

# 影響汽車購買的汽車屬性之研究

## THE ATTRIBUTES OF AUTOMOBILE AFFECT AUTOMOBILE PURCHASING BEHAVIOR

蘇高玄\*

高苑科技大學行銷與流通管理系講師

蔡佳蓉

高苑科技大學行銷與流通管理系學生

**Kao-Hsuan Su**

*Instructor, Department of Marketing and Distribution Management,  
Kao Yuan University*

**Chia-Jung Tsai**

*Student, Department of Marketing and Distribution Management,  
Kao Yuan University*

### 摘要

本研究欲探討從消費者購買汽車決定重要屬性的大數據資料中，利用 Relief-F 算法篩選出關鍵的重要汽車屬性；利用此汽車屬性製作成問卷加以發放，再針對問卷回收後執行統計逐步迴歸分析，得到符合國內消費者真正對購買車的重要屬性。研究結果從二資料庫共有 34 個特徵，經過 Relief-F 特徵過濾與逐步迴歸法篩選後，最後萃取出 9 個消費者心中認為最重要的汽車屬性。即可再進一步利用 ANOVA 檢定出顧客基本的特性與購買的重要因素之關係，即可分析出汽車市場的區隔。市場區隔有了清楚的定義，各家車廠就能善用資源針對自己的市場去分析，同時也能定義出直接競爭者、潛在競爭者及替代品競爭者，形成 SWOT 分析的基礎。

**關鍵字：**大數據、汽車屬性、逐步迴歸、購買行為、Relief-F

---

\*通訊作者，地址：高雄市路竹區中山路 1821 號，電話：0922-585178  
E-mail：t70027@cc.kyu.edu.tw

## ABSTRACT

In this study, we use Relief-F algorithm to filter the main attributes from the big data of consumers purchase automobile, and then use this key attributes to make into questionnaires. Then analysis the results by stepwise regression method, we got the really the 9 attributes of the consumers who cared in automobile purchase. Finally, we use ANOVA method to determine the basic characteristics of the customer and the purchase of the relationship between the important attributes, we can analysis the segment of the automobile market. Automobile marketing segments have a clear definition, each automobile-manufacturers can focus all resources to develop their marketing, to beat potential competitors.

**Keywords:** Big Data, Automobile Attributes, Stepwise Regression, Purchase Behavior, Relief-F

## 壹、研究動機與研究問題

### 一、研究背景

2012 年 3 月，美國在《大數據的研究和發展計畫》（White House, 2012）中把大數據的研發應用從以前的商業行為上升到國家戰略部署，以提高從龐大而複雜的數據中提取知識的能力，幫助解決一些國家最緊迫的挑戰。

在 2013 年所產生的數據量，是盤古開天至 2012 年所有數據的總和。接下來的每一年又比前一年更多出 3 倍以上，這就是巨量資料。巨量資料的精神在於，如何從看似無相關的大量資料，進行分析後獲得對於企業營運、服務等，提供不同面向的新見解，以做為未來提升的參考。長期以來，廣告界以嚴密的市場調查找出 TA（目標族群）的數據、資料，發展創意與精準廣告行銷方案，為廣告主創造利益（陳淑芬，2014）。

巨量資料的核心重點在於預測。一般看做是資訊工程「人工智慧」的分支「機器學習」（machine learning）的一部分。不過，巨量資料並不是要「教」機器人像人類一樣「思考」，而是經由計算大量的資料，以此推斷出「機率」。

近期的汽車大戰很精采，在國內汽車市場銷量受限的情況下，市占率就變成是兵家必爭的數字。台灣的汽車市場從高峰期的一年 50 餘萬輛，近年來已下滑到 38 萬輛左右，而國產車與進口車的比例，從最高曾拉開到 9：1，到現在已不到 3：1。由於進口車（含雙 B）不斷推出小車並拉低售價，從圖 1 上可以看出進口車的銷售量突破十萬輛，市占率逼近三成，從 2009~2013 逐年增加，顯示進口車的競爭力已明顯威脅到國產車市場的版圖。

## 二、研究動機與目的

隨著汽車產業的不斷發展，汽車相關的研究規模與數據也相對擴大，其中所蘊含的資訊量不斷增加。本研究想從不斷增加的數據中，利用特徵選擇（Feature Selection）方法來當做大數據分析的第一步，從中萃取出重要的汽車屬性（Automobile Attributes），其中以 Relief-F 方法為首選，因為 Relief-F 是公認效果最好的過濾式（filter）特徵選擇方法（Kononenko, 1994）。

但畢竟是利用國外的資料集來得出重要的汽車屬性（Automobile Attributes），這些屬性是不是獲得國內消費者的認可值得懷疑。因此本研究再實施第二階段分析，將第一階段利用 Relief-F 方法得到的少數重要汽車屬性製作成問卷、對廣大消費者施測，並將得到的數據利用逐步迴歸（Stepwise Regression）分析法，如此可得到更重要、顯著的自變數，也就代表是消費者更注重的、在意的汽車屬性。如此本研究即可再進一步利用 ANOVA 分析出顧客基本的特性與購買的重要因素之檢定，就可以分析出汽車市場的區隔。市場區隔有了清楚的定義，各家車廠就能善用資源針對自己的市場去分析，同時也能定義出直接競爭者、潛在競爭者及替代品競爭者，形成 SWOT 分析的基礎。

