

以資料探勘及分析層級程序法探討家戶購屋之購屋決策因素

翁頌舜

國立台北科技大學資訊與運籌管理研究所

wengss@ntut.edu.tw

連偉伶

國立台北科技大學資訊與運籌管理研究所

lianbarbara1006@gmail.com

摘要

大部分探討購屋決策因素之文獻僅考慮確定性資料對購屋決策制定的影響，卻忽略心理因素。另外，上述論文的購屋因素大多由主觀看法決定，例如：專家法。這將忽略其他可能的影響因素。本研究發展一個二階段實驗方法探討購屋決策因素。第一階段使用資料探勘技術的 *k*-means 演算法對購屋市場做區隔，再用關聯法則分析尋找潛藏的有趣樣式。第二階段結合階段一與心理因素調查的結果導入分析層級程序做決策制定。本研究將探討台北市行政區購屋決策因素之關聯性異同與重要程度之排序。希望藉此對具有關聯性的購屋決策因素有更深入的認識。

關鍵詞：購屋決策因素、資料探勘(DM)、分析層級程序法(AHP)

壹、導論

房地產景氣之波動是由於房地產市場供給與需求不均衡所帶來的結果。當房地產的需求大於供給時，將造成市場供不應求的現象，房地產市場即呈現繁榮的景氣，反之當供給大於需求時，供過於求的現象將使房地產市場呈現蕭條。上述是房地產市場的本身條件如何影響供需狀態，然而，總體經濟與金融市場對房地產供需的影響是相當顯著而且是不可忽視的。

過去房地產業被認為是「火車頭產業」，此非學術性的敘述包含兩層意涵：第一，因為房地產業被認為具有極大的「向後關聯」效果，即是說房地產業景氣可帶動建材、運輸、金融、仲介等週邊事業的發展；第二，既然房地產業能帶動總體經濟成長，也就代表房地產業活動具有領先總體經濟景氣的性質，可以將房地產業的活絡程度視為總體經濟景氣的領先指標，也就是說可將房地產景氣視為一般景氣變化的預期訊號(Hansen 1951; Matthews 1969)。然而，林秋瑾、王建安與張金鶚(1996)在探討房地產景氣與總體經濟景氣於時間上之領先或落後關係後，以實證分析發現以往認為房地產業是火車頭工業的說法並不成立，反而是「總體經濟景氣領先房地產景氣」的說法得到充分實驗證據的支持。彭建文與張金鶚(2000)檢視總體經濟變數對台北地區(包含新北市、台北市)等不同地區房地產景氣的影響，他們指出：在國內三次景氣大規模波動前，「台北縣市的預售屋房價與貨幣供給、空屋數、建照面積等變數均具有長期均衡關係存在，但在考量結構變遷下，預售屋房價與建照面積之間關係將會改變。其次，空屋數對於預售屋房價的影響，顯示要解決房地產市場近年不景氣的狀況還是要從消化空餘屋著手」。根據彭建文與張金鶚(1995)指出，台灣的空屋量相當龐大，當時統計國內總空屋數達到六十八萬之譜，國內

空屋率更是國外先進國家的三倍以上。如此大量多餘的供給量(空屋數)，卻使得社會中無殼蝸牛充斥，這是相當不合理的現象。國內空屋的敘述統計資訊如表 1 所示。本研究針對彭建文與張金鶚(1995)中提出造成台灣高空屋率的眾多可能原因中，針對住宅資訊不健全而造成待售、待租之空屋，提出以資料探勘技術找出未知的有趣的住宅消費行為樣式。

表 1 國內空屋的敘述統計資訊

表 臺灣地區各縣市歷年平均空屋率—1980至1992

縣市	最低	最高	平均	標準差
基隆市	15.08	21.94	18.74	1.8455
臺中市	15.06	21.28	19.38	1.5437
臺南市	14.77	19.21	17.57	1.3064
高雄市	12.59	17.17	15.57	1.2538
臺北縣	15.14	22.32	18.27	2.5314
宜蘭縣	15.67	20.39	17.81	1.6829
桃園縣	14.89	24.14	17.73	2.9555
新竹市	14.79	18.78	16.63	1.3037
新竹縣	7.41	11.79	9.50	1.6210
苗栗縣	7.12	9.43	8.19	0.8580
臺中縣	11.68	20.03	15.54	3.1670
彰化縣	13.53	18.16	15.67	1.6193
南投縣	11.33	14.64	12.92	1.2620
雲林縣	8.24	10.58	9.48	0.7821
嘉義市	13.20	21.77	18.85	2.1728
嘉義縣	11.03	13.96	12.59	1.0296
臺南縣	9.60	12.80	11.13	1.2420
高雄縣	10.57	13.36	12.30	0.9841
屏東縣	7.96	9.69	8.80	0.5153
臺東縣	8.46	10.56	9.62	0.6513
花蓮縣	12.62	17.06	15.46	1.2819
澎湖縣	12.87	17.10	15.19	1.2915
臺北市	8.56	13.42	10.87	1.8066
臺灣地區	12.96	16.23	14.68	1.2032

資料來源：張金鶚，彭建文(1995)

彭建文與張金鶚(2000)指出房地產本身具有許多異於一般產品的特性且房地產市場又具有相當強烈的地域性，在人口、區位、行政區特性，以及都市化程度...等不同條件下，不同地區房地產市場存在不同的供需條件。由於台北市是台灣政治與工商發展的中心，它具有引領和牽動台灣整體房地產經濟之地位。另外，由於台北市、新北市等居民所得、社經地位均高於其餘縣市，這也意味著台北市居民的購屋能力比較高，使得交易筆數較多，如此較易取得大量的樣本，也可以儘量避免因為抽樣樣本不足導致資料代表性失效的問題。當考量交易資料蒐集之過程，因為資料庫發生記錄錯誤或紀錄資料不全等非自然因素的問題不可避免。我們認為，若交易筆數夠多則較易取得完整的資料，雖然不能避免紀錄錯誤，但是當樣本足夠大時，將使得該錯誤的影響變小。因此綜合以上原因本研究選擇研究台北市的房地產交易狀況。

