

105-088-5438
MOTC-IOT-104-IEB017

公車動態資訊系統巨量資料 (Big Data) 蒐集與視覺化分析研究

著者：陳其華、吳東凌、陳翔捷、邱裕鈞、蘇昭銘、陳穆臻
張恩輔、黃笙玗、孫千山、張志鴻、謝志偉

交通部運輸研究所

中華民國 105 年 10 月

公車動態資訊系統巨量資料 (Big Data)

蒐集與視覺化分析研究

/ 陳其華等著. -- 初版. -- 臺北市 :

交通部運研所, 民105.10

面 ; 公分

ISBN 978-986-04-9941-4 (平裝)

公車動態資訊系統巨資料 (Big Data) 蒐集與視覺化分析研究

著 者：陳其華、吳東凌、陳翔捷、王晉元、邱裕鈞、蘇昭銘、陳穆臻、
張恩輔、黃笙玔、孫千山、張志鴻、謝志偉

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 (中文版>圖書服務>本所出版品)

電 話：(02)23496789

出版年月：中華民國 105 年 10 月

印 刷 者：九易數碼科技印刷有限公司

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：320 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：1010501745

ISBN：978-986-04-9941-4

(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/共同研究計畫出版品摘要表

出版品名稱：公車動態資訊系統巨量資料 (Big Data) 蒐集與視覺化分析研究			
國際標準書號 (或叢刊號) ISBN 978-986-04-8289-8 (平裝)	政府出版品統一編號 1010501745	運輸研究所出版品編號 105-088-5438	計畫編號 MOTC-IOT-104-IEB017
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：國立交通大學 計畫主持人：王晉元 研究人員：邱裕鈞、蘇昭銘、陳穆臻、張恩輔、黃笙玗、孫千山、張志鴻、謝志偉 地址：新竹市大學路 1001 號綜合一館 8 樓 聯絡電話：(03)5712121		研究期間 自 104 年 4 月 至 104 年 10 月
關鍵詞：公車動態資訊系統、巨量資料視覺化分析、資料探勘			
摘要： 近年來我國各地陸續完成市區汽車客運及公路客運動態資訊系統之建置作業，已累積相當程度之公車動態資訊系統(APTS)資料庫，將APTS資料庫所記錄之各項巨量資料，結合路線、班表、司機等資料進行巨量資料分析，可發現隱藏在各資料背後的重要訊息，並轉換為可應用於監督與管理之有效資訊。本計畫首先透過訪談中央與主要縣市政府汽車客運主管機關，彙整汽車客運管理作業上應用巨量資料分析之實務需求，並蒐集國內外公共運輸巨量資料應用之相關研究，再依據主管機關實務需求彙整結果，規劃本計畫巨量資料之分析與應用範圍。 在本計畫中，以大臺中地區為研究範圍，主要蒐集與視覺化分析以下巨量資料： 1. 公路客運或市區公車動態資訊系統資料庫(含靜態與動態資料) 2. 標竿示範客運業者自主管理資料庫(如駕駛員出勤、績效考核、肇事違規與油耗等資料) 3. 電子票證資料 運用本計畫所發展出之資料視覺化分析程序與儀表板，可協助公共運輸主管機關與公車業者分析客運之營運績效，本計畫亦透過實例闡述如何應用所開發之儀表板，以快速提供決策者更多管理實務上之資訊，減少過去仰賴經驗判斷之情況，未來主管機關欲運用視覺化分析掌握整體公共運輸服務狀況時，可參考本計畫之巨量資料蒐集與分析經驗與成果。			
出版日期	頁數	定價	本出版品取得方式
105 年 10 月	269	320	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。
機密等級： ☐ 密 ☐ 機密 ☐ 極機密 ☐ 絕對機密 (解密條件： ☐ 年 ☐ 月 ☐ 日解密， ☐ 公布後解密， ☐ 附件抽存後解密， ☐ 工作完成或會議終了時解密， ☐ 另行檢討後辦理解密) ☒ 普通			
備註：1.本研究之結論與建議不代表交通部之意見。 2.本研究係使用交通部公路總局經費辦理。			

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

TITLE: Big Data Visualization Analysis of Public Transit Data			
ISBN(OR ISSN) 978-986-04-8289-8	GOVERNMENT PUBLICATIONS NUMBER 1010501745	IOT SERIAL NUMBER 105-088-5438	PROJECT NUMBER MOTC-IOT-104-IEB017
DIVISION: Information System Division DIVISION DIRECTOR: Ci-Hwa Chen PRINCIPAL INVESTIGATOR: Chi-Hwa Chen PROJECT STAFF: Dong-Ling Wu, Siang-Jie Chen PHONE: 02-23496886 FAX: 02-25450426			PROJECT PERIOD FROM April 2015 TO October 2015
RESEARCH AGENCY: National Chiao Tung University PRINCIPAL INVESTIGATOR: Jin-Yuan Wang PROJECT STAFF: Yu-Jyun Ciou, Jhao-Ming Su, Mu-Jhen Chen, En-Fu Jhang, Sheng-Syuan Huang, Cian-Shan Sun, Jhih-Hong Jhang, Jhih-Wei Sie, ADDRESS: Rm. A803, Assembly Building 1, No.1001, Daxue Rd., East Dist., Hsinchu City 30010, Taiwan(R.O.C.) PHONE: 03-35716440			
KEY WORDS: Advanced Public Transit System , Big Data, Visualization Analysis			
ABSTRACT: Many cities in Taiwan recently have implemented the automatic vehicle location (AVL) and smart card system for bus fleets. Analyzing Big Data generated by both systems, along with route data, driver data and schedule data, will help the transit agencies to monitor and manage public transit. In this project, we interviewed transit agencies of central and local government to summarize their requirements of Big Data to facilitate the management of bus fleets. Other than realizing the demand of transit agencies, this project also studies previous relevant research, in order to identify the boundaries of this research . In this project, we focus on mining and visualizing the following three Big Data: 1. Automatic vehicle location data from intercity and city bus's advanced public transportation system (including steady and dynamic data) 2. Management data from bus companies.(including driver's performance, vehicle's oil consumption) 3. Smart card data We used data fusion technique to develop several dashboards that help transit agencies analyse the service levels of their bus fleets. This was done for the transit system in Taichung city, located in central Taiwan. Several practical examples are presented in this paper, which could offer precise information for decision makers of transit agencies, instead of relying on professional judgement. With user-friendly dashboards, our experience of using visualization to monitor the overall transit performance can easily be applied to other cities.			
DATE OF PUBLICATION 2016.10.	NUMBER OF PAGES 278	PRICE 320 NT dollars	CLASSIFICATION ☐ RESTRICTED ☐ CONFIDENTIAL ☐ SECRET ☐ TOP SECRET ☒ UNCLASSIFIED
1. The views expressed in this publication are not necessarily those of the Ministry of Transportation and Communications. 2. The budget of this research project is contributed by Directorate General of Highways, MOTC			

目錄

中文摘要.....	I
英文摘要.....	II
目 錄.....	III
圖 目 錄.....	VII
表 目 錄.....	XII
第一章 計畫概述.....	1-1
1.1 計畫背景.....	1-1
1.2 計畫目的.....	1-2
1.3 分析範圍.....	1-2
1.3.1 研究區域評估.....	1-2
1.3.2 研究對象評估.....	1-4
1.4 工作項目.....	1-5
第二章 文獻回顧.....	2-1
2.1 公車動態資訊系統發展.....	2-1
2.2 巨量資料與相關分析技術.....	2-3
2.2.1 巨量資料定義、特性與趨勢.....	2-3
2.2.2 巨量資料分析技術.....	2-6
2.3 公共運輸巨量資料分析案例.....	2-12
2.4 可視化應用工具及案例回顧.....	2-22
2.4.1 應用工具及技術.....	2-22
2.4.2 相關案例.....	2-26
第三章 需求彙整.....	3-1
3.1 主管機關需求座談辦理情形.....	3-1
3.2 客運業者需求座談辦理情形.....	3-4
3.3 需求彙整結果.....	3-7
第四章 系統架構規劃與工具.....	4-1
4.1 公共運輸巨量資料分析系統.....	4-1
4.2 相關分析工具.....	4-2
4.2.1 Tableau 簡介.....	4-2

4.2.2	R 語言簡介	4-5
4.2.3	IBM Watson 簡介	4-7
第五章	資料介接與處理	5-1
5.1	資料介接清單	5-1
5.2	電子票證資料處理與課題	5-3
5.3	公車動態資料處理與課題	5-6
5.3.1	定時資料	5-6
5.3.2	定點資料	5-7
5.3.3	資料品質與對策	5-8
5.4	數位式行車記錄器資料處理與課題	5-9
5.5	營運相關資料處理與課題	5-11
5.5.1	旅次需求資料	5-12
5.5.2	路線及班次資料	5-13
5.5.3	行政區 OD 運量資料	5-14
5.5.4	車型資料	5-15
5.5.5	油料資料	5-15
5.5.6	機務資料	5-17
5.6	其他運輸相關資料處理與課題（事故、計程車資料）	5-19
5.7	其他非運輸相關資料處理與課題（氣候）	5-19
第六章	探勘分析與視覺化呈現	6-1
6.1	電子票證分析	6-1
6.1.1	探勘分析目的	6-1
6.1.2	演算處理流程	6-4
6.1.3	視覺呈現暨應用	6-5
6.2	路線與班次重疊分析	6-19
6.2.1	探勘分析目的	6-19
6.2.2	演算處理流程	6-19
6.2.3	分析結果與視覺呈現	6-20
6.3	供需分析	6-23
6.3.1	探勘分析目的	6-23

6.3.2	演算處理流程	6-24
6.3.3	視覺呈現暨應用	6-26
6.4	結合電子票証與公車動態之分析	6-36
6.5	行車時間與班表分析	6-39
6.5.1	探勘分析目的	6-39
6.5.2	演算處理流程	6-40
6.5.3	視覺呈現暨應用	6-42
6.6	站牌班表產製分析	6-54
6.6.1	探勘分析目的	6-54
6.6.2	演算處理流程	6-55
6.6.3	站牌班表正確性分析	6-60
6.7	車輛相關分析(含機務、油料)	6-64
6.7.1	探勘分析目的	6-65
6.7.2	演算處理流程	6-65
6.7.3	分析結果與視覺呈現	6-72
6.8	駕駛員風險分析	6-86
6.8.1	探勘分析目的	6-86
6.8.2	演算處理流程	6-86
6.8.3	視覺呈現暨應用	6-87
6.9	車機穩定度分析	6-92
6.9.1	探勘分析目的	6-92
6.9.2	演算處理流程	6-92
6.9.3	視覺呈現暨應用	6-92
第七章	結論與建議	7-1
7.1	結論	7-1
7.2	建議	7-2
附錄 A	需求座談會議資料	A-1
附錄 B	教育訓練辦理情形	B-1
附錄 C	資料庫欄位說明	C-1
附錄 D	資料表欄位範例	D-1

附錄 E 參考文獻	E-1
附錄 F 期中報告審查意見回覆表	F-1
附錄 G 期末報告審查意見回覆表.....	G-197

圖目錄

圖 1-1 分析區域評估流程	1-4
圖 2-1 聰明公車系統架構圖	2-3
圖 2-2TransDec 應用	2-6
圖 2-3 資料探勘流程圖	2-10
圖 2-4 貓頭鷹深夜公車 N37 與公車電光板（中間）和深夜公車標誌 ...	2-12
圖 2-5 和欣客運行控中心	2-14
圖 2-6 大眾運輸資料分析架構	2-14
圖 2-7 鎮江公車停靠站	2-15
圖 2-8 英國航空改善使用者體驗	2-15
圖 2-9 旅運行程最佳化情境說明	2-16
圖 2-10 旅運資訊面板	2-17
圖 2-11 臺南市即時交通資訊網	2-18
圖 2-12 臺南市公車到站時刻管理系統之系統畫面	2-18
圖 2-13 MetroViz 系統示意圖	2-19
圖 2-14 公共運輸延誤視覺化之應用	2-20
圖 2-15 土耳其伊斯坦堡 BRT 各站運量及方向	2-20
圖 2-16 公車空間間距（S）與時間間距（h）視覺化示意圖 ...	2-21
圖 2-17 3D 公車間距視覺化分析軟體	2-21
圖 2-18 Google Maps API 類型圖	2-22
圖 2-19 Google Maps 標記叢集顯示	2-23
圖 2-20 Google Maps 熱力圖顯示	2-23
圖 2-21 D3 可視化範例	2-24
圖 2-22 D3 及 Google Maps API 顯示臺北市交通事故可視化 ..	2-24
圖 2-23 SAS Visual Analytics 顯示臺北市學生卡熱點可視化畫面	2-25
圖 2-24 U Bike App 使用介面示意圖	2-27
圖 2-25 Hubway 公共自行車地圖	2-27
圖 2-26 Hubway 公共自行車視覺分析應用之單車速率圖	2-28
圖 2-27 德國 DB Zugradar 軌道系統動態移動與時刻表地圖	2-28
圖 2-28 德國 DB Zugradar 手機電子票證與相關服務畫面	2-29
圖 2-29 德國 DB Zugradar 共乘與公共自行車租賃操作畫面 ...	2-29

圖 2-30 TRAVIC 平臺所呈現之澳洲墨爾本公共運輸營陸路網圖	2-30
圖 2-31 TRAVIC 平臺所呈現之瑞士蘇黎世全區之公車路網圖	2-31
圖 2-32 TRAVIC 呈現美國紐約公共運輸營陸路網與對應班次圖	2-31
圖 2-33 TRANSIT QUALITY AND EQUIT 舊金山地區轉乘績效圖	2-32
圖 3-1 主管機關需求座談會實況剪影	3-3
圖 4-1 巨量資料視覺化分析系統整合平臺架構示意圖	4-1
圖 4-2 Tableau 在巨量資料領域之定位	4-3
圖 4-3 Tableau 在 BI 領域的魔力象限（Magic Quadrant）定位 ..	4-3
圖 4-4 Tableau 支援之 Data Source 清單圖	4-4
圖 4-5 Tableau 視覺化呈現軟體支援之圖表列舉	4-5
圖 4-6 IBM Watson 介面	4-7
圖 5-1 公車動態定時資料欄位圖	5-6
圖 5-2 數位行車記錄器原始檔案名稱清單	5-10
圖 5-3 數位行車記錄器資料表欄位圖	5-10
圖 5-4 歷史資料日氣象資料收集來源示意圖	5-19
圖 5-5 氣象歷史資料收集來源	5-20
圖 5-6 氣象資料擷取欄位清單	5-21
圖 5-7 擷取/清理後之歷史氣象資料表	5-21
圖 6-1 電子票證分析指標與應用領域對應架構圖	6-3
圖 6-2 行政區運量統計圖	6-6
圖 6-3 行政區票卡別運量統計圖	6-7
圖 6-4 行政區間旅次需求統計圖	6-7
圖 6-5 豐原區分析旅次需求統計圖	6-8
圖 6-6 行政區間旅次需求示意圖	6-8
圖 6-7 路廊旅次示意圖	6-9
圖 6-8 整合地圖路廊旅次示意圖	6-10
圖 6-9 路線票種分析示意圖	6-11
圖 6-10 路線站牌運量分析示意圖	6-11
圖 6-11 站牌上下車人數分析示意圖	6-12
圖 6-12 使用者需求量分析示意圖	6-13
圖 6-13 前 100 位使用者需求量分析示意圖	6-13

圖 6-14 前單一使用者旅次分析示意圖	6-14
圖 6-15 路線 A 去回程各站牌每日平均車上人數(四月)散佈圖	6-15
圖 6-16 路線 A 去回程各站牌上下車人數與路線平均上下車人數之比	6-15
圖 6-17 計程車乘車上下車分佈圖	6-16
圖 6-18 上午時段豐客-東勢-谷關路廊乘車人數	6-17
圖 6-19 下午時段豐客-東勢-谷關路廊乘車人數	6-18
圖 6-20 晚上時段豐客-東勢-谷關路廊乘車人數	6-18
圖 6-21 路線與班次重疊分析流程圖	6-20
圖 6-22 小時次數累加示意圖	6-21
圖 6-23 全天次數累加示意圖	6-22
圖 6-24 供需分析演算處理流程	6-24
圖 6-25 各區全日供給與需求比例視覺呈現(以班次數為例)	6-27
圖 6-26 各區分時段供給與需求比例視覺呈現(中區)	6-27
圖 6-27 各區全日運量與需求比例視覺呈現	6-28
圖 6-28 各區分時段運量與需求比例視覺呈現(中區)	6-29
圖 6-29 OD 運量視覺呈現示意圖(總運量).....	6-30
圖 6-30 OD 運量視覺呈現示意圖(學生票).....	6-31
圖 6-31 OD 運量視覺呈現示意圖(敬老愛心票).....	6-32
圖 6-32 各區供給與預測需求比例視覺呈現	6-33
圖 6-33 各區分時段供給與預測需求比例視覺呈現(中區)	6-33
圖 6-34 大臺中地區旅次長度分布	6-34
圖 6-35 各行政區旅次長度	6-34
圖 6-36 低地板公車分析架構	6-35
圖 6-37 低地板公車搭乘比例	6-35
圖 6-38 結合電子票証與公車動態之視覺化介面	6-36
圖 6-39 行駛軌跡與載客量	6-37
圖 6-40 行駛路線與車輛位置	6-38
圖 6-41 行駛班次之基本資訊	6-38
圖 6-42 豐客 55 路公車一心市場到北新國車載客情形	6-39
圖 6-43 臺中地區公車站牌測站參照分類圖	6-41
圖 6-44 各站牌通過時間泡泡圖(以 55 路為例)	6-43

圖 6-45 各站牌通過時間樹圖(以 51 路為例)	6-44
圖 6-46 樹圖 ToolTip 功能示意圖	6-45
圖 6-47 站牌通過時間與時段及晴雨多維度比較圖	6-45
圖 6-48 各時段週間站牌通過時間暨晴雨多維度比較圖	6-46
圖 6-49 行車時間與溫度關係 (以 55 路為例)	6-46
圖 6-50 以 R 語言區域回歸取代圖 6 49 之分析	6-47
圖 6-51 預估到站時間準確性-整體水準	6-48
圖 6-52 預估到站時間準確性-時段因素分析	6-48
圖 6-53 預估到站時間準確性-站間數因素分析	6-49
圖 6-54 高密度班表視覺化呈現圖 (以 55 路為例)	6-49
圖 6-55 低密度班表視覺化呈現圖 (以 92 路為例)	6-50
圖 6-56 全日班表檢視範圍調整工具列	6-50
圖 6-57 班表連班/班車追越樣態視覺化呈現圖	6-51
圖 6-58 連班路線分析圖	6-51
圖 6-59 連班分佈分析 (時間角度)	6-52
圖 6-60 連班分佈分析 (空間角度)	6-52
圖 6-61 跨日班表比較視覺化呈現圖 (以週間為例)	6-53
圖 6-62 跨日班表比較視覺化呈現圖 (以週末為例)	6-53
圖 6-63 供給時空圖暨乘載率整合儀表板	6-54
圖 6-64 離峰下緣線示意圖	6-56
圖 6-65 尖峰下緣線示意圖	6-56
圖 6-66 離峰下緣線與各權重示意圖	6-57
圖 6-67 尖峰下緣線與各權重示意圖	6-58
圖 6-68 離峰最佳目標值結果	6-60
圖 6-69 尖峰最佳目標值結果	6-60
圖 6-70 油耗分析流程圖	6-66
圖 6-71 車輛風險分析流程圖	6-70
圖 6-72 車號油耗量圖	6-73
圖 6-73 各噸位數車輛數	6-74
圖 6-74 噸位數耗油量圖	6-74
圖 6-75 廠牌車輛數	6-75

圖 6-76 廠牌耗油數	6-76
圖 6-77 各出廠年份車輛	6-76
圖 6-78 車齡耗油圖	6-77
圖 6-79 各廠牌車輛載重數總計	6-78
圖 6-80 噸數廠牌耗油交叉分析	6-78
圖 6-81 各廠牌出廠年分車輛數	6-79
圖 6-82 車輛廠牌出廠年份之交叉分析	6-80
圖 6-83 車號車輛故障圖	6-80
圖 6-84 各廠牌車輛數與發生故障車輛數	6-82
圖 6-85 廠牌車輛故障	6-82
圖 6-86 年份廠牌車輛故障	6-83
圖 6-87 車齡/油耗/故障交叉分析圖	6-83
圖 6-88 高速駕駛空間分佈視覺化呈現	6-88
圖 6-89 由圖 6 88 右側儀表板篩選欲檢視業者	6-88
圖 6-90 高速駕駛空間分佈視覺化呈現	6-89
圖 6-91 高轉速駕駛空間分佈視覺化呈現示意圖	6-89
圖 6-92 車速與引擎轉速視覺化分析圖	6-90
圖 6-93 車速與引擎轉速勾稽路線視覺化呈現圖	6-91
圖 6-94 選取圖 6 93 速度與轉速資料即時顯示路線互動圖	6-91
圖 6-95 資料傳輸延遲發生數較高之車牌與路線圖	6-93
圖 6-96 車機延遲全年逐日分析圖	6-93
圖 6-97 車機延遲各路線逐月去/返多維度比較圖	6-94
圖 6-98 車機延遲各車輛逐月多維度比較圖	6-94
圖 6-99 車機資料傳輸延遲分析儀表板示意圖	6-95
圖 6-100 各傳輸延遲樣本百分率比較圖	6-97
圖 6-101 各傳輸漏失樣本百分率比較圖	6-98
圖 6-102 各傳輸延遲樣本累積比率比較圖	6-98
圖 6-103 各傳輸漏失樣本累積比率比較圖	6-99
圖 6-104 傳輸延遲影響因素比較分析圖	6-100
圖 6-105 指定路線比較各車機可靠度之儀表板	6-100
圖 6-106 預防檢修應用儀表板設計示意圖	6-101

表目錄

表 1.1 臺中市客運業者.....	1-5
表 2.1 目前的資料處理的數量	2-6
表 2.2 資料探勘平臺架構.....	2-8
表 3.1 主管機關需求座談會議程表	3-1
表 3.2 主管機關邀請清單.....	3-2
表 3.3 主管機關需求座談會簽到表	3-2
表 3.4 客運業者需求座談會議程表	3-5
表 3.5 客運業者邀請清單.....	3-5
表 3.6 客運業者需求座談會簽到表	3-6
表 3.7 主管機關需求彙整表	3-7
表 3.8 客運業者需求彙整表	3-8
表 3.9 主管機關需求處理方式彙整表	3-8
表 3.10 客運業者需求處理方式彙整表	3-9
表 5.1 蒐集之巨量資料彙總表	5-2
表 5.2 公車動態定時資料欄位說明表	5-7
表 5.3 公車動態定點資料欄位說明表	5-8
表 5.4 臺中市 104 年各行政區人口數	5-12
表 5.5 臺中市各時段旅次出發時間比例	5-13
表 5.6 汽車客運供給欄位說明	5-13
表 5.7 電子票證欄位說明.....	5-14
表 5.8 車輛清冊表.....	5-15
表 5.9 車輛零件統計表.....	5-16
表 5.10 電子票證資料表.....	5-16
表 5.11 車輛維修資料表.....	5-18
表 6.1 電子票證分析用欄位表	6-4
表 6.2 路線與班次重疊表.....	6-21
表 6.3 路段累積班次統計表	6-22
表 6.4 各車型搭乘比例 (豐原客運為例).....	6-36
表 6.5 各站到達時間與預估值之平均誤差及標準差表	6-61
表 6.6 周間/周末情境分析	6-61

表 6.7 周間/周末情境分析差異檢定	6-62
表 6.8 平時/春節假期情境分析	6-62
表 6.9 平時/春節假期情境分析差異檢定	6-63
表 6.10 晨峰/昏峰情境分析	6-63
表 6.11 晨峰/昏峰情境分析差異檢定	6-64
表 6.12 車輛零件統計表.....	6-67
表 6.13 車輛清冊表.....	6-68
表 6.14 電子票證資料表.....	6-68
表 6.15 車輛維修資料表.....	6-71
表 6.16 車輛特性欄位一覽表	6-72
表 6.17 車號油耗量.....	6-73
表 6.18 車號 685-FX 對應之駕駛編號表	6-73
表 6.19 噸位數耗油量表.....	6-75
表 6.20 廠牌耗油量表.....	6-76
表 6.21 車齡耗油量表.....	6-77
表 6.22 車號 536FH 對應駕駛編號表	6-81
表 6.23 檢修/故障平均公里數	6-85
表 6.24 TTIA 危險駕駛相關事件表	6-86
表 6.25 定時資料傳輸延誤比例比較表	6-96
表 6.26 定點資料傳輸延誤比例比較表	6-96
表 6.27 各月份定點資料傳輸延誤比較表	6-96
表 6.28 各時段定點資料傳輸延誤比較	6-97

第一章 計畫概述

1.1 計畫背景

先進公共運輸系統（APTS）近幾年在國內蓬勃發展，各地均紛紛建置相關系統，除了提供乘客公車動態資訊、預估車輛到站時間以及公車路線查詢之外，客運業者亦可透過追蹤車輛動態訊息以調整營運調度績效。經過這些年的努力，各地的系統已經漸趨成熟與穩定，民眾逐漸習慣使用公車動態系統所提供的資訊，並將其融入在日常生活中。而公共運輸主管機關與業者也開始思考如何利用公車動態資訊系統所提供的大量資料，產生具有管理意涵的資訊，使得公司的營運能夠更有績效。

然而在上述 APTS 中心所蒐集的龐大公車行車相關資料，除了上述應用外，卻鮮少進行更進一步的資料探勘（Data mining）與分析。這些龐大的資料，其實包含了許多有用的資訊，諸如行駛時間、駕駛員特性、車輛特性、公共運輸旅運特性等，如果能夠充分利用這些資料，整理出有用的資訊，應該可以協助業者進行有效的營運管理，更進一步提昇公共運輸的服務水準。對於政府單位而言，利用這些資料，也可以分析目前公共運輸的供需情形，作為整體公共運輸規劃的依據。

近年來，由於巨量資料分析技術日趨成熟，因此本計畫將針對 APTS 中心資料庫所記錄之各項巨量資料，結合客運業者營運路線、班表、電子票證、駕駛員等資料進行巨量資料分析，針對不同來源的資料進行探勘、比對與分析，發現隱藏在各資料背後的重要訊息，並轉換為可應用於監督或管理之有效資訊。

1.2 計畫目的

本研究以行政院「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」之政策為重要目標，並達成交通部「建置智慧型運輸系統，推動永續綠運輸，符合節能減碳」的任務。因應巨量資料分析工具與方法之演進，本研究探討如何從公共運輸巨量資料中挖掘有用資訊，整合實務監理管理需求，視覺化資料分析成果，俾利應用於政府公共運輸之資源配置與監督，改善公路（市區）客運業者（以下簡稱汽車客運業）之經營管理模式，提升整體汽車客運服務品質，以達成政府運用巨量資料創造前瞻施政之目標。

1.3 分析範圍

1.3.1 研究區域評估

本研究主要目的為將 APTS 中心資料庫所記錄之各項巨量資料，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，並轉換為可應用於監督與管理之有效資訊。為達到此目的，本研究考量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 1-1 所示，研究區域評估流程分為五階段，分述如下：

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，目前部份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，部份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，本研究即以臺中市作為本計畫研究區域。

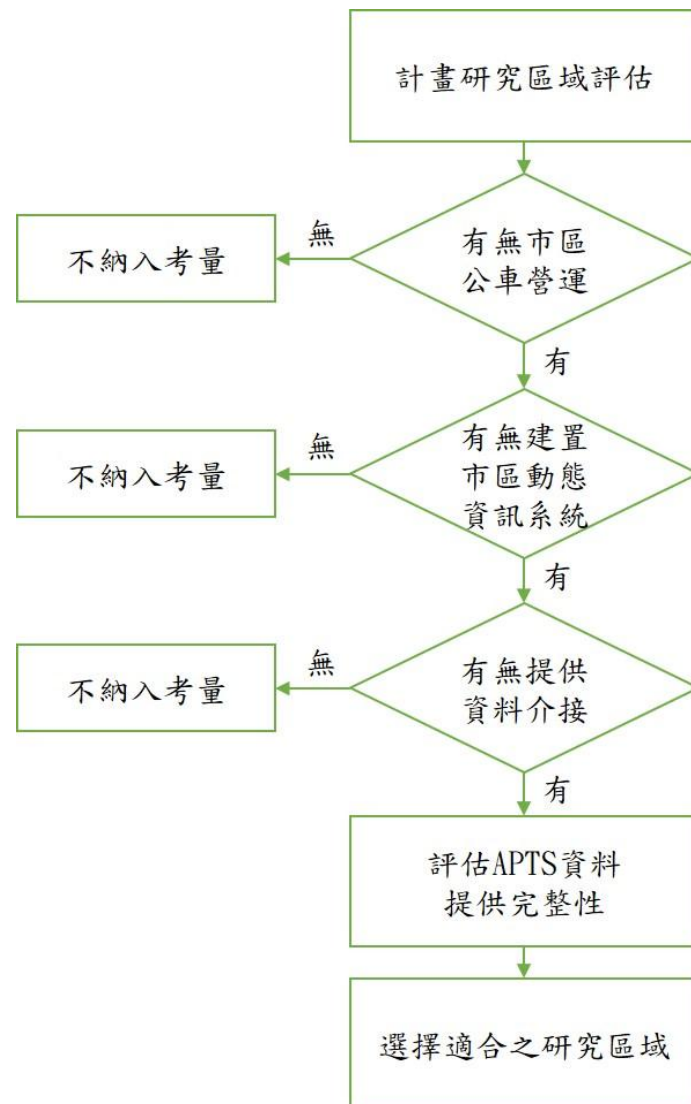


圖 1-1 分析區域評估流程

1.3.2 研究對象評估

根據前述評估流程，本研究係以臺中市作為分析區域，彙整行經臺中市汽車客運業者（表 1.1），比對客運業者經營路線，僅有豐原客運、全航客運、南投客運及臺中客運同時經營市區公車、公路客運及國道客運路線，其中，南投客運及全航客運於臺中市區公車營運路線較少，故不納入考量。而臺中客運於臺中市營運 42 條市區客運路線、豐原客運營運 81 條市區客運路線，且豐原客運涵蓋臺中市範圍較廣，故本研究係以豐原客運作為優先考量業者。

表 1.1 臺中市客運業者

客運業者	市區客運	公路客運	國道客運
豐原客運	◎	◎	◎
全航客運	◎	◎	◎
南投客運	◎	◎	◎
臺中客運	◎	◎	◎
苗栗客運	◎	◎	
仁友客運	◎	◎	
和欣客運	◎		◎
統聯客運	◎		◎
豐榮客運	◎		
中臺灣客運	◎		
東南客運	◎		
新竹客運		◎	◎
國光客運		◎	◎
臺西客運		◎	◎
巨業交通		◎	
花蓮客運		◎	
員林客運		◎	
彰化客運		◎	
杉林溪遊樂事業		◎	
總達客運		◎	
大有巴士			◎
中鹿客運			◎
阿羅哈客運			◎
建明客運			◎

1.4 工作項目

本研究的主要工作項目如下：

1. 訪談中央與主要縣市政府汽車客運主管機關，彙整汽車客運管理作業上應用巨量資料分析之實務需求。
2. 蒐集國內外公共運輸巨量資料應用之相關研究，並對應工作項目 1 中各主管機關實務需求彙整結果，規劃本計畫巨量資料之分析與應用範圍。

3. 探勘分析研究區域之公路客運或市區公車動態資訊系統資料庫(含靜態與動態資料)、標竿示範客運業者自主管理資料庫(如駕駛員出勤、績效考核、肇事違規與油耗等資料)、電子票証資料庫(如乘客旅次起迄資料)等相關資料。應用巨量資料技術，至少分析下列內容：
- (1) 研究區域中之汽車客運供給(路線分布、班次)與需求(如乘客旅次起迄資料)在不同時間與空間的分布狀況。
 - (2) 針對標竿示範客運業者，以營運路線為單位，針對不同班次、車輛、駕駛員之行車狀況(如旅行時間與延滯)進行比較分析，另協助業者建立營運路線各站牌之時刻表。
 - (3) 評估 APTS 各項設備(如車機)之穩定性與可靠度。
 - (4) 運用統計工具之資料探勘分析方法，針對標竿示範客運業者，建立駕駛員肇事或違規事件發生之風險分析。
 - (5) 運用巨量資料分析軟體，透過視覺化表達相關分析之重要資訊。
 - (6) 針對上述分析所觀察之問題或狀況，進行原因與趨勢探討，並研議對應之建議策略。

第二章 文獻回顧

2.1 公車動態資訊系統發展

先進公共運輸系統(Advanced Public Transportation System, APTS)係將 ATMS、ATIS 與 AVCSS 之技術應用於公共運輸系統，以改善公共運輸服務品質，提高營運效率，增加公共運輸系統之吸引力，且滿足各單位之功能需求。(運研所，2009)

國內對於 APTS 之推動係以「挑戰 2008：國家發展重點計畫」下之「聰明公車與國道客運計畫」為主。計畫主要內容為推動聰明公車之營運管理及服務資訊系統建置，進行都市公車動態資訊系統建置計畫、公路客運路線動態資訊系統推廣建置計畫、國道客運路線動態資訊系統建置計畫及大眾運輸車隊管理系統推廣建置計畫，以提昇大眾運輸服務品質及促進大眾運輸車隊管理的效能。

1. 發展歷程

國內最早 APTS 之研究，係由本所於 1989 年委託交通大學之研究計畫——「臺北市公車路線查詢系統之建立」，其後之相關研究，包括「無線電通訊在計程車及公車之應用與示範」、「臺灣地區先進公共運輸系統現況與引進先進技術之可行性研究」、「公車自動定位技術之研究」等，多屬小規模示範性計畫，直至 1994 年由本所及交通大學合作建置「公車動態資訊系統與高速公路替代道路選線即時導引系統」之實作計畫後，開始在新竹市、臺北市、以及金門縣建置較大規模的系統，再加上運研所提出「整合租用」的構想解決設備採購與升級的課題，國內 APTS 正式進入全面建置的階段。

近年 APTS 之建置計畫多以交通部「聰明公車」計畫下之都市公車系統建置為主。由於民眾對於公車動態資訊系統之反應良好，其動態資訊查詢使用率與服務滿意度皆尚稱滿意，因此，交通部公路總局於 2009 年參考聰明公車之成功經驗，導入公路汽車客運以建置公路客運動態資訊管理系統，目前已建置完成，並正式上線提供服務。

2. 建置概況

國內各縣市相關 APTS 計畫多著重於「聰明公車」之動態資訊系統建置及相關設備之擴充建置，目前已實施「聰明公車」之規劃建置縣市有：臺北市、新北市、基隆市、新竹市、新竹縣、桃園市、臺中市、臺南市、嘉義縣、高雄市、雲林縣、屏東縣、金門縣等縣市。而臺北市與新北市亦進行兩市公車動態的整合與交換。

3. 系統功能

就 APTS 之功能需求而言，應滿足民眾、業者、政府單位等不同需求。根據近年所執行之建置計畫，具體之系統功能項目分別如下所述。

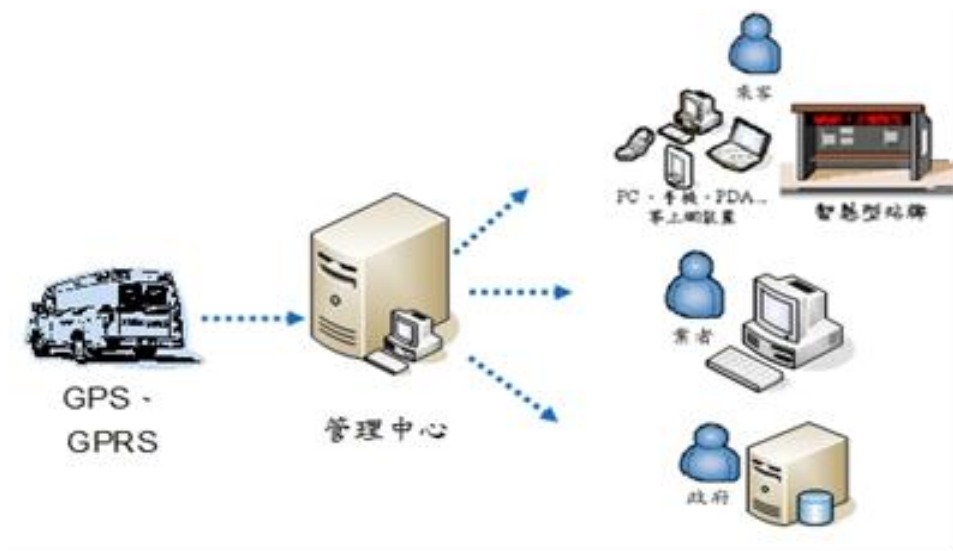
聰明公車之核心技術為透過車上單元之 GPS 定位及通訊(目前大多採用 GPRS)裝置，將相關動態資料回傳至管理中心，經管理中心進行資料處理後，再將資訊發佈給各資訊需求者，可提供：

- (1) 乘客面：乘客可透過網頁、行動裝置、語音電話、智慧型站牌等多元資訊介面，獲得車輛到站時間預測等時動態資訊。
- (2) 業者面：提供車輛即時追蹤、駕駛行為（怠速、急加減速、超速）行駛路徑、發車時間等營運調度資訊。
- (3) 政府面：提供監控車輛行駛路徑、緊急狀況通報等功能。

聰明公車之車輛動態資訊系統設備包括：

- (1) 管理中心：包含地理資訊系統、便民入口網站、動態資訊（車輛位置與到站時間預測）提供模組、行程規劃模組、分析模組等。
- (2) 車上單元（車機）：GPS 定位、GPRS（3G）通訊裝置等。
- (3) 查詢介面：網際網路、電話、行動裝置、智慧型站牌等。

其系統的運作方式為利用車上單元(車機)回傳資訊至管理中心，掌握車輛行駛資訊，提供管理者調度排班之決策支援，或縮短異常狀況處理時間與提昇車輛調度機動性；另一方面可透過中心發佈智慧型站牌、網頁、行動裝置、電話，提供民眾動態資訊，如圖 2-1 所示。



資料來源：運研所，2009

圖 2-1 聰明公車系統架構圖

2.2 巨量資料與相關分析技術

2.2.1 巨量資料定義、特性與趨勢

巨量資料 (Big Data) 是指因為資料量龐大，透過傳統的技術難以處理和分析，為新一代的技術和體系結構，透過大量取得、尋找或分析非常大量的各式各樣的資料。巨量資料的特徵在於三個方面，用 3Vs 定義大數據的 3 個特性，量 (Volume，資料大小)、速 (Velocity，資料輸入輸出的速度) 與多樣性 (Variety，資料具有多樣性)。巨量資料不只局限於上面提到三個特徵 (3Vs) 還可能擴大到四個 Vs，甚至 5 個 Vs，即量、速、多樣和價值 (Value) 及真實性 (Veracity)。這個定義被廣泛認可，因為它凸顯出巨量資料的必要性與意義 (Hashem 等人, 2015)。以下說明上述巨量資料之各項特性與分類。

1. **Volume**：所有類型的資料從不同來源大量產生且繼續擴大。蒐集大量的資料的好處包括創建隱藏的資訊和透過資料分析模式的特性。
2. **Variety**：非結構 (unstructured) 是巨量資料的重要特性，也是在關聯式資料庫上廣泛被使用的資訊探勘及統計分析等方法無法直接套用在巨量資料上的主要原因。關聯式資料庫的資料具有固定格式，巨量資料則包含各種不同格式的資料。異質性與多元性也

是巨量資料的重要特色。當巨量資料由不同型式的資料集合而成，如何利用不同型式的資料來提供整合性的服務？而當巨量資料由不同資料來源的資料集合而成，如何整合不同資料結構、不同的語意，亦是重要課題。

3. **Velocity**：由於資料流量太大，先存起來再慢慢分析並不足以因應作即時處理之需求。如何在大量資料流進來的時候直接執行過濾、或輔以適度的後設資料以加快處理速度是個技術挑戰。
4. **Value**：從多類型快速產生的巨量資料發現資料裡面隱藏的巨大價值。
5. **Veracity**：由於巨量資料可能來自不同來源、甚至來自廣大群眾，在利用資料的時候，不可避免地需要考慮資料的真實性。舉例來說，若有惡意的使用者利用社群網路散播假消息、或是不公正的使用者刻意於評價網站上詆毀特定賣方，都會影響到資料的真實性。

Wu 等人（2014）亦說明巨量資料的基本特性為資料量大且為異質（Heterogeneous），分散式主動（Autonomous）蒐集資料，並試圖探索資料間複雜（Complex）和不斷變化（Evolving）的關係，以這些特性為基礎提出 HACE Theorem。

巨量資料蒐集和分析大量的資料並應用各不同的領域，Kambatla 等人（2014）提出幾項巨量資料的應用領域，其分別說明如下：

1. **健康和人類福利**：大數據分析中最引人注意的應用之一就是在健康和人類福祉的領域，此領域也提出了些最大和快速增長的資料集，大數據分析對新的療法提供巨大潛力，此領域的大數據分析的成本效益也是廣泛認可的。
2. **自然環境**：這個時代緊迫的問題之一就是環境的長期影響變化，蒐集有關環保和其明顯衝擊的大量資料。
3. **公共部門**：政府和公共部門透過分析資料進而加快科學和工程的開發，保障國家安全。

4. **商務、商業和經濟體系**：最明顯的大數據分析的應用已經在企業上看到。據估計充分利用分析可以增加其營運獲利率達 60%。利用新的機會有很顯著的新收益。巨量數據分析作用在提高效率 and 指導決策過程。
5. **社會學**：描述社交網路資訊流動的速度和影響，人們每天透過社交網路溝通交流，由此產生的資訊被網路記錄下來，社會學家可以在這些資料的基礎上分析人類的行為模式、交往方式等。

此外，藉由交通巨量資料可能改變城市的運作規則。有研究發現，大約 30% 的汽車在城市交通正在尋找停車位，而巨量資料的分析可以幫助解決這個問題。Google 的無人駕駛汽車，搭載利用和快速分析大量的數據，已經記錄無事故行駛超過 70 萬英里。利用巨量資料以建立更高效的公共交通系統，提高使用的可能性，並可能顯著減少私人運具以及商用車輛的碳排放量。

在 20 世紀中期，當時網格計算（Grid Computing）是為重要的資源硬體，網格計算就是通過網際網路來共享強大的計算能力和資料儲存能力，網格計算技術注重於分享、選擇、各種不同地裡面貌的分佈來源，這些資源包含超級電腦、儲存和在科學、工程和商務方面計算大規模密集性問題，以上關鍵特性以透明的跨域管理和資源管理能力為主要特性。近年來許多大型公司推出雲端運算又分成三大類型，軟體即服務（Software as a Service, SaaS）、平臺即服務（Platform as a Service, PaaS）、基礎設施即服務（Infrastructure as a Service, IaaS），雲端技術的目標在於可靠和強大的服務能力，利用客戶電腦、平板電腦、手機和 I-Watch 等行動裝置來操作使用，從計算分析資源，使用可用的資料，最後是執行程式，都能在雲端上全部搞定，在使用雲端技術者能在硬體和管理成本方面能降低許多，總體資源的利用率能提高許多（Kambatla 等人, 2014）。

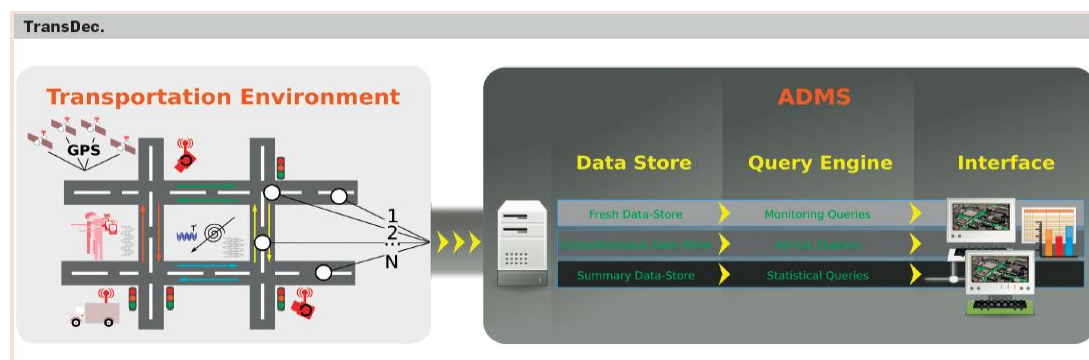
在 2008 年有研究指出全世界的企業服務系統的資料約 9.57×10^{21} ，並且預估每兩年會成長一倍，因此資料分析的規模逐年升高，表 2.1 顯示目前的資料處理的數量都是很龐大（Kambatla 等人, 2014）。

表 2.1 目前的資料處理的數量

案例	處理資料量
Walmart 每小時處理超過 100 萬交易質量	2.5 千兆位元
Facebook 每天使用者貼文資料和圖片分享	數百兆的百萬位元
Youtube	每分鐘上傳 100 小時的影片
MSN/即時通	每秒傳送 28,000 訊息資料
Twitter	服務超過 550 百萬使用者
Ebay	每天提供超過 100 千兆位元 資料處理

2.2.2 巨量資料分析技術

首先以洛杉磯大都會交通管理局(LAMetro)與南加州大學(USC)綜合多媒體系統中心(IMSC)的合作計畫說明交通運輸巨量資料分析的相關技術。2010 年秋季開始，LAMetro 與 USC IMSC 合作分析洛杉磯路網的高解析度時空交通資料。IMSC 開發了一套名為 TransDec (Transportation Decision-making) 的端到端系統以蒐集、儲存、分析和視覺化時空交通資料，其如圖 2-2 所示 (Jagadish 等人 2014)。



資料來源：Jagadish 等人, 2014

圖 2-2 TransDec 應用

以下說明時空交通資料的巨量資料分析方式：

1. **資料蒐集 (Acquisition)**：蒐集現有的即時資料，其包括線圈式車輛偵測器、公共汽車和軌道交通、匝道控制燈和可變標誌 (Changeable Message Signs)、事件 (如死亡人數、救護車到達時間)。

2. **資料淨化 (Cleaning)**：使用微軟的複雜事件處理 (Stream Insight) 刪除多餘的 XML，讓 46MB/分鐘的輸入資料縮減到 25MB/分鐘。然後將結果作成簡單的表儲存到微軟 Azure 雲端平臺上。
3. **整合/呈現**：使用 Oracle 11g 中的資料表將資料統合，並建立索引。例如，統合資料以建立綱要，以便支援預定的空間和時間查詢集合。
4. **分析**：應用多種機器學習技術已產生洛杉磯各路段的準確的交通模式，在一天內不同時段（例如，尖峰期），一周內的不同日期（例如，週末）及不同的季節。使用歷史事故資料來為新的事故分類，以預測事故排除時間和事故引發的交通停留的長度。
5. **解釋**：在複雜的系統中，很多事情都可能出錯，導致出現虛假的結果。例如，系統各（獨立的）元件的失效可能無人關注，導致資料遺漏。與此類似，有時候某個組織改變資料格式後，卻沒有通知下游的組織，導致分析錯誤。為了解決此類問題，業界開發了多個監控的程式，並配有相應的機制以獲取使用者的確認和修正。

巨量資料探勘的處理架構可以分為三個層面 (Wu 等人, 2014)，第一層為巨量資料探勘平臺，第二層為巨量資料應用，第三層為巨量資料探勘方法，其分別說明如下：

1. **巨量資料探勘平臺**：巨量資料通常儲存於不同的地方，資料量持續成長，因此需要一個有效率的平臺來儲存分散且巨大的資料量，資料探勘程式需要精密計算以進行資料分析和比較。
2. **巨量資料應用**：巨量資料應用有幾個重要的議題，資料分享是很多平臺系統所要做的目標，全世界關心巨量資料分析較敏感的資訊，像是交易、醫療紀律等。資料分享則會有隱私問題，重要資料被公開可能會有嚴重的後果，若要儲存以上資料需增加認證條款、限制某些使用者才能存取和匿名發出，以保障資料的安全，故需開發一個安全的認證或是存取的機制。知識的應用即需要了解每個工具背後的意義，才能在短時間內找到使用者所需的資料。

3. 巨量資料探勘方法：巨量資料的探勘系統可以交換資訊或是結合所有資訊以找出最佳的答案，可以分成三個層級：資料、模式及知識，如表 2.2 說明。

表 2.2 資料探勘平臺架構

層級	說明
資料	每個地方網站可以蒐集地方的資料，結合全球的地方網站就能得到全球資料分佈視圖之間的資料。
模式	每個地方網站地方的資料，作出有關地方的統計圖形或是模式，結合全球的地方網站就能得到全球資料分佈視圖和模式。
知識	模型相關性分析考察了從不同的資料來源，以確定哪些資料是有相關性以及如何準確的建立模型與模型之間的關聯性。

資料探勘大多都是探勘稀疏的、不確定的、不完整的資料。表 2.3 為資料型態會發生的問題及解決方法。巨量資料的資料量越來越多也越來越複雜，來源從網路、社群網路、通訊網站等等，對於巨量資料的應用，複雜的資料來源是一個重大的挑戰，像是 Facebook 全球使用者約十億人，要分析一個巨大的社群網路是一個巨大的挑戰，研究使用者與使用者之間的關係是非常困難。

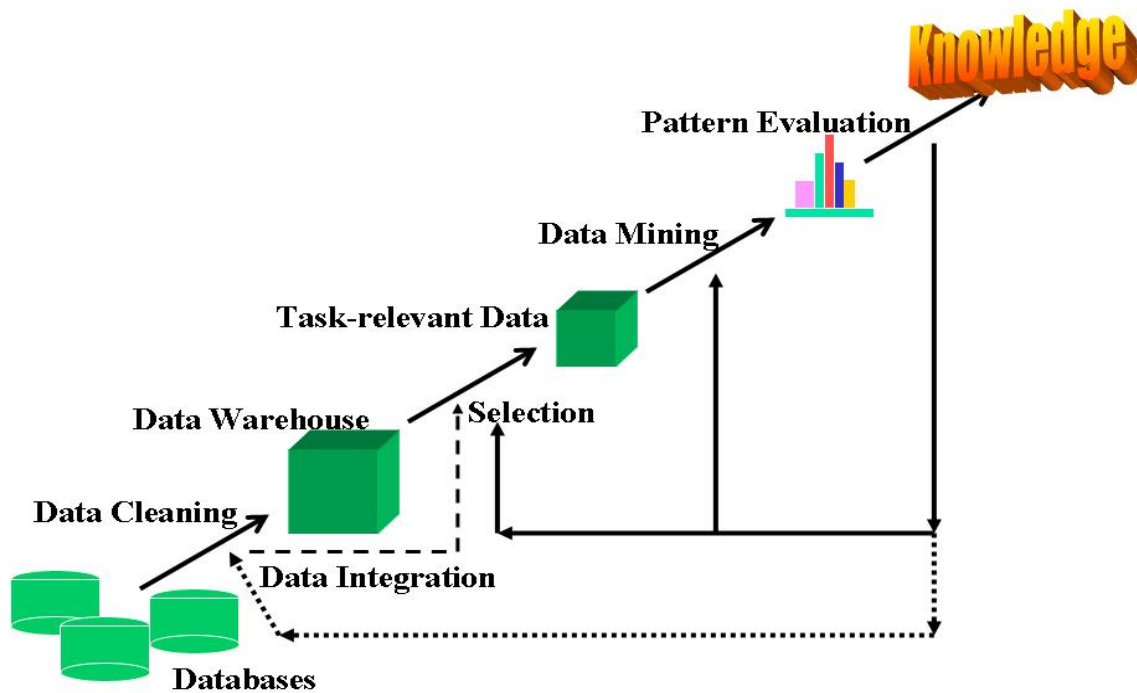
表 2.3 資料型態之解決方法

資料型態	問題	解決方法
稀疏性	通常發生在較高維度的資料，無法做出一個可靠的結論以及劃出趨勢圖或是分佈圖。	降低維度或是在增加樣本數。
不確定性	分析時使用錯誤的讀取方法或蒐集方法，因此會出現隨機或錯誤的分佈關係。	解決辦法為用資料的分佈去估計模式的參數。
不完整性質	缺少樣本資料的欄位值，缺少這些值可能會造成結果的偏誤。	新的資料探勘演算法可以處理這些問題，將資料做些處理，把缺少的值做些改進。

由於多源、大規模的、異構的和在分散式環境中，其所涉及的應用程式資料的動態特性，巨量資料的最重要特徵之一為進行複雜的計算過程以計算 petabyte (PB) 甚至是 exabyte (EB) 等級的資料 (Wu 等人, 2014)。因此利用一種平行計算的基礎設施、相應的程式設計語言支援，以及軟體模型以有效地分析和探勘分散式的資料，這是將巨量資料的處理從量改為品質的關鍵目標。Google 提出一個軟體架構，MapReduce，用於處理大量資料（大於 1TB）的平行運算，擁有可靠的容錯機制。Hadoop 是一款處理與保存大量資料的雲端運算平臺，主要應用在大量的資料運算與處理，其為目前最熱門用來處理、分析、與查詢巨量資料的平臺。Hadoop 是整合 MapReduce 演算法架構後，所開發出來的雲端演算法。之後也有人提出了分佈式協同聚合 (DisCo) 框架採用實用的分佈式數據前處理和協同聚合技術。在開放資源 MapReduce 的項目中，Hadoop 的應用說明 DisCo 具有完美的擴展性及可處理和分析大量的資料集。

許多資料探勘的技術也可以應用在巨量資料分析。Tan 等人 (2006) 將資料探勘定義為「整合所有從資料庫中挖掘知識的過程，而主要目的為將有原始資料轉換為有用的資訊。」而依 Han & Kamber (2001) 的定義「資料探勘主要從大量的資料中擷取或挖掘有意義的特徵，而這些資料來源可為儲存在資料庫、資料倉儲或其它資訊貯藏處」。

詳細的資料探勘流程如圖 2-3 所示，資料探勘主要程序包含資料淨化 (Data Cleaning)、資料選取 (Data Selection)、資料屬性選擇 (Feature Selection)、資料轉換及篩減 (Data Transformation and Reduction)、資料探勘 (Data Mining)、特徵過濾 (Pattern Filtering) 與特徵解釋 (Pattern Interpretation) 等步驟。



資料來源：Han and Kamber, 2001

圖 2-3 資料探勘流程圖

1. 概念/類別描述 (Concept/Class Description)：概念描述 (Concept Description)，有時又稱為類別描述 (Class Description)，是最簡易的一種描述性探勘 (Descriptive Mining)，其主要描述資料的特徵，或對資料作一般化 (Generalization) 與摘要化 (Summarization) 特徵描述。最常見的方法為屬性導向演繹 (Attribute-Oriented Induction, AOI)，透過主觀想法將資料之屬性欄位依可用性做適度篩減，並透過離散化方法將數值資料做類別化描述，以提高分析之正確性。
2. 關聯法則探勘 (Association Rule Mining)：關聯法則 (Association Rule) 主要找尋品項間的關聯程度，找出有趣且關聯性高的規則。過去廣泛應用於交易資料庫，分析顧客的購買行為是否存在品項間採購關聯之規則或特殊的購買模式，關聯法則又稱為購物籃分析 (Basket Analysis)。
3. 分類與預測 (Classification and Prediction)：分類 (Classification) 是藉由訓練資料 (Training Data) 學習出資料的分類規則，且資料具有目標值 (Target Value)，再依據事先定義的目標值，預測未知資料所屬的類別。並以測試資料驗證先前訓練的分類器 (Classifier) 分類準確程度之高低，因此，分類屬於監督式學習

(Supervised Learning)。常用分類方法包含決策樹(Decision Tree)、類神經網路(Neural Network)及支援向量機等。預測(Prediction)是根據對象屬性之過去觀察值推估該屬性未來之值。常用方法為迴歸分析、時間數列分析與類神經網路等。

4. 分群分析 (Cluster Analysis)：分群(Clustering)與分類最主要差異在於分類具有事先定義的目標值，而分群沒有。因此資料分群是透過分析資料間屬性特徵的相似程度，依循「群組內的資料相似度最高，群組間的資料相似度最低」之規則，將同質性較高的資料集結成群。由於分群並未事先定義目標值，因此，屬於非監督式學習(Unsupervised Learning)。
5. 推估分析(Estimation Analysis)：推估分析與分類方法類似，均為依據已定義之歷史資料，建立資料之分類模式。推估與分類最大差異在於分類是針對離散型資料作類別歸納，而推估則是針對連續型的資料。
6. 序列樣式探勘(Sequential Pattern Mining)：前述關聯法則探勘僅止於找尋出顧客購買品項間的關聯性與潛在特殊採購模式，但並未找尋品項間被採購的順序關係，若深入探討品項購買之時間性順序關係。該資料分析則稱為序列樣式探勘。
7. 離群值分析(Outlier Analysis)：離群值分析是資料分群的一種特殊資料分析應用，當資料均適度分群後，未被分群的離群值便有可能是異常資料。資料庫中經常包含此類資料，這些資料與其他大多數資料具有明顯差異，對資料分析準確性造成顯著影響。因此，通常會排除這些資料後才進行資料探勘，或是於特定領域透過異常分析找出不合理的資料。
8. 演化分析(Evolution Analysis)：演化分析主要以時間為維度，用以表示資料在持續期間下之趨勢或規律。

2.3 公共運輸巨量資料分析案例

目前大眾運輸已透過公車動態資訊系統及電子票證蒐集大量巨量資料，但除臺中市等少數縣市將公車動態資訊系統應用在補貼之監督

管理、及臺南市、臺北市將其應用在站站有班表之推估、與臺北市及臺中市應用在大眾運輸之起迄資料推估外，並未有大规模之應用，更未跟其他屬性資料庫進行空間與時間之綜合分析。茲就在國內外大眾運輸巨量資料之應用案例分別說明如下：

1. 大眾運輸路線設計

蒐集電子票證資料或手機通聯紀錄資料，巨量分析並推估旅次起迄對，設計滿足符合起迄需求的公共運輸路線。首爾市政府分析三十億筆通話紀錄，發現即使到了深夜，民眾的活動都相當活躍，且通話地點距帳單地址都有一段遙遠的距離，推測民眾仍有使用大眾運輸的需求，於是分析找出民眾最常活動的區域與時間，設計如圖 2-4 所示之午夜公車路線。



影像來源：首爾市廳；資料來源：財訊第 463 期

圖 2-4 貓頭鷹深夜公車 N37 與公車電光板（中間）和深夜公車標誌

2. 路線營運改善

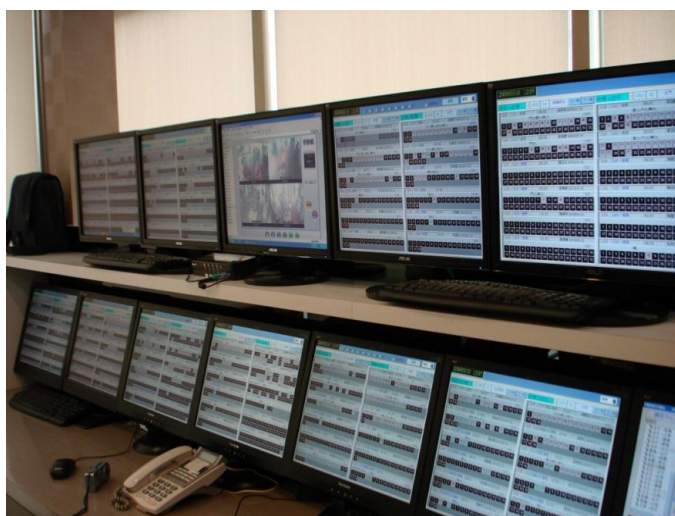
Pelletire 等人（2011）曾將智慧卡在大眾運輸規劃上之應用型態區分為戰略層級（strategic-level）、戰術層級（tactical-level）及營運層級（operational-level），研究中指出在戰略層級之應用研究（Agard 等人, 2006; Bagchi & White, 2005; Chu 等人, 2009）主要著重在營運機構制定行銷策略、營運計畫檢討等長期性規劃；戰術層級之應用研究（Bagchi & White, 2004; Chapleau & Chu, 2007）則偏重在發車頻率或班表的調整；營運層級的研究（Chapleau & Chu, 2007; Trepanier 等人, 2009）則偏重在如何利用電子票證資料進行運量統計及各種營運指標的產製。Seaborn 等人（2009）曾利用電子票證資料進行倫敦替區跨運具轉乘行為之研究，利用電子票證資料探討使用者在不同運具間

之轉乘時間。Morency 等人（2007）曾利用資料探勘技術分析加拿大不同大眾運輸使用者的時空特性分析。New York City Transit Authority (NYCTA) 利用 218 條公車路線每日上下車資料(約 2.8 百萬筆)，進行排班與路線規劃。藉由此資料可得到 128 條路線每日使用者之 OD，客運業者將此 OD 資料與天氣、交通資訊及特殊事件進行資料融合並進行分析，以作為隔日排班及路線規劃之依據(Zeng 等人, 2015)。

國內最早之研究為邱詩淳（2005）運用悠遊卡資料進行臺北市特定公車路線之營運改善方案；羅惟元（2008）同樣以悠遊卡資料探討公車路線之乘客起迄資料；臺北市公共運輸處曾於 2013 年結合公車動態資訊系統與悠遊卡刷卡資料進行乘客起迄資料的推估。惟目前此三篇研究較限於小規模之公車路線分析，且均侷限在乘客起迄資料之推估，對於大眾運輸規劃品質之提升較為有限。蘇昭銘等人（2014）曾針對臺灣之電子票證資料建構 52 項應用分析指標，同時以某個縣市之資料進行實際之應用，確認各項指標之實用性。

3. 環保駕駛教育

藉由數位行車紀錄器蒐集轉速、加速度、車輛速率等巨量資料，掌握駕駛行為，發現環保駕駛及危險駕駛行為模式，提供後續環保駕駛推動油耗狀況檢討，並能進一步有效預防與減低行車危險的發生。例如圖 2-5 為和欣客運公司透過行控中心監視車輛位置，並依其所回傳相關駕駛行為及油耗量數據，結合教育訓練，驗證了環保駕駛的節能減碳成效。

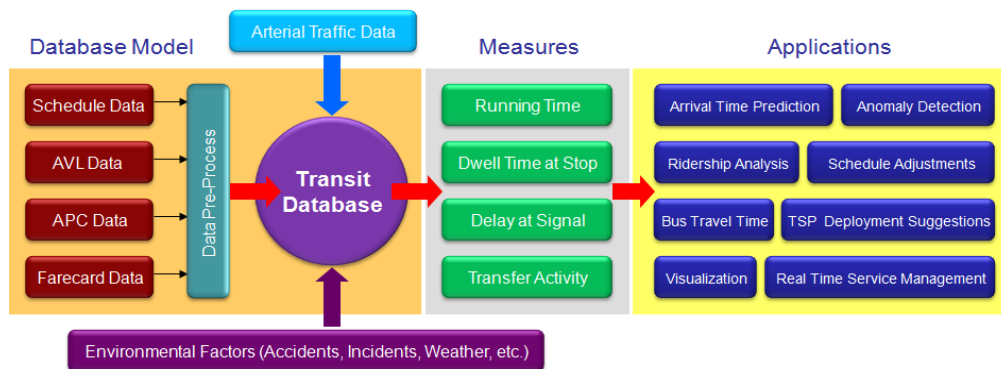


影像來源：本研究拍攝

圖 2-5 和欣客運行控中心

4. 動態車輛調度

美國許多公共運輸單位均建立自動資料蒐集系統(Automatic Data Collection Systems, ADCS)管理其車隊，藉由 ADCS 監控車隊績效、輔助班表之規劃及改善服務品質。Minneapolis 和 St. Paul 兩地區之 CBD 每日有超過 200 條路線。ADCS 資料庫包含車輛位置、載客數及電子票證交易紀錄，ADCS 可被用來取得改善大眾運輸的資訊與探勘出有價值的資料。Liao (2011) 建立一大眾運輸資料分析架構(圖 2-6)，此架構包含資料庫、估計方法及應用，資料庫來源包括大眾運輸資訊、道路交通資訊及其他環境影響因子，藉由這些資料計算公車行駛時間、上下車時間、號誌延滯時間及轉乘行為，並透過這些估計結果，將其應用到到達時間預估、需求分析、旅行時間估計及視覺化應用等。



資料來源：Liao (2011)

圖 2-6 大眾運輸資料分析架構

利用攝影機或各種感測器，調查停靠站或車上人數後，客運公司再進一步依據其數據分析調度車輛。例如圖 2-7，鎮江的公車停靠站會利用攝影機捕捉等車人數，以及電信公司會蒐集每天同樣時間的手機號碼，還有每幾分鐘後手機號碼會出現在哪個停靠站，再以巨量分析判斷乘客搭乘的路線，客運公司在透過電腦調整發車時間及密度，動態調度公車讓公車空駛耗損率下降 25%。

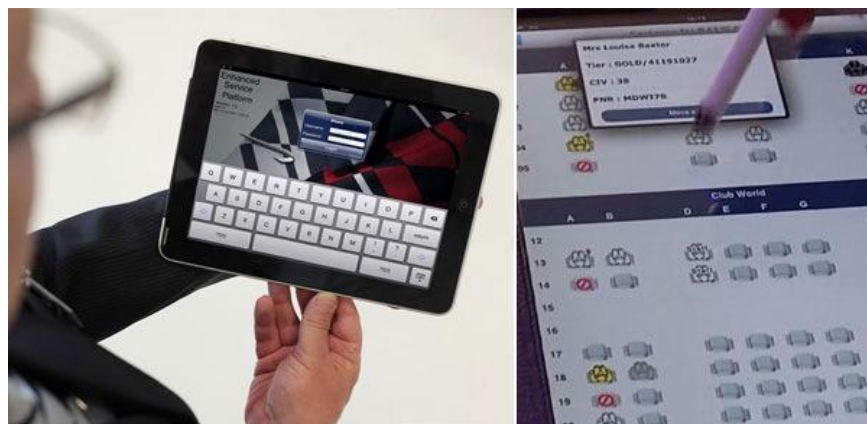


資料及影像來源：財訊第 463 期

圖 2-7 鎮江公車停靠站

5. 客戶關係管理與行銷

蒐集旅客在客運網站、行動應用、社交媒體上的行為數據，通過對巨量數據進行深度探勘，並在各種營銷活動中適時動態識別旅客及行為特徵和變化，提供客製化的產品或服務，用以改善服務品質，也能發掘新客源。例如圖 2-8 英國航空公司分析旅客購買習慣，發現乘客選擇靠窗的旅客多為短途航班，而選擇靠走道的旅客多為長途飛行的航班，當旅客下次再購買時自動為乘客推薦合適的座位。



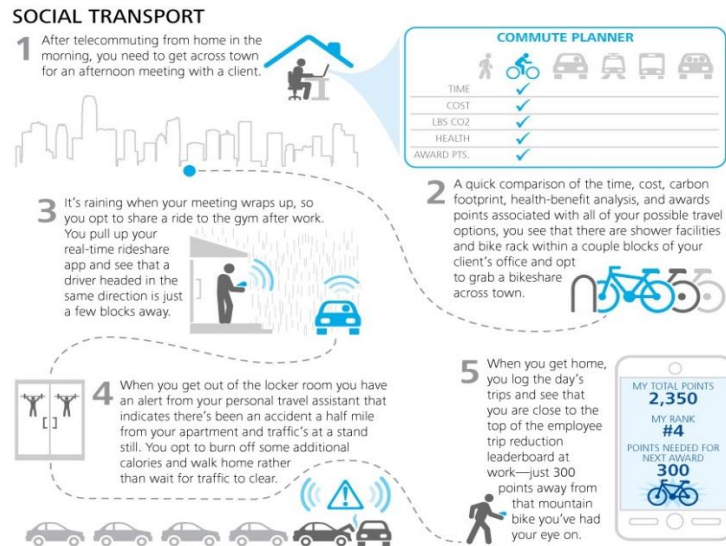
影像及資料來源：futuretravelexperience 網站

圖 2-8 英國航空改善使用者體驗

6. 旅運行程最佳化

通過巨量分析提供適地性服務。例如為旅客提供精準的廣告或資訊推送服務，提前為旅客個性化的需求進行滿足；為長時間等待公共運輸的旅客提供候機或候車服務；為到達目的地的旅客提供交通、餐飲、娛樂或景點等資訊等。例如圖 2-9 的情境說明，旅客下午與人有約，旅運行程最佳化系統分析各運具旅行時間、成本、二氧化碳排放

及健康因素相關資訊後，建議旅客自行騎腳踏車前往赴約，結束約會時正打算去健身中心卻發現下雨，系統這時建議旅客與人共乘車輛至健身中心，結束完活動回家，系統根據道路交通流量擁擠狀況，於是建議旅客採取步行方式。



資料來源：dupress 網站

圖 2-9 旅運行程最佳化情境說明

7. 運具資訊整合

旅客輸入起迄點及出發時間等資訊，透過巨量分析，而得到行前、行中等資訊服務。如圖 2-10 所示之運具資訊整合系統能分別分析步行、騎腳踏車、自行開車、搭乘公車、搭乘軌道運輸及共乘車輛所需花費的時間、距離與二氧化碳排放量，並能提供導引方向及興趣點位置等整合與跨運具的資訊。



