

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：羅孝賢 博士

運用資料採礦技術探討數位式行車紀錄器於公路
客運駕駛員異常操作行為管理之研究

A Study of Digital Tachograph in the Abnormal Driving Behaviors
Management of Bus Drivers Using Data Mining Techniques

研究生：高啟涵 撰

中華民國 95 年 6 月

論文名稱：運用資料採礦技術探討數位式行車紀錄器於公路客運駕駛員異常操作行為管理之研究

頁數：123

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：94 學年度第 2 學期碩士學位論文提要

研究生：高啟涵

指導教授：羅孝賢 博士

論文提要內容：

對於公路運輸服務業而言，異常駕駛行為除直接導致潛在行車肇事或交通違規，提高業者經營風險，亦會降低行車服務品質，造成車輛無謂耗損徒增保養維修費用。如何針對駕駛員異常操作情形作出有效管理以及增進行車安全，是客運業者亟待解決的問題。隨著科技進步，智慧型運輸系統（Intelligent Transportation System, ITS）發展日益蓬勃，國內公路客運業者已漸由傳統「機械式行車紀錄器」改用數位式行車紀錄器。惟數位式行車紀錄器蒐集項目眾多與資料量相當龐大，如能夠有效分析與利用，對於公路客運業者行車安全、油耗與保修費用以及駕駛員管理將有所助益。

本研究主要以公路客運為研究對象，蒐集國內某客運公司 61 部客運車輛數位式行車紀錄器資料、210 名駕駛員資料與肇事、交通違規、油耗與保養維修資料進行分析探討，資料蒐集時間為民國 93 年 6 月至民國 94 年 12 月。資料採礦首要步驟為定義與確認駕駛異常操作行為項目與門檻值，以萃取資料庫中有意義的資訊。運用多元迴歸分析與群集分析建立資料採礦模式，前者構建駕駛員異常行為關係模式，以釐清駕駛員異常行為對於行車安全、油耗與保修費用之影響關係；後者將相同駕駛特性駕駛員進行分群，使能與後續駕駛員管理配合運用。

研究結果發現，影響肇事次數為急減速與超速，影響交通違規次數為超速與電磁煞車操作異常；影響油耗費用為急加速、怠速過久以及引擎轉速異常，影響保養維修費用則為急減速、急加速、電磁煞車操作異常以及引擎轉速異常。群集分析將 210 名駕駛員分成三群，分別命名為一般等第 199 人，稍差等第 9 人、極差等第 2 人，分群結果符合一般現況，並構建判別模式可直接判別駕駛員所屬之群集等第。整合上述資料採礦結果提出可落實於駕駛員管理層面整合流程，與駕駛員管理獎懲案例與方法。最後評估以再教育訓練或相關管理方式提昇駕駛員素質水準之效益，結果得到一個月可以減少 13 次肇事、2 次交通違規，以及公司可節省 745,480 元額外油耗與保養維修費用，可利用節省費用提撥一定比例供再教育訓練用與新進駕駛員訓練，或增購數位式行車紀錄器進行車輛進行全面裝設，以充分掌握所有駕駛員駕駛行為資料，始能搭配相關管理措施以進行有效管理。

關鍵詞：資料採礦、多元迴歸分析、群集分析、判別分析、駕駛員管理

Title of Thesis:

Total Pages: 123

A Study of Digital Tachograph in the Abnormal Driving Behaviors Management of Bus Drivers Using Data Mining Techniques

Keywords: *Data Mining, Regression Analysis, Cluster Analysis, Discriminant Analysis, Driver Management*

Name of Institute:

Graduate Institute of Transportation Science, Tamkang University

Graduate Date: June 2006

Degree Conferred : *Master Degree*

Name of Student: *Chi-Han Kao*

高 啟 涵

Adviser: *Dr. Shiaw-Shyan Luo*

羅 孝 賢 博 士

Abstract:

For bus carriers, the abnormal driving behaviors will not only cause a higher risk of accident and traffic offence, but also deteriorate the vehicle worn-outs, which will cause bus service broken-down. Consequently, how to monitor and manage abnormal driving behaviors effectively and efficiently is an important issue to bus operators. With the progress of science and technology, Intelligent Transportation System is developed flourishingly. Today, many bus carriers have used digital tachographs to record the bus driving details. Particularly through data mining, the extraction of hidden predictive information from large databases will help us to find out and identify the relationship among abnormal driving behavior and driving safety, fuel consumption, and maintenance cost.

This study collected data from digital tachograph database, which include 61 buses, 210 drivers' data. In the meanwhile, the vehicles related accident and traffic offence records, fuel consumption and maintenance cost data were also collected from June 1, 2004 to December 31, 2005. The first step of data mining is to confirm and define the variables and related threshold values of abnormal driving behaviors in order to extract the meaningful information. The data mining techniques were used in this study, such as multiple regression analysis and cluster analysis. Multiple regression models were developed to establish the empirical relationship among abnormal driving behaviors and driving safety, fuel consumption, and maintenance cost. The cluster analysis was applied to categorize the sample of drivers which have similar driving characteristics. The discriminate analysis was used to determine the driver's level directly.

The findings of multiple regression models indicated that the emergent deceleration and speeding variables are the key determinants of the frequency of accident; speeding and abnormal operation of electromagnetic braking variables are the key determinants of the frequency of traffic offence; emergent deceleration and acceleration, long idle time of engine operation, abnormal operation of electromagnetic braking and abnormal engine rotation variables are the key determinants for fuel consumption and maintenance cost.

The cluster analysis has classified the 210 drivers into 3 category levels: fair, bad and very bad. In this case, there are 199 drivers in fair level, 9 drivers in bad level, and 2 drivers in very bad level. According to the data mining results, this study proposed an integrated driver management solution. After evaluation, it is shown that applying driver management strategies such as re-education, rewards and punishments, on a monthly basis, the frequency of accident can be reduced by 13 times, frequency of traffic offence by 2 times and a cost of NT\$745,480 for extra fuel consumption and maintenance cost are saved. The above savings could be used alternately for employees' re-education and training, or to equip digital tachographs on the bus fleets.



誌 謝

研究所的生涯轉眼間一逝而過，經過了這一個階段的學習與磨練，使得自己的能力更為充實，能夠順利完成論文與帶著滿滿的收穫畢業，首先要感謝恩師 羅孝賢教授的悉心教導與諄諄教誨，從論文研究方向與架構的確立，以及觀念的啟發，乃至於論文的寫作直到論文之完成，承蒙老師耐心指導與殷切斧正，使學生能夠順利的完成論文，除了專業領域外，還有做人處事態度方面的啟示，讓學生實為受益良多，僅致上十二萬分最高敬意與感激。在論文口試期間，承蒙 王中允教授與 陳苑蕙教授在繁忙之中特地撥冗前來審查，對於論文提供了許多極為寶貴的意見，使本論文更臻完備，在此致上由衷的感謝。此外在感謝 陶冶中教授與修習晶片學程的 廖述賢教授，兩位老師亦啟發了學生的多元思考空間，亦使得論文構想能順利產生，還要感謝系上所有的老師的教學薰陶以及兩位助教行政資源的協助，學生實為感激。

論文研究期間，感謝客運公司黎明義主任以及西門子公司的蔡聰毅先生，有你們的鼎力協助以及不厭其煩的提供資料與協助問題之解決，才能使得本論文研究進展得以相當順利且如期完成。

研究所期間，感謝高中好友上將、AK、小平在我作論文期間思緒困頓時，及時的拉我出去打球，使得常常看日出的我，還確保有健康的身體；還要感謝布仔長久以來對於我的加油打氣與文學氣質提升，適時地給予我支持與關懷，讓我在煎熬的論文製作過程獲得了些許的動力；並感謝學長卿豪對於 Paramics 軟體的幫忙教導，使我能夠快速的上手操作該軟體並能對於系上計劃案有所貢獻以及學長姐麗君與守琮的鼓勵與課業解惑以及學弟們遠凱與百里的協助與幫忙；也感謝各位研究所的同窗好友們紅豆甫、聖、哈利董董辰、瑋琦、書嫻、老賴、宏銘、漢瑄、正宏、士傑及鈺雯，跟你們一起奮鬥論文與課業切磋還有互相吐槽和鼓勵的那段時光真是令人難忘啊！

最後感謝父親與母親對於我在求學過程中的關心、支持與鼓勵，並讓我的求學之路無後顧之憂，有著媽媽親手煮的美味佳餚以及生活上無微不至的照顧，使得我作論文更是充滿了動力；還有我弟和我一起騎小白出去兜風還有一起機研真是太棒了，另外還有我門家最可愛的小喵咪，真的很高興妳在我熬夜的時候陪我，還逗我開心。僅以此論文獻給所有幫助、關心以及督促我的所有朋友，還有我摯愛的父親與母親，感謝你們！

高啟涵 謹誌

中華民國 95 年 6 月

目錄

第一章 緒論	1
1.1 研究背景與動機	1
1.2 研究目的	2
1.3 研究範圍	2
1.4 研究內容與流程	3
第二章 文獻回顧	6
2.1 數位式行車紀錄器相關回顧	6
2.1.1 國內外數位式行車紀錄器之發展	6
2.1.2 國內外數位式行車紀錄器產品類型	8
2.2 數位式行車紀錄器監控指標	14
2.2.1 行車安全方面	14
2.2.2 行車油耗與保養維修方面	16
2.3 資料採礦 (Data Mining) 相關回顧	17
2.3.1 資料採礦的意義	17
2.3.2 資料採礦步驟與資料庫知識發現(KDD)	18
2.3.3 資料採礦之功能與技術	21
2.3.4 資料採礦之應用	23
2.4 駕駛員管理相關回顧	25
2.5 小結	26
第三章 研究方法與理論基礎	28
3.1 研究架構	28
3.2 資料處理作業程序	29
3.2.1 資料蒐集	29
3.2.2 資料前置處理	30
3.3 資料採礦方法	31
3.3.1 多元迴歸分析	31
3.3.2 群集分析	34
3.4 判別分析	41
3.5 小結	45
第四章 駕駛異常操作指標設定與資料蒐集	46
4.1 數位式行車紀錄器駕駛異常操作監控指標與設定	46
4.1.1 行車安全方面	46

4.1.2 燃油消耗方面	50
4.1.3 機件磨損方面	53
4.1.4 行車品質方面	55
4.2 數位式行車紀錄器駕駛異常操作資料蒐集	58
4.3 駕駛員肇事與交通違規資料蒐集	59
4.3.1 駕駛員肇事資料	59
4.3.2 駕駛員交通違規資料	59
4.4 行車油耗與保養維修費用資料蒐集	60
4.4.1 車輛油耗費用	60
4.4.2 車輛保養維修費用	61
第五章 駕駛員異常操作行為關係模式構建	64
5.1 資料庫建立與分析	64
5.1.1 資料格式	64
5.1.2 基本資料分析	65
5.2 肇事與交通違規風險之模式構建	68
5.2.1 變數選取	69
5.2.2 模式設定	72
5.2.3 模式校估	73
5.3 行車油耗與保養維修費用之模式構建	77
5.3.1 變數選取	77
5.3.2 模式設定	79
5.3.3 模式校估	80
5.4 小結	84
第六章 駕駛員分群判別與管理	86
6.1 群集分析方法	86
6.2 駕駛員群集分析	86
6.2.1 群集數選擇	87
6.2.2 K-Means群集分析	88
6.2.3 群集假設檢定	90
6.2.4 駕駛員分群結果	92
6.3 駕駛員異常操作行為判別模式	93
6.3.1 判別模式架構	93
6.3.2 判別模式構建與結果分析	94
6.4 駕駛員分群管理	97
6.4.1 駕駛員評鑑等第	97
6.4.2 駕駛員行車安全風險與額外費用估計	100

6.4.3 駕駛員管理獎懲辦法	104
6.4.4 駕駛員素質提升之效益	111
6.5 小結	113
第七章 結論與建議	115
7.1 結論	115
7.2 建議	118
參考文獻	120



圖目錄

圖 1-1 研究流程圖	5
圖 2-1 德國 Siemens VDO 公司 DTCO 1381 數位式行車紀錄器	9
圖 2-2 澳洲 Circuitlink 公司 Tacholink Millenium 數位式行車紀錄器	10
圖 2-3 寶錄電子 BR6800 數位式行車紀錄器	11
圖 2-4 Siemens VDO FM200 PLUS 型數位式行車紀錄	13
圖 2-5 資料採礦技術流程圖	20
圖 3-1 研究架構圖	29
圖 3-2 群集分析之決策流程	36
圖 3-3 群集間距離示意圖	39
圖 3-4 群集分析架構圖	41
圖 3-5 三群體判別基準	43
圖 4-1 車輛加油流程圖	61
圖 5-1 交通違規項目分析圖	66
圖 5-2 肇事原因分析圖	67
圖 6-1 駕駛員異常操作行為群集數變化表	88
圖 6-2 各群集與變數顯著性檢定流程圖	91
圖 6-3 判別分析與模式構建流程圖	94
圖 6-4 駕駛員管理整合流程圖	103
圖 6-5 駕駛員責任歸屬判斷流程	104
圖 6-6 優良駕駛員決策圖	106

表目錄

表 2-1 資料採礦相關學者之定義	18
表 4-1 國道主要區段速限標準	47
表 4-2 急加減速門檻值對照表	50
表 4-3 一般路段之速度、引擎轉速與檔位對應關係 (FUSO 車)	51
表 4-4 上坡路段之速度、引擎轉速與檔位對應關係 (FUSO 車)	52
表 4-5 各項指標與監控項目	57
表 4-6 每月車輛油耗費用資料範例格式	61
表 4-7 公路客運營運成本及功能類別	62
表 5-1 駕駛員異常操作管理資料庫 (部分資料)	65
表 5-2 各項費用初步統計值	67
表 5-3 駕駛員異常操作資料統計	68
表 5-4 駕駛員異常操作資料統計(續)	68
表 5-5 複共線性判斷準則	69
表 5-6 第一次複共線性檢定	70
表 5-7 第一次複共線性檢定(續)	71
表 5-8 第二次複共線性檢定	72
表 5-9 第二次複共線性檢定(續)	72
表 5-10 模式設定各項變數意義與單位	73
表 5-11 模式 A 與 B 變異數分析(一)	74
表 5-12 模式 A 與 B 校估結果(一)	75
表 5-13 模式 A 與 B 變異數分析(二)	75
表 5-14 模式 A 與 B 校估結果(二)	76
表 5-15 七項變數複共線性檢定	78
表 5-16 七項變數複共線性檢定(續)	78
表 5-17 各項費用之相關分析	79
表 5-18 油耗與保養維修模式變數之意義與單位	80
表 5-19 模式 C 與 D 變異數分析(一)	81
表 5-20 模式 C 與 D 校估結果(一)	82
表 5-21 模式 C 與 D 變異數分析(二)	83
表 5-22 模式 C 與 D 校估結果(二)	83
表 6-1 各群集之起始中心點	89
表 6-2 各群集之疊代紀錄	89
表 6-3 各群集之最後中心點	90
表 6-4 各群集最後中心點間之距離	90
表 6-5 各駕駛員群集中之平均數	91

表 6-6 變異數分析表	92
表 6-7 駕駛員分群表	92
表 6-8 典型判別函數之係數值與結構矩陣	95
表 6-9 駕駛員判別分析之特徵值與 Wilks Λ 檢定統計分析表	96
表 6-10 費雪線性判別函數係數表	96
表 6-11 駕駛員分群判別分析結果分析表	97
表 6-12 各等第之基本統計值	98
表 6-13 駕駛員評鑑等第判別模式	99
表 6-14 異常操作行為風險表	100
表 6-15 多重比較檢定(一)	100
表 6-16 異常操作行為產生額外費用表	101
表 6-17 多重比較檢定(二)	101
表 6-18 駕駛員貢獻度計算表	105
表 6-19 優良表現駕駛員名單	106
表 6-20 不獎勵或懲處之駕駛員名單	108
表 6-21 D039 駕駛員之額外費用計算	109
表 6-22 稍差與極差等第駕駛員產生之額外費用	110
表 6-23 駕駛員素質提升節省金額計算(稍差等第)	112
表 6-24 駕駛員素質提升節省金額計算(稍差等第)	112

第一章 緒論

1.1 研究背景與動機

對於運輸服務業而言，行車安全是營運中最重要的項目之一，而駕駛員的駕駛行為攸關行車安全，根據國道警局分析指出大型車輛體積、噸位及馬力等，與小型車輛差距懸殊，一旦在高速公路發生事故，所造成的人員傷亡及車輛損壞程度，都較一般交通事故來得嚴重。2005 年高速公路上大型車輛（大貨車、大客車）的車流量僅佔國道車流量的 16.5%，但交通事故比例卻達 30.77%。就肇事原因來看，2005 年高速公路車禍事故以「酒後駕車」22 件最多；「未保持行車安全間距」20 件次之；「駕駛不當」15 件居第三，故可以得知，道路交通事故之發生多數都是駕駛人操作不當或駕駛習慣及態度不良等人為因素所引起的。故駕駛人對車輛的操作行為一般可由駕駛態度及駕駛習慣兩方面進行說明，駕駛態度係指駕駛人對行人、其他車輛路權的禮讓表現，以及對車前路況的注意程度，公路客運業還包括駕駛員對於乘客的表現；駕駛習慣則是駕駛操作車輛的慣用方式。一個優質的駕駛人應同時達到兩方面之水準。以公路客運業為例，異常的駕駛行為除直接導致潛在行車肇事或交通違規，提高業者之經營風險，損害業者之企業形象外，亦會降低行車服務品質，造成乘客負面的知覺感受，且容易造成車輛無謂的耗損，徒增保養維修費用。處於競爭激烈的環境中，若運輸服務業者未能針對異常之駕駛行為作出有效的約制與防範，長久下來，將使經營者陷入客源流失、成本增加、競爭力減弱之惡性循環。因此，如何針對駕駛員的異常操作情形作出有效的管理以及增進行車安全，是客運業者亟待解決的問題。

隨著科技的進步，智慧型運輸系統（Intelligent Transportation System, ITS）發展日益蓬勃，多數國內公路客運業者已由傳統之「機械式行車紀錄器」改用數位式行車紀錄器，俾對行車過程中之駕駛行為作詳細紀錄，並以圖表方式將資訊輸出。惟數位式行車紀錄器搜集的項目眾多與資料量相當龐大，其功能並不夠完備，輸出資訊大多停留在大量報表及數據階段，這些資料如能夠有效的分析與利用，對於公路客運業者之行車安全、油耗與保修費用以及駕駛員管理將有所助益。

駕駛員之異常操作行為除了將影響行車安全外，對於車輛之油耗與保養維修費用亦有潛在之影響，根據交通部運研所對國內 18 家汽車客運業者所提供之成本資料作歸納整理（2001），指出營運成本中管理費用佔 41%，業務費用佔營運

總成本的 28%，行車油耗費用佔 17%，保修費用佔 14%，故行車油耗與保修費用約為營運成本的三分之一，所以若能避免駕駛員之異常操作行為，將對於油耗或保養維修費用支出有所節省。因此本研究欲了解駕駛員之異常操作行為將會對行車安全與車輛油耗、保養維修費用造成何種程度之影響，以及哪些異常操作行為將會造成影響，並依據上述之關係，提出對於駕駛員管理方式，以增進行車安全與降低公司營運成本，並運用近年來應用越來越廣泛之資料採礦(Data Mining)技術。其目前被應用於各個領域（如醫療業、製造業、生化業），最重要的優點就是處理大量且複雜的資料庫資料，透過資料採礦技術將隱藏在資料中的資訊粹取出來。基於資料採礦技術之特色，運用在處理數位式行車紀錄器之資料，探討行車安全與駕駛員管理之課題，為本研究之主要動機。

1.2 研究目的

基於上述研究背景與動機，本研究主要目的有以下幾點：

- 對於國內外數位式行車紀錄器紀錄之資料項目與格式進行了解，蒐集駕駛員之相關資料，如肇事、交通違規、車輛油耗與保養維修費用，並進行資料採礦，定義駕駛員異常操作之項目與門檻值，以萃取出資料庫中未知且具有意義之資訊，並構建分析用資料庫。
- 構建資料採礦模式，以多元計量迴歸模式構建駕駛員異常操作行為與行車安全、燃油消耗與機件磨損方面之關係（包括肇事、交通違規、油耗與保養維修費用），以了解與說明其彼此間之因果關係。以及使用群集分析法，應用多元迴歸模式中重要之變數（異常操作項目），將駕駛員進行分群，並依分群結果命名各群作為駕駛員之評鑑等第。
- 運用判別分析，配合之前群集分析之結果，構建判別模式，以作為直接判別駕駛員評鑑等第之方式，並依據各等第提出駕駛員獎懲方式，進行有效管理。
- 估計駕駛員獎懲辦法與再教育訓練等適當的管理方式具有成效，即所有駕駛員皆再一般普遍之素質水準時，所能產生之效益。

1.3 研究範圍

本研究主要以公路客運為研究對象，蒐集國內某客運公司之 61 部客運車輛之數位式行車紀錄器資料、210 名駕駛員資料與肇事、交通違規、油耗與保養維

修資料進行分析探討，資料蒐集時間為民國 93 年 6 月至民國 94 年 12 月，共計 19 個月。

1.4 研究內容與流程

本研究主要內容與流程如下：

(1) 確立研究方向

本步驟主要是界定研究目標與範圍，以利後續研究之進行。

(2) 文獻回顧

在文獻回顧方面，本階段工作主要是收集國內外相關文獻，回顧對象依據研究主題，分別是數位式行車紀錄器之相關研究，資料採礦技術與應用領域之相關研究以及駕駛員管理之相關研究。

(3) 研究方法與理論基礎

透過文獻回顧，進行運用於本研究之資料採礦流程，以及找出符合本研究之資料採礦技術方法，並了解該資料採礦方法的理論基礎與分析進行方式。

(4) 數位式行車紀錄器監控設定與相關資料蒐集

進行資料採礦須先定義與蒐集數位式行車紀錄器異常操作項目之資料，並須先了解要設定之監控項目種類與其設定目的及原理，始能設定各項目監控之門檻值，以確保數位式行車紀錄器異常操作資料蒐集之正確性與合理性，始能從資料庫中萃取未知且有意義的資訊。並蒐集相關駕駛員資料，如駕駛員基本資料、肇事與交通違規資料以及車輛油耗與保養維修費用資料，以利後續異常操作關係模式之構建。

(5) 駕駛員異常操作行為關係模式構建

先將所蒐集之資料進行資料採礦前置處理作業，並使用處理完成之資料構建一個分析用之駕駛員異常操作管理資料庫。將所有資料進行初步統計分析，了解其特性與分佈情形。再透過資料採礦技術中之多元計量迴歸分析方法，將駕駛員異常操作項目作為解釋變數（自變數），肇事、交通違規、車輛油耗與保養維修費用為被解釋變數（因變數），構建駕駛員異常操作行為與肇事、交通違規、車輛油耗與保養維修費用之間的關係模式，探討其彼此間之因果關係與影響之程度。

(6) 駕駛員分群判別與管理

運用資料採礦技術中之群集分析，群集分析之分群輸入變數主要為之前構建多元計量迴歸模式中之解釋變數，將駕駛員依據其各自異常操作特性進行分群，並進行判別分析，構建各群集之判別模式，並依據各群集之特性分別進行命名，以作為駕駛員評鑑等第之依據。估計各等第駕駛員所產生之行車安全風險與額外油耗及保修費用，提出各等第之駕駛員管理獎懲辦法，後續估計管理與再教育訓練對於駕駛員水準素質提升後所能產生之效益。

本研究之研究流程圖如圖 1-1。



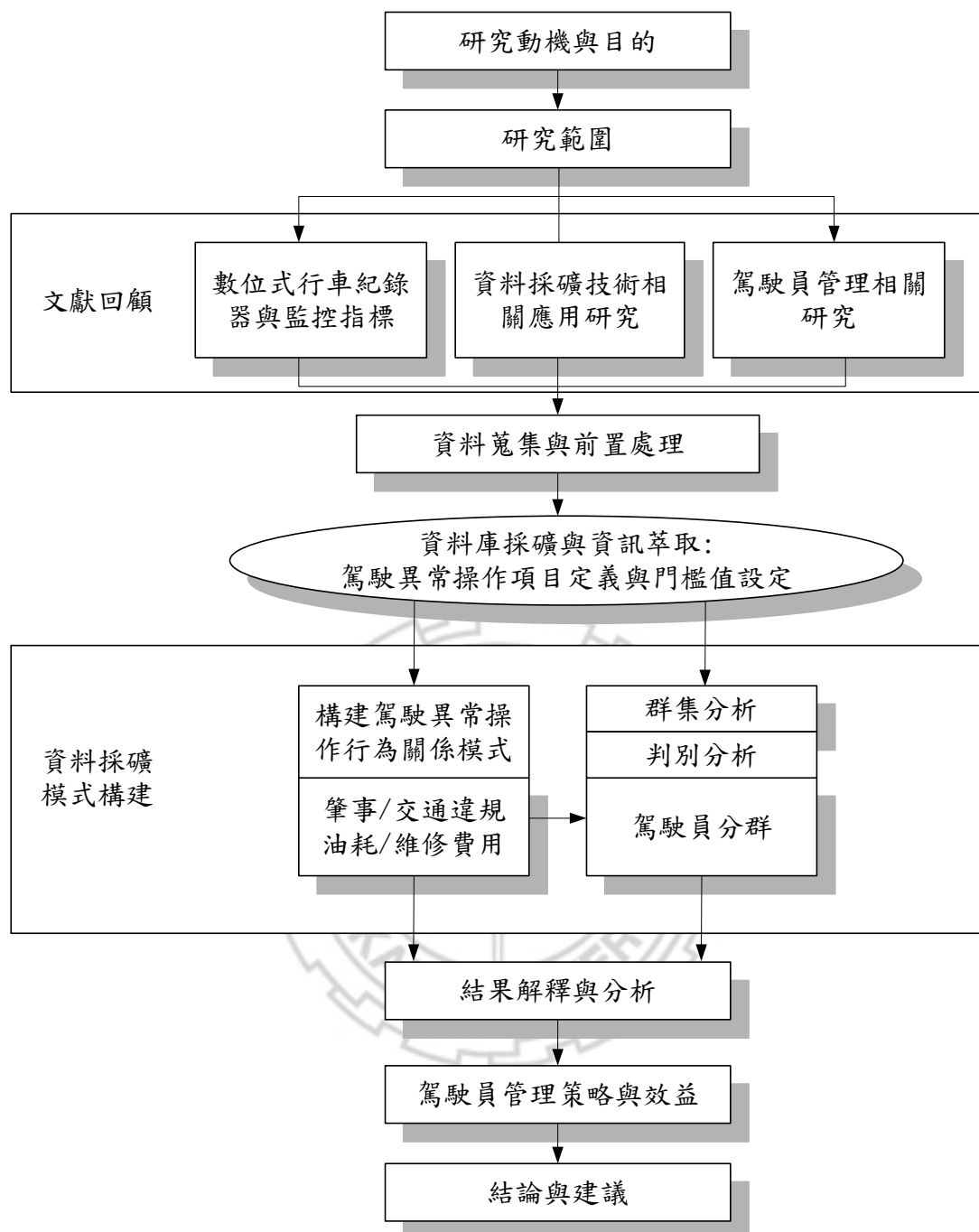


圖 1-1 研究流程圖

第二章 文獻回顧

依據第一章之研究目的與流程，本章節針對數位式行車紀錄器類型、數位式行車紀錄器監控指標、資料採礦以及駕駛員管理等四大方面進行相關文獻回顧。

2.1 數位式行車紀錄器相關回顧

國內開始使用行車紀錄器是於民國 59 年由日本引進台灣，傳統行車紀錄器屬於機械式，是利用機械式轉軸帶動指針繪製速度曲線，準確性低且需經過專業訓練之人員加以判讀，具有處理時間長及易被擅改等缺點。由於科技的發展，數位化的潮流，進而發展出數位式行車紀錄器，數位式行車紀錄器較傳統機械式行車紀錄器具有資料傳輸與管理之方便性，並可減少人為誤判，同時兼具擴充性、整合性及可依不同需求記錄不同組合之資料等多項優點，為後續建置智慧化車隊管理之重要基礎設備單元，本研究依據交通部運研所相關研究資料（民 93）以及國內外數位式行車紀錄器製造商網站資料，針對國內外數位式行車紀錄器之發展與產品類型進行回顧。

2.1.1 國內外數位式行車紀錄器之發展

行車紀錄器（Tachograph）於 1925 年在德國發明，到目前已有數十年的歷史，Tachograph 是從德文的回轉速度器（Tachometer）和紀錄（Graphik）兩個文字組合演變而成，最初的主要功能在於記錄駕駛的行車時間，以避免駕駛過度疲勞開車而影響安全，以及車輛過度運轉而引起事故。行車紀錄器主要使用在歐洲國家的車隊如巴士、重型貨車及緊急救援車輛，其他國家亦強制規定商用車輛必須裝設行車紀錄器，以日本為例，租用巴士、八噸以上貨車、以及部分地區的出租車輛及計程車應依規定安裝行車紀錄器。根據部分歐洲國家的實施經驗，行車紀錄器最初被引進時，由於具有監視駕駛行為的主要功能，因此造成駕駛及工會的反彈，駕駛將它戲稱為「車上的間諜（the spy in the cab）」，普遍存有抗拒的心態，然而近年來駕駛的心態有重大的改變，反而將行車紀錄器視為改善工作環境的利器，由於公路貨運市場的競爭激烈，加上許多公營客貨運公司民營化後，駕駛的工作時數常常超過法令規定的上限，造成行車安全性降低，行車紀錄器的工作時數記錄功能就成為駕駛的護身符，而數位式行車紀錄器更能減低行車紀錄被

擅改的機率，因此駕駛及工會多樂觀其成，希望能夠改善他們的工作環境。

數位式行車紀錄器的觀念在 1984 年即被提出，由於能夠提供種類更為繁多、資料更為詳細的數據，因此逐漸被商用車隊採用。歐盟已立法規定於 2004 年 8 月開始強制所有商用車輛的新車裝設數位式行車紀錄器，歐盟採用的數位式行車紀錄器是將數位式資料記錄在駕駛攜帶的 IC 卡（Driver Card）與車輛單元（Vehicle Unit）中，IC 卡記錄駕駛的基本資料、車輛使用資料、駕駛活動資料（由駕駛自行輸入）、起迄點資料等，IC 卡可保存 28 天的資料。車輛單元紀錄設備基本資料、IC 卡插入與抽出資料、駕駛活動資料、里程表資料、速度資料等，能夠保存 365 天的資料，車輛單元含有印表機，可將 IC 卡資料或車輛單元資料列印出來，而車隊管理者可利用公司 IC 卡（Company Card）下載車輛單元中記錄的資料。警察或交通監理單位路檢時需配備手提電腦及讀卡機，利用控制 IC 卡（Control Card）得到行車紀錄器資料，或是要求駕駛利用車輛單元的印表機列印資料，以接受檢查。

歐盟曾於 1998 至 2000 年進行一項大規模的商用車輛車上系統架構的計畫，稱為 COMETA（Commercial Vehicle Electronic and Telematic Architecture），主要原因是市場上商用車輛的車上設備越來越多，彼此功能重疊、無法相容的情況相當普遍，本計畫的目的在定義各種車上系統的功能（包括行車紀錄器），使各系統的介面能夠有效相容，計畫的成果為車上系統的系統架構（System Architecture），該架構可視為歐洲地區對於車隊管理系統的一項平台，所包含的商用車隊管理包括行動電子資訊交換、數位式行車紀錄器、駕駛輔助系統、電子收費設備等。在此架構中，行車紀錄器子系統擷取、儲存並保護相關的資訊，這些資訊包括駕駛受到規範的資訊如車速、工作時數以及其他車輛操作的資訊如加減速、耗油量等，行車紀錄子系統接受駕駛輸入及車輛感應器傳送的資訊，其資料庫可由駕駛、車隊管理者及相關單位進入讀取。

另一種與行車紀錄器功能類似的產品稱為「事故資料紀錄器」（Event Data Recorder, or EDR），EDR 記錄車輛撞擊前後一段時間內的車輛縱向、橫向加速度、行駛速度與行駛方向，以及大燈、左右轉燈、煞車、喇叭等相關資訊，利用撞擊前、撞擊時及撞擊後的資料判斷發生事故的原因，以 Siemens VDO 公司生產的 UDS（Umfall Data Schreiber）系統為例，系統可以記錄車速、橫向與縱向加速度，其記錄頻率達 500 Hz，即週期為 0.002 秒。發生事故時，系統記錄發生前 30 秒及發生後 15 秒共 45 秒時間內的資料。還有部份的 EDR 具有影像紀錄

的功能，如 SIS 公司生產之 MAC Box，以輔助事故鑑定及緊急救援之功能。

EDR 與行車紀錄器的差異主要在於 EDR 是專門使用在肇事鑑定上，並且可以利用無線通訊設備使得事故發生時自動通知緊急救援中心，而行車紀錄器雖然亦有協助肇事鑑定之功能，然而一般來說行車紀錄器在肇事鑑定的功能上不如 EDR，且行車紀錄器能夠長時間紀錄車輛操作狀態，此為 EDR 所沒有具備之功能。

國內廠商自民國 80 年代後期即陸續引進數位式行車紀錄器，並能夠自行生產，主要能夠生產數位式行車紀錄器的廠商有彙通、寶錄與康訊科技等公司，另外亦有廠商代理引進國外數位式行車紀錄器，主要為 Siemens VDO 的產品，國內外數位式行車紀錄器的產品與比較將於以下介紹。

2.1.2 國內外數位式行車紀錄器產品類型

以下介紹三種國外較為常用的數位式行車紀錄器型式。

一、德國 Siemens VDO 公司之 DTCO1381 數位式行車紀錄器

Siemens VDO 公司長久以來即生產機械式行車紀錄器，因應數位式行車紀錄器的需求，Siemens VDO 公司生產一套符合歐盟最新規範的數位式行車紀錄器—DTCO 1381，DTCO 1381 的外觀如圖 2-1，DTCO 1381 之大小與車上收音機相同，包括兩個 IC 卡讀卡機、列印機、螢幕、時鐘、操作控制及資料儲存之設備等，能夠記錄駕駛工作狀態、行駛速率與距離、相關參數如引擎轉速、車輛操作紀錄等，駕駛卡能夠記錄 28 天的工作狀態，而主機記憶體則可記錄 365 天車輛的所有操作紀錄以及 24 小時的速度資料。



