

### 第三章 資料倉儲維度擴增實作

台北花卉交易資訊系統，從建構成功運作至今年已經兩年多，系統逐漸趨於完備，但是對於只能得到交易的日期卻無法得到交易的確切時間感到小有遺憾，所以本章利用現有的台北花卉資料倉儲進行時間維度的擴增，將現有的年月日時間維度，擴增至可下挖到時分秒。實作的步驟牽涉到整個系統，為了清楚解釋維度擴增實作步驟，本章共分為五節，在第 3.1 節說明系統主架構以及原始資料所需的部份，第 3.2 節說明在三階正規化資料庫所修改的部份，第 3.3 節說明在資料倉儲所修改與新增的部份，第 3.4 節說明資料轉換服務的設計，第 3.5 節則說明架構線上分析處理系統的步驟。

#### 3.1 系統架構

對於台北花卉資料倉儲及線上分析系統來說，每天的資料流程是當花卉公司將當日的花卉交易資料透過檔案傳送服務(FTS)傳送至交通大學的伺服器後開始。當伺服器接收到交易資料後，透過第一次資料轉換服務(DTS-1)將資料萃取、淨化並將資料轉入預先建構完成的三階正規化資料庫中。在三階正規化資料庫中的資料在經由第二次資料轉換服務(DTS-2)的處理，把需要的資料轉入資料倉儲中。當資料倉儲建構完成後，在緊接著利用 Cognos 套裝軟體把資料倉儲中的資料轉化成超方體(Hyper Cube)，再透過線上分析處理系統(OLAP)介面將資料發佈至網路上以供使用者查詢使用。

花卉資料倉儲及線上分析系統的架構與資料流程如圖 3.1 所示如下。

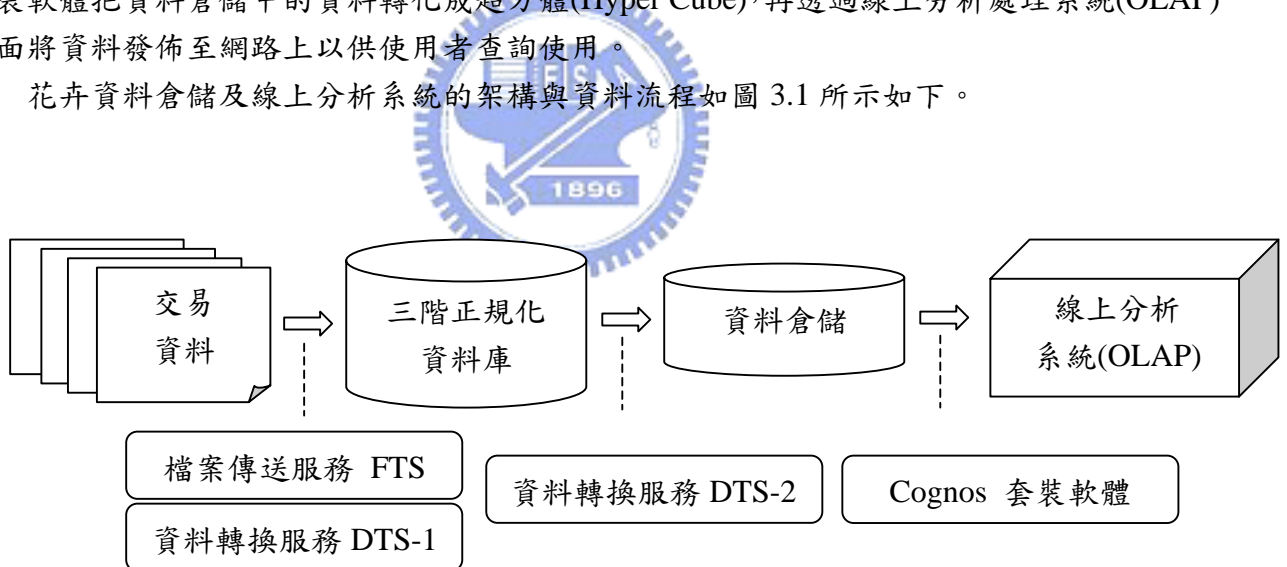


圖 3.1 系統架構與資料流程

### 3.2 三階正規化資料庫

本節將說明台北花卉資料倉儲系統的原始交易資料介紹以及三階正規化資料庫的架構。本節共分為兩小節，第 3.2.1 節介紹原始交易資料；第 3.2.2 節則說明三階正規化資料庫的架構。

#### 3.2.1 交易資料介紹

花卉公司將每日的交易資料經過初步整理後透過檔案傳送服務(FTS)，將五個交易資料檔案傳送至台北花卉資料倉儲伺服器，五個檔案分別為日交易資料表(DT\*.DBF)、供應人資料表(SU\*.DBF)、承銷人資料表(BU\*.DBF)、花卉資料表(PR\*.DBF)以及日進貨資料表(DP\*.DBF)，其檔案的格式如附錄二。

對於交易資料的五個檔案而言，日交易資料表中主要包含交易序號、成交日期、供應人代碼、承銷人代碼、花卉代號、花卉等級、花卉數量、拍賣價格等資訊。供應人資料表中主要包含供應人代碼、供應團體代號、供應人姓名、供應類別等資訊，其中供應類別可分為農會合作社、供應團體所屬花農、個人、進口商、行口商、花班種植面積六種類別。承銷人資料表中主要包含承銷人代碼、承銷人姓名、承銷區域等資訊，其中承銷區域可分為北市、北部、中部、雲嘉、台南、高雄市、鳳屏等七個區域。花卉資料表中主要包含花卉代號、花卉品名、花卉類別等資訊。日進貨資料表中主要包含進貨日期、交易序號、供應人代碼、花卉代號、花卉等級、花卉數量、底價等資訊。五個交易資料檔案內容如圖 3.2 所示。

	A	BC	D	E	F	G	H	I	J	K	
1	DT_DA	CI	IPM_NO	BU_N	SU_N	PR_NO	CL	PCS	BO	PRIC	AM
2	930201	C0	00101		X0023	FY000	B	40	5	0	
3	930201	C2	0010101	C0023	X0023	FY000	B	40	5	20	
4	930201	C0	00102		X0202	FY000	B	30	1	0	
5	930201	C2	0010201	C0202	X0202	FY000	B	30	1	20	
6	930201	C0	00103		X0078	FY000	B	30	1	0	
7	930201	C2	0010301	C0078	X0078	FY000	B	30	1	20	
8	930201	C0	00104		X0070	FY000	B	40	1	0	
9	930201	C2	0010401	C0070	X0070	FY000	B	40	1	20	
10	930201	C0	00105		X0084	FY000	B	40	1	0	
11	930201	C2	0010501	C0084	X0084	FY000	B	40	1	20	
12	930201	C0	00106		X0128	FY000	B	40	1	0	
13	930201	C2	0010601	C0128	X0128	FY000	B	40	1	20	
14	930201	C0	00107		X0063	FY000	B	40	6	0	
15	930201	C2	0010701	C0063	X0063	FY000	B	40	6	20	

	A	B	C	D	E	F	G
1	SU_LC	SU_N	SU_GR	SU GROUP	SU NAME	S	SU
2	00014	AB002	AB002			3	金
3	00016	ACA00	ACA00		會	3	金
4	00018	AD001	AD001	賴壯		3	金
5	00019	ADA00	ADA00	趙	會	3	
6	00020	ADA00	ADA00	林		3	
7	00022	ADA00	ADA00	林		3	
8	00023	ADA00	ADA04	會		3	
9	00025	ADA00	ADA00	林		3	
10	00026	AFA00	AFA00	羅	會	3	
11	00027	AFA01	AFA00	羅		3	
12	00038	AG001	AG001	陳		3	
13	00039	AG002	AG002	楊		3	
14	00041	AG005	AG005	陳		3	

	A	B	C	D	E	F	G
1	BU_L	BU_N	BU_NAME	BU_RE	BU_BD	BU_IDNO	BU_T
2		C0002			481026	女 A2	29976
3		C0003				F1	22811
4		C0004				Y1	28611
5		C0005				Q1	28118
6		C0006				A1	25011
7		C0007				N1	29823
8		C0008				Q1	25976
9		C0009			650215	女 F2	28668
10		C0010				F1	29721
11		C0011				A1	25978
12		C0012				N1	25011
13		C0013				T2	02-892
14		C0014				N1	25958

	A	B	C	D	E	F	G
1	PR_NO	PR_ABR	PR_NAME	PR_PR	PR_E	PR_N	
2	A0102	鐵莧	鐵莧	0	大宗	0	國產
3	A0104	變葉鐵莧	變葉鐵莧	0	大宗	0	國產
4	A0106	撒金鐵莧	撒金鐵莧	0	大宗	0	國產
5	A0108	旋葉鐵莧	旋葉鐵莧	0	大宗	0	國產
6	A0110	鑲邊旋葉鐵莧	鑲邊旋葉鐵莧	0	大宗	0	國產
7	A0112	紅穗鐵莧	紅穗鐵莧	0	大宗	0	國產
8	A0114	白穗鐵莧	白穗鐵莧	0	大宗	0	國產
9	A0116	紅葉鐵莧	紅葉鐵莧	0	大宗	0	國產
10	A0118	彩葉鐵莧	彩葉鐵莧	0	大宗	0	國產
11	A0120	紅邊鐵莧	紅邊鐵莧	0	大宗	0	國產
12	A0302	槲樹	槲樹	0	大宗	0	國產
13	A0502	沙漠玫瑰	沙漠玫瑰	0	大宗	0	國產
14	A0700	鐵線蕨	鐵線蕨	0	大宗	0	國產
15	A0702	大葉鐵線蕨	大葉鐵線蕨	0	大宗	0	國產

