

淡江大學運輸管理學系運輸科學碩士班碩士論文

指導教授：陶冶中 博士

悠遊卡大數據應用於大眾運輸乘客  
旅運型態之研究

**A Study on Applying the Big Data of Easy  
Card to Construct the Passengers' Travel  
Patterns of Public Transportation**

研究生：林浩瑋 撰

中華民國 105 年 6 月

## 謝辭

時光飛逝轉眼間研究所的生涯就這樣結束了，緊接而來的是即將面對社會所需面臨的挑戰。回顧在研究所的求學期間，不僅是需要完成所有的休業課程、學科考試以及督促我們研究成果的進度報告，這一路走來有失去也獲得了很多過往不曾體會過的經驗。

首先感謝指導教授-陶冶中老師，由大學的運輸專題課程開啟我對於運輸學門追求的熱情，並藉由老師的督促，使我從大三毅然決然以五年一貫的學制攻讀研究所，讓自己將有限的時間資源發揮至最大的效益，並在撰寫論文期間，不厭其煩地為我尋找可給予指導及資源的標的。使我在學習過程中，不僅磨練出依靠獨自思考能力，同時，不時的勉勵及激發我務必在期限內做出應有的研究成果，一路走來雖然不甚順遂但始終還是感謝老師一路的協助及支持，同時期盼自己將來在工作的職場上，能發揮陶家班的研究精神，盡心盡力工作更要大肆玩樂。

其次也要感謝陶家的傑智及簡睿學長，感謝你們在我轉寫論文的其間陪伴、鼓勵並協助我們撰寫論文過程，每當我們的簡報充滿了許多的疑問，你們仔細的觀察每個需要修正的環節，讓我們每次看似充滿危機的進度報告都能順利的將危機化為轉機，並藉由過去類似的經驗不時的鼓勵著我們別輕易放棄自己的目標。

謝謝創代科技的闕嘉宏老師，在我處理資料分析的過程中，給予我許多的技術指導，並不時的與我討論資料庫可以應用的研究範疇，十分感謝老師在百忙之中仍十分用心的抽空指導我做資料分析處理，在我即將卡關延畢的過程讓我看見準時畢業的曙光。

同時也要感謝與我共同完成兩年碩士生活的表弟兼室友，從小到大不論打電動、打球做蠢事總是有我們倆的份，就連熬夜做研究留守在學校至隔日我們都擁抱著共同應實踐的夢想，雖然我們的朋友一度都以為我們是好基友，十分感謝你

這兩年共同與我度過那些充滿樂趣及難熬的研究過程。

還記得當初大學畢業並沒有邀請親友參與我的畢業典禮，因為我承諾一年後希望他們能來參與我的碩士畢典，同時不時也成為我督促自己前進的動力，感謝在這段時間出現於我生命中的親友，不論是健身房的夥伴時常共同留守的碩班同學，讓我在面臨人生低谷時能有個抒發宣洩的窗口，並即時回神過來確認自己應當追求的標的為何。

最後要感謝我的家人爸爸、媽媽、阿姐、有緣的及無緣的女友的支持，雖然我每次都說我可能無法如期畢業，但是因著你們的信任及肯定讓我每次回家都充滿了信實的力量，已完成接下來一周所需面臨的挑戰，因為你們讓證明了自己所付諸的努力都別具意義。



林浩瑋 謹致

中華民國 105 年 6 月

論文名稱：

頁數：105

悠遊卡大數據應用於大眾運輸乘客旅運型態之研究

校系(所)組別：淡江大學

運輸管理學系

運輸科學碩士班

畢業時間及提要別：104 學年度第 2 學期

碩士學位論文提要

研究生：林浩瑋

指導教授：陶冶中 博士

論文提要內容：

近年來政府雖積極推動公共運輸相關建設，然而公共運輸市佔率比例始終不如預期。根據交通部統計處資料顯示，我國公共運輸市佔率連續三年雖呈現持續成長狀態，但每年成長幅度仍為逐年遞減，此顯示我國公共運輸服務與民眾實際之旅運需求仍有相當程度的落差，因此大多數民眾仍仰賴私人運具。

有鑑於此，本研究期望以大數據分析來探討使用者之實際旅運行為，並選定目前使用者人數最多之悠遊卡票證資料為大數據資料來源，然後瞭解國內公共運具市場提供服務之特質，再建立通勤族群之旅運特質時空間關聯模式。本研究係依照資料挖掘技術流程，依序將使用者通勤旅次長度及旅次起始時間等旅運特性資料進行分群，以此建立旅運特質之時空間關聯規則，並輔以問卷調查分析結果，進而探究民眾真實之旅運型態，以瞭解民眾對於公共運輸服務認知之缺口。

本研究經由時間分群之結果可得知多數通勤族群接受公共運輸服務之時間分布，同時亦透過通勤距離分群瞭解多數民眾之乘車旅運特性，而檢視出在若干之旅次長度或是起點下，民眾對於運輸系統的選擇偏好。根據研究結果顯示，在旅次長度落於 1 至 5 公里的區間內，民眾較偏好選擇公車；在旅次長度高於 15 公里以上，民眾則較偏好捷運。

本研究進一步進行旅運型態之問卷分析，結果得知在票價優惠、乘車次數上限及里程計費等三個乘車誘因情境，能有效改變機車族群使用公共運具之頻率，而在票價優惠的情境條件能吸引每週騎車通勤 1 至 2 次及搭乘大眾運具每週 1 至 2 次的族群；在乘車扣款上限的情境下，能吸引所得 1 萬元以下且幾乎每天騎車通勤的族群；而在里程計費的情境下則能吸引，所得 1 萬元以下、平均通勤時間為 40 至 50 分、每週搭乘大眾運具一至兩次的族群。

綜合上述，本研究透過悠遊卡電子票證乘車交易資料的挖掘，建立通勤族群對於選擇大眾運具之旅運時空間關聯規則，並改善機車族群使用公共運具之頻率。後續研究可再精進探討國內公共運輸服務缺口之分析程序，而提供兼具有效性與可行性之公共運輸服務缺口改善方案。

關鍵字：大數據、悠遊卡、資料挖掘、旅運型態

表單編號：ATRX-Q03-001-FM030-02



*Title of Thesis:*

*Total pages: 105*

***A Study on Applying the Big Data of Easy Card to Construct the Passengers' Travel Patterns of Public Transportation***

*Key word:*

*Big Data, Smart Card, Data Mining, Travel Pattern*

*Name of Institute :*

*Graduate Institute of Transportation Science, Tamkang University*

*Graduate date:*

*June 2016*

*Degree conferred:*

*Master Degree*

*Name of student: Hao-Wei Lin*

林浩瑋

*Advisor: Dr. Chi-Chung Tao*

陶冶中 博士

*Abstract:*

In recent years central and local governments have deployed public transportation systems; however, the growth rate of public transportation usage is still below the expected level. According to results of Ministry of Transportation and Statistics Department the proportion of public transportation usage has grown in recent three years but the scale of growth proportion decreases annually. It shows that a gap exists between public transportation service supply and travel demand. Therefore, most of people still rely on private transportation instead of public transportation.

In order to understand characteristics of public transportation service and to construct relevance models of travel pattern according to time and space for public transportation commuters, this study will focus on analyzing big data of EasyCard and clustering commuters by using trip length and trip time. Furthermore, a questionnaires survey is also conducted to explore travel patterns and service gap cognition of commuters in a real case.

Results show that most of commuters really travel according to trip time and trip length. It is also shown that commuters will take bus when their trip lengths are between 1 and 5 kilometers, while take MRT when their trip lengths are over 15 kilometers. The results of the questionnaires survey shows three incentive scenarios of concessionary fare, the maximum frequency of taking public transportation and the mileage calculation can effectively change usage frequencies of motorcyclists. Using the scenario of concessionary fare it will attract commuters who drive motorcycle once or twice every week or take public transportations once to twice every week. In the scenario of the maximum frequency of taking public transportation, it will attract commuters who drive motorcycle every week with income per month below 10,000 NTD. While in the scenario of mileage calculation it will attract commuters who take public transportation once or twice every week with income per month below 10,000 NTD and average travel time between 40 to 50 minutes.

In summary, this study constructs travel patterns of public transportation commuters by using big data from Easycard database. It is suggested to explore more detailed travel behavior relevance models in future.

表單編號：ATR-X-Q03-001-FM031-01



# 目錄

第一章、緒論 .....	1
1.1 研究動機及背景 .....	1
1.2 研究目的 .....	3
1.3 研究範圍 .....	4
1.4 研究架構與流程 .....	5
第二章、文獻回顧 .....	7
2.1 大眾運具及私人運具之關聯特性 .....	8
2.2.1 運具使用頻率改變之策略 .....	9
2.2 旅次回饋案例回顧 .....	11
2.2.1 香港案例 .....	12
2.2.2 新加坡案例 .....	18
2.2.3 各國城市比較小結 .....	25
2.3 大數據(Big Data) .....	26
2.3.1 大數據於交通類別的應用 .....	26
2.3.2 智慧卡票證相關應用 .....	28
2.4 資料採礦相關技術 .....	32
2.4.1 資料採礦(Data Mining)之定義 .....	33
2.4.2 資料採礦的功能與流程 .....	34
2.4.3 集群化 .....	35
2.4.4 關聯規則 .....	36



2.4.5 資料採礦之技術應用 .....	37
2.5 文獻小結及評析 .....	37
第三章 研究方法 .....	39
3.1 研究架構 .....	40
3.2.1 資料庫初步處理 .....	41
3.2.2 資料轉選及處理流程 .....	43
3.2.3 研究區域背景特性 .....	44
3.3 時距分析法 .....	45
3.3.1 時間分析-群聚分析法 .....	45
3.3.2 空間分析-資料分群 .....	47
3.3.3 關聯法則 .....	49
3.4 私人運具旅次調查 .....	52
3.4.1 問卷設計內容 .....	52
3.4.2 問卷發放對象與樣本數量 .....	53
3.4.3 問卷調查與抽樣方法 .....	56
3.4.4 問卷資料分析法 .....	58
3.5 研究假說 .....	65
3.6 資料分析工具 .....	66
3.6.1 R 語言 .....	66
3.7 小結 .....	68
第四章 實證分析-電子票證分析結果 .....	69
4.1 智慧卡資料庫資料採礦 .....	69

4.1.2 關聯規則檢定結果.....	76
4.2 小結.....	80
第五章 實證分析-私人運具旅次調查結果.....	81
5.1 敘述性統計結果.....	81
5.2.2 比較平均數法-變異數分析.....	86
5.2.3 單因子變異數分析小結.....	93
第六章 結論與建議.....	94
6.1 結論.....	94
6.2 建議.....	96
參考文獻.....	97
附件一.....	102



## 表目錄

表 2-1：世界主要都市捷運基本需求特性彙整表.....	11
表 2-2：香港人口資料彙整表.....	12
表 2-3：各車輛型態持有數量.....	19
表 2-4：新加坡大眾運輸平均每日旅次(千人).....	20
表 2-5：新加坡年度地鐵服務數據.....	20
表 2-6：新加坡 TSR 實施方案 .....	22
表 2-7：大眾運輸系統營運比較表.....	25
表 2-8：大數據於交通相關應用彙整.....	27
表 2-9：資料採礦於電子票證之相關研究彙整.....	30
表 2-10 各學者對資料採礦定義及方法.....	33
表 3-1：智慧卡欄位明細.....	42
表 3-2：續智慧卡欄位明細.....	42
表 3-3：資料前置處理結果.....	44
表 3-4：公共運具轉乘運量統計.....	45
表 3-5：時段區間分群表.....	46
表 3-6：發放手法比較表.....	54
表 3-7：抽樣對項統計表.....	57
表 3-8：多重比較法優缺點比較.....	64
表 4-1：平假日各族群間離峰時段使用比例.....	69
表 4-2：捷運族群旅次長度分佈表.....	71

表 4-3：公車族群旅次長度分佈表.....	71
表 4-4：捷運站點分群表.....	72
表 4-5：公車站點刪除項目.....	73
表 4-6：公車站點分群 1.....	74
表 4-7：公車站點分群 2.....	75
表 4-8：公車站點分群 3.....	76
表 4-9：捷運族群旅次關聯特質.....	77
表 4-10：通勤旅次量表.....	77
表 4-11：公車族群通勤旅次關聯規則.....	78
表 4-12：大眾運具族群通勤旅次關聯式.....	79
表 5-1：基本敘述性統計-性別.....	81
表 5-2：基本敘述性統計-年齡.....	82
表 5-3：基本敘述性統計-平均月收入.....	82
表 5-4：基本敘述性統計-家中機車數.....	83
表 5-5：基本敘述性統計-居住狀況.....	83
表 5-6：通勤旅次調查統計-每周騎車通勤數.....	84
表 5-7：通勤旅次調查統計-每日通勤時間.....	84
表 5-8：通勤旅次調查統計-通勤節省時間.....	85
表 5-9：通勤旅次調查統計-搭乘公共運輸頻率.....	85
表 5-10：變異數同值檢定表-收入水準-研擬方案.....	86
表 5-11：單因子變異數分析表-收入水準-研擬方案.....	86
表 5-12：多重比較分析表-收入水準-研擬方案.....	87

表 5-13：偏好比較分析-收入水準-研擬方案 .....	87
表 5-14：變異數同值檢定表-通勤次數-研擬方案 .....	88
表 5-15：單因子變異數分析表-通勤次數-研擬方案 .....	88
表 5-16：多重比較分析表-通勤次數-研擬方案 .....	89
表 5-17：偏好比較分析-通勤次數-研擬方案 .....	89
表 5-18：變異數同質性檢定-平均通勤時間-研擬方案 .....	90
表 5-19：單因子變異數分析-平均通勤時間-研擬方案 .....	90
表 5-20：多重比較分析表-平均通勤時間-研擬方案 .....	90
表 5-21：偏好比較分析-通勤次數-研擬方案 .....	91
表 5-22：變異數同質性檢定-搭乘頻率-研擬方案 .....	91
表 5-23：單因子變異數分析-搭乘頻率-研擬方案 .....	92
表 5-24：多重比較分析表-搭乘頻率-研擬方案 .....	92
表 5-25：多重比較分析表-搭乘頻率-研擬方案 .....	93
表 5-26：研擬方案比較表 .....	93

## 圖目錄

圖 1-1：台灣各縣市公共運輸使用率.....	1
圖 2-1：乘車依賴度級別.....	23
圖 3-1：研究架構圖.....	40
圖 3-2：智慧卡原始資料欄位.....	41
圖 3-3：資料前置處理流程.....	41
圖 3-4：智慧卡原始資料集交易欄位.....	43
圖 3-5：智慧卡預處理後交易欄位.....	43
圖 3-6：旅次變化趨勢圖.....	46
圖 3-7：K-means 演算流程.....	48
圖 3-8：站距分群換算示意.....	49
圖 3-9 Apriori 演算法運作架構.....	50
圖 3-10 數據挖掘程式語言市占調查結果.....	67
圖 4-1：時距分析趨勢圖-平日.....	70
圖 4-2：時距分析趨勢圖-假日.....	70

# 第一章、緒論

## 1.1 研究動機及背景

近年來隨著經濟發展之成長及國民所得的提高，國內汽機車持有率逐年上升，根據交通部交通統計資料庫結果(圖 1-1 所示)，過去 30 年間國內自小客車持有數成長了近 30 餘倍，機車持有數則成長了近 6 倍之多。在近年政府極力推動大眾運輸建設及政策下，其公共運輸之市占率已連續三年有逐漸成長之跡象，就縣市別觀之，以臺北市 37% 為首，其次為基隆市 32.4% 及新北市 31.1%，綜合觀察 20 縣市之公共運輸市占率，總計 5 個縣市達 10% 以上、10 個縣市公共運輸市占率創新高，顯示在政府極力推廣大眾運輸的政策下，使得民眾的乘車偏好已有移轉的現象，但在有限的道路資源下，還是無法有效分配充足的道路空間於每個用路人上，主要還是歸因於多數的民眾仍偏好高機動特質的私人運具，相較於如新加坡及香港等高度公共運輸使用率之國家，反觀國內僅近四成的公共運具市占率仍有很大的成長空間。

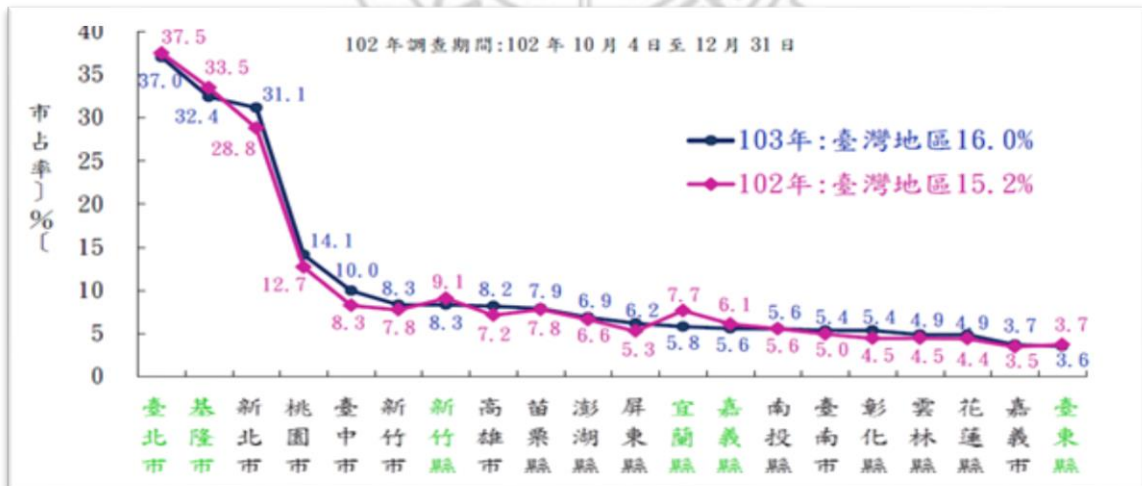


圖 1-1：台灣各縣市公共運輸使用率

資料來源：103 民眾日常使用運具狀況調查

另根據交通部統計處 103 年機車調查資料顯示，私人運具族群中有近五成之比例為使用機車之民眾，但有近 30% 之民眾具備兼用其他運輸系統之習慣，其最主要轉搭乘之交通工具為公共運輸工具（含市區公車、長程客運車、火車、交通車、捷運及高鐵）者占 87.4%。轉搭乘「捷運」及「市區公車」者呈增加趨勢，而轉搭乘「自用小客車」者則遞減，顯示提升公共運輸工具之便利性及普及性，有助抑制私人運具之成長。而在機車族群對於公共運具的使用觀點上，有 34.4% 之機車使用者表示，倘若政府實施機車管制措施或增加公共運輸服務之可及性，則有意願改用使用頻率（張新立 1995）。結果顯示在政府極力推動公共運具建設之條件下，能有效提高機車族群選搭公共運具的意願，相較於消極的管制措施而言，強化公共運具之便利性及普及性，將可提高民眾搭乘公共運具之意願，減緩私人機動運具之使用（交通部統計處 2014）。

由於淡水位於新北市郊區之故，其當地聯外交通除捷運系統外，僅有一條主要道路得以使用，以至於每逢尖峰時段主要的聯外道路時常產生道路壅塞之現象，主要的原因在於多數的通勤者選擇私人運具做為代步工具，以至於在有限的道路資源下，無法將其道路空間有效的分配到每位道路使用者上，進而增加每位通勤者的旅行時間及旅行成本，同時也產生了更多社會成本之耗損造成無謂損失。

為了提升公共運輸之使用率，及有效減緩尖峰時段所產生之交通壅塞現象，近年來可看見國外的交通單位已逐漸從硬體設施的建置上轉變成軟體平台的行銷及規劃，以智慧卡系統為例，其具備交易資料之儲存性及連續性兩種特色，在交易紀錄的同時亦能瞭解到使用者的使用區域若以公共運輸的應用為例，電子票證包含：選用運具別、旅次起點、旅次迄點、旅次起迄時間點及旅次鏈長度等相關資訊，可反應出公共運輸及都會空時空間變化情形。自 2000 年悠遊卡發行至今，其累積發卡量已突破 5,000 萬張，平均每日總交易紀錄逾 550 萬筆，其中在 2014 年台北市公車交易紀錄則達至近 200 萬筆之數量。有鑑於近來國內致力推動於如 NFC 及電子貨幣交易的趨勢興起，智慧卡的相關研究多涉獵於產品行銷、



廣告推廣等領域，同時也能看見應用於如公車站點設置及車輛發派等相關研究，由此我們可得知智慧卡資料之應用類別十分廣泛，故本研究期望透過電子票證之交易資料所累積龐大資料庫，應用大數據(Big Data)分析方法，以瞭解淡水區中學生族群之運具使用行為及旅次特性，並檢視在該區域時間面及空間面的變異。

故本研究欲探究國內通勤族群旅運特質，由於多數的私人運具族群皆有搭乘公共運輸之經驗後則不在偏好該運輸系統，故本研究針對國內北部電子票證系統進行分析，結合資料挖掘之技術探討民眾搭乘公共運輸之偏好，並對其輸出結果進行推論，其鑒於民眾乘車偏好眾多，在將乘客之起迄點資料、運具種類、旅次長度、乘車時間點等類別作為資料庫建置的基礎，以建立公共運輸族群旅次時空間關聯規則，檢視公車旅次及捷運旅次所提供服務之特質，並透過乘客運具偏好的問卷設計，歸納私人運具族群對於公共運輸服務偏好之差異，並期盼以情境模擬的手法檢視在何種特質的情境下，能增進私人運具族群對公共運具使用頻率。

## 1.2 研究目的

基於上述背景及動機，本研究將針對國內北部電子票證系統進行探討，本研究之目的如下列幾項：

- 一、以大數據分析電子票證資料庫，從大量的乘車票證資料當中探討日常通勤行為，及瞭解通勤旅次特質與公共運輸之偏好。
- 二、以公共運具通勤族為目標族群，瞭解使用者旅運型態特質，及對於時間及空間的關聯特性。
- 三、回顧國外公共運輸鼓勵政策，歸納誘因特質做為情境模擬理論之依據，探究通勤族群在何種情境誘因下，具備增進使用公共運輸頻率之特質，做為後續研究之參考依據。

### 1.3 研究範圍

淡水區因鄰近臺北市區，加以交通較為便利，為當地區帶來豐沛人潮與商機，民眾除可飽覽淡水老街風光與古蹟外，亦對陸域以外的遊程拓展增加誘因，使得淡水區深具休憩娛樂潛力，也因此每逢假日即吸引近二十萬的人潮湧入淡水，促使淡水成為商業服務與旅客聚集之集散地；以捷運淡水站為例，其周邊附有完善的教育學區之規劃，以大專院校為例其境內就包含了：淡江大學、真理大學、聖約翰科技大學及台北海洋技術學院，亦為北台灣大專院校分布最密集的行政分區之一。

根據捷運公司所提供 103 年捷運淡水站、紅樹林站及竹圍站之進站人數，與交通部統計處 103 年民眾日常使用運具狀況調查書中各公共運具市占率表格推估，可知隸屬淡水行政區內之三個捷運站，包含淡水站、紅樹林站及竹圍站，此三站之平均每日進出站旅客量人次為五萬七千餘人，約占新北市捷運總搭乘人次的 11.56%，行政院於 102 年 4 月 8 日核定辦理「修訂淡海新市鎮開發執行計畫」，淡水區因新市鎮持續開發而衍生更多旅運需求。然而，現有道路容量不足狀況恐將加速交通壅塞情形，幾經探討淡北路廊相關改善計畫，包括「台 2 線拓寬」、「淡水河北側平面道路計畫」、「淡海輕軌運輸系統」、「淡江大橋及其聯絡道計畫」等各改善方案間存在競合關係。經綜合分析結果顯示，為因應淡海新市鎮開發後之區域聯外運輸，「淡海輕軌運輸系統」建置計畫為最有效與首要進程之方案，主要原因為未來淡海輕軌運輸系統建置可適度提升淡水地區私人運具轉移使用公共運輸比率，更兼具政府致力推廣搭乘公共運輸之政策目標，故依此發展進程，若配合公共運輸旅次回饋機制使用，將使得淡水區在未來擁有更高程度之公共運輸發展潛力。

## 1.4 研究架構與流程

### 一、研究背景及動機目的

本研究在參閱 103 民眾日常使用運具狀況調查研究後，發現近年來大眾運輸之車站率逐年成長之趨勢漸於趨緩，並在國外相關案例佐證之情況下，探討國外政府以旅次回饋及乘車獎勵等方式推廣大眾運輸，進而構成本研究欲探討之研究背景，假設以同樣手法(類似機制)是否亦能增進通勤族群選搭大眾運輸的頻率。

### 二、現況分析及文獻回顧

針對上述研究標的進行現況分析及相關類別文獻彙整及評析，呼應其研究目的之可行性，以做為研究實證性之理論依據。其文獻回顧內容包含：1.旅次回饋發展現況及相關案例；2.大數據資料特性及應用；3.資料採礦技術等類型，在彙整其文獻評析及小結。

### 三、研究方法與問卷設計

本研究欲探勘大眾運輸乘客資料庫，分析其運具選擇偏好、旅次特性及相異運具間的使用規則，特過文獻歸納出運具偏好因素以分析相關影響類別，在資料採礦的研究設計上，本研究使用 MSSQL 對電子票證資料庫進行資料清理及轉置，在資料分群分析工具則使用 R 語言進行整合與處理，最後應用關聯規則演算法建構公車族群及捷運族群旅次時空間關聯規則，在問卷設計則參考旅次回饋的相關案例做為情境模擬的設計基礎，探討在何種情境特質下通勤族群能有增進使用公共運輸的頻率產生。

### 四、資料採礦及時空關聯建立

分別對公車族群及捷運族群之旅次資料依照旅次長度及旅次起始時間進行分群，再以關聯規則演算法選取其可信度較高之條件式，以建構公共運輸通勤族群之時空間關聯式，歸納公共運輸族群之旅行時間、旅行成本及旅次長度等相關偏好。

