



第 6 章

時態數據倉庫



時態數據倉庫是一種倉庫技術，其中時態（或歷史）方面of記錄被併入倉庫。例如，如果一個書是一個維度（在書店環境中），書價隨時間變化。要是我們w想在數據倉庫中跟踪圖書價格，然後數據倉庫技術不能使用。沒有這個時間或歷史方面of書價，數據倉庫的分析僅基於目前的書價。

本章重點介紹涵蓋星型模式的時態數據倉庫，其中保持時間I數據方面。時態數據倉庫也被稱為作為緩慢變化的維度(SCD)。本章還將介紹各種可以使用的技術處理 SCD。

6.1 書店案例研究

學習時態數據倉庫的概念，讓我們使用書店的例子作為案例研究。擁有多家分店的書店的管理將喜歡建立一個數據倉庫來分析他們的圖書銷售情況。他們已經存儲了所有圖書銷售交易在操作數據庫中。管理層將特別喜歡從各個角度分析他們的圖書銷售業績，如從按月、從賬簿和從分行開始。第一個任務這樣做就是定義一個星型模式。

該書店案例研究的簡單星型模式如圖 1 所示。6.1.它有 三個維度：時間維度、分支維度和書籍維度，事實有一個事實 測量，稱為Number 出售的書籍。有了這個星型模式，管理是 有能力的根據分店檢索按月售出的圖書數量 並根據書。

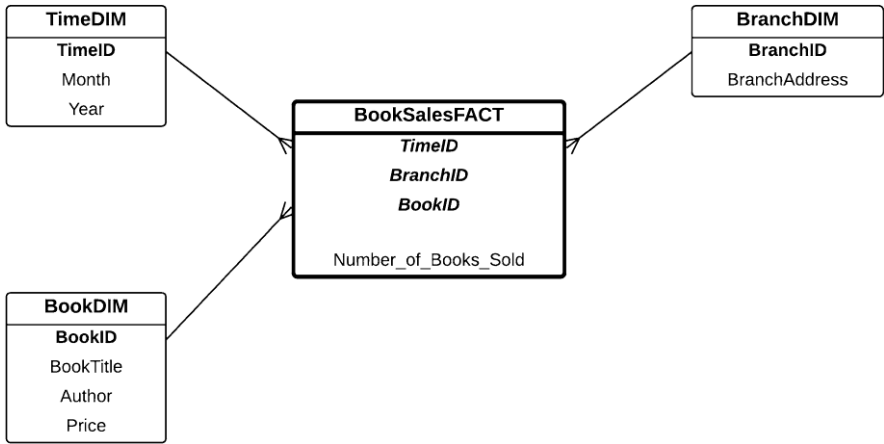


圖 6.1 書店星型模式

事實表可能如下所示（見表6.1）。注意銷售數字是僅用於說明目的，並在以下條目表格不一定全面。

事實表可以與維度結合以產生更全面的西我報告。在下面的示例中，事實表與Book Dimension 連接表，因此Book Dimension 表中的屬性將出現在報告。假設Book Dimension 表有以下數據（見表6.2）。

新報告如表所示6.3.注意三個附加屬性：

| Book Dimension | Book |
|----------------|----------|
| BookID | 1 |
| BookTitle | Book 1 |
| Author | Author 1 |
| Price | Price 1 |

噸圖書維度表中的標題、作者和價格。注意價格柱子在報表中。在這種情況下，價格列包含當前價格每本書，和很可能是原價。然而，時不時，一本書以折扣價出售。上述銷售報告並未反映這一點；因此，分析這份報告可能會產生誤導。

將時間融入 Price 的值，我們需要添加一個Bridge Table 到 Book Dimension 以存儲圖書價格的歷史記錄（即 Book Price 方面）。請注意，實現了賬面價格維度（橋接表）作為弱實體，具有 BookID、開始日期和結束日期的複合主鍵（參考到圖。6.2）。

| Book Dimension | Book | Price |
|----------------|------|---------|
| BookID | 1 | Price 1 |

圖書維度和圖書價格維度表如下（參考表6.4和6.5）。在 Book Price Dimension 表中，時間屬性 Price 管理，其中對於每個價格，期間（開始和結束日期）和備註被記錄下來。請注意，我們使用 Dec9999 表示結束日期仍處於打開狀態到現在。

表6.1 圖書銷售事實

timeID BranchID BookID 售出書籍數量
噸有能力的

| | | | |
|-------------|-------------------|--|--|
| 2008 年 3 月 | 城市 C1 5 | | |
| 2008 年 3 月 | 城市 H6 15 | | |
| 2008年3月 | 城市DV 23 | | |
| 2008年3月 | 市 | | |
| 2008 年 3 月 | 查斯頓 C1 15 | | |
| 2008 年 3 月 | 查斯頓 H6 3 | | |
| 2008 年 3 月 | Chadstone DV 2 | | |
| 2008年3月 | Chadstone | | |
| 2008 年 3 月 | 坎伯韋爾 C1 1 | | |
| 2008 年 3 月 | 坎伯韋爾 H6 1 | | |
| 2008 年 3 月 | 坎伯韋爾 DV 2 | | |
| 2008 年 3 月 | Camberwel | | |
| 2008 年 3 月 | | | |
| | ... | | |
| | ... | | |
| 2007 年 12 月 | 城市 C1 15 | | |
| 2007 年 12 月 | 城市 H6 6 | | |
| 2007 年 12 月 | 城市 DV 6 | | |
| 2007年12月 | 市 | | |
| 2007 年 12 月 | 查斯頓 C1 10 | | |
| 2007 年 12 月 | 查斯頓 H6 3 | | |
| 2007 年 12 月 | 查斯頓 DV 1 | | |
| 2007 年 12 月 | Chadstone | | |
| 2007 年 12 月 | 坎伯韋爾 C1 18 | | |
| 2007 年 12 月 | 坎伯韋爾 H6 3 | | |
| 2007 年 12 月 | 坎伯韋爾 DV 2 | | |
| 2007 年 12 月 | Camberwell | | |
| 2007 年 12 月 | | | |
| | ... | | |

表6.2 書籍尺寸

BookID 書名 作者 價格
桌子

| | | |
|----------------------|---------|---------------------|
| C1 CSIRO 飲食 CSIRO 團隊 | \$45.95 | H6 哈利波特 |
| 6 羅琳 | \$30.95 | DV 達芬奇密碼丹布朗 \$27.95 |
| | | |

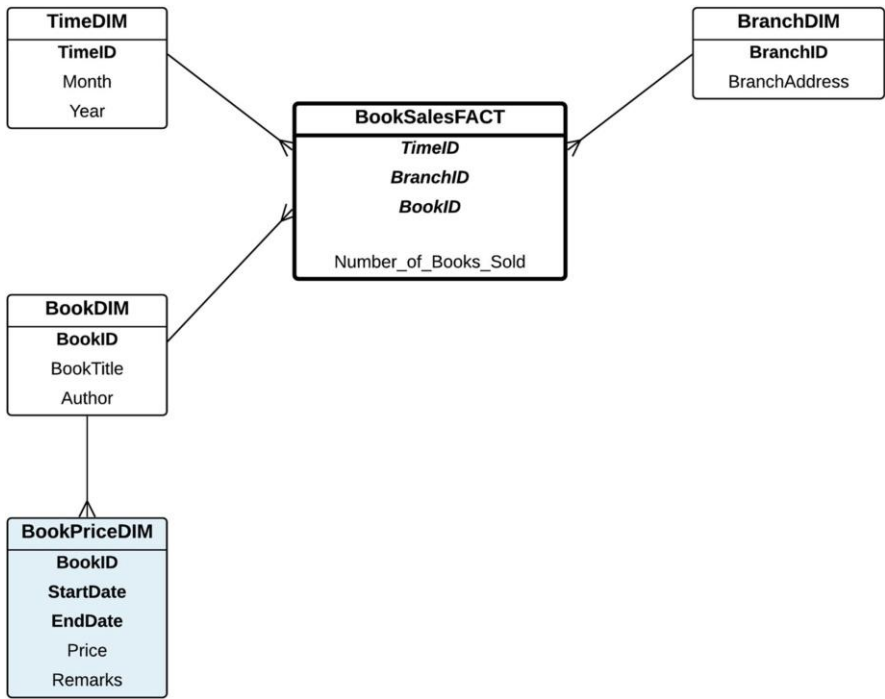


圖 6.2 時態I 數據倉庫星型模式

表6.4 書籍尺寸
BookID 書名 作者
桌子

| | | |
|----|----------|-----------|
| | | |
| C1 | CSIRO 飲食 | CSIRO 團隊 |
| H6 | 哈利波特6 | 羅琳 |
| DV | 達芬奇密碼 | 丹布朗 |

表6.5 噓k 價格維度表

| | | | | | | |
|--------|---------|----------|---------|--------|------------|--------------------|
| BookID | 開始日期 | 結束日期 | 價格 | 備註 | | |
| C1 | Jan2007 | July2007 | \$45.95 | 全價 | | |
| C1 | Aug2007 | Oct2007 | \$36.75 | 20% 折扣 | | |
| C1 | Nov2007 | Jan2008 | \$23.00 | 半價 | | |
| C1 | Feb2008 | Dec9999 | \$45.95 | 全價 | | |
| H6 | Jan2007 | Mar2007 | \$21.95 | | | |
| H6 | Apr2007 | Jan2008 | \$30.95 | 全價 | | |
| H6 | Feb2008 | Dec9999 | \$10.00 | 產品銷售結束 | DV Jan2007 | Dec9999 \$27.95 全價 |
| | | | | | | |

表6.6 報告t 2 具有正確的圖書價格

| 噸 | 噸ID | BranchID | BookID | 書名 | 作者 | 已售圖書價格 | 2008 年 3 月 | City C1 | CSIRO | 飲食 | CSIRO | 數的 |
|-------------|------------|----------|----------|-----------|----------|--------|------------|---------|-------|----|-------|----|
| 團隊 | \$45.95 | 5 | | | | | | | | | | |
| 2008 年 3 月 | 城市 | H6 | 哈利波特 6 | 羅琳 | \$10.00 | 15 | | | | | | |
| 2008 年 3 月 | 城市 | DV | 達芬奇密碼 | 丹布朗 | \$27.95 | 23 | | | | | | |
| 2008年3月 | 市 | | | | | | | | | | | |
| 2008 年 3 月 | Chadstone | C1 | CSIRO 飲食 | CSIRO 團隊 | \$45.95 | 15 | | | | | | |
| 2008 年 3 月 | Chadstone | H6 | 哈利波特 6 | 羅琳 | 10.00 美元 | 3 | | | | | | |
| 2008 年 3 月 | Chadstone | DV | 達芬奇密碼 | Dan Brown | \$27.95 | 2 | | | | | | |
| 2008年3月 | 查斯頓 | | | | | | | | | | | |
| 2008 年 3 月 | Camberwell | C1 | CSIRO 飲食 | CSIRO 團隊 | \$45.95 | 1 | | | | | | |
| 2008 年 3 月 | Camberwell | H6 | 哈利波特 6 | 羅琳 | \$10.00 | 1 | | | | | | |
| 2008 年 3 月 | Camberwell | DV | 達芬奇密碼 | Dan Brown | \$27.95 | 2 | | | | | | |
| 2008 年 3 月 | Camberwell | ... | ... | ... | ... | ... | | | | | | |
| 2008 年 3 月 | ... | ... | ... | ... | ... | ... | | | | | | |
| | | | | | | | | | | | | |
| | | | | | | | | | | | | |
| 2007 年 12 月 | City C1 | CSIRO 飲食 | CSIRO 團隊 | \$23.00 | 15 | | | | | | | |
| 2007 年 12 月 | 城市 | H6 | 哈利波特 6 | 羅琳 | 30.95 美元 | 6 | | | | | | |
| 2007 年 12 月 | 城市 | DV | 達芬奇密碼 | 丹布朗 | \$27.95 | 6 | | | | | | |
| 2007 年 12 月 | 市 | ... | ... | ... | ... | ... | | | | | | |
| 2007 年 12 月 | Chadstone | C1 | CSIRO 飲食 | CSIRO 團隊 | \$23.00 | 10 | | | | | | |
| 2007 年 12 月 | Chadstone | H6 | 哈利波特 6 | 羅琳 | 30.95 美元 | 8 | | | | | | |
| 2007 年 12 月 | Chadstone | DV | 達芬奇密碼 | 丹布朗 | 27.95 美元 | 1 | | | | | | |
| 2007 年 12 月 | Chadstone | ... | ... | ... | ... | ... | | | | | | |
| 2007 年 12 月 | Camberwell | C1 | CSIRO 飲食 | CSIRO 團隊 | \$23.00 | 18 | | | | | | |
| 2007 年 12 月 | Camberwell | H6 | 哈利波特 6 | 羅琳 | 30.95 美元 | 3 | | | | | | |
| 2007 年 12 月 | Camberwell | DV | 達芬奇密碼 | Dan Brown | \$27.95 | 2 | | | | | | |
| 2007 年 12 月 | Camberwell | ... | ... | ... | ... | ... | | | | | | |
| 2007 年 12 月 | ... | ... | ... | ... | ... | ... | | | | | | |
| | | | | | | | | | | | | |

因為每本書都有時間特徵（即時間屬性），前一個報告 1
可以修訂以包含該書的正確銷售價格。因此，新報告（報告
2）如表所示6.6.請注意，現在我們可以理解為什麼 更多副本f CSIRO Diet 於
2007 年 12 月以半價出售。還 請注意t Harry Potter
的銷售即將結束（例如需求量不大）。這 達芬奇密碼一直以全價出售。

6.2 時態數據倉庫的實現

本節將討論時態數據的 SQL 實現細節 倉儲。考慮一個操作數據庫的E/R圖如圖 6.3.

