

國立政治大學資訊管理研究所

100 學年度 第二學期

碩士學位論文

建立資料倉儲與資料採礦實現洞察力行

銷之研究

--以個案公司為例說明

指導教授：季延平 博士

研究生：劉映蘭

中華民國 101 年 7 月

致謝

在打下論文最後一個字之後，有很深刻的學生生涯告一段落，必須告別一個身分的感觸。大學四年加上碩士兩年共六年的時間，季老師給了我很多東西，除了學術上的指導，更多的是讓我累積經驗的機會，從做中學，在錯中成長，我好像漸漸知道資訊管理原來是怎麼一回事，以及從這越挖越深、越挖越廣的領域中獲得哪些樂趣，六年來的這些收穫很珍貴也很珍惜，感謝季老師。

這篇論文能完成要感謝很多人的配合與幫忙；接觸資料採礦的時間其實僅僅一年，它包含了太多領域的專業知識，尤其統計與數學，非本科系的我只能靠自修來自我充實，特別感謝台北大學資訊管理學系的陳宗天老師，給予我對於資料採礦的悉心指導，耐心回答所有不懂的邏輯和演算式，感謝陳老師。最後謝謝環域科技公司與台電公司在分析資料及專業知識上的幫助和訪談的配合。

謹以此成果獻給一路上支持我的家人和摯友們，尤其這段期間對翔玲有諸多叨擾，感謝你們的包容，我愛你們。

摘要

提供以顧客為導向的服務、提升顧客價值成為各企業經營策略的出發點，因此了解顧客特徵、需求、行為成為首要任務。

隨著經濟發展，台灣整體用電需求不斷攀升，同時電力產業亦須面對全球燃料價格上漲的現況，政府邁向低碳經濟時代，推出一系列相關政策，於我國電力結構中占有重要地位的台灣電力公司也開始思考如何在需求及成本皆提升的環境下，繼續維持穩定供電，提供以用戶為導向的服務。

本文即以台電公司為個案，以資料倉儲與資料採礦技術進行【用戶區隔】了解台電用戶特徵、【選擇目標用戶】找出優先服務的用戶對象及【產品定位】推薦適合目標用戶用電特徵之電價及優惠方案，實現行銷活動 STP 分析過程之研究。

關鍵詞：資料倉儲、資料採礦、洞察力行銷、市場區隔、STP

Abstract

Providing customer-oriented services and elevating customer value have become the starting point of business strategy. Therefore, understanding customer characteristics, requirements, and behavior has become the primary task.

As the overall demand for electricity rises constantly in Taiwan along with economic development, electric power industries have to face the surging fuel prices. To put low carbon economy into practice, the government has launched various policies. Taiwan Power Company, which plays an important role in the electric power structure in Taiwan, started thinking about how to provide stable power and customer-oriented services under the circumstances both demand and cost increase.

Based on the case of Taiwan Power Company, this research investigated customer segmentation - to understand customer characteristic, customer targeting - to find out the prior customer, and product positioning - to recommend target customer appropriate tariff and preferential tariff treatment through data warehouse and data mining technology. The purpose of the research was to achieve STP analysis in marketing process.

Key words: data warehouse, data mining, insight driven marketing, STP

目錄

第一章、緒論	8
第一節、研究背景	8
第二節、研究動機	10
第三節、研究目的	10
第四節、預期效益	11
第二章、文獻探討	12
第一節、洞察力行銷：市場區隔、目標市場與產品定位	12
第二節、資料倉儲	25
第三節、資料採礦與演算法介紹	31
第三章、研究方法與設計	44
第一節、研究方法	44
第二節、研究設計	45
第三節、研究架構	47
第四章、個案分析	49
第一節、我國電力環境現況	49
第二節、個案分析	51
第三節、個案公司需求訪談	57
第四節、STP 分析資料模型設計	61
第五節、STP 實做與分析	69
第五章、結論與建議	81
第一節、結論	81
第二節、未來發展	83
第六章、參考文獻	84

圖目錄

圖 1、洞察力：企業獲得差異化競爭優勢的關鍵之路	13
圖 2、洞察力行銷帶來的改變	14
圖 3、AMA 2004 年行銷概念圖	16
圖 4、不同市場區隔變數之實行難易程度	18
圖 5、合適/不合適的市場區隔變數之區隔結果	20
圖 6、不同行銷策略下之目標市場選擇方式	21
圖 7、三種目標市場選擇策略	22
圖 8、定位策略步驟	23
圖 9、啤酒市場的產品知覺圖	24
圖 10、資料倉儲特性-整合性	27
圖 11、資料倉儲特性-時間的變化性	28
圖 12、資料倉儲特性-不可更新性	28
圖 13、星狀模型	30
圖 14、資料立方體	30
圖 15、資料採礦步驟	32
圖 16、每個群集可視為一機率分配	36
圖 17、單一鏈結分群結果示意圖	38
圖 18、研究流程圖	46
圖 19、研究架構圖	47
圖 20、執行步驟	48
圖 21、台灣發電結構 (2012/1~4 月)	50
圖 22、台灣電力公司組織架構圖	52
圖 23、資料倉儲-高壓用戶屬性主檔	62
圖 24、資料倉儲-AMI 電表需量資料表	63
圖 25、預處理-AMI 缺漏值補正	64
圖 26、預處理-AMI 資料存儲結構轉換	66
圖 27、預處理-AMI 資料比例化前後數據	67
圖 28、預處理-AMI 資料比例化前用電趨勢線	67
圖 29、預處理-AMI 資料比例化後用電趨勢線	68
圖 30、K-MEANS 分群結果	73
圖 31、簡單貝氏機率分類建模結果	77
圖 32、STP 分析執行模式	81
圖 33、系統架構設計	82

表目錄

表 1、AMI 推動時程表	9
表 2、高低壓用戶佔比	9
表 3、AMA 針對行銷提出之定義	15
表 4、單一鏈結名目資料之計算範例資料	39
表 5、匹配矩陣	40
表 6、資料採礦於各領域之應用	43
表 7、台電公司發電結構	53
表 8、台電公司燃料成本	54
表 9、北市區處訪談結果	58
表 10、北北區處訪談結果	59
表 11、總部業務處費率組訪談結果	60
表 12、分析需求歸納	61
表 13、C4.5 用戶分類用戶篩選設定	69
表 14、C4.5 用戶分類分析結果 (部分)	70
表 15、本月份用電度數、尖離比分析	71
表 16、K-MEANS 用戶分群用戶篩選設定	71
表 17、K-MEANS 用戶分群分析結果	72
表 18、用戶 A 計畫性減少優惠方案試算結果	74
表 19、各區營業處尖離比 (由低至高)	75
表 20、簡單貝氏機率分類用戶篩選設定	76
表 21、簡單貝氏機率分類建模變數設定	77
表 22、建模結果-尖離比與方案參與關係	77
表 23、簡單貝氏機率分類預測用戶篩選設定	78
表 24、簡單貝氏機率分類預測結果	79
表 25、用戶 A 計畫性減少優惠方案一試算結果	79
表 26、用戶 B 計畫性減少優惠方案一試算結果	80

第一章、緒論

第一節、研究背景

近年來，全球燃料價格上漲、用電量亦隨著經濟發展而攀升，台電公司面臨如何提供民眾穩定用電的問題，因此不斷努力提升供電及服務品質，並致力以用戶服務導向經營目標。

另一方面，政府為邁向「低碳經濟」，於民國 99 年推動智慧電網 (Smart Grid) 節能減碳策略，由行政院核定「智慧型電表基礎建設推動方案」及各相關政府部門制定相關政策及計畫，故台電公司為因應國家策略之發展，陸續跟進智慧型電表基礎建設 (AMI, Advanced Metering Infrastructure) 裝設。

智慧型電表基礎建設 (AMI) 主要由智慧型電表 (Smart Meter)、通訊系統、電表資訊管理系統 (MDMS, Meter Database Management System) 所組成，可取代傳統人工抄表，每 15 分鐘自動傳回該電表所使用的用電度數。台灣 AMI 之設置對象可分為高壓用戶與一般家庭之低壓用戶。依據經濟部能源局規劃，優先推動占全國用電量 58% 之高壓用戶 2,3000 戶預計於民國 101 年全部裝設完畢；低壓用戶 (戶數高達 1200 萬戶) 由於用電量僅占 42% 且投資金額較大，故將以 50% 約 600 萬戶為目標。推動時程請參閱下表 1。

表 1、AMI 推動時程表

實施年度	高壓戶數	低壓戶數
民國 99 年	2 萬 3000 戶	300~500 戶
民國 100 年		1 萬戶
民國 101 年		
民國 102 年	--	100 萬戶
民國 103 年	--	
民國 104 年	--	
民國 105 年~	--	500 萬

資料來源：經濟部能源局民國 99 年 6 月 28 日發布，標題：我國智慧型電表基礎建設正式啟動

表 2、高低壓用戶佔比

	高壓用戶	低壓用戶
用戶數佔比	0.19%	99.81%
用電量佔比	58%	42%

資料來源：101 年台電公司

在用電戶方面，台電公司資料顯示，台灣的高壓用戶（約 23,000 戶）僅佔全體用戶數的 0.19%，但用電量卻占台電售電量的 58%，如上表 2 所示）；因此，若能善加利用 AMI 智慧型電表所收集的用電資料，台電公司資料倉儲將具備大量分析資料，做為高壓用戶動態負載分析中屬性、需量、計量等不同類型資料分析之基礎，便能最佳化資源配置、協調電能供需關係、引導電能消費行為、鼓勵用戶在負載低時段合理用電，尖峰負載時少用電。有效管理高壓用戶用電行為、

穩定供電及提高供電效率，以提供更佳的用戶服務。

第二節、研究動機

台電公司過去所從事的電力負載管理相關研究因為資料有限，分析資料顆粒度以月為單位，所涉及的成果都在某一特定範圍與時間，屬於靜態式研究，如今有了 AMI 所收集的資料來源，可充分掌握用戶特性、用電行為與趨勢等寶貴的資訊，未來負載管理相關分析可以更廣更動態，以協助策略之研擬規劃。為實現用戶導向服務，提供穩定供電品質，台電應確實掌握用戶需求（需求面）並最佳化資源配置（供應面）。

第三節、研究目的

本研究主要針對用戶需求面進行分析，試著回答以下問題：

- 台電用戶具備什麼特徵與行為？
- 對台電公司而言，誰是目標用戶？
- 應該提供目標用戶什麼服務？

運用資料探勘之分群（Clustering）、分類（Classification）方法，提出洞察力行銷-STP 分析模型，探討如何運用台電公司現有用戶服務資料倉儲（隸屬台電公司綜合研究所），針對高壓用戶實現負載管理之分析。

第四節、預期效益

由於過去受限於一至兩個月一次人工抄表的頻率，無法獲得即時且精準的分析資料，再加上分析工具的缺乏，故相關分析只能產生較為靜態的研究成果；今政策推廣 AMI 智慧電表裝設，透過自動傳回每 15 分鐘一筆的用電資訊，大幅度增加資料來源的完整度與精確度，本研究即以 AMI 智慧電表資料為基礎進行用戶負載有關研究，提升研究品質與範圍。

預期效益可分為公司與用戶兩方面，如下：

1、公司面效益

- 協助台電公司依據事實資料進行策略研擬：提出用戶區隔、目標用戶選擇及產品／服務定位之分析模型，以事實資料分析用戶行為資訊、掌握用戶特性、用電習慣與趨勢，提供負載策略研擬之基礎。
- 抑低尖峰用電：透過用戶區隔分析了解用戶特徵與行為，以電價優惠誘導用戶改變用電行為。

2、用戶面效益

- 協助用戶合理化用電與最佳化資源配置：誘導用戶移轉尖峰負載，朝向合理化用電，降低用電成本。

第二章、文獻探討

第一節、洞察力行銷：市場區隔、目標市場與產品定位

1、洞察力行銷（SAP，2011）

“洞察力：成長中的企業獲得差異化競爭優勢的關鍵之路。”在任何行業中，將最具競爭力的企業與最不具競爭力的企業相比，前者的員工受惠於洞察力的影響的比例是後者的 2 倍，相反的，後者將資料共用視為喪失控制力的比例是前者的 2.5 倍；這說明了洞察力對競爭優勢具有正向推動關係。

洞察力指的是對環境或任何情況做出預見的能力；因此，洞察力結合了經驗與分析能力，幫助實現更好的決策。洞察力能夠幫助企業尋找機會，發現問題，從而基於這些發現採取修正的行動。其關係如下圖 1 所示。

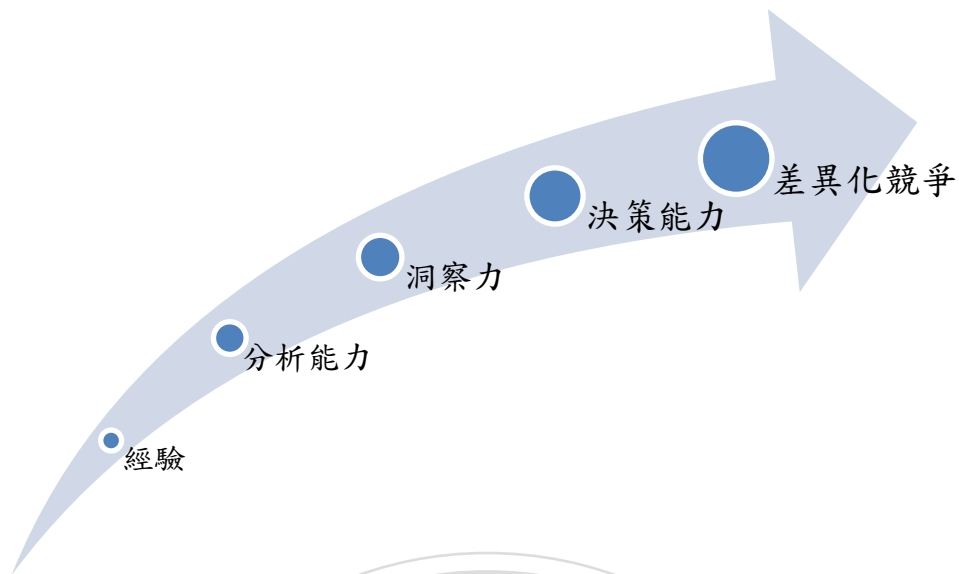


圖 1、洞察力：企業獲得差異化競爭優勢的關鍵之路

資料來源：本研究整理

洞察力行銷（Insight Driven Marketing）的目的地是企業透過對所掌握的顧客資料對顧客的價值或行為特徵進行深入分析，以獲得對顧客的洞察力，並利用洞察的結果制定差異化的產品，建置“以顧客為中心”的行銷模式。透過對現有顧客的區隔，把企業的顧客作為資源進行管理，建立基於各類顧客群的管理原則和具體方法，從而強化企業的行銷能力與市場競爭力。對企業帶來的影響是上自策略下至流程的整體性的改變。如下圖 2。

以洞察力行銷創造效益

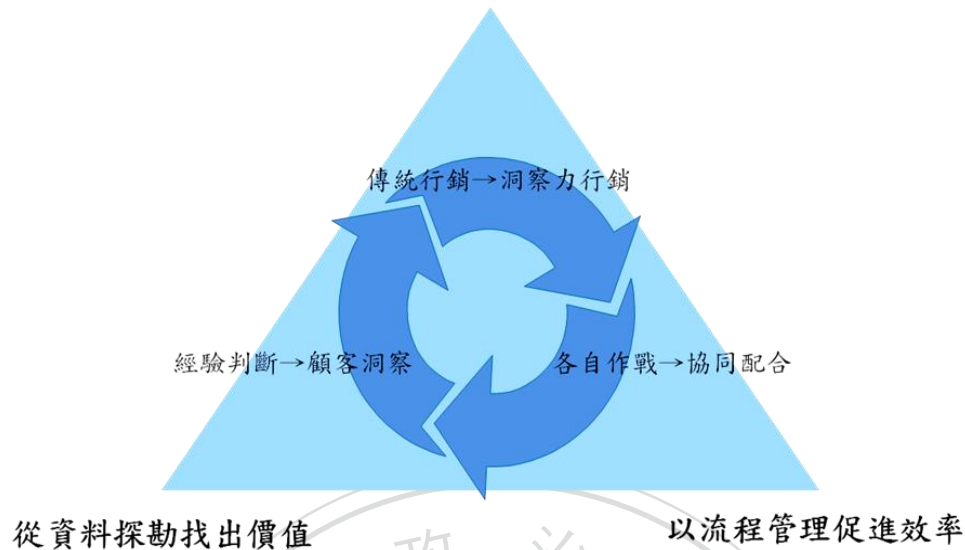


圖 2、洞察力行銷帶來的改變

資料來源：“*Insight: The Critical Path to Competitive Differentiation for Growing Companies*”, SAP, IDC executive brief, 2011

未來的市場行銷勢必以分析為策略的基礎，藉由顧客洞察力分析以獲得每一位顧客的特徵與行為資訊，針對每個顧客制訂差異化的行銷（Kotler, 2003）。

2、市場行銷

依據早期美國行銷學會（AMA, American Marketing Association）於 1960 年定義：行銷是引導物品與勞務從生產者流向消費者或使用者的企業活動。

然而，至 1985 年進一步修正，認為：行銷是規劃與執行產品或勞務的訂價、促銷與分配的一種過程，並創造交換行為，使個人與組

織滿意。

直到 2004 年針對現代蓬勃發展的行銷活動給與新的意義：行銷是組織的功能，行銷是企業一連串的活動，目的在創造價值、溝通價值，與傳遞價值予顧客，同時經營顧客關係使組織與其利益關係人獲利。整理如下表 3。

