

東海大學資訊管理研究所  
碩士學位論文

以經濟與社會面議題探討影響台中市房價之因素

**Inverstigating the factors to affect the housing prices in  
Taichung among the economic and social perspectives**

指導教授：姜自強 博士

研 究 生：吳昕穎 撰

中華民國 107 年 6 月

東海大學資訊管理學系碩士學位  
考試委員審定書

資訊管理學系研究所 吳昕穎 君所提之論文

以經濟與社會面議題探討影響台中市房價之因素

經本考試委員會審查，符合碩士資格標準。

學位考試委員會 召集人：陳柏言 (簽章)

委員：鄭永真

楊朝棟

姜再強

呂峻益

中華民國 107 年 6 月 19 日

## 誌謝

於東海大學資訊管理所求學期間，感謝許多人的支持和鼓勵，讓我能順利完成學業，首先要感謝我的父母支持我讀碩士班的決定，讓我能夠在沒有經濟壓力的情況下完成學業，也謝謝我的弟弟幫忙照顧、陪伴父母，讓我在異地念書也不用擔心父母健康以及家裡發生什麼事，非常開心可以在這個校園給姜自強博士當指導學生，老師總是苦口婆心的告訴我需要修改的地方以及論文的方向，引導我完成論文，也給予很多參與各項活動和比賽的機會及支持，在活動和比賽中學到了很多，讓我順利在時間內畢業。也謝謝同是碩士生的俊昭在熬夜的碩士生涯中一起熬夜努力、給予支持和鼓勵以及佳鴻、映瑋、家廣在論文上給予許多幫助，還有與我相伴兩年的同班同學彥齊、復榮、彥鋒、文彬、婉婷、煥豪、岳穎一起努力在很多個熬夜的夜晚完成各種報告、作業，還有各種活動讓我的碩士生涯並不孤單。另外也感謝我的好朋友美青、佩珊、映瑋、齡勻和嘉齡，在我想要偷懶放棄的時候，不斷督促我，為我加油打氣，聽我訴苦並給予許多幫助，很開心可以順利在時間內畢業，感謝大家。



論文名稱：以經濟與社會面議題探討影響台中市房價之因素

校所名稱：東海大學資訊管理學系研究所

畢業時間：2018年6月

研究生： 吳昕穎

指導教授：姜自強

論文摘要：

華人傳統觀念總是認為要擁有房產才有保障，因此在台灣，房地產一直是討論度極高的議題，而房價是人們在購置房地產中首先會考量的因素；而台中市處於台灣中部的精華地段，優點更是多到是不可勝言，在 2010 年縣市合併之後，台中房價的變動也成為人們關心的話題，因此本研究選擇台中做為探討影響房價因素的地區，下載內政部不動產交易實價查詢服務網提供之台中三區 2014-2016 年約 10 萬筆的房價資料和三個面向分別為房屋本身建物型態、社會面以及經濟面使用 SPSS 做 Apriori 演算法關聯分析以及決策樹分析，期望能幫助一般民眾在選購房地產上能夠有更多參考的指標和更多樣的選擇以及了解影響台中房價的因素為何。

關鍵詞：房價、Apriori 演算法、關聯分析、資料探勘、社會、經濟

Title of Thesis : Discussing the factors affecting housing prices in Taichung City on economic and social issues

Name of Institute: Tunghai University, Graduate Institute of Information Management

Graduation Time : ( 06 / 2018 )

Student Name : Hsin-Ying Wu

Advisor Name : Tzu-Chiang Chiang

Abstract :

Chinese Traditional ideas always believe that only real estate is necessary. Therefore, in Taiwan, the real estate has always been a highly discussed topic, and house prices are the factors that people will consider first when purchasing real estate; Taichung City is the best place in central Taiwan, but its advantages are even more worth mentioning. After the merger of counties and cities in 2010, the change in the house prices of Taichung become a topic of concern. Therefore, this study selected the Taichung as a factor to investigate the impact of house prices. Use SPSS and download the information about the house prices in Taichung from 2014 to 2016 of about 100,000 records provided by the Ministry of the Interior real estate transaction price inquiry service network, Use it for Apriori correlation analysis and Decision tree analysis. The expectation can help the general public to have more reference projects in purchasing real estate and understand the factors that affect the prices in Taichung.

Keywords : Housing prices, Apriori algorithm, correlation analysis, data mining, social, economic

# 目錄

目錄 .....	III
表次 .....	V
圖次 .....	VI
第一章 緒論 .....	1
第一節 研究背景與動機 .....	1
第二節 研究目的 .....	3
第三節 研究對象 .....	3
第二章 文獻探討 .....	4
第一節 房地產定義 .....	4
第二節 影響房價因素 .....	5
第三節 資料探勘 .....	7
第四節 不動產實價登錄說明 .....	9
第三章 研究方法 .....	10
第一節 研究架構 .....	10
第二節 資料處理 .....	10
第三節 Apriori 關聯分析 .....	12
第四節 CART 決策樹分析 .....	13
第四章 分析與討論 .....	14
第一節 資料集敘述統計和變數 .....	14
第二節 關聯分析 .....	18
壹、以房屋本身格局 .....	18
貳、以社會面 .....	25
第三節 決策樹分析 .....	26
壹、房屋格局、建物型態 .....	26

貳、家庭收支調查報告 .....	28
第五章 結論與建議 .....	29
參考文獻 .....	30



## 表次

表 2-1 不動產實價登錄彙總表.....	9
表 3-1 本研究資料處理表.....	11
表 4-1 六區分配筆數.....	17
表 4-2 六區平均每平方公尺價格.....	18
表 4-3 區域一關聯分析表.....	19
表 4-4 區域二關聯分析.....	20
表 4-5 區域三關聯分析表.....	21
表 4-6 區域四關聯分析表.....	22
表 4-7 區域五關聯分析表.....	23
表 4-8 區域六關聯分析表.....	24
表 4-9 結婚、離婚對數關聯分析表.....	25
表 4-10 社會移動人口關聯分析表.....	25

## 圖次

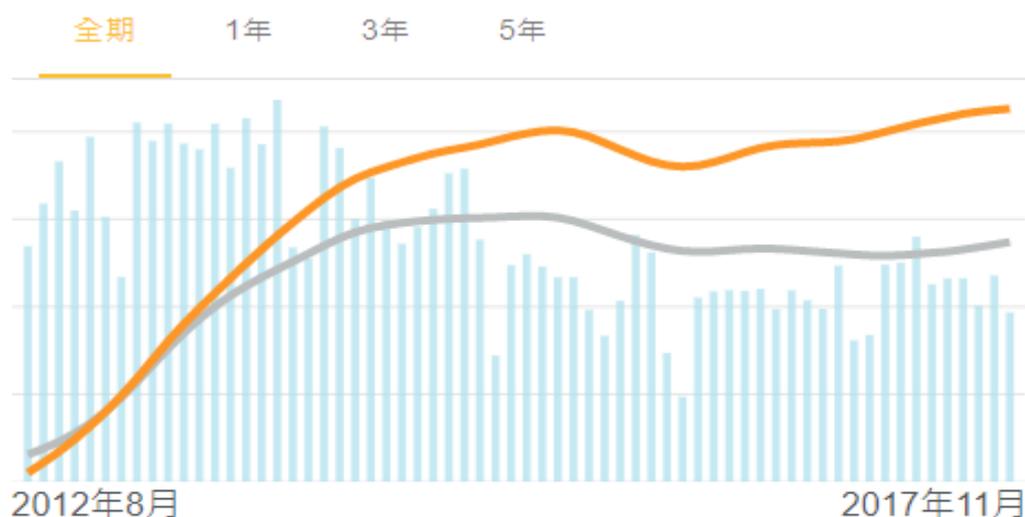
圖 1-1 台中市價量變動趨勢圖.....	1
圖 2-1 經濟景氣與房地產景氣對照圖.....	6
圖 2-2 監督式與非監督式資料探勘圖.....	7
圖 2-3 KDD 流程圖.....	8
圖 3-1 本研究架構圖.....	10
圖 4-1 2014 到 2016 建物型態圓餅圖.....	14
圖 4-2 房間數量直方圖.....	15
圖 4-3 客廳數量直方圖.....	15
圖 4-4 衛浴數量直方圖.....	16
圖 4-5 房屋格局型態決策樹分析圖.....	26
圖 4-6 預測變數重要性圖.....	27
圖 4-7 所得收入決策樹分析圖.....	28

