考將戶政的服務推廣至神經末梢；接著由本處說明「2018 年資料創新應用競賽—社會經濟空間資料暨戶政應用」的經過，希望透過競賽鼓勵產、官、學界以挖掘、重組、混搭等方式，產製實用又富創意之作品，經引用後提升政府及企業之服務品質。

續由浚鴻數據開發股份有限公司楊雅惠總經理演講「文字探勘技術在公部門的應用及陷阱」，開場先敘明大數據分析與統計調查、資料採礦間之差異，探究大數據分析之優勢及分析過程中可能遭遇的問題，並結合文字探勘與網路輿情分析所創造的優勢，以及公部門的應用案例，精湛的演講內容啟動了學員思考結合文字探勘與業務的應用靈感（圖 11）。



圖 11、107 年第 2 場內政大數據分析技術講座上課情形

3. 如何用對統計方法作對數據分析

從基礎統計概念將重要的統計觀念融入簡易的統計技術

操作，呈現多元面向易判讀的統計圖，輔以實例作說明，加強同仁在作統計分析時，透過圖表正確的呈現，增強數據分析的概念（圖 12）。



圖 12、107 年第 3 場內政大數據分析技術講座上課情形

3 場技術講座本部及所屬機關參加者超過 200 人次，深獲好評，滿意度皆超過 9 成，逾 9 成學員認為對業務有所助益，並有逾 8 成學員表達對後續技術講座有高度的參與意願，對本部未來如何開發有效議題、如何將輿情探勘技術與民意調查進一步結合運用，以及如何作出正確的數據分析成果咸具助益。

第2節 自行研究成果

一、道路交通安全肇事原因之地域與個人因素探討

藉由相關事故資料之串聯找出易肇事路段發生事故之各種危險因子，除擴大交通安全宣導及強化各項交通執法，作為

勤務規劃依據，更進一步提醒用路人小心駕駛，減少事故發生。先以國道交通事故分析，並以百萬車公里交通量進行相關統計分析，觀察事故發生主要路段為何及事故發生之原因、天候、車種別等；另交通事故與用路人違規是否存在相關性之分析，試以臺北市 A1 及 A2 類道路交通事故資料及違規資料進行大數據分析，提供防制道路交通安全及警察加強執法或制定交通罰則之參考（圖 13）。將分析結果製作成影片「道路交通安全肇事原因之地域與個人因素探討」，藉由簡潔易懂的統計圖表及活潑的內容告訴民眾道路交通安全肇事原因分析結果，並將影片置於網站（網址：<https://www.npa.gov.tw/NPAGip/wSite/public/Data/f1534908478321.mp4>）以達宣達效果（圖 14）。