## 五、私人運具旅次特質偏好調查

藉由問卷分析結果觀察私人運具通勤族群對於公共運輸服務之偏好，及瞭解在何種情境下會有增加使用公共運輸之頻率，並藉由旅次回饋情境模擬之方式詢問，其私人運具通勤族群對於各情境條件之偏好及選擇意願。

## 六、結論及建議

根據分析及結果，統整研究結論及後續相關作業事項，以歸納其國內公共運具通勤族群之旅運旅特質為何，並藉由私人運具偏好調查之結果，觀察使用者易受到何種乘車補助誘因而影響使用公共運具頻率之意願，以做為後續之研究範疇及值得探索的目標族群。

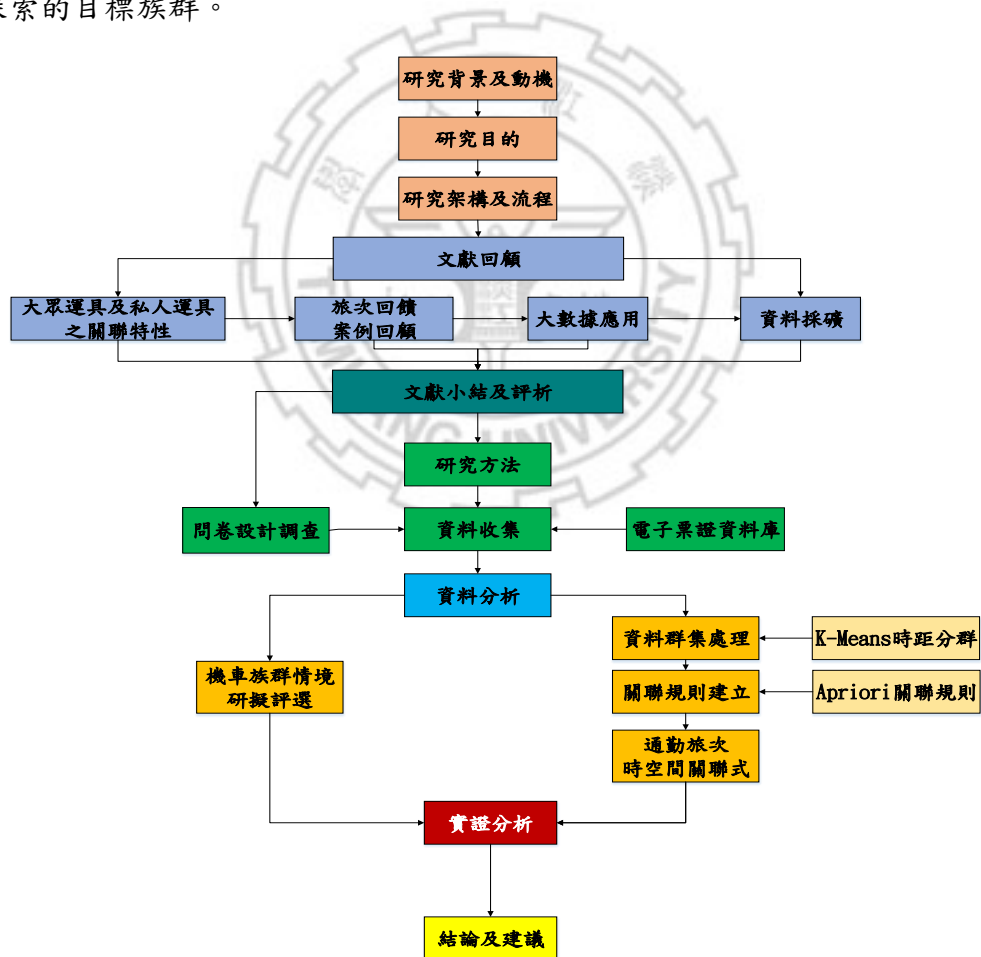


圖 1-2 研究架構圖

## 第二章、文獻回顧

近來大數據的議題正在世界延燒，其大數據擁有大量、多元、即時且真實等資料特性，目前透過手機載具、電子票證載具、刷卡交易購物皆有紀錄，透過這些紀錄可確實掌握民眾消費及旅運相關行為。倘利用這些資訊做為運輸基礎，更可以讓大眾運輸資訊運用更加有效率，亦可作為成本及風險相關評估之依據。

根據交通部統計處所提供資訊顯示出，其國內約有三成的機車使用族群擁有兼用其它運具之偏好，其多數為兼用公共運具之使用族群。國去研究顯示，政府應瞭解民眾對於旅運需求之偏好，提出相關配套於公共運輸服務之措施以抑制私人運具使用之成長(張新立1995)，目前國內外已逐漸從政令面之約束，逐漸轉化成以獎勵、鼓勵等方式推廣大眾運輸之市占率，以票價補貼之層面來看，一般來說分為：票價補貼及乘車次數補貼兩種優惠方式，前者優惠措施所推廣範圍十分廣泛，包含：相異族群差別取價(如：一般票、學童票、優待票等)，異質運具轉乘優惠(如：捷運系統轉乘公車系統)，早鳥乘車優惠等類型，後者優惠則是以乘車次數做為推廣之範疇，包含：一日乘車卷，月定乘車卷等類型，回顧如新加坡、澳洲及香港等高公共運輸市佔率等國下來說，近來以設計如乘車點數、車資支付上限等優惠，鼓勵民眾多選擇大眾運輸作為外出代步工具。

本研究係以探究私人運具族群及公共運輸之關聯特性，故從私人運具使用普及的狀況做為此主題的切入角度，期望基於私人運具使用的立場，以之討論相關的影響面向和特質，進而探討轉用行為如何可能。因此，本節針對相關文獻分為：大眾運具及私人運具之關聯特性、旅次回饋案例回顧、大數據應用、資料採礦等四部分進行探討。

## 2.1 大眾運具及私人運具之關聯特性

Kenworthy J. R (1999)探討世界各國主要城市的研究中，其結果顯示出私人運具的持有及當地居民之貧富並無顯著關聯，但與城市周邊之土地使用類別及密度型態有顯著之影響，而相較於以公共運輸為規劃導向之城市當中，其多數擁有高私人運具比例之城市中，在公共運輸營運回收之比例也較低，且在公共運輸之建置及維修成本的預算上則較為貧乏，故建議當地政府應降低都市中土地使用之密度，增進公共建設及基礎設施之投資並推廣公共運具之使用意象，以降低私人運具使用之比例。

Cullinane.S (2001)研究大眾運輸是否會影響到私人擁有汽車與使用的意願，以當地大學生為例，研究發現香港的大眾運輸工具，在費率既便宜且班次又充足下，汽車的擁有與使用情形，都有減緩的現象。由此可見，在良好大眾運輸工具的提供下，確實可以改變私人擁有汽車的意願。

徐瑞彬(2001)研究臺北都會區總體旅運型態，顯示男女使用運具的特性不同，也發現高所得的族群使用汽車的比例也越高，而大眾運輸工具是以低所得的族群使用率較高，可看出家戶所得的影響情形；另發現家戶運具持有率與運具次數關係顯著，小汽車持有數越高，則旅次發生比率也越高、大眾運輸使用率也較低。

陳鴻文(2002)研究顯示家戶內持有之汽、機車數為一競爭關係，且各種家戶型態持有汽、機車數量為不盡相同的類型，大眾運輸也確實能減少家戶持有機車數，因此政府應持續發展大眾運輸，但應配合機車管理策略才能減少機車使用量，讓家戶放棄持有機車，有效減少機車數量。

張正松(2002)表示小汽車轉乘停車場供過於求，平日容量閒置，假日則鼓勵民眾開車不坐捷運，與鼓勵大眾運輸目標矛盾，基於資源有效利用觀點，應將剩餘空間妥善規劃轉供其他轉乘運具，如腳踏車、機車等，以減少衍生的社會成本。

然而這來自於小汽車使用的特性，Moller and Thøgersen (2008)認為駕駛人使用小汽車的行為多具有習慣的特性，這主張來自於習慣性所形成的日常旅次，更會不斷在固定的環境脈絡中被重複的展現，此旅運特質也確實有益於個人在此日常生活環境下的需求，因此，習慣性的行為決策常自動產生，而且只有一種可供選擇的決定方案做為行為的指引，在此情況下，習慣常被認為難以改變，不過此課題也開展許多相關研究的討論和探究。

### 2.2.1 運具使用頻率改變之策略

其頻率改變策略多數分為兩種，第一種為政令面之影響，如透過建立大眾運輸系統、廣告行銷、道路縮減、道路封閉、票價優惠、課徵用路稅等(Fujii and Kitamura 2003)，不過此策略具有長久性和暫時性措施的討論，長久性策略例如減少道路容量即須考量交通工程的重新規劃和環境衝擊的配套擬定；用路稅的課徵則可能帶來龐大的社會衝擊和調適的壓力；大眾運輸系統的建置更是一筆可觀的預算支出，長久性策略均為龐大的政策成本，因此，暫時性的介入策略被提出討論，期望透過相對簡易經濟的手段予以改變既有運具使用習慣，如同Fujii and Kitamura (2003)透過一個月免費乘車票券的實驗，領有票券的實驗組提升20%的大眾運輸使用率。

第二種策略則訴諸直接對個人行為改變的勸說和引導，實際的作法可透過對社會組織(學校、社區、家庭等)的教育和行動計畫予以落實，此種策略的探討於(Fujii and Taniguchi 2005)的實驗被稱之旅運回饋計畫(travel feedback program)，同樣以減少家庭汽車使用並鼓勵大眾運輸和步行的行為改變目標，該研究期望藉此評估受訪者行為計畫的促發和認知的過程，並獲得有效的介入方式或考量，將樣本分群，其一透過口頭建議和資訊提供來引發家戶和個人減少小汽車使用的行為目標；其二則直接要求實驗的家戶和個人發展減少小汽車使用的具體行為計畫，

該研究顯示發展行為計畫的組別確實反映了他們減少使用汽車的程度，於全部旅次中估計減少了27.7%使用汽車的比例；而提供資訊和建議的組別則沒有產生顯著改善的介入效果。因此，(Fujii and Taniguchi 2005)認為內在的行為介入策略必須重視執行意圖對行為落實的重要影響，執行意圖為個人形成某一確實會執行的行動計畫，這對於實際行為是必需的，然而提供意見可能對個人執行意圖的發生僅有間接的影響，但是鼓勵個人發展行動計畫提供行為落實更有力且直接的執行動機。

鑑於以上兩類運具使用習慣改變策略的發展，本研究認為他們均可為大眾運輸促進的手段，但是必須更注意的，在於不同策略所實施的社會環境，這可能直接導致某些策略的成效與否，如在大眾運輸水準已經不錯的地方提供使用資訊和優惠的票價做為促進策略，即會有其實質的增進效果，但若於大眾運輸水準不佳且私人運具使用率高的環境下實施一樣的策略方案，其結果可能是不如預期的，此強調的，在於不同大眾運輸服務水準的環境下，其社會生活自然會產生特定的旅運行為，此即直接造成某項運具被選擇使用。



## 2.2 旅次回饋案例回顧

旅次回饋最初的概念來自於信用卡公司的促銷方案，藉由在消費的過程中累積一定的回饋金，其反應在航空公司的飛行里程數上，其後續的相關應用包含了旅行社的促銷方案，以增進消費者消費之意願，以大眾運輸的相關應用範疇上，近幾年包含：新加坡、澳洲及香港等國家皆採用藉由搭乘公共運輸來給予相當程度的旅次回饋獎勵，其實施上亦獲得相當程度的績效以創造高市占率之公共運具使用比例，下列彙整上述三種國家之回饋補貼案例陳列如下：

台灣目前有捷運的都市為台北都會區，根據都市及區域計畫統計彙編數據，截至 2015 年雙北市(台北、新北)人口 659 萬餘人，面積為 2,052.57 平方公里，人口密度為每平方公里 3216 人，而新加坡人口密度為每平方公里 6,389 人、香港的土地面積為 1,099 平方公里，密度為每平方公里 6,096 人。由此可見，香港、新加坡與台北都會區同為地狹人稠的地區，且捷運路網規模較為相近（詳見表 3-1），因此選取香港與新加坡作為本研究參考的案例對象。

表 2-1：世界主要都市捷運基本需求特性彙整表

都市	人口數(萬)	路線長度(公里)	路線數	每日旅次量	捷運市佔率
台北	659.6	131.2	5	196 萬	29.69%
香港	718.8	177.4	7	462 萬	64.25%
新加坡	547.9	148.8	5	265 萬	48.35%
巴黎	224.4	220.4	16	130 萬	58.03%

## 2.2.1 香港案例

香港從 1990 年至 2000 年汽車的數量成長了 68%，但汽車持有率仍非常低，每千人僅持有 49 輛車，同時期英國（UK）每千人持有 377 輛車、新加坡（Singapore）為 120 輛車，日本（Japan）為 325 輛車、美國（USA）為 485 輛車（Cullinane, 2003a），於 2005 年香港每千人僅持有 50 輛車（詳見表）。以車輛數目而言，每公里道路約有 275 輛領有牌照車輛（香港運輸署，2015）。

表 2-2：香港人口資料彙整表

年代	人口	私有車輛數	大眾運輸旅次(千)/每人
2005	6,965,900	351,000	5,849
2012	7,241,700.0	494,646	12,079
2013	7,305,700.0	517,997	12,350
2014	7,324,300.0	541,751	12,519

資料來源：香港運輸署

### （一）大眾運輸票價

大眾運輸票價非常低，也可以說使用者價格非常的低，儘管大眾運輸經營者沒有接受任何補助，卻有盈利的。香港地鐵的費用範圍從港幣 4 元到港幣 26 元。九廣鐵路（KCR）約 30 公里最大的旅次成本為港幣 31 元，大部分票價成本低於港幣 6 元，針對地區旅次的費用從港幣 3.5 元到港幣 9 元，到中國邊境要港幣 33 元。而都會旅次公車大致費用介於港幣 2.4 元至港幣 11 元，一直到新界（New Territories）最遠的點的旅次為港幣 34 元。依據(Gutmann et.al 2000)，在香港約 10 公里的大眾運輸旅次成本比倫敦（London）或東京（Tokyo）便宜，但比巴黎（Paris）或新加坡（Singapore）貴。

香港地鐵收費分成人票及特惠票，3 歲以下的小孩可免費乘搭，12 歲以下的小孩、65 歲或以上的長者、12 至 25 歲的全日制學生可使用特惠票。不同於一些地鐵系統，香港地鐵收費並非劃一，而是根據路程長短而定。所乘搭的站數越多，收費就會越高，一般路線的成人單程收費由港幣 3.8 元至港幣 23.1 元不等，特惠票價約為成人票價之一半。

付款方法共有四種：八達通、單程車票、旅客車票及紀念車票。單乘票為購票當日使用。八達通為一種通用磁卡，分 70、100、200 元，此較便利且有優惠可圖，搭乘多次不同車費的車程使用，並依據不同票面值多送 3-12 元票值，例如票價為 200 元的儲值卡實際票值是 212 元，亦可享有尾程優惠，縱使磁卡最後只剩一元，仍可繼續使用，並可乘車到最遠的地方，出站時驗票機會將磁卡自動收回，還可用於地鐵車站內的店鋪消費使用。又為了鼓勵更多的人搭乘地鐵，乘末班車和非高峰車時可以享受到打折優惠。

## （二）大眾運輸智慧卡—Octopus

八達通是香港通用的電子收費系統。晶片內置在信用卡大小的塑膠卡片，替卡片充值後放在接收器上即能完成付款過程。八達通在 1997 年 9 月 1 日開始使用，最初只應用在巴士、鐵路等公共運輸工具上，後來陸續擴展至其他行業，包括商店、食肆、停車場等業務，也用作學校、辦公室和住所的通行卡。充值的方法也由最初的充值機，擴展至商店付款處和以信用卡、銀行戶口自動轉賬。

八達通是全世界最早也是最成功的電子貨幣之一，普及程度也是全世界最高。截至 2005 年，香港總共流通 1240 萬張八達通，每日交易數量超過 800 萬。八達通卡（Octopus Card）可以使用於許多大眾運輸（所有的地鐵服務、所有的九廣鐵路服務、所有的主要渡輪服務與約 70% 的公車服務），為可加值的、非接觸式、塑膠製的電子智慧卡。八達通卡甚至可在鐵路系統裡的電話或照相館中使用，以及鐵路系統之外的一些超市與咖啡店。事實上，卡片本身是免費，可歸還已付

的保證金，並可以在任何時候在任何鐵道系統的車站以及其他方便的區位加值。八達通卡的使用簡單，有效鼓勵大眾運輸的使用，並改善了整合的容易性，購買票完全不需浪費時間。在尖峰時間與主要的轉運區位，感應票機能夠使民眾在每一站不論從前門或後門皆可搭上公車（Both Front and Middle Doors），而這也支撐了 White（1995）的觀點，其提出強調在 London 旅運卡（Travel-card）能保有大眾運輸乘客。

洪浩凱指出 1992 年香港大眾運輸鐵路公司（MRTC）針對非接觸智慧卡（CSC）當作車票使用進行可行性研究，研究結果可帶來相當的利益、可減少維修的勞工及材料成本、節省約 20% 的自動化收費維修編制。香港大眾運輸鐵路公司主管看到節省成本及方便乘客的種種好處之後，決定全力推動。而究其能帶來的營業效益而言，可從乘客和公司方面來看（夏明橋，1999）。

在乘客方面：

- （1）以一套共通票證系統進出各種運輸工具，提供經常使用的乘客最大的便利。
- （2）加速登車時間
- （3）工作人員有更多的時間處理顧客與票務無關的服務，如改善乘客服務品質。
- （4）以通用的儲值智慧卡讓各式大眾運輸方式之間轉乘更加容易
- （5）提升售票機的性能，並擴大服務業者網路，改善售票口的親和力
- （6）提升殘障人士與老人對系統的使用容易度
- （7）卡片交易可靠性更高

在業者方面：

- (1) 增加收入，並改善收款作業
- (2) 透過新的安全措施，維護收入安全
- (3) 因為聯運轉乘方便，因此增加大眾運輸的使用量
- (4) 交易時間更快，交易可靠性更高
- (5) 資訊供應量更完整，提供服務業者運輸工具的管理效率
- (6) 顧客接受意願高
- (7) 排班經過改善，減少業者成本
- (8) 加強查核系統及其他安全措施，改善服務業者資產與顧客的安全
- (9) 有效的改善管理及行政結構
- (10) 改善大眾運輸的形象
- (11) 目前使用的智慧卡，彈性變化大，可以應付未來各種應用需求

### (三) 利潤掛鉤車費優惠計劃

利潤掛鉤的車費優惠計劃為香港地鐵公司近年推廣之搭乘優惠，其根據當年及前年之營業盈餘，撥出 0.95% 至 1% 之利潤做為該年度之公共運輸補貼費用，以 2014 年為例，該年度香港地鐵公司之營利收入為 130 億元，按照與利潤掛鉤的車費優惠計劃，將要撥出一億二千五百萬元費用予以補貼，並配套著當日回程車資九折之優惠回饋乘客。

票價優惠機制內容簡述如下：

- 早尖峰乘車優惠：

為紓緩尖峰時段所造成之壅塞現象，香港地鐵公司於 2014 年九月推出之早尖峰乘車優惠試驗計畫，與新加坡 TSR 所推廣的方式相仿，主要是針對平日早上 7 點 15 分至 8 點 15 分間乘車民眾，總計共 29 個乘車及轉乘點，搭乘公共運輸即可享有 75 折之優惠。

- 都會票優惠：

港鐵公司於 2014 年 6 月推陳之方案，其主要提供乘客 30 日內，持有人可在購票當日起計 30 日內，乘港鐵來往香港島、九龍、葵青、荃灣及將軍澳區之間車程 40 次，僅需花費 400 圓之票價；以現時票價設計，乘客可享 7% 至 21% 折扣，調整後價格則可享 10% 至 24% 的車費折扣，港鐵估計，有 40% 在該範圍乘車的乘客可享有關優惠。

- 同車折返九折優惠：

結合利潤掛鉤的車費優惠計劃及服務表現安排之優惠，所有持八達通的乘客，只要於同一日內付費乘搭港鐵、輕軌、港鐵巴士車，同類交通工具回程均可享有九折優惠。

#### (四) 就業據點與接駁路網之整合

香港的公共運輸系統是以地鐵為基礎模式伴隨著地鐵的擴展，自 1999 年起地下鐵系統儼然成為當地居民外出代步的運具選擇，並隨著鐵路支線及服務站點的擴展，配合著政府與民間企業的合作，使當地絕大多數的居民就業範圍周邊一公里內皆有附設公共運輸的服務據點，以培養就業人有擁有搭乘公共運輸系統通勤的習慣。

當地政府在管制私有運具持有及土地使用類型上扮演其重要之角色，包含興建完善且優良的公共運輸系統並實施其股利乘車之優惠方案，並同時限制私有車輛之持有及使用。假如大眾運輸是一項綜合的、頻率高的、整合度夠的、高品質且夠便宜，則將會鼓勵民眾使用，這會使得汽車持有不是一種不可避免的慾望與期待 Cullinane( 2002 )。

由此可證，公共運輸所提供之服務為民眾選擇運具的標的之一，倘若能讓使用者感受到不亞於私有運具旅次之乘車成本及旅行銷率等考量因素，即可創造高公共運具市佔率之可能性。相較於其它各國城市香港政府選擇以軟硬兼施的方式管理其境內交通環境，並碩造當地通勤民眾對於公共運具擁有高度之依賴特性，以有效管理當地尖峰時段所產生之交通壅塞及道路事故等問題。



## 2.2.2 新加坡案例

### 一、新加坡與運輸

新加坡是一個小的、人口密度高、總人口約為 540 餘萬的城邦 (City-state)。它由一個主要的新加坡島以及 63 個群帶島嶼組成，主要的新加坡島面積為 648 平方公里，長約 42 公里，寬約 23 公里，伴隨著土地的供給受到限制與突增的交通需求，增加大眾運輸的使用似乎是解決未來交通問題的方法。新加坡政府組織出重要的、有效率、整合的土地交通網絡，能滿足民眾的需求與期望，並提供經濟與環境方面的目標。為了達到此目標，政府採取複合式方式 (Land Transport Authority, 1996)，這牽涉到：

1. 整合土地使用與交通計畫，以達到旅運需求最小化。
2. 擴展道路路網，容量最大化。
3. 城鎮與交通計畫。
4. 管理車輛持有與道路使用的需求，以緩和交通擁擠。
5. 改善大眾運輸系統。

新加坡陸路管理局 (Land Transport Authority, LTA) 為交通部 (Ministry of Transport) 底下法治的管理委員會，領導新加坡所有的土地運輸開發，它負責規劃、開發與管理土地運輸系統，企圖達到一個世界級的土地運輸系統 (a World Class Land Transport System)，它大致朝向提供一個品質的、整合的與有效率的土地運輸系統，以滿足人民的需求與期望，提供經濟與環境目標，與提供金錢價值 (LTA, 1998)。在朝向這些目標努力時，新加坡陸路管理局 (LTA) 與其他政府機構密切接觸以確認運輸系統有良善的規劃，並與都市開發完全整合。新加坡



積極改善它的整合運輸政策，所有的開發，不論是主要新市鎮或是購物中心，皆為運輸與土地使用整合的要素。然而，汽機車持有與使用及大眾運輸的議題，總是成為每個居民考量的核心，近年來，新加坡的汽機車持有數輛呈現成長趨勢(詳見表)。新加坡的地狹人稠否定了擴展道路路網的想法，因而朝向提升大眾運輸的效用。因此，當管控汽機車的持有與使用的同時，政府也企圖去改善大眾運輸系統，使得大眾運輸成為更具吸引力的私人汽車替代工具。政府朝向提供居民一個範圍更廣的整合性運輸選擇與一個提供高品質服務、便利的、具可及性、舒適的、安全的、快速的及對大部分民眾是可負擔的運輸系統。根據新加坡交通部，其採取嚴謹的運輸規劃，並儘可能地在問題尚未達到無法管理時前實施預防的步驟，運輸政策的結果雖然不能稱之為完美，但比起其他的城市，卻更成功地達到對民眾更有效率與更高的易行性，使民眾可以達到從一地到另一地無縫隙整合串連的通勤。

表 2-3：各車輛型態持有數量

年度	汽車	租車	計程車	公車	貨車	機車
2011	520,614	85,666	27,051	17,046	159,768	146,559
2012	535,233	84,778	28,210	17,162	160,417	144,110
2013	540,063	83,625	27,695	17,509	160,344	144,934
2014	536,882	82,141	28,736	17,554	161,698	145,026

資料來源：新加坡交通部

## 二、大眾運輸

在新加坡主要的大眾運輸服務包括：公車、計程車、輕軌運輸（LRT）與地鐵（MRT）。這幾年持續地努力改善大眾運輸的品質，並保持大眾運輸為可負擔的，使大眾運輸成為可取代汽車且吸引人的替代方案，其所做的改善不只運輸模式本身，也涉及改善所有媒介與端點設施，例如連接道（Linkway）、服務資訊、顧客服務等。在新加坡約有八成的旅次是搭乘大眾運輸（新加坡交通部 2014）。

表 2-4：新加坡大眾運輸平均每日旅次(千人)

年度	MRT		LRT		BUS		Taxi	
2011	2295	0.34	111	0.02	3385	0.5	933	0.14
2012	2525	0.35	124	0.02	3481	0.49	967	0.14
2013	2623	0.36	132	0.02	3601	0.49	967	0.13
2014	2762	0.36	137	0.03	3751	0.49	1020	0.13

資料來源：新加坡交通部

表 2-5：新加坡年度地鐵服務數據

項目	2011	2012	2013	2014
路線長度(公里)	89.4	89.4	89.4	89.4
總載客量(百萬)	394.25	391.46	402.59	413.77
旅客成長率(%)	1.18	0.71	2.84	2.78
平均日旅次量(千)	1139.3	1131.4	1171.9	1212.2
平日旅次成長(%)	1.27	0.69	3.58	3.45
年度總旅次量(百萬)	4988.9	4861.3	4928.3	5058.3
旅行距離成長率(%)	0.52	2.56	1.38	2.64