影像來源：dupress 網站

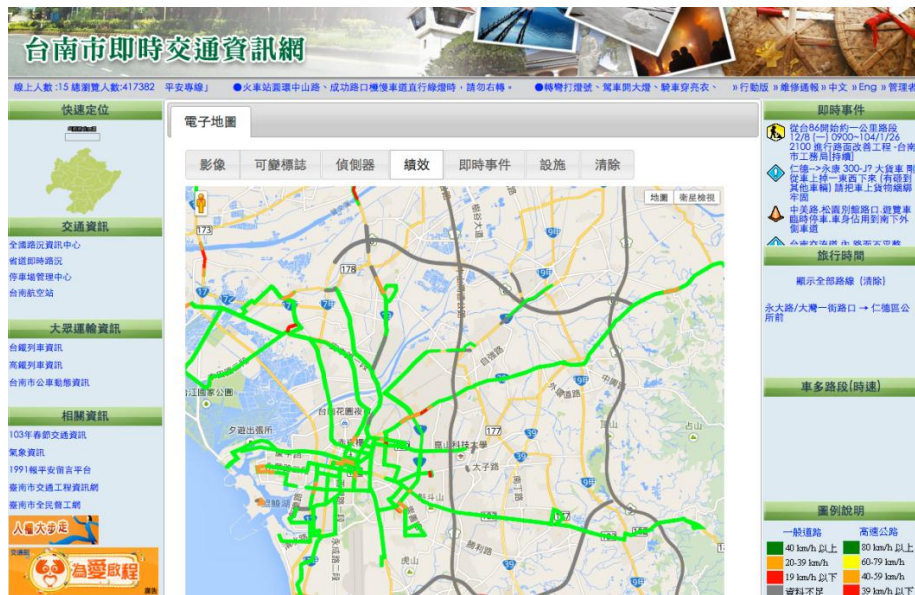
圖 2-10 旅運資訊面板

8. 安全管理

研究主要在以空間分析單元為基礎，探討交通意外事故與各單元社經變數之關聯性，近年來已有 Levine 等人 (1995)、Aguero-Valverde 及 Jovanis (2006)、Traynor (2008)、陳一昌等人 (2010) 均曾針對地區之肇事與空間因子之關聯性進行研究，其中陳一昌等人 (2010) 之研究僅提出觀念性架構。綜合前述探討，可發現結合地理資訊系統技術進行地區之交通意外事故嚴重性分析，並針對各地區之社經或環境因素進行探討已日益受到重視。本研究在將資料取得可行性納入分析後，發現目前並無法取得以里為單元之所得、車輛持有、交通屬性、就業特性及氣候條件等資料，故在後續研究終將僅就人口特性、道路等級等層面進行分析。

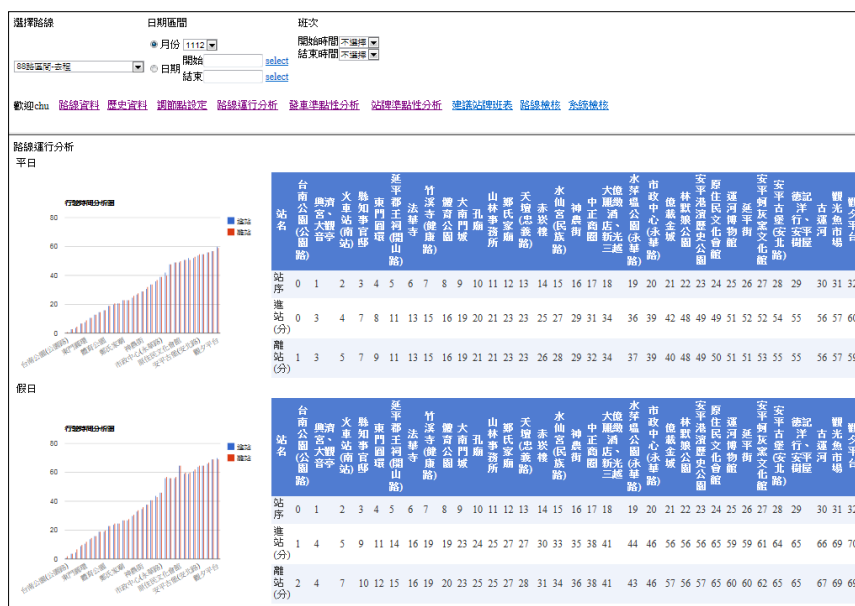
9. 路況預測

運用監測器或已安裝定位模組的客運車輛作為探針車 (Probe Vehicle)，蒐集車輛行駛的位置及車速等巨量資料，掌握道路即時路況，預測車輛行駛時間，並進一步結合可變標誌提醒或導引駕駛，藉此疏導車流，例如圖 2-11 顯示臺南市即時交通資訊網提供即時交通事件、預估旅行時間及車多路段等資訊；圖 2-12 則為臺南市公車到站時刻管理系統之系統畫面 (蘇昭銘等人，2011)。



影像來源：臺南市政府

圖 2-11 臺南市即時交通資訊網



資料來源：蘇紹銘等人（2011）

圖 2-12 臺南市公車到站時刻管理系統之系統畫面

10. 視覺化應用

Du 等人（2013）發展 MetroViz 系統藉由視覺化的分析，幫助使用者快速了解公共運輸資料及評估公共運輸服務品質。此系統公共運輸資料分為三個層級，第一為站牌/車站層級，第二為路線層級，第三層級為旅次層級，綜整這三個層級，此系統可呈現地圖視圖（map view）、路線視圖（route view）、日曆視圖（calendar view）。地圖視圖（map view）及路線視圖（route view）可幫助使用者了解站點位

置及路線資訊，日曆視圖（calendar view）則可反映每年的需求及承載狀況。此研究應用黑堡（維吉尼亞州）三年大眾運輸資料建立系統並分析，此資料庫包含 2 百萬筆資料。MetroViz 系統如圖 2-13 所示，左上角應用 Google map 展現黑堡（維吉尼亞州）地區公共運輸情形，點選站點則可顯示該站點費率分布情形（如學生票、全票、轉乘等）。左下角則顯示與該站點有關的路線。圖 2-13 右邊則為日曆視圖（calendar view），每一個方格為每日搭乘人數，點選每日日曆可進一步觀察每小時的搭乘人數。

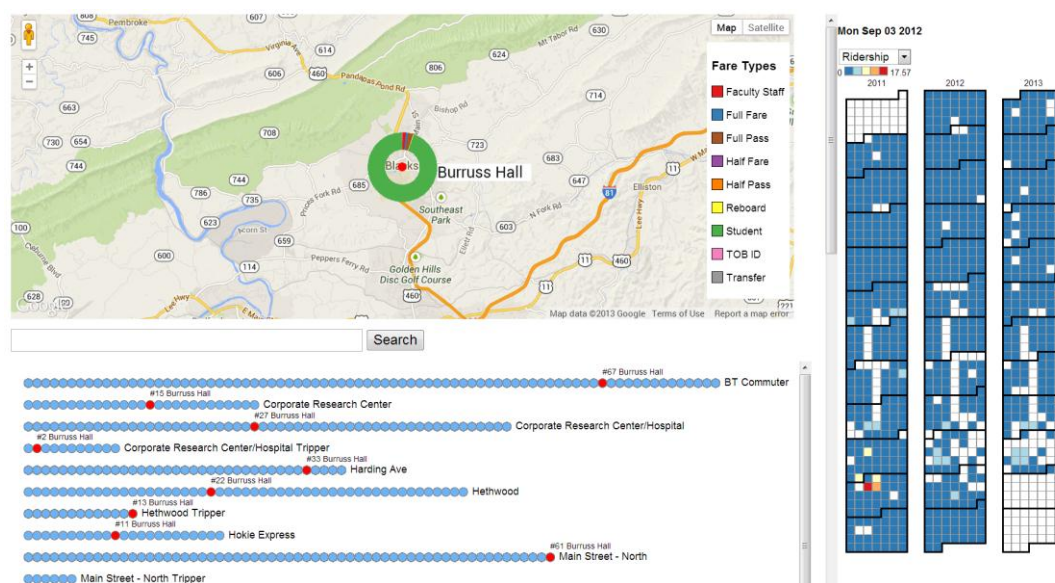


圖 2-13 MetroViz 系統示意圖

荷蘭為了找出無效率、瓶頸路線及有潛力的大眾運輸路線，應用自動車輛定位資料（Automatic Vehicle Location, AVL）分析各站點之延誤情形 van Oort 等人（2013）。荷蘭境內公共運輸資訊藉由 Borderless Public Transport Information 共享，各家業者將資訊提供給此平臺進行交換，其交換資訊包含班表、費率、車輛位置及準點率。為使各家業者資訊可順利交換，當地政府訂定了一個資料交換標準（BISON），藉由此資料平臺，提供了公共運輸巨量資料供研究單位分析。van Oort 等人（2013）之研究即應用 AVL 分析公車於各站點延誤之情形，蒐集每輛公車從起站點發車時間、抵達各站時間、離開各站時間、在路線上旅行時間及抵達終點站時間，與班表時間比較，分析其延誤情形，並透過地圖疊圖將其視覺化，如圖 2-14 所示，圖中綠色表示車輛早到，黃色表準時，紅色表延誤，藉由此圖形化的表示，可清楚了解哪一段容易造成車輛的延誤。

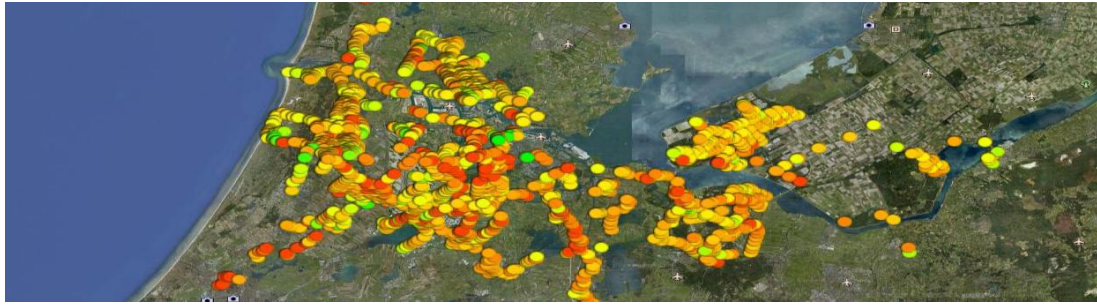
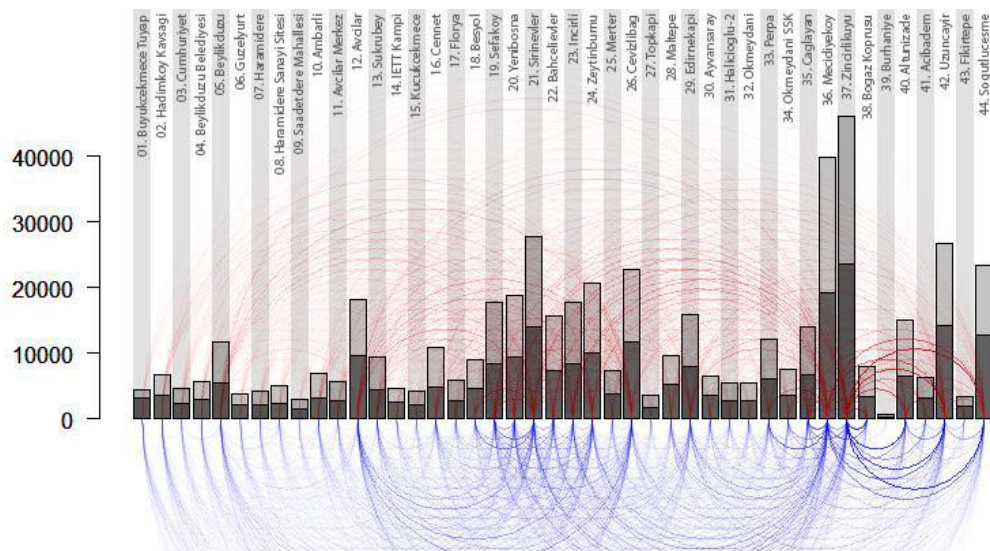


圖 2-14 公共運輸延誤視覺化之應用

土耳其伊斯坦堡透過電子票證系統蒐集大量的 BRT 公車路線營運資料，藉由此系統廣泛地蒐集電子票證交易資料（交易時間日期、付款狀態、卡片序號、費率、乘車路線等）及使用者旅運資料。Gokasar 等人（2015）空間與時間兩個維度分析旅運行為，並運用每日 800,000 筆的旅運資料及 460,000 筆電子票證 ID，分析多元運具的選擇、旅行及候車時間，進一步利用這些因子分析服務品質，並提供 BRT 系統管理及營運上的建議。以圖 2-15 為例，此圖藉由長條圖分析到達（灰色）及離開（深灰）的旅客數，並利用紅色（往亞洲方向）與藍色（往歐洲方向）代表方向，可清楚的了解哪一個站為重要的轉運點，進而針對這些站點研擬服務品質改善策略。

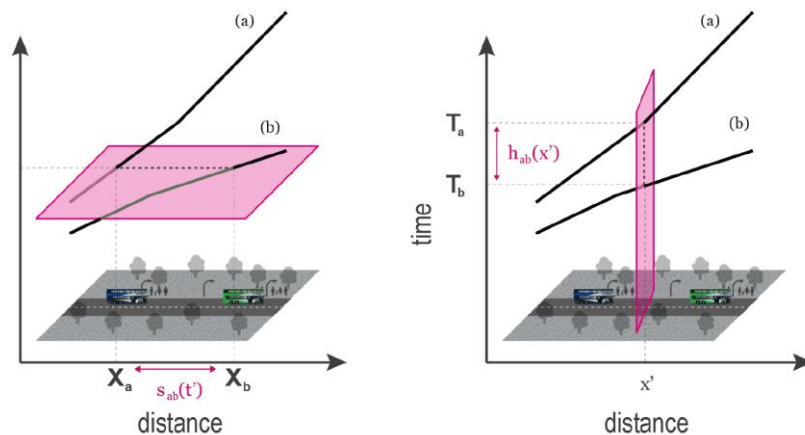


資料來源：Gokasar et al.（2015）

圖 2-15 土耳其伊斯坦堡 BRT 各站運量及方向

Anwar 等人（2015）應用新加坡陸運交通管理局（Land Transport Authority）監控 400 條公車路線資料（公車 GPS 資料、乘客載運資料）繪製公車路線時間與空間之 3D 視圖。每一筆資料包含紀錄時階、路線編號、站點編號、公車編號、行駛方向、最大載客量及載客數等。另外，藉由電子票證可蒐集到乘客的旅運行為。此研究將前述之資料

藉由 3D 視圖的方式呈現，藉由時間與空間的關係描繪公車移動情形如圖 2-16， x 與 y 標示公車位置（經緯度）， t 則表示時間，每輛公車在每一個時階（ t ）均會記錄一組（ x,y ）座標，應用公車在時空的關係，可進一步觀察公車之空間間距（ S ）與時間間距（ h ），如圖 2-16 所示。Anwar 等人（2015）將資料探勘概念應用到 R-tree，使探勘項目具有空間物件特性，此研究以 R-tree 作為空間指標進行分析。為使用使用者容易使用與分析，此研究建立一視覺化應用軟體供相關單位使用，如圖 2-17 所示，公共運輸營運單位或管理者，藉由圖 2-17 中 B 介面選擇欲分析之區域，由 C 部分選擇分析的設定如路線、時間、空間間距（ S ）或時間間距（ h ），A 介面則會將其結果透過視覺化的方式呈現，藉由空間間距（ S ）或時間間距（ h ）可檢視公車的發車班距是否恰當，民眾等候時間是否過長等訊息。



資料來源：Anwar et al.（2015）

圖 2-16 公車空間間距（ S ）與時間間距（ h ）視覺化示意圖



資料來源：Anwar et al.（2015）

圖 2-17 3D 公車間距視覺化分析軟體

2.4 可視化應用工具及案例回顧

2.4.1 應用工具及技術

在可視化應用部分，目前國內外已有幾種可視化顯示技術及工具，如 Google Maps API、D3.js(Data-Driven Documents)及 SAS Visual Analytics，以下本團隊先就這幾項技術及工具進行回顧及說明：

1. Google Maps API

Google Maps API 是一個以網路運用為基礎的高效能雲端服務平臺，除提供全球性的影像與向量資料外，更提供創新的加值運用及完整的 API，目前已提供超過 100 個國家之詳細路網資料，衛星影像涵蓋全球超過 30%之地表面積與 50%的人口涵蓋率，且已經提供數量超過 200 萬的興趣點資料。

現有 Google Maps API 為國內外普遍使用的技術，可使用之 API 有許多種類，如 Static Maps API、Places API、JS Maps API 等 API 項目，如圖 2-18 所示，而考量在可視化應用上，需能表示資料之密集程度，亦有常見的 2 種呈現方式，第 1 種為 Marker Cluster(標記叢集)，如圖 2-19 所示，第 2 種為 Heatmaps(熱力圖)，如圖 2-20 所示。



圖 2-18 Google Maps API 類型圖



圖 2-19 Google Maps 標記叢集顯示



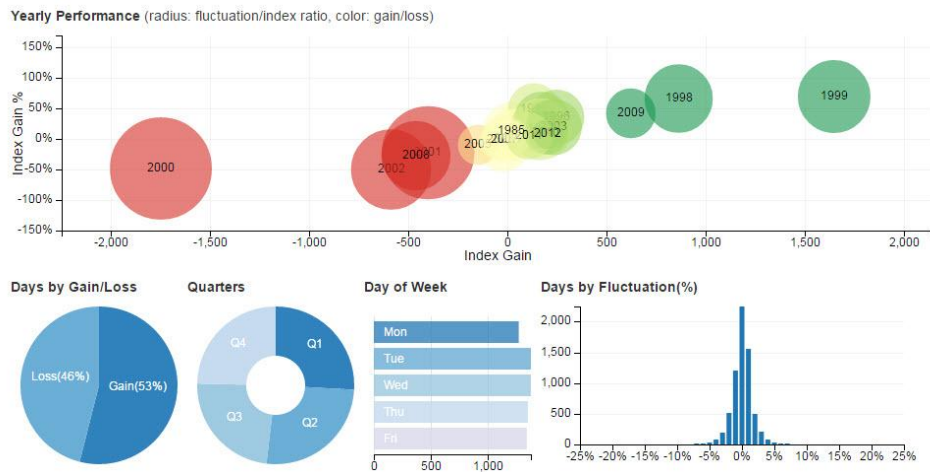
圖 2-20 Google Maps 熱力圖顯示

2. D3.js (Data-Driven Documents)

D3.js 是一個用動態圖形顯示數據資料的 JavaScript 函式庫，自 2011 年開始釋出，目前已經更新到 3.0.0 版，此工具可將資料進行可視化分析，並且廣泛引用 SVG、JavaScript 和 CSS 標準等技術，當資料載入後，利用資料的內容驅動 html 上的物件，甚至可將使用者的操作物件轉換成其他形狀顯示，效能上也提高許多。

D3 能於許多瀏覽器上執行，主要在於 D3 核心程式庫所需要的最低條件為瀏覽器能執行 JavaScript 和 W3C DOM API，目前 D3 可涵蓋

Firefox、Chrome、Safari、Opera、IE9+ 等瀏覽器上執行。此外，D3 提供許多種可視化呈現效果，如圓餅圖、長條圖等，其呈現效果強調互動性，如圖 2-21 所示，則為應用 D3 之可視化範例。



資料來源：<http://dc-js.github.io/dc.js/>

圖 2-21 D3 可視化範例

圖 2-22 為採用 D3 及 Google Maps API 技術顯示臺北市(A2)交通事故熱點資訊，在地圖上以 Google Maps API 熱力圖方式展示臺北市(A2)交通事故熱點位置，另外以 D3 所提供之長條圖等圖示呈現事故發生時間，其資料欄位可涵蓋時間、(A2)交通事故熱點空間位置。

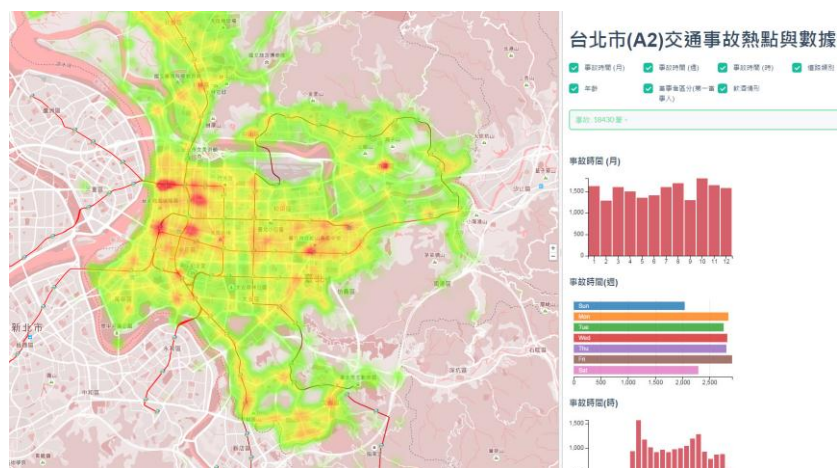


圖 2-22 D3 及 Google Maps API 顯示臺北市交通事故可視化

3. SAS Visual Analytics

SAS Visual Analytics 是一項容易使用、以 Web 為基礎且採用 SAS 高效能分析技術的軟體，可透過大量資料匯入，並結合 GIS 資料，以視覺化方式展示大量資料成果，圖 2-23 為 SAS Visual Analytics 顯示臺北市學生卡熱點可視化畫面，其資料欄位可涵蓋時間、公車站牌位置、公車站牌名稱及學生卡刷卡紀錄，而採用 SAS Visual Analytics 之優點如下：

(1)大資料分析應用(Big Data Analysis)

- ①可透過平行運作架構突破資料量的限制。
- ②使用者可以自己進行分析及產製報表。

(2)互動性視覺化分析(Interactive Data Visualization)

- ①以圖形方式展示，比數字更容易理解。
- ②使用者可彈性拖拉，以取得想看的資料內容

(3)行動商業智慧分析(Mobility BI)：

可利用行動裝置即時查詢取得分析內容及報表資料。

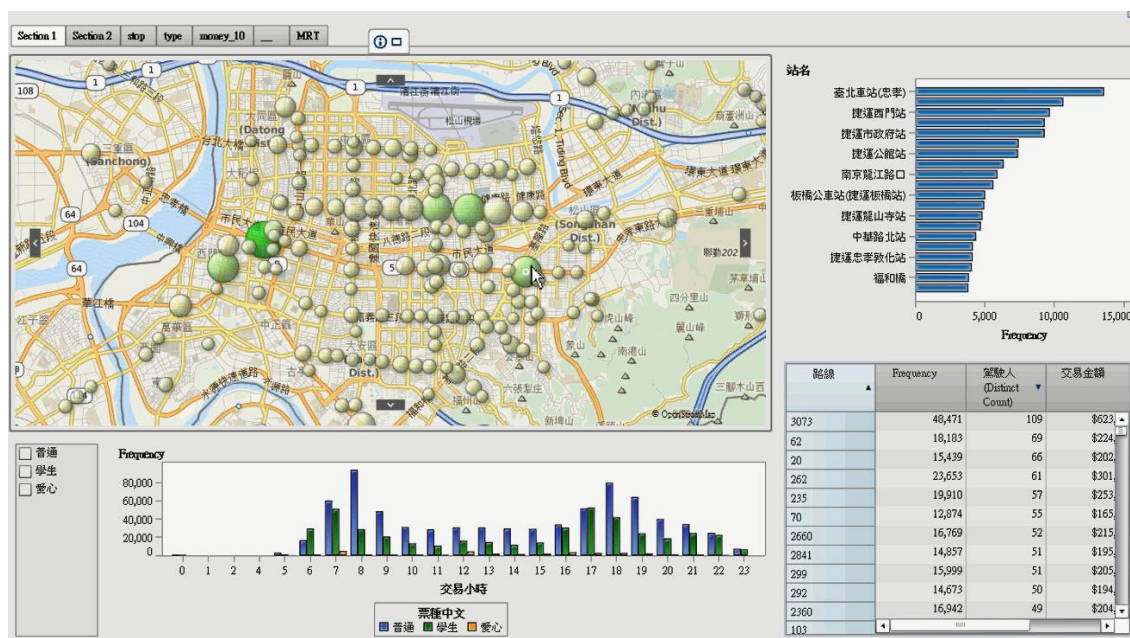


圖 2-23 SAS Visual Analytics 顯示臺北市學生卡熱點可視化畫面

上述回顧各項可視化應用工具，可發現在呈現上會採用圖表(如長條圖、圓餅圖)與地圖密度顯示(如熱力圖)以互動方式呈現，打破過往單純採用圖表呈現方式，能有助於管理層面之決策應用分析。

2.4.2 相關案例

每日生活中之運輸行為皆因為社會經濟活動的需求而起，以致運輸資訊應用服務興起之勢銳不可擋，其發展結果更帶動旅運者行程規劃、運輸工具選擇行為改變及提升運輸系統效率等直接的效益與貢獻。未來針對各種多元的交通服務情境對應的資訊服務內容，將逐漸滿足一般旅運者的資訊需求。目前彙整的相關案例中可知，結合應用大數據視覺化分析技術，進行公共運輸供給與需求在時間與空間維度上之應用研究，提供交通主管機關和業者決策參考依據，並提供使用者加值資訊應用服務已是普遍的主流應用方向，其相關案例彙整如下。

1. 臺北「UBike」App

UBike 能讓使用者方便掌握臺北市公共自行車各站點的可租借數量與停車格數量，更利用市政府釋出的開放資料預測平均等車時間，以利使用者預判熱門站點借還車所需的等待時間。該 App 每五秒即時更新公共自行車資訊，讓使用者便利查詢站點當下即時資訊，有效提高 U Bike 滿車站點出租率，提升空車站點還車率，並預測未來車輛數目。

「UBike」清新簡潔的使用者介面也是成功要素之一，圖 2-24 中代表站點的水滴符號佈滿地圖，綠色填得愈滿表示車子數量充足，反之水滴愈空就代表車子所剩無幾；另外也可以讓使用者紀錄騎乘的時間，並提供天氣、總騎乘人數等資訊。該 App 內建留言與社群廣播系統，蒐集騎友的使用資料，讓民眾可藉「UBike」取得路況回報與失物招領等資訊。如此創新的設計使此款 App 短時間內累計下載次數突破 16 萬人次，未來將與藝文單位配合提供 LBS 展覽與活動資訊刊登，發展更廣泛的市民應用服務。



圖 2-24 U Bike App 使用介面示意圖

2.美國波士頓「Hubway」公共自行車地圖

從圖 2-25 可見其地圖設計理念與臺北設計方式相近，惟目前當局僅提供網頁作為查詢環境。波士頓的 Hubway 單車共騎系統曾公佈大量的使用者騎乘歷史檔案，然後邀請單車愛好者和開發者參與共同做視覺化資訊的應用分析，最後得出各條單車路線的平均速率，供通勤族瞭解各路線之騎乘順暢度，如圖 2-26 所示。

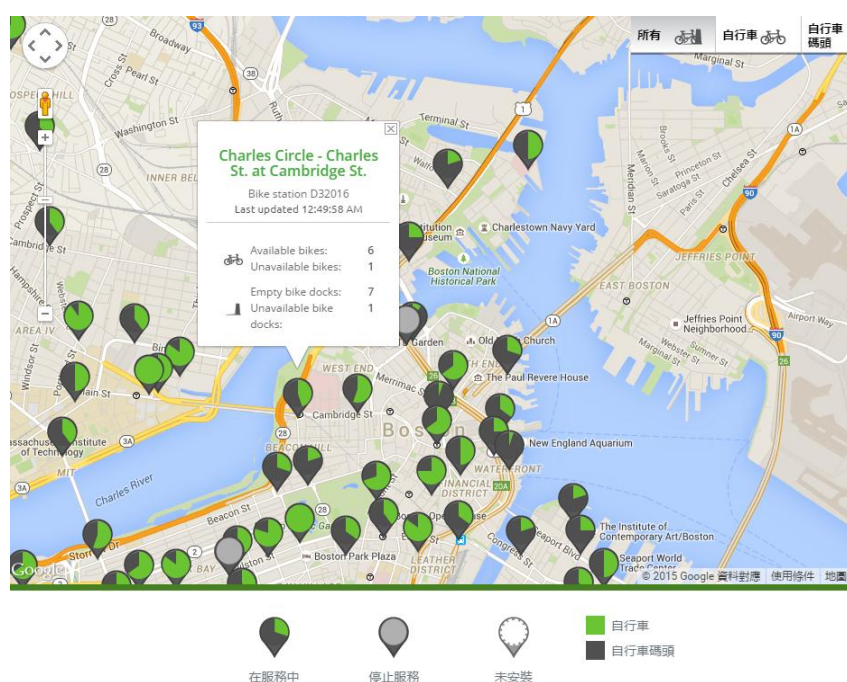


圖 2-25 Hubway 公共自行車地圖

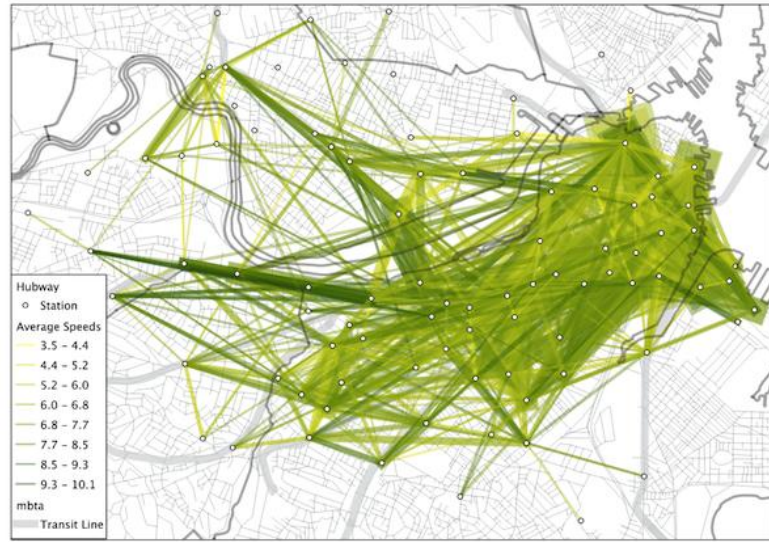


圖 2-26 Hubway 公共自行車視覺分析應用之單車速率圖

3. 德國「DB Zugradar (train radar)」公共軌道運輸系統視覺化 App

德國作為開放數據城市(Open Data City)所推動的公共運輸資訊視覺化項目，其目的是透過視覺化將民眾日常通勤所必需的旅運資訊以直觀易讀的圖形加以呈現。該地圖中呈現德國特定區域軌道系統的營運狀態，標示出軌道路線、列車運行方向(列車圖示的紅色箭頭)，以及各班次的延誤情形(以正負號數字表示)，若點選該運具圖示會跳出詳細的列車車次與停靠站點時刻表，如圖 2-27 所示。目前該局除火車動態雷達地圖外，也發展了共乘系統 App、公共自行車租賃系統 App，以及以手機為載體的電子票證程式，詳細操作畫面如圖 2-28~圖 2-29 所示。

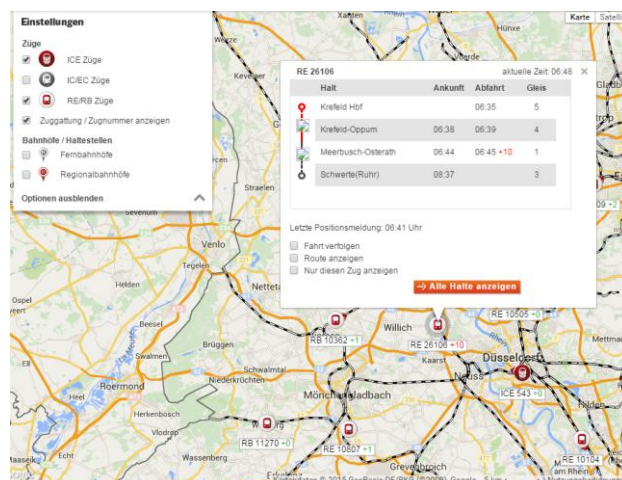


圖 2-27 德國 DB Zugradar 軌道系統動態移動與時刻表地圖

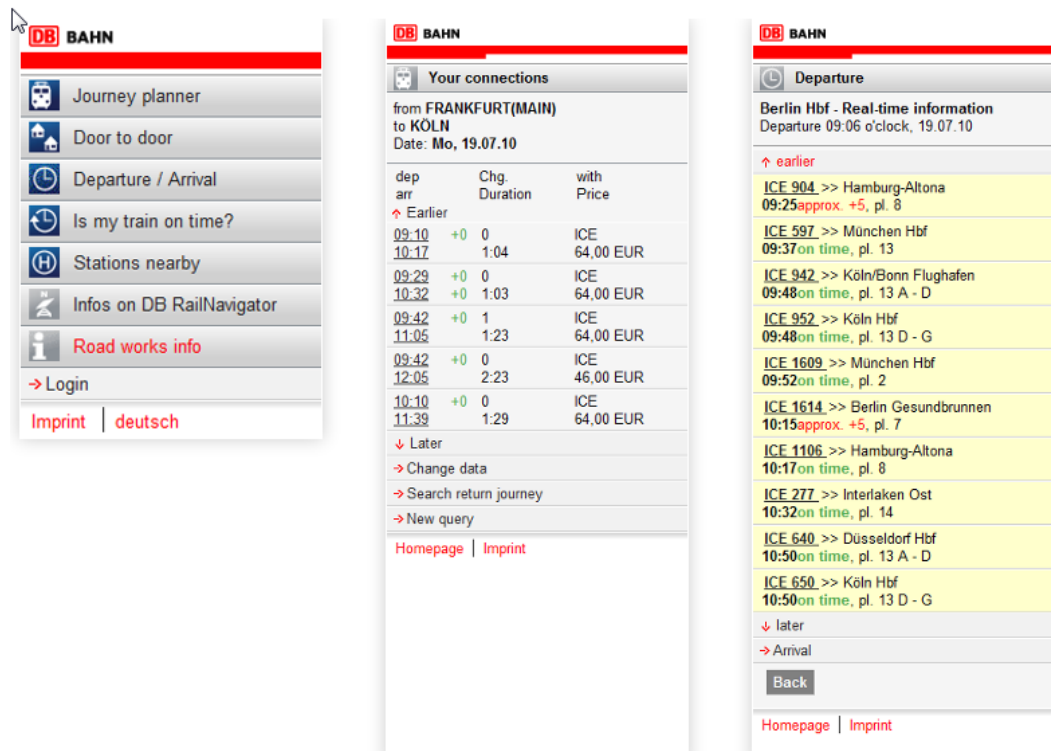


圖 2-28 德國 DB Zugradar 手機電子票證與相關服務畫面



圖 2-29 德國 DB Zugradar 共乘與公共自行車租賃操作畫面

4. 「TRAVIC」公共運輸路網活動繪製平臺

TRAVIC(Transit Visualization Client)是由德國佛萊堡大學所開發的視覺化工具，能繪製各種公共運輸的旅運模式。該平臺利用不同公共運輸固定的旅運時刻表與即時的延滯時間資訊，計算運具的位置並呈現在地圖上。該平臺目前已彙整來自全球五大洲，249 個公共運輸

路網圖整理在同一個動畫地圖上。TRAVIC 透過提供公共運輸營運網路圖，以分析不同運輸系統中運具運作的型態與模式。

以澳洲墨爾本市中心晨間通勤尖峰時段來觀察，如圖 2-30 所示，圖中紅點代表地鐵列車，黃點代表輕軌列車，而藍點則代表公車。即可發現輕軌系統的獨特性，因為墨爾本擁有全球最大的都市輕軌網路，可看出其輕軌系統高度與人行道區域相鄰，顯示出其不同公共運輸系統在轉乘營運的共同依賴性。

在 TRAVIC 圖層中也可根據過濾後的班次表及運具路線編號，產出單一區域或全市的路網圖，圖 2-31 顯示瑞士蘇黎世全區的公車路網圖，每一個對應的區塊都會包含多種的運具，因此可以依據使用者的需求篩選出特定時間經過特定區域的路線圖，作為旅運者轉乘時的參考依據。

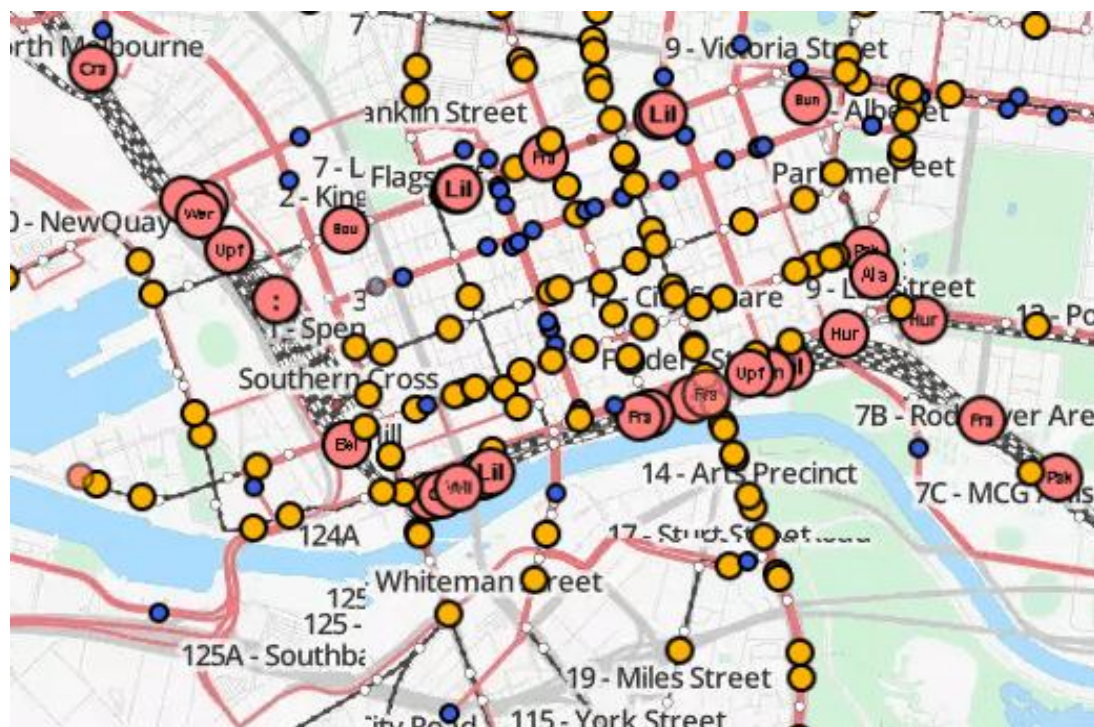


圖 2-30 TRAVIC 平臺所呈現之澳洲墨爾本公共運輸營運網路圖

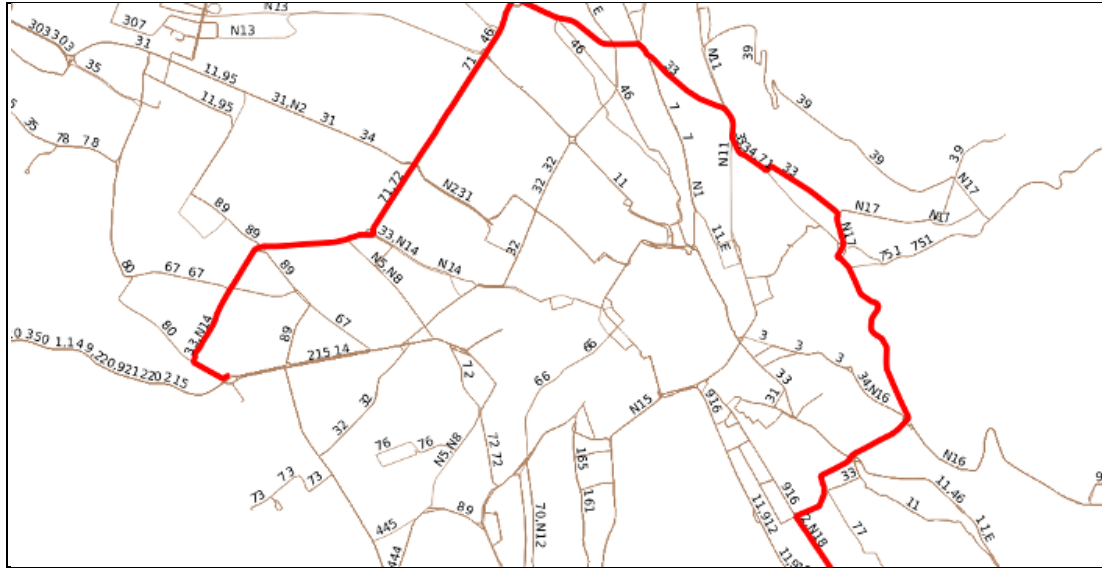


圖 2-31 TRAVIC 平臺所呈現之瑞士蘇黎世全區之公車路網圖

若使用者在該動態地圖上點選特定運具，則會顯示完整的路線與對應的抵達時間及離開時間(包含產生的延滯時間)，如圖 2-32 所示。TRAVIC 平臺上產生的圖層並未受限於地理的空間輸入資料(開放街道地圖)，每一種資料來源皆可以轉換成可被利用的圖層，例如 GIS 地理矩陣圖。該平臺未來將著重在行動裝置的動態資訊地圖之研發。

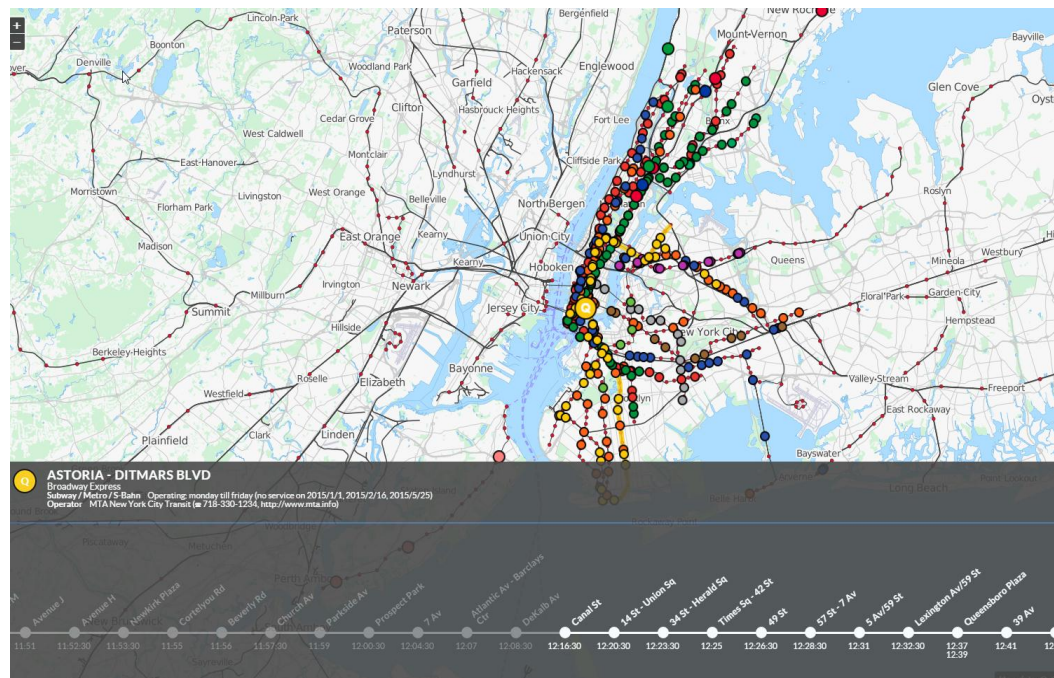


圖 2-32 TRAVIC 呈現美國紐約公共運輸營陸路網與對應班次圖

5.瑞士 Swissnex 都市資料競賽計畫(Urban Data Challenge)

公共運輸轉乘績效一般是由速度與效率來衡量，2013 年「TRANSIT QUALITY AND EQUITY」參賽作品藉由繪製轉乘密度資料結合不同區域的區民收入水準資料，以視覺化方式探討轉乘服務品質的價值影響性，由圖來分析赤貧地區是否享有同等品質的公車服務水準，此案例也可視為其他地方政府衡量其轉乘效益的借鏡。如圖 2-33 所示，紅點大小代表公車班次服務頻率、紅色深淺代表延誤時間長短(0-15 分鐘以上)、綠色方塊顏色深淺代表收入水準高低(5% - >25%居民水準)，若點選行政區可得知人種分佈比例表，右方的公車路線亦可經選擇後呈現在地圖上。

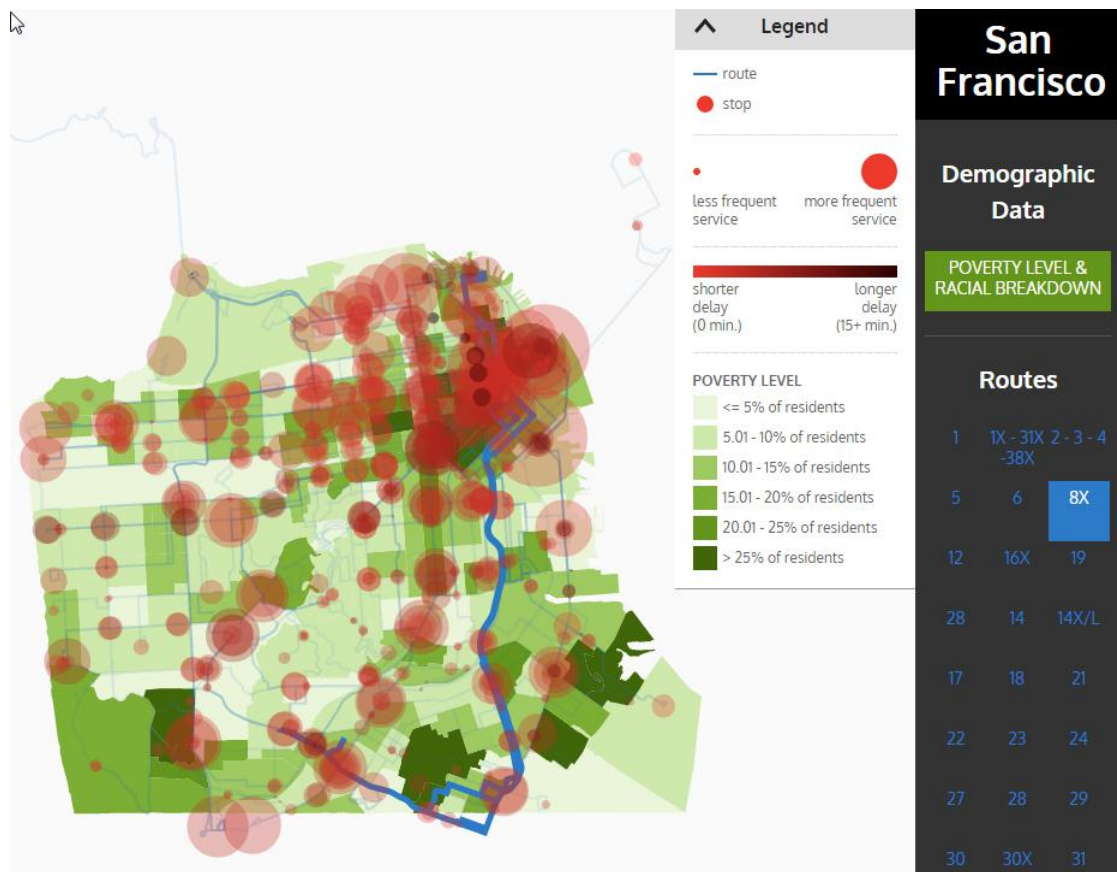


圖 2-33 TRANSIT QUALITY AND EQUITY 舊金山地區轉乘績效圖

第三章 需求彙整

為使分析項目符合相關單位之需求，必須對可能使用者進行需求座談與彙整，考量不同使用者的需求可能略有差異，本研究將使用者區分為(1)公共運輸主管單位與(2)客運業者分別召開需求座談會。在需求座談會舉辦前，透過工作會議的共識，初擬主管機關與業者可能的需求項目並製作相關的簡報（如附件 A），在座談會舉辦前分送受邀單位參考，座談會當天各單位參與之專業人員齊聚一堂，聽取主持人簡報後，透過集思廣益、腦力激盪的方式，討論 APTS 巨量資料分析及視覺化呈現的應用需求清單，本章節將分別說明辦理情形與彙整結果。

3.1 主管機關需求座談辦理情形

1. 辦理時間：104 年 05 月 08 日(星期五)上午 09:30~12:00
2. 辦理地點：交通大學臺北校區 3F 視聽教室(忠孝西路一段 118 號)
3. 會議議程：如表 3.1。

表 3.1 主管機關需求座談會議程表

時間	議程	主講者/與談者
09:30~10:00	報到	—
10:00~10:10	致詞	運研所運資組陳其華組長 交通大學王晉元教授
10:10~10:30	會議資料簡報	交通大學王晉元教授
10:30~12:00	腦力激盪/綜合座談	交通大學王晉元教授 中興工程 張恩輔 研究員
12:00	賦歸	

4. 邀請對象：共計 15 個單位如表 3.2。

表 3.2 主管機關邀請清單

邀請單位		
交通部公路總局	高雄市區監理所	彰化縣政府工務處
交通部路政司	高雄區監理所	臺中市政府交通局
交通部運輸研究所	基隆市府交通旅遊處	臺中區監理所
宜蘭縣政府建設處	新北市政府交通局	臺北市府交通局
花蓮縣政府建設處	新竹市政府交通處	臺北市區監理所
屏東縣府城鄉發展處	新竹區監理所	臺北區監理所
桃園市政府交通局	嘉義市府交通觀光處	臺東縣府觀光旅遊處
高雄市政府交通局	嘉義區監理所	臺南市政府交通局

註：依筆劃排序

5. 出席對象：簽到如表 3.3。

表 3.3 主管機關需求座談會簽到表

簽到單		
「公車動態系統巨量資料分析應用」 公路客運主管機關 需求座談會		
會議時間	104 年 5 月 8 日(星期五)上午 10 時 00 分	
會議地點	交通大學台北校區視聽教室	
會議主持人	交通部運輸研究所運輸資訊組 陳其華 組長 交通大學運輸與物流管理學系 王晉元 教授	
會議紀錄	王佳祺、蔡佳倫	
出席人員 (請簽名)	交通部路政司	
	交通部運輸研究所	陳其華
	交通部公路總局	潘其財 王章德
	臺北區監理所	謝其行
	新竹區監理所	吳成民
	臺中區監理所	陳美芳
	嘉義區監理所	
	高雄區監理所	李瑞奎
	臺北市區監理所	陳健成
	高雄市區監理所	林文雄
	臺北市政府交通局	王宗
	高雄市政府交通局	黃育庭
	基隆市政府交通旅遊處	
	新北市政府交通局	蘇育宜
	桃園市政府交通局	吳慧瑜
	新竹市政府交通處	
	臺中市政府交通局	吳凱 林漢編
	彰化縣政府工務處	
	嘉義市政府交通觀光處	
	臺南市政府交通局	
	宜蘭縣政府建設處	
	屏東縣政府城鄉發展處	
	花蓮縣政府建設處	
	臺東縣政府觀光旅遊處	
	交通大學研究團隊	王晉元 謝其華 張恩輔 李其仁

6. 活動剪影：如圖 3-1。



圖 3-1 主管機關需求座談會實況剪影

7. 需求座談摘要：

- (1) 利用巨量資料分析掌握連續假期運量，同時將公路客運車輛作為為流速探偵車。
- (2) 駕駛員工時與肇事不見得直接相關，過去曾經發生數次肇事司機在事發前均是休假狀態，此資訊提供團隊作分析參考。

- (3) 過去許多分析與管理都是以彙總資料或抽樣資料處理，如今已有詳細資料（幾近普查），分析與管理手段均可改變。
- (4) 除駕駛員管理（急加速/急減速）外，鐵公路轉乘規劃、補貼路線等議題均可考慮引入巨量分析應用。
- (5) 桃竹苗有 100 多條虧損補貼路線，希望巨量分析可讓虧損補貼的計算更精確且貼近真實。
- (6) 告知旅客即時路況甚至提供改道建議是未來更遠大理想目標。
- (7) 公部門除了提供 Data 外，API 的提供更為重要，可以加速產業的發展。
- (8) 若能以電信資料分析民眾在觀光景點的停留時間，將有助於觀光公車之排班。
- (9) 旅客查詢公車動態的查詢紀錄檔也可能蘊藏資料探勘價值。
- (10) 公車動態的資訊品質應提昇，包括預估時間是否準確？車機如何更可靠？準確率的提昇可以利用路段的平均行車時間處理，車機可靠度則必須找出可靠度低的時間與空間資訊，作為責成廠商改善的參考。
- (11) 過站不停、誤點脫班均是主管機關欲稽查的樣態。
- (12) 許多應用需求均須得知即時乘載率，相較於即時影像等技術，電子票證配合上下車均刷卡是較為經濟可行的方案，可作為未來公車動態系統的提昇目標。
- (13) 本計畫著重於歷史巨量資料之分析，而非即時少量資料的回應與傳遞。

3.2 客運業者需求座談辦理情形

1. 辦理時間：104 年 05 月 08 日(星期五)下午 14:00~16:00
2. 辦理地點：交通大學臺北校區 3F 視聽教室(忠孝西路一段 118 號)
3. 會議議程：如表 3.4。

表 3.4 客運業者需求座談會議程表

時間	議程	主講者/與談者
13:30~14:00	報到	—
14:00~14:10	致詞	運研所運資組陳其華組長 交通大學王晉元教授
14:10~14:30	會議資料簡報	交通大學王晉元教授
14:30~16:00	腦力激盪/綜合座談	交通大學王晉元教授 中興工程 張恩輔 研究員
16:00~	賦歸	

4. 邀請對象：共計 15 個單位如表 3.5。

表 3.5 客運業者邀請清單

邀請單位	
中華民國公共汽車客運商業同業公會全國聯合會	高雄汽車客運股份有限公司
	國光汽車客運股份有限公司
和欣汽車客運股份有限公司	統聯汽車客運股份有限公司
阿羅哈汽車客運股份有限公司	港都汽車客運股份有限公司
南臺灣汽車客運股份有限公司	新竹汽車客運股份有限公司
屏東汽車客運股份有限公司	鼎東汽車客運股份有限公司
首都客運股份有限公司	嘉義汽車客運股份有限公司
桃園汽車客運股份有限公司	豐原汽車客運股份有限公司

註：依筆劃排序

5. 出席對象：簽到如表 3.6。

表 3.6 客運業者需求座談會簽到表

簽到單		
「公車動態系統巨量資料分析應用」 公路客運業者 需求座談會		
會議時間	104年5月8日(星期五)下午14時00分	
會議地點	交通大學台北校區視聽教室	
會議 主持人	交通部運輸研究所運輸資訊組 陳其華 組長 交通大學運輸與物流管理學系 王晉元 教授	
會議紀錄	王佳祺、蔡佳倫	
出席人員 (請簽名)	中華民國公共汽車客運商業同業公會全國聯合會	
	國光汽車客運股份有限公司	王中
	統聯汽車客運股份有限公司	楊師宏 蕭信忠
	和欣汽車客運股份有限公司	
	阿羅哈汽車客運股份有限公司	
	首都客運股份有限公司	張淑文 蕭恩誠
	桃園汽車客運股份有限公司	林信介 何志華
	新竹汽車客運股份有限公司	連偉盛 林振民
	豐原汽車客運股份有限公司	黃朝勝 薛乙樓
嘉義汽車客運股份有限公司		

6. 需求座談摘要：

- (1) 「油耗分析」是很重要的分析應用，希望得知車隊或駕駛的耗油量是否在平均水準之上，若低於平均過多會予以檢討。一旦能作到油耗的節省，通常亦可節省輪胎成本。
- (2) 「潛在旅運需求」是業者最想從巨量資料分析得知的訊息，惟電子票證紀錄僅能提供已是公共運輸用戶的使用行為，未來仍有賴手機定位資料方可分析潛在旅客的需求潛力。
- (3) 站牌班表在沒有專用路權的公車系統不容易準點，此現象在公路客運路線尚不嚴重，但在市區客運極難掌控，特別是某些紅燈時間長達 90~140 秒，常引起民怨，甚至造成駕駛員分心導致事故率增加。
- (4) 週末/週間、晴/雨、偶發事件（例如廟會封路/重大事故）等均可能大幅影響行車時間，造成站牌班表準點率低，進而引

發民怨，建議利用巨量資料予以分析佐證，挑選出行車時間變異較小之路線辦理較為合適。

- (5) 數業者被要求延長營運時間，期待巨量資料分析可提供參考資訊。
- (6) 客訴資訊也建議納入分析，甚至包括社群網路上的正評或負評均是未來分析的方向。
- (7) 若能從手機定位來偵測乘客願意的步行距離，可作為未來業者設站之參考。
- (8) 部份行經國道的公路客運由於受限於國道上不可站立之規定，會導致上國道前滿載無法載客，加上班距較大，容易造成民怨，希望巨量資料分析能達到需求預測之目的。
- (9) 車距偵測器、車道偏離偵測器 胎壓胎溫偵測等均是駕駛安全分析可採用的設備與分析資料。

3.3 需求彙整結果

根據第 3.1 節與第 3.2 節之摘要紀錄，可彙整需求如表 3.7 與表 3.8。

表 3.7 主管機關需求彙整表

編號	需求簡述
A01	將公路客運車輛作為流速探偵車
A02	欲瞭解危險駕駛司機之特性
A03	駕駛員管理（急加速/急減速）
A04	分析民眾在觀光景點的停留時間，將有助於觀光公車之排班
A05	旅客的查詢紀錄檔也可能蘊藏資料探勘的價值
A06	瞭解預估時間是否準確
A07	探勘車機可靠度低之時間與空間，提供改善之參考
A08	擷取誤點脫班樣態

表 3.8 客運業者需求彙整表

編號	需求簡述
B01	分析油耗水準作為車隊或駕駛檢討之依據
B02	利用電信公司定位資訊挖掘潛在旅運需求
B03	分析適合(或不適合)採行站站班表的主客觀條件
B04	延長營業時間之潛在需求分析
B05	客訴資料分析
B06	社群網路上之正面/負面評價資訊
B07	車距/行車偏移/胎壓胎溫等偵測器之設備紀錄資料分析

經綜合討論與評估後，表 3.7 與表 3.8 之需求處理方式彙整如表 3.9 與表 3.10。

表 3.9 主管機關需求處理方式彙整表

編號	於本研究初探	納入未來研究方向(原因)
A01	行車時間分析 (CH4.4)	—
A02	駕駛風險分析 (CH4.9)	—
A03	駕駛風險分析 (CH4.9)	—
A04	—	需取得電信定位資料
A05	—	須取得歷史查詢紀錄檔
A06	行車時間分析 (CH4.4)	—
A07	車機穩定度分析 (CH4.10)	—
A08	班表分析 (CH4.6)	—

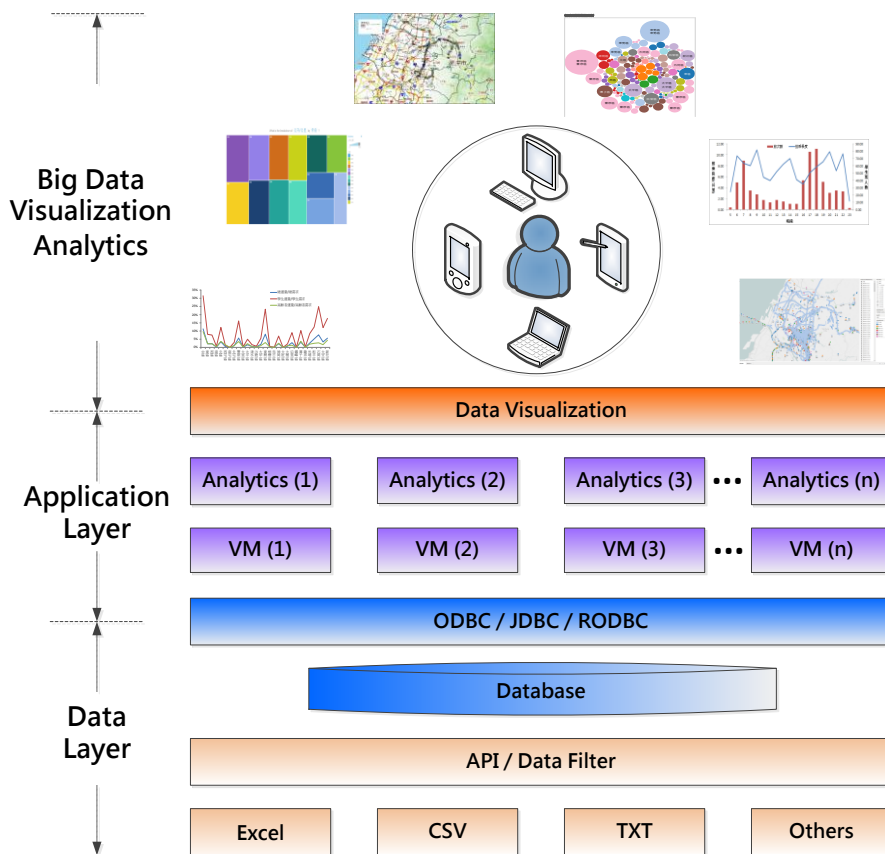
表 3.10 客運業者需求處理方式彙整表

編號	於本研究初探	納入未來研究方向(原因)
B01	駕駛員油耗分析 (CH4.8)	—
B02	—	需取得電信定位資料
B03	站牌班表產製 (CH4.5)	—
B04	—	需取得電信定位資料
B05	視資料取得狀況而定	—
B06	—	需較長時間研究
B07	—	需取得設備介接資料

第四章 系統架構規劃與工具

4.1 公共運輸巨量資料分析系統

由於從各單位所收集到的相關動態與靜態之巨量資料(Big Data)具有各自的格式與特性，為了能夠透過資料的探勘、比對與分析，建置視覺化分析系統整合平臺，建議可透過以下的三層架構來加以進行：資料層（Data Layer）、應用層（Application Layer）、與巨量資料視覺化分析層（Big Data Visualization Analytics）。本架構可彈性擴充系統與提升視覺化分析效能，整體系統架構示意圖如圖 4-1 所示：



資料來源：本研究自行整理與繪製

圖 4-1 巨量資料視覺化分析系統整合平臺架構示意圖

巨量資料視覺化分析系統整合平臺功能說明：

1. Data Layer

蒐集各相關單位的資料，諸如公車行駛路線、班表、油耗、維修資料、駕駛資訊與電子票證等各項原始資料，經由資料介接系統可讀取各種不同之資料格式，例如：Excel、CSV 與 Text 等，將資料進行過濾與整理後，再匯入資料庫平臺，提供視覺化分析應用系統之開發。

2. Application Layer

相關的應用系統開發透過 ODBC/ JDBC/RODBC 等介面與資料庫介接，存取各項所需資料，透過巨量資料探勘、演算法等相關技術分析研究，建置各項應用系統分析主題用之虛擬化整合平臺，並可依視覺化分析主題之需求，彈性擴充虛擬化平臺，提供多樣性巨量資料分析。

3. Big Data Visualization Analytics

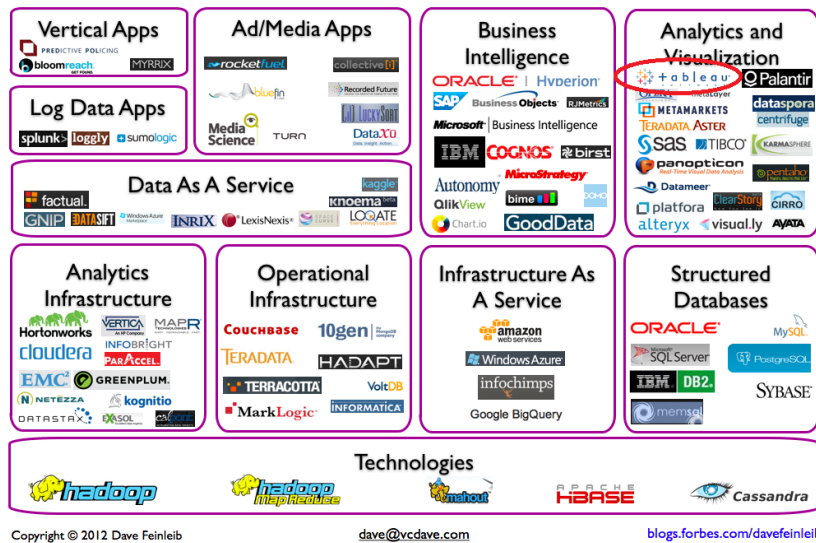
透過視覺化分析系統整合平臺，將各項分析主題，利用視覺化分析工具將複雜之巨量資料(Big Data)轉換成易讀之圖表呈現，進行公共運輸供給與需求在時間與空間維度上之相關應用研究，提供交通主管機關決策參考資訊、提升客運業者整體服務品質與未來民眾生活加值應用服務。

4.2 相關分析工具

4.2.1 Tableau 簡介

由圖 4-2 的架構圖可知，Tableau 在整個大數據分析中係屬於「分析暨視覺化呈現」類的商用軟體之一，雖然相類似軟體眾多，但若以圖 4-3 的魔力象限（Magic Quadrant）觀點評估，Tableau 屬於領導者地位。

Big Data Landscape



資料來源：<http://www.forbes.com/sites/davefeinleib>

圖 4-2 Tableau 在巨量資料領域之定位



資料來源：<http://www.informationweek.com/big-data/big-data-analytics/gartner-bi-magic-quadrant-2015-spots-market-turmoil/d/d-id/1319214>

圖 4-3 Tableau 在 BI 領域的魔力象限（Magic Quadrant）定位

Tableau 可開啟數十種資料來源如圖 4-4，舉凡本機端的 Office 家族、伺服器端各種商業版或開源版資料庫、新興的 Hadoop 系列，商業統計軟體檔案均在支援之列。



資料來源：<http://www.tableau.com/> 暨本研究截圖後製

圖 4-4 Tableau 支援之 Data Source 清單圖

Tableau 支援 10 餘種資料視覺化圖表如圖 4-5，同時可結合多圖表製作為儀表板，並設定圖表間資料的互動查詢，有助於使用者利用「看」圖表來完成資料分析作業，進而研擬對策。

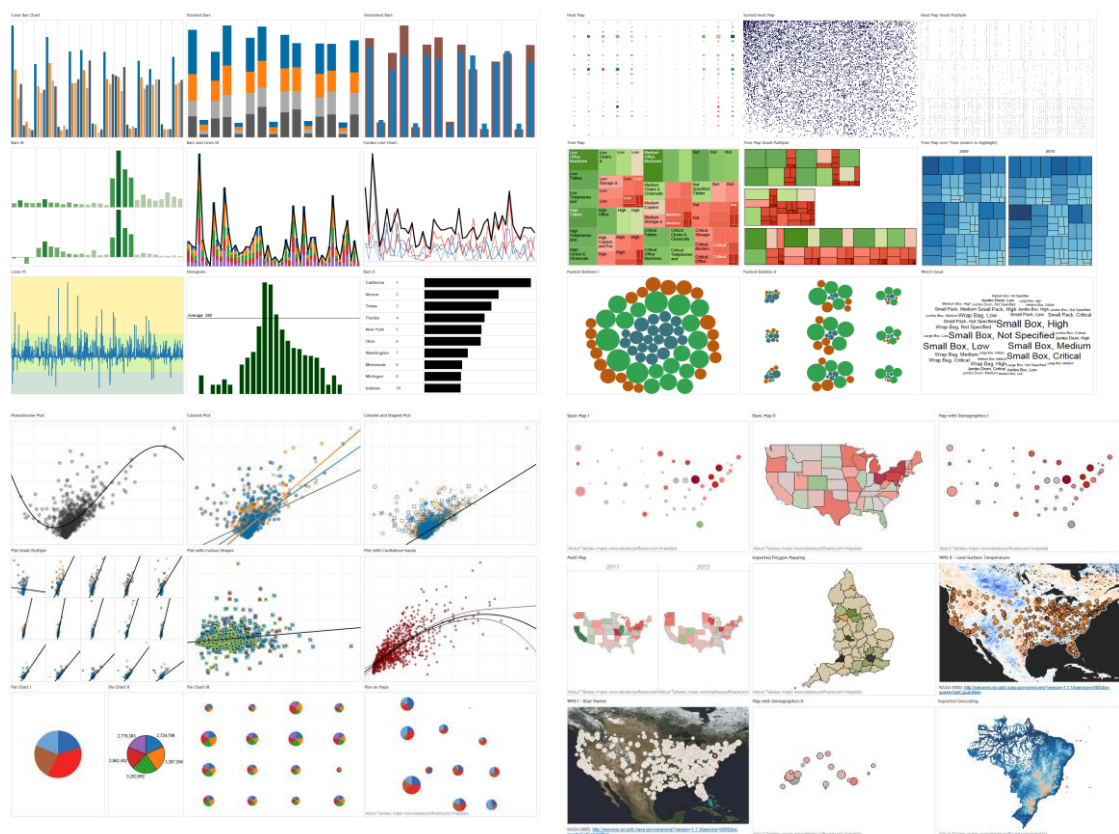


圖 4-5 Tableau 視覺化呈現軟體支援之圖表列舉

4.2.2 R 語言簡介

R 語言是一種基於 S 語言所發展出具備統計分析、繪圖與資料視覺化的程式語言。1976 年 AT&T 貝爾實驗室 John Chambers, Rick Becker, and Allan Wilks 三位博士共同研發 S 語言。R 語言是來自紐西蘭奧克蘭大學的 Ross Ihaka 和 Robert Gentleman 開發(也因此稱為 R)，1995 年，兩人決定以開放源碼方式發行。直至，1997 年 R 語言的核心開發團隊(R development core team)成立，專責 R 原始碼的修改與編寫。