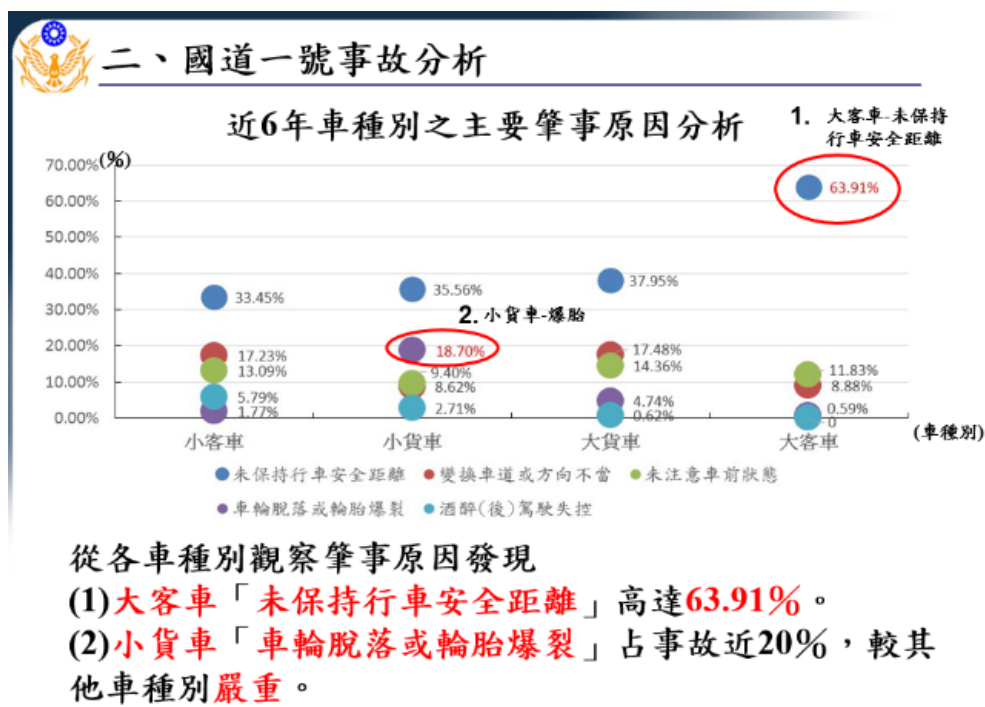


圖 13、「道路交通安全肇事原因之地域與個人因素探討」分析成果

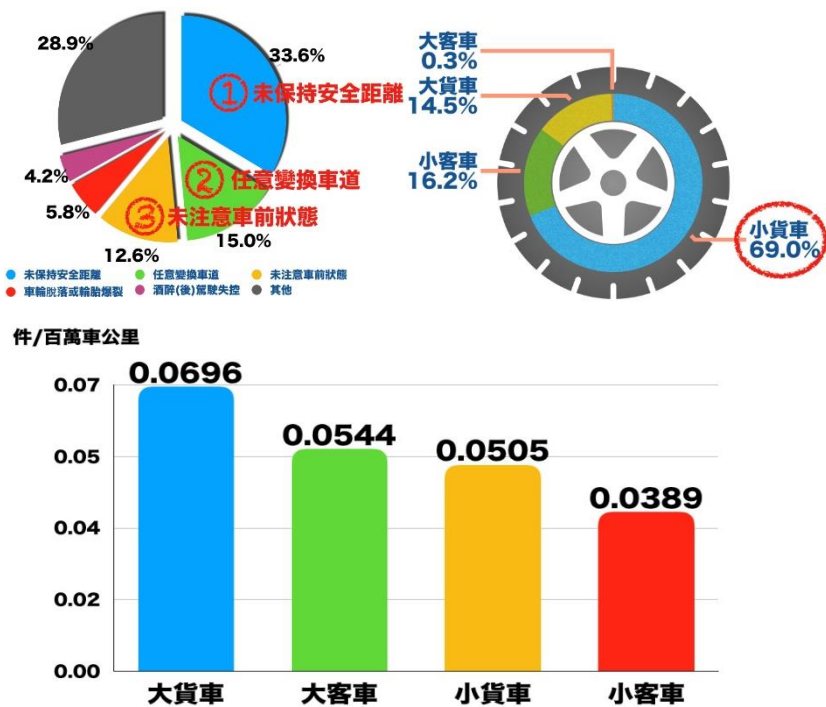


圖 14、「道路交通安全肇事原因之地域與個人因素探討」影片中統計圖

二、潛在風險住宅分布情形探討

連結戶籍檔、房屋稅籍檔、違建待拆檔、土壤液化潛勢區圖、順向坡圖、斷層帶圖，採用門牌比對及圖層套疊，找出潛在風險較高建築（屋齡大於 50 年、5 層樓公寓、非鋼筋結構、位於斷層帶、土壤液化高潛勢區、順向坡、頂樓加蓋違建、內住老年人口、附近房屋交易頻繁、中低收入戶）之分布。本分析以臺北市為例，找出位於活動斷層帶 500 公尺內之高土壤液化趨勢區，且建物結構為非鋼構、屋齡 40 年以上者，全臺北市共 113 棟（計 413 人居住），可供危老改建政策參用。未來可將資料拓展全國各縣市資料串接，或增加房地買賣熱區，找

