

106-080-5455

MOTC-IOT-104-IEB048

# 交通大數據分析與應用機制 先期規劃



交通部運輸研究所

中華民國 106 年 8 月



106-080-5455

MOTC-IOT-104-IEB048

# 交通大數據分析與應用機制 先期規劃

著者：張恩輔、王晉元、邱裕鈞、蘇昭銘、陳奕廷、  
鍾志成、黃笙玗、孫千山、吳明軒、施佑林、  
陳怡君、許家筠、游坤明、鄒慶士、郭奕宏、  
陳其華、吳東凌、陳翔捷

交通部運輸研究所

中華民國 106 年 8 月

國家圖書館出版品預行編目資料

交通大數據分析與應用機制先期規劃 / 張恩  
輔等著. -- 初版. -- 臺北市 : 交通部運  
研所, 民 106. 08  
面 ; 公分  
ISBN 978-986-05-3117-6(平裝)

1. 交通管理 2. 運輸系統

557

106012742

交通大數據分析與應用機制先期規劃

著 者：張恩輔、王晉元、邱裕鈞、蘇昭銘、陳奕廷、鍾志成、黃笙玟、孫千  
山、吳明軒、施佑林、陳怡君、許家筠、游坤明、鄒慶士、郭奕宏、  
陳其華、吳東凌、陳翔捷

出版機關：交通部運輸研究所

地 址：10548 臺北市敦化北路 240 號

網 址：[www.iot.gov.tw](http://www.iot.gov.tw) (中文版>圖書服務>本所出版品)

電 話：(02)23496789

出版年月：中華民國 106 年 8 月

印 刷 者：承亞興圖文印刷有限公司

版(刷)次冊數：初版一刷 65 冊

本書同時登載於交通部運輸研究所網站

定 價： 310 元

展 售 處：

交通部運輸研究所運輸資訊組 • 電話：(02)23496887

國家書店松江門市：10485 臺北市中山區松江路 209 號 • 電話：(02)25180207

五南文化廣場：40042 臺中市中山路 6 號 • 電話：(04)22260330

GPN：1010601039

ISBN：978-986-05-3117-6 (平裝)

著作財產權人：中華民國（代表機關：交通部運輸研究所）

本著作保留所有權利，欲利用本著作全部或部分內容者，須徵求交通部運輸研究所  
書面授權。



## 交通部運輸研究所合作研究計畫出版品摘要表

出版品名稱：交通大數據分析與應用機制先期規劃			
國際標準書號（或叢刊號） 978-986-05-3117-6(平裝)	政府出版品統一編號 1010601039	運輸研究所出版品編號 106-080-5455	計畫編號 104-IEB048
本所主辦單位：運輸資訊組 主管：陳其華 計畫主持人：陳其華 研究人員：吳東凌、陳翔捷 聯絡電話：(02)23496886 傳真號碼：(02)25450426	合作研究單位：財團法人中興工程顧問社 計畫主持人：張恩輔、王晉元、邱裕鈞 蘇昭銘、陳奕廷 研究人員：鍾志成、黃笙珵、孫千山、 吳明軒、施佑林、陳怡君、 許家筠、游坤明、鄒慶士、 郭奕宏 地址：臺北市南京東路5段171號 聯絡電話：(02)87919198 ext. 427 傳真號碼：(02)87912198		研究期間 自 105 年 1 月 至 105 年 11 月
關鍵詞：先進大眾運輸系統、大數據、資料探勘			
<p>摘要：</p> <p>本計畫針對公共運輸、公路、軌道等等領域進行研析，整理國外28個交通大數據應用案例以及訪談國內9個交通單位對大數據之需求及期待解決議題，提出交通主管機關應用大數據之建議。本計畫歸納整理出制定數據戰略、共享旅運資訊、強化資產管理、落實風險評估與翻轉營運服務等五大策略，期望達到「智慧數據引領智慧交通」，利用數據形成政策與決策。在大數據分析之實作上，本研究以臺中市區公車為分析對象，以電子票証與公車動態資訊系統資訊為公共運輸巨量資料分析之核心，分別自節點流量分析與節線流量分析為出發點，進行公車運輸大數據分析，研究成果可協助主管機關監測路線規劃與管理策略的盲點，並據以研擬具體策略並執行行動計畫。最後本研究另回顧歐美等先進國家公共運輸資料匯流平臺之標竿案例，透過盤點國內公共運輸巨量資料之種類與型態，據以提出匯流平臺之架構與機制。</p>			
出版日期	頁數	定價	本 出 版 品 取 得 方 式
106 年 8 月	262	310	凡屬機密性出版品均不對外公開。普通性出版品，公營、公益機關團體及學校可函洽本所免費贈閱；私人及私營機關團體可按定價價購。
<p>機密等級：</p> <p><input type="checkbox"/>密 <input type="checkbox"/>機密 <input type="checkbox"/>極機密 <input type="checkbox"/>絕對機密</p> <p>(解密條件：<input type="checkbox"/> 年 月 日解密，<input type="checkbox"/>公布後解密，<input type="checkbox"/>附件抽存後解密， <input type="checkbox"/>工作完成或會議終了時解密，<input type="checkbox"/>另行檢討後辦理解密)</p> <p><input checked="" type="checkbox"/>普通</p>			
備註：本研究之結論與建議不代表交通部之意見。			

**PUBLICATION ABSTRACTS OF RESEARCH PROJECTS  
INSTITUTE OF TRANSPORTATION  
MINISTRY OF TRANSPORTATION AND COMMUNICATIONS**

TITLE: Preliminary Planning on Big Data Analysis and Application Mechanism of Transportation Data			
ISBN(OR ISSN) ISBN978-986-05-3117-6(pbk.)	GOVERNMENT PUBLICATIONS NUMBER 1010601039	IOT SERIAL NUMBER 106-080-5455	PROJECT NUMBER 104-IEB048
DIVISION: Information System Division DIVISION CHIEF: Ci-Hwa Chen PRINCIPAL INVESTIGATOR: Ci-Hwa Chen PROJECT STAFF: Dong-Ling Wu, Siang-Jie Chen PHONE: 886-2-23496886 FAX: 886-2-25450426			PROJECT PERIOD  FROM January 2016 TO November 2016
RESEARCH AGENCY: Sinotech Engineering Consultants, Inc. PRINCIPAL INVESTIGATOR: En-Fu Chang, Jin-Yuan Wang, Yu-Jyun Ciou, Jhao-Ming Su, Yi-Ting Chen PROJECT STAFF: Jyh-Cherng Jong, Sheng-Hsuan Huang, Chian-Shan Suen, Ming-Hsuan Wu , You-Lin Shih, Yi-Jun Chen, Jia-Yun Xu, Kun-Ming You, Qing-Shi Zou, Yi-Hong Guo ADDRESS: 171 Nanking E. RD. SEC. 5, Taipei, Taiwan, R.O.C. PHONE: 886-2-87919198 ext. 427 FAX: 886-2-87912198			
KEY WORDS: Advance Public Transport System, Big Data, Data Mining			
<b>ABSTRACT:</b>  The purpose of this project is to propose a short-term plan of implementing Big Data application for MOTC. By analyzing 28 foreign traffic Big Data application cases in public transport, highway and railway fields and interviewing nine bureaus for their needs, this project introduces five policies, Strategy, Message, Asset, Risk and Transformation (SMART) to achieve the goal of making decision based on Big Data. This study use Taichung city as a role model to conduct big data analysis on bus transportation from node and line analysis prospective. The result of this study can help transit agency to monitor blind spot of their route planning , which is the foundation of proposing management strategy for improving bus transportation. Based on the role model of foreign public transportation data platform, this research also look back and study past research on construction and application of big data analysis platform, as well as identify domestic public transit data to propose related platform structure and framework.			
DATE OF PUBLICATION August 2017	NUMBER OF PAGES 262	PRICE 310	CLASSIFICATION <input type="checkbox"/> RESTRICTED <input type="checkbox"/> CONFIDENTIAL <input type="checkbox"/> SECRET <input type="checkbox"/> TOP SECRET <input checked="" type="checkbox"/> UNCLASSIFIED
The views expressed in this publication are not necessarily those of the Ministry of Transportation and Communications.			

# 目 錄

	頁 次
目 錄.....	III
圖 目 錄.....	VI
表 目 錄.....	XIV
<b>第一章 緒論.....</b>	<b>1 - 1</b>
1.1 研究背景.....	1 - 1
1.2 研究目的.....	1 - 3
1.3 工作項目.....	1 - 3
1.4 研究流程.....	1 - 4
<b>第二章 文獻回顧.....</b>	<b>2 - 1</b>
2.1 大數據定義與特性.....	2 - 1
2.2 建構導入流程.....	2 - 3
2.3 重要技術回顧.....	2 - 5
2.4 應用發展趨勢.....	2 - 11
<b>第三章 交通大數據優先領域發展規劃.....</b>	<b>3 - 1</b>
3.1 國外交通大數據案例回顧.....	3 - 1
3.1.1 公路運輸領域.....	3 - 1
3.1.2 鐵路運輸領域.....	3 - 10
3.1.3 航空運輸領域.....	3 - 16
3.1.4 航港運輸領域.....	3 - 24
3.1.5 小結.....	3 - 27
3.2 交通機關需求分析.....	3 - 28
3.2.1 公路運輸領域.....	3 - 29
3.2.2 鐵路運輸領域.....	3 - 32
3.2.3 航空運輸領域.....	3 - 35
3.2.4 航港運輸領域.....	3 - 38

3.2.5	綜合整理 .....	3 - 40
3.3	資料匯流協調機制建議 .....	3 - 40
3.3.1	現況問題描述 .....	3 - 41
3.3.2	協調機制建議 .....	3 - 42
3.4	優先推動領域規劃原則與策略 .....	3 - 46
3.4.1	規劃原則 .....	3 - 46
3.4.2	優先推動之策略 .....	3 - 48
3.5	五大策略之建議 .....	3 - 49
3.5.1	制定數據戰略 .....	3 - 49
3.5.2	共享旅運資訊 .....	3 - 51
3.5.3	強化資產管理 .....	3 - 54
3.5.4	落實風險評估 .....	3 - 55
3.5.5	翻轉營運服務 .....	3 - 57
<b>第四章</b>	<b>標竿案例分析 .....</b>	<b>4 - 1</b>
4.1	分析對象 .....	4 - 1
4.2	舒適面向分析 .....	4 - 3
4.2.1	節點流量分析 .....	4 - 4
4.2.2	節線流量分析 .....	4 - 9
4.3	密集面向分析 .....	4 - 14
4.3.1	區間車可行性分析 .....	4 - 14
4.3.2	班表優化調整 .....	4 - 20
4.4	信賴面向分析 .....	4 - 27
4.4.1	車機傳輸漏失與延誤 .....	4 - 27
4.4.2	數據異常樣態 .....	4 - 34
4.4.3	站牌班表推動可行性分析 .....	4 - 39
4.5	安全面向分析 .....	4 - 52
4.6	無縫面向分析 .....	4 - 62
4.6.1	電子票證跨運具分析效益 .....	4 - 62
4.6.2	跨運具分析可行性探討 .....	4 - 64
4.6.3	跨運具轉乘分析初探 .....	4 - 67

<b>第五章 資料匯流平臺規劃</b> .....	<b>5 - 1</b>
5.1 國外資料匯流平臺回顧.....	5 - 1
5.1.1 商業公司匯流平臺案例.....	5 - 1
5.1.2 城市層級匯流平臺案例.....	5 - 6
5.1.3 綜合比較.....	5 - 15
5.2 國內交通數據平臺回顧.....	5 - 16
5.2.1 盤點案例.....	5 - 16
5.2.2 綜合比較.....	5 - 27
5.3 基本需求規劃設計.....	5 - 28
5.3.1 匯流類別.....	5 - 28
5.3.2 資料倉儲規劃.....	5 - 30
5.3.3 服務網站規劃.....	5 - 36
5.4 軟硬體架構規劃.....	5 - 43
5.5 資料流通管理辦法草案規劃.....	5 - 49
5.6 其他配套規劃.....	5 - 56
5.6.1 資訊安全及效能規劃.....	5 - 56
5.6.2 系統備援與還原規劃.....	5 - 60
<b>第六章 結論與建議</b> .....	<b>6 - 1</b>
6.1 結論.....	6 - 1
6.2 建議.....	6 - 2
<b>附錄 A 教育訓練辦理情形</b> .....	<b>C - 1</b>
<b>附錄 B 參考文獻</b> .....	<b>D - 1</b>
<b>附錄 C 期中審查意見回覆表</b> .....	<b>E - 1</b>
<b>附錄 D 期末簡報資料</b> .....	<b>F - 1</b>
<b>附錄 E 期末審查意見回覆表</b> .....	<b>G - 1</b>

# 圖 目 錄

	頁 次
圖 1-1 「大數據交通管理與服務應用」研討會眾長官留影.....	1 - 1
圖 1-2 交通大數據三大優先應用領域.....	1 - 2
圖 1-3 研究流程圖.....	1 - 4
圖 2-1 大數據資料成長速度.....	2 - 1
圖 2-2 大數據 4Vs 特性示意圖.....	2 - 3
圖 2-3 大數據建構流程示意圖.....	2 - 4
圖 2-4 HDFS 架構示意圖.....	2 - 6
圖 2-5 Spark 架構示意圖.....	2 - 8
圖 2-6 資料視覺化 (Data Visualization) 示意圖.....	2 - 11
圖 2-7 物聯網(IoT)裝置成長預估.....	2 - 12
圖 2-8 視覺化的政治色彩地圖.....	2 - 14
圖 2-9 智慧城市應用巨量資料分析技術示意圖.....	2 - 15
圖 2-10 Google Flu Trends.....	2 - 16
圖 2-11 SmartBay Calway 監控資料收集示意圖.....	2 - 17
圖 2-12 美國 NASA 地球觀測站繪製全球空氣污染地圖.....	2 - 18
圖 2-13 Big Data Landscape 2016 (Version 3.0).....	2 - 19
圖 3-1 MyTransport.SG 手機應用程式.....	3 - 1
圖 3-2 MyTransport.SG 公車到站時間.....	3 - 2
圖 3-3 BusViz 公車站牌視圖.....	3 - 3
圖 3-4 BusViz 時空圖.....	3 - 3
圖 3-5 Transit App 介面.....	3 - 4
圖 3-6 RHTS 方法的 OD 調查結果.....	3 - 4
圖 3-7 App 蒐集的 OD 調查結果.....	3 - 5
圖 3-8 公車總旅行時間.....	3 - 6
圖 3-9 旅客總旅行時間比較.....	3 - 6

圖 3-10 MetroViz 介面 .....	3 - 7
圖 3-11 MetroViz 資料模型 .....	3 - 8
圖 3-12 伊斯坦堡 BRT 路線智慧卡使用比例 .....	3 - 9
圖 3-13 伊斯坦堡利用大數據分析票價型式圖型化示意圖 .....	3 - 9
圖 3-14 韓國首爾貓頭鷹公車服務 .....	3 - 10
圖 3-15 行車輛未停駛事故統計 .....	3 - 11
圖 3-16 文字探勘以軌跡類型分類 .....	3 - 11
圖 3-17 文字頻率分組事故 .....	3 - 12
圖 3-18 優化維修時間 .....	3 - 12
圖 3-19 SBB TIBCO 系統示意圖 .....	3 - 13
圖 3-20 通勤預報示意圖 .....	3 - 14
圖 3-21 瑞典鐵路大數據管理 .....	3 - 15
圖 3-22 時間序列叢集分析圖 .....	3 - 16
圖 3-23 感測器數據量 .....	3 - 17
圖 3-24 澳洲航空時刻恢復表 .....	3 - 17
圖 3-25 Taleris 示意圖 .....	3 - 19
圖 3-26 masFlight 資訊平臺 .....	3 - 20
圖 3-27 masFlight 運用大數據 .....	3 - 20
圖 3-28 以大數據分析進行溫室氣體排放量之演算過程 .....	3 - 21
圖 3-29 以大數據分析計算溫室氣體排放 .....	3 - 21
圖 3-30 樞紐機場座位供給趨勢 .....	3 - 22
圖 3-31 美國主要客運機場運量情勢 .....	3 - 22
圖 3-32 全美航空與美西航空相關樞紐機場座位供給趨勢變化 .....	3 - 23
圖 3-33 達美航空與西北航空相關樞紐機場座位供給趨勢 .....	3 - 23
圖 3-34 船舶安全距離示意圖 .....	3 - 24
圖 3-35 大船期搜尋頁面 .....	3 - 25
圖 3-36 TRAXENS 監控技術 .....	3 - 26
圖 3-37 SIMS 概觀 .....	3 - 27
圖 3-38 訪談進程一覽(以首次訪談為序) .....	3 - 28
圖 3-39 中華電信區域人潮客群分析系統 .....	3 - 30

圖 3-40 中華電信人潮統計表 .....	3 - 30
圖 3-41 內政資料開放平臺的交通事故資料 .....	3 - 31
圖 3-42 公共運輸整合資訊流通服務平臺上臺鐵與高鐵時刻表 ....	3 - 34
圖 3-43 氣象資料開放平臺 .....	3 - 37
圖 3-44 航港單一窗口服務平臺 .....	3 - 39
圖 3-45 諮詢評議委員會之組織架構圖 .....	3 - 43
圖 3-46 智慧資料引領智慧交通之發展策略 .....	3 - 48
圖 4-1 分析區域評估流程 .....	4 - 1
圖 4-2 乘載率分析架構示意圖 .....	4 - 3
圖 4-3 站牌上下車人數分布圖 .....	4 - 4
圖 4-4 土耳其 BRT 研究之 O/D 起迄呈現方式 .....	4 - 5
圖 4-5 結合 GIS 與 O/D 呈現 (以 208 路線為例) .....	4 - 5
圖 4-6 旅客起迄 Heat Map 圖 (以 100 路公車為例) .....	4 - 6
圖 4-7 具業者/日期/時段/起迄多重選擇之互動儀表板 .....	4 - 6
圖 4-8 於圖 4-7 右側之篩選過濾器選項 .....	4 - 7
圖 4-9 點選單一上車站牌之儀表板互動示意圖 .....	4 - 7
圖 4-10 選取特定点方圓 5 公里站牌操作示意圖 .....	4 - 8
圖 4-11 對應圖 4-10 之儀表板互動結果示意圖 .....	4 - 8
圖 4-12 車上乘載人數計算程序 I—產生上/下車分布圖 .....	4 - 9
圖 4-13 車上乘載人數計算程序 II—產生累積上/下車數分布 圖 .....	4 - 10
圖 4-14 車上乘載人數計算程序 III—累積上/下車數相減完成 .....	4 - 10
圖 4-15 單日車上旅客數分析探勘儀表板截圖 .....	4 - 11
圖 4-16 由圖 4-15 改良之載具時空軌跡套疊乘載人數儀表板 .....	4 - 11
圖 4-17 豐原-東勢路廊主要路線起迄圖 .....	4 - 12
圖 4-18 豐原-東勢路廊主要路線起迄圖 .....	4 - 12
圖 4-19 結合 GIS 與軌跡之乘載人數呈現 .....	4 - 13
圖 4-20 單日車上旅客總數隨時間變化趨勢圖 .....	4 - 13
圖 4-21 站牌間乘位損失示意圖 .....	4 - 14
圖 4-22 不同營運起迄區間數量之效果差異圖 .....	4 - 15

圖 4-23	以不同車站為界之運能損失差異示意圖 .....	4 - 15
圖 4-24	重疊區間中途折返之規劃方式比較 .....	4 - 16
圖 4-25	各路線全年乘載人數熱區套疊統計儀表板 .....	4 - 17
圖 4-26	各路線平假日乘載人數熱區比較儀表板 .....	4 - 17
圖 4-27	區間車衝擊評互動估儀表板 .....	4 - 18
圖 4-28	區間車衝擊評互動估儀表板的互動方式 .....	4 - 19
圖 4-29	以區間車衝擊評互動估儀表板協助設計區間車 .....	4 - 19
圖 4-30	資料探勘五步驟示意圖 .....	4 - 20
圖 4-31	各路線旅客數量時間序列預測儀表板 .....	4 - 21
圖 4-32	依時段/票種之旅客數量時間序列預測儀表板 .....	4 - 22
圖 4-33	點選上午七點時段一般票種之時間序列預測結果 .....	4 - 22
圖 4-34	點選下午五點時段學生票種之時間序列預測結果 .....	4 - 22
圖 4-35	分群數與平方誤差和關係圖 .....	4 - 23
圖 4-36	55 路全年各時段車上乘載人數集群結果圖 (2 群) .....	4 - 23
圖 4-37	55 路全年各時段車上乘載人數集群結果圖 (3 群) .....	4 - 24
圖 4-38	55 路全年各時段車上乘載人數集群結果圖 (4 群) .....	4 - 24
圖 4-39	班表調整績效模擬模工具外觀 .....	4 - 25
圖 4-40	均勻班距下各班次各區間車上乘載人數調整結果 .....	4 - 26
圖 4-41	模擬最佳發車時間各班次各區間車上乘載人數結果 .....	4 - 26
圖 4-42	資料傳輸漏失分析示意圖 .....	4 - 28
圖 4-43	車機可靠度儀表板右上方分析選項 .....	4 - 29
圖 4-44	不同起始日之各業者漏失累計比例比較圖 .....	4 - 30
圖 4-45	傳輸漏失之年分/月分因素比較圖 .....	4 - 30
圖 4-46	傳輸漏失之星期/時段比較圖 .....	4 - 31
圖 4-47	傳輸漏失之車牌比較圖 .....	4 - 31
圖 4-48	傳輸漏失之路線比較圖 .....	4 - 31
圖 4-49	傳輸漏失之電信商比較圖(無差異案例) .....	4 - 32
圖 4-50	傳輸漏失之車牌比較圖(有差異案例) .....	4 - 32
圖 4-51	不同客運業者傳輸延遲比較圖 .....	4 - 33
圖 4-52	不同電信商平均傳輸延遲比較圖 .....	4 - 33

圖 4-53 一般資料清理程序.....	4 - 34
圖 4-54 利用異常樣態分析提昇資料品質流程示意圖.....	4 - 34
圖 4-55 未成對票證紀錄範例圖.....	4 - 35
圖 4-56 未成對異常刷卡樣態分析-多維度分析儀表板設計圖.....	4 - 35
圖 4-57 未成對異常刷卡樣態分析-業者名稱與票種分佈圖.....	4 - 36
圖 4-58 未成對異常刷卡樣態分析-卡號與路分佈圖.....	4 - 36
圖 4-59 未成對異常刷卡樣態分析-車號與路線分佈圖.....	4 - 37
圖 4-60 未成對異常刷卡樣態分析-路線與司機代碼分佈圖.....	4 - 37
圖 4-61 未成對異常刷卡樣態分析-站牌分佈圖.....	4 - 38
圖 4-62 電子票證與公車動態資料勾機失敗差異比較圖.....	4 - 39
圖 4-63 站牌班表選定路線基本資料比較.....	4 - 40
圖 4-64 【153】路車輛行駛時空圖.....	4 - 40
圖 4-65 【51】路車輛行駛時空圖.....	4 - 41
圖 4-66 【51 路（往屯區藝文中心）】行駛時間下緣線.....	4 - 42
圖 4-67 【153 路（往谷關）】行駛時間下緣線.....	4 - 43
圖 4-68 【276 路（往臺中公園）】行駛時間下緣線.....	4 - 43
圖 4-69 【289 路（往東平社區）】行駛時間下緣線.....	4 - 44
圖 4-70 【51】莒光新城→屯區藝文中心.....	4 - 48
圖 4-71 【51】屯區藝文中心→莒光新城.....	4 - 48
圖 4-72 【153】市警局→谷關.....	4 - 49
圖 4-73 【153】谷關→市警局.....	4 - 49
圖 4-74 【276】東勢→臺中公園(雙十路).....	4 - 50
圖 4-75 【276】臺中公園(雙十路)→東勢.....	4 - 50
圖 4-76 【289】東平社區→臺中火車站.....	4 - 51
圖 4-77 【289】臺中火車站→東平社區.....	4 - 51
圖 4-78 肇事資料集中事故地點資料樣本截圖.....	4 - 59
圖 4-79 部分易混淆或不易解碼之資料截圖.....	4 - 59
圖 4-80 十公尺區塊範圍示意圖.....	4 - 59
圖 4-81 事故地址範例圖.....	4 - 59
圖 4-82 肇事地址經 Google Maps 轉換後之分佈圖.....	4 - 60

圖 4-83 發生事故的路段 Top50 橫條圖.....	4 - 61
圖 4-84 高風險路線樹圖.....	4 - 61
圖 4-85 高風險車輛樹圖.....	4 - 62
圖 4-87 臺中地區傳統鐵路車站 2014 各月分進出人數圖.....	4 - 63
圖 4-88 臺中地區 60 餘個 iBike 地理位置圖.....	4 - 63
圖 4-89 臺灣鐵路管理局多卡通交易紀錄主要欄位.....	4 - 64
圖 4-90 臺灣鐵路管理局多卡通卡號資料截圖.....	4 - 64
圖 4-91 桃園地區公車票證資料截圖.....	4 - 65
圖 4-92 同編碼之公共汽車票證資料截圖.....	4 - 65
圖 4-93 同編碼之臺鐵多卡通票證資料截圖.....	4 - 65
圖 4-94 臺鐵多卡通票證進站時間插補演算法示意圖.....	4 - 66
圖 4-95 同編碼之公共自行車票證資料截圖.....	4 - 66
圖 4-96 一至三月整體搭乘行為.....	4 - 68
圖 4-97 一至三月各票種搭乘行為.....	4 - 68
圖 4-98 一至三月各車站轉乘比例.....	4 - 69
圖 4-101 一至三月每日轉乘旅次量與平均轉乘時間變化.....	4 - 69
圖 4-99 三月一周各日平均轉乘旅次量與平均轉乘時間變化.....	4 - 70
圖 4-100 三月各臺鐵車站之公車轉乘臺鐵的轉乘時間盒鬚圖.....	4 - 71
圖 4-101 三月各臺鐵車站轉乘公車之轉乘時間盒鬚圖.....	4 - 71
圖 5-1 AirSage 資料分析及應用平臺示意圖.....	5 - 2
圖 5-2 INRIX 連結駕駛人網路示意圖.....	5 - 3
圖 5-3 INRIX 提供駕駛人服務示意圖.....	5 - 3
圖 5-4 ParkMe App 功能畫面.....	5 - 5
圖 5-5 Taxi Stockholm 計程車大數據應用平臺.....	5 - 6
圖 5-6 倫敦市區地鐵即時動態地圖.....	5 - 7
圖 5-7 IBM 協助打造交通及時預測系統示意圖.....	5 - 9
圖 5-8 往北方向平均旅行時間及速度 (2015 年 7 月 24 日).....	5 - 10
圖 5-9 SHRP2 在交通運輸可靠性上研究計畫流程示意圖.....	5 - 11
圖 5-10 交通運輸改善分析示意圖.....	5 - 12
圖 5-11 SHRP2 需求與供給研究示意圖.....	5 - 12

圖 5-12 貴州交易所交易數據類型 .....	5 - 13
圖 5-13 貴州交易所數據交易平臺 .....	5 - 14
圖 5-14 珠海智慧交通大數據架構圖 .....	5 - 15
圖 5-15 電子收費系統示意圖 .....	5 - 17
圖 5-16 臺灣高速公路電子收費系統 .....	5 - 18
圖 5-17 以運具分類查詢公共運輸資料 .....	5 - 19
圖 5-18 公共運輸資料服務內容檢視 .....	5 - 19
圖 5-19 以機關分類查詢公共運輸資料 .....	5 - 21
圖 5-20 以名稱查詢公共運輸資料 .....	5 - 21
圖 5-21 Easy Traffic 服務平臺入口網頁 .....	5 - 22
圖 5-22 Easy Traffic 服務平臺資料服務專區 .....	5 - 23
圖 5-23 Easy Traffic 服務平臺示範應用專區 .....	5 - 23
圖 5-24 示範應用範例說明 .....	5 - 24
圖 5-25 Easy Traffic 服務平臺示範模組應用專區 .....	5 - 24
圖 5-26 臺中市易肇事路口整合平臺建置計畫內容示意圖 .....	5 - 25
圖 5-27 肇事班點圖 .....	5 - 26
圖 5-28 肇事熱點圖 .....	5 - 27
圖 5-29 公共運輸服務水準評估方向分類圖 .....	5 - 28
圖 5-30 公車動態定點資料相關資料表及欄位圖 .....	5 - 31
圖 5-31 臺鐵載具軌跡相關資料欄位圖 .....	5 - 32
圖 5-32 公車電子票證資料欄位圖 .....	5 - 33
圖 5-33 臺鐵第三代票務系統對號車票資料欄位圖 .....	5 - 33
圖 5-34 臺鐵第三代票務系統自動售票機資料欄位圖 .....	5 - 34
圖 5-35 臺鐵第三代票務系統之定期票主要資料欄位圖 .....	5 - 34
圖 5-36 臺鐵第三代票務系統之多卡通主要資料欄位圖 .....	5 - 35
圖 5-37 監理系統運輸業車籍駕籍與違規資料欄位圖 .....	5 - 36
圖 5-38 匯流平臺功能架構示意圖 .....	5 - 36
圖 5-39 大數據資料彙集平臺入口網頁示意圖 .....	5 - 37
圖 5-40 資料集功能示意圖 .....	5 - 39
圖 5-41 問答專區示意圖 .....	5 - 39

圖 5-42 新增提問示意圖 .....	5 - 40
圖 5-43 管理系統入口網頁示意圖 .....	5 - 41
圖 5-44 角色權限管理示意圖 .....	5 - 42
圖 5-45 統計報表分析示意圖 .....	5 - 43
圖 5-46 軟體系統架構圖(基礎版) .....	5 - 44
圖 5-47 最小叢集部署與橫向擴充架構圖 .....	5 - 44
圖 5-48 大數據資料匯集運作架構 .....	5 - 45
圖 5-49 系統多層級架構圖 .....	5 - 47
圖 5-50 系統環境建置 .....	5 - 48
圖 5-51 整體軟體與資源應用規劃建議圖 .....	5 - 48
圖 5-52 資料集實體內容申請流程圖 .....	5 - 55
圖 5-53 SSO 作業流程圖 .....	5 - 58
圖 5-54 看門狗程式架構圖 .....	5 - 58
圖 5-55 平均回應時間與錯誤率關係圖 .....	5 - 59
圖 5-56 使用者數量與等待時間波譜圖 .....	5 - 59
圖 5-57 備援計畫分佈圖 .....	5 - 60
圖 5-58 系統備分示意圖 .....	5 - 61
圖 5-59 系統還原示意圖 .....	5 - 62

# 表 目 錄

	頁 次
表 1.1 臺灣交通大數據價值之六大核心共識一覽表 .....	1 - 2
表 3.1 國外運輸領域大數據應用目的彙整表 .....	3 - 27
表 3.2 國內交通機關大數據應用需求彙整表 .....	3 - 40
表 4.1 節點流量與節線流量分析比較表 .....	4 - 3
表 4.2 本研究蒐集近二年臺中地區公車動態定時回傳筆數表 ....	4 - 27
表 4.3 近二年臺中地區公車動態節定時資料漏失統計表 .....	4 - 29
表 4.4 公車動態歷史資料格式範例【289 路】 .....	4 - 41
表 4.5 採用下緣線基準法算出之【289 路】調整後之下緣線 ....	4 - 45
表 4.6 駕駛訓練管理建議 .....	4 - 47
表 4.7 不同資料來源電子票證跨運具分析優缺點比較表 .....	4 - 67
表 5.1 商業公司資料匯流應用目的比較表 .....	5 - 15
表 5.2 城市整合資料匯流應用目的比較表 .....	5 - 16
表 5.3 運輸分類資料彙整表 .....	5 - 20
表 5.4 公共運輸資料來源機關彙整表 .....	5 - 20
表 5.5 匯流平臺資料納選評估考慮因素表 .....	5 - 29
表 5.6 匯流平臺初創期間擬收納資料集表 .....	5 - 29
表 5.7 公車動態定時資料欄位清單 .....	5 - 31
表 5.8 公車動態定點資料欄位清單 .....	5 - 32
表 5.9 資料流通管理資料分級表 .....	5 - 52

# 第一章 緒論

## 1.1 研究背景

近年來國內外大數據分析應用逐漸普及，身為國內運輸前瞻領航的交通部運輸研究所（以下簡稱運研所）在 104 年初基於行政院「善用網路科技技術，透過資料開放及大數據概念創造前瞻施政」之政策方針下，推動了「公車動態資訊系統巨量資料（Big Data）蒐集與視覺化分析研究」計畫，該計畫由交通大學運輸研究中心、中興工程顧問社、中華大學等機構共同執行，該研究自公共運輸巨量資料中挖掘有價值之資訊，同時研擬各種視覺化呈現研究成果，成果豐碩。

同年 7 月 24 日交通部運研所更舉辦「大數據（Big Data）分析在交通管理與服務應用」研討會，交通部陳部長與六都交通局長（圖 1-1）在會中達成「創造臺灣交通大數據價值之六大核心共識」如表 1.1，也初擬未來交通大數據之優先應用領域如圖 1-2。



圖 1-1 「大數據交通管理與服務應用」研討會眾長官留影

為使交通主管機關未來將透過數據思維進行施政，優先達成政府和政府間資料開放，反映地方政府對於中央部會資料之需求，利用交通數據與其他領域資料之相互交流，啟發更多創新應用的可能，實有必要進一步規

劃未來交通大數據之分析與應用機制，並研析數據匯流平臺之內涵，讓大數據分析成果能支援交通政策。

表 1.1 臺灣交通大數據價值之六大核心共識一覽表

運輸好行，創造數據思維前瞻施政	安全提升，健全交通事故防治分析
數據共享，累積政府決策基礎資源	跨域整合，推動交通數據交流平臺
人才培育，建立資料科學合作機制	中央領航，群力六都數據顧問服務

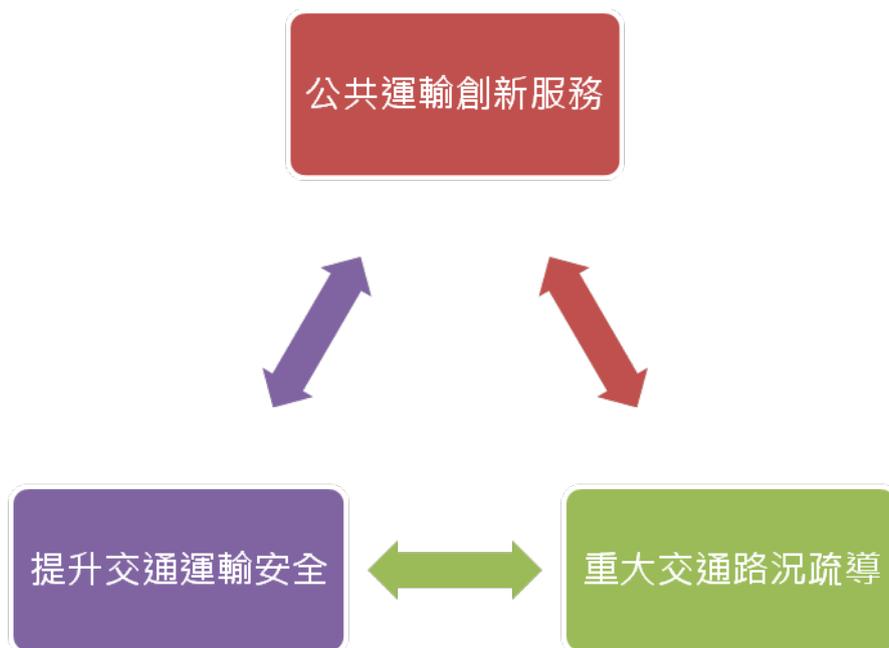


圖 1-2 交通大數據三大優先應用領域

## 1.2 研究目的

由前述背景說明可彙整本研究之目的如下：

1. 希冀透過本計畫建立之交通大數據分析與應用機制，為我國現階段交通領域大數據分析奠定良好的基礎，更對於我國未來交通領域大數據分析產生關鍵性的影響。
2. 透過公共運輸與其他建議優先領域，蒐集跨單位應用大數據之相關意見，提出交通主管機關未來應用大數據之建議。
3. 延續「公車動態資訊系統巨量資料（Big Data）蒐集與視覺化分析研究」計畫成果，擴大在公共運輸領域應用之分析範圍與深度。

### 1.3 工作項目

根據上述各項研究目的，本計畫之工作內容可彙整為下列幾項，其中與第 1 項研究目的相關有第 1、3、4、5、7 項工作項目；與第 2 項研究目的相關有第 1 項工作項目；與第 3 項研究目的相關有第 2 項工作項目；與第 4 項研究目的相關有 6 項工作項目。

1. 提出交通大數據之應用方向建議
2. 針對「公共運輸」領域，進行標竿案例分析研究
3. 提出交通部數據匯流平臺之建議
4. 進行深度訪談與辦理座談會、教育訓練
5. 發表 2 篇國內期刊或研討會論文
6. 指派 1 名專案經理、2 位駐點人員至運研所推動大數據應用工作
7. 驗竣後提供 1 年技術諮詢服務

### 1.4 研究流程

上述工作項目的研究重點為前三項，其研究流程如圖 1-3 所示，相關研究成果詳見第三～五章。

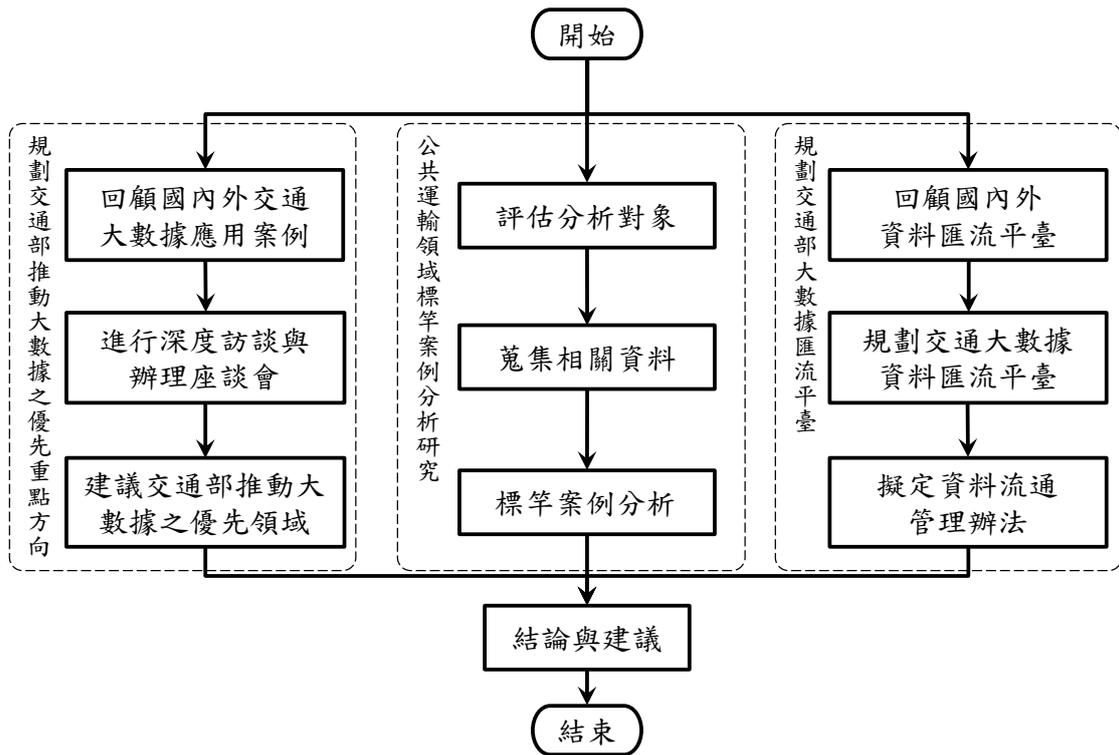


圖 1-3 研究流程圖

## 第二章 文獻回顧

### 2.1 大數據定義與特性

大數據 (Big Data) — 或稱巨量資料，顧名思義，是指大量的資訊，當資料量龐大到資料庫系統無法在合理時間內進行儲存、運算、處理，分析成能解讀的資訊時，就稱為大數據。目前的大數據資料增長數量與速度可用圖 2-1 來說明：

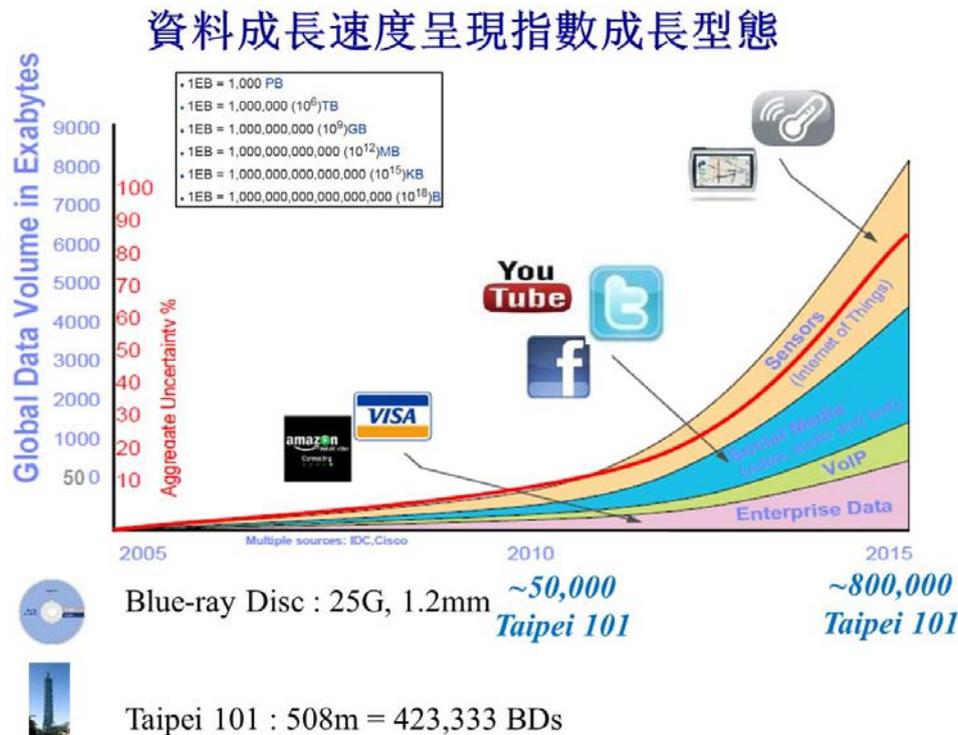


圖 2-1 大數據資料成長速度

在圖 2-1 中可看到至 2015 年為止，所有資料的數量（包括感測器、物聯網、社群媒體、企業資料等）加起來的總數極巨大，若將資料全部儲存在藍光 DVD 中，以一片 DVD 高度 1.2mm 來計算，每片 DVD 可存放 25GB 資料量來看，一棟台北 101 大樓約等於 423,333 片光碟高度，而至 2015 年止，資料量約等於 800,000 棟台北 101 大樓的高度。

這些巨量資料中有著珍貴的訊息，像是未知的相關性 (Unknown Correlation)、未顯露的模式 (Hidden Patterns)、市場趨勢 (Market Trend)，

可能埋藏著前所未見的知識跟應用等著被我們挖掘發現；但由於資料量太龐大，流動速度太快，現今科技無法處理分析，促使我們不斷研發出新一代的資料儲存設備及科技，希望從大數據中萃取出那些有價值的資訊。

「Big Data」這個詞最早由 IBM 提出，2010 年才真正開始受到注目，並成為專業用語登上維基百科，算是「大數據」的正式問世。而在 2012 年時，《紐約時報》的專欄文章「The Age of Big Data 2」更是宣告了「大數據時代」的來臨。

大數據幾乎無法使用大多數的資料庫管理系統處理，而必須使用「在數十、數百甚至數千台伺服器上同時平行運行的軟體」。大數據的定義取決於持有資料組的機構之能力，以及其平常用來處理分析資料的軟體之能力。

大數據由巨型資料集組成，這些資料集大小常超出人類在可接受時間下的蒐集、應用、管理和處理能力。大數據的大小經常改變，單一資料集的大小從兆位元組 (Tera Bytes, TB) 至數十兆億位元組 (Peta Bytes, PB) 不等。

在一分 2001 年的研究與相關的演講中，麥塔集團 (META Group, 現為 Gartner) 分析員道格·萊尼 (Doug Laney) 指出資料增長的挑戰和機遇有三個特性：量 (Volume, 資料大小)、速 (Velocity, 資料輸入輸出的速度) 與多變 (Variety, 多樣性)，合稱「3V」或「3Vs」。高德納與現在大部分大數據產業中的公司，都繼續使用 3V 來描述大數據。Gartner 於 2012 年修改對大數據的定義：「大數據是大量、高速、及/或多變的資訊資產，它需要新型的處理方式去促成更強的決策能力、洞察力與最佳化處理。」另外，有機構在 3V 之外定義第 4 個 V：真實性 (Veracity) 為第四個特性，以下簡介 4Vs 的定義：

#### 1. Volume：資料量

以前人們「手動」在表格中記錄、累積出數據；現在數據是由機器、網路、人與人之間的社群互動來生成。你現在正在點擊的滑鼠、來電、簡訊、網路搜尋、線上交易等等都是正在生成累積成龐大的數據，因此資料量很容易就能達到數 TB，甚至上看 PB 或 EB (Exabytes, 百萬兆位元組) 的等級。

#### 2. Velocity：資料輸入輸出速度

資料的傳輸流動 (Data Streaming) 是連續且快速的，隨著越來越多的機器、網路使用者，社群網站、搜尋結果每秒都在成長，每天都在輸出更多的內容。公司跟機構要處理龐大的資訊大潮向他們襲來，而回應、反應這些資料的速度也成為他們最大的挑戰，許多資料要能即時得到結果才能發揮最大的價值，因此也有人會將 Velocity 認為是「時效性」。

### 3. Variety：資料類型

大數據的來源種類包羅萬象，十分多樣化，如果一定要把資料分類的話，最簡單的方法是分兩類，結構化與非結構化。早期的非結構化資料主要是文字，隨著網路的發展，又擴展到電子郵件、網頁、社交媒體、視訊，音樂、圖片等等，這些非結構化的資料造成儲存 (Storage)、探勘 (Mining)、分析 (Analyzing) 上的困難。

### 4. Veracity：真實性

這個詞由在 Express Scripts 擔任首席數據官 (Chief Data Officer, CDO) 的 Inderpal Bhandar 在波士頓大數據創新高峰會 (Big Data Innovation Summit) 的演講中提出，認為大數據分析中應該加入這點做考慮，分析並過濾資料有偏差、偽造、異常的部分，防止這些「Dirty Data」損害到資料系統的完整跟正確性，進而影響決策。

以上大數據 4Vs 特性，彙整如圖 2-2 所示。

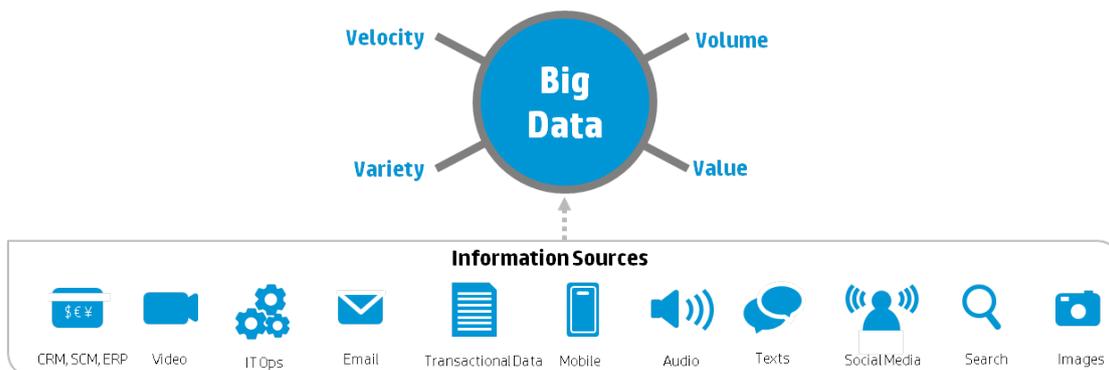


圖 2-2 大數據 4Vs 特性示意圖

## 2.2 建構導入流程

資料分析應用可分為幾個階段，從數據蒐集、傳輸、儲存、運算至應用，每階段皆有不同的導入邏輯，整體供應鏈也需有多方廠商支援與合作。數據來源包含行動裝置、各個感測設備、社群媒體、影音圖資等各種結構化與非結構化資訊，經由通訊設備傳輸到系統進行處理。在儲存端，以往

企業慣於把數據存在內部資料庫當中，需要時再進行批次運算，較為耗時且費用較高。

然而在巨量資料領域中，兼具橫向擴充和高度彈性的雲端分散式檔案系統(Distributed File System)，將大量數據分割成數個檔案以同時進行存取與備分，並在儲存結點上進行資料分析，為重要的技術趨勢。

大數據必須藉由計算機對資料進行統計、比對、解析方能得出客觀結果。美國在 2012 年就開始著手大數據，歐巴馬(Barack Obama)更在同年投入 2 億美金在大數據的開發中，更強調大數據會是之後的未來石油。

建構一個完整大數據用的時候，在資料蒐集、儲存、管理、處理、呈現、隱私控管、商業模式等環節，步驟流程圖說明，如圖 2-3 所示：

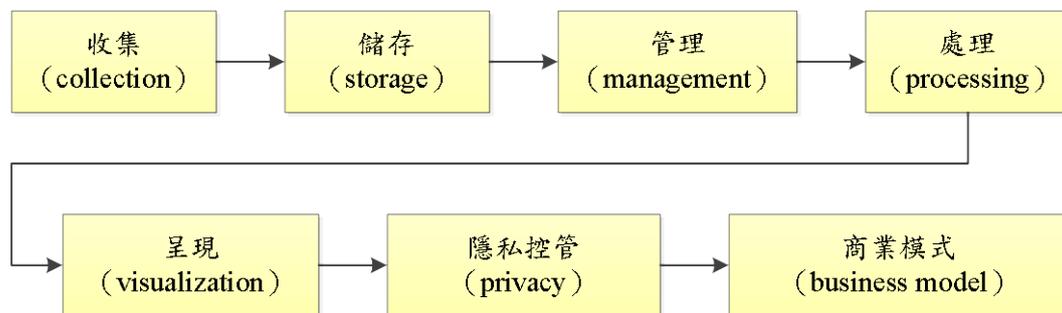


圖 2-3 大數據建構流程示意圖

### 1. 蒐集 (Collection)：

這是大數據的第一個門檻，沒有資料，一切免談。因此只有擁有資料的特定的公司和機構，或是設法建立蒐集資料的管道，才有機會發展大數據應用。

### 2. 儲存 (Storage)：

以現在的科技，只要有經費，儲存 PB 等級的資料，早已不是問題。Hadoop 是一個開放原始碼軟體計畫 (Open Source Software Project)，可以讓使用者將大型檔案切割成許多區塊分散放置在多台電腦上，也具備容錯、高效能等優點。

### 3. 管理 (Management)：

除了以檔案方式管理資料之外，將資料輸入資料庫系統，是資訊系統常見的作法，方便使用者將資料分門別類和搜索更新資料。然而在大數據的情況下，資料庫系統的效率成為應用開發的重點之一。因此，有別於傳統的關聯式資料庫系統 (Relational Database Management System)，專門

針對大數據應用的資料庫系統應運而生，有些藉著放寬一些傳統資料庫的規範，提昇資料處理的效率。

#### 4. 處理 (Processing) :

傳統的程式撰寫方法，是根據程式會被電腦一行一行循序執行 (Sequential Execution) 的思維。要快速處理大數據，程式必須能夠同時在多台電腦上執行。Google 所提出的 MapReduce 程式設計框架 (Framework)，方便使用者製作出能同時利用成千上萬台電腦上處理大數據的平行程式 (Parallel Program)。然而，並非每位應用開發者都願意將原有的程式改寫和平行化，尤其是套裝軟體的使用者，因此常用的 R、SPSS、Matlab 等軟體也設法支援大數據，以及以平行執行方式來加速處理。

#### 5. 呈現 (Visualization) :

資料處理的前後，往往需要專家或使用者的介入，去確認和觀察資料的意義。然而在大數據的情況下，需要透過一些專業、圖像化的方式有效地將資料快速地呈現在眼前。

#### 6. 隱私控管 (Privacy) :

實際上，許多資訊系統中的資料屬個人所有，負責保管資料的廠商即便在契約書上明定有分析資料的權利，但同時也有不可洩漏資料所有者個人隱私的義務。因此，大數據的控管，成為一個非常敏感、必須設立機制謹慎處理的議題。

#### 7. 商業模式 (Business Model) :

既然大數據應用的開發和系統的建構需要相當大的資源，如何讓花費許多資源所建構的大數據應用獲得足夠的利益，是發展新興大數據應用所必須的規劃。

## 2.3 重要技術回顧

精準挖掘出大數據真正價值是近年政府與企業關注的問題，分析大數據帶來有助政策與業務發展的寶貴資訊和分析，政府與企業可運用這些資訊預防產生高昂成本相關問題與即時作出反應，另外準確預測行為模式和民眾需求亦可提升企業營運蒐入，甚至找到提高利潤的商機。

大數據價值雖提供可觀回報，但要整合相關數據才會帶來政策與業務優勢，因此需要擴展資料庫系統和基礎架構，以便儲存 TB 甚至 PB 規模的即時數據，而傳統資料庫因軟硬體和成本高昂等問題，導致無法擴展或無法以經濟高效方式擴展。另外使用傳統資料庫，只可處理姓名、地址、電話號碼等結構化數據，但未能支援行動裝置、社交媒體和日誌文件等可變性非結構化數據，當政府與企業需要實際處理和應用這類大數據時，IT 部門就需要解決一連串問題，例如數據來源是否可靠、數據是否準確、是否具備足夠的儲存容量、硬體儲存會否讓數據管理變得困難、政府與企業組織架構是否能適應不斷發展的數據技術、是否有利用雲端運算、數據是否受到保護等需處理之議題。但是，隨著網路普及與科技進步之賜，許多處理資料分析與管理的技術因應而出，迎來了巨量資料的時代，以下簡介相關技術：

## 1. Apache Hadoop

Hadoop 是一個叢集系統 (Cluster System)，也就是由單一伺服器擴充到數以千計的機器，整合應用起來像是一台超級電腦。Hadoop 的兩大核心功能—儲存 (Store) 及處理 (Process) 資料，其使用 HDFS 分散式檔案系統與 MapReduce 平行運算架構，以下說明：

- (1) HDFS (Hadoop Distributed File System) 分散式檔案系統，其設計的概念是利用叢集系統中有數以千計的節點用來存放資料，透過 HDFS，Hadoop 能夠儲存上看 TB 甚至 PB 等級的巨量資料，不用擔心單一檔案的大小超過一個磁碟區的大小，而且也不用擔心某個機器損壞導致資料遺失，架構示意圖如圖 2-4 所示。

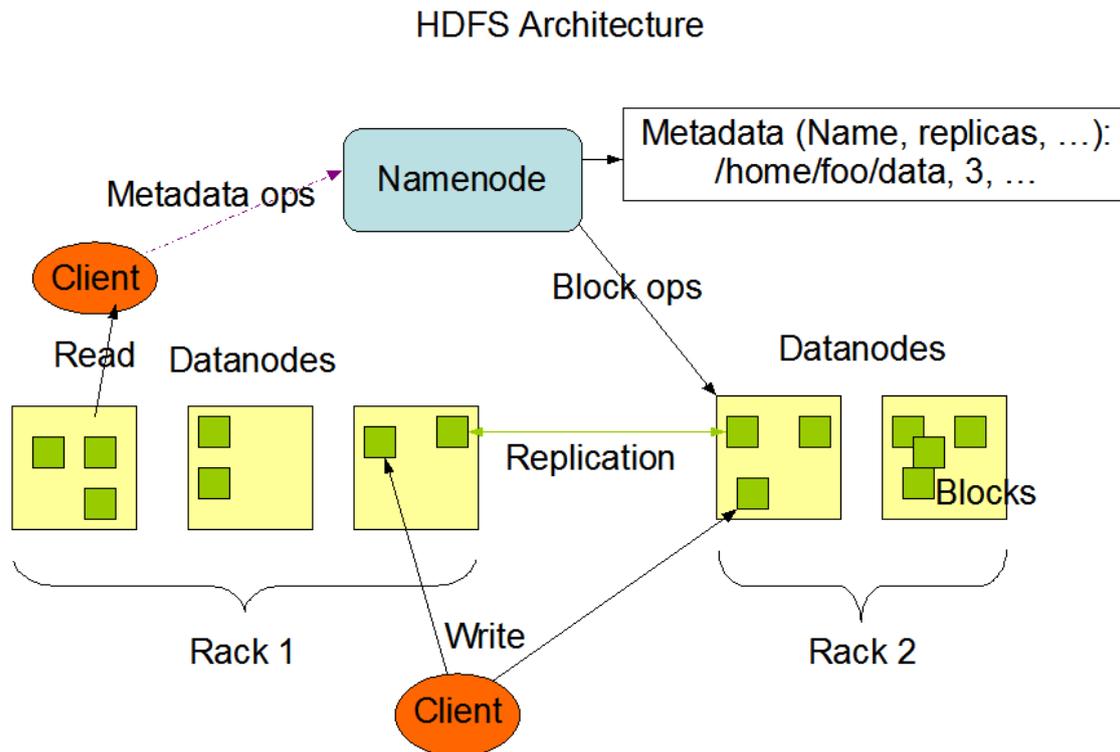


圖 2-4 HDFS 架構示意圖

- (2) MapReduce 顧名思義是以 Map 跟 Reduce 為基礎的應用程式。一般我們進行資料分析處理時，是將整個檔案丟進程式軟體中做運算出結果，而面對巨量資料時，Hadoop 的做法是採用分散式計算的技術處理各節點上的資料。在各個節點上處理資料片段，把工作分散、分佈出去的這個階段叫做 Mapping；接下來把各節點運算出的結果直接傳送回來歸納整合，這個階段就叫做 Reducing。這樣多管齊下、在上千台機器上平行處理巨量資料，可以大大節省資料處理的時間。

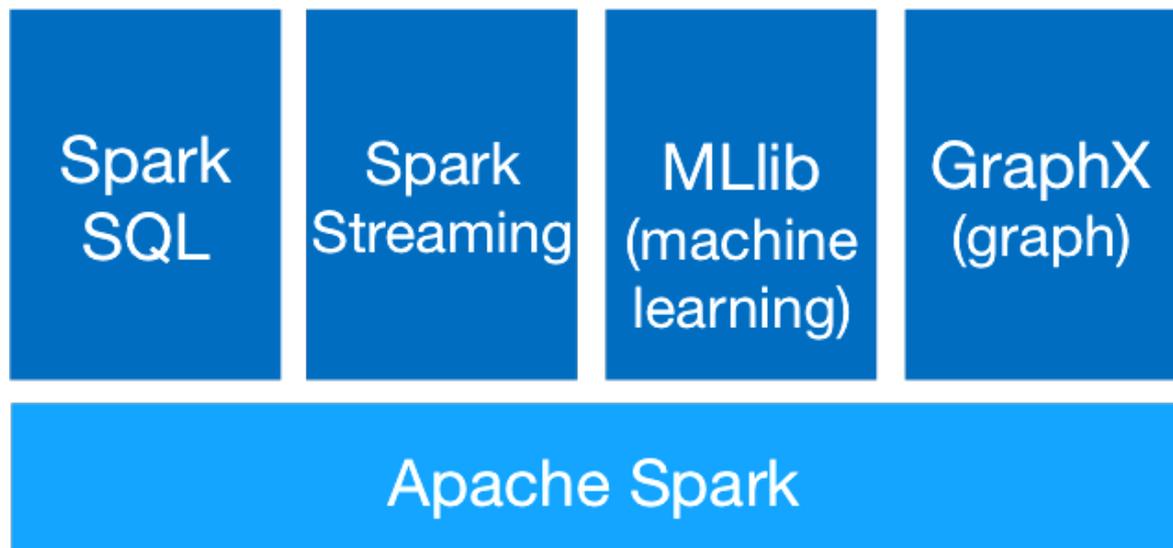
基於 Hadoop 處理巨量資料的種種優勢，像是解決了檔案存放、系統擴張、資料處理及備分等問題，因此 Hadoop 被廣泛應用於大數據儲存和大數據分析，成為大數據的主流技術。

在近幾年內，叢集運算 (Clusters) 在商用性與非商用性的領域也越來越普遍且應用的相當廣泛，2014 年無庸置疑成為了 Hadoop 的豐蒐年，對於 Hadoop 生態系例如 Flume、Oozie、Hive、Pig、HBase、YARN 等技能的市場需求也逐漸提高。

## 2. Apache Spark

由於 Hadoop 分析資料時需要將中間產生的數據存在硬碟中，因此會有讀寫資料的延遲問題；有別於 Hadoop 的 MapReduce 架構，Spark 使用了「記憶體內運算技術 (In-Memory Computing)」，能在資料尚未寫入硬碟時，就在記憶體內進行分析運算，號稱比原先的 Hadoop 快 100 倍。

使用 Spark 需要搭配叢集管理員和分散式儲存系統。Spark 支援獨立模式 (本地 Spark 叢集)、Hadoop YARN 或 Apache Mesos 的叢集管理。在分散式儲存方面，Spark 可以和 HDFS、Cassandra、OpenStack Swift 和 Amazon S3 等介面搭載。Spark 也支援偽分散式 (Pseudo-Distributed) 本地模式，不過通常只用於開發或測試時以本機檔案系統取代分散式儲存系統。在這樣的情況下，Spark 僅在一台機器上使用每個 CPU 核心執行程式，架構示意圖如圖 2-5 所示。



資料來源：[spark.apache.org](http://spark.apache.org)

圖 2-5 Spark 架構示意圖

Spark 構成要素：

- Spark SQL：For SL and unstructured data processing
- Spark Streaming：Stream processing of live data streams
- MLlib：Machine learning algorithms
- GraphX：Graph processing

Spark 特色：

- 能夠在記憶體內緩存資料集以進行互動式資料分析。
- Scala 或 Python 中的互動式命令列介面可降低橫向擴展資料探索的反應時間。

- Spark Streaming 對即時資料串流的處理具有可擴充性、高吞吐量、可容錯性等特點。
- Spark SQL 支援結構化和關聯式查詢處理 (SQL)。
- MLlib 機器學習演算法和 Graphx 圖形處理演算法的高階函式庫。

Spark 的核心支柱，是一個名為 RDD (Resilient Distributed Dataset) 的分散式計算模型，而其諸多優勢，主要也是源自於 RDD 本身的特性。

所謂的 RDD，乃是由 AMPLab 實驗室所提出的概念，類似一種分散式的記憶體。而且，RDD 是一種可跨群集 (Cluster) 被使用、可儲存於主記憶體中的 Immutable 的物件集合。這裡所謂的 Immutable 物件，乃是指在被產生之後，其狀態便無法被修改的物件。

在 RDD 之上，可以施加兩類型的操作，一種稱為「轉換 (Transformation)」，另一種則為「動作 (Action)」。其中，所謂的「轉換」，其操作結果為新的 RDD，意即其作用在於將 RDD 再轉換生成另一個 RDD。而所謂的「動作」，則是在 RDD 之上進行計算之後，將其結果返回 Spark 的驅動程序，或寫至檔案系統。

以 RDD 這種以儲存在記憶體為主的分散式物件來做計算，避免了 Hadoop MapReduce 在工作任務間必須透過檔案系統來溝通的效率問題。但是將計算資料儲存在檔案系統，就必須面臨磁碟 I/O 的緩慢效能，以及將資料做序列化、解序列化的計算負擔。Spark 因為以 RDD 為基礎，得以帶來大幅度的計算效能改善。

### 3. NoSQL

NoSQL 最早是指「No SQL」，號稱不使用 SQL 作為查詢語言的資料庫系統。但近來則普遍將 NoSQL 視為「Not Only SQL」，也就是「不只是 SQL」的意思，希望結合 SQL 優點並混用關聯式資料庫和 NoSQL 資料庫來達成最佳的儲存效果。

兩者存在許多顯著的不同點，其中最重要的是 NoSQL 不使用 SQL 作為查詢語言。其資料存儲可以不需要固定的表格模式，也經常會避免使用 SQL 的 JOIN 操作，一般有水平可延伸性的特徵。NoSQL 的實作具有二個特徵：使用硬碟，或者把隨機存取記憶體作儲存載體。

在大數據所帶動的潮流下，各種不同形態的 NoSQL 資料庫如雨後春筍般竄起，目前有 4 種比較受到關注的 NoSQL 資料庫，分別是 Key-Value

資料庫、記憶體資料庫(In-memory Database)、圖學資料庫(Graph Database)以及文件資料庫(Document Database)，簡介如下：

- (1) Key-Value 資料庫是 NoSQL 資料庫中最大宗的類型，這類資料最大的特色就是採用 Key-Value 資料架構，取消了原本關聯式資料庫中常用的欄位架構 (Schema)，每筆資料各自獨立，所以，可以打造出分散式和高擴充能力的特性。包括像 Google 的 BigTable、Hadoop 的 HBase、Amazon 的 Dynamo、Cassandra、Hypertable 等都是這類 Key-Value 資料庫。
- (2) 記憶體資料庫 (In-memory Database) 就是將資料儲存在記憶體的 NoSQL 資料庫，包括了 Memcached、Redis、Velocity、Tuple space 等。其實像 Memcached、Redis 都是一種 Key-Value 資料架構的資料庫，只是這類資料庫改將資料儲存在記憶體中來提高讀取效率，大多用來快取常用網頁，加快傳遞網頁的速度，減少讀取硬碟的次數，不過系統關機後就無法保存。
- (3) 圖學資料庫 (Graph Database) 這不是專門用來處理圖片的資料庫，而是指運用圖學架構來儲存節點間關係資料架構，例如用樹狀結構來組織從屬關係或網狀結構來儲存朋友關係，地理圖資系統通常也會用圖學資料庫來儲存地圖上每一點和鄰近點的關係，或用圖學資料庫來計算點與點之間最短的距離，也可以用同樣的概念來計算出人與人之間最短的交友距離。圖學資料庫最大的特性是對複雜性的擴充力，關係越複雜的資料越適合使用圖學資料庫。這類資料的資料結構沒有標準的作法，基本的圖學資料包括了節點 (Node)、關係 (Relation) 和屬性 (Property) 三種結構。例如用節點來記錄 Facebook 上的帳號、用關係來記錄朋友關係、用屬性來描述這個帳號的個人資料等。最後可以用網路圖來呈現出 Facebook 用戶之間的交友狀況。常見的圖學資料庫如 Neo4j、InfoGrid、AllegroGrph 等。
- (4) 文件資料庫 (Document Database) 主要是用來儲存非結構性的文件，例如最常見的非結構化資料就是 HTML 網頁。一個 HTML 網頁結構不像一般表格那樣有固定的欄位，每個欄位有特定資料類型和大小。例如網頁裡有 Head 和 Body 結構，Body 元素中可能會有 10 個段落，段落中會有文字、連結、圖片等。文件資料庫的資料結構往往是鬆散的樹狀結構。很多文件資料庫都是商用資料庫系統，文件資料庫的概念源自 IBM 的 Lotus Notes 儲存文件

的方式，XML 資料庫也是一種文件資料庫。常見的開源文件資料庫像是 CouchDB、MongoDB 以及 Riak 等。

#### 4. 結構化查詢語言 SQL

結構化查詢語言 (Structured Query Language)，簡稱 SQL，是專門用於關連式資料庫的一種查詢語言，可以用來定義資料庫結構、建立表格、指定欄位型態與長度，也能新增、異動或查詢資料。簡單來說，SQL 是一種用來從資料庫讀取與儲存資料的電腦語言。

SQL 歷經了四十多年的考驗仍然在蓬勃發展，雖然 NoSQL 的出現帶來了一些影響，但 SQL 仍然主導著市場，並在大數據領域贏得了很多投資與廣泛部屬。像是 Cloudera 推出了即時查詢開源工具 Impala—一款用來跑在 Hadoop 架構上的互動 SQL 查詢引擎，在這些工具發展下 SQL 在大數據領域中更是歷久不衰。

#### 5. 資料視覺化分析

大數據的重要性與日俱增，不少企業如電子商務、零售業及半導體製造業等，開始廣泛運用巨量資料為公司擬定企業策略，不過並不是人人都是數據專家、資料科學家，如果要讓主管跟客戶們清楚了解資料背後的意義，那倒不如讓他們「一目了然」。

資料視覺化 (Data Visualization) 是關於數據之視覺表現形式的研究，資料視覺化的技術可以幫助不同背景的工程人員溝通、理解，以達良好的設計與分析結果。市面上已經有許多工具、軟體為人們提供這方面的需求，像是 Tableau、QlikView 等工具就擁有絕佳的視覺化呈現效果，可以不限資料量、資料形式或主題，透過圖像化和便捷的操作介面製作出客製化報表，無需撰寫程式就能得到分析結果，應用範例如圖 2-6 所示。

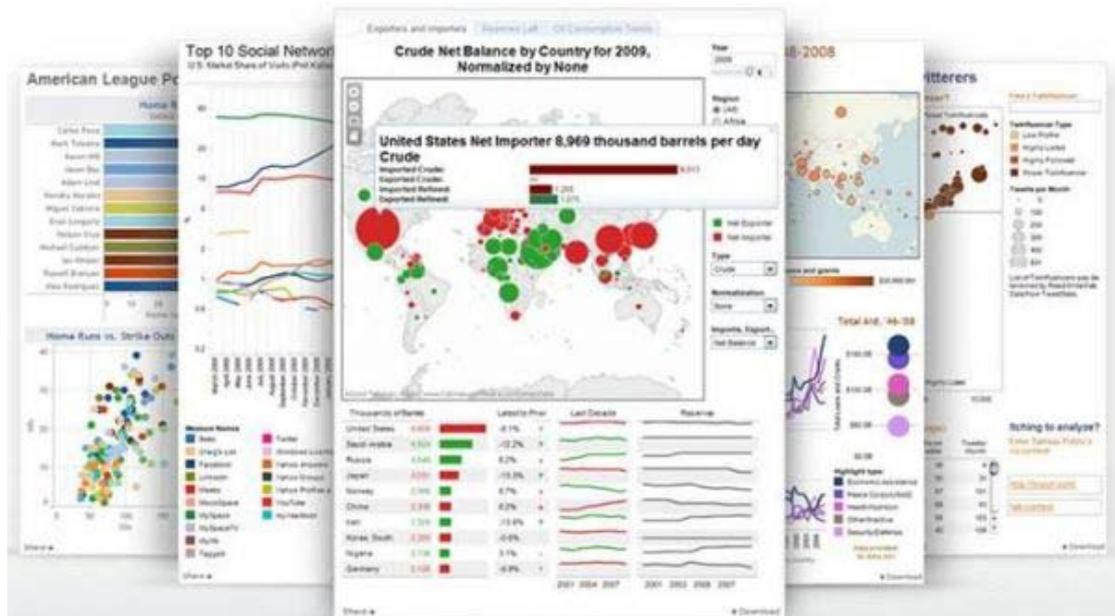


圖 2-6 資料視覺化 (Data Visualization) 示意圖

## 2.4 應用發展趨勢

大數據的應用範例包括大科學、RFID、感測裝置網路、天文學、大氣學、交通運輸、基因組學、生物學、大社會資料分析、網際網路檔案處理、製作網際網路搜尋引擎索引、通信記錄明細、軍事偵查、社群網路、通勤時間預測、醫療記錄、相片圖像和影像封存、大規模的電子商務等。然而，在現今網際網路跟雲端技術的發達，加上智慧型產品、行動裝置的普及下，整理大數據的發展重點與應用趨勢：

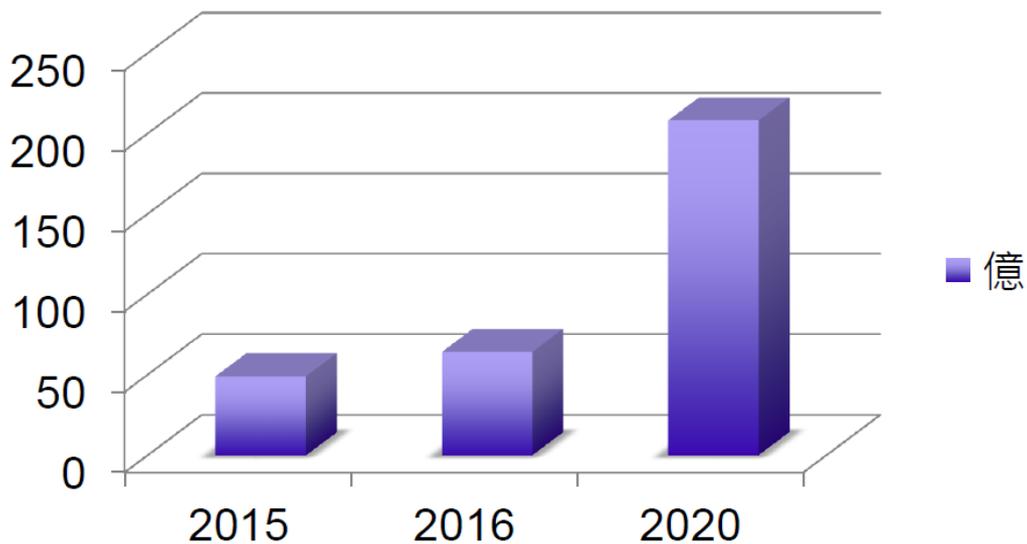
### 1. In-Memory 資料庫

當這世界上累積的資料量越來越龐大，各企業制訂商業策略所納入考量的資料也跟著大幅增加，這時候如何減少將資料 (Data) 轉換成資訊 (Information) 的時間變得尤其重要。

採用記憶體內運算技術(In-Memory Computing)減少原始資料的移動，僅搬移運算後的結果，加快處理的速度，並且透過壓縮技術減少資料量，能夠有效提升資料庫效能，應付企業對資料運算量及速度日益升高的要求，也使得企業得到的資訊更即時、能更快地回應各種市場需求及回饋，甚至開發出全新市場、開創出其獨一無二的價值。

### 2. 物聯網 (IoT) 資料之儲存處理與即時分析

在物聯網(IoT)概念起飛的時代，越來越多行動裝置、智慧型居家裝置被市場接受，進入到我們的生活中，根據 Gartner 的預測資料預估 2020 年時會有超過 208 億的物聯網(IoT)裝置，如圖 2-7 所示。



資料來源：Gartner

圖 2-7 物聯網(IoT)裝置成長預估

可以預期的是更多種類型的資料將以更多形式被感測、蒐集起來，而且這些大量且即時性的感測數據(Sensor Driven Data)屬於非結構化資料，也就是從未經處理的文字、聲音與影片導出的數據，如何儲存處理及分析成為現在十分重要的課題，可能從中挖掘出未知的趨勢並帶給人類生活重大的改變。

### 3. 大數據與雲端服務之整合

雲端(Cloud)從 2006 年被提出後，便廣為科技業所使用，各企業更是積極提出各種雲端服務。隨著國際廠商的積極投入與各國政府的政策推動，雲端運算技術愈臻成熟，企業對於雲端服務的採用亦越具規模；同時，伴隨著各式巨量資料的產生，透過雲端運算技術處理與分析，帶動業者打造各式加值應用及創新服務。根據國際研究暨顧問機構 Gartner 預測，2013 年全球雲端服務市場已達 341 億美元，預估 2014 年將可超過 400 億美元，2018 年更可突破 1,000 億美元。另外，根據市場研究公司 IDC 統計，全球雲端運算市場主要應用依序為通訊與媒體、政府與教育、製造、金融、流通、公共事業及醫療保健等七大領域。

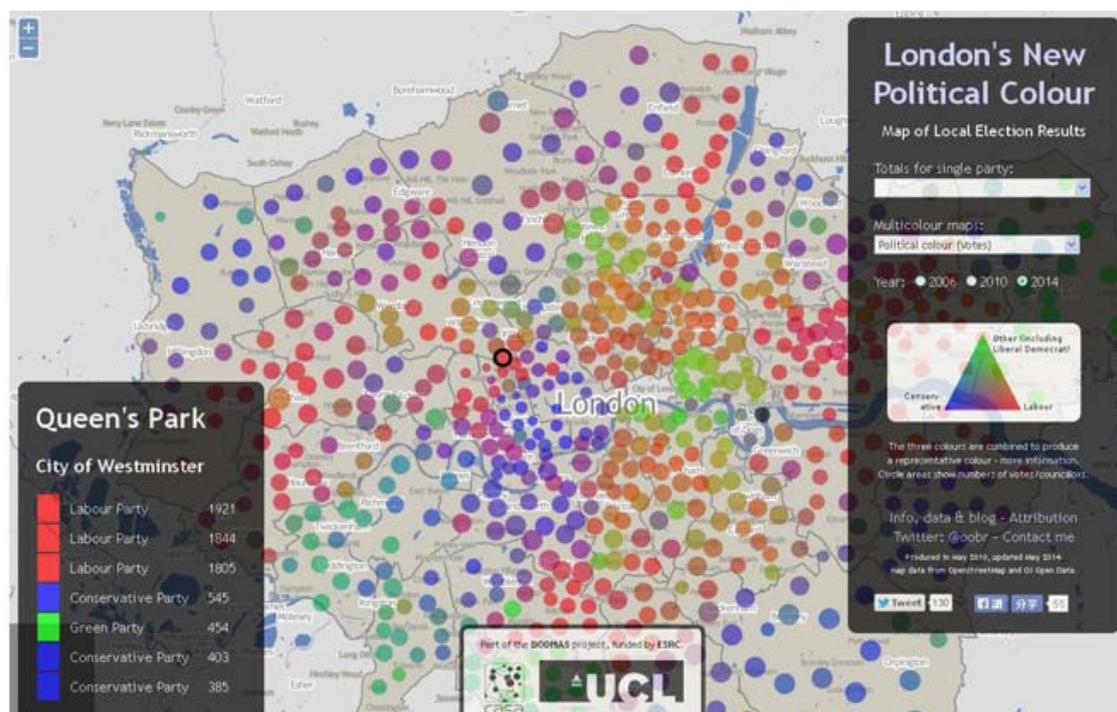
大數據與雲端技術可以說是相輔相成，大數據推動了雲端服務，而雲端服務的普遍也使得資料量攀升。2014 年全世界平均每天產生 23TB 的數據，大約是 2012 年的 920 倍，以這種情況來看，雲端服務在大數據時代相當於「公共設施」般不可或缺，不但用來儲存各式各樣的資料，還利用雲端運算來建構基於大數據的應用程式和 API，建立模型預測未來的事件。

#### 4. 巨量資料分析在社區政治之應用

一般對於政府開放的統計類資料，例如人口、勞工、教育、文化及選舉等資料較無法直接拿來應用，但倫敦政府則是找到了新應用方法，將地方選舉資料和社群開放街圖（Open Street Map）結合，創造出視覺化的政治色彩地圖。

這個選舉政治色彩地圖，是由倫敦政府將 32 個地區選舉資料，與開放街圖結合。這些選舉資料包括來自 2014 年倫敦地方選舉結果，以及對政黨委員的投票結果，如：英國勞工黨與保守黨或其他政黨，如綠黨、自由民主黨等，視覺化的政治色彩地圖如圖 2-8 所示。

此外，該地圖同時也能單獨顯示單一政黨、地區選舉的版圖分布。也讓民眾更了解每一地區政黨勢力分布，例如：2014 年勞工黨的勢力版圖偏向倫敦中心和東部，而保守黨則是偏向在外圍和西部。



資料來源：<http://casa.oobrien.com/london/>

圖 2-8 視覺化的政治色彩地圖

## 5. 善用巨量資料分析 打造智慧城市

從智慧建築到智慧城市的想像中，隨著聯網等技術的不斷發展，大樓建築可能同時具備相當多不同的終端需求，智慧建築所扮演的角色就是將這些需求加以整合，將大樓內個別的消防、門禁、冷凍空調、能源供應等子系統整合為智慧型連線網絡，各個系統在同一平台中共同運作，智慧網路管理系統便能以最佳化的方式將能源供應分配到各子系統，協同整合性運用，除了能達到智慧安全的管理，同時也發揮節能的最大效益。

以供電系統規劃來看，最基礎的部分就是電力狀態的監視，透過集合式多功能數位電錶並以通訊整合於中央監控系統，以進行電力的全面監控，用以分析及統計公共用電，來規劃管理營運應用；此外，亦可透過用電資訊與業主的客觀評估，提供「需量控制設備」來進行設備的運轉管控，甚至依據預先設定的條件或取得的參考信號，來解除或投入電量附載。

用電資訊的「可視化」，也被自動化產業視為智慧建築節能技術之一。近年來雲端及 IoT 議題發燒延燒，隨之而來的許多雲端應用，也如雨後春筍般的出現，自動化設備也逐步邁向雲端運算之流，而雲端整合資訊的功能，也成為自動化產業邁入雲端的切入點，節能可視化就是選項之一。

在 NEC 的智慧城市藍圖下，除了在防災與減災領域上有全方位的對策外，同時也兼顧了完備的環境保護機制。NEC 的智慧城市概念中，由最下層如地下水、天然氣、電力供給等核心基礎建設服務，在安全層面的緊急救護、防災系統等消防防災服務，以及公共安全、設施能源管理等整合管理基礎建設服務，最後是最上層的電子看板、e-learning 等通訊服務四階層，智慧城市應用巨量資料分析技術示意圖，如圖 2-9 所示。

新興的智慧大樓可自動控制大樓內數以千計的電力、空調、熱源等能源配置，透過使用量計測進行最佳的節能運轉，進一步透過區域能源的最佳化，以及各種能源的相互連結，達到智慧城市中的智慧能源使用的目標。



資料來源：DIGITIMES

圖 2-9 智慧城市應用巨量資料分析技術示意圖

## 6. 巨量資料分析在防疫管理之應用

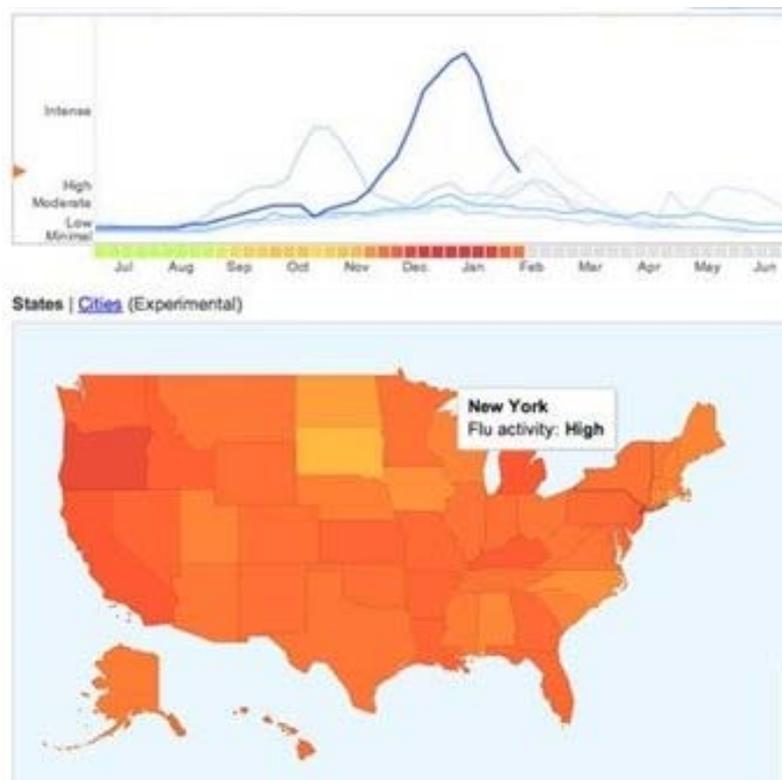
由於資訊化應用的普及，導致資料量的規模已經龐大到無法透過人工，在合理時間內達到擷取、管理、處理、並整理成為人類所能解讀的資訊，根據 IDC 統計，數位世界的資訊容量將會從 2009 年的 0.8ZB，在 2020 年成長到 35ZB，等於每 15 秒就成長 1PB，年複合成長率高達 40%，而且這些數據資料不僅巨大而且不同，如何優化資料，方便且容易的搜尋到所需要的資訊，也變得更加困難。

如新型禽流感疫情在 2009 年開始快速蔓延時，美國疾病管制局 (Centers for Disease Control ; CDC) 要求醫生必須通報新型流感的病例，但傳統的報送作業方式往往需要約兩個星期的時間，這樣的速度也會讓公共衛生機構錯失防疫的先機。

Google 工程師曾在科學期刊 Natural 發表有關流感的研究指出，他們將 Google 每天 300 萬個與禽流感相關的搜尋關鍵字，與 CDC 從 2007 到 2008 年的實際禽流感感染案例，透過數學演算法 (Algorism) 進行相關性 (correlation) 分析比對，最後找出 45 個與實際發生禽流感案例有很強相關

性的搜尋項目組合，只要在特定區域發現這些搜尋項目組合，發生禽流感的案例也會隨之增多。

Google 利用運算及線上查詢的優勢，如圖 2-10 所示，提供更快速而精確的預測，協助公共衛生機構掌握疫情資訊，可見單單只有儲存資料量龐大是沒有任何用處的，因為資料並不會自己進行分析，想要從大量的資訊中得到價值，必須要找到新技術。如位置訊息也是巨量資料分析一個非常重要的分析標的，巨量資料若能結合地理資訊系統(Geographic Information System；GIS)，除了「人」的位置資訊以外，「物體」的位置也可以利用這類的裝置進行大量的蒐集，這些訊息也成為重要的分析資料，讓用戶的地理位置變成非常寶貴的資料。