將上述整個敘述繪製成研究流程，如圖 2。本研究欲達成幾個目的：

- (一)利用 Relief-F 特徵選擇法，篩選國外汽車資料集的重要汽車屬性。
- (二)將上述(一)到的重要汽車屬性製程問卷，針對國內汽車消費者發放。
- (三)再利用逐步迴歸法對問卷結果分析出國內消費者所重視的汽車屬性。
- (四)針對消費者特性與汽車重要屬性，分析汽車市場的區隔特性。

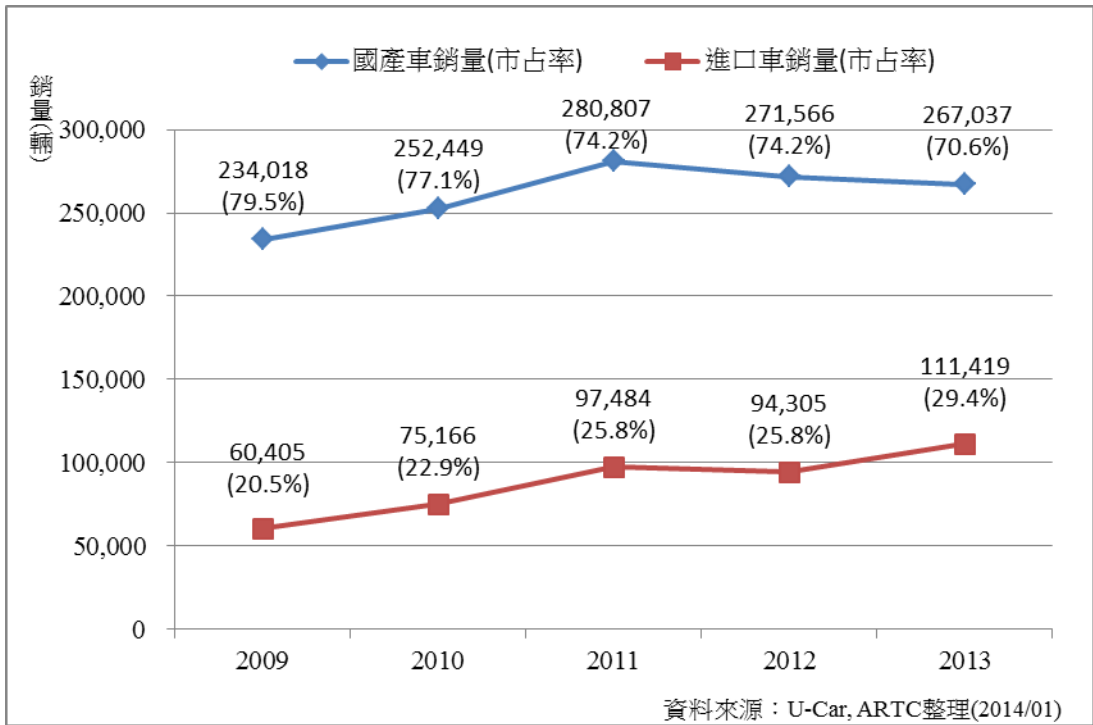


圖 1 近年我國國產車與進口車市占率概況

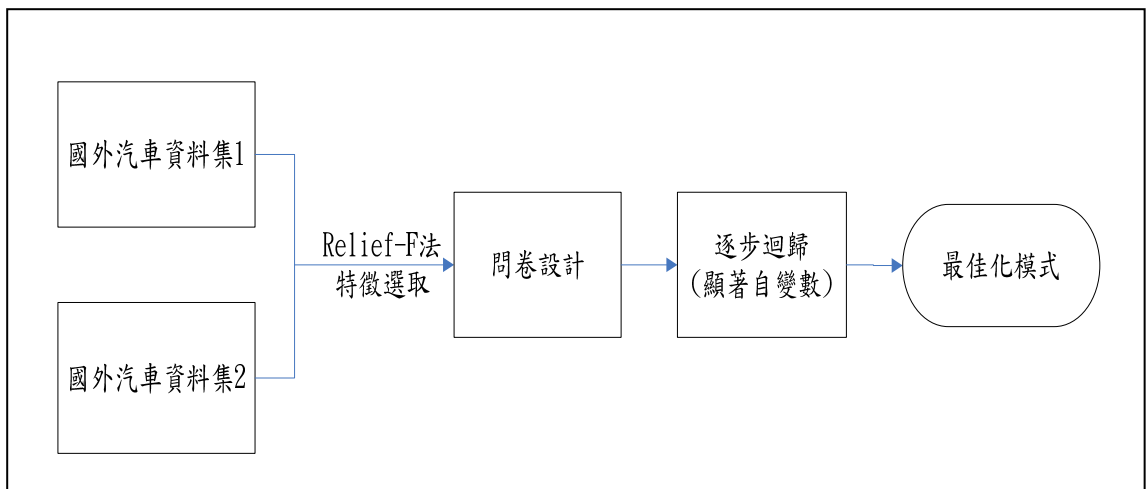


圖 2 研究流程圖

## 貳、文獻回顧

### 一、產品屬性

產品屬性則是產品所有外顯和內含的各種特徵性質的組合，而能為消費者所察覺者。企業推出的產品均有不同的產品屬性，至於這些屬性是否均能滿足消費者的需求，且為顧客所認同，則尚待進一步評估。