資料來源：新加坡交通部

新加坡地鐵成為大眾運輸系統的脊椎，路徑規劃行經新加坡人口最密集的區域，它所服務的不僅是連接許多衛星城鎮，也連接許多住宅區、市中心與工業區。現今，每日超過百萬乘客旅次搭乘地鐵，相較於 1987 年增加了 10 倍，超過 60 % 的新加坡人搭乘地鐵上班、回家與旅遊，許多設施改善地鐵的安全、便利與通勤友善：(1) 建造有遮蔽的步道來連接地鐵站與建物之間，以提供更流暢的旅次與轉乘，並整合運具模式，包括：計程車停等區、乘客上下車、腳踏車停放位置等也在地鐵站外提供；(2) 播放電影預告、運動消息與音樂新聞，減少通勤者在等待時的焦慮與無聊；(3) 旅遊資訊系統的顯示面版裝設於車站平台、廣場、入口與轉乘處，以呈現交通資訊，相關交通資訊例如：下兩班車的抵達時間、服務中止、延遲時間與車站最後的服務時間，讓通勤者在其旅運過程中進行決定；(4) 電子旅遊指示為一個網路，提供通勤者公車路徑與車費資訊，如同公車、輕軌與地鐵服務的時刻表。此項資訊可以使通勤者在許多資訊下計畫他們的旅遊，例如：車費、轉乘次數與最短的旅行時間；(5) 加強對老年人與不便者的可及性，包括：斜坡、電梯與可觸式的引導系統。

### 三、整合大眾運輸系統

整合是指行駛媒介設施友善地從一地到另一地的交叉連接。改善大眾運輸工具之間的整合有助民眾移動更容易，並減少成本與旅運上的不便。綜合性的資訊系統可以使民眾容易地與快速地發現與比較不同路徑，並從中選擇最適合的一個。因此，協調整合不同的大眾運輸模式將會減少道路的擁擠、便利通勤者、更有效率、成本有效率，以及主要策略在於邁向新加坡開發「世界級土地運輸系統」目標，近幾年來許多方案已經實施以強化新加坡的運輸整合。

#### 四、公共運輸回饋計畫

Travel Smart Rewards(TSR)是一個激勵新加坡通勤者的機制，參與者可以依照他們搭乘鐵路系統的里程數比例去換取點數，在離峰時段搭乘還能獲得額外的點數。點數的賺取是根據通勤者的旅次在車站開始時間以及旅行的距離，這些數據及資料都會記錄在通勤者的卡片中。

由於地鐵路網擴建時程緩慢，且造價費用過大，恐無法滿足新加坡日益快速增加之人口與旅運需求，故新加坡政府也積極思考改善通勤族搭乘經驗的對策，在營運面，陸運交通管理局嘗試了幾種旅運需求管理措施，推行了一系列的智慧旅行計畫 TSR（如表 2-7 所示），希望透過誘因手段鼓勵通勤族改變平日搭乘習慣（例如重新設定出門時間、通勤路徑等），甚至減少旅運需求，期盼能將旅運需求由尖峰時段引導至離峰時段，藉以舒緩尖峰時段之大量人潮。

表 2-6：新加坡 TSR 實施方案

新加坡智慧旅行公共運輸回饋計畫			
於早尖峰 07:45 前至指定車站轉乘/出站，則該旅次免費。	於早尖峰時段 07:45~08:00 至指定車站，則享有 0.5 圓折扣優惠。	於離峰時段使用月票，即享有不限里程之乘車優惠。	大眾運具旅次以乘車公里數計算累計點數，兌換獎勵。

智慧旅行計畫內容簡述如下：

##### （一）尖峰時段前免費旅程：

1. 只要在平日晨間尖峰時段 07:45 前，於 18 個指定車站任一站出站，則當次旅程免費，本措施於 2013 年 6 月 24 日施行，並將持續至 2016 年 6 月 30 日。
2. 旅客於平日 07:45~08:00 出站，可享有 0.5 元車資折扣。

## （二）智慧旅行獎勵：

票證系統彙整搭乘公共運輸通勤族群的乘車資料，並將搭乘依賴程度依序畫分為四等級，分別為：棕、銀、金及白金等四類，只要使用者於平日晨間次尖峰時段乘車，將可獲得更多之點數加乘(目前平日晨間次尖峰時段為 06:15~07:15 和 08:45~09:45)，每週五結算點數一次，其點數每 1,000 點即可兌換 1 元新加坡幣或等值獎品，每月即舉辦一次回饋禮品抽獎，最大獎金高達 1,500 元新加坡幣。





	 <b>Bronze</b>	 <b>Silver</b>	 <b>Gold</b>	 <b>Platinum</b>
<b>Monday - Friday</b>				
Before 6:15am off-peak	1x	1x	1x	1x
6:15 - 7:15am decongesting	3x	4x	5x	6x
7:15 - 8:45am peak	1x	1x	1x	1x

圖 2-1：乘車依賴度級別

資料來源：新加坡陸運管理局網站 2015

## （三）離峰時段優惠月票：

於 2015 年 7 月 5 日推行，適用於公車、地鐵和輕軌，不限里程，除平日尖峰時間(上午 06:30~09:00 和下午 17:00~19:30)外，其餘時間（含週末及國定例假日）皆可使用。

## 五、私人運具之管制

近幾年新加坡政府已經推廣許多實質的評估，去解決汽機車的使用與持有所導致的問題，早期的一項去抑制車輛持有成長的評估方式是增加道路稅（Tan，1976）。然而，道路稅的增加並沒有減緩車輛持有的增長或減少尖峰時間的擁擠。因而相繼發展出其他抑制車輛的方式，例如：額外限制費（the Additional

Registration Fee, ARE) 為在城鎮一種較高額的停車費、錯開上班尖峰時段、停車與轉乘大眾運輸 (Park-and-Ride) 方案與推出區域執照方案 (Area Licensing Scheme, ALS)。於 1975 年 6 月新加坡實施區域執照方案(ALS)，其定義為一個城市的限制區 (Restricted Zone)，除了遵照時刻表發車的公車、警車與救護車之外，在此區域所有的車輛必須事先購買有授權的執照，區域執照方案(ALS)營運時間為平日早上 7：30 至晚上 7：00；星期六與五大國定假日前夕為早上 7：30 至下午 2：00；星期日與國定假日不營運。在實施區域執照方案(ALS)之後始下降 45%的交通量，但是後來因為市中心就業人口增加，使此方案不永續。

另外，針對道路交通管理有兩項主要的需求管理工具，分別是車輛定額系統 (Vehicle Quota System, VQS) 與電子道路收費 (Electronic Road Pricing, ERP)。在 1990 年 3 月，政府藉著授權執照 (Certificate of Entitlement, COE) 推出車輛定額系統 (Vehicle Quota System, VQS)，更進一步控制新加坡汽機車數的增長。在此規定之下潛在的車輛持有者於登記車輛之前必須爭取加入會員，這項方案固定長期汽車持有人數維持在永續的成長率上。在此系統之下，新加坡陸路管理局 (LTA) 管理車輛持有人數的成長維持在每年 3%。然而，Foo (1998, 2000) 發現授權執照 (COE) 增加了車輛的價格，這對於車輛持有者的負擔造成不利影響。另外，透過入會程序提供會員更好的可及性資訊，因而可以做更好的資訊決策。

於 1998 年電子道路收費 (The Electronic Road Pricing, ERP) 系統的營運取代了區域執照方案，這項系統首次擴大規模使用以控制擁擠的交通，目標在於使汽機車使用者意識到行駛的真正成本，以管理交通擁擠。電子道路收費(ERP) 是根據交通水準隨著不同的車輛類型、不同的時段與區位允許彈性的變動費用，藉由科技予以支配道路使用的優先順序。在電子道路收費 (ERP) 監控下，有 17 %的交通量下降。

### 2.2.3 各國城市比較小結

香港地鐵營運較早、每日乘客量較多；新加坡地鐵系統較長；而台北的營運時間較晚，下表將香港地鐵、新加坡地鐵與台北捷運做一比較。

表 2-7：大眾運輸系統營運比較表

地區	香港	新加坡	台北
運輸系統	城市鐵道系統	城市鐵道系統	大眾捷運系統
營運日期	1979 年	1987 年	1996 年
系統長度	177.4 公里	148.9 公里	131.2 公里
車站數量	53(16 轉運站)	63(5 轉運站)	69(2 轉運站)
每日旅次量	約 462 萬	約 265 萬	約 196 萬
營運單位	港鐵公司	新加坡地鐵公司	台北大眾捷運公司
民眾旅運行為	50%公車市佔率	近八成公共運輸市佔率	高私人運具市佔率
	21%地鐵市佔率		多數為機車族群

根據上述個案分析後，可提供國內改善大眾運輸現況的參考，將說明如下：

1. 香港擁有多元化的公共運輸工具，且政府亦極力推動公共運輸之建設，在不同運輸服務上有客源量之差異，意味著公共運輸市場內存在著偏好公車族群及偏好捷運之族群，其整體的公共運輸環境與台北相似，皆為富有公共運輸設施的都市，其公共運輸市佔率之差異主要之差異點。
2. 由於新加坡境內具有對於私家車輛管制之政策，使民眾在擁有私人運具多了金錢成本之考量，使當地之自用車旅次量相對的低上許多，其市區大眾運輸路網長度也與台北都會區十分相似，透過了智慧旅行公共運輸回饋計畫吸引民眾對於公共運具使用之偏好，以創造近八成的公共運輸市站績效。

## 2.3 大數據(Big Data)

大數據(Big data)，或稱作：巨量資料、海量資料、大資料，係指所涉及之資料量規模巨大致無法在合理時間內以人工處理，進行分類、分析、擷取及管理等方式轉化成可解讀的資訊型式，在許多領域，由於資料集過度龐大，科學家經常在分析處理上遭遇限制和阻礙；領域包括氣象學、基因學、神經網路學、物理模擬，以及生物和環境研究。此限制也對網路搜尋、金融與經濟資訊學造成影響。資料集大小增長的部分原因來自於資訊持續從各種來源被廣泛收集，包括：搭載感測裝置的行動裝置、軟體記錄、相機、無線射頻辨識（RFID）和無線感測網路。

近來隨著大數據相關應用及研究範疇的普及，部分專家學者逐漸提出以大數據為導向的預測觀點，《紐約時報》2012 的一篇專欄中提及，「大數據」時代已經降臨，在商業、經濟及其他領域中，決策將日益基於資料和分析而作出，而並非基於經驗和直覺。因大數據具備完善的樣本條件下，與獨立小型的資料庫相比，亦可藉由資料交叉比對分析的手法下得出需多額外的資料關聯性，過往在資料採礦(Data Mining)技術下，有許多公司管理者藉由挖掘公司內部的人力配置數據及產品的銷售數據，藉以提升企業的營運績效及市場行銷策略的擬定，也有致力於醫療資源下的相關應用，隨著大數據意識抬頭，其分析的類別可略及察覺商業趨勢、判定研究品質、避免疾病擴散、打擊犯罪或測定即時交通路況等；這樣的用途正是大型資料興起的原因。

### 2.3.1 大數據於交通類別的應用

現今已有多數的公共收費採取電子支付系統，其每日所產生大量的交易數據顯示著多數民眾的消費習慣以逐漸轉變，同時營運者也能藉由此資料探索民眾的



消費偏好，(M.P.Pelletier et al. 2011)指出，智慧卡的資料數據之應用可於：戰略(公司長期的營運策略)、戰術(業務類別之調整及服務對象)及方針(系統搭乘人數及服務績效指標)等三個面向。

有別於以往的貨幣交易模式，智慧卡的交易方式不僅能減少我們現金不足的困擾及節省部分的交易時間，隨著互聯網絡及通訊科技的蓬勃發展，智慧卡的使用範圍已與我們生活息息相關(Blythe 2004)，早在 1995 年英國即開始發展智慧卡片的應用範疇，其涉及了郵政、電話、醫療產業、銀行及公共運輸等類別(Chen 1995)。以交通類別的應用為例，利用智慧卡每日產生的交易紀錄及乘車資料，營運者可有效管理及監控車資收入的安全性，同時也減少了以往駕駛人員收取車資的負擔(Trépanier et al. 2004)，在具備方便、效率、及便於管理的條件下，智慧卡所帶來的影響，營運者即可提供精確且創新的服務來創造商機(Dempsey 2008)。

表 2-8：大數據於交通相關應用彙整

作者/年代	分析變數	研究成果
Agard et al. (2006)	進出站點、進出站時間、卡種別	以乘車次數定義旅客使用級別，並更加瞭解旅客偏好是每日、每周或是每月而有變化。
Chu et al.(2009)	登機時間、登機地點、降落地點、乘客別	考慮旅次時間、候機時間及轉乘時間等變因計算其總旅次時間，並與相同起迄點的航線交互比較，對於飛航路線及班距進行調整。
Chapleau and Chu (2007)	登機時間、登機地點、降落地點、乘客別	觀察乘客於轉乘機場之行為及可選擇轉乘的航線，並對於對應機場的貨運資源配置進行調整。
林祥生、劉益豪(2008)	消費時間、消費次數、消費費用、艙等及旅次長度	以 RFM 法將旅客區分成四種族群，並針對不同族群別提供如：票價優惠、機場接駁服務等行銷方式。
陳民祐(2012)	購票機率、取票金額、購票時間及旅次長度	以 RFM 法將旅客根據購票頻率區分成 11 種族群並散佈於四象限內，並歸納出相異族群之行銷方式。

### 2.3.2 智慧卡票證相關應用

#### A. 國內相關研究

林祥生(1991)公車票證交易電腦化的實施，除可增進營運管理上的效益，同時亦提供記錄明確詳盡之電子票證資料，除了作為政府日後相關補貼政策或路線規劃之參考外，更重要的是，學術研究單位或經營者可依據電子票證交易紀錄內容，精確分析乘客起迄運量表，取代目前高成本、低效率的傳統運量調查方式；

Bagchi and White(2005)研究中以起迄點資料，來調整運輸服務、改進品質，Chapleau and Chu(2007)則是分析特定路線上乘客的變化及檢視運量的轉移，將路線最大運量上車點，瞭解到往返旅次之最大運量。Wei Wang(2011)電子票證研究針對路線旅運量負荷變動，找出每條路線尖峰時段可能無法滿足之區段，及轉乘時間分析來檢視轉乘旅次服務情況如：轉乘等待時間，上述票證資料相關研究皆為了能更瞭解乘客之旅次特性

林祥生等人(2005)研究指出，利用公車系統的票證資料構建分析模式，分別探究總體資料群之乘車需求及個體旅客搭乘行為之差異點，在總體資料分析的手法分別比較平日及假日公車系統的乘車旅次量差異，其平假日各時段尖離峰的旅次變化趨勢、及各乘客別的旅運特性差異；而在個體旅客行為的資料處理上，則是藉由比較一般族群及學生族群的旅次特質習性，提供營運者瞭解相異乘客別之旅運行為差異，能針對欲提升目標族群的服務水準同時亦能有效的配置安排車輛的班次供給及人力安排輪休。

邱詩淳(2006)則選用電子票證的乘車時間記錄與公車定位系統的到站資訊相互比對，以推估乘客旅次之 O-D 矩陣，利用乘客之旅次矩陣進行公車服務供給之服務水準改善，分別檢視區間車及直達車之最佳行駛路線，以建立最佳模式評比之關係式，包含：最小時間成本路線與時間成本路線、最佳車內時間路線及

最佳候車時間成本等類型，其主要的目的在於總體營運成本最小化。該研究選擇以 802 公車路線為例進行實例分析及應用，其研究結果顯示：直達車的於上午尖峰時段的實施效果最佳，而區間車的營運績效於任何時段內皆優於直達車之營運績效，而時間成本、行使成本及旅客車內及候車成本皆與模式評比關係式呈正向關係，此研究成果可提供營運者依照旅客需求及時段之特性調整服務供給之方式，增進營運績效。

羅惟元(2008)以 802 公車票證交易資料做研究對象，選定公車使用者之起迄比對表，根據悠遊卡資料特性，分別以公車系統面探究旅客乘車關係及以旅客資料的類別探討選搭公車的時點差異。藉由旅次客性起迄表得知何種使用者類別為經常性使用者，藉由出現於早尖峰及晚尖峰的出現頻率與以界定乘客之價值。

林良泰等人(2011)認為開發建置電子票證系統，除合理反應使用者付費之精神外，亦可收集旅次起迄點及旅客量資料，未來可藉此調整公車班次並開發新的公車路線，吸引潛在客源。

Wei Wang(2011)表示智慧卡(smart card)將成為可用來分析旅行者行為的新運輸資料來源。可更容易取得大量的個人旅行資料。彙整智慧卡之優點如下

- 1.能得到大量且單獨各人的旅次行為資料
- 2.能辨識旅次是搭乘哪種大眾運輸工具
- 3.透過感應刷卡的動作所取得的調查資料會比人工調查來的便捷 透過票證資料分析和應用可做為實務營運時之規劃依據

其國內智慧卡產出資料研究整理詳見如下：

表 2-9：資料採礦於電子票證之相關研究彙整

作者	林祥生等人(2005)	邱詩淳(2006)	羅惟元(2008)
研究對象	票證系統	票證系統	票證系統
研究方法	資料採礦	資料採礦 (群集化、關聯規則)	資料採礦
探究問題	瞭解使用者旅次特性及乘車行為	公車行駛最適路線	建立使用者旅次起迄比較表
研究成果	<ol style="list-style-type: none"> <li>1. 比較平假日旅次量及尖離峰旅次差異</li> <li>2. 歸納出相異族群尖離峰各自旅次特性</li> <li>3. 提供營運者個別族群旅運選擇偏好</li> </ol>	利用群集化與關聯規則分析公車最適路線規劃	<ol style="list-style-type: none"> <li>1. 建構運輸旅次起迄分析法</li> <li>2. 快速產生起迄分佈特性資料</li> </ol>

資料來源：謝萬興(2015)運用悠遊卡巨量資料分析公車乘客行為之研究

## B. 國外相關研究

智慧卡交易紀錄所產出之大數據相關應用於交通領域於國外行之有年，本研究欲瞭解國內使用者之旅運特質研究過程，故參閱國際上各城市以智慧卡進行之相關研究，分析比較及說明各差異點，藉由智慧卡研究之成果探究該國家之旅運行為特質，都市交通特性及時段需求之差異點，進而彙整其不同研究之成果、研究方法與產出。

Bagchi and White(2005)研究指出，智慧卡資料可應用於分析乘客對於新的運輸服務供給偏好來源，以克服現存營運資料不足之劣勢，藉由各統計指標中反應出乘客對於運輸服務之偏好及需求。藉由該系統業者可從中獲得下列資訊：

### 1. 大量旅客通勤旅運資料

2. 透過資料庫檢視各別乘客族群之每日通勤資訊
3. 藉由分群分類等技術可依照乘車比例區分乘客忠誠度價值

然而以現今的交通旅運資料庫分析乘客旅次偏好及特質是有相當挑戰的，由於資料庫類別資料不完整之故(如：缺乏旅次起迄點、旅次目的等)，這些問題點須仰賴相關的旅次端點推估及預測等調查方法予以克服彌補，故智慧卡具備大量之旅次資訊，但仍須與現有之資料收集方法配合以建立完整之旅次資料庫。

Long et al.(2012)研究北京公車智慧卡交易紀錄中指出，多數的旅運通勤資料皆存在著適地性(Location Based Service,LBS)資料雖然具備充足的時間及空間資訊，但礙於各資法之限制故缺乏個體使用者定位的類別資料，使其研究範圍受到相當之阻礙。該研究結合了該年度該區域之公車智慧卡交易資料與外出旅運偏好調查之結果，以及土地分區及社會經濟等相關資料交互比對，以區別使用者之居住地、就業分區及通勤旅次長度，以建立公共運輸旅次端點和交通分區(Traffic Analysis Zone,TAZ)之彙整，其研究結果如下述：

1. 交通勤資料以時間維度及空間(距離)維度兩面向進行探討，與外出旅運偏好調查結果對比，得出高吻合性之結果。
2. 取得居住分區及就業分區中旅次鏈之時間與空間相依性，進而與通勤族的旅運特性進行推估。
3. 對於該分區之公車通勤族進行視覺化分析，識別主要的交通路網動向。

Tao et al.(2014)以澳洲公車捷運系統(Bus Rapid Transit,BRT)智慧卡交易資料庫為研究案例，應用地理視覺法(Geo-Visualization-Based Method)，研究該運輸系統使用者之時間及空間之旅次變化，並發現 BRT 族群及公車族群之旅次差異變化點(平假日、國定假日乘車偏好)，其研究成果如下述：

1. 根據資料庫敘述性統計結果可之，BRT 族群之旅次長度比一般的公車使用族群之乘車距離還來得長。
2. 透過旅行模式分析法發現，北區之公車專用道與東南區之公車專用道為主要旅次群集的集散路網，以連接市區及郊區的接駁路線。

Eom and Sung(2011)以南韓首爾年長者智慧卡票證車行為例，資料顯示所有的交易紀錄中年長者僅佔 5%之比例，該研究指出，年長者之乘車時段多數落於上午九點至下午 5 點此區間當中，且並無明顯落於哪時段中，平日及假日之乘車偏好亦無明顯之變化。在運具選擇及轉乘的偏好上，相較於其他族群之運具組合年長者之交易紀錄中約有九成的交易紀錄皆為使用單一運具為主，且多數偏好軌道運輸系統而非公車系統，其轉乘逾三次比例不足 2%，即顯示年長者族群之運具偏好會隨著轉乘之次數而遞減，該研究成果可提供營運者改善行動不便之族群及高齡者族群對於轉乘行為意願不高之問題。

## 2.4 資料採礦相關技術

2015年公共交通國際聯會的Open big and smart data演講中提到：倫敦預期至2030年人口會成長至一千萬，現階段一周約有460萬公車旅次，而不含公車之公共運具則有250萬旅次。大數據資料來自於刷卡資料、交通資訊中心、顧客回饋、運具回饋資訊，在資訊複雜且眾多底下，如何把相關數據結合，透過演算法分析出可用數據，成為一大挑戰。日前台北市政府使用悠遊卡資料庫探索民眾乘坐大眾運具之偏好，透過如公用自行車、捷運及公車等歷史資料庫以探究民眾乘車行為，期盼能做為公共運輸路網規劃參考之依據，增進其大眾運輸路網服務效率，以透過大數據資料，可瞭解市界對於數據革命的脈動。

近年來透過物聯網的平台整合，使得許多資訊與網路後端的系統平台結合，所以各行各業普遍以電腦來匯集各種資料，但雜亂無章的資料是無法呈現即表達

義涵，故此透過資料採礦(Data Mining)的技術，即可將毫無組織、規則及排序義涵之資料進行排列及分析。本節針對資料挖掘進行相關文獻探討，首先概述資料挖掘之定義；再藉由文獻回顧之手法介紹各專家學者提出不同之資料挖掘手法及流程，並進行統整。

## 2.4.1 資料採礦(Data Mining)之定義

資料採礦是一種資料轉換的過程，先由沒有組織的數字與文字集合的資料，轉換為資訊，再轉換為知識最後產生決策；也就是從資料中發掘資訊或知識，有人稱為「資料考古學」(Data Archaeology)、「資料樣型分析」(Data Pattern Analysis)或「功能相依分析」(Functional Dependency Analysis)，目前也有許多相關的統計學者及專家，將資料庫系統結合並與電腦學習之技術聯合開發，也有許多產業學者認為，資料採礦為改變未來社會經濟行為的重要指標。下表為各學者對於資料採礦各自之定義及類別。

表 2-10 各學者對資料採礦定義及方法

學者	年期	資料採礦之定義
Frawley et al.	1992	為一大量自動處理資料的過程，運用統計分析將資料庫中重整，開發出有用、先前未知特徵資料趨勢。
Kleissner	1998	資料採礦是挖掘公司資料庫中所隱含的知識，並讓企業管理者瞭解內部及外部環境之變化趨勢，以做為營運決策支援的分析過程。
S.C. Hui and G. Jha	2000	為近來各專業領域挖掘知識，迅速興起的技術，了解資料中各面向需求與動機關聯性，透過分析、理解以及可視化等技術，對資料呈現結果進行應用。
YongSeog Kim	2004	從資料庫中挖掘出不明確及隱含於資料供的資訊，是一種預測未來的操作過程。
Han and Kamber	2010	傳統的資料採礦通常都採用結構化數據進行分類及預測，並以聚類分析之手法區別結構及非結構型資料。

## 2.4.2 資料採礦的功能與流程

一般來說，資料採礦之手法主要分為五大構面，其概述內容如下：

### 1. 分類

係指事先將資料依據其特徵性、同質性及類別進行區分處理動作，並依據已知之類別變數進行處理及運算，處理後可得資料間各類別變數間關聯及相關規則，再將其關連規則套入未知之類別資料，取得輸出分類結果。

### 2. 預測

應用資料內之歷史資料做為模型建立之基礎，透過時間之變化檢視資料觀察值，以預測或推估資料未來之變化趨勢。

### 3. 集群化

依據資料庫特徵值之特性進行分群，一般來說，透過演算法不斷之訓練及修正之過程中，資料群會依據其特性逐漸群集化，其處理之主要目的在於區分群體間之差異性，再對個別群體之相似樣本進行挑選，因為通常在進行群集分析前，研究者不會知道該資料應運用何種方式進行分析與判別。

### 4. 關聯規則

主要希望透過變數間之多屬性及多組合之關聯特性進行探討，以尋找使用者對於選擇偏好上其各變數間彼此之交互影響特性，而此手法多數應用於市場內商品銷售分析處理之範疇當中。

### 5. 順序分析

針對資料群之先後順序關係進行處理，主要希望透過資料變化之趨勢中取得關聯資訊，多數採用如時間變化順序、行為選擇之順序等類別。



### 2.4.3 集群化

集群化(clustering)，是應用於資料分析 (data analysis) 的非督導式技術之一。主要功能是將一群在多維度空間 (multi-dimensional space) 中的資料，依照資料的屬性或特徵來分類。因此，相似性高的會被分為一群，而分群後的資料，我們稱為集群化 (cluster)。