房地產景氣受到總體經濟景氣波動的影響相當巨大。前些年國內總體經濟景氣波動造成空屋數大量增加。我們認為深入研究家戶購屋因素對於瞭解房地產市場的供給面與需求面都有直接的正面意義。從供給面角度，藉由分析歷史購屋交易資料得知家戶購屋因素對於業界有相當重要的幫助。例如：對於建築業者而言，在知道家戶購屋因素的前提下，較能提供滿足家戶購屋需求的住宅物件，減少不必要的資金浪費，增加公司的競爭力。對於房仲業而言，如果知道家戶購屋因素，將可以更快且更好地確定家戶購屋需求，如此一來，可以減少每筆交易的時間成本(時間成本在本論文中定義為：家戶決策者從開始接受房屋仲介服務，包含實地看屋、討論合約內容等等到完成房屋交易所需的時間)，也能夠提升房仲人員的服務品質，增加房仲公司的競爭力。在這樣的背景下，本研究將對家戶的購屋因素深入探討，希望藉由資料探勘技術與分析層級程序法找出具有價值的資訊。

貳、 文獻探討

一、家戶購屋決策因素

家戶的購屋決策行為的相關文獻實可用汗牛充棟來形容，大部分文獻幾乎都是以使其家戶的效用最大化的個體經濟概念為最終目標，研究資料來源分別是過去的住宅普查資料、向房仲業者購得的歷史交易資料、或者是對已購屋者進行的問卷調查資料，近幾年的文獻則更進一步考慮總體經濟與社會條件對購屋決策產生的影響。其中，能夠處理效用的分析方法大多為統計迴歸模型或是各種羅吉特模型。探討的課題則包含住宅屬性、住宅區位與住宅外部環境與家戶本身對購屋決策行為的影響，也有文獻考慮租買選擇(住宅權屬)、遷移(區位選擇、遷移成本)、就業決策等與住宅需求的關係，而這些因素往往相互影響或同時被決定(Lancaster 1966; Rosen 1974; Follain & Jimenez 1985; Assadian & Ondrich 1993; 陳彥仲 1997; 陳彥仲、吳京玲 1998; 薛立敏等人 2002; Li 1977; 陳淑美、張金鶚 2002; 連經宇 2002; 連經宇、陳彥仲 2005)。曹葦如、張金鶚與林秋瑾(1999)以質化方法訪談不同家庭生命週期的已購屋者瞭解購屋決策的過程。他們發現已購屋者並沒有因為房屋價格較高而有較細緻的考量，所購得的房屋在價格與屬性上也未必符合心中所期待的。由此可知，在個體經濟學中「消費者是理性的」的假設顯然不常成立。謝文盛與林素菁(2000)透過研究民國八十五年的個人所得分配調查報告發現，台灣地區之財產稅賦對家戶租買選擇有相當程度的影響。當財產稅增加時，家戶則傾向以租賃方式取代購屋，減少購屋的需求;反之當恆常所得增加或房屋租賃價格上漲時，家戶傾向以自有方式消費住宅，減少租賃房屋的需求。McLeod, P. B.與 Ellis, J. R.(1983)證實所得因素是影響家戶購屋決策的關鍵因素。Michelson, R. S.(1977)針對家戶選擇住宅及搜尋住宅區位的過程做研究。他認為影響家戶遷移的主要因素為財務狀況，教育程度則會影響家戶的生活型態，上述兩項因素街和住宅區位選擇產生高度的相關性。張金鶚等人(1996)則認為家戶可能會因消費能力或其他特質差異而對不同住宅類型有不同的偏好。他們的實證結果顯示無論是家戶屬性或是住宅屬性的因素，都會顯著地影響家戶對不同住宅類型偏好的表現。林祖嘉(1990)則是針對自有及租賃兩種住宅類型的研究發現，兩種住宅類型真正地存有某種程度之替代性。根據上述文獻顯示，住宅單位的區位、位置、住宅價格、住宅類型都會影響購屋決策，所以我們把上述變數都加以考量。當消費者面對重要或是不確定性高的決策時，會傾向進行大量外部搜尋(Beatty & Smith 1987)。購屋搜尋者的相關文獻則針對正在進行的購屋搜尋行為或已購屋者的購屋搜尋行為進行深入探討，討論的主題包含租/購屋的經驗、購屋動機及有無透過仲介搜尋與搜尋期間的關係，以及影響搜尋努力程度的重要因素(李春長、張金鶚 1996; 周美伶 2005; 周美伶、張金鶚 2005; McCarthy 1982; Hempel & Jain 1978; Anglin 1997)。由此可知，購屋搜尋對於購屋決策也是影響因素之一。本研究限於研究資料涉及已購屋者的隱私，所以僅僅考慮購屋搜尋期間對購屋決策的影響。

本研究一開始即希望使用資料探勘技術在房地產海量般的資料庫中找出潛藏的有用資訊供往後研究發展，但是在財力與時間的考量下，本研究僅僅額外將住宅面積大小、購屋樓層高低、有無停車位、購屋時間等變數列入考量。

二、資料探勘

資料探勘(Data Mining, DM)觀念最早是由 Usama Fayyad 於 1991 年所提出，Berry 與 Linoff(2001)定義資料探勘:「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」。而資料探勘包含的技術有:群集分析(Cluster analysis)、分類(Classification)、推估(Estimation)、預測(Prediction)、關聯分析(Association analysis)。群集分析使用的方法有 *k*-means 和 agglomeration 法;分類使用的方法有決策樹(Decision tree)和記憶基礎推理(Memory-based reasoning);推估使用的方法有統計方法的相關分析、迴歸分析和類神經網路法;預測使用的方法有迴歸分析、時間數列分析和類神經網路法，最後關聯分析使用的方法有 Apriori 和交叉銷售(Cross-selling)法(韋端等人 2003)。