資料來源：Siemens VDO公司網站

圖2-1 德國Siemens VDO公司DTCO 1381數位式行車紀錄器

二、澳洲 Circuitlink 公司之 Tacholink Millenium 數位式行車紀錄器

澳洲 Circuitlink 公司所生產的數位式行車紀錄器，如圖 2-2，是利用記憶卡匣來儲存資料，安裝在儀表板上以利駕駛查看，所紀錄的資料包括速率、引擎轉速(RPM)、急煞車、急加速、停車時間、怠速時間、駕駛時間、行駛距離、事故資料等，並可選擇輸入包括喇叭、車門開關、水箱溫度、方向燈、車燈、油耗、載重等數位或類比式資料，分析軟體所產生的報告分為主要活動報告、異常駕駛行為報告、總旅次、每日總報告、每週總報告、每日駕駛報告、每月駕駛報告以及車隊車輛報告等。



資料來源：澳洲 Circuitlink 公司網站

圖 2-2 澳洲 Circuitlink 公司 Tacholink Millennium 數位式行車紀錄器

三、日本鈴木(ISUZU)公司之 MBCD 數位式行車紀錄器

日本鈴木公司之 MBCD (Mobil Business Computer Digital Tachograph) 是一套以數位式行車紀錄器為主體、結合 GPS 與 GIS 的車隊管理系統，該系統能夠記錄的資訊分為兩類：

1. 作業資訊：裝卸貨物、車門開關、待機、加油、休憩、路檢、洗車、實或空車等。
2. 行車資訊：車速、急加減速、時間、距離、位置、引擎轉速、高速或一般道路行駛、貨倉溫度等。

MBCD 主體為數位式行車紀錄器。MBCD 系統能提供國土交通省指定之三種紀錄，包含車速、行駛時間及距離，並能夠提供許多在車隊管理上有用報表，如運轉每日與每月報表、違反規定一覽表、安全管理總表等。

國內廠商自民國 80 年代後期即陸續引進數位式行車紀錄器，並能夠自行生產，國內數位式行車紀錄器的主要生產廠商之代表性產品介紹如下：

一、寶錄電子公司之 BR-6800 數位式行車紀錄器

該產品於民國 90 年推出，為 BR-92 的改良型，該產品已取得車輛研究測試中心 (ARTC) 之審驗合格報告，目前已有屏東客運、桃園客運等客運公司採用，阿羅哈客運先前也有採用。該系統由車上資料匣記錄行車資料，再透過資料匣讀

取機讀取記錄之資料，產品外觀如下圖 2-3 所示。



資料來源：寶錄電子股份有限公司網站

圖 2-3 寶錄電子 BR6800 數位式行車紀錄器

1. 行車紀錄器主機

記錄車輛行駛之重要資訊，包括車速、引擎轉速、GPS 訊號、時間、近遠燈、方向燈、煞車燈、倒車燈、車門開關、引擎異常、機油壓力、冷氣或冷藏溫度等，並可外接聲音及燈光警示，以告知駕駛員異常狀況。記憶卡匣具有 8MB 的容量，可記錄約 10,000 公里的行車資料，容量可擴充至 128MB。

2. 卡匣讀取機

主機與後台系統之傳輸工具，透過讀卡機將主機之卡匣資料讀取後，上傳至行車資料管理系統，以便進行資料分析。

3. 行車資料管理系統

主要功能在彙整、分析並管理車隊所有車輛之行車紀錄，提供管理者重要數據及駕駛績效考核之參考。本系統包括四個執行程式，CR6800r是自記憶卡匣讀取資料並轉存至資料彙集機；CRAssay是行車紀錄資料的彙整、分析及管理，並對個別紀錄進行檢視與操作，能夠列出如行車超速、行車路段、行車異常等表單；CRReport的功能是檢視行車紀錄摘要資料，並提供報表列印功能；CRDbMan則

是提供行車紀錄資料相關的資料維護，包括基本資料管理及GPS定位管理。

BR-6800 系統配備之記憶卡匣可擴充至 128Mbytes，可記錄至少 160,000 公里之資料，搭配 GPS 設備、無線通訊裝置後，可進行車隊即時之監控。

二、彙通科技公司之數位式行車紀錄器 ODR2002

1. 數位式行車紀錄器

安裝於車輛上，擷取車速、煞車、方向燈、大燈、霧燈、車內燈、冷氣、車門...等行車狀況，並記錄車號、路線編號、駕駛人員編號及電源開啟、關閉時間等行車資料，另有使用需求時可加裝其他感應器，以記錄引擎轉速及耗油量，該設備具超速警示功能。

2. PDA 資料收集及車機設定系統

數位式行車紀錄器上的資料透過USB連接埠經PDA連線傳送給資料處理分析電腦系統，並可利用PDA進行車機設定及檢測監看等功能。

3. 處理分析電腦系統

讀取數位式行車紀錄器之資料後，會自動處理分析車隊行車狀況，如車輛超速與否、班車營運狀況...等，具有駕駛行為管理、車輛肇事分析、車輛服務水準分析、車輛維護管理、車隊調度管理、車隊營運分析、道路服務水準分析等相關管理功能，供相關營運管理單位使用。

三、台灣西門子分公司威迪歐汽車電子部門代理之 VDO FM200 PLUS 型數位式行車紀錄器

1. 數位式行車紀錄器

台灣西門子分公司威迪歐汽車電子部門代理 Siemens VDO Automotive 生產之 FM200PLUS 數位式行車紀錄器，如圖 2-4，已經獲得車輛研究測試中心的審驗合格報告，目前有阿羅哈客運以及和欣客運等客運公司所採用。其主要功能為可設定駕駛異常操作監控項目與門檻值，以及配合車輛訊號輸入擷取煞車、方向燈、大燈、霧燈、車內燈、冷氣、車門開關等行車狀況。每位駕駛員配備獨立之啟動鑰匙，可針對個別駕駛原紀錄與蒐集駕駛資訊。VDO FM200 PLUS 配有 GPS 衛星定位模組，並採用五千分之一電子地圖，可追蹤車輛之行駛軌跡，進行車輛即時監控。

2. 資料蒐集方式

FM200 PLUS 採用隨身碟為媒介作為身分認證與資料下載用途，其中駕駛員使用藍色隨身碟，其記憶體為 128bytes EEPROM，具有駕駛員身分存取認證、日期時間輸入之功能，綠色隨身碟則為下載車輛旅程資料之用，其記憶體為 96~256 Kbytes EEPROM。亦可採用 GSM 無線通訊傳輸之方式，將車上之數位式行車紀錄器資料下載至管理者之主機裡。

3. 後端管理系統

Siemens VDO 系列之數位行車紀錄器亦搭配該公司所研發之後端資料庫管理軟體，可進行駕駛員管理、車輛維護管理、車隊調度管理等方面之管理，其所輸出資料格式為標準 SQL 資料庫格式，亦可使用其他資料庫軟體進行分析。



資料來源：Siemens VDO 網站

圖 2-4 Siemens VDO FM200 PLUS 型數位式行車紀錄器

四、康訊科技股份有限公司之偵航家 3010 行車紀錄系統

偵航家 3010 為一內含 GPS（全球衛星定位）的行車紀錄系統，記錄資料包含車輛之位置、速度、GPS 方向與收訊狀況等，使用者可將偵航家 3010 所記載之資料經由 RS232 介面下載至個人電腦或筆記型電腦。偵航家 3020 則將 3010 基本型行車紀錄器搭配無線數據通訊之功能，利用此功能，使用者不但可記錄行經之路徑軌跡，更可透過無線數據通訊將位置資料即時傳回監控中心，而無線數據通訊的種類則有 GSM、CDPD、DataTac、D-AMPS 等選擇。

為便於監控中心動態追蹤這些行車紀錄資料，偵航家 3010 與 3020 使用二套可即時顯示行車路徑於地圖上之軟體 TracStation 及 TracMate，運輸業者可利用此軟體隨時動態追蹤車輛之行進過程或接收車輛自動傳回之訊息，作為車輛監控、派遣管理、緊急救援等之用途。

2.2 數位式行車紀錄器監控指標

本研究欲探討公路客運駕駛異常操作行為對於行車安全與行車油耗、保養維修費用的影響，故分別就行車安全以及行車油耗與保養維修方面進行文獻回顧與探討，並作為後面章節駕駛操作行為監控項目與門檻值設定之依據。

2.2.1 行車安全方面

根據國道警局分析指出，大型車輛體積、噸位及馬力等，與小型車輛差距懸殊，一旦在公路發生事故，所造成的人員傷亡及車輛損壞程度，都較一般交通事故來得嚴重。例如在台灣地區 2005 年高速公路上大型車輛（大貨車、大客車）的車流量僅佔國道車流量的 16.5%，但交通事故比例卻達 30.77%。就肇事原因來看，2005 年高速公路車禍事故以「酒後駕車」最多；「未保持行車安全間距」次之；「駕駛不當」居第三。可見不當或異常的駕駛行為對於行車安全之影響相當大，以數位式行車紀錄器監控異常之駕駛操作行為是相當有必要的。

廖景元（民 69）研究指出在高速公路行駛時，由於超速、未保持行車安全間距以及任意變換車道等不當駕駛行為，因而產生造成肇事，其佔總肇事次數比例之 65%。

黃聖源（民 77）研究採用錄影觀察實地瞭解駕駛者外顯行為，其內容考量變換車道、車道選擇、違規轉向及違規使用車道等諸行為，對駕駛者旅行時間之影響。綜合分析結果所得結論為國內駕駛者發生變換車道行為之頻次相當高，其佔所觀察比例的 27.2%，且以行駛外側車道車輛較易發生該行為。未受紅燈影響車輛，任意變換車道或急切出、切入，除容易造成車流紊亂，影響道路服務水準以及旅行時間外，亦易因與前後車安全間距不足而發生車禍事故，而使道路更為堵塞。

林雅俐等學者（民 88）的研究使用「因素化駕駛行為問卷」(FDBQ)，對於駕駛者的駕駛行為進行資料蒐集與分析。其研究方法主要為利用統計中的集群分

析與邏輯斯迴歸模式對於駕駛者交通意外事故風險與駕駛者特性、駕駛行為進行關聯性分析，並建立駕駛者的肇事概率與風險評估方法，研究結果顯示，性別、年齡、駕車頻率、吊銷或吊扣駕照紀錄、超車違規紀錄與酒醉違規紀錄均與肇事風險有關，以及駕駛行為中之侵略性與超速性違規與交通意外事故風險有明顯的關聯性，侵略性高的駕駛者有較高機率超速違規以及較高的交通意外事故風險。

吳偉碩（民 90）蒐集肇事紀錄資料，並對於易肇事地點找出主要肇事原因，其採用逐步迴歸分析方法建立一肇事率預測模式，結果發現駕駛者在交叉路口附近未保持行車安全距離以及不當變換車道，以及在高速公路快車道內超速失控、車輛爆胎等造成交通事故的比例較高。

尹維龍（民 93）研究主要探討情緒壓力、駕駛技能對偏差駕駛行為與事故風險之間的相互關係，在偏差駕駛行為部分，以主軸因素分析、Varmax 轉軸方式萃取出潛在因素，結果顯示台灣地區偏差駕駛行為分別為違規、錯誤與疏忽等三個潛在因素，其中違規包含侵略性駕駛的部份；情緒壓力方面萃取出之潛在因素為情緒反應與壓力緊張；駕駛技能方面萃取出之潛在因素為駕駛技巧與安全感知。並以多元迴歸進行分析，以駕駛者屬性、情緒反應、壓力緊張、駕駛技巧與安全感知為自變數，結果顯示情緒反應、駕駛技巧與違規成正相關，安全感知、壓力緊張、年齡與違規成負相關。並以邏輯斯迴歸建立偏差駕駛行為與事故傾向之關係，分析結果顯示當違規與操作錯誤等行為都會影響事故發生，當違規與操作錯誤越高時，發生事故的機率有逐漸增加之趨勢。

Gully 等學者（1995）的研究其觀測駕駛者的駕駛行為發現，任意變換車道以及急減速會與肇事的發生頻率呈現相關，因此可知，某些特定駕駛行為的變化會與事故發生機率有所關聯性。

Quimby 等學者（1999）的研究對於駕駛者行駛速度選擇的影響因素進行實際調查分析，並且從該研究之調查發現，影響駕駛者速度選擇的因素有道路幾何設計與道路環境、法定速限以及駕駛者之教育訓練，並顯示駕駛人對於行駛速度的選擇是影響道路交通事故的次數與嚴重度的主要原因，選擇行駛速度越高的駕駛者，較有可能發生交通事故且其嚴重程度會較高。

Taylor 等學者（2000）研究指出較低的行駛速度能夠有效的降低車禍碰撞的次數，其研究發現平均行駛速度降低每小時 1 英哩，車禍碰撞事故就會減少 2%~7% 的比例，然而法定速率的訂定或是調整，仍需考量整體道路環境以及駕

駛人的意見。

2.2.2 行車油耗與保養維修方面

根據交通部運研所對國內汽車客運業者所提供之成本資料作歸納整理（民 89），指出客運業營運成本中管理費用佔 41%，業務費用佔營運總成本的 28%，行車費用佔 17%，保養維修費用佔 14%，故行車與保修費用約為營運成本的三分之一。而影響行車油耗與保養維修費用的因素除了車輛特性、道路交通特性以及環境因素特性外，駕駛行為因素亦為影響行車油耗與保養維修費用之重要因素之一，故本研究對於駕駛行為特性與行車油耗與保養維修費用之關係進行相關之文獻回顧與探討。

Waters and Laker（1980）以不同駕駛者進行油耗與駕駛習慣關係之實驗，在實驗過程中，所有駕駛人先以自己的開車習慣繞試驗場一週，再以建議的溫和駕駛行為繞試驗場一周，結果發現，後者較前者省下約 15% 的燃油消耗。

Ang and Fwa（1988）研究中指出，影響大眾運輸客運車輛的油耗因素分別為車輛特徵、車輛使用率與路線特徵、駕駛員行為以及道路環境等因素，而車輛特徵包含車輛廠牌、年份以及汽缸大小等；路線特徵包含旅客承載數、行駛速率、號誌停等數以及道路狀況等；駕駛員行為亦會影響車輛的油耗，一部車因兩班制通常會有兩位駕駛員來操作；道路環境包括道路幾何線型、坡度與路面狀況等。並採用迴歸分析之方式，將各項因素代入分析，並得迴歸模式進行油耗預測。其預測結果高承載低速度以及低承載高速度之油耗介於 1.52 km/l 至 3.06 km/l 間。

Hooker（1988）研究為對於單一車輛的最經濟省油的駕駛型態，主要影響燃油消耗的因素為換檔時機、行駛速度以及加、減速程度。駕駛者將檔位由低速檔變換至高速檔的時間的減少，以及在引擎經濟轉速下行駛，避免急踩油門或煞車，對於節省額外燃油消耗均有顯著的貢獻。

Decicco and Ross（1996）的研究指出改善交通管理措施及策略可節省約 5 % 的燃油消耗，但是當駕駛行為改善後，最高可節省燃油達 49% 的比例，由此可知，改善駕駛者的駕駛行為在燃油消耗的節省上效果相當顯著。

Vasiliev（1999）研究指出，當柴油引擎預熱或怠速時間不足時，將導致機油無法達到工作溫度，造成機油黏滯度不足，使得引擎產生偏磨的現象。

Voort 等學者（2001）研究中指出當車輛處於低檔高速、高檔低速、速度離

異過大以及急加減速時會因為引擎不在經濟轉速下運轉，而導致燃油之異常消耗。

蕭重威（民 83）研究模擬四行程車用引擎動力傳動系統，在行車型態下的動態性能。模擬方法以引擎循環模擬程式為基礎，增添引擎瞬態加速模式、傳動系統模式，其中包括換檔過程、行車型態及油門及剎車控制策略，結果發現油門及煞車控制策略對引擎動態性能影響最大。

胡俊生（民 85）指出車輛煞車時煞車鼓內部溫度及熱應力的升高，會對煞車鼓的材料造成不良的效果，先是產生熱變形影響煞車效果，繼而發生塑性變形，最後則會產生熱裂縫，若突然行經水坑，煞車鼓可能因急速冷卻而產生龜裂。此種現象對小車而言並不明顯，但是卻容易發生在大型巴士或卡車緊急煞車時會造成相當之危害。煞車鼓及來另片間的溫度也和煞車系統的性能有關。

2.3 資料採礦（Data Mining）相關回顧

2.3.1 資料採礦的意義

資料採礦（Data Mining）就是從大量的、不完全的、有雜訊的、模糊的、隨機的實際應用資料中，提取隱含在其中的、人們事先不知道的、但又是潛在有用的資訊和知識的過程。亦有人譯作「資料探勘」或是「資料挖掘」，也有人稱為「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）。目前已被許多研究人員視為結合資料庫系統與機器學習技術的重要領域，許多產業界人士也認為此一領域是一項增加各企業潛能的重要指標。這個領域蓬勃發展的原因是由於現代的公司企業經常蒐集了大量的資料，以及數位化與資料庫化所蒐集得來龐大的資訊，如數位式行車紀錄器之資料庫。但是資料超載與無結構化，使得相關決策單位無法有效的利用現有的資料，甚至使決策行為因無正確資訊可供參考而產生混亂與誤用。若能透過資料採礦的技術，從巨量的資料庫中，發掘出有意義或是令人驚奇的資訊與知識出來，做為決策支援之用，如此一來，必定能夠產生企業的競爭優勢或是提升管理效率。以下為資料採礦相關學者之定義，如表 2-1。

表 2-1 資料採礦相關學者之定義

學者（年代）	資料採礦的定義
Frawley（1991）	從資料中萃取出隱藏的、先前未知的有用資訊
Grupe and Owrang（1995）	從現存資料中剖析出專家們尚未發現之新事實及新關係
Fayyad（1996）	可萃取出資料中有效的、嶄新的、具潛在效益的資訊之非細瑣過程，其最終目標是瞭解資料的型樣（Pattern）
Berry、Linoff（1997）	為了發現有意義的模式或規則，以自動或半自動的方式，來勘查、分析大量資料所進行的流程
Cabena（1997）	將先前不知道、有效的資訊從大型資料庫抽出的過程，並且將萃取出來的有用資訊提供給主管做決定性的決策
Kim et. al.（2000）	提煉未知且可能有用的知識來幫助決策的一系列流程
Han and Kamber（2001）	從儲存在資料庫、資料倉儲或其它資訊儲存體的大量資料中，發現有趣特徵的過程

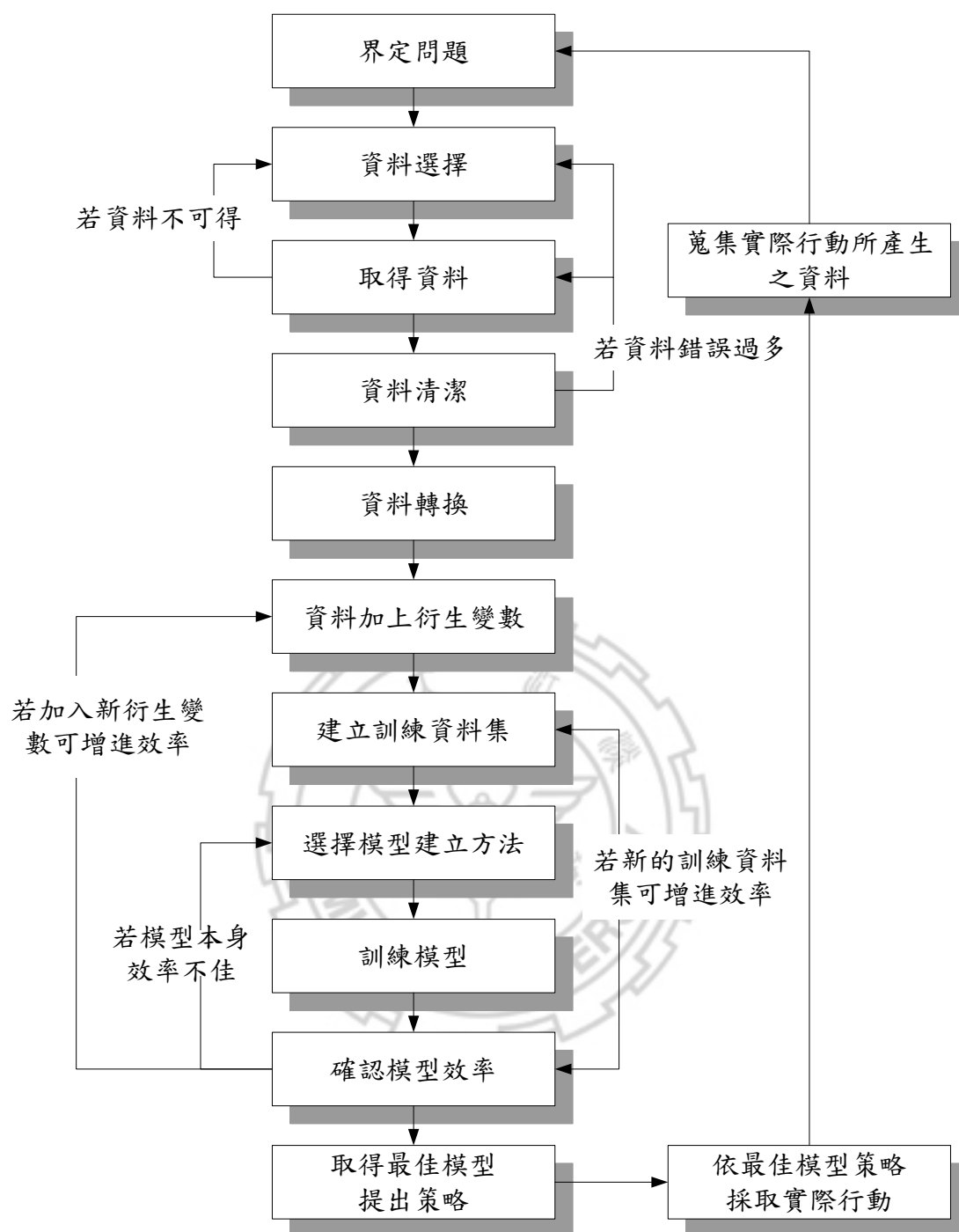
資料來源：許哲瑋（民 91）、本研究整理

2.3.2 資料採礦步驟與資料庫知識發現(KDD)

其實資料採礦為資料庫知識發現（Knowledge discovery in database, KDD）過程的一個步驟而已。依據 Fayyad（1996）對 KDD 的定義：它是一個從資料中找出合理、新奇及潛在效益的一個非瑣碎（Nontrivial）流程，其目的在瞭解資料的型樣（Patterns）。因此其流程為：先理解要應用的領域、熟悉相關的知識，接著建立目標資料集，並專注所選擇之資料子集；再從目的資料中做前置處理，去除錯誤或不一致的資料；然後做資料簡化與轉換工作；再由資料採礦的技術過程，建立各個採礦模型，找出分類型態或是進行預測分析；最後再經由解釋以及評估之後成為有用的知識。以上所述的程序呈現一個循環、一直重複的關係，最後才能得到相關有用的知識。

在資料採礦中亦有一連串的步驟，其中主要的進行步驟如下：

- (1) 界定問題：定義出要以資料採礦技術回答的問題或想要達成的目的。
- (2) 資料選擇：從龐大資料庫中選出部份資料，作為資料採礦的來源。
- (3) 取得資料：從資料的源頭取得所選擇的資料。
- (4) 資料清潔：確保取得的資料都是正確的，就是簡單、乾淨，沒有錯誤或多餘的資料。
- (5) 資料轉換：將資料轉換到正確的層級，層級就是資料被處理的單位，為的是確保資料都能在正確的層級下做加總等分析運算。
- (6) 資料加上衍生變數：除了原始資料外，也可以利用原始資料的值做運算後所得的衍生變數來進行資料採礦後續工作。
- (7) 建立訓練資料集：資料備齊後，至少要將其分為訓練及測試兩組，以因應後續步驟建模與評量所需，分組原則可以隨機分配、時間分配或地區分配，一般而言，為避免分組偏誤而造成模型不具代表性，通常以隨機分配方式分組為最佳。
- (8) 選擇模型建立方法：依據所面對的問題，選擇最恰當的資料採礦技術與演算法則來建立模型。
- (9) 訓練模型：透過選定的資料採礦技術以訓練資料集為資料來源，將預測模型建立起來或訓練起來。
- (10) 確認模型效率：以分為測試組的資料來檢驗前一步驟所建立的模型，可以利用錯差矩陣(Confusion matrix)，來驗證模型的預測能力，有多少是正確的及有多少是不正確的。
- (11) 選取最佳模型提出策略：驗證效率步驟的結果若是不佳時，可能會進行一些調整，如加入衍生變數、另選一組訓練資料或用別的採礦技術來重建模型，做幾次後比較出最佳模型來回答當初所提出的問題，並提出策略。
- (12) 依最佳模型策略採取實際行動：將資料採礦所得到的答案，實際運用在決策或管理方面。
- (13) 蒐集實際行動所產生的資料：最後依據行動執行後的結果來了解是否有解決問題或對組織有幫助，可以將這些新的資料當作下一次資料採礦的輸入，以便調整模型使其更有效率。



資料來源：本研究整理

圖 2-5 資料採礦技術流程圖

簡而言之，資料庫知識發現（KDD）是一種知識發現的一連串程序，而資料採礦為 KDD 中一重要的程序，而其最終的目的即為組織取得決策支援所需的資訊，這個資訊是突破盲點、能發現人所未見的知識和訊息，能替組織取得競爭優勢或提升管理效率。

2.3.3 資料採礦之功能與技術

依據尹相志（民 92）、倪達洸（民 91）等文獻，本研究綜合整理資料採礦之功能與技術。一般而言，Data Mining 功能主要包含下列六項，茲將這些功能的意義及使用的演算法簡述如下：

- 分類分析（Classification Analysis）

分類分析是從已知類別的物件集合中，依據其屬性（可能影響物件類別的變數）建立一個分類模型（如決策樹或決策法則）來描述物件屬性與類別之關係，然後再根據這個分類模型對其他未經分類或是新的資料進行預測。換言之，分類分析是一種依資料屬性建立類別的過程，通常從資料中產生「若則」法則，例如可藉由信用資料來預測客戶是否會呆帳，或藉由病歷來預測病人是否得到胃病。分類還可以透過分類結果的純度（Purity）來提供分類機率，這些分類機率可以作為名單排序或是篩選的重要依據。

分類分析中常用的演算法，包括決策樹（ID3、C4.5、CART、CHAID 等）以及類神經網路等。

- 預測分析（Prediction Analysis）

預測與分類的差異在於分類問題是推估「類別變數」，而預測問題主要是用來推估「連續變數」。以歷史連續數值來找出未來連續數值的量，例如利用過去降雨紀錄預測未來降雨量或是利用過去路段旅行時間紀錄預測未來的路段旅行時間等。

預測分析中常用的演算法，包括迴歸（線性迴歸、羅吉斯迴歸）、決策樹（CART）以及類神經網路等。

- 群集分析（Clustering Analysis）

群集分析係指所有的物件或資料分成若干群集的過程，也就是根據物件間的相似性（或不相似性），將所有的物件分成若干個群集，使得每個群集內的物件具有高度的相似性，而不同群集間具有高度的不相似性。群集分析的目的是要把群集與群集間的差異找出來，同時也要將群集中物件的相似性找出來。群集分析與分類分析的差別在於，群集分析是屬於非監督型學習的方式，使用者只須提供學習的學習資料組，系統必須自行發掘物件間的相似性和其他特性（如最佳群集數），以便建立模式和群集描述，作為未分群資料自動歸類的依據。

群集分析常用的演算法包括：K-平均分析法（K-Means Method）、EM 以及自我組織地圖（Self-Organization, SOM）等。

● 關聯規則分析（Association Rule Analysis）

關聯規則（Association Rule）就是於歷史資料中，找出哪些物件或是事件總是相伴發生，亦稱作「購物籃分析」（Market Basket Analysis, MBA）。通常運用於商業交易的作法為蒐集一組交易資料，每一交易包含若干交易項目，關聯分析的目的是由這些交易資料中，找出交易項目的關聯規則（Association Rule）。

關聯分析提供的資訊，若產品 A 是某事件的一部份，則產品 B 也出現在該事件的機率有多少，例如買牛奶的客戶，同時也買麵包的機率為 80%。此關聯的表示可能是正向，也有可能是負向，正向表示屬互補關聯，負向表示屬互斥關聯。關聯規則提供如下列分析描述「若 A、B、C 三種交易項目發生時，會發生交易項目 D 的機率為 p，機率越高表示關聯性越高」。就零售超級市場而言，係針對每一交易項目，分析會一起交易的項目有哪些？較著名的例子為分析超級市場交易中，發現可能會與某一商品一起購買的其他商品，例如尿布與啤酒、牛奶與麵包）。藉由這些消費者購物行為分析，業者可以調整貨架上的擺設位置，或調整存貨與訂單數量，進而設計促銷活動，以期更有效的銷售貨品。

在關聯規則分析常用的演算法包括：Apriori 以及通用規則建構（General Rule Instruction, GRI）。

● 序列分析（Sequential Pattern Analysis）

序列相關分析的目的即是由一群有次序性的交易中，找出經常循序出現的交易項目，進而了解顧客的長期購買行為。例如大多數的客戶購買洗衣機後會購買烘衣機；購買錄影機，在一段時間後會購買錄影帶；或當買新房子後一個月內買新爐具的比率為 45%，而二週內購買新冰箱的比率為 60%。換言之，即從「一段時間」的資料中，找出物品交易的相關性、相依性。經由長期分析購買者行為，業者可據以調整存貨與訂單數量，進而向已購買第一種產品的顧客促銷與其具有相依性的產品。

另外所謂時間序列曲線（Time-series Sequence）是指一連串實數所組成的記錄，而每個實數表示某一時間點所對應到的數值。例如，這一連串的實數可用來表示股價、銷售量、通話量、費率等會隨時間而改變的資料，在時間序列型資料上的應用，主要在於曲線的相似性查詢（Similarity Query），包括曲線完全相似

比對 (Whole Matching) 或子曲線相似比對 (Subsequence Matching)。假設有 n 個時間序列曲線集合 S_1, S_2, \dots, S_n 及一查詢曲線 (Query Sequence) Q ，完全相似比對是去尋找那些與 Q 相似的曲線，這些曲線的長度須相同；至於子曲線相似比對是指在任意長度的曲線中找出與查詢 Q 相近似的子曲線。探索時間序列型態資料 (Time-series Patterns) 的相似性關係，可用來偵測事件或者預測未來的行為發展，例如找出股價運行模式相當接近的股票、去年銷售模式雷同的商品、比較過去成長模式預測企業未來的發展等。

序列分析中常用的演算法包括：Apriori、時間序列、決策樹以及類神經網路等方法。

● 描述 (description)

描述是資料採礦過程中的一項重要附加價值，在進行資料採礦的前置作業時，透過資料視覺化，能夠將資料特性呈現在資料採礦者面前，不可否認的，人的眼睛有時是最好的資料採礦工具，透過良好的資料視覺化程序，可以發覺許多演算法無法判別的規則型態。

2.3.4 資料採礦之應用

Kuhnert 等學者 (2000) 為了避免車禍發生受傷的情形出現，透過透過非參數模式 (non-parametric models) 來分析，包括分類與迴歸樹 (Classification and Regression Tress, CART) 與多變異適應迴歸曲線 (Multivariate Adaptive Regression Splines, MARS)，分析出會受傷的高風險族群，其研究對象為澳洲布里斯本因交通意外而受傷的案例，在醫院所做的問卷調查，對同一地區持有駕照者抽樣，共有 689 個樣本，將資料分成個案與控制兩組，個案組為交通意外受傷之樣本，控制組為持有駕照者過去 2 年未發生交通意外受傷之樣本。資料收集上除了駕駛的年齡、性別、駕車經驗與駕駛習慣，還有安全帶的使用、酒後駕車、車速、疲勞駕駛等。在 CART 分析中可以從樹狀結構中發現，年齡在 27.5 歲以下在所有類別中最上層的節點被區分，其中小於 27.5 歲的駕駛人被分類屬於交通意外受傷類，而在大於 27.5 歲者，則又會因為未使用安全帶亦被分類成交通意外受傷類，最後才是性別與年紀較大者的分類。透過 MARS 分析結果可以發現，有交互作用的變數為年齡、有幾年的駕駛經驗與安全帶使用與否，在有使用安全帶的情況下，駕駛經驗較低者，各年齡成普遍都有高的肇事受傷風險，年齡在 30 到 45 者其駕駛經驗較低者，都會有較高的受傷風險。將 CART 重要分類變數與 MARS

有交互作用變數，再透過羅吉斯迴歸分析，結果發現，年齡在 27.5 歲以上，無使用安全帶者的受傷機會是高於使用安全帶者的 11 倍以上，但是否使用安全帶對於 27.5 歲以下卻不顯著。綜合結果可以瞭解到，發生交通意外時會受傷的風險族群有三類，分別為年齡 30 以下駕駛經驗較低者，年齡在 30 到 45 之間有較長的駕駛經驗者，以及老年人駕駛經驗較不足者，而駕駛經驗與是否使用安全帶之間的交互作用，可能與駕駛為冒險行為者有關。而就方法而言，CART 可以找到重要之分類變數，MARS 可分析出交互作用之變數，羅吉斯迴歸可分析出族群與嚴重程度間的關係，因此結合非參數之分析方法對實際應用為可行的方法。

