

***IBM SPSS Modeler 18.1.1 使用手冊***

**IBM**

**附註**

在使用本資訊及其支援的產品之前，請先閱讀第 193 頁的『注意事項』中的資訊。

**產品資訊**

本版適用於 IBM SPSS Modeler 18.1.1 版及所有後續版本和修正，直到新版中另有指示。

# 目錄

<b>第 1 章 關於 IBM SPSS Modeler . . . . .</b>	<b>1</b>	<b>第 5 章 建立串流 . . . . .</b>	<b>31</b>
IBM SPSS Modeler 產品 . . . . .	1	串流建立概觀 . . . . .	31
IBM SPSS Modeler . . . . .	1	建置資料串流 . . . . .	31
IBM SPSS Modeler Server . . . . .	1	使用節點 . . . . .	31
IBM SPSS Modeler Administration Console . . . . .	2	使用串流 . . . . .	36
IBM SPSS Modeler Batch . . . . .	2	串流說明 . . . . .	47
IBM SPSS Modeler Solution Publisher . . . . .	2	執行串流 . . . . .	48
IBM SPSS Collaboration and Deployment Services 的 IBM SPSS Modeler Server 配接器 . . . . .	2	使用模型 . . . . .	48
IBM SPSS Modeler 版本 . . . . .	2	新增備註和註解到節點和串流 . . . . .	49
說明文件 . . . . .	3	儲存資料串流 . . . . .	54
SPSS Modeler Professional 文件 . . . . .	3	載入檔案 . . . . .	55
SPSS Modeler Premium 文件 . . . . .	3	對映資料串流 . . . . .	56
應用程式範例 . . . . .	4	提示和快捷鍵 . . . . .	58
Demos 資料夾 . . . . .	4	<b>第 6 章 輸出的處理 . . . . .</b>	<b>59</b>
授權追蹤 . . . . .	4	檢視器 . . . . .	59
<b>第 2 章 IBM SPSS Modeler 18.1.1 新增功能 . . . . .</b>	<b>5</b>	顯示和隱藏結果 . . . . .	59
<b>第 3 章 產品概觀 . . . . .</b>	<b>7</b>	移動、刪除和複製輸出 . . . . .	60
新手啟動使用 . . . . .	7	變更起始對齊方式 . . . . .	60
啟動 IBM SPSS Modeler . . . . .	7	變更輸出項目的對齊方式 . . . . .	60
從指令行啟動 . . . . .	8	檢視器概要 . . . . .	61
連線到 IBM SPSS Modeler Server . . . . .	8	新增項目至檢視器 . . . . .	62
連線到 Analytic Server . . . . .	10	尋找並取代檢視器中的資訊 . . . . .	62
變更 temp 目錄 . . . . .	10	複製輸出到其他應用程式中 . . . . .	63
啟動多個 IBM SPSS Modeler 階段作業 . . . . .	11	互動式輸出 . . . . .	64
IBM SPSS Modeler 介面概覽 . . . . .	11	匯出輸出 . . . . .	65
IBM SPSS Modeler 串流畫布 . . . . .	12	HTML 選項 . . . . .	66
節點選用區 . . . . .	13	Web 報告選項 . . . . .	66
IBM SPSS Modeler 管理員 . . . . .	13	Word/RTF 選項 . . . . .	67
IBM SPSS Modeler 專案 . . . . .	15	Excel 選項 . . . . .	67
IBM SPSS Modeler 工具列 . . . . .	15	PowerPoint 選項 . . . . .	68
自訂工具列 . . . . .	17	PDF 選項 . . . . .	68
自訂 IBM SPSS Modeler 視窗 . . . . .	17	文字選項 . . . . .	69
變更串流的圖示大小 . . . . .	18	只有圖形選項 . . . . .	70
在 IBM SPSS Modeler 中使用滑鼠 . . . . .	19	圖形格式選項 . . . . .	70
使用快速鍵 . . . . .	19	檢視器列印 . . . . .	71
正在列印 . . . . .	20	若要列印輸出與圖表 . . . . .	71
實現 IBM SPSS Modeler 的自動化 . . . . .	20	預覽列印 . . . . .	71
<b>第 4 章 瞭解資料採礦 . . . . .</b>	<b>21</b>	頁面屬性：頁首與頁底 . . . . .	71
資料採礦概觀 . . . . .	21	頁面屬性：選項 . . . . .	72
評量資料 . . . . .	22	儲存輸出 . . . . .	72
資料採礦的策略 . . . . .	23	若要儲存檢視器文件 . . . . .	72
CRISP-DM 程序模型 . . . . .	23	樞紐表 . . . . .	74
模型類型 . . . . .	24	樞紐表 . . . . .	74
資料採礦範例 . . . . .	29	使用樞紐表 . . . . .	74
		使用階層 . . . . .	77
		顯示和隱藏項目 . . . . .	77
		表格外觀 . . . . .	78
		表格內容 . . . . .	79
		儲存格內容 . . . . .	81
		註腳和標題 . . . . .	81

資料格寬	83
變更行寬	83
顯示樞紐表中隱藏的邊框	84
選取樞紐表中的列、行與儲存格	84
列印樞紐表	84
從樞紐表建立圖表	85
舊版表格	85
選項	85
選項	85
一般選項	86
檢視器的選項	86
樞紐表選項	86
輸出選項	87
<b>第 7 章 處理遺漏值</b>	<b>89</b>
遺漏值概觀	89
處理遺漏值	89
處理帶有遺漏值的記錄	90
處理帶有遺漏值的欄位	90
處理具有遺漏值的記錄	90
歸因或填充遺漏值	92
用於遺漏值的 CLEM 函數	93
<b>第 8 章 建立 CLEM 表示式</b>	<b>95</b>
關於 CLEM	95
CLEM 範例	95
值和資料類型	96
表示式和條件	97
Stream、Session 及 SuperNode 參數	98
字串的處理	98
處理空白值和遺漏值	99
數字的處理	99
時間和日期的處理	99
多個欄位的彙總	100
多個回應資料的處理	101
表示式建置器	102
存取表示式建置器	102
建立表示式	102
選取函數	102
選取欄位、參數和廣域變數	105
檢視或選取值	105
檢查 CLEM 表示式	106
尋找/取代	106
<b>第 9 章 CLEM 語言參照</b>	<b>109</b>
CLEM 參照概觀	109
CLEM 資料類型	109
整數	109
實數	109
字元	110
字串	110
清單	110
欄位	110
日期	111
時間(M)	112
CLEM 運算子	112

函數參照	114
函數說明中的慣例	114
資訊函數	115
轉換函數	115
比較函數	116
邏輯函數	118
數值函數	118
三角函數	119
機率函數	120
空間函數	120
位元整數作業	121
隨機函數	122
字串函數	122
SoundEx 函數	125
日期和時間函數	126
序列函數	129
廣域功能	133
處理空白值與空值的函數	134
特殊欄位	134

## 第 10 章 使用具有儲存庫的 IBM SPSS Modeler . . . . . 137

關於 IBM SPSS Collaboration and Deployment Services 儲存庫	137
儲存和部署儲存庫物件	137
連接至儲存庫	138
輸入儲存庫的認證	138
瀏覽儲存庫認證	138
瀏覽儲存庫內容	139
在儲存庫中儲存物件	139
設定物件內容	139
儲存串流	141
儲存專案	141
儲存節點	142
儲存輸出物件	142
儲存模型和模型選用區	142
從儲存庫中擷取物件	143
選擇要擷取的物件	143
選取物件版本	144
在儲存庫中搜尋物件	144
修改儲存庫物件	145
編輯、重命名和刪除資料夾	145
鎖定和解除鎖定儲存庫物件	145
刪除儲存庫物件	146
管理儲存庫物件的內容	146
檢視資料夾內容	146
檢視和編輯物件內容	147
管理物件版本標籤	148
部署串流	148
串流部署選項	149
評分分支	150

## 第 11 章 匯出到外部應用程式 . . . . . 153

關於匯出到外部應用程式	153
開啟 IBM SPSS Modeler Advantage 中的串流	153
匯入並匯出 PMML 模型	154

支援 PMML 的模型類型 . . . . .	154
<b>第 12 章 專案和報告 . . . . .</b>	<b>157</b>
專案簡介 . . . . .	157
CRISP-DM 視圖 . . . . .	157
「類別」視圖 . . . . .	157
建立專案 . . . . .	158
建立新專案 . . . . .	158
新增到專案 . . . . .	158
將專案轉移至 IBM SPSS Collaboration and Deployment Services 儲存庫 . . . . .	159
設定專案內容 . . . . .	159
註解專案 . . . . .	160
物件內容 . . . . .	160
關閉專案 . . . . .	161
產生報告 . . . . .	161
儲存和匯出產生的報告 . . . . .	162
<b>第 13 章 自訂 IBM SPSS Modeler . . . . .</b>	<b>165</b>
自訂 IBM SPSS Modeler 選項 . . . . .	165
設定 IBM SPSS Modeler 選項 . . . . .	165
系統選項 . . . . .	165
設定預設目錄 . . . . .	166
設定使用者選項 . . . . .	166
自訂節點選用區 . . . . .	173
自訂選用區管理員 . . . . .	174
變更選用區標籤視圖 . . . . .	176
<b>第 14 章 串流和節點的效能考量 . . . . .</b>	<b>177</b>
節點的順序 . . . . .	177

節點快取 . . . . .	178
效能：處理節點 . . . . .	178
效能：建模節點 . . . . .	179
效能：CLEM 表示式 . . . . .	180

## 第 15 章 IBM SPSS Modeler 中的協助工具 . . . . . 181

IBM SPSS Modeler 協助工具概觀 . . . . .	181
協助工具支援的類型 . . . . .	181
針對視力受損用戶的協助工具 . . . . .	181
盲人使用者的協助工具 . . . . .	182
鍵盤協助工具 . . . . .	182
使用螢幕閱讀器 . . . . .	188
使用提示 . . . . .	189
與其他軟體的衝突 . . . . .	190
JAWS 和 Java . . . . .	190
在 IBM SPSS Modeler 中使用圖形 . . . . .	190

## 第 16 章 Unicode 支援 . . . . . 191

IBM SPSS Modeler 中的 Unicode 支援 . . . . .	191
--	-----

## 注意事項 . . . . . 193

商標 . . . . .	194
產品說明文件的條款 . . . . .	194

## 索引 . . . . . 197



---

## 第 1 章 關於 IBM SPSS Modeler

IBM® SPSS® Modeler 是一組資料採礦工具，通過這些工具可以採用商業技術快速建立預測性模型，並將其應用於商業活動，從而改進決策過程。IBM SPSS Modeler 參照行業標準 CRISP-DM 模型設計而成，可支援從資料到更優商業成果的整個資料採礦過程。

IBM SPSS Modeler 提供擷取自機器學習人工智慧以及統計資料的各種建模方法。「建模」選用區上提供的方法可讓您根據資料衍生新資訊，以及開發預測模型。每種方法都具有特定的強度且最適合因應特定類型的問題。

SPSS Modeler 可以作為單獨產品購買，也可以作為用戶端與 SPSS Modeler Server 一起使用。同時提供了大量其他選項，下列各節將對這些選項進行概述。有關進一步資訊，請參閱<https://www.ibm.com/analytics/us/en/technology/spss/>。

---

### IBM SPSS Modeler 產品

IBM SPSS Modeler 系列產品及關聯的軟體包括下列各項。

- IBM SPSS Modeler
- IBM SPSS Modeler Server
- IBM SPSS Modeler Administration Console (包含在 IBM SPSS Deployment Manager 中)
- IBM SPSS Modeler Batch
- IBM SPSS Modeler Solution Publisher
- IBM SPSS Collaboration and Deployment Services 的 IBM SPSS Modeler Server 配接器

### IBM SPSS Modeler

SPSS Modeler 是具有完整功能的產品，它安裝並執行於個人電腦上。您可以在本端方式作為單獨產品執行 SPSS Modeler，也可以在分佈方式下將其與 IBM SPSS Modeler Server 一起使用來提高大型資料集的效能。

借助 SPSS Modeler，您可以快速直接地建立準確的預測模型，而不進行程式設計。通過使用唯一可視介面，您可以輕鬆地查看資料採礦過程。借助該產品隨附的進階分析支援，您可以探索資料中先前隱藏的型樣和趨勢。您可以構建結果模型並瞭解影響結果的因素，從而利用業務機會並降低風險。

SPSS Modeler 推出了兩個版本：SPSS Modeler Professional 和 SPSS Modeler Premium。請參閱第 2 頁的『IBM SPSS Modeler 版本』主題，以取得更多資訊。

### IBM SPSS Modeler Server

SPSS Modeler 使用用戶端/伺服器架構將資源集約型作業的要求分發給功能強大的伺服器軟體，因而使大資料集的傳輸速度大大加快。

SPSS Modeler Server 是一個個別授權的產品，在分佈分析方式下，該產品在安裝了一個或多個 IBM SPSS Modeler 的伺服器主機上持續執行。這種運行方式大大提高了 SPSS Modeler Server 對大型資料集的處理速度，因為在伺服器上可以運行耗用記憶體體的作業，並且無需將資料下載到用戶端電腦上。IBM SPSS Modeler Server 還提供對 SQL 最佳化和資料庫內建模功能的支援，從而在效能和自動化方面帶來更多優勢。

## IBM SPSS Modeler Administration Console

Modeler Administration Console 是一個圖表使用者介面，用於管理多個 SPSS Modeler Server 配置選項，這些選項還可以通過選項檔案進行配置。主控台包含在 IBM SPSS Deployment Manager，可以用於監視和配置 SPSS Modeler Server 安裝，並且可供目前 SPSS Modeler Server 客戶免費使用。應用程式僅可以在 Windows 電腦上安裝；但它可以管理在任何受支援平台上安裝的伺服器。

## IBM SPSS Modeler Batch

資料採礦通常是交互過程，因此，還可以從指令行執行 SPSS Modeler 而不需要圖形使用者介面。例如，您可能具有長時間執行或重複作業，並且希望在使用者不進行人為介入的情況下執行這些作業。SPSS Modeler Batch 是該產品的一個特殊版本，可提供對 SPSS Modeler 完整分析性能的支援，而無需存取一般的使用者介面。要使用 SPSS Modeler Batch，需要 SPSS Modeler Server。

## IBM SPSS Modeler Solution Publisher

SPSS Modeler Solution Publisher 是一個支持建立 SPSS Modeler 串流的打包版本的工具，該版本的串流可以由外部執行時期引擎執行或內含到外部應用程式中。通過這種方式，您可以發行和部署完整的 SPSS Modeler 串流以用於未安裝 SPSS Modeler 的環境。SPSS Modeler Solution Publisher 作為 IBM SPSS Collaboration and Deployment Services - 評分 服務的組成部分分發，需要個別的授權。通過此授權，您可以接收 SPSS Modeler Solution Publisher Runtime，它使您能夠執行已發佈的串流。

有關 SPSS Modeler Solution Publisher 的進一步資訊，請參閱 IBM SPSS Collaboration and Deployment Services 文件。IBM SPSS Collaboration and Deployment Services Knowledge Center 包含名為 "IBM SPSS Modeler Solution Publisher" 和 "IBM SPSS Analytics Toolkit" 的部分。

## IBM SPSS Collaboration and Deployment Services 的 IBM SPSS Modeler Server 配接器

IBM SPSS Collaboration and Deployment Services 的一些配接器使 SPSS Modeler 和 SPSS Modeler Server 能夠與 IBM SPSS Collaboration and Deployment Services 儲存庫進行交互。通過這種方式，部署到儲存庫的 SPSS Modeler 串流可以由多個使用者共用，或者從瘦用戶端應用程式 IBM SPSS Modeler Advantage 進行存取。請將配接器安裝在管理儲存庫的系統上。

---

## IBM SPSS Modeler 版本

SPSS Modeler 推出了下列版本。

### SPSS Modeler Professional

SPSS Modeler Professional 提供處理大多數類型的結構化資料所需要的所有工具，例如 CRM 系統中追蹤的行為和互動、個人背景資訊、採購行為和銷售資料。

### SPSS Modeler Premium

SPSS Modeler Premium 是一個個別授權的產品，它對 SPSS Modeler Professional 進行了延伸，以便後者能夠處理專門的資料和非結構化文字資料。SPSS Modeler Premium 包含 IBM SPSS Modeler Text Analytics：

**IBM SPSS Modeler Text Analytics** 採用了先進語言技術和自然語言正在處理 (NLP)，以快速正在處理大量非結構化文字資料，擷取和群組織關鍵概念，以及將這些概念分為各式各樣的種類。擷取的概念和種類可以和現有結構化資料中進行已結合（例如人口統計學），並且可用於借助 IBM SPSS Modeler 的一整套資料採礦工具來進行建模，以此實現更好更集中的決策。



## IBM SPSS Modeler Subscription

IBM SPSS Modeler Subscription 提供的所有預測分析功能與傳統 IBM SPSS Modeler 用戶端提供的相同。使用訂閱版，您可以定期下載產品更新項目。

---

### 說明文件

文件可以從 SPSS Modeler 中的「說明」功能表獲取。這樣會開啟可在產品外部公開存取的 Knowledge Center。

作為產品下載的一部分，還會在個別的壓縮資料夾中以 PDF 格式提供每個產品的完整文件（包括安裝指示）。也可以從 Web 下載 PDF 文件，地址為：<http://www.ibm.com/support/docview.wss?uid=swg27046871>。

### SPSS Modeler Professional 文件

SPSS Modeler Professional 文件套組（安裝指示除外）如下。

- **IBM SPSS Modeler 使用者手冊**。使用 SPSS Modeler 的一般簡介，包括如何建置資料串流、處理遺漏值、建置 CLEM 表示式，處理專案和報告，將用於部署的串流打包到 IBM SPSS Collaboration and Deployment Services 或 IBM SPSS Modeler Advantage。
- **IBM SPSS Modeler Source、Process 和 Output 節點**。說明用於以不同格式讀取、處理和輸出資料的所有節點。實際上這表示所有節點而非建模節點。
- **IBM SPSS Modeler Modeling 節點**。說明所有用於建立資料採礦模型的節點。IBM SPSS Modeler 提供擷取自機器學習人工智慧以及統計資料的各種建模方法。
- **IBM SPSS Modeler 應用程式手冊**。本手冊中的範例旨在為具體的建模方法和技術提供具有針對性的簡介。還可以在「說明」功能表中查閱本手冊的線上版本。請參閱第 4 頁的『應用程式範例』主題，以取得更多資訊。
- **IBM SPSS Modeler Python Scripting 和自動化**。通過編寫 Python Script 實現系統自動化的相關資訊，其中包含可以用於操作節點和串流的內容的資訊。
- **IBM SPSS Modeler 部署手冊**。有關在 IBM SPSS Deployment Manager 下以正在處理工作的步驟形式執行 IBM SPSS Modeler 串流的資訊。
- **IBM SPSS Modeler CLEF 開發者手冊**。CLEF 提供了將第三方程式（例如，資料處理常式或建模演算法）作為節點整合到 IBM SPSS Modeler 的功能。
- **IBM SPSS Modeler 資料庫內挖掘手冊**。有關如何利用資料庫的功能通過第三方演算法來改進效能並增強分析功能的資訊。
- **IBM SPSS Modeler Server 管理和效能手冊**。提供有關如何配置和管理 IBM SPSS Modeler Server 的資訊。
- **IBM SPSS Deployment Manager 使用手冊**。有關使用 Deployment Manager 應用程式中包含的管理主控台使用者介面來監視和配置 IBM SPSS Modeler Server 的資訊。
- **IBM SPSS Modeler CRISP-DM 手冊**。借助 CRISP-DM 方法進行 SPSS Modeler 資料採礦的分步手冊。
- **IBM SPSS Modeler Batch 使用者手冊**。提供在批次模式下使用 IBM SPSS Modeler 的完整指導，包含批次模式執行和指令行引數的詳細資料。本手冊僅以 PDF 格式提供。

### SPSS Modeler Premium 文件

SPSS Modeler Premium 文件套組（安裝指示除外）如下。

- **SPSS Modeler Text Analytics 使用者手冊**。提供有關將文字分析與 SPSS Modeler 配合使用的資訊，包括文字採集節點、互動式工作台、範本和其他資源。

---

## 應用程式範例

SPSS Modeler 中的資料採礦工具可以說明解決很多業務和組織問題，應用程式範例將提供有關特定建模方法和技術的簡要的針對性說明。此處使用的資料集比某些資料挖掘器管理的大量資料儲存庫小得多，但涉及的概念和方法可擴展到實際應用程式。

要存取範例，請在 SPSS Modeler 中按一下「說明」功能表中的**應用程式範例**。

資料檔案和樣本串流安裝在產品安裝目錄下的 Demos 資料夾中。如需相關資訊，請參閱『Demos 資料夾』。

**資料庫建模範例**。請參閱 *IBM SPSS Modeler 資料庫內挖掘手冊* 中的範例。

**Scripting 範例**。請參閱 *IBM SPSS Modeler Script 編寫和自動化手冊* 中的範例。

---

## Demos 資料夾

與應用程式範例一起使用的資料檔案和樣本串流安裝在產品安裝目錄下的 Demos 資料夾中（例如：C:\Program Files\IBM\SPSS\Modeler\\Demos）。可以從 Windows「開始」功能表上的 IBM SPSS Modeler 程式群組存取此資料夾，也可以通過按一下**檔案 > 開啟串流對話框**中最近的目錄的清單上的 Demos 來進行存取。

---

## 授權追蹤

當您使用 SPSS Modeler 時，系統會定期追蹤並記錄授權使用情況。所記錄的授權度量值為 *AUTHORIZED\_USER* 和 *CONCURRENT\_USER*，並且記錄的度量值類型取決於您針對 SPSS Modeler 具有的授權類型。

產生的日誌檔可由 IBM License Metric Tool 處理，通過該工具可產生授權使用情形報告。

授權日誌檔建立在記錄 SPSS Modeler 用戶端日誌檔的目錄（依預設為 %ALLUSERSPROFILE%/IBM/SPSS/Modeler/<version>/log）中。

---

## 第 2 章 IBM SPSS Modeler 18.1.1 新增功能

IBM SPSS Modeler 在此次發行版中增加了下列功能。

- **E-Plot (測試版) 節點**。「節點」選用區的「圖形」標籤上提供新的測試版層次 E-Plot 節點。它使用直觀、現代且自訂程度極高的全新圖形介面，並且資料圖表是互動式的。使用這個新節點，可利用新的圖形功能。
- **Spark 節點**。節點選用區上的新 Spark 標籤提供了用於使用 Python 演算法的節點。這些新節點在 Windows 64 和 Mac 上支援。
  - **Isotonic-AS 節點**。新的 Isotonic-AS 節點在新的 Spark 標籤上提供使用。
  - **XGBoost-AS 節點**。新的 XGBoost-AS 節點在新的 Spark 標籤上提供使用。
  - **K-Means-AS 節點**。新的 K-Means-AS 節點在新的 Spark 標籤上提供使用。
- **超參數最佳化 (基於 Rbfopt)**。「一類 SVM」節點 (「專家」標籤)、「XGBoost 線性」節點 (「建置選項」標籤) 和「XGBoost 樹狀結構」節點 (「建置選項」標籤) 中加入了一個新選項。新的「超參數最佳化」選項會自動探索參數最佳組合，以便模型可對樣本達到預期或較小錯誤率。
- **隨機森林節點**。新的「隨機森林」節點在 Python 標籤上提供使用。
- **t-SNE 節點**。新的 t-Distributed Stochastic Neighbor Embedding (t-SNE) 節點在 Python 標籤及「圖表」標籤上提供使用。
- **CPLEX Optimization 節點的多資料來源**。現在，Optimization 專家可以將資料從多資料來源匯入至 CPLEX Optimization 節點，並將每一個資料來源配置給元組。



## 第 3 章 產品概觀

### 新手啟動使用

作為一種資料採礦應用程式，IBM SPSS Modeler 提供了用以尋找大資料集中有用關係的策略性方法。與更傳統的統計方法相比，您在開始時不必知道您要尋找什麼。您可以通過配適不同的模型和研究不同的關係來探索您的資料，直到發現有用的資訊。

### 啟動 IBM SPSS Modeler

要啟動此應用程式，請按一下：

開始 > [所有] 程式 > IBM SPSS Modeler <版本> > IBM SPSS Modeler <版本>

主視窗將在幾秒鐘後顯示。

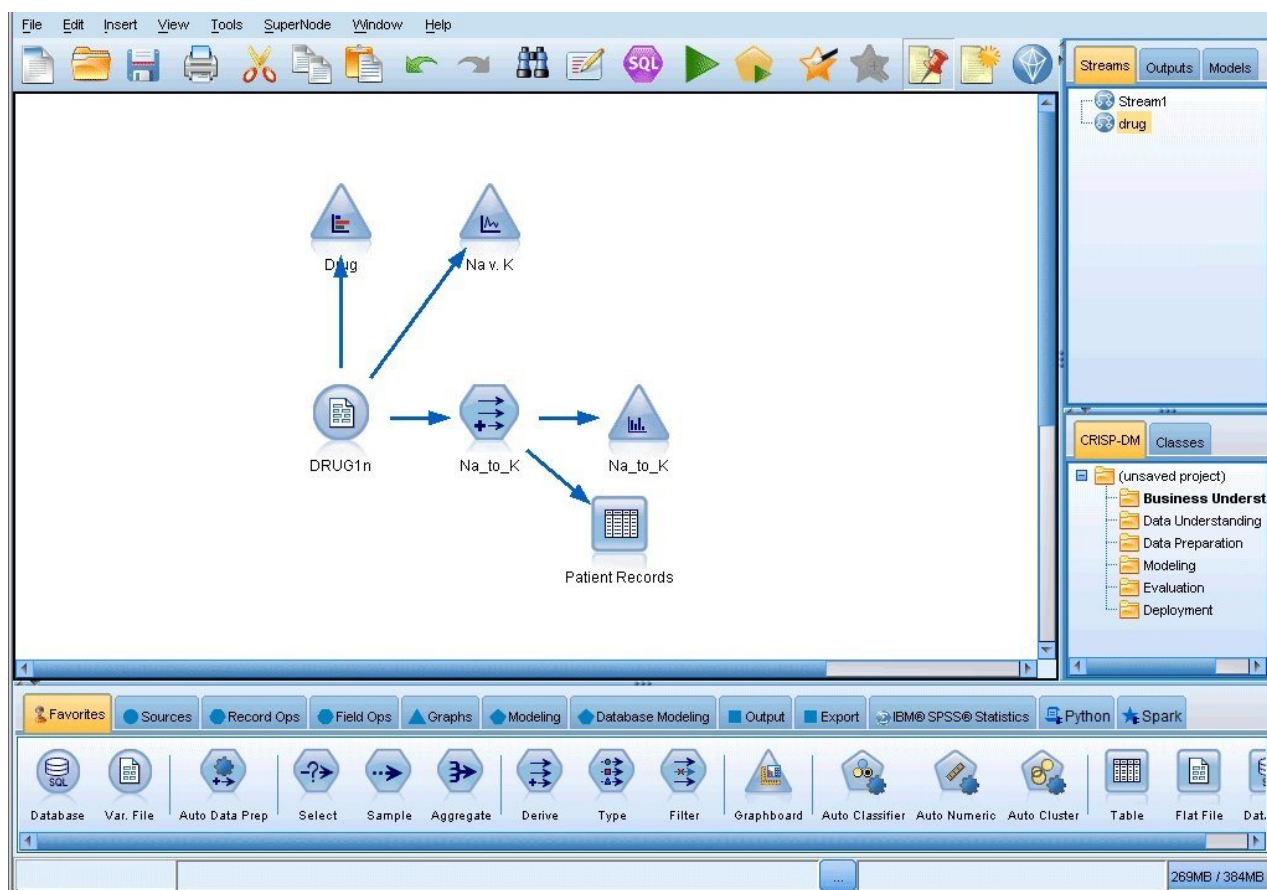


圖 1. IBM SPSS Modeler 主應用程式視窗

## 從指令行啟動

您可以使用作業系統的指令行來如下啟動 IBM SPSS Modeler：

1. 在安裝了 IBM SPSS Modeler 的電腦上，開啟 DOS 或指令提示符視窗。
2. 要以互動模式啟動 IBM SPSS Modeler 介面，請輸入 `modelerclient` 指令，然後輸入所需的引數；例如：  
`modelerclient -stream report.str -execute`

可用引數（旗標）容許您連接至伺服器、載入串流、執行 Script 或根據需要指定其他引數。

## 連線到 IBM SPSS Modeler Server

IBM SPSS Modeler 可作為單獨的應用程式執行，或作為直接連接至 IBM SPSS Modeler Server 的用戶端執行，或者作為通過處理程序協調器 (COP) 外掛程式從 IBM SPSS Collaboration and Deployment Services 連接至 IBM SPSS Modeler Server 或伺服器叢集的用戶端執行。現行連線狀態顯示在 IBM SPSS Modeler 視窗的左下角。

無論何時想連接至伺服器，都請手動輸入想要連接的伺服器的名稱或選取之前已定義的名稱。但是，如果您擁有 IBM SPSS Collaboration and Deployment Services，則可以從「伺服器登入」對話框搜尋伺服器清單或伺服器叢集清單。可以通過處理程序協調器執行瀏覽網路上執行的 Statistics 服務的功能。

### 連接至伺服器

1. 在「工具」功能表上，按一下**伺服器登入**。會開啟「伺服器登入」對話框。或者按兩下 IBM SPSS Modeler 視窗的連線狀態區域。
2. 使用該對話框指定要連接至本機伺服器電腦的選項或從表格中選取連線。
  - 按一下**新增或編輯**以新增或編輯連線。請參閱第 9 頁的『新增並編輯 IBM SPSS Modeler Server 連線』主題，以取得更多資訊。
  - 按一下**搜尋**以存取處理程序協調器中的伺服器或伺服器叢集。請參閱第 9 頁的『搜尋 IBM SPSS Collaboration and Deployment Services 中的伺服器』主題，以取得更多資訊。

**伺服器表格。**此表格包含已定義的伺服器連線集。該表格顯示預設連線、伺服器名稱、說明和埠號。您可以手動新增新的連線，以及選取或搜尋現有連線。要將特定的伺服器設定為預設連線，請在表格中「預設」欄中為此連線選取勾選框。

**預設資料路徑。**指定用於伺服器電腦上的資料的路徑。按一下省略號按鈕 (...)，以瀏覽至所需要的位置。

**設定認證。**不已勾選此複選框可啟用單一登入功能，該功能試圖使您使用本端電腦使用者名稱和密碼詳細資料登入伺服器。如果無法使用單一登入，或您勾選此方框以取消單一登入（例如，登入管理者帳戶），則啟用下列欄位讓您輸入您的認證。

**使用者 ID。**輸入用於登入伺服器的使用者名稱。

**密碼。**輸入與指定使用者名稱相關聯的密碼。

**網域。**指定用於登入伺服器的網域。只有伺服器電腦與用戶電腦處於不同的 Windows 網域時，才需要網域名稱。

3. 按一下**確定**以完成此連線。

切斷與伺服器的連接

1. 在「工具」功能表上，按一下**伺服器登入**。會開啟「伺服器登入」對話框。或者按兩下 IBM SPSS Modeler 視窗的連線狀態區域。
2. 在此對話框中，選取「本機伺服器」，然後按一下**確定**。

## 新增並編輯 IBM SPSS Modeler Server 連線

您可以在「伺服器登入」對話框中手動編輯或新增伺服器連線。按一下「新增」可以存取空白的「新增/編輯伺服器」對話框，在此對話框中可以輸入伺服器連線的詳細資料。在「伺服器登入」對話框中選取現有連線並按一下「編輯」，將開啟「新增/編輯伺服器」對話框，其中包含所選連線的詳細資料，以便可以進行任何變更。

註：不能編輯從 IBM SPSS Collaboration and Deployment Services 中新增加的伺服器連線，因為名稱、埠及其他詳細資料已在 IBM SPSS Collaboration and Deployment Services 中做過定義。最佳實踐指出，應該使用相同的埠與 IBM SPSS Collaboration and Deployment Services 和 SPSS Modeler Client 進行通訊。這些埠可以設定為 options.cfg 檔案中的 max\_server\_port 和 min\_server\_port。

### 新增伺服器連線

1. 在「工具」功能表上，按一下**伺服器登入**。會開啟「伺服器登入」對話框。
2. 在此對話框中，按一下**新增**。即會開啟「伺服器登入新增/編輯伺服器」對話框。
3. 輸入伺服器連線的詳細資料，然後按一下**確定**儲存此連線並傳回「伺服器登入」對話框。
  - **伺服器**。指定可用伺服器或從清單選取一個伺服器。伺服器電腦的名稱可以使用英數（例如 *myserver*）或指派給伺服器電腦的 IP 位址（例如，202.123.456.78）。
  - **埠**。指定伺服器要接聽的埠號。如果預設設置不可用，請向系統管理者索取正確的埠號。
  - **說明**。輸入此伺服器連線的選用說明。
  - **確保連線安全(使用 SSL)**。指定是否應該使用 SSL（安全的套接字層）連線。SSL 是一個一般通訊協定，用於確保通過網路傳送的資料的安全的。要使用此功能，必須在承載 IBM SPSS Modeler Server 的伺服器中啟用 SSL。必要時請與本端管理者聯絡，以瞭解詳細資料。

### 編輯伺服器連線

1. 在「工具」功能表上，按一下**伺服器登入**。會開啟「伺服器登入」對話框。
2. 在此對話框中，選取希望編輯的連線，然後按一下**編輯**。即會開啟「伺服器登入新增/編輯伺服器」對話框。
3. 變更伺服器連線詳細資料，然後按一下**確認**儲存變更內容並傳回至「伺服器登入」對話框。

## 搜尋 IBM SPSS Collaboration and Deployment Services 中的伺服器

在 IBM SPSS Collaboration and Deployment Services 中，可以使用處理程序協調器選取網路上可用的伺服器或伺服器叢集，從而代替手動輸入伺服器連線。伺服器叢集是一組伺服器，處理程序協調器從這群組伺服器中確定最適合對正在處理要求作出回應的伺服器。

雖然您可在「伺服器登入」對話框中手動新增伺服器，但搜尋可用伺服器能讓您不需要知道正確的伺服器名稱和埠號，即可連線至伺服器。此資訊會自動提供。但仍需輸入正確的登入資訊，如使用者名稱、網域和密碼。

註：如果您無權存取處理程序協調器功能，那麼仍然可以手動輸入要連接的伺服器名稱或選取先前已定義的名稱。請參閱『新增並編輯 IBM SPSS Modeler Server 連線』主題，以取得更多資訊。

### 搜尋伺服器和伺服器叢集

1. 在「工具」功能表上，按一下**伺服器登入**。會開啟「伺服器登入」對話框。

2. 在此對話框中，按一下**搜尋**開啟「搜尋伺服器」對話框。如果在試圖瀏覽處理程序協調器時未登入到 IBM SPSS Collaboration and Deployment Services，則系統會提示您執行此項操作。
3. 從清單中選取伺服器或伺服器叢集。
4. 按一下**確定**以關閉對話框，然後將此連線新增到「伺服器登入」對話框的表格中。

## 連線到 Analytic Server

如果有多個 Analytic Server 可用，可以使用「Analytic Server 連線」對話框來定義多個伺服器以在 IBM SPSS Modeler 中使用。您的管理者可能已經在 <Modeler\_install\_path>/config/options.cfg 檔案中設定預設 Analytic Server。但在定義之後，也可以使用其他可用伺服器。例如，使用 Analytic Server「來源」及「匯出」節點時，您可能想要在串流的不同分支中使用不同的 Analytic Server 連線，因此每一個分支執行時，它會使用自己的 Analytic Server，且不會將任何資料拉取到 IBM SPSS Modeler Server。請注意，如果分支包含多個 Analytic Server 連線，則會從 Analytic Server 拉取到 IBM SPSS Modeler Server。如需相關資訊（包括限制），請參閱第 41 頁的『Analytic Server 串流內容』。

要建立新的 Analytic Server 連線，轉至**工具 > Analytic Server 連線**並在對話框的下列部分中提供所需資訊。

### 連線

**URL**。以格式 `https://hostname:port/contextroot` 輸入 Analytic Server 的 URL，其中，hostname 是 Analytic Server 的 IP 位址或主機名稱，port 是其埠號，contextroot 是 Analytic Server 的環境定義根目錄。

**租戶**。輸入 IBM SPSS Modeler Server 所屬的租戶的名稱。如果不知道租戶，請聯絡您的管理者。

### 鑑別

**眾數**。從下列鑑別方式中進行選取。

- **使用者名稱和密碼**要求您輸入使用者名稱和密碼。
- **儲存的認證**要求您從 IBM SPSS Collaboration and Deployment Services 儲存庫 中選取認證。
- **Kerberos** 要求您輸入服務主體名稱和配置檔案路徑。如果不知道該資訊，請聯絡您的管理者。

**使用者名稱**。輸入 Analytic Server 使用者名稱。

**密碼**。輸入 Analytic Server 密碼。

**連接**。按一下**連線**以測試新的連線。

### 連線

在指定上述資訊並按一下**連線**之後，連線將新增到該「連線」表格。如果需要移除連線，請將其選取並按一下**移除**。

如果管理者在 options.cfg 檔案中定義了預設 Analytic Server 連線，也可以按一下**新增預設連線**以將其新增到可用連線中。將提示您輸入使用者名稱和密碼。

## 變更 temp 目錄

IBM SPSS Modeler Server 執行的某些作業可能需要建立暫存檔。依預設，IBM SPSS Modeler 在系統暫存目錄下建立暫時檔案。可通過下列步驟更改暫存目錄的位置。

1. 建立新目錄 `spss` 及其子目錄 `servertemp`。



2. 編輯 *options.cfg*，該文件位於 IBM SPSS Modeler 安裝目錄的 */config* 目錄下。在此檔案中編輯 *temp\_directory* 參數，將其更改為：*temp\_directory, "C:/spss/servtemp"*。
3. 完成此操作後，必須重新啟動 IBM SPSS Modeler Server 服務。可通過按一下 Windows 控制台中的服務標籤進行此服務重啟操作。只需停止該服務然後將其重新啟動即可啟動所作的變更。重新啟動機器也會重新啟動該服務。

所有臨時檔案此時將寫入該新目錄。

註：

- 必須使用正斜線。
- 當通過 IBM SPSS Collaboration and Deployment Services 工作執行評估串流時，不適用 *temp\_directory* 設定。當您執行這類工作時，會建立暫存檔。依預設，檔案會儲存至 IBM SPSS Modeler Server 的安裝目錄。當您在 IBM SPSS Modeler 中建立 IBM SPSS Modeler Server 連線時，可以變更儲存暫存檔的預設資料夾。

## 啟動多個 IBM SPSS Modeler 階段作業

如果需要同時啟動一個以上的 IBM SPSS Modeler 階段作業，則必須對 IBM SPSS Modeler 和 Windows 的設定做一些變更。例如，如果您有兩個獨立的伺服器授權，並且希望從同一台用戶端針對兩台不同的伺服器執行兩個串流，則需要對上述設置做一些更改。

要啟用多個 IBM SPSS Modeler 階段作業：

1. 按一下：

開始 > [所有] 程式 > **IBM SPSS Modeler**

2. 在 IBM SPSS Modeler 快捷鍵（帶箭頭的圖示）上用滑鼠右鍵按一下並選取內容。
3. 在目標文字框中，將 *-noshare* 新增到該字串的到結束。
4. 在 Windows 檔案總管中選取：

工具 > 資料夾選項...

5. 在「檔案類型」標籤上選取「IBM SPSS Modeler 串流」選項，然後按一下進階。
6. 在「編輯檔案類型」對話框中，選取「使用 IBM SPSS Modeler 開啟」，然後按一下編輯。
7. 在用於執行動作的應用程式文字框中，在 **-stream** 參數前新增 *-noshare*。

---

## IBM SPSS Modeler 介面概覽

在資料採礦程序中的每個階段，易於使用的 IBM SPSS Modeler 介面都會邀請您的特定商業專門知識。建模演算法（如預測、分類、分區段和關聯偵測）可確保得到強大而準確的模型。模型結果可以方便地部署和讀取入到資料庫、IBM SPSS Statistics 和各式各樣的其他應用程式中。

使用 IBM SPSS Modeler 是一個處理資料的三步驟程序。

- 首先，將資料讀取入 IBM SPSS Modeler。
- 接著，通過一系列處理來執行資料。
- 最後，將資料傳送至目的地。

這一作業序列稱為資料串流，因為資料以逐條記錄形式流動，從來源開始，流經每個操作，最終到達目的地（模型或某種資料輸出）。



圖 2. 簡單串流

## IBM SPSS Modeler 串流畫布

串流畫布域圖是 IBM SPSS Modeler 視窗的最大區域，也是您建立和操作資料串流的位置。

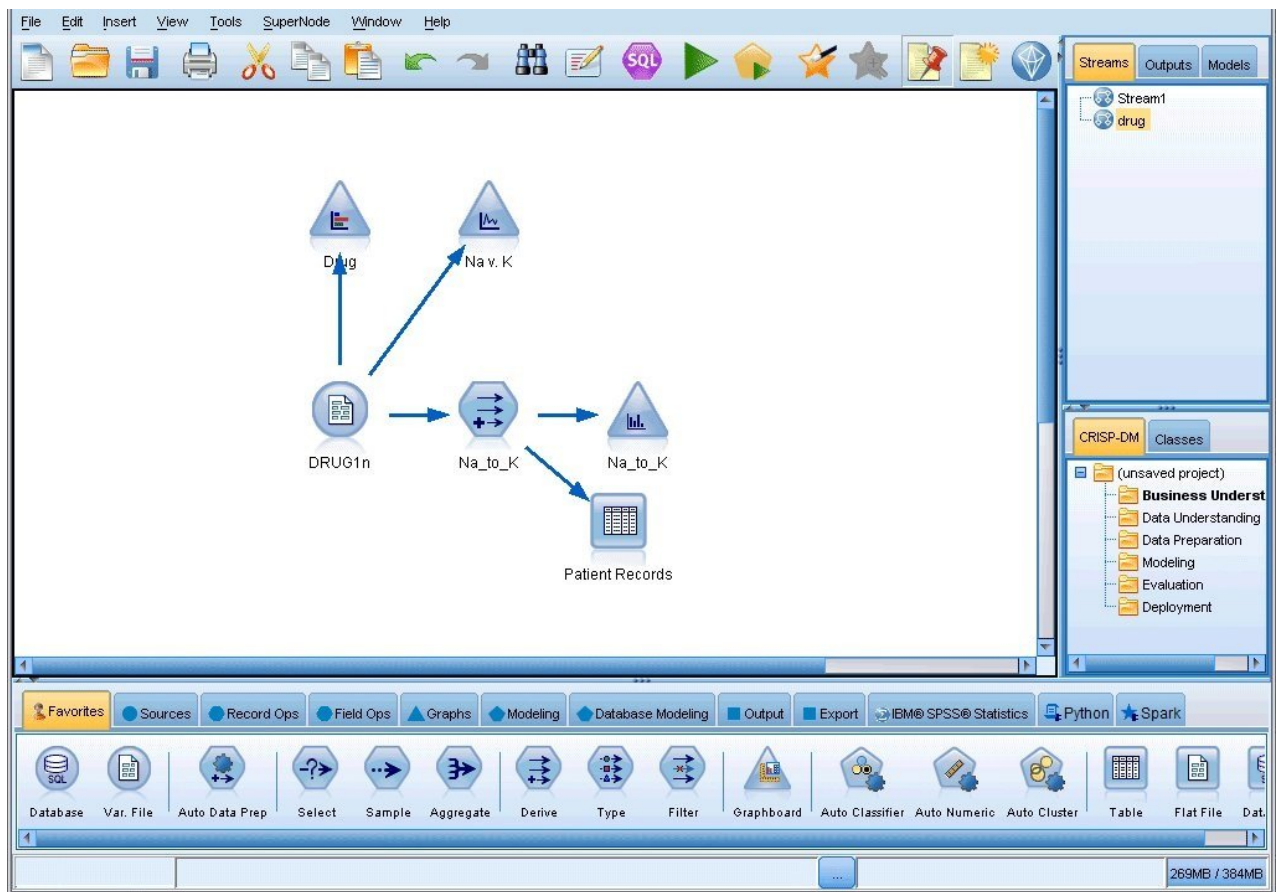


圖 3. IBM SPSS Modeler 畫布 (預設視圖)

串流是在介面的主畫布中通過繪製與業務相關的資料作業圖來建立的。每個作業都用一個圖示或節點代表，這些節點在串流中鏈結在一起，串流代表資料在各個作業中的流動。

在 IBM SPSS Modeler 中，可以在同一串流畫布或通過開啟新的串流畫布來一次處理多個串流。階段作業期間，串流儲存在 IBM SPSS Modeler 視窗右上角的「串流」管理器中。

註：如果使用啟用了內建觸控板的 **Force Click and haptic feedback** 設定的 MacBook，那麼從節點選用區將節點拖拽到串流畫布可能會導致向畫布新增複製節點。為避免發生此問題，建議取消 **Force Click and haptic feedback** 觸控板系統喜好設定。

## 節點選用區

SPSS Modeler 中的大部分資料和建模工具都可從橫跨串流畫布下方視窗的底部的節點選用區中獲取。

例如，記錄處理選用區標籤包含可用於對資料記錄執行作業（如選取、合併和附加）的節點。

要向畫布新增節點，請按兩下節點選用區中的圖示或將節點拖至畫布上。隨後可將各個圖標連接以建立一個代表資料流程的串流。

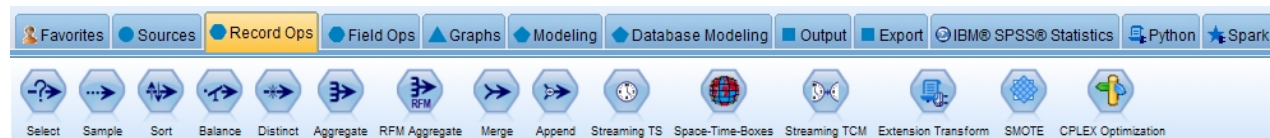


圖 4. 節點選用區中的「記錄處理」標籤

每個選用區標籤均包含一組不同的串流作業階段中使用的相關節點，如：

- 來源節點用於將資料引入到 SPSS Modeler 中。
- 記錄處理節點用於對資料記錄執行作業，例如選取、合併和附加。
- 欄位作業節點用於對資料欄位執行作業，例如過濾、衍生新欄位和確定給定欄位的測量層次。
- 圖形節點用於以圖形方式顯示建模前後的資料。圖形包含統計圖、直方圖、網絡節點和評估表。
- 建模節點使用 SPSS Modeler 中提供的建模演算法，例如神經網路、決策樹、叢集作業演算法和資料排序等。
- 資料庫建模節點使用 Microsoft SQL Server、IBM Db2 以及 Oracle 和 Netezza 資料庫中提供的建模演算法。
- 輸出節點用於為資料、圖表和模型結果生成可以在 SPSS Modeler 中檢視的各種輸出。
- 匯出節點用於生成可以在外部應用程式（例如，IBM SPSS Data Collection 或 Excel）中檢視的各種輸出。
- **IBM SPSS Statistics** 節點從 IBM SPSS Statistics 匯入資料或向其匯出資料，以及執行 IBM SPSS Statistics 程序。
- **Python** 節點可用於執行 Python 演算法。
- **Spark** 節點可用於執行 Spark 演算法。

隨著對 SPSS Modeler 的熟悉，您也可以自訂供自己使用的選用區內容。

在「節點選用區」的左端，您可以通過選取受監督、關聯或分區段對所顯示的節點進行過濾。

「節點選用區」下方是一個報告窗格，此窗格提供對各種作業進展的意見回饋，例如何時將資料讀入資料串流中。「節點選用區」下方還有一個狀態窗格，此窗格提供有關應用程式目前正在執行的操作的資訊以及何時需要使用者意見的指示資訊。

註：如果使用啟用了內建觸控板的 **Force Click and haptic feedback** 設定的 MacBook，那麼從節點選用區將節點拖拽到串流畫布可能會導致向畫布新增複製節點。為避免發生此問題，建議取消 **Force Click and haptic feedback** 觸控板系統喜好設定。

## IBM SPSS Modeler 管理員

管理程式窗格位於視窗右上角。此窗格包含用於管理串流、輸出和模型的三個標籤。

可以使用「串流」標籤開啟、更名、儲存和刪除階段作業中建立的串流。



圖 5. 「串流」標籤



圖 6. 「輸出」標籤

「輸出」標籤中包含由 IBM SPSS Modeler 中的串流作業生成的各類檔案，如圖形和表格。您可以顯示、儲存、更名和關閉此標籤上列出的表格、圖形和報告。



圖 7. 包含模型塊的「模型」標籤

在管理器標籤中，「模型」標籤具有最強大的功能。該標籤中包含所有模型塊，這些模型塊包含針對現行階段作業在 IBM SPSS Modeler 中產生的模型。可以直接從「模型」標籤瀏覽這些模型或者將它們新增到畫布內的串流中。

## IBM SPSS Modeler 專案

視窗右端底部是專案窗格，用於建立和管理資料採礦專案（與資料採礦作業相關的檔案群組）。可以通過兩種方法來檢視您在 IBM SPSS Modeler 中建立的專案：「類別」視圖和 CRISP-DM 視圖。

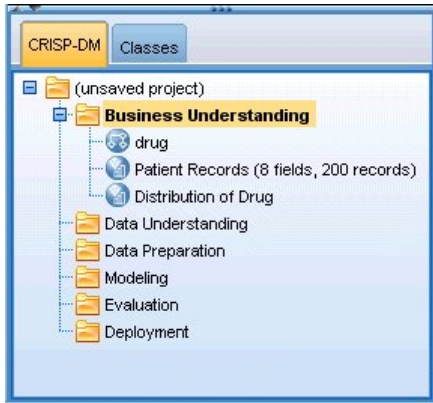


圖 8. CRISP-DM 視圖

依據業內認可的非專利方法「跨行業資料採礦標準程序」，CRISP-DM 標籤提供了一種專案組織方法。無論是有經驗的資料挖掘人員還是新手，使用 CRISP-DM 工具都會使您事半功倍。

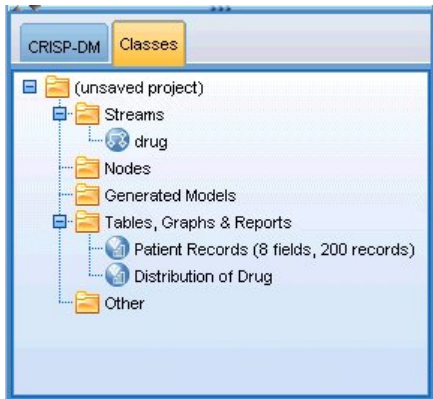


圖 9. 「類別」視圖

「類別」標籤提供了一種在 IBM SPSS Modeler 中按種類（即按照所建立物件的種類）組織您工作的方式。此視圖在盤點資料、串流和模型的庫存時非常有用。

## IBM SPSS Modeler 工具列

IBM SPSS Modeler 視窗頂部有一個圖示工具列，其中包含許多有用功能。下面是一些工具列按鈕及其功能。



建立新串流



開啟串流



儲存串流



列印目前串流



剪下並移到剪貼簿



複製到剪貼簿



貼上選擇範圍



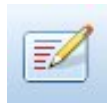
還原上一個動作



重做



搜尋節點



編輯串流內容



預覽 SQL 產生



執行目前串流



執行串流選擇



停止串流（僅在串流處於執行狀態時可用）



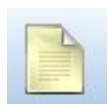
新增受



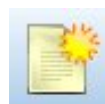
放大（僅限於 SuperNode）



縮小（僅限於 SuperNode）



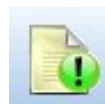
串流中無標示



插入備註



隱藏串流標記（如果有）



顯示隱藏的串流標記



在 IBM SPSS Modeler Advantage 中  
開啟串流

串流標記由串流備註、模型鏈結和評分分支指示組成。

在《IBM SPSS 建模節點》手冊中介紹了模型鏈結。

## 自訂工具列

您可以變更工具列的各個方面，例如：

- 是否顯示
- 圖示是否有可用工具提示
- 使用大或小圖示

要開啟或關閉工具列顯示，請執行以下操作：

1. 在主功能表上，按一下：

**視圖 > 工具列 > 顯示**

要變更工具提示或圖示大小設定，請執行以下操作：

1. 在主功能表上，按一下：

**檢視 > 自訂 > 顯示**

根據需要按一下顯示工具提示或大按鈕。

## 自訂 IBM SPSS Modeler 視窗

使用 SPSS Modeler 介面各部分之間的分界線，可以調整工具的大小或關閉某些工具以符合個人偏好。例如，如果要處理大型串流，那麼可以使用每條分界線上的小箭頭來關閉節點選用區、管理程式窗格和專案窗格。這樣可以最大化串流畫布，從而為處理大型串流或多個串流提供足夠的畫布。

此外，從「檢視」功能表上，按一下節點選用區、管理員或專案可開啟或關閉這些專案的顯示。

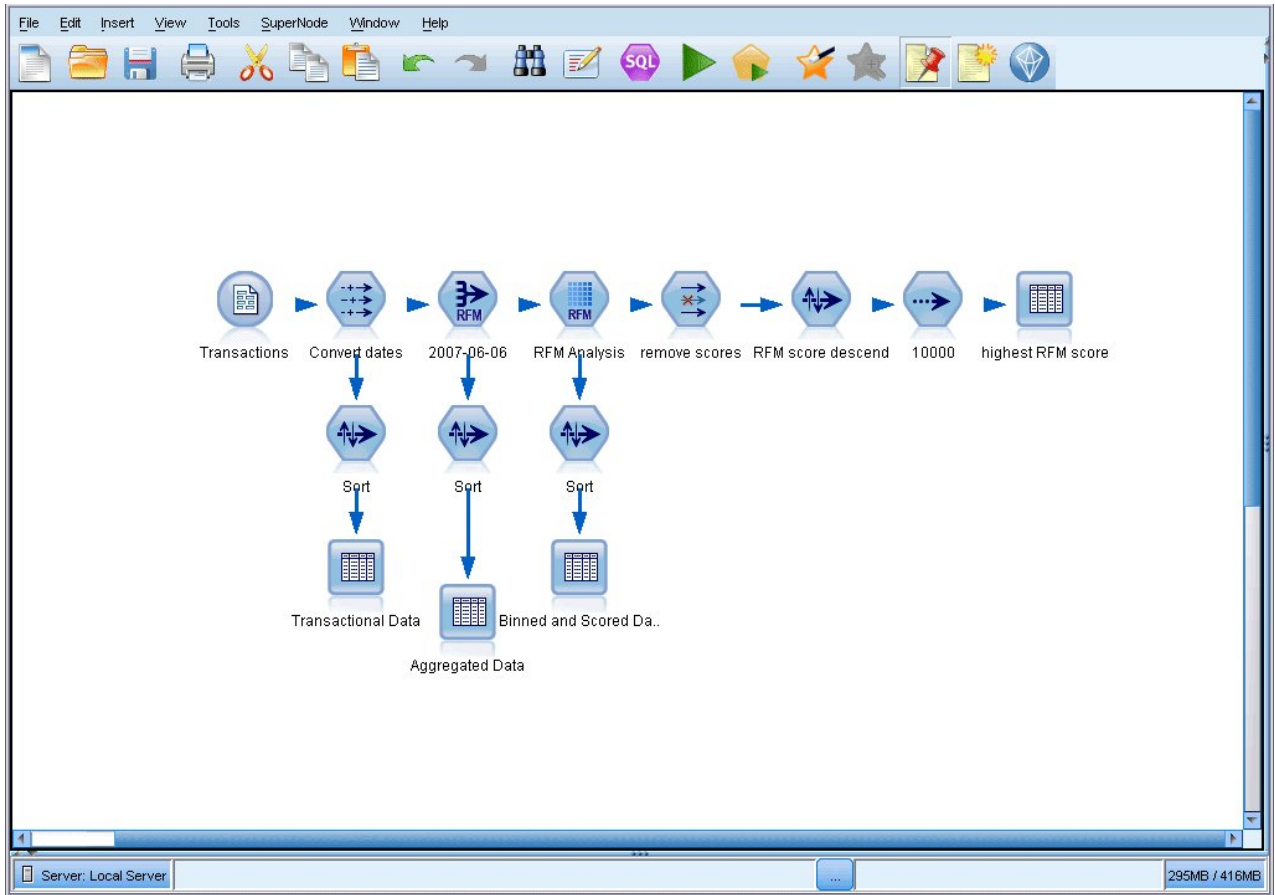


圖 10. 最大化的串流畫布

另外一種關閉節點選用區和管理員以及專案窗格的方法是：垂直或水平移動 SPSS Modeler 視窗端面或底部的捲動條，將串流畫布當作捲動頁使用。

您也可以控制畫面標示的顯示，此標示由串流備註、模型鏈結和評分分支指示組成。要開啟或關閉此顯示，請按一下：

視圖 > 串流標記

### 變更串流的圖示大小

您可以通過下列方式變更串流圖示的大小。

- 串流內容設定
- 串流中的蹦現功能表
- 使用鍵盤

您可以調整整個串流視圖的大小，將其調整為標準圖示大小的 8% 至 200% 之間的某個尺寸。

### 要調整整個串流的大小（串流內容方法）

1. 從主功能表中，選擇

工具 > 串流內容 > 選項 > 佈置。

2. 從「圖示大小」功能表中選擇所需大小。



3. 按一下**套用**以請參閱結果。
4. 按一下**確定**以儲存變更。

### 要調整整個串流的大小（功能表方法）

1. 用滑鼠右鍵按一下畫布上的串流背景。
2. 選擇圖示大小，並選擇所需的大小。

### 要調整整個串流的大小（鍵盤方法）

1. 同時按住主鍵盤上的 **Ctrl + [-]** 來縮小至下一個較小的尺寸。
2. 同時按住主鍵盤上的 **Ctrl + Shift + [+]** 來放大至下一個較大的尺寸。

獲取複合串流的整體視圖時，此功能尤其有用。您還可以使用此功能來最大程度地減少列印串流所需的頁數。

## 在 IBM SPSS Modeler 中使用滑鼠

IBM SPSS Modeler 中最常見的滑鼠用法如下所示：

- **按一下**。使用滑鼠右鍵或左鍵從功能表中選取選項、開啟蹦現功能表，以及存取其他標準控制項和選項。按一下並按住按鈕可移動和拖曳節點。
- **按兩下**。按兩下滑鼠左鍵可將節點放入於串流畫布中以及編輯現有節點。
- **按一下滑鼠中鍵**。按一下滑鼠中鍵並拖曳游標可連接串流畫布中的節點。按兩下滑鼠中鍵可切斷某個節點的連接。如果沒有三鍵滑鼠，可在按一下並拖曳滑鼠時通過按 **Alt** 鍵來模擬此功能。

## 使用快速鍵

IBM SPSS Modeler 中的多數可視化程式設計作業均有與之關聯的快速鍵。例如，可通過按一下某個節點並按鍵盤上的 **Delete** 鍵將此節點刪除。同樣地，可在按住 **Ctrl** 鍵的同時按 **S** 鍵來快速儲存某個串流。控制指令（例如此指令）由 **Ctrl** 和其他鍵的組合指定，例如 **Ctrl+S**。

標準 Windows 作業中使用了大量快速鍵，例如使用 **Ctrl+X** 來執行剪下作業。IBM SPSS Modeler 不僅受支援這些快捷鍵，而且還受支援下列應用程式特定的快捷鍵。

**註：**在某些情況下，IBM SPSS Modeler 中使用的舊快速鍵與標準 Windows 快速鍵相衝突。受支援將這些舊快捷鍵與 **Alt** 鍵組合使用。例如，可以使用 **Ctrl+Alt+C** 來開啟或關閉快取。

表 1. 支援的快速鍵

捷徑按鍵	函數
Ctrl+A	選擇全部
Ctrl+X	剪下
Ctrl+N	新串流
Ctrl+O	開啟串流
Ctrl+P	列印(P)
Ctrl+C	複製
Ctrl+V	貼上
Ctrl+Z	復原
Ctrl+Q	選取選定節點的所有下游節點
Ctrl+W	取消全選下游節點（使用 Ctrl+Q 進行切換）
Ctrl+E	從選定節點執行
Ctrl+S	儲存目前串流

表 1. 支援的快速鍵 (繼續)

捷徑按鍵	函數
Alt+方向鍵	向所使用的箭頭方向移動串流畫布上的選定節點
Shift+F10	開啟選定節點的蹦現功能表

表 2. 原有快速鍵支援的捷徑

捷徑按鍵	函數
Ctrl+Alt+D	複製節點
Ctrl+Alt+L	載入節點
Ctrl+Alt+R	更名節點
Ctrl+Alt+U	建立使用者輸入節點
Ctrl+Alt+C	切換快取開啟關
Ctrl+Alt+F	清除快取
Ctrl+Alt+X	展開受
Ctrl+Alt+Z	放大/縮小
刪除	刪除節點或連線