# 第一章 緒論

## 第一節 研究背景與動機

台中市處於台灣中部的精華地段，優點更是多到是不可勝言，不僅物產豐饒、氣候適中、商圈內容豐富且各有特色以及公共建設充足，交通建設也十分便利，藝術文化活動中心林立還有許多美麗的景點，新台灣國策智庫（2017）公布民調，台中市連續在 2016、2017 年在「宜居城市」中排名第一，超過八成的台中民眾選擇台中市是六都中最適合人居住的城市，另外台中市民對於台中市的光榮感也是最高的，不僅如此，美國有線電視網 CNN 也在 2016 年在官網刊登了一篇說明台中優點以及將台中評為台灣的最宜居都市，在 2010 年縣市合併之後，台中房價的變動也成為人們關心的話題。台中市建築開發商業同業公會理事長暨豐邑建設總經理邱崇喆（2017）認為台灣整體經濟環境有開始上揚，加上台中市的各項交通建設逐漸成型，房市可望平穩、樂觀。且 2008 年金融海嘯之後，各國的國際貨幣呈貶值狀態，造成通膨壓力大，不動產仍是最好的投資工具。台中

### 台中市價量變動趨勢



市建築經營協會理事長暨寶鯨建設總經理許枝旺（2017）

圖 1-1 台中市價量變動趨勢圖

資料來源:好時價 House+網

也認為相對於台北房市急凍情況，台中的房市雖然量縮，但房價還是維持平盤。本研究希望從各方面探討影響房價的因素。

此外在現今資訊流通的社會，不管我們做了什麼事情、買了什麼東西、看了什麼電視節目，甚至是身體狀況心跳等等，背後都會產生數據，現今資訊流通快速、網路發達便利，網路上有許多唾手可得的開放資料供我們查詢並且下載使用，開放資料（Open data）指的是一種經過許可認證並在網路上提供民眾下載的資料。主要來源可能來自不同領域，而這些資料不受著作權、專利權等所管制，可提供給一般社會民眾做參略，任何人都可以自由出版使用，不論是要拿來出版或是做其他的運用都不加以限制，這些開放資料的背後往往藏著許多肉眼看不出來的寶藏，我們可以使用「資料探勘」來使這些開放資料具有更多意義，現在連資料探勘的軟體及套件也都有很多種可以供我們免費或著付費下載做使用，本研究使用的第一種軟體 RStudio 為 R 語言的一種整合開發環境，而且是免費的，裡面有超過一萬以上的套件可使用，且每分每秒持續增加中，另外一種軟體則為 SPSS Modeler，SPSS 是統計產品與服務解決方案（Statistical Product and Service Solutions）的簡稱，為 IBM 推出的一系列用於數據挖掘、決策支持、統計學分析運算、預測分析等用途的軟體以及相關服務的其中一種，本研究使用 SPSS Modeler 以內政部不動產交易實價查詢服務網提供之全台中 2014-2016 年約 10 萬筆的房價資料做 K-means 集群分析再利用 RStudio 以台中市政府民政局提供之全台中 2014-2016 年人口遷入遷出統計資料和人口數資料、台中市政府主計處提供之 2014-2016 年台中家庭收支訪問調查資料和內政部不動產交易實價查詢服務網提供之全台中 2014-2016 年約 10 萬筆的房價資料中的建物型態以及房價資料做關聯分析並利用內政部不動產交易實價查詢服務網提供之全台中 2014-2016 約 10 萬筆的房價資料以房屋本身的各項條件來做決策樹分析，利用三個資料探勘的方法去探討影響房價之因素。

## 第二節 研究目的

在台灣，房價一直是討論度很高的議題，台灣人受中國文化「有土斯有財」觀念的影響，希望藉由不動產來堆積自己的財富，認為要掌握不動產在手裡才會有保障，不管是自住抑或著是投資，購置房產都是台灣人人生中幾乎可以說是必須的一件事，而房價在民眾買房時是很重要的一個因素，往往為考慮是否買房的重點之一，本研究期望透過此篇研究使用三種較常使用的資料探勘方式，去了解一部分影響台中房價的原因，並且期望讓台中民眾可以更加的了解台中房屋市場的變動趨勢以及影響台中房價的原因為何，並可以從中選擇房價與自己所期望的理想房屋平衡的房屋，另外也可以讓房屋建商了解民眾真正需要的房子是什麼樣的房子，讓建商朝大多台中民眾理想的房屋方式去建屋，不僅民眾可以買到自己理想的房屋，建商也可以了解什麼房子才能賣得出去，獲得雙贏的結果。

## 第三節 研究對象

本研究利用內政部不動產交易實價查詢服務網提供之全台中 2014-2016 約 10 萬筆的房價資料以房價為目標分別以三個面向來做關聯分析，第一個為房屋本身型態，使用房價資料和建物現況格局.房、建物現況格局.隔間、建物現況格局.衛、建物現況格局.廳、有無管理組織、建物型態以及有無車位做關聯分析；第二個則是使用房價資料和台中主計處提供之 2014-2016 年結婚對數、離婚對數和人口遷入遷出社會移動指標做關聯分析；最後將台中行政區以台中主計處家庭收支報告區分之六個區域和 2014-2016 年每月家庭收支報告統計資料做決策樹分析，期望能透過從三種面向去做探勘，能夠更加瞭解影響台中房價的因素。

## 第二章 文獻探討

### 第一節 房地產定義

房地產一般定義為覆蓋土地並永久附著於土地的一類實物，根據民法第 66 條規定：「稱不動產者，謂土地及其定著物。」也就是說包含了土地以及房屋本體，雖然在定義上略有不同，但一般而言房地產也會和不動產一詞互相通用，因此，本研究將內政部不動產交易實價查詢服務網提供之全台中 2014-2016 的房價資料中篩選出內含房屋+土地的資料約 10 萬筆的資料來做為主要的房價分析對象。另外，過去研究者蔡曜如（2003）也指出房地產一詞對於一般大眾及民間業者來說是房屋與土地的財產合稱，蔡育政（1999）於影響房地產價格因素之研究以台中市北屯區、西屯區、南屯區、中區、東區為例中將房地產的定義歸納為以下六項：

- 壹、 不可移動性:建築物或是產權依附在土地之上，土地本身有不可移動性
- 貳、 異質性:因為各種條件，讓房地產的特性會不一樣。顏聰玲(2004)指出房地產的價值，受到土地使用分區、密度管制及周邊公共設施多寡的影響，因為如此，房地產隨著地形地物關係以及效度上使用的不同，皆有其各自的價值以及差異性。
- 參、 耐久性:建築物大多有五、六十年以上的使用期，土地之使用期限也並不受限，此特性加上第一種的不可移動性，相對其他商品更多的保值價值。
- 肆、 昂貴性:房地產有高資本、投資的價值，一般家庭如需買房不足金額時，須向金融機構辦理購物貸款。
- 伍、 投資與消費之雙重性:房地產除了居住以外，很多人也會拿來當作投資的一種工具。
- 陸、 所有權之可分割性:房地產除了能夠屬於個人所有權之外，也可以由法律及契約形式的變更創造出其他所有權的型態。

## 第二節 影響房價因素

房地產界流傳著一句話：「沒有賣不掉的房子，只有賣不掉的價錢」，另外還有另外一句話則是：「房子真正的行情、價錢，只有最後實際成交的價格」在購買房屋民眾最主要考慮的因素大多為房價，影響房價的因素很多，本研究參考過去研究者文獻如下，林祖嘉&林素菁（1993）認為一般而言，影響房價之主要因素分做三大類：第一，與房屋本身所具之特性有關之因素，如坪數、房間數目等；第二，與房屋座落之該棟樓房有關，如獨棟式、公寓式、大樓式等等；第三，與房屋所在之鄰近環境特質有關。而 Lancaster（1966）則認為消費者購買財貨其效用是來自財貨所具有的各项特徵與數量，相同的物品可能有不同的特徵，不同的物品也可能有相同的特徵，不同的特徵會給消費者不同的效用，因而產生不同的價格，本研究以此為依據選出建物現況格局、房、建物現況格局、隔間、有無管理組織、建物型態以及有無車位來做 Apriori 關聯分析。且在經濟方面，楊鴻謙(2001)於不動產投資與市場分析一書中提到當國民之所得開始增加時，也會使國民的購買力開始增加，並且當國民提升生活品質後，會進而開始對不動產的需求增加，會帶動不動產的價格上升的情況，且經濟成長，民眾會開始將生活品質視為重要的一點，也會對不動產的需求增加產生一定的影響，另外，當經濟成長率為高時，像是民間的消費與產業的投資，也會帶動不動產市場之景氣。從景氣循環的觀點及不動產市場經驗中我們可以知道，經濟衰退和不動產有著時間落差，如圖 2-1 所示，當經濟景氣衰退 3-6 個月之後，不動產才會漸漸開始慢慢走入衰退，只有當經濟成長開始復甦時，不動產的價格接近同時上升，且陳幸宜(2003)於都市房價變動影響因素之系統動態模擬中提到通貨膨脹和物價水準變動是息息相關的，物價會影響貨幣的實際購買力，令消費者的實際購買力減低，另外當營建相關成本上漲時，融資的相關成本也在同時上升。會發生通貨膨脹的因素很多，主要將其分為「成本推動型」及「需求拉動型」兩種，最近幾年來國際原油價格上漲造成原物料之價格上漲，這屬於「成本推動型」；經濟的高成長之下，因為各種消費商品需求的上升，進而發生之通貨膨脹就屬於「需求拉動型」。經濟的成長會讓房地產、不動產的需求上升，和房地產、不動產相關的商品與勞務價格上揚，房價自然會上漲，這就是需求的拉動，也會帶動原物料像是鋼鐵等價格的上升，使得房屋生產的增加，並且引起成本推動的物品價格上升。由此可以看出經濟方面各項指