	A	B	C	D	E	F	G	H	I	J	K
1	DT_DATE	CP_ID	PM_NO	DTSTH	BU_N	SU_N	PR_NO	CL	PCS	SF	
2	890103	P 4	1839902		P0201	OBI01	D1702		30	09	
3	890103	P 9	1842702			JJK04	C9110		1	09	
4	890103	P 9	1888601			SVA45	A2118		1	21	
5	890103	P 4	1889202		P0304	GKI07	C5304		24	09	
6	890103	P 4	1933501		P0203	UGV0	D1723		4	03	
7	890103	P 4	1944301		P0306	DB004	B0524		24	09	
8	890103	P 4	1946603		P0360	UII01	A4102		30	18	
9	890103	P 4	1947101		P0405	LFI44	E0304		10	30	
10	890103	P 4	1959601		P0201	PMI01	O0303		30	09	
11	890103	P 4	1960001		P0310	LFI44	D1721		12	15	
12	890103	P 4	1962002		P0184	AAA0	H0200		24	09	
13	890103	P 4	1962102		P0201	AAA0	N3102		24	09	
14	890103	P 9	1965303			SVA04	G1310		1	13	
15	890103	P 4	1967203		P0304	EK015	P0102		1	12	

圖 3.2 原始交易資料檔案

### 3.2.2 三階正規化資料庫架構

在接收台北花卉公司所傳送的五個交易資料檔案後，必須透過第一次資料轉換服務將交易資料匯入 SQL Server2000 裡的資料庫，第一次轉換服務會在 3.4.1 節介紹。而資料在轉換的過程中須經過三階正規化的步驟，三階正規化的目的在於消除資料的重複性、相依性、遞移性。因此在經過三階正規化處理過後的資料可以大幅的減少所佔用的儲存空間，

同時可以藉由三階正規化資料庫建立關聯性，找出原始資料中重複以及錯誤的資料。

在系統當中是以 SQL Server2000 來架構台北花卉三階正規化資料庫。在原始的五個交易資料檔案經過三階正規化處理後可產生十一張資料表，十一張資料表為拍賣資料表、進貨資料表、供應人資料表、供應鄉鎮資料表、供應縣市資料表、拍賣線資料表、承銷人資料表、承銷區域資料表、花卉資料表、花卉等級資料表及花卉等級略表，其關聯圖如圖 3.3 所示。在附錄三中有十一張資料表的各個屬性詳細說明，包含資料表的主鍵、外鍵、欄位名稱、資料型別、允許空值與否、資料長度及附加說明。

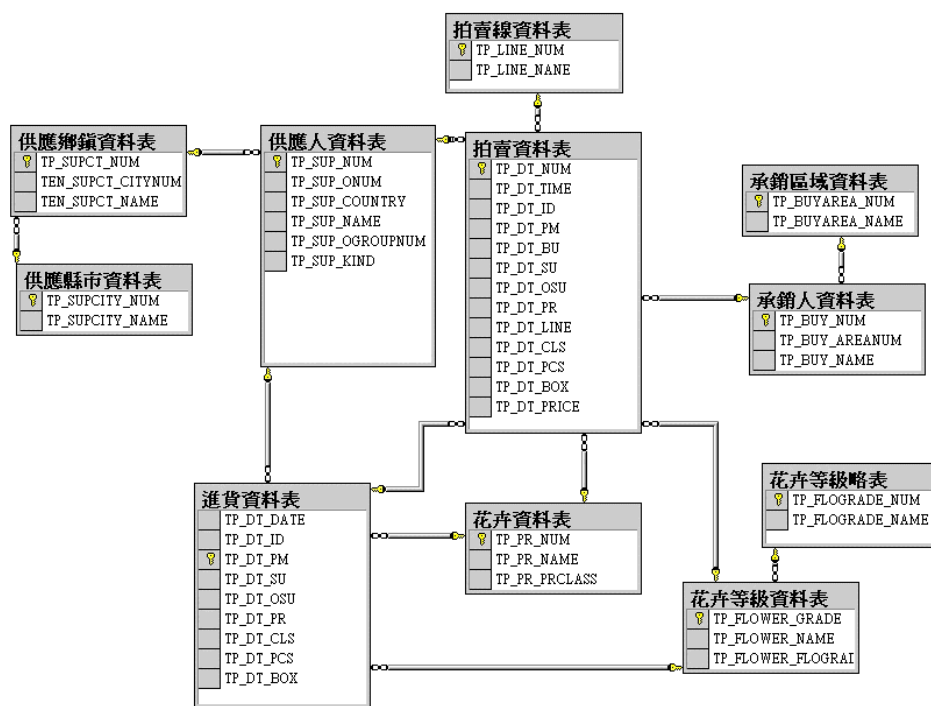


圖 3.3 台北花卉三階正規化資料庫



### 3.3 資料倉儲

在透過資料轉換服務將三階正規化資料庫內的資料匯入資料倉儲之前，必須先將資料倉儲建置完成。本節將說明時間維度擴增的實作部份以及擴增後的資料倉儲架構，在第 3.3.1 節將說明時間維度的架構；而第 3.3.2 節則說明擴增完成後的資料倉儲架構。

#### 3.3.1 時間維度的架構

由於台北花卉資料倉儲系統在時間維度的部份只能下挖到年月日。這部份導致供應人以及承銷人只能得知每筆交易資料的日期而無法得知確切交易時間，實屬台北花卉資料倉儲的小缺憾。若供應人以及承銷人能夠得知確切交易時間便能進行更精準的需求預測，因此產生了時間維度擴增的動機。

現今的台北資料倉儲系統由於 COGNOS 系統預設支援 Datetime 格式的資料自動作為時間維度，所以在台北資料倉儲系統中對於儲存交易時間資料的欄位都是採用 Datetime 格式。但也由於 COGNOS 系統的支援 Datetime 格式只支援年月日，如 2004/04/04；無法支援到時分秒。由於先前的台北資料倉儲系統並不需要呈現確切交易時間再加上 COGNOS 系統的原因，所以在台北資料倉儲系統中完全找不到儲存交易時間的欄位，在實作的第一步便是從資料儲存欄位動手，將三階正規化資料庫以及資料倉儲內有關交易時間的所有欄位更改成 Int 格式，以符合接下來的設計。

本小節將說明在資料倉儲中時間維度的架構，時間維度架構包含五個資料表，分別為月資料表、日資料表、時資料表、分資料表及秒資料表。表 3.1、表 3.2、表 3.3、表 3.4 及表 3.5 分別為月資料表、日資料表、時資料表、分資料表及秒資料表五個資料表屬性的詳密說明，包含主鍵、外鍵、欄位名稱、資料型別、空直允許與否、長度以及附加說明。五個資料表根據主鍵與外鍵，可建立關聯性如圖 3.4。再與資料倉儲事物表建立關聯性，即完成時間維度的架構。

表 3.1 月資料表

主鍵	外鍵	欄位名稱	資料型別	允許空值	長度	說明
◎		TP_DT_TIME	Char		10	月份+日期+時間，ex: 0928050505
	◎	TP_TIME_DATE	Char		8	日期+時間，ex: 28050505
		SEA_NAME	Char		4	月份中文，ex: 09月

表 3.2 日資料表

主鍵	外鍵	欄位名稱	資料型別	允許空值	長度	說明
◎		TP_TIME_DATE	Char		8	日期+時間，ex: 28050505
	◎	TP_TIME_HOUR	Char		6	時間，ex: 050505
		DATE_NAME	Char		4	日期中文，ex: 28日

表 3.3 時資料表

主鍵	外鍵	欄位名稱	資料型別	允許空值	長度	說明
◎		TP_DT_HOUR	Char		6	時間，ex: 050505
	◎	TP_TIME_MIN	Char		4	分+秒，ex: 0505
		HOUR_NAME	Char		4	時中文，ex: 5點

表 3.4 分資料表

主鍵	外鍵	欄位名稱	資料型別	允許空值	長度	說明
◎		TP_DT_MIN	Char		4	分+秒，ex: 0505
	◎	TP_TIME_SEC	Char		2	秒，ex: 05
		MIN_NAME	Char		4	分中文，ex: 05分

表 3.5 秒資料表

主鍵	外鍵	欄位名稱	資料型別	允許空值	長度	說明
◎		TP_DT_SEC	Char		2	秒，ex: 05
		SEC_NAME	Char		4	秒中文，ex: 05秒