## 正在列印

可在 IBM SPSS Modeler 中列印下列物件：

- 串流圖
- 圖形
- 表格(T)
- 報告 (來自報告節點和專案報告)
- Script (來自「串流內容」、「獨立式 Script」或「SuperNode Script」對話框)
- 模型 (模型瀏覽器、包含目前內容的對話框標籤、樹狀結構檢視器)
- 註解 (使用輸出的「註解」標籤)

要列印物件：

- 要不預覽就列印，請按一下工具列上的「列印」按鈕。
- 要在列印前設定頁，請選取「檔案」功能表中的版面設定。
- 要在列印前預覽，請選取「檔案」功能表中的預覽列印。
- 要檢視標準列印對話框中用於選取印表機以及指定外觀的選項，請選取「檔案」功能表中的列印。

## 實現 IBM SPSS Modeler 的自動化

由於進階資料採礦往往是一個冗長的複合過程，因此 IBM SPSS Modeler 包含對幾種類型的編碼和自動處理的支援。

- **表示式操作控制語言 (CLEM)** 是一種用於分析和操作在 IBM SPSS Modeler 串流中資料流程的語言。資料採礦者在串流作業中大量使用 CLEM 來執行作業，例如根據成本衍生利潤這種簡單的作業或是將 Web 日誌資料轉換成一組含有用資訊的欄位和記錄。
- **Script 編寫** 是用於在使用者介面上實現過程自動化的強大工具。Script 可以執行使用者使用滑鼠或鍵盤執行的同一類動作。還可以指定輸出並操作已產生的模型。

---

## 第 4 章 瞭解資料採礦

---

### 資料採礦概觀

透過各種技術，資料採礦可識別資料主體中有價值的資訊。資料採礦以這種方式來擷取資訊，這些資訊可用於決策支援、預測、預報以及估計之類的區域。資料通常量多價值低，以其原始形式提供的資料幾乎毫無用處。資料中的隱藏資訊才有價值。

在資料採礦中，成功來自您（或您的專家）對資料的瞭解與先進有效的分析技術的結合，電腦可使用這些技術來識別資料中的基礎關係和功能。資料採礦程序會根據歷程資料來產生模型，然後將模型用於預測、型樣偵測等等。用來建置這些模型的技術稱為機器學習或建模。

#### 建模技術

IBM SPSS Modeler 包括數種機器學習和建模技術，可以根據它們打算解決的問題類型進行大致的分組。

- 預測建模方法包括決策樹狀結構、神經網路以及統計模型。
- 叢集作業模型著重於識別相似記錄的群組，以及根據它們所隸屬的群組來標示記錄。叢集作業方法包括 Kohonen、*k*-means 和 TwoStep。
- 關聯規則將特定的結論（例如購買特殊產品）與一組條件（購買數個其他產品）相關聯。
- 篩選模型可用來篩選資料以尋找最有可能與建模相關的欄位和記錄，以及識別可能不符合已知型樣的離群值。可用的方法包括功能選擇和異常偵測。

#### 資料操作與探索

IBM SPSS Modeler 也包括數個機能可讓您將專門知識套用到資料：

- **資料操作。**建構衍生自現有資料項目的新資料項目，並將資料分解成有意義的子集。可以合併及過濾來自各種來源的資料。
- **瀏覽和視覺化。**使用「資料審核」節點顯示資料的各個方面以執行起始審核，其中包括圖形與統計資料。進階視覺化包括互動式圖形，可以匯出以併入專案報告中。
- **統計量。** 確認資料中變數之間的可疑關係。IBM SPSS Statistics 中的統計資料可用於 IBM SPSS Modeler。
- **假設檢定。**建構資料行為方式的模型並驗證這些模型。

通常，您將使用這些機能來識別資料中的一組有價值的屬性。這些屬性隨後可以供應給建模技術，建模技術將嘗試識別基礎規則和關係。

#### 一般應用程式

資料採礦的一般應用程式包括下列項目：

**直接郵件。**判定哪些人口統計群組的回應率最高。使用此資訊來最大化未來郵寄的回應。

**信用評分。**使用個人的信用歷程來制訂信用決策。

**人力資源。**瞭解過去的聘雇實務並建立決策規則來簡化聘雇程序。

**醫療研究。**建立決策規則以根據醫療證明來建議適當的程序。

**市場分析。**判定哪些變數（例如地理、價格和客戶性質）與銷售相關聯。

**品質控制。**根據產品製造分析資料並識別用來判定產品瑕疵的變數。

**產品原則。**使用意見調查資料，並透過套用決策規則以選取最重要的變數來制訂原則。

**醫療照護。**使用者意見調查和臨床資料可以結合，用來探索影響健康的變數。

名詞解釋

**屬性、欄位和變數**等術語指的是納入考量的所有案例共用的單一資料項目。表示特定案例的屬性值集合稱為記錄、範例或案例。

---

## 評量資料

除非您要使用的資料符合特定的準則，否則資料採礦可能不適用。下列小節顯示您應該考量的資料及其應用程式的幾個方面。

### 確保資料可用

這看起來似乎很明顯，但請注意，雖然資料可用，但採用的格式可能不易使用。IBM SPSS Modeler 可從資料庫（透過 ODBC）或從檔案匯入資料。但是，資料可能以無法直接存取的其他格式存放在機器上。需要先以適合的格式下載或傾出該資料，然後再使用。資料可能分散在不同的資料庫和來源中，需要將其聚集在一起。資料甚至可能不在線上。如果資料僅列印在紙張上，則必須先輸入資料再開始進行資料採礦。

### 檢查資料是否涵蓋相關的屬性

資料採礦的目標是識別相關的屬性，因此首先就併入此檢查看起來很奇怪。但是，它對於查看哪些資料可用以及嘗試識別未記錄之可能的相關因素很有用。例如，在嘗試預測冰淇淋的銷售時，您可能具備很多關於零售的銷售點或銷售歷程的資訊，這些資訊可能會擔任重要的角色。遺漏屬性不一定表示資料採礦不會產生有用的結果，但它們可能會限制產生的預測的正確性。

評量狀況的一個快速方式是對資料執行綜合性審核。繼續之前，請考量將一個「資料審核」節點附加至資料來源並執行它來產生完整報告。

### 注意雜訊資料

資料通常包含錯誤或者可能包含主觀意見，甚至是可變因素和判斷。這些現象總體稱為**雜訊**。有時，資料中有雜訊是正常的。很可能存在基礎規則，但不保證適用於 100% 的案例。

通常，資料中的雜訊越多，取得正確的結果就越困難。但是，IBM SPSS Modeler 中的機器學習方法能夠處理雜訊資料，並且已成功用於包含幾乎 50% 雜訊的資料集。

### 確保有充足的資料

在資料採礦中，資料集大小不一定很重要。資料集的代表以及其涵蓋的可能結果和變數組合重要的多。

通常，考量的屬性越多，需要提供給代表性資料涵蓋面的記錄也更多。

如果資料是代表性資料並且存在一般基礎規則，則很有可能包含數千筆（或者只有數百筆）記錄的資料樣本產生的結果與上百萬筆記錄產生的結果一樣好，但您獲得結果的速度卻更快。

## 向專案諮詢相關資料

在許多情況下，您將處理您的專有資料，因此將非常熟悉資料內容和意義。但是，如果您處理的是您組織其他部門的資料或客戶的資料，諮詢瞭解該資料的專家則很有必要。他們可指引您識別相關屬性，並且可協助您解釋資料採礦的結果、區分有價值資訊與無價值的資訊，或者因資料集中發生異常導致的構件。

---

## 資料採礦的策略

在大部分商業活動一樣，如果以規劃的系統化方式進行資料採礦會更有效。即使是使用最先進的資料採礦工具（例如 IBM SPSS Modeler），資料採礦中的大部分工作仍需要知識淵博的業務分析師來維持程序正常執行。若要引導您的規劃，請回答下列問題：

- 您要解決的實質性問題為何？
- 哪些資料來源可用，以及資料的哪些部分與目前的問題相關？
- 您在開始對資料進行採礦之前，需要進行哪些預先處理以及資料清理？
- 將使用哪些資料採礦技術？
- 您將如何評估資料採礦分析的結果？
- 您將如何充分利用擷取自資料採礦的資訊？

一般資料採礦程序很快可能變得複雜。需要掌握很多內容：複雜的業務問題、多重資料來源、資料來源間不同的資料品質、資料採礦技術的陣列、測量資料採礦是否成功的不同方式等等。

若要保持進展，它可協助明確定義用於資料採礦的程序模型。程序模型可協助您回答之前列在此區段的問題，並確保解決重要的事項。它充當資料採礦導覽圖，讓您在探索資料的複雜性時不會迷失方向。

建立用於 SPSS Modeler 的資料採礦程序為 Cross-Industry Standard Process for Data Mining (CRISP-DM)。從名字就能看出，此模型設計為一般模型，可套用至各種業界與業務問題。

---

## CRISP-DM 程序模型

一般 CRISP-DM 程序模型包括六個階段，用來解決資料採礦中的主要問題。這六個階段一起組裝在設計為將資料採礦合併至大型商業實務的循環程序中。

這六個階段包括：

- **商業理解**。這可能是資料採礦的最重要的階段。商業理解包括確定商業目標、評量狀況、確定資料採礦目標以及制訂專案計劃。
- **資料理解**。資料提供資料採礦的「原始資料」。這個階段滿足理解您的資料資源是什麼以及這些資源的性質的需求。它包括收集起始資料、說明資料、探索資料以及驗證資料品質。「輸出」節點選用區中提供的「資料審核」節點是理解資料不可或缺的工具。
- **資料準備**。對資料資源編目之後，您需要準備用於採礦的資料。準備包括選取、清除、建構、整合以及格式化資料。
- **建模**。指資料採礦的閃光部分，其中更準確的分析方法用來從資料擷取資訊。此階段涉及選取建模技術、產生測試設計以及建置和評量模型。
- **評估**。選擇模型之後，即準備好評估資料採礦結果如何協助您達成商業目標。這個階段的元素包括評估結果、檢查資料採礦程序以及確定後續步驟。
- **部署**。現在，您已完成這項工作的所有投入，是時候收穫成果了。這個階段專注於將您的新知識整合至每日商業程序來解決您的原始業務問題。這個階段包括計劃部署、監視及維護、產生定案的報告以及檢查專案。

在此程序模型中有幾個關鍵點。首先，程序以先前段落中概述的順序來執行步驟是一種普遍趨勢，另外，在數個位置還存在階段彼此之間以非線性方式互相影響的情況。例如，資料準備通常發生在建模之前。但是，在建模階段期間制訂的決策以及收集的資訊通常會引導您重新思考資料準備階段的各個部分，從而可能呈現新的建模問題。這兩個階段會互相反饋，直至已適當地解決這兩個階段為止。同樣地，評估階段可引導您重新評估您的原始商業理解，您可能發現您嘗試回答了錯誤的問題。此時，您可以修訂商業理解並使用想到的更好目標重新繼續處理程序的剩餘部分。

第二個關鍵點是資料採礦的疊代本質。您很少（甚至不會）只計劃資料採礦專案、完成它然後打包資料回家。解決客戶需求的資料採礦需要持續努力。從資料採礦的一個週期獲取的知識，或許一直會導致新的疑問和問題以及識別和滿足客戶需求的新機會。這些新的疑問、問題和機會通常可以透過重新對資料進行採礦來解決。這個採礦和識別新機會的程序應該成為您思考業務以及整體業務策略基礎的方法的一部分。

此簡介僅提供 CRISP-DM 程序模型的簡要概觀。如需模型的完整詳細資料，請參閱下列資源：

- CRISP-DM 手冊，可從安裝磁碟上的 \Documentation 資料夾中與其他文件一起存取。
- CRISP-DM 說明系統，可透過開始功能表或按一下 IBM SPSS Modeler 中「說明」功能表上的 **CRISP-DM 說明** 獲得。

---

## 模型類型

IBM SPSS Modeler 提供擷取自機器學習人工智慧以及統計資料的各種建模方法。「建模」選用區上提供的方法可讓您根據資料衍生新資訊，以及開發預測模型。每種方法都具有特定的強度且最適合因應特定類型的問題。

*IBM SPSS Modeler 應用程式手冊* 為上述多種方法提供了範例以及建模過程的一般簡介。本手冊作為線上指導教學提供，也有 PDF 格式。如需相關資訊，請參閱主題第 4 頁的『應用程式範例』。

建模方法分為以下種類：

- 受監督
- 關聯
- 分區段

### 受監督模型

受監督模型使用一個或多個輸入欄位的值來預測一個或多個輸出（或目標）欄位的值。這些技術的一些範例包括：決策樹（C&R 樹狀結構、QUEST、CHAID 和 C5.0 演算法）、迴歸方法（線性、logistic、通用性線性和 Cox 迴歸演算法）、神經網路、支援向量機器和貝葉斯網路。

「受監督」模型可說明組織預測已知的結果，例如顧客是否購買、流失或某交易是否符合某種已知的犯罪型態。其建模技術包含機器學習、規則歸納、子群組識別、統計技術和多模型產生。

### 受監督節點



「自動分類器」節點用於建立和對比二元結果（是或否，流失或不流失等）的若干不同模型，使用戶可以選擇給定分析的最佳處理方法。由於受支援多種建模演算法，因此可以對用戶希望使用的方法、每種方法的特定選項以及對比結果的準則進行選取。節點根據指定的選項產生一組模型並根據用戶指定的準則排列最佳候選項的順等級。



自動數值節點使用多種不同方法估計和對比模型的連續數值範圍結果。此節點和自動分類器節點的工作方式相同，因此可以選擇要使用和要在單個建模傳送中使用多個選項組合進行測試的演算法。受支援的演算法包含神經網路、C&R 樹狀結構、CHAID、線性迴歸、通用性線性迴歸以及受支援向量機器 (SVM)。可基於相關係數度、相對錯誤或已用變數數目對模型進行對比。



分類和迴歸方法 (C&R) 樹狀結構節點產生可用於預測或分類未來觀察的決策樹。該方法通過在每一個步驟最大限度降低不純潔度，使用遞歸分區來將訓練記錄分割為組。如果樹狀結構中某個節點中 100% 的觀察值都的目標欄位的一個特定種類，那麼該節點將被認定為「純潔」。目標欄位和輸入欄位可以是數值範圍或類別（標稱、序數或旗標）欄位；所有分割都是二元的（只有兩個子群組）。



QUEST 節點可提供用於建立決策樹的二元分類法，此方法的設計目的是減少大型 C&R 樹狀結構分析所需的處理時間，同時也減少在分類樹狀結構方法中發現的趨勢以便偏愛容許有多個分割的輸入。輸入欄位可以是數值範圍（連續），但目標欄位必須是種類。所有分割都是二元的。



CHAID 使用卡方統計資料來產生決策樹，以確定最佳的分割。與 C&R Tree 和 QUEST 節點不同，CHAID 可產生非二元樹狀結構，表示部分分割有兩個以上的分支。目標欄位和輸入欄位可以是數值範圍（連續）或類別欄位。「詳盡的 CHAID」是對 CHAID 的修改，它會徹底檢查所有可能的分割，但計算時間較長。



C5.0 節點建立決策樹或規則集。該模型的工作原理是根據在每個層次提供上限資訊收穫的欄位分割樣本。目標欄位必須為種類欄位。容許進行多次多於兩個子群組的分割。



決策清單節點可識別子群組或區段，顯示與整體相關的給定二元結果的概似度的高低。例如，您或許在尋找那些最不可能流失的客戶或最有可能對某個商業活動作出積極回應的客戶。通過自訂區段和並排預覽備選模型來比較結果，您可以將自己的業務知識體現在模型中。決策清單模型由一組規則構成，其中每個規則具備一個條件和一個結果。規則依順序套用，相符的第一個規則將決定結果。



線性迴歸模型基於目標和一個或多個預測值之間的線性關係預測連續目標。



PCA/因素節點提供強大的資料減少技術來減少資料的複雜性。主成份分析 (PCA) 可找出輸入欄位的線性組合，該組合最好地擷取了整個欄位集中的變異數，且組合中的各個成分相互正交（相互垂直）。因數分析則試圖識別底層因素，這些因素說明觀測的欄位集合內的相關性樣。對於這兩種方法，其共同的目標是找到可對原始欄位集中的資訊進行有效總結的少量衍生欄位。



功能選擇節點會根據某組準則（例如遺漏百分比）篩選可移除的輸入欄位，對於保留的輸入，隨後將對其相對於指定目標的重要性進行排等級。例如，假如某個給定資料集有上千個潛在輸入，那麼哪些輸入最有可能用於對病患結果進行建模呢？



判別分析所做的假設比邏輯迴歸方法的假設更嚴格，但在符合這些假設時，判別分析可以作為邏輯迴歸方法分析的有用替代項或補充。



邏輯迴歸是一種統計技術，它可根據輸入欄位的值對記錄進行分類。它類似於線性迴歸方法，但採用的是種類目標欄位而非數值範圍。



「通用性線性」模型對一般線性模型進行了擴展，這樣依變數通過指定的鏈結函數與因子和共變數線性相關。此外，此模式允許變數具有非常態分配。它包括統計模型大部分的功能，其中包括線性迴歸、logistic 迴歸方法、用於計數資料的對數線性模型以及區間刪失生存分析模型。



概化線性混合模型 (GLMM) 延伸了線性模型，使得目標可以有非常態分佈，通過指定的連接函數與因子和共變數線性相關，並且觀察可能相關。通用性線性混合模型涵蓋多種模式，從非常態縱向資料的簡單線性迴歸，到複雜的多層級模式。



使用 Cox 迴歸節點，您可以在已有的檢查記錄中建立時間事件的生存分析模型。該模型會生成一個生存分析函數，該函數可預測在給定時間 ( $t$ ) 內對於所給定的輸入變數值相關事件的發生機率。



使用支援向量機器 (SVM) 節點，可以將資料分為兩群組，而無需過度配適。SVM 可以與大量資料集配合使用，例如那些含有大量輸入欄位的資料集。



通過貝葉斯網路節點，你可以利用對真實世界認知的判斷力並結合所觀察和記錄的證據來建立機率模型。該節點重點應用了樹狀結構擴展簡單貝葉斯 (TAN) 和馬爾可夫覆蓋網路，這些算法主要用於分類問題。





自習回應模型 (SLRM) 節點可用於建立一個包含單個新觀察值或少量新觀察值的模型，通過此模型，無需使用全部資料對模型進行重新訓練即可對模型進行重新評估。



「時間序列」節點會為時間序列資料評估指數平滑化、單變量自身迴歸整合移動平均 (ARIMA)，以及多變量 ARIMA (或稱轉換函數) 模型，然後產生未來效能的預測。此「時間序列」節點類似於 SPSS Modeler 第 18 版中不推薦使用的先前「時間序列」節點。但是，此較新「時間序列」節點旨在利用 IBM SPSS Analytic Server 的能力來處理大資料，並在 SPSS Modeler 第 17 版中新增的輸出檢視器中顯示生成的模型。



The  $k$ -最近相鄰元素 (KNN) 節點將新的觀察值關聯到預測值空間中與其最鄰近的  $k$  個物件的種類或值 (其中  $k$  為整數)。相似的觀察值會彼此相鄰，相異的觀察值則會彼此相隔。



空間-時間預測 (STP) 節點使用包含位置資料、預測輸入欄位 (預測值)、時間欄位和目標欄位的資料。每個位置有數個資料列，用來代表在每次測量時每個預測工具的值。分析資料之後，它可用來在分析中所使用形狀資料內的任何位置，預測目標值。

## 關聯模型

關聯模型尋找您資料中的型樣，其中一個或多個實體 (如事件、購買或屬性) 與一個或多個其他實體相關聯。這些模型構建定義這些關係的規則集。資料中的欄位可以作為輸入和目標。您可以手動尋找這些關聯，但關聯規則演算法可以更快速地完成，並能探索更多複合的型樣。Apriori 和 Carma 模型是使用此類演算法的範例。另一種類型的關聯模型是序列偵測模型，後者可以在按時間建立結構的資料中尋找順序型樣。

關聯模型在預測多個結果時非常有用，例如，購買了產品 X 的顧客也購買了產品 Y 和 Z。關聯模型可以將特定結論 (如購買某些產品的決策) 與一組條件關聯起來。關聯規則演算法相對於更標準的決策樹演算法 (C5.0 和 C&RT) 的優勢在於，它可以找到任何屬性間存在的關聯。決策樹演算法將建置只有一個結果的規則，而關聯演算法會嘗試尋找許多規則，每個規則可能具有不同的結果。

## 關聯節點



「事前」節點從資料擷取一組規則，即擷取資訊內容最多的規則。Apriori 節點提供五種選取規則的方法並使用複雜的編製索引模式來高效地處理大資料集。對於較大的問題，Apriori 訓練的速度通常較快較快；它對可保留的規則數目量沒有任何限制，而且可處理最多帶有 32 個前提條件的規則。「事前」要求輸入和輸出欄位均為種類型欄位，但因為它專為處理此類型資料而進行最佳化，因而處理速度快得多。



CARMA 模型會從資料中擷取一組規則，而不需要您指定輸入或目標欄位。與 Apriori 不同，CARMA 節點提供建立規則設定支援 (前提條件和結果支援)，而不僅僅是前提條件支援。這就意味著產生的規則可以用於更多應用程式，例如用於尋找產品或服務 (前提條件) 的清單，這些產品或服務的結果為想在節日期間促銷的商品。



序列節點可探索循序資料或與時間有關的資料中的相關規則。序列是一系列可能會以可預測順序發生的項目集合。例如，一個購買了剃刀和須後水的顧客可能在下次購物時購買剃須膏。「序列」節點基於 CARMA 關聯規則演算法，使用有效的兩段式方法來尋找序列。



「相關規則」節點與 Apriori 節點類似；但是，與 Apriori 不同，「相關規則」節點能夠處理清單資料。另外，「關聯規則」節點可以與 IBM SPSS Analytic Server 配合使用，以正在處理大型資料以及利用更快的平行處理功能。

## 分區段模型

分區段模型將資料劃分為具有類似輸入欄位型樣的記錄區段或叢集。分區段模型只對輸入欄位感興趣，沒有輸出或目標欄位的概念。分區段模型的範例為 Kohonen 網路、K-Means 叢集、二階叢集和異常偵測等。

在不知道特定結果的情況下（例如，需要識別新犯罪型樣或在客戶群中識別利益群體時），分區段模型（也稱為「叢集模型」）非常有用。叢集作業模型著重於識別相似記錄的群組，以及根據它們所隸屬的群組來標示記錄。此技術的優點在於，不用提前瞭解這些群組及其特性就可以使用，它使叢集模型（其中沒有需要模型預測的預先定義輸出或目標欄位）區別於其他的建模技術。對於這些模型來說，沒有正確或錯誤的結果之分。其值根據它們的以下能力來決定：它們能夠擷取資料中的相關分組並為這些分組提供有用說明。叢集模型通常用於建立在後續分析中用作輸入的叢集或區段（例如，將潛在用戶分成幾個相似的子群組）。

## 分區段節點



自動叢集節點估計和比較識別具有類似特性記錄群組的叢集模型。節點工作方式與其他自動建模節點相同，使您在一次建模運行中即可試驗多個選項組合。可以使用基本測量比較模型，透過測量嘗試過濾並分級叢集模型的實用性，並基於特定欄位的重要性提供測量。



K-Means 節點將資料集叢集到不同群組（或叢集）。此方法將定義固定的叢集數目量，將記錄迭代分配給叢集，以及調整叢集中心，直到進一步優化無法再精確模型。*k*-means 節點作為一種非監督學習機制，它並不試圖預測結果，而是揭示隱含在輸入欄位集中的型樣。



Kohonen 節點會產生一種類神經網路，此類神經網路可用於將資料集叢集到各個差異群組。此網路訓練完成後，相似的記錄應在輸出對映中緊密地聚集，差異大的記錄則應彼此遠離。您可以通過查看模型塊中每個單位所擷取觀察的數量來找出規模較大的單元。這將讓您對叢集的相應數量有所估計。



TwoStep 節點使用二階叢集方法。第一步完成簡單資料製作，以便將原始輸入資料壓縮為可管理的子叢集集合。第二步使用層級叢集方法將子叢集一步一步合併為更大的叢集。TwoStep 具有一個優點，就是能夠為訓練資料自動估計最佳叢集數目。它可以高效處理混合的欄位類型和大型的資料集。



Anomaly Detection 節點確定不符合「正常」資料格式的異常觀察值（離群值）。即使離群值不匹配任何已知格式或用戶不清楚自己的查找目標，也可以使用此節點來確定離群值。

### 資料庫內資料採礦模型

IBM SPSS Modeler 支援與多家資料庫供應商的資料挖掘和建模工具整合，這包含 Oracle Data Miner 和 Microsoft Analysis Services。您可以在資料庫內建置及儲存模型以及為模型評分 — 所有這些作業都是在 IBM SPSS Modeler 應用程式中進行。有關完整的詳細資料，請參閱《IBM SPSS Modeler 資料庫內資料採礦手冊》。

### IBM SPSS Statistics 模型

如果您在電腦上擁有 IBM SPSS Statistics 安裝和軟體使用權的一個副本，您可以從 IBM SPSS Modeler 存取和執行某些 IBM SPSS Statistics 常式以建立模型和給模型分數。

---

## 資料採礦範例

從範例入手是瞭解資料採礦實際應用的最佳方式。《IBM SPSS Modeler 應用程式手冊》中包含很多應用程式範例，同時還提供了有關特定建模方法和技術的簡要針對性說明。請參閱第 4 頁的『應用程式範例』主題，以取得更多資訊。



---

## 第 5 章 建立串流

---

### 串流建立概觀

使用 IBM SPSS Modeler 進行的資料採礦重點關注通過一系列節點執行資料的過程，我們將這一過程稱為串流。這一系列節點代表要對資料執行的作業，而節點之間的鏈結指示資料流程的方向。通常，您可以使用資料串流將資料讀取入 IBM SPSS Modeler，通過一系列操作執行資料，然後將其傳送至某個地方，如表格或檢視器。

例如，假定您需要開啟某個資料來源、新增新欄位、根據新欄位中的值選取記錄，然後在表格中顯示結果。在這種情況下，您的資料串流應由以下四個節點組成：



「變數檔案」節點，設定此節點後可以讀取資料來源中的資料。



衍生節點，用於向資料集中新增計算的新欄位。



選取節點，用於設定選擇準則，以從資料串流中排除某些記錄。



表格節點，用於在畫面上顯示操作結果。

---

### 建置資料串流

通過唯一的 SPSS Modeler 介面，可以使用資料串流圖直觀地挖掘資料。基本上，您可採用下列步驟建立資料串流：

- 將節點新增到串流畫布。
- 連接節點以形成串流。
- 指定任意節點或串流選項。
- 執行串流。

本部分包含有關使用節點建立較複合資料串流的更多詳細資訊。還介紹了節點和串流的選項和設定。有關使用 SPSS Modeler 隨附的資料逐步建立串流的範例（位於程式安裝目錄的 Demos 資料夾中），請參閱第 4 頁的『應用程式範例』。

### 使用節點

在 IBM SPSS Modeler 中，節點用於協助您研究資料。畫布中的各種節點分別代表不同的物件和動作。IBM SPSS Modeler 視窗底部的選用區包含串流建立中可用的所有節點。

存在多種類型的節點。**來源節點**將資料導入到串流中，它位於節點選用區的「來源」標籤上。**處理節點**在單個資料記錄和欄位上執行作業，它位於選用區的「記錄處理」和「欄位作業」標籤上。**輸出節點**為資料、圖表和模型結果產生各種輸出，並包含在節點選用區的「圖形」、「輸出」和「匯出」標籤上。**建模節點**使用統計演算法建立模型塊，它位於節點選用區的「建模」標籤和「資料庫建模」（如啟動）標籤上。請參閱第 13 頁的『節點選用區』主題，以取得更多資訊。

連接節點會形成串流，當執行串流時，可以直觀地查看節點之間的關係並得出結論。串流類似於 Script - 可以儲存並將其重新用於其他資料檔案。

處理串流資料的可執行節點稱為**終端節點**。如果建模節點或輸出節點位於串流或串流分支的末端，則為終端節點。您不能將其他節點連接到終端節點。

註：可以自訂節點選用區。請參閱第 173 頁的『自訂節點選用區』主題，以取得更多資訊。

## 將節點新增至串流

可以通過多種方法將節點從節點選用區新增至串流：

- 按兩下選用區中的節點。註：按兩下某個節點會自動將其連接到目前串流。請參閱『連接串流中的節點』主題，以取得更多資訊。
- 將節點從選用區拖放到串流畫布。
- 按一下選用區中的節點，然後按一下串流畫布。
- 從 IBM SPSS Modeler 的「插入」功能表中選取合適的選項。

將節點新增至串流畫布後，按兩下該節點可顯示相應的對話框。可用選項取決於您要新增的節點的類型。有關對話框中特定控制項的資訊，請按一下其**說明**按鈕。

## 移除節點

要從資料串流中移除某個節點，請按一下該節點並按 Delete 鍵，或者用滑鼠右鍵按一下並從功能表中選取刪除。

## 連接串流中的節點

已新增到串流畫布的節點在連接之前不會形成資料串流。節點之間的連線指示資料從一項作業流程向下一項作業的方向。連接節點以形成串流的方法有以下幾種：按兩下、使用滑鼠中鍵或手動連接。

### 通過按兩下新增並連接節點

形成串流的最簡單的方法是按兩下選用區中的節點。此方法會自動將新節點連接到串流畫布中的選定節點。例如，如果畫布包含一個「資料庫」節點，則可以選中該節點，然後在選用區中按兩下下一個節點，如「衍生」節點。此動作會自動將「衍生」節點連接到現有的「資料庫」節點。您可以重複此過程，直到到達終端節點（如「直方圖」或「表格」節點），這時所有新節點都將連接至上游最後一個非終端節點。

### 使用滑鼠中鍵連接節點

在串流畫布中，可以使用滑鼠中鍵按一下某個節點並將其拖到另一個節點。（如果滑鼠沒有中鍵，可以通過按住 Alt 鍵的同時使用滑鼠從一個節點拖到另一個節點來模擬此操作。）

### 手動連接節點

如果沒有滑鼠中鍵並希望手動連接節點，您可以使用節點的**蹦現**功能表將其連接至畫布已有的另一個節點。

1. 用滑鼠右鍵按一下選擇連線的起始節點。這將開啟節點功能表。

2. 在功能表中按一下**連接**。
3. 此時，起始節點和游標處將同時顯示連線圖示。按一下畫布中的第二個節點以連接這兩個節點。

連接節點時，需要遵循以下幾項規則。如果試圖進行下列任何一種連線，您將收到錯誤訊息：

- 指向來源節點的連線
- 發自終端節點的連線
- 超出節點的上限輸入連線數
- 連接兩個已連線的節點
- 循環（資料傳回其從中流出的節點）

## 繞過串流中的節點

如果繞過資料串流中的某個節點，該節點的所有輸入和輸出連線都將取代為直接從其輸入節點通向其輸出節點的連線。如果該節點不同時具備輸入連線和輸出連線，則會刪除該節點的所有連線，而不是更改路線。

例如，您可能有一個串流，它衍生一個新欄位、過濾欄位，然後研究直方圖和表格中的結果。如果還希望在過濾欄位之前檢視相同的資料圖表和真值表格，您可以在串流中新增新的直方圖節點和表格節點，也可以繞過過濾器節點。繞過「過濾器」節點時，將直接通過「衍生」節點連線到圖表和表格。「過濾器」節點將從串流中中斷連線。

### 繞過節點

1. 在串流畫布中，使用滑鼠中鍵按兩下要繞過的節點。或者，您還可使用 Alt+按兩下。

附註：您可以按一下「編輯」功能表上的復原或按 Ctrl+Z 來復原此動作。

## 取消串流中的節點

可以停用串流中具有單個輸入的處理節點，這樣在串流執行期間將忽略此節點。這種方式使您不必刪除或繞過節點，並保持其連接至其餘節點。您還可以開啟和編輯節點設定，不過，除非重新啟用節點，否則任何變更均不會生效。

例如，您可能有一個過濾多個欄位，然後使用減少後的資料集建立模型的串流。如果還需要建立未過濾欄位的相同模型，以請參閱這些欄位是否改進模型結果，可以取消「過濾器」節點。取消「過濾器」節點時，到建模節點的連線將直接從「衍生」節點到「類型」節點。

### 取消節點

1. 在串流畫布中，用滑鼠右鍵按一下要取消的節點。
2. 在蹦現功能表中按一下**停用節點**。

或者可以在「編輯」功能表中按一下**節點 > 取消節點**。當需要重新在串流中包含節點時，可採用相同方式按一下**啟用節點**。

附註：您可以按一下「編輯」功能表上的復原或按 Ctrl+Z 來復原此動作。

## 在現有連線中新增節點

通過拖曳用於連接兩個節點的箭頭，您可以在這兩個已連接的節點之間新增新節點。

1. 使用滑鼠中鍵按一下並拖曳要向其中插入節點的連線箭頭。也可以在按住 Alt 鍵的同時按一下並拖曳滑鼠以模擬滑鼠中鍵。
2. 將連線拖曳到您要包含的節點，然後鬆開啟滑鼠按鈕。

註：可以通過繞過節點將新連線從節點中移除並還原原始節點。

## 刪除節點之間的連線

要刪除兩個節點間的連線：

1. 用滑鼠右鍵按一下連線箭頭。
2. 在功能表中按一下刪除連線。

要刪除從某個節點開始以及到該節點的所有連線，請執行下列一種操作：

- 選中節點並按 F3。
- 選取該節點，然後在主功能表中按一下：

編輯 > 節點 > 中斷連線

## 節點的設定選項

創建並連接節點後，可使用一些自訂節點的選項。用滑鼠右鍵按一下節點，然後選取某個功能表選項。

- 按一下編輯可開啟選定節點的對話框。
- 按一下連接可將一個節點手動連接到另一個節點。
- 按一下中斷連線可刪除出入節點的所有鏈結。
- 按一下更名並加註解可開啟「編輯」對話框的「註解」標籤。
- 按一下新增註解可新增節點相關的備註。請參閱第 49 頁的『新增備註和註解到節點和串流』主題，以取得更多資訊。
- 按一下停用節點可在正在處理期間「隱藏」節點。要使節點再次可見以供處理，可以按一下啟用節點。請參閱第 33 頁的『取消串流中的節點』主題，以取得更多資訊。
- 按一下剪下或刪除可從串流畫布中刪除選定節點。註：按一下剪下可以貼上節點，按一下刪除則不能。
- 按一下複製可生成不包含連線的節點副本。可以將此副本新增到新的或現有的串流中。
- 按一下載入節點可開啟先前儲存的節點，並將其選項載入目前選定的節點。這兩個節點必須的同一類型。
- 按一下擷取節點可從連接的 IBM SPSS Collaboration and Deployment Services 儲存庫中擷取節點。
- 按一下儲存節點可將節點的詳細資料儲存在檔案中。只能將節點詳細資料載入其他類型相同的節點。
- 按一下儲存節點可將所選節點儲存在連接的 IBM SPSS Collaboration and Deployment Services 儲存庫中。
- 按一下快取可展開包含對選定節點進行快取選項的功能表。
- 按一下資料對映可展開包含將資料對映到新來源或指定必備欄位選項的功能表。
- 按一下建立 SuperNode 可展開包含在目前串流中建立 SuperNode 選項的功能表。
- 按一下產生使用者輸入節點可取代選定節點。此節點產生的範例將包含與目前節點相同的欄位。
- 按一下從這裡執行可從選定節點向下執行所有終端節點。

## 節點的快取選項

為最佳化串流執行，可以在任何非終端節點上建立快取。如果已在節點上建立了快取，則在下一次執行資料串流時串流過節點的資料會充滿快取。從這時起，就會從快取（儲存在磁碟的暫存目錄中）而不是從資料來源中讀取資料了。

對於諸如排序、合併或聚集這樣比較耗時的作業，快取是最有用的。例如，假設已設定可從資料庫中讀取銷售資料的來源節點，和可按區域聚合銷售資料的「聚合」節點。可以在「聚合」節點上建立快取而不是在來源節點上建立快取，因為希望快取儲存的是已聚合的資料而不是整個資料集。



註：來源節點位置的快取僅用於儲存讀取 IBM SPSS Modeler 的原始資料複本，在大多數情況下，它不會帶來效能方面的提升。

對於已啟用快取的節點，會在右上角顯示一個小的文件圖示。當節點上的資料被快取時，文件圖示將變綠。

### 啟用快取

1. 在串流畫布上，用滑鼠右鍵按一下節點並按一下功能表上的**快取**。
2. 在快取子功能表中，按一下**啟用**。
3. 可以通過用滑鼠右鍵按一下節點，然後在快取子功能表中按一下**取消**來關閉快取。

### 資料庫中的快取節點

對於在資料庫中執行的串流，可於中遊將資料快取到資料庫的暫時表格中而不是檔案系統中。結合 SQL 最佳化使用時，這可能會造成效能顯著提升。例如，合併多個表格以建立資料採礦視圖的串流中的輸出在必要時可能會加以快取並重複使用。通過為所有下游節點自動產生 SQL，效能可得到進一步的提高。

為利用資料庫快取，必須同時啟用 SQL 最佳化和資料庫快取。注意，伺服器上的最佳化設定將置換用戶端上的最佳化設定。請參閱第 39 頁的『設定串流的最佳化選項』主題，以取得更多資訊。

如果已啟用資料庫快取，則只需用滑鼠右鍵按一下任意非終端節點就可以快取該點上的資料，並且在下次執行串流時，可以直接在資料庫中自動建立快取。如果未啟用資料庫快取或 SQL 最佳化，則快取將會改為寫入檔案系統。

註：下列資料庫支援用於快取的暫時表格：Db2、Oracle、SQL Server 和 Teradata。其他資料庫（例如 Netezza）將使用標準表格用於資料庫快取。可以針對特定資料庫自訂 SQL 代碼 - 請與服務人員聯絡以獲取協助。

### 清除快取

節點上的白色文件圖示表示其快取為空。如果快取已滿，則文件圖示會變成純綠色。如果要取代快取的內容，必須先清除快取，然後重新執行資料串流以對其進行重新填充。

1. 在串流畫布上，用滑鼠右鍵按一下節點並按一下功能表上的**快取**。
2. 在快取子功能表中，按一下**清除**。

### 儲存快取

可以將快取的內容儲存為 IBM SPSS Statistics 資料檔案 (\*.sav)。隨後您可以將該檔案作為快取重新載入，也可以設定將該快取檔用作其資料來源的節點。此外，還可以載入其他專案中儲存的快取。

1. 在串流畫布上，用滑鼠右鍵按一下節點並按一下功能表上的**快取**。
2. 在快取子功能表中，按一下**儲存快取**。
3. 在「儲存快取」對話框中，瀏覽至要儲存快取檔的位置。
4. 在「檔案名稱」文字框中輸入名稱。
5. 確保在「檔案類型」清單中已選取 **\*.sav**，然後按一下**儲存**。

### 載入快取

如果在將某個快取檔從節點中刪除之前已對其進行了儲存，則可以重新載入該檔案。

1. 在串流畫布上，用滑鼠右鍵按一下節點並按一下功能表上的**快取**。
2. 在快取子功能表中，按一下**載入快取**。

3. 在「載入快取」對話框中瀏覽至快取檔所在位置，選中該檔案，然後按一下**載入**。

## 預覽節點中的資料

為確保在建立串流時資料以預期的方式變更，可以在每個重要步驟上通過「表格」節點執行資料。為避免這樣做，可以從每個節點產生會顯示將要建立的資料樣本的預覽，從而減少建立每個節點所需的時間。

對於模型塊的上游節點，預覽顯示輸入欄位；對於模型塊或其下游節點（不含端節點），預覽顯示輸入和產生的欄位。

顯示的預設列數為 10；但是，您可以在串流內容中變更此值。請參閱第 37 頁的『設定串流的一般選項』主題，以取得更多資訊。

從產生功能表中，可以建立多種類型的節點。

註：預覽此節點產生的資料時，將對此節點套用所有內容變更且無法取消（此行為與按一下**套用**相同）。

## 鎖定節點

為了防止其他使用者修正串流中一個或多個節點的設定，您可以將節點封裝在稱為 SuperNode 的特殊類型的節點中，然後通過套用密碼保護鎖定 SuperNode。

## 使用串流

在串流畫布中連接來源、過程和終端節點後，便建立了一個串流。作為節點集合，可以對串流進行儲存、新增註解，還可以將其新增到專案。此外，還可以為串流設定各種選項，如最佳化、日期與時間設定、參數和 Script。這些內容將在下面的主題中進行介紹。

在 IBM SPSS Modeler 中，可以在同一 IBM SPSS Modeler 階段作業中使用和修改多個資料串流。主視窗的右端包含管理程式窗格，可協助您在目前開啟的串流、輸出和模型中進行導覽。如果您無法看到管理程式窗格，請在「檢視」功能表中按一下**管理員**，然後按一下**串流**標籤。

通過此標籤，可以執行下列操作：

- 存取串流。
- 儲存串流。
- 將串流儲存至目前專案。
- 關閉串流。
- 開啟新串流。
- 在 IBM SPSS Collaboration and Deployment Services 儲存庫中儲存和擷取串流（如果在您的網站上可用）。請參閱第 137 頁的『關於 IBM SPSS Collaboration and Deployment Services 儲存庫』主題，以取得更多資訊。

在「串流」標籤中的某個串流上按一下滑鼠右鍵可存取這些選項。

## 設定串流的選項

您可以指定眾多選項以套用至目前串流。還可將這些選項儲存為預設值以套用至所有串流。這些選項如下。

- **一般**。諸如要在串流中使用的符號和文字編碼等雜項。如需相關資訊，請參閱第 37 頁的『設定串流的一般選項』。
- **日期/時間**。有關日期和時間表示式格式的選項。如需相關資訊，請參閱第 38 頁的『設定串流的日期與時間選項』。

- **數字格式**。用於控制數值表示式格式的選項。如需相關資訊，請參閱第 39 頁的『設定串流的數字格式選項』。
- **最佳化**。這是用於提高串流效能的選項。如需相關資訊，請參閱第 39 頁的『設定串流的最佳化選項』。
- **記錄與狀態**。用於控制 SQL 日誌和記錄狀態的選項。如需相關資訊，請參閱第 40 頁的『設定串流的 SQL 日誌和記錄狀態選項』。
- **佈置**。有關串流在畫布上的佈置的選項。如需相關資訊，請參閱第 41 頁的『設定串流的佈置選項』。
- **分析伺服器**。有關將 Analytic Server 與 SPSS Modeler 配合使用的選項。如需相關資訊，請參閱第 41 頁的『Analytic Server 串流內容』。
- **地理空間**。有關對地理空間資料進行格式編排以便在串流中使用的選項。如需相關資訊，請參閱第 41 頁的『設定串流的地理空間選項』。

## 設定串流選項

1. 在「檔案」功能表上，按一下**串流內容**（或從管理程式窗格中的「串流」標籤選取串流，用滑鼠右鍵按一下，然後按一下**串流內容**）。
2. 按一下「**選項**」標籤。

或者，在「工具」功能表上，按一下：

**串流內容 > 選項**

**設定串流的一般選項：** 一般選項包含一組套用於目前串流的不同方面的雜項。

**基本部分**提供了下列基本選項：

- **小數點符號**。選取逗點 (,) 或句號 (.) 作為小數分隔字元。
- **分組符號**。針對為顯示格式編號，選取用於對值進行群組的符號（例如，3,000.00 中的逗點）。選項包括無、句點、逗點、空格以及語言環境所定義（在此情況下會使用現行語言環境的預設值）。
- **編碼**。指定文字編碼的串流預設方法。（註：適用於變量。檔案來源節點和純文字檔案匯出節點。其他節點不使用此設定；多數資料檔案帶有內含編碼資訊。您可以選擇系統預設或 UTF-8。系統預設值是在 Windows 控制台中指定，如果是分散式模式執行，則是在伺服器電腦上指定。請參閱第 191 頁的『IBM SPSS Modeler 中的 Unicode 支援』主題，以取得更多資訊。
- **規則組評估**。確定過長模型的評估方式。依預設，過長將採用投票方式綜合各項目規則產生的預測值，並確定最終預測值。為確保過長預設使用第一次命中規則，可選取**第一個符合項**。請注意，此選項不適用於決策清單模型，該模型始終套用由演算法定義的第一個符合項。

**在資料預覽中顯示的上限列個數**。指定當已要求某個節點的資料預覽時顯示的列個數。請參閱第 36 頁的『預覽節點中的資料』主題，以取得更多資訊。

**名義欄位的成員上限**。選取此選項可指定列名（集合）欄位的上限會員數，達到此數字後欄位的資料類型將變為無類型。當使用較大列名欄位時，此選項十分有用。註：當欄位的測量層次設定為無類型時，其角色將自動設定為無。這意味著這種欄位無法用於建模。

**限制用於 Kohonen 和 K-Means 建模的集大小**。選取此選項可指定用於 Kohonen 網路和 K-Means 建模的集合欄位的上限會員數。預設集合大小為 20，達到此數值後，將忽略該欄位並發出警告，其中顯示有關該欄位的錯誤資訊。

請注意，為了確保相容，此選項還將應用於在 IBM SPSS Modeler 第 14 版中已被取代的舊類神經網路節點；一些舊串流可能仍包含此節點。

**執行後重新整理來源節點。** 選取此選項將在執行目前串流時自動重新整理所有來源節點。此動作類似於針對某個來源節點點擊重新整理按鈕，區別在於此選項會自動重新整理目前串流的所有來源節點（使用者輸入節點除外）。

**註：**選取此選項將清除下游節點的快取，即使資料未發生變更也是如此。如果您使用工具列中的執行目前串流選項，但每次執行串流只進行一次清除，這意味著仍可使用下游快取作為單次執行時期的暫時儲存體。例如，假定您在進行某項複合的衍生作業後設定了中遊快取，且已為該衍生節點的下游附加若干圖形和報告。執行串流時，將清除衍生節點的快取，並僅重新填充第一個圖表或報告。後續終端節點將從該「衍生」節點的快取中讀取資料。請注意，如果您選擇個別執行每一個終端節點（如果有多個終端節點），而不是使用執行目前串流選項，那麼每一個執行終端節點時都會發生快取清除。

**在輸出中顯示欄位與值標籤。** 在表格、圖表和其他輸出中顯示欄位與值標籤。如果標籤不存在，則會改為顯示欄位名稱和資料值。依預設，標籤處於關閉狀態；但您可以在 IBM SPSS Modeler 的其他位置個別切換各個標籤。此外，還可以使用工具列中的切換按鈕選擇在輸出視窗中顯示標籤。



圖 11. 用於切換欄位和值標籤的工具列圖示

**顯示執行時間。** 執行串流之後，在「執行時間」標籤上顯示串流節點的相應執行時間。請參閱第 43 頁的『檢視節點執行時間』主題，以取得更多資訊。

**自動建立節點部分**提供了下列用於在各個串流中自動建立節點的選項。這些選項控制產生新建模塊時是否將這些塊插入到串流畫布中。依預設，這些選項僅套用至第 16 版或更高版本中建立的串流。在 IBM SPSS Modeler 第 16 版或更高版本中，如果開啟在第 15 版或之前版本中建立的串流並執行建模節點，那麼不會像舊版那樣將塊放入串流畫布。如果您使用 IBM SPSS Modeler 第 16 版或更高版本建立新串流並執行建模節點，那麼產生的塊將放入串流畫布。進行此設計的原因包括，在 IBM SPSS Collaboration and Deployment Services 中，以及在不出現 IBM SPSS Modeler Server 用戶端使用者界面的其他環境中，為新模型輸出建立模型套用節點選項有可能中斷第 16 版以前的版本中建立的那些以批次方式執行的串流。

- **為新模型輸出建立模型套用節點。** 自動地為新模型輸出建立模型套用節點。如果選中此選項，那麼還可以從**建立模型更新鏈結**中選擇是將這些鏈結設定為啟用、設定為停用還是不建立這些鏈結。

建立新的模型應用節點或來源節點時，下拉清單中的鏈結選項用於控制是否在建置器節點與新節點之間建立更新鏈結，以及建立這些鏈結後它們所處的模式。如果建立這些鏈結，那麼您有機會將其啟用，但這些選項使使用者能夠進行全面控制。

- **根據來源建置器建立來源節點。** 自動地根據來源建置器建立來源節點。與上一個選項相似，如果選中此選項，那麼還可以從**建立來源重新整理鏈結**下拉清單中選擇是將重新整理鏈結設定為啟用、設定為停用還是不建立這些鏈結。

**儲存為預設值。** 指定的選項僅套用至現行串流。按一下此按鈕以將這些選項設為所有串流的預設值。

**設定串流的日期與時間選項：** 這些選項指定目前串流中的各種日期與時間表示式要使用的格式。

**匯入日期/時間為** 選取是否將日期/時間用於日期/時間欄位或者是否將它們匯入為字串變數。

**日期格式** 選取要用於日期儲存欄位的日期格式或 CLEM 日期函數將字串解析為日期時使用的日期格式。

**時間格式** 選取要用於時間儲存欄位的時間格式或 CLEM 時間函數將字串解析為時間時使用的時間格式。

**輪替天數/分鐘數** 對於時間格式，請選取是否將負時差解釋為指示前一天或前一小時。

**日期基準線 (1 月 1 日)** 選取採用單日期的 CLEM 日期函數要使用的基年 (始終為 1 月 1 日)。

**兩位日期 (起始自)** 指定要為僅以兩位數表示的年度新增公元位的截止年度。例如, 指定 1930 為截止年度會假定 05/11/02 為 2002 年。相同的設定將針對 30 後的日期使用 20 世紀公元位, 因此 05/11/73 將假定為 1973 年。

**時區** 選取如何選取時區以用於 `datetime_now` CLEM 表示式。

- 如果您選取**伺服器**, 那麼時區將取決於下列各項目:
  - 如果目前串流使用 Analytic Server 資料來源, 那麼 `datetime_now` 表示式將使用 Analytic Server 中的時間; 預設情況下, 此伺服器將使用世界標準時間 (UTC)。
  - 如果目前串流使用資料庫來源節點, 那麼受支援的資料庫將使用 SQL 回送, 並且 `datetime_now` 表示式將使用資料庫時間。
  - 對於所有其他串流, 時區將使用 SPSS Modeler Server 中的時間。
- 如果您選取 **Modeler Client**, 那麼時區將反映安裝了 SPSS Modeler 的機器的時區詳細資料。
- 另外, 您可以對時區選取任何世界標準時間 (UTC) 值。

**儲存為預設值。** 所指定的選項僅適用於目前串流。按一下此按鈕以將這些選項設為所有串流的預設值。

**設定串流的數字格式選項：** 這些選項指定目前串流中的各種數值表示式要使用的格式。

**數字顯示格式。** 您可以選擇標準 (#####.###)、科學 (#.###E+##) 或貨幣顯示格式 (\$###.##)。

**小數位數 (標準、科學和貨幣)。** 針對數字顯示格式, 指定顯示或列印實數時要使用的小數位數。此選項將針對每個顯示格式個別指定。

**計算在。** 選取弧度或角度作為三角的 CLEM 表示式中要使用的測量單位。請參閱第 119 頁的『三角函數』主題, 以取得更多資訊。

**儲存為預設值。** 指定的選項僅套用至現行串流。按一下此按鈕以將這些選項設為所有串流的預設值。

**設定串流的最佳化選項：** 您可以使用「最佳化」設定來最佳化串流效能。請注意, IBM SPSS Modeler Server (如果使用的話) 上的效能和最佳化設定會置換用戶端上的任何相應設定。如果在伺服器中停用了這些設定, 那麼用戶端無法將其啟用。但是, 如果在伺服器中啟用它們, 則用戶端可以選擇將其停用。

註: 資料庫建模和 SQL 最佳化需要在 IBM SPSS Modeler 電腦上啟用 IBM SPSS Modeler Server 連接。通過啟用此設定, 您可以存取資料庫演算法, 直接從 IBM SPSS Modeler 回送 SQL 以及存取 IBM SPSS Modeler Server。要驗證目前授權的狀態, 請從 IBM SPSS Modeler 功能表中選擇下列項目。

說明 > 關於 > 其他詳細資訊

如果啟用了連接, 您可以在「授權狀態」標籤中看到選項**伺服器啟用**。

請參閱第 8 頁的『連線到 IBM SPSS Modeler Server』主題, 以取得更多資訊。

註: 是否受支援 SQL 回送和最佳化, 取決於使用的資料庫類型。有關使用 IBM SPSS Modeler 受支援和測試的資料庫和 ODBC 驅動程式的最新的資訊, 請參閱公司受支援網站 <http://www.ibm.com/support>。

**啟用串流重新寫入。** 選取此選項可啟用 IBM SPSS Modeler 中的串流重寫。提供了四種重寫類型, 您可以選取其中的一種或多種類型。串流重寫入會對串流中的節點在背景進行重新排序, 使其在不更改串流語義的情況下更高效的作業。

- **最佳化 SQL 產生。**利用此選項可對串流中的節點進行重新排序，以便可以使用 SQL 產生回送更多的作業，使其在資料庫中執行。在發現某個節點無法以 SQL 表達時，最佳化器將會向前檢查，確定是否有任何下游節點能夠以 SQL 表達並可以在不影響串流語義的情況下安全地移到問題節點前面。資料庫執行作業不但可以比 IBM SPSS Modeler 效率更高，而且這樣的回送作業可以減小傳回到 IBM SPSS Modeler 進行正在處理的資料集的大小。而這又可以進一步減小網路通信量，加快串流作業的速度。請注意，必須已選取產生 SQL 勾選框後 SQL 最佳化才能生效。
- **最佳化 CLEM 表示式。**此選項啟用最佳化器來搜尋能夠在串流執行之前被預正在處理的 CLEM 表示式，從而增加正在處理速度。舉個簡單的範例，如果您有一個像  $\log(\text{salary})$  這樣的表示式，最佳化器將計算實際工資值並通過傳送來對其進行正在處理。這可以用來提高 SQL 後推和 IBM SPSS Modeler Server 效能。
- **最適語法執行。**這種串流重寫入方法會增加合併多個包含 IBM SPSS Statistics 語法節點的作業的效率。通過將語法指令合併到單個作業中，而不是作為個別的作業分別執行，從而實現最佳化。
- **最佳化其他執行。**這種串流重寫入方法會增加無法委派給資料庫的作業的效率。最佳化是通過盡可能早地減少串流中的資料數量來獲得的。在保持資料完整性的同時，該串流會進行重寫入，以使作業距離資料來源更接近，因而減少了具有較高消耗的作業（如結合）的下游資料。

**啟用平行處理。**當在一台帶有多處理器的電腦上執行時期，此選項使得系統能夠平衡這些處理器之間的負載，從而可能產生最佳的效能。使用多個節點或者使用下列單個節點可能會由於平行處理而受益：C5.0、合併（按鍵進行）、排序、bin（秩法和分位法）以及聚合（使用一個或多個鍵欄位）。

**產生 SQL。**選取此選項可啟用 SQL 產生，從而可以通過使用 SQL 代碼使得串流作業回送到資料庫以產生執行處理，因此可能能夠改善效能。為了進一步改善效能，還可以選取最佳化 SQL 產生，從而實現將回送到資料庫的作業數最大化。節點的作業回送到資料庫之後，該節點將在串流執行時期以紫色強調顯示。

- **資料庫快取。**對於產生了要在資料庫中執行的 SQL 的串流，可以在中遊將資料快取到資料庫中的暫時表格，而不是快取到檔案系統。結合 SQL 最佳化使用時，這可能會造成效能顯著提升。例如，合併多個表格以建立資料採礦視圖的串流中的輸出在必要時可能會加以快取並重複使用。如果已啟用資料庫快取，則只需滑鼠右鍵按一下任意非終端節點就可以快取該點上的資料，並且在下次執行串流時，可以直接在資料庫中自動建立快取。如此一來，為下游節點產生 SQL 將成為可能，從而可進一步提高效能。此外，必要時可停用該選項，例如當策略或權限排除正被寫入資料庫的資料時。如果未啟用資料庫快取或 SQL 最佳化，則快取將會改為寫入檔案系統。如需相關資訊，請參閱主題第 34 頁的『節點的快取選項』。
- **使用寬鬆轉換。**利用此選項可以將資料從字串轉換為數字，或從數字轉換為字串（如果以適當的格式儲存）。例如，如果資料作為字串儲存在資料庫中，但實際包含有意義的數字，那麼可以轉換資料以便在回送時使用。

**註：**由於 SQL 實現的微小差別，在資料庫中執行的串流傳回的結果可能會與在 IBM SPSS Modeler 中執行時期傳回的結果有著輕微的不同。也是因為上述原因，這些差別還可能根據資料庫供應商的不同而有所區別。

**儲存為預設值。** 指定的選項僅套用至現行串流。按一下此按鈕以將這些選項設為所有串流的預設值。

**設定串流的 SQL 日誌和記錄狀態選項：** 這些設定包含用於控制串流產生的 SQL 陳述式的顯示，以及串流處理的記錄數量顯示的各種選項。

**執行串流期間在訊息日誌中顯示 SQL。** 指定在執行串流時產生的 SQL 是否傳送到訊息日誌。

**串流準備期間在訊息日誌中顯示關於 SQL 產生的詳細資料。** 指定在串流預覽期間要產生的 SQL 的預覽是否傳送到訊息日誌。

**顯示 SQL。** 指定日誌中顯示的任何 SQL 是否應該包含格式為 {fn FUNC(...)} 的原生 SQL 函數或標準 ODBC 函數（由 SPSS Modeler 產生）。前者依賴於可能沒有實現的 ODBC 驅動程式功能。

**重新設置 SQL 格式以增加可讀取性。** 指定是否應將在日誌中顯示的 SQL 格式化以提高可讀取性。

**顯示記錄狀態。** 指定在記錄到達終端節點後，報告記錄的時間。指定一個數字，用於表示每  $N$  個記錄更新一次狀態。

**儲存為預設值。** 指定的選項僅套用至現行串流。按一下此按鈕以將這些選項設為所有串流的預設值。

**設定串流的佈置選項：** 這些設定提供了有關串流畫布顯示和使用的眾多選項。

**下限串流畫布寬度。** 指定串流畫布的下限寬度（以像素為單位）。

**下限串流畫布高度。** 指定串流畫布的下限高度（以像素為單位）。

**串流捲動率。** 指定串流畫布的捲動速率，以控制在畫布中拖曳節點時串流畫布窗格的捲動速度。數字越大，捲動速率越高。

**圖示名稱上限長度。** 指定串流畫布中節點名稱的字元限制。

**圖示大小。** 選取適當選項，以便將整個串流視圖放大到標準圖示大小的 8% 至 200% 之間的某個尺寸。

**網格儲存格大小。** 從清單中選取網格儲存格大小。此數字用於使用不可見網格對串流畫布中的節點進行對齊。預設網格蜂巢 Cell 尺寸為 0.25。

**貼齊格線。** 選取此選項將按不可見網格型樣（預設選項）對齊圖示。

**產生圖示的擺放位置。** 選擇由模型塊產生的節點的圖示在畫布中的位置。預設為左上方。

**儲存為預設值。** 指定的選項僅套用至現行串流。按一下此按鈕以將這些選項設為所有串流的預設值。

**Analytic Server 串流內容：** 這些設定提供了用於與 Analytic Server 配合使用的眾多選項。

#### **Analytic Server 之外要處理的上限記錄數**

指定要從 Analytic Server 資料來源匯入 SPSS Modeler 伺服器的上限記錄數。

#### **節點無法在 Analytic Server 中進行處理時的通知**

此設定確定當要提交至 Analytic Server 的串流包含無法在 Analytic Server 中進行處理的節點時將發生的情況。指定要發出警告並繼續正在處理串流，還是擲出錯誤並停止正在處理。

#### **分割模型儲存體設定**

##### **模型大小 (MB) 超過時按照 Analytic Server 上的參照來儲存分割模型**

模型塊通常儲存為串流的一部分。帶有多個分割的分割模型可能產生大塊，在串流和 Analytic Server 之間來回移動塊可能影響效能。作為一種解決方案，當分割模型超出指定大小後，會將其儲存在 Analytic Server 上，而 SPSS Modeler 中的塊包含對該模型的參照。

##### **用於在執行完成後按照 Analytic Server 上的參照來儲存模型的預設資料夾**

指定要在 Analytic Server 上儲存分割模型的預設路徑。該路徑應以有效的 Analytic Server 專案名稱作為起點。

##### **用於儲存升級模型的資料夾**

指定要儲存「升級」模型的預設路徑。SPSS Modeler 階段作業結束後，不會清除升級模型。

**設定串流的地理空間選項：** 任何地理空間欄位都具有相關聯的座標系統，而無論它是形狀、座標還是單個軸值（例如  $X$  或  $Y$  或者緯度和經度）。此座標系統用於設定屬性，例如原點 (0,0) 以及與值相關聯的單位。

存在眾多的座標系統，它們分為兩類：地理座標系統和投影座標系統。SPSS Modeler 中的所有空間函數都只能與投影座標系統配合使用。

座標系統的本質決定了合併或附加來自兩個不同地理空間資料來源的資料時，要求這兩個資料來源使用同一個座標系統。因此，您必須對串流中使用的任何地理空間資料指定座標設定。

