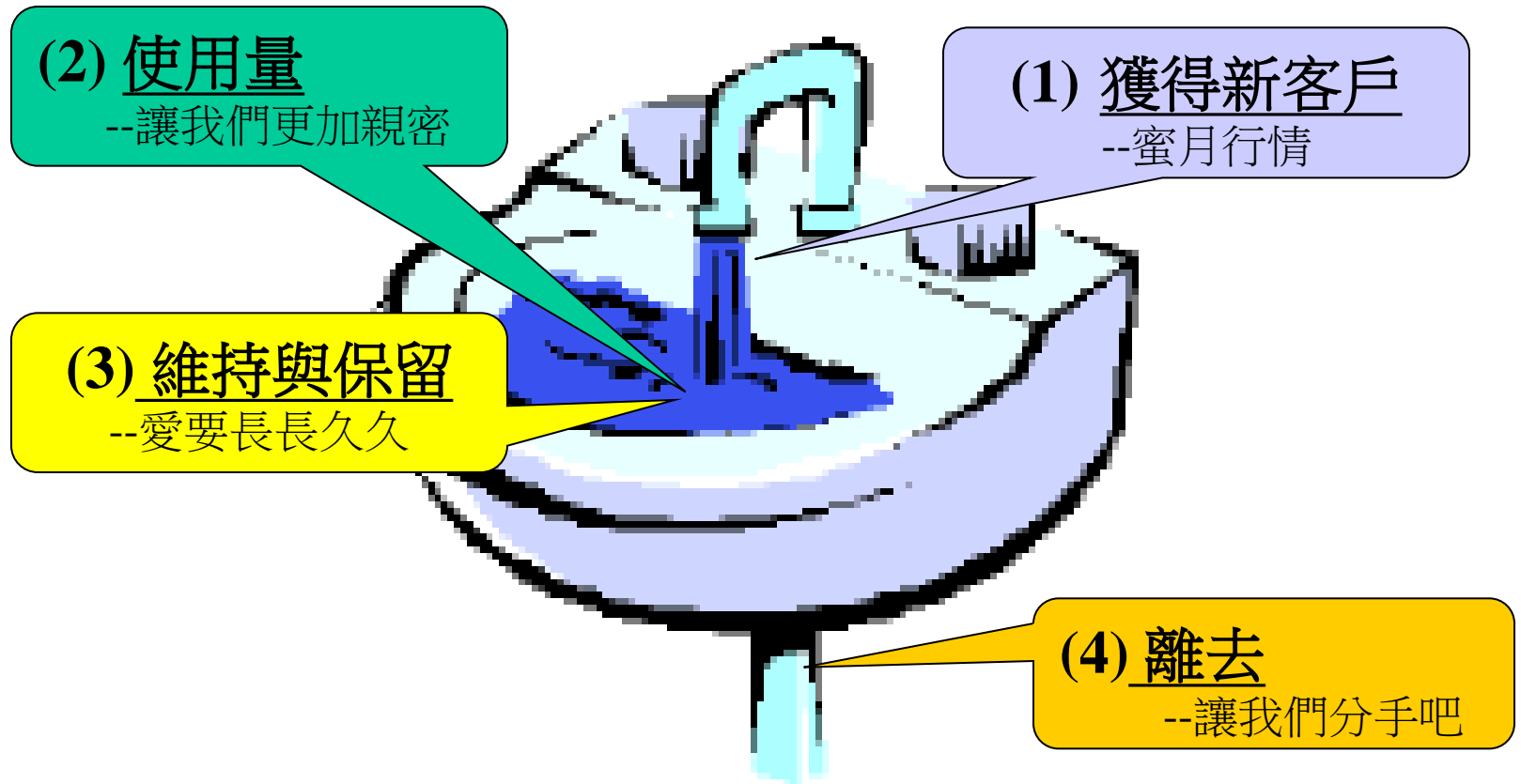


# 資料探勘在資料庫行銷上的應用

## Data Mining in Database Marketing



# 客戶的生命週期



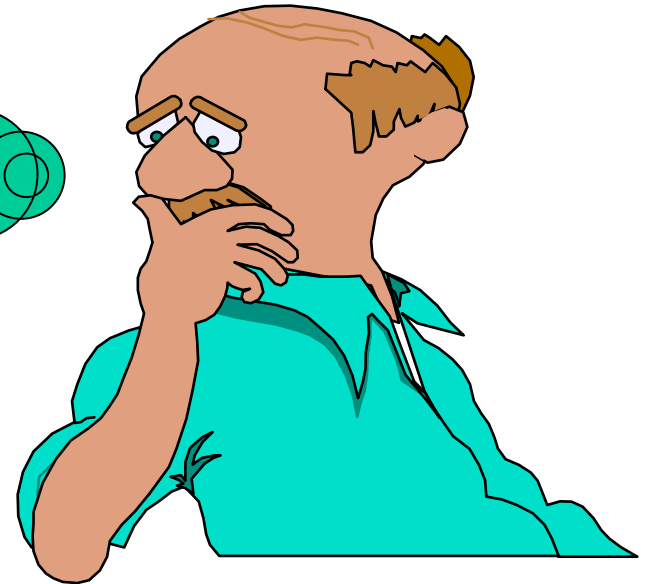
# 行銷經理人最想知道的事？

- 如何能抓住能讓我**獲利**的客戶？
- 什麼樣的**產品組合**客戶最喜歡？
- 誰最有可能對我的**廣告**郵件產生**反應**？
- 購買某**產品**的客戶是否有特定的**族群**？
- 某群消費者**未來的購買型態**為何？
- 誰最有可能產生**呆帳**？
- ...