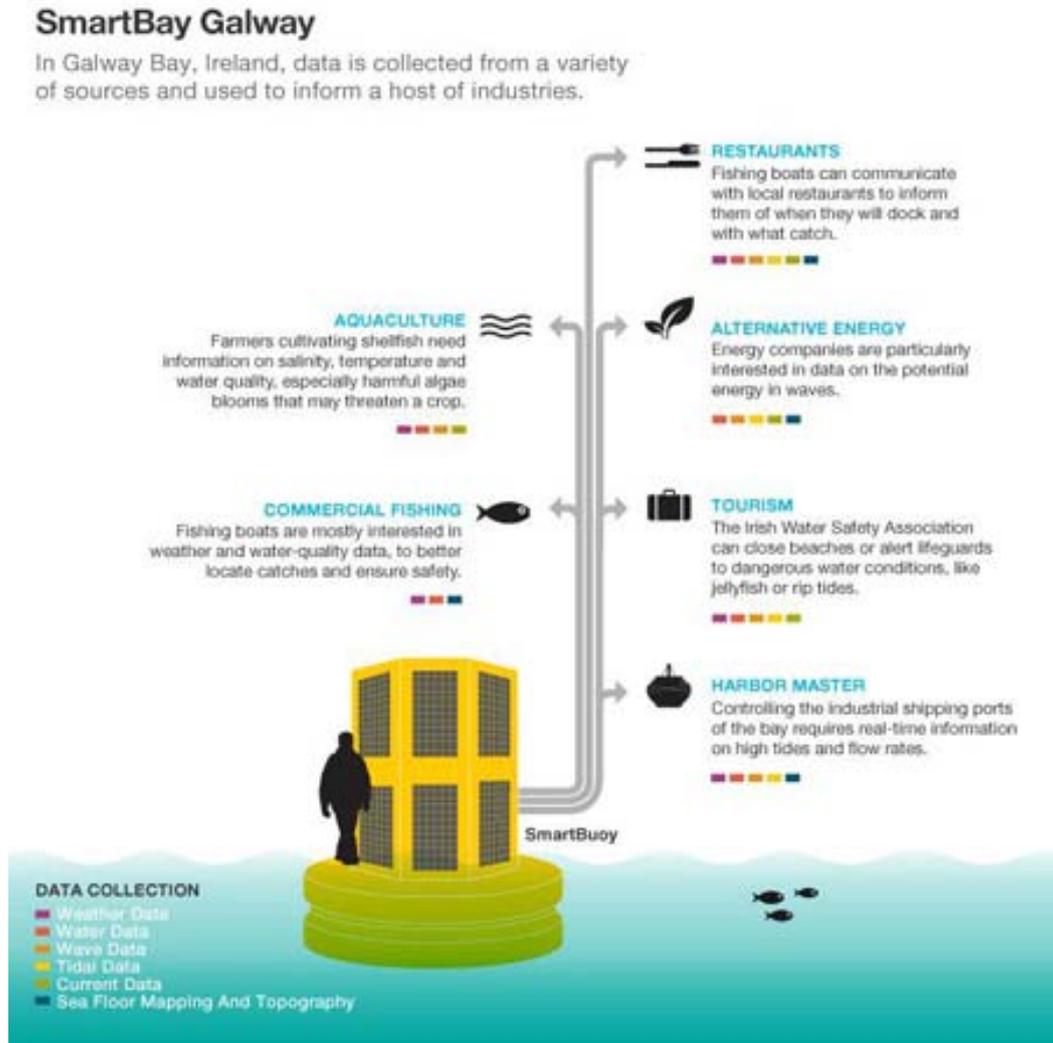
資料來源：Google.org

圖 2-10 Google Flu Trends

#### 7. 巨量資料分析在環境監控之應用

在資料蒐集技術與新興巨量資料應用方面，IBM 自 2009 年起，在資料較缺乏的領域如健康照護與環境應用，展開資料蒐集與建立的投資，其中包含與 Marine Institute Ireland 合作進行對海洋中海浪、污染和海產生命等資料即時監控與蒐集的「SmartBay Calway」先導計畫。

該計畫將蒐集到的巨量監控資料加以分析提供服務，並以 Web 或其他智慧通訊裝置傳遞給使用者，將資料多元提供不同領域的應用，此計畫資料蒐集與提供，如圖 2-11 所示。



資料來源：John Kennedy--Silicon Republic

圖 2-11 SmartBay Calway 監控資料蒐集示意圖

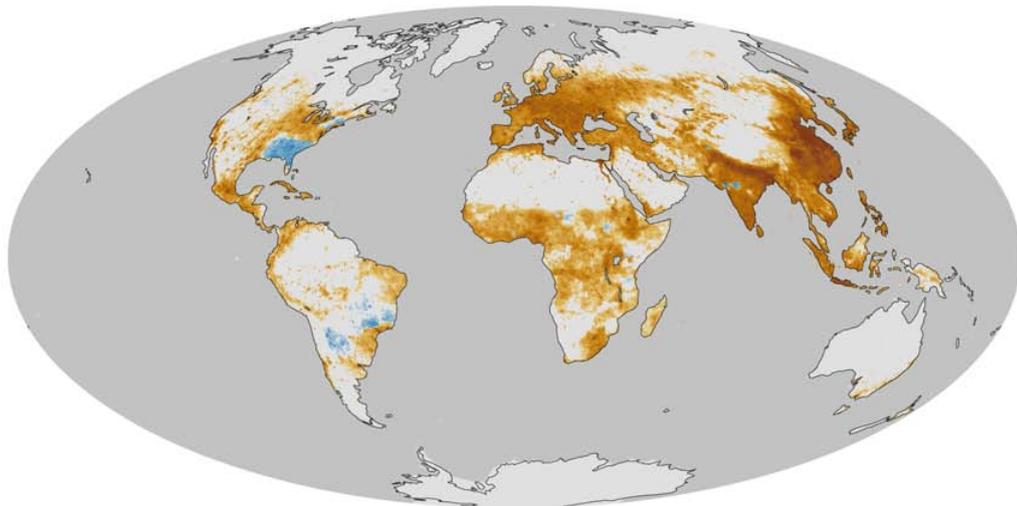
自動化資料蒐集及設備管理應用的技術實施，主要由資料蒐集與服務傳遞的閘道器，以及內建於感測器的資料處理軟體元件組成。在養殖應用案例上，布建於養殖水域的環境感測器，將感測到的水質溫度值、酸鹼值及溶氧值等數據資料，透過有線或無線通訊技術，將資訊傳送至後端的感測資料蒐集閘道器。感測資料蒐集閘道器解析各項水質環境資訊後，再依據場域環境，透過適切的通訊方式傳送環境資料至養殖應用平台儲存。

美國太空總署地球觀測站 (NASA Earth Observatory) 利用過去 150 年間全球年平均死亡率與空氣污染數據分析模型，進而製作出全球空氣污染地圖，顯示不同地區間的空氣污染對人體健康造成的直接影響。

近幾年各項研究皆顯示，不同國家地區的空氣污染程度與該地區人口的平均死亡率存有關聯性，而去年，北卡羅來納大學地球科學家 Jason West 與其團隊採用 6 台模擬大氣電腦模型分析計算後，發現每年全世界有 210 萬人口的直接死因是來自一種被稱為細懸浮微粒 (PM2.5) 的空氣污染源。

這種細懸浮微粒會導致健康問題，因為它們足夠小可通過呼吸系統進入人體肺部。而一般像是汽車所排放的廢氣，以及其它像是工業、家庭，如油煙釋出的氣體都會產生此有毒的細懸浮微粒。

而為了強調其風險性，NASA 地球觀測站則利用北卡羅來納大學的空氣污染數據分析模型，並結合 1850 年至 2000 年之間每千平方公里的年平均死亡人數，繪製出全球空氣污染地圖，如圖 2-12 所示。



資料來源：NASA Earth Observatory

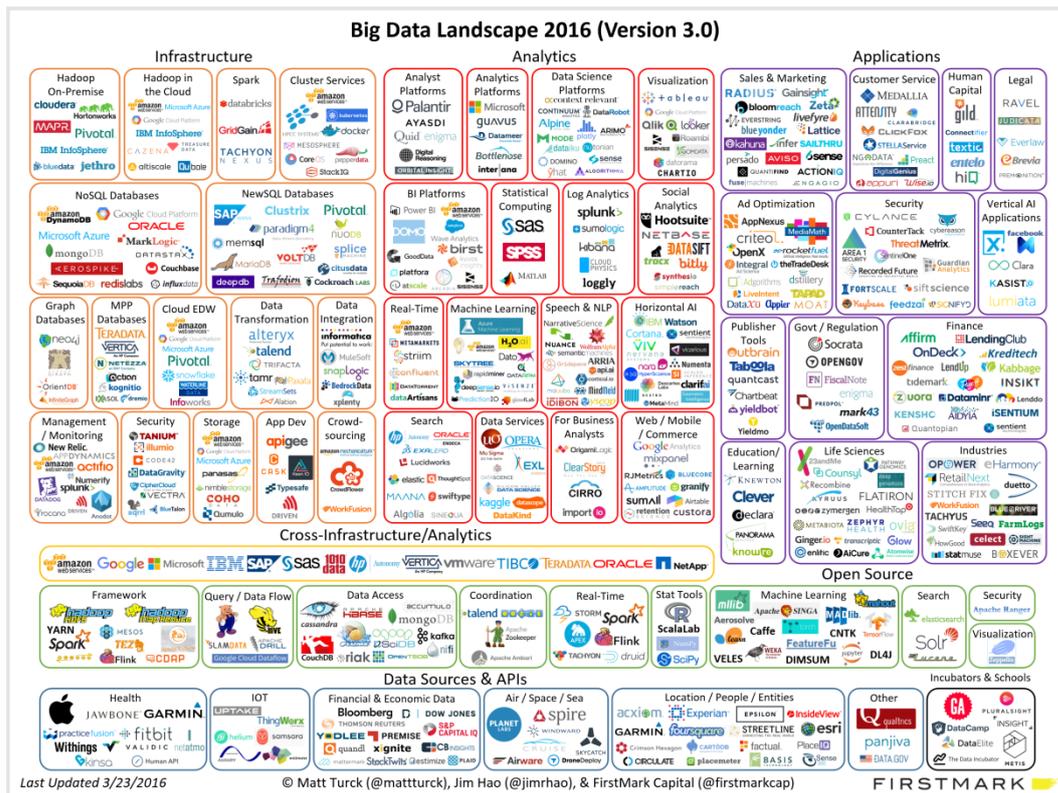
圖 2-12 美國 NASA 地球觀測站繪製全球空氣污染地圖

地圖能以視覺化方式呈現不同地區間的空氣污染程度對人體健康的直接影響，像是相較於淺棕色區域，地圖上越深褐色區塊則表示有越多人因空氣污染而死亡，至於藍色區塊則代表該地區空氣品質已獲得明顯改善。而目前在中國東部、印度北部以及歐洲城市皆顯示空氣污染影響最為嚴重。而在美國東南部的空氣污染濃度則有明顯改善。

## 8. 巨量資料之生態圈

由 Matt 製作的 Big Data 生態圈圖，如圖 2-13 所示。圖上的公司可分為架構、分析、應用、跨平台架構、開源解決方案、資料來源提供商六個領域。其中的公司如：IBM、Google、微軟、Amazon 等是跨平台解決方案提供廠商，統計軟體公司 SPA、SAS 也在裡頭佔一席之地。

因應 Big Data 的快速發展及帶動 ICT 產業的變革，台灣企業為能掌握產業趨勢及開拓市場新藍海，台北市電腦公會宣布「Big Data 跨域整合聯盟」於 2015 年 4 月 17 日成立，推動資料(data)應用與服務，將聚集國內資料擁有者、資料科學家及各領域專家的能量，共同建構大數據商業應用服務體系。Big Data 跨域整合聯盟的成立，就是希望帶動國內各大企業，將各自擁有不同領域的資訊，能結合成為有用的資訊庫。政府已提出開放資料、大數據及群眾外包等「科技三箭」，推動產業創新發展，聯盟將以建立跨界「資料交換機制」、培育產業「資料專家的人才」及組成「資料顧問服務團」，帶動各領域應用資訊來提升營運成效，協助產業升級與建立 Big Data 應用服務生態系，進而增加國際競爭力。



資料來源：@Matt Turck(@mattturck),Jim Hao(@jimrhao),&FirstMark Capital(@firstmarkcap)

圖 2-13 Big Data Landscape 2016 (Version 3.0)

## 第三章 交通大數據優先領域發展規劃

### 3.1 國外交通大數據案例回顧

#### 3.1.1 公路運輸領域

##### 1. 公車即時資訊發布

新加坡交通運輸部門（Land Transport Authority, LTA）推出手機應用程式 MyTransport.SG，透過與新捷運（SBS Transit）及 SMRT 巴士合作，取得更精確公車即時定位，旅客可確認車輛承載率及後續班次到站時間，以此提升旅運經驗。該 APP 會更新由車輛上傳的資訊，以色線表示車輛空間，如圖 3-1，綠色表示仍有座位、黃色表有站位，紅色則表示空間少。圖 3-2，旅客可藉由 APP 查詢車輛到站時間，以作旅運決策。藉由 LTA 的 MyTransport.SG 提供更多的公車資訊，旅客可更準確進行旅次規劃，更優化候車時間與轉乘等決策。



圖 3-1 MyTransport.SG 手機應用程式

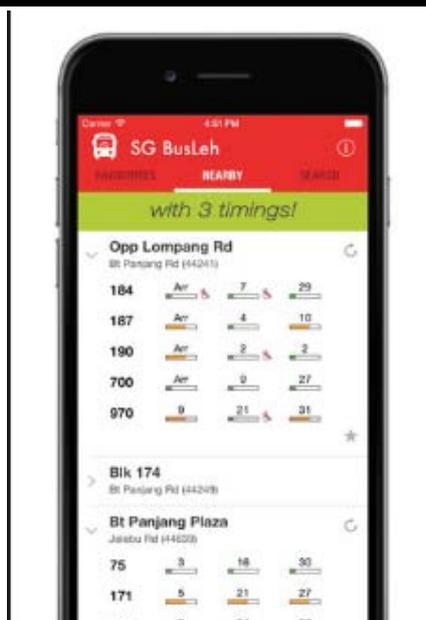


圖 3-2 MyTransport.SG 公車到站時間

## 2. 透過資訊視覺化輔助決策分析

大都市的交通運輸部門每天都需要針對如何有效分配資源做決策，包含如何使營運路線合適且可靠、何時增加熱門路線運能等，過去運輸規劃者依據可取得的資訊做出決策，而資訊來源包括調查及大眾反應回饋。新加坡陸路交通局（LTA）運用 BusViz 系統，以視覺化分析方式協助營運管理者和公共運輸監理者進行更有效監控及改善績效。

BusViz 為一基於網路的工具，由三部分組成：第一為演算層，執行多工運算以採掘 EZ-Link 票卡中的交易資料；第二為 API 層，負責介接 PLANET 資料庫以進行演算；第三為前端視覺化層，用以存取分析資料並呈現於瀏覽器上。透過乘客票卡資料產生的起迄 OD 創造出動態運輸系統，並由此紀錄推論擁擠程度、候車時間、公車到站頻率等，如圖 3-3，BusViz 公車站牌視圖中圓圈顏色代表該班次承載率，圓圈大小區別單、雙層巴士，而監理單位以此評估車隊績效以及作為補貼參考依據。

規劃者進行運輸系統設計與分析時常以時空圖為工具，如圖 3-4 展示公車服務路線績效，軌跡重疊表示有公車群聚現象，以顏色呈現車輛承載率。此外，BusViz 提供地圖視圖以更宏觀的呈現方式，展示所有通過選定站牌的公車路線。藉由 BusViz，LTA 可減少人力成本，並作為診斷工具辨識公車群聚、增加班次以提高運能及監控服務水準。

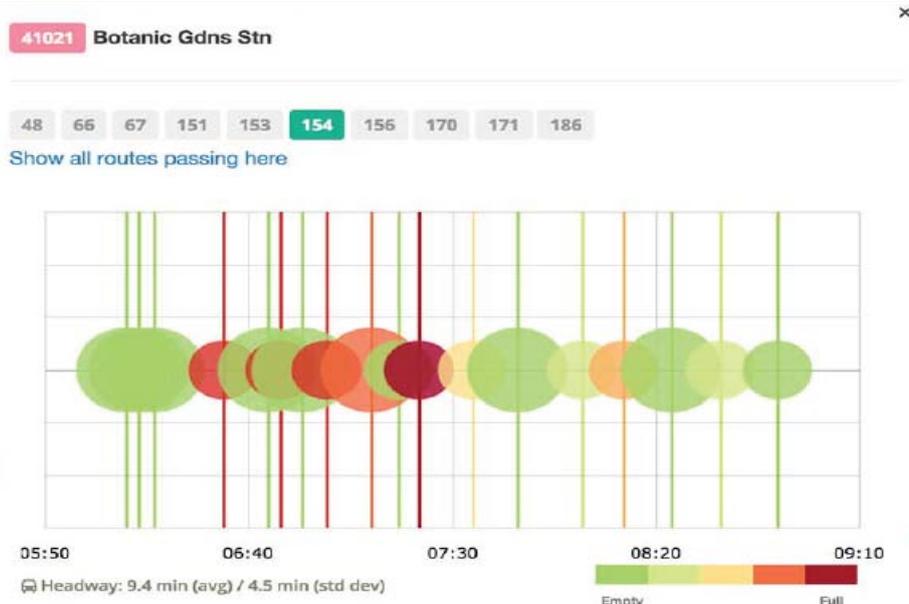


圖 3-3 BusViz 公車站牌視圖

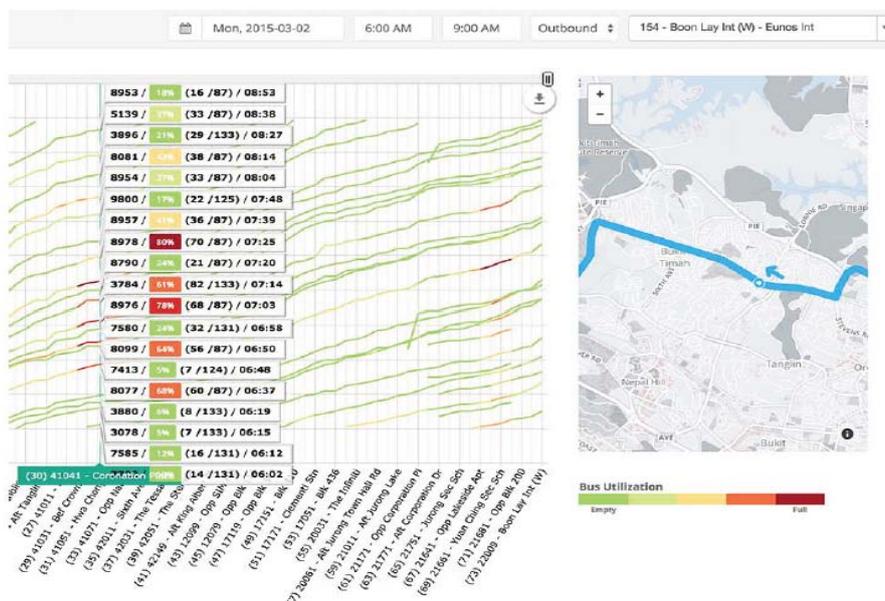


圖 3-4 BusViz 時空圖

### 3. 旅次起迄資料蒐集

為瞭解大量、連續由旅次規劃產生的資訊和傳統旅次調查法結果相比之差異，以紐約大都會運輸局（NYMTC）傳統家計旅次調查（RHTS）結果，對比智慧型手機應用程式 Transit App 資料，結果顯示儘管 Transit App 蒐集的資料維度少於 RHTS，然其巨量紀錄可與之抗衡。兩種調查方式皆可產生相同比率的旅次起迄對數，而 RHTS 時間及金錢成本遠高於 Transit App。

Transit App 是由蒙特婁一家新創公司所開發，超過 5 百萬人用來查詢當地即時交通時刻表，及規劃運輸旅次。以提供旅運資訊為核心，標示所有使用者附近的運輸服務最近班次，並提供自行車共享、汽車共享及共乘等訊息，如圖 3-5 顯示即時公車資訊（左）、旅次規劃（中）及共享運輸服務選項（右）。透過地理資訊等方法比對後可發現，如圖 3-6 和圖 3-7，兩組呈現結果相當相似，此外透過 App 蒐集的方式更能呈現出 RHTS 無法調查出的旅運起迄。

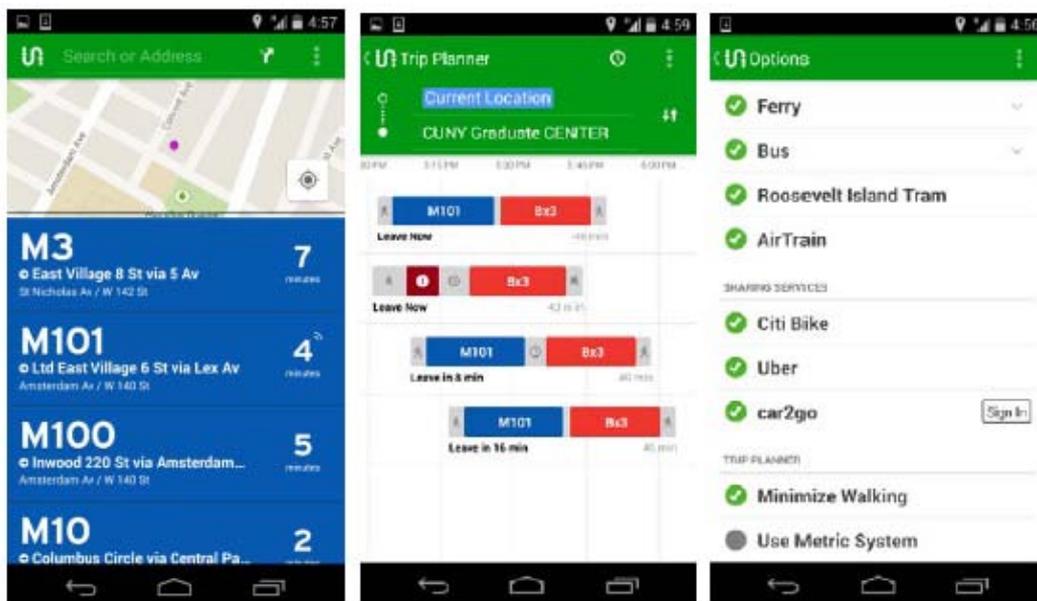


圖 3-5 Transit App 介面

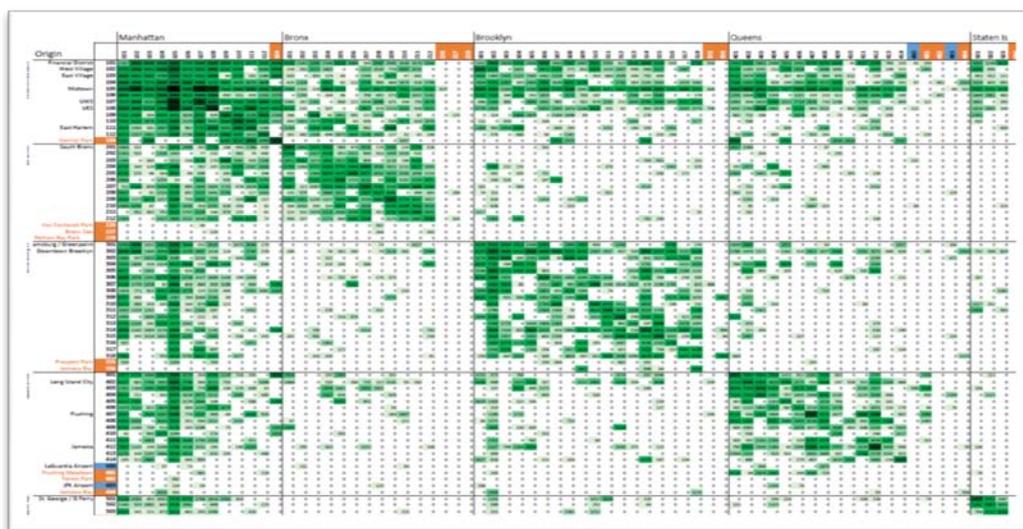


圖 3-6 RHTS 方法的 OD 調查結果



圖 3-7 App 蒐集的 OD 調查結果

#### 4. 公車服務可靠度改善

瑞典斯德哥爾摩為衡量交通政策在公車幹線實施效果，以車輛自動定位（AVL）與乘客自動計數（APC）資料進行大數據分析，結果證實當政策施行提升 7% 公車運行速度，可降低旅客 10% 的車內時間，行車週期縮減 8 分鐘，且車隊規模可減少 2 輛，此外規律性及群聚現象改善則可讓乘客減少 25% 的候車時間。

服務可靠度關係著服務水準、系統績效及效率，因此是旅客及業者關注的議題之一。由於車輛的站間運行時間和停留時間構成旅行時間及等候時間，所以可從三方面著手來改善服務可靠度：優先權（priority）、控制（control）及營運（operational）。以斯德哥爾摩熱門的公車路線為對象，實施改善規律性、提供更快速上下車、減少公車壅擠度、增加服務可靠度及減低旅運時間等措施。透過車輛自動定位和自動乘客計數資料以事前事後方法衡量服務品質，包含車輛及旅客各種旅行時間。政策實施前後的旅行時間變化如圖 3-8，而圖 3-9 分別表是南下和北上不同 OD 的旅行時間，綠色表示比表定時間短，紅色代表比表示時間長，經過改善後，全線只剩一小區域的旅行時間比表定時間長，表示平均對每位旅客而言，時間成本是降低的。

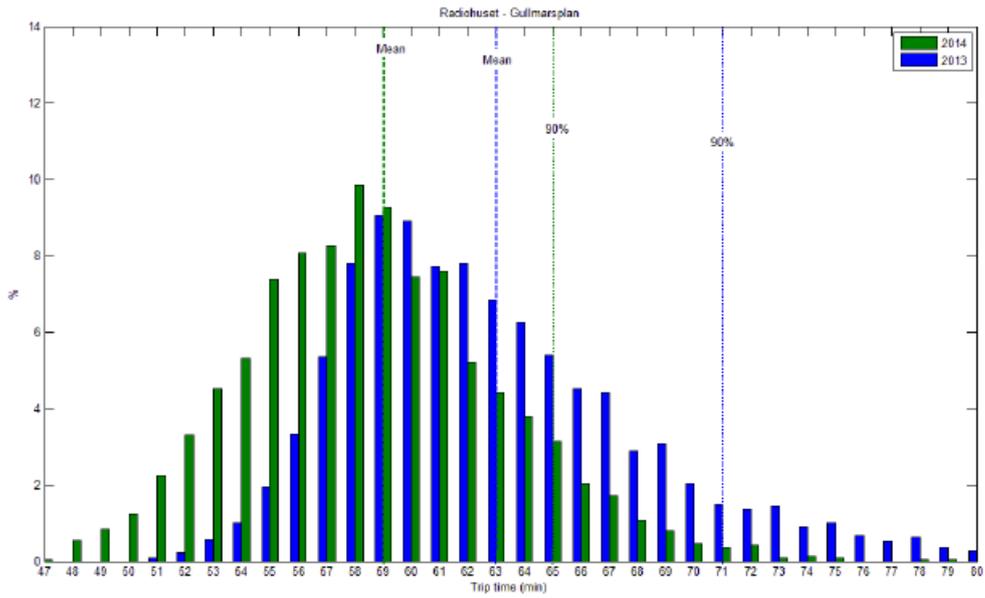


圖 3-8 公車總旅行時間

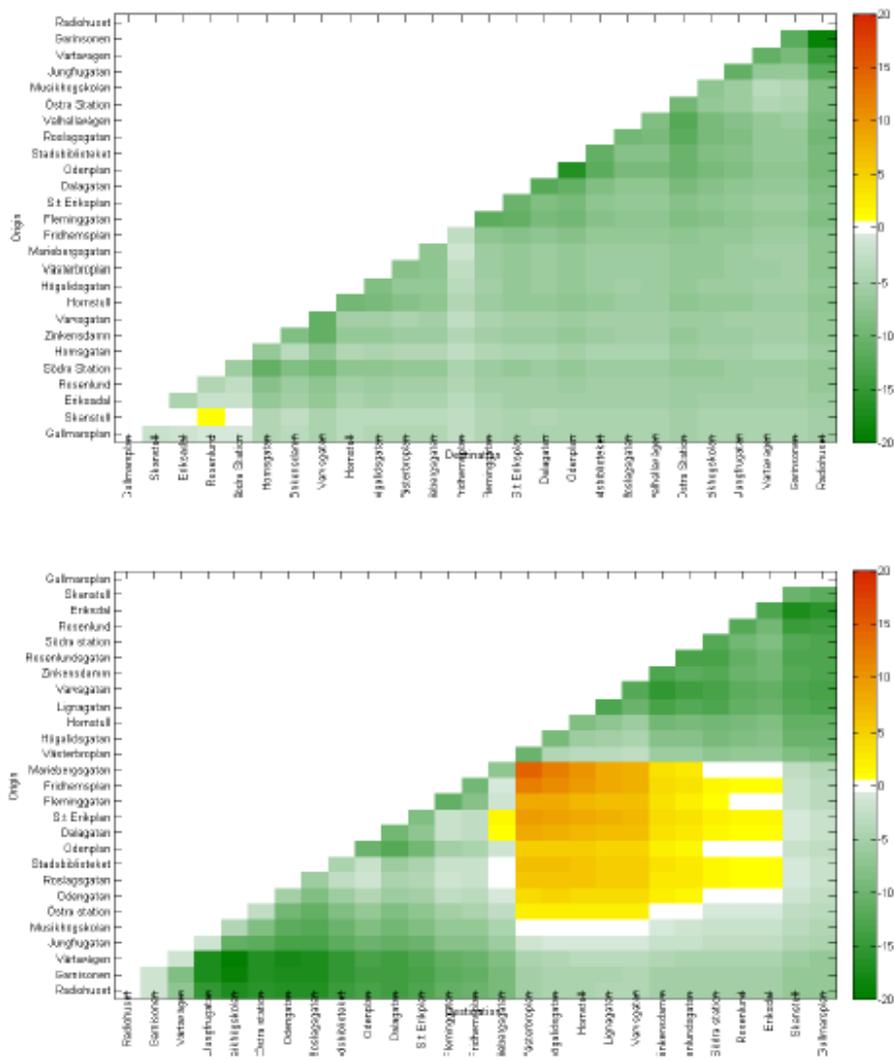


圖 3-9 旅客總旅行時間比較

## 5. 掌握公共運輸品質與使用率

瞭解公共運輸的品質及使用率對於優化排班和資源配置相當重要，其中乘客數和準點率乃評量服務品質的兩維度，其可透過自動車輛定位（AVL）、自動乘客計數（APC）及全球定位系統（GPS）取得資料。為了發展公共運輸資料視覺化，先進運輸技術中心（CATT）和馬里蘭大學透過MetroViz分析維吉尼亞三年的公車資料<sup>1</sup>。

MetroViz 為一可互動的視覺化分析工具，協助使用者探討公共運輸資料並評估系統服務品質。MetroViz 將介面分為三：地圖視圖、路線視圖及日曆視圖，如圖 3-10，其中前二項可讓使用者在地理或路線背景下清楚定位，而日曆視圖透過顏色和色塊數量分別呈現每個日期的準點率及乘客數。在資料處理部分，為了套用到不同運具上，將資料模型分為三層級，如圖 3-11 所示，(a)為場站層級、(b)以場站為基準的路線層級及(c)以場站和路線為基準的旅次層級。

該研究結果顯示透過準點率及乘客數的視覺化，規畫者能夠清楚定位、觀察異常事件，並發掘尖峰時間而未準點的車輛班次，以及進行其他交通管理項目。

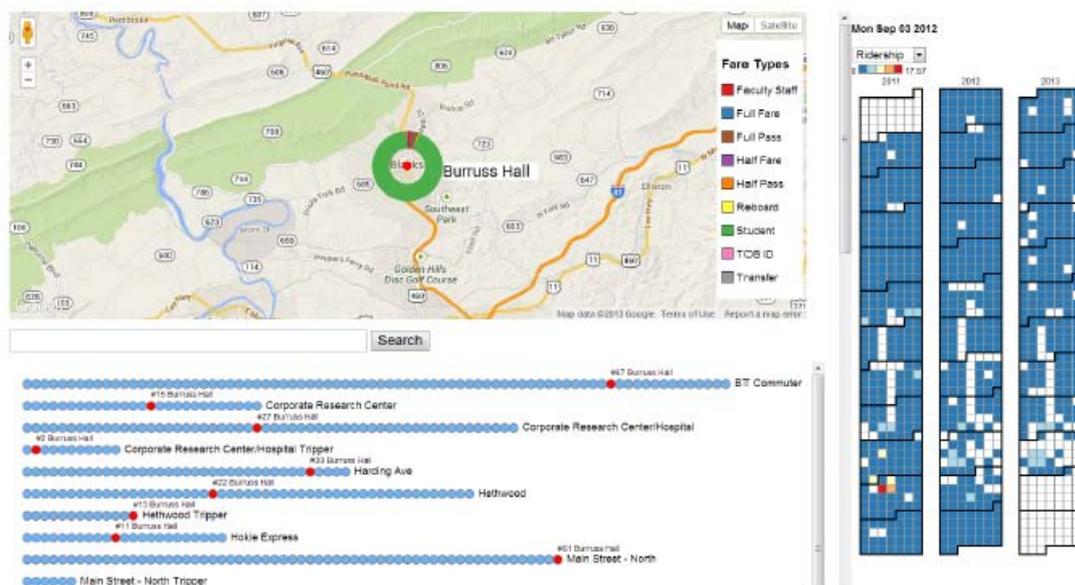


圖 3-10 MetroViz 介面

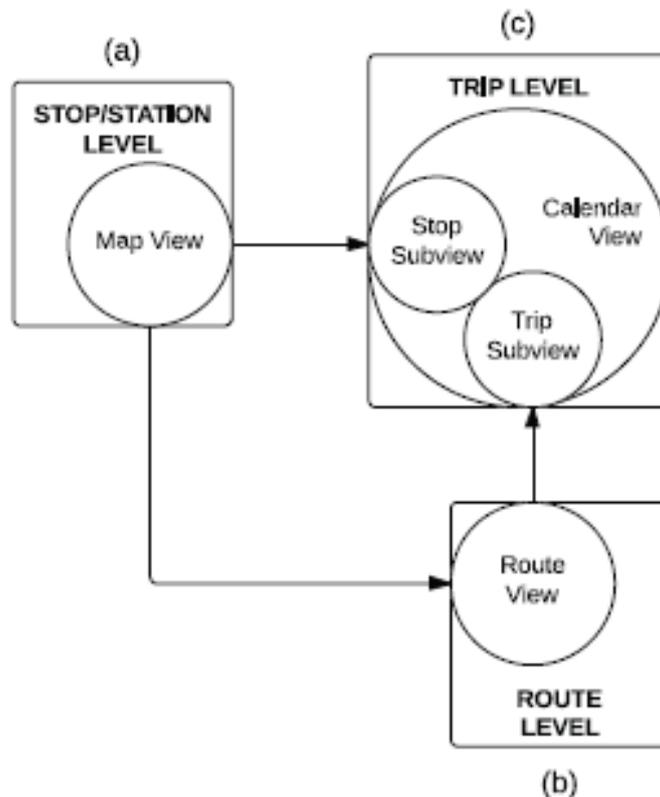


圖 3-11 MetroViz 資料模型

## 6. 公共運輸訂價

土耳其伊斯坦堡（Istanbul）公車捷運系統（Bus Rapid Transit，BRT）將大數據技術應用於自動蒐費系統（Automated Fare Collection System），將蒐集而來的旅運資料用於公共運輸訂價策略。另也透過旅運及研究途徑於運輸計畫管理，提供改善 BRT 計畫及管理建議，並提供圖表讓數據可更完成及有意義的呈現。

伊斯坦堡位於土耳其西北側，是土耳其最大城市，截至 2014 年底止居住人口為 1403 萬。伊斯坦堡的通勤者每天約有 280 萬旅次駕駛小汽車通勤，約占所有旅次 22%，為造成尖峰時段壅塞之因素。公共運輸部分，共有 8 家不同的公車業者，其他還有 Jitney、小巴士（Minibus）、計程車、通勤接駁車等等…。

BRT 系統透過智慧卡（Smart Card）：Istanbulkart 進行旅運資料蒐集，從 2004 年至今市占率為 95%，如圖 3-12 所示。伊斯坦堡公共運輸費率共有 37 種不同媒體類型（Media Type）及 69 種不同費率型態（Fare Type），蒐集不同費率靠人工判斷費時，因此當地透過驗票機（Validator）進行扣款、查驗的工作，並將資料透過網路傳至中央儲存系統進行應用。

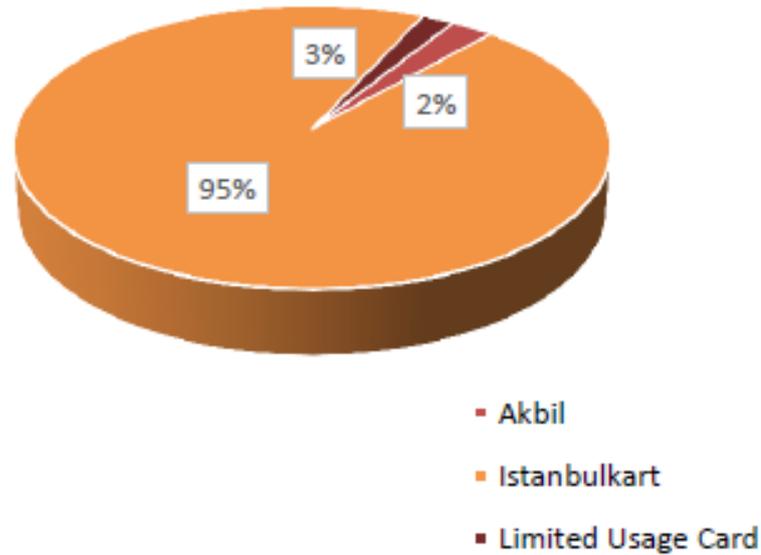


圖 3-12 伊斯坦堡 BRT 路線智慧卡使用比例

透過驗票機可以蒐集到的可用資訊有：數萬人的個人旅運資料、每卡的資訊、連續型的旅運資料、使用者的使用頻率，並連結同質性的運輸資料，透過套裝軟體加以分析，獲取分析結果中重要的資訊，如圖 3-13。

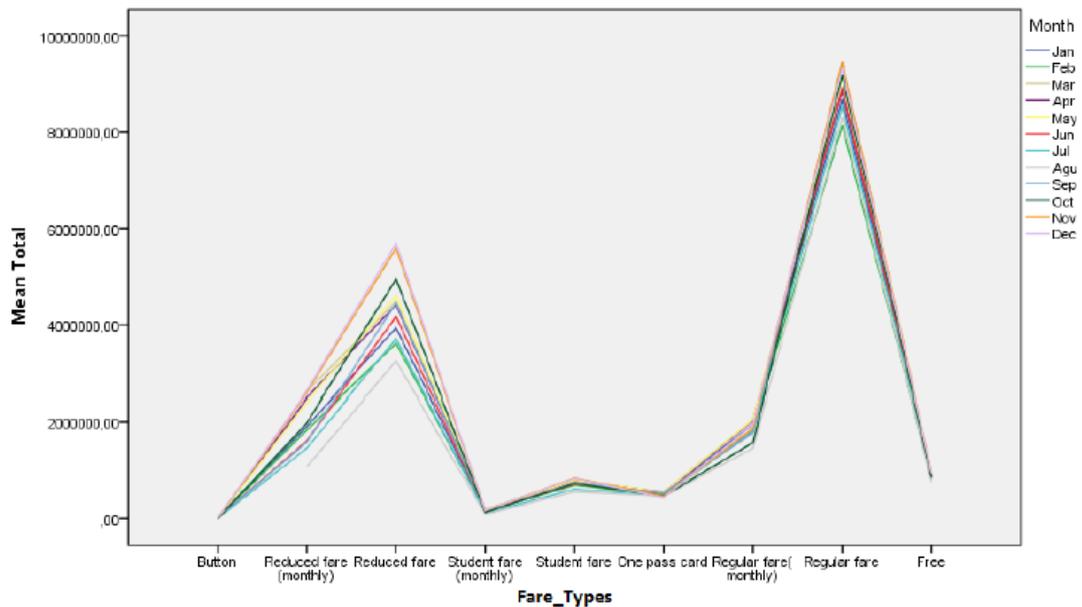


圖 3-13 伊斯坦堡利用大數據分析票價型式圖型化示意圖

## 7. 夜間公車需求與規劃

韓國首爾自 2013 年推出午夜公車服務，如圖 3-14，該項服務是由首爾市政府團隊與韓國電信公司（Korea Telecom）共同合作，透過調閱民眾深夜的通話及簡訊發送資料，比對每個用戶的「夜間發話地點」和「寄送帳單地址」，只要發現「發話地點」和「帳單寄送」地址不同，就表示用

戶人不在家，就可能有使用大眾運輸系統的需求，也能預測此些民眾返家之方向。



圖 3-14 韓國首爾貓頭鷹公車服務

在用大數據來分析午夜過後市民擁擠度最高的 5 個地區，以及 30 億通話紀錄及交通資訊後，所設計推出的午夜公車的路線，驗證了原本策畫團隊的假設推出「貓頭鷹公車」服務，從試辦計畫的 2 條路線(N26 和 N37)，經由深夜時間流動人口較多的弘大、鍾路、江南站等地，各路線共投入六輛公車，獲得優良的成效後，已擴增至目前的 9 條路線。蒐費以交通卡為基準，車資為 1850 韓元，並可透過公車站的公車動態資訊系統、交通資訊中心行動網站、首爾大眾交通應用程式等來得知深夜專用市內公車的抵達時間。

「貓頭鷹公車」行駛 50 天 累積搭乘人數就達 30 萬 3,940 人，每天平均搭乘人數有 6,079 人。這代表每輛公車每天平均有 138 人搭乘，每一個公車站約 5.8 人上車（總公車站數為 1,056 個）。深夜專用市區公車擴大營運之後，也使得拒絕載客的計程車申訴件數相對降低，是另一種計畫所帶來的成效。

由於公車路線晚上的時間和白天的時間不一樣，人的目的地也不一樣，首爾市政府例用手機大數據分析民眾移動的方式、距離及目的地，規劃出與日間不同路線的夜間公車，解決了夜間沒有地鐵服務的需求，也提升了市政管理之效益。

### 3.1.2 鐵路運輸領域

#### 1. 平交道事故報告分析

根據美國鐵路管理局（FRA）統計，1999~2013 年平交道事故發生近 4000 件，其中 243 件死亡，943 件受傷。羅格斯大學發展非統計工具進行數據、文字探勘以及資料視覺化，以協助鐵道工業（RR industry）對平交

道事故的執行與了解為目標，尤其對於卡車及拖板車事故發生的認識，提出研究結果：以鐵路干預降低平交道事故發生率、並提供各政府機關於標牌、速限與駕駛訓練改進降低事故。

羅格斯大學透過 Weave 以儀錶板呈現，如圖 3-15 所示，將事故統計以長條圖區分，可觀察「車輛於平交道未停駛」易發生處。文字探勘部分則採用 LDA 以主題模型 (topic modeling) 演算法進行，從事故報告中將原始文字歸類，並可進行搜索，如圖 3-16，"shoving" 事故發生在 "yard" 與 "side" 中。此外，將事故透過不同分類方式以視覺化呈現，如圖 3-17 以車輛類別區分，統計出文字頻率高低展示之，可直接觀察出目標車種發生事故比例，例如卡車事故、拖板車事故等。

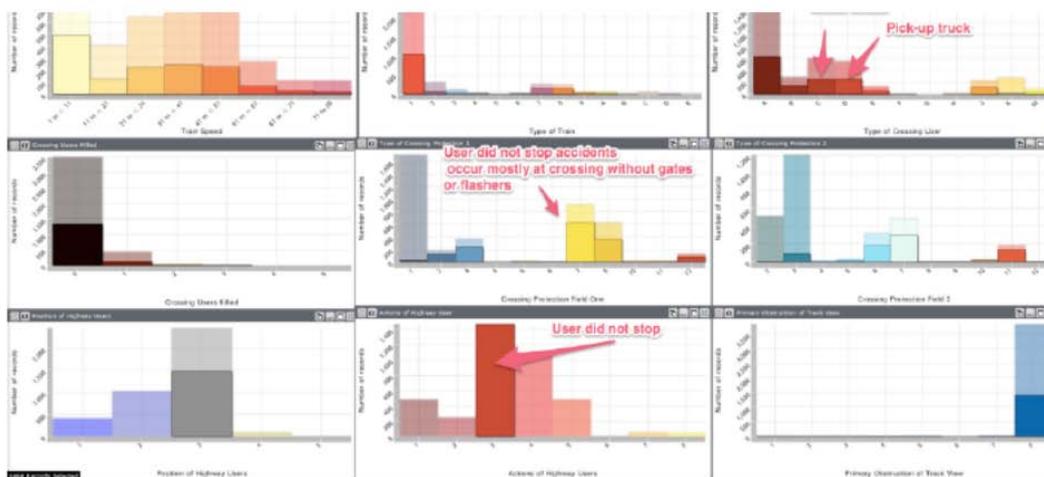


圖 3-15 行車輛未停駛事故統計

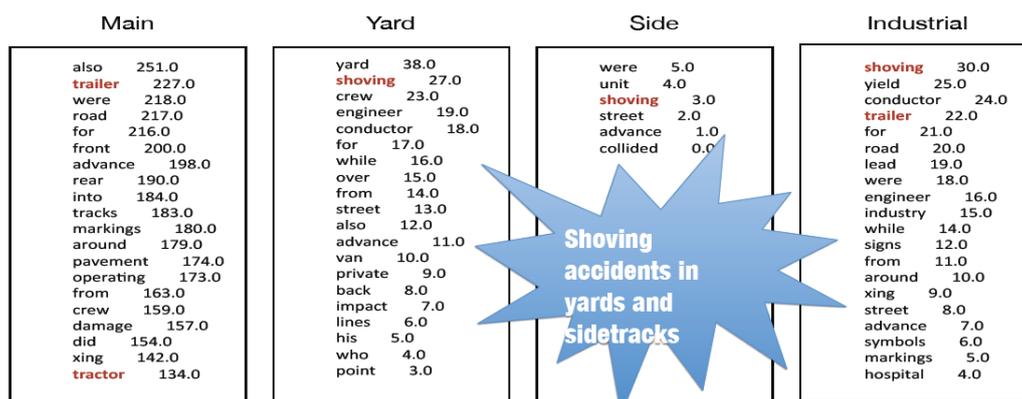


圖 3-16 文字探勘以軌跡類型分類

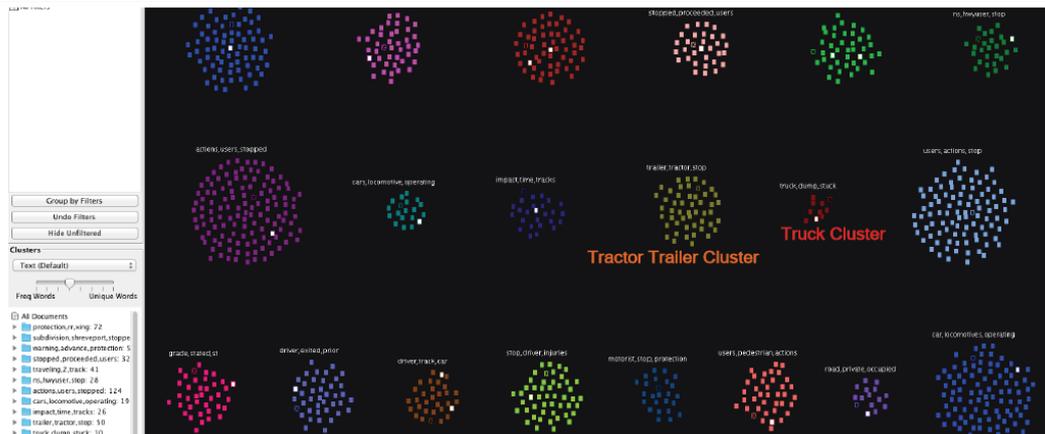


圖 3-17 文字頻率分組事故

## 2. 鐵路維修管理決策

鐵路延誤、中斷不僅提高營運成本也降低旅客滿意度，英國鐵路受天候因素影響造成號誌異常、列車故障及壅塞，因此，政府及營運單位為有效改善通勤的服務水準，運用大數據進行預防性維修管理（Predictive Maintenance, PM）。以該方法透過列車感測器蒐集資訊，達到辨識隱性故障，預測最佳維修時間，最終使營運者在列車故障和服務受影響前做出反應。如圖 3-18 所示，利用大數據進行維修時間優化，處理程序成本雖然高，相對而言效益更大，倘若能夠主動辨別出隱性故障，則可事先安排必要維修，將效率發揮到極大。

以航空業為借鏡，有效執行 PM 的步驟如下：(1)選擇正確的系統或子系統做預測，清楚界定範圍並且不預設立場地猜測，此為建立 PM 的關鍵。(2)取得所需資料，以此做為發展演算的基礎。(3)確定 PM 為維修策略的增值服務，PM 不僅以預測失敗為目標，更須在多變的企業系統中做出危機反應，作為維修計劃的支援。(4)加強鐵路專業知識方面的資料科學。(5)考量資料科學的風險，例如開發人員的經驗。

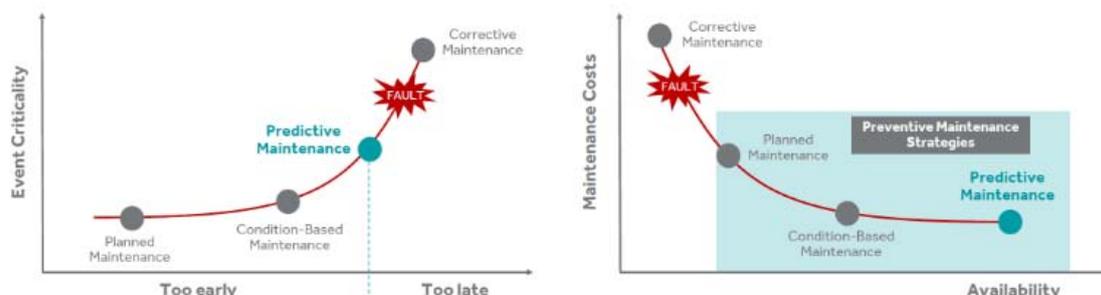


圖 3-18 優化維修時間

### 3. 鐵路列車延滯控制

未來鐵路產業的投資不僅於硬體設備，更需要數位基礎設施並提供即時營運資訊，因此瑞士聯邦鐵路（SBB）和 TIBCO 合作發展大數據應用。為改善旅客感受，SBB 透過 TIBCO 連結營運單位，降低延誤及事件排除時間帶給乘客使用上的優勢。

TIBCO 是設備和 BI 軟體領導者，專長於協助企業讓內部資料產生價值、並從隱藏事件中作出即時反應。SBB 利用 TIBCO 技術進行列車延滯控制，使抵達延誤時間不超過 90 秒，該系統提供精確、即時的列車定位，提高 SBB 系統中列車移動的掌握度，從而提高訊息正確性，如圖 3-19，透過 TIBCO 為中介以轉換、傳遞資訊至控制中心。此外，SBB 利用 TIBCO 模擬列車延誤情境，演練可能造成數小時誤點狀況的因應策略。

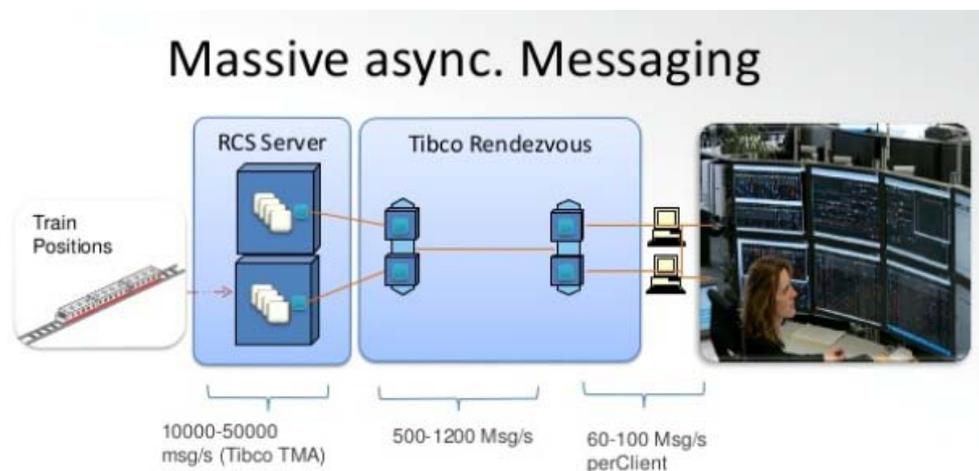


圖 3-19 SBB TIBCO 系統示意圖

### 4. 鐵路通勤預報

列車運行不順暢原因包含交通事故或列車誤點，通常一個班次延誤會造成後續班次誤點，甚至系統癱瘓，因此瑞典鐵路推出「通勤預報」（The Commuter Prognosis）。

Wilhelm Landerholm 與斯德哥爾摩交控中心共同研發演算法，透過數學模型解讀大數據以預測列車路網中的所有車輛，對未來 2 小時將發生的誤點在確實發生前提出警告，同時計算出該誤點對系統其他列車影響程度，並自動發出訊息給旅客，如圖 3-20 示意。

通勤預報以帶給旅客更多生活便利及旅運資訊為目的，提供乘客更精確的旅運管理，並減少旅客等候時間。此外藉由預測服務中斷，行控中心可預防漣漪效應造成更大的延誤，而此演算法未來亦可套用至其他公共運輸系統。



圖 3-20 通勤預報示意圖

## 5. 鐵路維修管理系統

瑞典鐵路局 (Swedish Railways) 中的基礎設施管理者將數百項數據進行處理，資料分別為結構化資料 (Structures)、類結構化資料 (Semi-Structured) 及非結構化 (Unstructured) 資料 3 大類，結構化資料包含：時間、速度及使用量；類結構化資料為：圖片、影像等；非結構化資料為維修紀錄。

在資料蒐集的過程中，將可能會有錯誤的數據被蒐集到，若是貿然使用將會造成較差的結果，為了獲取正確且有用的資料，移除或正規劃是必須的，或是透過偵測演算法也可以小規模的解決這些問題。透過新的管理工具，可進一步找出有用的關聯及移除不具意義的錯誤值，提供未來決策時更正確的資訊。

瑞典鐵路局透過路側設施或維修車輛蒐集所需資訊進行管理，雖然所蒐集的資訊來自於不同資料庫或系統，且為了分述不同問題，資料包含了存貨、軌道、環境資料及人力資料作為考量的資料，再使用大數據進行分析，提供做為維修的決策分析。

瑞典鐵路的基礎建設系統分為幾項子系統進行蒐集，分別為：列車誤點統計 (TFOR)、軌道資訊系統 (eBIS)、軌道位置測量系統 (Strix)、線上分析測量系統 (Optram)、軌道價格系統 (AGRESSO)、保密資訊系統 (BANSTAT)、錯誤偵測 (OFELIA)、維修檢測及安全系統 (BESSY)、營運及維修合約商管理系統 (Rufus)、Daily Graph、Maps、降速資訊 (HANNES)、天氣資訊 (RWIS)、列車位置追蹤 (Lupp)、列車追蹤 (Here and Now) 及營運及維修分析工具 (Duvan)，如圖 3-15。

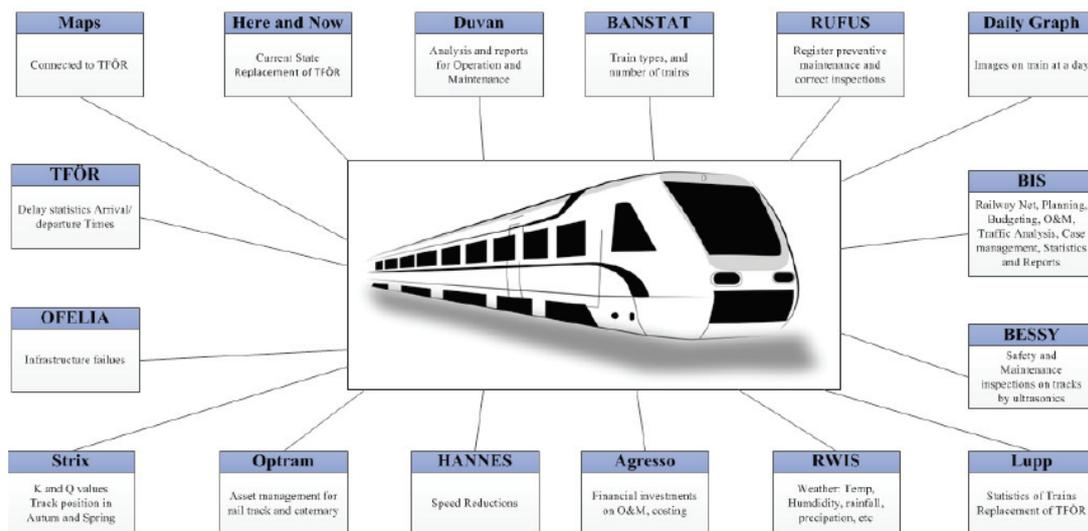


圖 3-21 瑞典鐵路大數據管理

## 6. 旅運型態分析

智慧票卡和資料自動搜集系統 (ADCS) 在交通運輸領域的應用，讓規劃者能夠分析關於通勤者行為及旅運模式大量連續的時間序列資料，而為了使時間相關資料成為知識，新加坡 MRT 進行時間序列資料的挖掘。

新加坡公共運輸包含計程車、公車和捷運，其交通政策採行軸輻式路網，系統中以公車服務城鎮到樞紐中心，捷運負責輸紐中心之間長距離運輸。新加坡公共運輸使用率雖以公車為大宗，但捷運持續增加，LTA 以捷運為發展目標，目前共 102 個 MRT 站與 21 個 LRT 站，路線長 149 公里，預計 2020 年覆蓋率達 278 公里。為了提升公共運輸使用，新加坡政府著墨於瞭解通勤行為及旅運模式。

研究以 LTA 的 EZ-Link 卡 2011 年 11 月交易資料為樣本，擷取起站、迄站以及進站時間記錄，欄位經排列後以 SAS Enterprise Miner 進行時空序列資料處理，經試誤法將一周數據分為 11 個叢集，如圖 3-22 所示，每個叢集分別呈現該時空下旅運特性，如 A 為晨峰強/昏峰緩和；B 為晨峰強，但與 A 相比可發現 B 旅客量較少，原因推測為 LRT 容量較小；C 為昏峰強，乃因其 MRT 通過工商業區域等。藉由時間序列採礦技術，運輸規劃者可以更快且更地深入分析旅運模式，而此時間序列架構未來可用於其他運輸系統中。

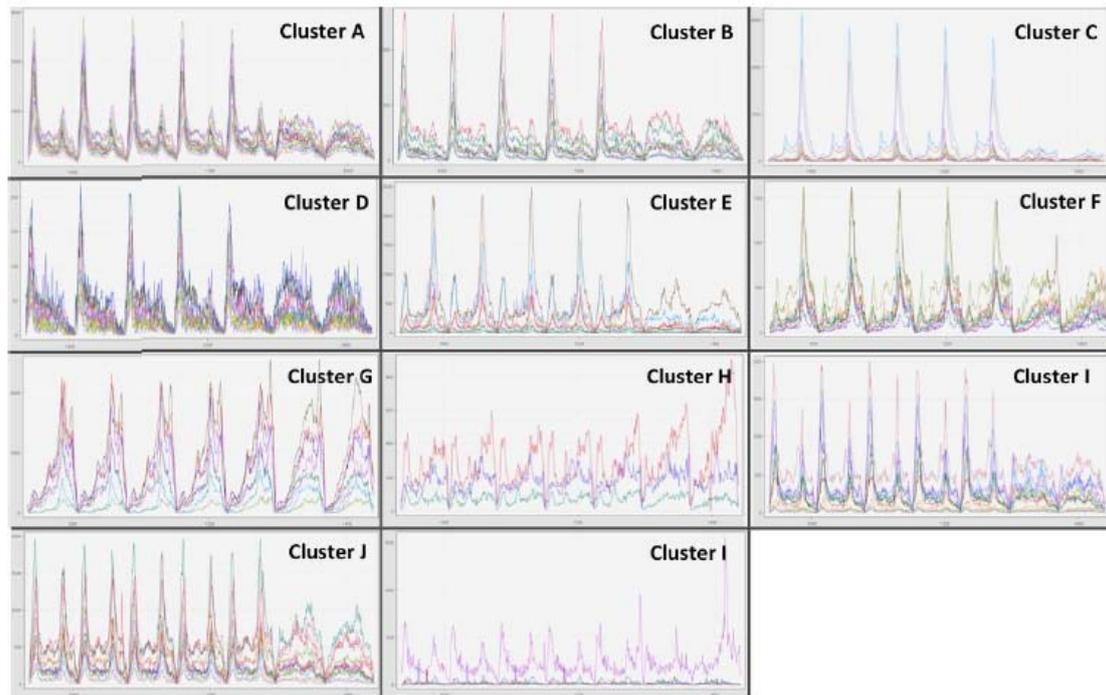


圖 3-22 時間序列叢集分析圖

### 3.1.3 航空運輸領域

#### 1. 航空數據分析研究

美國波音公司（Boeing）與卡內基梅隆大學（Carnegie Mellon）合作建立航空數據分析研究室（Aerospace Data Analytics Lab），利用人工智慧（AI）和大數據對波音飛機大改造，改善在空域活動時的品質。波音公司與卡內基梅隆大學的合作將大量數據透夠過機器學習法以優化飛機運行，讓數據導引未來工作，包含設計、建造和營運。

波音飛機上有數千個感測器，每個航行過程中會不斷產生大量數據，如圖 3-23 以波音 737 飛行 6 小時跨國航程為例，一架飛機每個引擎每小時可產生 20TB 的數據，一年共可取得將進 25 億 TB，包含飛行員和機器文字資料、結構工程數據庫等，經由卡內基梅隆大學進行數據處理，波音公司將能預測預防性維修可能出現的問題，而非等到問題出現或飛機降落才解決。透過機器學習幫助飛機保養和維護，預先決定維修時間，而非遵循固定保養週期，如此能避免飛行中發生故障，保護乘客安全。

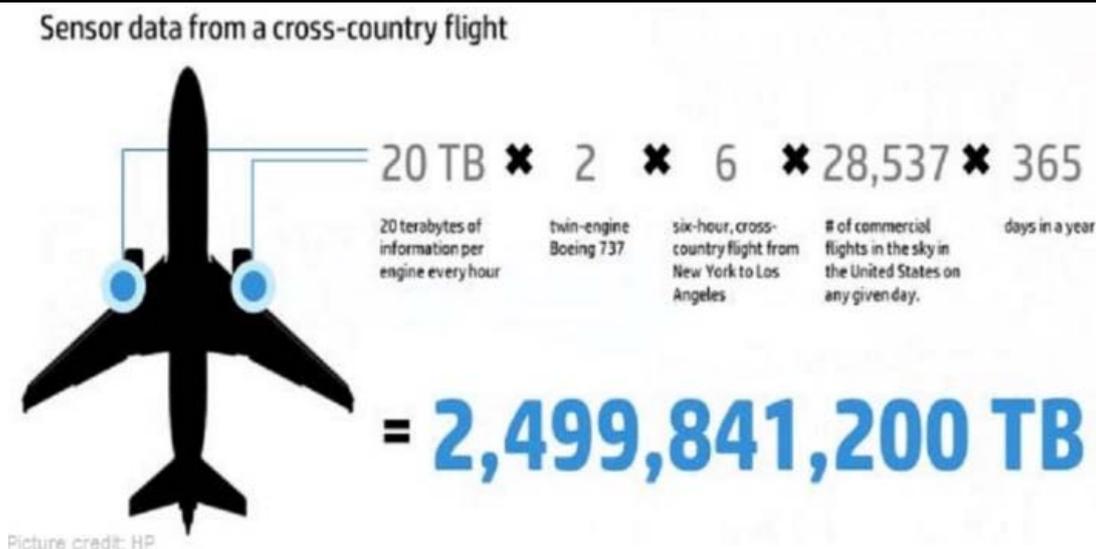


圖 3-23 感測器數據量

## 2. 航空時刻表恢復

澳洲航空（Qantas Airways）使用 Amadeus 推出的時刻表恢復系統（Schedule Recovery）解決天氣或交通管制等外部因素造成的航班計劃失衡問題，共分為四個構面，如圖 3-24 包含：飛機運作、顧客服務和成本控制、企業整合及決策支援。此技術基於一個推薦系統引擎（recommendation engine），該引擎透過數據分析，提醒航空公司可能發生的意外狀況，並辨別出緊急應該立即據此做出調整的問題，如延遲或取消航班、更換飛機或重新分配降落時段（Landing Slots）等。

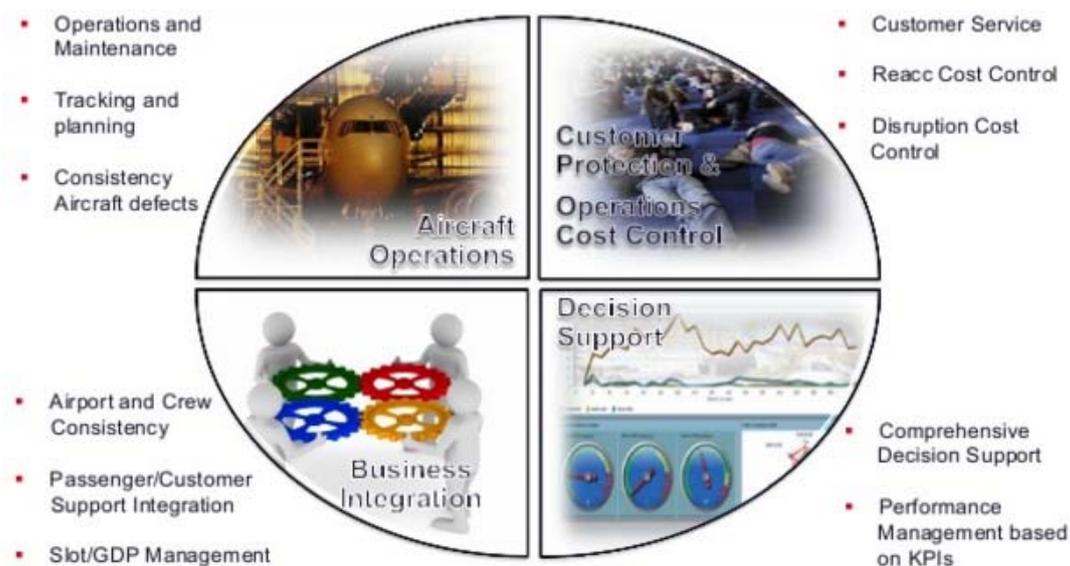


圖 3-24 澳洲航空時刻恢復表

時刻表恢復系統包含機場資源追蹤 (Airport Resource Tracker)、時刻表管理 (Schedule Manager) 及時刻表優化 (Schedule Optimiser) 三個模組。其中，機場資源追蹤讓航空公司有效率地與航空管制部門溝通，提高機場現有起降時刻的使用，亦能緩解機場起降能力不足問題，例如改善停機、登機口和行李傳輸帶的利用率。航班延誤是許多旅客的痛點，透過時刻表恢復系統提供更多訊息，降低突發事件造成的影響，減少航班延誤次數及時間，最終提升旅客體驗。

航班時刻表恢復系統能減少人工作業，在進行決策時將營運成本及對乘客影響同時納入考量範圍，因而能夠提升決策品質，亦達到控管航空公司營運成本。由於降低對乘客的影響程度，以及提升了在航班計劃失衡時恢復營運的速度，整體而言提升了決策品質及效率。

### 3. 營蒐管理與機隊控制

阿提哈德航空 (Etihad Airways) 為阿拉伯聯合大公國其中一國的國家航空公司，成立於 2003 年，總部設在阿布達比，航點遍布於中東、歐洲、亞洲、印度半島和北美等地。為了達成 21 世紀全球性航空公司的目標，阿提哈德航空採用大數據，從優化定價策略以獲得最大化蒐入、預防性維修到提升乘客的機上體驗。

阿提哈德航空運用 PROS 發展的定價及營蒐管理軟體來處理每天龐大的數據量，數據源自於線上比價、購票、報到、座位選擇、個人化資訊和優惠等過程。阿提哈德利用 PROS 發展新航線目的地，並增加航空公司與目標客群間的連結，透過旅客目的地搜尋和航空公司的連結，提供機位及價錢組合。

此外，阿提哈德航空利用 Taleris 對波音和空中巴士進行面的機隊控制，如圖 3-25 所示，藉感測器蒐集各數據回傳到機場管理中心，協助航空公司即時監控飛行器、管理、預防性維修、提前阻止問題發生、減少油耗及停滯於機場的週轉時間。藉由增加飛行器在空中的時間並減少在地面時間節省大量航空公司成本。

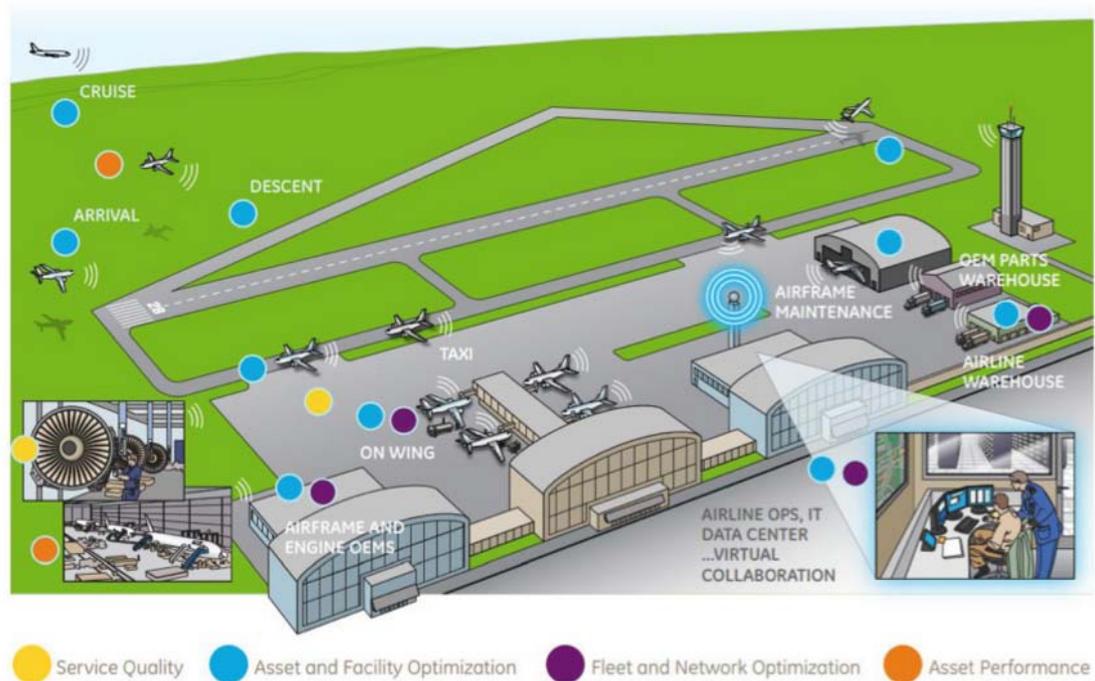


圖 3-25 Taleris 示意圖

#### 4. 優化航空營運

馬來西亞國際航空 (masFlight) 與航空情報機構 (OAG) 合作，提供空運訊息資料庫，包含：全球航班時刻、及時航班資料、飛機飛行路徑、美國機場基礎設施資料、全球機場天候及地圖、美國商業蒐益等資料，如圖 3-26。

副總裁 Tulinda Larsenn 博士分析指出，大數據分析可以解決航空管理問題，整合傳統平臺、綜合第三方資源、管理及時流量、分析大數據、付諸於 BI 及視覺化呈現，如圖 3-27 所示，大數據的運作過程，能讓營運效率增加。另外，雲端技術藉由網際網路進行資料的儲存、管理及處理。透過雲端技術及大數據分析產生更好的資訊以優化航空營運，減少航班誤點或取消航班的發生、有效降低油耗、改善資產利用率並增加營寬。



圖 3-26 masFlight 資訊平臺

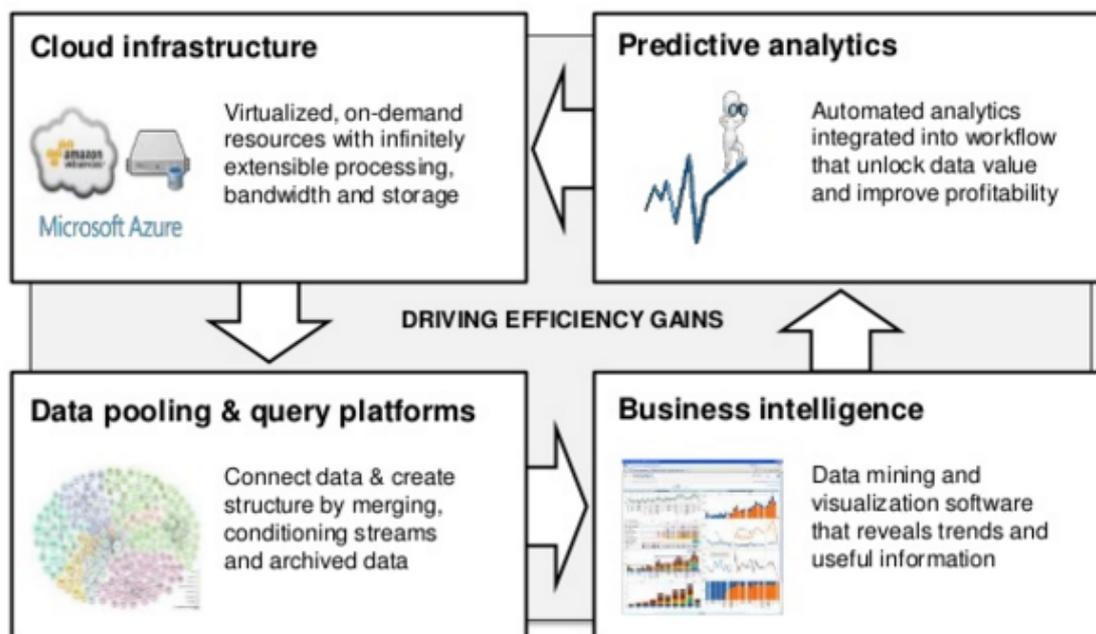


圖 3-27 masFlight 運用大數據

## 5. 航空領域環境衝擊研究

隨著科技進步，可用資訊量成指數增加，大數據分析方法儼然成為新一代資訊運用技術。作者以環境分析為例，機場管理當局進行永續發展規

劃或環境衝擊研究等，皆須報告碳排放清單，而過去文獻所採之推估方法，多數需要設定假設條件，例如假設飛機起飛時間為 2.9 分鐘、起飛推力為 100%（最大起飛推力）等，再透過燃油效率、排放系數等，估算出排放清單；然而藉由大數據分析，可提出更精確溫室氣體排放量之演算方法，如圖 3-28 和圖 3-29 所示。

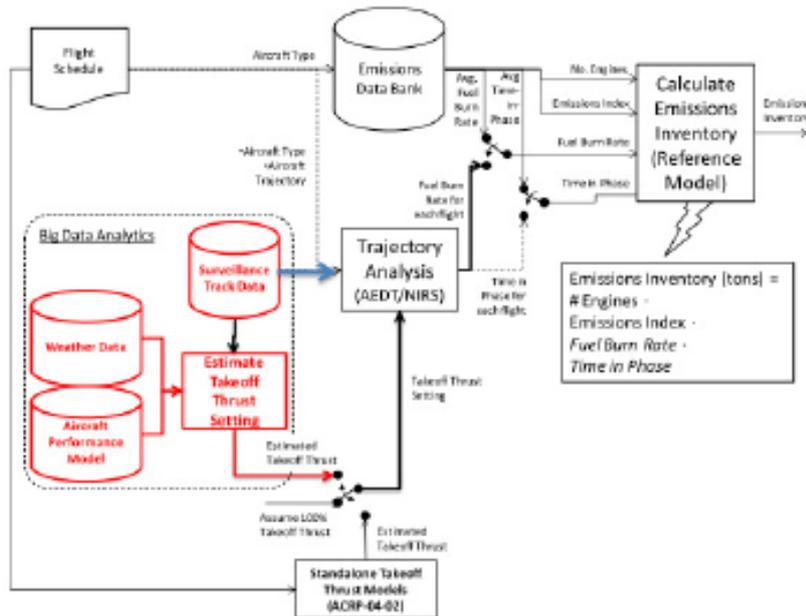


圖 3-28 以大數據分析進行溫室氣體排放量之演算過程

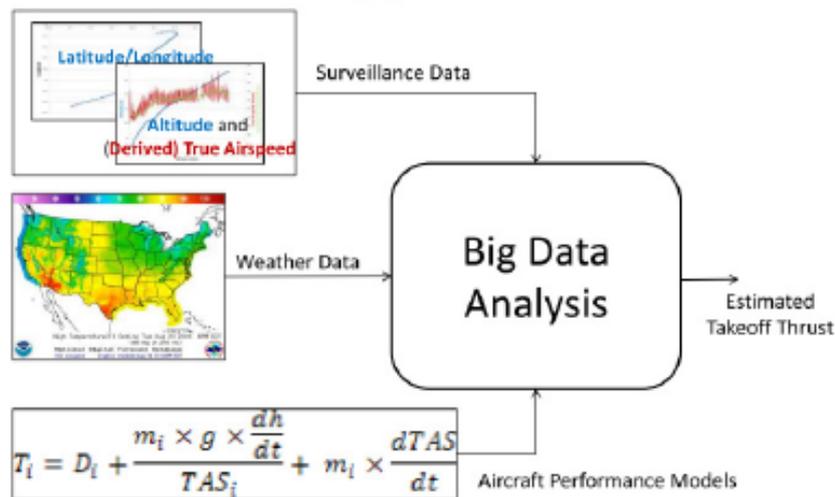


圖 3-29 以大數據分析計算溫室氣體排放

此外，作者針對機場登機門進行壅塞分析，透過 SOMAT（Surface Operation Management & Analysis Tool），可以針對個別機場探討其空側壅塞原因，藉此可與利害關係人進行協調解決，另可經由預測工具，估算可能發生壅塞或忙碌時段，以利管理單位提早因應。

## 6. 評估航空公司合併對機場之影響

大部分航空公司主管及華爾街人士對於航空公司合併抱持樂見的看法，將此視為產業的正向發展，然而對於中型及小型機場航空競爭力及票價有負面影響。作者檢視美國機場，因航空公司合併已產生贏者及輸者，且輸者將持續增加。

作者引用 OAG (Official Airline Guide) 資料庫資料，以 2000 年機場座位供給數為基準，彙整歷年各型樞紐機場座位供給趨勢，如圖 3-30，僅大部分大型樞紐機場於 2008 年金融危機復甦，並逐漸成長，對於多數中型及小型機場仍未回復，且大部分小型機場更有衰退趨勢。

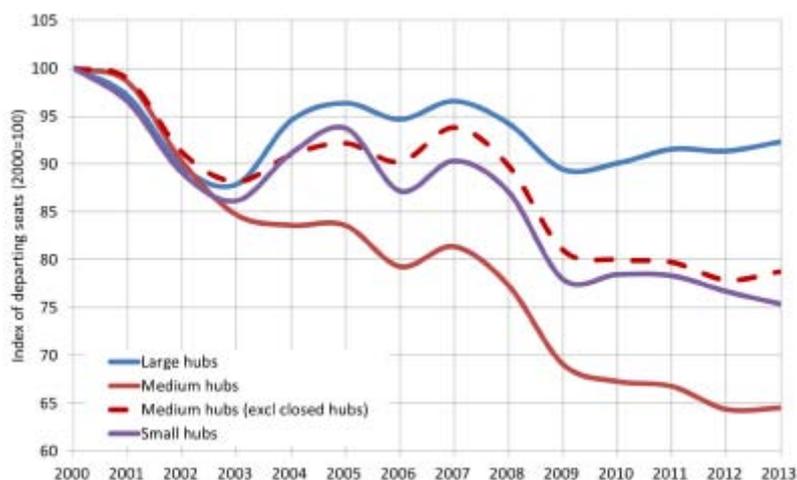


圖 3-30 樞紐機場座位供給趨勢

綜整會計年度 2001 年到 2003 年，美國主要客運機場客運量與排名，如圖 3-31，聖路易斯機場 (STL)、辛辛那提機場 (CVG)、匹茲堡機場 (PIT)、田納西機場 (MEM) 等客運量大幅衰退。

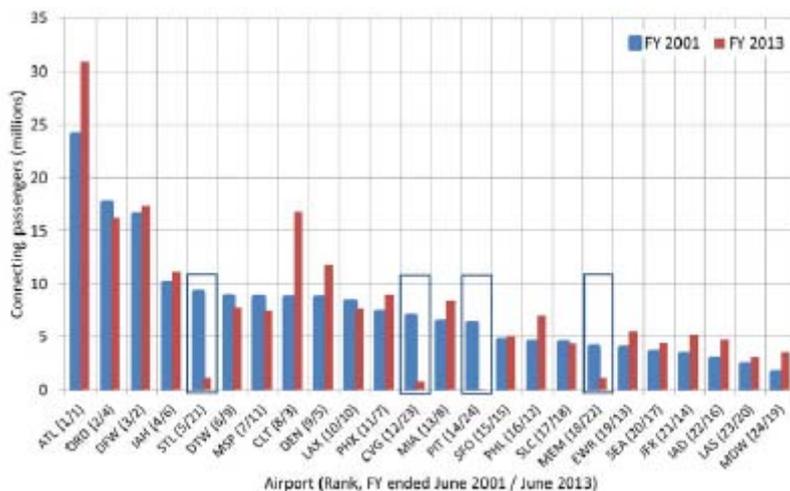


圖 3-31 美國主要客運機場運量情勢

全美航空 (US Airway) 於 2005 年被擁有美西航空 (America West Airlines) 的美西控股購併，2007 年美國政府核准以「美西航空」經營，由圖 3-32 可見，匹茲堡機場 (PIT) 與拉斯維加斯機場 (LAS) 明顯衰退，期中 LAS 為前美西航空樞紐機場。

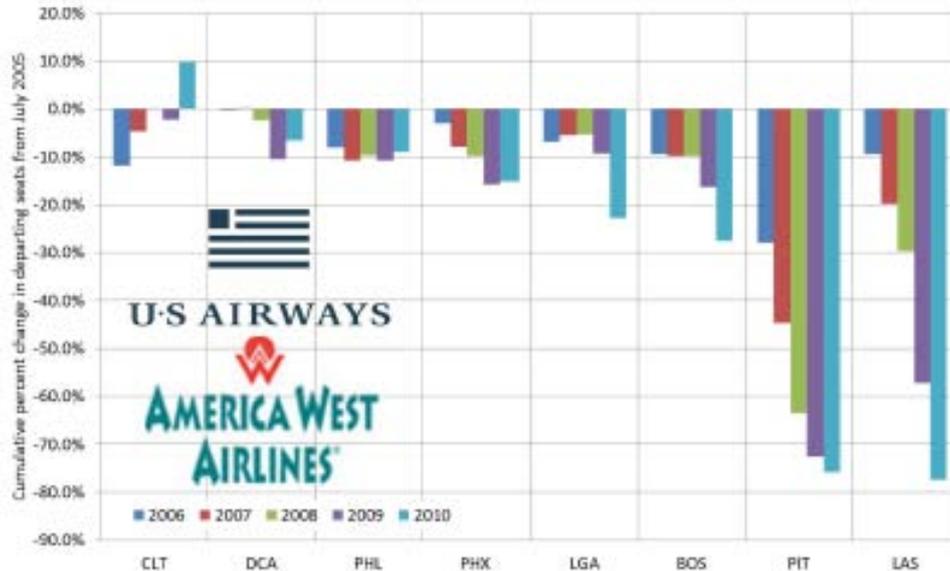


圖 3-32 全美航空與美西航空相關樞紐機場座位供給趨勢變化

另外，如圖 3-33 顯示達美航空 (Delta) 與西北航空合併，亦造成田納西機場 (MEM)、辛辛那提機場 (CVG) 等 2 機場明顯退，而紐約瓜拉迪亞機場 (LGA) 大幅成長。