星型模式先前顯示在圖 3 中。6.2.創建的 SQL 命令 維度表如下如下：

創建表 BranchDim 為
從分支中選擇 *；

創建表 BookDim 為
從書中選擇*；

創建表 TimeDim 為
選擇不同的
to_char(TransactionDate, 'MonYYYY') 作為 TimeID,
to_char(TransactionDate, 'Mon') 作為月份,
to_char(TransactionDate, 'YYYY') 作為年份
來自 BookTransaction；

創建表 BookPriceDim 為
從 BookPriceHistory 中選擇 *；

創建事實表的 SQL 如下：

創建表 BookSalesFact1 為
選擇
to_char(T.TransactionDate, 'MonYYYY') 作為 TimeID,
BK.BookID,
BR.BranchID,
sum(T.Quantity) as Number_of_Books_Sold
來自 BookTransaction T, Book BK, Branch BR
其中 T.BranchID = BR.BranchID
和 T.BookID = BK.BookID
通過...分組
to_char(T.TransactionDate, 'MonYYYY'),
BK.BookID,
BR.BranchID;

計算數字 of Books Sold 很簡單，即
這操作數據庫中賬簿事務表中的屬性數量。

認為我們希望有一個額外的事實度量，稱為總銷售額。

因此，新的星型模式如圖所示。6.4.

創建第二個 Fact Table 的 SQL 命令（我們稱之為 BookSalesFact2）如下：

創建表 BookSalesFact2 為
從 BookSalesFact1 中選擇 *；

更改表 BookSalesFact2
添加 (Total_Sales 數量) ；

```

宣布
 光標 PriceCursor 是
 選擇 *
 來自 BookPriceDim ;
開始
 對於 PriceCursor 循環中的項目
    -- 更新 BookSalesFact2 中 Total_Sales 的值
    更新 BookSalesFact2
    設置 Total_Sales = Number_Of_Books_Sold * Item.Price
    其中 BookID = Item.BookID
    和 to_date(TimeID, 'MonYYYY') >=
      to_date(Item.StartDate, 'MonYYYY')
    和 to_date(TimeID, 'MonYYYY') <=
      to_date(Item.EndDate, 'MonYYYY');
 結束循環;
結尾;
/
```

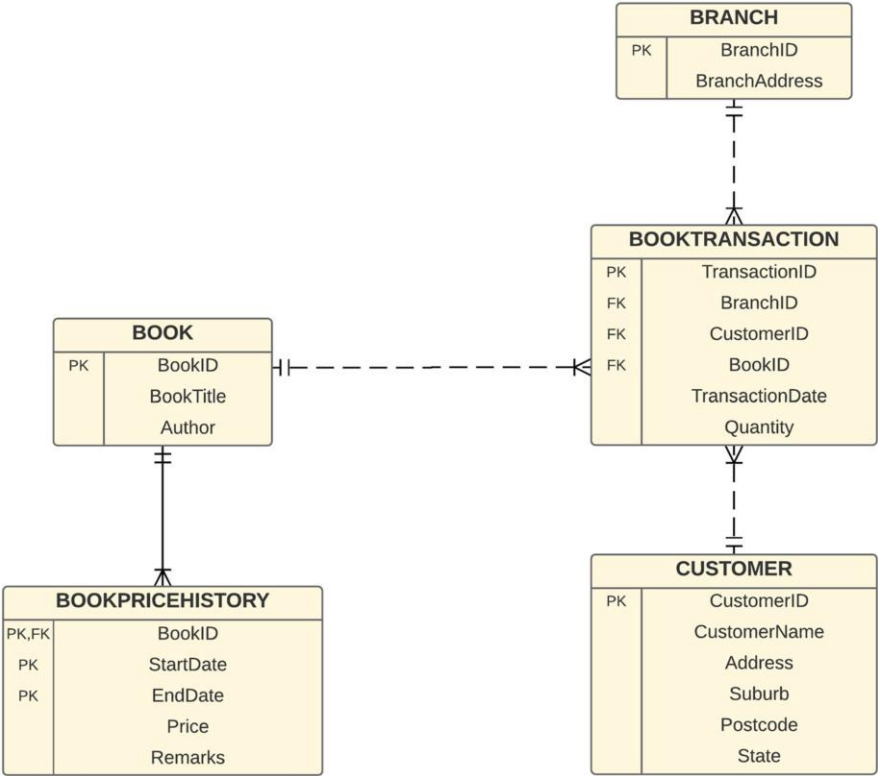


圖 6.3 操作數據庫的 E/R 圖

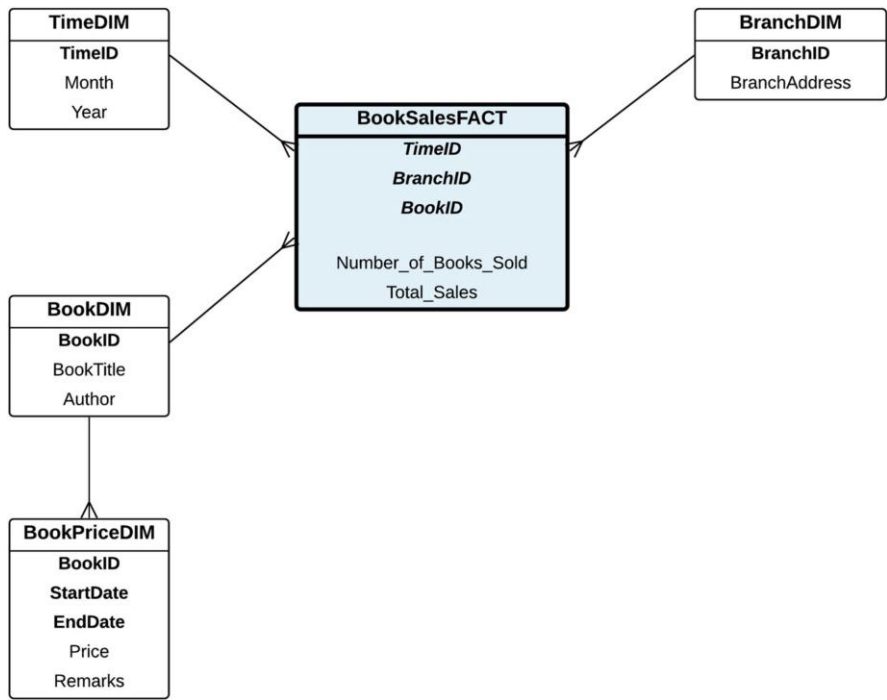


圖 6.4 站r 具有新事實屬性的模式：總銷售額

對於每條記錄在 `BookPriceDim` 表中，更新 `BookSalesFact2` 表併計算總銷售額乘以書中售出的書籍數量- `SalesFact2` 和 `BookPriceDim` 的“正確”價格（其中 `TimeID` 在事實上應該在書價的開始日期和結束日期之間）。上述方法假定 `BookSalesFact1` 已創建。圖書銷售事實2通過從 `BookSalesFact1` 複製創建，然後更新它 通過添加名為 `Total Sales` 的新列。但是，如果我們不假設 `BookSalesFact1` 已經創建，我們可以從創建 `BookSalesFact2` 從頭開始，使用以下創建表語句：

```
創建表 BookSalesFact2 為
選擇
    to_char(T.TransactionDate, 'MonYYYY') 作為 TimeID,
    BK.BookID,
    BR.BranchID,
    sum(T.Quantity) 作為 Number_Of_Books_Sold,
    sum(T.Quantity * BP.Price) 作為 Total_Sales
從
    BookTransaction T,
    預訂 BK,
    分支機構 BR,
    書價歷史BP
```

```

其中 T.BranchID = BR.BranchID
和 T.BookID = BK.BookID
和 BK.BookID = BP.BookID
和 T.TransactionDate >= to_date(BP.StartDate, 'MonYYYY') 和
T.TransactionDate <= to_date(BP.EndDate, 'MonYYYY')
通過...分組
    to_char(T.TransactionDate, 'MonYYYY'),
    BK.BookID,
    BR.BranchID;

```

6.3 時間屬性和時間維度

6.3.1 時間屬性

在上述案例研究中，Book Price Dimension 表中的 Book Price 稱為**顯屬性**。時間屬性是一個屬性，其中它的值屬性有生命週期。在這個例子中，每本書的價格都有一個生命週期，並且生命週期由 Book Price 中的 Start Date 和 End Date 屬性確定。尺寸表。例如，圖書 C1 的圖書價格 45.95 美元僅有效 2007 年 1 月至 2007 年 7 月。

Book Price Dimension 表是一個關係 DBMS (RDBMS) 實現時態數據倉庫化，它被實現為橋接表。書價維度與原書分離的原因方面是因為一本書在不同時期可能有很多價格。因為關係模型不允許嵌套表，有關信息書價有被分開到另一個表中。圖書價格維度表是笑wn 在表中6.5。

如上一節所述，當我們生成加入 Fact 的報告時，桌子使用維度表（例如，將 BookSalesFact 表與 Book 維度和書價維度表），我們可以生成報告 2（參考桌子6.6）其中包含 BookSalesFact 表中的所有屬性。增加書 TBook Dimension 表中的 title 和 Author 屬性以及價格Book Price Dimension 表中的屬性。本報告由這以下 SQL 語句：

```

選擇
    F.TimeID,
    F.BranchID,
    F.BookID,
    B.書名,
    B.作者,
    P.價格,
    F.Number_of_Books_Sold
來自 BookSalesFact F, BookDim B, BookPriceDim P
其中 F.BookID = B.BookID
和 B.BookID = P.BookID

```

```
和 to_date(F.TimeID, 'MonYYYY') >=
  to_date(P.StartDate, 'MonYYYY')
和 to_date(F.TimeID, 'MonYYYY') <=
  to_date(P.EndDate, 'MonYYYY');
```

此 SQL 命令連接三個提到的表（即 Book Sales Fact、Book 尺寸和書價尺寸）和支票如果 TimeID（即月和 年）BookSalesFact 表中的賬面價格屬於 開始日期和結束日期在圖書價格維度表中。結果，正確的價格將顯示在與月份（即 TimeID）匹配的報告中事實記錄。

但是，如果本書使用的時間粒度會出現問題 價格和事實不同。例如，如果 BookID C1 的價格是 23.00 美元，但 不是從 2007 年 11 月到 2008 年 1 月，但從 2007 年 11 月到 1 月 15 日 2008 年（結果y 價格從 2008 年 1 月 16 日起漲至 45.95 美元 而不是f 2008 年 2 月；參考表6.7），那麼在哪裡條款條件在上面的 SQL 變為：

```
和 to_date(F.TimeID, 'MonYYYY') >=
  to_date(P.StartDate, 'MonYYYY')
和 to_date(F.TimeID, 'MonYYYY') <=
  to_date(P.EndDate, 'MonYYYY');
```