R 語言是基於 S 語言的一個 GNU 計劃專案，所以也可以當作 S 語言的一種實現，通常用 S 語言編寫的代碼都可以不作修改的在 R 語言環境下執行。2013 年 KDnuggets 的調查顯示 R 語言高居資料解析／資料探勘／資料科學領域中編程／統計語言的第一位(60.9%)，其後分別是 Python(38.8%)，SQL(36.6%)，SAS(20.8%)，Java(16.5%) 與 MATLAB(12.5%)，除了 Python 遞增與 SQL 持平外，

其餘工具近年百分比均下滑。R 軟體橫跨 33 個領域，目前擁有套件數已由 2012 年的 3636 個成長到 2014 年的 5990 個。

整體而言，R 語言優勢可以彙整如以下 6 點：

- (1) 透過開放源碼的方式可以運用不同領域專家的知識共同開發，而且根據此方法所開發出的軟體更具彈性及可信度。
- (2) 擁有各種不同繪圖套件，透過多種圖形表達資料，即有強大的資料視覺化功能。
- (3) 除原生界面外，包含 Rstudio、Timm-R 等人性化的整合開發環境界面。
- (4) 支援跨平臺功能，可以在許多平臺上運作，像常見的 Windows、Mac OS 等都可以使用。
- (5) 透過 R 軟體可以根據需要分析的內容自行撰寫程式，且可與其它程式語言連結，具有高度相容性。且有許多熱心的支持者更新或增加新的套件，來擴充 R 軟體的功能。
- (6) 支援多種程式語言，其非封閉環境，可以呼叫 C 語言、Python、Java 等所撰寫的外部程式庫來輔助分析。

但沒有任何一套程式語言是完美的，一般坊間認為 R 語言有以下 3 點缺點待未來改進：

- (1) 較少商業性支援，須大量透過對此軟體的愛好使用者，針對各自所需的功能進行套件開發。
- (2) 承上之開發背景，各套件維護/升級較不一致，故在 R 主程式升級後，需要一段時間各套件方會陸續升級，常須處理套件間的版本相依性問題。
- (3) 原生程式未支援多工作業，須搭配套件方能在多 CPU 環境充分所有的計算資源。

4.2.3 IBM Watson 簡介

Watson Analytics 是結合 Watson 的機器學習以及旗下 SPSS、Cognos 等進階分析產品而成，適合於行銷、銷售、營運、金融和人力資源等資料密集型應用領域，由於 Watson Analytics 使用自然語言進行預測性分析和互動，可以為企業即時提供預測和視覺化分析工具。

Watson Analytics 和一般為資料科學家和分析人員設計的分析工具不同，它運用自動化步驟，帶使用者走過將所有資料輸入及分析過程，包括資料上傳與清洗、預測式分析、視覺式故事敘述等，另外能理解企業中習慣的對話方式所提出的問題，使用者獲取結果後，可以持續針對問題和資料加以微調，進而挖掘出相關度最高的事實，以及原本看不到的各種模式和關係，以協助決策。

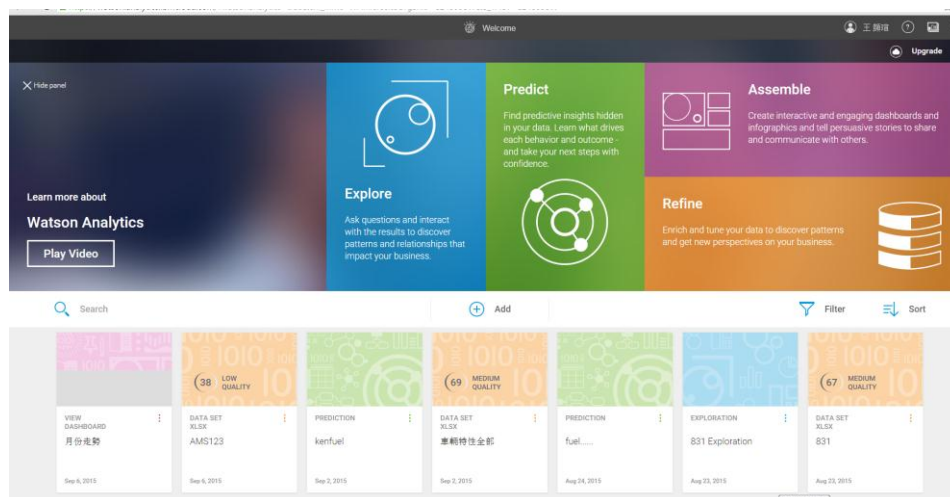


圖 4-6 IBM Watson 介面

第五章 資料介接與處理

5.1 資料介接清單

本案資料介接欄位共計 38 項目，共分為六大類，資料期間除公路客運之公車動態資料依其原資料保留規定為三個月，使得資料取得不足一年外，其餘資料期間皆為 103 年 5 月至 104 年 4 月，總筆數為 3,845,370,994 筆，資料空間超 363GB，其資料類型說明如下：

1. 電子票證資料：第一來源為豐原客運之電子票證資料，包含該公司所有電子票證公司之資料。其二為悠遊卡公司之資料，其為悠遊卡公司提供之不分客運公司之電子票證資料
2. 公車動態資料：因客運類型不同使得資料保存單位不相同，因此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；市區公車則透過臺中市政府的公車動態資訊系統取得。
3. 數位式行車記錄器資料：豐原客運公車因車齡不同，使得車上設備有傳統行車記錄器資料及數位式行車記錄器資料，因傳統行車記錄器判讀與存檔不易，因此本案蒐集之資料皆為由豐原客運所提供的數位行車記錄器。
4. 營運相關資料：包含有豐原客運之仕業、路線、站位、駕駛員、機務、油料、獎懲、維修及電動公車等資料。
5. 其他運輸相關資料：事故、計程車接單資料為本案之其他運輸相關資料。事故資料僅限於臺中市的 A1/A2 資料，計程車僅限於臺中市某家業者所提供的資料。
6. 其他非運輸相關資料：諸如氣候（包含雨量）等非運輸相關資料。

表 5.1 收集之巨量資料彙總表

編號	資料表中文名稱	資料名稱	容量(KB)	筆數
1	車機定時回傳資料	A1	21,590,424	235,107,005
2	車機定點回傳資料	A2	4,937,560	64,161,079
3	事故資料	accident	12,168	97,668
4	業者對照表	bus_company	24	21
5	公路客運歷史行車資料	bus_history_route	49,922,792	355,076,612
6	每日營運報表	DAILY_REPORT	2,184	32,414
7	電動公車充電資料	e_bus_avg_power	32	190
8	電動公車維修資料	e_bus_main_parts_fail	32	124
9	悠遊卡業者路線資料	Easycard_Provider_Route	64	856
10	悠遊卡電子票證資料	EasyCard-Taichung	13,711,304	97,263,994
11	悠遊卡票卡類型說明	Easycard_Ticket	24	10
12	豐原客運電子票證站牌對照	ETSTOPVILLAGE	1,752	9,346
13	油耗	FUEL	1,448	48,065
14	勤務資料表	fy_driver_roster	84,872	750,690
15	豐原客運電子票證資料	mcard	4,296,240	16,762,909
16	歷史預估到站時間	N1	19,988,832	311,500,390
17	行車記錄器	OBU	249,100,000	2,763,749,375
18	獎懲記錄	RECORDS_OF_BEHAVIOUR	40	174
19	路線一覽表	ROUTE	24	99
20	薪資明細	SALARY	256	3,518
21	臺中市區公車業者資料	TaichungPROVIDER	24	32
22	臺中市區公車業者資料	TaichungROUTE	40	252
23	臺中市區公車班表資料	TaichungSCHEDULE	288	7,851
24	臺中市區公車站牌資料	TaichungSTOP	1,560	21,256
25	臺中市區公車站牌資料	TaichungSTOP	1,576	21,256
26	公路客運業者資料	thbProvider	16	66
27	公路客運路線資料	thbRoute	856	2,346
28	公路客運班表名稱	thbSchedule	153,736	210,596
29	公路客運站牌資料	THBStop	10,440	79,140
30	公路客運票價資料	THBTicket	61,688	388,936
31	計程車訂單資料	TOrder	1,008	7,429
32	請假明細	VACATIONS	1,328	20,546
33	車輛清冊	VEHICLE	40	272
34	車輛維修	VEHICLE_MAINTENANCE	352	5,371
35	車輛零件統計表	VEHICLE_PARTS	64	351
36	天氣地區資料	weather_area	16	2
37	天氣地區資料	weather_city	16	1
38	天氣資料	weather_data	6,544	40,752
—	小計	—	363,889,664	3,845,370,994

對於所收集資料之介接來源與方式又可分為線上資料交換與檔案匯入，以下分別說明：

1. 線上資料交換：對於格式化資料，可透網路方式下載，利用演算邏輯判讀資料內容與型態，並直接匯入資料庫，其優點為節省人力之使用。可透過線上資料交換之資料主要有公車路線、班表、站牌、歷史行車資料等。
2. 檔案匯入方式：以 CSV 檔、二元檔、LOG 檔及 DBF 檔做為資料檔，利用自動資料庫匯入功能或自行撰寫資料匯入程式，將資料讀出後再匯入至資料庫中，其資料格式又可分為格式化資料及非格式化資料。
 - 格式化資料：在檔案內已有固定之資料欄位與長度，使得讀取資料時能直接利用資料庫匯入功能或自行撰寫資料匯入程式資料匯入至資料庫中。可透過此方法匯入之資料包含有歷史公車動態資料、電子票證資料、行車記錄器資料。
 - 非格式化資料：需利用人工方式資料整理至特定格式與內容，使得資料可以再利用資料庫匯入功能或自行撰寫資料匯入程式將資料匯入至系統中；如豐原客運提供之管理報告，因資料表內包含有其報表名稱、匯入日期及人員簽署欄位，皆不屬於分析項目，而且因其格式不一，因此需先透過人工處理後才能匯入資料庫內。

所有資料整理後之資料欄位與說明因表格眾多，因此相關之資料表說明請參閱附錄 C。

5.2 電子票證資料處理與課題

電子票證資料主要可用來分析地區性旅次需求資料及乘客搭乘類型，其資料分析特性包含有空間分析、時間分析及資料屬性分析，而在豐原客運、悠遊卡公車提供票證資料僅包含時間與資料屬性，因

此需將資料整理並加入空間資料後，才可開始後續資料分析課題。悠遊卡公司與豐原客運公司提供之電子票證資料為具格式化之 DBF 與 CSV 檔，因此在匯入過程中並無格式不符之問題。

惟在資料分析前處理及資料分析過程中，由於資料量大，無法預期將會有多少可能的潛在問題而無法進行資料分析。因此團隊採用模組化資料處理流程，針對不同分析主題所遇到的問題，以先分析再檢驗，最後處理問題的三階段處理方式，進行系統化之資料處理模式。

1. 分析資料：依運研所蘇昭銘等人於 2014 年運輸研究所「先進公共運輸系統整合資料庫加值應用系統維運及推廣計畫」研究案提出電子票證分析目的及流程，先進行資料分析，如路線分析即個別針對資料中之路線欄位進行不同資料區間之統計，以瞭解各路線之搭乘人數特性。
2. 檢驗資料：資料統計後，即可將統計結果進行比對，如前述路線分析後，即可針對路線數量、名稱、發車時間等資訊進行驗證，以比較與現有公車動態資料之差異，並標示無法比對或資料異常之資料。
3. 處理資料：針對標示為異常之資料尋找可能之處理方式，如路線無法與公車動態資料比對時，即可能有資料規則性差異、資料不規則性差異、及完全無比對資料等三種情況，資料規則性差異如路線代碼或名稱有前後資料固定差異、即可使用程式化方式消除其差異值。資料不規則性差異如路線名稱完全無法比對，即需要逐項處理以消除資料差異。最後則因無比對資料而需要其它資料，如歷史電子票證資料因路線調整或停駛而無法以現有公車資料比對其路線資訊，而公車動態資訊提供者也無法提此資料時，即需利用團隊歷年蒐集之公車資料進行比對以找出其相問題之路線資料。

資料處理過程因原始資料與輔助資料匹配性問題而衍生數個資料處理問題：

1. 上下車人數不同：在資料完整性分析過程中，利用電子票證上下車資料統計即時車上人數時發現可能因資料遺漏或不足，使得在同一班次公車資料中，即時車上人數出現負值（下車總人數大於上車總人數），因此後續在分析載客數時將無法獲得正確之結果。
2. 欄位資料內容缺漏或錯誤：在票證資料中，部份記錄中其無上車站序或下車站序，使得後續資料比對空間座標時，並無法利用路線、上下車站序判斷乘客上下車。
3. 空間資料無法比對：由於電子票證資料並無上下車座標資料，因此需另外再比對其輔助資料才可得其乘車上下車位置。由於本案分析資料期間為臺中市區公車因 BRT 上線使用而造成路線站牌變異過大，使得豐原客運提供之電子票證資料與路線座標資料，會因為站牌路線、座標調整或支線行駛等因素，使其無法比對所有之電子票證乘客上下車資料；而悠遊卡公司提供之比對路線資料中並無站牌座標，並且悠遊卡公司提供之比對路線名稱及編號與現行臺中市區公車資料並不相同，因此也無法利用臺中公車站牌資料比對悠遊卡公司。由此，本研究亦使用公車歷史行車資料與票證時間進比對，期望從中找到公車到站回傳之座標做為票證資料分析用，但又因資料格式不符(如車號 ABC123 與 ABC-123)或資料內容不符(部份客運公司公車車號在悠遊卡資料中以代號表示，與公車歷史行車資料無法比對)，使得成功比對之悠遊卡資料僅為原始資料量的 30%，因此後續分析主要仍以豐原客運提供之票證資料作為分析基礎。

針對前述資料問題，可歸納資料過濾與不過濾之處理準則：

1. 不過濾資料：資料易於修正或不影響分析結果之資料，如上下車人數不同之問題並非資料無法辨識，而其分析後之結果仍具有代

表意義因此並不過濾此問題資料，而是於資料分析時加註有此問題。

2. 過濾資料：資料無法辨識或資料經修正處理後容易發生錯誤，則將資料過濾不分析。如空間資料無法或難以比對時，即將資料過濾處理，以避免在修正後資料於電子票證時空分析時，發生站位錯誤的情況，而導致後續對策偏離原資料內容。

5.3 公車動態資料處理與課題

公車動態資料之處理包括兩大項：定時資料與定點資料，以下分節介紹處理欄位，最後說明相關課題與處理方式。

5.3.1 定時資料

本計畫擷取公車動態之定時資料（A1）如圖 5-1 作為分析依據，至於「定時資料」欄位之說明如表 5.2。

A1_06		
資料行名稱	資料類型	允許 Null
A1	char(2)	☒
Cmp	char(2)	☒
BusID	char(10)	☒
DutyStatus	tinyint	☒
BussStatus	tinyint	☒
Route	char(10)	☒
Goback	tinyint	☒
X	decimal(10, 7)	☒
Y	decimal(9, 7)	☒
Speed	tinyint	☒
Azimuth	int	☒
GPSTime	datetime	☒
Type	tinyint	☒
TrasTime	datetime	☒
SN	char(10)	☒
RecTime	datetime	☒
		☐

圖 5-1 公車動態定時資料欄位圖

表 5.2 公車動態定時資料欄位說明表

資料行描述	資料行名稱	資料類型	備註
訊息代碼	A1	char(2)	A1 定時
業者代碼	Cmp	char(2)	thbProvider
車牌號碼	BusID	char(10)	
勤務狀態	DutyStatus	tinyint	0 正常 1 開始 2 結束
行車狀況	BussStatus	tinyint	0:正常 1:車禍 2:故障 3:塞車 4:緊急求援 5:加油 100:客滿 99:非營運狀態 101:包車出租
路線代碼	Route	char(10)	thbRoute
去返程	Goback	tinyint	1 去程 2 回程 0 未知
經度座標	X	decimal(10, 7)	
緯度座標	Y	decimal(9, 7)	
車輛速度	Speed	tinyint	
車輛方位角	Azimuth	int	
GPS 時間(HH:mm:ss)GPSTime		datetime	
本訊息種類	Type	tinyint	1:定期、2:非定期
傳出時間	TrasTime	datetime	"年月日時 分秒"
發送序號	SN	char(10)	1~99999，流水號
接收時間	RecTime	datetime	"年月日時分秒"

5.3.2 定點資料

公車動態定點資料相關之資料表計有 4 個，其中最主要的資訊為 A2 表，該表各欄位之意義如表 5.3，另外三個包括 TaichungPROVIDER、TaichungROUTE 與 TaichungSTOP 資料表僅是用來描述 A2 紀錄值的標的或相關資訊而已，但這些資訊後續能提供許多維度分析之彈性。

表 5.3 公車動態定點資料欄位說明表

資料行描述	資料行名稱	資料類型	備註
訊息代碼	A2	char(2)	A2 定點
業者代碼	Cmp	char(2)	thbProvider
車牌號碼	BusID	char(10)	
勤務狀態	DutyStatus	tinyint	0 正常 1 開始 2 結束
行車狀況	BussStatus	tinyint	0:正常 1:車禍 2:故障 3:塞車 4:緊急求援 5:加油 100:客滿 99:非營運狀態 101:包車出租
路線代碼	Route	char(10)	thbRoute
去返程	Goback	tinyint	1 去程 2 回程 0 未知
站牌代碼	Stop	int	THBStop
進站離站	Leave	bit	0:離站、1:進站
GPS 時間(HH:mm:ss)GPSTime		datetime	
本訊息種類	Type	tinyint	1:定期、2:非定期
傳出時間	TrasTime	datetime	"年月日時 分秒"
發送序號	SN	char(10)	1~99999，流水號
接收時間	RecTime	datetime	"年月日時分秒"

資料來源：本研究整理

5.3.3 資料品質與對策

本節整理公車動態資料品質相關議題與因應對策如下：

1. 定點資料站牌代碼勾稽問題

錯誤！找不到參照來源。中 A2_06 表之 Stop 欄位理應能與 TaichungSTOP 表中的 STOP 欄位 100%勾稽對應，唯經比對在 2266 份站牌竟有近 8 成（1791 筆）無法對應。經向資料供應廠商諮詢後發現當站位調整或路線重整時，會調整 TaichungSTOP 表，但系統商僅維護最新（即時）版本之對應表，並未留下歷程。為解決此分析困境，本專案透過管道取得另一份歷史站牌對應表，成功將無法對應的站牌比例降低至 17%。

2. 車牌命名方式差異

本計畫取得之公車動態資料的車牌命名並不一致，部份資料格式為「123-AB」，但部份資料則為「123AB」，同時有許多車牌會加上「_0」或「_00」之尾碼，諸如「647-FX_00」，這樣的命名對一般使用者雖不會造成混淆，但就電腦系統欲做跨資料之勾稽時將產生困擾。而本計畫的處理方式是在視覺呈現程式忠實反應原貌，但在進行勾稽比對時動態用程式將這些異常符號去除並且統一格式。

5.4 數位式行車記錄器資料處理與課題

本計畫取得近 19 萬（189,083）筆數位行車記錄器壓縮檔，其檔名由客運公司代碼 2 碼（豐原客運 FY）、車牌 6 碼、路線 4 碼、日期所組成如圖 5-2。惟後續經解讀後發現「路線 4 碼」應僅為流水編號，與實際的營運許可證路線無直接關聯。

圖 5-2 中近 19 萬份檔案匯入資料庫後，計有近 8 億筆（790564435）資料，每筆資料均有 10 個欄位如圖 5-3。

	A	B	C
1	FY_012FV_1439_20150401055048.DAT.zip		
2	FY_012FV_1440_20150401070516.DAT.zip		
3	FY_012FV_1441_20150401070913.DAT.zip		
4	FY_012FV_1442_20150401082912.DAT.zip		
5	FY_012FV_1443_20150401115700.DAT.zip		
6	FY_012FV_1444_20150401130804.DAT.zip		
7	FY_012FV_1445_20150401131940.DAT.zip		
8	FY_012FV_1446_20150401141841.DAT.zip		
9	FY_012FV_1447_20150401161313.DAT.zip		
10	FY_012FV_1448_20150401173807.DAT.zip		
11	FY_012FV_1449_20150401173813.DAT.zip		
12	FY_012FV_1450_20150401190734.DAT.zip		
13	FY_013FV_1434_20150401115119.DAT.zip		
14	FY_013FV_1435_20150401220459.DAT.zip		
15	FY_015FV_2736_20150401054015.DAT.zip		
16	FY_015FV_2737_20150401083711.DAT.zip		
17	FY_015FV_2738_20150401113340.DAT.zip		

圖 5-2 數位行車記錄器原始檔案名稱清單

OBU_2

資料行名稱	資料類型	允許 Null
BUS	varchar(8)	☐
ROUTEID	varchar(8)	☐
GPSTime	datetime	☒
Longitude	float	☒
Latitude	float	☒
Speed	float	☒
RPM	float	☒
FrontDoor	smallint	☒
RearDoor	smallint	☒
Acc	smallint	☒
		☐

圖 5-3 數位行車記錄器資料表欄位圖

這 60 多 GB 近 8 億筆資料在分析前即發現數個問題，其中較關鍵的課題描述暨本專案的解決之道分述如下：

1. RouteID 並非路線許可證編號

RouteID 疑似流水號，計有 10257 種不重複數值，明顯大幅超過可能的路線數，加上許多數字為連號，且經查許多號碼皆非路線號碼。故本研究將忽略此欄位，當需要路線資訊時，則搭配公車動態或電子票證的時間與車牌來勾稽取得路線。

2. 速度欄位數值不合理

八億筆中計有 51910540 筆（6.5%）時速超過 150km/h，甚至這其中有 45324894 筆（87%）是超過 200km/h，明顯非一般公路車輛所能到達之速度，經請教設備原廠後，技術人員建議不該僅刪除超過 150km/h 之離群值，而是當個別車機若有 10% 以上的不合理資料時，整份車機資料全部放棄，同時這些資料不限於速度異常，包括以下幾個條件均在 10% 容忍統計範圍內：

- (1) 速度小於 0 km/h 或超過 150 km/h
- (2) 轉速小於 0 RPM
- (3) 位置座標偏移出臺灣本島

3. 同車牌同時間多筆資料

完成前項之資料清理後，資料筆數降低至 7 億筆（697198657），但其中發現有同車牌同秒間有多筆資料之情況，此議題則以「同車牌同時間多筆資料取平均」的方式處理，清理後的資料筆數約為 4.6 億筆作為後續分析的基礎。

5.5 營運相關資料處理與課題

營運相關資料包含營運地區需求、路線站位、駕駛員、機務及油料等資料，這些營運資料經由資料處理進行供需分析及車輛營運等分析。基此，本研究蒐集人口資料、汽車客運供給資料（路線長度、班次數）、搭乘運量（豐原客運電子票證資料）、車型資料、機務資料及油料資料，各項資料蒐集內容及格式說明如后。

5.5.1 旅次需求資料

本研究藉由人口資料反映該行政區之需求數，故蒐集臺中市人口資料及人口結構資料(學生人口數及高齡者人口數)，其中，學生人數係蒐集 5 歲-24 歲人口數，高齡者人口數為 65 歲以上人口數，資料內容如表 5.4 所示。

表 5.4 臺中市 104 年各行政區人口數

分區	總人口數	學生人口數	老年人口數	分區	總人口數	學生人口數	老年人口數
中區	19,231	5094	2762	神岡區	64,995	15855	6802
東區	75,044	16141	9179	潭子區	105,819	26793	8859
南區	120,919	28706	10894	大雅區	93,226	25654	7068
西區	115,824	28039	13604	新社區	25,153	5692	4134
北區	147,446	33228	17192	石岡區	15,309	3439	2286
西屯區	220,270	58015	17529	外埔區	31,746	7500	3688
南屯區	163,250	45497	11554	大安區	19,572	4458	2946
北屯區	262,204	65520	23291	烏日區	72,018	18049	7098
豐原區	166,665	42265	18355	大肚區	56,428	13784	5930
東勢區	51,313	11144	8615	龍井區	76,245	19634	6969
大甲區	77,640	20662	8796	霧峰區	64,563	14900	7794
清水區	85,971	21101	10323	太平區	182,783	44784	15656
沙鹿區	88,715	22204	7925	大里區	207,339	54928	15588
梧棲區	56,906	14764	5139	和平區	10,720	2254	1594
后里區	54,186	13364	6569				

資料來源：臺中市政府民政局

考慮人口結構資料並無法呈現分時段之需求分布，故本研究採用「臺中都會區住訪調查及運輸需求預測模式檢討與建立」調查之各時段旅次發生比例如表 5.5，學生人口係統計家-上學旅次，而老年人口係根據其調查結果，老年人比重較高之旅次(家-其他旅次)作為分時段比例參考。

表 5.5 臺中市各時段旅次出發時間比例

出發時間	0	1	2	3	4	5	6	7	8	9	10	11
總人口	0.27%	0.12%	0.10%	0.11%	0.26%	0.79%	5.89%	20.42%	8.71%	5.33%	4.05%	2.86%
學生人口	0.00%	0.00%	0.00%	0.00%	0.00%	0.33%	12.25%	28.91%	3.29%	1.13%	0.44%	0.24%
老人人口	0.17%	0.08%	0.06%	0.06%	0.18%	0.83%	4.74%	15.00%	9.56%	8.93%	7.25%	5.22%
出發時間	12	13	14	15	16	17	18	19	20	21	22	23
總人口	3.54%	1.62%	2.09%	2.63%	7.59%	14.07%	7.77%	3.79%	2.93%	2.98%	1.61%	0.48%
學生人口	2.33%	0.43%	0.24%	1.02%	13.24%	18.31%	8.06%	2.87%	1.91%	3.15%	1.70%	0.15%
老人人口	4.43%	1.84%	2.76%	3.95%	8.53%	7.75%	5.68%	4.40%	3.57%	3.25%	1.29%	0.47%

資料來源：臺中都會區住訪調查及運輸需求預測模式檢討與建立

5.5.2 路線及班次資料

汽車客運供給資料內容如表 5.6 所示，此資料內容主要來自公車動態系統之資料庫，此資料庫包含臺中市各業者之路線及班次資料，並記錄各路線之站位及到站時間，藉由這些資料後續可進一步分析各區於各時段之供給(路線數、班次數)。並藉由站點之距離，可進一步分析各區之路線長度。

表 5.6 汽車客運供給欄位說明

欄位項目	說明
業者編號	業者對應編號
路線編號名稱	該路線之名稱
站編號	紀錄各站點之編號，各站擁有其對應編號
站序	該站位在該路線之站位順序
站名	該站點之站名
區名	該站點所在之行政區
里名	該站點所在之村、里
預估距離(公尺)	該站點與前站之距離
路線編號	該路線之編號
首站發車時間	該班次首站發車時間
往返註記	往程=1 返程=2
預計到站時間	該班次到達各站時間

資料來源：本研究整理

5.5.3 行政區 OD 運量資料

行政區 OD 運量資料係根據電子票證資料彙整而來，電子票證資料由各客運公司保管，資料內容包含刷卡日期、搭乘路線編號、搭乘之車牌號碼、票種、上車時間及上下車地點，如表 5.7 所示。由於本研究以豐原客運作為分析對象，且電子票證受限於各公司的管制，故本研究僅針對豐原客運之電子票證進行分析。

本研究係以行政分區為單位進行區到區間的 OD 運量分析，故須將電子票證紀錄之上下車站位轉換為行政區，其站位轉換參考主要依據前述供給欄位站位與行政分區的關係，進而將電子票證站位資料，歸納為行政區資料。然在歸納過程中，發現部分電子票證站位名稱與汽車客運供給欄位站位名稱不一致，其原因在於站位因為路線調整、或站位移動，使得站位名稱更動。為克服此問題，嘗試利用站位編號進行配對，亦發現站位編號無法進行有效的配對。故為克服此問題，本研究則比對歷史站位資料，尋找站位名稱差異之站牌進行配對，以減少無法配對產生之誤差。

表 5.7 電子票證欄位說明

欄位項目	說明
日期	搭乘日期
路線編號	搭乘之路線編號
車輛編號	搭乘之車牌號碼
票種	使用票種
上車時間	上車時間
上車站名	上車地點
下車站名	下車地點

資料來源：豐原客運提供暨本研究整理

5.5.4 車型資料

資料庫中車型資料欄位包含車輛種類、廠牌、出廠年份、引擎編號、車輛噸位及座位數，如表 5.8。其中，車輛種類資料庫中僅呈現「大營客」，無法辨別出該車是一般大客車或低地板公車，故本研究進一步蒐集「公共運輸提升發展計畫」資料庫內，所記載的新闢路線、車輛汰舊換新之車牌編號及車型，與本資料庫配對。然「公共運輸提升發展計畫」之資料庫僅記載受補助之新購置車輛，車輛未汰換或未受補助則無法進行配對，必須透過另外的搜尋方式定義該車車型，需耗費大量的搜尋時間。

表 5.8 車輛清冊表

編號	資料行描述	屬性	資料行名稱	說明
1	車別	文字	VEHICLE_TYPE	營運車輛分類 (大營運客車)
2	車號	數字	BUSID	
3	廠牌	文字	BRAND	
4	年份	日期	YEARS	西元/月
5	引擎號碼	數字	ENGINE_ID	
6	噸位	數字	TON	
7	座位	數字	SEATS	
8	備註	文字	MEMO	

資料來源：豐原客運提供暨本研究整理

5.5.5 油料資料

本分析資料來自於客運業者的營運資料，包含有每日營運報表、車輛清冊和車輛維修等，本分析鎖定各車牌號碼一年的燃油資料以及車輛特性為主，由於客運業者暫未提供駕駛每天之耗油量，因此只能從各車牌號碼來整理哪些車號的駕駛和車輛特性是最耗油的原因。

1. 原始資料

對於油耗分析，本研究的原始資料包含三個部分：車輛零件統計表和電子票證資料。車輛零件統計表主要記錄營運車輛一年份的總耗油

量、各車牌號碼的公里柴油公升等資料(表 5.9)，電子票證資料主要記錄駕駛員編號、車牌號碼和路線編號等資料(表 5.10)。

表 5.9 車輛零件統計表

編號	資料行描述	屬性	資料行名稱	說明
1	流水號	數字	ID	
2	車號	數字	BUSID	
3	總油耗量	數字	TOTAL_OIL	
4	公里柴油公升	數字	DIESEL_FUEL_PER_KILOMETER	總公里數/柴油公升數
5	機油量	數字	ENGINE_OIL	公升
6	公里機油公升	數字	ENGINE_OIL_PER_KILOMETER	總公里數/機油公升數
7	柴油耗油金額	數字	AMOUNT_DIESEL_FUEL	
8	零件費	文字	AMOUNT_VEHICLE_PARTS	
9	輪胎費	數字	AMOUNT_TIRE	
10	機油費	數字	AMOUNT_ENGINE_OIL	
11	雜油費	數字	AMOUNT_OIL	
12	耗用總計	數字	AMOUNT_CONSUMPTION	
13	耗用公里	數字	AMOUNT_KILOMETER	耗用總計/總公里數,元/公里
14	年份	日期	YEARS	西元/月
15	廠牌	文字	BRAND_VEHICLE	

資料來源：豐原客運提供暨本研究整理

表 5.10 電子票證資料表

編號	資料行描述	屬性	資料行名稱
1	營運日期	日期	data
2	司機編號	數字	driver_id
3	路線編號	數字	route_id
4	車輛編號	數字	car_id
5		數字	d_no
6	更新時間	時間	uptime

資料來源：豐原客運提供暨本研究整理

2. 資料轉換與處理

資料前處理的動作，主要目的在於過濾並去除不需要的資料，有助於資料探勘的進行與最後歸責的品質。資料前處理的主要動作有：

- (1) 清除不需要的資料:將不需要的資料刪除。車輛清冊資料部分，本分析主要對象是各車輛特性及車牌號碼進行油耗分析，對於引擎號碼及座位數對本分析的探討並沒有幫助;車輛零件統計表資料部分，將機油和零件部分的資料刪除，這一部分的資料並非本分析的重點，因此未討論;電子票證資料部分，在這一部分只為了將車牌號碼及駕駛連結起來。
- (2) 去除遺漏的資料:刪除缺漏值及錯誤的資料，減少對資料探勘結束的影響。
- (3) 資料轉換:將資料的表示方法，轉換成相同的格式，其目的在協助之後探勘的進行。如:車輛出廠年份原本的 2003.01、2004.04，轉換成 2003 和 2004。
- (4) 耗油量:將車輛零件統計表內的公里柴油公升(總公里數/柴油公升數)轉換成每公里消耗多少油量(柴油公升數/總公里數)，主要原因公里柴油公升屬於燃油消耗(總公里數/柴油公升數)，並非直觀的耗油量(柴油公升數/總公里數)，轉換方式將公里柴油公升做倒數處理。

5.5.6 機務資料

分析資料來自於業者所提供之每日營運表、車輛清冊和車輛維修等，本分析以車輛進廠維修的次數和原因以及車輛特性為主。

1. 原始資料

對於機務分析，本研究的原始資料包含三個部分：車輛清冊、車輛維修和電子票證資料。車輛清冊主要記錄豐原客運營運車輛的基本資料，包含車別、車號、廠牌、年份和噸位等等(表 5.8)，車輛維修主要記錄一年內所有車輛進廠維修(表 5.11)。

表 5.11 車輛維修資料表

編號	資料行描述	屬性	資料行名稱
1	流水號	日期	ID
2	車號	數字	BUSID
3	維修日期	日期	MAINTENANCE_DATE
4	維修時間	日期	MAINTENANCE_TIME
5	故障原因	文字	SUBJECTS
6	進場日期	日期	REPAIRER_DATE
7	進場時間	日期	REPAIRER_TIME
8	完修日期	日期	COMPLETION_DATE
9	完修時間	日期	COMPLETION_TIME

2. 資料轉換與處理

資料前處理的動作，主要目的在於過濾並去除不需要的資料，有助於資料探勘的進行與最後歸責的品質。

- (1) 清除不需要的資料:將不需要的資料刪除。車輛清冊資料部分，本分析主要對象是各車輛特性及車牌號碼進行機務分析，因此無需使用引擎號碼及座位數。
- (2) 去除遺漏的資料:刪除缺漏值及錯誤的資料，減少對資料探勘結束的影響。
- (3) 資料轉換:將資料的表示方法，轉換成相同的格式，其目的在協助之後探勘的進行。如:車輛出廠年份轉換成車齡原本的2003.01、2004.04，轉換成12年和11年。
- (4) 資料歸類:將車輛維修資料內的故障原因分類，分成故障及維修兩大項目，故障的部分主要以車輛在行駛過程中發生拋錨，而維修則是其他原因。

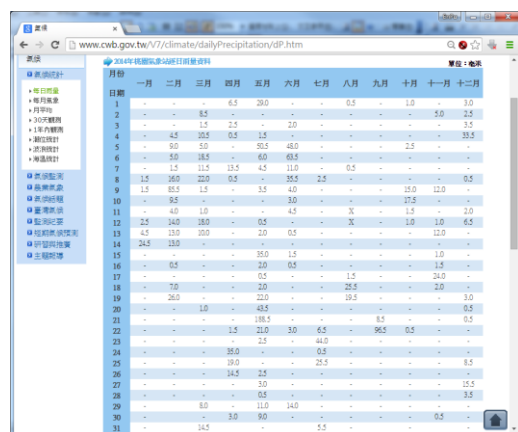
5.6 其他運輸相關資料處理與課題(事故、計程車資料)

本研究取得之計程車訂單資料因為在派車流程中，當乘客訂車時，系統會記錄乘車起點座標，但乘車下車時並未在訂車資料中記錄下車座標，車機僅回傳車乘下車訊息，因此若需乘客上下車時間及起訖座標，即需至行車軌跡資料中，將乘客上下車時間及座標資料取出才可獲得完整之乘車訂單資料。但由於計程車派遣系統中，行車軌跡資料因考量資料量、資料重要性及後續利用性，因此以簡易之文字檔資料存放，而本案所需之計程車資料，即需至文字檔中以人工方式取得資料，才可進行後續分析。

5.7 其他非運輸相關資料處理與課題(氣候)

相對於其他資料，氣象相關資料之介接與擷取較為單純，由於本計畫是分析公車動態暨電子票證的歷史資料，故並非介接即時天氣資訊，而是批次擷取歷史氣象資料來分析。

圖 5-4 中的氣象局歷史資料查詢雖然有詳細的氣溫資料，但雨量資料則是每天的匯總值，故本計畫另採同是氣象局維護的氣象歷史資料庫維護平臺(如圖 5-5)作為介接對象。



日期	一月	二月	三月	四月	五月	六月	七月	八月	九月	十月	十一月	十二月
1	-	-	6.5	20.0	-	-	0.5	-	1.0	-	-	3.0
2	-	-	8.5	-	-	-	-	-	-	5.0	2.5	-
3	-	-	1.5	2.5	2.0	-	-	-	-	-	3.5	-
4	-	4.5	20.5	0.5	1.5	-	-	-	-	-	35.5	-
5	-	9.0	5.0	-	50.5	48.0	-	-	-	2.5	-	-
6	-	5.0	18.5	-	6.0	68.5	-	-	-	-	-	-
7	-	1.5	11.5	15.5	4.5	11.0	-	0.5	-	-	-	-
8	1.5	36.0	22.0	0.5	-	35.5	2.5	-	-	-	-	0.5
9	1.5	83.5	1.5	-	3.5	4.0	-	-	25.0	12.0	-	-
10	-	8.5	-	-	-	3.0	-	-	-	12.5	-	-
11	-	4.0	1.0	-	-	4.5	-	X	-	1.5	-	2.0
12	2.5	14.0	18.0	-	0.5	-	-	X	-	1.0	1.0	6.5
13	4.5	13.0	10.0	-	2.0	0.5	-	-	-	-	2.0	-
14	24.5	13.0	-	-	-	-	-	-	-	-	-	-
15	-	-	-	-	35.0	1.5	-	-	-	-	1.0	-
16	-	6.5	-	-	3.0	8.5	-	-	-	-	1.5	-
17	-	-	-	-	0.5	-	-	1.5	-	-	24.0	-
18	-	7.0	-	-	2.0	-	-	25.5	-	-	2.0	-
19	-	26.0	-	-	23.0	-	-	19.5	-	-	3.0	-
20	-	-	-	-	48.5	-	-	-	-	-	0.5	-
21	-	-	-	-	188.5	-	-	-	8.5	-	-	0.5
22	-	-	-	1.5	24.0	3.0	6.5	95.5	0.5	-	-	-
23	-	-	-	-	2.5	-	44.0	-	-	-	-	-
24	-	-	-	38.0	-	-	0.5	-	-	-	-	-
25	-	-	-	33.0	-	-	26.5	-	-	-	6.5	-
26	-	-	-	14.5	2.5	-	-	-	-	-	-	-
27	-	-	-	-	3.0	-	-	-	-	-	12.5	-
28	-	-	-	-	6.5	-	-	-	-	-	3.5	-
29	-	-	-	8.0	-	11.0	14.0	-	-	-	-	-
30	-	-	-	3.0	9.0	-	-	-	-	-	0.5	-
31	-	-	-	14.5	-	-	5.0	-	-	-	-	-

(a)雨量資料



日期	測站氣壓(hPa)	溫度(°C)	相對濕度	風速(mph)	風向	降水量(mm)	日照時數(小時)
01/01	1015.5	26.9	84	2.3	東北東	0	0
01/02	1016.3	26.5	86	0.8	東北東	0	0
01/03	1016.8	26.2	90	0.9	西南西	0.5	0
01/04	1016.6	26.4	92	2	東	1	0
01/05	1017.7	24	87	4.3	東	0	0
01/06	1018.5	25.7	84	3.9	東	0.5	0
01/07	1018.9	23.9	83	4.3	東	0	0
01/08	1018.3	23.9	81	3.4	東	0	0
01/09	1019.5	24.7	81	2.9	東北東	0	0
01/10	1020.2	24.7	78	3.3	東	0	0
01/11	1020.5	24.3	76	3.6	東	0	0
01/12	1020	24	73	4.8	東	0	0
01/13	1019.4	24.1	71	1.9	東北東	0	0
01/14	1018.5	23.9	72	2.3	東北東	0	0
01/15	1018.2	23.4	73	3.1	東	0	0
01/16	1018.5	23.2	78	3.7	東	0	0
01/17	1018.3	23	75	2.8	東	0	0
01/18	1018.6	22.7	78	2.5	東	0	0
01/19	1018.6	22.7	75	3.1	東北東	0	0
01/20	1020.3	22.8	77	2.7	東	0	0
01/21	1020	23.0	75.3	2.4	東	0	0
01/22	1019.7	23.5	74	2.9	東	0	0
01/23	1019.6	23.5	74	2.3	東	0	0
01/24	1019.2	23.5	74	2.1	東北東	0	0

(b)氣溫資料

資料來源：<http://www.cwb.gov.tw/>

圖 5-4 歷史資料日氣象資料收集來源示意圖



資料來源：<http://e-service.cwb.gov.tw/HistoryDataQuery/index.jsp>

圖 5-5 氣象歷史資料收集來源

圖 5-5 所提供的資料欄位相當完整（如圖 5-6 計有 14 項資訊），時間也精細到每測站每小時均有一筆資料，由於該站臺並沒有提供 API 資料介接的申請程序，因此本計畫僅能採用「爬網」方式批次撈取資料，爬網程式主要是利用 .NET 類別庫中的「System.Net.WebClient」類別來模擬瀏覽器上網擷取資料，透過迴圈批次修改參數來達到批次背景抓取各測站各日各時段氣象資料之目的，程式撰寫完成後，批次撈取近 2 年（2013 年 1 月～2015 年 4 月）的歷史氣象資料匯入資料庫如圖 5-7。

觀測時間 (LST) ObsTime	海平面氣壓 (hPa) SeaPres	氣溫 (°C) Temperature	露點溫度 (°C) Td dew point	相對濕度 (%) RH	風速 (m/s) WS	風向 (360degree) WD	最大陣風 (m/s) WSGust	最大陣風風向 (360degree) WDGust	降水量 (mm) Precp	降水時數 (hr) PrecpHour	日照時數 (hr) SunShine	全天日照時量 (MJ/m²) GrossRad	能見度 (km) Visd
1	1011.4	1021.7	14.3	10.3	77	1.1	320.0	2.6	310.0	0.0	0.0	0.00	
2	1011.5	1021.9	13.0	9.5	79	0.2	0.0	2.9	350.0	0.0	0.0	0.00	10.0
3	1011.5	1021.9	13.0	9.4	79	0.7	10.0	2.5	360.0	0.0	0.0	0.00	
4	1011.6	1022.0	11.8	9.0	83	0.1	0.0	1.8	320.0	0.0	0.0	0.00	
5	1011.9	1022.3	12.1	8.6	79	0.2	0.0	2.7	340.0	0.0	0.0	0.00	10.0
6	1012.7	1023.1	13.0	8.0	72	1.5	10.0	5.3	10.0	0.0	0.0	0.00	
7	1013.2	1023.6	11.8	7.9	77	0.3	360.0	4.8	330.0	0.0	0.0	0.00	
8	1014.2	1024.6	13.4	8.0	70	1.7	10.0	4.5	310.0	0.0	0.0	0.7	0.23
9	1014.9	1025.2	15.3	7.9	61	1.4	320.0	3.7	310.0	0.0	0.0	1.0	0.86
10	1014.4	1024.6	17.6	8.5	55	1.1	360.0	4.1	360.0	0.0	0.0	1.0	1.39
11	1013.4	1023.6	19.4	8.1	48	1.4	360.0	5.6	340.0	0.0	0.0	1.0	1.80
12	1012.4	1022.5	20.1	6.8	42	2.5	340.0	6.2	350.0	0.0	0.0	1.0	1.98
13	1011.6	1021.7	20.2	6.6	41	2.0	350.0	7.2	320.0	0.0	0.0	1.0	2.20
14	1011.0	1021.1	19.4	6.0	42	3.1	340.0	7.9	350.0	0.0	0.0	1.0	1.80
15	1011.1	1021.3	18.4	5.4	42	3.9	350.0	10.1	350.0	0.0	0.0	1.0	1.38
16	1011.5	1021.7	17.2	5.2	45	3.1	350.0	8.9	350.0	0.0	0.0	1.0	0.84
17	1012.1	1022.4	15.2	4.7	49	3.1	340.0	8.5	310.0	0.0	0.0	0.7	0.26
18	1012.8	1023.1	14.0	4.9	54	3.4	350.0	6.5	330.0	0.0	0.0	0.0	
19	1013.7	1024.1	13.2	5.0	57	2.9	350.0	6.4	350.0	0.0	0.0	0.0	
20	1014.0	1024.4	12.6	4.9	59	2.6	350.0	6.9	360.0	0.0	0.0		12.0
21	1014.1	1024.5	12.4	5.0	61	1.9	340.0	5.4	10.0	0.0	0.0	0.00	12.0
22	1014.1	1024.5	11.9	4.8	62	1.4	340.0	5.4	360.0	0.0	0.0	0.00	
23	1014.1	1024.5	11.6	4.7	63	2.2	350.0	4.3	360.0	0.0	0.0	0.00	10.0
24	1014.3	1024.7	11.8	4.6	63	2.6	360.0	6.0	360.0	0.0	0.0	0.00	

資料來源：<http://e-service.cwb.gov.tw/HistoryDataQuery/index.jsp>

圖 5-6 氣象資料擷取欄位清單

weather_data		
資料行名稱	資料類型摘要	可否 Null
wd_index	decimal(18...	否
wd_wa_id	nvarchar(3)	是
wd_date	date	是
wd_hour	float	是
wd_hpa_1	float	是
wd_hpa_2	float	是
wd_temperature_1	float	是
wd_temperature_2	float	是
wd_humidity	float	是
wd_wind_ms	float	是
wd_wind_degree	float	是
wd_gust_ms	float	是
wd_gust_degree	float	是
wd_rain_mm	float	是
wd_rain_hr	float	是
wd_sunshine_hr	float	是
wd_sunshine_mj	float	是
wd_visibility	float	是

weather_area		
資料行名稱	資料類型	允許 Null
wa_id	nvarchar(3)	☐
wa_name	nvarchar(20)	☒
wa_memo	nvarchar(20)	☒
		☐

weather_city		
資料行名稱	資料類型	允許 Null
wc_id	nvarchar(1)	☐
wc_name	nvarchar(10)	☒
		☐

圖 5-7 擷取/清理後之歷史氣象資料表

氣象資料的介接相當完整，擷取 28 個月（2013 年 1 月～2015 年 4 月）僅有 2015 年 1 月 21 日為缺漏外，其餘數值均正常擷取匯入成功，相較於本計畫其他介接資料，算是資料品質最高之項目。

第六章 探勘分析與視覺化呈現

根據第三章的需求彙整結果，本研究將進行 9 項之巨量資料分析與成果視覺化，分別為：(1)電子票証分析、(2)路線與班次重疊分析、(3)供需分析、(4)結合電子票証與公車動態之分析(5)行車時間與班表分析、(6)站牌班表產製分析、(7)車輛相關分析（包含機務、油料）、(8)駕駛員分析風險分析與(9)車機穩定度分析。

6.1 電子票證分析

6.1.1 探勘分析目的

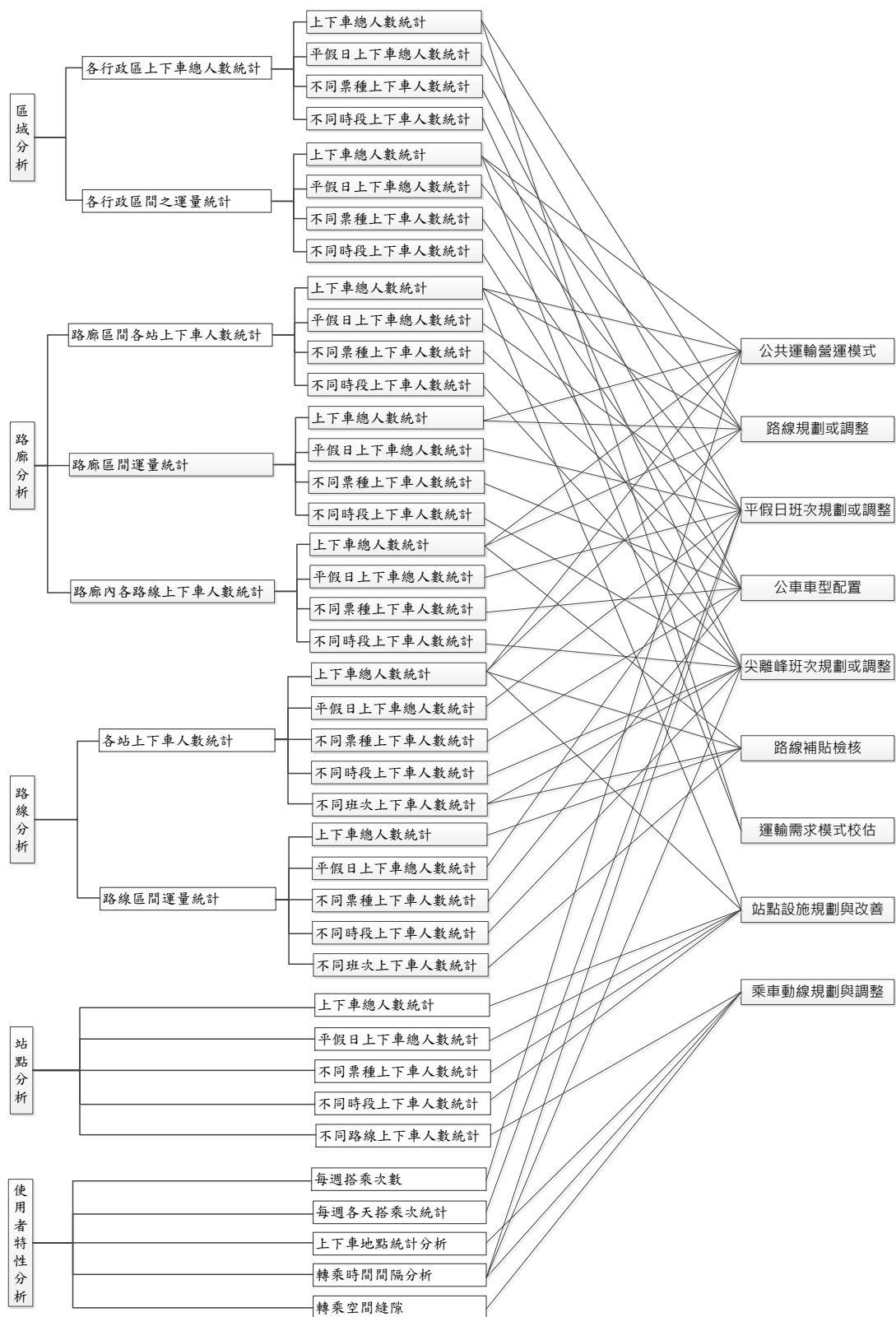
蘇昭銘等人於 2014 年運輸研究所「先進公共運輸系統整合資料庫加值應用系統維運及推廣計畫」研究案提出電子票證分析目的及流程，並提出電子票證資料之五大潛在應用範圍，包含區域、路廊、路線、站點、使用者等應用範圍。

1. 區域分析：以行政區為分析單元，主要在了解各行政區間大眾運輸系統之起訖量及行政區間之區間運量，行政區分析單元可為縣市、鄉鎮市區或村里等不同等級單元，其中區間運量分析方面，建議行政區間以同一層級，如：竹東鎮到橫山鄉、竹東鎮到新竹市東區、竹東鎮到五峰鄉、竹東鎮到北埔鄉之運量等；或是與上一層級間，如竹東鎮到新竹市、竹東鎮到苗栗縣、竹東鎮到臺北市之運量等。藉由行政區間之運量可檢討特定行政區之現行聯外大眾運輸路線是否會有供給過剩、供給不足等現象。
2. 路廊分析：以特定路廊（如臺中市臺灣大道路廊）為分析對象，使用者可自行設定路廊所涵蓋之大眾運輸路線，藉由路廊之分析，可了解路廊之主要承載區間及各路線間之運量差異。如臺東縣連接花蓮縣之山線路廊分析，即可了解乘客主要之搭乘區間為何？

是否以起訖點站牌之運量為主，或是中間特定區間站牌之運量較多，由這些運量指標，可作為路線調整、路線補貼等營運管理作業機制之參考。

3. 路線分析：以個別路線為分析單元，如新竹客運公司的 5625 公路客運路線，從該路線之運量分析中，可了解該路線主要之乘載區間為何？是否以起訖點站牌之運量為主，或是中間特定區間站牌之運量較多；亦可了解該路線乘客之主要上下車站點、不同班次運量班次，由該些運量指標，可作為路線調整、班次調整或路線補貼等營運管理作業機制之參考。
4. 站點分析：由於同一站點可能有多條路線經過，藉由不同路線在同一站點之上下車人數統計，了解該站點所產生或吸引之運量，如在新竹火車站之所有市區公車或公路客運路線上下車人數，即可了解一個地區中主要之大眾運輸服務區，可作為場站規劃、路線調整或相關營運計畫規劃之參考。
5. 使用者特性分析：在電子票證之分析中，亦可以票種為分析單元，了解每位使用者之搭乘行為，藉由使用者特性分析，除可掌握使用者搭乘地點與時間之時空分佈外，也可以藉由時空分佈情形，進一步了解轉乘行為之時間與空間縫隙，當發現有多數使用者具有類似之需求時，即可針對大眾運輸路線或班次，進行必要之調整或增加，以降低使用者在使用過程中之不方便性。

該研究並提出依五大潛在應用範圍共 52 項分析指標之內容，可作為決策與規劃人員進行交通分析與規畫時之參考數據。其應用領域之對應架構圖如圖 6-1 所示，其中可適用的應用領域則為「公共運輸營運模式」、「路線規劃或調整」、「平假日班次規劃或調整」、「低底板公車配置」、「加離峰班次規劃或調整」、「路線補貼檢核」、「運輸需求模式校估」、「站點設施規劃予改善」與「乘車動線規劃與調整」等相關領域。



資料來源：蘇昭銘等人，「先進公共運輸系統整合資料庫加值應用系統維運及推廣計畫」

圖 6-1 電子票證分析指標與應用領域對應架構圖

6.1.2 演算處理流程

本研究以豐原客運提供之電子票證資料做為研究對象，其資料欄位共計 50 欄，可用於資料分析欄位摘錄 19 欄，如表 6.1 所示。

表 6.1 電子票證分析用欄位表

編號	欄位名稱	說明
1	date	乘車日期
2	DRIVER_ID	駕駛員代碼
3	ROUTE_ID	路線代碼
4	CAR_ID	車輛編號
5	D_NO	勤務編號
6	T_CMP	票證公司別
7	KIND	票證類型
8	CARD	票證代碼
9	uptime	上車時間
11	upname	上車站名
13	dmtime	下車時間
15	dmname	下車站名
17	r_beg	上車站序
18	r_end	下車站序
19	r_mil	里程

本研究於取得之電子票證資料後，為確保資料正確性，採用之步驟包括下列三階段：

1. 資料完整性檢驗

取得電子票證資料後，需先檢驗個別欄位資料是否齊全，由於電子票證資料可能因取得之來源單位不同或資料轉換次數過多，可能有同一資料欄位內容不一致之情況，如欄位內容為空值或資料屬性不一致，則視為無效資料。

2. 資料加值

由於電子票證資料中，並無各上、下車站對應之經緯度座標，導至無法進行空間資料分析，因此利用所蒐集之公車路線、站牌等基本資料，利用路線編號、站牌名稱進行定位，以取得各站牌所在區域名稱。

3. 資料分析

經空間加值後，即可現有電子票證資料進行具有空間特性之資料分析，相關分析內容將於後續小節中加以說明。

6.1.3 視覺呈現暨應用

本節針對前述五大項應用領域及不同指標內容提出可能之應用情境，並利用相關圖表說明資料分析結果。

區域分析:圖 6-2 為行政區分時運量統計分析，透過此圖可知豐原區之分時旅次數皆高於其它行政區；並且多數行政區在晨昏尖峰時間，其旅次需求也比其它時段要高，比較特別的是上午 6 點即為旅次尖峰時間。圖 6-3 顯示在豐原區之電子票證使用數量高於其它地區，亦呼應圖 6-2 之統計結果，而一般卡搭乘次數在多數地區皆高於其它使用者。

1. 圖 6-4 可知行政區內之旅次需求仍是民眾搭乘公車主要目的，如豐原區、東勢區、太平區等，其次豐原區-潭子區、豐原區-東勢區、太平區-東區等跨區旅次需求，亦為民眾主要旅次行為。由分析可知，豐原區為民眾主要旅次發生地區，因此分析豐原區各時段民眾主要旅次發生的起迄地，圖 6-5 即可發現多數時段民眾之旅次行為仍為區域內旅次，僅有少數時段之神岡區、后里區、潭子區、北區等為民眾較常前往的地區。圖 6-6 則利用地理資訊軟體將旅次資料轉換成地理資訊，讓使用者可更容易了解旅次量及行政區之關係。

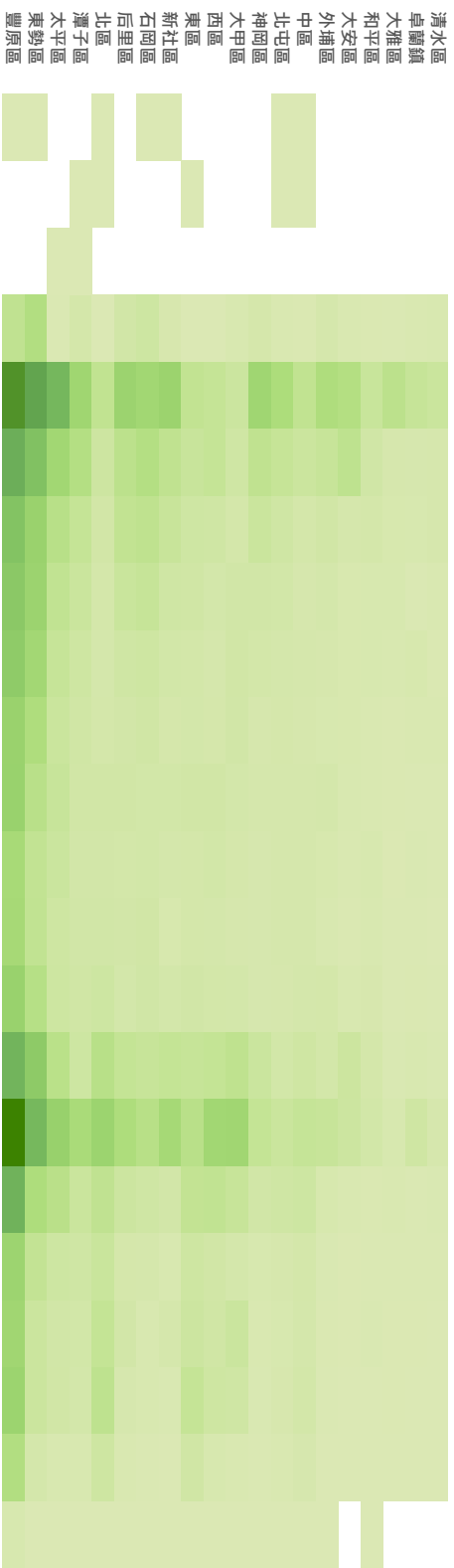


圖 6-2 行政區運量統計圖

行政區運量統計



圖 6-3 行政區票卡別運量統計圖

行政區間運量分析

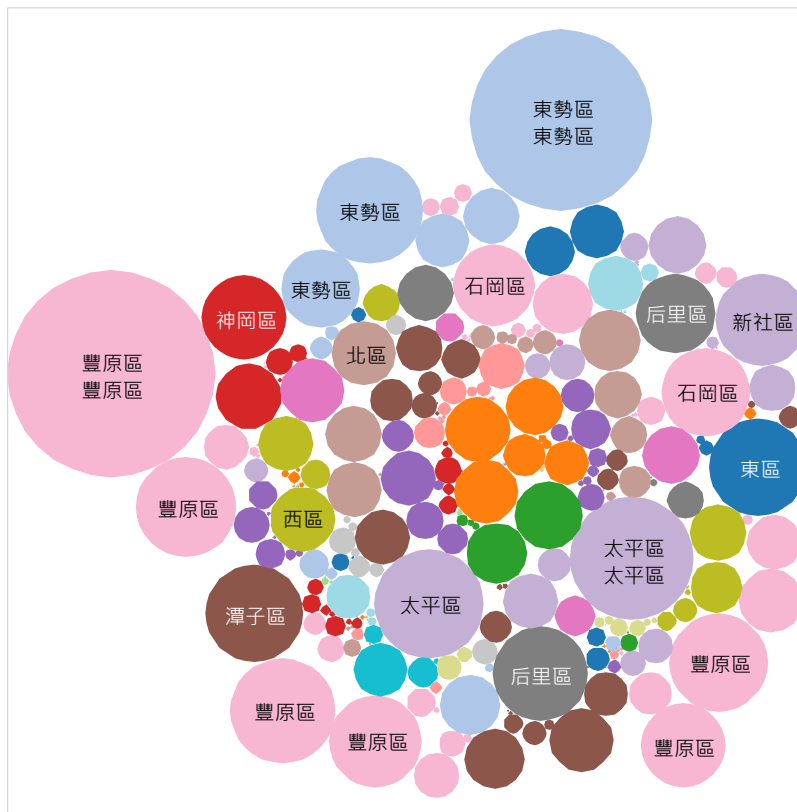


圖 6-4 行政區間旅次需求統計圖

行政區分時運量統計

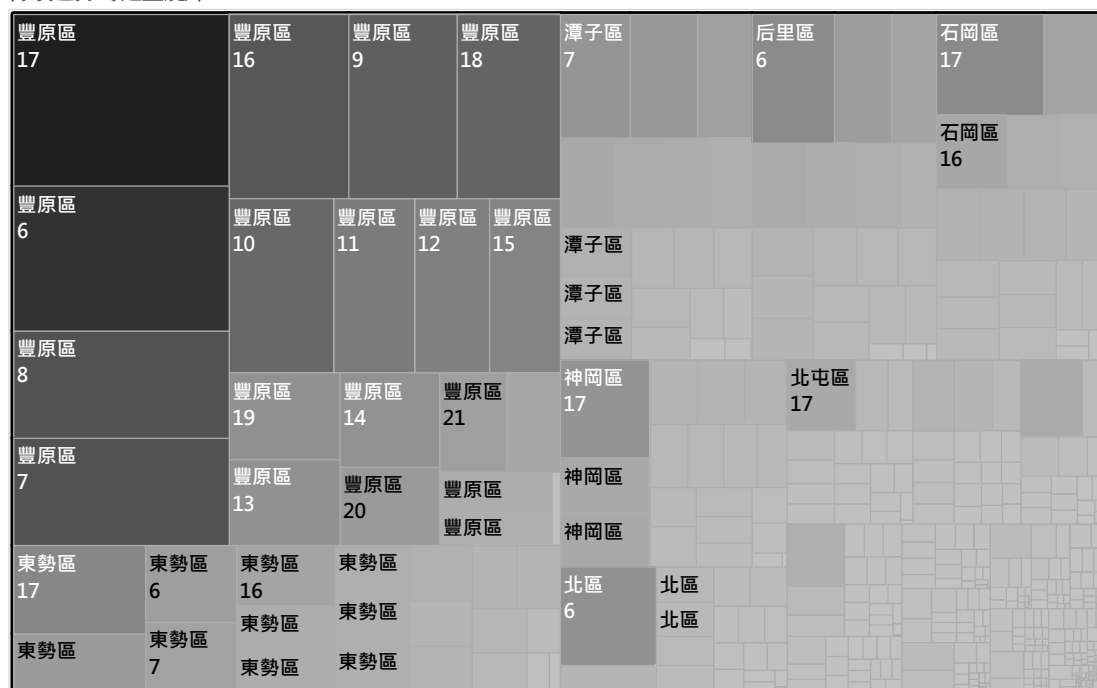


圖 6-5 豐原區分析旅次需求統計圖

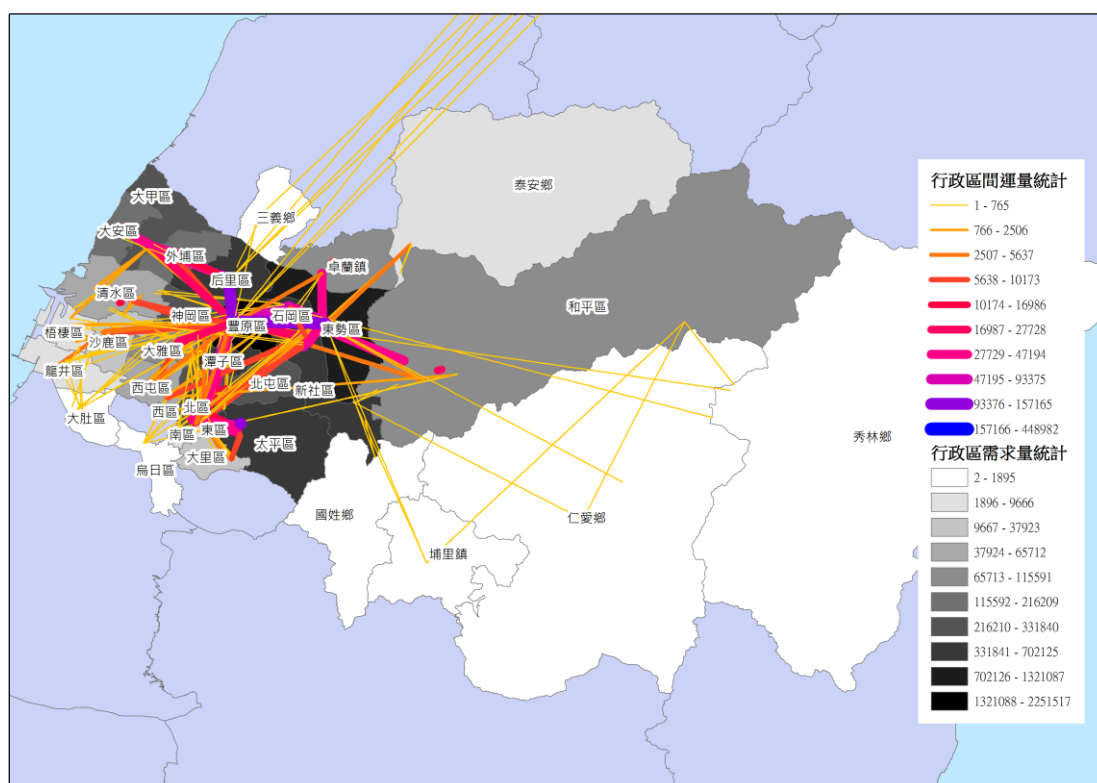


圖 6-6 行政區間旅次需求示意圖

路廊分析：圖 6-7 為利用電子票證各站上下車資料，進行空間定位、統計需求後所得之運量資料，圖中深色部份為需求較高之區域，淺色部份則為旅次需求較少之區域。由前述行政區間旅次分析雖可知行政區之間的旅次數，但無法明確指出其實際旅次發生之空間關係，因此必須進行路廊旅次示意圖，即可得旅次發生及與其空間關係，如

2. 圖 6-4 及圖 6-7 所示，豐原區-潭子區為主要旅次需求區間，其主要旅次行為主要沿省道臺 3 線(圖 6-8)，其資料即可提供使用者規劃路線、班次之用。

