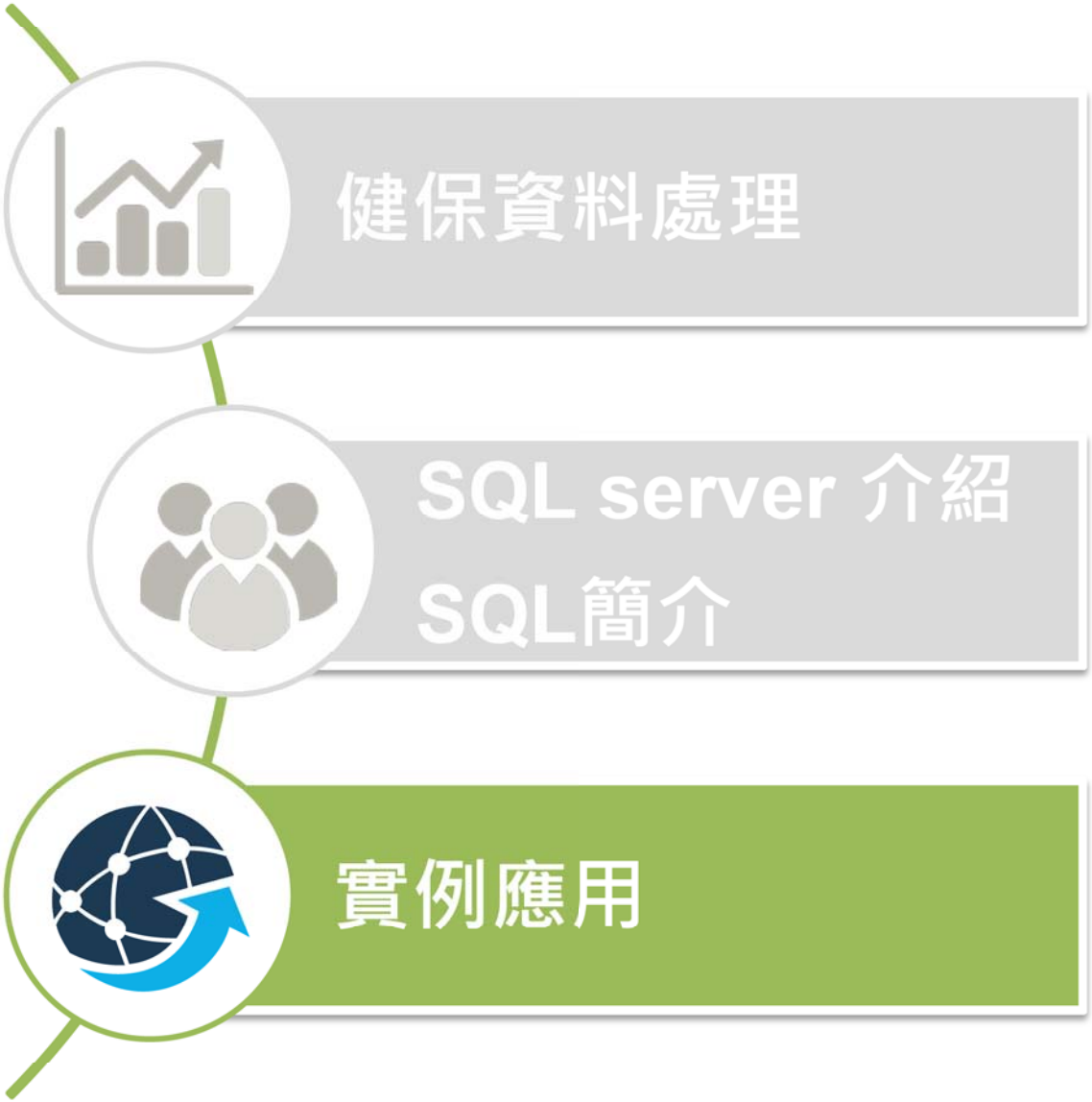
實例應用

SQL Server
健保資料處理



[範例] Doctor shopping 行為

- 逛醫院 (doctor shopping)，指病患因同一疾病到不同醫院就診。
- 影響：
  - 民眾：重複檢查、重疊用藥、喪失診斷先機
  - 資源：浪費醫療資源、浪費醫師時間
  - 1999年我國每人平均門診次數14.1次，經濟合作暨發展組織國家每人平均門診次數多為5~6次 (OECD 2001, 中央健保局 2002)
  - 台灣民眾會固定找同一基層醫療醫師的患者佔20%～24%
  - **臺醫學人士指SARS疫情暴露島內醫療體系弱點**

    台灣新增SARS病例和死亡人數再創記錄。台灣SARS疫情為什麼遲遲得不到控制？有醫學界人士認為，這與台灣的醫療體系和民眾的看病習慣有關。

1. Am J Manag Care. 2012 Sep;18(9):488-96.

Outpatient-shopping behavior and survival rates in newly diagnosed cancer patients.