這個在哪裡條款然而，條件會產生一個不正確的報告，因為 噸2007 年 1 月的 imeID 200701 將與 Book Price 中的兩條記錄匹配 尺寸表，一張一個價格為 23.00 美元，另一個價格為 45.95 美元。最後，在報告中，2008 年 1 月的賬簿 C1 將有兩條記錄， 在每個分店中，一個價格為 23.00 美元，另一個價格為 45.95 美元，僅僅因為 2008 年 1 月有兩個圖書價格（見報告 3在表中6.8）。

這個問題只有在事實中 TimeID 的粒度是 與開始日期和結束日期相同在圖書價格維度表中。如果時間 ID 是在月級別，那麼賬面價格的周期必須完全基於 月。另一方面，如果賬面價格的周期是一個日期粒度

表6.7 BookPriceDim 表

| BookID | 開始日期 | 結束日期 | 價格 | 備註 | | |
|--------|-----------|-----------|---------|--------|------------|--------------------|
| C1 | Jan2007 | Jul2007 | \$45.95 | 全價 | | |
| C1 | Aug2007 | Oct2007 | \$36.75 | 20% 折扣 | | |
| C1 | Nov2007 | 15Jan2008 | \$23.00 | 半價 | | |
| C1 | 16Jan2008 | Dec9999 | \$45.95 | 全價 | | |
| H6 | Jan2007 | Mar2007 | \$21.95 | | | |
| H6 | Apr2007 | Jan2008 | \$30.95 | 全價 | | |
| H6 | Feb2008 | Dec9999 | \$10.00 | 產品銷售結束 | DV Jan2007 | Dec9999 \$27.95 全價 |
| | ... | | | | | |

表6.8 報告t3：錯誤的報告

| TimeID | BranchID | BookID | 書名 | 作者 | 已售圖書價格 | Jan2008 City C1 | CSIRO | 飲食 | CSIRO |
|------------|---------------|--------|-----------|---------|------------|-----------------|-------|----|-------|
| Jan2008 | City C1 | CSIRO | 飲食 | CSIRO | 團隊 \$45.95 | 25 | | | |
| 2008 年 1 月 | City H6 | 哈利波特 6 | 羅琳 | \$30.95 | 10 | | | | |
| 2008 年 1 月 | 城市 DV | 達芬奇密碼 | 丹布朗 | \$27.95 | 7 | | | | |
| 2008年1月 | 市..... | | | | | | | | |
| Jan2008 | Chadstone C1 | CSIRO | 飲食 | CSIRO | 團隊 \$23.00 | 30 | | | |
| Jan2008 | Chadstone C1 | CSIRO | 飲食 | CSIRO | 團隊 \$45.05 | 30 | | | |
| 2008 年 1 月 | Chadstone H6 | 哈利波特 6 | 羅琳 | 30.95 | 美元 15 | | | | |
| 2008 年 1 月 | Chadstone DV | 達芬奇密碼 | 丹布朗 | 27.95 | 美元 5 | | | | |
| 2008 年 1 月 | Chadstone .. | | | | | | | | |
| Jan2008 | Camberwell C1 | CSIRO | 飲食 | CSIRO | 團隊 \$23.00 | 20 | | | |
| Jan2008 | Camberwell C1 | CSIRO | 飲食 | CSIRO | 團隊 \$45.05 | 20 | | | |
| 2008 年 1 月 | Camberwell H6 | 哈利波特 6 | 羅琳 | \$30.95 | 5 | | | | |
| 2008 年 1 月 | Camberwell DV | 達芬奇密碼 | Dan Brown | \$27.95 | 5 | | | | |
| 2008 年 1 月 | Camberwell .. | | | | | | | | |
| 2008 年 1 月 | | | | | | | | | |
| | | | | | | | | | |

級別，那麼事實中的 TimeID 也必須處於日期粒度級別。如果水平粒度不同，就會出現上述問題。另一種方式“解決”這個問題的方法是不允許報表與賬面價格結合 尺寸表。換句話說，Book Price Dimension 表是純粹處理的 作為補充 數據倉庫中的信息，以備管理層需要深入了解某些書籍的信息。

上述Book Price之間TimeID粒度不匹配的問題 維度表和 Fac如果我們同時顯示“兩個”價格，則可以解決記錄在報告中（見表6.9）。例如，在表的第一條記錄中6.9,它有兩本書的價格，因為那個月（例如 2008 年 1 月）發生了變化 書中價格從 23.00 美元到 45.95 美元不等。

報告 4 可以由以下 SQL 命令：

```
選擇
    F.TimeID,
    F.BranchID,
    F.BookID,
    B.書名,
    B.作者,
    listagg(P.Price, ';' 組內 (按 P.Price 排序)
        作為價格,
    F.Number_of_Books_Sold
來自 BookSalesFact F, BookDim B, BookPriceDim P
其中 F.BookID = B.BookID
和 B.BookID = P.BookID
和 to_date(F.TimeID, 'MonYYYY') >=
    to_date(P.StartDate, 'MonYYYY')
```

```
和 to_date(F.TimeID, 'MonYYYY') <=
to_date(P.EndDate, 'MonYYYY')
```

通過...分組

```
F.TimeID,
F.BranchID,
F.BookID,
B.書名,
B.作者,
F.Number_of_Books_Sold;
```

報告 3 中出現上述問題是因為我們將事實表與書價維度表；那裡事實表中的記錄沒有問題 本身。有人可能會問，事實表中的 **Total Sales** 屬性發生了什麼。 **Total Sales** 的計算仍然正確，如以下 SQL 所示陳述：

創建表 BookSalesFact2 為

選擇

```
to_char(T.TransactionDate, 'MonYYYY') 作為 TimeID,
BK.BookID,
BR.BranchID,
sum(T.Quantity) 作為 Number_of_Books_Sold,
sum(T.Quantity * BP.Price) 作為 Total_Sales
```

從

```
BookTransaction T,
預訂 BK,
分支機構 BR,
書價歷史BP
```

其中 T.BranchID = BR.BranchID

和 T.BookID = BK.BookID

和 BK.BookID = BP.BookID

和 T.TransactionDate >= BP.StartDate

和 T.TransactionDate <= BP.EndDate

通過...分組

```
to_char(T.TransactionDate, 'MonYYYY'),
BK.BookID,
BR.BranchID;
```

這條SQL連接了操作數據庫中的四張表，即Book交易，圖書、分行和圖書價格歷史記錄。在加入這些表時，它 獲得正確的t通過比較交易日期和開始日期來確定價格，以及圖書價格的結束日期。在這種情況下，我們假設開始日期和結束日期，如以及交易日期，具有日期粒度。一旦這些表被連接起來，分組然後基於月份。單本圖書總銷量正確，且分組然後基於月份。因此，錯誤報告的問題不是到期的不正確的事實，但由於事實和維度之間的連接 存儲時間的屬性。

桌子6.9 報告t 4：一個月的多個圖書價格

| timeID | BranchID | BookID | 書名 | 作者 | 價格書 | 已售出 | 2008 年 1 月 | City C1 CSIRO | 飲食 CSIRO | 數字的 |
|---|-----------------|----------------|-----------|----------|-----|----------|------------|---------------|----------|-----|
| 團隊 | \$23.00; | \$45.95 | 25 | | | | | | | |
| 2008 年 1 月 City H6 哈利波特 6 羅琳 \$30.95 10 | | | | | | | | | | |
| 2008 年 1 | 月城市 DV | 達芬奇密碼 | 丹布朗 | \$27.95 | 7 | | | | | |
| 2008年1月 | 市 | | | | | | | | | |
| 2008 年 1 | 月 Chadstone C1 | CSIRO 飲食 CSIRO | 團隊 | 23.00 美元 | ; | 45.95 美元 | 30 | | | |
| 2008 年 1 | 月 Chadstone H6 | 哈利波特 6 羅琳 | 30.95 美元 | 15 | | | | | | |
| 2008 年 1 | 月 Chadstone DV | 達芬奇密碼 | 丹布朗 | 27.95 美元 | 5 | | | | | |
| 2008 年 1 | 月 Chadstone | | ... | | | | | | | |
| 2008 年 1 | 月 Camberwell C1 | CSIRO 飲食 CSIRO | 團隊 | 23.00 美元 | ; | 45.95 美元 | 20 | | | |
| 2008 年 1 | 月 Camberwell H6 | 哈利波特 6 羅琳 | \$30.95 | 5 | | | | | | |
| 2008 年 1 | 月 Camberwell DV | 達芬奇密碼 | Dan Brown | \$27.95 | 5 | | | | | |
| 2008 年 1 | 月 Camberwell | | ... | | | | | | | |
| 2008 年 1 | 月 | ... | ... | | | | | | | |
| | | | | | | | | | | |

6.3.2 時間維度

如果時間屬性是值具有生命週期的屬性，則時間屬性方面是維度的記錄有生命週期的維度。為了例子，如果BookID C1在2007年出現，在2008年消失，然後重新出現再次2009年，那麼書本維度需要一個時間維度。在另一個情況，如果一個分支打開和關閉多次，那麼分支維度需要時間維度。然而，這兩個例子並不現實，因為書的出現和消失，樹枝開合的情況是非常罕見的。所以，在命令為了突出時間維度，我們將使用一個更現實的案例 研究日曆店。

日曆商店有許多分店。它出售不同類型的日曆和日記。有些分支是季節性的；他們開了一個攤位在各大商場，這些攤位是僅在特定月份開放（例如，一些攤位開放 從 10 月到 2 月，而其他可能從 11 月到 1 月開放）。這個 公司還有少量常設店鋪全年開放。

這家日曆商店公司想要分析他們的銷售額，類似於書店案例研究。日曆商店的星型模式具有三個維度：時間、分店和商品維度（而不是書籍維度在裡面書店案例研究）。對於商品維度，如果我們想跟踪的變化在價格方面，就像我們對 Book Price Dimension 所做的那樣，我們可以商品價格作為商品價格維度中的時間屬性。

對於分支維度，因為有些分支有有一定的壽命，分支維度需要臨時維度稱為分支歷史維度- 錫安。在分支歷史維度中，整個分支記錄都有一個時間

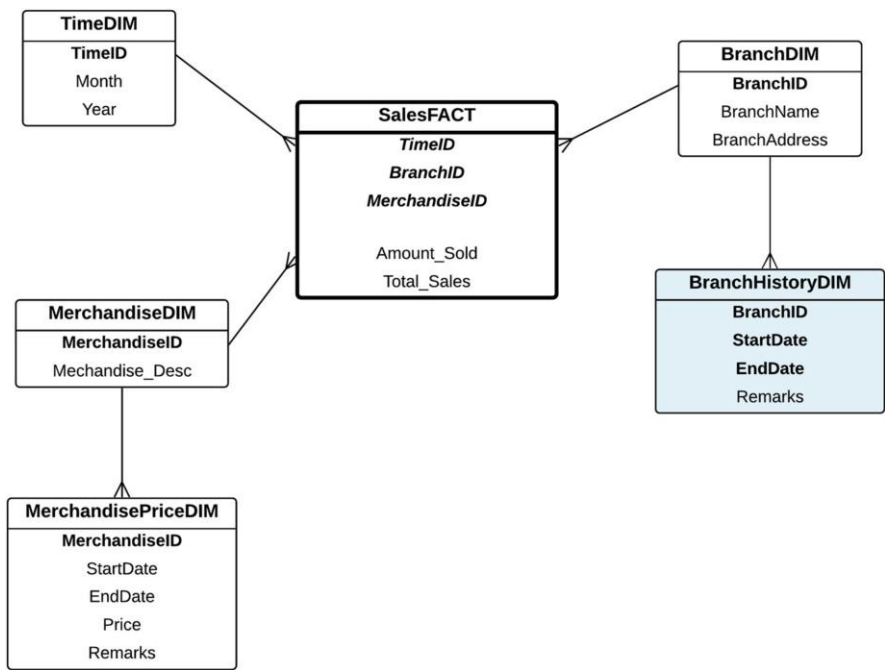


圖 6.5 站日曆商店案例研究的 r 架構

表6.10 分支
BranchID 分行名稱 分行地址
方面桌子

| | |
|------------------------------------|--|
| MEL 墨爾本中心 88 Lonsdale St | |
| CH Chadstone 購物中心 109 Dandenong Rd | |
| 陶氏唐卡斯特韋斯特菲爾德 75 Doncaster Rd | |
| | |

元素，由開始日期和結束日期指示。日曆的星型模式 店鋪如圖所示。6.5.

