

國立交通大學  
工業工程與管理學系

碩士論文

透過資料倉儲挖掘

台灣颱風與大宗花卉變化的關聯性法則

Data Warehouse Approach to Mining Association Rules  
between Typhoons and Main Flowers in Taiwan



研究生：陳昶年

指導教授：梁高榮博士

中華民國九十六年七月

透過資料倉儲挖掘

台灣颱風與大宗花卉變化的關聯性法則

Data Warehouse Approach to Mining Association Rules between Typhoons  
and Main Flowers in Taiwan

研 究 生：陳昶年

Student：Chang-Nian Chen

指導教授：梁高榮

Advisor：Gau-Rong Liang



碩 士 論 文

A Thesis Submitted to  
Department of Industrial Engineering and Management  
College of Management  
National Chiao Tung University  
In Partial Fulfillment of the Requirements  
For the Degree of Master of Engineering in  
Industrial Engineering and Management  
June 2007  
Hsinchu, Taiwan, Republic of China

研究生：陳昶年

指導教授：梁高榮博士

國立交通大學工業工程與管理學系

## 摘要

本文提出使用資料倉儲於台灣颱風與大宗花卉之間的關聯性法則探勘。每年颱風導致台灣的花卉產業重大的損失。然而，在花卉方面的損失與颱風之間的關聯性並不是如此明顯。主要的原因是為了分析這層關聯背後所要準確蒐集花卉資料的困難度，因此包括了許多高成本的風險，例如如何從不同的來源去蒐集花卉資料；如何判斷正確的花卉資料；如何將有用的花卉資料正確分類使其能夠進行統計分析...等。這份論文中採用資料倉儲系統解決上述艱難問題。換句話說，倉儲中的花卉資料來源於五個作為當地花卉批發市場交易資料目的地的資料超市。因此，資料蒐集的過程僅僅只需一個線上分析處理資料下載程序。接著運用 Apriori 演算法於台灣颱風與大宗花卉之間的關聯性法則探勘。並發現有 58 條種關聯性法則在其中。

關鍵詞：

颱風特性(The Property of Typhoon)

大宗花卉(Main Flowers)

資料倉儲(Data Warehouse)

關聯性法則(Association Rule, AR)

# Data Warehouse Approach to Mining Association Rules between Typhoons and Main Flowers in Taiwan

Student : Chang-Nian Chen

Advisor : Dr.Gau-Rong Liang

Department of Institute of Industrial Engineering & Management  
National Chiao Tung University

## Abstract

Data warehouse has been proposed for mining association rules between typhoons and main flowers in Taiwan. Typhoons bring huge damages in flower industry in Taiwan every year. However, the relation between flower damages and typhoons is not so clear. One major reason is the difficulty of collecting the right flower data for analyzing the relation because it includes many high cost tasks such as how to collect the flower data from the various sources, how to justify the correct flower data, how to sort out the valid flower data into the right category for statistical use, etc. In this thesis, a data warehouse was implemented for solving this hard challenge. In other words, the flower data in the data warehouse come from the five data marts which are distributed data achieves for storing transactional data in local flower wholesale markets. As a result, the data collection process becomes as a simple On-Line Analytical Processing (OLAP) data download procedure. Then Apriori algorithm was applied to mining the association rules between typhoons and major flowers. Also 58 association rules are found.

Keywords :

The Property of Typhoon

Main Flowers

Data Warehouse

Association Rule

## 誌謝

能夠順利完成此篇論文，首先要感謝恩師 梁高榮老師在我就讀研究所兩年的求學生涯中，不辭辛勞與諄諄教誨，特別是在論文指導期間中，時常給予我多方見解與不同方向的思維模式，讓我的思慮能夠更加慎密。而亦要感謝唐麗英老師、張永佳老師之不吝指導且對於論文審閱仔細並提供不少寶貴意見，以豐富本篇論文。

其次要感謝粗皮學弟、同窗好友端、計畫領導者新凱隊長在我寫論文的這段期間給予我實質上的幫助，讓我的論文能夠穩定下來，對於好友王天才、TOBE、MIYA學妹、學弟的感謝千言萬語也說不盡。

最後，我要感謝爺爺、奶奶與父母親對我的肯定與支持，讓我能夠專心一意的完成所有的學習。還有正就讀弘光的弟弟，謝謝你能夠在我最沮喪的時候，給予我鼓勵與打氣，讓我能夠再度回到跑道上，感謝你們在我身心憔悴時的體諒與容忍，讓我在兩年的研究所生涯中能全心全意的衝刺，謝謝您們的關心。

本論文在撰寫期間曾參與農糧署的研究計畫「建置高雄花卉批發市場資料倉儲與大型冷藏庫庫存管理系統」，計畫編號為「95I510」，接受支助，特此感謝。



# 目錄

摘要.....	i
Abstract.....	ii
誌謝.....	iii
目錄.....	iv
圖目錄.....	vii
表目錄.....	ix
第一章 緒論.....	1
1.1 研究動機.....	1
1.2 問題界定.....	3
1.2.1 整合 DM 為 DW.....	3
1.2.2 探勘颱風與大宗花卉資料.....	4
1.2.3 大宗花卉為研究範圍.....	5
1.3 研究目的.....	6
1.4 研究流程.....	7
1.5 論文架構.....	9
第二章 文獻回顧.....	10
2.1 花卉產業介紹.....	10
2.1.1 花卉切花種類之產地.....	10
2.1.2 大宗花卉.....	11
2.1.3 花卉種類與節慶.....	12
2.2 颱風特性分析.....	14
2.2.1 颱風路徑與降雨量.....	14
2.2.2 颱風強度與近中心最大風速.....	15
2.2.3 颱風移動速度.....	16
2.2.4 颱風強度與近中心最低氣壓.....	16
2.2.5 小結.....	16
2.3 資料倉儲.....	18
2.3.1 資料倉儲的建構模式.....	18
2.3.2 線上分析處理.....	21
2.4 資料探勘之關聯性法則.....	23
2.4.1 關聯性法則資料探勘技術.....	23
2.4.1.1 關聯性法則 Apriori 演算法.....	23
2.4.1.2 關聯性法則 DLG 演算法.....	25
2.4.1.3 關聯性法則 RARM 演算法.....	26
2.4.1.4 關聯性法則 DHP 演算法.....	27
2.4.1.5 關聯性法則 VIPER 演算法.....	28
2.4.2 關聯性法則演算法比較.....	28
第三章 台灣區花卉資料倉儲資料整合實作.....	30

3.1 系統架構與資料倉儲架構 .....	30
3.1.1 系統架構 .....	30
3.1.2 資料倉儲架構 .....	35
3.2 總資料倉儲整合工作 .....	37
3.2.1 總資料倉儲架構 .....	37
3.2.2 資料整合工作 .....	37
第四章 台灣區花卉資料倉儲資料轉換實作 .....	40
4.1 資料轉換服務 .....	40
4.1.1 實體關係模式至多維度模型轉換 .....	41
4.1.2 資料轉換服務 .....	43
4.2 自動化排程設計 .....	49
4.2.1 DOS 程式撰寫 .....	49
4.2.2 總資料倉儲自動化排程設計 .....	50
第五章 關聯性法則對大宗花卉產量之颱風特性實證分析 .....	52
5.1 研究架構與假設 .....	52
5.1.1 研究架構 .....	53
5.1.2 研究前提 .....	53
5.2 實證模型建立 .....	56
5.2.1 關聯性法則 .....	56
5.2.1.1 相關定義與假設 .....	56
5.2.1.2 主階段關聯性法則 .....	60
5.3 計算過程說明 .....	63
5.4 結果說明 .....	64
第六章 結論與建議 .....	72
6.1 研究結論 .....	72
6.2 未來研究方向 .....	74
參考文獻 .....	75
附錄 A：1996-2005 年侵台颱風一覽表 .....	77
附錄 B：1996-2005 年侵台颱風對應節慶分析表 .....	79
附錄 C：1996-2005 年侵台颱風數據轉換表 .....	81
附錄 D-1：1998-2005 年大宗花卉-玫瑰數據轉換表 .....	83
附錄 D-2：1998-2005 年大宗花卉-大菊數據轉換表 .....	84
附錄 D-3：1998-2005 年大宗花卉-小菊數據轉換表 .....	86
附錄 D-4：1998-2005 年大宗花卉-火鶴花數據轉換表 .....	88
附錄 D-5：1998-2005 年大宗花卉-香水百合數據轉換表 .....	90
附錄 D-6：1998-2005 年大宗花卉-葵百合數據轉換表 .....	92
附錄 D-7：1998-2005 年大宗花卉-非洲菊數據轉換表 .....	94
附錄 D-8：1998-2005 年大宗花卉-洋桔梗數據轉換表 .....	96
附錄 D-9：1998-2005 年大宗花卉-康乃馨數據轉換表 .....	98
附錄 D-10：1998-2005 年大宗花卉-劍蘭數據轉換表 .....	100

附錄 E：演算法軟體分析與應用 .....	102
附錄 F：颱風特性間之 Apriori 演算法軟體分析結果 .....	105



## 圖目錄

圖 1.1 完整總資料倉儲系統架構圖 .....	3
圖 1.2 關聯性法則方法與界定 .....	5
圖 1.3 研究流程圖 .....	7
圖 2.1 國產切花產業分布圖 .....	10
圖 2.2 菊花每年所有產地總產量比較圖 .....	11
圖 2.3 菊花每年各產地產量比較圖 .....	11
圖 2.4 2003 年玫瑰花之總成交量變化趨勢圖 .....	12
圖 2.5 颱風路徑分類圖 .....	15
圖 2.6 星狀綱要圖 .....	19
圖 2.7 雪花綱要圖 .....	20
圖 2.8 星座綱要圖 .....	20
圖 2.9 APRIORI 演算法流程圖 .....	23
圖 2.10 APRIORI 演算執行過程 .....	24
圖 2.11 DLG 演算執行過程 .....	25
圖 2.12 TIRE 資料結構 .....	26
圖 2.13 RARM 演算執行過程 .....	26
圖 2.14 DHP 演算執行過程 .....	27
圖 3.1 總倉儲實體系統架構圖 .....	31
圖 3.2 SQL SERVER 連線群組(A) .....	33
圖 3.3 SQL SERVER 連線 4 群組(B) .....	33
圖 3.4 SQL SERVER 連線群組(C) .....	34
圖 3.5 網路磁碟機畫面 .....	34
圖 3.6 檔案複製到網路磁碟機語法畫面 .....	34
圖 3.7 資料轉換時刻圖 .....	35
圖 3.8 台灣區花卉資料倉儲競筭架構圖 .....	36
圖 3.9 花卉批發資訊分享熱線架構圖 .....	36
圖 3.10 供應類別對應表 .....	38
圖 3.11 供應鄉鎮對應表 .....	38
圖 3.12 供應地區對應表 .....	38
圖 4.1 自動化更新轉換程序 .....	41
圖 4.2 實體關係至多維度模型架構 .....	41
圖 4.3 高雄資料超市內部轉換關係圖 .....	42
圖 4.4 拍賣資料表與總倉儲交易資料表實體關聯圖 .....	42
圖 4.5 總資料倉儲雪花綱要關聯圖 .....	43
圖 4.6 拍賣資料轉換工作 .....	43
圖 4.7 高雄拍賣資料轉換工作 .....	44
圖 4.8 高雄拍賣轉換資料工作來源查詢 .....	44
圖 4.9 高雄拍賣轉換資料工作屬性 .....	44

圖 4.10 高雄拍賣轉換資料工作轉換語法 .....	45
圖 4.11 訂貨資料轉換工作 .....	46
圖 4.12 高雄訂貨資料轉換工作 .....	46
圖 4.13 殘貨資料轉換工作 .....	47
圖 4.14 高雄殘貨資料轉換工作 .....	47
圖 4.15 議價資料轉換工作 .....	47
圖 4.16 台北議價資料轉換工作 .....	47
圖 4.17 自動化 MACRO 程式 .....	49
圖 4.18 自動化步驟語法編輯 .....	51
圖 4.19 自動化排程設定 .....	51
圖 4.20 自動化排程重複執行設定 .....	51
圖 5.1 流程圖 .....	52
圖 5.2 雨量觀測站選擇分佈圖 .....	54
圖 5.3 季節性花卉之玫瑰產量分析圖 .....	55
圖 5.4 颱風數據整理圖 .....	61
圖 5.5 大宗花卉產地供應量概略圖 .....	61
圖 5.6 大宗花卉產拍賣量與成交量概略圖 .....	61
圖 5.7 颱風特性與大宗花卉轉換數據圖 .....	62
附錄圖 1 輸入檔案 CP.TXT .....	102
附錄圖 2 輸入檔案 BCC.TXT .....	103
附錄圖 3 輸入檔案 ITEMS_NAME.TXT .....	103
附錄圖 4 輸出檔案 FI.TXT .....	103
附錄圖 5 輸出檔案 AR.TXT .....	104



## 表目錄

表 1.1 資訊分享與建置年份示意表 .....	6
表 2.1 前十名之大宗花卉排序表(2005 年).....	12
表 2.2 國曆特殊節慶 .....	13
表 2.3 農曆特殊節慶 .....	13
表 2.4 具有生產季節性花卉主要產品 .....	13
表 2.5 強度與近中心最大風速對照轉換表 .....	15
表 2.6 颱風強度與近中心最低氣壓之對照轉換表 .....	16
表 2.7 颱風七大路徑對應之降雨量程度表 .....	17
表 2.8 颱風強度對應近中心最大風速與最低氣壓之對照轉換表 .....	17
表 2.9 資料倉儲之四大特性 .....	18
表 2.10 OLAP 基本十大操作說明表.....	21
表 2.12 APRIORI 演算步驟說明表 .....	24
表 2.13 DLG 演算步驟說明表 .....	25
表 2.14 DHP 演算步驟說明表 .....	27
表 2.14 關聯性法則之演算法優缺點比較表 .....	28
表 3.1 各主機項目命名表 .....	33
表 3.2 總倉儲主機負責服務整理表 .....	35
表 3.3 總倉儲與各家資料超市維度比較表 .....	37
表 3.4 整合對照表 .....	39
表 4.1 總倉儲資料轉換程式 .....	40
表 4.2 拍賣資料轉換工作步驟表 .....	45
表 4.3 訂貨資料轉換工作步驟表 .....	46
表 4.4 殘貨資料轉換工作步驟表 .....	47
表 4.5 議價資料轉換工作步驟表 .....	48
表 4.6 COGNOS 元件名稱整理表 .....	50
表 5.1 颱風對應各家花市休市與總成交量分佈表 .....	53
表 5.2 季節性花卉對應大宗花卉對應表 .....	55
表 5.3 類別定義與相關項目假設表(上) .....	56
表 5.4 玫瑰花類別定義與相關項目假設表(中) .....	58
表 5.5 年份與五大市場資料對應表 .....	58
表 5.6 玫瑰花類別定義與相關項目假設表(下) .....	59
表 5.7 有意義關聯性法則的判斷 .....	60
表 5.8 九大颱風特性 .....	60
表 5.9 大宗花卉兩大部分剖析 .....	62
表 5.10 關聯性法則(六).....	64
表 5.11 關聯性法則(六).....	65
表 5.12 關聯性法則(七).....	65
表 5.13 關聯性法則(八).....	66

表 5.14 關聯性法則(九).....	66
表 5.15 關聯性法則整合表 .....	67
表 5.16 關聯性法則花種對應變項分佈表 .....	70
附錄表 1 關聯性法則(一).....	105
附錄表 2 關聯性法則(二).....	105
附錄表 3 關聯性法則(三).....	106
附錄表 4 關聯性法則(四).....	106
附錄表 5 關聯性法則(五).....	107



# 第一章 緒論

本章的內容共可分成五大部分，其主要目的在於陳述本篇論文的研究目的、動機、方法與架構。分別為第 1.1 節述說研究動機，第 1.2 節闡述問題的界定，第 1.3 節說明研究目的，第 1.4 節分析研究方法，第 1.5 節解說論文架構。

## 1.1 研究動機

國內生產的花卉以供應國內市場銷售為主，花卉之生產、消費者需透過運銷體系之運作，使生產者、消費者獲得最大的滿足。然而花卉之運銷體系以批發市場為軸心，擔負花卉集中、均衡與分散之運銷職能工作。政府農業單位為健全花卉運銷體系，讓花卉產業永續發展，自民國七十七年起陸續輔導花卉運銷者、生產者、農民團體，分別於台北、台中、彰化、台南等地興建完成四處花卉批發市場。

每年夏秋兩季之交，由於台灣處在最猛烈熱帶氣旋(Tropical Cyclone)的路徑要衝上，來自南太平洋西北部和南中國海之颱風時常侵襲本島，導致幾乎每年都要受到它的威脅，造成河川中下游流域嚴重的水患。有好幾年災情慘重，除了帶來人民喪失生命與失蹤、房屋的毀損、土壤的流失之外，更嚴重地還導致農漁業相當大的損失，這些都已直接影響到人民的生存發展。以過去十年天然災害造成農業損失的資料來看，台灣近十年間因天然災害所致的農業損失金額高達六百八十五億元以上[15]，其中僅僅颱風所造成的農業損失金額就占有所有損失金額的70.25%[15](約四百八十一億)，由此可見颱風為造成台灣農作物損失之最主要天然災害。以敏督利颱風農業災情報告指出[14]，農業災情估計損失金額約十四億元以上。花卉因深受敏督利颱風的影響，導致產量降低，如菊花產量減少21%，百合產量減少8%等。另外，依照「農業天然災害救助辦法」，每一位農民所受到之救助並非與其受損狀況成比例增加，而是依照目前農委會在實務上的做法。最高的就是果樹和『花卉』，每一公頃救助3萬元。最低的就是稻作和雜糧，一公頃救助8千元。由上述的花卉產量減少與救助金額來看，皆可明顯得知颱風對花卉損失比例有其相當大的影響程度存在。近年來雖然有氣象衛星對颱風的定位，並且有較完整的觀測、記錄、預報及防範措施，使得因颱風侵襲所引致災害據內政部網站(<http://www.moi.gov.tw>)統計有些許改善。但颱風仍是造成台灣地區最主要的氣象災害，如1996年賀伯颱風帶來的豪雨，光僅農產花卉影響部份就使得農民造成莫大的損失。因此，面臨一項重大的挑戰，就是如何去分析評估颱風對花卉造成的損害。對於此問題的存在，勢必要對颱風相關特性與花卉作進一步研究。

基於在2002年成功建置世界第一個台北花卉批發市場之花卉產業資料超市(Data Mart, DM)後，成功的經驗使得其他花卉批發市場也紛紛追隨著這資訊化的腳步，陸續完成了彰化、台中、台南三地之花卉批發市場之資料超市之建置。將每日拍賣作業的交易資料，定時地將資料轉入所屬的各家資料超市，放置資料在於花卉批發資訊分享熱線(Flower Wholesale Information Sharing Hotline, Flower-WISH)。伴隨著各家資料超市的成熟發展，整合四家資料超市的資料是必要的，不但各家批發市場資訊可以分享，也帶來整個台灣的花卉產業經濟莫大的效益。因此使得台灣區花卉總料倉儲(Data Warehouse, DW)因應而生，就先前而言，已完成整合四家資料超市的資料，初步完成了部份的總資料倉儲之建置。近年來，為帶動南部花卉產銷經濟力，高雄市政府及行政院農業委員會繼台北、台中、彰化及台南等地，於高雄市設立全台第五家花卉批發市場，是南台灣最大的內外銷集散中心。高

雄國際批發市場的設立，造成的各大批發市場的衝擊與競爭無非是一項重大的考驗。因此再於2006年建立完成高雄資料超市，為了達到將高雄資料超市資料整合先前四家資料超市資料，故需再次重新做進一步的整合。將原本只建置完成部份總資料倉儲的部份，做最終之整合與整頓，完成未完成的部分，此為本研究動機之一。

另外一方面，完全整合建置好的總資料倉儲，在預期上資料下載的效率比針對各家資料超市下載的效率還要來的提升，因此本研究利用總資料倉儲內的資料，來做資料挖礦技術的應用。因此將研究颱風特性以減少它對台灣區花卉產業之大宗花卉數量所造成的影響，此為本研究動機之二。



## 1.2 問題界定

整條花卉供應鏈所涉及到的相關決策者大體上可分為三大類別，分別是供應所有種類花卉的花卉供應人、負責中介買賣花卉的花卉批發市場經營者以及針對花卉有需求者之花卉承銷人。面對這些不同類型的決策者，其所對於花卉的行情資訊有著不一樣的需求。就以花卉供應人層面來說，他們希望透過來自不同資訊的傳遞與蒐集來選擇最適合的批發市場以及所要供應的花卉數量與種類；以花卉批發市場的經營者角度來看，不同前者的是思考層面上希望市場價格不要高低起伏不穩定以及在於任何情況事件下，殘貨數量的維持與減少；最後再於花卉承銷人方面當然是希望以最適合且有利的價格來購買品質相當且數量不差的花卉。

### 1.2.1 整合 DM 為 DW

以台灣地區花卉供應鏈之所有決策者角度來思考而言，如果能提供更多資訊並能夠分享其他地區供應鏈的相關情報，則能夠快速提升台灣地區花卉產業間之競爭力，進而促進台灣經濟的成長。故本論文延用陳家瑜[5]所採用的競籓(R. Kimball)提出的由下而上建置法(Bottom up)，將新建置的高雄資料超市與台北、台中、彰化、台南四家花卉資料超市進行最後統整程序，並整合出完整的總資料倉儲系統，由於台北、台中、彰化、台南四家超市已於先前完成並初步整合至總資料倉儲。因此本論文將不會針對此四家花市的部分進行詳細描述，圖 1.1 為完整總資料倉儲的系統架構。前段虛線框住的區域為本論文所要完成之部分，將台北、台中、彰化、台南、高雄五家資料超市之資料分別經由資料轉換服務機制，轉入總資料倉儲，並加入市場別的維度。此將再本論文第三章詳細說明其作法，再者，後段無虛線框住的區域則為黃俊端同學於同年完成之部分。

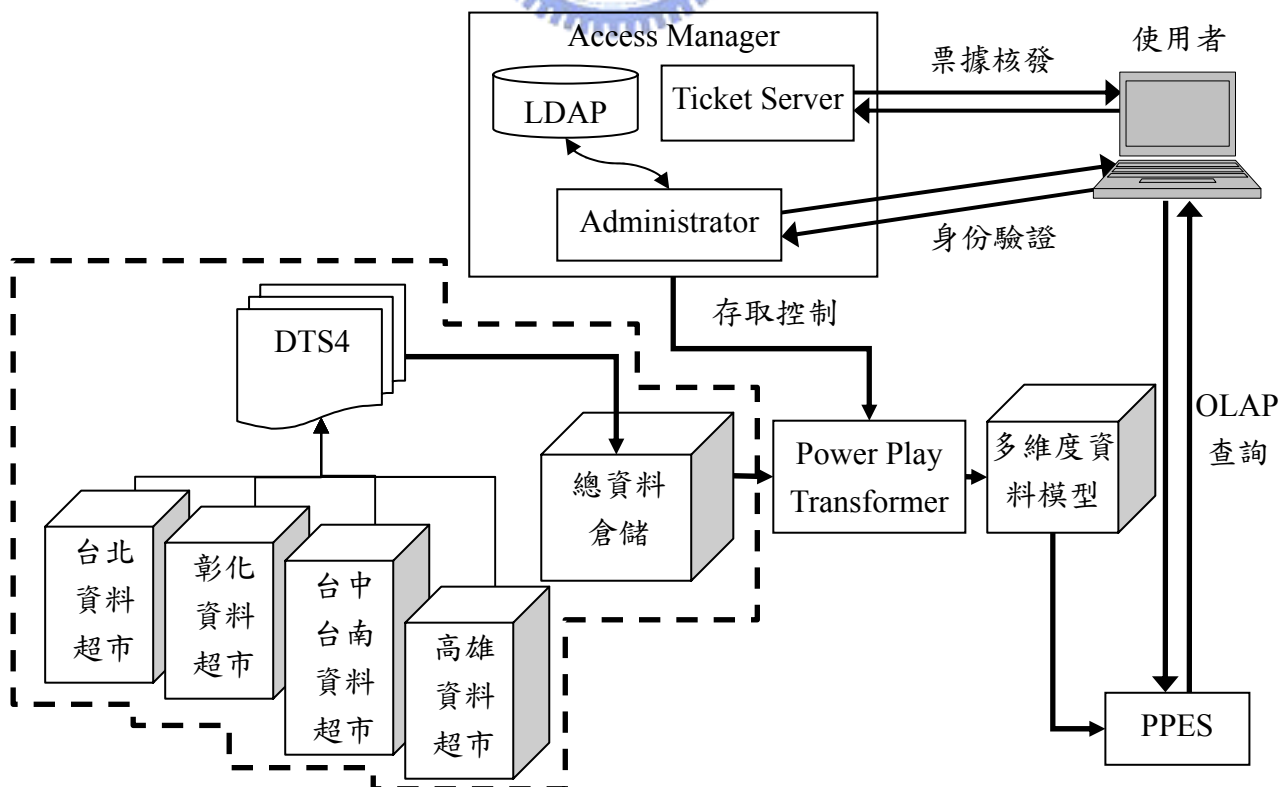


圖 1.1 完整總資料倉儲系統架構圖

## 1.2.2 探勘颱風與大宗花卉資料

颱風侵襲台灣地區的次數不在少數，每次一來就有一定機率會帶來狂風豪雨，而這些其實皆為颱風特性；颱風特性大體上分為颱風路徑、颱風雨量、颱風強度、颱風風速、颱風之最低氣壓等相關特性。這些特性說穿了都是可以作為不同研究方向的出發點，而在此研究將針對研究所需進而擷取選出所要用到的資訊與數據；另一方面，颱風特性將於第二章參考文獻做更進一步說明與探究。

將颱風特性相關資訊與花卉資料利用資料探勘技術進行一連串的分析與研究，若得知颱風侵襲台灣的某些颱風特性，便可進行大宗花卉量的分析。如圖 1.2 所示，本論文在第五章部分將利用總資料倉儲內的花卉資料，與颱風相關特性運用資料探勘技術進行階段性的研究。

在眾多資料探勘模式之中，可以將這些研究歸類出以下五種模式[11]：歸類(Classification)、分群(Clustering)、回歸(Regression)、關聯(Association)和時間序列(Sequence)，以下將針對資料的探勘進行概略說明。

1. 歸類(Classification)：根據「一些」變數的數值進行計算，然後依照計算的結果將資料分類於幾個離散的數值當中。例如企業可以根據客戶的收入、消費金額、年齡等資料進行分析，然後將顧客群分類為主力消費群或是次要消費群。
2. 分群(Clustering)：將資料依照特性的「相似度」進行自然的分組，然後找出群組間的相異以及群組內的相似之處。
3. 回歸(Regression)：在過去「連續性」的資料屬性下，此模式利用過去一系列的數值來預估一個連續數值的可能性。
4. 時間序列(Time series)：利用「現有數值」資料來預測將來可能的數值。此數值會有時間先後發生的順序，並且具有時間的階層性，例如每個月、每個季節、或者是某些假日等。
5. 關聯(Association)：是要找出某些事件同時發生的關聯性。例如颱風強度為中度颱風同時大宗花卉-玫瑰之當天四大市場拍賣量的比較(台北>彰化>台中>台南)。

本研究所要分析的變數(變項)，基本上超過45種以上。因此「分類」與「分群」此兩種模式的資料探勘可能會造成無法有效分析資料，並且變項中不只包含數值面(最大風速、降雨量、移動速度等)，還有包含文字面(近節慶區段、移動路徑、移動方向等)。如此一來，將無法進行數值運算。由於侵台颱風之日並非連續性的資料，如果用來分析侵台颱風這種間斷性資料(停留天數、近節慶個數、最短之近節慶天數等)，必定有其無法預測的困難度存在。同樣地，以時間數列來說，侵台颱風移動路徑並非數值型，故與回歸有著同樣預測的難度，也是不可行的。因此，本研究將選擇關聯性法則進行颱風對大宗花卉影響的一連串分析。由於颱風相關資料與資料倉儲中的大宗花卉是屬於兩種不同類型的資料，因此需要進行階段性關聯分析。加上資料倉儲的資料當中存在一些很難分析的變相存在(休市、節慶等)，所以採用關聯性法則來分析颱風對大宗花卉影響比較適合。故再此研究將不會針對其他四種模式予以深入探討之。

在一般的情況下，關聯法則的探勘方式是採用透過項目組合的方式，然後不斷檢查掃描資料，來判斷項目之間的關聯性。如此可以將颱風特性變相與大宗花卉量變項透過項目

組合的方式正確產生有用的結果，加上資料倉儲資料只有1996年至2007年。故對應侵台颱風相關資料時，可用資料倉儲資料只剩1996年至2005年。接著以各大花卉批發市場作為變項時，資料則會變的更少、更複雜、更加難以分析。

不過隨著「資料量」不斷成長，「資料變項」的增加，電腦系統掃描資料的工作的負擔也會越加沉重，導致每次探勘都必須耗費越來越冗長的時間，才能完成工作。因此近年來不斷有研究聚焦在面對關聯法則探勘效率的議題。利用不同的資料結構以及不同的項目關聯檢查機制，期望能夠有效改善探勘效率，並防止減少探勘效能的降低。

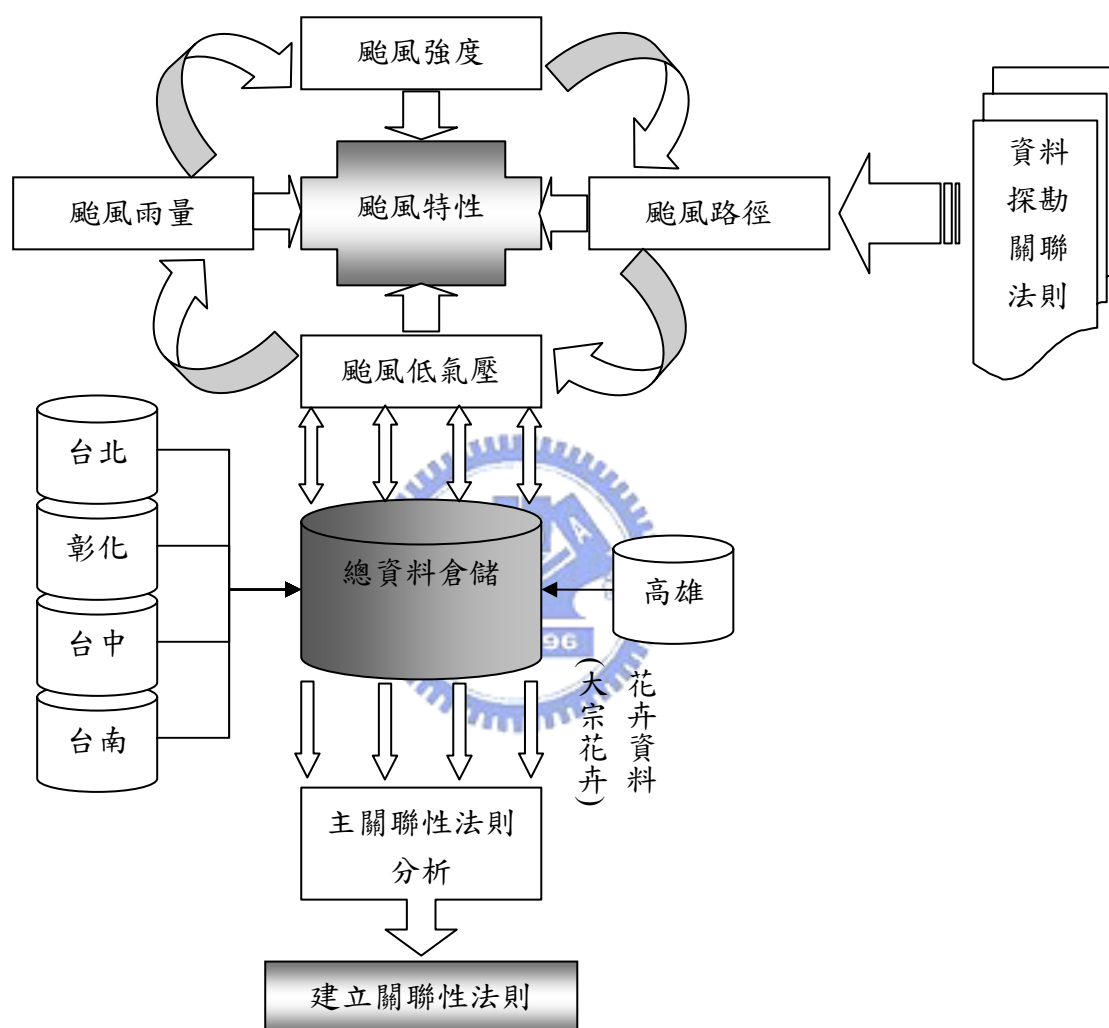


圖 1.2 關聯性法則方法與界定

### 1.2.3 大宗花卉為研究範圍

花卉種類數百種，有些花卉產量少而且不受歡迎，因此拿來做分析，並非是一項很好的選擇。基於以所要分析花卉為考量點。故本次研究指定大宗花卉所限定的花卉種類範圍來進行深入分析，將由文獻探討 2.1 節之 2.1.2 節清楚定義大宗花卉與選定之相關花卉種類範圍。

### 1.3 研究目的

本研究的主要目的之一，在於將新設立的高雄國際批發市場所建置的高雄資料超市，予以向上整合先前已建置好的台北、彰化、台中、台南四家花卉資料超市，使之成為最為完整的總資料倉儲。讓五條花卉供應鏈的決策者能夠整合五家花市的資訊，做最理想最完善的優良決策，基於機密與私密性資料的分類與控制，達到各類決策者無法間接侵犯到其他決策者之權益。如此一來，不但使得台灣區的花卉批發產業能夠增加整體經濟效益也能生生不息，永續經營下去。表 1.1 為最終總倉儲建置完成後，資訊的分享情形與完成建置年份整理。

表 1.1 資訊分享與建置年份示意表[4]

超市與倉儲 決策者	台北資 料超市	台中資 料超市	彰化資 料超市	台南資 料超市	總倉儲	高雄資 料超市	總倉儲
台北花市	◎				◎		
台中花市		◎			◎		
彰化花市			◎		◎		
台南花市				◎	◎		
高雄花市						◎	◎
完成建置年份	2002 年	2003 年	2004 年	2004 年	2005 年	2006 年	2006 年

近年來颱風的大肆侵襲，造成農產花卉業相當大的損失。會造成如此大的影響，主要是基於颱風特性的考量因素，其主要分為幾大部分，颱風路徑、颱風強度等。配合利用總資料倉儲來下載相關花卉資料。相對於分別由各資料超市下載資料還要來的快速，然後將颱風特性與花卉資料利用資料探勘技術，有效率的分析資料中所隱藏的訊息以供未來颱風來襲前之預測準則，此為本研究之主要目的之二。

基於以上種種解說與探討，本研究將針對以下幾個問題提出解決方法與方針：

1. 將最新建置完成的高雄花卉資料超市與先前四家花卉資料超市做最終整合，並完成建置總資料倉儲系統後半部份的工作。
2. 利用颱風特性之相關變數運用資料探勘技術分析颱風對花卉的影響。

## 1.4 研究流程

本論文的實施步驟與研究方法如圖 1.3 研究示意流程圖所示：

1. 業界訪談：先前訪談對象已針對台北、台中、彰化、台南四家花卉批發市場管理經營者，評估確認有其總資料倉儲需求之必要性，藉以初步整合四家資料超市之資料；此次訪談對象則針對新設立之高雄批發市場，透過直接訪談及當面溝通的過程中，了解到總倉儲再次的需求，因此重新針對五家資料超市的系統架構做一番解析，並再次整合各家資料超市之資料。
2. 問題界定與分析：分析歷年颱風各種特性對於大宗花卉數量的影響力以及針對颱風資訊情報進行有用的運銷決策。
3. 文獻回顧：參考相關文獻如資料倉儲、颱風特性分析、關聯性法則等，作為日後之研究理論基礎、系統基本架構及研究方法的依據。
4. 倉儲系統設計與實作：利用以前所發現並運用在花卉相關資訊所發展出的資料倉儲系統理論基礎架構，修改先前未完成部份。將新增的第五家資料超市予以整合，實作完整台灣區花卉資料倉儲系統。
5. 關聯性法則：利用近幾年在於台灣地區所發生之颱風情報，配合花卉業務情報網之資料蒐集，找尋颱風對於大宗花卉之影響，分析並建立關聯性法則。