尋都市更新最適區域（圖 15）。

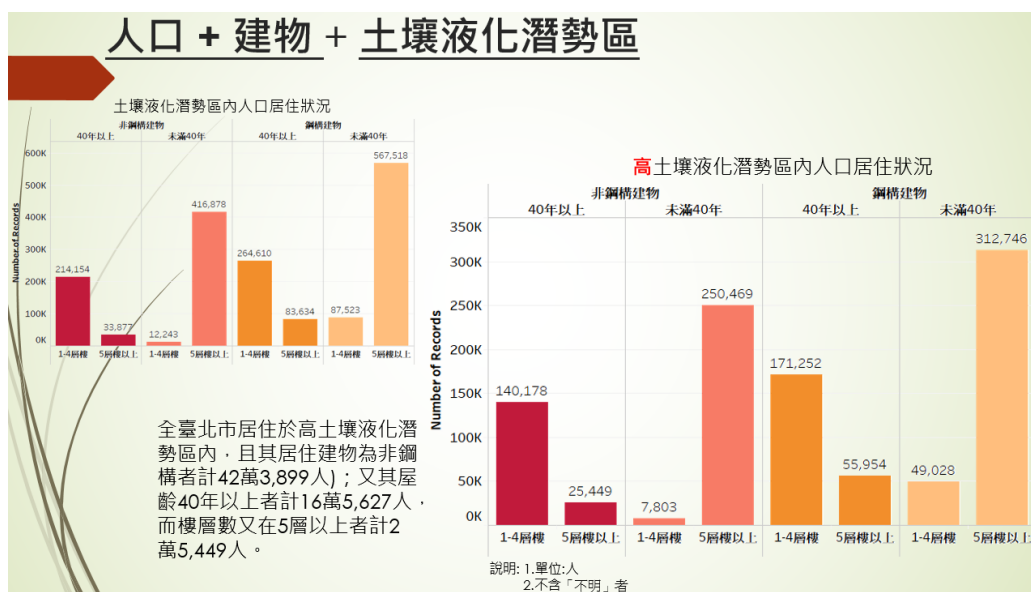


圖 15、「潛在風險住宅分布情形探討」分析成果

第3節 規劃大數據試辦委外案

建立議題與資料盤點作業模式：引進專業管理服務，協助本部進行議題與資料盤點、資料庫清洗、資料運用分析，建立數據與政策議題連結模式（SOP），並提供專案管理，有效推展本部及所屬機關跨單位、跨領域資料串聯應用。

一、議題設定需包含兩大重要面向：

（一）目標設定

目標設定的意義在於梳理現行分析需求，進而發想未來三到五年內可以進行的議題。

（二）機會掃描

基於實務上，時間上與人力上的限制，上述發想的議題需進行可行性之評估（如：議題之成熟度分析）及重要性與輕重緩急之盤點，方能集中現有資源，進而完成優先議題之執行。針對第一階段之執行項目—「議題設定」，規劃標準作業流程。

二、資料盤點

在執行各式資料分析前，應該對資料分析之基礎，也就是「資料」有完整的了解，因當無用的資料進入分析方法內，只會產出無用的分析結果，反而對其結果形成反效果。因此，對於議題研究時所需用到之資料的通盤了解，就更為重要。另外，為了因應例如「個人資料保護法」或「個人資料保護法施行細則」等法規之規範，更應對於所需使用之資料採取適當之解決對策或控制措施，以有效降低資料檔案遭受濫用或損害之風險。對於資料之管理、安全性、資料結構與其所有人或使用單位之管控，規劃資料盤點之方法與流程。