# 以前的區隔方法-行銷經理

1. 30~40歲之間
2. 已婚
3. 有7歲以下的小孩
4. 住在都會區
5. 大專(或以上)之教育程度

行銷經理

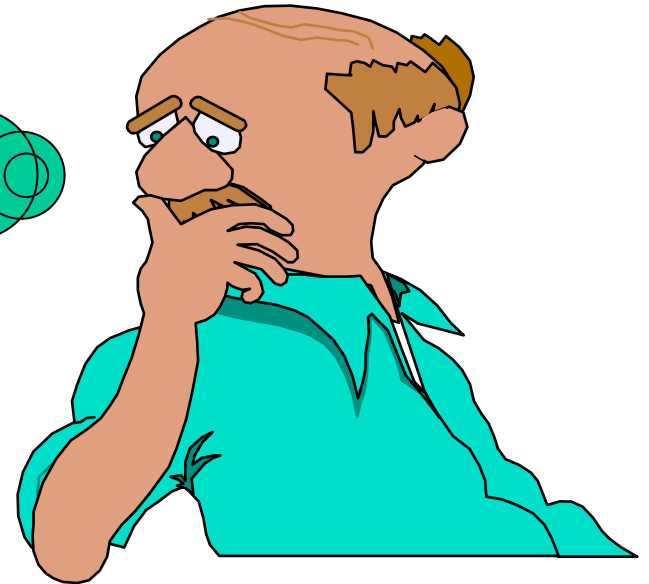


分類方式太粗略!

# 以前的區隔方法-行銷經理

1. Age 30~40
2. Married
3. Have 2 children
4. Live in metropolitan
5. College (or above) education degree

**Marketing Manager**



**Segmentation by intuition is  
too rough and effective!!**

# 她們的行為一樣嗎？

## 上班族

經常商務旅行  
網路與電腦的重度使用者  
雙薪家庭  
投資組合：基金、定存  
一個月去一次媚登峰



## 少奶奶

先生為自營商  
投資組合：股票、房地產  
三個月出國旅行一次  
SPA：一週三次

# 現今的區隔方法-資料庫行銷

- 資料庫行銷 (Database Marketing) 的目的就是希望藉由紀錄客戶的交易資料 (Transaction Database) 與客戶的基本資料 (Personal Database)，並加以分析，以了解客戶的消費習慣 (Consumer Behavior)，確定行銷目標，進而有效掌握一對一行銷 (One-to-One Marketing) 的關鍵
- 傳統的資料庫行銷分析中，利用三個指標 RFM對客戶分群

# RFM指標

- Recency (R)
  - 最近一次購買日期
- Frequency (F)
  - 購買頻率
- Monetary (M)
  - 消費金額
- R & F
  - 評估客戶忠誠度之指標
- M
  - 評估客戶利益高低之指標



# RFM指標

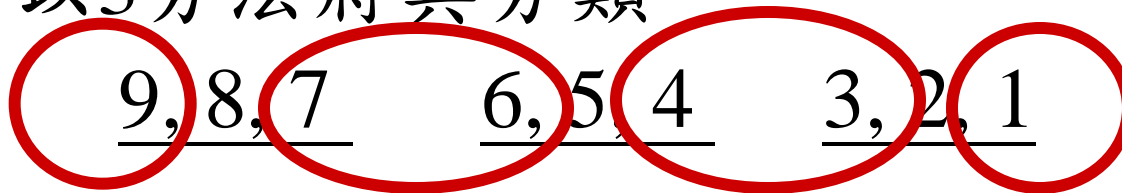
- RFM的方式是依**結果**而非**原因**
  - 預測未來會有相當大的誤差
- 難以將**新顧客**予以適當分類
- 如何分類、管理？
  - 5 分法
    - $5(F) * 5(R) * 5(M) = 125$  群組

# Frequency的考量

假設你有9位顧客，其購買次數如下：

9, 8, 7, 6, 5, 4, 3, 2, 1 (**Frequency**)