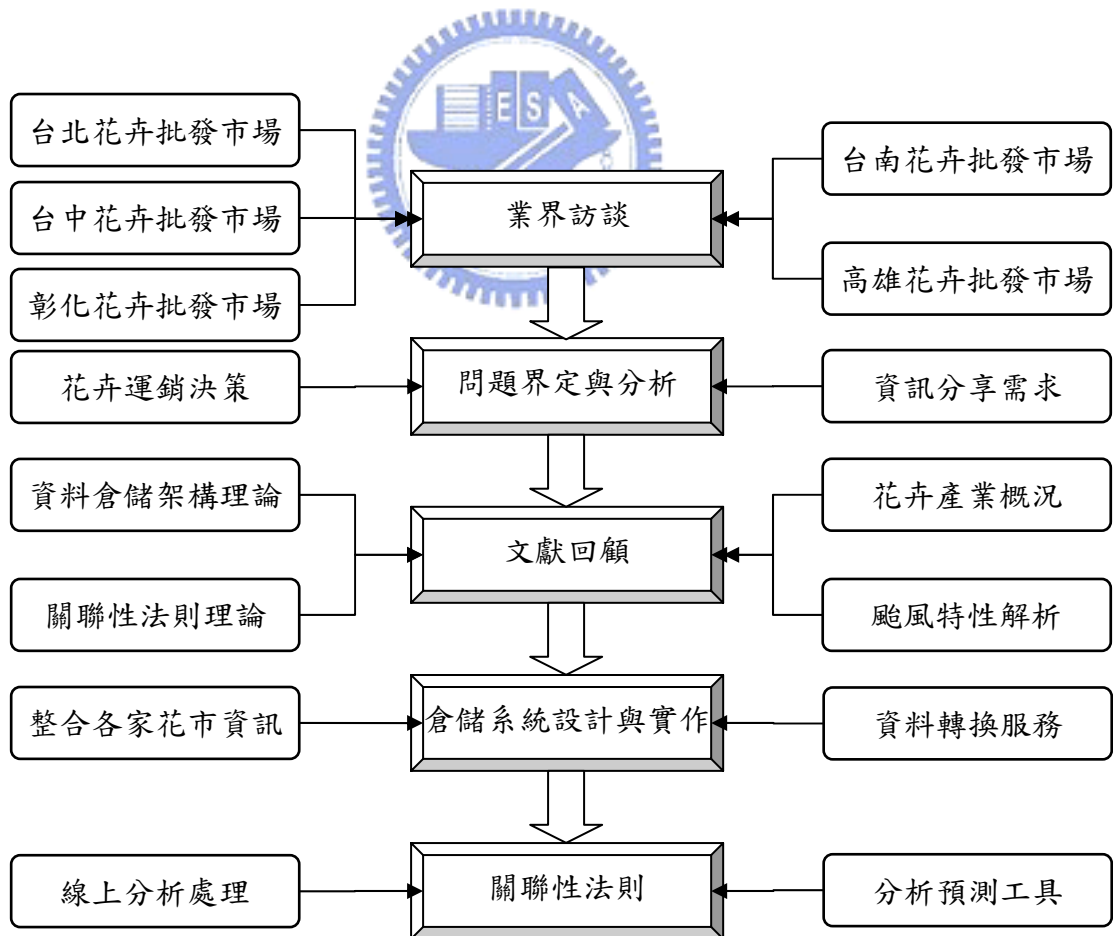


圖 1.3 研究流程圖

在關聯性法則部份，資料來源取決於總資料倉儲，基於下載總倉儲資料比個別下載各家資料超市資料還要來的有效率而且更省時。另一方面，在於分析過程中，首要使用初步關聯性法則從歷年颱風各種特性中找出颱風相關資訊情報強關聯規則(Strong Rule)，並同時做驗證。再來進行本研究重心的關聯性法則針對颱風特性與大宗花卉量的相關項目予以分析，建立關聯性法則找出並分析颱風對大宗花卉之影響。



## 1.5 論文架構

本論文的内容编排如下：

- 第一章：緒論—說明本論文之研究動機、問題界定、研究目的以及研究方法。
- 第二章：文獻回顧—包含資料倉儲的基本觀念與架構、颱風特性、關聯性法則演算法之相關理論。
- 第三章：總資料倉儲資料整合實作—先介紹總資料倉儲之系統基礎架構，接著再將高雄資料超市與先前整合好四家資料超市之資料做整合前之確認與剖析，共同利用資料轉換服務(1)轉入總資料倉儲系統內。
- 第四章：總資料倉儲資料轉換實作—將整合完成之五家資料超市資料進行資料轉換(2)與自動化排程建置。
- 第五章：關聯性法則—針對歷年颱風情報進行初步關聯性法則分析，再將颱風特性與花卉相關資料整合進行本研究重心的關聯性法則，以提供未來評估決策依據。
- 第六章：結論、建議與未來研究方向。



## 第二章 文獻回顧

本章主要的目的在於回顧資料倉儲的基礎理論與建構模式，並介紹關於花卉產業相關認知與颱風本身之特性分析、關聯性法則 APRIORI 演算法之理論基礎等相關文獻。本章共分為四節，第 2.1 節為花卉產業、節慶之相關介紹；第 2.2 節探討颱風幾大特性分析的相關介紹；第 2.3 節介紹資料倉儲的相關理論基礎與技術及線上分析處理的基本操作步驟；第 2.4 節說明計量方法：關聯性法則相關理論。

### 2.1 花卉產業介紹

花卉一詞依據我國古老文獻，花卉二字的花是指近似花瓣的苞葉，卉字是指具有欣賞價值的葉片和木本植物，所以花卉二字應包括所有具有觀賞價值的草本或木本植物在內。而現代所謂之花卉，乃是指除花朵外，其莖、葉、果等部份均可加以利用的植物。而依據台灣省政府農林廳出版之「台灣農業年報」，將花卉分成切花(cut flower)、盆花(potted flower)、球根(bulbs)、苗圃(nurseries)、種籽(herbaceous flower seed)等五大類，因此所謂花卉產業(flower industry)就是以這五類花卉為主要經營範圍的產業。

其中切花係指專為剪取花朵或枝葉供插花、花束、花籃、花園等應用者，目前為台灣產量最大用途最廣的花卉。本節主要分成三小節，為了初步了解花卉產地分布，因此於第 2.1.1 節解說位在台灣各地區分布切花種類的產地；由於大宗花卉所受影響，故於第 2.1.2 節解釋大宗花卉大體上明確的定義；第 2.1.3 節闡述花卉受節慶的影響。

#### 2.1.1 花卉切花種類之產地

台灣地區切花類花卉主要的生產地區，均集中在中南部，尤其以台中縣、彰化縣、南投縣、嘉義縣、台南縣、高雄縣以及屏東縣等為主要生產地區。總種植面積占全省切花類種植面積 95%以上，國產切花產地分佈如圖 2.1 所示。



圖 2.1 國產切花產業分布圖[19]

在於眾多花卉種類當中，就其針對菊花花卉產地為例，說明菊花每年所有產地之總產量與各地產量之比較，其說明如下。

近幾年以來，在於菊花的所有產地總產量有明顯的逐年下降的趨勢如圖 2.2 所示。92 年較 91 年總產量少掉將近 5000(千打)的總產量值，接下來於 92 年與 93 年之間緩和了下降的速度，直到 94 年又逐漸減少約 5000(千打)總產量值，由此可知菊花每年總產量的成長趨勢呈現負成長情形。

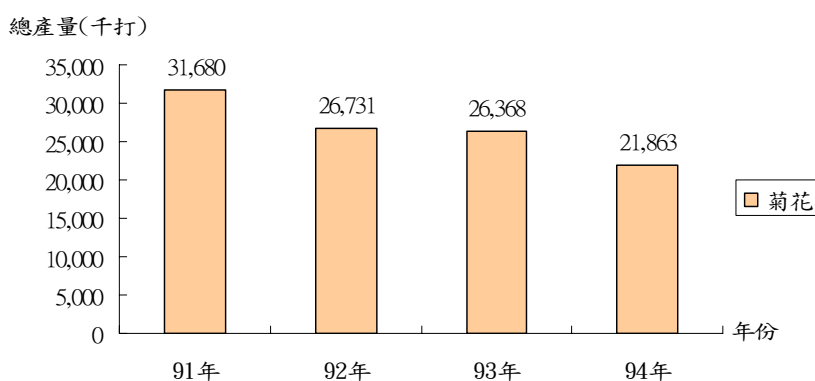


圖 2.2 菊花每年所有產地總產量比較圖

針對菊花負成長的趨勢，故在此深入探究台灣各個菊花產地產量值，並以 94 年菊花的各產地產量呈現如圖 2.3 所示。由圖 2.3 可明顯看出彰化縣產地產量高於其他任何一個縣市產地量，除了彰化縣以外的縣市產量加總，其總量還不到彰化縣產量的 10%。

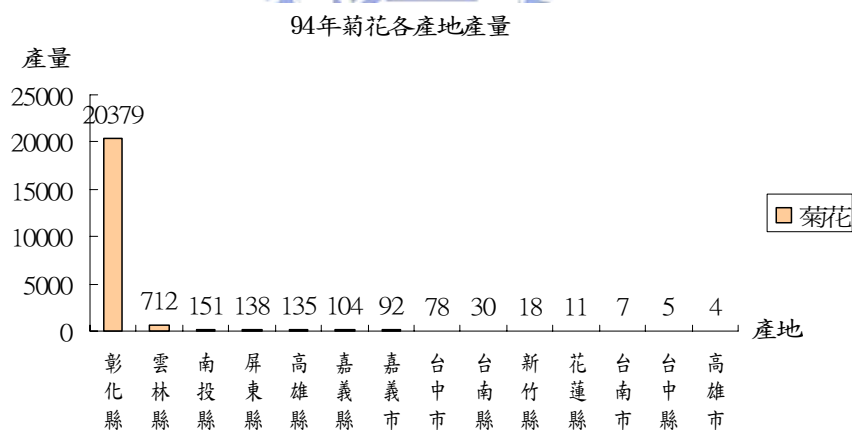


圖 2.3 菊花每年各產地產量比較圖

## 2.1.2 大宗花卉

大宗花卉[3]就是交易量排名順位比其他花卉還要領先，且少樣多量。在於訂單量也有一定程度以及在於所有花卉種類而言，不管針對數量或者是價格為考量出發點，予以前 10 或前 20 名的名次所排序出來的花卉結果稱之。相對於供給、需求量也就非常龐大。換而言之，消費者都能購買到的花種，也是買家最常最普遍喜歡的種類，不但最受歡迎，而且在於各種禮儀場合中也適用。其大宗花卉如菊花(大菊、小菊)、玫瑰、香水(葵)百合、康乃馨中輪、劍蘭、火鶴花、非洲菊、洋桔梗十種(如表 2.1 所示)。然而花卉價位有高價高檔花卉、中高檔花卉以及低檔平價容易購買到的「大宗花卉」。

表 2.1 前十名之大宗花卉排序表(2005 年)

排序	花卉名稱	總成交量(把)
1	火鶴花	98,442,563
2	小菊	91,489,375
3	香水百合	59,340,994
4	大菊	56,034,178
5	玫瑰	53,949,723
6	非洲菊	43,952,297
7	劍蘭	28,380,092
8	葵百合	26,042,755
9	康乃馨中輪	19,378,977
10	洋桔梗	11,346,182

### 2.1.3 花卉種類與節慶

花卉種類易受到節慶的影響，尤其是在於節慶前後幾天的間隔天數影響最為深遠。如圖 2.4 可以得知在 2003 年 2 月 14 日情人節將至的前幾天，圖形顯示出在於 2003 年 2 月 10 日此一時間點，玫瑰花的需求量就會開始慢慢有了提升的現象。也就是說，靠近節慶前 3~5 天這幾天為需求量之上升變化期。接著，依據花卉拍賣業者所累積之經驗指出，大約在節慶的前 2~3 天範圍左右時，玫瑰花的需求量則會達到最高峰的狀態。這與總倉儲所顯示的資料一致且吻合，到了情人節當天，需求量也就會慢慢地穩定下來，這也就得知玫瑰花為情人節節日中最具代表意義的象徵物。

節慶對應花卉[18]，如何選擇適合的花卉，讓過節過的更加有意義是一件相當重要的事。相對的，對於研究花卉節慶之間的關係也就會更加明確，讓預測更加準確，如表 2.2、表 2.3 所示。再者，生產季節性的花卉種類，隨季節性的變化，花卉產量也因此而上升或下降，如表 2.4 解說之。

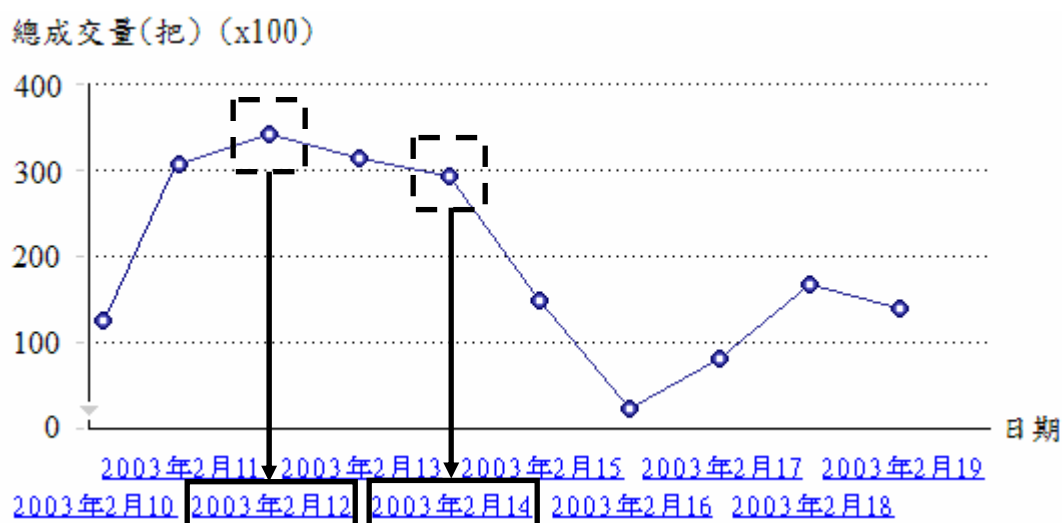


圖 2.4 2003 年玫瑰花之總成交量變化趨勢圖

表 2.2 國曆特殊節慶

編號	國曆日期	節日名稱	節慶對應花卉
1	1 月 1 日	開國紀念日	無特定花種
2	2 月 14 日	西洋情人節	玫瑰、百合、鬱金香、海芋
3	3 月 14 日	白色情人節	玫瑰、百合、鬱金香、海芋
4	4 月 5 日	清明節	劍蘭、菊花、百合、海芋
5	5 月 8 日	母親節	康乃馨
6	8 月 8 日	父親節	非洲菊、向日葵、石斛蘭
7	9 月 28 日	教師節	無特定花種
8	10 月 10 日	國慶紀念日	無特定花種
9	12 月 25 日	聖誕節	聖誕紅

表 2.3 農曆特殊節慶

編號	農曆日期	節日名稱	編號	農曆日期	節日名稱
1	1 月 1 日	農曆春節	11	6 月 23 日	關公生
2	1 月 9 日	天公生	12	7 月 7 日	七夕中國情人節
3	1 月 15 日	元宵節	13	7 月 15 日	中元節
4	2 月 2 日	土地公生	14	8 月 15 日	中秋節
5	2 月 19 日	觀音生	15	9 月 9 日	重陽節
6	3 月 3 日	三日節	16	9 月 19 日	觀音出家
7	3 月 23 日	媽祖生	17	12 月 16 日	尾牙
8	4 月 8 日	浴佛節	18	12 月 24 日	送神
9	5 月 5 日	端午節	農曆節慶適用花卉無特定花種 (七夕中國情人節是用玫瑰花)		
10	6 月 19 日	觀音得道			

表 2.4 具有生產季節性花卉主要產品

季節	月份	花卉 1	花卉 2
春	2 月	金魚草	玫瑰花
	3 月	非洲菊	玫瑰花
	4 月	洋桔梗	
夏	5 月	康乃馨	海芋
	6 月	火鶴花	滿天星
	7 月	星辰花	向日葵
秋	8 月	薑荷花	
	9 月	文心蘭	
	10 月	夜來香	
冬	11 月	百合	
	12 月	唐菖蒲	
	1 月	鬱金香	

## 2.2 颱風特性分析

西太平洋的颱風多半都發生在菲律賓以東的海上，如加羅林群島(Karolinen)等地。這些地方都是在北緯十度左右，也就是在東北信風帶內，同時太平洋上在夏季經常有高氣壓，颱風形成後便跟隨高氣壓環流向西至西北進行。有時也會走西北或更偏向北的路線，端視當時颱風周圍氣流方向而定。一般常是指向臺灣的方向，當颱風漸漸進行到北緯二十至三十度之間後，已到高氣壓的邊緣，而漸轉向北進入西風帶，受西風帶影響轉向東北。菲律賓、臺灣、琉球、日本等地，因在颱風必經之路附近，故常受颱風的侵襲而發生災害。據中央氣象局[16]統計熱帶氣旋可知颱風生成的條件有五大特性：

(1).表面溫度達攝氏 26 度(也有 26.5 度及 27 度的說法)以上的廣大溫暖海域，才能提供充足的水汽蒸發。(2).須在南北緯五度以下的地區，因為在此地球自轉偏向力(科氏力)才大得足以使氣流形成氣旋。(3).高低層的大氣之間的風速不能相差太大才不致將已發展的積雲吹散，如果沒有發展足夠的積雲，熱帶氣旋不會產生。(4).氣柱內明顯地不穩定或表面氣壓相對較低，並且通常在其上方有一反氣旋。(5).其他特性如近中心『最大風速』達到或大於 33m/s、『最低氣壓』達到或小於 980mb。

經由上述對於颱風特性的認知，得知颱風特性之多，並非在短時間就可以弄清楚的。所以為了能夠更清楚地了解每種颱風特性，故針對颱風路徑、降雨量、強度、近中心最大風速、移動速度、近中心最低氣壓六大特性做進一步分析與解說。

本節主要分成四小節，第 2.3.1 節說明颱風路徑與降雨量的關係；第 2.3.2 節述說颱風強度與近中心最大風速的特性轉換；第 2.3.3 節解說颱風移動速度；第 2.3.4 節闡述颱風強度與近中心最低氣壓的相關。

### 2.2.1 颱風路徑與降雨量

颱風侵襲常會帶來豪雨，但即使是同樣路徑與強度的颱風，降雨量亦可能大不相同。颱風降雨量的多寡須視路徑、強度、水汽含量、雲雨分布、地形、移動速度等而定。根據路徑分析[16]，各地降雨情況可歸納出下面七種情形：

1. 第一類路徑：『通過台灣北部或北部海上，向西或西北進行者。』颱風的降雨以北部地區最嚴重，中部山區雨量亦多。如當時大陸上有高氣壓南下，更能增強雨勢，常導致北部的大水災。
2. 第二類路徑：『穿過本省中部，向西或西北進行者。』颱風在登陸前，北部及東部地區雨勢較強，穿過本省中部後，南部因西南風吹入導致雨勢增強，但以中南部山區雨勢最烈，尤以阿里山地區為最嚴重。
3. 第三類路徑：『通過台灣南部或南部海上，向西或西北進行者。』颱風從台灣南端通過，除東部雨量較多之外，其他地區雨量不多。
4. 第四類路徑：『沿東岸或東部海上北上者。』颱風沿東岸或東方海面北上，以東部降雨最多，有時北部及東北部亦有較強雨勢。
5. 第五、六類路徑：『沿西岸或台灣海峽北上者以及通過中南部，再向東北出海者。』颱風對西南部及東南部影響較大，雨量最多雨勢亦強烈，東部、北部及東北部雨量不多。
6. 第 7 類路徑：為不屬於以上六類之特殊路徑者。

以上所描述之七類路徑可以由圖 2.5[16]清楚看出其分布情形：

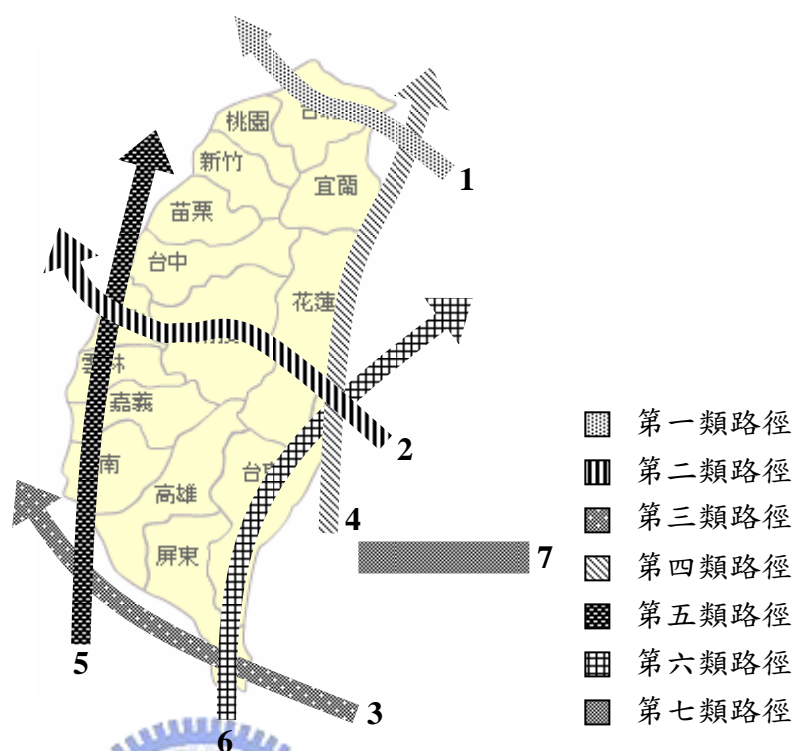


圖 2.5 颱風路徑分類圖[16]

### 2.2.2 颱風強度與近中心最大風速

颱風強度由強至弱可分為超級強烈颱風、強烈颱風、中度颱風、輕度颱風及熱帶低壓五種強度類型如表 2.5 所示。但是一般來說都只會分成強烈颱風、中度颱風、輕度颱風三種主要強度類型。一般人通常誤以為以肉眼所見外圍環流大小區分颱風強度，其實應該以近颱風中心附近最大平均風速來劃分才有其真正價值[17]。相對的，依颱風近中心最大風速來區別，其包括平均風速和瞬間最大陣風，兩者相差可達一、二倍之多。另一方面，近中心最大風速如果越大，則所產生的破壞也就相對的越大，甚至會發生山洪暴發、河水暴漲，淹水等問題。

表 2.5 強度與近中心最大風速對照轉換表

颱風強度類型	近中心最大風速		
	公尺/秒	海浬/時	風級
熱帶低壓	17.2 以下	34 以下	8 級以下
輕度颱風	17.2~32.6	34~63	8 級~11 級
中度颱風	32.7~50.9	64~99	12 級~15 級
強烈颱風	51.0~66.9	100~129	16 級~19 級
超級強烈颱風	67 以上	130 以上	20 級以上

### 2.2.3 颱風移動速度

移動速度意指颱風中心移動的速度。由於颱風移動速度的快慢會直接的影響颱風對於所在位置的災害程度，故將颱風移動速之研究列入考慮。假設地球為球形，且時間間隔二十四小時之兩點經緯度已知，則移動速度  $V$  的定義如下[6]：

$$V = \frac{1}{24} [6371 \times \cos^{-1}(\sin \phi_1 \sin \phi_2 + \cos \phi_1 \cos \phi_2 \cos(\lambda_2 - \lambda_1))]$$

其中  $\phi_1$ 、 $\phi_2$ ：時間間隔二十四小時之兩緯度點

$\lambda_1$ 、 $\lambda_2$ ：時間間隔二十四小時之兩經度點

地球的半徑：7371(單位：km)

移動速度： $V$  (km/hr)

### 2.2.4 颱風強度與近中心最低氣壓

目前颱風中心最低氣壓的高低是根據衛星雲圖決定最大風速，再由經驗公式換算而來。由此可知中心最低氣壓與中心最高風力並非絕對的線性關係。過去根據飛機觀測顯示，兩個颱風的最低氣壓一樣，而最大風速卻並非亦完全相同。相對的，當一個熱帶氣旋有相當低的氣壓時，都是集中在中心形成強烈的氣壓梯度的，這其實是一般熱帶氣旋的特性。因此一個熱帶氣旋氣壓極低時可以推論風力亦強，這在大部份情況都會成立。

基於颱風分級的標準各家說法皆無一定共識，故在眾多分級的方法中進而選擇其中一個由香港熱帶氣旋追擊站(<http://hkciw.uhome.net/>)所定義的標準，如表 2.6[6]所示：

表 2.6 颱風強度與近中心最低氣壓之對照轉換表

颱風類型	颱風強度(對應)	近中心最低氣壓(mb)
一級颱風	熱帶低壓	$P > 980$
二級颱風	輕度颱風	$980 \geq P > 965$
三級颱風	中度颱風	$965 \geq P > 945$
四級颱風	強烈颱風	$945 \geq P > 920$
五級颱風	超級強烈颱風	$P \leq 920$

### 2.2.5 小結

綜合以上所有颱風特性之解析與探討，了解到單從單一颱風特性進行一連串分析是無法從中獲得有用的資訊。必須找到其所互相對應的颱風特性，如此一來，在於研究颱風特性層面上，才有一定程度上的價值。如颱風七大路徑對應之降雨量而言，每當颱風侵襲台灣時，相對的，一定順勢帶來大風大雨。因此，也就產生了一個有價值的資訊，那就是針對哪些颱風路徑導致台灣哪部份雨量最多最為嚴重，故在此為了更清楚了解其關係，所以進一步整理在表2.7所示：

表 2.7 颱風七大路徑對應之降雨量程度表

降雨量 颱風路徑	北部地區	東部地區	中部地區	南部地區 (西南與東南)	降雨量偵測點
第一類路徑	◎	□	△	□	北部
第二類路徑	◎	◎	◎	◎	全部
第三類路徑	□	◎	□	□	東部
第四類路徑	△	◎	□	□	東部
第五類路徑	□	□	□	◎	南部
第六類路徑	□	□	□	◎	南部
第七類路徑					

◎ 降雨大且嚴重    △ 降雨適中    □ 降雨不大

由表2.7可以清楚得知，每種颱風路徑所對應之台灣地區降雨量的程度。對其降雨程度大小，在此可以選定所要偵測降雨量的點，可分為北部、東部與南部。再者，也可以進一步將相似路徑予以整合並合併之，例如第三類路徑與第四類路徑可以歸納成同一類。同樣地，第五類路徑與第六類路徑也可以歸納成另一類，因此七類颱風路徑也就可以縮減成三到四類左右(第七類路徑可以忽略不管、第二類路徑可以依照不同要求，給定所要針對的降雨量偵測點)。如此縮減動作，不但可以減少分析之維度數，也可以更容易拿來進行一連串的研究依據。

接者，颱風特性去掉上述所探討的颱風路徑與降雨量之後，剩下颱風強度、颱風近中心最大風速、颱風風速及颱風近中心最低氣壓四種。其中颱風風速尚未有其對應的颱風特性，本身又為數值型資料，故在此不予探討。颱風強度對應颱風近中心最大風速，另一方面，颱風強度又對應颱風近中心最低氣壓，故可以選擇颱風近中心最大風速或颱風近中心最低氣壓來取代颱風強度(如表2.8所示)以方便計算。

表 2.8 颱風強度對應近中心最大風速與最低氣壓之對照轉換表

颱風強度類型	近中心最大風速與最低氣壓		
	公尺/秒	海浬/時	近中心最低氣壓(mb)
熱帶低壓	17.2 以下	34 以下	$P > 980$
輕度颱風	17.2~32.6	34~63	$980 \geq P > 965$
中度颱風	32.7~50.9	64~99	$965 \geq P > 945$
強烈颱風	51.0~66.9	100~129	$945 \geq P > 920$
超級強烈颱風	67 以上	130 以上	$P \leq 920$

表2.8顯示出近中心最大風速或最低氣壓兩者皆可以取代颱風強度。由於此兩者皆為數值型資料，因此有效取代颱風強度有利於在做資料分析時，數值型資料能夠更接近真實，更方便做分析。

## 2.3 資料倉儲

伴隨著電腦及資訊科技的時代來臨與進步的快速，早期一般所使用傳統的關聯式資料庫已經無法滿足大多數的使用者。人們對於資訊的需求不停高漲，需要的是更加快速且具有彈性與多元化的資料分析。單憑藉著以往的資料庫技術來進行大量資料分析已不敷使用，資料倉儲也因應此背景的需求而誕生於世。資料倉儲能夠運用的範圍非常廣泛，不但能夠透過擷取與清理分散於不同地區及不同環境下的資料庫資料，而且還可以對資料庫資料進行轉換與整合，將資料儲存在資料倉儲內，以供未來做決策之用。

在 1990 年代，殷默(B. Inmon)提出「資料倉儲」一詞，並於兩年後的 1992 年定義資料倉儲是具有主題導向(Subject-oriented)、整合性(Integrated)、隨時間變化(Time-Variant)及非揮發性(Non-Volatile)四種特性之資料庫[20]。此四種特性簡單敘述如表 2.9 所示。

表 2.9 資料倉儲之四大特性

特性	內容簡述
主題導向	資料倉儲是基於分析特定的主題上而建立的，所有與主題無關的資料皆會被刪除。例如高雄花卉批發市場所建立的資料倉儲，是以在高雄拍賣的花卉銷售為主，事實上並不會存在高雄以外的拍賣資料
整合性	於資料倉儲內的資料都須具有一定相當程度的整合性，主要是整合企業中可能同時存在的多個資料庫，並且跨越不同的平台，透過資料轉換過程機制，使欄位顯示名稱、編碼表示方式、日期時間形式等相關主題屬性使之具有一致性的格式。
隨時間變化	當新資料被增加儲存進日常的作業系統中，基於維持資料倉儲的可用性，因此須在某些特定的時間點到作業系統中擷取新資料，如此一來方可確保資料它的資料是最具時效性的。
非揮發性	資料一旦被新增儲存至資料倉儲之後，便難以被作任何更動，只會被使用者查詢，因此它具有唯讀和累積的性質。

本節主要分成兩小節，第 2.3.1 節說明資料倉儲的建構模式；第 2.3.2 節介紹線上分析處理的模式基礎。

### 2.3.1 資料倉儲的建構模式

在 1990 年殷默提出由上而下(Top-down)的資料倉儲建立模式之後，接著競菠又於 1996 年提出由下而上(Bottom-up)的建構方法[20]。換句話說，前者所主張由上而下法，主要是建立在以企業資料倉儲為頂點概念下的一種傳統關聯式資料庫工具，向下分支延伸擴展至不同部門類型所屬的資料庫，讓決策上的需求能夠獲得最大的滿足。然而競菠所提倡的由下而上法[23][22]，簡單的來說，是針對各種不同的商業程序來建立其所屬的資料超市(Data Mart)，並藉由資料通道(Data bus)及制式化維度(Conformed Dimensions)的相關概念，將下層所有資料超市利用最有效率的方式集結成為一個主要的總資料倉儲。由於花卉資料倉儲概念建立在由下而上的建置方式，故以下只針對兩種資料倉儲建構之其中一種競菠模式詳述其內容。

### 競簇模式：

競簇的資料倉儲模式架構主要是以維度資料建模法(Dimensional Data Modeling)來建置出資料倉儲。然而，於維度資料建模法中，它採用資料表來做為建構模式的基礎，而非運用以往傳統式的實體關聯圖。在競簇模型中，其資料表之構成要素包含事物表(Fact Table)及維度資料表(Dimensional Table)。

事物表內所存放的資料列是大量且不允許重覆的。歷史資料往往都成千上萬筆，較少的欄位，且只有部分為了做為關聯索引所必須利用文字資料呈現外，大部分的資料皆為數值型態。另外，表內存在一定需求對應之維度資料表所作為關連的外鍵(Foreign Key)。屬於靜態資料內容，一經載入之後便不能再做任何異動修改動作。然而維度資料表相較於事物表之間差異程度上，僅上百或上千筆資料其資料量極小，表內擁有較多的欄位，且大部份皆為文字型態。將所有於事物表內資料的屬性皆儲存在維度資料表內，有效率並有彈性的管理維度內之階層關係，屬於動態資料型態與修改異動是被允許的。

緊接著，競簇的維度資料模型依其所屬結構之不同可進一步區分成三種主要的綱要：分別為圖 2.6 的星狀綱要(Star Schema)、圖 2.7 的雪花綱要(Snowflake Schema)、圖 2.8 的星座綱要(Fact Constellation Schema)[12]。

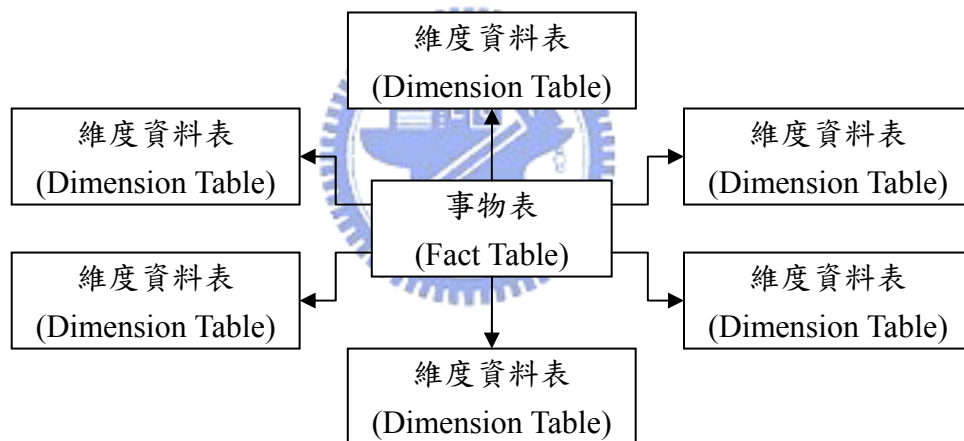


圖 2.6 星狀綱要圖

星狀綱要圖是大部分最多人所會採用的多維度模型。它主要是以事物表及多個維度資料表所組合而成，以事物表為中心並由多個相關的維度資料表環繞著，其形狀有如星形。圖 2.6 為星狀綱要示意圖。而第二種綱要圖由於結構形狀像似雪花的結晶體，所以稱之為雪花綱要圖。圖 2.7 為雪花綱要之示意圖。可為星狀綱要圖的延伸或者是變化，因為事實表格的資料來源有可能是來自企業已存在的作業型資料庫，而這些資料庫當初或許是以正規劃的議題所規劃出的結構，為了達成企業關心主題的查詢目的。因此在某個維度下再次建立其維度的延伸維度，以符合企業目的。

星狀綱要圖與雪花星狀綱要圖類似但經仔細分析與研究，發現與雪花是綱要圖最主要的地方在於：

1. 星狀綱要圖的結構較於簡潔查詢效率比雪花綱要圖好。
2. 雪花綱要圖比較接近正規劃資料庫格式，所以轉換成本來說相對的就會比星狀綱

要還要來的低。

而應用層面來說則是必須視主題的複雜度及轉換成本加以考量，並無哪種綱要圖比較好的說法。不過當事實表格的資料是需要做加總或是其他轉換運算後資料，則必須規劃在最接近事實表格的維度以利轉換處理。

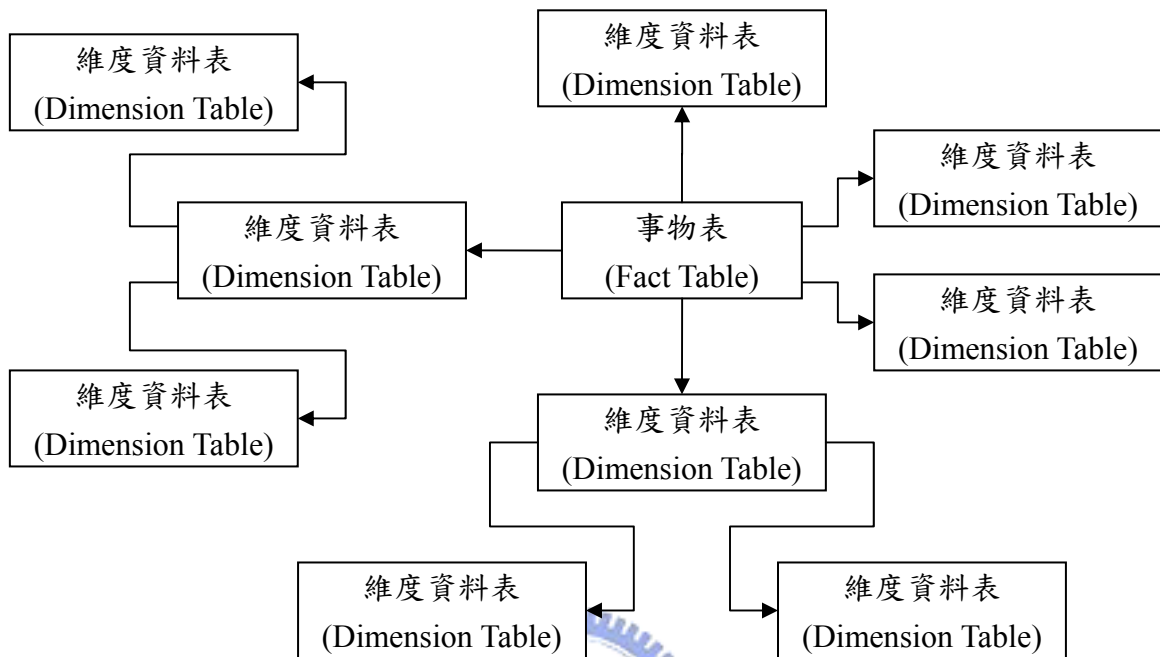


圖 2.7 雪花綱要圖

第三種綱要架構它是由多個不同的事實表格去分享所有可能用到的各種維度，一個事實表格可以想像成是為了達成某個企業關心的主題所設計出來的表格。當這些事實表格所需分析的維度有被共用時，就可以用此種綱要來表示。換言之，這種可以看成是數個星狀式綱要的結合方式，有人亦稱為星系綱要 (Galaxy Schema) 或是事實星座 (Fact Constellation)，統稱為星座綱要圖 (Fact Constellation Schema)。如圖 2.8 所示。