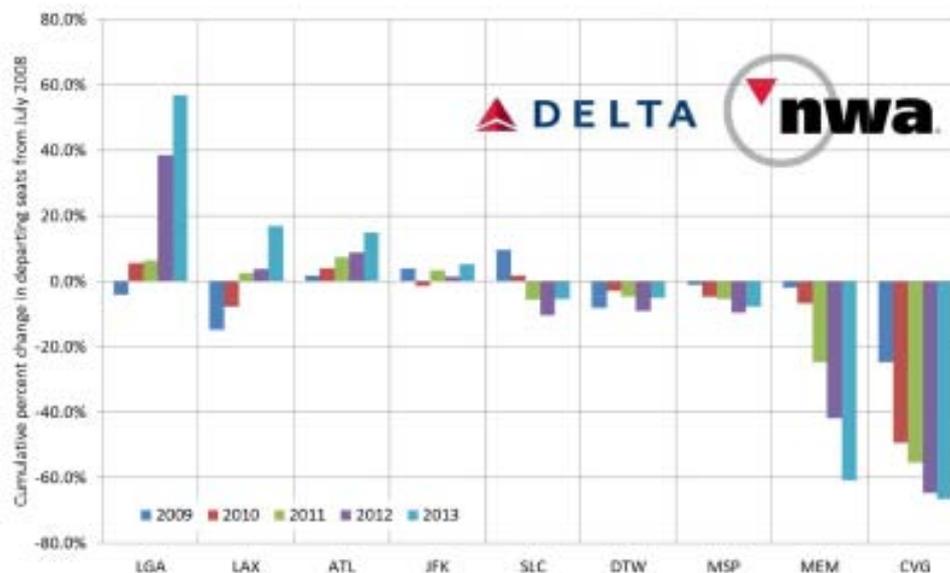


圖 3-33 達美航空與西北航空相關樞紐機場座位供給趨勢

### 3.1.4 航港運輸領域

#### 1. 利用 AIS 資料安全距離估計 SSD

由於國際海運貿易持續增加，造成大型貨櫃港區水域的交通密度高，例如新加坡。而水域中船舶運行的行為多變，船舶在進出港時必須保持一最小安全距離以避免碰撞。因此新加坡大學（NUS）發表一篇研究，蒐集新加坡水域內船舶自動辨識系統（AIS）資料，評估安全距離的機率分配。其中將船隻類型和規模、能見度（白天和夜晚）、及移動方向（交叉、正遇或追越）分為 36 個航行類別加以研究。圖 3-34 為船舶 SSD 示意圖，透過 AIS 統計後將船舶分為交叉、正遇及追越三種方向，以圖示模型呈現。

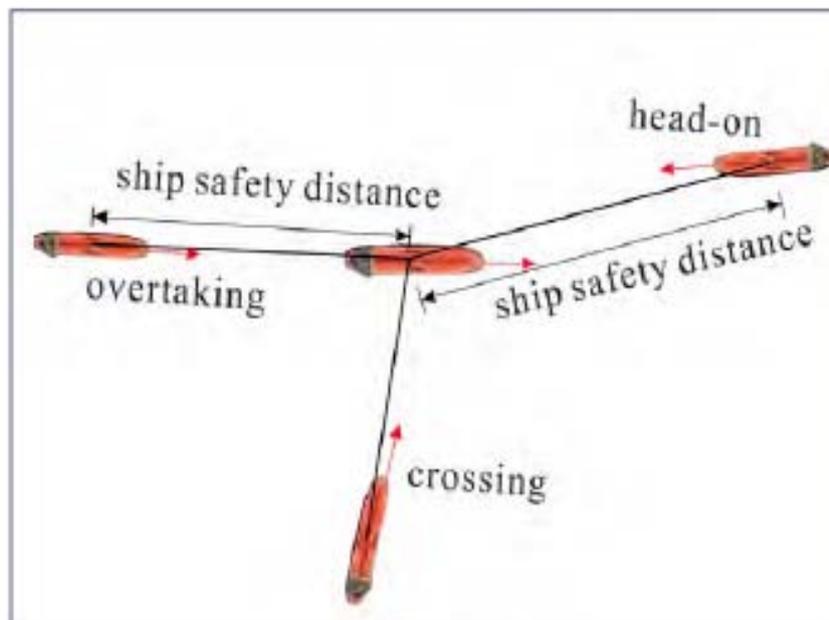


圖 3-34 船舶安全距離示意圖

研究結果顯示，對數常態分配及伽瑪分配為兩項適合對新加坡水域安全距離估計的方法，此外，安全距離分布可以反映運行行為，包含船舶運行交叉向的安全距離明顯受能見度狀態（日和夜）影響，以及船隻規模越大的安全距離亦越大。

#### 2. 船運管理

貨訊通（CargoSmart）提供全球船運管理軟體，幫助托運人、蒐貨人、物流供應商、無船公共運送人和航商優化供應鏈規劃並提高貨運準時率。貨訊通與全球 30 多家海運承攬人連結，利用大數據和雲端運算平臺為客戶提供船運管理方案，包含船期表、船運視野、合約管理、貿易法規和標竿解決方案等。

全球每月會發生超過五百萬次船期變動，舉凡天候不佳、港口罷工等都會影響到船期運行，為解決船期混亂及延誤的情形，貨訊通採用了 TIBCO 的技術，進行複雜事件處理、先進分析和數據視覺化，分析即時海上交通狀況，和多家船運公司航行情況，為客戶提供有用資訊、幫助作出更佳決策，進而優化期船運管理、提升效率及降低成本。

船期變動若無法掌握可能導致高額成本，包含訂艙不準確、滯期費增加或錯過貨櫃交付截止日等。為了提供客戶即時的船期狀態，貨訊通推出大船期（Big Schedules）在線搜尋平臺，如圖 3-35，搜尋頁面涵蓋最近七天船舶動態地圖、海運事件、天氣標記和即時新聞。大船期整合不同來源的數據進行分析，幫助客戶簡化搜尋流程，讓拖運人和物流供應商更快取得即時的船期數據，以及提供智能分析幫助其船運規劃。大船期智能功能包括：即時數據、視覺化、個人化搜尋及洞察力（Insights），在使用者蒐尋當下可以立即查出海運承攬人公布的船期和即時的船舶位置，以圖示方式提供快速、個人化的搜尋結果，幫助托運人和物流商做出更正確的供應鏈決策。

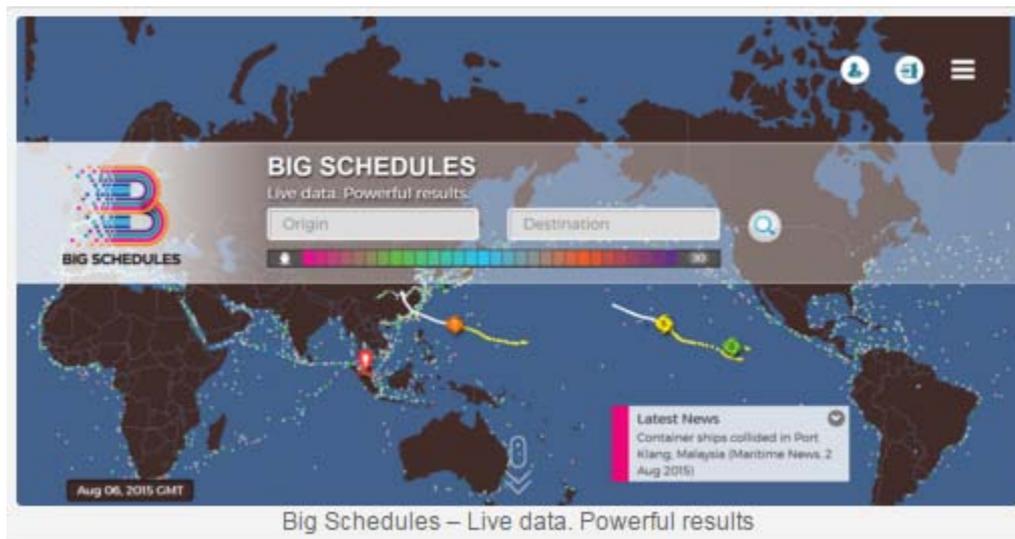


圖 3-35 大船期搜尋頁面

### 3. 聰明貨櫃

達飛海運集團（CMA CGM）乃法國貨櫃承運公司，總部位於馬賽，成立於 1978 年，現為世界上第三大貨櫃航運公司。全球百分之九十的商業產品經由海洋運輸各地，達飛海運除了載運各類別商品外，亦開發物流解決方案，在 2015 年啟動 TRAXENS，此投資開發創新貨櫃監理、地理定位和多方面合作系統，達飛海運透過 TRAXENS 將大數據導入運輸事業中，提供客戶全球追蹤貨物管道和即時資料的蒐集。

達飛海運 BOUGAINVILLE 輪，如圖 3-36 所示，貨輪上裝載配置 TRAXENS 技術的聰明貨櫃（Smart Containers），其利用內建的接蒐器能夠進行彼此間，及對船上基礎設施之通信。TRAXENS 系統透過高技術裝置監控貨櫃，廣泛地從每一個貨櫃蒐集資料，包含位置、溫度、濕度、振動、衝擊、企圖竊盜及通關狀態等，資料傳送到達飛集團後增加航線上資訊價值，並傳給客戶、保險公司及海關。



圖 3-36 TRAXENS 監控技術

經由 TRAXENS 技術，達飛海運把貨櫃變得智慧化、具連結性、且將複合運輸系統引領進入大數據時代。擁有每年一千兩百萬貨櫃量的運輸，大量資訊的蒐集與分析，也為終端消費者提供了服務的改善。

#### 4. 優化船舶操作節省能源

日本郵船(NYK)是日本具代表性的海運公司之一，與旗下 Monohakobi 技術研究所 (MTI) 及日本聯合造船廠 (JMU) 合作，針對一批 14000TEU 貨櫃船進行大數據研究。其中包含從 JMU 加入日本郵船的「NYK BLUE JAY」號，該船是第一艘加入此計劃的貨櫃船。

依計畫，日本郵船提供 JMU 船舶技術建議、蒐集與分析「NYK BLUE JAY」各種運作大數據，並將結果回饋給 JMU。此合作研究主要目的在於藉由分析船舶推進性能節省能源，通過檢測船體應力提高安全性，預防嚴重動力裝置事故。

日本郵船在船舶上安裝 MTI 研發的 SIMS，其資訊流程如圖 3-37 所示，船員和岸上操作人員可以即時交換數據，包含每小時精準的操作情形和燃料消耗狀況，對船速、燃料消耗和氣候等運行狀態數據的優化評估可以使船舶操作和配置效率更佳。

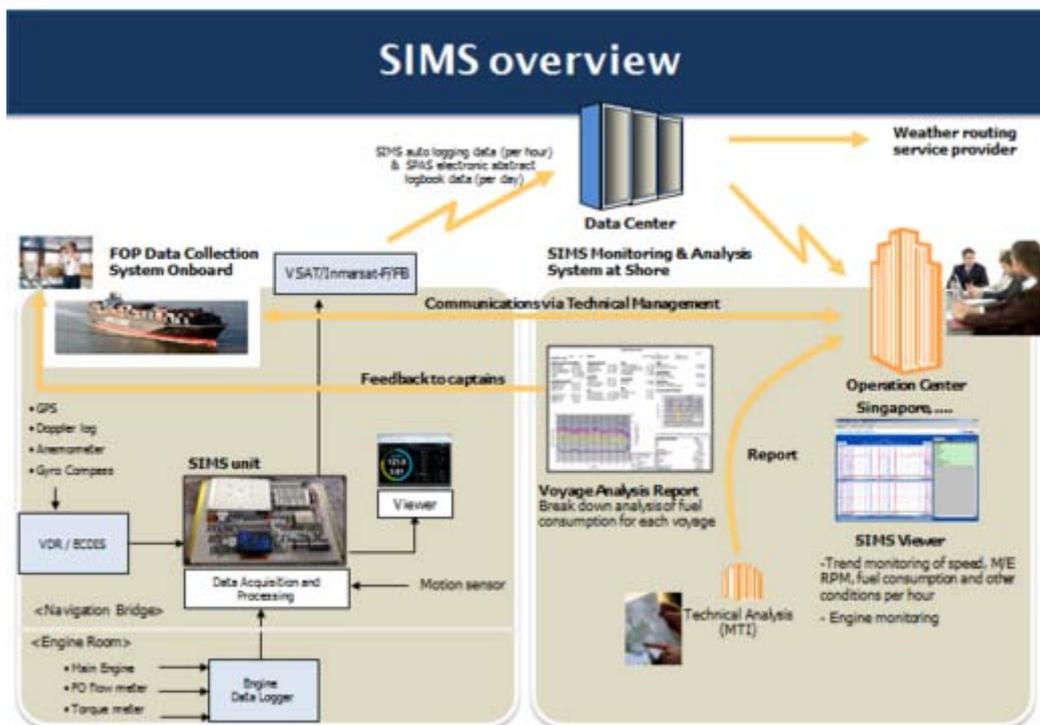


圖 3-37 SIMS 概觀

### 3.1.5 小結

綜整上述案例回顧，可知交通大數據的應用目的不外乎是蒐集需求、便民資訊、提升安全、資產管理、延誤改善、輿情分析、節能減碳、追蹤動態以及效能提升等，但各運輸領域所著重的項目略有不同，如表 3.1。

表 3.1 國外運輸領域大數據應用目的彙整表

目的\領域	H公路運輸	R鐵路運輸	F航空運輸	P航港運輸
1 蒐集需求	●	○	○	○
2 便民資訊	●	●	◎	○
3 提升安全	◎	●	●	●
4 資產管理	◎	●	●	◎
5 延誤改善	◎	●	●	○
6 輿情分析	◎	●	◎	○
7 節能減碳	◎	◎	◎	●
8 追蹤動態	◎	◎	◎	●
9 績效提昇	●	◎	●	●

應用重點：●主要 | ◎次之 | ○較低

資料來源：本研究整理

在各項應用目的中又以安全與績效的提升是多數運輸領域的應用重點，然而除了這些應用目的，從案例中可發覺資料視覺化在大數據應用中佔有重要的地位，在資訊發布上，適當的視覺化呈現可讓民眾快速獲得所需資訊；在資料分析上，藉由資料視覺化工具可讓專家容易掌握資料特性，並做進一步分析。

透過國外案例的回顧，可作為國內發展交通大數據應用之借鏡，但通常無法直接採用，除了國情不同所關注議題可能不同外，資料來源也是主重因素之一，例如新加坡的公車即時乘載率，國內公車並無配備相關感測器來偵測與提供車上乘載狀況。因此，在本研究提出國內交通大數據推動發展建議之前，必須要了解國內交通機關的需求與相關資料集狀況。

### 3.2 交通機關需求分析

在回顧國外交通大數據案例後，本研究藉由訪談讓交通部所屬相關單位了解其業管資料可能衍生之應用，並蒐集各單位在交通大數據的發展上所關注的課題與需求，以及相關資料集的狀況。如圖 3-38 所示，本研究前後共訪談了九個單位多次，包含桃園機場公司、高鐵局、臺鐵局、民航局、路政司、公路總局、基隆港務分公司、航政司以及航港局，每場訪談皆由該機關的長官與同仁共同與會，本節則以彙整的方式，說明公路、鐵路、航空、航港等四個領域的訪談結果。



圖 3-38 訪談進程一覽(以首次訪談為序)

### 3.2.1 公路運輸領域

#### 3.2.1.1 關注議題與需求

##### 1. 找尋潛在旅運需求

為提升公共運輸使用率，除了提供符合需求的公路客運路線，掌握潛在需求開闢新路線也是重要工作之一。若能從 eTag 資料或行動電信資料中，藉由大數據分析技術探討民眾旅運行為，則可有助於找尋潛在運輸需求，而進一步規劃提供公共運輸路線。

##### 2. 提昇公車動態資訊品質

目前過內的公車動態資訊系統相當普及，但仍有改善之處，例如山區路線通常只提供始發站的發車時間或頻率，而難以預估其到達時間，對於回程旅客不易掌握搭車時間，即便僅能根據歷史資料提供粗估資訊，對旅客來說已經非常有幫助，可增加民眾搭乘公共運輸之意願。

##### 3. 駕駛員風險管理

公共運輸駕駛員的駕駛行為除了影響周邊交通，更直接攸關車上旅客安全，透過大數據技術分析駕駛員特質、交通違規與肇事紀錄之關聯，可對風險較高之駕駛員施加教育訓練等措施，以達到預防事故發生之效果。

##### 4. 掌握油耗水準

近年來氣候變遷與全球暖化是世界各國最關注的議題之一，而公共運輸發展則是減少碳排放量的重要策略之一，除了鼓勵民眾使用公共運輸外，如果能掌握公共運輸車輛之油耗情況並加以改善，則可進一步達到節能減碳之目標，亦能減少客運業者營運成本。

#### 3.2.1.2 相關資料集狀況

##### 1. eTag 資料

eTag 為一種被動無線射頻辨識 (RFID) 的標籤貼紙，由於國內高速公路電子蒐費 (ETC) 採用 RFID 系統，許多車輛大多都裝設了 eTag。當車輛通過感應門架時，系統可獲取該車輛之車牌號碼、過站時間與地點等資訊，因此可根據車輛的起迄地點來推估民眾的旅運需求，但由於此資料含有個人資訊，屬於個人機敏資料而難以取得。

## 2. 行動電信資料

有別於 eTag 資料僅能分析車輛的旅運起迄點，行動電信資料則可以用來分析個人的旅運起迄點，而且隨著行動電話的普及，幾乎可用來推估全國人民的旅運行為。目前國內電信業者已可根據基地臺資料進行區域人潮統計之分析，如圖 3-39 和圖 3-40。

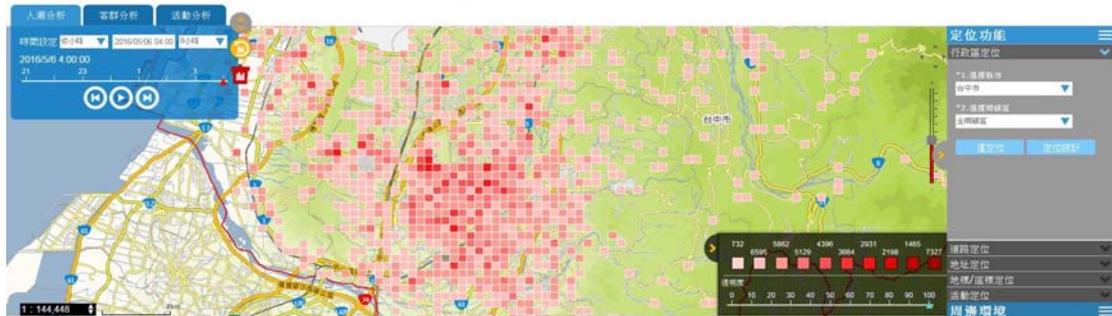


圖 3-39 中華電信區域人潮客群分析系統



圖 3-40 中華電信人潮統計表

## 3. 公車動態資料庫

公車動態資料隨客運類型不同，資料保存單位亦不相同，可分為公路客運資料與市區公車資料，公路客運資料主要由公路總局的聰明公車系統提供，而市區公車則從各縣市政府的公車動態資訊系統取得。

兩種客運類型的公車動態資料都包含定時回傳資料與定點回傳資料，定時資料係車輛於每分鐘回傳三次資料，其主要內容包含業者代碼、車牌號碼、路線代碼、去返程、經緯度座標、車速、GPS 時間等；定點資料則是車輛於各站牌進站離站時所回傳的資料，主要內容包含業者代碼、車牌號碼、路線代碼、去返程、站牌代碼、進站離站、GPS 時間等。

#### 4. 公路監理資料

隨著第三代公路監理資訊系統(M3)的啟用,除了可提供民眾更便利的監理相關業務服務之外,並透過人車歸戶系統整合車、駕籍資訊,同時,在路檢聯稽攔查業務執行上,亦可利用行動裝置針對違規個案立即開單入案。

根據「各機關應用公路監理資訊連結作業及管理要點」,目前 M3 資料庫可提供車籍基本資料、駕籍基本資料、車輛違規明細資料、運輸業駕駛員及公司資料,與運輸業車輛資料等五大類資料供外界做大數據分析應用,但因其資料涉及個人資訊,須審核通過後才能取得。

#### 5. 交通事故資料

雖然交通事故與交通安全息息相關,但交通事故資料非交通部所有,而是歸屬於內政部警政署。警政署將交通事故資料分為人員當場或 24 小時內死亡之交通事故(A1),以及人員受傷或超過 24 時死亡之交通事故(A2)等兩類,由於涉及個人機敏資料,目前在內政部資料開放平臺上(如圖 3-41)僅揭露事故發生時間、發生地點、死傷人數和車種等四個欄位,且發生地點之內容為一般文字敘述,而非地理座標資料,因此此開放之交通事故資料不易與駕駛人、肇事車輛或地理資訊系統整合作進一步大數據應用。

事故類別: A1 類	發生時間	發生地點	死亡受傷人數	車種
	103年01月01日 04時30分	彰化縣伸港鄉彰化縣伸港鄉新港村新港路56號前	死亡1;受傷0	普通重型-機車
	103年01月01日 06時00分	彰化縣彰化市臺1線188公里100公尺處南向內側車道彰化縣彰化市國聖里中山路3段554巷口	死亡1;受傷0	腳踏自行車-慢車;自用-小客車
	103年01月01日 06時00分	嘉義市西區興業西路4 8 5 號前	死亡1;受傷0	營業用-大貨車;腳踏自行車-慢車
	103年01月01日 06時30分	花蓮縣秀林鄉和仁村(和仁明隧道內南下車道)(附近)台9線164公里700公尺處南向向外側車道	死亡1;受傷0	普通重型-機車

圖 3-41 內政資料開放平臺的交通事故資料

#### 6. 公共運輸車輛油耗或加油紀錄

為了掌握公共運輸車輛之油耗情況，需要車輛的油耗或加油紀錄，一般屬於客運業者的營運資料，但通常是每輛車一年分的總油耗量，或平均每公里油耗量這類的統計資料，因此在大數據分析應用上，能再挖掘的資訊不多，若能精確到每輛車每天每一趟服務，甚至是隨時隨地的監控油耗狀況，則能夠找出油耗特性與影響因素，業者也能對症下藥來降低油料支出成本。

#### 7. 電子票證資料

電子票證是用來掌握旅客起迄相當重要的資料，在本研究的標竿案例中亦會嘗試引用，相關的資料格式或應用方向可參考第四章。

### 3.2.2 鐵路運輸領域

#### 3.2.2.1 關注議題與需求

##### 1. 掌握連續假期旅運需求

每逢連續假期，臺鐵總是肩負運送民眾返鄉或旅遊的任務，然而受到連續假期長度與性質的不同，其旅運需求不易掌握，例如春節假期每天的需求都不同，而且六天假或九天假的每日需求變化也迥異；又例如中秋節民眾可能返鄉為主，但國慶日可能就以旅遊為主，不同的旅次目的影響不同 OD 的需求量。因此希望可從過去歷史票務資料與民眾購票情況，找出連續假期的旅運需求特性，以擬定更符合需求的行車計畫或自由座配置。

##### 2. 資料視覺化呈現

在鐵路監理業務方面，高鐵局雖然僅能就高鐵公司提供之資料進行趨勢分析，但若利用大數據的視覺化 (Visualization) 技術，將有助於以更直覺的方式瞭解資料背後的意義，並可能發掘過去透過傳統報表無法察覺的潛在問題。而對臺鐵局來說，即便只是將現有基礎資料做視覺化呈現，對於營運管理與決策輔助上也非常有用。

##### 3. 應用物聯網提供旅客精緻化服務

大數據應該要配合物聯網的應用方能提供政府與民眾更多便利，例如在通勤列車的各車廂中裝設感測器，偵測車內乘載狀況並即時發布給旅客，以供其搭乘時選擇較舒適的車廂，不僅提高服務品質，亦可降低乘載變異。

#### 4. 提升平交道安全

平交道碰撞事故長年以來都是臺鐵局在安全上最大的威脅，而約有九成的平交道事故是由公路端所造成的，除了持續增加平交道的防護措施之外，臺鐵局亦希望能結合其行車安全資料庫與公路端相關資料，發展具體的大數據應用來協助降低事故的發生。

#### 5. 以預防性維修取代矯正性維修

臺鐵局近年已計畫在不久的將來建置維修資訊管理系統（Maintenance Management Information System, MMIS），來協助臺鐵人員維護設備，同時透過該系統所記錄的設備維修歷史資料，結合生命週期的概念，將可達到預防性維護之目的。而臺灣高鐵已有維修管理資訊系統，但有關維修週期上，雖系統供應商與零件商會提供預防維修的週期建議，但仍需營運經驗歸納，例如高鐵道岔的維修週期就與原廠建議有所差異。

#### 6. 改善列車誤點情況

列車誤點一直是臺鐵持續改善但旅客仍不滿意的項目，主要因為誤點通常是由多重原因所造成，所以不易釐清癥結所在，而難以對症下藥加以改善，或許能透過大數據分析來協助找出列車誤點的根本原因。

#### 7. 供給需求匹配

根據預測的需求量提供適當的供給量，可避免供過於求造成運輸資源浪費，或供給過少造成一票難求、乘載服務水準過低的現象。在需求預測的方法上，除了利用時間序列做短期預測外，亦可根據車票預售情況進行推估。而在供給量方面，可從過去銷售狀況檢討配位，甚至隨預售情況動態調整以符合需求。

### 3.2.2.2 相關資料集狀況

#### 1. 列車時刻表

臺鐵與高鐵皆有公告列車時刻表供民眾搭乘時參考，其主要內容包含每天各車次的資料、停靠車站順序以及到站離站時間，除了可於營運單位官方網站下載之外，亦可從交通部「公共運輸整合資訊流通服務平臺」透過應用程式介面（Application Programming Interface, API）取得（如圖 3-42）。然而目前列車時刻表的公開內容並沒有包含不停靠之車站的通過時間，因

此部分大數據應用課題，例如列車服務可靠度，可能必須以內插的方式補足通過車站的時間，方可進行後續分析。

The screenshot shows the 'Public Transport Integration Information Exchange Platform' (公共運輸整合資訊流通服務平台) interface. The main content is a table listing various train services and their timetables. The table has columns for Service Name, Vehicle Type, Data Type, Data Provider, Citation Count, Description, and URL. The data is as follows:

服務名稱	運具類型	資料類型中文	資料提供單位	引用次數	說明	網址
臺鐵每日車次資料(近60天)	臺鐵	時刻表	交通部臺灣鐵路管理局	37,479		
高鐵每日車次資料(近28天)	高鐵	時刻表	台灣高速鐵路股份有限公司	37,465		
臺鐵每日時刻表資料(近60天)	臺鐵	時刻表	交通部臺灣鐵路管理局	37,610		
高鐵每日時刻表資料(近28天)	高鐵	時刻表	台灣高速鐵路股份有限公司	37,793		

圖 3-42 公共運輸整合資訊流通服務平臺上的臺鐵與高鐵時刻表

## 2. 票務系統資料

臺鐵的票證資料共有五類：對號票、非對號票、自動售票機、定期票、電子票證，每種票證在資料庫中所記載的欄位，除了乘車起迄車站外，則視票證的特性而各有不同，例如對號車含有車次欄位、定期票則含有刷進與刷出時間，值得注意的是電子票證僅有紀錄刷出時間而無刷進時間，相對之下，高鐵的票證資料則單純許多。

## 3. 行車保安資料庫

為了因應複雜之列車運轉環境，並擺脫以人工登錄、查詢行車事故之模式，臺鐵局建置了行車保安資訊系統，透過電腦系統進行統計分析並產生量化風險值，以風險管理的概念來確保鐵路行車安全。行車保安系統的資料庫紀錄了自民國 81 年以來的各項行車事故，內容包含事故發生的概況與原因，以及相關的人、事、時、地、物等細節，雖然此系統目前僅供臺鐵局內部使用，但未來對於鐵路安全方面之大數據應用極具潛力。

而高鐵的事件事故紀錄資料較為敏感，並非屬於 Open Data 的範疇，就監理的角度而言，高鐵局僅能依法監理，而無法源可要求高鐵公司提送全部原始資料，欲蒐集資料可嘗試以研究性質計畫專用的方式向高鐵公司申請局部期間內的資料。

#### 4. 維修管理資訊系統紀錄

高鐵的維修管理資訊系統係運用電腦化管理制度，以控管高鐵機電及軌道系統之維修作業，達成降低營運維修成本並提升效率之目標，而臺鐵路則正在推動維修管理系統。

維修管理系統不僅記錄著鐵路系統所有設備，如基地車站建物、機電系統、軌道、號誌、電力及通訊等的基本資料與維修歷程，通常也提供各項管理功能，如資產管理、工單管理、庫存管理、採購管理、維修排程等，甚至是預防性維修。

#### 5. 列車實際運行情況

中央行車控制系統（Central Traffic Control System, CTC）之主要任務在於監控系統全線列車運轉狀況，該系統中記錄著每天各車次於各車站的實際到達、離開或通過時間，是掌握列車誤點狀況的重要資料。

#### 6. 誤點原因紀錄

唯有掌握造成列車延誤的原因，才能有效提出改善方案，目前臺鐵路列車誤點紀錄資料中，將延誤原因分為七大類，共 45 項，但此資料並非在列車延誤的當下即時記錄，而是事後才由人員透過系統填報，因此該紀錄可能與實際有落差，此外誤點原因的分類是否夠周延，也影響著此資料在大數據分析應用上的品質。

### 3.2.3 航空運輸領域

#### 3.2.3.1 關注議題與需求

##### 1. 提升聯外運輸服務品質

地面聯外運輸為機場服務重要的一環，但因入境航班旅客量不易精準掌握，常發生突如其來的大量旅客造成客運車班跟計程車班不足的問題，為了確保聯外運輸系統的服務品質，或許可根據航班與旅客資料透過大數據分析預估並發布短期旅客量變化，讓聯外運輸能即時因應。

##### 2. 飛機和機場設備的妥善率與預防性維護

近年來預防性設施維護已逐漸成為維持設施正常運作的主要手段，如今現代化飛機上均有完整的電子偵測設備，能自行做健康管理，產生預防性維修的建議報表供航空公司參考。而在機場方面，將來可在建築資訊模

型 (Building Information Modeling, BIM) 的基礎上, 結合大數據視覺化技術來呈現設施的使用率或維修汰換次數等資訊, 協助人員發現高負荷的設備, 並事先做預防性準備。

### 3. 天候對於班機延誤或取消之關聯

天氣狀況對於班機延誤或取消之影響為民航局關注的課題之一, 其中班機延誤有兩種定義: (1)以飛機於跑道起飛或降落著地的時間來評估準點, (2)以飛機開關艙門的時間來評估準點, 兩種延誤定義所探討的議題有所不同, 但均有其價值與意義。此外, 若要進一步擬定班機延誤改善方針, 除了天候資訊, 還必須結合延誤當下的其他背景資訊, 如資源運用和設施調度的情況等, 透過大數據分析技術找出確切原因, 甚至是預防、預測等進階運用。

### 4. 航站人流管理

當前桃園機場最關心的議題之一就是人潮壅塞問題, 從旅客來到機場至搭上飛機, 或下飛機至離開機場, 過程中不只是班機的準點情況, 機場中各項空間設施的安排運用都影響著旅客的感受, 如服務櫃檯、行李系統、登機門、安檢、通關線、候機室等, 若能將其使用狀況數據化, 未來可以利用大數據做檢討和改善旅客等候時間, 甚至是預警讓機場能提前準備因應措施。

### 5. 機場設施使用績效

掌握機場各項設施的使用效率以及實際使用情況, 機場公司才能更有效地安排運用這些設施。例如了解報到櫃檯的服務速度以及航空公司實際使用櫃檯的狀況, 將有助於報到櫃檯的安排, 以避免部分櫃檯太過壅擠, 進而影響出境大廳旅客動線。又例如從登機門過去的使用情況, 可研究不同機型或旅客量等各項因素對登機門使用效率的影響, 可更精準分配與調度登機門。

## 3.2.3.2 相關資料集狀況

### 1. 航班資料

航班資料可分為預先公告與即時公告兩類, 預先公告的航班資料為定期航班的預訂班表, 內容包含各班機的所屬航空公司、班機編號、目的地、預定地起飛時間與抵達時間等; 而即時公告的航班資料為即時出入境航班

資訊，內容則含有即將出入境班機的所屬航空公司、班機編號、機型、起點或目的地機場、所使用的登機門或報到櫃臺、運行狀態等。其中若能定期蒐集並儲存即時出入境航班資訊，則有助於班機可靠度或空側運作績效方面的大數據分析。此外，有關各航班上的旅客相關資訊，機場公司僅能掌握其數量，而旅客的特性如性別、是否為常客等詳細資訊則是在航空公司手中。

## 2. 機場設施使用紀錄

從機場各項設施的使用紀錄中，可分析出其使用效率甚至可以了解人潮流動狀況，然而機場公司在此部分能掌握的資訊不多，其中許多設備的使用紀錄是在系統設備商手上，欲取得通常須以高價金額向其購買。而報到櫃臺的服務人數與服務效率則是在航空公司手中，目前桃園機場正在與航空公司協調取得。至於旅客進、出關的資料則在移民署，由於該資料有機敏性，取得上有其困難度。

## 3. 氣候資料

國內中央氣象局配合行政院開放資料政策，建置「氣象資料開放平臺」，如圖 3-43，開放氣象、地震、海象等資料供外界加值運用，共分為預報、觀測、地震海嘯、氣候及天氣警特報等 5 大類資料集。

The screenshot shows the '氣象資料開放平臺' (Weather Data Open Platform) website. The page title is '氣象資料開放平臺 OPEN DATA'. The URL is 'opendata.cwb.gov.tw/datalist'. The page content includes a navigation menu and a list of data items under the heading '資料清單' (Data List). The list includes various weather forecast and observation data items with their corresponding IDs.

中央氣象局開放資料平臺提供的資料包括下列項目與對應的資料編號	
<b>預報</b>	
一週農業氣象預報-未來天氣概況、各地天氣預報及農事建議	F-A0010-001
海面天氣預報-海面天氣預報	F-A0012-001
海面天氣預報-海面天氣預報(英文版)	F-A0012-002
長期天氣預報-月長期天氣展望	F-A0013-001
長期天氣預報-季長期天氣展望	F-A0013-002
波浪預報模式資料-台灣海域預報資料	F-A0020-001
潮汐預報-未來1個月潮汐預報	F-A0021-001
滿潮預報圖-今日滿潮預報影像圖	F-A0022-001
滿潮預報圖-明日滿潮預報影像圖	F-A0022-002
滿潮預報圖-後天滿潮預報影像圖	F-A0022-003
滿潮預報圖-第4天滿潮預報影像圖	F-A0022-004

圖 3-43 氣象資料開放平臺

### 3.2.4 航港運輸領域

#### 3.2.4.1 關注議題與需求

##### 1. 整合海關、港棧、貿易資料分析客貨運趨勢

為健全港埠未來的長期發展，建議港務局的航港單一窗口服務平臺（MTNet）能整合海關、港棧、貿易等資料庫，甚至是鄰近港口的轉口貨物狀況或世界各國社經條件，透過大數據分析深入探討國際航運的發展脈動趨勢，作為港埠發展決策參考。

##### 2. 港區週邊交通/重車追蹤管理

進出港區的大型貨櫃車容易影響附近交通狀況，透過貨櫃車進出港區的門哨紀錄，有助於擬定相關措施來減少貨櫃車對港區周邊交通的影響，以及透過重車追蹤管理來避免載貨車輛行經超過道路限重之路段。

##### 3. 臺灣周邊海域船隻防災減災分析

從過去的海事資料中可分析臺灣海域周邊容易發生事故的海域，甚是結合海氣象、船舶航跡等其他資料探討不同因素與海事的關聯，以便能做到海難的預警、減災或其他更積極的作為。

##### 4. 船舶港內碼頭運作效率與擁擠密度

強化營運效能是港務公司的營運方針之一，根據船舶於港內的運作、裝卸貨狀況等資料，能以大數據技術分析碼頭運作效率、運轉速度、密度擁擠度等，最後在地理資訊系統上展現分析結果，有助於提升港口作業績效。

#### 3.2.4.2 相關資料集狀況

##### 1. 航港單一窗口服務平臺資料

航港單一窗口服務平臺（MTNet）為國內航港資訊系統之單一入口平臺，如圖 3-44，除了整合國內航政監理作業、港灣棧埠作業、港口國管制作業等各項服務，也提供多元化的航港相關資訊，其資料庫包含進出港預報資料、國內外航運業、航線資料、各式船舶資料、海事報告等，同時也因作業需要，介接了衛福部疾管署、經濟部商業司、農委會漁業署、財政部關務署、內政部移民署等機關的資料。



圖 3-44 航港單一窗口服務平臺

## 2. 海事資料

航港局的海事案件記錄內容包含涉入之船舶名稱、事件發生時間、肇事地點、海事原因、人員傷亡、船體損害等資訊。其中事件發生時間僅有日期而無時分，未來較難作更細緻的分析，而肇事地點的描述方式不一，雖大多是已用經緯度標示，但仍有不少是用簡單文字敘述，因此在與地理資訊系統整合分析上較為麻煩。

## 3. 自動識別系統資料

自動識別系統 (Automatic Identification System, AIS) 是安裝在船舶上的設備，會自動廣播船舶識別碼 (含水上行動通訊識別碼、船名、無線電呼號、IMO 編號)、航儀動態 (GPS 船位經緯度、對地航向航速、艏向、轉向速率)、船舶靜態與航程資料 (船舶與貨載種類、天線位置、船舶長度、目前最大靜態吃水、定位系統、目的港、預計抵達時間) 等資訊，並顯示在 AIS 的螢幕或電子海圖上，可協助船舶人員及海事主管單位追蹤與監視船舶動向。

## 4. 港內船舶動態資料

港務公司對於船舶在港內的運作紀錄有較完整的資料，除了進出港時間，也包含引領、帶解纜工作、港勤船舶作業、移泊狀況與派工情況等資

料，幾乎可掌握船舶於港內的各項動態，惟獨貨物裝卸的掌握度較低，必須參考起重機的使用紀錄。

### 3.2.5 綜合整理

將上述各項關注議題與需求對應至表 3.1 之交通大數據的應用目的，如表 3.2，可發現主要集中在蒐集需求、便民資訊、提升安全與績效提升方面，不過雖然目的相同，各運輸領域所著重的內容則有所不同，例如在蒐集需求上，公路運輸主要在找尋潛在運輸需求，鐵路則是旅運需求預測，而航港是以貨物為主而非旅客。

表 3.2 國內交通機關大數據應用需求彙整表

目的\領域	H公路運輸	R鐵路運輸	F航空運輸	P航港運輸
1蒐集需求	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
2便民資訊	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
3提升安全	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>		<input checked="" type="checkbox"/>
4資產管理		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
5延誤改善		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	
6輿情分析				
7節能減碳	<input checked="" type="checkbox"/>			
8追蹤動態				
9績效提昇		<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>

在綜整各國內交通機關對大數據應用的需求後，本研究於民國 105 年 5 月 25 日召開第一次座談會，除交通部所屬單位外，亦邀請六都交通局與專家學者出席，共同交流討論交通大數據之相關課題；亦於 105 年 9 月 29 日及 10 月 7 日分別在臺灣鐵路管理局、公路總局共辦理兩場座談會彙整相關單位之意見。

### 3.3 資料匯流協調機制建議

完整且高品質的資料集是發展交通大數據的重要關鍵，本節將彙整各單位深度訪談過程中發現之資料匯流問題，並據以設計相關的匯流協調機制供未來推動時參考。

### 3.3.1 現況問題描述

目前交通相關單位資料匯流相關問題計有七類，分述如下：

#### 1. 資料分析者與擁有者態度落差

許多資料擁有者為第一線營運單位，必須直接面對營運壓力，業務繁忙以致於無暇釐清本身擁有的資料及其可能的應用範圍，更遑論執行大數據分析應用相關研究，甚至連擷取資料提供資料分析者都成為額外負擔；但資料分析者則專注於資料科學，積極渴望取得資料並從中提煉出有價資訊，兩者態度立場不同且無互利機制，間接造成資料流通與取得困難重重。

#### 2. 資料需求者逐專案重複向權責單位介接資料

承上，資料擁有者或權責單位常須面對眾多資料分析者（通常搭配專案）要求提供資料，造成許多重複作業，包括多位需求者要求同種資料集或同一需求者在逐年期申請同一資料集等，再再都造成資料擁有權責單位不小負擔。

#### 3. 資料擁有者提供資料時缺乏充分授權

國內大數據方興未艾，許多規定尚不明確，故提供資料過程的相關權利義務也尚未明確，特別是個資或機敏資料的處理更需明確授權與規範才能讓一線人員安心提供，包括機敏資料範疇、去識別化處理方式（遮罩/編碼）等均無明確規範，間接導致資料擁有者會以相對保守的方式面對分析者提出的需求。

#### 4. 資料擁有者擔心資料輾轉外流

除機密敏感資料外，資料擁有者亦擔心資料輾轉外流，特別是被民代斷章取義於議事廳披露，造成不必要困擾，降低資料匯流意願，因此若有資料匯流平臺或機制可同時考量資料整合、介接、流通、去識別規範及資料安全性等議題，將可有助於大數據資料流通應用。

#### 5. 分析者無法提出明確資料欄位需求

資料分析者並非原始資料的管理者，因此無法瞭解資料集全貌，難以提出確切的欄位，但資料擁有者基於前述幾項因素，通常不便或不願提供資料欄位清單，其中部分原因是資料分析者在未分析前，往往會有「完整取得後再評估用途」心態，造成資料提供者提供的資料集遠大於實際所需，

但這也包含資料分析者必須經過探勘使用方知欄位價值等因素，因此雙方溝通時常有落差，若能有制度或平臺居中協調則可大幅降低差距，讓資料擁有者願意提供清單，而資料分析者申請時也應充分說明需求與理由，並在統一的資料匯流平臺作業。

#### 6. 部分資料集由交通部外之行政機關管轄

欲產生高價值之大數據應用成果往往須依賴不同源之資料集融合，而這些資料往往已經跨越單一部會管轄，例如交通肇事資料係由內政部警政署管轄，部分運輸系統由民間單位營運...等，雖然亦可透過政府 Open Data 平臺取得相關資料，但其資料完整性或詳細程度可能不足以讓大數據應用產生高價值，故必須以交通部層級高度與其他部會用互惠方式交流資料集，或者以主管機關或監理角色介入協調方能事半功倍。

#### 7. 資料匯流其他問題

除上述六點外，長期而言資料匯流尚有以下議題須探討或處理：

- (1) 資料擁有者現階段亦未釐清或認知本身擁有何種資料及其應用之範圍。
- (2) 資料匯流平臺亦需考量整合各機關資料時，資料其格式、流通及去識別規範、介接方式以及資料安全性。
- (3) Open Data 平臺之資料完整度以及後續維護問題。

### 3.3.2 協調機制建議

考量 3.3.1 節所述之現況，由於大數據分析所需資料集隨著時空環境需求不斷變化，加上物聯網 (IoT) 時代即將來臨，未來將持續有新創資料集問世，到底哪些資料有分析需求與分析價值？哪些資料應該以交通部資源統一介接匯流？均是需要持續動態決策的課題。

有鑑於此，除初期資料匯流建議外，更需要一個制度性的運作機制方能讓達成目標，本研究建議可在交通部指導下成立「交通大數據諮詢評議委員會」來執行協調作業，同時建構資訊平臺來奠定國內交通大數據分析基石，以下將說明此委員會之成立目的與職責、組織架構、運作方式及相配套的資訊平臺等建議供參，唯未來成立時應針對實際需求及時空環境做適當調整。

### 3.3.2.1 目的與職責

交通大數據諮詢評議委員會係由交通部成立，其成立目的在於定期對國內大數據資料匯流與大數據執行計畫進行通盤檢視，用以評估各種資料集是否需要由交通部內統一做資料蒐集匯流與發布，同時肩負協調大數據相關計畫之預算與執行標的，提昇橫向聯繫，促使部內資源發揮綜效，同時能與他部會資料匯流的統一窗口。主要的職責如下：

1. 協調優先納入匯流之資料集或優先推動之大數據相關計畫供部屬機關遵循執行。
2. 審核部內各機關提出之資料匯流需求暨大數據相關建置計畫。
3. 代表交通部作為跨部會資料匯流之行政窗口。

### 3.3.2.2 組織架構

諮詢評議委員會建議設主任委員/副主任委員各一名，負責召開會議，主委與副主委可由次長/主秘擔任或由諮詢評議委員會委員互選之。諮詢評議委員會則分為三大類，第一類為機關代表，包含交通部內單位（含部屬機關）及地方政府交通相關局處，第二類則為專家代表，第三類則為民營公司代表如圖 3-45。

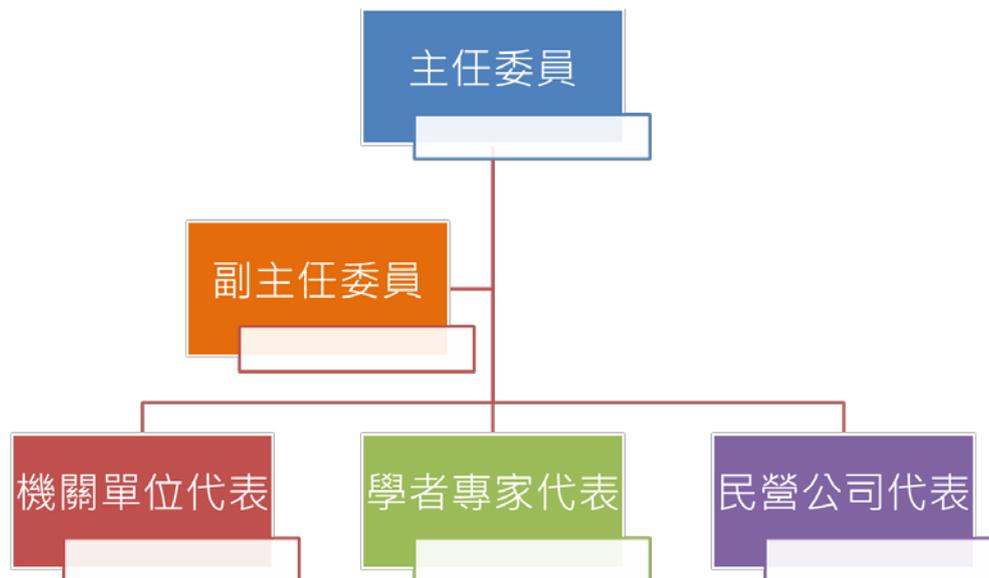


圖 3-45 諮詢評議委員會之組織架構圖

機關代表之數量應由交通部視需求擬定機關清單，初步評估部內單位可能包含路政司、航政司、郵電司、科技顧問室、管理資訊中心等五單位，部屬機關則可能包括臺灣區國道高速公路局、民用航空局、臺灣區國道新建工程局、觀光局、公路總局、中央氣象局、鐵路改建工程局、航港局、

高速鐵路工程局、運輸研究所、中華郵政股份有限公司、臺灣鐵路管理局、桃園國際機場股份有限公司、臺灣港務股份有限公司等。凡列入交通部擬定清單之機關，即可由機關推薦代表委員乙位，由交通部聘任。為利會議順利召開，機關委員得設第 1 順位及第 2 順位代理人，其職權與委員相同。機關委員無任期限限制，但機關得依需要更換。除了中央政府的機關代表外，地方政府機關也掌握許多有助於交通大數據分析的資料集，尤其是直轄市，因此建議可邀請六都交通/觀光局(處)加入委員會的組成中。

專家代表委員部分則可聘請交通運輸、資料科學、巨量資料、物聯網領域之專業學者數名，由交通部聘任，任期為一年但可續聘。

民營公司代表之設立係考量部分交通資料係由民營公司維運，典型例子如電子票證公司或電子蒐費公司，甚至部分運輸系統本身就由民營公司經營，典型例子如臺灣高速鐵路股份有限公司。

委員會成立初期可先尋求高公局、公路總局、高鐵局、臺鐵局、桃機公司、民航局、港務公司與航港局等和民眾生活相關之機關運作，隨著資料匯流需求擴大後再逐一邀請其他機關參與。

### 3.3.2.3 運作方式

本節將說明 5 項運作機制，前 3 項運作機制可決定資料匯流平臺蒐納新資料集的優先順序，同時讓資料匯流平臺的資料集持續擴充成長，需要透過諮詢評議委員會會議召開來驅動，而第 4 第 5 項則屬於資料匯流平臺的共享機制，建議可由資料匯流平臺的管理單位根據諮詢評議委員會制定的原則執行即可，毋需每項資料申請均由諮詢評議委員會決議。

#### 1. 匯流資料集新增機制 1：由上而下 (Top-Down) 目標導向機制

此運作方式係由主任委員或副主任委員召開會議，由專家學者以宏觀視野推薦高潛力、高價值之資料集作為優先納入匯流平臺之候選清單，但由於專家學者係以大數據上位政策願景目標角度提出方向建議，因此並無資料集的確切內容，故經諮詢評議委員會決定納入後，仍須成立工作小組與資料擁有單位商討匯流清單細節，包含各欄位適合匯流與否等，甚至若有關鍵重要的欄位在現有系統尚未蒐集保存，則必須編列計畫經費蒐集建置。此種由上而下目標導向的資料匯流機制建議一年至少召開乙次擷取學者專家意見敲定該年度目標資料集即可，後續則視目標推動的進展與需要召開後續的討論協調會議，但不會再討論新擴增的匯流資料集。

## 2. 匯流資料集新增機制 2：由下而上 (Bottom-Up) 需求導向機制

有別於前一項目標導向的資料匯流機制，需求導向的運作機制是開放國內各公家機關向諮詢評議委員會申請新增資料匯流集，暫時不開放大專院校、民營公司等非公家機關申請匯流機制，非公家單位若有執行公部門專案而衍生資料需求時必須由專案計畫委託單位（例如教育部或地方政府）提出申請，各機關申請時至少須包括：

- (1) 資料集簡稱暨原始資料維運單位
- (2) 資料集詳細內容（例如資料表清單、資料欄位型態...等）
- (3) 資料集應用計畫（例如專案計畫目的、可回饋資料集項目等）
- (4) 申請機關過去提供匯流平臺的新資料集或回饋資料集內容
- (5) 其他備註說明（例如跨域整合分析等其他有利於審核通過之說明）

由於此機制是需求導向之資料集新增機制，故會議召開的頻率理應比目標導向式更為頻繁，建議以每年至少兩次，但不超過四次（每季一次）為原則，召開會議時會討論當下已送件之申請，經會議討論通過的資料集將納入匯流清單，並與資料維運單位的委員代表協商後續介接匯流細節，包含技術窗口與預計完成時程等，後續視需要召開工作會議。

## 3. 匯流資料集新增機制 3：資料維運單位自主加入機制

除前述目標導向與需求導向機制之外，若資料擁有者（資料維運單位）認為過去重複面對各方資料分析者需求而疲於奔命，可在諮詢委員會主動提出欲加入匯流的資料集清單並提供聯繫窗口，經會議共識同意後由資料匯流平臺管理單位（例如交通部管理資訊中心或運輸研究所）與窗口聯繫完成介接，未來各界的資料分析需求者就直接與匯流平臺管理單位申請資料，原始資料維運單位不用重複回應各界需求。

## 4. 匯流資料共享機制 1：資料集清單與資料描述共享機制

為使匯流資料能被廣為運用，匯流平臺會有專區羅列目前所有資料集清單、資料描述、範例資料若干筆，這個專區將供公家單位內瀏覽瞭解匯流平臺資源，但此專區僅能下載範例資料，欲實際取得資料須另行申請方能抓取資料。

## 5. 匯流資料共享機制 2：資料實體內容申請機制

前項共享之目的在於讓各界瞭解匯流平臺所涵蓋的資料範圍，但如欲取得實體資料則必須另行申請，建議本平臺初始階段僅提供公家機關申請，非公務單位若基於執行政府採購之專案計畫欲使用資料時，必須透過發包單位申請，同時必須遵守資料匯流平臺所規範的資料保密要求。

申請時至少必須提供以下資料：

- (1) 欲申請之資料集名稱與欄位
- (2) 前項申請內容之理由或應用
- (3) 資料範圍(例如時間或空間,ex:2015年~2016年/臺中地區...等)
- (4) 本次申請預期回饋資料集內容
- (5) 申請機關過去提供匯流平臺的新資料集或回饋資料集內容

### 3.3.2.4 資訊平臺

除資料匯流與使用權等制度面協調機制外，亦需要資訊系統之匯流平臺作為資料實質儲存之平臺，考量目前交通部轄下各機關之職掌與專業，資料匯流平臺建議可由管理資訊中心規劃建置之。

## 3.4 優先推動領域規劃原則與策略

### 3.4.1 規劃原則

本研究延續去年中央地方交通首長提出的六大共識與三大優先應用領域(詳見第 1.1 節)，除了參考國外大數據應用案例(詳見第 3.1 節)外，也訪談各交通機關瞭解需求(詳見第 3.2 節)，作為優先推動領域之規劃依據如圖 3-46，規劃時計有五點原則說明如下：

#### 1. 以「解決問題之需求導向」為最高指導原則

大數據分析僅是「工具」而非「目的」，故任何策略或計畫一定要經相關機關認定對其職掌業務有所助益或能解決其關注問題，避免僅為了導入大數據分析應用技術而忽略最基本之初衷。

#### 2. 異中求同研擬跨單位共同議題合併推動

各交通領域因運具差異各有不同挑戰與關注議題，但仍有少數性質具一致性，因此相同性質或需求的議題可考慮合併並優先推動。

### 3. 各單位均應啟動資料品質提昇

訪談中過程中得知多單位均存在若干「資料無留存」、「資料未數位化」、「資料未統一匯流」等資料缺乏或品質不佳之問題，但這些資料往往均與各單位現行的作業流程或資訊系統高度相依，僅能由各單位各自提昇改善。

### 4. 不只「分析」資料，更要「開創」資料源

若只將交通大數據的發展侷限在「分析目前已存在資料集」，未來將很容易遇到發展瓶頸，因此將既有資料集做最好的應用分析之餘，也必須建立新的資料源，才能和現有資料源產生加乘效果，提昇交通大數據之整體價值。

### 5. 依資料完備性決定優先計畫性質

當關注議題分析所需之資料集較為完備時，優先啟動「探勘應用」性質計畫；當關注議題分析所需之資料集較為缺乏時，優先啟動「資料蒐集」性質計畫。

## 3.4.2 優先推動之策略

近 10 年來經歷網際網路、地理資訊定位、行動通訊、平行運算、雲端計算、物聯網等各階段的資訊革命，造就大數據 (Big Data) 有利的發展條件，而在交通領域也隨著這股浪潮發展，包括載具軌跡、電子蒐費系統等都順勢發展，未來應從 Big Data 昇華為 Smart Data，進而達到 Smart Transportation，意即「智慧數據引領智慧交通」，利用數據形成政策與決策。

在「智慧數據引領智慧交通」目標引領下，前述章節表 3.1 與表 3.2 中所整理的 9 項需求 (1 蒐集需求、2 便民資訊、3 提昇安全、4 資產管理、5 延誤改善、6 輿情分析、7 節能減碳、8 追蹤動態、9 績效提昇) 可以歸納彙整成四大策略如表 3.3，再結合 3.3 節的資料匯流協調機制以「制定數據戰略」稱之後組成 SMART 如圖 3-46，各項策略之建議計畫將說明於下節。

表 3.3 原表 3.1 與表 3.2 之應用目歸納表

應用目的	歸納之發展策略
便民資訊	共享旅運資訊

資產管理/節能減碳	強化資產管理
提昇安全/追蹤動態	落實風險評估
蒐集需求/延誤改善/輿情分析/績效提昇	翻轉營運服務

資料來源：本研究整理



圖 3-46 智慧資料引領智慧交通之發展策略

### 3.5 優先推動策略建議

本節分別說明上述各項策略的推動建議。

#### 3.5.1 制定數據戰略

##### 3.5.1.1 策略目的

資料（數據）係大數據發展之基礎，越廣泛之資料融合可以激盪出更多創意火花，因此各機關內的資料均應盡量整合匯流，切勿讓各個資訊系統形成孤島，同時須有完善的機制來確保交通大數據的永續發展，故本策略之目的在於建構多個健全的資料匯流與共享之環境，以奠定交通大數據發展深厚的基石。

##### 3.5.1.2 成功關鍵因素

本策略之成功關鍵因素包括：

1. 數據匯流需要高層充分授權，基層人員方能積極執行，而「交通大數據諮詢評議委員會」更是關鍵，惟有交通部層級由上而下推動方能讓各部屬機關同步齊心運作。
1. 「交通大數據諮詢評議委員會」亦可考慮併入目前既有的類似組織，例如擴充「Open data 資料諮詢小組」之職責，允許部分現階段不宜完全 Open Data 之資料集亦能在有限的單位內匯流介接。
2. 部分機關目前已本於自身業務需求建立機關內共通之資料匯流平臺或系統，或者業務資料庫本身即具有高度分析價值，未來均可在既有的資訊系統基礎下擴充或與外界交換資料。

### 3.5.2 共享旅運資訊

#### 3.5.2.1 策略目的

由於運輸的不可儲存性，無論供不應求造成民怨或供過於求形成資源浪費均非主管機關所樂見，因此供需之間的資料不僅應共享，更應該即時傳遞，本策略欲利用大數據技術讓資訊的價值發揮到極致，以達到供需平衡的理想狀態，這也是大數據定義中 Velocity 的精神之一。

#### 3.5.2.2 成功關鍵因素

本策略之成功關鍵因素包括：

1. 即時乘載人數基於不同運用會有不同精確度需求，以新加坡（第 3.1.1 節圖 3-1 與圖 3-2）案例而言，不須精確數值，只須概分三個擁擠等級即可，但僅有密度高之幹線公車較有機會推動；反之部分路線行經國道之市區公車因為國道不允許站立之規定，須能克服投幣旅客或中途上/下車等關建議題。
2. 共享資訊立意良善，但除讓旅客可以輕鬆取得即時資訊外，也要設計適當的誘因讓旅客願意回饋資訊，甚至能建立商業模式永續營運。
3. 資訊共享時須注意進出量預報資訊可能影響部分旅客的選擇行為。

### 3.5.3 強化資產管理

#### 3.5.3.1 策略目的

許多成功的大數據應用案例均是以「開源節流」的角度切入，此方向值得運輸系統營運單位或公司效法，開源節流的首部曲就是強化資產管理，應用大數據技術協助運輸系統營運單位掌握其資源運用狀況，進而從中找出成本節約方案，或思考如何將閒置資源充分利用，甚至是挪作他用來創造額外價值。

#### 3.5.3.2 成功關鍵因素

本策略之成功關鍵因素包括：

1. 由於資產管理可讓執行單位短期內坐蒐「開源節流」效益，因此相對於其他策略而言較容易推動，但若單位內無成果管控或責任(利潤)中心制度，就會缺乏執行動力，故本策略成功的前提在於要先創造有利於自發開源節流的環境。
2. 開源節流雖立意良善，但從主管機關監理角度仍須關注其發展，避免過度節流而導致服務水準低落，特別是安全領域更不容許因經費節流而犧牲打折。

### 3.5.4 落實風險評估

#### 3.5.4.1 策略目的

「安全」是交通運輸最基礎的前提與價值，故在大數據應用中也不會缺席，透過大數據分析與風險管理的整合運用，達到提升交通運輸安全、防患於未然之目的。綜觀各種運輸系統，公路由於未配置專有路權且一般社會大眾均有機會取得駕駛資格，但經驗與習慣良窳不齊，反觀鐵路、航空、航港不僅路（航）權高度管制，其駕駛也須有專業資格限制與監理，故本策略以公路安全為優先。

#### 3.5.4.2 成功關鍵因素

本策略之成功關鍵因素包括：

1. 目前第三代公路監理資訊系統擁有完整車輛及駕駛人相關資料，是執行本策略最重要的資料集之一，但由於監理資料庫之運用相當敏感，切勿造成違反個資或侵犯隱私之疑慮，故建議由公共運輸駕駛等具公眾利益價值的課題優先。
2. 長期目標可將分析逐漸擴充至一般駕駛，甚至終極理想應設法與第三責任險或相關保險費用關聯勾稽。
3. 除公路系統外，其他運輸系統亦有安全風險評估議題，例如鐵路平交道的鐵公路兩側曝光量計算在不久的將來均可能因為資料蒐集技術的進步（例如電子票證、電子蒐費）讓風險評估更為精確。

### 3.5.5 翻轉營運服務

#### 3.5.5.1 策略目的

過去運輸系統的經營管理上僅能根據少數資料進行決策，而這些有限的資料往往又是彙總後的結果，難以再深入探索或做其他加值應用，同時受限於電腦科技，也無法對龐大資料進行有效率的分析計算。近年來隨著資訊科技的進步，許多資料的蒐集已是常態，如電子票證紀錄等，此外大數據分析技術的發展也已成熟，如集群分析，因此本策略希望能藉此為營運管理帶來劃時代的思維與方法。

本策略之成功關鍵因素包括：

1. 現階段非典型交通時空資料係以電信定位資訊最具發展潛力，但電信定位的運用仍有許多非技術面的課題須釐清確認與突破。
2. 許多交通問題源自尖離峰供需比例失衡，鉅細靡遺的大數據雖能精確「描述問題」，但卻無法單用大數據「解決問題」，須結合營運調度、蒐益管理等手段，故現有的法規制度也應重新檢視是否有鬆綁需要，例如允許業者有條件彈性排班等。

## 第四章 標竿案例分析

本計畫的標竿案例分析主要在探討公共運輸的服務品質，共包含「舒適」、「密集」、「信賴」、「安全」與「無縫」等五個分析方向，本章首先介紹本標竿案例的分析對象，隨後就五個分析方向分別說明其研究成果。

### 4.1 分析對象

本案主要目的為將公共運輸大數據之各項巨量資料，結合客運業者營運路線、班表、司機等資料進行巨量資料分析，透過不同來源資料的探勘、比對與分析，發現隱藏在各資料背後的重要訊息，並轉換為可應用於公、私部門之有效管理資訊。為達到此目的，本團隊依據各地區主管機關、汽車客運與 APTS 營運現況，根據資料取得可行性與完整性進行評估，評估流程如圖 4-1 所示，研究區域評估流程分為五階段，分述如下：

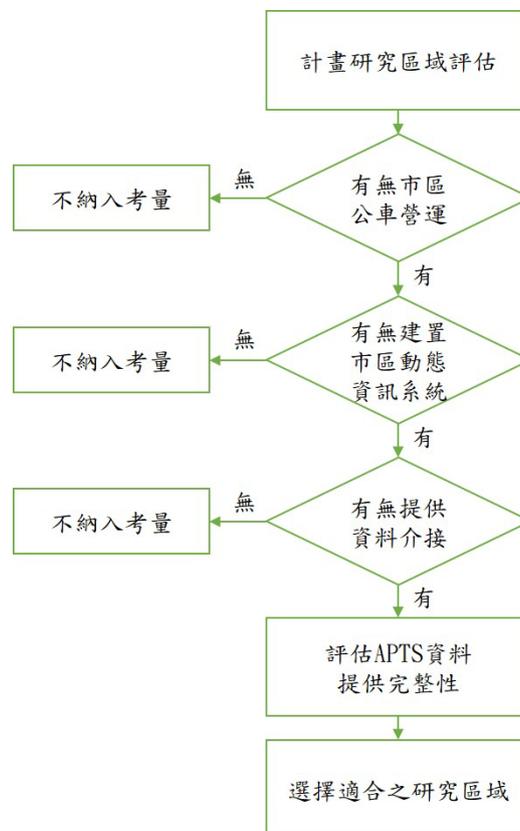


圖 4-1 分析區域評估流程

第一階段評估有無市區公車，目前公路總局已完成公路客運動態資訊管理系統建置，針對公路客運已建立完整之業者、路線、站牌資訊、以及時刻表資訊。然部分地區尚無市區客運，如苗栗縣、南投縣、雲林縣、臺東縣，無法取得市中心公車動態資訊，故本研究排除無市區客運之地區。

第二階段評估有無建置市區動態資訊系統，考量本研究需介接研究區域汽車客運主管機關之 APTS 動態與靜態資料，故本研究進行研究區域評估時，以該區域是否已建置 APTS 為優先考量。目前已建置市區公車動態資訊系統之縣市有彰化縣、嘉義市、澎湖縣、連江縣、基隆市、新北市、桃園市、嘉義縣、屏東縣、高雄市、金門縣、臺北市、臺中市、臺南市。

第三階段評估該區域有無提供資料介接，此階段考量該區域主管機關之配合度，部分縣市礙於設備及配合廠商之限制，使得無法提供外部單位介接公車動態資料，故該地區則不納入考量，如彰化縣、嘉義市、澎湖縣、連江縣。

第四階段則評估該區域 APTS 資料提供完整性，如前述之評估，市區公車資料包括業者資料、路線資料、站位資料及時刻表資料，目前部分縣市已完成市區公車動態資訊系統建置，在其範圍內市區公車已建立完整之業者、路線、站牌資訊、以及時刻表資訊，部分縣市則是將資訊建置於公路客運動態資訊管理系統；其餘縣市則是尚未建置系統。然有些縣市雖已建置完成，但提供的完整度不一，如基隆市、新北市、桃園市、嘉義縣、屏東縣、高雄市、金門縣僅提供業者、路線、站位資料介接服務，而臺北市、臺中市、臺南市則提供業者、路線、站位、時刻表資料介接服務。囿於本研究希冀結合客運業者營運路線、班表、司機等資料進行巨量資料分析，故經過評估，此階段符合本研究需求之區域為臺北市、臺中市、臺南市等三個縣市。

第五階段為選擇適合之研究區域，根據前述評估過程，篩選出適合本研究計畫之區域為臺北市、臺中市、臺南市等三個縣市，這三個縣市其汽車客運發展均有其特性。臺北市雖公共運輸使用率為全國最高，其汽車客運資料多著重於市區公車，轄區內公路客運路線較少，較難分析市區客運及公路客運之營運狀況。臺南市公共運輸正在發展中，雖已建置 APTS，但囿於公車路線數及運量資料量較少，較難以運用巨量資料進行分析。臺中市近年來公共運輸快速發展，搭乘數屢創新高，值得深入探討之地區，且其市區內客運業者多營運市區客運、公路客運及國道客運，此特性有助於藉由巨量資料分析方法探討城際運輸及區內運輸旅次行為之差異。另一

個特點為臺中市位處西部城際走廊中心，為國道客運交會處，係為我國重要的轉運點，故本研究以臺中市作為本計畫研究區域。

## 4.2 舒適面向分析

過去許多研究均曾分析過公車動態或電子票證兩個資料集，前者的分析價值在於運輸供給面（班次數、班距...等）；後者的價值在於使用人數（上車人數/下車人數）等，但若能以演算法融合兩個資料集如圖 4-2，即可分析供給與需求匹配的狀態，由於公共運輸車輛的乘位（座位＋立位）有限，因此可利用乘載率（或稱利用率）來描述供需匹配狀態，此種分析可利用微觀的方式呈現「供過於求形同運能浪費」或「供不應求服務水準低落」等兩種極端狀況，這種分析結果無論就主管機關或經營者的角度而言都相當重要。



圖 4-2 乘載率分析架構示意圖

無論各式公共運輸，電子票證相關之分析可分為兩大類，其一簡稱為節點流量分析，係分析旅客搭乘之起迄地點，其二簡稱為節線流量分析，係分析車上搭乘人數，兩者兼之比較摘要整理如表 4.1。

表 4.1 節點流量與節線流量分析比較表

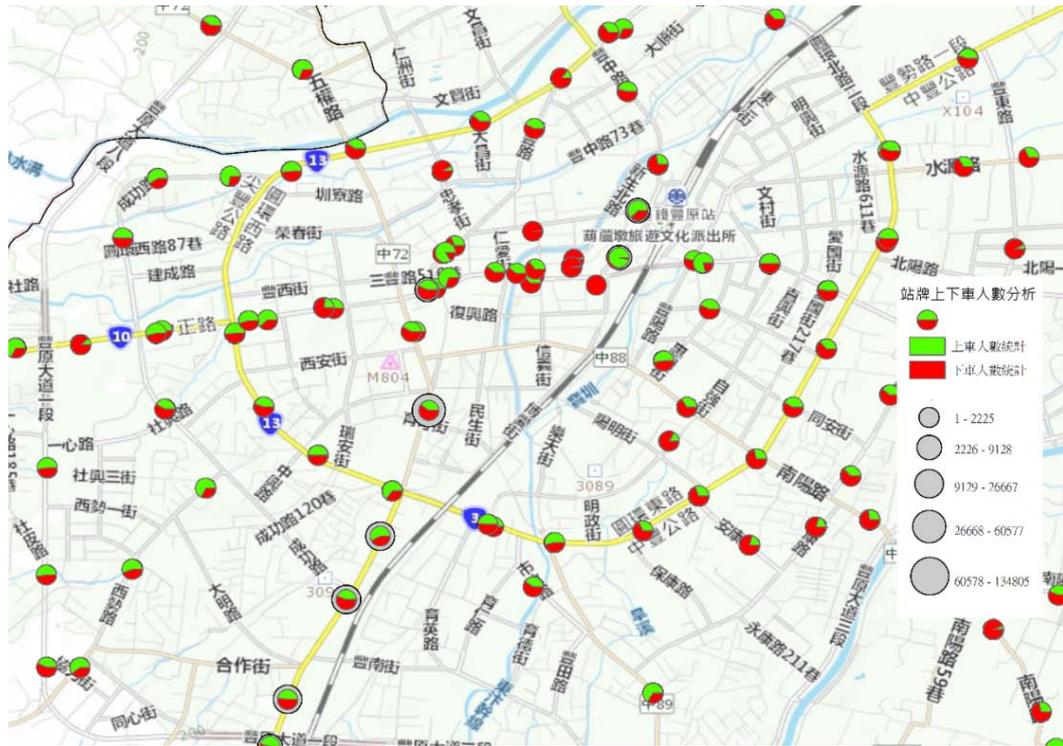
	節點流量	節線流量
分析標的	各站牌上車/下車人數 各車站進站/出站人數	載具上旅客數
應用目的	瞭解旅客起迄實績	掌握乘載服務水準
所需資料	票證資料	票證資料＋載具軌跡
相關指標	旅次數、票箱蒐入	延人公里、延座公里

資料來源：本研究整理

## 4.2.1 節點流量分析

節點流量分析目的在於分析旅客經常上車/下車之站牌位置，上/下車總人數除影響站牌停等上下客時間，進而影響總行車時間外，也與票箱蒐入高速相關。以下將逐一說明節點流量分析之數據探索方式。

節點流量最常見的靜態呈現方式如圖 4-3，除能清楚表達地理位置之餘，可利用圖塊大小來表達數量等級，同時用圓餅圖來比例來呈現上/下車比例。



資料來源：公車動態資訊系統巨量資料 (Big Data) 蒐集與視覺化分析期末報告

圖 4-3 站牌上下車人數分布圖

圖 4-3 僅分別呈現上/下車旅客總數層級，並無法描述旅次起迄呈現方式（詳見圖 4-4），將該案例之橫軸改良為地理資訊系統座標後如圖 4-5，此呈現方式雖可明確表達各站牌間的起迄關聯，但各起迄之間旅客數多寡的判別能力較為低落，雖然圖 4-5 左下角圖例說明不同粗細弧線所代表的旅客數，但在判讀上不甚容易，還是需要借助滑鼠以互動 Tooltip 方式輔助之。

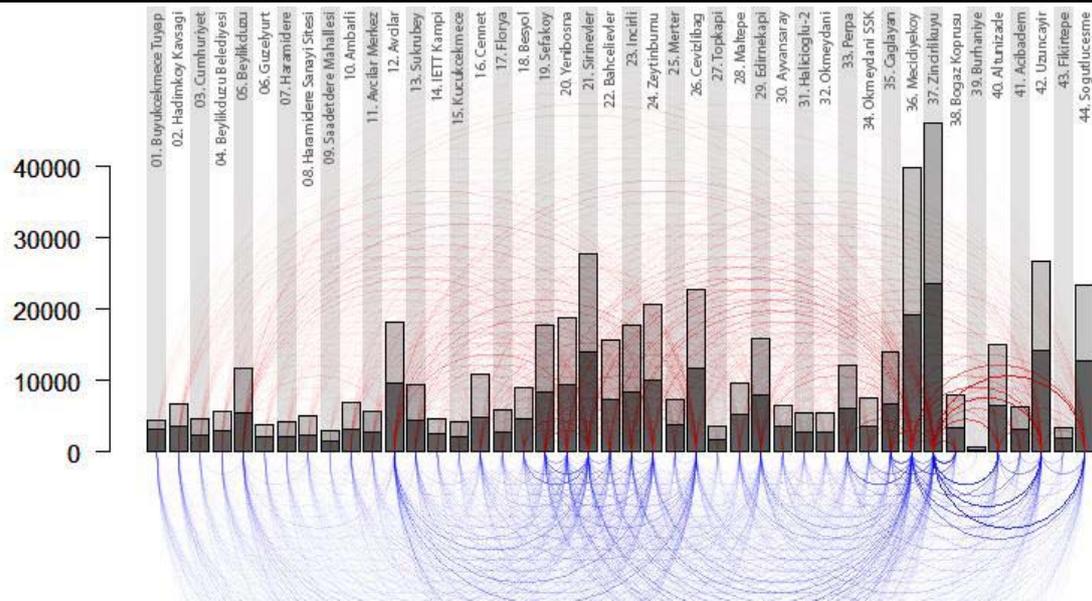


圖 4-4 土耳其 BRT 研究之 O/D 起迄呈現方式



圖 4-5 結合 GIS 與 O/D 呈現 (以 208 路線為例)

相對於圖 4-5 這種相對花俏但實用性低的起迄呈現方式，本研究建議利用運輸領域 OD 起迄表概念搭配熱圖 (Heat Map) 即可得到更佳效果，主要原因在於站牌與起迄組合眾多，後者的數量會依據前者的數量平方倍成長，受限於紙張或螢幕大小限制，不適合呈現起迄數值，故本研究改採用熱圖 (Heat Map) 呈現如圖 4-6，除解決篇幅大小問題外，熱圖利用顏色申請來判斷起迄數多寡也較為便捷，需要讀取確切數值時，則可利用移動滑鼠到關切起迄後以 ToopTip 呈現數值。

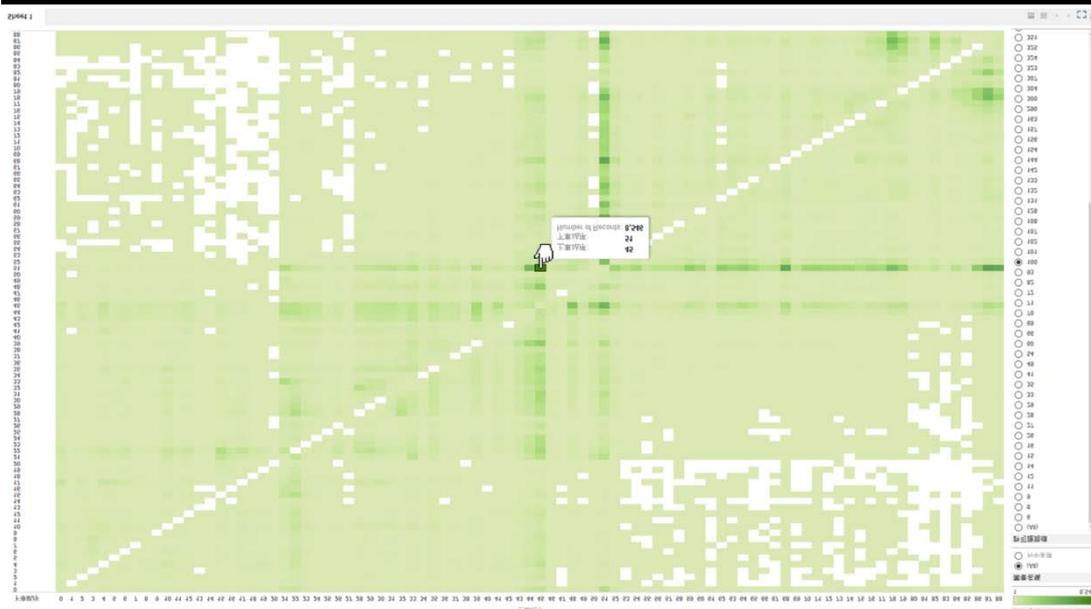


圖 4-6 旅客起迄 Heat Map 圖（以 100 路公車為例）

但無論是圖 4-5 或圖 4-6 均只能檢視單一路線之起迄行為，無法探索路廊、區域等多路線之旅客起迄樣態，在比較各種資料視覺化探索方案後，最適當的探索方案是將上車旅次、下車旅次分別用地理座標呈現如圖 4-7。

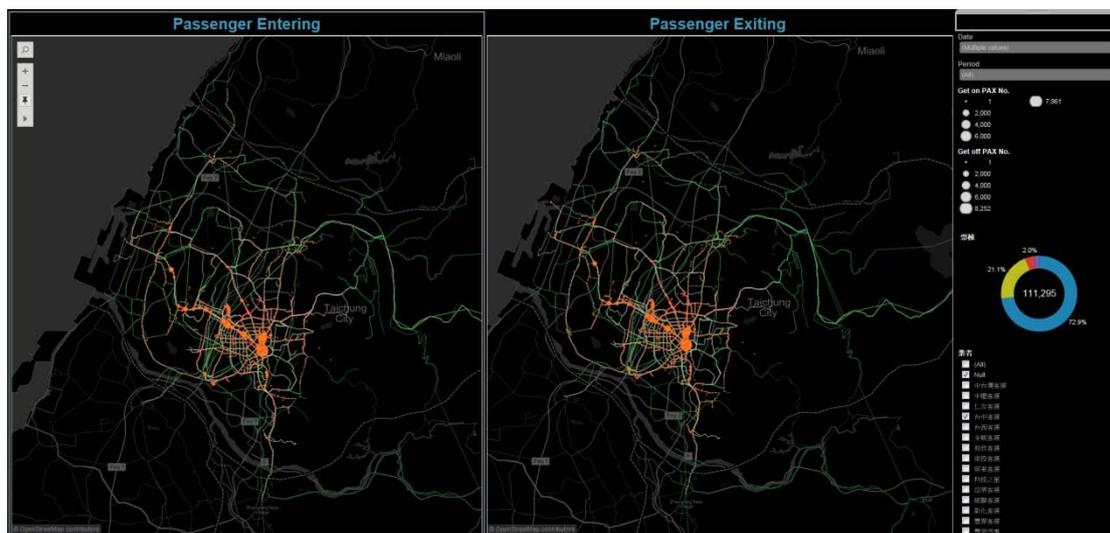


圖 4-7 具業者/日期/時段/起迄多重選擇之互動儀表板

圖 4-7 右側提供日期、時段、業者三種篩選條件，同時均支援多重選擇如圖 4-8 以滿足觀察特定日期範圍或尖離峰連續時段之用。

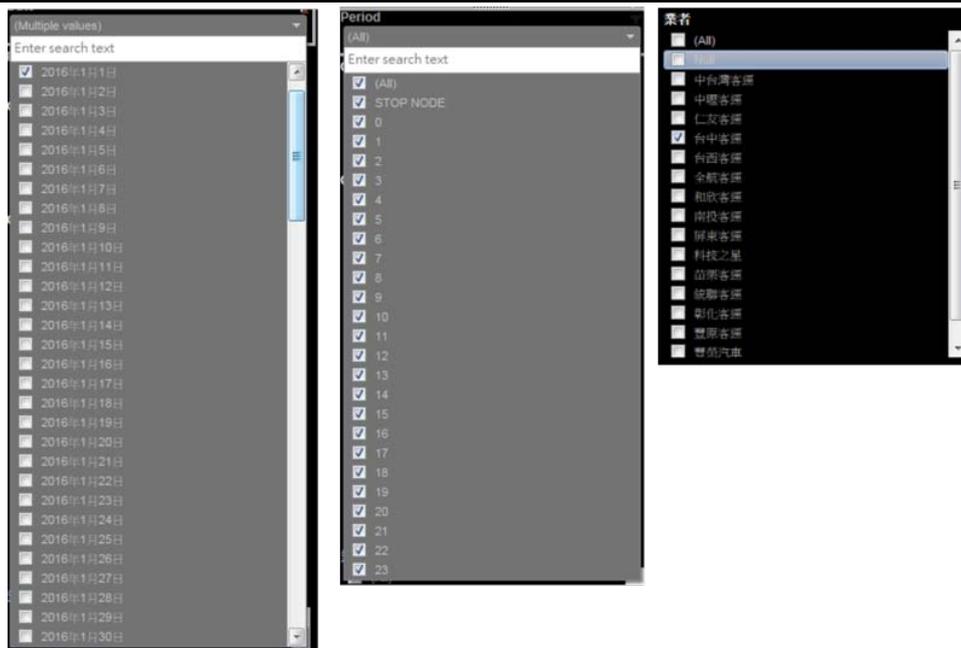


圖 4-8 於圖 4-7 右側之篩選過濾器選項

圖 4-7 中地理資訊系統中除以「●」描述站牌位置之外，亦以「●」大小來呈現數量多寡，同時點選特定上車站牌如圖 4-9 時，右側的下車站牌、票種圓餅圖均會同步更新，有助於探索旅客起迄樣態。



圖 4-9 點選單一上車站牌之儀表板互動示意圖

在探索過程中，單一站牌之選取可能稍有不足，因此儀表板應該保留選取指定範圍，例如圖 4-10 係以靜宜大學附近 5 公里為站牌圈選範圍，選取確認後之儀表板則如圖 4-11，用以瞭解開區域搭乘臺中客運之旅客之下車目的的分佈。

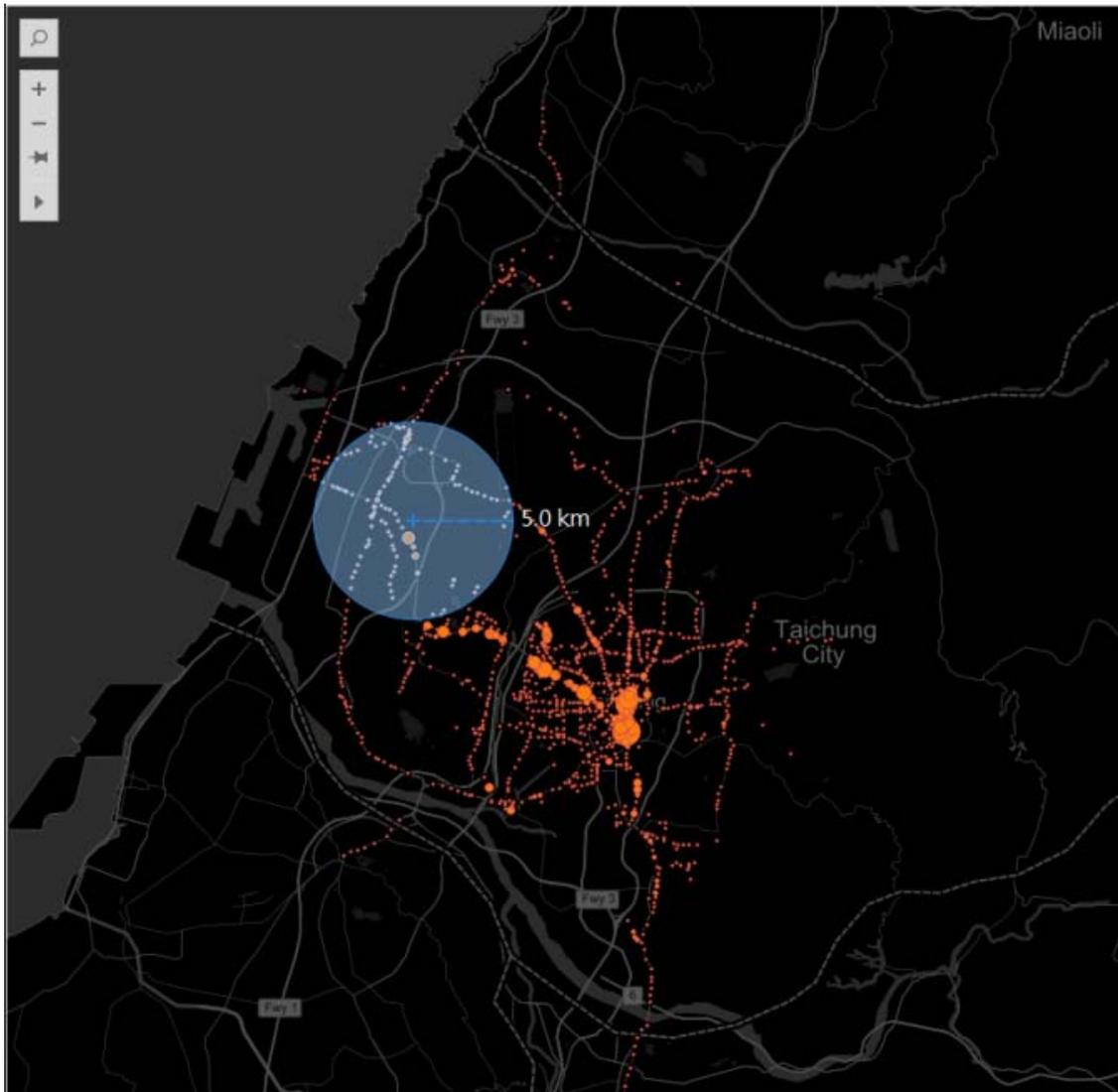


圖 4-10 選取特定點方圓 5 公里站牌操作示意圖



圖 4-11 對應圖 4-10 之儀表板互動結果示意圖

## 4.2.2 節線流量分析

### 4.2.2.1 分析流程概要

有別於節點流量，節線流量分析必須考慮以下兩個因素：

#### 1. 旅客搭乘起迄（含空間與時間）

旅客起迄的時間和空間若同時集中，該節線上的乘載人數就會提高，同時旅客搭乘距離亦是重要影響因素，同是 100 位旅客上下車，平均搭乘 5 個站間和平均搭乘 30 個站間相比，雖然旅次總數相同(甚至票蒐也相同)，但是後者的延人公里將遠高於前者。

#### 2. 旅客搭乘車輛（載具）

舒適面向只考慮延人公里是不夠的，必須納入供給端的客座公里因素，同樣的延人公里，若由少數車輛提供服務，在舒適度的表現將不如多數車輛服務之情境，換言之納入電子票證「車輛車號」欄位來分析。

綜合以上兩點，考量各式公共運輸之乘位（座位＋立位）在營運過程中是不會任意變動的，故乘載率演算法之關鍵就在於車上旅客數之計算，求算的方法可先取得特定路線特定車輛單向的各站牌旅客上下車人數如圖 4-12，再分別針對上車旅客與下車旅客做累積旅客數繪製如圖 4-13，由於上車與下車旅客總數會維持物質不滅平衡，故圖 4-13 最後累積的加總會相同。最後將各路段的累積上車數扣除累積下車數可得到如圖 4-14 之各區間乘載量，而其中數值最高者即為最大乘載區間（Maximum Load Section，MLS）。以上基本理論在實作時必須考慮車上蒐費/車外蒐費特性、資料闕漏、多路線多天期等因素適當編寫程式碼。