劉水深（1984）將產品屬性依其表現方式之不同，可分為四種，由內而外分別為：原生屬性、形式屬性、知覺屬性、衍生屬性，如圖 3 示。

#### 1.原生屬性（Essential Attribute）

原生屬性系指產品的物理、化學、機械等各項功能，係賦予產品實質效能的屬性。具體言之，即為實際使用該產品時，該產品所提供之各種基本性能，而廣為顧客所察覺、瞭解、接受者。例如汽車的原生屬性就是板金、零組配件等。

#### 2.形式屬性（Essential Attribute）

為了滿足顧客的要求，產品必需以一種實體的形式出現，可以看見觸摸、購買、使用與感覺。通常所謂的產品即以此階段出現。例如汽車的形式屬性是車子的外型、色彩搭配、配備的多寡..等。

#### 3.知覺屬性（Perceived Attribute）

在形式屬性的外圍，是顧客的知覺，包括了顧客對產品的態度和本身的認知。一般而言，著重在顧客的心理層面，偏向顧客和產品的互動（interaction）關係。知覺屬性不以實質形式彰顯於外，而是隱藏於產品之中，是顧客對產品的期待或是信念。例如對汽車品牌的認知、汽車的獨特性、價格的合理性..等。

#### 4.衍生屬性（Augmented Attribute）

舉凡與產品有關的服務活動、特性皆屬之。衍生屬性係延展產品長度、寬度、深度的一種屬性，最明顯的例子即為汽車的售後服務、品質保證等。

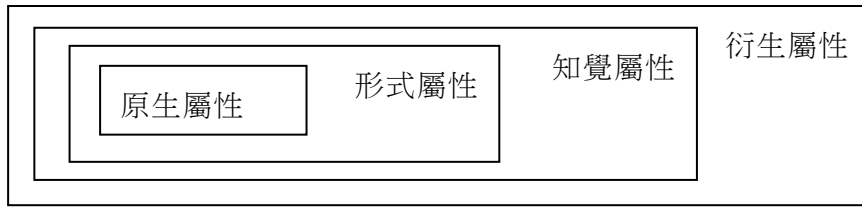


圖 3 產品的四種屬性

綜合以上所述，產品屬性可分四類，但不一定每項產品均包含此四者。本研究在汽車屬性探討裡面，只包含了前二項：原生屬性與形式屬性。

## 二、購買行為

顧客購買行為之研究分析，主要目的在於了解目標市場中，顧客購買行為的基礎、特性及動向，以作為市場管理的依據。Spears and Singh (2004) 提出購買意願為消費者願意購買該產品的可能性，它也是個人對於產品的行動傾向，亦是指個人想要購買產品的計劃。

經過各方案評估之後，顧客會選擇一個最能解決原來問題的方案，並採取購買行動。一般而言，當顧客對於某一產品或品牌的態度頗佳時，其購買意願愈高，顧客選擇該產品或品牌的機會愈大。本研究操作型定義為：『購買意願』是指廠商根據顧客問卷分析得到顯著的汽車屬性，然後推出顧客喜愛的汽車產品。並配合顧客特性來區隔市場，搭配車商的行銷策略，來吸引顧客購買該汽車的意願增加。

## 三、Relief-F 特徵選取演算法

特徵選取是一種基於最佳化標準的過程，從一個擁有大量特徵的集合中，挑選出有助於提高分類效能的特徵組合來做為影像分析、影像切割和影像分類等用途的依據。特徵選取這個概念已經發展已久，也發展出很多特徵選取的方法。

Relief-F 是公認效果較好的評估方法，由 White (1994) 提出，他擴展了基本的 Relief 演算法，改進了 Relief 演算法只能解決兩類的分類問題，使得 Relief-F 可以解決多類的分類問題 Relief-F。

Wang and Makedon (2004) 利用 Relief-F 演算法在微陣列基因表現 (microarray gene expression) 分類器上，現階段在微陣列資料過於龐大，使得進行分類上會花費

過多的時間，亦會影響分類器的效能表現。作者透過 Relief-F 方法與其他特徵選擇法，包括資訊增益法(Information Gain)、增益比例法(Gain Ratio)與卡方統計法(Chi-square Statistic)，將過濾得到的少數特徵放入 SVM、KNN 分類器中，得到利用 Relief-F 方法所選擇出的特徵，會讓分類器的效能表現比其他特徵選擇法來的好。