Chiou SJ, Wang SI, Liu CH, Yaung CL.

Source

Asia University, 500 Lioufeng Rd, Wufeng Taichung, Taiwan 41354, Taiwan. E-mail: chiou@asia.edu.tw.

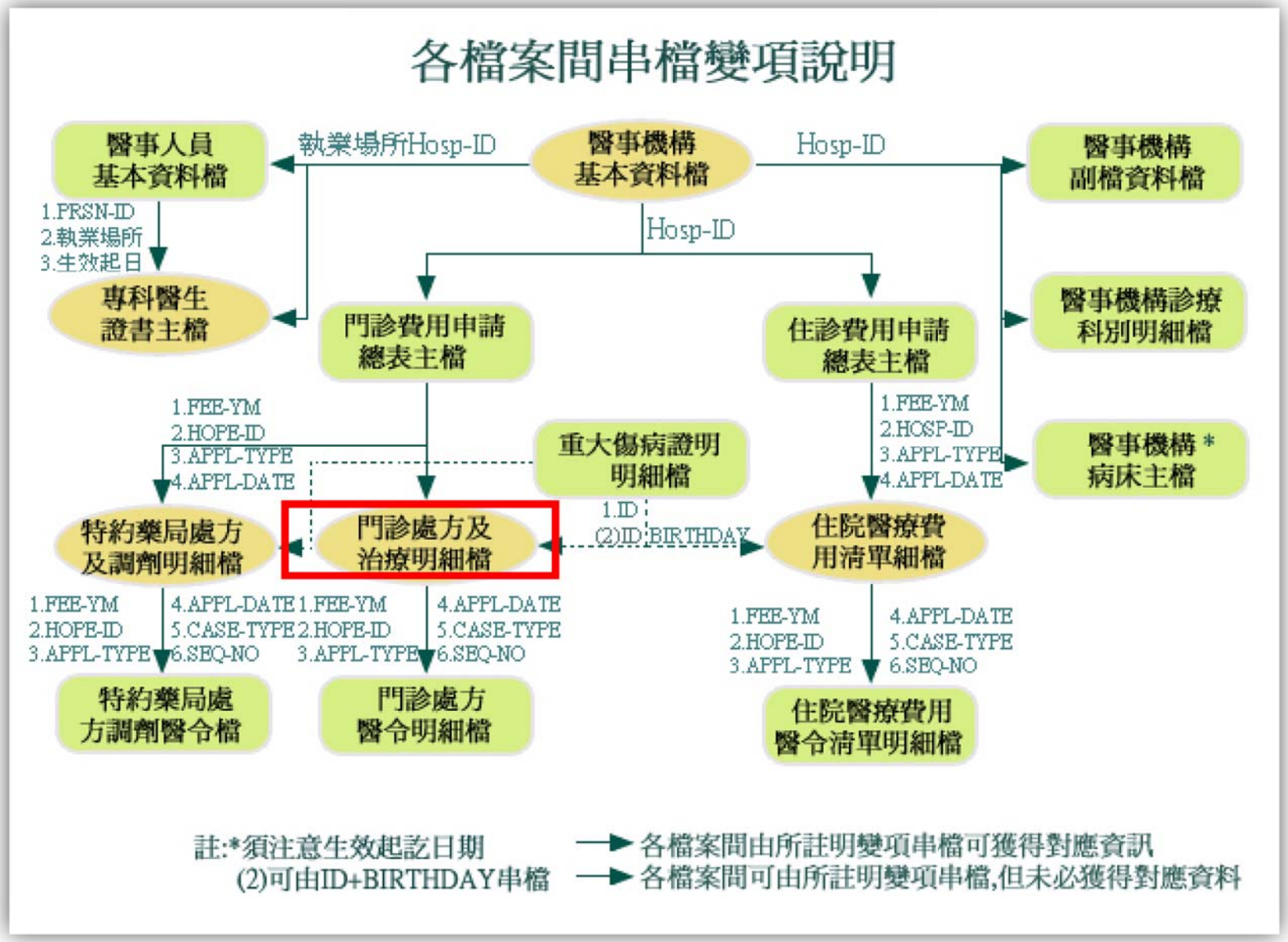
Abstract

Objectives: To evaluate the appropriateness of the definition of outpatient-shopping behavior in Taiwanese patients. Study Design: Linked study of 3 databases (Taiwan Cancer Registry, National Health Insurance [NHI] claim database, and death registry database). Methods: Outpatient shopping behavior was defined as making at least 4 or 5 physician visits to confirm a cancer diagnosis. We analyzed patient-related factors and the 5-year overall survival rate of the outpatient-shopping group compared with a nonshopping group. Using the household registration database and NHI database, we determined the proportion of outpatient shopping, characteristics of patients who did and did not shop for outpatient therapy, time between diagnosis and start of regular treatment, and medical service utilization in the shopping versus the nonshopping group. Results: Patients with higher incomes were significantly more likely to shop for outpatient care. Patients with higher comorbidity scores were 1.4 times