透過以上論述架構，得到本次 2 項主題下之各項分析議題與 KPI（圖 16）及所需連結資料項目簡表（圖 17）。

主題	政策	編號	議題	KPI(子議題)	較表	備註
無殼蝸牛	社會住宅、包租代管	S01	青年無殼情形是否嚴重？	青年無殼族人數、青年無殼比率	無自有住宅人數按年齡組別分	1.無殼定義有多種 2.其他子議題：30歲內有殼者人數、寄放學區大戶(一宅多戶、未成年戶長)
		S02	住宅權屬是否有性別傾斜現象？	住宅權屬性別比例、無殼青年性別比例	住宅權屬按性別分、S01+按性別分	
		S03	哪些地區無殼情形較為嚴重？	6都青年無殼比率、全國青年無殼比率最高之縣市、鄉鎮市區、村里	S01+按區域別分	
		S04	哪些家庭無殼情形較為嚴重？	核心家庭(夫妻+小孩)無殼比率	S01+按家庭組織型態別分	
		S05	弱勢者與非弱勢者無殼差距？	弱勢者無殼比率	S01+按弱勢身分別分	
人口少子化	生育補助、育嬰津貼	S06	有房者較(不)願意生育小孩？	已婚有殼者與已婚無殼者平均幼年子女數比	家庭15歲以下子女數按住宅權屬及家庭組織型態分	1.結婚後幾年生第一胎、結婚後幾年有房子按已婚者年齡分，可分析歷年變化趨勢。 2.另可研究城鄉人口變遷議題(遷入最多之鄉鎮市區村里、遷出最多之鄉鎮市區村里及其人口特性、遷入後幾年有房子、與當地房價之關聯性)
		S07	生育率影響因素為何？	生育率影響因素勝算比	生育率影響因素回歸分析(住宅權屬、住宅價格、族群別、婚齡..)	
		S08	哪些地區少子化情形較為嚴重？	6都已婚者平均幼年子女數、全國平均數最低之縣市、鄉鎮市區、村里	S06+按區域別分	
人口高齡化	銀髮照顧	S09	老年人口上下樓方便性？	住5樓公寓老年人口數	老年人口數按門牌樓層及建物樓層分	
		S10	老年人口家人在宅照顧？	與子女同住之老年人口數	老年人口數按家庭組織型態分	
		S11	老年人口家人鄰近照顧？	未與子女住同但同縣市之老年人口數	S10+按是否和子女戶籍同縣市分	
		S12	哪些地區老人上下樓較不方便？	6都住5樓公寓老年人口比率、全國比率最高之縣市、鄉鎮市區、村里	S09+按區域別分	
		S13	哪些地區老人人口家人照顧較多？	6都與子女同住之老年人口比率、全國比率最高之縣市、鄉鎮市區、村里	S11+按區域別分	
建築高齡化	危老改建、都市更新	S14	高齡建築之影響？	高齡建築人口涵蓋率？	人口數按屋齡、樓層、結構別分	潛在風險指數Top100建築僅供內部參考，不對外公佈。
		S14-1	潛在風險建築之影響？	潛在風險建築人口涵蓋率？	S14+按斷層帶、土壤液化、連連區分	
		S15	危老改建熱區？	潛在風險指數Top100建築	依S14-1計算每棟樓之潛在風險指數	
		S16	都市更新熱區？	都市更新指數Top100建築	依S15+每坪房價計算最小統計區都市更新指數	
新住民無殼蝸牛	社會住宅、包租代管	S17	青年無殼情形是否嚴重？	青年無殼族人數、青年無殼比率	S01+新住民別	可藉比較新住民與非新住民之KPI值，了解新住民與非新住民之差異。
		S18	住宅權屬是否有性別傾斜現象？	住宅權屬性別比例、無殼青年性別比例	S02+新住民別	
		S19	哪些地區無殼情形較為嚴重？	6都青年無殼比率、全國青年無殼比率最高之縣市、鄉鎮市區、村里	S03+新住民別	
		S20	哪些家庭無殼情形較為嚴重？	核心家庭(夫妻+小孩)無殼比率	S04+新住民別	
		S21	弱勢者與非弱勢者無殼差距？	弱勢者無殼比率	S05+新住民別	
新住民人口少子化	生育補助、育嬰津貼	S22	有房者較(不)願意生育小孩？	已婚有殼者與已婚無殼者平均幼年子女數比	S06+新住民別	可由S23了解新住民生育率影響因素與非新住民之差異
		S23	生育率影響因素為何？ (哪些地區少子化情形較為嚴重？)	生育率影響因素勝算比	S07+新住民別	
新住民人口高齡化	銀髮照顧	S24	老年人口上下樓方便性？	住5樓公寓老年人口數	S09+新住民別	
		S25	老年人口家人在宅照顧？	與子女同住之老年人口數	S10+新住民別	
		S26	老年人口家人鄰近照顧？	未與子女住同但同縣市之老年人口數	S11+新住民別	
		S27	哪些地區老人上下樓較不方便？	6都住5樓公寓老年人口比率、全國比率最高之縣市、鄉鎮市區、村里	S12+新住民別	
新住民建築高齡化	危老改建、都市更新	S29	高齡建築之影響？	高齡建築人口涵蓋率？	S14+新住民別	潛在風險指數Top100建築僅供內部參考，不對外公佈。
		S29-1	潛在風險建築之影響？	潛在風險建築人口涵蓋率？	S14-1+新住民別	
		S30	危老改建熱區？	潛在風險指數Top100建築	S15+新住民別	
		S31	都市更新熱區？	都市更新指數Top100建築	S16+新住民別	
新住民群聚	文化推展	S32	是否有群聚特性？	新住民人口密度最高之縣市、鄉鎮市區、村里	人口數及密度按新住民別及鄉鎮市區別分	新住民指與國人結婚之陸港澳、外籍、無國籍及已入我國籍者
新住民婚姻	家庭輔導	S33	是否離婚再嫁率高？	離婚率、再嫁率	離婚率、再嫁率按新住民別分	