在下列狀況下，將自動地重新投影資料以使用選擇的串流座標系統：

- 對於空間函數（例如 area、closeto 和 within），傳送到函數的參數將自動進行重新投影；但是，原始列資料保持不變。
- 在空間-時間預測 (STP) 中使用建立或評分（塊）節點時，將對位置欄位自動進行重新投影。進行評分時，來自於塊的位置是原始位置。
- 使用「地圖視覺化」節點時。

**串流座標系統。**只有在選中勾選框後才可用。請按一下**變更**以顯示一系列可用的投影座標系統，然後選取要用於目前串流的投影座標系統。

**儲存為預設值。**您選取的座標系統將僅應用於目前串流。要選取將此坐標系用作所有串流的預設坐標系，請按一下此按鈕。

**選取地理空間座標系統：** SPSS Modeler 中的所有空間函數都只能與投影座標系統配合使用。

「選取串流座標系統」對話框中有一個座標系統清單，其中包含您可以為串流中使用的任何地理空間資料選取的所有投影座標系統。

對於每個座標系統，列出了下列資訊。

- **WKID** 每個座標系統的唯一知名 ID。
- **名稱** 座標系統的名稱。
- **單位** 與座標系相關聯的測量單位。

除了包含所有座標系統的清單以外，此對話框還提供了過濾控制項。如果您知道所需座標系的完整名稱或部分名稱，請將其輸入到對話框底部的**名稱**欄位中。可供您選擇的一系列座標系統將自動過濾出的，以便僅顯示名稱包含所已輸入文字的座標系統。

## 檢視串流作業訊息

使用「串流內容」對話框中的「訊息」標籤，可以輕鬆檢視有關執行、最佳化和模型建置和評估所用時間等串流作業的訊息。該表格還會報告錯誤訊息。

### 檢視串流訊息

1. 在「檔案」功能表上，按一下**串流內容**（或從管理程式窗格中的「串流」標籤選取串流，用滑鼠右鍵按一下，然後按一下**串流內容**）。
2. 按一下**訊息**標籤。

或者，在「工具」功能表上，按一下：

### 串流內容 > 訊息

除有關串流作業的訊息之外，此處還報告錯誤訊息。當串流執行因錯誤而終止時，此對話框將在「訊息」標籤中開啟，並顯示錯誤訊息。此外，出錯節點將在串流畫布中強調顯示為紅色。

如果在「使用者選項」對話框中啟用了 SQL 最佳化和日誌選項，則還會顯示有關產生的 SQL 的資訊。請參閱第 39 頁的『設定串流的最佳化選項』主題，以取得更多資訊。



可以通過從儲存按鈕下拉清單（位於左側「訊息」標籤的下方）中按一下儲存訊息為串流儲存此處報告的訊息。

可以通過從下拉清單中選取清除全部訊息清除給定串流的所有訊息。

請注意，CPU 時間是伺服器處理程序使用 CPU 的時間長度。經歷時間是執行開始與執行結束之間的總時間，因此還包含傳輸檔案和呈示輸出等操作所耗用的時間。當串流使用了多個 CPU（平行執行）時，CPU 時間可能會大於經歷時間。如果串流完全推回到用作資料來源的資料庫中進行處理，那麼 CPU 時間將為零。

## 檢視節點執行時間

在「訊息」標籤上可以選擇顯示「執行時間」，您可以在其中請參閱執行在 IBM SPSS Modeler Server 上的串流中所有節點的相應執行時間。注意執行在其他區域中的串流的時間可能並不準確，例如 R 或 Analytic Server。

註：要使此功能有效，必須在選項標籤的一般設定中已選取顯示執行時間勾選框。

在節點執行時間表格中，各欄如下所示。按一下欄標題可以按遞增或遞減對各項目進行排序（例如，以便請參閱哪些節點的執行時間最長）。

**終端節點。**節點所屬的分支的 ID。此 ID 是分支末位處的終端節點的名稱。

**節點標記。**執行時間的相關節點的名稱。

**節點 ID。**執行時間的相關節點的唯一 ID。此 ID 在節點建立時由系統產生。

**執行時間。**執行此節點所耗用的時間。

## 設定串流和工作階段參數

可以定義用於 CLEM 表示式和 Script 編寫的參數。這些參數在效果上等同於使用者定義變數，這些變數可以在目前串流、階段作業或 SuperNode 中儲存和保留，並且可以從使用者介面和 Script 中存取。例如，如果儲存串流，則為此串流設定的所有參數也將儲存。（此特點使這些參數區別於僅在 Script 中使用和宣告的本端 Script 變數。）通常將參數用於 Scripting 來控制 Script 的行為，可通過提供無需在 Script 中寫在程式中的欄位和值的相關資訊。

參數的範圍取決於設定參數的位置：

- 串流參數可以在串流 Script 中或在串流內容對話框中設定，且可用於串流中的所有節點。這些參數顯示在表示式建置器中的「參數」清單中。
- 工作階段參數可以在獨立式 Script 中或在工作階段參數對話框中設定。這些階段作業參數可用於現行階段作業中使用的所有串流（即在管理程式窗格的「串流」標籤中列出的所有串流）。

也可以為 SuperNode 設定參數，在這種情況下，這些參數僅對於該 SuperNode 中封裝的節點可見。

通過使用者介面設定串流和工作階段參數

1. 要設定串流參數，請在主功能表中按一下：

工具 > 串流內容 > 參數

2. 要設定工作階段參數，請在「工具」功能表中按一下設定工作階段參數。

提示？。如果您希望在執行時期提示使用者為該參數輸入值，請勾選此方框。

**名稱。** 參數名稱在這裡列出。可以通過在本欄位中輸入名稱來建立新的參數。例如，要為下限溫度建立參數，可以鍵入minvalue。請勿包含表示 CLEM 表示式中的參數的 \$P- 字首。該名稱也用於在 CLEM 表示式建置器中顯示。

**長名稱。** 列出每個所建立參數的描述性統計量名稱。

**儲存體。** 在清單中選取儲存類型。儲存類型表明資料值在參數中如何儲存。例如，當所使用的值包含希望保留的先導 0 時（例如 008），應選取字串作為儲存類型。否則，先導 0 將從值中剝離。有效的儲存類型為字串、整數、實數、時間、日期及時間戳記。對於日期參數，注意其值必須用下一段落所示的 ISO 標準符號指定。

**值。** 列出每個參數的目前的值。根據需要調整參數。注意對於日期參數，其值必須用 ISO 標準符號（即 YYYY-MM-DD）指定。系統不接受以其他格式指定的日期。

**類型（選用）。** 如果計劃將該串流部署到外部應用程式，則從清單中選取測量層次。否則，建議保留 類型欄的值。如果您想為參數指定值限制（如數值型範圍的上限或下限），請從清單中選取指定。

注意只能通過使用者介面為參數設定長名稱、儲存類型和類型選項。不能用 Script 設定這些選項。

按一下位於右側的箭頭可使已選取的參數在可用參數清單中上下移動。使用刪除按鈕（標示為 X）可刪除已選取的參數。

## 指定參數值的執行時期提示

如果串流位於需要在不同情況下為相同參數輸入不同值的位置，您可以為一個或多個串流或階段作業參數值指定執行時期提示。

**參數。**（選用）為參數輸入值，或保留預設值（如有）。

**關閉這些提示。** 如果您不希望這些提示在執行串流時顯示，請選取此複選框。您可通過選取定義參數的「串流內容」或「階段作業內容」對話框中的提示？勾選框來重新顯示這些提示。請參閱第 43 頁的『設定串流和工作階段參數』主題，以取得更多資訊。

## 指定參數類型的值限制

在將串流部署到讀取資料建模串流的外部應用程式期間，您可以啟用參數的值限制。通過此對話框，可以指定執行串流的外部使用者可用的值。對話框中的值限制會呈動態變化，具體取決於資料類型。此處顯示的選項即類型節點的可用選項值。

**類型。** 顯示目前選定的測量層次。您可以根據想要在 IBM SPSS Modeler 中使用參數的方式變更此值。

**儲存體。** 顯示已知的儲存類型。儲存類型不受您選擇在 IBM SPSS Modeler 中採用的測量層次（連續、列名或旗標）影響。可以在「參數」主標籤中更改儲存類型。

根據在類型欄位中選定的測量層次，對話框的下半部分會呈動態變化。

連續測量層次

**低值。** 指定參數值的下限。

**高值。** 指定參數值的上限。

**標籤。** 您可以為範圍欄位的任何值指定標籤。按一下標籤按鈕，可開啟個別對話框來指定值標籤。

列名測量層次

值。使用此選項可指定將用作列名欄位的參數值。值不會在 IBM SPSS Modeler 串流中強制使用，但會在外部部署應用程式的下拉清單中使用。使用箭頭及刪除按鈕，您可以修改現有值以及重新排序或刪除值。

旗標測量層次

**True**。指定符合條件時參數的旗標值。

**False**。指定符合條件時參數的旗標值。

標籤。可以為旗標欄位的值指定標籤。

## 串流部署選項

通過串流內容對話框的「部署」標籤，您可以指定用於在 IBM SPSS Collaboration and Deployment Services 內部署串流以進行模型更新、自動化工作調度或供 IBM Analytical Decision Management 進一步使用的選項。所有串流都需要指定評分分支才能進行部署。請參閱第 137 頁的『儲存和部署儲存庫物件』主題，以取得更多資訊。

## 串流的迴圈執行

使用「串流內容」對話框中的「執行」標籤可以在目前串流中設定迴圈條件來自動執行重複性作業。

設定完這些條件後，可以將其用於簡介 Scripting，因為其使用串流的基本 Scripting 來移入 Script 視窗。隨後，您可以對該基本 Scripting 進行修改 - 可能用作建立更好 Script 的基礎。請參閱第 133 頁的『廣域功能』主題，以取得更多資訊。

設定串流的迴圈

1. 在「檔案」功能表上，按一下串流內容（或從管理程式窗格中的「串流」標籤選取串流，用滑鼠右鍵按一下，然後按一下蹦現功能表上的串流內容）。
2. 按一下執行標籤。
3. 選取迴圈/條件式執行執行方式。
4. 按一下迴圈標籤。

或者，在「工具」功能表上，按一下：

串流內容 > 執行

作為另一種替代方案，也可以用滑鼠右鍵按一下節點，然後從快速功能表中按一下：

迴圈/條件式執行 > 編輯迴圈設定

疊代。您無法編輯此列數字，但是您可以使用表格右側的按鈕來新增、刪除、上移或下移疊代。

表格頭。設定迴圈時，這些迴圈反映了您建立的疊代金鑰和任何疊代變數。

## 檢視串流的廣域值

使用「串流內容」對話框中的「廣域量」標籤可以檢視目前串流的廣域值集合。廣域值使用設定廣域節點建立，用於確定選取的欄位的平均數、合計或標準差等統計資料。

執行「設定廣域」節點後，這些值即可用於執行各式各樣的串流作業。請參閱第 133 頁的『廣域功能』主題，以取得更多資訊。

檢視串流的廣域值

1. 在「檔案」功能表上，按一下**串流內容**（或從管理程式窗格中的「串流」標籤選取串流，用滑鼠右鍵按一下，然後按一下**串流內容**）。
2. 按一下 **廣域** 標籤。

或者，在「工具」功能表上，按一下：

**串流內容 > 廣域**

**可用的廣域屬性。**此表格將列出可用的廣域屬性。您不能在此更改廣域值，但可以使用表格右側的「清除所有值」按鈕清除串流的所有廣域值。

## 搜尋串流中的節點

通過指定一定的搜尋準則（如節點名稱、種類和 ID），可以搜尋串流中的節點。該功能對包含大量節點的複合串流特別有用。

搜尋串流中的節點

1. 在「檔案」功能表上，按一下**串流內容**（或從管理程式窗格中的「串流」標籤選取串流，用滑鼠右鍵按一下，然後按一下**串流內容**）。
2. 按一下「**搜尋**」標籤。

或者，在「工具」功能表上，按一下：

**串流內容 > 搜尋**

可以指定多個選項來限制搜尋範圍，但如果按節點 ID 搜尋（即使用**ID 等於欄位**），將排除其他選項。

**節點標記包含。**選中此勾選框，輸入整個或部分節點標記以搜尋特定節點。搜尋不區分觀察值，且將多個單字視為一段文字。

**節點種類。**選中此勾選框，在清單中按一下某個種類以搜尋特定類型的節點。**處理節點**表示位於節點選用區的「記錄處理」或「欄位作業」標籤上的節點；**套用模型節點**是指一個模型塊。

**關鍵字包含。**選中此勾選框，輸入一個或多個完整關鍵字以搜尋節點，這些節點在節點對話框的「註解」標籤上的「關鍵字」欄位中存在該文字。您輸入的關鍵字文字必須是完全相符。使用分號分隔多個關鍵字，以搜尋選項（例如，輸入 proton;neutron 將尋找帶有這些關鍵字的所有節點。請參閱第 53 頁的『註解』主題，以取得更多資訊。

**註解包含。**選中此勾選框，輸入一個或多個單詞，以搜尋那些在其節點對話框「註解」標籤上的主文字區域中包含此文字的節點。搜尋不區分觀察值，且將多個單字視為一段文字。請參閱第 53 頁的『註解』主題，以取得更多資訊。

**產生名稱如下的欄位。**勾選此方框，輸入某個產生欄位的名稱，例如 \$C-Drug。可以使用此選項來搜尋產生某個特定欄位的建模節點。只輸入一個必須是完全相符的欄位名稱。

**ID 等於。**選中此勾選框，輸入節點 ID 以搜尋具有此 ID 的特定節點（選取此選項將取消前面的所有選項）。節點 ID 是在建立節點時由系統分配的，可用於參照節點以編寫 Script 或實現自動化。只輸入一個必須是完全相符的節點 ID。請參閱第 53 頁的『註解』主題，以取得更多資訊。

**在 SuperNode 中搜尋。**此複選框為預設已勾選，表示在 SuperNode 內外的節點上執行搜尋。如果只打算在串流的最高級處的 SuperNode 外面的節點上執行搜尋，請取消選中此框。

**尋找。**在指定了所有的選項後，按一下此按鈕開始搜尋。

與指定選項相符的節點在對話框下方列出。從清單中選取節點並在串流畫布中強調顯示此節點。

## 更名串流

使用「串流內容」對話框中的「註解」標籤，可以為串流新增描述性統計量註解並建立自訂名稱。這些選項在針對新增至專案窗格的串流產生報告時特別有用。請參閱第 53 頁的『註解』主題，以取得更多資訊。

## 串流說明

對於您建立的每個串流，IBM SPSS Modeler 會為其生成包含串流內容相關資訊的串流說明。如果您想知道串流的功能但又未安裝 IBM SPSS Modeler，例如通過 IBM SPSS Collaboration and Deployment Services 存取串流時，這非常有用。

串流說明以 HTML 文件格式顯示，其中包含多個部分。

### 一般串流資訊

該部分包含串流名稱，以及建立和上次儲存串流時的詳細資料。

### 說明與備註

這部分包含任何：

- 串流註釋（請參閱第 53 頁的『註解』）
- 未連線到特定節點的備註
- 連接至串流的建模和評分分支中的節點的備註

### 評分資訊

該部分包含有關串流的評分分支的各種標題下的資訊。

- **備註。** 包含僅連接到評分分支中節點的備註。
- **輸入。** 列出輸入欄位及其儲存類型（例如，字串、整數、實數等等）。
- **輸出。** 列出輸出欄位及其儲存類型，包含建模節點產生的其他欄位。
- **參數。** 列出任何與串流的評分分支相關的參數，以及每次模型評分時可以檢視或編輯的參數。按一下串流內容對話框的部署標籤上的評分參數按鈕，可以 ID 這些參數。
- **模式節點。** 顯示模型名稱和類型（例如，類神經網路以及 C&R 樹狀結構等等）。這是在串流內容對話框的部署標籤上為模式節點選取的模型塊。
- **模型詳細資料。** 在各種標題下顯示 ID 的模型塊詳細資料。如有可能，還併入模型的預測值重要性和評估表。

### 建模資訊

包含串流的建模分支相關資訊。

- **備註。** 列出任何連接至建模分支內節點的備註或註解。
- **輸入。** 列出輸入欄位及其在建模分支中的角色（以欄位角色值形式，例如，「輸入」、「目標」、「分割」等）。
- **參數。** 列出任何與串流的建模分支相關的參數，以及每次模型更新時可以檢視或編輯的參數。按一下串流內容對話框的部署標籤上的模型建立參數...按鈕，可以 ID 這些參數。
- **建模節點。** 顯示用於產生或更新模型的建模節點的名稱和類型。

## 預覽串流說明

您可以在串流內容對話框按下一個選項，從而在 Web 瀏覽器中檢視串流說明的內容。說明內容取決於您在對話框的「部署」標籤上所指定的選項。請參閱第 149 頁的『串流部署選項』主題，以取得更多資訊。

要檢視串流說明，請執行下列操作：

1. 在主 IBM SPSS Modeler 功能表中，按一下：

**工具 > 串流內容 > 部署**

2. 設定部署類型、指定的評分節點以及任何評分參數。
3. 如果部署類型為「模型更新」，則可以選取：
  - 建模節點以及任何模型建立參數...
  - 在串流的評分分支上的模型塊
4. 按一下**預覽串流說明**按鈕。

## 匯出串流說明

可以將串流說明內容匯出到 HTML 檔案。

要匯出串流說明：

1. 在主功能表上，按一下：

**檔案 > 匯出串流說明**

2. 輸入 HTML 檔案名並按一下 **儲存**。

## 執行串流

指定串流的所需選項並連接所需節點後，便可以通過在串流中的節點之間執行資料來執行串流。有多種在 IBM SPSS Modeler 中執行串流的方法。您可以：

- 在「工具」功能表中，按一下**執行**。
- 在工具列上按一下其中一個**執行...**按鈕。使用這些按鈕，您可以執行整個串流，也可以僅執行所選的終端節點。請參閱第 15 頁的『IBM SPSS Modeler 工具列』主題，以取得更多資訊。
- 通過用滑鼠右鍵按一下某個終端節點並在蹦現功能表中按一下**執行**來執行單個資料串流。
- 通過用滑鼠右鍵按一下任意非終端節點並在蹦現功能表中按一下**從這裡執行**來執行部分資料串流。這將只執行那些在所選節點之後的作業。

若要中止執行進行中的串流，可以按一下工具列中的紅色「停止」按鈕，也可以在「工具」功能表中按一下**停止執行**。

如果任何串流的執行時期間超過三秒鐘，便會顯示「執行回饋」對話框以指示進度。

某些節點會進一步顯示有關串流執行的更多資訊。方法為在對話框中選取對應的列。系統將自動已選取第一列。

## 使用模型

如果串流包含一個建模節點（即，來自節點選用區的「建模」或「資料庫建模」標籤的節點），則在串流執行時期會建立**模型塊**。模型塊是模型的儲存器，即規則、公式或方程式的集合，允許您針對來源資料產生預測，並且位於預測分析的核心。



圖 12. 模型區塊

當您成功執行建模節點時，對應的模型塊被置於串流畫布，由金色菱形圖示代表（因此名為「塊」）。您可以開啟塊並瀏覽其內容以檢視模型的詳細資料。要檢視預測，可附加並執行一個或多個終端節點，其輸出將以可讀取的形式顯示預測。

一般的建模串流由兩個分支組成。**建模分支**包含建模節點以及在其之前的來源節點和處理節點。**評分分支**在執行建模節點時建立，包含模型塊和終端節點，或是用於檢視預測的節點。

有關進一步資訊，請參閱《IBM SPSS Modeler 建模節點》手冊。

### 新增備註和註解到節點和串流

您可能需要向組織中的其他人說明串流。為了幫您進行，您可以將解釋備註附加到串流、節點和模型塊中。

其他人然後可以在畫面上檢視這些備註，或者您可以列印包含備註的串流影像。

您可以列出串流或 SuperNode 的所有備註，變更清單中備註的順序，編輯註解文字和變更備註的前景色或背景顏色。請參閱第 52 頁的『列出串流備註』主題，以取得更多資訊。

您還可以通過串流內容對話框、節點對話框或模型塊視窗的「附註」標籤以文字附註的形式將附註新增到串流、節點和塊中。只有當「註解」標籤開啟時才能看到這些備註，例外情況是串流註解也可顯示為畫面上的備註。請參閱第 53 頁的『註解』主題，以取得更多資訊。

### 備註

備註採用文字框的形式，在此可以輸入任何數量的文字，還可新增任何數量的備註。備註可以是獨立的（不附加到任何串流物件），或者可以連接至串流中的一個或多個節點或模型塊。獨立備註通常用於說明串流的整體目的；連接的備註說明附加到的節點或塊。節點和塊可以附加多個備註，串流可以有任何數量的獨立備註。

註：雖然無法將串流註解附加到節點或模型塊，但可將串流註解顯示為畫面上的備註。請參閱第 52 頁的『將註解轉換為備註』主題，以取得更多資訊。

文字框的外觀變更為指示目前備註模式（或顯示為備註的註解），如下表格所示。

表 3. 註解及註釋文字框模式

註解文字框	註解文字框	模式	表示	獲得方式...
		編輯	備註已開啟，可進行編輯。	新建備註或註解，或按兩下現有備註或註解。
		上次選定的	可以移動或刪除備註，也可調整其大小。	在編輯後按一下串流背景，或按一下現有備註或註解。
		視圖	編輯完成。	編輯後按一下另一節點、備註或註解。

當您建立新獨立備註時，最初顯示在串流畫布的左上角落。

如果您將備註附加到節點或塊中，備註最初顯示在附加串流物件的上方。

文字框顯示為白色，以顯示可以已輸入文字。當您已輸入文字時，按一下文字框的外部。備註背景變為黃色，以顯示文字輸入完成。一直已選取備註，您可以對其進行移動、刪除或調整大小。

再次按一下時，邊框變為實行以表示編輯完成。

按兩下備註將文字框變更為編輯模式 -- 背景變為白色，註解文字可以編輯。

您還可以將備註附加到 SuperNode。

**涉及備註的作業：** 您可以對備註執行大量作業。您可以：

- 新增獨立備註
- 將備註附加到節點或塊
- 編輯備註
- 調整備註大小
- 移動備註
- 切斷備註連接
- 刪除備註
- 顯示或隱藏串流的所有備註

#### 新增獨立備註

1. 確保串流上未選取任何項目。
2. 請執行下列其中一項：
  - 在主功能表上，按一下：  
**插入 > 新增註解**
    - 在串流背景上用滑鼠右鍵按一下，然後按一下蹦現功能表上的**新建註解**。
    - 按一下工具列中的**新建註解**按鈕。
3. 輸入註解文字（或從剪貼簿貼上文字）。
4. 在串流中按一下節點以儲存備註。

#### 將備註附加到節點或塊

1. 在串流畫布上選取一個或多個節點或塊。
2. 請執行下列其中一項：
  - 在主功能表上，按一下：  
**插入 > 新增註解**
    - 在串流背景上用滑鼠右鍵按一下，然後按一下蹦現功能表上的**新建註解**。
    - 按一下工具列中的**新建註解**按鈕。
3. 輸入註解文字。
4. 在串流中按一下另一節點以儲存備註。

您也可以：



5. 插入獨立備註（請參閱上一節）。
6. 請執行下列其中一項：
  - 選取註解，按 F2，然後選取節點或區塊。
  - 選取節點或區塊，按 F2，然後選取註解。
  - （僅限三個按鈕的滑鼠）將滑鼠指標移至註解上，按住中間按鈕，並將滑鼠指標拖曳至節點或區塊上，然後鬆開滑鼠按鈕。

#### 將備註附加到其他節點或塊

如果備註已附加到節點或塊，或者如果目前在串流層次，您想將其附加到其他節點或塊，進行下列其中一個操作：

- 選取註解，按 F2，然後選取節點或區塊。
- 選取節點或區塊，按 F2，然後選取註解。
- （僅限三個按鈕的滑鼠）將滑鼠指標移至註解上，按住中間按鈕，並將滑鼠指標拖曳至節點或區塊上，然後鬆開滑鼠按鈕。

#### 編輯現有備註

1. 請執行下列其中一項：
  - 按兩下註解文字框。
  - 選中文字框，然後按下 Enter 鍵。
  - 用滑鼠右鍵按一下文字框顯示其功能表，然後按一下「編輯」。
2. 編輯註解文字。您可以在編輯時使用標準 Windows 快速鍵，例如使用 Ctrl+C 複製文字。在蹦現功能表中列出編輯期間的其他備註選項。
3. 按一下一次文字框外部，以顯示調整大小控制項，然後再次完成備註。

#### 調整註解文字框大小

1. 選取備註，以顯示調整大小控制項。
2. 按一下並拖曳控制項以調整文本框大小。
3. 按一下文字框外部，儲存變更。

#### 移動現有備註

如果您想移動備註，而非其附加的物件（如有），進行下列其中一個操作：

- 在備註上移動滑鼠指標，按下左側滑鼠按鈕，將備註拖至新位置。
- 選取備註，按下 Alt 鍵，使用方向鍵移動備註。

如果您想移動備註連同任何附加備註的節點或塊：

1. 選取您想移動的所有物件。
2. 請執行下列其中一項：
  - 將滑鼠指標移至其中一個物件上方，按下滑鼠左鍵，將物件拖至新位置。
  - 選取一個物件，按下 Alt 鍵，使用方向鍵移動物件。

#### 將備註從節點或塊中斷連線

1. 選取要中斷連線的一個或多個備註。
2. 請執行下列其中一項：

- 按 F3 鍵。
- 用滑鼠右鍵按一下選定備註並從其功能表中按一下「切斷」。

#### 刪除備註

1. 選取要刪除的一個或多個備註。
2. 請執行下列其中一項：
  - 按 Delete 鍵。
  - 用滑鼠右鍵按一下選定備註並從其功能表中按一下「刪除」。

如果備註附加到節點或塊，連線行也被刪除。

如果備註原來為串流或 SuperNode 註解，且已轉換成獨立備註，備註將從畫布中刪除，但其文字保留在串流或 SuperNode 的「註解」標籤中。

#### 顯示或隱藏串流的備註

1. 請執行下列其中一項：
  - 在主功能表上，按一下：

##### 檢視 > 備註

- 按一下工具列中的顯示/隱藏備註按鈕。

**列出串流備註：** 您可以檢視已為某個特定串流或 SuperNode 添加的所有備註的清單。

在此清單上，您可以

- 變更備註的順序
- 編輯註解文字
- 變更備註的前景色或背景顏色

#### 列出備註

要列出已為串流添加的備註，請執行下列一種操作：

- 在主功能表上，按一下：

##### 工具 > 串流內容 > 備註

- 在管理程式窗格中用滑鼠右鍵按一下串流，並按一下串流內容，然後按一下備註。
- 用滑鼠右鍵按一下畫布上的串流背景，並按一下串流內容，然後按一下備註。

**文字。** 備註文字。按兩下文字將欄位變更為可編輯文字框。

**鏈結。** 要將備註附加到的節點的名稱。如果該欄位為空白，則備註套用到串流。

**定位按鈕。** 這些按鈕可在清單中上下移動已選取的備註。

**備註顏色。** 要變更備註的前景色或背景顏色，選取備註，選取自訂顏色複選框，然後從背景或前景清單（或兩個清單）中選取一種顏色。按一下套用，然後按一下串流背景請參閱變更效果。按一下確定以儲存變更。

**將註解轉換為備註：** 可將對串流或 SuperNode 的註解轉換為備註。

如果是串流，則註解轉換為串流畫布上的獨立備註（即不附加到任何節點）。

SuperNode 註解轉換為備註時，備註不附加到串流畫布上的 SuperNode，但在放大到 SuperNode 時可見。

#### 轉換串流註解為備註

1. 在「工具」功能表中，按一下**串流內容**。（或者，您可以在管理程式窗格中的串流上用滑鼠右鍵按一下，然後按一下**串流內容**。）
2. 按一下**註釋標籤**。
3. 選取將註釋顯示為註解勾選框。
4. 按一下**確定**。

#### 轉換 SuperNode 註解為備註

1. 按兩下畫布上的 SuperNode 圖示。
2. 按一下**註釋標籤**。
3. 選取將註釋顯示為註解勾選框。
4. 按一下**確定**。

### 註解

可以採用多種方式為節點、串流和模型添加註解。您可以新增描述性統計量註解並指定自訂名稱。這些選項在針對新增至專案窗格的串流產生報告時特別有用。對於節點和模型塊，還可以新增工具提示文字，以說明識別串流畫布中的相似節點。

#### 新增註解

編輯節點或模型塊時將開啟一個帶有標籤的對話框，其中包含用於設定各式各樣註解選項的「註解」標籤。您也可以直接開啟「註解」標籤。

1. 要為節點或模型塊添加註解，請用滑鼠右鍵按一下串流畫布中的節點或模型塊，然後按一下**更名並加註解**。此時將開啟「編輯」對話框並顯示「註解」標籤。
2. 要為串流添加註解，請在「工具」功能表中按一下**串流內容**。（或者，您可以在管理程式窗格中的串流上用滑鼠右鍵按一下，然後按一下**串流內容**。）按一下「註解」標籤。

**名稱**。選取自訂可調整自動產生的名稱，或建立在串流畫布中顯示的唯一節點名稱。

**工具提示文字**。（僅適用於節點和模型塊）輸入用作串流畫布上工具提示的文字。使用大量相似節點時，此功能特別有用。

**關鍵字**。指定要在專案報告中以及在搜尋串流中節點時或追蹤儲存在儲存庫中的物件時使用的關鍵字（參閱第 137 頁的『關於 IBM SPSS Collaboration and Deployment Services 儲存庫』）。可使用分號分隔多個關鍵字 - 例如 income; crop type; claim value。將去除每個關鍵字啟動和結尾處的空格 - 例如，income ; crop type 與 income;crop type 得到的結果相同。（但是，關鍵字中的空格不會去除。例如，具有一個空格的 crop type 和有兩個空格的 crop type 不同。）

主文字區域可用於輸入有關節點作業或節點中所做決策的詳細註解。例如，在共用和重用串流時，對某些決策（如使用過濾器節點捨棄包含許多空白值的欄位）進行註解將很有用。為節點添加註解會將此類資訊與節點一起儲存。您也可以選擇將這些註解包含在從專案窗格建立的專案報告中。請參閱第 157 頁的『專案簡介』主題，以取得更多資訊。

**註解顯示為備註**。（僅適用於串流和 SuperNode。註解）選中此框將註解轉換為在串流畫布上可見的獨立備註。請參閱第 49 頁的『新增備註和註解到節點和串流』主題，以取得更多資訊。

**ID。** 顯示唯一的 ID，該 ID 可用於參照節點以編寫 Script 或實現自動化。此值將在建立節點時自動產生，且不會變更。另請注意，為避免與字母 "O" 的混淆，節點 ID 中不使用數字零。使用右側的複製按鈕可根據需要將 ID 複製並貼上到 Script 中或其他位置。

## 儲存資料串流

建立串流後，可以對其進行儲存以備日後繼續使用。

### 儲存串流

1. 在「檔案」功能表中按一下**儲存串流或串流另存新檔**。
2. 在「儲存」對話框中，瀏覽至要在其中儲存串流檔的資料夾。
3. 在「檔案名稱」文字框中輸入串流的名稱。
4. 如果要將儲存的串流新增到目前專案，請選中**新增到專案**。

按一下 **儲存** 會在指定目錄中以副檔名 `*.str` 儲存串流。

**自動備份檔。** 每次儲存串流時，都會將以前儲存的檔案版本自動保留為副本，並為檔名附加一個連字號（例如 `mystream.str-`）。要還原已備份的版本，只需刪除連字號並重新開啟檔案即可。

### 儲存狀態

除串流以外，您還可以儲存狀態，其中包含目前顯示的串流圖和所有已建立模型塊（在管理程式窗格的「模型」標籤中列出）。

#### 儲存狀態

1. 在「檔案」功能表中，按一下：  
**狀態 > 儲存狀態或狀態另存為**
2. 在「儲存」對話框中，瀏覽至要在其中儲存狀態檔案的資料夾。

按一下**儲存** 會在指定目錄中以副檔名 `*.cst` 儲存狀態。

### 儲存節點

您也可以儲存單個節點，方法是用滑鼠右鍵按一下串流畫布中的節點，然後在蹦現功能表中按一下**儲存節點**。使用副檔名 `*.nod`。

### 儲存多個串流物件

如果在多個物件（如串流、專案或模型塊）未儲存的情況下結束 IBM SPSS Modeler，系統將提示您在完全關閉軟體之前進行儲存。如果選擇儲存項目，則會開啟一個對話框，顯示儲存每個物件的選項。

1. 只需選中要儲存物件的勾選框即可。
2. 按一下**確定**在所需的位置儲存每個物件。

隨後系統將為您顯示每個物件的標準「儲存」對話框。在儲存完成後，應用程式將會關閉。

### 儲存輸出

通過 IBM SPSS Modeler 輸出節點產生的表格、圖形和報告可以採用輸出物件（`*.cou`）格式進行儲存。

1. 檢視要儲存的輸出時，請在輸出視窗功能表中按一下：

**檔案 > 儲存**

2. 指定輸出檔案的名稱和位置。

3. (選用) 選取「儲存」對話框中的**將檔案新增到專案**可以在目前專案中包含該檔案。請參閱第 157 頁的『專案簡介』主題，以取得更多資訊。

或者，也可以用滑鼠右鍵按一下管理程式窗格中列出的任意輸出物件，然後在蹦現功能表中選取**儲存**。

## 加密和解密資訊

在儲存串流、節點、專案、輸出檔案或模型塊時，您可以對其進行加密以防止有人在未經授權的情況下使用它。為此，應在儲存時選取一個額外選項，並為要儲存的項目新增密碼。可以為您儲存的任何項目設定密碼，從而進一步增加其安全性；此加密法與在 IBM SPSS Modeler 與 IBM SPSS Modeler Server 之間傳送檔案時使用的 SSL 加密不同。

當您嘗試開啟加密項目時，系統會提示您輸入密碼。輸入正確的密碼後，項目將自動解密並正常開啟。

### 加密項目

1. 在「儲存」對話框中，針對要加密的項目按一下**選項**。此時將開啟「加密選項」對話框。
2. 選中 **加密此檔案**。
3. (選用) 為提高安全性，選用中 **遮罩密碼**。此操作會將您輸入的所有內容顯示為一連串的点。
4. 輸入密碼。警告：如果忘記密碼，則無法開啟檔案或模型。
5. 如果已選取**遮罩密碼**，請再次已輸入密碼以確保已輸入無誤。
6. 按一下**確定**傳回「儲存」對話框。

註：如果儲存任何加密受保護的項目的副本，除非您變更「加密選項」對話框中的設定，否則新項目將自動採用原有密碼以加密格式儲存。

## 載入檔案

可以在 IBM SPSS Modeler 中重新載入以下幾種已儲存的物件：

- 串流 (. str)
- 狀態 (. cst)
- 模型 (. gm)
- 模型選用區 (. gen)
- 節點 (. nod)
- 輸出 (. cou)
- 專案 (. cpj)

## 開啟新檔案

可以直接通過「檔案」功能表載入串流。

- 在「檔案」功能表中，按一下**開啟串流**。

使用「檔案」功能表中提供的子功能表項目可以開啟所有其他類型的檔案。例如，要載入模型，可以在「檔案」功能表中按一下：

**模型 > 開啟模型或載入模型選用區**

## 開啟最近使用的檔案

為快速載入最近使用的檔案，可以使用「檔案」功能表底部的選項。

選取最近串流、最近專案或最近狀態可展開最近使用的檔案清單。

## 對映資料串流

使用對映工具，可以將新資料來源連接到先前存在的串流。對映工具不僅會設定連線，還將協助您指定新來源中的欄位如何取代現有串流中的欄位。您只需連接到現有的串流，而不必為新資料來源重新創建整個資料串流。

使用資料對映工具，可以將兩個串流段結合在一起，並確保所有（必要）欄位的名稱均正確相符。實際上，對映資料將直接導致建立新的過濾器節點，該操作會通過更名與相應欄位相符。

以下兩種方法都可以對映資料：

**選取取代節點。**此方法從要取代的節點入手。首先，用滑鼠右鍵按一下要取代的節點；然後，使用蹦現功能表中的資料對映 > 選取取代節點選項，選取要用於取代該節點的節點。

**對映至。**此方法從要引入串流中的節點入手。首先，用滑鼠右鍵按一下要引入的節點；然後，使用蹦現功能表中的資料對映 > 對映至選項，選取它要結合到的節點。此方法通常適用於對映到終端節點。註：不能對映至合併節點或附加節點，而只需採用常態方式將串流連接到合併節點即可。

資料對映已緊密整合到串流的建立過程中。如果嘗試連接至已具有連線的節點，您可以選擇取代連線或對映至該節點。

## 將資料對映到範本

要將某個範本串流的資料來源取代為將您自己的資料帶入 IBM SPSS Modeler 的新來源節點，應使用「資料對映」蹦現功能表中的選取取代節點選項。此選項可用於所有節點，但合併節點、「聚合」節點和所有終端節點除外。使用資料對映工具執行此作業有助於確保現有串流作業與新資料來源之間的欄位正確符合。下列步驟簡單介紹資料對映過程。

**步驟 1：指定原始來源節點中的必要欄位。**為了正確執行串流作業，需要指定必要欄位。請參閱第 57 頁的『指定必要欄位』主題，以取得更多資訊。

**步驟 2：將新資料來源新增到串流畫布。**使用某個來源節點帶入新的取代資料。

**步驟 3：取代範本來源節點。**使用範本來源節點的蹦現功能表中的「資料對映」選項，按一下選取取代節點，然後選取取代資料的來源節點。

**步驟 4：檢查對映的欄位。**在開啟的對話框中，檢查軟體是否正從取代資料來源中的欄位正確對映到串流中的欄位。所有未對映的必要欄位都顯示為紅色。這些欄位用於串流作業中，並且必須取代為新資料來源中的類似欄位，以正常進行下游作業。請參閱第 57 頁的『檢查對映的欄位』主題，以取得更多資訊。

使用對話框確保所有必要欄位都正確對映後，原有資料來源將中斷連線，新資料來源將使用稱為對映的「過濾器」節點連接至串流。該過濾器節點可在串流中實現真正的欄位對映。串流畫布中還併入取消對映 過濾器節點。取消對映 過濾器節點可用於取消欄位名稱對映，方法是將其新增到串流中。該節點將復原已對映的欄位，但請注意，您必須編輯所有下游終端節點才能重新選取欄位和套版欄位。

## 串流之間的對映

與連接節點類似，這種資料對映方法不要求您預先設定必要欄位。採用這種方法，只需使用「資料對映」蹦現功能表中的對映至從一個串流連接到另一個串流即可。對於對映至終端節點和在串流之間複製和貼上的操作，這種類型的資料對映十分有用。註：使用對映至選項時，不能對映至合併節點、附加節點和所有類型的來源節點。

在串流之間對映資料

1. 用滑鼠右鍵按一下要用於連接到新串流的節點。
2. 在功能表上，按一下：

資料對映 > 對映至

3. 使用游標選取目標串流中的目的地節點。
4. 在開啟的對話框中確保欄位完全符合，然後按一下**確定**。

## 指定必要欄位

通常，對映至現有串流時，串流創建者會指定必要欄位。這些必要欄位指出是否在下游作業中使用某個特定欄位。例如，現有串流可能會建置一個使用流動欄位的模型。在該串流中，流動是必要欄位，因為不能在沒有該欄位的情況下建立模型。同樣，用於操控節點（如「衍生」節點）的欄位是衍生新欄位的必要欄位。將此類欄位明確設定為必要欄位有助於確保將新來源節點中的相應欄位對映至這些欄位。如果未對映某些必備欄位，您將收到錯誤訊息。如果確定不必要使用某些操作或輸出節點，則可以將這些節點從串流中刪除，並將相應欄位從「必要欄位」清單中刪除。

設定必要欄位

1. 用滑鼠右鍵按一下將取代的範本串流的來源節點。
2. 在功能表上，按一下：

資料對映 > 指定必要欄位

3. 使用「欄位選擇器」，可以在清單中新增或刪除欄位。要開啟「欄位選擇器」，請按一下欄位清單右側的圖示。

## 檢查對映的欄位

選定要將一個資料串流或資料來源對映至另一個資料串流或資料來源的位置後，將顯示一個對話框，您可以在其中選取對映欄位或確保系統預設值對映正確。如果已為串流或資料來源設定了必要欄位，但它們不符合，這些欄位將顯示為紅色。資料來源中的所有未對映欄位將以原樣通過過濾器節點，但請注意，您也可以對映非必要欄位。

**原始。**列出範本或現有串流中的所有欄位，即顯示在較下游位置的所有欄位。新資料來源中的欄位將對映至這些欄位。

**已對映。**列出已選取用於對映至範本欄位的欄位。這些欄位的名稱可能必須進行變更，以與串流作業中使用的原始欄位比對。按一下某個欄位的表格資料格可啟動包含可用欄位；的清單。

如果不確定要對映哪些欄位，在對映之前仔細檢查來源資料會很有用。例如，可以使用來源節點中的「類型」標籤檢查來源資料的摘要。

---

## 提示和快捷鍵

通過熟悉下列快捷鍵和提示，您可以快速而輕鬆地操作：

- 通過按兩下快速建立串流。只需按兩下選用區中的某個節點，即可新增該節點並將其連接到目前串流。
- 使用組合鍵選取下游節點。按 Ctrl+Q 和 Ctrl+W 可切換所有下游節點的選擇。
- 使用快捷鍵對節點進行連接和中斷連線。當在畫布中已選取某個節點時，按 F2 可開始連線，按 Tab 可移至所需的節點，按 Shift+空格可完成連線。按 F3 可切斷對所選節點的所有輸入與輸出連接。
- 在「節點選用區」標籤中自訂常用節點。在「工具」功能表中按一下**管理選用區(M)...**可開啟一個對話框，用於新增、刪除或移動「節點選用區」中顯示的節點。
- 更名節點和新增工具提示。每個節點對話框都包含一個「註解」標籤，您可在其中為畫布中的節點指定自訂名稱，以及新增工具提示以協助您管理串流。此外，還可以添加詳細註解以追蹤進度、儲存過程詳細資料，以及表示所需或已實現的任何商業決策。
- 自動在 **CLEM** 表示式中插入值。使用「表示式建置器」（可通過用於衍生節點和充填節點等的對話框存取此工具），可以在 CLEM 表示式中自動插入欄位值。可以按一下「表示式建置器」中的值按鈕在現有欄位值中選擇。



圖 13. 「值」按鈕

- 快速瀏覽檔案。在「開啟」對話框中瀏覽檔案時，使用「檔案」清單（按一下對話框頂部「查找」欄位旁的黃色菱形按鈕）可存取先前使用的目錄，以及 IBM SPSS Modeler 預設目錄。使用向前和向後按鈕可捲動瀏覽存取的目錄。
- 最小化輸出視窗群。使用所有輸出視窗右上角的紅色 X 按鈕可快速關閉並刪除輸出。這樣，您可以在管理程式窗格的「輸出」標籤中僅保留所期望或感興趣的結果。

全部鍵盤快捷鍵都適用於此軟體。請參閱第 182 頁的『鍵盤協助工具』主題，以取得更多資訊。

您是否知道可以...

- 使用滑鼠拖曳並選取串流畫布中的一群組節點。
- 將節點從一個串流複製並貼上到另一個串流。
- 通過每個對話框和輸出視窗存取「說明」。
- 獲得有關 CRISP-DM（跨行業資料採礦標準處理程序）的說明。（在「說明」功能表中按一下 **CRISP-DM** 說明。）



---

## 第 6 章 輸出的處理

當您執行部分串流時，可透過模型塊節點的**模型標籤**或**進階標籤**，在「檢視器」中檢視結果。在「檢視器」中，您可以輕易地導覽您想要查看的輸出。您也可以處理輸出，建立一份只包含您所需之輸出資料的文件。部分圖形輸出也可以使用「檢視器」。

「檢視器」可用於 IBM SPSS Modeler 中的下列輸出中：

- TCM 模型塊
- STP 模型塊
- TwoStep-AS 叢集模型塊
- GSAR 模型塊
- 地圖視覺化圖形節點

---

### 檢視器

結果會出現在「檢視器」中。您可以利用「檢視器」，執行下列動作：

- 瀏覽結果
- 顯示或隱藏選用的表格、圖表
- 移動選用項目，以變更結果的顯示順序
- 在「檢視器」和其他應用程式之間，移動項目

「檢視器」包含兩個框框：

- 左框包含內容的概要輸出。
- 右框則是放置統計表格、圖表、文字輸出。

您可以按一下概要中的項目，直接跳到對應的表格或圖表上。您可以按一下大綱窗格的右邊框，再予以拖曳，即可改變大綱窗格的寬度。

### 顯示和隱藏結果

在「檢視器」中，您可以選擇性地顯示或隱藏整個程序的個別表格跟結果。這個程序可以大幅減少在目錄框中顯示的輸出量。

### 若要隱藏表格和圖表

1. 在「檢視器」的大綱窗格中連按兩下有書本圖樣的項目。

或

2. 按一下項目以表示選取。
3. 在功能表上，選擇：

檢視 > 隱藏

或

4. 按一下「概要」工具列上，合起來的書本（「隱藏」）圖像。

打開的書本（「顯示」）圖像會變成作用中圖像，代表項目現在已經隱藏起來。

### 若要隱藏程序結果

1. 在大綱窗格中，按一下程序名稱左側的方框。

這樣就會隱藏程序中所有的結果，並收合概要檢視。

### 移動、刪除和複製輸出

您可以利用複製、移動、或刪除項目或項目群組的方式，來重新排列結果。

#### 若要移動「檢視器」中的輸出

1. 選取概要或目錄框中的項目。
2. 將選取的項目拖放到不同的位置。

#### 若要刪除「檢視器」中的輸出

1. 選取概要或目錄框中的項目。
2. 按下「刪除」鍵。

或

3. 在功能表上，選擇：

**編輯 > 刪除**

### 變更起始對齊方式

依預設，所有結果一啟動都是向左對齊的。若要變更新輸出項目的起始對齊方式：

1. 在功能表上，選擇：

**編輯 > 選項**

2. 按一下「檢視器」標籤。
3. 在「初始輸出狀態」組別中，選取項目類型（例如，樞紐表、圖表、文字輸出）。
4. 選取您要的對齊方式選項。

### 變更輸出項目的對齊方式

1. 在概要或目錄框中，選取您要對齊的項目。
2. 在功能表上，選擇：

**格式 > 靠左對齊**

或

**格式 > 置中**

或

**格式 > 靠右對齊**

## 檢視器概要

大綱窗格會提供「檢視器」裡面的文件目錄。您可以使用大綱窗格，瀏覽不同的結果，並且控制其顯示方式。您在大綱窗格所執行（大部分）的動作，目錄框會隨之反應出來。

- 在大綱窗格中選擇項目，就會在目錄框中顯示相對應的項目。
- 如果在大綱窗格中移動項目，則目錄框中的相對應項目，會隨之移動。
- 如果合併概要輸出，則合併階層中，所有項目的結果都會被隱藏起來。

控制大綱顯示。若要控制概要的顯示方式，您可以：

- 展開、合併概要輸出
- 變更選用項目的概要階層
- 變更概要顯示中項目的大小
- 變更概要顯示中所使用的字型

## 若要收合及展開概要輸出

1. 在您想收合或展開之概要項目，其左側的方框上按一下。

或

2. 按一下概要中的項目。
3. 在功能表上，選擇：

檢視 > 收合

或

檢視 > 展開

## 如何變更概要階層

1. 按一下大綱窗格中的項目。
2. 在功能表上，選擇：

編輯 > 大綱 > 升級

或

編輯 > 大綱 > 降級

## 若要變更概要項目的大小

1. 在功能表上，選擇：

視圖 > 大綱大小

2. 選取概要大小（「小」、「中」或「大」）。

## 若要變更概要中的字型

1. 在功能表上，選擇：

視圖 > 大綱字型...

2. 選取字型。

## 新增項目至檢視器

您可以在「檢視器」中，加入很多項目，像是標題、新文字、圖表、或其他應用程式的資料等。

### 若要新增標題或文字

您可以在「檢視器」中，加入新的文字項目，該項目跟表格或圖表都不相連。

1. 按一下表格、圖表、或其他項物件（需放在標題或文字之前）。
2. 在功能表上，選擇：

插入 > 新標題

或

插入 > 新文字

3. 連按兩下新物件。
4. 輸入文字。

### 若要新增文字檔案

1. 在「檢視器」的大綱窗格或目錄框中，按一下表格、圖表或其他需要放在文字之前的物件。
2. 在功能表上，選擇：

插入 > 文字檔...

3. 選取文字檔。

若要編輯文字，請在文字上連按兩下。

### 將物件貼入檢視器

其他應用程式中的物件，也可以貼入「檢視器」中。您可以使用「貼在之後」或是「選擇性貼上」其中任何一個選項。兩種貼上類型，都會將新物件放在「檢視器」中，目前選用物件的後面。當您想選擇物件的貼入格式時，請使用「選擇性貼上」。

### 尋找並取代檢視器中的資訊

1. 若要在「檢視器」中尋找或取代資訊，請從功能表選擇：

編輯 > 尋找

或

編輯 > 取代

您可以使用「尋找和取代」執行下列作業：

- 搜尋整個文件或只搜尋選定項目。
- 從目前的位置往下或往上搜尋。
- 搜尋目錄框和大綱窗格，或限制搜尋目錄框或大綱窗格。
- 搜尋隱藏項目。這些包括目錄框中的任何隱藏項目（例如，預設隱藏的「附註」表格），以及樞紐表中的隱藏列和行。
- 限制搜尋條件要區分觀察值。
- 限制樞紐表中的搜尋條件要符合整個儲存格內容。

- 在樞紐表中將搜尋條件限制為僅限註腳標記。若「檢視器」中的選項包括非樞紐表的其他項目，則此選項無法使用。

## 隱藏項目和樞紐表階層

- 多向度樞紐表中目前可見階層下的階層並不被視為隱藏階層，並且即使搜尋中並未包含隱藏項目，這些項目也將納入搜尋區中。
- 隱藏項目包括目錄框中的隱藏項目（大綱窗格中書本圖示是闔起來的項目，或是包含在大綱窗格收合區塊中的項目）以及樞紐表中隱藏的列與行，不論是預設隱藏（例如，預設隱藏的空白列與行），或是以編輯表格並選擇性隱藏特定列或行的方式手動隱藏的列與行。如果您有明確選取「包含隱藏項目」，則搜尋中只會包含隱藏選項。
- 在這兩種情況中，當系統找到含有搜尋文字或數值的隱藏或不可見元素時，會將它們顯示出來，但該項目之後便會回到原始狀態。

## 在樞紐表中尋找值範圍

若要在樞紐表中尋找落於特定數值範圍內的值：

1. 在「檢視器」中啟動樞紐表，或選取一或多個樞紐表。確認僅選取樞紐表。若選取了任何其他個體，則「範圍」選項無法使用。
2. 在功能表上，選擇：

編輯 > 尋找

3. 按一下「範圍」標籤。
4. 選取範圍類型：介於、大於或等於，或小於或等於。
5. 選取定義範圍的一或多個值。
  - 若值之一包含非數值字元，會將兩個值都視為字串。
  - 若兩個值均為數字，僅會搜尋數值。
  - 您無法使用「範圍」標籤以取代值。

此功能不適用於舊版表格。如需相關資訊，請參閱主題第 85 頁的『舊版表格』。

---

## 複製輸出到其他應用程式中

輸出物件可以複製並貼入其他應用程式中（如文書處理程式或試算表等）。您可以用各種格式來貼上輸出。依目標應用程式與所選輸出個體的不同，您可以使用下列部分或所有格式：

**meta 檔。** WMF 和 EMF meta 檔格式。這些格式僅適用於 Windows 作業系統。

**RTF（純文字格式）。** 多重所選物件、文字輸出和樞紐表，均可複製並貼上為 RTF 格式。對於樞紐表，在大部分的應用程式中，這表示表格會貼入為表格，並可於其他應用程式中編輯。若樞紐表寬度大於文件寬度，會換行、配合文件寬度縮減尺度或保持不變，視樞紐表選項設定而定。請參閱第 86 頁的『樞紐表選項』主題，以取得更多資訊。

註：Microsoft Word 可能無法正確地顯示極寬表格。

**影像。** JPG 和 PNG 影像格式。

**BIFF。** 樞紐表和文字輸出可貼上至 BIFF 格式的試算表。樞紐表中的數字會維持數值精確度。此格式僅適用於 Windows 作業系統。

文字。樞紐表和文字輸出可複製並貼上為文字。這種處理程序對只能接收或傳送文字的應用程式 (如電子郵件) 非常有用。

**Microsoft Office 圖形物件。** 支援此格式的圖表可以複製到 Microsoft Office 應用程式，並在那些應用程式中像原生 Microsoft Office 圖表一樣進行編輯。由於 SPSS Statistics/SPSS Modeler 圖表與 Microsoft Office 圖表之間存在差異，複製版本中不會保留 SPSS Statistics/SPSS Modeler 圖表的部分功能。不支援複製多個「Microsoft Office 圖形物件」格式的選定圖表。

如果目標應用程式支援多種可用格式，它可能會提供「選擇性貼上」功能表項目讓您選取格式，或者會自動顯示可用格式清單。

## 複製並貼上多個輸出物件

將多重輸出物件貼上到其他應用程式時，適用下列限制：

- **RTF 格式。** 在大部分的應用程式中，樞紐表會貼上為可在該應用程式編輯的表格。圖表、樹狀結構和模型視圖會貼入為影像。
- **meta 檔和影像格式。** 所有所選輸出物件會貼入為其他應用程式中的單一物件。
- **BIFF 格式。** 不包括圖表、樹狀結構和模型視圖。

## 選擇性複製

複製並貼上大量輸出 (尤其是非常大的樞紐表) 時，您可以使用「編輯」 > 「選擇性複製」來限制複製到剪貼簿的格式數量，以提升操作速度。

亦可將所選格式儲存為要複製到剪貼簿的預設格式組。此設定會留存於階段作業中。

## 複製為

您可以在「輸出檢視器」中的所選物件上按一下滑鼠右鍵，然後選取編輯 > 複製為以複製到最普及的格式 (例如，全部、影像檔或 **Microsoft Office 圖形物件**)。選取編輯 > 複製會複製全部。請注意，如果所選物件的複製為變成灰色或不存在，則此複製格式不適用於該特定物件。

---

## 互動式輸出

互動式輸出物件包含多個相關的輸出物件。一個物件中的選項可以變更另一個物件中所顯示或強調顯示的內容。在表格中選取一列可能會強調顯示地圖中的一個區域或顯示不同的種類的圖表。

互動式輸出物件不支援編輯功能，例如變更文字、顏色、字型或表格框線。可從互動式物件將個別物件複製到「檢視器」。可在樞紐表編輯器中編輯從互動式輸出複製的表格。

## 從互動式輸出複製物件

選取檔案>複製到檢視器可將個別輸出物件複製到「檢視器」視窗。

- 可用的選項視互動式輸出內容而定。
- 圖表和地圖建立圖表物件。
- 表格建立可在樞紐表編輯器中編輯的樞紐表。
- **Snapshot** 建立現行視圖的影像。
- 模型建立現行互動式輸出物件的副本。

選取編輯>複製物件可將個別輸出物件複製到剪貼簿。

- 將複製的物件貼至「檢視器」等於選取**檔案>複製到檢視器**。
- 將物件貼至另一個應用程式會將物件作為影像貼上。

## 縮放與平移

對於地圖，您可以使用**檢視>縮放**來縮放地圖的視圖。在已縮放的地圖視圖中，您可以使用**檢視>平移**來移動視圖。

## 列印設定值

選取**檔案>列印設定值**可控制如何列印互動式物件。

- **僅列印可見視圖**。僅列印目前顯示的視圖。此選項是預設值。
- **列印所有視圖**。列印互動式輸出中包含的所有視圖。
- 選取的選項還可以判定用於匯出輸出物件的預設動作。

---

## 匯出輸出

「匯出輸出」會將「檢視器」輸出存成 HTML、文字、Word/RTF、Excel、PowerPoint (需要 PowerPoint 97 或更新的版本) 和 PDF 格式。您也可以用多種不同的圖形格式匯出圖表。

註：「匯出至 PowerPoint」僅適用於 Windows 作業系統。

## 匯出輸出

1. 讓「檢視器」變成作用中視窗（在視窗的任何地方上按一下）。
2. 按一下工具列上的**匯出**按鈕，或在輸出視窗中用滑鼠右鍵按一下並選取**匯出**。
3. 輸入檔名（或圖表的字首）並選取匯出格式。

**要匯出的物件**。您可以匯出「檢視器」中的所有物件及看得到的物件，或只匯出選定物件。

**文件類型**。您可以使用的選項如下：

- **Word/RTF**。樞紐表匯出為 Word 表格，並完整保留所有格式屬性，例如資料格邊框、字型樣式及背景顏色。文字輸出則是匯出為 RTF 格式。PNG 格式中包括圖表、樹狀結構圖和模型視圖。請注意，Microsoft Word 可能無法正確地顯示極寬表格。
- **Excel**。樞紐表列、直欄及資料格將匯出成 Excel 列、直欄及資料格，並完整保留所有格式屬性，例如資料格邊框、字型樣式及背景顏色。文字輸出則是保留所有字型屬性匯出。文字輸出中的每一行在 Excel 檔中都是一列，且該行的全部內容都在單一資料格內。PNG 格式中包括圖表、樹狀結構圖和模型視圖。輸出可匯出為 *Excel 97-2004* 或 *Excel 2007* 和更新版本。
- **HTML**。樞紐表將匯出成 HTML 表格。文字輸出是以預先格式化的 HTML 格式匯出。圖表、樹狀結構圖和模型視圖以選取的圖形格式內嵌於文件中。檢視以 HTML 格式匯出的輸出需要與 HTML 5 相容的瀏覽器。
- **Web 報告**。Web 報告是一種互動式文件，它與大部分瀏覽器相容。許多可以從「檢視器」中取得的樞紐表互動式特性也可以從 Web 報告中取得。您也可以將 Web 報告匯出為 IBM Cognos® Active Report。
- **可攜式文件格式**。所有輸出的匯出會與「列印預覽」中所顯示的相同，也會保留所有的格式屬性。
- **PowerPoint 檔案**。樞紐表匯出成 Word 表格，並分別嵌入 PowerPoint 檔案的各投影片之中，每張投影片嵌入一個樞紐表。樞紐表的所有格式屬性都保留不變，例如，資料格邊框、字型樣式和背景顏色。圖表、樹狀結構圖和模型視圖以 TIFF 格式匯出。不包括文字輸出。

「匯出至 PowerPoint」僅適用於 Windows 作業系統。

- **文字。** 文字輸出格式包括純文字、UTF-8 及 UTF-16。樞紐表的匯出格式有兩種：以分隔符號分隔，或以空白分隔。所有的文字匯出，其輸出格式都是用空白鍵分隔。針對圖表、樹狀結構圖表以及模型視圖，會將表示影像檔名的一行插入每個圖形的文字檔中。
- **無（僅限於圖形）。** 可用的匯出格式包括：EPS、JPEG、TIFF、PNG 及 BMP。在 Windows 作業系統上，也可使用 EMF（加強型 meta 檔）格式。

開啟包含的資料夾。開啟包含匯出建立之檔案的資料夾。

## HTML 選項

HTML 匯出需要與 HTML 5 相容的瀏覽器。您可以用 HTML 格式匯出輸出的選項如下：

**樞紐表中的層。** 依預設，樞紐表階層的包含或排除是由各樞紐表的表格內容來控制。您可以改寫此設定值，並包含或排除目前可見階層以外的所有階層。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

**以互動式匯出分層表格。** 分層表格以其在「檢視器」中出現的方式顯示，且您可以使用互動方式變更瀏覽器中顯示的層。如果未選取此選項，則每一個表格層都顯示為個別表格。

**表格匯出為 HTML。** 此控制匯出樞紐表的樣式資訊。

- **以樣式及固定欄寬匯出。** 所有樞紐表樣式資訊（字型樣式、背景顏色等）和欄寬皆保留。
- **無樣式匯出。** 樞紐表轉換為預設的 HTML 表格。不保留任何樣式屬性。系統自動決定直欄寬度。

**包含註腳和標題。** 控制所有樞軸表註腳及標題的包含或排除。

**模型視圖。** 依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

附註：對於 HTML，您也可以控制匯出圖表的影像檔案格式。請參閱第 70 頁的『圖形格式選項』主題，以取得更多資訊。

若要設定 HTML 匯出選項

1. 選取「HTML」作為匯出格式。
2. 按一下「變更選項」。

## Web 報告選項

Web 報告是一種互動式文件，它與大部分瀏覽器相容。許多可以從「檢視器」中取得的樞紐表互動式特性也可以從 Web 報告中取得。

**報告標題。** 報告標頭中顯示的標題。依預設，使用檔名。您可以指定要使用的自訂標題，而非檔名。

**格式。** 有兩個報告格式選項：

- **SPSS Web 報告 (HTML 5)。** 此格式需要與 HTML 5 相容的瀏覽器。
- **Cognos Active Report (mht)。** 此格式需要支援 MHT 格式檔案或 Cognos Active Report 應用程式的瀏覽器。