數都對房地產市場價格的變動具有相當大的影響。

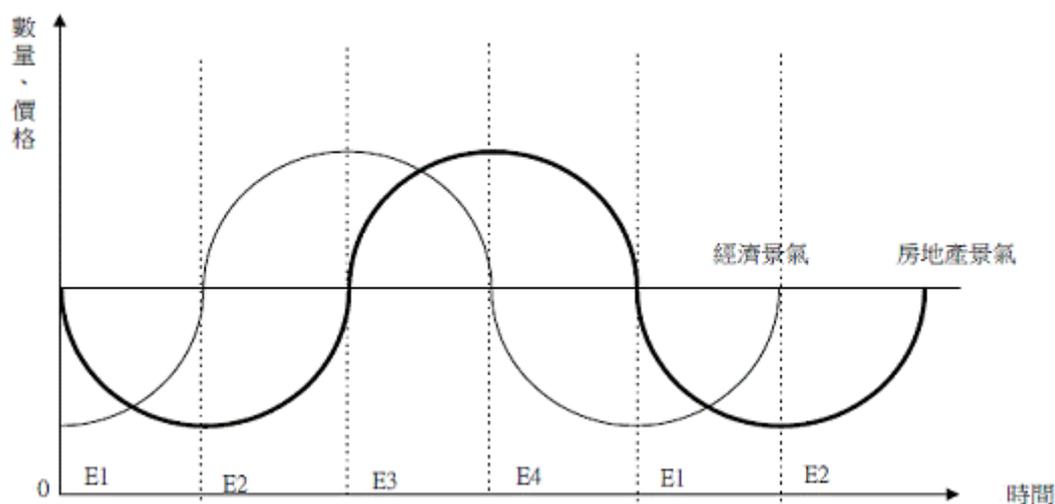


圖 2-1 經濟景氣與房地產景氣對照圖

資料來源: 楊鴻謙(2001), 不動產投資與市場分析

另外在社會面方面，顏聰玲(2004)於不動產市場分析與預測一書中提到當一個區域內的人口量增加則不動產的需求也會增加，人口會一直增加，但土地並不會，當人口慢慢增加且集中在某個地方，因為不動產的「不可增加性」的關係，會導致人口和房地產產生供給以及需求之失衡，因此會使得該地區不動產之價格上升；另外，在人口組成方面，人口組成包含了婚姻狀態、性別以及年齡層的分布，如果這個地方的幼年、孩童人口與老人人口比例比較多的話，則不具有購買房屋的能力，如果青年男女超過結婚年齡卻還沒有結婚，就會增加像是套房之類的小坪數住宅等等的的需求；至於家戶結構，是指組成一個家庭的人口數，在家戶結構中是不是為三代同堂或著是小家庭的家庭處於多數，上述這些原因都可以用來推估住宅的需求量。另外在人口分布與遷徙方面，社會移動人口應會對房地產有影響，台灣分為北部、中部、南部以及東部，東部因大多為山區，資源有限且開發不易，工商業在當地並不發達，是人口外流最嚴重的區域，當地的房價相對來說也較低，而北、中、南三區，在西部較易開發且為政治經濟中心、工商產業的發展重地以及金融服務業中心，分別在三區形成高雄、台中及台北三大都會區，特別是台北為政治經濟中心，房價更是高得嚇死人。

### 第三節 資料探勘

資料探勘為一種挖掘資料背後含意的技術，學者定義如下：Frawley (1911) 認為是資料庫中挖掘出明確、潛在而且是很有用之資訊的過程。而 Grupe & Owrang (1995) 則認為是從已經存在的資料庫之中挖掘出專家仍然不知道的新資訊、事實。Berry & Linoff (1997) 則認為是使用半自動或自動的方法，對海量、大量資料分析，從中尋找量有意義之法則或關係。資料探勘的方式概分為監督式資料探勘以及非監督式探勘，監督式使用在已經知道要挖掘的方向時，

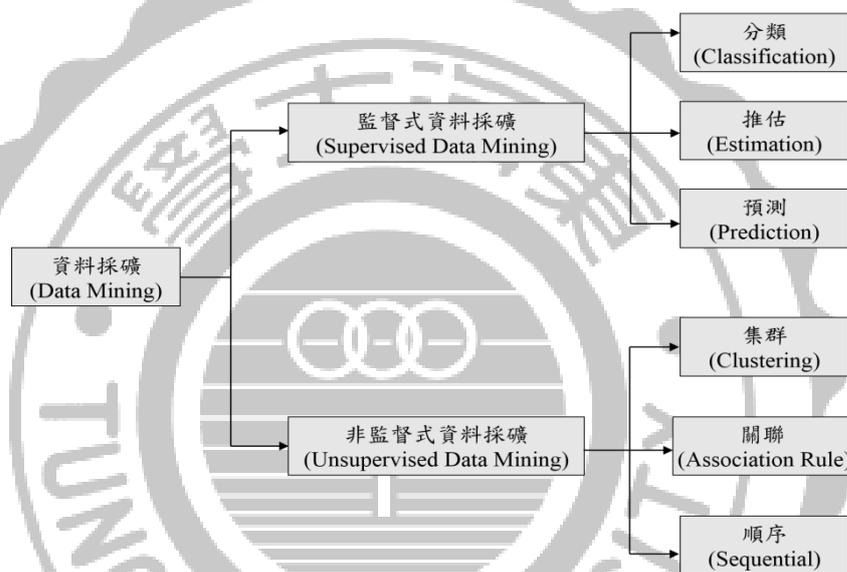


圖 2-2 監督式與非監督式資料探勘圖

來源：廖述賢&溫志皓,資料探勘理論與應用 P2-2

而非監督式則用於探詢一個問題的背後，究竟被那些變數影響（廖述賢&溫志皓,2012），本研究使用的 Apriori 演算法屬於非監督式的關聯方法中的一種，CART 決策樹則為監督式分類方法之一。另外，Han&Kamber（2000）將資料探勘廣義定義為資料庫之知識發掘（Knowledge Discovery in Databases，簡稱 KDD），下圖為 KDD 的流程圖：

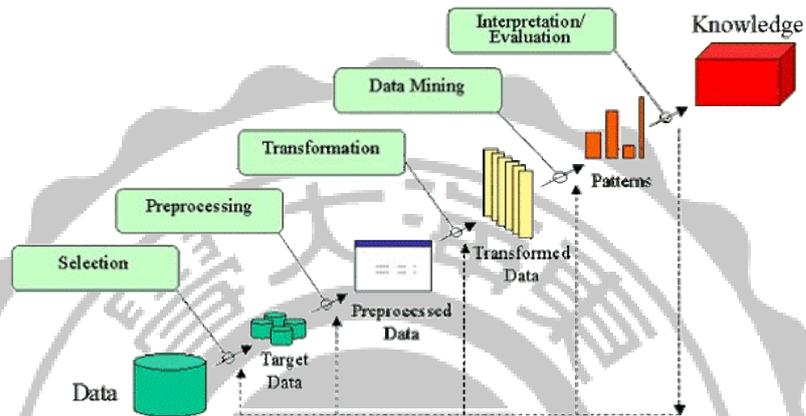


圖 2-3 KDD 流程圖

資料來源:Han&Kamber DATA MINING Concepts and Techniques

#### 第四節 不動產實價登錄說明

蔡育政(1999)於影響房地產價格因素之研究以台中市北屯區、西屯區、南屯區、中區、東區為例中提到過去的房地產價格資訊不易透明化，因過去房地產分散於各個地區，交易大多為買賣雙方自己私底下議價，使得交易過程不透明，近年來，透過資訊及媒體發展，房地產的買賣資料才逐漸透明化。林崇詠(2015)於影響房地產價格因素之研究以臺中市西屯區為例中提到內政部地政司一直在不斷的修法，像是地政士法、平均地權條例以及不動產經紀業管理條例等法，推動著不動產實價登錄的制度產生，規範不動產的交易應該在買、賣結束後，完成所有權轉移登記的 30 天以內，跟主管機關登記建物及土地成交的案件之實際交易價格等資料，全面一步一步獲得不動產買賣交易的實價資料，並整理出如圖 2-4 之不動產實價登錄彙總表：

表 2-1 不動產實價登錄彙總表

項目	內容
登錄範圍	地政事務所所收之買、賣登記申請的案件。 不動產仲介經紀業簽訂租賃契約書的案件。 不動產代銷經紀業簽有起造人或建築業委託代銷的成交案件。
登錄者	代銷業者、不動產經紀業、地政士若沒有委託地政士登記也沒有委託代理成交者或經紀業居間，將由權利人（買方）申報，賣方沒有申報之義務
登錄資訊	標的資訊、價格資訊及交易標的
登錄期限	買、賣辦理移轉登記後的三十日以內
處罰規定	違反申報建物成交案件及登錄土地實際資訊義務，將處以三萬至十五萬元罰鍰。
開放查詢	開放查詢去識化、去區段化方式