表 3、AMA 針對行銷提出之定義

年份	定義
1960 年	行銷是引導物品與勞務從生產者流向消費者或使用者的企業活動。
1985 年	行銷是規劃與執行產品或勞務的訂價、促銷與分配的一種過程，並創造交換行為，使個人與組織滿意。
2004 年	行銷是組織的功能，行銷是企業一連串的活動，目的在創造價值、溝通價值，與傳遞價值予顧客，同時經營顧客關係使組織與其利益關係人獲利。

可以發現 AMA 對行銷的定義由單純的連繫產品、消費者的關係，擴大範圍至訂價、促銷、使用者滿意導向，最終創造顧客價值及企業獲利。其 2004 年提出行銷概念圖，如下圖 3 所示。

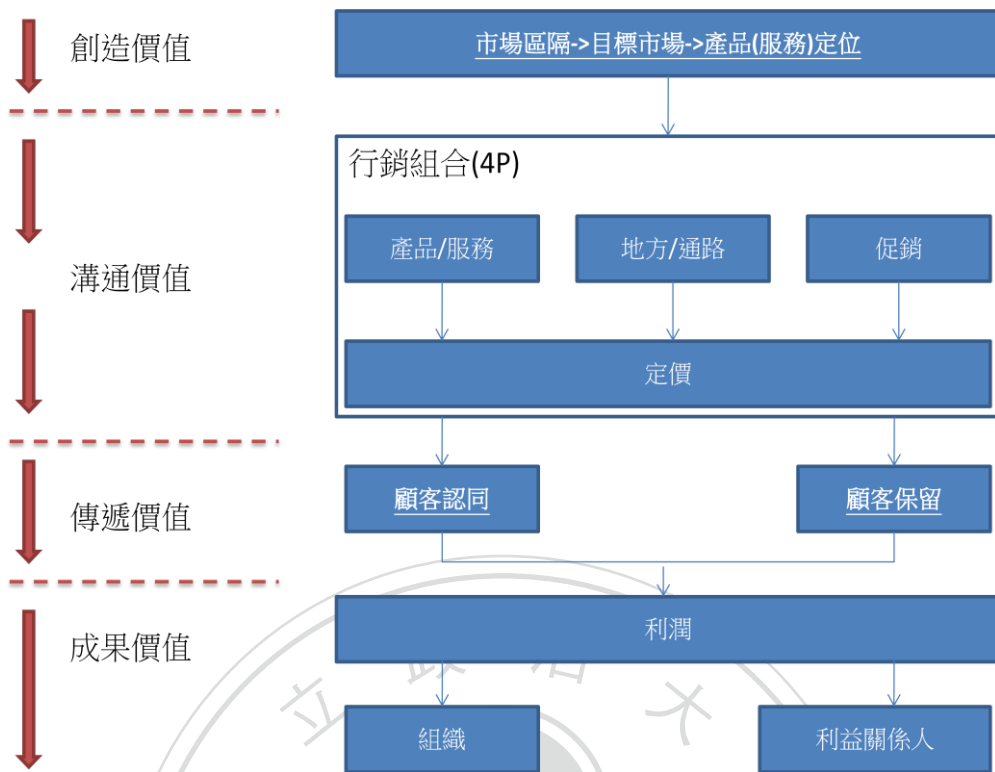


圖 3、AMA 2004 年行銷概念圖

資料來源：現代行銷學理論與實務，范惟翔，新加坡商湯姆生亞洲私人有限公司出版，2007

由 AMA 於 2004 年提出之定義可知，行銷活動是以創造價值為始點，首以市場區隔（Segmentation）、目標市場選擇（Targeting）、產品（服務）定位（Positioning）做為開端。以下分別介紹。

2.1、市場區隔（Segmentation）

市場區隔的過程即是將大市場根據某些特定的區隔變數，如顧客特徵、行為、喜好、需求等屬性，將大市場切割成幾個有明顯差異區隔市場，針對每個區隔市場進行不同層次的策略制定、設計產品或其行銷組合，增加產品豐富性以降低單一風險，以滿足不同區隔市場消

費者的需求，使企業因市場的區隔而增加獲利。

一般用來進行市場區隔的變數有以下幾類：

- 地理變數
- 人口統計變數：如年齡、生命週期、家庭形態、所得、職業、教育程度、其他（宗教、種族、國籍）。
- 生活型態：消費者對週遭環境的觀察與感受。如建設公司以生活便利或遠離塵囂來吸引不同生活型態的顧客。
- 心理變數：指消費者的人格特質、習慣、態度。如 B&Q 的消費型態強調自我實現夢想的滿足感。
- 產品相關的行為變數：如使用時機、為消費者帶來的利益、使用狀態（分為過去使用、初次使用、未來使用）、使用率、品牌忠誠度、消費者對產品的態度。

市場區隔變數越能精確描述顧客心理、行為者行銷愈容易成功，但相對的，卻越不容易收集與實行；不同的變數在實行市場區隔之難易程度如下圖 4 所示。

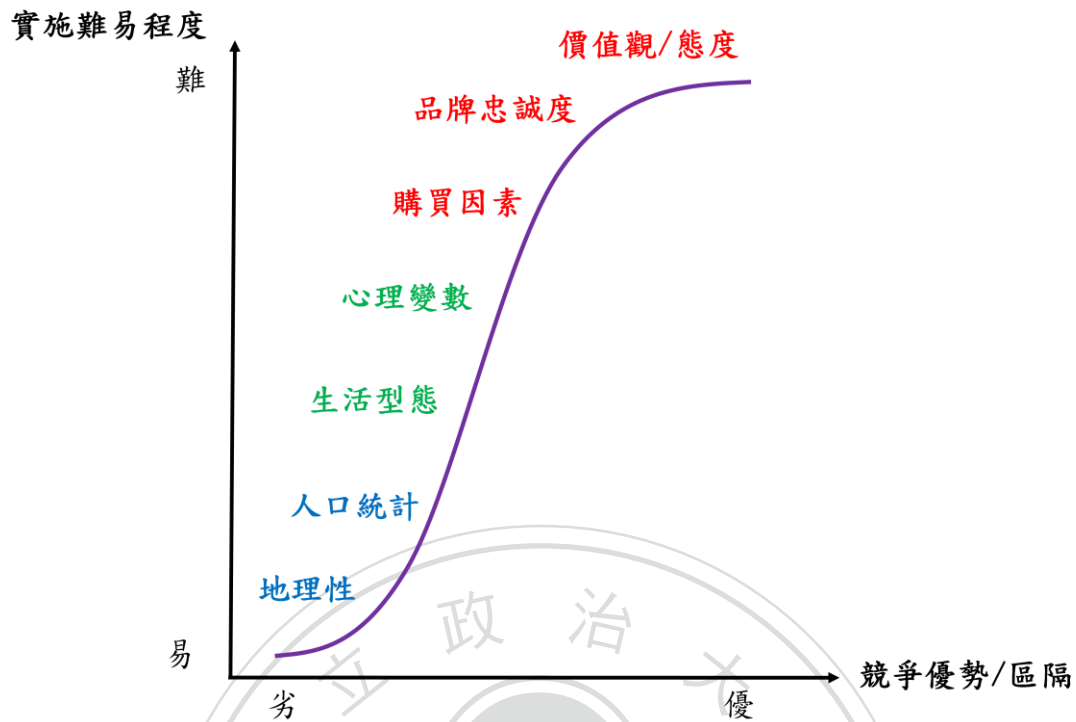


圖 4、不同市場區隔變數之實行難易程度

不同的區隔變數所區隔出的市場結果會得出不同意義，因此在進行市場區隔時，需考慮到該區隔方法是否有效。有效的市場區隔或者是良好的市場區隔必須具備以下幾個特性：

- 可測量性：區隔市場的規模、購買力和基本情況等是可以利用數字表示其大小、強弱與好壞，並且可以測量出其數值。
- 可接近性：區隔市場的客戶族群是行銷通路能夠接觸的到，並且對其提供服務。如在學校附近設置安親班、文具用品店；商業辦公大樓外食人口多，故多簡餐、咖啡店。
- 可辨別性：區隔市場要在概念上容易區分，對於不同的行銷

活動和方案應該要有顯著不同的反應。

- 可操作性：必須能夠設計出有效的行銷方案吸引所區隔出的市場，並提出適合該區隔市場所需要的服務。

- 重要性/獲利性：區隔市場必須足夠大，或者大到能帶來足夠的盈利，使公司有制訂行銷策略並投入資源滿足區隔市場需求的動機。

成功的市場區隔變數可以將顧客聚合為群間相異、群內相似的特定群體。如下圖 5 例子所示，以橫軸為購買太陽能熱水器的機率，左圖以“變量 A：氣候”為區隔變數，而右圖以“變量 B：教育程度”為區隔變數。

可以看出對於購買太陽能熱水器的潛在顧客而言，顧客所在的氣候會明顯影響購買的機率，具有“所在氣候區越寒冷購買機率越大”的現象；反觀右圖，不同教育程度的顧客對太陽能熱水器的需求並沒有反映在購買機率上，可知教育程度對購買太陽能熱水器的潛在顧客便不是一個有效區隔的變數，或稱他為不相關的描述性變數。

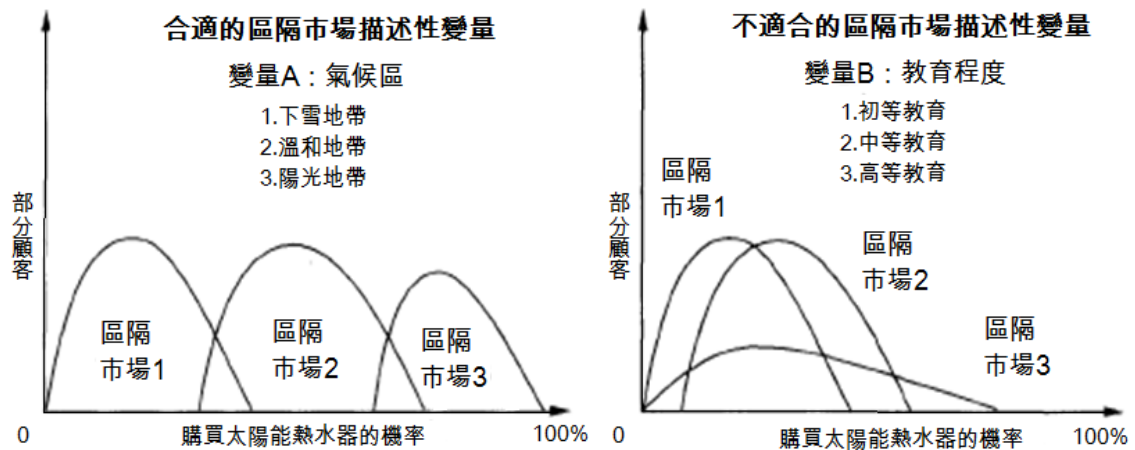


圖 5、合適/不合適的市場區隔變數之區隔結果

資料來源：“Marketing Engineering: Computer-Assisted Marketing Analysis Planning”, Gary L. Lilien, Arvind Rangaswamy, Addison Wesley, 1997

2.2、目標市場選擇 (Targeting)

目標市場選擇即是在市場區隔之後對各種型態眾多顧客族群或區隔市場依企業策略進行評量，以確定哪些區隔市場是企業所需且值得進入的目標市場，並選擇一個或多個顧客群做為企業的目標市場。在選擇評估方面，需仔細比較各區隔市場是否具備購買力、成長潛力、以及企業本身是否具有能力服務該群顧客、企業目標和資源的是否匹配等因素。

目標市場的選擇可依行銷策略分為以下幾種方式：1.規模最大的大眾行銷(無差異行銷)2.規模最小的一對一行銷(或稱微市場行銷、本地行銷)以及介於兩者之間的3.差異化行銷(區隔市場行銷)和4.集中行銷(利基行銷)。如下圖 6 所示。

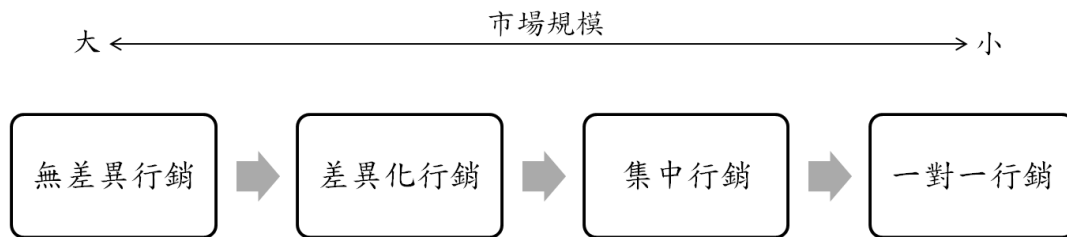


圖 6、不同行銷策略下之目標市場選擇方式

- 無差異行銷

無差異行銷（undifferentiated marketing）或大眾行銷（mass marketing）的策略是專注於顧客的共同需求，而不是需求的差異性，企業對整個市場提供相似的產品或服務，故有區隔市場間差異化小、同質性產品、低行銷成本等特徵。

- 差異化行銷

差異化行銷（differentiated marketing）的策略是決定幾個區隔市場作為目標市場，並依每個區隔市場之需求提供差異化的產品或服務，甚至規畫不同的市場行銷策略。通常需要額外的市場研究、預測、銷售分析以及通路管理等成本，才可能在目標市場內創造更大的銷售利潤，相對無差異行銷之行銷成本高，因此企業必須在成本與銷售同時增加之間求取平衡或做出取捨。

- 集中行銷

集中行銷(concentrated marketing)或利基行銷(niche marketing)的策略在企業資源有限條件下，針對一個或多個區隔市場，集中資源提供相同的產品或服務，追求該區隔市場的市場占有率及投資報酬率。集中行銷策略的目的是瞭解這些區隔市場的需求，並且在所提供服務的區隔市場上維繫著很好的聲譽，以便取得主導與強勢地位。

- 一對一行銷

一對一 (one-to-one marketing) 是指透過與顧客一對一的互動，為其個別需求量身訂作，大量客製化；以大量生產的技術，提供低成本並趨近於個別訂製的產品與服務。

各行銷策略如下圖 7 示意。

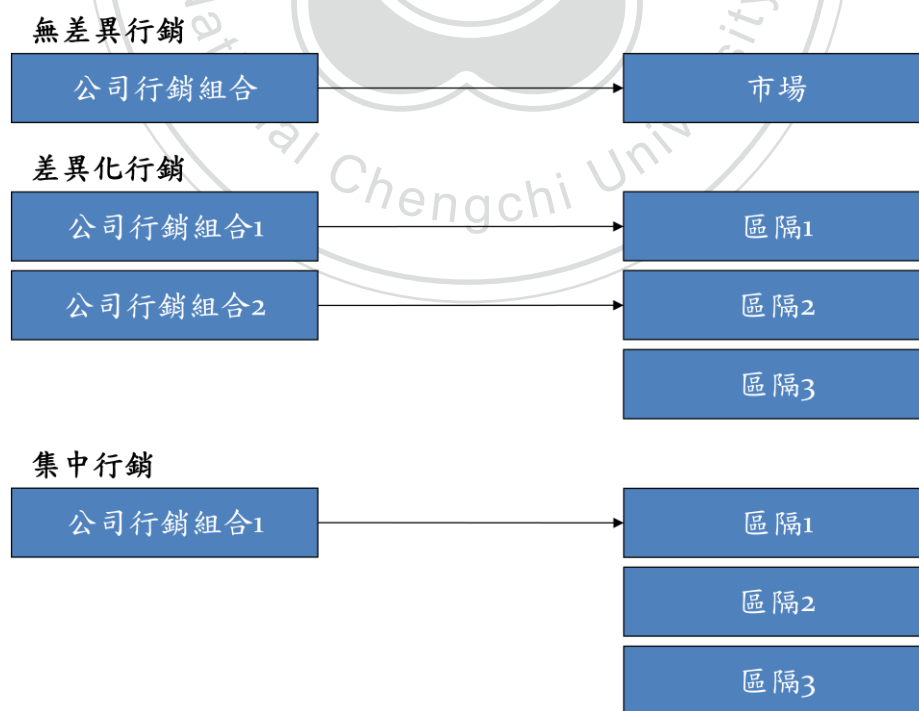


圖 7、三種目標市場選擇策略

資料來源：現代行銷學理論與實務，范惟翔，新加坡商湯姆生亞洲私人有限公司
出版，2007

2.3、產品定位 (Positioning)

產品定位是為了在目標市場之消費群心中建立與其他競爭者不同的形象，根據目標市場所需，決定提供何種商品或服務，並設計合適的市場行銷計畫；故產品定位也可以稱之為品牌效應，指企業利用行銷手法將某一產品在目標客戶群的心目中建立相對於競爭產品而言的形象、個性等差異處而形成購買偏好 (Blattberg, 2009)。

企業在做產品或品牌定位時需考慮的不只產品本身及消費者需求，更重要是認清同樣位於市場上的現有競爭對象，因此，在制定定位策略時首要確認競爭者及其在市場上的定位，分析顧客次之，最後才決定公司產品位置的選擇；一旦產品經過定位，隨著時間及市場的變化需進行不斷的調整，因此有必要監控產品的位置，以評估定位策略。其定位策略步驟如下圖 8。