以3分法將其分類



其差異有真有  
如此顯著嗎？

# F & R的考量

## 時間序列的問題?

Customer B						
Jan	Feb	Mar	Apr	May	Jun	total
6	5	4	3	2	1	21

Customer D						
Jan	Feb	Mar	Apr	May	Jun	total
5	3	4	1	2	6	21

Customer A						
Jan	Feb	Mar	Apr	May	Jun	total
1	2	3	4	5	6	21

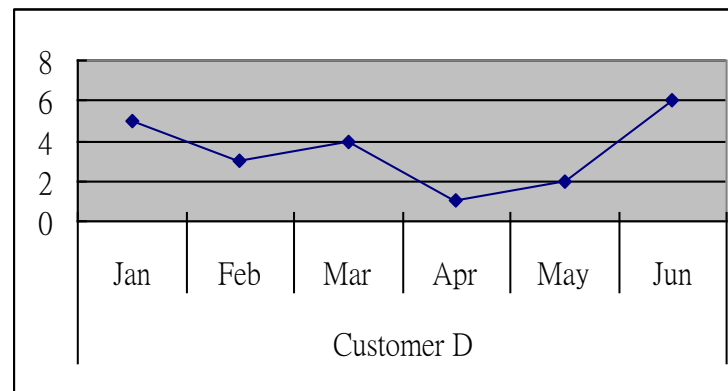
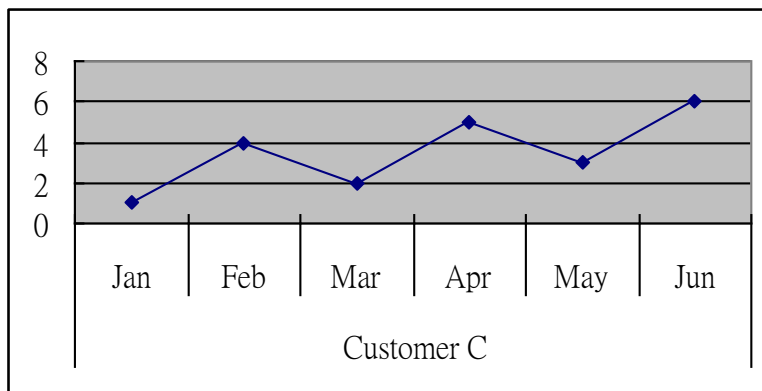
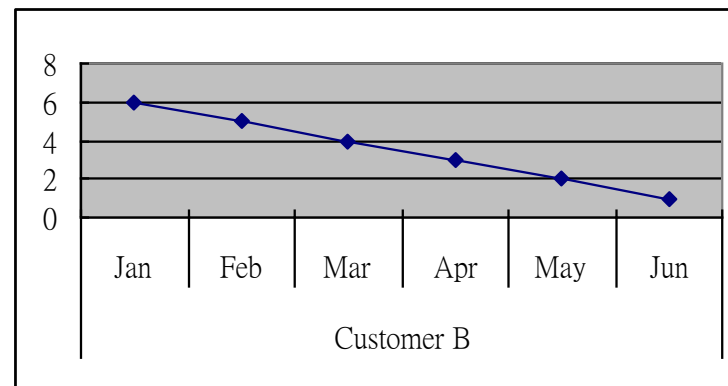
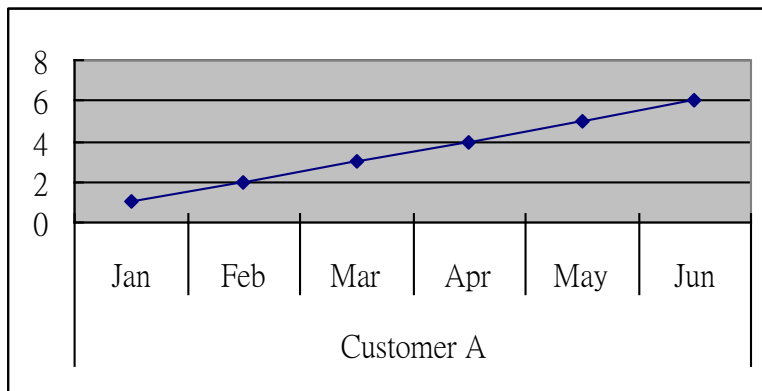
Customer C						
Jan	Feb	Mar	Apr	May	Jun	total
1	4	2	5	3	6	21

F

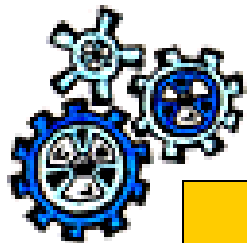
R

積分相同,但在時間上的趨勢卻沒考慮到!

# F & R的考量



這四個顧客的行為不同!



什麼是資料探勘？

# 什麼是資料探勘？

- 資料探勘是電腦應用領域的新名詞。然而當人類還在茹毛飲血的上古時代早已進行著資料探勘的行為
  - － 為了快速並準確捕獲獵物，人類的祖先必須細心觀察獵物的習性、預測獵物的行為，才能戰勝獵物、存活下去
- 同樣的，在講究即時、競爭激烈的網路時代，若能事先破解消費者的行為模式，將會是贏得電子商務的關鍵因素之一

# 什麼是資料探勘？

- 資料探勘
  - Data Mining
- 定義
  - William Frawley & Gregory Piatetsky Shapiro, 1991
  - 從現有的大量資料中，擷取不明顯的、之前未知的、可能有用的資訊
- 目標
  - 建立起決策模型
    - 哪一類的使用者對我的產品有興趣？
  - 根據過去的行動來預測未來的行為



# 什麼是資料探勘？

## 大量的資料

信用卡消費資料: 假設每人平均有1.5張信用卡, 每月平均消費10筆, 該銀行約有150萬的客戶, 就資料量而言, 每月約有2250萬筆消費記錄, 每年約有2億7千萬筆消費記錄



裏面有礦!

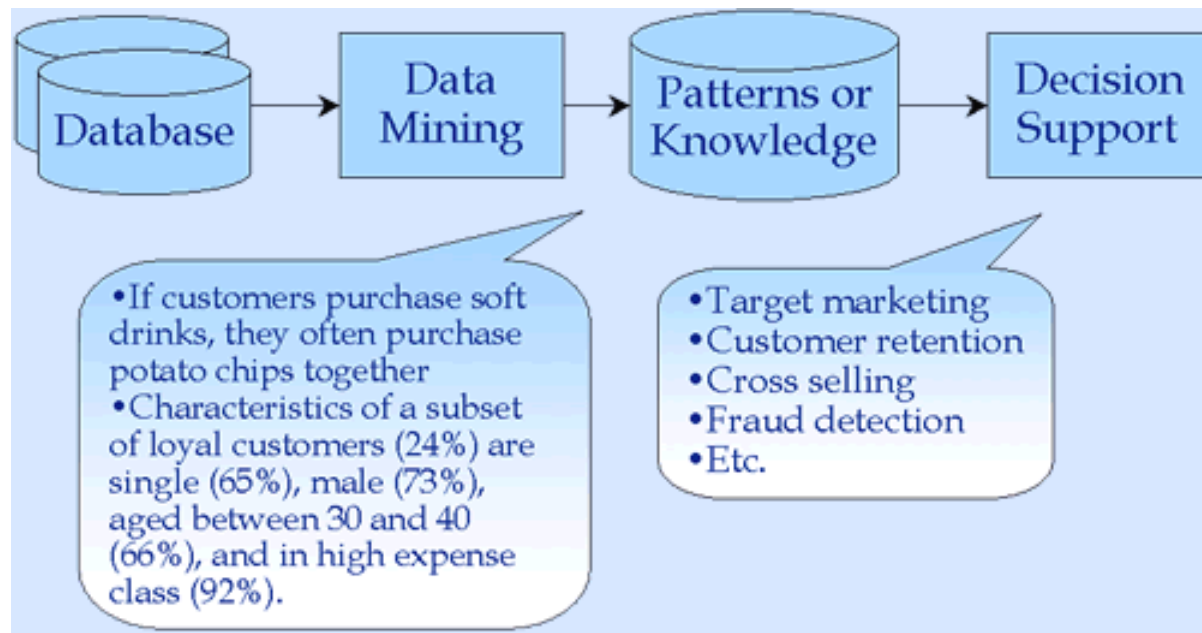


## 型態或規則

客戶的消費型態或規則



# 什麼是資料探勘？



# 資料探勘的發展歷程

- 1960s

- 資料蒐集

- 磁帶、軟碟、  
硬碟、...