集群分析是資料探勘的一種，是根據物件資料以及物件之間的關係資訊為基礎來計算相似度，並以彼此的相似度來計算並分群，而將各相似物件集群到相同群集中，在集群分析中如何衡量物件間的相似性是很重要的而目前也依不同的理論有各種相似度的計算和分群，如以雛型為基礎、以圖形為基礎、以密度為基礎、共享屬性等分群方式(施雅月、賴錦慧 2008)，相似度的計算可採用 ManhattanEuclidean 與 Minkowski 等數種距離衡量法測量距離遠近以定義相似度，也可以採用不同的機率分配定義相似度。然後依據所需相似度的要求分成不同數量的群集，相似度越小群集就會越少。近來集群分析依據不同的分群理論基礎和不斷的改良分成以下幾種集群分析的方式：

- (1)分割方式分群法：包括 k-means 法、k-medoids 法、CLARANS 法。
- (2)階層式分群法：包括凝聚式階層分裂法、分裂式階層分群、階層平衡式重複縮減與分群(BIRCH)、類別屬性的階層分群(ROCK)、使用動態模型的階層分割法(Chameleon)。
- (3)密度式分群法：包括根據相連高密度區域的密度式分群法(DBSCAN)、透過資料點排序來找出群組結構(OPTICS)、根據密度分佈函數進行群組。
- (4)方格式分群法：包括統計訊息方格、使用小波轉換進行分群。
- (5)模型式分群法：包括期望最大化分群法(EM)、概念分群、網路類神經分群。

(6)高維度資料分群：包括維度增長子空間分群法(CLIQUE)、維度縮減子空間分群法(PROCLUS)、頻繁樣式分群法。

(7)限制式分群分析：包括有障礙個體分群、使用者限制的群組分析、半監督式分群。(王派洲 2008)集群分析需要依據相似度來群聚，較常用的資料點距離計算方式和步驟(林鼎浩 2000)。包括 Minkowski 距離、Euclidean 距離、Manhattan Mahalanobis 距離、謝比雪夫距離、相關係數。機率方面則是較常使用高斯分配(GaussDistribution)來計算相似度。

群集法是為非監督方式所以並不知道將會分出什麼樣的集群，以相似度來集群的群集，對於探勘者來說必須自己去解讀和詮釋分析出的集群意義，重點是探勘者必須很了解資料分析前和分析後所代表的意義，以及探勘者對整個問題和探勘原因必須非常了解。我們將會使用集群分析來將時間較密集中被取出的查詢和資料表的組合予以集群，作為一個或數個決策的所需資訊。

#### 2.4.4 關聯規則

關聯法則也被稱為購物籃法則，是由於美國最大超級市場連鎖業者 Wal Mart 所使用來分析顧客購買商品之間關聯性而得名，是使用顧客購買的商品為資料以關聯法則找出關聯性較高的商品組合，依關聯性高的商品組合作為商品擺設的參考，比如啤酒和尿布會在同一次購物裡出現的機率很高就可以將兩樣商品擺在鄰近位置促進購買機率或者擺在較遠位置，以使顧客走更多的路增加更多商品被瀏覽的機會。關聯法則的演算法在後來發展出許多種類也不斷發展改良。關聯法則較具代表性的是 Agrawal(1994)提出的 Apriori 法則，為使用支持度、信賴度和增益找出關聯的演算法，在找出所有項目組後由支持度求出高頻項目組。

## 2.4.5 資料採礦之技術應用

整體來看，資料採礦為應用大量的資料數建立可應用於實際環境模擬之技術，以萃取出具研究意義之模式關聯規則。截至目前，資料採礦技術仍時常應用於市場行銷、醫療資源管理及公司內部管理等類別中，主要以提高企業之營運績效及減少人力資源耗損為應用標的。

陳民祐、王建富(2013)的研究則應用資料採礦的技術，挖掘台鐵電話及網路訂票的資料庫，以 RFM 法將顧客需分為四種消費族群，並以積點回饋的會員計畫來刺激顧客消費頻率與使公司獲利得以提升，同時增進與維繫顧客的關係。林祥生、劉益豪(2007)的研究則將資料採礦的技術應用於國際線乘客的購票偏好上，同樣是應用 RFM 法區別顧客的市場定位，藉由乘客乘搭機倉之類型與平均飛行里程，等五項消費特質予以進行群集化，以建構乘客分類的預測模式，最後再以不同的旅次長度及特性，分析航線間的關聯規則。

## 2.5 文獻小結及評析

### 1. 完善之公共運具環境及相關誘因推廣，能降低私人運具比例

私人運具的持有與使用，主要還是與一個都市的大眾運輸系統是否完善最為相關，若在一個大眾運輸系統完備下的都會中，公共運具提供高效率之旅運服務，並配套相關之公共運具乘車優惠之鼓勵下，其民眾選擇私人運具及持有也會降低。

### 2. 多元的旅次回饋手法增進民眾對於公共運輸之偏好

前述回顧城市為例，其當地皆具備完整之公共運輸路網，且實施不同旅次回饋的政策鼓勵民眾搭乘大眾運具，以創造高公共運輸市佔率之績效，得

以做為國內運輸市場效法之對象。

3. 大數據資料庫具備大量且多元之消費交易記錄

透過物聯網的資訊技術整合，網路所產生的資訊被電腦系統記錄下來，使用者無形之交易資料反應最直接之消費型態，過往之研究嘗試藉由不同之資料處理手法探究使用者交易記錄，並建立其應對策略以做為營運者及管理人員之決策依據考量。

4. 以集群及關聯規則手法探究交通大數據通勤旅次特質

資料採礦為將雜亂無章且難以辨識的資訊整合分析之技術，並透過不同的資料處理分群等技術處理不同性質之資料類別，而集群分析的處理概念為將性質相仿的資料類別與以分群，而關聯規則的處理方式則是藉由挖掘龐大的資料中，探索資料項目間的相關性。

5. 以捷運乘車資料探究通勤族群之旅運特質

過往相關研究多數以探究公車之乘車資料，以消彌公車服務缺口及提升公車發車班次及服務水準做為研究標的，根據資料顯示台北捷運每日載客量近兩百萬人次，可見得捷運系統對於雙北通勤族群之重要性，故可嘗試以捷運乘車資料檢視其通勤族群對於公共運具之旅運需求偏好與特質。

6. 以乘車優惠做為誘因檢視實施於國內之可行性

在回顧國外乘車誘因相關案例中，可知在完善的公共運輸路網配套其公共運輸乘車鼓勵政策，能有效提升其公共運輸市佔率之成長空間，回顧國內雙北區公共運輸網不亞於新加坡及香港等城市公共運輸服務水準，而在乘車補助上僅有偏遠地區之公車接駁服務才有相關的配套措施，其將乘車誘因實施於私人運具較密集之區域，希冀能改善通勤族對於公共運輸的依賴程度。

### 第三章 研究方法

大數據的議題正在世界延燒，目前透過手機載具、電子穿戴票證載具、刷卡交易購物皆有紀錄，透過這些紀錄可確實掌握民眾消費及旅運相關行為。倘利用這些資訊做為基礎，透過資料分析及乘客偏好的關連探討可讓大眾運輸資訊運用更加有價值，亦可作為成本及風險相關評估之依據。日前台北市政府使用悠遊卡資料庫探究民眾乘坐大眾運具之偏好，透過如公用自行車、捷運及公車等歷史資料庫以探究民眾乘車行為，研擬各運具使用者之時空間偏好關聯式，期盼能做為公共運輸路網規劃參考之依據，進而增進其大眾運輸瞭解服務效率。

故本研究以悠遊卡乘車資料做為研究基底，並選用2014年6月1日至2014年6月30日共三十日之公共運具票證資料，其包含了捷運及公車兩運輸系統，由於公車之票證資料僅能從單一的上車或下車記錄進行旅次起迄點之推估，固本研究引用創代科技-台北市公車旅次OD資料研究一文中進行引用，以在推估公車使用者之旅次長度條件下進行後續相關之研究工作。

本研究欲效法台北市政府探勘悠遊卡資料庫之手法觀察使乘客乘車偏好，選定旅次起始時間欄位及起始站點欄位，推估多數乘客的旅次長度及乘車時段偏好，運用資料探勘(Data Mining)技術中常用的K-means群聚法相資料相似的群體進行群聚，再藉由分群之結果進行Apriori關聯法則的分析，以瞭解大眾運輸的使用者旅次特質之時空關聯性。

由於悠遊卡票證資料僅能記錄其使用者使用公共運具之相關資料，受限於個資法及隱私保護之限制下，本研究所取得之有效資料欄位並未能有效鎖定單一族群之所有旅運資訊及私人運具之使用變化，故在私人運具旅運特質的探索當中，本研究欲以問卷發放探究機車使用族群每日通勤之旅次資料，及對於公共運輸服務偏好，藉由不同情境乘車誘因模擬能增加機車族群使用公共運輸特質為何。

### 3.1 研究架構

實證研究架構流程演變如圖所示，最後分為四大階段，第一階段係先結合「資料清洗處理」及「欄位合併歸納」，則當中也需確定資料轉換之格式，則將此階段命名為「資料庫初步處理」；其第二階段首先將「時間欄位」及「距離欄位」進行標準化集群處理，此階段欲利用資料採礦中之集群化技術，將欲分析變數依造給定條件定義成新的類別群集變數，因此將此階段命名為「時距分析法」，則第三階段一開始皆為選擇研究方法、演算法、分析方法，而本研究選定使用關聯規則方法選定具備乘客旅次特質之關連組合分析之，所以將此階段命名為「旅次特質關聯規則分析」；最後，第四階段則是以私人運具調查問卷的分析結果做為基礎，藉由發放問卷詢問學生機車通勤族群之相關特質，以瞭解機車族群每日通勤旅次的相關聯結，因此將此階段命名為「私人運具旅次調查」。

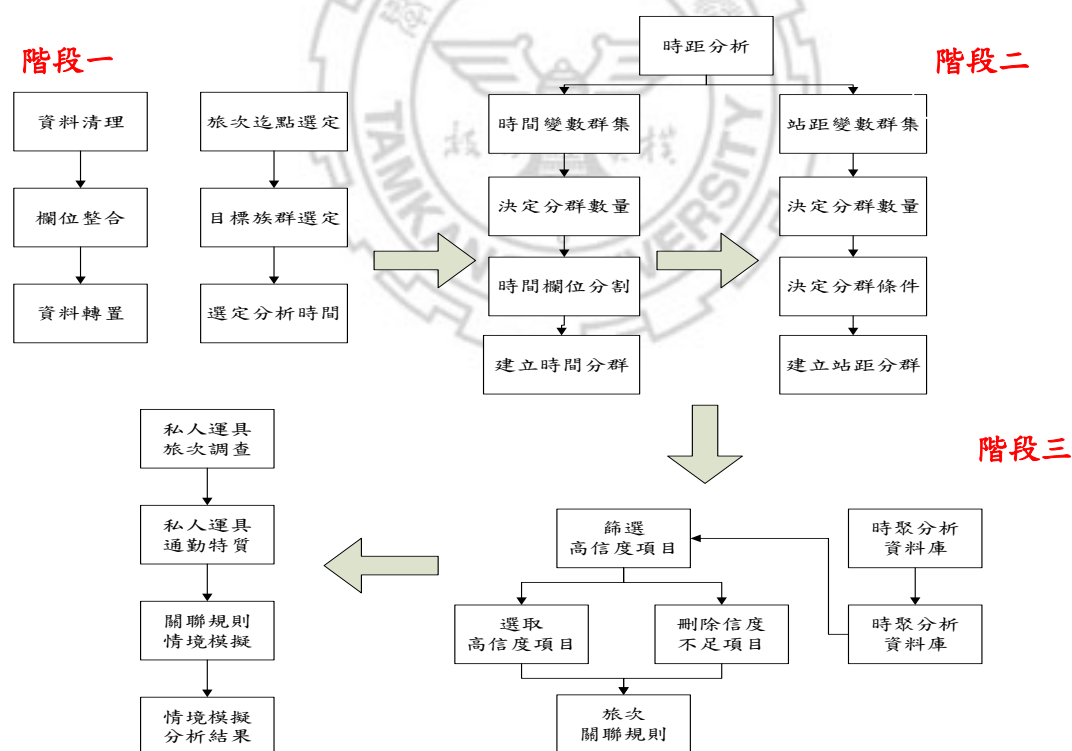


圖 3-1：研究架構圖

### 3.2.1 資料庫初步處理

本研究所取得之資料量十分龐大，多數記錄欄位與格式主要以記錄及儲存交易資料為優先，並非設計做為學術分析適用格式，故在進行分析前必須先進行資料前置處理之步驟，將原本儲存於記事本內的資料轉換至各自對應的欄位內容中。在悠遊卡公司所提供2015年6月1日至2015年6月30日共30天的原始乘客交易資料(如下圖所示)，共31,457,280筆票證交易記錄。

檔案(F)	編輯(E)	格式(O)	檢視(V)	說明(H)
1673917781	2	1	2015-06-15	08:30:17.000.46.100.16
3577553427	2	1	2015-06-15	19:32:58.000.50.36.20
236031110	2	1	2015-06-15	10:41:31.000.50.39.16
1651308117	2	1	2015-06-15	17:18:44.000.57.52.16
557425327	2	1	2015-06-15	18:23:35.000.52.34.24
1651439189	2	1	2015-06-15	14:49:42.000.86.51.16
2117837487	2	1	2015-06-15	21:38:26.000.54.127.8
1530897071	2	1	2015-06-15	21:44:41.000.93.52.16
1651439189	2	1	2015-06-15	10:16:50.000.42.86.16
2066910086	2	1	2015-06-15	09:53:49.000.38.131.24
2017436335	2	1	2015-06-15	18:13:27.000.91.86.16
1651439189	2	1	2015-06-15	08:21:39.000.86.42.16
3661007750	2	1	2015-06-15	07:35:42.000.71.52.40
2943767196	2	1	2015-06-15	08:24:02.000.35.56.28
2943767196	2	1	2015-06-15	14:59:15.000.56.35.28
2075480732	2	1	2015-06-15	19:45:19.000.82.77.20
3189188911	2	1	2015-06-15	18:34:05.000.52.77.28
2119994246	2	1	2015-06-15	18:24:00.000.101.55.20
1568184687	2	1	2015-06-15	19:01:54.000.24.91.24
648583535	2	1	2015-06-15	08:52:44.000.46.131.24
630074818	2	1	2015-06-15	07:32:00.000.56.61.20
2779289967	2	1	2015-06-15	10:28:55.000.86.131.16
2779289967	2	1	2015-06-15	21:04:19.000.93.86.20
637349314	2	1	2015-06-15	06:54:14.000.64.41.28
2071169414	2	1	2015-06-15	13:21:52.000.34.50.24
2122615174	2	1	2015-06-15	21:23:23.000.109.82.28
2122615174	2	1	2015-06-15	10:46:33.000.82.109.28

圖 3-2：智慧卡原始資料欄位

從原始資料集中，僅能得知卡片序號(經亂數處理)、乘客別、運具別、交易日期及時間點、對應區段等幾項資料，在經過初步的資料整合及清理，將資料中之異常值及極端值清除及合併相關欄位後，以便於進行後續的資料處理及分析。

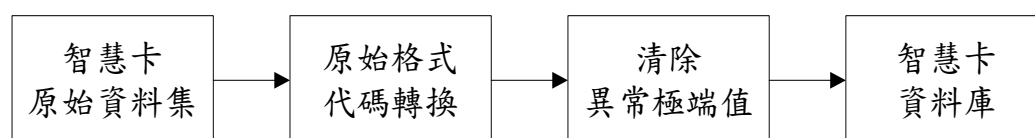


圖 3-3：資料前置處理流程

其經過初步的篩選及處理後整合至9項欄位，其資料格式如表所示，而其中與本研究有關的欄位解釋分別說明如下。

表 3-1：智慧卡欄位明細

編號	欄位名稱	欄位描述	編號	欄位名稱	欄位描述
1	FareID	序號	6	Zone_Num	行駛區段
2	Subtype	票種	7	Entry_flag	上/下車
3	TranDate	交易日期	8	Entry_Zone	上車區段
4	CarID	卡號	9	Area_Code	區段代碼
5	Line_Num	路線代號			

經過篩選資料內容及評估資料正確性後，其真正能保留下來且應用的資料只剩七種欄位類別：持卡人卡別、搭乘運具種類、乘客進出站時間、公車行駛路線、進出站點、站點編號、站點分群編號等欄位變數。其保留原因如下：

票種：辨識乘客身分種類、運具別：可得知乘客所搭乘運具之種類、交易時間：得知乘客刷卡時間、運具行駛路線：可得知乘客所搭乘之公車行駛路線、進出站點：有助於瞭解乘客旅次之起迄點、進出站點編號：有助於得知進出站點所對應之編號代碼、進出站分群編號：以站牌為基準將周邊所涵蓋之站點位置整合分群之代碼

表 3-2：續智慧卡欄位明細

編號	欄位名稱	欄位描述
1	Card Type	持卡人使用卡別(一般、學生、敬老)
2	Vehicle Type	搭乘運具種類(捷運、公車)
3	txn_timestamp1	乘客進站時間
4	line_no1	公車行駛路線編號(捷運無此資料)
5	stop_name1	乘客進站點(旅次起點)
6	stop_id1	乘客進站(站點編號)
7	stop_id2	乘客出站(站點編號)



	A	B	C	D	E	F	G	H	I	J	K	L
1	2.05E+09	22	1	50:08.0	0	1	1	132	3018	705FU	15	
2	1.22E+09	22	1	50:10.0	0	1	1	107	3018	705FU	15	
3	1.81E+09	22	1	50:12.0	0	1	1	165	3018	705FU	15	
4	3.59E+09	22	1	50:13.0	0	1	1	116	3018	705FU	15	
5	7.17E+08	22	6	50:15.0	0	1	1	36	3018	705FU	12	
6	2E+09	22	1	50:16.0	0	1	1	83	3018	705FU	15	
7	6.52E+08	22	1	50:25.0	0	1	1	204	3018	705FU	15	
8	2.08E+09	22	1	53:02.0	0	1	1	147	3018	705FU	15	
9	1.16E+09	22	2	03:36.0	0	1	1	92	2433	012FU	0	
10	3.38E+09	22	9	03:42.0	0	1	1	108	2433	012FU	8	
11	2.54E+09	22	9	03:43.0	0	1	1	136	2433	012FU	8	
12	1.6E+09	22	6	04:14.0	0	1	1	36	2433	012FU	12	
13	3.26E+09	22	2	05:11.0	0	1	1	130	2433	012FU	0	
14	2.61E+09	22	6	08:37.0	0	1	1	3	2433	012FU	12	
15	4.29E+09	22	6	33:57.0	0	2	2	57	2433	012FU	12	
16	4.23E+09	22	2	35:48.0	0	2	2	9	2433	012FU	0	
17	2.39E+09	22	2	37:30.0	0	2	2	75	2433	012FU	0	
18	2.92E+08	22	1	13:13.0	0	1	1	251	2433	012FU	15	
19	2.88E+09	22	4	15:51.0	0	1	1	129	2433	012FU	0	
20	5.6E+08	22	1	16:59.0	0	1	1	190	2433	012FU	15	
21	2.02E+09	22	6	18:12.0	0	2	2	217	2433	012FU	12	
22	2.03E+09	22	4	21:28.0	4	2	2	5	2433	012FU	0	

圖 3-4：智慧卡原始資料集交易欄位

### 3.2.2 資料轉選及處理流程

在進行資料採礦的過程中，首先需要一個完整的資料集，透過將資料群集的格式正規劃處理及轉置，以減少資料格式無法被演算法所讀取的問題；但是，處理資料的程序不僅複雜且十分耗時耗力，故本研究欲瞭解公共運輸族之旅運時間及空間關聯性，則需先鎖定欲分析的資料欄位進行處離及轉置。

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	card type	vehicle_type	txn_timestamp	line_no1	stop_name1	stop_id1	parent_id	vehicle	txn_timestamp	line_no2	stop_name2	stop_id2	parent_id2
7	2 M	2014/6/1 06:04	NULL	淡水	71 M71	B	2014/6/1 11:31	756 捷運淡水站	129872 M71				
21	2 M	2014/6/1 06:18	NULL	關渡	68 M68	B	2014/6/1 13:01	1504 捷運淡水站	151763 M71				
442	6 M	2014/6/1 07:14	NULL	淡水	71 M71	B	2014/6/1 15:25	756 捷運淡水站	129872 M71				
1402	1 M	2014/6/1 08:21	NULL	竹圍	69 M69	B	2014/6/1 08:52	1504 捷運淡水站	151763 M71				
1938	6 M	2014/6/1 08:44	NULL	關渡	68 M68	B	2014/6/1 22:11	1504 捷運淡水站	151763 M71				
2350	1 M	2014/6/1 09:03	NULL	紅樹林	70 M70	B	2014/6/1 13:32	756 捷運淡水站	129937 M71				
6581	1 M	2014/6/1 10:40	NULL	劍潭	57 M57	B	2014/6/1 11:12	756 捷運淡水站	129872 M71				
6989	6 M	2014/6/1 10:47	NULL	竹圍	69 M69	B	2014/6/1 17:33	756 捷運淡水站	129937 M71				
9964	1 M	2014/6/1 11:27	NULL	紅樹林	70 M70	B	2014/6/1 15:34	1504 捷運淡水站	151763 M71				
9988	1 M	2014/6/1 11:28	NULL	紅樹林	70 M70	B	2014/6/1 15:34	1504 捷運淡水站	151763 M71				
11373	1 M	2014/6/1 11:43	NULL	民權西路	55 M55	B	2014/6/1 18:22	756 捷運淡水站	129872 M71				
11918	1 M	2014/6/1 11:51	NULL	紅樹林	70 M70	B	2014/6/1 15:34	1504 捷運淡水站	151763 M71				
14044	1 M	2014/6/1 12:19	NULL	淡水	71 M71	B	2014/6/1 12:30	756 捷運淡水站	129872 M71				
14200	6 M	2014/6/1 12:20	NULL	淡水	71 M71	B	2014/6/1 17:19	1504 捷運淡水站	151698 M71				
16083	1 M	2014/6/1 12:44	NULL	淡水	71 M71	B	2014/6/1 12:52	756 捷運淡水站	129872 M71				
18319	6 M	2014/6/1 13:09	NULL	三和國中	177 M177	B	2014/6/1 18:22	756 捷運淡水站	129872 M71				
18322	6 M	2014/6/1 13:09	NULL	三和國中	177 M177	B	2014/6/1 18:22	756 捷運淡水站	129872 M71				
19414	1 M	2014/6/1 13:24	NULL	淡水	71 M71	B	2014/6/1 13:43	1504 捷運淡水站	151698 M71				
19928	1 M	2014/6/1 13:32	NULL	淡水	71 M71	B	2014/6/1 16:11	756 捷運淡水站	129872 M71				
19932	1 M	2014/6/1 13:32	NULL	淡水	71 M71	R	2014/6/1 16:10	756 捷運淡水站	129872 M71				

圖 3-5：智慧卡預處理後交易欄位

在票證乘客族群組成的部分，以一般卡使用者使用比例54.9%為最高，其次為學生族群的23.7%、敬老族群的15.46%、愛心卡使用族群的3.67%、優待卡使用者的2.07%以及愛心陪伴卡的0.68%，如下表所示。在運具類別資料數量中，包含了捷運、公車及公共自行車等三種資料類型，以捷運的交易數量16,681,380筆(佔整體之53.02%)最高，其次為公車的13,637,820(佔整體之43.35%)，最後為公共自行車的1,138,030(佔整體之3.63%)由此可知捷運及公車為北部地區多數民眾所仰賴的大眾運輸工具。

表 3-3：資料前置處理結果

運具種類	資料數量	百分比
捷運	16,681,380	53.02%
公車	13,637,820	43.35%
公共自行車	1,138,030	3.63%
總體資料數	31,457,280	100%

### 3.2.3 研究區域背景特性

由於新北市淡水區為北部地區大專院校分布最密集之行政分區之一，其區內包含了：淡江大學、真理大學、聖約翰科技大學及台北海洋技術學院等四所大專院校，根據教育部資料顯示，四所大專院校之學生總數約有47,300人，以新北市公共運輸市佔率31.1%水準為例，每日搭乘公共運輸往返淡水大學生族群就有近一萬四千餘人之通勤旅次；又因低收入族群或學生對機車使用有較高之依賴性，可推估淡水選擇私人運具學生族群有相當比重，不論是捷運轉乘公車或是公車轉乘轉乘旅次量，淡水站皆為各站之最，故本研究之研究區域選定為新北市淡水區

表 3-4：公共運具轉乘運量統計

捷運轉乘公車		公車轉乘捷運	
站名	運量	站名	運量
捷運淡水站	18,138	捷運淡水站	16,934
捷運新埔站	8406	捷運新埔站	7651
南港展覽館站	8242	捷運圓山站	6916
捷運景安站	6991	板橋區公所站	6454
捷運永寧站	4696	南港展覽館站	5505

### 3.3 時距分析法

#### 3.3.1 時間分析-群聚分析法

在對資料進行初步的清理及轉置後，時間欄位中其乘客進出站時間需依照時間之特性進行區別，以時段區間的方式進行群體分群，故在時間點處理的部分，本研究利用不同的時間切割方式來進行乘客間的分群，將乘客的乘車時間資料在系統前端即利用不同的機制優先處理未整理的資料，並將零亂的資料整理成較有規律性的資料，可避免過多的未處理資料在系統末端處理改善，增加系統資料分群之完整度。在資料前處理後，資料皆有經過篩選，可提供系統更多可用且正確的資料，增加系統效能，在系統中、後段的路線構建上，即可利用先前所分析好的資料來構建路線，可減少在路線構建時及路線完成後之改善動作，降低系統的複雜性，提升時間點分群結果的精確度。