more likely to shop for outpatient care than patients with lower scores. Patients diagnosed with more advanced cancer were more likely to shop than those who were not. Patients might be more trusting of cancer diagnoses given at higher-level hospitals. The nonshopping groups had a longer duration of survival over 5 years. Conclusions: Health authorities should consider charging additional fees after a specific outpatient- shopping threshold is reached to reduce this behavior. The government may need to reassess the function of the medical sources network by shrinking it from the original 4 levels to 2 levels, or by enhancing the referral function among different hospital levels.

PMID: 23009299 [PubMed - in process]

http://www.ajmc.com/publications/issue/2012/2012-9-vol18-n9/Outpatient-Shopping-Behavior-and-Survival-Rates-in-Newly-Diagnosed-Cancer-Patients

2. From Study to Data



2000年承保抽樣歸人檔

全民健康保險研究資料庫
National Health Insurance Research Database
ENGLISH

簡介 | 資料庫內容 | 相關規定 | 申請作業 | 成果 | 學術活動 | On-site服務 | 常見問題 | 回首頁 | 意見箱 | 國衛院首頁

查詢

資料清單

下載全部清單(Excel 檔)

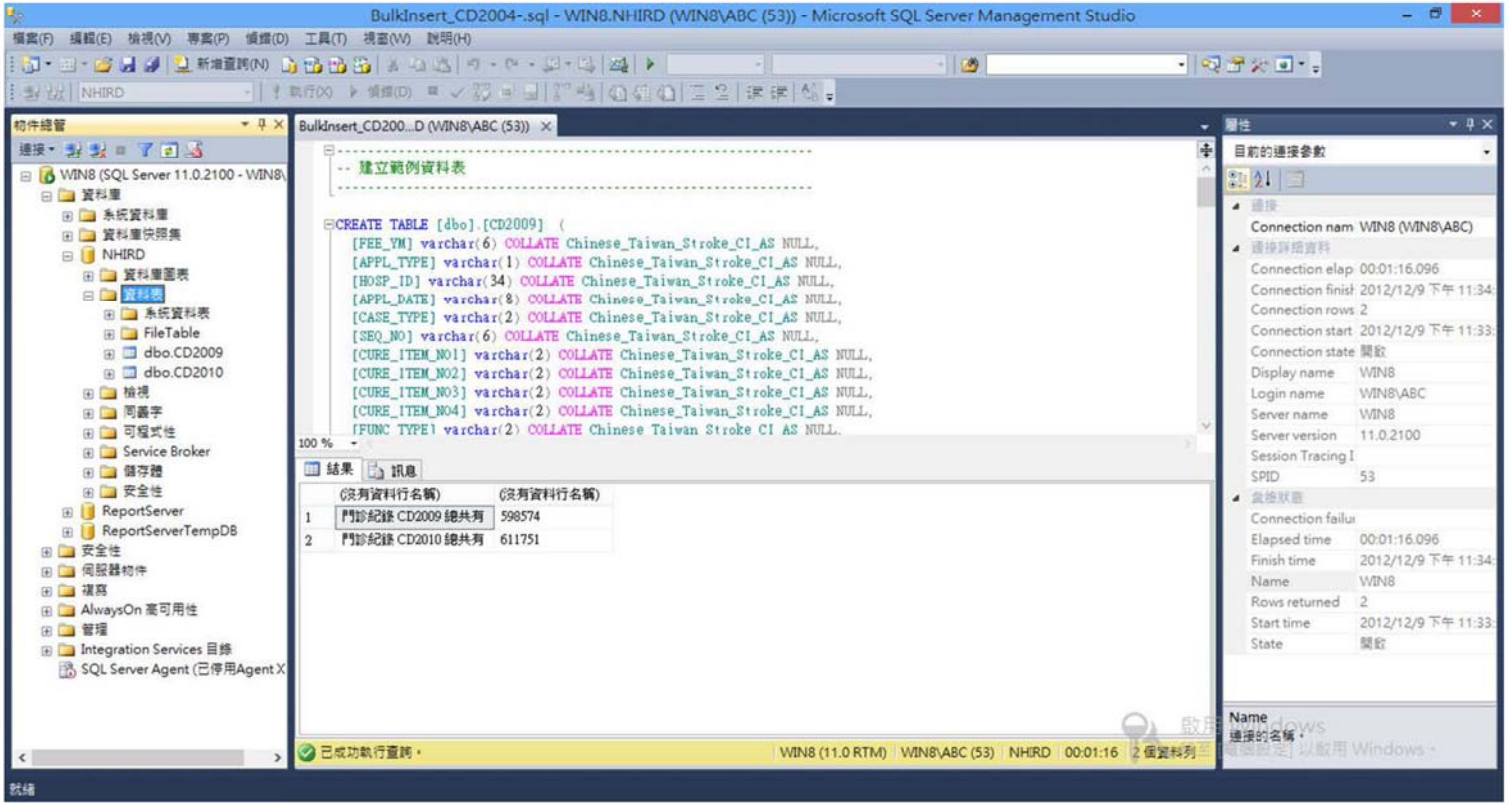
抽樣歸人檔清單
2000年承保抽樣歸人檔
2005年承保抽樣歸人檔
NEW 2010年承保抽樣歸人檔
各年資料清單
85年
86年
87年
88年
89年