分支維度和分支歷史維度的樣本記錄是 笑wn 在表中6.10和6.11.

在分支歷史維度表中，開始日期和結束日期指的是 BranchID 例如, Chadstone 或整個記錄，而不是特定的屬性。這意味著，對於 分公司 (BranchIDCH) 在過去的幾次中打開了幾次 年，如開始日期和結束日期所示，而墨爾本中心 店鋪 (BranchIDMEL) 全年開放。備註屬性僅用於 評論對於分支歷史維度中每個出現的分支條目 桌子。Contact Number 屬性顯示特定分支的聯繫號碼 在特定的壽命。在此示例中，它顯示了 Chadstone 分支 (CH) 有一個每年不同的聯繫電話號碼（一個手機號碼），而

表6.11 BranchHistoryDim 表

| BookID | 開始日期 | 結束日期 | 備註 | 聯繫電話 | | |
|--------------|------------------------|------------------------|----------------|-------|------------------------|--------------|
| MEL Jan0000 | Dec9999 | 總店 | (03) 9859 8070 | CH | 2007 年 10 月 2008 年 3 月 | 0411 848 821 |
| CH Oct2008 | Feb2009 | 正在重建 | 中 0413 356 66 | CH | 2009 年 10 月 2010 年 2 月 | 0412 313 |
| 313 陶氏化學 | 2007 年 11 月 2008 年 2 月 | 0427 | 123 456 | 陶氏 | 2008 年 11 月 2009 年 2 月 | |
| 0427 123 456 | 陶氏化學 | 2009 年 10 月 2010 年 2 月 | 0427 123 456 | | | |

唐卡斯特分公司（DOW）從那以後每年都有相同的聯繫手機號碼 它開了。

當銷售事實加入了分支維度和分支歷史維度-

錫安，該報告可能看起來如表中所示6.12.

為了簡單，商人-

疾病ID留空。該報告於 2007 年 10 月開始，當時只有兩家商店開業（Melbourne Central 和 Chadstone Mall），然後在 2007 年 11 月，商店開業（唐卡斯特韋斯特菲爾德）。然後2008 年 4 月，僅墨爾本中心被打開，作為查斯頓購物中心於 2008 年 10 月重新開業。

如果我們比較這個日曆商店中的時間屬性和時間維度

案例研究，兩者看起來非常相似。唯一的區別是，在

時態

屬性，在歷史維表中有一個時間屬性（例如

書書價維度表中的價格），而在時間維度中

案例，有是沒有這樣的屬性。時間維度中的屬性只有

這父維度的關鍵標識符和可能的任何備註屬性。或者

換句話說，我們可以說時間維度是 維度是所有，統稱為時間屬性。

6.3.3 另一個時間維度

天寶

維度或具有時間屬性的維度通常是實現的

mented作為橋接表，如前幾節所述。在本節中，我們

要去查看未將時間維度實現為 橋表，而不是作為一個正常的維度。

假設那裡是一個簡單的星型模式，用於維護課程/學位列表

（例如計算機科學學士、碩士信息技術等）。一個

課程或學位結構發生變化時不時地，只是由於自然

學科，尤其是那些正在以相當高的速度移動，

比如信息技術。每門課程都有一個課程代碼。當課程

變化重要的是，大學或學院通常會引入一門新課程

（或學位）並逐步淘汰舊課程（或學位），而不是“更新”

表6.12 報告：SalesFact 加入了分支維度和分支歷史維度

| Time Branch Branch Start End Contact 金額 合計 | | | | | | | | | |
|---|--|--|--|--|--|--------------------------------|--|--|----------|
| ID ID 姓名 地址 日期 備註編號 商品銷售 銷售額 | | | | | | | | | |
| Oct2007 MEL Melbourne 88 Lonsdale Jan0000 Dec9999 Main | | | | | | | | | |
| Oct2007 CH Chadstone 109 Dandenong Oct2007 Mar2008 | | | | | | | | | |
| Nov2007 MEL Melbourne 88 Lonsdale Jan0000 Dec9999 Main | | | | | | | | | |
| Nov2007 CH Chadstone 109 Dandenong Oct2007 Mar2008 | | | | | | | | | |
| | | | | | 2007 年 11 月 DOW/唐卡斯特 75 | 唐卡斯特 Nov2007 2008 年 2 月 | | | 南城路..... |
| | | | | | Dec2007 MEL 墨爾本 88 Lonsdale Jan0000 Dec9999 Main | W埃斯特菲爾德路..... | | | |
| | | | | | 2007 年 12 月 CH Chadstone 109 丹德農 2007 年 10 月 2008 年 3 月 | 中央大街小店..... | | | |
| | | | | | 2007 年 12 月 DOW/唐卡斯特 75 唐卡斯特 Nov2007 2008 年 2 月 | 南城路..... | | | |
| | | | | | Jan2008 MEL Melbourne 88 Lonsdale Jan0000 Dec9999 Main | Westfield 路..... | | | |
| | | | | | Jan2008 CH Chadstone 109 Dandenong Oct2007 Mar2008 | 中央大街小店..... | | | |
| | | | | | Jan2008 DOW Doncaster 75 Doncaster Nov2007 Feb2008 | 南城路..... | | | |
| | | | | | | Westfield 路..... | | | |

(繼續)

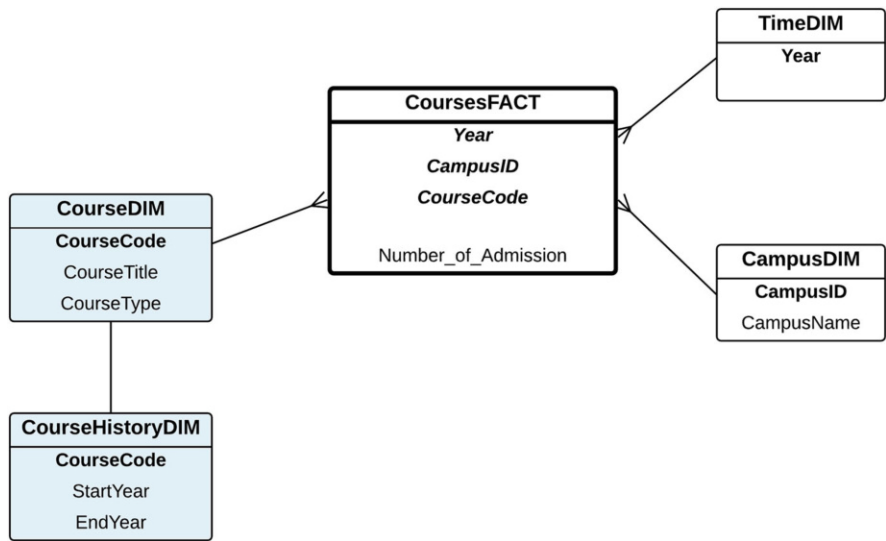


圖 6.6 課程歷史作為時間維度：使用橋接表

e現有學位。發生這種情況時，進入該課程/學位的學生將錄取新課程（而不是舊課程），而現有學生將仍然使用舊的學位結構教授，直到沒有更多的學生老課程。新課程將有一個新的課程代碼，即使該課程姓名可能保持不變。因此，課程或學位通常具有課程代碼的數量，但是其中只有一個可以招收新生。

數字6.6笑是一個具有時間維度的簡單星型模式，稱為歷史維度，以課程代碼屬性作為標識符。什麼時候課程介紹存儲在StartYear屬性中。當課程沒有不再招收新學生（換句話說，該課程正在被逐步淘汰或關閉），記錄結束年份。沒有結束年份的課程意味著該課程是一門活躍的當前課程。星型圖如圖。6.6能夠分析“新”學生（入學）參加課程，每年和每個校園。因為一門課程有只有“一個”的歷史，Course和課程歷史維度是一對一的關係。

這個案例研究與日曆商店案例研究有很大不同。上一節，其中一家分店（或商店）每年營業數月，並且是關閉在今年剩下的時間裡。不過，這家店明年又開張了。這表示商店打開和關閉，因此，一家商店在歷史上有很多實例商店。另一方面，當課程已關閉，它永遠不會重新打開。因此，課程有歷史上僅此一例。如果沒有指定結束年份，課程仍然是積極的。如果有結束年份（通常是過去的一年），則該課程已經關閉，不招收任何新學生。因為關係基數

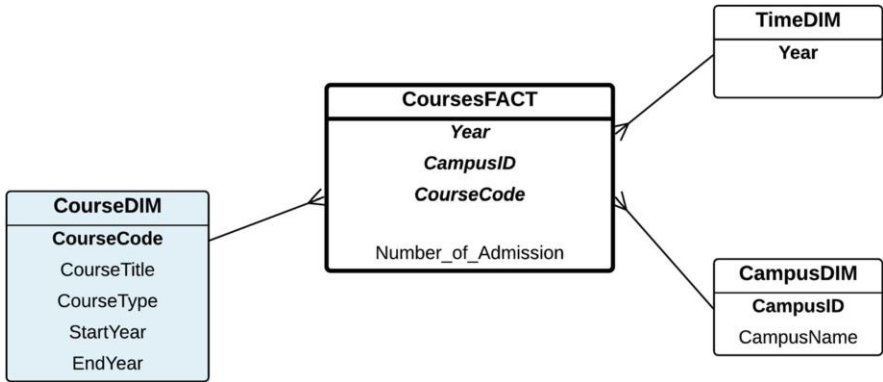


圖 6.7 課程作為時間維度的歷史：不是作為橋樑表

課程與課程之間歷史是一對一的關係，兩個維度都可以結合進入一個稱為課程維度的維度（見圖 6.7）。因此，時間維度，在這種情況下，課程維度是直接與事實相關聯，所以這個時間維度實際上是一個“正常”維度，即使有是此課程維度表中每條記錄的生命週期。

6.4 漸變維度

時態數據倉庫通常稱為緩慢變化的尺寸或緩慢變化的維度是這些維度的記錄的維度在一段時間內。使用本文前面的例子，書籍維度有價格信息，而且價格很常見，一本書的變化隨著時間的推移“慢慢地”，這就是為什麼 Book Dimension 是一個例子。緩慢變化的維度。它被稱為“慢慢地”，因為改變為價格不會發生，例如，每小時甚至每週；而是書價在更長的時間內發生變化，例如數月甚至數年。

漸變維度的主題不同於屬性或記錄，那是“快速”變化，例如移動出租車的位置屬性，變化非常頻繁（例如變化到位置坐標記錄每操作數據庫中的分鐘），或股票市場中的股票價格，每隔幾秒改變一次等。這些快速變化的話題記錄或價值觀經常在實時數據倉庫或流數據領域進行研究倉儲，這不是本章的範圍。

有三種基本類型 SCD 稱為 Type 1、Type 2 和 Type 3。這些類型以不同的方式對待 SCD 的實現。然而，最近，數據倉儲從業人員哈添加了新類型，稱為 0 型、4 型和 6 型，這豐富了 SCD 的實施選項。

6.4.1 SCD 0 型和 1 型

SCD 0 型和 1 型非常相似；他們實際上並沒有記錄歷史 變化在維度。

在類型 0 中，維度存儲記錄的“原始或初始”值當數據倉庫建成。如果維度屬性的值發生變化，沒有記錄更改。記錄保持與記錄時相同 首先插入數據倉庫。

使用 Bookshop 案例研究，這些值是“全價”。對於很多書籍原價或初始價格是全價，但對於某些書籍，價格列出最初可能不是全價。因此，使用 SCD 的 Book Dimension 表 噸type 0 列出了全價，如表中所示6.13. 如果圖書價格發生變化在此之後，新價格不會記錄在數據倉庫中。

對於其他系統，更希望存儲原始值或初始值，而不是 比所謂的本例中的全價。

因為 SCD Type 0 不記錄歷史書價，星 模式母鹿s 沒有時間維度。書店的星型模式 案例分析如圖所示。6.1.