然而，國內、外應用資料探勘於「家戶購屋決策因素」的研究卻相對較少，僅見國內學者以群集分析技術做市場區隔，再利用 EKB 模式針對購屋動機、購屋資訊的尋求及購屋評估準則做研究(戚靜玟 1998)。其中 EKB 模式由 Engel et al(1968)所提出，EKB 模式將消費者行為視為一個連續過程，以消費者決策過程為主幹，再考慮影響各決策階段的相關因素，以及這些因素間的互動關係。嚴格來說，戚靜玟(1998)僅以群集分析做市場區隔，並未將完整的資料探勘技術應用到購屋決策問題。因此本研究欲使用家戶購屋歷史交易資料，並應用資料探勘中的「群集分析」和「關聯法則」兩種方法進行實證研究。

(一) 群集分析

Cheng 與 Chen(2009)認為要使知識發現的過程更準確、更容易，在交易資料準備妥當後，需做資料預處理(Data preprocessing)。而資料預處理目的，正是為了提高資料探勘的品質，畢竟「有高品質的資料，才有高品質的探勘結果」(曾憲雄等人 2007)。資料預處理程序包括資料清理(Data cleaning)、資料整合(Data integration)和資料轉換(Data transformation)、資料化約(Data integration)，及資料離散化和概念階層生成(Data discretization and concept hierarchy generation)(Jiawci & Micheline 2003)。

當我們對於問題的目標類別無法確切得知時，通常會採用群集分析此一分群方法來解決問題。將資料做分群的方法有以下幾種:劃分方法(Partitioning methods)、階層方法(Hierarchical methods)、基於密度的方法(Density-Based methods)、基於網格的方法(Grid-Based methods)，及基於模型的方法(Model-based methods)。本研究即使用劃分方法中的經典劃分方法「*k*-means 演算法」(Jiawci & Micheline 2003)。而 *k*-means 演算法執行步驟依序為:決定群集數量、計算各群集的中心、資料點距離的計算、將資料點分到特定的群集中，重複執行直到各群集中心點不再改變為止。

(二) 關聯法則

關聯法則被用來從大量的資料中發現關係和潛在的關聯性(Agrawal, Imielinski, & Swami 1993; Sanchez, Vila, Cerda, & Serrano 2009; Wesley Changchien & Lu 2001)。而所被探勘出的法則能夠有效地發現未知的關係，且對於預測和決策制定而言能夠提供一些結果(Rajasethupathy, Scime, Rajasethupathy, & Murray 2009)。其中，最被廣為人知的即是由 Agrawal 與 Srilkant 於 1995 年所提出的 Apriori 演算法，此演算法能夠有效地發現具有

意義的項目集合(itemset)並且在一個交易資料庫中建置關聯法則。Apriori 演算法主要目的，是希望從結論句的項目集合去找出頻繁項目集合(frequent itemset)，以此來歸納出強關聯法則，本研究即採用 Apriori 演算法。而關聯法則大致上可分為三種型式:屬性值型式的關聯法則-布林型關聯法則(Boolean association rule)和定量型關聯法則(Quantitative association rule);資料維度型式的關聯法則-單一維度關聯規則(single-dimensional association rule)和多維度關聯規則(multi-dimensional association rule);抽象層級型式的關聯法則-單一層級關聯規則(single-level association rule)和多層級關聯規則(multi-level association rule)。

三、分析層級程序法

陳慶益(2010)應用模糊層級分析法針對首次購屋者對於外部環境對於購屋決策的影響。賀振宇(2008)以層級分析法對高雄市的中產階級購屋者做研究，其探討的課題是購屋者對「住宅本身條件」、「交通設施」、「公共設施」、「嫌惡設施」、「整體經濟因素」、「政府政策」等六個評估購面做評估，希望藉此主觀性的購屋因素執行購屋決策。林士軒(2008)利用多準則決策技術(MCDM)的層級分析法建立計有國宅高齡者居住環境的改造評估系統。顏祈福(2007)利用模糊層級分析法對其家人做問券調查，其選擇的評估準則為總體環境因素、環境因素、生活機能、主觀、建築物條件(在本論文中為住宅屬性)。賴慧蓉(2007)認為過去建商主導的房地產市場已經無法滿足消費者需求，在「顧客導向」的趨勢下，將消費者購屋行為考量的主觀與客觀因素列入考量有其必要。賴慧蓉(2007)利用專家的建議，以四大構面要素、排名前五項影響購屋決策的評估準則做分析層級程序。王彥傑(2005)利用層級分析法與模糊層級分析法建立土地開發評估的層級結構，並求出各評估準則之權重。上述文獻多數使用分析層級程序法與模糊分析層級程序法做決策制定，而且其評估構面的選擇多數為專家訪談。本研究希望以較客觀的方法決定評估準則，所以利用資料探勘技術從歷史資料庫找出潛藏的價值資訊做為評估準則。

參、研究架構

本研究最終目標是建立一個「家戶購屋之購屋決策模式」，研究對象設定為台北市松山區、大安區及中正區三個行政區的已購屋者，透過實證分析與專家問卷調查兩種方式進行實驗，最後根據實驗結果建置出家戶購屋模式。

研究架構主要分為兩個部分，相對於國內、外多篇研究文獻是將可能影響購屋決策行為的因素細分為多個購屋決策因素，大部分文獻僅考慮確定性資料(deterministic data)(如:家戶戶籍資料、家戶人口數資料)對於家戶購屋決策行為的影響，因此本研究第一部分目的是從已購屋者的歷史交易資料中分析家戶購屋決策因素。第一部分使用家戶購屋歷史交易資料並應用兩種資料探勘相關技術-群集分析與關聯法則分析，進行市場區隔並尋找較受家戶重視的購屋決策因素及家戶購屋決策潛藏的有趣的樣式。然而，國內、外多數研究文獻僅考慮確定性資料對於家戶購屋決策行為的影響外，卻忽略心理因素對於家戶購屋決策制定可能所產生的影響。研究架構第二部分目的是將家戶購屋決策時可能會存有的心理因素列入考量，期望能以更全面的觀點探討家戶購屋決策。第二部分首先納入第一部分的實驗結果，將實驗結果設定為家戶購屋評估因素的一環，再透過