2000年承保抽樣歸人檔

請注意：
2000年承保抽樣歸人檔(簡稱LHID2000)，將五萬人一組一年的資料統一以一片CD光碟片發行，因有數個年度五萬人的資料量接近一片CD光碟片之最大容量，易發生燒錄不良狀況，故自即日起LHID2000不再提供CD光碟片，改為燒製DVD光碟片提供，而收費標準不變，以每一片CD光碟片的資料新台幣500元計價。

STEP1. 資料轉換

- 將健保資料庫原始檔案匯入
- 1) 開啟新 Database, 2) 開啟新 table 3) 匯入資料



STEP 2a.變數處理: 跨年度串連 (>”<)

SQLQuery1.sql - 1...N-NB\zCBRAIN (53))

```sql
--- 如果將逛醫院定義為一年看超過 4 位醫師
--- 針對 2009 年逛醫院者: 請問 2009 年看病次數 與 2010 年看病次數 各為何 ??
with shopper as (
  select id, count(tid) as [2009年門診次數] from CD2009 group by id having count(distinct PRSN_ID) > 4
), visit2010 as (
  select id, count(tid) as [2010年門診次數] from cd2010 group by id
)
  select shopper.ID, [2009年門診次數], [2010年門診次數]
  from shopper join visit2010 on shopper.ID=visit2010.ID;
```

100 %

Results | Messages

	ID	2009年門診次數	2010年門診次數
1	0000000000000000000000000000028195	15	16
2	0000000000000000000000000000019412	24	9
3	0000000000000000000000000000018663	30	18
4	0000000000000000000000000000024961	14	21
5	0000000000000000000000000000025187	9	7
6	0000000000000000000000000000034876	6	9
7	0000000000000000000000000000027226	25	34
8	0000000000000000000000000000007407	12	26
9	0000000000000000000000000000033313	14	8
10	0000000000000000000000000000039516	49	45