- 1980s

- 資料查詢

- 資料庫  
(SQL語言)

- 1990s

- 資料統計

- 資料倉儲  
(OLAP)

- 2000s

- 資料分析

- 資料探勘

**所有的公司都需要資料探勘？**

**（從公司的規模來看）**

# 小型公司

- 靠著對客戶的了解，使客戶建立忠誠度
  - 買酒
    - 葡萄酒桶
      - 提供買酒的建議是基於對我的了解及對酒的豐富知識
      - 一家家打折的烈酒專賣店可以聘請一群專業的品酒師，但他們需要花上數月的時間才能像葡萄酒桶那麼懂我
  - 理髮
- 小型公司靠著留心客戶的需要，記住客戶的喜好，從過去的互動中擷取經驗以便在未來做得更好

# 大型公司

- 大型公司的職員多數無法和客戶有直接親身的互動
  - 單一客戶可能每次所面對的都是不同且不知名的客服人員
  - 這些客服人員如何從中留心、記住並學習
- 小型公司靠著卓越的直覺能夠立刻認出客戶的姓名、臉孔、聲音以及喜好
- 在這種情形下，大型公司要用什麼東西取代這些能力？
  - No!

# 大型公司

- 大型公司需要資料探勘
- 資料探勘可能是唯一的選擇
- 大型公司也有實行資料探勘的環境及條件
  - 大部分的作業均已自動化
  - 利用這些自動化所產生的資料，透過資料探勘，我們可以從這些經年累月的大量客戶資料中，分析客戶的族群，客戶的興趣，找出消費者的行為模式，了解哪些類型的客戶需要哪些方面的服務，並進一步預測消費者的行為
- 資料探勘可能是了解客戶的唯一機會

# 客戶、服務、獲利

- 當客戶再也得不到那種貼心的服務時，最直接的效應就是客戶開始流失
- 大型公司經常投下數億元卻也無法有效解決客戶流失的問題
- 大型公司應停下來思考獲利與客戶之間的關係(服務)，提出有效的解決之道(資料探勘)，並將最有價值的客戶篩選出來，以有效掌握一對一行銷的關鍵
  - 典型企業中80%的利潤是由20%的客戶所創造出來的

# 資料倉儲與資料探勘 的關係



# 資料倉儲

- 資料探勘和資料倉儲(Data Warehousing)
  - 要了解客戶先要從不同來源資料的收集及有效的組織開始，此一步驟稱之為資料倉儲
  - 接下來資料需要分析、了解，進而轉換為行銷的訊息，這就是資料探勘
- 資料倉儲可以說是資料探勘的前置步驟
- 資料倉儲可以說是線上分析處理(OLAP)的前置步驟
- 線上分析處理
  - 簡單的資料探勘

# 資料倉儲

- 資料倉儲提供企業所需的記憶，但記憶對智慧是沒有很大的幫助
- 資料探勘的目的就是要經由資料倉儲產生商業智慧進而獲利

# 資料探勘的進行方式

# 資料探勘的進行方式

- 資料探勘的方法是經由自動或半自動的方法探勘及分析大量的資料，以建立有效模式(Pattern)及規則(Rule)
- 資料探勘的進行方式
  - 由上而下(Top-Down)
    - 稱為假說檢定
  - 由下而上(Button-UP)
    - 稱為知識發掘(Knowledge Discovery)
    - 知識發掘不需要事先的假設，而是直接讓資料說話

# 資料探勘的應用

# 資料探勘的應用

- 美國聯邦政府利用資料探勘來找出罪犯
  - 從數以千計的相關報導中，篩選可能的關聯性以及可能的頭緒
  - 是最早使用資料探勘的單位之一
  - 是屬於文件探勘的應用之一
- 商品擺設
  - 美國安道超市 (Safeway Corporation) 紀錄每一筆顧客的交易，他們藉此找出哪些商品應放在一起賣，並找出容易對特定商品有回應的消費者

# 資料探勘的應用

- 客戶招攬
  - Firefly(螢火蟲)公司要求他的會員對音樂及電影評分，透過評分所顯示的好惡程度自動將喜好相似的人分群，並建立虛擬社群 (Virtual Community)。在將每個會員分派至對應的社群後，自動計算出這個社群大多數人評分最高的音樂及電影，並推薦給這社群的成員
- 交叉銷售(Cross Sale)
  - USAA針對軍人及眷屬開設保險公司，它可透過資料探勘來預測客戶的生命週期中，何時需要哪些產品？

# 資料探勘的應用

- 客戶挽留

- 開發一個新客戶比維持一個舊客戶所花的成本更高
- 美國南加州天然氣公司利用資料探勘後發現哪些類型的客戶對價錢較敏感？哪些類型的客戶對贈品較有興趣？
  - 不要根據客戶花在公司消費額來決定要給客戶的服務
  - 有些優良的客戶即使沒有禮物也會繼續使用公司的產品
  - 注重流行的客戶可能就需要禮物來留住他

- 客戶區隔

- 信用卡銀行可利用資料探勘去分析哪些客戶會對哪些企劃案有興趣？



# 資料探勘的應用

- 客戶流失

- 有些公司流失的客戶遠比新加入的客戶還多，而新客戶所帶來的利潤卻不如以往
- 為彌補客戶流失所帶來的損失，而暫停一些不賺錢的服務，並不一定是個好的策略
  - 例如某家銀行決定停止部分賠錢的轉帳繳費服務，然而根據研究指出大部分高價值及忠誠度高的客戶是使用該服務的族群，冒然停止這項服務可能會導致這群客戶的流失
- 為避免客戶流失而提供更多服務，有可能只保住了一小部分的客戶，而讓公司花了太多的錢
- 公司可利用資料探勘針對已經離去的客戶，找出可能的離去原因，以發動預防措施