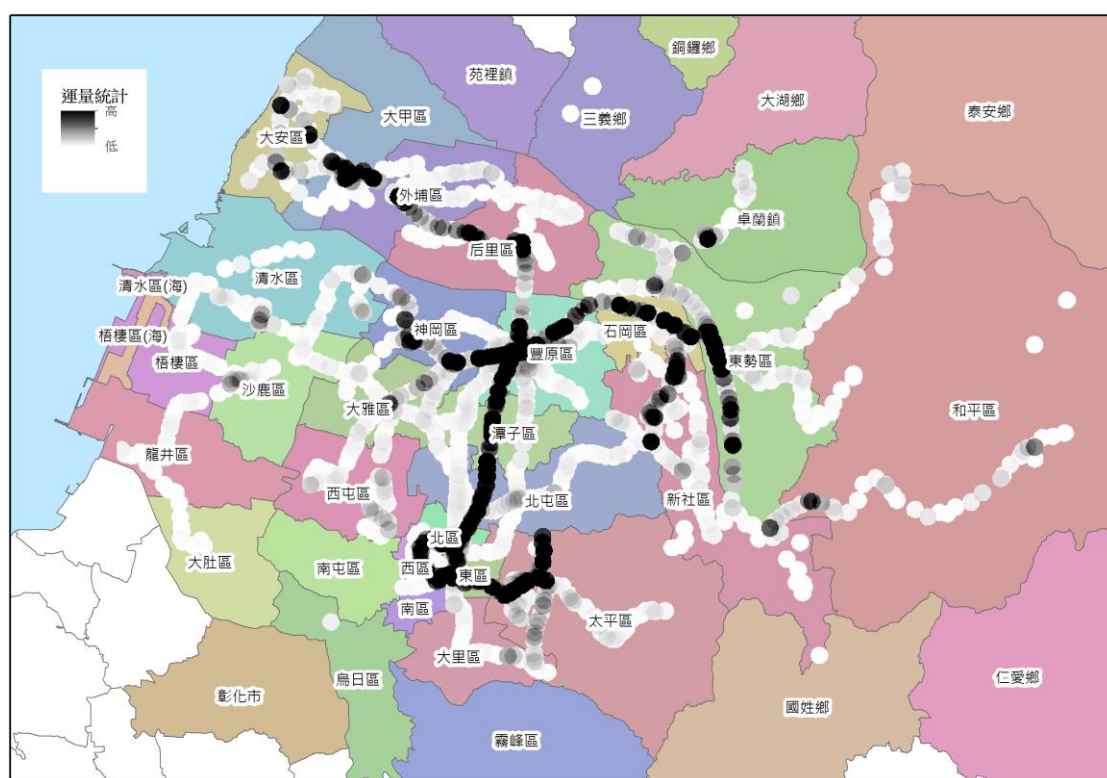


圖 6-7 路廊旅次示意圖



圖 6-8 整合地圖路廊旅次示意圖

3. 路線分析：若利用傳統表格式統計資料，雖可獲得各路線之統計數據（如圖 6-9），但站點分佈及公車行駛路線之空間關聯性並無法由表格式資料獲得，圖 6-10 路線分析即利用公車行駛路線空間資料結合各路線運量資料，以同時獲得空間及運量等資料，由圖中公車路線之站牌大小比較各路線之行駛方向及運量大小，其結果除可利用在公車路線調整、車隊車型調整之外，亦可比對路線行徑方向與空間/時間供給與需求量，以檢驗供需資料是否符合實際情況。

路線票種分析

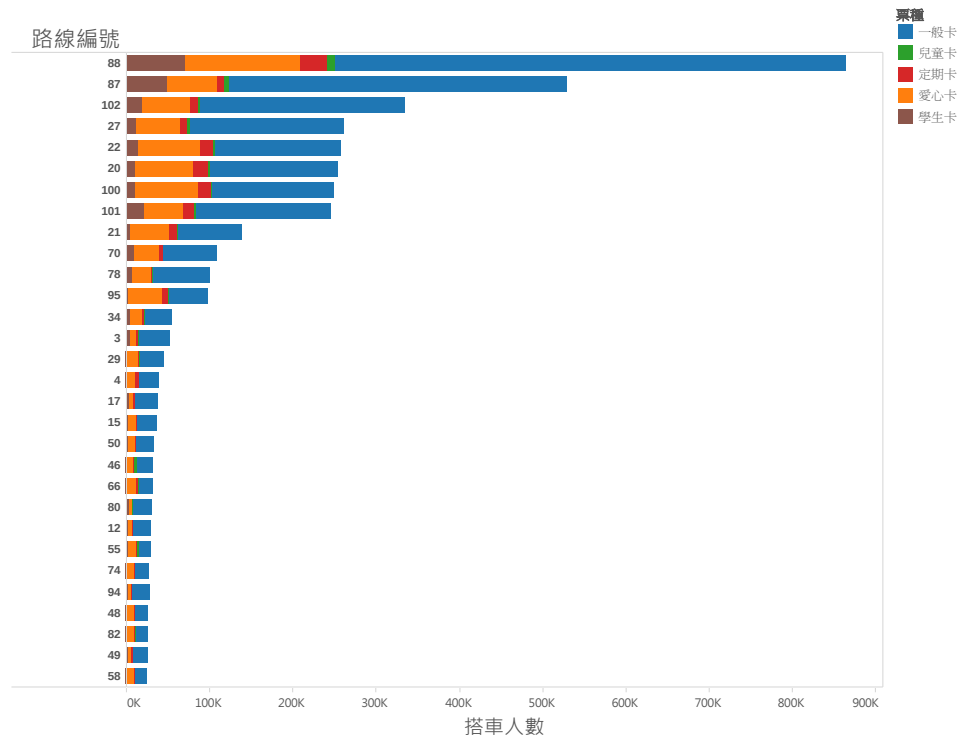


圖 6-9 路線票種分析示意圖

路線站牌運量分析

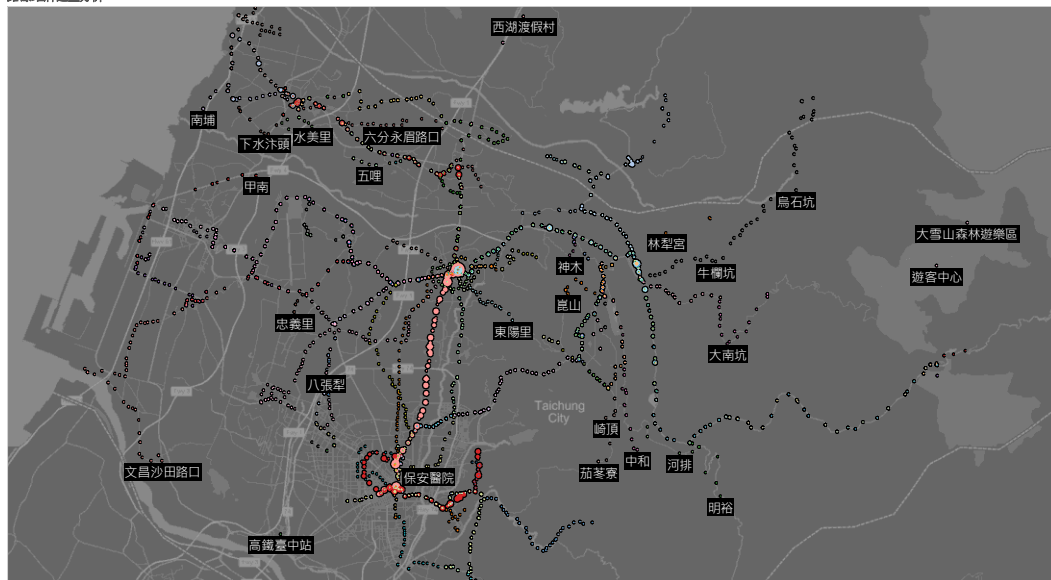


圖 6-10 路線站牌運量分析示意圖

- 站牌分析：利用傳統表格式統計資料，雖可獲得各站點之統計數據，但站位分佈及公車行駛路線之空間關聯性並無法由表格式資

料獲得，圖 6-11 站牌分析即利用站位座標及公車行駛路線，並利用圖形大小及圓餅圖方式表示其站點需求量及上下車比例，其分析結果亦可在候車人數較多之站牌改善候車設施。

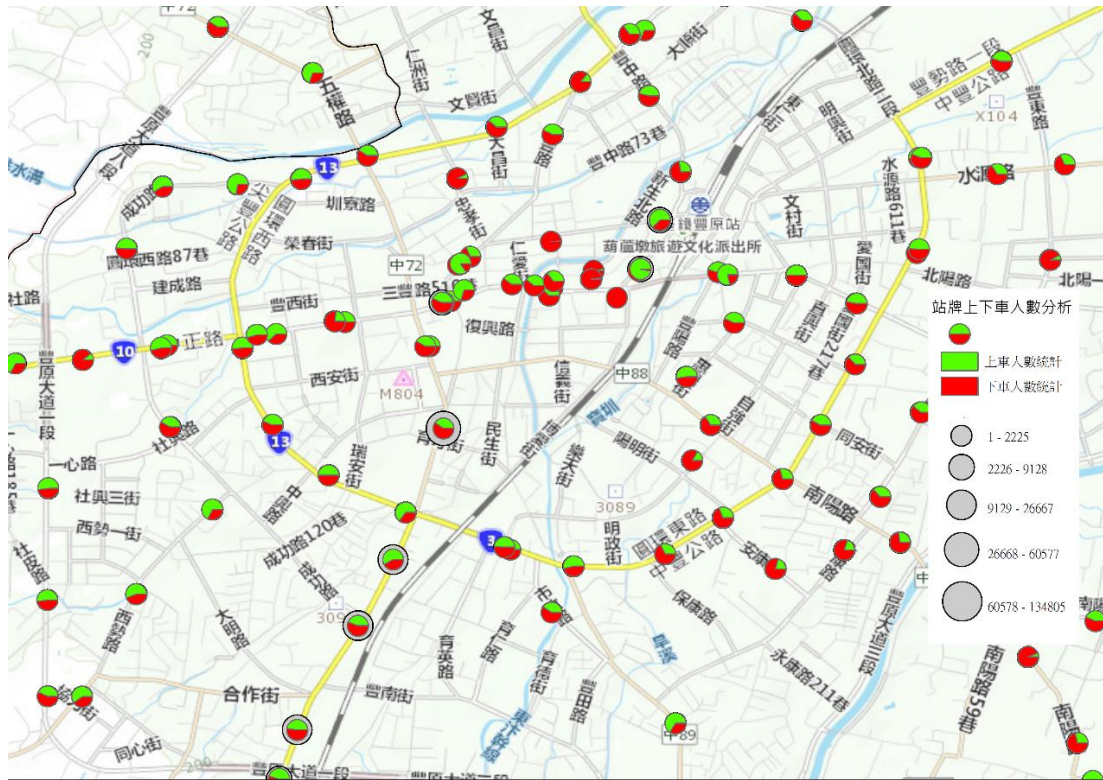


圖 6-11 站牌上下車人數分析示意圖

5. 使用者分析：主要分析使用者搭乘率及使用者行為，可視為已實現之需求資料，並可做為路線規劃及站點設置。圖 6-12 即為使用者搭乘數量統計，由資料中發現，搭乘次數前 1.1% 使用者，其使用者搭乘總數佔總電子票證數量 25%。圖 6-13 中顯示前使用量前 100 名之使用者，以一般卡及愛心卡，而使用量最高之使用者，其旅次行為如圖 6-13 所示，該使用者平均每天搭乘 6.5 次，並有明顯固定旅次行為，若多數使用者皆有相同之固定旅次行，管理者即可依此行為進行路線調整、站牌設置等，使公共運輸供給更符合實際需求。

使用者分析

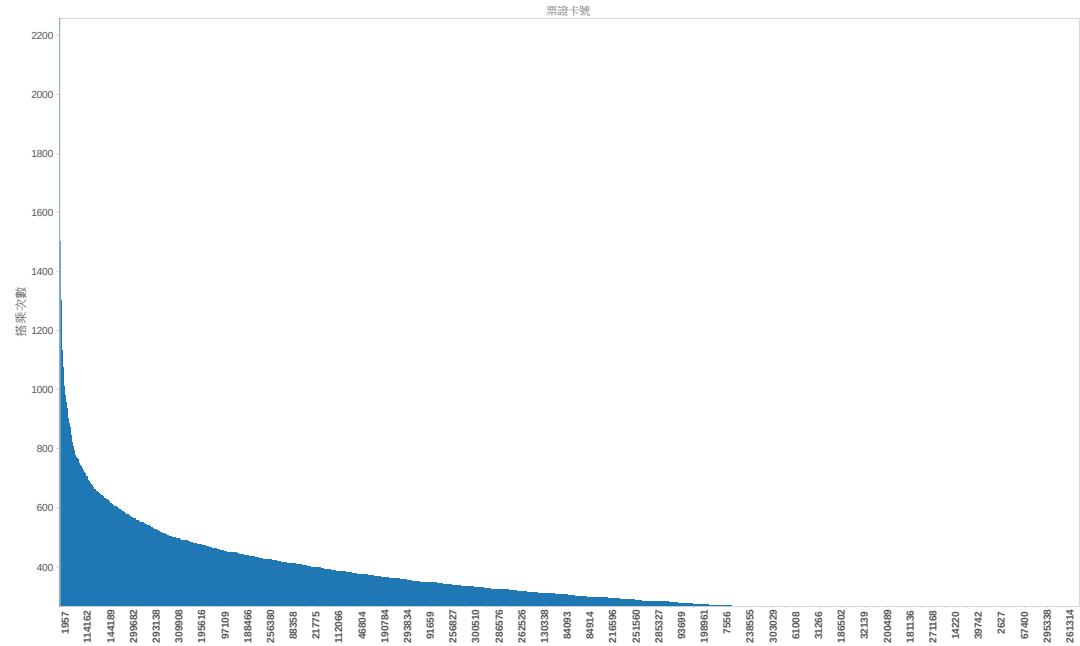


圖 6-12 使用者需求分析示意圖

使用者分析

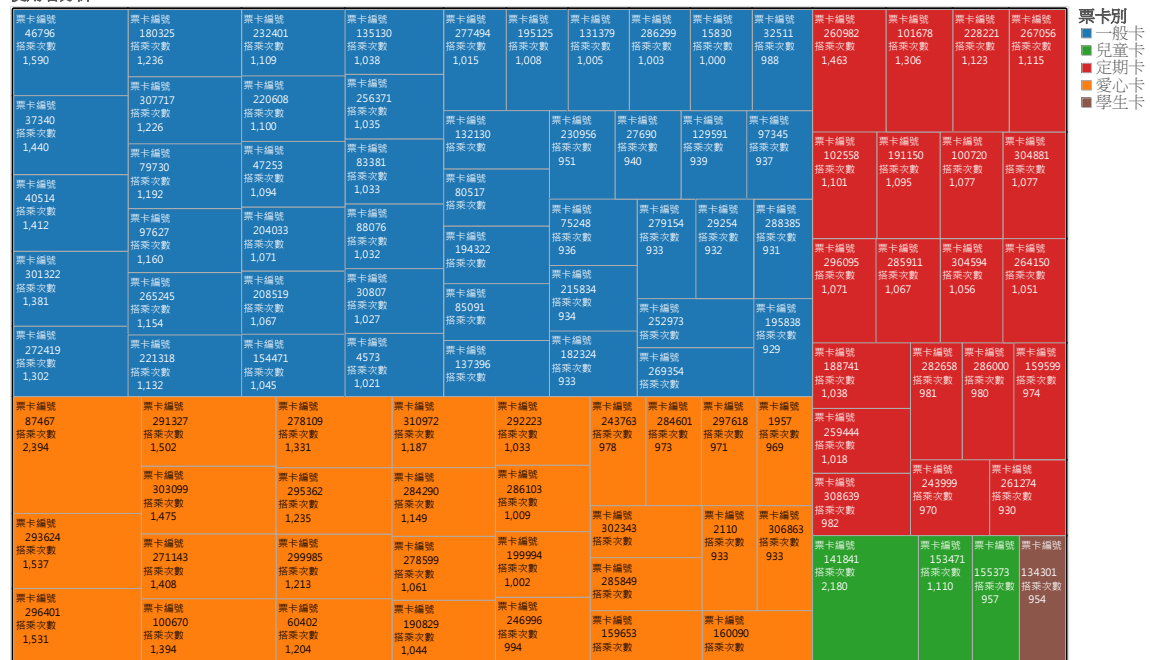


圖 6-13 前 100 位使用者需求量分析示意圖

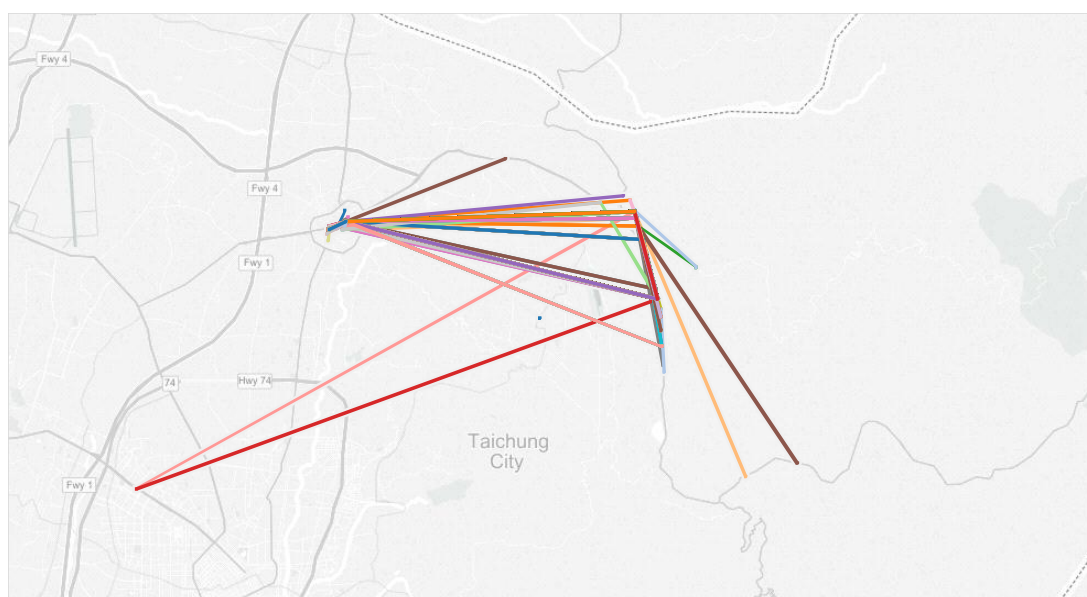


圖 6-14 前單一使用者旅次分析示意圖

6. 路線調整分析：路線乘載量若不均衡散佈代表路線經營效率差，未能載客所剩下的空座位數由於不能保留造成資源浪費，意味著路線設計上還有改善的可能性，最常用的改善策略為採取區間車營運方式或裁減路線來提高營運效率。如圖 6-15 及圖 6-16 所示，以路線 A 為例，明顯路線前後兩端乘載量較少，站序 2 站其後的乘載量急遽上升，而站序 51 後乘載量急遽降低，由於班次數最少要滿足發生於站序 11 至站序 12 最大乘載區間車上人數 476 人，這代表灰色斜線區域為運能大於運量的狀況，為了節省成本，提高單位投入獲得的效益，業者可評估開行站序 2 至站序 51 區間車的可能性。斷點一般選擇設在乘載量較小及上下車人數較大的場站，在該場站斷開路線，可相對減少路線客流的換乘。除此之外，決定裁減路線的評估原則還必須考量替代路線是否能滿足縮短路線段的運輸需求，以及公車場站滿足車輛停放與調度的需求才行。

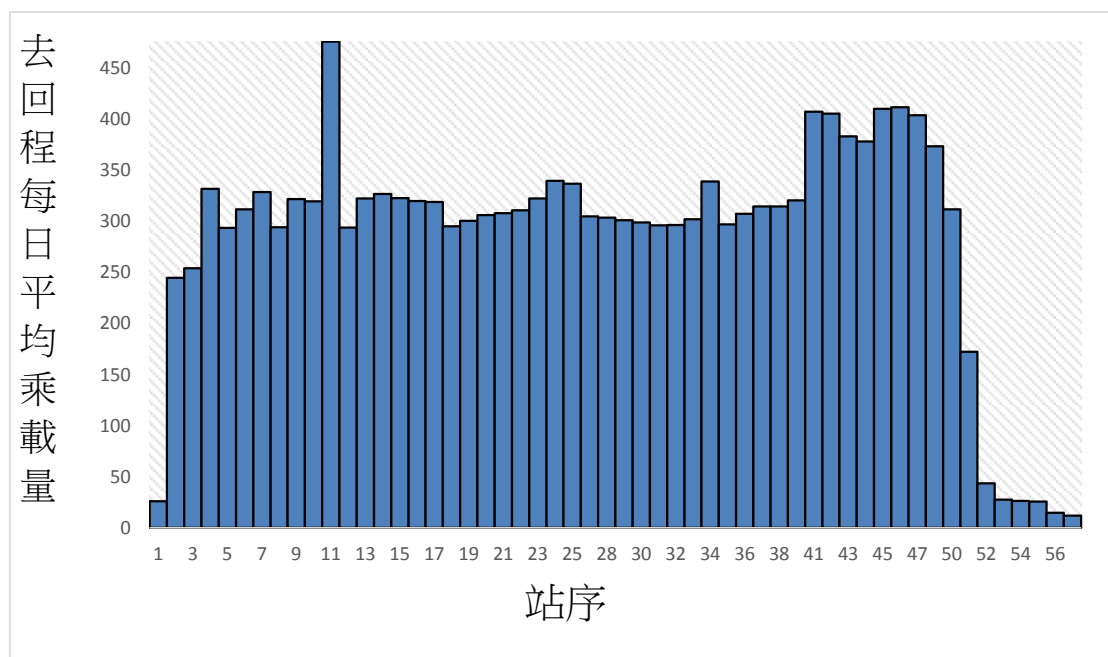


圖 6-15 路線 A 去回程各站牌每日平均車上人數(四月)散佈圖

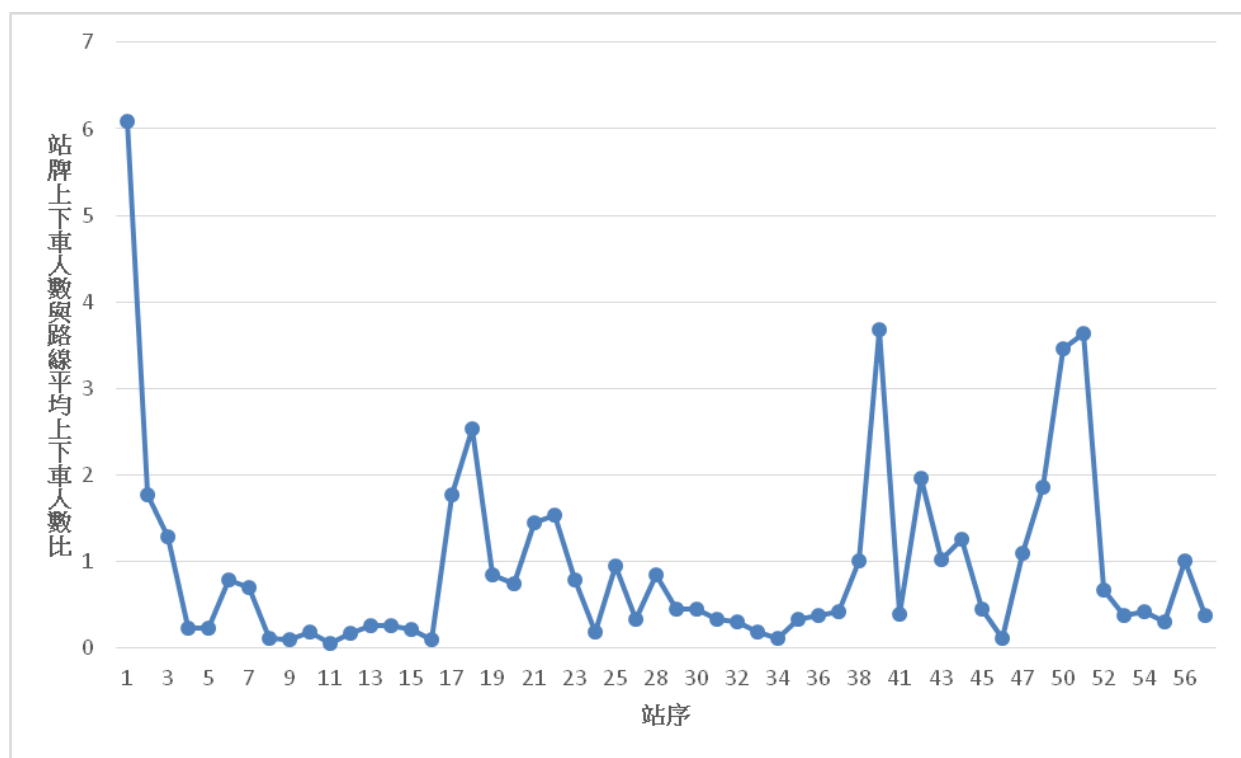


圖 6-16 路線 A 去回程各站牌上下車人數與路線平均上下車人數之比

7. 公車路線與計程車旅次之關係：研究團隊透過計程車派遣系統廠商所提供之訂單及計程車軌跡資料，利用 GIS 分析可知經篩選後 7429 筆訂單資料中，以計程車乘客上車位置週邊 100 公尺為搜尋範圍可得 621 筆乘車資料為搭乘公車直達可到之起點起訖點，僅站訂單資料的 8.3%，再篩選公車服務時間內之訂單筆數，其結果更僅為 432 筆資料(5.8%)，其最大旅次長度為 9 公里，最小為 504 公尺，平均旅次長度為 2.1 公里。另外，使用起迄點無公車服務之資料亦可找出不同訂單相同起迄之筆數，最多為 114 筆，表示在相同起迄點下，一年內共有 112 筆訂單記錄，此結果亦可做為路線調整之參考。



圖 6-17 計程車乘車上下車分佈圖

以豐客-東勢-谷關路廊為範圍，利用 SQL server 及 GIS 軟體進行空間與時間之分析。分析方式為以一週分時段為資料統計基礎，統計路廊中路線各站牌間每日各時段之上下車之搭乘人數，即可找到在路廊中不同路線，但相同起訖站中之乘客上下車數；再利用相同星期平均其乘車人數，最後利用 GIS 軟體以時間、空間方式將資料輸出成圖片或影片(圖 6-18、圖 6-19、圖 6-20)。即可得到不同星期、時間及路段之上下車人數，後續即可再利用初步成果透過 Tableau 等其它分析軟體做資料交叉分析以找出問題或特性，以做為管理決策之用。



圖 6-18 上午時段豐客-東勢-谷關路廊乘車人數



圖 6-19 下午時段豐客-東勢-谷關路廊乘車人數



圖 6-20 晚上時段豐客-東勢-谷關路廊乘車人數

6.2 路線與班次重疊分析

6.2.1 探勘分析目的

在過去公車路線審議時，審議委員大多僅能針對當下的路線進行審查，較難以全面性的方式檢核路線重疊之問題，尤其是市區公車與公路客運重疊之問題，因審查單位不同更能凸顯此問題。因此，透過公車資訊巨量資料分析，將各站點與路線資料進行建檔，透過資料探勘的方式進行分群，檢視屬性類似的路線，並透過視覺化的方式即可輔助路線審議單位避免路線過度重疊之問題。

6.2.2 演算處理流程

當取得各站牌、班表以及臺中市街道資料後，研究分析資料的流程如下：

1. 道路分段

取得街道資料後，由於要計算某時段內公車經過某路段的頻率，故將街道資料套入地理資訊系統軟體中進行路網結構分析，以路口做為節點，節點與節點間為一路段。

2. 公車路線

將取得的站牌資料中的路線、方向與站序資料新增欄位資料「路線_方向_站序」做為求取站點間最短路徑之路線名稱，將整理過後的資料套入地理資訊系統軟體中以站點至站點間進行最短路徑分析，分析的結果視為公車行進路線。

3. 圖層套疊

在地理資訊系統軟體中將分段後的道路圖層以及求取最短路徑後的公車路線進行套疊，以選取圖層的方式找出每一條公車路線所對應到的所有街道路段。

4. 班次分析

本研究以每小時某路段公車的經過頻率做為呈現結果，故將取得的班表以一小時為一單位分割。再整合每小時內所有公車班次所經過的路段，彙整出每小時內某路段上公車經過的頻率。

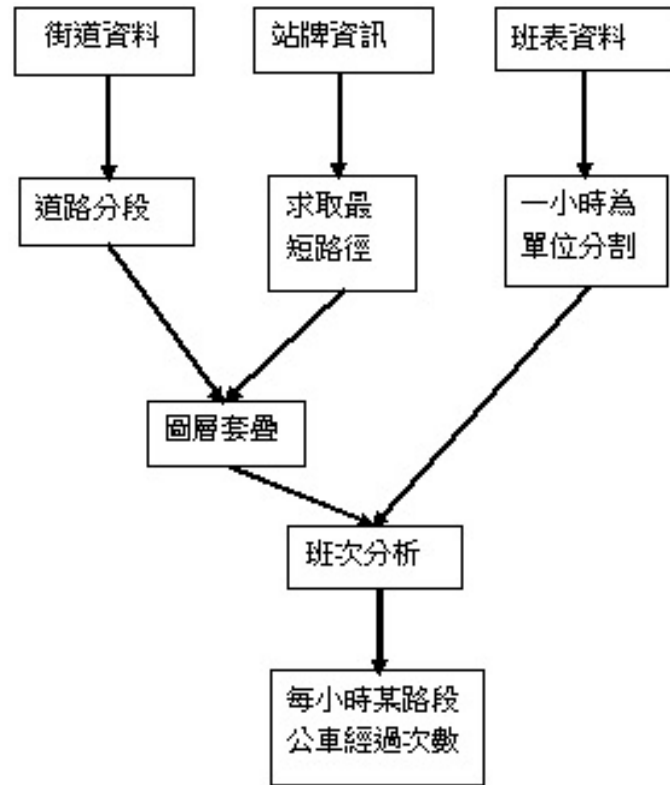


圖 6-21 路線與班次重疊分析流程圖

6.2.3 分析結果與視覺呈現

經由路線與班表重疊分析後，所得到的結果為每小時內某路段上公車經過的次數，表 6.2 為上午 16:00 至 16:59 所累加的部分結果，數據內容由原本的站點資料與班擴增出每小時路段上所通過的公車班次，呈現公車供給方面路線層級的資訊。

在視覺化展現的部分，以 16:00 至 16:59 為例，利用 GIS 軟體進行視覺化的呈現，以顏色區分累積次數的多寡，如圖 6-22，其中最高累積次數為 155 次。

表 6.2 路線與班次重疊表

道路編號	次數累加
6600000306467	155
6600000055248	151
6600000027989	151
6600000195951	145
6600000295572	131
6600000306598	129
6600000295496	125
6600000295761	122
6600000295449	117
6600000054662	116

資料來源：暨本研究整理

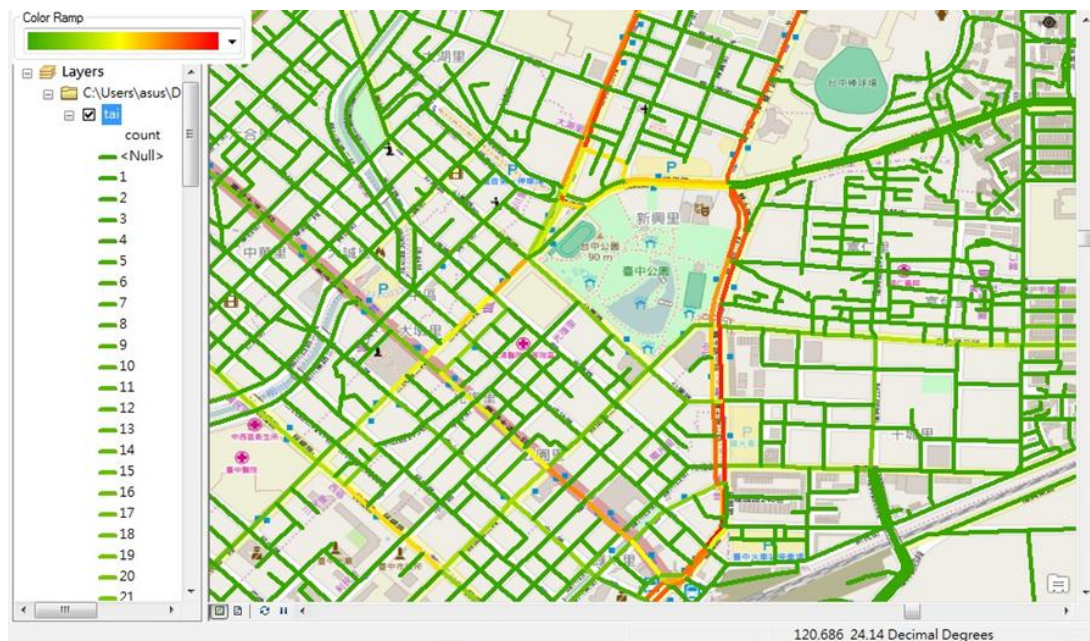


圖 6-22 小時次數累加示意圖

圖 6-23 為臺中市道路一天的累加次數，綜合兩張累加圖，我們可以得知從臺中火車站經由雙十路一段至一中商圈的路廊供給相當密集，一天的最大公車經過累積班次數大於 1000 班。表 6.3 為計算路段平均公車經過班次，其中含 0(累積次數為 0)為所有路段總和，包含公車路線沒有經過之路段；不含 0，則表示只計算有公車路線經過之路段。此外，平均每路段之班次數，是以公車路線經過之路段數乘上班次數再除以總路段數，得到之結果為每天每一路段會有幾班公車經過。

其計算結果以有公車經過之道路而言，一天平均會有 127 班公車經過。但在特定路段卻有大於 1000 的班次，這也顯現了路線班次重複率較高的問題！

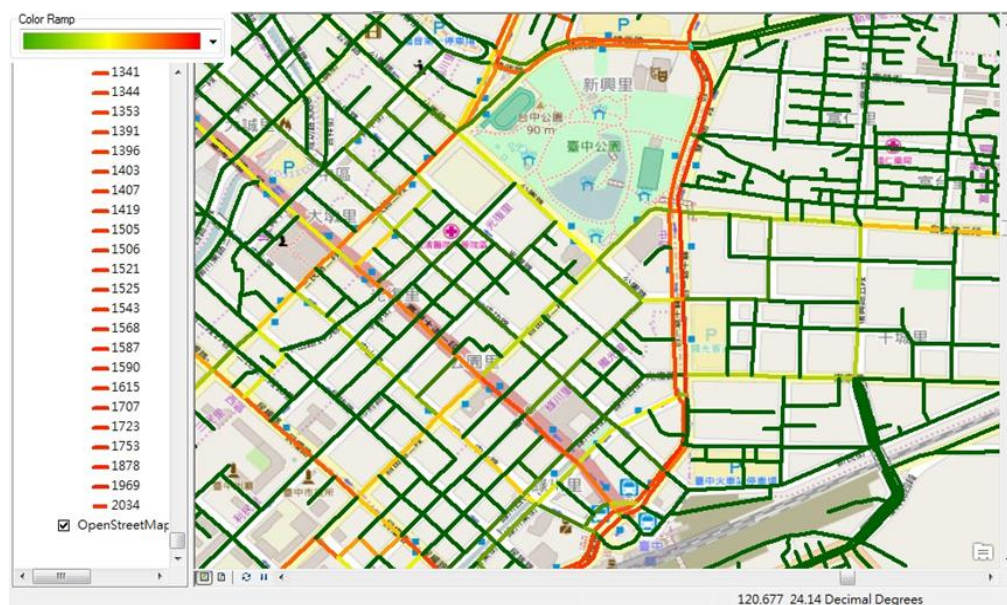


圖 6-23 全天次數累加示意圖

表 6.3 路段累積班次統計表

	路段數	路線*經過路段數*班次	平均每路段經過班次數 (班/段)
含 0	148880	1824563	12.26
不含 0	14342	1824563	127.22

資料來源：本研究整理

6.3 供需分析

透過供給與需求分析可了解目前公車提供之服務是否可滿足現況之需求，進一步藉由時間與空間的切割，更可得知公車之路線及班次是否滿足特定的族群（如學生及高齡者）。基此，此節將探討供需分析之目的、目前蒐集到之資料、供需分析演算流程以及視覺呈現暨應用。

6.3.1 探勘分析目的

臺中市區內包含 29 個行政區，然各行政區人口結構及汽車客運供給狀況均不相同，為使各區域之供給與需求達到公平的目的，並了解公車之供給是否可滿足學生族群及高齡族群，以建議相關單位規劃路線及設計班次時，可將有限的資源提供給需要的族群。交通部運輸研究所過去委託中華大學及交通大學執行「先進公共運輸系統整合資料庫增值應用與示範計畫(2/2)」、「國家區域公共運輸服務指標調查示範計畫 (2/2)」，然這兩個計畫礙於資料限制，僅分析各行政區之供給及總人口的關係，並未針對學生族群及高齡族群進行分析。本研究藉由巨量資料的蒐集，針對研究區域中之汽車客運供給(路線分布、班次)與需求（如乘客旅次起迄資料）在不同時間與空間的分布狀況進行分析。

研究供給與需求探勘分析目的說明如下：

1. 了解汽車客運供給與旅次需求的關係，著重於高齡人口與學生人口之供需狀況。
2. 分析實際運量與旅次需求間之關係，探討各行政區公車運量與旅次需求之比例。
3. 分析供給與預測旅次需求之關係，藉由臺中市運輸規劃預測之結果，以了解目前汽車客運之供給是否能滿足未來之運輸需求。

6.3.2 演算處理流程

本研究供需分析演算處理流程如圖 6-24 所示。此演算處理流程中，包含 6 個資料庫，分別為人口統計資料、路線資料、班次資料、車輛資料、運量資料及預測需求資料。整合各項資料進行資料演算處理，可得到各區全日供給與需求比例、各區分時段供給與需求比例、各區全日運量與需求比例、各區分時段運量與需求比例、各區各時段 OD 運量、各區分時段運量與需求比例、各區各時段 OD 運量(分票種)、各區全日供給與預測需求比例及各區分時段供給與預測需求比例等結果，各項產出說明如下：

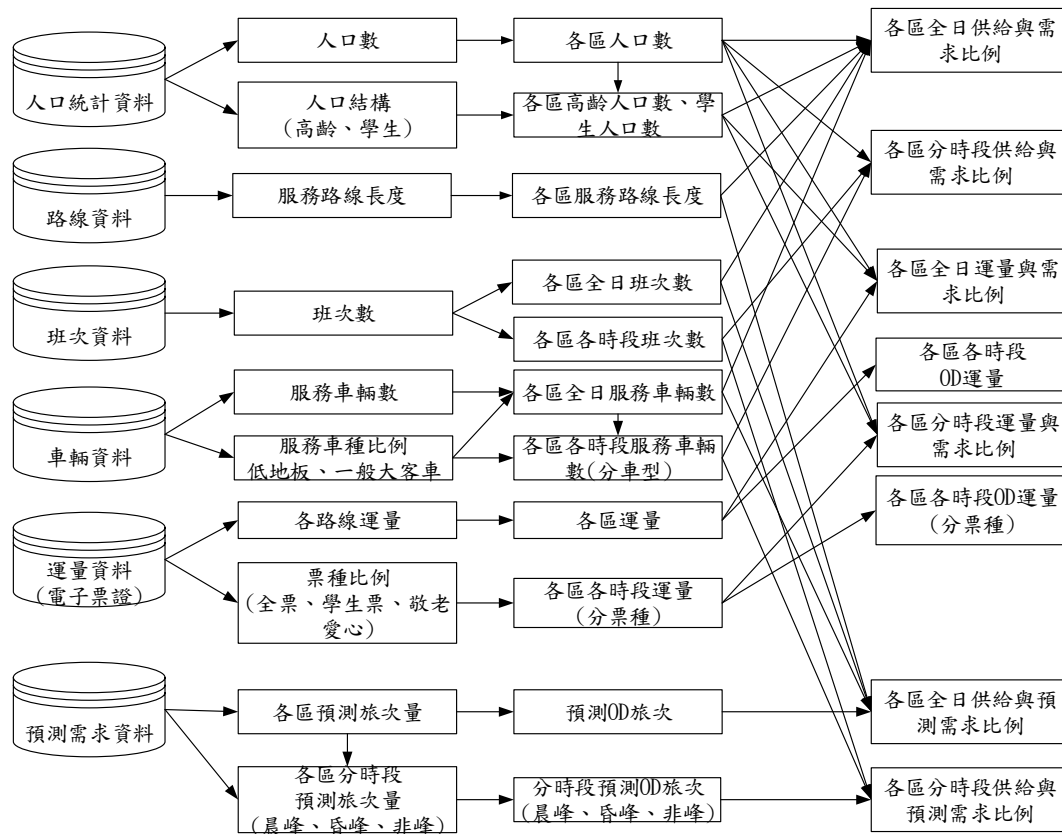


圖 6-24 供需分析演算處理流程

1. 各區全日供給與需求比例

此項指標係結合人口數與人口結構資料，推估各行政區高齡人口數及學生人口數。進一步結合路線資料、班次資料及車輛資料估算各行政區每日班次數、路線長度及服務班次數，其中，服務班次數可參

考車輛資料庫進一步探討車種之服務比例。結合前述各項資料，各區全日供給與需求比例可產出「各行政區每人享有路線數」、「各行政區每人享有班次數」、「各行政區每人享有路線長度」等指標，另各項指標參考人口結構亦可細分為「各行政區每學生/高齡者享有路線數」、「各行政區每學生/高齡者享有班次數」、「各行政區每學生/高齡者享有路線長度」等指標。

2. 各區分時段供給與需求比例

各區分時段供給與需求比例係參考前述之結果，進一步將供給與需求根據時間進行劃分，以了解尖峰與離峰之供需特性。另外，不同旅運者(學生及高齡者)特性，其旅次發生時間亦不同，因此，本研究參考臺北市捷運工程局委託執行之「臺中都會區住訪調查及運輸需求預測模式檢討與建立」之調查結果推估各時段需求。基此，此部分各區分時段供給與需求比例可產出「各行政區各時段每人享有路線數」、「各行政區各時段每人享有班次數」、「各行政區各時段每人享有路線長度」等指標，另各項指標參考人口結構亦可細分為「各行政區各時段每學生/高齡者享有路線數」、「各行政區各時段每學生/高齡者享有班次數」等指標。

3. 各區全日(分時段)運量與需求比例

分析運量與需求比例之目的在於分析搭乘汽車客運人數占總旅運需求之關係。過去運輸規劃分析中，欲了解汽車客運市占率，大多透過抽樣調查的方式進行。近年來電子票證技術成熟，透過電子票證可以調查公共運輸搭乘情形。故此部分將分析各區全日(分時段)運量與需求比例，以了解汽車客運於該行政區之市占率。

4. 各區 OD 運量

臺中市目前係採里程計費方式，故藉由電子票證資料之分析，可了解各行政區公共運輸之 OD 運量。電子票證記錄欄位亦包含票種，故可藉由票種的分析，探討學生與高齡人口旅次發生時間，已調派適當的車型。

5. 各區全日(分時段)供給與預測需求比例

前述各項資料推估係藉由現況資料進行分析，然對於未來的需求仍有推估分析需要。故本研究參考臺北市政府捷運工程局委託執行之「臺中都會區住訪調查及運輸需求預測模式檢討與建立」之預測模式所推估之旅次發生數及旅次吸引數，推估各行政區之旅次需求及旅次OD分佈，進而了解目前的各項供給是否可滿足未來之需求。

6.3.3 視覺呈現暨應用

根據前述演算處理流程，各項分析項目視覺呈現如下：

1. 各區全日(分時段)供給與需求比例

各區供給與需求比例視覺化呈現如圖 6-25 所示，本研究以各區全日供給為最上層之分析結果，藉由圖 6-25 可清楚了解各行政區班次數服務情形，其值越高表示該區每人/學生/高齡者享有之班次數服務則越多。進一步比較各區、各時段供給量與潛在需求量之分佈，可發現比值愈高公車服務愈便利。此外，中區因臺中車站位於此區，公車路線均以此地為起訖點或行經站點，故每人享有班次數全市最高。比值低之行政區表示該區公車服務便利性相對較低，例如，西屯、北屯、清水、沙鹿等。

另外，可進一步利用分時段之資料分析各時段之供需情形如所圖 6-26 示，此圖係以臺中市區為例，以總人口數來說，各時段供給與需求比偏低，但以學生或高齡者來說，在部分時段則有供給大於需求的情形。將人口數依時段按運量票種比例劃分，觀察不同時段的供需比。此外，本研究發現尖峰時段供需比較低，而離峰時段則較高，其中，學生族群於離峰時段供需比高，但尖峰時段則較低，可考量調整部份班次，以提高尖峰時段供給。尖峰時段高齡族群之供需比低於學生族群，顯示該時段公車使用者主要以高齡者為主，於該時段應進一步分析高齡者旅次目的，改善其搭乘方便性。另外，觀察中區學生人口全日班次數與總人口數(千人)之比值約 213(班次數/千人)，而分時段最高達約 5,234(班次數/千人)，其原因為非上下學時間，學生 10 時-11

時及 13 時-14 時分別占全日旅次 0.68%及 0.67%，顯示學生需求很低，但班次仍需正常供給服務其他族群之使用者，故造成全日與分時有較大的差異。

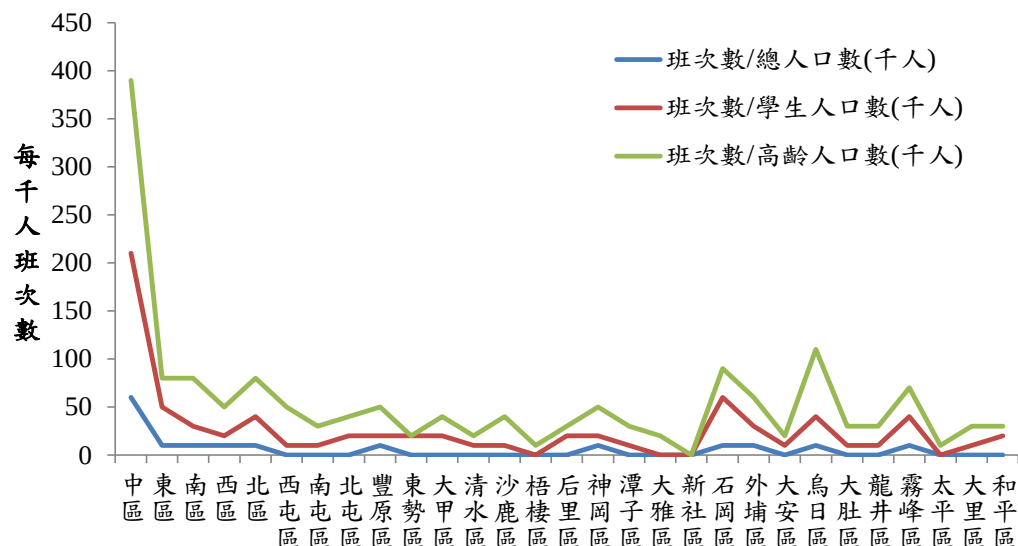


圖 6-25 各區全日供給與需求比例視覺呈現(以班次數為例)

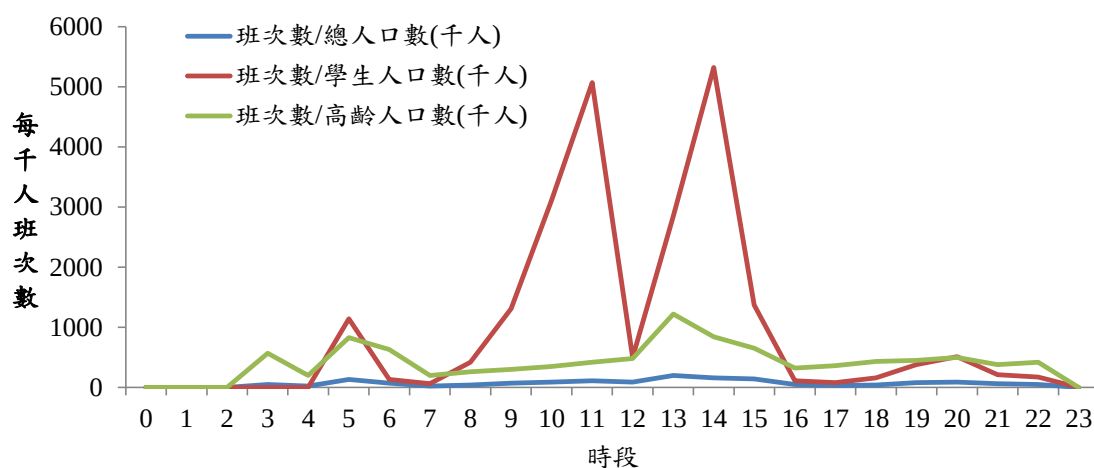


圖 6-26 各區分時段供給與需求比例視覺呈現(中區)

2. 各區全日(分時段)運量與需求比例

各行政區運量與需求比例視覺呈現如圖 6-27，其中，運量係指電子票證刷卡次數，需求則以人口數作為需求。圖中比較總運量、學生運量及高齡者運量與其需求之比例，藉由此結果可分析該行政區公車

市占率之情形。囿於本研究僅採豐原客運電子票證資料，運量集中在豐原客運營業區域，因此，圖 6-27 可了解該地區使用公共運輸之比例。並結合前述供需比之結果，可了解運量與需求之比例與供需比之關係，進而了解需改善的地區。由分析結果發現，運量與需求比例以中區、北區、東勢區、豐原區及石岡區較高。以豐原客運運量資料為例，此結果說明豐原客運在各區的市占率，可提供客運業者進行市場區隔或路線規劃時參考。

進一步探析中區分時段運量與需求比例，如圖 6-28 所示，藉由此圖可觀察高齡者與學生族群之運量比例於各時段均有不同之趨勢，如高齡者於上午 8 點至下午 4 點之比例較高，而學生族群則集中在晚上 6 點至晚上 10 點之間。由此可知，中區尖峰時間公車市占率約 5%-15% 間，而離峰時間約 20%-25%，顯示該區仍未將公車作為主要的通勤工具。但在離峰時段，高齡者有很高的比例使用公車。

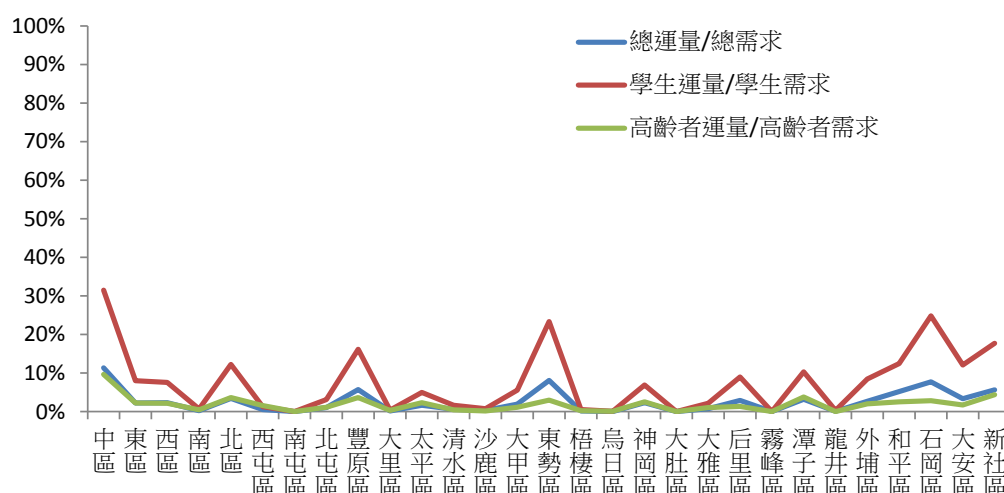


圖 6-27 各區全日運量與需求比例視覺呈現

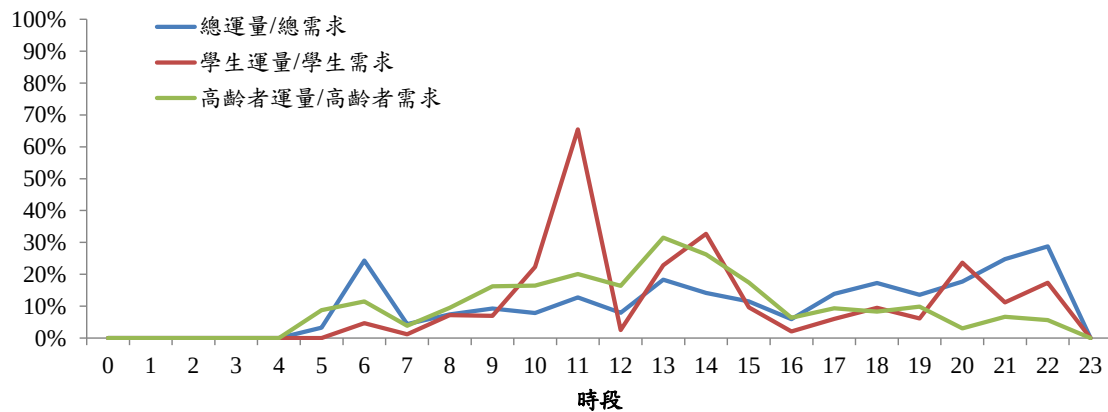


圖 6-28 各區分時段運量與需求比例視覺呈現(中區)

3. 各區 OD 運量

本研究蒐集豐原客運電子票證資料，亦可藉由電子票證之分析，透過視覺呈現方式，了解豐原客運各區 OD 之運量，如圖 6-29、圖 6-30 及圖 6-31 所示，藉由各區 OD 運量視覺化表示，可以了解搭乘豐原客運之起訖點大多為行政區內之旅次(顏色較深之區塊)，此資訊可做為業者路線規劃或行銷策略安排時之運用。由此分析結果可得知，大多運量 OD 為同一行政區內或鄰近行政區。此外，本研究採以豐原客運全日運量資料進行分析，故由圖可觀察到大多旅次之起訖點為豐原區，顯示豐原客運主要客源為豐原區之居民。

地區	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
神岡區	15	2	0	1	82	16	0	8	545	33	38	83	38	10	0	5	0	503	0	23	1	0	1	22	9	0	0	26	29	0
大肚區	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	3	0	0	7	0	0	0	0	0	0	0	0
大雅區	4	0	0	0	52	166	0	3	199	0	0	53	19	0	0	0	0	33	0	163	0	18	0	0	0	0	0	0	0	0
后里區	0	0	0	0	2	0	0	1	663	0	0	0	0	261	4	0	0	1	0	0	0	0	0	117	0	0	0	2	8	0
霧峰區	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
潭子區	139	17	78	3	798	28	0	418	1069	0	0	0	0	0	3	0	0	2	0	17	0	849	0	0	0	0	0	0	0	0
龍井區	0	0	0	0	0	0	0	0	2	0	0	0	85	0	0	0	0	1	7	0	0	2	0	0	0	0	0	0	0	0
外埔區	1	0	0	0	1	0	0	0	63	0	7	0	0	233	12	0	0	3	0	0	115	0	0	368	0	126	0	3	40	0
和平區	0	0	0	0	0	102	0	0	35	0	0	0	0	0	276	0	16	0	0	0	0	0	0	0	1	0	0	0	0	0
石岡區	3	0	0	0	2	25	0	1	456	0	0	0	0	0	350	0	0	0	0	0	0	0	0	1	3	282	0	54	0	0
大安區	0	0	0	0	0	0	0	0	7	0	0	0	0	255	2	0	0	17	0	0	2	0	0	4	0	0	371	0	0	0
新社區	63	5	0	0	52	0	0	94	187	3	1	5	0	0	250	0	0	20	0	0	3	0	0	29	1	60	0	652	0	0

圖 6-29 OD 運量視覺呈現示意圖(總運量)

[illegible]

4. 各區全日(分時段)供給與預測需求比例

以預測需求分析之供需關係，如圖 6-32 所示，其採用「臺中都會區住訪調查及運輸需求預測模式檢討與建立」研究所推估之旅次發生數作為預測旅次之需求。另外，觀察分時段之趨勢，如圖 6-33 所示，由圖可發現，於晨峰時間預測需求量較大，然班次之供給較少，可檢視於該時段是否有供給不足之現象。

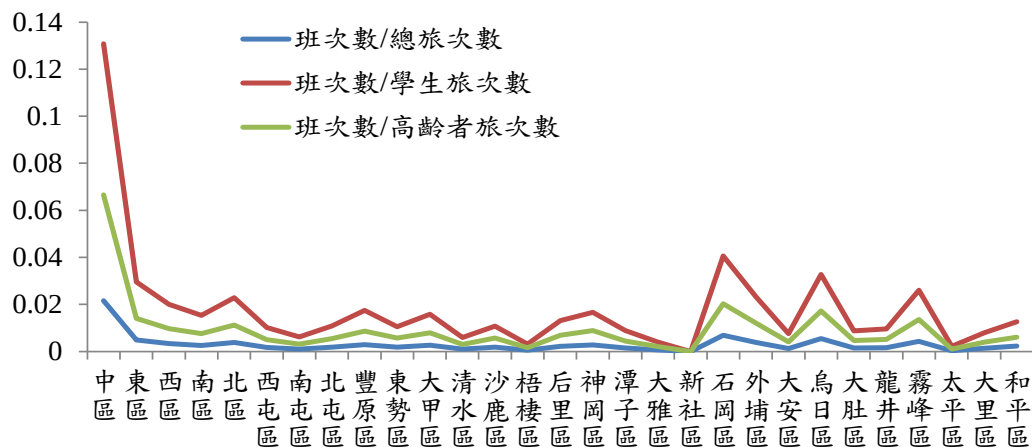


圖 6-32 各區供給與預測需求比例視覺呈現

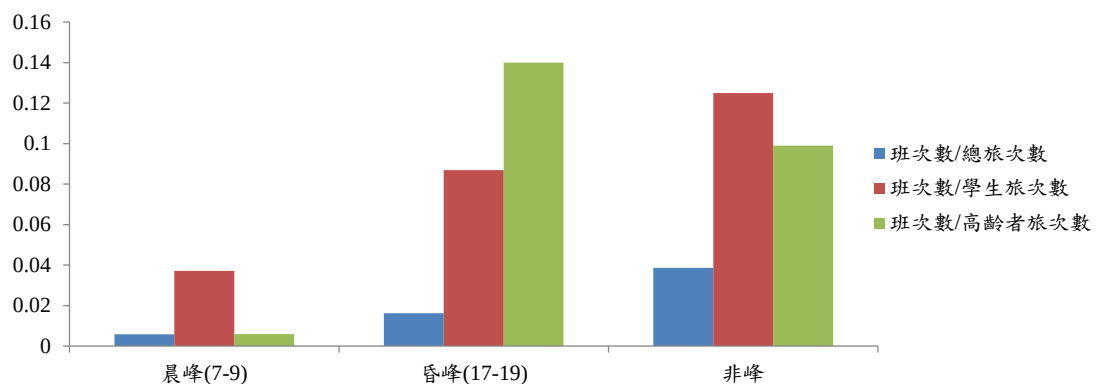


圖 6-33 各區分時段供給與預測需求比例視覺呈現(中區)

5. 搭乘里程長度分析

搭乘旅次長度為制定或分析票價結構時重要的參考依據，故本研究利用電子票證資料分析臺中地區平均旅次長度。本研究採用豐原客運電子票證資料，旅次長度分布如圖 6-34 及圖 6-35，由圖可知大臺中地區平均旅次長度為 10.9 公里，進一步分析各行政區平均旅次長度，以烏日區最長(75.08 公里)，其原因為烏日區路線包含高鐵站至谷關路線，使得其路線長度最長。

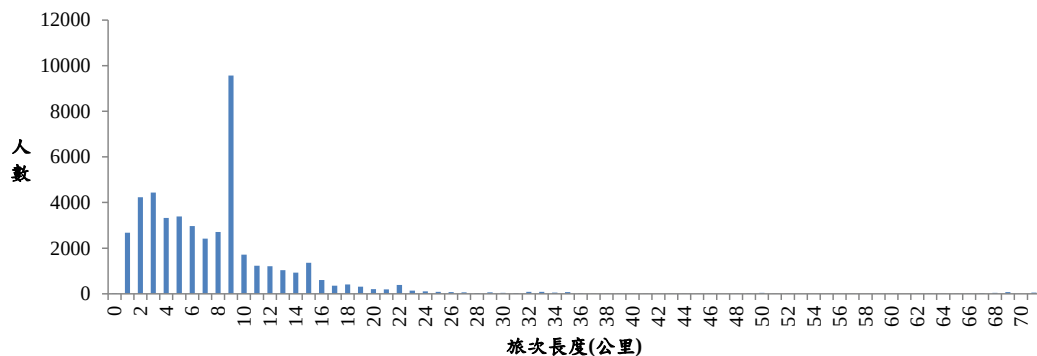


圖 6-34 大臺中地區旅次長度分布

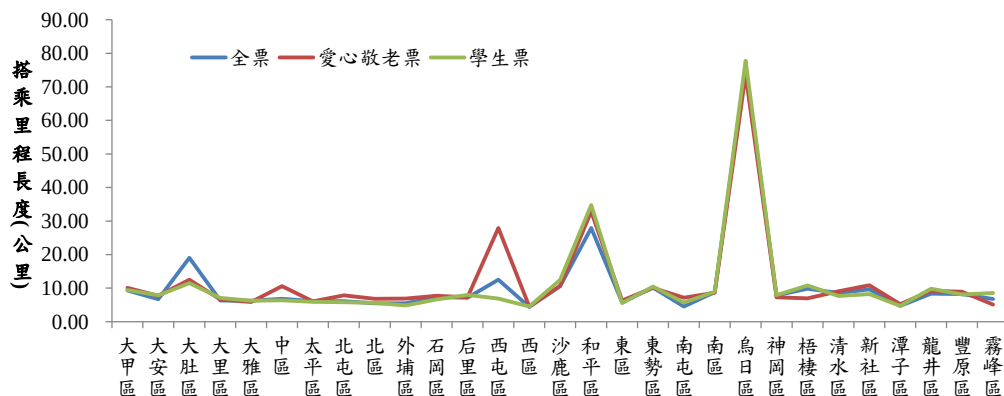


圖 6-35 各行政區旅次長度

6. 老年人、殘疾人士使用低地板公車比例

本團隊在過去公共運輸相關政策參與中，發現各業者均致力於購買低地板公車服務行動不便之老年人及殘疾人士。然目前尚無相關統計分析，低地板公車行駛路線，是否可以滿足目標族群使用。本研究以結合電子票證資料(敬老愛心票)及車型資料如圖 6-36，探討持有愛心敬老票之民眾，搭乘低地板公車之比例。分析結果，以豐原客運為例，如圖 6-37 所示，整體而言，民眾接受到低地板公車服務比例約 25%，其中，僅有 3%持敬老愛心票搭乘公車的人接受到低地板公車的服務如表 6.4。另外，本研究發現，低地板公車佔豐原客運車隊車輛數約 9%，而其服務人數為 25%，顯示低地板公車服務路線多為黃金路線，而偏遠低區路線則較少使用低地板公車服務，且偏遠地區老年人比例亦較都市地區高，可能使得需要低地板公車的族群無法收到應有的服務。

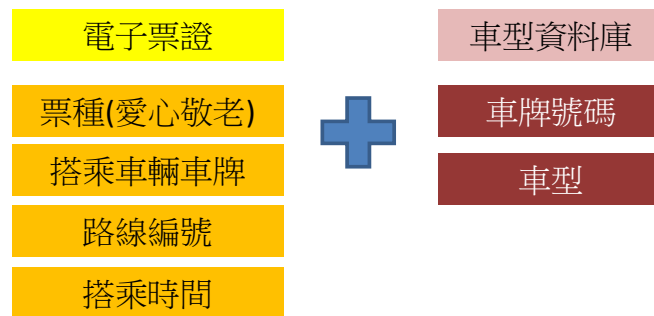


圖 6-36 低地板公車分析架構

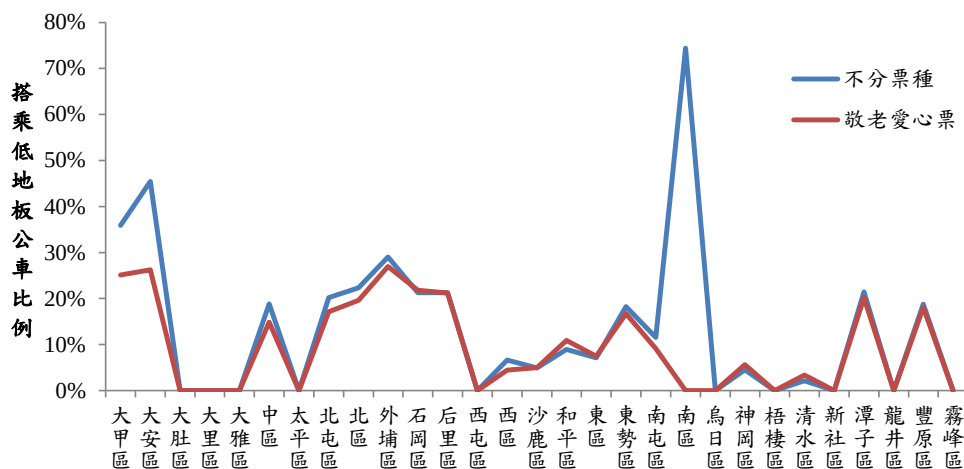


圖 6-37 低地板公車搭乘比例

表 6.4 各車型搭乘比例 (豐原客運為例)

車型	全票	愛心敬老票	學生票	總計
低地板公車	19%	3%	3%	25%
一般大客車	55%	13%	7%	75%
總計	74%	16%	10%	100%

6.4 結合電子票証與公車動態之分析

將電子票証與公車動態的資訊加以整合，可以分析在每一時段中車輛的密集程度與車上人數的關係，據此可以作為班距調整的基礎，讓有限的運輸資源發揮最大的效果。

本研究針對本部分的需求，設計如下圖（圖 6-38）之視覺化呈現界面。



圖 6-38 結合電子票証與公車動態之視覺化介面

左半部（範圍 1）主要呈現路線的發車軌跡，縱軸由上而下代表時間軸，橫軸由左而右代表該路線的各站牌名稱。從這部份的圖形，可以清楚觀察到該路線每班次的行駛軌跡。線條越陡，代表行駛速度越慢；線條越靠近，代表兩班次越接近。在本部分的路線軌跡中，特別使用不同的顏色代表車上的載客量，分別代表車內人數小於 30 人（均有座位）、大於 30 人小於 50 人（雖有人站立，但還不擁擠）、以及大於 50 人（車內擁擠），如圖 6-39 所示。

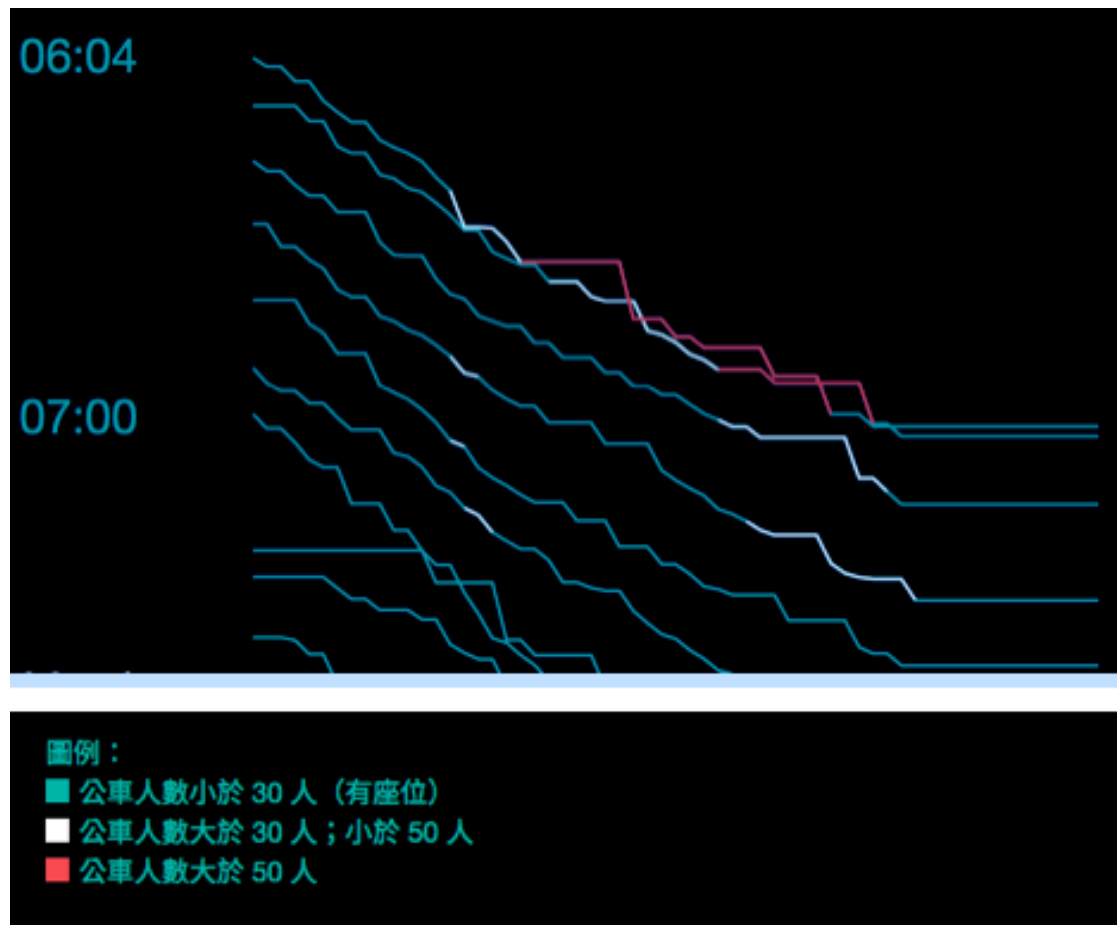


圖 6-39 行駛軌跡與載客量

右上部分（範圍 2）主要呈現路線的地圖，該路線的軌跡會特別以顏色標示，當滑鼠游標移動到左半部時，會在地圖上顯示該路線在該時間點所有公車的位置，如圖 6-40 所示。



圖 6-40 行駛路線與車輛位置

右下部分（範圍 3）為背景資料的呈現區域，包含路線名稱、日期、班次行駛時間、與平均行駛時間的差異等，如圖 6-41。



圖 6-41 行駛班次之基本資訊

根據以上的視覺化呈現介面，可以看出一些有用的資訊。如圖 6-42 所示，可以發現 55 路公車，在晚間 8:00~10:00 時段，於一心市場到北新國中之間的车上人數相當多，甚至有些班次之车上人數超過 50 人，建議可以在該時段加強輸運。