資料來源：林崇詠(2015)，影響房地產價格因素之研究以臺中市西屯區為例整理

## 第三章 研究方法

### 第一節 研究架構

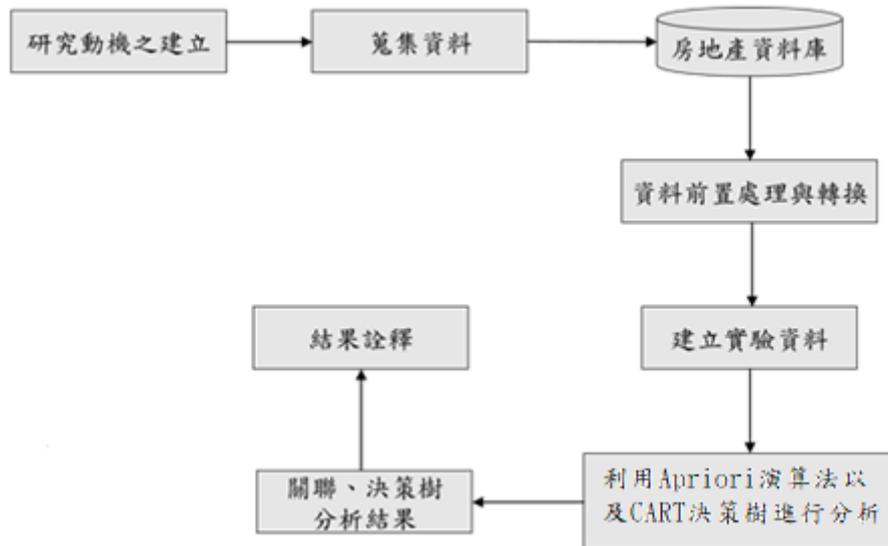


圖 3-1 本研究架構圖

資料來源:本研究根據 KDD 模型繪製

本研究根據 Han & Kamber (2000) 提出的資料選擇、前置處理、資料轉換、資料探勘解釋或評估資料探勘流程來設計本研究架構:

本研究架構圖 3-1 將列出以下探勘流程:研究動機之建立、蒐集資料、房屋資料庫、資料前置處理與轉換、建立實驗資料、運用資料探勘之關聯規則進行分析、分析結果和結果詮釋。經由研究之動機奠定,接著開始蒐集及尋找資料,並篩選出資料中有遺漏值或是不一致性的再建立出實驗資料,並透過 Apriori 關聯法進行分析,再將結果進行詮釋。

### 第二節 資料處理

本研究依據台中政府開放平台中下載之台中主計處提供的家庭收支調查報告將台中 29 個行政區區分為六個區域,將中區、東區、南區、西區、北區、西屯區、南屯區分為第一區域;北屯區、豐原區、大里區、太平區分為第二區域;沙鹿區、神岡區、潭子區、大雅區分為第三區域;梧棲區、烏日區、大肚區、龍井區、霧峰區分為第四區域;后里區、大甲區、外埔區、大安區、清水區分為第

五區域；最後將東勢區、新社區、石岡區、和平區分為第六區域，使用這六個區域來做為關聯分析的目標。

在我國房地產市場之發展－影響暨政府因應對策(2003)指出房地產一詞對於一般大眾及民間業者來說是房屋與土地的財產合稱，本研究因此篩選出內政部不動產交易實價查詢服務網台中 2014 到 2016 年的開放資料中屬性為交易標的中的房屋+土地的類別後約 10 萬筆，將此 10 萬筆資料標示上六個區域。另外，在台灣地區環境品質與公共設施對房價與房租影響之分析(1993)中有提到影響房價的因素會與房屋本身所具之特性以及與房屋座落之該棟樓房有關，本研究將內政部不動產交易實價查詢服務網台中 2014 到 2016 年六個區域約 10 萬筆的開放資料中挑出建物現況格局.房、建物現況格局.隔間、有無管理組織、建物型態以及有無車位和單價每平方公尺指標來做關聯分析，以此看出究竟房屋本身什麼樣的特質會影響到房價。

表 3-1 本研究資料處理表

資訊:內政部不動產交易實價查詢服務網台中 2014 到 2016 年房屋資料			
原本欄位	原本欄位內容	新欄位	轉為
交易筆棟數	內容中除了車位 0 以外	有無車位	有
	內容中有車位 0		無
單價每平方公尺	經過四分位數排列	單價每平方公尺指標	低/高

資料來源:本研究整理

### 第三節 Apriori 關聯分析

資料探勘的過程中，關聯法則是最常被應用的模式之一（Han&Fu,1999），經常會使用關聯法則來做分析，關聯法則又稱為「購物籃分析」，而在關聯法則中，Apriori 演算法是最著名而且最常被拿來使用的，最早是由 Agrwal & Srikant(1994) 提出，關聯法則的產生必定先定義最低支持度以及最低可靠度再由這兩者產生：

壹、支持度(Support)即某項目集在資料庫中出現的頻率（frequency），通常以 Support(X)表示之，以本研究為例公式如下：

$$\text{Sup(房價)} = \frac{\text{房價高(低)的項目在房屋資料庫中出現的總次數}}{\text{房價資料庫中全部資料筆數}}$$

貳、信賴度(Confidence)指兩個項目組之間的信賴程度，也就是在 X 出現的機率下，Y 也會出現的條件機率，通常以  $\text{Support}(X \cap Y) / \text{Support}(X)$  表示之，以本研究為例公式如下：

$$\text{Conf(房價} \rightarrow \text{公寓)} = \frac{\text{Sup(房價和公寓)}}{\text{Sup(房價)}}$$

算出支持度與信賴度之後，從房價資料庫中找出高頻率的項目集，另外它必須大於最低支持度，利用它產生關聯法則後信賴度須高於最低信賴度，此法則才算真正的確定成立，再進行相關分析，相關分析即為增益值。

參、增益值(lift)

$$\text{Lift} = \frac{\text{Conf(房價} \rightarrow \text{公寓)}}{\text{Sup(公寓)}}$$

若  $\text{Lift} > 1$ ，規則才具有實用性

## 第四節 CART 決策樹分析

決策樹分析有許多不同類的演算法，包括 ID3、CHAID、C5.0、CART 等，本研究使用其中的 CART 決策樹來進行分類分析，它又稱做分類迴歸樹，是由 Breiman (1984) 在 <Classification and regression trees> 一書中提出，黃俊傑 (2011) 提到 CART 決策樹是一種二元(binary)分割的演算法，二元分割顧名思義是將資料集合分割成左右兩個子集合，再從每一個子集合中在分割兩個子集合，反覆演算分割直到整個決策樹建構完成，而針對 CART 決策樹分割所需最小雜質改變輻模型採用 Gini 索引法，以下以本研究為例：

本房價資料中含有 106160 筆資料，此將價位以平均值區分為低和高，房價為高的有 53080 筆資料，房價為低的有 53080 筆資料，高房價出現的頻率為  $p_1$  為  $53080/106160$ ，也就是 0.5；反之低價位出現的頻率  $p_2$  也是 0.5，此時 Gini 索引值就為：

$$\text{Gini(房價)} = 1 - \sum_{j=1}^n P_j^2$$

$$\text{Gini(房價)} = 1 - (0.5^2 + 0.5^2) = 0.5$$

而 Gini 索引值越大，樣本的類別分布會越平均(曾憲雄,2005;高克志,2006)

## 第四章 分析與討論

### 第一節 資料集敘述統計和變數

#### 壹、資料來源

##### 一、房屋本身形態

本研究使用的資料為內政部不動產交易實價查詢服務網台中 2014 到 2016 年約 10 萬筆的開放資料依據台中主計處提供之台中家庭收支調查報告分為六區資料中的建物現況格局.房、建物現況格局.隔間、有無管理組織、建物型態以及有無車位等欄位和單價每平方公尺指標。