類型 1也不記錄變化的歷史；相當它只記錄 最新值 of 記錄。使用 Book Dimension 示例，圖書價格 圖書維度將是最新價格。這意味著當發生變化時 of price, Book Dimension中的舊價格將被新價格覆蓋 桌子。桌子6.14笑ws 使用 SCD 類型的 Book Dimension 表的內容 1.在這種情況下，BookID H6 的 10.00 美元價格是最新價格。請注意，另一個 此示例中的兩本書的價格與 SCD 類型 0 中的全價相同，但 這並不意味著價格從未改變。這本書尺寸表只告訴我們這些是目前書籍的價格。

與 SCD 類型 0 類似，SCD 類型 1 星型模式不維護時間方面。星型模式如圖所示。6.1.

因為 SCD Type 0 和 Type 1 不維護圖書價格的歷史，如果我們需要生成將 Book Dimension 表與 Book Sales 連接起來的報告

表6.13 噓k 維度

表 (SCD T是的 0)

| BookID | 書名 | 作者 | 價格 | | |
|--------|----------|----------|---------|-----|---------|
| C1 | CSIRO 飲食 | CSIRO 團隊 | \$45.95 | H6 | 哈利波特 |
| 6 | 羅琳 | DV 達芬奇密碼 | \$30.95 | 丹布朗 | \$27.95 |
| | | | | | |

表6.14 噓k 維度

表 (SCDT是的 1)

| BookID | 書名 | 作者 | 價格 | | |
|--------|----------|----------|---------|-----|---------|
| C1 | CSIRO 飲食 | CSIRO 團隊 | \$45.95 | H6 | 哈利波特 |
| 6 | 羅琳 | DV 達芬奇密碼 | \$10.00 | 丹布朗 | \$27.95 |
| | | | | | |

事實表，我們需要注意不要在書價之間建立關聯 來自 Book Dimension 和來自 Book Sales Fact 的 TimeID，因為這本書 報告中的價格不一定是該特定 TimeID 的圖書價格。 在這樣的報告中，書價的列標題可以更改為“原始 書價”（適用於 SCD 類型 0）或“最新書價”（適用於 SCD 類型 1）以避免任何誤解在解釋報告時。

6.4.2 SCD 類型 2

SCD Type 2 跟踪歷史，但不將歷史與主要尺寸；相反，新記錄是不斷地添加到維度中。使用噓以 k 維為例，當一本書的價格發生變化時，創造“另一本書”具有相同的詳細信息，但使用新的 BookID，當然還有新價格。除了開始日期和結束日期，它還有一個當前標誌表明是否有記錄是當前或過去的記錄。任何附加信息，例如備註，也可以包括在內。SCD 類型 2 的書本尺寸表 如表所示6.15。

請注意，同一本書有不同的 BookID 對應不同的圖書價格，並且時期。通常情況下，BookID 屬性被實現為代理鍵，但在此 例子，我們只是在原始 BookID 上添加了一個序列號來區分 同一本書在不同的時間段。由於這些多個 BookID 同一本書，報告說連接 Book Sales Fact 和 Book Dimension 表將顯示如下（見表中的報告 3 SCD 類型 26.16）。

請注意，報告 3 中的第一條記錄 BookID 是 C1_4，因為這是 2008 年 3 月 CSIRO 飲食書的 BookID。因為 SCD 類型 2 僅更改 這原始 Book Dimension 表，星型模式類似於圖 1. 6.1，但是 Book Dimension 表有額外的屬性（見圖6.8）。

表6.15 噓k 尺寸表（SCD 類型 2）

| BookID | 書名 | 作者 | 日期 | 日期 | 價格 | 備註 | 標誌 | 開始 | 結束 | 當前 | | | |
|--------|--------------|-----------|------------|------------|-------------|----------------|----------|--------|----|----|--|--|--|
| C1_1 | CSIRO飲食 | CS | RO | 團隊 | Jan2007 | Jul2007 | \$45.95 | 全價 | N | | | | |
| C1_2 | CSIRO飲食 | CS | RO | 團隊 | Aug2007 | Oct2007 | \$36.75 | 20% 折扣 | N | | | | |
| C1_3 | CSIRO飲食 | CS | RO | 團隊 | Nov2007 | Jan2008 | \$23.00 | 半價 | N | | | | |
| C1_4 | CSIRO飲食 | CS | RO | 團隊 | 2008 年 12 月 | 9999 美元 | 45.95 美元 | 全價 | Y | | | | |
| H6_1 | 哈爾y 波特 6 | 羅琳 | 2007 年 1 月 | 2007 年 3 月 | 21.95 美元 | 推出 | N | | | | | | |
| H6_2 | 哈爾y 波特 6 | 羅琳 | 2007 年 4 月 | 2008 年 1 月 | 30.95 美元 | 全價 | N | | | | | | |
| H6_3 | 哈爾y Potter 6 | Rowling | Feb2008 | Dec9999 | \$10.00 | End of Product | Y | | | | | | |
| DV_1 | 達芬奇密碼 | Dan Brown | Jan2007 | Dec9999 | \$27.95 | 全價 | Y | | | | | | |
| | | | | | | | | | | | | | |

表6.16 報告t 3（SCD 類型 2）

[illegible]

6.4.3 SCD 類型 3

SCD 類型 3 是類型 2 的簡化版。與類型 2 不同，類型 2 維護多個同一本書的記錄，T 類型 3 沒有同一本書的多條記錄。一本書有一個條目在 Book Dimension 表中。對於價格，它只記錄當前價格和之前的價格。換句話說，它不維護整個歷史y 價格變化；相反，它只保留最後兩個價格。書 尺寸表對於 SCD 類型 3 顯示在表 6.17。請注意，對於一本書 沒有ve 任何先前價格，Null 值將記錄在先前價格中 柱子。

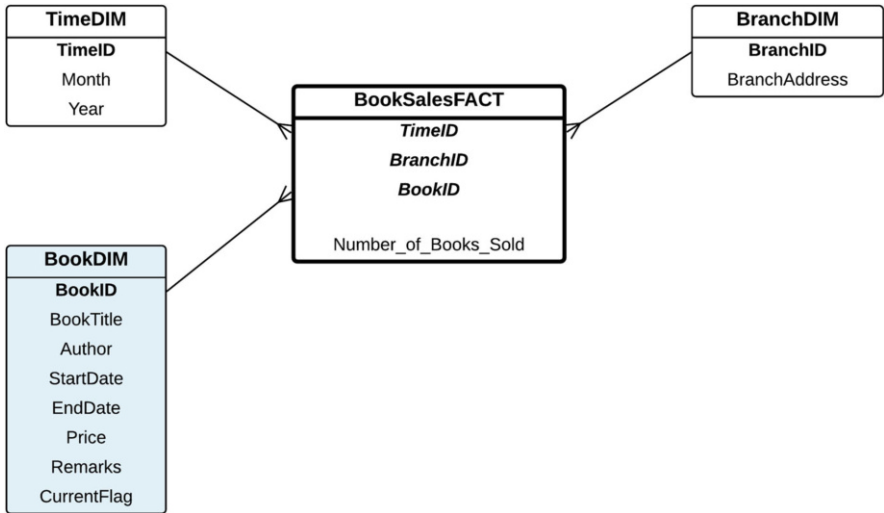


圖 6.8 Book Dimension 具有 SCD 類型 2 的書店星型模式

表6.17 噓k 尺寸表 (SCD 類型 3)

| BookID | 書名 | 作者 | 當前價格 | 先前價格 | | |
|--------|----------|----------|---------|---------|--|--|
| C1 | CSIRO 飲食 | CSIRO 團隊 | \$45.95 | \$23.00 | | |
| H6 | 哈利波特 6 | 羅琳 | \$10.00 | \$30.95 | | |
| DV | 達芬奇密碼 | 丹布朗 | \$27.95 | 無 | | |
| | | | | | | |

採用 SCD 類型 3 的主要理由是大多數分析將應基於當前價格，最多一個過去的價格，這是以前的價格價格。也許，這樣就可以與定價趨勢進行比較。假設分析完整的歷史沒有必要。

此外，雖然我們可以存儲有關價格何時更改的信息，這原始 SCD Type 3 沒有記錄這一點。它只記錄當前和以前的價格。如果不保留價格更改的日期，則不是可能的將圖書價格與圖書中的 TimeID 信息相關聯銷售情況表。因此，可以生成的報告將需要支付顆粒物r 注意書中信息之間可能存在的不匹配 price 列（來自 Book Dimension 表）和 TimeID 列（來自圖書銷售情況表）。

SCD Type 3 的星型圖如圖 1 所示。6.9.

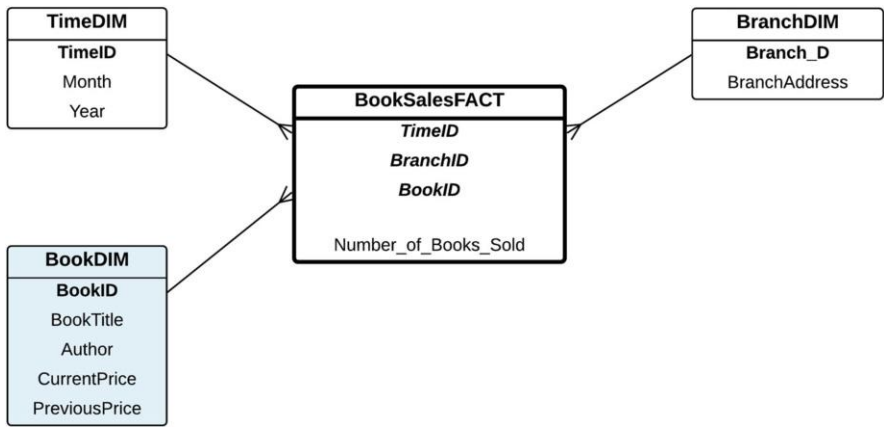


圖 6.9 圖書維度的 SCD 類型 3 的書店星型模式

表6.18 噓k 維度
BookID 書名 作者
桌子

| | | |
|----|----------|-----------|
| | | |
| C1 | CSIRO 飲食 | CSIRO 團隊 |
| H6 | 哈利波持 | 6羅琳 |
| DV | 達芬奇密碼 | 丹布朗 |

表6.19 噓k 價格維度表

| BookID | 開始日期 | 結束日期 | 價格 | 備註 |
|--------|---------|---------|---------|--------|
| C1 | Jan2007 | Jul2007 | \$45.95 | 全價 |
| C1 | Aug2007 | Oct2007 | \$36.75 | 20% 折扣 |
| C1 | Nov2007 | Jan2008 | \$23.00 | 半價 |
| C1 | Feb2008 | Dec9999 | \$45.95 | 全價 |
| H6 | Jan2007 | Mar2007 | \$21.95 | |
| H6 | Apr2007 | Jan2008 | \$30.95 | 全價 |
| H6 | Feb2008 | Dec9999 | \$10.00 | 產品銷售結束 |
| DV | Jan2007 | Dec9999 | \$27.95 | 全價 |
| | | | | |

6.4.4 SCD 類型 4

SCD 型 4 是一種治療 SCD 的新方法，它不在原始 SCD 理論中。在 SCD Type 4，我們創建一個新維度來維護屬性值的歷史 變化。基本上，時間維度開頭有描述 本章是 SCD Type 4。在 Type 4 中，保留了原始 Book Dimension 表 沒有價格屬性。價格屬性，開始日期和結束日期是分開的 進入另一個表，即 Book Price Dimension 表。圖書維度 並且圖書價格維度表是 shown 在表中6.18和6.19.