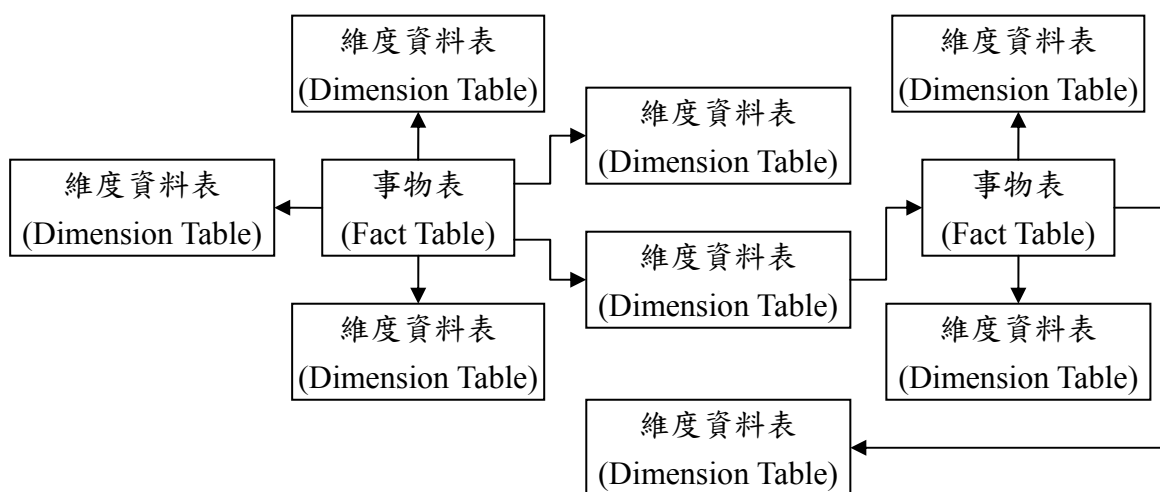


圖 2.8 星座綱要圖

競簇模式之資料倉儲建構概念是由下往上的方式，基於單一的商業流程立足點因而建

立一個資料倉儲，競箴將此種架構的資料倉儲稱之為資料超市。對許多企業組織而言，其內部同時包含多個商業流程可能性很高。因此在建置資料倉儲的過程中，需要針對每個不同商業流程建立所屬的資料超市，爾後再將這些數個資料超市利用最有效率的方式進行向上整合出一個整體的資料倉儲。

競箴建議最好一次建置單一個資料超市的方式做考量為其發展方法論，並針對單一個資料超市的建構，分別提出四個維度化設計之流程步驟。即(1).選擇商業流程：競箴賦與商業流程非常廣泛的定義，其定義可能包含有零售的銷售點、存貨、訂單等。故建議選擇一個最有影響力的流程作為資料倉儲的首要商業程序的依據，其最有影響力在這裡是指當選擇此流程對於最迫切的商業問題可以被順利解決。(2).定義資料細度：定義資料細度在此可以解釋為決定資料倉儲中所包含資料詳細程度的流程。另外最為底層且詳細的資料細度稱為原子(Atomic)，也就表示不能再被細部分割的資料。針對如何才能選擇出最適當的資料細度是非常重要的一門學問，因為不適當的選擇資料細度，會影響資料倉儲的使用者對於資料分析上受到非常大的限制，例如某資料倉儲內僅包含每月資料而受限於缺少每月的資料。(3).選擇維度：完成資料細度的定義後，接下來將選擇適當的維度資料表。在於大量的屬性維度資料表內，它主要提供使用者進行更多面向的資料查詢。(4).確認事物表中的衡量值：在於流程的最後一步則是決定事物表中的衡量值。衡量值的選擇對於系統開發者而言，不僅具有更多的查詢結果同樣地也決定了資料倉儲系統的效能。

以競箴所提倡的維度資料建模法來說，把資料表作為其架構，其主要包含事物表及維度資料表兩個基本組成要素。另外，競箴重視於流程導向的資料模式，嘗試地去找尋不同商業流程間存在的交互作用。換而言之，就是找到資料倉儲中重要的衡量值(Measures)與屬性(Dimensions)。選擇建構一個競箴模式的資料倉儲，花的時間也就比較少，效率也就相對提升。

### 2.3.2 線上分析處理

在 1993 年 Codd E.F 提出線上分析處理(On-Line Analytical Processing, OLAP)的概念 [13][10]，簡單的來說，它是一種應用程式有著能夠直接存取資料倉儲且本身具有資料分析的功能，而非僅僅只是作為報表工具的替代品。超方體(Cube)為資料倉儲裡資料的表示型態，使用者透過線上分析處理機制對維度及衡量值進行操作，則可以對資料做適當的加值應用，藉此獲得一定程度的資訊需求。

線上分析處理的基本十大操作方式，可分為兩個主要層面，第一個層面與維度有關的操作，例如上捲(Roll-Up)、下挖(Drill-Down)、切片(Slicing)、切丁(Dicing)與轉軸(Pivoting)五種。第二個層面與衡量值有關的操作，有視覺化(Visualize)、篩選(Selecting)、輸出(Output)、排序(Sorting)及計算(Computing)五種，整理如表 2.10 所示。

表 2.10 OLAP 基本十大操作說明表

層面	操作方式	解說
第一層面 (與維度相關)	上捲(Roll-Up)	將維度的單位向上放大
	下挖(Drill-Down)	將維度的單位向下縮小
	切片(Slicing)	限制某維度的上下限範圍
	切丁(Dicing)	限制多維度的上下限範圍
	轉軸(Pivoting)	把水平維度與垂直維度做交換(90 度)
第二層面 (與衡量值相關)	視覺化(Visualize)	把數值化的衡量值利用圖形表達
	篩選(Selecting)	限制住衡量值的範圍
	輸出(Output)	將衡量值另儲存為檔案並供加值應用
	排序(Sorting)	將衡量值由大到小或由小到大排列
	計算(Computing)	用特定程式來修改衡量值

線上分析處理大體上主要可分為三種資料倉儲儲存體的型態，分別為多維度線上分析處理(Multidimensional On-Line Analytical Processing, MOLAP)、關聯式線上分析處理(Relational On-Line Analytical Processing, MOLAP)、混合式線上分析處理(Hybrid On-Line Analytical Processing, MOLAP)。將於下列分別解說其意義，其意義分述如下。

1. ROLAP：利用傳統的關聯式資料庫來存放並彙總資料，其好處是利用現有儲存架構，因此比較容易完成，但是查詢效率可能較差與所能儲存的空間較小的種種缺失。
2. MOLAP：真正的多維度資料庫結構來存放彙總資料，好處是查詢效率相對之下比ROLAP 較高、儲存空間較大、維護成本也低，但其缺點是必須重新開發與系統之擴充性低。
3. HOLAP：為 ROLAP 與 MOLAP 的混合體，受益於 ROALP 較大的可擴延性和 MOLAP 的快速計算，有效綜合了兩種模式的特性，換句話說，也就是在於前兩者的優、缺點之間取得一個平衡。

## 2.4 資料探勘之關聯性法則

本節主要分成兩小節，第 2.4.1 主要述說關聯性法則資料探勘技術介紹；第 2.4.2 節說明相關之關聯性法則概略比較。

### 2.4.1 關聯性法則資料探勘技術

關聯法則探勘的模型是由Agrawal[8]等學者在1993年所提出。它用來呈現資料中特定項目同時發生的關聯性，也就是說在發生項目A的情況下，同時也會發生項目B的可能性。根據產生關聯法則的步驟中依照工作程序特性，大致上可以分成兩個階段，第一個階段是從大量的資料中尋找符合支持度的高頻項目集合(Frequent Patterns)；第二的階段就是將第一個階段產生的高頻項目集合，進行信賴度驗證的工作。由於第一個階段必須利用繁複的工作去掃描資料，並且透過不斷組合、刪減、驗證來產生高頻項目集。因此第一階段成為探勘關聯法則最費時的地方，而第二個階段是利用第一階段的結果，推演出關聯法則。

關聯性法則的形式為  $X \rightarrow Y$ ，其中  $X, Y \subset I$ ，且  $X \cap Y = \Phi$ 。支持度(Support)為支援關聯規則交易(發生狀態)次數百分比。信賴度(Confidence)為發生左邊交易(狀態)中，同時又發生右邊交易(狀態)的概率，也可說是每一條規則(Rule)之強度的度量單位。而  $\text{Confidence}(X \rightarrow Y) = \text{Support}(X \cup Y, D) / \text{Support}(X, D)$ 。依照條件機率，若某關聯性法則的信賴度超過一定限度時，其意義為若此交易(狀態)包含  $X$ ，有很高的機率會包含  $Y$ 。因此，挖掘關聯性法則也就是要找出所有  $X \rightarrow Y$  形式的關聯性法則，並且滿足下列條件： $\text{Support}(X \cup Y, D) \geq \text{Min\_Support}$  且  $\text{Confidence}(X \rightarrow Y) \geq \text{Min\_Confidence}$ 。

#### 2.4.1.1 關聯性法則 Apriori 演算法

1994 年 Agrawal[8]等學者提出 Apriori 演算法。其演算法流程如圖 2.9 所示。

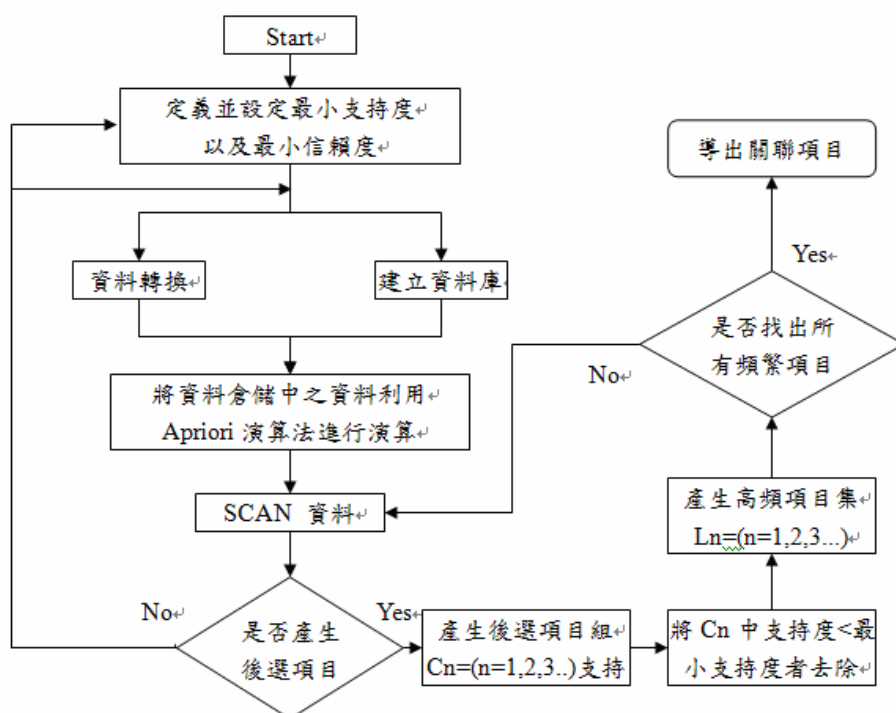


圖 2.9 Apriori 演算法流程圖[8]

上述說明關聯性法則 Apriori 進行資料採擷之演算法，然而對其演算法流程之解析，很難了解其所表達的意涵。故在此以另一種方式解說其演算步驟，其主要歸類分五大步驟，其演算步驟說明如表 2.11 所示。

表 2.12 Apriori 演算步驟說明表

步驟	執行演算方式
1	先定義最小支持度和最小信賴度
2	Apriori 演算法使用之候選項目集合(Candidate Itemset)的觀念，若候選項目集合的支持度 > 最小支持度(Minimun Support = 1)，則該候選項目集合為高頻項目集合。
3	首先由資料庫讀取所有的記錄，求出候選 C1 的支持度，再找出高頻項目集合 L1，並利用這些高頻項目集合的結合，產生候選 C2。
4	然後再 SCAN 資料倉儲資料，得出 C2 的支持度後，找出高頻 L2，並利用這些 L2 的結合，產生候選 C3。
5	重覆 SCAN 資料倉儲資料，與最小支持度比較，產生高頻項目集合，再結合產生下一級候選項目集合，直到不能結合出產生新的候選項目集合為止。

由圖2.10得知， $C_i = L_{i-1} * L_{i-1}$ ，會將 $L_{i-1}$ 中具有相同 $i-2$ 個項目的 $(i-1)$ -高頻項目集予以組合，以產生一個新的 $i$ -項目集。且" $C_i$ 中的每一項目集之任一子集合須在 $L_{i-1}$ 中出現，否則此 $C_i$ 項目集須加以刪除"。如此scan資料至無法組合出更高頻項目集時，停止搜尋。

根據表 2.12 演算法的步驟，進行相關例子推導，利用高頻項目集產生候選項目集的情形。假定資料庫中存在相關資料，其中包含 Event\_1(Item(1)、Item(2))、Event\_2(Item(3)、Item(4)、Item(5))五個項目(事件(Event)為項目之集合，如圖 2.10 左上方所示)，透過掃描資料庫可以得到項目在資料庫中的次數。設定支持度為 2，當項目發生次數大於或等於 2 時，方能成為高頻項目。詳細候選項目與高頻項目產生流程如圖 2.10 所示。演算到最後，可以產生四個一階高頻項目：Item(1)、Item(2)、Item(3)、Item(5)；四個二階高頻項目：Item(13)、Item(23)、Item(25)、Item(35)；一個三階高頻項目：Item(235)。如果要產生關聯法則，就將這些高頻項目集進行信賴度的判斷，如果符合信賴度的條件，就能夠成為關聯法則。

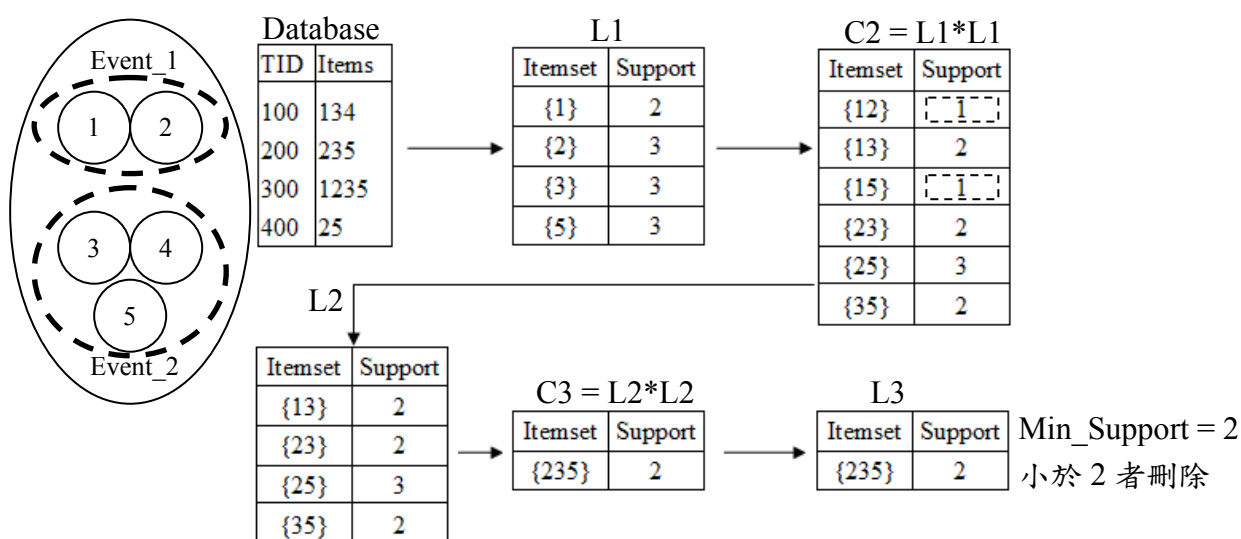


圖 2.10 Apriori 演算執行過程[21]

### 2.4.1.2 關聯性法則 DLG 演算法

Yen, S. J.等學者[26]提出DLG演算法[1]，主要在改進尋找高頻項目集的步驟，其概念是在產生高頻序列步驟中，只掃描資料庫一次並利用Bit Vector進行集合運算建立有向圖(Directed Graph)。利用有向圖指出項目間的關聯，並且追蹤有向圖產生large itemsets及 sequential patterns。DLG演算法可分為三個步驟，如表2.13所示。由表2.13得知步驟有三，但是步驟中又有小步驟，因此以圖2.11例子來說明此步驟。

表 2.13 DLG 演算步驟說明表

步驟	執行演算方式
1	產生 large 1-itemsets 並記錄相關資訊。 此步驟有下列的工作： (a) 在演算法實施之前對每一個 item 編一個整數。 (b) 掃描資料庫計算每一個 item 的 support值(出現次數)。 (c) 對每一個 large item 建立一個 bit vector。若此 item 出現在第 i 筆交易，則此 bit 設定為 1，否則為 0。Bit vector的長度為交易的數目。
2	建構一個關聯圖，指出 large items 之間的關聯。產生 large 2-itemset。
3	利用關聯圖產生 k-itemsets ( $k > 2$ )。

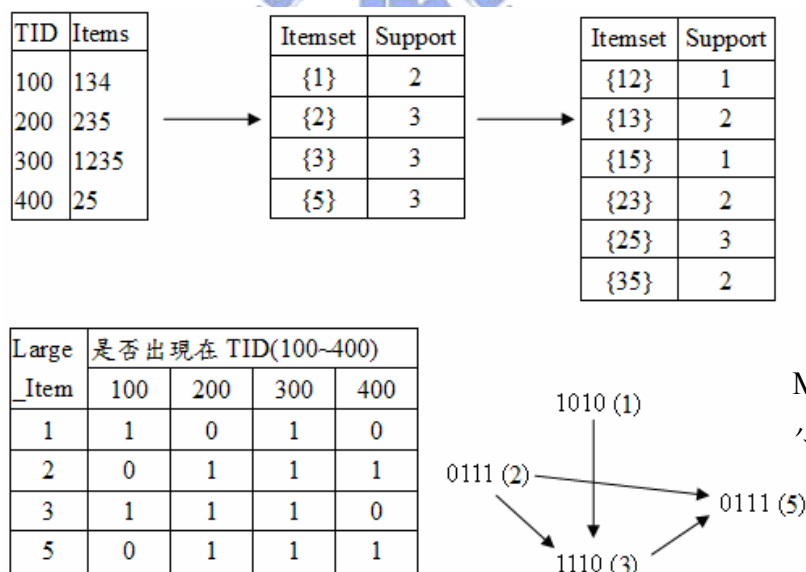


圖 2.11 DLG 演算執行過程

假定Min\_Support值等於2。Itemset  $\{i_1, i_2, \dots, i_k\}$  的Support值為  $BV_1 \cap BV_2 \cap BV_3 \cap \dots \cap BV_k$  所形成位元向量中"1"的個數，其中" $\cap$ "表示布林函數中AND集合運算。然後求出圖左下方的位元向量表格，如果想要得到itemset $\{1,2\}$ 的支持度值，其結果為  $BV_1 \cap BV_2$ ，即  $(1010) \cap (0111)$  結果為  $(0010)$ ，則支持度值為1。經過第一階段步驟之後就不再掃描資料庫，利用位元向量表來算產生圖之右下角有向圖(資料結構採linked list)。若  $BV_i \cap BV_j$  之支持度值大於Min\_Support則  $i$  至  $j$  的有向邊被建立，同時  $\{i,j\}$  是一個large 2-itemset。最後large 2-itemsets為  $\{1,3\}$ 、 $\{2,3\}$ 、 $\{2,5\}$ 、 $\{3,5\}$ 。

DLG演算法主要的優點是產生candidate itemsets容易，且僅需掃描資料庫一次。但如果項目過多會形成一個複雜的有向圖形，如此以linked list實施在空間及時間上會比array方式更加的負擔。

### 2.4.1.3 關聯性法則 RARM 演算法

Amitabha Das等人[9]提出了RARM演算法，是將原始的資料庫經由前置處理，存成一個Trie的資料結構，如圖2.12所示。這個資料結構依Support Count由大到小儲存所有的1-項目組(1-itemsets)及2-項目(2-itemsets)；之後就在這個Trie上面進行尋找高頻-1項目組(Frequent-1-itemsets)與高頻-2項目(Frequent-2-itemsets)的操作，尋找其它長度的高頻項目組，仍然需要掃描交易資料庫。經由上述解說之後，在此用一個例子來說明會更清楚，如圖2.13 RARM演算執行過程。

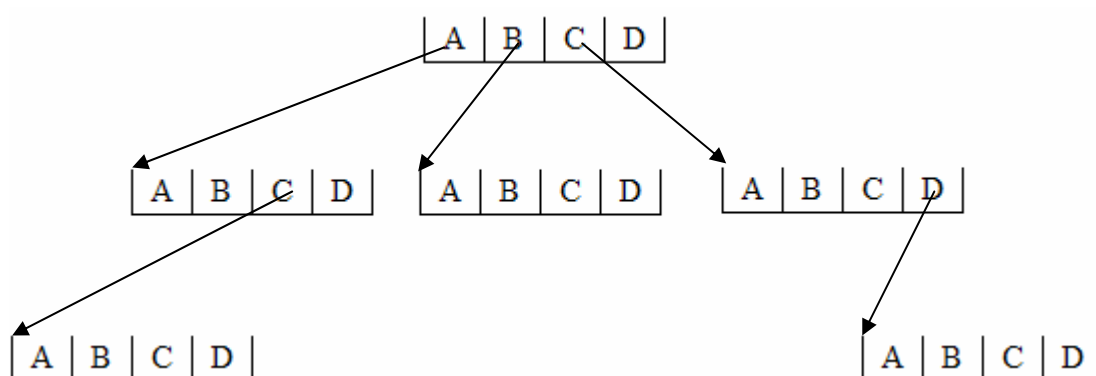


圖 2.12 Trie 資料結構

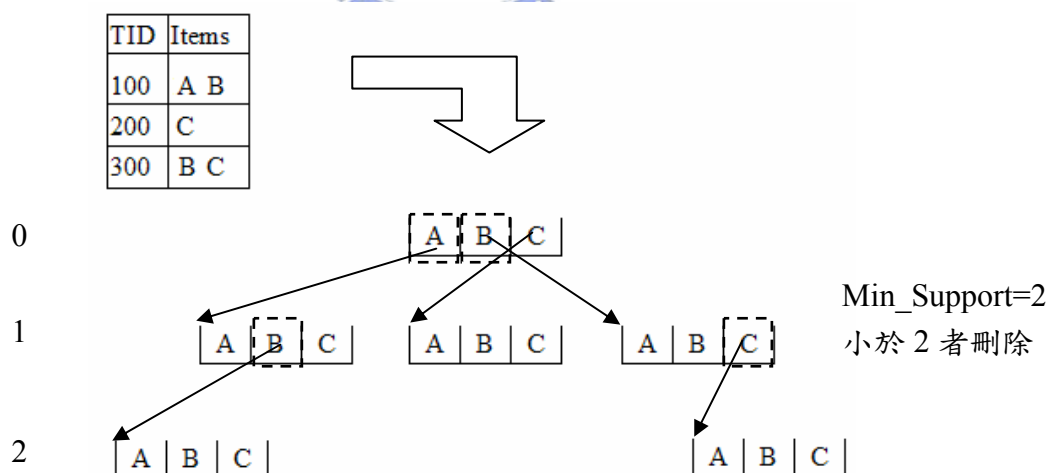


圖 2.13 RARM 演算執行過程

由資料庫進行前置處理得到圖下方Trie資料結構，之後再從Trie資料結構尋找高頻項目集。由Trie資料結構圖可以清楚得知，高頻-1項目組為B與C並無高頻-2項目組。RARM演算法提出一個前置處理的機制，但是若遇到使用者想要尋找高頻3-項目組(Frequent 3-itemsets)或更長的高頻項目組時，還是必須去重新掃描資料庫。且為結構中只儲存1-項目組與2-項目組，無法將完整的交易資料載入。

#### 2.4.1.4 關聯性法則 DHP 演算法

Park, J. S.[24]於1995年提出DHP(Direct Hashing and Pruning)演算法，在候選-2項目集合事先使用雜湊函數(Hash function)過濾部分不可能成為高頻項目集合的候選項目集合，以減少比對的次數。其步驟有四，如表2.14所示。

表 2.14 DHP 演算步驟說明表

步驟	執行演算方式
1	將每個交易下的項目使用與Apriori高頻項目集合產生候選項目集合相同的方式建立Hash table。
2	針對資料庫下的所有項目建立對映編號，如 1、2、3
3	使用雜湊函數計算Hash table下所有的集合，將所得出來的直放入相對映的Bucket。 p.s. 雜湊函數為 $h\{x,y\} = ((\text{order of } x) * 10 + (\text{order of } y)) \bmod 7$ 。
4	使用Apriori方法找出第2層的候選項目集合。

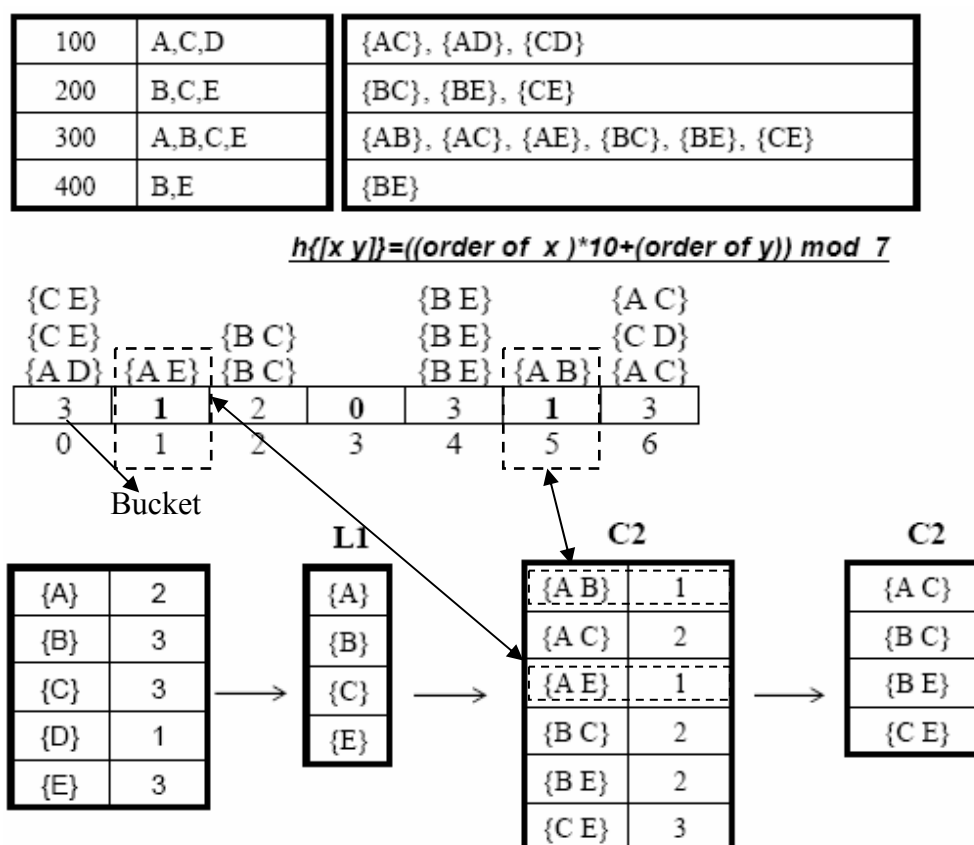


圖 2.14 DHP 演算執行過程

如圖2.14的資料庫含有4筆紀錄，有A、B、C、D、E等5個不同的項目，最小支持度設定為50%。也就是說候選項目集合的次數要超過2(含)次以上才能成為高頻項目集合，採用的雜湊函數為 $h\{x,y\} = ((\text{order of } x) * 10 + (\text{order of } y)) \bmod 7$ 。

步驟一首先將每個交易下的項目使用與Apriori高頻項目集合產生候選項目集合相同的方式建立Hash table。因此Tid = 100下的項目有A、C、D結合之後就會產生{AC}、{AD}、

{CD}等3個集合，Tid = 200下的項目有B、C、E 結合之後就會產生{BC}、{BE}、{CE}等3個集合，Tid = 300下的項目有A、B、C、E結合之後就會產生{AB}、{AC}、{AE}、{BC}、{BE}、{CE}等6個集合，Tid = 400下的項目有B、E結合之後就會產生{BE}等1個集合。

步驟二對資料庫下的所有項目建立對映編號，因此A、B、C、D、E建立的對映編號為1、2、3、4、5。與步驟三使用雜湊函數計算Hash table下所有的集合，將所求得出來的值放入相對映的Bucket列。如集合{AB}，則對映到的編號為{1,2}，雜湊函數 $H(1,2) = (1*10+2) \bmod 7 = 5$ (餘數)，故{AB}是放入第5 bucket；集合{AC}，則對映到的編號為{1,3}，雜湊函數 $H(1,3) = (1*10 + 3) \bmod 7 = 6$ (餘數)，所以{AC}是放入第6 bucket，依序以上的方法得知，第0~6 bucket分別放入3、1、2、0、3、1、3個物件集合。

步驟四使用Apriori方法找出第2層的候選項目集合分別為{AB}、{AC}、{AE}、{BC}、{BE}、{CE}等6個集合，經雜湊函數計算分別對映到的bucket位址為5、6、1、2、4、0。其各bucket的位址所含的集合個數分別為1、2、1、2、2、3，經過最低支持度檢測只有{AB}、{BC}、{BE}、{CE}，成為第2層的候選項目集合C2，{AB}、{AE}則成功的被雜湊函數事先去除。

DHP演算法雖然利用簡單的雜湊函數減少了第2層的候選項目集合，但是其雜湊函數會有碰撞的情形。如果碰撞太厲害，其執行效率會比Apriori更差。

#### 2.4.1.5 關聯性法則 VIPER 演算法

Pradeep Shenoy 等人[25]提出 VIPER 演算法，主要是提出一套壓縮的方式，希望能減少資料儲存的空間。但是 VIPER 演算法在產生高頻項目組時，每次均要將前一個高頻項目組解壓縮，才能產生候選項目組，並將結果再以壓縮的方式儲存在磁碟中。這一壓一解的動作，降低了執行的效率。

#### 2.4.2 關聯性法則演算法比較

經由2.4.1節對關聯性法則演算法Apriori、DLG、RARM、DHP、VIPER五種的說明，了解到各種演算法有其不一樣的優缺點。如表2.14所示。

表 2.14 關聯性法則之演算法優缺點比較表

演算法種類	優點	缺點
Apriori	可以找出高頻項目集 3 以上	反覆性高
DLG	使用 Bit-Vector 來紀錄交易資料	花費相當大的記憶體空間
RARM	存成一個 Trie 的資料結構	尋找高頻 3-項目組難度高
DHP	利用雜湊函數減少了第 2 層候選項目集合	耗用大量的記憶體
VIPER	以壓縮的方式，希望能減少資料儲存的空間	需要高頻率的一壓一解，效率不佳

由表2.14可以得知，每種演算法之優缺點。在此本研究將選用以Apriori演算法為基礎的演算軟體(附錄E)來進行分析，雖然反覆性高，但是有著可以進行高頻項目集3以上的尋找效率。VIPER演算法如果碰到需要產生高頻項目組時，就會進行一次壓縮。經由一次再一次的壓縮，必定帶來電腦沉重的負擔，嚴重的話可能導致當機的情況發生。RARM演算法碰到尋求高頻項目組也是如此。DLG演算法本身就需耗費相當大的記憶體空間，因此在

運算時當機產生的次數一定高於前述兩種演算法。最後，DHP為了要能過濾大量的候選項目集合必須使用更多的bucket才能達到目的，但是會耗用大量的記憶體且每增加的bucket數目愈多，其所過濾候選項目集合的效能成長就愈少。此時執行因執行過程中要比Apriori多一個計算雜湊函數的負擔，其執行的效率反而比Apriori還要差。



### 第三章 台灣區花卉資料倉儲資料整合實作

由第 1.1 節可知目前總資料倉儲系統的建立，已整合了台北、彰化、台中、台南四家花市的資料，其最新於 2006 年完成的高雄資料超市系統仍然為獨立的個體。雖然其他四家花市資料超市的使用者都能夠藉由總資料倉儲系統來互相擷取並分享資訊，但對整條花卉供應鏈來說，整個台灣區的每一個花卉市場都是重要不可或缺的個體。因此高雄資料超市系統的建置需要再次加入先前所整合好的四家花卉供應鏈，進行完成最後整合工作，也是當前首要的任務。此一任務能為整個台灣地區花卉供應鏈補足缺了一角的資訊分享管道，也帶來其他四家花市正面與負面的連鎖效應。

本系統架構重在整合性。將高雄資料超市加入先前所整合好的台北、彰化、台中、台南四個資料超市並完成最終整合。本章將說明台灣區花卉資料倉儲的系統架構組織。本章共分為兩小節，第 3.1 節細部說明系統與資料倉儲的架構。第 3.2 節說明如何整合五家資料超市的檔案。

#### 3.1 系統架構與資料倉儲架構

資料倉儲與系統機能要求兩者為花卉批發資訊分享熱線的核心技術與系統穩健度的保證。本節將介紹花卉總資料倉儲建立的系統基本架構，與資料倉儲建構模式。第 3.1.1 節介紹系統建置在 IBM 主機上的方式；第 3.1.2 節介紹建置資料倉儲所採用的架構模式。

##### 3.1.1 系統架構

建置完成總資料倉儲系統，在於硬體層面上，將資料處理部份與檔案發佈部份分別由兩台主要 IBM e-server 來進行處理。首先經由負責資料傳輸與轉換的 e-server 1-2 主機，接著再由 e-server 1-1 負責權限控管與檔案發佈網路。再者，對於台北、彰化、台中、台南、高雄各家資料超市完整硬體設施建置如圖 3.1 所示，依序由上而下分別代表第一台 e-server 2-1 為存放台北資料超市之主機、第二台 e-server 2-2 為存放彰化資料超市之主機、第三台 e-server 2-3 為存放台中與台南資料超市之主機以及最後建置完成之第四台 e-server 2-4 為存放高雄資料超市之主機。然而最右邊名為 e-server 1-1 主機主要負責將總資料倉儲的檔案進一步的發佈於網路上，最後中間名為 e server 1-1 的主機，則是存放總資料倉儲的資料庫檔案\*.mdf。表 3.1 將予以各主機之命名及其他相關屬性做詳細的整理與概要介紹。

總倉儲實體系統架構如圖 3.1 所示，每台 IBM 主機皆以 Hub 的方式接到防火牆(如圖 3.1 實線部分)，其中 e-server 1-2 透過 SQL Server 建立起遠端伺服器連線機制。SQL Server 可以透過防火牆內部網路，直接讀取 Domain Name 來達到對應其他各伺服器資料的存取依據。如圖 3.2 所示，SQL Server 的群組底下在於先前已建立完成四個伺服器連線，加上新設立的伺服器連線，共五大伺服器連線。FII 為存放總資料倉儲之本機伺服器，FLOWERDB2 為彰化資料超市之伺服器，FLOWERDB3 為台中與台南資料超市共用之伺服器，SCFS01 為台北資料超市之伺服器，KSFLOWERDB 為高雄資料超市之伺服器。e-server 1-2 透過遠端伺服器連線(如圖 3.1 虛線部分)，進行每日自動化資料轉換服務，並將五家資料超市的資料自動轉換至總資料倉儲中。此部分連線需透過資料轉換服務進行，如圖 3.3 分別為左上角台北拍賣 DTS1-1、左下角台中拍賣 DTS1-2、右上角彰化拍賣 DTS1-3、

右下角台南拍賣 DTS1-4 之連線畫面與設定。加上圖 3.4 高雄拍賣 DTS1-5 之連線畫面與設定，然後當 e-server1-2 透過 cognos 套裝軟體建立報表檔案\*.pyi 與超方體檔案\*.mdc，將超方體檔案透過防火牆通過網路連線時，複製超方體檔案存放至 e server 1-1，此部分運作是由 e-server1-2 透過網路磁碟機的設定完成連線。

如圖 3.5 所示網路磁碟機 F 槽，Fdatabase 為 e-server 1-1 所屬的 Domain Name；在自動化更新轉換程序當中，需要將特定語法設定在自動化 Macro(DTS\_5)，圖 3.6 為複製檔案至網路磁碟機之語法畫面；e-server 1-1 為存放權限控管相關資料之檔案，以及將 e server 1-2 所產生之超方體檔案，透過 cognos 的套裝軟體 PPES(PowerPlay Enterprise Server)，將超方體即時發佈至網路上，讓使用者可以隨時透過網路並藉由網頁進行一連串之線上分析處理。

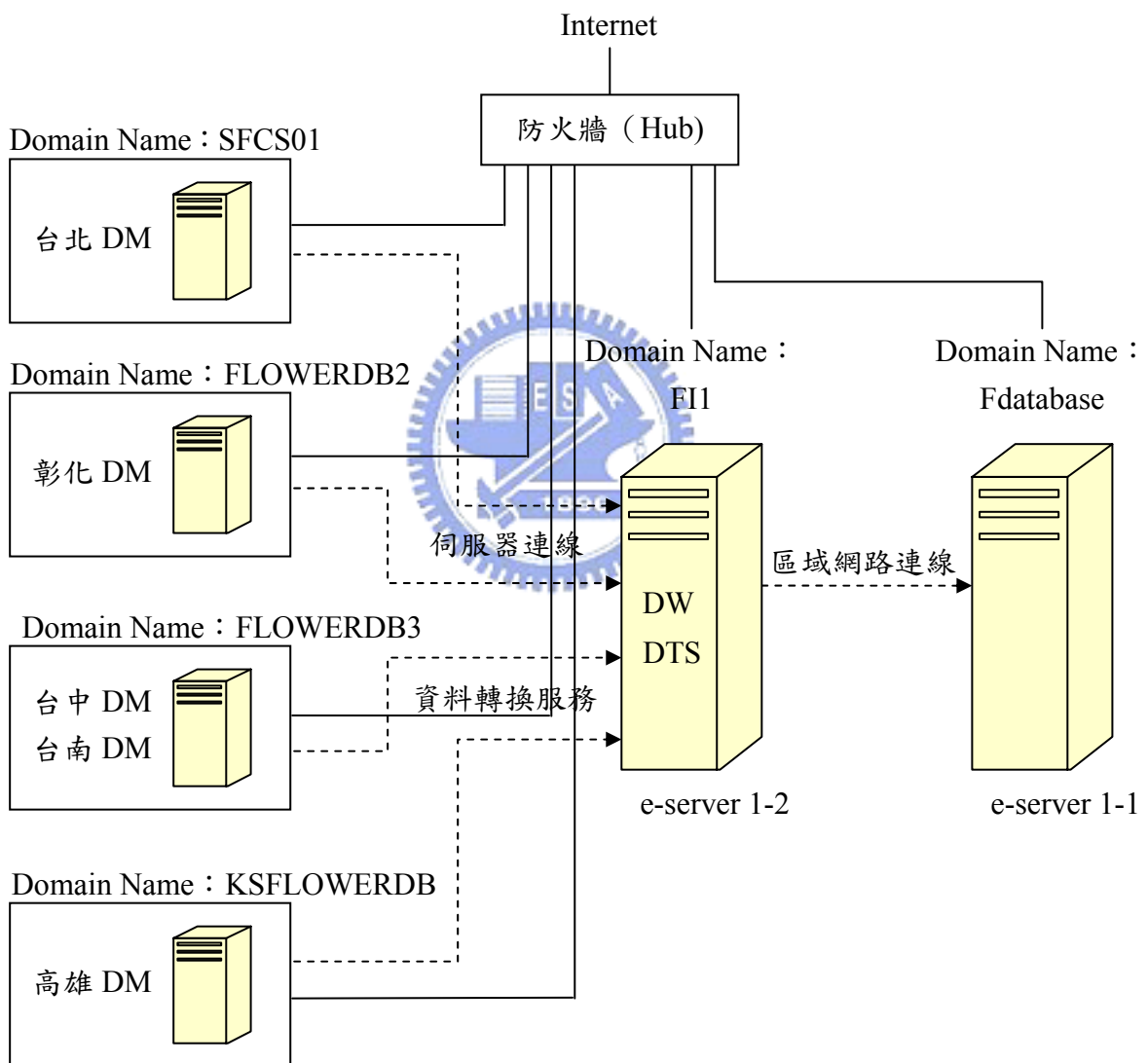


圖 3.1 總倉儲實體系統架構圖

然而，一個系統網路安全機制是有某種重要性存在的，因此，在選擇防火牆軟硬體方面，需要特別謹慎並多做比較之。在此系統中所使用 NetScreen 防火牆系列，主要可分為兩個層面：