- 風險評估及詐欺察覺

- 銀行可利用資料探勘去分析公司或個人的財務狀況，以作為是否貸款給他們的參考
- 信用卡公司可利用資料探勘去分析信用卡在哪些情況下會被盜刷？

# 統一超商砸四十億找出 氣溫和銷售關係



統一超商砸下一年的獲利打造情報高速公路

# 七五三感冒指數

- 根據統計，氣溫差三度，某些商品的銷售就可能相差一倍以上
- 在這個氣象經濟的時代，掌握溫度變化就等於掌握了商機
- 日本零售商對於「七五三感冒指數」就很敏感，如果一天當中溫度相差七度、今天和昨天的溫度差到五度、且溼度差大於三〇%的話，代表感冒的人會增加，商家就要考慮把感冒藥、溫度計和口罩之類的用品上架

# 溫度和暢銷商品

- 除了感冒指數，溫度和暢銷商品之間也有密不可分的相關性
  - 氣溫在二十四度到二十七度之間時
    - 鰻魚、冰品和防曬乳會賣得好
  - 溫度在二十二度到二十五度時
    - 涼飲、冰咖啡和殺蟲劑就不可少
  - 溫度在十七度到二十度當中時
    - 布丁、沙拉和優格則很受歡迎
- 統一超商雖提供門市詳細的氣溫與商品銷售資訊做為店長訂貨的依據，但是也相對要求各門市對營運成績負更大的責任

# Amazon.com

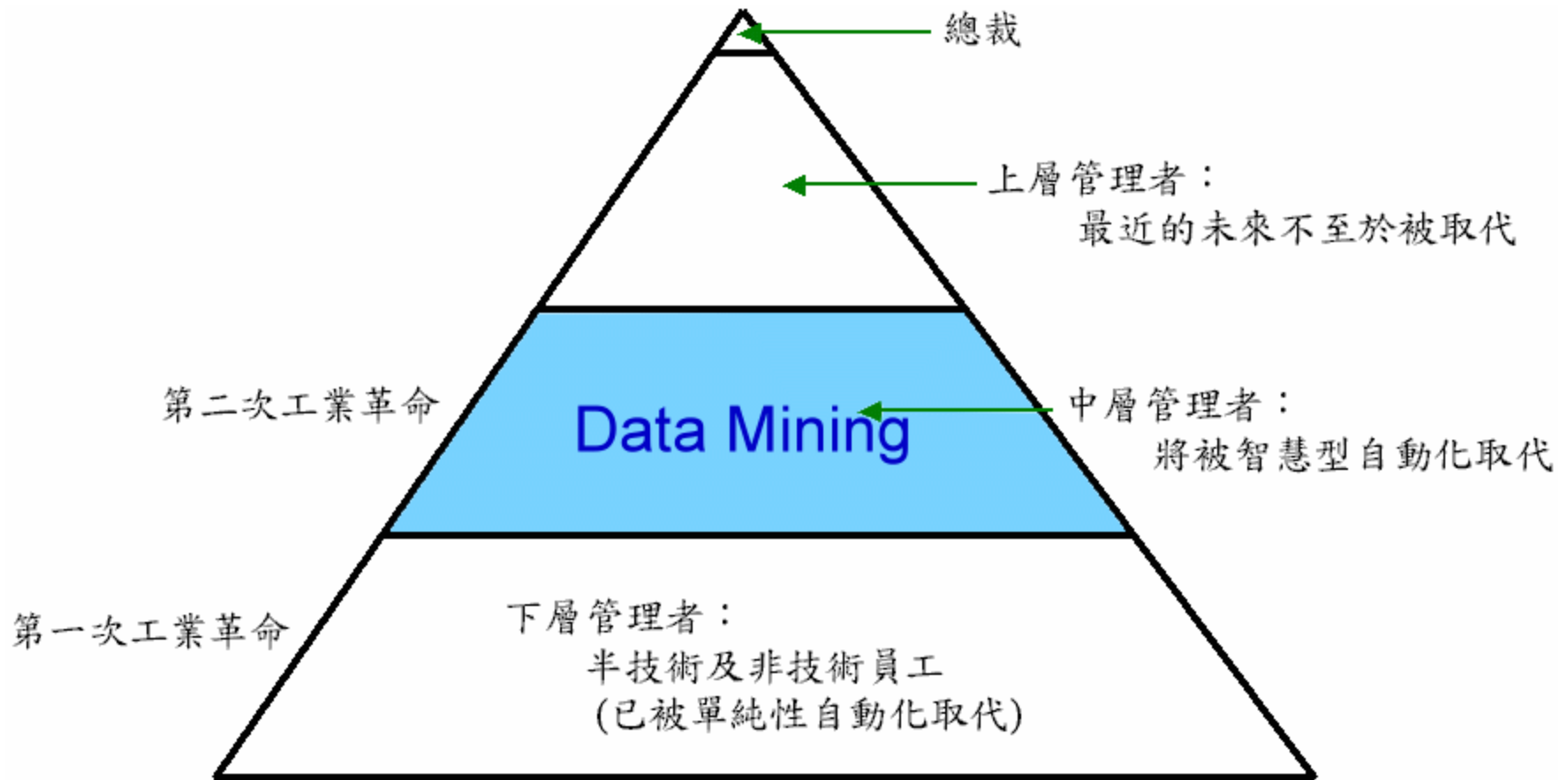
- 全球最大的網路書店亞馬遜 (Amazon.com) 也有一套強大的顧客管理系統 (CRM)，能夠迅速處理每日高達二十萬人次流量的龐大資訊
- 當顧客在亞馬遜網站買東西之後，銷售系統就會紀錄下顧客買了什麼、甚至還在哪些商品上面做過停留
- 當下次這位顧客再次進入亞馬遜書店時，系統會主動識別出顧客身分、並根據顧客的喜好推薦商品，這顯然對維持顧客忠誠度有很大幫助
  - CRM 讓亞馬遜擁有高達六五%的顧客回流率