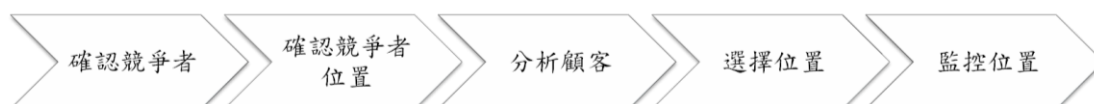


圖 8、定位策略步驟

資料來源：現代行銷學理論與實務，范惟翔，新加坡商湯姆生亞洲私人有限公司
出版，2007

在認清競爭對象及消費者需求時，除了依靠管理者直觀的方法來了解市場結構，亦可以應用知覺圖產品定位工具來幫助決策。知覺圖由點、向量、縱橫兩軸三元素組成，分別代表該市場上產品、顧客感官屬性與產品的基本屬性。如下圖 9 為啤酒市場之知覺圖。

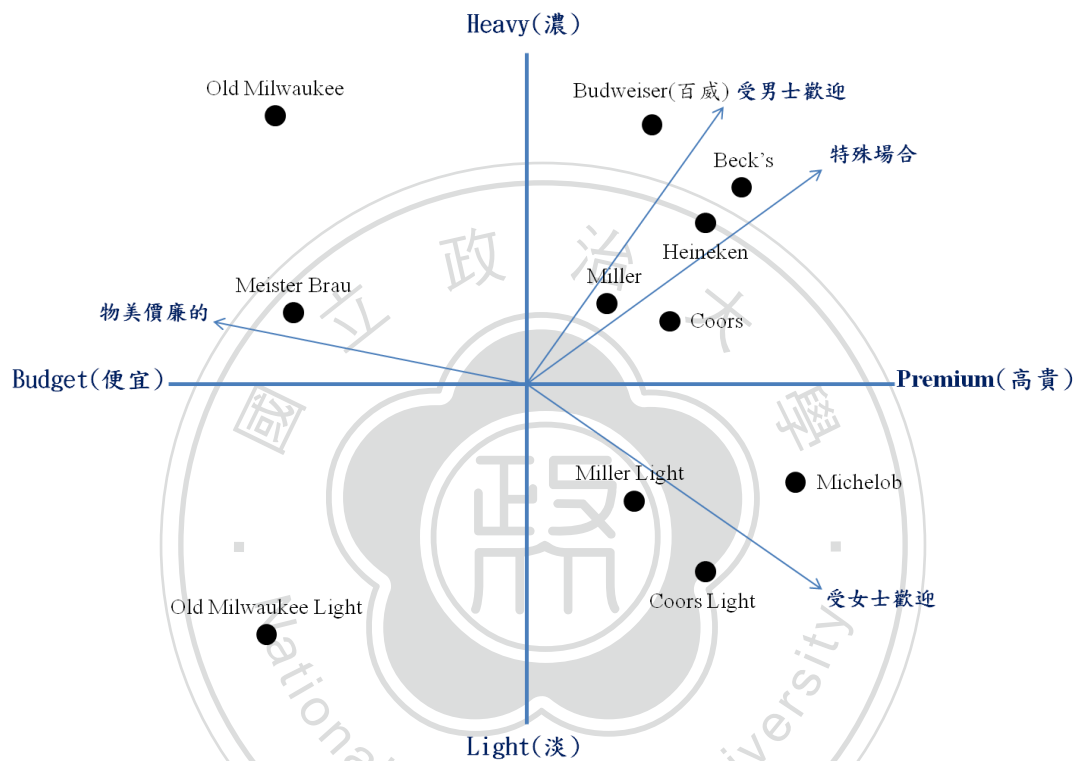


圖 9、啤酒市場的產品知覺圖

資料來源：“Marketing Engineering: Computer-Assisted Marketing Analysis Planning”, Gary L. Lilien, Arvind Rangaswamy, Addison Wesley, 1997

每個點分布在以啤酒口味（淡-濃）與價格（貴-便宜）所圍出的二維空間中，不同品牌的啤酒依其產品本身屬性座落在不同象限中，如圖中百威啤酒是屬於口味較濃且價格偏高的啤酒，點與點之間距離越近代表代消費者心中兩者相似程度、互相替代的程度越高，因此兩

產品間競爭也越激烈；另外，向量標示出市場上消費者對不同啤酒的感官資訊，如特殊場合適合喝的啤酒特徵是口味偏濃且價格偏高的，如圖中 Beck's 啤酒，而百威啤酒是男士心中最喜愛的啤酒。

由上例可知，知覺圖可以使管理者能夠將產品市場結構中的關鍵要素進行歸納和形象化，進而進行策略的思考。

第二節、資料倉儲

W. H. Inmon 定義資料倉儲為支援管理決策過程的、主題導向的 (subject orientation)、整合的 (integrated)、穩定的 (nonvolatile)、隨時間不斷變化 (time variant) 的資料集合 (Inmon, 2002)。

另一位學者 Kimball 則提出資料倉儲中的資料是自日常交易系統中抽取出來的，是一群靜態的、可支援決策的資料庫。

從技術的角度來看，具有下列三種觀點 (Kimball, 2004)：

- 資料倉儲是一個過程，而不是一個場所。
- 資料倉儲是一種架構，並不是一種產品。
- 資料倉儲是需要建造的，公司無法購買資料倉儲。

資料倉儲集合了企業資料之基礎建設，可視為環境架構的核心，以便讓使用者能快速、簡單的查詢資訊，以及各種分析來幫助管理者策略的訂定，甚至更進一步延伸資訊成為企業知識。

1、資料倉儲的特性

1.1、主題導向

資料倉儲是以企業關心的主題為導向進行資料的組織，所對於這些有興趣的主題儲存、提供相關的資料作為分析及多維度的查詢。“主題”是一個抽象的概念，主要將企業資訊系統中的資料綜合、歸類並進行分析，它對應了企業中某一分析領域所涉及的分析對象。

1.2、整合性

資料倉儲的資料來自企業操作型的交易處理系統，再加上如前所述，資料倉儲儲存資料的方式是以主題為導向，對於每一個主題所對應的資料來源分散在各個系統資料庫中，有許多重複且不一致的現象，其中可能包括了上百種不同的資料結構、不同的格式、不同的屬性、不同的命名等等；另外，資料倉儲在資料的應用上，有許多資料需要經過加總、轉換的處理而無法直接自來源資料取得。對資料倉儲的建置而言，整合是一項最重要也最複雜的工程。如下圖 10。

企業作業型環境

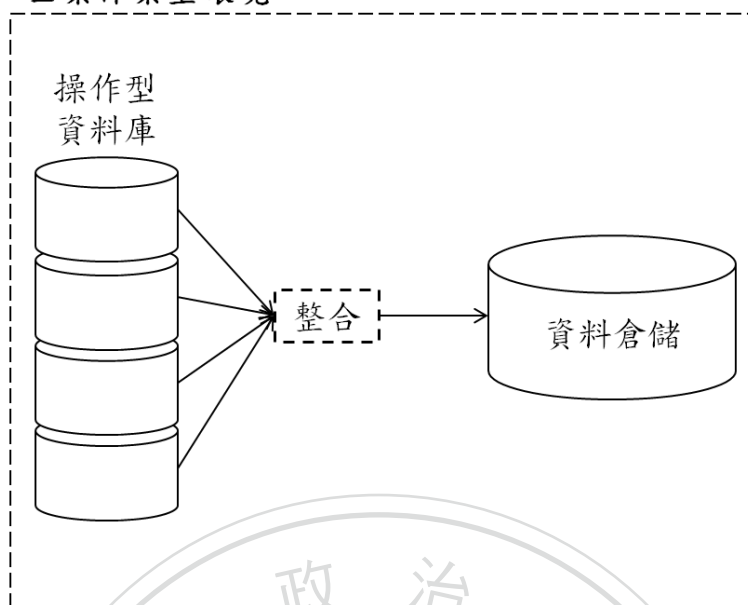


圖 10、資料倉儲特性-整合性

1.3、隨著時間不斷變化（時間差異性）

“時間”在資料倉儲中是重要的元素，可視為資料的戳記，每一筆資料是在某時刻產生的快照，累積快照使資料倉儲保留了每個時間點活動和事件的歷史記錄；故資料倉儲往往累積了 5~10 年長時間的資料，並且以時間為資料表主鍵，一筆正確的快照資料一旦被累積就不會被修改，如下圖 11 所示。

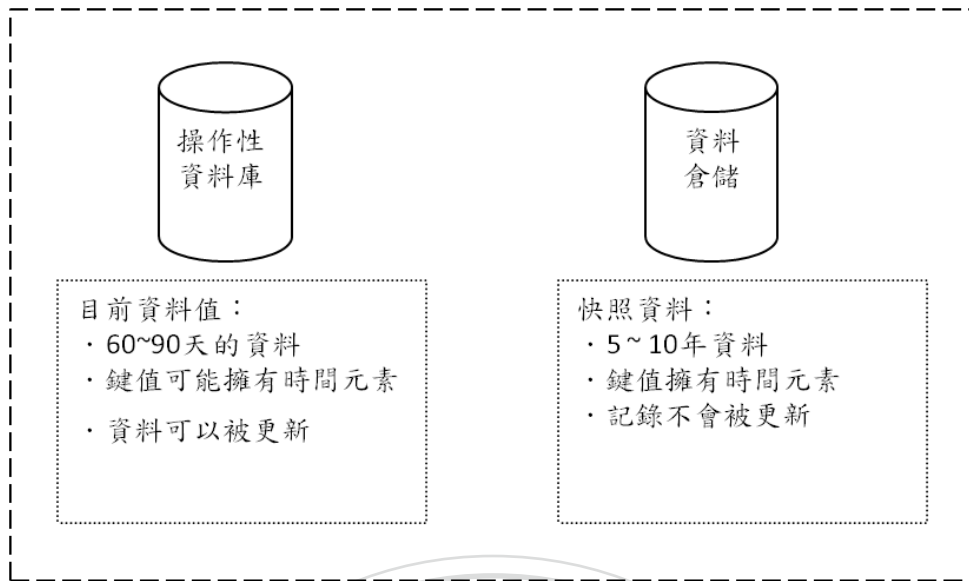


圖 11、資料倉儲特性-時間的變化性

資料來源：Building the Data Warehouse, William H. Inmon, Wiley, 3 edition, 2002.

1.4、不可更新性

傳統作業處理環境會不斷新增、刪除、修改每筆資料，如顧客基本資料，但在資料倉儲中只有載入與存取兩種動作。當快照轉換到資料倉儲內，其內容就不會再被改變，以確保資料所反應當時的狀況。

如下圖 12。

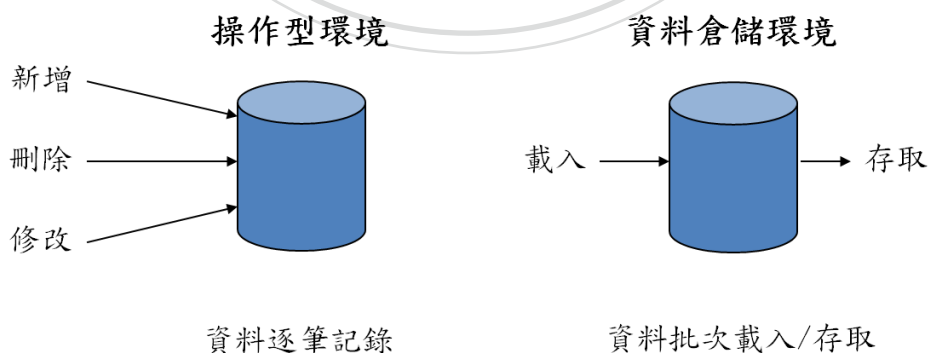


圖 12、資料倉儲特性-不可更新性

資料來源：Building the Data Warehouse, William H. Inmon, Wiley, 3 edition, 2002.

2、資料倉儲的應用

在應用上，資料倉儲可以作為線上分析處理（Online Analytical Processing，縮寫為 OLAP）和資料採礦（Data Mining）的資料來源。

OLAP 主要功能在於操作靜態資料倉儲的資料，以星狀模型（Star - Schema）方式作為資料儲存的結構，見圖 13，將資料組成資料立方體（Cube）來做儲存，見圖 14。

星狀架構是由事實資料表（Fact Table）儲存各維度資料表的主鍵，以及欲觀察的測量值（Measure），這些測量值通常具有可累加的特性；及一連串記錄各個資料表屬性的維度資料表（Dimension Tables）所組成。

這樣的資料結構讓使用者可以方便的進行多種組合、多種角度、動態多維度分析，並提供視覺化操作、上捲（Roll - Up）、下鑽（Drill - Down）、切片及切丁（Slice and Dice）以及轉軸（Pivot）等資料操作功能與展示介面。

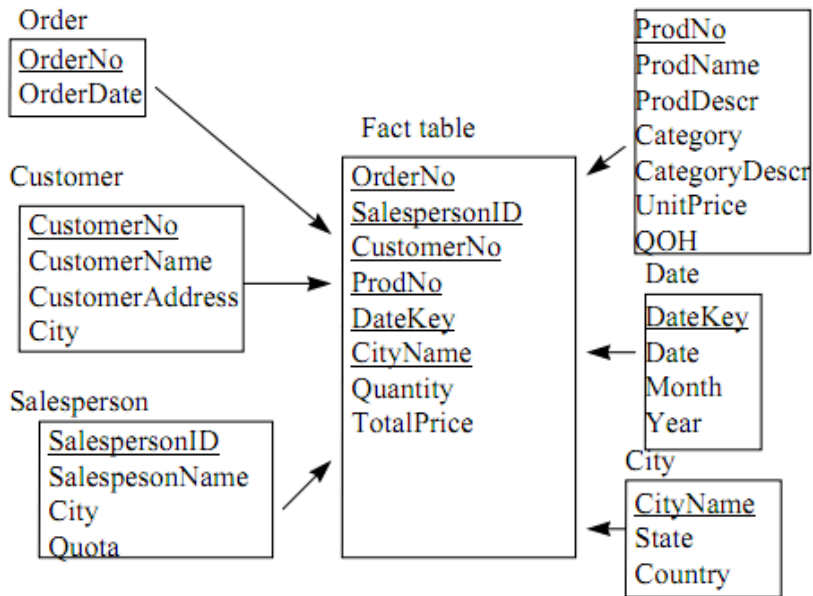


圖 13、星狀模型

資料來源：“An overview of data warehouse and OLAP technology”, S. Chaudhuri and U. Dayal, ACM SIGMOD Record, 1997.

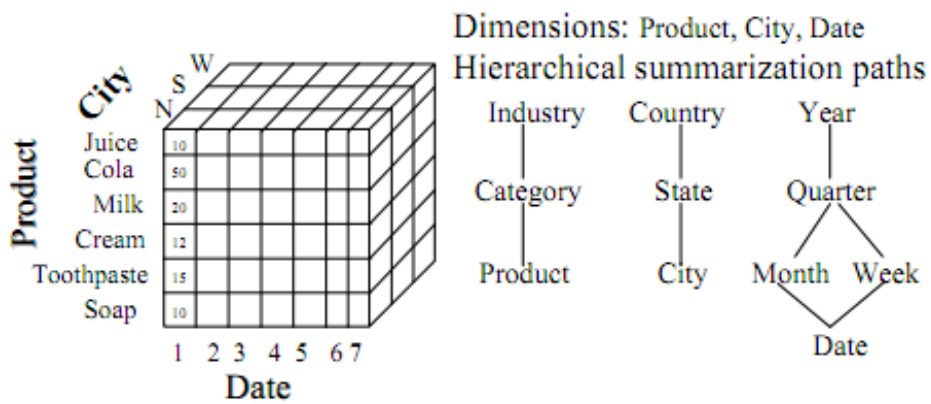


圖 14、資料立方體

資料來源：“An overview of data warehouse and OLAP technology”, S. Chaudhuri and U. Dayal, ACM SIGMOD Record, 1997.