##### (一) 建物型態

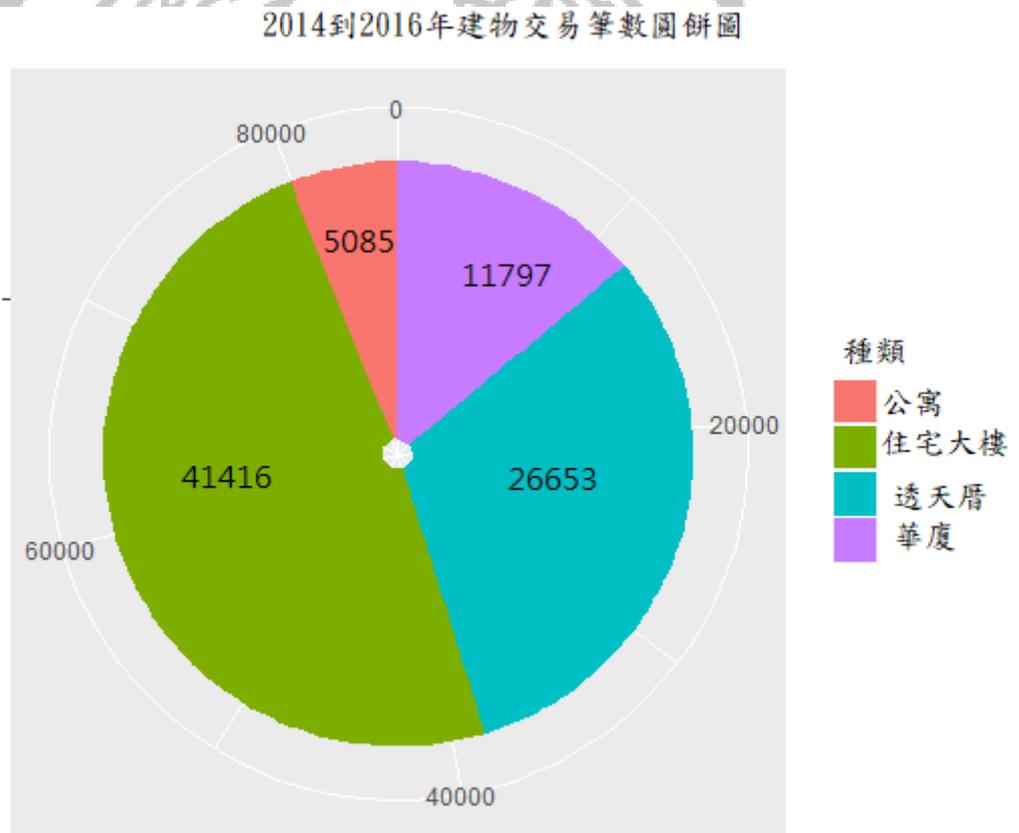


圖 4-1 2014 到 2016 建物型態圓餅圖

在 2014 到 2016 年之中所有資料包含了公寓 5085 筆、住宅大樓 41416 筆、透天厝 26653 筆以及華廈 11797 筆，在台中房屋資料中佔最多的建物型態為住宅大樓，最少的則為公寓。

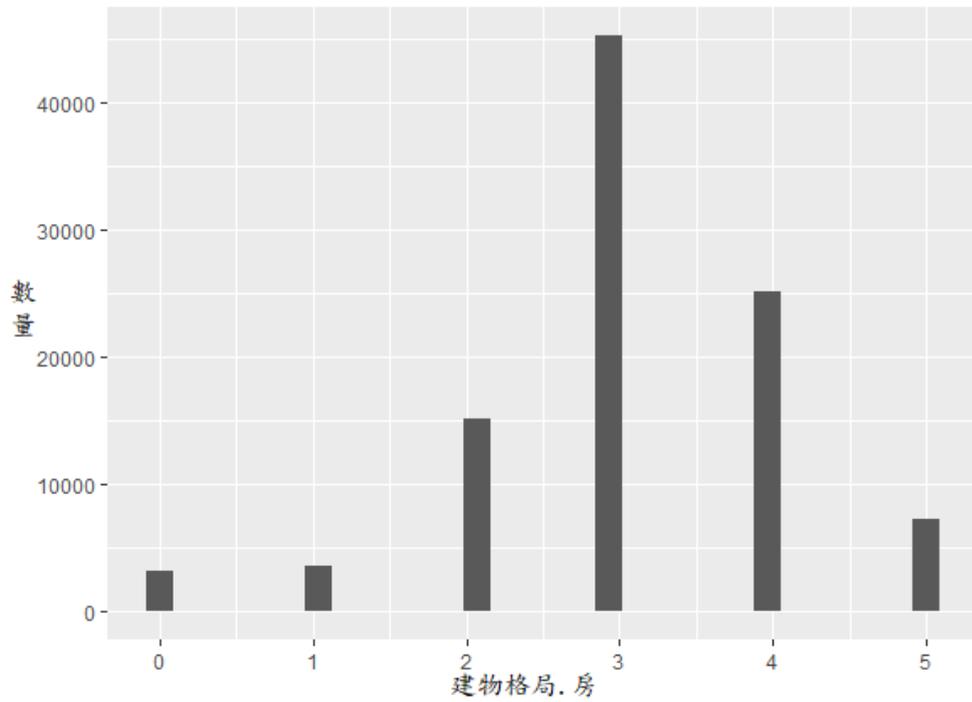


圖 4-2 房間數量直方圖

由圖 4-2 房間數量直方圖中可以看出在建物格局有幾個房間中三個房間占最多，其次為四個房間，最低則為零個房間及一個房間。

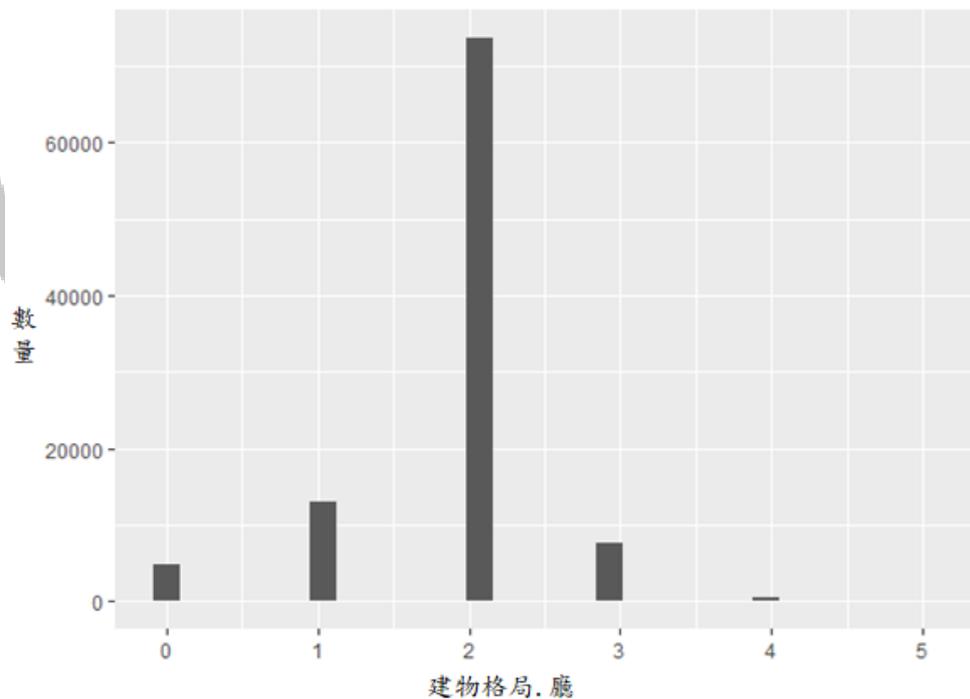


圖 4-3 客廳數量直方圖

建物格局中有幾廳中，二廳的占有最多，遠遠超出其他，而最低為四廳。

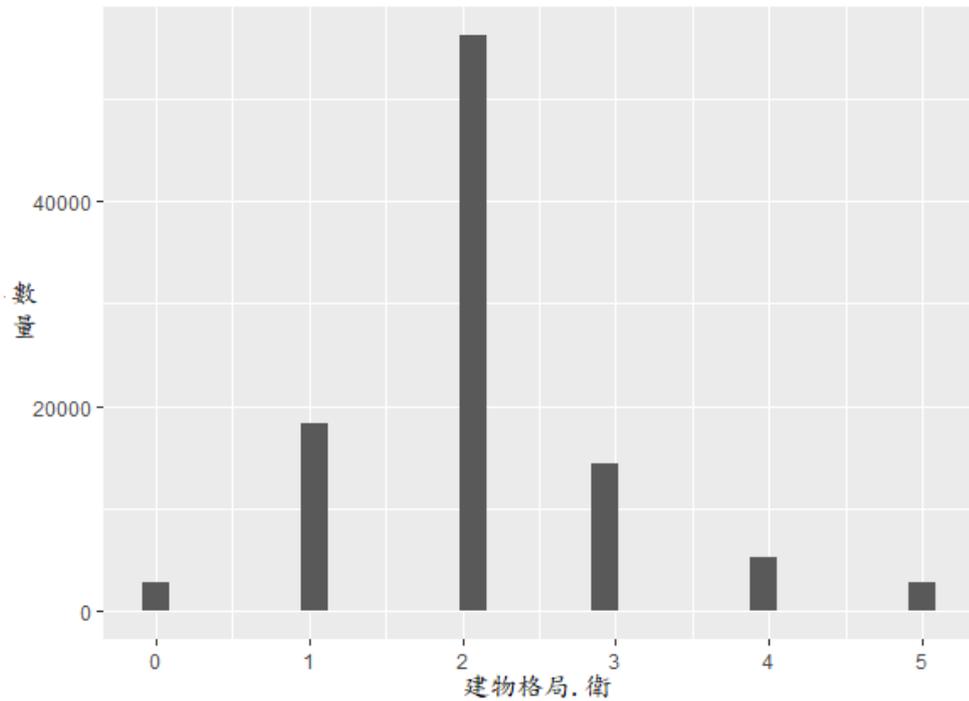


圖 4-4 衛浴數量直方圖

衛浴設備中，佔筆數最高為兩衛浴，最低為零及五衛浴。本研究以此資料中八個參數做關聯以及決策樹分析。

## 二、社會面

本研究在社會面使用台中政府開放平台中由台中民政局提供之 2014-2016 年結婚對數、離婚對數以及人口遷入遷出之社會移動人口兩項資料來和 2014-2016 年平均房價做 Apriori 演算法關聯分析。