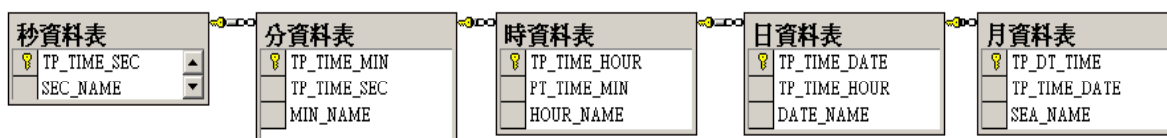


圖 3.4 時間維度資料表關聯圖

以下透過實際資料以及關聯來說明，可更清楚的表達時間維物的架構。如圖 3.5 所示。透過月資料表中的 TP\_TIME\_DATE 欄位與日資料表中的 TP\_TIME\_DATE 建立關聯；再透過日資料表中的 TP\_TIME\_HOUR 與時資料表中的 TP\_TIME\_HOUR 建立關聯，則此同時月資料表則與時資料表也建立起關聯性，同理往下進行與分資料表、秒資料表的關聯建立。以在事物表中 TP\_DT\_TIME 欄位「0101030001」為例，對照月資料表中即可得知此欄位屬於 1 月，且 TP\_DT\_DATE 欄位為「01030001」，接著對照日資料表中即可得知此欄位屬於 1 日，且 TP\_DT\_HOUR 欄位為 030001，同理向下進行，就可得知 TP\_DT\_TIME 欄位「0101030001」是屬於 1 月 1 日 3 時 0 分 01 秒交易的資料。

月資料表

TP_DT_TIME	TP_TIME_DATE	SEA_NAME
0101030000	01030000	01月
0101030001	01030001	01月
0101030002	01030002	01月
0101030003	01030003	01月
0101030004	01030004	01月
0101030005	01030005	01月
0101030006	01030006	01月
0101030007	01030007	01月
0101030008	01030008	01月
0101030009	01030009	01月
0101030010	01030010	01月
0101030011	01030011	01月
0101030012	01030012	01月
0101030013	01030013	01月
0101030014	01030014	01月
0101030015	01030015	01月
0101030016	01030016	01月
0101030017	01030017	01月

日資料表

TP_TIME_DATE	TP_TIME_HOUR	DATE_NAME
01030000	030000	01日
01030001	030001	01日
01030002	030002	01日
01030003	030003	01日
01030004	030004	01日
01030005	030005	01日
01030006	030006	01日
01030007	030007	01日
01030008	030008	01日
01030009	030009	01日
01030010	030010	01日
01030011	030011	01日
01030012	030012	01日
01030013	030013	01日
01030014	030014	01日
01030015	030015	01日
01030016	030016	01日
01030017	030017	01日

時資料表

TP_TIME_HOUR	TP_TIME_MIN	OUR_NAME
000000	0000	00點
000001	0001	00點
000002	0002	00點
000003	0003	00點
000004	0004	00點
000005	0005	00點
000006	0006	00點
000007	0007	00點
000008	0008	00點
000009	0009	00點
000010	0010	00點
000011	0011	00點
000012	0012	00點
000013	0013	00點
000014	0014	00點
000015	0015	00點
000016	0016	00點
000017	0017	00點

分資料表

TP_TIME_MIN	TP_TIME_SEC	MIN_NAME
0000	00	00分
0001	01	00分
0002	02	00分
0003	03	00分
0004	04	00分
0005	05	00分
0006	06	00分
0007	07	00分
0008	08	00分
0009	09	00分
0010	10	00分
0011	11	00分
0012	12	00分
0013	13	00分
0014	14	00分
0015	15	00分
0016	16	00分
0017	17	00分

秒資料表

TP_TIME_SEC	SEC_NAME
00	00秒
01	01秒
02	02秒
03	03秒
04	04秒
05	05秒
06	06秒
07	07秒
08	08秒
09	09秒
10	10秒
11	11秒
12	12秒
13	13秒
14	14秒
15	15秒
16	16秒
17	17秒

圖 3.5 時間維度五資料表內容關聯

### 3.3.2 資料倉儲的架構

在三階正規化資料庫架構完成後，必須先進行台北花卉資料倉儲系統的建置才能透過第二次轉換服務將資料匯入倉儲內。系統中是使用 SQL Server 2000 搭配雪花綱要 (Snowflake Schema) 來架構資料倉儲。圖 3.6 為台北花卉資料倉儲系統的雪花綱要圖。雪花綱要圖主要是利用事物表(Fact Table)來存放要分析的衡量值(Measure)，而事物表的周圍則是進行線上分析處理時所需要的維度表(Dimension Table)，如圖 3.6 的拍賣資料表即為此雪花綱要圖的事物表，周圍則是維度表資料表，包含容器別、時間、拍賣線、花卉等級、承銷人、花卉資料、供應人、供應地區等八個維度十九個資料表。相較於三階正規化資料庫裡的資料表，除了每張資料表中進一步的篩選外，也增加了幾張新的維度表以供查詢如供應團體資料表、供應類別資料表、花卉品名資料表、容器資料表、花卉種類資料表等，其屬性的詳細說明如附錄四，包含主鍵、外鍵、資料欄名稱、資料型別、長度、可否為空值及相關說明。

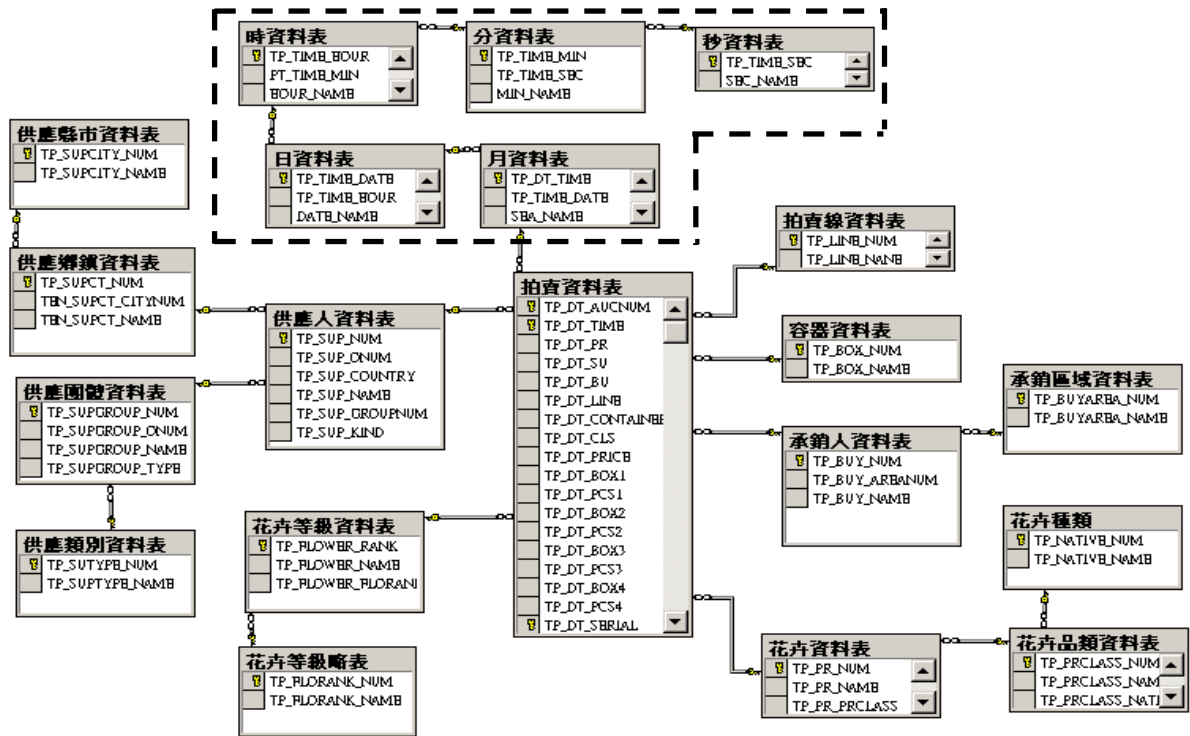


圖 3.6 台北資料倉儲雪花網要圖





### 3.4 資料整合轉換

對於一個資料倉儲系統的建置過程來說，資料整合轉換（ETL）通常是最花費時間與人力的，以歷史經驗來看資料整合轉換工作平均可以佔去專案的 60% - 70% 的時間。一個系統的資料來源，不僅僅侷限於來源資料庫，也可能來自於文字檔或是其他不同的檔案格式，所以資料整合轉換工作就是將各個不同的資料檔案或是來源資料庫中所擷取出的資料，根據專案的需求及資料倉儲的設計，將不需要或是重複的資料進行篩選，再存入資料倉儲預設的欄位內。資料整合轉換工作的內容如圖 3.7 所示，主要包含資料取出（Data Extraction）、資料轉換（Data Transformation）、資料載入（Data Loading）三個步驟。首先，資料取出就是從來源資料檔案或是來源資料庫中所讀取並擷取出所需的資料；第二步資料轉換則是將前步驟所取出的資料，根據需求設定規則、索引欄位、或是跟其他資料進行合併，轉換出最終所需要的資料型態，同時若含有錯誤、資料不足或不符合規則資料進行清除或更正；第三步資料載入則是將取出、轉換後的資料載入資料倉儲中所對應的欄位。