Giudici 等人（2002）從消費者購買行為上瞭解到商品與購買的關係，透過關聯式的結構來描述消費者行為（Consumer behavior），這些相關的商品可做為賣場促銷與銷售時做為參考，分析資料為義大利的大型超市，在 1997 年共 52 週，每週所賣出的 26 種食品類產品，研究方法先以購物籃分析（Market basket analysis, MBA），找出產品與產品之間的關聯，再計算出這些關聯中產品的損益比（odds ratio），找出最大的 5 個商品集合（cluster），而此關聯性的產品可做為組合促銷之依據，但由於解釋上雖可行，卻沒有實際的數據可以了解到這樣的促銷所能帶來實際的增加的銷售量，因此，結合馬可夫鏈蒙地卡羅演算法（Markov Chain Monte Carlo composition algorithm, MCMC）與貝氏定理（Bayesian Model），對相同的資料進行模擬分析，其中各個商品之間的關係，以機率值表示能提昇銷售量之大小。從本研究發現的產品組合皆可在實際銷售上有所幫助，如季節性的銷售產品，或是飲料與食品之組合等，另外，在方法上由於購物籃分析只能分析出消費者有同時購買之產品關係，因此透過貝氏定理與 MCMC 演算法，才能發現產品組合所帶來的效益，但貝氏定理與 MCMC 需編碼與撰寫程式，對於實質應用上較不方便。

Rygielski 等人（2002）為瞭解資料挖掘技術應用在顧客關係管理（CRM）的成效，透過類神經網路（Neural Network）、卡方自動互動檢視法（Chi-square Automated Interaction Detection, CHAID）兩種方法，分別探討兩個實際的案例。類神經網路部分是以預測一零售商最低存貨量，使其能滿足最多的消費者，以達到最低成本與獲得最高利潤，結果發現降低 2% 存貨量會可增加 11.6% 的利潤；CHAID 部分則是分析持有信用卡的族群，結果發現可分為 16 個族群。比較類神經網路與 CHAID 中發現，在結果解釋、模式應用的難易與使用的成本來看 CHAID 是較佳的，但預測的準確度卻是以類神經網路較好；對於資料需求方面，

類神經網路需要將資料重新編碼，而 CHAID 需要將連續性資料轉成離散型資料才能處理，而目前各領域應用類神經網路較 CHAID 廣泛，但處理上較繁雜，所花費的時間成本高，因此在 CRM 領域中應用資料挖掘工具時，類神經網路與 CHAID 之間的取捨就需要依實際案例來考量。

Anand 等人 (2005) 因近年來對於事故資料庫的建立，且由於資料量越來越大，決策者不能再由以往的直覺判斷來做決策，而是要以這些豐富的資料來進行決策，故利用資料採礦的技術來進行事故原因的分析。其中主要將資料採礦分為兩類，一類為傳統之統計技術，另一類為新一代的技術如決策樹(Decision Trees)與關聯規則(Association Rules)。本研究所主要使用的資料採礦技術即為決策樹法與關聯規則。該研究主要使用美國國家緊急回報中心(National Response Center, NRC)事故回報資訊系統(Incident Reporting Information System, IRIS)之資料庫，並擷取 1990 年至 2002 年在哈瑞思城(Harris County, TX)中的石化工廠所發生的事故，共有 7265 筆事故為其研究的範圍。資料採礦取得資料後得先對於資料先進行初步的檢視分析，依據決策樹的步驟，初步分析事故的發生原因主要為設備的失效，其次為未知的原因；接下來在對於設備失效其中的原因再進行分類，可得知儲存桶所造成之比例最高，其次為管線。依據關聯規則解釋其原因，其中 NOx 對於電子設備危害機率最大，其 Lift 值高達將近 4 次，同樣地，在管線事故中油類相關的發生機率，Lift 值高達 4.84 次，皆比其他的情況高出不少，可利用此結果進一步進行預測。

因此資料採礦的優勢是可以對於「乾淨」的資料進行分析和預測，以及後續的驗證，故資料的篩選對於資料採礦而言是相當重要的一環。

2.4 駕駛員管理相關回顧

劉建良 (民 89) 研究利用「PZB 服務品質概念性模式」並以公車服務特性為基礎，建構出適合公車運輸系統的服務品質模式與衡量方法，並發現影響公車服務品質的因素很多，其中影響服務品質的最重要關鍵因素為駕駛員的管理，不論是「服務前」的職前訓練，或是「服務中」的參與、提供服務，甚至是「服務後」的獎懲制度都會是影響公車服務品質優劣之主要因素。分析影響駕駛員服務品質的原因並對於駕駛員管理提出相關的對策，其主要提出的對策分為三大項，「服務前」：改善人員晉用與訓練制度、強化場站人員對駕駛員之管理權責與技能，「服務中」：裝置數位式行車紀錄器，「服務後」：薪資結構之調整以及紓解駕

駛員壓力與提高駕駛員工作滿足感。

陳國樑（民 93）之研究蒐集中部某家經營國道公路及市區客運公司從民國 88 年至 91 年四年間共有 138 筆的肇事、保險理賠資料進行統計整理，並以該公司的駕駛員為研究對象採問卷調查方式，發放問卷分別以教育程度、服務年資、年齡分組等針對於公司管理制度、情緒反應/駕駛行為、駕駛員車輛管理制度及道路環境三個構面進行一般統計分析與交叉分析，並以 LISREL 線性結構軟體探討分析該公司管理制度對駕駛員、情緒反應、駕駛行為與車輛管理制度的因果關係，經由其模式參數意義與校估結果得知以下四個結果：

- 公司管理制度與車輛管理制度成顯著正向的直接影響關係。
- 公司管理制度與情緒反應、駕駛行為成顯著正向的直接影響關係。
- 公司管理制度與駕駛員成顯著正向的直接影響關係。
- 駕駛員與情緒反應、駕駛行為之關係為顯著正向的直接影響關係。

2.5 小結

在數位式行車紀錄器產品與特性方面主要可以區分為國外與國內的產品，國內廠商主要有寶錄、彙通與康訊公司，前兩者其數位式行車紀錄器皆有感測器，可直接量測車輛速度，並記錄煞車、車燈、方向燈、車門開關等訊號，康訊公司之設備則由 GPS 訊號計算車輛速度，並無煞車、車燈等之紀錄。寶錄公司之行車紀錄器利用記憶卡匣讀取主機所紀錄之資料，主機具有按鈕可設定功能，彙通公司則由 PDA 設定主機與下載資料，康訊公司則利用 RS-232 與電腦連線下載資料與設定主機。

國內數位式行車紀錄器與歐盟主要不同之處，在於歐盟是利用 IC 卡做資料下載與上傳，且 IC 卡專屬於個人，可追蹤駕駛個人的駕駛時間與紀錄，其概念偏向於個人駕駛管理，可應用於非商用車隊之車輛。紀錄器並附有印表機，可列印相關資料以供路邊臨檢，異常狀況的警告屬於標準功能。

國外屬車隊管理之數位行車紀錄器，如 Siemens VDO 公司旗下的 FM200 或 FM300 系列產品，其主要的形式類似飛行器的黑盒子，沒有印表機功能但有警示功能，資料的下載與傳輸以隨身碟(記憶卡)或是 GSM/GPRS 簡訊傳送下載至資料庫之方式。國內廠商所開發之產品與國外後者較為類似，國內目前的產品是利用記憶卡匣或 PDA 下載主機資料，沒有印表機，警告功能並非標準配備。

整體而言，國內外的數位式行車紀錄器能夠紀錄與輸出的資料項目大致都類似，能夠直接即時監控與紀錄駕駛員異常操作行為的主要有德國 Siemens VDO 與澳洲 Circuitlink 公司的數位式行車紀錄器，其可藉由事先設定駕駛員異常操作的項目與門檻值，直接藉由該設定來判定與監控紀錄，而一般其他數位式行車紀錄器需記錄大量之資料並完全仰賴後端軟體之判讀。這樣的設計可以大量減低後端判讀分析之資料量與增加易讀性。

本研究欲探討公路客運駕駛異常操作行為對於行車安全與油耗、保養維修費用的影響，故分別就行車安全、燃油消耗、機件耗損等方面進行文獻回顧與探討，並作為後面章節駕駛操作行為監控項目與門檻值設定之依據。

在行車安全方面可以得知超速、未保持行車安全間距以及任意變換車道等不當駕駛行為，會增加造成肇事的風險，以及急加減速亦會對於車輛碰撞的機率增加。在燃油消耗方面，可以得知不正常之駕駛行為如使引擎轉速過高或是急加減速等行為會造成油料的額外損耗。在機件磨損方面，不當之油門與煞車操作，會造成機件磨損的增加，甚至損壞。由上述之相關文獻回顧可知，駕駛操作的異常會對於行車安全、燃油消耗以及機件磨損甚至行車舒適度有顯著的影響，故使用數位式行車紀錄器監控相關之項目，並依據合理之範圍定訂門檻值以界定正常與異常之駕駛行為，並利用數位行車紀錄器所記錄蒐集之資料，進一步對於駕駛員之異常駕駛行為進行分析與探討，以釐清駕駛員之問題所在。

由相關之資料採礦研究可以得知，資料採礦可以應用的層面相當廣泛，目前主要應用的範圍分別於金融服務業、保險業以及電信業方面有客戶貢獻度分析、信用評分、風險評估、客戶區隔、交叉行銷等研究；製造業方面有客戶貢獻度分析、品質管制、行銷績效分析、生產分析和存貨分析等研究；以及零售業方面的客戶忠誠度、客戶區隔、購物籃分析、定價分析、交叉行銷和銷售預測等另外還有生物科技、醫療保健、航太空業、環境、法律等各領域，唯有在交通運輸方面的運用仍屬偏少，故本研究將以數位式行車紀錄器所擁有的龐大資料庫特性，運用資料採礦的技術來進行數位式行車紀錄器資料分析，以釐清駕駛員異常操作行為對於肇事、交通違規、油耗與保養維修費用之關係，並對於駕駛員進行分群與管理分析。

駕駛員管理相關研究，公司之管理方式與制度，會對於駕駛員之情緒反應與駕駛行為有正面之直接影響，故可以得知公司有良好的管理制度，是能夠提升駕駛員的素質與表現，進而提升行車安全與減少異常操作之駕駛行為。

第三章 研究方法與理論基礎

本章針對前述研究目的與相關文獻的探討，提出本研究之研究方法，說明資料採礦作業處理之程序，以及所使用資料採礦方法之理論基礎。

3.1 研究架構

本研究資料採礦分析架構包括選取輸入資料、轉換處理資料、執行採礦方法及檢定解釋結果，其中資料採礦方法包括多元迴歸分析以及群集分析，多元迴歸分析是用來構建預測模式，而群集分析是用來對於駕駛員進行分群與管理。

一、資料選取與設定：

資料採礦首要的步驟，為選定欲採礦和分析的資料，並選定資料來源，而且，資料不一定為特定資料庫中之所有資料，採礦之資料可能從一個或多個資料庫中，取得表格、概略表或記錄文字檔。因此，使用者可根據本身所需設定條件，以篩選出所需的資料。

二、資料前置處理：

為確認資料正確與完整性並順利進行採礦步驟，首先須將已收集的資料進行整理、清除重複或無效的資料記錄，確認所需資料的完整。並且將資料轉換為採礦軟體所能讀取之存取格式，以利進行資料庫構建。

三、資料採礦：

依研究之目的，選定合適的資料採礦方法，本研究所選用之資料採礦方法為多元迴歸分析與群集分析，利用這些資料採礦方法從大型數位式行車紀錄器資料庫中發掘存在的多種特徵及資訊並進行預測。

四、解釋結果：

經過資料採礦方法之分析後，可以得到相關之模式，並從模式中可以得到變數之間的關係，利用其結果作為模式完成的重要基礎。

五、駕駛員分群與管理

以群集分析方法，並輸入之前多元迴歸分析之相關變數作為分群決策變數，對於駕駛員進行分群，並依據模式所得到之結果，經解釋與說明後，提出相關的

駕駛員管理方式與對策。

本研究之研究架構流程如下圖 3-1。

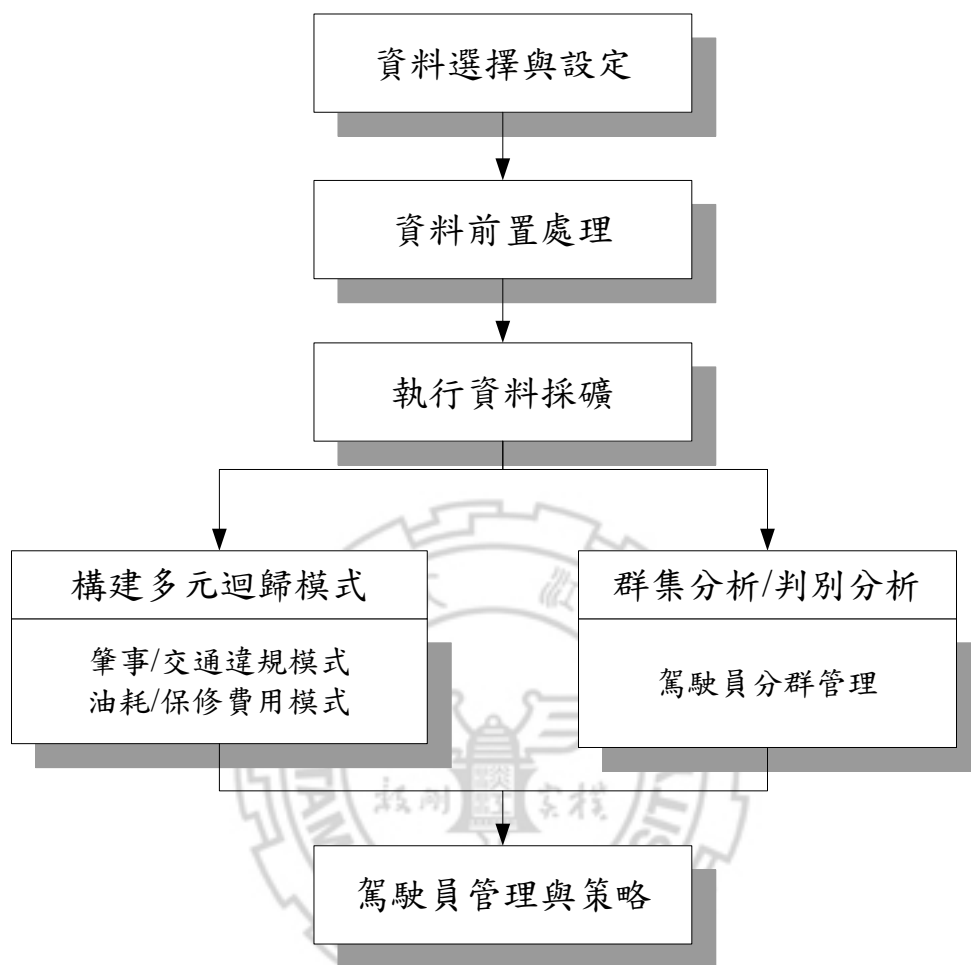


圖 3-1 研究架構圖

3.2 資料處理作業程序

在進行資料採礦之前，首先是要進行資料的蒐集與取得，有了與研究目的相關的資料後，便需要對於資料進行相關前置處理，以確保資料範圍以及資料的可信度，並保證資料的清潔度。要構建出好的資料採礦模式，就是必須將資料的相關前置處理謹慎的進行。於本節介紹本研究資料蒐集之項目、內容與範圍，以及資料的前置處理過程。

3.2.1 資料蒐集

本研究蒐集國內某客運公司於民國 93 年 6 月至民國 94 年 12 月的數位行車

紀錄器資料庫以及相關駕駛與車輛之資料，並釐清各項資料格式與內容。

- 研究期間為民國 93 年 6 月至民國 94 年 12 月，一共 19 個月。
- 數位式行車紀錄器資料庫，SQL 資料庫格式。

該客運公司所採用之數位式行車紀錄器為德國西門子 VDO 公司所生產的 VDO FM200 Plus 數位式行車紀錄器，所紀錄之資料以 SQL 資料庫格式儲存，數位式行車紀錄器紀錄之資料表項目相當多，依據本研究之目的，選定所需之相關資料表，主要有駕駛員資料(Drivers)、車輛資料(Vehicle)、駕駛事件資料(Event Data)以及事件定義敘述(Event Description)等，並於後續將這些資料表之資料進行彙整與篩選。

- 駕駛員基本資料

資料項目包含駕駛員姓名、員工編號、生日、駕照審驗日期以及駕駛車輛等，資料儲存格式為 Excel 格式。

- 駕駛員交通違規資料

資料項目包含違規日期、時間、車號、駕駛員姓名、違規事項以及罰款金額等，資料儲存格式為 Excel 格式。

- 駕駛員肇事資料

資料項目包含肇事日期、車號、駕駛員姓名、肇事狀況敘述、維修項目以及建議懲處等，資料儲存格式為 Excel 格式。

- 車輛油耗與維修費用資料

資料項目包含車號、每月定期保養與工資費用、車廠維修耗材金額、附屬油品金額、油量總額費用以及平均每公里之費用等，資料儲存格式為 Excel 格式。

3.2.2 資料前置處理

前置處理為在進行資料採礦前的一個重要處理過程，不同的採礦領域所需的前置處理技術也不盡相同(Bhandari, 1997; Simoudis, 1996)。本研究之資料前置處理部分主要有下列幾個步驟：

一、資料選擇

於已經蒐集之資料中，將各項資料所包含之項目，依據本研究目的，篩選所

需之項目。

二、資料雜訊清除：

將各項資料中含有亂碼或是有部分遺漏之資料與以清除，並且查核字元之正確性，以及刪除不需要之資料。

三、資料屬性核對與合併輸入

核對各項資料之屬性，並且將對應之資料與以合併輸入，如駕駛員之資料合併輸入。

四、資料格式轉換

由於資料格式有SQL資料庫格式與Excel格式，故需將資料格式統一為SQL資料庫格式，以利後續資料庫之構建。

五、資料庫構建

將原始資料經合併、轉換與篩選完成之資料，構建成本研究所需要的新資料庫，始能進行後續分析。

3.3 資料採礦方法

資料經由相關之處理程序後，即可針對研究目的選取適用之資料採礦方法，資料採礦可使用之技術或方法相當多，如線性迴歸、邏輯斯迴歸、類神經網路、決策樹、群集分析或是關聯規則等，本研究以多元迴歸方法構建駕駛員異常操作行為對於行車安全以及行車油耗與保養維修費用關係之模式，並進而依據迴歸模式所輸入與校估後之變數，進行駕駛員群集分析，並提出相關管理對策。

3.3.1 多元迴歸分析

（一）迴歸分析之意義

迴歸（Regression）是一種試圖以一個或多個獨立自變數（Independent variable）來解釋另一個相依因變數（Dependent Variable），然後利用所獲得之樣本資料去估計模型中參數的計量分析方法。而獨立變數與相依變數的迴歸分析模型中，一般又可分為線性與非線性關係，本研究於後續迴歸模式之構建是採用線性關係的多元迴歸模式。

多元迴歸分析係指單一反應變量Y對k個解釋變數（ X_1, X_2, \dots, X_k ）之統計模

式。其探討重點在於 k 個斜率項 $(\beta_1, \beta_2, \dots, \beta_k)$ ，探討 X_k 之解釋能力是否顯著。模式如下：

$$Y_n = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \dots + \beta_k X_{nk} + e_n$$

其中 Y_n 為因變數， $X_{n1}, X_{n2}, \dots, X_{nk}$ 為自變數，

若有 n 個樣本資料，則可表示成：

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \beta_0 + \beta_1 \chi_{11} + \beta_2 \chi_{21} + \dots + \beta_m \chi_{m1} + \varepsilon_1 \\ \beta_0 + \beta_1 \chi_{12} + \beta_2 \chi_{22} + \dots + \beta_m \chi_{m2} + \varepsilon_2 \\ \vdots \\ \beta_0 + \beta_1 \chi_{1n} + \beta_2 \chi_{2n} + \dots + \beta_m \chi_{mn} + \varepsilon_n \end{bmatrix}$$

多元迴歸模式 $Y = X \cdot \beta + \varepsilon$ 可採用最小平方方法求得迴歸係數 β 的數值，最小平方方法目的是要找出未知母數的數值，使誤差平方和(Error Sum-of Squares, SSE)為最小。

(二) 迴歸分析基本假設

1. 自變數與應變數間皆應為線性關係，若不符線性條件，則應變數須轉換型態為對數、平方、三次方或其他影響變數與因變數間呈線性關係之型態。
2. 誤差項之期望值為零，即 $E(e_n) = 0$ 。
3. 誤差項變異數為一常數：誤差項變異數相同（均質性），即 $V(e_i) = V(e_j) = \sigma^2, \forall i \neq j$ 。
4. 觀測值互相獨立，無自我相關現象：即 $E(e_i, e_j) = 0, \forall i \neq j$ 。各機率分配之隨機變數 Y_i 間相互獨立，若違反此假設，可能產生自我相關或事件相依情況。
5. 誤差項為常態分配：即 $e_i \sim N(0, \sigma^2)$

(三) 迴歸變數解釋能力

迴歸分析中，其自變數 X 對應變數 Y 有無解釋能力，可分為兩步驟說明，一為有無解釋能力之問題，即自變數與因變數有無相關，可採用相關分析；或檢定 β_i 是否為零，可採用 t 檢定探討個別變數是否顯著。二為解釋能力之高低。其做法如下：

● 總變異之分解

由ANOVA表可知，因變數 Y 之總平方和可分解為模式平方和與誤差平方

和之加總，即總變異為可解釋變異與不可解釋變異之和， $SST = SSR + SSE$ 。

● 判定係數 R^2 值

判定係數之意義相當於總變異中可被解釋之百分比比例，為模式配適度（Goodness of Fit）之指標，即

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}, \quad 0 \leq R^2 \leq 1$$

其中判定係數 R^2 值越大，表示迴歸模式配適度越高，亦即預測值越接近觀察值。若引入一新解釋變數，若其變數具解釋能力，則模式判定係數必然提高。然而，解釋變數每增加一個，也代表誤差自由度隨之減少一個，由於判定係數無法顯示模式之解釋能力可能降低，有鑑於此，實務上多採調整後判定係數，以彌補其缺點。

$$Adj. R^2 = 1 - \frac{SSE/df_e}{SST/df_t} = 1 - \frac{SSE/(n-p-1)}{SST/(n-1)}, \quad 0 \leq R^2 \leq 1$$

（四）迴歸模式檢定

迴歸模式建立後，應就其統計顯著性加以檢定。最常見之顯著檢定分為兩種：t檢定（t-test）與F檢定。t檢定為測定準則變數Y與個別預測變數 X_j 間之統計關係，F檢定則將所有預測變數視為一個整體，而測定所有預測變數X之間是否有顯著統計關係存在。

● t檢定（t-test）：

檢定迴歸係數是否具有統計上之顯著意義，即檢定每一係數是否為零。其虛無假設與對立假設為：

$$H_0: \beta_j = 0$$

$$H_1: \beta_j \neq 0$$

$$t_j = \frac{b_j - \beta_j}{S_{b_j}} = \frac{b_j}{S_{b_j}}$$

其中 β_j 為迴歸係數， S_{b_j} 為 b_j 之估計標準差。

若t檢定結果，拒絕虛無假設： $H_0: \beta_j = 0$ ，即 β_j 不等於零，即Y與 X_j 之間存在有顯著性直線關係；若t檢定結果，接受虛無假設： $H_0: \beta_j = 0$ ，表示無足夠統計證據支持Y與 X_j 之間有顯著關係存在，即可將 X_j 項目之變數去除。

● F檢定（F-test）：

F值檢定自變數集合與因變數之間是否具有顯著關係。以最小平方法所估計之迴歸式，須以F統計量來檢定整個迴歸式是否有意義，即檢定所有係數是否為零，若所有係數均為零，則所校估之迴歸式並無意義。其虛無假設與對立假設為：

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_n = 0$$

$$H_1 : \beta_i \text{ 不全為 } 0, i = 1, 2, \dots, n$$

$$F = \frac{MSR}{MSE} = \frac{SSR/(k-1)}{SSE/(n-k)}$$

其中SSR為迴歸平方和，SSE為誤差平方和。若以0.05為顯著水準，則F值大於 $F_{(1-0.05, k-1, n-k)}$ ，則拒絕虛無假設 H_0 。

基於上述理論，迴歸分析除可用以預測外，尚可解釋變數之間影響關係，探討自變數與因變數有無相關，本研究係將使用迴歸分析之基本假設、目的與模式設定說明如下：

- 基本假設

本研究使用迴歸分析用以解釋如急減速、急加速、超速、電磁煞車操作異常、檔位異常、怠速過久以及引擎轉速異常等駕駛員異常操作行為，與客運車輛的燃油費、附屬油品以及修車材料費用之間的相互影響關係，以及這些異常操作行為與肇事和交通違規的相互關係。

- 目的

由於客運之營運需求，故每部客運車輛通常皆會有兩位或兩位以上之駕駛員輪流排班執勤，使得各類變數難以直接察覺相互影響關係，若要了解單一駕駛員之異常操作行為對於行車油耗與保養維修費用之影響關係，必須先了解整體駕駛員異常操作行為對於相關費用之影響關係與趨勢，故本研究乃以迴歸分析建構解釋其關係之模式。

- 模式設定

先以強迫進入法將所有的異常操作行為的變數輸入，在依據最小平方法以及相關的統計指標，並去除不顯著的輸入變數，找尋駕駛員異常操作行為對於肇事、交通違規以及油耗與保養維修費用解釋力較高的模式，選定模式後即可解釋其對應關係與各變數影響之程度。

3.3.2 群集分析

(一) 群集分析之意義

群集分析（Cluster Analysis）係指根據一組準則變數（Criterion Variables）將N個個案集成K個群別的統計方法，其中 $K \leq N$ 。其目的在辨認在某些特性上相似的事物，並將這些事物按照這些特性劃分成K個群集，使在同一群集內的事物具有高度的同質性，而不同群集間則具有高度的異質性。

（二）決策流程

群集分析的決策流程由圖 3-2 所示，包括研究問題、選擇變數、衡量相似性、選擇群集方法、決定群集數、解釋群集等六步驟。

1. 研究問題：

首先要確定群集分析的主要目的是探索性的或是驗證性的。傳統上，群集分析最常用於探索性的用途，即將一組事物作客觀的分類。

2. 變數的選擇：

不論研究的目的是探索性的或驗證性的，群集分析的結果都受到所選變數的限制，變數的選擇必須兼顧理論、觀念和實務的考量，不應該毫無選擇的把不相關的變數納進來，因為群集分析技術無法區分相關和不相關的變數，把與研究目的不相關的變數選進來將增加產生異常變數（Outliers）的機會，對於群集分析的結果會有重大的影響。

3. 相似性的衡量：

衡量事物間的相似性（Interobject Similarity），亦即事物與事物間的相似程度。各事物間相似程度的衡量方法有許多種，大致可分為（1）距離衡量、（2）關聯衡量、（3）相關衡量等三類。

4. 群集方法的選擇：

將成對事物間的相似性加以衡量後，接著就可利用群集方法將各樣本單位歸入群集中。群集的方法可分成階層式群集方法（Hierarchical Methods）和非階層式群集方法（Nonhierarchical Methods）兩大類。

5. 群集數的選擇：

群集數目的決定，是群集分析的一項重要決策，如群集數目較多，群集內事物的同質性會較高，但較不容易找出資料的結構；如群集數目較少，雖較能看出資料的簡單結構，但群集內事物的同質性則較低。但有關群集數目的決定尚無一客觀的標準程序可供遵循，通常可以各連續群集步驟下群集間的距離作為參考。

6. 群集的解釋：

一旦經由群集分析而找出群集後，應設法描述這些群集。常用的一種方法是以群集之重心（Centroid）——即群集內的各事物點在各變數上的平均值來描述該群集。除此之外，我們亦可計算群集的變異情形，如群集內各點間的平均距離或各點與重心間之平均距離來輔助描述該群集。

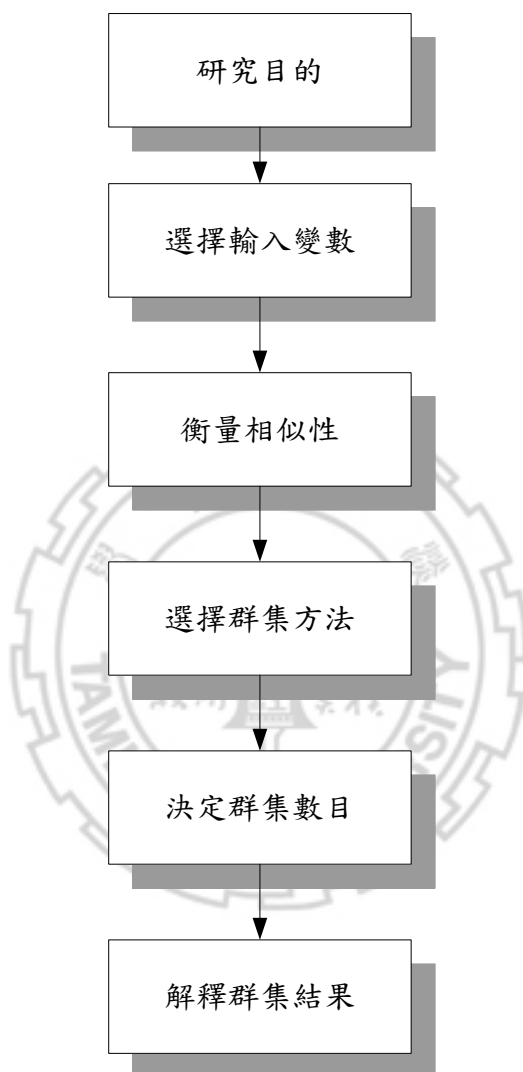


圖3-2 群集分析之決策流程

（三）相似性衡量

在選擇了相關的變數、剔除異常事物以及檢查假定之後，接著要衡量事物間的相似性亦即定義事物與事物間的相似程度，即個體間的距離。大致可分為距離衡量、關聯衡量以及相關衡量等三大類。其中，相關衡量是先計算成對事物間的相關係數，然後以各事物間相關係數的型態而非大小來判定事物間的相似性。由於大多數群集分析的應用都強調數值的大小而非型態，因此很少使用相關衡量，

以下將分別介紹距離衡量和關聯衡量兩類方法。

1. 距離衡量：

主要是以點與點間的距離為代表。點距間的距離則有歐基里德距離 (Euclidean Distance)、馬氏距離 (Mahalanobis Distance)、城市街道距離 (City-Block Distance)、混合式距離。一般分類常用為歐基里德距離，由於本研究也是採用歐基里德距離 (Euclidean Distance) 為相似性之衡量基礎，因此便依此法進行介紹。設有 n 件事物，每件事物有 m 個屬性，則第 k 件事物與第 l 件事物間的歐幾里德距離為

$$d_{ij} = \left[\sum_{p=1}^m (x_{kp} - x_{lp})^2 \right]^{1/2}$$

上式中 x_{ip} 和 x_{jp} 分別是在 m 度空間中， k 點和 l 點對變數 p ($p=1, 2, \dots, m$) 的投影。

2. 關聯衡量：

若事物的屬性全部以名目尺度或虛變數來表示，則兩事物的相似性可用配合係數 (Matching Coefficient) 或相似比 (Similarity Ratio) 來衡量。

● 配合係數 (Matching Coefficient)

配合係數的公式如下

$$S_{kl} = \frac{a+b}{m},$$

$$a = \sum_{p=1}^m x_{kp} x_{lp}$$

$$b = \sum_{p=1}^m (1 - x_{kp})(1 - x_{lp})$$

上式中， a 為 k 和 l 這兩個事件共同具有的屬性數目， b 為 k 和 l 共同不具有的屬性數目， m 為屬性總數。

● 相似比 (Similarity Ratio)

若不考慮事件 i 和 j 都不具有的屬性，則可得相似比的公式如下

$$SR_{kl} = \frac{a}{m-b}$$

上式中， a 為 k 和 l 這兩個事件共同具有的屬性數目， b 為 k 和 l 共同不具有的屬性數目， m 為屬性總數。

(四) 階層式群集方法

最常用的幾種階層是群集方法為最短距離法（又稱單一連結法 Single Linkage）、最遠距離法（又稱完全連結法 Complete Linkage）、平均連結法（Average Linkage）、重心法（Centroid）以及華德法（或稱最小變異數法 Ward Method or Minimum-Variance Method）。以下就 Omi 等學者（1979）所採用之最短距離法、Radloff and Betters（1978）所採用之平均連結法、最遠距離法以及 Briggs and France（1983）採用之華德法進行介紹。

- 最短距離法(Nearest Neighbor)

最近法對 P、Q 兩群距離定義是以 P 群內每一點到 Q 群內每一點的距離最小值當做 P、Q 兩群的距離，即 $d_{P,Q} = \underset{\substack{k \in P \\ l \in Q}}{\text{Min}} d_{k,l}$

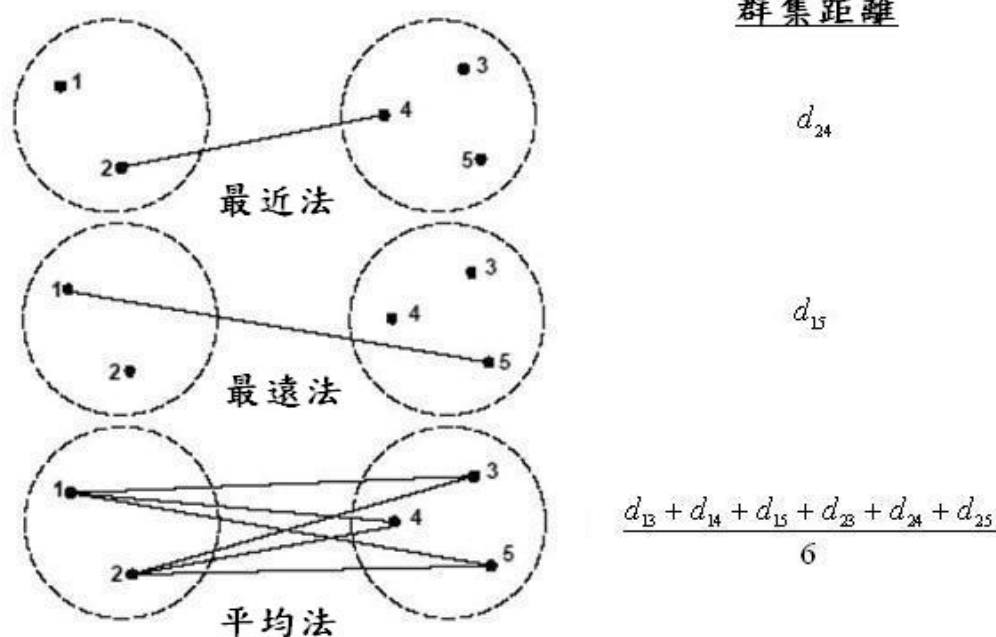
- 最遠距離法(Furthest Neighbor)