## 三、經濟面

本研究使用台中政府開放平台中由台中主計處提供之 2014-2016 年家庭收支報告中挑選出所得收入總計(百萬)、消費支出(百萬)以及儲蓄(百萬)等三個參數來和 2014-2016 年平均房價資料來做 CART 決策樹分類分析。

## 貳、資料統計

本研究依據臺中市主計處家庭收支訪問調查報告將全台中區分為六個區域，將中區、東區、南區、西區、北區、西屯區、南屯區分為第一區域；北屯區、豐原區、大里區、太平區分為第二區域；沙鹿區、神岡區、潭子區、大雅區分為第三區域；梧棲區、烏日區、大肚區、龍井區、霧峰區分為第四區域；后里區、大甲區、外埔區、大安區、清水區分為第五區域；最後將東勢區、新社區、石岡區、和平區分為第六區域，並將這六個區域的房屋筆數以及平均每平方公尺價格統計圖表如下：

表 4-1 六區分配筆數

分區	地區	筆數				
		2014	2015	2016	2017	總共
一	中、東、南、西、北、西屯、南屯	16864	12561	10288	10109	49822
二	北屯、豐原、大里、太平	13760	8386	7681	7733	37560
三	沙鹿、神岡、潭子、大雅	3767	2787	2329	2161	11044
四	梧棲、烏日、大肚、龍井、霧峰	3315	2324	2287	1667	9593
五	后里、大甲、外埔、大安、清水	1302	1151	960	944	4357
六	東勢、新社、石岡、和平	247	228	196	191	862
	總共	39255	27437	23741	22805	113238

資料來源:本研究整理

由表 4-1 可以看出台中六個區域的房屋成交量均逐年下降

表 4-2 六區平均每平方公尺價格

分區	地區	平均每平方公尺價格				
		2014	2015	2016	2017	總共
一	中、東、南、西、 北、西屯、南屯	61456.24	62265.7	63479.27	65028.22	252229.4
二	北屯、豐原、大里、 太平	55467.36	58796.1	56619.49	58855.68	229738.6
三	沙鹿、神岡、潭子、 大雅	52491.98	51522.7	52129.09	54631.38	210775.2
四	梧棲、烏日、大肚、 龍井、霧峰	45233.73	47640.98	49096.62	48071.77	190043.1
五	后里、大甲、外埔、 大安、清水	40413.03	41234.34	43028.13	43332.28	168007.8
六	東勢、新社、石岡、 和平	30373.6	35372.57	32236.54	40671.79	138654.5
	總共	285435.9	296832.4	296589.1	310591.1	1189449

資料來源:本研究整理

由表 4-2 可以看出台中六個區域的平均每平方公尺價格大多逐年上升

## 第二節 關聯分析

本研究以二個面向利用 Apriori 演算法做關聯分析，分別為房屋本身型態、社會面，首先依據臺中市主計處家庭收支訪問調查報告將全台中區分為六個區域，並利用此六個區域中的 2014-2016 年的房價資料為對象分為建物格局和建物型態等兩種分開做 Apriori 關聯分析，使用 SPSS 套裝軟體採用 Apriori 演算法做為探勘方法，本研究因資料筆數較多，故支持度並不會太高，將最低信賴度設為 0.5，以及增益值 $>1$ ，規則才具有實用性，得到結果如下：

### 壹、以房屋本身格局

名詞解釋：

1. 建物現況格局.房:代表此建築物的房屋格局有幾個房間

2. 建物現況格局.廳:代表此建築物的房屋格局有幾個客廳
3. 建物現況格局.衛:代表此建築物的房屋格局有幾個衛浴設備
4. 有無車位:代表此建築物的車位情況
5. 有無管理組織:代表此建築物的管理組織強況
6. 建物型態:說明此建築物的建物型態為公寓、住宅大樓、透天厝、華廈中的哪一種
7. 單位每平方公尺指標:本研究將六個區域依據四分位數排列分別將六個區域的平均每平方公尺房價分為低和高。

利用此房屋資料投入六項參數分別為建物現況格局.房、建物現況格局.廳、建物現況格局.衛、有無車位、有無管理組織、建物型態來和單位每平方公尺指標做關聯分析:

一、 區域一:中區、東區、南區、西區、北區、西屯區、南屯區

表 4-3 區域一關聯分析表

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=高	建物型態標示=住宅大樓 建物現況格局.廳=2	51.99	56.883	1.138
單位每平方公尺指標=高	建物型態標示=住宅大樓	63.371	54.699	1.094
單位每平方公尺指標=高	建物現況格局.衛=2	57.631	53.478	1.07
單位每平方公尺指標=高	建物現況格局.衛=2 建物現況格局.廳=2	53.665	52.756	1.055
單位每平方公尺指標=高	建物現況格局.廳=2	74.411	51.007	1.02

根據表 4-3 區域一關聯分析表結果顯示，在本研究台中分區區域一中如果建物型態為住宅大樓且客廳有兩間的話，支持度 51.99%的情況下，信心度有 56.883%，在此情況下，單價每平方公尺的價格會較高，另外如果房間格局中衛浴有兩間且客廳也有兩間的話，在支持度 53.665%且信心度

為 52.756% 的情況下會使單價每平方公尺指標為低，依照結果本研究推斷應是由於區域一大多為市中心發展重點，地狹人稠，且現今單身獨住的人越來越多，因此衛浴設備兩間及客廳兩間可能房屋坪數較大，且設施太多較不實用，造成相對來說衛浴兩間的房價偏低。

## 二、 區域二：北屯區、豐原區、大里區、太平區

表 4-4 區域二關聯分析

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=低	建物現況格局.衛=2	59.067	55.718	1.114
單位每平方公尺指標=低	建物現況格局.衛=2 建物現況格局.廳=2	52.838	54.979	1.1
單位每平方公尺指標=低	建物現況格局.廳=2	74.893	51.342	1.027
單位每平方公尺指標=高	建物型態標示=住宅大樓	50.673	50.307	1.006

根據表 4-4 顯示在本研究台中分區區域二中當建物格局有兩個衛浴的房屋在支持度為 59.067% 且信心度為 55.718% 的情況下，單位每平方公尺指標為低，在支持度為 74.893% 且信心度為 51.342% 的情況下，房屋格局為兩個客廳也是造成房價降低的原因，而在房屋格局中有兩個衛浴且有兩個客廳的房屋格局在支持度 52.838% 且信心度為 54.979% 的情況下，單位每平方公尺指標也是為低，猜測在台中分區區域二中居住的民眾較多是離家工作自己獨居，而非家庭式居住，太大的空間反而造成不必要的困擾，另外在支持度 50.673% 且信心度為 50.307% 的情況下，建物型態為住宅大樓則單位每平方公尺指標為高。

### 三、 區域三：沙鹿區、神岡區、潭子區、大雅區

表 4-5 區域三關聯分析表

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=高	建物型態標示=透天厝	61.447	62.787	1.256
單位每平方公尺指標=高	建物型態標示=透天厝 無車位	57.392	62.181	1.244
單位每平方公尺指標=高	無管理組織	56.332	56.495	1.13
單位每平方公尺指標=高	無管理組織 無車位	51.373	54.987	1.1
單位每平方公尺指標=低	建物現況格局.廳=2	60.046	54.259	1.085
單位每平方公尺指標=高	無車位	72.295	51.339	1.027

根據表 4-5 顯示在本研究台中分區區域三中，在支持度 61.447%且信心度為 62.787%的情況下，建物型態為透天厝會使單位每平方公尺指標為高，另外可能因為透天厝本身的建物型態就沒有車位也沒有管理組織，為獨棟建築，因此在支持度為 72.295%且信心度為 51.339%的情況下無車位及在支持度為 56.332%且信心度為 56.495%的情況下無管理組織也會和房價指標高有關聯，另外也可能因為在區域三本身透天厝型態的建物型態就較多也可能造成此結果。

#### 四、 區域四：梧棲區、烏日區、大肚區、龍井區、霧峰區

表 4-6 區域四關聯分析表

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=低	無車位	64.103	56.75	1.135
單位每平方公尺指標=低	無管理組織	55.821	56.141	1.123
單位每平方公尺指標=低	無管理組織 無車位	51.021	56.08	1.121
單位每平方公尺指標=高	建物現況格局.廳=2	68.149	52.221	1.045
單位每平方公尺指標=高	建物型態標示=透天厝	51.824	50.072	1.002