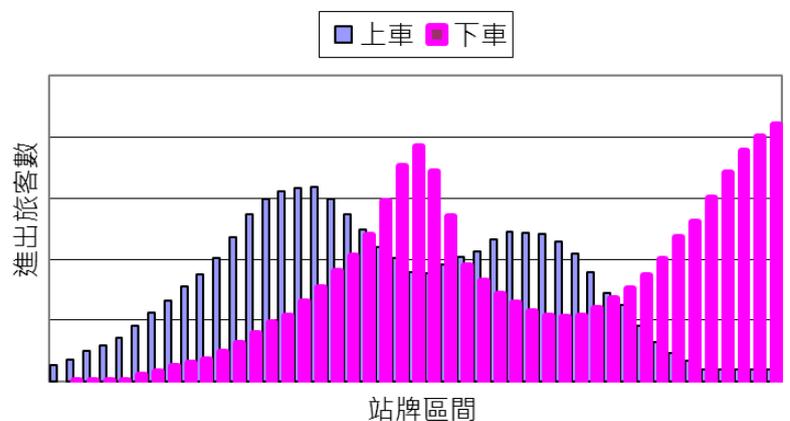


圖 4-12 車上乘載人數計算程序 I—產生上/下車分布圖

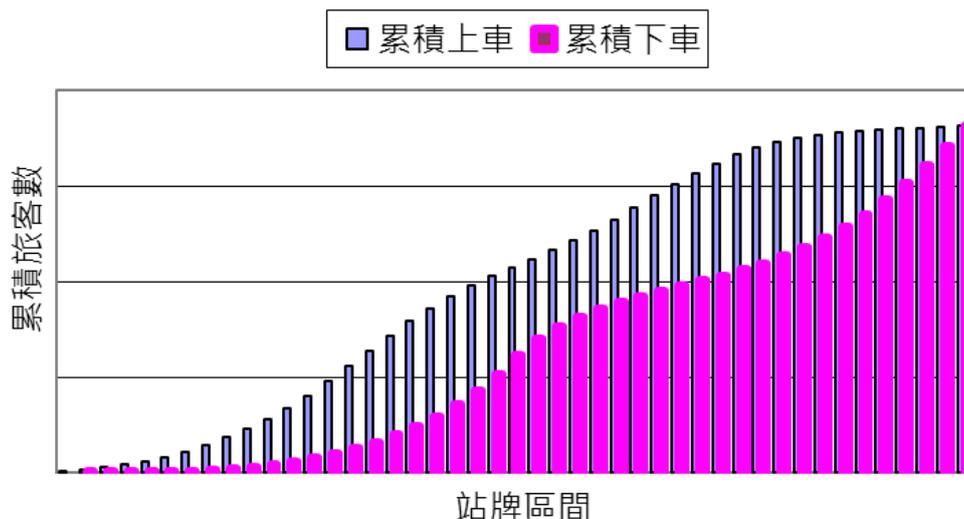


圖 4-13 車上乘載人數計算程序 II—產生累積上/下車數分布圖

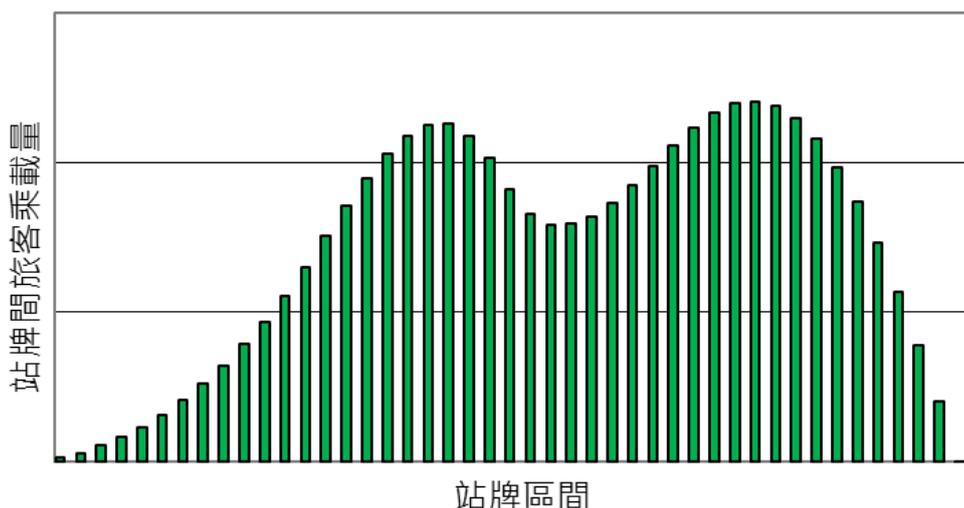


圖 4-14 車上乘載人數計算程序 III—累積上/下車數相減完成

#### 4.2.2.2 單路線乘載人數檢視

在實作圖 4-12~圖 4-14 概念之演算法後即可從公車動態和電子票證融合出臺中地區各業者/各路線/每日任何時間點之「車上旅客人數」，同時利用圖 4-15 的儀表板讓研究人員可以任意挑選業者、路線、日期來檢視，圖 4-15 截圖為 33 路公車 2016-4-29 日當天共計 13 輛車全天候的車上人數紀錄。

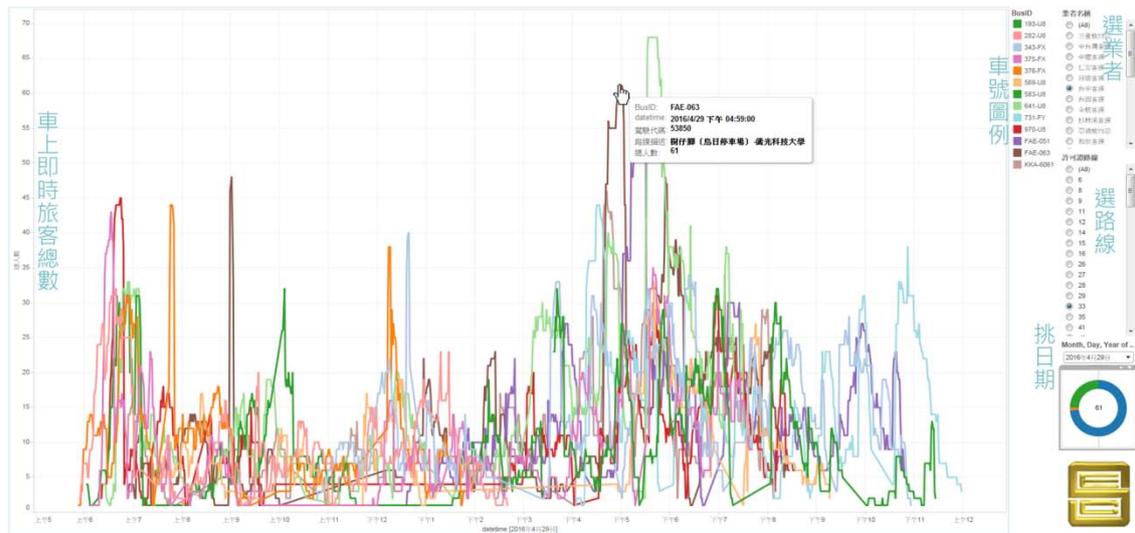


圖 4-15 單日車上旅客數分析探勘儀表板截圖

圖 4-15 係以特定路線所有車輛全天之車上旅客數總覽，但從中不利觀察路線上各車輛班距，同時僅能知道高載客之時段，無法判讀高載客之路段。因此本研究將圖 4-15 之 Y 座標改為車站站牌，原圖 4-15 中 Y 座標之總旅客人數改以顏色呈現如圖 4-16。

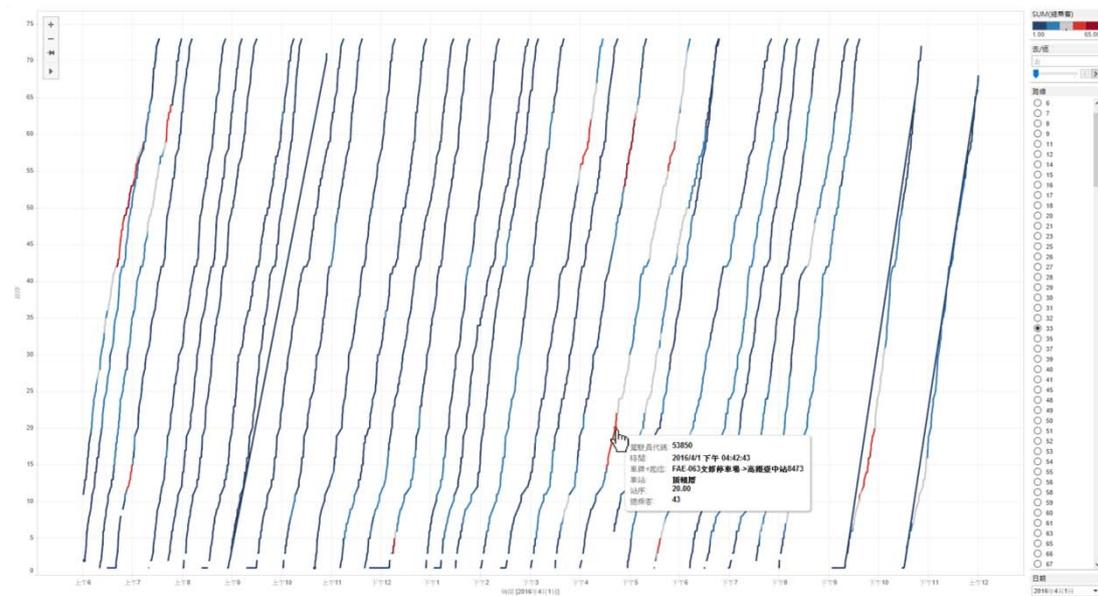


圖 4-16 由圖 4-15 改良之載具時空軌跡套疊乘載人數儀表板

#### 4.2.2.3 多路線乘載人數檢視

圖 4-15 與圖 4-16 之優勢是可針對指定路線做微觀、高解析的檢視，但分析的標的僅限單一路線，本節將近一步介紹多路線乘載人數資料呈現方式。

對於總班次較少之路廊，以豐原～東勢間為例，主要有 206、207、208 與 209 四條路線重疊如圖 4-17，故可擴充原圖 4-16 之車輛時空暨乘載圖如圖 4-18，考慮路廊上有多條路線，因此圖 4-18 之設計係以顏色做為路線區隔識別，乘載人數多寡則改由時空線粗細來描述。

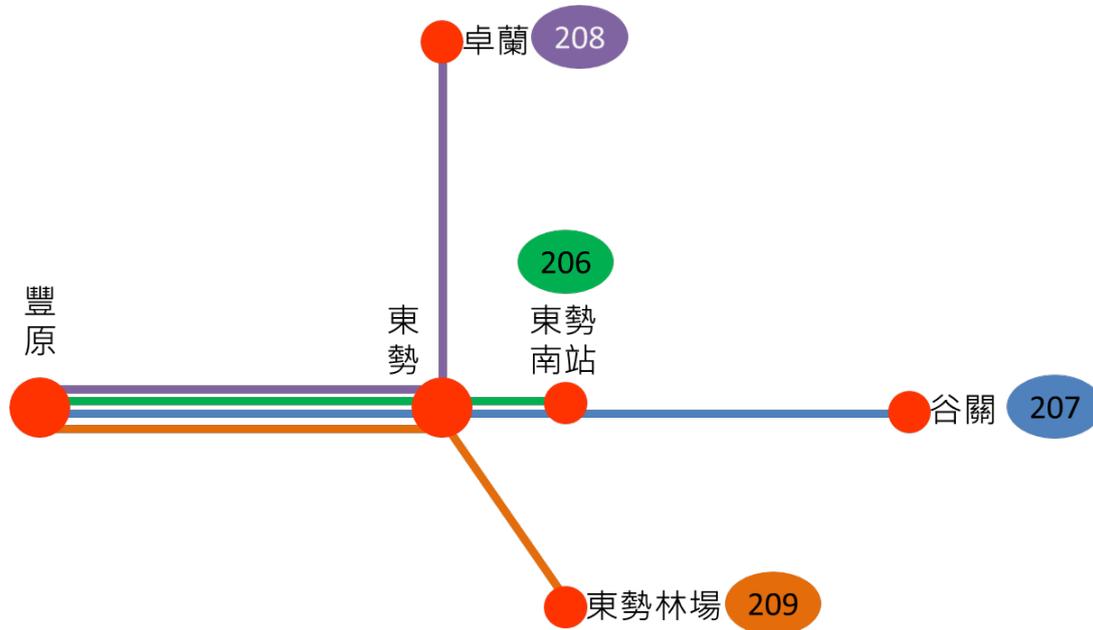


圖 4-17 豐原-東勢路廊主要路線起迄圖

由圖 4-18 發現若單看單一路線(例如 208)班距並不平均並略有不足，但若以整個路廊角度而言，豐原東勢之間的供給相當充分，欲考慮多路線組成之路廊供需時，就必須採用圖 4-18 之儀表板探索，將比圖 4-16 單一路線之檢視更為完備。



圖 4-18 豐原-東勢路廊主要路線起迄圖

圖 4-18 跨路線之探索仍有其限制，例如單一路線就已是高密度路線，或者路線重疊變化較多，甚至是平行路線等情境均不適合採時空圖方式呈現乘載人數，而必須改採地理資訊系統基礎之呈現如圖 4-19，一旦採地理資訊系統方式呈現後，就無法在同一時間描述時間維度，必須改以動畫的方式呈現不同時間之車輛位置與乘載人數變化。

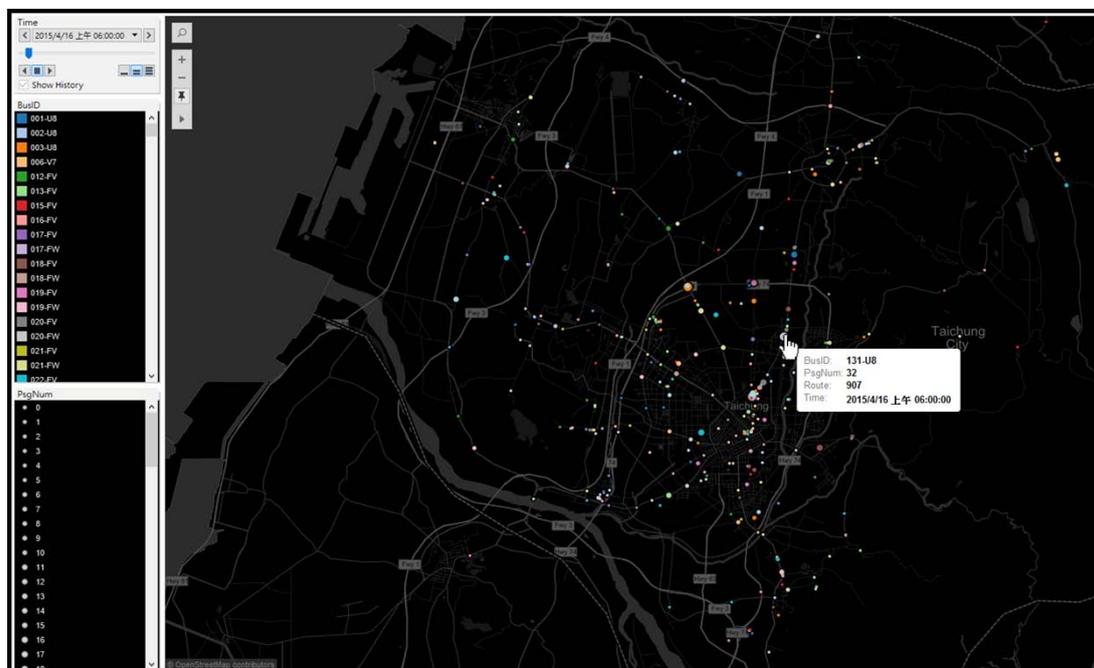


圖 4-19 結合 GIS 與軌跡之乘載人數呈現

圖 4-19 之資訊相當充分，但由於車上乘載人數的變化隨著時間變化，較難觀察出不同時段之數量差異，故可以類似圖 4-20 之方式呈現全日變化，以此圖為例，係呈現今年(2016 年)4 月 29 日臺中客運（多路線多車輛）在各時段的車上旅客數彙總趨勢。

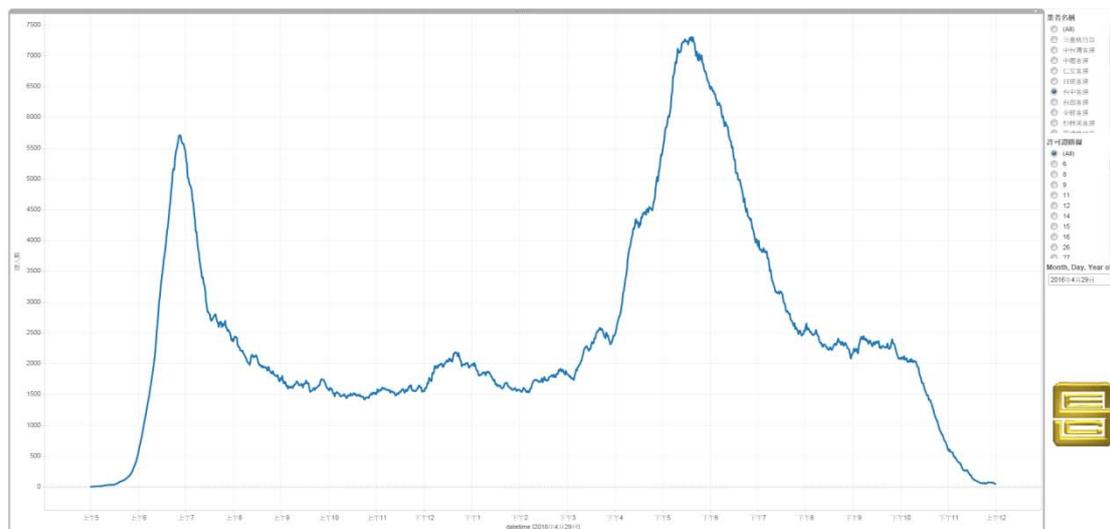


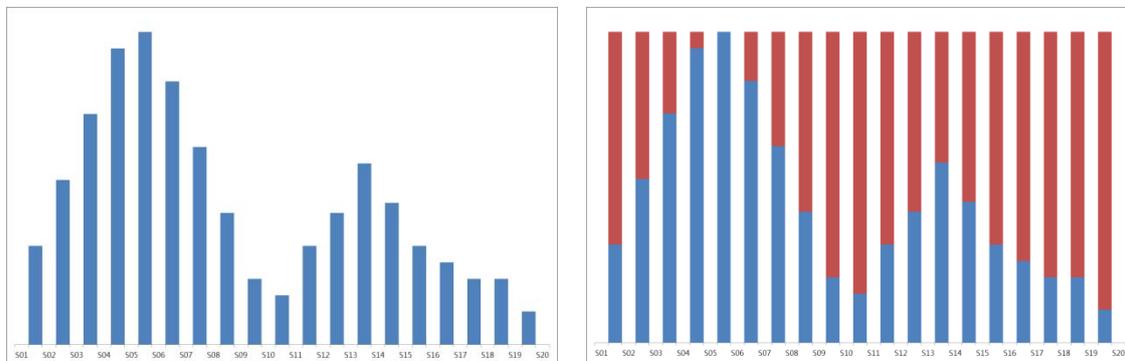
圖 4-20 單日車上旅客總數隨時間變化趨勢圖（以臺中客運為例）

## 4.3 密集面向分析

### 4.3.1 區間車可行性分析

#### 4.3.1.1 區間車設計理論與原則

在掌握各路線站牌間車上乘載人數後，可嘗試調整路線的營運區間，從圖 4-12~圖 4-14 的分析流程可知，由於各站牌進出的乘客數量存在不一致的現象，進而導致站牌間通過量也會有差異，假設 S01~S20 共 20 站牌 19 個區間的站牌間運量如圖 4-21 左側，若以「滿足最大乘載區間(S05~S06)之需求」為目標規劃班次數，在其他區間 (S01~S05 與 S06~S20) 則勢必會有運能的損失如圖 4-21 右側紅色柱狀部分。



S01~S20 各站牌間運量示意圖

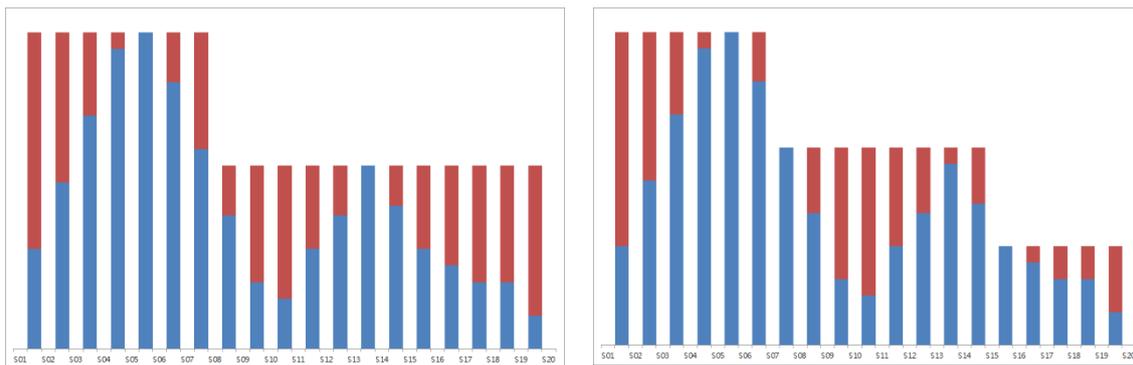
S01~S20 站牌間乘位損失示意圖

圖 4-21 站牌間乘位損失示意圖

如欲減少圖 4-21 右側之紅色損失運能，理論上可將 S01~S20 間以 S08 為界切分為兩段如圖 4-22 左側，兩段營運起迄的最大乘載區間則分別在「S05~S06」與「S13~S14」間，若基於這兩營運路段的最大乘載區間規劃公車班次數量，相較於圖 4-21 右側則可減少運能之浪費(紅色柱狀減少)，以上即是切割路線營運起迄的主要效益。

同理，若將 S01~S20 間以 S07、S15 為界切分為三段如圖 4-22 右側，三段營運起迄的最大乘載區間則分別在「S05~S06」、「S07~S08」與「S15~S16」間，若基於這三個營運路段的最大乘載區間規劃公車班次，則可再度降低運能之損失，但管理成本也相對更提高，須做適當權衡。

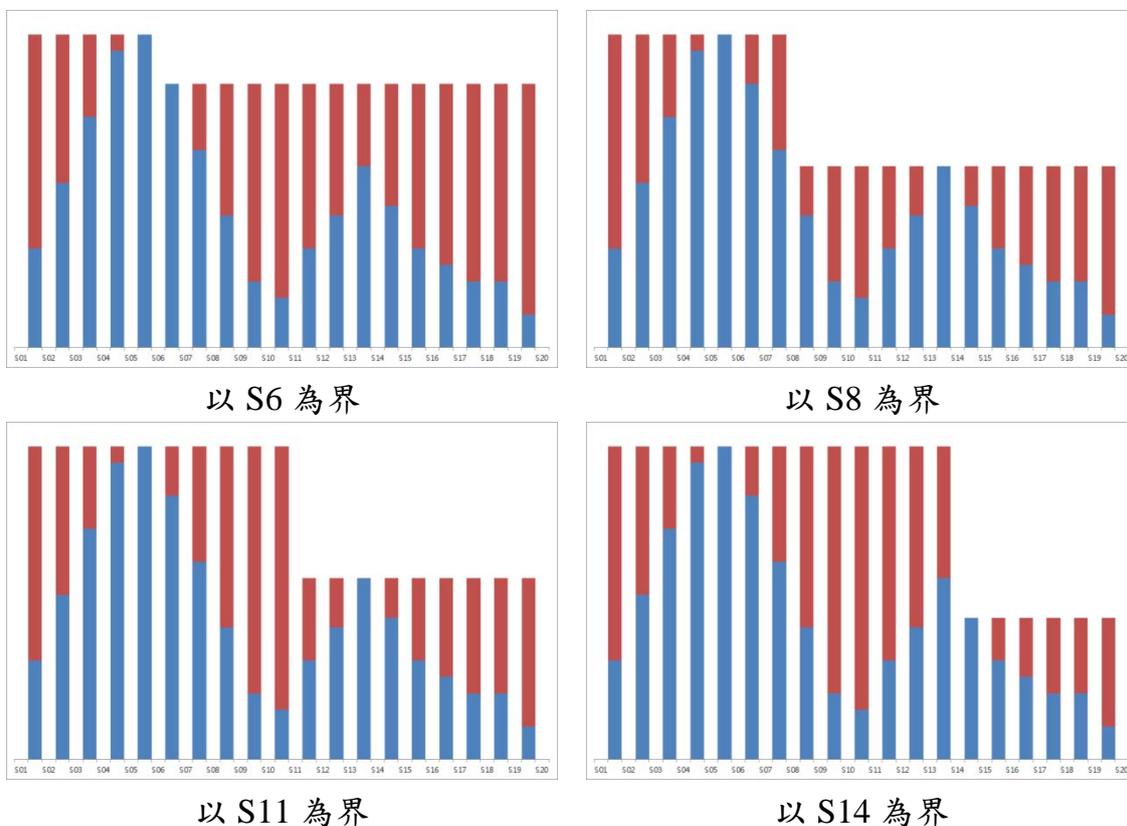
此外，路線營運起迄規劃不單純劃分幾個區間之議題，即使劃分區間數量相同，不同劃分處也會產生不同的轉乘衝擊及運能運用效率，同樣以圖 4-21 左側的站間運量為例，即使同樣劃分為兩個區間下，圖 4-23 分別以 S06、S08、S11 與 S14 呈現不同劃分方式之運能損失狀況。



2 種區間營運型態

3 種區間營運型態

圖 4-22 不同營運起迄區間數量之效果差異圖



以 S6 為界

以 S8 為界

以 S11 為界

以 S14 為界

圖 4-23 以不同車站為界之運能損失差異示意圖

綜合前述，本研究未來執行路線營運區間調整時，為避免業者管理成本過高，初步會以兩個區間來規劃，暫時不嘗試同路線切割為三個營運區間，至於切分點則會以最佳化方式來分析評估。

此外，切分路線營運起迄除了管理成本提高外，尚有其他副作用，同樣以圖 4-21 為例，一旦以 S08 為界分區營運，勢必導致部分起迄點跨區(例如由 S03 往 S11)之旅客無法直達，根據本團隊過去之規劃經驗，會建議

捨棄如圖 4-24 上方的切分營運，而是採用如圖 4-24 下方的重疊區間營運設計，如此可在達成「減少運能損失」目標下，同時又不會增加跨區旅客之不便（轉乘）。

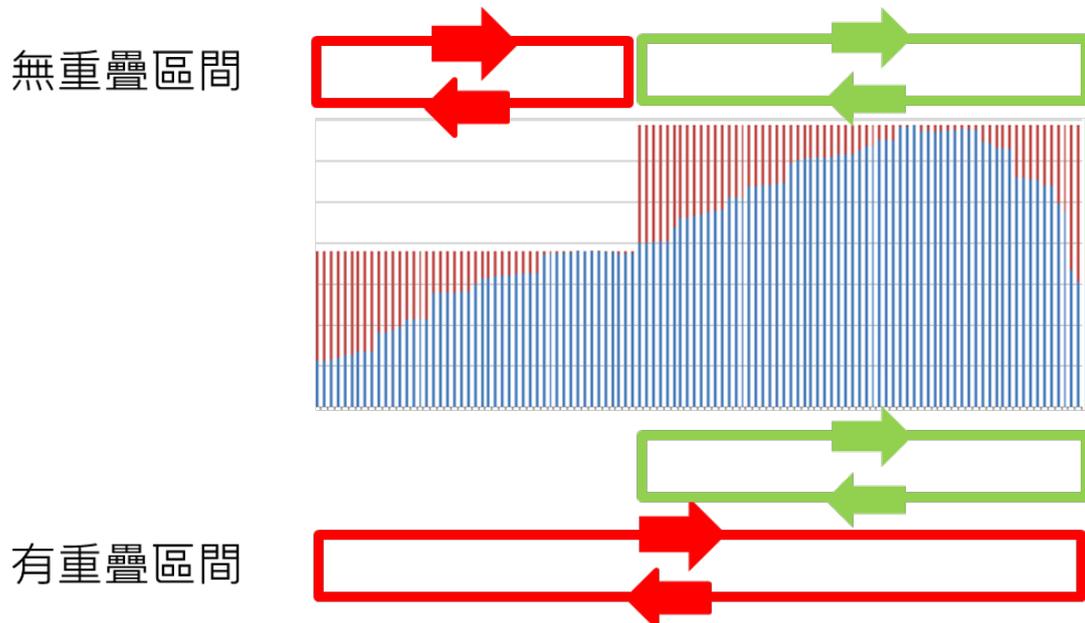


圖 4-24 重疊區間中途折返之規劃方式比較

#### 4.3.1.2 重疊區間決策輔助工具

從圖 4-21～圖 4-24 之分析可知，區間營運以重疊區間方式較為合適，同時區間車不應過於複雜，除原先核定的完整路線外，僅適合增加 1～2 個重疊區間，重疊區間之決定雖可根據圖 4-16 來輔助決策，但圖 4-16 係呈現單一日細部乘載細節之儀表板，規劃人員不可能以單日乘載細節作為決策依據，因此本研究進一步開發了整合全年熱區的互動儀表板如圖 4-25，除可挑選路線外，亦可調整「熱區」定義，預設值為 30 人，儀表板會即時更新使用者指定路線中，全年度裡超過熱區定義人數之「重現天數」，重現天數越高的站牌空間即是考慮區間之處，以圖 4-25 為例，上午的空間熱區係以潭子火車站附近為主，一年約重現 180 次，而下午的空間熱區則在新民高中附近，使用者可任意調整路線、熱區定義、分析月分、平常日/例假日等，意即數億筆的公車動態與電子票證已經融合成一個簡單亦用的決策輔助系統，協助規劃人員決定區間車之範圍。

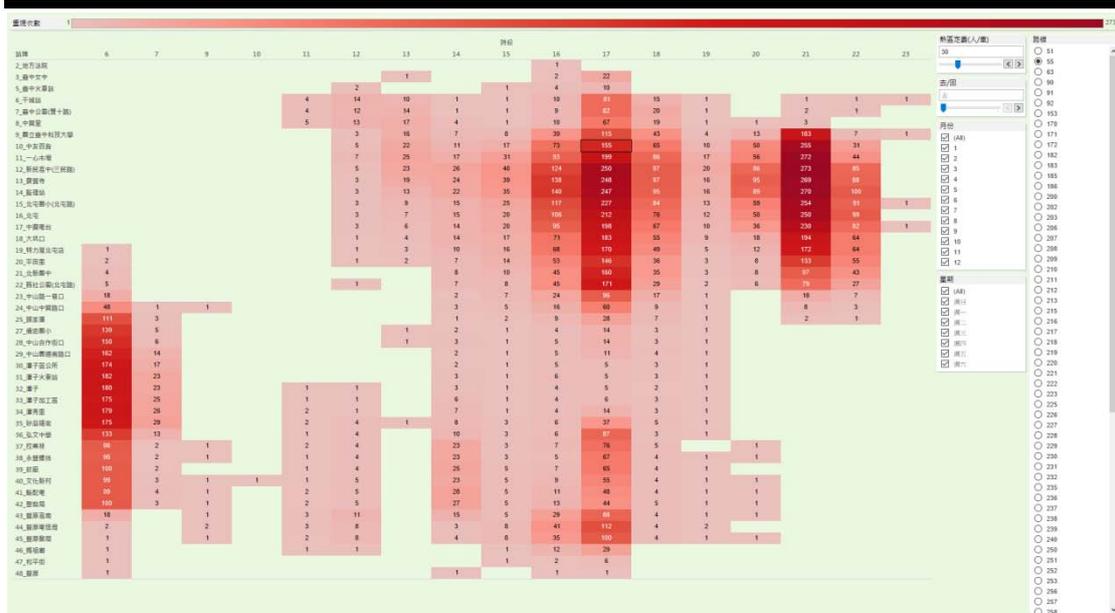


圖 4-25 各路線全年乘載人數熱區套疊統計儀表板

除區間車之輔助決策外，圖 4-25 也可改善強化如圖 4-26，用來判斷週間（週一～週五）與週末（週六週日）的乘載狀況是否有顯著不同，作為是否需要建構不同班表或者各時段應增/減班次之參考。



圖 4-26 各路線平假日乘載人數熱區比較儀表板

#### 4.3.1.3 轉乘衝擊評估輔助工具

區間車雖可達到減少運能浪費的目的，但卻也造成跨區旅客的不便，即便採取圖 4-24 下方的重疊區間營運設計，由於原本所開行的班次中有一部分改為區間車，使得跨區旅客的等待時間變長。因此除了上述協助規劃

人員決定區間車範圍的儀表板外，本研究亦針對此課題開發了衝擊評估的互動儀表板。

圖 4-27 為區間車衝擊評估互動儀表板，主要有四個區域：

1. 第 I 區：以公車時空圖顯示車上旅客人數，其縱軸為空間、橫軸為時間，圖中每條線代表一臺公車在時空中行駛的軌跡，並以線的顏色顯示公車上的旅客人數。
2. 第 II 區：以直條圖顯示各站間通過的旅客人數，其縱軸為空間、橫軸為旅客人數，同時圖中亦會根據第 IV 區的參數設定顯示區間車服務範圍及運能供給情況。
3. 第 III 區：顯示各起迄的旅客人數，其縱軸為空間，每一直條矩形表示一種旅次起迄組合的起迄車站的位置，其長度即表示行經站數，而矩形的顏色則表示該起迄組合的旅客人數。
4. 第 IV 區：規劃人員可透過調整此區中的各項參數，來評估區間車所帶來的影響。

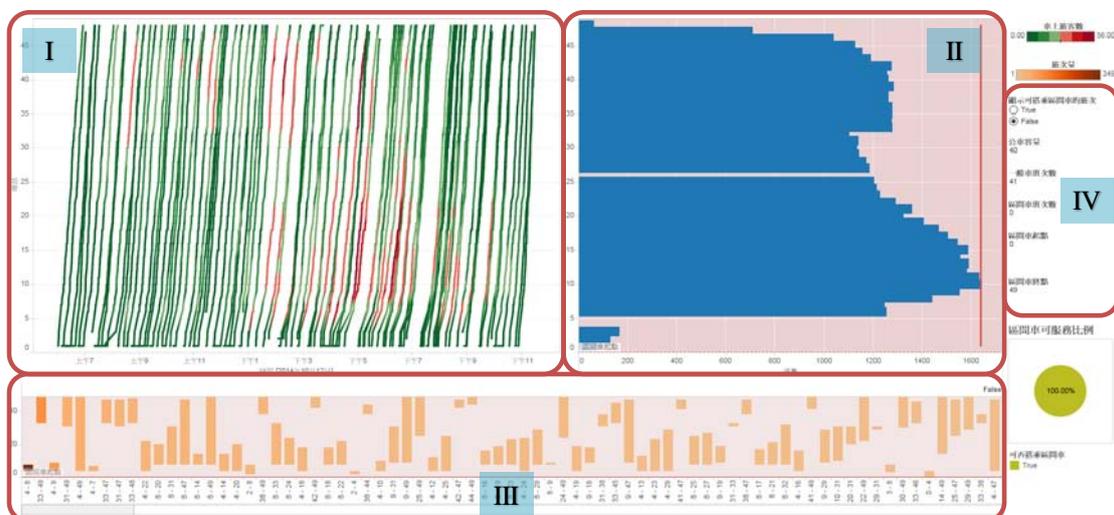


圖 4-27 區間車衝擊評估互動儀表板

在區間車衝擊評估互動儀表板中，第 II 區會依照規劃人員在第 I 區選擇的公車，自動加總計算其站間旅客人數，藉此掌握最大乘載區間所在，以及站間運量變化情形；而第 III 區則會根據規劃人員在第 II 區所選擇的區間，自動統計通過該區間之旅客的起迄組合，進而了解通過該區間的熱門旅次起迄為何，以及區間車可能影響的旅次，如圖 4-28 所示。

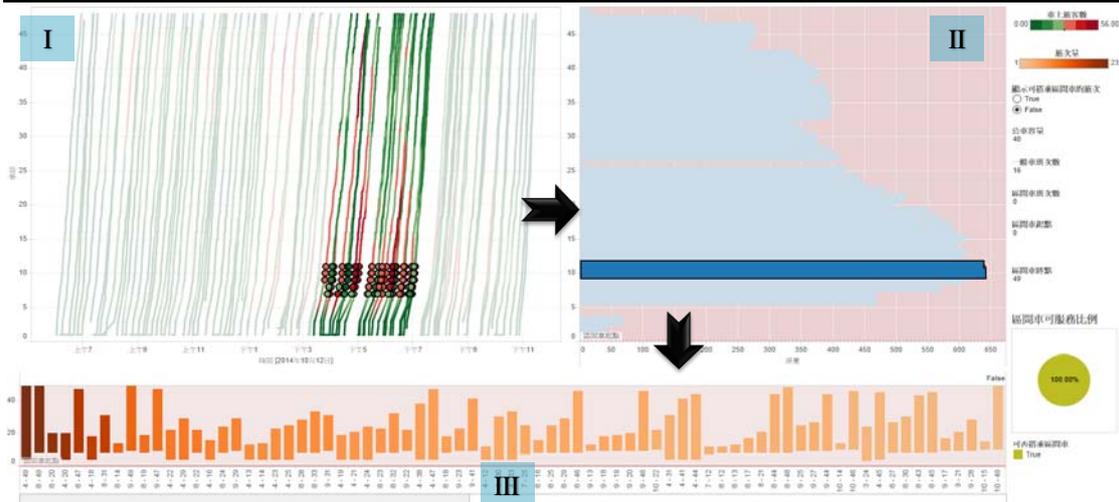


圖 4-28 區間車衝擊評互動估儀表板的互動方式

最後，規劃人員可利用第 IV 區的參數調整，包含區間車起終點、公車容量與班次數等，來決定如何開行區間車，其規劃結果亦會在第 II 區中呈現，如圖 4-29，第 II 區將會根據第 IV 區的參數設定，顯示公車供給量的情況與區間車可服務的旅次，如此便能協助規劃人員決定較佳的區間車範圍，讓區間車能服務到較多的旅次，同時一般公車也要提供足夠運能來服務不在區間車服務範圍內的旅客。

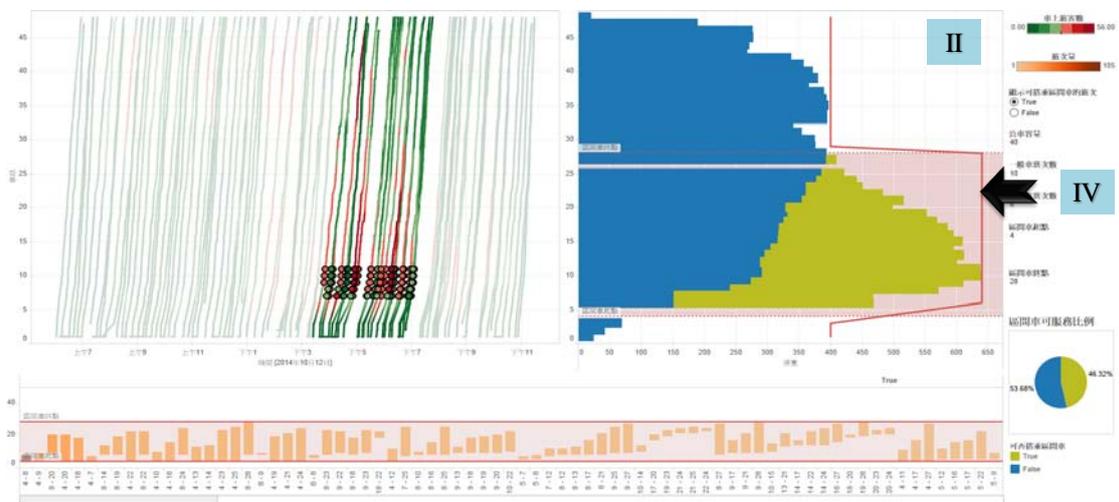


圖 4-29 以區間車衝擊評互動估儀表板協助設計區間車

#### 4.3.1.4 區間車服務模式配套

區間車服務模式不僅可補足尖峰時段部分區間運能不足問題，亦可減少運能損失，透過上述儀表板可協助營運單位找尋合適的區間車服務範圍，同時也能將旅客轉乘的衝擊降至最低，然而主管機關對於業者在營運上規定，須給予足夠的彈性。

首先在區間車的服務時間方面，不應侷限僅能在上下班之尖峰時段開行區間車，尤其是在假日的上下班尖峰時段可能不需要區間車，而是需要在其他時段開行。此外，在連續假期或有特殊活動期間，也應讓業者能視需求提供區間車服務。

其次在區間車的服務範圍方面，隨著區間車的服務時間不同，其服務範圍也會不同，但原則上是以將原路線分兩個區間為上限，而且區間的里程長度不應過短，若原路線的總里程長度本身就不長，則不應再規劃區間車。當業者欲申請的區間車服務範圍和其他現有路線重疊時，主管機關可能須做整體路廊之分析，或和其他業者共同探討，以確認是否有區間車服務之必要。

而在區間車的服務頻率方面，雖然為顧及原有旅客權益，應該是在原路線的運輸服務不變的前提下額外提供區間車服務，但若考量業者營運成本以及減少不必要的運能浪費，當原路線的運輸需求不高時，可考慮適度減少原路線班次，提高區間車班次。

最後，為了避免旅客混淆，在業者能彈性規劃區間車的同時，也必須能清楚標示每輛公車的行車資訊，最後能以明顯的色差以利旅客識別，但這也可能衝擊業者車輛交互流用之彈性。

### 4.3.2 班表優化調整

大數據分析除檢討過去營運績效外，也能進一步預測未來並調整營運策略，若就圖 4-30 之資料探勘五個層次而言，第 4.2 節與第 4.3.1 節等均已經分別介紹第二層級（查詢）與第三層級（加值）的應用分析，本節將近一步介紹推論預測與策略優化之應用。

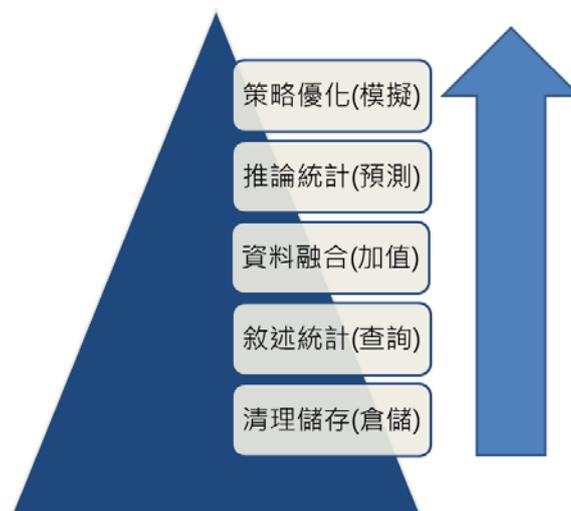


圖 4-30 資料探勘五步驟示意圖

欲調整未來營運策略，將包含預測未來旅客數並依據預期的旅客時空分佈來規劃班表，故以下分別說明這兩類分析程序的相關應用。

#### 4.3.2.1 旅客數預測

考量一般公共汽車運輸較常運用的班表改版均為短期需求預測，故建議採用時間序列預測方法（Time Series Forecasting Method），而非諸如旅次產生、吸引、分配與指派的運輸規劃方式，考量後續的應用會逐路線調整班表，因此在時間序列的分析也設計成逐路線分析預測，此外，票證資料中蘊含了票種，故可考慮逐票種執行預測。

綜合上述需求，可設計如圖 4-31 之旅客數預測儀表板，一般時間序列包含基礎量(level)、趨勢(trend)、週期性(periodic pattern)等元素，以本案例而言，以週循環的週期性較為明顯，使用者除了可在右側挑選欲分析路線外，亦可調整右上角的「起始日」來決定欲使用近 n 個月的資料執行時間序列預測分析，在使用者操作控制項後，左側圖表會即時更新分析結果。

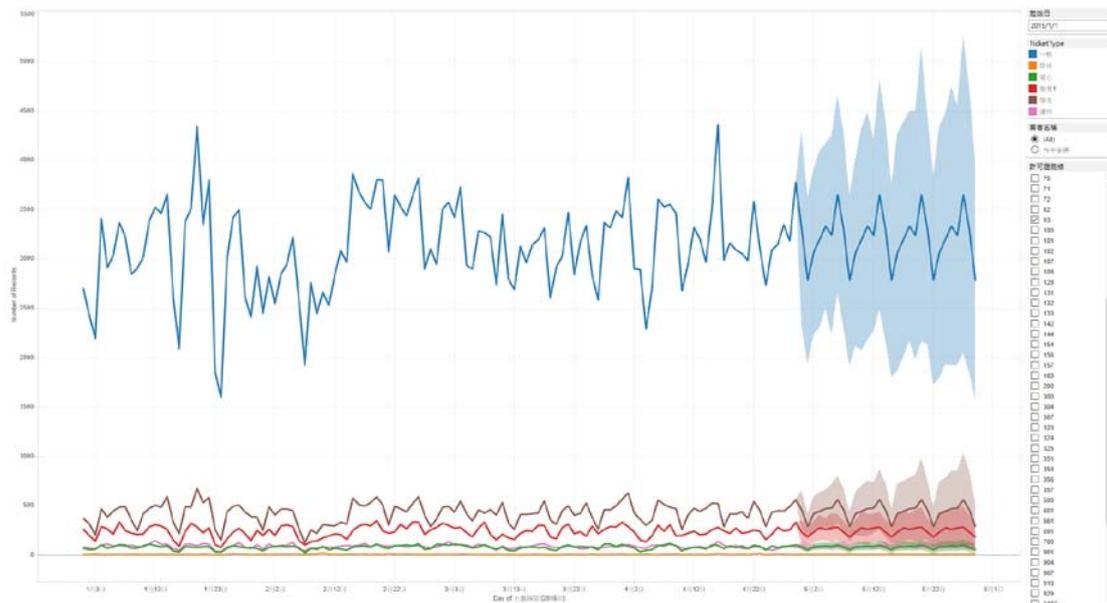


圖 4-31 各路線旅客數量時間序列預測儀表板

考量後續做班表調整時需要分時段的運量，故求得全路線的運量之後，必須依歷史紀錄比例（如圖 4-32 上方直條圖）分配到各時段，甚至必須根據類似圖 4-6 的比例分配到各起迄。

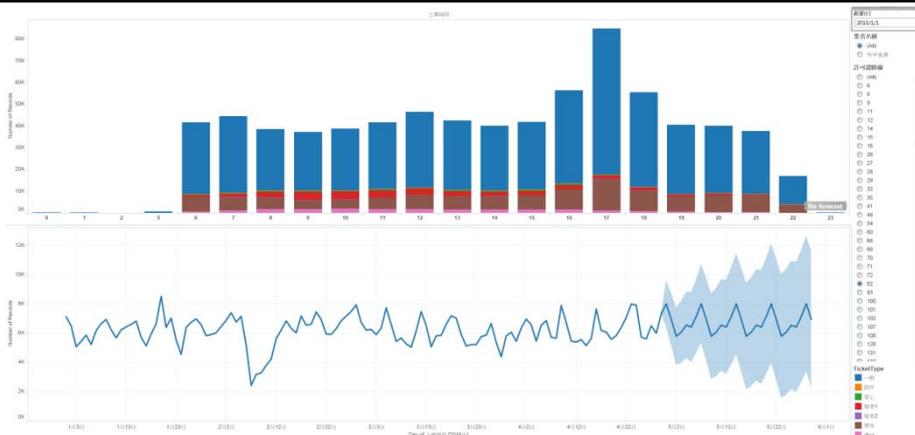


圖 4-32 依時段/票種之旅客數量時間序列預測儀表板

除了上述依比例的分配方式外，亦可透過點選圖 4-32 上方直條圖的區塊，由儀表板過濾資料針對使用者選取之時段與票種分別進行時間序列預測分析，以圖 4-33 為例，係點選上午七點時段一般票種所得之預測結果，圖 4-34 則為點選下午五點時段學生票種所得到預測結果，兩相比較即可知道兩者的週期樣態明顯不同。



圖 4-33 點選上午七點時段一般票種之時間序列預測結果

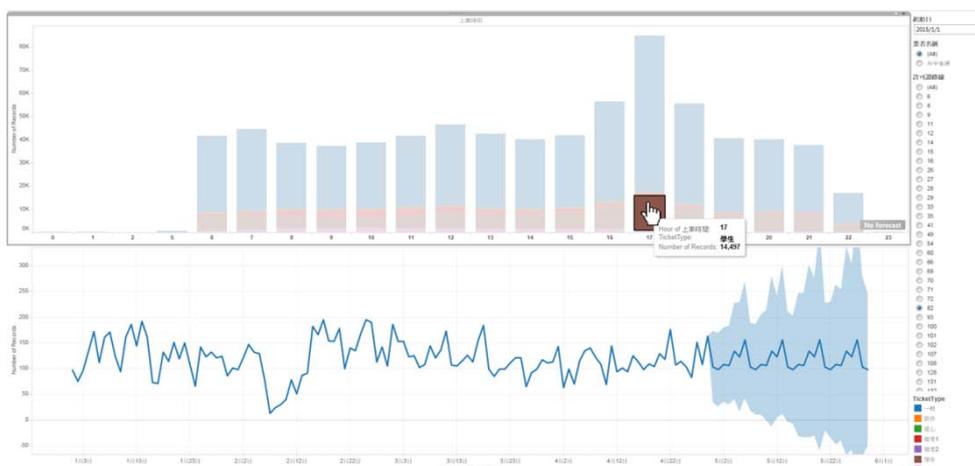


圖 4-34 點選下午五點時段學生票種之時間序列預測結果

### 4.3.2.2 最適班表規劃

公車供給是一種有限資源，加上運輸的不可儲存性，理應適當得在高強度需求時段提供服務，因此本研究將圖 4-25（重疊區間決策輔助工具）中全年各路線各時段各站牌的平均車上乘載人數執行集群分析（Cluster Analysis），用以作為各時段發車班次數之參考。

本研究使用 R 語言環境搭配 NbClust 套件，計算各種分群數所對應的平方誤差和關係如圖 4-35。

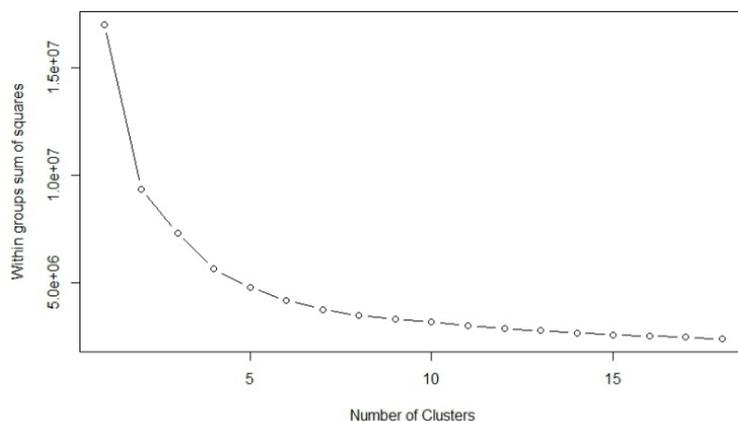


圖 4-35 分群數與平方誤差和關係圖

從圖 4-35 中可發現 2 群/3 群/4 群的平方誤差和下降較快，到了 5 群之後就緩慢下降，故在指定群數的前提下利用 K-means 方法將時空予以分群，最後再以日曆搭配時段的方式呈現分群結果如圖 4-36、圖 4-37 與圖 4-38。

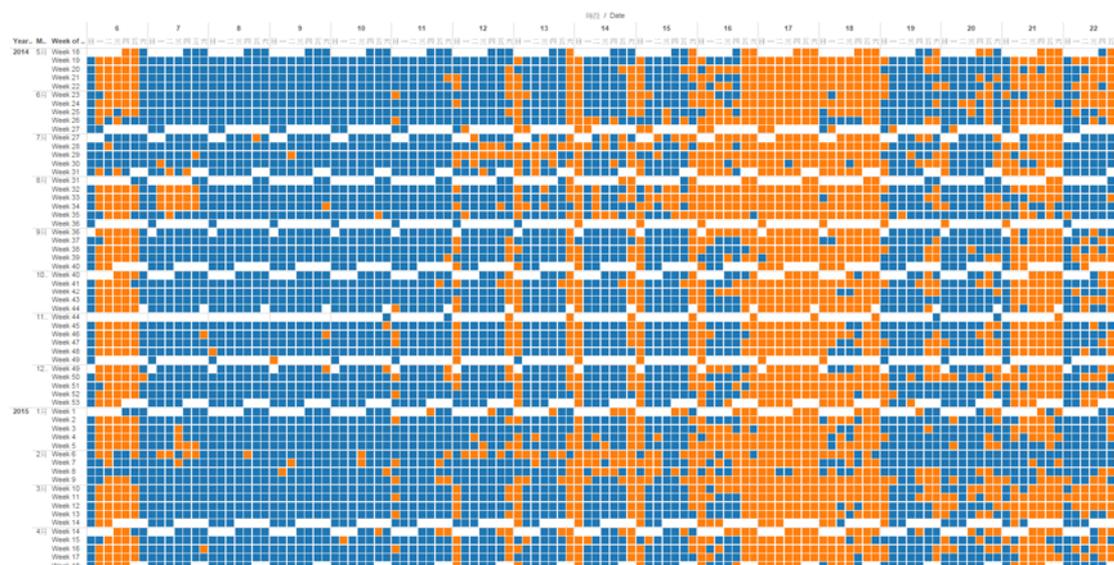


圖 4-36 55 路全年各時段車上乘載人數集群結果圖（2 群）

若以分 2 群的圖 4-36 來看，上午六點、下午五/六點及九點是需求較高的族群，而且上午六點是週間/週末不同群，但下午五/六點則不分週間週末均屬高需求。

一旦將分群數提高到 3 群如圖 4-37，上午六點與下午五點為一群，而下午六點已後為另外一群，其餘時段則為另一群，但下午 1~5 點則較為混雜，但由於二/七/八 3 個月分較為特殊，研判與學生寒暑假有關聯。至於將分群數由 3 提高到 4 群如圖 4-38 時，與圖 4-37 差異甚小。



圖 4-37 55 路全年各時段車上乘載人數集群結果圖 (3 群)

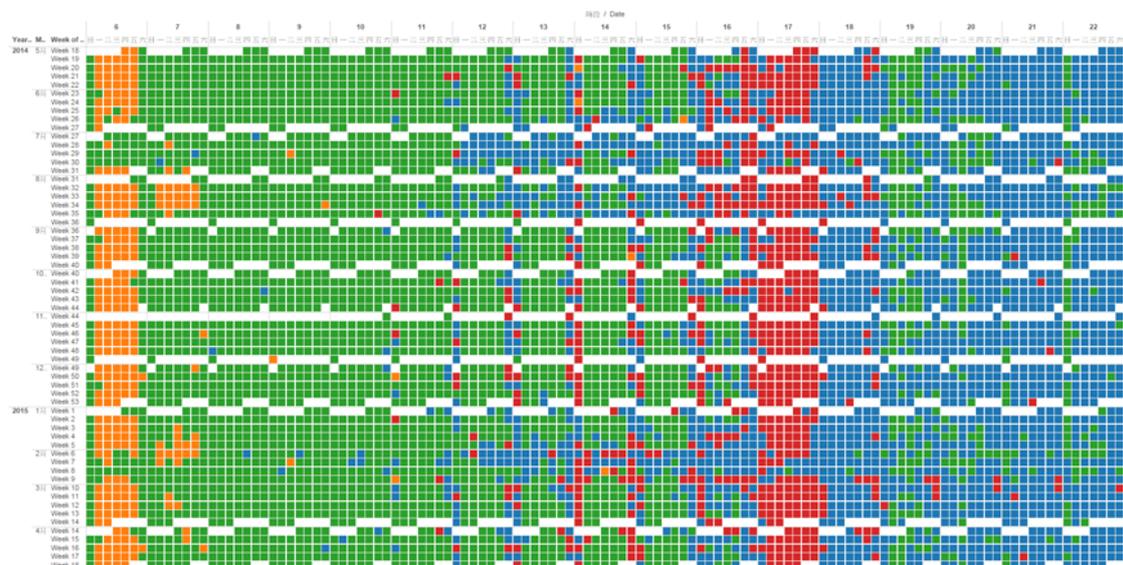


圖 4-38 55 路全年各時段車上乘載人數集群結果圖 (4 群)

圖 4-36~圖 4-38 之集群分析成果雖可作為不同日期、時段發車班次數多寡之參考，但上述資訊作為班表決策輔助資訊恐有不足，包括確切的理

想發車數量及理想發車時間均沒有辦法從圖 4-36~圖 4-38 的預測結果找到答案，因此本研究進一步發展班表調整績效模擬模式來補其不足。

考量公共汽車客運均有營運計畫，故全日發車總數可視為模式已知，而旅次起迄需求就短期而言也可視為已知（可直接使用歷史票證紀錄或時間序列預測之結果），由模式根據總搭乘人次或總旅客搭乘區間數等兩項績效指標來決定每個班次之最適發車時間。

當模式僅考慮搭乘總人次時，意即模式會選擇理想發車時間來讓每個班次均能載到最多人次之旅客，若以總旅客搭乘區間數做為績效指標，則模式除了考量人次外，亦會把旅客搭乘區間數（距離）納入考量，也就是客座公里最大化之概念。

整個模擬工具外觀如圖 4-39，除須輸入前述之票證紀錄與單向發車總數外，尚須設定「車廂容量」與「等待時間」兩個參數，前者係用來控制服務水準，避免模擬時發生單一車輛載運過多旅客之不合理行為，此數值設定越低，就需要開行更多班次才能疏運所有旅客，但服務水準也相對提高。至於「等待時間」係指每個班次可以載運該行車時空軌跡前/後多少時間範圍之旅客。

模擬模式會依序列出績效指標最高之理想發車時間，這些發車時間欲發布之前可考慮先以 5 分鐘或 10 分鐘進整方符合實務發車時間非破數的需求。

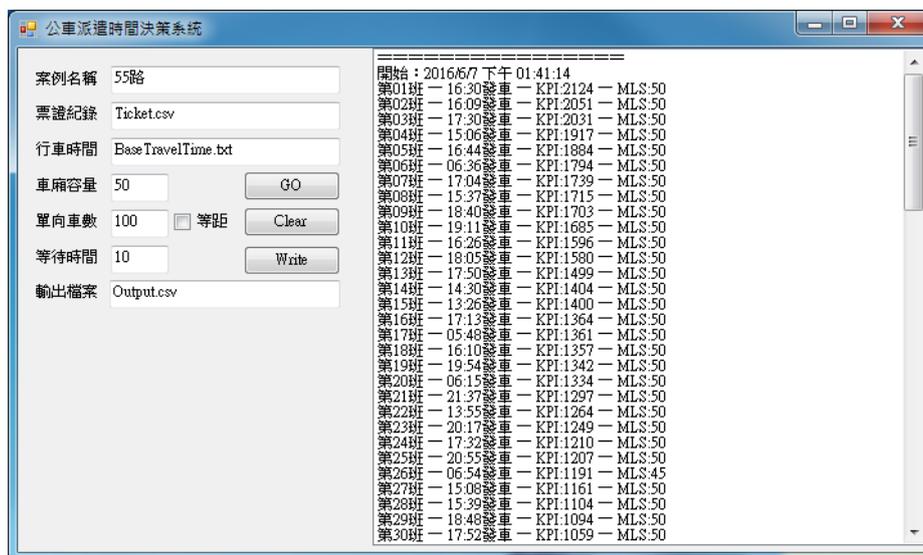


圖 4-39 班表調整績效模擬模工具外觀

圖 4-40 為均勻班距下每班次之乘載狀況調整結果，可發現若將班次供給用平均的方式提供，事實上許多班次全程均是低乘載的綠線，僅有少數

班次中的少數區段會有紅色，但因為模式設定了車廂容量上限，故有 1049 筆電子票證旅客尚無法使用服務。

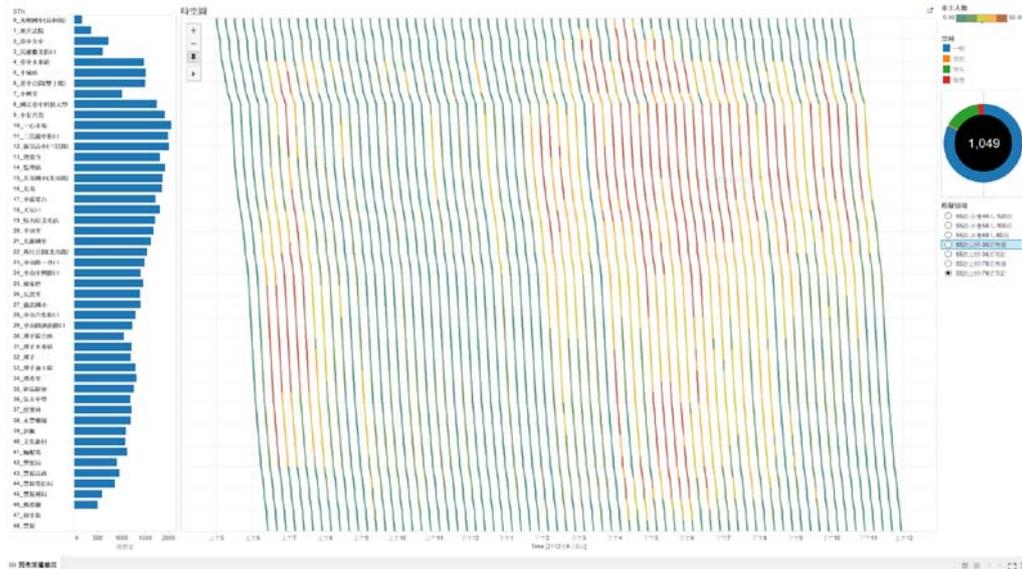


圖 4-40 均勻班距下各班次各區間車上乘載人數調整結果

一旦開啟最佳發車班距的功能後，就可從圖 4-41 觀察到同樣的發車數在全日內並不均勻，但相較於圖 4-40，在相同的旅運需求條件下，所有車輛的乘載均變高了，代表資源的有效運用，惟須助益的是這種高度以需求導向來決定發車時段的模式可能會在離峰時段產生超過營運計畫書之最長班距限制，這部分必須要作業人員手動補足調整。

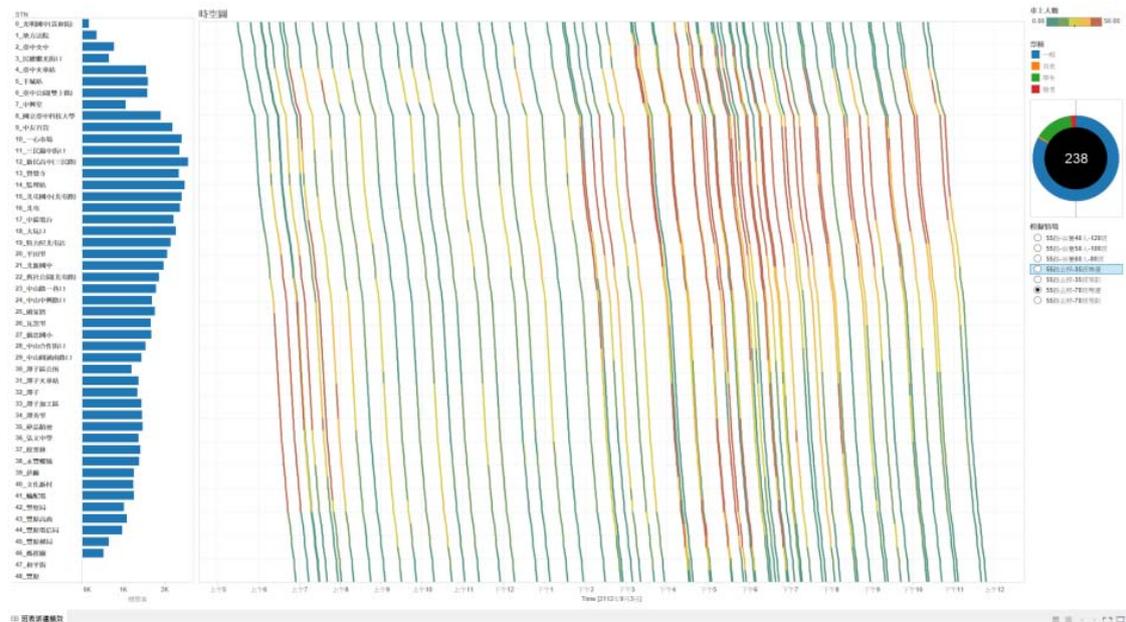


圖 4-41 模擬最佳發車時間各班次各區間車上乘載人數結果

## 4.4 信賴面向分析

### 4.4.1 車機傳輸漏失與延誤

#### 4.4.1.1 分析目的與構想

車機在 APTS 系統中扮演相當重要角色，其穩定性自然成為關注議題，整個公車動態雖由眾多資通訊設備所組成，資料傳輸的過程至少會經「車機→電信業者→伺服器→資料庫」等過程，每個程序均可能導致漏失或延誤，但就應用面之需求而言，並不會探討每個子元件或流程的各別可靠度，而是將所有資通訊設備視為一個完整系統分析最終產出，提供管理者作為後續汰換車機的參考數據之一。

本研究蒐集了連續兩年的公車動態定時紀錄共約 30 億筆，各月分資料筆數整理如表 4.2。

表 4.2 本研究蒐集近二年臺中地區公車動態定時回傳筆數表

	2014 年(筆數)	2015 年(筆數)	2016 年(筆數)
01 月	—	120,736,338	140,923,064
02 月	—	98,197,619	126,445,859
03 月	—	111,706,500	139,918,460
04 月	—	105,377,187	135,243,033
05 月	116,270,155	121,508,421	—
06 月	114,097,875	116,472,664	—
07 月	117,896,146	128,279,627	—
08 月	126,589,481	141,548,276	—
09 月	123,088,080	137,770,264	—
10 月	127,585,266	146,314,291	—
11 月	120,281,538	143,245,270	—
12 月	127,538,310	143,949,280	—

資料來源：本研究整理

#### 4.4.1.2 傳輸漏失分析

回傳失敗（漏失）是呈現公車動態系統可靠度與穩定度的重要指標，漏失頻率越低代表可靠度越高，反之若漏失比率偏高，將衝擊即時資料的正確性與持續性。

理論上公車動態定時回傳的頻率為每分鐘三次，意即每 20 秒一次如圖 4-42(a)所示，但若蒐到資料間隔時間達 40 秒如圖 4-42(b)，即代表漏失乙筆，同理，若間隔達 60 秒圖 4-42(c)即代表漏失 2 筆，餘類推。換言之，須把公車動態 A1 定時資料根據每車輛逐一檢測，先經歸納為同一趟次並依時間排序後計算第 N 筆與第 N+1 筆之時間差值，正常情況此差值應為 20 秒左右，若數值明顯大於 20 秒，每差距 20 秒可視為漏傳乙次。

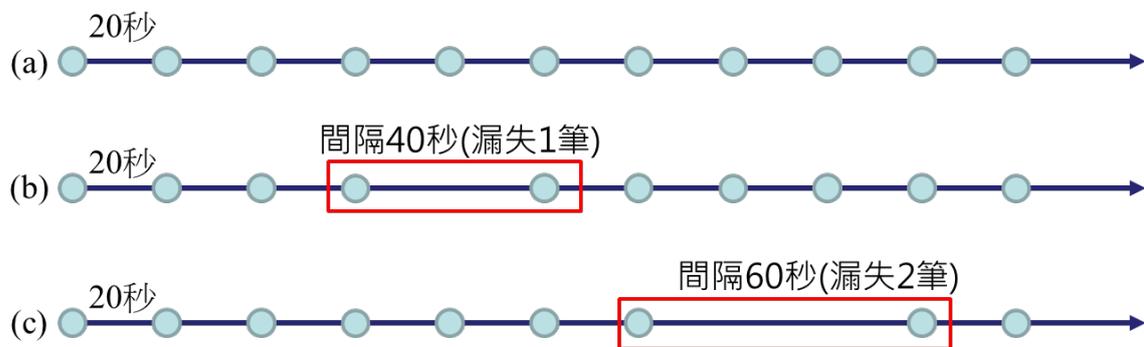


圖 4-42 資料傳輸漏失分析示意圖

整體而言，近二年 30 億筆公車動態定時資料的漏失分析結果如表 4.3，間隔秒數在 0~19 者屬於相對理想優質無漏失的比例佔約 68%，而間隔超過 20 秒但小於 39 秒的資料近 32%，落在此區塊的資料雖不如前類資料理想，但仍尚未漏「足」一次（即 40 秒），換言之，若以較為寬鬆的標準判定，99.55047%的資料均可視為無漏失資料，額外的 0.5%則為漏失資料。需要進一步分析這些漏失情況是分佈於何處（年分/月分/星期/客運業者/路線/車輛/通訊業者..等）。

後續欲比較各種因子或情境的資料漏失狀況時，因為各因子的母體樣本數不可能均一，故表 4.3 中「資料筆數」欄位應須先標準化，而「累計比例」就是一種適當的標準化方式，基於漏失數是越少越好，因此「累計比例」相較於各種漏失數佔母體比例提供更直覺的判讀價值，例如表 4.3 中的 99.84871%可解讀為漏失數量在 3 次以下佔全母體的比例，故後續欲進行各種因此分析時，均將資料採「累計比例」方式呈現。

表 4.3 近二年臺中地區公車動態節定時資料漏失統計表

漏失數	間隔秒數	資料筆數	佔母體比例	累計比例
0	0~19	1999583596	67.88249%	67.88249%
1	20~39	932829187	31.66798%	99.55047%
2	40~59	6020935	0.20440%	99.75487%
3	60~79	2764076	0.09384%	99.84871%
4	80~99	946107	0.03212%	99.88082%
5	100~119	1047126	0.03555%	99.91637%
>=6	>120	2463383	0.08363%	100.00000%

資料來源：本研究整理

本研究設計之探勘儀表板右側計有兩種分析選項如圖 4-43，右側供使用者切換分析維度，以利分析人員快速切換不同的觀測面向，用以檢視各種因素對車機可靠度是否敏感，這些分析維度包括客運業者、車牌、路線、年分、月分、星期、時段、電信業者、通訊型態（2G/3G）等，未來可根據新的資料欄位取得而擴充。

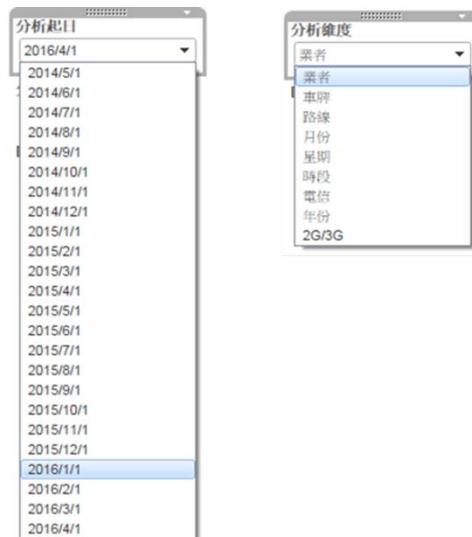


圖 4-43 車機可靠度儀表板右上方分析選項

圖 4-43 左側則是調整分析起始時間，本研究蒐集了兩年完整資料，在應用時若欲分析整體長期趨勢變化時，可能須將 2014 年 5 月起的所有資料納入分析，但若就輔導業者提出潛在車機故障清單的應用而言，可能就必須將分析起始日設定於今年(2016)一月或四月，如此分析結果方是近期車機可靠度表現，避免已修復或已改善的車機影響到分析結果。

以營業者分析為例，分析起日設定為 2014/5/1 時與 2016/1/1 就有明顯不同如圖 4-44 意即在去年以前淺藍色業者曾有段時間有漏失數較高之狀況，但若僅觀測今年之樣本，則各家業者之數值相當一致。

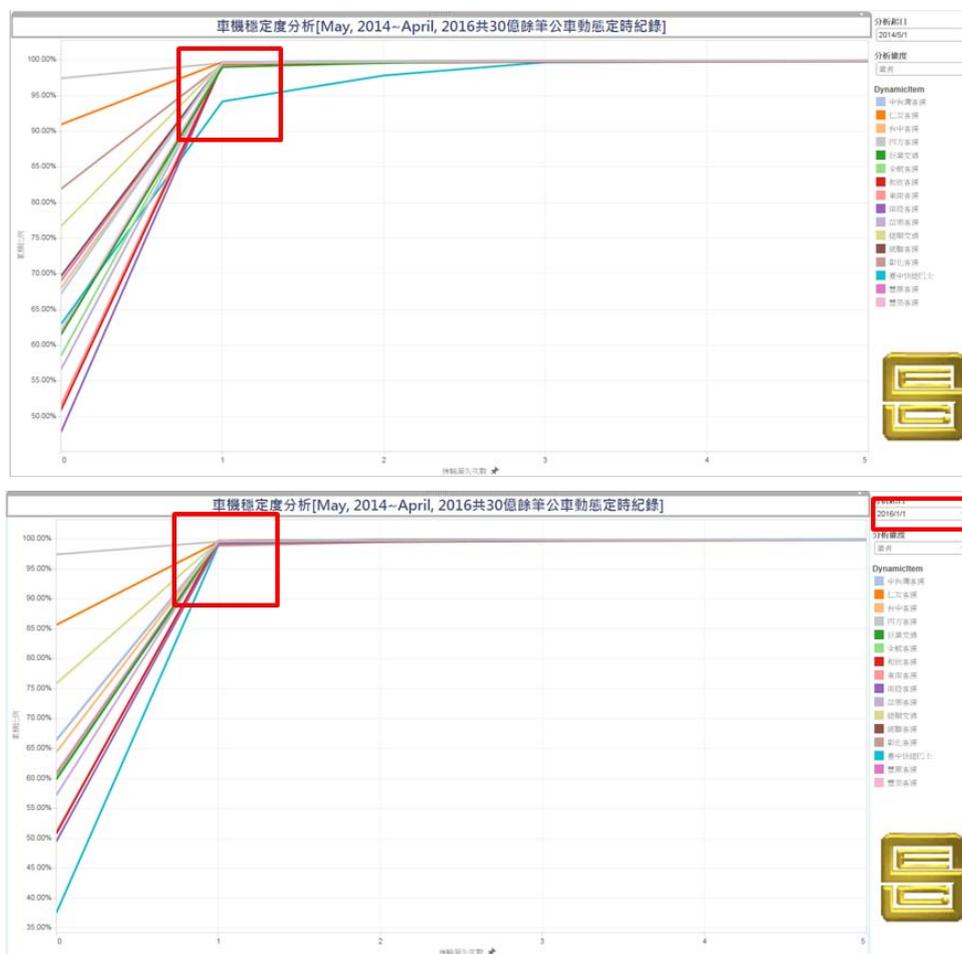


圖 4-44 不同起始日之各業者漏失累計比例比較圖

從圖 4-45 可得到無論從年分、月分觀點而言，這兩個因素均對傳輸漏失沒有顯著的影響。

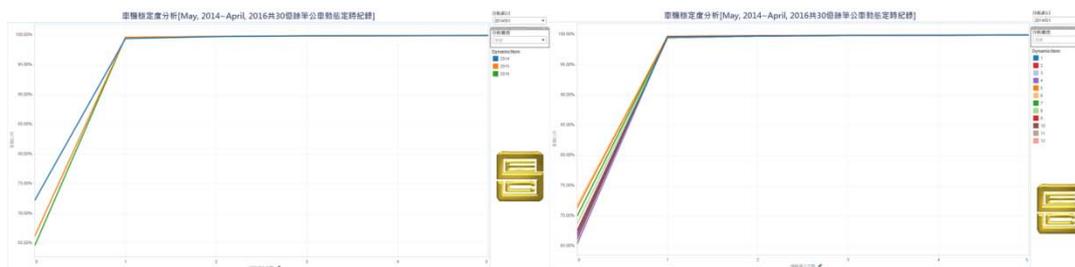


圖 4-45 傳輸漏失之年分/月分因素比較圖

同理，從圖 4-46 亦可得知星期、時段這兩個因素均對傳輸漏失沒有顯著的影響。

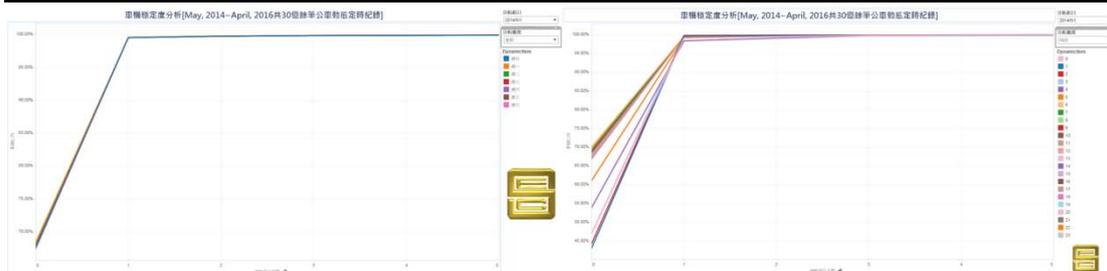


圖 4-46 傳輸漏失之星期/時段比較圖

就如同一般先驗知識，當分析維度切至「車牌」如圖 4-47 時，就能讓部分績效值較差的車牌號碼現形，這些車牌將可提供給業者參考。

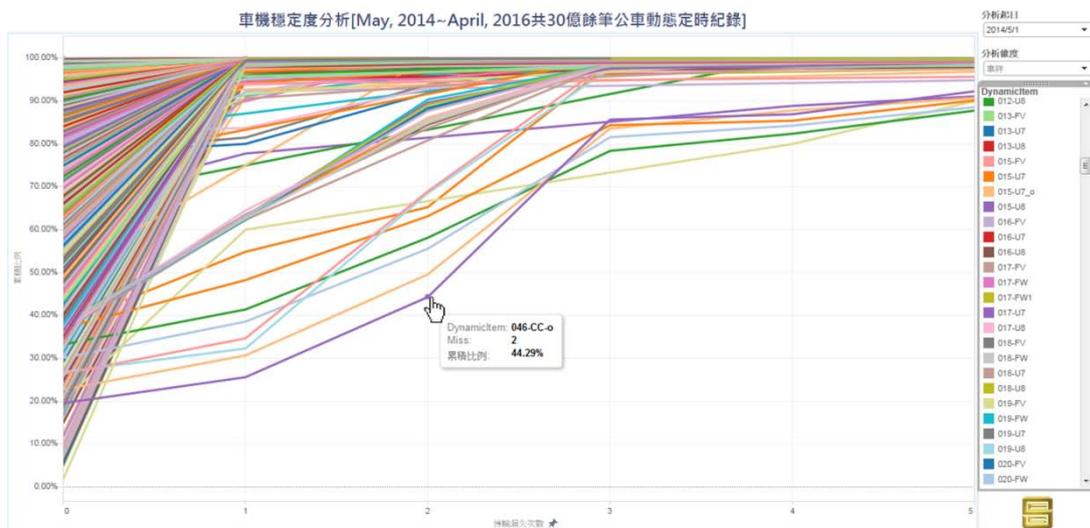


圖 4-47 傳輸漏失之車牌比較圖

同理，當分析維度切至「路線」如圖 4-48 時也能提供明顯較為不同的異常路線資訊供業者參考。

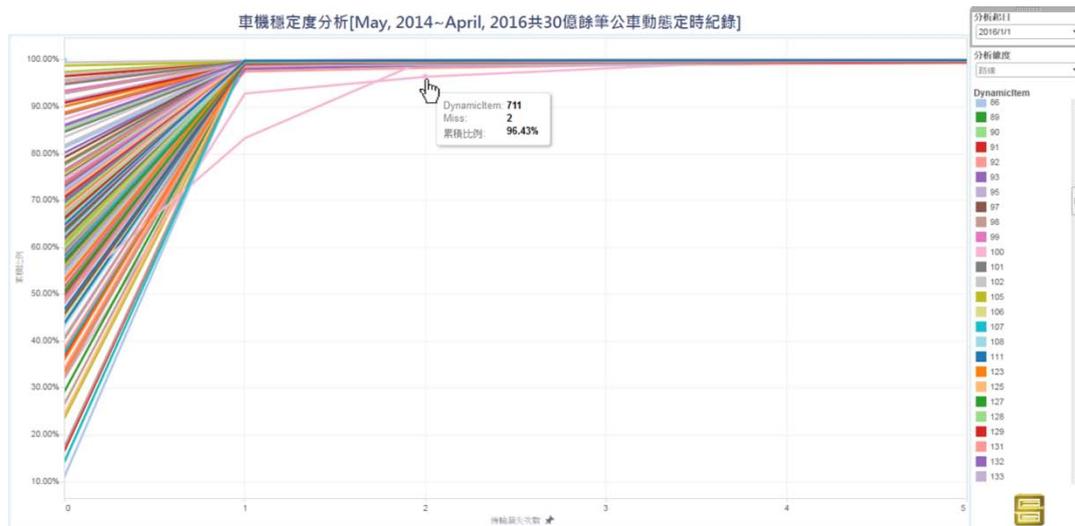


圖 4-48 傳輸漏失之路線比較圖

除了公車動態資訊外，本研究也蒐集了客運公司自主管理的車機 SIM 卡通訊業者資訊，可進一步逐路線分析不同業者之傳輸漏失表現，例如圖 4-49 呈現兩通訊業者在 12 路公車路線上幾乎沒有顯著差異。

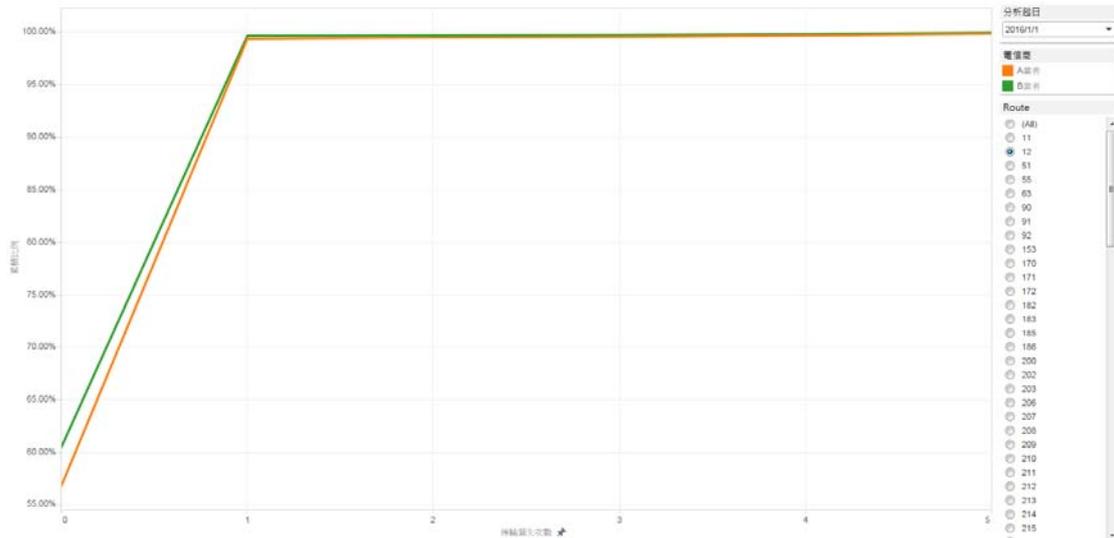


圖 4-49 傳輸漏失之電信商比較圖(無差異案例)

但同樣的分析當切換至路線 253 時，就可以發現兩家業者之表現顯著不同如圖 4-50。

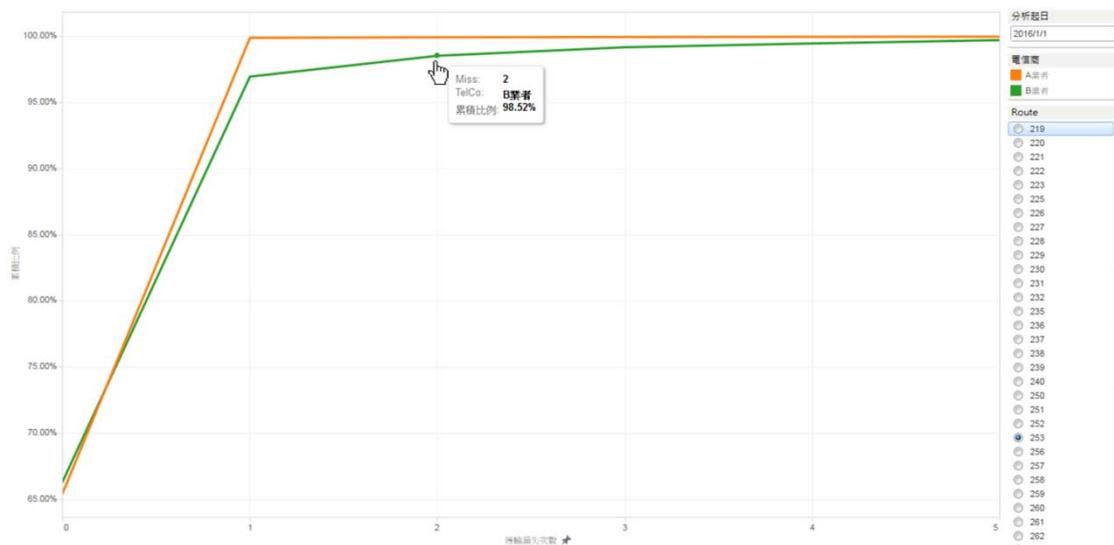


圖 4-50 傳輸漏失之車牌比較圖(有差異案例)