#### 1. 企業內部網路(Intranet)

NetScreen 的出現改寫了防火牆的效能標準，傳輸速率則破紀錄地達到 100 Mbps 全線速。並且它所採用的是具革命性的『軟硬體架構』所設計的 Netscreen 防火牆，其設計目的在讓各企業能利用最為經濟的方式來擁有最安全、最容易使用與最高效能的新一代防火牆，消除一般傳統式軟體防火牆所造成的傳輸瓶頸。

NetScreen 不僅能在您連接網際網路時，保護公司內部網路的安全，也可做為公司內部不同子網路(Individual Subnetwork)之間的網路安全防護設備。NetScreen 可以在您的區域網路中，保護重要的網路設備或資源，確保各部門中機密資料或重要內部網路伺服器的安全，甚或保衛總裁(CEO)及高層主管個人電腦中敏感的機密資訊。

## 2. 企業外部網路(Extranet)

NetScreen 的虛擬私人網路(Virtual Private Network)功能，允許您利用網際網路上的安全隧道來節省公司的成本。藉著 NetScreen 強而有力的處理能力，VPN 功能讓企業可以安全地『複製遠端資料庫』、執行遠端『備份』、及『管理配置伺服器』的能力。NetScreen 的 VPN 功能，使得一個要求安全及高性能網路環境的企業體，能藉著與供應商、合作夥伴及分公司間網際網路，來達到符合成本效益的端對端通訊目的。最後進一步來看，VPN 除了能節省租用昂貴的私有資料通訊線路，也能提供安全的企業網路，將企業的支出做最佳的安排。

綜合以上所敘述之兩大層面，其所選用的防火牆，並非使用軟體防火牆，也非硬體防火牆，而是結合硬體與軟體防火牆之軟硬體防火牆。它不但有著軟體防火牆之使用彈性(企業可以將軟體防火牆安裝在現有閘道伺服器(gateway)上，而不需要再變更網路架構)與擴充性(依照目前的安全需求來選擇合適的產品，即使日後需要增加功能，也僅需隨之更新軟體即可)。而且也具備硬體防火牆之高效能(硬體與作業系統最佳化之組合)與容易安裝的特性。故選擇它作為花卉計畫使用之軟硬體防火牆，其簡述功能如下所示：

1. 每一個 trust 埠之使用人數最高可達 10 人 (無授權使用人數限制)。
2. 整體防火牆處理效能應可於滿載時達到各介面 75Mbps 的吞吐量或總處理能力 Throughput。
3. 發生告警或攻擊事件時，可以自動利用電子郵件(E-Mail)通知系統管理者。
4. 提供 5 埠 10/100Mbps 之標準高速乙太網路介面。
5. 提供 2000 個(含)以上之同時連線數(Concurrent Sessions)及每秒 2000 新增 Session 的處理能力。
6. 最少可提供 100 個(含)以上之存取過濾條件設定能力。
7. 支援網路位址轉譯功能(Network Address Translation)及連接埠位址轉譯功能(Port Address Translation)，並且支援即時告警與記錄(Real time alarming & Logging)等功能。
8. 支援動態路由 BGP(Border Gateway Protocol)/OSPF(Open shortest Path First)標準，也就是說邊界閘道協定與先開啟最短的路徑兩者為不同的路由協定標準。

表 3.1 各主機項目命名表

硬體代號 (e-server)	主機型號	IP 位置	Domain Name 及資料庫名稱	功能
1-1	IBM e-server X series 236	140.113.58.186	Fdatabase	提供線上存取服務
1-2	IBM e-server X series 236	140.113.58.185	FI1 Flower2	總資料倉儲
2-1	IBM e-server X series 235	140.113.58.162	SFCS01 TPCDB	台北資料超市
2-2	IBM e-server X series 235	140.113.58.180	FLOWERDB2 TenWay1	彰化資料超市
2-3	IBM e-server X series 235	140.113.58.181	FLOWERDB3 flowers、flower4	台中與台南資料超市
2-4	IBM e-server X series 235	140.113.58.187	KSFLOWERDB Kaoshiung	高雄資料超市

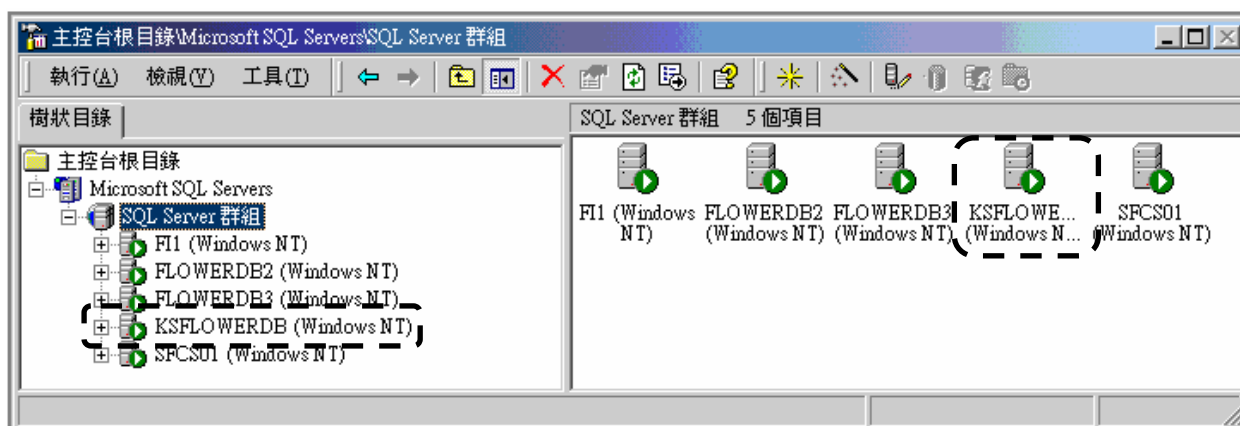


圖 3.2 SQL Server 連線群組(a)



圖 3.3 SQL Server 連線 4 群組(b)

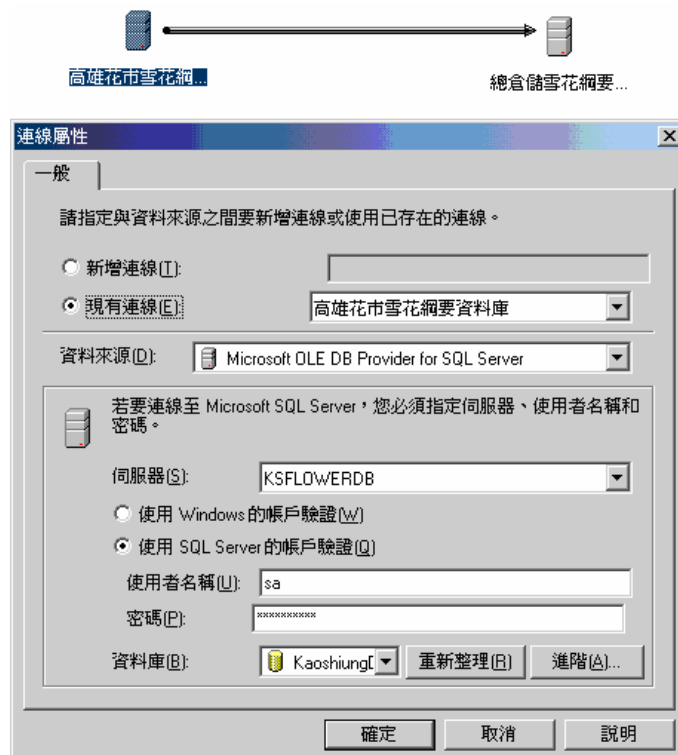


圖 3.4 SQL Server 連線群組(c)



圖 3.5 網路磁碟機畫面



圖 3.6 檔案複製到網路磁碟機語法畫面

從圖 3.1 中，由五個資料超市與第二台主機之間的所有資料轉換流程，皆需透過每日進行的自動化資料轉換服務來完成。在實際時間運作上可參照圖 3.7 資料轉換時刻圖，每日於五家資料超市個別執行完資料轉換服務後，方能於 e-server 1-2 進行總資料倉儲的自動化資料轉換服務。表 3.2 將負責總資料倉儲的兩台 IBM 主機，分別以存放檔案以及執行服務兩方面做出整理。

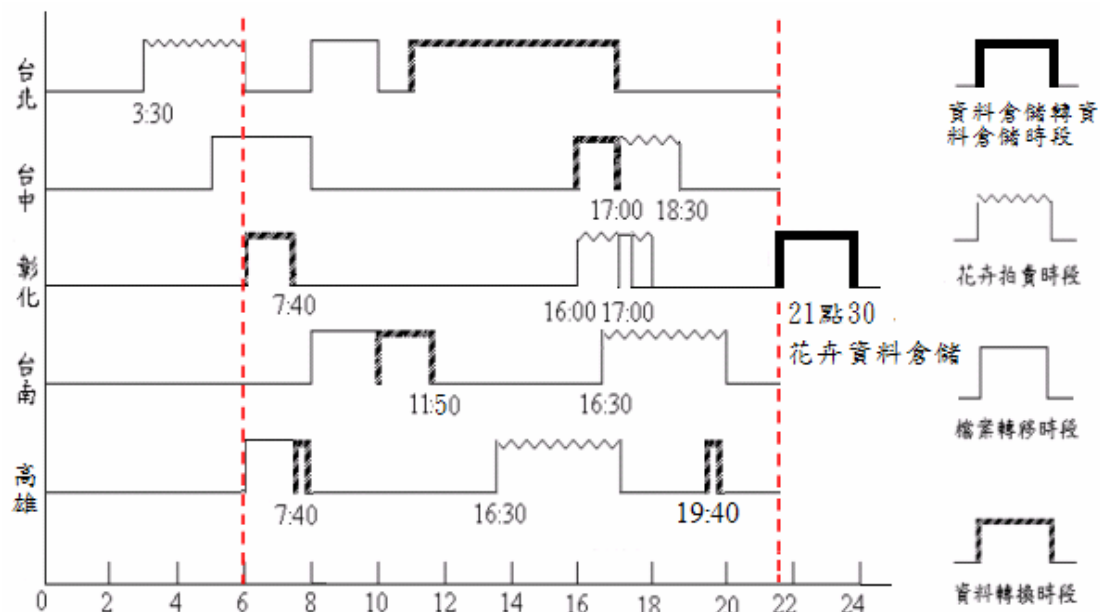


圖 3.7 資料轉換時刻圖

表 3.2 總倉儲主機負責服務整理表

主機名稱	IBM e-server			
	第一台		第二台	
存放檔案	檔案	檔案類型	檔案	檔案類型
	<ul style="list-style-type: none"> <li>資料庫檔案</li> <li>DTS 檔案</li> <li>報表檔案</li> <li>超方體檔案</li> </ul>	mdf dts pyi mdc	<ul style="list-style-type: none"> <li>權限控管檔案</li> <li>超方體檔案</li> </ul>	lae mdc
執行服務	總資料倉儲之自動化轉換服務 <ul style="list-style-type: none"> <li>進行資料轉換服務</li> <li>建立超方體</li> <li>將超方體透過網路磁碟機複製於第二台主機</li> </ul>		提供線上分析處理服務 <ul style="list-style-type: none"> <li>權限控管運作</li> <li>檔案發佈至網路上</li> </ul>	

### 3.1.2 資料倉儲架構

以台灣區花卉資料倉儲而言，採取競簇教授所提出之資料倉儲建置架構，由台北、台中、彰化、台南和高雄五家資料超市由下而上的模式建立而成。其基本架構如圖 3.8 台灣區花卉資料倉儲競簇架構圖所示。其中台北、台中、彰化、台南花卉資料超市已於之前完成建置並穩定運作中，也整合過一次，故本系統架構將對前四家資料超市不加以贅述。

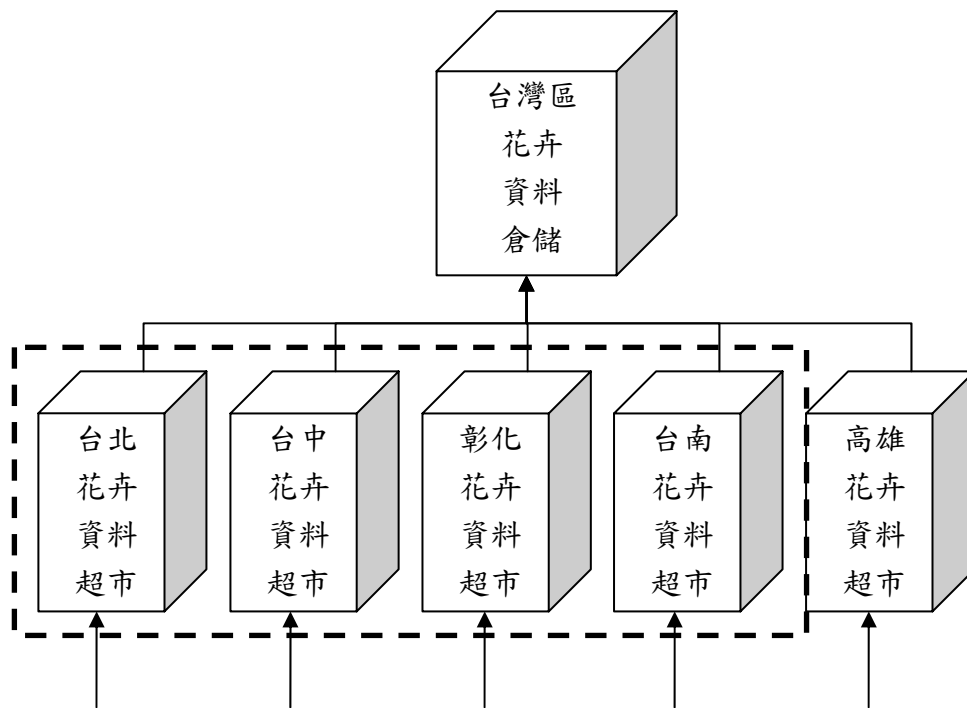


圖 3.8 台灣區花卉資料倉儲競渡架構圖[5]

整體資料倉儲架構如圖 3.9 花卉批發資訊分享熱線架構圖所示，本章將著重在右邊粗虛線框內範圍的資料轉換服務，分別從各家花市進行資料轉換服務，彙整成總資料倉儲。

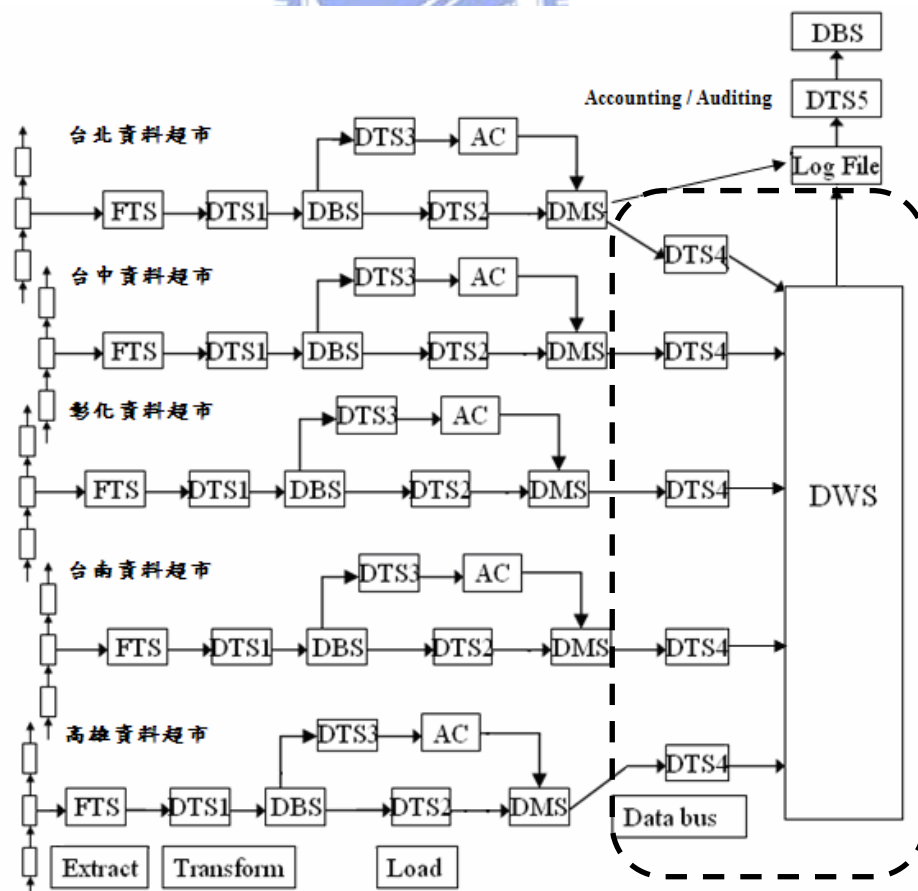


圖 3.9 花卉批發資訊分享熱線架構圖[7]

## 3.2 總資料倉儲整合工作

總資料倉儲的設計關鍵要點，主要是希望能同時顯示出台北、台中、彰化、台南及高雄五家資料超市的資訊。但是基於此五家超市各為不同設計者所設計，且在於各家資料超市的資料內容存在些許差異。因此首要工作就是將相關欄位名稱以及資料表內容都先行一步做整合，如此一來才能使總資料倉儲的資料轉換工作順利地繼續進行。本節共分兩小節，第 3.2.1 節介紹總資料倉儲之設計架構；第 3.2.2 節說明資料整合方式及其結果。

### 3.2.1 總資料倉儲架構

資訊共享是一種達到雙贏的理念，但是並非所有資訊皆能分享。總資料倉儲為了避免一些私密性的資料外洩流通在外，以及基於資料傳輸速度考量上和眾多使用者訴求，必定要刪減不必要的維度。與先前各家資料超市互相比較，總資料倉儲明顯多了特殊節慶以及市場別的維度，而在供應單位中之所以減少供應人層級部份，目的是為了避免個人資料遭致外洩的可能，表 3.3 為總倉儲與各家資料超市維度架構之比較。

表 3.3 總倉儲與各家資料超市維度比較表

維度	總倉儲	各家資料超市
供應單位	無供應人層級	有供應人層級
節慶	有	無
市場別	有	無
供應地區	有	有
日期	有	有
星期	有	有
農曆	有	有
拍賣線	有	有
花卉種類	有	有
花卉等級	有	有
容器	有	有
承銷地區	有	有
判斷	12 有	2 無 10 有
統計	共 12 個維度	共 10 個維度

### 3.2.2 資料整合工作

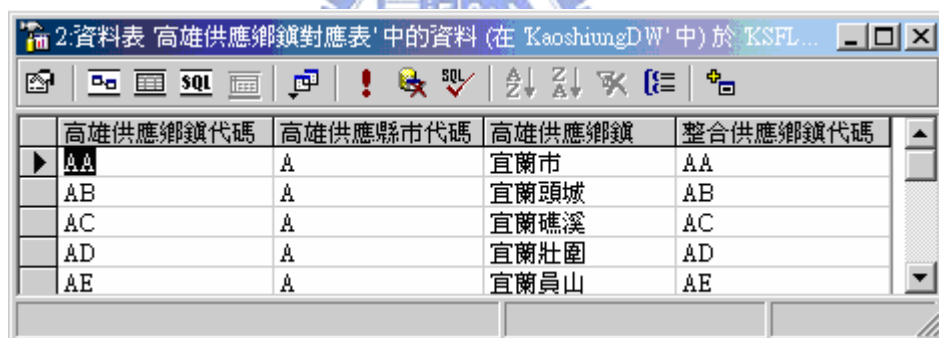
當進行資料轉換服務的前置工作時，必須先將台北、台中、彰化、台南、高雄五家資料超市的資料表做進一步整合，以制訂出共同標準。此外，除了日期、星期、農曆、花卉等級四個維度的資料表所屬欄位內容不需加以整合外，以及新增的節慶、市場別維度以外，其他維度皆需要採取聯集方式進行統整。

表 3.4 整合對照表為整合各家資料超市的資料表後的情形，每個欄位所具代表意義各不相同，如欄位一表示維度類別，然而在最後一列另外加入 Fact Table，並非維度之意；欄位二內容為五家資料超市內的原始資料表，基於方便了解之考量，故以統一代稱命名，不加贅述；欄位三為總資料倉儲經過聯集整合後的資料所新建立出的資料表；欄位四為因應整合需求所新建在各家資料超市的轉換資料表如圖 3.10 高雄供應類別對應表、圖 3.11 高雄供應鄉鎮對應表、圖 3.12 高雄承銷地區對應表，以高雄資料超市為例，之所以新增了高雄供應類別對應表，目的是要做為總資料倉儲內供應單位與其對應的高雄資料超市內供應類別的對照表，以便有效地進行資料轉換的動作。



高雄供應類別代碼	高雄供應類別	整合供應類別代碼
0	不知為何	6
1	合作社	2
2	供應團體所屬花農	10
3	個人	1
4	進口商	7

圖 3.10 供應類別對應表



高雄供應鄉鎮代碼	高雄供應縣市代碼	高雄供應鄉鎮	整合供應鄉鎮代碼
AA	A	宜蘭市	AA
AB	A	宜蘭頭城	AB
AC	A	宜蘭礁溪	AC
AD	A	宜蘭壯圍	AD
AE	A	宜蘭員山	AE

圖 3.11 供應鄉鎮對應表



承銷地區代碼	承銷地區	整合對應碼
1	台北市	1
12	基隆市	2
13	台北縣	2
14	新竹縣市	2
15	桃園縣市	2

圖 3.12 供應地區對應表

表 3.4 整合對照表

維度類別	各花市初始資料表	總倉儲最終整合資料表	各花市轉換資料表
供應單位	供應團體資料表、供應類別	供應單位	供應類別對應表
供應地區	供應行政區資料表、供應縣市資料表、供應人資料表	供應縣市、供應鄉鎮	供應鄉鎮對應表
日期			
星期	星期表	星期	
農曆	農曆對照表、農曆年、月	農曆對照表、農曆年、月	
節慶		節慶範圍資料表、節慶類別資料表	
市場別		市場別	
拍賣線	拍賣線資料表	拍賣線	
花卉種類	花卉資料表、花卉品名資料表、花卉種類	花卉種類、花卉品類、花卉品名	
花卉等級	花卉等級略表、花卉等級資料表	花卉等級、花卉等級二碼	
容器	容器資料表	容器	
承銷地區	承銷人與承銷區域資料表	承銷地區	承銷地區對應表
事實表	拍賣資料表	總倉儲交易資料表	

## 第四章 台灣區花卉資料倉儲資料轉換實作

第三章已整合了台北、彰化、台中、台南、高雄五家花市的資料，接著將實作整合總倉儲與五家資料超市之資料轉換部分。讓五家資料超市的使用者皆能夠藉由總資料倉儲系統來互相擷取並分享資訊。本章共分為兩小節，第 4.1 節介紹資料轉換服務的設計。第 4.2 節說明如何設計資料倉儲資料轉換自動化機制。

### 4.1 資料轉換服務

資料轉換服務建構在資料轉換模組上，將交易程序依序分類為拍賣、訂貨、殘貨與議價四種，分別從台北、台中、彰化、台南和高雄五個雪花網要資料庫，轉換至總倉儲的雪花網要資料庫。本節將說明此模組的資料轉換步驟；另一方面，由於前四家資料轉換模組已於先前建立完成，故在此不加贅述。

完整的資料轉換模組如圖 4.1 自動化更新轉換程序所示，主要是由五個資料轉換程式所組成。其中拍賣、訂貨、殘貨、議價資料轉換工作為自動化更新轉換程序下的四個封包，再加上自動化 Macro 的運作，也就組成了自動化更新轉換程序。透過資料轉換服務過程，可以將五家資料庫的交易資料予以三階正規化，並自動轉入至資料倉儲。本節將自動化更新轉換程序編號為 DTS\_0，拍賣資料轉換工作編號為 DTS\_1，訂貨資料轉換工作編號為 DTS\_2，殘貨資料轉換工作編號為 DTS\_3，議價資料轉換工作編號為 DTS\_4 與自動化 Macro 編號為 DTS\_5。如表 4.1 總倉儲資料轉換程式所示，此外，表 4.1 還依序列出此四個封包之下的其他子資料轉換服務程式，在此將給予 DTS 子編號使之清楚識別。

本節分為兩小節，第 4.1.2 節介紹實體資料如何轉換至多維度模型；第 4.2.2 節說明資料轉換服務應如何建置。

表 4.1 總倉儲資料轉換程式

DTS 編號	DTS_1 (拍賣)	DTS_2 (訂貨)	DTS_3 (殘貨)	DTS_4 (議價)	DTS_5 (自動化 Macro)
清空總倉儲 交易資料表	DTS_1-0				
	子資料轉換服務程式				
台北	DTS_1-1	DTS_2-1	DTS_3-1	DTS_4-1	
台中	DTS_1-2	DTS_2-2	DTS_3-2		
彰化	DTS_1-3	DTS_2-3	DTS_3-3		
台南	DTS_1-4	DTS_2-4	DTS_3-4		
高雄	DTS_1-5	DTS_2-5	DTS_3-5		



圖 4.1 自動化更新轉換程序

#### 4.1.1 實體關係模式至多維度模型轉換

總倉儲的實體關係模式至多維度模型轉換方式，見圖 4.2 實體關係至多維度模型架構圖完整呈現出本節的模型架構。虛線框住部分為前者所完成的一部分，加上新建立好的高雄部分，把已存在的四家花市拍賣資料表加上新建立好的高雄拍賣資料表，經由內部轉換過程，依序地將資料轉入總倉儲的總倉儲交易資料表。另一方面，再加上先前建好的整合維度資料表，便構成了總倉儲之雪花網要關聯圖。有了此架構後，再運作其資料轉換工作的服務，如此便可完整地呈現出完成之後的總倉儲資料轉換。

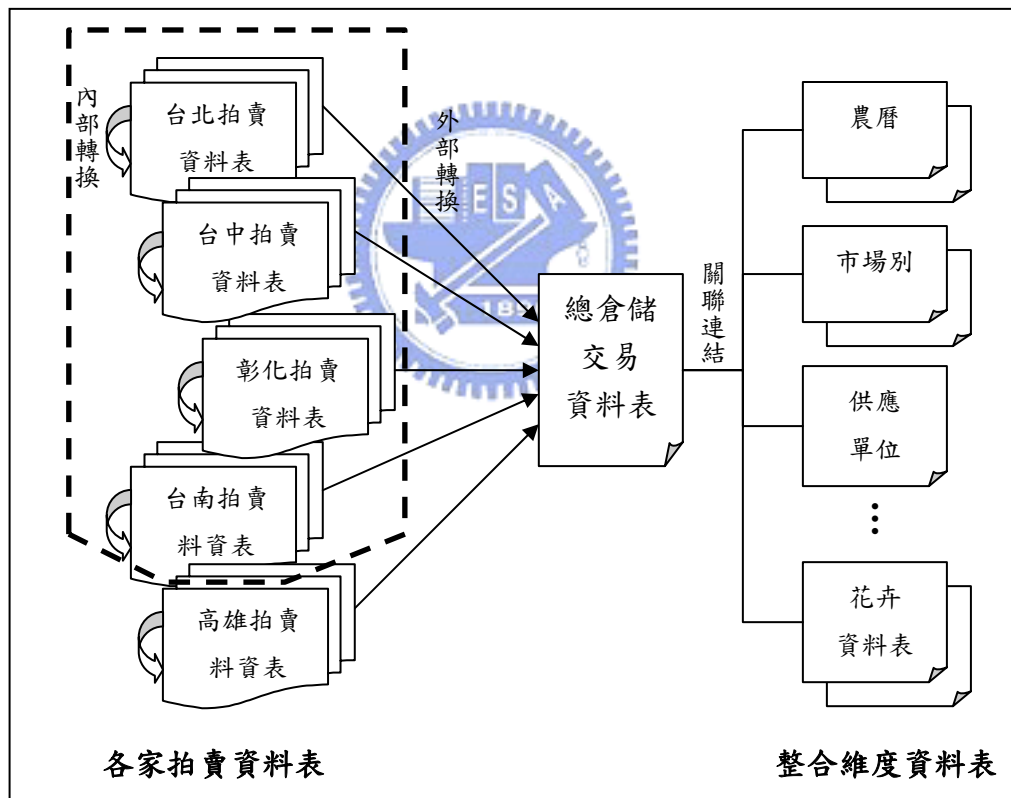


圖 4.2 實體關係至多維度模型架構

此將以高雄資料超市為例，首先必須在高雄資料超市的資料庫內，新建三個整合用的資料表，分別為高雄承銷地區對應表、高雄供應鄉鎮對應表、高雄供應類別對應表。此三張資料表主要為與承銷人資料表、供應人資料表做資料整合之用，如圖 3.16 高雄資料超市內部轉換關係圖右上方虛線框住部分。從拍賣資料表的供應人代號欄位 AUC\_SUPPLYNUM，對應到供應人資料表的供應人代號 SUP\_NUM；供應人資料表的鄉鎮代碼 SUP\_LOCATION，對應到高雄供應鄉鎮對應表的高雄供應鄉鎮代碼；供應人資料

表的供應團體代碼 SUP\_GROUP，對應到供應團體資料表的供應團體代碼 SUPGROUP\_GROUPNUM；供應團體資料表的供應團體代碼 SUPGROUP\_GROUPNUM，再對應至高雄供應類別對應表的高雄供應類別代碼。圖 4.3 左下方虛線框住部分，從拍賣資料表的承銷人代號欄位 AUC\_BUYNUM，對應到承銷人資料表的承銷人代碼 BUY\_BUYNUM，再由承銷人資料表的承銷地區代碼 BUY\_AREANUM，對應到高雄承銷地區對應表的承銷地區代碼。

高雄資料超市的拍賣資料表，經資料轉換工作在內部轉換資料完成後，資料則會加入至總倉儲交易資料表中，如圖 4.4 拍賣資料表與總倉儲交易資料表實體關聯圖所示。在黑線連結對應部分表示欄位資料轉換形式，如此便完成了實體連結。經由五家花市資料表的轉換動作後，便可轉為雪花網要架構圖，如圖 4.5 總資料倉儲雪花網要關聯圖。

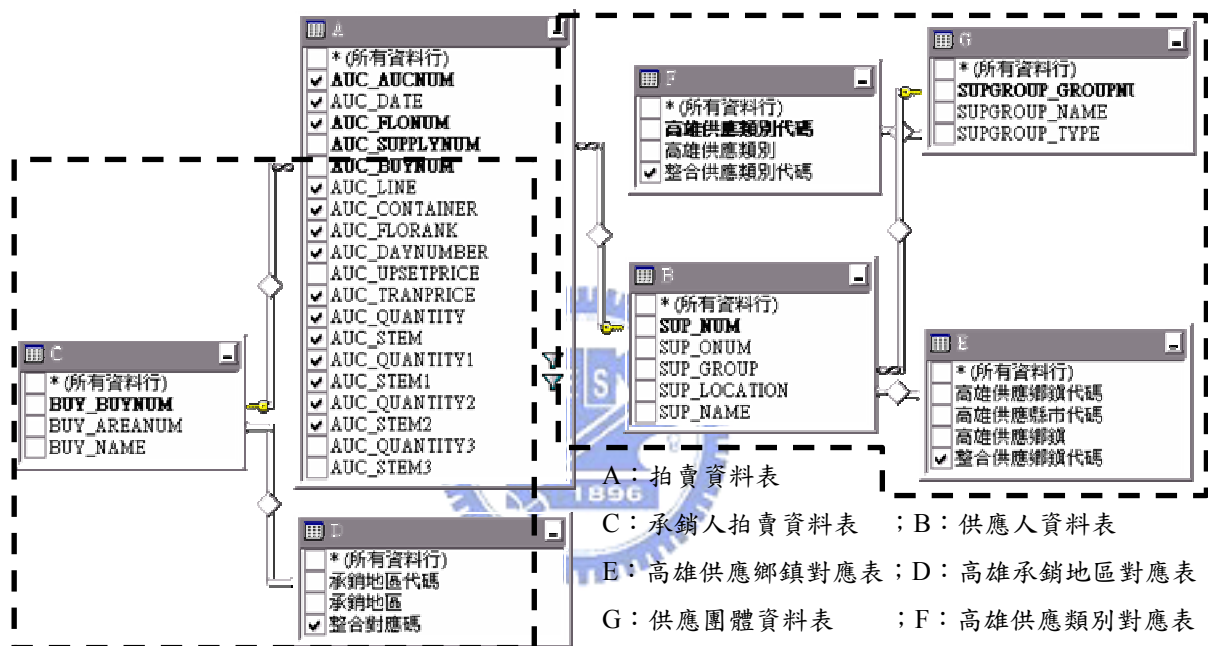


圖 4.3 高雄資料超市內部轉換關係圖

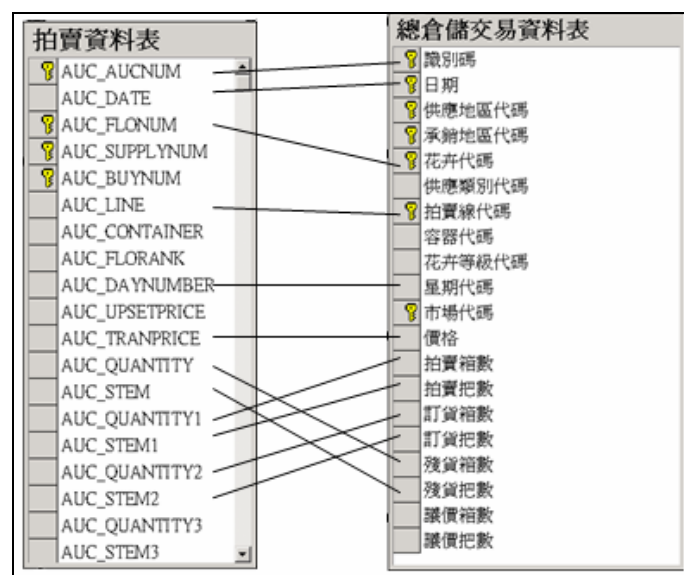


圖 4.4 拍賣資料表與總倉儲交易資料表實體關聯圖

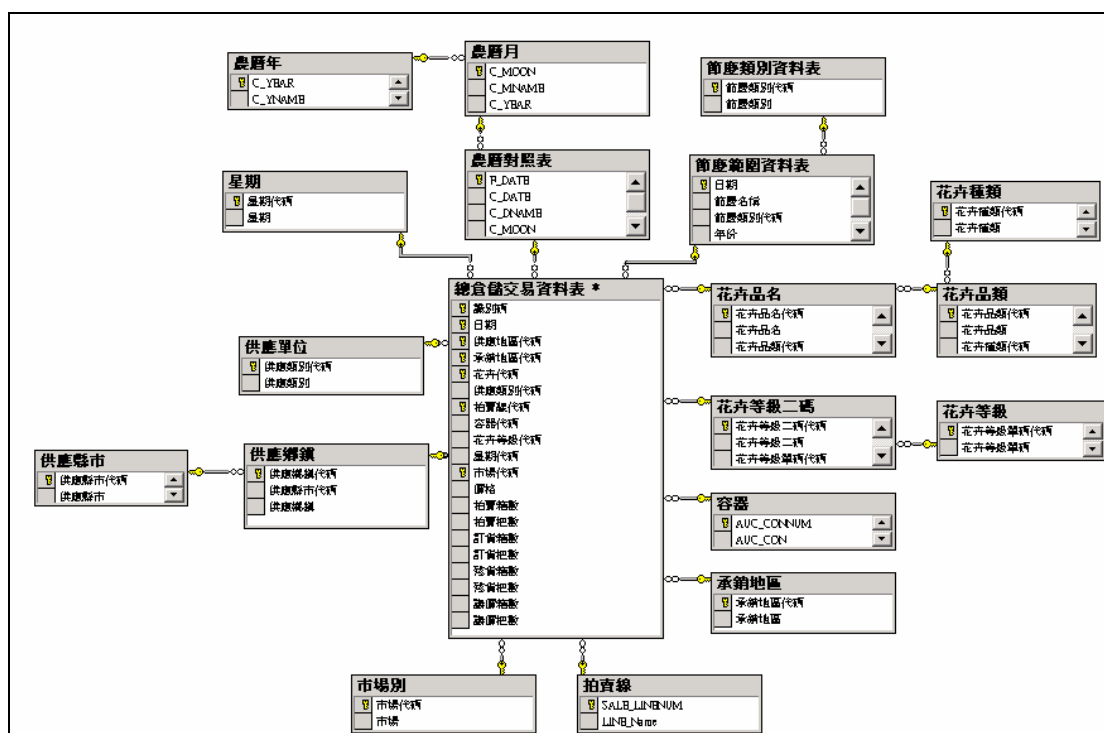


圖 4.5 總資料倉儲雪花網要關聯圖

## 4.1.2 資料轉換服務

一個完整的自動化更新轉換程序，經由仔細衡量評估後，需要按照拍賣、訂貨、殘貨、議價順序來進行資料轉換工作。當各家花市之四個原始交易檔案[2]，經個別自動化資料轉換服務轉入所屬資料超市後，首先應予進行拍賣資料轉換工作，如圖 4.6 拍賣資料轉換工作所示。



圖 4.6 拍賣資料轉換工作

當總倉儲伺服器連線至總倉儲的雪花網要資料庫時，需先執行清空交易資料表的動作，然後再依台北、台中、彰化、台南、高雄的順序進行拍賣資料轉換工作。在此將以高雄拍賣為例子，如圖 4.7 高雄拍賣資料轉換工作。首先連線至高雄資料超市的雪花網要資料庫，利用查詢方式有條理地將拍賣資料表當中的拍賣資料，轉換至總倉儲雪花資料庫中的總倉儲交易資料表，此部分 SQL 查詢語法建立如圖 4.8 高雄拍賣轉換資料工作來源查詢

所示。另外圖 4.9 高雄拍賣轉換資料工作屬性可明顯看出來源與目的地之轉換模式；此部分的轉換語法如圖 4.10 高雄拍賣轉換資料工作轉換語法所示，主要是以 VB script 撰寫而成。



圖 4.7 高雄拍賣資料轉換工作

This screenshot shows the '來源' (Source) tab of the '轉換資料工作屬性' (Data Conversion Job Properties) dialog box. The '描述(D):' (Description) field contains '高雄花市拍賣資料轉換'. The '連線:' (Connection) field shows '高雄花市雪花網要資料庫'. Under the 'SQL 查詢(S):' (SQL Query) section, a complex SELECT statement is displayed, mapping various auction data fields like 'AUC\_AUCNUM' to '識別號' (Identification Number) and 'AUC\_DATE' to '日期' (Date). Buttons at the bottom include '參數(M)...', '預覽(P)...', '建立查詢(Q)...', '瀏覽(B)...', '剖析查詢(A)', '確定', '取消', and '說明'.