以國內大眾運輸系統之營運時間為例，其營運時間為早上的六點至晚間十二點，本研究將資進行分群的手法有兩種：

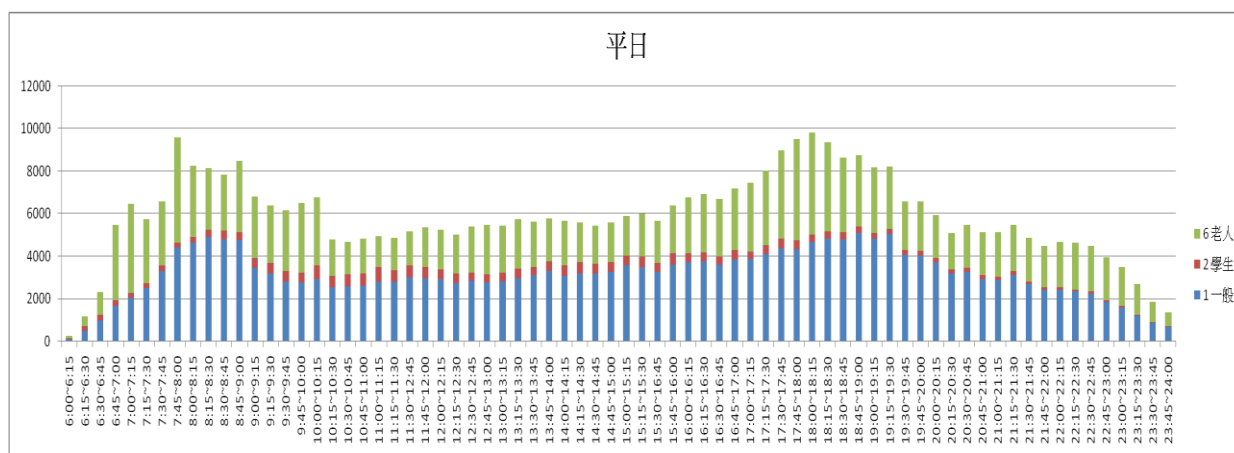


圖 3-6：旅次變化趨勢圖

1. 以十五分鐘做為一個分群基準亦可將一日十八小時的時段區間切分為七十二組時間分群，並區分平日及假日的兩種資料類別，以長條圖分析的方式觀察在各時間點中，各乘客類別使用大眾運輸的偏好趨勢。
2. 其二之分群方式則是將一日十八小時的營運時間區分為六等分：分別以A(6時~9時) B(10時~13時) C(14時~16時) D(17時~19時) E(20時~22時) F(23時~00時)等六個時段之分群手法，其主要目的是希望能夠藉由與其他的變數以關聯規則等方式做出各式運族群之通勤旅次特質。

表 3-5：時段區間分群表

時間變數	時段區間	時間變數	時段區間
A	06:00~09:59	D	17:00~19:59
B	10:00~13:59	E	20:00~22:59
C	14:00~16:59	F	23:00~00:00

### 3.3.2 空間分析-資料分群

在空間分析之部分，其欲分析之欄位則包含旅次起點、旅次迄點兩資料類型，故在選定旅次迄點之情況下，可依照旅次鏈之長度特性做為區別基準，以捷運接駁路網為例，其旅次鏈長度之差異範圍變化不大(以捷運淡水站為迄點其旅次長度至多40公里)，故在分群中無須考量極端離群值之疑慮，故本研究選擇以於切割式之手法對資料進行分群，切割式資料分群是最常被使用之分群技術，而目前最常被使用之分群技術是K-means演算法，由於它具有方法簡單、容易實作、執行速度極快（幾乎無其它分群技術比它快速）等之優點，是頗受歡迎及廣泛被使用於企業界與學術界之技術 (Tsai et al, 2006)。

#### 一、 K-means 分群方法

K-means (McQueen 1967) 在1967 年被提出，是第一個分群演算法，此方法具有分群速度較快之特性，但它有分群不均、雜訊無法濾除、分群結果不穩定等缺點 (Tsai et al. 2004)。K-means 的分群執行步驟為：(1)隨機從資料集裡選擇形心點(Centroid)，(2)根據形心點開始進行資料分群，(3)重新計算形心點，(4)若新的形心點和舊的形心點相同，資料分群即結束。所謂的分群，它是一種資料探勘之技術，其用途在於將資料依照某種條件為依據，將符合條件之資料進行資料歸類之動作。分群的方法可分成(1)切割式分群，(2)階層式分群，(3)密度式分群與(4)網格式分群，而本研究則是屬於切割式分群之方法。切割式分群是四種分群方法裡最為常見的一種分群演算法。它藉由切割空間的觀點來歸類資料，將資料空間分割成數個大小不一的子空間，在同一個子空間中的資料點便視為同一群。此類的分群演算法主要是找尋各群集之中心點，利用各資料點間距離之遠近達到分群之效果，如圖1 所示。這一類的演算法主要的優點為分群快速，但缺點為結果不穩定與無法過濾雜訊點。

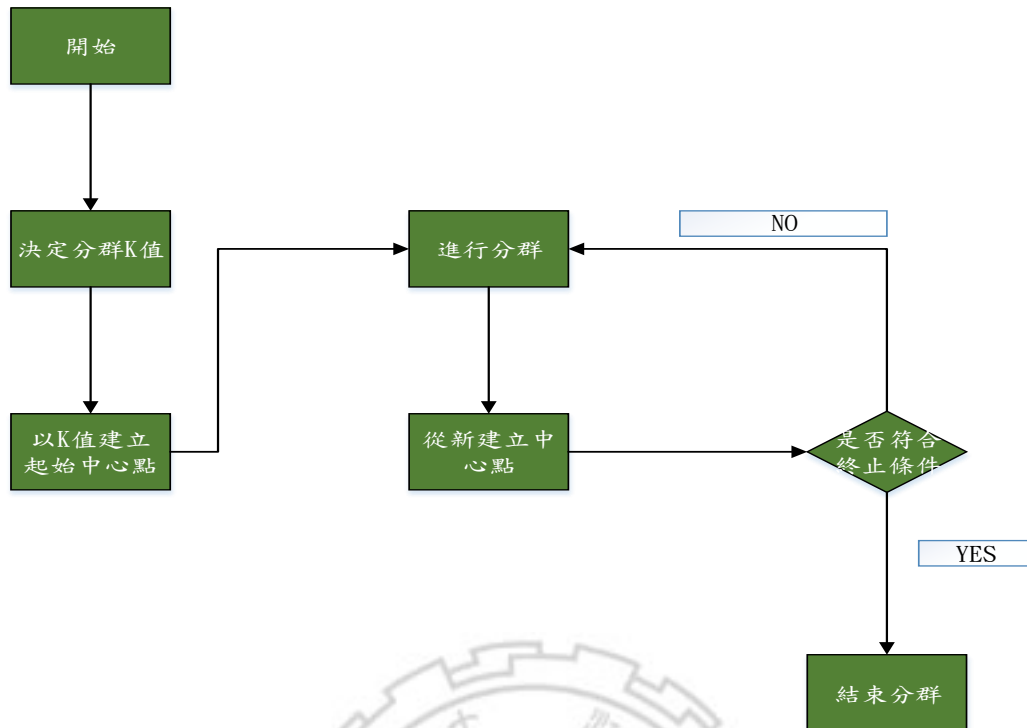


圖 3-7：K-means 演算流程

資料來源：胡建勳(2008)應用資料群聚與關聯法則於網路拍賣使用者之特性研究

故本研究根據台北捷運公司所提供之捷運各站點間距查詢資訊系統所提中之資訊，故可將台北捷運共140餘個站點資料依照其旅次長度歸納旅次鏈分群(如下圖所示)。以本研究探討捷運淡水站學生族群之通勤特質為例，將旅次迄點其將旅次的迄點定在捷運淡水站，根據各站點所對應之編號，以取代成與捷運淡水站之距離格式，以紅樹林站為例則標記為2.7KM，竹圍站則標記成6.7KM以此類推，即可將各站點至捷運淡水站的旅次長度欄位標準化，對於該欄位進行集群分析處理。

而本研究欲使用上述所介紹之K均值分群法(K-Means)進行分群演算，其K均值法並不適用於所有類型的資料集，由於該演算法是藉由資料欄位與搜尋半徑所配對的條件進行分群，假設資料欄位中涵有極端值或是離群值的存在及可能影響分群的質量，而本研究再根據旅次長度欄位標準化的處理手法，其距離欄位內所

對應的數值介於3.3至35.6，故不存在著極值影響其分群結果。

設定K(搜尋半徑數量)為7並將搜尋半徑定為5(5公里定義為一群集群體)，其原因在於旅次鏈至高長度為35.6公里，在搜尋半徑為5公里之條件下，將K值定為7則可完整將所有旅次進行分群，以助於以K-Means分群方式處理，以至淡水站點距離五公里為一組類別進行站距分群的處理手法，以10公里內的站點分群為例，其中即包含了：忠義站、復興崗站、北投站及新北投奇岩站等站點，並將其分群定義成M1的變數類別，其中M則表示為捷運的資料類別，數值5則代表站距長度5公里的群集含意，依此類推，故根據此分群手法將共將台北捷運各站點區分了M5~M35的七種站間距離分群。

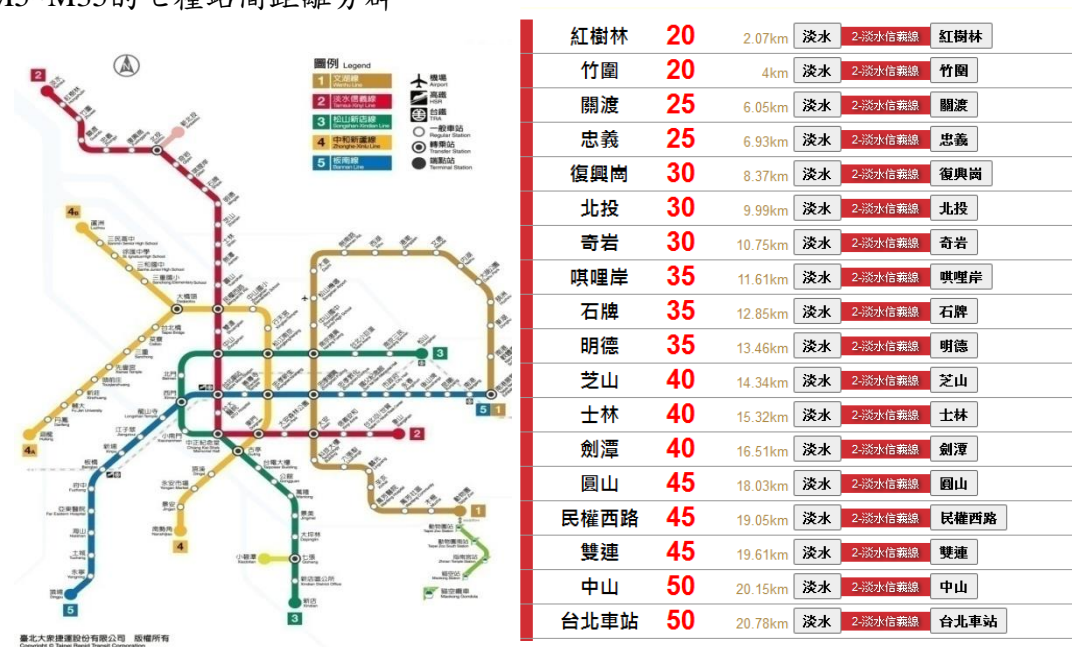


圖 3-8：站距分群換算示意

### 3.3.3 關聯法則

關聯法則主要在協助尋找資料庫中資料間的相互關係，最早由Agrawal等人於1994年所提出，是關聯規則演算法中最原始經典的計算原型。假設在交易資料庫中存在多筆的交易記錄，每筆交易記錄皆會紀錄顧客所購買的項目，關聯法則



就是用來表示項目之間的關係。在探勘關聯規則的過程中，主要分成以下兩個階段：第一階段中，先找出滿足最小支持度的項目組，這些滿足最小支持度的項目組，稱之為高頻項目組，若一個項目組包含有 $k$ 個項目，稱之為 $k-1$ 項目組，若某 $k-1$ 項目組滿足最小支持度，稱之為高頻 $k-1$ 項目組。第二階段中，以最小信賴度為條件，計算高頻項目組所形成的關聯規則，若滿足最小信賴度，則關聯規則成立。其主要為分為下述五步驟：

1. 設定門檻值。
2. 完整的比對資料庫以取得所有符合門檻值的項目集合。
3. 不為空集合時將持續進行步驟4與5。
4. 交叉組合項目集裡的所有項目來產生新的候選項目集合。
5. 候選項目集合來比對資料庫以取得所有符合門檻值的新項目集合。
6. 利用所得符合門檻值項目集合來分析得到最後的關聯式法則。

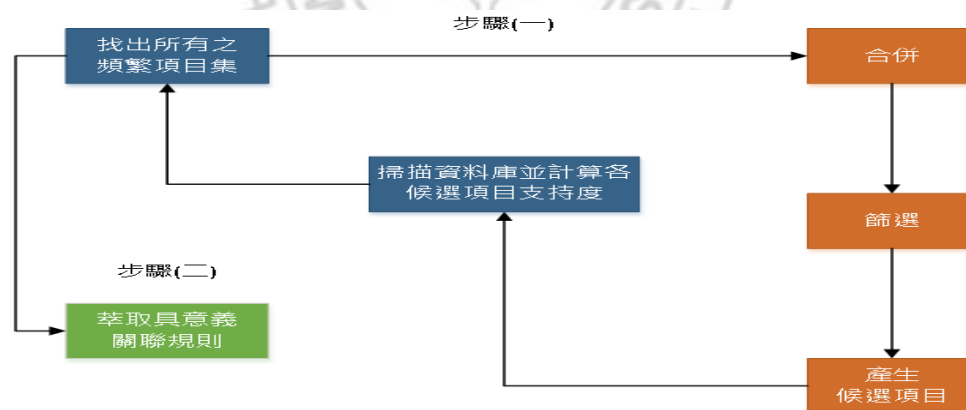


圖 3-9 Apriori 演算法運作架構

資料來源：鄭景邁(2007)資料探勘技術應用學童補助教學研究



上述所提及支持度與可靠度，其關聯性規則的產生由支持度(Support)與可靠度(Confidence)兩個參數來決定，而支持度與可靠度皆必須大於使用者訂定之最低限制，才能判定為有意義的關聯式規則(Kouris et al.,2005)，則支持度與可靠度解釋如下：

#### 一、 支持度：

定義為決策變數在資料庫中出現的比例，以 $\text{Sup}(X)$ 表示，支持度越高越值得重視。

$$\text{Support}(A \Rightarrow B) = P(A \cap B)$$

#### 二、 可靠度：

定義為關聯式規則可信的程度，亦即當某決策變數 $X$ 已確定或成立時，另一決策變數 $Y$ 發生或成立的機率，表示為 $\text{ConF}(X \rightarrow Y)$ 。

$$\text{Confidence}(A \Rightarrow B) = P(B \mid A)$$

#### 三、 相關性：

Wang et al.(2004)指出當中為了降低由支持度與可靠度可能造成的偏誤，則必須再考慮其相關性(Correlation)，進行相關性分析，即增益值(Lift)。

當增益值 $>1$ ，表示 $X$ 與 $Y$ 為正相關，規則具有實用性；

當增益值 $=1$ ，表示 $X$ 與 $Y$ 為正相關，結果與亂數取得方式相似；

當增益值 $<1$ ，表示 $X$ 與 $Y$ 為負相關，比亂數取得之結果更差。

### 3.4 私人運具旅次調查

本研究希望藉由挖掘大眾運輸電子票證資料庫後，瞭解各使用之通勤族群對於公共運輸使用的偏好及旅次特質，其分析結果可得知各族群平日旅次之進出站時間點、通勤旅次之長度、乘車時間及旅運成本票價等結果，在此運輸服務的特質下，是否具備私人運具族群不願選擇之因素，故在此分析的理論基礎下，以相同的建構條件下，利用顯示性與敘述性偏好的方式設計問卷，且希望真實呈現私人運具通勤族群在面對公共運輸回饋方案時的反應。

#### 3.4.1 問卷設計內容

本研究利用顯示性與敘述性偏好的方式設計問卷，以迄點為捷運淡水站為例，其主要概念是希望得之私人運具通勤族群每日外出的旅次起點為何，問卷主要分為四個部份，第一部份「基本資料」、第二部份為「私人運具旅次特質」、第三部份為「大眾運輸願付價格評量」與第四部份為「運量提升方案研擬」。「運量提升方案研擬」之設計基準為參照新加坡、澳洲及香港等國家之乘車回饋方案所研擬而出，以檢視出何種情境對於私人運具族群的吸引誘因是較為顯著的。「高鐵價格促銷方案選擇行為」則是蒐集小汽車與高鐵價格促銷方案偏好資料，利用旅行時間、費用、使用期限與購票通路等屬性，利用交設計產生不同的情境，分析小汽車駕駛人面臨不同高鐵價格促銷方案時的運具與促銷方案選擇行為。

##### 一、 基本資料：

藉由基本資料調查的手法得知旅運者之性別、年齡、職業、教育程度、平均月所得、家中小汽車持有數，作為分析不同社會經濟特性對私人運具與大眾運輸旅次回饋的影響；並以詢問離家最近之捷運站點，以便後續的旅運特質分析。

## 二、 私人運具旅次特質：

其主要欲藉由此題型的問向瞭解受測者每週騎車通勤的次數，及每日通勤的時數為何，平均每週加油的次數及價格為何，以立估算私人運具族群每日的旅運成本多寡，在與大眾運輸的接駁服務相較之下，其選擇私人運具下能節省多少的通勤時間，期盼能藉由分析結果彙整出私人運具族群之旅次特質關係。

## 三、 通勤旅次服務水準

此部份是衡量私人運具族群對於大眾運輸服務特性的相關看法，參考過去文獻評估大眾運輸服務品質的變數與考量本研究的相關性。根據私人運具族群過去搭乘大眾運輸的印象與經驗，對於服務品質問項進行滿意度的評估。問卷以正向方式描述，內容包含：候車時間、乘車時間、轉乘時間、接駁便利性、乘車舒適度、旅行成本考量等類別。

## 四、移轉誘因方案研擬

以新加坡、澳洲及香港等國家所實施的旅次回饋方案為設計基準，歸納下列共六種方案研擬，分別為：票價折扣優惠、學生通勤月票、紅利點數兌換、尖離峰定價、里程計價及乘車扣款上限等措施，並根據智慧卡資料庫分析之結果，將使用者之時空間分析結果轉換成各情境回饋之條件，觀察私人運具族群對於各方案之偏好變動為何，期望藉由反應最為顯著的情境下做為後續補助的研究標的。

### 3.4.2 問卷發放對象與樣本數量

#### (一) 問卷發放方式

問卷調查代表一個普遍而具體化的操作化過程，必須透過客觀、有系統的科學方法，在應用上必須瞭解實施的過程與程序，才能蒐集到可靠且有效的資料(吳明清 1991)。以問卷方式來收集資料，常使用的調查進行方式有：郵寄給

受訪者自填(Postal Questionnaires)、訪員以電話訪談(Telephone Interviewing)、訪員當面進行訪談(In-Person Interviewing)，根據上述幾種問卷調查方法可簡略綜整成表3-15進行比較，並可得知由訪員當面進行訪談調查方式雖然投入人力以及金錢花費較高，但優點眾多，不僅可以控制調查範圍及對象、回覆率也較高、訪員與受訪者較容易建立關係並適時解釋題項內容，亦可運用輔助工具，例如：利用文字、實物或圖示來幫忙表達辭義，當中由於較有彈性，就容易失去標準化的訪談方式，因此本研究將使用由訪員當面進行訪談，並且會進行訪員訓練，避免問卷調查結果產生誤差。

表 3-6：發放手法比較表

手法	郵寄給受訪者	電話訪談法	當面訪談法
調查範圍	較廣	較窄	介於中間值
調查對象	可控制 代表性難估計	可控制 代表性較強	可控制 代表性較強
回覆率	低	中	高
問項內容	簡單	簡單	深入
所需時間	長	短	短

## (二) 決定樣本數量

樣本數的決定在研究設計中，受訪者的數量要多少才算具有代表性，且要趨近於多少才算夠大，至目前似乎無一致標準。Nix和Barnette(1998)認為，如果樣本太小，容易導致統計考驗力下降，錯失發現重要研究結果；但若樣本過大，也會導致統計考驗力過當，造成無實質效益卻達到統計上的顯著水準。李茂能(2002)也指出，當樣本大小趨於無限大時，統計量不管實際效果值為多少，一定會達到統計上的任何顯著水準；而當研究的樣本很小時，統計量無論實際效果值有多大，常無法達到統計上的任何顯著水準。另外，Gorsuch(1983)建議樣本數最少為變數項的5倍，且大於100，才能進行因素分析；Comrey(1973)也提出樣本數在100以下不適合進行因素分析，一般進行因素分析的樣本最好

能夠大於300；而Gay(1992) 對於問卷樣本數的多寡提出看法，分別為下列四點：

1. 描述研究(descriptive research)時，樣本數最少占母群體的10%，如果母群體較小，則最小樣本數最好為母群體的20%。
2. 相關研究(correlational studies)之目的在於探究變項間有無關係存在，受試者至少在30 位以上。
3. 因果比較研究與許多實驗研究，各組之人數至少要有30 位。
- 4.如果實驗研究設計得宜，有嚴密的實驗控制，每組受試者至少在15 人以上，但權威學者仍認為每組受試者最少應有30 人，最為合適調查人數。則Roscoe(1975)提出四項決定樣本大小的原則：

- 一、適合做研究的數目在 30到500之間比較洽當。
- 二、當樣本被分為次樣本時，每個次樣本不應少於 30個。
- 三、從事多變量研究時，樣本數目應該大於變數數目，以10倍為佳。
- 四、對於有實驗控制的簡單實驗研究，樣本數在10至20個之間。

由於考量到實際情境與研究間差異，抽取樣本時如果侷限於樣本的「絕對數量」多寡，則較為不適宜，因為各母群體性質不同，異質性很高，包含個體數差異甚大。所以可接受之受試樣本數準則只是一種參考指標，因此針對樣本多少的認知與實務應用上，應把握「選取具代表性之樣本比選取未具代表性而眾多的樣本人數來的重要(吳明隆 2005)」之準則。綜合上述，發現在可容許誤差愈小或信心水準愈高時，則所需樣本數也愈大。因此，本研究假定可容許誤差為5%之情況下，所需樣本數經由公式(1)計算後為385份，可正確反映整個母體之信心水準達95%，其中， $n$ 為樣本數； $Z$ 為標準常態分配值； $e$ 為抽樣誤差； $p$ 為最大值。

### 3.4.3 問卷調查與抽樣方法

#### (一) 資料蒐集

本研究之主要問卷調查地點為新北市淡水區，調查時間自2016年3月15日起，至2015年4月20日止，歷時35個調查工作日，而在研究對象上需有條件限制，說明如下：

1. 18歲以上持有機車駕駛執照之男女族群
2. 就讀於淡江大學、真理大學、聖約翰大學及台北海院等四所學校學生
3. 曾經有或是現在有騎乘機車通勤上學的族群
4. 每週通勤至淡水數量達兩次以上之族群

#### (二) 抽樣方法選擇

在樣本數不變的情況，樣本異質性愈大，則代表性愈高，因為多元化的樣本可互相中和、避免過度集中的原則下，本研究係以「分層比例隨機抽樣法」作為抽樣方法，其有五個步驟可遵循：

1. 分析並決定母群的層。
2. 蒐集抽樣架構及各層之元素個數的數據資料。
3. 計算各層的元素比率。
4. 決定樣本大小。
5. 計算各層的抽樣個數。

按照上述流程，首先本研究的分層標準是依照淡水區的大專院校劃分為淡江大學、真理大學、聖約翰科技大學及台北海洋技術學院，其詳細區分如表所示，根據交通部統計部所提供之機車市佔比例及教育部104學年度全國各大專校院學生人數所提供資料，分別計算各分各校學生在學人數比例換算，其淡水

區的學生母體數為47300人，分別求得各校區之所占比率，而比率分別為淡江大學57.08%、真理大學17.98%、聖約翰科技大學11.20%以及台北海洋技術學院13.74%。因此，淡江大學應抽樣本數為213份，真理大學為67份，聖約翰科技大學42份，台北海洋技術學院則為52份，但考慮到可能有無效問卷與抽樣誤差等因素影響，所以本研究決定發放500份問卷樣本。

$$N = z^2 \times \frac{p(1-p)}{e^2} = (1.96)^2 \times \frac{0.5-(1-0.5)}{0.05^2} = 376$$

表 3-7：抽樣對項統計表

學校	學生人數	學生人數	機車市佔率	應發數量	預計發放量
淡江大學	27000	47300	38%	213	239
真理大學	8500			67	80
聖約翰大學	5300			42	50
海洋學院	6500			52	60

### 3.4.4 問卷資料分析法

本研究使用 SPSS 軟體10.0 版為主要電腦輔助分析工具。以下就本研究中所使用的統計方法予以簡要的介紹，基本資料分析部分主要是利用敘述統計，針對調查問卷的各項特性進行討論，其內容包括：性別、居住地、年齡、教育程度、職業與平均月收入等，欲計算出各項分類的統計量，並藉由上述基本資料的分析結果，了解回收樣本的分佈型態，以給予對於研究樣本的基本特質有初步瞭解。

#### A. 變異數分析

變異數分析（Analysis of variance，簡稱ANOVA）統計分析中常用的方法之一，主要應用於探討連續型（Continuous）資料型態之因變量（Dependent variable）與類別型資料型態之自變數（Independent variable）的影響關係，當自變數的因子中包含及超過三個類別條件下，即可檢定其各類別所評比之平均數是否相等的統計模式。廣義上，可將T檢定中的變異數是否相等（Equality of variance）的合併T檢定（Pooled T-test）視為變異數分析的一種，其基於T檢定為分析兩組平均數是否相等，並採用相同的統計概念，且實際上當變異數分析套用於T檢定分析時，當中產生之F值會等於T檢定的平方項。