圖 16、內政大數據分析主題、議題與 KPI 彙整表

一、人：戶政+移民+弱勢 (76)

項目	欄位名稱	格式	說明
A01	縣市	X2	
A02	鄉鎮市區	X4	
A03	村里	X5	
A04	戶流水號	X8	
A05	人流水號	X4	
A06	稱謂	X2	簡化(參考家庭收支調查)
A07	性別	X1	
A08	出生年	X3	
A09	出生月	X2	
A10	出生地	X2	
A11	教育程度	X2	
A12	婚姻狀況	X2	
A13	原住民身分	X2	
A14	新住民身分	X4	含國別
A15	家庭組織型態	X2	參考戶口普查
A16	身心障礙	X2	
A17	中低收入戶	X2	
A18	遷徙次數	X2	
A19	最近遷入年	X3	
A20	最近遷入月	X2	
A21	最近遷出縣市	X2	
A22	最近遷出鄉鎮市區	X4	
A23	結婚次數	X2	
A24	最近結婚年	X3	
A25	最近結婚月	X2	
A26	子女數	X2	含非同戶籍子女
A27	7-14 歲子女數	X1	含非同戶籍子女
A28	未滿 6 歲子女數	X1	含非同戶籍子女
A29	門牌樓層	X2	
A30	子女戶籍同縣市	X1	
K01	身分證字號		資料處理連結完成後刪除
K02	門牌地址		資料處理連結完成後刪除

圖 17、內政大數據人房地資料連結簡表-1

二、房：房屋稅籍+違建(雙北) (47)

項目	欄位名稱	格式	說明
B01	建照年	X3	
B02	建物樓層	X2	
B03	建物結構	X3	
B04	建物坪數	X8	
B05	土地坪數	X8	
B06	房屋權屬	X2	1. 本人擁有 2. 其他戶內人口擁有 3. 非本人及戶內人口擁有，但為戶內人口之直系親屬擁有 4. 以上皆非
B07	違建註記	X2	
B08	擁有房屋棟數合計	X3	
B08	擁有房屋坪數合計	X8	
B08	擁有土地坪數合計	X8	
K01	身分證字號		資料處理連結完成後刪除
K02	門牌地址		資料處理連結完成後刪除

三、地：最小統計區+斷層+土壤液化+順逆向坡+易淹水區+實價登錄+緊急避難處所 (25)

項目	欄位名稱	格式	說明
C01	最小統計區	X2	
C02	斷層帶註記	X2	
C03	土壤液化註記	X2	
C04	順逆向坡註記	X2	
C05	易淹水區註記	X2	
C06	每坪房屋價格	X8	
C07	緊急避難處所註記	X7	離最近緊急避難處所公尺數
K02	門牌地址		資料處理連結完成後刪除