**排除物件。** 您可以從報告中排除所選取的物件類型：

- **文字。** 文字物件不是日誌。此選項包括包含作用中資料集之相關資訊的文字物件。
- **日誌。** 包含已執行之指令語法清單的文字物件。日誌項目還包括由不產生任何「檢視器」輸出的指令發現的警告與錯誤訊息。



- **注意事項表格。**統計及製表程序中的輸出包括「注意事項」表格。此表格包含所使用資料集、遺漏值、用來執行程序之指令語法的相關資訊。
- **警告與錯誤訊息。**來自統計及圖表程序的警告與錯誤訊息。

重新製做表格與圖表的樣式以符合 **Web 報告**。此選項會將標準 Web 報告樣式套用到所有表格與圖表。此設定會置換輸出中的任何字型、顏色或其他樣式，如「檢視器」中所顯示。您無法修改標準 Web 報告樣式。

**Web 伺服器連線。**您可以包括一或多個正在執行「IBM SPSS Statistics Web 報告應用程式伺服器」的應用程式伺服器的 URL 位置。Web 應用程式伺服器提供功能以放置表格、編輯圖表以及儲存已修改的 Web 報告。

- 針對您要併入 Web 報告的每個應用程式伺服器選取**使用**。
- 如果 Web 報告包含 URL 規格，則 Web 報告會連接至該應用程式伺服器以提供其他編輯功能。
- 如果您指定多個 URL，則 Web 報告會嘗試以指定 URL 的順序來連接至每個伺服器。

「IBM SPSS Statistics Web 報告應用程式伺服器」可從 <http://www.ibm.com/developerworks/spssdevcentral> 下載。

## Word/RTF 選項

您可以用 Word 格式匯出輸出的選項如下：

**樞紐表中的層。**依預設，樞紐表階層的包含或排除是由各樞紐表的表格內容來控制。您可以改寫此設定值，並包含或排除目前可見階層以外的所有階層。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

**寬樞紐表。**控制處理對定義的文件寬度來說太寬的表格的方式。依預設，表格會換行配合。表格分成幾個區域，列標籤會在各表格的區域重複。或者，你可以縮小寬表格或不改變寬度表格，讓它們延伸超過定義的文件寬度。

**保留岔斷點。**如果您已定義中斷點，這些設定會保留在 Word 表格中。

**包含註腳及標題。**控制所有樞紐表註解和標題的包含或排除。

**模型視圖。**依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

**匯出的頁面設定。**開啟對話框，您可以針對匯出文件定義頁面大小和頁面邊緣。用來決定換行和縮小行為的文件寬度等於頁寬減左和右邊緣。

若要設定 Word 匯出選項

1. 選取「**Word/RTF**」作為匯出格式。
2. 按一下「**變更選項**」。

## Excel 選項

您可以用 Excel 格式匯出輸出的選項如下：

**建立工作表或活頁簿，或是修改現有工作表。**依預設，新的工作簿已建立。如果已經有包含指定名稱的檔案，就會被改寫。如果您選擇選項來建立工作表，如果指定檔案裡已經有包含指定名稱的工作表，就會被改寫。如果您選擇選項來修改現有的工作表，您必須指定工作表名稱。（對於建立工作表，這是選擇性的。）工作表名稱不能超過 31 個字元，也不能包含正斜線或反斜線、方括號、問號或星號。

匯出至 Excel 97-2004 時，如果您修改現有工作表，圖表、模型視圖和樹狀結構圖不會包含於匯出輸出。

**工作表位置。** 控制對匯出輸出的工作表位置。依預設，從第一列啟動，匯出輸出會新增在有任何內容的最後一行，不會修改任何現有的內容。新增新行至現有的工作表，這是一個好的選擇。對新增新行至現有的工作表，在最後一列新增匯出輸出是個好選擇。在指定儲存格位置新增匯出輸出將會改寫任何新增匯出輸出區域現有的內容。

**樞紐表中的層。** 依預設，樞紐表階層的包含或排除是由各樞紐表的表格內容來控制。您可以改寫此設定值，並包含或排除目前可見階層以外的所有階層。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

**包含註腳及標題。** 控制所有樞紐表註解和標題的包含或排除。

**模型視圖。** 依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

若要設定 Excel 匯出選項

1. 選取「**Excel**」作為匯出格式。
2. 按一下「**變更選項**」。

## PowerPoint 選項

您可以用於 PowerPoint 的選項如下：

**樞紐表中的層。** 依預設，樞紐表階層的包含或排除是由各樞紐表的表格內容來控制。您可以改寫此設定值，並包含或排除目前可見階層以外的所有階層。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

**寬樞紐表。** 控制處理對定義的文件寬度來說太寬的表格的方式。依預設，表格會換行配合。表格分成幾個區域，列標籤會在各表格的區域重複。或者，你可以縮小寬表格或不改變寬度表格，讓它們延伸超過定義的文件寬度。

**包含註腳和標題。** 控制所有樞軸表註腳及標題的包含或排除。

**使用「檢視器」大綱項目作為投影片標題** 在匯出時所建立的每張投影片上加上標題。每一張投影片都包含從檢視器匯出的單一項目。標題是由檢視器大綱窗格中項目的概要項目形成。

**模型視圖。** 依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

**匯出的頁面設定。** 開啟對話框，您可以針對匯出文件定義頁面大小和頁面邊緣。用來決定換行和縮小行為的文件寬度等於頁寬減左和右邊緣。

若要設定 PowerPoint 匯出選項

1. 選取「**PowerPoint**」作為匯出格式。
2. 按一下「**變更選項**」。

附註：「匯出至 PowerPoint」僅適用於 Windows 作業系統。

## PDF 選項

您可以用於 PDF 的選項如下：

**嵌入式書籤。** 此選項包含 PDF 文件中對應於「檢視器」大綱項目的書籤。與「檢視器」大綱窗格類似，書籤能讓您更輕鬆地在大量輸出物件中導覽文件。

**嵌入式字型。** 嵌入式字型能確保 PDF 文件將在所有電腦中看起來相同。否則，如果用來檢視（列印）PDF 文件的電腦中沒有文件中使用的某些字型，字型替換可能不會產生最適結果。

**樞紐表中的層。** 依預設，樞紐表階層的包含或排除是由各樞紐表的表格內容來控制。您可以改寫此設定值，並包含或排除目前可見階層以外的所有階層。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

**模型視圖。** 依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

若要設定 PDF 匯出選項

1. 選取「可攜式文件格式」作為匯出格式。
2. 按一下「變更選項」。

影響 PDF 輸出的其他設定

**頁面設定/頁面屬性。** 頁面大小、方向、邊緣、頁首及頁底的内容與顯示、PDF 文件的印出圖表大小等都是由頁面設定和頁面屬性選項來控制。

**表格內容/表格式集。** 寬表格和/或長表格的尺度，以及表格階層的列印都是由每個表格的表格內容來控制。這些性質亦可儲存於「表格式集」。

**預設/現行印表機。** PDF 文件的解析度 (DPI) 是預設印表機或目前所選印表機（可使用「頁面設定」進行變更）的目前解析度設定。最高解析度為 1200 DPI。如果印表機的設定較高，PDF 文件解析度仍然是 1200 DPI。

附註：以較低解析度的印表機進行列印時，高解析度文件的品質會降低。

## 文字選項

您可以用於文字輸出的選項如下：

**樞紐表格式。** 樞紐表的匯出格式有兩種：以分隔符號分隔，或以空白分隔。對於以空白分隔的格式，您也可以控制：

- **欄寬。** 自動配適不會將任何欄內容換行，每一欄的欄寬都和最寬的標籤或該欄的值相同。自訂會設定最大欄寬，會套用至表格中所有欄，超過該寬度的值會於該欄中換至下一行。
- **列/欄邊框字元。** 控制用來建立列及欄邊框的字元。若要隱藏列及欄邊框的顯示，請為值輸入空格。

**樞紐表中的層。** 依預設，樞紐表階層的包含或排除是由各樞紐表的表格內容來控制。您可以改寫此設定值，並包含或排除目前可見階層以外的所有階層。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

**包含註腳和標題。** 控制所有樞紐表註腳及標題的包含或排除。

**模型視圖。** 依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

若要設定文字匯出選項

1. 選取「文字」作為匯出格式。
2. 按一下「變更選項」。

## 只有圖形選項

您可以用於只有圖形輸出的選項如下：

**模型視圖。** 依預設，模型視圖的包含或排除是由各模型的模型特質來控制。您可以改寫此設定值，並包含或排除目前可見視圖以外的所有視圖。請參閱模型內容主題，以取得更多資訊。（附註：所有模型視圖，包括表格，都匯出為圖形。）

## 圖形格式選項

如果只是 HTML、文字文件和輸出圖表，您可以選取圖形格式，並控制各個圖形格式的不同選擇性設定。

若要選取圖形格式和輸出圖表的選項：

1. 選取「HTML」、「文字」或「無 (只有圖形)」作為文件類型。
2. 再從下拉清單中，選取圖形檔案格式。
3. 按一下「變更選項」變更選定圖形檔案格式的選項。

## JPEG 圖表匯出選項

- 影像大小。 原始圖表大小的百分比，最高可達 200%。
- 轉換成灰階。 將顏色轉換成灰色的陰影。

## BMP 圖表匯出選項

- 影像大小。 原始圖表大小的百分比，最高可達 200%。
- 壓縮影像以減少檔案大小。 這是一種無損失壓縮技術，可建立較小的檔案，但不會影響影像品質。

## PNG 圖表匯出選項

**影像大小。** 原始圖表大小的百分比，最高可達 200%。

**顏色深度。** 可決定匯出圖表中顏色的數目。不論是以何種深度儲存的圖表，都會有實際使用的最低顏色數目，以及該深度所允許的最高顏色數目。例如，如果圖表包含三種顏色 -- 紅色、白色和黑色，而您將它儲存為 16 色，則該圖表仍將維持三種顏色。

- 如果圖表中的顏色數目超過該深度的顏色數目，則會遞色顏色以複製該圖表中的顏色。
- 目前的畫面深度是您電腦的螢幕目前所顯示顏色數目。

## EMF 和 TIFF 圖表匯出選項

**影像大小。** 原始圖表大小的百分比，最高可達 200%。

附註：：EMF（加強型 meta 檔）格式僅適用於 Windows 作業系統。

## EPS 圖表匯出選項

**影像大小。** 您可以用百分比的方式來指定原始影像的大小 (最高可達 200%)，或是您可以用像素來指定影像寬度 (高度由寬度值與圖面比決定)。匯出的影像一定與原始影像成比例。

**包含 TIFF 預覽影像。** 儲存 TIFF 格式的 EPS 影像預覽，以在無法顯示 EPS 影像的應用程式中顯示。

**字型。** 控制 EPS 影像的字型處理。

- 使用字型參照。 如果圖表中所用的字型無法在輸出裝置上使用，就使用此字型。否則輸出裝置就會使用替代字型。

- **以曲線取代字型。** 將字型轉換成 PostScript 曲線資料。文字本身在可編輯 EPS 圖形的應用程式中，將無法再編輯成文字。如果輸出裝置無法使用圖表中所用的字型，此選項就很有用。

---

## 檢視器列印

若要列印「檢視器」視窗的內容，有兩種選擇：

**所有顯示輸出。** 只列印目錄框中目前所顯示的項目。隱藏的項目（大綱窗格中，項目的書本圖像如果是闔起來的，或者項目隱藏於合併概要層裡面）將不會印出來。

**選擇。** 只列印目前在大綱窗格及/或目錄框中選擇的項目。

## 若要列印輸出與圖表

1. 讓「檢視器」變成作用中視窗（在視窗的任何地方上按一下）。
2. 在功能表上，選擇：

檔案 > 列印...

3. 選取您想要的列印設定。
4. 按一下「確定」就可以啟動列印。

## 預覽列印

「預覽列印」會告訴您「檢視器」文件的每一頁上將列印些什麼。通常在實際列印「檢視器」文件之前，最好先檢查「預覽列印」，因為「預覽列印」會讓您看到在「檢視器」的目錄框中可能看不到的項目，其中包括：

- 分頁符號
- 樞紐表的隱藏階層
- 寬表格的分頁
- 列印在每一頁上的頁首和頁底

如果「檢視器」上，目前已經選取任何項目的話，預覽會只顯示所選用的輸出。若要預覽所有輸出，請確定在檢視器中未選取任何項目。

## 頁面屬性：頁首與頁底

頁首和頁底是印在每一頁頂端和底端的資訊。您可以輸入任何文字，當作頁首和頁底。您也可以使用對話框中央的工具列，來插入下列各個項目：

- 日期和時間
- 頁碼
- 檢視器檔名
- 標題概要標記
- 頁面標題和次標題。
- 「設成預設值」使用這裡指定的設定值作為新「檢視器」文件的預設設定值。（附註：這樣會使現行「頁首/頁底」標籤和「選項」標籤的設定成為預設設定。）
- 概要標題標記可指出各頁面上第一個項目的第一、第二、第三、和/或第四層的概要標題。
- 頁標題和次標題會印出目前頁面的標題和次標題。您可以用「檢視器」中「插入」功能表上的「插入新頁面標題」，或是 TITLE 和 SUBTITLE 指令來建立這些標題。如果您未指定任何頁標題或次標題的話，這個設定將被忽略。

附註：您可以在「選項」對話框（選擇「編輯」功能表上的「選項」）的「檢視器」標籤上，控制新頁面標題和副標題的字型特性。在「檢視器」中編輯標題，也可以變更現有頁面標題和次標題的字型特性。

若要查看頁首和頁底在列印頁面上的外觀，請在「檔案」功能表上選擇「預覽列印」。

## 插入頁面標頭及標底

1. 讓「檢視器」變成作用中視窗（在視窗的任何地方上按一下）。
2. 在功能表上，選擇：

檔案 > 標頭和標底...

3. 輸入您想出現在每一頁上的頁首及/或頁底。

## 頁面屬性：選項

這個對話框控制印出的圖表大小、所列印輸出項目之間的空間，以及頁碼。

- **印出圖表大小。**可控制印出圖表與定義的頁面大小之間相對的大小。圖表的圖面比（寬對高的比例）不會受所印出圖表大小的影響。圖表整個印出的大小會同時受到它的長度和寬度的限制。一旦圖表的外框到達頁面的左側或右側邊框時，圖表大小就無法再增加，也就無法填滿更多頁面高度。
- **項目之間的空間。**控制列印項目之間的空間。每個樞紐表、圖表和文字物件都是個別的項目。這個設定不會影響檢視器中各個項目的顯示。
- **頁碼起始。**以指定的數字啟動循序編頁碼。
- **設為預設值。**此選項使用這裡指定的設定值作為新「檢視器」文件的預設值。（請注意，此選項會將現行標頭/標底設定與選項設定設為預設值。）

## 若要變更列印圖表大小、編頁碼、及列印項目間的空間

1. 讓「檢視器」變成作用中視窗（在視窗的任何地方上按一下）。
2. 在功能表上，選擇：

檔案 > 頁面屬性...

3. 按一下「選項」標籤。
4. 變更設定，再按一下「確定」。

---

## 儲存輸出

可以儲存「檢視器」的內容。

- **輸出物件 (\*.cou)**。這種格式可儲存整個輸出儲存器，包括圖形、標籤、註釋等。這種格式可在 IBM SPSS Modeler 中開啟與檢視、新增至專案，並使用 IBM SPSS Collaboration and Deployment Services 儲存庫進行發佈與追蹤。這種格式與 IBM SPSS Statistics 不相容。
- **檢視器檔案 (\*.spv)**。用來在「檢視器」視窗中顯示檔案的格式。當您將 IBM SPSS Modeler 中的模型塊以這種格式儲存時，只會儲存模型標籤中的「檢視器」內容。

若要控制儲存 Web 報告的選項，或者將結果儲存為其他格式（例如，文字、Word、Excel），請使用檔案功能表上的匯出。

## 若要儲存檢視器文件

1. 從「檢視器」視窗功能表選擇：

檔案 > 儲存

2. 輸入文件名稱，然後按一下「儲存」。

您也可以選擇執行下列動作：

#### 鎖定檔案以避免在 IBM SPSS Smartreader 中進行編輯

若鎖定「檢視器」文件，您可以處理樞紐表（對調列與欄、變更顯示的階層等），但不能編輯任何輸出或是儲存任何變更到 IBM SPSS Smartreader（與「檢視器」文件搭配使用的獨立產品）中的「檢視器」文件。這項設定不影響在 IBM SPSS Statistics 或 IBM SPSS Modeler 中開啟的「檢視器」文件。

#### 使用密碼來加密檔案

使用密碼加密檔案可保護儲存在「檢視器」文件中的機密資訊。加密之後，只有提供了密碼才能開啟文件。IBM SPSS Smartreader 使用者也需要提供密碼才能開啟檔案。

若要加密「檢視器」文件：

- a. 在「另存輸出」對話框中選取「使用密碼加密檔案」。
- b. 按一下「儲存」。
- c. 在「加密檔案」對話框中提供密碼，並在「確認」密碼文字方框中重新輸入一次。密碼長度上限為 10 個字元，區分字母觀察值。

警告：密碼一旦遺失將無法復原。密碼一旦遺失，檔案將無法開啟。

#### 建立強式密碼

- 使用 8 個以上的字元。
- 在您的密碼中包含數字、符號甚至是標點符號。
- 避免使用連續的數字或字元，例如 "123" 及 "abc"，也避免重複的情況，例如 "111aaa"。
- 不要建立使用個人資訊的密碼，例如生日或是暱稱。
- 定期變更密碼。

附註：不支援將加密檔案儲存至 IBM SPSS Collaboration and Deployment Services 儲存庫。

#### 修改已加密的檔案

- 如果您開啟已加密的檔案，對其進行修改並選擇「檔案 > 儲存」，系統將以相同的密碼儲存已修改的檔案。
- 若要變更已加密檔案的密碼，請開啟檔案，重複上述的加密檔案步驟後，在「加密檔案」對話框中指定不同的密碼。
- 開啟已加密的檔案，選擇「檔案 > 另存新檔」，並在「另存輸出」對話框中取消選取「使用密碼加密檔案」，即可儲存已加密檔案的未加密版本。

附註：加密資料檔案及輸出文件無法在第 21 版之前的 IBM SPSS Statistics 版本中開啟。加密語法檔案無法在第 22 版之前的版本中開啟。

#### 將所需模型資訊與輸出文件儲存在一起

只有在輸出文件中的模型檢視器項目需要輔助資訊才能啟用部分互動式特性時，此選項才適用。按一下**進一步資訊**以顯示這些模型檢視器項目及需要輔助資訊之互動式特性的清單。將此資訊與輸出文件儲存在一起可能會大幅增加文件大小。如果您選擇不儲存此資訊，則您仍然可以開啟這些輸出項目，但指定的互動式特性將會不可用。

---

## 樞紐表

### 樞紐表

許多結果都可以表示成互動式的樞紐表。也就是說，您可以重新排列列、行、和層。

### 使用樞紐表

您可以透過下列選項來處理樞紐表，包括：

- 轉置列和行
- 移動列和行
- 建立多維度層
- 分組及解散列和行
- 顯示和隱藏列、行和其他資訊
- 旋轉列和欄標籤
- 尋找用語的定義

### 啟動樞紐表

您必須先**啟動**表格，才能處理或修改樞紐表。若要啟動表格：

1. 連按兩下表格。  
  
或
2. 在表格上用滑鼠右鍵按一下，然後從蹦現功能表中選擇**編輯內容**。
3. 從子功能表中選擇「**在檢視器中**」或「**在不同的視窗中**」。

### 樞紐分析表格

表格有三個維度：列、行及層。一個維度可以包含多個元素 (或沒有任何元素)。您可以在不同維度之間移動元素，藉以變更表格的組織。若要移動元素，只需拖放元素到您想要的位置即可。

### 在維度內變更元素的顯示順序

若要變更表格維度 (列、行或階層) 內元素的顯示順序：

1. 若尚未開啟樞紐分析匣，請從「樞紐表」功能表選擇：  
  
    **樞軸 > 樞紐分析匣**
2. 在樞紐分析匣中拖放維度內的元素。

### 在維度元素內移動列和行

1. 在表格本身 (非樞紐分析匣) 中，按一下您要移動其列或行的表格。
2. 拖曳標記到新的位置。

### 轉置列和行

如果您只是想要翻轉列與行，可用另一種簡單的方法來使用樞紐分析匣：

1. 在功能表上，選擇：  
  
    **樞紐 > 轉置列和行**

這樣做的效果，相當於將所有列元素拖曳到直欄維度上，並將所有直欄元素拖曳到列維度上。



## 分組列或行

1. 選取列或欄標籤，將它們分在同一組 (按一下並拖曳滑鼠，或者按住 Shift 鍵，以選取數個標記)。
2. 在功能表上，選擇：

編輯 > 群組

會自動插入群組標記。如果連按兩下群組標記，就可以編輯標記文字。

附註：若要將列或欄新增至現有群組，必須先取消群組群組中目前的項目。然後建立包含其他項目的新組別。

## 解散列或行的群組

解散群組會自動刪除群組標記。

## 旋轉列或欄標籤

您可以在表格最內層欄標籤與最外層列標籤的水平和垂直顯示之間旋轉標記。

1. 在功能表上，選擇：

格式 > 旋轉內層欄標籤

或

格式 > 旋轉外層列標籤

您只能旋轉最內層的欄標籤，和最外層的列標籤。

## 排序列

若要排序樞紐表列：

1. 啟動表格。
2. 選取該直欄中要用來排序的任何儲存格。若只要排序一排選取的列群組，請選取該直欄中二或多個要用來排序的連續儲存格。
3. 在功能表上，選擇：

編輯 > 排序列

4. 從子功能表選取「遞增」或「遞減」。
  - 如果列維度包含群組，則排序只會影響包含該選項的群組。
  - 您無法排序群組範圍以外的項目。
  - 您無法針對列維度中具有多個項目的表格進行排序。

## 插入列和欄

若要在樞紐表中插入列和欄：

1. 啟動表格。
2. 選取表格中的任何儲存格。
3. 在功能表上，選擇：

前置插入

或

## 後置插入

從子功能表中，選擇：

列

或

欄

- 在新列或新行的每個儲存格中插入加號 (+) 可避免新列或新欄因空白而自動隱藏。
- 在含有巢狀或階層維度的表格中，會依對應的維度層級插入欄或列。

## 控制變數和數值標籤的顯示

如果變數包含描述性變數或數值標籤，就可以控制樞紐表中的變數名稱、變數標籤、數值和數值標籤顯示。

1. 啟動樞紐表。
2. 在功能表上，選擇：

檢視 > 變數標籤

或

檢視 > 值標籤

3. 從子功能表選取下列其中一個選項：
  - 名稱或值。僅顯示變數名稱（或數值）。不顯示描述性標記。
  - 標籤。僅顯示描述性標記。不顯示變數名稱（或數值）。
  - 兩者。顯示名稱（或數值）和描述性標記。

## 變更輸出語言

若要變更樞紐表的輸出語言：

1. 啟動表格
2. 在功能表上，選擇：

檢視 > 語言

3. 選取其中一個可用語言。

變更語言只會影響應用程式產生的文字，如：表格標題、列和直欄標籤以及註腳文字。變數名稱、描述性變數和數值標籤不會受到影響。

## 瀏覽大型表格

若要使用瀏覽視窗來瀏覽大型表格：

1. 啟動表格。
2. 在功能表上，選擇：

檢視 > 導覽

## 復原變更

您可以復原對於啟動的樞紐表之最近變更或所有變更。這兩個動作都只會套用至自從最近一次啟動表格以來所做的變更。

若要復原最近變更：

1. 在功能表上，選擇：

編輯 > 復原

若要復原所有變更：

2. 在功能表上，選擇：

編輯 > 還原

## 使用階層

您可以將每個類別 (或類別組合)，以不同的二維表格顯示出來。這些表格可以被視為層而且層層相疊，所以您只看到最頂端的那一層。

## 建立和顯示階層

若要建立階層：

1. 如果尚未開啟樞紐分析匣，請從樞紐表功能表中選擇：

樞軸 > 樞紐分析匣

2. 將元素從列維度或直欄維度，拖曳到層級維度。

將元素移動至層級維度會建立多維表格，但只會顯示一個二維的「圖塊」。您可以看到的表格是最頂層的那個表格。例如，如果 yes/no 類別變數是在層級維度中，則多維表格便會有兩個層級：一個代表 yes 種類，另一個代表 no 種類。

變更顯示的階層

1. 從階層的下拉清單中選擇類別 (在樞紐表本身，而不是樞紐分析匣)。

## 到圖層類別

您可以利用「到層類別」，在樞紐表中變更層。當有許多階層、或選定階層有許多類別時，此對話框特別有用。

## 顯示和隱藏項目

您可以隱藏下列儲存格類型，包括：

- 維度標記
- 類別 (包括列或行中，其標記儲存格和資料格)
- 類別標記 (不隱藏資料格)
- 註腳、標題和標題

## 隱藏表格中的列和行

### 顯示表格中隱藏的列和行

1. 在功能表上，選擇：

檢視 > 顯示所有種類

如此會將表格中所有隱藏列和行都顯示出來。(如果這個表格的「表格內容」中，已經選取「隱藏空白的列與行」的話，則全部空白的列或行仍然會被藏起來)。

## 隱藏和顯示維度標記

1. 選取維度標記，或維度裡面任何一個類別標記。
2. 從「檢視」功能表或蹦現功能表中選擇隱藏維度標籤或顯示維度標籤。

## 隱藏和顯示表格標題

若要隱藏標題：

1. 啟動樞紐表。
2. 選取標題。
3. 從「檢視」功能表中，選擇「隱藏」。

若要顯示隱藏的標題：

4. 從「檢視」功能表中，選擇「全部顯示」。

## 表格外觀

「表格式集」是一組性質，用來定義表格外觀。您可以選取先前定義的「表格式集」，也可以自行設計「表格式集」。

- 套用「表格式集」之前或之後，您都可以透過儲存格性質，來變更某個儲存格或某組儲存格的格式。即使您套用新的「表格式集」，編輯過的資料格格式仍保持不變。
- 除此之外，您可以將所有的儲存格，重設成目前「表格式集」所定義的儲存格格式。此選項會重設所有被編輯過的資料格。如果「表格式集檔案」清單中，已經選取如顯示的話，任何被編輯過的儲存格，都會根據目前表格的性質，加以重設。
- 只有在「表格內容」對話框中定義的表格內容才會儲存在「表格式集」中。「表格式集」並不包括個別儲存格修改。

## 若要套用表格式集

1. 啟動樞紐表。
2. 在功能表上，選擇：

格式 > 表格式集...

3. 從檔案清單選取「表格式集」。若要從別的目錄選取檔案，請按一下瀏覽。
4. 按一下**確定**，將「表格式集」套用到選定的樞紐表上。

## 若要編輯或建立表格式集

1. 在「表格式集」對話框中，從檔案的清單中選取「表格式集」。
  2. 按一下**編輯外觀**。
  3. 調整表格內容，將它們變成理想的屬性之後，然後按一下**確定**。
  4. 按一下**將目前表格設定存檔**以儲存剛才編輯的「表格式集」，或按一下**另存新檔**，將它存成新的「表格式集」。
- 您對「表格式集」所作的修改，只會影響到選定的樞紐表。編輯過的「表格式集」無法套用到任何使用該「表格式集」的其他表格上，除非您再選取那些表格，然後重新套用「表格式集」。
  - 只有在「表格內容」對話框中定義的表格內容才會儲存在「表格式集」中。「表格式集」並不包括個別儲存格修改。

## 表格內容

您可以利用「表格內容」來設定表格的一般性質、和為表格的不同部分設定儲存格樣式。您可以：

- 控制一般性質，如隱藏空白列 (或行)，或者調整列印性質。
- 控制註腳標記的格式和位置。
- 判斷資料區域中的儲存格、列、欄標籤，或者表格其他區域的指定格式。
- 控制線寬和顏色，這些線條組成表格每個區域的邊框。

## 若要變更樞紐表內容

1. 在功能表上，選擇：

格式 > 表格內容...

2. 選取標籤 (一般、註腳、儲存格格式、框線 或 列印)。
3. 選擇您要的選項。
4. 按一下**確定**或**套用**。

新性質會套用到選定的樞紐表上。

## 表格內容：一般

下列幾項性質會套用到整個表格上。您可以：

- 顯示或隱藏空白列或行。(空白的列或行，代表其任何資料格中，都不含資料。)
- 控制列標籤的放置位置，可以在左上角或嵌入巢狀結構中。
- 控制最大和最小行寬 (以點表示)。

若要變更一般的表格內容：

1. 按一下**一般**標籤。
2. 選擇您要的選項。
3. 按一下**確定**或**套用**。

設定要顯示的列：

註：此功能僅適用於舊版表格。

依預設，含有許多列的表格會顯示在列數為 100 的區域。若要控制表格中顯示的列數：

1. 選取「依列顯示表格」。
2. 按一下「設定要顯示的列」。

或

3. 從啟動的樞紐表的「檢視」功能表選擇「依列顯示表格」和「設定要顯示的列」。

**要顯示的列。** 控制一次顯示的列數上限。瀏覽控制項可讓您移至表格的不同區域。下限值為 10。預設值是 100。

**最少行列數容差。** 控制表格最深處之列維度的列數下限，以分割顯示的表格檢視。例如，若每個最深處列維度群組中有六個種類，指定值 6 可避免在顯示的檢視中分割任何群組。此設定可能會造成顯示的檢視中的總列數超過指定顯示的列數上限。

## 表格內容：注意事項

「表格內容」對話框的「注意事項」標籤會控制註腳格式化及表格註解文字。

**註腳。**註腳標記的性質包括：文字樣式和文字位置。

- 註腳標記的樣式可以是數字 (1、2、3...) 或字母 (a、b、c...)。
- 註腳標記可以加到文字上，變成上標或下標。

**註解文字。**您可以將註解文字新增至每一個表格。

- 當您移至「檢視器」中的表格上方時，會在工具提示中顯示註解文字。
- 當表格具有焦點時，螢幕閱讀器會讀取註解文字。
- 「檢視器」中的工具提示只會顯示註解的前 200 個字元，但是螢幕閱讀器會讀取整個文字。
- 當您將輸出匯出至 HTML 或 Web 報告時，註解文字會用作替代文字。

## 表格內容：儲存格格式

就格式而言，表格可分成以下區域：標題、階層、角落標籤、列標籤、欄標籤、資料、標題及註腳。而對表格的每個區域來說，您可以修改其相關的儲存格格式。這些儲存格格式包括文字特性（例如字型、大小、顏色、樣式）、水平和垂直對齊、背景顏色，以及內層儲存格邊緣。

儲存格格式會套用到區域圖上（資訊的類別）。這不是個別儲存格的特性。當您執行表格的樞紐分析時，這項區別是很重要的。

例如，

- 如果您指定欄標籤的儲存格格式為粗體字型，則無論目前直欄維度所顯示的是何種資訊，欄標籤都將顯示粗體。如果您將直欄維度中的項目移到另一個維度，該項目並不會保留欄標籤的粗體特性。
- 如果您只利用下列方式將欄標籤改成粗體：在啟動的樞紐表中，將儲存格反白，然後再按一下工具列上的「粗體」按鈕，那麼無論您將那些儲存格移到那個維度，它們都會保持粗體，至於其他移到直欄維度中的項目，直欄標籤就不會繼續保持粗體了。

若要變更儲存格的格式：

1. 選取儲存格的格式標籤。
2. 從下拉清單中選取一個區域，或在樣本的任一個區域上按一下。
3. 選取區域圖的特性。您的選擇會反映在範例上。
4. 按一下**確定**或**套用**。

替代的列色彩

若要套用不同的背景和/或文字色彩到表格「資料」區域的替代列：

1. 選取「區域」下拉清單中的「資料」。
2. 選取(勾選)「背景色彩」群組中的「替代的列色彩」。
3. 選取替代的列背景和文字要使用的色彩。

替代的列色彩只會影響表格的「資料」區域。不會影響列或欄標籤區域。

## 表格內容：邊框

您可以替表格中的每個邊框位置，選取線條樣式和顏色。如果您選取無作為樣式，則選用部分將沒有線條。

若要變更表格邊框：

1. 按一下**邊框**標籤。
2. 如果想選取邊框位置，請從清單中，在它的名稱上按一下，或者在「樣本」區域中，按一下線條。
3. 選取線條樣式或選取無。
4. 請選取顏色。
5. 按一下**確定**或**套用**。

## 表格內容：列印

您可以控制的樞紐表列印性質如下：

- 列印所有階層、或只列表格的最頂層，並且將每一層印在不同頁面上。
- 將表格水平或垂直摺起來，以配合列印頁面的大小。
- 當表格超過所定義的頁面大小（不論太寬或太長）時，透過控制印出表格之任何一部分中所包含的最小行列數，來控制最小行列數。

附註：如果表格太大無法顯示在現行頁面上，但頁面上方的其他輸出，又符合所定義頁面長度的話，此表格會自動列印在新的頁面上，不論最小行數的設定為何。

- 當表格無法印在同一頁時，仍然將剩下的文字保留起來。您可以在每頁的底部、頂端，顯示待續文字。如果兩個選項都未選取的話，就不會顯示待續文字。

若要控制樞紐表列印性質：

1. 按一下「**列印**」標籤。
2. 選取您要的列印選項。
3. 按一下**確定**或**套用**。

## 儲存格內容

「儲存格」性質會套用到所選的儲存格上。您可以變更字型、數值格式、對齊方式、邊緣和顏色。儲存格性質會置換表格內容；因此，如果您變更表格內容的話，任何單獨套用的儲存格性質並不會隨之變更。

若要變更儲存格的性質：

1. 選取表格中的儲存格。
2. 從「格式」功能表或跳現功能表中，選擇**儲存格內容**。

## 字型及背景

「字型和背景」標籤可控制表格中，選定的儲存格的字型樣式、顏色和背景顏色。

## 格式值

「格式化數值」標籤可控制選定的儲存格數值格式。您可以選取數字、日期、時間、或貨幣的格式，也可以調整螢幕上所顯示之小數位數。

## 對齊方式及邊距

「對齊方式和邊緣」標籤可控制選定儲存格中數值的水平和垂直對齊，以及該儲存格的上、下、左和右邊緣。「混合」水平對齊方式可將根據各儲存格的類型來對齊內容。例如，日期向右對齊，而文字數值則向左對齊。

## 註腳和標題

您可以在表格中新增註腳和標題。也可以隱藏註腳或標題、變更註腳標記和將註腳重新編號。

## 新增註腳和標題

若要在表格中新增標題：

1. 從「插入」功能表中，選擇「標題」。

表格上的任何一個項目都可以加入註腳。若要新增註腳：

1. 在啟動的樞紐表中，按一下標題、儲存格、或者標題。
2. 從「插入」功能表中，選擇「註腳」。
3. 在提供的區域插入註腳文字。

## 若要隱藏或顯示標題

若要隱藏標題：

1. 選取標題。
2. 從「檢視」功能表中，選擇「隱藏」。

若要顯示隱藏的標題：

1. 從「檢視」功能表中，選擇「全部顯示」。

## 若要隱藏或顯示表格中的註腳

若要隱藏註腳：

1. 在包含註腳參考的儲存格上用滑鼠右鍵按一下，並從蹦現功能表中選取隱藏註腳

或

2. 在表格的註腳區域中選取註腳，並從蹦現功能表中選取隱藏。

附註：對於舊版表格，請選取表格的註腳區域，從蹦現功能表中選取編輯註腳，然後取消選取（清除）您要隱藏的任何註腳的「可見」內容。

如果儲存格中包含多個註腳，請使用後者的方法來選擇性地隱藏註腳。

若要在表格中隱藏所有註腳：

1. 選取表格註腳區域中的所有註腳（使用按一下並拖曳滑鼠，或者按住 Shift 的方式來選取註腳），並從「檢視」功能表中選取「隱藏」。

若要顯示隱藏的註腳：

1. 從「檢視」功能表中選取「顯示所有註腳」。

## 註腳標記

註腳標記會變更可用來標記註腳的字元。依預設，根據表格內容設定，標準註腳標記可能是依序排列的字母或數字。您也可以指派特殊標記。當您對註腳重新編號或針對標準標記在數字與字母之間切換時，特殊標記並不會受到影響。標準標記的數字或字母顯示，以及註腳標記的下標或上標位置，皆由「表格內容」對話框的「註腳」標籤控制。

若要變更註腳標記：

1. 選取註腳。
2. 從格式功能表中，選擇註腳標記。

特殊標記限制為 2 個字元。在表格的註腳區域中，具有特殊標記的註腳會優先於具有依序排列的字母或數字的註腳，因此變更特殊標記可能會重新排列註腳清單中的順序。



## 註腳重新編號

當您切換列、行、階層以轉換表格時，註腳順序可能會亂掉。若要將註腳重新編號：

1. 從「格式」功能表中，選擇「註腳重新編號」。

## 編輯舊版表格中的註腳

對於舊版表格，您可以使用「編輯註腳」對話框來輸入及修改註腳文字與字型設定、變更註腳標記，以及選擇性地隱藏或刪除註腳。

當您在舊版表格中插入新註腳時，「編輯註腳」對話框會自動開啟。若要使用「編輯註腳」對話框來編輯現有註腳（而不建立新註腳）：

**標記。**依預設，根據表格內容設定，標準註腳標記可能是依序排列的字母或數字。若要指派特殊標記，只需在「標記」行中輸入新標記值。當您對註腳重新編號或針對標準標記在數字與字母之間切換時，特殊標記並不會受到影響。標準標記的數字或字母顯示，以及註腳標記的下標或上標位置，皆由「表格內容」對話框的「註腳」標籤控制。請參閱第 80 頁的『表格內容：注意事項』主題，以取得更多資訊。

若要將特殊標記變更為標準標記，請用滑鼠右鍵按一下「編輯註腳」對話框中的標記，並從蹦現功能表中選取註腳標記，然後在「註腳標記」對話框中選取「標準標記」。

**註腳。**註腳的內容。顯示內容反映出了目前字型與背景設定。使用「字型」子對話框，可變更個別註腳的字型設定。如需相關資訊，請參閱主題『註腳字型與顏色設定』。在「儲存格內容」對話框的「字型及背景」標籤中，可將單一背景顏色套用至所有註腳，也可對其進行變更。請參閱第 81 頁的『字型及背景』主題，以取得更多資訊。

**可見。**依預設，所有註腳皆為可見。取消選取（清除）「可見」勾選框可隱藏註腳。

**註腳字型與顏色設定：**對於舊版表格，您可以使用「格式」對話框來變更一或多個所選註腳的字型系列、樣式、大小及顏色：

1. 在「編輯註腳」對話框中，選取(按一下)「註腳」網格中的一或多個註腳。
2. 按一下「格式」按鈕。

會將所選字型系列、樣式、大小與顏色套用至所有所選註腳。

背景顏色、對齊方式與邊緣可在「儲存格內容」對話框中設定，並套用至所有註腳。您無法為個別註腳變更這些設定。請參閱第 81 頁的『字型及背景』主題，以取得更多資訊。

## 資料格寬

「設定資料格寬度」是用來將所有資料儲存格設成相同的寬度。

若要設定所有資料儲存格的寬度：

1. 在功能表上，選擇：

格式 > 設定資料格寬度...

2. 輸入格寬度的值。

## 變更行寬

1. 按一下並拖曳行的邊框。

## 顯示樞紐表中隱藏的邊框

如果表格邊框多半被隱藏起來的話，您可以將它們顯示出來。這樣做，可以變更行寬之類的工作，變得比較容易。

1. 從「檢視」功能表中，選擇「網格線」。

## 選取樞紐表中的列、行與儲存格

您可以選取完整的列或行，或指定的資料與標記儲存格之集合。

若要選取多個儲存格：

選取 > 資料及標籤儲存格

## 列印樞紐表

有幾個因素會影響樞紐表的列印結果，但是您可以變更樞紐表的屬性，來控制這些因素。

- 對於多向度樞紐表（有多個階層的表格）而言，您可以列印所有階層或者只列印最頂層（看得見的那一層）。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。
- 對於長或寬樞紐表而言，您可以自行調整表格大小，以配合頁面，或者控制表格的分隔符號、分頁符號的位置。請參閱第 81 頁的『表格內容：列印』主題，以取得更多資訊。

您可以使用「檔案」功能表上的「預覽列印」，來查看樞紐表的列印結果。

## 控制寬和長表格的表格分行

如果樞紐表太寬或太長，以致無法列印在定義頁面上時，樞紐表會自動分割，以便印成很多段。您可以：

- 控制大型表格從哪裡啟動分割列或行。
- 分割表格時，要求哪些列或行必須連在一起。
- 調整表格大小，以符合定義頁面。

若要指定印出樞紐表的列和欄岔斷，請執行下列動作：

1. 啟動樞紐表。
2. 在您要插入岔斷的位置左側的直欄中按一下任意資料格，或在您要插入岔斷的列之前的列中按一下任意資料格。
3. 在功能表上，選擇：

格式 > 岔斷點 > 垂直岔斷點

或

格式 > 岔斷點 > 水平岔斷點

1. 啟動樞紐表。
2. 在您要插入岔斷的位置左側的直欄中按一下任意資料格，或在您要插入岔斷的列之前的列中按一下任意資料格。
3. 在功能表上，選擇：

格式 > 岔斷點 > 垂直岔斷點

或

格式 > 岔斷點 > 水平岔斷點

若要指定要保持在在一起的列或直欄，請執行下列動作：

1. 選取需要連在一起的列或欄標籤。按一下並拖曳或按住 Shift 並點按來選取多個列或直欄標籤。
2. 在功能表上，選擇：

**格式 > 岔斷點 > 保持在在一起**

若要檢視岔斷點與保持在在一起群組，請執行下列動作：

1. 在功能表上，選擇：

**格式 > 岔斷點 > 顯示岔斷點**

中斷點會顯示為垂直或水平行。保持在在一起群組會顯示成由黑色邊框括住的灰色矩形區。

附註：舊式表格不支援顯示岔斷點與保持在在一起群組。

若要清除中斷點與保持在在一起群組

若要清除中斷點：

1. 按一下垂直中斷點左側行中的任何儲存格，或按一下水平中斷點上方列中的任何儲存格。
2. 在功能表上，選擇：

**格式 > 岔斷點 > 清除岔斷點或群組**

若要清除保持在在一起群組：

3. 選取指定群組的行或列標籤。
4. 在功能表上，選擇：

**格式 > 岔斷點 > 清除岔斷點或群組**

當您旋轉或重新排列任何列或行時，則會自動清除所有中斷點與保持在在一起群組。此行為不適用於舊版表格。

## 從樞紐表建立圖表

1. 請連按兩下樞紐表，來啟動它。
2. 選取您要顯示在圖表中的列、行或儲存格
3. 在選定區域中的任何位置上，用滑鼠右鍵按一下。
4. 從蹦現功能表中選擇**建立圖形**，並選取圖表類型。

## 舊版表格

您可以選擇將表格作為遺存表格來呈現（稱為版本 19 中的完整表格），它們與 20 之前的版本 IBM SPSS Statistics 完全相容。遺存表格呈現可能較為緩慢，僅在需要與 20 之前的版本相容時推薦使用。如需有關如何建立舊版表格的資訊，請參閱第 86 頁的『樞紐表選項』。

---

## 選項

### 選項

這些選項可以控制各種設定。

## 若要變更選項設定

1. 在功能表上，選擇：

編輯 > 選項...

2. 按一下您想要變更設定的標籤。
3. 變更設定。
4. 按一下**確定**或**套用**。

## 一般選項

### 執行緒數目上限

計算結果時多執行緒程序使用的執行緒數目。**自動**設定基於可用的處理核心數目。如果您想要在多執行緒程序執行時，讓更多處理字元可用於其他應用程式，請指定較低的值。在分散式分析模式下，停用此選項。

### 輸出

**顯示小數值的前導零。**顯示僅包含小數部分之數值的前導零。例如，顯示前導零時，值 .123 會顯示為 0.123。此設定不適用於使用貨幣或百分比格式的數值。除了固定的 ASCII 檔案 (\*.dat) 之外，將資料儲存到外部檔案時，不會包括前導零。

**測量系統。**使用度量系統 (點、英吋或公分)，來指定樞紐表在列印時的儲存格邊緣、儲存格寬度、表格間距等多項屬性。

### 檢視器的選項

檢視器輸出顯示選項只會影響變更設定之後產生的新輸出。這些設定中的變更不會影響「檢視器」中已顯示的輸出。

**初始輸出狀態。**控制每一次執行程序時會自動顯示或隱藏的項目，以及項目的初始對齊方式。您可以控制下列項目的顯示：日誌、警告、註解、標題、樞紐表、圖表、樹狀結構圖和文字輸出。您也可以顯示或隱藏日誌中的指令。您可以複製日誌中的指令語法，再把它存入語法檔。

註：「檢視器」中顯示的所有輸出項目都是靠左對齊。版面調整設定只會影響到列印輸出時的對齊方式。置中或靠右對齊的項目都以小符號來識別。

**標題。**控制新輸出標題的字體、大小和顏色。

**頁面標題。**控制由 TITLE 與 SUBTITLE 指令語法產生或由插入功能表上的新頁面標題建立的新頁面標題與頁面標題的字型樣式、大小及顏色。

**文字輸出** 用於文字輸出的字型。文字輸出已設計為使用單一空間位置 (固定點數) 的字型。如果您選取固定比例的字型，則表列式輸出不會適當地對齊。

**預設頁面設定。**控制列印之方向與邊距的預設選項。

### 樞紐表選項

「樞紐表」選項可設定各種樞紐表的顯示選項。

## 表格式集

請從檔案清單上選取「表格式集」，再按一下「確定」或「套用」。您可以使用 IBM SPSS Statistics 所提供的其中一個「表格式集」、或在「樞紐表編輯器」中建立您自己的圖表（從「格式」功能表中選擇「表格式集」）。

- 瀏覽。讓您可以從其他目錄選取「表格式集」。
- 設定表格式集目錄。讓您可以變更預設的「表格式集」目錄。使用「瀏覽」以瀏覽至您要使用的目錄，選取該目錄中的「表格式集」，然後選取「設定表格式集目錄」。

附註：在舊版 IBM SPSS Statistics 中建立的表格式集無法在 16.0 或更新版本中使用。

## 直欄寬度

這些選項會控制樞紐表中欄寬的自動調整。

- 僅依標記調整。根據欄標記的寬度調整行寬。這會產生更精簡的表格，但資料值寬度如果超過標記的話，就會被截斷。
- 調整所有表格的標記與資料。可依欄標籤或最大資料值來調整欄寬，以較寬者為主。如此產生出來的表格會比較寬，但是所有的數值都會顯示出來。

## 預設編輯模式

此選項會控制在「檢視器」視窗或個別視窗中啟動樞紐表。根據預設，連按兩下樞紐表時便可以在「檢視器」視窗中啟動所有表格，但大型表格除外。您可以選擇在另一個視窗中啟動樞紐表，也可以選取大小設定，使較小的樞紐表開啟於「檢視器」視窗，而較大的樞紐表則開啟於另一個視窗。

## 以 Rich Text 格式將寬表格複製到剪貼簿

當樞紐表以 Word/RTF 格式貼上時，文件太寬的表格會換行、配合文件寬度縮減尺度或保持不變。

## 輸出選項

輸出選項控制多種輸出選項的預設設定。

**螢幕閱讀器協助工具。**控制螢幕閱讀器如何讀取樞紐表列和欄標籤。您可以讀取每一個資料格的完整列和欄標籤，或者僅讀取在表格中的資料格之間移動時變更的標籤。



---

## 第 7 章 處理遺漏值

---

### 遺漏值概觀

在資料採礦的資料預備階段，通常需要取代資料中的遺漏值。**遺漏值**是資料集中未知、未收集或已輸入不正確的值。通常，這些值不可用於欄位中。例如，性別欄位應包含值應包含值 *M* 和 *F*。如果在此欄位中探索值 *Y* 或 *Z*，則完全可以確定此值無效，並且應將其解釋為空白值。同樣地，年齡欄位出現負值也毫無意義，應將其解釋為空白值。此類明顯錯誤通常是由於問卷過程中人為已輸入或保留欄位為空白以示拒絕回答造成的。有時候，您可能會進一步檢查這些空白字段，以弄清拒絕提供本人年齡等行為是否會影響具體預測結果。

某些建模技術在處理遺漏值方面具有明顯的優勢。例如，C5.0 和 Apriori 可以很好地處理在類型節點中明確宣告為「遺漏」的值。其他建模技術在處理遺漏值時比較麻煩，並且需要較長的訓練時間，且生成的模型不夠精確。

IBM SPSS Modeler 可識別的遺漏值類型有以下幾種：

- **空值或系統遺漏值。**這兩種類型是資料庫或來源檔案中留空白、並且尚未在來源節點或類型節點中專門定義為「遺漏」的非字符串值。系統遺漏值顯示為 `$null$`。請注意，空字符串在 IBM SPSS Modeler 中不被視為空白值，但它們可能會被某些資料庫視為空白值。
- **空字符串和空白。**空字符串值和空白（帶有不可見字元的字符串）不被視為空值。對於大多數用途，空字符串都視為相當於空白。例如，如果您選取在來源節點或類型節點中將空白視為空白值的選項，則此設定也套用到空字符串。
- **空白值或使用者定義的遺漏值。**這些是在來源節點或類型節點中被明確定義為遺漏的值（例如 unknown、99 或 -1）。您還可以將空白和空白視為空白值，這樣將使得它們被旗標為進行特殊處理並排除在大多數計算之外。例如，您可以使用 @BLANK 函數將這些值以及其他類型的遺漏值處理為空白值。

**混合資料中的讀數。**注意，當您讀取以數值形式（整數、實數、時間、時間戳記或日期）儲存的欄位時，所有非數值型欄位都將設定為空值或系統遺漏值。這是因為 Clementine 與其他應用程式不同，它不容許使用混合儲存類型的欄位。為避免發生這種情況，可以根據需要變更來源節點或外部應用程式中的欄位儲存類型，以字符串的形式讀取包含混合資料的欄位。

**從 Oracle 中讀取空字符串。**在 Oracle 資料庫中進行值的讀取寫入時，要注意，與 IBM SPSS Modeler 及大多數其他資料庫不同，Oracle 將字符型空值等同於空值對待並儲存。這表示同樣的資料從 Oracle 資料庫中解壓縮和從檔案或其他資料庫中解壓縮其表現可能有所不同，可能會傳回不同的結果。

---

### 處理遺漏值

您應根據自己所從事的業務或經營領域常識來確定如何處理遺漏值。為了減少訓練時間和增加精確度，可能需要移除資料集中的空白值。從另一方面講，空白值的出現還可能會帶來新的業務機會或其他靈感。選擇最佳技術時，應考慮資料的下列幾個方面：

- 資料集的大小
- 包含空白值的欄位個數
- 遺漏資訊數量

通常有兩種方法可供選擇：

- 可以排除帶有遺漏值的欄位或記錄

- 可以使用各式各樣的方法歸因、取代或強制遺漏值

使用資料審核節點可以在很大程度上實現上述兩種方法的自動化。例如，可以產生過濾器節點來排除具有過多遺漏值的欄位，這在建模中很有用，並產生 SuperNode 為任何或所有保留欄位歸因遺漏值。這便是審核帶來的實際功能，從而不僅可以評量資料的現行狀態，還可以基於評量結果採取動作。

## 處理帶有遺漏值的記錄

如果大部分遺漏值都集中在少量記錄中，您只需排除這些記錄。例如，銀行通常會儲存詳細而完整的貸款客戶記錄。但是，如果銀行在核准內部職員的貸款時管制不嚴，則收集到的員工貸款資料中可能會存在多個空白欄位。此種情況下，有兩種方法可以處理遺漏值：

- 可以使用選取節點刪除員工記錄。
- 如果資料集較大，可以放棄所有帶有空白值的記錄。

## 處理帶有遺漏值的欄位

如果大部分遺漏值都集中在少量欄位中，您可以在欄位層次而不是記錄層次處理它們。此方法還容許您先檢驗特定欄位相對於建模的重要性，然後確定如何處理遺漏值。如果某個欄位對於建模而言並不重要，那麼可能不值得保留，而與其中包含的遺漏值數目無關。

例如，某市場調查公司可能會從包含 50 個問題的普通問卷中收集資料。很多人拒絕提供年齡和政治派別資訊。此種情況下，年齡和 政治派別就會有很多遺漏值。

### 欄位測量層次

確定要採用的方法時，您還應考慮帶有遺漏值的欄位的測量層次。

**數字欄位。**對於數值型欄位類型（如連續），您應在建立模型前清除所有非數值，因為如果數值型欄位中包含空白值，很多模型將無效。

**種類欄位。**對於種類欄位（如列名和旗標），雖然不必更改遺漏值，但更改後可以增加模型的精度。例如，使用 性別欄位的模型即使含有無意義值（如 Y 和 Z）也仍然有效，不過如果刪除除 M 和 F 以外的值將增加模型的精度。

### 篩選或刪除欄位

要篩選帶有大量遺漏值的欄位，您可以採用以下幾種方法：

- 使用資料審核節點根據品質過濾欄位。
- 您可以使用功能選擇節點來篩選遺漏值超過指定百分比的欄位，並根據相對於特定目標的重要性來對欄位進行排等級。
- 除刪除欄位以外，還可以使用「類型」節點將欄位角色設定為無。這會將這些欄位保留在資料集中，但將它們排除在建模處理程序之外。

## 處理具有遺漏值的記錄

### 什麼是系統遺漏值？

系統遺漏值代表未知或不適用的資料值。在資料庫中，這些值通常稱為空值。

系統遺漏值不同於空白值。空白值通常在「類型」節點中定義為特定值或值的範圍，可以將其視為使用者定義的遺漏。空白值在建模環境定義中以不同方式處理。



## 構造系統遺漏值

在從資料來源讀取的資料中可能出現系統遺漏值（例如，資料庫表格可能包含 *NULL* 值）。

可以通過在表示式中使用值 **undef** 來構造系統遺漏值。例如，下列 CLEM 表示式傳回年齡（如果小於或等於 30）或遺漏值（如果大於 30）：

```
if Age > 30 then undef else Age endif
```

在以下狀況下也可建立遺漏值：執行外部結合時，數字除以零時，計算負數的平方根時，以及其他狀況。

## 顯示系統遺漏值

系統遺漏值在表格和其他輸出中顯示為 `$null$`。

## 正在測試系統遺漏值

如果引數值是系統遺漏值，請使用特殊函數 **@NULL** 傳回 `true`，例如：

```
if @NULL(MyFieldName) then 'It is null' else 'It is not null' endif
```

## 傳送到函數的系統遺漏值

傳送到函數的系統遺漏值通常將遺漏值傳播到輸出。例如，如果欄位 `f1` 的值在特定列中是系統遺漏值，那麼表示式 `log(f1)` 也求值為該列的系統遺漏值。**@NULL** 函數除外。

## 涉及算術運算子的表示式中的系統遺漏值

將算術運算子套用於包含系統遺漏值的值將產生系統遺漏值。例如，如果欄位 `f1` 的值在特定列中是系統遺漏值，那麼表示式 `f1 + 10` 也求值為該列的系統遺漏值。

## 涉及邏輯運算子的表示式中的系統遺漏值

處理涉及邏輯運算子的表示式中的系統遺漏值時，三值邏輯規則（*true*、*false* 和 *missing*）適用並可在事實表格中進行說明。下表格中顯示了常見邏輯運算子 *not*、*and* 和 *or* 的事實表格。

表 4. 針對 *NOT* 的事實表格

運算元	NOT 運算元
true	false
false	true
遺漏	遺漏

表 5. 針對 *AND* 的事實表格

Operand1	Operand2	運算元 1 AND 運算元 2
true	true	true
true	false	false
true	遺漏	遺漏
false	true	false
false	false	false
false	遺漏	false
遺漏	true	遺漏
遺漏	false	false

表 5. 針對 AND 的事實表格 (繼續)

Operand1	Operand2	運算元 1 AND 運算元 2
遺漏	遺漏	遺漏

表 6. 針對 OR 的事實表格

Operand1	Operand2	運算元 1 OR 運算元 2
true	true	true
true	false	true
true	遺漏	true
false	true	true
false	false	false
false	遺漏	遺漏
遺漏	true	true
遺漏	false	遺漏
遺漏	遺漏	遺漏

## 涉及比較運算子的表示式中的系統遺漏值

比較系統遺漏值和非系統遺漏值時，結果求值為系統遺漏值而非 true 或 false 結果。系統遺漏值可以相互比較；兩個系統遺漏值視為相等。

## if/then/else/endif 表示式中的系統遺漏值

使用條件式表示式並且條件式表示式傳回系統遺漏值時，else 子句中的值從條件式表示式傳回。

### 「選取」節點中的系統遺漏值

如果針對特定記錄，選擇表示式求值為遺漏值，那麼記錄不是從「選取」節點輸出（此動作適用於「包含」和「捨棄」方式）。

### 「合併」節點中的系統遺漏值

使用索引鍵進行合併時，不會合併在索引鍵欄位中具有系統遺漏值的任何記錄。

## 聚集中的系統遺漏值

聚合欄中的資料時，在計算中不包含遺漏值。例如，在具有三個值 { 1、2 和 undef } 的欄中，欄中值的總和計算為 3；平均數計算為 1.5。

## 歸因或填充遺漏值

在僅有幾個遺漏值的情況下，可以用插入值取代空白值。可以在資料審核報告中實現上述操作，在此報告中您可以為特定欄位指定相應選項，然後產生一個 SuperNode 採用多種方法對值進行歸因。這種方法最為靈活，還可以指定在單個節點中處理大量欄位。

下列方法可用於輸入遺漏值：

**固定。** 替換為固定值（可以欄位平均數、範圍中間值，或者您指定的常數）。

隨機。替換為基於正常分佈或均勻分佈產生的隨機值。

表示式。用於指定自訂表示式。例如，您可以使用設定廣域節點建立的廣域變數取代數值。

演算法。基於 C&RT 演算法替換為模型預測的值。對於使用此方法輸入的每個欄位，都會有一個個別的 C&RT 模型，還有一個填入節點會使用該模型預測的值取代空值與無效。然後使用過濾器節點刪除該模型產生的預測欄位。

如果還要為特定欄位強制指派值，則可以使用類型節點來確保欄位類型僅包含合法值，然後將需要取代空白值欄位的 檢查欄設定為強制。

---

## 用於遺漏值的 CLEM 函數

有多個函數可用於處理遺漏值。選取節點和填入節點中經常會用下列函數來放棄或填入遺漏值：

- `count_nulls(LIST)`
- `@BLANK(FIELD)`
- `@NULL(FIELD)`
- `undef`

@ 函數可以與 @FIELD 函數一起使用來識別一個或多個欄位中是否存在空白值或空值。這些欄位可以予以標示（存在空白值或空值時），也可以用取代值填充，或者用於其他各種作業。

如下所示，您可以計算欄位清單中的空值數目：

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

如果要使用接受輸入類型的欄位清單的函數，則可以使用特定的 @FIELDS\_BETWEEN 和 @FIELDS\_MATCHING 函數，如下例所示：

```
count_nulls(@FIELDS_MATCHING('card*'))
```

可以使用 `undef` 函數來填充帶有系統遺漏值的欄位，系統遺漏值顯示為 **\$null\$**。例如，取代數值型時可以使用條件式陳述式，如：

```
if not(Age > 17) or not(Age < 66) then undef else Age endif
```

此操作將用系統遺漏值來取代所有不在該範圍內的值，系統遺漏值顯示為 **\$null\$**。借助 `not()` 函數，您可以獲取所有其他數值型，包含任何負值。請參閱第 134 頁的『處理空白值與空值的函數』主題，以取得更多資訊。

### 放棄記錄時的注意事項

請注意，使用選取節點放棄記錄時，語法將使用三值邏輯並在選取陳述式中自動包含空值。要排除選取表示式中的空值（系統遺漏值），必須在表示式中使用 `and not` 明確指出此意圖。例如，要選取並包含處方藥類型為藥品 C 的所有記錄，您可以使用下列選取陳述式：

```
Drug = 'drugC' and not(@NULL(Drug))
```

早期版本遇到上述狀況時會排除空值。



---

## 第 8 章 建立 CLEM 表示式

---

### 關於 CLEM

表示式操作控制語言 (CLEM) 是一種用於分析和操作在 IBM SPSS Modeler 串流中資料流程的功能強大的語言。資料採礦者在串流作業中大量使用 CLEM 來執行作業，例如根據成本衍生利潤這種簡單的作業或是將 Web 日誌資料轉換成一組含有用資訊的欄位和記錄。

在 IBM SPSS Modeler 中，CLEM 用於：

- 對加載記錄欄位上的條件進行比較和評估。
- 為生成新欄位衍生值。
- 為現有欄位衍生新值。
- 合理安排記錄序列。
- 將記錄中的資料插入到報告中。

CLEM 表示式對於 IBM SPSS Modeler 中的資料預備是不可或缺的，此外它還廣泛適用於從記錄和欄位作業（選取、平衡和填入器）到統計圖和輸出（分析、報告和表格）的一系列節點。例如，您可以在衍生節點中使用 CLEM 來根據公式（如率）建立一個新欄位。