# WAL-MART

- 世界大零售商威名百貨 (WAL-MART) 也很善用「購物籃分析」(Market Basket Analysis)，讓毫不起眼的資訊變黃金
- 威名百貨在一次分析客戶資料庫時發現，跟尿布一起購買最多的商品竟然是啤酒
  - 後來透過市場調查才得知，原來美國的太太常叮嚀丈夫在下班前幫嬰兒買尿布，而四〇%的先生在買完尿布後，又會隨手拎兩罐啤酒
  - 既然尿布和啤酒一起購買的情形特別多，威名百貨就在各個賣店把啤酒和尿布擺在一起，結果尿布和啤酒的銷售量雙雙增加三成！

# 資料探勘的定位及展望

# 資料探勘的定位





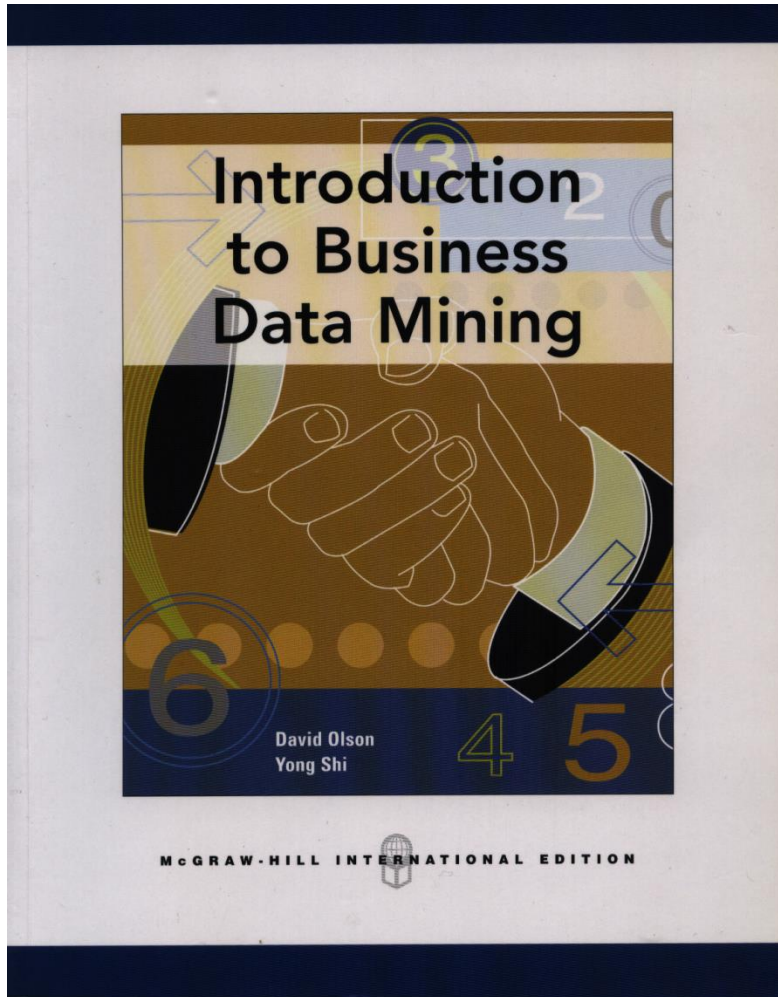
# 資料探勘的展望

- 美國時代雜誌將資料探勘列為21世紀五大新興行業之一，可看出資料探勘未來在商業界的重要性
  - 幫助企業導入CRM的公司越來越多
- 未來的行銷重心將由產品轉為客戶
  - 你的客戶可能不再是你的客戶，因為不知道何時競爭對手會提供更好的服務，把你的客戶搶走。
  - 誰掌握了最多有關客戶的知識，誰就擁有最多的資本
  - 對客戶了解越多，就越能深化品牌的獨特性，競爭也越強
- 只有透過將資料轉為知識，將知識轉為行動，才能徹底將行動轉為極大的利潤

# DM vs. Fortune-Telling

- Data Mining
  - Data: Attributes
  - Algorithm: Concept Description, Association, ...
  - Predict Future Trends
- Fortune-Telling
  - Data: 生辰八字, 面相, 手相, ...
  - Algorithm: 紫微斗數, 四柱推命, 鐵板神算, ...
  - Predict Future Fortune