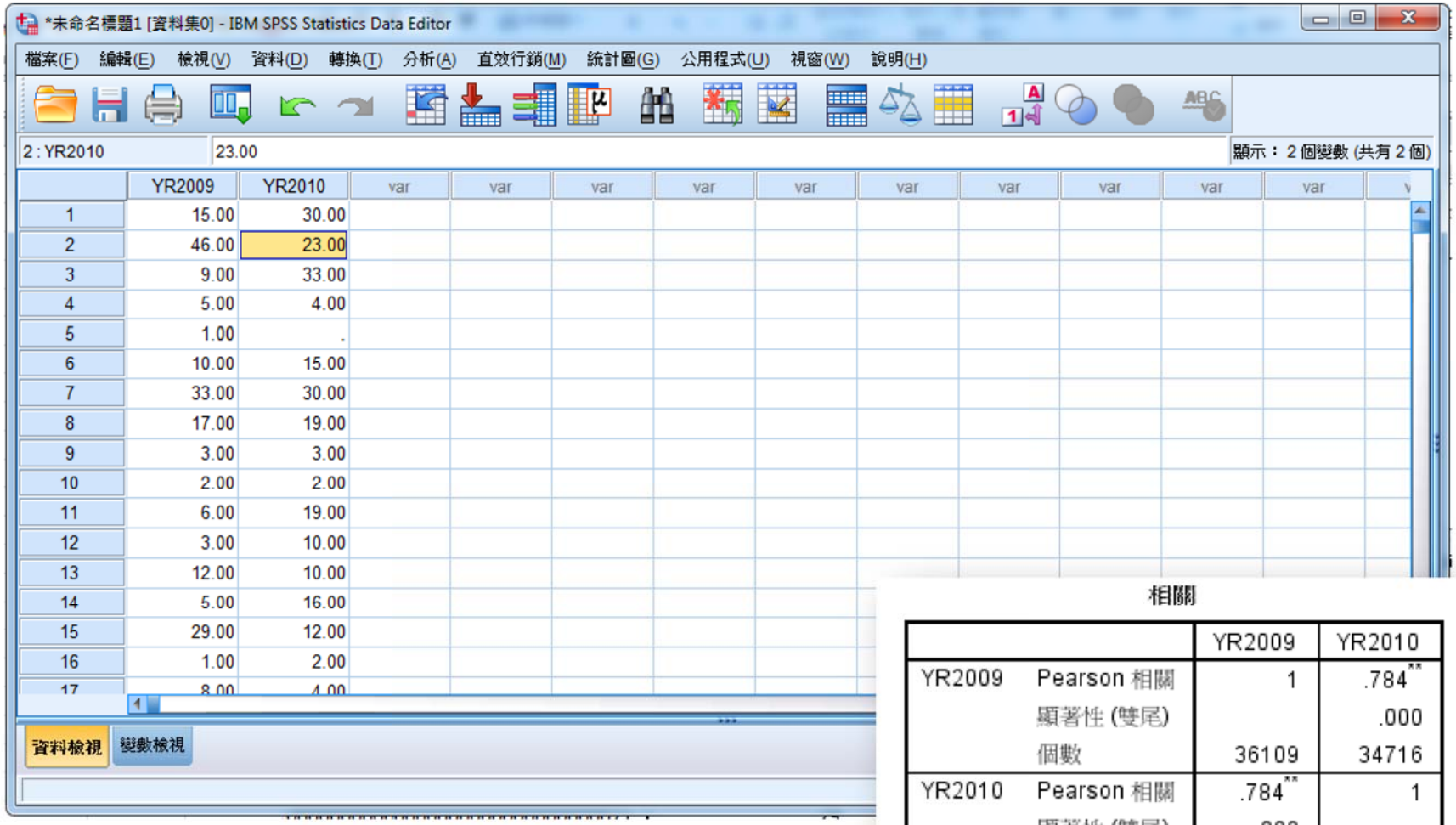
STEP 3. 統計分析

1. Copy-Paste 複製貼上

   → to SPSS, EXCEL

2. 匯出成 文字檔案 (CSV檔)

3. 透過 ODBC 連接



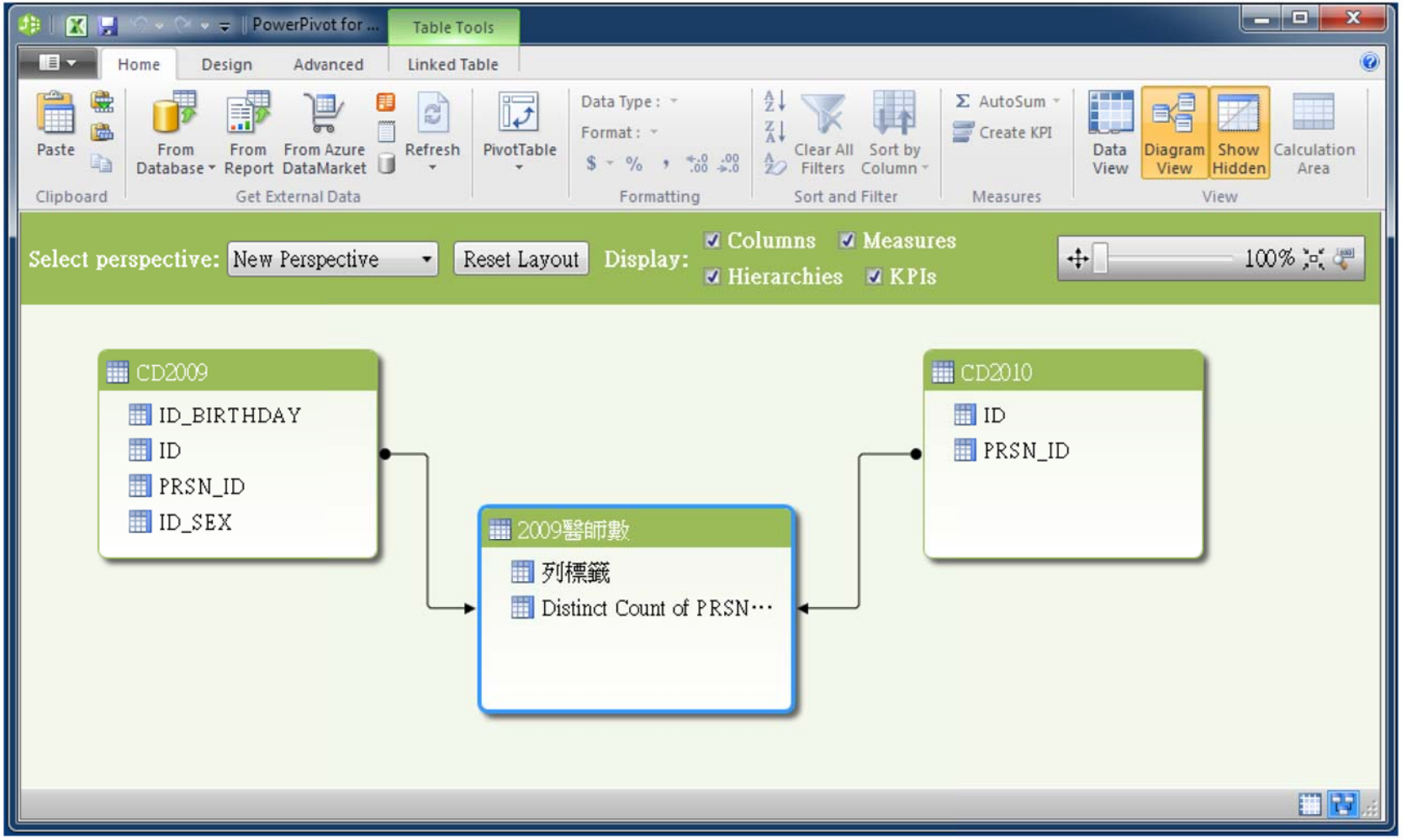
	YR2009	YR2010
1	15.00	30.00
2	46.00	23.00
3	9.00	33.00
4	5.00	4.00
5	1.00	.
6	10.00	15.00
7	33.00	30.00
8	17.00	19.00
9	3.00	3.00
10	2.00	2.00
11	6.00	19.00
12	3.00	10.00
13	12.00	10.00
14	5.00	16.00
15	29.00	12.00
16	1.00	2.00
17	8.00	4.00