Spolaor, Cherman, Monard, and Lee (2013) 透過特徵選擇過程旨在選擇模型構建中所使用相關特徵的子集，以通過消除不相關和冗餘特徵(irrelevant and redundant features)來降低大量數據的維度(data dimensionality)。本文通過擴展單標籤特徵選擇 Relief-F 算法，提出了一種新的多標籤特徵選擇(Multi-label Feature Selection)算法 RF-ML，利用合成數據集，將所提出的算法與 Relief-F 算法進行實驗比較，結果顯示將排序過的相關特徵與所提出的算法結合，可以得到最好的效果。

從以上文獻得知 Relief-F 應用在很多方面領域，當該領域有很多重要特徵需要選取時，都會使用此法。本研究亦是從國外文獻蒐集大量有關顧客在選購汽車時的考量因素，因此使用 Relief-F 法先將重要的購買決定因素選取出來，再依據該重要因素設計一份問卷以讓國內消費者來填寫，從而得知國內消費者的習性。

#### 四、逐步迴歸法

逐步迴歸(Stepwise Regression)為運用廣泛的複迴歸分析法之一，其結合了順向選擇法(Forward Addition)與反向剔除法(Backward Elimination)的優點，其步驟如下：

- 1.將解釋變數逐一代入迴歸式分析，一次僅放入一個自變數。
- 2.首先進入之解釋變數應與被解釋變數相關性最大者，即最大正相關或最大負相關者，接著剩餘解釋變數與解釋變數間之相關性高者，可能會有高度線性重合現象，宜取捨挑選其一。
- 3.在逐一解釋變數代入的過程中，若  $\text{adj}R^2$  減少或呈現該解釋變數為不顯著，則該解釋變數應汰除，反之則該留下。

Wong and Huang (2006) 以網路入侵檢測系統做研究，若能夠消除不顯著、無用的特徵，即可讓檢測系統產生更高的性能。因此他們結合了逐步迴歸分析與支持向量機(Support Vector Machine)用以檢測網路入侵，實際結果證明，使用逐步迴歸法所獲得的最佳特徵模型即可以獲得與全部特徵集合相同的性能。

Chen, Shi, Shu, and Gao (2013) 根據 2005 年至 2009 年在華東地區的數據，建立了 PM10 (粒徑小於 10 $\mu\text{m}$  的顆粒物) 濃度預測模型，並結合了小波分析和逐步迴歸來預測模型，並使用 2010 年的測試數據，計算出區域模型預測結果的準確率。研究指出該結合的模型預測 PM10 濃度的優點具有廣泛的空間和時間適應性，在中國東部地區 PM10 濃度預報的準確率和精度最高均有顯著提高。

從以上文獻得知在問卷回收後，能利用逐步迴歸法來找出顯著自變數。本研究根據第一階段：針對大數據下顧客對汽車購買的因素，使用 Relief-F 法篩選出重要的決定因素；再依據此重要因素設計問卷、發放、回收，建立多元的迴歸方程式。最後再利用逐步迴歸法將找出該方程式中顯著自變數，並與問卷中消費者的基本資料進行 ANOVA 分析，得到顧客基本的特性與購買的重要因素之檢定，即可分析出汽車市場的區隔。

## 參、研究方法

### 一、研究架構

本研究架構如圖 4，說明如下：

- (一)蒐集國外二筆大數據汽車資料，先做資料清除動作，將有遺漏值的資料去除。
- (二)使用 Relief-F 法，先將重要的購買汽車屬性篩選出來。
- (三)再依據篩選出之重要汽車屬性設計問卷、發放與回收。
- (四)建立多元的迴歸方程式，再利用逐步迴歸法將找出該方程式中顯著自變數。
- (五)得到最佳化模式，並與問卷中消費者的基本資料進行 ANOVA 分析，得到顧客基本的特性與購買的重要汽車屬性，即可分析出汽車市場的區隔。