圖 4.8 高雄拍賣轉換資料工作來源查詢

This screenshot shows the '轉換' (Conversion) tab of the '轉換資料工作屬性' (Data Conversion Job Properties) dialog box. The '名稱(N):' (Name) field is set to 'DTS Transformation'. The '類型:' (Type) is 'ActiveX Script'. Below, two lists are shown: '來源' (Source) on the left and '目的地' (Destination) on the right. The '來源' list includes fields like '識別號', '日期', '花卉代碼', and '承銷地區整合對應碼'. The '目的地' list includes '識別碼', '日期', '供應地區代碼', and '承銷地區代碼'. A mapping line connects the '識別號' field from the source to the '識別碼' field in the destination. Buttons at the bottom include '新增(W)', '編輯(E)', '刪除(D)', '測試(T)', '全選(S)', '全部刪除(A)', '確定', '取消', and '說明'.

圖 4.9 高雄拍賣轉換資料工作屬性

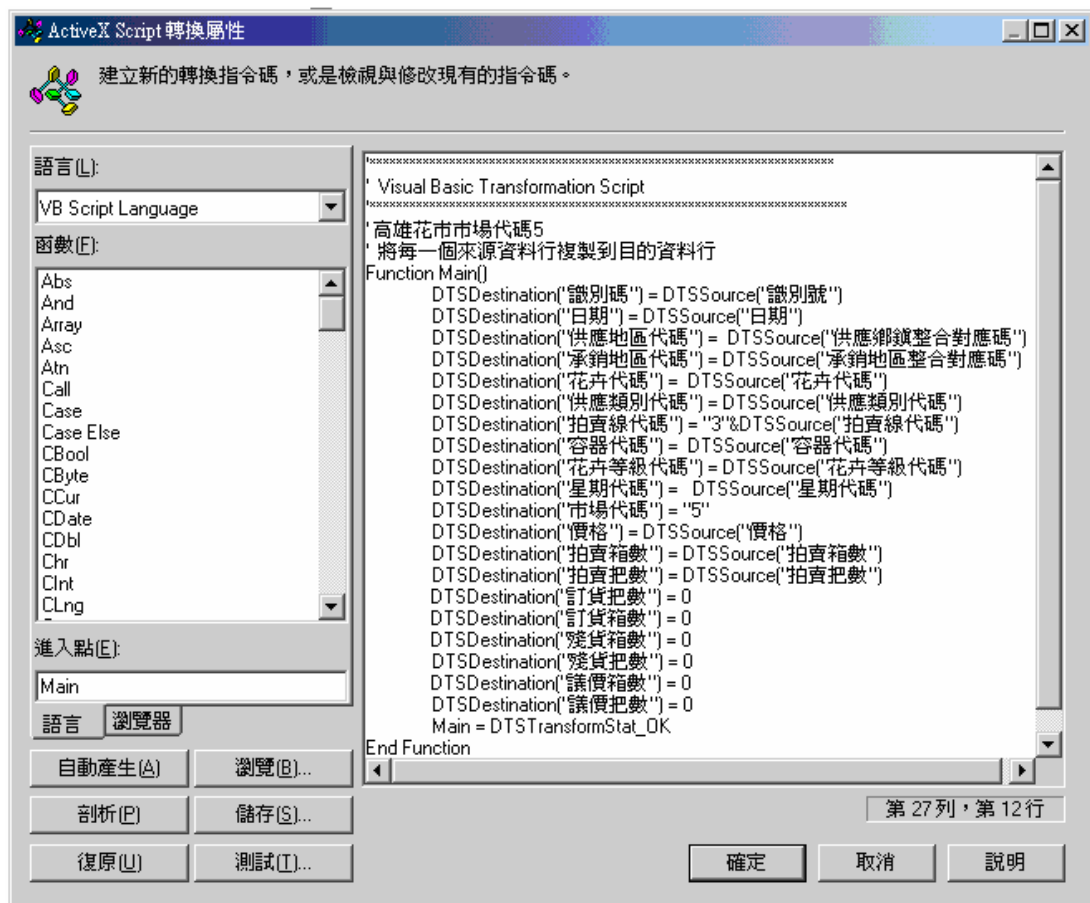


圖 4.10 高雄拍賣轉換資料工作轉換語法

基於台北、台中、彰化、台南拍賣的資料轉換四大部分已於先前建立完成，故僅需針對高雄拍賣的資料轉換進行著手。除了一些特殊狀況需要針對前者進行新增、整合動作之外，如按照每一個步驟便能完成拍賣資料轉換工作。表 4.2 拍賣資料轉換工作步驟表詳細地將拍賣資料轉換工作步驟做完整描述。

表 4.2 拍賣資料轉換工作步驟表

步驟	工作名稱	資料轉換動作
DTS_1-0	清空總倉儲 交易資料表	清空交易資料表
DTS_1-1	台北拍賣	1. 連線台北、台中、彰化、台南花市的雪花綱要資料庫與總倉儲雪花綱要資料庫 2. 將台北、台中、彰化、台南花市拍賣資料表的拍賣資料轉入總倉儲交易資料表
DTS_1-2	台中拍賣	
DTS_1-3	彰化拍賣	
DTS_1-4	台南拍賣	
DTS_1-5	高雄拍賣	1. 連線高雄花市的雪花綱要資料庫與總倉儲雪花綱要資料庫 2. 將高雄花市拍賣資料表的拍賣資料轉入總倉儲交易資料表

在完成拍賣資料轉換後半段工作後，緊接著進行修改訂貨資料轉換工作。由於一開始進行拍賣資料轉換工作時，就已經先行清空總倉儲交易資料表，故在訂貨資料轉換工作則不需要做清空的動作，如圖 4.11 訂貨資料轉換工作所示。依序按照台北訂貨、台中訂貨、

彰化訂貨、台南訂貨以及新增的高雄訂貨，便完成修改訂貨資料轉換工作。

各家訂貨的詳細程序如圖 4.12 高雄訂貨資料轉換工作所示。先連線至高雄資料超市的雪花網要資料庫，將拍賣資料表當中的訂貨資料，轉換至總倉儲雪花資料庫中的總倉儲交易資料表，表 4.3 訂貨資料轉換工作步驟表將訂貨資料轉換工作步驟詳細描述。



圖 4.11 訂貨資料轉換工作



圖 4.12 高雄訂貨資料轉換工作

表 4.3 訂貨資料轉換工作步驟表

步驟	工作名稱	資料轉換動作
DTS_2-1	台北訂貨	1. 連線台北、台中、彰化、台南花市的雪花網要資料庫與總倉儲雪花網要資料庫 2. 將台北、台中、彰化、台南花市拍賣資料表的訂貨資料轉入總倉儲交易資料表
DTS_2-2	台中訂貨	
DTS_2-3	彰化訂貨	
DTS_2-4	台南訂貨	
DTS_2-5	高雄訂貨	1. 連線高雄花市的雪花網要資料庫與總倉儲雪花網要資料庫 2. 將高雄花市拍賣資料表的訂貨資料轉入總倉儲交易資料表

在修改完成訂貨資料轉換工作後，緊接著進行修改殘貨資料轉換工作，如圖 4.13 殘貨資料轉換工作所示。同等於 DTS\_1 與 DTS\_2 的仿效模式，按照台北殘貨、台中殘貨、彰化殘貨、台南殘貨以及新增的高雄殘貨的順序，便完成了修改殘貨資料轉換工作。

各家殘貨的詳細程序如圖 4.14 高雄殘貨資料轉換工作所示。先連線至高雄資料超市的雪花網要資料庫，將拍賣資料表當中的殘貨資料，轉換至總倉儲雪花資料庫中的總倉儲交易資料表，表 4.4 殘貨資料轉換工作步驟表將殘貨資料轉換工作步驟詳細描述。



圖 4.13 殘貨資料轉換工作



圖 4.14 高雄殘貨資料轉換工作

表 4.4 殘貨資料轉換工作步驟表

步驟	工作名稱	資料轉換動作
DTS_3-1	台北殘貨	1. 連線台北、台中、彰化、台南花市的雪花網要資料庫與總倉儲雪花網要資料庫 2. 將台北、台中、彰化、台南花市拍賣資料表的殘貨資料轉入總倉儲交易資料表
DTS_3-2	台中殘貨	
DTS_3-3	彰化殘貨	
DTS_3-4	台南殘貨	
DTS_3-5	高雄殘貨	1. 連線高雄花市的雪花網要資料庫與總倉儲雪花網要資料庫 2. 將高雄花市拍賣資料表的殘貨資料轉入總倉儲交易資料表

在修改殘貨資料轉換工作完成後，接著將進行議價資料轉換工作，如圖 4.15 議價資料轉換工作所示。由於議價交易行為只有在台北花市進行，所新設立之高雄花市同樣沒有議價的問題存在，因此資料轉換工作也僅僅只有台北議價一個封包，如圖 4.16 台北議價資料轉換工作所示。故在此針對殘貨轉換工作無須做任何修改與更動，表 4.5 議價資料轉換工作步驟表將殘貨資料轉換工作步驟詳細描述。



圖 4.15 議價資料轉換工作



圖 4.16 台北議價資料轉換工作

表 4.5 議價資料轉換工作步驟表

步驟	工作名稱	資料轉換動作
DTS_4-1	台北議價	1. 連線台北花市的雪花綱要資料庫與總倉儲雪花綱要資料庫 2. 將台北花市拍賣資料表的議價資料轉入總倉儲交易資料表

最後在於 DTS\_5 層面，將於第 4.2 節詳細說明其程式撰寫、命令語法與自動化排程設計。



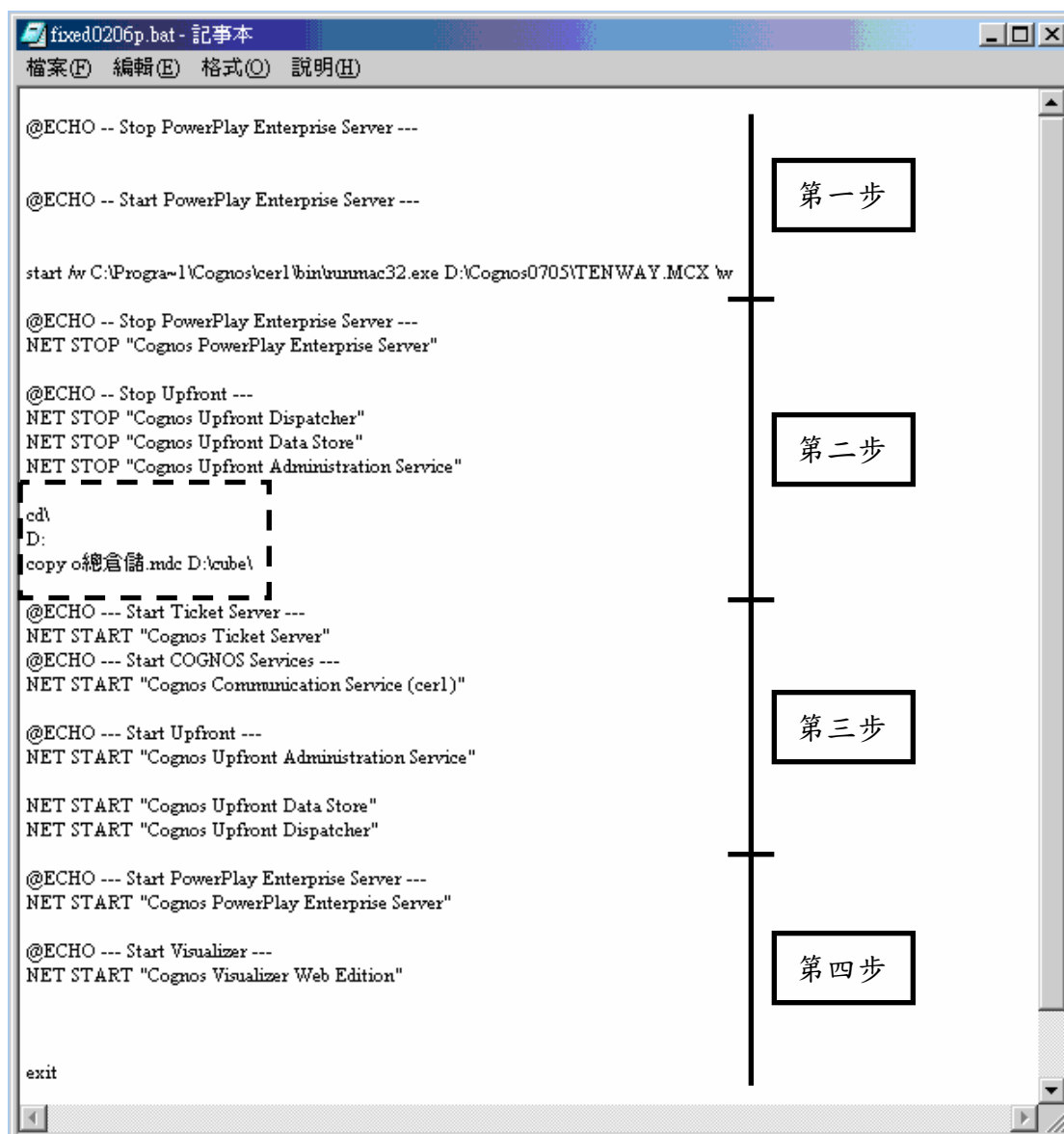
## 4.2 自動化排程設計

在於修改完成拍賣、訂貨、殘貨、議價資料轉換工作之後，仍需進一步考量自動化因素。因此，接下來將建立並設定自動化作業排程，其目的是為了讓資料不需手動就能自動轉進資料倉儲的功能機制，以節省不必要的人力與時間浪費，來達到最高效益。

本節分為兩小節，第 4.2.1 節介紹 DOS 程式撰寫自動化；第 4.2.2 節說明總資料倉儲自動化排程之建置。

### 4.2.1 DOS 程式撰寫

利用 DOS 程式撰寫而成的自動化 Macro 程式，讓 cognos 套裝軟體能夠自動開啟並執行多維度模型之轉換，使總資料倉儲可以每日定時更新舊有資料。然而自動化 Macro 程式所夾帶的檔案類型為 MS-DOS 之批次檔案-fixed0206p.bat，其程式撰寫如圖 4.17 所示。



```
@ECHO -- Stop PowerPlay Enterprise Server ---

@ECHO -- Start PowerPlay Enterprise Server ---

start /w C:\Program Files\Cognos\cer1\bin\runmac32.exe D:\Cognos0705\TENWAY.MCX /w

@ECHO -- Stop PowerPlay Enterprise Server ---
NET STOP "Cognos PowerPlay Enterprise Server"

@ECHO -- Stop Upfront ---
NET STOP "Cognos Upfront Dispatcher"
NET STOP "Cognos Upfront Data Store"
NET STOP "Cognos Upfront Administration Service"

cd\
D:
copy o總倉儲.mdc D:\cube\

@ECHO --- Start Ticket Server ---
NET START "Cognos Ticket Server"
@ECHO --- Start COGNOS Services ---
NET START "Cognos Communication Service (cer1)"

@ECHO --- Start Upfront ---
NET START "Cognos Upfront Administration Service"

NET START "Cognos Upfront Data Store"
NET START "Cognos Upfront Dispatcher"

@ECHO --- Start PowerPlay Enterprise Server ---
NET START "Cognos PowerPlay Enterprise Server"

@ECHO --- Start Visualizer ---
NET START "Cognos Visualizer Web Edition"

exit
```

圖 4.17 自動化 Macro 程式

由圖 4.17 可以看出此 MS-DOS 批次檔內使用幾個 DOS 指令，其可分為 ECHO、NET STOP、NET START 三種。接下來將簡單描述 DOS 指令功能的代表意義，以初步了解撰寫 DOS 程式的目的。在正常的情況下，批次檔中的每道命令執行前都會先出現螢幕上。有了 ECHO 命令，就可以任意控制是否列印出命令列。NET STOP 及 NET START 為停止與開啟服務的 DOS 指令，其主要是為了預防執行中的服務衝突以及停止與啟動 PowerPlay Enterprise Server 的子命令。

MS-DOS 之批次檔案-fixed0206p.bat 其運作步驟將進行更深入剖析，首先基於正式設計程式之基本要求，狀態顯示字串(@ECHO)。有了此字串狀態顯示，可以很清楚得知目前進行的狀態與步驟。接著在於解說程式內容之前，先了解相關元件名稱意涵，如表 4.6 所示。

表 4.6 Cognos 元件名稱整理表

編號	元件名稱	元件說明
1	Cognos PowerPlay Enterprise Server	線上分析處理工具並為 Cognos 子元件
2	Upfront Dispatcher	提供使用者一種簡單的查訪資訊和發佈資訊的服務機制為 Cognos 子元件 (分派、資料儲存與使用者服務)
3	Upfront Data Store	
4	Upfront Administration Service	
5	Cognos Ticket Server	Access Management 軟體中的一個子元件
6	Cognos Communication Service(cerl)	Cognos 溝通服務子元件
7	Cognos Visualizer Web Edition	Cognos Visualizer 軟體中的一個子元件，可將資訊以互動式、多指標的視覺化展現。(圖表)

得知相關元件名稱後，下一步將依序解說程式流程。第一步先進行開啟 Cognos 資料夾之啟動應用程式(runmac32.exe)與 Powerplay Transformer(TENWAY.MCX)並製作轉成超方體。第二步停止所有服務元件(Cognos PowerPlay Enterprise Service、Upfront Dispatcher、Upfront Data Store 與 Upfront Administrator Service)，虛線框為複製 D 槽超方體至 cube 資料夾做為備份並等待傳送至 IBM e-server1-1(將檔案發佈至網路的主機)。第三步進行 Access Management 的權限控管與 PowerPlay Enterprise Server 中 Upfront 入口網站使用者服務、資料儲存與分派。第四步將啟動 Cognos Visualier 軟體之子元件並結束批次檔程式。

## 4.2.2 總資料倉儲自動化排程設計

經 MS-DOS 之批次檔案-fixed0206p.bat 程式撰寫後，則完成自動化更新程序的資料轉換服務工作，接下來將利用 SQL Server2000 Agent 進行排程工作。首先點選並編輯作業步驟，如圖 4.18 自動化步驟語法編輯使用作業系統命令語法。代表在此路徑下執行自動化更新程序.dts 的檔案，接著再繼續編輯作業排程部分，如圖 4.19 自動化排程設定。命名此排程為自動化更新程序 snowfianl，於排程類型選項選擇重複執行方式便可以設定每日執行運作時間，如圖 4.20 所示。依序前幾個動作流程將可以順利完成自動化的設計。

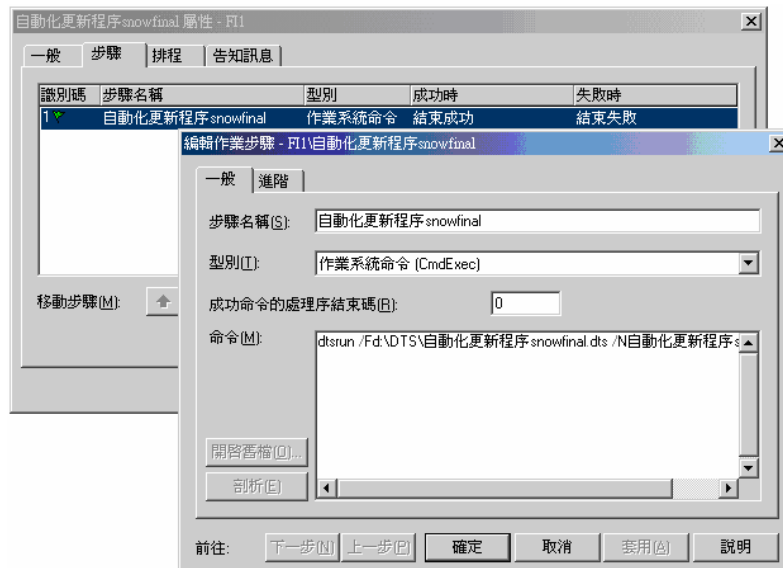


圖 4.18 自動化步驟語法編輯

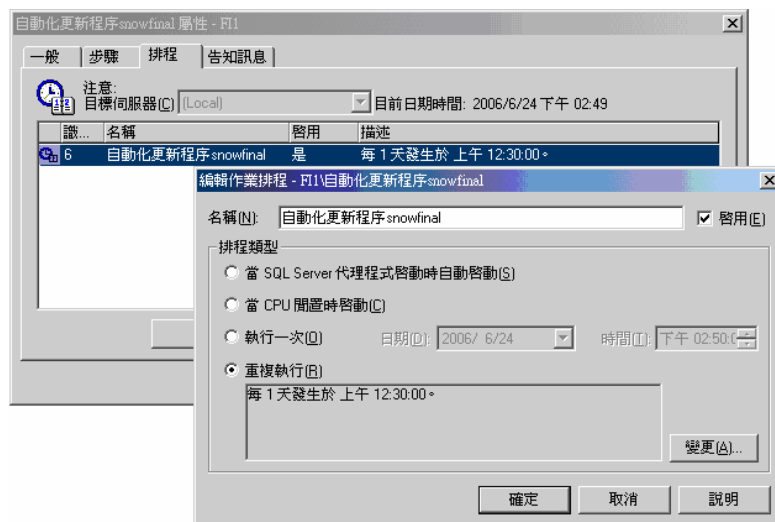


圖 4.19 自動化排程設定

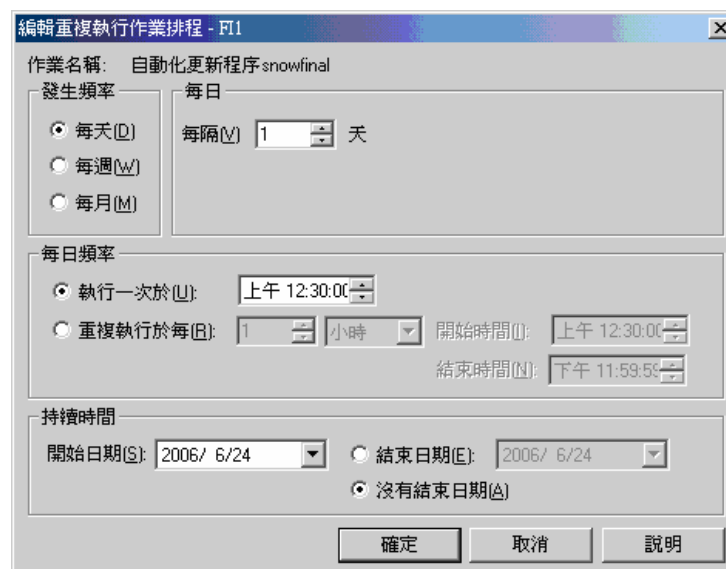


圖 4.20 自動化排程重複執行設定

## 第五章 關聯性法則對大宗花卉產量之颱風特性實證分析

在第二章之第 2.2 節所述的各種颱風特性解析為其利基點，並依照此方向收集歷年侵台颱風之相關資料，並進行資料前置處理(如附錄 A)。再來由第 2.1 節了解到大宗花卉相關資訊後，得知相關影響花卉(如節慶、生產季節性等)變項，同樣對資料作前置處理。最後再經由 2.4 節比較之後決定使用關聯性法則 Apriori 進行階段性分析並進一步探討颱風特性對大宗花卉變化之影響。本章分為 5.1 節研究架構與假，5.2 節實證模型之建立及 5.3 節結果說明等三節進行討論。

### 5.1 研究架構與假設

為了初步了解整個流程的運作，因此於圖 5.1 流程圖解說之。

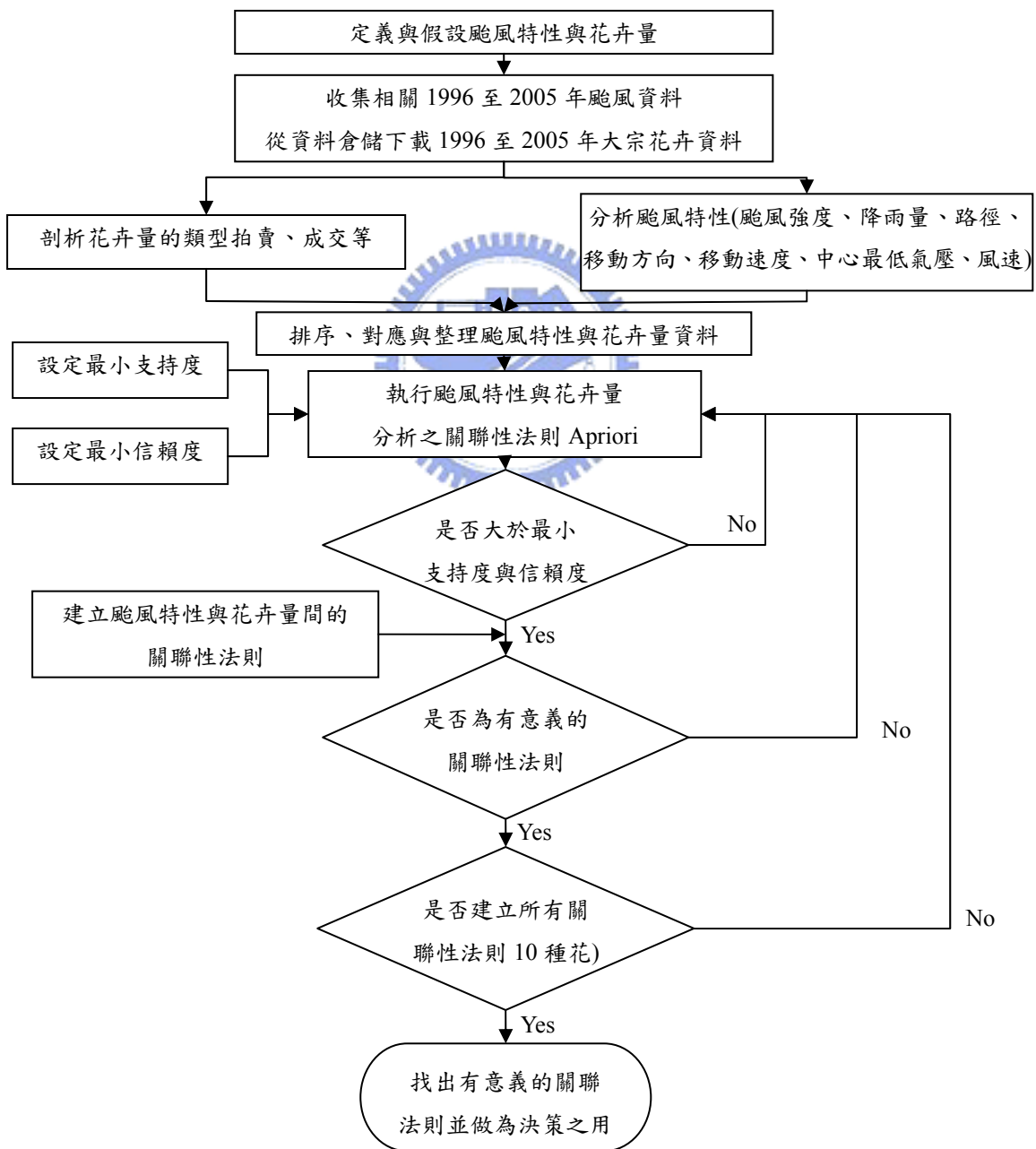


圖 5.1 流程圖

### 5.1.1 研究架構

將第 2.1.2 節所定義好的大宗花卉進而排序出前十名之後，根據第二章 2.2 節的颱風特性分析之文獻回顧，適當加以縮減為 4 大颱風特性。接著，再把颱風特性與前十大之大宗花卉產地量與總成交數量進行同步關聯分析。建立有意義的關聯性法則。

### 5.1.2 研究前提

依照第 2.2 節相關文獻回顧分析結果顯示，得知影響大宗花卉數量之颱風特性因素有颱風強度、颱風移動路徑、最大總雨量、颱風最低氣壓。但以實際資料驗證推估時，若干變數會因為資料欠缺而產生難以估計的問題，必須予以適當修正或剔除。例如：颱風侵襲台灣的天數範圍內，如果發生所要分析的資料數據遇到花卉批發市場營業狀態為休市時，則會導致「總成交量為零」的問題發生。另外還有「如何選擇適當雨量偵測站問題」與「資料筆數之建構」三大問題以及「節慶與大宗花卉無相關性假設」和「大宗花卉合乎其季節性假設」兩大假設。此時，不但會造成研究過程中的一大障礙；情況嚴重時，可能會讓整個研究因此而無法繼續進行下去。為了探究三大相關問題與兩大研究假設，故將更進一步於以下深入探討之，其三大問題與兩大研究假設如下所示。

#### 1. 問題類型一：各家花卉批發市場遇到總成交量為零的問題

每家花卉批發市場皆有其休市日，於休市日時，則停止營業，導致相關玫瑰花數量數值呈現零單位。剛好休市日當天對應到颱風侵襲日，如此情形發生(如表 5.1 的陰影區所示)，將會產生資料無法有效對應問題產生，因此需要予以適當的修正與刪除規範機制。

表 5.1 颱風對應各家花市休市與總成交量分佈表

項目 編號	颱風侵台期間	玫瑰總成交量(把)					花市正常營運與休市		問題
	資料有問題日	台北	台中	彰化	台南	市場別	營業狀態	休市市場	型式
1	1998/7/8-10	0	4340	5327	3072	12739	正常休市	台北	型 1
2	1998/8/3-5	7001	1981	3599	2290	14871	否	否	
3	1998/9/27-29	6313	2518	3187	1683	13701	否	否	
4	1998/10/14-16	2980	453	367	831	4631	否	否	
5	1998/10/25-27	6109	2561	3455	1531	13656	否	否	
6	1999/6/5-6	0	7091	10816	3932	21839	否	否	型 2
7	1999/10/6-10	0	0	0	0	0	正常休市	四家市場	型 3

由表 5.1 可以清楚看出有三種不同形式的問題，其問題分析與改善如下所示。

型 1 問題分析：編號 1 於台北批發市場總成交量顯示 0。經查證結果，發現 7 月 10 日恰好為台北批發市場之休市日，因此需適當刪除或改變分析型態。

型 1 問題改善：

- 刪除表 5.1 之相關總成交量資料，減少缺乏資料因素。
- 將缺乏玫瑰花總成交量之台北花市資料，往前一天、往後一天或內

插法方式求得台北花市玫瑰花的總成交量。

- c. 四家花市皆改變分析型態為颱風開始至結束日所增加或減少的玫瑰花總成交量。

型2問題分析：編號6於台北批發市場總成交量同樣也顯示0，再經查證結果，確定6月6日並非為台北批發市場之休市日，因此判斷結果為未顯示出玫瑰花總成交量。

型2問題改善：通知維修人員進行修復或刪除此列資料。

型3問題分析：編號7之所有批發市場總成交量皆顯示0，又恰好是休市日，故與型1問題有部分類似的地方。

型3問題改善：可遵循型1問題來進行問題的改善。

## 2. 問題類型二：如何選擇適當雨量偵測站問題

每逢颱風侵襲時所帶來是大量豪雨，因此會有很多偵測站專門在測量颱風的降雨量，其偵測站從南到北都有不同設站地點，故選擇適當雨量偵測站來記錄颱風降雨量才會使分析結果更有其可靠信。

由第2.2.1節文獻整理得知，可以選定所要偵測降雨量的點，可分為北、東、中與南部，其所對應的觀測站可為台北、花蓮、南投、高雄雨量觀測站。在此研究將考量以台北雨量觀測站為準，如圖5.2所示。

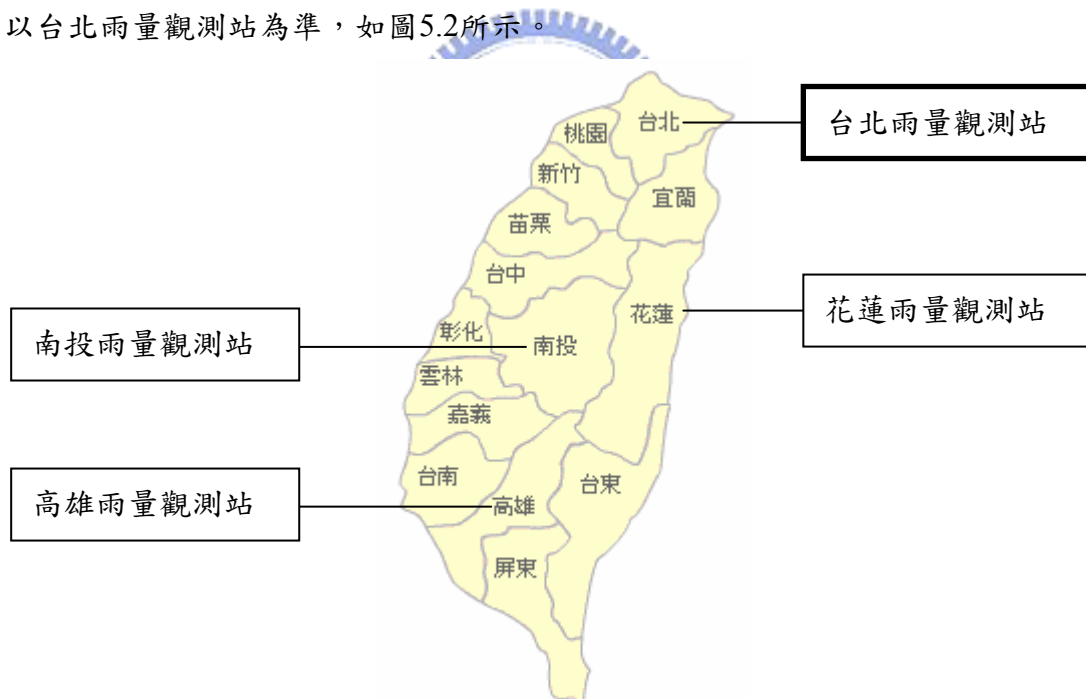


圖 5.2 雨量觀測站選擇分佈圖

## 3. 問題類型三：資料筆數之建構問題

基於受限於資料筆數的關係，故對於每筆颱風資料都必須更進一步探討與分析，避免隨意的刪除資料筆數，最後讓分析結果不具代表性且無意義。

探討完三大問題後，接著將針對「節慶與大宗花卉無相關性假設」和「大宗花卉合乎其季節性假設」兩大假設進行詳細解說。

### 1. 假設一：節慶與大宗花卉無相關性假設

何種節慶應該送哪一種花，事實上並沒有一定之準則。依照文獻回顧得知在於情人節節日來臨時，一般人固然都會立即聯想到要送情人美麗的玫瑰花，但是花卉行銷工作者和花店業者聯盟仍可依照當季時令盛產的花卉作行銷宣傳促銷活動。除了玫瑰花之外，亦可以用花中后的鬱金香，或婀娜多姿的金魚草替代價昂的玫瑰花，作為情人間表達深切愛意的另外選擇。除了可增加當季花卉的消費量之外，同時消費者又可享受價廉物美的花卉產品。因此假設節慶對應大宗花卉種類無一定對應的花卉標準。

### 2. 假設二：大宗花卉合乎其季節性假設

花卉依其種類的不同，也有其不同的開花及生產季節，而呈現出明顯地季節循環變動性。通常依大自然的條件下，花卉生產仍以春季較多。如表5.2所示。有些花卉可以找到其對應月份，有些花卉只能利用月份範圍。同樣地，也可以用圖形來解說，如圖5.3所示。