分析層級程序法將已購屋者的心理因素建置成層級架構來做一科學性的深入探討。

本研究主要目的是探討台北市部分行政區家戶購屋決策因素之關聯性異同與重要程度之排序。所以本研究最後將利用研究架構的兩部分實驗結果，建立一個貼近實際房地產交易市場的「家戶購屋之購屋決策模式」，並提供給房地產業者做為公司市場行銷策略擬定的參考指標，希望藉此幫助房地產業者對於具有關聯性的購屋決策因素有更深入的認識，同時幫助公司增加自身的競爭優勢。本研究之研究架構如圖 1 所示：

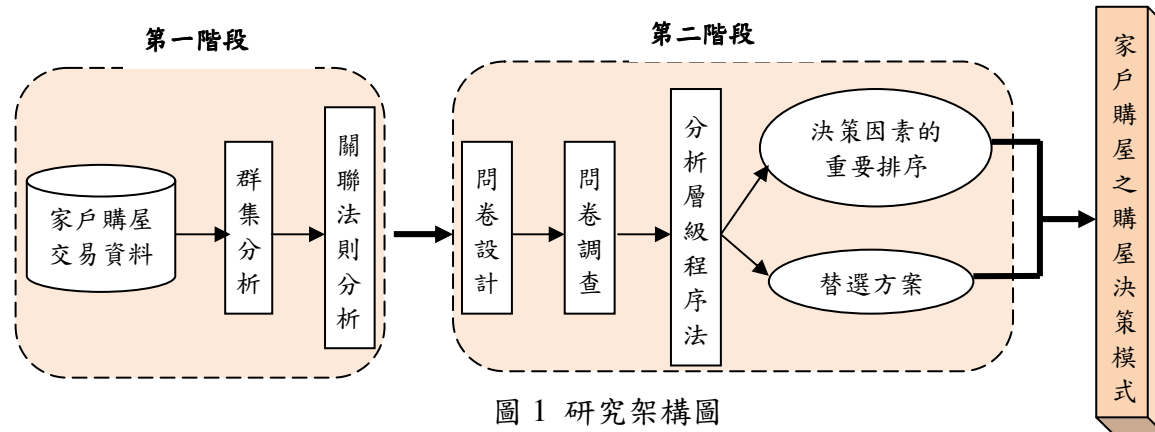


圖 1 研究架構圖

一、研究流程

本研究流程是將研究架構的兩大部分，分別設定為兩個階段以進行後續實驗。首先第一階段進行資料探勘，其中第一階段程序依序為資料蒐集、群集分析和關聯法則分析；第二階段則進行分析層級程序法，其中第二階段程序依序為問卷設計、問卷調查、執行分析層級程序法並對評估因素做重要程度的排序和選定替選方案。以下將對各個階段的研究程序做詳細的介紹。

(一) 第一階段實驗研究程序

本研究透過「台灣不動產交易中心」和台灣某房仲企業兩種管道，蒐集台北市松山區、中正區，和大安區三個行政區自 2010 年至 2011 年共二年的家戶購屋歷史交易資料，兩種管道的資料筆數總計 2024 筆，並且將資料建置到資料庫中。家戶購屋歷史交易資料內容包括行政區、物件類別、物件位置、權狀(坪)、成交價(萬)、格局、樓別、成交日、屋齡(年)、銷售天數，和車位等，但無包括人口統計變數資訊和安全便利性等可能會影響家戶購屋的決策因素，因此本研究將這些可能具有影響力的家戶購屋決策因素納入第二階段研究作為家戶購屋決策的評估準則。

本研究依據相關文獻-(曾憲雄等人 2007)和房仲業專員的專業知識進行資料轉換，透過不同的資料轉換方法進行市場區隔，以知曉各個行政區的目標顧客群及顧客群的購屋特性。最後由本研究者比較各種 *k*-means 分群結果並從中選出最佳的分群結果，實驗結果發現「運用資料探勘軟體」此資料轉換方式，依照同一路段門牌號碼(例如:1-50 號、51-100 號、101-150 號)將地址資料轉變成類別資料(例如:八德路二段 1-150 號資料轉換後變為 A1、八德路三段 1-150 號資料轉換後變為 A2)所產生的分群結果最佳。

期望能尋找較受家戶重視的購屋決策因素及家戶購屋決策潛藏的有趣的樣式，因此接續進行關聯法則分析，本研究採用的是 Apriori 演算法，希望透過關聯法則分析能找尋出哪些購屋因素是多數家戶購屋者會考慮的因素、哪些購屋因素之間具有關聯性，

又哪些購屋因素是多數家戶購屋者可能會同時考慮的。

(二) 第二階段實驗研究程序

第二階段首先納入第一階段的實驗結果，將實驗結果設定為「住宅內部屬性」次目標下的評估準則，並依據第二章「家戶購屋因素之探討」文獻回顧及研究者與房仲業專員訪談後的結果，透過分析層級程序法將已購屋者可能具有的心理因素建置成層級架構並設計成問卷。本研究問卷訪談對象設定為相關企業、房仲業專員和欲購屋一般民眾三種對象，進行專家知識與經驗的主觀判定，以驗證並確保本研究問卷之合理性與效用性，使本研究實驗架構更臻完整。