而變異數分析中的機率分布依據，怎採用平方和（Sum of square）與自由度（Degree of freedom）所計算的組間數及組內均方（Mean of square）推算F值，若換算結果具顯著差異時，則進行事後多重比較（Multiple comparison），較常見的手法包含Scheffé's method、Tukey-Kramer method與Bonferroni correction，多數均應用於探討其各組之間的差異。

在變異數分析的基本理念下，其因子之分析手法可分為單因子變異數分



析、雙因子變異數分析及多因子變異數分析三大類型，依照資料特性不同而衍生出三種不同之分析法。

在分析法的比較上，變異數分析之結果較優於兩組比較的T檢定在於，後者研究成果會導致多重比較（multiple comparisons）的問題使型一誤差（Type one error）的比例增高，因此，比較可藉由多組平均數的差異性檢視則為變異數分析的主要目的。

故本研究中採用的單因子變異數分析來檢視，機車通勤族群之個人基本資料與(個人所得)旅次特性(通勤次數、通勤時間、搭乘公共運具頻率)資料與轉移誘因方案的「滿意程度」兩者間的平均數是否具有一致性。

一般而言，假說檢定的程序有下列六個：

（一）設立虛無假說與對立假說，並依對立假說的方向決定檢定。

（二）選統計檢定的方法。

（三）選擇需要的顯著水準，以  $p$  值為其拒絕虛無假說的機率值，用來表示實際資料的顯著水準。一般而言，檢定的顯著水準採用0.05，稱為「型一誤差」，為拒絕真實虛無假說的誤差機率值，本研究大部份皆設定型一誤差為0.01 與0.05，本研究在上下游通路成對樣本t檢定時則調整為0.1，再以 $p$  值來檢定其差異顯著水準。

（四）找出臨界值、確定拒絕域。

（五）計算實際檢定統計量。

（六）比較實際統計量與臨界值，加以檢定，做出結論。

## B. 單向變異數分析的事後檢定

當變異數分析結果顯示變之間有顯著差異時，則必須採用事後比較檢定（Post hoc tests）以檢視測試處理組與對照組兩者是否有顯著變化。透過多重比較測定（Multiple comparison analysis testing）即可檢視其結果。常用的分析技術有Tukey、Newman-Keuls、Scheffee、Bonferroni與Dunnnett法，上述每個統計技術都有其對應之資料類型及優缺點。有些統計技術用以進行比較檢定，有些手法則適用於探究新的研究成果。

### 一、使用多重比較分析檢定

當完成變異分析後，假使結果顯示出具顯著差異。則顯示顯著之變數須需要持續進行後續分析。但變異數分析不能顯示各處理間是否不同，故在執行變異數分析之後，則持續採用的檢定稱為事後比較檢定（Post hoc tests）。

在事後比較檢定中的多種比較方法，最常用的方法有Bonferroni、Sheffee、Tukey、Newman-Keuls與Dunnnett方法。每一個檢定技術都有其優點及研究限制。有些檢定技術可應用於成對比較，有些檢定技術則是由分析人員主動選擇所需要比較的處理數據群。每個方法各有其 $\alpha$ 值其特殊的檢定結果。因此分析人員必須依據自己的研究目及限制的選擇適用的處理手法。

研究資料的特性往往會影響比較方法的選擇。例如各組數據的樣本數目可能不相同。不同的檢定其檢定力的要求可能不同。有些研究要結合一些處理組數據與結合一些控制組數據，再相互比較。因此多重比較之技術其選擇是取決於探討比較的目的，有興趣的處理群與數據的型式。

## 二、比較數據組之類別

對比代表對於變異數分析後，其中兩組平均值是否具有顯著不同的檢定工作。類別有兩型：一般（Simple）與混和（Complex）。一般對比代表兩組數據之比較，例如試驗組1與控制組2。混和對比代表比較兩個不同組合之數據組是否有顯著差別。例如實驗組中第1、2、4組之結合數據，與對照組中1、3、4組之結合數據，兩組新數據相互比較。

執行變異數分析的目的也有測試原理或產生新理論。針對不同的研究目的，必須選用不同的多重比較檢定技術。

## 三、成對比較檢定

### 1. Tukey法

Tukey多重比較法是將每一個試驗組與每一個試驗組逐一比較。如果試驗組與控制組的樣本數目不相等，則結果為是最佳選擇。首先對於成對數據群具有最大不同平均值的兩組數據進行檢定。以一個q統計量以判別各組平均評比不同是否顯著。q統計量來自最大的平均值減最小的平均值，再除以所有族群其平均值之標準差。所有族群平均值的標準差之總和除以樣本數目稱為族群之均方，在許多統計軟體均有計算並加以呈現。q值再與特定的 q表查對其臨界值（Critical q-value）加以比較。如果計算之q值大於q表中的臨界值，即是代表有統計顯著性差異。

在平均值具有最大差異的兩組數據完成比較之後，再進行第二對數據之比較。此比較作業持續進行，一直到q值不高於臨界q值。

使用Tukey法的實例列舉如下：在一項研究中，以四種不同的抗生素以進行MRSA試驗（Multiple-Drug Resistant Staphylococcus Auras infection）。控制組

只有以Vancomycin處理。其他三種抗生素用於三個處理組。因為樣本數目不同，因此選用Tukey法。另一個考慮因子是因為第一型錯誤的考量比第二型錯誤更為重要。研究的重點在於新藥是否優於傳統用藥Vancomycin。

如果犯了第一型錯誤，醫生將使用新藥以取代舊藥Vancomycin，但是新藥的藥效不如舊藥，而且成本更高。如果犯了第二型錯誤，只是導致用藥不改變。Tukey法可能引起較多的第二型錯誤，但是可以協助研究者避免第一型錯誤。

Tukey法的優點是成對比較，計算方式簡單，對不相同樣本數目之處理也可進行比較。其缺點是檢定力不如其他方法，而且無法進行複雜化數據群的比較。

## 2. Newman – Keuls法

Newman-Keuls法與Tukey法非常相似，但是Newman-Keuls法對於每一個比較可以都有各自的 $\alpha$ 值。此方法因此可以執行更多的比對，因此其檢定力比Tukey法更強大。在第一次比對中，也是與Tukey法相同。然而再使用每一組數據之平均值與全部數據平均值比對。其q臨界值隨著連續之比對而降低，而Tukey的q臨界值都是維持一定。因而Tukey法對 $\alpha$ 值並不擴增，Newman-Keuls對 $\alpha$ 值則逐漸增加，第一型錯誤的機率也逐漸增加。

如果兩組數據之間相對較小的差異值也是十分重要，Newman-Keuls法則是十分適用。尤其是進行一些新的，缺乏資訊的研究項目。例如HIV是一種新症，對感染無藥物可以處理。因此只要藥物的處理結果有微小不同，也是重要的發現。因為Newman-Keuls法的較大的檢定力，此技術十分適合。而且第一型錯誤對醫療的影響並不大。

總而言之，Newman-Keuls統計技術對於微小差異之鑑定十分重要。第二型錯誤的風險大於第一型錯誤。對於新的科學領域而且缺乏研究資料之研究主

題十分有用。但是使用此方法，其各組樣本數目必須相同。

#### 四、多重組群之比較

##### 1. Scheffee 法

Scheffee法可用以比較所有可能的對比，包括簡單與複雜。所有處理都進行比較時，Scheffee法與上述兩個方法的檢定力差別不大。只有一些特定的組合需要進行比對時，Bonferroni多種比較法則是最佳方法。

如果一個理論已發展完成，Scheffee法是十分適合。可用以比對所有試驗群與各組合的不同。換言之，試驗的調查數據與一個理論評估是否有顯著不同，Scheffee法最適合驗證其是否不同（顯著差異）。

##### 2. Bonferroni 法

Bonferroni法對 $\alpha$ 值並不產生擴增，也可以進行複雜性的分析。但是Bonferroni法不適用於探索性的研究分析。研究者必須預先進行各種可能的判別。研究者對於研究主題的基本原理必須具備。因此Bonferroni主要用以確定證實原理。因為此方法有此限制， $\alpha$ 值的擴充不顯著。

##### 3. Dunnett 法

Dunnett法主要用以進行控制組之試驗設計。對於處理組與控制組的細小差異可以使用此方法加以發現。研究人員以兩組或兩組以上的試驗組與對照組進行比對。以每一個試驗組之平均值與控制組之平均值相比較。其他的檢定方法以每組平均值比較所有數據之平均值。因此Dunnett法比較容易有顯著差異。不同試驗組中，平均值與控制組平均差別愈大者，顯著性愈高。

故本研究再根據私人運具旅次調查之條件下，其問項較廣類型較多元，在資料處理上是用於多重比較的分析方式，在檢視其各方法之優缺點比較後，選擇社會科學的研究經常使用、且各組樣本數相等或不相等時可使用之多重群組比較之Scheffee做為多重比較之研究方法。

其各比較分析之優缺點比較彙整如下表所示：

表 3-8：多重比較法優缺點比較

統計技術	測試目的	優點	缺點
Tukey	對 Simple 處理群加以比對	1.數據群無意義，適用差異比較 2. 統計套裝軟體均有此功能 3. 降低第一型錯誤之風險 4.樣本數目不同可以採用	1. 無法測試複雜比較 2. 探索性試驗並不理想
Newman-Keuls	Simple 處理群加以比對	1. 比 Tukey 方法更有檢定效力 2. 在一些商業軟體可使用 3. 降低第二型錯誤 4. 易發現微小顯著性的差異	1. 無法執行複雜比對 2. 處理樣本數需要相同 3. 第一型錯誤之機率大
Scheffee	所有的處理都可檢定，包括，Simple, Complex	1.適用於探索性統計範疇 2.更具有檢定力 3.許多統計軟體均有 4.減少第二型錯誤	1.對試驗比對用途不大 2.型一誤差的相當大
Bonferroni	對選擇性之處理進行比對，包括 Simple,Complex	1. $\alpha$ 值不擴增 2. 針對控制組與不同之處理組進行比對 3. 許多統計軟體具有此功能。	1. 處理樣本數目須相同 2.須自行定義其對照組 3. 不適用於探索統計
Dunnett	在 ANOVA test 顯示有顯著差異時，以對照組與各處理組。	1. 具有高檢定力，尤其微小差異，但是顯著性存在 2. 針對處理與對照組進行測試，以確定其不同	1. 許多軟體無此功能 2. 對試驗組群無法測試。 3. 探索性研究並不理想。

### 3.5 研究假說

#### 一、「收入水準」與「情境研擬方案」

張新立(2005) 低收入之年輕族群或學生則為對機車使用有較高之依賴性，故年輕族群之收入水準會影響其運具選擇之偏好，根據第二章的國外旅次回饋案例回顧中可歸納出幾種回饋方式，本研究在參考並研擬出：乘車票價優待、里程計費優惠、尖離峰定價優惠、乘車紅利積點回饋、學生通勤月票、乘車票價上限等六種情境回饋方案，瞭解機車通勤族群對於各方案偏好及評價為何。

因此，藉由上述文獻回顧，本研究推論「收入水準」對「情境模擬方案」偏好評價會有所影響。

H1：「收入水準」對「情境研擬方案」有顯著相關

#### 二、「通勤次數」及「通勤時間」與「情境研擬方案」

蕭潤君(2014)不同運具的通勤成本包含固定成本、時間成本、燃料成本及車禍成本，顯示出通勤過程中所產生之成本影響，另根據吳羽姍(2015) 通勤成本指的不只是金錢上的支付，亦包括通勤所耗費之時間，即可推估通勤族對於成本的支出不僅在於金錢上之消耗，並涵蓋了時間上的支出。

因此，藉由上述文獻回顧，本研究推論「通勤次數」及「通勤時間」對「情境模擬方案」偏好評價會有所影響。

H2：「通勤次數」對「情境研擬方案」有顯著相關

H3：「通勤時間」對「情境研擬方案」有顯著相關

#### 三、「搭乘捷運頻率」及「搭乘公車頻率」與「情境研擬方案」

賴孟聰(2012) 指出不同涉入程度的乘客對於運具偏好亦有顯著性的差異，

黃俊擇(2014) 為提升旅運者的公車搭乘行為意向，應致力於提升其公車涉入程度，讓一般民眾體認到大眾運輸系統與生活密切地結合，故對於運具使用頻率與涉入程度具備某種關聯性。

因此，藉由上述文獻回顧，本研究推論「搭乘捷運頻率」及「搭乘公車頻率」對「情境模擬方案」偏好評價會有所影響。

H4：「搭乘捷運頻率」對「情境研擬方案」有顯著相關

H5：「搭乘公車頻率」對「情境研擬方案」有顯著相關

## 3.6 資料分析工具

### 3.6.1 R 語言

#### A.基本介紹

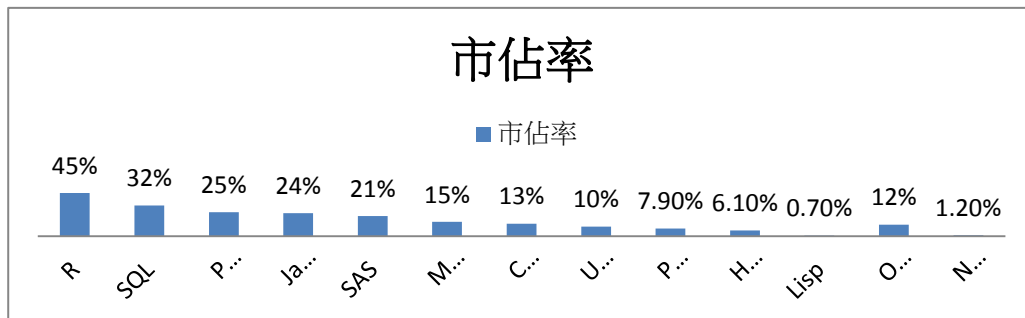
R語言是目前在統計運算和預測分析領域中最熱門的程式語言，全球有超過200萬名使用者，是專門為統計分析設計，軟體本身已經事先內建多種統計與數字分析工具，另外也可以透過安裝套件增加其他功能，如統計技術、繪圖功能，以及編程介面和資料輸出／輸入功能等等。近來R語言被廣泛使用在經濟、財經、生物分析領域中，關鍵在於軟體擁有處理大量資料的能力，加上不需要支付授權費用，自然往往成為許多研究人員的最愛。

其下列歸納幾項R語言軟件的特色：

1. 高效率的資料處理效能和存取機制。
2. 完整的數據分析功能，包含矩陣、向量、函數、因子等數種類別。
3. 出色的統計繪圖功能。



4. 多元化處理功能，透過豐富的外接擴充函數包(packages)可使用多元的資料分析方式求解。



資料來源：R數據分析方法與案例詳解

## B. 應用類別

本研究欲應用R語言進行資料採礦之研究，其中包含群集化及建立關聯規則兩種類別，在群集化的操作上以K-Means法進行分類，在關聯規則處理上則引用Apriori演算法建立其關係，而在arules外接擴充包中於本研究中則扮演相當重要的角色，arules是由Hahsler等人於 2005至2010年陸續開發出來的研究方法，主要應用於探討關聯性質的義題，同時提供了Apriori和Eclat兩種資料挖掘的演算方式。

### 3.7 小結

#### 1. 淡水區為新北市大專院校最密集的行政區之一

過往相關研究顯示，18 至 25 歲的青少年對於機車擁有高度的依賴性，由於該年齡區間多數為就讀大專院校的學生族群，而淡水當地設有淡江大學、真理大學、聖約翰大學及台北海洋技術學院等四所院校，又因捷運淡水站為雙北地區公車轉乘捷運及捷運轉乘公車資料最密集的站點之一，故本研究選擇淡水區為本研究之研究標的。

#### 2. 依照時間特性及空間資料特性進行群聚分析

以公共運輸每日服務時間 18 小時為例，依照其時段特性以三小時做為一區間，將每日服務時間切分為時段 A 至時段 F 等六種類別；在空間資料上本研究選用以 K-Means 法進行群聚分群處理，並依據通勤通勤族旅次長度之特性，每五公里做為一分群類別將旅次長度切分為 M5 至 M35 等七種類別。

#### 3. 以關聯規則演算法求得通勤族群之旅運時空偏好關聯式

關聯規則主要在協助尋找資料庫中資料間的相互關係，本研究藉由關聯規則之演算檢視通勤族群對於公共運輸服務中，其對於時間特性及空間特性之旅運偏好，藉以觀察通勤族群在特定時段下及在若干的旅次長度條件中，會選擇何種公共運具做為外出選擇，以建構旅運時空偏好關聯式。

#### 4. 檢視乘車誘因對於學生族群選用公共運具之頻率影響

根據交通部統計處所提供資料顯示，有三成的機車族群有兼用其它運具之偏好，多數為兼用公共運具族群，又因學生對於機車擁有高度之使用偏好，同時學生族群對於旅運成本變動上反應是較為敏感的，故本研究以收入水準、每週通勤次數、平均通勤時間及選搭公共運具之頻率分析乘車誘因對於通勤族影響為何。

## 第四章 實證分析-電子票證分析結果

### 4.1 智慧卡資料庫資料採礦

本研究將資料庫共分為平日及假日兩種類別，將旅次迄點定於捷運淡水站以使用者族群做為類別區分，概略分析在各時段區間內各族群之乘車偏好變化，以2014年6月為例，其假日就包含6/1、6/7、6/8、6/14、6/15、6/21、6/22、6/28及6/29共九日共2,407,636筆交易紀錄，而平日天數共有二十一日共11,428,727筆交易紀錄，其初步的研究統計結果如下：

表 4-1：平假日各族群間離峰時段使用比例

時段	各族群交易比例(平日)	各族群交易比例(假日)
<b>06:00-09:59</b>	一般 47% 學生 15% 老人 31% 其它 6%	一般 46% 學生 26% 老人 20% 其它 9%
<b>10:00-13:59</b>	一般 51% 學生 21% 老人 27% 其它 8%	一般 52% 學生 15% 老人 25% 其它 8%
<b>17:00-19:59</b>	一般 51% 學生 8% 老人 35% 其它 6%	一般 53% 學生 12% 老人 28% 其它 8%
<b>20:00-22:59</b>	一般 52% 學生 6% 老人 37% 其它 5%	一般 55% 學生 8% 老人 31% 其它 6%
<b>23:00-00:00</b>	一般 1% 學生 0% 老人 0% 其它 0%	一般 1% 學生 0% 老人 0% 其它 0%

其可觀察出一般族群之平日的尖峰時段落於06:00-09:59，而在10:00-13:59的時段中，學生乘車偏好較早晨時段來的高，而持敬老卡的使用者除了23:00-00:00的時段，其餘時段乘車比例皆相差不大，由於此分群手法僅能得知在特定的時間當中，各使用族群的偏好趨勢變化走勢，故本研究以時距分析將18小時的公共運輸營運時間中，以15分鐘做為一段區間其觀察乘車偏好變化。

### 總旅次量分析：

由下圖可之平日的乘車趨勢呈現雙峰的曲線變化，縱軸表示智慧卡交易數量橫軸表示以15分鐘為一個時段區間的時間距離變數，其圖中三種顏色各象徵不同的族群，藍色表示一般卡別使用者、黃色表示學生綠色怎表示為敬老票卡的使用者，而在整天的資料中學生族群在各時間點所呈現的變化明顯較其他兩個族群少得許多，故此可以推論在的通勤族群中，學生選搭大眾運輸的比例十分的低。

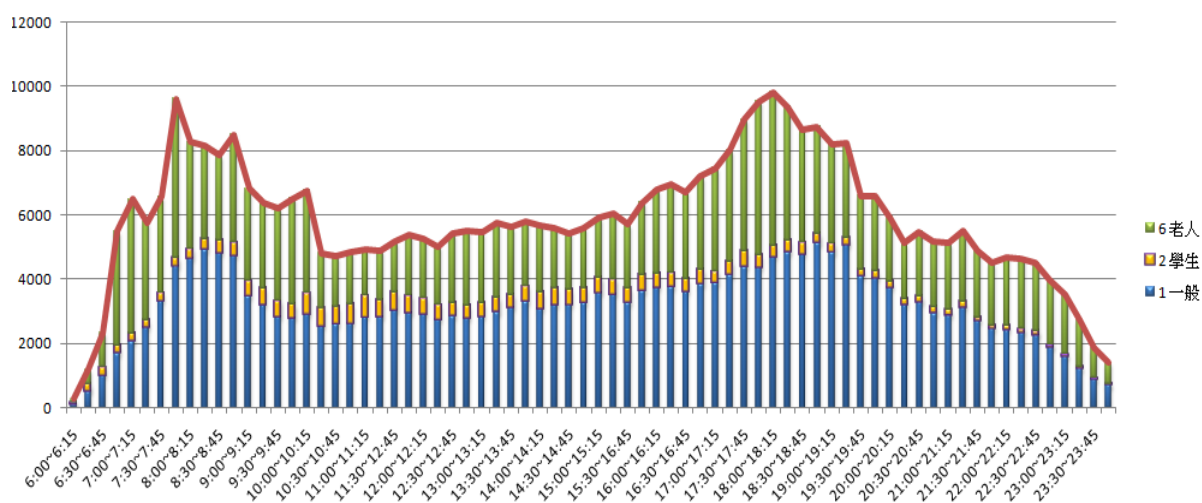


圖 4-1：時距分析趨勢圖-平日

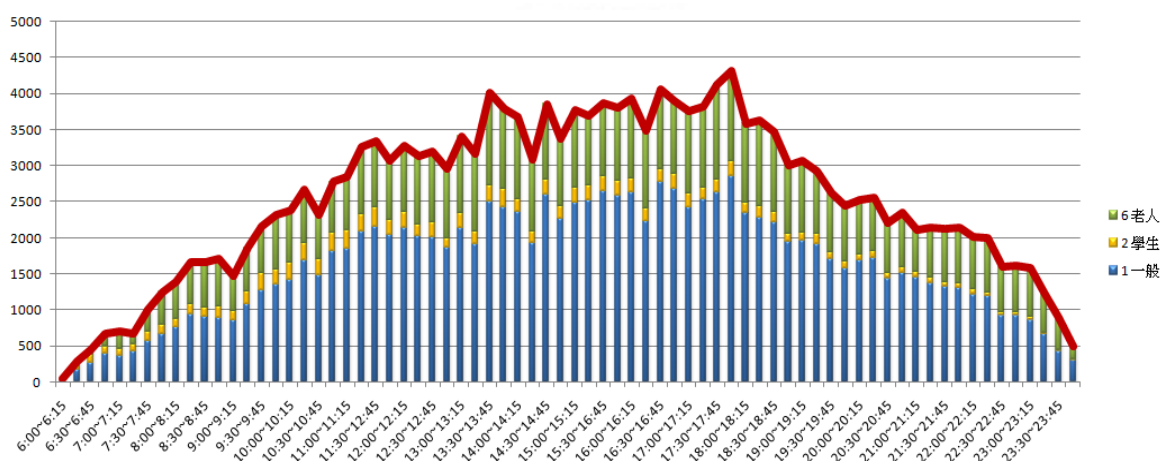


圖 4-2：時距分析趨勢圖-假日

#### A. 捷運族群

本研究以2014年6月電子票證交易紀錄做為資料分析基礎，故在當時的時空背景下僅有144個捷運站點(因迄點為捷運淡水站故不納入其中)，再根據資料分群後各旅次交易比例如下表所示。

由表中得知旅次長度最顯著的群集在十五公里的區間，其次為二十五公里的區間有26%比例，之後依序為二十公里的14%、五公里及十公里的11%，最後為三十公里及三十五公里的6%及2%，而在二十五公里的區間擁有較高比例，可能是因為該區間涵蓋較多之站點數(涵蓋34個站點)，使得分群比例較高。

表 4-2：捷運族群旅次長度分佈表

旅次長度	乘車交易量	比例	旅次長度	乘車交易量	比例
五公里	6408	11%	二十五公里	13382	26%
十公里	6208	11%	三十公里	3276	6%
十五公里	13776	27%	三十五公里	1108	2%
二十公里	7684	14%	總計	51842	100%

#### B. 公車族群

根據上述刪除靠近淡水站週邊之站點後其修正後之公車站點共253站(如下表所示)，依序將各站點的站距分群別區別為：五公里、十公里、十五公里、二十公里、二十五公里、三十公里及三十五公里(含以上)七種群聚類型。

結果顯示各分群所佔之比重依序為：五公里的72%，其次為十公里的11%、十五公里的6%、二十公里的4%、而二十五公里以上的群集僅佔了7%的比例

表 4-3：公車族群旅次長度分佈表

旅次長度	乘車交易量	比例	旅次長度	乘車交易量	比例
五公里	160068	72%	二十公里	9580	4%
十公里	24964	11%	其它	16284	7%
十五公里	12716	6%	總計	223612	100%

表 4-4：捷運站點分群表

紅樹林	M5	北門	M25	松江南京	M25
竹圍	M5	西門	M25	動物園	M35
關渡	M5	小南門	M25	萬芳社區	M30
忠義	M10	古亭	M25	萬芳醫院	M30
復興崗	M10	台電大樓	M25	辛亥	M30
北投	M10	市政府	M25	麟光	M30
新北投	M10	國父紀念	M25	六張犁	M30
奇岩	M10	忠孝敦化	M25	科技大樓	M30
唹哩岸	M15	善導寺	M25	中山國中	M30
石牌	M15	先嗇宮	M25	松山機場	M30
明德	M15	三重	M25	大直	M30
芝山	M15	菜寮	M25	劍南路	M30
士林	M15	台北橋	M25	西湖	M35
劍潭	M15	頂溪	M25	港乾	M35
圓山	M20	永安市場	M30	蘆洲	M25
民權西路	M20	景安	M30	三民國中	M25
雙連	M20	南勢角	M30	徐匯中學	M25
中山	M20	公館	M25	三和國中	M25
中山國小	M20	萬隆	M30	三重國小	M25
行天宮	M20	景美	M30	大橋頭	M25
松江南京	M25	大坪林	M30	迴龍	M35
忠孝新生	M25	七張	M30	丹鳳	M30
臺大醫院	M25	區公所	M30	輔大	M30
中正紀念	M25	新店	M35	新莊	M30
東門	M25	南港展覽	M30	亞東醫院	M30
大安森林	M25	南港	M30	海山	M30
大安	M25	昆陽	M30	土城	M35
信義安和	M25	後山埤	M30	永寧	M35
台北 101	M30	永春	M30	板橋	M30
中山	M20	龍山寺	M30	文德	M35
軟體園區	M35	江子翠	M30	內湖	M35
南京三民	M25	新埔	M30	木柵	M35