CLEM 表示式還可用於廣域搜尋和取代作業。例如，可在填入器節點中使用表示式 @NULL(@FIELD) 將系統遺漏值 取代為整數值 0。（如果要取代 使用者遺漏值，即空白值，則可使用 @BLANK 函數。）

還可以建立更多較為複合的 CLEM 表示式。例如，可以根據一組有條件規則來衍生新欄位，例如使用下列表示式建立的新值種類：If: CardID = @OFFSET(CardID,1), Then: @OFFSET(ValueCategory,1), Else: 'exclude'。

此範例使用 @OFFSET 函數來表達如下的意思：如果某一給定記錄的 CardID 欄位值和先前記錄的該欄位值相同，則傳回先前記錄的名為 ValueCategory 欄位的值。否則，將為其賦予字串 "exclude"。換言之，如果鄰近記錄的 CardID 值相同，就應該賦予它們相同的 ValueCategory。（帶有字串 exclude 的記錄可在稍後通過選取節點呼叫。）

---

### CLEM 範例

為闡釋使用 CLEM 的正確語法和可能的表示式類型，我們給出了表示式的範例。

#### 簡單表示式

下面這個公式可根據欄位 *After* 和 *Before* 的值衍生一個新欄位，您的公式也可以像這個公式一樣簡單。

$(After - Before) / Before * 100.0$

注意：在參照欄位值的時候不需要給欄位名稱加引號。

類似的，如下的表示式只是簡單地傳回欄位 *salary* 的對數值。

$\log(salary)$

#### 複式表示式

表示式也可以很長而且非常複合。如果兩個欄位 (*\$KX-Kohonen* 和 *\$KY-Kohonen*) 的值落在指定範圍內，如下表示式將傳回 *true*。注意：由於此處的欄位名稱包含有特殊字元，因此給它們加上了單引號。

```
('KX-Kohonen' >= -0.2635771036148072 and 'KX-Kohonen' <= 0.3146203637123107  
and 'KY-Kohonen' >= -0.18975617885589602 and  
'KY-Kohonen' <= 0.17674794197082522) -> T
```

少數幾個函數（如字串函數）需要使用正確的語法鍵入幾個參數值。下面例子中使用的函數 *subscrs* 將傳回 *produce\_ID* 欄位的第一個字元，用以指明該物品究竟是有的有機、轉基因還是常規。表示式的結果將由 `-> `result`` 來說明。

```
subscrs(1,produce_ID) -> `c`
```

類似的表示式有：

```
stripchar(`3`,`123`) -> `12`
```

需要注意的是字元要用反單引號括起來。

表示式中的組合函數

通常，CLEM 表示式是由組合函數構成的。下列函數將 *subscr* 和 *lowertoupper* 結合使用，以傳回 *produce\_ID* 的第一個字元並將其轉換為大寫。

```
lowertoupper(subscr(1,produce_ID)) -> `C`
```

同樣的表示式還可以簡寫入為：

```
lowertoupper(produce_ID(1)) -> `C`
```

另一個經常使用的函數組合是：

```
locchar_back(`n`, (length(web_page)), web_page)
```

該表示式用以確定從 *web\_page* 欄位值的最後一個字元往前數起字元 ``n`` 所在的位置。通過引入 *length* 函數，表示式可動態計算目前值的長度從而避免了使用可能無效的靜態數字（以 7 為例，如果值的長度少於 7 個字元，將導致無效）。

特殊函數

有許多可用的特殊函數（函數名前加有符號 `@`）。一般的函數包含：

```
@BLANK('referrer ID') -> T
```

通常，特殊函數通過組合方式加以使用，這是一次在多個欄位中標記空白的常用方法。

```
@BLANK(@FIELD)-> T
```

CLEM 文件中還將討論更多的範例。請參閱第 109 頁的『CLEM 參照概觀』主題，以取得更多資訊。

---

## 值和資料類型

CLEM 表示式類似於由值、欄位名稱、運算子和函數構成的公式。最簡單的有效 CLEM 表示式是一個值或一個欄位名稱。有效值的範例如下：

```
3 1.79  
'banana'
```

欄位名稱的範例如下：

Product\_ID'\$P-NextField'

其中, *Product* 是市場購物籃資料集中的一個欄位名, '*\$P-NextField*' 是一個參數名, 而表示式的值則是一個指定欄位的值。通常, 欄位名稱以一個以字母開頭, 並可以包含數字和底線 (`_`)。如果使用引號將名稱括起來, 您不必遵守上述命名規則。CLEM 值可以是如下所列的任何一種類型:

- 字串 - 例如, "c1"、"類型 2" 以及 "a piece of free text"
- 整數 - 例如 12、0 和 -189
- 實數 - 例如 12.34、0.0 和 -0.0045
- 日期/時間欄位 - 例如, 05/12/2002、12/05/2002 和 12/05/02

還有可能使用到下列元素:

- 字元碼 - 例如 `a` or 3
- 項目清單 - 例如, [1 2 3] 和 ['類型 1' '類型 2']

字元碼和清單一般不作為欄位值出現。通常, 它們用作 CLEM 函數的引數。

### 參照規則

雖然軟體在確定 CLEM 表示式中某部分是欄位、值、參數還是字串時非常靈活, 但在建立表示式時遵循下列的一般規則將十分有利於您的實踐操作:

- **字串** - 在撰寫字串時始終使用雙引號 ("類型 2" 或 "value")。可改為使用單引號, 但可能會與用引號括住的欄位混淆。
- **字元** - 始終使用反單引號, 例如 `。例如, 請注意函數 `stripchar(`d`,`drugA")` 中的字元 `d`。此類情形中的唯一例外狀況是使用整數來參照字串中的特定字元。例如, 請注意函數 `lowertoupper("druga"(5)) -> "A"` 中的字元 `5`。註: 在標準的 U.K. 和 U.S. 鍵盤上, 反引號字元 (重音符, Unicode 0060) 鍵就在 Esc 鍵的正下方。
- **欄位** - 在 CLEM 表示式中使用欄位時, 通常不需要加引號 (`subscr(2,arrayID)`) ->字元)。您只有在需要將空格或其他特殊符號括入時才使用單個 (如 'Order Number')。用引號括住但未定義在資料集中的欄位將被誤讀為字串。
- **參數** - 始終使用單引號 ('\$P-threshold')。

---

## 表示式和條件

CLEM 表示式可以傳回一個結果 (在衍生新值時使用) - 例如:

```
Weight * 2.2Age + 1sqrt(Signal-Echo)
```

或者, 它們也可判斷 *true* 或 *false* (根據某一條件進行選取時使用) - 例如:

```
Drug = "drugA"Age < 16not(PowerFlux) and Power > 2000
```

您可以在 CLEM 表示式中任意組合使用運算子和函數 - 例如:

```
sqrt(abs(Signal)) * max(T1, T2) + Baseline
```

括號和運算子優先級將確定表示式求值的順序。本例中, 求值的順序為:

- 首先計算 `abs(Signal)`, 隨後再將 `sqrt` 應用於上一步計算結果。
- 計算 `max(T1, T2)`。
- 上述兩個結果相乘: **x** 的優先級高於 **+**。
- 最後, 再將 `Baseline` 與結果相加。

優先級按遞減排列（即，從第一個執行的作業到最後一個作業）如下：

- 函數引數
- 函數呼叫
- **xx**
- **x / mod div rem**
- + -
- > < >= <= /== == = /=

如果您希望置換優先級，或者您對求值順序不確定，利用圓括弧可使求值順序固定下來 - 例如：

```
sqrt(abs(Signal)) * (max(T1, T2) + Baseline)
```

---

## Stream、Session 及 SuperNode 參數

可以定義用於 CLEM 表示式和 Script 編寫的參數。這些參數在效果上等同於使用者定義變數，這些變數可以在目前串流、階段作業或 SuperNode 中儲存和保留，並且可以從使用者介面和 Script 中存取。例如，如果儲存串流，則為此串流設定的所有參數也將儲存。（此特點使這些參數區別於僅在 Script 中使用和宣告的本端 Script 變數。）通常將參數用於 Scripting 來控制 Script 的行為，可通過提供無需在 Script 中寫在程式中的欄位和值的相關資訊。

參數的範圍取決於設定參數的位置：

- 串流參數可以在串流 Script 中或在串流內容對話框中設定，且可用於串流中的所有節點。這些參數顯示在表示式建置器中的「參數」清單中。
- 工作階段參數可以在獨立式 Script 中或在工作階段參數對話框中設定。這些階段作業參數可用於現行階段作業中使用的所有串流（即在管理程式窗格的「串流」標籤中列出的所有串流）。

也可以為 SuperNode 設定參數，在這種情況下，這些參數僅對於該 SuperNode 中封裝的節點可見。

在 CLEM 表示式中使用參數

CLEM 表示式中的參數由 \$P-pname 代表，其中 pname 是參數名。在 CLEM 表示式中，使用的參數必須用單引號括起來 - 例如 '\$P-scale'。

可用的參數可直接通過表示式建置器來檢視。檢視目前的參數：

1. 在任何可使用 CLEM 表示式的對話框中，按一下「表示式建置器」按鈕。
2. 從「欄位」清單中，選取參數。

從清單中選取要插入項目到 CLEM 表示式的參數。請參閱第 105 頁的『選取欄位、參數和廣域變數』主題，以取得更多資訊。

---

## 字串的處理

有很多可用於字串的作業，包含：

- 將字串轉換為大寫或小寫 - `uppertolower(CHAR)`。
- 消除某個字串變數中指定的字元（例如 `ID\_` 或 `\$`） - `stripchar(CHAR,STRING)`。
- 確定字串變數的長度（字元的數量） - `length(STRING)`。
- 檢查字串值的按字母順序 - `alphabefore(STRING1, STRING2)`。
- 消除值中的前端或尾部空白 - `trim(STRING)`、`trim_start(STRING)` 或 `trimend(STRING)`。



- 從字串中擷取開頭的或末尾的  $n$  個字元 - `startstring(LENGTH, STRING)` 或 `endstring(LENGTH, STRING)`。例如，假設您有一個名為 *item* 的欄位，它將產品名和一個 4 位 ID 代碼組合在一起（如 ACME CAMERA-D109）。要建立僅含 4 位代碼的新欄位，請在衍生節點中指定下列公式：

```
endstring(4, item)
```

- 相符特定的格式 - `STRING matches PATTERN`。例如，要選取工作職稱中任意位置具有 "market" 的職員，則可以在選取節點中指定下列內容：

```
job_title matches "*market*"
```

- 取代字串內出現的全部某子字串 - `replace(SUBSTRING, NEWSUBSTRING, STRING)`。例如，要在文字採集之前使用分號取代不受支援字元（如豎線 ( | )）的所有實例，需使用填入節點中的 `replace` 函數。在填入欄位中，選取可能出現該字元的所有欄位。對於取代：條件，請選取始終，並在取代為：中指定下列條件：

```
replace('|',';',@FIELD)
```

- 根據顯示的特定子字串衍生一個旗標欄位。例如，您可以使用衍生節點中的字串函數為每個回應產生個別的旗標欄位，所使用的表示式如下：

```
hassubstring(museums,"museum_of_design")
```

請參閱第 122 頁的『字串函數』主題，以取得更多資訊。

---

## 處理空白值和遺漏值

對資料挖掘人員而言，取代空白值和遺漏值是常見的資料預備作業。CLEM 向您提供了多種自動處理空白值的工具。最常用來處理空白值的是填入節點；除此以外，在任何可以使用 CLEM 表示式的節點中還可使用下列函數：

- `@BLANK(FIELD)` 可用於在特定欄位（如 *Age*）中找出值為空白的記錄。
- `@NULL(FIELD)` 可用於確定指定欄位中值為系統遺漏值的記錄。在 IBM SPSS Modeler 中，系統遺漏值顯示為 **\$null\$** 值。

請參閱第 134 頁的『處理空白值與空值的函數』主題，以取得更多資訊。

---

## 數字的處理

在 IBM SPSS Modeler 中，有很多可用於數值的標準作業，如：

- 計算指定角度的正弦值 - `sin(NUM)`
- 計算數值型欄位的自然對數值 - `log(NUM)`
- 計算兩個數字的和 - `NUM1 + NUM2`

請參閱第 118 頁的『數值函數』主題，以取得更多資訊。

---

## 時間和日期的處理

時間和日期格式可能會因為您資料來源和語言環境的不同而有所不同。可在「串流內容」對話框中設定具體到每個串流的日期與時間格式。下面的範例是處理日期/時間欄位時常用的函數。

計算已流逝的時間

利用與下列函數相似的一系列函數，您可以輕鬆的計算出從基準線日期開始的時間流逝。該函數傳回一個實數值，用以代表從基準線日期至日期字串 `DATE` 所代表的日期之間的時間長度（以月為計算單位）。這是基於每月 30.0 天的近似數字。

date\_in\_months(日期)

日期/時間值的比較

利用與下列函數類似的函數，可以在記錄之間比較日期/時間欄位的值。如果日期字串 DATE1 代表的日期早於日期字串 DATE2 代表的日期，函數將傳回 *true*。否則，此函數會傳回值 0。

date\_before(DATE1,DATE2)

計算差分

您可以使用函數來計算兩個時間或兩個日期之間的差分，如：

date\_weeks\_difference(DATE1,DATE2)

該函數將傳回一個實數值，用以代表從日期字串 DATE1 代表的日期至日期字串 DATE2 代表的DATE之間的時間長度（以週為計算單位）。這基於一週 7.0 天。如果 DATE2 早於 DATE1，該函數將傳回負值。

當前日期

可使用函數 @TODAY 將現行日期新增到資料集之中。當前日期可作為字串，以「串流內容」對話框中選定的日期格式新增到指定欄位或新欄位中。請參閱第 126 頁的『日期和時間函數』主題，以取得更多資訊。

---

## 多個欄位的彙總

CLEM 語言包含了一組可傳回涉及多個欄位的彙總統計資料的函數。這些函數在分析意見調查資料時特別有用，因為意見調查資料中某個問題的多個答案會儲存在多個欄位中。請參閱第 101 頁的『多個回應資料的處理』主題，以取得更多資訊。

比較函數

您可以使用 min\_n 和 max\_n 函數在多個欄位值之間進行比較，例如：

```
max_n(['card1fee' 'card2fee' 'card3fee' 'card4fee'])
```

您還可以使用一些計數函數來獲得符合特定準則的值的計數，即便這些值儲存在多個欄位中也無妨。例如，計算持有時間在五年以上的卡片的數量：

```
count_greater_than(5, ['cardtenure' 'card2tenure' 'card3tenure'])
```

若要計算相同欄位集中的空值數量：

```
count_nulls(['cardtenure' 'card2tenure' 'card3tenure'])
```

注意：本例中的計數是持有的卡片的數量，並非持有卡片的人數。請參閱第 116 頁的『比較函數』主題，以取得更多資訊。

要計算指定值在多個欄位中出現的次數，您可以使用 count\_equal 函數。下例計算在清單中包含值 Y 的欄位個數。

```
count_equal("Y",[Answer1, Answer2, Answer3])
```

假設清單中的欄位具有下列值，則該函數對於值 Y 傳回所示結果。

表 7. 函數值。

Answer1	Answer2	Answer3	計數
Y	N	Y	2

表 7. 函數值 (繼續).

Answer1	Answer2	Answer3	計數
Y	N	N	1

## 數值函數

您可以使用 `sum_n`、`mean_n` 和 `sdev_n` 函數來獲取涉及多個欄位的統計資料，例如：

```
sum_n(['card1ba1' 'card2ba1''card3ba1'])  
mean_n(['card1ba1' 'card2ba1''card3ba1'])
```

請參閱第 118 頁的『數值函數』主題，以取得更多資訊。

## 產生欄位清單

在使用任何可以將欄位清單作為輸入的函數時，特定的函數 `@FIELDS_BETWEEN(start, end)` 和 `@FIELDS_MATCHING(pattern)` 可用作輸入。例如，假定欄位順序依照之前 `sum_n` 範例所示，則下面的表示式等效：

```
sum_n(@FIELDS_BETWEEN(card1ba1, card3ba1))
```

或者，要計算所有以 `card` 起始的欄位中的空值的數量：

```
count_nulls(@FIELDS_MATCHING('card*'))
```

請參閱第 134 頁的『特殊欄位』主題，以取得更多資訊。

---

## 多個回應資料的處理

有許多比較函數可用於分析多個回應資料，其中包含：

- `value_at`
- `first_index/last_index`
- `first_non_null / last_non_null`
- `first_non_null_index / last_non_null_index`
- `min_index / max_index`

例如，假設有一個多回應問題，內容是在個人購買過程中最看重的三個因素（如價格、他人推薦、評論、本端供應商等）。在這種情況下，您可以通過衍生首次併入價格的欄位索引，來確定價格重要性：

```
first_index("price", [Reason1 Reason2 Reason3])
```

同樣地，假定您已讓客戶按照購買的可能性對三輛汽車進行了排等級，並且已經在三個不同的欄位中對這些答案進行了編碼，如下所示：

表 8. 汽車分級範例.

客戶 ID	car1	car2	car3
101	1	3	2
102	3	2	1
103	2	3	1

在這種情況下，您可以使用 `min_index` 函數來確定他們最喜歡（排名第一或最低排名）的汽車欄位索引：

```
min_index(['car1' 'car2' 'car3'])
```

請參閱第 116 頁的『比較函數』主題，以取得更多資訊。

參照多重回應集

@MULTI\_RESPONSE\_SET 特殊函數可用於參照多重回應集中的所有欄位。例如，若前例中的三個 *car* 欄位都包含在名為 *car\_rankings* 的多重回應集中，則下面的函數將傳回相同的結果：

```
max_index(@MULTI_RESPONSE_SET("car_rankings"))
```

---

## 表示式建置器

您可手動鍵入 CLEM 表示式，也可使用表示式建置器，後者將顯示一個 CLEM 函數和運算子的完整清單以及目前串流中的資料欄位，因此，您無需記住欄位或函數的準確名稱即可快速建立表示式。此外，「建置器」控制項自動為欄位和值新增適當的參照，使得建立語法正確的表示式變得更加容易。

註：「表示式建置器」在 Script 編譯或參數設定中並不受支援。

註：如果要變更資料來源，那麼在變更資料來源之前，您應檢查表示式建置器是否仍支援選取的函數。由於並非所有資料庫都支援所有函數，因此對新資料來源執行時期，可能會發生錯誤。

## 存取表示式建置器

所有使用 CLEM 表示式的節點中都有「表示式建置器」，這些節點包含選取、平衡、衍生、填入器、分析、報告和表格。按一下公式欄位右側的「計算機」按鈕即可將其開啟。

## 建立表示式

「表示式建置器」不僅提供了欄位、函數和運算子的完整清單，如果您的資料已經實例化，它還可以用來存取資料值。

使用「表示式建置器」建立表示式

1. 參照函數和欄位清單，在表示式欄位鍵入公式。

或

2. 從捲動的清單中選取所需的欄位和函數。
3. 按兩下或按一下黃色箭頭按鈕可將欄位或函數新增到表示式欄位中。
4. 使用位於對話框中心的運算元按鈕可將作業插入到表示式中。

## 選取函數

函數清單顯示畫面了所有可用的 CLEM 函數和運算子。捲動清單以從中選取函數，為了更方便搜尋，也可使用下拉清單來顯示函數或運算子的子集合。可用的函數按種類群組，以便更方便的搜尋。

在 CLEM 語言說明的參照部分中說明了其中的多數種類。如需相關資訊，請參閱第 114 頁的『函數參照』。

其他種類如下所示：

- 一般函數包含精選的最一般函數。
- 最近使用包含了目前對話中所使用的 CLEM 函數的清單。
- @ 函數包含所有特殊函數的清單，其名稱前面帶有 "@" 符號。
- 資料庫函數。如果串流包含資料庫連線（通過「資料庫來源」節點），該選項列出在此資料庫中可用的函數，包含使用者定義的函數 (UDF)。如需相關資訊，請參閱第 103 頁的『資料庫函數』。

- **資料庫聚集。**如果串流包含資料庫連線（通過「資料庫來源」節點），該選項列出在此資料庫中可用的聚集選項。這些選項在「聚合」節點的表示式建置器中提供。
- **資料庫視窗聚集。**如果串流包含資料庫連線（通過「資料庫來源」節點），該選項列出在此資料庫中可以使用的視窗聚集選項。這些選項僅在欄位作業選用區中的節點內的表示式建置器中提供。

註：由於 SPSS Modeler 從資料庫系統視圖中獲取 **Window 聚合函數**，因此可用的選項取決於資料庫行為。

儘管稱這些選項為「聚合」，但其並不是設計用於「聚合」節點；這些選項更適用於「衍生」或「選取」這樣的節點。原因是這些選項的輸出是標數量而不是 true 正的聚合，也就是說，它們與「聚合」節點一樣不會減少輸出中顯示的資料的數量。例如，您可以使用此種聚集在資料列中提供諸如「目前列和所有上一列的平均值」這樣的移動均數。

- **內建聚集。**包含可用聚集的可能模式的清單。
- **運算子**列出在建立表示式時可使用的所有運算子。另外，也可以通過對話框中央的按鈕來使用運算子。
- **全部函數**包含了可用 CLEM 函數的完整清單。

選定了一群組函數之後，按兩下即可將函數插入到表示式欄位中游標指示的位置。

## 資料庫函數

資料庫函數可以在多數不同位置列出；下表格顯示了 SPSS Modeler 查找函數詳細資料時搜尋的位置。資料庫管理者可以使用此表格來確保使用者有權存取所需區域以能夠使用不同的函數。

另外，此表格列出了一些條件，這些條件用於在某個函數可供使用時根據資料庫和函數類型進行過濾。

註：如果使用 Amazon Redshift 中的資料庫函數，您的資料庫管理者可能需要授權您對下列六個資料庫物件的許可權。前四個是系統型錄表格，最後一欄兩個是綱目。

- pg\_type
- pg\_proc
- pg\_namespace
- pg\_aggregate
- information\_schema
- pg\_catalog

表 9. 表示式建置器中的資料庫函數

資料庫	函數類型	函數的尋找位置	用於過濾函數的條件
Db2 LUW	UDF	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE 為 F，而 FUNCTIONTYPE 為 S
Db2 LUW	UDA	SYSCAT.ROUTINES SYSCAT.ROUTINEPARMS	ROUTINETYPE 為 F，而 FUNCTIONTYPE 為 C
Db2 iSeries	UDF	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE 為 F，而 FUNCTION_TYPE 為 S
Db2 iSeries	UDA	QSYS2.SYSROUTINES QSYS2.SYSPARMS	ROUTINE_TYPE 為 F，而 FUNCTION_TYPE 為 C
Db2 z/OS	UDF	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE 為 F，而 FUNCTIONTYPE 為 S
Db2 z/OS	UDA	SYSIBM.SYSROUTINES SYSIBM.SYSPARMS	ROUTINETYPE 為 F，而 FUNCTIONTYPE 為 C

表 9. 表示式建置器中的資料庫函數 (繼續)

資料庫	函數類型	函數的尋找位置	用於過濾函數的條件
SQL Server	UDF	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE 為 FN 或 FS
SQL Server	UDA	SYS.ALL_OBJECTS SYS.ALL_PARAMETERS SYS.TYPES	TYPE 為 AF
Oracle	UDF	ALL_ARGUMENTS ALL_PROCEDURES	滿足下列所有條件： <ul style="list-style-type: none"> <li>OBJECT_TYPE 為 FUNCTION</li> <li>AGGREGATE 為 NO</li> <li>PLS_TYPE 不為 NULL</li> </ul>
Oracle	UDA	ALL_ARGUMENTS ALL_PROCEDURES	滿足下列所有條件： <ul style="list-style-type: none"> <li>ARGUMENT_NAME 為 NULL</li> <li>AGGREGATE 為 YES</li> <li>PLS_TYPE 不為 NULL</li> </ul>
Teradata	UDF	DBC.FUNCTIONS DBC.ALLRIGHTS	滿足下列所有條件： <ul style="list-style-type: none"> <li>FUNCTIONTYPE 為 F</li> <li>COLUMNNAME 為 RETURN0</li> <li>SPPARAMETERTYPE 為 0</li> <li>ACCESSRIGHT 為 EF</li> </ul>
Teradata	UDA	DBC.FUNCTIONS DBC.ALLRIGHTS	滿足下列所有條件： <ul style="list-style-type: none"> <li>FUNCTIONTYPE 為 A</li> <li>COLUMNNAME 為 RETURN0</li> <li>SPPARAMETERTYPE 為 0</li> <li>ACCESSRIGHT 為 EF</li> </ul>
Netezza	UDF	####.._V_FUNCTION NZA.._V_FUNCTION INZA.._V_FUNCTION	對於 ####.._V_FUNCTION，下列條件適用： <ul style="list-style-type: none"> <li>RESULT 不包含具有值的字串，例如 TABLE%</li> <li>FUNCTION 不包含具有值的字串，例如 '/_%' escape '/'</li> <li>VARARGS 為 FALSE</li> </ul> 下列條件適用於 NZA.._V_FUNCTION 和 INZA.._V_FUNCTION： <ul style="list-style-type: none"> <li>RESULT 不包含具有值的字串，例如 TABLE%</li> <li>FUNCTION 不包含具有值的字串，例如 '/_%' escape '/'</li> <li>BUILTIN 為 f</li> <li>VARARGS 為 FALSE</li> </ul>
Netezza	UDA	####.._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	同時滿足下列兩個條件： <ul style="list-style-type: none"> <li>AGGTYPE 為 ANY 或 GROUPED</li> <li>VARARGS 為 FALSE</li> </ul>

表 9. 表示式建置器中的資料庫函數 (繼續)

資料庫	函數類型	函數的尋找位置	用於過濾函數的條件
Netezza	WUDA	####._V_AGGREGATE NZA.._V_FUNCTION INZA.._V_FUNCTION	對於 ####._V_AGGREGATE，下列條件適用： <ul style="list-style-type: none"> <li>• AGGTYPE 為 ANY 或 ANALYTIC</li> <li>• AGGREGATE 不為 MAX_LABEL</li> <li>• VARARGS 為 FALSE</li> </ul> 下列條件適用於 NZA.._V_FUNCTION 和 INZA.._V_FUNCTION： <ul style="list-style-type: none"> <li>• AGGTYPE 為 ANY 或 ANALYTIC</li> <li>• BUILTIN 為 f</li> <li>• VARARGS 為 FALSE</li> </ul>

### 表格中使用的項目關鍵字

- UDF 使用者定義的函數
- UDA 使用者定義的聚集函數
- WUDA 使用者定義的視窗聚集函數
- #### 目前連接至的資料庫。

### 選取欄位、參數和廣域變數

欄位清單顯示了資料串流的這個位置上的所有可用欄位。捲動清單以從中選取一個欄位。按兩下或按一下黃色箭頭按鈕可將欄位新增到表示式中。

請參閱第 98 頁的『Stream、Session 及 SuperNode 參數』主題，以取得更多資訊。

除了這些欄位，您還可以從下列項目中進行選擇：

**多回應集。**如需相關資訊，請參閱 *IBM SPSS Modeler Source, Process, and Output Nodes* 手冊。

**最近使用**包含了目前對話中所使用的欄位、多重回應集、參數和廣域值的清單。

**參數。**請參閱第 98 頁的『Stream、Session 及 SuperNode 參數』主題，以取得更多資訊。

**廣域值。**如需相關資訊，請參閱 *IBM SPSS Modeler Source, Process, and Output Nodes* 手冊。

### 檢視或選取值

系統的很多位置都可以用來檢視欄位值，包含表示式建置器、資料審核報告以及在時間間隔節點中編輯未來值時。注意：要使用該功能，資料必須在來源或類型節點中完全實例化，以便掌握資料的儲存、類型和值。

要從「表示式建置器」或時間間隔節點檢視欄位值，請選取要檢視的欄位並按一下「值選取器」按鈕，這將開啟一個列出選取的欄位欄位值的對話框。隨後可選取一個值並按一下**插入**已將值貼上到目前表示式或清單中。



圖 14. 「值選取器」按鈕

對於旗標和列名欄位，將列出所有定義的值。對於連續（數值型範圍）欄位，顯示下限和最大值。

## 檢查 CLEM 表示式

按一下「表示式建置器」中的檢查（右下角）以驗證表示式。尚未檢查的表示式顯示為紅色。如果遇到錯誤，將顯示一條指示起因的訊息。

將檢查下列項目：

- 值和欄位名稱的正確參照
- 參數和廣域變數的正確使用
- 運算子的有效使用
- 參照的欄位存在
- 參照廣域的存在及定義

如果您遇到語法錯誤，請嘗試使用清單和運算子按鈕來建立表示式，而不要手動鍵入表示式。這種方法可自動為欄位和值新增適當的參照。

在 IBM Analytical Decision Management 中建立表示式時，請注意下列限制。表示式不得包含下列任意項目：

- 對 IBM SPSS Modeler 串流參數的參照
- 對 IBM SPSS Modeler 串流廣域量的參照
- 對資料庫函數的參照
- 對下列其中一個特殊欄位或欄位值 @ 函數的參照：
  - @TARGET
  - @PREDICTED
  - @FIELD
  - @PARTITION\_FIELD
  - @TRAINING\_PARTITION
  - @TESTING\_PARTITION
  - @VALIDATION\_PARTITION

---

## 尋找/取代

可在編輯 Script 或表示式文字的位置（包括 Script 編輯器和 CLEM 表示式建置器）或定義「報告」節點中的範本時使用「尋找/取代」對話框。在上述任一區域編輯文字時，按 Ctrl+F 可存取此對話框，並確保游標的對焦點位於文字區域中。例如，處理填入節點時，可以通過「設定」標籤的任一文字區域或表示式建置器中的文字欄位存取此對話框。

1. 使用文字區域中的游標，按 Ctrl+F 可存取「尋找/取代」對話框。
2. 輸入要搜尋的文字，或從最近搜尋項目下拉清單中選擇。
3. 輸入取代文字（如果有的話）。
4. 按一下尋找下一個開始搜尋。
5. 按一下取代取代目前選定的內容，或按一下全部置換更新所有項或選定的實例。
6. 每次作業完成後，此對話框將關閉。從任一文字區域中按 F3 鍵，可重複上一次尋找作業，或按 Ctrl+F，可再次存取該對話框。

### 搜尋選項

區分觀察值。指定尋找作業是否區分觀察值；例如 *myvar* 是否與 *myVar* 相符。無論怎樣設定，取代文字始終完全按照已輸入插入。



**僅尋找全字。**指定尋找作業是否相符單詞中內部的文字。如果已選取，*spider* 的搜尋結果將不會包括 *spiderman* 或 *spider-man*。

**一般表示式。**指定是否使用正規表示式語法（請參閱下一節）。如果已選取，**僅尋找全字**選項將停用並且會忽略其值。

**僅選取的文字。**控制項使用全部置換選項時的搜尋範圍。

### 正則表示式語法

使用一般表示式，您可以搜尋特殊字元（如標籤或新行）、字元的類別或範圍（如 *a* 到 *d*）、任何數字或非數字以及範圍（如行首或行尾）。受支援的表示式類型如下。

表 10. 字元相符項.

字元	相符項
x	字元 x
\\	反斜線字元
\0n	含八進位值的字元 0n (0 <= n <= 7)
\0nn	含八進位值的字元 0nn (0 <= n <= 7)
\0mnn	含八進位值的字元 0mnn (0 <= m <= 3, 0 <= n <= 7)
\xhh	含十六進位值的字元 0xhh
\uhhhh	含十六進位值的字元 0xhhhh
\t	定位點 ('\u0009')
\n	換行符 ('\u000A')
\r	回車符 ('\u000D')
\f	換頁符 ('\u000C')
\a	警示 (蜂鳴) 符 ('\u0007')
\e	跳出符 ('\u001B')
\cx	x 對應的控制字元

表 11. 符合字元類別.

字元類別	相符項
[abc]	a、b、或 c (簡單類別)
[^abc]	除 a、b、或 c 之外的所有字元 (相減法)
[a-zA-Z]	a 到 z 或 A 到 Z，包含 (範圍)
[a-d[m-p]]	a 到 d 或 m 到 p (合併)。也可指定為 [a-dm-p]
[a-z&&[def]]	a 到 z 和 d、e、或 f (交集)
[a-z&&[^bc]]	a 到 z，除 b 和 c 外 (相減法)。也可指定為 [ad-z]
[a-z&&[^m-p]]	a 到 z，而非 m 到 p (相減法)。也可指定為 [a-lq-z]

表 12. 預先定義的字元類別.

預先定義的字元類別	相符項
.	任意字元 (可能或不可能與行終止符相符)
\d	任意數字：[0-9]
\D	非數字：[^0-9]

表 12. 預先定義的字元類別 (繼續).

預先定義的字元類別	相符項
\s	空格字元：[ \t\n\x0B\f\r]
\S	非空格字元：[^\s]
\w	單詞字元：[a-zA-Z_0-9]
\W	非單詞字元：[^\w]

表 13. 界限相符.

界限比對符	相符項
^	行首
\$	行尾
\b	單詞界限
\B	非單詞界限
\A	輸入的開頭
\Z	除最後終止符外 (如果有) , 輸入的結尾
\z	輸入的結尾

---

## 第 9 章 CLEM 語言參照

---

### CLEM 參照概觀

本節介紹表示式操作控制語言 (CLEM)，它是一個用來分析和操控 IBM SPSS Modeler 串流中所使用資料的強大工具。可以在節點內使用 CLEM 來執行從評估條件或推導值到將資料插入報告等一系列作業。

CLEM 表示式由值、欄位名稱、運算子和函數組成。使用正確的語法，可以建立各式各樣功能強大的資料作業。

---

### CLEM 資料類型

CLEM 資料類型可以由下列任意一項構成：

- 整數
- 實數
- 字元
- 字串
- 清單
- 欄位
- 日期/時間

#### 參照規則

雖然 IBM SPSS Modeler 在確定 CLEM 表示式中使用的欄位、值、參數和字串時非常靈活，但在建立表示式時遵循下列一般規則將有利於您的實踐操作：

- 字串 - 撰寫入字串時一律使用雙引號，如 "Type 2"。可改為使用單引號，但可能會與用引號括住的欄位混淆。
- 欄位 - 只有需要將空格或其他特殊符號括入時才使用單引號，如 'Order Number'。用引號括住但未定義在資料集中的欄位將被誤讀為字串。
- 參數 - 使用參數時始終用英文單引號，如 '\$P-threshold'。
- 字元 - 始終使用反單引號 (`)，如 stripchar(`d`, "drugA")。

下列主題詳細介紹了這些規則。

#### 整數

整數代表為十進制數字序列。可以選擇在整數前加負號 (-) 來表示負數，例如 1234、999 和 -77。

CLEM 語言可以處理任意精準度的整數。上限整數大小取決於您的平台。如果值太大而無法以整數欄位顯示，將欄位類型變更為 Real 通常可以解決這一問題。

#### 實數

實數是指浮點數字。實數的形式為一個或多個數字後跟一個小數點，小數點後再跟一個或多個數字。CLEM 實數為倍精準度數。

可以選擇在實數前加負號 (-) 表示負數，例如 1.234、0.999 和 -77.001。可以使用 <數字> e <指數> 格式以指數表示法表示實數，例如 1234.0e5、1.7e 和 2。當 IBM SPSS Modeler 應用程式從檔案中讀取數字字串並自動將它們轉換為數字時，它接受小數點之前無前導數字或小數點之後無數字的數字，例如，999. 或 .11。不過，這些形式在 CLEM 表示式中是非法的。

註：在 CLEM 表示式中參照實數時，無論目前串流或語言環境的設定如何，都必須使用期間作為小數分隔字元。例如，指定

```
Na > 0.6
```

而不是

```
Na > 0,6
```

即使在串流內容對話框中選取了逗點作為小數符號，此規則也適用，它與代碼語法應獨立於任何特定語言環境或慣例的一般準則一致。

## 字元

字元（通常顯示為 CHAR）通常在 CLEM 表示式中用於對字串進行測試。例如，可以使用函數 `isuppercode` 來驗證某個字串的第一個字元是否為大寫。下面的 CLEM 表示式使用一個字元來表示應對字串的第一個字元進行這樣的測試：

```
isuppercode(subscrs(1, "MyString"))
```

要在 CLEM 表示式中表示特定字元的代碼（相對於位置），請使用 ``<character>`` 形式的反單引號，例如 ``A`` 和 ``Z``。

註：對於欄位而言，沒有 字元 儲存類型，如果使用會產生 字元 儲存類型的表示式衍生或填充欄位，則結果將轉換為字串。

## 字串

通常，應該將字串使用雙引號括住。例如，字串 "c35product2" 和 "referrerID"。要在字串中表示特殊字元，請使用反斜線 - 例如，"`\"`65443"。（要表示反斜線字元，應使用兩個反斜線 `\\`。）也可以使用英文單引號將字串括起，但這樣會導致無法與加引號欄位（'referrerID'）相區別。請參閱第 122 頁的『字串函數』主題，以取得更多資訊。

## 清單

清單是元素的依序排列，其中的元素可以是混合類型。清單以方括弧 ([]) 括住。例如，清單 [1 2 4 16] 和 ["abc" "def"]。清單並不用作 IBM SPSS Modeler 欄位的值。它們為函數提供引數（如 `member` 和 `oneof`）。

註：清單可以僅包含靜態物件（例如字串、數字或欄位名稱），而不包含函數呼叫。

## 欄位

CLEM 表示式中的名稱，如若不是函數名稱，則視其為欄位名稱。這些名稱可以簡單地寫入作 `Power`、`val27` 和 `state_flag` 等，但是，如果名稱以數字開頭或包含非英文字母字元，如空格（底線除外），請用單個將其括起來 - 例如，'`Power Increase`'、'`2nd answer`'、'`#101`'、'`$P-NextField`'。

註：加引號但資料集中未定義的欄位會誤讀取為字串。

## 日期

日期將以在「串流內容」對話框中指定的「基準線」日期為依據進行計算。預設基準線日期為 1900 年 1 月 1 日。

CLEM 語言支援下列日期格式。

表 14. CLEM 語言日期格式

格式	範例
DDMMYY	150163
MMDDYY	011563
YYMMDD	630115
YYYYMMDD	19630115
YYYYDDD	前四位數代表年，後三位數代表一年中的某一天 - 例如，2000032 代表 2000 年的第 32 天，即 2000 年 2 月 1 日。
DAY	目前語言環境下的星期幾， 例如，英文環境下的 Monday 和 Tuesday 等等。
MONTH	目前語言環境下的月 - 例如 January 和 February 等等。
DD/MM/YY	15/01/63
DD/MM/YYYY	15/01/1963
MM/DD/YY	01/15/63
MM/DD/YYYY	01/15/1963
DD-MM-YY	15-01-63
DD-MM-YYYY	15-01-1963
MM-DD-YY	01-15-63
MM-DD-YYYY	01-15-1963
DD.MM.YY	15.01.63
DD.MM.YYYY	15.01.1963
MM.DD.YY	01.15.63
MM.DD.YYYY	01.15.1963
DD-MON-YY	15-JAN-63、15-jan-63 和 15-Jan-63
DD/MON/YY	15/JAN/63、15/jan/63 和 15/Jan/63
DD.MON.YY	15.JAN.63、15.jan.63 和 15.Jan.63
DD-MON-YYYY	15-JAN-1963、15-jan-1963 和 15-Jan-1963
DD/MON/YYYY	15/JAN/1963、15/jan/1963 和 15/Jan/1963
DD.MON.YYYY	15.JAN.1963、15.jan.1963 和 15.Jan.1963
MON YYYY	Jan 2004
q Q YYYY	日期依次由代表季度的數字 (1-4)、字母 Q 以及四位數的年份代表 - 例如，2004 年 12 月 25 日代表為 4 Q 2004。
ww WK YYYY	前兩位數字代表一年中的某周，後跟字母 WK 以及代表年的四位數字。計算一年中的某周時，假設這周的第一天是星期一且第一周至少有一天。

## 時間(M)

CLEM 語言支援下列時間格式。

表 15. CLEM 語言時間格式

格式	範例
HHMMSS	120112、010101 和 221212
HHMM	1223、0745 和 2207
MMSS	5558 和 0100
HH:MM:SS	12:01:12、01:01:01 和 22:12:12
HH:MM	12:23、07:45 和 22:07
MM:SS	55:58 和 01:00
(H)H:(M)M:(S)S	12:1:12、1:1:1 和 22:12:12
(H)H:(M)M	12:23、7:45 和 22:7
(M)M:(S)S	55:58 和 1:0
HH.MM.SS	12.01.12、01.01.01 和 22.12.12
HH.MM	12.23、07.45 和 22.07
MM.SS	55.58 和 01.00
(H)H.(M)M.(S)S	12.1.12、1.1.1 和 22.12.12
(H)H.(M)M	12.23、7.45 和 22.7
(M)M.(S)S	55.58 和 1.0

## CLEM 運算子

CLEM 語言提供了下列運算子。

表 16. CLEM 語言運算子.

運算	備註	優先級 (參閱下一節)
or	在兩個 CLEM 表示式之間使用。如果其中一個表示式為 true 或兩個表示式同時為 true，則傳回 true 值。	10
and	在兩個 CLEM 表示式之間使用。如果兩個表示式同時為 true，則傳回 true 值。	9
=	用於任兩個可比較項目之間。若 ITEM1 等於 ITEM2，則傳回 true 值。	7
==	與 = 相同。	7
/=	用於任兩個可比較項目之間。若 ITEM1 不等於 ITEM2，則傳回 true 值。	7
/==	與 /= 相同。	7
>	用於任兩個可比較項目之間。若 ITEM1 完全大於 ITEM2，則傳回 true 值。	6
>=	用於任兩個可比較項目之間。若 ITEM1 大於或等於 ITEM2，則傳回 true 值。	6
<	用於任兩個可比較項目之間。如果 ITEM1 嚴格少於 ITEM2，則傳回 true 值。	6

表 16. CLEM 語言運算子 (繼續).

運算	備註	優先級 (參閱下一節)
<=	用於任兩個可比較項目之間。若 ITEM1 小於或等於 ITEM2，則傳回值為 true。	6
&&=_0	用於兩個整數之間。相當於布林表示式 INT1 && INT2 = 0。	6
&&/=_0	用於兩個整數之間。相當於布林表示式 INT1 && INT2 /= 0。	6
+	將兩個數字相加：NUM1 + NUM2。	5
><	連接兩個字串，例如 字串1 ><字串2.	5
-	將兩個數字相減：NUM1 - NUM2。也可以用在一個數字之前：- NUM。	5
*	用於將兩個數字相乘：NUM1 * NUM2。	4
&&	用於兩個整數之間。結果為 INT1 和 INT2 的位元 "and"。	4
&&~	用於兩個整數之間。結果為 INT1 的位元 "and" 和 INT2 的位元補碼。	4
	用於兩個整數之間。結果為 INT1 和 INT2 的位元 "inclusive or"。	4
~	用於整數之前。生成整數的位元補碼。	4
/&	用於兩個整數之間。結果為 INT1 和 INT2 的位元 "exclusive or"。	4
INT1 << N	用於兩個整數之間。產生整數向左偏移 N 個位置的位元型樣。	4
INT1 >> N	用於兩個整數之間。產生 INT 向左偏移 N 個位置的位元型樣。	4
/	用於將兩個數相除：NUM1 / NUM2。	4
**	用於兩個數字之間：BASE ** POWER。傳回 BASE 的 POWER 次幂。	3
rem	用於兩個整數之間：INT1 rem INT2。傳回餘數，INT1 - (INT1 div INT2) * INT2。	2
div	用於兩個整數之間：INT1 div INT2。執行整數除。	2

### 運算子優先級

優先級決定複式表示式 (尤其是含有不止一個中置運算子且未加括號的表示式) 的剖析順序。例如，

$$3 + 4 * 5$$

剖析為  $3 + (4 * 5)$  而不是  $(3 + 4) * 5$ ，因為相對優先級規定 \* 應先於 + 剖析。CLEM 語言中的每個運算子均有一個與其關聯的優先級值。此值越大，運算子在剖析清單中的重要性越高，即它將先於優先級值比它大的其他運算子處理。

## 函數參照

下列 CLEM 函數可用於在 IBM SPSS Modeler 中處理資料。您可以在各種對話框（例如「衍生」節點和「設為旗標」節點）中輸入這些函數作為代碼，也可以使用表示式建置器來建立有效的 CLEM 表示式，而後一種方式無需記憶函數清單或欄位名稱。

表 17. 用來搭配 IBM SPSS Modeler 資料使用的 CLEM 函數

函數類型	說明
資訊	用於深入瞭解欄位值。例如，函數 <code>is_string</code> 針對類型為字串的所有記錄傳回 <code>true</code> 值。
轉換	用於構建新欄位或轉換儲存類型。例如，函數 <code>to_timestamp</code> 會將選定欄位轉換為時間戳記。
比較	用於欄位值的相互比較或與指定字串進行比較。例如， <code>&lt;=</code> 用來比較兩個欄位的值是否有一個較小或者兩者是否相等。
邏輯	用來進行邏輯作業，例如， <code>if</code> 、 <code>then</code> 、 <code>else</code> 作業。
數值型(N)	用來進行數值型計算，例如對欄位值取自然對數。
三角的	用來進行三角計算，例如指定角度的反餘弦。
機率	傳回基於各種分佈的機率，例如學生氏 <i>t</i> 分佈中某個值少於特定值的機率。
空間	用於對地理資料執行空間計算。
位元	用於以位元型樣操作整數。
隨機	用於隨機選取項目或產生數字。
字串	用於對字串執行各種作業，例如 <code>stripchar</code> 用來移除指定的字元。
SoundEx	用於在不知道字串準確拼字的情況下根據某些字母的假設發音尋找字串。
日期和時間	用於對日期、時間和時間戳記欄位執行各種作業。
序列	用於深入瞭解資料集的記錄序列，或者根據該序列執行作業。
廣域量	用於存取由「設定廣域」節點建立的廣域值。例如， <code>@MEAN</code> 用於參照某個欄位在整個資料集中所有值的平均值。
空白和空值	用於存取、旗標或填充使用者指定的空白值或系統遺漏值。例如， <code>@BLANK (FIELD)</code> 用於為出現空白值的記錄添加一個 <code>true</code> 旗標。
特殊欄位	用於標示檢查中的特定欄位。例如，在衍生多個欄位時使用 <code>@FIELD</code> 。

## 函數說明中的慣例

涉及函數中的項目時，下列慣例通用於本手冊。

表 18. 函數中的使用慣例說明。

慣例	說明
<code>BOOL</code>	布林或旗標，例如 <code>true</code> 或假。
<code>NUM</code> 、 <code>NUM1</code> 和 <code>NUM2</code>	任意數字。
<code>REAL</code> 、 <code>REAL1</code> 和 <code>REAL2</code>	任意實數，例如 1.234 或 -77.01。
<code>INT</code> 、 <code>INT1</code> 和 <code>INT2</code>	任意整數，例如 1 或 -77。
<code>CHAR</code>	字元碼，例如 <code>`A`</code> 。
字串	字串，例如 <code>"referrerID"</code> 。
<code>LIST</code>	項目清單，例如 <code>["abc" "def"]</code> 。
<code>ITEM</code>	欄位，例如 <code>Customer</code> 或 <code>extract_concept</code> 。



表 18. 函數中的使用慣例說明 (繼續).

慣例	說明
DATE	其值具有 DD-MON-YYYY 這樣格式的日期欄位，例如 start_date。
TIME	其值具有 HHMMSS 這樣格式的時間欄位，例如 power_flux。

本手冊中的函數以函數、結果類型（整數、字串等）和說明（如果有）各佔一欄的形式列出。例如，函數 rem 的說明如下。

表 19. rem 函數說明.

函數	結果	說明
INT1 rem INT2	數字	傳回 INT1 除以 INT2 的餘數。例如，INT1 - (INT1 div INT2) * INT2。

有關使用慣例的詳情（例如，如何在函數中列出項目或指定字元）將在其他文檔中介紹。請參閱第 109 頁的『CLEM 資料類型』主題，以取得更多資訊。

## 資訊函數

資訊函數用於深入瞭解特定欄位的值。它們通常用於衍生旗標欄位。例如，可以使用 @BLANK 函數來建立一個旗標欄位，以指示選定欄位的值為空白值的記錄。同樣，可以使用儲存類型函數（如 is\_string）來檢查某個欄位的儲存類型。

表 20. CLEM 資訊函數.

函數	結果	說明
@BLANK(FIELD)	布林	根據上游類型節點或來源節點（「類型」標籤）中所設定的空白值處理規則，對所有空白值的記錄傳回 true 值。
@NULL(ITEM)	布林	為未定義值的所有記錄傳回 true 值。未定義值是系統空值，在 IBM SPSS Modeler 中顯示為 \$null\$。
is_date(ITEM)	布林	為所有日期類型的記錄傳回 true 值。
is_datetime(ITEM)	布林	為所有日期、時間或時間戳記類型的記錄傳回 true 值。
is_integer(ITEM)	布林	為所有整數類型的記錄傳回 true 值。
is_number(ITEM)	布林	為所有數字類型的記錄傳回 true 值。
is_real(ITEM)	布林	為所有實數類型的記錄傳回 true 值。
is_string(ITEM)	布林	為所有字串類型的記錄傳回 true 值。
is_time(ITEM)	布林	為所有時間類型的記錄傳回 true 值。
is_timestamp(ITEM)	布林	為所有時間戳記類型的記錄傳回 true 值。

## 轉換函數

轉換函數可用來構建新欄位和轉換現有檔案的儲存類型。例如，可通過將字串結合在一起或分拆字串來形成新字串。要結合兩個字串，請使用運算子 >>。例如，如果欄位 Site 的值為 "BRAMLEY"，則 "xx" >> Site 將傳回 "xxBRAMLEY"。即使引數不是字串，>> 的結果也始終是字串。因此，如果欄位 V1 為 3，欄位 V2 為 5，則 V1 >> V2 將傳回 "35"（字串而非數字）。

轉換函數（及其他要求指定類型輸入（如日期或時間值）的函數）取決於「串流選項」對話框中指定的現行格式。例如，如果想轉換值為 Jan 2003、Feb 2003 等的字串欄位，請選取相符的日期格式 **MON YYYY** 作為串流的預設日期格式。

表 21. CLEM 轉換函數

函數	結果	說明
ITEM1 >< ITEM2	字串	將兩個欄位的值連結，並傳回結果字串 ITEM1ITEM2。
to_integer(ITEM)	整數	將指定欄位的儲存類型轉換為整數。
to_real(ITEM)	實數	將指定欄位的儲存類型轉換為實數。
to_number(ITEM)	數字	將指定欄位的儲存類型轉換為數字。
to_string(ITEM)	字串	將指定欄位的儲存類型轉換為字串。使用此函數將實數轉換為字串時，將傳回在基點後存在 6 位數的值。
to_time(ITEM)	時間(M)	將指定欄位的儲存轉換為時間。
to_date(ITEM)	日期	將指定欄位的儲存轉換為日期。
to_timestamp(ITEM)	時間戳記	將指定欄位的儲存轉換為時間戳記。
to_datetime(ITEM)	日期時間	將指定欄位的儲存轉換為時間、時間或時間戳記值。
datetime_date(ITEM)	日期	傳回數字、字串或時間戳記的日期值。請注意，這是唯一能夠讓您將數字（秒）轉換回日期的函數。如果 ITEM 是字串，請透過剖析現行日期格式的字串來建立日期。在串流內容對話框中指定的日期格式必須正確無誤，才能順利完成此函數。如果 ITEM 為數字，將解譯為自基準日期（或新紀元）以來的秒數。日期的分數會被截斷。如果 ITEM 為時間戳記，將傳回時間戳記的日期部分。如果 ITEM 為日期，則它會依原樣傳回。
stb_centroid_latitude(ITEM)	整數	傳回與 geohash 引數的質心相對應的緯度的整數值。
stb_centroid_longitude(ITEM)	整數	傳回與 geohash 引數的質心相對應的經度的整數值。
to_geohash(ITEM)	字串	傳回對應於緯度和經度並使用密度的指定位元數的 geohash 字串。geohash 是一個代碼，用於根據緯度和經度詳細資料識別一組地理座標。to_geohash 的三個參數為： <ul style="list-style-type: none"> <li>• 緯度：範圍 (-180, 180)，並且在 WGS84 座標系統中單位為度。</li> <li>• 經度：範圍 (-90, 90)，並且在 WGS84 座標系中單位為度。</li> <li>• 位元：要用於儲存雜湊的位元數。範圍為 [1,75]。這將同時影響所傳回的字串的長度（對每 5 個位元使用 1 個字元）以及雜湊的精確度。例如，5 個位元（1 個字元）代表近似 2500 公里或 45 個位元（9 個字元）代表近似 2.3 公尺。</li> </ul>

## 比較函數

比較函數用於欄位值的相互比較或與指定字串進行比較。例如，可以使用 = 來檢查字串是否相等。字串相等驗證範例：Class = "class 1"。

對數值型比較來說，大於 表示離正無窮更近，小於 表示離負無窮更近。即，所有負數均少於任意正數。

表 22. CLEM 比較函數

函數	結果	說明
count_equal(ITEM1, LIST)	整數	傳回欄位清單中等於 ITEM1 的值的個數，如果 ITEM1 為無效，則傳回無效。
count_greater_than(ITEM1, LIST)	整數	傳回欄位清單中大於 ITEM1 的值的個數，如果 ITEM1 為無效，則傳回無效。
count_less_than(ITEM1, LIST)	整數	傳回欄位清單中少於 ITEM1 的值的個數，如果 ITEM1 為無效，則傳回無效。

表 22. CLEM 比較函數 (繼續)

函數	結果	說明
count_not_equal(ITEM1, LIST)	整數	傳回欄位清單中不等於 ITEM1 的值的個數，如果 ITEM1 為無效，則傳回無效。
count_nulls(LIST)	整數	傳回欄位清單中空值的個數。
count_non_nulls(LIST)	整數	傳回欄位清單中非空值的個數。
date_before(DATE1, DATE2)	布林	用於檢查日期值的排序。DATE1 早於 DATE2，則傳回 true 值。
first_index(ITEM, LIST)	整數	傳回欄位清單中包含 ITEM 的第一個欄位的指標，如果找不到該值，則傳回 0。只有字串、整數與實數類型才支援。
first_non_null(LIST)	任意	傳回所提供欄位清單中的第一個非空值。支援所有儲存體類型。
first_non_null_index(LIST)	整數	傳回欄位清單中包含非空值的第一個欄位的指標，如果所有值都為空值，則傳回 0。支援所有儲存體類型。
ITEM1 =ITEM2	布林	為 ITEM1 等於 ITEM2 的記錄傳回 true 值。
ITEM1 /=ITEM2	布林	兩個字串不完全相同時傳回 true 值；完全相同時傳回 0。
ITEM1 <ITEM2	布林	為 ITEM1 少於 ITEM2 的記錄傳回 true 值。
ITEM1 <=ITEM2	布林	為 ITEM1 少於等於 ITEM2 的記錄傳回 true 值。
ITEM1 >ITEM2	布林	為 ITEM1 大於 ITEM2 的記錄傳回 true 值。
ITEM1 >=ITEM2	布林	為 ITEM1 大於等於 ITEM2 的記錄傳回 true 值。
last_index(ITEM, LIST)	整數	傳回欄位清單中包含 ITEM 的最後一個欄位的指標，如果找不到該值，則傳回 0。只有字串、整數與實數類型才支援。
last_non_null(LIST)	任意	傳回所提供欄位清單中的最後一個非空值。支援所有儲存體類型。
last_non_null_index(LIST)	整數	傳回欄位清單中包含非空值的最後一個欄位的指標，如果所有值都為空值，則傳回 0。支援所有儲存體類型。
max(ITEM1, ITEM2)	任意	傳回 ITEM1 和 ITEM2 兩ITEM中較大的一個。
max_index(LIST)	整數	傳回數值型欄位清單中包含最大值的欄位指標，如果所有值為空值，則傳回 0。例如，如果所列的第三個欄位包含最大值，將傳回指標值 3。如果有多個欄位包含最大值，則將傳回第一個列出的欄位（位於最左側）。
max_n(LIST)	數字	傳回數值型欄位清單中的最大值，如果所有欄位值為無效，則傳回無效。
member(ITEM, LIST)	布林	如果 ITEM 是指定清單的會員，則傳回 true 值。否則傳回 false。另外，還可以指定欄位名稱清單。
min(ITEM1, ITEM2)	任意	傳回 ITEM1 和 ITEM2 兩ITEM中較小的一個。
min_index(LIST)	整數	傳回數值型欄位清單中包含最小值的欄位指標，如果所有值為空值，則傳回 0。例如，如果所列的第三個欄位包含最小值，將傳回指標值 3。如果有多個欄位包含最小值，則將傳回第一個列出的欄位（位於最左側）。
min_n(LIST)	數字	傳回數值型欄位清單中的最小值，如果所有欄位值為空值，則傳回無效。
time_before(TIME1, TIME2)	布林	用於檢查時間值的排序。TIME1 早於 TIME2，則傳回 true 值。
value_at(INT, LIST)		在偏移 INT 處傳回每個列出欄位值；如果偏移超過有效值範圍（即小於 1 或大於所列欄位個數），則傳回空值。支援所有儲存體類型。

## 邏輯函數

CLEM 表示式可用來執行邏輯作業。

表 23. CLEM 邏輯函數

函數	結果	說明
條件1 and 條件2	布林	此作業為邏輯鏈結作業，在 條件1 和 條件2 同時為 true 時傳回 true 值。如果條件1 為假，則不求條件2 的值；此時可以構造這樣的鏈結，讓條件1 首先測試條件2 中作業是合法的。例如，length(Label) >=6 and Label(6) = 'x'。
條件1 or 條件2	布林	此作業為邏輯（包含）聯集運算，在 COND1 或 COND2 為 true 或者這兩者同時為 true 時傳回 true 值。如果條件1 為 true，則不求條件2 的值。
not(條件)	布林	此作業為邏輯否作業，如果條件 為假，則傳回 true。否則，此作業將傳回值 0。
if 條件 then 表示式1 else 表示式2 endif	任意	此作業為條件求值。如果條件 為 true，此作業將傳回表示式1 的結果。否則，將傳回表示式2 的求值結果。
if 條件1 then 表示式1 elseif 條件2 then 表示式2 else 表示式_N endif	任意	此作業為多分支條件求值。如果 條件1 為 true，那麼此作業將傳回 表示式1 的結果。否則，如果條件2 為 true，此作業將傳回表示式2 的求值結果。否則，將傳回表示式_N 的求值結果。

## 數值函數

CLEM 包含許多一般數值型函數。

表 24. CLEM 數值函數

函數	結果	說明
-NUM	數字	用來使 NUM 無效。傳回具有相反符號的對應數字。
NUM1 + NUM2	數字	傳回 NUM1 與 NUM2 相加所得的和。
NUM1 - NUM2	數字	傳回 NUM2 減去 NUM1 所得的差。
NUM1 * NUM2	數字	傳回 NUM1 乘以 NUM2 所得的值。
NUM1 / NUM2	數字	傳回 NUM1 除以 NUM2 所得的商。
INT1 div INT2	數字	用於進行整數除法。傳回 INT1 除以 INT2 所得的商。
INT1 rem INT2	數字	傳回 INT1 除以 INT2 的餘數。例如，INT1 - (INT1 div INT2) * INT2。
INT1 mod INT2	數字	此函數已不推薦使用。請改為使用 rem 函數。
BASE ** POWER	數字	傳回 BASE 的 POWER 次冪，底和冪次可以是任意數字（一種情況除外：如果 POWER 為除整數 0 之外其他類型的 0，則 BASE 不能為 0）。如果 POWER 為整數，則可通過將 BASE 連續相乘 POWER 次計算出結果。因此，如果 BASE 為整數，則結果為整數。如果 POWER 為整數 0，則結果將始終是與 BASE 類型相同的 1。如果 POWER 不是整數，則用公式 $\exp(\text{POWER} * \log(\text{BASE}))$ 計算結果。
abs(NUM)	數字	傳回 NUM 的絕對值（始終為同一類型的數字）。
exp(NUM)	實數	傳回 e 的 NUM 次冪，其中 e 是自然對數的底數。
fracof(NUM)	實數	傳回 NUM 的小數部分，定義為 NUM-intof(NUM)。
intof(NUM)	整數	將其參數截為整數。傳回與 NUM 符號相同的整數，取滿足 $\text{abs}(\text{INT}) \leq \text{abs}(\text{NUM})$ 的最大值。

表 24. CLEM 數值函數 (繼續).