1. 問卷設計

本研究最終共整理出 22 個購屋評估因素，依照各個評估因素性質的不同分為五個次目標來探討，分別為購買管道、住宅內部屬性、住宅外部環境、家戶性質與住宅保值性。對於此五個次目標及各次目標的評估準則建立層級結構並設計成專家問卷，以進行後續的專家問卷調查。本研究的層級架構如圖 2 所示：

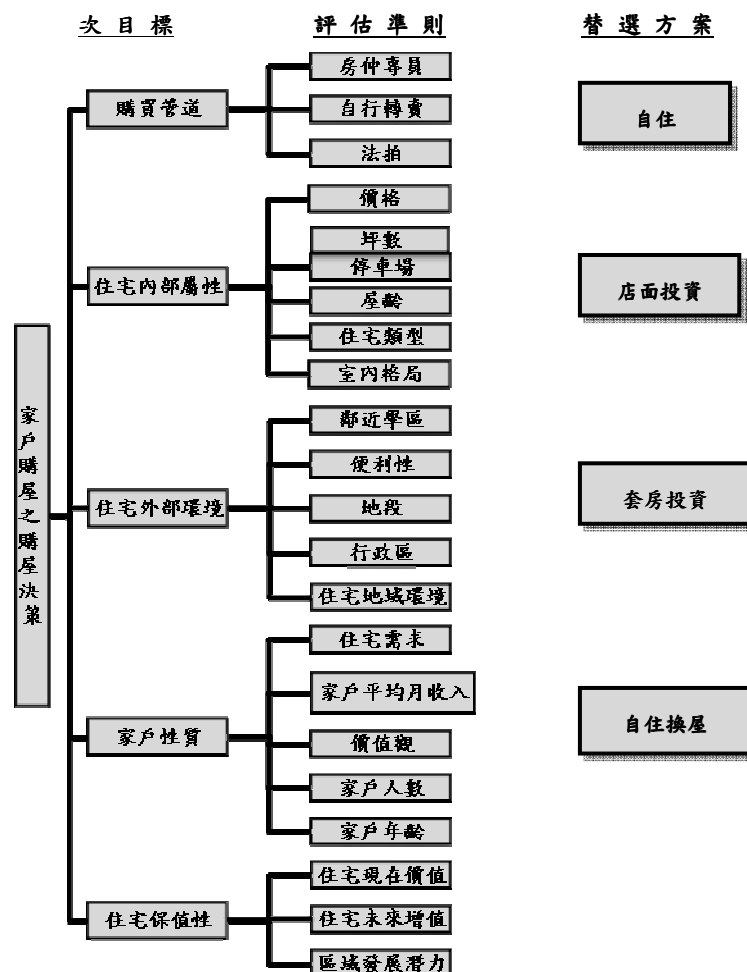


圖 2 整體層級架構圖

(1) 各層級要素的說明

a. 次目標

購買管道

購買管道是指家戶透過某種形式(如:電視廣告、報紙刊登、街頭發放的傳單或網路行銷)收集資訊並購得房屋，最後本研究整理家戶購屋管道發現，家戶購屋最終仍多偏向透過房仲業專員的協助，但也會透過其他的購屋管道(如:自行轉賣或法拍)。

住宅內部屬性

住宅內部屬性即房屋物件本身所帶有的條件，也是最與居住者息息相關的部份。而此次目標下的評估準則是來自於第一階段實驗之結果，本研究將關聯法則實驗結果剔除較無具關聯性的購屋評估因素(如:樓層、銷售日期，和交易天數)。

住宅外部環境

住宅外部環境主要指的是房屋物件所在地的週邊區域，較屬於房屋物件附帶的條件，本研究最後統整出鄰近學區、便利性、地段、行政區，和住宅地域環境等五個主要的評估準則。

家戶性質

家戶性質探討範圍包含價值觀、教育程度、先前生活環境、家戶購屋能力及家戶成員結構等，以上因素皆有可能影響最終家戶購屋決策的制定。然而有些因素較屬於主觀性因素(如:價值觀、教育程度，和先前生活環境);有些因素(如:家戶購屋能力)雖屬於量化資料，因牽涉到家戶隱私較難以取得真實資料。綜合上述原因，本研究將購屋者的家戶性質分為五個評估準則，透過問卷調查方式做深入探討。

住宅保值性

依據與房仲業專員訪談結果，知曉住宅保值性是投資購屋與首購自住者最明顯的區分點。因為多數首購自住的家戶購屋者較重視與居住相關的評估因素(如:房屋所在地點、價格，及坪數)，但投資的家戶購屋者卻較重視房屋的價值性，此是緣由多數投資購屋者屬於非首次購屋顧客，所以在購買第二屋時會考慮的購屋決策因素較有別於購買第一屋時會考量的決策因素。

b. 評估準則

次目標	評估準則	定義	準則屬性
購買管道	房仲專員	透過房仲業專員協助買賣房屋，若成功成交需付手續費給房仲業專員	質化
	自行轉賣	賣方自行尋找買方或買方自行連絡賣方，過程中無透過房仲業專員之協助	質化
	法拍	週轉不靈的房屋變由法院拍賣。此管道的特徵是房屋產權絕對清楚且房屋價格也較市價便宜	質化
住宅內部屬性	價格	房屋物件的出售價格	量化
	坪數	房屋所涵蓋的坪數，包括室內坪數、室外坪數	量化
	停車場	房屋是否包含停車位	質化
	屋齡	房屋的年齡，由屋齡低到高可將房屋分為全新屋、新屋與中古屋	量化