#### 4.4.1.3 傳輸延誤分析

除傳輸漏失之外，傳輸延誤亦是另一分析重點，傳輸延誤係比較車機 GPS 時間與資料寫入系統時間之差，透過大數量的車機回傳時間資料進行平均值計算，藉以得知兩者之間的時間差異，此分析標的亦可進行如圖 4-44

～圖 4-50 之各種分析，以圖 4-51 為例即為不同客運業者傳輸延遲累計比例比較圖。

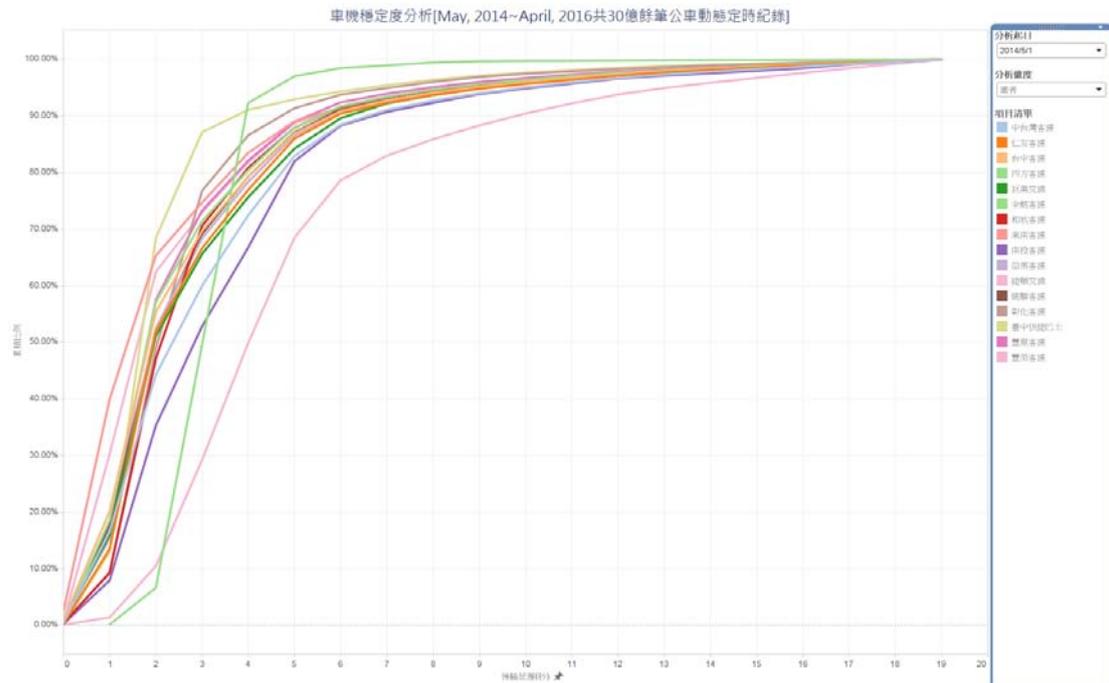


圖 4-51 不同客運業者傳輸延遲比較圖

受限於篇幅，本節將不重複類似圖 4-44～圖 4-50 之各種分析成果，但額外設計儀表板如圖 4-52，右上方可調整傳輸延遲門檻值，左右兩張圖地圖分別代表兩家通信業者，藍色區塊代表平均延遲低於門檻值之行政區（里），橘色區塊代表平均傳輸延遲高於門檻值之行政區。

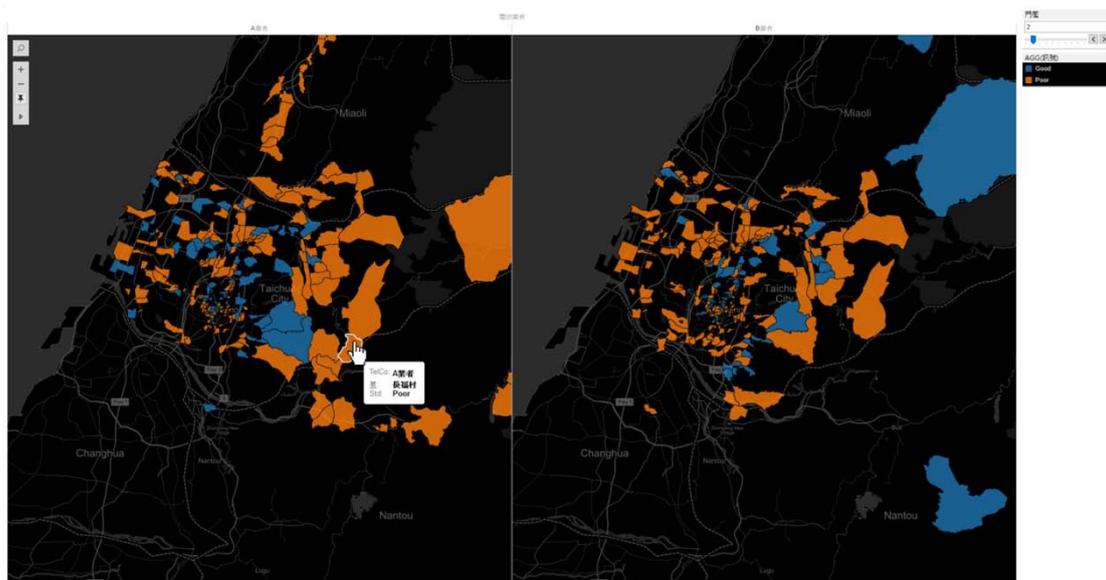


圖 4-52 不同電信商平均傳輸延遲比較圖

### 4.4.2 數據異常樣態

#### 4.4.2.1 分析目的

圖 4-53 係一般數據分析前執行「資料清理」與「資料融合」之程序，清理與融合過程除相當耗時費工外，在「清理」的瞬間也等同「捨棄」許多資料，甚至把 Dirty Data 背後的價值也一併放棄。

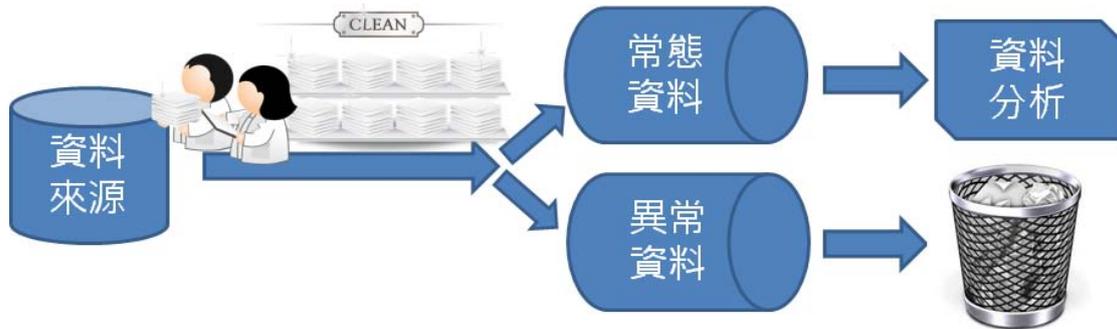


圖 4-53 一般資料清理程序

在大多數的情況下，異常樣態的資料背後一定有其背景與原因，追溯其原因或者分佈亦是資料採礦的價值之一，特別是現代化的資料蒐集機制幾乎都以資通儀器設備來完成，異常的資料樣態極可能代表設備故障、設定配置錯誤等狀況，故若調整如圖 4-54 之流程，可透過異常樣態的分析可追溯/釐清異常源頭，進而採取設備更換/設定調整等作為，將有助於提昇資料蒐集品質。

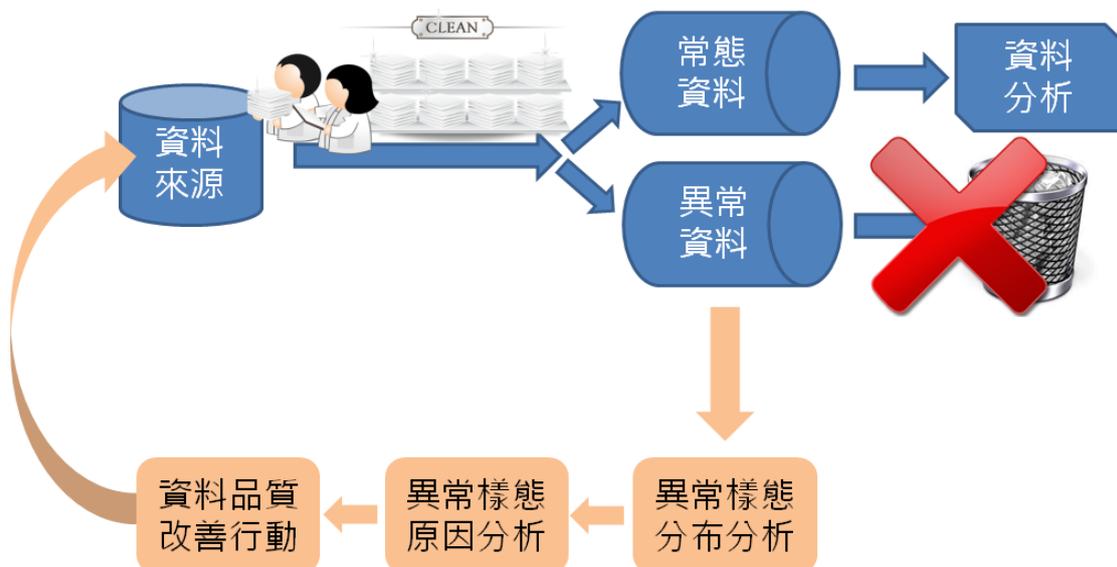


圖 4-54 利用異常樣態分析提昇資料品質流程示意圖

綜合上述，本專案挑選公共運輸中最重要之兩項資料集（公車動態與電子票證）作為分析範例。

#### 4.4.2.2 未成對刷卡之電子票證紀錄

本研究取得某電子票證公司 2016 年 1 月~4 月臺中地區公車交易資約 5010 萬筆，其中約有 4.2%（210 萬筆）未成對如圖 4-55，意即該旅次下車未刷卡，故原始資料提供者以「1900-01-01」與「0」等預設值填入資料表。但在資料解讀上不宜根據資料表欄位名稱為「下車時間」認定是下車未刷卡導致旅次未成對，因為只要「上車刷卡」與「下車刷卡」有任何一次因機器失誤或旅客疏忽而導致漏失，後端接蒐資料時就會將第一次刷卡視為「上車時間」，但實際上該刷卡行為可能在旅客下車時產生。

票證公司	客運公車代碼	卡號	票種編號	票種代碼	路線公路編號	司機編號	車號	上車時間	上車站序	上車計費站序	下車時間	下車站序	下車計費站序	
1	EASYCARD	033041	0Q71u5wqJ2j6	1	1	75	13376	200-FS	2016-03-08 10:05:00	54	54	1900-01-01 00:00:00	0	0
2	EASYCARD	033041	7TAcK2vAalZ9	1	1	73	13782	199-FS	2016-03-14 17:26:00	12	12	1900-01-01 00:00:00	0	0
3	EASYCARD	033041	2wUVW1KTT1d2	2	2	73	13782	199-FS	2016-03-14 13:56:00	10	10	1900-01-01 00:00:00	0	0
4	EASYCARD	033041	q85Q7.lLgKZ9	1	1	73	13782	199-FS	2016-03-20 19:06:00	3	3	1900-01-01 00:00:00	0	0
5	EASYCARD	033041	gseOzGwS65e/	1	1	73	13782	199-FS	2016-03-20 13:41:00	5	5	1900-01-01 00:00:00	0	0
6	EASYCARD	033041	3nPZg7pVVC	1	1	73	13782	199-FS	2016-03-20 08:03:00	18	18	1900-01-01 00:00:00	0	0
7	EASYCARD	033041	Yd8VmaWIXV86	1	1	73	13782	199-FS	2016-03-20 18:59:00	1	1	1900-01-01 00:00:00	0	0
8	EASYCARD	033041	DoQmqr2n.eI7	6	6	73	13782	199-FS	2016-03-20 12:28:00	13	13	1900-01-01 00:00:00	0	0
9	EASYCARD	033036	FAmBc3Bfo8H3	1	1	17	52550	198-U8	2016-03-09 08:42:00	47	94	1900-01-01 00:00:00	0	0
10	EASYCARD	033036	uQYc9Z8Bqa9	2	2	17	52550	198-U8	2016-03-09 21:51:00	22	44	1900-01-01 00:00:00	0	0
11	EASYCARD	033036	WJNEVHIdThD	6	6	17	52550	198-U8	2016-03-09 08:26:00	59	118	1900-01-01 00:00:00	0	0
12	EASYCARD	033036	jaJHHW3Qte09	6	6	17	52550	198-U8	2016-03-09 14:47:00	22	44	1900-01-01 00:00:00	0	0
13	EASYCARD	033036	Btu4krvmo5	1	1	17	52550	198-U8	2016-03-10 07:09:00	36	72	1900-01-01 00:00:00	0	0
14	EASYCARD	033036	Gal2JNFyqQL7	1	1	17	52550	197-U8	2016-03-31 18:36:00	26	52	1900-01-01 00:00:00	0	0
15	EASYCARD	033036	2TR.F5lmhQK9	1	1	18	53041	197-U8	2016-03-04 06:29:00	0	0	1900-01-01 00:00:00	0	0

圖 4-55 未成對票證紀錄範例圖

本研究擷取 210 萬筆未成對之票證紀錄並以多維度方式(包括原始欄位及衍生欄位)展開，同時在 BI 軟體中分別以兩組 ComboBox 並設計儀表面板如圖 4-56，透過儀表面板中兩個 ComboBox 的控制項變化即可任意切換 100 種（ $10 \times 10 = 100$ ）異常資料樣態分佈分析。

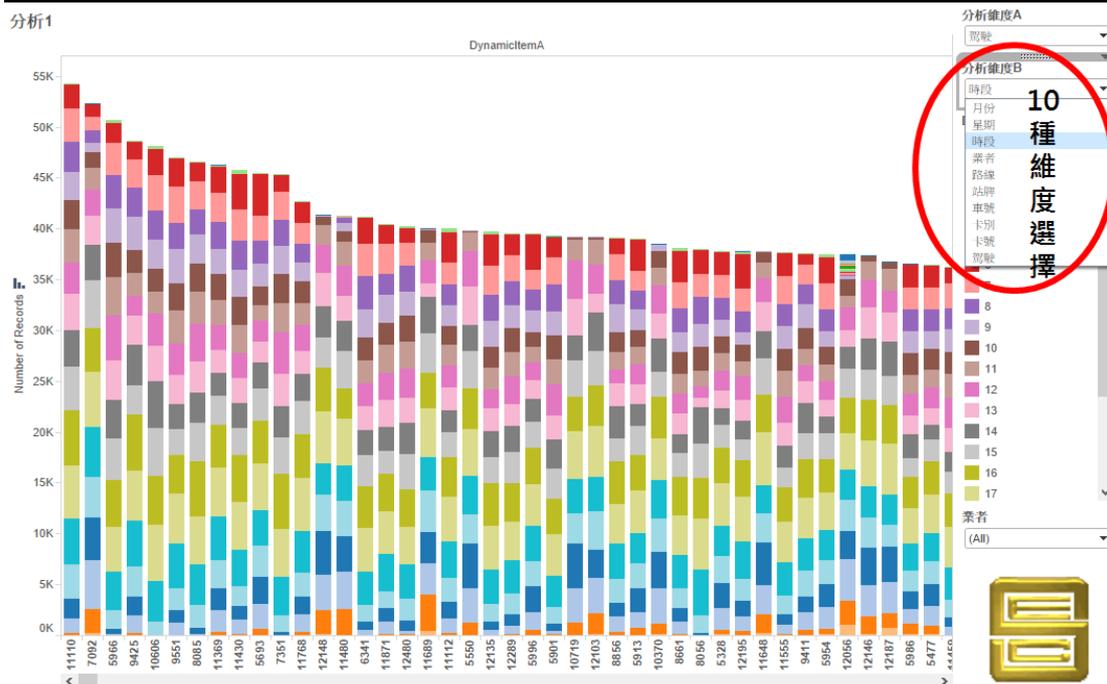


圖 4-56 未成對異常刷卡樣態分析-多維度分析儀表板設計圖

透過兩個 ComboBox 控制項的調整過程中即可得到若干訊息，例如從圖 4-57 可發現未成對的交易紀錄集中在某 R 業者，同時一般票的比例約是學生票的三倍多等。

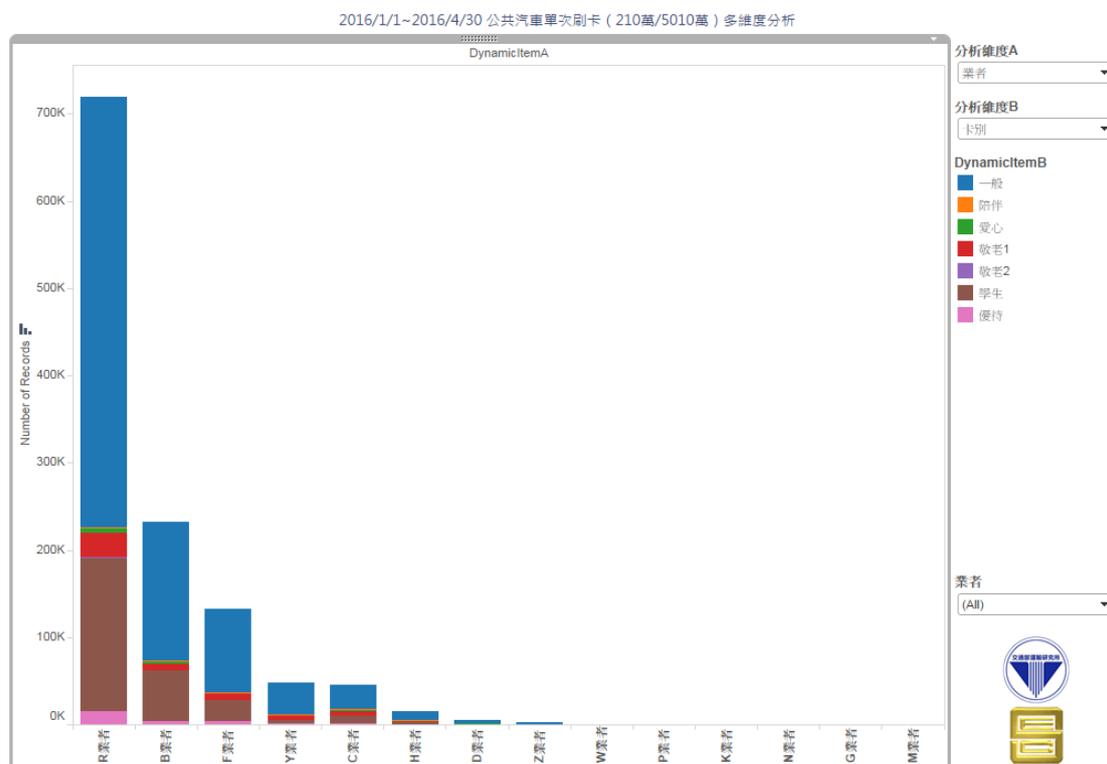


圖 4-57 未成對異常刷卡樣態分析-業者名稱與票種分佈圖

從圖 4-58 則進一步挖掘異常次數集中在特定的卡號上，且次數極高，同時這些異常卡號所搭乘路線名稱的比例也在右側的路線號碼，以滑鼠 hover 即可取得該卡號搭乘之路線資訊，利用此方式可快速找出異常樣態的分佈方式，近而追蹤問題來源。

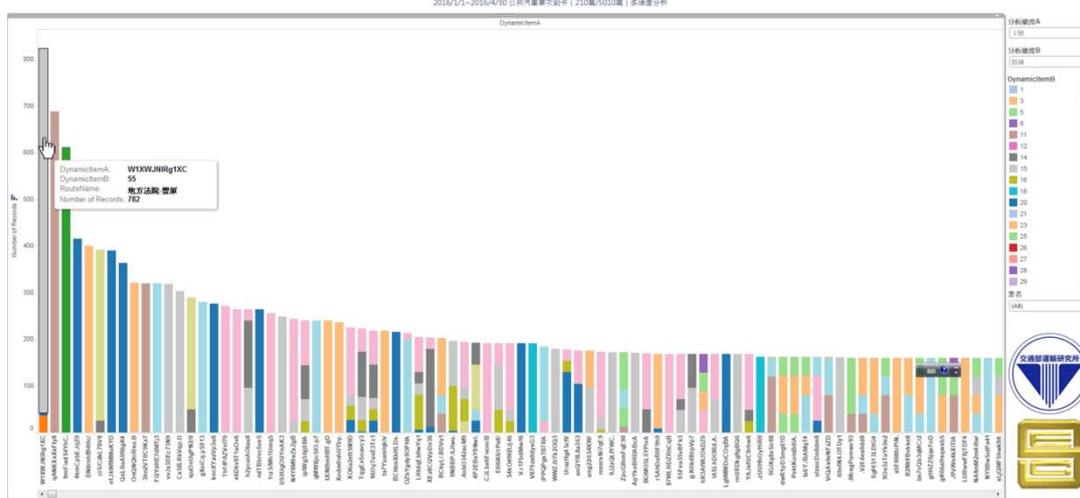


圖 4-58 未成對異常刷卡樣態分析-卡號與路分佈圖

同理，圖 4-59 係由車牌號碼與路線觀點分析，可快速從 3600 餘輛車牌中掌握高頻率重現之車牌，以利後續檢核刷卡機是否異常。

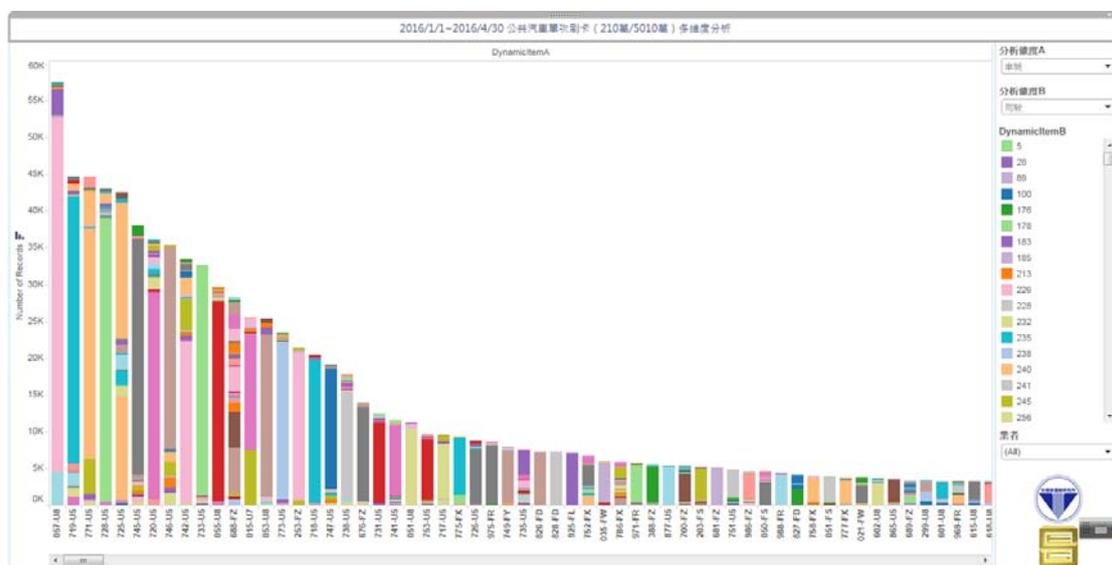


圖 4-59 未成對異常刷卡樣態分析-車號與路線分佈圖

此外，如圖 4-60 的路線、司機代碼為切入的分析也有助於瞭解異常資料之分佈。



圖 4-60 未成對異常刷卡樣態分析-路線與司機代碼分佈圖

最後，本研究也分析這分未成對刷卡紀錄所留下之唯一刷卡站點位置如圖 4-61，從中可知未成對刷卡雖然遍佈整個大臺中地區，但其中某些路線的某些站牌發生頻率較為明顯。

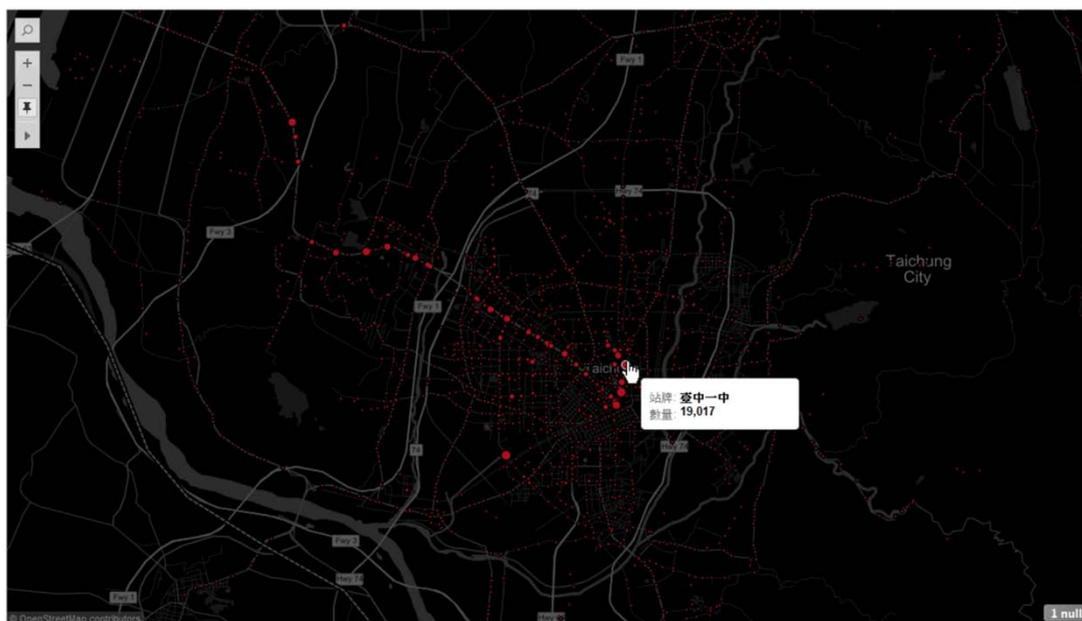


圖 4-61 未成對異常刷卡樣態分析-站牌分佈圖

以上各儀表板均連結資料源，只要更新資料即可得到新的分析報表，可在圖 4-54 的反覆修正回饋中持續監控異常資料樣態。

#### 4.4.2.3 公車動態與電子票證比對

第 4.4.2.2 節之異常樣態比對是指單一資料集內部的合理性比對，但部分異常樣態必須跨資料集比對方能發現異常，本研究以公車運輸系統中最

大的兩個資料集合（公車動態與電子票證）作為跨資料集異常樣態比對案例。

本研究從兩資料集中擷取不重複的「日期」、「車牌」與「時段」資料先予以聯集處理，接著分別對應兩個原始資料集，作為判斷資料漏失的參考。

圖 4-62 係以 2016 年 4 月 29 日為例，最左側係公車動態與電子票證聯集後的日期、車牌與時段資訊，但從第 2921~2933 列就可發現存在公車動態資訊但卻沒有電子票證紀錄，以異常資訊除電子票證漏失外，亦不排除是該時段並無任何旅客搭乘所致，即使是後者，在資料分析上亦有其價值。

此外如圖 4-62 中 2917 列顯示 241-U8 車輛在七點有電子票證交易紀錄但卻不見公車動態資料，此類異常樣態除公車動態漏失外，亦有可能是校時因素所導致的時間差異所致。

但若如圖 4-62 中 2936~2941 列，車號 246U8 在 6/7 兩個時段均有資訊，但是 8 時之後就只剩下票證資訊而沒有定點的公車動態，公車動態漏失的可能性就大幅提昇。

1	日期	聯集車牌	聯集時段	票證路線	票證路線描述	公車動態路線	公車動態路線描述
2911	2016/4/29	237U5	14	3	東山高中-台中高工	3	東山高中->高鐵臺中站
2912	2016/4/29	237U5	15	3	東山高中-台中高工	3	東山高中->高鐵臺中站
2913	2016/4/29	237U5	16	3	東山高中-台中高工	3	東山高中->高鐵臺中站
2914	2016/4/29	237U5	17	3	東山高中-台中高工	3	東山高中->高鐵臺中站
2915	2016/4/29	237U5	18	3	東山高中-台中高工	3	東山高中->高鐵臺中站
2916	2016/4/29	237U5	19	3	東山高中-台中高工	3	東山高中->高鐵臺中站
2917	2016/4/29	241U8	7	157	外埔區公所-向上惠文路口	NULL	NULL
2918	2016/4/29	241U8	8	157	外埔區公所-向上惠文路口	157	向上惠文路口->大甲區公所
2919	2016/4/29	241U8	9	157	外埔區公所-向上惠文路口	157	向上惠文路口->大甲區公所
2920	2016/4/29	241U8	10	157	外埔區公所-向上惠文路口	157	向上惠文路口->大甲區公所
2921	2016/4/29	245XH	5	NULL	NULL	98	神岡庄尾->普濟寺
2922	2016/4/29	245XH	6	NULL	NULL	98	神岡庄尾->普濟寺
2923	2016/4/29	245XH	7	NULL	NULL	98	神岡庄尾->普濟寺
2924	2016/4/29	245XH	8	NULL	NULL	98	神岡庄尾->普濟寺
2925	2016/4/29	245XH	10	NULL	NULL	98	神岡庄尾->普濟寺
2926	2016/4/29	245XH	11	NULL	NULL	98	神岡庄尾->普濟寺
2927	2016/4/29	245XH	12	NULL	NULL	98	神岡庄尾->普濟寺
2928	2016/4/29	245XH	13	NULL	NULL	98	神岡庄尾->普濟寺
2929	2016/4/29	245XH	14	NULL	NULL	98	神岡庄尾->普濟寺
2930	2016/4/29	245XH	15	NULL	NULL	98	神岡庄尾->普濟寺
2931	2016/4/29	245XH	16	NULL	NULL	98	神岡庄尾->普濟寺
2932	2016/4/29	245XH	17	NULL	NULL	98	神岡庄尾->普濟寺
2933	2016/4/29	245XH	18	NULL	NULL	98	神岡庄尾->普濟寺
2934	2016/4/29	246U8	6	82	高鐵台中站-台中水湳	82	高鐵台中站->台中水湳
2935	2016/4/29	246U8	7	82	高鐵台中站-台中水湳	82	高鐵台中站->台中水湳
2936	2016/4/29	246U8	8	82	高鐵台中站-台中水湳	NULL	NULL
2937	2016/4/29	246U8	9	82	高鐵台中站-台中水湳	NULL	NULL
2938	2016/4/29	246U8	10	82	高鐵台中站-台中水湳	NULL	NULL
2939	2016/4/29	246U8	11	82	高鐵台中站-台中水湳	NULL	NULL
2940	2016/4/29	246U8	12	82	高鐵台中站-台中水湳	NULL	NULL
2941	2016/4/29	246U8	13	82	高鐵台中站-台中水湳	NULL	NULL

圖 4-62 電子票證與公車動態資料勾機失敗差異比較圖

整體而言，各種數據異常樣態僅是最終的結果，異常的原因通常包羅萬象，從一開始的硬體設備、偵測、解碼、傳輸、儲存與後處理等均可能導致資料異常，故從最下游（甚至已是經處理過的二手資料）資料執行異

常樣態分析僅能發現、回報問題，較無法確認、釐清問題原因，但仍對資料品質的改善有一定程度之助益。

### 4.4.3 站牌班表推動可行性分析

#### 4.4.3.1 示範路線選定

在站牌班表推動可行性分析前，可以先選定示範路線並對其進行站牌班表之試算，為了達到此目的，這些示範路線必須選取性質差異較大之路線，以確保分析方法可適用於不同性質之路線。

所選定的示範路線可分為市區與郊區路線。根據過往文獻分析，市區公車運行的平均速度約為 10-15 km/hr，因此本研究將運行速度小於 15 km/hr 者為市區路線、大於 15 km/hr 者為郊區路線。

本研究以公車動態歷史資料進行分析，並輔以視覺化軟體 Tableau 檢視各路線之車輛行駛時空圖進行分析，初步選定兩條市區路線 51 路及 289 路、兩條郊區路線 153 路及 276 路等共四條路線，並整理該四條路線之基本資料如圖 4-63。

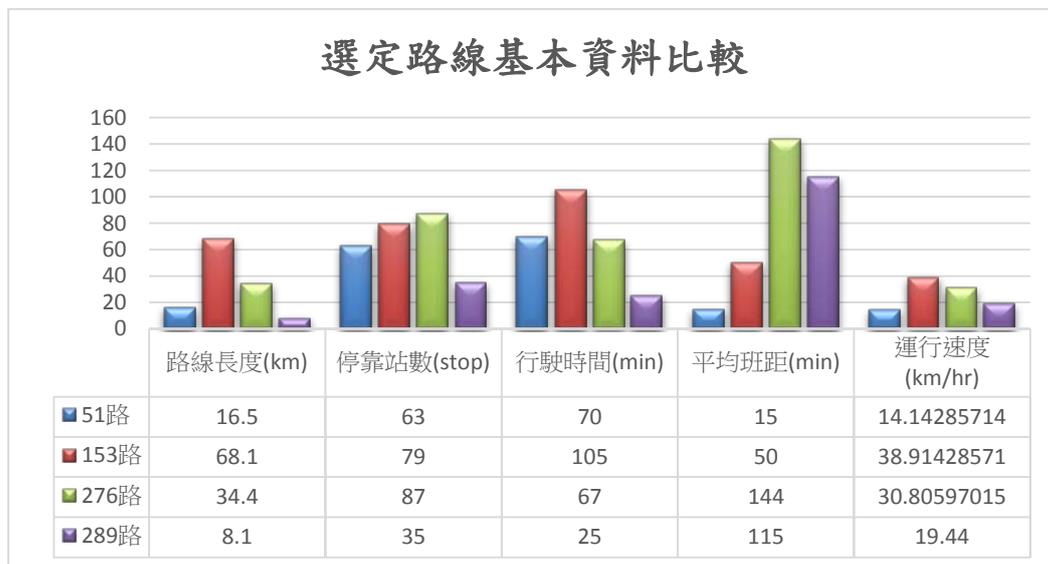


圖 4-63 站牌班表選定路線基本資料比較

由圖 4-63 所示，可以發現郊區路線（153/276）的平均運行速度比市區路線（51/289）快上許多，且交通擁擠的程度比較不嚴重。如圖 4-64 所示，153 路之車輛行駛時空圖幾乎找不到水平線（代表路況擁塞造成車輛靜止）。此外 153 路與 276 路行經郊區的長度遠較行經市區長，導致這兩條路線的平均站距比市區公車平均站距要長，其中 153 路平均站距大於 800 m，276 路的平均站距約為 400 m，顯然較不受道路壅塞影響。

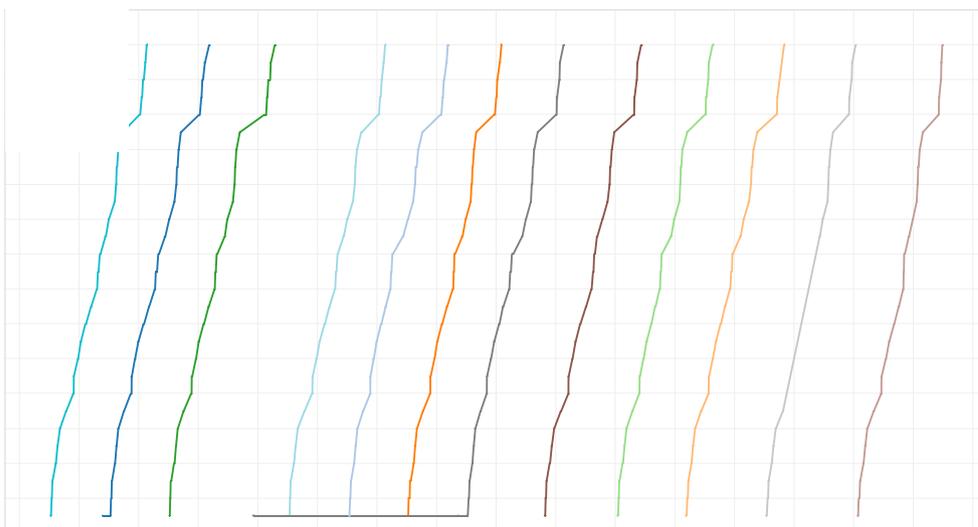


圖 4-64 【153】路車輛行駛時空圖

相較之下，市區路線的運行速度便與郊區路線有一段落差，尤其 51 路公車平均站距僅約 250 m 左右、平均運行速度小於 15 km/hr，且行經市區佔總路線長度的比例偏高，因此比較容易出現交通擁擠的現象。如圖 4-65 所示，在所顯示的車輛行駛時空圖上應出現較多代表走走停停的水平線。

跨日檢視

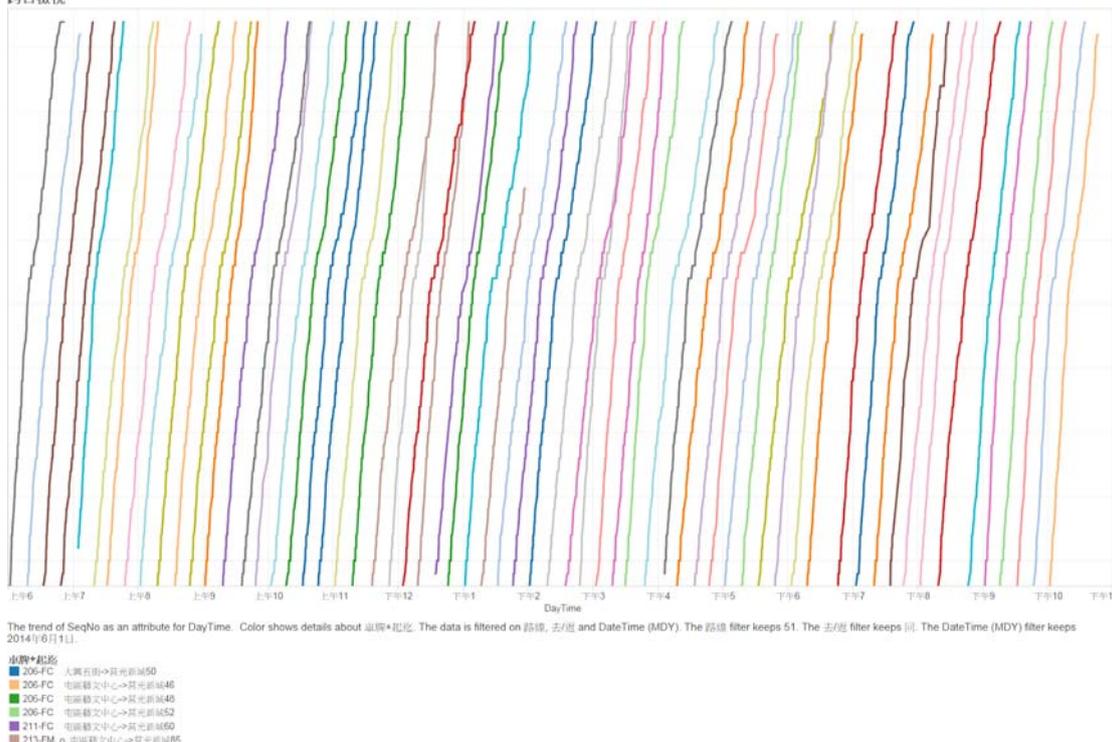


圖 4-65 【51】路車輛行駛時空圖

#### 4.4.3.2 推動機制建議

由於近年市區公車普遍已安裝車機、可定時將公車 GPS 定位資料回傳至公車動態資料庫，故可以這些公車動態歷史資料作為站牌班表產製的基礎。考量上述公車動態歷史資料之資料格式係依循臺灣車載資通訊產業協會（TTIA）制訂之《營業大客車車載機產業標準 1.6 版本》規範，故進行資料處理時應可避免因格式差異產生資料處理上的困擾。

表 4.4 公車動態歷史資料格式範例【289 路】

2	Date	Route	BusID	Goback	NameZh	Leave	Time	UniqueBu	driver_id
3	2015/1/1	289	609-FX	0	臺中家商	TRUE	2015/1/1 07:50	609-FX_o	1458
4	2015/1/1	289	609-FX	0	後火車站	FALSE	2015/1/1 07:50	609-FX_o	1458
5	2015/1/1	289	609-FX	0	後火車站	TRUE	2015/1/1 07:51	609-FX_o	1458
6	2015/1/1	289	609-FX	0	新時代購	FALSE	2015/1/1 07:52	609-FX_o	1458
7	2015/1/1	289	609-FX	0	新時代購	TRUE	2015/1/1 07:52	609-FX_o	1458
8	2015/1/1	289	609-FX	0	大智忠孝	FALSE	2015/1/1 07:53	609-FX_o	1458
9	2015/1/1	289	609-FX	0	大智忠孝	TRUE	2015/1/1 07:53	609-FX_o	1458
10	2015/1/1	289	609-FX	0	建成大振	FALSE	2015/1/1 07:55	609-FX_o	1458
11	2015/1/1	289	609-FX	0	建成大振	TRUE	2015/1/1 07:55	609-FX_o	1458

資料來源：豐原客運提供暨本研究整理

本研究建議推動程序可參考以下數項機制：

##### 1. 由公車動態資料庫取得公車動態歷史資料

自資料庫取得公車動態歷史資料後，必須進行前置處理、本研究使用 C 語言儲存包括路線、車號、到離站時間、駕駛員代號等資訊，以備後續計算之用。

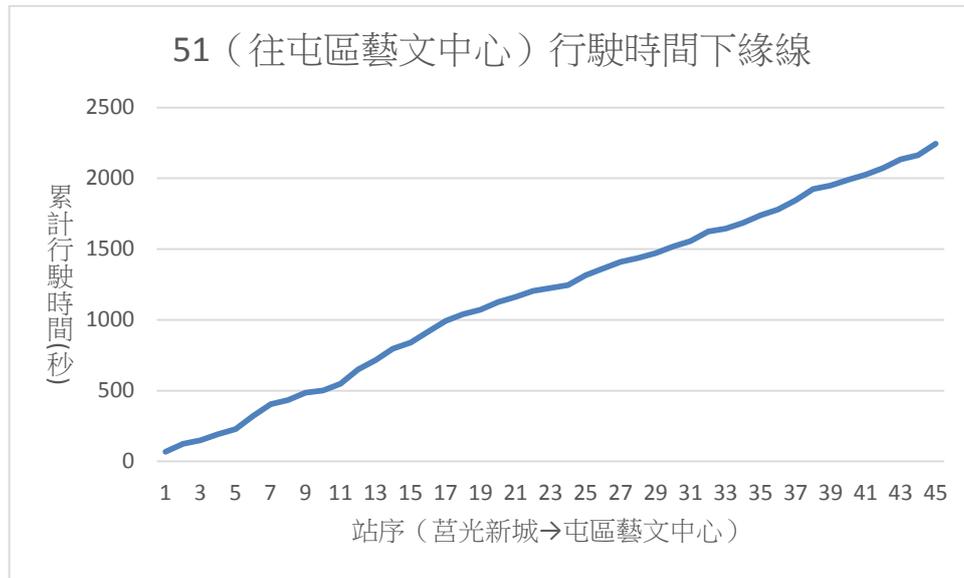
##### 2. 輔導駕駛行為

根據所取得的歷史資料，可以針對每一條路線的每一時段繪製時空圖，然後根據此時空圖，可以找出行駛速度偏快與偏慢的駕駛員，可建議業者針對這些駕駛員進行相關的關懷或是輔導，了解其駕駛行為與大多數人不同的原因，並進一步輔導駕駛行為，讓同一時段的班次行駛時間或是速度趨向一致。

##### 3. 站牌班表產製

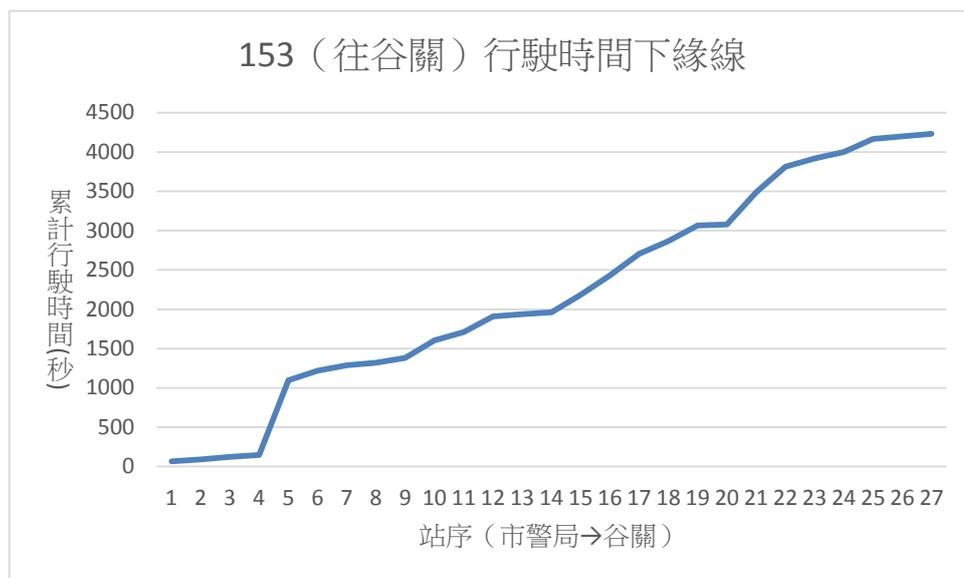
本研究建議採取下緣線基準法，此方法乃是以該路線中 fastest 一班公車作為下緣線，並以下緣線為基礎假設多條候選基準線（通常與下緣線呈倍數關係），對偏離候選基準線者訂出懲罰值，最後選擇總懲罰值最小者為新基準線。以此方法訂出之新基準線，其所代表的意義為呈現出大部分駕駛之駕駛習性，如此一來只須改變少數駕駛之駕駛習性，即可讓該路線之

行駛時間變得可受控制、達到控管班距的目的，如此一來，站牌班表即可由各班次取其相鄰各站之行駛時間平均值來獲得。



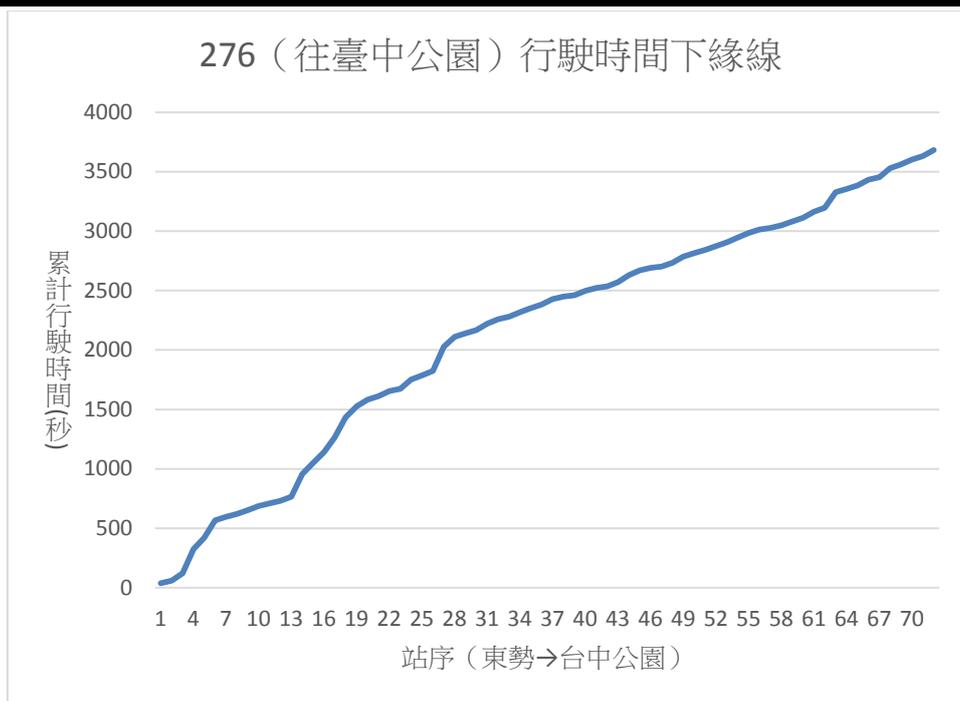
資料來源：豐原客運提供暨本研究整理

圖 4-66 【51 路 (往屯區藝文中心)】行駛時間下緣線



資料來源：豐原客運提供暨本研究整理

圖 4-67 【153 路 (往谷關)】行駛時間下緣線



資料來源：豐原客運提供暨本研究整理

圖 4-68 【276 路 (往臺中公園)】 行駛時間下緣線



資料來源：豐原客運提供暨本研究整理

圖 4-69 【289 路 (往東平社區)】 行駛時間下緣線

以 289 路為例（如圖 4-69），以不同日的第 N 班車或固定每天的同一時段（如 7-8 點）選出行駛時間最短之一班車，此即為初始下緣線；參考運輸計劃季刊第 41 卷第 4 期「利用公車動態資訊系統產製站位班表之研究」，以初始下緣線為基準乘以一權重調整值（=1.25）即獲得調整後之下緣線，如下表 4.5 所示：

表 4.5 採用下緣線基準法算出之【289 路】調整後之下緣線

起站	下一站	初始 下緣線	權重 調整	調整後之 下緣線(秒)
臺中火車站	臺中家商	66	1.25	82.5
臺中家商	後火車站	89	1.25	111.25
後火車站	新時代購物中心	109	1.25	136.25
新時代購物中心	大智忠孝路口	143	1.25	178.75
大智忠孝路口	建成大振街口	172	1.25	215
建成大振街口	大東紡織(振興路)	198	1.25	247.5
大東紡織(振興路)	漢聲電臺	215	1.25	268.75
漢聲電臺	東門橋	229	1.25	286.25
東門橋	早溪	264	1.25	330
早溪	太平橋前	284	1.25	355
太平橋前	下部仔	317	1.25	396.25
下部仔	中油太平站	333	1.25	416.25
中油太平站	下太平	355	1.25	443.75
下太平	太平農會	388	1.25	485
太平農會	太平中平路口	413	1.25	516.25
太平中平路口	中平國中	446	1.25	557.5
中平國中	精美社區	470	1.25	587.5
精美社區	太平游泳池	495	1.25	618.75
太平游泳池	長億六永成北路口	522	1.25	652.5
長億六永成北路口	長億六太平一街口	565	1.25	706.25
長億六太平一街口	長億高中	588	1.25	735
長億高中	太平精美路口	725	1.25	906.25
太平精美路口	三豐社區	780	1.25	975
三豐社區	東平社區	823	1.25	1028.75

資料來源：豐原客運提供暨本研究整理

#### 4.4.3.3 駕駛訓練管理建議

自公車動態資料庫取得公車動態歷史資料之後，必須先進行資料處理，具體做法羅列如下述之四步驟：

##### 1. 資料清理

計算車上時間：排除車輛停靠時間，僅探討司機開車快慢。

##### 2. 計算車上時間中位數

計算各路線路徑各司機各停靠站的車上旅行時間中位數：計算中位數，避免離群值影響。

##### 3. 計算門檻值

計算各路線路徑各停靠站的車上旅行時間中位數最大值與最小值：作為駕駛行為調整建議用。

##### 4. 駕駛行為調整建議

計算各路線路徑各司機各停靠站旅行時間中位數等於最大值與最小值的比例：在同一路線路徑中，若各停靠站旅行時間中位數等於最大值的個數佔總停靠站數量的比例愈高，則建議加快速度；若各停靠站旅行時間中位數等於最小值的個數佔總停靠站數量的比例愈高，則建議放慢速度。

以選定的路線為例，表 4.6 試列舉「個別駕駛抵達各停靠站的旅行時間中位數最大值或最小值」佔比重最高者之駕駛代碼、與對應之駕駛訓練管理建議。

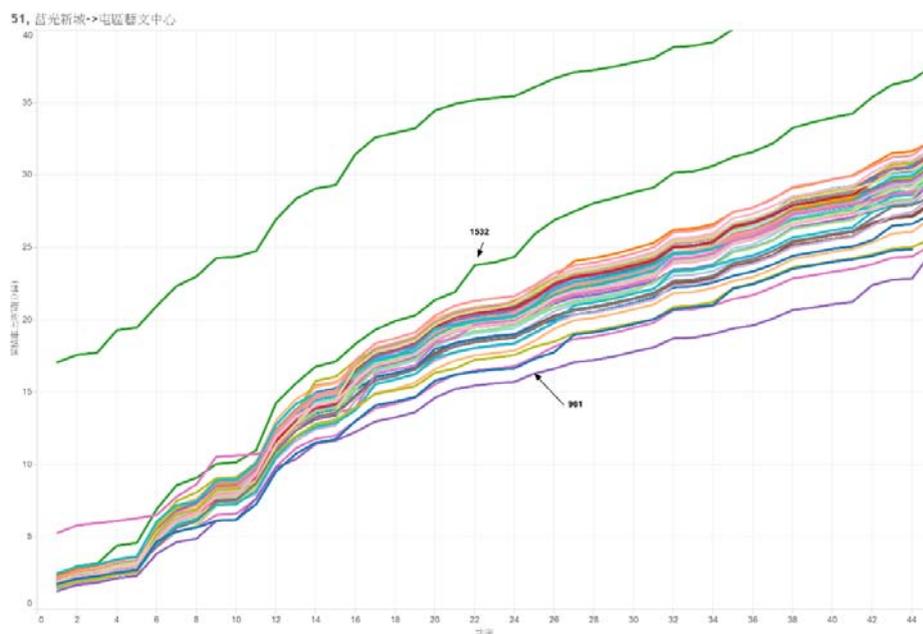
表 4.6 駕駛訓練管理建議

路線名稱	編號	駕駛代號	最大值 佔比	最小值 佔比	駕駛訓練 管理建議
莒光新城 → 屯區藝文中心	51	961	4%	42%	放慢
	51	1532	27%	0%	加快
屯區藝文中心 → 莒光新城	51	961	0%	40%	放慢
	51	1528	25%	0%	加快
市警局	153	1264	0%	46%	放慢

→谷關	153	1324	30%	4%	加快
谷關	153	936	4%	48%	放慢
→市警局	153	1310	27%	4%	加快
東勢→	276	871	1%	49%	放慢
臺中公園(雙十路)	276	1459	26%	4%	加快
臺中公園(雙十路)	276	1468	3%	29%	放慢
→東勢	276	1510	17%	6%	加快
東平社區	289	1469	0%	46%	放慢
→臺中火車站	289	1160	25%	8%	加快
臺中火車站	289	1489	8%	29%	放慢
→東平社區	289	1510	22%	0%	加快

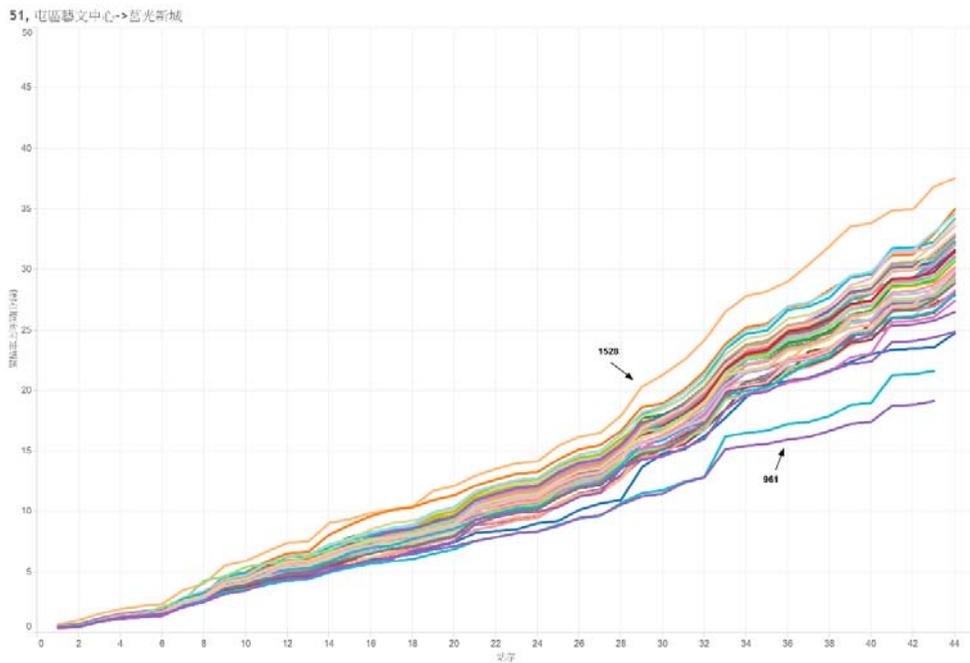
資料來源：豐原客運提供暨本研究整理

表 4.6 中各路線所對應的旅行時空圖如圖 4-70~圖 4-77，每一駕駛員之駕駛行為皆以一條線作為代表、該線為駕駛員於各站間最快之行駛紀錄，故只要從圖中就可輕易判斷哪些駕駛員之駕駛行為必須調整：時空圖中斜率越大者代表該駕駛行駛時間比其他駕駛更長，建議駕駛加快；反之，斜率越小者建議駕駛放慢。



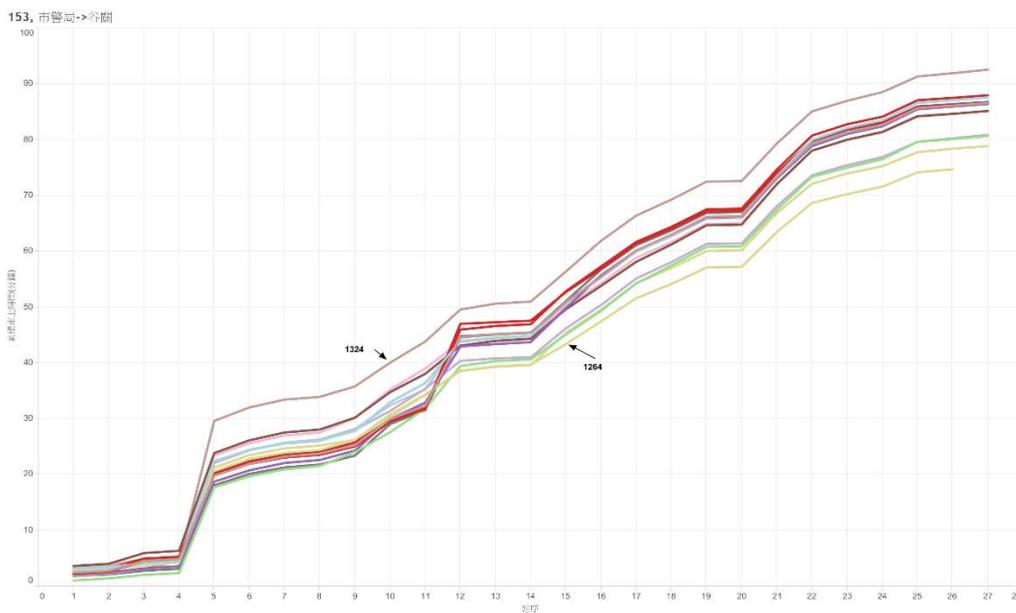
資料來源：本研究整理

圖 4-70 【51】莒光新城→屯區藝文中心



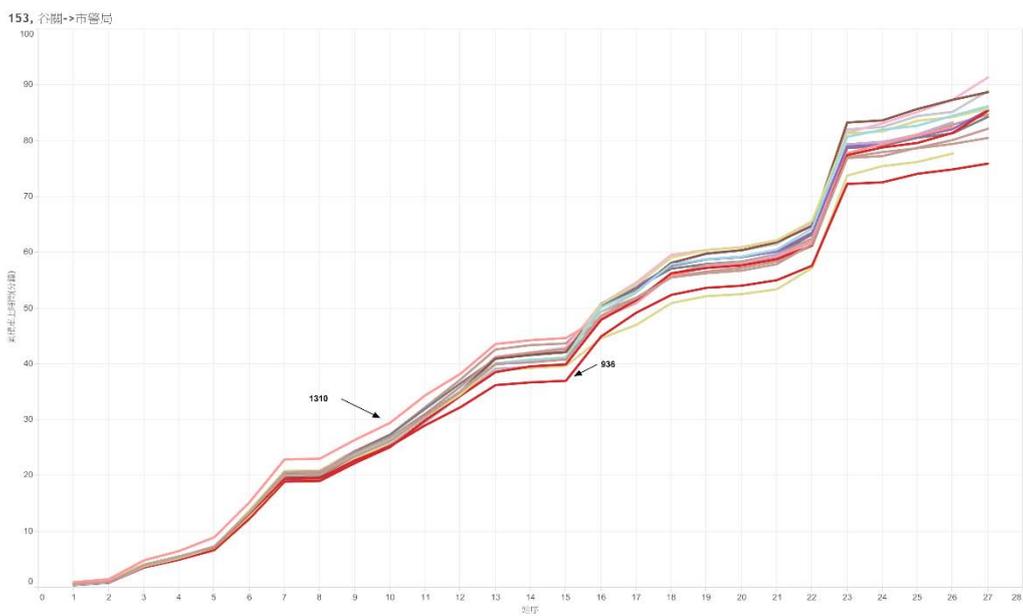
資料來源：本研究整理

圖 4-71 【51】屯區藝文中心→莒光新城



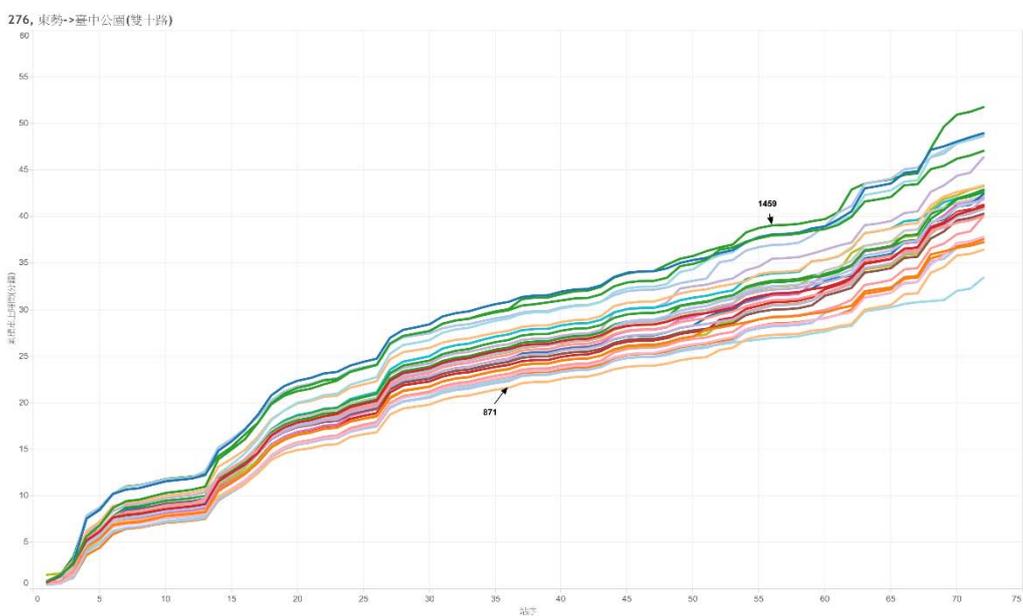
資料來源：本研究整理

圖 4-72 【153】市警局→谷關



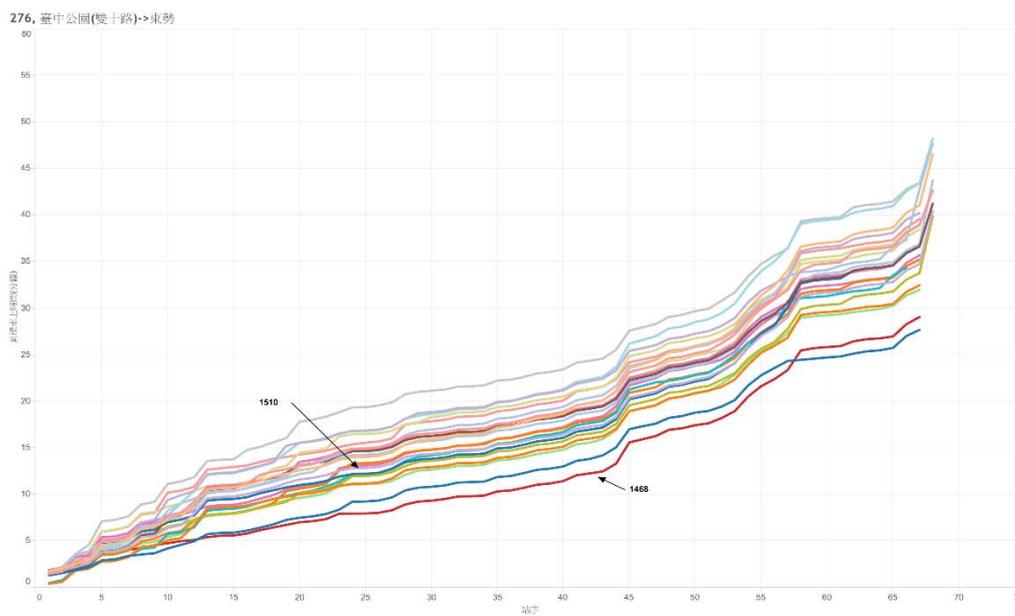
資料來源：本研究整理

圖 4-73 【153】 谷關→市警局



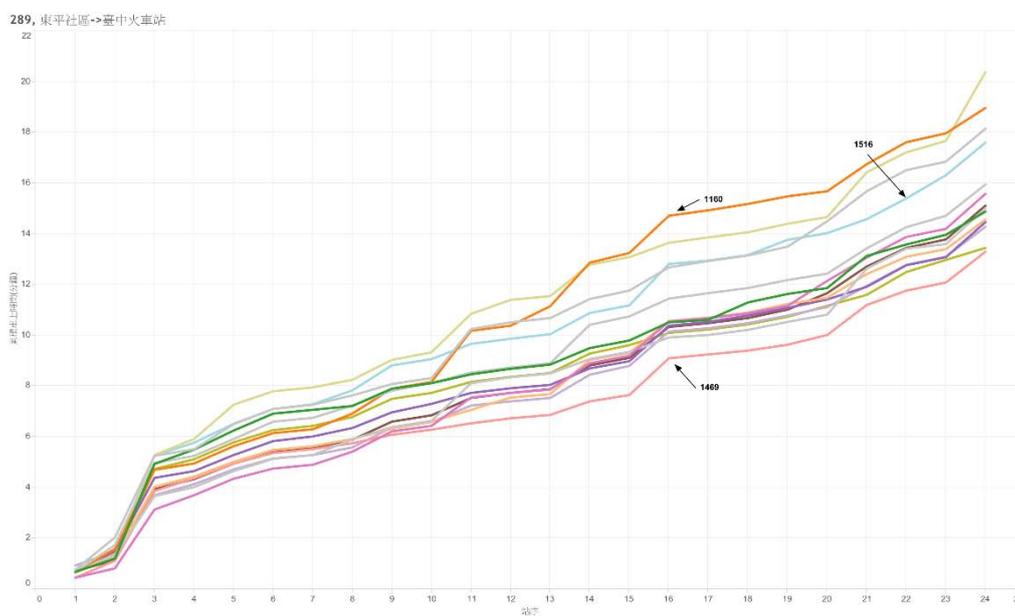
資料來源：本研究整理

圖 4-74 【276】 東勢→臺中公園(雙十路)



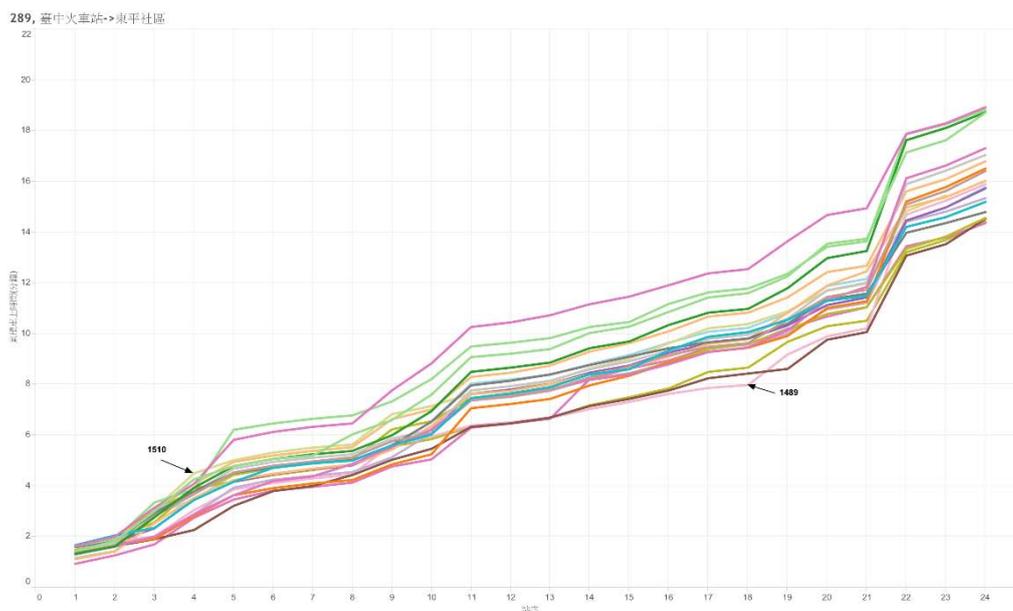
資料來源：本研究整理

圖 4-75 【276】臺中公園(雙十路)→東勢



資料來源：本研究整理

圖 4-76 【289】東平社區→臺中火車站



資料來源：本研究整理

圖 4-77 【289】臺中火車站→東平社區

## 4.5 安全面向分析

安全是大眾運輸系統最重要的基本條件，本研究使用公部門肇事紀錄資料集來試作安全面向分析，根據肇事地點資料建構肇事風險地圖，再與公車動態 A1 紀錄套疊，找出行經高肇事風險路段最頻繁之路線或車輛，可作為駕駛員宣導或關注其風險之用。

本研究蒐集 2014 年 10 月 01 日起至 2015 年 04 月為期共七個月之臺中市交通事故資料共 87,260 筆資料，後續利用 Google Map Geocode API 將資料集之事故地址轉為經緯度座標，同時以 10 公尺×10 公尺為切割基本單位在臺中地圖上劃分 xx 區塊並予以編碼，再套疊 2015 全年之公車動態 A1 之座標位址，意即每 20 秒抽樣抓取市區公車之位置，分析其出現在肇事地點方塊之頻率，以作為風險高低之判斷，進而得到高風險客運路線與車輛，以下分五步驟說明分析程序。

### 1. 步驟一

由於原資料僅大致紀錄事故地點的相關資料(如圖 4-78)，並無經緯度資料，因此本研究撰寫 C#程式彙整事故發生地址並呼叫 Google Map Geocode API，將事故地點傳至 Google Map Geocode API，並從 API 回傳的

資料中，取得事故地點的經緯度。但由於部分地點的相關資料不夠精確不利於轉譯經緯度座標 (如圖 4-79)，因此轉換後經緯度結果會與實際發生事故地點有誤差，此點可作為未來資料紀錄/蒐集或資料匯流時可改善之項目。

Type	發生地點	死亡人數	受傷人數	車種
A2	臺中市西區臺灣大道二段297號	0	2	普通重型-機車,普通輕型-機車,自用-小客車
A2	臺中市沙鹿區臺灣大道路七段200號	0	2	普通重型-機車,普通重型-機車
A2	臺中市潭子區民族路一段30號	0	2	普通重型-機車,普通重型-機車
A2	臺中市大里區瑞成一街125巷15號	0	2	普通重型-機車,普通重型-機車
A2	臺中市南區復興路一段406號	0	3	普通重型-機車,普通重型-機車,普通重型-機車
A2	臺中市北區金龍里金龍街62號	0	2	普通重型-機車,普通輕型-機車
A2	臺中市北屯區軍功里東山路一段236-1號	0	2	普通重型-機車,普通重型-機車
A2	臺中市太平區振福路397號	0	2	普通重型-機車,普通重型-機車,乘客-人
A2	臺中市霧峰區吉峰路155-3號	0	3	普通輕型-機車,普通重型-機車,普通重型-機車
A2	臺中市南屯區河南路四段508號	0	2	普通重型-機車,普通重型-機車,普通重型-機車
A2	臺中市東勢區豐勢路115號	0	2	普通重型-機車,普通重型-機車

圖 4-78 肇事資料集中事故地點資料樣本截圖

Type	發生地點	死亡人數	受傷人數	車種
A1	臺中市南屯區豐樂里永春東二路6657號路燈附近(附近)	1	1	自用-小貨車,自用-小貨車,自用-小客車
A2	臺中市烏日區五光里五光路五中巷12470號路燈附近(附近)	0	2	自用-小客車,普通重型-機車,普通重型-機車
A2	臺中市烏日區五光里環中路七段橋樑車道18346號路燈附近(附近)	0	2	普通重型-機車,普通重型-機車
A2	臺中市烏日區烏日里中山路三段集泉橋16494路燈附近(附近)	0	2	普通重型-機車,腳踏自行車-慢車
A2	臺中市烏日區成功西路14848號路燈附近(附近)	0	2	自用-小客車,自用-小客車
A2	臺中市西屯區科園路R26號路燈附近(附近)	0	2	自用-小貨車,普通重型-機車,普通重型-機車,自用-小貨車,自用-小貨車
A2	臺中市龍井區福田里中彰大橋路17K-L08路燈附近(附近)	0	2	普通重型-機車,普通重型-機車
A2	臺中市烏日區學田里學田路14804號路燈附近(附近)	0	2	普通重型-機車,行人-人
A2	臺中市烏日區湖日里中彰快速道路北向下烏日匝道橋樑車道2175號路燈附近(附近)	0	2	普通重型-機車,普通重型-機車
A2	臺中市烏日區三和里成功東路15008號路燈附近(附近)	0	2	普通重型-機車,普通重型-機車

圖 4-79 部分易混淆或不易解碼之資料截圖

## 2. 步驟二

依據事故資料的經緯度，取出最大與最小經緯度組成一個矩形，以每 10 公尺的距離劃分區塊，並依據給予位置代碼。如圖 4-80 所示，A 點為最小經緯度、B 點為最大經緯度，C 點則是該區塊的位置代碼。

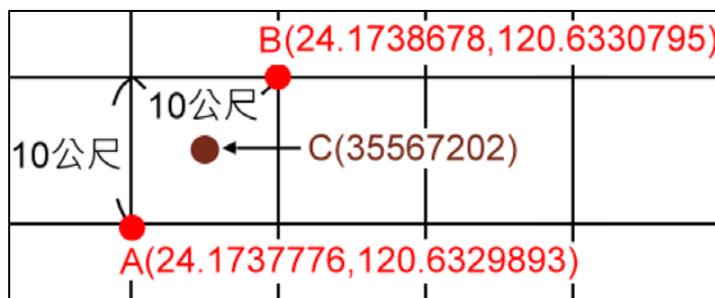


圖 4-80 十公尺區塊範圍示意圖

## 3. 步驟三

將每筆事故資料對應至符合範圍經緯度的區塊。圖 4-81 中該地址經緯度落在於圖 4-80 中的 C 點位置。

seat	address	latitude	longitude
35567202	臺中市西屯區臺灣大道	24.1738022	120.6329917

圖 4-81 事故地址範例圖

#### 4. 步驟四

統計各區塊中的事故資料，依照區塊中發生事故次數大到小進行排序，並取前 50 筆區塊作為後續分析之標的。

#### 5. 步驟五

將步驟 4 結果的 50 筆區塊範圍經緯度與公車動態 A1 資料進行對應，找出經過這 50 個路段的路線與車號，同時統計其經過的次數頻率，藉此評估高風險之路線與車輛。

##### 4.5.1.1 分析結果

由於本分析係根據歷史肇事風險地點作為公共運輸駕駛者之安全提示依據，而非針對公共運輸司機個人的駕駛行為稽核，故不會引起司機心理潛在之抗拒，以下將摘要說明前述各階段所得之分析結果。

經步驟 1~3 及 Google Map Geocode API 將事故地址轉換為經緯度後，全臺中地區的事故地點分佈如圖 4-82，多數事故均屬無人受傷之事故（藍色），僅少數事故有 1 人受傷（橘色），至於 2~8 人受傷或死亡者數量極少，在圖 4-82 中不易發現。

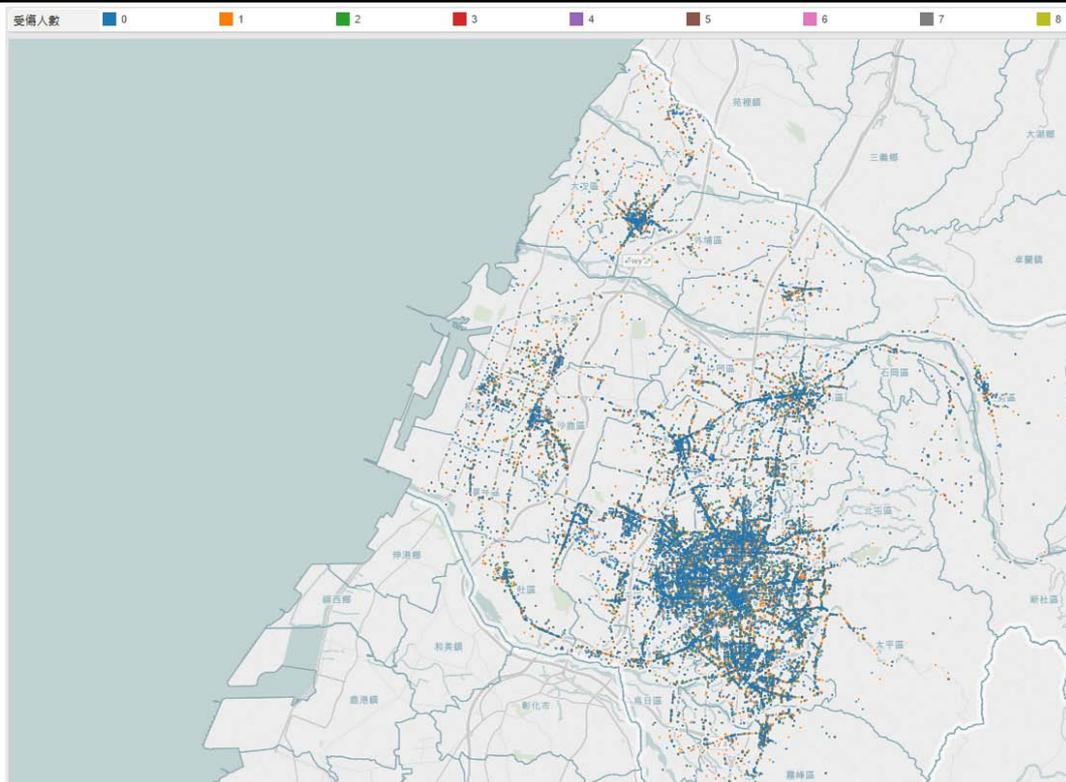


圖 4-82 肇事地址經 Google Maps Geocoding API 轉換後之分佈圖

根據步驟 4 所取得之前 50 名區塊的地址則如圖 4-83 所示。

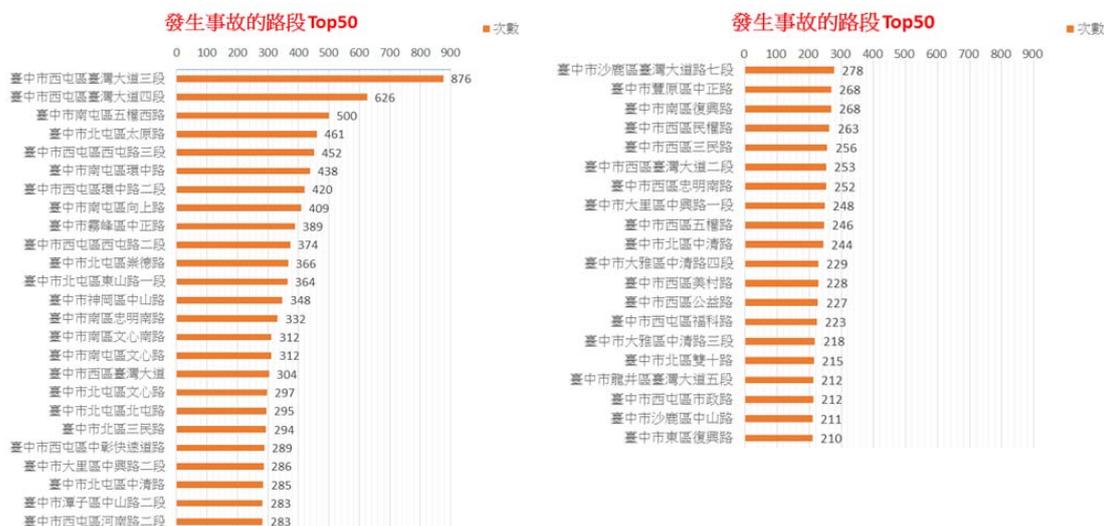


圖 4-83 發生事故的路段 Top50 橫條圖

最後利用步驟 5 所分析之高風險路線與車輛則分別如圖 4-84 與圖 4-85，從中可知 63、55、12 等共 53 條路線高頻率經過前述危險路段，而 672-FZ、252-FZ 與 679-FZ 等共 200 輛車則是高頻率經過前述危險路段，圖 4-84 與圖 4-85 中方塊之大小代表通過頻率的相對值，可作為司機安全宣導標的之參考。



圖 4-84 高風險路線樹圖



圖 4-85 高風險車輛樹圖

## 4.6 無縫面向分析

### 4.6.1 電子票證跨運具分析效益

跨運具的電子票證分析可以協助研究團隊串連不同運具的起/迄點，勾勒出民眾實際的起迄需求，亦即將兩段起迄拼湊成一組，若發現這些「重組」後的起迄存在高度的一致性時，主管機關或客運業者即可考慮開聘新的路線來滿足這些民眾，使其運輸過程中毋需轉乘或以更短的距離或效率（例如行經快速道路）來完成其旅次。

就本計畫的標竿案例場域（臺中市）而言，公共運輸的主要成員除公車外，尚包括臺灣鐵路管理局「縱貫線日南～追分」及「臺中線（俗稱山線）泰安～成功」共計 18 站如圖 4-86、以及 60 餘個公共腳踏車 iBike 站點如圖 4-87，均是跨運具分析之首選。

車站名稱	進出	月份											
		1	2	3	4	5	6	7	8	9	10	11	12
1日南	進站	9,892	9,405	10,087	11,043	9,539	9,178	8,504	8,946	8,661	9,426	9,019	8,875
	出站	9,600	8,827	9,280	10,875	9,433	8,563	7,941	8,214	8,163	9,318	8,556	8,300
2大甲	進站	76,718	90,062	78,652	100,335	73,452	71,751	69,737	73,734	69,637	73,548	71,753	69,498
	出站	77,076	85,716	74,562	99,277	71,400	67,126	66,832	71,615	65,665	70,230	66,925	64,325
3台中港	進站	2,014	2,097	2,205	3,596	2,104	1,901	1,815	1,738	1,831	2,255	2,050	1,992
	出站	2,140	2,250	2,274	3,746	2,236	2,047	1,830	1,745	1,960	2,553	2,067	2,058
4清水	進站	31,913	33,162	35,991	37,896	34,917	33,252	32,850	34,234	34,152	40,464	35,228	33,030
	出站	33,455	32,361	35,071	37,892	36,145	32,171	32,081	33,660	33,537	40,873	34,951	33,735
5沙鹿	進站	73,026	72,547	79,122	80,418	75,175	73,215	68,303	71,794	69,657	76,554	73,802	69,911
	出站	69,603	69,468	77,841	77,663	76,687	72,525	66,907	69,221	66,473	73,377	70,950	69,681
6龍井	進站	13,192	13,591	14,307	14,684	14,029	12,784	11,647	12,651	12,730	14,152	12,964	13,119
	出站	13,343	13,891	14,407	15,274	14,734	12,968	11,808	12,289	13,142	14,733	13,164	13,711
7人社	進站	22,362	21,867	22,797	23,517	23,233	20,414	22,389	21,677	21,251	22,941	21,678	21,470
	出站	22,439	21,684	22,560	24,413	22,712	20,276	19,578	21,069	21,161	23,145	21,901	22,209
8追分	進站	25,918	16,512	17,483	34,541	26,323	13,878	11,330	12,711	10,521	14,200	13,816	21,381
	出站	8,598	8,213	7,931	8,628	8,028	7,166	7,115	7,606	7,543	8,128	7,591	7,384
9泰安	進站	9,295	14,715	8,989	8,188	8,075	7,652	8,694	9,122	7,657	8,244	7,925	7,750
	出站	10,646	16,776	10,921	10,002	9,254	9,718	10,145	11,192	9,315	9,916	10,374	9,824
10后里	進站	57,021	60,520	57,689	56,785	52,498	58,012	65,021	70,963	52,367	54,330	54,040	50,357
	出站	60,003	62,486	60,834	60,377	57,868	60,268	64,684	69,575	55,918	59,073	58,862	55,647
11豐原	進站	235,288	240,091	236,964	213,007	209,414	210,791	224,310	242,736	205,870	211,576	228,868	208,990
	出站	249,154	244,456	243,402	223,912	224,204	216,017	231,507	250,897	212,462	225,252	238,558	219,789
12潭子	進站	68,486	67,006	67,444	66,905	68,821	68,191	69,189	74,341	67,049	70,840	70,725	68,804
	出站	70,731	69,889	75,006	74,325	76,296	71,283	69,615	74,818	71,387	77,187	75,435	76,486
13太原	進站	75,877	76,249	83,151	81,853	81,615	75,766	77,066	80,954	77,814	83,413	83,992	82,609
	出站	73,370	74,105	80,351	78,696	79,129	73,970	75,143	79,913	76,314	81,674	82,940	81,258
14台中	進站	794,331	794,116	792,081	773,872	777,181	755,133	794,764	882,501	763,498	813,473	794,345	786,236
	出站	758,696	793,589	775,908	753,098	745,510	745,674	789,325	878,008	738,373	788,238	767,977	757,923
15人慶	進站	63,689	60,428	71,892	70,362	68,709	62,452	61,600	64,143	65,326	71,160	67,076	69,569
	出站	61,237	59,574	70,509	68,905	66,198	62,425	61,182	63,443	64,203	69,635	66,103	68,428
16烏日	進站	20,831	20,990	22,182	21,302	21,172	19,832	19,749	21,184	19,689	20,776	20,992	20,743
	出站	24,962	27,236	27,645	27,353	27,290	25,443	26,131	28,168	25,205	26,657	27,772	27,437
17新烏日	進站	147,840	160,448	152,095	155,677	165,654	150,628	165,206	182,262	158,462	164,159	162,498	161,824
	出站	118,210	152,576	139,615	137,166	140,692	142,200	151,014	167,804	143,312	147,149	146,627	148,540
18成功	進站	15,197	14,796	16,169	14,836	15,020	14,543	15,421	17,279	16,244	16,065	15,740	15,747
	出站	33,297	23,539	26,632	41,444	35,564	21,985	20,810	23,858	20,258	23,367	23,535	30,997