關於資料倉儲於資料採礦的應用詳見下個章節，3、資料採礦與演算法介紹。

第三節、資料採礦與演算法介紹

1、資料採礦

前一章節所介紹的 OLAP 主要是透過以定義好的靜態資料進行動態多維度的檢視，與 OLAP 不同的是，資料採礦著重在運用演算法，在資料中找出隱藏的規則，並利用這樣的規則來預測未來。

隨著資料庫技術的迅速發展與成熟，以及資料庫管理系統的廣泛應用，企業累積的資料越來越多，對公司而言資料是一大資產，如何從資料中得到有價值的資訊便成為一大課題。目前的資料庫系統可以高效率地實現資料新增、查詢、修改、統計等功能，但無法發現資料間存在的關係和規則，亦無法根據現有的資料預測未來的發展趨勢。

資料採礦 (Data Mining) 便是從大量的、不完全的、有雜訊的、模糊的、隨機的實際應用資料中，找出隱含在其中的、人們事先不知道的、但又是潛在有用的資訊和知識的過程 (Jiawei Han, 2000)。

1.1、資料採礦的步驟，參考下圖 15：

- 資料清理 (消除噪音或不一致資料)
- 資料整合 (多種資料來源可以組合在一起)
- 資料選擇 (從資料庫中擷取與分析任務相關的資料)
- 資料變換 (資料變換或統一成適合的形式；如，資料彙總)
- 資料採礦 (使用智慧方法擷取資料模式)

- 模式評估（根據某種衡量方法，識別提供的知識是否真正有效）
- 知識表示（提供視覺化結果，向用戶提供資料採礦後的知識）。

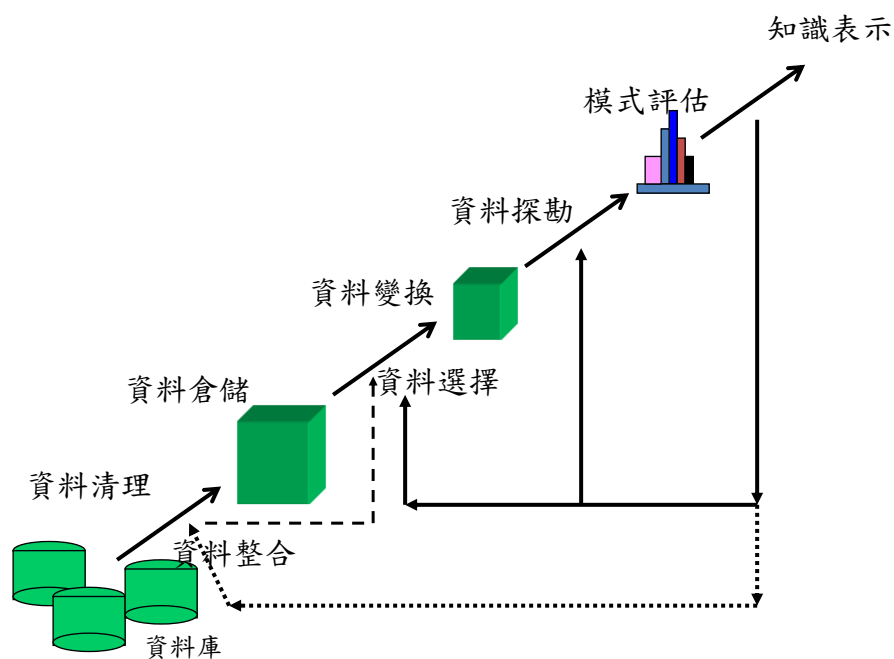


圖 15、資料採礦步驟

1.2、資料採礦的種類：

依活動內容的不同，資料採礦可分為：分類（Classification）、推估（Estimation）、預測（Prediction）、同質分組或關聯規則（Affinity Grouping or Association Rule）、群集化（Clustering）、描述及視覺化（Description and Visualization）（Michael J. A. Berry, Gordon S. Linoff,

2001)。

另外，以採礦方法分類，資料採礦可分為監督式資料採礦 (Supervised Data Mining) 與非監督式資料採礦 (Unsupervised Data Mining) 兩種。

- 監督式資料採礦

屬於由上而下 (top - down) 的方法，通常是在已知要預測的目標下進行，以演算法找出資料所屬類別的規則，並以此規則預測新資料所屬類別。分類、推估、預測三類資料採礦屬之。

- 非監督式資料採礦

屬於由下而上 (bottom - up) 的方法，演算法在資料間發現規則，並讓使用者決定是否採用採礦後的規則。同質分組或關聯規則、群集化、描述及視覺化三類資料採礦屬之。

2、演算法介紹

2.1、K-means 分群演算法

K-Means 演算法是 J. B. MacQueen 於 1967 年所提出。是把 N 個物件分為 K 個聚集，使類別內具有較高的相似度，而類別間的相似度最低，其中相似度根據一個聚集中物件的平均數 (被看作聚集的重心) 來計算。

K-Means 演算法的處理流程如下。首先，隨機地選擇 K 個物件，每個物件初始地代表了一個聚集中心，對剩餘的每個物件，根據其與各個聚集中心的距離，將它賦給最近的聚集。然後重新計算每個聚集的平均數。

這個過程不斷重複，直到目標函式收斂。目標函式的定義如下：

$$E = \sum_{i=1}^k \sum_{p \in c_i} |p - m_i|^2$$

- 此處 E 是資料集中所有物件平方誤差的總和。
- k 表共分為 k 群。
- c_i 為第 i 群資料集合。
- p 指包含於 c_i 中的所有資料點。
- m_i 是資料群 c_i 的平均數。

這個目標函式以達到 E 最小為目標，表示各群內資料盡可能的相似，群間資料盡可能的獨立。對處理大資料集，該演算法是相對可伸縮的和高效率的。

其進行的步驟如下：

- 使用者給定分群數 k
- 系統隨機指派群中心

- 系統產生初始分群結果
- 系統產生新的群中心
- 系統修正分群邊界
- 重複第三和第四步驟直到資料不再變動

但 K-Means 在應用上有一定的限制，其一，使用者必須事先給定分群群數（即 k 值）；其二，K-Means 收斂的依據是各群的平均數，因此若資料存在少量的極端值將會對平均數產生極大的影響，使分群結果產生誤差。

2.2、EM 分群演算法

EM（全名為 Expectation Maximization，期望值最大）演算法是以 K-Means 為基礎加以改善及擴充的方法。EM 不強調將個別資料分配給一個特定的群集，而是根據資料與群集之間隸屬關係發生的機率來分派，換句話說，在聚集之間沒有嚴格的界線。其目標函式是以加權後的衡量值來計算。

每個聚集都可以用機率分配來做數學描述，整個資料集可視為這些分配的混合，如圖 16 中有兩個群集，分別遵循高斯分配 $g(m_1, \sigma_1)$ 和 $g(m_2, \sigma_2)$ ，並具有各自的平均數與標準差。

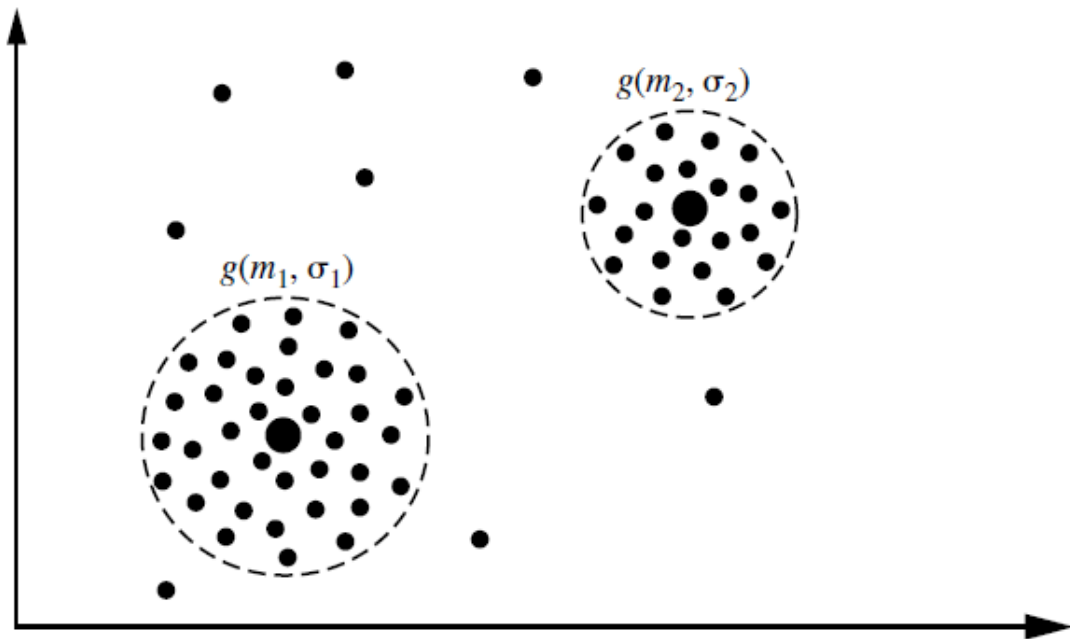


圖 16、每個群集可視為一機率分配

資料來源：資料探勘 - 概念與方法 (*Data Mining: Concepts and Techniques, 2/e*),
 Jiawei Han, Micheline Kamber, CS, 2008.

EM 首先對整個資料集的混合模型進行初始的估計，反覆根據資料集產生的混合密度對每個資料點重新計算目標函式值，每個資料點都賦予一個機率以假設它是特定群集的成員，具有該群集屬性的可能性。演算法邏輯步驟描述如下：

- E-step 期望值 (Expectation) 步驟：用以下機率將每個物件 X_i 指派到聚集 C_k 。

$$P(x_i \in C_k) = p(C_k | x_i) = \frac{p(C_k)p(x_i | C_k)}{p(x_i)}$$

其中 $p(x_i | C_k)$ 服從常態分配 (即高斯)，此一步驟是針對每個群集計算資料點 x_i 的隸屬機率。

● M-step 最大化 (Maximization) 步驟：尋找一個讓條件最大化的機率模型，利用前面得到的機率估計重新估計模型參數。

$$m_k = \frac{1}{n} \sum_{i=1}^n \frac{x_i P(x_i \in C_k)}{\sum_j P(x_i \in C_j)}$$

EM 演算法比較簡單且容易實現。在實務中，它收斂很快，但是可能達不到全域最佳。

2.3、單一鏈結分群演算法

單一鏈結演算法 (Single linkage clustering) 又稱最短距離法，屬於階層法 (Hierarchical method) 中的聚合 (agglomerative) 分群法，繼承其分群之後不做重新分群、分群之後的結果可以用樹狀圖呈現等特性，分別計算兩個群集中距離最近的兩個項目的距離，分群結果示意如圖 17。

其品質受限於：一旦群集間互相聚合就不能再回頭修正，也就是說，如果某個聚合決策在後來證明不是好的選擇，該方法無法退回並更正。

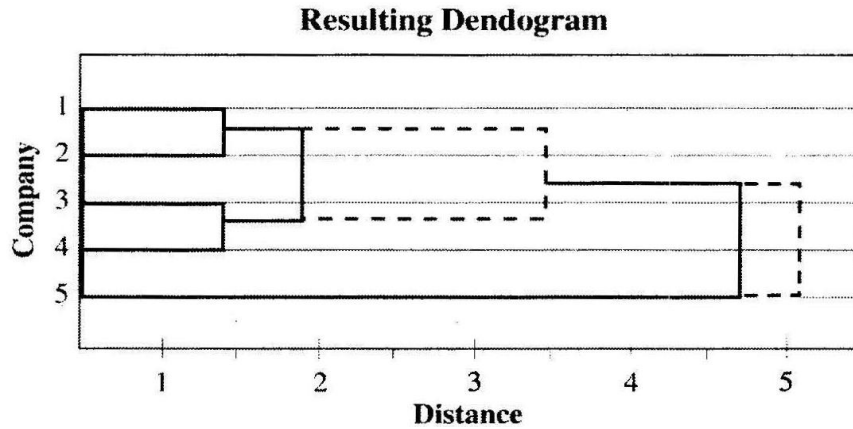


圖 17、單一鏈結分群結果示意圖

演算法邏輯步驟：

- 將所有資料視為單獨群集。
- 依據距離衡量函式所衡量出的距離大小聚合兩個最接近的群集。
- 繼續聚合相近的群集直到得到目標群集數。

距離衡量函式的方法會因資料型態有所不同，有以下幾種：

若資料型態為可數連續或離散型資料，其相似度距離可以：相關係數 (Correlation coefficient)、歐幾里得距離、曼哈頓距離、Minkowski 距離來衡量計算，以相關係數、歐幾里得距離兩種計算方法最常用。

- 相關係數：

$$D(x,y) = \frac{x_1y_1 + \dots + x_ny_n}{\sqrt{(x_1^2 + x_2^2 + \dots + x_n^2) (y_1^2 + y_2^2 + \dots + y_n^2)}}$$

其中 $x_1 \sim x_n$ 表 x 群集中的資料點， $y_1 \sim y_n$ 表 y 群集中的資料點。

● 歐幾里得距離：

$$D(x_i, x_j) = \sqrt{(x_{1i} - x_{1j})^2 + \dots + (x_{ni} - x_{nj})^2}$$

其中 i 和 j 表示一對觀測到的資料點， x_{kj} 表第 k 個變數的 i 值，

$k=1 \sim n$ 個變數。

若資料型態為不可數名目行資料，其相似度距離則以匹配衡量法計算。計算範例如下：

下表 4 有 A~D 五組織，分別針對 F1~F8 八種特徵進行調查，若具備該特徵以 Y 紀錄，反之以 N 紀錄。

表 4、單一鏈結名目資料之計算範例資料

Essential Features? (Yes or No)								
	F1	F2	F3	F4	F5	F6	F7	F8
Organization A	Y	Y	N	N	Y	Y	Y	Y
Organization B	N	Y	N	N	N	Y	Y	Y
Organization C	Y	N	Y	Y	Y	N	N	N
Organization D	Y	N	N	N	Y	Y	Y	Y

完成資料收集後，對兩兩組織進行比較，計算匹配值。以 A、B 兩組織為例，F1~F8 特徵中 F1 與 F5 兩筆記錄不同，其餘 F2、F3、F4、F6、F7、F8 等 6 筆記錄相同，故 A、B 兩組織匹配值為 6/8。以

相同方式兩兩計算可得一對稱的匹配矩陣，見表 5。

表 5、匹配矩陣

	A	B	C	D
A	1			
B	6/8	1		
C	2/8	0	1	
D	7/8	5/8	3/8	1

2.4、C4.5 分類演算法

J.Ross Quinlan 於 1993 年提出，是決策樹的一種，所產生的決策樹為多元樹。C4.5 的目標是找出哪一種規則最能夠減低分類結果的亂度（Entropy，熵值）。計算資訊增益率（Gain Ratio）來決定決策樹建立過程中以哪個分類屬性作為較上層的分類節點。

- 熵值計算：

$$E(X) = \sum_{p=1}^k -C_p \log_2 C_p$$

其中 X 表事件，k 為事件發生的結果種類， C_p 為發生該結果的機率值。

- 資訊增益率：

$$\text{Gain Ratio}(A) = \frac{E(X) - E(A)}{E(S)}$$

其中 X 表事件，A 為事件結果的屬性，E(S) 為結點分割後整

體的熵值。訊增益率越大代表該節點分割後的子集合中類別越單純，類別代表性及解釋性高，所得的分類結果越好。

2.5、CART 分類演算法

CART 全名為 Classification and Regression Tree，分類與迴歸樹，於 1984 年由 L. Brieman 提出。與 C4.5 相異處在於 CART 為二元樹，並以吉尼指標（Gini index）做為建樹的準則。

- 吉尼指標

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

其中 D 代表決策樹某節點， p_j 表節點 D 中第 j 個類別的機率。

Gini 值越小代表該節點分割後的子集合中類別越單純，所得的分類結果越好。

2.6、簡單貝氏分類演算法

貝氏分類方法是以貝氏定理為基礎所作的延伸應用，計算類別的條件機率，在統計學裡有廣泛地應用。特別適用於當輸入的屬性維度多時，並以屬性間互相獨立為前提假設，儘管這項假設過於理想與簡單，但它往往能解決複雜的現實世界中的分類預測問題。

運作原理是透過訓練樣本學習與記憶分類所使用屬性間的關係，

再用學習得到的規則對未歸類的資料進行分類預測，因此具有只需要少量的訓練資料就能準確的估計資料屬性歸類的優點。

- 計算各屬性的條件機率

貝氏定理：

$$\begin{aligned}P(c_j | a_1, a_2, \dots, a_n) &= \frac{P(a_1, a_2, \dots, a_n | c_j) P(c_j)}{P(a_1, a_2, \dots, a_n)} \\ &= P(a_1, a_2, \dots, a_n | c_j) P(c_j)\end{aligned}$$

屬性獨立：

$$P(a_1, a_2, \dots, a_n | c_j) P(c_j) = \prod_{i=1}^n P(a_i | c_j)$$

- 分類預測模型

$$\begin{aligned}C_{NB} &= \operatorname{argmax} P(c_j | a_1, a_2, \dots, a_n) \\ &= \operatorname{argmax} P(c_j) \prod_i P(a_i | c_j)\end{aligned}$$

3、資料採礦於商業應用

早期關於資料採礦的研究多著重於技術面，如針對不同型態資料的處理、演算法的效率性及有用性、以及隱私權與資料安全性方面的議題。隨著資料採礦技術的成熟，越來越多關於商業應用的討論議題。

一項 META Group 顧問公司調查顯示，自 1996 年以來，越來越

多企業投入資料採礦及知識發現技術，包括 Amazon 與 Yahoo 兩大具前瞻性的企業，皆投入資金進行顧客區隔、尋找目標顧客、提高並維持顧客保留、交叉銷售等研究，所應用的產業範圍更拓及零售、保險、金融、電信、證券 (Lach, 1999)。各領域應用情形如下表 6。

表 6、資料採礦於各領域之應用

應用領域	應用	說明
零售	類同定位 (Affinity positioning)	有效的定位產品
	交叉銷售 (Cross-selling)	為顧客找到更多產品
銀行	顧客關係管理 (Customer relationship management)	確認顧客價值，並發展方案獲取最大收入
信用卡管理	舉力 (Lift)	確認有效的市場區隔
	顧客流失 (Churn)	確認顧客可能的流動率
保險	詐欺偵測 (Fraud detection)	確認保單應受調查
電信	顧客流失 (Churn)	確認顧客可能的流動率
電話行銷	線上資訊 (Online information)	幫助電話行銷員容易取得資料
人力資源管理	顧客流失 (Churn)	確認潛在的員工流動率

資料來源：“資料探勘- Introduction to Business Data Mining”，郭志隆、張芳菱(譯)，麥格羅·希爾 (民 97)

可知資料採礦已廣泛的應用於各行業領域，幫助了解顧客區隔、找出目標客群、定位產品，甚至延伸於顧客關係管理，了解顧客流動並確保顧客忠誠。

第三章、研究方法與設計

第一節、研究方法

本研究的研究方法，主要採取個案研究法並加以系統實作進行。

1、個案研究法

個案研究是社會科學研究之方法之一。Yin (1985) 認為個案研究是一種實證研究，在現實生活環境中研究當前的現象。

並提出個案研究法的特性如下：

- 在自然環境 (natural setting) 中進行現象的研究。
- 使用多種資料收集方法。
- 分析單位 (資料收集的對象) 可能是一個 (單一個案) 或多個 (多個案) 實體。
- 對於每一個實體都深入瞭解其複雜性。
- 個案研究比較適合運用在問題仍屬探索性階段，且尚未有前人研究可循的情況下，甚至分類性研究及假設或命題的提出。
- 沒有變數操弄 (manipulation)、實驗設計或控制。
- 研究當前的現象，解決當前的問題。
- 個案研究比較適合研究「為什麼」或「如何做」的問題，並且可以做為未來相關研究的基礎。