## B.公車站點分群結果

在考量通勤者的旅次起點多數非捷運淡水站周邊之站點(如:淡江大學站、真理大學站)，故將淡水站附近之站點予以刪除(共刪除120站)，以增進本研究站距分群及關聯規則分析之精確度，其欲刪除之站點歸納如下表所示。

表 4-5：公車站點刪除項目

淡江大學	濱海沙崙路口	真善美社區	獅仔頭
淡水漁市	崩山口	真理大學	港西
新市二路二段	華翠	淡金新市路口	台電北展館
台電宿舍	埔尾	台北灣社區	大福社區
聖約翰科技大學	三安里	新興國小	港西
永樂巷口	登星豪景	行忠宮	名來新城
夏威夷社區	大崙國小	櫻花村社區	乙天宮
美人嶼社區	滬尾砲台停車場	新春街	溫莎公爵
沙崙新市二路口	東元廠	自然公園	海誓山盟社區
小白宮(淡水分局)	慈雲寺	福港街	北華
鳳翔區	楊厝	渡船頭	淡水農會自強
福德宮	謝厝	濱海義山路口	高爾夫山莊
宏泰新村	小坪頂	山仔頂	桃源國中
出磺口	黑得公仔	滬尾砲台	西盛
綠堤街口	重建街口	仁愛之家	大漢
板國中	德茂	育英街口	常春藤社區
三塊厝	華南銀行	佛朗明哥	萬居
大崛	小北街	觀海社區	大坑
觀光果園	水流公	立體停車場	茂長
許厝	水源國小	甜水郡社區	自強路口
林子	八德中華路口	破布子腳	景新街
正德國中	山海關	北淡水	捷運淡水站
天元宮	體育場	義山里	坪頂國小
海景園中園	大度立德路口	埤島里	新市路
真理大學	大華柏園	商工路口	康厝
台北海院淡水校區	書香大地	紅柿子腳	德街口
淡水新城	田寮	民興里	聖心別墅

表 4-6：公車站點分群 1

公車站點	分群代號	公車站點	分群代號	公車站點	分群代號
捷運竹圍站	B5	大屯橋	B5	大片頭	B10
東宇花園城	B5	關渡	B5	牛車寮	B10
觀景台	B5	成泰路口	B5	北投國小	B10
一信幼稚園	B5	育英國小	B5	五府千歲	B10
天生國小	B5	新天母社區	B5	新興公寓	B10
外寮	B5	紅柿子腳	B5	大業中央北路	B10
果嶺大道	B5	竹圍國中	B5	成德公園	B10
天生社區	B5	八勢里	B5	中央南路	B10
米粉寮	B5	安老所	B5	後厝	B15
大溪橋	B5	新海路	B5	八里服務站	B15
頂好	B5	凱德聖道院	B5	五股服務社	B15
公 23 公園	B5	家畜所	B5	東山	B15
淡海	B5	大庄	B5	古庄	B15
淡水行政中心	B5	公司田溪橋	B5	八里國小	B15
油車口	B5	加州大廈	B5	南勢崗	B15
北勢溪口	B5	關渡大橋	B5	三和國中	B15
興化店	B5	三芝	B10	福成橋頭	B15
捷運紅樹林站	B5	承德路七段	B10	劍潭國小	B15
太子廟	B5	實踐街口	B10	四棧橋	B15
灰瑤子	B5	後洲子	B10	德音國小	B15
中山路口	B5	芝蘭路口	B10	大同路口	B15
八勢一街	B5	白沙灣	B10	捷運菜寮站	B15
大河文明	B5	田心子	B10	大窠橋	B15
下圭柔山	B5	張厝	B10	坑口	B15
淡水漁市	B5	洲仔	B10	民族路口	B15
淡金新市一路口	B5	安公廟	B10	幸福戲院	B15
台北灣社區	B5	烏山頭	B10	洲仔橋	B15
新興國小	B5	淺水灣	B10	空中大學	B15
行忠宮	B5	龍行	B10	士林簡易庭	B15
櫻花村社區	B5	灰瑤	B10	蓬萊坑	B15
新春街	B5	立農國小	B10	中廣公司	B15
竹圍國小	B5	成德社區	B10	圓山皇宮大廈	B15
福港街	B5	興華國小	B10	捷運三和國中站	B15
馬偕醫院	B5	參觀台	B10	茂興	B15



表 4-7：公車站點分群 2

公車站點	分群代號	公車站點	分群代號	公車站點	分群代號
明日世界	B15	埔坪	B20	西盛館	B20
錫板	B15	薪水居易	B20	營盤口	B20
三民高中	B15	德林寺(三和路)	B20	明志國小	B20
西雲寺	B15	北觀風景管理處	B20	大鵬里	B20
建和新村	B15	成功里	B20	明志里	B20
埤頭里	B15	新十八王公	B20	永興里	B20
民和公寓	B15	龍門路口	B20	關渡大橋一	B20
溪墘	B15	長壽西街口	B20	下泰山巖	B20
舊城里	B15	三重客運	B20	崎頭	B20
大南路口	B15	建成公園	B20	新樹民安路口	B20
八里	B15	星海別墅	B20	金陵女中	B20
大崁里	B15	就業服務處	B20	下坡角	B20
車路崎	B15	民生西路口	B20	三洋	B20
廖添丁廟	B15	圓環(鈕釦街)	B20	銘德一村	B25
安老院	B15	中山國中	B20	捷運徐匯中學	B25
葫東重慶路口	B15	成淵高中	B20	老梅	B25
明志和平路口	B15	格致中學	B20	濟安宮	B25
後店	B15	厚德派出所	B20	捷運蘆洲站	B25
挖仔尾	B15	防癆協會	B20	橫溪橋	B25
圓山頂	B15	民族承德路口	B20	中興街口	B25
承德公館路口	B15	中正智成街口	B20	成州郵局	B25
民族重慶路口	B15	南京西路口	B20	臺北車站	B25
大龍洞崎	B15	聖心女中	B20	三民高中站	B25
龜子山橋	B15	頂新莊	B20	富貴角燈塔	B25
情人湖路口	B15	酒泉重慶路口	B20	菜寮(重新路)	B25
士林監理站	B15	聯邦新城	B20	金山	B25
富基漁港	B20	新庄子	B20	頭前	B25
鄉野俱樂部	B20	橋頭	B20	昌吉重慶路口	B25
社后	B20	中山藝術公園	B20	鶯江國小	B25
大同國小	B20	大庭新村	B20	西門	B25
崎頂	B20	三芝遊客中心	B20	新海橋頭	B25
正義重新路口	B20	同榮里	B20	十八王公	B25

表 4-8：公車站點分群 3

公車站點	分群代號	公車站點	分群代號	公車站點	分群代號
草里	B30	丹鳳一	B30	泰山巖	B35
石門洞	B30	捷運輔大站	B30	泰山區公所	B35
野柳地質公園	B30	捷運丹鳳站	B30	中角	B35
樹林後火車站	B30	大武崙	B30	武嶺街口	B35
板橋公車站	B30	清水路 41 巷口	B35	加投里	B35
文化街二	B30	金山區公所	B35	武聖街口	B35
山佳國小	B30	金山中學	B35	加油站	B35
石門區公所	B30	跳石	B35	大武崙工業區	B35
東榮街口	B30	筠園	B35	文程路口	B35
新興街口	B30	噴水米廠	B35	龜吼漁港	B35
樹林農會	B30	金山青年活動中心	B35	忠聯倉儲	B35
捷運新埔站	B30	中角國小	B35	萬里區公所	B35
聯合醫院 (板橋院區)	B30	乾華	B35	捷運泰山貴和站 (明志路)	B35
石門	B30	野柳	B35	興寮里	B35
板橋花市	B30	大鵬國小	B35		
尖子鹿	B30	泰山高中	B35		

#### 4.1.2 關聯規則檢定結果

本研究欲藉由挖掘智慧卡資料庫建立其大眾運輸族群之旅次特質關係式，以瞭解在相同的旅次長度之服務下，大眾運輸族群選擇運具之偏好因素，期望能蒐尋到可能會影響運具選擇之變因為何。

##### ● 捷運族群旅次特質關聯結果

從捷運族群之關聯條件式我們可以看到，多數的通勤族群旅次長度落在10公里及15公里這兩種區間當中，根據運具組合與旅次偏好之組合；亦可從中得知幾種偏好組合條列如下：

1. 在旅次起始時間中，有41.3%的學生族群外出時段為6點至9點；有30.2%的學生族群外出時段為10點至13點。
2. 在旅次長度的部分，十五公里的有25.2%、二十公里的有15.1%、二十五公里有25.5%的比例。
3. 故學生捷運族群的旅次特質為，早上六點至九點出門，且屢次長度落在十五公里的長度區間當中。

表 4-9：捷運族群旅次關聯特質

配對	組合	support	confidence	lift
時段=06-09	族群=學生	0.413	1	1.17
時段=10-13	族群=學生	0.302	1	1.04
旅次長度=15 公里	族群=學生	0.364	1	1.12
旅次長度=20 公里	族群=學生	0.151	1	1.02
旅次長度=25 公里	族群=學生	0.255	1	1.09
族群=學生,時段 06-09	旅次長度=15 公里	0.273	1	1.43

故此我們可以推論出搭乘捷運通勤之學生族群，其通勤旅次之起點多位於北投區及士林區一帶，根據關聯規則分析結果可以看出，多數的學生通勤旅次長度皆落15公里的範圍當中，在15公里的站點則有：捷運唹哩岸站、捷運石牌站、捷運明德站、捷運芝山站、捷運士林站及捷運劍潭站等站點，再藉由旅次量表的結果得知多數的學生通勤族的旅次起點為捷運石牌站、接著依序為捷運士林站、捷運芝山站及捷運劍潭站。

表 4-10：通勤旅次量表

站名	旅次量	站名	旅次量
唹哩岸	872	芝山	2028
石牌	4756	士林	2836
明德	1268	劍潭	2016

資料來源：本研究整理

根據台北捷運公司網站所提供之票價及乘車時間資料換算，在通勤費用的類別上，在15公里的通勤族群則需花費35至40元的費用才可以達捷運淡水站，在通勤時間的類別上，15公里的通勤族群平均之乘車時間約為13.17分鐘。

## ● 公車族群旅次特質關聯結果

由公車通勤族群的旅次關聯規則中，可以出多數的旅次長度皆落於5公里的區間中，且隨著旅次長度的增加，使用者搭乘的比例也相對的減少，且多數的乘車時間點則落在6點至9點、10點至13點、14點至16點及17點至19點四個時段區間當中，符合原先預期，故可以推論出多數的公車族群的旅次長度為中短程的接駁性質，且在早上及下午各時段之乘車比例差異不大。

1. 在學生公車旅次中分析結果顯示，有71.2%的旅次長度為5公里，13.8%的旅次長度為10公里；而在5公里的旅次長度下，有15.7%的民眾於時段6點至9點搭車、19.8%的民眾於時段10點至13點搭車、16.4%的民眾於時段14點至16點搭車、13.6%的民眾於時段17點至19點搭車。
2. 故可以得知公車族群的乘車時段無間離峰之差異性，且較於擁有A級路權的捷運軌道系統，公車仰賴的是一般的市區道路，容易受到道路系統中其它運輸工具及號誌等影響而增加旅行時間，對於時間偏好變動較敏感的通勤族來說，其選擇比例會低於捷運系統是可想而知的。

表 4-11：公車族群通勤旅次關聯規則

運具組合	偏好	Support	Confidence	lift
學生族群	旅次長度=5 公里	0.712	1	1.058
學生族群	旅次長度=10 公里	0.138	1	1.043
時段=6-9	旅次長度=5 公里	0.157	1	1.043
時段=10-13	旅次長度=5 公里	0.198	1	1.043
時段=14-16	旅次長度=5 公里	0.164	0.745	1.047
時段=17-19	旅次長度=5 公里	0.136	0.789	1.109

根據台北市聯營公車所提供之票價及乘車時間資料換算，在通勤費用的類別上，5公里的通勤族群則需花費12至15元的費用可以達捷運淡水站，在10公里及15公里的族群則需要兩段票費用24至30元的費用才可抵達捷運淡水站。

## ● 公共運輸族群旅次特質關聯結果

在分析公共運輸運輸旅次特質的關聯的處理上，本研究選用上述兩種公共運輸系統(捷運、公車)之通勤旅次資料處理及合併，在資料欄位的處理上，以上述分群處理的概念相通，在乘車時段一樣是分為六個時段，在旅次長度的部分也同樣分為七種級別，期盼在相同的資料格式下能求得淡水區大眾運輸學生通勤族群之旅次特質關聯。

由下表可之在旅次長度為五公里時，通勤族較偏好選擇公車，而在旅次長度為十五公里以上時，則較偏好選擇捷運，其主要的原因可能在於相較於公車系統，捷運所提供的是線的服務，對於中長程的通勤旅次上捷運的服務水準比公車來得高，不論是乘車時間、候車時間、轉乘時間或是舒適乘度皆比公車系統來得優秀。

由此我們可以推估，在公共運輸使用者當中，假使通勤旅次長度為短程時，較偏好短距離的公車接駁服務做為旅次選擇；而在中長程通勤的旅次上，民眾則偏好選擇捷運作為外出的代步工具。

表 4-12：大眾運具族群通勤旅次關聯式

組合	偏好	confidence	lift
旅次長度=5 公里	公車	0.715	2.072
旅次長度=15 公里	捷運	1	1.932
旅次長度=25 公里	捷運	1	1.932

## 4.2 小結

### 1. 學生族群搭乘公共運輸比例仍有一定成長空間

由平假日之各族群乘車使用比例得知不論平日及假日各時段區間中，於捷運淡水站選搭公共運具之學生族群仍不及持一般卡之民眾及敬老卡的族群，由於平日搭乘捷運進出淡水知民眾理應主要為學生族群為主，其僅近兩成的公共運輸使用比例可以見得，學生族群選搭公共運具之比例仍有持續成長之空間。

### 2. 多數捷運族群之通勤長度介於十至十五公里區間

在空間面分析，根據各站點至淡水站之旅次長度進行分群，在捷運通勤資料顯示多數通勤族群の旅次長度落在15公里的區間當中，在此區間包含了：唹哩岸、石牌、明德、芝山、士林及劍潭等站點，其中以石牌站的乘車資料最為顯著(佔34%比重)，其次依序為士林、芝山、劍潭、明德、唹哩岸；而在公車族群的資料顯示，多數通勤族之旅次長度則落於5公里的區間當中。

### 3. 捷運族群及公車族群各有不同的旅次時間特性

在時間面分析上，其捷運通勤族群の旅次起始時間落在6點至9點及10點至13點的時段區間當中；而在公車通勤族的起始時間則分佈在6點至9點、10點至13點、14點至16點及17點至19點四個時段區間當中，其中以10點至13點(19.8%)之乘車比例最高。

## 第五章 實證分析-私人運具旅次調查結果

### 5.1 敘述性統計結果

本研究於2016年3月5日起，至2016年4月5日止，共計發放503份問卷，經刪除填答錯誤、遺漏等不符合問卷規範之無效問卷後，本研究有效問卷數量為474份，整體問卷回收率為94.23%。以下將進行受訪者之樣本結構描述，在個人基本資料內容包括性別、年齡、平均收入、家中機車持有數、居住狀況等問項、在通勤旅次調查則包含了每周騎車通勤次數、平均騎乘時間、騎車可節省時間、每周搭乘公共運具(公車、捷運)之頻率，描述方式以樣本數百分比呈現。

#### ● 個人基本資料

##### 一、性別

由交通部統計處自用小客車使用調查報告顯示，私人運具持有人有69.8%是男性、30.2%是女性，由於不確定實際在馬路上出現的小客車駕駛人的性別比例是多少，但可以預期必定是男性比例會高於女性。

在有效樣本中，性別方面，男性人數為311人，女性為163人，男性占總樣本比例為65.61%，女性占總樣本比例為34.39%，由總樣本數中可得，由總樣本數中可得出每日往返淡水機車通勤族群之男性比例高於女性，此表示為合理，統計資料如表所示。

表 5-1：基本敘述性統計-性別

受訪者基本資料		次數	百分比(%)
性別	男性	311	65.61%
	女性	163	34.39%
	總計	474	100%

## 二、年齡

在有效樣本中，18-20歲人數有161人，21-25歲人數有227人，26-30歲人數有86人，其主要騎乘機車的通勤族年齡層落於18-20歲及21-25歲這區間當中，共占總樣本比例81.86%，同時也驗證了張新立(2005)指出低收入之年輕族群或學生對機車使用依賴較高之研究成果，統計資料如表所示。

表 5-2：基本敘述性統計-年齡

受訪者基本資料		次數	百分比
年齡	18-20 歲	161	33.97%
	21-25 歲	227	47.89%
	26-30 歲	86	18.14%
	總計	474	100%

## 三、平均月收入

平均月收入分布情形方面，1萬元以下為218人，1萬至2萬元為171人，2萬至3萬元為51人，3萬元以上元為34人，由於抽樣對象以學生族群為主，且包含日間部及進學部的兩種族群，故在月平均收入的類別增至2-3萬及3萬以上兩類別，而結果顯示多數所得水準皆落於1萬元以下及1至2萬這兩類別中，此月收入人數占總樣本比例分別為45.99%及36.08%，如表所示。

表 5-3：基本敘述性統計-平均月收入

受訪者基本資料		次數	百分比
平均月收入	1 萬以下	218	45.99%
	1-2 萬	171	36.08%
	2-3 萬	51	10.76%
	3 萬以上	34	7.17%
	總計	474	100%



#### 四、家中機車持有數

在家中機車持有數分布面，持有一輛為172人，兩輛為193人、三輛為83人，四輛及以上有26人，結果顯示多數通勤族群家中車輛持有數皆落於一輛及兩輛這兩類別當中，佔總樣本比例77%之比重，結果如表所示。

表 5-4：基本敘述性統計-家中機車數

受訪者基本資料		次數	百分比
家中機車數	1 輛	172	36.29%
	2 輛	193	40.71%
	3 輛	83	17.51%
	4 輛及以上	26	5.48%
	總計	474	100%

#### 五、居住狀況

在居住狀況的部分，根據結果顯示多數的通勤族群皆非外宿族群，在與家人與親友同住的選項有377人，而住在外租屋(包含共宿、獨自租屋)僅有97人，結果如下表所示。

表 5-5：基本敘述性統計-居住狀況

受訪者基本資料		次數	百分比
居住狀況	與家人親友同住	377	79.53%
	在外租屋	97	20.47%
	總計	474	100%

#### ● 通勤旅次特性調查

##### 一、每周騎車通勤

在每周騎車通勤中，可以看出多數的通勤族群每周使用機車往返淡水的次數皆在三次以上，其佔了總體比例近85%之比例，由此可以看出多數的機車通勤族群十分依賴機車做為日常通勤的運具選擇，結果如下表所示。

表 5-6：通勤旅次調查統計-每周騎車通勤數

通勤旅次特質調查		次數	百分比
每周騎車通勤	幾乎每天	260	54.85%
	每周 3-4 次	147	31.01%
	每周 1-2 次	67	14.14%
	總計	474	100%

## 二、平均通勤時間

在通勤時間的結果當中，可以看出多數通勤族群能夠接受最久的乘車時間落於40分鐘內，在30分鐘以內及30至40分鐘兩類別中及佔據了73.4%比重，由於通勤時間與旅次長度成正向變動之關聯，故可以假設通勤時間越長期通勤旅次也相對越長，其結果如下表所示。

表 5-7：通勤旅次調查統計-每日通勤時間

通勤旅次特質調查		次數	百分比
每日通勤時間	30 分鐘以內	185	39.02%
	30 分至 40 分	163	34.38%
	40 分至 50 分	87	18.35%
	50 分鐘以上	39	8.22%
	總計	474	100%

在平均通勤節省時間上，主要是詢問通勤族群認為相較於搭乘公共運具，平均每日騎乘能節省您多少的旅次時間，而過半的使用者認為節省時間為30分鐘以內，其機車族群會因為個人騎車速度而影響，又因公共運輸總旅次時間包含了候車時間、乘車時間及轉乘時間，故會因為旅次起點週邊之公共運輸可及性而影響到通勤節省時間之差異，其分析結果如下。

表 5-8：通勤旅次調查統計-通勤節省時間

通勤旅次特質調查		次數	百分比
通勤節省時間	30 分鐘以內	238	50.21%
	30 分至 40 分	113	23.84%
	40 分至 50 分	87	18.35%
	50 分鐘以上	36	7.59%
	總計	474	100%

### 三、搭乘公共運輸頻率

在搭乘頻率上，主要是以瞭解機車通勤族群對於使用大眾運具之態度，藉由不同乘車頻率的差異以推估機車族群是否具備選擇公共運具之可能性，結果顯示近半數的機車族的使用頻率為每個月三至四次，其受測結果如下表所示。

表 5-9：通勤旅次調查統計-搭乘公共運輸頻率

通勤旅次特質調查		次數	百分比
搭乘公共運輸 頻率	每週一至兩次	85	17.93%
	每月三至四次	220	46.81%
	每月一至兩次	167	35.26%
	總計	474	100%

## 5.2.2 比較平均數法-變異數分析

### i. 單因子變異數分析：「收入水準」與「研擬方案」

1. 首先在變異數同質性檢定的部分，票價優惠、里程優惠及乘車上限的統計量分別為 1.63、2.18、0.45，對應顯著值依序為 0.181、0.097 及 0.721，皆大於 0.05 表未達顯著水準，屬各組變異數相等，可進行變異數分析。

表 5-10：變異數同質性檢定表-收入水準-研擬方案

變異數同質性檢定		
	Levene 統計量	顯著性
票價優惠	1.63	0.181
里程計費	2.18	0.097
乘車上限	0.45	0.721

2. 在變異數分析摘要表內， $F=4.544$ (里程優惠)， $F=8.482$ (乘車上限)，兩方案之  $P$  值均小於 0.05，達到顯著水準。因此可得到以下結論：「不同收入水準之族群，對於方案選擇之評價，有顯著差異」，須再進行事後多重比較。

表 5-11：單因子變異數分析表-收入水準-研擬方案

單因子變異數分析					
		平方和	平均平方和	F	顯著性
里程優惠	組間	15.947	5.316	4.544	0.004
	組內	520.609	1.17		
票價回饋	組間	5.876	1.959	1.807	0.145
	組內	482.297	1.084		
乘車上限	組間	27.786	9.262	8.482	0**
	組內	485.902	1.092		

1. 接著由事後多重比較可知，在里程計費的方案研擬當中，平均收入(J)兩萬至三萬與平均收入(I)一萬以下有顯著差異( $P=0.006<0.05$ )，及平均收入(I)一萬至兩萬有顯著差異( $P=0.046<0.05$ )；在乘車上限優惠的方案研擬當中，平均收入

(J)三萬以上與平均收入(I)一萬以下有顯著差異( $P=0.002<0.05$ )，及平均收入(J)三萬元以上與平均收入(I)兩萬至三萬有顯著差異( $P=0.007<0.05$ )。

2. 平均收入一萬元以下族群(平均數=3.92)，比平均收入三萬至五萬之族群(平均數=3.84)，對於更偏好里程優惠之選擇；平均收入一萬以下族群(平均數=3.43)，比平均五萬以上之族群(平均數=3.25)，對於更偏好里程計費選擇。

表 5-12：多重比較分析表-收入水準-研擬方案

研擬方案	(I)平均收入	(J)平均收入	顯著性	95% 信賴區間	
				下界	上界
里程計費	兩萬至三萬	一萬元以下	0.006	-0.92	-0.11
		一萬至兩萬	0.046	-0.96	0.01
		三萬以上	0.982	-0.99	0.73
乘車上限	三萬以上	一萬元以下	0.002	-1.86	-0.3
		一萬至兩萬	0.155	-1.45	0.15
		兩萬至三萬	0.007	-1.77	-0.1

3. 故將平均收入一萬元以下定義為族群 1，收入一萬至三萬定義為 2，三萬至五萬定義為 3，五萬元以上定義為 4。在里程計費的情境下，相較於所得三萬至五萬之族群，所得一萬元以下之族群較偏好優惠方案；而在乘車上限的情境下，所得三萬至五萬之族群之偏好高於所得五萬之族群，其比較結果如表所示。

表 5-13：偏好比較分析-收入水準-研擬方案

		平均評分	F	顯著性	比較結果
里程計費	一萬元以下	3.92	4.544	0.004	1>3
	三萬至五萬	3.84			
乘車上限	一萬元以下	3.43	8.482	0	1>4
	五萬元以上	3.25			

ii. 「每週騎乘機車通勤次數-與「研擬方案」

- 變異數同質性檢定，里程計費及乘車上限的統計量分別為 5.46、1.87、2.14，對應顯著值依序為 0.11、0.16、0.12 皆大於 0.05 表未達顯著水準，故各組變異數可視為相等，可持續進行變異數分析。

表 5-14：變異數同值檢定表-通勤次數-研擬方案

變異數同質性檢定		
	Levene 統計量	顯著性
票價優惠	5.46	0.11
里程計費	1.87	0.16
乘車上限	2.14	0.12

- 在變異數分析摘要表內，票價優惠、乘車上限之顯著性值皆小於 0.05，達到顯著水準，而里程計費之顯著值大於 0.05 表是個群體間隻平均數無差異。因此可得到以下結論：「相異通勤次數族群，對於票價優惠及乘車上限兩方案之評價有顯著差異」，故須再進行事後多重比較。