P、Q 最遠法距離的計算是以 P 群內每一點到 Q 群內每一點的距離中之最大值當做 P、Q 兩群的距離，即 $d_{P,Q} = \underset{\substack{k \in P \\ l \in Q}}{\text{Max}} d_{k,l}$

- 平均連結法(Average Linkage)

P、Q 兩群平均法的計算是以 P 群內每一點到 Q 群內每一點距離的平均，即 $d_{P,Q} = \frac{\sum_{k \in P} \sum_{l \in Q} d_{k,l}}{n}$ ，圖 3-3 為用圖形表示群集間距離。

群集距離



資料來源：黃俊英（2001）

圖3-3 群集間距離示意圖

● 華德法或稱最小變異數法（Ward Method or Minimum-Variance Method）

P、Q 兩群距離是以 P 群中心點 \bar{x}_P 到兩群合併中心點 \bar{x} 距離平方乘以 P 的點數，與 Q 中心點 \bar{x}_Q 到總中點的 \bar{x} 距離平方乘以 Q 的點數之和，即

$$d_{P,Q} = n_P \times \left\| \bar{x}_P - \bar{x} \right\|^2 + n_Q \times \left\| \bar{x}_Q - \bar{x} \right\|^2$$

（五）非階層式群集方法

非階層式集群分析中最常使用的方法為 K-Means 群集分析法，與其他的分群法最大的差異在於，若樣本的分群數為已知的情形之下，則 K-Means 群集分析法可以將所有的樣本分為指定的群數，而且 K-Means 群集分析法比較不容易受到不適當之屬性（構面）、異常點與不同的相似度（距離）的影響。

K-Means 群集分析法之樣本分群演算步驟如下：

1. 依據預先設定樣本分群的群數，先將樣本分為 N 個群組，而且每個群組的樣本成員選擇是經過資料庫中隨意選取所產生的。

2. 計算所有的樣本向量的歐基里德距離測度矩陣。

$$u_{i,j} = \{1, \text{if } \|x_j - N_i\|^2 \leq \|x_j - N_k\|^2, \text{ for each } k \neq i, \\ 0, \text{otherwise}\}$$

其中，1代表在歐基里德距離測度矩陣中，假如第j個樣本向量 x_j 屬於群組i，而0表示第j個向量 x_j 不屬於群組i。

3. 計算所有的樣本資料向量到各個群組中心樣本向量的成本函數。

$$J = \sum_{i=1}^N J_i = \sum_{i=1}^N \left(\sum_{k, x_k \in G_i} \|x_k - N_i\|^2 \right)$$

其中， $J_i = \sum_{k, x_k \in G_i} \|x_k - N_i\|^2$ 為在群組i內部的最小成本函數。

通常距離函數 $d(x_k, N_i)$ 可以被應用於在i群組的 x_k 向量，因此符合全部成本。函數可以寫成：

$$J = \sum_{i=1}^N J_i = \sum_{i=1}^N \left(\sum_{k, x_k \in G_i} \|x_k - N_i\|^2 \right)$$

並且將每個樣本資料分派到與其距離最近的中心樣本資料向量所屬的群組。

4. 重新計算每個群組樣本資料向量的中心點。

$$N_i = \frac{1}{|G_i|} \sum_{k, x_k \in G_i} x_k$$

其中， $|G_i|$ 為各個樣本群集所包含的樣本函數的大小。亦為 $|G_i| = \sum_{j=1}^n \mu_{ij}$

5. 重複步驟2、3 直到各群組沒有樣本資料向量被重新分派的情形出現為止。

基於上述之理論方法，本研究欲使用群集分析之基本假設、目的與作法架構說明如下：

1. 基本假設

使用迴歸分析中建立模式主要之變數，尋求各項異常操作情形類似之駕駛員，並將其分群，並假設駕駛員群數無大小關係。

2. 目的

將相同特性之駕駛員歸類為同一群集中，依照所選入之異常操作項目，包括行車安全、燃油消耗與機件磨損等層面，以期能夠將駕駛員分群並進行管理。

3. 群集架構

將各分群變數輸入如急減速、急加速、超速、電磁煞車操作異常、怠速過久以及引擎轉速異常等項目變數，透過群集分析方法，得到各個群集，並依照各群集之特性進行命名。分析架構如圖3-4。

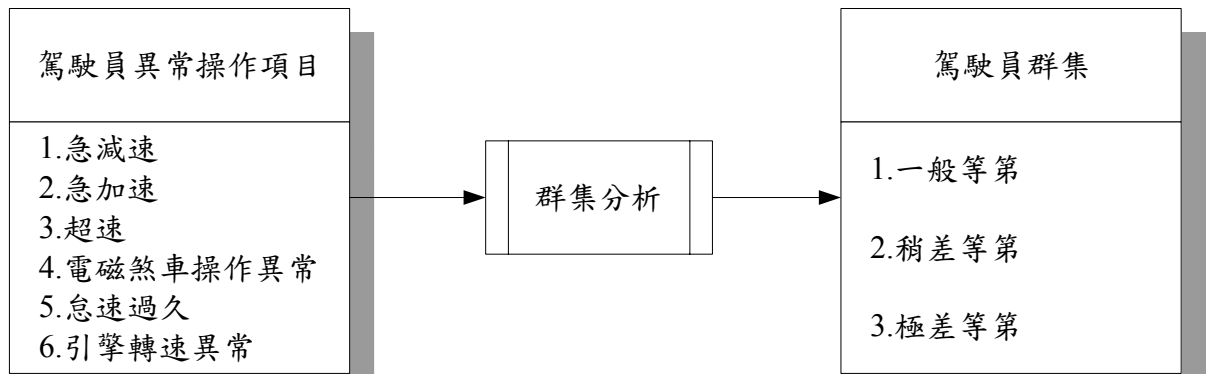


圖3-4 群集分析架構圖

3.4 判別分析

(一) 判別分析之意義

判別分析是一種判別和分類的技術，屬於多變量分析中用於判別樣本所屬類型的一種統計分析方法。係用以檢定 p 個群體在 i 個預測變項所構成的重心方面之差異是否達到顯著水準，如果達到顯著水準，還可進一步探討這 p 個群體之差異是在哪些向度方面。進行判別分析之過程，可求得 j 組 ($j < i$) 的判別函數係數 (Discriminant Function Coefficients)，將 i 個預測變項予以加權後而後得到 j 組的判別函數。其所要解決的問題是在一些已知研究對象用某分析方法分成若干群集的情況下，此時取數個變量的組合，求得適合的判別函數時，即可決定該群體是屬於哪個群集。

判別分析的基本模型就是判別函數，亦即判別變數的線性函數關係，其數學式如下：

$$D_x = d_0 + d_1 X_1 + d_2 X_2 + \cdots + d_i X_i + \cdots + d_k X_k$$

其中， D_x 為判別函數值； X_i 為各判別變數； d_i 為相對應之判別係數。

判別函數值 (Discriminant Function Value) D_x 又常稱為判別分數 (Discriminant Score)。而 d_i 代表判別係數或權重 (Discriminant Coefficient or

Weight)，表示各判別變數對於判別函數值的影響，其中 d_0 為常數項。

上述判別函數之方程式為線性判別函數。因此判別函數係由兩個以上的群體中抽取多種變量之資料，以建立判別關係式，用以判定未知新樣本的歸類。例如有數個群體，取數個變量組，作為適當的判別函數時，即可辨別該群體之歸類。根據這些判別函數，可以看出各變量對群體歸屬的影響強度，當取一新樣本時，求其判別函數值，依其值之大小，即可以判別該樣本群體之歸類。

(二) 理論基礎

判別函數是以 D_x 作為各指標 (X_i) 之一次線性函數。在各群體間， D_x 值為最容易分離時，求取所選定指標之權數 (d_i) 之方法。在數學上，即為求 D_x 軸之群間變動與群內變動之比值最大時的 d_i 值。其基本假設如下：

1. 一個判別變數不能為其他判別變數之線性組合。
2. 各類組間的組內變異數—共變異數矩陣應為相等。
3. 各組判別變數之間具有多變量常態分配。

● 兩群體判別分析

設有兩母體群 G_1 與 G_2 ，其最佳判別式為：

$$D_x = d_0 + d_1 X_1 + d_2 X_2 + \cdots + d_i X_i + \cdots + d_p X_p$$

而 G_1 與 G_2 群之平均值即為其判別得點 $\bar{D}^{(1)}$ 與 $\bar{D}^{(2)}$ ，而兩群體之判別得點平均值為

$\bar{D} = 1/2(\bar{D}^{(1)} + \bar{D}^{(2)})$ ，即為其判別點。因此，樣本 \underline{X}_0 之判別基準為：

(1) $\bar{D}^{(1)} > \bar{D}^{(2)}$ 時

若 $D_0 > \bar{D}$ ，則 \underline{X}_0 屬於 G_1 群

若 $D_0 < \bar{D}$ ，則 \underline{X}_0 屬於 G_2 群

(2) $\bar{D}^{(1)} < \bar{D}^{(2)}$ 時

若 $D_0 > \bar{D}$ ，則 \underline{X}_0 屬於 G_2 群

若 $D_0 < \bar{D}$ ，則 \underline{X}_0 屬於 G_1 群

若兩群體間之平均值相差很小時，由判別函數所求得之誤判機率將增大，則所求得之結果毫無意義。因此應對於兩群體平均值進行顯著性檢定，虛無假設為兩母群體之平均值相等，對立假設則為兩母群體之平均值不相等，即為：

$$H_0: \mu^{(1)} = \mu^{(2)}$$

$$H_1: \mu^{(1)} \neq \mu^{(2)}$$

並假設為常態分配時，則使用 F 分配來檢定兩群體是否有顯著的差異，即為自由度 $(P, n_1 + n_2 - P - 1)$ 之 F 分配：

$$F = \frac{n_1 + n_2 - P - 1}{P} \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 2)} D_{12}^2$$

亦可使用Hotelling之 t^2 檢定，其原理如同單變量兩群體之 t 檢定，且與 F 分配有關，如下所示：

$$t^2 = \frac{n_1 n_2}{(n_1 + n_2)} D_{12}^2 = F$$

● 三個群體以上之判別分析

判別分析共同的特徵為事前須就分析的對象加以分解，分析的目的在于客觀的分解，使分解後的變量更具代表性。三群以上之分析通常有兩種方法，分別為

(1) 作成一個判別式來判別。

(2) 就每兩群作成判別式來判別，若有 S 群，其所需之判別式共有 $S(S-1)/2$ 個。在判別效率上，第1種較第2種方法差，但第1種方法的優點能為用一個判別式判定群體之歸屬，第二個方法的缺點為判別式過多。

在判別檢定的時候，分別就每兩個群體進行分析，如在三個群體分析時，依照 G_1 、 G_2 ， G_2 、 G_3 以及 G_1 、 G_3 分別求得其判別式 D_{12} 、 D_{23} 以及 D_{13} ，若三群體之平均值大小分別為 $G_1 < G_2 < G_3$ 時，則下式成立：

$$D_{13} = D_{12} + D_{23}$$

然後求得各群之判別得點平均值與每兩群之判別點，若樣本 \underline{X}_0 之判別得點值為 D_0 時，其判別基準如下：

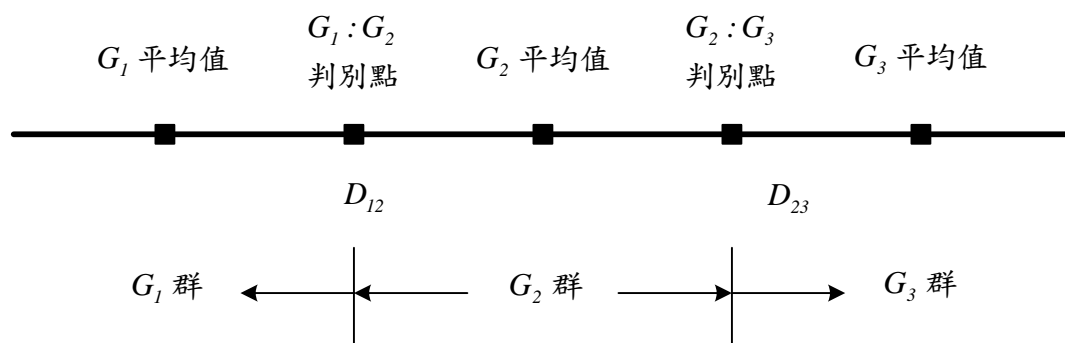


圖3-5 三群體判別基準

即為 $D_0 < D_{12}$ ，則 \underline{X}_0 屬於 G_1 群。

$D_{12} < D_0 < D_{23}$ ，則 \underline{X}_0 屬於 G_2 群。

$D_{23} < D_0$ ，則 \underline{X}_0 屬於 G_3 群。

● Wilks之 Λ 檢定

當一個變量為 t 檢定時，多個變量則為 Hotelling 之 t^2 檢定；同樣地，當一個變量為 F 檢定時，多個變量則為 Wilks 之 Λ 檢定，下列為 Λ 統計量之意義。

$$\Lambda = \frac{S_W}{S_T} = \frac{S_T - S_B}{S_T}$$

其中 S 為群體數， S_B 為群間變動， S_W 為群內變動， S_T 為總變動（ $S_T = S_B + S_W$ ）。群中變動之有無，以 Λ 值比 1 小來決定，值越小越有判別力，及分類中心間有相當大的分離，並且相對於類組內部之分散程度而言非常明顯。當變量 $P \geq 2$ 時，上式之 Λ 統計量可以寫成為

$$\Lambda = \frac{|S_W|}{|S_T|} = \frac{|S_W|}{|S_B + S_W|} = \frac{|W|}{|W + B|} = \frac{1}{\prod_{l=1}^m (1 + \theta_l)}$$

其虛無假設與對立假設為

H_0 ：群間無差異。

H_1 ：群間有差異。

而 F 統計量為， $F = \frac{1 - \Lambda^{\frac{1}{s-1}}}{\Lambda^{\frac{1}{s-1}}}$

（三）費雪判別函數

線性判別可分為線性判別函數或是費雪判別函數，費雪判別函數又稱為典型判別函數，其是尋找判別變數的線性組合之最佳權重，使組內變異數比值最大，即 F 值最大。費雪考慮將兩個母體的多變量資料經由線性組合之方式，轉換成單變量的資料，而且希望在這個單變量的資料所對應的兩個母體之中心點相距最遠，也就是在這個線性組合的方向能將兩個群體的資料點區分的最清楚。

當導出費雪判別函數（典型變量）後，就可以得到每個判別變數的典型權重（費雪判別係數），標準化的費雪判別係數值越大，代表此判別變數對群體的判別能力越強。而費雪判別係數可以用來直接進行一個觀察值的判別，也就是說，可以將觀察值之各變數值，代入費雪判別係數計算，即可得到判別之結果。

基於上述之理論方法，本研究欲使用判別分析之基本假設、目的與流程說明如下：

1. 基本假設

使用群集分析之分群結果，並將原本作為分群之各項變數，作為判別分析輸入做為判別模式中之變項，依據這些變項進行判別分析，並檢驗判別結果是否具顯著性。

2. 目的

使用判別分析構建判別模式，主要目的是為了能夠建立判別模式直接能夠將駕駛員判別屬於哪個群集（等第），對於實務應用上將可提高便利性。

3. 判別分析流程

將各判別變數輸入如急減速、急加速、電磁煞車操作異常、怠速過久以及引擎轉速異常等項目變數，透過判別分析得到判別函數與係數，並使用費雪判別係數作為各群集之線性判別模式。

3.5 小結

本研究主要運用資料採礦之相關方法與技術作為研究方法，以對於所蒐集到之駕駛員資料與數位式行車紀錄器資料進行資料採礦，首先運用多元迴歸分析法以構建駕駛員異常操作行為之關係模式，主要欲探討駕駛員之異常操作行為對於駕駛員行車安全方面：交通違規與肇事風險，以及行車油耗與保養維修費用之關係，以期能夠以多元迴歸分析之資料採礦方法，了解駕駛員各項異常操作情形會對於行車安全與車輛保修費用所產生之個別主要影響，且所得到之模式可運用於後續駕駛員管理之依據。再運用資料採礦之群集分析法將駕駛員資料進行分群，了解現有駕駛員之群集特性以及分布狀況，並依據各群之特性進行命名，與作為駕駛員評鑑與獎懲管理之依據，並使用判別分析，建立判別模式，可對於衛之群組織駕駛員直接進行分群，以增進管理之效率。

第四章 駕駛異常操作指標設定與資料蒐集

本研究欲以資料採礦方法探討公路客運駕駛員異常操作行為對於行車安全與油耗、保養維修費用的影響，故必須釐清駕駛正常操作與異常操作之標準，才能判定哪些駕駛操作行為是異常會影響行車安全與造成燃料浪費與機件耗損。故必須先定義出異常操作之項目以及門檻值，並依據前述項目與設定對於資料庫進行資料採礦，剔除不屬研究範圍之資料以縮減資料量，始能夠從龐大的資料庫中篩選出有意義且可解釋的資訊，以利後續資料採礦模式之建立。

4.1 數位式行車紀錄器駕駛異常操作監控指標與設定

本研究之研究對象為國內某客運公司，其於國道客運路線之 61 部車輛裝置數位式行車紀錄器，採用機種為 VDO FM200-PLUS 型數位式行車紀錄器，藉以監測車輛行駛資訊，如：GPS 車輛定位、行車速率、煞車、引擎轉速以及旅次行駛里程等，可進一步獲知駕駛操作行為。該款數位式行車紀錄器是屬於較為新型之機種，其可先設定駕駛操作監控項目與門檻值，並依照所設定之門檻值進行監控與紀錄，如此一來可簡化輸出資料之繁雜性並增加易讀性，資料輸出採標準資料庫格式，需配合後端資料庫管理系統以及其搭配之軟體。但對於駕駛操作行為之管理仍需做進一步的分析，故本研究參考張季倫(2002)提出之構想建立架構，分成行車安全、燃油消耗、機件耗損及行車品質等四種駕駛異常操作監控類型作為數位式行車紀錄器監控記錄指標與門檻值設定之依據，各類指標定義、內容與門檻值設定說明如下。

4.1.1 行車安全方面

行車安全方面是為了預防與監控駕駛員是否有不當的駕駛行為而造成行車安全的風險，而進行監控項目與門檻值之設定，主要有違規超速與急加減速等兩大項指標，相關說明如下。

一、違規超速指標

- 定義：車輛行車速率超過設定速限。
- 內容：車輛行駛速率(Km/h)。

● 功能：

根據道路交通事故原因統計，超速為車輛肇事的前三大主因之一，且近幾年來呈現上升的趨勢，因此超速是造成肇事之主要原因。因此就行車安全層面而言，超速是最基本且最重要的指標之一。為了衡量駕駛是否有超速之行為，本研究以設定數位行車紀錄器門檻值之方式，紀錄駕駛超過設定超速門檻值之次數，即可判定駕駛之超速行為。

● 原理：

超速指標不論是在一般傳統類比式行車紀錄器或是數位式行車紀錄器都是一個相當基本之指標，判定駕駛是否超速。在過去傳統類比式行車記錄器中，以直接偵測紀錄速率的方式，並裝置蜂鳴器，當行駛速率超過設定速限時，即發出警告音。在數位式行車紀錄器中，行車速率資料更是基本紀錄之一，並可依據公司營運管理單位的需求，設定超速監控的門檻值，行駛速率一但超過設定的門檻值，即直接紀錄之，輸出可呈現一趟旅次超速之次數。的因此本研究直接利用數位行車紀錄器之速率原始資料，偵測行車速率，並以公部門提供之速限資料作為駕駛是否超速之門檻。

● 門檻值設定：

超速門檻值主要以政府機關提供之速限資料作為駕駛超速之門檻設定標準，依據交通部於92年4月15日核定之國道主要路段速限資料，高速公路各主要路段速限標準如表4-1。

表 4-1 國道主要區段速限標準

國道別	路段	設計速率 (公里/小時)	速限 (公里 /小時)
國道一號	基隆端—內湖交流道	100	100
	內湖交流道—五股交流道	120	
	汐止—五股高架段	120	
	五股交流道—桃園交流道	100	
	桃園交流道—湖口交流道	120	
	湖口交流道—鳳山溪橋	100	

	鳳山溪橋—苗栗交流道	120	
	苗栗交流道—泰安服務區	100	
	泰安服務區—楠梓交流道	120	
	楠梓交流道—高雄端	100	
國道三號	基隆端—汐止系統交流道	100	90
	汐止系統—中和交流道	隧道內 90 隧道外100	
	中和交流道—土城交流道	100	100
	土城交流道—龍井交流道	120	110
	快官交流道—九如交流道	120	
國道三甲	辛亥路端—新光路端	80	80
國道五號	南港端—石碇端	80	70
國道二號	機場端—機場系統交流道	90	90
	機場系統—鶯歌系統交流道	100	100
國道四號	清水端—中港系統交流道	90	90
	中港系統交流道—豐原端	90	
國道八號	台南端—南133 鄉道	80	80
	南133 鄉道—新化端	100	100
國道十號	左營端—鼎金系統交流道	80	80
	鼎金系統交流道—燕巢系統交流道	80	
	燕巢系統交流道—旗山端	90	90

資料來源：民國93年高速公路年報

雖然超速即會增加行車安全之風險，但若門檻值設定過於嚴格，一超過速限標準即記錄超速一次，將影響駕駛員的駕駛專注力，容易造成疲勞，反而對行車安全性不增反減，失去指標意義。因此考量駕駛員心理層面與政府機關違規取締之標準，以及車輛行駛路線為台北高雄線行駛於國道一號之範圍，以法定速限之5%作為緩衝區，並考量駕駛員因一時疏忽而不小心超速，超速紀錄僅限超過法

定速限7%且長達5秒以上者，超速每過5秒則加計超速一次，待車速返回速限以下為止。門檻值設定標準如下：

1. 速率等於 105km/h 時，蜂鳴器響 1 秒提醒駕駛。
2. 速率大於等於107km/h且持續5秒時，紀錄一次，超速每過5秒加計一次，依此類推。

二、急加減速指標

- 定義：駕駛員急踩油門或是急踩煞車。
- 內容：加減速度對時間之變化率，稱為衝度 $da(t)/dt$ 。
- 功能：

根據文獻，急加速與急減速是造成車輛追撞和碰撞的主因之一，因此偵測車輛行駛時急加減速的發生次數為指標的主要功能，為了衡量駕駛員是否有急加減速之行為，以加速度對時間的變化量做為判定依據。可偵測駕駛員是否有踩油門或踩煞車之行為，依據正常範圍之加減速度變化量設定該指標之門檻值，判定駕駛員是否有急加減速之行為產生。

- 原理：

當加減速行為產生時，車輛會因慣性作用而產生突向前傾或後頓之現象，觀察單位時間內加速度之變化量稱為衝度，若衝度為正值，則表示駕駛員正進行加速行為，反之，若衝度為負值，則表示駕駛員正產生減速行為，急加減速指標即利用此一特性，配合衝度值大小，衡量當駕駛加減速行為發生時，是否有過急的情形產生，同時亦能根據衝度值大小判斷當時急加減速之程度，如急煞車與緊急煞車。

- 門檻值設定：

依據張季倫（2002）安排受測者填答於測試路段急加減速感受程度問卷，感受程度共分大、中、小三級。將記錄衝度值及感受程度之問卷資料以簡單線性迴歸後，採模糊德非法分析，以各資料點衝度之幾何平均數代表共識，以求出共識區之衝度值，研究結果感受範圍之衝度門檻值如下表4-2。

表4-2 急加減速門檻值對照表

感受程度	小	中	大
衝度範圍(m/s^3)	1.16 ~ 1.62	1.62 ~ 2.06	> 2.06
	-1.16 ~ -1.62	-1.62 ~ -2.06	< -2.06

資料來源：張季倫（民90）、本研究整理

故本研究之數位行車紀錄器急加減速門檻值設定，分別參考上述依據將急加速與急減速之衝度分別依照車輛行駛速度之低速、中速與高速範圍設定三個等級，因為大客車低速之扭力表現較佳，故駕駛員於低速時若油門控制不當，易造成急加速之情況，故急加速之門檻值設定因不同速度區間而設定不同數值，低速較高速之衝度門檻值嚴格，以及另外設定緊急煞車判斷之門檻值，一但有發生超過設定門檻值，即依發生次數記錄之。相關監控項目依據客運公司管理需求與實際駕駛操作調整後門檻值設定如下：

1. 急加速（低速）：每秒加速度大於 2.7m/s^3 。
2. 急加速（中速）：車輛速度大於 40km/h 、小於 70km/h ，每秒加速度大於等於 1.38m/s^3 。
3. 急加速（高速）：車輛速度大於 70km/h ，每秒加速度大於等於 0.83m/s^3 。
4. 急煞車（輕度）：車輛速度小於 40km/h ，每秒減速度大於 1.38m/s^3 。
5. 急煞車（中度）：車輛速度大於 40km/h 、小於 70km/h ，每秒減速度大於 2.2m/s^3 。
6. 急煞車（重度）：車輛速度大於 70km/h ，每秒減速度大於 2.7m/s^3 。
7. 緊急煞車：每秒減速度大於 3.33m/s^3 。

4.1.2 燃油消耗方面

燃油消耗方面主要針對不當駕駛行為因而導致額外的油料損耗訂定監控項目與門檻值，主要監控之三項指標分別為異常轉速、急加減速以及怠速過久等三大項指標，相關說明如下。

一、異常轉速與檔位異常指標

- 定義：引擎在轉速超過合理或安全之轉速區間，以及不在檔位對應之速度區間如低檔位高轉速或高檔位低轉速下之運轉。

- 內容：引擎轉速、速率與檔位之對應關係。

- 功能：

根據文獻，引擎轉速過高或過低會造成燃油額外的消耗以及造成機件之磨損，因此為偵測造成引擎異常轉速之狀態，引擎轉速過高或是過低，以引擎轉速與速率的關係作為異常轉速指標主要之判定依據，檔位區間判斷作為輔助，並可以對駕駛員不當換檔行為提出建議。

- 原理：

當引擎處於在低檔位高轉速狀態下，轉速一定會超過正常範圍之上限；同理，在高檔位低轉速狀態下，轉速一定會低於正常範圍之下限。因此，若要監測引擎不正常轉速，主要需界定一般平路與上下坡之正常引擎轉速範圍，並配合對應之行駛速率推斷檔位之區間，可判斷檔位是否正確。

- 門檻值設定：

參考相關文獻所進行之客運車輛行駛速度區間與轉速區間之對應，分為一般路段與上坡路段，因為上坡路段車輛須以較大之動力爬坡，故與平路比較相同之速度區間與檔位下會有較高的轉速產生，必須另外界定。

本研究之車輛為變速箱為五檔之設計，而上述研究車輛為變速箱為八檔之設計，參考相關引擎動力傳動系統之文獻，始可得知，變速箱檔位較少之車輛相較於檔位較多之車輛，其單一檔位之速度區間會較大，尤其是越高之檔位兩者速度區間差異越大，較少檔位者每檔位會涵蓋較大之速度範圍，故依據相關文獻以及該款車輛日本三菱FUSO原廠建議之說明，本研究之五檔變速箱客運車輛速度、引擎轉速與檔位對應關係整理如下表4-3與表4-4。

表4-3 一般路段之速度、引擎轉速與檔位對應關係（FUSO車）

檔位	引擎轉速（RPM）	速率（Kph）
1	800～1200	5～11
2	1000～1600	11～26
3	1000～1600	17～41
4	1000～1600	38～59
5	1000～1600	52～99

資料來源：日本三菱FUSO、本研究整理

表4-4 上坡路段之速度、引擎轉速與檔位對應關係（FUSO車）

檔位	引擎轉速（RPM）	速率（Kph）
3	1000~1800	15~38
4	1000~1800	34~48
5	1000~1800	47~85

資料來源：日本三菱FUSO、本研究整理

依據日本三菱FUSO原廠的相關建議與國內某客運公司維修人員之說法，以及上述之速度、引擎轉速與檔位對應關係，配合數位式行車紀錄器之GPS定位模組，判斷車輛位於上下坡之路段，可避免監控紀錄之誤判。監控項目與門檻值之說明與設定如下：

1. 引擎轉速過高：引擎轉速大於 2000 rpm 且持續 2 秒，即記錄之。提醒引擎轉速大於 2000 且不在三義坡或林口下坡段、三重交流道或高雄九如交流道其中之一，即以蜂鳴器響 1 秒警示。
2. 異常轉速區間：引擎轉速小於1200 rpm或大於1850 rpm，且速度大於25 km/h，即記錄之。
3. 未使用一檔：一檔開關未啟動，且引擎轉速小於800，速度小於2 km/h，加速度大於等於1 km/h，即記錄之。
4. 異常4檔檔位區間：速度大於40 km/h小於90 km/h之區間、速度小於50 km/h減速度大於1 km/h、轉速小於1000 rpm、速度大於80 km/h以及轉速大於2000 rpm加速度大於1km/h，且不在三義坡或林口下坡段路段、三重交流道或高雄九如交流道其中之一，即視為4檔檔位異常紀錄之。
5. 異常5檔檔位區間：速度大於60 km/h小於80 km/h轉速小於1300 rpm、減速度大於1 km/h以及速度大於109 km/h加速度大於1 km/h轉速大於1900 rpm，且不在三義坡或林口下坡段路段、三重交流道或高雄九如交流道其中之一，即視為5檔檔位異常紀錄之。

二、急加減速指標

- 定義：駕駛員急踩油門或是急踩煞車。
- 內容：加減速度對時間之變化率，稱為衝度 $da(t)/dt$ 。
- 功能：

根據文獻，急加速和急減速是造成異常燃油消耗的主因之一，因此偵測車輛行駛時急加減速的發生次數為指標的主要功能，為了衡量駕駛員是否有急加減速之行為，以加速度對時間的變化量做為判定依據。可偵測駕駛員是否有踩油門或踩煞車之行為，依據正常範圍之加減速度變化量設定該指標之門檻值，判定駕駛員是否有急加減速之行為產生。

- 原理：與 4.1.1 急加減速指標相同。
- 門檻值設定：與 4.1.1 急加減速指標相同。

三、怠速過久指標

- 定義：駕駛員於同一地點停留超過一定時間且引擎為運轉狀態。
- 內容：GPS 定位與紀錄時間以及引擎運轉狀態。
- 功能：

在廠區或是固定地點怠速過久會造成燃油消耗浪費以及車輛引擎積碳等缺點，以及判斷司機是否仍在廠區未出車，即為此監控項目之功能。

- 原理：

由於此數位式行車紀錄器有 GPS 全球衛星定位模組，故可以設定車輛之定位，配合時間之紀錄，車輛若在特定地點或是某固定地點，以及引擎在怠速之運轉條件下，即可判斷駕駛員是否於特定地點或是某固定地點是將車輛發動怠速運轉過久沒有行駛車輛。

- 門檻值設定：

依據客運公司營運與管理之需求，分別設定廠區怠速過久與怠速過久之監控項目與門檻值，說明如下：

1. 廠區怠速過久：引擎轉速大於 0 rpm、速度小於 40 km/h、指定位置內（車廠及停車場）時間超過 40 分鐘，即記錄之。
2. 怠速過久：引擎轉速小於 800 rpm、轉速大於 0 rpm、速度小於 2 km/h、時間超過 40 分鐘，即記錄之。

4.1.3 機件磨損方面

機件磨損方面主要針對因不當駕駛行為可能導致額外的機件磨損進行監控

項目訂定與門檻值設定，分別有異常轉速、急加減速以及電磁煞車使用異常等三大項指標，相關說明如下。

一、異常轉速與檔位異常指標

- 定義：引擎在轉速超過合理或安全之轉速區間，以及不在檔位對應之速度區間如低檔位高轉速或高檔位低轉速下之運轉。
- 內容：引擎轉速、速率與檔位之對應關係。
- 功能：

根據文獻，引擎轉速過高或過低會造成燃油額外的消耗以及造成機件之磨損，因此為偵測造成引擎異常轉速之狀態，引擎轉速過高或是過低，以引擎轉速與速率的關係作為異常轉速指標主要之判定依據，檔位區間判斷作為輔助，並可以對駕駛員不當換檔行為提出建議。

- 原理：與 4.1.2 異常轉速指標相同。
- 門檻值設定：與 4.1.2 異常轉速指標相同。

二、急加減速指標

- 定義：駕駛員急踩油門或是急踩煞車。
- 內容：加減速度對時間之變化率，稱為衝度 $da(t)/dt$ 。
- 功能：

相關文獻指出，急加速和急減速會容易造成車輛引擎之損耗、加速引擎機油之劣化，而且煞車作用時間過久、煞車壓力過大、高速率下煞車以及煞車頻率過高，均易造成煞車鼓發熱變形，甚至造成煞車來令額外之磨損。因此偵測車輛行駛時急加減速的發生次數為指標的主要功能，為了衡量駕駛員是否有急加減速之行為，以加速度對時間的變化量做為判定依據。可偵測駕駛員是否有踩油門或踩煞車之行為，依據正常範圍之加減速度變化量設定該指標之門檻值，判定駕駛員是否有急加減速之行為產生。

- 原理：與 4.1.1 急加減速指標相同。
- 門檻值設定：與 4.1.1 急加減速指標相同。

三、電磁煞車操作異常指標

- 定義：駕駛員是否有電磁煞車使用不當或是過久。
- 內容：電磁煞車啟動時機與時間。
- 功能：

電磁煞車之使用時機主要是在使用腳煞車之前，先以電磁煞車將速度盡可能降至最低，主要的目的是輔助車輛的氣壓鼓式煞車或是碟式煞車系統，降低煞車來令的磨損以及煞車時所產生的高溫，但是電磁煞車使用不當或是過久，有可能會造成電磁線圈燒毀。因此電磁煞車使用異常指標的主要功能在偵測駕駛員使用電磁煞車之行為，以判別駕駛是否有使用電磁煞車不當或是過久而導致電磁煞車燒毀之不良行為。