住宅內部屬性	住宅類型	房屋之用途，主要可區分為公寓、電梯大廈、套房、辦公室、店面，及土地	質化
	室內格局	房屋包含的房間間數、客廳廳數，和衛浴間數	量化
住宅外部環境	鄰近學區	房屋所在區域歸屬於哪一國小或國中之學區	質化
	便利性	房屋所在區域生活機能、購物、就醫、和搭乘交通工具的便利性	質化
	地段	房屋所在路段	質化
	行政區	房屋位於台北市之行政區域	質化
	住宅地域環境	房屋位於商業區、住宅區或住商混合區，此與房屋所在區域的安全性也具有相關性	質化
家戶性質	住宅需求	購屋者的購屋動機與目的	質化
	家戶平均月收入	評估購屋者能負擔的房貸之基準	量化
	價值觀	購屋者的家庭教育、成長背景及人文素養，此會影響購屋者看屋時的觀感	質化
	家戶人數	欲居住的人數，此會影響購屋者對坪數、格局或停車位之要求	量化
	家戶年齡	欲居住者之年齡會影響購屋者之住宅需求	量化
住宅保值性	住宅現在價值	房屋物件現在的房價	量化
	住宅未來價值	房屋物件未來的增值性，即房屋物件未來幾年的房價	量化
	區域發展潛力	房屋所在區域未來的發展潛力，例如：都市更新、捷運開通或地區繁榮性	質化

c. 替選方案

自住

自住的主要特色是家戶購屋者多關注於自身的購屋和居住需求，即較多家戶購屋者會考慮的是住宅內部屬性、住宅外部環境與個人性質等因素，反而較不會像投資的家戶購屋者對於購屋此一決策會存有更多的考慮因素(如:購買管道或住宅的保值性)。

店面投資

此種類型的家戶購屋者購屋主要目的即是投資，店面是投資最直接獲利的方式，因此房仲業專員會建議欲投資的家戶購屋者可直接購買店面。而店面投資最重要也最主要的特色是承租方可幫購屋者付房屋租賃稅，這也是為何投資的家戶購屋者會選擇店面投資方式的原因，所以說店面投資者多會考慮住宅外部環境及住宅的保值性等因素。

套房投資

此種類型家戶購屋者主要目的仍是投資，只是和店面投資不同處在於套房投資的承租方式是由多個承租方承擔。套房投資的特色在於多數家戶購屋者會購買大坪數且價格較便宜的一樓店面或購買電梯大廈建物中較小坪數的房屋，再由購屋者自行裝潢後分租給學生或社會人士。

自住換屋

和前述的「自住」替選方案不同的是，多數家戶購屋者會考慮自住換屋此一替選方案主要是因為居住空間不足或家庭人口增加等原因，才欲購買第二屋或是換較大坪數的住宅自住。所以說家戶購屋者會考慮自住換屋通常是因其自身具有相當的經濟能力且擁有購屋之經驗，是故自住換屋和自住兩替選方案相比家戶購屋者的購屋目的就已不同。

問卷調查結果

問卷調查結果方面是利用決策軟體求算各評估因素的重要程度，並且客觀地分析和評選出最佳的家戶購屋替代方案。最後利用兩階段的實驗結果，建立「家戶購屋之購屋決策模式」，提供給房地產業者做為公司市場行銷策略擬定的參考指標。

肆、 實驗結果

一、 第一階段實驗結果

本研究者使用「運用資料探勘軟體」此資料轉換方式進行後續的群集分析，使用 Clementine 軟體進行資料探勘分析，我們發現將 k 值設為 4 會有最佳的分群結果。經由執行 k -means 群集分析實驗後，軟體最後篩選出「坪數」、「價格」、「住宅類型」、「樓別」、「停車場」和「銷售日期」，此點說明無論是松山區、大安區或中正區，家戶購屋的購屋決策因素是相似的。

本研究再分別對於松山區、大安區及中正區進行關聯法則分析，首先將最低支持預設值設為 15%、最低信心度預設值設為 80%，利用 Apriori 演算法執行關聯法則分析，最後以 Lift 值由高至低將關聯法則作排序，且所有法則的 Lift 值皆大於 1，探勘後的實驗結果統整如下：

- (一) 從家戶購屋歷史交易資料中，我們發現最近兩年台北市此三個行政區家戶購屋的住宅類型較偏向購買「電梯大廈」，其次為「公寓」。而資料探勘後的結果也發現，在所有關聯法則中含有「電梯大廈」的法則數量明顯多於其他法則，其次為含有「公寓」的關聯法則。此點與創市際市場研究顧問公司於 2011 年底對「台灣地區購屋意向與物件需求網路調查」的調查結果相符合，更說明家戶購屋的重點有偏向「電梯大廈」的趨勢。
- (二) 在含有「電梯大廈」的關聯法則中，我們發現家戶購屋若決策購買電梯大廈的話，家戶的購屋決策因素會與以下因素具關聯性：屋齡-「全新屋(屋齡介於 5 年以下)」或「中古屋(屋齡介於 21 至 50 年之間)」；房屋結構-「1/1/1.0(房/廳/衛)」或「3/2/2.0(房/廳/衛)」；坪數-「小(25 坪以下)」或「中(26 至 45 坪以下)」；停車位-「無」或「有」，與價格-「中(1000 至 3000 萬元)」。

依據購屋因素在關聯法則中的排序性而言，本研究發現若法則前項的家戶購屋決策因素為「屋齡-全新屋」的話，則法則後項會為「坪數」、「價格」和「銷售日期」，此點說明家戶若決策購買「電梯大廈」，則購屋決策因素會先考慮「屋齡」，即家戶會先決策住宅給家戶自身的內、外部環境觀感，其次才為「坪數」和「價格」；若法則前項的家戶購屋決策因素為「房屋結構(房/廳/衛)」的話，則法則後項會為「坪數」和「屋齡」，此說明家戶購屋若先決策欲居住的空間是否足夠，則購屋決策因素會

先考慮「房屋結構」，其次才為「坪數」和「屋齡」；若法則前項的家戶購屋決策因素為「價格」的話，則法則後項會為「坪數」和「停車位」，此說明家戶購屋若先決策自身是否足夠負擔住宅價錢的話，購屋決策因素會先考慮「價格」，其次才會考慮「坪數」和「停車位」。