根據表 4-6 顯示在本研究台中分區區域四中在支持度為 51.824%且信心度為 50.072%的情況下當建物型態標示為透天厝時，單價每平方公尺指標為高，而分別在支持度 64.103%且信心度為 56.75%的情況下無車位、支持度 55.821%且信心度為 56.141%的情況下無管理組織以及在支持度 68.149%且信心度 52.221%的情況下房屋格局為兩個客廳的房屋則單價每平方公尺指標為低，透天厝在區域四可能是屬於較高價的建築型態，但猜測並沒有和區域三一樣較多住在透天厝這個建物型態中，因此在此情況下如果以其他建物型態像是華廈、住宅大樓或著公寓這三個建物型態的話，如果沒有車位及沒有管理組織的會較低價。

五、 區域五：后里區、大甲區、外埔區、大安區、清水區

表 4-7 區域五關聯分析表

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=高	建物型態標示=透天厝	77.635	54.06	1.081
單位每平方公尺指標=低	建物現況格局.廳=2	62.195	53.903	1.078
單位每平方公尺指標=高	建物型態標示=透天厝 無車位	74.942	53.546	1.071
單位每平方公尺指標=低	建物現況格局.廳=2 無車位	71.813	52.131	1.043
單位每平方公尺指標=高	無管理組織	79.245	52.033	1.041
單位每平方公尺指標=高	建物型態標示=透天厝 無管理組織 無車位	69.42	51.409	1.028
單位每平方公尺指標=高	無車位	85.32	51.133	1.023
單位每平方公尺指標=高	無管理組織 無車位	74.137	50.403	1.008

根據表 4-7 顯示在本研究台中分區區域五第五區來看在支持度 77.635% 且信心度為 54.06% 的情況下，透天厝的建物型態較高價，支持度高達 77.635% 可以看出在 2014 到 2016 年中房屋交易在第五區域中大多為透天厝，在支持度為 62.195% 且信心度為 53.903% 的情況下建物現況格局為兩廳的房屋，單位每平方公尺指標為高，而此區也是因為透天厝的建物型態，而導致在支持度為 79.245 且信心度為 51.409 的情況下無管理組織和支持度 85.32% 且信心度為 51.133% 的情況下無車位會和單位每平方公尺指標為高有關聯，並且在此區域中也有一項關聯為透天厝、無管理組織和無車位在支持度 69.42% 且信心度為 51.409% 的情況下單位每平方公尺指標為高。

## 六、 區域六：東勢區、新社區、石岡區、和平區

表 4-8 區域六關聯分析表

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=高	建物型態標示=透天厝 無車位	81.628	55.983	1.122
單位每平方公尺指標=高	建物型態標示=透天厝	82.907	55.961	1.122
單位每平方公尺指標=高	建物現況格局.廳=2 無車位	53.837	55.724	1.117
單位每平方公尺指標=高	建物現況格局.廳=2	55.698	55.324	1.109
單位每平方公尺指標=高	建物型態標示=透天厝 無管理組織 無車位	78.721	54.505	1.093
單位每平方公尺指標=高	建物型態標示=透天厝 無管理組織	9.767	54.373	1.09

根據表 4-7 顯示在本研究台中分區區域六中，在支持度為 78.721%且信心度為 54.505%下建物型態為透天厝且無車位及無管理組織會較為高價，猜測應該也是因為在此區域透天厝的建物型態較多且透天厝本身的建物型態就是無車位且無管理組織的關係，另外在此區域中在支持度為 55.698%且信心度為 55.324%的情況下建物現況格局有二個客廳的單位每平方公尺指標會為高。

## 貳、以社會面

在以社會面看房價方面，挑選了台中民政局提供之 2014-2016 年結婚對數、離婚對數以及人口遷入遷出三項資料來和房價做關聯分析：

### 一、結婚、離婚對數

表 4-9 結婚、離婚對數關聯分析表

前項	後項	支持度	信心度
單位每平方公尺指標=高	平均結婚對數=少	77.778	42.857
單位每平方公尺指標=高	平均離婚對數=少	72.222	38.462
單位每平方公尺指標=高	平均離婚對數=少 平均結婚對數=少	72.222	38.462

依據結婚、離婚資料和房價來做關聯分析之後發現，當結婚對數少的時候，房價會比較高，本研究推斷這也有可能是現在結婚率下降以及房價持續上漲的關係，結、離婚對數對房價的影響的結果和本來預期的結果不同，預期是結婚對數高，則房價高，因猜測結婚成家需購置房產可能須購買較大的住宅，但研究結果顯示結婚對數低時房價會較高，推斷應是因為現在結婚率下降，獨身主義者逐年增加，而購置房產可以投資存老年安養資金的緣故。另外推斷離婚對數少會跟房價有關聯的原因可能是因為離婚對數少就不會有急著脫手房產，而房價自然不會變低的關係。

### 二、社會移動人口

表 4-10 社會移動人口關聯分析表

後項	前項	支持度	信心度	增益值
單位每平方公尺指標=高	社會移動人口=負	46.429	69.231	1.491
單位每平方公尺指標=高	社會移動人口=負	46.429	30.769	0.574

依據台中民政局 2014-2016 年人口遷入遷出統計資料中，遷入人口數以及遷出人口數，本研究以遷入人口減掉遷出人口後，如值為正則社會移動人口為正，如值為負則社會移動人口為負，由表中可看出在 2014-2016 間，當社會移動人口為負值時平均每平方公尺價格會為高，但第二項也說明當社會移動人口為負值時

平均每平方公尺價格會為低，但由增益值來看，第二項規則的增益值並沒有 $>1$ ，因此此規則的可信度並不高，社會移動人口為負值卻房價高的原因有可能是因為房價高，但並不是所有人都買得起房價較高的房屋，因此遷入的人比遷出的人少。