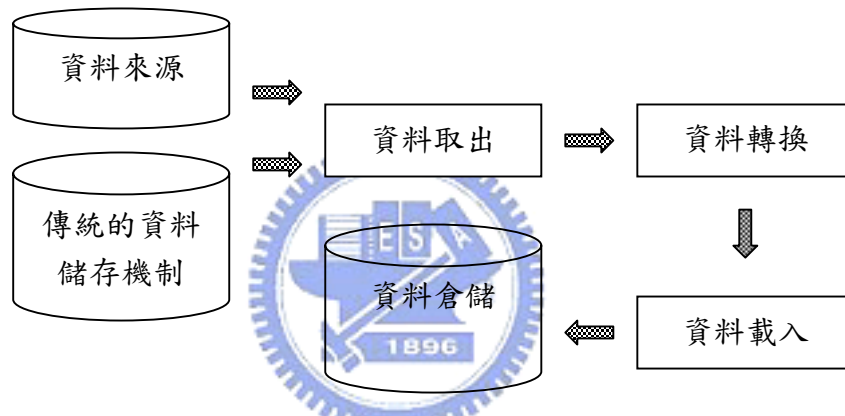


圖 3.7 資料整合轉換工作流程圖

對於本論文所使用的資料整合轉換工具而言，是採用 SQL Server2000 中的資料轉換服務（Data Transformation Service，DTS）去負責處理資料轉換與資料載入的工作；而資料取出的步驟是採用檔案轉移服務（File Transfer Service，FTS）軟體[12]來負責將花卉交易資料檔傳送至資料倉儲系統所在；在論文中主要就資料轉換與資料載入部分特別說明，FTS 軟體在本論文中不予說明。在系統中就資料轉換部分共可分為兩個封裝，如圖 3.8 所示，分別為三階正規化 1、三階正規化 2。而資料載入部份可分為三個封裝，如圖 3.8 所示，為倉儲事物表部份、倉儲供應團體維度以倉儲其餘維度部份三個封裝。三階正規化 1、三階正規化 2 相對於系統架構圖 3.1 中的資料轉換服務 DTS-1，而倉儲事物表部份、倉儲供應團體維度以及倉儲其餘維度部份三個封裝則對應於系統架構中的資料轉換服務 DTS-2。

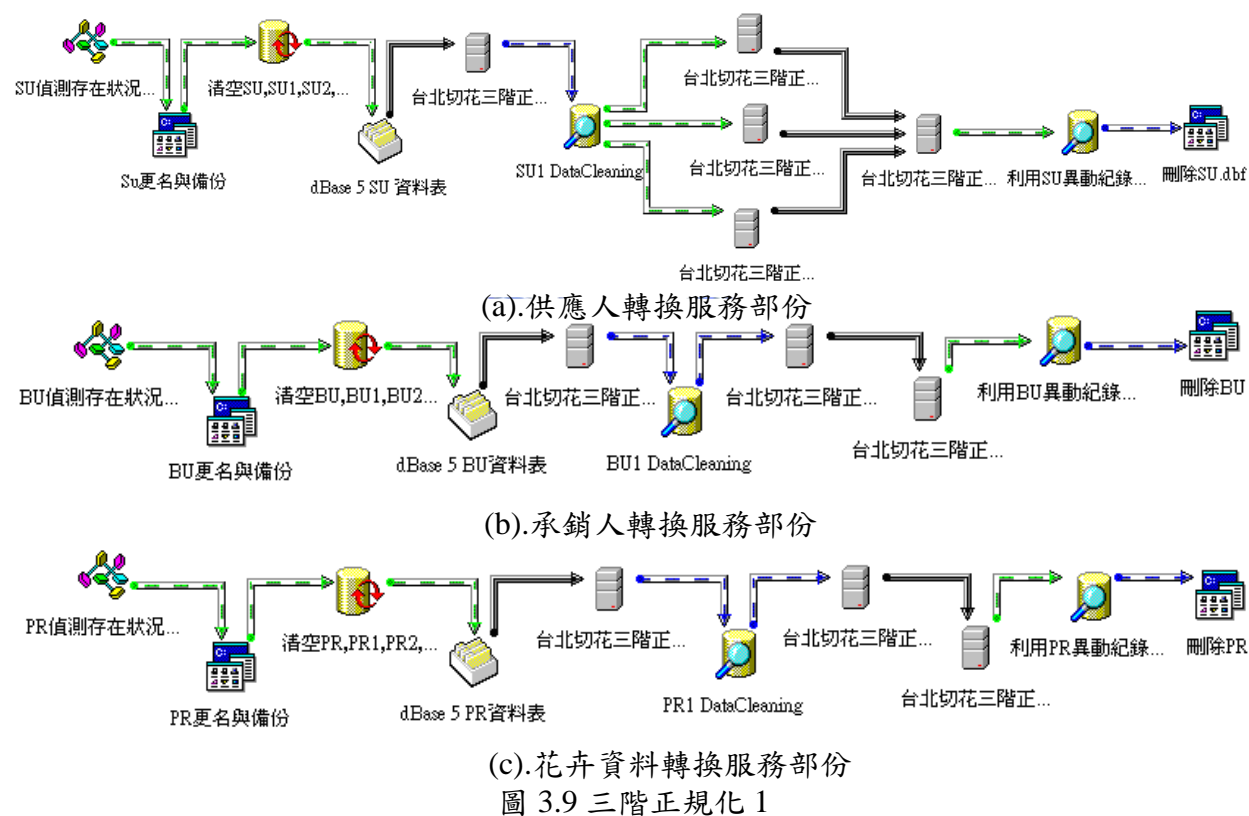
三階正規化1	FLOWERB...	2005/4/11...
三階正規化2	FLOWERB...	2005/4/11...
倉儲事物表部份	FLOWERB...	2005/4/11...
倉儲供應人團體維度	FLOWERB...	2005/4/11...
倉儲其餘維度部份	FLOWERB...	2005/4/11...

圖 3.8 資料轉換服務封裝

本節共分為兩小節；在 3.4.1 節說明資料轉換步驟的執行，如何透過資料轉換服務 DTS-1 將五個原始交易資料進行轉換後匯入三階正規化資料庫中。在 3.4.2 節說明資料載入步驟的執行，如何透過資料轉換服務 DTS-2 在三階正規化資料庫中載入所需的資料建構資料倉儲。

### 3.4.1 資料轉換：DTS-1

就資料轉換服務 DTS-1 而言，主要的目的為將五個原始交易資料檔匯入三階正規化資料庫中。由於日進貨資料表與日交易資料表中包含有外鍵的資料，需參考供應人資料表、承銷人資料表或花卉品名資料表中的資料，若有資料擁有外鍵的屬性卻在與其有關聯性的資料表中沒有相對應的資料存在，在資料轉換的過程中會導致錯誤的發生而無法完成資料的轉換。故將資料轉換服務 DTS-1 分為兩個部分，首先須完成被參考資料表(供應人資料表、承銷人資料表或花卉品名資料表)的轉換工作—三階正規化 1，如圖 3.9 所示；接著將日進貨資料表與日交易資料表資料匯入三階正規化資料庫中--三階正規化 2，如圖 3.11 所示。



在三階正規化 1 封裝中，首先就圖 3.9 (a).供應人轉換服務部分說明，步驟一開始系統先偵測來源資料夾內是否有 SU.DBF 檔案之存在。若無則跳出錯誤結束資料轉換服務；若有 SU.DBF 檔案存在繼續進行以下程序。呼叫一外部.bat 檔將原始 SU 資料檔更名並備份資料至備份資料夾中。進行清空 SU 資料表以及 SU1、SU2 暫存資料表中資料的動作。將資料初步轉移至 SU1 暫存資料表中後透過資料淨化比對供應人資料，判斷供應人資格是否停用並對於有誤的資料，如供應人代碼、名稱、團體代碼等允以紀錄暫存於例外資料表中，

待資料確認後再重新轉換。在資料轉換部分共可分為新增、無異動(修改)、停用三種程序；先就「停用」的程序說明，當供應人太久未供貨至市場拍賣，市場經營者為了方便管理變會取消其供應人資格以便管理，在資料轉換過程中將供應人資料使用狀況欄位更改為「N」並將 SACTTYPE 欄位設為 3，其餘欄位由原始資料檔直接匯入相對應之欄位，選取停用供應人資格之 SQL 語法如表 3.6 所示。

表 3.6 選取停用供應人資格之 SQL 語法

```
SELECT      G.*
FROM        SU1 AS S, SUTable AS G
WHERE       ((S.SNUM = G.TP_SUP_ONUM) AND (S.SNAME <> G.TP_SUP_NAME) AND
            (G.TP_SUP_IFUSE = 'Y'))
```

就「新增」的程序說明，當一新供應人加入時或是當供應人更改姓名時，在資料轉換過程中將供應人資料表 SACTTYPE 欄位設為 1，其餘欄位由原始資料檔直接匯入相對應之欄位，選取新增供應人資格之 SQL 語法如表 3.7 所示。

表 3.7 選取新增供應人資格之 SQL 語法

```
SELECT      *
FROM        SU1 AS S
WHERE       (S.SNUM NOT IN
            (SELECT TP_SUP_ONUM FROM SUTable)) OR
            (S.SNUM IN (SELECT TP_SUP_ONUM FROM SUTable WHERE TP_SUP_IFUSE = 'Y' AND TP_SUP_NAME NOT IN
            (SELECT SNAME FROM SU1 WHERE SNUM = TP_SUP_ONUM))))
```

就「無異動」的程序說明，當一新供應人資料並無改變或是當供應人只更改姓名以外的資料時，在資料轉換過程中將供應人資料表 SACTTYPE 欄位設為 2，其餘欄位由原始資料檔直接匯入相對應之欄位，選取新增供應人資格之 SQL 語法如表 3.8 所示。

表 3.8 選取無異動供應人資格之 SQL 語法

```
SELECT      *
FROM        SU1 AS S
WHERE       (S.SNUM NOT IN
            (SELECT TP_SUP_ONUM FROM SUTable)) OR
            ((S.SNUM IN (SELECT TP_SUP_ONUM FROM SUTable WHERE TP_SUP_IFUSE = 'Y' AND TP_SUP_NAME IN
            (SELECT SNAME FROM SU2 WHERE SNUM = TP_SUP_ONUM))))
```