Type 4 的主要優點是我們不需要有不同的 BookID。對於同一本書。此外，還會保留整個更改歷史記錄。如圖所示。早些時候，這種方法將保證加入信息的報告從圖書銷售事實表和維度表將是準確的，並將反映正確的書在某個 TimeID 的價格。顯示了 SCD 類型 4 的星型模式之前在圖 6.2。

6.4.5 SCD 類型 6

SCD Type 6 是 Type 2 和 Type 3 的組合。在 Type 3 中，只有當前價格和記錄了以前的價格，但不是整個歷史記錄。在類型 2 中，整個歷史的更改被保留，但有一個單獨的標識符（例如代理鍵）是需要的。在類型 6 中，不需要為同一本書使用單獨的標識符（如噸類型 3），但保留整個歷史記錄（類似於類型 2）。本書的 SCD 類型 6 尺寸表如表所示 6.20.

在類型 6 中，無需維護單獨的歷史記錄表。歷史本身保存在原始維度表中。顯示了 SCD 類型 6 的星型模式在圖 6.10。請注意，Book Dimension 表有一個複合鍵，包括 BookID、開始日期和結束日期。

表6.20 噓k 尺寸 (SCD 類型 6)

| 書名 | 開始日期 | 結束日期 | 當前頁數 | 前一頁ID | 當前ID | 標題 | 作者 | 日期 | 日期 | 價格 | 價格 | 備註 | 標誌 |
|----------------------|-------------|------------|----------|----------|------|----|----|----|----|----|----|----|----|
| C1 CSIRO 飲食 CSIRO 團隊 | 2007 年 1 月 | 2007 年 7 月 | 45.55 美元 | 空 | 全價 | N | | | | | | | |
| C1 CSIRO 飲食 CSIRO 團隊 | Aug2007 | Oct2007 | \$36.75 | \$45.95 | 20% | 折扣 | N | | | | | | |
| C1 CSIRO 飲食 CSIRO 團隊 | Nov2007 | Jan2008 | \$23.00 | \$36.75 | 半價 | N | | | | | | | |
| C1 CSIRO 飲食 CSIRO 團隊 | 2008 年 12 月 | 9999 美元 | 45.95 美元 | 23.00 美元 | 全價 | Y | | | | | | | |
| H6 哈利波特 6 羅琳 | 2007 年 1 月 | 2007 年 3 月 | 21.95 美元 | 空 | 發射 | N | | | | | | | |
| H6 哈利波特 6 羅琳 | 2007 年 4 月 | 2008 年 1 月 | \$30.95 | \$21.95 | 全價 | N | | | | | | | |
| H6 哈利波特 6 羅琳 | 2008 年 12 月 | 9999 美元 | 10.00 美元 | 30.95 美元 | 產品結束 | Y | | | | | | | |
| DV 達芬奇密碼丹布朗 | Jan2007 | Dec9999 | 27.95 美元 | 空 | 全價 | Y | | | | | | | |

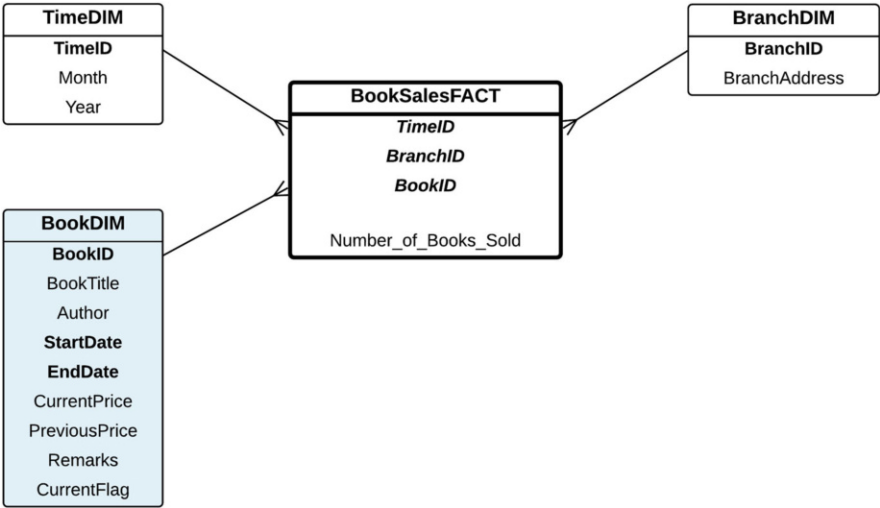


圖 6.10 圖書維度的 SCD 類型 6 的書店星型模式

如果我們檢查這個星型模式（和 Book Dimension 表），Book 維度表有一個複合鍵，包括 BookID、開始日期和 結束日期，而事實表使用只有 BookID 作為對 Book 的引用 維度表。服用以BookID C1為例，C1有4條記錄 Book Dimension 表中，Fact Table 中有很多 C1 的記錄。Book Dimension 表和 Fact Table 之間的基數關係為 不再 1-m，但毫米。通常，有三種可能的選擇來解決這個問題 問題：

- 1.向 Book Dimension 表添加一個新的代理鍵，因此 Fact 桌子將使用 Book Dimension 中的此代理鍵作為參考。這個 代理key 也可以簡單地是 BookID、Start Date 之間的連接 和結束日期。或者，代理鍵是帶有序列號的 BookID，如 SCD 類型 2。
- 2.如果我們不想要 Book Dimension 表中的代理鍵，那麼 Fact 表還必須除了 BookID 之外，還包括開始日期和結束日期。但這會很混亂，因為 Fact 已經有了 TimeID，引用 噸ime 維度表。
- 3.最後一個選項是在 Book 之間有一個關聯表（或橋接表） Dimension 和 Fact，因為 Book Dimension 和 Fact 之間的基數 是毫米。這個連接 Book Dimension 和 Fact 的中間表將有一個 合成的鍵包含來自 Book Dimension 和 Fact 的標識符。這個也是一個亂七八糟的選擇，因為我們試圖用關係理論來解決 這個mm問題。

解決這個問題的最簡單方法是採用第一個選項，使 SCD 類型 6 與 SCD Type 2 幾乎相同。換句話說，SCD Type 6 是不必要的或無用。但是，如果我們堅持實施 SCD 類型 6，如星號所示圖式在圖 6.10，因為沒有 PK-FK 在表中明確實現，它是 Book Dimension 和 Fact 這兩個表仍然可以獨立存在。但是，當我們想要生成一個連接 Fact 和 Book Dimension 的報告時表 join 必須包括以下條件：(i) 兩個表的 BookID 必須是同樣，並且(ii) 事實的 TimeID 必須在開始日期和結束日期之間書籍維度的日期。這些連接條件與連接非常相似使用條件在 SCD 類型 4 或前面討論的時間數據倉庫中。在本章中，日期是 Fact、Book 之間連接條件的一部分維度和書價維度。上這個註釋，如果我們檢查 SCD 類型 6 更深刻的是，那裡 SCD Type 6 和 SCD Type 4 之間有很強的相似性。SCD Type 4 實際上是 SCD Type 6 的規範化版本，其中 Book 維度表行為就像 Book Price Dimension 和事實。SCD Type 6 相反，類型 6 是 SCD 類型 4 的非標準化版本，其中所有東西都集中在一個表中，即 Book Dimension 表。加入條件加入 Book Dimension 和 Fact 是相同的。

6.4.6 SCD 在 SQL 中的實現

SCD 的概念很容易理解。但是，在 SQL 中實現 SCD 有幾個具有挑戰性的技術方面。在本節中，我們將檢查 SQL 創建六種 SCD 類型。

操作數據庫中的源表是 Book 表和 Book 價格歷史表，如 E/R 圖所示（見圖 6.1）。結構這些兩張表是：

- 書籍 (BookID、BookTitle、作者)
- BookPriceHistory (BookID、StartDate、EndDate、價格、備註)

SCD Type 0 是一個非時間維度。使用 SCD 的 Book Dimension 表與類型 0 見表 6.13。請注意，“原價”可能不是“原價”價格，因為有些書哈哈推出價格可以大幅打折。因此，Book Dimension 表使用 SCD Type 0 檢查備註“全價”。

```
創建表 SCD0 為
選擇不同的
    B.BookID, B.BookTitle,
    B.Author, H.Price 作為 OriginalPrice
來自 B 書, BookPriceHistory H
其中 B.BookID = H.BookID
和 H.Remarks = '全價';
```

SCD Type 1 使用最新價格（參考表6.14）。我們可以選擇一本書結束日期是“Dec9999”。或者，我們可以按開始日期和選擇“最新”開始日期。在下面的 **SQL** 中，內部查詢對書籍進行排名基於開始日期的記錄按降序排列，外部查詢選擇那些書排名為 1。

```

創建表 SCD1 為
選擇
    T.BookID, T.BookTitle,
    T.Author, T.Price 作為 CurrentPrice
從 (
    選擇
        B.BookID, B.BookTitle, B.Author,
        to_date(H.StartDate, 'MonYYYY'), H.Price,
        rank() over( 按 B.BookID 劃分
            按 to_date(H.StartDate, 'MonYYYY') desc) 排序
        作為等級
    來自 B 書, BookPriceHistory H
    其中 B.BookID = H.BookID) T
其中 T.Rank = 1;

```

SCD Type 2 是 **Book** 表和 **Book Price History** 表之間的連接，所以那我們可以獲得書籍的詳細信息以及開始日期和結束日期。然而，一個書與之前一樣，應添加具有相同 **BookID** 的序列號 **seqnum** 在表中6.15。

SCD 類型 2 的 SQL 使用 **rank** 函數以及分區子句，所以同一本書將根據他們的開始日期排名。當前標誌柱子將識別賬簿記錄是否具有當前價格。

```

創建表 SCD2 為
選擇 B.BookID || '_' ||
    rank() over(按 B.BookID 分區
        按 to_date(H.StartDate, 'MonYYYY') asc) 排序
    作為 BookID,
    B.BookTitle, B.Author, H.StartDate,
    H.EndDate, H.Price, H.Remarks,
    case H.EndDate 當 'Dec9999' 然後 'Y' 否則 'N'
    以 CurrentFlag 結尾
來自 B 書, BookPriceHistory H
其中 B.BookID = H.BookID;

```

SCD T類型 3 相當複雜，因為它不僅有當前價格（類似於 **SCD 類型 2**）也是以前的價格。如果 **SCD Type 2** 使用 **rank 1**，**SCD Type 3** 需要 排名 2（對於以前的價格）也是如此。因此，**SCD 類型 3** 以排名加入表 1和表排名為 2。由於有些書沒有以前的價格，我們需要使用外連接而不是內連接。

```

創建表 SCD3 為
選擇
    T1.BookID, T1.BookTitle, T1.Author,
    T1.CurrentPrice, T2.CurrentPrice 作為 PreviousPrice
從 (

```

```

選擇
    T.BookID, T.BookTitle,
    T.Author, T.Price 作為 CurrentPrice
從 (
    選擇
        B.BookID, B.BookTitle,
        B.Author, to_date(H.StartDate, 'MonYYYY'),
        H.價格,
        rank() over( 按 B.BookID 劃分
            按 to_date(H.StartDate, 'MonYYYY') desc) 排序
            作為等級
        來自 B 書, BookPriceHistory H
        其中 B.BookID = H.BookID) T
其中 T.Rank = 1) T1,
(選擇
    T.BookID, T.BookTitle,
    T.Author, T.Price 作為 CurrentPrice
從 (
    選擇
        B.BookID, B.BookTitle, B.Author,
        to_date(H.StartDate, 'MonYYYY'), H.Price,
        rank() over( 按 B.BookID 劃分
            按 to_date(H.StartDate, 'MonYYYY') desc) 排序
            作為等級
        來自 B 書, BookPriceHistory H
        其中 B.BookID = H.BookID) T
其中 T.Rank = 2) T2
其中 T1.BookID = T2.BookID(+);

```

SCD 類型 4 實際上是第一部分討論的時態數據倉庫 的這一章。創建 Book Dimension 的 SQL 是從 書價來自操作數據庫的歷史表。

```

創建表 SCD4 為
從 BookPriceHistory 中選擇 *;

```

SCD T類型 6 是 SCD 類型 2 和 SCD 類型 3 的組合。但是，SC 的實施 SQL 中的 D 類型 6 可能非常具有挑戰性。我們不能簡單地 加入 SCD Type 2 和 SCD Type 3，因為 SCD Type 2 擁有完整的歷史 價格，而 SCD 類型 3 只有當前和以前的價格。此外，SCD Type 2 和 SCD Type 3 之間的連接不能使用等值連接。相反，我們 需要檢查 SCD 類型 3 的 BookID 是否是 SCD 類型 2 的 BookID 的一部分，使用喜歡操作員。

創建 SCD TSCD Type 2 和 SCD Type 3 的 ype 6 需要幾個步驟。第一步是連接 SCD Type 2 和 SCD Type 3 使用喜歡操作員。連接結果使用 SCD 類型 3 的 BookID 屬性，其餘的 屬性來自 SCD 類型 2。此外，訂單號屬性將具有 序號對於每組 BookID。此連接查詢的結果是 笑wn 在表中6.21.