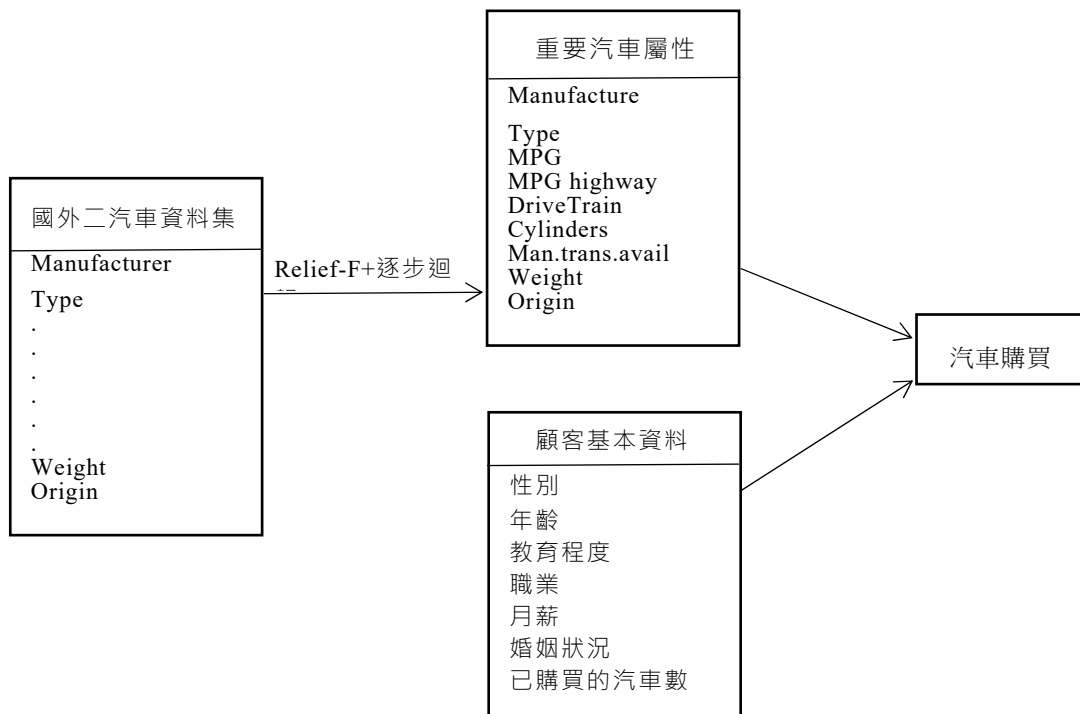


圖 4 研究架構圖

## 二、研究資料說明

本研究先蒐集到二筆國針對汽車特徵、各車款銷售量（出廠數）及消費者屬性資料來做分析。說明如下：

### (一)汽車資料集 1

該汽車資料是取自美國，針對 1993 乘客對汽車消費與購買需求所蒐集到的資料，共包含 93 筆資料與 25 個特徵，如表 1。

### (二)汽車資料集 2

該汽車資料是取自卡內基·梅隆大學（Carnegie Mellon University）圖書館，共包含 392 筆資料與 9 個特徵，如表 2。

表 1 汽車資料集 1 的特徵說明

Attributes	特徵中文解釋	Attributes	特徵中文解釋
Manufacturer	製造商	Rev.per.mile	發動機轉速
Type	車型	Man.trans.avail	手排/自排
Min.Price	基本款售價	Fuel.tank.capacity	油箱容量
Price	平均售價	Passengers	載客數
Max.Price	旗艦款售價	Length	車長
MPG.city	市區里程/每加侖	Wheelbase	軸距
MPG.highway	高速公路里程/每加侖	Width	車寬
AirBags	氣囊數	Turn.circle	迴轉半徑
DriveTrain	傳動系統	Rear.seat.room	後座空間
Cylinders	汽缸數	Luggage.room	行李空間
EngineSize	引擎尺寸	Weight	車重
Horsepower	引擎馬力	Make	汽車生產地(應變數)
RPM	轉速		

表 2 汽車資料集 2 的特徵說明

Attributes	特徵中文解釋	Attributes	特徵中文解釋
MPG	英哩/每加侖	Acceleration	加速性能
Cylinders	汽缸數	Year	出產年份
Displacement	發動機排氣量	Origin	汽車原始產地
Horsepower	引擎馬力	Name	汽車名稱(應變數)
Weight	車重		

### 三、評估準則

#### (一)Relief-F 評估準則

Relief-F 方法其主要的思想是，好的特徵應該使同類的樣本接近，而使不同類的樣本之間遠離。該法會給每個特徵一個權重，並指出與類別的相關性，然後根據特徵對鄰近樣本的區分能力來更新該特徵權重，權重較高的特徵會被選取出來，即

$$W[i]=W[i]-\sum \text{diff}(i,R,H)/m+\text{diff}(i,R,M)/m$$

式中  $W[i]$  為特徵  $i$  的權重， $R$  是訓練集中隨機選取的樣本， $H$  是  $R$  中來自不同類別最鄰近的樣本， $M$  是  $R$  中來自同類別最鄰近的樣本，參數  $m$  是  $m$  次隨機從資料集中抽取的樣本。Relief-F 運算效率高，對數據類型沒有限制，屬於一種特徵權重算法，該演算法會賦予所有和類別相關性高的特徵較高的權重。

## (二)逐步迴歸評估準則

本研究逐步迴歸法是利用自由軟體 R 中的 `step` 指令來執行，其準則是採用 Akaike(1974)所提出的 AIC (Akaike Information Criterion)，其統計量定義如下：

$$AIC = -2 * \log(L(M)) + 2 * k$$

其中  $L(M)$  為配適參數模型  $M$  下的概似函數 (likelihood function) 在 MLE 的值， $k$  為參數個數。AIC 主要是評估統計模型的複雜度和衡量統計模型「擬合」資料優良性的一種標準，它的優點在於其具有有效性 (efficiency)，由於其參數懲罰項相較於其他模型準則少，故容易有參數過度配適 (over fitting) 的問題，但會隨著樣本數的增加而改善。所以優先考慮的模型應是 AIC 值最小的那一個。