- 原理：

電磁煞車之作動原理是在傳動軸的位置加裝一電磁線圈，一般行駛時並無任何動作，而駕駛員使用時會按下開關，使電磁線圈產生一逆向磁場，使傳動軸會產生反轉的作用力，而達到減緩速度，類似煞車的效果。要判斷電磁煞車是否作動，由電磁煞車開關是否啟動即可得知。

- 門檻值設定：

參考日本三菱FUSO原廠車輛的使用建議與國內某客運公司維修人員之說明，將電磁煞車使用異常指標的監控項目與門檻值設定如下：

1. 電磁煞車使用過久：駕駛員拉下電磁煞車、轉速大於 600 且持續超過 12 秒，即記錄之。
2. 電磁煞車使用不當：駕駛員拉下電磁煞車、轉速大於 600、每秒加速度大於等於 1 km/h 時，即記錄之。

4.1.4 行車品質方面

一、急加減速指標

- 定義：駕駛員急踩油門或是急踩煞車。
- 內容：加減速度對時間之變化率，稱為衝度 $da(t)/dt$ 。
- 功能：

當加減速行為產生時，車輛會因慣性作用而產生突向前傾或後頓之現象，而急加速和急減速會使這個現象更為明顯，易會造成乘客的不舒適感。因此偵測車輛行駛時急加減速的發生次數為指標的主要功能，為了衡量駕駛員是否有急加減速之行為，以加速度對時間的變化量做為判定依據。可偵測駕駛員是否有踩油門或踩煞車之行為，依據正常範圍之加減速度變化量設定該指標之門檻值，判定駕駛員是否有急加減速之行為產生。

- 原理：

當駕駛員有急加減速行為時，車輛會因慣性作用而產生突向前傾或後頓之現象，車內的乘客也會隨車體做前後擺盪。因此可以前述之衝度做為指標，觀察單位時間內加速度之變化量，可有效判斷駕駛是否有急踩油門或煞車之行為導致車體突向前傾或後頓，造成乘客前後俯仰而感到不適或受傷。

- 門檻值設定：與 4.1.1 急加減速指標相同。

二、車燈使用異常指標

- 定義與內容： 駕駛員是否有閃遠光燈。

- 功能與原理：

為了要了解駕駛員是否有不當的使用遠光燈，因為不當的使用遠光燈會使得對向來車的駕駛人的視線受到干擾，甚至影響他人的行車安全，或是不當的對前車閃遠光燈，亦會對於前車駕駛人造成干擾，甚至因為不當的使用遠光燈而造成其他駕駛人的不愉快感，而降低了客運公司行車的品質，甚至影響形象。

- 門檻值設定：啟動遠光燈開關即記錄。

將上述各項指標與監控項目彙整如下表4-5。

表4-5 各項指標與監控項目

指標類型	指標名稱	監控項目
行車安全	違規超速指標	違規超速
	急加減速指標	急加速 急煞車 緊急煞車
燃油消耗	異常轉速與檔位異常指標	引擎轉速過高 異常轉速區間
		未使用一檔 4檔檔位區間異常 5檔檔位區間異常
	急加減速指標	急加速 急煞車 緊急煞車
	怠速過久指標	廠區怠速過久 一般怠速過久
機件磨損	異常轉速與檔位異常指標	引擎轉速過高 異常轉速區間
		未使用一檔 4檔檔位區間異常 5檔檔位區間異常
	急加減速指標	急加速 急煞車 緊急煞車
	電磁煞車操作異常指標	電磁煞車使用不當 電磁煞車使用過久
行車品質	急加減速指標	急加速 急煞車 緊急煞車
	車燈使用異常指標	閃遠光燈

4.2 數位式行車紀錄器駕駛異常操作資料蒐集

本研究蒐集國內某客運公司61部大客車上數位式行車紀錄器之資料，藉由設定該公司車輛之數位式行車紀錄器的相關監控項目與門檻值，可以偵測駕駛員操控車輛的行為，並判斷是否有異常操作之情形。數位式行車紀錄器監控原理就是由車輛之引擎或變速箱將速率和引擎轉速等類比訊號傳導至行車紀錄器，予以數位化後加以儲存，同時亦可將車輛上各種裝置的運作狀況以數位化之資訊儲存於記錄器之記憶體中。管理者可以應用傳輸介面將紀錄器資料下載至電腦中進行深入管理分析。數位式行車紀錄器所監控與記錄的資料，除了較機械式紀錄器精確外，也因為數位化與電子化的緣故，可監控管理之項目較傳統機械式的來的多，能夠蒐集較多的資料，對於經營管理者方面亦因數位資料可迅速自動轉換成簡明之資訊，大幅提高可讀性且節省人力與時間。

在該客運公司之61部大客車中，每部車上裝設有數位式行車紀錄器主機，駕駛配備有數位式行車紀錄器車輛的駕駛員，皆領有一支專屬藍色隨身碟（內建128bytes記憶體之儲存裝置），為啟動數位式行車紀錄器駕駛員身分存取認證、日期時間輸入之功能，駕駛員之藍色隨身碟內建有特定編號。駕駛員於出勤前須向管理組長領取專屬隨身碟，插入車上主機並啟動，發動車輛後紀錄器便開始運作，於收班後取出隨身碟交由管理人員。

該款數位式行車紀錄器主機之資料下載，是採用無線通訊之方式，是利用其GSM模組進行資料自動下載，管理者每天於電信業者具有優惠之時段，將一天中數位式行車紀錄器之資料下載至電腦之資料庫中，另外該數位式行車紀錄器還配備有綠色隨身碟則為下載車輛旅程資料之用，其記憶體大小為96~256 Kbytes。資料可自動下載且採無線傳輸相較於需使用卡匣或是其他儲存媒體手動下載之方式，對於公司的管理方面確實是有效的減少下載資料之人力以及減低儲存媒體損壞或遺失而造成記錄資料之缺少。

● 資料範圍：

本研究蒐集之數位式行車紀錄器資料，主要包含有車輛資料、駕駛員資料以及駕駛異常操作之監控項目與門檻值設定等資料，駕駛異常監控項目與門檻值設定如4.1節中所敘述。以上這些資料的原始資料格式，為該數位式行車紀錄器下載至管理電腦後所建立之關聯式資料庫結構，為標準SQL資料庫之格式，故本研

究必須對於這些資料再進行資料格式的轉換與篩選整理，以利後續與其他資料之整合與分析之用。

- 資料期間：

本研究配合該客運公司開始於車輛上裝設數位式行車紀錄器之時間，蒐集數位式行車紀錄器之資料於民國93年6月1日至民國94年12月31日。

4.3 駕駛員肇事與交通違規資料蒐集

根據國內外的調查顯示，絕大部分的肇事事故都是駕駛人操作不當（人為因素）所引起的，而違規即是不良好之駕駛習慣或是態度所造成的。故本研究除了數位式行車紀錄器之監控資料外，另外蒐集該客運公司駕駛員之肇事與交通違規之資料，肇事與違規資料之蒐集時間為民國93年6月1日至民國94年12月31日，範圍與定義說明如下。

4.3.1 駕駛員肇事資料

- 肇事定義：

本研究將肇事之範圍界定在該客運公司之駕駛員於車輛營運期間，駕駛客運車輛與其他非該公司車輛或是其他人員發生事故，造成車輛損壞或是人員傷亡之情形。而肇事原因調查之結果歸咎於駕駛員責任者。

- 資料範圍：

蒐集該客運公司駕駛員之肇事資料，篩選出資料之範圍為駕駛數位式行車紀錄器車輛的駕駛員，並發生肇事之事件，即列入資料之範圍。

4.3.2 駕駛員交通違規資料

- 交通違規定義：

本研究將違規之範圍界定在該客運公司之駕駛員於車輛營運期間，駕駛客運車輛於駕駛行進中所產生之違反交通法令行為而受到主管單位舉發開罰之情形，且該項行為屬駕駛員本身自發之行為，非公司允許或下令之行為，且該項違

規罰款由駕駛員自付者。

- 資料範圍：

蒐集該客運公司駕駛員之違規資料，篩選出資料之範圍為駕駛數位式行車紀錄器車輛之駕駛員，其違規並被舉發開罰的罰單資料，即列入資料之範圍。

4.4 行車油耗與保養維修費用資料蒐集

根據交通部運研所對國內汽車客運業者所提供之成本資料作歸納整理（民89），指出客運業營運成本中管理費用佔41%，業務費用佔營運總成本的28%，行車費用佔17%，保養維修費用佔14%，故行車與保修費用約為營運成本的三分之一。而駕駛行為因素亦為影響行車油耗與保養維修費用之重要因素之一，故本研究亦對於該客運公司營運車輛之油耗與保養維修費用進行資料蒐集，資料蒐集的期間為民國93年6月1日至民國94年12月31日。相關資料之蒐集時間與內容定義說明如下。

4.4.1 車輛油耗費用

- 油耗費用計算方式：

該客運公司車輛加油有一定的程序，在每輛車上皆備有一份駕駛員加油登錄表，其用途為當駕駛員加油時，必須登記車號、加油日期、時間、里程錶里程數及加油量。每月月底將登錄表交由會計部門進行電腦建檔統計，由此可計算出該車輛於當月之行駛總里程及總加油量，以此估算當月燃料費用。在保養場內由委外石油公司提供燃料，於保養場內設置有加油機，而加油機亦附有一份加油登錄表，駕駛人需填寫加油時間及加油量，以統計該設備之出油量。車輛加油流程如下圖4-1所示。

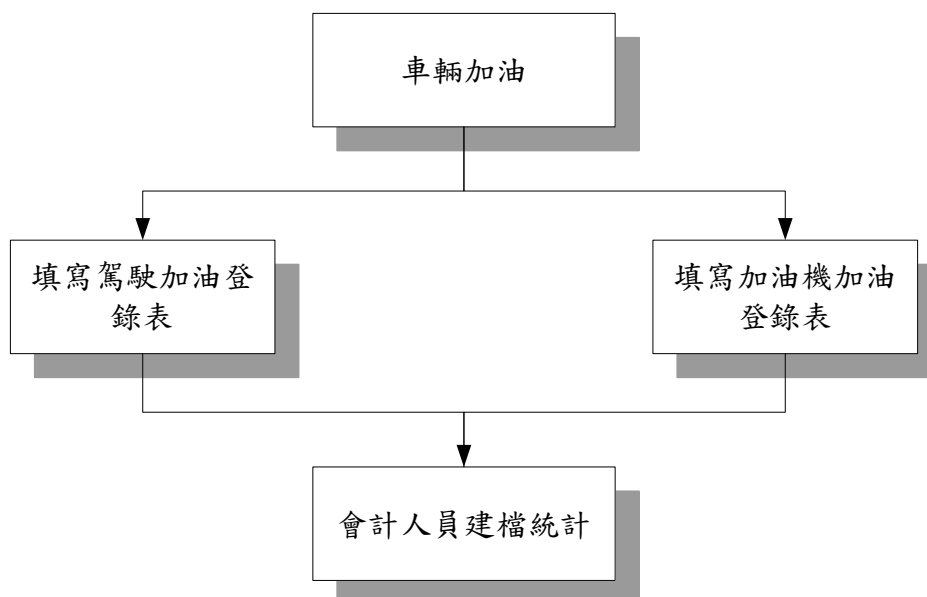


圖4-1 車輛加油流程圖

- 資料範圍：

車輛油耗資料為資訊管理部門所統計之資料，油耗資料主要為記錄每個月每部車的總行駛里程與總加油量，每公升的柴油價格採石油公司與客運公司的簽約價格，故可得到每部車每千公里的油耗費用。資料之範圍篩選裝設數位行車紀錄器車輛之每月油耗費用。下表4-6為每月車輛油耗費用資料之範例格式。

表4-6 每月車輛油耗費用資料範例格式（車號為隨機假設）

車號	初里程數	末里程數	行駛里程	總加油量	平均 km/L	總金額
AA-123	899541	931017	31476	8420.33	3.738	138604
AB-321	900347	935450	35103	11176.76	3.141	184422
...
AZ-567	884318	921155	36837	10044.77	3.667	165747

4.4.2 車輛保養維修費用

- 保養維修費用定義與計算方式：

本研究探討的主題之一是駕駛員異常操作行為所影響之行車油耗與保養維修費用，故僅就燃料、附屬油品、輪胎、修車材料等進行討論。至於其他項功能

類別雖歸類為行車或保養維修類，如車輛折舊、行車附支、通行費等，因與駕駛操作行為無直接關聯，故不列入考量。表4-7為交通部訂定之18項公路客運營運成本及功能類別。

表4-7 公路客運營運成本及功能類別

成本項目	功能類別
1. 燃料	行車
2. 附屬油品	行車
3. 輪胎	行車
4. 車輛折舊	行車
5. 行車人員薪資	行車
6. 行車附支	行車
7. 修車材料	保養維修
8. 修車員工薪資	保養維修
9. 修車附支	保養維修
10. 業務員工薪資	業務
11. 業務費用	業務
12. 各項設備折舊	保養維修、業務、管理
13. 管理員工薪資	管理
14. 管理費用	管理
15. 稅捐費用	行車、保養維修、業務、管理
16. 站場租金	保養維修、業務
17. 通行費	行車
18. 財務費用	營業外

註：功能類別係依各成本項目所對應之會計科目而異。

資料來源：交通部（民86），「汽車客運業統一會計科目」

1. 附屬油料類：

附屬油料主要有機油、差速器油、變速箱油、煞車油與輪軸黃油等，都是按車輛行駛里程更換。附屬油料以及相關機件的更換都是具有固定的時間進行，以防止油品劣化而導致車輛零件受損。會計人員彙總附屬油品領料單據，並輸入電

腦，作為月結費用之依據，以及製成每月報表。

2. 輪胎類：

由於該客運公司之大客車輪胎配置為前輪左右各一，而後輪則左右各二。依照輪胎之換裝作業順序，車輛通常會將新胎置於前輪，約行駛三個月後，將其卸下置於後輪外側，再於三個月後換裝於內側，這是因為在各個位置輪胎磨耗的部分略有不同，調換位置可避免固定某部位過度摩擦而破損。在每次裝卸時，除了記錄裝卸時間及行駛里程數，對於輪胎的胎紋及胎溝，維修人員需一併量測紀錄。若是輪胎吃胎、破損或是胎紋過淺，則須報請會計進行補胎或換胎作業。但由於該客運公司之輪胎更換是外包給輪胎公司，並以一年固定之金額總包所有客運車輛之輪胎更換維修作業，故本研究不將輪胎費用列入計算。

3. 修車材料類

該客運公司車輛簡易的保養維修作業多於自有之保養場進行。車輛需作定期的保養檢修，按照行駛里程對零件做不同程度的清理保養或更換，主要分為四種等級，分別為一級、二級、三級以及輪軸保養。此外，駕駛員於發車前或收班後若查覺車輛行駛不順或機件故障，須通報維修人員並填寫維修單。若要更換零件，則須報請會計開立領料單據至倉庫領取所需零件。其相關零件種類、價格及數量全部輸入電腦建檔，會計開立領料單時，便是從資料庫中進行零件及數量選取，一方面方便統計每月保養維修費用，另一方面為有效控管零件的數量。保養維修的零件項目相當的多主要有煞車來令片、空氣濾、機油濾、柴油濾、乾燥劑修包、排檔桿橡皮、扭力桿套墊、水箱水管、剎車除污劑、黃油、水精、油封、噴油嘴以及其他消耗零件等。

● 資料範圍：

保養維修費用包含附屬油料與保養維修零件等費用，資料之範圍為裝設數位式行車紀錄器車輛之資料，為資訊管理部門之每月統計報表。

第五章 駕駛員異常操作行為關係模式構建

經由前述章節確定駕駛異常操作監控指標項目與設定後，即可確保數位式行車紀錄器之駕駛異常操作行為資料紀錄之可靠性，故即可對於所蒐集之數位式行車紀錄器資料進行資料採礦，於本章節中，將以資料採礦技術中之多元迴歸分析法，對於數位式行車紀錄器所紀錄之駕駛異常操作資訊進行探討與發掘，以期能夠釐清駕駛員異常操作行為對於行車安全與行車油耗與保修費用之間的影響項目與關係。

5.1 資料庫建立與分析

將前述所蒐集與經過前置處理之資料整理完成後，為了後續研究分析之需要，將駕駛員異常操作資料、肇事、交通違規、車輛油耗與相關保養維修費用整理輸入構建成為一個新的駕駛員異常操作管理資料庫。

5.1.1 資料格式

為了後續分析研究的進行，本研究再進行資料採礦技術之運用前，首要工作即為將原始數位式行車紀錄器中的駕駛員異常操作資訊，進行資料篩選與清潔之動作，將有遺漏或是錯誤之資料，予以清除，並且檢查資料之正確性，以確保資料之可信度，且為能夠提供後續資料採礦技術運用後能夠得到有用之資訊，故在進行完成前述作業後，將篩選與整理完成之資料，建立一個新的駕駛員異常操作管理資料庫，並將各項異常操作資料予以標準化，將駕駛員各項異常次數除以千公里，以便於分析與查詢，主要格式如表5-1，所顯示之資料為部分駕駛員資料。

表5-1 駕駛員異常操作管理資料庫（部分資料）

駕駛員 編號	里程(千 Km)	肇事	違規	急減速	急加速	超速	閃遠燈	電磁煞車 操作異常	怠速 過久	檔位 異常	引擎轉 速異常
D001	262.8	0	0	9.52	4.20	0.71	3.69	1.43	0.60	157.53	4.51
D002	280.685	0.0036	0	2.24	1.69	0.05	0.86	0.06	0.53	17.95	1.40
D003	24.455	0	0	0.78	0.70	0.00	2.86	0.00	0.00	8.79	1.14
D004	57.67	0	0	10.40	10.28	0.55	18.28	3.40	2.25	168.35	11.79
D005	106.58	0	0	4.03	2.13	0.02	4.72	0.18	1.14	26.82	4.32
D006	101.105	0	0	2.69	1.65	0.00	2.13	0.06	1.12	22.11	6.13
D007	80.665	0	0	4.21	2.78	0.27	6.91	0.21	1.95	55.49	9.58
D008	98.915	0	0	5.22	3.61	0.00	1.90	0.20	1.07	25.82	6.01
D009	108.77	0	0	7.37	2.31	0.00	0.91	0.11	1.55	38.55	5.03
D010	278.13	0.0036	0.0108	2.36	1.09	0.00	6.66	1.01	0.55	13.59	2.17
D011	60.59	0	0.0165	8.30	3.55	0.15	16.95	2.23	2.06	45.59	10.56
D012	14.965	0	0	7.35	2.54	0.00	26.33	0.40	0.74	43.23	7.48
D013	257.69	0	0	4.78	3.31	10.24	1.76	0.24	0.86	38.71	7.00
D014	108.04	0	0	3.76	1.48	0.00	2.26	4.19	0.83	17.34	4.30
D015	114.61	0	0	1.20	0.58	0.00	2.32	0.05	0.25	17.05	1.68

註：單位為次/千公里

5.1.2 基本資料分析

本研究除了蒐集數位式行車紀錄器之資料外，另外蒐集了該客運公司駕駛員以及車輛之相關資料，主要包括駕駛員之交通違規與肇事紀錄、客運車輛之油耗與保養維修費用等，有了上述之資料並且進行資料之前置處理後，對於資料進行基本之統計分析，以了解相關之趨勢或是問題所在。

● 交通違規與肇事

（一）交通違規

在駕駛員交通違規部份，經過資料之前置處理後，將違規停車等非行駛中之交通違規行為罰單資料剔除，故最後得到共87筆交通違規紀錄，各項統計整理如圖

5-1。其中主要之項目有未依車道行駛45筆、超速18筆、行駛路肩8筆、未保持安全車距5筆、闖紅燈5筆以及未遵守交通標線指示4筆，其中以未依車道行駛所佔比例最高將近52%，其次是超速佔21%。

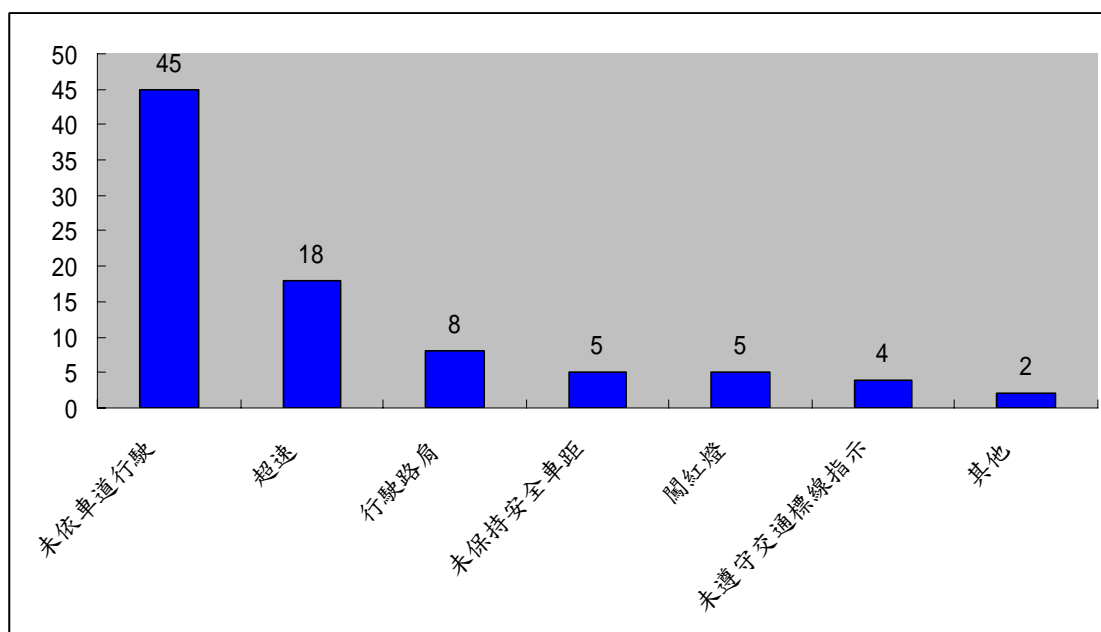


圖5-1 交通違規項目分析圖

(二) 肇事

駕駛員肇事部份，經過資料前置處理後，將該客運公司自家車輛於停車場擦撞等事故資料剔除，篩選之資料主要為與外界他人或車輛發生之事故，各項統計資料整理如圖5-2。主要包含一般擦撞35筆、追撞23筆、撞到路人2筆以及車輛翻覆1筆，總共為61筆駕駛員肇事資料，主要為擦撞與追撞之事故。

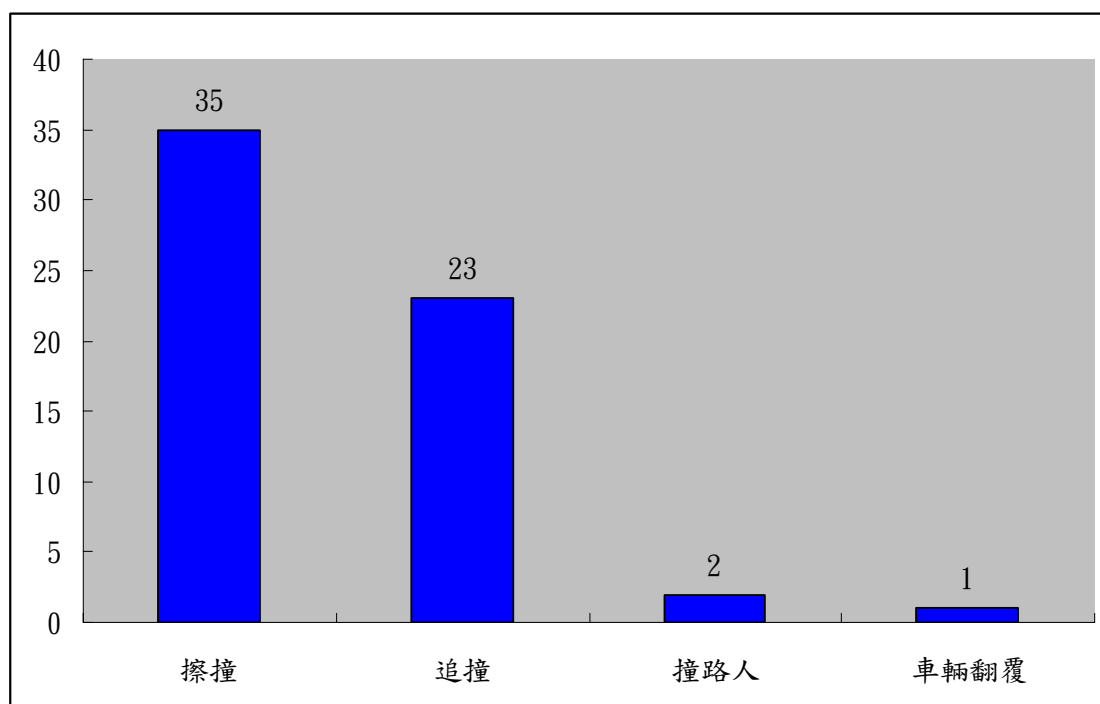


圖5-2 肇事原因分析圖

● 油耗與保養維修費用

本研究之對象主要為該客運公司61輛裝設有數位式行車紀錄器之大客車，在行車油耗與保養維修費用部份，進行燃料、附屬油料以及修車材料費用之資料蒐集。配合數位式行車紀錄器紀錄資料蒐集之時間。將所蒐集之燃料費、附屬油料費用以及修車材料費用資料進行初步統計，以燃料費所佔比例最高，其次為修車材料費，最後為定期更換之附屬油料費整理如表5-2。

表5-2 各項費用初步統計值

範圍 \ 項目	燃料費 (元/千公里)	附屬油料費 (元/千公里)	修車材料費 (元/千公里)
最大值	5453.209	196.068	1359.755
最小值	4652.088	128.349	437.122
平均值	5118.895	169.139	772.640
標準差	176.879	13.129	165.492
百分比	84.461%	2.791 %	12.748 %

● 駕駛員異常操作情形

本研究經由前述之駕駛員異常操作管理資料庫構建後，初步分析駕駛員之各項異常操作資料，各項目為急減速、急加速、超速、閃遠燈、電磁煞車操作異常、怠速過久、檔位異常以及引擎轉速異常等，各駕駛員異常操作項目之資料，均已經過將原始異常次數除以駕駛員行駛里程之資料標準化程序，以便於比較分析與後續迴歸模式構建。各項駕駛員異常操作資料整理如表5-3與表5-4。

表 5-3 駕駛員異常操作資料統計

範圍 \ 項目	急減速 (次/千公里)	急加速 (次/千公里)	超速 (次/千公里)	閃遠燈 (次/千公里)
最大值	286.96	31.17	52.27	86.17
最小值	0.76	0.39	0.00	0.00
平均值	13.09	4.26	3.25	6.26
標準差	31.98	4.62	7.95	9.23
百分比	10.74	3.50	2.67	5.14

表 5-4 駕駛員異常操作資料統計(續)

範圍 \ 項目	電磁煞車操作異常 (次/千公里)	怠速過久 (次/千公里)	檔位異常 (次/千公里)	引擎轉速異常 (次/千公里)
最大值	12.73	11.92	1262.97	136.58
最小值	0.00	0.00	4.46	0.86
平均值	1.35	0.97	81.01	11.62
標準差	2.16	1.18	147.99	17.47
百分比	1.11	0.79	66.50	9.54

5.2 肇事與交通違規風險迴歸模式構建

由於行車安全對於客運公司影響甚大，且亦為客運公司首重目標，故在駕駛員管理方面，必須先從行車安全方向進行，而駕駛員肇事對於行車安全有重大影響，交通違規亦為駕駛員不佳之駕駛行為或態度所造成，不但駕駛員受罰會增加罰單費

用之負擔，亦會影響公司之企業形象，更有可能影響行車安全，故本研究著手探討對於駕駛員發生肇事或交通違規與數位式行車紀錄器中監控之異常操作行為是否有所關聯，並藉由多元迴歸模式之構建，尋求其中有關係與影響之變數。其中肇事迴歸模式構建樣本數為43筆駕駛員資料，交通違規迴歸模式構建樣本數為52筆駕駛員資料。

5.2.1 變數選取

將所有駕駛異常操作項目：急加速、急減速、緊急煞車、超速、電磁煞車操作異常、怠速過久、檔位異常以及引擎轉速異常等項目，作為迴歸模式初步輸入之預測變數。但在迴歸模式中，最容易發生預測變數間具有高度相關，即複共線性(Multicollinear)問題的產生，如此將會使迴歸模式出現不合理的現象，如相關係數正負符號不合或是估計值不穩定等問題，將嚴重影響迴歸結果。

在複共線性檢定中，可由變異數膨脹因素(Variance Inflation Factors, VIF)、容忍值(Tolerance)、特徵值(Eigenvalue)以及條件指標(Condition Index)等方面進行判斷。若變數間具有複共線性問題，其變異數膨脹因素的極大值將大於 10，允差的極小值將小於 0.1，特徵值的極小值將小於 0.01 或條件指標的極大值大於 30，表 5-5 為複共線性判斷之準則。而複共線性問題為程度上嚴重與否，通常多以變異數膨脹因素(VIF)或是容忍值作為判定參考，變異數膨脹因子與容忍值之關係說明如下：令 R_i^2 為預測變數 X_i 對於其他變數 $X_1, X_2, \dots, X_{i-1}, \dots, X_j$ 之複相關係數，故 X_i 之變異數膨脹因子與容忍值之定義如下：

$$\text{容忍值：} Tolerance = 1 - R_i^2 \quad (5.1)$$

$$\text{變異數膨脹因子：} VIF = \frac{1}{Tolerance} = \frac{1}{1 - R_i^2} \quad (5.2)$$

表 5-5 複共線性判斷準則

變異數膨脹因素 (VIF)	> 10
容忍值 (Tolerance)	< 0.1
特徵值(Eigenvalue)	< 0.01
條件指標 (C.I.)	> 30

故本研究即對於急減速、緊急煞車、急加速以及超速等四項影響行車安全之變數，進行複共線性檢定。檢定結果如表5-6與表5-7，結果可知，緊急煞車與急減速具有高度的複共線性，由其原理可知緊急煞車亦為減速行為的一種，故會存在高度的相關性。

表5-6 第一次複共線性檢定

複共線性統計量		
變數項目	容忍值	變異數膨脹因素 (VIF)
(常數)	0.000	0.000
急減速	0.007	141.461
緊急煞車	0.010	100.735
急加速	0.056	17.811
超速	0.448	2.233
電磁煞車操作異常	0.562	1.780
怠速過久	0.256	3.908
檔位異常	0.276	3.620
引擎轉速異常	0.354	2.821

表5-7 第一次複共線性檢定（續）

複共線性診斷											
維度	特徵值	條件指數	變異數比例								
			(常數)	急減速	緊急煞車	急加速	超速	電磁煞車操作異常	怠速過久	檔位異常	引擎轉速異常
1	6.090	1.000	0.004	0.000	0.000	0.000	0.005	0.006	0.002	0.003	0.002
2	1.295	2.168	0.009	0.001	0.004	0.000	0.005	0.013	0.001	0.000	0.006
3	0.729	2.889	0.029	0.000	0.000	0.001	0.281	0.034	0.004	0.023	0.002
4	0.458	3.647	0.007	0.000	0.000	0.000	0.144	0.532	0.002	0.001	0.025
5	0.181	5.805	0.685	0.000	0.000	0.003	0.111	0.014	0.017	0.022	0.096
6	0.118	7.198	0.021	0.000	0.000	0.001	0.278	0.293	0.002	0.858	0.116
7	0.086	8.391	0.038	0.001	0.005	0.016	0.152	0.016	0.259	0.089	0.547
8	0.041	12.250	0.000	0.004	0.011	0.181	0.020	0.092	0.652	0.003	0.037
9	0.003	47.255	0.207	0.994	0.979	0.798	0.005	0.000	0.062	0.000	0.168

由於其原理與特性之關係，故將緊急煞車納入急減速的計算範圍中，並再次進行第二次複共線性檢定。檢定結果如下表5-8與表5-9所示。

表5-8 第二次複共線性檢定

複共線性統計量		
變數項目	容忍值	變異數膨脹因素 (VIF)
(常數)	0.000	0.000
急減速	0.469	2.134
急加速	0.181	5.536
超速	0.452	2.211
電磁煞車操作異常	0.562	1.778
怠速過久	0.279	3.581
檔位異常	0.276	3.620
引擎轉速異常	0.400	2.497

表5-9 第二次複共線性檢定（續）

複共線性診斷										
維度	特徵值	條件指數	變異數比例							
			(常數)	急減速	急加速	超速	電磁煞車 操作異常	怠速 過久	檔位 異常	引擎轉 速異常
1	5.838	1.000	0.005	0.005	0.001	0.005	0.006	0.002	0.004	0.003
2	0.778	2.739	0.050	0.014	0.003	0.269	0.002	0.005	0.018	0.008
3	0.546	3.270	0.000	0.318	0.001	0.001	0.286	0.001	0.004	0.001
4	0.434	3.666	0.020	0.184	0.000	0.165	0.297	0.002	0.001	0.048
5	0.177	5.749	0.817	0.006	0.008	0.109	0.006	0.016	0.018	0.165
6	0.116	7.083	0.023	0.056	0.005	0.332	0.270	0.000	0.912	0.070
7	0.076	8.751	0.075	0.177	0.029	0.088	0.014	0.497	0.041	0.563
8	0.034	13.017	0.010	0.239	0.953	0.031	0.118	0.477	0.002	0.142

5.2.2 模式設定

經由上述的複共線性檢定後，可以確認模式之輸入變數間不存在複共線性。接

著就是須將設定模式中之變數為自變數或因變數，基於此部分所探討之目的，本研究將肇事與交通違規作為因變數（ Y ），將急加速、急減速、超速、電磁煞車操作異常、怠速過久、檔位異常以及引擎轉速異常作為自變數（ X ），並分別構建兩個迴歸式。各變數意義與單位如表5-10，迴歸模式架構如下：

$$\text{模式A: } Y_A = \sigma_0 + \sigma_1 X_1 + \sigma_2 X_2 + \sigma_3 X_3 + \sigma_4 X_4 + \sigma_5 X_5 + \sigma_6 X_6 + \sigma_7 X_7 + \varepsilon_A$$

$$\text{模式B: } Y_B = \tau_0 + \tau_1 X_1 + \tau_2 X_2 + \tau_3 X_3 + \tau_4 X_4 + \tau_5 X_5 + \tau_6 X_6 + \tau_7 X_7 + \varepsilon_B$$