### 第三節 決策樹分析

#### 壹、 房屋格局、建物型態

最後本研究將房屋本身格局、建物型態、有無車位等等所有和房屋本身因素有關資料投入決策樹：

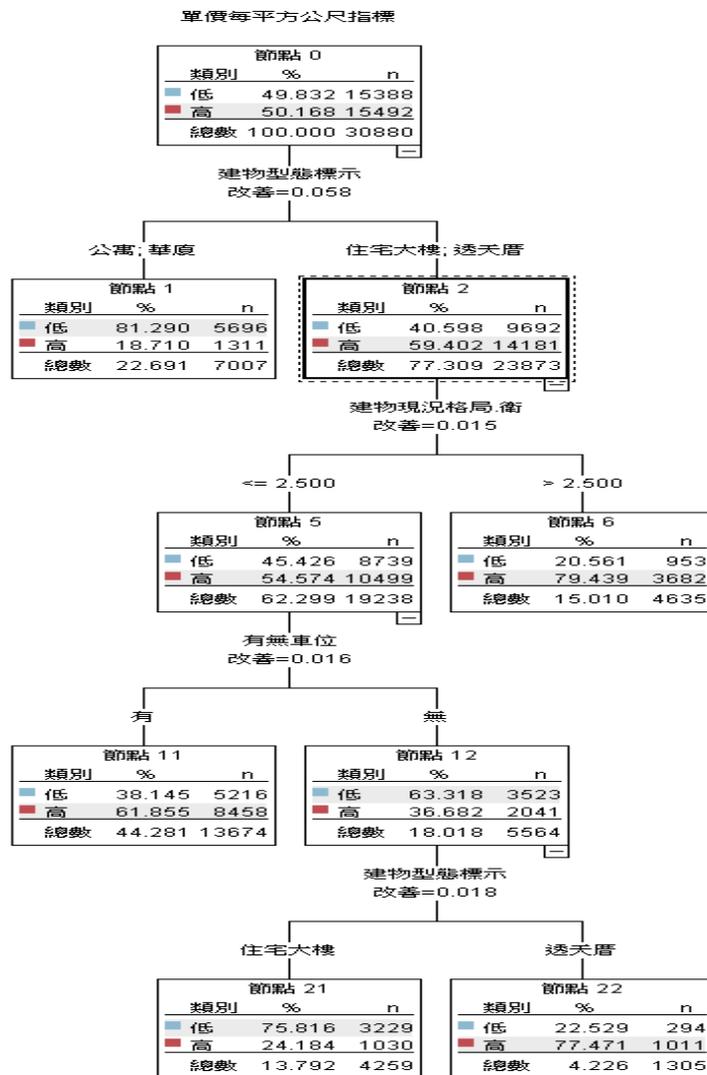


圖 4-5 房屋格局型態決策樹分析圖

決策樹分析出來的結果單價每平方公尺指標為低的規則有 2 項，而指標為高的規則有 3 項：

一、單價每平方公尺指標=低

(一) 如果建物型態為公寓及華廈

(二) 如果衛浴 $\leq 2.5$ 、為住宅大樓且無車位

二、單價每平方公尺指標=高

(一) 如果衛浴 $\leq 2.5$ 、建物型態為住宅大樓及透天厝且有車位

(二) 如果衛浴 $\leq 2.5$ 、建物型態為透天厝且無車位

(三) 如果衛浴 $\leq 2.5$ 、建物型態為住宅大樓及透天厝

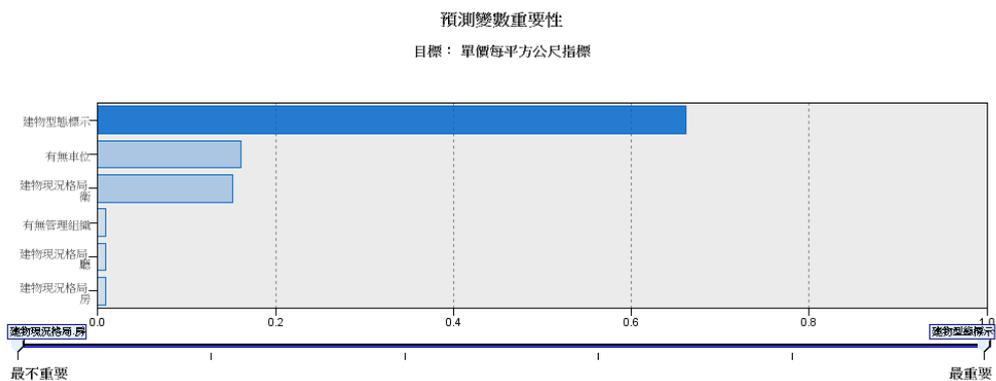


圖 4-6 預測變數重要性圖

且從圖 4-6 可以看出在影響單價每平方公尺指標中，最重要的參數為建物型態標示，可以知道在建物格局中影響房價最主要的為建物型態，而第二重要的為有無車位，接下來則是建物現況格局有幾個衛浴設施。

## 貳、家庭收支調查報告

本研究將家庭收支調查報告中四個參數分別為所得收入總計(百萬)、消費支出(百萬)、儲蓄(百萬)與實價網提供之房屋資料中的平均房價投入做決策樹分析出現結果如下:

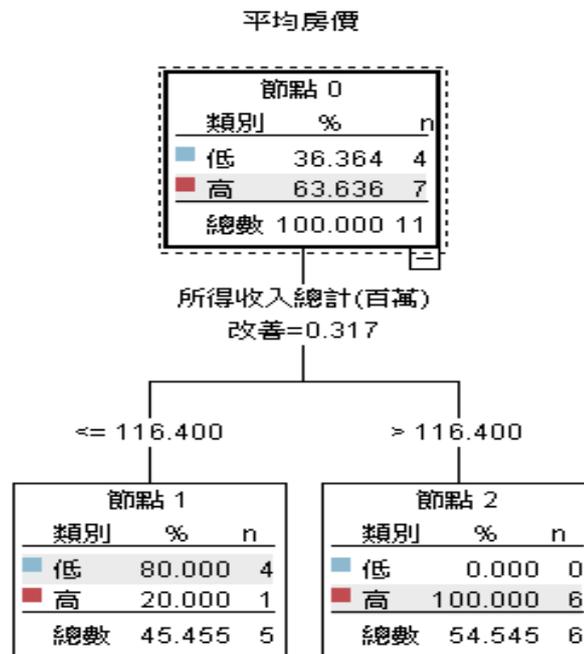


圖 4-7 所得收入決策樹分析圖

根據決策樹分析結果來看，產生單價每平方公尺指標為高和低的規則各一項：

- 一、單價每平方公尺指標=低
  - (一) 如果所得收入總計(百萬) $< 116.4$
- 二、單價每平方公尺指標=高
  - (一) 如果所得收入總計(百萬) $> 116.4$

可以看出當國民收入較高時，房價也會較高；而當國民收入較低時，房價也較低。

## 第五章 結論與建議

首先由房屋本身建物型態來看，由關聯分析規則和決策樹結果顯示，建物型態為住宅大樓及透天厝的價格相對來說會比較高，而有無管理組織及有無車位可能會受到透天厝的建物型態影響，較難看出是否影響到了房價高低，另外某些區域中客廳個數、房間個數以及衛浴個數也會影響房價，但在有的地區中不造成影響，而有無隔間則均無顯示在規則內，對於房價應是沒有太大的影響，日後研究可以另外提出其他備有車位以及管理組織的公寓及華廈等其他型態的建物單獨出來探討各個建物型態分別是否受有無車位以及有無管理的影響。

從社會和經濟方面探討影響房價因素，本研究限制為台中改制六都以後，台中縣市合併，很多資料並不容易比對在一起，且資料中並無明顯的行政區分區，資料大多為統計資料，筆數太少，並無完整資料庫。社會面方面以資料較齊全的結、離婚對數以及社會移動人口兩項資料做關聯分析，結、離婚對數對房價的影響的結果和本來預期的結果不同，預期是結婚對數高，則房價高，但結果反之，而依照社會人口移動對房價的影響的結果則也和預期相反。最後以經濟面做決策樹分析結果就和預期相同，收入越高則房價會越高。

未來希望能找到更齊全的社會資料例如：台中市地區之福利機構位置、車載運輸量等等以及經濟指標資料例如：台中市地區的國民生產毛額(GDP)、消費者物價指數、歲出等等，目前這些社會方面資料都只有全台中統計一筆資料並沒辦法探討以行政區探討台中地區而經濟指標則大多只有全台灣統計每年一筆，希望未來能有更齊全的台中資料庫，可以從更多外在因素去探討房價這個問題，並使用範圍更廣大的區域去做分析更佳，期望能讓一般社會大眾在買房子時，能夠有更多的參考指標並提供更多樣化的選擇和思考空間。

## 參考文獻

### 中文文獻

1. 林祖嘉、林素菁 (1993 年 1 月)，台灣地區環境品質與公共設施對房價與房租影響之分析，住宅學報，第一期，第 22 頁。
2. 林崇詠 (2015)，《影響房地產價格因素之研究以臺中市西屯區為例》，朝陽科技大學財務金融系，未出版。
3. 陳幸宜 (2003)，《都市房價變動影響因素之系統動態模擬》，成功大學都市計劃學系碩士論文，未出版。
4. 新台灣國策智庫 (2017 年 6 月)，「六都市長就任兩週年施政滿意度暨主要政治人物好感度民調」解析，新台灣國策智庫通訊，第五十五期，第 7 頁。
5. 楊鴻謙 (2001)，《不動產投資與市場分析》，台北：文笙。
6. 廖述賢、溫志皓 (2012)，《資料探勘理論與應用》，台北：博碩文化，P2-2.2-3
7. 銀行季刊，第二十五期，第 31-64 頁。
8. 蔡育政 (1999)，《影響房地產價格因素之研究以臺中市北屯區、西屯區、南屯區、中區、東區為例》，朝陽科技大學財務金融系，未出版。
9. 蔡曜如 (2003 年 12 月)，我國房地產市場之發展—影響暨政府因應對策，中央銀行季刊，第四期，第 32 頁。
10. 顏聰玲 (2004)，《不動產市場分析與預測》，台北：新文京開發。

### 英文文獻

1. Agrawal, R. and Srikant, R. (1994) .Fast Algorithms for Mining Association Rules in Large Database, Proceedings of the 20th International Conference on Very Large Data Bases , pp. 487-499.
2. Berry, M. J. A. and Linoff, G. S. (2004) .Data Mining Techniques: For Marketing, Sales, and Customer Relationship Management, 2nd Ed., New York: John Wiley.
3. F. H. Grupe, and Owrang, M. M. (1995) .Database Mining Discovering New

- Knowledge and Cooperative Advantage, *Information Systems Management*, 12(4), pp. 26-31.
4. Frawley, W., and Piatesky-Shapiro, C., and Matheus, C. (1992). Knowledge discovery in database: an overview, *AI Magazine*, 16(3), pp. 57-70
  5. Han & Kamber (2000). *Data mining: concepts and techniques*, USA: Morgan Kaufmann Pub
  6. Han, J. and Y. Fu (1999). Mining Multiple-Level Association Rules in Large Databases, *IEEE Transactions on Knowledge and Data Engineering*, 11(5), pp. 789-805.
  7. Lancaster, K. (1996). A new approach to consumer theory, *Journal of Political Economy*, 74(2), 132-157.

#### 網路

1. CNN, 2016年4月, <https://edition.cnn.com/travel/gallery/taichung-most-liveable-city/index.html>
2. 內政部不動產交易實價查詢服務網, 2014年1月-2017年12月, <http://lvr.land.moi.gov.tw/homePage.action>
3. 葉思含, 〈房市春燕何時來? 台中5巨頭看法喜憂參半〉, 《自由時報》, 民 106.1.23, <http://estate.ltn.com.tw/article/2383>。