# 肆、研究結果

## 一、Relief-F 演算法執行結果

### (一)Relief-F 演算法於汽車資料集 1

從表 3 可以得到利用 Relief-F 演算法所算出資料集 1 重要的特徵值，其中 Manufacturer、Type 二項的特徵值 (屬極端值) 較高，因此提高了整個平均值 (Mean) 為 0.3199。我們以中位數 (Median) 為標準，將值小於 0.2251 的特徵刪除。所以資料集 1 從原本有 25 個特徵再刪除後只剩下 12 個特徵。

### (二)Relief-F 演算法於汽車資料集 2

從表 4 可以得到利用 Relief-F 演算法所算出資料集 2 重要的特徵值，因該特徵值並無特別較高的特徵值發生，因此我們以平均值 (Mean) 0.2644 為基準，將值小於 0.2644 的特徵刪除。所以資料集 2 從原本有 9 個特徵在刪除後只剩下 5 個特徵。

所以二個資料集利用 Relief-F 法篩選過後，剩下的重要汽車屬性應該有 17 個；其中有 1 個重要屬性在二個資料集均有被選取，那就是 Cylinders (汽缸數)。因此正確的重要汽車屬性是 16 個。

表 3 資料集 1 重要特徵值

	attr_importance	attr_importance information
Manufacturer	0.973913	
Type	0.8456522	
Min.Price	0.1910572	
Price	0.1503032	
Max.Price	0.1325363	
MPG.city	0.2273492	
MPG.highway	0.2592754	
AirBags	0.5923913	
DriveTrain	0.4065217	
Cylinders	0.5369565	Min.:0.1325
EngineSize	0.2	1stQu.:0.1935
Horsepower	0.1880834	Median:0.2251
RPM	0.2228261	Mean:0.3199
Rev.per.mile	0.1942661	3rdQu.:0.3616
Man.trans.avail	0.4423913	Max.:0.9739
Fuel.tank.capacity	0.2076637	
Passengers	0.2041667	
Length	0.1625279	
Wheelbase	0.2596327	
Width	0.1977053	
Turn.circle	0.270903	
Rear.seat.room	0.1859335	
Luggage.room	0.3466712	
Weight	0.2778031	

表 4 資料集 2 重要特徵值

	attr_importance	attr_importance information
mpg	0.2226912	
cylinders	0.3216621	Min.:0.1337
displacement	0.2739512	1stQu.:0.2234
horsepower	0.223642	Median:0.2711
weight	0.2681915	Mean:0.2644
acceleration	0.1337248	3rdQu.:0.3069
year	0.3020277	Max.:0.3693
origin	0.3693403	

## 二、逐步迴歸法執行結果

根據上一節所述，二個資料集利用 Relief-F 法篩選過後，剩下的重要因素有 16 個。其建立初步多元迴歸模式如下：

$$\begin{aligned} \text{Car} = & \beta_0 + \beta_1 \text{cylinders} + \beta_2 \text{displacement} + \beta_3 \text{weight} + \beta_4 \text{year} + \beta_5 \text{origin} + \\ & \beta_6 \text{Manufacturer} + \beta_7 \text{Type} + \beta_8 \text{MPG.city} + \beta_9 \text{MPG.highway} + \beta_{10} \text{AirBags} + \\ & \beta_{11} \text{DriveTrain} + \beta_{12} \text{Man.trans.avail} + \beta_{13} \text{Wheelbase} + \beta_{14} \text{Turn.circle} + \\ & \beta_{15} \text{Luggage.room} + \beta_{16} \text{Weight} + \varepsilon \end{aligned}$$

茲將這 16 個重要汽車屬性與填表人的基本資料做成問卷並於網路上發放。會選擇於網路上發放問卷是希望借助網路無遠弗屆的力量外，亦希望透過網路病毒式的傳播，無論性別、地區別、年齡別..等皆能填答，使問卷的發放更接近於隨機抽樣。這樣會使得問卷結果與問卷分析更具代表性。

表 5 顯示逐步迴歸執行結果，從結果可以看出在 16 個迴歸變數裡面，最後篩選出最重要的汽車屬性有 9 個，分別是：汽車品牌（製造商）、汽車造型、是否省油、高速公路是否省油、汽車傳動系統、汽車汽缸數、汽車手／自排、汽車重量與汽車原產地。

## 三、變異數分析執行結果

根據上一小節逐步迴歸所得到的結果，共有 9 個重要汽車屬性。本研究將這些重要的汽車屬性與問卷上的基本資料進行變異數分析（ANOVA）。問卷上的基本資料有：性別、年齡、教育程度、職業、月薪、婚姻狀況、已購買的汽車數等七項。

本研究以顯著水準 $\alpha=0.05$ ，若 p-value 值小於 0.05 則代表其具有顯著性。若基本資料選項超過 3 個，再依據 Scheffe 事後分析法，找出是哪些選項具有顯著不同，茲將分析結果整理如表 6，說明如下：