圖 4-86 臺中地區傳統鐵路車站 2014 各月分進出人數圖

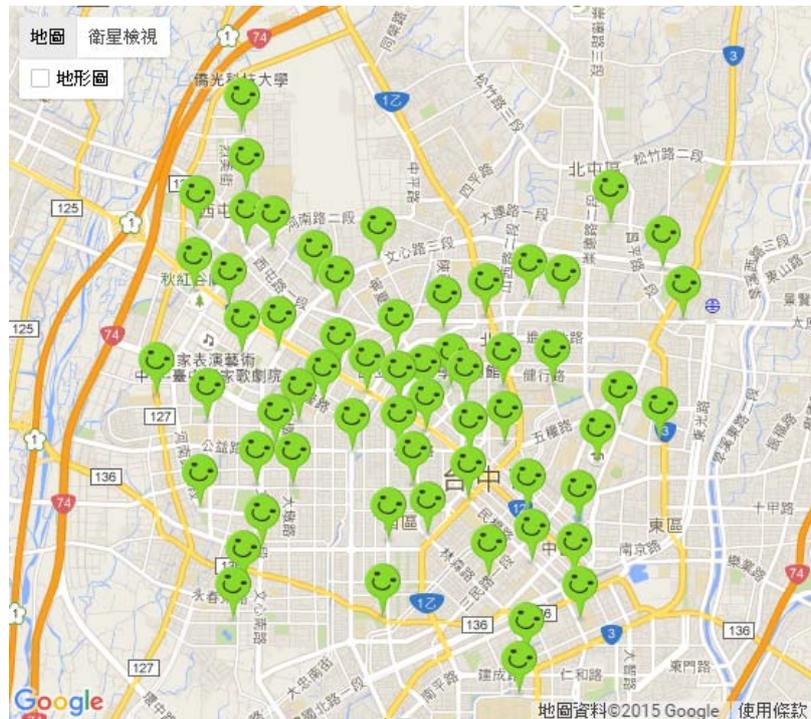


圖 4-87 臺中地區 60 餘個 iBike 地理位置圖

## 4.6.2 跨運具分析可行性探討

### 4.6.2.1 可行關鍵因素

整體而言，跨運具分析之成功關鍵因素在於能勾稽同一客體連續的旅運行為，因此所謂的「卡號」或其他類似名稱的識別碼即是關鍵之最。然而，基於個人資料議題因素，通常資料保管者在對外提供資料時，有可能是提供晶片內碼或利用「遮罩/編碼/碎片」等手段混淆，例如以臺灣鐵路管理局為例，其多卡通交易紀錄主要欄位如圖 4-88，其實際資料呈現如圖 4-89，但反觀市區公車交易紀錄之卡號就明顯不同如圖 4-90。

欄位ID	欄位名稱	型態
CARD_TXN_TYPE_ID	交易類別	number(3)
CARD_TXN_SUBTYPE_ID	交易次類別	number(3)
DEV_ID	設備編號	number(9)
SP_ID	業者別	number(3)
TXN_DATE	交易日	number(8)
TXN_TIME	交易時間/出站時間	number(6)
CARD_PHYSICAL_ID	晶片卡號(內碼)	number(10)
ISSUER_ID	卡片發行者	number(3)
CARD_TXN_SEQ_NO	交易序號	number(3)
TXN_AMT	交易金額	number(5)
ELECTRONIC_VALUE	卡片餘額	number(7,2)
SVCE_LOC_ID	場站代碼	number(3)
PROCESSING_DATE	處理日	date
SETT_DATE	清分日	date
BUSINESS_DATE	營運日	date
DEV_TYPE_ID	設備種類	number(3)
ENTRY_LOC_ID	進站場站代碼	number(3)
XFER_CODE	轉乘代碼	varchar2(4)
XFER_DISC	轉乘優惠	number(7,2)
PERSONAL_DISC	優惠補助款	number(7,2)
PENALTY	罰款金額	number(5)
LOYALTY_COUNTER	忠誠點數	number(5)
AGENT_NO	站務員代碼	number(5)

資料來源：臺鐵路提供&本研究整理（原 61 個欄受限篇幅僅列 23 個欄位）

圖 4-88 臺灣鐵路管理局多卡通交易紀錄主要欄位

	dev_id	sp_id	txn_date	txn_time	card_physical_id	issuer_id	card_txn_seq_no	amt
1	100063728	100	20120101	3	2245366493	255	39	28
2	100063784	100	20120101	5	1511239282	255	130	64
3	100060184	100	20120101	7	2240096386	255	184	94
4	100063777	100	20120101	21	3571478595	255	51	14
5	100060182	100	20120101	23	1624675155	255	220	14
6	100063728	100	20120101	25	4071209166	255	26	18
7	100063742	100	20120101	25	3573112195	255	59	51
8	100063728	100	20120101	28	3589149283	255	108	18
9	100063770	100	20120101	31	1116679389	255	106	64
10	100063728	100	20120101	36	4175734083	255	221	18
11	100060201	100	20120101	43	3113447117	255	110	14
12	100063733	100	20120101	45	3229214723	255	203	51
13	100061568	100	20120101	50	1401413822	255	54	103
14	100060184	100	20120101	52	2891223275	255	141	14
15	100063733	100	20120101	54	2669067936	255	3	51

圖 4-89 臺灣鐵路管理局多卡通卡號資料截圖

若無法取得圖 4-89 與圖 4-90 之間的勾機對應關係，就會造成分析之瓶頸，換言之，若兩項資料源對於「卡號」能有一致的編碼方式，就存在電子票證跨運具分析的可行性。

路線	駕駛	車輛	卡號	發卡公司	卡種	上車日期	上車時間	上車站	下車日期	下車時間	下車站
1	43010	998990 309-UU	2F121B26	ECC	一般票	2015-12-10	13:54:13.0000000	林口長庚	2015-12-10	14:19:02.0000000	蘆竹區公所
2	47020	036536 869-FL	2F12E10E	ECC	敬老1票	2015-12-25	09:00:18.0000000	崎頂	2015-12-25	09:05:52.0000000	新街尾
3	42160	998957 698-FD	2F13010F	ECC	敬老1票	2015-12-09	08:20:02.0000000	平鎮區公所	2015-12-09	09:30:46.0000000	湧光里集會所
4	47050	036536 869-FL	2F13C50E	ECC	敬老1票	2015-12-10	10:00:48.0000000	大溪區公所	2015-12-10	10:23:35.0000000	建漢公司
5	41090	36603 035-FN	2F13CD0E	ECC	敬老1票	2015-12-14	11:22:54.0000000	桃園區公所	2015-12-14	11:32:31.0000000	龍安街
6	47030	041808 278-WD	2F13E40E	ECC	敬老1票	2015-12-02	09:29:00.0000000	新街尾	2015-12-02	09:37:52.0000000	仁美街口
7	42030	044632 115-FV	2F13EA0E	ECC	敬老1票	2015-12-11	10:16:05.0000000	忠孝廣場	2015-12-11	10:20:38.0000000	游泳路口
8	47090	036536 869-FL	2F14080F	ECC	敬老1票	2015-12-17	09:58:45.0000000	國軍醫院	2015-12-17	10:09:53.0000000	員樹林農會
9	41030	36603 035-FN	2F14C00E	ECC	敬老1票	2015-12-12	14:23:38.0000000	莒光街口	2015-12-12	14:56:46.0000000	天祥六街口
10	41061	36603 035-FN	2F14C00E	ECC	敬老1票	2015-12-11	13:56:40.0000000	桃園區公所	2015-12-11	14:22:24.0000000	聖保祿醫院
11	41070	035741 866-FP	2F14C00E	ECC	敬老1票	2015-12-11	16:01:17.0000000	聖保祿醫院	2015-12-11	16:16:00.0000000	陽明公園
12	46020	35787 557-U7	2F156C0E	ECC	學生票	2015-12-31	18:29:12.0000000	富岡火車站	2015-12-31	18:50:00.0000000	李厝
13	46031	43210 558-U7	2F156C0E	ECC	學生票	2015-12-31	16:14:38.0000000	大坡(一)	2015-12-31	16:59:13.0000000	富岡火車站
14	42120	044543 103-FV	2F15F4A8	ECC	孩童票	2015-12-12	16:04:15.0000000	平鎮高中	2015-12-12	16:18:27.0000000	福德路
15	46050	017776 859-FL	2F16215D	ECC	一般票	2015-12-28	16:11:36.0000000	富泰公司	2015-12-28	16:29:01.0000000	高鐵桃園站

圖 4-90 桃園地區公車票證資料截圖

#### 4.6.2.2 可行性深入分析

為了更深入探討電子票證跨運具分析的可行性，本研究商請電子票證公司用一致的卡號編碼方式提供公共汽車、傳統鐵路以及公共自行車之資料如圖 4-91、圖 4-92 與圖 4-94。

客運公車代碼	卡號	票種編號	票種代碼	路線公路編號	司機編號	車號	上車時間	上車站序	上車計費站序	下車時間	下車站序	下車計費站序
1	033036	/gL3RChKV528	1	1	17	51264	199-U8 2016-03-24 14:06:00	47	94	2016-03-24 14:14:00	39	78
2	033036	9/wCuN4C1JKA	1	1	17	51264	199-U8 2016-03-24 17:45:00	22	44	2016-03-24 17:52:00	19	38
3	033036	OLcWJCOgei1	1	1	17	51264	199-U8 2016-03-24 19:58:00	51	102	2016-03-24 20:08:00	59	118
4	033036	i65UQx3.H4G0	1	1	17	51264	199-U8 2016-03-24 16:55:00	54	108	2016-03-24 17:04:00	44	88
5	033036	qxLjphMveZ3	1	1	17	51264	199-U8 2016-03-24 15:42:00	22	44	2016-03-24 16:25:00	49	98
6	033036	gg9HUUO3wp/3	1	1	17	51264	199-U8 2016-03-24 20:15:00	59	118	2016-03-24 20:25:00	50	100
7	033036	KUEp2HU4L3E/	1	1	17	51264	199-U8 2016-03-24 11:23:00	34	68	2016-03-24 11:35:00	28	56
8	033036	cbv5VQDjZzyc	1	1	17	51264	199-U8 2016-03-24 18:31:00	0	0	2016-03-24 18:52:00	11	22
9	033036	w2CCyCC.bRR/	1	1	17	51264	199-U8 2016-03-24 11:26:00	32	64	2016-03-24 11:33:00	29	58
10	033036	tuJ3hkU5pO/A	1	1	17	51264	199-U8 2016-03-24 17:44:00	22	44	2016-03-24 18:09:00	10	20

圖 4-91 同編碼之公共汽車票證資料截圖

從圖 4-91 可發現資料欄位完整，但上/下車(計費)站序共四個欄位雖有編號但並無實際對應，換言之，此資料集並無實際站牌空間資訊可供分析使用，故本研究將另行發展插補推估演算法來彌補此資訊漏失，但插補的機制必須依賴正確完整的公車動態資訊。

運具	卡號	票種編號	票種代碼	進站時間	進站地點	車站時間	車站地點	
1	台鐵	zLvC.DAAnCW.	1	1	NULL	9	2015-01-01 07:21:00	15
2	台鐵	g1tQKLYmbsl7	1	1	NULL	12	2015-01-01 18:30:00	11
3	台鐵	1tR/Dp13971.	1	1	NULL	12	2015-01-01 11:21:00	11
4	台鐵	1tR/Dp13971.	1	1	NULL	11	2015-01-01 21:14:00	12
5	台鐵	0m0LJQsGEL.3	1	1	NULL	28	2015-01-01 13:32:00	12
6	台鐵	e2o0BGJkJuU6	1	1	NULL	9	2015-01-01 13:11:00	16
7	台鐵	k0th1fwAAHM7	1	1	NULL	15	2015-01-01 16:00:00	8
8	台鐵	ZNJpYnl/p2C	1	1	NULL	15	2015-01-01 11:32:00	9
9	台鐵	6JOAMYSIL.3	1	1	NULL	15	2015-01-01 14:54:00	9
10	台鐵	6JOAMYSIL.3	1	1	NULL	9	2015-01-01 18:50:00	15

圖 4-92 同編碼之臺鐵多卡通票證資料截圖

至於圖 4-92 所呈現的臺鐵多卡通票證資料可清楚發現缺乏「進站時間」資訊，但「進站地點」與「出站地點」之代碼則是有明確對應表，因此相對於圖 4-91 的公車交易紀錄而言，鐵路的刷進刷出空間資訊是完整的，但是刷入時間付之闕如，因此本研究另行發展插補推估演算法來彌補此資訊漏失，插補的機制必須依賴臺灣鐵路管理局之時刻表，插補演算法概念如圖 4-93，假設票卡資料為 D 站搭乘至 B 站並於  $T_1$  時刷出，考慮可能的步行時間得到  $T_2$  後向前搜尋符合上/下車停靠的可行車次，即可由得  $T_3$ ，本研究已取得與所蒐集之多卡通資料日期對應的臺鐵路時刻表並完成插補。

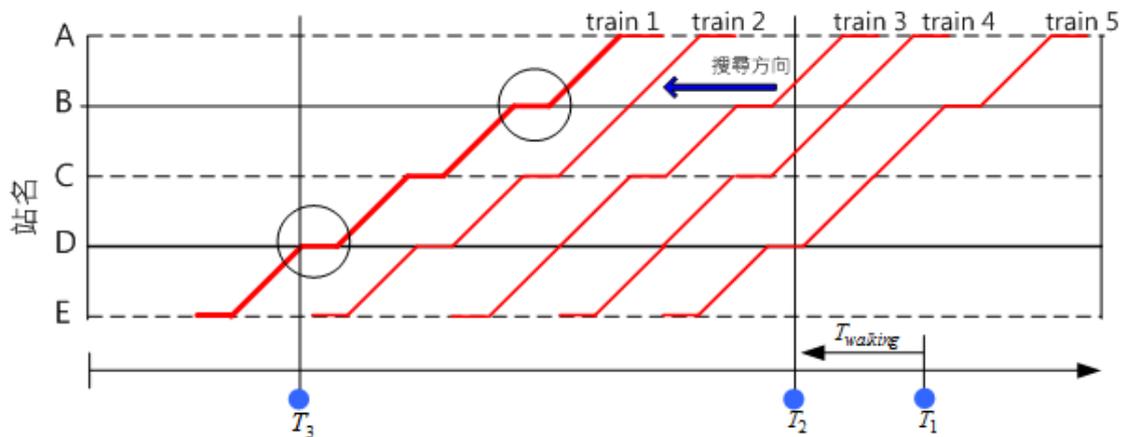


圖 4-93 臺鐵多卡通票證進站時間插補演算法示意圖

最後，如圖 4-94 的公共自行車的資料則在分析上最不具可行性，因為除借車時間闕漏且無法像臺鐵多卡通可利用其他資料（例如時刻表）插補之外，其借車還車地點之代碼亦無明確之地理座標資訊，且無法像公車票證可透過其他資料（例如公車動態）插補，屬於可行性最低的運具類別。

	運具	卡號	票種編號	票種代碼	借車時間	借車地點	還車時間	還車地點
1	公共自行車	cy1A8Y2z7.M8	1	1	NULL	3051	2016-03-02 01:00:00	3051
2	公共自行車	b5cJ8UWEBsP0	1	1	NULL	3052	2016-03-02 02:33:00	3051
3	公共自行車	HFTcClErywqD	6	6	NULL	3010	2016-03-02 11:09:00	3051
4	公共自行車	2Klor42VvuS.	1	1	NULL	3070	2016-03-02 12:09:00	3051
5	公共自行車	SPhJ13672lt1	1	1	NULL	3070	2016-03-02 12:09:00	3051
6	公共自行車	Ybxq2C.3ibi/	1	1	NULL	3053	2016-03-02 15:42:00	3053
7	公共自行車	Rp5Ma4q5Bh28	1	1	NULL	3027	2016-03-02 16:32:00	3053
8	公共自行車	tnojXG.sz7B5	6	6	NULL	3017	2016-03-02 17:00:00	3053
9	公共自行車	rL7PoEbjFaz5	1	1	NULL	3022	2016-03-02 17:07:00	3053
10	公共自行車	n0NE0tWXOR5	1	1	NULL	3024	2016-03-02 17:12:00	3053

圖 4-94 同編碼之公共自行車票證資料截圖

綜合上述分析探討，欲利用電子票證執行跨運具分析時，可能的資料來源包括票證公司與營運單位，不同蒐集來源各有優缺點整理如表 4.7。本研究採「票證公司」途徑蒐集，雖可輕易勾稽同一使用者（卡號）之各式運具交易紀錄，但相對也面臨部分營運細節資訊缺乏之挑戰，所幸本研究已分別針對公車、鐵路兩種運具闕如的欄位發展插補演算法作為後續初探分析之用。

表 4.7 不同資料來源電子票證跨運具分析優缺點比較表

資料來源	票證公司	營運單位
優點	易取得一致編碼之紀錄	資訊更為充分（例如站牌資訊、進站時間、開門位置等）
缺點	恐缺乏部分營運細節資訊	須逐單位蒐集 不易勾稽同一使用者行為

#### 4.6.3 跨運具轉乘分析初探

除了探討跨運具轉乘分析的可行性之外，本研究亦針對此一課題進行初步分析，以了解未來可能之應用潛力。同樣以本標竿案例的研究區域為對象，本研究根據悠遊卡公司提供之民國 105 年 1 月~3 月分票證資料，利用視覺化分析方法，分析臺鐵日南~追分及泰安~成功等 18 座車站，與臺中市公車之間跨運具轉乘的情況。

整體而言，使用悠遊卡於臺中市搭乘臺鐵的旅次中，約有 35% 的旅次有轉乘行為，如圖 4-95，其中搭乘臺鐵前後都有轉乘行為的旅次僅有 2.53%，先搭乘公車再轉乘臺鐵的有 9.93%，搭乘臺鐵後再轉乘公車的有 21.83%，是前者兩倍，可能是因為公車班次密度較高，民眾較易轉乘的緣故。此外，進一步從圖 4-96 可發現，不同票種之旅客的轉乘行為皆相似。

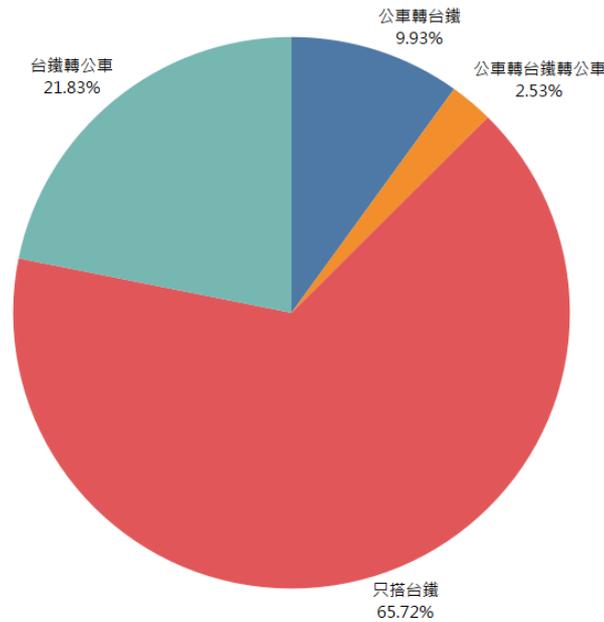


圖 4-95 一至三月整體搭乘行為

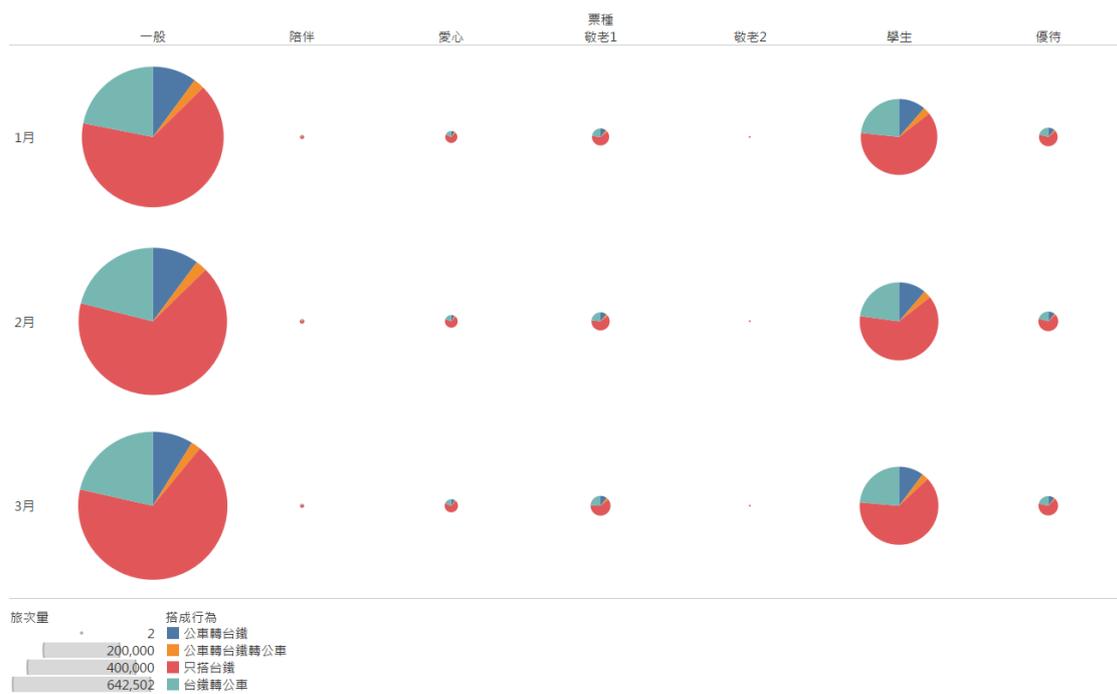


圖 4-96 一至三月各票種搭乘行為

從車站的角度來看，如圖 4-97，臺中站是轉乘比例最高的車站，以悠遊卡於臺中站搭乘臺鐵的旅次中，約有 25% 的旅次來自於公車轉乘，而使用悠遊卡搭乘臺鐵至臺中站出站之後，有將近 50% 的旅次會再轉乘公車，與前述臺鐵轉公車的旅次量是公車轉臺鐵的兩倍情況相符。

## 公車轉台鐵比例



## 公車轉台鐵

False  
True

## 台鐵轉公車比例



## 台鐵轉公車

False  
True

圖 4-97 一至三月各車站轉乘比例

若以時間的觀點，圖 4-98 顯示了 1 月到 3 月每天轉乘旅次量與平均轉乘時間的變化，在轉乘旅次量方面，一、二月分受到連續假期和選舉的影響，旅次量變化較為不定，而三月分由於沒有這些因素的影響，每週呈現規律性的變化；在轉乘時間方面，公車轉乘臺鐵普遍比臺鐵轉乘公車有較長的轉乘時間，研判是公車的班距較密，因此轉乘公車所需的時間較短。

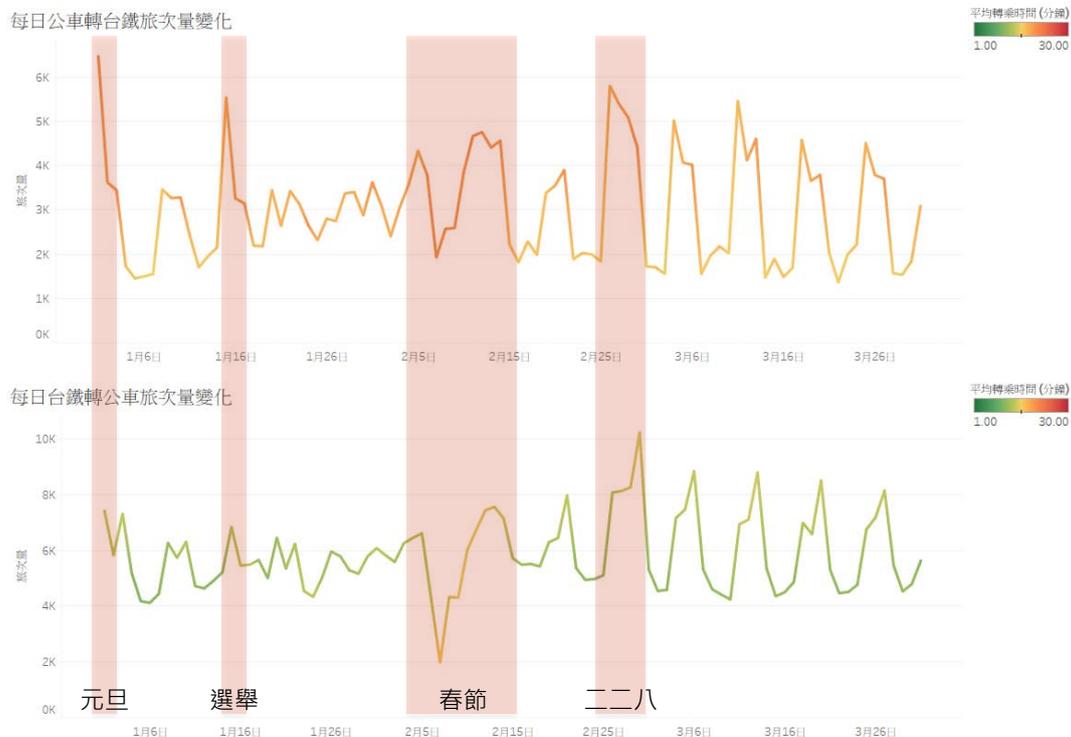


圖 4-98 一至三月每日轉乘旅次量與平均轉乘時間變化

由於三月分有週期性變化，因此針對該資料進一步檢視一周各日每小時的平均轉乘時間與轉乘旅次量的變化，如圖 4-99 所示。在轉乘旅次量方面，公車轉乘臺鐵有平日的下午尖峰，尤其週五的尖峰更為明顯，是其他平常日的兩倍以上，臺鐵轉乘公車則有平日的上下午尖峰，通常上午尖峰高於下午尖峰，只有週五的下午尖峰高於其他尖峰許多。在轉乘時間方面，幾乎在所有的時段中，公車轉乘臺鐵同樣比臺鐵轉乘公車有較長的轉乘時間，但在部分深夜時段是呈現相反的狀況，可能原因是公車在深夜時班次驟減。



圖 4-99 三月一周各日平均轉乘旅次量與平均轉乘時間變化

改以車站的維度來檢視各站的轉乘時間，其結果如圖 4-100 和圖 4-101，圖 4-100 為公車轉臺鐵的轉乘時間分布狀況，可發現臺中港站和泰安站的轉乘時間較其他車站要長；圖 4-101 則為臺鐵轉公車的轉乘時間分布狀況，泰安站同樣是轉乘時間比其他車站還要長。

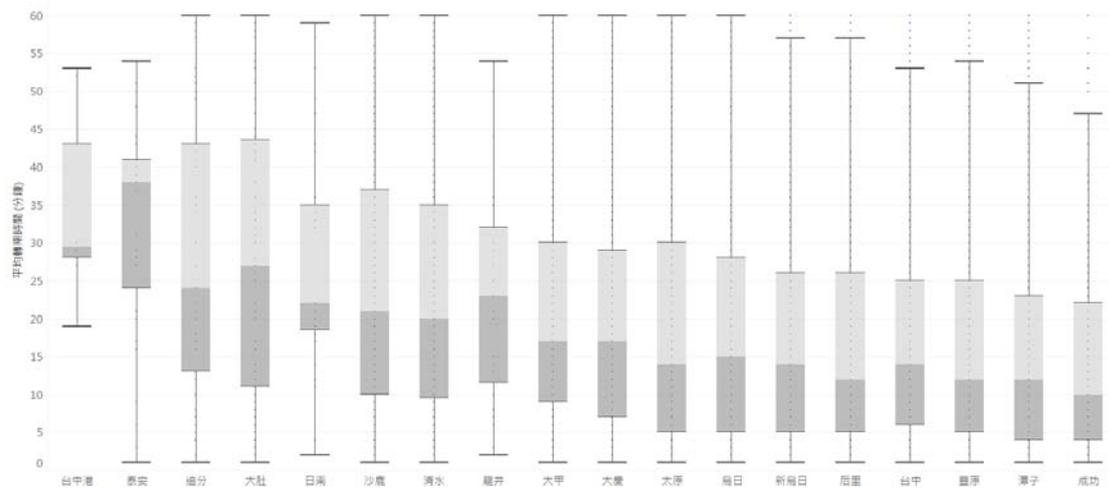


圖 4-100 三月各臺鐵車站之公車轉乘臺鐵的轉乘時間盒鬚圖

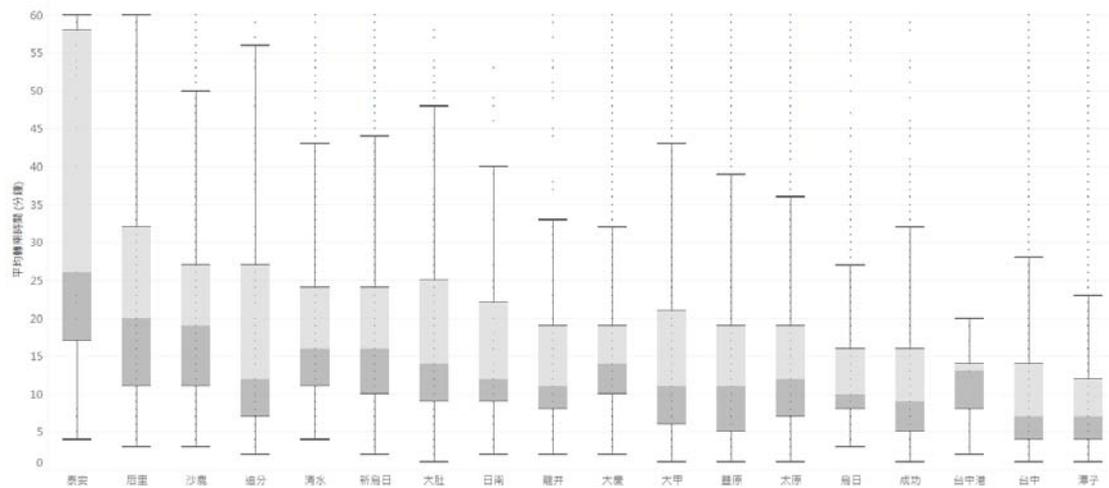


圖 4-101 三月各臺鐵車站轉乘公車之轉乘時間盒鬚圖

## 第五章 資料匯流平臺規劃

### 5.1 國外資料匯流平臺回顧

#### 5.1.1 商業公司匯流平臺案例

以下將依字母順序分別介紹 AirSage、INRIX、ParkMe 與 Taxi Stockholm 等四家商業公司匯流平臺之應用目的與功能。

##### 1. 案例一：AirSage 人/車流移動分析平臺

AirSage 是一間 2000 年成立於美國亞特蘭大的軟體公司，在過去數年中，跟美國兩家電信公司協議合作，希望擁有將硬體設備放進公司的防火牆內的獨家專利，以實時蒐集、匿名化、加密和分析通信數據，並於 2012 年完成協議，目前 AirSage 整合美國兩家電信公司中超過 1 億臺的用戶行動裝置所產生的資料，藉由其專利技術 WiSE (Wireless Signal Extraction) 每日擷取超過 150 億筆行動資料，並識別化處理成為匿名資訊，再結合人口統計數據等資料，提供即時的群眾流量彙總分析報告，提供政府部門或企業用戶的位置與移動分析情報，創造出嶄新的應用模式。

例如，一些手機機主在某個商業園區裡從早上 9 點待到下午 5 點，可以推斷出他們很可能是在那裡工作，因此，公路工程師就可以估計出，在該處高速公路出口的車流量中，有多少是由這些員工上下班貢獻的。如 Moore County 就透過 AirSage 的流量分析報告，透過演算分析做進一步的推估預測，掌握民眾交通模式，如每個月的特定時間，有多少人會經過哪些路段，再透過這些人的動線及居住位置，推估平均蒐入等資訊，進而作為城市中交通規劃決策參考。AirSage 除用於交通運輸規劃之外，也應用於旅遊觀光分析，以及展店評估、廣告行銷等零售業。

AirSage 產品所採用的數據挖掘功能已存在於行動電話網路中，只需安裝兩個不同的套裝軟體就可以實現，一個安裝在用戶系統中，實時訪問網路的數據，並將自己的移動紀錄發送到 AirSage 公司，另一個則安裝在 AirSage 公司，將蒐集到的手機用戶資訊和 GIS 系統融合，用來進行空間分析和製圖。

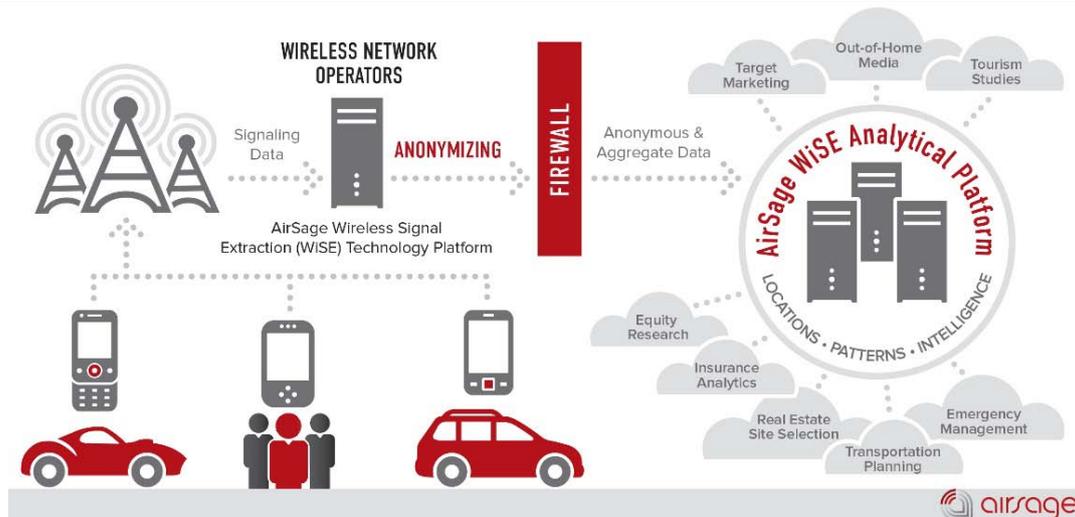


圖 5-1 AirSage 資料分析及應用平臺示意圖

諸如 AirSage 的巨量資料分析企業，已經成為許多研究及投資機構關注的對象，如 AirSage 已被國際研究機構 Gartner、科技網站 FierceWireless、美國成長最快速的私人公司排名 The Inc. 500|5000、美國行動通訊暨網路協會 CTIA 等單位，列為 2013 年值得關注的一家新創公司。AirSage 執行長史密斯表示，美國各級政府為了路況，每年花費 10 億美元（約臺幣 329 億 4500 萬元），但能掌握的路況不到全國大小道路的百分之一，以同樣的經費，該公司能服務的道路可增為百分之十。

## 2. 案例二：INRIX 交通路況預測平臺

INRIX 是一家位於西雅圖郊外的路況分析公司，專門蒐集整理北美及歐洲約一億臺車輛的即時地理位置資料，包括 BMW、福特、豐田等廠牌，也包括計程車及貨運車等車隊資料，除車輛資料以外，也包含智慧型手機、攝影機、道路感應器與車載導航系統等裝置。

INRIX 透過且善用所蒐集到資訊，並且推出免費應用程式，以匿名方式獲取 App 用戶所回傳之駕駛行為，包括地理位置座標、車速及發動頻率等即時訊息，每分鐘更新即時交通資訊，編譯成每九十天更新一次的歷史資料庫，透過這些龐大的交通數據，再將其他可能影響交通的因素，如節慶、假日、當前及未來的天氣、比賽、事故、道路建設等變數納入資料庫進行統計分析與比對，產出即時有效的交通資訊（包含事件資料、車流量、平均車速、自由車速、即時油價等），並且提供路況預測（包含預估旅行時間、最快行車路徑、事件預估解除時間等），將分析結果反饋給 App 用戶。



圖 5-2 INRIX 連結駕駛人網路示意圖

INRIX 將這些交通數據銷售給 GPS 生產商和各國的交通公共部門，在全球已累積超過兩百個合作伙伴及客戶，例如奧迪、ADAC、ALK、ANWB、Coyote、福特、英國廣播公司、MapQuest、微軟等。對於一般用戶而言，透過 INRIX Traffic 的數據分析服務，可即時獲取交通新聞、地圖、影像、天氣、封路、施工及意外事故等訊息，並可查詢行駛路段臨近之加油站、車位以及燃油及停車費比價資訊，讓用戶有效節省交通時間與花費。

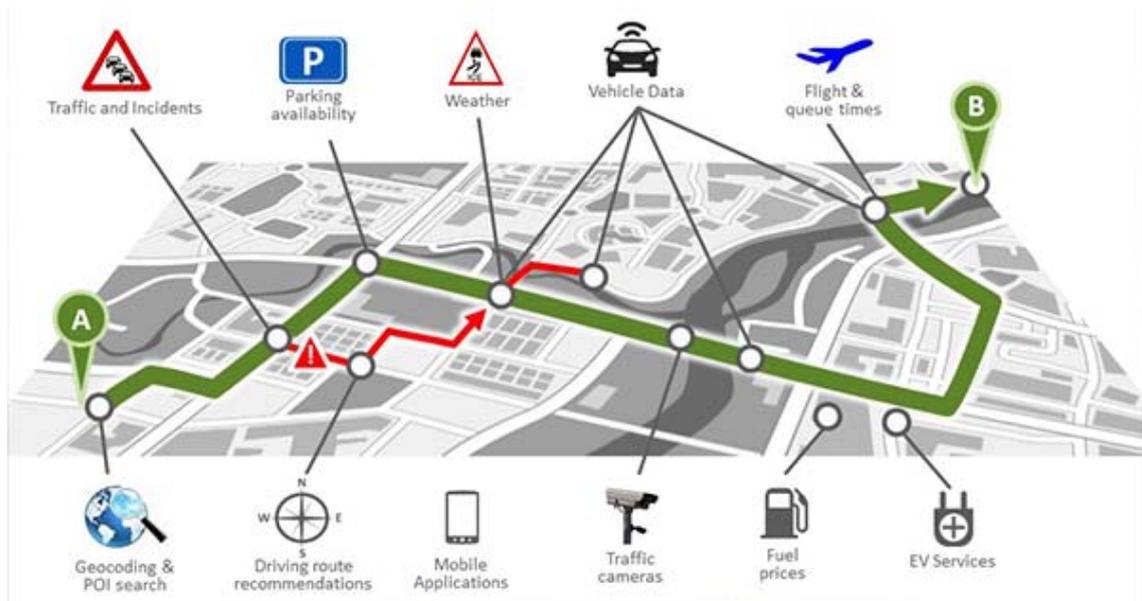


圖 5-3 INRIX 提供駕駛人服務示意圖



目前 INRIX 的資訊服務範圍以北美及歐洲為主，橫跨全球 40 個國家，涵蓋產業包含汽車及電子通訊設備製造商、政府機構、行動應用供應商、無線運營商、房地產業以及媒體公司等。25 個美國的公共運輸部門在日常工作、旅行服務、交通阻塞評估和道路管理規劃時都會使用 INRIX 提供的交通資訊。2012 年，北美、英國、德國和奧地利使用 INRIX 交通資訊和分析幫助他們用更低的成本來管理和監測路網。

### 3. 案例三：ParkMe 智慧停車平臺

隨著全球都市化程度越來越高，都市人口集中，經濟發展帶動汽車數量，停車成了一門大生意，根據 Frost & Sullivan 的報告，全球停車商機高達 1000 億美元（約合 3 兆臺幣），約等於全球半導體銷售額的三分之一，在未來 3 到 5 年將吸引多達 2.5 億美元的投資金額。根據交通部統計，臺北市有 16.5% 的駕駛人每天要花半小時找車位，平均每個人每天要花 14.5 分鐘，找車位所需時間居全國之冠，2014 年，臺灣人平均花 9.2 分鐘在住家附近找車位。

ParkMe 創辦於 2009 年，是全球規模最大的停車 App，ParkMe 可以顯示全球 64 個國家、超過 3200 個城市，多達 9 萬個地點內超過 3 千萬個停車位資訊，去年（2015 年）9 月被交通數據分析公司 INRIX 併購，為奧迪和豐田等車廠內建導航系統提供停車資訊，被美國知名科技媒體 Business Insider 譽為年度市場上最熱門 15 個應用之一。顧客可以透過 GPS 及車內導航系統、ParkMe 官網（<https://www.parkme.com/los-angeles-parking>）或 App（包含 Android、iOS 及 Windows Phone）得到 ParkMe 提供的資訊。

ParkMe 共同創辦人 Sam Friedman 起初為了蒐集停車位資訊，跟夥伴們騎著腳踏車四處在洛杉磯市內晃來晃去，看到停車位就拍照作記錄，用最土法煉鋼的方式充實車位資料庫，從 2009 年發展至今，Parkme 已與全美幾乎所有連鎖停車場建立合作關係，直接抓取大型停車場的即時數據，早已不需要再像當年一樣騎著腳踏車到處拍照，但 ParkMe 仍會不時派專人到每個停車場拍照，定期上傳各設施的狀態，檢查車位與停車場廠商提供的資訊是否有出入。

目前 ParkMe 的數據來源除了連鎖停車場，還包括市政府和用戶自主回報，後兩者主要用於路邊停車位，在裝有攝影機的城市，ParkMe 會分析攝影機街拍影像，即時更新路邊車位資訊，其他資訊蒐集工具還包括路邊感測器和智慧停車蒐費表（smart meter），用戶若自主回報停車位，可獲得車位預約折扣點數。目前與 ParkMe 合作提供車位資訊的公司約 500 家，

絕大多數的停車場業者和政府單位都免費將資訊提供給他們，有的還反過來付錢給 ParkMe，因為這些業者發現跟 ParkMe 合作可有效增加停車流量，而市政府發現這有助於解決市區交通壅塞問題。

ParkMe 極盡所能地彙整最大量且即時的停車位數據，不使用感應器查明空位，取而代之的是用一預測算法來計算可能會有空位的地方，提供車位快照讓使用者在需要停車時，只要打開 App 就能找到離自己最近而且價格最理想的停車位，而且除了車位照片，ParkMe 也提供了鄰近區域停車平均價位、車位評價、剩餘車位和聯絡資訊，駕駛人如果覺得現場找車位壓力太大，還可以事先在 App 上預約車位和付款，ParkMe 以綠、橙、紅三種不同顏色的圖示說明停車位的狀況以及價格，點選圖示後可以檢視該停車位的資訊。

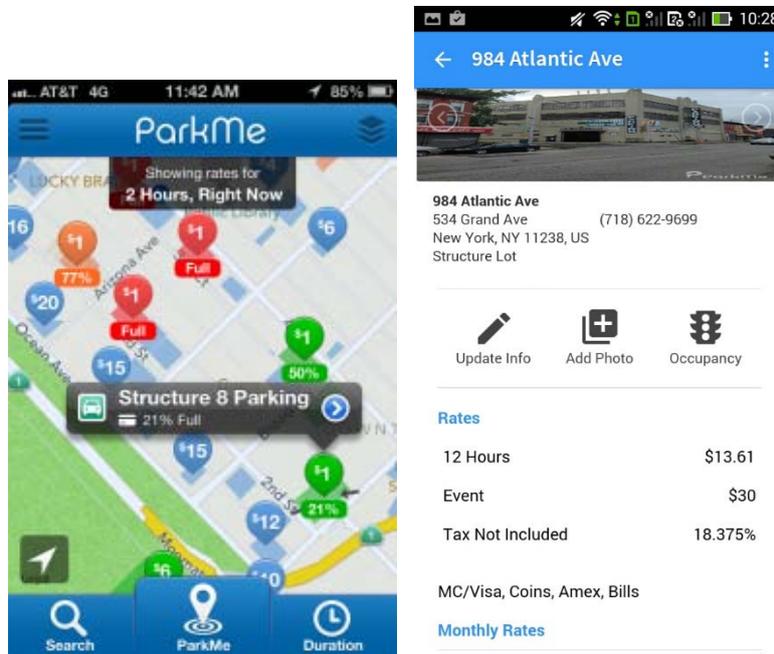
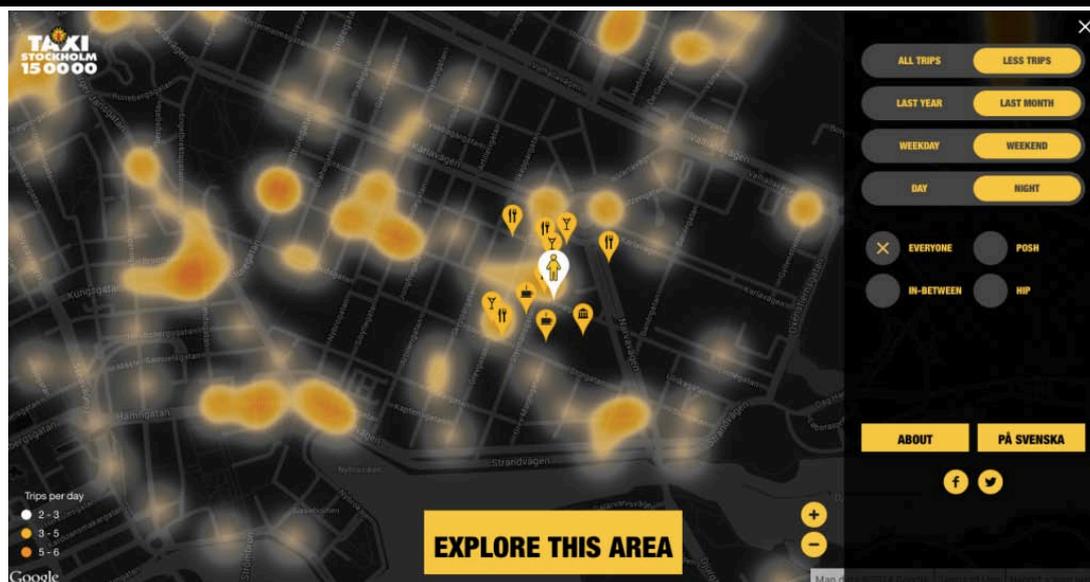


圖 5-4 ParkMe App 功能畫面

#### 4. 案例四：Taxi Stockholm 計程車資訊平臺

Taxi Stockholm 是一家位於斯德哥爾摩的計程車營運軌跡資料平臺，此平臺係利用公司車隊每年高達 8 百萬筆載客路徑的 GPS 資料所設計而成，在地圖上記錄當地人以及觀光客們在斯德哥爾摩實地造訪的景點和路線，如圖 5-5，再搭配熱度圖（Heat Map）按時間條件分類，使用者可以找出去年或上個月、白天或夜晚、週間或週末甚至各種類型名人的私房旅遊路線；找出熱門旅遊景點之後，使用者可以接著利用 Google 街景功能，探索周遭環境以及熱門景點。



資料來源：<http://www.taxitrails.se/en>

圖 5-5 Taxi Stockholm 計程車大數據應用平臺

### 5.1.2 城市層級匯流平臺案例

以下將分別介紹英國(倫敦)、瑞典(斯德哥爾摩)、新加坡(全國)、菲律賓(宿霧)、美國(全國)與中國(貴州、珠海)等城際間推動的資訊平臺與應用目的功能。

#### 1. 案例一：倫敦智慧交通相關平臺

倫敦當局認為發展數位城市注重的是通信技術，必須對整個城市的生態系統發揮正向作用，並為人們生活與工作帶來切實影響。在倫敦市裡，所有地鐵站都有免費 Wi-Fi，乘客可以用手機蒐到各種在地便利資訊；不論是博物館、藝術中心、歌劇院或是酒吧，大多都會有對應之 App 服務，讓更多人可透過網路體驗場館的各種服務；公共圖書館不論是實體或網路都有豐富圖書資源；幾乎所有的市政府區級機構，公共事業組織都會有官方網站，提供政務訊息和線上服務窗口，加上官方 Facebook、Twitter、YouTube 與 Digg 等最基本之設置；所有文化、藝術、科學等資源，都可以通過網路找尋相關頁面，獲得豐富詳盡的資料訊息。

##### (1) 史上最高科技的奧運：利用 Big Data 智慧網解決交通混亂

倫敦交通局 (Transport for London, TFL) 為了倫敦奧運的舉行，除了啟用奧運聯合路網 (Olympic Route Network) 之外，還利用 Big Data 協調大批人流的行動。

精密的交通控制管理「Split Cycle Offset Optimisation Technique (Scoot)」應用於 2,100 個重要路段，並將 1,300 個路段的號誌重新設定為依據即時交通流量進行適應性調整，可監測與控制倫敦交通系統立即做出疏導方案。例如：奧運主要幹道若需要疏導，在某些匯集路段會延長紅燈時間。

各路段 CCTV 影像也隨時利用軟體分析各類元素。其中一個資料來源「Oyster Card」智慧卡，提供人流資料幫助倫敦交通局預測通勤尖峰與旅行動態，直接對用路人提供路況預報，如在未來的幾小時內某路段的通行狀況是否良好。

另外，運用免費的「Inrix」交通 App 能輕易的在奧運時期觀賽與城市觀光。Inrix 是研究交通資訊的大數據專家，在奧運期間提供即時交通訊息，利用多元資料來源整合交通資訊，協助駕駛人規劃最有效率的路徑，幫助旅客了解路況，避開壅塞路段，做出旅行路線的最佳選擇。

## (2) TUBE：倫敦市區即時動態地圖線上查詢系統

由 Matthew Somerville 根據倫敦市政府線上開發數據中心的 API Transport for London，然後將數據與 OpenStreetMap 一起建構而成。該網站能免費查到倫敦市區所有的地鐵線、地鐵站，並即時更新，紅色的點代表的是地鐵站，處於不斷移動黃色的點代表正穿梭於倫敦市中心的地鐵。通過此地圖可以查詢，附近的地鐵站靜態位置以及列車動態位置並即時更新。



資料來源：<http://traintimes.org.uk/map/tube/>

圖 5-6 倫敦市區地鐵即時動態地圖

### (3) 尖峰時段預約下載停車位

在倫敦交通高峰時間，停車位通常是一位難求，為此人們開始通過網路即時查詢停車場空位情況，並進行車位預訂。JustPark 公司與 BMW 旗下之 MINI 品牌合作，提供一款行動 App，與 MINI Connected 完成連結，再輸入目的地後，系統便會找尋目的地周遭的停車場，預約車位並提供導航，大幅省去找尋車位的麻煩，同時確保抵達時有車位可停。

### (4) 數位政府：城市開放型數據中心

城市網路數據中心的建立，對現代化城市的發展具有重大戰略意義。高速網路資料中心可以使城市的政府，企業，及公司客戶以較低的技術成本傳輸、使用和複製數據，並且能安全可靠地來傳輸加密數據。倫敦市政府早在 2000 年即提出打造「電子政府」概念，逐漸消除數位鴻溝。隨著雲計算時代以及智慧城市遠景規劃的逐步成熟，英國政府已經開始逐步轉向智慧資料庫的共享與開放應用。

英國大倫敦市政府 (GLA) 指揮跨行政區的部門—如交通、安全、經濟發展和旅遊業等執行「城市資料中心共享」方案。它指定倫敦市的各級機構、公務員和其他數據捐助者把數據積累到一個公共資料庫。許多數據格式並非可立即被使用，但網路科技玩家們迅速將這些數據轉化為民眾可以理解的圖形與圖表、地圖和文件，經過切割和分流將其發佈到個人電腦和其他電子裝置。有一個稱為「開放知識基金」的組織，創建名為「WhereDoesMyMoneyGo.org」（我的金錢流向）的網站，另一個組織，「科技推桿有限公司」，目前已在倫敦地區建成「OpenlyLocal.com」（開放當地數據）的網站。他們都是從這個資料儲存中心介接龐大數據。

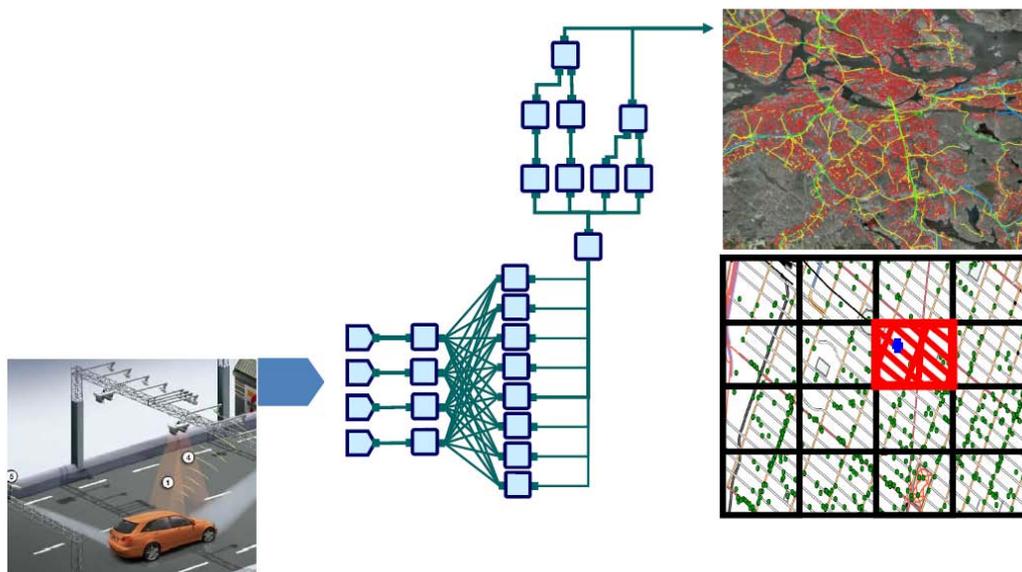
即使是業餘愛好者也正在利用開放資料編寫有用的手機應用程式，有效促進倫敦融入自由而創新之發展體系。資料開放鼓勵更多創新應用，把數據開放給能夠把事情做好的人，這可能比政府部門提供相同服務更有效率。

## 2. 案例二：斯德哥爾摩之旅行時間預測與道路蒐費系統

斯德哥爾摩為瑞典首都，也是瑞典第一大城市，是北歐的金融、商業、文化中心，瑞典政經中心，人口約有 82 萬，斯德哥爾摩委託 IBM 利用交通大數據打造交通即時預測系統如圖 5-7 所示，透過蒐集探偵車 GPS、線圈偵測器（速率與流量）、即時天氣資料以及員警提

供之資料作為輸入，進入數據分析與運算邏輯後，可產出旅行時間預測資料，進而作為徵收城市擁擠費之計算基礎。

運用大數據分析建置道路收費系統，利用 RFID 與車牌辨識技術，監測車輛進入市中心收費區（Charging Zones）之進入與離開時間，並據以收取不同的費用；收費後之市中心交通流量銳減 25%，使用公共運輸使用人數增加，且停車罰款降低了 29%，道路交通廢氣排放量減少了 8% 至 14%，市中心的溫室氣體排放量也下降了 40%。



資料來源：<http://www-07.ibm.com/tw/dp-cs/smartercity/solutions.html>

圖 5-7 IBM 協助打造交通及時預測系統示意圖

### 3. 案例三：新加坡 FASTER 計畫

為舒解新加坡民眾受地鐵列車故障、延遲事件困擾，新加坡陸路交通管理局（The Land Transport Authority, LTA）、SMRT、IBM 以及電信公司星和（StarHub）於 2014 年 6 月初展開 2 年的 Fusion AnalyticS for Public Transport Emergency Response（FASTER）計畫，應用大資料來減低列車故障等狀況對民眾造成的影響。首先 LTA 和 IBM 得到 StarHub 所蒐集的通勤者的去個資化、定位資料，還有 SMRT 提供的重要地鐵站錄影資料、儲值卡交易資料，再分析蒐集到的資料，來幫助 LTA 和 SMRT 處理人潮擁擠問題，像是透過發送警報告知乘客，哪一個區間的通勤狀況可能會有延遲，讓乘客能改變他們的交通路線。

透過儲值卡資料能得到有多少人正進出地鐵、電信蒐集的資料則能提供民眾是從何處來，還有民眾在離開地鐵站後，是前往何處的訊息。這使

得 FASTER 計畫能對交通網路提供更有效的管理，例如，在哪一個月臺、哪一輛列車的人數擁擠程度，還有民眾要花多少時間才能搭上列車。

#### 4. 案例四：菲律賓宿霧 Open Traffic

由 2011 年起世界銀行與宿霧市政府，共同發展出利用開放原始碼平臺去蒐集、視覺化及分析計程車上行動電話的速度資料，在完成初期的專案成果，後續在世界銀行 Big Data 創新獎助支持下，本專案過去在六個國家進行測試並改善原先的專案成果，目前已在宿霧市上線使用。其交通數據的蒐集有別於其它地方建置車輛偵測，而是利用既有的 WiFi 及 GPS 建構一個大型的感測網路，來廣泛蒐集市區道路資訊。

其資訊蒐集方式，是以 Open Street 圖塊 (tile) 為基礎單元，在於圖塊上設置虛擬的偵測區域，再透過演算法配合蒐集到的速度位置資訊，計算出跨偵測區的旅行時間，並提供計算後的交通資訊供民眾使用。

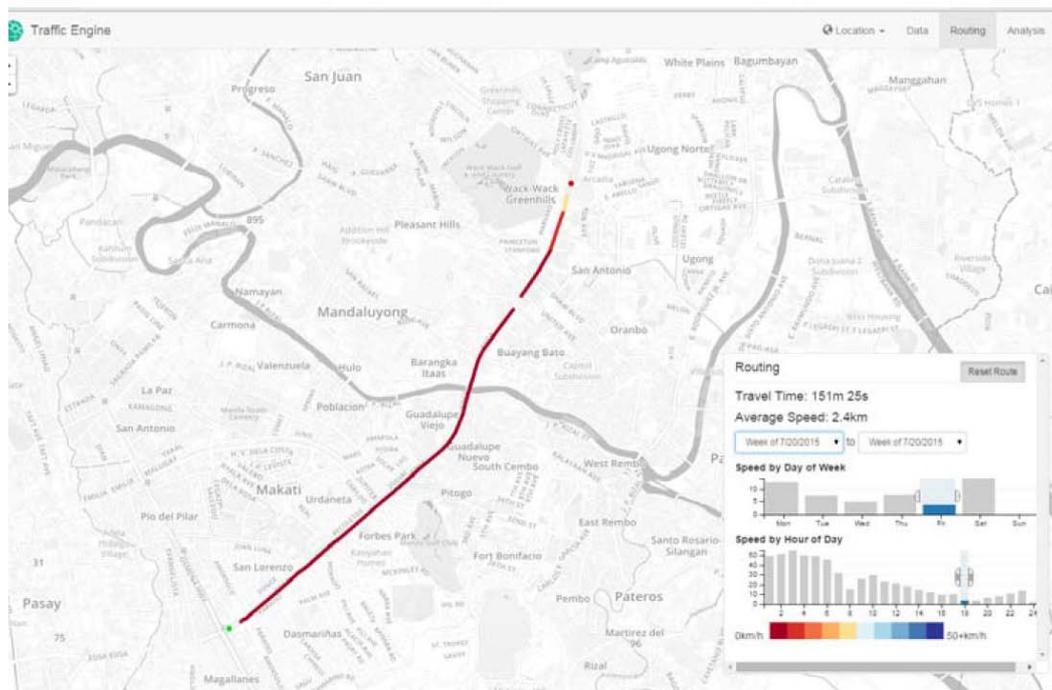


圖 5-8 往北方向平均旅行時間及速度 (2015 年 7 月 24 日)

#### 5. 案例五：美國公路戰略研究計畫 (Strategic Highway Research Program)

SHRP 美國公路戰略研究計畫(1989-1993)，是一項五年產品為導向的重點在瀝青、路面性能、混凝土與結構和公路運輸，四個領域的研究計畫，因為在 20 世紀末期公路的科學研究投入較為薄弱，因此 SHRP 的目標是為 21 世紀準備更好的公路技術，讓公路能更安全、更耐久。

該平臺資料蒐集方式係透過由各單位自主性提供相關研究數據於平臺中提供需求者進行分析、探討與回饋，且平臺中的相關研究計劃是公開的，透過研究報告、研討會、影片、相關網站連結，將研究的過程、成果、與實際應用情形等呈現在美國聯邦公路局與美國運輸研究委員會等相關網站上。

而近期 SHRP 2 (2006-2015 年)是 SHRP 研究計畫的延伸，該研究計畫主要是透過美國聯邦公路管理局、美國公路及運輸員協會、美國運輸研究委員會與交通相關產業、學界等，彼此間的相互合作，創造一個更具創新、實用與更成熟的工具，來為老齡化的國家公路之安全、交通堵塞、道路和橋樑的革新方法尋找解決的戰略方案，研究主要集中在安全、革新、可靠性、乘載量四大主題，以保護生命、減少壅塞、提高生活品質為目標，透過團隊的探討與分析，找出其解決方式。

- (1) 安全：透過公路安全的分析數據，為道路安全提供先進技術。
- (2) 革新：利用橋樑、路面、鐵路等相關交通技術性的管理，來提升更好的運輸效能。
- (3) 可靠性：利用研究人員間相互討論與實際的數據分析，透過其結果分析，進而研擬提高更具效率的運輸流程，並運用於交通相關部門，其研究流程如圖 5-9 所示。
- (4) 乘載量：透過經濟影響分析、網路、高速公路的養護規劃與環境整合，創造一個更好的運輸動態流程。

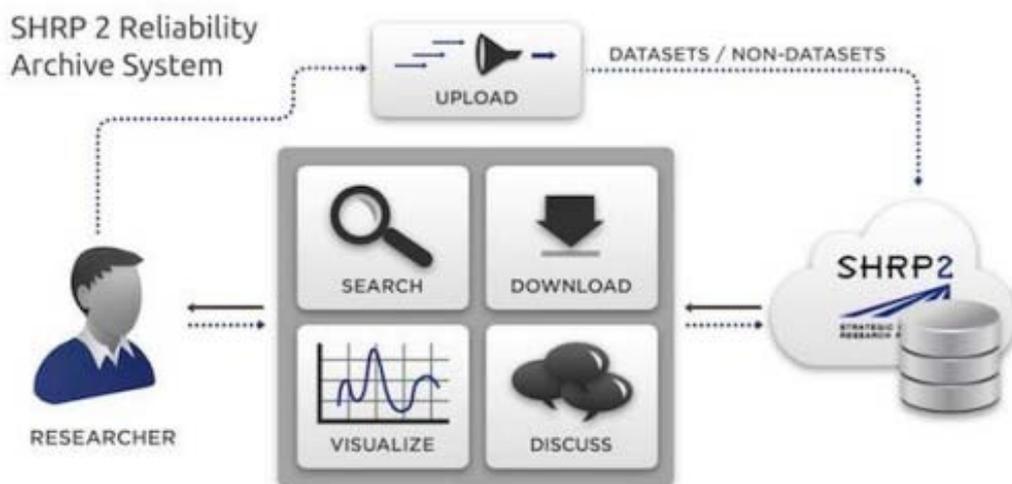


圖 5-9 SHRP2 在交通運輸可靠性上研究計畫流程示意圖

在該研究中概括了各種的運輸領域，包含風險管理、運輸系統管理與營運、公共事業、鐵路、路面材料與施工等。例如，提高交通運輸(包括交通事件的反應以及交安響應)，來減少交通堵塞。或是提供如何興建更具效用的路面建設等，讓全新的技術能更廣泛且有效地應用於交通運輸上，其交通運輸改善計畫分析流程如圖 5-10 所示。

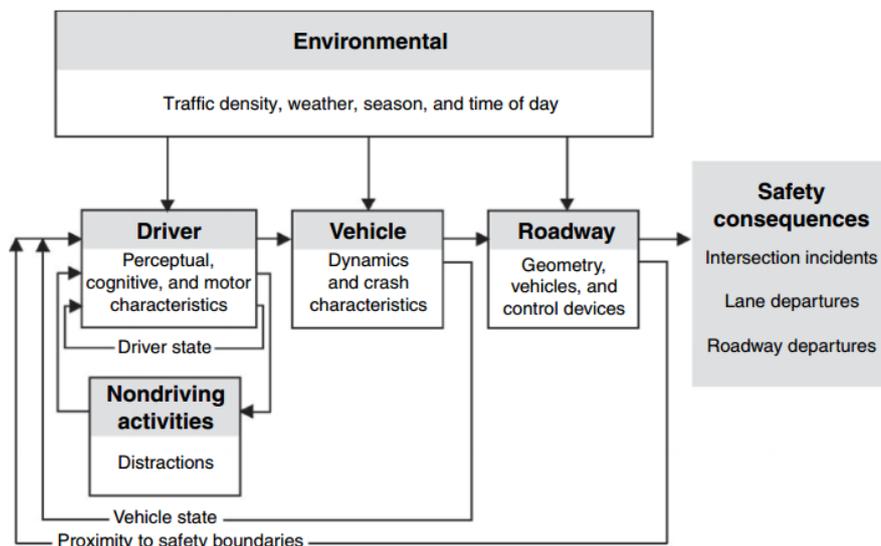


圖 5-10 交通運輸改善分析示意圖

如遇到道路的蒐費問題，想要了解如何透過交通運輸的疏導來解決道路乘載量的增加等問題，該問題由提問者拋出問題後，對此議題有興趣者則可提出多種解決方案進行討論，其中包括可透過交通與旅客需求的數據分析等，提升交通營運等方式，逐步蒐斂議題、歸納方案、提出建議、解決問題，其目標在於歸納出共識進而提出解決方案，此研究作法關聯方式，如圖 5-11 所示。

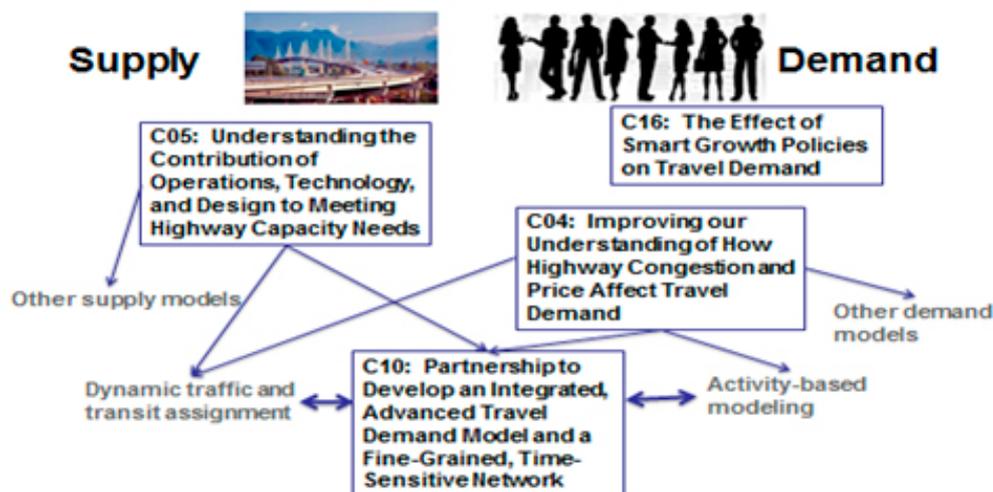


圖 5-11 SHRP2 需求與供給研究示意圖

## 6. 案例六：中國貴陽大數據交易所（GBDEX）平臺

貴陽大數據交易所於 2014 年 12 月 31 日成立，為全球首個大數據交易所，交易所發展的三大要素為大數據的清洗、大數據的分析與建模以及大數據的交易，而數據的定價將由數據賣方與交易所進行協議定價，最後出售給買方，因此交易所定位為數據交易輔助機構，協助完善交易機制，協助會員完交易，未來隨著物聯網的發展與應用，屆時數據交易亦將成為常態。

目前交易平臺規劃有 30 種類型的交易數據，包含有金融、政府、醫療、能源交通等，其中以銀行徵信需求數據為主，並強調交易內容非原始數據（Raw Data），而是經過清洗、建模、分析及可視化後的結果，如圖 5-12 所示，數據交易系統如圖 5-13 所示；交易資格則須審核通過成為會員才具有於平臺數據買賣資格，而賣方要保證數據所有權、合法及可信，買方則須保證不濫用且尊首賣方使用授權，目前數據交易以內資企業為主，因此臺灣企業僅據有賣方資格，無法成為買方。而數據的交易形式則包含了 API、終端及在線三種，其中 API 對於買方而言可累積歷史數據，因此售價最高；數據定價則由買賣雙方協議、拍賣與集合競價三種方式進行，而交易所則向賣方蒐取交易交額 40% 作為手續費，買方則無須給付手續費。

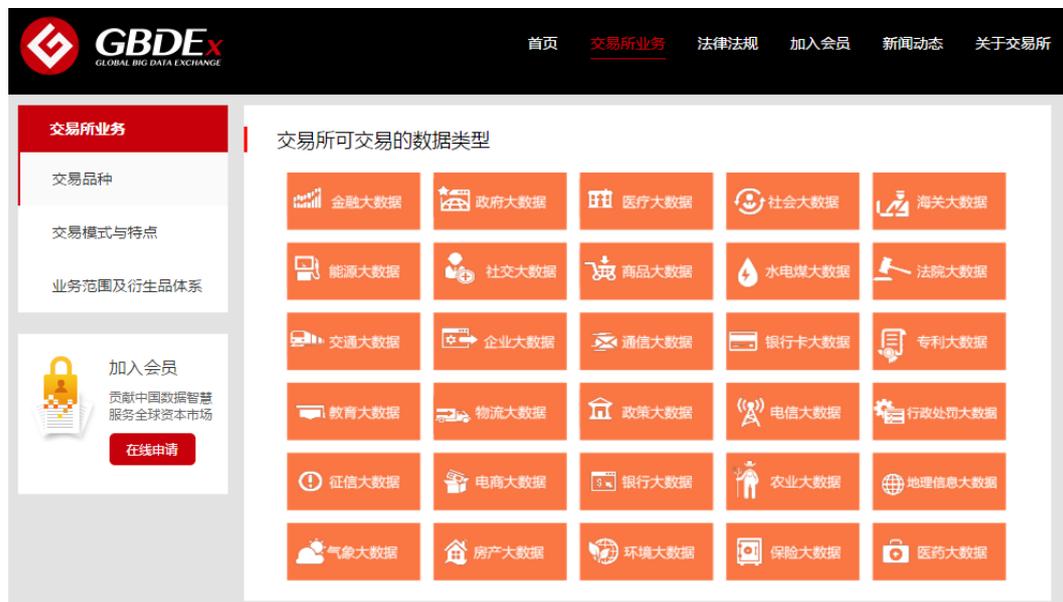


圖 5-12 貴州交易所交易數據類型



圖 5-13 貴州交易所數據交易平臺

### 7. 案例七：中國珠海智慧交通系統

珠海為中國最早實行對外開放政策的 4 個經濟特區之一，據珠海市統計局數據顯示，珠海人口在 2009 年至 2014 年間，只有 8% 的增長，但城市擁有汽車的數量卻增加了 75%，隨著經濟的深化發展，日益凸顯城市交通所面對的挑戰，珠海以智慧交通來解決問題，由西門子公司配合中國重點發展綠色交通，提供智慧交通解決方案，投資發展綜合交通管理項目。整個項目投資涉及 1,500 萬人民幣，資金主要用作開發硬體，例如大數據處理服務器，以及應用軟體等。

為達到智慧管理西門子與珠海交通運輸局合作，整合了交通管理部門、建設部門、營運部門等，實地蒐集分布在珠海大街小巷的攝影機、感測器、車載 GPS 等大數據，再加以篩選、分析，將這些數據轉化成有用的智慧數據，協助交通管理者施行交通控制、路況分析、交通預測、公車調度、事故管理等工作，減低交通擁擠帶來的壓力。



表 5.2 城市整合資料匯流應用目的比較表

城市	系統/專案簡稱	應用目的
英國(倫敦)	開放型數據中心	人/車流移動分析平臺
	JustPark & BMW	尖峰時段預約停車
	TUBE	利用市府 API 提供動態地鐵地圖查詢
瑞典 斯德哥爾摩	—	旅行時間預測與道路蒐費系統
新加坡	FASTER	列車故障/延遲之通報與旅客應變
菲律賓宿霧	Open Traffic	開放源碼平臺蒐集資料並預測旅行時間
美國	SHRP	為交通運輸的革新尋找解決的戰略方案
中國貴州	GBDEx	大數據交易所
中國珠海	智慧交通系統	路況分析/交通預測/公車調度/事故管理

資料來源：本研究整理

從上述回顧這些平臺所蒐集的資料與應用中都有共同之特色，利用大量的數據資料進行分析、比較，與應用於交通管理方面，將匯集與應用兩方面同時進行考量，並非單一僅蒐集數據之用。

此外，兩類平臺也略有差異，商業公司平臺的應用目的通常較為集中，且都會搭配手持裝置 APP 提供服務，而城市層級的平臺則偏向提供原始資料或 API 供開發者應用，即使有應用分析也是較具有監理、交管等公共利益性質之應用，這些特性都將納入本研究之參考。

## 5.2 國內交通數據平臺回顧

### 5.2.1 盤點案例

#### 1. 案例一：交通部國道高速公路局交通資料庫

本資料庫彙整了國道即時及歷史交通資訊，免費對外開放使用，提供用路人更多資訊取得管道，交通資訊內容包含了「高快速公路整體路網交通管理系統」與「高速公路電子蒐費系統」之資料。其中電子蒐費系統由多套子系統整合，協同運作以完成電子蒐費。系統的運作包含車輛偵測與車型辨識子系統（包含車速判斷）、扣款系統（紅外線或微波）、執法

系統（包含照相機與攝影機執法模組）、前端的車上單元（OBU）以及後端的帳務資訊系統等，系統示意圖如圖 5-15 所示。

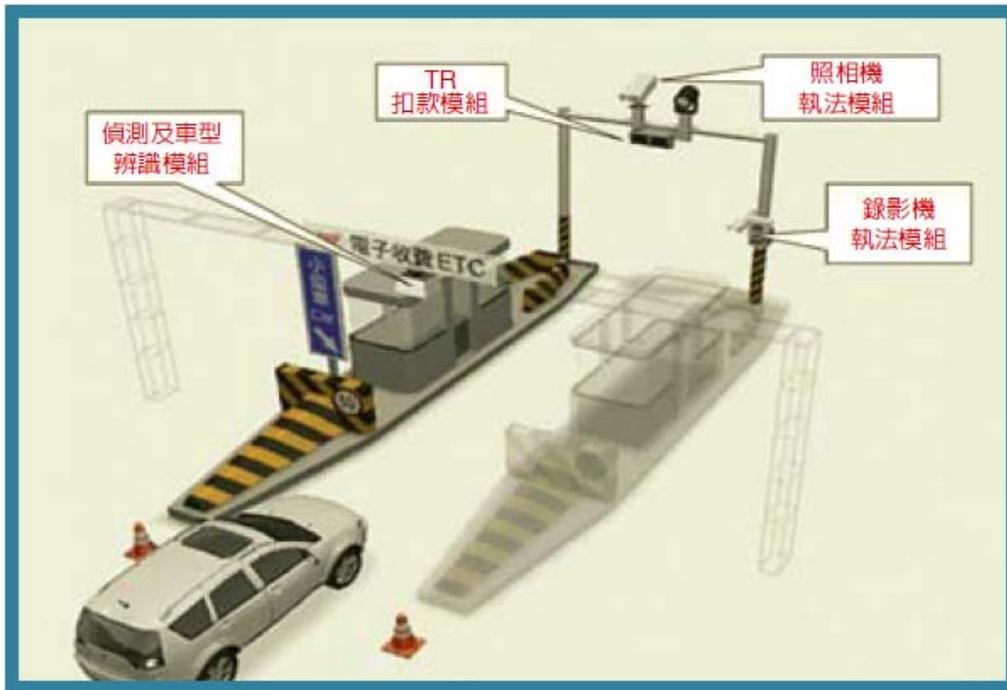


圖 5-15 電子蒐費系統示意圖

配置 OBU 的車輛通過車道系統時，在極短的時間內，透過短距離通訊（DSRC）協定可完成電子蒐費卡片扣款。當車輛通過車道系統（RSU）時，車型辨識模組的雷達偵測器會感應到車輛已進入蒐費區，架設在車道上方的車輛感知偵測器（TOM 2000）就以六道雷射波來偵測反射波的波形，以判斷車輛的車型，同時計算車輛的速度，並通知扣款模組、執法模組、錄影存證模組等開始動作，使各模組產生的資料有一個同步可以匹配的參考點。當各個子系統都完成動作後，再把資料送交主控模組（Main Control Unit, MCU），主控模組負責整合每個動作，把每一輛車通過車道系統的資料包含扣款資料、執法資料與車型辨識資料整合成一筆，產出交易紀錄檔提供後臺帳務系統進行後續的資料處理。

高速公路 ETC 電子蒐費系統，目前高速公路上的 319 座蒐費門架，每天記錄約 1,400 萬筆交易資料，迄 2015 年 9 月底已累積記錄 89 億筆通行費資料，ETC 大數據的運用方向包含：

- (1) 分析車輛起迄點：了解各交流道車輛出入數量、找出尖峰時段
- (2) 指揮匝道儀控設備：進行車流量進入高速公路管制、避免塞車
- (3) 分析路段運量：據以制定路段或時段差別費率

## (4) 預測行車時間：預測交通狀況、行車時間



圖 5-16 臺灣高速公路電子蒐費系統

## 2. 案例二：交通部公共運輸整合資訊流通服務平臺

交通部公共運輸整合資訊流通服務平臺（<http://transport.tw/PTX/>）以雲端架構與服務模式為基礎，發展一公共運輸資訊與資料的共享平臺，使得各種運輸相關之資訊與資料得以透過標準格式相互流通。

由於各項公共運輸旅運資訊系統由各縣市交通主管機關或相關管理單位建置及維護管理，此平臺由交通部制定相關資料標準，促進各政府機關所發布公共運輸資訊格式之統一，提供未來相關單位以及民間增值廠商得以統一格式接蒐公共運輸旅運資料，以下針對此平臺提供之服務進行介紹：

## (1) 資料服務查詢

- A. 運輸分類查詢：提供平臺使用者以不同運具類型進行屬於該類型相關資料內容的查詢，如圖 5-17 圖 5-18，目前平臺提供之資料類型包含公車、高鐵、臺鐵及航空四類運具，各項運具資料內容彙整如表 5.3 所示。



The screenshot shows the PTX website interface. At the top, there is a search bar and navigation links. The main content area is titled '資料服務查詢' (Data Service Query). On the left, there is a sidebar with '運具類型' (Transport Mode) and a list of options: 航空 (Air), 基本服務 (Basic Service), 公車 (Bus), 高鐵 (High-Speed Rail), and 臺鐵 (Taiwan Railway). The '航空' option is selected. The main area displays a table of search results for '航空' (Air).

服務名稱	運具類型	資料類型中文	資料提供單位	引用次數	說明	網址
航空公司資料	航空	業者資料	交通部民用航空局	20,646		
機場即時出境航班	航空		交通部民用航空局	19,989		
機場即時入境航班	航空		交通部民用航空局	19,913		
機場即時出入境航班	航空	到離站	交通部民用航空局	138,035		
機場資料	航空	站點	交通部民用航空局	18,947		

圖 5-17 以運具分類查詢公共運輸資料

The screenshot shows the API documentation for the 'Airline' service. It includes the endpoint 'GET /v1/Air/Airline', implementation notes, response class details, and a table of query parameters.