相關

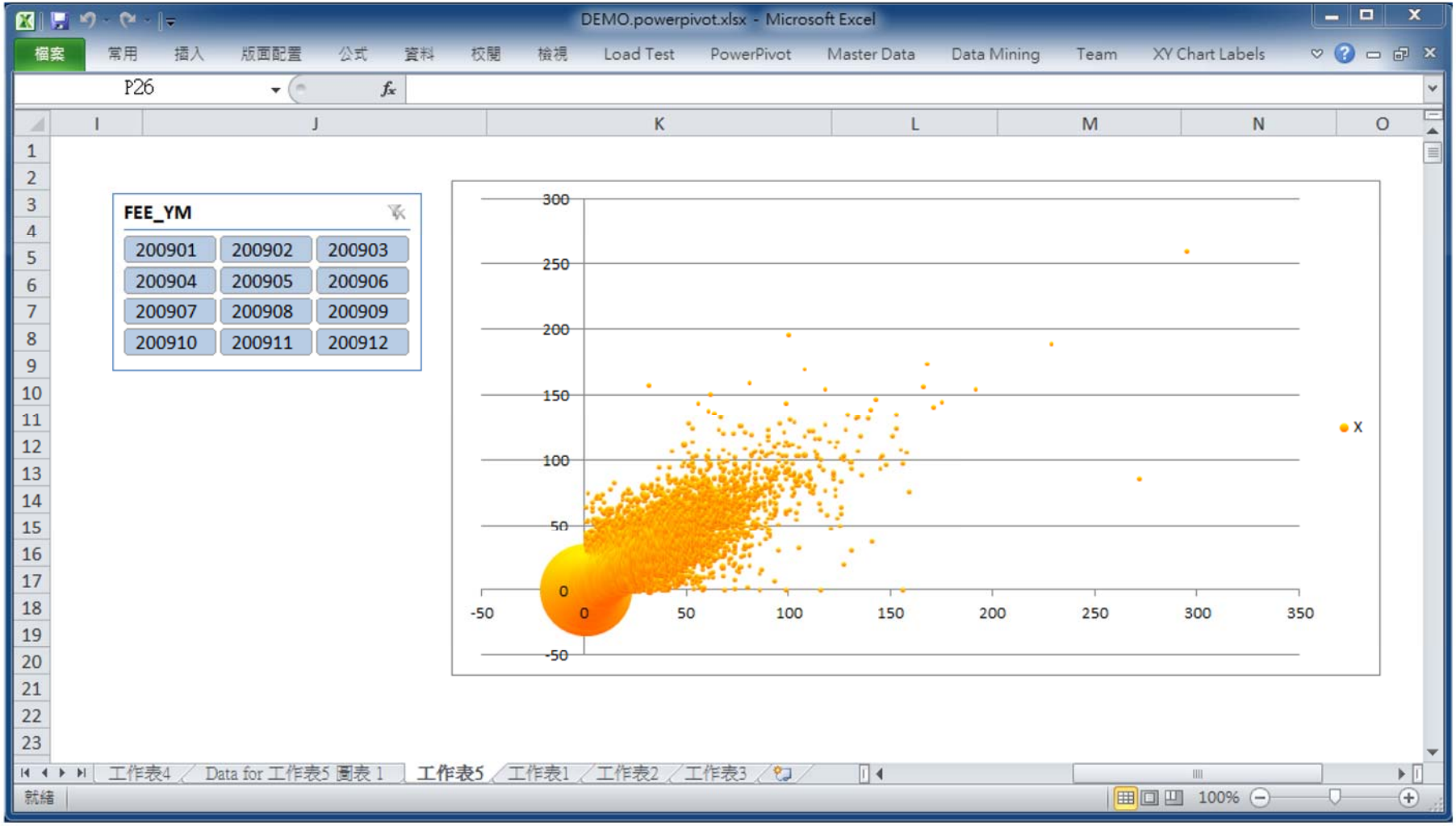
		YR2009	YR2010
YR2009	Pearson 相關	1	.784**
	顯著性 (雙尾)		.000
	個數	36109	34716
YR2010	Pearson 相關	.784**	1
	顯著性 (雙尾)	.000	
	個數	34716	34716