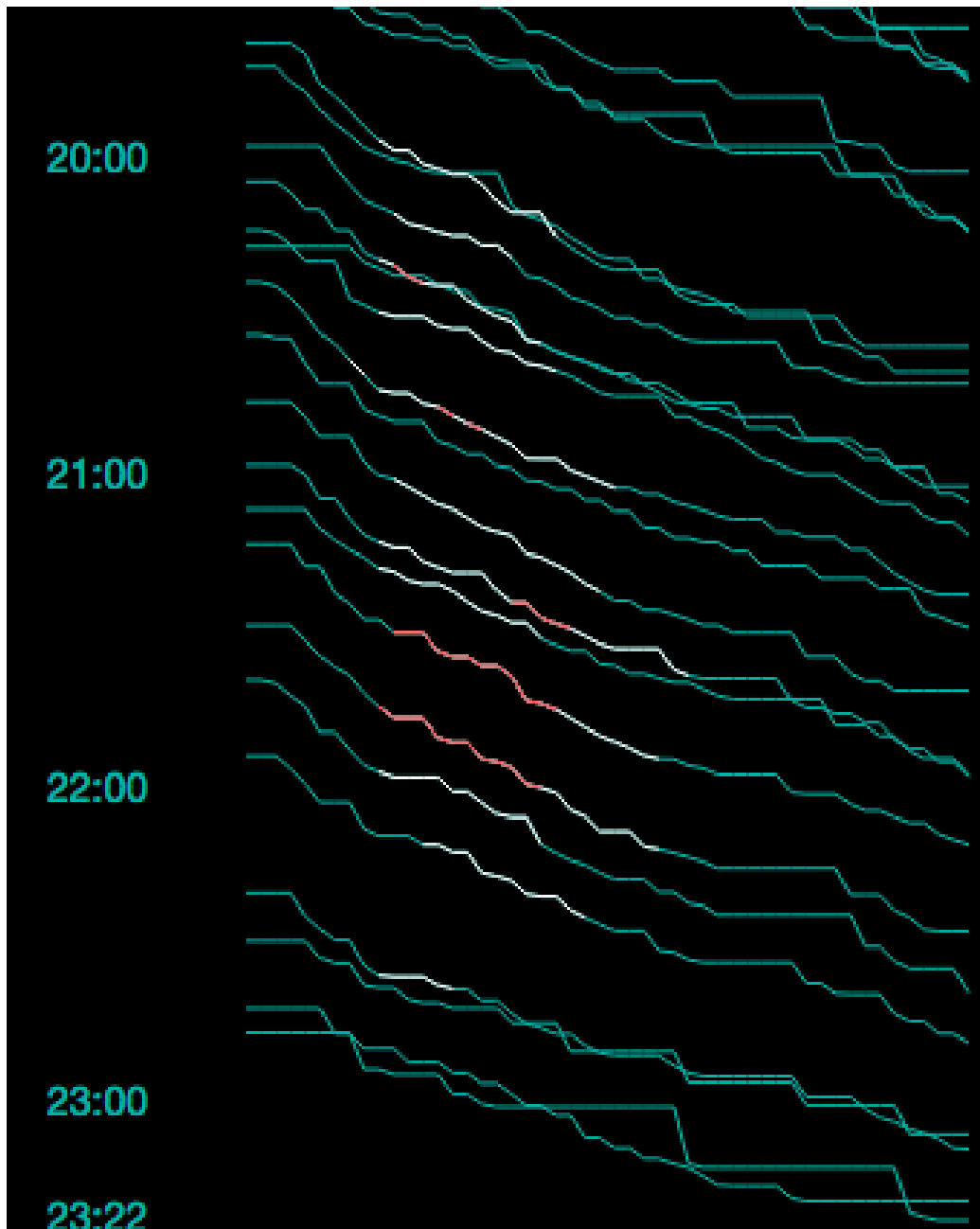


圖 6-42 豐客 55 路公車一心市場到北新國車載客情形

6.5 行車時間與班表分析

6.5.1 探勘分析目的

公車運輸系統存在的目的在於將旅客從 A 站輸送至 B 站，除最基本的安全需求外，代表運輸效率的「行車時間」是主管機關暨社會大眾最關心的議題，包括：

- (1) 各路線總行車時間
- (2) 路線中通過各站牌時間
- (3) 各種影響因素（月份/時段/氣溫/雨量）
- (4) 預估到站時間之準確性暨影響因素

「行車時間」僅討論每一車輛各別行車狀況或績效，而「班表分析」近一步探討車輛與車輛之間的關係，通常會關心以下課題：

- (1) 班次總數（含平均班距）
- (2) 班距變異（尖離峰差異）
- (3) 連班（含多輛追越、兩兩相互追越）等
- (4) 從路線、時段、空間（站牌）等維度分析連班樣態
- (5) 分析班表中的乘載情形

6.5.2 演算處理流程

本分析中較重要的流程包括：

1. 根據氣象測站位置決定每個站牌欲參考之測站

由於臺中地區幅員較大，氣象局計有以下兩個觀測站資訊，因此先計算站牌與每一測站距離，找出最近的測站資訊，圖 5-2 中 TaichungSTOP 資料表約有兩萬餘筆站牌，扣除重複位置者約 6 千百餘站牌可大致分屬兩個觀測站如圖 6-43。

- (1) 臺中：編號 17131 位於臺中市精武路 295 號
- (2) 梧棲：編號 4125 位於臺中市臺灣大道 10 段 2 號

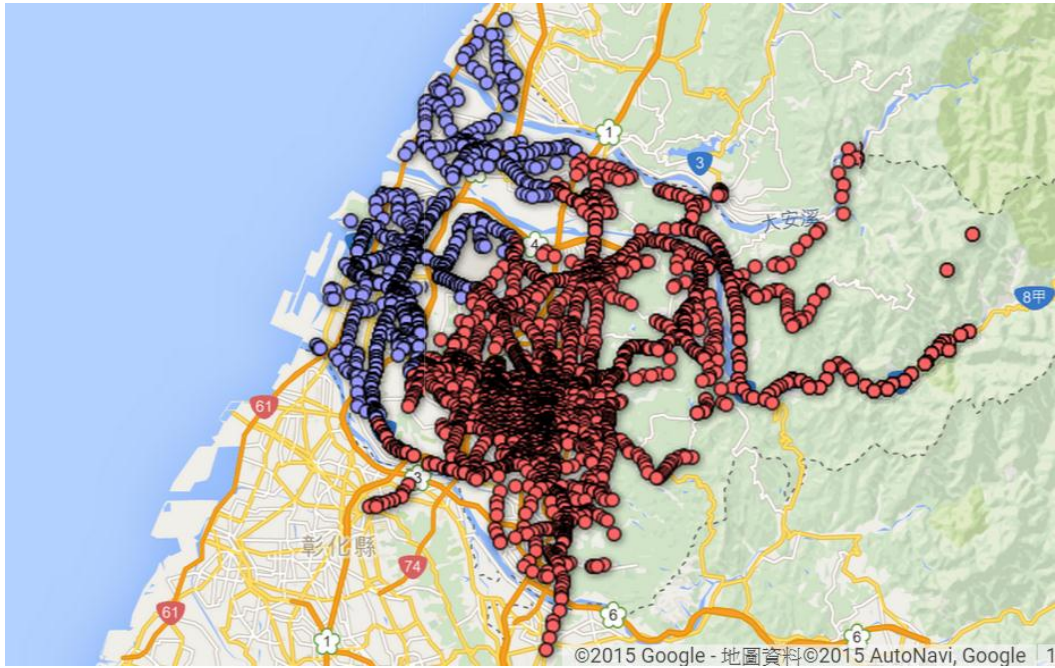


圖 6-43 臺中地區公車站牌測站參照分類圖

2. 將公車動態資料依趟次分組

公車動態資訊僅詳實紀錄每一傳回資訊之細節，但並未將資料群組為可直接判讀行車時間資訊之格式，特別關聯式資料庫的先天的儲放設計就未排序，須搭配後天的索引建立或查詢語法來執行。依趟次分組的演算法主要流程是將公車動態資訊先依公車車牌、GPS 時間依先後排序，接著檢視這些依時間排序後的資料，若發現有路線或去返程的改變時，就可視為新趟次的分界點。以上僅是基本原則，程式實作時尚須考慮資料清理、不同日期&業者、批次大量計算效率等細節。

3. 計算公車行經每個站牌之通過時間

前述分組過程中僅擷取單趟次首尾時間，若欲進一步瞭解公車通過各站牌時間可依據圖 5-2 中的進站離站欄位 (Leave) 媒合不同筆資料計算時間差。換言之，係透過計算公車觸發 InRadius (通常為接近站牌約 40 公尺) 與 OutRadius (通常為離開站牌約 50 公尺) 之時間差。

4. 預先計算各筆資料之屬性維度

為方便後續多維度分析需求，除了公車動態資料必帶的屬性（如路線/車牌/站牌）之外，舉凡資料所屬的月份/日/星期/時段/溫度/晴雨/雨勢等資料也均先計算並連結完成，將有助於後續的多維度探索分析。

5. 製作班表分析相關資料

考慮班表分析須呈現「Time-Space Diagrams」來描述公車彼此之間的行為，因此需要能自動產生一個不重複的識別碼來區分每趟次公車，以利後續能呈現每趟次之軌跡，此外，Time-Space 需要時間與空間兩種資訊，時間的部份較易處理，空間的部份要分別根據去/返程擷取站序資訊來取代原有的車站經緯度，方能順利呈現 Time-Space Diagrams。

6. 搜尋追越或連班樣態

追越樣態的偵測可搜尋「後發先至」之行為，除檢視連續兩站牌間是否有追越外，也須檢核在同一站牌內發生追越。至於連班偵測則是逐站牌檢核計算兩相鄰公車之間距寫出結果後續分析使用。

7. 結合電子票證資料呈現乘載旅客變化

搭配同時期之電子票證紀錄計算車上乘載旅客數，並充分融合兩項資料來源設計儀表板來達到同時檢視車輛分佈暨乘載率之目的。

6.5.3 視覺呈現暨應用

以下分別就七個主題以視覺化呈現的方式來介紹分析結果。

6.5.3.1 站牌通過時間探討

本計畫將全年度(2014-05-01~2015-04-30)之資料批次清理與融合後，以「路線」挑選作為檢視的主要依據如圖 6-44 右側，而在平均通過時間的呈現上，本計畫捨棄了需要花時間判讀的數字呈現，而改以

人類眼睛較容易辨識的性質-「圖形大小」來呈現如圖 6-44 左側，此種呈現方式可以讓使用者快速聚焦在最大泡泡（臺中火車站）。



圖 6-44 各站牌通過時間泡泡圖(以 55 路為例)

事實上，人眼除了容易辨識「圖形大小」之外，對於顏色深淺變化也相當直覺敏銳，因此本研究進一步將圖 6-44 進化為圖 6-45 的樹圖，圖 6-45 相對於圖 6-44 有三項提昇：

1. 右側路線改以複選方式處理

將原圖 6-44 之單選架構（Radio Box）改為圖 6-45 之多選架構（Check Box），可保有原來顯示單一路線功能外，等於增加多路線比較的彈性，系統可隨探索者的需求或喜好彈性點選，左側的資訊會即時更新，大幅提昇探勘效率。

2. 左側多納入「顏色深淺」作為數值大小的呈現

人眼對於「大小」的辨識雖比「數值」敏銳，但「顏色辨識」也是人眼能高效率判讀的因子，因此圖 6-45 同時結合「大小」與「顏色」元素，讓使用者對於站牌通過時間之大小差異有更直覺的檢視方式。

3. 左側依站牌停留時間之數值大小排序

原圖 6-44 之圖形僅能讓使用者辨識「通過時間」差距很大的站牌，較為細部的排序比較恐無法直接判讀，故圖 6-45 採用樹圖(Tree Map)技術來呈現，等同是在視覺化呈現中內嵌排序資訊，大幅增加圖形所能透露的資訊。

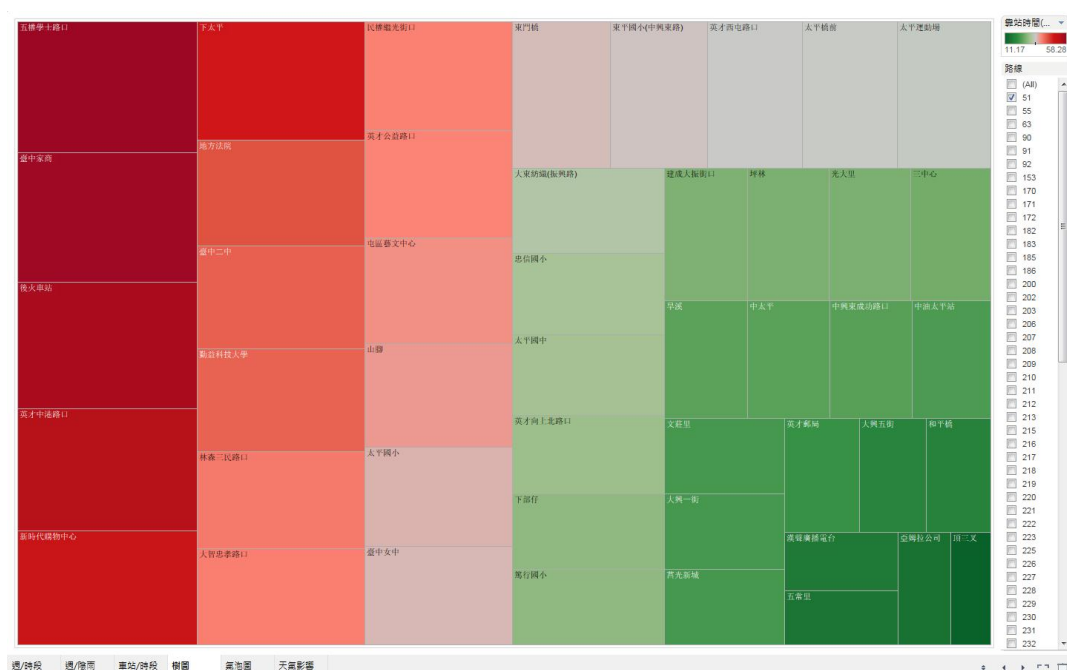


圖 6-45 各站牌通過時間樹圖(以 51 路為例)

不論是圖 6-44 或圖 6-45 之呈現，若使用者想讀取細部數值時，均可透過滑鼠移動到該區域（Mouse Hover）的方式讀取資訊，系統會以 ToolTip 方式提供細部資訊如圖 6-46。



圖 6-46 樹圖 ToolTip 功能示意圖

以上圖 6-44~圖 6-46 均只考慮單一維度（路線），當需要分析三個以上維度時，就不再適合採用「圖形大小」，僅須使用「顏色」呈現，例如圖 6-47 可分析使用者選定之路線於各站牌/各時段/晴天&雨天之平均站牌通過時間比較，此種探索方式可以快速的以視覺化呈現、比較不同因子（維度）是否所差異，協助使用者在眾多的數據中決定挖掘方向。



圖 6-47 站牌通過時間與時段及晴雨多維度比較圖

同理，透過圖 6-48 的視覺化呈現設計可根據使用者挑選的路線（含複選）分析不同時段/星期/晴雨之平均站牌通過時間差異。

以圖 6-48 的設定情境而言，顯然傍晚時段的數值顯著高於其他時段，而週間的比較裡，週日的數值有較其他週間日還低的現象。



圖 6-48 各時段週間站牌通過時間暨晴雨多維度比較圖

6.5.3.2 路線行車時間探討

路線行車時間亦可採用如圖 6-45～圖 6-48 般的多維度分析概念，在此不再贅述，本節另以 XY 散佈圖探討各路線全年度行車時間與溫度關係如圖 6-49，系統會自動繪製套疊線性回歸線，當滑鼠移動至回歸線時會以 ToolTip 顯示回歸線重要訊息，諸如係數或 P-Value 值等。

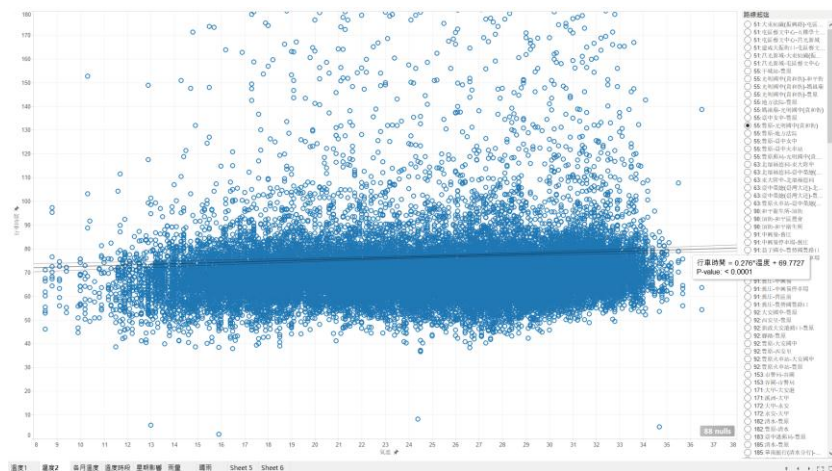


圖 6-49 行車時間與溫度關係（以 55 路為例）

然而，Tableau 內建的趨勢線功能係以線性回歸為主，本計畫也額外嘗試以 R 語言使用 Loess 函數（Local Regression）搭配 ggplot2 套件（<http://ggplot2.org/>）繪製局部回歸曲線如圖 6-50。

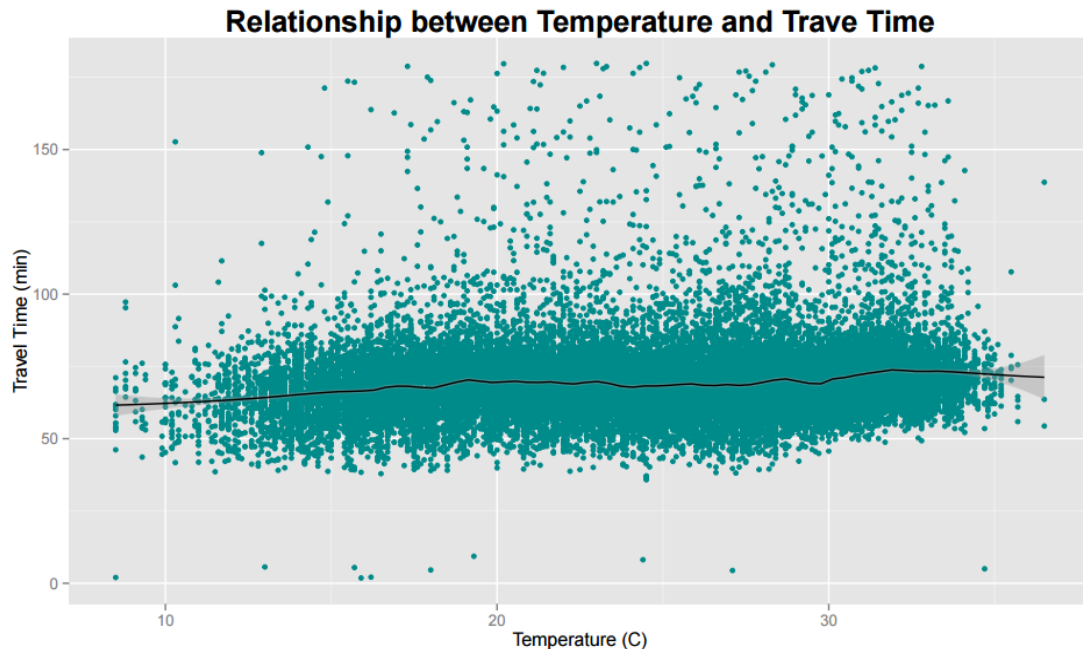


圖 6-50 以 R 語言區域回歸取代圖 6-49 之分析

6.5.3.3 預估到站時間準確性分析

由於公車系統並非專有路權，從前述兩節的行車時間分析可觀察到行車時間變異相當高，這種「不確定感」對車上(已上車)旅客與車外（候車）衝擊不同，前者由於已在車上，能直接感受到路況條件，僅有不確定能否如預期時間到達自己的目的站，但候車旅客之不確定感則更高，旅客久候過程尚可能換運具之權衡。而公車動態系統大幅降低此不確定感，旅客可利用網路（含電腦、手持裝置）得知班車預計到達時間，此資訊有助於旅客選擇運具及安排行程。由於民眾相當需要甚至依賴此資訊，故預估時間正確性也值得探討，故本研究以豐原客運在臺中市的路線為例執行分析。

圖 6-51 係由「週」因素來檢視預估到站時間正確性，從中可觀察到兩個現象：

1. 七條線相當接近，代表此因素並非影響正確性之因素，但整個軌跡可視為整體平均，
2. 整體而言，若以正負 3 分鐘為標準，正確率達 80%，若以正負 4 分鐘為標準，正確率達 85%，若以正負 5 分鐘為標準，正確率達 90%。

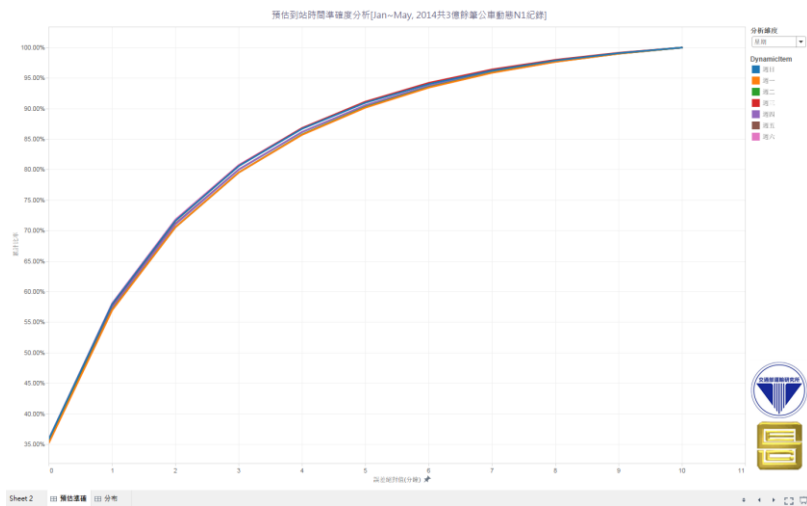


圖 6-51 預估到站時間準確性-整體水準

圖 6-52 則是改變分析維度為「時段」，從中可觀察出各時段略有差異，越是右下方曲線代表準確度越差，從顏色對應為下午 4 點/5 點，越是左上方曲線代表準確度越佳，從顏色對應為下午 1 點/2 點。

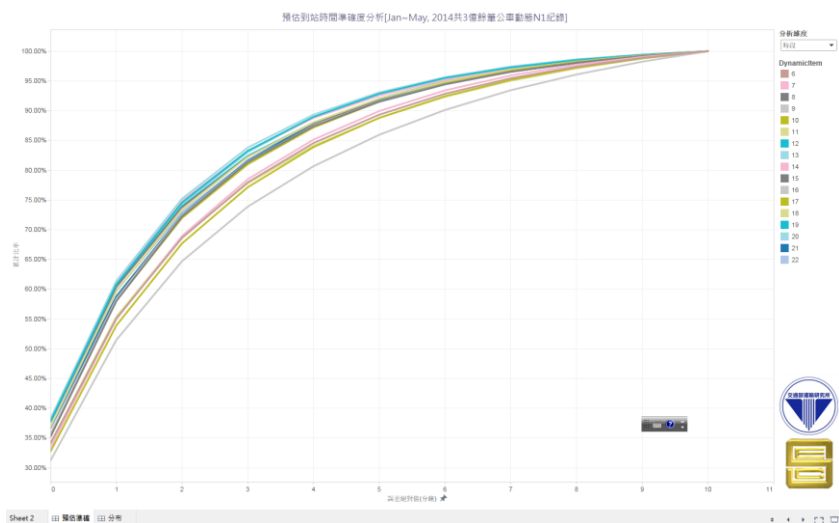


圖 6-52 預估到站時間準確性-時段因素分析

圖 6-53 改以「站間數」作為分析維度，就如先驗知識，公車位置與預測點位置越接近時，準確率極高，但隨著距離或站間數越長，準確度就越來越低。對於遠距離的預測雖然準確度低，但由於系統會隨著公車行進持續更新預測，待距離越來越近時即可修正到相當正確的預估到站時間。

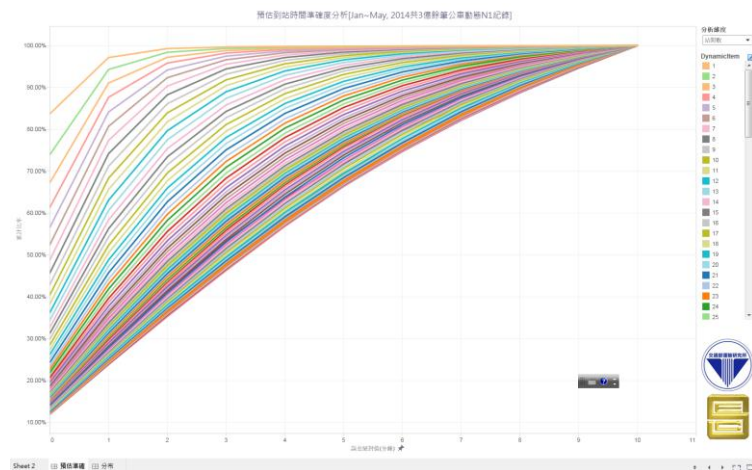


圖 6-53 預估到站時間準確性-站間數因素分析

6.5.3.4 班次數及分布視覺化呈現

「Time-Space Diagrams」係呈現班次數與分佈最合適之視覺化呈現方式，例如圖 6-54 與圖 6-55 係分別針對 55 路與 92 路公車，可明顯發現兩路線班次數之差異與各自於全天的分佈狀況。



圖 6-54 高密度班表視覺化呈現圖（以 55 路為例）

無論圖 6-54 與圖 6-55 均可透過右側選擇 2014 年 5 月 1 日~2015 年 4 月 30 日任一日、任一路線及往/返程，左側的時空圖會即時更新。

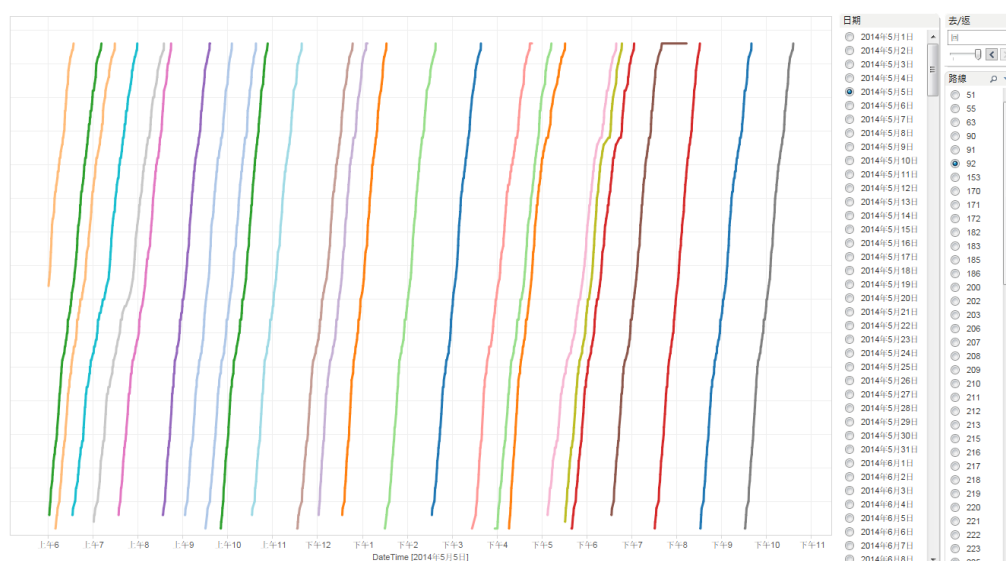


圖 6-55 低密度班表視覺化呈現圖（以 92 路為例）

6.5.3.5 連班樣態視覺化呈現

圖 6-54 與圖 6-55 係以全日檢視來瞭解一日班表之全貌，但針對班次較多之路線（例如圖 6-54 之 55 路線），可利用如圖 6-56 之工具調整檢視範圍，除了「+」「-」等比例放大外，也可以用滑鼠指定放大範圍或移動檢視範圍。

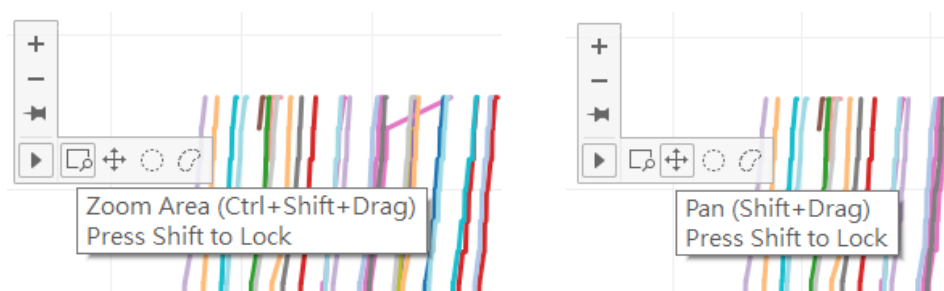


圖 6-56 全日班表檢視範圍調整工具列

透過圖 6-56 之工具，將可清楚觀察公車連班或追越的詳細狀況如圖 6-57，如欲追蹤細部的發生時間或站牌名稱，只要將滑鼠移至觀測點，系統會立刻以 ToolTip 方式顯示詳細資訊如圖 6-57。

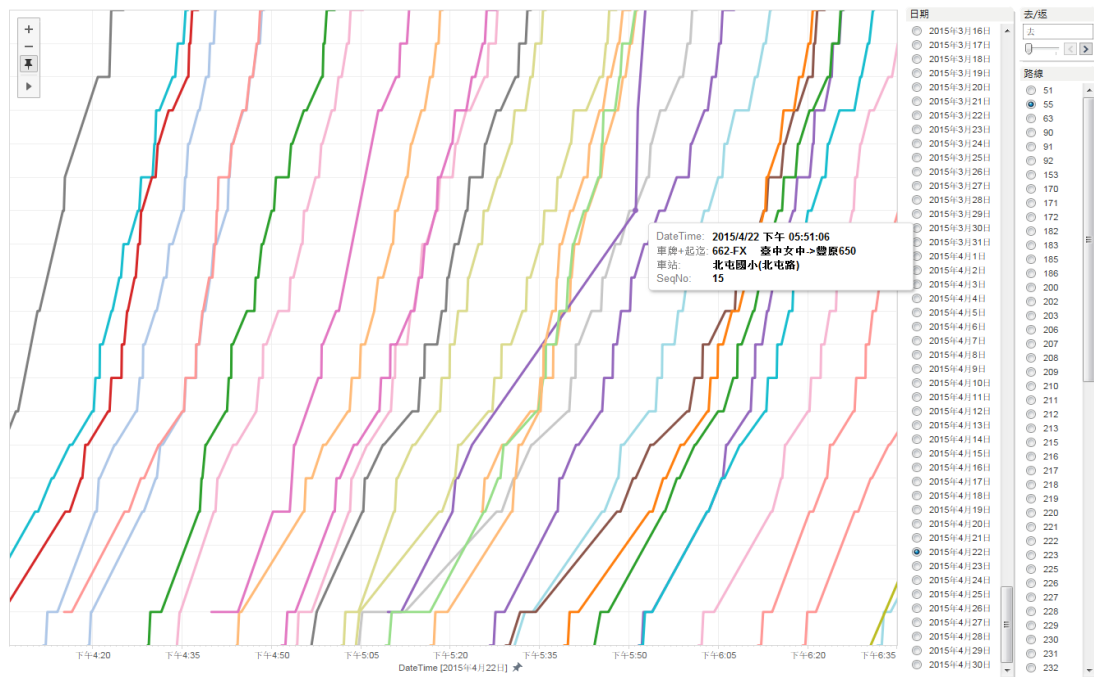


圖 6-57 班表連班/班車追越樣態視覺化呈現圖

圖 6-57 雖能精確檢視連班或追越行為，但若要瞭解連班的統計仍有所不足，故本研究另以圖 6-58 統計全年各路線發生連班的數量比較，可觀察班次數較高之路線比較容易產生連班狀況。

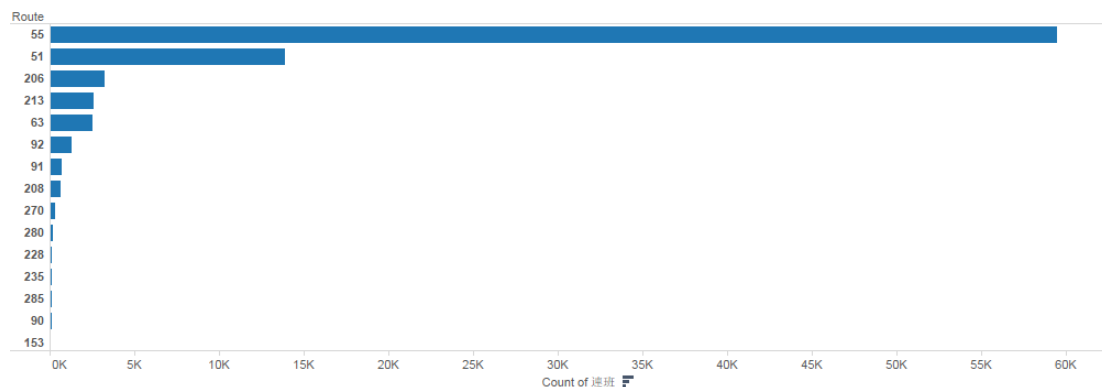


圖 6-58 連班路線分析圖

至於連班發生的時間或地點亦可深入分析，例如圖 6-59 與圖 6-60 均可讓使用者選擇欲分析路線，圖 6-59 統計連班的時段分佈比例，如一般先驗知識，上/下午尖峰較常發生連班，且下午尖峰會比上午尖峰

還明顯，此外，圖 6-59 除時段外亦含去/返程維度，以圖 6-59 為例，傍晚去程之連班會比返程明顯。

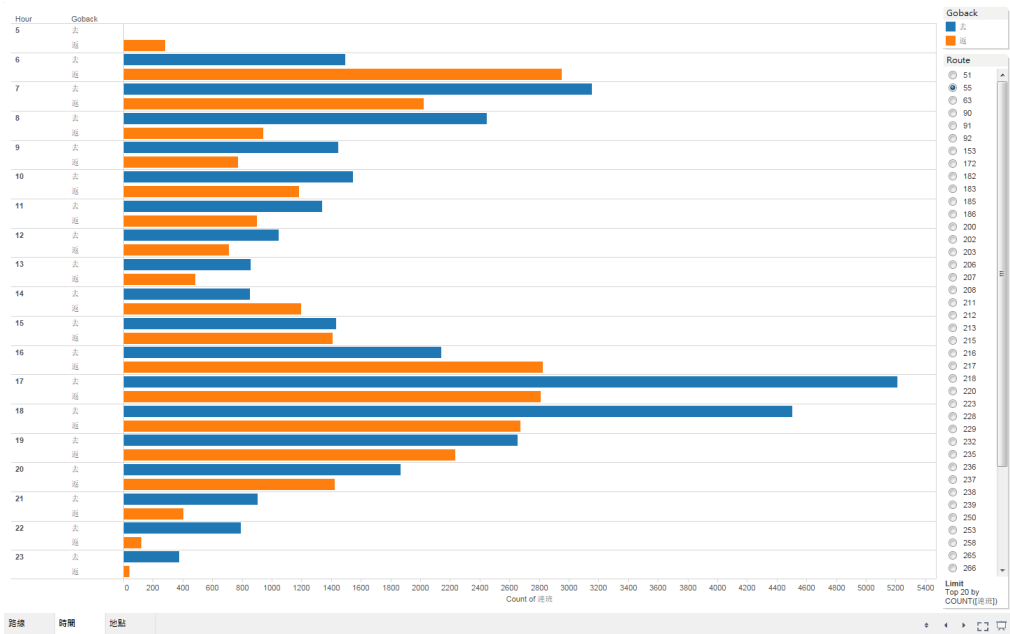


圖 6-59 連班分佈分析（時間角度）

圖 6-60 則是根據使用者指定之路線來統計連班發生的站牌名稱，以瞭解最頻繁發生連班之地點，同樣區分去程與返程。

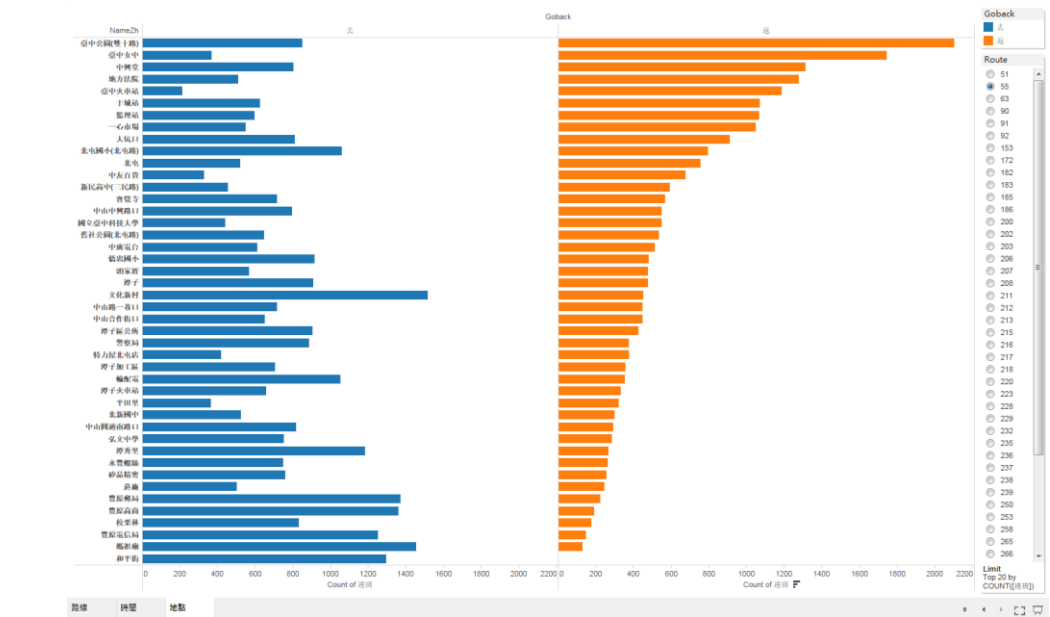


圖 6-60 連班分佈分析（空間角度）

6.5.3.6 跨日班次班表軌跡比較

前述圖 6-54～圖 6-57 均著眼於同日班次車輛之分析，若將分析焦點在於不同天之同班次公車軌跡，則可設計作為評估站牌班表推動可行性之用，以圖 6-61 與圖 6-62 為例，係以 277 路線去程為例，分別繪製週間與週末之路線軌跡套疊圖。

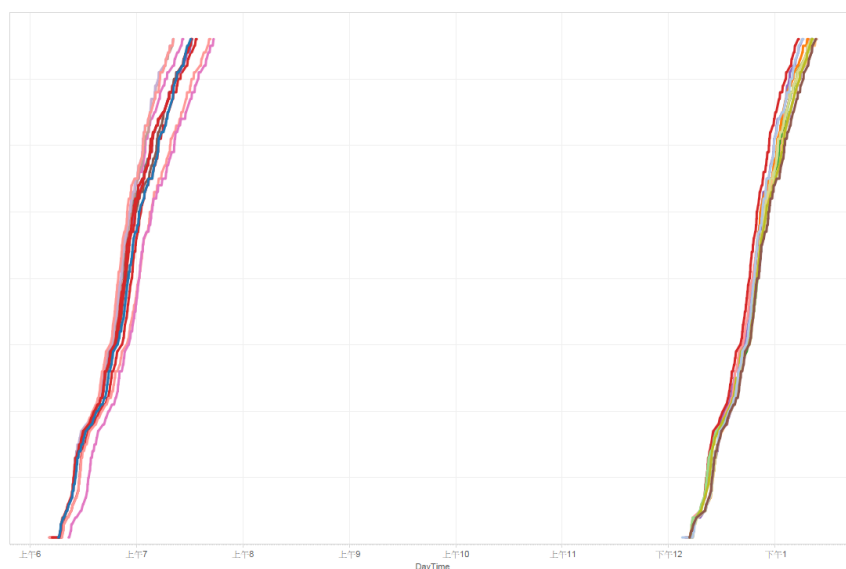


圖 6-61 跨日班表比較視覺化呈現圖（以週間為例）

從圖形上方的開口大小即可判斷行車時間變異，當變異越小，代表推動站牌班表的可行性越高。

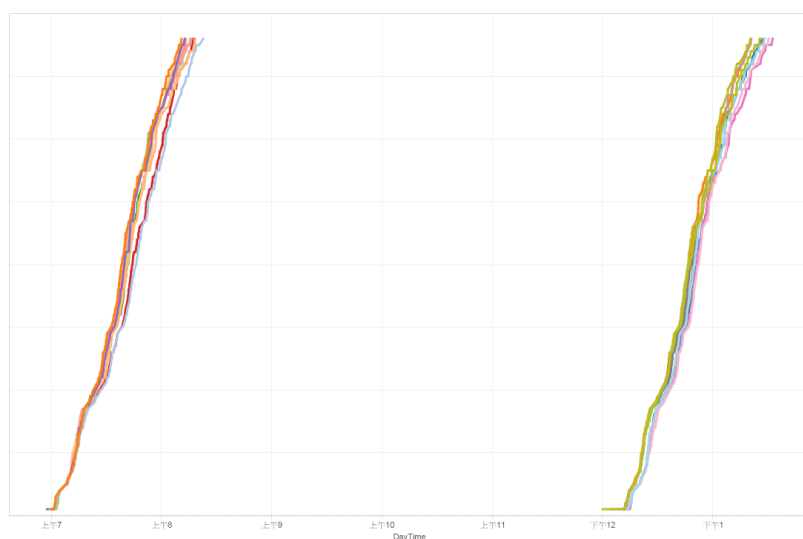


圖 6-62 跨日班表比較視覺化呈現圖（以週末為例）

6.5.3.7 班表暨乘載率聯動儀表板

第 6.5.3.4 節與第 6.5.3.5 節係以每一趟次不同顏色來呈現班次數與分佈情形，本節進一步用乘載人數來繪製公車軌跡之顏色如圖 6-63，利用這種上色方式即可讓使用者觀察乘載率之全貌，以圖 6-63 為例，下午五點～七點階段係屬於乘載率較高之時段，配合滑鼠移動可瞭解每一個時空點之詳細資訊，包括車牌、站牌、確切時間、車上載客人數等資訊，同時右下角之甜甜圈圖會同步顯示總人數以及票種（如一般票、學生票、敬老票等）分佈比例。換言之，此儀表板同時可呈現行車時間、班表、班距、連班、追越、乘載量、票種組成比例等多樣資訊。

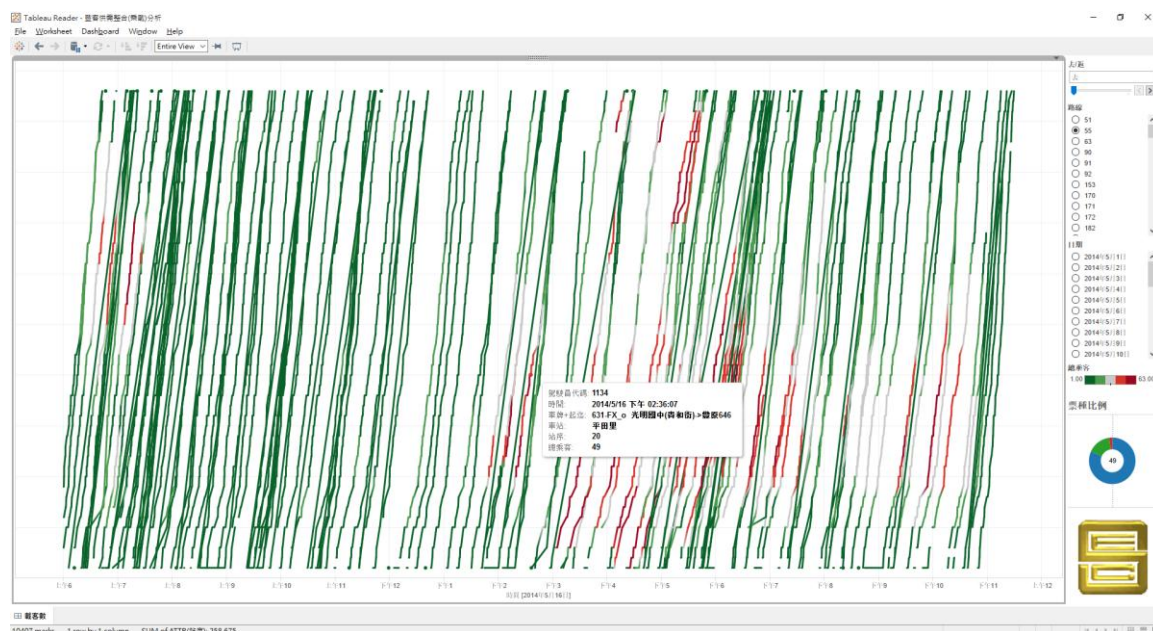


圖 6-63 供給時空圖暨乘載率整合儀表板

6.6 站牌班表產製分析

6.6.1 探勘分析目的

站牌班表產製之目的，為透過歷史公車動態資料數據，建立每條路線上的各站班表，以利旅客掌握公車到達各站位時間，節省旅客等

待時間，且可供轉乘參考，有助於提升整體服務水準。

6.6.2 演算處理流程

本研究所預計使用之站牌班表產製方式，主要著重在簡單、易執行、無需傳輸費用且沒有複雜的模式，對於已建置完成公車動態資訊系統之地區，尤其是偏遠地區路線而言，可輕易複製本作法並自行操作，因此，在實務面上的推廣較為可行。本研究所提出的站位班表產生方式，基本上可以分為五個步驟，以下以 51 路公車為例說明：

1. 蒐集歷史資料

透過已經建置完成的公車動態資訊系統，蒐集每條路線上每一班次的所有站位之到離站時間，因為本研究以每條路線班次為單位建立歷史資料庫，因此不同路線班次將分開蒐集，而路線班次相同但車次不同則合併蒐集。除此之外，由於一天當中的交通狀況會隨著時間而有所差異，因此本研究將尖峰及離峰時刻的資料予以分開蒐集，即預估到達各站位之旅行時間分為尖離峰兩種，以制定更符合實際需求的班表。其中尖、離峰時刻之定義，隨著不同之路線將會有所差異，因此應由實地調查或對歷史資料之研判而加以定義，在此路線中，尖峰時間定為上午六點半至八點半，以及下午五點至七點。

2. 刪除異常值

由於已蒐集完的歷史資料庫中，可能會因為設備故障、交通狀況異常（如事故等）或者人為疏失而導致極端值的出現。若是將這些極端值也納入後續產製站位班表的計算中，將會顯著影響所產生班表的正確性，因此必須將這些異常值予以刪除。

為了將公車於各站位停等之時間納入考量，因此建議將各站的到站時間減去前一站的離站時間，即可計算出相鄰各站間旅行時間，並使用兩站間之距離，計算平均旅行速度以評估旅行時間之合理性，若兩站間之平均旅行速度超出公車車速上限(每小時 40 公里)，或低於步行速度(每小時 4 公里)，則予以刪除，接著計算相鄰兩站旅行時間之

平均值及標準差，又根據統計學上離群值的定義，由平均值加減 3 倍標準差，所包含的資料約為 99.7%，因此，一般而言，若超出平均值 3 倍標準差外的相對應資料稱為離群值，故本研究將這些離群值予以刪除。

3. 找出歷史資料之下緣

為避免車輛提早駛離導致民眾搭不上車的情形發生，本研究將針對每一條路線之離站時間找出其下緣線以進行分析。下緣線乃針對每條路線的歷史資料，繪製如圖 6-64、圖 6-65 之時空圖，橫軸為站位編號，縱軸為離站時間，圖中每一條線分別代表每個車次之時空軌跡，而每一點則代表該車次於各站位之離站時間。如此即可找出每個站位車輛最早之離開時間，構成如圖 6-64、圖 6-65 所示的下緣線。

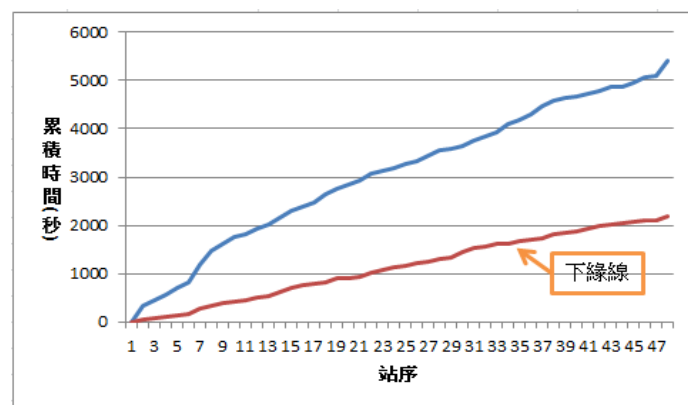


圖 6-64 離峰下緣線示意圖

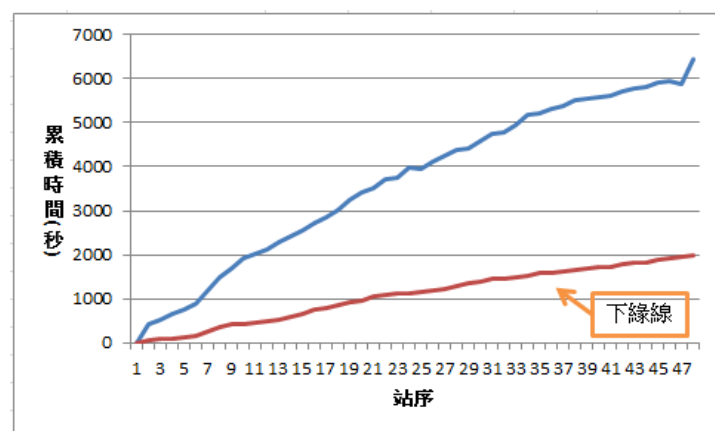


圖 6-65 尖峰下緣線示意圖

若採用每個站位的下緣（最早）時間來制定班表，在大部分的情況下，車輛將會比公告的班表晚到，如此一來可能會因此產生多餘的等候時間，因此如何降低車輛過早駛離使得準時到站的民眾無法上車的情況發生，並將等候時間控制在民眾所能夠接受的範圍內，本研究將進行下一步驟：下緣線基礎法。

4. 下緣線基礎法（下緣線的調整）

前一步驟已利用下緣線找出各站位車輛最早之離開時間，因此在多數情況下車輛會比下緣線（最早）的時間還要晚到。由於民眾不可能完全準時依照公告的班表抵達招呼站，故民眾比班表提早個幾分鐘抵達是可以預期的情況。因此本研究假設若車輛只比班表提早幾分鐘駛離，應該不會對民眾造成太大的困擾。

基於以上的觀察，本研究建議可將前一步驟所找出的下緣線略微向上調整，也就是將下緣線加上些許的時間（修正量），以減少民眾的候車時間。本研究所提出的修正量為在歷史資料庫中，找出每個站位最晚離站時間（最上緣）與最早離站時間（最下緣）的差值（稱為時間差）後，再乘以一個適合權重值。為了能夠在上、下緣間找出最合適的權重值，本研究建議以時間差的固定比例（如 10%），分別計算各比例下站位班表（如圖 6-66、圖 6-67 所示），並透過目標值來判別其優劣。

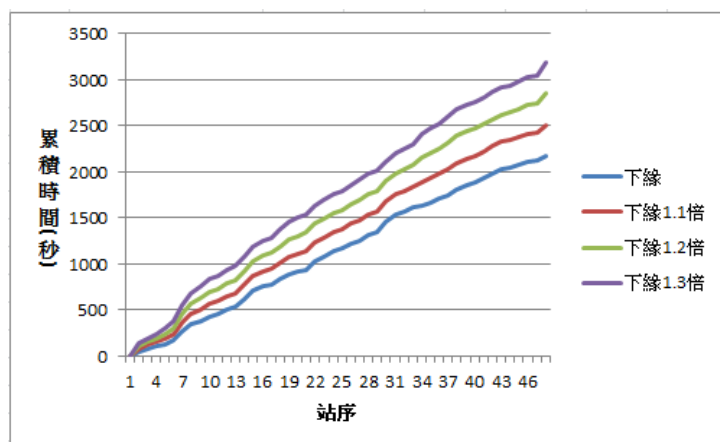


圖 6-66 離峰下緣線與各權重示意圖

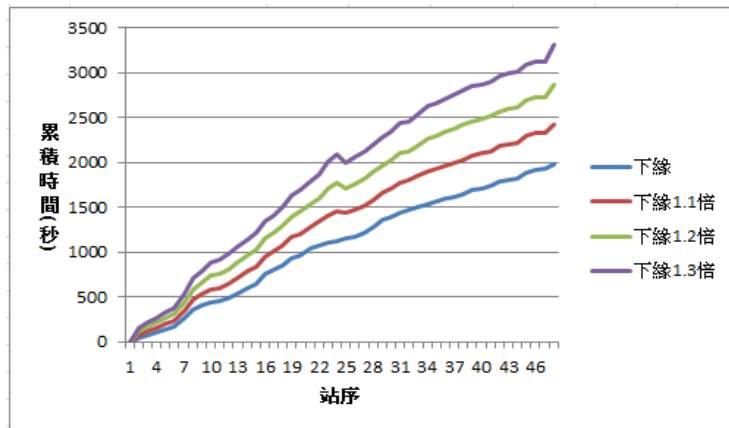


圖 6-67 尖峰下緣線與各權重示意圖

其中本研究對目標值之設定，將依照車輛早到與晚到的時間長短，分別給予不同的懲罰值。由於公車過早到達可能會讓準時到達的民眾搭不上車，這種情況是非常不樂見的，因此必須有較大的懲罰值。至於公車晚到的情況，通常隨著等待時間的增加，乘客的不滿意程度會隨之急遽提高，因此對應的懲罰值採用指數型態來加以呈現，懲罰值乃為可調整之參數，各地方可以依據實際情況自行設定。本研究之預設值主要是說明懲罰建議呈現指數分布的形態。本研究所訂定之目標值計算方式如下：

- (1) 若車輛早到 2 分鐘以上，則準時到達的民眾可能無法搭上車，因此懲罰值設定為 300，或任何一個大數。
- (2) 若車輛早到 2 分鐘至晚到 5 分鐘以內視為準點，此範圍內對民眾的搭乘權益沒有顯著影響，因此懲罰值為零。
- (3) 若車輛晚到 5 分鐘以上 10 分鐘以下，懲罰值設定為 $2 \times$ （晚到時間）。
- (4) 若車輛晚到 10 分鐘以上 15 分鐘以下，懲罰值設定為 $5 \times$ （晚到時間）。
- (5) 若車輛晚到 15 分鐘以上，由於時間過長，嚴重影響民眾的權益，因此將懲罰值設定為 $20 \times$ （晚到時間）。（懲罰區間與懲罰值之設定，可依照地區與使用者之不同進行微調，本研究僅將此一方法架構提出，參數部分則可因地制宜。）

依此所得到目標值之數學式如式(5.1)所示：

$$\text{Min} \sum_{i=1}^{i=n} \alpha \quad (6.1)$$

$$\text{其中：} \begin{cases} +300 & , \text{if } x_i - \beta Y \leq 2 \\ 0 & , \text{if } 2 < x_i - \beta Y < 5 \\ (x_i - \beta Y) \times 2 & , \text{if } 5 \leq x_i - \beta Y < 10 \\ (x_i - \beta Y) \times 5 & , \text{if } 10 \leq x_i - \beta Y < 15 \\ (x_i - \beta Y) \times 10 & , \text{if } 15 \leq x_i - \beta Y \end{cases}$$

n = 所蒐集之資料比數

x_i = 當天第 i 班車實際到站時間

Y = 下緣線

β = 下緣線修正權重

雖然如此即可求得建議之修正權重值，但本研究為求修正權重值之最佳性，將得出之建議權重值再各增減 5% 後，再度進行比較，以得出最後建議之權重值。由於在時刻表已公告之情形下，旅客到站行為近似常態分配，因此將下緣線略為向上調整，以找出目標值最小的修正權重，將只會影響到少數民眾之權益，但卻能夠有效降低多數民眾的候車時間，因此本研究採此方法來找尋最佳的下緣修正權重值。此外由操作經驗可以得知，本研究欲降低車輛過早駛離使得準時到站的民眾無法上車的情況發生，因此由下緣線向上修正能較快找出最合適之修正權重值。

5. 確定建議權重值

根據前述作法，以站為單位，計算各種不同調整比例的目標值，對應到最小目標值情境所使用之比例，該比例即為本研究所建議之修正比例，此案例所得結果如圖 6-68、圖 6-69 所示。

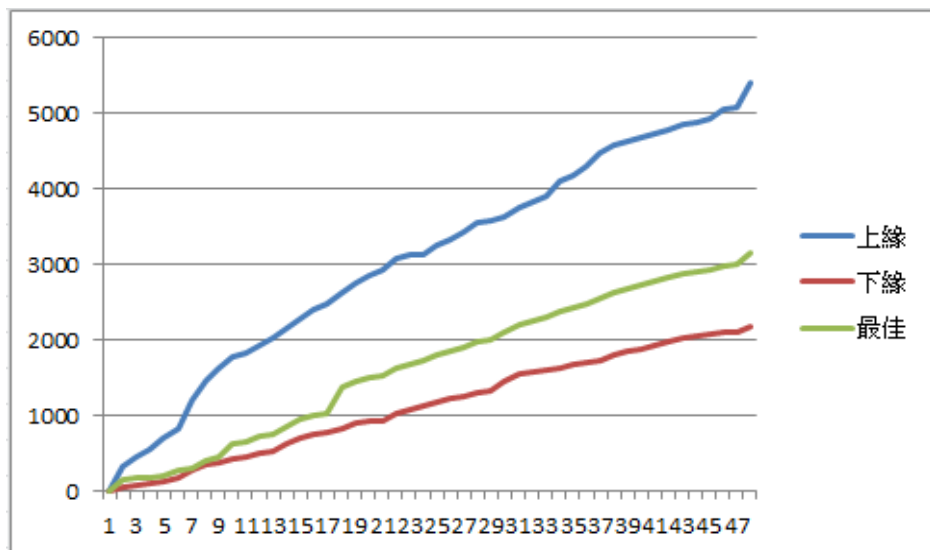


圖 6-68 離峰最佳目標值結果

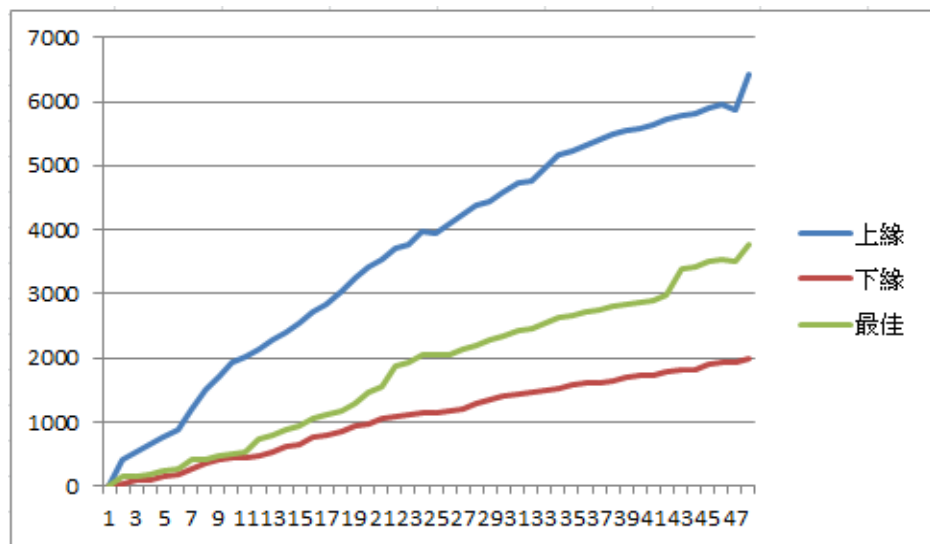


圖 6-69 尖峰最佳目標值結果

6.6.3 站牌班表正確性分析

自豐原客運歷史公車定點資料庫中，可得共 75 條路線，其中以 51 路公車為例，使用 2014 年 11 月至 2015 年 3 月之資料，計算預估各站位之累積旅行時間，並使用 2015 年 4 月之資料進行精準度評估，計算各站到達時間與預估值之平均誤差及標準差，結果如表 6.5 所示。

表 6.5 各站到達時間與預估值之平均誤差及標準差表

站序	非尖峰 旅行 時間	非尖峰 平均誤 差(秒)	非尖峰 標準差 (秒)	尖峰旅 行時間	尖峰平 均誤差 (秒)	尖峰標 準差(秒)
1	00:00:00	0	0	00:00:00	0	0
2	00:01:42	73.34	50.40	00:01:38	101.79	67.79
3	00:02:24	25.27	28.75	00:02:19	20.15	25.45
4	00:02:52	33.75	31.70	00:02:48	45.29	36.94
5	00:03:20	62.34	38.38	00:03:20	60.93	40.40
6	00:03:46	14.32	22.79	00:03:46	18.37	25.37
7	00:04:45	80.04	50.09	00:04:46	92.77	51.41
8	00:05:38	56.61	28.69	00:05:38	53.31	20.96
9	00:06:13	82.63	61.80	00:06:12	96.14	58.88
10	00:09:08	-79.10	28.59	00:09:15	-88.88	31.51

註解：取前 10 站為例

接著對產製之站排班表進行情境分析，分別針對周間/周末、平時/春節假期、晨峰/昏峰三種情境分析，結果如表 6.6 至表 6.11 所示。

表 6.6 周間/周末情境分析

	周間		周末	
站序	平均誤差	標準差	平均誤差	標準差
1	0	0	0	0
2	10.92	52.27	9.50	51.17
3	45.66	50.41	46.00	49.12
4	117.59	67.15	115.59	63.89
5	169.12	70.81	167.54	69.72
6	147.61	80.08	145.36	77.28
7	245.42	111.97	245.39	109.49
8	273.13	133.57	265.56	125.20
9	351.75	159.26	331.83	153.56
10	266.34	169.13	251.27	165.77

註解：以秒為單位

表 6.7 周間/周末情境分析差異檢定

	周間	周末
平均數	291.7081	225.3399
變異數	19506.34	18470.27
觀察值個數	196	196
皮耳森相關係數	0.885347	
假設的均數差	0	
自由度	195	
t 統計	14.06092	
P(T<=t) 單尾	9.09E-32	
臨界值：單尾	1.652705	
P(T<=t) 雙尾	1.82E-31	

表 6.8 平時/春節假期情境分析

	平時		春節假期	
站序	平均誤差	標準差	平均誤差	標準差
1	0	0	0	0
2	11.16	51.61	-7.93	58.34
3	46.33	49.93	29.85	50.40
4	118.10	65.98	86.45	65.89
5	169.94	69.66	133.28	83.28
6	148.40	78.60	106.89	87.40
7	247.24	111.06	194.62	104.50
8	273.08	131.15	210.78	119.30
9	349.57	157.36	244.60	138.11
10	265.47	167.71	163.20	154.68

註解：以秒為單位

表 6.9 平時/春節假期情境分析差異檢定

	平時	春節假期
平均數	280.6009	27.90521
變異數	18868.57	40647.01
觀察值個數	196	196
皮耳森相關係數	0.339817	
假設的均數差	0	
自由度	195	
t 統計	17.53725	
P(T<=t) 單尾	2.98E-42	
臨界值：單尾	1.652705	
P(T<=t) 雙尾	5.95E-42	

表 6.10 晨峰/昏峰情境分析

	昏峰		晨峰	
站序	平均誤差	標準差	平均誤差	標準差
1	0	0	0	0
2	85.46	63.51	10.63	56.28
3	148.30	68.97	82.12	51.11
4	189.63	83.60	93.83	69.44
5	237.61	77.81	148.26	79.97
6	243.02	88.39	137.66	87.77
7	280.03	108.90	163.14	113.78
8	374.84	128.76	243.58	118.81
9	478.78	163.10	280.27	145.95
10	567.45	181.15	310.35	160.80

註解：以秒為單位

表 6.11 晨峰/昏峰情境分析差異檢定

	昏峰	晨峰
平均數	459.9242	53.9039
變異數	45585.17	36602.37
觀察值個數	98	98
皮耳森相關係數	-0.10977	
假設的均數差	0	
自由度	97	
t 統計	13.31283	
P(T<=t) 單尾	6.44E-24	
臨界值：單尾	1.660715	
P(T<=t) 雙尾	1.29E-23	

在上述三種情境中，經檢定後兩兩間皆有明顯差異，因此在實際運作上，需針對各情境建置特定班表，無法使用單一班表滿足所有情況，然而，使用多種班表容易造成旅客及各站點班表張貼混亂，因此不建議實施各站班表。

雖然目前不建議全面實施站牌班表，但未來若能夠根據所收集到的資料，針對某些特性駕駛員進行積極的輔導與關懷，減少不同駕駛員駕駛行為的差異後，站牌班表的制定仍然是一個可以努力的方向。

6.7 車輛相關分析(含機務、油料)

近年來隨著都市地區人口成長與經濟發展帶動世界各地運輸產業的發展，客運業者也不例外，面對眾多的同業競爭，客運業者為了提高自身營運效率，對於人力與資產的投資也不遺餘力，車輛的購買與人員的訓練目的在於強化公司的營運效率，然而大量的營運車輛產生的油耗勢必增加公司的成本支出，因此如何有效的減少燃油量已達

到營利的最大化和全球普遍節能減碳的情況下，確實是一門重要的課題。

6.7.1 探勘分析目的

燃油消耗一直是運輸產業相當大的支出部分，為了減少不必要的成本，各大運輸產業都極力考慮如何降低油料支出的成本。本研究將駕駛、車號和車輛特性進行油耗分析，找出哪些車號和車輛特性是耗油量較大的族群，盡可能連結到駕駛，對這族群給予客運業者意見及改善方向。

另外，車輛是客運公司重要的資產，需要隨時掌控車輛的使用狀況，因此機務分析就十分的重要。本研究主要針對車輛特性及進廠維修資料進行機務分析，找出哪些車號和車輛特性是容易發生車輛維修的族群以及哪些特性是容易造成車輛故障，針對這一部分給予業者意見及改善方向。

6.7.2 演算處理流程

一、油料分析:

本次研究主要針對車輛進行油耗分析，油耗受到駕駛員開車行為、車輛特性等原因所影響，由於資料內沒有每日駕駛耗油量，因此主要以車輛為主體進行分析，研究架構如圖 6-70 所示，主要分成三個部分：

(1) 蒐集資料

透過所蒐集的資料，整理對油耗分析有所影響的因素，如：車輛載重量、廠牌、出廠年份等。

(2) 油耗分析

分別對車輛特性及車號做分析。

(3) 油耗模式應用

利用整理油耗分析結果，討論其管理上的使用與適用性。

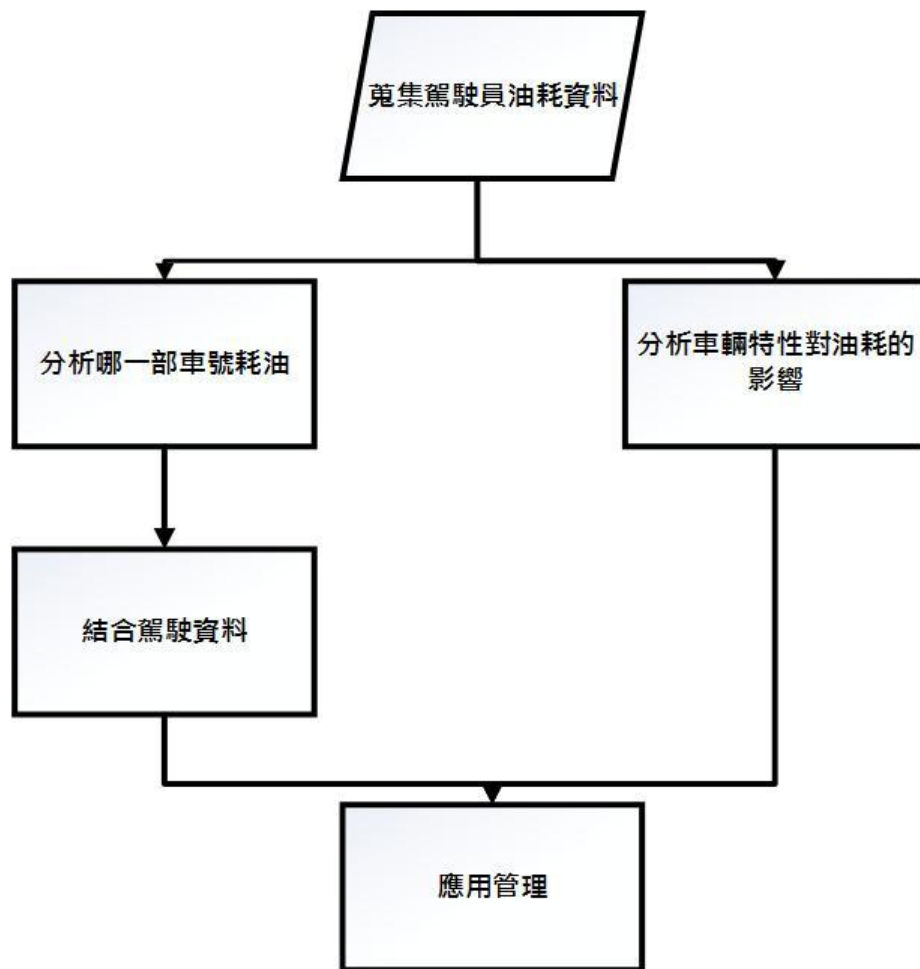


圖 6-70 油耗分析流程圖

本分析資料來自於客運業者的營運資料，包含有每日營運報表、車輛清冊和車輛維修等，本分析鎖定各車牌號碼一年的燃油資料以及車輛特性為主，由於客運業者暫未提供駕駛每天之耗油量，因此只能從各車牌號碼來整理哪些車號的駕駛和車輛特性是最耗油的原因。