**Implementation Notes**  
取得所有航空公司資料

**Response Class (Status 200)**  
Model | Model Schema

**Inline Model [ Airline ]**

- Airline {**
- AirlineID (string):** 航空公司IATA國際代碼,
- AirlineName (NameType, optional):** 航空公司名稱,
- AirlineNameAlias (NameType, optional):** 航空公司簡稱,
- AirlineIATA (string, optional):** 航空公司IATA國際代碼,
- AirlineICAO (string, optional):** 航空公司ICAO國際代碼,
- AirlineEmail (string, optional):** 航空公司電子信箱,
- AirlineAddress (string, optional):** 航空公司地址,
- AirlinePhone (string, optional):** 航空公司聯繫電話,
- AirlineNationality (string, optional):** 航空公司國籍,
- UpdateTime (string):** 資料更新日期時間(ISO8601格式:yyyy-MM-ddTHH:mm:sszzz)
- }**
- NameType {**
- Zh\_tw (string):** 中文繁體名稱,
- En (string, optional):** 英文名稱
- }**

**Parameters**

Parameter	Value	Description	Parameter Type	Data Type
ticket	<input type="text"/>	預設為guest的ticket	query	string
\$select	<input type="text"/>	挑選	query	string
\$filter	<input type="text"/>	過濾	query	string
\$orderby	<input type="text"/>	排序	query	string
\$top	30	取前幾筆	query	string
\$skip	<input type="text"/>	跳過前幾筆	query	string
\$format	JSON • 列舉詳細資料	指定來源格式	query	string

圖 5-18 公共運輸資料服務內容檢視

表 5.3 運輸分類資料彙整表

運具類型	資料內容
公車	公車線型資料
	公車路線資料
	公車站牌資料
	公車定時資料
	公車定點資料
	公車預估到站時間資料
高鐵	高鐵時刻表資料
	高鐵車次資料
	高鐵基本資料
臺鐵	臺鐵車次資料
	臺鐵時刻表資料
	臺鐵車站基本資料
航空	機場即時出入境航班資料
	航空公司資料
	機場資料

資料來源：<http://transport.tw/PTX/>與本研究整理

- B. 機關分類查詢：提供平臺使用者以不同資料來源機關進行屬於該機關提供之相關資料內容的查詢，如圖 5-19，包含各縣市政府交通局以及其他公共運輸管轄單位，彙整如表 5.4 所示。

表 5.4 公共運輸資料來源機關彙整表

項次	機關名稱	提供資料類型
1.	臺北市公共運輸處	公車資料
2.	新北市政府交通局	公車資料
3.	桃園市政府交通局	公車資料
4.	臺中市政府交通局	公車資料
5.	臺南市政府交通局	公車資料
6.	高雄市政府交通局	公車資料
7.	交通部臺灣鐵路管理局	臺鐵資料
8.	臺灣高速鐵路股份有限公司	高鐵資料
9.	中華民國交通部公路總局	公車資料



10.	交通部民用航空局	航空資料
-----	----------	------

The screenshot shows the 'Public Transport Data Exchange' (PTX) website. The search results are filtered by agency. The table below represents the data shown in the screenshot:

服務名稱	通具類型	資料類型中文	資料提供單位	引用次數	說明	網址
公路總局公車路線與站牌資料	公車	站點,路線	交通部公路總局	15,929		
公路總局公車預估到站時間資料	公車	到離站	交通部公路總局	1,632,843		
臺北市公車預估到站時間	公車	到離站	臺北市政府交通局	11,825,781		
臺北市公車站牌資料	公車	站點	臺北市政府交通局	33,680		
臺北市公車線型資料	公車	路線,路線GIS線型	臺北市政府交通局	28,466		
公路總局公車路線資料	公車	路線	交通部公路總局	26,126		
公路總局公車定點資料	公車	動態點位	交通部公路總局	104,311		
臺北市公車路線與站牌資料	公車	站點,路線	臺北市政府交通局	21,101		
臺北市公車定時資料	公車	動態點位	臺北市政府交通局	11,425,586		
公路總局公車定時資料	公車	動態點位	交通部公路總局	523,910		

圖 5-19 以機關分類查詢公共運輸資料

C. 名稱查詢：提供平臺使用者輸入欲查詢資料名稱之關鍵字進行相關資料內容的查詢，如圖 5-20。

The screenshot shows the PTX website with search results for '公車'. The table below represents the data shown in the screenshot:

服務名稱	通具類型	資料類型中文	資料提供單位	引用次數	說明	網址
公路總局公車路線與站牌資料	公車	站點,路線	交通部公路總局	15,929		
公路總局公車預估到站時間資料	公車	到離站	交通部公路總局	1,632,843		
臺北市公車預估到站時間	公車	到離站	臺北市政府交通局	11,825,781		
臺北市公車站牌資料	公車	站點	臺北市政府交通局	33,680		
臺北市公車線型資料	公車	路線,路線GIS線型	臺北市政府交通局	28,466		
公路總局公車路線資料	公車	路線	交通部公路總局	26,126		
公路總局公車定點資料	公車	動態點位	交通部公路總局	104,311		
臺北市公車路線與站牌資料	公車	站點,路線	臺北市政府交通局	21,101		
臺北市公車定時資料	公車	動態點位	臺北市政府交通局	11,425,586		
公路總局公車定時資料	公車	動態點位	交通部公路總局	523,910		

圖 5-20 以名稱查詢公共運輸資料

## (2) 加值應用服務

提供加值業者應用平臺所提供的資料，並於加值應用專區中提供相關資料服務，讓資料應用者可以透過統一的資料標準下載與介接平臺上的公共運輸資料，包含：資料標準文件、API 介接範例、線上 API 說明、加值作業說明、資料服務引用範例。

## 3. 案例三：Easy Traffic 簡單交通簡單生活

Easy Traffic 簡單交通簡單生活為一民間企業營運之大數據服務平臺，其網頁入口畫面如圖 5-21 所示，其開發動機係因目前各機關資料豐富，且部分資料已開放，但多著重於營運管理層面或是平臺查詢服務，缺乏規劃與評估功能，導致掌握大量資料卻缺乏完整分析資訊，也並未加以應用與延伸，故在現有各單位資源豐富之下，如何應用這樣的資料內容加以延伸應用，為當下最重要的事情，該平臺除介紹本計畫可應用的範圍內容外，也透過實作的方式告知相關應用者應用方向，以下針對此平臺提供之服務進行介紹：

- (1) 資料服務專區：提供現有平臺資料狀況說明供查詢，如圖 5-22 圖 5-22 Easy Traffic 服務平臺資料服務專區
- (2) 所示；目前因資料正規化定義未明，暫不提供使用者自行上傳資料，以免資料錯誤，但保留其擴充性。
- (3) 示範應用專區：揭露目前合作配合之示範資訊，呈現資料應用成果，並公開所應用之資料，如圖 5-23 及圖 5-24 所示。
- (4) 模組應用專區：包含三大應用模組，使用者分析模組、天災決策模組、活動決策模組，依據不同模組使用不同的表單組合，透過 GIS 地圖進行呈現，如圖 5-25 所示。



圖 5-21 Easy Traffic 服務平臺入口網頁

序號	資料名稱	資料提供者	更新頻率	檔案類型	更多資訊
1	公車業者	TMS	1季	JSON	查看
2	縣市代碼	TMS	有新增時才更新	JSON	查看
3	活動資訊	TMS	1天	JSON	查看
4	旅館資訊	TMS	1天	JSON	查看
5	餐廳資訊	TMS	1天	JSON	查看
6	景點資訊	TMS	1天	JSON	查看
7	政府代碼	TMS	1天	JSON	查看
8	航空站	TMS	有新增時才更新	JSON	查看
9	航空到站	TMS	1分鐘	JSON	查看
10	航空離站	TMS	1分鐘	JSON	查看
11	雨量	TMS	10分鐘	JSON	查看
12	空氣汙染	TMS	1小時	JSON	查看

圖 5-22 Easy Traffic 服務平臺資料服務專區



圖 5-23 Easy Traffic 服務平臺示範應用專區



圖 5-24 示範應用範例說明

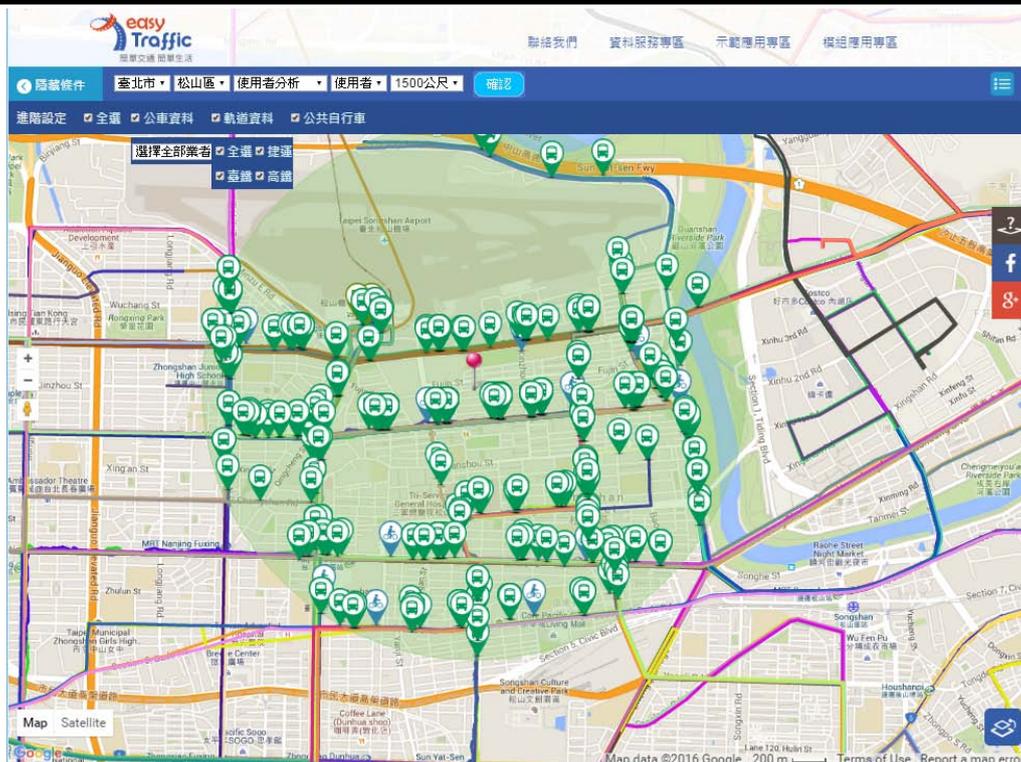


圖 5-25 Easy Traffic 服務平臺示範模組應用專區

#### 4. 案例四：臺中市易肇事路口整合平臺

臺中市政府為提升交通事故分析，於 2015 年提出「易肇事路口整合平臺建置計畫」，整合交通事故統計、事故熱點分析、交通警察勤務派遣、交通宣導區位，針對肇因來達到交通工程改善、交通宣導標語、交警各路口取締重點型態，以及 APP 推播高風險路口。平臺建置內容如圖 5-26 所示，以下針對內容說明：

- (1) 交通事故統計分析：每月匯入警政署的交通事故資料庫，依據機動車事故、高齡者事故、自行車事故、酒駕事故、大客車事故等加以分類，透過事件分類，統計各時段、車種的交通事故肇因，並製作肇事斑點圖。
- (2) 事故熱點分析：根據每個路口所發生事故之頻率，透過分析找出臺中市交通肇事熱點，並依肇事程度以不同顏色呈現。
- (3) 交通警察勤務派遣：藉由肇事斑點圖與事故熱點分析，來規劃交通勤務警察派遣路口及頻率。
- (4) 交通宣導區位：依肇事斑點圖與事故熱點分析，針對各區域道路設置，以達到宣教育的提醒效果。
- (5) 符合肇因之交通工程改善與交通宣導標語：特別針對臺中市年度前 30 大肇事路口及每月前 10 易肇事路口進行車輛碰撞構圖分析，

藉由碰撞型態、碰撞位置的分析，了解違規駕駛型態，對照型態研擬交通工程改善作為。再透過碰撞型態，了解主要肇事原因，來決定應宣導行車需注意內容及行車時需注意的駕駛行為。

- (6) 交通警察依路口特性區分取締重點型態：根據車輛碰撞型態或碰撞位置，提高重點路口取締，藉此降低違規的交通事故。
- (7) APP 推播高風險路口：根據熱點分析結果，透過 APP 推播告知駕駛者，前方路口為高風險路口，應注意常發生肇事主因。藉由 APP 提醒以達到降低交通事故發生。

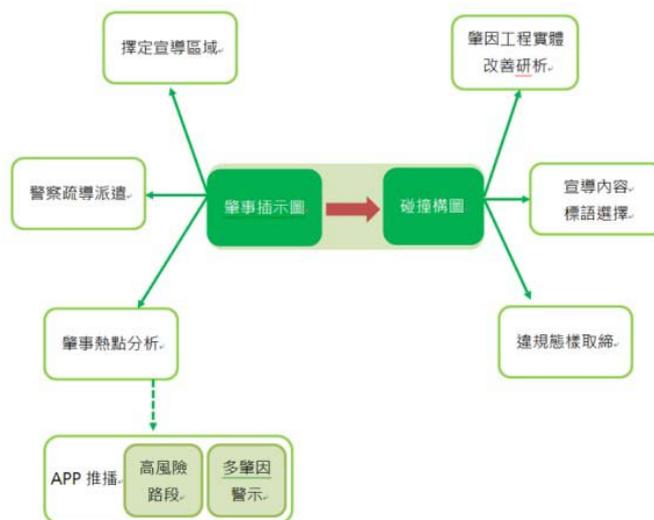


圖 5-26 臺中市易肇事路口整合平臺建置計畫內容示意圖

肇事斑點圖呈現方式主要為透過 Google GIS 系統與臺中市交通事故統計資料整合，利用不同標記符號代表單一交通事故，並定位於 GIS 系統；透過斑點圖可清楚了解該區域哪些路段、路口於何時發生何種交通事故，也能夠依據密集程度判斷交通事故發生頻率較高為哪些區域。

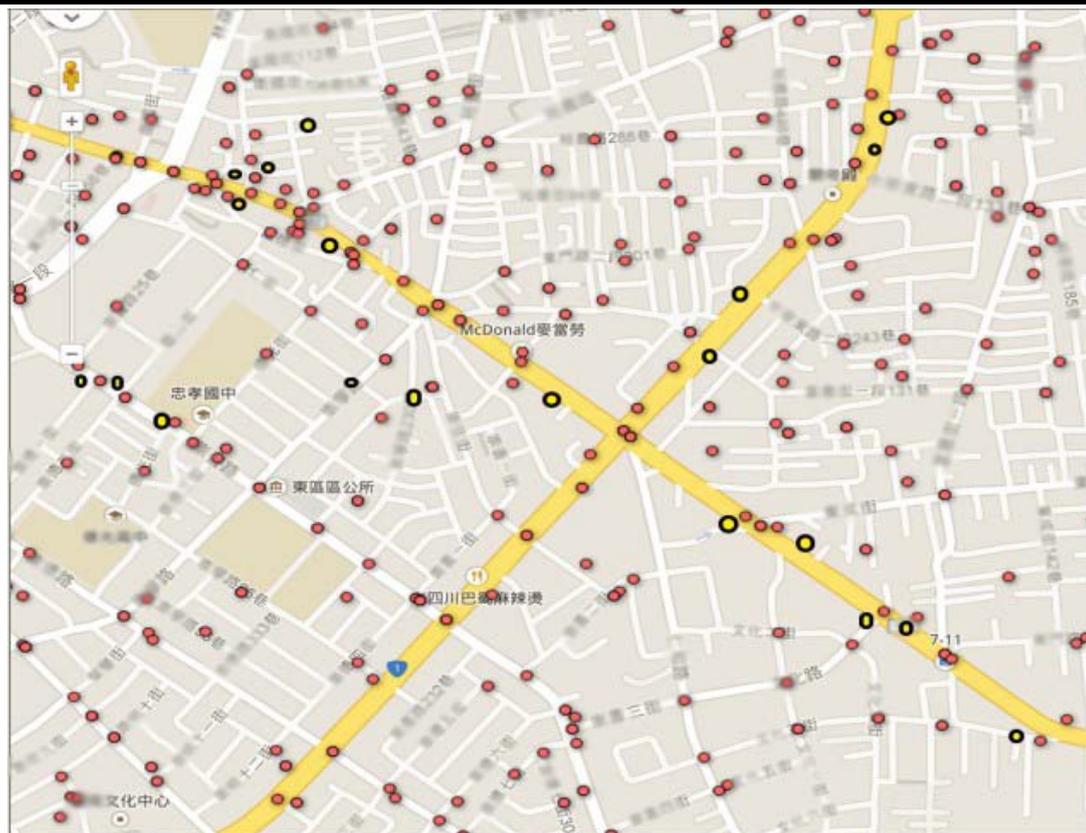


圖 5-27 肇事班點圖

肇事熱點分析是將斑點圖標記符號於 GIS 圖資上交通事故件數所呈現密集程度數值化的數據透過交通肇事當量值換算，換算路口肇事當量值，並依照高低程度以不同顏色質化方式呈現於圖資中，肇事頻率越高的區域則以較鮮豔顏色顯示；反之，肇事事件頻率較低的區域則以較淺色或是不著色方式顯示。

車輛碰撞構圖主要針對已發生交通事故路口，定義一定時間內，將發生於該路口肇事型態重新繪製於同一圖中，可直接於圖中呈現肇事車種、不同碰撞型態與發生件數、時段及地點。透過車輛碰撞構圖可針對肇事主因進行改善，分析路口發生哪幾類碰撞型態事故，肇事的主要方向性、時段與車種。

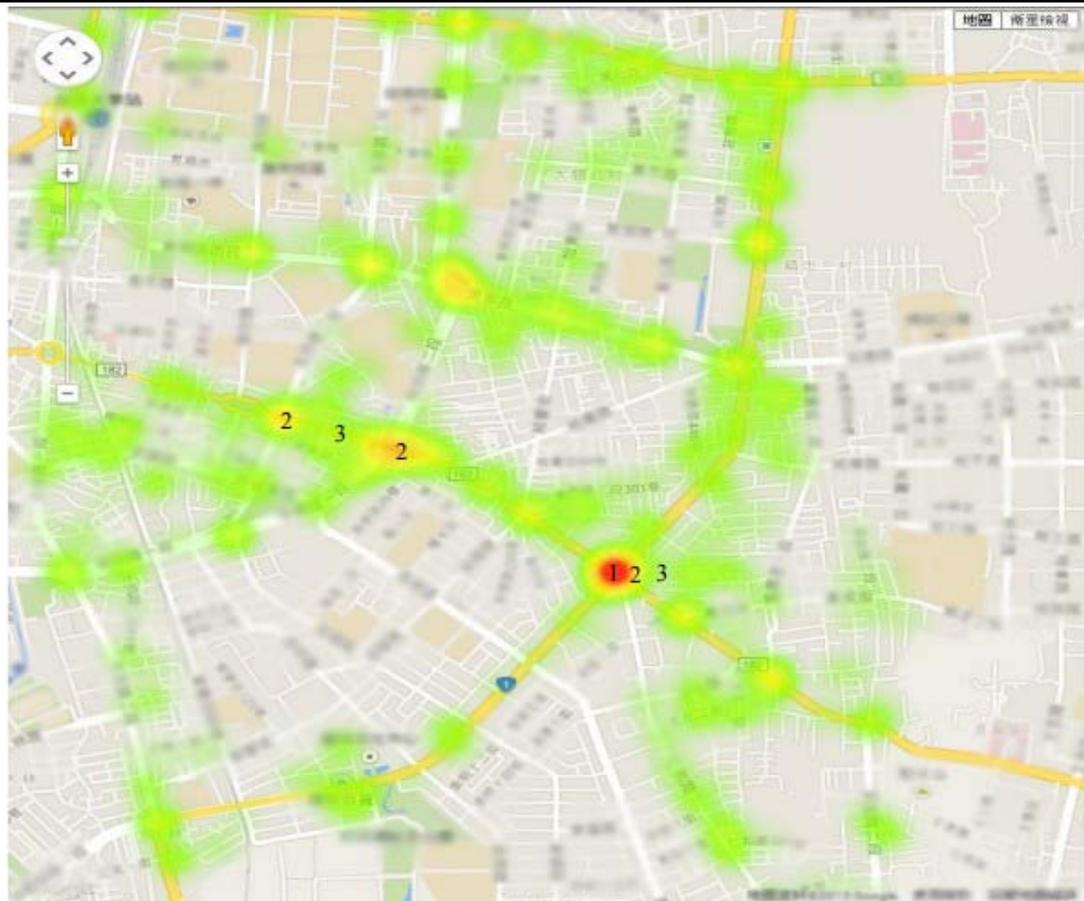


圖 5-28 肇事熱點圖

## 5. 案例五：交通部交通數據匯流平臺

為了因應運輸需求和供給的迅速變遷，產、官、學、研各方面均需交通相關數據資料作為分析與後續發展之基礎，因此交通部管理資訊中心於 105 年度辦理「交通數據匯流平臺建置案」，未來該平臺將可作為交通數據應用需求與資料來源之橋樑，希望達到產製有價分析資料、作為政策形成基礎、瞭解整體運輸績效，以及解決民眾關切議題等目標，該計劃的正式報告尚未出版，故在此僅做初步回顧。

### 5.2.2 綜合比較

目前國內大數據正在起步階段，有多個平臺均在進行蒐納資訊，對外提供開放式的 API 服務，以利有需要的加值業者取得與應用，對國內大數據發展極有幫助，然而就交通部角度而言，除持續提供資源並輔導這類 Open Data 平臺永續發展外，也可考慮建構機關內政策評估研究使用之資料匯流平臺，此資料平臺將蒐集保留更長期、更詳細之資料，同時以機關間資料共享運用為主，非公家機關欲使用時必須結合公部門計畫專案申請

之。此定位除可與目前已蓬勃發展的 Open Data 平臺有所區隔外，亦可提高各部屬機關提供業務細節資料之意願，進而讓政府部門之間各種政策評估、交通管理、公共運輸供給調整、壅塞預警等施政評估專案計畫有統一的資料匯流與成果回饋平臺。

## 5.3 基本需求規劃設計

### 5.3.1 匯流類別

考量平臺未來運用的情境，資料類別之設計應允使用者可透過多種不同分類方式均可快速找到資料集為佳，可能的資料集分類包括：運具別、行政區（縣市）、營運單位等，可考慮以巢狀標籤（Nested Label）來實作之。

除前述幾種顯而易見的分類方式外，亦可考慮從公共運輸服務水準的評估方向來分類如圖 5-29，包含快速、密集、舒適、信賴、安全與無縫等面向。



圖 5-29 公共運輸服務水準評估方向分類圖

至於匯流資料集內容，基於表 5.5 之考慮因素，較具大數據分析價值潛力的特性包含：動態資料、數位化資料、交易明細而非彙整報表、能由機器自動產生...等，在此條件下，建議資料匯流平臺第一階段應蒐納的資料如表 5.6 四大類。

表 5.5 匯流平臺資料納選評估考慮因素表

資料屬性	紀錄方式	資料時間	自動介接	資料格式	外部應用
動態資料	電子記錄	交易紀錄	有API	XML,JSON	有潛力
靜態資料	紙本記錄	日彙整值 月彙整值 年彙整值	無API	EXCEL CSV 其他(如二進位檔)	無潛力 (如MMIS)

表 5.6 匯流平臺初創期間擬蒐納資料集表

匯流類別	資料範例	服務水準評估面向
載具軌跡	公路運輸（公車動態定時/定點/預估） 鐵路運輸（行控中心 Event Log）	快速/密集/信賴
電子票證	票證公司(悠遊卡/一卡通等)票卡資料 營運單位(客運公司/臺鐵局等)票務資料	舒適/無縫
監理資料	公路運輸業車籍資料/客運駕籍資料 違規紀錄	安全
肇事資料	A1/A2/A3 各類事故紀錄	安全

資料來源：本研究整理

除原始資料集外，資料匯流平臺也將規劃增值回饋資料集，包括但不限以下兩種增值方式：

### 1. 資料集融合結果

所謂資料融合，係指將 A 資料集與 B 資料運算後得出新的資料集，例如第四章標竿案例第 4.3.2 節之分析，以演算法結合公車動態、電子票證得到各車輛於各時間各空間上之車上乘載總人數就屬於融合後的資料集，這些資料集回饋至匯流平臺後可讓其他共享者直接在這個基礎下繼續其他分析。

### 2. 資料插補結果

所謂資料插補，係指原始資料集有些欄位缺漏，而分析者裡用演算法搭配其他資料插補出新的資訊，例如第四章標竿案例第 4.7.2.2 節中電子票證公司提供的資料有若干闕漏，包括臺鐵局票證無進站時間、公車票證之車站代碼無確切位置資訊等，均可撰寫演算法插補資料後回饋至匯流平臺，以利後續分析者可直接存取較高品質的二級資料源。

匯流平臺應該鼓勵用戶回饋資料集至平臺，某機關單位若能累積越多回饋資料集紀錄，將有助於未來申請下載其他資料集的審核，形成良性循環。

### 5.3.2 資料倉儲規劃

在資料倉儲方面之規劃初步建議採用關聯式與非關聯式資料庫的混合方式進行，關聯式資料庫的使用則將再區分為讀取為主的資料庫（On-Line Analytical Processing, OLAP）與寫入為主（On-Line Transaction Processing, OLTP）兩大類，前者佔整各服務中的絕大多數，後者則僅有少數資料集回饋或系統維運相關資料使用之。

對於須大量蒐尋的資訊，交由非關聯式資料庫進行快取記憶，以提供快速查詢使用。檔案伺服器的部分，將固定備分存放來源資料的內容資訊，以及資料庫與系統相關備分的資料。

以下將依表 5.6 說明各類蒐納資料之主要欄位概觀，實際欄位也應與資料提供單位做最後確認。

首先在載具軌跡方面，國內公車動態相當成熟，但各縣市的公車動態資料格式看似相同類似表 5.7 表 5.8，但各系統間仍存在不少差異，為利於後續整合分析因考慮統一格式後再蒐納，建議可依交通部 104 年初頒訂之「公共運輸旅資料標準 v1.0」蒐納。

此外，表 5.7 表 5.8 僅是子資料表，實際蒐納時必須連同父資料表（如圖 5-30）一併蒐納，而且必須逐日以日期為 Key 封存成一分，會強調此原則的原因在於定時定點資料表本身內含日期時間欄位，但父資料表則是持續維護更新而無日期時間欄位，因此常有歷史定時定點資料無法找到匹配之站牌號碼（實務增刪或移設）之困境，故建議蒐納時必須以「日」為單位蒐納各歷史資料日所對應之父資料表資料內容。

資料行名稱	資料類型	允許 Null
A2	char(2)	<input checked="" type="checkbox"/>
Cmp	char(2)	<input checked="" type="checkbox"/>
BusID	char(10)	<input checked="" type="checkbox"/>
DutyStatus	tinyint	<input checked="" type="checkbox"/>
BussStatus	tinyint	<input checked="" type="checkbox"/>
Route	char(10)	<input checked="" type="checkbox"/>
Goback	tinyint	<input checked="" type="checkbox"/>
Stop	int	<input checked="" type="checkbox"/>
Leave	bit	<input checked="" type="checkbox"/>
GPSTime	datetime	<input checked="" type="checkbox"/>
Type	tinyint	<input checked="" type="checkbox"/>
TrasTime	datetime	<input checked="" type="checkbox"/>
SN	char(10)	<input checked="" type="checkbox"/>
RecTime	datetime	<input checked="" type="checkbox"/>

資料行名稱	資料類型	允許 Null
STOP	nchar(4)	<input checked="" type="checkbox"/>
ID	nchar(4)	<input checked="" type="checkbox"/>
ROUTEID	nvarchar(5)	<input checked="" type="checkbox"/>
NAMEZH	nvarchar(15)	<input checked="" type="checkbox"/>
SEQNO	int	<input checked="" type="checkbox"/>
PGP	tinyint	<input checked="" type="checkbox"/>
TERMINAL	tinyint	<input checked="" type="checkbox"/>
DISTRICTID	tinyint	<input checked="" type="checkbox"/>
GOBACK	tinyint	<input checked="" type="checkbox"/>
LATITUDE	decimal(10, 8)	<input checked="" type="checkbox"/>
LONGITUDE	decimal(10, 7)	<input checked="" type="checkbox"/>

資料行名稱	資料類型	允許 Null
PROVIDER	nvarchar(5)	<input checked="" type="checkbox"/>
ID	nvarchar(5)	<input checked="" type="checkbox"/>
NAMEZH	nvarchar(11)	<input checked="" type="checkbox"/>
TYPE	tinyint	<input checked="" type="checkbox"/>

圖 5-30 公車動態定點資料相關資料表及欄位圖

表 5.7 公車動態定時資料欄位清單

車牌號碼	PlateNumb	M	xs:string
營運業者資料	Operator	M	OperatorType
路線資料	Route	M	BusRouteType
班次資料	Trip	O	BusTripType
車輛位置	BusPosition	M	PointType
行駛速度	Speed	M	xs:double
方位角	Azimuth	M	xs:double
車輛狀態	BusStatus	M	CodeList
去返程	Direction	M	CodeList
勤務狀態	DutyStatus	M	CodeList
車輛種類	CarType	O	xs:string
資料送出日期時間	TrasTime	O	xs:dateTime
GPS 記錄時間	GPSTime	O	xs:dateTime
資料收到日期時間	RecTime	M	xs:dateTime

資料來源：公共運輸旅資料標準 v1.0

表 5.8 公車動態定點資料欄位清單

站牌/位資料	Stop	M	BusStopType
車牌號碼	PlateNumb	M	xs:string
營運業者資料	Operator	M	OperatorType
路線資料	Route	M	BusRouteType
班次資料	Trip	O	BusTripType
去返程	Direction	M	CodeList
資料型態種類	MessageType	M	CodeList
車輛種類	CarType	O	xs:string
資料送出日期時間	TrasTime	O	xs:dateTime
GPS 記錄時間	GPSTime	O	xs:dateTime
資料收到日期時間	RecTime	M	xs:dateTime

資料來源：公共運輸旅資料標準 v1.0

至於鐵路之載具軌跡則須蒐納綜合調度所（或稱行控中心）的延誤資料，主因在於鐵路系統為公告時刻表之公共運輸系統，故蒐集載具軌跡時將連同計畫性（時刻表）與實際軌跡（對比時刻表之延滯）一併蒐集方可得到確切軌跡，相關資料表可整理如圖 5-31，其中 CTC\_TRN\_ANALYZE 表即含延誤資訊，比對 TRA\_PTRN\_TT\_EXPAND 之表定時間即達到蒐納載具軌跡之目的。

TRA_PTRN_TT_EXPAND		TIME_EXPAND		CTC_TRN_ANALYZE	
資料行名稱	資料類型摘要	資料行名稱	資料類型摘要	資料行名稱	資料類型摘要
識別碼	int	識別碼	int	識別碼	int
DEP_DATE	date	DEP_DATE	date	ID	nvarchar(255)
TRN_CODE	nvarchar(255)	TRN_CODE	nvarchar(255)	DEP_DATE	date
STN_ID	int	TRN_CLASS_ID	nvarchar(255)	TRN_CODE	nvarchar(255)
SCHE_DEP_TIME	time(0)	MAIN_ROUTE_ID	int	STN	smallint
SCHE_ARR_TIME	time(0)	VIA_ROUTE_ID	int	DELAY_TIME	smallint
ARR_PURPOSE	int	SEA_LINE_FLAG	int	REASON	nvarchar(255)
TRN_STN_SN	int	TRN_DIR_FLAG	int	NOTE	nvarchar(255)
OUT_ROUTE_ID	int	OVER_NIGHT_STN_ID	int	CREATED	nvarchar(255)
REASON	nvarchar(255)	S_STN_ID	int	ARR_DELAY_TIME	smallint
NOTE	nvarchar(255)	E_STN_ID	int	TRN_STN_SN	smallint
MARK	nvarchar(255)	SCHE_S_DEP_TIME	time(0)		
		SCHE_E_ARR_TIME	time(0)		
		RUN_DATE	nvarchar(255)		
		CRIPPLE_FLAG	nvarchar(255)		
		PACKAGE_FLAG	nvarchar(255)		
		DINNING_FLAG	nvarchar(255)		
		OPERATION_REMARK	nvarchar(255)		

TR_CLASS	
資料行名稱	資料類型摘要
Class_Code	varchar(50)
Class_Name	varchar(50)
Class_EName	varchar(50)

資料來源：臺鐵局提供與本研究整理

圖 5-31 臺鐵載具軌跡相關資料欄位圖

接下來說明電子票證相關欄位，公共汽車電子票證將依蒐集對象（營運單位或票證公司）有約略差異如圖 5-32，未來實際蒐納時，各營運業者

(例如不同客運公司) 或各票證公司 (例如悠遊卡與一卡通) 也都會存在些許差異。

營運單位		票證公司	
資料行名稱	資料類型摘要	資料行名稱	資料類型摘要
路線	int	票證公司	char(8)
駕駛	nvarchar(10)	客運公車代碼	varchar(6)
車輛	nvarchar(8)	卡號	char(12)
卡號	nvarchar(255)	票種編號	tinyint
發卡公司	nvarchar(10)	票種代碼	tinyint
卡種	nvarchar(10)	路線公路編號	int
上車日期	date	司機編號	int
上車時間	time(7)	車號	varchar(8)
上車站	nvarchar(50)	上車時間	smalldatetime
下車日期	date	上車站序	int
下車時間	time(7)	上車計費站序	int
下車站	nvarchar(50)	下車時間	smalldatetime
		下車站序	int
		下車計費站序	int

資料來源：業者提供與本研究整理

圖 5-32 公車電子票證資料欄位圖

在鐵路的電子票證方面，臺灣鐵路管理局目前正在建置第四代票務系統，在新系統尚未上線前，建議以第三代之格式欄位蒐納資料，包括對號車票、自動售票機、定期票與多卡通等分別如圖 5-33 圖 5-34 圖 5-35 圖 5-36。

欄位ID	欄位名稱	型態	欄位ID	欄位名稱	型態	欄位ID	欄位名稱	型態
STOP_CODE	站號	char	RIDING_TYPE	行程別	char	TICKET_KIND	卡別	int
TERM_NO	終端機號	char	SURCHARGE	加價別	char	TIMES	回數	int
FBEG_DATE	當班年月日	int	RCV_OLD_TKT	收回原票符號	char	RES_NO	個人預約號碼	int
FBEG_TIME	當班時間	int	CERTIFICATE_NO	憑證號碼	char	SEAT_KND	座位別	char
SELLER_NO	售票員代號	char	FARE	總價	int	ACC_SEQ	科目序號	char
TRNS_DATE	交易年月日	int	PRICE	總價	int	PROM_NO	多元票價案號	char
TRNS_TIME	交易時間	int	RFND_PRICE	手續費	int	LIST_PRICE	原價金額	int
TKT_NO	票號	char	RFND_FREETAX	免稅手續費	int	DPT_DATE	發車日期	int
TRNS_TYPE	交易序號	int	CHARGE_PAYMENT	手續費付款方式	char	SEQNO	流水號	char
BOARD_DATE	乘車年月日	int	CHARGE_MARK	退費別	char	ADD_DATE	ADD日期	date
BOARD_TIME	乘車時刻	int	KILO	公里數	real			
OP_TYPE	離連線別	char	DIRECT	方向	char			
TRNS_CODE	交易別	char	ROUTE_CODE	路徑	char			
PAYMENT	付款方式	char	PTY_NO	團體預約號	char			
TRAIN_TYPE	車種	char	STEPS	找座位STEP	char			
TRAIN_NO	車次	char	FARE_ADJUCT	退票調整金額	int			
TKT_BEG_STOP	票面起站	char	NO_LINES	車箱別	int			
TKT_END_STOP	票面到站	char	SECTION	前/後段	char			
SEAT_BEG_STOP	座位起站	char	SYSTEM_TYPE	系統類別	char			
SEAT_END_STOP	座位到站	char	TRNS_STATUS	訂票取票 或購票	char			
NO_TKTS	票數	int	TKT_TYPE	票種別	int			
PASN_TYPE	身份別	char	CAR_NO	車號	char			
AGE_TYPE	年齡別	char	SEAT_NO	座號	int			

資料來源：臺鐵局提供與本研究整理

圖 5-33 臺鐵第三代票務系統對號車票資料欄位圖

欄位ID	欄位名稱	型態	欄位ID	欄位名稱	型態
ADATE	記帳日期	Int	T_COUNT	10元找錢數	Int
TRANS_DATE	交易日期	Int	INDATE	轉入日期	Int
TRANS_TIME	交易時間	Int	INTIME	轉入時間	Int
BSTOP	起站號	char	KIND	原因	char
ATVM_ID	機號	char			
TRAIN_TYPE	車種	char			
PASN_TYPE	身份別	char			
AGE_TYPE	年齡別	char			
RIDING_TYPE	行程別	char			
SURCHARGE	加價別	char			
NO_TKTS	票數	int			
SEQ		int			
PAYMENT	付款方式	int			
ESTOP	到達站	char			
CASH_TOT_AMOUNT	現金交易金額	Int			
CARD_TOT_AMOUNT	磁卡交易金額	Int			
CASH_AMOUNT	現金單向金額	Int			
CARD_AMOUNT	磁卡單向金額	Int			
DISTANCE	距離	Int			
TF_COUNT	15元找錢數	Int			
O_COUNT	1元找錢數	Int			
F_COUNT	5元找錢數	Int			

資料來源：臺鐵局提供與本研究整理

圖 5-34 臺鐵第三代票務系統自動售票機資料欄位圖

欄位ID	欄位名稱	型態
STOP_CODE	發售站代碼	char
TKM_NO	發售機編號	tint
SALE_DATE	發售年月日	int
TKS_NO	票流水號	sint
OST_CD	進站碼	char
OD_CNT	進出累計數	int
DST_CD	出站碼	char
OD_DATE	進站年月日	int
TRAIN_TYPE	車種代號	char
PASN_TYPE	身份別	char
AGE_TYPE	年齡別	char
TK_TYPE	票別	char
EBABLE	可使用次數	sint
CAR_NO	車次	char
PASN	人數	int
MAN_KILO	延人公里	real
REVENUE	收入	int
DIRECT	方向	int
INTIME	進入時間	int
OUTTIME	輸出時間	int
ADD_DATE	轉檔日期	date

資料來源：臺鐵局提供與本研究整理

圖 5-35 臺鐵第三代票務系統之定期票主要資料欄位圖

欄位ID	欄位名稱	型態
CARD_TXN_TYPE_ID	交易類別	number(3)
CARD_TXN_SUBTYPE_ID	交易次類別	number(3)
DEV_ID	設備編號	number(9)
SP_ID	業者別	number(3)
TXN_DATE	交易日	number(8)
TXN_TIME	交易時間/出站時間	number(6)
CARD_PHYSICAL_ID	晶片卡號(內碼)	number(10)
ISSUER_ID	卡片發行者	number(3)
CARD_TXN_SEQ_NO	交易序號	number(3)
TXN_AMT	交易金額	number(5)
ELECTRONIC_VALUE	卡片餘額	number(7,2)
SVCE_LOC_ID	場站代碼	number(3)
PROCESSING_DATE	處理日	date
SETT_DATE	清分日	date
BUSINESS_DATE	營運日	date
DEV_TYPE_ID	設備種類	number(3)
ENTRY_LOC_ID	進站場站代碼	number(3)
XFER_CODE	轉乘代碼	varchar2(4)
XFER_DISC	轉乘優惠	number(7,2)
PERSONAL_DISC	優惠補助款	number(7,2)
PENALTY	罰款金額	number(5)
LOYALTY_COUNTER	忠誠點數	number(5)
AGENT_NO	站務員代碼	number(5)

資料來源：臺鐵局提供與本研究整理

圖 5-36 臺鐵第三代票務系統之多卡通主要資料欄位圖

至於監理資料部分，公路總局 M3 系統具有相當豐碩之資料即可供分析之用，初步僅建議先由蒐集運輸業車輛與駕駛之相關資料如圖 5-37。



圖 5-37 監理系統運輸業車籍駕籍與違規資料欄位圖

### 5.3.3 服務網站規劃

匯流平臺網站之功能架構可概分為前臺服務提供與後臺系統管理兩大類如圖 5-38，以下將分節介紹之。

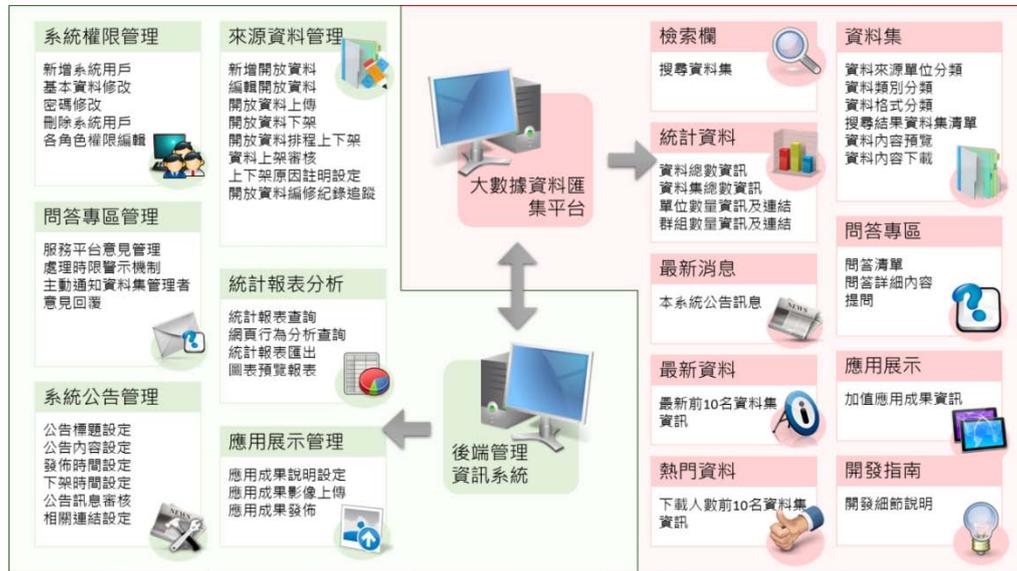


圖 5-38 匯流平臺功能架構示意圖

#### 5.3.3.1 前臺服務規劃

前臺入口網站示意如圖 5-39，以檢索介面為入口網站主軸，並搭配主題式查詢內容如所示，除以「資料集檢索模組」提供最直覺性的操作流程，透過對資料內容的分析整理、發展關鍵字建議、後分類、標籤雲等機制外，可提供使用者以最關心、最在乎的議題下去搜尋相關資訊，提供懶人包資訊下載等，故建議思考之方向列點如下：

1. 分眾：採用現有資料有查詢或下載數量統計，分析使用族群，列出前 5~7 大類，再歸納資料懶人包，其餘對象則以熱門查詢或因應時事提供資訊。
2. 分地：針對地方性的資訊，分析使用者所在位置，依縣市提供相關數據或資訊。
3. 建立資料相關連結，分析使用者行為，首頁即時回饋「可能有興趣」的資訊。
4. 其他功能部分則規劃有檢索欄、最新消息、最新資料、熱門資料、統計資料、資料集、問答專區、應用展示、開發指南、線上申請等功能。



圖 5-39 大數據資料彙集平臺入口網頁示意圖

### (1) 主題專區

利用蒐集而來且可開放的資源資料進行分析與規劃，設計與勞工、民眾切身相關主題，以輪播方式呈現於本系統平臺頁面上，當民眾進入該頁面時則可以索取相關的資料集內容，打包下載加以應用。

### (2) 檢索欄

提供使用者縮小資料集範圍，加速查詢並取得所需資料，因此功能為使用者最直覺性的操作模式，除入口網站以此為主軸外，更於所有功能介面下皆保留檢索欄，供使用者隨時能快速查詢，其查詢結果則以資料集介面顯示，資料集介面說明細述如(5)。

### (3) 最新消息

發佈由管理者欲告知使用者之公告訊息，其公告訊息內容係由作業管理系統中撰入並同步於服務平臺。

### (4) 最新資料及熱門資料

為利用更新時間、瀏覽下載人數統計資料提供最新資料及熱門資料的快速檢索方式，依據瀏覽統計數前三名搭配圖示的方式吸引使用者注意力。

### (5) 統計資料

提供本系統最新的資料總量、資料集數量、分類數量、提供資料的單位數量。

#### (6) 資料集

本功能介面分為左側類別分類區塊及右側資料集列表區塊，資料集示意畫面如圖 5-40 所示，兩區塊內容說明如下：

- A. 左側類別分類區塊：依據資料處理後的各資料類別分類，例如：來源資料單位類別、來源資料群組類別、開放資料格式類別等，點選後則將該類別下所有資料集列表於右側區塊。
- B. 右側資料集列表區塊：提供所有資料集、檢索後資料集或類別下資料集，供使用者查詢內容，並於右上方顯示目前本區塊的所有資料集筆數。



圖 5-40 資料集功能示意圖

#### (7) 問答專區

提供使用者與系統管理者的溝通管道，專區介面提供每筆問答資料，依發表時間排序，顯示欄位有標題、提問者、瀏覽人數及管理者回覆情形等，專區示意畫面如圖 5-41 所示，點選資料則提供該筆問答的詳細內容及管理者的回覆內容，除可解決提問者問題外，亦可提供給其他使用者參考；若使用者尚有其他問題則可新增提問，填入提問者、標題、意見內容等必填欄位，或更詳細指出反應問題之相關資料

集名稱，提問示意畫面如圖 5-42 所示，以利管理者更快速的修正問題。

發表日期	標題	提問者	瀏覽人數	回覆情形
2016-03-15	連結下載空白	*m*	2	未回覆
2016-03-11	無法下載	*宗*	1	未回覆
2016-02-27	請增加中文說明，以利使用	*jo*	8	處理中
2016-02-17	資料亂碼	*mech*	5	處理中
2016-02-02	針對OpenTime的改進	*世*	12	已回覆

圖 5-41 問答專區示意圖

問答專區 > 提問

\*提問人

\*標題

資料集名稱

\*意見內容

圖 5-42 新增提問示意圖

#### (8) 應用展示

為鼓勵各界創新加值應用本系統開放資料，特建置此平臺空間發佈其應用成果，以引領更多企業加值應用。

#### (9) 開發指南

提供使用、下載及介接本系統開放資料說明，以利指引使用者充分使用。

### (10) 線上申請

依據開發指南中所述提供申請使用、下載及介接本系統開放資料作法，提供相關表單下載申請用。

#### 5.3.3.2 後端管理規劃

後端作業管理系統規劃以資料內容管理介面為主軸，示意畫面如圖 5-43 所示，其核心為「系統權限管理模組」及「來源資料管理模組」，並包含統計報表分析、問答專區管理、系統公告管理及應用展示管理功能，而各功能說明如下：



圖 5-43 管理系統入口網頁示意圖

#### 1. 系統權限管理

依據本案需求系統用戶權限至少須分為資料提供者、資料審查者、系統管理者、系統稽核者等 4 種，此功能則是針對系統用戶及其權限做管理，子功能細分有系統用戶清單及角色權限管理，分別說明如下：

- (1) 系統用戶清單：列出目前系統所有用戶清單，並提供新增帳戶、刪除帳戶、基本資料修改及密碼修改等功能。
- (2) 角色權限管理：進行資料提供者、資料審查者、系統管理者、系統稽核者等 4 種角色對本系統各功能的權限管理，示意畫面如圖 5-44 所示，而左側功能選單會動態依角色變動，僅提供該角色權限下的功能列表。



圖 5-44 角色權限管理示意圖

## 2. 來源資料管理

本功能為此系統核心，為進行開放資料編輯、新增、刪除的專區，因系統使用者須細分為資料提供者與審核者，因此功能分為資料上架登錄與資料上架審核兩功能，此管理模組中的三功能說明如下：

- (1) 資料上架登錄：本功能為欲提供資料者使用，未來將依據資料盤點處理後，規劃必填或選填欄位、資料上傳等於此功能，而資料將暫存於資料庫，須待審核通過後方可上架。
- (2) 資料上架審核：本功能則為資料審核者使用，提供開放資料建檔及編修的紀錄追蹤等功能，並可對不同類別、項目的資料進行立即或排程上下架、審核作業及註記上下架原因。

## 3. 統計報表分析

提供資料統計報表查詢及網頁行為分析數據，並可自行選擇統計年、季、月或區間之數據，統計結果以圖表方式具體呈現，另亦可提供統計資料匯出功能，其示意畫面如圖 5-45 所示，子功能說明如下：

- (1) 統計報表查詢：此為提供相關管理報表，如資料集總量及異動數量統計、各資料集中資料的總量及異動數量統計、各來源單位資料總量及異動數量統計。

- (2) 網路行為分析：提供開放資料服務平臺的各功能使用分析，如各功能的使用統計、資料下載及瀏覽人數統計、熱門瀏覽或下載資料排名、檢索查詢關鍵字統計等。

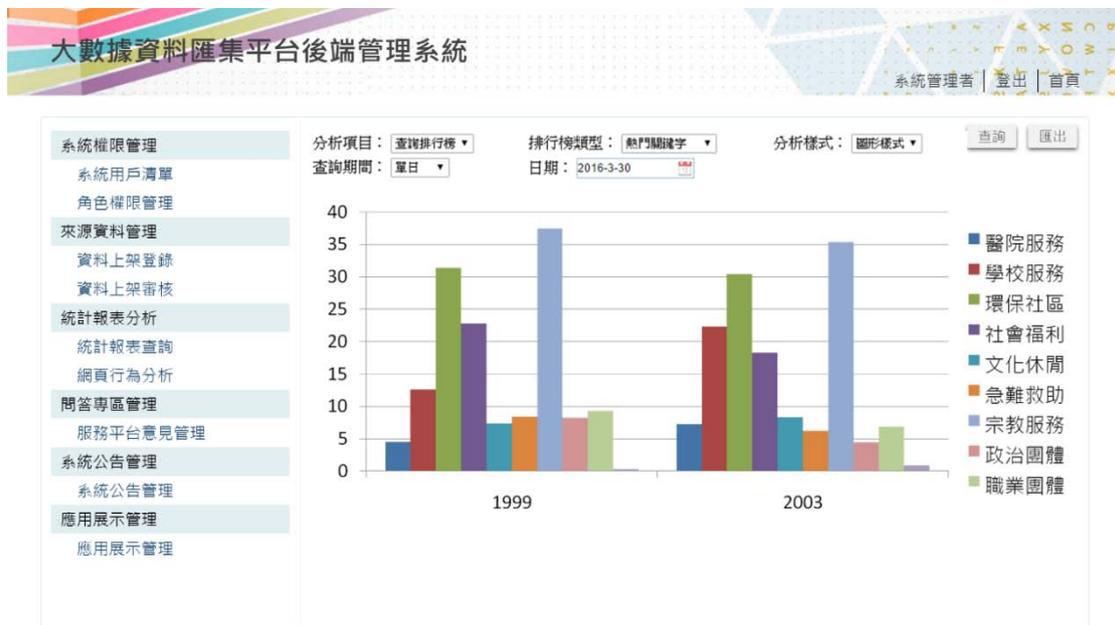


圖 5-45 統計報表分析示意圖

#### 4. 問答專區管理

本功能係針對使用開放資料者所提出的問題進行管理，規劃將意見分為未回覆清單與已回覆清單，並含主動通知資料集管理者、處理時限警示及意見回覆等功能。

#### 5. 系統公告管理

本功能提供系統管理者發佈最新消息於開放資料服務平臺，其細部含訊息發佈時間與下架時間、訊息主旨與內容、相關連結等功能。

#### 6. 應用展示管理功能

本功能為開放資料服務平臺的應用展示區內容管理，提供系統管理者可自行發佈增值應用成果，其細部含應用成果說明、成果影像上傳、相關連結等功能。

### 5.4 軟硬體架構規劃

一個完整性的大數據彙整平臺，應規劃採用雲端服務與實體存放為原則，其中包括資料處理模組、資料儲存模組、資料來源模組、叢集式管理模組、管理員介面等，其軟體系統架構圖，如圖 5-46；而在叢集式管理方

面，需考量後續應適應使用單位著重效益評估，由小漸大的投資擴充模式，最小的叢集部部屬所需的節點資訊，用以應付 Big Data 處理 Workload 的成長，以增加 Worker Node 的方式做橫向擴充（Scale-Out），擴充過程不必中斷執行中的服務，如圖 5-47，用以維持穩定性成長；而整體運作架構則如圖 5-48 所示。



圖 5-46 軟體系統架構圖(基礎版)

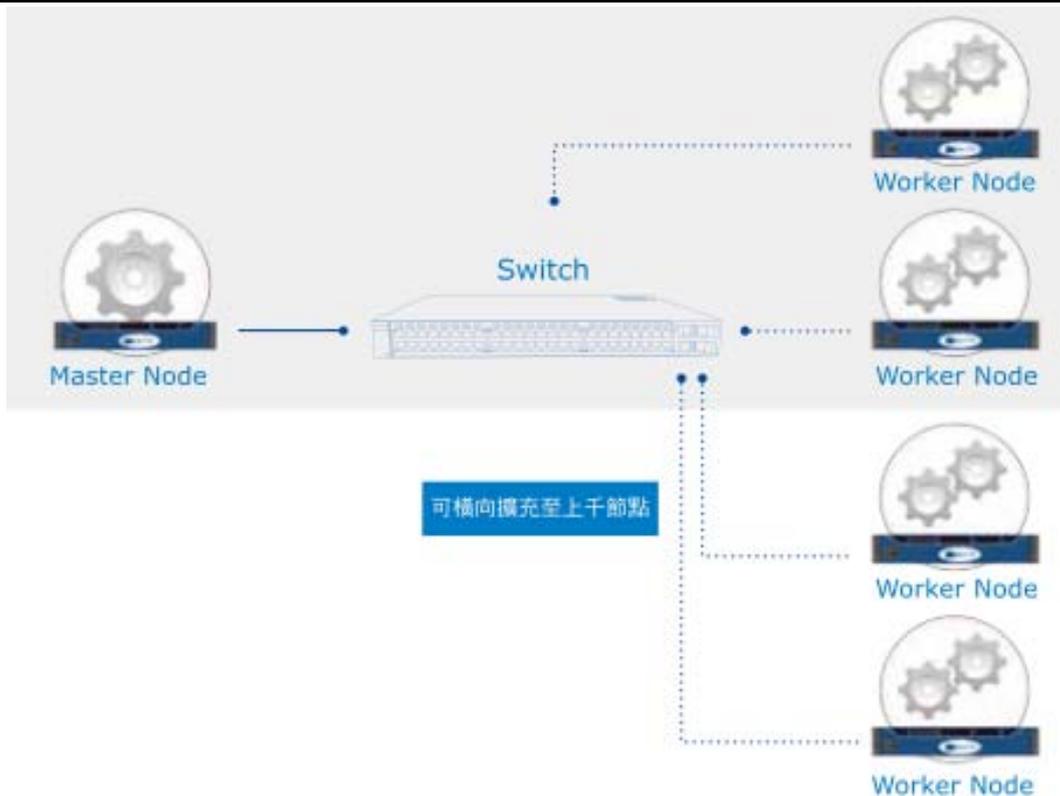


圖 5-47 最小叢集部署與橫向擴充架構圖

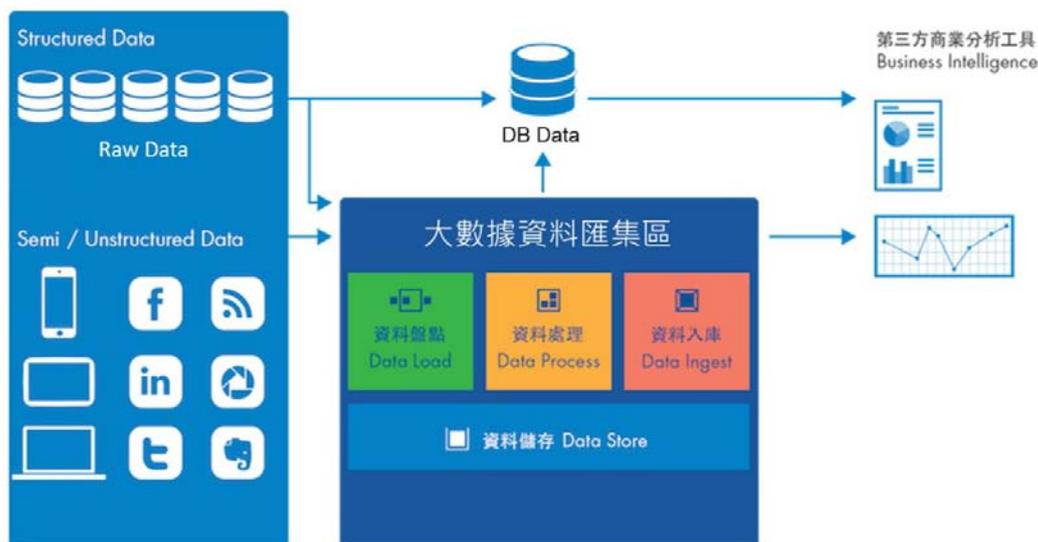


圖 5-48 大數據資料匯集運作架構

除上述所提及之系統架構外，進行系統建置外，仍應考量未來主要服務對象、實體與雲端混合應用上，以現行公共運輸資訊內容欲建立一大數據彙整平臺來說，若將其進行調整與修正，應考量以下內容：

1. 平臺安全性：資訊的取用除網路設備的防火牆管理，亦考量使用者的驗證與加密的處理機制。



2. 系統穩定性：由於各式運輸資料的大量存放與應用分析，對於系統的監控上則進行多方面的監控機制，並設計資訊快取及分流機制，以因應大量 Query 的處理及回應。
3. 來源資料蒐集：為提昇服務品質，除交通部內機關資料外，亦保留納入來自於交通部屬機關、縣市政府、公共運輸業者、民間票證公司等各單位資料蒐納之彈性。
4. 分析模組規劃：透過資料的蒐集，將資料交叉混合分析，產生各式效益的資訊，以提供不同的成果。
5. 資料蒐納：雖雲端技術發展下，系統可以有彈性的應用，但實體設備原有的穩定性及效率仍不可忽視，因此在本系統上，資料的蒐納存放，則考量實體與雲端的混合搭配。
6. 效能資安：含伺服反向代理、負載平衡服務、資料快取機制、資訊安全防護機制。

#### 7. 反向代理伺服器 (Reverse Proxy Server)

該平臺應採反向代理伺服器使其具備攻擊預防功能、反向代理緩存靜態內容、資訊加密、非常態存取防止功能、負載平衡功能、網址重定向功能、高可用性等特性。

- (1) 攻擊預防：反向代理伺服器主要是代為存取內部受到保護的服務，使其外部使用者無法直接接觸內部的服務網站，進而得以保護存心不良的駭客攻擊或入侵。
- (2) 緩存服務：透過代理伺服器的緩存處理，預先 Cache 相關靜態資料，減少重複性的需求，及系統的負擔。
- (3) 負載平衡：為降低網頁伺服器的負擔，透過代理伺服器的處理，平均分擔網頁伺服器的負擔，使其用戶進行網路存取時的品質與效率提升。

#### 8. 監控服務機制說明

平臺中應針對資訊服務、系統設備、後端程式與儲存進行相關監控，採用定時輪巡的機制，監控所屬的服務運作情形，並紀錄相關的 Log，以便於異常產生時，進行相對應的排解處理。

- (1) 服務監控：判斷各項服務的運作情形，包含連線數、服務情形等狀態。
- (2) 設備監控：判斷各設備的資源使用情形，如：CPU 使用率、記憶體使用率、硬碟使用率、網路使用率等。
- (3) 程式監控：監控後端資料蒐集、分析模組等相關程式的運作情形。
- (4) 儲存監控：監控資料庫使用率、存放空間的狀態。

## 9. 權限驗證服務說明

本平臺應朝向後續進階發展為主，故蒐納大量的資料進行彙整、分析外，仍應有民眾應用查詢服務、政府及加值業者進行相關資訊分析結果介接。因此在權限驗證方面，將透過特定申請機制，檢核使用者身分及可用權限，提供資訊介接或推送資訊的服務，藉以確保服務品質與系統資源利用的有效管理與查核。

## 10. 模組分析說明

依據深度訪談結果後進行模組內容規劃，該模組內容應包含加值應用之實用性、資料分析結果正確性方向為出發。

## 11. 紀錄服務說明

紀錄服務將紀錄各項服務項目、使用者存取資源、事件紀錄、監控紀錄，並分類存放至特定之紀錄服務設備之中。

依據上述所提，擬建議系統多層級架構圖、環境建置、軟體架構及資源運用架構圖繪製如後：

### 1. 多層級架構設計：

多層級架構如圖 5-49，簡介如下：

- (1) 服務對象：以政府（包含專案合作夥伴）為主。
- (2) Presentation：資訊開放平臺、後端管理平臺、資訊發佈服務 API。
- (3) Business：處理各項資料之蒐集、清理、分析、驗證、轉換、及發佈。
- (4) Data Access：各項屬性資料(Metadata)存放、紀錄存放等。
- (5) 效能資安：含資料快取機制、資訊安全防護機制。
- (6) 系統監控：服務監控、設備監控、程式監控、儲存監控。

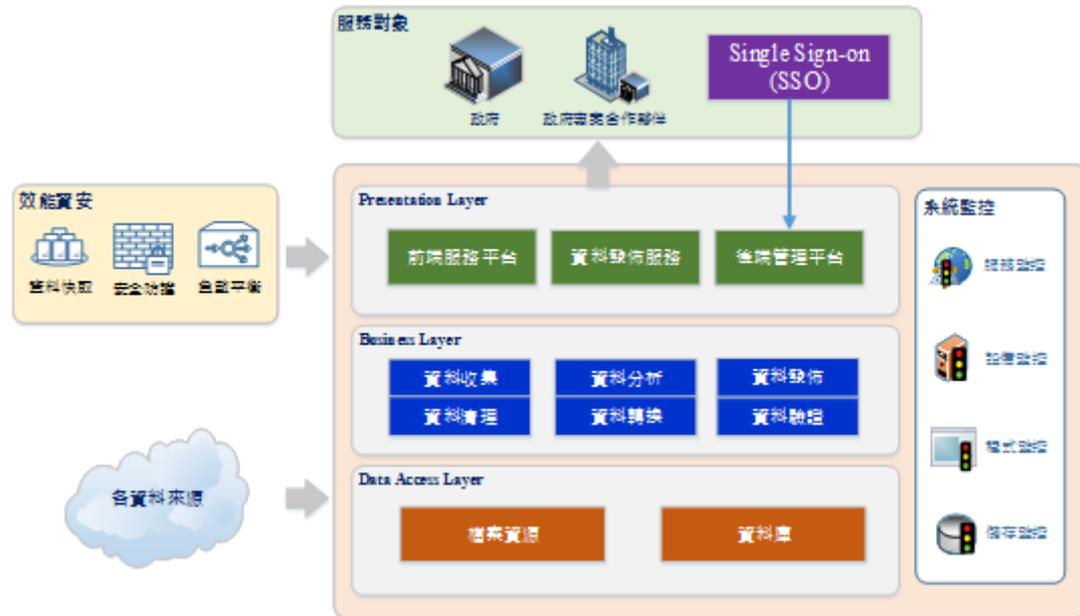


圖 5-49 系統多層級架構圖

## 2. 系統環境建置：

因本研究需考量以開源軟體進行建置，但開源軟體在運作上有其效能限制，故提出開源軟體與商業軟體兩種建議，其實際開發採用之類別，則依據實際資料量、運作內容、服務項目而進行變更，如圖 5-50。

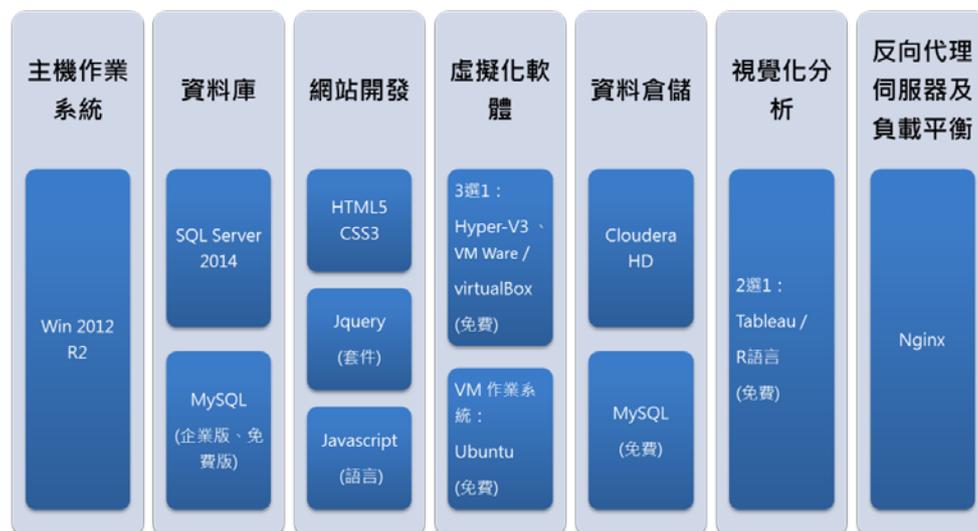


圖 5-50 系統環境建置

## 3. 整體軟體架構與資源應用：

每一 server 中均須包含資料蒐集平臺、資料發布平臺以及管理後臺。透過 load balancer 來平衡使用者的流量與備援機制。資料倉儲則使用 iSCSI

做存取以及採用 RAID10 來做到資料的保護。異地備援部分屬於可選部分，以避免發生機房受損的天災意外，如圖 5-51 中所述之項目。

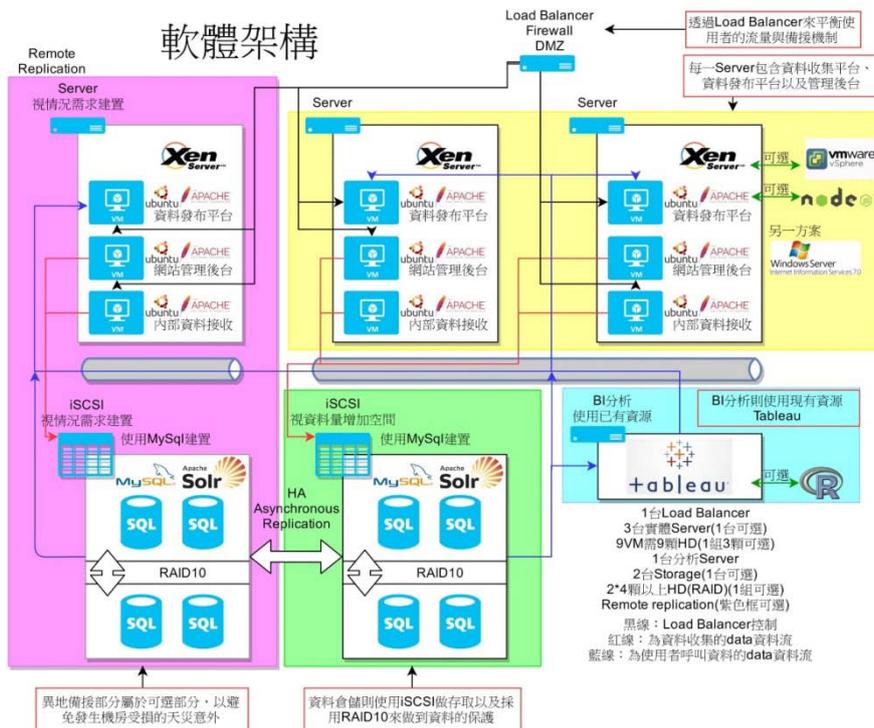


圖 5-51 整體軟體與資源應用規劃建議圖

#### 4. 應用範圍限制及可行性評估

現行流通資料類型可大致歸納為結構性資料、半結構性資料、以及非結構性資料三種，分述如下：

- (1) 結構性資料：指在來源網站/資料提供單位可提供具有部分或完整的詮釋資料欄位，且該詮釋資料可做為檢索點使用。
- (2) 半結構性資料：主要由文字組成，但較無規則可循或格式不固定，例如部落格資源、HTML 網頁內容等，檔案呈現型態為文字檔。
- (3) 非結構性資料：例如照片、影像、掃瞄文件等，檔案呈現型態有圖片檔及影音檔。

在本階段暫不規劃半結構性及非結構性資料內容匯入，待未來圖像、影音資料集之蒐納需求較為明確後再擴充。

此外，以目前大數據之發展趨勢而言，在不久的將來在蒐納資料量的成長上，可能必須面臨單日資料量以數十、數百 TB 的速度增加，總資料量也達到了 PB(Petabyte)等級，如此規模的資料量將使傳統的資料庫難以處理，屆時可能需要導入分散式的儲存架構，例如開源軟體 Hadoop，惟此軟體 Map-Reduce 架構雖可大量運算，但此架構較無法因應即時反應運算



結果之需求，故若對於反應時間有較高的需求，則可考慮如 Spark 等開源軟體，這類軟體將使用記憶體內運算的技術來提昇反應效率。上述這些分散式運算架構亦都有商業版本，例如 IBM InfoSphere Streams、HPE Vertica 等均是未來資料量達到 PB(Petabyte)等級時的可行方案。

## 5.5 資料流通管理辦法草案規劃

本節初擬「交通部巨量資料匯流平臺資料流通管理辦法 草案」供參，惟實際執行仍需待「交通大數據諮詢評議委員會」經交通部同意成立且辦法核定通過後為準。

### 「交通部巨量資料匯流平臺資料流通管理辦法 草案」

中華民國 106 年 0 月 0 日交通部字第 00000 號函訂定

#### 第一章 前言

- 一、 為推動中央/地方交通領域相關機關與事業單位（以下簡稱各單位）業務資料相互流通、勾稽分析發揮綜效、翻轉施政服務，故由交通部成立「交通大數據諮詢評議委員會」（以下簡稱委員會），透過交通部管理資訊中心建置之「交通巨量資料匯流平臺資訊系統」（以下簡稱匯流平臺）共享數據資訊，同時訂定本辦法作為資料流通規範之依據。
- 二、 「交通大數據諮詢評議委員會」組織與運作詳見本辦法第二章。
- 三、 有關資料流通分級制度，詳見本辦法第三章。
- 四、 流通資料集之上架/下架辦法與程序請見本辦法第四章。
- 五、 資料集介接申請辦法詳見本辦法第五章。

#### 第二章 交通大數據諮詢評議委員會

- 一、 委員會成立目的與職責在於定期通盤檢視國內交通大數據資料匯流與大數據執行計畫，評估各種資料集是否需要由交通部內統一做資料蒐集匯流與發布，提昇橫向聯繫，促使部內資源發揮綜效，同時能與他部會資料匯流的統一窗口，主要的職責如下：

(一) 協調優先納入匯流之資料集供部屬機關遵循執行。

- (二)審核各單位提出之資料匯流需求暨相關建置計畫。
- (三)代表交通部作為跨部會資料匯流之行政窗口。
- 二、諮詢評議委員會建議設主任委員/副主任委員各一名，負責召開會議，主委與副主委可由交通部次長/主秘擔任或由諮詢評議委員會委員互選之。諮詢評議委員會則分為三大類，第一類為機關代表，第二類則為專家代表，第三類則為民營公司代表。
- 三、承上，第一類機關代表包括中央與地方交通相關機關，機關名單由委員會擬定，機關代表由各機關指派，更換時亦同。第二類專家代表則由委員會聘任交通運輸、資料科學、巨量資料、物聯網領域之專業學者數名，任期為一年但可續聘。第三類民營公司則由委員會視需要邀請從事運輸系統營運或交通資料維運之公司參與。

### 第三章 資料流通權限分級制度

- 一、本辦法遵循「行政院及所屬各機關政府資料分類及授權利用蒐費原則」中規範之甲類/乙類/丙類三大類之定義落實資料權限分級，以下引述該原則之分類說明：

#### (一)甲類資料：

為開放資料，指以開放格式提供，且以無償方式、不可撤回，並得再轉授權方式授權利用為原則者。

#### (二)乙類資料：

為有限度利用資料，指以開放格式提供，且符合下列條件之一者：

1. 有償提供。
2. 保留撤回權。
3. 其他限制條件授權利用。

#### (三)丙類資料：

為不開放資料，指依法律規定不得開放、因資料敏感或有其他特殊情形，經各機關首長核可不予開放者。

- 二、本辦法將「乙類資料」依序限制條件授權的不同，再細分

為「乙1」,「乙2」與「乙3」三子類,分述如下:

(一)「乙1」類資料:

本子類資料僅供公務機關申請流通介接,同時使用單位申請時須簽署資料擁有單位制定之「含罰則之資料保密協定」作為申請要件,同時資料擁有單位保有審核通過與否之權利,即使審核通過後亦保留撤回權利。

(二)「乙2」類資料:

本子類資料與「乙1」相似,但擴大開放國內登記有案之法人(公司行號/學術機關等)申請流通介接。

(三)「乙3」類資料:

本子類資料與「乙2」相似,但擴大開放國內自然人申請流通介接,同時申請審核權交由匯流平臺管理員執行。

三、資料分級由原資料集擁有單位初擬,經委員會同意後執行,惟各單位認定為甲類資料前,應整體考量該資料提供使用之公益性、是否涉及個人隱私或營業秘密、資料提供單位是否有完整權利及得否再轉授權等因素。各單位亦可視資料的流通限制需求認定為乙1~乙3類。若資料敏感或現階段不宜流通,則認定為丙類。

四、綜合以上四點,本辦法之資料分級彙整如表 5.9。

表 5.9 資料流通管理資料分級表

等級代號	甲類	乙1類	乙2類	乙3類	丙類
簡要說明	允許介接 (免申請審核)	允許介接 (限制申請)	允許介接 (擁有者審核)	允許介接 (平台審核)	現階段 不宜介接
允許申請單位	毋需申請	公務機關	公務機關 公司/學校 法人/團體	C級允許單位 + 自然人	公務機關
申請審核單位	毋需審核	資料擁有者 自行審核	資料擁有者 自行審核	授權平台 管理者審核	—
資料集欄位 說明&範例	對外公開	僅對公務機 關帳號公開	對外公開	對外公開	—
備註	可考慮推薦 至其他平台 (OpenData)	申請單位須 簽署保密協 定(含罰則)	資料擁有者 可拒絕申請	可考慮推薦 至其他平台 (ex:PTX)	公文往返 專簽處理 去識別化 資料彙總

- 五、 匯流平臺僅作為資料流通之用，原始資料中若含個資或機敏資料，資料擁有單位本於善盡保護管理職責應於匯流平臺蒐納前以遮罩或去識別化方式處理之。

#### 第四章 資料流通種類分級制度

- 一、 第三章敘述之資料分級係由流通管理權限與限制角度制定之分級制度，本章說明依實際資料種類區分為四級資料，分述如下。
- 二、 一級資料：即原始資料，依照來源單位提供之格式保存，供後續清洗、檢核、統計與彙整使用。
- 三、 二級資料：即正規化資料，資料經過清洗、檢核與正規化後，如有異質資料串接之需求，可依需求增加串接欄位，並以共通資料表綱要格式保存，供後續資料統計與彙整使用。
- 四、 三級資料：即統計資料，資料經過統計彙整後，或來源單位直接提供統計性資料，則以設計確認之格式保存，供使用者查詢與後續彙整使用。
- 五、 四級資料：即主題性資料，設定特定主題，並預先串接相關之異質資料，以多元資料集合方式保存與提供查詢。
- 六、 一般而言，一、二級資料由資料擁有單位提供，三、四級資料由業務單位、研究單位產製回饋。

#### 第五章 資料集之上架/下架辦法

- 一、 為因應需求多樣性，匯流平臺流通的資料集清單並非固定，本辦法明定三種資料集新增方式，包括(1)目標導向型新增機制(2)需求導向型新增機制(3)自主加入型三大類。以下分別說明之。
- 二、 目標導向型之新增機制係由委員會聘任的專家學者驅動，委員會每年將至少召開乙次會議蒐集目標導向流通資料集候選清單，專家委員以其宏觀視野推薦高潛力、高價值之資料集作為流平臺新納入資料集候選清單，經委員會討論篩選納入確定蒐納範疇後，再由相關資料擁有單位成立工

作小組商討流通欄位、分級與上架期程，若現有系統尚未保存關鍵重要欄位，則須編列計畫經費蒐集建置。

三、需求導向型之新增機制係由資料需求單位（限公家機關）提出新資料集上架申請，委員會每年將至少召開兩次會議討論需求導向流通資料集蒐納清單，需求單位提出申請時，至少必須包括以下 5 點資訊作為委員會審核之用。通過者再由相關資料擁有單位成立工作小組商討流通欄位、分級與上架期程。

(一)資料集簡稱暨原始資料維運單位

(二)資料集詳細內容(例如資料表清單、資料欄位型態...等)

(三)資料集應用計畫(例如專案目的、可回饋資料集項目等)

(四)申請機關過去提供匯流平臺的新資料集或回饋資料集

(五)其他備註(例如跨域整合等其他有利於審核通過之說明)

四、前項需求導向新增機制暫時不開放民營公司或大專院校等非公務單位申請，非公家單位若有執行公部門專案而衍生資料需求時必須由專案計畫委託單位提出申請。

五、資料擁有者（或資料維運單位）亦可主動在委員會主動提出欲加入匯流的資料集清單並提供聯繫窗口，經委員會同意後由資料匯流平臺管理單位與窗口聯繫完成介接上架，資料擁有者未來即可不再重複回應各界資料申請需求。

六、各單位提供各項資料上架時，應配合資料集新增機制之資料分級表進行各類資料之授權，且該資料內容有限制使用之必要時，應於使用規範中定明其使用限制。前項使用規範，應包括使用者承諾事項、政府機關責任之限制及終止條件等，其涉及著作權之授權利用者，並應定明授權方式及範圍。

七、匯流平臺資料集除甲類資料不可撤回外，其餘資料均可由資料擁有單位提出撤回要求，經委員會同意後，由匯流平臺進行相關設定。

八、匯流平臺端之資料上架程序主要處理流程為產製資料項目

服務所需的介接/下載檔案/分析後成果頁面 URL 聯結，之後於資料集目錄中增列項目並公開資料集資訊完成上架作業。資料上架時，靜態資料可以產製查詢連結提供使用者預覽、下載，動態資料網站提供瀏覽服務，資料服務需由維運人員可分時空條件進行打包、壓縮後，提供資料下載。

- 九、 匯流平臺端之資料下架程序主要處理程序為刪除相關使用授權以及將資料項目由資料集目錄中移除。

## 第六章 資料集介接申請辦法

- 一、 根據本辦法第二章資料流通分級，甲類資料係屬公開資料毋需申請即可下載流通；丙類資料則因機敏或其他特殊因素不提供匯流與流通，至於乙類資料則根據不同子類在允許申請者、申請程序(保密協定之有無)、審核者有不同限制，其流程差異如圖 5-52。

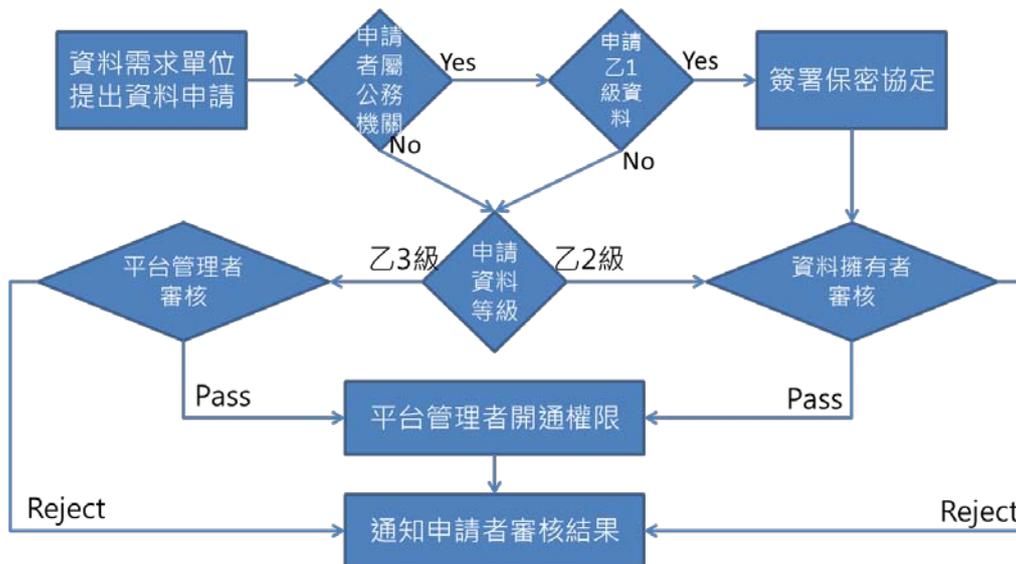


圖 5-52 資料集實體內容申請流程圖

- 二、 為使匯流資料能被廣為流通，匯流平臺會有專區羅列目前所有資料集清單、資料描述、範例資料若干筆，專區將根據資料分類供可存取之單位帳號瀏覽匯流平臺資源，但此專區內僅能下載範例資料，欲實際取得資料須另行申請方能抓取或介接原始資料。
- 三、 匯流平臺之基本流通申請規範包括以下四點，至於乙 1 類之保密協定(含罰則)則由資料擁有者提供。

(一)提供介接的資源內容

介接內容範圍本平臺可授權之應用資料內容，其中可能包含資料庫及網站所蒐集之可公開資料。

(二)在使用時資源來源需註明清楚

介接資料需註明為「OOO 機關來源」之字樣，確保介接單位取得後做不法之用途。

(三)相關法律釐清條款

包含授權資料不得再對外授權、蒐費，另本授權資料僅供參考，申請單位不得進行任何法律求償及訴訟。

(四)介接資料使用期限或終止介接規則

本平臺提供之介接服務皆以一年為限，則當申請期限屆滿時，申請單位需重新進行申請；另外若申請單位在使用上有損機關之形象時，平臺有權立即終止該服務，並且永久不得申請。

- 四、申請介接者須在匯流平臺網頁之「開發指南」閱讀相關申請辦法與注意事項後，進入「線上申請」下載相關表單進行後續作業。申請者需先閱讀申請規範並同意後才可進入申請單頁面；但因本申請單上需蓋有申請單位之印信，因此申請者除線上填註外，仍須下載申請書（含申請規範或保密協定）用印後寄送匯流平臺管理小組，平臺小組蒐件後將視資料級別送交資料擁有者執行審查作業或逕行審查。通過審查者即可取得介接所發佈之資訊。

## 第七章 其他

- 一、本辦法中各單位間流通資料以互惠無償為原則，若有特殊情形可經委員會同意後酌蒐資料處理費。
- 二、本辦法自發布日施行，若有未盡事宜，將依相關規定或委員會討論決議辦理。

## 5.6 其他配套規劃

### 5.6.1 資訊安全及效能規劃

#### 1. 資訊安全及保密規劃

在蓬勃發展網路環境下，隨著資訊便利而來的資訊安全問題，為維護整體資訊安全，強化各項資訊資產之安全管理，確保其具機密性、完整性、可用性，因此本案將資訊安全及保密規劃有以下二類方式來確保資訊不外流。

##### (1) 硬體安全

包含硬體環境控制及人為管理控制等，均須防止資訊外流洩漏。

##### A. 安全機制

每一臺虛擬主機應該要建立應有的防禦機制，如防火牆、防毒軟體、入侵偵測系統等。另外，本案也建議在虛擬化平臺直接加上核心層的防火牆，彼此之間的流量都會可受到檢視，不僅安全得以確保，再加上虛擬主機自身的防禦，可達到雙層的防護。

##### B. 加密及監控

對於敏感資料及安全要求性高的虛擬主機，無論在關機或運作時都應該加密，雖然此作法較繁瑣，也會降低主機的效能，但為了安全性考量，本案建議還是需要以此方式來運作。

##### C. 存取限制

為避免資訊資產因未授權之存取而使機密性或敏感性資料遭不當使用，應考量人員職務授予相關權限，必要時得採行加密及身分鑑別機制，以加強資料之安全。

##### (2) 個人安全防護

包含人身安全、個人隱私權安全等，由於本案需要執行使用者登入之動作，因此建議使用單一簽入(Single Sign On, SSO)系統機制的方式來進行使用者登入。

單一簽入是一種資訊系統登入的控管機制，整合組織內的各項應用系統，集中管理登入的身分辨認作業，同時也配合目錄服務所提供

的功能來控管使用者的權限，當使用者透過網路連接到具有單一簽入的系統時，則具有單一簽入系統在處理登入的控管作業可讓使用者簽入系統後只需通過身分辨認一次，就能讓使用者依據目錄服務所核定的權限，進入或使用經整合後的各個應用系統，不再需要重複作登入時的身分辨認，減少使用者記憶不同系統多組帳號密碼的麻煩，提高使用者操作系統的便利性，其作業流程圖如圖 5-53。

目前我的 E 政府已建有單一簽入系統，本案可直接應用介接該系統來做使用者登入，除了第一次登入者需要先分別到「我的 E 政府」進行帳號申請，以及到本案申請窗口申請登入者層級設定外，之後使用時，即可在輸入帳號密碼，本案之系統會自動連結至單一簽入平臺去進行帳號密碼核對再回傳至本案系統做登入動作，本案系統中並不需要做任何個人保密資料的維護，更加能確保使用者的個資安全。

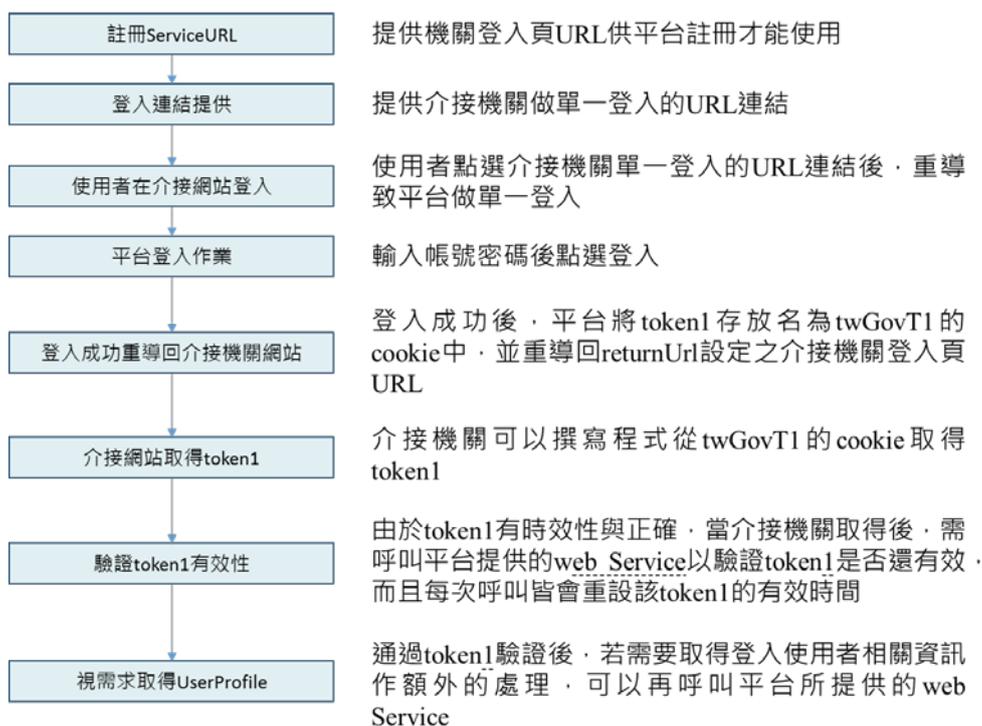


圖 5-53 SSO 作業流程圖

## 2. 系統效能規劃

### (1) 建立看門狗機制(watch dog)，加強系統穩定性

因系統資料龐大，設計較為複雜，為提升系統之穩定性，規劃進行監控試。監控方法在系統中建立看門狗機制(watch dog)進行管理。

該機制將監控所有受託管之程式，當該程式沒有正常報到時，看門狗程式將會視該程式已發生異常，因此會自動將異常的程式重新啟動，如有任何無法重新啟動的異常狀況，將於系統介面當中呈現異常狀況。看門狗程式之架構如圖 5-54 所示。

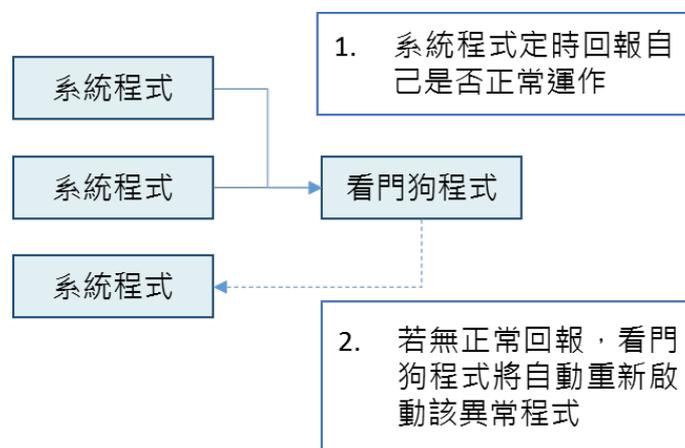


圖 5-54 看門狗程式架構圖

## (2) 系統負載與壓力測試

壓力測試要做到點對點(end-to-end)的監控與分析，從模擬使用者的行為到後端的資料庫，釐清每一個環節的問題。本案將利用 Webserver Stress Tool 軟體進行網站極限測試(ramp test)。

Webserver Stress Tool 是一個可以模擬使用者瀏覽網頁的程式。測試者可以設定欲模擬瀏覽端的使用者人數與點擊網頁連結的頻率，透過 Webserver Stress Tool 進行系統流量的測試。可以提供客觀的測試數據，減輕許多自行測試的麻煩。在測試階段，測試系統會記錄並計算平均回應時間、出錯率與模擬使用者數量。此外，Webserver Stress Tool 亦可提供相當多樣化的網頁流量測試報表，而且可以根據測試使用者或是測試目標 URL 為基準產生報表。管理人員最關心的圖形化報表等，測試結果之範例可參照圖 5-55、圖 5-56 所示。