當完成以上的資料的判斷與匯入的動作後，再利用 SU 異動紀錄將資料匯入 SU Table，當供應人資料表 SACTTYPE 欄位為 1 時，新增資料並將 IFUSE 設為 Y，欄位為 2 時將資料更新但 IFUSE 不作更改，欄位為 3 時將 IFUSE 設為 N；SU 異動紀錄 SQL 語法所示如表 3.9。最後步驟為刪除 SU.DBF 後結束供應人轉換服務部分之流程，圖 3.10 為供應人轉換服務部分中部分動作之說明。

表 3.9 SU 異動紀錄 SQL 語法

```

'利用SACTTYPE三種情況1,2,3
'分別決定如何處理資料的方式
'*****
Function Main()

    DTSDestination("TP_SUP_ONUM") = DTSSource("SNUM")
    DTSDestination("TP_SUP_GNUM") = DTSSource("SGNUM")
    DTSDestination("TP_SUP_NAME") = DTSSource("SNAME")
    DTSDestination("TP_SUP_ID") = DTSSource("SID")
    DTSDestination("TP_SUP_VOICE") = DTSSource("SVOICE")
    DTSDestination("TP_SUP_KIND") = DTSSource("SKIND")
    DTSDestination("TP_SUP_STOP") = DTSSource("SSTOP")

    Select Case Trim(DTSSource("SACTTYPE"))
' 新增 -->Ifuse欄位寫入為'Y' (Insert)
        CASE "1"
            DTSDestination("TP_SUP_IFUSE")="Y"
            Main = DTSTransformstat_InsertQuery
' 修改 -->Ifuse欄位不變 (Update)
        CASE "2"
            Main = DTSTransformstat_UpdateQuery
' 停用 -->Ifuse欄位改為'N' (User)
        CASE "3"
            DTSDestination("TP_SUP_IFUSE")="N"
            Main = DTSTransformstat_UserQuery
    End Select
End Function
    
```

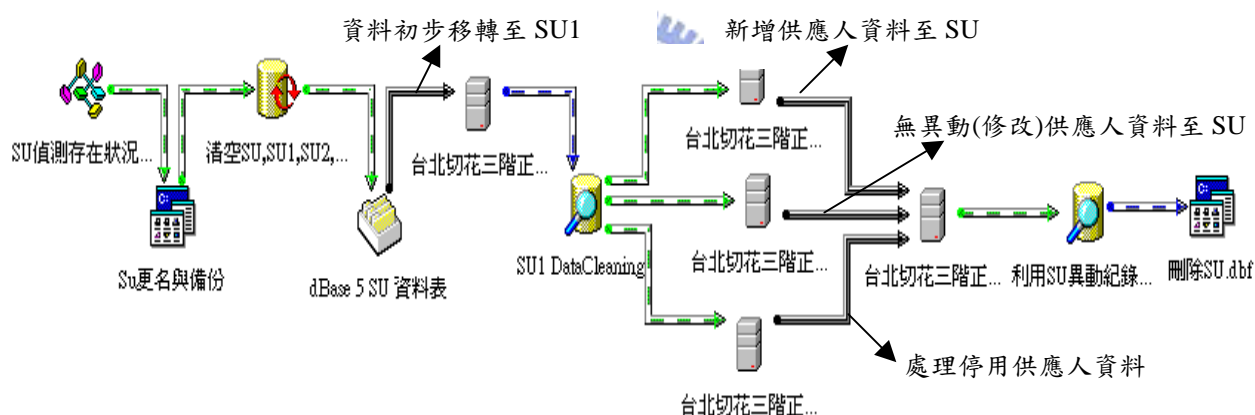


圖 3.10 供應商轉換服務部分

由於圖 3.9 (b).承銷人轉換部分和圖 3.9 (c).花卉資料轉換部分與圖 3.9 (a).供應商轉換部分的流程相似，兩者在資料的轉換部分上只有新增與修改的動作，並不需要判斷承銷人資格或花卉資料停用與否，因此不在贅敘。

在三階正規化 DTS-2 封裝中，包含有日進貨資料轉換服務、日交易資料轉換服務兩部分。日進貨資料轉換服務、日交易資料轉換服務兩部份的轉換過程相似，僅在部份資料轉換的步驟上有所差異，如圖 3.11 中圓形虛線註明的部分。在圖 3.12 說明日交易資料轉換服務部份中重要流程的主要目的，包含資料初步移轉、移除供應商對應失敗的資料、移除承銷人對應失敗的資料、移除花卉資料對應失敗的資料、交易資料之異動紀錄等步驟，而圖 3.11 虛線部分的差異在於日進貨資料轉換服務並無移除承銷人對應失敗的資料的步驟。

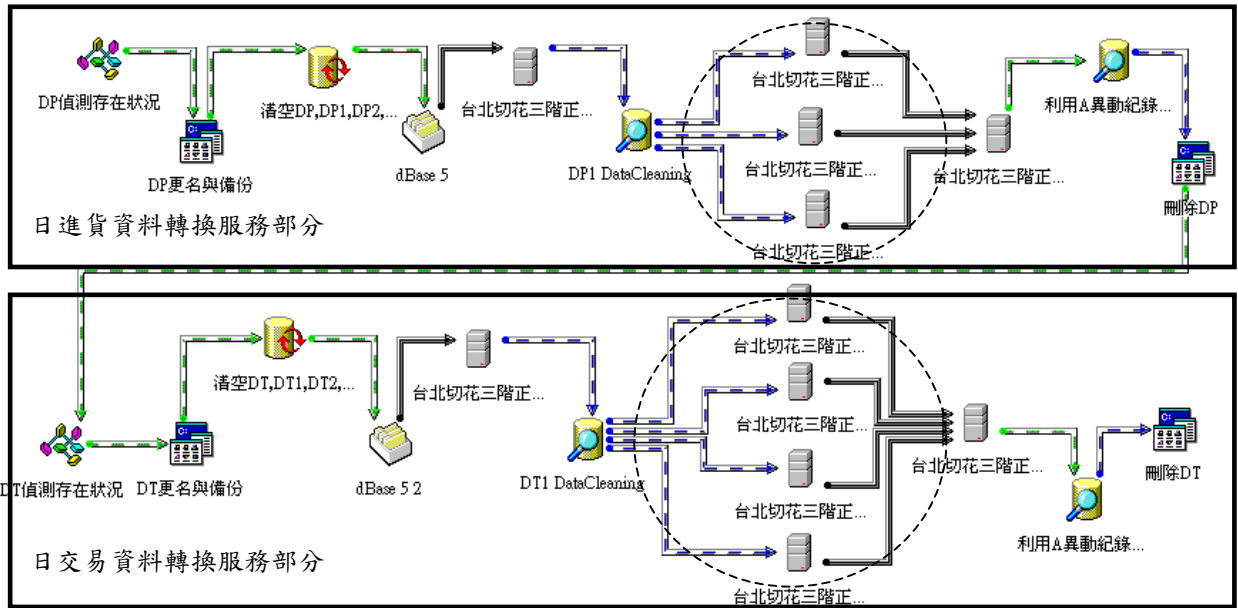


圖 3.11 三階正規化 2

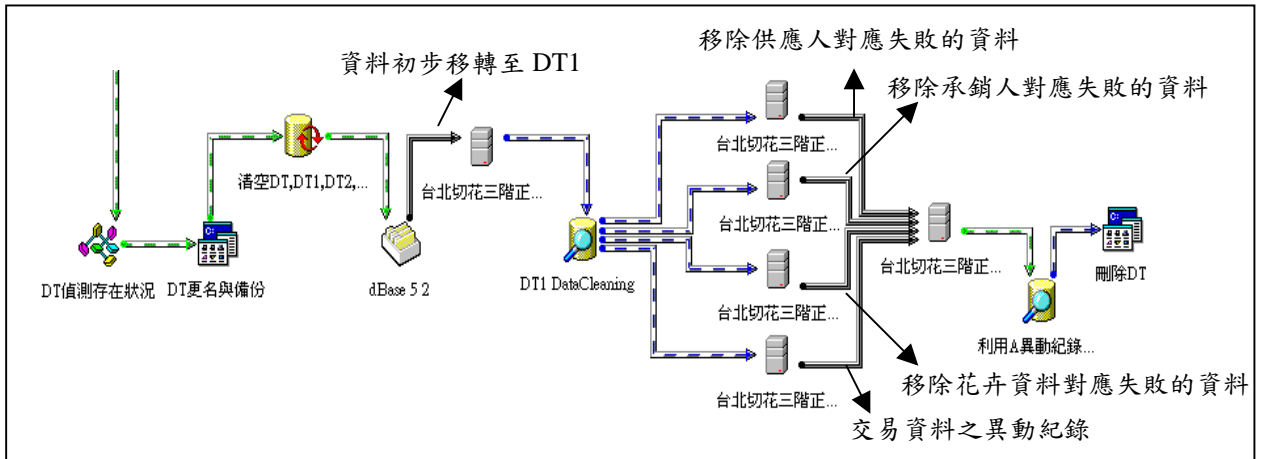


圖 3.12 日交易資料轉換服務部分

在三階正規化 DTS-2 封裝中主要就日交易資料轉換服務部份的轉換過程說明，在日進貨資料轉換完成後，首先偵測來源資料夾內是否有 DT.DBF 檔案之存在。若無則跳出錯誤結束資料轉換服務；若有 DT.DBF 檔案存在繼續進行以下程序。呼叫一外部.bat 檔將原始 DT 資料檔更名並備份資料至備份資料夾中。進行清空 DT 資料表以及 DT1、DT2 暫存資料表中資料的動作。將資料初步轉移至 DT1 資料夾中，此步驟中排除欄位 TP\_DT\_ID 為 0 與 7 的資料(TP\_DT\_ID 為紀錄交易類別之欄位，0 為進貨資料、7 為取消交易)，並將其餘資料依其交易類別儲存，其資料選擇之 SQL 語法如表 3.10。