以「行政區」而言，本研究發現家戶購屋若選在「松山區」，較偏向選購位於電梯大廈內全新屋、坪數小且無停車位的房屋；而若家戶購屋選在「大安區」，較偏向選購位於電梯大廈內全新屋或中古屋、坪數中的房屋，卻無特別要求住宅是否附有停車位；而若家戶購屋選在「中正區」，較偏向選購位於電梯大廈內全新屋、坪數小、價格中且附有停車位的房屋。

- (三) 在含有「公寓」的關聯法則中，我們發現家戶購屋若決策購買公寓的話，家戶的購屋決策因素會與以下因素具關聯性：屋齡-「中古屋(屋齡介於 21 至 50 年之間)」；停車位-「無」，與價格-「中(1000 至 3000 萬元)」。

依據購屋因素在關聯法則中的排序性而言，本研究發現若法則前項的家戶購屋決策因素為「屋齡」的話，則法則後項會為「價格」，此點說明家戶若決策購買「公寓」，則購屋決策因素會先考慮「屋齡」，雖然依據資料我們瞭解公寓多偏向中古屋，但住宅在屋齡方面仍有差異，因此家戶仍會先決策住宅給家戶自身的內、外部環境觀感，其次才為「價格」；若法則前項的家戶購屋決策因素為「價格」的話，則法則後項會為「屋齡」，此說明家戶購屋會先決策自身是否足夠負擔公寓住宅價錢，所以購屋決策因素會先考慮「價格」，其次才會考慮住宅「屋齡」家戶自身是否能接受。上述是若「屋齡」或「價格」為關聯法則前項時的狀況，但根據我們探勘後的結果發現，包含「公寓」的關聯法則中，「住宅類型-公寓」為法則前項在所有法則中佔多數。而若法則前項的家戶購屋決策因素為「住宅類型-公寓」的話，則法則後項會為「屋齡」、「價格」、「停車位」和「銷售日期」。

以「行政區」而言，本研究發現家戶購屋若選在「松山區」，其次會選購住宅類型為公寓內的中古屋、價格中的房屋；而若家戶購屋選在「大安區」，其次會選購住宅類型為公寓，卻無特別要求住宅是否附有停車位。根據探勘的結果，本研究發現家戶購屋選在「中正區」的法則極少，此點說明購屋若選在「中正區」，較偏向選購電梯大廈的住宅類型。

- (四) 若關聯法則無含有任何「住宅類型」的法則，則我們發現家戶購屋的購屋決策因素會與以下因素具關聯性：屋齡-「全新屋(屋齡介於 5 年以下)」或「中古屋(屋齡介於 21 至 50 年之間)」；房屋結構-「1/1/1.0(房/廳/衛)」或「3/2/2.0(房/廳/衛)」；坪數-「小(25 坪以下)」或「中(26 至 45 坪以下)」；停車位-「無」，與價格-「小(1000 萬元以下)」或「中(1000 至 3000 萬元)」。

伍、 結論

有鑑於房地產市場失衡的狀態，促使家戶的購屋決策行為漸成為重要的研究課題，其中關於購屋決策因素即刻受到重視，因此本研究主要是探討台北市家戶購屋的購屋決

策因素，設定研究對象為台北市的松山區、大安區及中山區此三個行政區域。

本研究發展一個二階段實驗方法以探討購屋決策因素。第一階段群集分析的實驗結果中，本研究者發現「樓別」在關聯法則分析中並無被探勘出來，此點說明「樓別」在松山區、大安區或中正區是家戶購屋會考慮的購屋決策因素，但「樓別」此一購屋決策因素和其他購屋決策因素之間並無具備排序性。從第一階段關聯法則分析的實驗結果中，我們發現若家戶購屋首重「行政區」、次重「住宅類型」，則依序會重視的購屋因素為「屋齡」、「坪數」和「價格」；若家戶購屋首重「行政區」、次重「住宅類型」，則依序會重視的購屋因素為「屋齡」、「坪數」和「價格」，而此兩點闡述「屋齡」與「房屋結構」具有相輔相成的關係，家戶購屋會先在兩點間決策其一後，才會再考慮「坪數」和「價格」等的購屋決策因素。而此結果與創市際網路顧問公司在 2009 年底對於「台灣地區購屋意向與物件需求網路調查」的調查結果相異，因此我們可以說隨著時間的推移，台北市松山區、大安區及中正區家戶購屋第三重視的購屋決策因素為「屋齡」或「房屋結構」，也就是說現代人重視房屋的住屋內、外部環境，勝過過去家戶購屋較重視住屋「坪數」此一購屋決策因素。但若家戶購屋無特別指定所需求的「住宅類型」，則家戶購屋的決策因素會變為，首重「行政區」、次重「屋齡」，依序會重視的購屋因素則為「坪數」、「價格」和「停車場」。因此我們統整台北市此三個行政區的家戶購屋決策因素，發現家戶購屋首重「行政區」，其次為「住宅類型」、「屋齡」、「房屋結構」、「坪數」、「價格」和「停車位」。

接續，首先將第一階段資料探勘的實驗結果納入第二階段實驗，設定為家戶購屋評估因素的一環，透過分析層級程序法將已購屋者的心理因素建置成層級架構並設計成問卷，我們將發放問卷給企業、房仲業專員及預購屋一般民眾以探討家戶購屋的主觀性心理因素，並將問卷調查結果透過分析層級程序法來做一科學性的深入探討與分析，本研究期望能把實驗結果建置成「家戶購屋之購屋決策模式」，並提供給房地產業者做為市場策略擬定之參考。