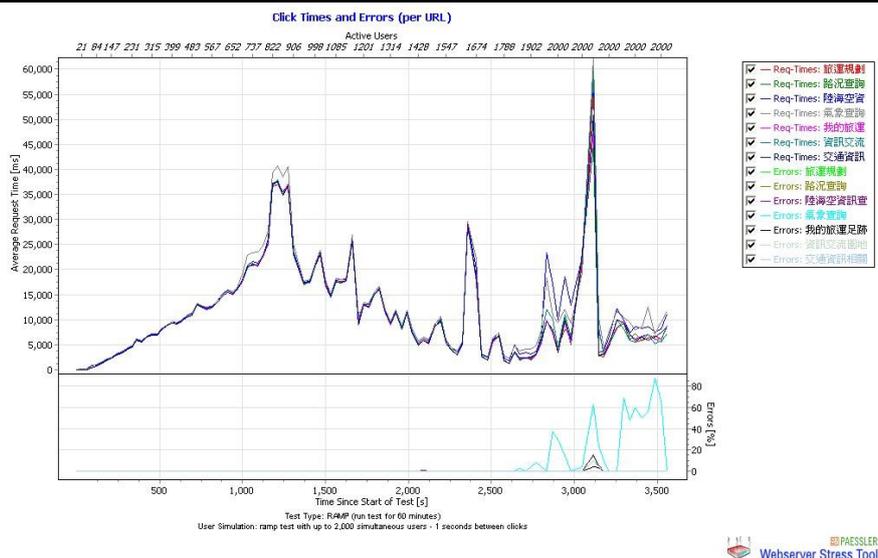


圖 5-55 平均回應時間與錯誤率關係圖

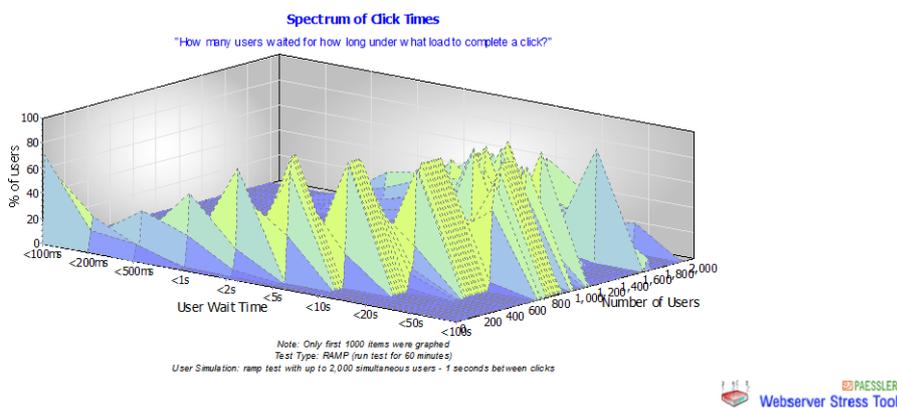


圖 5-56 使用者數量與等待時間波譜圖

## 5.6.2 系統備援與還原規劃

一個永續經營之系統除了在系統建置階段中，需考量各項資料備分作為外，當其進入開發完成進入維運階段以後，備援及備分更形重要，此作法將對於可能會發生的風險預先進行準備防範於未然，能夠維持其高可用性，減少系統受到損害時所造成的損失。

### A. 系統備援計畫

為確保系統資料處於最佳狀態，因此本案建議定期備分系統資料庫，並在必要時之額外備分作業。但若要確保系統資料處理最佳狀態下，應採分臺備援的狀況，預設需有多臺伺服器可以相互運用，再利用設定資料庫排程備分維護計畫進行執行。其執行方式與流程應如圖 5-57 所示，後續當實際系統建置完成，

陸陸續續匯入大量資料可再行研議。

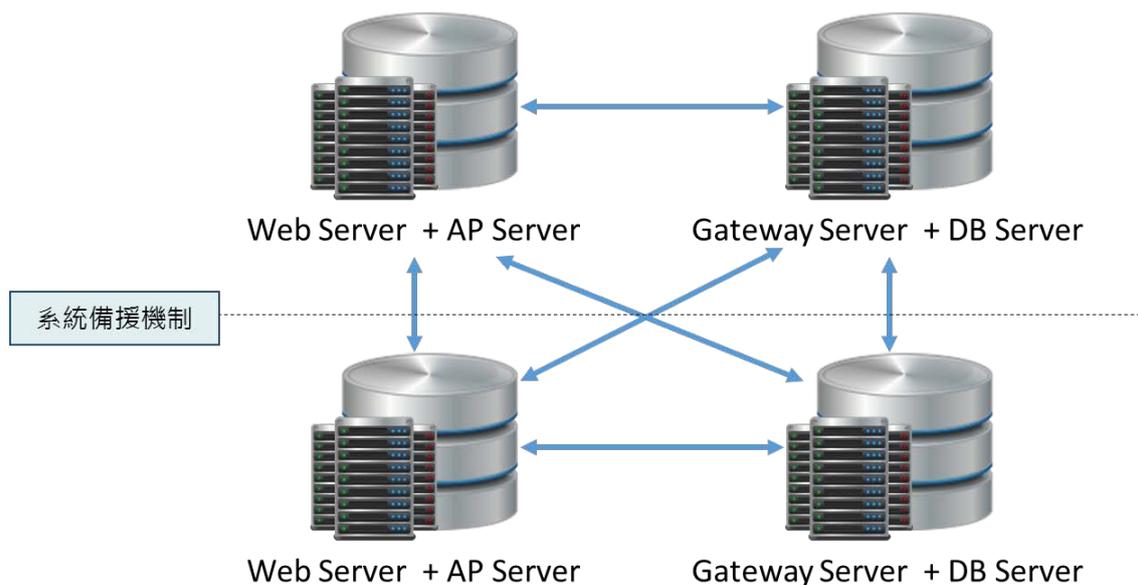


圖 5-57 備援計畫分佈圖

針對系統資料之備分，建議採用「三種備分方式綜合使用」。利用每週末做「完整資料庫備分」、每天清晨做「異動式資料備分」，以及每天八小時做一次「交易記錄備分」的方式實行資料庫備分的工作。以下將分別介紹方式：

(a) 完整資料庫備分

此方式需要相當龐大的儲存空間來存放備分資料，所以不建議每日做如此的備分方式，不過此方法的好處是將來回存資料的時候，只需將整個資料庫從一分資料庫備分回存到資料庫即可。

(b) 異動式資料備分

此方式所備分的資料是從上一次執行「完整資料庫備分」後有更動過的資料，因此所需的時間和儲存空間會比「完整資料庫備分」的方式來的少。

(c) 交易記錄備分

此方式則只備分交易記錄檔的內容，並記錄對資料庫所做的異動過程，也就是只記錄資料庫的異動情形。

在此三種方式互相搭配下，可取其長處，減少短處，使得當發生問題時，也能讓資料庫回到最近八個小時的狀態，讓損失

的程度縮小，但同時又能兼顧到回復過程不致於太複雜，以及耗用過多備分資源(時間與空間)等問題。針對運行的系統，設定完整的資料庫備分計畫，並定期將備分檔案備分至磁碟、磁帶機或其他資料儲存裝置。

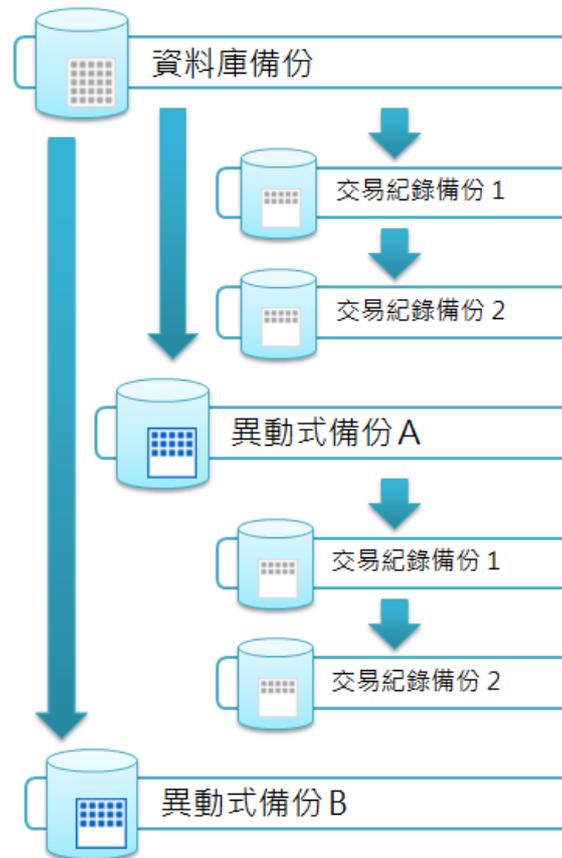


圖 5-58 系統備分示意圖

## B. 系統還原計畫

進行系統與資料庫的備分，主要目的在於當系統發生損害時，能夠即時回復正常的運作，因此對於日常維護中所進行的備分，我們必須要有一套復原計畫來讓系統迅速地回復正常運作，將系統資料還原至資料庫發生損害前之最終狀態。

在進行還原計畫之前時，必須先進行系統備分計畫。需先回存「完整資料庫備分」，如此才可將「異動式備分」和「交易記錄備分」回存；然後再將最近一次的資料庫備分回存到伺服器後，再繼續回存後續的「異動式備分」和「交易記錄備分」的資料。假設，星期三中午發生資料庫損毀的情況，則第一個步驟必須先回存上週末的「完整資料庫備分」，然後再回存當天

清晨所做的「異動式備分」，最後一步則是回存早上到中午的所有「交易記錄備分」，如此，可進可能的將損害降到最低。其流程如圖 5-59 所示：

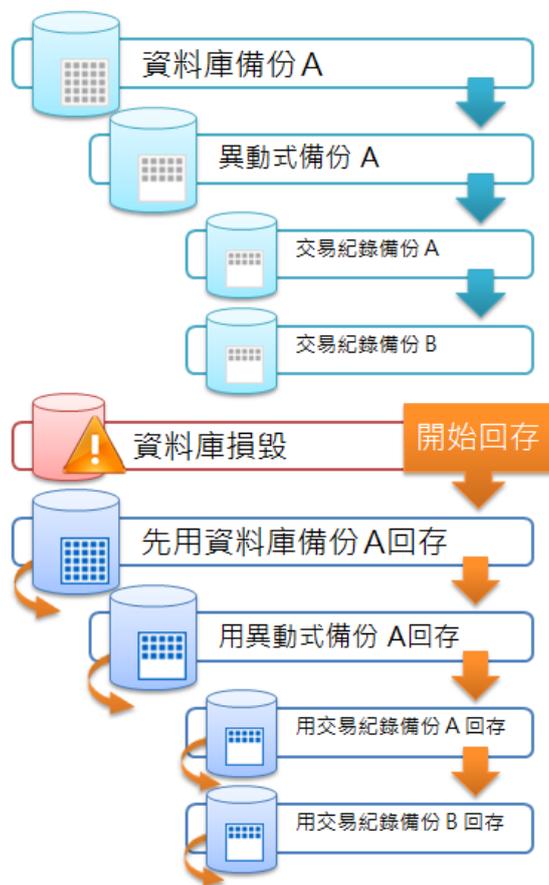


圖 5-59 系統還原示意圖

a. 災害復原計畫

一個完善且有效率的災難復原計畫(Disaster Recovery; DR)，可以在遇到各種停機狀況(如同伺服器損毀、資料遺失)時，快速地回復系統運行，降低資料損毀的可能性。

而依據不同的災難發生原因，未來之實際作業方式將與 貴部資訊系統中心進行討論，以下為本案之初步建議：

(a) 硬體伺服器的損毀

硬體伺服器可以透過備援主機或是建立虛擬主機方式備援、於主要伺服器運作失效時，可接手處理最快時間復原。

(b) 異動式資料備分

資料的遺失，可利用「備分暨還原計畫」來做保護。

## 第六章 結論與建議

本研究針對國外交通領域大數據應用作完整回顧，透過訪談國內相關單位之需求，據以提出建議，其中資料匯流平臺是推動大數據分析應用之核心基礎，故本研究亦針對資料匯流平臺進行初步規劃並研擬資料流通管理辦法草案；同時以公共運輸為例執行標竿案例分析。本研究所得之結論與建議整理如下，可做為未來推動交通大數據分析應用或政策創新之重要參考依據。

### 6.1 結論

1. 本研究分別整理國外 28 個交通大數據應用案例，涵蓋公路、鐵路、航空、航港四大類運具，其應用目的可歸納為蒐集需求、便民資訊、提升安全、資產管理、延誤改善、輿情分析、節能減碳、追蹤動態以及效能提升等九大類，詳細內容可參閱第 3.1 節。
2. 本研究共計訪談桃園機場公司、高鐵局、臺鐵局、民航局、路政司、公路總局、基隆港務分公司、航政司以及航港局等九個交通單位對大數據之關注議題，詳細內容可參閱第 3.2 節。
3. 綜合前述需求，本研究歸納整理出(1)制定數據戰略(Stratgy)、(2)共享旅運資訊(Message)、(3)強化資產管理(Asset)、(4)落實風險評估(Risk)、(5)翻轉營運服務(Transformation)等五大策略，期能從 Big Data 昇華為 Smart Data，進而達到 Smart Transportation，意即「智慧數據引領智慧交通」，利用數據形成政策與決策。
4. 本研究與各交通單位訪談時深切感受實務上資料蒐集流通之困難，除彙整現況問題外，亦建議一套資料流通之協調機制供參，相關內容詳見第 3.3 節。
5. 資料匯流平臺之軟硬體及功能規劃設計如第五章，其中第 5.5 節草擬資料流通管理辦法草案，未來需經交通部同意成立「交通大數據諮詢評議委員會」後之辦法為準。

6. 本研究以臺中市區公車為標竿案例對象，從旅客觀點出發，分別針對「舒適」、「密集」、「信賴」、「安全」與「無縫」等五個面向進行交通大數據分析之研究，可提供主管機關未來在營運管理與提昇公共運輸服務水準上之參考。

## 6.2 建議

1. 由於各機關資料數據均龐大，在研究時程的限制下，本研究建議以「需求導向」確認機關欲解決的問題後再盤點相對應之資料集後再提出資料匯流共享機制，未來可以此為基礎，逐年拓展資料集，並探索有無機會進行大數據分析應用。
2. 電子票証是公共運輸巨量資料分析之核心，甚至有機會能串連多項運具甚至非交通的生活連結，但目前電子票証的格式設計主要為蒐費之用，因此應用上仍存在諸多課題，例如公車站牌資訊闕漏、公共自行車借用時間闕漏、臺鐵進站時間闕漏等。面對這些問題除短期發展演算法插補推估外，建議長期應系統性補足解決，方能讓電子票証資料在交通大數據應用上產生更多價值。
3. 電子票証與公車動態資料為目前最主要的兩項資料來源，但這兩項的資料，無論是在內部或是這兩項資料之間，均存在著不一致性，尤其是兩項資料在車牌、站牌的一致性對應部分，往往造成分析的費時，也減少了可分析的資料數量，建議未來應該針對車牌、站牌的編碼與異動更新應該有一套系統的方式，若以上改善均完成，將大幅提昇原始資料品質，同時亦可讓後續研究針對轉乘行為做深入分析。
4. 本研究受限於時間，僅針對旅客觀點之服務品質角度分析「舒適」、「密集」、「信賴」、「安全」與「無縫」等五個面向，未來建議可進一步深入分析，例如在無縫方面，可分析各車站在離峰之公車轉臺鐵或臺鐵轉公車之平均轉乘時間，除此之外，亦可嘗試引入營運者觀點或偏遠路線補貼等不同角度之分析。
5. 搭配電信定位資料來分析公共運輸課題是未來極具潛力的研究方向，有助於現有公共運輸路線的調整、大眾運輸使用率的提升以及潛在目標族群的開發等。

## 附錄 A 教育訓練辦理情形

本專案依照合約要求辦理三場教育訓練，前兩場之主題為進階 R 語言實作暨應用實例講座，第三場則為 Tableau 商業軟體之使用訓練，此軟體是一套新商業智慧(BI)與資料科學(Data Science)分析軟體，提供強大的巨量資料(Big Data)處理與視覺化分析能力。以上訓練分別在 2016 年 3/31、9/8 與 10/28 辦理完成。以下彙整教育訓練辦理情形。

### R 語言進階實作辦理情形

1. 訓練時間：2016 年 03 月 31 日(星期四) 09:00~12:00
2. 訓練地點：交通部運研所 5 樓會議室 (台北市敦化北路 240 號)
3. 訓練講師：臺北商業大學資訊與決策科學研究所 鄒慶士教授
4. 訓練內容：如表 A.1。

表 A.1 R 語言進階實作訓練課程表

主題	時間	教學案例
時間序列預測	AM09:00~AM09:40	衛星影像序列資料、氣象測站溫度序列資料、各地區選舉結果序列資料、序列資料、人畜移動軌跡序列資料、無線網路 AP 熱點資料、疾病和火山爆發序列資料等
時間序列集群	AM09:40~AM10:20	
休息	AM10:20~AM10:30	
空間點型態分析	AM10:30~AM11:10	
空間屬性分析	AM11:10~AM11:50	
問答	AM11:50~PM12:00	

5. 訓練對象：簽到如表 C.2。

表 C.2 R 語言進階實作課程簽到表



交通部運研所 R 軟體課程簽到暨講義領取表

一、課程時間：2016/03/31 AM 9:00~ PM 12:00

二、課程地點：交通部運輸研究所 5F

服務單位	職稱	姓名	服務單位	職稱	姓名
交通大學	學生	蔡志修			
中興社	研究員	天明軒			
景翔	工程師	林翰			
中興社	研究員	黃笠珽			
運研所	助理研習	傅珺			
運研所	助理研習	劉齊甄			
運研所	組長	洪朝傑			
"	研究員	周文靜			
"		鄭德偉			
"	助理研究員	呂恩慧			
"	研究員	劉坤澤			
"	副研究員	柏立仁			
"	研習員	陳汝和			
景翔	工程師	郭偉			
景翔	工程師	白家欣			
運研所	研究員	賴靜慧			
		史習平			
		陳佩蓉			
		江芷瑛			
運研所	研究員	蘇欽同			
中興社	研究員	張頌楠			

6. 活動剪影：如圖 A-1。

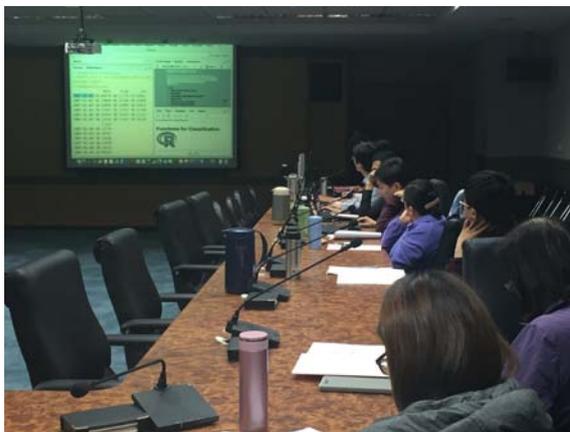


圖 A-1 R 語言進階實作訓練課程實況剪影

## R 語言應用實例講座辦理情形

1. 訓練時間：2016 年 09 月 8 日(星期四) 09:00~12:00
2. 訓練地點：國立台北商業大學行政大樓五樓演講廳
3. 訓練講師：臺北商業大學資訊與決策科學研究所 鄒慶士教授
4. 訓練內容及現場電子告示牌：如表 C.3。

表 C.3 R 語言應用實例講座訓練課程表

時間	活動內容
09:00-10:20	氣象資料分析案例
10:20-10:40	休息
10:40-12:00	臺中公車轉乘分析案例
12:00	賦歸

演講人：鄒慶士 教授  
活動時間：2016年09月08日  
活動地點：國立臺北商業大學 行政大樓五樓演講廳

主辦單位：中興工程顧問社  
協辦單位：中華R軟體學會、臺灣資料科學與商業應用協會

5. 訓練對象：簽到如表 A.4。

表 A.4 R 語言應用實例講座課程簽到表

「MOTC-IOT-104-IEB048 交通大數據分析與應用機制先期規劃」  
R 語言應用實例講座  
簽到表

一、時間：中華民國 105 年 9 月 08 日(星期四)上午 09 時 00 分

二、地點：台北商業大學行政大樓 5 樓演講廳

三、主講人：鄒教授慶士

四、出席單位及人員：

出席單位	簽到
交通部運輸研究所	
交通部運輸研究所	陳翔捷
財團法人中興工程顧問社	張恩楠
財團法人中興工程顧問社	胡仲璋
財團法人中興工程顧問社	黃竺琄
財團法人中興工程顧問社	劉雪乳
財團法人中興工程顧問社	蔡淑沂

楊世均 梁景濤 陳奕瑄 鄧曉雅 卓怡君

莊舒媛

6. 活動剪影：如圖 A-2。

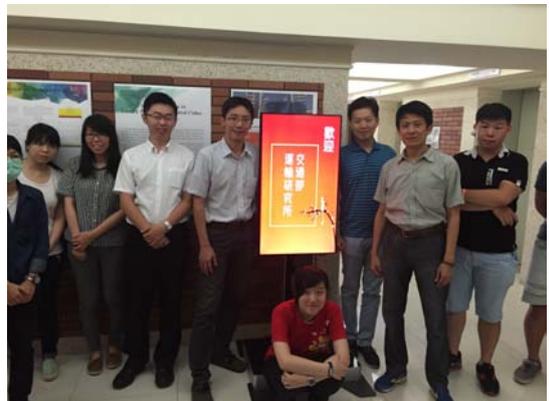


圖 A-2 R 語言應用實例講座課程實況剪影

## 商業智慧軟體訓練辦理情形

1. 訓練時間：2016 年 10 月 28 日(星期五) 09:30~17:00
2. 訓練地點：交通部運研所 5 樓會議室 (台北市敦化北路 240 號)
3. 訓練講師：美商天睿台灣分公司郭柏森工程師
4. 訓練內容：課程如表 C.5，演練案例如表 A.6。

表 C.5 Tableau 訓練課程表

Agenda	Schedule	Duration
Introduction to Tableau	09:30 - 10:00	30 min
The Tableau Desktop Demo	10:00 - 10:30	30 min
Quick Break	10:30 - 10:50	20 min
Review pre-workshop question	10:50 - 11:00	10 min
Tableau Desktop Hands-On	11:00 - 12:00	1 hour
Lunch Break	12:00 - 13:30	1 hour
Tableau Desktop Hands-On (continued...)	13:30 - 15:30	2 hour
Quick Break	15:30 - 15:50	20 min
Tableau Desktop Hands-On (continued...)	15:50 - 16:50	1 hour
Supporting Documents	16:50 - 17:00	10 min

表 A.6 Tableau 訓練演練案例表

題目	說明	子題	子題
Challenge 1	折線圖	a: Calculated Field	
Challenge 2	Group By		
Challenge 3	drill down		
Challenge 4	交叉分析表		
Challenge 5	交叉分析表且輔以顏色		
Challenge 6	地理資訊分析		
Challenge 7	多商品視覺化分析	a: Data Blending	b: Set
Challenge 8	子彈圖	a: Parameter	
Challenge 9	儀錶板		
Challenge 10	多圖層堆疊分析	a: Actions	
Challenge 11	利潤曲線		
Challenge 12	功能發布		

5. 訓練對象：簽到如表 A.7。

表 A.7 Tableau 軟體訓練課程簽到表

「Tableau」視覺化分析軟體教育訓練  
報名表

時間:105 年 10 月 28 日(五)上午 9:30-12:00 13:30-17:00

地點:本所五樓會議室

授課人:Teradata 公司

姓名	職稱	身分證字號	簽到
黃明正	研究員		
林暹耀	助理研究員		林暹耀
陳佩茶	研究員		陳佩茶
史習平	研究員		史習平
張恩輔			張恩輔
黃竺珺			黃竺珺
胡仲璋			胡仲璋
陳翔捷			陳翔捷
李水川			李水川
曾騰毅			曾騰毅
蔣明峰			蔣明峰
張宏正			張宏正
林郁雯			林郁雯
白亞菁			白亞菁
劉雲雅			劉雲雅

6. 活動剪影：如圖 A-3。



圖 A-3 Tableau 軟體訓練課程實況剪影

## 附錄 B 參考文獻

1. 鄔德傳，出席第 93 屆 TRB 年會報告，民國 103 年 4 月。
2. Anwar, A., Odoni, A., and Toh, N., “BusViz: Big Data for Bus Fleets”, *TRB 95th Annual Meeting Compendium of Papers*, 2016.
3. Cats, O., and Fadaei Oshyani, M., “An empirical evaluation of improve bus service reliability: Performance metrics and a case study in Stockholm”, *Conference on Advanced Systems in Public Transport*, 2015.
4. Davidson, A., “Big Data Exhaust for Origin-Destination Surveys: Using mobile trip-planning data for simple surveying”, *TRB 95th Annual Meeting Compendium of Papers*, 2016.
5. Du, F., Brul'e, J., Enns, P., Manjunatha, V., and Segev, Y., “MetroViz: Visual Analysis of Public Transportation Data”, *CoRR abs/1507.05215*, 2015.
6. <http://dataconomy.com/predictive-maintenance-big-data-on-rails/>
7. <http://pt.slideshare.net/capgemini/presentation-of-tibco-architecture-week-2013-sweden>
8. <http://www.airline.ee/qantas/qantas-deploys-amadeus-analytics-to-reduce-impact-of-flight-disruptions/>
9. <http://www.cargosmarton.com/look-out-for-global-vessel-disruptions-on-big-schedules/>
10. <http://www.computing.co.uk/ctg/news/2431585/qantas-airways-deploys-a-madeus-analytics-to-reduce-impact-of-flight-disruptions>
11. <http://www.i4u.com/2015/09/94906/swedish-train-operator-can-see-future-delays-using-big-data>
12. <http://www.lta.gov.sg/apps/news/page.aspx?c=2&id=a759467d-4ff4-4bd0-9a78-79f7fb7772a4>
13. <http://www.nyk.com/english/release/4208/004274.html>
14. <http://www.slideshare.net/Dell/big-data-use-cases-36019892/>



15. <http://www.tibco.com/assets/blt90217e2671599659/ss-cargosmart.pdf>
16. <https://share.inside.com.tw/posts/20169>
17. [https://www.cargosmart.com/sc/company/about\\_us.htm](https://www.cargosmart.com/sc/company/about_us.htm)
18. [https://www.porttechnology.org/news/insight\\_smart\\_ships\\_talk\\_to\\_smart\\_containers](https://www.porttechnology.org/news/insight_smart_ships_talk_to_smart_containers)
19. <https://www.washingtonpost.com/news/innovations/wp/2015/10/06/how-artificial-intelligence-could-lead-to-self-healing-airplanes/>
20. Lee, R., K., W., and Kam, T., S., “Time-Series Data Mining in Transportation: A Case Study on Singapore Public Train Commuter Travel Pattern”, *International Journal of Engineering and Technology*, 2014.
21. Tulinda, L., “Cloud-Based data warehousing to power aviation analytics”, *Transportation Research Board Annual Meeting*, 2014.
22. Williams, T., and Betak, J., “Using Text mining and Data Visualization to Analyze Railroad Grade Crossing Accidents”, *The Newsletter of the Railway Applications Section*, 2015.
23. Zhang, L., Wang, H., and Meng, Q., “Big data based estimation for ship safety distance distribution in port waters”, *Journal of the Transportation Research Board*, Volume 2479, 2015.

## 附錄 C 期中報告審查意見回覆表

一、開會時間：105 年 7 月 20 日下午 2 時

二、開會地點：運研所五樓

三、主持人：林所長信得

紀錄：陳翔捷

四、出席單位及人員：(略)

五、主席致詞：(略)

六、簡報：(略)

七、討論：

與會代表 (依發言順序)	審查意見	回覆辦理情形	主辦單位 查核意見
淡江大學運輸管理學系張教授勝雄	1. 建議報告書中第 3.1 節(國外交通大數據應用)之案例標題應以案例應用方式呈現而非國家或營運單位。	遵照評審意見辦理。	同意辦理
	2. 認同大數據應採用需求導向切入，但研究單位做相關訪談時，單一受訪人會侷限某個關注議題及需求。	圖 3.44 中的訪談對象僅是主要聯絡窗口，實際訪談時，該窗口會召集該機關相關同仁共同與會。	同意辦理
	3. 建議盤點出交通部既有的數據，包含公路、航空、海運、鐵路，並探索有無機會取得其他資料進行結合。	受限各機關資料數據均龐大，本研究以「需求導向」確認機關欲解決問題後再盤點相對應之資料集，過程中已有許多困境須面對，若要在無任何需求基礎下取得各單位所有數據，執行難度過高，建議後續研究分領域、分機關逐年執行。	同意辦理
	4. 有關標竿案例分析，報告書呈現係以資料探索及視覺化成果為主，與最嚴謹之大數據定義未盡相同。此外，標竿案例偏重於電子票證公車動態資訊系統，但客	本研究利用離線歷史資料進行分析，的確未符合大數據 3V 中 Velocity 的嚴謹定義，主因在於部份資料(如電子票證)未有即時回傳機	同意辦理

	<p>運業的電子票證資料同時已含站位資料，亦即沒有公車動態資訊應也完成分析。另建議借鏡韓國貓頭鷹夜間公車，其乃結合手機基地台資料而開闢新路線，代表若能連結其他資料即可發展更多應用。</p>	<p>制，即使離線取得也是在運研所長官大力協助下完成。電子票證的確可取代公車動態資料的部份功能，唯獨在偏遠路線或路線起終點車上無乘客或僅有投現乘客時，就須仰賴公車動態資料。至於結合電信資料分析，本案在運研所長官的引領下，曾拜訪國內某電信公司討論合作研究事宜。</p>	
<p>淡江大學運輸管理學系羅副教授孝賢</p>	<p>1. 本報告書為先期規劃，對於架構、策略、機制等上位概念應多加著墨。另外強調機制的重要性，應先定義出大方向以利後期執行。</p>	<p>遵照評審意見辦理。</p>	<p>同意辦理</p>
	<p>2. Big data 導入流程等建構有 7 個步驟，其中隱私與商業模式很重要，由於個資法致資料取得難度增加，因此隱私定義及權責劃分須更明確。此外建議應建立 Business model 以利機制永續發展，並訂定規則導入民間創意進行加值更是關鍵。</p>	<p>評審觀點極為關鍵，將在報告中適當放入相關論述。</p>	<p>同意辦理</p>
	<p>3. 建議盤點既有資料，讓資料變資訊才有價值。大數據分析應用須從服務角度而非管理角度切入，因此不僅要符合政府施政需求，更應滿足一般使用者或業者之需求。</p>	<p>評審觀點極為關鍵，將在報告中適當放入相關論述。有關盤點資料議題則同張委員審查意見 3 之回覆。</p>	<p>同意辦理</p>
	<p>4. 計畫書中優先推動原則，建議以由下而上，從使用者觀點出發。例如，臺北市公車改革，民眾的需求不會是上/下車刷卡紀錄起迄，而是候車時間更短、費</p>	<p>遵照評審意見辦理。</p>	<p>同意辦理</p>

	用更低。		
	5. 在翻轉營運服務方面，除了透過大數據讓服務更有效率外，也要檢視現有的法規或制度是否鬆綁或允許業者能彈性(動態)排班。	遵照評審意見辦理。	同意辦理
	6. 在落實評估風險管理方面，建議透過工作日誌、相關工時，檢視有無疲勞駕駛、不良駕駛行為。	遵照評審意見辦理。	同意辦理
逢甲大學運輸科技與管理學系李副教授克聰	1. 建議在報告書第一章補充研究流程(Flow Chart)圖，最後一章則應放入後續工作項目。	遵照評審意見辦理。	同意辦理
	2. 標竿案例內容主要為資料探勘及視覺化呈現，與嚴謹的大數據定義仍有落差。	本研究利用離線歷史資料進行分析，的確未符合大數據3V中Velocity的嚴謹定義，主因在於部份資料(如電子票證)未有即時回傳機制，即使離線取得也是在運研所長官大力協助下完成。	敬悉
	3. 建議思考如何透過大數據提升搭乘大眾運輸使用率，除盤點現有資料之外，也要開發潛在目標族群。	本研究在運研所長官帶領下，曾拜訪國內電信公司，欲利用電信資料來探索路廊中未使用公共運輸之民眾屬性，應可作為開發潛在目標族群之基礎。	敬悉
	4. 標竿案例中談到五個分析面向，請補充說明選擇原因。其中舒適面向應定義合理的供需範圍值。	遵照評審意見辦理。	同意辦理
	5. 除區間營運之調整外，可考慮採更積極之營運調整手段，例如路線調整。	路線調整考慮因素眾多，故擬於最末章節建議後續以專案方式執行路線調整。	同意辦理
	6. 「預約式大眾運輸」是最理想大眾運輸模式，期待隨著大數據	欲達到「預約式大眾運輸」，目前的法規及軟硬體環境均需要調整，有賴各界	敬悉

	的進步能早日落實。	努力。	
臺北醫學大學 醫務管理學系 謝教授邦昌	1. 建議報告書中文獻參考須清楚標示對應。	遵照評審意見辦理。	同意辦理
	2. 訪談單位過於侷限，建議與其他相關單位共同討論。以健保資料大數據為例，包括醫生、護士、病人，各方看法不一致，皆須列入考量。	本研究除期中階段訪談單位外，亦已於5月邀請相關單位與專家學者參與座談會，下半年會持續訪談並再舉辦一場座談會。	同意辦理
	3. 建議五個面向中安全議題要加以著墨，例如，當客運載客量過多，人潮流量造成客運設備損害的疑慮。	遵照評審意見辦理。	同意辦理
	4. 建議報告書中方法論應加強整合，不宜僅以單一方法分析呈現。	遵照評審意見辦理，以第4.3節為例將整合時間序列預測、集群分析及系統模擬三種方法處理班表調整議題。	同意辦理
	5. 本案應強調Big data中Value的重要性，建議可整合非交通資料，找出更有意義的價值，及發現更多問題所在。	遵照評審意見辦理。	同意辦理
交通部公路總局	1. 建議資料共享機制分級化，依資料安全分為普中高等級，若各機關使用資料等級是普級，不需重複申請。	未來資料流通管理辦法擬制定兩種等級，上架前由資料擁有者決定分級，一般性資料級採類開放資料處理，另一級別則須簽署保密協定並配合相關運用管制。	同意辦理
	2. 本報告書中大數據應蒐集的資料數量、資料來源及蒐集手法，再詳細確認。回饋的資料應該要更即時以便能讓使用者即時運用。	即時介接資料確實符合大數據3V中Velocity之要求，惟此目標並非一蹴可及，部份資料集可能先求有(批次離線)再求好(即時)為宜。	敬悉

	3. 第三代公路監理資訊系統(M3)資料機敏性高，各單位申請使用必須經交通部同意，總局會以去識別化方式提供。	感謝評審提供之資訊。	敬悉
	4. 將來各部屬機關提供給交通數據匯流平臺之資料，應明確統一異質資料之格式。	實務上在介接資料時，平台與資料提供者會有多次協商，以兼顧格式一致及提供者窒礙難行之顧慮。	敬悉
交通部管理資訊中心	1. 資料蒐集的方式，團隊是以委員會的方式進行，建議把六都代表放入委員會組成中。另外請釐清 Open data 資料諮詢小組和大數據諮詢小組是否整併？	遵照評審意見辦理。	同意辦理
	2. 報告書 P5-27 提及有關公共運輸資料倉儲規劃，國道客運及公路客運已有新的規範標準，請團隊修正。	遵照評審意見辦理	同意辦理
	3. 有關臺鐵、高鐵等電子票證規格及欄位，是否要統一？另建議有關標竿案例，未來也可釋出資料給其他單位做研究。	擬於最末章節建議未來研究彙整臺鐵第四代票務系統(建置中)與高鐵、捷運票務規劃統一欄位。資料釋出部份則須尊重資料集擁有單位的意願辦理。	同意辦理
	4. 交通部交通數據匯流平臺之現況分為三大類(安全、路況、運輸)及支援性一大類(社經資料等)。	感謝評審提供之資訊。	敬悉
	5. 目前管理資訊中心與運研所兩個專案同時進行，為避免有重工疑慮，建議運研所主要做資料需求分析，管理資訊中心做交通數據匯流平臺資料蒐集。	感謝評審提供之意見。	敬悉
運計組書面審查意見	1. P3-32，本案除了解國內交通機關之需求及相關資料狀況外，建議亦應探討各運具系統在大數據	如同審查會議中眾委員之意見，大數據適合以需求(問題)導向方式規劃，故本	同意辦理

	領域發展方向、需要何類資料產出及資料整合問題。	訪談並綜整各運具之需求提出交通大數據發展策略，詳見期中報告 3.4 節。	
	<p>2. P3-45 現況問題描述部分，提供相關建議如下：</p> <p>(1) 資料擁有者現階段亦未釐清或認知本身擁有何種資料及其應用之範圍。</p> <p>(2) 資料匯流平台亦需考量整合各機關資料時，資料其格式、流通及去識別規範、介接方式以及資料安全性。</p> <p>(3) Open Data 平台之資料完整度以及後續維護問題。</p>	感謝評審提供之意見，將全數納入期末報告 ch3.3.1。	同意辦理
	3. 本案除擴大前案公車動態資訊系統數據之分析範圍及深度，建議往後應著重於交通領域應用大數據技術分析之規劃及優先推動之建議。	遵照評審意見辦理。	同意辦理
	4. 本報告書廣蒐整理文獻，其內容豐富多元值得肯定，惟期中報告初稿共 5 個章節中，3 個章節皆有文獻回顧，建議將文獻回顧統整為一個章節	三個章節之文獻回顧定位有明顯區隔，CH2 首重非運輸領域之大數據應用，CH3 則依鐵/公/空/港領域逐一介紹國外應用案例，CH5 則係以資料匯流平台觀點回顧文獻，每章節均會帶出小結引導後續的規劃方式，故不建議將之合併。	同意辦理
運資組書面審查意見	<p>1. 第三章「國外交通大數據案例回顧」部分</p> <p>(1) 請修正或刪除 3.1 節中公路運輸領域第 6、7、8、9、10 與 11 案例，上述案例部分非關大數據應用，僅是資料視覺化呈現或資訊發佈，若經重新評估後仍為大</p>	遵照評審意見辦理。	同意辦理

	數據應用案例，請團隊強化該案例之說明(數據種類、資料勾稽整合方式或演算法)。		
	(2)請強化說明3.2.2節中臺鐵局「預防性維修」與大數據應用之關聯性。	遵照評審意見辦理。	同意辦理
	(3)請補充3.3.2.3節中有關匯流資料共享機制之資料實體內容申請方式，研提流通管理規範內容，以使各資料需求單位能透過交通部管理資訊中心所建置之數據匯流平臺申請資料集。	此部份將於第五章中以「資料流通管理辦法」專節說明之。	同意辦理
	2. 第四章「標竿案例分析」部分 (1)請補充4.3.1.3節區間車衝突互動評估儀表板之設計開發流程，包含運用 SQL 資料庫整合運算後之資料欄位說明、Tableau 第 I-IV 區連動圖表運用之方法。	將於儀表板交付時隨付之系統文件中說明欄位清單。	同意辦理
	(2)請補充4.3.2.2節最適班表演算法流程，包含 SQL 資料庫之轉換、C 語言模擬工具之計算與 Tableau 乘載狀況模擬之互動關係。	遵照評審意見辦理。	同意辦理
	(3)請補充4.3.2.2節最適班表 R 語言之 K-mean 演算法分群後之視覺化分析結果，以及將 R 語言輸出至 Tableau 之變數名稱，並說明兩種工具整合分析之流程。	遵照評審意見辦理。	同意辦理
	(4)請修正4.4.3節「站牌班表推動可行性分析」中選定標的分析路線(市區路線與郊區路線)之緣由，以及4-66-4-73車輛行駛時空圖中，由 Tableau 儀表板勾選	遵照評審意見辦理。	同意辦理

	之跨日時間範圍。		
	(5)請強化說明4.6.1節中運用電子票證跨運具分析未來之議題(如班表整合、旅客搭乘軌道運輸、汽車客運與公共自行車之轉乘旅次分析)。	遵照評審意見辦理，將於期末報告說明。	同意辦理
	3.第五章「資料匯流平臺」部分 (1)5.1節之國外資料匯流平臺回顧案例請增加美國SHRP「InSight&InDepth」資料服務平臺模式。	遵照評審意見辦理。	同意辦理
	(2)5.3.3節之服務網站規劃，所使用之資料內容與功能規劃示意圖，請依本案與管理資訊中心聯合工作會議中所討論第一階段匯入之資料欄位，進行呈現，其內容應包含分級資料屬性、資料開放查詢條件(時間條件、空間條件)與資料欄位。	遵照評審意見辦理，將以載具軌跡資料、電子票證、監理資料及肇事資料之欄位為設計原則。	同意辦理
	(3)請以公車動態資訊系統、電子票證資料及臺鐵資料為例，依3.3.2.3節中所規劃之匯流資料共享機制1、2，說明申請者自匯流平臺瀏覽清單與下載範例資料之流程，並在網站功能中規劃各資料權責單位如何利用該平臺，審核各單位之提案申請，以使平臺之實體審核機制與本案未來所研提之流通管理規範能無縫接軌。	遵照評審意見辦理，實際作為與流程規劃將與團隊內部人員進行討論後提出。	同意辦理

主席結論	1. 有關交通部未來推動之大數據計畫，請計畫執行團隊針對 SMART 五大面向，找出具體工作項目並細部規劃辦理方式。	遵照主席裁示辦理。	同意辦理
	2. 請將本案未來訂定之流通管理辦法與交通部管理資訊中心所建置之數據匯流平臺進行整合。	遵照主席裁示辦理。	同意辦理
	3. 本案期中報告原則審查通過，請中興工程顧問社針對會議中與會委員及本所運資組意見進行檢討修正，於會議紀錄文到 1 週內研提回覆意見辦理情形說明表（並納為期末報告附件）。	感謝主席與諸位評審之肯定。	敬悉

## 附錄 D 期末報告審查意見回覆表

一、開會時間：105 年 12 月 21 日下午 2 時

二、開會地點：運研所五樓

三、主持人：吳副組長東凌

紀錄：陳翔捷

四、出席單位及人員：(略)

五、主席致詞：(略)

六、簡報：(略)

七、討論：

與會代表 (依發言順序)	審查意見	回覆辦理情形	主辦單位 查核意見
交通大學運輸與物流管理學系盧教授宗成	1. 1-3 頁之研究目的建議強化說明與本案成果之關聯性	根據評審意見，於 1.3 節加強說明各項工作成果與研究目的之關聯。	同意辦理
	2. 建議第三章補充國內大數據分析之案例	有關國內大數據分析之案例請參閱 5.2 節。	同意辦理
	3. 3-29 頁資料集之收納狀況建議納入電子票證資料	感謝評審意見，將會納入電子票證資料。	同意辦理
	4. SMART 策略著重在於公部門之方向，可否強化私部門如何參與之作為	本案工作項目係要求研究團隊研擬公部門之計畫方向，未來公部門執行時會透過計畫發包或創意競賽的方式引入私部門人力資源投入交通大數據領域。	同意辦理
	5. 4-26 建議將模擬結果改成優化或調整結果	遵照評審意見辦理，改為調整結果。	同意辦理
	6. 4-25 頁之數學式建議強化說明，包含目標式與限制式，以及旅客數不變之假設	該模式並非最佳化數學模式，故改以文字描述其演算邏輯與前提假設。	同意辦理
	7. 請補充說明站牌班表選定路	遵照審查意見辦理，已於期末報告圖 4-63 (第 4.4.3.1	同意辦理

	線之班距	節) 予以補充。	
	8. 5-16 頁請思考 ETC 電子收費系統放入國內交通數據會流平台之妥適性	ETC 電子收費資料收納於交通部國道高速公路局交通資料庫平台之中，將於報告中補充說明並改用更妥適之用詞。	同意辦理
	9. 請檢視參考文獻格式之合適性	將依照運研所的規定辦理。	同意辦理
淡江大學運輸管理學系張教授勝雄	1. 4-45 有關站牌班表，請問如何權衡旅行時間延長之問題？	已於期末報告第 4.4.3.2 節中加強說明，並修正誤植的部份。	同意辦理
	2. 匯流平臺將來的目的與使用者如何規劃？資料擁有者如何定義(如票證資料是屬於業者或是票證公司?)	匯流平臺將來的目的在於提供政府部門之間的資料共享運用，結合大數據分析技術進行交通管理、公共運輸供給調整、壅塞預警等施政評估，主要使用者為公家機關。本案對於資料擁有者之定義為提供資料給匯流平臺的單位。	同意辦理
	3. 資料去識別化的過程建議再強化	資料去識別化並非本研究之重點，標竿案例分析所使用之資料，在取得之時皆已去識別化，因此本研究不需再進行資料去識別化之工作。	同意辦理
	4. 資料擁有者空有資料未必資料如何應用，建議未來交通部向各主管機關討論可開放至	同意評審看法。	敬悉
交通部管理資訊中心審查意見	1. 有關 SMART 數據戰略中之資料匯流協調機制與平臺規劃內容，可提供本中心未來平臺規畫參考，請主辦單位與中興工程顧問社於期末審查後再次與交通部就本案成果進行討論	遵照評審意見辦理。	同意辦理

	2. 有關 PTX 公共運輸流通平台，管理資訊中心對預售票務資料之收納有意願，預報系統之建置有待未來進行研議	感謝評審的支持。	敬悉
交通部公路總局審查意見	1. 有關共享旅運資訊，建議未來有標準格式	同意評審意見，將納入共享旅運資訊相關計畫的工作項目中。	同意辦理
	2. 希望未來相關風險分析計畫之推動，能回饋資料至三代監理系統	同意評審意見，將納入落實風險評估相關計畫的工作項目中。	同意辦理
交通部航港局審查意見	1. 建議對於旅客數預測之過程加入其他變數	同意評審意見，將納入本研究的後續建議之中。	同意辦理
	2. 請教有關團隊所運用的視覺化分析介面採用內容與原因？	本團隊承襲過去研究之建議，採用 Tableau 做為本案之視覺化分析工具。	同意辦理
交通部台灣鐵路管理局	1. 建議考量臺鐵與客運之轉乘行為分析	本計畫受限於研究經費與時程限制，僅針分析台中地區臺鐵與客運之轉乘行為，詳見報告 4.6.3 節，若臺鐵有此需求，建議未來可進行全面性的分析研究。	同意辦理
	2. 本研究客運班表之視覺化分析未來可提供臺鐵參考	本團隊可提供臺鐵相關視覺化分析之技術支援。	同意辦理
臺北市政府交通局書面審查意見	在蒐集資料上，是否有遇到個資法之狀況？	為了避免違反個資法，本研究所蒐集之資料皆是去識別化後的資料。	同意辦理
本所運計組書面審查意見	1. 4.4.3 節有關車輛行駛時空圖，由於每一條客運路線行駛路徑及停靠站皆不同，建議結合內建之 OpenStreet Map，連結呈現互動式顯示儀錶板，除可提供選線功能、觀察各班次其行駛路徑，益能保留原來時空圖分析之內容。	有關選線功能與觀察各班次行駛路徑之功能，此功能與本研究產製站牌班表之目的不相同，因此將本功能納入未來系統擴充之考量。	同意辦理

	2. P4-42, 修正駕駛行為部分, 由相關資料應可計算出每一條路線及車輛之行駛速率, 請教為何找出行駛速度偏快與慢之駕駛員, 係採用時空圖而非直接用速率直觀變數是否有特別用意?	已於在第 4.4.3.3 節之「駕駛訓練管理建議」內文與表 4.4.2 講述如何尋找偏快和偏慢的駕駛員(行駛時間最小值/最大值佔比最高者), 委員提問之時空圖並非用以尋找極端行為之駕駛員、而僅作示意之用, 為避免日後造成混淆故已於期末報告中予以刪除, 並於文中加強說明。	同意辦理
	3. P4-65 借還車地點之代碼無地理座標資訊一節, 是否已確認資料用有者及其資料庫裡並無地點代碼對照表?	和悠遊卡公司確認過無地點代碼對照表。	同意辦理
	4. P4-68~69, 本報告在無縫轉乘方面之分析相當具有參考性, 惟建議進一步思考分析資料其目的及用意。如轉乘時間應可用來觀察分析各車站在不同時間點之「平均」轉乘時間, 包含在不同的時間點如尖離峰之公車轉臺鐵與臺鐵轉公車之平均所需轉乘時間, 已供研判其無縫無務是否到位?以及該服務水準是否可滿足當下之需求?又抑或是哪些車站及公車路線在哪些時間點可供改進或者調整。	感謝評審肯定, 惟受限於研究時程無法再繼續深入分析, 評審之意見將納入本研究的後續建議之中。	同意辦理
	5. 建議在結論與建議章節就公共運輸領域或其他優先領域, 規納整理並提出較具體且可於未來執行之大數據計畫之相關建議, 包含初步想法、執行策略等具體工作項目。	依評審意見將未來執行大數據計畫之策略規納整理於報告的結論與建議中, 至於其他細部內容請詳見 3.5 節。	同意辦理
運資組書面審查意見	1. 有關 3.2 節公路、鐵路、航空與航港各運輸領域之關注議題與	有關各單位提出之議題可參閱報告。	同意辦理

	需求，請補充說明各項議題之提出單位，以利對應。		
	2. 資料匯流平臺回顧案例請增加交通部管理資訊中心 105 年建置之「交通數據匯流平臺」。	遵照評審意見辦理。	同意辦理
	3. 服務網站規劃，提及台鐵局刻正建置第四代票務系統，請補充說明第三代與第四代系統在各式票務資料欄位之差異。	據聞第四代票務系統目前正開發中，須待建置完成才能得知確切的資料欄位。	同意辦理
主席結論	本案期末報告原則審查通過，請中興工程顧問社團隊配合運資組規劃，針對 SMART 五大應用領域之計畫，在 1 月份另召開說明會，俾利各計畫推動之相關單位瞭解與參考本案規劃成果。請針對今日會議中與會委員及本所運資組意見進行檢討修正，請中興工程顧問社團隊依審查意見研提處理情形答覆意見，送本所審查後，做為修正報告之依據，並請於 105 年 12 月 26 日完成修正定稿之提送。	遵照主席裁示辦理。	同意辦理

# 附錄 E 期末簡報資料

交通部運輸研究所  
Institute of Transportation, MOTC

## 交通大數據分析與應用機制先期規劃

### 期末簡報

財團法人中興工程顧問社

簡報內容

交通大數據分析與應用機制先期規劃

優先領域  
發展規劃

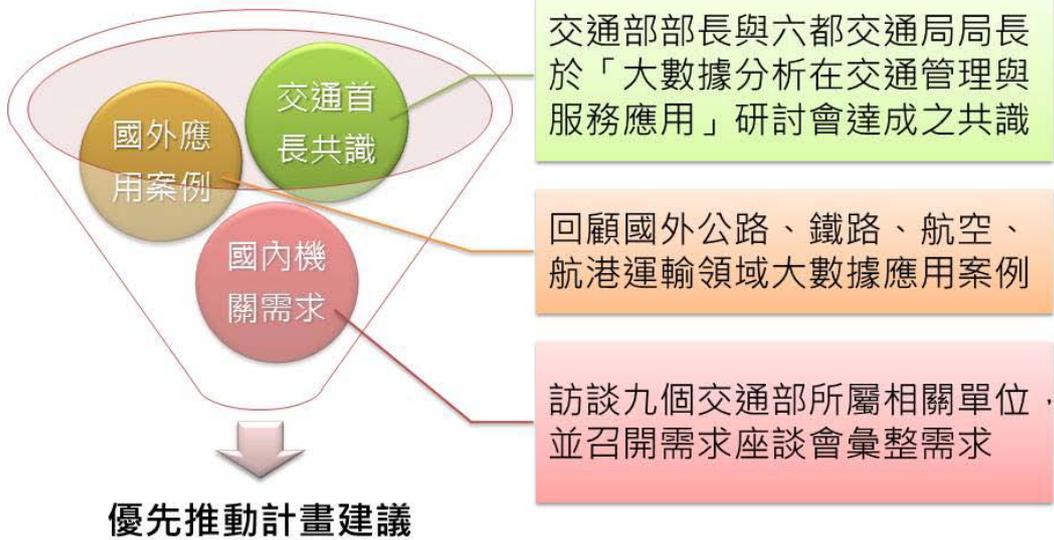
資料匯流  
平台規劃

公共運輸  
案例分析

財團法人中興工程顧問社

# 第一篇 優先領域發展規劃

HIGH  
Priority



## 優先領域發展規劃

3



- 國外大數據應用案例  
– 蒐整公路(12)鐵路(6)航空(6)航港(4)共28例



財團法人中興工程顧問社



# 優先領域發展規劃



## • 國外交通應用案例比較

目的\領域	H公路運輸	R鐵路運輸	F航空運輸	P航港運輸
1蒐集需求	●	○	○	○
2便民資訊	●	●	◎	○
3提升安全	◎	●	●	●
4資產管理	◎	●	●	◎
5延誤改善	◎	●	●	○
6輿情分析	◎	●	◎	○
7節能減碳	◎	◎	◎	●
8追蹤動態	◎	◎	◎	●
9績效提昇	●	◎	●	●

應用重點：●主要 | ◎次之 | ○較低

財團法人中興工程顧問社



# 優先領域發展規劃



## • 國內陸運領域機關訪談



國內航運領域機關訪談



關注議題		關注議題	
蒐集需求	H1找尋潛在旅運需求	蒐集需求	P1整合海關/港口/貿易分析客貨運趨勢
便民資訊	H2提昇公車動態資訊品質	便民資訊	P2港區週邊交通/重車追蹤管理
提升安全	H3駕駛員風險管理	提升安全	P3台灣週邊海域船隻防災/減災分析
節能減碳	H4掌握油耗水準	績效提升	P4船舶港內碼頭運作效率與擁擠密度

關注議題		關注議題	
便民資訊	F1提升聯外運輸服務品質	蒐集需求	R1掌握連續假期旅運需求
資產管理	F2飛機設備妥善率	便民資訊	R2業管資料視覺化呈現
	F3預防性設施維護	便民資訊	R3應用物聯網提供旅客精緻化服務
延誤改善	F4天候對於延誤及班機取消關聯	提升安全	R4提升平交道安全
	F5供給需求匹配與訊息發布	資產管理	R5以預防性維修取代矯正性維修
績效提升	F6航站人流管理	延誤改善	R6改善列車誤點情況
	F7報到櫃檯使用績效	績效提升	R7供給需求匹配
	F8登機門使用績效	績效提升	R8自由座數量配置
資料倉儲	F9建立內部資料匯流	資料倉儲	R9建立內部資料匯流



# 交通部大數據優先領域發展規劃

8



## • 國外應用案例與國內機關需求彙整表

目的\領域	H公路運輸	R鐵路運輸	F航空運輸	P航港運輸
1蒐集需求	● ✓	○	○	○ ✓
2便民資訊	● ✓	● ✓	◎ ✓	○ ✓
3提升安全	◎ ✓	● ✓	●	● ✓
4資產管理	◎	● ✓	● ✓	◎
5延誤改善	◎	● ✓	● ✓	○
6輿情分析	◎	●	◎	○
7節能減碳	◎ ✓	◎	◎	●
8追蹤動態	◎	◎	◎	●
9績效提昇	●	◎ ✓	● ✓	● ✓

應用重點：●主要 | ◎次之 | ○較低

✓：受訪單位關注議題所對應之應用目的

財團法人中興工程顧問社



# 優先領域發展規劃

9



## • 規劃原則



- 1 問題需求導向
- 2 異中求同推動
- 3 資料品質提昇
- 4 開創資料來源
- 5 考量資料現況

財團法人中興工程顧問社



## • SMART發展策略

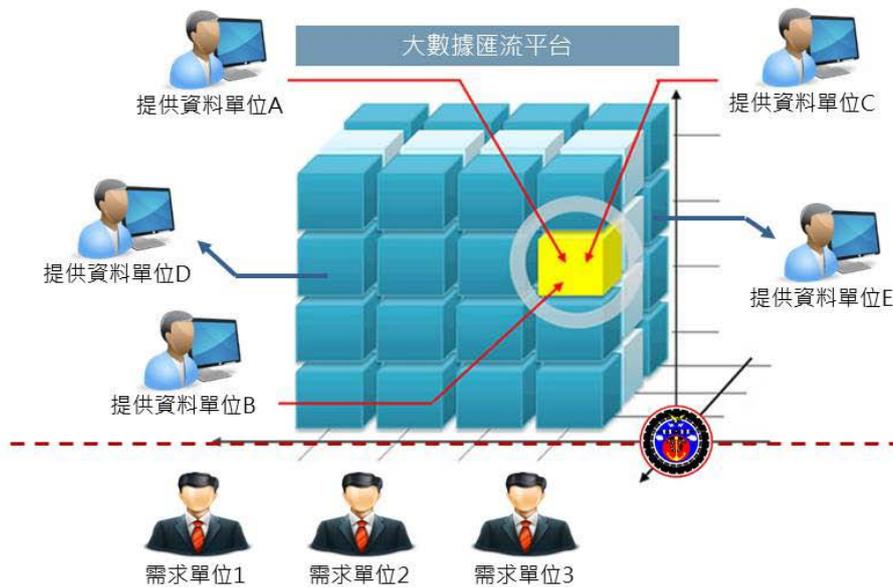


## • 各策略之目標

發展策略	目標方向
制定數據戰略 Strategy	為建構大數據相關系統，需要建立一個健全的資料匯流與共享機制，以奠定交通大數據發展深厚的基石
共享旅運資訊 Message	透過運輸供給與需求之間的即時資料傳遞和共享，將資訊的價值發揮到極致，維持供需平衡的理想狀態。
強化資產管理 Asset	以「開源節流」的角度切入，應用大數據技術協助運輸系統營運單位掌握其資產運用狀況，甚至將閒置資源充分利用創造額外價值
落實風險評估 Risk	結合大數據分析與風險管理，達到提升交通運輸安全防患於未然之目的。
翻轉營運服務 Transformation	改變過去供給導向之運輸系統經營方式，透過大數據分析技術為營運管理帶來新的思維與方法。

財團法人中興工程顧問社

## 第二篇 資料匯流平台規劃



### 資料流通管理辦法

24



- 流通機制清單

#### 資料集新增機制

目標導向

需求導向

自主加入

#### 資料集共享機制

資料集清單與  
資料描述共享

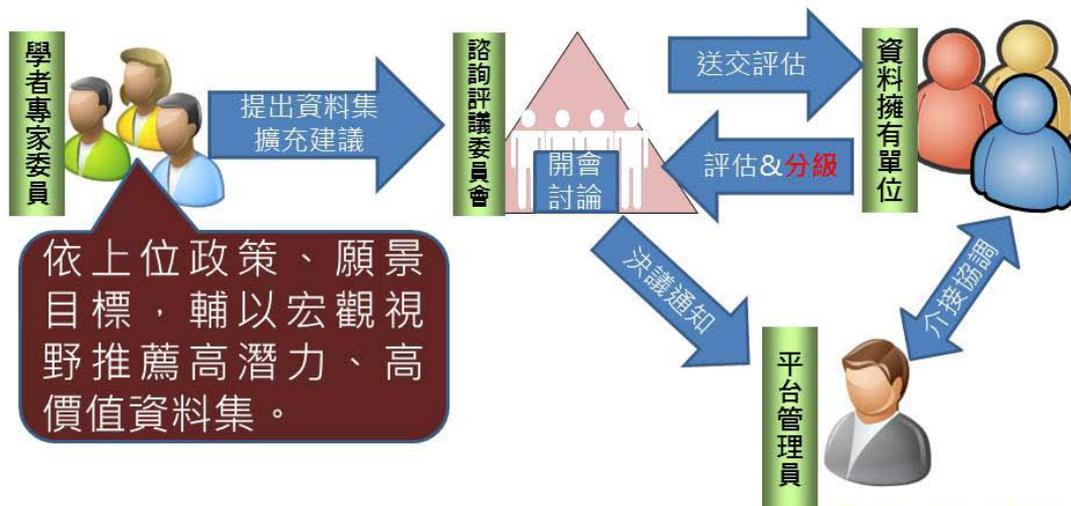
資料集實體  
內容申請



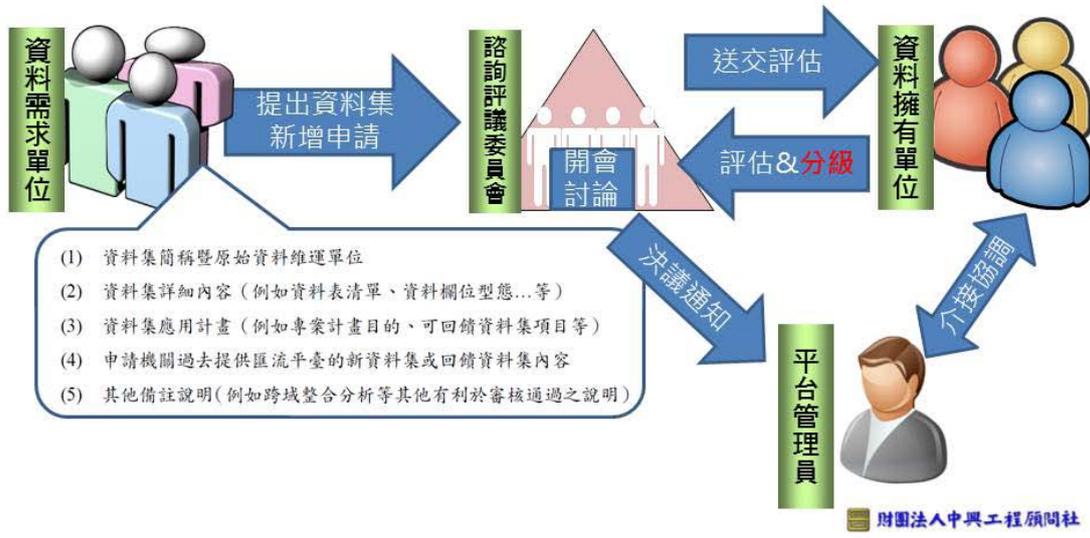
## • 流通角色清單



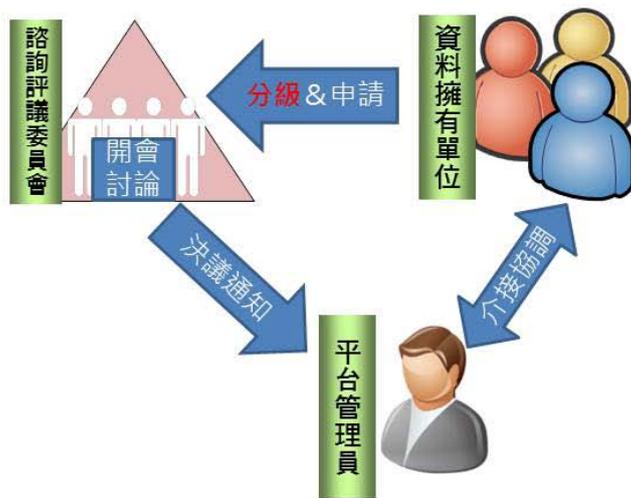
## • 資料集新增機制I-目標導向 - 每年乙次



資料集新增機制II-需求導向  
- 每半年乙次



資料集新增機制II-自主加入





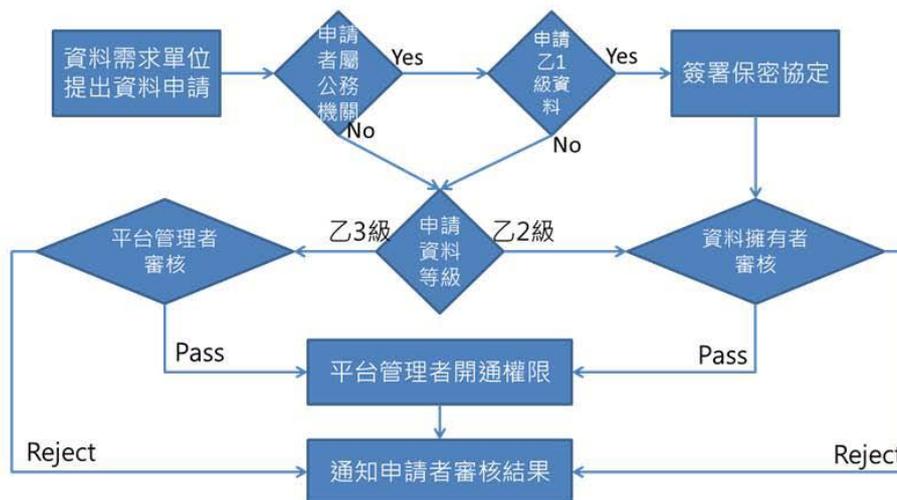
## 資料分級說明表

– 依「行政院及所屬各機關政府資料分類原則」擴充

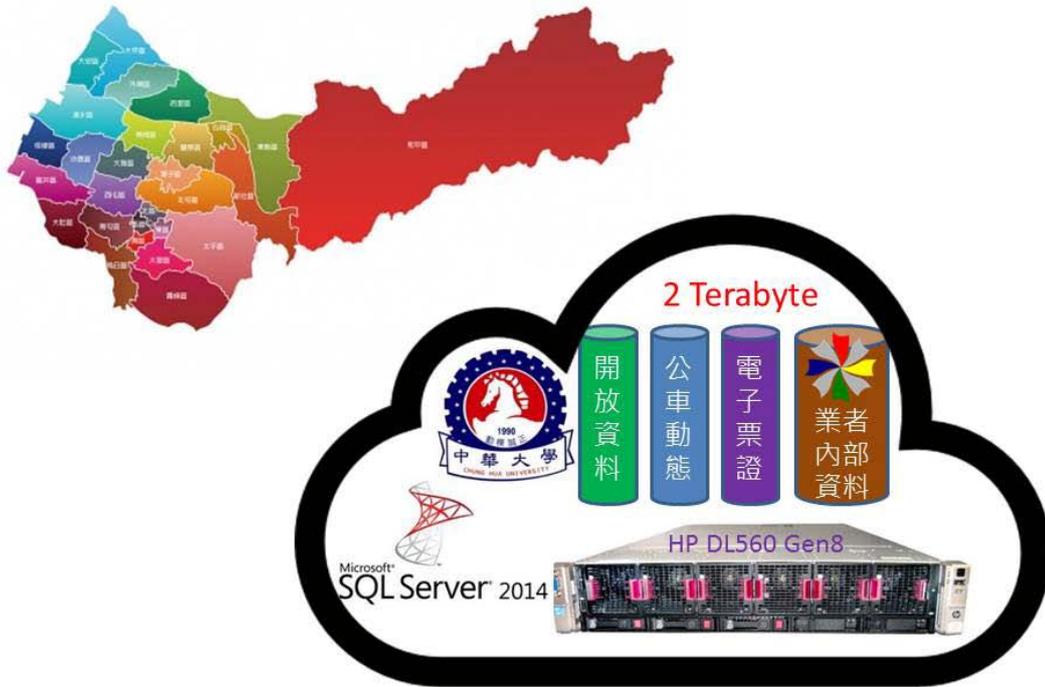
等級代號	甲類	乙1類	乙2類	乙3類	丙類
簡要說明	允許介接 (免申請審核)	允許介接 (限制申請)	允許介接 (擁有者審核)	允許介接 (平台審核)	現階段 不宜介接
允許申請單位	毋需申請	公務機關	公務機關 公司/學校 法人/團體	C級允許單位 + 自然人	公務機關
申請審核單位	毋需審核	資料擁有者 自行審核	資料擁有者 自行審核	授權平台 管理者審核	—
資料集欄位 說明 & 範例	對外公開	僅對公務機 關帳號公開	對外公開	對外公開	—
備註	可考慮推薦 至其他平台 (OpenData)	申請單位須 簽署保密協 定(含罰則)	資料擁有者 可拒絕申請	可考慮推薦 至其他平台 (ex:PTX)	公文往返 專簽處理 去識別化 資料彙總



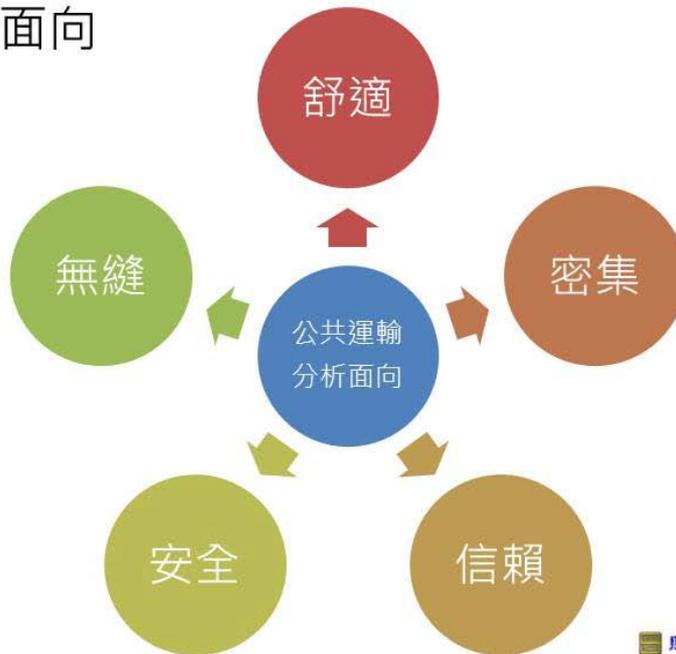
## 資料集實體內容申請流程



# 第三篇 公共運輸分析案例



- 分析面向





# 公共運輸分析案例

- 主要資料 & 分析標的



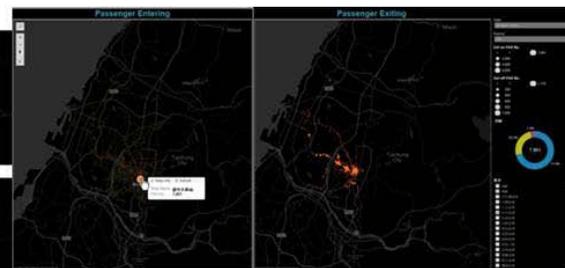
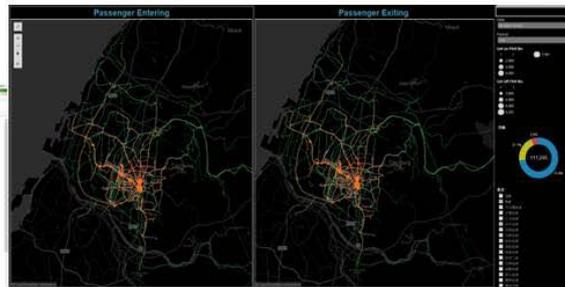
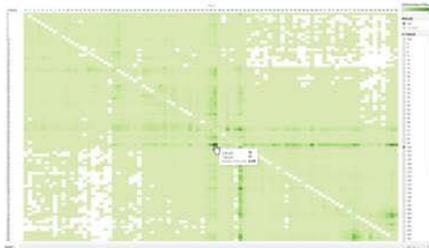
	節點流量	節線流量
分析標的	站牌上/下車人數	車上旅客總人數
應用目的	瞭解旅客起訖實績	掌握乘載服務水準

- 舒適
- 密集
- 信賴
- 安全
- 無縫



# 公共運輸分析案例

- 節點流量

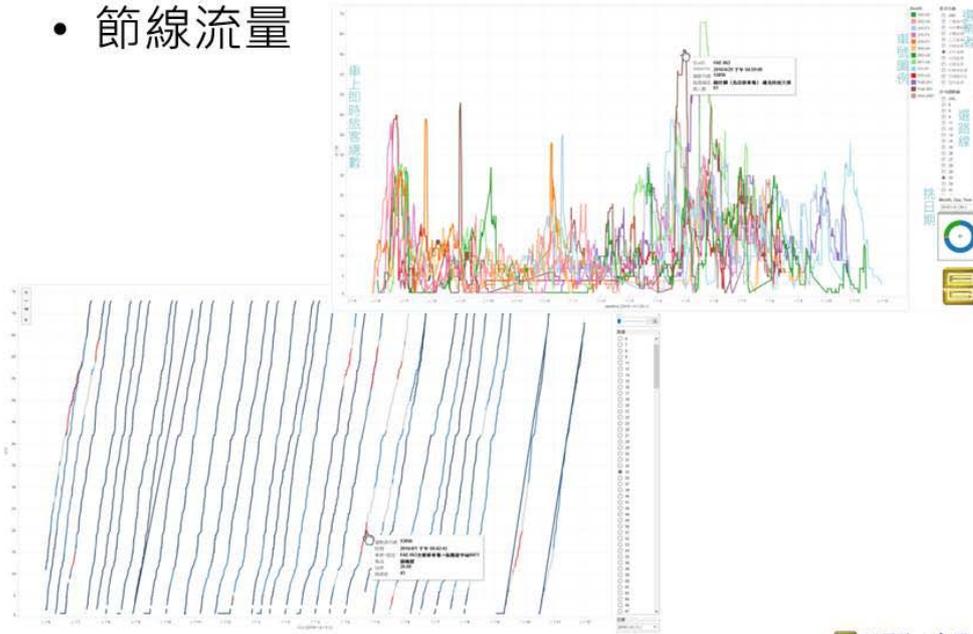


- 舒適
- 密集
- 信賴
- 安全
- 無縫



# 公共運輸分析案例

## • 節線流量

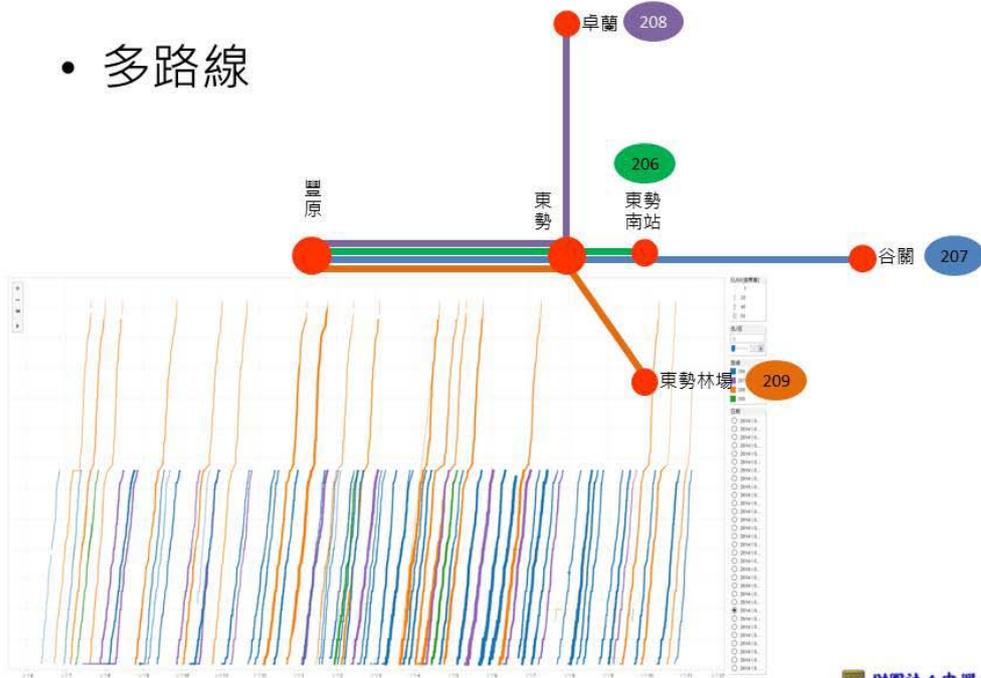


- 舒適
- 密集
- 信賴
- 安全
- 無縫



# 公共運輸分析案例

## • 多路線

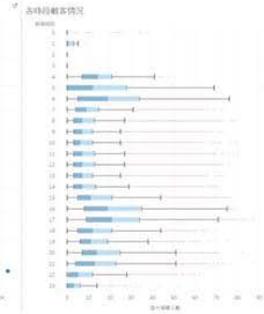
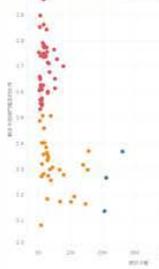
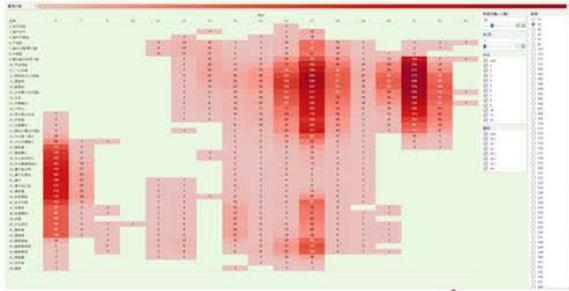


- 舒適
- 密集
- 信賴
- 安全
- 無縫



# 公共運輸分析案例

## • 區間車時空評估工具



財團法人中興工程顧問社

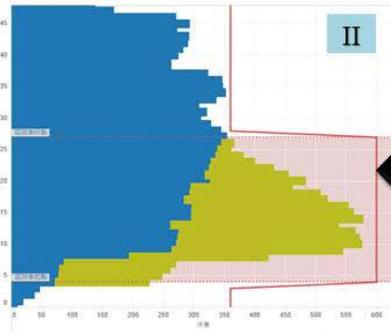
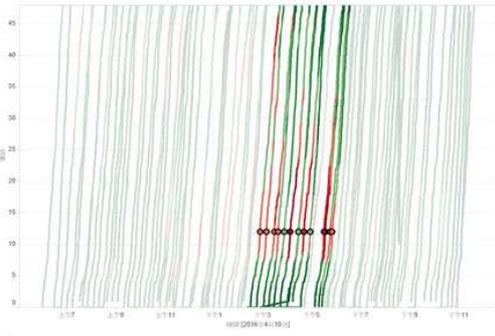


- 舒適
- 密集
- 信賴
- 安全
- 無縫



# 公共運輸分析案例

## • 轉乘衝擊評估輔助工具

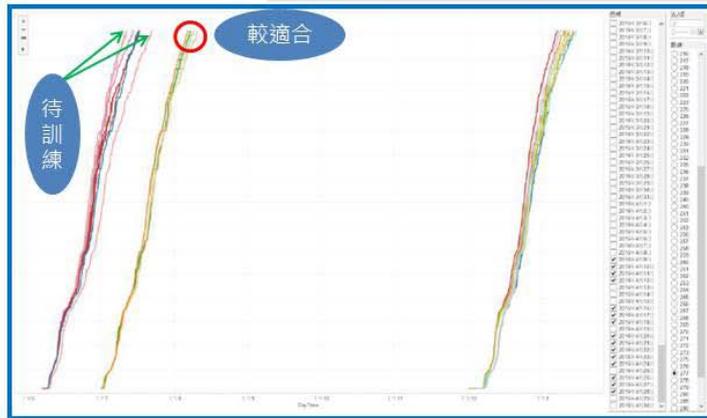


財團法人中興工程顧問社



- 舒適
- 密集
- 信賴
- 安全
- 無縫

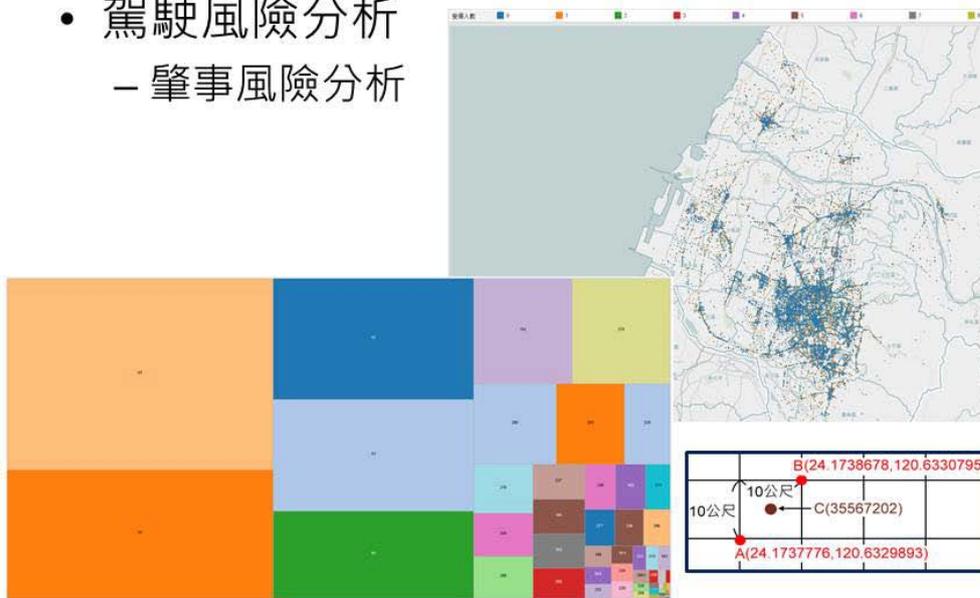
站牌班表



舒適  
密集  
信賴  
安全  
無縫

財團法人中興工程顧問社

駕駛風險分析  
- 肇事風險分析



舒適  
密集  
信賴  
安全  
無縫

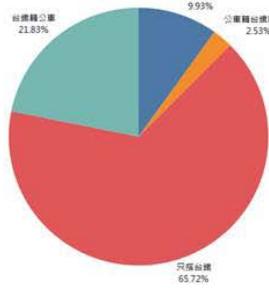
財團法人中興工程顧問社

# 公共運輸分析案例

- 分析對象
  - 臺中市公車(市區客運&公路客運)
  - 臺灣鐵路管理局 (海線&山線)



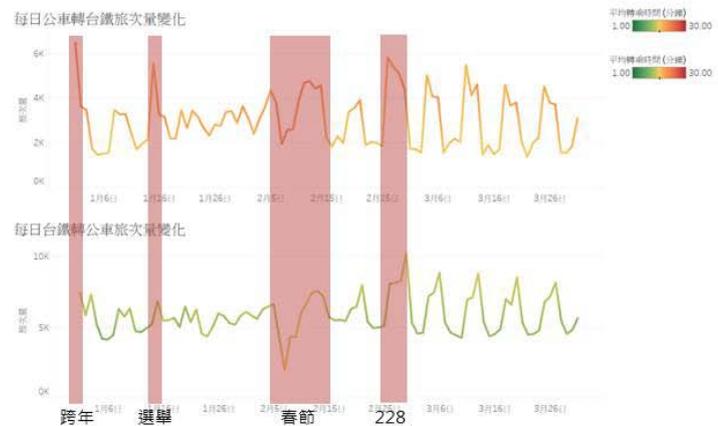
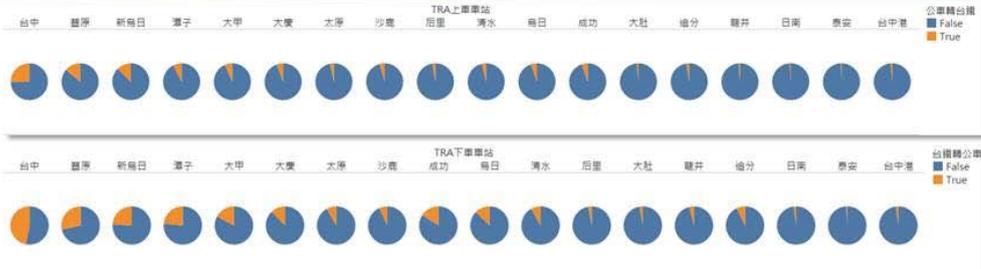
舒適  
密集  
信賴  
安全  
無縫



線路名稱	站名	1	2	3	4	5	6	7	8	9	10	11	12
台中線	台中	8,802	8,400	19,887	18,243	8,228	8,713	8,264	8,246	8,881	8,428	8,218	8,275
	成功	9,600	8,827	8,288	16,875	9,433	8,563	7,941	8,214	8,163	9,318	8,556	8,368
海線	台中	78,718	90,962	78,652	188,338	71,452	73,791	69,737	73,734	68,937	73,548	71,753	68,458
	成功	77,878	85,716	74,362	185,277	71,488	67,726	68,832	71,815	67,111	72,825	70,829	64,225
山線	台中	2,814	2,587	2,205	5,596	2,104	1,981	1,875	1,735	1,800	1,902	1,802	1,802
	成功	2,140	2,258	2,274	5,746	2,238	2,047	1,830	1,730	1,800	1,802	1,802	1,802

財團法人中興工程顧問社

# 公共運輸分析案例

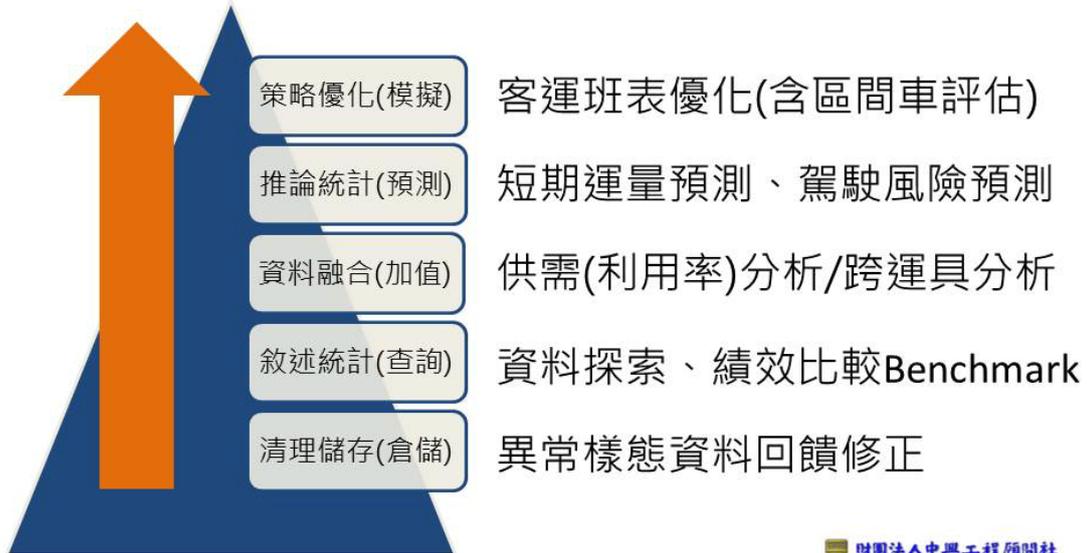


財團法人中興工程顧問社

舒適  
密集  
信賴  
安全  
無縫



- 應用潛力彙整



財團法人中興工程顧問社

# 簡報結束 敬請指教



交通大數據分析與應用機制先期規劃