表 3.10 資料選擇之 SQL 語法

SELECT	*
FROM	DT
WHERE	(TP_DT_ID NOT IN (0, 7))

在完成資料初步轉移之後，進行移除供應人對應失敗的資料、移除承銷人對應失敗的



資料、移除花卉資料對應失敗的資料、交易資料之異動紀錄等步驟，表 3.11 為選擇供應人對應失敗資料之 SQL 語法；表 3.12 為選擇承銷人對應失敗資料之 SQL 語法；表 3.13 為選擇花卉資料對應失敗資料之 SQL 語法；交易資料之異動紀錄步驟需先判斷資料為新增或修改後再進行資料的載入；若已有資料存在在 AACTTYPE 欄位設定為 2，若無則欄位設定為 1；表 3.14 為選擇除了對應失敗的資料外其餘資料的 SQL 語法。

表 3.11 選擇供應人對應失敗資料之 SQL 語法

```
SELECT *
FROM DT1
WHERE ASU NOT IN
      (SELECT TP_SUP_ONUM FROM SUTable WHERE TP_SUP_IFUSE = 'Y')
```

表 3.12 選擇承銷人對應失敗資料之 SQL 語法

```
SELECT *
FROM DT1
WHERE ABU NOT IN
      (SELECT TP_BUY_NUM FROM BUTable)
```

表 3.13 選擇花卉資料對應失敗資料之 SQL 語法

```
SELECT *
FROM DT1
WHERE APR NOT IN
      (SELECT TP_PR_NUM FROM PRTable)
```

表 3.14 選擇其餘資料之 SQL 語法

```
SELECT *,是否存在=(SELECT COUNT(*) FROM DTTable WHERE TP_DT_NUM=RTRIM(LTRIM(&DATE))+RTRIM(LTRIM(APM)))
FROM DT1
WHERE (ASU IN (SELECT TP_SUP_ONUM FROM SUTable WHERE TP_SUP_IFUSE = 'Y')) AND
      (ABU IN (SELECT TP_BUY_NUM FROM BUTable)) AND
      (APR IN (SELECT TP_PR_NUM FROM PRTable))
```

完成以上動作後，進行利用 A 異動紀錄將資料寫入拍賣資料表之步驟；在此步驟中根據前一步驟所對應之 AACTTYPE 欄位進行新增或修改的動作，欄位為 1 進行 InsertQuery 新增動作，欄位為 2 進行 UpTIMEQuery 更新動作。最後刪除 DT.DBF 檔案後結束三階正規化 DTS-2 封裝。

### 3.4.2 資料載入：DTS-2

當原始交易資料檔成功匯入關聯式資料庫後，便可著手進行從資料庫將資料載入至資料倉儲的動作，此動作對應於系統的架構中資料轉換服務 DTS-2 的部分。資料轉換服務 DTS-2 主要包含三個封裝：倉儲供應團體維度、倉儲其餘維度部份以及倉儲事物表部分三個封裝。與資料轉換服務 DTS-1 相同，由於受到主鍵外鍵的限制，所以須先完成倉儲供應團體維度封裝，接著倉儲其餘維度部份封裝後，才能進行倉儲事物表部分的封裝。

就「倉儲供應團體維度」封裝而言，在雪花網要供應團體資料表中包含 4 個欄位除了 TP\_SUPGROUP\_NAME 為 TP\_SUP\_ONUM+TP\_SUP\_NAME 組成外，其餘皆可從資料庫

供應人資料表中對應的欄位直接載入，倉儲供應團體維度封裝如圖 3.13 所示。



圖 3.13 倉儲供應人團體維度

就「倉儲其餘維度部分」封裝而言，在供應人資料載入部分，雪花網要供應人資料表中 TP\_SUP\_NAME 為 TP\_SUP\_ONUM+TP\_SUP\_NAME 所組成，TP\_SUP\_COUNTRY 取其供應人代碼前兩碼；在花卉資料載入部分，雪花網要花卉資料表中 TP\_PR\_NAME 為 TP\_PR\_NUM+TP\_PR\_NAME 所組成；在倉儲其餘維度部分封裝中除了以上三個欄位需進行資料的整合轉換外，其餘欄位皆可從資料庫中對應的欄位直接載入，倉儲其餘維度部分封裝如圖 3.14 所示。

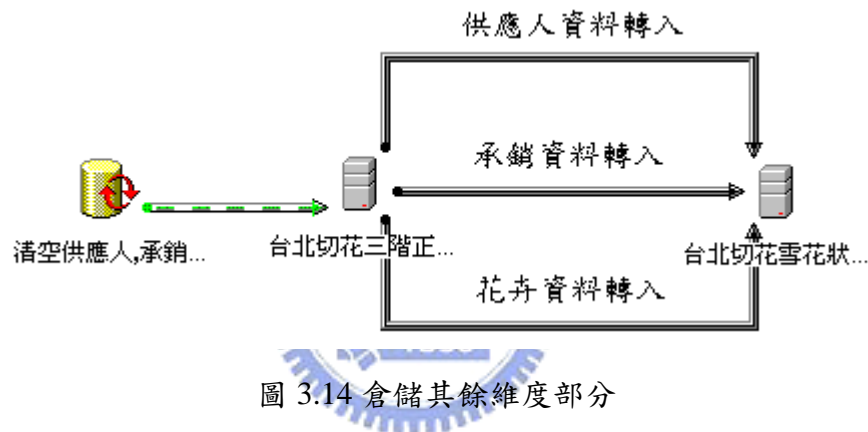


圖 3.14 倉儲其餘維度部分

就「倉儲事物表部分」封裝而言，需先判斷交易別 TP\_DT\_ID 後，根據不同的交易別將 TP\_DT\_PCS(拍賣把數)與 TP\_DT\_BOX(拍賣箱數)存入相對應的事物表衡量值中，其餘欄位皆可從資料庫中對應的欄位直接載入，倉儲事物表部分封裝如圖 3.15 所示。



圖 3.15 倉儲事物表部分

在資料轉換 DTS-1 以及資料載入 DTS-2 完成後，資料倉儲的建構大抵完成，後續動作還需建立資料倉儲資料轉取的自動化，但因為非本論文所需實作的部分，故不予說明。

### 3.5 線上分析處理系統

在本節中，將說明如何利用 Cognos 商業情報系統中三套軟體 (Impromptu、Transformer、PowerPlay Web Server) 來建構線上分析處理系統。步驟主要分為目錄設計、報表製作、建置超方體、發佈超方體四步驟，詳密步驟說明如下。

首先即是目錄設計的部份，此部份的工作重點為與資料來源的 SQL Server 資料倉儲建立連結，並將其設定儲存以供作為報表製作的基礎。實作方法為利用 Impromptu Administrator 建立目錄 (Catalog) 並儲存成 \*.cat 檔，再連結台北花卉資料倉儲 (FLOWERBASE) 伺服器中的 TPCDW 資料庫，如圖 3.16。接著選取需要在線上分析中展示的资料表並將其移至目錄資料表 (Catalog Tables)，如圖 3.17 所示。