1. 在基本資料－『性別』上與重要汽車屬性-『汽車造型』的 p-value 值為 0.031，代表男、女對汽車造型有顯著的差異。

表 5 逐步迴歸執行結果

Step : AIC=-208.86				
Car ~ Manufacturer + Type + MPG + MPG.highway + DriveTrain + Cylinders + Man.trans.avail + Weight + Origin				
	Df	Sum of Sq	RSS	AIC
<截距>			26.790	-208.86
汽車重量	1	0.48224	27.272	-208.38
汽車傳動系統	1	0.75131	27.541	-207.01
汽車造型	1	0.78328	27.573	-206.85
汽車品牌（製造商）	1	0.86510	27.655	-206.44
是否省油	1	1.00122	27.791	-205.76
汽車原產地	1	1.00477	27.794	-205.74
高速公路是否省油	1	1.23506	28.025	-204.59
汽車汽缸數	1	1.82769	28.617	-201.68
汽車手／自排	1	1.97708	28.767	-200.96

表 6 ANOVA 分析表

	重要汽車屬性	p-value 值	Scheffe 法
性別	汽車造型	0.031 <sup>*</sup>	
年齡	汽車品牌（製造商）	0.043 <sup>*</sup>	20~25 歲、51 歲以上
教育程度	高速公路是否省油	0.032 <sup>*</sup>	大學、碩士
月薪	汽車手／自排	0.015 <sup>*</sup>	18000~30000、50001 以上
婚姻狀況	汽車造型、高速公路是否省油	0.000 <sup>**</sup>	
已購買的汽車數	汽車造型	0.033 <sup>*</sup>	一輛、三輛

(本研究整理)

- 2.在基本資料－『年齡』上與重要汽車屬性－『汽車品牌』的 p-value 值為 0.043，因年齡具有六個級距，依據 Scheffe 事後分析得知在 20~25 歲、51 歲以上二個年齡層對汽車品牌有顯著的差異。
- 3.在基本資料－『教育程度』上與重要汽車屬性－『高速公路是否省油』的 p-value 值為 0.032，因教育程度具有四個級距，依據 Scheffe 事後分析得知大學、碩士二個教育程度對高速公路是否省油有顯著的差異。
- 4.在基本資料－『月薪』上與重要汽車屬性－『汽車手／自排』的 p-value 值為 0.015，

因月薪具有五個級距，依據 Scheffe 事後分析得知在月薪 18000~30000、50001 以上二個級距對汽車手／自排有顯著的差異。

- 5.在基本資料－『婚姻狀況』上與重要汽車屬性－『汽車造型』、『高速公路是否省油』的 p-value 值為 0.000，代表已婚、未婚對汽車造型與高速公路是否省油有顯著的差異。
- 6.在基本資料－『已購買的汽車數』上與重要汽車屬性－『汽車造型』的 p-value 值為 0.033，因已購買的汽車數具有四個級距，依據 Scheffe 事後分析得知在家庭已購買的汽車數為一輛車、三輛車時對汽車造型有顯著的差異。

## 伍、結論

### 一、總結

(一)本研究先針對二筆國外的所蒐集到的汽車屬性，先利用第一階段：Relief-F 法將有用的汽車屬性篩選出來。從原本二資料庫共有 34 汽車屬性中，以 Relief-F 過濾後只保留下 16 個重要屬性。利用這 16 個重要屬性製作成問卷、並發放。網路上填寫問卷目的是能更接近隨機抽樣方式。問卷回收後，再第二階段利用逐步迴歸法求出國內消費者認為最有顯著差異的汽車購買屬性。分別是：汽車品牌(製造商)、汽車造型、是否省油、高速公路是否省油、汽車傳動系統、汽車汽缸數、汽車手／自排、汽車重量與汽車原產地等九項因素。

(二)這九項汽車購買重要屬性再與問卷基本資料做變異數分析 (ANOVA)，得到以下結論：

- 1.在性別上，男、女生對於汽車造型上會有顯著差異。建議汽車廠商針對男、女生分別開發不同車型來吸引這二個不同族群，以增加產品購買率。
- 2.在年齡上，『20~25 歲』、『51 歲以上』這二個族群對於汽車品牌認定上是具有顯著差異的。『20~25 歲』這族群屬於剛進入社會可能是首購族，應該會對於價格、