其中 $\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6, \sigma_7$ 與 $\tau_0, \tau_1, \tau_2, \tau_3, \tau_4, \tau_5, \tau_6, \tau_7$ 為迴歸係數。

表5-10 模式設定各項變數意義與單位

變數代號	意義	單位
Y_A	肇事	次/千公里
Y_B	交通違規	次/千公里
X_1	急減速	次/千公里
X_2	急加速	次/千公里
X_3	超速	次/千公里
X_4	電磁煞車操作異常	次/千公里
X_5	怠速過久	次/千公里
X_6	檔位異常	次/千公里
X_7	引擎轉速異常	次/千公里

5.2.3 模式校估

迴歸模式中選擇變數的方法有強迫進入法(Enter Method)、向前選取法(Forward Method)、向後選取法(Backward Method)以及逐步迴歸法(Stepwise Method)，本研究採用逐步迴歸法，先將所有變數輸入再逐步刪除不顯著之變數，並觀察模式中判定係數R-Square（ R^2 ）值檢視模式之解釋能力，F值（F-test）判定是否具統計意義，以及各項變數係數之t值是否顯著，以構建與校估迴歸模式。將所有變數輸入構建肇事迴歸模式與交通違規迴歸模式，校估結果如表5-11與表5-12。

● 判定係數 R^2 值

由社會科學研究之相關參考書籍得知，一般計量經濟之迴歸模式研究，當 R^2 值 ≥ 0.3 時，表示分析結果為可採信的。當然， R^2 值越高越好，代表迴歸模式的配適

度以及解釋能力越佳，即可解釋變異（SSR）佔總變異（SST）之比例越高。故在模式 A 與模式 B 中，所得到之 R^2 值：模式 A 為 0.552，模式 B 為 0.454，皆在可接受範圍內。

● F 值

在 $\alpha=0.05$ 的顯著水準下，模式 A 之 F 值檢定結果 $F=15.998 > F_{(1-0.05)}=2.839$ ，模式 B 之檢定結果 $F=8.469 > F_{(1-0.05)}=2.799$ ，因此 A 與 B 兩個模式皆拒絕 $H_0: \beta=0$ 之假設，故表示該兩模式皆具有統計上之意義。

● t 值

在 $\alpha=0.05$ 的顯著水準下，模式 A 之各項係數之 t 值檢定應 $|t| > t_{(1-0.05/2)}=2.023$ ，模式 B 之各項係數之 t 值檢定應 $|t| > t_{(1-0.05/2)}=2.011$ ，始能拒絕 $H_0: \beta_i=0$ 之假設。結果發現，模式 A 中的急加速變數（ X_2 ）係數為負值，以及其 t 值過小未能通過檢定，表示該係數並不具顯著性；而模式 B 中之急減速（ X_1 ）與急加速（ X_2 ）變數之 t 值亦過小未通過檢定，表示這兩個係數不具顯著性。

經由上述 t 檢定中，兩模式中皆有不具顯著性之變數係數，故需要再次進行模式校估之作業。

表 5-11 模式 A 與 B 變異數分析（一）

變異數分析						
模式		平方和	自由度	均方和	F 值	機率值(p)
模式 A	迴歸	0.007045	7	0.001	41.845	0.000*
	殘差	0.000842	35	0.000		
	總和	0.007887	42			
模式 B	迴歸	0.000279	7	0.000	6.078	0.000*
	殘差	0.000288	44	0.000		
	總和	0.000567	51			

註：*代表達 $\alpha=0.05$ 之顯著水準

表5-12 模式A與B校估結果（一）

統計值		迴歸模式			
		模式A		模式B	
R 值		0.945		0.701	
R ² 值		0.893		0.492	
調整後 R ² 值		0.872		0.411	
常數	係數 (t值)	0.005115	(3.496)	0.005485	(7.408)
X_1	係數 (t值)	0.000832	(6.957)	0.000054	(1.157)
X_2	係數 (t值)	-0.002086	(-2.419)	-0.000144	(-0.399)
X_3	係數 (t值)	-0.000567	(-1.840)	0.000101	(2.294)
X_4	係數 (t值)	0.000341	(0.539)	0.000806	(2.672)
X_5	係數 (t值)	-0.005875	(-2.323)	-0.001187	(-0.997)
X_6	係數 (t值)	0.000023	(1.805)	0.000007	(0.682)
X_7	係數 (t值)	0.001217	(9.824)	0.000041	(0.495)

在第二次校估結果中，模式A與模式B之判定係數R²值、F值與各項迴歸係數之t值皆合理具解釋力以及具顯著性，以及迴歸係數之正負值皆合理，皆為正值，校估之結果與各項數值內容整理如表5-13與表5-14。

表5-13 模式A與B變異數分析（二）

變異數分析						
模式		平方和	自由度	均方和	F 值	機率值(p)
模式 A	迴歸	0.004348	2	0.002174	24.572	0.000*
	殘差	0.003539	40	8.85E-05		
	總和	0.007887	42			
模式 B	迴歸	0.000257	2	0.000129	20.392	0.000*
	殘差	0.000309	49	6.3107E-06		
	總和	0.000567	51			

註：*代表達 $\alpha=0.05$ 之顯著水準

表5-14 模式A與B校估結果（二）

統計值		迴歸模式			
		模式A		模式B	
R 值		0.742		0.674	
R ² 值		0.552		0.454	
調整後 R ² 值		0.529		0.432	
常數	係數 (t值)	0.005551	(3.301)	0.004916	(11.135)
X_1	係數 (t值)	0.000500	(4.263)		
X_3	係數 (t值)	0.000893	(3.456)	0.000127	(4.114)
X_4	係數 (t值)			0.000778	(3.321)

經由第二次校估後，我們可以得到模式A與模式B之迴歸模式，模式A是駕駛員肇事次數（ Y_A ）為因變數，校估得到之模式中有急減速（ X_1 ）與超速（ X_3 ）為其自變數，由此可以得知駕駛員肇事次數與其急減速與超速之駕駛行為是有相關的，且為正向的相關，配合之前對於肇事原因之初步分析，迴歸模式是可以解釋其原因的，因為駕駛員若越常有急減速與超速之異常操作行為，發生肇事（擦撞或追撞）之比率會越高，急減速顯示出駕駛員之行車速率不夠平穩，超速顯示出駕駛員較愛開快車，此皆為容易發生肇事之駕駛行為。而在模式B中，結果得到是以交通違規次數（ Y_B ）與超速（ X_3 ）以及電磁煞車操作異常（ X_4 ）之迴歸關係式，此結果亦屬合理可解釋，配合之前交通違規之項目說明，超速為交通違規受到舉發開單之次多項目，由此可以得到平時容易發生超速駕駛行為之駕駛員，其交通違規被開單之比率亦會越高；另外，經由模式構建亦發現電磁煞車操作異常行為之發生與交通違規發生次數是有關聯的，經由模式得到結果可知，電磁煞車操作異常與交通違規次數是成正相關的，表示發生電磁煞車操作異常次數越頻繁之駕駛員，亦越容易發生交通違規。模式A與模式B之迴歸模式如下：

- 肇事模式（A）： $Y_A = 0.005551 + 0.0005X_1 + 0.000893X_3$
 $R^2=0.552$ (4.263) (3.456)
 - 交通違規模式（B）： $Y_B = 0.004916 + 0.000127X_3 + 0.000778X_4$
 $R^2=0.454$ (4.114) (3.321)
- 肇事模式與交通違規之 R^2 值解釋能力尚可，兩者之 R^2 值不高之原因可能為肇

事與交通違規之發生尚有其他因素導致，如交通違規部分，僅由超速與電磁煞車操作異常變數來解釋尚有不足，不過確實發現平常會超速的駕駛操作行為與受交通違規舉發（超速舉發）有所關聯，不過也因為超速的行為不一定會受到交通違規舉發，所以也可能為模式解釋能力稍差之原因。

在迴歸模式中，與發生肇事有關的異常操作行為是急減速與超速，若駕駛員未發生任何異常操作行為，迴歸模式中之常數項即為每千公里駕駛員可能發生肇事次數為0.0056次，而迴歸式中之係數即可代表異常操作行為發生後所產生之效果，若急減速每千公里發生一次，肇事發生次數則可能增加0.0005次，超速每千公里發生一次，肇事發生次數則可能增加0.0009次，由此可以得知超速對於發生肇事之影響比急減速來的大。至於與交通違規有關的則是超速與電磁煞車操作異常，常數項為0.0049代表未發生任何異常操作行為每千公里可能產生交通違規之次數，超速在交通違規之受罰原因中為次多之項目，最多為未依車道行駛應無相關之異常操作監控項目可衡量，故得到超速與交通違規有關聯是合理的，代表平時容易發生超速之駕駛員，亦較容易受到交通違規罰單，且每千公里增加一次之超速行為，交通違規受罰之次數可能會增加0.00013次，而每千公里增加一次電磁煞車操作異常行為，交通違規次數會增加0.00078次。

5.3 行車油耗與保養維修費用之模式構建

駕駛員之異常操作行為除了對於行車安全有所影響之外，還可能對於客運車輛之行車油耗以及相關保養維修費用造成影響，本研究欲經由樣本資料蒐集與篩選，利用多元迴歸模式檢測各駕駛員異常操作項目以及車輛油耗與保養維修費用之變數間影響程度與關係，以解釋駕駛員之異常操作行為對於車輛之油耗與保養費用之影響與關係。該客運公司裝設數位式行車紀錄器之車輛共有61部，於本模式構建中，扣除有一部分車輛維修或是調動至別路線，為確保資料正確性，取47個樣本用來構建行車油耗與維修費用之迴歸模式。

5.3.1 變數選取

本研究所訂定的異常操作行為監控項目主要有急加速、急減速（已將緊急煞車併入其中）、超速、電磁煞車操作異常、怠速過久、檔位異常以及引擎轉速異常等，首先選取上述所有的監控項目作為輸入之預測變數。再以前述七項變數進行複共

線性檢定，判定變數間是否具有複共線性，以確保迴歸模式之正確性。由分析的結果可見，該七項變數之間沒有存在複共線性之問題，結果整理如表 5-15 與表 5-16。

表 5-15 七項變數複共線性檢定

複共線性統計量		
變數項目	容忍值	變異數膨脹因素 (VIF)
(常數)	0.000	0.000
急減速	0.278	3.594
急加速	0.190	5.257
超速	0.514	1.945
電磁煞車操作異常	0.264	3.791
怠速過久	0.470	2.128
檔位異常	0.313	3.197
引擎轉速異常	0.227	4.405

表 5-16 七項變數複共線性檢定 (續)

複共線性診斷										
維度	特徵值	條件指數	變異數比例							
			(常數)	急減速	急加速	超速	電磁煞車操作異常	怠速過久	檔位異常	引擎轉速異常
1	6.882	1.000	0.000	0.001	0.000	0.004	0.002	0.000	0.001	0.001
2	0.724	3.083	0.002	0.000	0.000	0.426	0.007	0.002	0.001	0.000
3	0.200	5.872	0.025	0.085	0.002	0.319	0.137	0.006	0.000	0.004
4	0.073	9.678	0.004	0.369	0.004	0.035	0.485	0.002	0.073	0.053
5	0.055	11.139	0.003	0.177	0.021	0.044	0.206	0.003	0.337	0.124
6	0.043	12.644	0.052	0.266	0.004	0.007	0.125	0.006	0.188	0.422
7	0.015	21.321	0.521	0.026	0.001	0.036	0.000	0.781	0.076	0.007
8	0.007	30.638	0.393	0.075	0.968	0.130	0.038	0.200	0.325	0.389

5.3.2 模式設定

在本部分主要探討之課題為駕駛員異常操作行為與車輛油耗、保養維修費用間之關係，故就是須將各項變數視為自變數或因變數，本研究在此將燃油費用與保養維修費用作為因變數（ Y ），將急加速、急減速、超速、電磁煞車使用不當、怠速過久、檔位異常以及引擎轉速異常作為自變數（ X ）。將燃料費、附屬油料費以及修車材料費進行相關分析，發現燃料費與修車材料費和附屬油料費間無顯著之相關性，但修車材料費與附屬燃料費間存在相關性，相關係數為0.338，對於此現象是可以解釋的，因為附屬油料除了定期里程更換以外，尚會因為車輛故障進場維修而可能因為維修而需要更換，故兩者才會出現相關性，分析結果如表5-17。

表5-17 各項費用之相關分析

相關分析				
費用項目	統計值	修車材料費	附屬油料費	燃料費
修車材料費	Pearson 相關係數 (Sig.)	1	0.338* (0.02)	-0.135 (0.367)
	顯著性	—	顯著	不顯著
附屬油料費	Pearson 相關係數 (Sig.)	0.338* (0.02)	1	0.011 (0.943)
	顯著性	顯著	—	不顯著
燃料費	Pearson 相關係數 (Sig.)	-0.135 (0.367)	0.011 (0.943)	1
	顯著性	不顯著	不顯著	—
註：* 代表具相關性者達到 $P < 0.05$ 之顯著水準(雙尾)				

另外，附屬油料費用與駕駛員之操作行為無明顯關聯，因為附屬油料之更換幾乎是由定期里程來決定，不會因為駕駛員有較多之異常操作行為而造成額外之費用，就如前述之分析，惟有維修時有可能需額外更換附屬油料，故將附屬燃料費用併入修車材料費中，兩者合併稱為保養維修費。因此，燃料費與保養維修費之間並無相關性，故可分別建立油耗與保養維修費用的多元迴歸方程式，分別為模式C與模式D，並進行模式檢定與校估作業。表5-18為模式各項變數之意義與單位，各迴歸模式架構如下：

$$\text{模式C: } Y_C = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 + \alpha_5 X_5 + \alpha_6 X_6 + \alpha_7 X_7 + \varepsilon_C$$

$$\text{模式D: } Y_D = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \varepsilon_D$$

其中 $\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6, \alpha_7$ 與 $\beta_0, \beta_1, \beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7$ 為迴歸係數。

表 5-18 油耗與保養維修模式變數之意義與單位

變數代號	意義	單位
Y_C	燃料費	元/千公里
Y_D	保養維修費	元/千公里
X_1	急減速	次/千公里
X_2	急加速	次/千公里
X_3	超速	次/千公里
X_4	電磁煞車操作異常	次/千公里
X_5	怠速過久	次/千公里
X_6	檔位異常	次/千公里
X_7	引擎轉速異常	次/千公里

5.3.3 模式校估

油耗與保養維修部分之迴歸模式構建亦採用強迫進入法先將所有變數輸入，並觀察模式中判定係數R-Square (R^2) 值檢視模式之解釋能力，以及F值 (F-test) 判定是否具統計意義以及各項變數係數之t值是否顯著，以構建與校估迴歸模式。將所有變數輸入構建油耗費用迴歸模式與保養維修費用迴歸模式，校估結果如表5-19與表5-20。

- 判定係數 R^2 值

在 C 模式與 D 模式中，所得到之 R^2 值 C 模式為 0.855，D 模式為 0.890， R^2 值相當高，模式解釋能力佳。

- F值

在 $\alpha=0.05$ 的顯著水準下，C 模式之F值檢定結果 $F=32.942 > F_{(1-0.05)}=2.248$ ，D 模式之檢定結果 $F=44.874 > F_{(1-0.05)}=2.248$ ，因此C與D兩個模式皆拒絕 $H_0: \beta=0$ 之假設，故表示該兩模式皆具有統計上之意義。

- t值

在 $\alpha=0.1$ 的顯著水準下，C 模式之各項係數之t值檢定應 $|t| > t_{(1-0.05/2)}=1.682$ ，D 模式之各項係數之t值檢定應 $|t| > t_{(1-0.05/2)}=1.682$ ，始能拒絕 $H_0: \beta_i=0$ 之假設。結果

發現，C模式中之超速（ X_3 ）與檔位異常（ X_6 ）係數為負值並不合理，以及其t值過小未能通過檢定，表示該兩個係數並不具顯著性；而D模式中檔位異常（ X_6 ）係數為負值並不合理，且超速（ X_3 ）與檔位異常（ X_6 ）變數之t值亦未通過檢定，表示這兩個係數不具顯著性。

經由上述t檢定中，兩模式中皆有不具顯著性之變數係數，故需要再次進行模式校估之作業。

表5-19 模式C與D變異數分析（一）

變異數分析						
模式		平方和	自由度	均方和	F 值	機率值(p)
模式 C	迴歸	991042.861	7	141577.552	32.942	0.000*
	殘差	167611.498	39	4297.731		
	總和	1158654.359	46			
模式 D	迴歸	637770.842	7	91110.120	44.874	0.000*
	殘差	79184.148	39	2030.363		
	總和	716954.990	46			

註：*代表達 $\alpha=0.05$ 之顯著水準

表5-20 模式C與D校估結果（一）

統計值		迴歸模式			
		模式C		模式D	
R 值		0.925		0.943	
R ² 值		0.855		0.890	
調整後 R ² 值		0.829		0.870	
(常數)	係數 (t值)	4652.873	(74.097)	558.337	(12.936)
X ₁	係數 (t值)	4.177	(1.343)	4.563	(2.134)
X ₂	係數 (t值)	65.248	(2.989)	47.644	(3.175)
X ₃	係數 (t值)	-0.027	(-0.019)	0.413	(0.411)
X ₄	係數 (t值)	14.162	(0.786)	18.583	(1.500)
X ₅	係數 (t值)	81.890	(1.230)	54.292	(1.187)
X ₆	係數 (t值)	-0.053	(-0.090)	-0.036	(-0.090)
X ₇	係數 (t值)	8.695	(2.154)	5.094	(1.836)

由上述結果得到，在燃料費模式中，急減速（X₁）、超速（X₃）、電磁煞車操作異常（X₄）以及檔位異常（X₆）之迴歸係數t檢定結果係數值不顯著以及有兩項係數值為負不合理。保養維修費的模式中，超速（X₃）、怠速過久（X₅）與檔位異常（X₆）之迴歸係數t檢定結果係數值不顯著，且檔位異常係數為負值並不合理。本研究探究其原因，急減速、電磁煞車操作異常與油耗較無關係屬合理之情況，因為油耗的增加駕駛員油門控制有關，而怠速過久實際上亦無明顯與保養維修有關係，除非長期的怠速過久造成車輛嚴重積碳才會造成車輛可能需要維修。超速的部分因為監控門檻值是依照法定規定速限再加上5%之緩衝區進行設定，在時速達105km/h時會以蜂鳴器提醒駕駛員，在時速達107km/h以上超過5秒才會紀錄，故實際上一般之駕駛員受到蜂鳴器之提醒即會減速，所以超過門檻值之超速行為會較少，即使超速有可能是因為下坡或是負載較輕且引擎轉速並未過高，除非是嚴重異常之超速行為才有可能對於油耗或保養維修費用有顯著之影響。至於檔位異常部分，因為該公司所採用之客運車輛為日本三菱FUSO之大客車，其變速箱無法輸出檔位訊號提供數位式行車紀錄器參考，故僅能由一般速度與引擎轉速之區間，設定檔位異常之門檻值，此項目可能會與實際情況有所差異，應為造成模式中迴歸係數

不合理以及不顯著之情況。故將上述各項不合理與不顯著之變數排除，並作進一步的校估。第二次校估結果整理如表5-21與表5-22。

表5-21 模式C與D變異數分析（二）

變異數分析						
模式		平方和	自由度	均方和	F 值	機率值(p)
模式 C	迴歸	972278.8	3	324092.9	74.774	0.000*
	殘差	186375.5	43	4334.315		
	總和	1158654	46			
模式 D	迴歸	634275.7	4	158568.9	80.551	0.000*
	殘差	82679.25	42	1968.554		
	總和	716955	46			

註：*代表達 $\alpha=0.05$ 之顯著水準

表5-22 模式C與D校估結果（二）

統計值		迴歸模式			
		模式C		模式D	
R 值		0.916		0.941	
R ² 值		0.839		0.885	
調整後 R ² 值		0.828		0.874	
(常數)	係數 (t值)	4604.711	(81.938)	586.9	(17.871)
X ₁	係數 (t值)	-	-	4.917	(2.445)
X ₂	係數 (t值)	79.526	(4.505)	52.805	(4.224)
X ₄	係數 (t值)	-	-	19.377	(1.867)
X ₅	係數 (t值)	92.807	(1.691)	-	-
X ₇	係數 (t值)	11.696	(3.336)	5.151	(2.014)

經由第二次之校估後，模式中之係數值與各項統計值皆為合理之情況，模式 C 之 R² 值為 0.839，F 值為 74.774；而模式 D 之 R² 值為 0.885，F 值為 80.551，R² 值與 F 值皆依然相當高，模式解釋能力佳。

迴歸模式結果發現，模式C中最後得到之變數為急加速（ X_2 ）、怠速過久（ X_5 ）以及引擎轉速異常（ X_7 ），模式D中之變數為急減速（ X_1 ）、急加速（ X_2 ）、電磁煞車操作異常（ X_4 ）以及引擎轉速異常（ X_7 ），即代表上述這些異常操作行為，會對於車輛油耗與保養維修費用產生影響，這些異常操作行為次數越多，油耗與保養維修之費用亦會越高，其中各項異常操作行為之影響程度可以由迴歸係數之大小來判斷。模式C與模式D之多元迴歸模式如下所示：

- 油耗模式（C）：

$$Y_C = 4604.711 + 79.526X_2 + 92.807X_5 + 11.696X_7$$

$$R^2 = 0.839 \quad (4.505) \quad (1.691) \quad (3.336)$$

- 保修模式模式（D）：

$$Y_D = 586.9 + 4.917X_1 + 52.805X_2 + 19.377X_4 + 5.151X_7$$

$$R^2 = 0.885 \quad (2.445) \quad (4.224) \quad (1.867) \quad (2.014)$$

迴歸模式中之各項係數大小即為該項變數對於油耗或保養維修費用的影響程度，係數值越大，影響程度越大。在油耗費用模式中，常數項代表駕駛員於一般正常行駛無異常操作行為下每千公里會消耗4605元的燃料費，各項係數值代表駕駛員每增加一次異常操作行為會產生額外的燃料費用，每千公里增加一次急加速會增加79.53元、每千公里增加一次怠速過久會增加92.81元、每千公里增加一次引擎轉速異常會增加11.7元，由此可見，怠速過久影響最大，其次為急加速，最後為引擎轉速異常。在保養維修費用模式中，常數項代表駕駛員完全於無異常操作行為下每千公里會有586.9元的保養維修費，各項係數值代表駕駛員每增加一次異常操作行為會產生額外的保養維修費用，每千公里增加一次急減速會增加4.917元、每千公里增加一次急加速會增加52.805元、每千公里增加一次電磁煞車操作異常會增加19.377元、每千公里增加一次引擎轉速異常會增加5.151元，由此可知，急加速之影響最大，其次為電磁煞車操作異常，以及引擎轉速異常，最後為急減速。

5.4 小結

本研究運用資料採礦技術，將數位式行車紀錄器中駕駛員異常操作行為之資料進行整理與分析，駕駛員之異常操作行為主要會影響行車安全、燃油消耗以及機件磨損等方面，於本章節對於駕駛員異常操作行為對於發生肇事與交通違規以及對於油耗與保養維修費用之影響進行探討，以資料採礦技術中屬於預測方面之迴歸方法

進行多元迴歸模式構建，以尋求其彼此之關係或趨勢。

行車安全部分，與發生肇事有關的異常操作行為是急減速與超速，由迴歸模式之係數大小來看，可以得知超速對於發生肇事之影響比急減速來的大。至於與交通違規有關的則是超速與電磁煞車操作異常，超速在交通違規之受罰原因中為次多之項目，最多為未依車道行駛應無相關之異常操作監控項目可衡量，故得到超速與交通違規有關聯是合理的，代表平時容易發生超速之駕駛員，亦較容易受到交通違規罰單，但因為超速行為不一定會受到違規舉發，故可能為模式解釋能力稍差之原因。亦發現電磁煞車操作異常與交通違規呈正相關，表示電磁煞車操作亦常行為越頻繁之駕駛員，其越容易發生交通違規。

油耗與保養維修部份，影響油耗費用的異常操作行為是急加速、怠速過久與引擎轉速異常等三個影響因素，係數值大小可以判斷彼此間影響之程度，分別為怠速過久影響最大，急加速次之，最後是引擎轉速異常，怠速過久將造成油料的浪費，急加速攸關駕駛員油門之控制行為，油門急踩自然增加油耗，引擎轉速越高油耗也自然越大。影響保養維修費用之駕駛員異常操作行為是急減速、急加速、電磁煞車操作異常以及引擎轉速異常，其中影響程度由大至小分別為急加速、電磁煞車操作異常、引擎轉速異常，最後為急減速。異常或頻繁的急加速、急減速與引擎轉速異常的確是會造成車輛的負擔進而造成零件故障，亦會使附屬油品加速劣化而可能導致零件故障，而電磁煞車異常操作會增加電磁線圈故障之機率。

第六章 駕駛員分群判別與管理

本研究於前述第五章運用多元迴歸分析構建駕駛員異常操作行為關係模式，探討駕駛員異常操作行為對於行車安全與行車油耗與保養維修費用之影響項目與關係，可以得知相關之影響變數。在本章節將運用資料採礦之群集分析方法以及使用判別分析方法，將駕駛員進行分群，再配合前述章節所構建模式所得到之結果，對於駕駛員管理提出相關之方法與對策案例，與評估可能產生之效益。

6.1 群集分析方法

群集分析為資料採礦技術中的其中一種，本研究將使用群集分析來對於駕駛員進行分群。依據先前所述，群集分析方法主要有二種，一為「階層群集分析法」如平均連結法或是華德法，二為「非階層群集分析法」之 K-Means 法(K-means methods)。由於實際應用上階層群集分析法所耗資源與時間過於龐大，因此在實務上以非階層群集分析法為主。吳明隆（2002）也提出：若觀察值的個數較多或資料非常龐大（樣本數超過 200 個以上），以採「K-平均群集分析法」較為適宜。本研究之駕駛員樣本有 210 個，且於資料庫中駕駛員異常操作監控項目內有 3,757,420 筆資料，故以資料採礦技術中群集分析之 K-Means 分析法較為適合亦較有效率。且非階層式群集法的 K-Means 法較階層式群集法不受異常事物之存在、距離衡量之誤差及距離計算方法之選擇的影響。若以非隨機性方式選取起始點（即研究資料為已知，或是先經由階層式群集分析獲得上述資料），則 K-Means 法之效果要比華德法及平均連結法好，而華德法和平均連結法又較其他的階層群集分析法來的好。因此採用 K-Means 分析法為本研究之資料採礦群集分析法，以歐氏平方距離為距離衡量之方式。

6.2 駕駛員群集分析

本研究欲透過群集分析之過程，探討整體駕駛員之異常操作行為，是否為某些駕駛員特性相似，而不同群之駕駛員有明顯差異之情形，或是為一平均類似之現象。群集分析之分群參考依據，主要為之前迴歸分析結果所得到具顯著性以及不具有共

線性之變數，主要有急減速、急加速、超速、電磁煞車操作異常、怠速過久以及引擎轉速異常等六項監控指標，這些指標主要關於行車安全、油耗與保養維修費用。本研究將以 SPSS 13.0 統計軟體作為運算工具，並採用非階層群集分析法中之 K 平均法（K-Means）來進行群集分析。

6.2.1 群集數選擇

根據 K-Means 群集分析法之演算規則，使用 K-Means 法時，要事先訂定群集個數，在一般研究應用上，群集數的決定可依實務上的需要而設定，一般而言以 3 到 10 個群集較易溝通處理。根據本研究之資料性質與研究目的，以及管理之評等需求，決定將駕駛員分群於 3 到 7 個群集之範圍進行選擇。乃配合多變量統計法之判別分析，找出最合適群集數之門檻值進行分析，作為進行 K-Means 分析法時群集數之選擇依據。

由圖 6-1 的駕駛員異常操作行為分群之群集數選擇變化表情形可發現，經過不同群集數的分群結果，本研究主要以奇數群進行選擇且群數不宜過多，以便後續駕駛員分群評等之用，發現奇數群中以分為三群時正判率為最高，隨後正判率即下降，分為六群又回升，隨後又立即下降，並隨著群集數的增加而小幅變動，因此駕駛員異常操作行為分群以三群為最佳分群數。

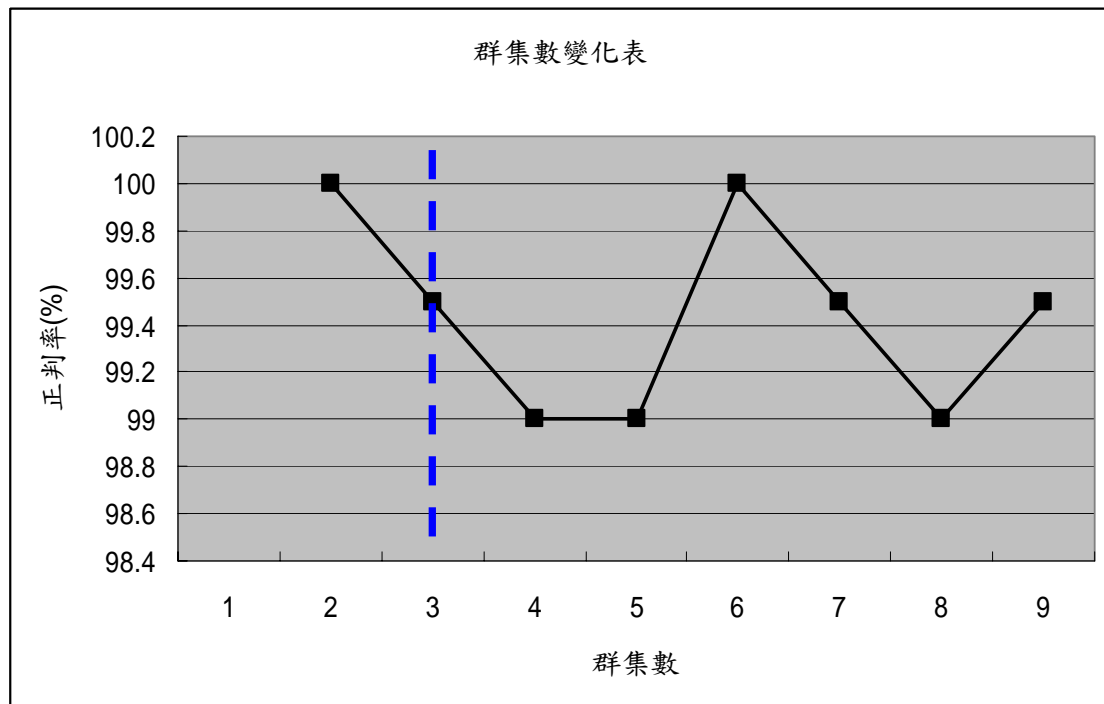


圖 6-1 駕駛員異常操作行為群集數變化表

6.2.2 K-Means 群集分析

本階段將利用非階層群集分析法中採用較廣之 K-Means 分析法，其演算步驟如下：

1. 將某一事物點分割成 K 個原始群集。
2. 計算某一事物點到各群集中心（平均數）的距離，研究中採用歐基里德平方距離，然後將一些事物點分派到距離最近的群集。重新計算得到新事物點的群集和喪失該事物點的群集兩者的重心。
3. 重複第二步驟，直到各事物點皆不須重新分派到其他群集為止。

依據其演算概念，以及之前所做的群集數選擇後，將所蒐集到的 210 名駕駛員樣本分割成三個原始群集（即為 $K=3$ ）進行群集分析。經由 SPSS 程式運算，K-Means 分析法之計算流程與最後計算結果如表 6-1 至表 6-3。

表 6-1 即 K-Means 分析法初期將定出三群的起始重心。接著進行第 2 步驟的計算。表 6-2 則為 SPSS 內部運算各事物到群集重心的流程。

表 6-1 各群集之起始中心點

群集起始中心點			
指標變數	群集		
	1	2	3
急減速	86.64	286.96	0.76
急加速	29.67	24.39	0.39
超速	19.28	16.28	0.00
電磁煞車操作異常	12.73	11.60	0.17
怠速過久	5.11	11.92	0.08
引擎轉速異常	73.44	87.40	0.97

表 6-2 各群集之疊代紀錄

疊代紀錄			
疊代	群集中心點之變更		
	1	2	3
1	54.38	54.38	49.62
2	7.21	7.21	0.93
3	3.32	3.32	0.26
4	0.00	0.00	0.00

註：因為距離沒有改變或距離改變很小而達成的收斂。

任何中心點已經變更的最大距離為 0.00。目前的疊代次數為 4。

表 6-3 各群集之最後中心點

群集最後中心點			
指標變數	群集		
	1	2	3
急減速	93.96	246.08	7.09
急加速	16.17	25.43	3.51
超速	21.33	34.27	2.12
電磁煞車操作異常	5.74	6.85	1.10
怠速過久	2.54	9.91	0.80
引擎轉速異常	63.82	83.19	8.54

最後計算得到三個群集各自中心點之數值，由此可得各群集的指標平均值。表 6-3 中，若以平均異常操作次數多寡來分優劣，其駕駛員分群之優劣順序為第 3 群、第 1 群、第 2 群。並且可由表 6-4 中心點間的距離瞭解群集間之差異。

表 6-4 各群集最後中心點間之距離

群集最後中心點間之距離			
群集	1	2	3
1	-	154.35	105.62
2	154.35	-	253.61
3	105.62	253.61	-

6.2.3 群集假設檢定

根據上述群集結果，以變異數分析各駕駛員異常操作變項在不同組群集之平均數是否確實存在顯著差異。根據此項研究主題，提出下列假設（Hypotheses），茲陳述如下。流程圖如下圖 6-2 所示。