1. 原始資料

對於油耗分析，本研究的原始資料包含三個部分：車輛清冊、車輛零件統計表和電子票證資料。車輛清冊主要記錄營運車輛的基本資料，包含車別、車號、廠牌、年份和噸位等等(表 6.13)，車輛零件統計表主要記錄營運車輛一年份的總耗油量、各車牌號碼的公里柴油公升等資料(表 6.12)，電子票證資料主要記錄駕駛員編號、車牌號碼和路線編號等資料(表 6.14)。

表 6.12 車輛零件統計表

編號	資料行描述	屬性	資料行名稱	說明
1	流水號	數字	ID	
2	車號	數字	BUSID	
3	總油耗量	數字	TOTAL_OIL	
4	公里柴油公升	數字	DIESEL_FUEL_PER_KILOMETER	總公里數/柴油公升數
5	機油量	數字	ENGINE_OIL	公升
6	公里機油公升	數字	ENGINE_OIL_PER_KILOMETER	總公里數/機油公升數
7	柴油耗油金額	數字	AMOUNT_DIESEL_FUEL	
8	零件費	文字	AMOUNT_VEHICLE_PARTS	
9	輪胎費	數字	AMOUNT_TIRE	
10	機油費	數字	AMOUNT_ENGINE_OIL	
11	雜油費	數字	AMOUNT_OIL	
12	耗用總計	數字	AMOUNT_CONSUMPTION	
13	耗用公里	數字	AMOUNT_KILOMETER	耗用總計/總公里數,元/公里
14	年份	日期	YEARS	西元/月
15	廠牌	文字	BRAND_VEHICLE	

資料來源：豐原客運提供暨本研究整理

表 6.13 車輛清冊表

編號	資料行描述	屬性	資料行名稱	說明
1	車別	文字	VEHICLE_TYPE	營運車輛分類 (大營運客車)
2	車號	數字	BUSID	
3	廠牌	文字	BRAND	
4	年份	日期	YEARS	西元/月
5	引擎號碼	數字	ENGINE_ID	
6	噸位	數字	TON	
7	座位	數字	SEATS	
8	備註	文字	MEMO	

資料來源：豐原客運提供暨本研究整理

表 6.14 電子票證資料表

編號	資料行描述	屬性	資料行名稱
1	營運日期	日期	data
2	司機編號	數字	driver_id
3	路線編號	數字	route_id
4	車輛編號	數字	car_id
5		數字	d_no
6	更新時間	時間	uptime

資料來源：豐原客運提供暨本研究整理

2. 資料轉換與處理

資料前處理的動作，主要目的在於過濾並去除不需要的資料，有助於資料探勘的進行與最後歸責的品質。

資料前處理的主要動作有：

- (1) 清除不需要的資料:將不需要的資料刪除。車輛清冊資料部分，本分析主要對象是各車輛特性及車牌號碼進行油耗分析，對於引擎號碼及座位數對本分析的探討並沒有幫助;車輛零件統計表資料部分，將機油和零件部分的資料刪除，這一部分的資料並非本分析的重點，因此未討論;電子票證資料部分，在這一部分只為了將車牌號碼及駕駛連結起來。
- (2) 去除遺漏的資料:刪除缺漏值及錯誤的資料，減少對資料探勘結束的影響。
- (3) 資料轉換:將資料的表示方法，轉換成相同的格式，其目的在協助之後探勘的進行。如：車輛出廠年份原本的 2003.01、2004.04，轉換成車齡，12 年、11 年;車輛廠牌匿名化。
- (4) 耗油量:將車輛零件統計表內的公里柴油公升(總公里數/柴油公升數)轉換成每公里消耗多少油量(柴油公升數/總公里數)，主要原因公里柴油公升屬於燃油消耗(總公里數/柴油公升數)，並非直觀的耗油量(柴油公升數/總公里數)，轉換方式將公里柴油公升做倒數處理。

二、機務分析:

本次研究主要針對所有車輛進行機務分析，車輛風險受到駕駛員開車行為、車輛特性等原因所影響，本研究討論車輛風險的特性，研究架構如圖 6-71 所示，主要分成三個部分：

(1) 蒐集資料

透過所蒐集的資料，整理對機務分析有所影響的因素挑出，如:車輛載重量、廠牌、出廠年分等。

(2) 機務分析

由資料做分析，整理機務分析結果，分別有對車輛特性及車號做分析。

(3) 機務模式應用

利用整理機務分析結果，討論其管理上的使用與適用性。

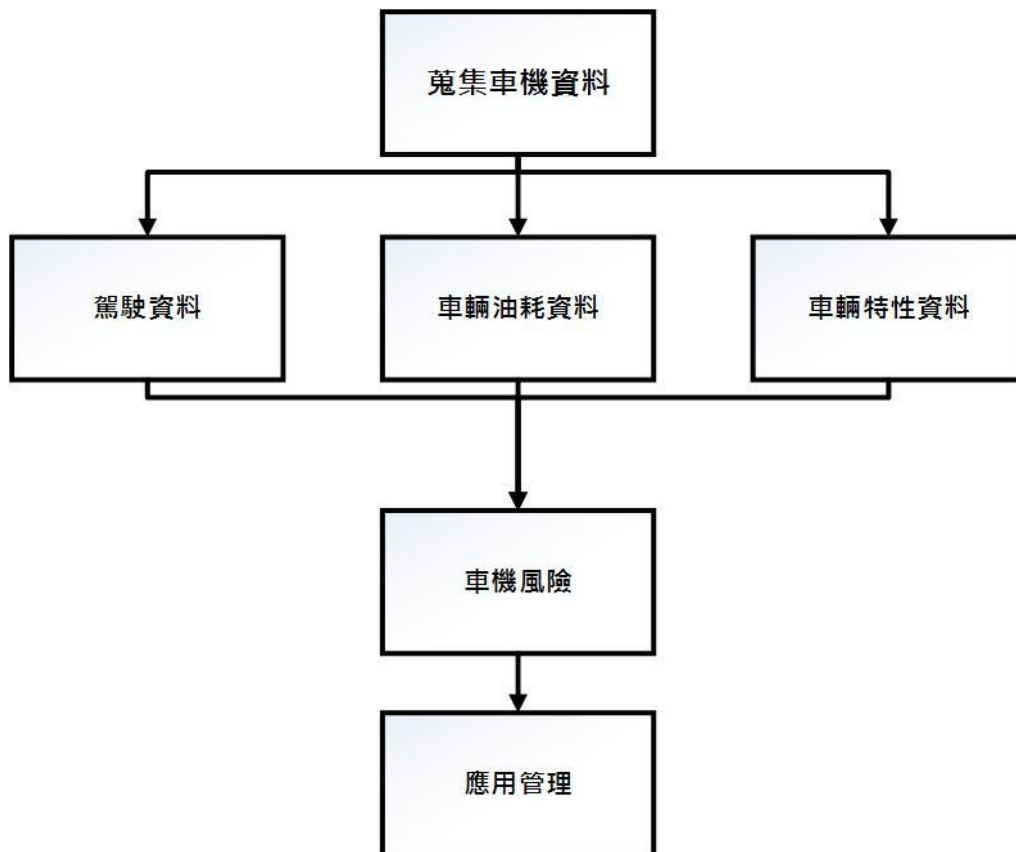


圖 6-71 車輛風險分析流程圖

3. 原始資料

對於機務分析，本研究的原始資料包含三個部分：車輛清冊、車輛維修和電子票證資料。車輛清冊主要記錄豐原客運營運車輛的基本資料，包含車別、車號、廠牌、年份和噸位等等(表 6.13)，車輛維修主要記錄一年內所有車輛進廠維修(表 6.15)。

表 6.15 車輛維修資料表

編號	資料行描述	屬性	資料行名稱
1	流水號	日期	ID
2	車號	數字	BUSID
3	維修日期	日期	MAINTENANCE_DATE
4	維修時間	日期	MAINTENANCE_TIME
5	故障原因	文字	SUBJECTS
6	進場日期	日期	REPAIRER_DATE
7	進場時間	日期	REPAIRER_TIME
8	完修日期	日期	COMPLETION_DATE
9	完修時間	日期	COMPLETION_TIME

資料來源：豐原客運提供暨本研究整理

4. 資料轉換與處理

資料前處理的動作，主要目的在於過濾並去除不需要的資料，有助於資料探勘的進行與最後歸責的品質。

- (1) 清除不需要的資料:將不需要的資料刪除。車輛清冊資料部分，本分析主要對象是各車輛特性及車牌號碼進行機務分析，因此無需使用引擎號碼及座位數。
- (2) 去除遺漏的資料:刪除缺漏值及錯誤的資料，減少對資料探勘結束的影響。
- (3) 資料轉換:將資料的表示方法，轉換成相同的格式，其目的在協助之後探勘的進行。如：車輛出廠年份原本的 2003.01、2004.04，轉換成車齡，12 年、11 年;車輛廠牌匿名化。

- (4) 資料歸類: 將車輛維修資料內的故障原因分類，分成故障及檢修 兩大項目，故障的部分主要以車輛在行駛過程中發生拋錨，而檢修則是只要車子出現異常就歸類於此。

三、欄位定義

透過資料前處理，將資料轉換、刪除資料本身欄位值不完整性及豐原客運所提供的資料定義，整理出本分析進行探勘所需之欄位如表 6.16。

表 6.16 車輛特性欄位一覽表

欄位名稱	說明
車號	五碼
廠牌	A、B、C、D 和其他
年份	16、14、13、12、11、10、9、8、7、6、5、4、3、2、1 (單位:年)
噸位	5.4、14、14.5、16、17(單位:公噸)

資料來源：豐原客運提供暨本研究整理

6.7.3 分析結果與視覺呈現

利用 Tableau，針對車輛特性和車號進行油耗分析和機務分析，挑出哪些特點是油耗和機務的主因，並提出相關的處理。

一、油料分析:

1. 哪一車號較耗油?

把公司內所有的車號耗油量(一公里平均消耗多少柴油公升數)做排名，圖 6-72 和表 6.17 所示，把前 15 名耗油車號列出，得到 685-FX 此車號耗油為平均為 4.55 公升/公里，與其他車號相比高出許多。

哪一部車號較耗油

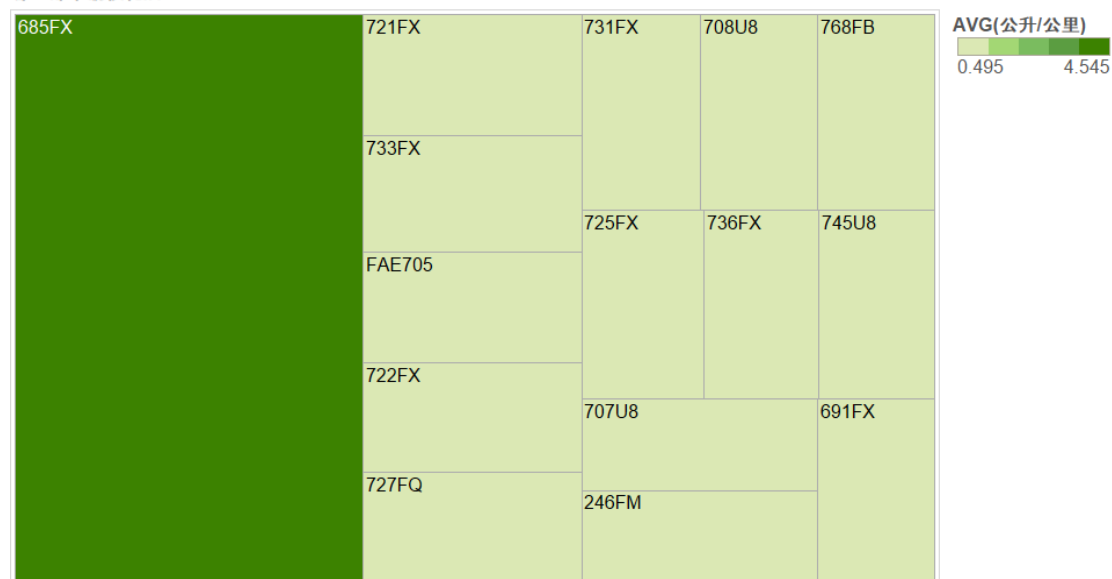


圖 6-72 車號油耗量圖

表 6.17 車號油耗量

車號	685FX	721FX	733FX	FAE705	722FX	727FQ	731FX
油耗量	4.55	0.61	0.59	0.56	0.55	0.55	0.53

註:油耗量單位為 L/KM

表 6.18 將車號 685-FX 列出，發現此車號都是同一駕駛(編號 1404)所開。由於車耗 685-FX 一年平均耗油量約 4.55 公升/公里，此車輛為國瑞 2005 年出場的車輛，已經使用超過 10 年，在其他 2005 年出廠的車大多油耗量約 0.51 公升/公里，高出其他車輛近十倍以上，建議應與該駕駛(編號 1404)討論其開車方式或是車輛是否有異常。

表 6.18 車號 685-FX 對應之駕駛編號表

日期	車號	駕駛編號	日期	車號	駕駛編號
1030501	685FX	1404	1040304	685FX	1404
1030501	685FX	1404	1030503	685FX	1404
1030503	685FX	1404	1030501	685FX	1404

1040304	685FX	1404
---------	-------	------

2. 哪一部車重較耗油？

首先，將所有噸數的車輛數以下圖 6-73 所示。

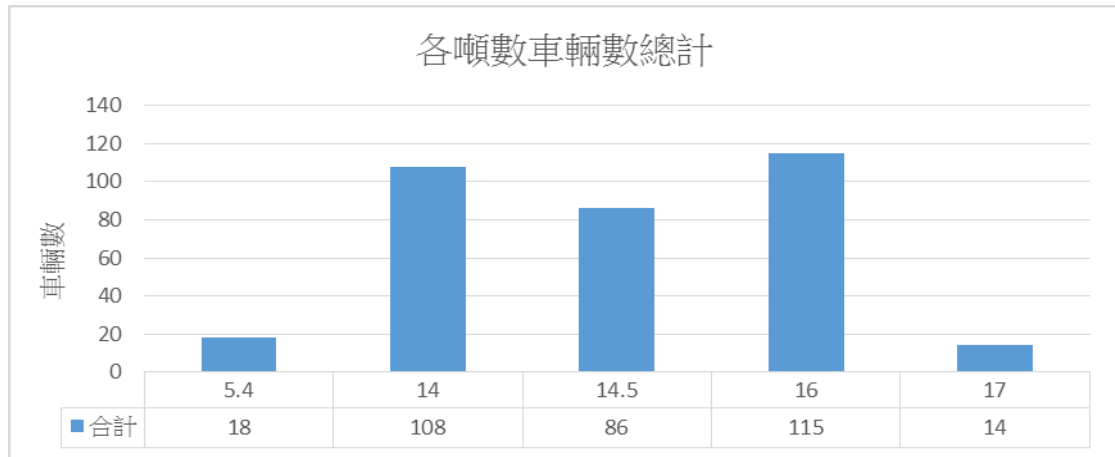


圖 6-73 各噸位數車輛數

將所有車輛的噸位數列出，可觀察發現以 14 公噸、14.5 公噸、16 公噸為主要的營運車輛，從圖 6-74 和表 6.19 得知，14 公噸、14.5 公噸、16 公噸和 17 公噸平均耗油為 0.37 公升/公里到 0.46 公升/公里之間，差距不大，而噸位數最輕的 5.4 公噸的車載耗油最少，一般情況，車載越輕耗油量越少。

哪一部車重較耗油

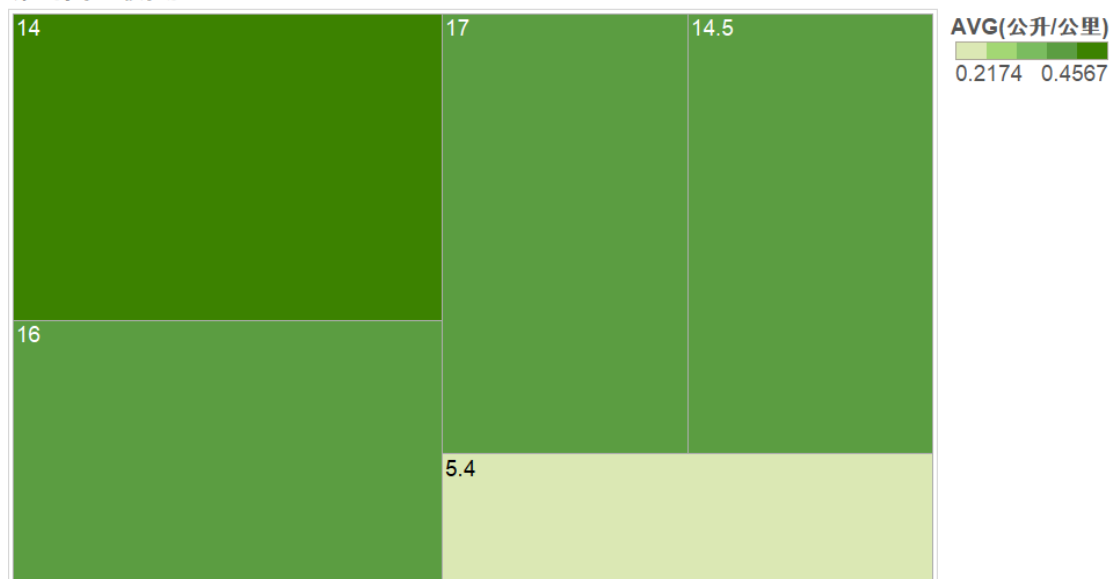


圖 6-74 噸位數耗油量圖

表 6.19 噸位數耗油量表

噸位數	5.4	14	14.5	16	17
油耗量	0.22	0.46	0.37	0.39	0.38

註：油耗量單位為(L/KM)

3. 哪一廠牌較為耗油？

首先，介紹豐原客運內所有廠牌的車輛數，如圖 6-75 所示。

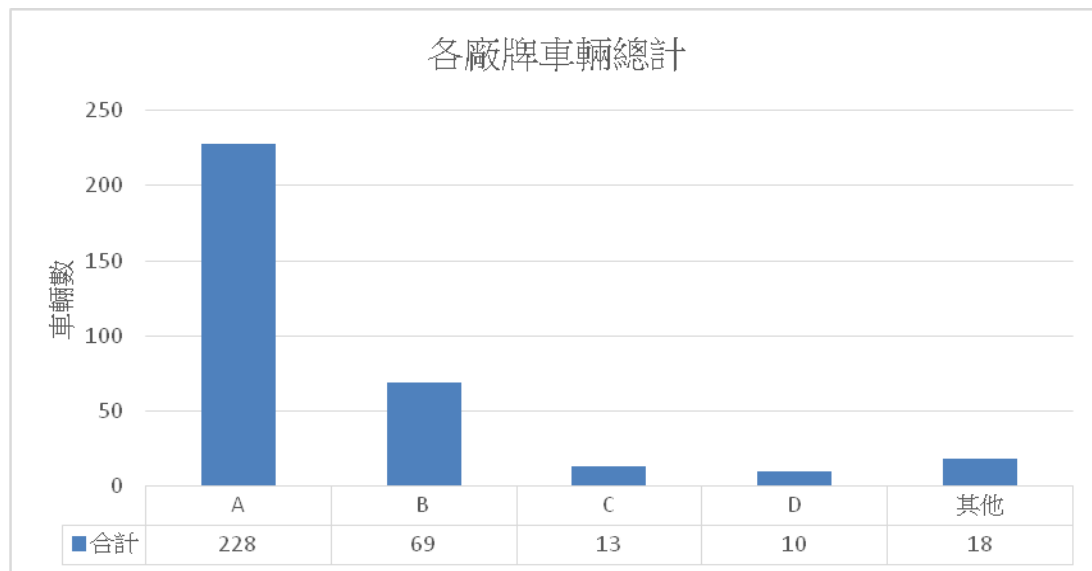


圖 6-75 廠牌車輛數

將所有車輛的廠牌列出，主要以 A 牌、B 牌、C 牌、D 牌和其他為主要營運車輛，從圖 6-76 和表 6.20 得知，呼應前一個分析哪一部車重較耗油，得到車載越輕的耗油量越少，而貴公司其他廠牌的耗油量大多相同，平均耗油約 3.4 公升/公里到 4.1 公升/公里之間。

哪一廠牌較為耗油

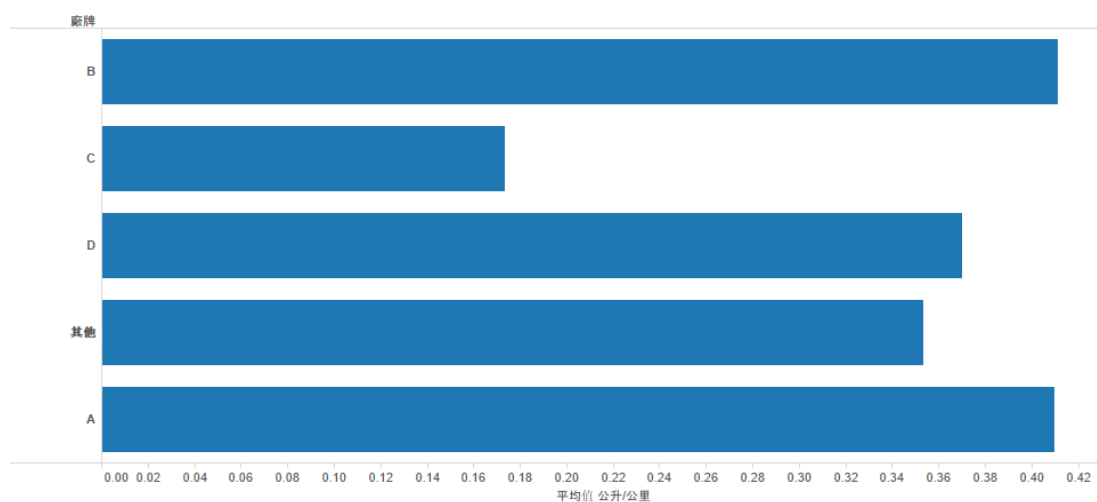


圖 6-76 廠牌耗油數

表 6.20 廠牌耗油量表

廠牌	C	D	其他	B	A
油耗量	0.17	0.37	0.34	0.41	0.41

註：油耗量單位為(L/KM)

4. 哪一車輛出廠年份較耗油？

首先，將所有車輛以車齡進行分類，以圖 6-77 所示。

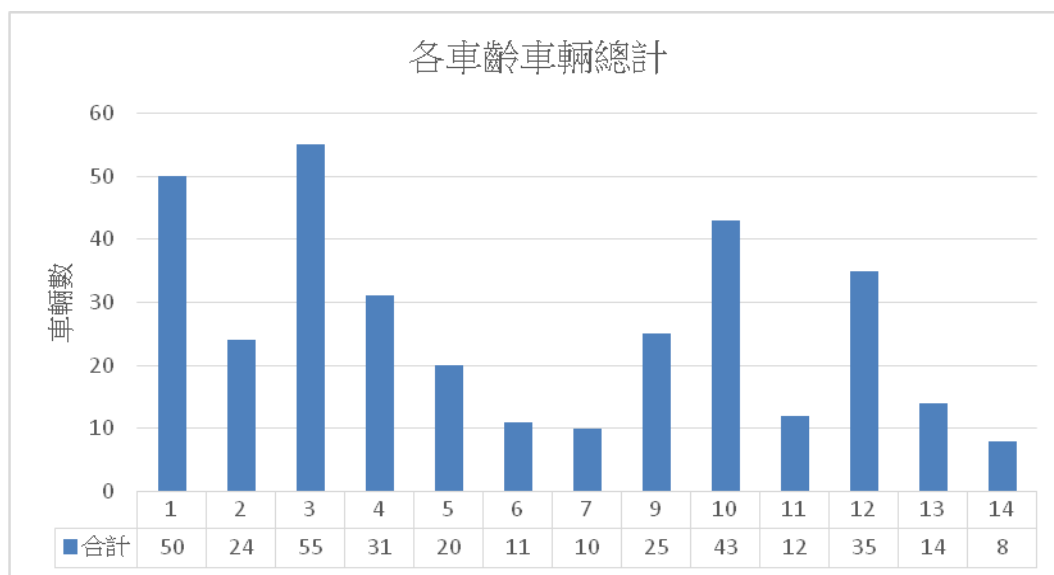


圖 6-77 各出廠年份車輛

從圖 6-78 和表 6.21 得知，只有 10 年車齡的車輛較耗油，平均 0.51 公升/公里，而其他出廠的年份都大多是 0.4 公升/公里左右。

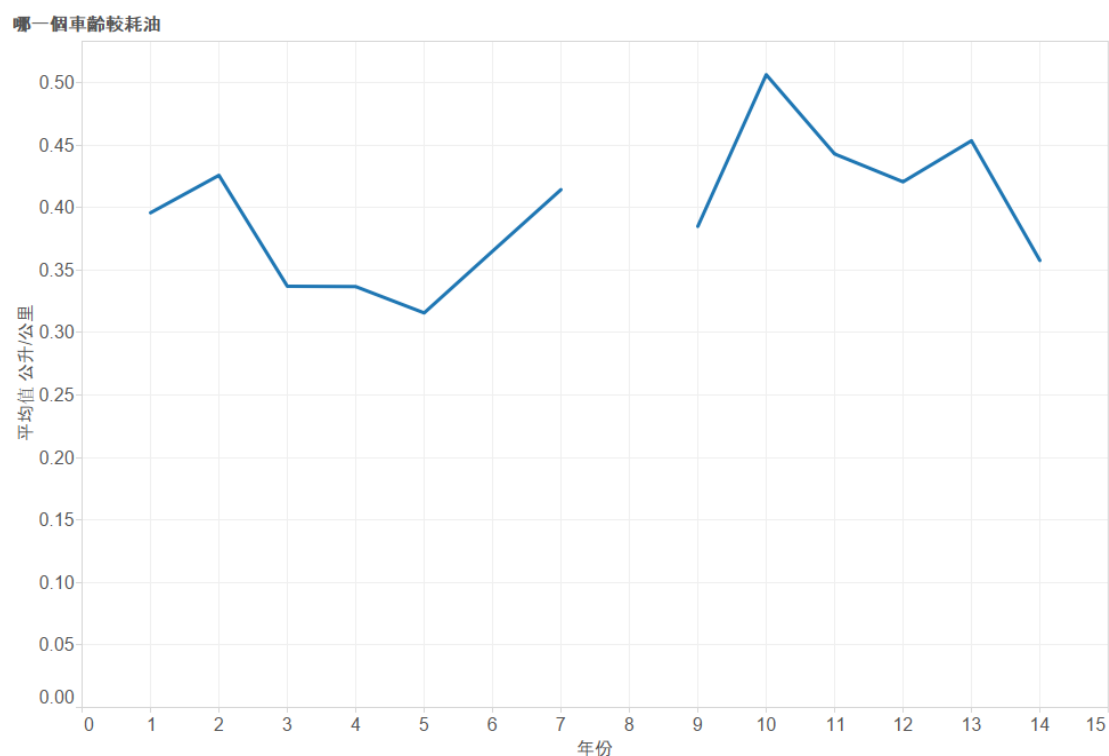


圖 6-78 車齡耗油圖

表 6.21 車齡耗油量表

車齡	10	13	11	2	12	7	1
油耗量	0.51	0.45	0.44	0.43	0.42	0.41	0.4

註：油耗量單位為(L/KM)

5. 廠牌和車載重做耗油量交叉分析

首先，將各廠牌車輛噸數列出如圖 6-79，橫坐標第一列為車輛載重，橫坐標第二列為廠牌，得知，廠牌國瑞 14 公噸、14.5 公噸和廠牌唐榮 16 公噸是豐原客運主要的營運車輛。

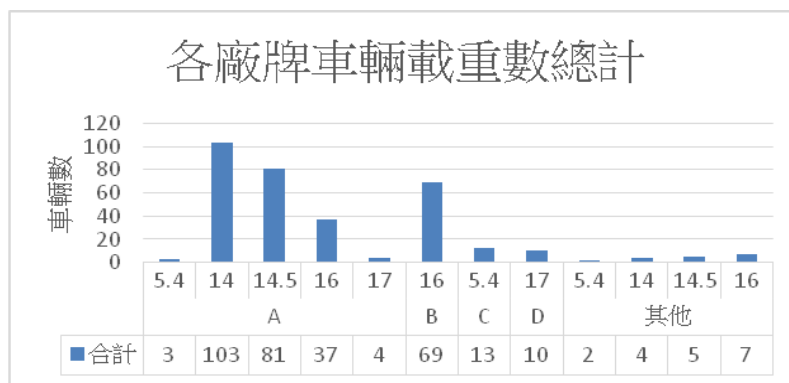


圖 6-79 各廠牌車輛載重數總計

從圖 6-80 得知，縱軸為車載重，橫軸為耗油量，藍色圓圈為廠牌 C 屬於中小型巴士，耗油量較少，約耗油量 0.17 公升/公里，唯一 17 公噸的 D，平均耗油量約 0.37 公升/公里。業者主力營運車種為 A，此廠牌的車子大多都是 14 公噸以上，耗油量 0.4 公升/公里；此外之該公司擁有 69 臺 B 的車子，耗油量平均約 0.4 公升/公里。

依照路線的特性來看，建議偏遠深山的路線應建議使用中小型巴士 C，由於噸位數少，耗油量較低，並且偏遠深山地區搭乘人數較少，因此能節省營運成本，另外，中長途路線應旅客量視情況改變巴士大小，雖說耗油量差不大，但是數量龐大的話能節省的開銷是不容忽視的。

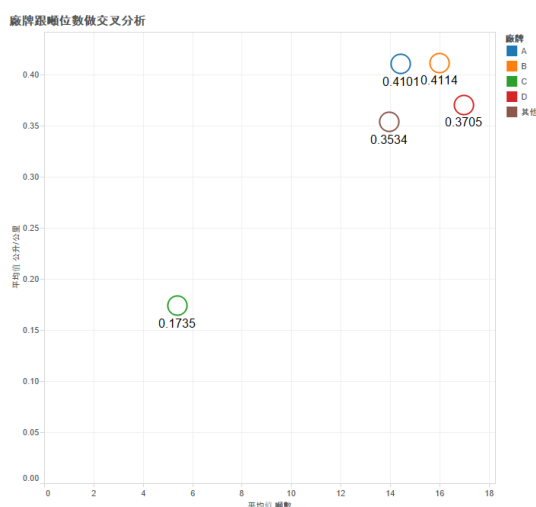


圖 6-80 噸數廠牌耗油交叉分析

6. 車齡和廠牌做油耗交叉分析

首先將各車齡的車輛數以圖 6-81 所示，橫坐標第一列為車輛出場年份，橫坐標第二列為車輛廠牌，業者營運的主力車種為廠牌 A，主要以 2003 年(車齡 12 年)、2005 年(車齡 10 年)和 2012 年(車齡 3 年)出場的車子為主，超過 12 年車齡的有 50 臺。

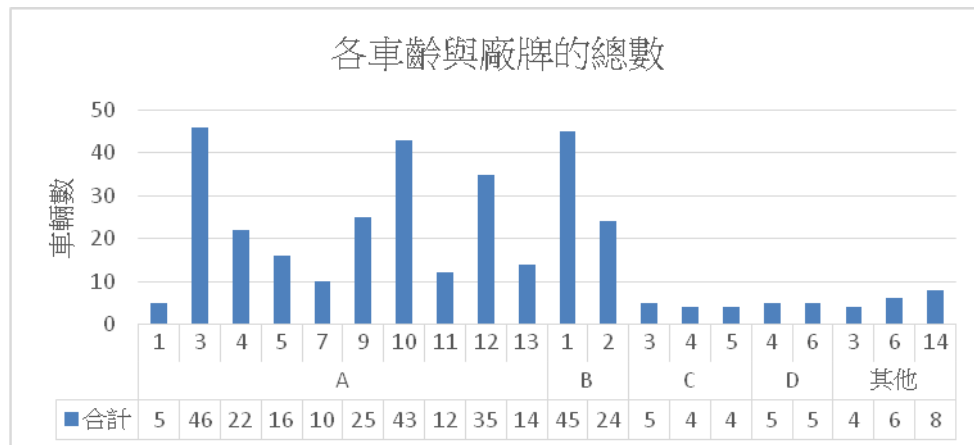


圖 6-81 各廠牌出廠年分車輛數

從圖 6-82 得知，A 廠牌的車子主要買 2002 年後出廠，在 2008 年(車齡 7 年)以前出廠的車耗油量較大，尤其是 2005 年(車齡 10 年)出廠的車子平均耗油量高達 0.5 公升/公里，而近 5 年內出廠的車子耗油量就較低，B 廠牌是營運車輛第二多的廠牌，也是較新的車子，使用期間大約二到三年，平均耗油量約 0.4 公升/公里。

建議可將 2005 年(車齡 10 年)以前的出場的車子能陸續淘汰，由於耗油量大並且車子老舊容易出現故障的機率高。

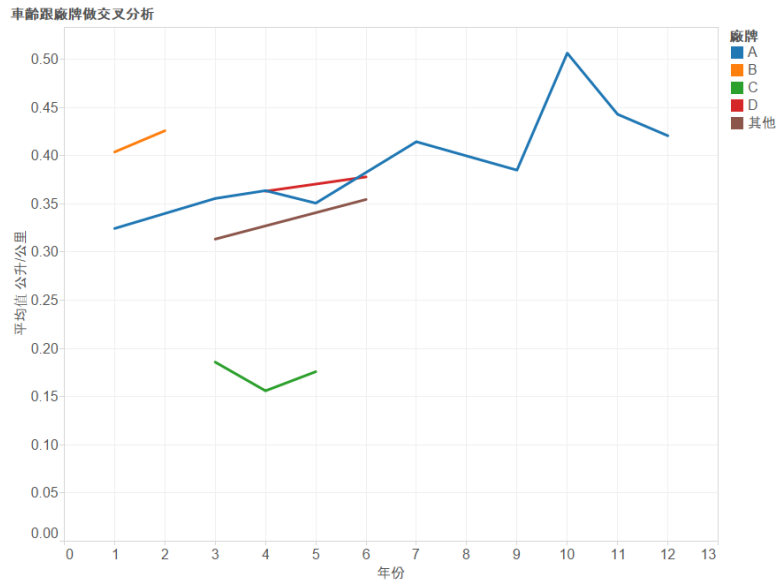


圖 6-82 車輛廠牌出廠年份之交叉分析

二、機務分析:

1. 哪一部車號經常發生車輛故障(路上拋錨)?

圖 6-83 將 2014/5 到 2015/4 所有車輛故障(路上拋錨)的車號進行分析，將前 20 名的車號列出，發現 536-FH 車號在這一段期間內已發生 5 起車輛故障，平均兩個月就發生一次，頻率極高，其餘的車號都小於 3 次。

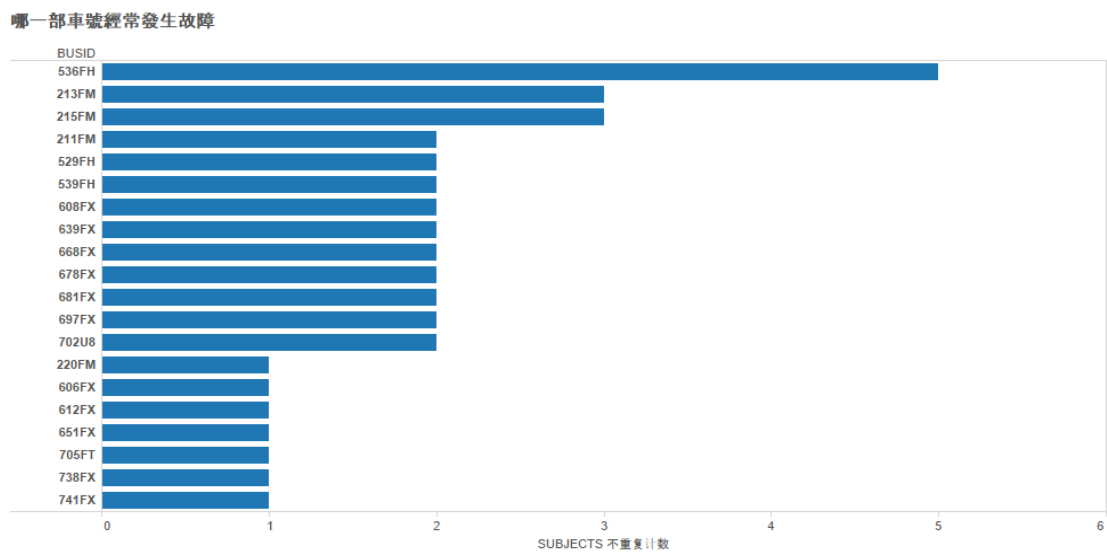


圖 6-83 車號車輛故障圖

將 536-FH 在 2014/5 到 2015/4 所有發生車輛故障(路上拋錨)列出，並與當日駕駛做連結可得表 6.22，發現駕駛編號以下幾位駕駛員開 536-FH 時，經常發生車輛故障(路上拋錨)。

建議未來應深入瞭解 536FH 所有駕駛的行為及調閱形車紀錄器，是否有操作不當等原因造成車輛容易故障。

表 6.22 車號 536FH 對應駕駛編號表

車號	駕駛編號	車號	駕駛編號
536FH	263	536FH	1428
536FH	1013	536FH	1448
536FH	1055	536FH	1449
536FH	1136	536FH	1461
536FH	1172	536FH	1464
536FH	1419	536FH	1468

資料來源：豐原客運提供暨本研究整理

2. 哪一廠牌經常發生車輛故障(路上拋錨)?

從圖 6-84 與圖 6-85 得知，廠牌國瑞車款經常發生車輛故障(路上拋錨)，公司內有 228 輛，有 27 臺曾發生車輛故障，發生車輛故障的比例約 0.11，另一個廠牌唐榮擁有 69 臺，其中有 6 臺車曾發生車輛故障，發生車輛故障的比例約 0.087，從以上兩個數字得知，車輛發生的故障(路上拋錨)的機率還算低。

建議：從此分析得知，購買車輛的廠牌對於車輛故障沒有直接的關聯。

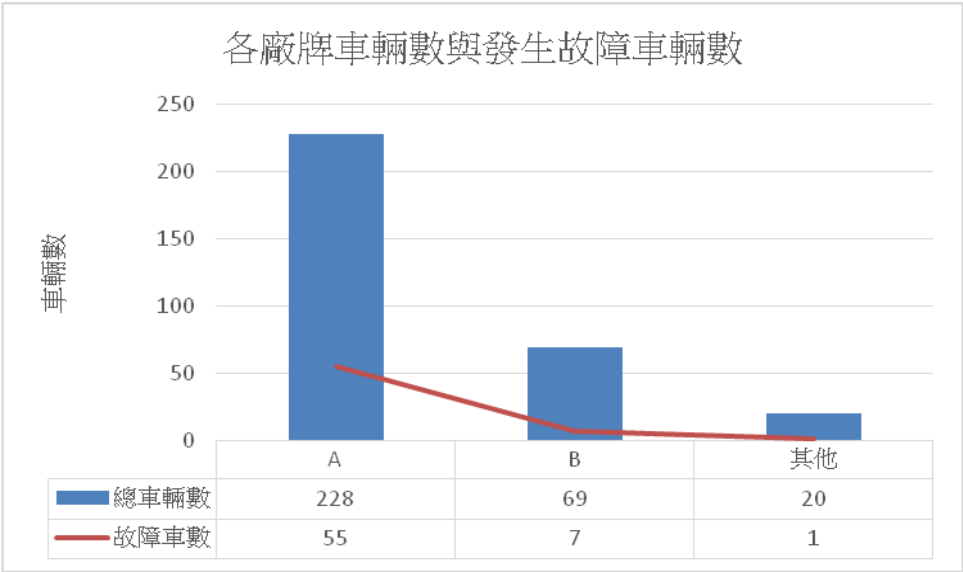


圖 6-84 各廠牌車輛數與發生故障車輛數

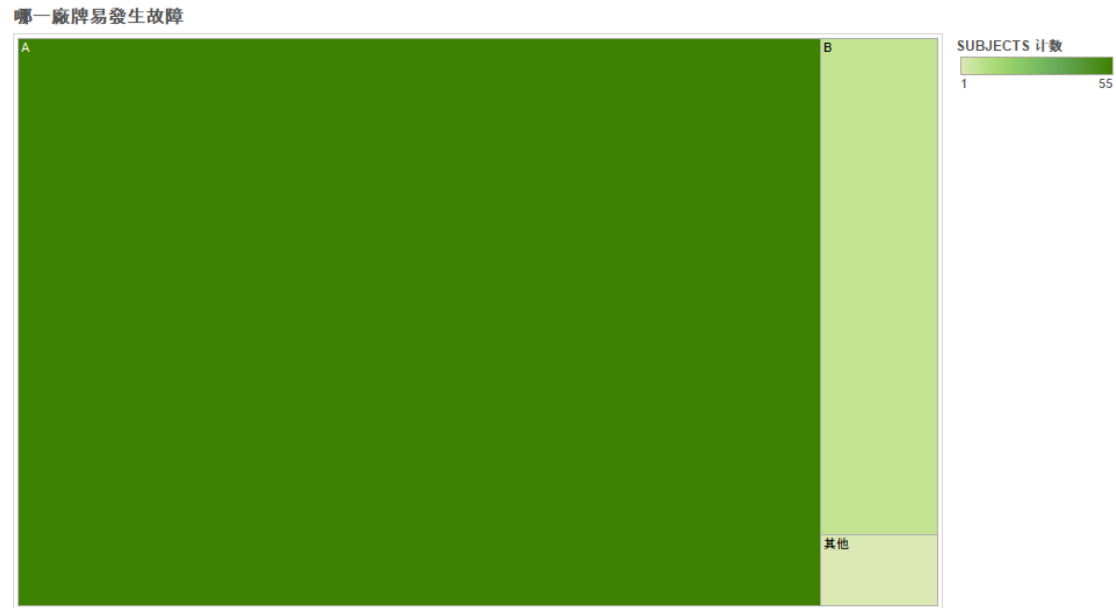


圖 6-85 廠牌車輛故障

3. 哪一出場時間的廠牌經常發生車輛故障(路上拋錨)?

圖 6-86 的縱座標為故障發生次數，而橫坐標為車齡，從圖 6-86 得知使用大約使用超過 7 年以上的車輛較經常發生車輛故障，這三年都超過 8 次以上。

建議未來業者可對這些車輛優先考慮汰換，或是建議 5 年以上車齡需要進行詳細的修檢。

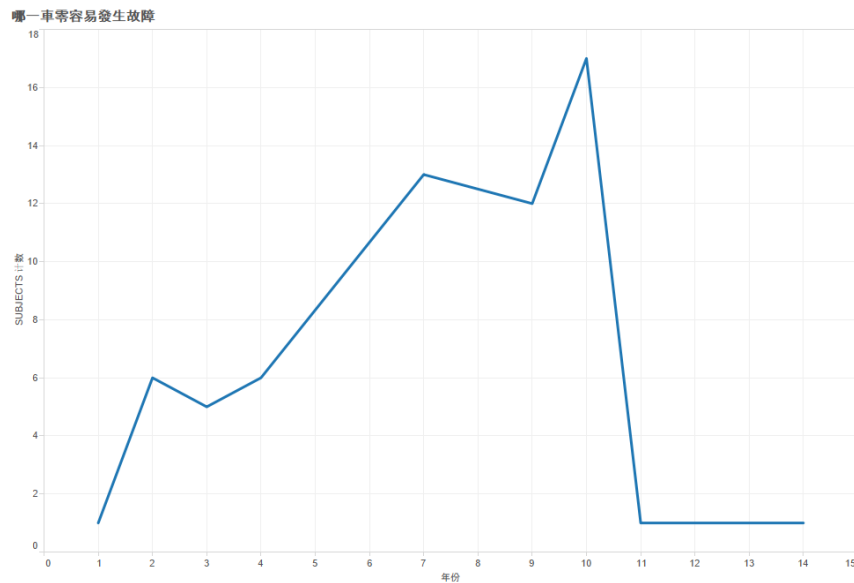


圖 6-86 年份廠牌車輛故障

4. 車輛故障進行車齡和耗油量的交叉分析

圖 6-87 的橘色線為故障次數，而藍色長條圖為車齡平均油耗，車齡 10 年(2010 年購買)和 7 年(2008 年購買)出廠的車款車輛故障頻率高且耗油較高。

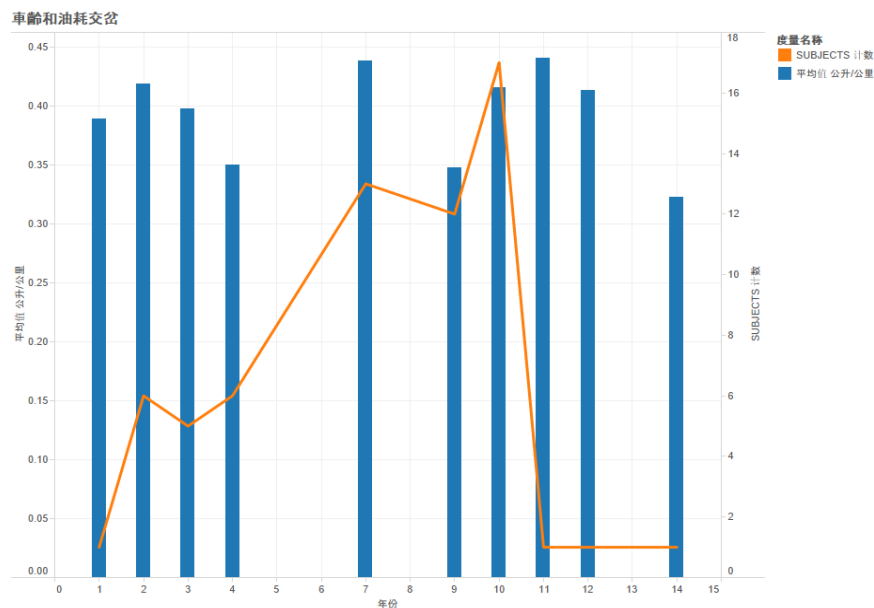


圖 6-87 車齡/油耗/故障交叉分析圖

5. 哪一臺車號較容易發生異常(檢修)?

將 2014/5 到 2015/4 所有車輛異常(檢修)的車號進行分析，將前 20 名的車號列出，發現車號 651FX 經常發生車輛異常，一年內多達 40 次，平均一週快一次。

建議未來應深入瞭解 651FX 所有駕駛的行為及調閱行車紀錄器，是否有操作不當等原因造成車輛容易車輛異常。

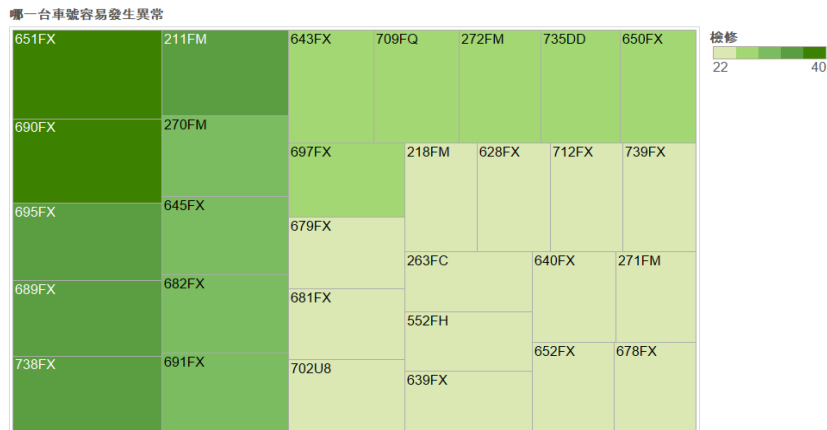


圖 6-69 車輛檢修

6. 每臺車平均檢修公里次數

在此依照各車輛一年的行駛距離來換算每臺車每多少公里要檢修一次，在此能看出豐客車子維修保養的好壞，若能不要在短距離內經常發生進廠維修的情況，既能節省成本也能提高車輛的壽命。

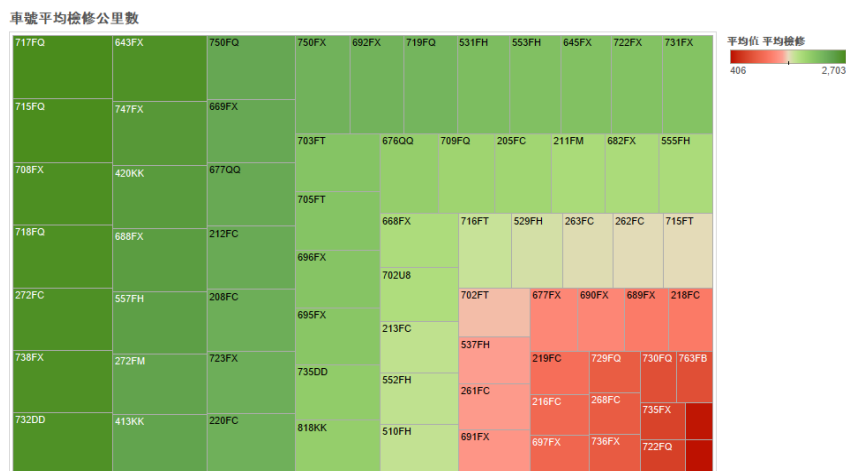


圖 6-70 平均檢修公里數

7. 車輛檢修進行廠牌和耗油量的交叉分析

在此發現廠牌與車輛檢修有直接的關係。

車齡與油耗交叉分析

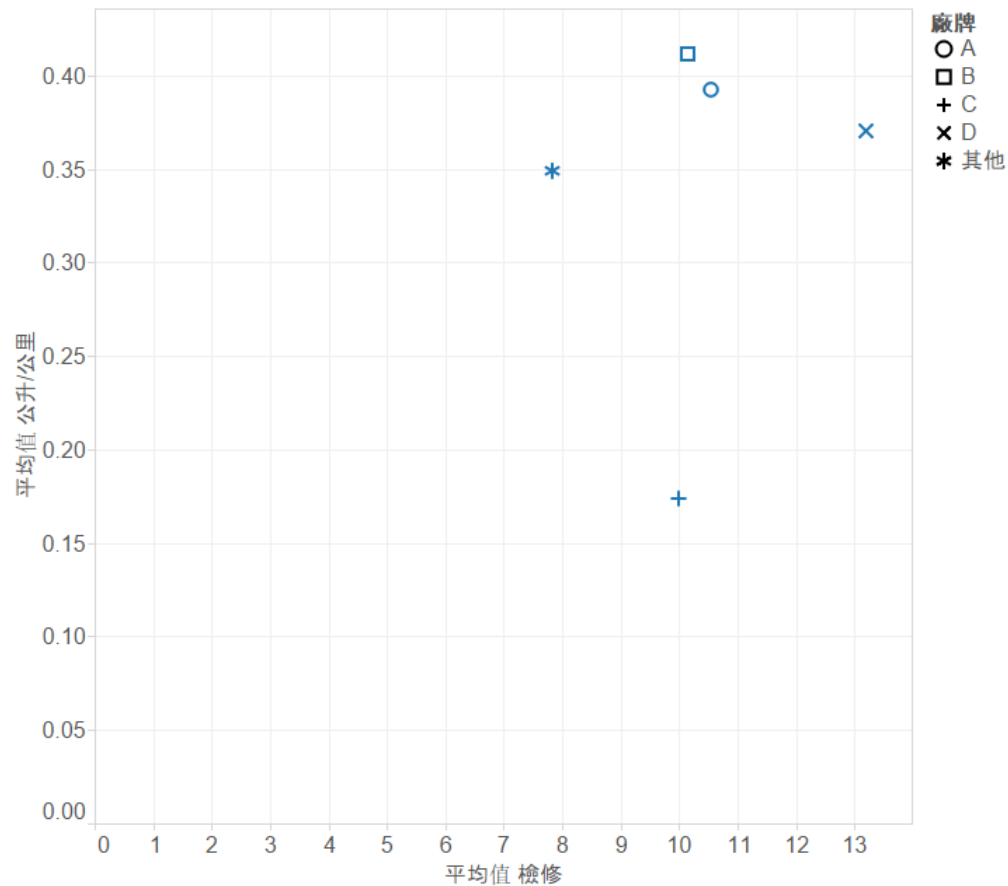


圖 6-71 車齡/油耗/檢修交叉分析圖

本研究有大約計算豐原客運整個營運車輛平均多少公里會發生一次故障及檢修。

表 6.23 檢修/故障平均公里數

公司平均故障公里數	公司平均檢修公里數
45957.19km	7881.08km

6.8 駕駛員風險分析

6.8.1 探勘分析目的

安全是大眾運輸系統最重要的基本條件，公車運輸系統的安全性雖遠比私人運具高，但基於「精益求精」的原則，本計畫亦利用巨量資料對已發現或潛在駕駛風險進行分析，期能增進公車運輸系統之安全。考慮的危險駕駛包括：(1)超速駕駛(2)超轉駕駛(3)急加速駕駛(4)急減速駕駛。

6.8.2 演算處理流程

本計畫並非使用如表 6.24 的車機內建事件偵測紀錄，而是根據第 5.4 節介接、資料清理數位行車記錄器原始記錄(RAW DATA)來分析。

表 6.24 TTIA 危險駕駛相關事件表

事件代碼	事件名稱	說明
0x0002	超轉超速	瞬時轉速超過特定值
		瞬時時速超過特定值
0x0002	急加/減速	單位時間加速/減速超過限制速度範圍

資料來源：營業大客車車載機產業標準 V1.5

經第 5.4 節的資料清理後雖已無明顯的異常資料，但仍根據以下四項原則做駕駛風險樣態擷取：

1. 超(高)轉：擷取每分鐘轉數超過 3000 者，一分鐘內取最大轉數值。
2. 超(高)速：擷取時速超過 70km/h 者，一分鐘內取最大時速值。
3. 急加速：3 秒內（含 0～3 秒）時速遽增超過 30km/h 者
4. 急減速：3 秒內（含 0～3 秒）時速遽減超過 30km/h 者

其中第 1,2 項可直接從「RPM」與「Speed」欄位之數值判斷，第 3,4 項則須根據時間排序後跨行比對，會耗用較多的計算資源。另外，數位行車記錄器時間資訊精確至秒以下，故一旦偵測到危險駕駛樣態時，通常該時點的前後幾秒亦同屬危險狀態，故需適當的彙整，本分析以一分鐘內取最危險值來分析。

此外，由於僅有標竿廠商-豐原客運配合本研究提供數位行車紀錄器，為能夠執行較全面之分析，本研究額外採用公車動態定時資料（每 20 秒一次）中的速度做分析。至於在駕駛代碼的連結上，數位行車記錄器並沒有與司機做連結，故本研究勾稽電子票證中的司機代碼來呈現。

6.8.3 視覺呈現暨應用

駕駛風險樣態經偵測擷取後，在視覺化呈現上主要有兩大類，分述如下：

6.8.3.1 地理空間分佈

以下用高速駕駛與高轉速駕駛為例說明地理空間視覺化呈現方式，其中圖 6-88 為高速駕駛分佈，由於係採全年度公車動態資料為分析依據，資料筆數會讓畫面過於複雜，故搭配圖 6-89 的業者篩選，以及暨有的日期篩選、路線篩選來呈現，當滑鼠移至每一點時，即時產生該點的詳細資訊，包括行駛路線、車牌、車速以及最重要的「司機代碼」，此代碼並非公車動態資訊系統暨有的資料，而是本計畫融合公車動態與電子票證紀錄所建置。

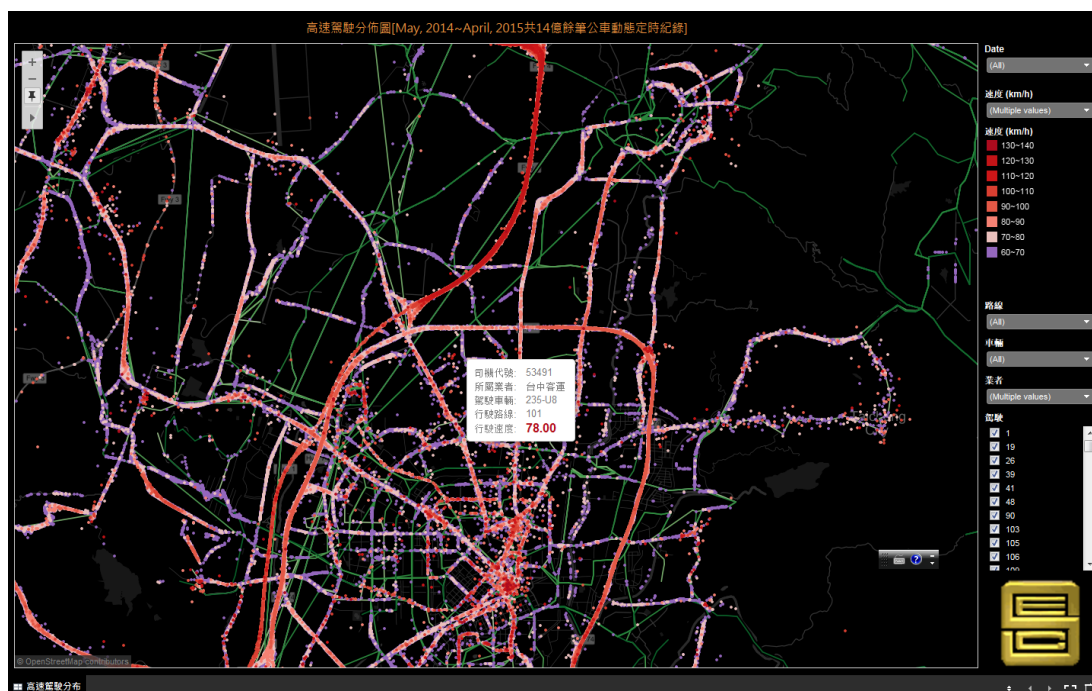


圖 6-88 高速駕駛空間分佈視覺化呈現

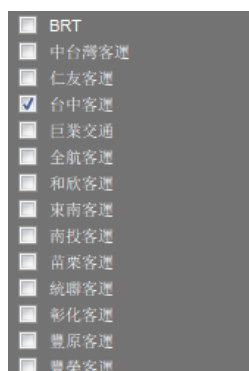


圖 6-89 由圖 6-88 右側儀表板篩選欲檢視業者

由於資料已是數字代碼（以圖 6-89 為例為 53491），故本報告如實呈現並無個資疑慮，未來若要應用在業者自主管理時，可根據各業者保管的駕駛代碼對應表得知確切的駕駛自然人，為便於追蹤管理，儀表板也支援「僅顯示指定編號司機」功能，例如圖 6-90 右側勾選「代號 53491」，左側就只會顯示該司機員的高速駕駛資料點。

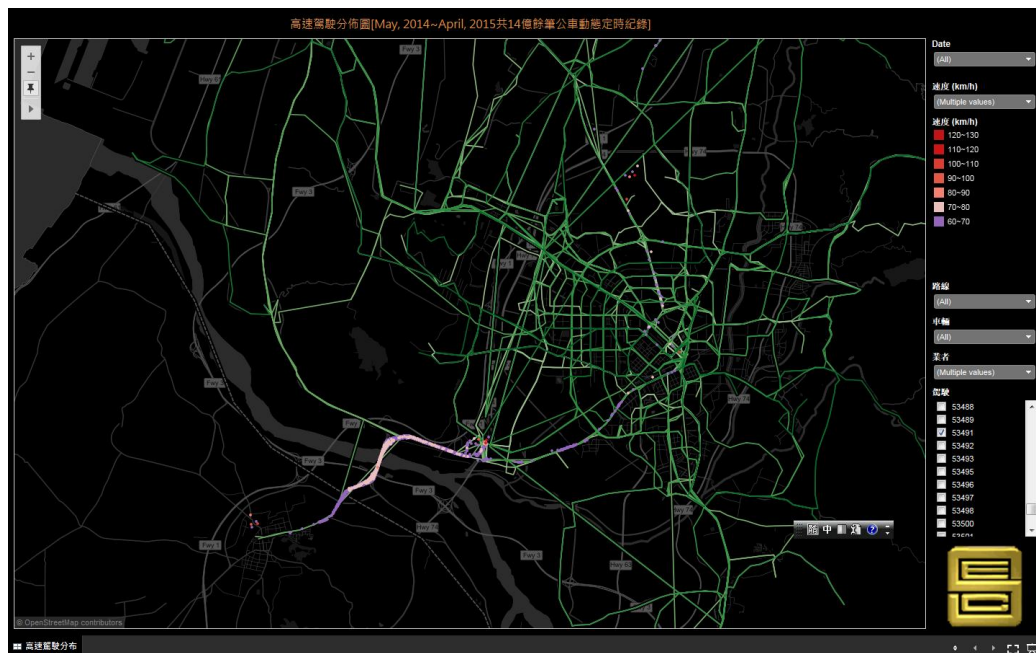


圖 6-90 高速駕駛空間分佈視覺化呈現

同樣的呈現方式亦可運用在高轉速駕駛的分佈如圖 6-91，由於本計畫僅有標竿業者-豐原客運提供數位行車記錄器資料，故圖 6-91 僅呈現豐原客運的高轉速駕駛分佈，至於駕駛員的部份仍是以電子票證上的司機代碼呈現，應用上亦可透過右側的儀表板指定（篩選）特定司機的高轉速駕駛分佈。

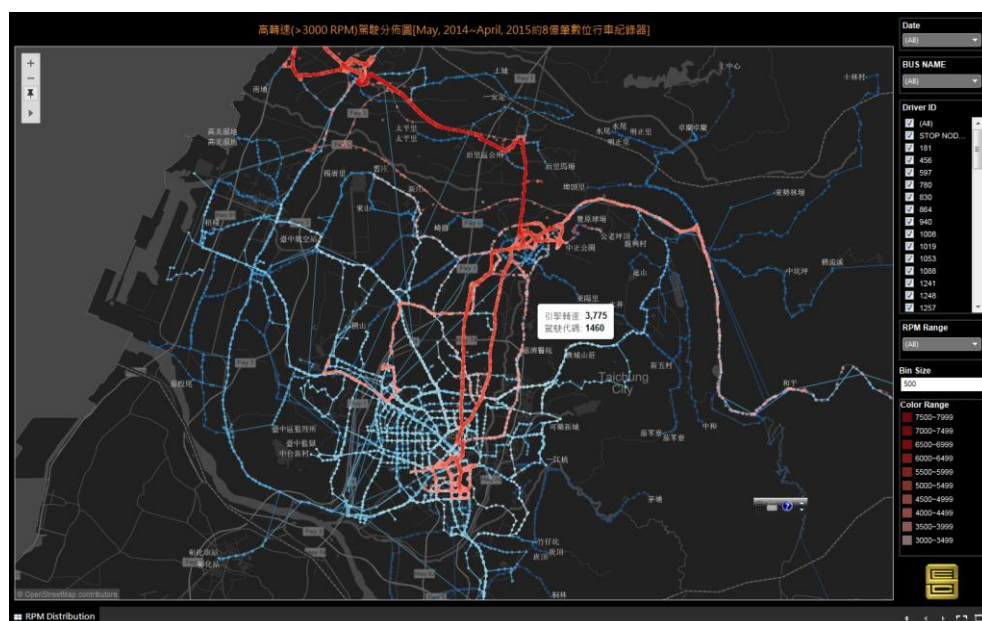


圖 6-91 高轉速駕駛空間分佈視覺化呈現示意圖

6.8.3.2 引擎轉速/車速/行駛軌跡對應分析：

圖 6-92 為全年(2014/05/01~2015/04/30)各車輛之引擎轉速與車速對應圖，本計畫將原始約 60 餘 GB 資料去蕪存菁、格式轉換、索引壓縮為適合 Tableau 呈現之格式後，透過圖 6-92 可作為數位行車記錄器資料異常比對、不同車輛駕駛行為差異比較。

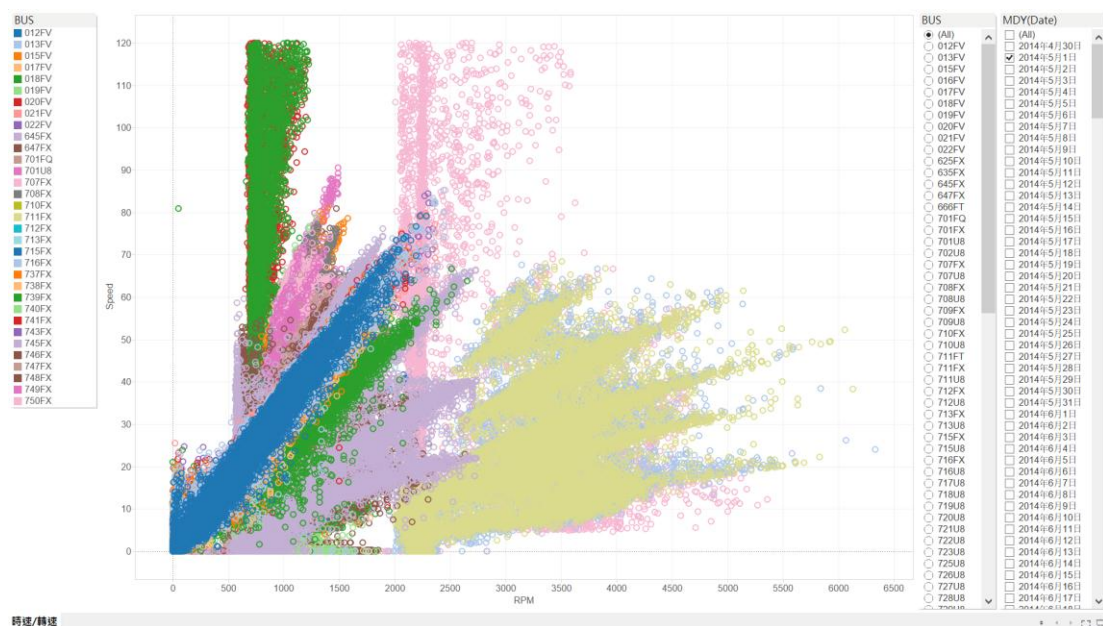


圖 6-92 車速與引擎轉速視覺化分析圖

為使圖 6-92 在使用上能更便利，本研究近一步將資料與地圖勾稽如圖 6-93，用統一的顏色來代表每一車輛之行駛軌跡及速度轉速對應，同時當使用者在圖 6-93 點選任一資料點時，互動儀表板會即時聚焦該車輛當天的行駛路線軌跡如圖 6-94，此儀表板將更有助於檢視不同車輛駕駛行為差異。