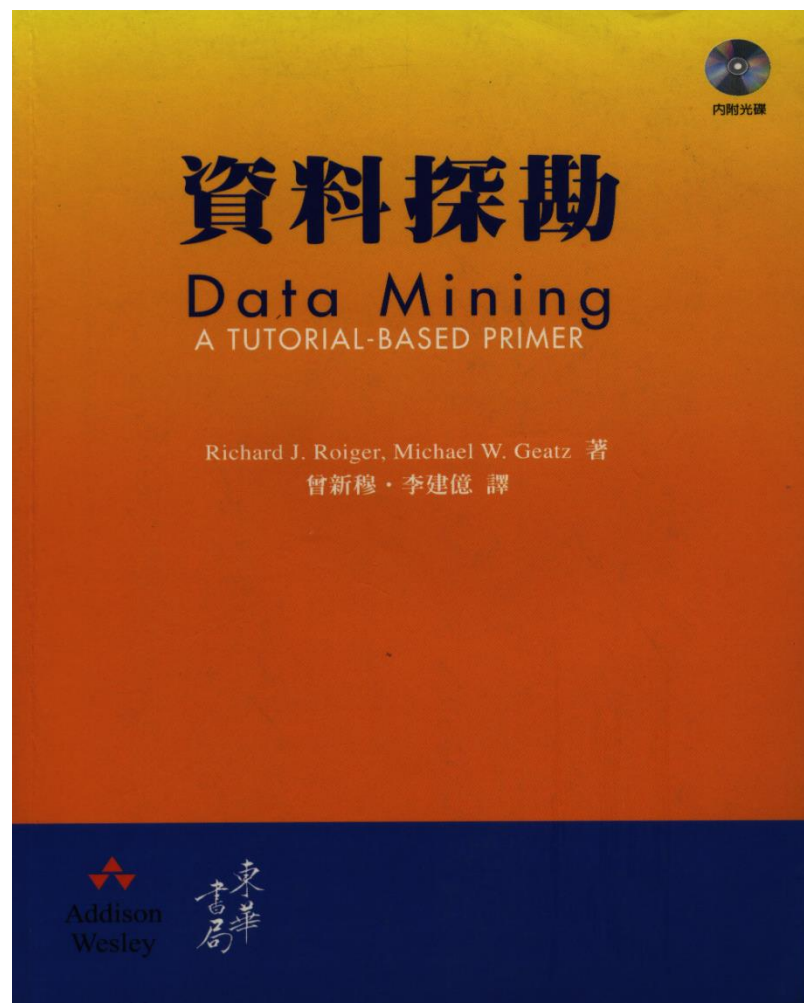
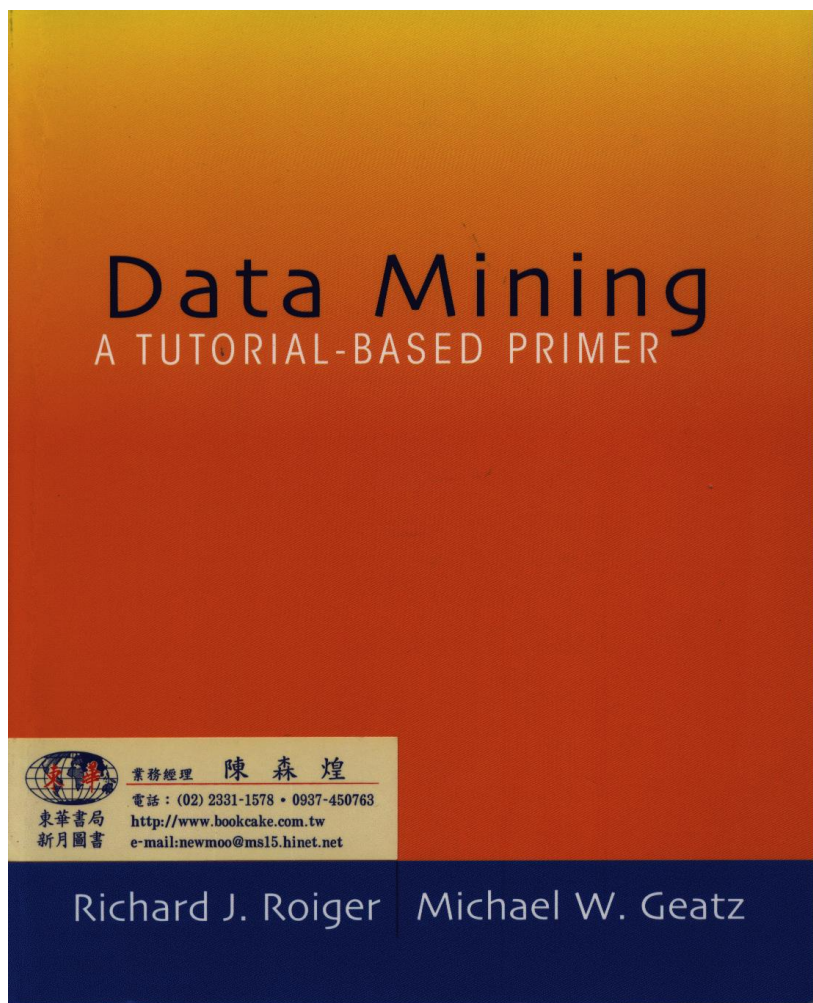
# 教科書