表 5.2 季節性花卉對應大宗花卉對應表

大宗花卉對應季節性					
花卉名稱	火鶴花	小菊	香水百合	大菊	玫瑰
月份	6 月	11 月~3 月	3 月~5 月	11 月~3 月	3 月~5 月
花卉名稱	非洲菊	劍蘭	葵百合	康乃馨中輪	洋桔梗
月份	3 月	10 月~4 月	3 月~5 月	5 月	4 月

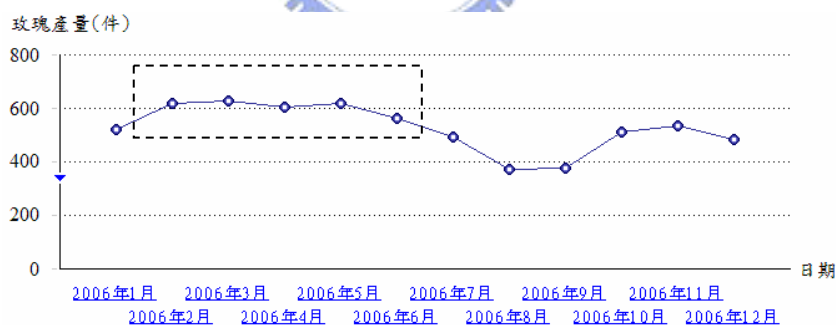


圖 5.3 季節性花卉之玫瑰產量分析圖

由圖5.3可以了解到，在於3至6月份之玫瑰產量相對偏高的狀態，代表玫瑰在此月份範圍內具季節性花卉的現象，因此在做關聯分析時，應該加以區分之，讓分析能夠更加精確。

經由以上四大問題與兩大假設之深入探討與解析後，面對可能發生的所有非對應資料皆有其解決的方法。如此一來，方能進行下一節所要探究的關聯性法則之建立。

## 5.2 實證模型建立

針對相關侵台颱風資料作初步的關聯性法則分析，分析並找出侵台颱風內在隱藏的有用資訊，建立關聯颱風特性法則(附錄F)。接著再把對應之大宗花卉相關資訊做更進一步整合，作為本研究重心的關聯性法則分析，挖掘出颱風特性對應至大宗花卉之關聯性法則，作為評估颱風對大宗花卉產地量與總成交量變化的重要依據。

### 5.2.1 關聯性法則

在關聯性法則建立之前，首先要定義清楚所要分析的相關項目，以利於分析過程中能夠更加快速有效率，接著才可以進行主要階段關聯性法則。以上述說之相關定義與假設及主關聯性法則將於以下闡述。

#### 5.2.1.1 相關定義與假設

相關變項的定義與假設是分析前首要的工作；因此於選擇颱風特性層面上，需對颱風特性予以清晰的定義與假設。其颱風特性為颱風強度、最大總雨量(降雨量)、颱風移動路徑(方向與通過地區)、颱風移動速度、颱風中心最低氣壓、颱風中心最大風速以及侵台颱風時間年份、季節、旬、停留時間以及侵台颱風近節慶數(前後五天)等。另一方面，此研究將針對「10種」大宗花卉進行一連串分析。基於花卉產地對大宗花卉量有著一定之影響程度，因此大宗花卉產地供應量也應予以在花卉產地供應量排序(前三)、產量供應波動狀態、最大起伏變動大小(量)給予適當的定義與假設。最後在大宗花卉成交量方面，為了讓五大批發市場能夠有進一步比較的空間，故需多加考量其各家批發市場資料的存在與否。適當將上述變項加以整合初步分析出有用的資訊，其相關定義與假設如表 5.3、5.4、5.6 所示。

表 5.3 類別定義與相關項目假設表(上)

類別	編號	事件	假設項目			單位
時間	1	年份	1996 ~ 2005			年
	2	季節	春(3 ~ 5)、夏(6 ~ 8)、秋(9 ~ 11)、冬(12 ~ 2)			月
	3	旬	上旬(1 ~ 10)、中旬(11 ~ 20)、下旬(21 ~ 31)			日
	4	停留天數	1 ~ 2、3 ~ 4、5 以上			天
	5	近節慶個數	1、2、不符(超過 5 天)			天
	6	近節慶區段	前、後、前與後、超出區段			N
	7	最短之近節慶天數	1~3、4~6、7 以上			天
颱風特性	8	颱風強度	輕度颱風、中度颱風、強烈颱風			N
	9	最大總雨量	0 ~ 500、501 ~ 1000、1001 ~ 1500、1501 ~ 2000			公釐
	10	最大總雨量發生地點	(1)	台北	彭佳嶼、鞍部、竹子湖	N
			(2)	宜蘭	宜蘭、蘇澳	
			(3)	花蓮	花蓮	

			(4)	台東	成功、台東、大武、蘭嶼			
			(5)	嘉義	阿里山			
			(6)	台南	台南			
			(7)	高雄	高雄			
			(8)	屏東	恆春			
	11	路徑分類	(1)	第一類型	(a)	第一類路徑		N
					(b)	第二類路徑		
			(2)	第二類型	(c)	第三類路徑		
					(d)	第四類路徑		
			(3)	第三類型	(e)	第五類路徑		
					(f)	第六類路徑		
	(4)	其他類型	(g)	第七類路徑				
12	通過地區	(1)	北部	北部、北中南部、北部中部		N		
		(2)	東北部	東北部(角)				
		(3)	東部	東部、東部近海				
		(4)	西部	西岸中南部、西部近海				
		(5)	中部	中部、中南部				
		(6)	南部	南部				
		(7)	其他(海上)	台灣海峽、(北部、東北部、東部、西部、南部)海上				
	13	移動方向	(1)	北	北、北(北)東、北(北)西		公里/小時	
			(2)	西	西、西北西、西南西			
			(3)	東	東北東			
			(4)	南	南南西			
	14	移動速度	慢(0~15)、中(16~30)、快(31~45)					
	15	中心最低氣壓	(1)	第一種最低氣壓類型	P>980		百帕-hpa	
					980≥P>965			
			(2)	第二種最低氣壓類型	965≥P>945			
					945≥P>920			
					P≤920			
16	中心最大風速	(1)	第一種最大風速類型	17.2 以下		米/秒		
				17.2~32.6				
		(2)	第二種最大風速類型	32.7~50.9				
				51.0~66.9				
		(3)	第三種最大風速類型	67.0 以上				

由表5.3可以清楚定義出編號1至16號所屬的十六種不同事件，其中在於颱風特性層面編號8至16號分別為颱風強度、颱風最大總雨量、颱風最大總雨量發生地點、路徑(分類)、通過地區、移動方向、移動速度、中心最低氣壓、中心最大風速九大項進行有效率事件內

之小項目個數縮減，如編號11路徑分類之七小項縮減至第一類型、第二類型、第三類型、其他四項，如此假設的方式可以利於關聯性法則進行之時間縮減與消耗記憶體容量大小。

要進行大宗花卉成交量相關資料分析前，其生產源頭之花卉產地資訊也必定占整個分析很大的影響程度存在。因此，預先將產地供應相關資料作初步整理與假設，以玫瑰花為例(其餘九種花卉請參照附錄D-2~D-10)，作為主階段分析前之重要基礎。如表5.4所示。

表 5.4 玫瑰花類別定義與相關項目假設表(中)

類別	編號	事件	假設項目				單位
玫瑰花 產地供應	1	產地 供應 排序	花卉產地供應量排序 (前三名)	(1)	彰南中	彰化→南投→台中	N
				(2)	南彰中	南投→彰化→台中	
			定義：颱風對應花卉產地 相關供應縣市之前三名 排序	(3)	南中彰	南投→台中→彰化	
				(4)	南中北	南投→台中→台北	
				(5)	南彰雲	南投→彰化→雲林	
				(6)	南彰苗	南投→彰化→苗栗	
	2	產地 供應 波動	產量供應波動狀態	(1)	上升狀態	持續上升	N
						先上升後下降	
			定義：颱風期間對應花 卉產量供應波動狀態	(2)	下降狀態	持續下降	
						先下降後上升	
	3	最大起伏變動大小(量)	定義：颱風期間對應花 卉供應最大起伏波動大 小(量)	0~250、251~500、500 以上			件

由表5.4可以清楚得知玫瑰花產地供應受侵台颱風影響之兩大事件，其一花卉產地供應之前三名排序，在於初步資料分析後，發現產地排名隱藏著非確定性，故列入分析項目。其次，可以了解到當侵台颱風來臨時之供應波動狀態與波動大小，讓相關的資料更具有分析價值。

然而，再進行主階段分析前，侵台颱風年份與相關市場資料對應性也是很重要的，其對應結果如表 5.5 所示。

表 5.5 年份與五大市場資料對應表

市場 年份	台北	台中	彰化	台南	高雄	備註
1996 年	●	●	◎	◎	●	分析難度高
1997 年	◎	●	◎	◎	●	分析難度高
1998 年至 2002 年	◎	◎	◎	◎	●	分析難度中
2003 年至 2005 年	◎	◎	◎	◎	◎	分析難度低

◎ 資料可對應性 ● 資料不可對應性

經由資料對應後，很明顯地發現到台北在 1996 年無對應資料，台中情況比台北更糟，1996 與 1997 年皆無對應資料，無對應資料相對的則無分析的價值。因此，將 1996 與 1997 年之所有資料予以刪除，並把 1998 至 2005 年的資料，分成四大市場與五大市場兩部分分析。為了要在於侵台颱風來襲時分析其拍賣量與成交量在於各大市場之間所隱藏的相關重要資訊，故要對其量(拍賣與成交)作進一步定義與假設，以玫瑰花為例(其餘九種花卉請參照附錄 D-2~D-10)，如表 5.6 所示。

表 5.6 玫瑰花類別定義與相關項目假設表(下)

類別	編號	事件	假設項目						單位
數 量	1	拍 賣 量	颱風前一天與 颱風當天	(1)	北>彰>中>南			北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
				(2)	北>彰>南>中				
				(3)	北>中>彰>南				
				(4)	中>北>彰>南				
				(5)	中>北>彰>雄>南				
				(6)	中>北>雄>彰>南				
				(7)	過半休市				
			拍賣量 變化狀態	(1)	北(降)			降:下降程度為 四大市場之最	把
				(2)	彰(降)				
				(3)	中(降)				
				(4)	南(降)				
				(5)	無法判斷				
	2	成 交 量	颱風前一天	(1)	北>彰>中>南			北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
				(2)	北>彰>南>中				
				(3)	北>中>彰>南				
				(4)	雄>北>彰>中>南				
				(5)	雄>北>彰>南>中				
				(6)	雄>北>中>彰>南				
				(7)	過半休市				
			颱風當天	(1)	北>彰>中>南			北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
				(2)	北>彰>南>中				
				(3)	北>南>中>彰				
				(4)	雄>北>彰>中>南				
				(5)	雄>北>中>彰>南				
				(6)	過半休市				
			成交量 變化狀態	(1)	北(降)	(6)	無法 判斷	降:下降程度為 四大市場之最	把
				(2)	彰(降)				
(3)				中(降)					
(4)				南(降)					
(5)				雄(降)					

表5.6所示拍賣與成交量皆以颱風來襲之前一天與當天作為衡量基準，如此可以很清楚得知量的變化與各家市場量的比較，對於過半休市的情況也予以清楚假設。讓不確定性能夠降到最低，分析的結果更有可信度。

在眾多關聯性法則中，須找出具有意義的關聯規則，以提供決策者參考依據。如表5.7所示。

表 5.7 有意義關聯性法則的判斷

關聯性法則	是否為有意義的關聯性法則
路徑分類 AND 通過地區 ==> 颱風強度	是
路徑分類 AND 最大總雨量發生地點 ==> 當天拍賣量的比較	是
路徑分類 AND 當天拍賣量的比較 ==> 當天成交量的比較	是
颱風強度 AND 產量供應波動 ==> 颱風移動速度	否
前一天拍賣量比較 AND 最大總雨量 ==> 最大總雨量發生地點	否
當天拍賣量比較(中>北>彰>雄>南) AND 最大總雨量發生地點 ==> 最大總雨量	否

由表5.7得知有意義關聯性法則的判別，不外乎關聯性法則所屬之左邊與右邊項目應予以限制，左邊應為颱風特性，右邊應為大宗花卉相關項目，而非同時存在左邊或同時出現於右邊。對於最小支持度與最小信賴度的值也需符合，如果不符合其設定值，則予以刪除不考慮。接著將於下兩小節進行主階段關聯性法則(颱風特性與大宗花卉量)。

### 5.2.1.2 主階段關聯性法則

初步關聯性法則分析(參照附錄F)，主要是針對颱風相關特性進行更深入的分析與縮減，其颱風特性可分為九項，如表5.8所示。

表 5.8 九大颱風特性

事件(Event)	項目(Item)
1 颱風強度	輕度颱風、中度颱風、強烈颱風
2 最大總雨量	0~500、501~1000、1001~1500、1501~2000
3 最大總雨量發生地點	台北、宜蘭、花蓮、台東、嘉義、台南、高雄、屏東
4 颱風移動路徑	第一類型(1,2)、第二類型(3,4)、第三類型(5,6)、其他類型(7)
5 颱風通過地區	北部、東北部、東部、西部、中部、南部、其他(海上)
6 颱風移動方向	北、西、東、南
7 颱風移動速度	慢(0~15)、中(16~30)、快(31~45)
8 颱風近最低中心氣壓	第一種最低氣壓、第二種最低氣壓、第三種最低氣壓
9 颱風近中心最大風速	第一種最大風速、第二種最大風速、第三種最大風速

由表5.8可以得知事件與項目在於颱風特性之清楚區別，9種事件所對應的總項目共有高達39項。而此39項可作為關聯性法則分析之相關變項，詳細說明可參照5.2.1.1節。詳細颱風資料(Excel整理)如圖5.4颱風數據整理圖。

編號	年份	期間	發生點	結束點	期間 (停留天數)	颱風名稱	颱風強度	路徑分類 代號(7)	通過地區	移動 方向	移動速度 (Km/hr)	中心最低 氣壓(百 帕-hpa)	最大風速 (米(m)/秒 (s))	最大總雨 量(mm)
1	2005	9-10	9月30日	10月3日	4	LONGWANG 龍王	強烈颱風	2	中部	W	23	925	51	181
2	2005	8-9	8月30日	9月1日	2	TALIM 泰利	強烈颱風	2	中部	W	21	920	53	666
4	2005	8	8月3日	8月6日	4	MATSA 馬莎	中度颱風	1	北部海上	NW	14	955	40	666
5	2005	7	7月16日	7月20日	5	HAITANG 海葵	強烈颱風	1	北部	WNW	15	912	55	1215.3
6	2004	12	12月3日	12月4日	2	NANMADOL 南瑪都	中度颱風	6	南部	ENE	27	958	38	296.5

圖 5.4 颱風數據整理圖

接著大宗花卉方面可分為兩大部份，第一部份為大宗花卉產地供應量的探討與分析，詳細資料如圖5.5大宗花卉產地供應量概略圖。第二部份為大宗花卉拍賣量與成交量的剖析，詳細資料如圖5.6大宗花卉拍賣量與成交量概略圖。兩部分皆擁有其所屬的變項，如表5.9所示。

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
18			宜蘭	花蓮	台北縣	台北市	新竹	桃園	苗栗	台中市	台中縣	彰化	南投	雲林	嘉義	台南縣	台南市	高雄市	高雄縣	屏東	台東	供應縣市
19	1996年7月	1996年7月29日	0	0	0	0	0	0	0	0	0	13	81	134	9	0	0	17	0	0	0	254
20	1996年7月	1996年7月30日	0	0	0	0	0	0	0	0	0	8	75	128	0	3	0	0	0	0	0	214
21	1996年8月	1996年8月1日	0	0	0	0	0	0	0	0	0	1	5	12	0	0	0	0	0	0	0	18
22	加總		0	0	0	0	0	0	0	0	0	22	161	274	9	3	0	17	0	0	0	486
23	百分比		0	0	0	0	0	0	0	0	0	0.0453	0.3	0.6	0	0	0	0.035	0	0	0	1
24																						

圖 5.5 大宗花卉產地供應量概略圖

A	B	C	D	E	F
136	彰化	高雄	台南	台北	台中
137	總成交量(把)	總成交量(把)	總成交量(把)	總成交量(把)	總成交量(把)
138	國產花	國產花	國產花	國產花	國產花
139	FR玫瑰	國產花	FR玫瑰	FR玫瑰	FR玫瑰
140	2004年6月28日	3147	39255	1487	14486
141	2004年6月29日	3100	62520	1444	8478

圖 5.6 大宗花卉產拍賣量與成交量概略圖

表 5.9 大宗花卉兩大部分剖析

事件(Event)		項目 (Item)
1	產地供應(大宗花卉)	花卉產地供應量排序(前三名)(圖 5.5 左下框)
		產量供應波動狀態(圖 5.5 右框)
		最大起伏變動大小(量)(圖 5.5 右框)
2	大宗花卉數量變化 (拍賣量與成交量)	拍賣量變化(前一天、當天)、拍賣量變化狀態
		成交量變化(前一天、當天)(圖 5.6 下框)、成交量變化狀態(圖 5.6 右框)

由圖5.5與圖5.6可以清楚得知大宗花卉產地供應與大宗花卉數量變化(市場)的情形，配合表5.4與表5.6將可以完整表達出所屬意義。此外，表5.9顯示出大宗花卉之事件與項目之區分，2大事件包含9個項目。有效的將表5.8(颱風特性)與表5.9(大宗花卉量變化)部分予以適當地將相關項目加以整合後，進行主階段關聯性法則分析。其步驟如下。

步驟一首先利用資料倉儲下載所限定大宗花卉產地供應量相關資料(1996年至2005年大宗花卉產地資料)與大宗花卉拍賣量及成交量(1996年至2005年大宗花卉量的資料)。同樣地，收集相關歷年颱風資料，以對應總資料倉儲之1996年至2005年侵台颱風資料為準的颱風筆數有「34筆」。

步驟二將設定「最小支持度15%」與「最小信賴度門檻值80%」。

步驟三將整合完成的資料(颱風特性與花卉量變化)依照轉換數據表形式轉換成數據並符合國際標準規定形式，如圖5.7或(參照附錄C、附錄D-1~D-10與附錄E)。

步驟四進行Apriori演算法軟體(如附錄E)。

步驟五找出颱風特性與大宗花卉變化間有意義的關聯性法則，共需進行十次演算(玫瑰、大菊、火鶴花、香水百合、葵百合、非洲菊、洋桔梗、康乃馨、劍蘭十種大宗花卉)並將找出的關聯性法則作為日後決策之用。

編號	年份	颱風強度	最大總雨量	出現地點	路徑分類	通過地區	移動方向	移動速度	中心最低氣壓	最大風速	玫瑰平均供應量(前三名排序)	最大起伏變動大小	產量波動	颱風前一天拍賣量比較	颱風當天拍賣量比較	拍賣量變化量狀態	颱風前一天成交量比較	颱風當天成交量比較	成交量變化量狀態
1	1	1	1	7	3	4	1	1	1	1	5	3	4	1	1	3	1	1	2
2	2	1	1	5	1	5	1	2	1	1	5	1	4	1	1	1	1	1	1
3	3	1	1	1	2	7	1	1	1	1	5	3	3	7	1	5	7	1	6
4	4	1	3	2	1	2	7	1	1	3	3	2	3	2	1	1	1	1	1
5	5	1	2	2	3	7	1	1	2	2	6	1	1	2	1	1	3	1	1

圖 5.7 颱風特性與大宗花卉轉換數據圖

### 5.3 計算過程說明

由 2.4 節之 2.4.1 節得知關聯性法則的形式為  $X \rightarrow Y$ 。其中  $X$  可定義為颱風特性項或大宗花卉量項， $Y$  可定義為颱風特性項或大宗花卉量項。支持度(Support)為支援關聯規則颱風特性發生次數百分比。信賴度(C Confidence)為左邊颱風特性發生中，同時又發生颱風特性或大宗花卉量狀態(產地量、拍賣量、成交量相關比較變項)的概率。而  $Confidence(X \rightarrow Y) = Support(X \cup Y, D) / Support(X, D)$ 。找出所有  $X \rightarrow Y$  形式的關聯性法則，並且滿足下列條件： $Support(X \cup Y, D) \geq Min\_Support$  且  $Confidence(X \rightarrow Y) \geq Min\_Confidence$ 。

以分析颱風特性近中心最大風速與中度颱風兩項之間是否存在關係，其計算過程如下所示。

計算一：

設定  $X$  為第二種中心最低氣壓(965 以下且大於 945)， $Y$  為中度颱風， $Min\_Support = 35\%$ ， $Min\_Confidence = 95\%$

$Support(X) = 0.359$  (颱風特性  $X$  發生次數百分比)

$Support(XY) = 0.359$  (颱風特性  $X$  與颱風特性  $Y$  發生次數百分比)

$Confidence(X \rightarrow Y) = P(Y|X) = \frac{Support(X \cup Y)}{Support(X)} = \frac{0.359}{0.359} = 1$  (颱風特性  $X$  發生時，同時又發生颱風特性  $Y$  的概率)

以分析颱風特性最大總雨量與大宗花卉成交量變化狀態兩項之間是否存在關係，其計算過程如下所示。

計算二：

設定  $X$  為最大總雨量(1001-1500mm)， $Y$  為大宗花卉成交量變化狀態(台北降)， $Min\_Support = 10\%$ ， $Min\_Confidence = 95\%$

$Support(X) = 0.147$  (颱風特性  $X$  發生次數百分比)

$Support(XY) = 0.147$  (颱風特性  $X$  與大宗花卉成交量變化狀態  $Y$  發生次數百分比)

$Confidence(X \rightarrow Y) = P(Y|X) = \frac{Support(X \cup Y)}{Support(X)} = \frac{0.147}{0.147} = 1$  (颱風特性  $X$  發生時，同時又發生大宗花卉成交量變化狀態  $Y$  的概率)

經由以上計算過程可以清楚了解其 Support 和 Confidence 如何計算出來，但是如果碰到並非 1 對 1 的情況發生時，就必須藉由 Apriori 演算軟體進行演算(如附錄 E)，而非利用計算可得到。

## 5.4 結果說明

進行主要階段關聯性法則分析，在於最小支持度門檻值及最小信賴度門檻值予以設定不同，因而產生不同層次的關聯性法則。如果兩大門檻值隨意設定，將會產生數種關聯性法則，法則之多並非本研究之主要目標，而是要追求具有相當意義有用的法則，也就是值得且令人感到興趣的法則。接著會以同樣方式進行相關颱風特性與大宗花卉產地供應量及颱風特性與大宗花卉成交量的關聯性法則分析，其演算結果，將會於以下作深入剖析與解釋。

### 1. 針對颱風強度與最大總雨量找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 10%，最小信賴度門檻值 = 80%，關聯性法則如表5.10所示：

表 5.10 關聯性法則(六)

項目組數	支持與信賴度	關聯性法則	值
2	(a)支持度	輕度颱風 最大總雨量(0-500mm) 產量變動狀態(先下降後上升)	0.147
	(b)信賴度	輕度颱風 AND 最大總雨量(0-500mm) ==> 產量變動狀態(先下降後上升)	0.833
	(c)支持度	輕度颱風 最大總雨量(0-500mm) 當天成交量比較(北>彰>中>南)	0.147
	(d)信賴度	輕度颱風 AND 最大總雨量(0-500mm) ==> 當天成交量比較(北>彰>中>南)	0.833
有意義的關聯性法則： 1. 輕度颱風 AND 最大總雨量(0-500mm) ==> 產量變動狀態(先下降後上升) 2. 輕度颱風 AND 最大總雨量(0-500mm) ==> 當天成交量比較(北>彰>中>南) 3. 輕度颱風 AND 最大總雨量(0-500mm) ==> 產量變動狀態(先下降後上升) AND 當天成交量比較(北>彰>中>南)			

上述關聯性法則解釋：

在侵台颱風與影響產地及成交量資料全部相關記錄中，最小支持度與最小信賴度分別設為10%和95%。

項目組數為2者：

(a) 由表5.13可以清楚得知，侵台颱風強度為輕度颱風、最大總雨量範圍在0-500mm與產量變動狀態(先下降後上升)同時發生次數百分比為0.147(支持度)。

(b) 在侵台颱風為輕度颱風且最大總雨量(0-500mm)下，同時發生颱風侵襲花卉產地後，玫瑰花產地產量變動狀態(先下降後上升)，其概率為1。

(c) 侵台颱風強度為輕度颱風、最大總雨量範圍在0-500mm與當天成交量比較(北>彰>中>南)同時發生次數百分比為0.147(支持度)。

(d) 在侵台颱風為輕度颱風且最大總雨量(0-500mm)下，同時發生颱風侵襲後，當天成

交易比較(北>彰>中>南)，其概率為1。

由以上所述，發現(b)與(d)左邊條件皆一致，故可以有效的進行適當合併產生新的關聯性法則。當侵台颱風為輕度颱風且最大總雨量範圍在0-500mm條件下，同時發生玫瑰產地產量為下降後上升與當天之成交量比較(北>彰>中>南)的可能。可以初步判斷颱風在某些條件下直接影響玫瑰產地產量與得知四家批發市場當天成交量的比較結果。

## 2. 針對最大總雨量地點與移動路徑找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 10%，最小信賴度門檻值 = 95%，關聯性法則如表5.11所示：

表 5.11 關聯性法則(六)

項目組數	支持與信賴度	關聯性法則	值
2	支持度	出現台北地點 玫瑰產地排序(南投>台中>彰化) 第一類型(1,2)	0.118
	信賴度	出現台北地點 AND 第一類型(1,2) ==> 玫瑰產地排序 (南投>台中>彰化)	1
有意義的關聯性法則：			
1. 出現台北地點 AND 第一類型(1,2) ==> 玫瑰產地排序(南投>台中>彰化)			

上述關聯性法則解釋：

在侵台颱風與影響產地資料全部相關記錄中，最小支持度與最小信賴度分別設為10%和95%。

項目組數為2者：

- 由表5.13可以清楚得知，侵台颱風於最大總雨量出現在台北、移動路徑為第一類型(1,2)與玫瑰產地排序(南投>台中>彰化)同時發生次數百分比為0.118(支持度)。
- 在侵台颱風之台北為最大總雨量且第一類型(1,2)移動路徑下，同時發生颱風侵襲花卉產地後，玫瑰產地排序的大小以南投產地最大，台中居中，彰化為最小，其概率為1。

由以上所述，發生玫瑰產地排序(南投>台中>彰化)情況下，可以初步判斷颱風影響玫瑰產地嚴重性為彰化比南投嚴重。

## 3. 針對颱風強度與最大總雨量找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 10%，最小信賴度門檻值 = 95%，關聯性法則如表5.12所示：

表 5.12 關聯性法則(七)

項目組數	支持與信賴度	關聯性法則	值
2	支持度	輕度颱風 最大起伏變動大小(251-500) 最大總雨量 (501-1000mm)	0.147

	信賴度	輕度颱風 AND 最大總雨量(501-1000mm) ==> 最大起伏變動大小(251-500)	1
有意義的關聯性法則：			
1. 輕度颱風 AND 最大總雨量(501-1000mm) ==> 最大起伏變動大小(251-500)			

4. 針對颱風最大總雨量與最大風速找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 10%，最小信賴度門檻值 = 95%，關聯性法則如表5.13所示：

表 5.13 關聯性法則(八)

項目組數	支持與信賴度	關聯性法則	值
2	支持度	最大總雨量(1001-1500mm) 最大風速(32.7-50.9m/s) 成交量變化下降狀態(台北降)	0.118
	信賴度	最大總雨量(1001-1500mm) AND 最大風速 (32.7-50.9m/s) ==> 成交量變化下降狀態(台北降)	1
有意義的關聯性法則：			
1. 最大總雨量(1001-1500mm) AND 最大風速(32.7-50.9m/s) ==> 成交量變化下降狀態(台北降)			

5. 針對最大總雨量地點與移動速度找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 10%，最小信賴度門檻值 = 95%，關聯性法則如表5.14所示：

表 5.14 關聯性法則(九)

項目組數	支持與信賴度	關聯性法則	值
2	支持度	出現嘉義地點 移動速度(0-15km/hr) 當天成交量比較(雄>北>彰>中>南)	0.118
	信賴度	出現嘉義地點 AND 移動速度(0-15km/hr) ==> 當天成交量比較(雄>北>彰>中>南)	1
有意義的關聯性法則：			
1. 最出現嘉義地點 AND 移動速度(0-15km/hr) ==> 當天成交量比較(雄>北>彰>中>南)			

經由以上之有意義關聯性法則剖析後，接著將運用相同方式求出大菊、火鶴花、香水百合、葵百合、非洲菊、洋桔梗、康乃馨、劍蘭之關聯性法則。首先將整理好的大宗花卉資料，予以定義假設並進行整合，接著執行數據轉換(如附錄D-1至D-10)並進行Apriori演算軟體的運作。此研究挖掘出關聯性法則共有「58條」，其相關結果如表5.15所示。

表 5.15 關聯性法則整合表

有意義之關聯性法則(支持度 15%、信賴度 80%)			
玫瑰	1	玫瑰花產地排序	中度颱風 AND 出現台北地點 ==> 玫瑰產地排序(南投>台中>彰化)
			中度颱風 AND 第一類型(1,2) ==> 玫瑰產地排序(南投>台中>彰化)
	2	最大起伏變動量大小	最大總雨量(501-1000mm) AND 出現台北地點 ==> 最大起伏變動量大小(251-500)
			最大總雨量(501-1000mm) AND 最大風速(17.2 以下,17.2-32.6m/s) ==> 最大起伏變動量大小(251-500)
	3	產量變動狀態	最大總雨量(0-500mm) AND 第二類型(3,4) ==> 產量變動狀態(先下降後上升)
			移動方向西 AND 移動速度(16-30Km/hr) ==> 產量變動狀態(先下降後上升)
	4	拍賣量變化下降狀態	中度颱風 AND 最大總雨量(1001-1500mm) ==> 拍賣量變化下降狀態(台北降)
	5	當天成交量比較	最大總雨量(0-500mm) AND 最大風速(17.2 以下,17.2-32.6m/s) ==> 當天成交量比較(北>彰>中>南)
	6	成交量變化下降狀態	最大總雨量(1001-1500mm) AND 移動速度(0-15Km/hr) ==> 成交量變化下降狀態(台北降)
			出現台北地點 AND 最大起伏變動量大小(251-500) ==> 成交量變化下降狀態(台北降)
大菊	1	大菊產地排序	輕度颱風 AND 第二類型(3,4) ==> 大菊產地排序(彰化>雲林>嘉義)
			最大總雨量(0-500mm) AND 移動方向西 ==> 大菊產地排序(彰化>雲林>嘉義)
	2	最大起伏變動量大小	輕度颱風 AND 第二類型(3,4) ==> 最大起伏變動量大小(0-250)
			中度颱風 AND 第二類型(3,4) ==> 最大起伏變動量大小(251-500)
			出現台北地點 AND 第一種中心最低氣壓(980 以上,980 以下且大於 第二種中心最低氣壓(965 以下且大於 945)) ==> 最大起伏變動量大小(0-250)
	3	產量變動狀態	沒有
	4	拍賣量變化下降狀態	沒有
	5	當天成交量比較	最大總雨量(0-500mm) AND 通過海上 ==> 當天成交量比較(彰>南>北>中)
	6	成交量變化下降狀態	出現嘉義地點 AND 移動方向北 ==> 成交量變化下降狀態(台北降)
			成交量變化下降狀態(台北降)

小菊	1	小菊產地排序	沒有
	2	最大起伏變動量大小	最大總雨量(501-1000mm) AND 第一類型(1,2) ==> 最大起伏變動量大小(501 以上)
			出現嘉義地點 AND 第一類型(1,2) ==> 最大起伏變動量大小(501 以上)
			移動方向北 AND 第二種中心最低氣壓(965 以下且大於 945) ==> 最大起伏變動量大小(501 以上)
	3	產量變動狀態	沒有
	4	拍賣量變化下降狀態	沒有
	5	當天成交量比較	沒有
火鶴花	6	成交量變化下降狀態	沒有
	1	火鶴花產地排序	強烈颱風 AND 第一類型(1,2) ==> 火鶴花產地排序(雲林>南投>高雄)
	2	最大起伏變動量大小	出現台北地點 AND 移動速度(0-15Km/hr) ==> 最大起伏變動量大小(501 以上)
			通過海上 AND 移動方向北 ==>最大起伏變動量大小(501 以上)
			移動速度(0-15Km/hr) AND 最大風速(32.7-50.9m/s) ==> 最大起伏變動量大小(501 以上)
	3	產量變動狀態	移動速度(16-30Km/hr) AND 最大風速(17.2 以下,17.2-32.6m/s) ==> 產量變動狀態(持續下降)
	4	拍賣量變化下降狀態	沒有
香水百合	5	當天成交量比較	最大總雨量(501-1000mm) AND 出現台北地點 ==> 當天成交量比較(北>彰>南>中)
	6	成交量變化下降狀態	沒有
	1	香水百合產地排序	強烈颱風 AND 第一類型(1,2) ==> 香水百合產地排序(南投>台中>嘉義)
			出現嘉義地點 AND 移動速度(0-15Km/hr) ==> 香水百合產地排序(南投>台中>嘉義)
	2	最大起伏變動量大小	最大總雨量(0-500mm) AND 移動方向西 ==> 最大起伏變動量大小(501 以上)
	3	產量變動狀態	第一類型(1,2) AND 移動方向西 ==> 產量變動狀態(先上升後下降)
	4	拍賣量變化下降狀態	最大總雨量(501-1000mm) AND 出現台北地點 ==> 拍賣量變化下降狀態(台北降)
香水百合	5	當天成交量比較	輕度颱風 AND 移動速度(16-30Km/hr) ==> 當天成交量比較(北>彰>中>南)
			最大總雨量(501-1000mm) AND 出現台北地點 ==> 當天成交量比較(北>彰>中>南)
	6	成交量變化下降狀態	沒有

蔡百合	1	蔡百合產地排序	沒有
	2	最大起伏變動量大小	中度颱風 AND 移動方向北 ==> 最大起伏變動量大小(0-250)
			最大總雨量(0-500mm) AND 移動速度(16-30Km/hr) ==>
			最大起伏變動量大小(0-250)
	3	產量變動狀態	最大總雨量(501-1000mm) AND 移動方向北 ==>
			最大起伏變動量大小(0-250)
			沒有
非洲菊	4	拍賣量變化下降狀態	沒有
	5	當天成交量比較	輕度颱風 AND 最大總雨量(0-500mm) ==>
	6	成交量變化下降狀態	當天成交量比較(北>彰>南>中)
			沒有
			沒有
	1	非洲菊產地排序	最大總雨量(0-500mm) AND 移動方向北 ==>
	2	最大起伏變動量大小	非洲菊產地排序(彰化>南投>台中)
			第二類型(3,4) AND 移動方向北 ==>
			非洲菊產地排序(彰化>南投>台中)
洋桔梗	3	產量變動狀態	中度颱風 AND 移動方向北 ==> 最大起伏變動量大小(0-250)
			中度颱風 AND 移動速度(16-30Km/hr) ==>
			最大起伏變動量大小(0-250)
	4	拍賣量變化下降狀態	最大總雨量(501-1000mm) AND 移動方向北 ==>
			最大起伏變動量大小(0-250)
			沒有
	5	當天成交量比較	輕度颱風 AND 移動方向北 ==> 當天成交量比較(北>彰>南>中)
洋桔梗	6	成交量變化下降狀態	最大總雨量(0-500mm) AND 最大風速(17.2 以下,17.2-32.6m/s) ==>
			當天成交量比較(北>彰>南>中)
			沒有
	1	洋桔梗產地排序	最大總雨量(0-500mm) AND 移動方向西 ==>
	2	最大起伏變動量大小	洋桔梗產地排序(彰化>嘉義>雲林)
			輕度颱風 AND 移動方向北 ==> 最大起伏變動量大小(0-250)
			中度颱風 AND 移動方向北 ==> 最大起伏變動量大小(0-250)
洋桔梗	3	產量變動狀態	中度颱風 AND 移動速度(0-15Km/hr) ==>
			最大起伏變動量大小(0-250)
			輕度颱風 AND 最大總雨量(501-1000mm) ==>
	4	拍賣量變化下降狀態	產量變動狀態(持續下降)
			最大總雨量(501-1000mm) AND 出現台北地點 ==>
			產量變動狀態(持續下降)
	5	當天成交量比較	沒有
洋桔梗	6	成交量變化下降狀態	沒有

康乃馨	1	康乃馨產地排序	出現嘉義地點 AND 移動速度(0-15Km/hr) ==> 康乃馨產地排序(彰化>南投>台南)
	2	最大起伏變動量大小	中度颱風 AND 最大總雨量(0-500mm) ==> 最大起伏變動量大小(0-250)
			中度颱風 AND 移動方向北 ==> 最大起伏變動量大小(0-250)
	3	產量變動狀態	沒有
	4	拍賣量變化下降狀態	沒有
	5	當天成交量比較	中度颱風 AND 最大總雨量(501-1000mm) ==> (過半休市)
	6	成交量變化下降狀態	沒有
劍蘭	1	劍蘭產地排序	中度颱風 AND 第一類型(1,2) ==> 劍蘭產地排序(台中>彰化>苗栗)
			最大總雨量(501-1000mm) AND 出現台北地點 ==> 劍蘭產地排序(台中>彰化>苗栗)
	2	最大起伏變動量大小	第一類型(1,2) AND 移動方向西 ==> 最大起伏變動量大小(501 以上)
	3	產量變動狀態	通過中部地區 AND 移動速度(16-30Km/hr) ==> 產量變動狀態(持續下降)
	4	拍賣量變化下降狀態	沒有
	5	當天成交量比較	沒有
	6	成交量變化下降狀態	沒有

由表5.15可以很清楚得知，此關聯性法則可分花卉產地排序、最大起伏變動量大小、產量變動狀態、拍賣量變化下降狀態、當天成交量比較、成交量變化下降狀態六種不同形式說明。依每個形式所找出對應之有意義關聯性法則會隨著門檻值之設定不同，而產生不同的結果。此外顯示「沒有」的字眼代表此形式對應之關聯性法則受門檻值的限制而進行刪除無意義的關聯性法則(可參照表5.7)。接著將上述十大大宗花卉關聯性法則予以整合分析，讓法則之分佈情形更明顯化，如表5.16所示。