- 研究結論的好壞與研究者的整合能力有相當大的關係。
- 改變研究對象與資料收集方法可以發展一些新的假說。
- 個案法不預設研究變數、及何者為自變數、依變數。

2、採取個案研究之原因

根據前述電力產業現況描述，可知目前仍以台電公司為主要供電來源，幾乎以獨佔者姿態立足市場，故以台電公司為個案研究對象進行分析與推論可充分代表市場現況，因此採以個案研究法，從中收集台電公司內部負載管理想法與需求，並深入探討執行面與影響面等實務運作方面的見解，方能奠立具可行性與務實性的研究結果，進而有效推行。

第二節、研究設計

1、研究對象之選擇

本研究主要配合 AMI 之裝設時程，如表 1，預計今年（民國 101 年）底可完成高壓用戶 23,000 戶之裝設進度，再加上高壓用戶雖佔總體用戶只有 0.19%，但用電量卻占台電售電量的一半以上（58%），見表 2。

故以（1）用戶資料完整性以及（2）分析成效與影響程度為優先考量，決定選擇台電公司高壓用戶為研究對象。

2、研究流程

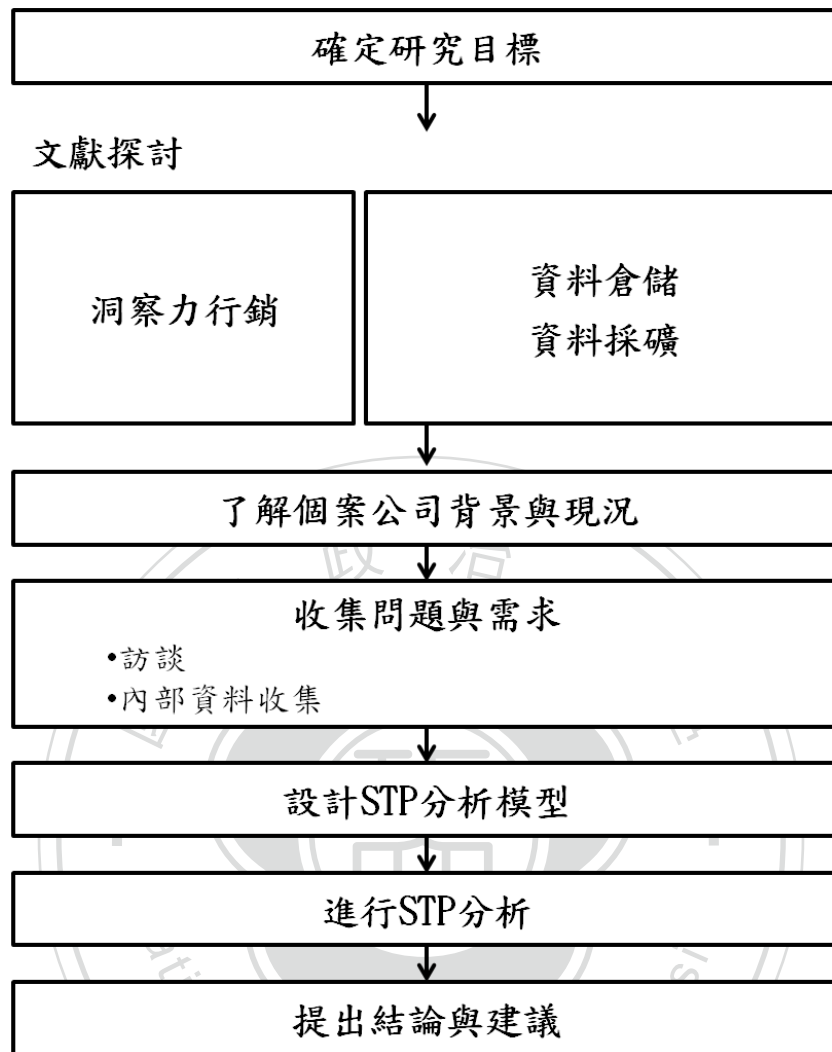


圖 18、研究流程圖

資料來源：本研究整理

本研究流程規劃如上圖 18，首先確定研究目標與問題，針對目標進行環境背景與洞察力行銷理論兩方面之文獻探討，另一方面透過訪談與資料收集了解個案公司現況，初步了解內部需求後以資料採礦技術進行 STP 分析。

第三節、研究架構

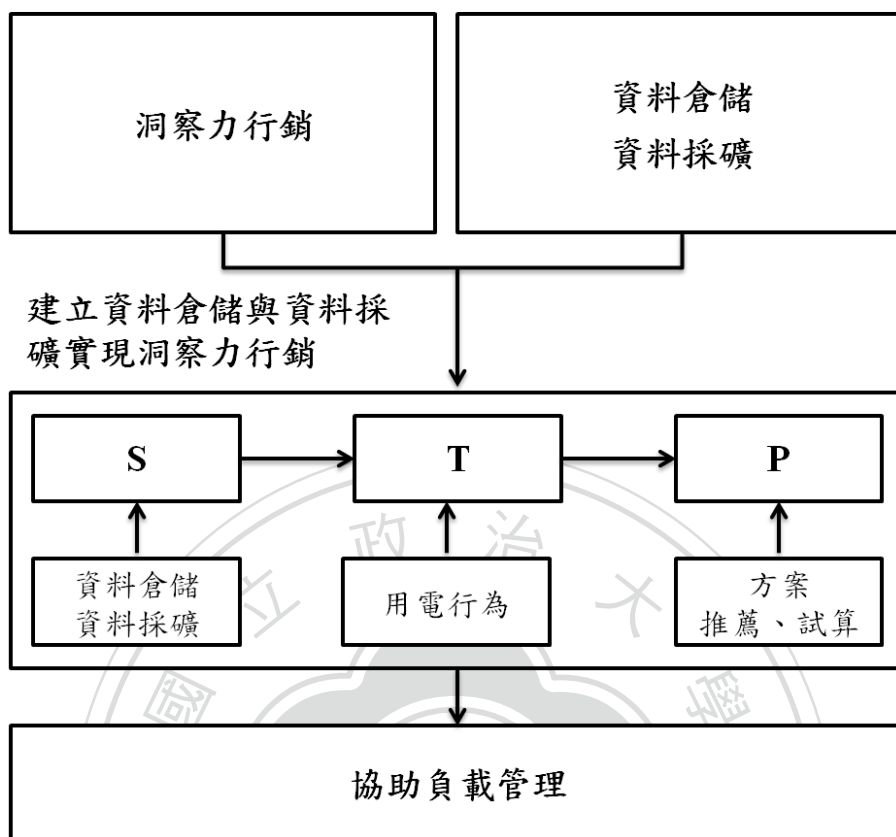


圖 19、研究架構圖

資料來源：本研究整理

如上圖 19 所示，本研究以 SQL Server 2008 Business Intelligence 為分析工具，既有用戶服務資料倉儲為基礎，進行台電公司高壓用戶洞察力行銷之實作與分析，因此研究架構包含：1.洞察力行銷 - 市場區隔、目標市場選擇、產品定位，2.資料倉儲，3. 資料採礦 - 分群、分類方法等三大部分。詳細執行步驟如下圖 20 所示。

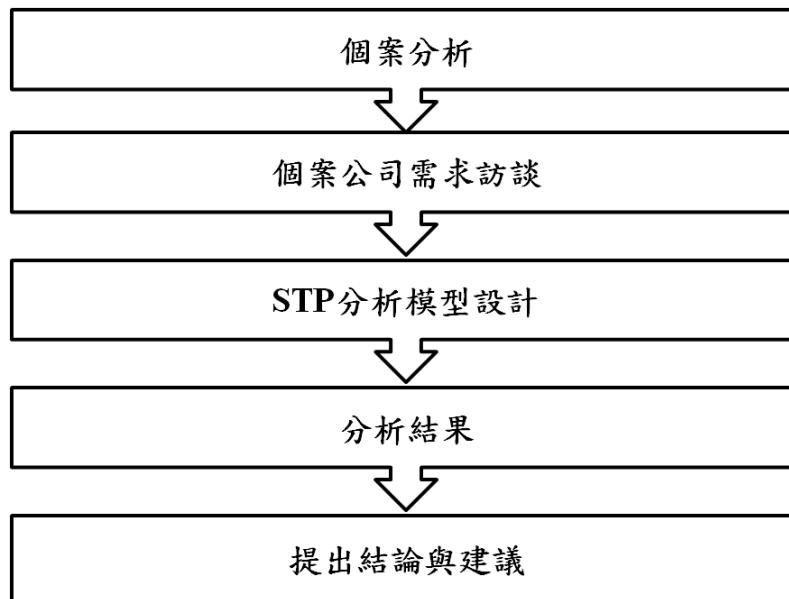


圖 20、執行步驟

資料來源：本研究整理

執行步驟說明：

- 個案公司背景分析，並進一步訪談相關單位，收集需求資訊。
- 納入訪談需求意見，藉由 STP 分析模型了解如何以資料倉儲與資料採礦技術實現市場區隔之洞察力行銷。
- S：根據資料採礦之分群、分類功能進行市場區隔，解釋與評估用戶區隔結果。T：針對各個用戶區隔進行用電行為之 KPI 試算，如尖峰用電量、尖離比，以找出若干目標市場。P：試算並推薦目標用戶合適的優惠方案，藉由優惠電價吸引用戶願意調整用電行為，以達到尖峰用電移轉效果。
- 歸納分析結果並提出建議。

第四章、個案分析

第一節、我國電力環境現況

近來福島核災滿周年，核電廠安全與核廢料問題成為討論話題，不少民間團體聯合提出廢核要求，對於核四是否順利運作的問題有不少爭論。

“針對廢核要求，台電公司表示，台灣若立即廢核，會有停限電之可能。”（中央社，2012/3/11）

1、電力供給面-台灣發電結構

台灣用電需求攀升，逐漸高於目前所能供應的電量。在電力供應面，台電統計民國 101 年 1~4 月份發電結構資料顯示，目前台灣用電來自台電供電佔 78.4%，其於 21.6% 來自購電（台灣電力公司，2012）。

台電供電又以火力發電為主要供電來源，佔總發購電 54.1%，其次為核能發電，佔 20.3%，再生能源發電第三，佔 2.5%，以抽蓄水力發電最少，僅佔 1.5%。

購電來源以民營火力發電為最大宗，佔 18.1%；汽電共生次之，佔 2.6%；民營再生能源發電最少，佔 0.9%，包含託營水力發電(0.3%)、民營水力發電(0.1%)、民營風力發電(0.5%)、太陽能發電(0.0%)。

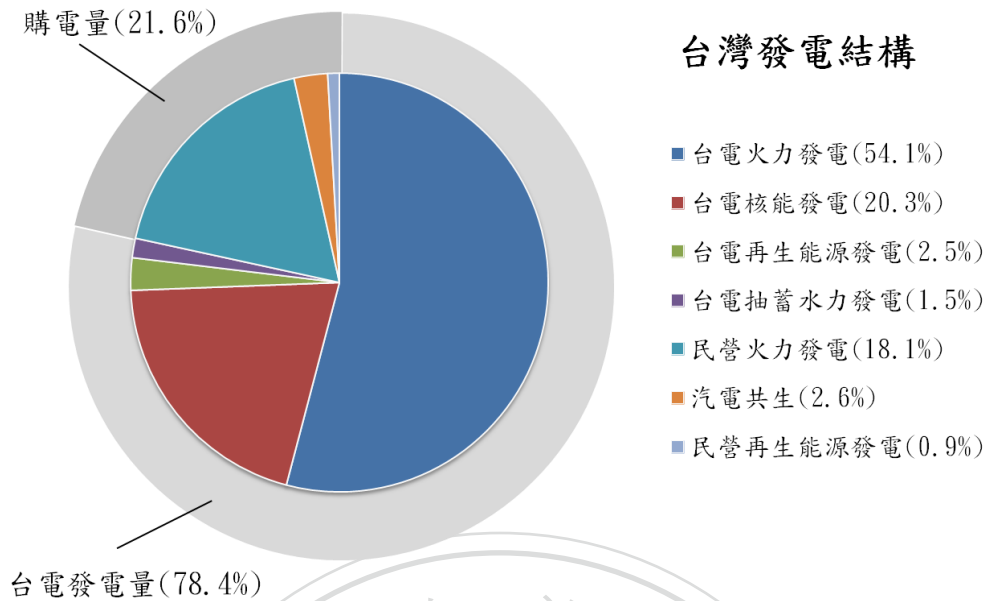


圖 21、台灣發電結構 (2012/1~4 月)

資料來源：本研究整理

2、電力需求面-台灣用電結構

經濟部能源局資料顯示，台灣主要用電需求來自：工業及能源部門用電最多占 54.49%，服務業部門次之占 19.37%。(經濟部能源局，2012/1)

台灣為獨立能源系統的海島國家，能源供給的先天條件極為脆弱，自主能源匱乏，目前約 99.4% 的能源仰賴進口，且欠缺有效的能源備援系統，由上圖 21 可知，目前台灣供電依然以台電公司為主 (78.4%)。

第二節、個案分析

1、個案公司簡介

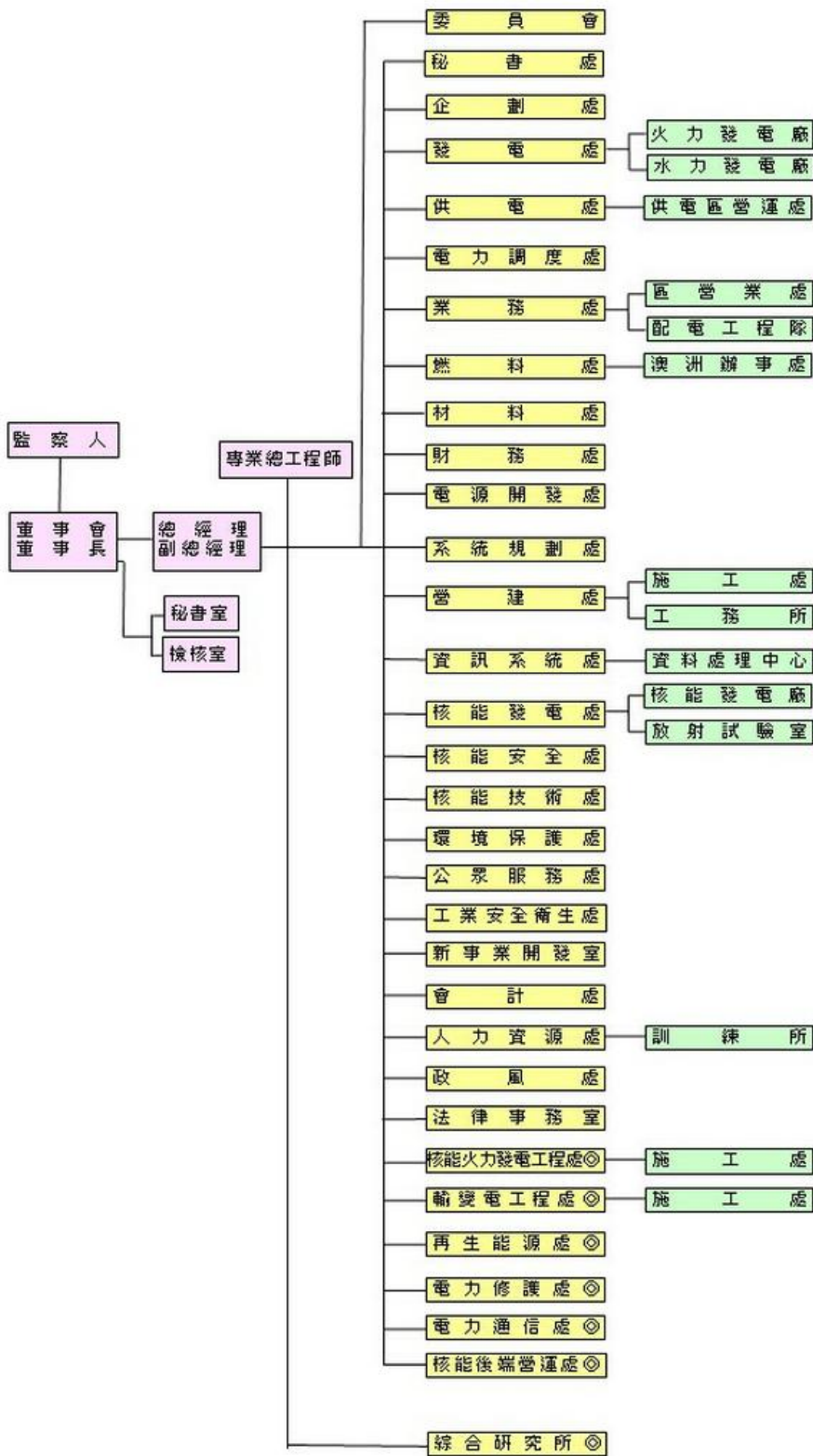
台灣電力公司，全名台灣電力股份有限公司（Taipower，簡寫為 TPC），於民國 35 年成立，負責台灣本島、澎湖、金門、馬祖的電力需求，為台灣國營電力事業，先後發展水力、火力、多能源等供電來源；民國 75 年起，因政治經濟環境改變，國內用電需求大增，台電始意識供應與需求平衡問題，直到今日，不斷思考並調整需求管理策略。