表6.21 SCD T之間的連接類型 2 和 SCD 類型 3

| Book | Book | Start | End | Price | 當前訂單ID | 標題 | 作者 | 日期 | 日期 | 備註 | 標誌編號 | C1 | CS | RO | 飲食 | CS | RO | 團隊 |
|---------|---------|-----------|-------------|---------------|---------|-------|---------|---------|---------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| Jan2007 | Jul2007 | \$45.95 | 全價 | N 1 | | | | | | | | | | | | | | |
| C1 | CS | RO | 飲食 | CS | RO | 團隊 | Aug2007 | Oct2007 | \$36.75 | 20% | 折扣 | N 2 | | | | | | |
| C1 | CS | RO | 飲食 | CS | RO | 團隊 | Nov2007 | Jan2008 | \$23.00 | 半價 | N 3 | | | | | | | |
| C1 | CS | RO | 飲食 | CS | RO | 團隊 | Feb2008 | Dec9999 | \$45.95 | 全價 | Y 4 | | | | | | | |
| DV | 達芬奇密碼 | Dan Brown | Jan2007 | Dec9999 | \$27.95 | 全價 | Y 1 | | | | | | | | | | | |
| H6 | 哈利波特 6 | 羅琳 | 2007 年 1 月 | 2007 年 3 月 21 | .95 | 美元 | 推出 | N 1 | | | | | | | | | | |
| H6 | 哈利波特 6 | 羅琳 | Apr 2007 | Jan2008 | \$30.95 | 全價 | N 2 | | | | | | | | | | | |
| H6 | 哈利波特 6 | 羅琳 | 2008 年 12 月 | 9999 年 10.00 | 美元 | 產品 | 結束 | Y 3 | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | |

選擇

```
SCD3.BookID,  
SCD2.書名, SCD2.作者,  
SCD2.StartDate, SCD2.EndDate,  
SCD2.Price, SCD2.Remarks, SCD2.CurrentFlag,  
row_number() 結束  
  (按 SCD3.BookID 分區  
   按 SCD3.BookID、SCD2.StartDate 排序) 為  
   訂單號  
從 SCD2、SCD3  
其中 SCD2.BookID 類似於 SCD3.BookID||'|_-%'  
按 SCD3.BookID、SCD2.StartDate 排序;
```

下一步是做一個表的自連接6.21.第一張表的價格
將成為當前價格，第二張表中的價格將成為
預惡性價格。此自聯接查詢有兩個聯接條件。第一個 條件是 BookID
必須相同。為了從第二個價格
桌子要成為上一個價格，我們需要在第二個中添加訂單號
一個表，然後是第二個加入條件是匹配訂單號 這些兩張桌子。自連接的 SQL
如下，而結果顯示 在表中6.22.

選擇

```
T1.BookID, T1.BookTitle, T1.Author,  
T1.開始日期, T1.結束日期,  
T1.Price 作為 CurrentPrice,  
T2.Price 作為 PreviousPrice,  
T1.Remarks, T1.CurrentFlag  
從 (  
  選擇 SCD3.BookID,  
        SCD2.BookTitle、SCD2.Author、SCD2.StartDate、
```

SCD2.EndDate, SCD2.Price, SCD2.Remarks,
SCD2.CurrentFlag,
row number() 結束
 (按 SCD3.BookID 分區
 按 SCD3.BookID、SCD2.StartDate 排序) 為
 訂單號
從 SCD2、SCD3
其中 SCD2.BookID 類似於 SCD3.BookID||'_'
按 SCD3.BookID、SCD2.StartDate) T1 排序,
(
選擇 SCD3.BookID,
 SCD2.BookTitle、SCD2.Author、SCD2.StartDate、
 SCD2.EndDate, SCD2.Price, SCD2.Remarks,
 SCD2.CurrentFlag,
 row number() 結束
 (按 SCD3.BookID 分區
 按 SCD3.BookID、SCD2.StartDate) + 1 排序
 訂單號
從 SCD2、SCD3
其中 SCD2.BookID 類似於 SCD3.BookID||'_'
按 SCD3.BookID、SCD2.StartDate 訂購) T2
其中 T1.BookID = T2.BookID
和 T1.OrderNumber = T2.OrderNumber
按 T1.BookID、T1.StartDate 排序;

然而，自連接結果排除了每本書的第一條記錄，因為那裡不是每本書的第一個記錄的先前價格。為了解決這個問題，自連接查詢需要使用外連接，而不是內連接。使用外連接的 `SCD` 類型 6 的正確 `SQL` 查詢與高於 `SQL` 使用 `inner join` 的查詢，只有一個區別，就是 `join`

表6.22 自加入結果

[illegible]

表6.23 正確的自連接結果（SCD 類型 6）

| 書 | 開始 | 結束 | 當前 | 上一頁 | 當前 | | | | | | | |
|-------|----------|----------|-------------|------------|----------|----------|--------|-------|--|--|--|--|
| ID | 標題 | 作者 | 日期 | 日期 | 價格 | 價格 | 備註 | 標誌 | | | | |
| C1 | CSIRO 飲食 | CSIRO 團隊 | 2007 年 1 月 | 2007 年 7 月 | 45.95 美元 | 空 | 全價 | N | | | | |
| C1 | CSIRO 飲食 | CSIRO 團隊 | Aug2007 | Oct2007 | \$36.75 | \$45.95 | 20% 折扣 | N | | | | |
| C1 | CSIRO 飲食 | CSIRO 團隊 | Nov2007 | Jan2008 | \$23.00 | \$36.75 | 半價 | N | | | | |
| C1 | CSIRO 飲食 | CSIRO 團隊 | 2008 年 12 月 | 9999 美元 | 45.95 美元 | 23.00 美元 | 全價 | Y | | | | |
| DV | 達芬奇密碼 | 丹布朗 | Jan2007 | Dec9999 | 27.95 美元 | 空 | 全價 | Y | | | | |
| H6 | 哈利波特 6 | 羅琳 | 2007 年 1 月 | 2007 年 3 月 | 21.95 美元 | 空 | 發射 | N | | | | |
| H6 | 哈利波特 6 | 羅琳 | 2007 年 4 月 | 2008 年 1 月 | \$30.95 | \$21.95 | 全價 | N | | | | |
| H6 | 哈利波特 6 | 羅琳 | 2008 年 12 月 | 9999 美元 | 10.00 美元 | 30.95 美元 | 產品結束 | Y | | | | |
| | | | | | | | | | | | | |

條件現在是一個外連接（在每個連接條件的末尾都有一個 (+) 號），如下：

```
其中 T1.BookID = T2.BookID (+)
和 T1.OrderNumber = T2.OrderNumber (+)
```

最終結果如表所示6.23.

6.4.7 創建事實表

在解決了 SCD 的實施挑戰之後，下一步是檢查 平方L 命令來創建事實表。對於 SCD 類型 0、1、3 和 4（參見樣本記錄在表格中6.13,6.14,6.17,6.18和6.19），BookID 屬性為 e與數據倉庫的“非臨時”版本中的 BookID 完全相同（參考到表6.2屆時本章的軋花）。因此，事實表將 時不受影響我們使用 SCD 類型 0、1、3 或 4。原始書本尺寸， 作為笑wn 在非臨時版本中，具有 BookID 的三本書：C1（CSIRO 飲食），H6（哈利波特）和DV（達芬奇密碼）。SCD 中的書本尺寸 0、1、3 和 4 型也有ve 與原始 Book Dimension 相同的 BookID，和記錄數在這些表中是相同的。

但是，在生成報告時，通過加入事實表和書籍 維度表，當所有的存在書籍維度的屬性（包括價格），我們無法在 Fact 中繪製 TimeID 之間的相關性 和書價在 Book Dimension 表中，因為 Book Price 顯示在報告上可能不是 TimeID 上的價格。這個問題在 桌子6.3. 這就是為什麼我們不使用 SCD 類型 0、1 或 3，而是使用 SCD 類型 4，這在本章前面已經討論過了。使用 SCD Type 4，我們不僅保持書價的完整歷史；我們還能夠在以下情況下生成正確的報告 加入事實表。

對於 SCD T類型 2 (參見表中的示例記錄6.15), 記錄數 與原始 Book Dimension 表不同。這是因為每次 噓k 價格被改變, 一個“新”的書被插入到維度表中。這個 對內容有影響事實表。例如, 在事實表中, 它 不該是BookID C1 為 CSIRO 飲食書, 但 C1_1 或 C1_2 等, 可參考 到歷史中不同的 Book 實例。因此, 當使用 SCD 類型 2 時, Fac的內容t 表必須包含正確的 BookID。SQL 命令 創建事實表對於 SCD 類型 2 如下:

```
創建表 BookSalesFactWithSCD2 為
選擇
    to_char(T.TransactionDate, 'MonYYYY') 作為 TimeID,
    BK.BookID,
    BR.BranchID,
    sum(T.Quantity) as Number_of_Books_Sold
來自 BookTransaction T, SCD2 BK, Branch BR
其中 T.BranchID = BR.BranchID
和 BK.BookID 像 T.BookID||'_%'
和 to_date(BK.StartDate, 'MonYYYY') <= T.TransactionDate
和 T.TransactionDate <= to_date(BK.EndDate, 'MonYYYY')
通過...分組
    to_char(T.TransactionDate, 'MonYYYY'),
    BK.BookID,
    BR.BranchID;
```

有三種注意事項: (i) 它與 Book Dimension SCD Type 2 結合。原因是操作數據庫中的Book Table沒有 新的 BookID。新的 BookID 僅存在於 SCD 類型 2 中, 因為新的 BookID BookID 是在更改圖書價格時創建的。(ii) 連接條件使用 一個LIKE 和通配符以匹配 SCD 類型 2 中的新 BookID 和 BookID 在操作數據庫的 Boo Transaction 表中。(iii) 加入條件 必須包括檢查日期, 以便在 Book Dimension 中正確記錄 SCD T與 Book Transaction 表中的記錄連接時使用 ype 2 從操作數據庫。事實表如表所示6.24.