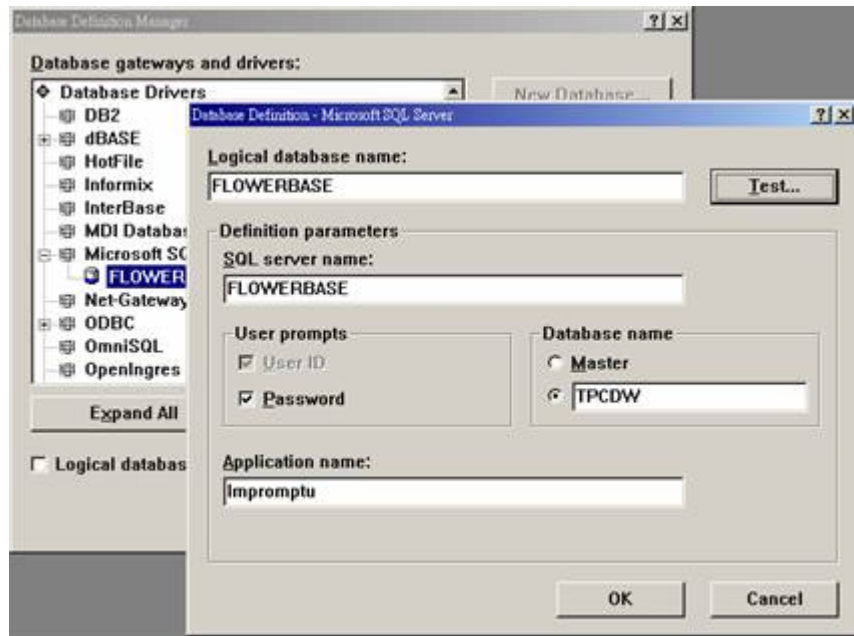


圖 3.16 資料庫的選取與聯結

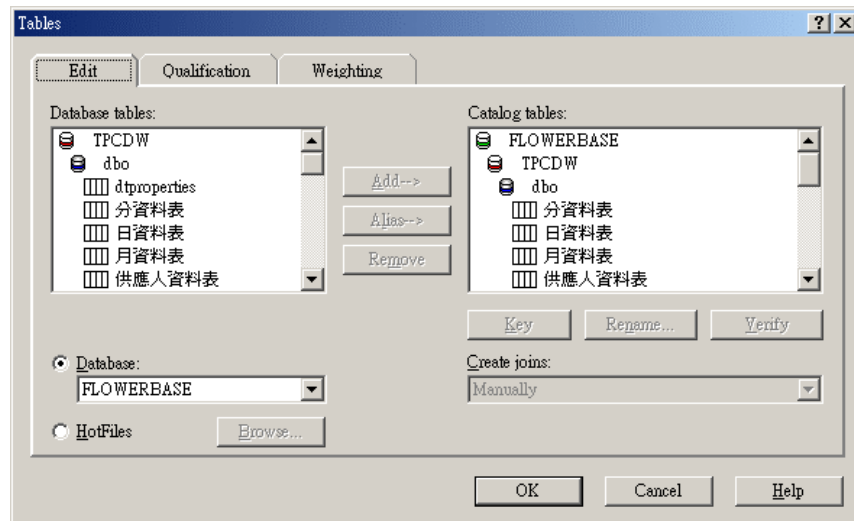


圖 3.17 資料表選取

在選取完需要的資料表後，利用 Impromptu Administrator 內預設的 Join 功能設定各資料表中的關聯性，在關聯性部分可以參照台北資料倉儲的雪花網要圖。下圖 3.18 為供應人資料表與供應鄉鎮資料表的關聯性設定。

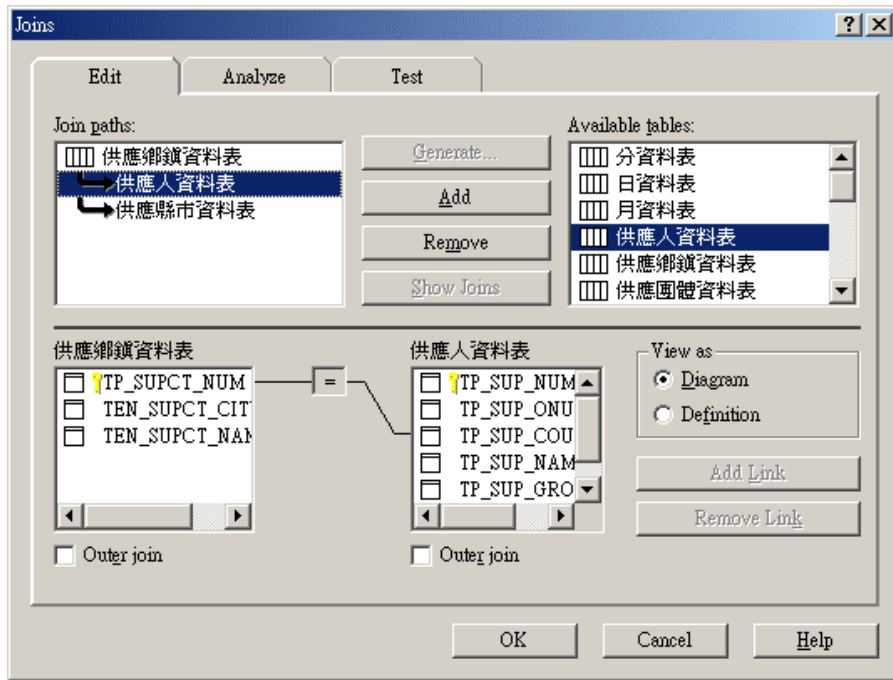


圖 3.18 資料表關聯設定

目錄設定完成之後，如圖 3.19 選取需要呈現在簡易報表的維度與衡量值。如此即可製成簡易的報表以供查詢，如圖 3.20。將報表以內建的\*.imr 格式存檔以外，還須另存\*.iqd 檔以作為製作超方體(Hyper Cube)工具 PowerPlay Transformer 的資料來源。

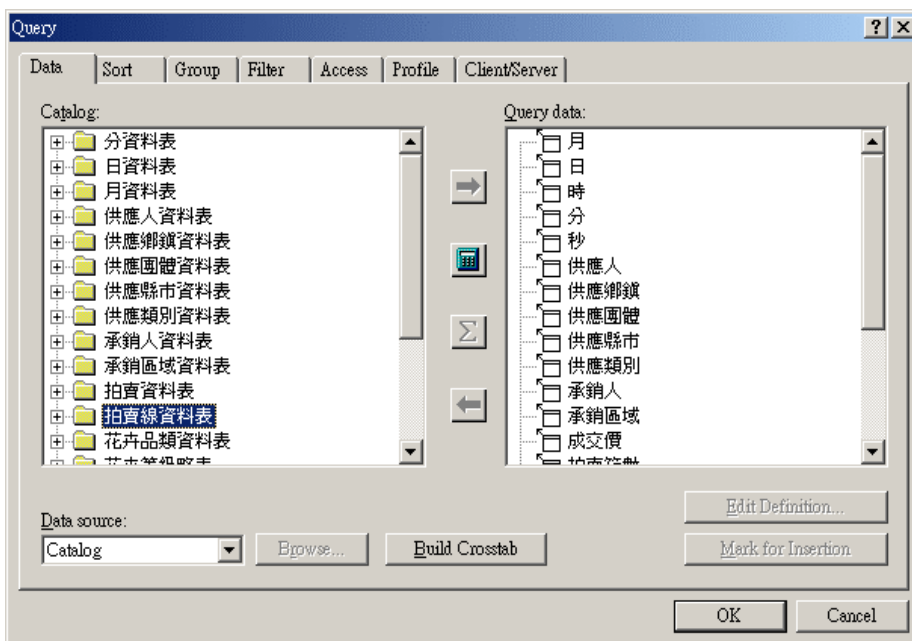


圖 3.19 維度與衡量值的選取

## &lt;時間維度擴增實作&gt;

月	日	時	分	秒	供 應 人	供 應 縣 市	供 應 鄉 鎮	供 應 類 別
01	19	00	00	00	JLA26	台中縣	神岡鄉	農會合作社(場)
01	19	00	00	02	JFL16	台中縣	后里鄉	農會合作社(場)
01	19	00	00	02	JFL16	台中縣	后里鄉	農會合作社(場)
01	18	00	00	03	GKX67	桃園縣	大溪鎮	進口商
01	18	00	00	03	GKX67	桃園縣	大溪鎮	進口商
01	18	00	00	03	GKX67	桃園縣	大溪鎮	進口商
01	19	00	00	03	JFL16	台中縣	后里鄉	農會合作社(場)
01	18	00	00	04	JFE00	台中縣	后里鄉	農會合作社(場)
01	19	00	00	10	JLA26	台中縣	神岡鄉	農會合作社(場)
01	19	00	00	16	LFQ89	南投縣	埔里鎮	農會合作社(場)
01	19	00	00	16	LFQ89	南投縣	埔里鎮	農會合作社(場)
01	18	00	00	20	JFR10	台中縣	后里鄉	農會合作社(場)
01	19	00	00	20	LFQ89	南投縣	埔里鎮	農會合作社(場)
01	18	00	00	21	GKX67	桃園縣	大溪鎮	進口商
01	18	00	00	21	GKX67	桃園縣	大溪鎮	進口商
01	19	00	00	21	JLA11	台中縣	神岡鄉	農會合作社(場)
01	19	00	00	21	JLA11	台中縣	神岡鄉	農會合作社(場)
01	19	00	00	31	LFC78	南投縣	埔里鎮	農會合作社(場)
01	18	00	00	36	GKX67	桃園縣	大溪鎮	進口商
01	19	00	00	37	LFC78	南投縣	埔里鎮	農會合作社(場)
01	18	00	00	38	MPAA1	雲林縣	北港鎮	花旗
01	19	00	00	48	LFC82	南投縣	埔里鎮	農會合作社(場)
01	18	00	00	53	NPJ00	嘉義縣	新港鄉	農會合作社(場)
01	18	00	00	55	GKX67	桃園縣	大溪鎮	進口商

圖 3.20 簡易報表

利用前項動作所產生的\*.iqd 檔案為資料來源來設計超方體，由於系統的預設值會自動設計超方體，為了避免系統因自動設計而跟預期的設定不一致，因此建議不勾選自動產生超方體(Run AutoDesign)。進入 PowerPlay Transformer 後，主畫面如圖 3.21。其中有五個小視窗，前四個視窗為本研究的主要設計內容，分別為維度視窗(Dimension Map)、資料來源視窗(Data Source)、衡量值視窗(Measures)、超方體視窗(Power Cubes)。而登入視窗主要為控制使用者權限，並不在本研究的討論範圍。