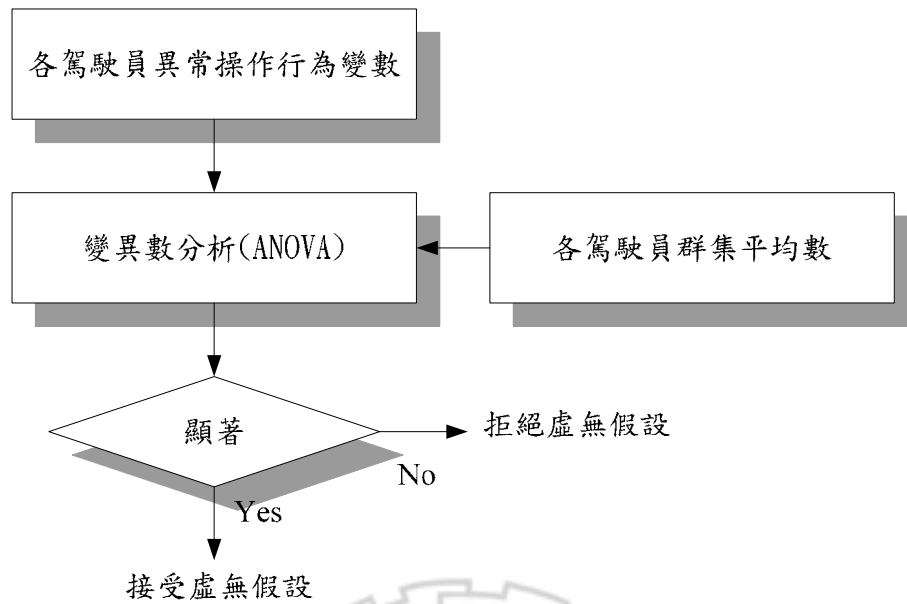


圖 6-2 各群集與變數顯著性檢定流程圖

依據上述由 K-Means 分析法將駕駛員所分成之群集，檢定各異常操作變數與駕駛員分群之顯著性，各駕駛員群集中之平均數如表 6-5。

表 6-5 各駕駛員群集中之平均數

指標變數	群集		
	1	2	3
急減速	93.96	246.08	7.09
急加速	16.17	25.43	3.51
超速	21.33	34.27	2.12
電磁煞車操作異常	5.74	6.85	1.10
怠速過久	2.54	9.91	0.80
引擎轉速異常	63.82	83.19	8.54

經由變異數分析後，如表 6-6，可以得知上述三個群集中之平均數皆與急減速、急加速、超速、電磁煞車操作異常、怠速過久以及引擎轉速異常等六個駕駛員異常操作變數上皆達顯著差異，意思就是該六個駕駛員分群變數可以顯著地將 210 個駕駛員樣本分為三群。

表 6-6 變異數分析表

指標變數	群集		誤差		F 值	機率值(p)
	均方	自由度	均方	自由度		
急減速	87298.52	2	189.11	207	461.64	0.00*
急加速	1142.08	2	10.55	207	108.27	0.00*
超速	2558.92	2	39.09	207	65.47	0.00*
電磁煞車操作異常	123.42	2	3.54	207	34.89	0.00*
怠速過久	93.77	2	0.49	207	190.36	0.00*
引擎轉速異常	18325.18	2	130.97	207	139.92	0.00*

註：*代表達 $\alpha=0.05$ 之顯著水準

6.2.4 駕駛員分群結果

經由上述的 K-Means 分析以及檢定各項駕駛員分群變數是否能夠有效的將 210 個駕駛員樣本確實的分成三群，其結果是顯著的，即可以確定 K-Means 分析法將駕駛員分成三群是正確的，故可將駕駛員分成三群分別命名為一般、稍差、極差，因為群集 3 中共有 199 名駕駛員，且其各項異常操作項目之平均值較小，故可視為該客運公司中一般駕駛員之情況，故將該群集命名為一般，其餘兩個群集因為異常操作之平均值皆偏高且人數較少，故依情節輕重分別命名為稍差與極差。並以該三個評等用來做為後續駕駛員管理獎懲與再教育訓練之依據。各分群之駕駛員如表 6-7。

表 6-7 駕駛員分群表

駕駛員分群 等級	駕駛員 ID	總計
一般	D001、 D002、 D003、 D004、 D005、 D006、 D007、 D008、 D009、 D010、 D011、 D012、 D013、 D014、 D015、 D016、 D018、 D019、 D020、 D021、 D022、 D023、 D024、 D025、 D026、 D027、 D028、 D030、 D031、 D032、 D033、 D034、 D035、 D036、 D037、 D038、 D040、 D041、 D042、 D043、 D044、 D045、 D046、 D047、 D048、 D049、 D050、 D051、 D053、 D054、 D055、 D056、 D057、 D058、 D059、 D060、 D061、 D062、 D063、 D064、 D065、 D066、 D068、	199

	D069、D070、D071、D072、D073、D074、D075、D076、D077、D078、D079、D080、D081、D082、D083、D084、D085、D086、D087、D088、D089、D090、D091、D092、D093、D094、D095、D096、D097、D098、D099、D100、D101、D102、D103、D104、D105、D106、D107、D109、D110、D111、D112、D113、D114、D115、D116、D117、D118、D119、D120、D121、D122、D123、D124、D125、D126、D127、D128、D129、D130、D132、D133、D134、D135、D136、D137、D138、D139、D140、D141、D142、D146、D147、D148、D149、D150、D151、D152、D154、D155、D156、D157、D158、D159、D160、D161、D162、D163、D164、D165、D166、D167、D168、D169、D170、D171、D172、D173、D174、D175、D176、D177、D178、D179、D180、D181、D182、D183、D184、D185、D186、D187、D188、D189、D190、D191、D192、D193、D194、D195、D196、D197、D198、D199、D200、D201、D202、D203、D204、D205、D206、D207、D208、D209、D210	
稍差	D017、D029、D039、D052、D067、D108、D131、D143、D153	9
極差	D144、D145	2

6.3 駕駛員異常操作行為判別模式

6.3.1 判別模式架構

判別分析所建立之判別關係式可提供研究者對於一群未知狀況之個體，透過將其特性因素投入判別關係式後得到其所希望區別分類之目的，應用於本研究，即為透過投入一些影響駕駛員異常操作行為之重要變項資料，結合前述群集分析得到之駕駛員分群情況，以構建完成駕駛員分群判別模式。

由之前的群集分析結果可知，影響駕駛員分群之顯著變項有急減速、急加速、超速、電磁煞車操作異常、怠速過久以及引擎轉速異常等六項異常操作指標，並採用該六項變項進行判別分析，以構建判別模式以及求得判別係數，驗證之前群集分析之準確度，並且作為未來有駕駛員之異常操作資料時，但不知道其會屬於那一個

評等的群集，即可利用判別模式加以歸類。本研究之判別分析架構如圖6-3。

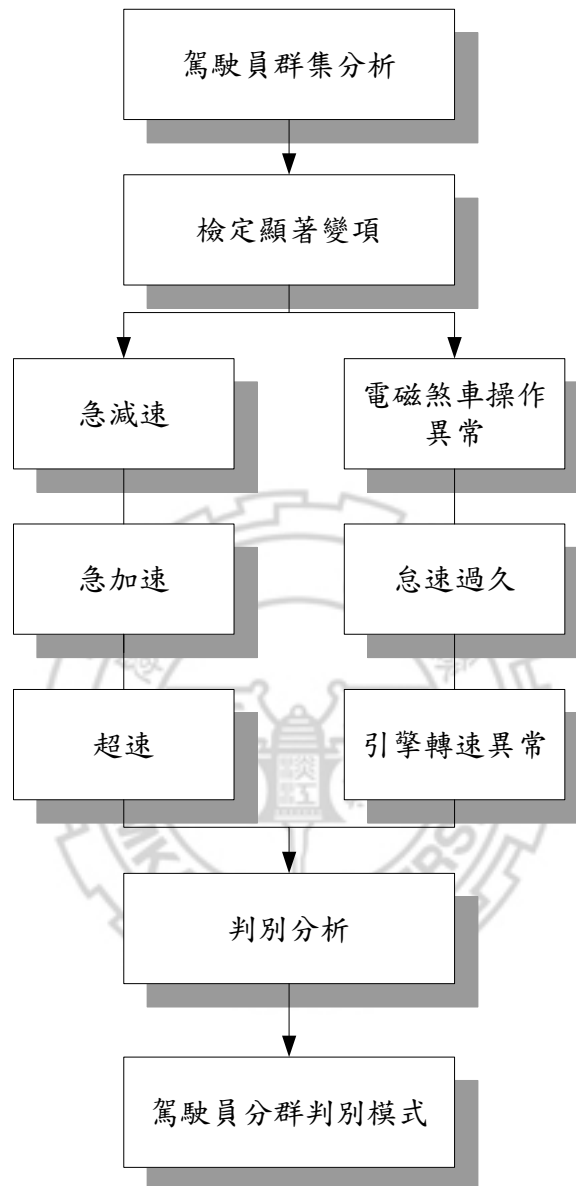


圖6-3 判別分析與模式構建流程圖

6.3.2 判別模式構建與結果分析

本研究之判別模式架構，以駕駛員群集為判別函數值，各項駕駛異常操作指標為判別變數，如下式所示：

$$D_x = d_0 + d_1 X_1 + d_2 X_2 + d_3 X_3 + d_4 X_4 + d_5 X_5 + d_6 X_6$$

其中 D_x 為駕駛員群集判別函數

X_1 為急減速

X_2 為急加速

X_3 為超速

X_4 為電磁煞車操作異常

X_5 為怠速過久

X_6 為引擎轉速異常

經由之前使用K-Means分析法將駕駛員分群後，將其群集分析之結果經由判別分析建立下列判別函數分別可得典型判別函數，包含兩個判別函數，其各項係數如表6-8所示，以及費雪線性判別函數，其由費雪判別係數構成三項線性判別模式如表6-10，並經由Wilks Λ 值與卡方檢定結果顯示所構建之判別函數是能夠有效的分群使得群集間有顯著的差異，如表6-9所示。本研究以費雪線性判別函數作為駕駛員各群之間的判別模式，後續將駕駛員異常次數直接代入即可分群，有利於實務上之運用。

表6-8 典型判別函數之係數值與結構矩陣

	未標準化		標準化		結構矩陣	
	1	2	1	2	1	2
急減速	0.0701	0.0037	0.9639	0.0508	0.8093	0.3089
急加速	-0.1760	0.2798	-0.5716	0.9088	0.5200	-0.1930
超速	0.0305	0.0075	0.1909	0.0466	0.3960	0.5790
電磁煞車操作異常	-0.2604	0.1190	-0.4897	0.2238	0.3685	0.3919
怠速過久	1.0030	-1.7583	0.7040	-1.2341	0.2846	0.3178
引擎轉速異常	0.0232	0.0461	0.2649	0.5281	0.1934	0.3101
(常數)	-1.1517	-0.2653				

表6-9 駕駛員判別分析之特徵值與Wilks Λ 檢定統計分析表

判別函數	各項統計檢定值						
	特徵值	解釋變異數 (%)	典型相 關係數	檢定的判 別函數	Wilk's Lambda	卡方值	機率值(p)
1	6.68	88.03	0.93	1&2	0.07	548.96	0.00*
2	0.91	11.97	0.69	2	0.52	132.12	0.00*

註：*代表達 $\alpha=0.05$ 之顯著水準

表6-10 費雪線性判別函數係數表

分類函數係數			
	群集觀察值個數		
	1	2	3
(常數)	-34.3815	-297.4743	-1.8694
急減速 (X_1)	0.4486	1.6070	-0.0049
急加速 (X_2)	0.1701	-5.2528	0.1342
超速 (X_3)	0.2659	0.7196	0.0446
電磁煞車操作異常 (X_4)	-1.2110	-6.6684	-0.0658
怠速過久 (X_5)	0.3218	32.6506	1.1915
引擎轉速異常 (X_6)	0.3469	0.3388	0.0146

可由上述得到的費雪判別函數係數值，可得到三組駕駛員判別函數如下：

- 群集-3 (一般) =
 $-1.8694 - 0.0049 X_1 + 0.1342 X_2 + 0.0446 X_3 - 0.0658 X_4 + 1.1915 X_5 + 0.0146 X_6$
- 群集-1 (稍差) =
 $-34.3815 + 0.4486 X_1 + 0.1701 X_2 + 0.2659 X_3 - 1.2110 X_4 + 0.3218 X_5 + 0.3469 X_6$
- 群集-2 (極差) =
 $-297.4743 + 1.6070 X_1 - 5.2528 X_2 + 0.7196 X_3 - 6.6684 X_4 + 32.6506 X_5 + 0.3388 X_6$

依據所得之判別函數進行分類時，將每一觀察值帶入上述三組函數，分類函數大者即屬於該群集，及完成駕駛員分群之判別。再經由群集分析所得到的實際駕駛員分群結果，與判別分析進行交叉比對，顯示結果如表6-11所示，可以得到整體預

測分群正確率達99.5%，正確判別率佳。因此，經由群集分析的結果再配合判別分析的檢定，此群集結果是可以接受的。

表6-11 駕駛員分群判別分析結果分析表

		駕駛員群集	預測駕駛員類組			總計
實際類組	個數		1 (稍差)	2 (極差)	3 (一般)	
		1 (稍差)	8	0	1	9
		2 (極差)	0	2	0	2
		3 (一般)	0	0	199	199
	百分比	1 (稍差)	88.89	0	11.11	100
		2 (極差)	0	100	0	100
		3 (一般)	0	0	100	100

註：正確判別率99.5% = (8+2+199)/210。

6.4 駕駛員分群管理

為何需要將駕駛員分群？因為駕駛員之異常操作行為會影響行車安全、油耗與保養維修費用，這三大方面攸關客運公司之營運，故可以得知，有必要依據異常操作之情形，對於駕駛員進行分群，有了適當的分群後，即可判斷駕駛員是屬於哪一群組，便可以對於不同群組進行不同之管理方式。

6.4.1 駕駛員評鑑等第

經由資料採礦技術群集分析將210名駕駛員分成三個群集，並分別命名為一般、稍差與極差共三個等第。之前的分群結果，一般等第有199名駕駛員，稍差等第有9名駕駛員，而極差等第有2名駕駛員，再配合資料庫中各個駕駛員之異常操作次數，分別計算各個等第中異常操作項目之範圍、平均值與標準差，可以得知各等第之基本統計值，若有駕駛員相關異常操作數據，若恰好其各項異常操作值皆落於某等第之範圍內，即可馬上判斷其屬於哪個等第。各等第與異常操作項目統計值整理如表6-12。

表6-12 各等第之基本統計值

項目 等第		急減速	急加速	超速	電磁煞車 操作異常	怠速過久	引擎轉速 異常
一般	範圍	0.76~ 62.76	0.39~ 15.09	0~ 39.79	0~ 9.21	0~ 2.92	0.86~ 75.04
	平均值	7.09	3.51	2.12	1.10	0.80	8.54
	標準差	10.95	2.54	5.10	1.56	0.60	8.99
稍 差	範圍	39.24~ 160.48	4.94~ 31.17	0~ 51.95	0.24~ 12.73	0.78~ 5.11	10.86~ 136.58
	平均值	93.96	16.17	21.33	5.74	2.54	63.82
	標準差	38.86	10.65	16.90	5.04	1.65	37.22
極 差	範圍	205.19~ 286.96	24.39~ 26.47	16.28~ 52.27	2.10~ 11.60	7.91~ 11.92	78.97~ 87.40
	平均值	246.08	25.43	34.27	6.85	9.91	83.19
	標準差	57.82	1.47	25.45	6.72	2.83	5.96

由上表可以看出，該客運公司落於一般等第之駕駛員各項異常操作項目平均每千公里之次數值並不會很大，此屬於正常之情形，因為一般正常之駕駛行為是不會超出異常操作監控之門檻值的，但難免會發生異常操作行為，可能是因路況或是其他環境因素影響，此亦顯示於該客運公司對於駕駛員之管理恰當，並未出現許多駕駛員落於稍差或極差之等第。而分群的目的主要是用來考核駕駛員，以一般、稍差與極差三等第評鑑駕駛員操作行為，並訂定在一般等第中表現優於平均值之駕駛員為優良駕駛員，並予以額外之獎勵。可採每週或每月等固定週期統計駕駛員行車表現，對於考核成績為稍差或極差者，管理人員須進行瞭解及監控管理，若有被評等為極差的駕駛人，公司必須讓該名駕駛員立即接受再教育訓練，且若情節重大或多次未改善應予以免職。

由於每位駕駛員的駕駛特性不盡相同，從表6-4-1各異常操作項目之範圍中可發現，某駕駛員的急減速次數與超速落於一般等第，而其他四個項目次數可能落在其他等第的範圍值內，故此時並不能立即研判該評等該名駕駛員等第。基於此問題即可利用之前判別分析所構建出之判別模式，運用判別模式將駕駛員之異常操作項目次數代入，所得到之數值較大者即屬於該等第，即可判斷駕駛員之等第。例如有某

位駕駛員之異常操作行為平均每千公里急減速280次、急加速5次、超速17次、電磁煞車操作異常10次、怠速過久6次以及引擎轉速異常20次，將這些操作數值分別代入駕駛員評鑑等第判別模式中，如表6-13，並計算其值。

表6-13 駕駛員評鑑等第判別模式

等第	駕駛員評鑑等第判別模式
一般	$-1.8694 - 0.0049 X_1 + 0.1342 X_2 + 0.0446 X_3 - 0.0658 X_4 + 1.1915 X_5 + 0.0146 X_6$
稍差	$-34.3815 + 0.4486 X_1 + 0.1701 X_2 + 0.2659 X_3 - 1.2110 X_4 + 0.3218 X_5 + 0.3469 X_6$
極差	$-297.4743 + 1.6070 X_1 - 5.2528 X_2 + 0.7196 X_3 - 6.6684 X_4 + 32.6506 X_5 + 0.3388 X_6$

$$-1.8694 - 0.0049 * 280 + 0.1342 * 5 + 0.0446 * 17 - 0.0658 * 10 + 1.1915 * 6 + 0.0146 * 20 = 4.97$$

$$-34.3815 + 0.4486 * 280 + 0.1701 * 5 + 0.2659 * 17 - 1.211 * 10 + 0.3218 * 6 + 0.3469 * 20 = 92.96$$

$$-297.4743 + 1.607 * 280 - 5.2528 * 5 + 0.7196 * 17 - 6.6684 * 10 + 32.6506 * 6 + 0.3388 * 20 = 274.45$$

經由上列之計算得知，由極差等第之判別式所計算得到之數值最大，故該駕駛員落於極差之等第。將駕駛員評鑑等第判別模式落實於實際駕駛員管理層面，可以有以下作法：

- 新進駕駛員管理：駕駛員是客運公司的命脈，為了解決駕駛員不足的問題客運公司常會招募新進駕駛員，對於新進駕駛員的訓練與教育相當重要，除了以訓練和教育提升新進駕駛員素質外，亦需要有效的管理，故可利用駕駛員評鑑等第判別模式，判斷新進駕駛員於職前訓練與教育後實際上路之表現，若異常操作情況嚴重應予以再教育訓練，若屢次無法改善應予免職。
- 在職駕駛員管理：一般在職駕駛員指的是已經在客運公司工作一段時間，對於公司之情形與相關規定有一定程度的了解，不過仍會有少數的駕駛員發生較多的異常操作行為，此時可利用駕駛員評鑑等第判別模式判斷是否有超出一般等第的情況，若落於稍差或極差等第，相關管理人員必須對於該駕駛員之情況進行了解與溝通，試圖了解原因，並予以進行再教育訓練課程，若屢次情節重大者，應評估予以免職，以確保公司之營運正常。

6.4.2 駕駛員行車安全風險與額外費用估計

本研究將210名駕駛員進行分群並歸類於各等第，配合第五章所構建之肇事與交通違規迴歸模式，可以計算出各個等第駕駛員之平均每千公里之肇事次數與交通違規次數，即為駕駛員之肇事風險與交通違規風險。並由多重比較檢定可以得知各等第之間的肇事與交通違規次數在統計上有明顯的差異，在肇事風險方面，一般等第為平均每千公里0.011次、稍差等第為0.0713次、極差等第為0.1592次；交通違規風險方面，一般等第為平均每千公里0.006次、稍差等第為0.0121次、極差等第為0.0146次各等第平均數據與統計檢定整理如表6-14與表6-15。

表6-14 異常操作行為風險表

	一般	稍差	極差	總平均
肇事風險	0.011	0.0716	0.1592	0.015
交通違規風險	0.006	0.0121	0.0146	0.0063
人數(n)	199	9	2	210

註：單位（次/千公里）

表6-15 多重比較檢定（一）

等第(i-j)	肇事風險			交通違規風險		
	平均差	標準誤	機率值(p)	平均差	標準誤	機率值(p)
1-2	-0.088	0.007	0.000*	-0.002	0.001	0.009*
1-3	0.061	0.003	0.000*	0.003	0.000	0.000*
2-3	0.148	0.007	0.000*	0.005	0.001	0.000*

註：1、2、3分別指一般、稍差、極差；*代表達 $\alpha=0.05$ 之顯著水準

該客運公司目前約有480名駕駛員，故本研究將其等比例放大至各等第中，分別為一般等第456人、稍差等第19人以及極差等第5人，並依據上表之數據可以估計目前公司每千公里之肇事與交通違規風險如下：

- 肇事風險：

$$0.011*(456/480)+0.0716*(19/480)+0.1592*(5/480)=0.0149 \text{ (次/千公里)}$$

- 交通違規風險：

$$0.006*(456/480)+0.0121*(19/480)+0.0146*(5/480)=0.0063 \text{ (次/千公里)}$$

各等第之駕駛員異常行為所產生之平均額外油耗與保養維修費用，額外燃料費用部份其一般等第為平均每千公里453.81元、稍差等第為2267.71元、極差等第為3915.27元；額外保養維修費用部份一般等第為平均每千公里285.55元、稍差等第為1755.83元、極差等第為3113.93元，相關數據整理如下表6-16。並透過多重比較檢定確認其各項費用彼此之間在統計上有顯著之差異，如表6-17所示。

表6-16 異常操作行為產生額外費用表

費用 \ 等第	一般	稍差	極差	總平均
額外燃料費用	453.81	2267.71	3915.27	564.52
額外保修費用	285.55	1755.83	3113.93	375.5
額外總費用	739.36	4023.54	7029.2	940.01
人數(n)	199	9	2	210

註：單位（元/千公里）

表6-17 多重比較檢定（二）

等第 (i-j)	額外燃料費用			額外保修費用			額外總費用		
	平均差	標準誤	機率值	平均差	標準誤	機率值	平均差	標準誤	機率值
1-2	-1813.9	134.87	0.000*	-1470.28	92.54	0.000*	-3284.18	225.03	0.000*
1-3	-3461.46	281.24	0.000*	-2828.39	192.97	0.000*	-6289.85	469.26	0.000*
2-3	-1647.56	309.38	0.000*	-1358.10	212.27	0.000*	-3005.66	516.2	0.000*

註：1、2、3分別指一般、稍差、極差；*代表代表達 $\alpha=0.05$ 之顯著水準

依據上表之數據可以估計目前公司每千公里之額外燃料與保修費用如下：

- 燃料費用：

$$453.81*(456/480)+2267.71*(19/480)+3915.27*(5/480)=561.67 \text{ (元/千公里)}$$

- 保修費用：

$$285.55*(456/480)+1755.83*(19/480)+3113.93*(5/480)=373.21 \text{ (元/千公里)}$$

依據交通部公路客運營運量統計，該客運公司年營運里程為45,746,436公里，平均每月營運里程約為3,812,203公里，故可以推算出每月與每年估計之油耗與保養維修費用，計算如下：

- 月燃料費估計： $561.67 \times 3812 = 2,141,086$ （元）
- 年燃料費估計： $561.67 \times 45746 = 25,694,156$ （元）
- 月保修費估計： $373.21 \times 3812 = 1,422,676$ （元）
- 年保修費估計： $373.21 \times 45746 = 17,072,865$ （元）
- 月總費用估計：3,563,762（元）；年總費用估計：42,767,021（元）

● 駕駛員管理整合模式

將本章駕駛員群集與判別分析以及第五章所構建之肇事、交通違規、油耗與保養維修迴歸模式進行整合，並落實於實際之駕駛員管理層面，本研究提出駕駛員管理之流程，架構如圖6-4所示。如有客運公司現有一個或一群駕駛員，且尚未經由前述過程之分群與分析，即可利用下圖中之流程，分別將其判別等第以及估計出其可能發生之肇事或交通違規風險以及油耗與保養維修費用，即可了解駕駛員對於公司之影響或是貢獻程度，當然，肇事、交通違規風險與油耗、保養維修費用越小，茲然相對貢獻度越大，並可依公司之獎懲辦法進行獎勵或是懲處。

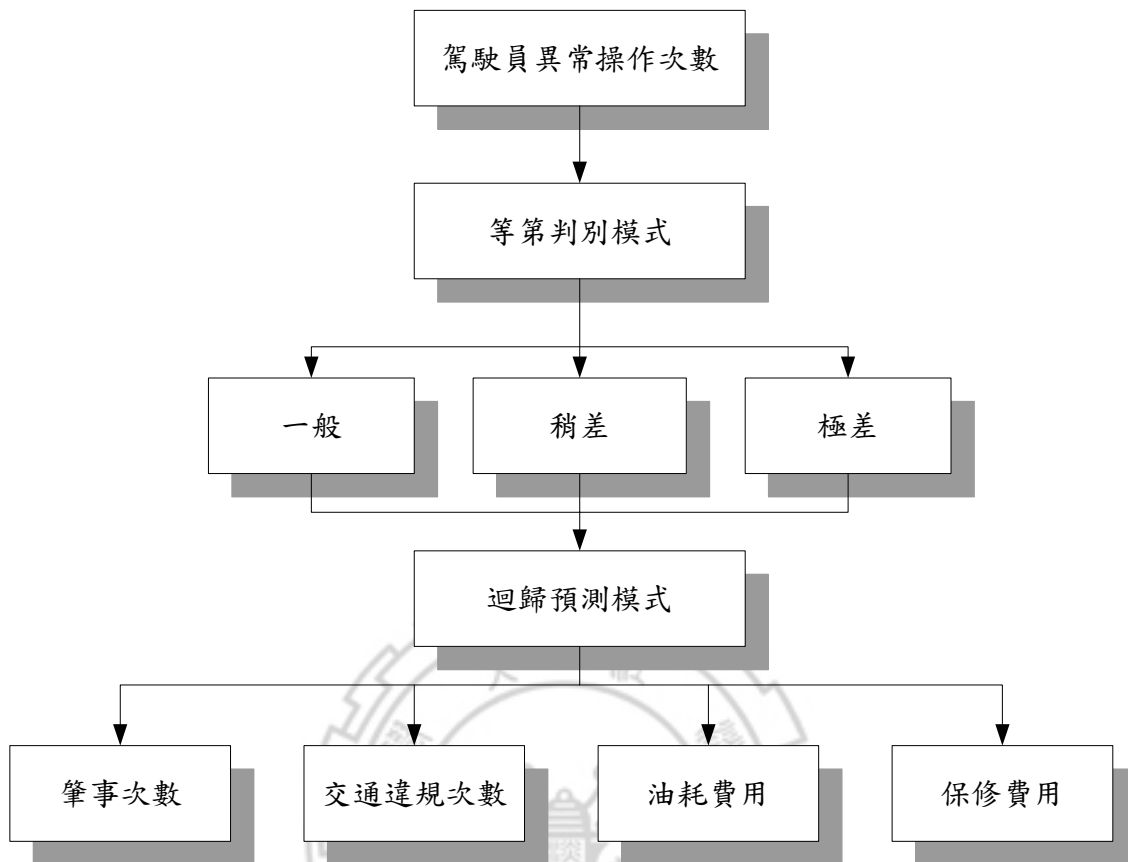


圖6-4 駕駛員管理整合流程圖

● 車輛油耗與保養維修之駕駛員責任歸屬

另外，在客運公司之駕駛員排班方面，一部車通常會有兩位正班駕駛員進行輪班，偶而會因為有駕駛員請假或是其他因素，會由代班駕駛替代，所以一部車輛可能會有兩位以上之駕駛員使用或操作。若車輛之月統計發現油耗過大或是車輛之保養維修費用增多，此時對於判斷哪個駕駛員所應負之責任較大是一個重要之管理問題。

數位式行車紀錄器之運用的確可以提高駕駛員責任歸屬之判斷，因為於該款 VDO FM200Plus 之數位行車紀錄器，每位駕駛員會配發一把身分辨識鑰匙，駕駛員欲駕駛該車輛時，需先使用該鑰匙起動數位式行車紀錄器使能行駛。所以一部車有不同駕駛員行駛操作，數位式行車紀錄器皆會依身分紀錄其各項異常操作項目與次數。

再配合本研究第五章所構建之迴歸模式，於模式中之變數即為影響油耗或是保

養維修之重要變數，故可將使用同一部車的各駕駛員異常操作次數代入模式中，即可估計出各駕駛員的油耗或是保修費用，雖可能會與實際數據有落差，可採用權重之方式，即可計算出各駕駛員異常操作所產生於該車之費用，進而由各駕駛員所產生費用大小判斷駕駛員之責任歸屬。流程如圖6-5。

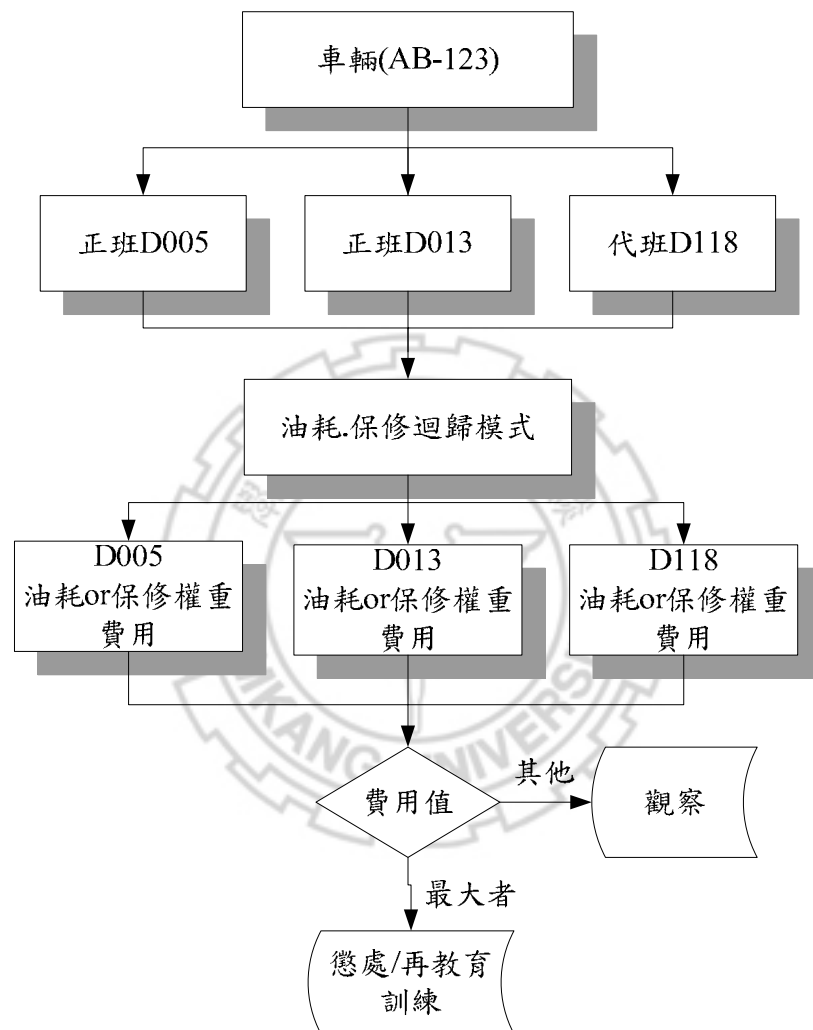


圖6-5 駕駛員責任歸屬判斷流程

6.4.3 駕駛員管理獎懲辦法

由於駕駛員難免會因為路況、天候等因素出現異常操作的情形，若要求達到完全零失誤無非過於嚴苛，但一味地消極處罰，卻在無形中會增加駕駛人工作壓力。本研究透過資料採礦技術，將該客運之 210 名駕駛員之異常操作情形進行分析，結果發現大多數之駕駛員之異常操作次數平均值相對於另外兩組較小許多，故可視為

其為一般普遍正常駕駛員之情形，故接下來可對於各等第之駕駛員進行獎懲。公司若能妥善訂出賞罰標準，嘉獎表現良好者具有鼓勵作用，同時對於應懲處的駕駛員也具規範效用。

（一）獎勵辦法

本研究所提出之駕駛員獎勵辦法是從一般等第之駕駛員中，選擇出各項異常操作之表現優於該等第中之平均值者，經由計算如 D050 駕駛員各項異常操作次數減去一般等第異常操作之平均值，再乘以迴歸模式之係數，即可得到各項異常操作次數所造成之額外費用，再將這些費用加總得到額外支出總金額，該名駕駛員 D050 所得到金額-601.14 元為負值，即若以一般等第平均值為標準，代表該名駕駛員較平均值能夠替公司節省 601.14 元，此即為該駕駛員對於公司之貢獻度。計算整理如表 6-18。

表 6-18 駕駛員貢獻度計算表

異常操作項目	(D050-一般等第平均值) ($E_D - E_0$)	迴歸係數 (β)	額外支出金額 $\beta*(E_D - E_0)$
急減速	(1.32-7.09)	4.917	-28.371
急加速	(0.44-3.51)	132.331	-406.256
電磁煞車操作異常	(0.44-1.10)	19.377	-12.789
怠速過久	(0.22-0.80)	92.807	-53.828
引擎轉速異常	(1.10-8.54)	16.847	-125.342
額外支出總金額(元/千公里)			-626.586
貢獻度 (= -額外支出總金額)			626.586

將一般等第之駕駛員計算其貢獻度，發現共有 128 名駕駛員之貢獻度為正值(即額外支出總金額為負值)，符合標準的這些駕駛員即視為優良表現駕駛員，整理如表 6-19。決策流程如圖 6-6。