函數	結果	說明
log(NUM)	實數	傳回 NUM 的自然對數 (以 $e$ 為底), NUM 不能為零 (無論何種類型)。
log10(NUM)	實數	傳回 NUM 的對數 (以 10 為底), 不能為零。此函數定義為 $\log(\text{NUM}) / \log(10)$ 。
negate(NUM)	數字	用來使 NUM 無效。傳回具有相反符號的對應數字。
round(NUM)	整數	用於將 NUM 四捨五入取整數, 如果 NUM 為正數, 則採用公式 $\text{intof}(\text{NUM}+0.5)$ ; 如果 NUM 為負數, 則採用公式 $\text{intof}(\text{NUM}-0.5)$ 。
sign(NUM)	數字	用於確定 NUM 的符號。如果 NUM 為整數, 此作業將傳回 -1、0 或 1。如果 NUM 為實數, 將傳回 -1.0、0.0 或 1.0, 這取決於 NUM 是負數、零還是正數。
sqrt(NUM)	實數	傳回 NUM 的平方根。NUM 必須是正數。
sum_n(LIST)	數字	傳回數值型欄位清單中欄位值的和, 如果所有欄位均為無效, 則傳回空值。
mean_n(LIST)	數字	傳回數值型欄位清單中欄位值的平均數, 如果所有欄位值均為空值, 則傳回空值。
sdev_n(LIST)	數字	傳回數值型欄位清單中欄位值的標準差, 如果所有欄位值均為空值, 則傳回空值。

### 三角函數

本部分介紹的函數以角度為參數或傳回結果為角度。無論哪種情況, 角度單位 (弧度或度數) 均由相關串流選項的設定控制。

表 25. CLEM 三角函數

函數	結果	說明
arccos(NUM)	實數	計算指定角度的反餘弦值。
arccosh(NUM)	實數	計算指定角度的反雙曲餘弦值。
arcsin(NUM)	實數	計算指定角度的反正弦值。
arcsinh(NUM)	實數	計算指定角度的反雙曲正弦值。
arctan(NUM)	實數	計算指定角度的反正切值。
arctan2(NUM_Y, NUM_X)	實數	計算 $\text{NUM\_Y} / \text{NUM\_X}$ 的反正切值, 然後使用兩個數的符號推衍生象限資訊。結果是 $-\pi < \text{ANGLE} \leq \pi$ (弧度) - $180 < \text{ANGLE} \leq 180$ (度數) 範圍內的實數。
arctanh(NUM)	實數	計算指定角度的反雙曲線正切值。
cos(NUM)	實數	計算指定角度的餘弦值。
cosh(NUM)	實數	計算指定角度的雙曲餘弦值。
pi	實數	此常數是最接近 pi 的實數。
sin(NUM)	實數	計算指定角度的正弦值。
sinh(NUM)	實數	計算指定角度的雙曲正弦值。
tan(NUM)	實數	計算指定角度的正切值。
tanh(NUM)	實數	計算指定角度的雙曲線正切值。

## 機率函數

機率函數傳回基於各種分佈的機率，例如，學生氏  $t$  分佈中某個值將會少於特定值的機率。

表 26. CLEM 機率函數

函數	結果	說明
<code>cdf_chisq(NUM, DF)</code>	實數	傳回具有指定自由度的卡方測試中某個值將會少於指定數字的機率。
<code>cdf_f(NUM, DF1, DF2)</code>	實數	傳回自由度為 $DF1$ 和 $DF2$ 的 $F$ 分佈中某個值將會少於指定數字的機率。
<code>cdf_normal(NUM, 平均值, STDDEV)</code>	實數	傳回具有指定平均數和標準差的常態分佈中某個值將會少於指定數字的機率。
<code>cdf_t(NUM, DF)</code>	實數	傳回具有指定自由度的學生氏 $t$ 分佈中某個值將會少於指定數字的機率。

## 空間函數

空間函數可以與地理空間資料配合使用。例如，它們使您能夠計算兩個點之間的距離以及多邊形的面積等資訊。並且，還存在要求對多個基於空間預測（處於和接近於，等等）的地理空間資料集進行合併的狀況，此任務可以通過合併條件完成。

這些空間函數與工具 > 串流內容 > 選項 > 地理空間中指定的座標系統配合使用。

註：這些空間函數不適用於三維度資料。如果將三維度資料匯入到串流中，那麼這些函數將只使用前兩個維度，而忽略  $Z$  軸的值。

表 27. CLEM 空間函數

函數	結果	說明
<code>close_to(SHAPE, SHAPE, NUM)</code>	布林	測試兩個形狀之間的相互距離是否不超過特定的 $DISTANCE$ 。如果使用了投影座標系統，那麼 $DISTANCE$ 以公尺為單位。如果未使用任何座標系統，則其為任意單位。
<code>crosses(SHAPE, SHAPE)</code>	布林	測試兩個形狀是否相交。此函數適用於兩個線串形狀或者一個線串和一個多邊形。
<code>overlap(SHAPE, SHAPE)</code>	布林	測試兩個多邊形之間是否存在交集並且該交集是否在這兩個形狀之內。
<code>within(SHAPE, SHAPE)</code>	布林	測試是否整個 $SHAPE1$ 包含在 $POLYGON$ 之內。
<code>area(SHAPE)</code>	實數	傳回所指定 $POLYGON$ 的面積。如果使用了投影系統，那麼此函數將傳回以平方公尺為單位的結果。如果未使用任何座標系統，則其為任意單位。形狀必須是 $POLYGON$ 或 $MULTIPOLYGON$ 。
<code>num_points(SHAPE, LIST)</code>	整數	傳回點欄位 ( $MULTIPOINT$ ) 中包含在 $POLYGON$ 的範圍內的點數。 $SHAPE1$ 必須是 $POLYGON$ 或 $MULTIPOLYGON$ 。
<code>distance(SHAPE, SHAPE)</code>	實數	傳回 $SHAPE1$ 與 $SHAPE2$ 之間的距離。如果使用了投影座標系統，那麼此函數將傳回以公尺為單位的結果。如果未使用任何座標系統，則其為任意單位。 $SHAPE1$ 和 $SHAPE2$ 可以是任何地理測量類型。

## 位元整數作業

這些函數可以按代表二進制補碼值的位元型樣（其中位元位置  $N$  的加權為  $2^{*N}$ ）來操作整數。位元從 0 開始往上編號。這些作業就好像是把整數的符號位元向左無限延伸。因此，在最高有效位元上的所有地方，正整數有 0 位元，負整數有 1 位元。

表 28. CLEM 位元整數作業。

函數	結果	說明
<code>~ INT1</code>	整數	產生整數 $INT1$ 的位元補碼。即， $INT1$ 的對應位元為 0 的結果均為 1。 $\sim\sim INT = -(INT + 1)$ 始終成立。
<code>INT1    INT2</code>	整數	此作業的結果是 $INT1$ 與 $INT2$ 的位元 "inclusive or"。即， $INT1$ 和 $INT2$ 其中一個的對應位元上是 1 或兩者的對應位元上同時是 1 時「或」後的結果為 1。
<code>INT1   /&amp; INT2</code>	整數	此作業的結果是 $INT1$ 與 $INT2$ 的位元 "exclusive or"。即， $INT1$ 和 $INT2$ 其中一個的對應位元是 1 但非同時為 1 時「異或」後的結果為 1。
<code>INT1 &amp;&amp; INT2</code>	整數	產生整數 $INT1$ 和 $INT2$ 的位元 "and"。即，結果中的每個位元位置僅當 $INT1$ 和 $INT2$ 中的對應位元位置均為 1 時才為 1。
<code>INT1 &amp;&amp;~ INT2</code>	整數	將對 $INT1$ 和 $INT2$ 位元求補後的結果進行位元 "and" 運算。即， $INT1$ 的對應位元是 1 且 $INT2$ 的對應位元是 0 時進行這一運算後的結果才為 1。此運算與 $INT1 \&\& (\sim\sim INT2)$ 運算的結果相同，可用於清除在 $INT2$ 中設定的 $INT1$ 的位元。
<code>INT &lt;&lt; N</code>	整數	產生 $INT1$ 向左偏移 $N$ 個位置的位元型樣。如果 $N$ 為負值，那麼向右移。
<code>INT &gt;&gt; N</code>	整數	產生 $INT1$ 向右偏移 $N$ 個位置的位元型樣。如果 $N$ 為負值，那麼向左移。
<code>INT1 &amp;&amp;=_0 INT2</code>	布林	相當於布林表示式 $INT1 \&\& INT2 \neq 0$ ，但效率更高。
<code>INT1 &amp;&amp;/=_0 INT2</code>	布林	相當於布林表示式 $INT1 \&\& INT2 == 0$ ，但效率更高。
<code>integer_bitcount(INT)</code>	整數	計算 $INT$ 的二補數代表中 1 位元或 0 位元的號碼。如果 $INT$ 為非負數，那麼 $N$ 為 1 位元的號碼。如果 $INT$ 為負數，那麼 $N$ 為 0 位元數目。由於符號延伸的原因，非負整數中將有無數個 0 位元，而負整數中將有無數個 1 位元。對於此運算，始終有 $integer\_bitcount(INT) = integer\_bitcount(-(\sim\sim INT))$ 成立。
<code>integer_leastbit(INT)</code>	整數	傳回整數 $INT$ 中最低有效位元組的位元位置 $N$ 。 $N$ 是保證可以對 $INT$ 進行整除的條件下，2 的最高冪次。
<code>integer_length(INT)</code>	整數	以二補整數形式傳回 $INT$ 的長度（以位元計）。即， $N$ 是使 $INT < (1 \ll N)$ if $INT \geq 0$ $INT \geq (-1 \ll N)$ if $INT < 0$ 成立的最小整數。如果 $INT$ 非負，那麼將 $INT$ 代表為無符號整數需要一個長度至少為 $N$ 位元的欄位。另外，最少需要用 $N+1$ 個數位來代表有符號整數 $INT$ （不論符號是正是負）。
<code>testbit(INT, N)</code>	布林	測試整數 $INT$ 中位置 $N$ 處的位元，並以布林形式傳回第 $N$ 位元的狀態（true 表示 1，false 表示 0）。

## 隨機函數

下列函數用於隨機選取項目或隨機產生數字。

表 29. CLEM 隨機函數.

函數	結果	說明
oneof(LIST)	任何	傳回一個從 LIST 中隨機選取的元素。應以 [ITEM1,ITEM2,...,ITEM_N] 的形式已輸入清單ITEM。注意，還可以指定欄位名稱清單。
random(NUM)	數字	傳回類型相同 (INT 或 REAL) 的均勻分配的隨機數字，其範圍從 1 至 NUM。如果使用的是整數，則僅傳回整數。如果使用 (十進制) 實數，則傳回實數 (由串流選項確定的十進制精準度)。此函數所傳回的最大亂數可以等於 NUM。
random0(NUM)	數字	此函數的內容與 random(NUM) 相同，但將從 0 開始取值。它所傳回的最大亂數決不會等於 NUM。

## 字串函數

在 CLEM 中您可以對字串執行下列作業：

- 比較字串
- 建立字串
- 存取字元

在 CLEM 中，是一對雙引號之間的任何字元序列 ("string quotes")。這裡所用的字元 (字元) 可以是任何一個英數字元。這些字元已在 CLEM 表示式中使用反單引號並以 `` 的形式宣告過，例如 `z`、`A` 或 `2`。如果字元超出範圍或為字串中的負指數，則會導致未定義的行為。

附註。在使用與不使用 SQL 回送的字串之間的對比可能產生存在尾部空格的不同結果。

表 30. CLEM 字串函數.

函數	結果	說明
allbutfirst(N,字串)	字串	傳回一個字串，其為移除了前 N 個字元的 STRING。
allbutlast(N,字串)	字串	傳回一個字串，其為移除了最後幾個字元的 STRING。
alphabefore(字串1,字串2)	布林	用於檢查字串的按字母順序的排序。如果字串1 在字串2 之前，則傳回 true 值。
endstring(LENGTH,字串)	字串	從指定字串中擷取最後 N 個字元。如果字串長度小於或等於指定的長度，則字串保留不變。
hasendstring(字串, SUBSTRING)	整數	此函數的功能等同於 isendstring(SUBSTRING,字串)。
hasmidstring(字串, SUBSTRING)	整數	此函數的功能等同於 ismidstring(SUBSTRING,字串) (內含的子字串)。
hasstartstring(字串, SUBSTRING)	整數	此函數的功能等同於 isstartstring(SUBSTRING,字串)。
hassubstring(字串, N, SUBSTRING)	整數	此函數的功能等同於 issubstring(SUBSTRING, N,字串)，其中 N 的預設值為 1。



表 30. CLEM 字串函數 (繼續).

函數	結果	說明
count_substring(字串, SUBSTRING)	整數	傳回字串中指定子字串出現次數。例如，count_substring("foooo.txt", "oo") 傳回結果為 3。
hassubstring(字串, SUBSTRING)	整數	此函數的功能等同於 issubstring(SUBSTRING, 1, 字串)，其中 N 的預設值為 1。
isalphacode(字元)	布林	如果字元是字元碼為字母的指定字串（常為欄位名稱）中的某個字元，則傳回 true 值。否則，此函數的傳回值為 0，例如 isalphacode(produce_num(1))。
isendstring(SUBSTRING, 字串)	整數	如果字串 字串 以子字串 SUBSTRING 結尾，則該函數將傳回字串 字串 中子字串 SUBSTRING 的整數下標。否則，此函數會傳回值 0。
islowercode(字元)	布林	如果字元是指定字串（通常為欄位名稱）中的某個小寫字母字元，則傳回 true 值。否則，該函數的傳回值為 0，例如 islowercode(``) 和 islowercode(country_name(2)) 均為有效表示式。
ismidstring(SUBSTRING, 字串)	整數	如果 SUBSTRING 是 字串 的子字串，但首字元或尾端字元與 字串 不同，則該函數將傳回此子字串的起始下標。否則，此函數會傳回值 0。
isnumbercode(字元)	布林	如果指定字串（通常為欄位名稱）的字元為數字字元，則傳回 true 值。否則，此函數的傳回值為 0，例如 isnumbercode(product_id(2))。
isstartstring(SUBSTRING, 字串)	整數	如果字串 字串 以子字串 SUBSTRING 開頭，則該函數將傳回下標 1。否則，此函數的傳回結果為 0。
issubstring(SUBSTRING, N, 字串)	整數	搜尋字串 STRING（從其第 N 個字元啟動）中等於字串 SUBSTRING 的子字串。找到之後，此函數會傳回整數子字串，相符子字串在此處啟動。否則，此函數會傳回值 0。如果未給定 N，則此函數預設為 1。
issubstring(SUBSTRING, 字串)	整數	搜尋字串 STRING（從其第 N 個字元啟動）中等於字串 SUBSTRING 的子字串。找到之後，此函數會傳回整數子字串，相符子字串在此處啟動。否則，此函數會傳回值 0。如果未給定 N，則此函數預設為 1。
issubstring_count(SUBSTRING, N, 字串):	整數	傳回 字串 中第 N 次出現的 SUBSTRING 的指標。如果 SUBSTRING 的出現次數少於 N 次，則傳回結果為 0。
issubstring_lim(SUBSTRING, N, STARTLIM, ENDLIM, 字串)	整數	此函數的功能與 issubstring 相同，但比對必須從下標 STARTLIM 或其之前開始，並且必須在下標 ENDLIM 或其之前結束。通過對任意一個參數提供 false，可以將 STARTLIM 或 ENDLIM 限制停用 - 例如，issubstring_lim(SUBSTRING, N, false, false, 字串) 函數功能等同於 issubstring。
isuppercode(字元)	布林	如果字元是大寫字母字元，則傳回 true 值。否則，此函數的傳回值為 0，例如 isuppercode(``) 和 isuppercode(country_name(2)) 均為有效表示式。

表 30. CLEM 字串函數 (繼續).

函數	結果	說明
last(字元)	字串	傳回 字串 (長度至少為一個字元) 的最後一個字元字元。
length(字串)	整數	傳回字串 字串 的長度，即字串內的字元數。
locchar(字元, N, 字串)	整數	用於識別符號欄位中的字元位置。此函數將在字串 字串 中搜尋字元 字元 (從 字串 的第 N 個字元開始搜尋)。此函數的傳回值表示 (從 N 開始) 找到字元的位置。如果找不到字元，則函數傳回結果為 0。如果此函數存在無效偏移 (N) (例如，某個偏移超出了字串長度)，則函數傳回結果為 \$null\$。例如，locchar(`n`, 2, web_page) 在名為 web_page 的欄位中從第二個字元開始搜尋字元 `n`。附註：請務必使用單左引號來封裝指定的字元。
locchar_back(字元, N, 字串)	整數	與 locchar 函數類似，所不同的是搜尋反向進行並且從第 N 個字元開始。例如，locchar_back(`n`, 9, web_page) 將從 web_page 的第 9 個字元開始搜尋並且是從後向前移動搜尋。如果此函數存在無效偏移 (例如，某個偏移超出了字串長度)，則函數傳回結果為 \$null\$。理想情況下，您將 locchar_back 與 length(<field>) 函數結合使用，以便動態地使用欄位目前的值的長度。例如 locchar_back(`n`, (length(web_page)), web_page)。
lowertoupper(字元) lowertoupper(字串)	字元 或字串	可以輸入字串或字元，此函數將傳回一個類型相同項目，小寫字元將轉換為相應的大寫字元。例如，lowertoupper(`a`)、lowertoupper("My string") 和 lowertoupper(field_name(2)) 均為有效表示式。
matches	布林	如果字串與指定型樣相符，則將傳回 true 值。型樣必須是字串文字，而不能是包含型樣的欄位名稱。問號 (?) 可以併入型樣以僅符合一個字元；(*) 符合零個以上字元。若要符合一個問號或星號文字 (而不是使用它們作為萬用字元)，可將反斜線 (\) 用作跳出字元。
replace(SUBSTRING, NEWSUBSTRING, 字串)	字串	在指定 字串 中，用 NEWSUBSTRING 取代 SUBSTRING 的所有實例。
replicate(COUNT, 字串)	字串	傳回一個由原始字串指定個數的複本所構成的字串。
stripchar(字元, 字串)	字串	借助此函數，您可從某個字串 或欄位中移除指定字元。例如，您可以使用此函數從資料中移除多餘符號 (如貨幣符號) 以獲取一個簡單的數字或名稱。例如，通過語法 stripchar(`\$`, 'Cost') 可以移除所有值中的美元符號並傳回一個新欄位。附註：請務必使用單左引號來封裝指定的字元。

表 30. CLEM 字串函數 (繼續).

函數	結果	說明
skipchar(字元, N, 字串)	整數	在字串 字串 中搜尋除 字元 以外的字元 (從第 <i>N</i> 個字元開始搜尋)。此函數傳回一個表示發現字元位置的整數子字串, 如果從第 <i>N</i> 個位置開始的每個字元是 字元, 則傳回 0。如果此函數存在無效偏移 (例如, 某個偏移超出了字串長度), 則函數傳回結果為 \$null\$。 locchar 函數經常與 skipchar 結合使用以確定 <i>N</i> (啟動搜尋字串的位置) 的值。例如 skipchar(`s`, (locchar(`s`, 1, "MyString")), "MyString")。
skipchar_back(字元, N, 字串)	整數	與 skipchar 函數類似, 所不同的是搜尋 反向 進行並且從第 <i>N</i> 個字元開始。
startstring(LENGTH, 字串)	字串	從指定字串中擷取前 <i>N</i> 個字元。如果字串長度小於或等於指定的長度, 則字串保留不變。
strmember(字元, 字串)	整數	等同於 locchar(字元, 1, 字串) 函數。此函數傳回整數子字串或 0, 傳回的整數子字串表示 字元 第一次出現的位置。如果函數存在無效偏移 (例如, 某個偏移超出了字串長度), 則函數傳回結果為 \$null\$。
subscrs(N, 字串)	CHAR	傳回輸入字串 字串 的第 <i>N</i> 個字元 字元。此函數也可採用簡寫入形式 字串( <i>N</i> )。例如, lowertoupper("name"(1)) 是有效的表示式。
substring(N, LEN, 字串)	字串	傳回從位於下標 <i>N</i> 的字元開始的字串 SUBSTRING, 此子字串由字串 字串 的 <i>LEN</i> 個字元組成。
substring_between(N1, N2, 字串)	字串	傳回 字串 的子字串, 該字串從下標 <i>N1</i> 開始並在下標 <i>N2</i> 結束。
trim(字串)	字串	移除指定字串的前導和尾部空格。
trim_start(字串)	字串	移除指定字串的前導空格。
trimend(字串)	字串	移除指定字串的尾部空格。
unicode_char(NUM)	CHAR	輸入必須是十進制值, 而不能是十六進位值。傳回 Unicode 值為 <i>NUM</i> 的字元。
unicode_value(字元)	NUM	傳回 字元 的 Unicode 值。
uppertolower(字元) uppertolower(字串)	字元 或字串	可以輸入字串或字元, 此函數將傳回一個類型相同項目, 大寫字元將轉換為相應的小寫字元。 註: 指定字串時一定要用英文雙引號, 而指定字元時一定要用反向單引號。簡單欄位名稱不應使用引號。

## SoundEx 函數

SoundEx 是在知道字串的發音卻不知其準確拼字的情況下用來尋找字串的一種方法。此種方法在 1918 年提出, 它根據某些字母如何發音的相關語音假設, 搜尋出類似發音的單字。此種方法可用於在資料庫中搜尋名稱, 例如相似名稱的拼字和發音可能有所不同。儘管存在限制 (例如, 即便前導字母組合 ph 和 f 的發音相同, 但它們仍然不相符), 基本的 SoundEx 演算法在多個來源中均有歸檔, 並且絕大多數資料庫均以某種形式受支援此種演算法。

表 31. CLEM 聲音編碼函數

函數	結果	說明
soundex(字串)	整數	傳回指定字串的四字元 SoundEx 代碼。
soundex_difference(字串1,字串2)	整數	傳回 0 和 4 之間的某個整數，此整數表示兩個字串中 SoundEx 編碼相同的字元的個數，0 表示不相似，4 表示兩個字串極相似或完全相同。

## 日期和時間函數

CLEM 包含一系列函數，用來處理具有字串變數（代表日期與時間）的日期時間儲存的欄位。使用的日期和時間格式是每個串流特定的，並在「串流內容」對話框中指定。日期與時間函數根據目前選定的格式來剖析日期與時間字串。

如果用兩位數指定日期中的年（即未指定世紀），則 IBM SPSS Modeler 將採用「串流內容」對話框中所指定的預設世紀。

註：如果將資料函數後推到 SQL 或 IBM SPSS Analytic Server，那麼在 Analytic Server 資料來源後面的分支中，該資料內的任何日期格式字串 (to\_date) 都必須與 SPSS Modeler 串流中指定的日期格式相符。

表 32. CLEM 日期和時間函數.

函數	結果	說明
@TODAY	字串	如果在串流內容對話框中選中了轉換日/分鐘，那麼此函數將採用現行日期格式的現行日期字串。如果您採用兩位數日期格式且未選取轉換日/分鐘，該函數將在目前伺服器上傳回 \$null\$。
to_time(ITEM)	時間	將指定欄位的儲存轉換為時間。
to_date(ITEM)	日期	將指定欄位的儲存轉換為日期。
to_timestamp(ITEM)	時間戳記	將指定欄位的儲存轉換為時間戳記。
to_datetime(ITEM)	日期時間	將指定欄位的儲存轉換為時間、時間或時間戳記值。
datetime_date(ITEM)	日期	傳回數字、字串或時間戳記的日期值。請注意，這是唯一能夠讓您將數字（秒）轉換回日期的函數。如果 ITEM 是字串，請透過剖析現行日期格式的字串來建立日期。在串流內容對話框中指定的日期格式必須正確無誤，才能順利完成此函數。如果 ITEM 為數字，那麼將解釋為自基準日期（或新紀元）以來的秒數。日期的分數會被截斷。如果 ITEM 為時間戳記，將傳回時間戳記的日期部分。如果 ITEM 為日期，則它會依原樣傳回。
date_before(DATE1, DATE2)	布林	如果 DATE1 所代表的日期或時間戳記在 DATE2 之前，那麼傳回 true 值。否則，此函數會傳回值 0。
date_days_difference(DATE1,DATE2)	整數	以整數的形式傳回從 DATE1 所代表日期或時間戳記到 DATE2 所代表日期或時間戳記之間的天數。如果 DATE2 早於 DATE1，則此函數會傳回一個負數。
date_in_days(DATE)	整數	以整數形式傳回從基準線日期到 DATE 所代表日期或時間戳記之間的天數。如果 DATE 早於基準線日期，則此函數會傳回負數值。必須包含有效日期，才能適當進行計算。例如，您不應該指定 2001 年 2 月 29 日這個日期。因為 2001 不是閏年，所以此日期不存在。

表 32. CLEM 日期和時間函數 (繼續).

函數	結果	說明
date_in_months( <i>DATE</i> )	實數	以實數形式傳回從基準線日期到日期或時間戳記 <i>DATE</i> 的月數。這是基於一個月 30.4375 天的大概圖例。如果 <i>DATE</i> 早於基準線日期，則此函數會傳回負數值。必須包含有效日期，才能適當進行計算。例如，您不應該指定 2001 年 2 月 29 日這個日期。因為 2001 不是閏年，所以此日期不存在。
date_in_weeks( <i>DATE</i> )	實數	以實數形式傳回從基準線日期到日期或時間戳記 <i>DATE</i> 的周數。這基於一週 7.0 天。如果 <i>DATE</i> 早於基準線日期，則此函數會傳回負數值。必須包含有效日期，才能適當進行計算。例如，您不應該指定 2001 年 2 月 29 日這個日期。因為 2001 不是閏年，所以此日期不存在。
date_in_years( <i>DATE</i> )	實數	以實數形式傳回從基準線日期到日期或時間戳記 <i>DATE</i> 的年數。這是基於一年 365.25 天的大概圖例。如果 <i>DATE</i> 早於基準線日期，則此函數會傳回負數值。必須包含有效日期，才能適當進行計算。例如，您不應該指定 2001 年 2 月 29 日這個日期。因為 2001 不是閏年，所以此日期不存在。
date_months_difference( <i>DATE1</i> , <i>DATE2</i> )	實數	以實數形式傳回從日期或時間戳記 <i>DATE1</i> 到日期或時間戳記 <i>DATE2</i> 的月數。這是基於一個月 30.4375 天的大概圖例。如果 <i>DATE2</i> 早於 <i>DATE1</i> ，則此函數會傳回一個負數。
datetime_date( <i>YEAR</i> , <i>MONTH</i> , <i>DAY</i> )	日期	建立給定 <i>YEAR</i> 、 <i>MONTH</i> 、 <i>DAY</i> 的日期值。引數必須是整數。
datetime_day( <i>DATE</i> )	整數	傳回給定 <i>DATE</i> 或時間戳記中的月份日期。結果為範圍 1 到 31 內的一個整數。
datetime_day_name( <i>DAY</i> )	字串	傳回給定 <i>DAY</i> 的全稱。該引數必須是介於 1 (星期日) 到 7 (星期六) 範圍內的整數。
datetime_hour( <i>TIME</i> )	整數	傳回某個 <i>TIME</i> 或時間戳記的小時部分。結果為 0 到 23 範圍內的整數。
datetime_in_seconds( <i>TIME</i> )	實數	傳回 <i>TIME</i> 儲存的秒鐘部分。
datetime_in_seconds( <i>DATE</i> ) 和 datetime_in_seconds( <i>DATETIME</i> )	實數	傳回從目前 <i>DATE</i> 或 <i>DATETIME</i> 到基準線日期 (1900-01-01) 之間的差分的累計秒數。
datetime_minute( <i>TIME</i> )	整數	傳回某個 <i>TIME</i> 或時間戳記的分鐘部分。結果為 0 到 59 範圍內的整數。
datetime_month( <i>DATE</i> )	整數	傳回某個 <i>DATE</i> 或時間戳記的月部分。結果為 1 到 12 範圍內的整數。
datetime_month_name ( <i>MONTH</i> )	字串	傳回給定 <i>MONTH</i> 的全稱。該引數必須是介於 1 到 12 範圍內的整數。
datetime_now	時間戳記	以時間戳記形式傳回目前時間。
datetime_second( <i>TIME</i> )	整數	傳回某個 <i>TIME</i> 或時間戳記的秒鐘部分。結果為 0 到 59 範圍內的整數。
datetime_day_short_name ( <i>DAY</i> )	字串	傳回給定 <i>DAY</i> 的簡稱。該引數必須是介於 1 (星期日) 到 7 (星期六) 範圍內的整數。
datetime_month_short_name ( <i>MONTH</i> )	字串	傳回給定 <i>MONTH</i> 的簡稱。該引數必須是介於 1 到 12 範圍內的整數。
datetime_time( <i>小時</i> , <i>分鐘</i> , <i>SECOND</i> )	時間	傳回給定 <i>HOURL</i> 、 <i>MINUTE</i> 和 <i>SECOND</i> 的時間值。引數必須是整數。
datetime_time( <i>ITEM</i> )	時間	傳回給定 <i>ITEM</i> 的時間值。

表 32. CLEM 日期和時間函數 (繼續).

函數	結果	說明
datetime_timestamp(YEAR, MONTH, DAY,小時,分鐘, SECOND)	時間戳記	傳回指定 YEAR、MONTH、DAY、小時、分鐘 以及 SECOND 的時間戳記值。
datetime_timestamp( DATE, TIME)	時間戳記	傳回給定 DATE 和 TIME 的時間戳記值。
datetime_timestamp (NUMBER)	時間戳記	傳回給定秒數的時間戳記值。
datetime_weekday( DATE)	整數	對給定 DATE 或時間戳記中的星期。
datetime_year( DATE)	整數	傳回某個 DATE 或時間戳記的年份部分。結果為一個整數，如 2002。
date_weeks_difference ( DATE1, DATE2)	實數	以實數形式傳回從日期或時間戳記 DATE1 到日期或時間戳記 DATE2 的周數。這基於一週 7.0 天。如果 DATE2 早於 DATE1，則此函數會傳回一個負數。
date_years_difference ( DATE1, DATE2)	實數	以實數形式傳回從日期或時間戳記 DATE1 到日期或時間戳記 DATE2 的年數。這是基於一年 365.25 天的大概圖例。如果 DATE2 早於 DATE1，則此函數會傳回一個負數。
date_from_ywd(YEAR, WEEK, DAY)	整數	將年份、年中的第幾個星期以及星期幾轉換為使用 ISO 8601 標準的日期。
date_iso_day( DATE)	整數	傳回使用 ISO 8601 標準的日期中的星期幾。
date_iso_week( DATE)	整數	傳回使用 ISO 8601 標準的日期中的年中第幾個星期。
date_iso_year( DATE)	整數	傳回使用 ISO 8601 標準的日期中的年份。
time_before( TIME1, TIME2)	布林	如果 TIME1 所代表的時間或時間戳記在 TIME2 之前，則傳回 true 值。否則，此函數會傳回值 0。
time_hours_difference ( TIME1, TIME2)	實數	以實數形式傳回 TIME1 與 TIME2 所代表的時間或時間戳記之間的小時差。如果在「串流內容」對話框中選取轉換日/分鐘，則用 TIME1 的較大值表示前一天。如果您未選取輪替選項，則 TIME1 的較高值會導致傳回的值為負數。
time_in_hours( TIME)	實數	以實數的形式傳回 TIME 所代表的小時數。例如，採用時間格式 HHMM 時，表示式 time_in_hours('0130') 的返回值為 1.5。TIME 可以代表時間或時間戳記。
time_in_mins( TIME)	實數	以實數的形式傳回 TIME 所代表的分鐘數。TIME 可以代表時間或時間戳記。
time_in_secs( TIME)	整數	以實數的形式傳回 TIME 所代表的秒數。TIME 可以代表時間或時間戳記。
time_mins_difference( TIME1, TIME2)	實數	以實數形式傳回 TIME1 與 TIME2 所代表的時間或時間戳記之間的分鐘差。如果您在串流內容對話框中選取輪替天數/分數，則會採取 TIME1 的較高值來參照前一天（或前一小時，如果僅以現行格式指定了分鐘或秒數的話。）如果不選取翻轉日選項，TIME1 的較大值會造成回覆值為負數。
time_secs_difference( TIME1, TIME2)	整數	傳回 TIME1 和 TIME2 所代表的時間或時間戳記之間的時差秒數（整數）。如果您在串流內容對話框中選取輪替天數/分數，則會採取 TIME1 的較高值來參照前一天（或前一小時，如果僅以現行格式指定了分鐘或秒數的話。）如果您未選取輪替選項，則 TIME1 的較高值會導致傳回的值為負數。

## 轉換日期與時間值

請注意，轉換函數（及其他要求指定類型輸入（如日期或時間值）的函數）取決於「串流選項」對話框中指定的現行形式。例如，如果某個名為 *DATE* 的欄位是以值為 *Jan 2003*、*Feb 2003* 等的字串進行儲存的，則可以將它轉換為日期儲存，如下所示：

```
to_date(DATE)
```

為了使此轉換生效，請選取相符的日期格式 **MON YYYY** 作為串流的預設日期格式。

如需使用填充節點將字串值轉換為日期的範例，請參閱串流 *broadband\_create\_models.str*，它安裝在 `\Demos` 資料夾中的 *streams* 子資料夾下。

**儲存為數字形式的日期。**請注意，前述範例中的 *DATE* 是欄位的名稱，而 *to\_date* 是 CLEM 函數。如果已將日期儲存為數字，則可以使用 *datetime\_date* 函數進行轉換，其中數字解析為自基線日期（或紀元）以來的秒數。

```
datetime_date(DATE)
```

通過將日期轉換為秒數（再反向轉換），您可以進行用現行日期加號減固定天數的計算，例如：

```
datetime_date((date_in_days(DATE)-7)*60*60*24)
```

## 序列函數

對於某些作業而言，事件序列很重要。通過該應用程式，您可以處理下列記錄序列：

- 序列和時間序列
- 序列函數
- 記錄編製索引
- 求值的平均數、總和以及對值進行比較
- 監視變更--區別
- @SINCE
- 偏移值
- 其他序列工具

對於多數應用程式，每個通過串流傳送的記錄都可以視為獨立於所有其它記錄的個別案例。此種狀況下，記錄的順序通常並不重要。

然而對於某些類別的問題，記錄序列非常重要。這些一般是發生在時間序列中的狀況，其中記錄序列代表事件的依序排列。每條記錄代表著時間中的某個特定時刻的 Snapshot；但是，大部分最重要的資訊可能並未以瞬時值形式包含，而是以此類值隨時間不斷變化和運行的方式來包含。

當然，相關參數可能並不是時間。例如，記錄可以代表對沿著直線的距離所執行的分析，但也將套用相同的原則。

通過下列特性可以立即識別出序列和特殊函數：

- 這些函數的字首均為 @。
- 函數名稱採用大寫。

序列函數可以參照節點目前處理的記錄、已通過節點的記錄、甚至是（某種情況下）尚未通過節點的記錄。序列函數可與其他 CLEM 表示式部分自由組合，雖然某些函數對引數有所限制。

## 範例

您會發現瞭解自某一特定事件發生或條件為 true 以來的時間長度非常有用。使用函數 @SINCE 可實現上述目的，例如：

```
@SINCE(Income > Outgoings)
```

該函數傳回最後一欄一條滿足此條件的記錄的偏移量，即在此記錄之前滿足條件的記錄數。如果條件從不為 true，則 @SINCE 傳回 @INDEX + 1。

有時您可能需要在 @SINCE 所用的表示式中參照現行記錄的值。此時可以套用函數 @THIS，它指定了一個始終套用至現行記錄的欄位名稱。若要找出 Concentration 欄位值是現行記錄兩倍多的最後一條記錄的偏移，您可以使用下列表示式：

```
@SINCE(Concentration > 2 * @THIS(Concentration))
```

在有些情況下，現行記錄的 @SINCE 函數的定義條件為 true，例如：

```
@SINCE(ID == @THIS(ID))
```

基於這種原因，@SINCE 函數將不對現行記錄的條件求值。如果要對現行記錄以及前面記錄的條件求值，請使用類似函數 @SINCE0；如果現行記錄的條件為 true，則 @SINCE0 傳回值為 0。

表 33. CLEM 序列函數

函數	結果	說明
平均值(欄位)	實數	傳回指定欄位 欄位 或 <i>FIELDS</i> 的值的平均數。
@MEAN(欄位,表示式)	實數	傳回目前節點所接收的最後一欄 表示式 條記錄（包含現行記錄）中的 欄位 值的平均數。 <i>FIELD</i> 必須是數值欄位的名稱。 <i>EXPR</i> 可能是求值為大於 0 的整數的任何表示式。如果省略 <i>EXPR</i> ，或者它超出到目前為止收到的記錄數，則會傳回到目前為止收到的所有記錄的平均值。
@MEAN(欄位,表示式, INT)	實數	傳回目前節點所接收的最後一欄 表示式 條記錄（包含現行記錄）中的 欄位 值的平均數。 <i>FIELD</i> 必須是數值欄位的名稱。 <i>EXPR</i> 可能是求值為大於 0 的整數的任何表示式。如果省略 <i>EXPR</i> ，或者它超出到目前為止收到的記錄數，則會傳回到目前為止收到的所有記錄的平均值。 <i>INT</i> 指定要回顧的數目上限。這遠比只使用兩個引數有效得多。
@DIFF1(欄位)	實數	傳回 欄位 的一階微分差。單一引數形式只會傳回欄位的現行值與先前值之間的差異。如果相關的先前記錄不存在，則傳回 <b>\$null\$</b> 。
@DIFF1(欄位1,欄位2)	實數	兩參數形式給出了 欄位1 對 欄位2 的一階微分差。如果相關的先前記錄不存在，則傳回 <b>\$null\$</b> 。此值計算為 @DIFF1(欄位1)/@DIFF1(欄位2)。
@DIFF2(欄位)	實數	傳回 欄位 的二階微分差。單一引數形式只會傳回欄位的現行值與先前值之間的差異。如果相關的先前記錄不存在，則傳回 <b>\$null\$</b> 。@DIFF2 計算為 @DIFF(@DIFF(欄位))。
@DIFF2(欄位1,欄位2)	實數	兩參數形式給出了 欄位1 對 欄位2 的二階微分差。如果相關的先前記錄不存在，則傳回 <b>\$null\$</b> 。這是一項複合計算 -- @DIFF1(欄位1)/@DIFF1(欄位2) - @OFFSET(@DIFF1(欄位1),1)/@OFFSET(@DIFF1(欄位2)) / @DIFF1(欄位2)。



表 33. CLEM 序列函數 (繼續)

函數	結果	說明
@INDEX	整數	傳回現行記錄的指標。索引在記錄到達目前節點時配置給這些記錄。分配給第一條記錄的指標為 1，各個後續記錄的指標以 1 遞增。
@LAST_NON_BLANK(FIELD)	任意	傳回 FIELD 的最後一個非空白值，如上游來源或「類型」節點中所定義。如果到目前為止的記錄中的 FIELD 沒有非空白值，則會傳回 \$null\$。請注意，空白值（也稱為使用者遺漏值）可以個別針對每個欄位進行定義。
@MAX(欄位)	數字	傳回指定 欄位 的最大值。
@MAX(欄位, 表示式)	數字	傳回目前為止所接收的最後一欄 表示式 條記錄（包含現行記錄）中的 欄位 欄位的最大值。FIELD 必須是數值欄位的名稱。EXPR 可能為求值為大於 0 之整數的任何表示式。
@MAX(欄位, 表示式, INT)	數字	傳回目前為止所接收的最後一欄 表示式 條記錄（包括現行記錄）中 欄位的最大值。FIELD 必須是數值欄位的名稱。表示式 可以是任何值為大於 0 的整數的表示式。如果 表示式 被忽略或超出了目前已接收的記錄數，則傳回目前為止接收到的所有記錄數的最大值。INT 指定要回顧的數目上限。這遠比只使用兩個引數有效得多。
@MIN(欄位)	數字	傳回指定 欄位 的最小值。
@MIN(欄位, 表示式)	數字	傳回到目前為止收到之最後 EXPR 筆記錄（包括現行記錄）內 FIELD 的最小值。FIELD 必須是數值欄位的名稱。EXPR 可能為求值為大於 0 之整數的任何表示式。
@MIN(欄位, 表示式, INT)	數字	傳回到目前為止收到之最後 EXPR 筆記錄（包括現行記錄）內 FIELD 的最小值。FIELD 必須是數值欄位的名稱。表示式 可以是任何值為大於 0 的整數的表示式。如果 表示式 被忽略或超出了目前已接收的記錄數，則傳回目前為止接收到的所有記錄數的最小值。INT 指定要回顧的數目上限。這遠比只使用兩個引數有效得多。

表 33. CLEM 序列函數 (繼續)

函數	結果	說明
@OFFSET(欄位,表示式)	任意	<p>傳回現行記錄中值為 表示式 的記錄偏移量中的 欄位 值。正偏移參照已經經過的記錄 (即 "lookback")，而負偏移指定尚未到達的記錄 (即 "lookahead")。例如，@OFFSET(Status, 1) 傳回前一條記錄中 Status 欄位的值，而 @OFFSET(Status, -4) 「超前查看」序列的四條記錄 (即未通過節點傳送的記錄)，以獲取欄位值。請注意，必須將負 (向前檢查) 偏移指定為常數。(僅限於正偏移) 表示式 還可以是任意的 CLEM 表示式，將通過此表示式給出現行記錄的偏移值。在此情況下，採用該函數的三參數版本可以提高效能 (參閱下個函數)。如果表示式傳回的不是非負整數，將出現錯誤 - 即，計算的期待偏移量不合法。</p> <p>註：自參照的 @OFFSET 函數無法使用文字前瞻。例如，在填入節點中，不能使用 @OFFSET(field1,-2) 之類的表示式來取代 field1 的值。</p> <p>註：在「填充」節點中填入欄位時，實際上存在該欄位的兩個不同值，即填入前值和填入後值。當 @OFFSET 參照自身時，將參照填充後的值。僅對於過去的列，才存在填充後的值，因此自參照 @OFFSET 只能參照過去的列。由於自參照 @OFFSET 不能參照未來的行，因此將對偏移量執行下列檢查：</p> <ul style="list-style-type: none"> <li>• 如果偏移量為文字值，並且指向未來，那麼將在執行開始前報告錯誤。</li> <li>• 如果偏移量為表示式並且在執行時期求值為未來，那麼 @OFFSET 將傳回 \$null\$。</li> </ul> <p>附註：不支援同時在一個節點內使用 "lookahead" 和 "lookback"。</p>
@OFFSET(欄位,表示式, INT)	任意	<p>與 @OFFSET 函數執行相同的作業，但附帶第三個參數 <i>INT</i> (用於指定向後查看值的最大數目量)。當從表示式計算偏移時，採用第三個參數可以提高效能。</p> <p>例如，通過表示式 @OFFSET(Foo, Month, 12)，系統將知道僅保留 Foo 的最後一欄 12 個值，否則，系統將儲存所有滿足條件的值。如果偏移量為常數包括必須為常數的負「期待」偏移量第三個參數將毫無意義，此時應當使用該函數的雙參數版本。另請參閱在前述雙參數版本中自參照函數的說明。</p> <p>附註：不支援同時在一個節點內使用 "lookahead" 和 "lookback"。</p>
@SDEV(欄位)	實數	傳回指定的 欄位 或 <i>FIELDS</i> 值的標準差。
@SDEV(欄位,表示式)	實數	傳回現行節點收到之最後 <i>EXPR</i> 筆記錄 (包括現行記錄) 內 <i>FIELD</i> 值的標準偏差。 <i>FIELD</i> 必須是數值欄位的名稱。 <i>EXPR</i> 可能是求值為大於 0 的整數的任何表示式。如果省略 <i>EXPR</i> ，或者它超出到目前為止收到的記錄數，則會傳回到目前為止收到的所有記錄的標準偏差。
@SDEV(欄位,表示式, INT)	實數	傳回現行節點收到之最後 <i>EXPR</i> 筆記錄 (包括現行記錄) 內 <i>FIELD</i> 值的標準偏差。 <i>FIELD</i> 必須是數值欄位的名稱。 <i>EXPR</i> 可能是求值為大於 0 的整數的任何表示式。如果省略 <i>EXPR</i> ，或者它超出到目前為止收到的記錄數，則會傳回到目前為止收到的所有記錄的標準偏差。 <i>INT</i> 指定要回顧的數目上限。這遠比只使用兩個引數有效得多。

表 33. CLEM 序列函數 (繼續)

函數	結果	說明
@SINCE(表示式)	任意	如果 表示式 (可以是任何的 CLEM 表示式) 的結果為 true，則傳回已傳送的記錄數。
@SINCE(表示式, INT)	任意	新增第二個參數 (INT) 時，會指定要回顧的上限記錄數。如果表示式 結果不為 true，則 INT 為 @INDEX+1。
@SINCE0(表示式)	任意	對於現行記錄，@SINCE 不傳回任何值；而如果現行記錄的表示式 結果為 true，則 @SINCE0 傳回 0。
@SINCE0(表示式, INT)	任意	新增第二個參數 (INT) 時，會指定要回顧的上限記錄數。
@SUM(欄位)	數字	傳回指定欄位 欄位 或 FIELDS 的值總和。
@SUM(欄位, 表示式)	數字	傳回現行節點收到之最後 EXPR 筆記錄 (包括現行記錄) 內 FIELD 的值總和。FIELD 必須是數值欄位的名稱。EXPR 可能是求值為大於 0 的整數的任何表示式。如果省略 EXPR，或者它超出到目前為止收到的記錄數，則會傳回到目前為止收到的所有記錄總和。
@SUM(欄位, 表示式, INT)	數字	傳回現行節點收到之最後 EXPR 筆記錄 (包括現行記錄) 內 FIELD 的值總和。FIELD 必須是數值欄位的名稱。EXPR 可能是求值為大於 0 的整數的任何表示式。如果省略 EXPR，或者它超出到目前為止收到的記錄數，則會傳回到目前為止收到的所有記錄總和。INT 指定要回顧的值數目上限。這遠比只使用兩個引數有效得多。
@THIS(欄位)	任意	傳回現行記錄中名為 欄位 的欄位的值。僅用於 @SINCE 表示式。

## 廣域功能

函數 @MEAN、@SUM、@MIN、@MAX 和 @SDEV 多數時候用於處理所有已讀取的記錄 (包含目前記錄)。但在某些時候，它可用於檢測如何將現行記錄中的值與整個資料集中的值進行比較。如果使用設定廣域節點來產生整個資料集中的值，則您可在 CLEM 表示式中使用廣域函數存取這些值。

例如，

@GLOBAL\_MAX(Age)

傳回資料集中的最大 Age 值，同時表示式

(Value - @GLOBAL\_平均值(Value)) / @GLOBAL\_SDEV(Value)

將給出該記錄 Value 和作為標準離差的廣域平均數之間的差。僅當設定廣域節點計算出廣域值後，您方可使用它們。通過按一下「串流內容」對話框「廣域」標籤中的清除廣域值按鈕，您可以取消目前所有廣域值。

表 34. CLEM 廣域函數.

函數	結果	說明
@GLOBAL_MAX(欄位)	數字	傳回先前由「設定廣域」節點產生的整個資料集的 FIELD 最大值。FIELD 必須是數值欄位、日期/時間/日期時間欄位或字串欄位。如果未設定對應的廣域值，則會發生錯誤。
@GLOBAL_MIN(欄位)	數字	傳回先前由「設定廣域」節點產生的整個資料集的 FIELD 最小值。欄位必須是數值欄位、日期/時間/日期時間欄位或字串欄位。如果未設定對應的廣域值，則會發生錯誤。

表 34. CLEM 廣域函數 (繼續).

函數	結果	說明
@GLOBAL_SDEV(欄位)	數字	傳回先前由「設定廣域」節點產生的整個資料集的 <i>FIELD</i> 標準差。 <i>FIELD</i> 必須是數值欄位的名稱。如果未設定對應的廣域值，則會發生錯誤。
@GLOBAL_平均值(欄位)	數字	傳回先前由「設定廣域」節點產生的整個資料集的 <i>FIELD</i> 值的平均數。 <i>FIELD</i> 必須是數值欄位的名稱。如果未設定對應的廣域值，則會發生錯誤。
@GLOBAL_SUM(欄位)	數字	傳回先前由「設定廣域」節點產生的整個資料集的 <i>FIELD</i> 值總和。 <i>FIELD</i> 必須是數值欄位的名稱。如果未設定對應的廣域值，則會發生錯誤。

## 處理空白值與空值的函數

借助 CLEM，您可以指定在某個欄位中用作「空白值」或遺漏值的特定值。下列函數用於處理空白值。

表 35. CLEM 空白和空值函數.

函數	結果	說明
@BLANK( <i>FIELD</i> )	布林	根據上游類型節點或來源節點（「類型」標籤）中所設定的空白值處理規則，對所有空白值的記錄傳回 true 值。
@LAST_NON_BLANK( <i>FIELD</i> )	任何	傳回 <i>FIELD</i> 的最後一個非空白值，如上游來源或「類型」節點中所定義。如果到目前為止的記錄中的 <i>FIELD</i> 沒有非空白值，則會傳回 \$null\$。請注意，空白值（也稱為使用者遺漏值）可以個別針對每個欄位進行定義。
@NULL( <i>FIELD</i> )	布林	如果欄位的值為系統遺漏值 \$null\$，則傳回 true 值。對於所有其他值（包含使用者定義的空白），則傳回 false。如果想同時檢查空白和空值，則請使用 @BLANK( <i>FIELD</i> ) 和 @BLANK( <i>FIELD</i> )。
undef	任何	一般用於在 CLEM 中輸入 \$null\$ 值，例如，在充填節點中用空值填滿空白值。

在充填節點中可以「填寫」空白值欄位。在填入器節點和衍生節點（僅多種模式）中，特殊的 CLEM 函數 @FIELD 將指向正在檢查的目前欄位。

## 特殊欄位

特殊函數用於指示所檢查的具體欄位，或用於產生輸入欄位清單。例如，當一次衍生多個欄位時，應使用 @FIELD 函數來指示「對所選取的欄位執行衍生動作」。使用表示式 log(@FIELD) 來為每個選定欄位衍生一個新的日誌欄位。

表 36. CLEM 特殊欄位.

函數	結果	說明
@FIELD	任何	對表示式環境定義中指定的所有欄位執行動作。
@TARGET	任何	當 CLEM 表示式用於使用者定義的分析函數時，@TARGET 代表目標欄位或正在分析的目標/預測對的「正確值」。此函數通常用於「分析」節點。

表 36. CLEM 特殊欄位 (繼續).

函數	結果	說明
@PREDICTED	任何	當 CLEM 表示式用於使用者定義的分析函數時，@PREDICTED 代表正在分析的目標/預測對的預測值。此函數通常用於「分析」節點。
@PARTITION_FIELD	任何	替換目前分割區欄位的名稱。
@TRAINING_PARTITION	任何	傳回目前訓練分割區的值。例如，要用選取節點選取訓練記錄，請使用 CLEM 表示式： <code>@PARTITION_欄位 = @TRAINING_PARTITION</code> 此表示式將確保無論資料中的每個分割區使用哪個值代表，選取節點都將照常工作。
@TESTING_PARTITION	任何	傳回目前測試分割區的值。
@VALIDATION_PARTITION	任何	傳回目前驗證分割區的值。
@FIELDS_BETWEEN(start, end)	任何	根據資料中欄位的自然（即插入）順序，傳回指定開始和結束欄位（包含）之間的欄位名稱清單。
@FIELDS_MATCHING(pattern)	任何	傳回與指定型樣相相符的欄位名稱的清單。問號 (?) 可以併入型樣以僅符合一個字元；(*) 符合零個以上字元。若要符合一個問號或星號文字（而不是使用它們作為萬用字元），可將反斜線 (\) 用作跳出字元。 註：這需要以字串文字為引數；它不能使用巢套表示式來產生該引數。
@MULTI_RESPONSE_SET	任何	傳回已命名的多重回應集的欄位清單。



---

## 第 10 章 使用具有儲存庫的 IBM SPSS Modeler

---

### 關於 IBM SPSS Collaboration and Deployment Services 儲存庫

SPSS Modeler 可以與 IBM SPSS Collaboration and Deployment Services 儲存庫結合使用，使您能夠管理資料採礦模型和相關預測物件的生命週期，並使這些物件可供企業應用程式、工具和解決方案使用。可以通過此方式共用的 IBM SPSS Modeler 物件包含串流、節點、串流輸出、專案和模型。這些物件儲存在中央儲存庫中，可以和其他應用程式共用，也可以使用延伸的版本、meta 資料和搜尋功能進行追蹤。

在您將 SPSS Modeler 與儲存庫搭配使用之前，需要在儲存庫主機上安裝一個配接器。沒有該配接器，您在試圖從某些 SPSS Modeler 節點或模型存取儲存庫物件時會看到下列訊息：

可能需要更新儲存庫以支援新節點、模型和輸出類型。

有關安裝配接器的指示，請參閱產品下載過程中以 PDF 檔案形式提供的《SPSS Modeler 部署安裝》手冊。在 SPSS Modeler 部署手冊中提供如何從 IBM SPSS Deployment Manager 存取 IBM SPSS Modeler 儲存庫物件的詳細資料。

下列部分提供從 SPSS Modeler 中存取儲存庫的資訊。

### 可擴展的版本和搜尋支援

儲存庫提供了全面的物件版本和搜尋功能。例如，假定建立了一個串流，並將其儲存到可與其他部門的研究人員共用的儲存庫中。如果稍後在 SPSS Modeler 中更新了此串流，則可以將此更新後的版本新增到儲存庫中而不用改寫以前的版本。可以始終存取所有的版本，並可以按照名稱、標籤、使用的欄位或其他屬性搜尋這些版本。例如，可以搜尋所有將淨營收用作輸入的模型版本，或搜尋所有由特定程序設計員建立的模型。（要在傳統的檔案系統中執行此操作，必須以不同的檔名儲存每個版本，且各個版本間的關係對軟體而言是未知的。）

### 單一登入

單一登入功能可讓使用者連接至儲存庫，而不必每次都輸入使用者名稱和密碼詳細資料。使用者現有的本端網路登入詳細資料提供了 IBM SPSS Collaboration and Deployment Services 所需的必要鑑別資訊。此功能取決於下列條件：

- IBM SPSS Collaboration and Deployment Services 必須配置以使用單一登入提供者。
- 使用者必須登入到與提供者相容的主機中。

如需相關資訊，請參閱第 138 頁的『連接至儲存庫』。

---

### 儲存和部署儲存庫物件

在 IBM SPSS Modeler 中建立的串流可以儲存在儲存庫中，就像儲存為具有副檔名 .str 的檔案一樣。這樣，單個串流便可由企業中的多個使用者存取。請參閱第 139 頁的『在儲存庫中儲存物件』主題，以取得更多資訊。

另外，還可以在儲存庫中部署串流。部署的串流將儲存為具有附加 meta 資料的檔案。部署的串流可以充分利用 IBM SPSS Collaboration and Deployment Services 的企業級功能，例如自動化評分和模型更新。例如，可以按定期計劃間隔在新資料可用時對模型進行自動更新。另外，也可以部署一組串流用於 Champion Challenger 分析，該分析將比較各個串流以確定哪個串流包含最有效的預測模型。

註：關聯規則、STP 及 TCM 建模節點不支援 IBM SPSS Collaboration and Deployment Services 中的「模型評估」或「冠軍挑戰者」步驟。

您可以部署串流（副檔名為 .str）。部署為串流將使該串流能夠由小型用戶端應用程式 IBM SPSS Modeler Advantage 使用。請參閱第 153 頁的『開啟 IBM SPSS Modeler Advantage 中的串流』主題，以取得更多資訊。

如需相關資訊，請參閱第 149 頁的『串流部署選項』。

## 其他部署選項

IBM SPSS Collaboration and Deployment Services 不僅為管理企業內容提供了極其廣泛的功能，還提供部署或匯出串流的其他一些機制，包含：

- 出串流和模型，以供 IBM SPSS Modeler Solution Publisher Runtime 稍後使用。
- 匯出一個或多個 PMML（用於對模型資訊編碼的基於 XML 的格式）中的模型。請參閱第 154 頁的『匯入並匯出 PMML 模型』主題，以取得更多資訊。

---

## 連接至儲存庫

1. 要連接至儲存庫，請在 IBM SPSS Modeler 主功能表中按一下：

工具 > 儲存庫 > 選項...

2. 在儲存庫 **URL** 欄位中，輸入或選取想要存取的儲存庫安裝的目錄路徑或 URL。一次只能連接一個儲存庫。

每個網站或安裝的設定都是特定的。有關特定的登入詳細資料，請與本端系統管理者聯絡。

**設定認證。** 不已勾選此複選框可啟用**單一登入**功能，該功能試圖使您使用本端電腦使用者名稱和密碼詳細資料登入。如果無法使用單一登入，或您勾選此方框以取消單一登入（例如，登入管理者帳戶），則會顯示另外窗口讓您輸入您的認證。

### 輸入儲存庫的認證

根據設定不同，在「儲存庫：認證」對話框中，下列欄位可能是必填項：

**使用者 ID 和密碼。** 指定用於登入的有效使用者名稱和密碼。如果需要，請聯絡本端管理者以獲取進一步資訊。

**提供者。** 選擇安全提供者以進行鑑別。可配置儲存庫以使用不同的安全提供者；必要的話，請聯絡您的本端管理者以取得相關資訊。

請記住儲存庫和使用者 **ID**。將目前設定儲存為預設設定，以便在每次想要連接時不必再次加入這些設定。

### 瀏覽儲存庫認證

從 Analytic Server、Cognos、ODBC 或 TM1 來源節點連接至儲存庫時，可以選取先前記錄的認證來連接至儲存庫。這些認證列出在「選取儲存庫認證」對話框中。要選取此對話框，請按一下認證欄位旁邊的「瀏覽」。

在「選取儲存庫認證」對話框中，強調顯示提供的清單中的認證，然後按一下「確定」。如果清單過大，請使用**過濾器**欄位輸入名稱或名稱的一部分，以尋找所需的認證。



---

## 瀏覽儲存庫內容

在儲存庫中，可以按照與 Windows 檔案總管類似的方式瀏覽其中儲存的內容；也可以瀏覽每個已儲存物件的版本。

1. 要開啟 IBM SPSS Collaboration and Deployment Services 儲存庫 視窗，在 SPSS Modeler 功能表中，按一下：

工具 > 儲存庫 > 探索...

1. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。

瀏覽器視窗初始情況下顯示資料夾階層的樹狀結構形視圖。按一下資料夾名稱可顯示其內容。

此時，符合目前選擇或搜尋準則的物件將列在右側窗格中，已選取版本的詳細資訊將顯示在右下窗格中。顯示的屬性套用至最新的版本。

---

## 在儲存庫中儲存物件

您可以將串流、節點、模型、模型選用區、專案和輸出物件儲存在儲存庫中，其他使用者和應用程式可從這裡對這些內容進行存取。

您也可以其他使用者通過使用 IBM SPSS Collaboration and Deployment Services 部署入口網站在網際網路上能夠檢視的格式將串流輸出發行到儲存庫。

## 設定物件內容

儲存物件時，將顯示「儲存庫：儲存」對話框，以使您能夠設定該物件的眾多內容的值。您可以：

- 選擇用於儲存物件的名稱和儲存庫資料夾
- 新增物件的相關資訊，例如版本標籤及其他可搜尋內容
- 指定一個或多個分類主題到物件
- 設定物件的安全選項

下列幾節將詳細介紹您可設定的內容。

## 選擇儲存物件的位置

在「儲存庫：儲存」對話框中，輸入下列內容。

**儲存位置。**顯示目前資料夾，即物件將儲存到的位置。按兩下清單中的資料夾名稱，將此資料夾設為現行資料夾。使用「上層資料夾」按鈕導覽到上層資料夾。使用「新資料夾」按鈕在現行層次中建立一個資料夾。

**檔案名稱。**儲存物件時使用的名稱。

**儲存。**在目前的位置儲存物件。

## 新增儲存物件的相關資訊

「儲存庫：儲存」對話框的「資訊」標籤上的所有欄位都是選用的。

**作者。**在儲存庫中建立物件的使用者名稱。依預設，將顯示用於儲存庫連接的使用者名稱，但可以在此變更該名稱。

**版本標籤。** 從清單中選取標籤以指示物件版本，或按一下新增以建立新標籤。避免在標籤中使用 "[ ]" 字元。如果不打算為此物件版本指定標籤，請確保未勾選任何複選框。請參閱第 147 頁的『檢視和編輯物件內容』主題，以取得更多資訊。

**說明。** 物件的說明。使用者可以按說明搜尋物件（見附註）。

**關鍵字。** 與物件相關的一個或多個關鍵字，可用於搜尋目的（見附註）。

**到期日。** 物件在此日期之後不再對一般使用者可見，但其擁有人和儲存庫管理者仍然可以請參閱。要設定到期日，選擇日期選項並輸入日期，或使用行事曆按鈕選擇日期。

**儲存。** 在目前的位置儲存物件。

註：說明和關鍵字欄位中的資訊被當做和物件的「註解」標籤上的 SPSS Modeler 中已輸入的任何資訊不同。按說明或關鍵字的儲存庫搜尋沒有從「註解」標籤傳回資訊。請參閱第 144 頁的『在儲存庫中搜尋物件』主題，以取得更多資訊。

## 為儲存物件指定主題

主題是面向儲存庫中所儲存內容的層次分類系統。您可以在儲存物件時從可用主題中進行選擇，使用者還可按主題搜尋物件。可用的主題清單由具備適當專用權的儲存庫使用者設定（如需相關資訊，請參閱《Deployment Manager 使用手冊》）。

要對該物件指定主題，請在「儲存庫：儲存」對話框的「主題」標籤上執行以下操作：

1. 按一下新增按鈕。
2. 從可用主題清單中按一下主題名稱。
3. 按一下「確定」。

要刪除主題指定：

4. 在已指定的主題清單中選取主題。
5. 按一下「刪除」。

## 設定儲存物件的安全選項。

可以在「儲存庫：儲存」對話框的「安全」標籤上針對已儲存的物件設定或變更眾多安全選項。對於一個或多個主體（即使用者或使用者群組），您可以：

- 分配物件的存取權
- 修改物件的存取權
- 刪除物件的存取權

**主體。** 對此物件具有存取權的使用者或群組的儲存庫使用者名稱。

**許可權。** 該使用者或群組對物件的存取權。

**新增。** 允許您新增一個或多個使用者或群組到對該物件具有存取權的使用者/群組清單。請參閱第 141 頁的『新增使用者到權限清單』主題，以取得更多資訊。

**修改。** 允許您修改選定使用者或群組對該物件的存取權。依預設授權讀取存取權。使用此選項可以授權其他存取權，包括擁有人、寫入、刪除和修改權限。

**刪除。** 從該物件的權限清單中刪除選定使用者或群組。

**新增使用者到權限清單：** 在「儲存庫：儲存」對話框的「安全」標籤上選取新增時，下列欄位可用：

**選取提供者。**選擇安全提供者以進行鑑別。可配置儲存庫以使用不同的安全提供者；必要的話，請聯絡您的本端管理者以取得相關資訊。

**尋找。**輸入您要新增的使用者或群組的儲存庫使用者名稱，並按一下**搜尋**以在使用者清單中顯示該名稱。要一次新增多個使用者名稱，請保留該欄位為空白，並按一下**搜尋**即可顯示所有儲存庫使用者名稱的清單。

**使用者清單。**從清單中選取一個或多個使用者名稱，按一下「確定」以將其新增到權限清單中。

**修改物件的存取權：** 在「儲存庫：儲存」對話框的「安全」標籤上選取修改時，下列欄位可用：

**擁有人。**選取此選項為該使用者或群組授予對此物件的擁有人存取權。擁有人對物件具有完全控制權限，包含「刪除」和「修改」存取權。

**讀取。**依預設，非物件擁有者的使用者或使用群組僅具有物件的讀取存取權。選中適當的勾選框，以便為該使用者或群組新增「寫入」、「刪除」和「修改」權限等存取權。

## 儲存串流

您可以在儲存庫中將串流儲存為 *.str* 檔案，其他使用者可以從這裡存取它。

註：有關部署串流以充分利用儲存庫附加功能的詳細資訊，請參閱第 148 頁的『部署串流』。

要儲存目前串流：

1. 在主功能表上，按一下：

檔案 > 儲存 > 儲存為串流...

2. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
3. 在儲存庫中：儲存對話框，選擇要用來儲存物件的資料夾，指定要記錄的任何其他資訊，並按一下**儲存**按鈕。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

## 儲存專案

您可以在儲存庫中將整個 IBM SPSS Modeler 專案儲存為 *.cpj* 檔案，以便其他使用者從這裡存取它。

因為專案檔案是其他 IBM SPSS Modeler 物件的儲存器，因此需要告訴 IBM SPSS Modeler 將專案的物件儲存到儲存庫中。通過使用「專案內容」對話框中的設定可以執行此操作。請參閱第 159 頁的『設定專案內容』主題，以取得更多資訊。

一旦將專案配置為在儲存庫中儲存物件，則不論何時在專案中新增新物件，IBM SPSS Modeler 都可自動提示儲存該物件。

結束 IBM SPSS Modeler 階段作業後，必須儲存專案檔案的新版本，以便此專案檔案可記錄所新增的內容。專案檔案可自動包含（及擷取）其物件的最新版本。如果在 IBM SPSS Modeler 階段作業過程中沒有向專案新增任何物件，則不必重新儲存專案檔案。但是，必須儲存已變更的專案物件（串流、輸出等等）的新版本。

要儲存專案

1. 在 IBM SPSS Modeler 的管理程式窗格的 CRISP-DM 或「類別」標籤上選取專案，然後在主功能表中按一下：

檔案 > 專案 > 儲存專案...

2. 必要的话指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
3. 在儲存庫中：儲存對話框，選擇要用來儲存物件的資料夾，指定要記錄的任何其他資訊，並按一下儲存按鈕。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

## 儲存節點

您可以在儲存庫中將目前串流的個別節點定義儲存為 *.nod* 檔案，其他使用者可以從這裡存取它。

要儲存節點：

1. 用滑鼠右鍵按一下串流畫布中的節點，然後按一下儲存節點。
2. 必要的话指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
3. 在儲存庫中：儲存對話框，選擇要用來儲存物件的資料夾，指定要記錄的任何其他資訊，並按一下儲存按鈕。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

## 儲存輸出物件

您可以在儲存庫中將目前串流的輸出物件儲存為 *.cou* 檔案，其他使用者可以從這裡存取它。

要儲存輸出物件：

1. 在 SPSS Modeler 的管理程式窗格的「輸出」標籤上按一下物件，然後在主功能表中按一下：  
**檔案 > 輸出 > 儲存輸出...**
2. 此外，可用滑鼠右鍵按一下「輸出」標籤中的物件，然後按一下儲存。
3. 必要的话指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
4. 在儲存庫中：儲存對話框，選擇要用來儲存物件的資料夾，指定要記錄的任何其他資訊，並按一下儲存按鈕。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

## 儲存模型和模型選用區

您可以在儲存庫中將個別模型儲存為 *.gm* 檔案，其他使用者可以從這裡存取它。您還可在儲存庫中將「模型」選用區的全部內容儲存為 *.gen* 檔案。

儲存模型：

1. 在 SPSS Modeler 的「模型」選用區上按一下物件，然後在主功能表中按一下：  
**檔案 > 模型 > 儲存模型...**
2. 此外，可用滑鼠右鍵按一下「模型」選用區中的物件，然後按一下儲存模型。
3. 從「完成儲存程序」繼續。

儲存模型選用區：

1. 用滑鼠右鍵按一下模型選用區的背景。
2. 在蹦現功能表中，按一下儲存選用區。
3. 從「完成儲存程序」繼續。

完成儲存程序：

1. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
2. 在儲存庫中：儲存對話框，選擇要用來儲存物件的資料夾，指定要記錄的任何其他資訊，並按一下儲存按鈕。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

---

## 從儲存庫中擷取物件

可以擷取已儲存在儲存庫中的串流、模型、模型選用區、節點、專案和輸出物件。

註：除使用此處介紹的功能表選項外，還可以在 SPSS Modeler 視窗右上角的管理程式窗格的適當標籤上按一下右鍵以擷取串流、輸出物件、模型和模型選用區。

1. 要擷取串流，在 IBM SPSS Modeler 主功能表中，按一下：

檔案 > 擷取串流...

2. 要擷取模型、模型選用區、專案或輸出物件，在 IBM SPSS Modeler 主功能表中按一下：

檔案 > 模型 > 擷取模型...

或

檔案 > 模型 > 擷取模型選用區...

或

檔案 > 專案 > 擷取專案...

或

檔案 > 輸出 > 擷取輸出...

3. 此外，可用滑鼠右鍵按一下管理員或專案窗格，並在蹦現功能表中按一下擷取。
4. 要擷取節點，在 IBM SPSS Modeler 主功能表中，按一下：

插入 > 儲存庫中的節點（或 SuperNode）...

- a. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
5. 在「儲存庫：擷取」對話框中，瀏覽到物件，將其選中並按一下擷取按鈕。請參閱主題，以取得更多資訊。

## 選擇要擷取的物件

在「儲存庫：擷取/搜尋」對話框中，存在下列欄位：

**查找範圍。**顯示現行資料夾的資料夾階層。要導覽到其他資料夾，請從清單中選中並直接導覽到相應位置，或使用該欄位下方的物件清單進行導覽。

「上一層資料夾」按鈕。導覽到層次中目前資料夾的上一層。

「新資料夾」按鈕。在階層中的目前層次上建立新資料夾。

**檔案名稱。**選定物件的儲存庫檔案名稱。要擷取該物件，按一下擷取。

**檔案類型。** 您選擇要擷取的物件類型。在物件清單中只顯示該類型的物件以及資料夾。要顯示不同類型的物件以進行擷取，請從清單中選取對應的物件類型。

**鎖定時開啟。** 依預設，當擷取物件時，它將在儲存庫中被鎖定，以便不會被其他用戶更新。如果不希望在擷取物件時將其鎖定，請取消勾選此複選框。

**說明和關鍵字。** 如果在儲存物件時定義了物件的其他詳細資訊，都會顯示在這裡。請參閱第 139 頁的『新增儲存物件的相關資訊』主題，以取得更多資訊。

**版本。** 要擷取最新的版本以外的物件版本，請按一下此按鈕。此時顯示出所有版本的詳細資訊，因此有助於選擇所需的版本。

## 選取物件版本

要選取儲存庫物件的特定版本，請在「儲存庫：選取版本」對話框中執行以下操作：

1. (選用) 按版本、標籤、大小、建立日期或建立使用者排序清單，請按一下相應欄的標題。
2. 選取您要使用的物件版本。
3. 按一下「繼續」。

---

## 在儲存庫中搜尋物件

可以依名稱、資料夾、類型、標籤、日期或其他準則搜尋物件。

### 依名稱搜尋物件

1. 在 IBM SPSS Modeler 主功能表中，按一下：

工具 > 儲存庫 > 探索...

- a. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
2. 按一下「搜尋」標籤。
  3. 在搜尋已命名的物件欄位中，指定要尋找的物件名稱。

當依名稱搜尋物件時，可以用星號 (\*) 作為萬用字元相符任何字串，用問號 (?) 相符任何單個字元。例如，\*cluster\* 將符合在名稱中任意位置包含字串 cluster 的所有物件。搜尋字串 m0?\_\* 相符 M01\_cluster.str 和 M02\_cluster.str 但不相符 M01a\_cluster.str。搜尋不區分觀察值 ( cluster 與 Cluster 和 CLUSTER 都相符)。

註：如果物件數量巨大，那麼搜尋可能需要較長時間。

### 按其他準則搜尋

可以根據標題、標籤、日期、作者、關鍵字、指標內容或說明執行搜尋。只能找到與所有指定的搜尋準則相符的物件。例如，可以找到所有包含一個或多個叢集模型的串流，這些串流還應用了特定標籤，並且修改於特定日期後。

**物件類型。** 您可以將搜尋限於模型、串流、輸出、節點、SuperNode、專案、模型選用區或其他類型的物件。

- **模型。** 可以按種類 (分類、近似、叢集類別) 或特定建模演算法 (例如 Kohonen) 搜尋模型。

還可以按所使用的欄位進行搜尋，例如，搜尋將名稱為 *income* 的欄位用作輸入或輸出 (目標) 欄位的所有模型。

- **串流**。對於串流，可以按使用的欄位或串流中包含的模型類型（種類或演算法）限定搜尋。

**主題**。可以搜尋具有相關特定主題的模型，可用主題清單由具有適當專用權的儲存庫使用者進行設定（有關進一步資訊，請參閱《部署管理員使用手冊》）。要獲得此清單，請勾選此方框，然後按一下出現的「新增主題」按鈕，從清單中選取一個或多個主題，並按一下「確定」。

**標籤**。用特定物件版本標籤限定搜尋。

**日期**。可以指定建立或修改日期，然後搜尋指定日期範圍之前、之後或者之間的物件。

**作者**。限定搜尋特定使用者建立的物件。

**關鍵字**。搜尋特定關鍵字。在 IBM SPSS Modeler 中，可在「註解」標籤中為串流、模型或輸出物件指定關鍵字。

**說明**。在說明欄位中搜尋特定項目。在 IBM SPSS Modeler 中，可在「註解」標籤中為串流、模型或輸出物件指定說明。多個搜尋詞組可用分號隔開，例如，income; crop type; claim value。（注意在搜尋詞組中，空格很重要。例如，具有一個空格的 crop type 和有兩個空格的 crop type 不同。）

---

## 修改儲存庫物件

您可以直接從 SPSS Modeler 修改儲存庫中的現有物件。您可以：

- 建立、更名或刪除資料夾
- 鎖定或解除鎖定物件
- 刪除物件

## 編輯、重命名和刪除資料夾

1. 要對儲存庫中的資料夾執行作業，在 SPSS Modeler 主功能表中按一下：

工具 > 儲存庫 > 探索...

- a. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
2. 確保資料夾標籤處於作用中狀態。
  3. 要建立新的資料夾，可用滑鼠右鍵按一下上層資料夾並按一下新資料夾。
  4. 要重新命名資料夾，可用滑鼠右鍵按一下該資料夾並按一下重新命名資料夾。
  5. 要刪除資料夾，可用滑鼠右鍵按一下該資料夾並按一下刪除資料夾。

## 鎖定和解除鎖定儲存庫物件

您可以鎖定一個物件，以防止其他使用者更新任一現有版本或建立新的項目版本。物件圖示上的掛鎖符號指示物件被鎖定。



圖 15. 鎖定的物件

要鎖定物件

1. 在儲存庫總管視窗中，用滑鼠右鍵按一下所需物件。
2. 按一下鎖定。

要解除鎖定物件

1. 在儲存庫總管視窗中，用滑鼠右鍵按一下所需物件。
2. 按一下**解除鎖定**。

## 刪除儲存庫物件

在刪除儲存庫中的物件之前，必須確定是要刪除物件的所有版本還是僅刪除其特定版本。

刪除物件的所有版本

1. 在儲存庫總管視窗中，用滑鼠右鍵按一下所需物件。
2. 按一下**刪除物件**。

刪除物件的最新版本

1. 在儲存庫總管視窗中，用滑鼠右鍵按一下所需物件。
2. 按一下「**刪除**」。

刪除物件的以前版本

1. 在儲存庫總管視窗中，用滑鼠右鍵按一下所需物件。
2. 按一下**刪除版本**。
3. 選取要刪除的版本並按一下 **確定** 。

---

## 管理儲存庫物件的內容

可以從 SPSS Modeler 中控制各類物件內容。您可以：

- 檢視資料夾的內容
- 檢視和編輯物件的內容
- 建立、套用和刪除物件的版本標記

### 檢視資料夾內容

要在儲存庫視窗中檢視任何資料夾的內容，可用滑鼠右鍵按一下要檢視的資料夾。按一下**資料夾內容**。

「一般」標籤

此標籤將顯示資料夾的名稱、建立和修改日期。

「權限」標籤

在此標籤中，您可以指定資料夾的讀取許可權和寫入權。將列出有權存取上層資料夾的所有使用者與群組。許可權遵循階層。例如，如果您沒有讀取權，則不能有寫入權。如果您沒有寫入權，則不能有刪除權。

**使用者和群組**。列出對此資料夾至少具有「讀取」存取權的儲存庫使用者與群組。選取「寫入」和「刪除」勾選框，將此資料夾的這些存取權新增到特定使用者或群組。按一下「許可權」標籤右側的**新增使用者/群組**圖示，以將存取權指派給其他使用者和群組。可用的使用者和群組清單由管理者控制。

**串聯權限**。選擇一個選項以控制現行資料夾的變更如何應用到其子資料夾（如果有）。

- **串聯所有權限**。將權限設定從目前資料夾串聯到所有子資料夾。這是一次為多個資料夾快速設定權限的方法。根據需要為上層資料夾設定權限，然後根據需要進行串聯。



- **僅串聯變更。**僅串聯上次應用變更之後所做的變更。例如，新增了一個新群組，並希望授予該群組對銷售分支下所有資料夾的存取權，則可授予該群組對根銷售資料夾的存取權並將變更串聯到所有子資料夾。現有子資料夾的所有其他權限儲存不變。
- **不要階式排列。**所有變更僅套用至現行資料夾，而不串聯到子資料夾。

## 檢視和編輯物件內容

在「物件內容」對話框中，您可以檢視和編輯內容。雖然某些內容無法變更，但一律可以通過新增新版本來更新某個物件。

1. 在儲存庫視窗中，用滑鼠右鍵按一下所需的物件。
2. 按一下物件內容。

### 一般標籤

**名稱。** 儲存庫中顯示的物件名稱。

**建立日期。**物件（而非版本）的建立日期。

**最近一次修改日期。**最新版本的修改日期。

**作者。** 使用者的登入名稱。

**說明。** 依預設，此處包含在 SPSS Modeler 的此物件的「註解」標籤上指定的說明。

**鏈結的主題。**通過儲存庫可在需要時根據主題組織模型和相關物件。可用的主題清單由具備適當專用權的儲存庫使用者設定（如需相關資訊，請參閱《*Deployment Manager 使用手冊*》）。

**關鍵字。** 可在「註解」標籤上為串流、模型或輸出物件指定關鍵字。多個關鍵字之間應由空格進行分隔，且最多可包含 255 個字元。（如果關鍵字中包含空格，則使用引號對其進行分隔。）

### 版本標籤

儲存在儲存庫中的物件可能有多個版本。「版本」標籤中可顯示有關每個版本的資訊。

可為已儲存物件的特定版本指定或修改下列內容：

**版本。** 根據儲存版本的時間產生該版本的唯一 ID。

**標籤。** 版本的目前標籤（如果有的話）。與版本 ID 不同，標籤可從物件的一個版本移動到另一個版本。

還為每個版本顯示了檔案大小、建立日期和建立者。

**編輯標記。**按一下「版本」標籤右上角的編輯標記圖示以定義、套用或刪除儲存物件的標籤。請參閱第 148 頁的『管理物件版本標籤』主題，以取得更多資訊。

### 權限標籤

在「權限」標籤上，可以為物件設定讀取寫入權限。將列出有權存取現行物件的所有使用者與群組。許可權遵循階層。例如，如果您沒有讀取權，則不能有寫入權。如果您沒有寫入權，則不能有刪除權。

**使用者和群組。** 列出對此物件至少具有「讀取」存取權的儲存庫使用者與群組。選取「寫入」和「刪除」勾選框，將此物件的這些存取權新增到特定使用者或群組。按一下「許可權」標籤右側的**新增使用者/群組**圖示，以將存取權指派給其他使用者和群組。可用的使用者和群組清單由管理者控制。

## 管理物件版本標籤

使用「編輯版本標記」對話框可以：

- 套用標籤到選取的物件
- 從選取的物件刪除標籤
- 定義新標籤並將其套用到物件

套用標籤到物件

1. 在 **可用的標記** 清單中選取一個或多個標籤。
2. 按一下向右箭頭按鈕移動所選標籤到 **適用的標記** 清單。
3. 按一下**確定**。

從物件刪除標籤

1. 在**已套用標籤** 清單中選取一個或多個標籤。
2. 按一下向左箭頭按鈕移動所選標籤到 **可用的標記** 清單。
3. 按一下**確定**。

定義新標籤並將其套用到物件

1. 在 **新標記** 欄位中鍵入標籤名稱。
2. 按一下向右箭頭按鈕移動新標籤到 **適用的標記** 清單。
3. 按一下**確定**。

---

## 部署串流

為使串流可用於小型用戶端應用程式 IBM SPSS Modeler Advantage，必須在儲存庫中將其部署為串流（.str 檔案）。

註：無法部署評分分支中存在多個來源節點的串流。

串流可以充分利用 IBM SPSS Collaboration and Deployment Services 的企業級功能。請參閱第 137 頁的『儲存和部署儲存庫物件』主題，以取得更多資訊。

### 要部署目前串流（檔案功能表方法）

1. 在主功能表上，按一下：

**檔案 > 儲存 > 部署**

2. 選擇部署類型並根據需要填寫剩餘對話框內容。
3. 按一下**部署為串流**以部署要與 IBM SPSS Modeler Advantage 或 IBM SPSS Collaboration and Deployment Services 配合使用的串流。
4. 按一下**儲存**。有關進一步資訊，請按一下**說明**。
5. 從「完成部署過程」繼續。

### 要部署目前串流（工具功能表方法）

1. 在主功能表上，按一下：

**工具 > 串流內容 > 部署**

2. 選擇部署類型，根據需要完成「部署」標籤的其餘部分，然後按一下儲存。請參閱『串流部署選項』主題，以取得更多資訊。

## 完成部署過程

1. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。
2. 在「儲存庫：儲存」對話框中，選擇要儲存物件的資料夾，指定要記錄的任何其他資訊，然後按一下儲存。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

## 串流部署選項

「串流選項」對話框中的「部署」標籤可指定用於部署串流的選項。

在部署串流時，您可以在瘦用戶端應用程式 IBM SPSS Modeler Advantage 中開啟和修改該串流。串流以副檔名為 .str 的檔案儲存在儲存庫中。

通過部署串流，可以充分利用 IBM SPSS Collaboration and Deployment Services 附帶的其他功能，如多使用者存取、自動評分、模型更新和 Champion Challenger 分析。

註：關聯規則、STP 及 TCM 建模節點不支援 IBM SPSS Collaboration and Deployment Services 中的「模型評估」或「冠軍挑戰者」步驟。

從「部署」標籤中，還可以預覽 IBM SPSS Modeler 為串流建立的串流說明。請參閱第 47 頁的『串流說明』主題，以取得更多資訊。

**部署類型。**選擇您要部署串流的方式。所有串流都需要指定評分分支才能進行部署；其他要求和選項取決於部署類型。

註：關聯規則、STP 及 TCM 建模節點不支援 IBM SPSS Collaboration and Deployment Services 中的「模型評估」或「冠軍挑戰者」步驟。

- **<無>**。串流將不會被部署到儲存庫。除串流說明預覽外的所有選項均被停用。
- **僅適用於評分**。當您按一下儲存按鈕時，串流被部署到儲存庫。可使用在評分節點欄位中指定的節點對資料進行評分。
- **模型更新**。同「僅適用於評分」，除此之外，還可使用在建模節點和模型塊欄位中指定的物件於儲存庫中更新模型。請注意，依預設在 IBM SPSS Collaboration and Deployment Services 中不支援自動模型更新，因此如果要在從儲存庫執行串流時使用此功能，那麼必須選擇此部署類型。請參閱第 151 頁的『模型更新』主題，以取得更多資訊。

**評分節點。**選取圖表、輸出或匯出節點，以識別用於對資料評分的串流分支。雖然串流實際上可以包含任何號碼的有效分支、模型和終端節點，但必須指定一個且唯一一個評分分支以進行部署。這是部署任何串流時最基本的要求。

**評分參數。**通過此選項可指定參數，在執行評分分支時可對這些參數進行修改。請參閱第 150 頁的『評分和建模參數』主題，以取得更多資訊。

**建模節點。**對於模型更新，指定用於在儲存庫中重新產生或更新模型的建模節點。此建模節點的類型必須與為模型塊指定的類型相同。

**模型建立參數...**通過此選項可指定參數，在執行建模節點時可對這些參數進行修改。請參閱第 150 頁的『評分和建模參數』主題，以取得更多資訊。

**模型塊。**對於模型更新，指定將在儲存庫中每次串流更新（通常作為排定的工作的一部分）時更新或重新產生的模型塊。模型必須位於評分分支上。儘管評分分支可能存在多個模型，但只能指定一個模型。請注意，最初建立串流時，這可能實際上是一個佔位模型，此模型將在新資料可用時進行更新或重新產生。

**部署為串流。**如果要將串流與 IBM SPSS Modeler Advantage 或 IBM SPSS Collaboration and Deployment Services 配合使用，請按一下此選項。

**檢查。**按一下此按鈕以查看此串流是否為可部署的有效串流。所有串流都需要指定評分分支才能進行部署。如果未滿足這些條件，那麼將顯示錯誤訊息。

**儲存。**如果串流有效，則將其部署。否則將顯示錯誤訊息。按一下**修正程式**按鈕，更正錯誤並重試。

**預覽串流說明。**允許您檢視 IBM SPSS Modeler 為串流建立的串流說明內容。請參閱第 47 頁的『串流說明』主題，以取得更多資訊。

**註：**關聯規則、STP 及 TCM 建模節點不支援 IBM SPSS Collaboration and Deployment Services 中的「模型評估」或「冠軍挑戰者」步驟。

## 評分和建模參數

將串流部署到 IBM SPSS Collaboration and Deployment Services 時，可以選擇每次對模型進行更新或分數時，可檢視或編輯哪些參數。例如，您可以指定每次執行工作時的上限和最小值，或可變更的其他值。

1. 要使某個參數可見以便在部署串流後對其進行檢視或編輯，請在「評分參數」對話框的清單中將其選中。

可用參數的清單在串流內容對話框的「參數」標籤中進行定義。請參閱第 43 頁的『設定串流和工作階段參數』主題，以取得更多資訊。

## 評分分支

如果您正部署串流，必須將串流的一個分支指定為**評分分支**（即包含評分節點的分支）。在您將某個分支指定為評分分支後，該分支將在串流畫布上強調顯示，就像評分分支上指向模型塊的模型鏈結那樣。這種視覺化表示法在出現多個分支的複合串流中特別有用，此時評分分支可能並不會立即顯示。

**註：**只能指定單個串流分支作為評分分支。

如果串流已定義有評分分支，則新指定的分支將取代它成為評分分支。通過「自訂顏色」選項，可以設定評分分支的指示顏色。請參閱第 167 頁的『設定顯示選項』主題，以取得更多資訊。

通過「顯示/隱藏串流標記」工具列按鈕，可以顯示或隱藏評分分支指示。



圖 16. 「顯示/隱藏串流標記」工具列按鈕

## ID 評分分支以供部署

您可以從終端節點的蹦現功能表或「工具」功能表中指定評分分支。如果使用蹦現功能表，則會在串流內容的「部署」標籤中自動設定評分節點。

要將分支指定為評分分支（蹦現功能表）

1. 將模型區塊連接至終端節點（來自區塊的下游處理節點或輸出節點）。
2. 用滑鼠右鍵按一下終端節點。
3. 在功能表中，按一下**用作評分分支**。

要將分支指定為評分分支（工具功能表）

1. 將模型區塊連接至終端節點（來自區塊的下游處理節點或輸出節點）。
2. 在主功能表上，按一下：

**工具 > 串流內容 > 部署**

3. 在**部署類型**清單中，根據需要按一下**僅評分或模型更新**。請參閱第 149 頁的『串流部署選項』主題，以取得更多資訊。
4. 按一下**評分節點欄位**，並從清單中選取終端節點。
5. 按一下**確定**。

## 模型更新

模型更新是使用較新資料重新建立串流中現有模型的過程。在儲存庫中的串流本身不會變化。例如，演算法類型和串流相關設定保持不變，但模型會採用新資料；如果新版本的模型工作性能優於舊的模型，則會進行更新。

在一個串流中只能將單個模型塊設定為重新整理，這稱為**更新模型**。如果在串流內容的「部署」標籤上按一下**模型更新**選項（參閱第 149 頁的『串流部署選項』），則您所指定的模型塊將成為更新模型。您還可以從模型塊的蹦現功能表中指定模型為更新模型。要完成此操作，模型塊必須已位於評分分支上。

如果關閉模型塊的「更新模型」狀態，這等效於將串流的部署類型設定為「僅評分」，同時「串流內容」對話框的「部署」標籤將相應更新。通過目前評分分支上的模型塊的蹦現功能表中的**作為更新模型**選項，可以開啟和關閉此狀態。

從評分分支中刪除模型塊的模型鏈結時，還將刪除該模型塊的「更新模型」狀態。您可通過「編輯」功能表或工具列復原對模型鏈結的移除；這樣還會恢復模型塊的「更新模型」狀態。

## 如何選取更新模型

與評分分支相同，在串流中也將強調顯示到更新模型的鏈結。選為更新模型的模型塊，以及由此強調顯示的鏈結，取決於串流中存在多少個模型塊。

### 串流中僅單個模型

如果在 ID 評分分支時其上面只有單個鏈結模型塊，則該模型塊將成為串流的更新模型。

### 串流中有多個模型

如果串流中存在多個鏈結模型塊，則按以下方式選擇更新模型。

如果在「串流內容」對話框的「部署」標籤中定義了模型塊，且該模型塊也存在於串流中，則該模型塊成更新模型。

如果未在「部署」標籤上定義模型塊，或定義的模型塊不在評分分支上，則最接近終端節點的模型塊成為更新模型。

如果您後面取消選取這兩個模型鏈結作為重新整理鏈結，則只會強調顯示評分分支，而不包括鏈結。部署類型被設定為「僅評分」。

註：您可以選擇設定一個鏈結為「取代」狀態，而不是其他鏈結。在這種情況下，選為更新模型的模型塊具有重新整理鏈結，並且在指定評分分支時最接近終端節點。

### 串流中無模型

如果串流中不存在模型，或所有模型均未帶有模型鏈結，部署類型將被設定為「僅評分」。

### 檢查評分分支是否存在錯誤

在指定評分分支時，將檢查其是否存在錯誤。

如果發現了錯誤，評分分支將以評分分支錯誤顏色強調顯示，同時顯示錯誤訊息。通過「自訂顏色」選項，可以設定評分分支的錯誤顏色。請參閱第 167 頁的『設定顯示選項』主題，以取得更多資訊。

如果發現了錯誤，請執行以下操作：

1. 根據錯誤訊息的內容糾正錯誤。
2. 在主功能表上，按一下：

**工具 > 串流內容 > 部署**

並按一下**檢查**。

3. 如有必要，重複此過程直到不再發現錯誤。

---

## 第 11 章 匯出到外部應用程式

---

### 關於匯出到外部應用程式

IBM SPSS Modeler 提供了若干機制，用以將整個資料採礦過程匯出到外部應用程式，以便您所作的準備資料和建立模型工作還可在 IBM SPSS Modeler 之外得以利用。

在上一部分中介紹了如何將串流部署到 IBM SPSS Collaboration and Deployment Services 儲存庫以充分利用其多使用者存取、工作計劃和其他功能。類似地，IBM SPSS Modeler 串流也可以與以下應用程序結合使用：

- IBM SPSS Modeler Advantage
- 能夠以 PMML 格式匯入並匯出檔案的應用程式

有關在 IBM SPSS Modeler Advantage 中使用串流的進一步資訊，請參見『開啟 IBM SPSS Modeler Advantage 中的串流』。

有關將模型以 PMML 檔形式匯出和匯入，以便與其他支援此格式的應用程式共用模型的資訊，請參閱第 154 頁的『匯入並匯出 PMML 模型』。

---

### 開啟 IBM SPSS Modeler Advantage 中的串流

IBM SPSS Modeler 串流可與小型用戶端應用程式 IBM SPSS Modeler Advantage 一起使用。您可以完全在 IBM SPSS Modeler Advantage 中建立自訂應用程式，還可以使用在 IBM SPSS Modeler 中建立的串流作為應用工作流程的基準。

要開啟 IBM SPSS Modeler Advantage 中的串流：

1. 在 IBM SPSS Collaboration and Deployment Services 儲存庫中部署串流，確保按一下部署為串流選項。請參閱第 148 頁的『部署串流』主題，以取得更多資訊。
2. 按一下「在 IBM SPSS Modeler Advantage 中開啟」工具列按鈕，或者在主功能表中按一下：

#### 檔案 > 在 IBM SPSS Modeler Advantage 中開啟

1. 必要的話指定儲存庫的連線設定。請參閱第 138 頁的『連接至儲存庫』主題，以取得更多資訊。如需特定的埠、密碼及其他連線詳細資料，請與您的本端系統管理者聯絡。

註：儲存庫伺服器還必須安裝有 IBM SPSS Modeler Advantage 軟體。

1. 在儲存庫中：儲存對話框，選擇要用來儲存物件的資料夾，指定要記錄的任何其他資訊，並按一下儲存按鈕。請參閱第 139 頁的『設定物件內容』主題，以取得更多資訊。

進行此操作可啟動 IBM SPSS Modeler Advantage，此時串流已經開啟。串流在 IBM SPSS Modeler 中已關閉。

---

## 匯入並匯出 PMML 模型

PMML（也稱為預測模型標記語言）是一種 XML 格式，用於描述資料採礦和統計模型，包含模型的輸入、用於為資料採礦準備資料的轉換，以及定義模型自身的參數。IBM SPSS Modeler 可匯入並匯出 PMML，這使得其能夠與其他支援此格式的應用程式（如 IBM SPSS Statistics）共用模型。

有關 PMML 的詳細資訊，請參閱資料採礦群組網站 (<http://www.dmg.org>)。

### 匯出模型

PMML 匯出受支援大多數模型類型，這些模型類型產生在 IBM SPSS Modeler 中。請參閱『支援 PMML 的模型類型』主題，以取得更多資訊。

1. 用滑鼠右鍵按一下模型選用區上的模型塊。（或者，按兩下畫布上的模型塊並選取「檔案」功能表。）
2. 在功能表上，按一下**匯出 PMML**。
3. 在「匯出」（或「儲存」）對話框中，指定此模型的目標目錄及唯一名稱。

注意：可在「使用者選項」對話框中為 PMML 匯出變更選項。在主功能表上，按一下：

**工具 > 選項 > 使用者選項**

然後按一下 PMML 標籤。

請參閱第 168 頁的『設定 PMML 匯出選項』主題，以取得更多資訊。

### 匯入以 PMML 格式儲存的模型

以 PMML 格式從 IBM SPSS Modeler 或其他應用程式中匯出的模型可以匯入到模型選用區中。請參閱『支援 PMML 的模型類型』主題，以取得更多資訊。

1. 在模型選用區上，用滑鼠右鍵按一下選用區並從功能表中選取**匯入 PMML**。
2. 選取要匯入的檔案並根據需要為變數標籤指定選項。
3. 按一下「開啟」。

**使用變數標籤（如果模型中出現這些標籤）。**PMML 可為資料字典中的變數同時指定變數名稱和變數標籤（例如，Referrer ID，簡稱 *RefID*）。如果在最初匯出的 PMML 中出現變數標籤，則選中此選項可以使用這些變數標籤。

如果已選取變數標籤選項但在 PMML 中沒有變數標籤，則按常態使用變數名稱。

## 支援 PMML 的模型類型

### PMML 匯出

**IBM SPSS Modeler 模型。**在 IBM SPSS Modeler 中建立的下列模型都可匯出為 PMML 4.0 格式：

- C&R 樹狀結構
- QUEST
- CHAID
- 線性迴歸
- 神經網路
- C5.0



- 羅吉斯迴歸
- Genlin
- SVM
- Apriori
- Carma
- K-Means
- Kohonen
- TwoStep
- GLMM (僅支援 Fixed Effect Only GLMM 模式)
- 決策清單
- Cox
- 序列 (不支援序列 PMML 模型評分)
- Statistics 模型

資料庫原生模型。對於使用資料庫原生演算法產生的模型，PMML 匯出不可用。無法匯出使用 Microsoft 的 Analysis Services 或 Oracle Data Miner 建立的模型。

## PMML 匯入

IBM SPSS Modeler 可以匯入並分數由所有 IBM SPSS Statistics 產品的目前版本產生的 PMML 模型，包括從 IBM SPSS Modeler 匯出的模型和由 IBM SPSS Statistics 17.0 或以後版本產生的模型或變換 PMML。這實質上意味著評分引擎可評分的任何 PMML，下列除外：

- 無法匯入 Apriori、CARMA、異常偵測、序列和關聯規則模型。
- 將 PMML 模型匯入到 IBM SPSS Modeler 中後，雖然可以對其進行評分，但不能進行瀏覽。（注意，其中包含最初從 IBM SPSS Modeler 中匯出的模型。為避免此限制，可將模型按產生的模型檔 (\*.gm) 匯出而不是按 PMML 匯出。）
- 在匯入時會執行有限的驗證，但在試圖對模型分數時會執行全面驗證。因此有可能匯入成功，但評分卻失敗或產生不正確的結果。

註：對於匯入到 IBM SPSS Modeler 中的第三方 PMML，IBM SPSS Modeler 將試圖對可以識別並進行評分的有效 PMML 進行評分。但是，無法保證將對所有 PMML 進行評分，也無法保證以應用程式產生 PMML 的方式對 PMML 進行評分。



---

## 第 12 章 專案和報告

---

### 專案簡介

專案是一組與資料採礦作業相關的檔案。專案包含資料串流、圖形、已產生的模型、報告以及在 IBM SPSS Modeler 中建立的任何其他內容。乍一看，IBM SPSS Modeler 專案似乎只是一種組織輸出的方式，但實際上其作用遠不止此。利用專案，可以：

- 註解專案檔案中的每個物件。
- 使用 CRISP-DM 方法指導資料採礦工作。專案還包含 CRISP-DM 說明系統，該系統針對使用 CRISP-DM 進行的資料採礦提供詳細資料和現實範例。
- 將非 IBM SPSS Modeler 物件新增到專案中，如用於表示資料採礦目標的 PowerPoint 投影片放映或與計劃使用的演算法相關的白皮書。
- 根據註解生成全面更新報告和簡單更新報告。為了便於在組織的內聯網上發行，可以用 HTML 的形式產生這些報告。

註：如果 IBM SPSS Modeler 視窗中未顯示專案窗格，請在「檢視」功能表中按一下**專案**。

可通過兩種方式檢視新增到專案中的物件：**類別視圖**和 **CRISP-DM 視圖**。新增到專案的任何內容都會新增到這兩個視圖中，並且您可以在這兩個視圖之間切換以建立效果最佳的組織。

### CRISP-DM 視圖

通過支援跨行業資料採礦標準過程 (CRISP-DM)，IBM SPSS Modeler 專案提供業已證明的非專有方式來組織資料採礦工作的各個方面。CRISP-DM 使用六個階段來說明從開始（收集業務需求）到結束（部署結果）的整個過程。雖然某些階段通常並未涉及 IBM SPSS Modeler 中的工作，但專案窗格仍舊包含全部六個階段，以便您有一個中心位置來儲存和追蹤與專案關聯的所有材料。例如，業務瞭解階段通常包括收集需求並與同事會談以確定目標，而不是在 IBM SPSS Modeler 中處理資料。使用專案窗格，可以將會談記錄儲存在 *Business Understanding* 資料夾中，以便將來參照及納入到報告中。

專案窗格中的 CRISP-DM 視圖還配備了自己的說明系統，該系統可全程指導您完成整個資料採礦過程。在 IBM SPSS Modeler 中，可通過在「說明」功能表中按一下 **CRISP-DM 說明**來存取此說明。

附註：如果視窗中未顯示專案窗格，請按一下「檢視」功能表上的**專案**。

### 設定預設專案階段

新增到專案的物件會被新增到 CRISP-DM 的預設階段。這表示，您需要根據使用物件的資料採礦階段來手動組織物件。最好將預設資料夾設定為您目前正在使用的階段。

選取要作為預設文件夾的階段：

1. 在 CRISP-DM 視圖中，用滑鼠右鍵按一下要設定為預設資料夾的階段的資料夾。
2. 在功能表中，按一下**設為預設值**。

預設資料夾將以粗體的顯示。

### 「類別」視圖

專案窗格中的「類別」視圖根據所建立物件的類型，將您的工作分門別類地組織在 IBM SPSS Modeler 中。可將儲存後的物件新增到下列任何種類中：

- 串流
- 節點
- 模型
- 表格、圖形、報告
- 其他（非 IBM SPSS Modeler 檔案，如與資料採礦工作相關的投影片放映或白皮書）

將物件新增到類別視圖也會將它們新增到 CRISP-DM 視圖的預設階段資料夾中。

附註：如果視窗中未顯示專案窗格，請按一下「檢視」功能表上的**專案**。

---

## 建立專案

專案本質上是一個檔案，包含對與該專案關聯的所有檔案的參照。這表示專案項目既個別儲存又作為專案檔案 (.cpj) 中的參照儲存。因為這種參照結構，請注意下列幾點：

- 必須先個別儲存專案項目，然後才能將其新增到專案中。如果未儲存某項目，系統將提示您先儲存它，然後再將它新增到目前專案中。
- 個別更新的物件（如串流）也會在專案檔案中更新。
- 手動從檔案系統中移動或刪除物件（如串流、節點和輸出物件）會導致專案檔案中的鏈結無效。

## 建立新專案

在 IBM SPSS Modeler 視窗中新建專案非常簡單。您可以從建立新專案開始（如果沒有專案開啟），也可以關閉現有專案並從頭開始。

在主功能表上，按一下：

**檔案 > 專案 > 新建專案...**

## 新增到專案

建立或開啟專案後，便可以使用多種方法來新增物件，如資料串流、節點和報告。

從管理員新增物件

使用 IBM SPSS Modeler 視窗右上角的管理員，可以新增串流或輸出。

1. 從其中一個管理器標籤中選取一個物件（如表格或串流）。
2. 用滑鼠右鍵按一下，然後按一下**新增到專案**。

如果以前儲存過該物件，則會自動將它新增到適當的物件資料夾（在類別視圖中）或預設階段資料夾（在 CRISP-DM 視圖中）。

3. 或者，您可以將物件從管理員拖放到專案窗格中。

註：系統可能會讓您先儲存物件。儲存時，一定要在「儲存」對話框中已選取將檔案**新增到專案**。這樣，在您儲存物件後，將會自動將它新增到專案中。

從畫布新增節點

您可以使用「儲存」對話框從串流畫布新增單個節點。

1. 在畫布中選取一個節點。
2. 用滑鼠右鍵按一下，然後按一下**儲存節點**。或者，在主功能表中，按一下：

### 編輯 > 節點 > 儲存節點...

3. 在「儲存」對話框中，選中將檔案新增到專案。
4. 為節點建立名稱，然後按一下儲存。

這樣會儲存檔案並將它新增到專案中。節點會被新增到類別視圖中的 *Nodes* 資料夾以及 CRISP-DM 視圖中的預設階段資料夾。

### 新增外部檔案

您可以將變化非 IBM SPSS Modeler 物件新增到專案中。這在 IBM SPSS Modeler 中管理整個資料挖掘過程時十分有用。例如，可以將指向資料、附註、簡報和圖片的鏈結儲存在專案中。在 CRISP-DM 視圖中，可將外部檔案新增到您選取的資料夾。在類別視圖中，只能將外部檔案儲存到 *Other* 資料夾。

### 將外部檔案新增到專案：

1. 將檔案從桌面拖到專案中。
- 或
2. 用滑鼠右鍵按一下 CRISP-DM 視圖或類別視圖中的目標資料夾。
  3. 在功能表中，按一下新增到資料夾。
  4. 在對話框中選取一個檔案，然後按一下開啟。

這會在 IBM SPSS Modeler 專案內新增對選取的物件的參照。

## 將專案轉移至 IBM SPSS Collaboration and Deployment Services 儲存庫

可以一步將整個專案（包含所有元件檔案）轉移到 IBM SPSS Collaboration and Deployment Services 儲存庫。所有已經在目標位置的物件均不移動。此功能也可以反過來使用：可以將整個專案從 IBM SPSS Collaboration and Deployment Services 儲存庫 轉移到本端檔案系統。

### 轉移專案

確保在專案窗格中開啟要轉移的專案。

### 要轉移專案，請完成下列步驟：

1. 用滑鼠右鍵按一下根專案資料夾，然後按一下轉移專案。
2. 如果系統提示，請登入 IBM SPSS Collaboration and Deployment Services 儲存庫。
3. 為專案指定新的位置，然後按一下確定。

### 設定專案內容

可以使用專案內容對話框來自訂專案的內容和文件。要存取專案內容，請執行以下操作：

1. 用滑鼠右鍵按一下專案窗格中的物件或資料夾，然後按一下專案內容。
2. 按一下專案標籤以指定基本專案資訊。

**建立日期。**顯示專案的建立日期（不可編輯）。

**摘要。**您可以輸入將在專案報告中顯示的資料採礦專案的摘要。

**內容。**列出專案檔案參照的元件的類型和數量（不可編輯）。

將未儲存的物件儲存為。指定是應將未儲存的物件儲存到本端檔案系統還是儲存在儲存庫中。請參閱第 137 頁的『關於 IBM SPSS Collaboration and Deployment Services 儲存庫』主題，以取得更多資訊。

**載入專案時更新物件參照。**選中此選可更新專案對其元件的參照。註：新增到專案中的檔案不會儲存在專案檔案本身中。更確切地說，專案中儲存的是對這些檔案的參照。這意味著移動或刪除檔案會將該物件從專案中刪除。

## 註解專案

專案窗格提供了多種方式來註解資料挖掘工作。專案級註解常用於追蹤「宏觀」目標和決策，而資料夾或節點註解提供其他詳細資訊。「註解」標籤提供了足夠空間供您記錄專案級詳細資料，比如具有不可恢復丟失資料的資料排除、資料探索過程中形成的可能假設等。

### 註解專案：

1. 選取 CRISP-DM 或「類別」視圖中的專案資料夾。
2. 用滑鼠右鍵按一下該資料夾，然後按一下**專案內容**。
3. 按一下**註釋**標籤。
4. 輸入用於說明專案的關鍵字和文字。

## 資料夾內容和註解

可以對 (CRISP-DM 視圖和類別視圖中的) 單個專案資料夾進行註解。在 CRISP-DM 視圖中，這可能是記錄組織在每個資料採礦階段的目標的一種極為有效的方式。例如，對 *Business Understanding* 資料夾使用註解工具，就可以進行這樣的記錄：「此研究的商業目標是減少高價值客戶的流失」。然後，可通過選中**納入報告**選項自動將此文字併入在專案報告中。

### 註解資料夾：

1. 在專案窗格中選定一個資料夾。
2. 用滑鼠右鍵按一下該資料夾，然後按一下**資料夾內容**。

在 CRISP-DM 視圖中，用每個階段的目的摘要以及完成相關資料採礦作業的指引註解資料夾。您可以刪除或編輯這些註解中的任何一個。

**名稱。** 此區域顯示選取的欄位的名稱。

**工具提示文字。** 建立將滑鼠指標懸停在專案資料夾上時將顯示的自訂工具提示。這在 CRISP-DM 視圖中非常有用，例如，快速概覽每個階段的目標或標示某個階段的狀態 (如「進行中」或「完成」)。

**註解欄位。** 此欄位用於進行可在專案報告中詳細核對的更長註解。CRISP-DM 視圖在註解中包含每個資料採礦階段的說明，但您可自由地為自己的專案自訂此說明。

**併入報告。** 要在報告中納入註解，請選中**納入報告**。

## 物件內容

您可以檢視物件內容並選擇是否在專案報告中納入單個物件。要存取物件內容，請執行以下操作：

1. 用滑鼠右鍵按一下專案窗格中的物件。
2. 在功能表中，按一下**物件內容**。

**名稱。** 此區域列示已儲存的物件的名稱。

**路徑。** 此區域列示已儲存的物件的位置。

併入報告。選中此選項將在產生的報告中納入物件詳細資料。

## 關閉專案

當您結束 IBM SPSS Modeler 或開啟新專案時，現有專案檔案 (.cpj) 將被已關閉。

一些與專案（如串流、節點或圖形）關聯的檔案仍可能開啟。如果您希望使這些檔案開啟，請對訊息...**是否要儲存並關閉這些檔案？**回覆否。

如果在關閉專案之後修改並儲存任何關聯檔案，則下次開啟專案時，會把這些更新過的檔案併入進來。如果不希望將更新過的檔案包括進來，請將該檔案從專案中刪除或用其他檔名儲存它。

---

## 產生報告

專案最有用的一項功能是能夠根據專案項目和註解產生報告。正如 CRISP-DM 方法通篇所討論的，這是有效資料採礦的關鍵元件。您可以直接產生若干種檔案類型的報告，也可以直接輸出到畫面視窗以便立即檢視。從輸出窗口中，可以列印、儲存或檢視 Web 瀏覽器中的報告。還可以將儲存的報告分發給組織中的其他人。

在資料採礦過程中，經常會根據專案檔案多次產生報告以分配給專案參與人員。報告從專案檔案以及所建立的全部註解中精選有關所參照物件的資訊。您可以基於類別視圖或 CRISP-DM 視圖建立報告。

### 產生報告：

1. 選取 CRISP-DM 或「類別」視圖中的專案資料夾。
2. 用滑鼠右鍵按一下該資料夾，然後按一下**專案報告**。
3. 指定報告選項並按一下**產生報告**。

報告對話框中的選項提供多種方式來產生所需類型的報告：

**輸出名稱。**如果選擇將報告輸出傳送到畫面，請指定輸出視窗的名稱。您可以指定自訂名稱，或讓 IBM SPSS Modeler 自動為您命名視窗。

**輸出至畫面。**選取此選項可產生報告並在輸出視窗中進行顯示。請注意，您可以選擇將報告從輸出視窗匯出為各種類型的檔案。

**輸出至檔案。**選取此選項可產生報告並將其儲存為在「檔案類型」清單中指定類型的檔案。

**檔名。**為產生的報告指定檔名。依預設，這些檔案將儲存到 IBM SPSS Modeler \bin 目錄中。使用省略號按鈕 (...) 可指定其他位置。

**檔案類型。**可用檔案類型包括：

- **HTML 文件。**報告儲存為單個 HTML 檔案。如果報告包含圖形，則會將它們儲存為 PNG 檔案並被 HTML 檔案參照。在網際網路上發行報告時，一定要上傳 HTML 檔案以及其參照的所有影像。
- **文字檔。**報告儲存為單個文字檔。如果報告包含圖形，則報告中只併入檔名和路徑參照。
- **Microsoft Word 文件。**報告儲存為單個文件，且所有圖形均直接內含在該文件中。
- **Microsoft Excel 文件。**報告儲存為單個試算表，且所有圖形均直接內含在該試算表中。
- **Microsoft PowerPoint 文件。**每個階段均顯示在一張新的投影片上。所有圖形均直接內含在 PowerPoint 投影片中。
- **輸出物件。**在 IBM SPSS Modeler 中開啟此檔案 (.cou) 時，它與報告格式群組中的**輸出至螢幕**選項相同。

附註：若要匯出至 Microsoft Office 檔案，您必須已安裝對應的應用程式。

**標題。** 指定報告的標題。

**報告結構。** 選取 **CRISP-DM** 或**類別**。CRISP-DM 視圖提供狀態報告，其中包含「宏觀」用法概要以及有關每個資料採礦階段的詳細資料。類別視圖是基於物件的視圖，它更適合對資料和串流進行內部追蹤。

**作者。** 顯示預設使用者名稱，但您可以變更它。

**報告包含。** 選取在報告中包含物件的方法。選取**所有資料夾和物件**將包含新增到專案檔案中的所有項目。您也可以根據是否在物件內容中已選取了**納入報告**來包含項目。或者，要檢查未報告項目，可以選擇只包含標示為排除的項目（此時未選取**納入報告**）。

**選取。** 使用此選項，可通過只選取報告中的**最近專案**來提供專案更新。或者，可通過為**舊有項目**設定參數來追蹤較舊並且可能尚未解析的問題。選取**所有項目**即是在生成報告時不考慮時間因素。

**順序依據。** 您可以選取下列物件特性的組合以便在資料夾中為它們排序：

- **類型。** 按類型將物件群組。
- **名稱。** 字母的組織物件。
- **新增日期。** 根據物件新增到專案中的日期為物件排序。

## 儲存和匯出產生的報告

在畫面上產生的報告顯示在新輸出視窗中。此報告中併入的所有圖形都會顯示為內嵌影像。

### 報告項目

報告中列出了每一個串流的節點總數。該數顯示在下列標題下，這些標題使用的是 IBM SPSS Modeler 項目，而不是 CRISP-DM 項目：

- **資料讀取器。** 來源節點。
- **資料的作者。** 匯出節點。
- **模型建立器。** 建立或建模節點。
- **模型應用器。** 已產生的模型（也稱為塊）。
- **輸出建置器。** 圖表節點或輸出節點。
- **其他。** 任何其他與專案相關的節點。例如，節點選用區上「欄位作業」標籤或「記錄選項」標籤上的可用節點。

### 儲存報告：

1. 在「檔案」功能表中，按一下**儲存**。
2. 指定檔名。

報告即儲存為輸出物件。

### 匯出報告：

3. 在「檔案」功能表中，按一下**匯出**以及要匯出為的檔案類型。
4. 指定檔名。

報告即以您選擇的格式儲存。

您可以將報告匯出為如下類型的檔案：

- HTML



- 文字
- Microsoft Word
- Microsoft Excel
- Microsoft PowerPoint

附註：若要匯出至 Microsoft Office 檔案，您必須已安裝對應的應用程式。

使用視窗頂部的按鈕可以：

- 列印報告。
- 在外部 Web 瀏覽器中以 HTML 形式檢視報告。



---

## 第 13 章 自訂 IBM SPSS Modeler

---

### 自訂 IBM SPSS Modeler 選項

您可以通過多種作業來根據需要自訂 IBM SPSS Modeler。本自訂的內容主要包括設定特定的使用者選項，如記憶體配置、預設目錄以及聲音和顏色的用法。您也可以自訂位於 IBM SPSS Modeler 視窗底部的節點選用區。

---

### 設定 IBM SPSS Modeler 選項

自訂和設定 IBM SPSS Modeler 選項有以下幾種方式：

- 通過在工具 > 選項功能表中按一下系統選項來設定系統選項，如記憶體使用和語言環境。
- 通過在工具 > 選項功能表中按一下使用者選項來設定使用者選項，如顯示字型與顏色。
- 通過在工具 > 選項功能表中按一下說明應用程式，來指定與 IBM SPSS Modeler 協同工作的應用程式的位置。
- 通過在「檔案」功能表中按一下設定目錄或設定伺服器目錄，來指定在 IBM SPSS Modeler 中使用的預設目錄。

另外，還可以設定套用至某些或全部串流的選項。請參閱第 36 頁的『設定串流的選項』主題，以取得更多資訊。

### 系統選項

可以通過在工具 > 選項功能表中按一下系統選項為 IBM SPSS Modeler 指定首選語言或語言環境。還可以在此設定 SPSS Modeler 的最大記憶體使用量，並指定自動儲存串流的頻率。請注意，在此對話框中所做的變更只有在重新啟動 SPSS Modeler 後才會生效。

**最大記憶體量。**選取此選項可限制 IBM SPSS Modeler 的記憶體使用量 (MB)。在某些平台上，如果電腦資源有限或負載過重，則 SPSS Modeler 將限制其處理程序大小以減少電腦的資源耗費。如果所處理的資料數量很大，這有可能導致「記憶體不足」錯誤。可以通過指定新臨界值來減輕記憶體負載。

例如，試圖顯示非常大的決策樹可能會引起記憶體錯誤。在這種情況下，我們建議您將記憶體量增大到最大值 4096Mb。在有可能正在處理非常大數量的資料的此類情況下，增大記憶體限額之後，請關閉 SPSS Modeler，然後從指令行將其啟動，以確保使用上限限度的記憶體數量來正在處理資料。

要從指令行啟動（假定 SPSS Modeler 安裝在預設位置中），請在命令提示字元視窗中輸入下列指令：

```
C:\Program Files\IBM\SPSS\Modeler\18.1.1\bin\modelerclient.exe" -J-Xss4096M
```

**使用系統區域環境。**依預設，此選項處於已選取狀態並被設定為「英文(美國)」。取消選中此選項可從可用語言和語言環境的下拉清單中選取其他語言。

**串流自動儲存間隔 (分鐘)。**指定希望 SPSS Modeler 自動儲存串流的頻率。最大值為 60 分鐘，最小值為 1 分鐘，預設值為 5 分鐘。

### 管理記憶體

除了在「系統選項」對話框中指定最大記憶體設定外，還可以通過其他方式最佳化記憶體使用：

- 在串流內容對話框中調整標準欄位數量上限選項。此選項指定欄位的測量層次變為無類型後列名欄位的上限會員數。請參閱第 37 頁的『設定串流的一般選項』主題，以取得更多資訊。

- 按一下視窗右下角 (此處顯示 IBM SPSS Modeler 正使用的記憶體和已配置的記憶體 ( *xx* MB/ *xx* MB))，強迫 IBM SPSS Modeler 釋放記憶體。按一下此區域使其變暗，隨後記憶體配置數將下降。一旦此區域恢復正常顏色，則表示 IBM SPSS Modeler 已釋放了所能釋放的全部記憶體。

## 設定預設目錄

通過選取「檔案」功能表中的**設定目錄**或**設定伺服器目錄**，可以指定用於檔案瀏覽器和輸出的預設目錄。

- **設定目錄**。可以使用此選項來設定工作目錄。預設工作目錄基於您 IBM SPSS Modeler 版本的安裝路徑，或者來自於用於啟動 IBM SPSS Modeler 的指令行路徑。在本機模式下，工作目錄是用於所有用戶端作業和輸出檔案 (如果是通過相對路徑參照檔案) 的路徑。
- **設定伺服器目錄**。只要有遠端伺服器連線，「檔案」功能表中的「設定伺服器目錄」選項都處於啟用狀態。使用此選項可以為所有伺服器檔案和用於輸入或輸出的資料檔案指定預設目錄。預設伺服器目錄為 *\$CLEO/data*，其中 *\$CLEO* 是伺服器版本 IBM SPSS Modeler 的安裝目錄。在指令行模式下，還可以通過使用帶有 *modelerclient* 指令行參數的 *-server\_directory* 旗標來置換此預設目錄。

## 設定使用者選項

可通過選取**工具 > 選項**功能表中的**使用者選項**，來設定 IBM SPSS Modeler 的一般選項。這些選項適用於 IBM SPSS Modeler 中套用的所有串流。

通過按一下相應標籤可以設定下列類型的選項：

- 通知選項，如模型改寫和錯誤訊息。
- 顯示選項，如圖表和背景顏色。
- 語法顏色顯示選項。
- PMML 匯出選項，將模型匯出為預測模型標示語言 (PMML) 時使用。
- 使用者或作者資訊，例如姓名、姓名首字母和電子郵件位址。此資訊可能會顯示在節點以及您建立的其他物件的「註解」標籤上。
- 在傳統方式和 Analytic Server 模式之間切換。

要設定特定於串流的選項，例如小數點、時間格式及資料格式、最佳化、串流佈置以及串流 Script，請使用可以從「檔案」和「工具」功能表中開啟的「串流內容」對話框。

## 設定通知選項

使用「使用者選項」對話框的「通知」標籤，可以設定有關 IBM SPSS Modeler 中警告和確認視窗的出現及類型的各種選項。您還可以指定產生新的輸出和模型時管理程式窗格中「輸出」和「模型」標籤的行為。

**顯示資料串流執行回饋對話框** 選中此項可在串流執行三秒時顯示包含進度指標的對話框。此對話框還包含有關串流所建立的輸出物件的詳細資料。

- **完成時關閉對話框** 依預設，當串流完成執行時期關閉該對話框。如果您希望此對話框在串流完成後保持可見，請取消選中此勾選框。

**當節點改寫檔案時發出警告** 選中此項將在節點作業改寫現有檔案時發出警告並顯示錯誤訊息。

**當節點改寫資料庫表格時發出警告** 選中此項將在節點作業改寫現有資料庫表格時發出警告並顯示錯誤訊息。

## 音效通知

使用該清單指定是否在發生事件或錯誤時用音效進行通知。這裡提供了多種音效。使用「播放」（揚聲器）按鈕可以播放選取的聲音。使用省略號按鈕 (...) 可瀏覽和選取一種聲音。

註：IBM SPSS Modeler 中用於建立音效的 .wav 檔案儲存在安裝位置的 /media/sounds 目錄下。

- **全部靜音** 選中此項將關閉所有事件的音效通知。

## 視覺通知

此群組中的選項用於指定產生新項目時顯示的內容中右上角的管理程式窗格中「輸出」和「模型」標籤的行為。從清單中選取**新建模型**或**新建輸出**可指定相應標籤的行為。

下列選項適用於**新建模型**：

**取代上一個模型** 如果已選取（預設），將改寫「模型」標籤和串流畫布中該串流的現有模型。如果取消勾選此框，那麼該模型將新增到標籤和畫布上的現有模型中。請注意，此設定將被模型鏈結上的模型取代設定置換。

下列選項適用於**新建輸出**：

**輸出超出 [n] 時覆蓋** 選取當「輸出」標籤上的項目數超過預先指定的數量時是否顯示警告。預設數量為 20；但是，可以根據需要變更此預設值。

下列選項在所有情況下都可用：

**選取標籤** 選擇在串流執行過程中有相應物件產生時是否切換到「輸出」或「模型」標籤。

- 選取始終將切換到管理程式窗格中的相應標籤。
- 選取如果由目前串流產生則只為由畫布中目前可見串流產生的物件切換到相應標籤。
- 選取從不將限制軟體以切換到相應標籤的方式來通知您產生了輸出或模型。

**閃爍標籤** 選取當已產生新的輸出或模型時是否讓管理程式窗格中的「輸出」或「模型」標籤閃動。

- 選取如果未被選取將在管理程式窗格中有新物件產生時閃動相應標籤（如果尚未選定）。
- 選取從不將限制軟體通過防雨板相應標籤來通知您有物件產生。

**捲動選用區使其可見（僅限新建模型）**。選取是否自動捲動管理程式窗格中的「模型」標籤以使最新生成的模型可見。

- 選取始終將啟用捲動。
- 選取如果由目前串流產生則僅對由畫布中目前可見串流產生的物件捲動。
- 選取從不將限制軟體自動捲動「模型」標籤。

**開啟視窗（僅限新建輸出）**。選取是否在產生後自動開啟輸出視窗。

- 選取始終將總是開啟新的輸出視窗。
- 選取如果由目前串流產生則為由畫布中目前可見串流產生的輸出開啟一個新視窗。
- 選取從不將限制軟體自動為產生的輸出開啟新視窗。