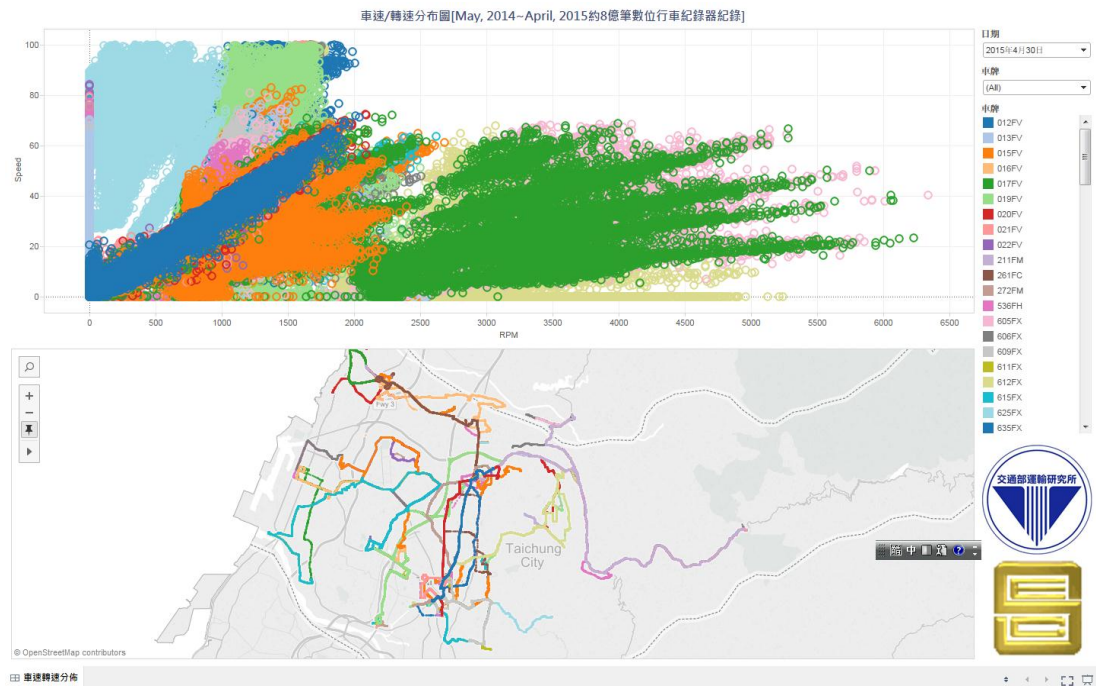


圖 6-93 車速與引擎轉速勾稽路線視覺化呈現圖

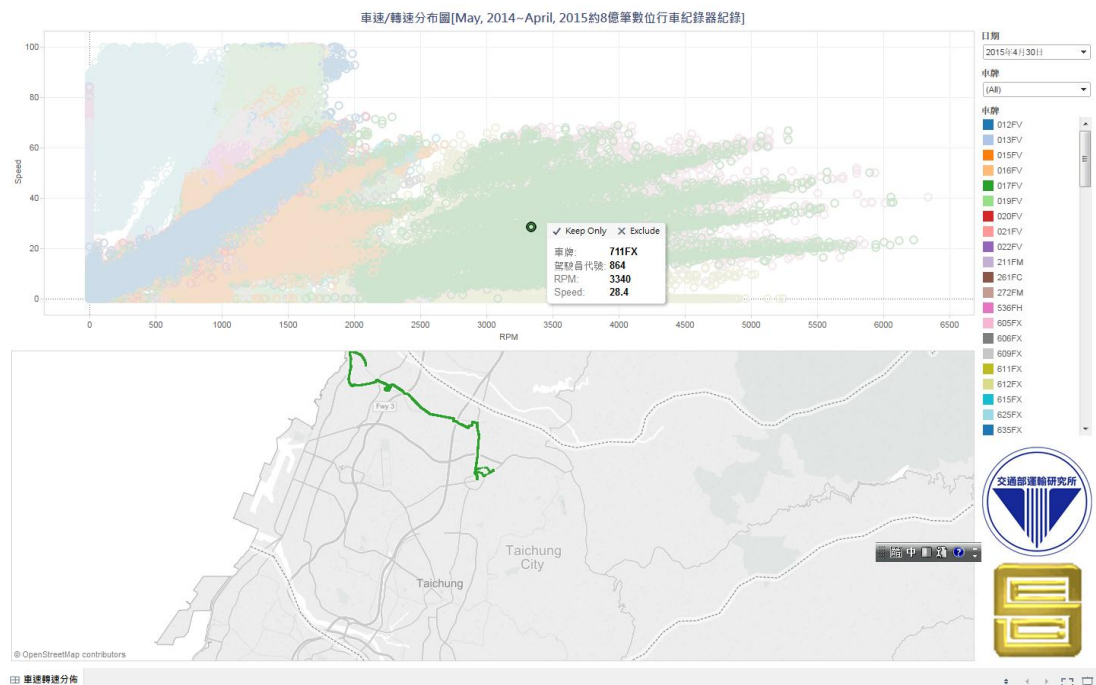


圖 6-94 選取圖 6-93 速度與轉速資料即時顯示路線互動圖

6.9 車機穩定度分析

6.9.1 探勘分析目的

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曾經「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。

6.9.2 演算處理流程

回傳延誤之計算係比較 GPS 時間與資料寫入系統時間做比對，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，建議這兩者之間的時間差應在 3 秒以內視為正常，除計算延誤比例外，也要以多維度（包括月份/星期/時段/站牌/車輛）分析延誤發生的車輛。除回傳延誤外，回傳失敗是另一分析重點，分析流程係分析同車輛 A1 的定時資料，先排序後計算兩筆時間差值，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。

6.9.3 視覺呈現暨應用

本節先以單一業者為例分析，再擴充至臺中地區各業者各路線資料，最後說明影響因素分析以及未來的運用建議。所有分析均以臺中地區全年（2014/05/01~2015/04/30）公車動態資料約 14 億筆。

6.9.3.1 單一業者分析

本節先針對單一業者介紹 5 種不同分析維度來檢視資料傳輸延遲之分佈狀況。

1. 統計經常發生傳輸延誤之車牌或路線如圖 6-95。

	BUSID	NUM	Percentage		Route	NUM	Percentage
1	641-FX	215596	1.81205284602215	1	55	1271686	10.6883348278564
2	712-FQ	187032	1.57197660391295	2	51	961025	8.07727456144104
3	708-FQ	177172	1.48910474607802	3	12	748701	6.2927223968424
4	536-FH	160629	1.35006325072679	4	213	545491	4.58477206919178
5	651-FX	155373	1.30588733949146	5	206	517398	4.34865451319213
6	709-FQ	142077	1.19413640422035	6	63	493651	4.14906445152824
7	627-FX	131570	1.10582660601836	7	270	468810	3.94027947987739
8	616-FX	128176	1.07730053243908	8	207	445137	3.74131137738995
9	648-FX	118772	0.998261287907676	9	91	406858	3.41958198123751
10	686-FX	104380	0.877298632942135	10	208	403693	3.39298061916372
11	643-FX	104352	0.877063297037533	11	280	383020	3.21922707788366
12	640-FX	101487	0.852983391084484	12	92	335820	2.8225179815542
13	649-FX	100420	0.844015412148392	13	90	318817	2.67961025348451
14	650-FX	97151	0.816539945286083	14	653	312324	2.625037538178
15	711-FQ	96571	0.811665130119323	15	200	255081	2.14391849577996
16	666-FX	92847	0.780365454807228	16	220	224887	1.89014234208141
17	679-FX	88514	0.743947223570034	17	6606	189313	1.59114807528429
18	710-FQ	84415	0.709495728107016	18	237	186927	1.57109409427069
19	652-FX	84192	0.707621445723934	19	288	147233	1.2374718300821
20	632-FX	82822	0.696106796105897	20	235	146726	1.23321056923805

圖 6-95 資料傳輸延遲發生數較高之車牌與路線圖

2. 以年/月的方式呈現 365 天各日資料傳輸延遲平均值，檢視者可在右側挑選（可多選）欲分析之路線如圖 6-96。

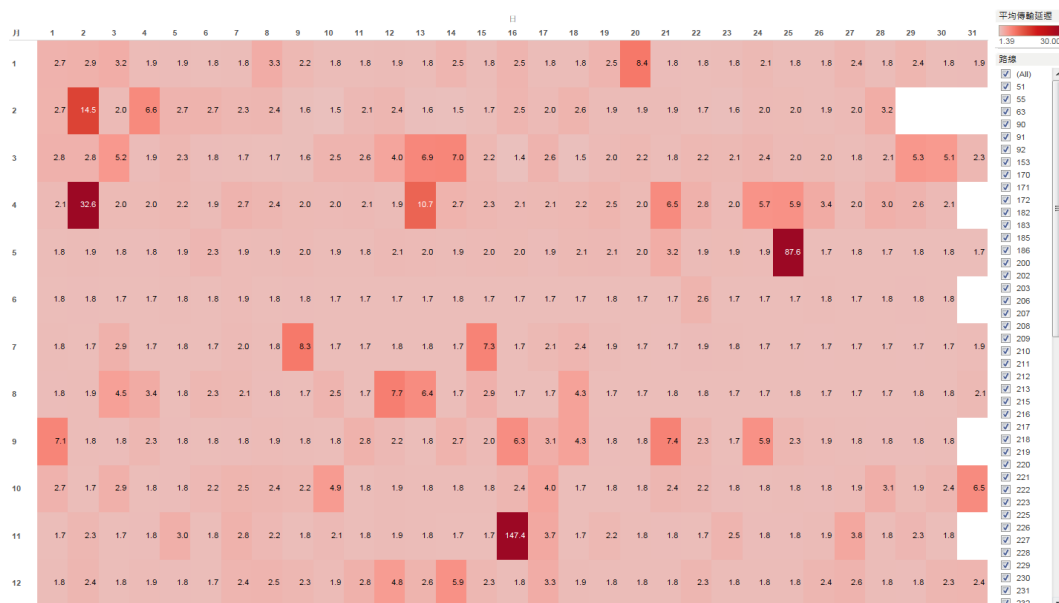


圖 6-96 車機延遲全年逐日分析圖

3. 以路線/月份/去回程維度分析傳輸延誤程度如圖 6-97。



圖 6-97 車機延遲各路線逐月去/返多維度比較圖

4. 以車牌/月份維度分析傳輸延誤程度圖 6-98。

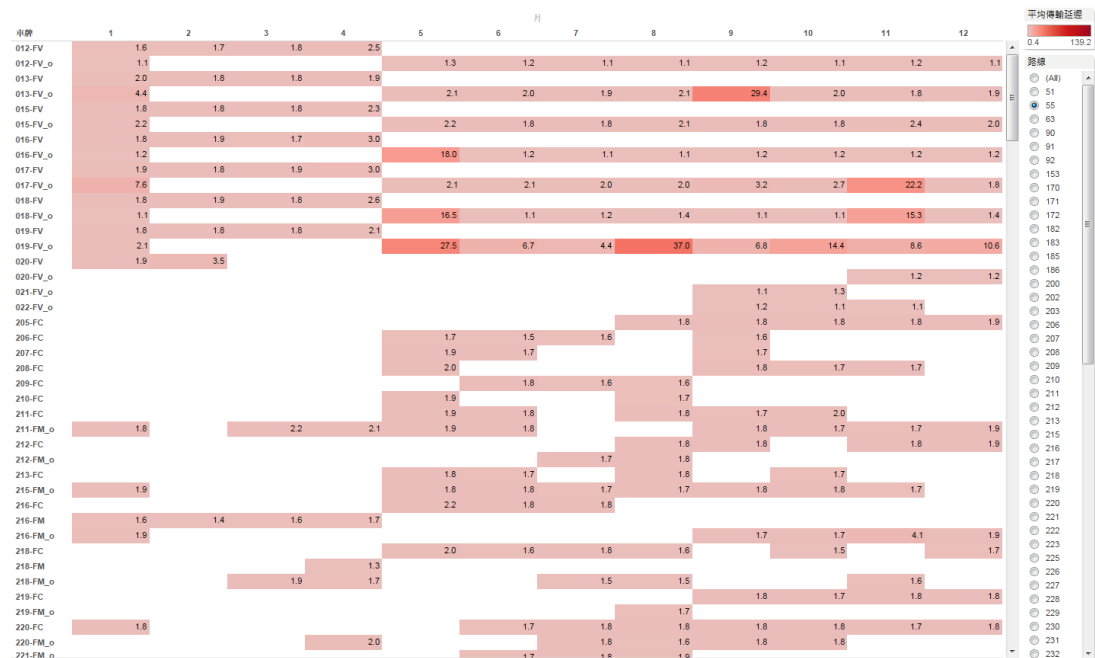


圖 6-98 車機延遲各車輛逐月多維度比較圖

5. 儀表板樣板：建議以圖 6-96 為基礎，在使用者選擇（可多選）路線後，除呈現全年各日之平均值外，當使用者點取某日期（例如圖 6-99 中的 5 月 25 日）時，下側動態依延遲量順序顯示該日期中資料傳輸延遲最高的車輛，將有助於動態查詢各異常日期之異常車輛。

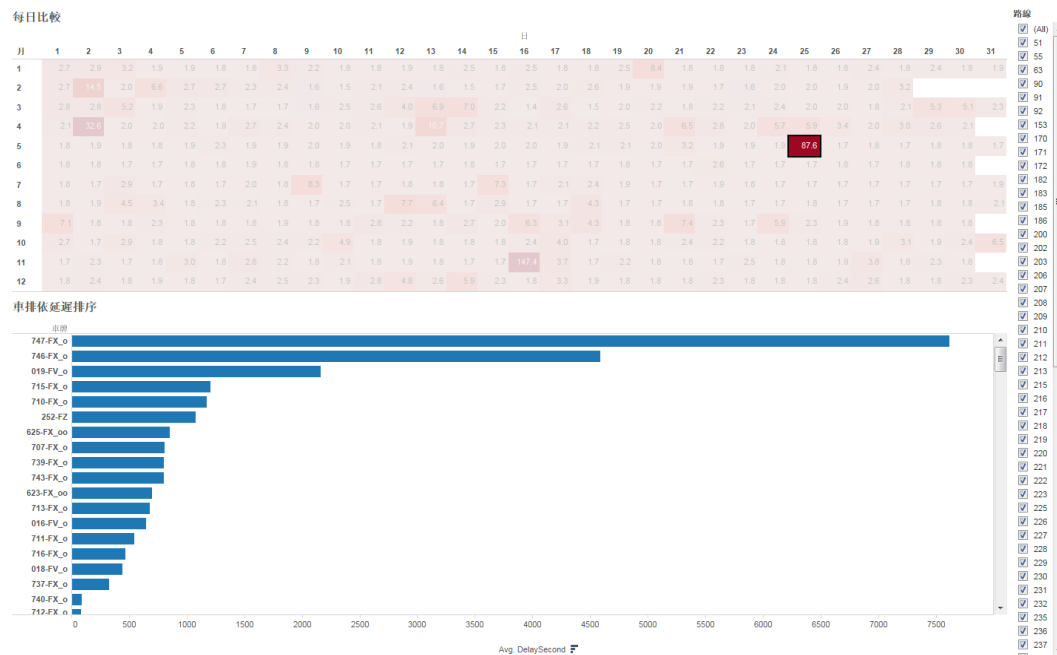


圖 6-99 車機資料傳輸延遲分析儀表板示意圖

6.9.3.2 整體穩定度分析

1. 固定門檻值分析

表 6.25 與表 6.26 分別就「定時資料」與「定點資料」的傳輸延遲做比較，若以 3 秒作為延誤標準，穩定度約在 90%~97%之間，而表 6.27 與表 6.28 則是針對表 6.26 中的延誤數分別做月份與時段的分析比較。

表 6.25 定時資料傳輸延誤比例比較表

	NAMEZH	總數	延誤數	延誤比率
1	彰化客運	11659989	1520716	0.130421735389
2	仁友客運	49936911	4828773	0.096697470934
3	和欣客運	16589470	1435138	0.086508972257
4	統聯客運	320558674	20488996	0.063916523438
5	東南客運	28925047	1804600	0.062388835530
6	巨業交通	64272476	3948569	0.061434835651
7	BRT	15159077	881949	0.058179597610
8	苗栗客運	1807656	101492	0.056145638329
9	台中客運	403649499	21872753	0.054187489527
10	中台灣客運	151696603	7866949	0.051859757202
11	豐原客運	235107005	12033391	0.051182613635
12	豐榮客運	20816313	950737	0.045672689491
13	南投客運	184950	7654	0.041384157880
14	全航客運	89000825	3254150	0.036563144218

表 6.26 定點資料傳輸延誤比例比較表

	NAMEZH	總數	延誤數	延誤比率
1	彰化客運	2881136	250841	0.087063227837
2	和欣客運	3470668	295380	0.085107535494
3	東南客運	7476299	551361	0.073747853048
4	苗栗客運	529822	38725	0.073090585139
5	巨業交通	17851...	1176...	0.065929378785
6	豐榮客運	7096498	450380	0.063465106310
7	台中客運	11292...	6426...	0.056904488823
8	南投客運	41583	2086	0.050164730779
9	統聯客運	86528...	4256...	0.049192986301
10	中台灣...	42764...	2005...	0.046905695143
11	豐原客運	64161...	2811...	0.043822361528
12	仁友客運	11935...	515488	0.043190558751
13	全航客運	24736...	976597	0.039480222635
14	BRT	1416672	52173	0.036827861353

表 6.27 各月份定點資料傳輸延誤比較表

	NAMEZH	1	2	3	4	5	6	7	8	9	10	11	12
1	BRT	NULL	NULL	NULL	NULL	NULL	NULL	2338	14709	15042	20084	NULL	NULL
2	中台灣客運	164913	176417	230401	221916	156505	114664	138699	151795	152018	156434	162557	179594
3	仁友客運	48171	50696	64569	73999	32350	27032	30959	32340	32781	30548	43072	48971
4	台中客運	563130	561916	672374	717802	465800	362812	389820	461088	484845	570704	572906	602913
5	巨業交通	94837	69390	70097	72164	127681	110566	118706	120040	98725	91429	97676	105595
6	全航客運	82882	92445	110757	115453	73867	56245	61873	72190	71498	76877	80927	81583
7	和欣客運	19489	14394	14364	15856	30779	29850	42852	35433	31856	20350	19640	20517
8	東南客運	56550	57492	71178	53619	41418	31437	32384	31671	31205	38417	54386	51604
9	南投客運	NULL	NULL	NULL	NULL	NULL	NULL	NULL	49	1093	944	NULL	NULL
10	苗栗客運	4042	3501	3510	3402	801	2717	3597	3491	3247	3247	3250	3920
11	統聯客運	373666	374701	487792	469321	312326	238076	269931	314470	319777	340934	362680	392907
12	彰化客運	22791	27470	36205	38929	15445	14833	15972	14711	14351	16075	16169	17890
13	豐原客運	240434	254036	314094	336415	203027	170304	177614	197851	220202	229849	231527	236337
14	豐榮客運	28855	31498	38197	38815	42432	34840	35780	39991	39319	40621	41326	38706

表 6.28 各時段定點資料傳輸延誤比較

	NAMEZH	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1	BRT	3435	3149	3120	2697	2763	2600	2642	2685	2846	2697	3257	3546	3113	2951	2830	2488	2369
2	中台灣客運	92101	113653	125395	121572	122954	131473	129047	134639	134131	147374	146230	131532	115091	100459	77967	68415	53482
3	仁友客運	26179	34757	37782	39776	33296	27821	21370	27348	28887	30987	35398	37356	32755	26447	23944	22757	21531
4	台中客運	323645	396187	408616	408609	404158	400742	365221	360412	375530	409854	429579	406494	399211	385976	319392	272165	222706
5	巨業交通	56273	71469	73158	67732	68884	68640	73822	67876	71562	67410	76265	73277	73051	75720	70492	60857	39091
6	全航客運	51716	61323	60417	56887	59992	58733	57901	55411	56381	56516	61952	63456	60883	57113	52318	47108	48907
7	和欣客運	7401	8310	15658	17020	19300	17906	18101	17206	17558	20590	20529	19595	18140	18755	17339	15952	14018
8	東南客運	32768	40222	41910	32718	33473	32644	30970	35110	31794	31166	39268	37511	36918	34120	23436	18142	16588
9	南投客運	66	143	196	194	180	175	91	139	190	2	127	140	191	224	28	NULL	NULL
10	苗栗客運	3439	2470	2042	3528	2628	3510	3781	1602	1054	3419	3512	3407	3452	880	NULL	NULL	NULL
11	統聯客運	203161	242631	254889	252035	254821	276545	268963	278567	283132	302659	310636	288025	254361	220275	178758	160896	130664
12	彰化客運	11362	20686	21868	16569	12876	14454	14462	14216	13694	12394	17935	18603	19653	16067	10911	7326	7505
13	豐原客運	189839	177678	160464	160227	166086	171385	177332	159248	151587	153641	187755	204470	196812	159234	128319	113186	100868
14	豐榮客運	29301	29855	25824	26496	28147	23048	24977	21435	22467	23910	33061	31229	30060	28705	24167	22609	19287

2. 各門檻值比例分佈比較

前項分析有兩個可改進方向，第一：表 6.25～表 6.26 係以固定的門檻值（3 秒）呈現資料，無法瞭解其他門檻值的分佈。第二：表 6.27～表 6.28 僅顯示延誤數量，並未考慮曝光量差異。圖 6-100 與圖 6-101 分別針對 14 億筆定時資料進行傳輸延遲、傳輸漏失之分析，不僅計算各門檻值之數值，也透過除以總母體數達到標準化目的，每一條線之所有樣本點加總均為 100%，由於 X 座標為負面指標（延遲秒數、漏失次數），故這些曲線越是左偏者代表車機穩定度越高，意即有越高比例的低延遲、低漏失表現。

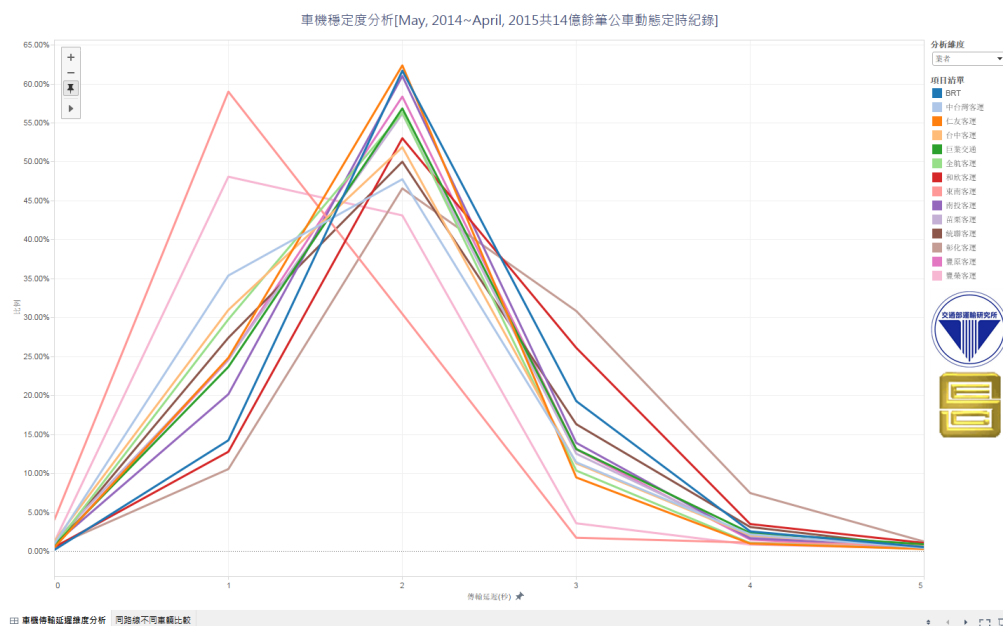


圖 6-100 各傳輸延遲樣本百分率比較圖

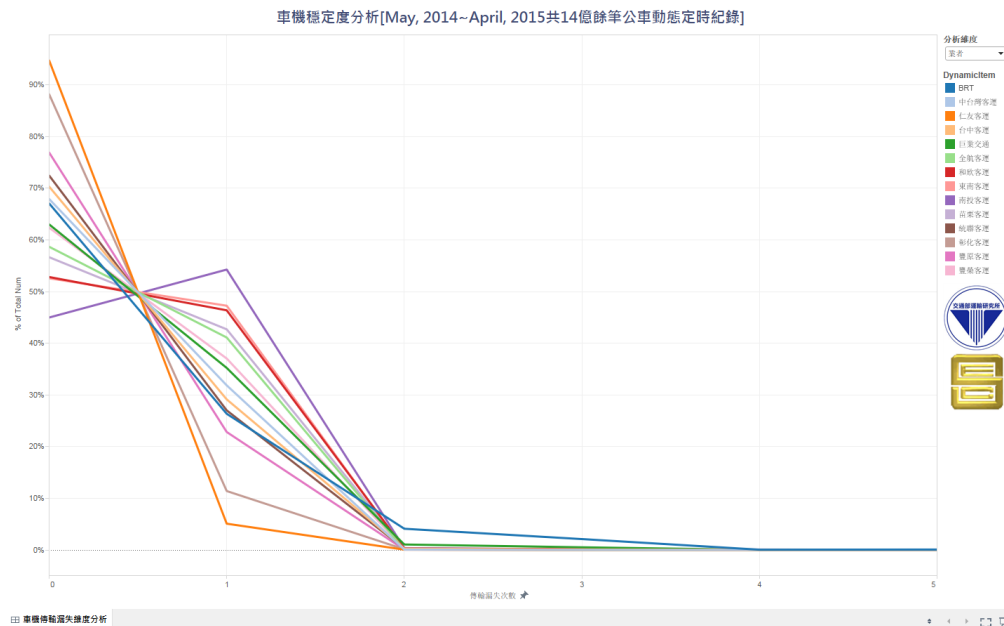


圖 6-101 各傳輸漏失樣本百分率比較圖

3. 各門檻值累積比例分佈比較

前述圖 6-100 與圖 6-101 雖可比較不同標的之車機穩定度，但若各門檻值起伏變化過大就很難判斷比較，故本研究近一步將圖 6-100 與圖 6-101 改為累積比例方式呈現如圖 6-102 與圖 6-103。

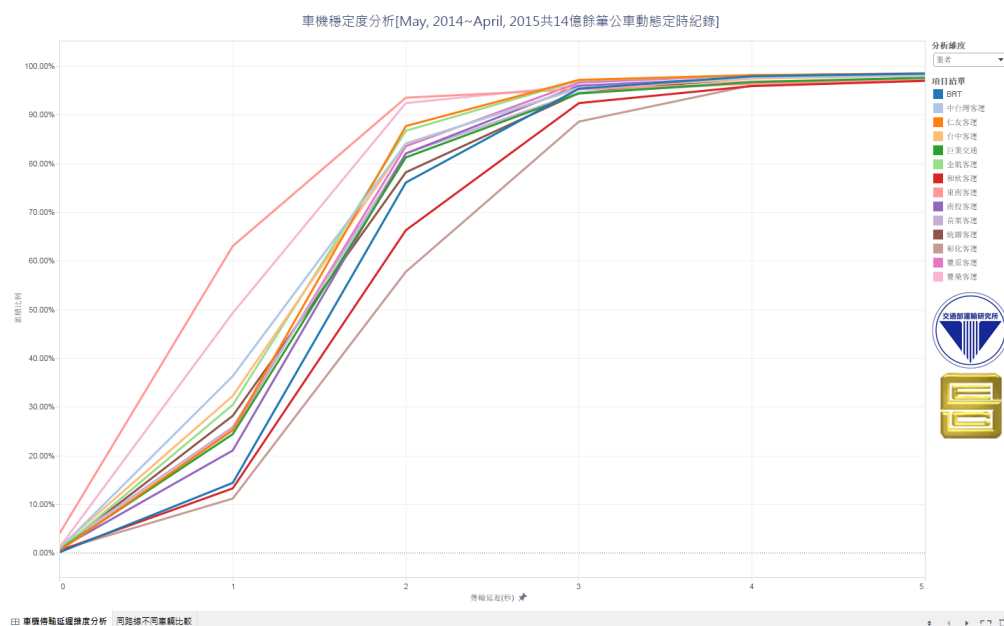


圖 6-102 各傳輸延遲樣本累積比率比較圖

經過累積比率的換算後，越是左上角之曲線代表穩定度越高，而越是右下角曲線就是穩定度較低的標的。圖 6-102 與圖 6-103 僅是以「業者」為例做展示，事實上在儀表板的右上角設計有一個 ComboBox 可切換 6 種不同分析維度來比較不同因素對車機穩定度之影響。

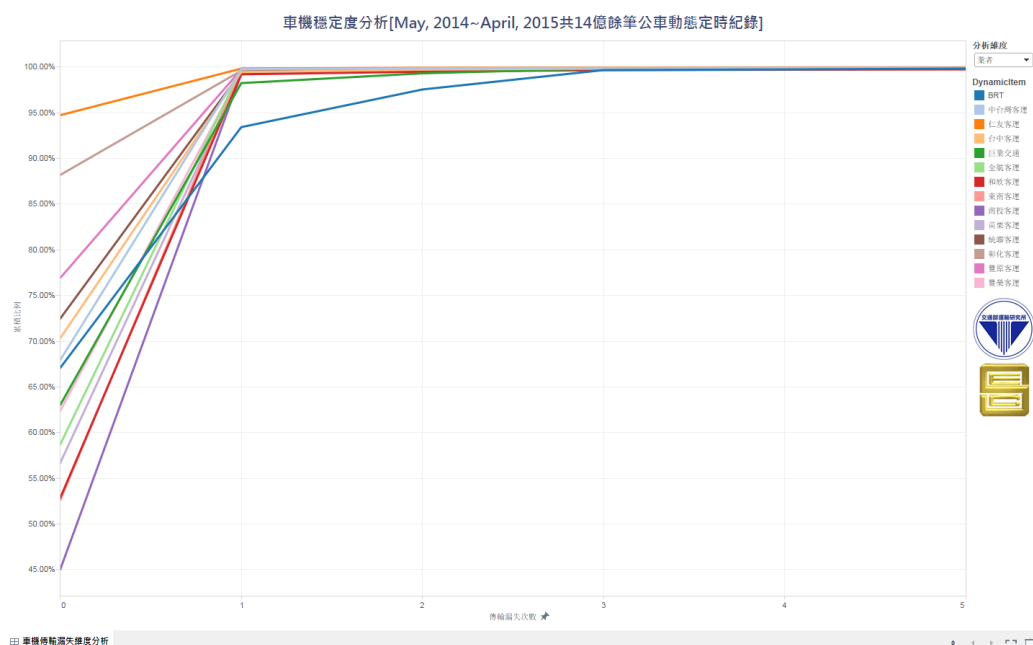


圖 6-103 各傳輸漏失樣本累積比率比較圖

6.9.3.3 影響因素分析

第 6.8.3.2 節展示之儀表板支援多種維度切換，配合「累積比例分佈」之設計，只要切換分析因素即可從線條的分散程度來瞭解該因素是否是影響車機穩定度之因素。

圖 6-104 係以「傳輸延遲」為例說明儀表板切換各種分析維度之截圖，從圖中可觀察出「月份」與「星期」明顯非影響因素，但「時段」與「業者」就有些微差異，至於路線與車輛則是明顯影響很大。從此視覺化呈現將可導引未來研究應朝向路線、車輛、業者等較大影響因素做更進一步之分析，至於月份/星期/時段等因素可先忽略。

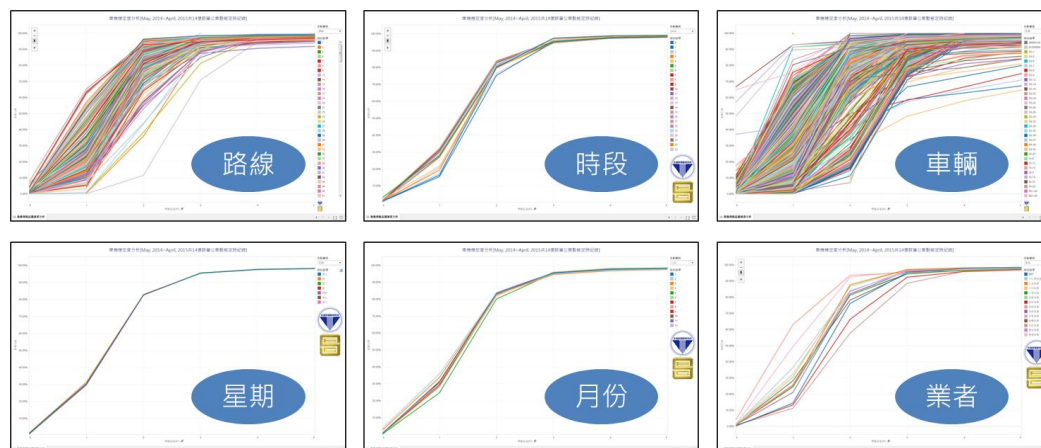


圖 6-104 傳輸延遲影響因素比較分析圖

由於「路線」和「車輛」兩因素很可能與車輛行駛於市中心或郊區有關，因此圖 6-105 之設計係讓使用者可根據路線篩選各車輛的績效曲線，以圖 6-105 為例，在挑選 55 路固定行駛範圍的情況下，各車輛的表現仍明顯有所差異，從中可推論出車輛（機）應是比路線（地區）的更為關鍵的影響因素。同時右下角曲線（例如橘色）可視為已故障或即將故障之車機，建議儘快報修瞭解原因並修復之。

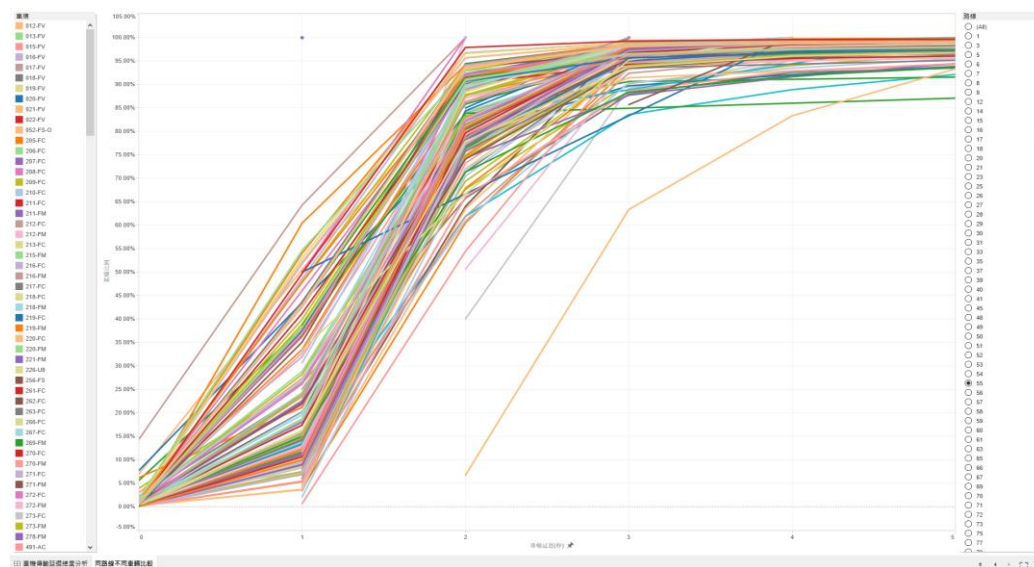


圖 6-105 指定路線比較各車機可靠度之儀表板

綜合以上分析，可歸納出車機本身才是最主要關鍵影響因素，為落實預防保養，建議未來可利用類似如圖 6-106 之儀表板追蹤檢核，左側為各業者之績效比較，點選左側任一業者曲線時，右側會即時顯示該業者所有車輛之績效，以圖 6-106 為例即可觀察出車機績效區分為三大區塊，越是右下側即是該追蹤檢修之對象，當滑鼠移至右側任一曲線時，均會即時顯示該車輛之車牌。此外，圖 6-106 係以全年度之資料呈現，未來應用時，可以考慮僅呈現最近一月/一週/一日來滾動檢討。

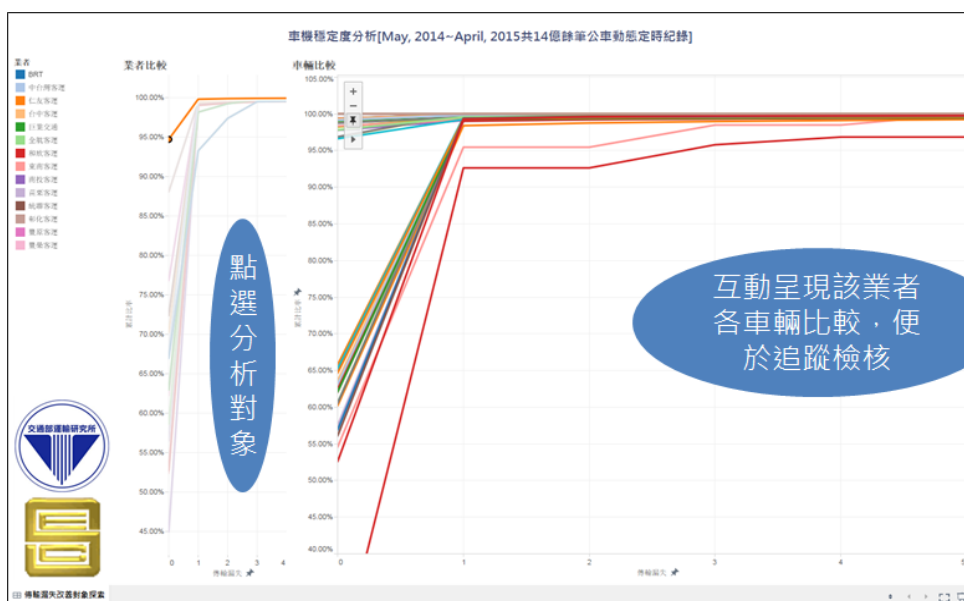


圖 6-106 預防檢修應用儀表板設計示意圖

第七章 結論與建議

本研究主要蒐集包含公車動態資訊系統在內的各項相關資料，建立大數據資料庫，並透過適合的資料探勘分析方法，針對所蒐集到的資料進行探索性的分析，探討如何可以協助業者進行有效的營運管理，更進一步提昇公共運輸的服務水準。對於政府單位而言，也可探討如何利用這些資料，分析目前公共運輸的供需情形，作為整體公共運輸規劃的依據。

根據本研究的結果，可以得到以下的結論與建議。

7.1 結論

1. 本研究透過各種管道蒐集大量的公車相關資料，包含 38 大項資料，超過 38 億筆資料，資料量超過 360GB，這些大量的資料，可以作為後續相關研究的基礎。
2. 對於這些所蒐集到的大量資料，本研究針對每一種資料的特性，分別開發不同的資料清洗方式，將不合理（錯誤）以及無法比對的資料予以過濾，以確保分析結果的合理性。在本研究對各項資料處理的方法，除了可以作為未來類似資料處理的依據外，更可以作為改善未來資料蒐集方法的基礎。
3. 本研究以所蒐到的資料為基礎，分別進行了電子票証分析、路線與班次重疊分析、供需分析、行車時間與班表分析、站牌班表產製分析、車輛相關分析（包含機務、油料）、駕駛員分析風險分析與車機穩定度分析等分析。根據分析的結果，可以得到一些對與營運管理相當有用的資料，無論從主管端或是業者端都具有相當的意義，可以作為提昇公共運輸服務水準的依據。

4. 承上，由於各縣市政府公車動態與電子票證均大同小異，故前述成果僅須微幅調整即可推廣應用至其他城市。
5. 本研究對於分析的結果，使用不同的工作設計不同的視覺化呈現界面，透過視覺化的呈現，可以讓使用者很快就可以掌握資料的特性，清楚且具體了解問題的所在，並據以制定相關的決策。本研究視覺化的呈現分別使用了 Tableau、R 語言、GIS 軟體，以及自行開發的 Java Script 系統。
6. 本研究根據業者與主管的單位的實際需求，分別進行許多不同面相的分析，在分析的過程中，多次與實務單位討論，充分了解業者實際面臨的課題，以求分析結果能夠滿足需求。
7. 由於本研究的資料來源繁多，分別具有不同的格式與特性，再加上所使用的分析與視覺化展項工具也有所不同，因此提出一套具有三個層級的系統架構，可以整合不同的資料特性與分析工具。本架構具有足夠的彈性與整合能力，可作為未來我國全面發展公共運輸巨量資料分析的基礎。
8. 本研究分別舉辦了 2 場的需求座談會與 1 場專家座談會，廣邀主管單位與客運業者與會，對於巨量資料分析進行了深入的介紹與展示本研究的成果。與會人士對公共運輸的巨量資料分析均表達了高度的肯定與興趣，也贊同巨量分析對業務有絕對的助益。

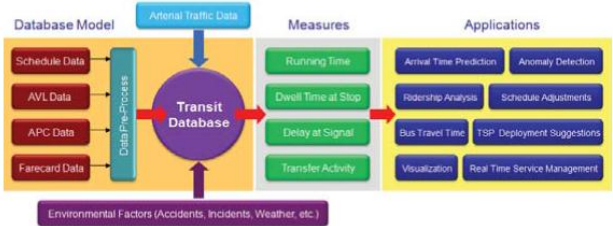
7.2 建議

1. 電子票証是公共運輸巨量資料分析的核心，有完整的電子票証資料可以讓分析的成果更加豐富。但目前由於各種限制，電子票証的取得並不容易，再加上目前電子票証的格式設計主要為蒐費之用，因此仍存在站序（站名）等諸多課題。建議未來除了可以透過各種協商管道定時取得電子票証資料外，更可以積極規範所提

供的資料格式，以減少資料取得後的分析工作，也可以顯著增加可以分析的資料量。

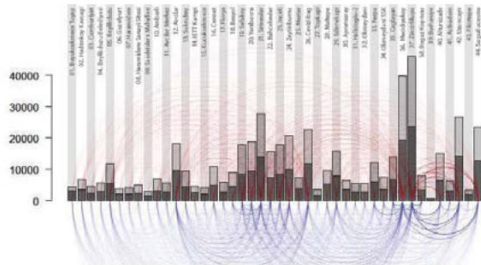
2. 電子票証與公車動態資料為目前客運巨量資料研究最主要的兩項資料來源，但這兩項資料，無論是在內部或是這兩項資料之間，均存在著不一致性，尤其是兩項資料在車牌、站牌的一致性對應部分，往往造成分析費時，也減少了可分析的資料數量，建議未來應該針對車牌、站牌的編碼與異動更新應該有一套系統方式，同時增加轉乘勾稽資訊，若以上改善均完成，將大幅提昇原始資料品質，同時亦可讓後續研究針對轉乘行為做深入分析。
3. 本研究受限於時間，在服務水準的分析上以國內過去未見的「車上乘載率」為主，建議未來可以加上多目標的決策參考指標。
4. 本研究受限於時間，僅能夠進行探索性的分析，因此使用的工具各異，也缺少一套整合性的系統介面。建議未來可以開發整合式分析介面（包括可匯入資料、可加入其他業者資料...等），以具有親和力的方式提供給各單位使用。
5. 由於公共運輸的巨量資料分析乃是各級政府之所需，也為各民間業者之所需，建議可以在中央成立相關的發展機制或是發展中心，整合各級地方政府（尤其是六個直轄市）的資料，發展共同平臺，並分享分析的成果。
6. 建議未來公共運輸資料部分可由交通部資訊中心所建置之公共運輸資料流通平臺統一彙整蒐集，以免各單位或是各計畫重複建置類似的資料庫，造成資源的浪費與資料的不一致。

附錄 A 需求座談會議資料

 交通部運輸研究所 Institute of Transportation, MOTC 公車動態資訊系統巨量資料(Big Data) 蒐集與視覺化分析研究 主管機關需求座談 會議資料 國立交通大學運輸研究中心 中華民國104年05月08日	會議主題 - 公共運輸主管機關對巨量資料分析之需求 - 車輛軌跡紀錄 - 電子票證紀錄 - 其他資料紀錄  1 國立交通大學運輸研究中心
借鏡國外案例 - 城市：韓國首爾 - 目的：大眾運輸路線設計（深夜公車）  2 國立交通大學運輸研究中心	借鏡國外案例 城市：美國明尼蘇達州-明尼亞波利斯  3 國立交通大學運輸研究中心
借鏡國外案例 - 城市：江蘇省鎮江市 - 目的：動態調度公車  4 國立交通大學運輸研究中心	借鏡國外案例 - 城市：荷蘭 - 目的：顯示延誤程度與分佈  5 國立交通大學運輸研究中心

借鏡國外案例

- 城市：土耳其伊斯坦堡
- 目的：顯示各站進出量/起訖改善服務品質



6

國立交通大學運輸研究中心

腦力激盪原則

避免批判

不批判他人發言，才能讓所有成員自由發想。

自由發想

使成員無拘束地暢所欲言，在輕鬆的氣氛下，愉悅流暢地思考。

量重於質

用撒網捕魚的方式，先進行圍繞主題的發想，以徵求大量點子。

結合改善

以彼此的發想為基礎，不斷激盪出更好的新點子，甚至允許團隊中出現與他人相似的發想。

7

國立交通大學運輸研究中心

主管機關需求目的列舉



8

國立交通大學運輸研究中心

客運監理需求

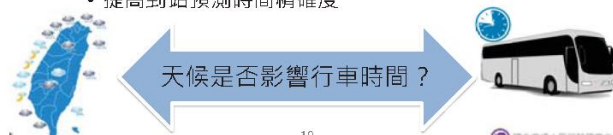
- 服務水準監理
 - 最少班次數/最大班距
 - 平均速度/班距分佈
 - 延誤/脫班/早開
 - 舒適度（乘載率）
 - 路線調整
- 行車安全監理
 - 違規駕駛(超速/急駛)
 - 異常駕駛(急加/減速)

9

國立交通大學運輸研究中心

民眾服務需求

- 提昇「公車動態資訊」品質
 - 增加資訊項目
 - 車輛即時乘載率
 - 其他項目？
 - 提高資訊正確
 - 提高APTS設備(車機)可靠度
 - 提高到站預測時間精確度



10

國立交通大學運輸研究中心

簡報結束
敬請指教



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

公車動態資訊系統巨量資料(Big Data)
蒐集與視覺化分析研究

客運業者需求座談
會議資料

國立交通大學運輸研究中心
中華民國104年05月08日

會議主題

- 公路客運業者對巨量資料分析之需求？
 - 車輛軌跡紀錄
 - 電子票證紀錄
 - 其他資料紀錄

1

借鏡國外案例

- 城市：韓國首爾
- 目的：大眾運輸路線設計（深夜公車）

2

借鏡國外案例

- 城市：美國明尼蘇達州-明尼亞波利斯

Database Model

Arterial Traffic Data

Measures

Applications

Transit Database

Environmental Factors (Accidents, Incidents, Weather, etc.)

3

借鏡國外案例

- 城市：江蘇省鎮江市
- 目的：動態調度公車

4

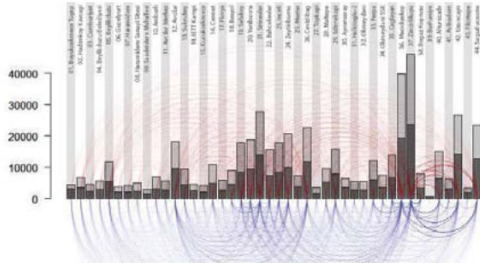
借鏡國外案例

- 城市：荷蘭
- 目的：顯示延誤程度與分佈

5

借鏡國外案例

- 城市：土耳其伊斯坦堡
- 目的：顯示各站進出量/起訖改善服務品質



6

國立交通大學運輸研究中心

腦力激盪原則

避免批判

不批判他人發言，才能讓所有成員自由發想。

自由發想

使成員無拘束地暢所欲言，在輕鬆的氣氛下，愉悅流暢地思考。

量重於質

用撒網捕魚的方式，先進行圍繞主題的發想，以徵求大量點子。

結合改善

以彼此的發想為基礎，不斷激盪出更好的新點子，甚至允許團隊中出現與他人相似的發想。

7

國立交通大學運輸研究中心

客運業者需求目的列舉

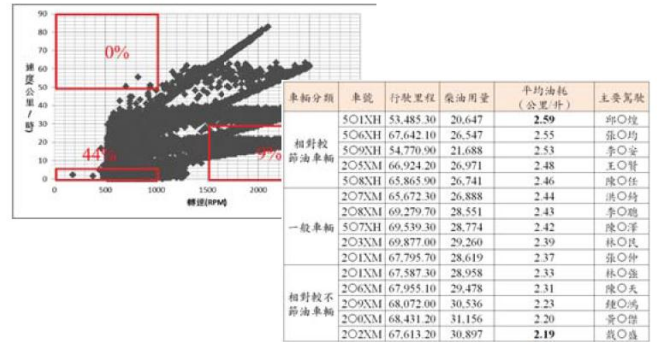


8

國立交通大學運輸研究中心

客運業者需求-降低成本

- 行車記錄器與油料資料分析



9

國立交通大學運輸研究中心

客運業者需求-民眾服務

- 製作理想之站班表

表一：台中一日月曆		表二：台中一日月曆	
站名	時間	站名	時間
臺中火車站	07:50	臺中火車站	07:50
臺中火車站	08:10	臺中火車站	08:10
臺中火車站	08:30	臺中火車站	08:30
臺中火車站	08:50	臺中火車站	08:50
臺中火車站	09:10	臺中火車站	09:10
臺中火車站	09:30	臺中火車站	09:30
臺中火車站	09:50	臺中火車站	09:50
臺中火車站	10:10	臺中火車站	10:10
臺中火車站	10:30	臺中火車站	10:30
臺中火車站	10:50	臺中火車站	10:50
臺中火車站	11:10	臺中火車站	11:10
臺中火車站	11:30	臺中火車站	11:30
臺中火車站	11:50	臺中火車站	11:50
臺中火車站	12:10	臺中火車站	12:10
臺中火車站	12:30	臺中火車站	12:30
臺中火車站	12:50	臺中火車站	12:50
臺中火車站	13:10	臺中火車站	13:10
臺中火車站	13:30	臺中火車站	13:30
臺中火車站	13:50	臺中火車站	13:50
臺中火車站	14:10	臺中火車站	14:10
臺中火車站	14:30	臺中火車站	14:30
臺中火車站	14:50	臺中火車站	14:50
臺中火車站	15:10	臺中火車站	15:10
臺中火車站	15:30	臺中火車站	15:30
臺中火車站	15:50	臺中火車站	15:50
臺中火車站	16:10	臺中火車站	16:10
臺中火車站	16:30	臺中火車站	16:30
臺中火車站	16:50	臺中火車站	16:50
臺中火車站	17:10	臺中火車站	17:10
臺中火車站	17:30	臺中火車站	17:30
臺中火車站	17:50	臺中火車站	17:50
臺中火車站	18:10	臺中火車站	18:10
臺中火車站	18:30	臺中火車站	18:30
臺中火車站	18:50	臺中火車站	18:50
臺中火車站	19:10	臺中火車站	19:10
臺中火車站	19:30	臺中火車站	19:30
臺中火車站	19:50	臺中火車站	19:50
臺中火車站	20:10	臺中火車站	20:10
臺中火車站	20:30	臺中火車站	20:30
臺中火車站	20:50	臺中火車站	20:50
臺中火車站	21:10	臺中火車站	21:10
臺中火車站	21:30	臺中火車站	21:30
臺中火車站	21:50	臺中火車站	21:50
臺中火車站	22:10	臺中火車站	22:10
臺中火車站	22:30	臺中火車站	22:30
臺中火車站	22:50	臺中火車站	22:50
臺中火車站	23:10	臺中火車站	23:10
臺中火車站	23:30	臺中火車站	23:30
臺中火車站	23:50	臺中火車站	23:50
臺中火車站	00:10	臺中火車站	00:10
臺中火車站	00:30	臺中火車站	00:30
臺中火車站	00:50	臺中火車站	00:50
臺中火車站	01:10	臺中火車站	01:10
臺中火車站	01:30	臺中火車站	01:30
臺中火車站	01:50	臺中火車站	01:50
臺中火車站	02:10	臺中火車站	02:10
臺中火車站	02:30	臺中火車站	02:30
臺中火車站	02:50	臺中火車站	02:50
臺中火車站	03:10	臺中火車站	03:10
臺中火車站	03:30	臺中火車站	03:30
臺中火車站	03:50	臺中火車站	03:50
臺中火車站	04:10	臺中火車站	04:10
臺中火車站	04:30	臺中火車站	04:30
臺中火車站	04:50	臺中火車站	04:50
臺中火車站	05:10	臺中火車站	05:10
臺中火車站	05:30	臺中火車站	05:30
臺中火車站	05:50	臺中火車站	05:50
臺中火車站	06:10	臺中火車站	06:10
臺中火車站	06:30	臺中火車站	06:30
臺中火車站	06:50	臺中火車站	06:50
臺中火車站	07:10	臺中火車站	07:10
臺中火車站	07:30	臺中火車站	07:30
臺中火車站	07:50	臺中火車站	07:50

10

國立交通大學運輸研究中心

客運業者需求-提高可靠度

- 異常車輛預警



11

國立交通大學運輸研究中心

客運業者需求-提高可靠度



不同剩餘電量



每單位電量之充電時間

每單位電量之行駛時間



妥善率

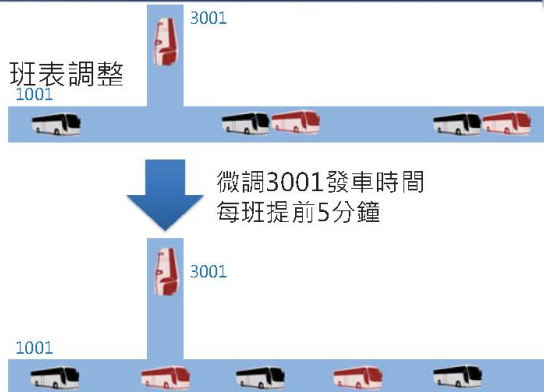
使用率

12

國立交通大學運輸研究中心

客運業者需求-提昇效率

- 班表調整



13

國立交通大學運輸研究中心

簡報結束
敬請指教

公車動態資訊系統巨量資料 (Big Data) 蒐集與視覺化分析研究

附錄 B 教育訓練辦理情形

本專案計畫教育訓練共計分兩部分進行，第一部分重點在於 R 語言之程式實作訓練。第二部分的教育訓練則為 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。分別在 2015 年 5/29 與 7/1 辦理完成。以下彙整教育訓練辦理情形。

R 語言訓練辦理情形

1. 訓練時間：104 年 05 月 29 日(星期五) 09:00~16:30
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：上午場課程如表 B.1，下午場課程如表 B.2。

表 B.1 上午場 R 語言訓練課程表

課程單元	課程主題
R 軟體基礎	● 試算表資料匯入與網頁資料擷取案例-以 瀕臨危機之世界遺產為例
資料視覺化	● 多變量條件式繪圖(lattice)案例-以校友捐款資料為例 ● 圖形文法繪圖(ggplot2)案例-以保險客戶 資料探索式分析為例 ● 熱繪圖(heat map)應用-相關係數矩陣之 視覺化

表 B.2 下午場 R 語言訓練課程表

課程單元	課程主題
資料探勘	● k-means 集群分析案例-社群網站青少年學習用戶市場區隔 ● 關聯規則探勘案例-線上音樂城購物籃分析
機器學習	● 樹狀模型之分類樹案例-銀行授信風險管理 ● 貝式分類(文字資料處理)案例-垃圾簡訊/郵件過濾

5. 訓練對象：簽到如表 B.3。

表 B.3 R 語言訓練課程簽到表



公車動態資訊系統巨量資料蒐集與視覺化分析研究



交通部運研所 R 軟體課程簽到暨講義領取表

一、課程時間：2015/05/29 AM 9:00~ PM 16:30

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
中華大學	學生	游耀翔	交通大學		王聖元
中華大學	學生	紀詩詩	中興	研究員	張恩輔
"	組長	李慶霖	中興	研究員	黃學珍
中華大學	助理	林立凡	運研所	副組長	吳東谷
中華大學	助理	許榮高	中興	研究員	吳明軒
"	"	黃瑞	中興	研究員	孫千山
統計處	專員	崔培均	運研所	副研習	郭松有
"	編審	周淑芬	"	助理研習	王怡婷
統計處	專員	鍾靜涵	"	研究員	史習華
中華大學	學士	孫煥博	"	助理研習	陳志岳
統計處	專員	葉登菁	"		張淑文
交通大學	學生	王佳淇	"		朱桐雲
運研所	助理研習	何新捷	黑湖科技	工程師	鄧瑞
交通大學	學生	蔡仁倫	"	"	張家駒
			台灣大學	學生	陳桂豪

6. 活動剪影：如圖 B-1。



圖 B-1 R 語言訓練課程實況剪影

商業智慧軟體訓練辦理情形

1. 訓練時間：104 年 07 月 01 日(星期三) 09:00～16:30
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：
 - (1) 昊青公司工程師：黃曉菁、賴彥叡
 - (2) 中興工程顧問社研究員：張恩輔
4. 訓練內容：上午場課程如表 B.4，下午場課程如表 B.5。

表 B.4 上午場 Tableau 軟體訓練課程表

時間	課程內容
1 hour	Tableau Desktop 安裝與基本操作
10 min	Break
1 hour	Tableau Desktop 基礎案例練習
10 min	Break
1 hour	Tableau Desktop 基礎案例練習

表 B.5 下午場 Tableau 軟體訓練課程表

時間	課程內容
1 hour	Tableau Desktop 進階功能練習
10 min	Break
1.5 hour	Tableau Desktop 進階功能練習
10 min	Break
0.5 hour	Tableau 展示(MOTC-IOT-104-IEB017公車動態資訊系統巨量資料 (big data) 蒐集與視覺化分析案初步結果展示)

5. 訓練對象：簽到如表 B.6。

表 B.6 Tableau 軟體訓練課程簽到表

「Tableau Desktop 教育訓練」報名表		
時間:104年7月1日(三)上午9:00-12:00下午1:30-4:30(共6小時)		
地點:本所五樓會議室 授課人:Tableau 代理商		
單位:交通部管理資訊中心、運研所運資組		
姓 名	身分證字號	簽 到
李霞		
王國琛		
黃月貞		黃月貞
陳其華		
吳東凌		
洪銘揚		
劉仲潔		
周家慶		周家慶
邱佩諱		
何毓芬		
王詠萱		
陳致仲		
林秋錦		林秋錦
呂恩慈		
陳翔捷		陳翔捷

「Tableau Desktop 教育訓練」報名表		
時間:104年7月1日(三)上午9:00-12:00下午1:30-4:30(共6小時)		
地點:本所五樓會議室 授課人:Tableau 代理商		
單位:交通大學團隊(交通大學、中華大學、台北商大、中興社)		
姓 名	身分證字號	簽 到
李慶霖		李慶霖
黃瑋		黃瑋
許葉彥		許葉彥
謝勳遠		謝勳遠
沈美慧		沈美慧
張志鴻		張志鴻
林靜芬		林靜芬
黃笙琰		黃笙琰
吳明軒		吳明軒
張思輔		張思輔
陳桂豪		陳桂豪
王晉元		
王佳琪		王佳琪
蔡佳倫		蔡佳倫
鄒慶士		鄒慶士
莊舒媛		莊舒媛
曹祐嘉		曹祐嘉
周淑樺		周淑樺

「TABLEAU DESKTOP 教育訓練」		
報名表		
時間:104年7月1日上午9:00-12:00 下午1:30-4:30(共6小時)		
地點:本所五樓會議室		
授課人:Tableau 代理商		
單位:運工組、運計組、運安組		
姓 名	身分證字號	簽 到
翟慰宗		翟慰宗
林通耀		
鄒德傳		鄒德傳
王怡婷		王怡婷
許修豪		
張明子		張明子
符雲梅		符雲梅
黃曉菁(吳育)		黃曉菁
賴彥敏(吳育)		賴彥敏

「TABLEAU DESKTOP 教育訓練」		
報名表		
時間:104年7月1日上午9:00-12:00 下午1:30-4:30(共6小時)		
地點:本所五樓會議室		
授課人:Tableau 代理商		
單位:桃園市政府		
姓 名	身分證字號	簽 到
劉廣堂		劉廣堂
洪敏琛		洪敏琛
林彥鋒		林彥鋒
陳錦星		陳錦星
陳盈潔		陳盈潔
張新福		張新福

6. 活動剪影：如圖 B-2。



圖 B-2 Tableau 軟體訓練課程實況剪影

附錄 C 資料庫欄位說明

資料表名稱	A1		
說明	車機定時回傳資料		
編號	中文說明	欄位名稱	型態
1	業者別	Cmp	char(2)
2	車號	BusID	char(10)
3	勤務 ID	DutyStatus	tinyint
4	車輛狀態	BussStatus	tinyint
5	路線	Route	char(10)
6	往返程	Goback	tinyint
7	緯度	X	decimal(10, 7)
8	經度	Y	decimal(9, 7)
9	速度	Speed	tinyint
10	GPS 時間	GPSTime	datetime
11	類型	Type	tinyint
12	傳輸時間	TrasTime	datetime
13	接收時間	RecTime	datetime

資料表名稱	A2		
說明	車機定點回傳資料		
編號	中文說明	欄位名稱	型態
1	業者別	Cmp	char(2)
2	車號	BusID	char(10)
3	勤務編號	DutyStatus	tinyint
4	車輛狀態	BussStatus	tinyint
5	路線	Route	char(10)
6	往返程	Goback	tinyint
7	站號	Stop	int
8	到離站	Leave	bit
9	GPS 時間	GPSTime	datetime
10	類型	Type	tinyint
11	傳輸時間	TrasTime	datetime
12	接收時間	RecTime	datetime

資料表名稱	accident		
說明	事故資料		
編號	中文說明	欄位名稱	型態
1	編號	ID	int
2	年	Year	int
3	月	month	int
4	日	date	int
5	時	hour	int
6	分	minute	int
7	發生時間	Atime	datetime
8	縣市	County	varchar(10)
9	鄉鎮	Town	varchar(10)
10	道路	Road	varchar(10)
11	路段	Sec	varchar(10)
12	車道	Lane	varchar(10)
13	弄	Alley	varchar(10)
14	號	NO	varchar(10)
15	公尺	Meter	int
16	道路	Road1	varchar(10)
17	路段	Sec1	varchar(10)
18	方向	Side	varchar(10)
19	備註	Near	varchar(5000)
20	死亡人數	Die	int
21	受傷人數	Injury	int

資料表名稱	bus_company		
說明	業者對照表		
編號	中文說明	欄位名稱	型態
1	編號	bc_id	nvarchar(5)
2	名稱	bc_name	nvarchar(30)
3	備註	bc_memo	nvarchar(30)

資料表名稱	bus_history_route		
說明	公路客運歷史行車資料		
編號	中文說明	欄位名稱	型態
1	公司別	bhr_bc_id	nvarchar(5)
2	路線別	bhr_driving_route	nvarchar(10)
3	路線名稱	bhr_route_name	nvarchar(100)
4	支線	bhr_driving_path	nvarchar(10)

5	時間	bhr_record_time	datetime
6	經度	bhr_longitude	float
7	緯度	bhr_latitude	float
8	車速	bhr_speed_up	decimal(3, 0)
9	平均車速	bhr_speed_avg	decimal(3, 0)
10	狀態	bhr_status	nvarchar(20)

資料表名稱	Easycard_Provider_Route		
說明	悠遊卡業者路線資料		
編號	中文說明	欄位名稱	型態
1	業者編號	ProviderID	varchar(10)
2	業者名稱	ProviderName	nvarchar(10)
3	路線編號	RouteID	varchar(10)
4	路線名稱	RouteName	nvarchar(50)

資料表名稱	EasyCard-Taichung		
說明	悠遊卡電子票證資料		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int
2	票證公司別	cmp	varchar(8)
3	業者編號	provider	varchar(6)
4	卡片編號	card	varchar(12)
5	卡片類型	cardno	varchar(1)
6	卡片類型編號	cardid	varchar(1)
7	路線編號	routeid	varchar(5)
8	駕駛員編號	driverid	varchar(5)
9	車號	busid	varchar(8)
10	上車時間	ontime	datetime
11	上車站序	onseq	int
12	上車站名	onname	varchar(20)
13	計價站起站	onseqp	int
14	計價站名稱	onnamep	varchar(20)
15	下車時間	offtime	datetime
16	下車站序	offseq	int
17	下車站名	offname	varchar(20)
18	計價站起站	offseqp	int
19	計價站名稱	offnamep	varchar(20)

資料表名稱	Easycard_Ticket		
說明	票卡類型說明		
編號	中文說明	欄位名稱	型態
1	票卡類型	TicketID	int
2	說明	TicketType	nvarchar(10)

資料表名稱	ETSTOP		
說明	台中公車站名對照表		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int
2	路線編號	RouteID	int
3	站別	Type1	nvarchar(254)
4	站序	Type2	int
5	站牌編號	StopID	nvarchar(254)
6	站名	StopName	nvarchar(254)
7	往返程	Goback	int
8	經度	Longitude	numeric(38, 8)
9	緯度	Latitude	numeric(38, 8)

資料表名稱	fy_driver_roster		
說明	勤務資料表		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int
2	勤務日期	FDR_date	nvarchar(20)
3	往返程	FDR_goback	tinyint
4	名稱	FDR_name	nvarchar(20)
5	星期	FDR_week	tinyint
6	勤務序號	FDR_seqno	tinyint
7	路線名稱	FDR_route	nvarchar(20)
8	起站名稱	FDR_departure_stop	nvarchar(20)
9	迄站名稱	FDR_departure_time	nvarchar(20)
10	經過站 1	FDR_via_stop_2	nvarchar(20)
11	經過站時間 1	FDR_via_time_2	nvarchar(20)
12	經過站 2	FDR_via_stop_3	nvarchar(20)
13	經過站時間 2	FDR_via_time_3	nvarchar(20)
14	經過站 3	FDR_via_stop_4	nvarchar(50)
15	經過站時間 3	FDR_via_time_4	nvarchar(20)
16	經過站 4	FDR_via_stop_5	nvarchar(50)
17	經過站時間	FDR_via_time_5	nvarchar(50)

18	結速	FDR_destination	nvarchar(20)
19	總里程	FDR_route_mileage	decimal(6, 1)
20	總時間	FDR_route_time	int

資料表名稱	mcardf		
說明	電子票證		
編號	中文說明	欄位名稱	型態
1	日期	date	nvarchar(7)
2	駕駛員編號	driver_id	nvarchar(6)
3	路線編號	route_id	nvarchar(4)
4	車號	car_id	nvarchar(6)
5	勤務編號	d_no	nvarchar(5)
6	電子票證公司別	t_cmp	nvarchar(4)
7	票種	kind	nvarchar(20)
8	票種名稱	t_name	nvarchar(20)
9	卡號	card	nvarchar(10)
10	上車時間	uptime	time
11	上車代號	upno	decimal(5, 0)
12	上車站名	upname	nvarchar(26)
13	計費站名	upbeg	nvarchar(26)
14	下車時間	dmtime	time
15	下車代號	dmno	decimal(5, 0)
16	下車站名	dmname	nvarchar(26)
17	計費站名	dmend	nvarchar(26)
18	上車站序	r_beg	nvarchar(2)
19	下車站序	r_end	nvarchar(2)
20	總里程	r_mil	decimal(6, 1)

資料表名稱	N1		
說明	歷史預估到站時間		
編號	中文說明	欄位名稱	型態
1	站牌編號	STOPID	int
2	路線名稱	ROUTEID	varchar(50)
3	車號	BUSID	varchar(8)
4	即時站名	CURRENTSTOP	int
5	終點站名	DESTINATIONSTOP	int
6	末班車	ISLASTBUS	varchar(2)
7	預估時間	ESTIMATE TIME	int
8	預估距離	STOPDISTANCE	varchar(2)

9	往返程	DIRECTION	smallint
10	類型	TYPE	smallint
11	傳輸時間	TRASTIME	datetime
12	接收時間	RECTIME	datetime

資料表名稱	OBU		
說明	行車記錄器		
編號	中文說明	欄位名稱	型態
1	車號	BUS	varchar(8)
2	GPS 時間	GPSTime	datetime
3	經度	Longitude	float
4	緯度	Latitude	float
5	速度	Speed	float
6	轉速	RPM	float
7	前門	FrontDoor	smallint
8	後門	RearDoor	smallint
9	電門	Acc	smallint

資料表名稱	TaichungPROVIDER		
說明	台中市區公車業者資料		
編號	中文說明	欄位名稱	型態
1	編號	ID	nvarchar(5)
2	名稱	NAMEZH	nvarchar(11)
3	類型	TYPE	tinyint

資料表名稱	TaichungROUTE		
說明	台中市區公車業者資料		
編號	中文說明	欄位名稱	型態
1	路線編號	ID	nvarchar(5)
2	業者名稱	PROVIDERID	int
3	路線名稱	NAMEZH	nvarchar(5)
4	備註	DDESC	nvarchar(23)
5	起點	DEPARTUREZH	nvarchar(9)
6	迄點	DESTINATIONZH	nvarchar(11)

資料表名稱	TaichungSCHEDULE		
說明	台中市區公車班表資料		
編號	中文說明	欄位名稱	型態

1	路線編號	Route	nvarchar(50)
2	發車時間	departureTime	datetime
3	往返程	Goback	tinyint
4	班表日期	Schedule_date	datetime

資料表名稱	TaichungSTOP		
說明	台中市區公車站牌資料		
編號	中文說明	欄位名稱	型態
1	站牌編號	ID	nchar(4)
2	路線編號	ROUTE	nvarchar(5)
3	站牌名稱	NAMEZH	nvarchar(15)
4	站序	SEQNO	int
5	上下車	PGP	tinyint
6	場站	TERMINAL	tinyint
7	地區編號	DISTRICTID	tinyint
8	往返程	GOBACK	tinyint
9	緯度	LATITUDE	decimal(10, 8)
10	經度	LONGITUDE	decimal(10, 7)

資料表名稱	thbProvider		
說明	公路客運業者資料		
編號	中文說明	欄位名稱	型態
1	業者編號	Provider	int
2	流水號	Id	int
3	英文名稱	nameEn	nvarchar(30)
4	中文名稱	nameZh	nvarchar(9)
5	類型	type	tinyint

資料表名稱	thbRoute		
說明	公路客運路線資料		
編號	中文說明	欄位名稱	型態
1	路線編號	Route	nvarchar(6)
2	流水號	Id	nvarchar(6)
3	業者編號	providerId	int
4	英文名稱	nameEn	nvarchar(300)
5	中文名稱	nameZh	nvarchar(70)
6	起點英文名稱	departureEn	nvarchar(54)
7	起點中文名稱	departureZh	nvarchar(13)

8	迄點英文名稱	destinationEn	nvarchar(54)
9	迄點中文名稱	destinationZh	nvarchar(13)
10	駛車憑證開始日期	opBD	nvarchar(22)
11	駛車憑證結束速日期	opED	nvarchar(22)
12	狀態	status	tinyint
13	類型	type	int
14	站數	stopNums	int