**Introduction to Business Data  
Mining by David Olson & Yong  
Shi, McGRAW-HILL, 2007**

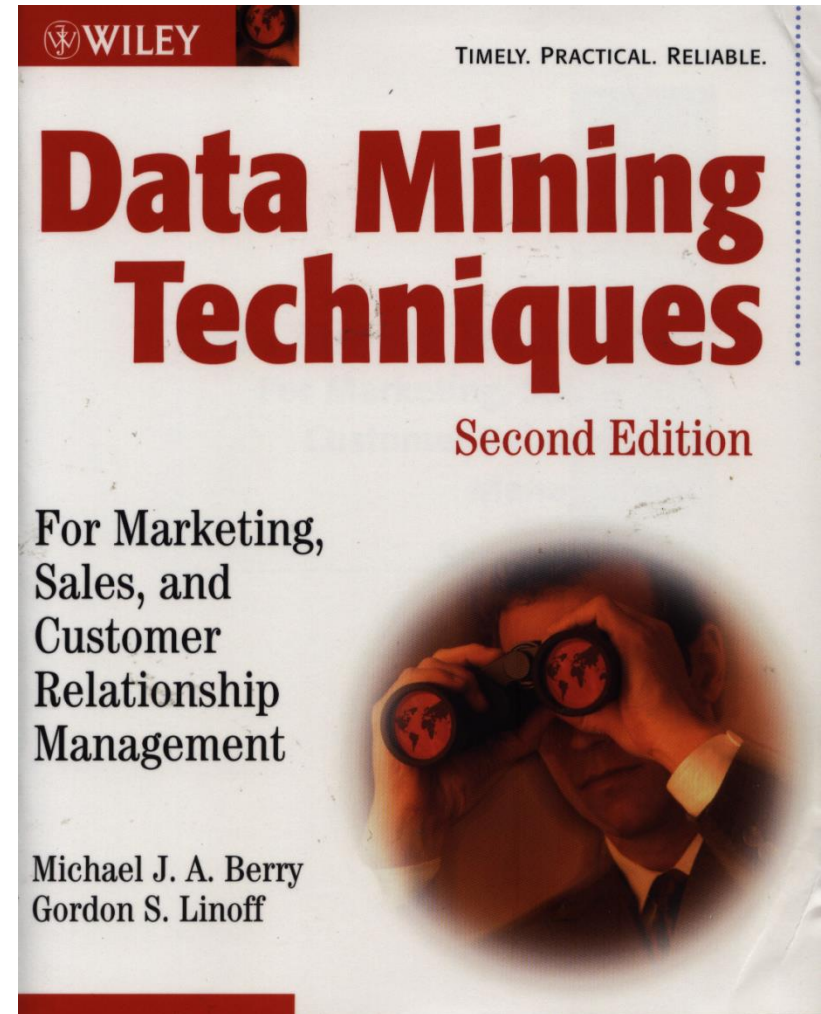
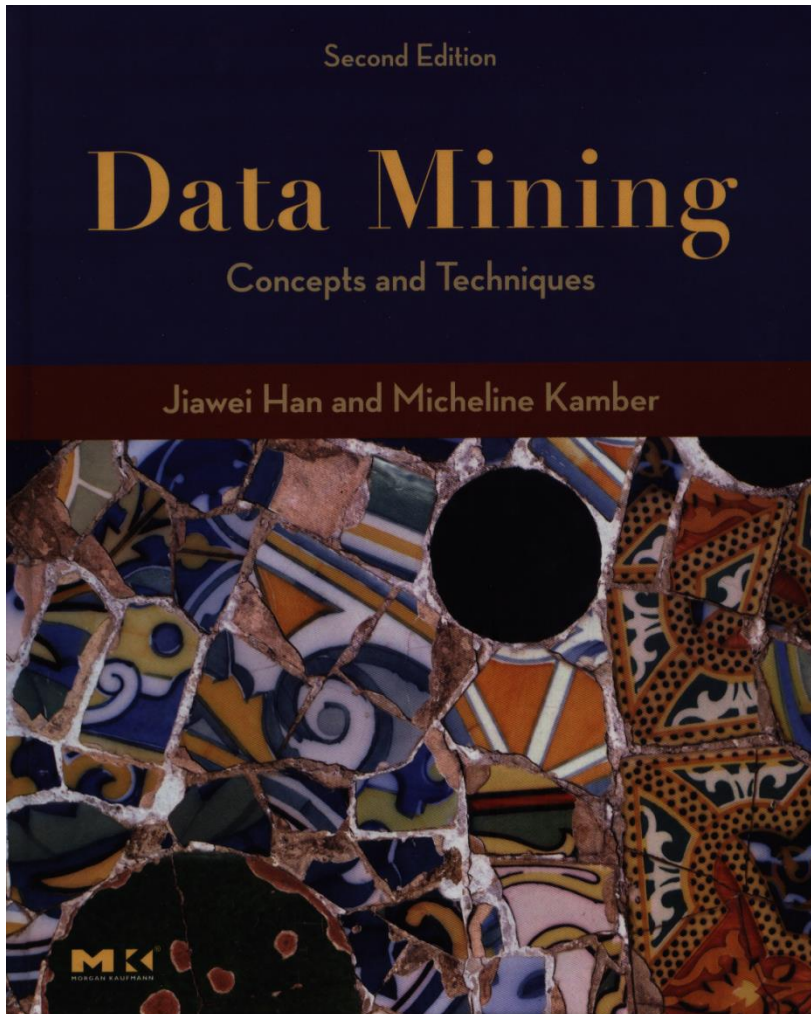
**電話：0912933790 康代育**

# 參考書





# 參考書



# A Brief History of DM Society

- 1989 IJCAI Workshop on Knowledge Discovery in Databases (Piatetsky-Shapiro)
  - Knowledge Discovery in Databases (G. Piatetsky-Shapiro and W. Frawley, 1991)
- 1991-1994 Workshops on Knowledge Discovery in Databases
  - Advances in Knowledge Discovery and Data Mining (U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy, 1996)
- 1995-1998 International Conferences on Knowledge Discovery in Databases and Data Mining (KDD'95-98)
  - Journal of Data Mining and Knowledge Discovery (1997)
- 1998 ACM SIGKDD, SIGKDD'1999-2001 conferences, and SIGKDD Explorations
- More conferences on data mining
  - PAKDD, PKDD, DaWaK, (IEEE) ICDM, etc.

# Papers for Data Mining

- Proceedings
  1. Proceedings of the International Conference on Data Mining (ICDM)
  2. Proceedings of the International Conference on Data Engineering (ICDE)
  3. Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining
  4. Proceedings of the International Conference on Very Large Data Bases (VLDB)
  5. Proceedings of ACM SIGMOD International Conference on Management of Data
  6. Proceedings of the International Conference on Database Systems for Advanced Applications (DASFAA)
  7. Proceedings of the International Conference on Database and Expert Systems Applications (DEXA)

8. Proceedings of the International Conference on Data Warehousing and Knowledge Discovery (DaWak)
9. Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD)
10. European Conference on Principles of Data Mining and Knowledge Discovery (PKDD)

- **Journals**

1. IEEE Transactions on Knowledge and Data Engineering (TKDE)
2. Journal of Intelligent Information Systems
3. Data Mining and Knowledge Discovery
4. ACM SIGMOD Record
5. The International Journal on Very Large Database
6. Knowledge and Information Systems
7. Data & Knowledge Engineering
8. International Journal of Cooperative Information Systems