表 5.16 關聯性法則花種對應變項分佈表

變項 花種	颱風 強度	路徑 分類	最大 雨量	出現 地點	通過 地區	移動 方向	移動 速度	型式 1	型式 2	型式 3	型式 4	型式 5	型式 6
玫瑰													
大菊													
小菊													
火鶴													
香水													
葵百合													
非洲菊													
洋桔梗													
康乃馨													
劍蘭													

由表5.16可得知在於型式1~6之分佈中，玫瑰花資料之探勘能力(所能找出有意義的關聯性法則)，比其他九種花卉還要容易篩選且資料穩定度高。相反的，小菊與葵百合兩種花卉之探勘能力顯得比較弱。扣除玫瑰花、小菊和葵百合之其他七種花卉介於前三者之間。

然而對篩選有意義的關連法則而言，由於所能產生之法則上百條可能含許多不合理或無意義的法則，故應該予以適當控制。或著是和花卉公司相關工作人員討論以了解法則背後的涵義。若能合理解釋這些法則，則可以將這些相關知識與供應單位分享以適當掌握颱風對大宗花卉量的影響。



## 第六章 結論與建議

本章主要可分兩個部分，第一節為研究結論，運用資料轉換服務機制成功將高雄資料超市整合前四大資料超市，完成最終總資料倉儲之建置並說明資料探勘過程與 Apriori 搜尋關聯性法則之結果。第二節為研究建議，以利未來研究方向之用。

### 6.1 研究結論

資訊技術的發展造就近幾年來資料倉儲的技術之進步，在於效率講究的前提下，快速獲取資料並確保其資料正確性是可否否認的事實也是值得進一步探討的課題。

基於高雄批發市場之崛起，在於高雄資料超市之建置完畢後。假若要進行分析不同花卉供應鏈的互動情形，各家花卉資料超市權限之取得是必須的，並且分別蒐集不同資料超市之資料，方能進行資料的分析與探勘。因此使用者對於分析不同花卉供應鏈互動情形之重視程度相當的高。在於完成最終總資料倉儲後，可以透過總資料倉儲的市場維度機制直接一次下載各個花卉供應鏈之資料。時間上的考量，相對的使整個分析資料的過程變的更簡短。

另一方面，透過 Apriori 演算法軟體搜尋之颱風對應大宗花卉關聯性法則之結果，了解到花卉之間的探勘能力之差別與強弱。並根據六大型式(花卉產地排序、最大起伏變動量大小、產量變動狀態、拍賣量變化下降狀態、當天成交量比較、成交量變化下降狀態)篩選出適當之有意義的關聯性法則。當侵襲颱風來臨時，隨著颱風特性之剖析，進行一連串之評估與法則之取捨，使各家花卉批發市場與生產者有所幫助。

本研究總共進行 10 種花卉的演算法運算，挖掘出有意義的關聯性法則有多達 58 條規則(詳細結果請參照表 5.15 關聯性法則整合表)。其詳細分布於 10 種花卉的規則為玫瑰花 10 條(六大型式皆有)、大菊 7 條(產量變動狀態與拍賣量變化狀態沒有挖掘出相關有意義之關連性法則)、小菊 3 條(僅挖掘出最大起伏變動大小相關之關連性法則)、火鶴花 6 條(拍賣量變化下降狀態與成交量變化下降狀態無法挖掘出相關有意義之關連性法則)、香水百合 7 條(僅無法挖掘出成交量變化下降狀態相關之關連性法則)、葵百合 4 條(僅挖掘出最大起伏變動大小與當天成交量比較之相關的關連性法則)、非洲菊 7 條(產量變動狀態、拍賣量變化下降狀態與成交量變化下降狀態無法挖掘出有意義的關連性法則)、洋桔梗 6 條(拍賣量變化下降狀態、當天成交量比較與成交量變化下降狀態沒辦法挖掘出有意義的關連性規則)、康乃馨 4 條(挖掘產量變動狀態、拍賣量變化下降狀態與成交量變化下降狀態並無結果)、劍蘭 4 條(拍賣量變化下降狀態、當天成交量比較與成交量變化下降狀態挖掘不出來有意義的關連性規則)。接著，針對這 58 條有意義的關聯法則當中可以發現幾條有趣的法則，例如花卉種類為玫瑰並且為型 1 法則第二條(參照表 5.15)。由於南投位於山區，彰化為平地，加上台灣南至北有著台灣五大脈之中央山脈擋住，因此侵台颱風(中度颱風與第一類型路徑)對於玫瑰產地的侵襲是不能用直覺去判斷的。花卉種類為大菊並且為型六法則，得知颱風最大總雨量出現在嘉義地區與移動方向為北的情況(現象)下，台北花卉批發市場成交量下降的最為嚴重。由此可剖析出此颱風對於產地與市場皆有影響。花卉種類為劍蘭並且為型三法則，發現當颱風路徑通過中部地區且移動方向為西時，劍蘭產地產量變動狀態為持續下降。雖然有中央山脈擋住，但是此種類型的颱風，還是會造成劍蘭產地量的減少與影響。如果仔細去剖析其它法則，則會發現還有許多有趣的法則在其中。

發現在 Apriori 關聯性法則的探勘過程中，將最小支持度門檻值設愈低，找出的關聯規則也就會跟著多。這是因為愈小的支持度，其滿足條件程度會愈多，且時間上耗費也很大。反之，若是最小支持度門檻值設定的太高，要找出存在的關聯性法則也就更加困難。最後，探討有時找出的關聯性法則。可能與花卉之相關人員可以接受之程度還有一段距離，為了讓所發現的關聯性法則更加有其價值，因此需予以適當篩選法則並與花卉之相關人員討論以供決策之用。



## 6.2 未來研究方向

未來研究層面，可分為兩大層面進行研究。第一個層面是針對總資料倉儲維修自動化，第二個層面是資料探勘關聯性法則再研究。以下將為此兩層面進一步闡述。

### 一、總資料倉儲維修自動化

總資料倉儲為各家資料超市資料之匯整區，利用資料轉換機制，完成對五家資料超市拍賣資料表之轉換。因此總資料倉儲很容易受到下端資料超市之影響。問題發生之頻繁導致總資料倉儲管理員需要花費比各家資料超市管理員還要多時間，在維修上顯的很沒有效率。未來各家資料超市如果能夠更加穩定，將可以運用維護自動化機制，適時進行總資料之監控與維護。如此一來，將可以減輕管理者些許負擔。

### 二、資料探勘關聯性法則再研究

目前本研究為找出颱風對大宗花卉(成交量)的關聯性法則分析，換個思考方向，可以針對成交價來進行分析。雖然大宗花卉以成交量為其依據，成交價卻沒有所對應之主題。但是相信此方向的研究有很大的可行度以及價值。

隨著資料串流(Data Stream)的重要性與日俱增，如何將傳統的高頻項目集(建立關聯性法則之其中一環)探勘問題應用於新興之資料串流環境(Data Stream Environment)中，變成了一個十分有趣且重要的議題。基於資料串流本身的到達速度十分快速並限制記憶體空間隨時保持可利用的條件下，而有其困難度。以花卉批發市場拍賣作業為例，拍賣作業自動化之拍賣速度每筆皆以接近「5」秒的速度完成交易，在此種交易速度之下，如果能夠配合 Apriori 演算軟體本身變項(項目)與筆數新增快速的優勢進一步改良，將可以達到快速有效率的資料探勘。

## 參考文獻

- [1] 邱麗如，「以布林代數為基礎的資料挖掘方法及其應用」，私立靜宜大學資訊管理學系碩士論文，2005。
- [2] 柯珮婕，梁高榮「台中花市的花卉交易資料超市設計」，機械工業雜誌，5月，84-95頁，2005。
- [3] 郭孟涵，梁高榮，「利用線上分析處理技術建構花卉的線上產銷年報」，機械工業雜誌，六月，81-86頁，2002。
- [4] 陳佳佑，「花卉資料倉儲的三種改進方案：權限控管、審計資訊與季節性資料分析」，國立交通大學工業工程研究所碩士論文，2006。
- [5] 陳家瑜，「台灣花卉供應鏈的資料倉儲設計與量測變數回歸應用」，國立交通大學工業工程研究所碩士論文，2006。
- [6] 彭雅琴，「侵台颱風的最低中心氣壓與最大風速之極值統計」，國立中興大學土木工程研究所碩士論文，2001。
- [7] 黃俊端，「設計與實作花卉批發資訊分享熱線整合型審計資訊系統」，國立交通大學工業工程研究所碩士論文，2007。
- [8] Agrawal, R. and Srikant, R., "Fast Algorithms for Mining Association Rules in Large Database", Proceedings of 20th International Conference on Very Large Database, pp.487-499, 1994.
- [9] Amitabha, D., Ng, W. K. and Woon, Y. K., "Association Rule Min-ing: Rapid Association Rule Mining," Pro-ceedings of the Tenth International Confer-ence on Information and Knowledge Man-agement, Oct. 2001.
- [10] Codd, E. F., Franks, S. B., and Salley, C. T., "Providing OLAP to User Analyst: An IT Mandate", www.hyperion.com, 1993.
- [11] Fu, Y., "Data mining Tasks, techniques and applications," IEEEPOTENTIALS, pp.18-20, 1997.
- [12] Gray, P. and Watson, H., Decision Support in the Data Warehouse, Prentice Hall, 1998.
- [13] Han, J. and Kamber, M., Data Mining: Concepts and Techniques, Academic Press, 2001.
- [14] <http://www.aptcn.com/APAGRI/33.nsf>，兩岸農業商務網。
- [15] [http://www.coa.gov.tw/show\\_index.php](http://www.coa.gov.tw/show_index.php)，行政院農業委員會。
- [16] <http://www.cwb.gov.tw/>，中央氣象局。
- [17] [http://www.paps.kh.edu.tw/p5/st\\_homepage/88609-1/%E9%A2%B1%E9%A2%A8%E5%B0%8F%E7%99%BE%E7%A7%91.htm](http://www.paps.kh.edu.tw/p5/st_homepage/88609-1/%E9%A2%B1%E9%A2%A8%E5%B0%8F%E7%99%BE%E7%A7%91.htm)，颱風百科。
- [18] [http://www.tflower.com.tw/member/intranet/tflower\\_web/book/book\\_07.htm](http://www.tflower.com.tw/member/intranet/tflower_web/book/book_07.htm)，台北花卉批發市場網之節慶花。
- [19] [http://www.tflower.com.tw/member/intranet/tflower\\_web/map/map\\_01.htm](http://www.tflower.com.tw/member/intranet/tflower_web/map/map_01.htm)，台北花卉批發市場網之花卉產地分布。
- [20] Inmon, W. H., Building the Data Warehouse, Addition-Wesley, 1992.
- [21] Kaufman, L. and Rousseuw, P. J., Finding Groups in Data: An Introduction to Cluster Analysis, John Wiley & Sons, NY, 1990.
- [22] Kimball, R. and Ross, M., The Data Warehouse Toolkit : The Complete Guide to Dimensional Modeling (2nd Edition), John Wiley & Sons, 2000.
- [23] Kimball, R., "A Dimensional Modeling Manifesto," DataBase Manager System, pp. 58-70, Aug. 1997.
- [24] Park, J. S., Chen, M. S. and Yu, P. S., "An Effective Hash Based Algorithm for Mining

- Association Rules,” Proceedings of the ACM SIGMOD International Conference on Management of Data, San Jose, USA, pp. 175-186, 1995.
- [25] Shenoy, P., Haritsa, R. J., Sudarshan, S., Bhalotia, G., Bawa, M. and Shah, D., “Turbo-charging Vertical Mining of Large Databases,” ACM SIGMOD Record, Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data, May 2000.
- [26] Yen, S. J. and Chen, A. L., "An Efficient Approach to Discovering Knowledge from Large Databases", In Proc. of 4th Int. Conf. on Parallel and Distributed Information Systems, Miami Beach, Florida, USA, Dec., pp. 8-18, 1996.



附錄 A：1996-2005 年侵台颱風一覽表

編號	年份	期間	發生點	結束點	停留 天數	颱風名稱	颱風強度	路徑分類 代號(7)	通過地區	移動 方向	移動 速度 (Km/hr)	中心最低 氣壓(百 帕-hpa)	最大風速 (米(m)/秒 (s))
1	2005	9-10	9 月 30 日	10 月 3 日	4	LONGWANG 龍王	強烈颱風	2	中部	W	23	925	51
2	2005	8-9	8 月 30 日	9 月 1 日	3	TALIM 泰利	強烈颱風	2	中部	W	21	920	53
3	2005	8	8 月 3 日	8 月 6 日	4	MATSA 馬莎	中度颱風	1	北部海上	NW	14	955	40
4	2005	7	7 月 16 日	7 月 20 日	5	HAITANG 海棠	強烈颱風	1	北部	WNW	15	912	55
5	2004	12	12 月 3 日	12 月 4 日	2	NANMADOL 南瑪都	中度颱風	6	南部	ENE	27	958	38
6	2004	10	10 月 24 日	10 月 25 日	2	NOCK-TEN 納坦	中度颱風	1	東北角	NNW	21	945	43
7	2004	9	9 月 11 日	9 月 12 日	2	HAIMA 海馬	輕度颱風	4	東北海上	N	14	998	18
8	2004	8	8 月 23 日	8 月 26 日	4	AERE 艾利	中度颱風	1	北部海上	W	11	960	38
9	2004	6-7	6 月 29 日	7 月 2 日	4	MINDULLE 敏督利	中度颱風	4	東北部	N	10	942	45
10	2003	11	11 月 2 日	11 月 3 日	2	MELOR 米勒	輕度颱風	4	東部近海	NNE	10	987	25
11	2003	8-9	8 月 31 日	9 月 2 日	3	DUJUAN 杜鵑	中度颱風	3	南部海上	W	26	950	43
12	2003	8	8 月 2 日	8 月 4 日	3	MORAKOT 莫拉克	輕度颱風	3	南部	W	19	990	23
13	2002	9	9 月 4 日	9 月 8 日	5	SINLAKU 辛樂克	中度颱風	1	北部海上	W	11	955	40
14	2002	7	7 月 9 日	7 月 10 日	2	NAKRI 納克莉	輕度颱風	6	中部北部	NE	15	987	18
15	2001	9	9 月 24 日	9 月 28 日	5	LEKIMA 利奇馬	中度颱風	2	中南部	NW	5	965	35
16	2001	9	9 月 8 日	9 月 10 日	3	NARI 納莉	中度颱風	7	北部海上	NE	5	970	33
17	2001	9	9 月 13 日	9 月 19 日	7	NARI 納莉	中度颱風	7	北中南部	WSW	6	960	40
18	2001	7	7 月 28 日	7 月 31 日	4	TORAJI 桃芝	中度颱風	2	北部	NW	17	962	38
19	2001	7	7 月 10 日	7 月 11 日	2	TRAMI 潭美	輕度颱風	3	南部	NW	10	995	20
20	2001	7	7 月 3 日	7 月 5 日	3	UTOR 尤特	中度颱風	3	南部海上	WNW	29	960	38
21	2001	6	6 月 22 日	6 月 24 日	3	CHEBI 奇比	中度颱風	5	西部近海	NNW	25	965	35
22	2001	5	5 月 11 日	5 月 13 日	3	CIMARON 西馬隆	輕度颱風	4	東部近海	NNE	19	990	23

23	2000	10-1 1	10 月 30 日	11 月 1 日	3	XANGSANE 象神	中度颱風	4	東部近海	NNE	15	960	38
24	2000	9	9 月 8 日	9 月 10 日	3	BOPHA 寶發	輕度颱風	7	東部海上	SSW	17	990	23
25	2000	8	8 月 28 日	8 月 30 日	3	PRAPIROON 巴比倫	輕度颱風	1	東北部海上	NW	19	970	30
26	2000	8	8 月 21 日	8 月 23 日	3	BILIS 碧利斯	強烈颱風	2	中部	NW	22	930	53
27	2000	7	7 月 8 日	7 月 9 日	2	KAI-TAK 啟德	中度颱風	4	東部	NNE	17	965	35
28	1999	10	10 月 6 日	10 月 10 日	5	DAN 丹恩	中度颱風	5	西部海上	N	19	968	38
29	1999	6	6 月 5 日	6 月 6 日	2	MAGGIE 瑪姬	中度颱風	3	南部海上	WNW	20	965	38
30	1998	10	10 月 25 日	10 月 27 日	3	BABS 芭比絲	中度颱風	5	台灣海峽	NE	13	965	35
31	1998	10	10 月 14 日	10 月 16 日	3	ZEB 瑞伯	強烈颱風	4	東部海上	N	15	920	55
32	1998	9	9 月 27 日	9 月 29 日	3	YANNI 楊妮	輕度颱風	4	東部海上	NNW	13	996	23
33	1998	8	8 月 3 日	8 月 5 日	3	OTTO 奧托	輕度颱風	2	中部	NW	25	985	30
34	1998	7	7 月 8 日	7 月 10 日	3	NICHOLE 妮蔻兒	輕度颱風	5	西岸中南部	NE	5	998	18
35	1997	8	8 月 27 日	8 月 29 日	3	AMBER 安珀	中度颱風	2	中部	NW	18	945	48
36	1997	8	8 月 16 日	8 月 18 日	3	WINNIE 溫妮	中度颱風	7	北部海上	WNW	19	945	43
37	1996	7-8	7 月 29 日	8 月 1 日	4	HERB 賀伯	強烈颱風	1	北部	WNW	19	930	53
38	1996	7	7 月 25 日	7 月 27 日	3	GLORIA 葛樂禮	中度颱風	3	南部	NW	15	965	35
39	1996	5	5 月 21 日	5 月 24 日	4	CAM 凱姆	輕度颱風	7	南部海上	NE	15	995	20

附錄 B：1996-2005 年侵台颱風對應節慶分析表

編號	年	颱風名稱	發生點	結束點	節慶一	節慶二	節慶三
1	1996	CAM 凱姆	5 月 21 日	5 月 24 日	後 3 天浴佛節		
2	1996	GLORIA 葛樂禮	7 月 25 日	7 月 27 日	後 8 天觀音得道		
3	1996	HERB 賀伯	7 月 29 日	8 月 1 日	後 4 天觀音得道	後 8 天關公生	
4	1997	WINNIE 溫妮	8 月 16 日	8 月 18 日	前 8 天父親節	前 7 天七夕情人節	後 1 天中元節
5	1997	AMBER 安珀	8 月 27 日	8 月 29 日	前 10 天中元節		
6	1998	NICHOLE 妮蔻兒	7 月 8 日	7 月 10 日	前 10 天端午節		
7	1998	OTTO 奧托	8 月 3 日	8 月 5 日	後 5 天父親節	後 7 天觀音得道	
8	1998	YANNI 楊妮	9 月 27 日	9 月 29 日	後 1 天教師節	後 8 天忠秋節	
9	1998	ZEB 瑞伯	10 月 14 日	10 月 16 日	前 4 天國慶	前 9 天中秋節	
10	1998	BABS 芭比絲	10 月 25 日	10 月 27 日	後 3 天重陽節		
11	1999	MAGGIE 瑪姬	6 月 5 日	6 月 6 日	N		
12	1999	DAN 丹恩	10 月 6 日	10 月 10 日	後 4 天國慶		
13	2000	KAI-TAK 啟德	7 月 8 日	7 月 9 日	N		
14	2000	BILIS 碧利斯	8 月 21 日	8 月 23 日	前 7 天中元節		
15	2000	PRAPIROON 巴比倫	8 月 28 日	8 月 30 日	N		
16	2000	BOPHA 寶發	9 月 8 日	9 月 10 日	後 4 天中秋節		
17	2000	XANGSANE 象神	10 月 30 日	11 月 1 日	N		
18	2001	CIMARON 西馬隆	5 月 11 日	5 月 13 日	後 2 天母親節	後 4 天媽祖生	
19	2001	CHEBI 奇比	6 月 22 日	6 月 24 日	後 3 天端午節		
20	2001	UTOR 尤特	7 月 3 日	7 月 5 日	前 8 天端午節	前 8 天觀音得道	
21	2001	TRAMI 潭美	7 月 10 日	7 月 11 日	N		
22	2001	TORAJI 桃芝	7 月 28 日	7 月 31 日	後 11 天父親節	後 11 天觀音得道	
23	2001	NARI 納莉	9 月 13 日	9 月 19 日	前 11 天中元節		
24	2001	NARI 納莉	9 月 8 日	9 月 10 日	前 6 天中元節		

25	2001	LEKIMA 利奇馬	9 月 24 日	9 月 28 日	後 4 天教師節	後 7 天中秋節	
26	2002	NAKRI 納克莉	7 月 9 日	7 月 10 日	N		
27	2002	SINLAKU 辛樂克	9 月 4 日	9 月 8 日	N		
28	2003	MORAKOT 莫拉克	8 月 2 日	8 月 4 日	後 2 天七夕情人節	後 6 天父親節	後 10 天中元節
29	2003	DUJUAN 杜鵑	8 月 31 日	9 月 2 日	後 11 天中秋節		
30	2003	MELOR 米勒	11 月 2 日	11 月 3 日	N		
31	2004	MINDULLE 敏督利	6 月 29 日	7 月 2 日	前 7 天端午節		
32	2004	AERE 艾利	8 月 23 日	8 月 26 日	前 1 天七夕情人節	後 7 天中元節	
33	2004	HAIMA 海馬	9 月 11 日	9 月 12 日	N		
34	2004	NOCK-TEN 納坦	10 月 24 日	10 月 25 日	前 2 天重陽節	後 8 天觀音出家	
35	2004	NANMADOL 南瑪都	12 月 3 日	12 月 4 日	N		
36	2005	HAITANG 海棠	7 月 16 日	7 月 20 日	後 7 天觀音得道		
37	2005	MATSA 馬莎	8 月 3 日	8 月 6 日	後 5 天父親節	後 8 天七夕情人節	
38	2005	TALIM 泰利	8 月 30 日	9 月 1 日	前 11 天中元節		
39	2005	LONGWANG 龍王	9 月 30 日	10 月 3 日	前 2 天教師節	後 10 天國慶紀念日	

附錄 C：1996-2005 年侵台颱風數據轉換表

編號	年份	季節	旬	停留 天數	近節慶個 數(5 天)	節慶時 機(前後)	最短近節 慶天數	颱風 強度	路徑分類 代號	最大總 雨量	出現 地點	通過 地區	移動方 向	移動 速度	中心最 低氣壓	最大 風速
1	1	1	3	2	1	2	1	1	4	1	4	7	1	1	1	1
2	1	2	3	2	3	4	3	2	2	1	4	6	1	1	2	2
3	1	2	3	2	1	2	2	3	1	4	5	1	2	2	3	3
4	2	2	2	2	1	2	1	2	4	2	1	7	2	2	3	2
5	2	2	3	2	3	4	3	2	1	1	3	5	1	2	3	2
6	3	2	1	2	3	4	3	1	3	1	7	4	1	1	1	1
7	3	2	1	2	1	2	2	1	1	1	5	5	1	2	1	1
8	3	3	3	2	1	2	1	1	2	1	1	7	1	1	1	1
9	3	3	2	2	1	1	2	3	2	2	1	7	1	1	3	3
10	3	3	3	2	1	2	1	2	3	2	2	7	1	1	2	2
11	4	2	1	1	3	4	3	2	2	1	4	7	2	2	2	2
12	4	3	1	3	1	2	2	2	3	1	4	7	1	2	1	2
13	5	2	1	1	3	4	3	2	2	2	8	3	1	2	2	2
14	5	2	3	2	3	4	3	3	1	2	4	5	1	2	3	3
15	5	2	3	2	3	4	3	1	1	2	1	7	1	2	1	1
16	5	3	1	2	1	2	2	1	4	2	1	7	4	2	1	1
17	5	3	3	2	3	4	3	2	2	3	1	3	1	1	2	2
18	6	1	2	2	2	3	1	1	2	1	2	3	1	2	1	1
19	6	2	3	2	1	2	1	2	3	1	6	4	1	2	2	2
20	6	2	1	2	3	4	3	2	2	1	3	7	2	2	2	2
21	6	2	1	1	3	4	3	1	2	2	7	6	1	1	1	1
22	6	2	3	2	3	4	3	2	1	2	5	1	1	2	2	2
23	6	3	2	3	3	4	3	2	4	3	1	1	2	1	2	2

24	6	3	1	2	3	4	2	2	4	3	1	7	1	1	1	2
25	6	3	3	3	1	2	2	2	1	2	1	5	1	1	2	2
26	7	2	1	1	3	4	3	1	3	2	1	1	1	1	1	1
27	7	3	1	3	3	4	3	2	1	1	1	7	2	1	2	2
28	8	2	1	2	1	2	1	1	2	1	4	6	2	2	1	1
29	8	2	3	2	3	4	3	2	2	1	4	7	2	2	2	2
30	8	3	1	1	3	4	3	1	2	1	8	3	1	1	1	1
31	9	2	3	2	3	4	3	2	2	3	5	2	1	1	3	2
32	9	2	3	2	1	1	1	2	1	2	5	7	2	1	2	2
33	9	3	2	1	3	4	3	1	2	2	1	7	1	1	1	1
34	9	3	3	1	1	1	1	2	1	1	1	2	1	2	3	2
35	9	4	1	1	3	4	3	2	3	1	3	6	3	2	2	2
36	10	2	2	3	3	4	3	3	1	3	5	1	2	1	3	3
37	10	2	1	2	1	2	2	2	1	2	5	7	1	1	2	2
38	10	2	3	1	3	4	3	3	1	2	5	5	2	2	3	3
39	10	3	3	2	1	1	1	3	1	1	3	5	2	2	3	3

附錄 D-1 ：1998-2005 年大宗花卉-玫瑰數據轉換表

玫瑰平均 供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
5	3	4	1	1	3	1	1	2
5	1	4	1	1	1	1	1	1
5	3	3	7	1	5	7	1	6
2	3	2	1	1	1	1	1	1
6	1	1	1	2	1	1	3	1
6	3	1	1	7	5	1	6	6
6	1	1	1	2	2	1	2	2
6	1	1	1	7	5	1	6	6
2	3	2	3	3	1	3	1	1
3	3	2	1	1	1	1	1	1
2	1	2	1	1	3	1	2	3
2	3	2	1	1	1	1	1	1
2	3	2	1	1	3	1	1	3
2	3	3	1	7	1	1	6	1
2	2	2	1	1	1	2	2	1
3	1	2	1	1	1	2	2	1
2	1	4	1	7	1	1	6	1
3	3	3	2	2	1	2	2	1
3	3	4	1	7	1	1	6	1
3	2	4	1	2	5	1	2	6
2	3	2	1	2	1	2	2	1
3	2	3	1	2	3	1	2	3
3	3	2	4	4	4	1	1	4
3	3	3	4	4	3	1	1	3
3	3	1	7	5	5	7	4	6
3	3	2	5	5	1	4	4	1
3	3	4	7	5	5	7	4	6
3	1	2	6	6	3	5	4	1
3	2	1	7	6	5	7	4	6
3	3	2	6	6	3	6	5	3
2	3	4	5	5	3	4	4	1
3	3	2	5	6	4	5	4	5
4	3	3	6	5	1	4	4	1
3	3	4	6	6	5	6	5	6

附錄 D-2 ：1998-2005 年大宗花卉-大菊數據轉換表

大菊平均 供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
1	2	4	2	3	2	2	3	4
1	2	3	2	1	1	2	1	1
1	1	3	7	1	6	7	1	6
2	2	2	1	2	4	2	2	4
2	1	1	1	1	1	1	1	1
3	2	1	1	7	6	1	7	6
2	2	3	1	1	2	1	1	2
2	2	1	1	1	6	1	1	6
1	1	4	2	2	1	2	2	2
4	2	3	1	2	1	1	3	2
1	1	2	1	1	6	1	1	6
1	2	3	2	2	2	2	2	2
1	1	2	1	1	1	1	1	1
5	3	2	1	7	1	2	7	1
1	2	3	1	1	4	1	1	4
1	1	4	2	2	1	2	2	1
4	3	2	1	7	6	1	7	6
1	3	2	2	1	1	1	1	6
6	1	1	2	7	1	2	7	1
4	3	4	1	2	4	1	1	4
2	1	3	1	2	4	1	2	4
1	3	3	1	1	1	1	1	4
1	1	2	3	3	2	3	3	4
1	1	2	1	1	4	1	1	4
5	2	1	7	4	6	7	4	6
5	2	3	4	4	1	4	4	1
6	2	1	7	4	6	7	4	6
1	1	4	4	4	3	4	4	3
1	2	1	7	5	6	7	5	6
2	1	3	6	4	3	4	4	3
1	3	2	4	6	3	4	5	3
1	2	3	5	7	2	5	7	5
1	2	3	4	4	1	4	4	3
1	2	3	5	6	3	5	6	3

類別	編號	事件	假設項目					單位		
大菊 產地供應	1	產地供應 排序	花卉產地供應量排序 (前三名)	(1)	彰雲嘉	彰化→雲林→嘉義		N		
				(2)	彰雲屏	彰化→雲林→屏東				
			定義：颱風對應花卉產地 地相關供應縣市之前 三名排序	(3)	彰雲投	彰化→雲林→南投				
				(4)	彰雲中	南投→台中→台北				
				(5)	彰中雲	彰化→台中→雲林				
				(6)	彰雲北	彰化→雲林→台北				
	2	產地供應 波動	產量供應波動狀態	(1)	上升狀態	持續上升		N		
						先上升後下降				
			定義：颱風期間對應花卉 產量供應波動狀態	(2)	下降狀態	持續下降				
						先下降後上升				
3	最大起伏變動大小量	0~250、251~500、500 以上					件			
								定義：颱風期間對應花卉 供應最大起伏波動 大小(量)		
類別	編號	事件	假設項目					單位		
數 量	1	拍賣量	颱風前一天 與颱風當天	(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件		
				(2)	彰>北>南>中					
				(3)	南>彰>北>中					
				(4)	彰>雄>南>北>中					
				(5)	雄>彰>南>北>中					
				(6)	南>北>彰>雄>中					
				(7)	過半休市					
		拍賣量 變化狀態		(1)	北(降)		降：下降程度為四大 市場之最	把		
				(2)	彰(降)					
				(3)	中(降)					
				(4)	南(降)					
				(5)	無法判斷					
				(1)	彰>南>北>中				北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
				成交量	颱風前一天	(2)				
	(3)	南>彰>北>中								
	(4)	彰>雄>南>北>中								
	(5)	雄>彰>南>北>中								
	(6)	南>雄>彰>北>中								
	(7)	過半休市								
	颱風當天	(1)	彰>南>北>中							
		(2)	彰>北>南>中							
		(3)	南>彰>北>中							
		(4)	彰>雄>南>北>中							
		(5)	南>彰>雄>北>中							
	成交量 變化狀態		(1)		北(降)	(4)	南(降)			
			(2)	彰(降)	(5)	雄(降)				
			(3)	中(降)	(6)	無法判斷				

附錄 D-3 ：1998-2005 年大宗花卉-小菊數據轉換表

小菊平均 供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
5	2	4	3	2	4	3	2	4
5	3	3	1	1	4	1	1	1
5	2	3	6	1	5	6	1	6
3	3	2	1	1	1	1	1	1
3	2	3	1	1	1	1	1	1
5	3	1	1	7	5	1	7	6
1	3	3	1	3	2	1	3	2
5	3	1	1	7	5	1	7	6
1	2	3	2	2	1	6	7	3
1	3	3	1	2	4	1	2	4
6	1	2	1	1	1	1	1	1
1	3	3	2	2	2	2	2	2
1	2	3	1	1	2	1	1	2
3	3	2	1	7	1	1	7	1
1	3	3	1	1	2	1	1	2
1	1	3	2	2	1	2	2	1
1	3	2	1	7	5	1	7	6
4	3	2	1	1	2	1	1	6
1	2	1	1	7	4	2	7	6
5	3	4	1	1	1	1	1	1
1	2	3	1	2	2	1	2	2
1	3	3	1	1	4	1	1	2
1	2	2	2	1	1	6	1	1
4	1	2	1	1	2	1	1	4
5	3	1	6	4	5	6	4	6
1	3	3	4	4	1	4	4	1
4	3	1	6	4	5	6	4	6
6	2	3	5	5	4	4	3	4
1	1	1	6	5	5	6	3	6
1	1	3	4	4	4	4	4	4
1	3	2	4	6	3	4	6	3
1	3	3	4	4	2	6	6	4
5	3	3	4	6	1	4	6	6
1	3	4	5	5	5	5	5	6

類別	編號	事件	假設項目					單位	
小菊 產地供應	1	產地供應 排序	花卉產地供應量排序 (前三名)	(1)	彰雲中	彰化→雲林→台中		N	
				(2)	彰雲嘉	彰化→雲林→嘉義			
			定義：颱風對應花卉產地 地相關供應縣市之前 三名排序	(3)	彰嘉中	彰化→嘉義→台中			
				(4)	彰中投	彰化→台中→南投			
				(5)	彰嘉雲	彰化→嘉義→雲林			
				(6)	彰中嘉	彰化→台中→嘉義			
	2	產地供應 波動	產量供應波動狀態	(1)	上升狀態	持續上升		N	
						先上升後下降			
			定義：颱風期間對應花卉 產量供應波動狀態	(2)	下降狀態	持續下降			
						先下降後上升			
3	最大起伏變動大小量 定義：颱風期間對應花卉 供應最大起伏波動 大小(量)	0~250、251~500、500 以上					件		
類別	編號	事件	假設項目					單位	
數 量	1	拍賣量	颱風前一天 與颱風當天	(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件	
				(2)	彰>北>南>中				
				(3)	南>彰>北>中				
				(4)	彰>雄>南>北>中				
				(5)	雄>彰>南>北>中				
				(6)	南>彰>雄>北>中				
				(7)	過半休市				
		拍賣量 變化狀態	(1)	北(降)		降：下降程度為四 大市場之最	把		
			(2)	彰(降)					
			(3)	中(降)					
			(4)	南(降)					
			(5)	無法判斷					
			(1)	彰>南>北>中				北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
			2	成交量	颱風前一天				
	(3)	南>彰>北>中							
	(4)	彰>雄>南>北>中							
	(5)	雄>彰>南>北>中							
	(6)	過半休市							
	颱風當天	(1)				彰>南>北>中			
		(2)			彰>北>南>中				
		(3)			南>彰>北>中				
		(4)			彰>雄>南>北>中				
		(5)			雄>彰>南>北>中				
		(6)			南>彰>雄>北>中				
		(7)			過半休市				
	成交量 變化狀態	(1)			北(降)	(4)	南(降)		降：下降程度為四 大市場之最
		(2)		彰(降)	(5)	雄(降)			
		(3)		中(降)	(6)	無法判斷			

附錄 D-4 ：1998-2005 年大宗花卉-火鶴花數據轉換表

火鶴花平均供應量 (前三名排序)	最大起伏 變動大小	產量波動	颱風前一天拍賣量 比較	颱風當天拍賣量 比較	拍賣量變化量狀態	颱風前一天成交量 比較	颱風當天成交量 比較	成交量變化量狀態
2	3	4	2	2	2	2	2	2
2	3	3	2	2	3	2	2	1
2	3	2	6	3	5	6	3	6
2	3	3	2	2	2	2	2	2
5	1	2	3	3	1	3	3	1
2	2	1	2	7	5	2	7	6
1	3	2	2	2	5	2	3	6
6	1	1	2	7	5	2	7	6
3	3	3	2	2	1	2	2	2
1	3	3	2	2	1	2	2	1
2	1	3	2	2	5	2	2	6
2	3	3	2	2	1	2	2	1
6	3	3	2	2	3	2	2	1
1	3	2	6	7	5	6	7	6
1	3	3	2	2	1	2	2	1
1	1	1	2	2	1	2	2	1
1	2	4	3	7	5	3	7	6
1	3	2	1	1	2	1	1	6
1	3	1	2	7	1	2	7	1
1	3	4	2	2	3	2	2	1
1	1	3	2	2	1	2	2	1
6	3	2	1	1	5	6	7	6
6	3	3	2	3	1	3	3	1
1	2	1	1	7	2	1	7	4
1	2	1	6	5	5	6	7	6
1	3	1	5	6	5	5	5	6
3	3	1	6	4	5	6	4	6
1	3	2	5	6	1	5	5	5
1	2	1	6	6	5	6	5	6
2	3	3	5	4	4	5	4	4
3	3	2	4	4	1	4	4	1
1	3	3	4	6	2	4	5	2
3	3	2	4	4	5	4	4	6
3	3	3	6	6	1	6	6	3

類別	編號	事件	假設項目					單位				
火鶴花產地供應	1	產地供應排序	花卉產地供應量排序(前三名)	(1)	雲投中	雲林→南投→台中		N				
				(2)	投雲中	南投→雲林→台中						
			定義：颱風對應花卉產地相關供應縣市之前三名排序	(3)	雲投雄	雲林→南投→高雄						
				(4)	投中雲	南投→台中→雲林						
				(5)	投雲嘉	南投→雲林→嘉義						
				(6)	雲中投	雲林→台中→南投						
	2	產地供應波動	產量供應波動狀態	(1)	上升狀態	持續上升		N				
			先上升後下降									
			定義：颱風期間對應花卉產量供應波動狀態	(2)	下降狀態	持續下降						
			先下降後上升									
3	最大起伏變動大小量	0~250、251~500、500 以上					件					
								定義：颱風期間對應花卉供應最大起伏波動大小(量)				
類別	編號	事件	假設項目					單位				
數量	1	拍賣量	颱風前一天與颱風當天	(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件				
				(2)	北>彰>南>中							
				(3)	北>南>中>彰							
				(4)	北>彰>南>雄>中							
				(5)	北>雄>南>彰>中							
				(6)	北>南>雄>彰>中							
				(7)	過半休市							
		拍賣量變化狀態	(1)	北(降)		降：下降程度為四大市場之最	把					
			(2)	彰(降)								
			(3)	中(降)								
			(4)	南(降)								
			(5)	無法判斷								
			2	成交量	颱風前一天			(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
								(2)	北>彰>南>中			
	(3)	北>中>彰>南										
	(4)	北>彰>南>中>雄										
	(5)	北>南>彰>中>雄										
	(6)	過半休市										
	颱風當天	(1)			彰>南>北>中							
		(2)			北>彰>南>中							
		(3)			北>南>彰>中							
		(4)			北>彰>南>雄>中							
		(5)	北>南>雄>彰>中									
		(6)	南>北>雄>彰>中									
		(7)	過半休市									
	成交量變化狀態	(1)	北(降)	(4)	南(降)		降：下降程度為四大市場之最	把				
		(2)	彰(降)	(5)	雄(降)							
		(3)	中(降)	(6)	無法判斷							