資料表名稱	thbSchedule		
說明	公路客運班表名稱		
編號	中文說明	欄位名稱	型態
1	業者編號	Provider	int
2	路編號	Route	nvarchar(6)
3	站牌編號	Stop	int
4	班表編號	Schedule	int
5	班次代號	serviceId	int
6	發車時間	aTimeReg	char(150)
7	發車間距	dTimeReg	char(150)
8	尖峰發車時間	aTimeRush	char(150)
9	尖峰發車間距	dTimeRush	char(150)
10	星期一	mon	nvarchar(5)
11	星期二	tue	nvarchar(5)
12	星期三	wed	nvarchar(5)
13	星期四	thu	nvarchar(5)
14	星期五	fri	nvarchar(5)
15	星期六	sat	nvarchar(5)
16	星期日	sun	nvarchar(5)
17	備註	special	int
18	特別日	dates	nvarchar(150)
19	狀態	status	int

資料表名稱	THBStop		
說明	公路客運站牌資料		
編號	中文說明	欄位名稱	型態
1	業者編號	Provider	int
2	路線編號	Route	nvarchar(6)
3	站牌編號	Stop	int
4	站牌英文名稱	nameEn	nvarchar(98)
5	站牌中文名稱	nameZh	nvarchar(18)

6	站序	seqNo	int
7	上下車	Pgp	int
8	場站	Terminal	tinyint
9	地區編號	districtId	tinyint
10	緯度	Latitude	decimal(10, 8)
11	經度	Longitude	decimal(10, 7)
12	地址	Address	nvarchar(64)

資料表名稱	THBTicket		
說明	公路客運票價資料		
編號	中文說明	欄位名稱	型態
1	業者編號	Provider	int
2	路線編號	Route	nvarchar(6)
3	票價編號	Tickets	int
4	起站站序	departureSeqNo	nvarchar(20)
5	迄站站序	destinationSeqNo	nvarchar(20)
6	票價	price	decimal(10, 2)
7	幣別	currency	nvarchar(3)
8	計程收費	period	nvarchar(800)

資料表名稱	TOrder		
說明	計程車訂單資料		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int
2	上車時間	UPDATE	datetime
3	上車緯度	Lat1	varchar(50)
4	上車經度	Lon1	varchar(50)
5	下車時間	DMDATE	datetime
6	下車緯度	Lat2	varchar(50)
7	下車經度	Lon2	varchar(50)

資料表名稱	weather_area		
說明	天氣地區資料		
編號	中文說明	欄位名稱	型態
1	地區編號	wa_id	nvarchar(3)
2	地區名稱	wa_name	nvarchar(20)
3	備註	wa_memo	nvarchar(20)

資料表名稱	weather_city		
說明	天氣地區資料		
編號	中文說明	欄位名稱	型態
1	縣市代號	wc_id	nvarchar(1)
2	備註	wc_name	nvarchar(10)

資料表名稱	weather_data		
說明	天氣資料		
編號	中文說明	欄位名稱	型態
1	流水號	wd_index	decimal(18, 0)
2	縣市代號	wd_wa_id	nvarchar(3)
3	日期	wd_date	date
4	小時	wd_hour	float
5	PA 值 1	wd_hpa_1	float
6	PA 值 2	wd_hpa_2	float
7	溫度 1	wd_temperature_1	float
8	溫度 2	wd_temperature_2	float
9	濕度	wd_humidity	float
10	風速	wd_wind_ms	float
11	風向	wd_wind_degree	float
12	最大陣風	wd_gust_ms	float
13	最大陣風風向	wd_gust_degree	float
14	雨量	wd_rain_mm	float
15	小時	wd_rain_hr	float
16	日照時速	wd_sunshine_hr	float
17	全天空日射量	wd_sunshine_mj	float
18	能見度	wd_visibility	float

資料表名稱	DAILY_REPORT		
說明	每日營運報表		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	日期	DATE	date
3	路線	ROUTE	nvarchar(4)
4	名稱	ROUTE_NAME	nvarchar(11)
5	路線里程	ROUTE_LENGTH	decimal(8,1)
6	行駛班次	SCHEDULES	int
7	行駛里程	MILEAGE	decimal(8,1)
8	人數合計	PASSENGERS	decimal(8,1)

9	延人公里	PASSENGER_KILOMETER	decimal(8,1)
---	------	---------------------	--------------

資料表名稱	FUEL		
說明	油耗		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	日期	DATE	date
3	車號	BusID	varchar(7)
4	油量	FUEL	decimal(28,0)

資料表名稱	RECORDS_OF_BEHAVIOUR		
說明	獎懲記錄		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	發文日期	POSTING_DATE	date
3	發文字號	POSTING_ID	nvarchar(7)
4	部門別	STATION	nvarchar(3)
5	員工姓名	DRIVER_NAME	nvarchar(3)
6	到職日期	FIRST_DAY_OF_WORK	date
7	離職日期	LAST_DAY_OF_WORK	date
8	獎懲記錄	REWARD_TYPES	nvarchar(4)
9	獎懲事由	RECORDS_OF_BEHAVIOUR	nvarchar(90)

資料表名稱	ROUTE		
說明	路線一覽表		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	序號	Permit_No	int
3	許可證字號	Permit_ID	nvarchar(18)
4	路線編號	ROUTE_ID	nvarchar(4)
5	營運路線名稱	ROUTE_NAME	nvarchar(18)
6	里程小計	Length	decimal(5,1)
7	里程往程	LENGTH_TOWARD	decimal(5,1)
8	里程返程	LENGTH_RETURN	decimal(5,1)
9	班次平日小計	TRIP_WEEKDAYS	int
10	班次平日往	TRIP_WEEKDAYS_TOWARD	int
11	班次平日返	TRIP_WEEKDAYS_RETURN	int

12	班次假日小計	TRIP_WEEKENDS	int
13	班次假日往	TRIP_WEEKENDS_TOWARD	int
14	班次假日返	TRIP_WEEKENDS_RETURN	int
15	備註	MEMO	nvarchar(33)
16	實施日期	DATE	nvarchar(9)

資料表名稱	SALARY		
說明	薪資明細		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	月份	PAY_PERIOD	varchar(5)
3	司機代號	DRIVER	varchar(4)
4	出勤天數	DAYS	tinyint
5	行駛里程	MILEAGE	decimal(8,1)
6	里程津貼	MILEAGE_ALLOWANCE	decimal(8,1)
7	載客人數	PASSENGERS	decimal(8,1)
8	載客津貼	PASSENGERS_ALLOWANCE	decimal(8,1)
9	售票金額	TICKET_AMOUNT	decimal(8,1)
10	售票津貼	TICKET_ALLOWANCE	decimal(8,1)
11	延長工時	OVERTIME	decimal(8,1)
12	應領金額	SALARY	decimal(8,1)

資料表名稱	VACATIONS		
說明	請假明細		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	部門	STATION	nvarchar(3)
3	司機代號	DRIVER_ID	nvarchar(4)
4	假別	VACATIONS_TYPE	nvarchar(2)
5	日期	VACATIONS_DATE	varchar(50)
6	起始時間	STARTING_TIME	time
7	結束時間	END_TIME	time
8	天數	DAYS	tinyint
9	小時	HOURS	decimal(3,1)

資料表名稱	VEHICLE		
說明	車輛清冊		
編號	中文說明	欄位名稱	型態

1	車別	VEHICLE_TYPE	nvarchar(3)
2	車號	BUSID	nvarchar(6)
3	廠牌	BRAND	nvarchar(6)
4	年份	YEARS	nvarchar(7)
5	引擎號碼	ENGINE_ID	nvarchar(17)
6	噸位	TON	decimal(4,1)
7	座位	SEATS	nvarchar(6)
8	備註	MEMO	varchar(50)

資料表名稱	VEHICLE_MAINTENANCE		
說明	車輛維修		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	車號	BUSID	varchar(50)
3	維修日期	MAINTENANCE_DATE	date
4	維修時間	MAINTENANCE_TIME	time
5	故障原因	SUBJECTS	varchar(500)
6	進場日期	REPAIRER_DATE	date
7	進場時間	REPAIRER_TIME	time
8	完修日期	COMPLETION_DATE	date
9	完修時間	COMPLETION_TIME	time

資料表名稱	VEHICLE_PARTS		
說明	車輛零件統計表		
編號	中文說明	欄位名稱	型態
1	流水號	ID	int IDENTITY(1,1)
2	車號	BUSID	nvarchar(6)
3	總耗油量	TOTAL_OIL	decimal(10,2)
4	公里柴油公升	DIESEL_FUEL_PER_KILOMETER	decimal(10,2)
5	機油量	ENGINE_OIL	decimal(10,2)
6	公里機油公升	ENGINE_OIL_PER_KILOMETER	decimal(10,2)
7	柴油耗油金額	AMOUNT_DIESEL_FUEL	decimal(10,2)
8	零件費	AMOUNT_VEHICLE_PARTS	decimal(10,2)
9	輪胎費	AMOUNT_TIRE	decimal(10,2)
10	機油費	AMOUNT_ENGINE_OIL	decimal(10,2)
11	雜油費	AMOUNT_OIL	decimal(10,2)
12	耗用總計	AMOUNT_CONSUMPTION	decimal(10,2)
13	耗用公里	AMOUNT_KILOMETER	decimal(10,2)

14	年份	YEARS	nvarchar(6)
15	廠牌	BRAND_VEHICLE	nvarchar(6)

資料表名稱	e_bus_avg_power		
說明	電動公車充電資料		
編號	中文說明	欄位名稱	型態
1	流水號	ebap_no	decimal(18,0)
2	日期	ebap_date	nvarchar(7)
3	車號	ebap_bus	nvarchar(5)
4	開始月份	ebap_month_start	int
5	結束月份	ebap_month_end	int
6	總電量	ebap_month_total_power	float

資料表名稱	e_bus_main_parts_fail		
說明	電動公車維修資料		
編號	中文說明	欄位名稱	型態
1	流水號	ebmpf_no	decimal(18,0)
2	車號	ebmpf_bus	nvarchar(5)
3	開始月份	ebmpf_time_start	datetime
4	結束月份	ebmpf_time_end	datetime
5	公里數	ebmpf_drive_km	int
6	地區	ebmpf_location	nvarchar(50)
7	事件說明	ebmpf_item	nvarchar(200)
8	後續情況	ebmpf_repair	nvarchar(200)

資料表 A1

--

資料表 A2

資料表:accident

附錄 D 資料表欄位範例

資料表:bus_company

bc_id	bc_name	bc_memo
bus01	大有巴士	
bus02	中鹿客運	
bus03	仁友客運	

資料表:bus_history_route

bhu_code	bhu_date	bhu_bc_id	bhu_driving_route	bhu_route_name	bhu_driving_path	bhu_record_time	bhu_longitude	bhu_latitude	bhu_speed_up	bhu_speed_avg	bhu_
1376308	910-FM	bus19	9015	北港→斗南交流道→國道1號高速公路→中港交流道→臺中車站	182	2015-01-28 16:59:45.000	120.30169	23.56959	9	9	定係
1376308	910-FM	bus19	9015	北港→斗南交流道→國道1號高速公路→中港交流道→臺中車站	281	2015-01-28 17:00:05.000	120.30134	23.56911	10	20	定係
1376308	910-FM	bus19	9015	北港→斗南交流道→國道1號高速公路→中港交流道→臺中車站	287	2015-01-28 17:00:25.000	120.29993	23.5692	10	25	定係

資料表:Easycard_Provider_Route

ProviderID	ProviderName	RouteID	RouteNO	RouteName
033030	巨業客運	16355	6355	沙鹿-龍井奉天宮
033030	巨業客運	26355	6355	沙鹿-龍井奉天宮

資料表:EasyCard-Taichung

ID	cmp	provider	card	cardno	cardid	routeid	driverid	busid	ontime	onseq	onname	onseqp	onnamep	offtime	offseq	offname	offseqp	offname
4542	EASycARD	033031	oxd5Q8yRqVC	1	1	55	940	013FY	2014-05-10 14:16:17.000	7		14		2014-05-10 14:20:43.000	9		18	
4543	EASycARD	033031	IZMykAMVC...	1	1	55	940	013FY	2014-05-10 14:21:22.000	9		18		2014-05-10 14:35:21.000	19		38	
4544	EASycARD	033031	YTMaTmoSO...	1	1	55	940	013FY	2014-05-10 15:16:38.000	46		92		2014-05-10 15:38:23.000	25		50	

資料表:Easycard_Ticket

TicketID	TicketType
0	無
1	一般
2	敬若1

資料表:ETSTOP

ID	RouteID	Type1	Type2	StopID	StopName	GoBack	Longitude	Latitude
1	1086	F	0	FT013	豊勢	1	120.62861800	24.25741500
2	1086	D	0	0618A	豊勢老人会館	1	120.62695500	24.26107700
3	1086	D	0	754	東新園中	1	120.62663200	24.26272000

資料表:fv_driver_roster

ID	FDR_date	FDR_goBack	FDR_name	FDR_weak	FDR_sqrno	FDR_route	FDR_departure_stop	FDR_departure_time	FDR_via_stop_2	FDR_via_time_2	FDR_via_stop_3	FDR_via_time_3	FDR_via_stop_4	FDR_via_time_4	FDR_via_stop_5	FDR_via_time_5	FDR_destination	FDR_route_mileage	FDI
1	1030527	2	514	2	6	210	土城	15:35									大甲	10.7	25
2	1030527	1	514	2	7	216	大甲	16:00									南埔	7.2	20

資料表:mcardf

date	driver_id	route_id	car_id	d_no	t_comp	kind	t_name	card	uptime	upno	up_mpy	up_mnt	up_cp	up_ct	up_fm	up_ts	upname	upbeg	up_no	up_tran	dtime						
1040412	1387	91	642FX	236	TWSC	代格全票	61-0	33R8F140	11:24:47.00000000	111	0.000	0.000	0.000	0.000	0.000	0.000	豊原	豊原	1	0	12:00:56.000000C						
1040412	1387	91	642FX	236	TWSC	代格全票	61-0	33R8F140	13:15:31.00000000	113	0.000	0.000	0.000	0.000	0.000	0.000	豊原	豊原	1	0	13:26:53.000000C						
1040412	1387	91	642FX	236	TWSC	敬老愛殘	台中...	3ABAD...	10:35:47.00000000	4636	11.000	0.000	11....	11....	0.000	0.000	福德祠	福...	1	1	11:23:50.000000C						
dtime	din_mpy	din_mnt	din_cp	din_ct	din_tm	din_ts	dname	dmerk	din_no	din_tran	total	r_beg	r_end	r_mnt	tex	texmoney	no_tex	del	r_mntb	subtime	stime	same_st	out	text	cnt	tot	modintr
112	33.000	33.000	0.000	0.000	0.000	0.000	電信局前	電信局前	0	0	33.000			13.6	2	31.000	0		0.0		0			0			
114	20.000	20.000	0.000	0.000	0.000	0.000	社皮靴...	社皮靴...	0	0	20.000			8.5	1	19.000	0		0.0		0			0			

資料表:N1

N1	STOPID	ROUTEID	BUSID	CURRENTSTOP	DESTINATIONSTOP	ISLASTBUS	ESTIMATETIME	STOPDISTANCE	DIRECTION	TYPE	TRASTIME	SN	RECTIME
N1	13814	92	708-FT	13814	13800		720		1	2	2014-01-01 00:55:13.000		2014-01-01 00:55:13.00
N1	13814	92	708-FT	13814	13799		780		1	2	2014-01-01 00:55:13.000		2014-01-01 00:55:13.00
N1	13814	92	708-FT	13814	13798		780		1	2	2014-01-01 00:55:13.000		2014-01-01 00:55:13.00

資料表:OBU

BUS	GPSTime	Longitude	Latitude	Speed	RPM	FrontDoor	RearDoor	Acc
012FV	2015-04-01 05:50:48.000	0	0	0	0	0	0	1
012FV	2015-04-01 05:50:49.000	0	0	0	9	0	0	1
012FV	2015-04-01 05:50:50.000	0	0	0	0	0	0	1

資料表:TaichungPROVIDER

PROVIDER	ID	NAMEZH	TYPE
2000	2000	台中市交通處	0
9000	9000	銓鼎科技	0
6	6	豐原客運	0

資料表:TaichungROUTE

ROUTE	ID	PROVIDERID	NAMEZH	DDESC	DEPARTUREZH	DESTINATIONZH
藍1	藍1	15	藍1	統聯轉運站-臺中工業區	臺中工業區	統聯轉運站
藍1	藍1	3	藍1	統聯轉運站-臺中工業區	臺中工業區	統聯轉運站
藍2	藍2	5	藍2	綠能環線街車	臺中火車站(...	臺中火車站(左環)

資料表:TaichungSCHEDULE

Route	departureTime	Goback	Schedule_date
100	1900-01-01 05:30:00.000	1	2015-05-18 10:14:56.673
100	1900-01-01 05:40:00.000	1	2015-05-18 10:14:56.677
100	1900-01-01 05:52:00.000	1	2015-05-18 10:14:56.677

資料表:TaichungSTOP

STOP	ID	ROUTEID	NAMEZH	SEQNO	PGF	TERMINAL	DISTRICID	GOBACK	LATITUDE	LONGITUDE
2444	2444	128	東港椰	10	0	0	1	2	24.26446800	120.5489780
3895	3895	128	糠椰	11	0	0	1	2	24.26361500	120.5456780
4810	4810	128	竹園(港...	12	0	0	1	2	24.26180000	120.5415800

資料表:hbProvider

STOP	ID	ROUTEID	NAMEZH	SEQNO	PGF	TERMINAL	DISTRICID	GOBACK	LATITUDE	LONGITUDE
2444	2444	128	東港椰	10	0	0	1	2	24.26446800	120.5489780
3895	3895	128	糠椰	11	0	0	1	2	24.26361500	120.5456780
4810	4810	128	竹園(港...	12	0	0	1	2	24.26180000	120.5415800

資料表:hbRoute

Provider	Route	Id	provided	nameEn	nameZh	departureEn	departureZh	destinationEn	destinationZh	opED	opED	status	atTime	seTime	STD	TST	type	stopTimes
18	000601	000601	18	Hsuehu County Zoubai City Office—THSR Hsuehu S...	竹北市公所—高雄新竹站	Hsuehu County Zoubai City Office	竹北市公所站	THSR Hsuehu Station	新竹高鐵站	2014/6/6 上午 12:00:00	2017/6/6 上午 12:00:00	1	NULL	NULL	NULL	NULL	1	11
18	000602	000602	18	THSR Hsuehu Station—Hsuehu County Zoubai City ...	高雄新竹站—竹北市公所	THSR Hsuehu Station	新竹高鐵站	Hsuehu County Zh...	竹北市公...	2014/6/6 上午 12:00:00	2017/6/6 上午 12:00:00	1	NULL	NULL	NULL	NULL	1	13
18	000701	000701	18	Ku Chi Peak—Kuanhsung Station	古奇峰—新莊車站(先行...	Ku Chi Peak	古奇峰	Ku Chi Peak	古奇峰	2013/11/22 上午 12:0...	2014/12/13 上午 12:0...	1	NULL	NULL	NULL	NULL	1	47

資料表:hbSchedule

Provider	Route	Stop	Schedule	stopId	serviceId	aTimeReg	dTimeReg	aTimeRush	dTimeRush	mon	tue	wed	thu	fri	sat	sun	special	dates	status
13	079001	295708	20894	295708	1	0610	0610	0610	0610	True	True	True	True	True	True	True	0		1
13	079001	295708	20894	295708	2	0900	0900	0900	0900	True	True	True	True	True	True	True	0		1
13	079001	295708	20894	295708	3	1500	1500	1500	1500	True	True	True	True	True	True	True	0		1

資料表:THBStop

Provider	Route	Stop	nameEn	nameZh	seqNo	Pgp	Terminal	districtId	Latitude	Longitude	Address
13	5617A2	126687	Longyuan	龍源	41	0	0	0	24.82819000	121.2058300	桃園市龍潭區中豐路與龍源路口前50M對面
13	5617A2	100647	Geoping	高平	42	0	0	0	24.82379000	121.2016300	桃園市龍潭區中豐路牛欄分線16支電桿
13	5617A2	126688	Shueang	水倉	43	0	0	0	24.81768000	121.2017800	新竹縣關西鎮中豐路72號

資料表:THBTicket

Provider	Route	Tickets	Ticket	Id	RouteId	departureSeqNo	destinationSeqNo	price	currency	period
13	001001	0	半票_不分時段_不分車種	半票_不分時段_不分車種	001001	0_火車站	1_竹中	8.00	NTD	
13	001001	0	全票_不分時段_不分車種	全票_不分時段_不分車種	001001	0_火車站	1_竹中	15.00	NTD	
13	002001	0	半票_不分時段_不分車種	半票_不分時段_不分車種	002001	0_火車站	1_交大光復校...	8.00	NTD	

資料表:TOrder

ID	UPDATE	Lat1	Lon1	DMDATE	Lat2	Lon2	TPGEO	DMGEO
1	2014-06-05 17:17:32.000	24.1421630	120.6854400	2014-06-05 17:34:08.000	24.1421580	120.6847570	0xE61000000010C85B3BCB64243840B2D7BB3FDE2B5E40	0xE61000000010C61DD787764243840412D060FD32B5E4
2	2014-06-05 17:34:10.000	24.1421580	120.6847570	2014-06-05 17:34:12.000	24.1421580	120.6847570	0xE61000000010C61DD787764243840412D060FD32B5E40	0xE61000000010C61DD787764243840412D060FD32B5E4
3	2014-06-05 17:36:25.000	24.1422820	120.6850170	2014-06-05 17:36:27.000	24.1422820	120.6850170	0xE61000000010C3ACFD8976C2438400R0D8B51D72B5E40	0xE61000000010C3ACFD8976C2438400R0D8B51D72B5...

資料表:weather_area

wa_id	wa_name	wa_memo
A01	臺中	觀測站
A02	梧棲	觀測站

資料表:weather_city

wc_id	wc_name
A	台中地區

資料表:weather_data

wd_index	wd_wa_id	wd_date	wd_hour	wd_hpa_1	wd_hpa_2	wd_temperature_1	wd_temperature_2	wd_humidity	wd_wind_mps	wd_wind_degree	wd_gust_mps	wd_gust_degree	wd_rain_mmm	wd_rain_hr	wd_sunshine_hr	wd_sunshine_mj	wd_visibil
1	A02	2013-01-01	1	1017	1020.3	10	6.6	79	3	40	4.4	40	0	0	0	0	0
2	A02	2013-01-01	2	1016.9	1020.2	10.2	6.9	80	2.4	40	4.3	30	0	0	0	0	0
3	A02	2013-01-01	3	1016.6	1019.9	10.7	7.3	79	2.7	40	4	40	0	0	0	0	0

資料表:DAILY_REPORT

ID	DATE	ROUTE	ROUTE_NAME	ROUTE_LENGTH	SCHEDULES	MILEAGE	PASSENGERS	PASSENGER_KILOMETER
6738	2014-05-18	992	豐原調度	NULL	120	521.4	0.0	0.0
6739	2014-05-18	993	台中調度	NULL	145	437.6	0.0	0.0
6740	2014-05-18	994	沙鹿調度	NULL	15	90.3	0.0	0.0

資料表:FUEL

id	DATE	BusID	FUEL
10566	2015-04-25	270FM	188
10567	2015-04-25	708FT	140
10568	2015-04-25	688FX	133

資料表:RECORDS_OF_BEHAVIOUR

ID	POSTING_DATE	POSTING_ID	STATION	DRIVER_NAME	FIRST_DAY_OF_WORK	LAST_DAY_OF_WORK	REWARD_TYPES	RECORDS_OF_BEHAVIOUR
1	2014-05-02	總字第099號	東勢站	謝	2013-06-13	NULL	小過1次	該員於1003.17重擊乘客F4號,由谷龍站開08時30分往東勢266班車,未依核定路線行駛
2	2014-05-02	總字第100號	卓蘭站	江	1998-11-05	NULL	小過1次	該員於103.04.07重擊乘客F1號,由水尾開19時往卓蘭班車,未依核定路線開至卓蘭市區
3	2014-05-02	總字第101號	豐原站	張	2001-06-12	NULL	小過1次	該員於103.03.19重擊乘客F7號,由光明開13時20分往豐原站55路班車,行經導秀

資料表:ROUTE

Id	Permit_No	Permit_ID	ROUTE_ID	ROUTE_NAME	Length	LENGTH_TOWARD	LENGTH_RETURN	TRIP_WEEKDAYS	TRIP_WEEKDAYS_TOWARD	TRIP_WEEKDAYS_RETURN	TRIP_WEEKENDS	TRIP_WEEKENDS_TOWARD	TRIP_WEEKENDS_RETURN	MEMO
1	1	中市交通稽字第100034349號	12	明港高中-豐原高中	45.4	22.7	22.7	204	102	102	156	78	78	全國定期日
2	1	中市交通稽字第100034349號	12	豐香班次(與全航、中港聯營)上半年	45.4	22.7	22.7	48	24	24	38	19	19	
3	1	中市交通稽字第100034349號	12	豐香班次(與全航、中港聯營)下半年	45.4	22.7	22.7	54	27	27	40	20	20	

資料表:SALARY

ID	PAY_PERIOD	DRIVER	DAYS	MILEAGE	MILEAGE_ALLOWANCE	PASSENGERS	PASSENGERS_ALLOWANCE	TICKET_AMOUNT	TICKET_ALLOWANCE	OVERTIME	SAL.
1841	10311	009	25	2878.3	5761.0	9221.0	3801.0	195814.0	9323.0	14.5	457.
1842	10311	002	26	132.9	266.0	244.0	99.0	5347.0	254.0	0.5	448.
1843	10311	007	27	11032.6	21313.0	884.0	375.0	87785.0	4179.0	0.0	593.

資料表:VACATIONS

ID	STATION	DRIVER_ID	VACATIONS_TYPE	VACATIONS_DATE	STARTING_TIME	END_TIME	DAYS	HOURS
18698	015	0928	公休		103070	17:00:00.00000000	1	0.0
18699	015	0928	公休		103070	17:00:00.00000000	1	0.0
18700	015	0928	公休		103071	17:00:00.00000000	1	0.0

資料表:VEHICLE

VEHICLE_TYPE	BUSID	BRAND	YEARS	ENGINE_ID	TON	SEATS	MEMO
大營客	205-FC	國瑞	2003.01	J08C-TP16429	14.0	47	
大營客	206-FC	國瑞	2003.01	J08C-TP16408	14.0	47	
大營客	207-FC	國瑞	2003.01	J08C-TP16428	14.0	47	

資料表:VEHICLE_MAINTENANCE

ID	BUSID	MAINTENANCE_DATE	MAINTENANCE_TIME	SUBJECTS	REPAIRER_DATE	REPAIRER_TIME	COMPLETION_DATE	COMPLETION_TIM
1	536FH	2014-05-01	06:20:00.0000000	於台關站無法發動前往救援	2014-05-01	06:20:00.0000000	2014-05-01	09:40:00.0000000
2	607FX	2014-05-01	09:35:00.0000000	儀表燈檢修	2014-05-01	09:35:00.0000000	2014-05-01	10:00:00.0000000
3	681FX	2014-05-01	01:10:00.0000000	後視鏡檢修前左輪取換	2014-05-01	01:10:00.0000000	2014-05-01	01:55:00.0000000

資料表:VEHICLE_PARTS

ID	BUSID	TOTAL_OIL	DIESEL_FUEL_PER_KILOMETER	ENGINE_OIL	ENGINE_OIL_PER_KILOMETER	AMOUNT_DIESEL_FUEL	AMOUNT_VEHICLE_PARTS	AMOUNT_TIRE	AMOUNT_ENGINE_OIL	AMOUNT_OIL	AMOUNT_CONSUMPTION	AMOUNT_KILOMETER	YEARS	BRAND_YEHIC
1	012FY	0.00	0.00	0.00	0.00	11486.00	0.00	0.00	0.00	11486.00	0.39	2013	豐裕	
2	013FY	0.00	0.00	0.00	0.00	16139.00	0.00	0.00	0.00	16139.00	0.39	2013	豐裕	
3	015FY	0.00	0.00	0.00	0.00	24117.00	0.00	0.00	0.00	24117.00	0.39	2013	豐裕	

資料表:e_bus_avg_power

ebsp_no	ebsp_date	ebsp_bus	ebsp_month_start	ebsp_month_end	ebsp_month_total_power
1	2013/12	012FY	472	1282	1254.2
2	2013/12	013FY	633	1783	1537.5
3	2013/12	015FY	58	931	1192.5

資料表:e_bus_main_parts_fail

ebspf_no	ebspf_bus	ebspf_time_start	ebspf_time_end	ebspf_drive_km	ebspf_location	ebspf_item	ebspf_repair
1	021FY	2013-12-21 20:20:00.000	2013-12-21 23:00:00.000	563	台中火車站	車輛異常,無法行駛	現場更換整車控制器後開回充電站
2	013FY	2014-01-02 13:00:00.000	2014-01-02 17:00:00.000	1943	台中女中	1.無動力轉向功能 2.冷氣無送風功能	1.動力方向機馬達保險絲燒損 2.冷氣63A 700W保險絲燒損
3	017FY	2014-01-07 14:00:00.000	2014-01-07 20:00:00.000	2280	豐原站	空壓機無法打氣	空壓機本體溫度感知器異常

附錄 E 參考文獻

1. ITS 整體發展規劃，2009，交通部運輸研究所。
2. 邱詩淳（2005）。運用悠遊卡及資料探勘求解公車營運改善方案。中華大學運輸科技與物流管理學碩士班論文。
3. 陳一昌、黃維信、賴靜慧（2010），道路交通事故空間定位資訊之應用探討，國土資訊系統通訊，Vol. 75，頁 53-67。
4. 羅惟元（2008）。以悠遊卡交易資料探索公車路線之旅客起迄。淡江大學運輸管理學系碩士班論文。
5. 蘇昭銘、何文基、陳俐融、陳允中、劉時勳（2014），新竹市交通事故之空間特性關聯分析，103 年道路交通安全與執法研討會。
6. 蘇昭銘、游坤明、張靖、林至康（2014），先進公共運輸系統整合資料庫加值應用系統維運及推廣計畫，交通部運輸研究所。
7. 蘇昭銘等人（2011），台南市公車動態資訊在站點時刻表管理之研究，華夏科技股份有限公司。
8. Agard, B., Morency, C., and Trépanier, M. (2006) *Mining public transport user behaviour from smart card data*. 12th IFAC Symposium on Information Control Problems in Manufacturing – INCOM 2006, Saint-Etienne, France, May 17–19.
9. Aguero-Valverde, J. and Jovanis, P. P. (2006), Spatial analysis of fatal and injury crashes in Pennsylvania, *Accident Analysis & Prevention*, 38 (3), 618-625.
10. Anwar, A., Zeng, W., and Arisona, S.M. (2015) The Time Space Diagram Revisited. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2442.

11. Bagchi, M. and White, P.R. (2004) . What role for smart-card data from bus system? *Municipal Engineer* 157, March 2004, pp. 39-46.
12. Bagchi, M., and White, P.R. (2005). The potential of public transport smart card data. *Transport Policy*, 12, pp. 464-474.
13. Chapleau, R., and K.K. Chu. (2007). *Modeling transit travel patterns from location-stamped smart card data using a disaggregate approach*, 11th World Conference on Transportation Research CDROM, Berkeley, California.
14. Chu, K. K., Chapleau, R., and Trépanier, M. (2009). *Driver-Assisted Bus Interview (DABI) : Passive Transit Travel Survey using Smart Card Automatic Fare Collection System and its Applications*. 88th Annual Meeting of the Transportation Research Board CDROM, Washington, 21 p.
15. Du, F., Brulé, J., Enns, P., Manjunatha, V., and Segev, Y. (2013) *MetroViz: Visual Analysis of Public Transportation Data*. Website : https://wiki.cs.umd.edu/cmsc734_f13/images/6/60/MetroViz_final_paper.pdf.
16. Gokasar, I., Simsek, K., and Ozbay, K. (2015) *Using Big Data Of Automated Fare Collection System For Analysis And Improvement of Brt- Bus Rapid Transit Line in Istanbul*. Proceedings of Transportation Research Board 94th Annual Meeting, Washington, DC.
17. Han, J., and Kamber, M., (2001) , *Data Mining: Concepts and Techniques*, San Francisco: Morgan Kaufmann.
18. Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A. and Khan, S. U., (2015) , The rise of big data on cloud computing: Review and open research issues, *Information Systems*, 47, 98-115.
19. Jagadish, H. V., Gehrke, J., Labrinidis, A., Papakonstantinou, Y., Patel, J. M., Ramakrishnan, R., and Shahabi, C., (2014) , Big data

- and its technical challenges, *Communications of the ACM*, 57 (7) , 86-94.
20. Kambatla, K., Kollias, G., Kumar, V. and Grama, A., (2014) , Trends in big data analytics, *Journal of Parallel and Distributed Computing*, 74 (7) , 2561-2573.
 21. Levine, N., Kim, K. E., and Nitz, L. H. (1995) . Spatial analysis of Honolulu motor vehicle crashes: I. Spatial patterns. *Accident Analysis & Prevention*, 27 (5) , 663-674.
 22. Liao, C.F. (2011) Data-Driven Support Tools for Transit Data Analysis, Scheduling and Planning, *Technical Report (CTS 11-15)* , Intelligent Transportation Systems Institute Center for Transportation Studies, University of Minnesota.
 23. Morency, C., Trépanier, M., and Agard, B., (2007) . Measuring transit use variability with smart-card data. *Transport Policy*, Volume 14, Issue 3, pp. 193-203.
 24. Pelletier, M. P., Trépanier, M., and Morency, C. (2011) . Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19 (4) , 557-568.
 25. Seaborn, C., Wilson, N. H., and Attanucci, J. (2009) . *Using Smart Card Fare Payment Data To Analyze Multi-Modal Public Transport Journeys (London, UK)*. 88th Annual Meeting of the Transportation Research Board CDROM, Washington, 16 p.
 26. Tan, P. N., Steinbach, M., and Kumar, V., (2006) , *Introduction to Data Mining*, Boston: Addison Wesley.
 27. Traynor, T. L. (2008) . Regional economic conditions and crash fatality rates—a cross-county analysis. *Journal of safety research*, 39 (1) , 33-39.

28. Trépanier Martin, Morency Catherine, and Agard Bruno (2009) .
Calculation of Transit Performance Measures Using Smartcard Data,
Journal of Public Transportation, 12 (1) , pp. 79–96.
29. van Oort, N., Sparing, D., Brands, T., and Goverde, R.M.P. (2013)
*Optimizing Public Transport Planning and Operations Using
Automatic Vehicle Location Data: The Dutch Example*. MT-ITS
conference Dresden.
30. Wu, X., Zhu, X., Wu, G.-Q., and Ding, W., (2014) , Data Mining
with Big Data, *IEEE Transactions on Knowledge and Data
Engineering*, 26 (1) , 97-107.
31. Zeng, Q., Reddy, A., Lu, A. and Levine, B. (2015) *Develop New
York City Surface Transit Boarding and Alighting Ridership Daily
Production Application Using Big Data*. Proceedings of
Transportation Research Board 94th Annual Meeting, Washington,
DC.

附錄 F 期中報告審查意見回覆表

一、開會時間：104 年 08 月 14 日下午 2 時

二、開會地點：運研所五樓會議室

三、主持人：陳副所長茂南

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 依發言序	審查意見	回覆辦理情形	主辦單位 查核意見
健行科技大學 行銷與流通管 理系 朱副教 授松偉	1. 建議國道客運、公路客運 與市區客運路線三種業 別的分析應分開討論。	本計畫人員將與運研 所討論可執行之項目。	
	2. 臺中市日前廢除 BRT， 請說明本案對於相關路 線變更的分析方式。	本研究之供給以班次 數做為分析，需求以電 子票證資料分析。路線 變更後仍以此兩種資 料進行供需分析，與 BRT 存廢無關。	
	3. 兩次座談會的會議和簽 到紀錄顯示出席比例不 是很高，請團隊條列式說 明座談會結論與需求，並 探討計畫後續研究方向 是否有配合座談會業者 需求。建議團隊期末可考 慮增加辦理一場需求座 談會。	本計畫受限於時程較 短，目前致力於數據分 析，如果時間許可，即 會增加需求座談會場 次。	
	4. 報告書中無法看出期中 簡報彩色內容，建議團隊	遵照辦理，將於期末報 告時上傳報告書至雲	

	期末報告書放至雲端資料讓委員自行下載，讓委員較能了解分析結果。	端供委員下載。	
	5. 目前本案相關分析項目，部分延續過去運研所計畫所產出的結果，如路廊分析與區域分析，請團隊說明這些項目是否能滿足本案？	前期計畫已確認需求。	
台灣大學土木工程系 康教授 仲	1. P33 頁車輛故障圖無法讓人一目了然，建議團隊應思考用方塊圖呈現的清晰度。	遵照辦理，將修訂於期末報告中。	
	2. P28 頁連班圖，放著不動的靜態圖應該要能很清楚呈現，視覺化呈現的每個階段都要清楚。	遵照辦理，將修訂於期末報告中。	
	3. P22 頁的視覺化呈現，除了趨勢展現，更應注意離群值所代表的意義。	遵照辦理，將修訂於期末報告中。	
	4. P23 頁的視覺化解析應以更直觀的方式闡述專業判斷，並強化假設與論點的關聯，以達到外行人亦能輕鬆解讀視覺化資訊的目標。	遵照辦理，將修訂於期末報告中。	
	5. 請團隊強化說明分析項目的處理流程，包含資料如何處理與補遺邏輯。	遵照辦理，將於期末報告內提出。	

逢甲大學運輸科技與管理系 李副教授克聰	1. 請團隊暫先不要從本案衍生太多結論，因為豐客有許多路線尚有其他業者經營。	敬悉。	
	2. 建議以問題導向延續後續研究方向，先定義清楚我們必須先找到哪些非常有價值的資訊。	將與豐客討論欲達成目標，以此為依據精進分析。	
	3. 建議第二章的視覺化分析案例，可以將外國應用公共運輸大數據案例的效果和效益進行彙整。	遵照辦理，將於期末報告內提出。	
	4. 建議第四章每小節的分析項目皆應補上小結進行討論。	遵照辦理，將於期末報告內提出。	
臺北醫學大學醫務管理學系暨研究所 謝教授邦昌	1. 請團隊強化說明本案整體資料分析架構(資料蒐集，清理至資訊呈現)。	遵照辦理，將修訂於期末報告中。	
	2. 請團隊說明目前掌握的資料，和目前各分析項目需要但短缺的資料。	遵照辦理，將於期末報告內提出。	
	3. 建議團隊探討路線轉乘的關聯性，以及人的關聯和區域的關聯，再進行群聚分析，並納入人口結構和商業結構進行更多價值性的探索。	會視電子票證資料取得完整性，決定分析內容。	

	4. 建議團隊提供本案整體軟體與硬體之分析架構，包含本案使用各軟體的結合方式。	遵照辦理。	
交通部公路總局公路人員訓練所王所長在莒	1. 建議團隊修改期中報告撰寫方式，避免使用服務建議書撰寫方式。	遵照辦理，將修訂於期末報告中。	
	2. 建議團隊思考公路客運和市區客運的競合。	視與運研所討論業別可分開分析之項目結果，進行競合。	
	3. 視覺化分析項目應思考為誰設計？使用者的不同視覺化分析內容應有所不同，政府機關關心的是風險管理，特別是歷史資料和異常事件的視覺化資訊。	本計畫主要設計對象為主管機關及業者，後續將加強視覺化設計。	
	4. 建議團隊檢視油耗分析結論撰寫方式，目前未納入累積里程數分析，應謹慎推導出廠牌影響的關聯性結論。	遵照辦理，廠牌影響部分將模糊化處理。	
運研所運計組	1. 依報告書 2.2.1 節巨量資料定義具有 3 個 Vs 甚至擴大至 5 個 Vs 特徵，本計畫蒐集公車動態資訊是否都具有 5 個 Vs 特徵？在公共運輸管理中如何使用，建請補充說明。	遵照辦理，將於期末報告內提出。	
	2. 報告書 4.1.3 演算處理流程資料完整性檢驗是否有標準作業程序？資料檢核是否有使用電腦程式進行預檢作業或僅為人工判讀；另是否有「無效資料」之判定標準，建請補充說明。	遵照辦理，將於期末報告內提出。	
	3. 報告書第 4 章各項分析	遵照辦理，將於期末報	

	原始資料格式及名稱均不相同，建議研議整併之可行性，俾利後續資料蒐集及應用。	告內提出。	
	4. 第2章文獻回顧部分，建議與節末參考文獻對照表編號聯結。	遵照辦理，將修訂於期末報告中。	
	5. P3-7~10，建議將需求部分加以權重化，將各需求分析資料取得難易、處理時間、對於分析問題之重要性納入權重表補述說明需求排序。	需求部分資料可取得之項目皆已納入計畫分析，故無需進行排序。	
	6. P4-5 資料完整性檢驗部分，請補述視為無效資料的比率及是否有其他判讀或補足資料之方法。	遵照辦理，將於期末報告內提出。	
	7. 建議補述各圖之文字分析部分；如圖4-5分析旅次需求之起迄點如何判斷及圖內文字代表意義。	遵照辦理，將修訂於期末報告中。	
	8. 建議於本計畫中，補述處理何種問題及分析目的、角度與蒐集取得資料之關聯；如4-3路線與班次重疊分析，增述進行探討問題本質及處理問題所需資料（現可取得何種資料、尚需何種資料以補強完善）。	遵照辦理，將於期末報告內提出。	
運研所運資組	第二章「蒐集與分析國外公共運輸巨量資料視覺化案例」部分 (1) 「公共運輸巨量資料視覺化分析」是本研究案重點工作，目前國內相關研究甚少，但國外已	遵照辦理，將修訂於期末報告中。	

	有相當多研究成果與案例，本研究蒐集之案例與探討不多，希望後續研究工作能針對此部分增加比重，以提高視覺化設計參考。		
	(2) 請特別針對電子票證資料加值應用之視覺化分析增加國外案例說明及補充資料，包含自如何自路網整體點、線與面角度進行電子票證資料視覺化分析，分析各案例之視覺化設計元素，以及國外案例如何與跨領域資料包含人口、收入等資料庫進行套疊分析。	遵照辦理，詳細演算流程將於期末報告提出。	
	2.第四章「公共運輸電子票證分析、供需分析視覺化介面設計」部分 (1) 電子票證視覺化分析之時間維度請配合透過商業智慧視覺化分析軟體至少以每小時為單位，進行時間空間暨旅客量三維資料呈現、特定站牌間各時間乘客量分布圖、特定區間乘客量分布圖、特定班次車內乘載量分布圖，並請於4.1.3節加強說明後續如	遵照辦理，將修訂於期末報告並提出詳細演算流程。	

	何將原始資料欄位(上下車時間與站名)進行資料清洗，以流程圖說明完整分析邏輯。		
	(2) 請修正 4.1.3 節演算處理流程說明方式，請配合針對 4.1.5 節呈現之視覺化分析所有成果以流程圖說明資料分析過程。	遵照辦理，將於期末報告修正並提出詳細演算流程。	
	(3) 請針對 4.2.5 節成果利用視覺化分析軟體以地理資訊方式呈現區域 OD 運量探勘之結果。	遵照辦理，將於期末報告內提出。	
	(4) 請強化 4.3 節說明路線班表重疊分析結果，並增加視覺化分析功能，另以案例說明如何透過視覺化檢視公車路線重疊狀況，並請以流程圖完整說明操作流程。	遵照辦理，詳細演算流程將於期末報告提出。	
	(5) 請強化 4.5 節說明站牌班表正確性分析結果，目前該節僅呈現到達時間與預估值平均誤差與標準差。請以路線為單位進行案例說明，增加尖峰與非尖峰各一案例，將 4.5.3 節演算處理流程完整進行案例說明，該案例應包含該路	詳細演算流程將於期末報告提出，並加入精確度之情境分析。	

	線時空圖之繪製與下緣線之調整，求解數學模型之流程，再說明班表產製分析之結果。		
	3.第四章「班表與行車時間、車機穩定度、油耗風險分析」部分 (1) 班表與行車時間、車機穩定度分析請更新研究成果視覺化呈現圖表。	遵照辦理，將修訂於期末報告中。	
	(2) 油耗風險分析請將目前WASTON 初探分析轉換為 Tableau 展示，轉換為以人/車為主體的油耗分析儀表板，請參考車機延遲分析儀表板之設計理念，針對不同資料維度進行鑽探性分析。	遵照辦理，將修訂於期末報告中。	
	4. 後續研究工作建議 (1) 目前本研究電子票證分析成果多為區域分析觀點，並無法了解旅運行為全貌，特別自 3.1 節主管機關需求訪談可得知，許多大數據應用需求均須得知即時乘載率，日前本所舉辦大數據六都論壇，臺灣搭乘公車需上下車皆刷卡之城市其交通首長更強烈	遵照辦理，將於期末報告內提出。	

	表達揭露旅運時空資訊之重要性與急迫性，且本研究目前後續眾多分析亦皆與車輛即時乘載率相關，後續研究請配合以大數據及微觀觀點彙整分析電子票證資料庫，並透過商業智慧視覺化分析軟體設計地理資訊呈現。		
	(2) 請配合將電子票證資料加值應用分析後之資訊(各時間維度下之站牌運量、上下車人數)，與本研究目前已階段性完成公車動態資訊系統之行車時間資訊進行勾稽，並進一步視覺化呈現各時空維度下之乘載率。	遵照辦理，將於期末報告內提出。	
	5. 共通性修訂意見 (1) 報告中許多圖示解析度均不足，以致圖內文字無法辨識，將影響讀者對於內容之理解，後續請針對以上情形全面改善。	遵照辦理，將修訂於期末報告中。	

附錄 G 期末報告審查意見回覆表

一、開會時間：104 年 11 月 13 日下午 2 時

二、開會地點：運研所十樓會議室

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 依發言序	審查意見	回覆辦理情形	主辦單位 查核意見
中央研究院統計科學研究所丘副所長政民	1. 請說明資料處理之系統性方法，以及清除和補遺之相關規則。	已於 ch5 專章說明。	
	2. 本案之資料分析經驗應該要記錄下來，並在報告書中呈現，以利後續相關計畫參考。	遵照評審意見辦理。	
	3. 報告書中無法看出視覺化分析彩色內容，建議乙方往後應彩色印刷或在報告書後附上光碟，讓委員較能了解分析結果。	根據合約規定，後續定稿報告將由甲方印製，故建請運研所印製時做整體考量。	
健行科技大學行銷與流通管理系 朱副教授松偉	1. 請附上期中報告審查意見處理情形。	遵照評審意見辦理	
	2. 有關第二章文獻回顧，請併同相關參考文獻，放到本報告書最後。	遵照評審意見辦理	

	3. 透過本計畫可將大數據資料品質問題找出，建議應追溯至到資料品質供給端，協調電子票證公司或相關單位配合修正資料品質問題。	遵照評審意見辦理，已於 ch7 加強說明。	
政治大學 統計學系 鄭副教授 宇庭	1. 建議往後可透過焦點座談，根據所設計之儀表板，再次將結果回饋予需求端使用者。	受限於專案期限，無法額外辦理超過合約項目之焦點座談，建議後續研究辦理。	
	2. 建議應思考本案如果要給運研所進行成果推廣，其他地方單位之資料結構是否具備相同之資料欄位。	由於各縣市政府公車動態與電子票證均大同小異，故本案成果僅須微幅調整即可推廣應用。	
逢甲大學 運輸科技 與管理系 李副教授 克聰	1. 請豐原客運說明，本案研究成果跟以前客運公司的分析方法，有何具體的差異與效益。	豐原客運羅經理已於審查會議中詳細向委員報告。	
	2. 請問未來台中地區之客運審議機制、補貼機制，能否從研究中得到參考依據。	本研究成果可作為客運審議機制、補貼機制的參考資訊之一。	
	3. 分析維度是否能採取多目標方式，以提供乘載率之外之決策參考指標。	已納入 ch7 未來研究建議第 4 點。	
交通部公路總局 張東閔課長	1. 本案之研究成果對於公路主管機關之監理有非常大之幫助。	感謝評審之肯定。	
	2. 未來汽車客運之審議與路線規劃可參採本案相關成果。	感謝評審之肯定。	
臺北市 政府交通 局書面 審查意見	1. 非引述文字之台北市、台中市與台南市等建議改為臺北市、臺中市與臺南市。	遵照評審意見修改。	

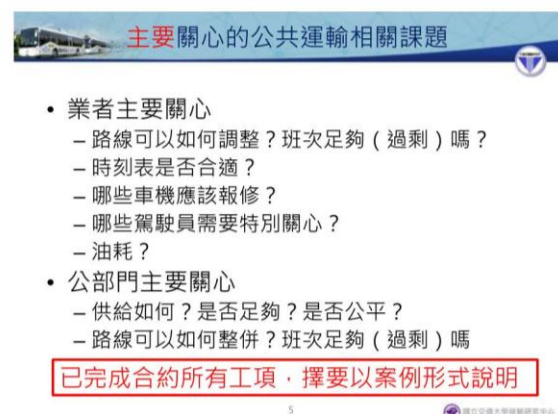
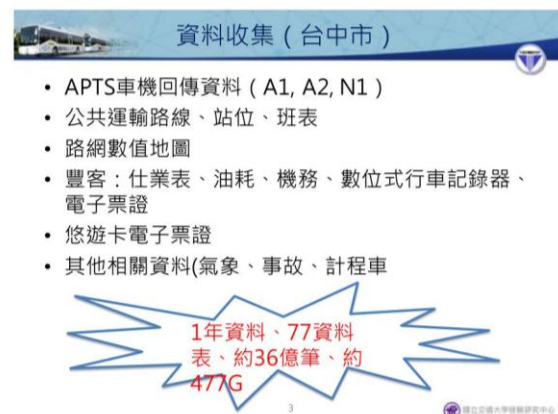
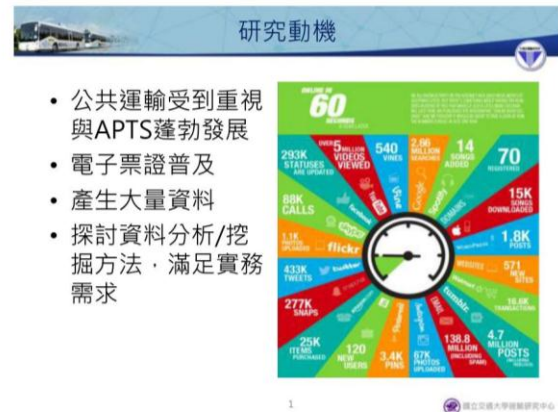
	2. P2-6 現圈式或線圈式車輛偵測器建議釐清。	遵照評審意見釐清並修正。	
	3. P6-14.7 誤植，另 P6-16.8 應為 7。	遵照評審意見修改。	
	4. P6-77 業者風的主力…請修正相關用字。	遵照評審意見修改。	
	5. 資料界接與資料介接用字建議釐清。	遵照評審意見釐清並修正。	
	6. 附錄 C 資料庫欄位說明建議增加附註欄，並提供欄位說明與資料範例。	遵照評審意見另闢附錄 D 提供各資料表範例資料。	
	7. 本案分析成果建議增加結果驗證。	本案乃採全部資料（母體）之分析，不似抽樣方法論需要校估或驗證。	
	8. 建議增加本案後續應用說明（如擴充匯入資料分析或加入其他業者資料）	遵照評審意見於 ch7 說明。	
臺南市政府 交通局 書面審查 意見	1. 本報告書有部份文字誤植，如 P5-1，第 2 點略以「…公路客運資料主要由『公路』總局的聰明公車系統提供…」、第 3 點略以「…豐原客運所提供的數位行車『記』錄器…」，建議定稿時再予詳核。	遵照評審意見釐清並修正。	
	2. 第五章有關資料界接與處理，在 TTIA 所定格式規範下，各縣市政府因應轄內公車運行狀況而有衍生適宜之公車動態系統，資料欄位與定義說明上略有不同。爰此，交通部已有著手研究並	感謝評審之建議，已於 ch7 補充說明此一建議。	

	制訂統一之格式規範，建議可納供參		
	3. 電子票證資料可用以作為相關 OD 資料分析，如圖 6-11 所示。惟本研究因資料取得不易，僅能分析豐原客運上下車資料，未能進一步考量不同業者營駛重複路線、站位時的影響，建議爾後如能取得更多資料，可納供研究分析。	未來研究若能取得完整資料，將根據評審建議辦理。	
	4. 本研究以電子票證作 OD 起訖點分析，是否有考量乘客轉乘行為？如否，可否說明分析困難點以提供各縣市研究參考。	未考量轉乘因素，主要瓶頸在於電子票證資料僅有站序資訊，未能勾稽實際站牌，次要原因則是原始資料缺轉乘碼，次要因素可開發演算法依據卡號補足。以上資訊供各縣市政府參考，同時也在 ch7 說明。	
	5. P6-57，本研究下緣線基礎法的懲罰值採指數分配的型態進行模式建構，並有設定懲罰值。因等待時間增加而使乘客不滿意程度提高，懲罰值採指數分配建構，可否進一步說明為何使用指數分配，與懲罰值參數設定「2、5、20」理由，供各縣市參考。	遵照評審意見辦理	
運資組書 面審查意見	1. 請修正圖 6-11 站牌運量分析圖表說明，改為站牌上下車人數分析。	遵照評審意見修改。	

	2. 請修正圖 6-11 站牌運量分析圖表說明，改為站牌上下車人數分析。	遵照評審意見修改。	
	3. 請針對表 6.5 路段累積班次統計表，補充含 0 與不含 0 之定義，以及表 6.5 的計算方式。	遵照評審意見辦理	
	4. 請補充說明圖 6.23 與圖 6.24 縱軸每千人班次數尺度差異甚大之情形，以及圖 6-26 出現運量為 0 之異常狀況(10-11, 13-14 時)之相關原因。	觀察中區學生人口全日班次數與總人口數(千人)之比值約 213(班次數/千人)，而分時段最高達約 5,234 (班次數/千人)，其原因為非上下學時間，學生 10 時-11 時及 13 時-14 時分別占全日旅次 0.68% 及 0.67%，顯示學生需求很低，但班次仍需正常供給服務其他族群之使用者，故造成全日與分時有較大的差異。圖 6-26 已修正。	
	5. 圖 6-59 跨日班表比較圖，勾選的時間並非同一基準點，不宜以此闡述駕駛員行車狀況差異大之論述，請修正報告書中相關圖表說明。	已更新為同基準（同是平常日或週末）圖形。	
	6. 請說明圖 6-85 以車齡 10 年和 7 年出廠說明故障頻率高和耗油高，但第 11 年後 pattern 不同之原因。	本研究主要利用資料分析之方法挖掘可能產生的問題。然受限於時間與客觀環境，並無法針對問題產生的原因進行探討。	

	7. 有關圖 6-102，時段、星期、月份對傳輸延遲影響不大，請以路線、車輛、業者說明影響因素。	圖 6-102 之原意即是讓讀者判讀何種因素影響大或小？將補充說明該圖之意義與價值。	
主席結論	本案期末報告原則審查通過，請交通大學團隊針對資料處理流程強化說明，並針對資料缺漏問題說明，以讓相關行政機關進行處理。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，於請交通大學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 104 年 11 月 24 日完成修正定稿之提送。	遵照評審意見辦理	

附錄 H 期末報告簡報資料

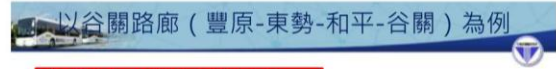




分析案例：路線調整、班次分析、時刻表檢討

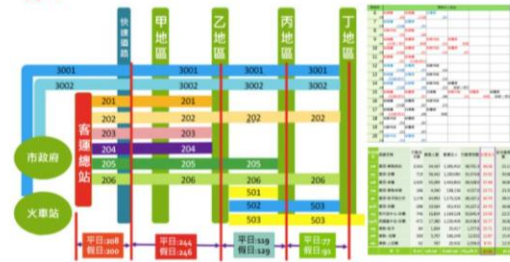
6

國立交通大學運輸研究中心



以谷關路廊（豐原-東勢-和平-谷關）為例

班次過多？過少？還是剛好？



7

國立交通大學運輸研究中心



谷關路廊現行路線



8

國立交通大學運輸研究中心



載容量
(票証)



座位數

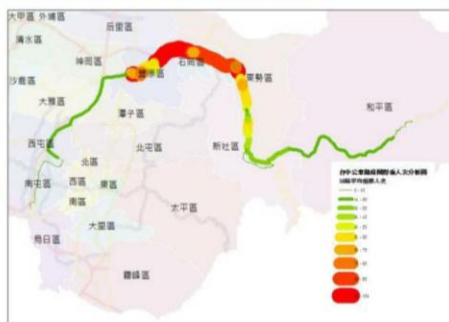
乘載率
(分區段、
分時段)

9

國立交通大學運輸研究中心



16:00~16:59平均搭乘人數



10

國立交通大學運輸研究中心



16:00~16:59班次數



11

國立交通大學運輸研究中心

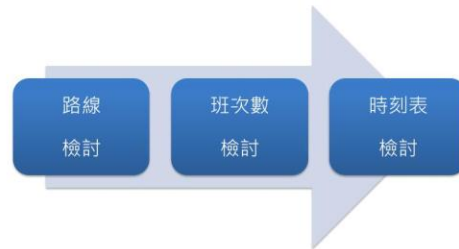
16:00~16:59平均乘載率



12

國立交通大學運輸研究中心

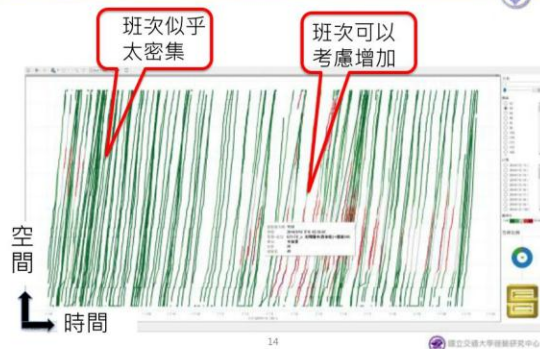
路綫 班次數 時刻表



13

國立交通大學運輸研究中心

班表時空圖連動乘載率儀表板

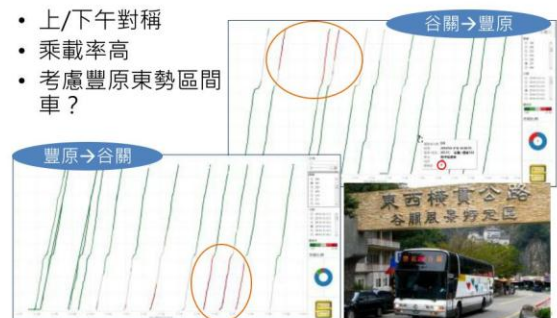


14

國立交通大學運輸研究中心

以207路為例

- 上/下午對稱
- 乘載率高
- 考慮豐原東勢區間車？



獲得市府與業者高度肯定

- 與台中市政府與豐原客運說明
 - 運資組、研究團隊
- 視覺化的呈現方式深獲肯定
- 豐客表示對其營運有極大的助益

16

國立交通大學運輸研究中心

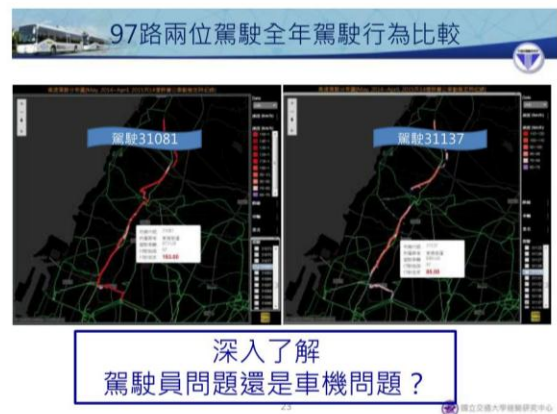
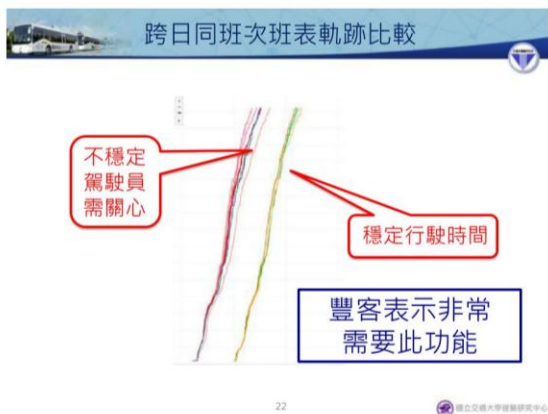
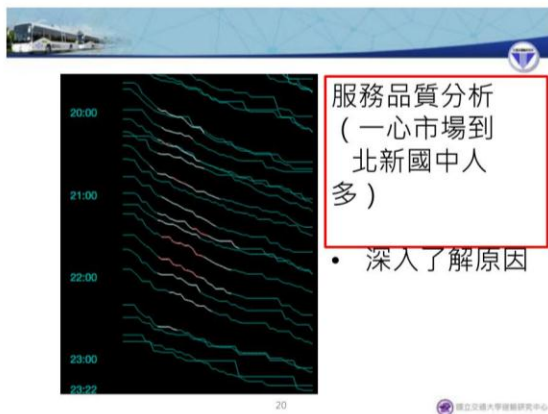
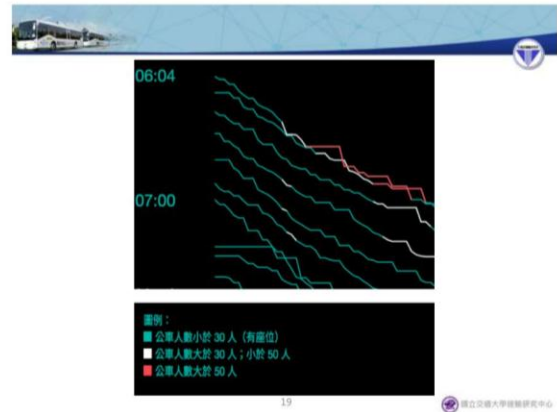
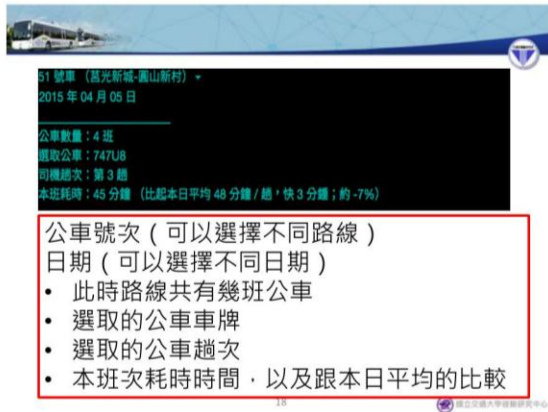
嘗試不同的視覺化呈現方式與工具



提供互動式操作

17

國立交通大學運輸研究中心





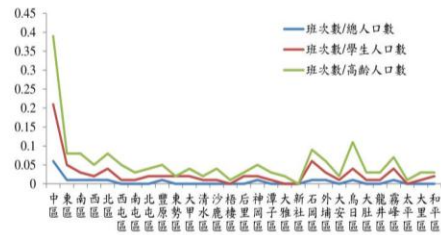
分析案例：各區供給是否公平？特殊車輛配置是否合適？提昇公運使用率的瓶頸？

24

國立交通大學運輸研究中心



各行政區每千人班次數



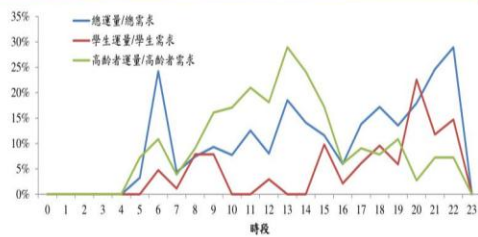
- 較低之區域：資源分配不均，公車之便利性較低，可能無法吸引該區潛在旅次搭乘

25

國立交通大學運輸研究中心



中區分時段運量與潛在需求比例



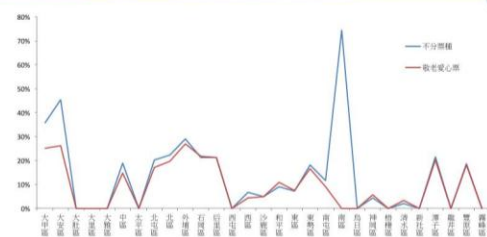
- 尖峰市占率約5%-15%間，離峰時間約20%-25%，仍未將公車作為主要的通勤工具
- 離峰時段，高齡者有很高的比例使用公車

26

國立交通大學運輸研究中心



低地板公車服務比例



- 敬老愛心票種服務比例，南區、大甲區差異大
- 需要搭乘的族群未接受到其服務

27

國立交通大學運輸研究中心



分析案例：車機穩定度分析

故障後
再來報修

VS.

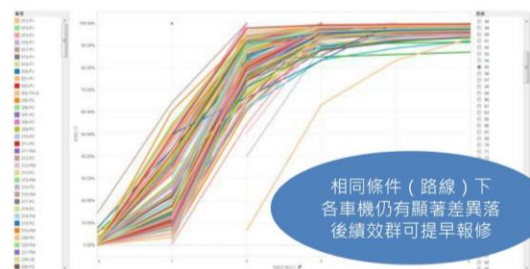
即將故障前
預防性報修

28

國立交通大學運輸研究中心



整體穩定度分析(傳輸漏失)



相同條件（路線）下
各車機仍有顯著差異
後續族群可提早報修

29

國立交通大學運輸研究中心



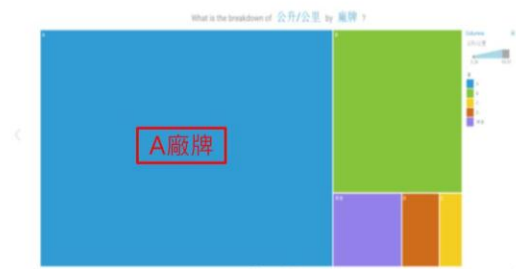
分析案例：階層式油耗分析

30

國立交通大學運輸研究中心



第一階層：廠牌



31

國立交通大學運輸研究中心



第二階層：年份 針對A廠牌下鑽探



32

國立交通大學運輸研究中心

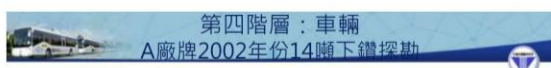


第三階層：噸數 針對A廠牌2002年份下鑽探



33

國立交通大學運輸研究中心



第四階層：車輛 A廠牌2002年份14噸下鑽探



34

國立交通大學運輸研究中心



主要資料處理課題

- 多種資料來源整合問題
 - 描述：公車動態/電子票證車牌無法完全對應
 - 原因：AB-123, AB123, AB-123_o, AB-123_oo
 - 解法：字串處理後再勾稽核，但仍有少數例外
 - 建議：制定統一標準並納入品質管理程序

35

國立交通大學運輸研究中心



- 電子票證上/下車未成對問題
 - 描述：僅有上車之資訊，無下車資訊
 - 原因：忘記刷卡
 - 解法：不納入分析（約10%）
 - 建議：未來可分析異常資料之分佈狀況

36

國立交通大學運輸研究中心



- 悠遊卡公司提供之電子票證僅有站牌代號
 - 描述：僅有站牌代號，缺乏勾稽至實際站牌位置資訊
 - 解法：本案以營運業者提供資料分析
 - 建議：應制定統一標準(含時空)並納入品質管理程序

37

國立交通大學運輸研究中心



- 電子票證缺乏轉乘代號
 - 描述：此次蒐集資料未提供轉乘(路線)欄位
 - 原因：原始欄位理應存在但未提供
 - 解法：利用「卡號識別法」來重組轉乘行為
 - 建議：未來應制定統一標準並納入品質管理程序

38

國立交通大學運輸研究中心



- 公車動態站牌資料未定義
 - 原因：站牌代碼隨路線重整持續修改，系統商僅維護最新版本，故歷史站牌代碼與最新站牌總表不一致
 - 解法：以其他專案之對應表提高對應率
 - 建議：應制定統一標準並納入品質管理程序，並留存站牌對應表異動紀錄

39

國立交通大學運輸研究中心



資料收集與處理建議

- 制定資料規格標準供遵循
 - 統一各縣市/系統/廠商之資料格式
 - 強化各資料來源間勾稽整合
- 在保護隱私前提下強化資料收集範圍
 - 電子票證其他交通行為(ex:雙鐵/腳踏車/停車付費)
 - 電子票證其他消費行為(超商購物地點)
 - 電信定位(內政部最小統計區移轉量)

40

國立交通大學運輸研究中心



相關分析工具比較

	R	Tableau	Watson	GIS	JavaScript
分析能力	●	●	●	▲	●
圖形展示	●	●	▲	▲	●
價格	●	▲	X	X	●
支援平台	●	●	▲	X	●
學習門檻	X	●	▲	▲	X

41

國立交通大學運輸研究中心

巨量資料整合平台架構圖



- 44 國立交通大學管理研究中心

建議

- 43 國立交通大學傳播研究中心

系統展示主題

45 國立交通大學通識研究中心

簡報結束
敬請指教

公車動態資訊系統巨量資料 (Big Data) 蒐集與視網化分析研究

COLLEGE OF BUSINESS
BIG DATA
ANALYTICS