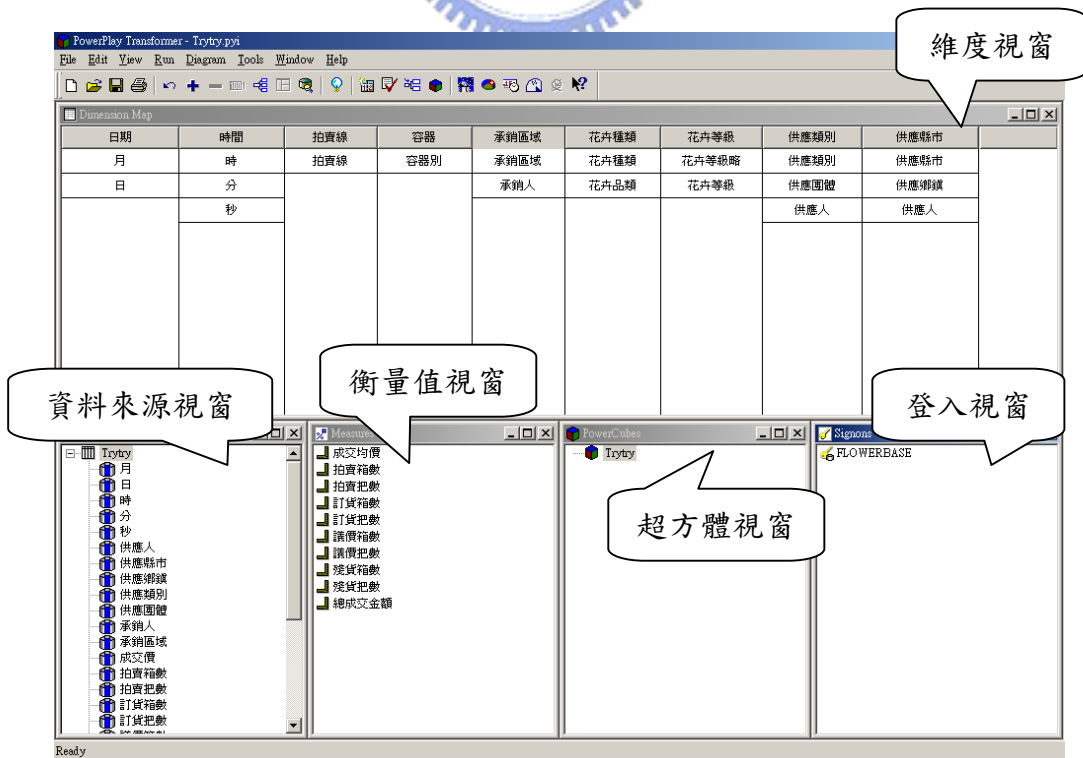


圖 3.21 PowerPlay Transformer 架構之超方體



從維度視窗以及衡量值視窗可以看出，本系統建構出 10 個衡量值及 9 個維度，衡量值分別為成交均價、拍賣箱數、拍賣把數、訂貨箱數、訂貨把數、議價箱數、議價把數、殘貨箱數、殘貨把數、總成交金額，而維度的詳細資料如表 3.10 所示。在超方體視窗中的超方體完成後，需將其儲存為\*.mdc 檔以供 PowerPlay Enterprise Server 發佈至 Internet 中。

表 3.10 維度說明

編號	維度名稱	階層式單位說明
1	日期	分為月、日等時段。
2	時間	分為時、分、秒等狀態。
3	拍賣線	分為 C、D、E、F 四條拍賣線與未經拍賣線。
4	容器別	分為紙箱、立式容器兩大類。
5	承銷區域	分為北部、中部、雲嘉、其他區域等 8 個區域。
6	花卉種類	分為國產花與進口花，可再細分到各花卉種類。
7	花卉等級	分為特級、優級、良級、未分級四大類。
8	供應類別	分為農會、合作社、花班、個人花農等多種。
9	供應縣市	分為台灣各縣市，可再細分至各鄉鎮。

在完成超方體的維度與衡量值架構後，利用先前轉換出的超方體(\*.mdc)透過 PowerPlay Enterprise Server 發佈至 Internet 中，以供使用者可以進行線上使用。圖 3.22 為發佈超方體至 Internet 的動作。圖 3.23 為線上分析處理系統的畫面。

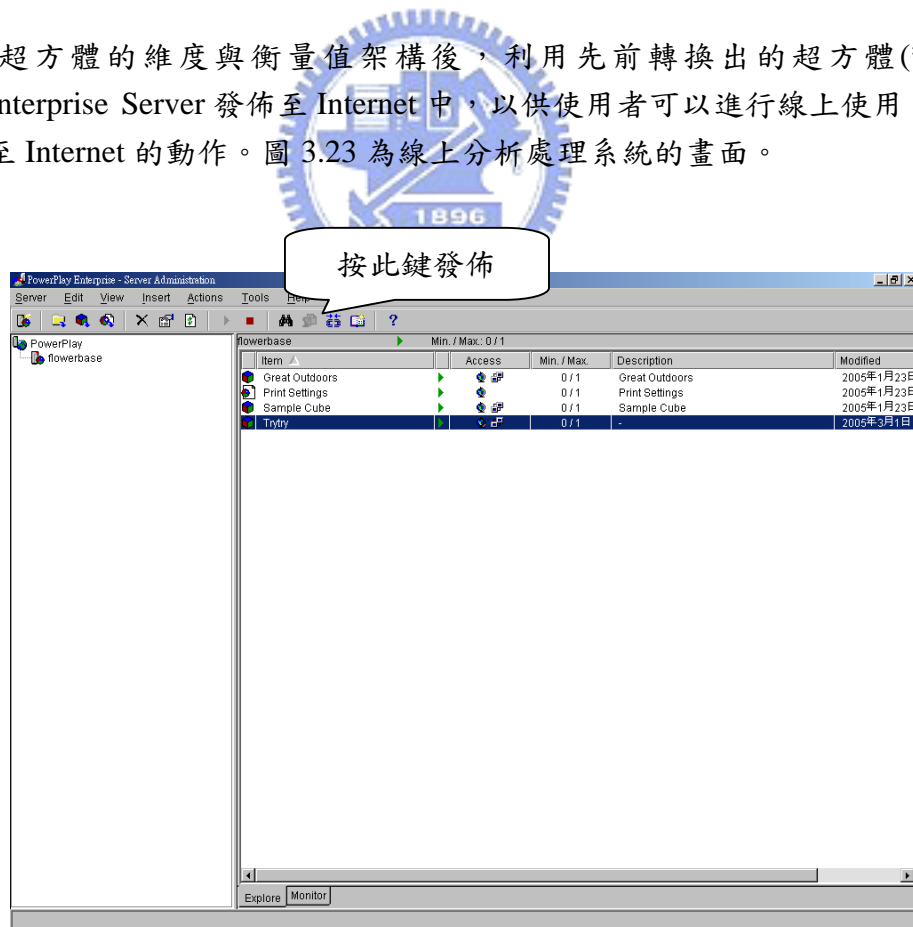


圖 3.22 發佈超方體至 Internet

Trytytry確定版 - Cognos PowerPlay Web Explorer - Microsoft Internet Explorer

檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)

網址(D) http://lowerbase/cognos/cgi-bin/ppds.cgi.exe?DC=Q&E=Trytytry%BDT%AA%A9&LA=tw&LO=tw&BACK=cognos/cgi-bin/ppds.cgi.exe?bc=%26LA%3Dtw%26L...

Cognos PowerPlay Web Explorer Trytytry確定版

日期 時間 拍賣線 容額 承銷區域 花卉種類 花卉等級 供應類別 供應縣市

成交價 作為數值	▽ 00點	▽ 01點	▽ 02點	▽ 03點	▽ 04點	▽ 05點	▽ 06點	▽ 07點	▽ 08點	▽ 18點	▽ 19點	▽ 20點	▽ 21點	▽ 22點	▽ 23點	時間
台中縣	85	128	124	69	103	80	55	52	52	118	111	102	105	97	98	89
桃園縣	99	0	76	51	71	86	90	86	20	99	125	79	86	94	121	79
南投縣	76	62	84	45	95	73	46	23	7	76	85	74	50	64	52	77
雲林縣	80	70	81	59	56	59	41	25	0	80	60	74	65	55	50	56
嘉義縣	70	41	59	51	84	65	39	26	13	111	54	66	72	41	56	69
彰化縣	44	48	61	40	47	43	37	28	23	41	57	60	60	58	48	44
台北市	46	76	103	53	57	61	59	28	9	36	74	155	61	110	114	57
台北縣	31	0	42	28	58	75	33	15	17	0	94	0	13	144	0	57
台南縣	19	14	31	19	33	24	17	15	12	10	31	19	29	15	30	24
屏東縣	75	67	48	52	44	42	42	37	91	61	67	90	37	46	67	44
高雄縣	52	74	98	241	106	62	42	23	28	0	70	96	9	24	140	103
苗栗縣	45	45	85	45	94	67	57	28	99	51	76	76	65	40	83	75
宜蘭縣	14	11	17	16	28	21	15	17	6	0	10	10	11	14	17	19
新竹縣	0	0	181	47	47	41	75	73	15	0	155	94	133	0	64	49
花蓮縣	0	0	107	46	99	86	76	0	0	0	0	0	0	0	0	87
高雄市	0	0	0	528	20	10	15	0	0	0	0	0	0	0	0	96
供應縣市	62	67	82	47	75	61	47	35	29	74	77	72	69	72	66	64

近端內部網路

圖 3.23 線上分析處理系統