**. 在顯著水準為0.01時 (雙尾)，相關顯著。

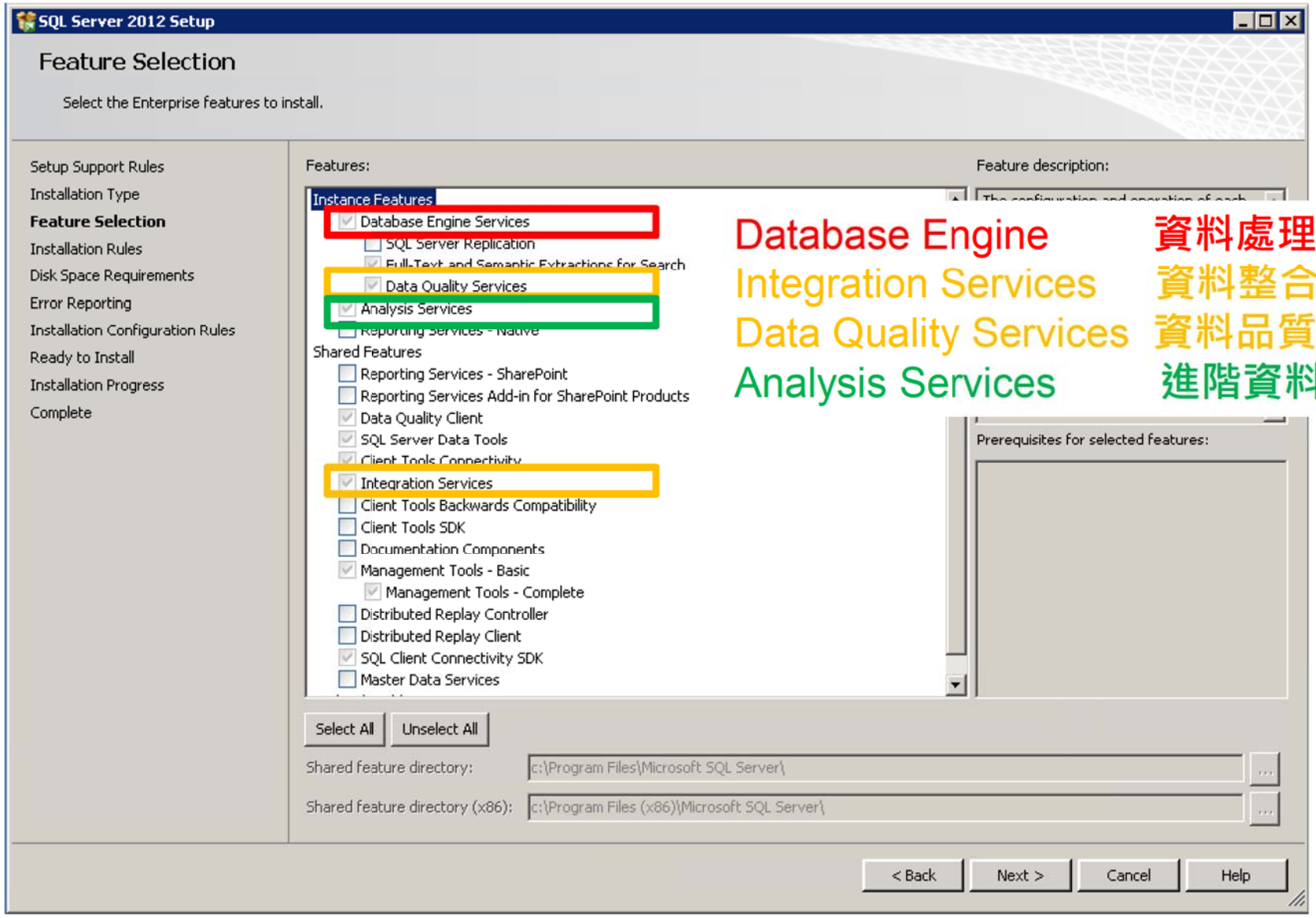
STEP 2b: 利用 PowerPivot 串連資料



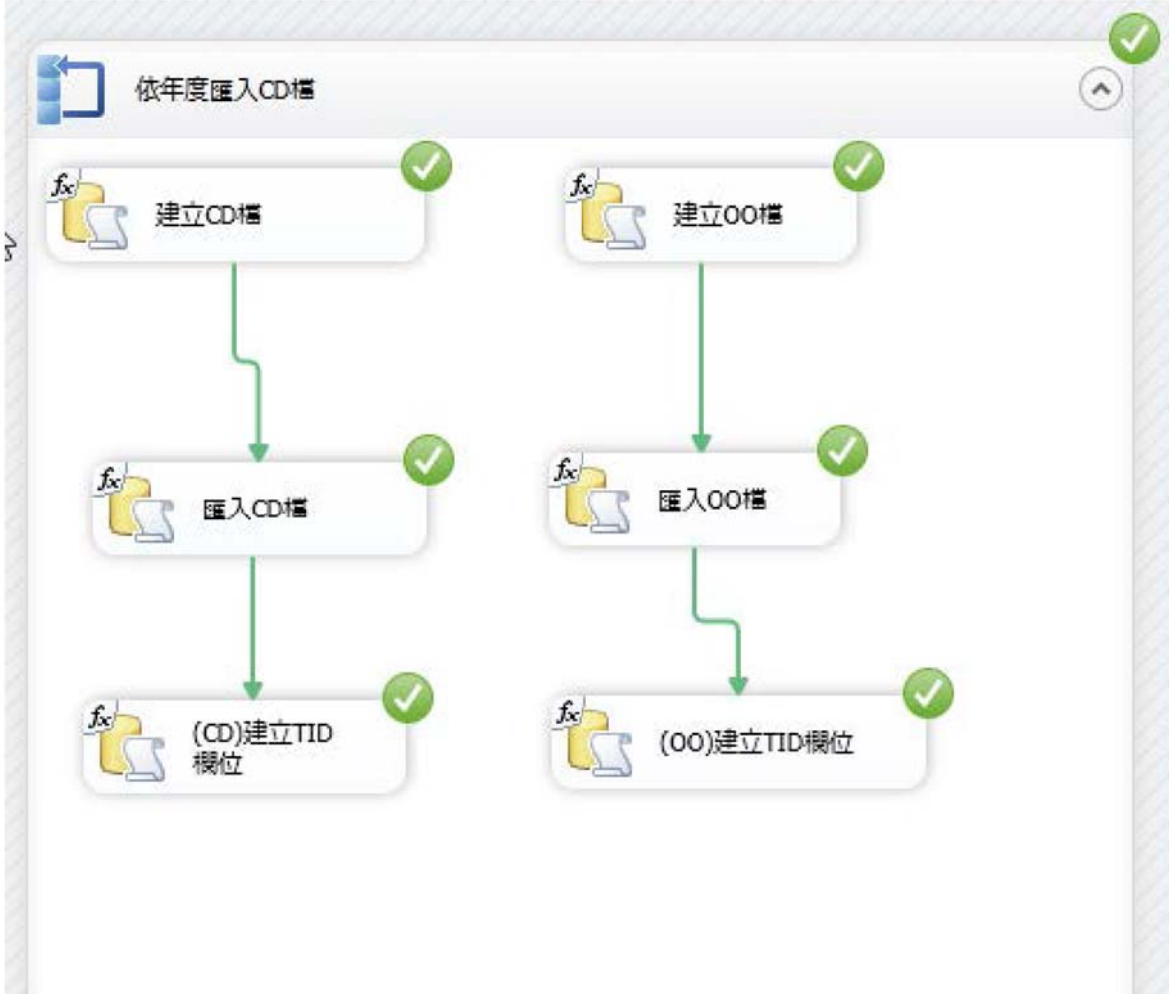
利用 PowerPivot 動態分析資料



New features in SQL server 2012



SSIS: 資料整合轉換



Pros

- 圖形化介面
- 自動進行批次(大量)資料匯入
- 資料匯入一目了然

Cons

- 需要時間學習介面
- 需要時間開發

謝謝敬請指教

- SQL 與 RDBMS 是專門用來處理資料的工具。
- SQL 相當於資料處理的標準語言，適用複雜資料串連，到處通用。
- 資料庫處理軟體與現有統計軟體結合互補提昇品質與生產效能。