圖 17、內政大數據人房地資料連結簡表-2

第5章 效益評估

本專案計畫之目的為「充分發揮數據支援決策功能，提升各業務單位核心資料庫品質與自主分析能量，及拓展跨單位大數據連結應用，達到業務、資訊、統計三合一統合綜效，以強化本部政策議定、法案推動及議題辯護能力。」推動迄今將近一年，功能效益漸次顯現：

第1節 創新性

1. 首創部級大數據工作小組：本案為政府機關中第一個由統計單位主導成立之部會級大數據工作小組，制度化提升執行層次、強化執行力道，顯示內政部對大數據分析工作之重視。
2. 首創部級大數據工作專案計畫：本案為政府機關中第一個由統計單位主導成立之部會級大數據專案計畫，詳列工作項目、推動方式及獎勵措施，顯示內政部全力推動大數據分析工作之決心。
3. 建立系統性資料連結作業模式：為發揮資料連結綜效，避免重複建置、清洗資料、浪費國家資源，資料連結管理是必然的趨勢。本案從內政部及所屬機關資料連結開始，建立一個有效率、安全的應用環境，並創造資料連結應用最大價值。

第2節 應用性

1. 藉本計畫充實相關公務統計資料：公務統計原為支援政府決策之重要項目，應隨施政重點不斷更新修正。惟實務上，業務單位基於諸多原因，無法迅速充實相關統計項目；今透過本專案計畫之實施，統計單位可藉共同參與，站在決策高度，掌握相關業務數據品質及產製情形，適時建立相關公務統計資料。
2. 藉本計畫建立大數據人員訓練機制：本計畫非常重視人才培育，分析技術（know-how）生根，統計處規劃了嚴謹完整之教育訓練課程，並安排4位專員擔任輔導員，全程實際參與解決問題，深獲各單位好評。長官並指示「建立機制、持續辦理」，表達肯定。
3. 藉本計畫強化數據支援論述功能：本計畫從「議題盤點」開始，以「強化政策議定、法案推動及議題辯護能力」為目的，強調數據分析不是僅將資料（data）研析萃取所得到的資訊（information）為滿足，更是要將其轉化為決策智慧（wisdom），以發揮支援本部政策論述功能為最高指導原則。

第3節 效益性

1. 發揮組織橫向統合綜效：本計畫將各單位之專長，包括業務單位 domain knowledge（業務知識）、資訊單位 information

technology (資訊技術)、統計單位 statistics methodology (統計方法), 以「成立工作小組、共同執行專案」3合1方式緊密連結, 充分發揮團隊戰力。

2. 提升各單位自主分析能量：透過自辦工作坊教育訓練、多場技術講座, 培訓來自各單位之大數據分析種子學員 40 人, 並引進簡易視覺化分析軟體, 將訓練知識、分析技巧實際運用於本身業務, 每個人都可成為數據分析者, 目前已有 9 篇初步成果報告即為明證。
3. 強化資料分析效率及效能：藉由各項 SOP 設計, 從議題診斷、資料清洗到資料分析, 按部就班「以對的方式做事 (do the things right)」提高分析效率, 也留下完整作業軌跡及經驗; 另單位間 3 合 1 緊密合作及資料集中化的設計, 減少散彈打鳥、各自為政、重複浪費的情形, 確保「做對的事 (do the right things)」, 發揮資料連結效益, 提升整體分析效能, 強化政府證據建立能力。

第6章 結論與建議

第1節 結論

數據分析究竟是否屬於「大數據」分析，其實並不重要，重要的是它有沒有應用價值。話雖如此，不可否認的，具4V特性的大數據分析的確帶給我們前所未有的重要發現與應用，在國內外工、商業界，這類案例已多如牛毛，無庸贅述。去年11月美國眾議院通過「證據基礎決策基本法(Foundations for Evidence-Based Policymaking Act)」，規範聯邦機關應提出證據建立計畫，建構高效能的證據建立工作小組，以統籌機關之政策評估與分析作業，提醒我們政府統計重視並發展大數據分析之必要性。