近年來，在能源來源的限制下，加以國際能源價格上漲、用電需求攀升的情況下，台電提出以下未來發展目標：

- 開發多元電源
- 加強供電品質
- 推動節能減碳
- 落實顧客服務
- 善盡社會責任

以因應台灣經濟與社會的發展，提供穩定、高品質的電力，在執行上，除了積極配合政府相關政策推動各種電力建設外，更注重質量並重的電力服務。

公司組織架構如下圖 22：



註：1.◎表示該單位為不屬總管理處之直屬單位。
 2.綜合研究所所長直接向總經理報告。

圖 22、台灣電力公司組織架構圖

資料來源：台灣電力公司網站

2、個案公司發電結構

台電供電可分火力、核能、再生能源、水力三大類。各項供電資訊如下表 7 所示：

表 7、台電公司發電結構

發電方式		發電度數 (百萬度)	發電佔比 (%)
水力發電	水力抽蓄	954	1.5
火力發電	合計	35,368	54.1
	燃油	1,770	2.7
	燃煤	20,333	31.1
	燃氣	13,265	20.3
核能發電	合計	13,271	20.3
再生能源發電	合計	1,656	2.5
	慣常水力	1,320	2.0
	風力	333	0.5
	太陽能	3,865	0.0
台電發電量合計		51,250	78.4

資料來源：101 年 5 月號台電月刊

承前述，火力發電為主要電力來源，佔台電總發購電量 54.1%，由上表數據可知，火力發電主要以燃煤與燃氣為大宗，分別佔 31.1% 及 20.3%，以燃油最少，僅 2.7%（台灣電力公司，2012）。煤、天然氣、石油皆屬化石燃料，所對應的燃料成本整理如下表 8：

表 8、台電公司燃料成本

燃料種類	成本（元/度）	發電佔比（%）
煤	1.4790	31.1
天然氣	3.1263	20.3
石油	5.5657	2.7

資料來源：101 年 5 月號台電月刊

觀察各燃料成本與發電佔比的對應關係，發現成本與發電量呈現負相關，可知台電公司在火力發電管理上，以成本低者為優先供電來源。

3、負載管理思維

電力是一種不具型體的能源形式，不像石油、瓦斯可以事先儲存，台電電力調度中心必須時時刻刻掌握用電需求，確保電廠的發電產能永遠要高於當下的用電量，以避免缺電的問題(天下雜誌, 2012/3/7)。因此，在用電尖峰時段，台電必須備足機組與發電系統以供應尖峰用電需求，在離峰時段則必須保養與維運停擺的設備，顯示台電平時有四到七成的發電機組，都是閒置的。此處所指尖峰時段一般為六到九月的夏季，此為每年的用電高峰期，而上午十時至下午五時之間則為每日的用電尖峰時段。

面對不斷攀升的用電需求，尖峰時間用電量不斷攀升，台電公司必須備足發電機組，並且使用更多化石燃料，啟動相對成本高昂的發

電設備以因應尖峰時刻發電，導致成本居高不下。

面對這樣的用電環境，不斷增購發電設備雖然可滿足尖峰時段用電，但卻造成離峰時段更多機組閒置的情況，隨之而來的是閒置機組日常維護的成本開銷，可見此法治標不治本，有效的方式應是降低尖峰時間用電量，即降低該時段的電力負載，倡導用戶節約用電，亦搭配時間電價、優惠方案等策略，以電費優惠為誘因，鼓勵用戶用電轉移尖峰時段至離峰時段。

目前除了政府政策加持、AMI 智慧電表積極推廣與裝設，台電公司方面應著重思考負載管理還能如何調整，以解決穩定供電與成本居高不下的問題。目前電力公司為了滿足用戶的用電需求，避免尖峰時段供電不足，正積極的推廣尖峰負載管理，減少尖、離峰負載的差距。

4、負載管理實施內容

針對負載管理，台電公司提出兩大策略方向：

- 針對電價基礎（Price-Based）制訂

電業根據各個時段之供電成本或價值，訂定各種隨時間變動之差別費率，用戶可根據電業公佈之不同時段費率，決定是否在某些特定時段減少用電以避免支付較高之電價，如季節電價、時間電價。

- 針對激勵基礎 (Incentive-Based) 制訂

用戶根據與電業事先約定之生效條件，於電力系統高載期間或電力供應調度需要時，配合降低用電以獲得電費扣減或補償，如減少用電方案、需量反應計畫，以及儲冷式空調、空調冷氣週期性暫停用電優惠電價等措施。

各項負載管理辦理內容如下 (台電公司網站，2011)：

- 時間電價

時間電價將一日時間分為尖峰、離峰、半尖峰時段，依不同時段發電成本不同，分別訂定各時段電價，尖峰費率最高，半尖峰費率次之，離峰費率最低。以時間電價反映不同時段之供電成本差異，借流動電費之費率高低鼓勵用戶轉移尖峰負載到離峰以減輕電費負擔。

- 季節電價

季節電價將全年畫分成夏月電費 (6/1~9/30) 與非夏月電費 (其他月份)，每年夏月期間流動電費費率高於其他非夏月期間流動電費費率。

- 減少用電方案

分為計畫性與臨時性兩種方案類型。計畫性減少用電方案係指，符合特定條件的用戶於台電所指定的期間內 (通常為夏月或尖峰時段)，依契約內容抑低用電；臨時性減少用電方案為當電力公司系統

供電不足時，請用戶配合暫停部分用電。

此兩種方案實施方式皆為用戶與台電公司簽訂契約，若用戶於約定或臨時通知的時段確實減少用電，則流動電費可享有優惠。

- 需量反應計畫

用戶配合緊急通知時減少用電，可選擇抑低用電前 15 分鐘、30 分鐘或前 1 小時通知之抑低用電方式，視依約降載量程度享有基本與流動電費折扣。適用於經常契約容量 500 瓩以上之高壓或特高壓用戶選用。

- 儲冷式空調系統

為抑低成長快速冷氣負載，台電公司鼓勵企業用戶或辦公大樓裝置夜間儲冷、白天釋冷的儲冷式空調系統，充分利用夜間離峰時段廉價的電力及優惠折扣，轉移尖峰負載。

- 空調系統週期性暫停用電

為抑低造成夏季尖峰的冷氣負載，對中央空調系統用戶，若達到每運轉 60 分鐘暫停 15 分鐘，便給予用戶電費優惠。

第三節、個案公司需求訪談

AMI 電表上線後，為更有效進行負載管理，找出重點目標用戶為首要工作，故選以與用戶直接接觸的前線單位為需求訪談對象：北

市區處大用戶服務專員、北北區處大用戶服務專員；以及業務處訪談對象：總部業務處費率組。

1、北市區處大用戶服務專員

表 9、北市區處訪談結果

項目	訪談回覆
1.	<p>用戶希望台電告知選用二段式或三段式電費何者較為划算，目前業務處僅有 Excel 檔案可進行試算，但試算結果是較粗略的估計值，且專員要自行輸入很多參數，使用不便。</p>

資料來源：本研究整理

承上表 9，項目 1.說明：

由於過去缺乏詳細用電資料，僅得知每月尖離峰總用電度數，因此大用戶服務專員在流動電費試算上只能粗略估算，被動的由用戶提出試算需求後給予建議。未來以 AMI 每 15 分鐘一筆的資料做為試算基礎，可確實掌握各大用戶用電行為，針對不同行為者進行差異化管理，將各類用戶用電資料套用各時點時間電價費率，以及該用戶可能簽訂之各契約條件，主動給予方案建議。

需求歸納：(1) 用戶分群、分類 (2) 流動電費試算 (3) 減少用電方案試算。

2、北北區處大用戶服務專員

表 10、北北區處訪談結果

項目	訪談回覆
1.	建議 AMI 資料處理時應填補缺漏資料。
2.	系統分析用戶群的用電方案選擇，要能以電費試算結果呈現才較有說服力。
3.	建議可依照供電饋線和變電所區隔用戶，並分析區隔用戶之用電行為及尖峰用電抑低移轉之試算。

資料來源：本研究整理

承上表 10，項目 1.說明：

AMI 於收集時，會因資料傳輸問題導致該用戶於某時點資料的缺漏，導致資料在後續分析時，如分群、分類，造成分析結果偏差，因此應設計資料填補的機制。

項目 2.說明：

用戶分群、分類後資訊僅能表達該群/類特性，尚不足以提供用戶服務的決策資訊，因此應整合後續電費試算，以突顯分群、分類後用戶面與公司面的效益。

項目 3.說明：

不同使用者會有不同的分析需求，北北區處大用戶服務專員主要

以饋線別與用戶所屬變電所為條件管理並服務各大用戶，因此再進行分群、分類前應予以使用者資料篩選的功能。

需求歸納：(1) 填補資料缺漏值 (2) 分析流程可設計為【資料篩選】→【分群、分類】→【電費試算】。

3、總部業務處費率組

表 11、總部業務處費率組訪談結果

項目	訪談回覆
1.	是否能以單一條件篩選出用戶，再進行分群。
2.	是否能於分完群後，依照群組的屬性條件，找出資料中其他有相同特質的潛在用戶，例如對目前有參加「計減一」的用戶分群，藉由這些用戶群的特質找出其他尚未參加方案的用戶，並向其推薦該措施。

資料來源：本研究整理

承上表 11，項目 1.說明：

同北北區處訪談結果之項目 3.說明。

項目 2.說明：

目前計畫性減少用電方案參與狀況不一，某些用戶主動參與方案

爭取電費折扣，但某些用戶可能因為相關資訊不全，尚未考慮方案的簽訂，故總部業務處費率組提出是否藉由分群、分類分析歸納出相同性質的用戶，再進一步從中探討各群/類用戶參與方案的情形，主動推薦那些尚未加入方案的用戶，以電費折扣為誘因，達到尖峰用電抑低的効果。

需求歸納：分析流程可設計為【資料篩選】→【分群、分類】→【優惠方案參與情況分析】→【電費試算】→【推薦優惠方案】。

4、用戶區隔分析需求歸納

歸納以上訪談結果如下表 12：

表 12、分析需求歸納

單位	需求內容
北市區處 大用戶服務專員	(1) 用戶分群、分類分析需求 (2) 流動電費試算需求 (3) 減少用電方案試算需求
北北區處 大用戶服務專員	(1) 資料預處理 - 填補缺漏值 (2) 分析流程可設計為【資料篩選】→【分群、分類】→【電費試算】
總部業務處 費率組	(1) 分析流程可設計為【資料篩選】→【分群、分類】→【優惠方案參與情況分析】→【電費試算】→【推薦優惠方案】。

資料來源：本研究整理

第四節、STP 分析資料模型設計

1、資料倉儲

本研究分析資料來源為台電用戶服務資料倉儲(隸屬單位：台電綜合研究所)，依據前述研究需求找出倉儲中相關資料欄項，以高壓用戶屬性主檔，及AMI電表需量資料表為資料來源，其中前者高壓用戶屬性主檔提供用戶基本屬性資料，共計123欄位，目前有24,185戶高壓用戶資料，而後者AMI電表需量資料提供用戶用電行為資料。兩表之實體資料模型圖如下圖23及圖24：

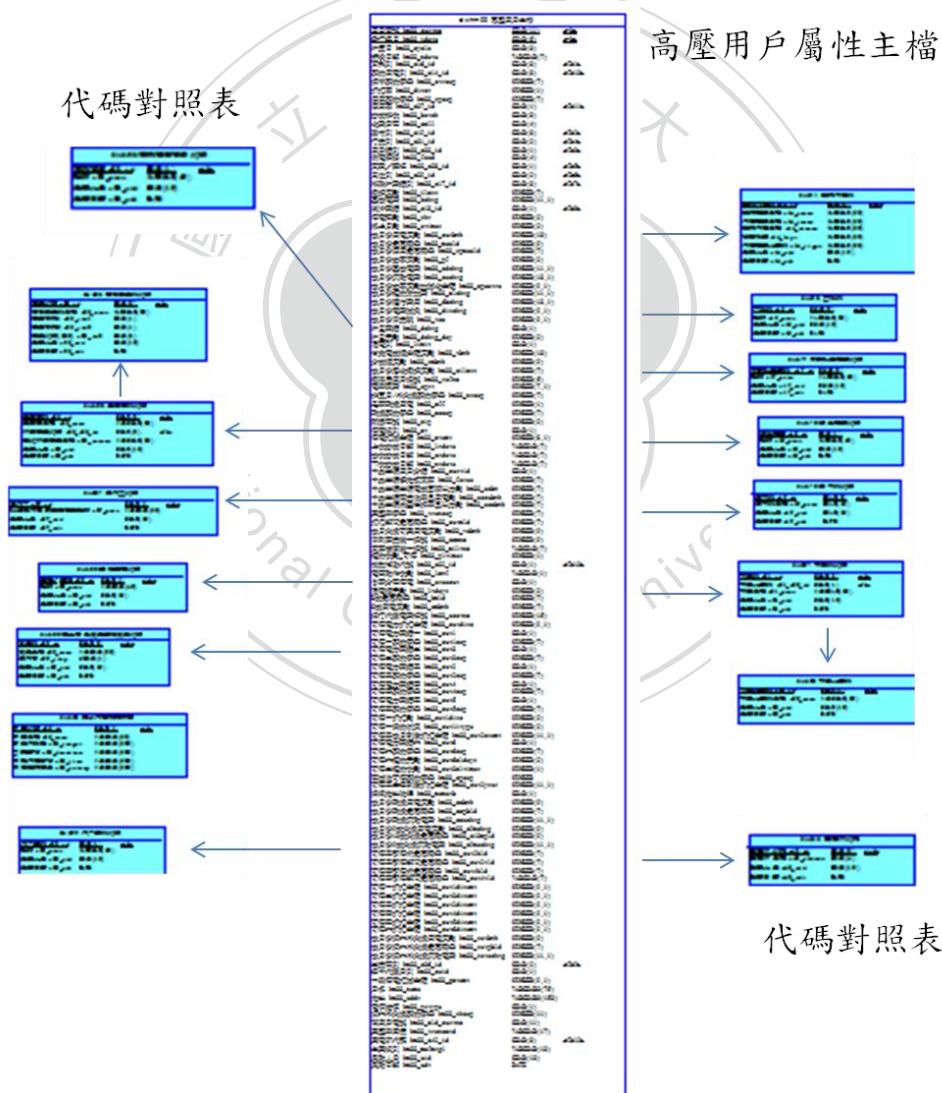


圖 23、資料倉儲-高壓用戶屬性主檔

資料來源：台電綜合研究所提供



圖 24、資料倉儲-AMI 電表需量資料表

資料來源：台電綜合研究所提供

2、資料超市

在後續分析的資料應用上，可直接抓取高壓用戶屬性主檔進行分群、分類，但 AMI 電表需量資料表的儲存結構不利分析，故需經過結構調整與轉換，以資料倉儲-AMI 電表需量資料表為基礎，抽取並轉換資料，建置為資料超市。

● 缺漏值補正

資料預處理首先設計 AMI 每 15 分鐘一筆用電度數資料缺漏值之填補機制。資料倉儲中與用電度數相關欄位有五項，除了 AMI 資料

外，高壓用戶屬性主檔會記錄該用戶於該月份總用電度數，以及該月份離峰時段用電總和、半尖峰時段用電總和、尖峰時段用電總和等，記錄於主檔是為方便台電電費計算而用，因此加總某用戶於某月分AMI所記錄的各筆資料會與主檔總用電數值相等，並且記錄該總值分布於離峰、半尖峰、尖峰時段之用電度數。示意圖如下圖 25：

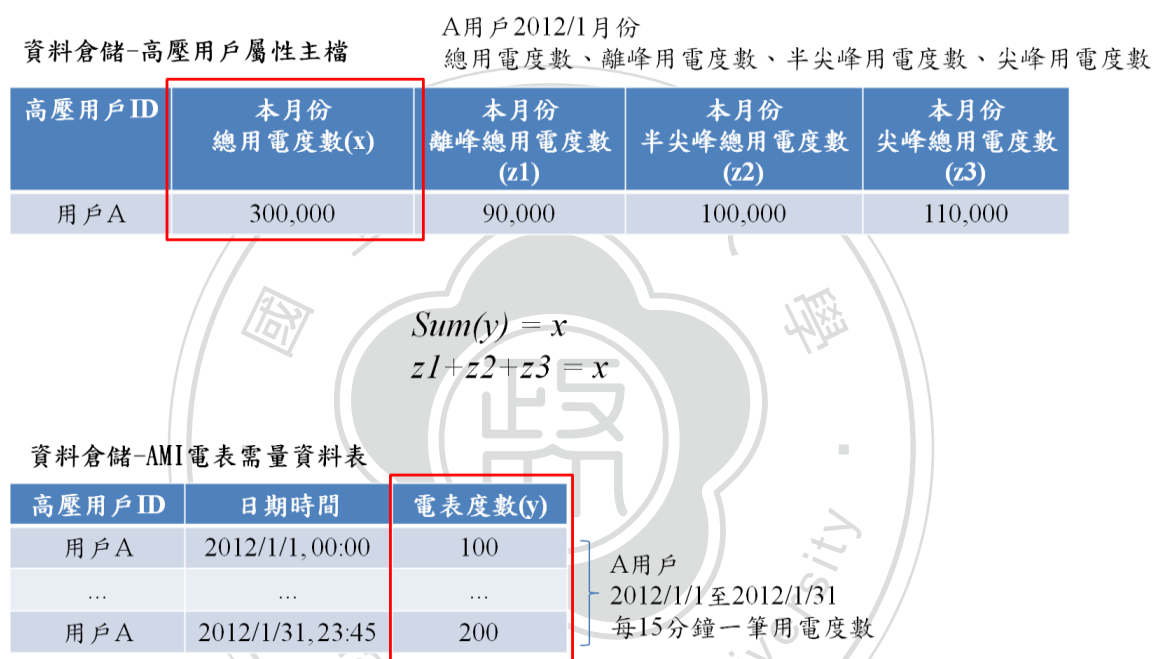


圖 25、預處理-AMI 缺漏值補正

資料來源：本研究整理

承上，可知總用電度數 (x) 與 AMI 電表度數總和 ($Sum(y)$) 之差值即為資料缺漏總值，此差值依據該時段佔總用電度數比例給予一加權乘數，平均分配給該時段缺漏值。以上圖 25 數值為例，若用戶 A 於 2012 年 1 月 AMI 表中共缺漏 5 筆資料，其中 3 筆 (a、b、c) 發生於尖峰時段，離峰與半尖峰各佔 1 筆 (d、e)，加總 AMI 資料表

中用戶 A 於 2012/1 用電度數為 298,000，數值補正計算如下：

$$\text{總用電度數} - \text{AMI 本月分用電總和} = 300,000 - 298,000 = 2,000$$

➤ 尖峰時段缺漏值補正數值

$$= 2,000 * (110,000/300,000) * (1/3) = 244.4$$

➤ 半尖峰時段缺漏值補正數值

$$= 2,000 * (100,000/300,000) * (1/1) = 666.6$$

➤ 離峰時段缺漏值補正數值

$$= 2,000 * (90,000/300,000) * (1/1) = 600$$

故資料缺漏值 a、b、c 填補為 244.4，缺漏值 d 填補為 666.6，缺漏值 e 填補為 600。

● 資料存儲結構轉換

資料缺漏補正後，為適用分群、分類分析，將原分散於 AMI 電表需量資料各列的用戶用電資料，轉存為單筆完整用戶用電資料，如下圖 26 所示：

高壓用戶ID	日期時間	電表度數(kw)
用戶A	2012/1/1, 07:00	100
用戶B	2012/1/1, 07:00	200
用戶C	2012/1/1, 07:00	150
...
用戶A	2012/1/1, 07:15	120
用戶B	2012/1/1, 07:15	180
用戶C	2012/1/1, 07:15	200
...