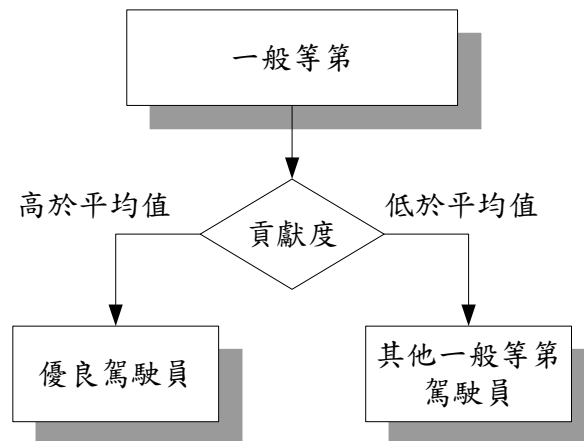


圖 6-6 優良駕駛員決策圖

表 6-19 優良表現駕駛員名單

優良表現駕駛員（貢獻度高於一般等第之平均值）					
駕駛員編號	貢獻度	駕駛員編號	貢獻度	駕駛員編號	貢獻度
D002	430.55	D066	150.90	D140	412.58
D003	623.68	D070	570.20	D141	13.62
D005	254.74	D073	475.42	D146	149.19
D006	298.79	D075	435.92	D147	403.82
D008	30.98	D076	609.09	D148	464.07
D009	166.09	D077	109.99	D149	286.61
D010	475.78	D078	448.80	D150	208.92
D012	164.54	D080	473.32	D151	59.94
D013	75.11	D081	488.56	D152	245.04
D014	293.26	D084	291.32	D154	389.12
D015	602.82	D085	235.29	D155	235.37
D018	393.90	D086	205.13	D156	314.92
D019	253.63	D087	302.79	D157	143.45
D020	381.86	D088	356.96	D159	117.32
D021	267.36	D090	526.12	D160	261.38
D022	156.58	D092	90.64	D162	66.11
D023	95.51	D093	227.40	D163	186.31

D024	317.06	D096	111.95	D166	206.14
D025	363.98	D098	331.45	D168	611.91
D026	138.42	D099	149.62	D169	168.00
D027	417.39	D100	125.36	D170	57.07
D028	613.10	D101	234.88	D171	61.97
D030	272.51	D103	53.71	D172	148.03
D031	244.66	D105	204.83	D173	207.12
D032	556.16	D110	150.98	D174	215.72
D033	512.48	D112	86.46	D175	163.60
D034	444.55	D115	132.54	D179	257.35
D036	459.51	D116	84.55	D180	265.47
D037	542.45	D117	58.57	D181	194.25
D038	551.64	D118	164.11	D182	44.51
D040	480.10	D119	268.45	D187	55.74
D041	655.66	D121	149.04	D188	135.93
D042	331.68	D123	358.27	D191	385.55
D043	411.06	D125	238.26	D195	336.03
D044	316.90	D129	431.42	D196	376.75
D046	595.13	D130	226.61	D199	271.32
D047	52.82	D132	173.78	D201	225.63
D049	386.65	D133	84.70	D203	239.16
D050	626.58	D134	257.90	D207	241.44
D051	626.50	D135	126.69	D208	258.95
D053	534.20	D136	374.29	D209	524.96
D054	249.40	D138	222.22	D210	440.30
D056	90.15	D139	214.69		
總計：128 人					

基於駕駛員管理相關文獻回顧可知，公司之管理制度對於駕駛員有直接正面之影響，亦會影響駕駛員之情緒反應與駕駛行為，客運公司應考量駕駛員之心理感受，給予適當之獎勵，以激勵駕駛員士氣，並提升效率與減少異常操作行為之發生，故

優良駕駛員獎勵辦法案例如下：

- 獎勵期間：每一次獎勵以一個月為基準。
- 參與辦法資格：
 - 駕駛員須以獎勵計算期間內正常出勤行駛。
 - 駕駛員如有期間內發生有則肇事或有責故障，該獎勵則不列入。
- 獎金：
 - 於獎勵期間內符合優良駕駛員標準者，予以獎金 1500 元以茲鼓勵（參考該客運公司之假設金額），並記嘉獎一次。
 - 優良駕駛員前五名者，再加發 1000 元獎金（假設金額）。
- 獎金費用估計：以 128 名符合標準之優良駕駛員計算，獎金費用每月約需 $128*1500+5*1000=197,000$ 元。若以公司目前之所有駕駛員推算優良駕駛員將有 292 人，故分別可計算一個月與一年所需之獎金費用如下：
 - 月獎金費用： $292*1500+5*1000=443,000$ （元）
 - 年獎金費用： $443,000*12=5,316,000$ （元）

上述獎勵辦法前五名者加發獎金是為了刺激已經於優良駕駛員標準內之駕駛員，使其產生良性之競爭。至於於一般等第中未達優良標準的其他駕駛員有 71 人則不獎勵亦不懲處，整理如表 6-20。

表 6-20 不獎勵或懲處之駕駛員名單

一般等第其他駕駛員（貢獻度低於一般等第之平均值）	
駕駛員編號	總計
D061、D068、D060、D057、D072、D062、D198、D059、D165、 D035、D095、D189、D004、D063、D186、D016、D064、D045、 D083、D185、D177、D058、D197、D069、D048、D079、D071、 D114、D065、D109、D202、D120、D128、D074、D200、D111、 D183、D011、D082、D192、D104、D137、D094、D206、D194、 D204、D091、D161、D113、D106、D190、D184、D127、D122、 D102、D193、D124、D089、D167、D001、D126、D158、D007、 D097、D178、D205、D107、D164、D142、D176、D055	71 人

(二) 懲處辦法

對於稍差與極差等第之駕駛員應進行適當懲處，以對於駕駛行為或習慣不良而產生較多較嚴重異常操作情形之駕駛員有所警惕與規範。經由計算如 D039 駕駛員各項異常操作次數減去一般等第異常操作之平均值，再乘以迴歸模式之係數，即可得到各項異常操作次數所造成之額外費用，再將這些費用加總得到額外支出總金額，該名駕駛員 D039 所產生之額外總費用金額 1547.55 元（千公里）。由此可知，稍差或極差等第所造成之額外費用會較一般等第的額外費用增加許多，顯見異常操作行為將對於客運公司造成相當嚴重之負擔。計算過程整理如表 6-21 與 6-22。

表 6-21 D039 駕駛員之額外費用計算

異常操作項目	(D039-一般等第平均值) ($E_D - E_0$)	迴歸係數 (β)	額外支出金額 $\beta*(E_D - E_0)$
急減速	(69.95-7.09)	4.917	309.083
急加速	(5.05-3.51)	132.331	203.790
電磁煞車操作異常	(0.95-1.10)	19.377	-2.907
怠速過久	(1.19-0.80)	92.807	36.195
引擎轉速異常	(66.76-8.54)	16.847	980.832
額外支出總金額(元/千公里)			1526.993

表 6-22 稍差與極差等第駕駛員產生之額外費用

等第	駕駛員編號	額外支出總金額 (元/千公里)	人數
稍差	D017	3740.76	9
	D029	2119.63	
	D039	1526.10	
	D052	587.08	
	D067	2335.05	
	D108	6043.12	
	D131	1756.15	
	D143	5883.40	
	D153	5571.60	
極差	D144	5878.04	2
	D145	6702.82	

由稍差與極差等第中各駕駛員異常操作行為所產生之額外總費用可知，以極差等第之駕駛員所產生之額外費用最多，此為相當嚴重之情形，故提出懲處辦法案例如下：

- 懲處期間：以每月為基準。
- 極差等第懲處：由於該等第駕駛員其異常操作之次數最多，對於公司所造成之額外費用較為嚴重，基於避免增加駕駛員之心理壓力，第一次落於該等第者僅以口頭警告，並令其接受再教育訓練。而累犯者（兩次以上），每次懲罰金額為1500元（假設金額），並記警告一次，罰金從駕駛員薪資中扣除，並令駕駛員馬上接受每月定期之行車安全與駕駛觀念再教育訓練，若情節重大或是連續落於該等第完全無改善者，應予免職。
- 稍差等第懲處：對於第一次落於該等第之駕駛員，予以口頭警告。而兩次以上者，每次記警告一次，令其接受每月之行車安全與駕駛觀念再教育訓練。
- 管理單位監控追蹤：對於稍差與極差等第之駕駛員由於其異常操作情況會造成車輛一定程度之影響，故管理單位需要對於這些駕駛員與其駕駛之車輛進行觀察與追蹤，如下列做法。

- 急加速、怠速過久與引擎轉速異常數據過高之駕駛員，應由所屬車場管理人員控管該員所駕車輛之油耗狀況。
- 急減速異常數據過高之駕駛員，應由所屬車場管理人員控管該員所駕車輛煞車來令片汰換之頻率，另由相關稽核人員針對該員跟車查核是否行車安全間距不足。
- 檔位異常數據過高之駕駛員，應由所屬車場管理人員控管該員所駕車輛之離合器組使用里程是否合理正常。
- 電磁煞車駛操作異常數據過高之駕駛員，應由所屬車場管理人員控管該員所駕車輛之電瓶與電磁煞車控制盒、線圈更換頻率是否正常。

（三）數位式行車紀錄器鑰匙使用管理規範

- 駕駛員每人皆配置一支個人專屬數位式行車紀錄器鑰匙，規定勿借他人使用以免個人資料不正確，若有該情事發生，經查證後記警告一次，且有嚴重異常操作之情況，兩者皆須受懲處。
- 駕駛員自行保管數位式行車紀錄器鑰匙，如有遺失依規定賠付。
- 代班駕駛、車廠保養及原廠維修保養一律使用備用鑰匙。

6.4.4 駕駛員素質提升之效益

有了前述之駕駛員評等以及獎懲辦法，以及對於稍差與極差等第之駕駛員進行再教育訓練，以上這些都是駕駛員之管理方式，目的就是為了希望駕駛員能夠符合客運公司的期望需求以及增進行車安全，以期至少能夠在一般普遍的水準之內，避免有駕駛員有嚴重的素質差異，故我們可以預期推算若稍差與極差等第之駕駛員素質皆提升至一般等第之水準，勢必可以減少一些由過多異常操作為所造成的肇事或交通違規次數以及所產生之額外費用，此即為駕駛員素質提升之效益。依照該客運公司現有之駕駛員人數 480 人估計，並以等比例推算各等第人數。費用計算方式如表 6-23 與 6-24。在行車安全效益提升方面，一個月肇事與交通違規減少次數計算如下：

- 月減少肇事次數：
 $[0.0716 \times (19/480) + 0.1592 \times (5/480)] \times 3812 = 13 \text{ (次)}$

- 月減少交通違規風險：

$$[0.0121*(19/480)+0.0146*(5/480)]*3812 = 2 \text{ (次)}$$

表 6-23 駕駛員素質提升節省金額計算（稍差等第）

異常操作項目	(稍差-一般等第平均值) ($B_1 - N_0$)	迴歸係數 (β)	節省金額 $\beta*(B_1 - N_0)$
急減速	(93.96-7.09)	4.917	427.1398
急加速	(16.17-3.51)	132.331	1675.31
電磁煞車操作異常	(5.74-1.10)	19.377	89.90928
怠速過久	(2.54-0.80)	92.807	161.4842
引擎轉速異常	(63.82-8.54)	16.847	931.3022
素質提升節省總金額(元/千公里)			3285.146

表 6-24 駕駛員素質提升節省金額計算（極差等第）

異常操作項目	(極差-一般等第平均值) ($B_2 - N_0$)	迴歸係數 (β)	節省金額 $\beta*(B_2 - N_0)$
急減速	(246.08-7.09)	4.917	1175.114
急加速	(25.43-3.51)	132.331	2900.696
電磁煞車操作異常	(6.85-1.10)	19.377	111.4178
怠速過久	(9.91-0.80)	92.807	845.4718
引擎轉速異常	(83.19-8.54)	16.847	1257.629
素質提升節省總金額(元/千公里)			6290.327

經由上述之計算，且依照該客運公司現有駕駛員人數估計，可以得到以教育訓練或管理方式提昇駕駛員素質（等第）所能節省之一個月及一年之效益，推估過程如下：

- 月節省金額：

$$[3285.146*(19/480)+6290.327*(5/480)]*3812 = 745,480 \text{ (元)}$$

- 月節省金額（扣除獎勵獎金）：

$$745,480 - 443,000 = 302,480 \text{ (元)}$$

- 年節省金額：

$$[3285.146 \times (19/480) + 6290.327 \times (5/480)] \times 45746 = 8,946,146 \text{ (元)}$$

- 年節省金額（扣除獎勵獎金）：

$$8,946,146 - 5,316,000 = 3,630,146 \text{ (元)}$$

由此可見提升駕駛員素質，除了可以增加乘客的滿意度以外，還可替客運公司節省油耗與保養維修的費用，在扣除獎勵獎金後仍有不少的節省金額。因此可以得知，對於駕駛員進行教育訓練或是以合適的管理方式提昇駕駛員之水準素質是很重要的，故公司方面可以依照所節省之費用提撥一定額度作為獎勵獎金以及駕駛員教育訓練與新進駕駛員訓練，直接提昇新進人員至一定水準，或增購數位式行車紀錄器裝設於尚未裝設之車輛，掌握充分之駕駛員行車資料數據，才能對於駕駛員進行適當的管理。

6.5 小結

對於駕駛員進行分群與管理，本研究採用資料採礦技術中之群集分析方法，對於 210 名駕駛員樣本進行分群，分群結果得到三個群集，分別依其特性命名為一般、稍差以及極差群集，再利用判別分析，構建出判別模式，始能採用駕駛員判別模式直接由駕駛員之異常操作次數判斷其落於哪個群集。並於後續將該三個群集作為駕駛員評鑑等第之依據，之後即稱為一般、稍差與極差等第，駕駛員於一般等第有 199 人、稍差等第有 9 人、極差等第有 2 人，若有未知其應屬何種等第之駕駛員，即可利用本研究所構建的駕駛員評鑑等第判別模式與以判斷並歸類。

配合第五章所構建迴歸模式之係數，對於各等第駕駛員異常操作行為所產生之平均肇事風險、交通違規風險、油耗以及保養維修費用進行估計，並確定其各等第之間有統計上之顯著差異，再將判別模式與迴歸模式之運用整合，提出可以落實於駕駛員管理層面之方法整合流程，並解決一部車多位駕駛之責任歸屬問題。後續提出獎勵與懲處辦法，將一般等第中表現優於一般等第之平均值者選出為優良駕駛員，因為其貢獻度大於平均值，並對優良駕駛員給予獎金獎勵。稍差與極差等第之駕駛員主要著重於再教育訓練，規定落於該等第之駕駛員必須參加該月之再教育訓練，以提升行車安全與駕駛觀念。最後並依照該客運公司現有之駕駛員約 480 人，推估以再教育訓練或是相關管理方式提昇駕駛員素質水準之效益，結果得到一個月可以減少 13 次肇事、2 次交通違規，以及公司可節省 745,480 元之額外油耗與保養維修費用，總計一年可節省 8,946,146 元的額外費用，在扣除獎勵獎金後一年尚有

3,630,146 元之節省費用，此即為駕駛員素質提升所能帶來之效益，公司可利用節省下來之費用提撥一定比例供再教育訓練用與新進駕駛員訓練，直接提昇新進人員素質至一定水準，或是增購數位式行車紀錄器進行車輛全面裝設，以充分掌握所有駕駛員之駕駛行為資料，始能有效的管理。



第七章 結論與建議

本研究運用資料採礦技術，以數位式行車紀錄器紀錄駕駛員異常操作情形為主體，定義駕駛員異常操作項目與設定門檻值，以萃取出資料庫中未知且具有意義之資訊，並構建資料採礦模式，探討駕駛員異常操作行為對於行車安全（肇事、交通違規風險）以及行車油耗與保養維修費用之關係，透過蒐集國內某客運公司 61 部車輛之數位式行車紀錄器資料，以及駕駛員肇事與交通違規紀錄以及車輛油耗與保養維修費用，並經由資料採礦程序前置處理作業，將資料進行適當篩選與整理，構建駕駛員異常操作行為之關係模式，即利用輸入經由資料庫採礦所得到之各異常操作項目作為自變數，構建肇事、交通違規、油耗與保養維修費用之多元迴歸模式。對於駕駛員進行群集分析，使用迴歸模式中重要異常操作項目作為 K-Means 群集分析之分群變數，結果將 210 名駕駛員樣本分成三個群集，並使用判別分析，構建三個群集之判別模式，並依據各群集命名為一般、稍差與極差等第，以進行後續之駕駛員管理與獎懲，並得到駕駛員管理提升駕駛員素質後所得到之效益。

依據前述各章節之分析與探討，本研究綜合提出下列之結論與建議：

7.1 結論

（一）駕駛員異常操作監控項目定義與門檻值設定

本研究依據相關駕駛員異常操作文獻之研究與客運公司營運需求以及客運車輛原廠建議設定值，定義出駕駛員異常操作之項目以及門檻值，主要內容分成四大項目為行車安全、燃油消耗、機件磨損以及行車品質，細部指標包含急減速、急加速、違規超速、電磁煞車操作異常、怠速過久、引擎轉速異常、檔位異常以及車燈使用異常等指標項目，並依據前述指標項目與設定門檻值對於資料庫進行資料採礦，去除不屬研究範圍之資料以縮減資料量，始能夠從龐大的資料庫中篩選出有意義且可解釋的資訊，以利後續資料採礦模式之建立。

（二）駕駛員異常操作行為關係模式部份

基於前述資料採礦所得到資料庫中具有意義之駕駛員異常操作資訊，本研究運用多元迴歸分析構建資料採礦模式，探討駕駛員異常操作行為對於行車安全，油耗與保養維修費用之關係，以下為本研究此部份之結論：

1. 行車安全方面

- 影響肇事次數之異常操作行為是急減速與超速，若駕駛員未發生任何異常操作行為，迴歸模式中之常數項即為每千公里駕駛員可能發生肇事次數為 0.0056 次，而迴歸式中之係數即可代表異常操作行為發生後所產生之效果，若急減速每千公里發生一次，肇事發生次數則增加 0.0005 次，超速每千公里發生一次，肇事發生次數則增加 0.0009 次，由此可以得知超速對於發生肇事之影響比急減速來的大，此結果與國內外相關肇事原因研究調查是能夠吻合的。
- 影響交通違規次數之異常操作行為則是超速與電磁煞車操作異常，常數項為 0.0049 代表未發生任何異常操作行為每千公里可能產生交通違規之次數，超速在交通違規之受罰原因中為次多之項目，最多為未依車道行駛應無相關之異常操作監控項目可衡量，故得到超速與交通違規有關聯是合理的，代表平時容易發生超速之駕駛員，亦較容易受到交通違規罰單，且每千公里增加一次之超速行為，交通違規受罰之次數會增加 0.00013 次，而每千公里增加一次電磁煞車操作異常行為，交通違規次數會增加 0.00078 次。交通違規迴歸模式解釋能力 R^2 值尚可，可能因為超速行為無法解釋所有交通違規之情形，以及超速行為發生不一定會受到違規舉發所致。

2. 油耗與保養維修方面

- 影響油耗費用之駕駛員異常操作行為是急加速、怠速過久以及引擎轉速異常，迴歸模式中常數項代表駕駛員於一般正常行駛無異常操作行為下每千公里會消耗 4605 元的燃料費，各項係數代表駕駛員每增加一次異常操作行為會產生額外的燃料費用，每千公里增加一次急加速會增加 79.53 元、每千公里增加一次怠速過久會增加 92.81 元、每千公里增加一次引擎轉速異常會增加 11.7 元，由前述數值即可知各項異常操作行為影響之程度與重要性，且急加速行為是由駕駛員油門控制，油門越是急踩加速度越大，引擎轉速也會增高，亦越耗油，而怠速過久將導致燃料浪費無效率。
- 影響保養維修費用之駕駛員異常操作行為則是急減速、急加速、電磁煞車操作異常以及引擎轉速異常，迴歸式中之常數項代表駕駛員於一般正常行駛無異常操作行為下每千公里會有 586.9 元的保養維修費，各項係數代表駕駛員每增加一次異常操作行為會產生額外的保養維修費用，每千公里增加一次急減速會增

加 4.917 元、每千公里增加一次急加速會增加 52.805 元、每千公里增加一次電磁煞車操作異常會增加 19.377 元、每千公里增加一次引擎轉速異常會增加 5.151 元，由前述數值可知各項變數對於保養維修費用之影響程度與重要性，各項變數項目亦皆為會影響保養維修費用，符合先驗知識。

（三）駕駛員分群與管理部份

- 運用群集分析建立駕駛員分群之資料採礦模式，對於 210 名駕駛員樣本進行 K-Means 分析法分群，分群結果得到三個群集依其特性命名為一般等第、稍差等第以及極差等第，並得到一般等第有 199 人、稍差等第有 9 人、極差等第有 2 人。並運用判別分析，構建各群集之判別模式，可對於未知等第之駕駛員直接判別歸類。
- 配合所構建迴歸模式之係數，對於各等第駕駛員異常操作行為所產生之平均肇事風險、交通違規風險、油耗以及保養維修費用進行估計，並確定其各等第之間有統計上之顯著差異，再將判別模式與迴歸模式之運用整合，提出可以落實於駕駛員管理層面之方法整合流程，並解決一部車多位駕駛之責任歸屬問題。
- 提出獎勵與懲處管理辦法案例，將一般等第中表現優於一般等第之平均值者選出為優良駕駛員，因為其貢獻度大於平均值，並對優良駕駛員給予獎金以及記功嘉獎等獎勵。稍差與極差等第之駕駛員主要著重於再教育訓練，規定落於該等第之駕駛員必須參加公司該月之再教育訓練，以提升行車安全與駕駛觀念。若有持續情節重大者，應予免職處分。
- 依照該客運公司現有之駕駛員約 480 人，以等比例放大估計各等第人數，分別為一般等第 456 人，稍差等第 19 人，極差等第 5 人，推估以再教育訓練或是相關管理方式提昇駕駛員素質水準之效益，結果得到一個月可以減少 13 次肇事、2 次交通違規，以及公司可節省 745,480 元之額外油耗與保養維修費用，總計一年可節省 8,946,146 元的額外費用，在扣除獎勵獎金後一年尚有 3,630,146 元之節省費用，此即為駕駛員素質提升所能帶來之效益。
- 該客運公司應可利用產生效益所節省之費用提撥一定比例供再教育訓練用與新進駕駛員訓練，將直接提昇新進人員素質至一定水準，或是增購數位式行車記錄器進行車輛進行全面裝設，以充分掌握所有駕駛員之駕駛行為資料，始能搭配相關管理措施並施行有效的管理。

- 本研究依據相關之結果可以得知，將數位式行車紀錄器資料進行資料採礦，尋求其中重要之資訊，並搭配運用於實際駕駛員管理層面，對於客運公司之營運成本節省以及增進行車安全將有所助益。

7.2 建議

- 本研究所蒐集之資料主要為數位式行車紀錄器資料與相關駕駛員肇事、交通違規紀錄、油耗與保養維修費用資料，在相關行車費用部份由於蒐集資料對象緣故未將輪胎成本費用納入，亦未納入其他成本項目如行車附支或修車附支等費用，故建議後續研究可將各項費用擴大納入研究範圍，應可使研究層面更加完備。
- 本研究所蒐集資料對象為國內某客運業者，路線為國道客運路線，其所採用之數位式行車紀錄器為 Siemens VDO FM200-Plus 型，該數位式行車紀錄器主要之特色即為其可事先設定異常操作監控項目與門檻值，故若要直接運用本研究之模式與結果，所設定之監控項目必須相同，門檻值不可差異過大，由於該路線為國道客運路線，市區公車或其他省道客運業之門檻值設定應不同，因市區公車或省道客運車輛之行駛速率較低，且常會有減速之行為，故不宜直接套用。建議後續研究可朝市區公車等方向進行研究，探討市區公車之門檻值訂定與駕駛員異常操作之影響關係，或是運用於貨運車輛（若有裝設數位式行車紀錄器）駕駛員異常操作行為之探討。
- 由於本研究主要是利用資料採礦之架構進行數位式行車紀錄器駕駛員異常操作行為對於行車安全與相關費用層面之影響關係探討，故後續駕駛員管理方面亦是著力於駕駛員異常操作之管理與獎懲部份，對於其他的駕駛員管理如服務態度、行車品質或是個人品德操守等未納入考量，故建議未來客運公司可採用本研究對於駕駛員異常操作行為之管理架構，整合其他層面之駕駛員管理，以作為實際應用於客運公司之駕駛員管理制度。
- 後續研究可嘗試其他資料採礦技術，可探討其他資料採礦技術是否對於更龐大的數位式行車紀錄器資料庫，能以更有效率的進行資料處理與分析模式構建，可使用如類神經網路、決策樹或關聯規則等其他資料採礦方法，進行分類或是預測之研究。

- 根據高速公路肇事原因調查顯示，除了超速，未保持安全車距是肇事的主因之一，且根據本研究對於該客運公司之肇事原因統計，追撞事故亦占肇事原因中有 40% 之比例，但礙於需搭配前車距離感應雷達等裝置，且需配合數位式行車紀錄器之訊號接頭之擴充性，目前僅能由駕駛員急減速之情形大致推斷是否未保持安全車距，建議未來客運業者可增加未保持安全車距之監控項目，可使數位式行車紀錄器對於行車安全之功能發揮的更為完備。
- 本研究對於駕駛員所產生之油耗之估計是以該客運公司資訊管理部門之統計資料並配合構建模式進行分析，建議未來可採用更精確之駕駛員油耗估計，應可以數位式行車紀錄器搭配數位式之油量感知裝置，以取得各旅次更為詳盡之駕駛員油耗資料。



參考文獻

一、英文部分

1. B. W. Ang and T. F. Fwa , “A study on the fuel-consumption characteristics of public buses Energy,” Volume 14, Issue 12, Pages 797-803, December 1989.
2. Decicco, J. and M. Ross, “Recent advances in automotive technology and the cost-effectiveness of fuel economy improvement, ” Transportation Research Part D, Vol. 1, No. 2, pp. 79-96, 1996.
3. Giudici, P. and Passerone, G., “Data mining of association structures to model consumer behavior,” Computational Statistics and Data Analysis 38(4), 533-541, 2002.
4. Hooker, J. N., ”Optimal driving for single-vehicle fuel economy, ”Transportation Research Part A, Vol. 22, No. 3, pp. 183-201, 1988.
5. Mascha van der Voort, Mark S. Dougherty and Martin van Maarseveen, “A prototype fuel-efficiency support tool,” Transportation Research Part C, Vol. 9, pp. 279-296, 2001.
6. Quimby A., Maycock G., Palmer C. and Butteress S., “The factors that influence a driver's choice of speed—a questionnaire study,” Report 325, TRL, Crowthorne, 1999.
7. Kuhnert P. M., Do K.A. and McClure R., “Combining non-parametric models with logistic regression: an application to motor vehicle injury data,” Computational Statistics and Data Analysis 34(3), 371-386, 2000.
8. Rygielski C., Wang, J. C. and Yen, D.C., “Data mining techniques for customer relationship management,” Technology in Society 24(4), 483-502, 2002.
9. Stanley M. Gully, David J. Whitney and Fred E. Vanosdall, “Prediction of police officers' traffic accident involvement using behavioral observations,” Accident Analysis & Prevention, Volume 27, Issue 3, Pages 355-362, June 1995.
10. Sumit Anand, Nir Keren, Marietta J. Tretter, Yanjun Wang, T. Michael O'Connor and M. Sam Mannan , “Harnessing data mining to explore incident databases, ”Journal of Hazardous Materials, Volume 130, Issues 1-2, Pages 33-41,

17 March 2006.

11. Taylor, M.C., Lynam, D.A. and Baruya, A., "The effects of driver's speed on the frequency of road accidents," TRL Report 421, prepared for the Road Safety Division, DETR, 2000.
12. Ulla E. and Lars K. Thomsen, "Safety effects of speed reducing measures in danish residential areas," Accident Analysis & Prevention, Volume 24, Issue 1, Pages 17-28, February 1992.
13. Vasiliev, L. L., V. S. Burak, A. G. Kulakov, D. A. Mishikinis, and P. B. Bohan, "Heat storage device for pre-heating internal combustion engines at start-up," International Journal Thermodynamic Science, Vol. 38, pp. 98-104, 1999
14. Waters, M. H. L. and I. B. Laker, "Research on Fuel Conservation for Cars," Report No. 921, Transportation Research Laboratory, Crowthorne, England, 1980.
15. U. Fayyad, G. Piatetsky-Shapiro, and P. Smyth, "From Data Mining to Knowledge Discovery in Databases," American Association for Artificial Intelligence, 1996.
16. "Reducing Traffic Injuries Resulting From Excess and Inappropriate Speed," European Transport Safety Council Rue du Cornet 34 B-1040 Brussels, 1995.

二、中文部份

1. 謝邦昌,「多變量分析(3)－判別分析法」,中國統計通訊,第九期,頁 28-38,民國 87 年 9 月。
2. 林震岩,「多變量分析－SPSS 的操作與應用」,智勝圖書,民國 95 年 1 月。
3. 邱浩政,「社會與行為科學的量化研究與統計分析」,五南圖書,民國 93 年 5 月。
4. 張建邦,「多變量分析」,三民書局,民國 86 年 7 月。
5. 黃俊英,「多變量分析」,華泰文化,民國 90 年 1 月。
6. 王鴻儒,「SQL Server 2005 資料庫設計建置管理實務」,金禾圖書,民國 94 年 12 月。
7. 尹相志,「SQL 2000 Analysis Service 資料採礦服務」,維科圖書,民國 92 年。
8. 胡百敬、尹相志,「SQL Server 商業智慧聖經」,學貫行銷,民國 93 年。
9. 「車輛油耗指南」,經濟部能源局,民國 93 年 2 月。

10. 「高速公路年報」，國道高速公路局，民國 93 年。
11. 鄭子珊、林維信等，「數位式行車紀錄器功能技術規範建立與示範應用之研究」，交通部運輸研究所與財團法人中華顧問工程司合作辦理，民國 93 年 4 月。
12. 「汽車客運業統一會計科目」，交通部運輸研究所，民國 86 年 2 月。
13. 張杏珍、邱穀等，「公路車輛行車成本調查」，交通部運輸研究所，民國 89 年 1 月。
14. 林繼國、顏漏有等，「汽車客運業路線別成本計算制度」，交通部運輸研究所，民國 88 年 9 月。
15. 「顧客關係管理與資料採礦」，EC 研究報告，ComerceNet Taiwan，民國 91 年 2 月。
16. 林雅俐、劉正華，「駕駛行為與肇事概率之關聯性研究」，工業工程學刊，第 16 卷第 1 期，頁 161-172，民國 88 年 1 月。
17. 楊東麟、洪明傳，「資料探勘在資料倉儲的應用」，資訊與教育雜誌，第 84 期，民國 90 年。
18. 倪達洸，「資料採礦方法論」，RUN!PC，第 104 期，民國 91 年 9 月。
19. 黃木泳，「資料採礦技術及發展趨勢」，RUN!PC，第 104 期，民國 91 年 9 月。
20. 林大煜，「道路交通肇事資料分析方法之探討」，運輸計劃季刊，第 8 卷第 3 期，頁 379-413，民國 68 年。
21. 謝邦昌，「資料探勘技術發展趨勢」，軟體與應用，民國 91 年 9 月。
22. 謝邦昌，「多變量分析（4）—集群分析法」，中國統計通訊，第 9 卷第 10 期，頁 31-42，民國 87 年 10 月。
23. 廖少威、張文亮，「關渡濕地底泥性狀對植被分類判別函數之建立」，水土保持學報，第 34 期，頁 209-226，民國 91 年 9 月。
24. 吳冠宏、郭廣洋、吳信宏，「應用資料挖掘於交通事故資料分析」，警學叢刊，第 36 期，頁 131-146，民國 94 年 9 月。
25. 王立心，「集群分析的理論與應用」，高師教育研究，第 6 期，頁 91-116，民國 87 年 6 月。
26. 陳俊傑，「應用資料探勘技術於道路交通事故資料分析之研究」，中央警察大學資訊管理研究所碩士論文，民國 91 年 6 月。
27. 許哲瑋，「資料挖掘與統計方法應用於資料庫行銷之實證研究-以美妝保養品業為例」，台北大學企管研究所碩士論文，民國 91 年 6 月。
28. 楊進彥，「汽車行車記錄器應用於交通管理之研究」，淡江大學運輸科學研究所

碩士論文，民國 89 年 6 月。

29. 薛如珊，「使用自組織映射網路進行資料群集和資訊樣型採擷的資料探勘法」，臺灣大學工業工程學研究所碩士論文，民國 90 年 6 月。
30. 張季倫，「公路客運行車監控之研訂及駕駛與車輛資料庫管理系統之研發--數位式行車紀錄器之應用」，交通大學交通運輸研究所碩士論文，民國 90 年 6 月。
31. 廖景元，「高速公路巡邏力對行車安全之影響研究」，交通大學管理研究所碩士論文，民國 69 年 6 月。
32. 黃聖源，「駕駛行為因素對市區幹道系統服務水準影響分析之初探」，台灣大學土木工程研究所碩士論文，民國 77 年 6 月。
33. 吳偉碩，「台南環線高快速公路肇事特性分析與安全改善之研究」，交通大學交通運輸研究所碩士論文，民國 90 年 6 月。
34. 尹維龍，「應用駕駛行為量表探討偏差駕駛行為與事故傾向關係之研究」，交通大學運輸科技與管理研究所碩士論文，民國 93 年 6 月。
35. 蕭重威，「車輛行車型態駕駛模擬」，清華大學動力機械研究所碩士論文，民國 83 年 6 月。
36. 胡俊生，「剎車鼓內部溫度和熱應力分析及性能改善」，逢甲大學機械工程研究所碩士論文，民國 85 年 6 月。
37. 戚培芳，「中山高速公路肇事分析模式之研究」，交通大學運輸研究所碩士論文，民國 86 年 6 月。
38. 劉建良，「公車駕駛員之薪資滿足感與工作壓力、工作滿足與服務品質之關係探討」，淡江大學運輸科學研究所碩士論文，民國 89 年 6 月。
39. 陳國樑，「公路客運公司管理制度與駕駛員行為特性關係之研究」，逢甲大學交通工程管理研究所碩士論文，民國 93 年 6 月。