SCD Type 6 與 SCD Type 2 類似, 即包含相同數量的 記錄。唯一的區別是在 SCD 類型 6 中, BookID 不會改變, 並且 這使用原始 BookID。因此, 創建 Fact Table 的 SQL 略 更簡單。在連接條件中, 它只是比較來自 SCD 類型 6 的 BookID 和 噓k 事務表 (請注意, 在 SCD 類型 2 事實表中, 它使用 LIKE 和通配符)。為 SCD 類型 6 創建事實表的 SQL 命令如下如下:

```
創建表 BookSalesFactWithSCD6 為
選擇
    to_char(T.TransactionDate, 'MonYYYY') 作為 TimeID,
    BK.BookID,
    BR.BranchID,
    sum(T.Quantity) as Number_of_Books_Sold
來自 BookTransaction T, SCD6 BK, Branch BR
其中 T.BranchID = BR.BranchID
```

表6.24 流式細胞儀表 (SCD 類型 2)

| TimeID | BranchID | BookID | 售出書籍數量 |
|----------|-------------------|--------|--------|
| 2008年3月 | 城市C1_4 | 5 | |
| 2008年3月 | 城市H6_3 | 15 | |
| 2008年3月 | 城市DV_1 | 23 | |
| 2008年3月 | 市 | | |
| 2008年3月 | 查斯頓C1_4 | 15 | |
| 2008年3月 | 查斯頓H6_3 | 3 | |
| 2008年3月 | ChadstoneDV_1 | 2 | |
| 2008年3月 | Chadstone | | |
| 2008年3月 | 坎伯韋爾C1_4 | 1 | |
| 2008年3月 | 坎伯韋爾H6_3 | 1 | |
| 2008年3月 | 坎伯韋爾DV_1 | 2 | |
| 2008年3月 | Camberwell | | |
| 2008年3月 | | | |
| | ... | | |
| | ... | | |
| 2007年12月 | 城市C1_3 | 15 | |
| 2007年12月 | 城市H6_2 | 6 | |
| 2007年12月 | 城市DV_1 | 6 | |
| 2007年12月 | 市 | | |
| 2007年12月 | 查斯頓C1_3 | 10 | |
| 2007年12月 | 查斯頓H6_2 | 8 | |
| 2007年12月 | ChadstoneDV_1 | 1 | |
| 2007年12月 | Chadstone | | |
| 2007年12月 | 坎伯韋爾C1_3 | 18 | |
| 2007年12月 | 坎伯韋爾H6_2 | 3 | |
| 2007年12月 | 坎伯韋爾DV_1 | 2 | |
| 2007年12月 | Camberwell | | |
| 2007年12月 | | | |
| | ... | | |

和 BK.BookID = T.BookID
和 to_date(BK.StartDate, 'MonYYYY') <= T.TransactionDate
和 T.TransactionDate <= to_date(BK.EndDate, 'MonYYYY')
通過...分組
to_char(T.TransactionDate, 'MonYYYY'),
BK.BookID,
BR.BranchID;

事實表如表所示6.25. 它的記錄幾乎與 桌子6.24, 除了這裡的保留原始 BookID。

值得注意的是, 當我們加入 SCD Type 2 和 Fact Table 時, 或者什麼時候我們加入 SCD 類型 6 和事實表, 正確的賬面價格將是 笑wn (因為 Book Price 與 TimeID 匹配), 因此不會有 一個 報告中這兩個屬性之間的相關性不正確。的原因

表6.25 流式細胞儀表 (SCD
類類型 6)

| imeID | BranchID | BookID | 售出書籍數量 |
|-------------|------------------|--------|--------|
| 2008 年 3 月 | 城市 C1 | 5 | |
| 2008 年 3 月 | 城市 H6 | 15 | |
| 2008年3月 | 城市DV | 23 | |
| 2008年3月 | 市 | | |
| 2008 年 3 月 | 查斯頓 C1 | 15 | |
| 2008 年 3 月 | 查斯頓 H6 | 3 | |
| 2008 年 3 月 | Chadstone | DV 2 | |
| 2008年3月 | Chadstone | | |
| 2008 年 3 月 | 坎伯韋爾 C1 | 1 | |
| 2008 年 3 月 | 坎伯韋爾 H6 | 1 | |
| 2008 年 3 月 | 坎伯韋爾 DV | 2 | |
| 2008 年 3 月 | Camberwel | | |
| 2008 年 3 月 | | | |
| | ... | | |
| | ... | | |
| 2007 年 12 月 | 城市 C1 | 15 | |
| 2007 年 12 月 | 城市 H6 | 6 | |
| 2007 年 12 月 | 城市 DV | 6 | |
| 2007年12月 | 市 | | |
| 2007 年 12 月 | 查斯頓 C1 | 10 | |
| 2007 年 12 月 | 查斯頓 H6 | 3 | |
| 2007 年 12 月 | 查斯頓 DV | 1 | |
| 2007 年 12 月 | Chadstone | | |
| 2007 年 12 月 | 坎伯韋爾 C1 | 18 | |
| 2007 年 12 月 | 坎伯韋爾 H6 | 3 | |
| 2007 年 12 月 | 坎伯韋爾 DV | 2 | |
| 2007 年 12 月 | Camberwell | | |
| 2007 年 12 月 | | | |
| | ... | | |

這是 SCD Type 2 和 Type 6 都保持了 Book 的完整歷史 價格; 他們保留每個書價的開始日期和結束日期。

SCD Type 4 (我們的時態數據倉庫版本) 也保持完整 歷史f 賬面價格, 因此, 加入事實表時的報告和 噓k維表也是正確的, 如前表所示6.6. SCD 類類型 2 和 SCD 類型 6 也將生成相同的報告。唯一的區別是 噓的方式k 價格歷史被保留。SCD Type 4 使用兩個表 (即 Book Dimension 和 Book Price Dimension), 而 SCD 類型 2 和 SCD 類型 6 使用 一張桌子僅, 在 Book Dimension 中維護 Book Price 歷史記錄。

6.5 總結

在本章中，我們重點關注將歷史數據合併到數據倉庫中。這個
稱為時態數據倉庫。時態數據倉庫使用這個概念 of
橋表（或弱實體），其中歷史保存在橋中
桌子。維護某些屬性的歷史是很重要的，以便使
聯想在分析事實產生的報告時，分析更準確
和尺寸。但是，在加入時需要一定程度的謹慎 流式細胞儀
表和時間維度，尤其是當粒度級別為 事實和時間維度之間的時間是不一樣的。
時態數據倉庫也稱為緩慢變化的尺寸 (SCD)。 本章介紹了 SCD
的各種治療方法和類型。不同種類 將用於數據倉庫的不同目的。

6.6 練習

6.1 大學裡的教授從事諮詢是很常見的
項目與一家公司。這項諮詢工作通常被視為一項研究活動，
作為教授的研究專長某公司正在使用特定的
項目。大學的研究辦公室跟踪所有諮詢工作 教授承擔。執行此操作的簡單 E/R
圖如圖 1 所示。6.11. 每個諮詢項目都有項目編號和T它，以及公司誰
假如進行諮詢的金額、金額和年份。
出於研究報告目的，研究大學的某位負責人分配了一個
研究每個諮詢工作的字段。一些研究領域的例子是人工智能，
大數據、可視化、Web 開發、移動應用程序等。
另請注意，教授的職位歷史也會被保留。這是 看教授說的金額走勢特別有用
安全的s 從公司晉升為教授級別的階梯 （例如助理教授、副教授和教授）。
任務是設計一個星型模式來跟踪 Total Consulting
每年，每個研究領域以及每個教授的金額。編寫 SQL 命令創建維度和事實表。
6.2 員工維度是一個緩慢變化的維度（SCD）。為簡單起見，
維度包含只有這些屬性：EmployeeID、Name 和 Salary。SCD 有六種類型：類型
0、1、2、3、4 和 6。為員工繪製樣本表 使用每種 SCD
類型進行標註。假設工資屬性是 顯屬性,這可能會不時改變，由於促銷，增加 of
薪水等。在 SCD 需要時向表中添加更多屬性。為了
每種類型，您需要至少有兩個員工記錄，這是兩個 僱員：

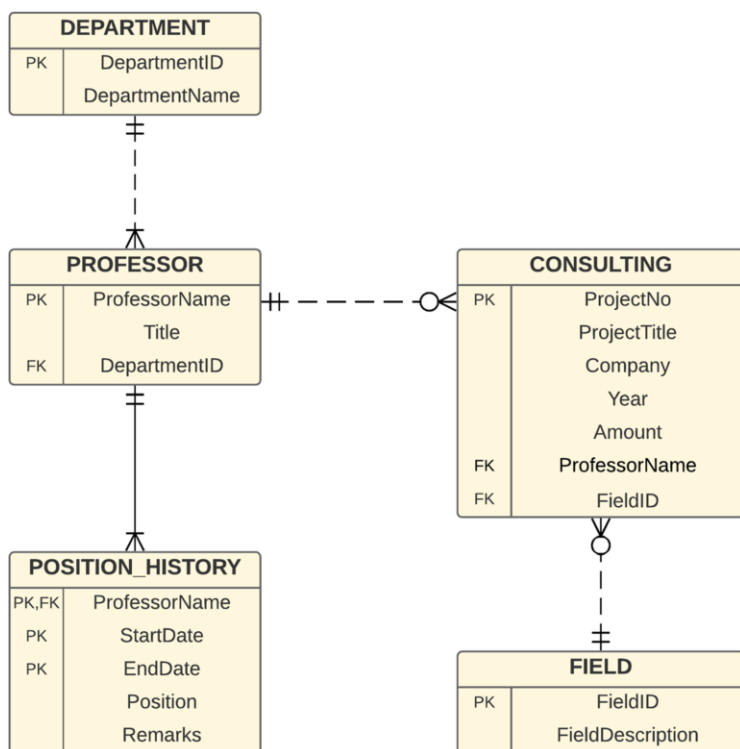


圖 6.11 諮詢E/R 圖

- 第一個員工是 Adam，其 `EmployeeID` : A1。他於 1 年開始工作 2016 年 3 月，工資 3900 美元。在他的試用期於 2016 年 5 月 31 日結束後，他的薪水增加到4300美元。2017年初，他收到了一份工資增加，並且他的薪水變成了 5000 美元。從 2017 年 7 月 1 日（至今），他的薪水已經是5500美元。
- 第二位員工是 Ben，其 `EmployeeID` : B2。他於 1 年開始工作 2017 年 2 月初始工資為 4000 美元。在他的試用期結束後 31 2017 年 5 月，他的薪水變成了 4750 美元，這是他目前的薪水。

6.7 進一步閱讀

斯洛夫ly 變化維度 (SCD) 已在各種數據中進行了討論和解釋 倉儲書籍，例如[1-7] 和 [8]。已經討論過延長 SCD通過組合幾個基本的 SCD，從 6 型到更複雜的 SCD。

[9] 討論時態數據倉庫的對象關係實現，使用多值屬性、對象和類。有關對象的更多詳細信息-

的關係實現RDBMS，例如 Oracle，可以在 下本書：[10]。
很早時態數據庫上的 `ork` 可以在 [11–14] 和[15]。

參考

1. C. Adamson, *Star Schema The Complete Reference* (麥格勞-希爾奧斯本媒體, 2010 年)
2. R.拉貝熱, *數據倉庫導師：實用數據倉庫和商業智能見解* (麥格勞-希爾, 紐約, 2011)
3. R. Golfarelli, S. Rizzi, *數據倉庫設計：現代原理和方法* (麥格勞-希爾, 紐約, 2009 年)
4. R. Kimball, M. Ross, *數據倉庫工具包：維度的權威指南造型* (威利, 倫敦, 2013 年)
5. R. Kimball, M. Ross, W. Thorntwaite, J. Mundy, B. Becker, *數據倉庫生命週期噸工具包* (威利, 倫敦, 2011 年)
6. W. H. Inmon, *構建數據倉庫。ITPro 系列* (威利, 倫敦, 2005 年)
7. M. Jarke, *數據倉庫基礎*, 第 2 版。 (施普林格, 柏林, 2003)
8. A. Vaisman, E. Zimányi, *數據倉庫系統：設計和實施。以數據為中心系統和應用程序* (施普林格, 柏林, 2014 年)
9. E. Malinowski, E. Zimányi, *高級數據倉庫設計：從傳統到空間和天時 應用。以數據為中心的系統和應用程序* (施普林格, 柏林, 2008 年)
10. J.W.拉哈尤, D. Tanar, E. Pardede, *面向對象的 Oracle。IT 專業人員的解決方案* (IRM 出版社, 2006 年)
11. RT斯諾德格拉斯, I. Ahn, *時態數據庫*. 計算機 19(9), 35–42 (1986)
12. RT Snodgrass, *Temporal 數據庫, 在時空推理的理論和方法中在地理空間, 國際會議地理信息系統——從空間到領土：理論和時空推理方法*, 比薩, 意大利, 1992 年 9 月 21-23 日, 訴訟, 編。經過AU Frank, I. Campari, U. Formentini. 計算機科學講義, 卷。639 (施普林格, 柏林, 1992), 第 22-64 頁
13. CS Jensen, RT Snodgrass, *時間數據管理*. IEEE Trans. 知道。數據工程 11 (1), 36–44 (1999)
14. RT斯諾德格拉斯, I. Ahn, *數據庫中的時間分類*, 1985 年 ACM 會議記錄 SIGMOD國際數據管理會議, 美國德克薩斯州奧斯汀, 5 月 28 日至 31 日, 1985年, 編。由某人Navathe (ACM 出版社, 紐約, 1985 年), 第 236-246 頁
15. RT斯諾德格拉斯, *使用 SQL 開發面向時間的數據庫應用程序* (摩根考夫曼, 洛斯阿爾托斯, 1999)