↓ 預處理

高壓用戶ID	日期	00:00	00:15	...	23:30	23:45
用戶A	2012/1/1	100	120	...	100	90
用戶B	2012/1/1	200	180	...	170	190
用戶C	2012/1/1	150	200	...	160	150
...
用戶A	2012/1/2	150	100	...	120	100
...

圖 26、預處理-AMI 資料存儲結構轉換

資料來源：本研究整理

● 資料比例化

比例化公式：某時點用電度數 / Max(該資料列所有用電度數)。

比例化後數值將介於 0~1 間，目的在於消去資料間量值的差距，使資料趨勢容易被比較。

如下圖 27 及圖 28 用戶 A、B 之數據及趨勢線所示，發現比例化前用戶 A 於各時間點用電量皆高於用戶 B，在分群、分類上容易被視為不同群/類，但經資料比例化後（見圖 29），可得 A、B 兩用戶在用電行為趨勢是上相似的。

高壓用戶ID	日期	00:00	00:15	...	23:30	23:45
用戶A	2012/1/1	2515	2370	...	2950	1584
用戶B	2012/1/1	1950	1834	...	2905	1683

↓ 比例化

高壓用戶ID	日期	00:00	00:15	...	23:30	23:45
用戶A	2012/1/1	0.8525	0.8034	...	1	0.5369
用戶B	2012/1/1	0.6713	0.6313	...	1	0.5793

圖 27、預處理-AMI 資料比例化前後數據

資料來源：本研究整理

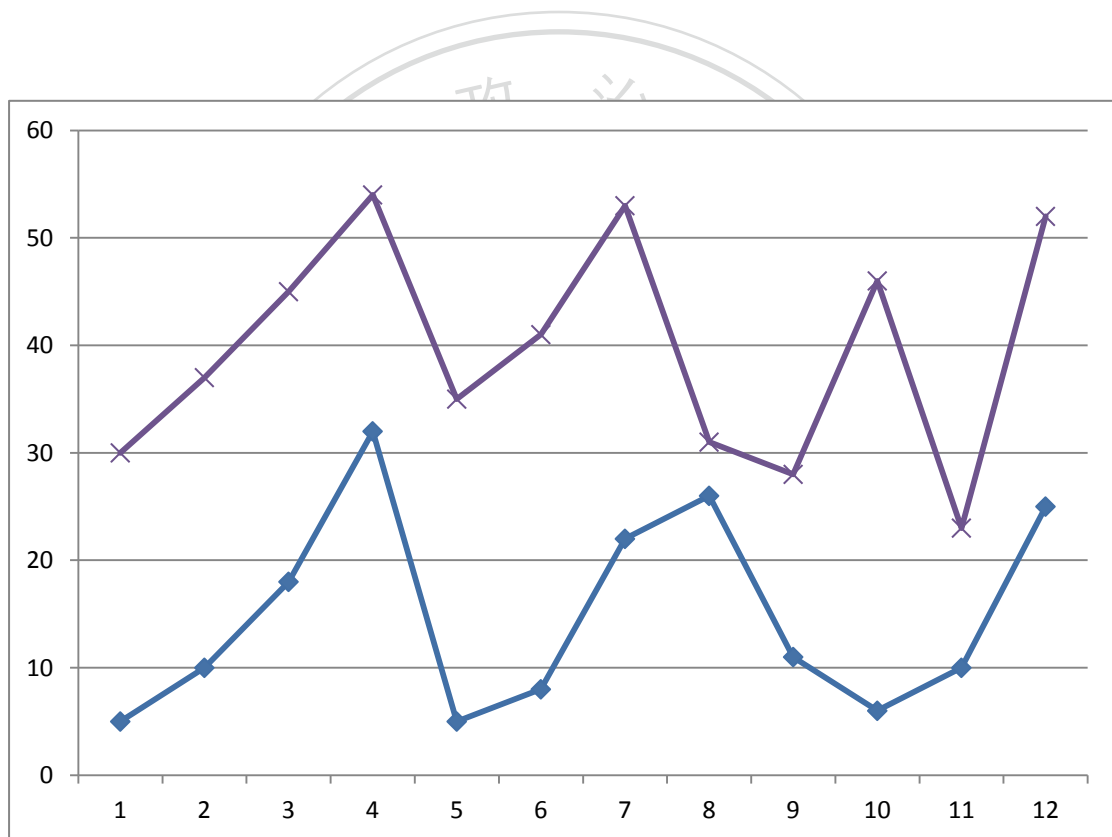


圖 28、預處理-AMI 資料比例化前用電趨勢線

資料來源：本研究整理

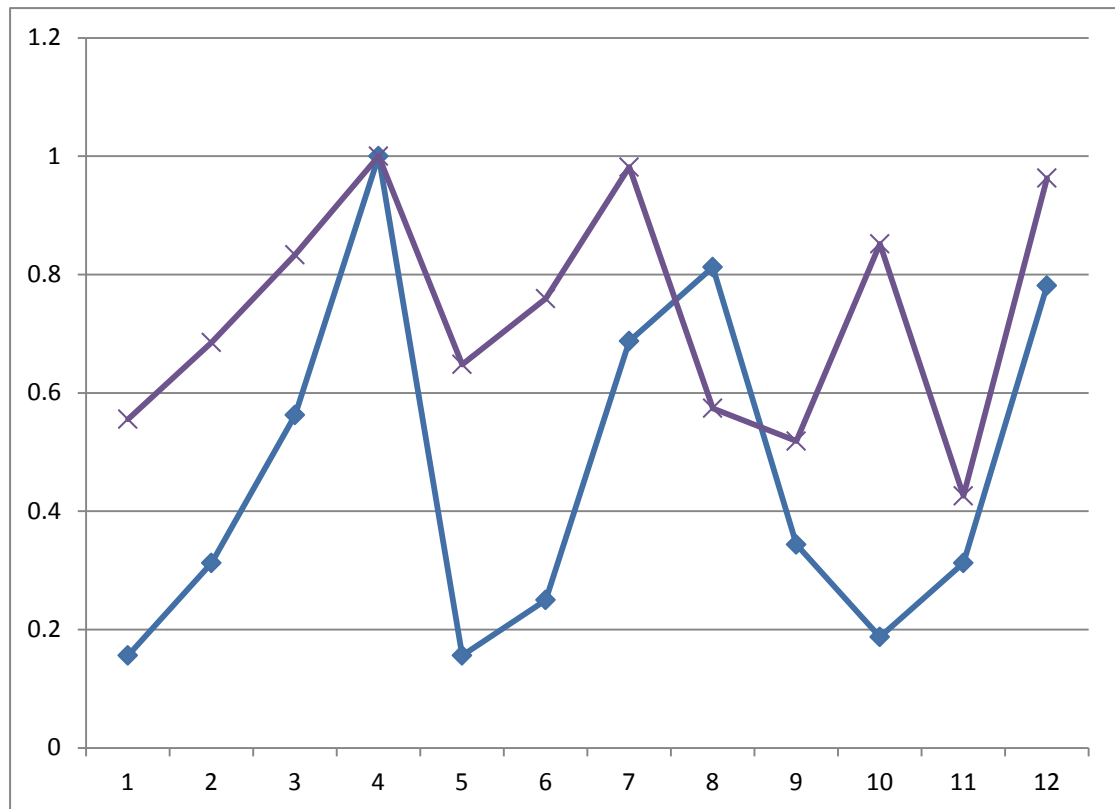


圖 29、預處理-AMI 資料比例化後用電趨勢線

資料來源：本研究整理

因此，在資料預處理時，同時保留原始用電資料與比例化後用電資料，可提供使用者在分析上視需求選擇較適當的資料形式。

● 資料超市現況

據台電綜合研究所提供資料，目前取得 AMI 資料筆數共計 85,708 筆，共計有 1,190 戶高壓用戶，資料時間範圍橫跨 2011/7/1 ~ 2011/8/30 及 2012/1/2 ~ 2012/1/30 共 3 個月份。

● 新增用電行為指標【尖離比】

用電行為指標：尖離比 = 尖峰用電量 / 離峰用電量（陳宗天，台電綜研所會議簡報，100年12月16日）。尖離比越高表示用戶用電集中在尖峰時段，反之，尖離比越低表示用電分散在離峰時段。此處尖峰時段定義為每日早上7時~下午5時，其餘時間屬離峰時段。

以尖離比來識別用戶，比值越高者越為優先改善對象，優先鼓勵調整用電行為。於分析應用上可做為目標用戶選擇的依據，或做為用戶分群、分類的屬性之一。

第五節、STP 實做與分析

本研究分為三個分析面向：用戶分布分析（以 C4.5 為例）、用戶行為分析（以 K-MEANS 為例）、建議優惠方案（以貝氏機率分類為例），探討各資料採礦模型於 STP 之應用。

1、用戶分布分析

表 13、C4.5 用戶分類用戶篩選設定

分類模型	C4.5
資料來源	高壓用戶屬性主檔
日需量	2011/7/18
測試資料	30%

分類變數	區處別、行業別、經常契約容量、尖離比
觀察變數	計畫性減少用電方案類別
條件篩選後共 1,172 戶	

資料來源：本研究整理

承上表 13 各參數設定，30% 資料為測試資料，得母體 821 戶，挑選高雄區營業處為分析對象，共 62 戶，觀察該區營業處所轄用戶，參與計畫性減少用電方案情形如下表 14 所示。

表 14、C4.5 用戶分類分析結果（部分）

區處	經常契約容量	無	計一	計二	計三	計四	總計
高雄區營業處	小於 31,400	46 (81.03%)	3 (6.9%)	1 (3.45%)	0 (1.72%)	3 (6.9%)	53
	9,420~	3 (33.33%)	0 (8.33%)	1 (16.67%)	0 (8.33%)	3 (33.33%)	7
	大於 31,400	1 (28.57%)	0 (14.29%)	0 (14.29%)	1 (28.57%)	0 (14.29%)	2

資料來源：本研究整理

由分析結果可得，高雄區營業處所轄之高壓用戶簽訂經常契約容量大多低於 9,420，且參與計畫性減少用電方案情形尚不普遍（佔約 20%）。進一步以分析高雄區營業處，經常契約容量小於 9,420 之用戶（共 53 戶），以本月份用電度數與尖離比觀察其用電情況。

表 15、本月份用電度數、尖離比分析

本月份用電度數	戶數佔比	平均尖離比
~1,200,000	66%	1.92
1,200,000~2,500,000	13%	1.17
2,500,000~3,800,000	11%	1.22
3,800,000~5,000,000	6%	1.15
5,000,000~	4%	1.00

資料來源：本研究整理

由上表 15 統計，此區隔(高雄區營業處、經常契約容量小於 9,420) 7 月份用電度數主要分布於 1,200,000 度以下，佔 66%，平均尖離比為 1.92，顯示尖峰用電量為離峰用電量 1.92 倍，用電行為存在可改善的空間。得此資訊，高雄區營業處大用戶服務專員可針對該用戶群試算各別適用之優惠方案，以誘導降低尖峰用電需量。

2、用戶行為分析

依上例延伸分析，觀察該區隔用電行為，並選擇出優先改善之目標用戶。

表 16、K-MEANS 用戶分群用戶篩選設定

分群模型	K-MEANS
資料來源	比例化後日需量
用戶篩選條件	高雄區營業處、經常契約容量小於 9,420、本月分用電度數小於 1,200,000

日需量	2011/7/18
測試資料	30%
群數	3
條件篩選後共 48 戶	

資料來源：本研究整理

承上表 16 各參數設定，同以 30% 資料為測試資料，篩選後得母體 33 戶，以 K-MEANS 分為三群，觀察各群用戶用電行為(見表 17)，並繪製用電曲線圖(見圖 30)。

表 17、K-MEANS 用戶分群分析結果

母體		33 戶			
群 1		群 2		群 3	
戶數	尖離比	戶數	尖離比	戶數	尖離比
19	<u>0.85</u>	13	<u>1.29</u>	1	<u>1</u>

資料來源：本研究整理

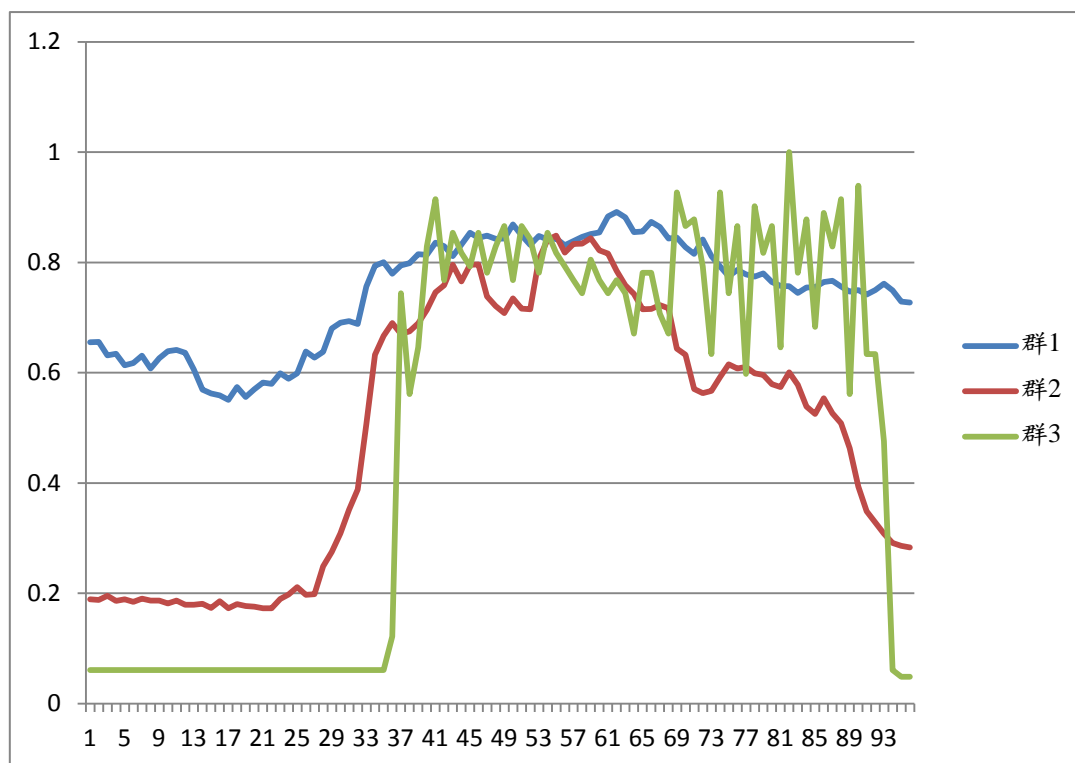


圖 30、K-MEANS 分群結果

資料來源：本研究整理

分析結果顯示，群 1 用電行為明顯於各時段皆偏高，此群佔母體約 58%，尖峰用電量與離峰用電量差距不大，且各時段用電量大，推測該用戶區隔於各時段均有用電需求，但是由群 2、群 3 得知，此區隔中依然有約 42% 用戶於尖離峰時段用電行為差異大，推測其用電行為具有調整空間。

觀察三者尖離比，群 2 尖離比最大 (= 1.29)，具有明顯尖峰用電高於離峰用電行為特徵，可進一步觀察尖峰抑低可能，故選為優先觀察對象，試算並比較不同計減方案參與前後用電抑低情形，下表 18 以用戶 A 試算結果為例。(註：此處設定抑低月份為 7、8 月，唯計

減四因契約內容規定，抑低月份為 8 月。)