性能..等為考量點；反觀『51 歲以上』族群可能在社會上是有身份地位的，好的汽車品牌是能彰顯其在社會上的地位。

- 3.在教育程度上，以『大學』、『碩士』二個族群對『高速公路是否省油』有顯著差異。建議汽車廠商可以開發更省油的車型來吸引這二個不同教育程度的族群，以增加產品購買率
- 4.在月薪上，以月薪落在『18,000~30,000』與『50,001 以上』這二個級距，對汽車是否為手／自排會有顯著差異。月薪高、低通常代表著社會上的成年人與新鮮人；手排車駕駛的樂趣與操控感會高於自排車，駕駛手排車以實現自我滿足為主。建議汽車廠商可以以操控樂趣、自我滿足的訴求，針對這二個月薪級距推出手／自排的車款，來吸引購買意願。
- 5.在婚姻狀況上，『已婚』與『未婚』對於汽車造型、高速公路是否省油有顯著差異。『已婚』者會以家庭人口數、舒適、安全性、與家庭支出方面為考量的重點。因此建議汽車廠商針對已婚者推出以舒適、安全性的汽車造型；與經濟省油的車款來區隔已婚、未婚的市場區隔，來增加購買意願。
- 6.在已購買的汽車數上，其中購買『一輛車』與已經購買『三輛車』對於汽車造型上有顯著差異。顧客在購買第一輛車與已經購買三輛車時，對汽車的造型上需求、想法應該會不同。建議汽車廠商可以針對購買第一輛車與要購買第三輛車以上的顧客，開發不同造型的車已符合顧客的需求。

## 二、管理意涵

在競爭激烈的汽車市場中，無論國內、外廠商都希望能知道顧客對汽車的需求為何，以便能開發符合顧客需求的車型，進而獲得顧客消費的青睞。本研究在管理的意涵可得到下列三點：

- (一)汽車的屬性太多了，若能知道少數重要的汽車屬性，汽車公司便能利用 80/20 法則，將公司多數資源分配給這些符合顧客需求的重要屬性，減少車商大規模開發每個汽車屬性，以便減少成本、提高公司利潤。
- (二)不是每個客戶，都是你的顧客。將重要汽車屬性與顧客基本資料做分析，以便知道什麼樣的客戶著重在汽車哪一方面的屬性。這符合行銷 STP 原則，透過市場區



隔 (segmentation) → 選擇目標市場 (targeting) → 定位 (positioning) 的過程，集中行銷力道在願意來購買汽車的人身上。

(三)市場區隔有了清楚的定義，各家車廠就能善用資源針對自己的市場去分析，同時也能定義出直接競爭者、潛在競爭者及替代品競爭者，形成 SWOT 分析的基礎。

## 致謝

本研究由科技部計畫支持（計畫編號：MOST 104-2815-C-244-002-H）。

## 參考文獻

### 一、中文部分

1. 劉水深(1984)，產品規格化與策略應用，台北：華泰。
2. 陳淑芬(2014)，大數據時代來臨，禪天下，117，62-65。

### 二、英文部分

1. Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716-723.
2. Chen, Y., Shi, R., Shu, S., & Gao, W. (2013). Ensemble and enhanced PM 10 concentration forecast model based on stepwise regression and wavelet analysis. Atmospheric Environment, 74, 346-359.
3. Kononenko I. (1994). Estimating Attributes: Analysis and Extensions of RELIEF. Paper presented at the Proc. of the 1994 European Conference on Machine Learning, Italy.

4. White House (2012), Big data fact sheet, Retrieved August 23, 2014, from [http://www.whitehouse.gov/sites/default-/files/microsites/ostp/big\\_data\\_fact\\_sheet\\_final.pdf](http://www.whitehouse.gov/sites/default-/files/microsites/ostp/big_data_fact_sheet_final.pdf).
5. Spears, N., & Singh, S. N. (2004). Measuring attitude toward the brand and purchase intentions. Journal of Current Issues & Research in Advertising, 26(2), 53-66.
6. Spolaor, N., Cherman, E. A., Monard, M. C., & Lee, H. D. (2013). ReliefF for Multi-label Feature Selection. Paper presented in Proceedings of the Brazilian Conference on Intelligent Systems (BRACIS), Brazil.
7. Wang, Y., Makedon, F. (2004). Application of Relief-F Feature Filtering Algorithm to Selecting Informative Genes for Cancer Classification Using Microarray Data. Paper presented in IEEE Computational Systems Bioinformatics Conference, USA.
8. Wong, W. T., & Huang, W. C. (2006). Toward the best feature model for network intrusion detection using stepwise regression and support vector machine. International Computer Symposium, 2, 843-848.

106年02月06日收稿

106年02月07日初審

106年04月23日複審

106年05月16日接受

## 作者介紹

### Author's Introduction

姓名 蘇高玄  
Name Kao-Hsuan Su  
服務單位 高苑科技大學行銷與流通管理系講師  
Department Instructor, Department of Marketing and Distribution Management, Kao Yuan University  
聯絡地 82151 高雄市路竹區中山路 1821 號  
Address No.1821, Jhongshan Rd., Lujhu Dist, Kaohsiung City 82151, Taiwan  
E-mail t70027@cc.kyu.edu.tw  
專長 資料探勘，機器學習  
Speciality Data Mining, Machine Learning

姓名 蔡佳蓉  
Name Chia-Jung Tsai  
服務單位 高苑科技大學行銷與流通管理系大學生  
Department Student, Department of Marketing and Distribution Management, Kao Yuan University  
聯絡地 82151 高雄市路竹區中山路 1821 號  
Address No.1821, Jhongshan Rd., Lujhu Dist, Kaohsiung City 82151, Taiwan  
E-mail l224556118@gmail.com  
專長 行銷管理，消費者行為  
Speciality Marketing Management, Consumer Behavior