本部於去年底成立大數據工作小組，以專案計畫方式，有系統的推動大數據分析相關工作，與美國之發展方向不謀而合，並已獲初步成果。未來將賡續利用建立之議題與資料盤點模式，找出2項政府及民眾迫切需要之政策案例來實作，如將戶籍與地籍資料作串聯，探討無殼青年家庭組織型態變遷情況，藉由瞭解近幾年家庭組織型態變化，發掘相應問題(例：比較近幾年無殼青年中家庭組織型態變化與一般家庭是否一致？其可能差異原因為何？)，更可運用於空間統計，應用於空間大數據之展現。未來更期待透過業務、資訊與統計深度合作

的運作模式，開啟政府數據應用分析的新風貌。

第2節 建議

本計畫為長期性推動，透過標準流程建立，並不斷充實各單位自主分析能量，統合內政大數據，以強化內政部證據建立能力，計畫推行未滿一年，仍需持續積極推動，本年下半年將持續根據本專案計畫進程，適時安排相關講座，賡續辦理內政大數據相關分析及教育訓練，積極協助同仁解決大數據分析遭遇之困難，並啟發實作新思維。

(一) 大數據分析：

- 1.下半年重點為全力完成既訂分析工作(包括2篇委外試辦案例、2篇自主分析)，強化與政策之連結應用。
- 2.另藉試辦案完成之「戶政、地政及移民連結資料庫」，內容完整詳實，亟具應用價值，擬去識別化後，併同開發完成之視覺化軟體 (Tableau) 模組，提供部內單位試用 (使用辦法另訂)，預計將可有效提升本部自主分析能量。

(二) 教育訓練：依據試辦委外服務案所建立之「議題與資料盤點」、

「資料庫清洗」及「資料分析」3項標準作業模式 (SOP) 範本，規劃 108 年基礎分析班及主管應用班課程。

1. 基礎分析班：擬於 108 年第 1 季間辦理，對象為本部暨所屬機關單位承辦人。課程內容包括 (1) 跨領域、跨單位之議題資

料串聯應用作業模式；(2)資料庫清洗之優先次序、輕重緩急，及評估所需技術、時程之作業模式；(3)統計應用分析能力及軟體操作教育訓練。藉由課程講授與實務操作，提升承辦人對議題設定、資料清洗及數據分析之思維與能力。

2. 主管應用班：擬於 108 年第 1 季間辦理，對象為本部暨所屬機關單位之科長以上主管。課程內容包括 (1) 內政大數據之視覺化軟體 (Tableau) 模組操作應用 (2) 資料與議題互動設計之面向、步驟與思維邏輯。藉由簡易拖拉點選之統計分析工具，運用於已建置完成之「戶政、地政及移民連結資料庫」，供本部主管作為政策制定、法案推動及議題辯護之參考。

- (三) 建置內政大數據資料庫：盤點本部及所屬各大資料庫，建置內政大數據資料庫進行分析，藉由跨領域、跨單位資料之連結，進行資料萃取 (extract)、轉置 (transform)、載入 (load) 至資料倉儲系統，針對本部重大政策或民眾關注議題，產製關鍵數據指標，進而在資訊向上集中後，軟硬體設備具備的條件下，定期更新資料，建立預測模型及決策管理系統，提供多面向統計資訊，並以大數據文字探勘方式蒐集具體民意，增廣支援決策應用。並規劃「智慧型數據分析平臺」決策支援應用系統，提供內政決策分析。預計 108 年完成 6 項內政議題分析，並清

洗與匯入議題所需之戶政、地政、營建等資料集數據。

參考文獻

- 1.社論，證據基礎決策之挑戰－從美國談起，主計月刊第 748 期，2018.04[民 107.04]，頁 2-3
- 2.饒志堅，交通統計大數據首次應用分析省思，主計月刊第 719 期，2015.11[民 104.11]，頁 78-82