表 18、用戶 A 計畫性減少優惠方案試算結果

方案名稱	原電費	優惠額度	優惠後金額	優惠比例	抑低電量
計減一	\$1,672,200	\$33,540	\$1,638,660	2.01%	375kW*2
計減二	\$1,672,200	\$234,108	\$1,438,092	14.00%	300kW*2
計減三	\$1,672,200	\$67,080	\$1,605,120	4.01%	375kW*2
計減四	\$1,672,200	\$40,248	\$1,631,952	2.41%	300kW*2

資料來源：本研究整理

參考以上試算結果，可得用戶 A 之於台電最大抑低電量為 750kW / 年 (375kW*2)，簽訂計減一與計減三皆可得此效果，其中用戶 A 由計減三預期可獲得之優惠額度 (\$67,080) 大於由計減一預期可獲得之優惠額度 (\$33,540)，故高雄區營業處大用戶服務專員於用戶 A 可推薦參與計畫性減少用電方案三，以優惠電費為誘因，誘導其改變用電行為，則台電可得抑低電量共 750kW / 年 (375kW*2)，用戶 A 可獲得 4.01% 優惠 / 年。

3、建議計畫性減少用電方案

以簡單貝氏分類演算法建立計畫性減少用電方案推薦模型，首先統計及觀察台灣各區營業處平均尖離比，從中找出用電行為優良者與優先改善者。(註：採用 2011/7/18 尖離比數據)

下表 19 列出各營業處之平均尖離比，由低至高排列，用電行為最佳者為宜蘭區營業處，平均尖離比 0.8，顯示尖離峰時段用電平均；用電行為最優先改善者為台北西區營業處，平均尖離比 2.07，顯示尖峰用電為離峰時段之兩倍，因此以台北西區營業處為此範例欲推薦計畫性減少用電方案之對象，以改善用電行為。

宜蘭區多農業，屬該營業處之大用戶僅 16 戶，考慮各營業處之用戶數與各工業區特徵，以新竹區營業處用戶做為優惠方案推薦模型之建模母體，平均尖離比 1.164，共 93 戶。

表 19、各區營業處尖離比（由低至高）

區營業處	平均尖離比
宜蘭區營業處	0.8094375
苗栗區營業處	0.965357143
新營區營業處	1.110705882
新竹區營業處	1.163870968
台中區營業處	1.17747619
花蓮區營業處	1.330777778
鳳山區營業處	1.33160177
台南區營業處	1.398559524
桃園區營業處	1.408116438
台東區營業處	1.438
屏東區營業處	1.449666667
高雄區營業處	1.534850575
基隆區營業處	1.590771429

南投區營業處	1.63
嘉義區營業處	1.667978261
澎湖區營業處	1.771241379
台北市區營業處	1.813217391
台北北區營業處	1.879555556
台北南區營業處	1.928808824
雲林區營業處	1.936923077
彰化區營業處	1.94347191
台北西區營業處	<u>2.072677419</u>

資料來源：本研究整理

以新竹區營業處所轄用戶為計畫性減少用電方案推薦建模資料來源，篩選條件設定如下表 20，及建模變數設定如下表 21。

表 20、簡單貝氏機率分類用戶篩選設定

分類模型	簡單貝氏機率分類
資料來源	高壓用戶屬性主檔、比例化前日需量
日需量	2011/7/18
區營業處	新竹區營業處
測試資料	30%
條件篩選後共 93 戶	

資料來源：本研究整理

表 21、簡單貝氏機率分類建模變數設定

建模變數	工業區別、契約用電別、縣市別、行業別、用戶類別、經常契約容量、本月分基本電費、本月分流動電費、本月分用電度數、計畫性減少用電方案類型、本月分最高需量、日需量、尖離比
預測變數	計畫性減少用電方案類型

資料來源：本研究整理

30%測試資料，得母體 66 戶，建模結果發現尖離比與計畫性減少用電方案有高關聯性，見圖 31，各尖離比區間參與方案情形如表 22。



圖 31、簡單貝氏機率分類建模結果

資料來源：本研究整理

表 22、建模結果-尖離比與方案參與關係

母體		66 戶					
屬性	無	計一	計二	計三	計四	備註	
戶數	59	3	2	2	0	總計 66	
尖離比	<u>~0.7</u>	0.068 (4 戶)	0.333 (1 戶)	1 (2 戶)	0 (0 戶)	-- (0 戶)	機率值 (戶數)
	<u>0.7~1.2</u>	0.78 (46 戶)	0 (0 戶)	0 (0 戶)	0.5 (1 戶)	-- (0 戶)	
	1.2~1.7	0.119 (7 戶)	0.333 (1 戶)	0 (0 戶)	0 (0 戶)	-- (0 戶)	

	1.7~2.36	0.017 (1 戶)	0 (0 戶)	0 (0 戶)	0 (0 戶)	-- (0 戶)	
	2.36~	0.017 (1 戶)	0.333 (1 戶)	0 (0 戶)	0.5 (1 戶)	-- (0 戶)	

資料來源：本研究整理

由建模結果可知新竹區營業處所轄用戶參與方案情形仍以無參與居多，計減一有 3 戶，計減二 2 戶，計減三 2 戶，又以參與計減二之用戶尖離比最佳。以此模型結果針對台北西區營業處進行方案推薦，篩選條件設定如下表 23。

表 23、簡單貝氏機率分類預測用戶篩選設定

分類模型	貝氏機率分類
資料來源	高壓用戶屬性主檔、比例化前日需量
日需量	2011/7/18
區營業處	台北西區營業處
測試資料	30%
條件篩選後共 62 戶	

資料來源：本研究整理

表 24、簡單貝氏機率分類預測結果

方案類型	機率
無參與	0.842857143
計減一	<u>0.071428571</u>
計減二	0.042857143
計減三	0.042857143

資料來源：本研究整理

經模型結果如上表 24，預測出計減一為適當的建議方案，因此針對此區隔中經常契約容量高於 500kw 且尖離比大於 2 者進行試算，方案參數統一設定於夏月（六~九月）中選六、七月進行抑低，每月依經常契約容量大小抑低百分之 50，試算區隔中 A、B 兩用戶為例，結果分別見下表 25 及表 26。

表 25、用戶 A 計畫性減少優惠方案一試算結果

月份	預計 可抑低比例	預計 可抑低容量	基本電費 扣減金額
6	50.00%	499.5kW	\$22,337.64
7	50.00%	499.5kW	\$22,337.64
原基本電費金額（試算月份）			\$446,752.80
預估基本電費扣減金額（試算月份）			\$44,675.28
預估基本電費扣減後金額（試算月份）			\$402,077.52
預估基本電費扣減比例（試算月份）			10.00%

資料來源：本研究整理

表 26、用戶 B 計畫性減少優惠方案一試算結果

月份	預計 可抑低比例	預計 可抑低容量	基本電費 扣減金額
6	50.00%	1050kW	\$46,956.00
7	50.00%	1050kW	\$46,956.00
原基本電費金額（試算月份）			\$939,120.00
預估基本電費扣減金額（試算月份）			\$93,912.00
預估基本電費扣減後金額（試算月份）			\$845,208.00
預估基本電費扣減比例（試算月份）			10.00%

資料來源：本研究整理

經試算，台電可分別得 A、B 抑低效果 999kW (499.5kW * 2) 與 2100kW (1050kW * 2)，並且用戶面可獲得\$44,675 與\$93,912 之電費折扣。

第五章、結論與建議

第一節、結論

1、研究結論

行銷 STP 理論可借由資料倉儲與資料採礦技術實現，經研究提出 STP 分析執行模式如下圖 32。

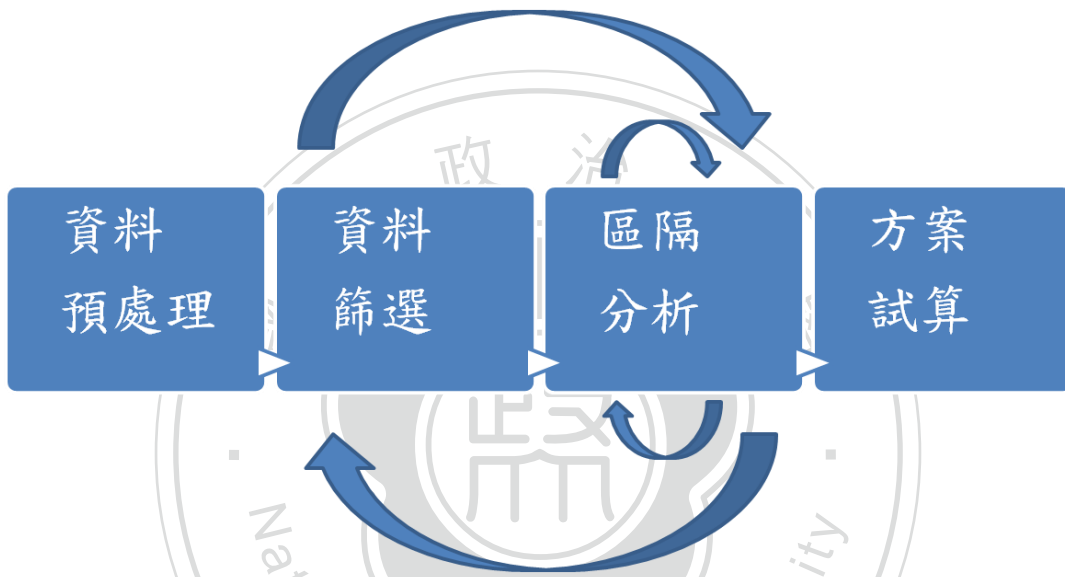


圖 32、STP 分析執行模式

資料來源：本研究整理

模式說明：

市場區隔、目標市場選擇、市場定位並非一次性可完成的過程，粗略的步驟可分為：資料預處理、分析對象篩選、用戶區隔、方案試算，步驟需經過重複執行並反覆實驗的過程；因此在 STP 過程中是需要多次進行區隔、對象篩選、目標試算才能從若干區隔中找出目標並定位產品。

本研究即說明如何以資料採礦技術執行 STP，以此分析模式設計以下系統架構（見圖 33），建議個案公司建置系統落實區隔分析，以協助負載管理之長期目標。

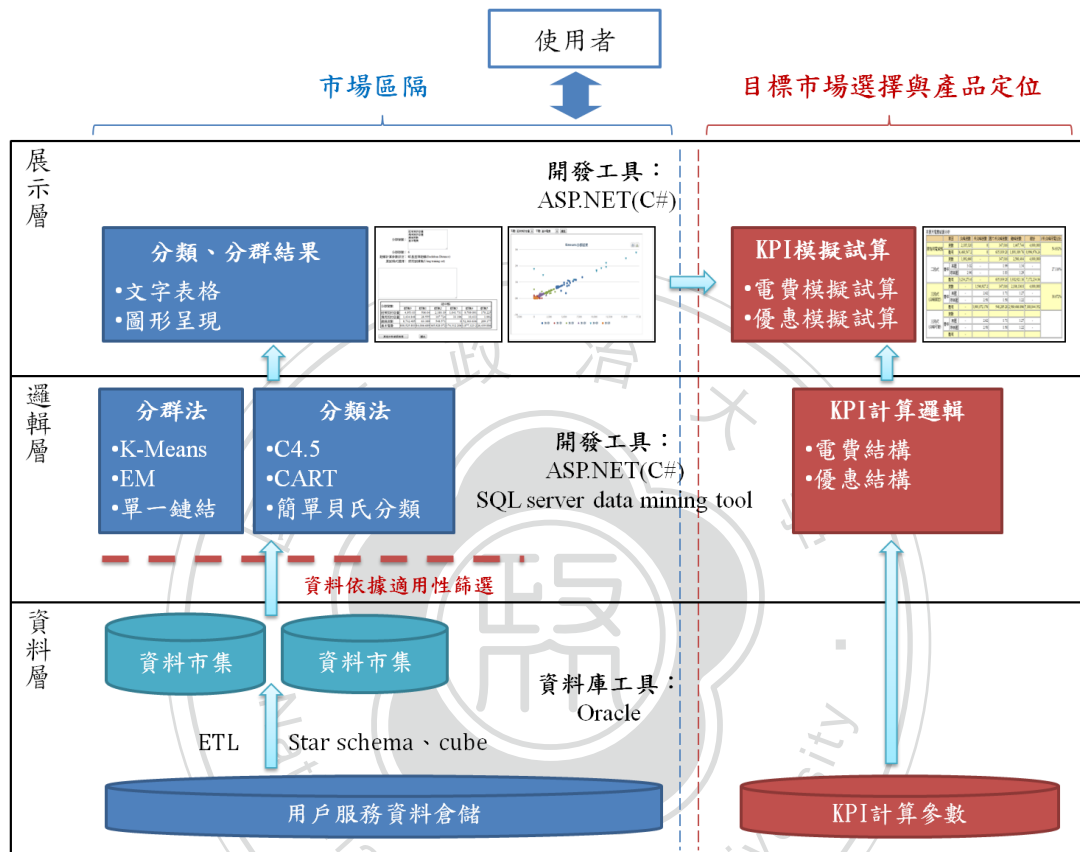


圖 33、系統架構設計

資料來源：本研究整理

2、研究限制

由於 AMI 裝置甫上線，再加上 AMI 資料收集端與資料倉儲資料拋轉機制尚未成熟，導致資料倉儲所累積之資料量不足，使可分析資料不齊全，因此本研究主要提出 STP 實踐之模式，多處數據僅供參考。

另一方面，由於研究時間限制，文獻所探討若干採礦模型無法一一設計適用情境並實際分析，故依用戶分布分析、用戶行為分析、計畫性減少用電方案推薦分析，等三大分析目標，挑選分群之 K-MEANS 方法，分類之 C4.5、簡單貝氏機率分類方法作為示範模型。

第二節、未來發展

負載管理策略如何制訂需透過經驗累積，STP 亦是，透過不斷實驗與嘗試的過程，才能描繪出詳細的台電用戶生態，當對於用戶特徵與行為有豐富且完整的認知，才能夠進一步思考服務該如何改善。

因此，在用戶特徵與行為的探索上，十分仰賴完整的資料來源，在資料面除了高壓用戶資料的完備更可以擴及低壓用戶群，以及長時間歷史資料累積，甚至結合發電端的發電成本資料，使得未來在負載分析範圍更加延展與深入。

另一方面，除了用戶知識的累積，個案公司應重新思考更多的用戶服務，計畫性減少用電方案行之有年，是否舊有的方案內容依然適用現今市場環境，未來電價策略應走向更多元更彈性，除了符合市場需求外，也為更能達成負載管理目標而設計。

第六章、參考文獻

中文參考文獻

1. 王派洲(譯)(民97)。資料探勘 - 概念與方法(原作者: Jiawei Han, Micheline Kamber)。滄海。(原著出版年: 2006)
2. 台灣電力公司(民99)。99年台灣電力公司顧客服務白皮書。
3. 台灣電力公司(民101)。台電月刊101年5月號。第59-60頁。
4. 台灣電力公司(民99)。台灣電力公司歷史發展與組織架構。101年6月, 取自 <http://www.taipower.com.tw>。
5. 台灣電力公司。台灣電力公司負載管理說明。101年6月, 取自 http://www.taipower.com.tw/big/Knowledge_world/FAQ/part6.htm。
6. 吳旭智、賴淑貞(譯)(民90)。資料採礦理論與實務 - 顧客關係管理的技巧與科學。(原作者: Michael J. A. Berry, Gordon S. Linoff)。維科圖書有限公司。(原著出版年: 1997)
7. 郭志隆、張芳菱(譯)(民97)。資料探勘- Introduction to Business Data Mining。(原作者: David Louis Olson)。麥格羅·希爾。(原著出版年: 2007)
8. 范惟翔(民96)。現代行銷管理理論與實務。新加坡商湯姆生亞洲私人有限公司。第7-8頁。

9. 國家政策研究基金會 (民 98)。國政研究報告-因應地球暖化之台灣能源政策規劃建議。101 年 3 月，取自
<http://www.npf.org.tw/post/2/5952>。
10. 經濟部能源局 (民 99)。經濟部能源局 2010 年 12 月能源報導。
11. 經濟部能源局 (民 101)。經濟部能源局 2012 年 3 月能源統計月報。

英文參考文獻

1. Gary L. Lilien, Arvind Rangaswamy, “Marketing Engineering: Computer-Assisted Marketing Analysis Planning”, Addison Wesley, 1997, pp.57-58.
2. Jiawei Han, Micheline Kamber, “Data Mining: Concepts and Techniques”, Morgan Kaufmann Publishers, August 2000.
3. Kotler, Philip, “Marketing Insights from A to Z: 80 Concepts Every Manager Needs to Know”, John Wiley & Sons, Inc., 2003, pp.39.
4. Lach J., “Data mining digs in”, American Demographics, volume 21, number 7, July 1999, pp. 38-45.
5. Ralph Kimball, Joe Caserta , “The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning”, Wiley, 2004.
6. Robert C. Blattberg, Byung-Do Kim, Scott A. Neslin, “Database Marketing: Analyzing and Managing Customers”, Springer, 1 edition, 2009.
7. SAP , “Insight: The Critical Path to Competitive Differentiation for Growing Companies”, IDC executive brief, Jun 30, 2011.
8. Surajit Chaudhuri, Umeshwar Dayal, “An Overview of Data Warehouse and OLAP Technology”, SIGMOD Record, Vol. 26, No. 1, 1997, pp. 65-74.
9. William H. Inmon, “Building the Data Warehouse”, Wiley, 3 edition, 2002.
10. Yin, Robert K., “Case Study Research: Design and Methods”, Sage, 2002.