附錄 D-5 ：1998-2005 年大宗花卉-香水百合數據轉換表

香水百合 平均供應 量(前三 名排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
2	2	3	1	2	2	2	2	2
2	2	3	1	2	1	2	2	1
1	2	2	6	2	5	6	2	6
1	2	2	1	2	1	2	2	4
1	1	3	2	2	1	2	2	1
1	1	1	1	7	5	2	6	6
1	1	3	3	3	3	3	3	3
5	1	1	1	7	5	2	6	6
1	2	3	2	2	1	2	2	2
5	1	2	2	2	1	2	2	2
3	2	1	2	2	1	2	2	1
1	2	3	2	2	1	2	2	2
1	3	3	2	2	3	2	2	3
2	3	2	2	7	1	2	6	6
2	3	3	1	2	2	2	2	2
2	1	3	2	2	1	2	2	1
3	2	4	1	7	5	2	6	6
2	2	4	1	1	1	1	1	4
2	2	4	2	7	5	2	6	6
2	1	4	2	2	1	2	2	3
2	2	3	1	2	4	2	2	4
2	3	2	1	2	1	3	2	1
2	3	3	2	2	3	2	2	2
1	3	2	1	2	1	2	2	2
1	3	1	6	7	5	6	5	6
1	3	2	4	4	1	4	4	1
1	1	2	6	4	5	6	4	6
3	1	1	5	4	1	5	4	6
1	3	1	6	4	5	6	4	6
6	1	3	4	7	2	4	4	6
1	3	2	4	4	3	4	4	3
1	3	3	4	5	2	4	5	2
1	3	2	4	6	1	4	4	3
1	3	2	4	4	5	4	4	6

類別	編號	事件	假設項目					單位							
香水百合產地供應	1	產地供應排序	花卉產地供應量排序 (前三名)	(1)	投中嘉	南投→台中→嘉義		N							
				(2)	投嘉中	南投→嘉義→台中									
			定義：颱風對應花卉產地 相關供應縣市之前 三名排序	(3)	中中彰	中縣→中市→彰化									
				(4)	投嘉中	南投→嘉義→台中									
				(5)	投彰中	南投→彰化→台中									
				(6)	中投嘉	台中→南投→嘉義									
	2	產地供應波動	產量供應波動狀態		(1)	上升狀態	持續上升		N						
			定義：颱風期間對應花卉 產量供應波動狀態				先上升後下降								
					(2)	下降狀態	持續下降								
							先下降後上升								
3		最大起伏變動大小量		0~250、251~500、500 以上				件							
		定義：颱風期間對應花卉 供應最大起伏波動 大小(量)													
類別	編號	事件	假設項目					單位							
數 量	1	拍賣量	颱風前一天 與颱風當天	(1)	北>彰>南>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄		件						
				(2)	北>彰>中>南										
				(3)	北>南>彰>中										
				(4)	北>彰>雄>南>中										
				(5)	北>雄>南>彰>中										
				(6)	北>南>彰>雄>中										
				(7)	過半休市										
		拍賣量 變化狀態		(1)	北(降)		降：下降程度為四 大市場之最		把						
				(2)	彰(降)										
				(3)	中(降)										
				(4)	南(降)										
				(5)	無法判斷										
				2	成交量	颱風前一天				(1)	彰>北>南>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄		件
										(2)	北>彰>中>南				
										(3)	北>南>彰>中				
	(4)	北>彰>雄>南>中													
	(5)	北>雄>彰>南>中													
	(6)	過半休市													
	颱風當天	(1)	彰>北>中>南												
		(2)	北>彰>中>南												
		(3)	北>南>彰>中												
		(4)	北>彰>雄>南>中												
		(5)	北>雄>彰>南>中												
		(6)	過半休市												
	成交量 變化狀態		(1)	北(降)	(4)	南(降)		降：下降程度為四 大市場之最	把						
			(2)	彰(降)	(5)	雄(降)									
			(3)	中(降)	(6)	無法判斷									

附錄 D-6：1998-2005 年大宗花卉-葵百合數據轉換表

葵百合平均供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
1	2	4	2	3	2	2	2	2
3	1	3	2	3	1	3	2	1
5	2	2	6	2	5	6	2	6
1	1	3	3	3	1	3	2	1
1	1	4	3	3	1	3	2	1
3	1	1	3	7	5	3	6	6
1	1	2	3	2	1	3	1	1
2	1	1	2	7	5	2	6	6
3	1	2	3	3	1	3	2	1
2	1	4	2	7	2	1	6	2
2	1	4	3	7	2	3	2	2
6	1	3	3	3	1	3	2	4
3	1	3	2	2	2	2	2	1
2	1	2	3	7	5	7	6	6
5	1	4	3	2	1	3	1	1
2	1	3	3	3	1	3	2	1
6	1	4	3	7	5	3	6	6
1	2	2	1	1	1	1	1	4
1	1	4	3	7	5	7	6	6
2	1	4	3	3	1	3	2	3
1	1	3	3	3	4	3	2	4
2	2	2	3	3	3	7	2	3
1	1	3	3	3	4	2	2	4
1	1	2	3	3	1	3	2	3
3	1	1	6	5	5	7	6	6
1	1	2	4	6	1	4	3	1
1	2	4	6	6	5	7	4	6
1	1	3	5	5	5	5	5	6
1	2	1	6	4	5	7	3	6
3	1	3	4	6	3	4	4	3
1	1	4	4	5	3	4	3	3
4	2	3	6	5	2	6	5	2
3	1	2	5	5	3	4	4	3
1	2	2	4	4	5	4	5	6

類別	編號	事件	假設項目						單位
葵百合產地供應	1	產地供應排序	花卉產地供應量排序(前三名)	(1)	投嘉中	南投→嘉義→台中		N	
				(2)	投嘉彰	南投→嘉義→彰化			
			定義:颱風對應花卉產地相關供應縣市之前三名排序	(3)	投中嘉	南投→台中→嘉義			
				(4)	投嘉中	南投→嘉義→台中			
				(5)	投嘉苗	南投→嘉義→苗栗			
				(6)	投彰嘉	南投→彰化→嘉義			
	2	產地供應波動	產量供應波動狀態		(1)	上升狀態	持續上升		N
			定義:颱風期間對應花卉產量供應波動狀態				先上升後下降		
					(2)	下降狀態	持續下降		
							先下降後上升		
3		最大起伏變動大小量		0~250、251~500、500 以上				件	
		定義:颱風期間對應花卉供應最大起伏波動大小(量)							
類別	編號	事件	假設項目						單位
數量	1	拍賣量	颱風前一天與颱風當天	(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄		件
				(2)	彰>北>南>中				
				(3)	北>彰>南>中				
				(4)	北>彰>南>雄>中				
				(5)	雄>彰>北>南>中				
				(6)	北>南>彰>雄>中				
				(7)	過半休市				
		拍賣量變化狀態		(1)	北(降)		降:下降程度為四大市場之最		把
				(2)	彰(降)				
				(3)	中(降)				
				(4)	南(降)				
				(5)	無法判斷				
	2	成交量	颱風前一天	(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄		件
				(2)	彰>北>南>中				
				(3)	北>彰>南>中				
				(4)	北>彰>南>雄>中				
				(5)	雄>北>彰>南>中				
				(6)	彰>北>雄>南>中				
				(7)	過半休市				
			颱風當天	(1)	彰>北>南>中		降:下降程度為四大市場之最		把
				(2)	北>彰>南>中				
				(3)	北>彰>雄>南>中				
				(4)	北>南>雄>彰>中				
				(5)	北>雄>彰>南>中				
				(6)	過半休市				
			成交量變化狀態		(1)	北(降)	(4)	南(降)	降:下降程度為四大市場之最
					(2)	彰(降)	(5)	雄(降)	
					(3)	中(降)	(6)	無法判斷	

附錄 D-7：1998-2005 年大宗花卉-非洲菊數據轉換表

非洲菊平均供應量 (前三名排序)	最大起伏 變動大小	產量波動	颱風前一天拍賣量 比較	颱風當天拍賣量 比較	拍賣量變化量狀態	颱風前一天成交量 比較	颱風當天成交量 比較	成交量變化量狀態
1	2	4	2	2	5	2	2	6
1	1	3	2	2	1	2	2	1
1	1	2	6	2	5	6	2	6
1	1	3	2	2	4	3	2	4
1	1	2	2	2	3	2	2	3
1	1	1	1	6	5	2	6	6
1	1	1	3	2	4	3	2	6
1	1	1	1	6	5	2	6	6
1	2	3	1	6	1	2	2	2
1	1	3	2	2	1	2	2	2
1	1	2	1	2	4	2	2	3
1	1	3	2	2	2	2	2	2
1	1	3	2	2	4	2	2	4
1	1	2	1	6	5	2	6	6
3	1	3	2	2	1	2	2	1
3	1	1	1	2	1	2	2	1
3	1	4	2	6	5	2	6	6
3	1	2	2	1	1	1	1	1
3	1	4	1	6	5	2	6	6
3	1	4	2	2	4	2	2	3
3	1	3	2	2	1	2	2	1
3	1	2	2	2	4	2	2	3
3	1	3	2	2	4	2	2	1
1	2	2	2	3	2	1	2	2
1	2	1	6	6	5	6	2	6
1	1	3	4	4	1	4	3	1
3	1	4	6	4	5	6	3	6
1	1	3	5	5	2	5	3	2
3	1	1	6	5	5	6	4	6
1	1	3	4	4	2	4	3	2
1	1	4	4	4	2	4	3	3
3	2	3	4	5	2	4	4	2
3	1	2	4	4	1	4	5	1
3	2	2	4	4	3	4	3	3

類別	編號	事件	假設項目				單位	
非洲菊產地供應	1	產地供應排序	花卉產地供應量排序(前三名)	(1)	彰投中	南投→嘉義→台中	N	
				(2)	彰投南	南投→嘉義→彰化		
			定義:颱風對應花卉產地相關供應縣市之前三名排序	(3)	投彰中	南投→台中→嘉義		
	2	產地供應波動	產量供應波動狀態	(1)	上升狀態	持續上升	N	
			定義:颱風期間對應花卉產量供應波動狀態			先上升後下降		
				(2)	下降狀態	持續下降		
						先下降後上升		
3	最大起伏變動大小量	0~250、251~500、500以上				件		
							定義:颱風期間對應花卉供應最大起伏波動大小(量)	
類別	編號	事件	假設項目				單位	
數量	1	拍賣量	颱風前一天與颱風當天	(1)	北>彰>中>南		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
				(2)	北>彰>南>中			
				(3)	北>南>彰>中			
				(4)	北>彰>雄>南>中			
				(5)	彰>北>雄>南>中			
				(6)	北>雄>彰>南>中			
				(7)	過半休市			
		拍賣量變化狀態	(1)	北(降)		降：下降程度為四大市場之最	把	
			(2)	彰(降)				
			(3)	中(降)				
	(4)	南(降)						
	(5)	無法判斷						
	2	成交量	颱風前一天	(1)	彰>北>南>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
				(2)	北>彰>南>中			
				(3)	北>南>彰>中			
				(4)	北>彰>雄>南>中			
				(5)	彰>雄>北>南>中			
				(6)	過半休市			
		成交量	颱風當天	(1)	彰>北>南>中		降：下降程度為四大市場之最	把
				(2)	北>彰>南>中			
				(3)	北>彰>雄>南>中			
				(4)	北>雄>彰>南>中			
				(5)	彰>北>南>雄>中			
(6)				過半休市				
成交量變化狀態		(1)	北(降)	(4)	南(降)			
	(2)	彰(降)	(5)	雄(降)				
	(3)	中(降)	(6)	無法判斷				

附錄 D-8 ：1998-2005 年大宗花卉-洋桔梗數據轉換表

洋桔梗平均供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
2	1	3	1	2	2	2	2	2
2	1	2	1	3	3	2	3	3
1	1	2	5	2	5	7	2	6
2	1	3	1	2	1	2	2	1
1	1	3	1	2	3	2	2	3
1	1	3	1	6	5	2	7	6
4	1	2	1	2	5	2	2	6
5	1	1	1	6	5	2	7	6
6	1	2	1	2	3	2	2	3
6	1	3	1	6	2	2	7	2
3	1	3	1	2	5	2	2	6
5	1	3	2	3	2	3	3	2
5	2	3	2	2	2	3	3	2
2	1	2	2	6	1	7	7	4
3	1	3	2	3	1	3	3	1
3	1	3	1	3	1	2	3	1
1	1	4	1	3	2	2	7	6
2	1	2	1	1	5	1	1	6
5	1	4	1	6	2	2	7	1
3	1	4	1	2	3	2	2	3
1	1	3	1	2	2	2	2	2
1	1	2	2	3	3	3	3	3
1	2	2	1	2	5	2	2	6
1	1	2	2	2	1	3	2	4
3	1	1	5	5	5	7	5	6
1	1	2	3	4	1	4	4	1
5	1	4	5	5	5	7	4	6
1	1	3	4	5	4	5	5	4
3	2	1	4	6	5	4	7	2
1	1	3	3	4	2	4	4	2
1	1	4	3	4	1	4	4	3
6	1	2	3	5	4	6	7	5
6	1	3	3	4	2	7	6	4
1	1	2	4	5	5	6	7	1

類別	編號	事件	假設項目						單位
洋桔梗產地供應	1	產地供應排序	花卉產地供應量排序 (前三名)	(1)	彰嘉雲	彰化→嘉義→雲林		N	
				(2)	彰嘉投	彰化→嘉義→南投			
			定義：颱風對應花卉產地 地相關供應縣市之前 三名排序	(3)	彰嘉南	彰化→嘉義→台南			
				(4)	彰嘉中	彰化→嘉義→台中			
				(5)	嘉彰南	嘉義→彰化→台南			
				(6)	嘉彰北	嘉義→彰化→台北			
	2	產地供應波動	產量供應波動狀態	(1)	上升狀態	持續上升		N	
			定義：颱風期間對應花卉 產量供應波動狀態			先上升後下降			
	3	最大起伏變動大小量 定義：颱風期間對應花卉 供應最大起伏波動 大小(量)	(2)	下降狀態	持續下降				
					先下降後上升				
				0~250、251~500、500 以上				件	
類別	編號	事件	假設項目						單位
數量	1	拍賣量	颱風前一天 與颱風當天	(1)	彰>北>中>南		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件	
				(2)	北>彰>中>南				
				(3)	北>彰>南>中				
				(4)	北>彰>雄>中>南				
				(5)	北>雄>中>彰>南				
				(6)	過半休市				
		拍賣量 變化狀態	(1)	北(降)		降：下降程度為四 大市場之最	把		
			(2)	彰(降)					
			(3)	中(降)					
			(4)	南(降)					
	(5)	無法判斷							
	2	成交量	颱風前一天	(1)	彰>北>中>南		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件	
				(2)	北>彰>中>南				
				(3)	北>彰>南>中				
				(4)	北>彰>雄>南>中				
				(5)	北>雄>中>彰>南				
				(6)	雄>彰>南>北>中				
				(7)	過半休市				
			颱風當天	(1)	彰>北>中>南				
				(2)	北>彰>中>南				
				(3)	北>彰>南>中				
		成交量 變化狀態	(4)	北>彰>雄>南>中		降：下降程度為四 大市場之最	把		
			(5)	北>雄>彰>中>南					
			(6)	過半休市					
			(1)	北(降)	(4)			南(降)	
			(2)	彰(降)	(5)			雄(降)	
(3)			中(降)	(6)	無法判斷				

附錄 D-9 ：1998-2005 年大宗花卉-康乃馨數據轉換表

康乃馨平均供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
3	1	1	6	6	5	3	1	4
3	1	4	6	6	5	7	1	6
3	1	2	6	6	5	7	6	6
3	1	2	6	6	5	7	1	2
3	1	4	6	6	5	7	2	2
2	1	1	3	6	5	2	6	6
2	1	4	2	6	5	1	6	6
2	1	1	6	6	5	1	6	6
2	1	2	6	6	5	7	6	6
2	1	2	6	6	5	7	6	6
2	1	4	6	6	5	7	6	6
2	1	3	3	1	2	2	2	2
2	3	3	3	1	4	2	2	4
6	1	2	1	6	5	1	6	6
6	1	3	3	3	2	2	3	2
1	1	1	3	6	5	2	6	6
5	1	2	1	6	5	1	6	6
5	1	3	6	6	5	7	6	6
5	1	3	6	6	5	7	6	6
5	1	4	6	6	5	7	6	6
6	1	3	3	2	2	2	3	2
6	1	2	6	6	5	7	6	6
1	1	4	4	3	4	4	3	4
1	1	2	4	6	5	3	6	6
1	1	1	6	6	5	7	6	6
1	1	2	6	6	1	6	6	1
1	1	4	6	6	5	7	6	6
1	1	1	6	6	5	7	6	6
1	1	1	6	6	5	7	5	6
1	1	3	5	4	2	5	4	2
1	1	4	6	6	5	7	6	6
1	1	3	6	6	5	7	6	6
1	1	3	6	4	5	7	4	6
1	1	2	6	5	1	7	5	5

類別	編號	事件	假設項目					單位				
康乃馨產地供應	1	產地供應排序	花卉產地供應量排序(前三名)	(1)	彰投南	彰化→南投→台南		N				
				(2)	彰嘉中	彰化→嘉義→台中						
			定義:颱風對應花卉產地相關供應縣市之前三名排序	(3)	彰中投	彰化→台中→南投						
				(4)	彰嘉雲	彰化→嘉義→雲林						
				(5)	彰桃屏	彰化→桃園→屏東						
				(6)	彰南桃	彰化→台南→桃園						
	2	產地供應波動	產量供應波動狀態		(1)	上升狀態	持續上升		N			
			定義:颱風期間對應花卉產量供應波動狀態				先上升後下降					
			(2)	下降狀態	持續下降							
					先下降後上升							
3	最大起伏變動大小量	0~250、251~500、500 以上						件				
									定義:颱風期間對應花卉供應最大起伏波動大小(量)			
類別	編號	事件	假設項目					單位				
數量	1	拍賣量	颱風前一天與颱風當天	(1)	彰>北>南>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件				
				(2)	彰>南>北>中							
				(3)	北>彰>南>中							
				(4)	彰>北>南>雄>中							
				(5)	北>彰>南>雄>中							
				(6)	雄>彰>中>南>北							
				(7)	過半休市							
		拍賣量變化狀態	(1)	北(降)		降:下降程度為四大市場之最	把					
			(2)	彰(降)								
			(3)	中(降)								
			(4)	南(降)								
			(5)	無法判斷								
			2	成交量	颱風前一天			(1)	彰>南>北>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件
								(2)	北>彰>南>中			
	(3)	南>彰>中>北										
	(4)	南>北>中>彰										
	(5)	彰>北>南>雄>中										
	(6)	北>雄>中>彰>南										
	(7)	過半休市										
	颱風當天	(1)			彰>南>中>北							
		(2)			北>彰>南>中							
		(3)			北>南>彰>中							
		(4)	北>彰>南>雄>中									
		(5)	彰>雄>南>中>北									
		(6)	過半休市									
	成交量變化狀態	(1)	北(降)	(4)	南(降)	降:下降程度為四大市場之最	把					
		(2)	彰(降)	(5)	雄(降)							
		(3)	中(降)	(6)	無法判斷							

附錄 D-10：1998-2005 年大宗花卉-劍蘭數據轉換表

劍蘭平均 供應量 (前三名 排序)	最大起伏 變動大小	產量波動	颱風前一 天拍賣量 比較	颱風當天 拍賣量 比較	拍賣量變 化量狀態	颱風前一 天成交量 比較	颱風當天 成交量 比較	成交量變 化量狀態
2	2	4	3	3	2	3	3	2
3	3	3	1	2	1	1	2	1
4	2	2	7	3	5	6	3	6
3	3	2	3	1	1	3	1	1
2	1	2	3	3	3	1	1	4
2	1	1	3	6	5	3	7	6
3	2	3	2	2	5	2	2	6
2	1	1	1	6	5	1	7	6
3	2	3	3	3	3	3	3	4
3	3	3	3	3	2	3	3	2
3	3	1	2	2	5	2	2	6
2	3	3	3	3	4	3	3	4
3	1	1	1	1	2	1	1	2
5	3	2	3	6	2	3	7	2
3	3	3	2	1	2	2	1	2
3	1	3	3	3	1	3	3	1
3	1	2	1	6	5	1	7	6
3	3	2	2	2	5	2	2	6
6	2	4	3	6	5	3	7	2
3	2	4	3	3	3	3	3	3
3	3	3	3	3	2	3	3	2
3	3	3	2	2	3	2	2	4
3	2	2	3	3	1	3	3	1
3	1	2	1	1	2	1	1	2
4	2	1	7	4	5	6	4	6
3	3	3	4	4	1	4	5	1
3	3	4	7	4	5	6	4	6
3	1	3	5	5	2	6	6	4
3	1	1	7	4	5	6	5	6
4	1	3	7	6	4	6	7	4
3	3	2	7	4	3	4	4	3
3	2	3	5	5	4	6	7	5
3	3	3	4	5	1	4	6	1
3	3	3	6	5	2	5	6	2

類別	編號	事件	假設項目					單位	
劍蘭 產地供應	1	產地供應 排序	花卉產地供應量排序 (前三名)	(1)	中彰南	台中→彰化→台南		N	
				(2)	中彰投	台中→彰化→南投			
			定義：颱風對應花卉產地 地相關供應縣市之前 三名排序	(3)	中彰苗	台中→彰化→苗栗			
				(4)	中苗彰	台中→苗栗→彰化			
				(5)	中苗雲	台中→苗栗→雲林			
				(6)	中彰南	台中→彰化→台南			
	2	產地供應 波動	產量供應波動狀態		(1)	上升狀態	持續上升		N
			定義：颱風期間對應花卉 產量供應波動狀態				先上升後下降		
				(2)	下降狀態	持續下降			
						先下降後上升			
3	最大起伏變動大小量 定義：颱風期間對應花卉 供應最大起伏波動 大小(量)	0~250、251~500、500 以上					件		
類別	編號	事件	假設項目					單位	
數 量	1	拍賣量	颱風前一天 與颱風當天	(1)	彰>北>南>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件	
				(2)	彰>南>北>中				
				(3)	北>彰>南>中				
				(4)	彰>北>雄>中>南				
				(5)	雄>彰>北>南>中				
				(6)	彰>雄>北>南>中				
				(7)	過半休市				
		拍賣量 變化狀態		(1)	北(降)		降：下降程度為四大 市場之最	把	
				(2)	彰(降)				
				(3)	中(降)				
				(4)	南(降)				
				(5)	無法判斷				
	2	成交量	颱風前一天	(1)	彰>北>南>中		北→台北 彰→彰化 中→台中 南→台南 雄→高雄	件	
				(2)	彰>南>北>中				
				(3)	北>彰>南>中				
				(4)	彰>北>雄>南>中				
				(5)	彰>雄>南>北>中				
				(6)	過半休市				
			颱風當天	(1)	彰>北>南>中				
				(2)	彰>南>北>中				
				(3)	北>彰>南>中				
				(4)	彰>北>南>雄>中				
				(5)	彰>雄>南>北>中				
				(6)	雄>彰>南>北>中				
				(7)	過半休市				
		成交量 變化狀態	(1)	北(降)	(4)	南(降)		降：下降程度為四大 市場之最	把
			(2)	彰(降)	(5)	雄(降)			
			(3)	中(降)	(6)	無法判斷			

## 附錄 E：演算法軟體分析與應用

Profession\_ARP 之 Apriori 演算法軟體是由美國伊利諾大學香檳分校吳建文副教授所發展出建構在 Windows 作業系統中資料探勘關聯性法則演算軟體。它是由大型電腦系統所撰寫出來的演算法軟體，經過多次修改之後，從原本僅能探勘 100 多筆資料數，不斷改善與修改，目前能探勘資料數已經高達 1000 多筆以上，其效能與運算速度都大為提昇，本軟體在於運算過程中，需要耗掉電腦記憶體，作為演算速度之驅動器。

在於本軟體主要分為三大部分：輸入檔案、輸出檔案、主要執行檔，將於下面幾點完整地敘述其功用。

### 1. 輸入檔案

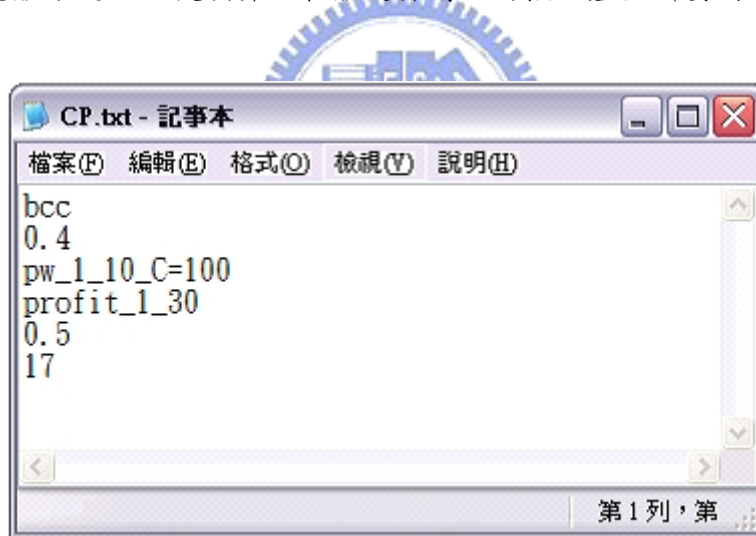
輸入檔案形式以存文字檔為基準，包含 CP.txt、bcc.txt、items\_name.txt 三種不同輸入格式，分別為設定資料檔名、資料庫資料內容及資料項目名稱，如附錄圖 1、附錄圖 2、附錄圖 3 所示。

### 2. 輸出檔案

3. 輸出檔案形式同樣以存文字為準則，包含 FI.txt、AR.txt 兩種不同輸出畫面各為高頻項目集檔案與關聯性法則檔案，如附錄圖 4、附錄圖 5 所示。

### 4. 主要執行檔

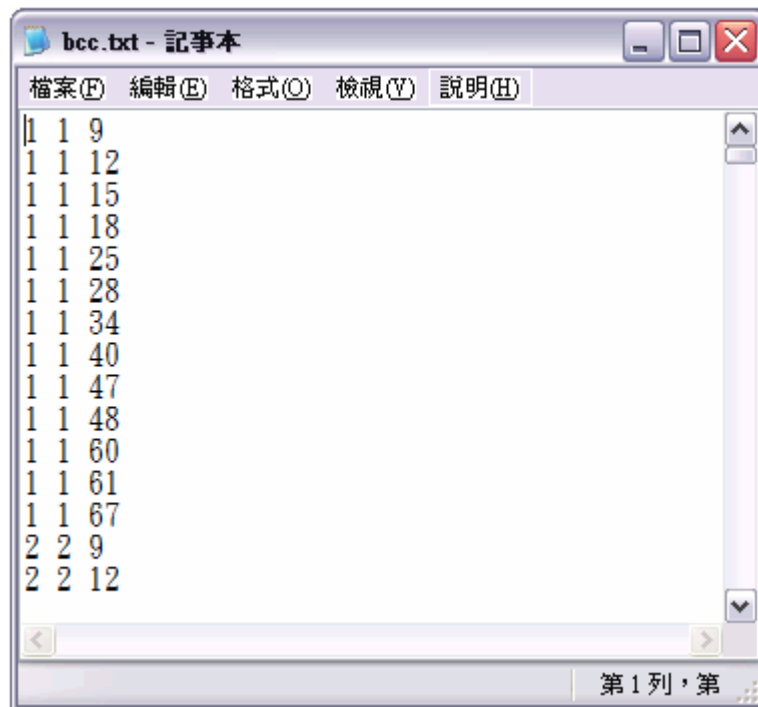
以執行檔的方式呈現，會伴隨著輸入檔案資料的不同，而執行不同的演算方式，如果未按照輸入檔案格式進行輸入資料者，將無法運行演算功能。另外一方面，假如記憶體不足以驅使演算法軟體運算時，將無法產生所要的結果。



附錄圖 1 輸入檔案 CP.txt

由附錄圖 1 可清楚得知輸入檔案包含輸入資料檔名與兩種參數，其兩種參數分別為支持度(Support)與信賴度(Confidence)。資料庫中有  $s$  筆交易、事件同時包含一群資料項目，則說項目此一群資料項目有 support  $s$ ，另一方面。如果一群資料項目  $\{A, D\}$  為 Frequent Pattern。則  $A \rightarrow D$  為一關聯性法則且其 Confidence 為  $\frac{Sup(A \cap D)}{Sup(D)}$ 。上述輸入資料檔名為 bcc，Support = 0.4，Confidence = 0.5。

接下來由附錄圖 2 發現每一行皆包含三種數字，前兩種必須一樣且一致(國際標準格式)，後者數字為資料轉換所得數字。例如第一位數字為 1 則表示第一筆交易或事件，後者數字為 1，代表此為項目 1 之第一種，如果後者數字為 5，代表為項目之累計數字轉換所得之。



附錄圖 2 輸入檔案 bcc.txt



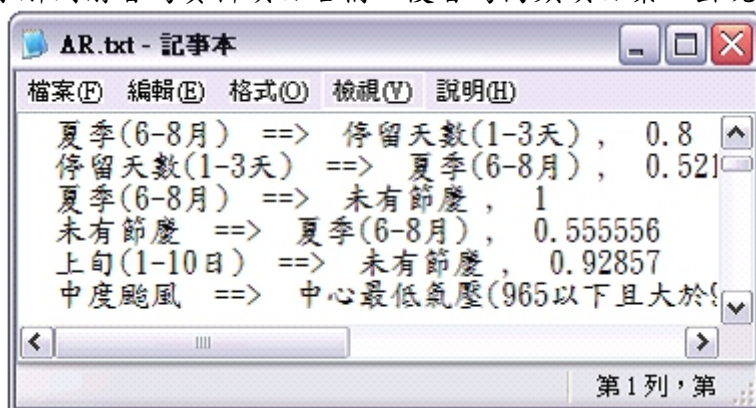
附錄圖 3 輸入檔案 items\_name.txt

再者，如附錄圖 3 所示，看出每行所顯示之資料皆為所要分析之資料項目，唯有第一行為國際標準格式設定為 0 外，其他行皆表示資料項目名稱。



附錄圖 4 輸出檔案 FI.txt

經由上附錄圖 4 了解到前者為資料項目名稱，後者為高頻項目集，出現次數頻率值。



附錄圖 5 輸出檔案 AR.txt

最後如附錄圖 5 可認知到關聯性法則的基本模式，當左邊條件成立時，右邊條件跟著成立的機率(IF..Then)，也就是所謂的信賴程度，上述為上旬且未有節慶的機率約為信賴度 0.93。



## 附錄 F：颱風特性間之 Apriori 演算法軟體分析結果

1. 針對近節慶個數(前後5天)與近節慶區段找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 50%，最小信賴度門檻值 = 95%，關聯性法則如附錄表1所示：

附錄表 1 關聯性法則(一)

項目組數	支持與信賴度	關聯性法則	值
1	(a)支持度	不符(超過 5 天) 超出區段	0.564
	(b)信賴度	不符(超過 5 天) ==> 超出區段	1
2	(c)支持度	不符(超過 5 天) 超出區段 最短近節慶 7 天以上	0.538
	(d)信賴度	不符(超過 5 天) AND 超出區段 ==> 最短近節慶 7 天以上	0.955
有意義的關聯性法則：			
1. 不符(超過 5 天) ==> 超出區段			
2. 不符(超過 5 天) AND 超出區段 ==> 最短近節慶 7 天以上			

上述關聯性法則解釋：

在侵台颱風與節慶資料全部相關記錄中，最小支持度與最小信賴度分別設為50%和95%。

項目組數為1者：

(a) 每次侵台颱風所伴隨而來的節慶數目(前後五天)與來臨時之前、後、前後同時發生不會相會節慶日的次數，顯示次數百分比為0.564(支持度)。也就是說明每逢侵台颱風來臨時，有超過一半的機率前後五天內不會有節慶產生與相會節慶的可能性。

(b) 當五天內不會有節慶產生時，同時不會發生前、後面臨節慶日的概率為1。即為信賴度。

項目組數為2者：

(c) 不符(超過5天)、超出區段與最短近節慶7天以上同時發生次數百分比為0.538。

(d) 在不符(超過5天)與超出區段下，同時發生最短近節慶七天以上的概率為0.955。

經由上述之初步關聯分析發現到幾個重要訊息，侵台颱風來臨時，節慶這個因素影響力顯的不是很重要。不管在於節慶發生天數(前五天)、最短之近節慶天數或進節慶前後區段，皆清楚說明「節慶因素」是可以不去考量的。因此可以予以適當刪減，如此一來不但可以減少分析之複雜度，也可以使後續分析能夠更容易進行。

2. 針對最低中心氣壓與近中心最大風速找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 35%，最小信賴度門檻值 = 95%，關聯性法則如附錄表2所示：

附錄表 2 關聯性法則(二)

項目組數	支持與信賴度	關聯性法則	值
1	(a)支持度	近中心最大風速(32.7-50.9m/s) 中度颱風	0.359
	(b)信賴度	近中心最大風速(32.7-50.9m/s) ==> 中度颱風	1
2	(a)支持度	第二種中心最低氣壓(965 以下且大於 945) 近中心最大風速(32.7-50.9m/s) 中度颱風	0.359
	(b)信賴度	第二種中心最低氣壓(965 以下且大於 945) AND 近中心最大風速(32.7-50.9m/s) ==> 中度颱風	1

有意義的關聯性法則：

1. 近中心最大風速(32.7-50.9m/s) ==> 中度颱風
2. 第二種中心最低氣壓(965 以下且大於 945) AND 近中心最大風速(32.7-50.9m/s) ==> 中度颱風

上述關聯性法則解釋：

在最低中心氣壓與近中心最大風速資料全部相關記錄中，最小支持度與最小信賴度分別設為35%和95%。

項目組數為1者：

(a) 以侵台颱風特性近中心最大風速(32.7-50.9m/s)與中度颱風次數百分比為0.359，可以清楚了解到，颱風約有1/3的機率為近中心最大風速(32.7-50.9m/s)與中度颱風的可能性。

(b) 當侵台颱風為近中心最大風速(32.7-50.9m/s)，同時也為發生中度颱風的概率為1。

項目組數為2者：

(c) 第二種中心最低氣壓、近中心最大風速(32.7-50.9m/s)與颱風強度為中度颱風同時發生次數百分比為0.359。

(d) 在第二種中心最低氣壓(965以下且大於945)與近中心最大風速(32.7-50.9m/s)下，同時出現颱風強度為中度颱風的概率為1。

為了驗證資料是否有一定的可信度，因此進行上述驗證性的關聯性法則(二)分析，其結果呈現與相關文獻有一定的吻合度，也與之前判斷的情況差不了多少，幾乎一樣。雖然資料筆數不是很多，但是仍可以經由此方式證實資料的可靠度與準確度。接著將以相同方式依序找出以下有意義的關聯性法則。

3. 針對侵台颱風移動方向與近中心最大風速找出之關聯性法則舉例說明：

設定最小支持度門檻值 = 30%，最小信賴度門檻值 = 95%，關聯性法則如附錄表3所示：

附錄表 3 關聯性法則(三)

項目組數	支持與信賴度	關聯性法則	值
2	(a)支持度	移動方向北 近中心最大風速(32.7-50.9m/s) 中度颱風	0.308
	(b)信賴度	移動方向北 AND 近中心最大風速(32.7-50.9m/s) ==> 中度颱風	1
有意義的關聯性法則：			
1. 移動方向北 AND 近中心最大風速(32.7-50.9m/s) ==> 中度颱風			

4. 針對侵台颱風移動路徑與移動速度找出之關聯性法則舉例說明：

設定最小支持度門檻值=20%，最小信賴度門檻值=95%，關聯性法則如附錄表4所示：

附錄表 4 關聯性法則(四)

項目組數	支持與信賴度	關聯性法則	值
2	(a)支持度	第二類型(3,4) 移動速度(0-15km/hr) 移動方向北	0.205
	(b)信賴度	第二類型(3,4) AND 移動速度(0-15km/hr) ==> 移動方向北	1
有意義的關聯性法則：			
1. 第二類型(3,4) AND 移動速度(0-15km/hr) ==> 移動方向北			

5. 針對最大總雨量與近中心最大風速找出之關聯性法則舉例說明：

設定最小支持度門檻值=20%，最小信賴度門檻值=95%，關聯性法則如附錄表5所示：

附錄表 5 關聯性法則(五)

項目組數	支持與信賴度	關聯性法則	值
2	(a)支持度	最大總雨量(0-500mm) 近中心最大風速(32.7-50.9m/s) 中度颱風	0.205
	(b)信賴度	最大總雨量(0-500mm) AND 近中心最大風速 (32.7-50.9m/s) ==> 中度颱風	1
有意義的關聯性法則：			
1. 最大總雨量(0-500mm) AND 近中心最大風速(32.7-50.9m/s) ==> 中度颱風			