參考文獻

1. 吳旭智、賴淑貞譯，Berry, M. J. A., & Linoff, G. S., 著，2001，資料採礦理論與實務-顧客關係管理的技巧與科學，台北:維科圖書有限公司。
2. 林秋瑾、王健安、張金鶚，1996，『房地產景氣與總體經濟景氣於時間上領先、同時、落後關係之探討』，國科會人文及社會科學彙刊，第七卷，第一期:35~56 頁。
3. 韋端、鄭宇庭、鄧家駒、匡宏波、謝邦昌，2003，Data Mining 概述-以 Clementine 7.0 為例，台北:中華資料採礦協會。
4. 陳彥仲，1997，『住宅選擇之程序性決策模式』，住宅學報，第五期:37~49 頁。
5. 陳彥仲、吳京玲，1998，『家戶住宅區位選擇與地方財政分配之實證研究』，都市與計劃，第二十五卷，第二期:223~238 頁。
6. 陳淑美、張金鶚，2002，『家戶遷移決策與路徑選擇之研究——台北縣市的實證研究』，住宅學報，第十一卷，第一期:1~22 頁。
7. 連經宇，2002，『家戶購屋決策影響因素之初探研究——結合模糊語意的因素分析法

- 之實證比較』，台灣土地研究，第四期:29~51 頁。
8. 連經宇、陳彥仲，2005，『結合模糊語意尺度法與不連續選擇理論建立家戶住宅消費選擇行為模式之研究』，都市與計劃，第三十二卷，第一期:57~81 頁。
 9. 戚靜玟，1998，『住宅市場區隔與購屋行為之研究』，台灣科技大學管理技術研究所碩士論文，未出版。
 10. 曾憲雄、蔡秀滿、蘇東興、曾秋蓉、王慶堯，2007，資料探勘(Data Mining)，台北:旗標出版股份有限公司。
 11. 曾龍譯，Jiawei, H. & Micheline, K. 著，2003，資料採礦-概念與技術，台北:維科圖書有限公司。
 12. 彭建文、張金鶚，2000，『總體經濟對房地產景氣影響之研究』，國家科學委員會研究會刊:人文及社會科學，第十卷，第三期:330~343 頁。
 13. 創市際市場研究顧問公司，11 月，2011 (<http://www.insightexplorer.com/index.html>)
 14. 薛立敏、曾喜鵬、陳雅君，2002，『區域人口遷入台北都會區地點選擇之實證研究』，住宅學報，第十一卷，第二期:159~178 頁。
 15. 謝文盛、林素菁，2000，『租稅效果對住宅租買選擇影響之分析』，住宅學報，第九卷，第一期:1~17 頁。
 16. Agrawal, R., Imielinski, T., and Swami, A., "Mining association rules between sets of items in large databases," In Proceeding of the *ACM SIGMOD international conference on management of data*, 1993, pp. 207-216.
 17. Assadian, A. & Ondrich, J., "Residential Location, housing Demand and Labor Supply Decisions of One- and Two-Earner Household: The Case of Bogota, Colombia," *Urban Studies*(30:1), 1993, pp.73-86.
 18. Beatty, S. E. & S. M. Smith, "External Search Effort: An Investigation Across Several Product Categories," *Journal of Consumer Research*(14), 1987, pp. 83-95.
 19. Cheng, C. H., & Chen, Y. S., "Classifying the segmentation of customer value via RFM model and RS theory," *Expert Systems with Applications*(36), 2009, pp. 4176-4184.
 20. Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R. (Eds.), "Advances in Knowledge Discovery and Data Mining," MA: The MIT Press, 1996.
 21. Follain, Jr. J. R. and Jimenez, E., "The Demand for Housing Characteristics in Development Countries," *Urban Studies*(22:4), 1985, pp. 421-432.
 22. Hansen, A.H. "Business cycle and national income", 1951, p. 40.
 23. Lancaster, K. J., "A New Approach to Consumer Theory," *Journal of Political Economy*(74:2), 1966, pp.132-157.
 24. Li, M. M., "A logit model of homeownership", *Econometrica*(45:5), 1977, pp. 1081-1091.
 25. Matthews, R.C.O., "The trade cycles," 1969, pp. 104-106.
 26. McLeod, P. B. & Ellis, J. R., "Alternative Approaches to the Family Life Cycle in the Analysis of Housing Consumption," *Journal of Marriage and the Family*(45:3), 1983, pp. 699-708.

27. Rajasethupathy, K., Scime, A., Rajasethupathy, K. S., and Murray, G. R., "Finding persistent rules: Combining association and classification results," *Expert Systems with Applications*(36:3), 2009, pp. 6019-6024.
28. Rosen S., "Hedonic Prices and Implicit Markets: Product Differentiation Pure Competition," *Journal of Political Economy*(32:1), 1974, pp. 34-35.
29. Snchez, D., Vila, M. A., Cerda, L., and Serrano, J. M. "Association rules applied to credit card fraud detection," *Expert Systems with Applications*(36:2), 2009, pp. 3630-3640.
30. Wesley Changchien, S., & Lu, T-C., "Mining association rules procedure to support on-line recommendation by customers and products fragmentation," *Expert Systems with Applications*(20:4), 2001, pp. 325-335.

A study of house buying factor by applying Data Mining technique and Analytical Hierarchy Process

SUNG-SHUN WENG

Graduate Institute of Information and Logistics Management
National Taipei University Technology
wengss@ntut.edu.tw

WEI-LING LIEN

Graduate Institute of Information and Logistics Management
National Taipei University Technology
lianbarbara1006@gmail.com

Abstract

Most of the literatures about house buying factor only considered deterministic data to prompt to make house buying decision, however neglected the influence incurred by physical-factors. Moreover, many literatures used subjective factors, for example, expert method. Such selection incurred ignoring potential possible factors. This manuscript develops a two-phase experiment to investigate house buying factor. Phase one, we use the k -means algorithm based on Data Mining (DM) technique to do customer-division, then search for potential interested-patterns by the Associating Analysis based on DM. Phase two, the Analytical Hierarchy Process (AHP) is applied to make the house buying decisions with the result of phase one and the one of physical factor. This manuscript intends to investigate related housing factors and its corresponding ranking for further application.

Keywords: House buying factors, Data Mining (DM), Analytical Hierarchy Process (AHP)