表 5-15：單因子變異數分析表-通勤次數-研擬方案

單因子變異數分析	平方和		平均平方和	F	顯著性
票價折扣	組間	7.645	3.823	7.98	0**
	組內	237.613	0.479		
	總和	245.259			
里程計費	組間	2.609	1.305	1.709	0.182
	組內	378.625	0.763		
	總和	381.234			
乘車上限	組間	9.227	4.613	6.817	0.001
	組內	335.679	0.677		
	總和	344.906			

- 在票價優惠的比較上，從表中可以得知每週通勤次數 1 至 2 天族群與幾乎每天通勤的族群有顯著關聯；每週通勤次數 1 至 2 天及每週 3 至 4 天之通勤族群有顯著關聯性。

4. 而在乘車上限比較上，每週通勤 1 至 2 天族群與幾乎每天的族群有顯著關聯；  
每週通勤 3 至 4 天族群與每週 1 至 2 天族群有顯著關聯。

表 5-16：多重比較分析表-通勤次數-研擬方案

Scheffe 法					
依變數	(I) 通勤次數	(J) 通勤次數	顯著性	95%信賴區間	
				下界	上界
票價折扣	每週 1 至 2 天	幾乎每天	0.001	0.1	0.46
		每週 3 至 4 天	0.008	0.06	0.51
乘車上限	每週 1 至 2 天	幾乎每天	0.001	0.11	0.54
		每週 3 至 4 天	0.042	0.02	0.43

1. 在票價優惠類別中，每週通勤 1 至 2 天族群(平均數=4.09)，幾乎每天之族群(平均數=2.57)，更偏好學生通勤票之選擇；在乘車上限中，每天通勤族群(平均數=3.76)評價，比通勤 1 至 2 天族群(平均數=2.68)來得高。
2. 故將每天通勤定義為族群 1，每週 3 至 4 天定義為 2，每週 1 至 2 天定義為 3，其比較結果如表所示。

表 5-17：偏好比較分析-通勤次數-研擬方案

		平均評分	F	顯著性	比較結果
票價折扣	幾乎每天	2.57	7.98	0	2>1 3>2
	每週 3 至 4 天	3.69			
	每週 1 至 2 天	4.09			
乘車上限	幾乎每天	3.76	6.817	0.001	1>2 3>2
	每週 3 至 4 天	2.57			
	每週 1 至 2 天	2.68			

### iii. 「平均通勤時間」與「研擬方案」

1. 變異數同質性檢定里程計費統計量為 11.74，對應顯著值依序為 0.13 大於 0.05 未達顯著水準，故各組變異數可視為相等，可持續進行變異數分析。

表 5-18：變異數同質性檢定-平均通勤時間-研擬方案

變異數同質性檢定		
	Levene 統計量	顯著性
票價折扣	8.45	0.045
里程計費	11.74	0.13

2. 在變異數分析摘要表內，票價優惠、里程計費之顯著性值皆小於 0.05，達到顯著水準。因此可得到以下結論：「相異通勤時間族群，對於票價優惠及里程計費兩方案之評價有顯著差異」，故須再進行事後多重比較。

表 5-19：單因子變異數分析-平均通勤時間-研擬方案

單因子變異數分析	平方和		平均平方和	F	顯著性
票價優惠	組間	8.243	2.748	1.984	0.116
	組內	338.706	0.761		
	總和	346.949			
里程計費	組間	12.487	4.162	3.611	0.013
	組內	513.018	1.153		
	總和	525.506			

3. 在里程計費比較上，從表中可以得知通勤時間 40 分至 50 分族群與 30 分鐘以內的族群有顯著關聯。

表 5-20：多重比較分析表-平均通勤時間-研擬方案

依變數	(I) 通勤時間	(J) 通勤時間	顯著性	95%信賴區間	
				下界	上界
里程計費	30 分以內	30 至 40 分	0.946	-0.47	0.3
		40 至 50 分	0.03	-0.86	-0.03
		50 分以上	0.226	-0.64	0.09

- 在里程計費的類別中，40 分至 50 分(平均數=3.4)族群之評價，高於 30 分以內(平均數 2.96)之評價。
- 故將 30 分內族群定義為族群 1，30 分至 40 分定義為 2，40 分至 50 分定義為 3，50 分鐘以上定義為 4，其比較結果如表所示。



表 5-21：偏好比較分析-通勤次數-研擬方案

		平均評分	F	顯著性	比較結果
里程計費	30 分鐘內	2.96	3.611	0.013	3>1
	30 至 40 分	3.04			
	40 至 50 分	3.4			
	50 分以上	3.23			

iv. 「搭乘大眾運輸頻率」與「情境研擬方案」

- 變異數同質性檢定，票價優惠里程計費及乘車上限統計量分別為 1.632、2.118、0.445，對應顯著值依序為 0.181、0.097、0.721 皆大於 0.05 表未達顯著水準，故各組變異數可視為相等，可接續進行變異數分析。

表 5-22：變異數同質性檢定-搭乘頻率-研擬方案

變異數同質性檢定		
情境方案	Levene 統計量	顯著性
票價優惠	1.632	0.181
里程計費	2.118	0.097
乘車上限	0.445	0.721

- 在變異數分析摘要表內，票價優惠、里程計費顯著性值皆小於 0.05，達到顯著水準。因此可得到以下結論：「相異選搭大眾運具之族群，對於票價優惠及里程計費兩方案之評價有顯著差異」，故須再進行事後多重比較。

表 5-23：單因子變異數分析-搭乘頻率-研擬方案

單因子變異數分析	平方和		平均平方和	F	顯著性
票價優惠	組間	15.947	5.316	4.544	0.004
	組內	520.609	1.17		
	總和	536.557			
里程計費	組間	27.786	9.262	8.482	0
	組內	485.902	1.092		
	總和	513.688			
乘車上限	組間	5.876	1.959	1.807	0.145
	組內	482.297	1.084		
	總和	488.174			

3. 在票價優惠類別中，每週一至兩次(平均數=3.98)族群之評價，高於每月三至四次族群(平均數=3.25)，在里程計費的類別中，每週一至兩次(平均數=3.81)族群之評價，高於每月三至四次族群(平均數=3.13)。

表 5-24：多重比較分析表-搭乘頻率-研擬方案

Scheff 法					
依變數	(I) 搭乘捷運次數	(J) 搭乘捷運次	顯著性	95%信賴區間	
				下界	上界
票價優惠	每月一至兩次	每週一至兩次	0.006	-0.92	-0.11
		每月三至四次	0.057	-0.88	0.01
里程計費	每月三至四次	每週一至兩次	0.003	0.11	0.75
		每月一至兩次	0.064	0.3	1.86

4. 故將每週一至兩次族群定義為族群 1，每月三至四次定義為 2，每月一至兩次定義為 3，其比較結果如表所示。

表 5-25：多重比較分析表-搭乘頻率-研擬方案

		平均評分	F	顯著性	比較結果
票價折扣	每週一至兩次	3.98	4.544	0	1>3
	每月三至四次	3.25			
	每月一至兩次	2.97			
里程計費	每週一至兩次	3.81	8.482	0001	1>2
	每月三至四次	3.13			
	每月一至兩次	2.2			

### 5.2.3 單因子變異數分析小結

根據上述各分析結果其歸納，票價優惠的情境條件能吸引每周騎車通勤一至兩次及搭乘大眾運具每週一至兩次的族群；在乘車扣款上限的情境下，能吸引所得一萬元以下幾乎每天騎車通勤的族群；而在里程計費的情境下則能吸引，所得一萬元以下、平均通勤時間為四十至五十分、每周搭乘大眾運具一至兩次的族群，其歸納結果如下表所示。

表 5-26：研擬方案比較表

方案研擬	所得水準	平均通勤時間	每週通勤次數	搭乘大眾次數
票價優惠	X	X	每週一至兩次	每週一至兩次
乘車上限	一萬元以下	X	幾乎每天	X
里程計費	一萬元以下	四十至五十分	X	每週一至兩次

根據問卷調查結果我們得知，在不同乘車誘因下能吸引不同特質的通勤族群，以里程計費為例其能吸引包含了低所得、長距離通勤族及鮮少使用大眾運具之族群，固本研究建議應優先給予高通勤需求及長距離通勤距離之族群較大之優惠空間，較能吸引通勤族增加使用公共運具之意願。

## 第六章 結論與建議

本研究運用電子票證大數據資料瞭解新北市淡水區學生大眾運輸通勤族群在時間面及空間變的變化情形，建立一道從資料清理轉置、時間面群集、空間面群集、通勤旅次關聯性分析與運具偏好分析之流程與架構。首先，由整體面的資料分析對所有大眾運輸族群的乘車偏好進行探討，瞭解平假日各族群之偏好差異為何、平假日尖離峰旅次變化現象、各時段區間不同使用族群的乘車比例。其次以歸納分析大眾運輸族群於各站點轉乘比例變化行為，以觀察在各大轉乘站中其當地交通環境背景與通勤使用者的關聯性。最後，本研究選定大眾運輸轉乘比例最高之捷運淡水站為研究地點，同時，淡水區為北區大專院校分布最密集之集散地，故選擇乘車旅次量最低之學生族群為研究對象，建立時間分群及距離分群予以分析學生通勤族群對於大眾運具之偏好及旅次特質，並歸納其研究成果，謹將本研究獲得之具體結論與建議說明如下：

### 6.1 結論

1. 本研究建立一套以電子票證大數據資料探討大眾運輸通勤族群行為與特性之流程，以新北市淡水區淡水捷運站之公共運輸乘車資料為分析標的，資料起始時間於2014年6月1日至2014年6月30日前後共30日之時間長度包含捷運及市區公車之交易記錄，其研究成果能反應其一重要運輸走廊之各通勤族群對於大眾運具選擇於時間與空間的使用特性，並提供欲探究票證資料的研究人員一套資料處理及分群邏輯之參考依據。
2. 在平日與假日之乘車行為變化方面，研究結果顯示一般族群之平日的尖峰時段落於上午6點至9點，而學生族群的乘車高峰落在上午10點至下午1點的時段中較早晨時段來的高，而持敬老卡的使用者除了傍晚11點至12點的時段以

外，其餘時段的乘車比例皆相差不大，根據本研究結果可推估，敬老卡族群在平日及假日使用規律十分相似，與一般族群及學生族群變化不同。

3. 在空間面分析，本研究將乘車旅次迄點定為捷運淡水站，根據各站點至淡水站之旅次長度進行分群，在捷運通勤資料顯示多數通勤族群的旅次長度落在15公里的區間當中，在此區間包含了：唹哩岸、石牌、明德、芝山、士林及劍潭等站點，其中以石牌站的乘車資料最為顯著(佔34%比重)，其次依序為士林、芝山、劍潭、明德、唹哩岸；而在公車族群的資料顯示，多數通勤族之旅次長度則落於5公里的區間當中。
4. 在時間面分析上，其捷運通勤族群的旅次起始時間落在6點至9點及10點至13點的時段區間當中；而在公車通勤族的起始時間則分佈在6點至9點、10點至13點、14點至16點及17點至19點四個時段區間當中，其中以10點至13點(19.8%)之乘車比例最高。
5. 本研究以Apriori關聯規則演算法探究通勤資料欄位間之關係，首先分別探勘捷運族群及公車族群對於旅次特質偏好之變化，其次則將捷運通勤資料集公車的通勤資料合併於相同資料庫當中，以分析大眾運輸通勤族群對於旅次變化之差異點，跟去研究結果可以推論出，當通勤族的旅次長度高於10公里時會較偏好捷運系統，而在旅次長度落於5公里以內時，則較偏好公車做為外出的代步工具。
6. 在旅次回饋情境模擬的結果顯示，多數的機車族群較偏好票價優惠、乘車扣款上限優惠及里程計費優惠之回饋方式，票價優惠的情境條件能吸引每周騎車通勤一至兩次及搭乘大眾運具每週一至兩次的族群；在乘車扣款上限的情境下，能吸引所得一萬元以下幾乎每天騎車通勤的族群；而在里程計費的情境下則能吸引，所得一萬元以下、平均通勤時間為四十至五十分、每周搭乘大眾運具一至兩次的族群。

7. 故此可以推估在旅次回饋的兩種方案當中，由通勤時間的長短可以推論，其方案較吸引通勤旅次較長的機車族群，故在選擇研擬回饋方案之對象時，可以選擇捷運做為設計基礎及研擬標的。

## 6.2 建議

1. 本研究證實能藉由探勘電子票證大數據資料庫而瞭解公共運具通勤族群之旅運行為及特性，藉以提供政府管理部門執行決策之考量，而國內應更積極投入數具科學探討之範疇當中，以不同的演算方式從大量且多元的數據中，探索更多使用者對於交通運具之偏好，並整合其執行策略以針對不同之通勤族群進行差異化鼓勵措施，以增加國內公共運具市場之市佔空間。
2. 資料品質對於分析結果影響甚遠，本研究之資料時間長度前後僅30日，在後續研究可考量增加資料庫之時間長度，由於民眾外出之旅運行為容易受到時間差異、氣候因素、連續假日或節慶及環境地理背景等影響，故後續建議可考量將資料庫依照時間、月份及天氣等差異分門別類加以分析，可望能取得更多元之研究成果。
3. 本研究選擇以新北市淡水區的學生通勤族群做為調查對象，無法將研究成果及誘因偏好做為所有學生族群偏好依據，後續建議可將調查對象以學校為單位與以區分，以檢視在不同之行政分區及地理環境條件下，學生族群對於運具使用頻率之偏好及對於乘車誘因之觀點差異為何。
4. 由於電子票證無法記錄私人運具相關資料，故本研究選用問卷調查方式詢問旅次偏好，文獻中所提及票價成本變動為影響低所得及學生族群偏好選擇之因素之一，故後續可嘗試以願付價格的方式，詢問其價格變動最敏感之區間，以實際價格的變化觀察吸引運具偏好移轉之比例。

## 參考文獻

### 中文部分

1. 吳秀玲(2001)，台北市信義區大眾運輸路線規劃之研究，交通大學交通運輸研究所碩士論文。
2. 邱瑋亭(2006)，消費者著迷對象、消費行為、產品與代言人選擇關聯性探勘之研究，淡江大學管理科學研究所碩士論文。
3. 邱詩淳(2006)，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系碩士論文。
4. 吳沛樵(2007)，非接觸式智慧卡對政府效益之分析，國立臺灣大學土木工程學研究所碩士論文。
5. 林祥生、劉益豪(2007)，應用資料採礦探討航空旅客之線上購票行為，中華大學運輸科技與物流管理學系碩士論文。
6. 林良泰、陳乃萁.(2010)，K-means集群分析法應用於號誌定時時制時段劃分之研究，運輸學刊第二十二卷第四期。
7. 林祥生、邱詩淳、劉益豪(2005)，應用悠遊卡資料挖掘公車乘客之需求特性，中華民國運輸學會第20屆論文研討會論文集，367-387。
8. 胡建勳、陳宗亮(2008)，應用資料群聚與關聯法則於網路拍賣使用者之特性與偏好研究，2008知識社群與系統發展研討會。
9. 郭奕杖等人(2012)，探討高鐵實施不同票種之價格促銷方案對小汽車駕駛人運具選擇行為的影響，運輸學刊第二十四卷第四期。

10. 黃中怡(2002)，策略性人力資源管理-組織策略、人力資源策略、人資部門涉入程度、核心競爭力、國家文化與組織績效之關聯，成功大學國際企業研究所碩士論文。
11. 陳鴻文(2002)，家戶特性與汽機車持有數及使用量關係之研究-以台北市為例，國立交通大學運輸科技與管理科學系碩士論文。
12. 張新立(2004)，民眾機車使用依賴性影響因素之探討---以台北市為例，行政院國家科學委員會專題研究計畫成果報告。
13. 張恭碩(2007)，非接觸式智慧卡於停車收費之效益分析，國立臺灣大學土木工程學研究所碩士論文。
14. 張智勝(2013)，以質性研究方法探究大學生使用機車與大眾運輸工具之態度-以高雄師範大學為例，國立成功大學都市計劃研究所碩士論文。
15. 陳民祐(2012)，應用資料探勘探討線上訂票乘客行為以台鐵西幹線為例，中興大學行銷學研究所碩士論文。
16. 陳偉穎(2012)，通勤運具選擇行為異質性之研析：混合羅吉特模式應用，交通大學交通運輸研究所碩士論文。
17. 莊依珊(2015)，滿意度指數為基礎之國到計乘電子收費服務預警機制之研究，淡江大學運輸科學研究所碩士論文。
18. 黃俊擇(2015)。公車使用者涉入程度與運具選擇行為意向之研究-台南市實證分析，成功大學交通管理科學系學位論文。
19. 楊志文、段良雄(2004)，考慮選擇集合及異質性之個體城際客運選擇模式，運輸計畫學刊第三十三捲。
20. 蕭珣君(2008)，消費者行為理論於城際公共運輸旅客行為意象與搭乘因素之



研究，淡江大學運輸科學所碩士論文。

21. 蕭宇軒(2010)，影響大眾運輸因素之研究，成功大學都市計畫研究所碩士論文。
22. 謝萬興、張學孔(2015)，運用悠遊卡巨量資料分析公車乘客行為之研究，台灣大學土木研究所交通組碩士論文。
23. 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學位論文。
24. 尹相志，「SQL Server 2005 Data Mining資料挖掘與Office 2007資料挖掘增益集」，台北市：悅知文化，民國96年。
25. 交通部統計處，「民眾日常使用運具狀況調查摘要分析」，民國104年。
26. 交通部統計處，「機動車輛登記數」，民國104年02月。
27. 吳明清，「教育研究法—基本觀念與方法分析」，臺北：五南，民國80年。
28. 吳明隆，「SPSS與統計應用分析」，臺北：知城數位科技，民國94年。
29. 吳明隆，「SPSS操作與應用：問卷統計分析實務」，臺北：五南，二版，民國98年。
30. 王正林、黃文，「利用R語言打通大數據經脈」，臺北：佳魁，二版，民國105年。

## 英文部分

1. Agrawal, R., Imieliński, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record*, 22(2), 207-216.
2. Berson, A., & Smith, S. J. (2002). *Building data mining applications for CRM*. McGraw-Hill, Inc.
3. Bagchi, M., & White, P. R. (2005). The potential of public transport smart card data. *Transport Policy*, 12(5), 464-474.
4. Champney, W. S., & Burdine, R. (1995). Macrolide antibiotics inhibit 50S ribosomal subunit assembly in *Bacillus subtilis* and *Staphylococcus aureus*. *Antimicrobial agents and chemotherapy*, 39(9), 2141-2144.
5. Cullinane, S., & Cullinane, K. (2003). Car dependence in a public transport dominated city: evidence from Hong Kong. *Transportation research part D: Transport and environment*, 8(2), 129-138.
6. Chapleau, R., & Chu, K. (2007). Modeling transit travel patterns from location-stamped smart card data using a disaggregate approach. In *11th World Conference on Transport Research*.
7. Farhi, E., Goldstone, J., Gutmann, S., & Sipser, M. (2000). Quantum computation by adiabatic evolution. *arXiv preprint quant-ph/0001106*.
8. Fong, A.M., Hui, S. C., & Jha, G. (2002). Data mining for decision support. *IT professional*, 4(2), 9-17.
9. Hui, S. C., & Jha, G. (2000). Data mining for customer service support. *Information & Management*, 38(1), 1-13.
10. Hsieh, N. C. (2004). An integrated data mining and behavioral scoring model for

- analyzing bank customers. *Expert systems with applications*, 27(4), 623-633.
11. Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, 31(8), 651-666.
  12. Kenworthy, J. R., & Laube, F. B. (1999). Patterns of automobile dependence in cities: an international overview of key physical and economic dimensions with some implications for urban policy. *Transportation Research Part A: Policy and Practice*, 33(7), 691-723.
  13. Kleissner, C. (1998). Data mining for the enterprise. In *System Sciences, 1998., Proceedings of the Thirty-First Hawaii International Conference on* (Vol. 7, pp. 295-304). IEEE.
  14. Kim, Y., & Street, W. N. (2004). An intelligent system for customer targeting: a data mining approach. *Decision Support Systems*, 37(2), 215-228.
  15. Kim, Y., & Street, W. N. (2004). An intelligent system for customer targeting: a data mining approach. *Decision Support Systems*, 37(2), 215-228.
  16. Rygielski, C., Wang, J. C., & Yen, D. C. (2002). Data mining techniques for customer relationship management. *Technology in society*, 24(4), 483-502.
  17. Piateski, G., & Frawley, W. (1991). *Knowledge discovery in databases*. MIT press.
  18. Pelletier, M. P., Trépanier, M., & Morency, C. (2011). Smart card data use in public transit: A literature review. *Transportation Research Part C: Emerging Technologies*, 19(4), 557-568.
  19. Thøgersen, J., & Møller, B. (2008). Breaking car use habits: The effectiveness of a free one-month travelcard. *Transportation*, 35(3), 329-345.

## 附件一

### 「淡水區私人運具旅次特性」研究調查問卷

各位同學您好：

此問卷之目的主要針對淡水區使用私人運具通勤族群進行旅次調查。

由於淡水區為北區大專院校集中最密集之行政分區之一，其包含了：淡江大學、真理大學、聖約翰大學及海洋技術學院等四所學校，以淡江大學近 1 萬 1 千名外宿學生為例，校內就有近 45% 的學生是通勤族群，其每日通勤的學生就佔了相當大的比重；而本研究在挖掘台北悠遊卡的資料發現，其每日搭乘大眾運輸的學生族群僅有不到 4000 筆的乘車資料。有鑑於此，本研究期望以「探討學生族群私人運具旅次特性」觀點，瞭解「私人運具之通勤旅次偏好」，期望以瞭解私人運具族群對於公共運輸之服務偏好之差異性，以建立私人運具偏好移轉之方案研擬。感謝您的填答，使本研究結果完善及精確。祝 平安順心

淡江大學運輸科學研究所

指導教授：陶冶中 博士

研究生：林浩瑋 敬上

連絡電話：0930-026-770

E-Mail：604660075@s04.tku.deu.tw

## 《問卷翻頁開始作答》

## 一、基本資料

- 1.職業：☐學生 ☐上班族
- 2.年齡：☐18~20歲 ☐21~25歲 ☐26~30歲
- 3.教育程度：☐高中 ☐大專院校 ☐研究所以上
- 4.個人每月平均收入約(含零用錢)：☐1萬以下 ☐1~2萬 ☐2~3萬 ☐3萬以上
- 5.您家裡有幾輛機動車(汽車/機車)? ☐1輛 ☐2輛 ☐3輛 ☐4輛及以上
- 6.居住狀況：☐與家人或親友同住 ☐在外租屋 ☐其他\_\_\_\_\_
- 7.居住地行政分區：☐台北市或新北市 ☐雙北以外
- 8.距離住家最近的捷運站點/火車站點：\_\_\_\_\_

## 二、旅客旅次交通運具資料

- 1.請問您每日出門前往淡水時段為：  
☐6點~9點 ☐10點~12點 ☐13點~15點 ☐16點~18點
- 2.自行騎車往返淡水的次數：☐幾乎每天 ☐每週3~4天 ☐每週1~2天
- 3.搭乘大眾運具往返淡水的頻率：  
☐每週3~4天 ☐每週1~2天 ☐每月1~2天或更少
- 4.平均每次加油金額：☐100元以內 ☐101~120元 ☐121~150元 ☐150元以上
- 5.平均每天騎車至淡水的通勤時間：  
☐30分鐘以內 ☐30分至40分 ☐40分至50分 ☐50分至60分 ☐60分以上
- 6.相較於大眾運輸您認為騎車至淡水大約能節省您多少的通勤時間?  
☐30分鐘以內 ☐30分至40分 ☐40分至50分 ☐50分至60分 ☐60分以上
- 7.請問在何種情況下會使您選搭大眾運輸工具?(至多選三項)  
☐天氣狀況不佳 ☐同行友人搭大眾運具 ☐通勤距離過長 ☐不易尋找停車空間  
☐停車成本考量 ☐目的地周邊設有大眾運具站點 ☐其它\_\_\_\_\_

### 三、大眾運輸服務水準偏好調查

各位填答者您好，為有效改善私人運具族群選搭大眾運輸之意願，下列歸納各影響選搭大眾運輸之因素，請根據個人觀點填入其意願選項

請問以下這些大眾運具乘車優惠方案，能提升您的選搭意願？

	同意	非常同意	普通	不同意	非常不同意
1.候車時間長短會影響我選擇大眾運具意願	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2.乘車時間長短會影響我選擇大眾運具意願	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3.轉乘時間長短會影響我選擇大眾運具意願	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4.接駁便利性會影響我選擇大眾運具意願	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5.乘車舒適性會影響我選擇大眾運具意願	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6.乘車票價考量會影響我選擇大眾運具意願	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

《請繼續翻頁作答》

#### 四、運量提升方案研擬

各位填答者您好，為有效改善私人運具族群選搭大眾運輸之意願，以下陳列出各種情境，以情境模擬概念做為運具選擇決策之方案研擬，請根據吸引誘因最高回饋方案，填答能使您改變運具選擇偏好之選項。

請問以下這些大眾運具乘車優惠方案，能提高您的選搭意願？

非常  
沒意願  
普通  
有意願  
非常  
有意願

1.公共運輸乘車費用給予八折之優惠  
(悠遊卡折扣再八折優惠)

☐ ☐ ☐ ☐ ☐

2.隨著乘車距離越長給予較多之票價優惠  
(里程計價優惠)

☐ ☐ ☐ ☐ ☐

離峰時段搭乘票價八折優惠  
(尖離峰定價法)

☐ ☐ ☐ ☐ ☐

4.隨乘車距離給予紅利點數，以紅利兌換現金或  
禮品回饋(里程紅利累積回饋)

☐ ☐ ☐ ☐ ☐

5.發行學生專屬通勤月票  
(如以每月2000元之金額即可享有乘車60次優惠)

☐ ☐ ☐ ☐ ☐

6.實施當週乘車金額上限優惠  
(當週乘車金額至定額後其餘免費)

☐ ☐ ☐ ☐ ☐

《問卷到此結束，由衷感謝您的填寫》