若要回復至此標籤的系統預設值，請按一下**預設值**。

## 設定顯示選項

使用「使用者選項」對話框的「顯示」標籤，可以設定 IBM SPSS Modeler 中字型與顏色的顯示選項。

**啟動時顯示歡迎使用對話框** 選中此選項表示在啟動時顯示歡迎使用對話框。歡迎使用對話框提供了選項，這些選項可以啟動應用程式範例指導教學、開啟示範資料流或者現有串流或專案以及建立新串流。

**顯示串流和受標記** 如果已選取此選項，那麼依預設將顯示串流和 SuperNode 中的標記（如果有）。標示包含串流備註、模型鏈結以及評分分支強調顯示。

**標準字型和顏色（重新啟動後生效）** 此控制項框中的選項用於指定 IBM SPSS Modeler 畫面設計、色系以及顯示字型大小。此處選取的選項直到您關閉並重新啟動 IBM SPSS Modeler 後才會生效。

- **外觀與操作方式** 請選取標準色系和畫面設計。您可以從下列各項中進行選擇：

**SPSS Standard**（預設），這是 IBM SPSS 產品的一般設計。

**SPSS Classic**，這是 IBM SPSS Modeler 較早版本的使用者熟悉的設計。

**Windows**，此 Windows 設計有助於加強串流畫布和選用區的對照。

- **節點的預設字型大小** 指定要用於節點選用區以及用於串流畫布中顯示的節點的字型大小。
- **指定固定寬度字型** 要選取固定寬度字型以及相關聯的字型大小以用於 Script 和 CLEM 表示式控制項，請選中此勾選框。預設字型為 Monospace plain；按一下變更... 將顯示可以選取的其他字型的清單。

註：您可以在串流內容對話框的「選項」標籤的「佈置」窗格中設定串流的節點圖示大小。從主功能表中，選擇工具 > 串流內容 > 選項 > 佈置。

**自訂顏色** 此表格列出用於各種顯示項目的目前所選顏色。針對表格中列出的每一個項目，您可以透過按兩下顏色直欄中的對應列並從清單中選取顏色來變更現行顏色。若要指定自訂顏色，請捲至清單底端，並按一下顏色... 項目。

**圖表種類顏色順序** 此表格列出用於新建立的圖形中的顯示的目前所選顏色。這些顏色的順序反映了它們用於圖表的順序。例如，如果使用了列名欄位作為包含四個唯一值的顏色覆蓋，那麼將只使用此處列示的前四種顏色。針對表格中列出的每一個項目，您可以透過按兩下顏色直欄中的對應列並從清單中選取顏色來變更現行顏色。若要指定自訂顏色，請捲至清單底端，並按一下顏色... 項目。此處所作的變更不會影響先前建立的圖形。

若要回復至此標籤的系統預設值，請按一下預設值。

## 設定語法顯示選項

使用「使用者選項」對話框的「語法」標籤，可以設定在 IBM SPSS Modeler 中建立的 Script 中的字型屬性和顯示顏色的選項。

**語法強調顯示。**此表格列出了目前所選的用於各種語法項目（包含字型和可在其中顯示字型的視窗）的顏色。對於表格中所列的每個項目，可通過按一下列中對應下拉清單，並從清單中選取一種顏色，以變更顏色。此外，對於字型項目，可以選擇新增粗體的和斜體字強調。

**預覽。**此表格顯示了一個使用在語法強調顯示表格中選取的顏色和字型屬性的語法顯示範例。此預覽將在您變更任何選擇後更新。

按一下預設值可將此標籤回復為系統預設值。

## 設定 PMML 匯出選項

在 PMML 標籤上，可以控制 IBM SPSS Modeler 如何將模型匯出為預測模型標記語言 (PMML)。請參閱第 154 頁的『匯入並匯出 PMML 模型』主題，以取得更多資訊。

**匯出 PMML。** 可以在此配置最適合目標應用程式的 PMML 變體。

- 選取帶有延伸容許將 PMML 副檔名用於沒有等效標準 PMML 的特例。注意，多數情況下，這會得到與標準 PMML 相同的結果。

- 選取作為標準 **PMML...** 則匯出與 **PMML** 標準盡可能一致的 **PMML**。

**標準 PMML** 選項。如果選擇了作為標準 **PMML...** 選項，則可以選擇兩種有效方式中的一種來匯出線性和邏輯迴歸模型：

- 作為 **PMML < 一般迴歸方法 > 模型**
- 作為 **PMML < 迴歸方法 > 模型**

有關 **PMML** 的詳細資訊，請參閱資料採礦群組網站 (<http://www.dmg.org>)。

## 設定使用者資訊

**使用者/作者資訊**。您在此處輸入的資訊，可以顯示在您建立的節點和其他物件的「註解」標籤上。

## 設定方式

在模式標籤上的 **Modeler** 模式設定中，您可以從下列模式中進行選擇：

- **傳統 SPSS Modeler** 模式在使用者介面中顯示所有可用節點和表示式。
- **Analytic Server** 模式僅顯示 **Analytic Server** 受支援的節點和表示式。但請注意，某些節點和某些 **CLEM** 表示式即使不受 **Analytic Server** 的完全受支援也仍然會顯示。下表格提供了哪些節點受到 **Analytic Server** 受支援、局部受支援以及不支援的一般資訊。

有關更多詳細資料，請參閱受支援的節點。

請參閱 **Analytic Server** 文件以瞭解有關 **Analytic Server** 的進一步資訊。

請注意，如果您配置 **SPSS Modeler** 以在「資料庫建模」選用區上顯示資料庫節點，則在切換模式時這些節點不會受到影響。一律會顯示資料庫節點。如果您使用 **IBM Db2 for z/OS** 或 **IBM Netezza** 整合，則在某些情況下，在切換成 **Analytic Server** 模式之後那些節點可能會從「資料庫建模」選用區中消失。如果發生此情況，請前往工具 > 選項 > **Helper** 應用程式並重設勾選框。

表 37. 節點支援

節點類型 (選用區名稱)	受 Analytic Server 受支援	受 Analytic Server 局部受支援	不受 Analytic Server 受支援
來源	<ul style="list-style-type: none"> <li>• Analytic Server 來源節點</li> </ul>		<ul style="list-style-type: none"> <li>• 資料庫</li> <li>• Var. 檔案</li> <li>• 固定檔案</li> <li>• 統計資料檔案</li> <li>• 資料收集</li> <li>• IBM Cognos</li> <li>• TWC 匯入</li> <li>• TM1 匯入</li> <li>• SAS 檔案</li> <li>• Excel</li> <li>• XML 使用者輸入</li> <li>• Sim Gen</li> <li>• 資料視圖</li> <li>• 地理空間</li> <li>• 物件儲存體</li> <li>• 延伸匯入</li> <li>• R 匯入</li> <li>• SNA (擴散分析和群組分析)</li> </ul>
記錄作業	<ul style="list-style-type: none"> <li>• 選取</li> <li>• 排序</li> <li>• 平衡</li> <li>• 不同</li> <li>• RFM 聚集</li> <li>• 附加</li> <li>• 時間序列</li> <li>• 延伸轉換</li> <li>• 串流式 TCM</li> </ul>	<ul style="list-style-type: none"> <li>• 樣本 (僅對簡單方法支援隨機百分比。複合不受支援。)</li> <li>• 合併 (僅支援按鍵和條件合併)</li> <li>• 聚合 (第一個分位數、第三個分位數和中值不受 Analytic Server 受支援)</li> </ul>	<ul style="list-style-type: none"> <li>• 空間時間限制</li> <li>• CPLEX Optimization</li> </ul>



表 37. 節點支援 (繼續)

節點類型 (選用區名稱)	受 Analytic Server 受支援	受 Analytic Server 局部受支援	不受 Analytic Server 受支援
欄位作業	<ul style="list-style-type: none"> <li>• 類型</li> <li>• 過濾器</li> <li>• 衍生</li> <li>• 填充器</li> <li>• 重新分類</li> <li>• 整合</li> <li>• SetToFlag</li> <li>• 重組</li> <li>• 欄位重新排序</li> <li>• 重新投影</li> <li>• 時間間隔</li> </ul>	<ul style="list-style-type: none"> <li>• 自動資料準備 (僅支援轉換)</li> <li>• 分級 (如果已選取「同分值」屬性保留在目前分級中，那麼不支援 Binning 方法 equalFreq)</li> <li>• RFM 分析 (如果已選取「同分值」屬性保留在目前分級中，那麼不支援 Binning 方法 Tiles)</li> <li>• 分割區 (不受 Analytic Server 受支援，除非使用唯一欄位重複地向分割區分配列)</li> </ul>	<ul style="list-style-type: none"> <li>• 匿名</li> <li>• 歷史</li> <li>• 轉置</li> </ul>
圖形	<ul style="list-style-type: none"> <li>• 統計圖</li> <li>• 多點</li> <li>• 時間圖</li> <li>• 分佈</li> <li>• 直方圖</li> <li>• 集合</li> <li>• 網路</li> <li>• 評估</li> <li>• 地圖視覺化</li> <li>• E-Plot (測試版)</li> <li>• t-SNE</li> </ul>	<ul style="list-style-type: none"> <li>• 圖版 (僅對測量層次為離散、列名、序數或旗標的欄位支援聚集方式函數)</li> </ul>	

表 37. 節點支援 (繼續)

節點類型 (選用區名稱)	受 <b>Analytic Server</b> 受支援	受 <b>Analytic Server</b> 局部受支援	不受 <b>Analytic Server</b> 受支援
建模	<ul style="list-style-type: none"> <li>• 時間序列</li> <li>• TCM</li> <li>• Isotonic-AS</li> <li>• 隨機樹狀結構</li> <li>• Tree-AS</li> <li>• Linear-AS</li> <li>• GLE</li> <li>• LSVM</li> <li>• STP</li> <li>• TwoStep-AS</li> <li>• 關聯規則</li> <li>• XGBoost-AS</li> <li>• Random Forest</li> <li>• K-Means-AS</li> </ul>	<ul style="list-style-type: none"> <li>• 自動分類器</li> <li>• 自動數值 (這兩個節點僅支援分割，當使用「自動分類器」選項啟用分割的情況下在 <b>Analytic Server</b> 上執行時必須提供具有分割角色的欄位)</li> <li>• 延伸 (R 語法建立模型不受 <b>Analytic Server</b> 受支援)</li> <li>• 下列節點僅支援分割和 PSM：               <ul style="list-style-type: none"> <li>- C&amp;R 樹狀結構</li> <li>- 線性</li> <li>- 神經網路</li> <li>- CHAID</li> <li>- Quest</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• 自動叢集</li> <li>• 決策清單</li> <li>• C5.0</li> <li>• 迴歸方法(R)</li> <li>• PCA/因素</li> <li>• 功能選擇</li> <li>• 區別</li> <li>• 羅吉斯</li> <li>• GenLin</li> <li>• GLMM</li> <li>• Bayes 網路</li> <li>• Apriori</li> <li>• Carma</li> <li>• 序列</li> <li>• K-Means</li> <li>• Kohonen</li> <li>• TwoStep</li> <li>• Anomaly</li> <li>• KNN</li> <li>• R</li> <li>• 下列節點具有 asl 但只是讀取/寫入 asl：               <ul style="list-style-type: none"> <li>- Cox</li> <li>- SVM</li> <li>- SLRM</li> </ul> </li> </ul>
輸出	<ul style="list-style-type: none"> <li>• 表格</li> <li>• 矩陣</li> <li>• 分析</li> <li>• 資料審核</li> <li>• 轉換</li> <li>• Statistics</li> <li>• 平均數</li> <li>• 報告</li> <li>• 設定廣域</li> </ul>	<ul style="list-style-type: none"> <li>• 延伸輸出</li> <li>• R 語法輸出</li> </ul>	<ul style="list-style-type: none"> <li>• 模擬評估</li> <li>• 模擬擬合</li> <li>• R 輸出</li> </ul>

表 37. 節點支援 (繼續)

節點類型 (選用區名稱)	受 Analytic Server 受支援	受 Analytic Server 局部受支援	不受 Analytic Server 受支援
匯出	<ul style="list-style-type: none"> <li>Analytic Server 匯出節點</li> </ul>	<ul style="list-style-type: none"> <li>延伸匯出</li> <li>R 語法匯出</li> </ul>	<ul style="list-style-type: none"> <li>資料庫</li> <li>純文字檔案</li> <li>統計資料匯出</li> <li>資料收集</li> <li>Excel</li> <li>IBM Cognos 匯出</li> <li>TM1 匯出</li> <li>SAS</li> <li>XML 匯出</li> <li>物件儲存體</li> <li>R 匯出</li> </ul>
IBM SPSS Statistics			<ul style="list-style-type: none"> <li>統計資料檔案</li> <li>統計資料轉換</li> <li>統計資料統計資料</li> <li>統計資料輸出</li> <li>統計資料匯出</li> </ul>
IBM SPSS 文字分析	<ul style="list-style-type: none"> <li>文字鏈結分析</li> <li>文字採礦</li> <li>語言節點</li> </ul>		<ul style="list-style-type: none"> <li>檔案清單</li> <li>Web 資訊來源</li> <li>文字鏈結分析</li> <li>翻譯</li> <li>文字採礦</li> <li>檔案檢視器</li> </ul>
Python(Y)	<ul style="list-style-type: none"> <li>t-SNE</li> <li>Random Forest</li> </ul>		<ul style="list-style-type: none"> <li>SMOTE</li> <li>一類 SVM</li> <li>XGBoost 樹狀結構</li> <li>XGBoost 線性</li> </ul>
Spark	<ul style="list-style-type: none"> <li>Isotonic-AS</li> <li>XGBoost-AS</li> <li>K-Means-AS</li> </ul>		<ul style="list-style-type: none"> <li>SMOTE</li> <li>一類 SVM</li> <li>XGBoost 樹狀結構</li> <li>XGBoost 線性</li> </ul>

## 自訂節點選用區

串流是通過使用節點建立的。IBM SPSS Modeler 視窗底部的節點選用區包含建立串流時可能需要使用的所有節點。請參閱第 13 頁的『節點選用區』主題，以取得更多資訊。

可以通過以下兩種方式重新組織節點選用區：

- 自訂選用區管理員。請參閱第 174 頁的『自訂選用區管理員』主題，以取得更多資訊。

- 變更包含子選用區的選用區標籤在節點選用區上的顯示方式。請參閱第 175 頁的『建立子選用區』主題，以取得更多資訊。

## 自訂選用區管理員

根據 IBM SPSS Modeler 對您的用途，您可以自訂選用區管理員。例如，如果您經常分析資料庫中的時間序列資料，則可能需要確保在單個選用區標籤中可同時使用「資料庫來源」節點、「時間間隔」節點、「時間序列」節點，以及「時間繪圖圖形」節點。通過選用區管理員，只要在節點選用區中建立自訂選用區標籤，您就可以輕鬆進行這些調整。

通過選用區管理員，您可以執行以下多種作業：

- 控制串流畫布中節點選用區上所顯示的選用區標籤。
- 變更選用區標籤在節點選用區上的顯示順序。
- 建立並編輯您自己的選用區標籤以及任何相關聯的子選用區。
- 編輯標籤上的預設節點選擇。

要存取選用區管理器，請在「工具」功能表中按一下**管理選用區(M)...**。

**選用區名稱。**列出了所有可用的選用區標籤，無論其是否顯示在節點選用區上。這包含已建立的任何選用區標籤。請參閱『建立選用區標籤』主題，以取得更多資訊。

**節點數。**每個選用區標籤上所顯示的節點的數量。此處的號碼較大意味著建立子選用區以便對標籤上的節點進行劃分會更為方便。請參閱第 175 頁的『建立子選用區』主題，以取得更多資訊。

**是否顯示** 選取此欄位可顯示節點選用區上的選用區標籤。請參閱第 175 頁的『顯示節點選用區上的選用區標籤』主題，以取得更多資訊。

**子選用區。**要選取在選用區標籤上顯示的子選用區，請強調顯示所需的**選用區名稱**，然後按一下此按鈕以顯示「子選用區」對話框。請參閱第 175 頁的『建立子選用區』主題，以取得更多資訊。

**還原預設值。**要完全刪除您對選用區和子選用區所做的所有變化和新增，並恢復預設的選用區設定，請按一下此按鈕。

## 建立選用區標籤

要建立自訂選用區標籤：

1. 從「工具」功能表中，開啟選用區管理程式。
2. 在已顯示？直欄的右側，按一下「新增選用區」按鈕；即會顯示「建立/編輯選用區」對話框。
3. 鍵入唯一的**選用區名稱**。
4. 在**可以使用的節點區域**，選取要新增至選用區標籤的節點。
5. 按一下「新增節點」右箭頭按鈕，將強調顯示的節點移動到**選定節點區域**。重複該操作，直到新增完所需的所有節點。

在新增所有必要節點之後，可以變更這些節點在選用區標籤上的顯示順序：

6. 使用簡單的箭頭按鈕可將節點上移或下移一列。
7. 使用列箭頭按鈕可將節點移至清單底部或頂部。
8. 要從選用區中刪除節點，請強調顯示該節點，然後按一下**所選節點區域**右側的「刪除」按鈕。

## 顯示節點選用區上的選用區標籤

在 IBM SPSS Modeler 中有一些可用的選項，您可能從來都不需要使用它們；在此情況下，可以使用「選用區管理員」隱藏包含這些節點的標籤。

要選取將在節點選用區上顯示的標籤：

1. 從「工具」功能表中，開啟選用區管理程式。
2. 使用已顯示？直欄中的勾選框，選取是否包括或隱藏每一個選用區標籤。

若要永久地從「節點選用區」中移除選用區標籤，請強調顯示該節點，然後按一下已顯示？直欄右側的「刪除」按鈕。一旦刪除，選用區標籤將無法回復。

註：除「最愛」標籤以外，您無法刪除 IBM SPSS Modeler 提供的任何預設選用區標籤。

### 變更節點選用區上的顯示順序

在選取要顯示的選用區標籤之後，可以變更它們在「節點選用區」上的顯示順序：

1. 使用簡單的箭頭按鈕可將選用區標籤上移或下移一列。將這些選用區標籤上移可將它們移至節點選用區的左側，反之亦然。
2. 使用列箭頭按鈕可將選用區標籤移至清單底部或頂部。位於清單頂部的選用區標籤將顯示在節點選用區左側。

## 在選用區標籤上顯示子選用區

以控制選用區標籤在「節點選用區」上的顯示順序的方式，可以控制哪些子選用區顯示在它們的父選用區標籤中。

要選取子選用區以便將其顯示在選用區標籤上，請執行以下操作：

1. 從「工具」功能表中，開啟選用區管理程式。
2. 選取所需的選用區。
3. 按一下「子選用區」按鈕；即會顯示「子選用區」對話框。
4. 使用已顯示？直欄中的勾選框，選取是否包括每一個選用區標籤。所有子選用區會始終顯示，且無法刪除。
5. 若要永久地從選用區標籤中移除子選用區，請強調顯示子選用區，然後按一下已顯示？直欄右側的「刪除」按鈕。

註：無法刪除「建模」選用區標籤隨附的預設子選用區。

### 變更選用區標籤上的顯示順序

在選取要顯示的子選用區之後，可以變更它們在父選用區標籤上的顯示順序：

1. 使用簡單的方向鈕將子選用區向上或向下移動一列。
2. 使用「行箭頭」按鈕可以將子選用區移至清單底部或頂部。

當您選取母項選用區標籤時，您建立子選用區會在「節點」選用區上顯示。請參閱第 176 頁的『變更選用區標籤視圖』主題，以取得更多資訊。

## 建立子選用區

由於可以將任何現有的節點新增到您建立的自訂選用區標籤中，因此您可以選取更多的節點，這樣無需捲動畫面就可以在畫面上查看這些節點。為了避免必須進行捲動，您可以建立子選用區，並將您為選用區標籤選擇的

節點放入其中。例如，如果您建立了一個選用區標籤，其中包含通常用於建立串流的節點，那麼可以建立 4 個子選用區以便按來源節點、欄位作業、建模和輸出來劃分選擇。

註：您只能從那些已新增到父選用區標籤的子選用區節點中進行選取。

要建立子選用區，請執行以下操作：

1. 從「工具」功能表中，開啟選用區管理程式。
2. 選取要將子選用區新增到的選用區。
3. 按一下「子選用區」按鈕；即會顯示「子選用區」對話框。
4. 在已顯示？直欄的右側，按一下「新增子選用區」按鈕；即會顯示「建立/編輯子選用區」對話框。
5. 鍵入唯一的子選用區名稱。
6. 在可以使用的節點區域，選取要新增至子選用區的節點。
7. 按一下「新增節點」右箭頭按鈕，將所選節點移動到選定節點區域。
8. 新增完所需的節點後，按一下**確認**以傳回至「子選用區」對話框。

當您選取母項選用區標籤時，您建立的子選用區會在「節點」選用區上顯示。請參閱『變更選用區標籤視圖』主題，以取得更多資訊。

### 變更選用區標籤視圖

由於 IBM SPSS Modeler 中存在很多可用的節點，因此如果不捲動至節點選用區左側或右側，則可能無法在較小的畫面上顯示所有的節點；這種情況在建模選用區標籤上尤其明顯。要減少捲動操作，您可以選擇僅顯示子選用區（如果存在）中包含的節點。請參閱第 175 頁的『建立子選用區』主題，以取得更多資訊。

要變更選用區標籤上顯示的節點，請選取選用區標籤，然後從左側的功能表中選取顯示所有節點或者僅顯示特定子選用區中的節點。

---

## 第 14 章 串流和節點的效能考量

可對串流進行設計以獲得最佳效能，主要方法有：使用最有效的配置排列節點；必要時啟用節點快取；以及注意本節中詳細列出的其他考量。

除了此處討論的考量外，您通過有效地使用資料庫（尤其是通過對 SQL 進行最佳化製作），通常可以獲得額外的更加實質性的效能提高。

---

### 節點的順序

即使未在使用 SQL 最佳化，串流中節點的順序也會影響效能。通常的目標是最大程度地減少下游正在處理；因此當您具有減少資料數量的節點時，請將它們放到串流的開頭附近。IBM SPSS Modeler Server 可在編譯期間自動套用一些重新排序規則以將某些節點前置（如果經證實這樣做是安全的）。（此功能預設處於啟用狀態。請向系統管理者核實，以確保它在您的安裝中處於啟用狀態。）

使用 SQL 最佳化時，您會希望上限程度地提高它的可用性和效率。由於當串流包含無法在資料庫中執行的作業時最佳化會停止，因此，最好在串流的開頭將 SQL 最佳化的作業群組成一組。此策略可將大多數正在處理保持在資料庫中進行，因此進入 IBM SPSS Modeler 中的資料較少。

在大多數資料庫中，均可執行下列作業。嘗試在串流的開頭將它們群組為一組：

- 按索引鍵合併（結合）
- 選取
- 聚集
- 排序
- 樣本
- 附加
- 在包含模式下執行的可相異的作業，在該模式下將選取所有欄位
- 填入器作業
- 使用標準算術或字串操作的基本衍生作業（具體取決於資料庫受支援哪些作業）
- 設為旗標

在大多數資料庫中，均無法執行下列作業。應在完成先前清單中的作業之後將下列作業放置在串流中：

- 針對任何非資料庫資料（如純文字檔案）的作業
- 按順序合併
- 平衡
- 在捨棄模式下或只選取了部分欄位作為可相異的欄位的情況下執行的可相異的作業
- 需要存取來自記錄（正在處理之中的記錄除外）的資料的任何作業
- 狀態和計數欄位衍生
- 歷史節點作業
- 涉及「@」（時間序列）函數的作業
- 類型檢查模式 警告 和 放棄
- 模型建構、應用和分析

註：樹狀結構、結果集、線性迴歸和因素已產生的模型可以產生 SQL，因此可以將其推送回到資料庫中。

- 在除正在正在處理資料的資料庫以外的任何位置進行的資料輸出

---

## 節點快取

為最佳化串流執行，可以在任何非終端節點上建立快取。如果已在節點上建立了快取，則在下次執行資料串流時串流過節點的資料會充滿快取。從這時起，就會從快取（儲存在磁碟的暫存目錄中）而不是從資料來源中讀取資料了。

對於諸如排序、合併或聚集這樣比較耗時的作業，快取是最有用的。例如，假設已設定可從資料庫中讀取銷售資料的來源節點，和可按區域聚合銷售資料的「聚合」節點。可以在「聚合」節點上建立快取而不是在來源節點上建立快取，因為希望快取儲存的是已聚合的資料而不是整個資料集。

註：來源節點位置的快取僅用於儲存讀取入 IBM SPSS Modeler 的原始資料複本，在大多數情況下，它不會帶來效能方面的提升。

對於已啟用快取的節點，會在右上角顯示一個小的文件圖示。當節點上的資料被快取時，文件圖示將變綠。

### 啟用快取

1. 在串流畫布上，用滑鼠右鍵按一下節點並按一下功能表上的**快取**。
2. 在快取子功能表中，按一下**啟用**。
3. 可以通過用滑鼠右鍵按一下節點，然後在快取子功能表中按一下**取消**來關閉快取。

### 資料庫中的快取節點

對於在資料庫中執行的串流，可於中遊將資料快取到資料庫的暫時表格中而不是檔案系統中。結合 SQL 最佳化使用時，這可能會造成效能顯著提升。例如，合併多個表格以建立資料採礦視圖的串流中的輸出在必要時可能會加以快取並重複使用。通過為所有下游節點自動產生 SQL，效能可得到進一步的提高。

為利用資料庫快取，必須同時啟用 SQL 最佳化和資料庫快取。注意，伺服器上的最佳化設定將置換用戶端上的最佳化設定。請參閱第 39 頁的『設定串流的最佳化選項』主題，以取得更多資訊。

如果已啟用資料庫快取，則只需滑鼠右鍵按一下任意非終端節點就可以快取該點上的資料，並且在下次執行串流時，可以直接在資料庫中自動建立快取。如果未啟用資料庫快取或 SQL 最佳化，則快取將會改為寫入檔案系統。

註：下列資料庫支援用於快取的暫時表格：Db2、Oracle、SQL Server 和 Teradata。其他資料庫（例如 Netezza）將使用標準表格用於資料庫快取。可以針對特定資料庫自訂 SQL 代碼 - 請與服務人員聯絡以獲取協助。

---

## 效能：處理節點

**排序。**排序節點必須讀取整個輸入資料集才能為該集進行排序。資料儲存庫在記憶體中，直到達到某個限制，並且超出限制的資料會溢出到磁碟上。排序演算法是一種組合演算法：將資料讀取到記憶體中（直到達到某個限制）並使用 Hybrid 式快速排序演算法為其排序。如果所有資料均在記憶體中，則排序完成。否則，會套用合併排序演算法。將排序後的資料寫入至檔案中，並將下一個資料組塊讀取到記憶體中，進行排序，然後寫入到磁碟上。此處理會一直重複，直到讀取了所有資料；然後，合併排序後的組塊。合併可能需要多次轉移儲存在磁碟上的資料。在使用高峰期，排序節點將在磁碟上擁有兩個完整的資料集副本：已排序和未排序。



該演算法的總執行時間依賴於  $N \cdot \log(N)$  的階數，其中  $N$  是記錄數。在記憶體中排序比從磁碟進行合併速度快，因此通過配置更多記憶體來進行排序可減少實際執行時間。該演算法為自己配置一部分實體 RAM（由 IBM SPSS Modeler Server 配置選項記憶體使用倍增器控制）。要增加用於排序的記憶體，請提供更多實體 RAM 或增加此值。請注意，當使用的記憶體比率超過過程的工作集從而導致部分記憶體分頁到磁碟上時，效能會下降，原因是：記憶體內排序演算法的記憶體存取型樣是隨機的，可導致過度分頁。排序演算法由排序節點以外的多個節點套用，但是套用的效能規則相同。

**bin。** Bin 節點讀取整個輸入資料集來計算 Bin 範圍，然後為記錄 Bin。計算範圍時會快取資料集；然後，重新掃描資料集以進行配置。當 Binning 方法是定寬 或平均數+標準差 時，直接將資料集快取到磁碟上。這些方法具有線性執行時間，並且需要足夠的磁碟空間來儲存整個資料集。當 Binning 方法為秩 或分位顯示 時，使用之前介紹的排序演算法為資料集排序，並且將排序後的資料集作為快取使用。排序為這些方法提供執行時間  $M \cdot N \cdot \log(N)$ ，其中  $M$  是 bin 欄位個數，並且  $N$  是記錄數；它要求磁碟空間是資料集大小的兩倍。

根據產生的 Bin 產生衍生節點將提高後續傳送的效能。衍生作業比 bin 速度快很多。

**按索引鍵合併（結合）。** 當合併方法為鍵值時（等效於資料庫結合），合併節點按鍵欄位為它的每個輸入資料集排序。程序的此部分的執行時間為  $M \cdot N \cdot \log(N)$ ，其中  $M$  是輸入數， $N$  是最大輸入中的記錄數；它要求有足夠的磁碟空間來儲存它的所有輸入資料集以及最大資料集的第二個副本。合併本身的執行時間與輸出資料集的大小成比例，輸出資料集的大小取決於比對鍵值的頻率。在「輸出是輸入的卡氏乘積」這一最糟糕的情況下，執行時間可能會接近於  $NM$ 。這種情況很少見，大多數結合的相符鍵值很少。如果一個資料集相對大於其他資料集，或者，如果已經按鍵欄位為輸入資料排序，則可以使用「最佳化」標籤提高此節點的效能。

**分類聚合。** 未設定鍵是連續的選項時，此節點先讀取（但不儲存）它的整個輸入資料集，然後生成任意聚合的輸出。在其他極端的狀況下（聚合資料的大小達到某個限制，該限制由 IBM SPSS Modeler Server 配置選項記憶體使用倍增器確定），將會為資料集的其餘部分進行排序並對其進行處理，就像設定了鍵是連續的選項一樣。如果設定了此選項，則不會儲存任何資料，原因是：聚合輸出記錄在讀取輸入資料時生成。

**相異的。** 「相異的」節點將所有唯一索引鍵欄位儲存在輸入資料集中；如果所有欄位都是鍵欄位且所有記錄都是唯一的，則其儲存整個資料集。依預設，「相異的」節點排序鍵欄位上的資料，然後從每個群組選取（或捨棄）第一個相異的記錄。對於包含少量相異的鍵值的較小型資料集，或者對於預先進行了排序的資料集，可以選擇選項以提高正在處理速度和正在處理效率。

**類型。** 在一些情況下，類型節點在讀取值時快取輸入資料；快取用於下游的正在處理。快取需要足夠的磁碟空間來儲存整個資料集，但會加速正在處理。

**評估。** 評估節點必須為輸入資料排序才能計算分位顯示。因為在每種情況下分值和所產生的記錄順序是不同的，所以會對所評估的每個模型重複進行排序。執行時間為  $M \cdot N \cdot \log(N)$ ，其中  $M$  是模式個數， $N$  是記錄數。

---

## 效能：建模節點

**類神經網路和 Kohonen。** 類神經網路訓練演算法（包括 Kohonen 演算法）會多次傳送訓練資料。資料儲存在記憶體中，直到達到某個限制，並且超出限制的資料會溢出到磁碟上。從磁碟存取訓練資料的成本很高，原因是：存取方法是隨機的，這可能會導致磁碟活動過多。您可以針對這些演算法取消磁碟儲存，強制將所有資料儲存在記憶體中，方法是：在節點對話框的「模型」標籤上選取速度最佳化選項。請注意，如果儲存資料所需的記憶體數量大於伺服器過程的工作集，則會將資料的一部分分頁到磁碟上，並且效能將相應地降低。

當最佳化記憶體啟用時，會根據 IBM SPSS Modeler Server 配置選項 建模記憶體限制百分比 的值將一定百分比的實體 RAM 配置給演算法。要使用更多記憶體來訓練神經網路，請提供更多 RAM 或增加此選項的值，但是請注意，將該值設定得太大將導致分頁。

類神經網路演算法的執行時期間取決於所需的精確度層次。您可以通過在節點的對話框中設定停止於條件來控制執行時期間。

**K-Means**。K-Means 叢集演算法具有與類神經網路演算法相同的記憶體使用控制選項。但是，因為對資料的存取是按順序進行的，所以儲存在磁碟上的資料的效能會較好。

---

## 效能：CLEM 表示式

對資料串流進行回顧的 CLEM 序列函數 (@ 函數) 必須儲存足夠的資料才能滿足最長的回顧。對於其回顧度未作限制的作業，必須儲存欄位的所有值。未限制的作業是指偏移值不是整數的作業；例如，@OFFSET(Sales, Month)。偏移值是欄位名稱 *Month*，在執行之前其值是未知的。伺服器必須儲存 *Sales* 欄位的所有值以便確保結果準確。如果上限是已知的，則應將它作為附加參數提供；例如，@OFFSET(Sales, Month, 12)。此作業指示伺服器最多儲存 *Sales* 的 12 個最新值。序列函數 (限制的或未限制的) 幾乎始終禁止 SQL 產生。

---

## 第 15 章 IBM SPSS Modeler 中的協助工具

---

### IBM SPSS Modeler 協助工具概觀

IBM SPSS Modeler 為所有使用者提供協助工具支援，還為視力和有其他功能性損傷的使用者提供針對性支援。本節說明了使用協助工具加強功能（例如，螢幕閱讀器和鍵盤快速鍵）的操作特點和方法。

---

### 協助工具支援的類型

無論您視力受損還是依賴於鍵盤進行操作，都可以通過多種替代方案來使用此資料採礦工具箱。例如，您無需使用滑鼠即可建立串流，指定選項以及讀取輸出。後續主題列出了可用的鍵盤快速鍵。另外，IBM SPSS Modeler 還廣泛支援螢幕閱讀器（例如 JAWS for Windows）。您還可以最佳化色系，以提供更高的對照。下列主題討論了這些支援類型。

### 針對視力受損用戶的協助工具

您可以在 IBM SPSS Modeler 中指定多項內容，它們將增強使用此軟體的能力。

#### 顯示選項

可以為圖形顯示選取顏色。也可以選擇將特定的 Windows 設定用於此軟體。這樣可能有助於增強視覺對照。

1. 若要設定顯示選項，請在「工具」功能表上，按一下**使用者選項**。
2. 按一下**顯示標籤**。此標籤上的選項包含軟體色系、圖表顏色和節點字型大小。

註：螢幕閱讀器無法讀取圖形，因此有視力障礙的使用者無法存取圖形。

#### 針對通知使用音效

通過開啟或關閉音效，您可以控制接收針對軟體中特定作業的警示的方式。例如，您可以針對節點建立和刪除或產生新輸出或模型等事件啟動音效。

1. 若要設定通知選項，請在「工具」功能表上，按一下**使用者選項**。
2. 按一下**通知標籤**。

#### 控制新視窗的自動啟動

「使用者選項」對話框上的「通知」標籤還用於控制是否在另一個視窗中啟動新產生的輸出，例如表格和圖表。對您而言，取消此選項並且僅在需要時開啟輸出視窗可能更為簡單。

1. 若要設定這些選項，請在「工具」功能表上，按一下**使用者選項**。
2. 按一下**通知標籤**。
3. 在該對話框中，從**視覺化通知**群組中的清單選取新輸出。
4. 在開啟視窗下面，選取**永不**。

#### 節點大小

可以使用標準大小或較小大小顯示節點。您可能希望調整這些大小以滿足您的需求。

1. 要設定節點大小選項，請在「檔案」功能表上按一下**串流內容**。
2. 按一下**佈置標籤**。

3. 從圖示大小清單中，選取標準。

## 盲人使用者的協助工具

對盲人使用者的支援主要通過螢幕閱讀器實現，例如 JAWS for Windows。為了最佳化螢幕閱讀器與 IBM SPSS Modeler 的配合使用，您可以指定多項設定。

### 顯示選項

畫面的視覺對照越高，螢幕閱讀器的運行效果越好。如果已具有高對比的 Windows 設定，那麼可以選擇將這些 Windows 設定用於軟體自身。

1. 若要設定顯示選項，請在「工具」功能表上，按一下**使用者選項**。
2. 按一下**顯示標籤**。

註：螢幕閱讀器無法讀取圖形，因此有視力障礙的使用者無法存取圖形。

### 針對通知使用音效

通過開啟或關閉音效，您可以控制接收針對軟體中特定作業的警示的方式。例如，您可以針對節點建立和刪除或產生新輸出或模型等事件啟動音效。

1. 若要設定通知選項，請在「工具」功能表上，按一下**使用者選項**。
2. 按一下**通知標籤**。

### 控制新視窗的自動啟動

「使用者選項」對話框上的「通知」標籤還用於控制是否在另一個視窗中啟動新產生的輸出。對您而言，停用此選項並視需要開啟輸出視窗可能更為簡單。

1. 若要設定這些選項，請在「工具」功能表上，按一下**使用者選項**。
2. 按一下**通知標籤**。
3. 在該對話框中，從**視覺化通知**群組中的清單選取**新輸出**。
4. 在開啟視窗下面，選取**永不**。

## 鍵盤協助工具

可以通過鍵盤存取產品功能。基本上，您可以按 Alt 鍵加上相應鍵啟動視窗功能表（例如，按 Alt+F 可以存取「檔案」功能表）或者按 Tab 鍵在對話框控制項間進行捲動。此外還有與各種產品的主要視窗相關的**特殊議題**，以及有助於瀏覽對話框的實用提示。

本節將介紹了鍵盤協助工具的亮點，包括開啟串流、使用節點對話框以及對輸出執行操作等。另外，還提供了一個鍵盤快速鍵清單，以提高瀏覽效率。

### 用於瀏覽主視窗的快速鍵

大部分資料採礦工作都在 IBM SPSS Modeler 的主視窗中完成。該主要區域稱為**串流畫布**，用於建立和執行資料串流。該視窗的底部包含**節點選用區**，其中包含了所有可用節點。選用區群組織在標籤上，這些標籤與每個節點群組的資料採礦作業類型相對應。例如，用於將資料引入 IBM SPSS Modeler 的節點在「來源」標籤上進行群組，而用於衍生、過濾欄位或對欄位設置類型的節點在「欄位作業」標籤上進行群組。

視窗的右端包含一些用於管理串流、輸出和專案的工具。右側的上半部分包含**管理員**以及分別用於管理串流、輸出和產生的模型的三個標籤。通過選取標籤，然後從清單中選取物件，您可以存取這些物件。右側的下半部分包含**專案窗格**，您可以通過它將工作分組為專案。此區域中有兩個標籤，它們反映了專案的兩個不同視圖。

類別視圖會按類型對專案物件進行排序，而 **CRISP-DM** 視圖按相關的資料採礦階段（例如，資料預備或建模）對物件進行排序。說明系統和《使用者手冊》中介紹了該 IBM SPSS Modeler 視窗的這些不同部分。

下列是用於在 IBM SPSS Modeler 主視窗中移動以及建立串流的快捷鍵表格。後續主題中列出了用於對話框和輸出的快捷鍵。請注意，只有在主視窗中才可以使用這些快捷鍵。

表 38. 主視窗捷徑

捷徑按鍵	函數
Ctrl+F5	將對焦點移至節點選用區。
Ctrl+F6	將對焦點移至串流畫布。
Ctrl+F7	將對焦點移至管理程式窗格。
Ctrl+F8	將對焦點移至專案窗格。

表 39. 節點及串流捷徑

捷徑按鍵	函數
Ctrl+N	建立新的空白串流畫布。
Ctrl+O	顯示「開啟」對話框，您可以從中選取並開啟現有串流。
Ctrl+數字鍵	將對焦點移至視窗或窗格上的相應標籤。例如，在標籤窗格或視窗中，Ctrl+1 用於將焦點移至左起第一個標籤，Ctrl+2 用於將焦點移至左起第二個標籤，依此類推。
Ctrl+向下方向鍵	在節點選用區中，用於將對焦點從選用區標籤移至該標籤下的第一個節點。
Ctrl+向上方向鍵	在節點選用區中，用於將對焦點從某個節點移至其選用區標籤。
Enter	如果在節點選用區中選取了某個節點（包括已產生的模型選用區中的優化模型），那麼按下此鍵會將該節點新增至串流畫布。在畫布中選取了某個節點後，按 <b>Enter</b> 將開啟該節點的對話框。
Ctrl+Enter	在選用區中選取某個節點後，此快捷鍵會將該節點新增至串流畫布而不選取它，並且會將對焦點移至選用區中的第一個節點。
Alt+Enter	在選用區中選取某個節點後，此快捷鍵會將該節點新增至串流畫布並選取它，同時會將對焦點移至選用區中的第一個節點。
Shift+ 空白鍵	對焦點位於選用區中的節點或備註上時，此快捷鍵用於在選中與取消選中該節點或備註之間進行切換。 如果還已選取了任何其他節點或備註，那麼這會將它們取消選取。
Ctrl+Shift+空格鍵	對焦點位於串流中的節點或備註上，或者位於選用區中的節點或備註上時，此快捷鍵用於在選中與取消選中該節點或備註之間進行切換。 這不會影響任何其他所選節點或備註。
向左/向右方向鍵	如果對焦點位於串流畫布，那麼此快捷鍵用於在畫面上水平移動整個串流。 如果對焦點位於選用區標籤上，那麼此快捷鍵用於在標籤之間週期移動。如果選用區節點具有焦點，請在選用區中的節點之間移動。
向上/向下方向鍵	如果對焦點位於串流畫布中，那麼此快捷鍵用於在畫面上垂直移動整個串流。如果選用區節點具有焦點，請在選用區中的節點之間移動。如果對焦點位於某個子選用區上，那麼此快捷鍵用於在此選用區標籤的其他子選用區之間進行移動。
Alt+向左/向右方向鍵	按方向鍵的方向在串流畫布中水平移動所選節點和備註。
Alt+向上/向下方向鍵	按方向鍵的方向在串流畫布中垂直移動所選節點和備註。
Ctrl+A	選取串流中的所有節點。

表 39. 節點及串流捷徑 (繼續)

捷徑按鍵	函數
Ctrl+Q	對焦點位於某個節點上時，此快捷鍵用於選取該節點及下游的所有節點，並用於取消選取上游的所有節點。
Ctrl+W	對焦點位於所選節點上時，此快捷鍵用於取消選取該節點以及下游的所有所選節點。
Ctrl+Alt+D	複製所選節點。
Ctrl+Alt+L	如果選取了串流中的模型塊，那麼此快捷鍵用於開啟「插入」對話框，以便您將已儲存的模型從 .nod 檔案載入串流中。
Ctrl+Alt+R	針對所選節點顯示「註釋」標籤，從而使您能夠將該節點更名。
Ctrl+Alt+U	建立「使用者輸入」來源節點。
Ctrl+Alt+C	開啟或關閉節點的快取。
Ctrl+Alt+F	清空節點的快取。
Tab	在串流畫布上，循環切換現行串流中的所有來源節點及註解。在節點選用區上的各個節點之間移動。在所選子選用區上，移至該子選用區中的第一個節點。
Shift+Tab	按逆序執行 Tab 鍵所執行的作業。
Ctrl+Tab	在對焦點位於管理程式窗格或專案窗格的情況下，將對焦點移至串流畫布。在對焦點位於節點選用區上時，此快捷鍵用於在節點與其選用區標籤之間移動對焦點。
任何英文字母鍵	當對焦點位於目前串流中的某個節點上時，此快速鍵會將對焦點和週期移至其名稱以所按鍵開頭的下一個節點。
F1	開啟說明系統並轉到與對焦點相關的主題。
F2	對畫布中選取的節點啟動連線過程。使用 Tab 鍵可以移至畫布上的所需節點，然後按 Shift+空格鍵完成連線。
F3	刪除畫布中所選節點的所有連線。
F6	在管理程式窗格、專案窗格和節點選用區之間移動對焦點。
F10	開啟「檔案」功能表。
Shift+F10	開啟節點或串流的蹦現功能表。
刪除	從畫布中刪除所選節點。
Esc	關閉蹦現功能表或對話框。
Ctrl+Alt+X	展開 SuperNode。
Ctrl+Alt+Z	放大 SuperNode。
Ctrl+Alt+Shift+Z	縮小 SuperNode。
Ctrl+E	當對焦點位於串流畫布中時，此快捷鍵用於執行目前串流。

IBM SPSS Modeler 中還使用了一些標準快速鍵，例如使用 Ctrl+C 進行複製。請參閱第 19 頁的『使用快速鍵』主題，以取得更多資訊。

## 對話框和表格的快捷鍵

對對話框、表格和對話框中的表格進行操作時，一些快捷鍵和螢幕閱讀器鍵非常有用。特殊鍵盤快捷鍵和螢幕閱讀器快捷鍵的完整清單如下所示。

表 40. 對話框及表示式建置器捷徑.

捷徑按鍵	函數
Alt+4	用於關閉所有已開啟的對話框或輸出視窗。可以從管理程式窗格的「輸出」標籤中擷取輸出。
Ctrl+End	在對焦點位於表示式建置器中的任何控制項上時，此快捷鍵會將插入項目點移到表示式的末尾處。
Ctrl+1	在表示式建置器中，用於將對焦點移至表示式編輯控制項。
Ctrl+2	在表示式建置器中，用於將對焦點移至函數清單。
Ctrl+3	在表示式建置器中，用於將對焦點移至欄位清單。

## 表格捷徑

表格快捷鍵用於節點（例如，「類型」、「過濾」和「合併」節點）的輸出表格和對話框中的表格控制項。通常，您將使用 Tab 鍵在表格資料格之間移動，並使用 Ctrl+Tab 離開表格控制項。註：有時候，螢幕閱讀器可能不會立即開始閱讀 Cell 內容。按一次或兩次方向鍵將重設軟體並啟動語音。

表 41. 表格捷徑.

捷徑按鍵	函數
Ctrl+W	對於表格，閱讀所選列的簡短說明。例如，「所選的第 2 列的值為性別、旗標、m/f 等。」
Ctrl+Alt+W	對於表格，閱讀所選列的詳細說明。例如，「所選的第 2 列的值為欄位 = 性別、類型 = 旗標、性別 = m/f 等。」
Ctrl+D	對於表格，閱讀所選區域的簡短描述。例如，「所選區域為一列六欄。」
Ctrl+Alt+D	對於表格，提供所選區域的詳細描述。例如，「所選區域為一列六欄。所選欄為欄位、類型和遺漏值。所選列為 1。」
Ctrl+T	對於表格，提供所選欄的簡短說明。例如，「欄位、類型和遺漏值。」
Ctrl+Alt+T	對於表格，提供所選欄的詳細說明。例如，「所選欄為欄位、類型和遺漏值。」
Ctrl+R	對於表格，提供表格中的記錄數。
Ctrl+Alt+R	對於表格，提供表格中的記錄數及直欄名稱。
Ctrl+I	對於表格，閱讀具有對焦點的 Cell 的資訊或內容。
Ctrl+Alt+I	對於表格，閱讀具有對焦點的 Cell 的資訊（直欄名稱及 Cell 內容）的詳細說明。
Ctrl+G	對於表格，提供簡短的基本選擇資訊。
Ctrl+Alt+G	對於表格，提供詳細的基本選擇資訊。
Ctrl+Q	對於表格，提供表格資料格之間快速切換。當您使用方向鍵在表格中移動時，按 Ctrl+Q 將閱讀詳細說明，例如「性別 = 女性」。再次按 Ctrl+Q 將切換到簡短說明（儲存格內容）。
F8	對於表格，當對焦點在表格中時，將對焦點設定到欄標題。
空白鍵	對於表格，當對焦點在欄標題上時，啟用欄排序。

## 備註的快捷鍵

處理畫面上的備註時，您可以使用下列快捷鍵。

表 42. 註解捷徑

捷徑按鍵	函數
Alt+C	切換顯示/隱藏備註功能。
Alt+M	插入新備註（如果目前顯示了備註）；顯示備註（如果備註目前處於隱藏狀態）。
Tab	在串流畫布上，循環切換現行串流中的所有來源節點及註解。
Enter	備註具有對焦點時，指示編輯起點位置。
Alt+Enter 或 Ctrl+Tab	結束編輯並儲存編輯變更。
Esc	取消編輯。在編輯期間所作的變更將丟失。
Alt+Shift+向上方向鍵	將文字區域的高度縮小一個網格 Cell（或一個像素），前提是對齊網格功能處於開啟（或關閉）狀態。
Alt+Shift+向下方向鍵	將文字區域的高度增大一個網格 Cell（或一個像素），前提是對齊網格功能處於開啟（或關閉）狀態。
Alt+Shift+向左方向鍵	將文字區域的寬度縮小一個網格 Cell（或一個像素），前提是對齊網格功能處於開啟（或關閉）狀態。
Alt+Shift+向右方向鍵	將文字區域的寬度增大一個網格 Cell（或一個像素），前提是對齊網格功能處於開啟（或關閉）狀態。

## 叢集檢視器和模型檢視器的快速鍵

可以使用快速鍵在叢集檢視器和模型檢視器視窗中進行瀏覽。

表 43. 一般捷徑 - 叢集檢視器及模型檢視器

捷徑按鍵	函數
Tab	將對焦點移到下一個畫面控制項。
Shift+Tab	將對焦點移到上一個畫面控制項。
向下方向鍵	如果下拉清單具有對焦點，那麼將開啟該清單或者移至該清單中的下一項目。 如果功能表具有對焦點，那麼將移至該功能表中的下一項目。 如果縮圖具有對焦點，那麼將移至集合中的下一項（如果最後一個縮圖具有對焦點，那麼將移至第一項）。
向上方向鍵	如果已開啟下拉清單，那麼此鍵用於移至該清單中的上一項目。 如果功能表具有對焦點，那麼此鍵用於移至該功能表中的上一項目。 如果縮圖具有對焦點，那麼此鍵用於移至集合中的下一項（如果第一個縮圖具有對焦點，那麼此鍵用於移至最後一欄一項）。
Enter	關閉已開啟的下拉清單，或者在開啟的功能表中進行選擇。
F6	在視窗的左右窗格之間切換對焦點。
左右方向鍵	如果標籤具有對焦點，那麼此鍵用於移至上一個或下一個標籤。 如果功能表具有對焦點，那麼此鍵用於移至上一個或下一個功能表。
Alt+字母	選取名稱中包含這個帶底線字母的按鈕或功能表。
Esc	關閉已開啟的功能表或下拉清單。

僅限「叢集檢視器」



叢集檢視器包含一個叢集視圖，該視圖包含叢集-特徵網絡。

要選擇該叢集視圖而不是「模型摘要」視圖：

1. 重複按 Tab 鍵直到已選取視圖按鈕。
2. 按向下方向鍵兩次以選中叢集。

您可以在此處選取網絡中的單個 Cell：

3. 重複按 Tab 鍵，直到您到達視覺化工具列中的最後一個圖示為止。



圖 17. 顯示「視覺化樹狀結構」圖示

4. 再按一次 Tab 鍵，然後按空格鍵，再按方向鍵。

現在，下列鍵盤快速鍵可用：

表 44. 叢集檢視器捷徑

捷徑按鍵	函數
方向鍵	在網絡中的各個單元之間移動對焦點。右側窗格中的 Cell 分佈顯示將在對焦點移動時發生變化。
Ctrl+, (逗點)	選取或取消選取網絡中 Cell 具有對焦點的整欄。 要將某欄新增到選擇中，請使用方向鍵瀏覽到該欄中的某個 Cell，然後再次按 Ctrl+,。
Tab	使對焦點移出網絡並移至下一個畫面控制項。
Shift+Tab	使對焦點移出網絡並移回到上一個畫面控制項。
F2	進入編輯模式（僅限於標籤和說明單元）。
Enter	儲存編輯變更並結束編輯模式（僅限於標籤和說明單元）。
Esc	結束編輯模式而不儲存變更（僅限於標籤和說明單元）。

## 快速鍵範例：建立串流

為了讓依賴鍵盤或螢幕閱讀器的使用者更清晰地瞭解串流建置程序，下面提供了一個不使用滑鼠建立串流的範例。在此範例中，您將通過下列步驟，建立一個包含變數檔案節點、衍生節點和直方圖節點的串流：

1. 啟動 **IBM SPSS Modeler**。第一次啟動 IBM SPSS Modeler 時，對焦點位於節點選用區「我的最愛」標籤上。
2. 按 **Ctrl+下方向鍵**。將對焦點從該標籤移至標籤主體。
3. 按**右方向鍵**。將對焦點移至「變數檔案」節點。
4. **空格鍵**。選取「變數檔案」節點。
5. 按 **Ctrl+Enter** 鍵。將「變數檔案」節點新增至串流畫布。此組合鍵還會保持選中「變數檔案」節點，以便新增的下一個節點將連接至該節點。
6. **Tab**。將焦點向後移到節點選用區。
7. 按**向右箭頭鍵 4 次**。移至「衍生」節點。
8. **空格鍵**。選取「衍生」節點。
9. 按 **Alt+Enter** 鍵。將該「衍生」節點新增至畫布，並將所選內容移至該「衍生」節點。現在，此節點可以連接至下一個所新增的節點。

10. **Tab**。將焦點向後移到節點選用區。
11. **按向右箭頭鍵 5 次**。將對焦點移至選用區中的「直方圖」節點。
12. **空格鍵**。選取「直方圖」節點。
13. **輸入**。將該節點新增到串流，並將對焦點移至串流畫布。

繼續完成下一個範例，或者如果您希望稍後再嘗試下一個範例，請儲存串流。

### 快速鍵範例：編輯節點

在此範例中，您將使用上一範例中建立的串流。該串流包含一個「變數檔案」節點、一個「衍生」節點和一個「直方圖」節點。在指示啟動時，焦點位於串流中第三個節點（即，「直方圖」節點）上。

1. **按 Ctrl+左方向鍵 2 次**。將對焦點移回「變數檔案」節點。
2. **輸入**。開啟「變數檔案」對話框。一直按 Tab 鍵以轉到「檔案」欄位，然後輸入文字檔路徑和名稱以選取該檔案。按 Ctrl+Tab 以瀏覽到對話框的下半部分，並一直按 Tab 鍵導覽至「確定」按鈕，然後按 Enter 鍵關閉對話框。
3. **Ctrl+右方向鍵**。將對焦點移至第二個節點（即，「衍生」節點）。
4. **輸入**。開啟「衍生節點」對話框。一直按 Tab 鍵以選取欄位，並指定衍生條件。按 Ctrl+Tab 以瀏覽到「確定」按鈕，然後按 Enter 鍵關閉對話框。
5. **Ctrl+右方向鍵**。將對焦點移動到第三個節點（即，「直方圖」節點）。
6. **輸入**。開啟「直方圖節點」對話框。一直按住 Tab 鍵以選取欄位，並指定圖表選項。對於下拉清單，請按向下方向鍵以開啟該清單並強調顯示某個清單項目，然後按 Enter 鍵以選取該清單項目。一直按 Tab 鍵以瀏覽到「確定」按鈕，然後按 Enter 鍵關閉對話框。

此時，您可以新增更多節點或執行目前串流。建立串流時，請記住下列提示：

- 手動連接節點時，使用 F2 建立連線起點，並按 Tab 鍵移至終點，然後使用 Shift+空格鍵完成連線。
- 使用 F3 鍵可切斷畫布中所選節點的所有連線。
- 建立串流之後，使用 Ctrl+E 可以執行目前串流。

提供了一個完整的快速鍵清單。請參閱第 182 頁的『用於瀏覽主視窗的快速鍵』主題，以取得更多資訊。

### 使用螢幕閱讀器

市場上有各種各樣的螢幕閱讀器。IBM SPSS Modeler 配置為支援 JAWS for Windows，後者使用隨 IBM SPSS Modeler 一起安裝的 Java Access Bridge。如果已安裝 JAWS，那麼只需在啟動 IBM SPSS Modeler 之前啟動 JAWS 即可使用此產品。

註：我們建議您至少留出 6GB 的空間，以便將 JAWS 與 SPSS Modeler 一起執行。

由於 IBM SPSS Modeler 具備獨特的以圖形方式代表資料採礦過程的特徵，因此可通過最佳可視化方式使用圖表和圖形。但是，您可以根據使用螢幕閱讀器以文字方式檢視的輸出和模型來瞭解並作出決策。

註：對於 64 位元用戶端機器，某些輔助技術功能無效。這是因為 Java Access Bridge 並非針對 64 位元作業設計。

使用 IBM SPSS Modeler 字典檔案

隨 JAWS 提供了可用的 IBM SPSS Modeler 字典檔案 (*Awt.JDF*)。要使用此檔案，請完成下列步驟：

1. 瀏覽至 IBM SPSS Modeler 安裝版本的 */accessibility* 子目錄，然後複製該字典檔 (*Awt.JDF*)。
2. 將其複製到包含 JAWS Script 的目錄。

如果其他 Java 應用程式在執行中，那麼您的機器可能已包含名為 *Awt.JDF* 的檔案。在這種情況下，您可能無法使用此字典檔，而必須手動編輯該字典檔。

## 將螢幕閱讀器用於 HTML 輸出

在 IBM SPSS Modeler 中使用螢幕閱讀器檢視顯示為 HTML 的輸出時，可能會遇到一些困難。受影響的一些輸出類型包含：

- 在迴歸、邏輯迴歸和因子/PCA 節點的「進階」標籤中檢視的輸出
- 報告節點輸出

在上述每個視窗或對話框中，工具列上提供了可用於在預設瀏覽器（提供標準螢幕閱讀器支援）中啟動輸出的工具。然後，您可以使用螢幕閱讀器傳達輸出資訊。

## 交互樹狀結構視窗中的協助工具

「互動式樹狀結構」視窗中決策樹的標準顯示可能會導致螢幕閱讀器的問題。要存取易存取版本，請在「互動式樹狀結構」功能表中按一下：

### 檢視 > 可存取視窗

此選項將顯示一個類似於標準樹狀對映圖的視圖，但該視圖可由 Jaws 正確閱讀。您可以使用標準的方向鍵向上、向下、向左、向右移動。當您瀏覽可存取視窗時，「互動式樹狀結構」視窗中的對焦點將相應地移動。使用空格鍵可以變更選擇，使用 Alt+/ 可以延伸目前選擇。

---

## 使用提示

一些提示可幫助您更方便地存取 IBM SPSS Modeler 環境。下列是在 IBM SPSS Modeler 中工作時的一般提示。

- **結束展開的文字框。**使用 Ctrl+Tab 可結束展開的文字框。請注意，Ctrl+Tab 也用來結束表格控制項。
- **使用 Tab 鍵而不是方向鍵。**選取對話框的選項時，請使用 Tab 鍵在選項按鈕之間進行移動。方向鍵在該環境定義中不起作用。
- **下拉清單。**在對話框的下拉清單中，您可以使用 Esc 鍵或空格鍵選取項目，然後關閉下拉清單。也可以使用 Esc 鍵關閉在跳轉到另一個控制項時未關閉的下拉清單。
- **執行狀態。**對大型資料庫執行串流時，JAWS 可能會在閱讀串流狀態時出現滯後現象。定時按 Ctrl 鍵可更新狀態報告。
- **使用節點選用區。**第一次進入節點選用區的標籤時，JAWS 有時會閱讀「群組框」，而不是該節點的名稱。在這種情況下，可以按 Ctrl+向右方向鍵，然後按 Ctrl+向左方向鍵以重設螢幕閱讀器並收聽節點名稱。
- **閱讀功能表。**有時，在您第一次開啟某個功能表時，JAWS 可能不會閱讀第一個功能表項目。如果您懷疑發生了這種情況，請使用向下方向鍵，然後使用向上方向鍵收聽功能表中的第一個功能表項目。
- **階層式功能表。**JAWS 不會閱讀階層式功能表的第一級功能表。如果您在功能表中移動時聽到發音中斷，請按向右方向鍵收聽子功能表項目。

另外，如果已安裝 IBM SPSS Modeler Text Analytics，那麼下列提示可使互動式工作台介面更易於使用。

- **進入對話框。**進入對話框時，您可能需要按 Tab 鍵以便將對焦點置於第一個控制項上。
- **結束展開的文字框。**使用 Ctrl+Tab 可結束展開的文字框並移至下一個控制項。請注意，Ctrl+Tab 也用來結束表格控制項。
- **在樹狀結構清單中鍵入首字母以尋找元素。**在「種類」窗格、「已擷取的結果」窗格或程式庫樹狀結構中查找某個元素時，您可以在對焦點位於此窗格上的情況下鍵入該元素的首字母。此操作將選中下一個出現的以您所已輸入字母開頭的元素。

- **下拉清單。** 在對話框的下拉清單中，您可以使用空格鍵選取項目，然後關閉該清單。

## **與其他軟體的衝突**

使用螢幕閱讀器（例如，JAWS）正在測試 IBM SPSS Modeler 時，我們的開發小組探索在貴組織內使用 Systems Management Server(SMS) 可能與 JAWS 的功能發生衝突，而無法閱讀基於 Java 的應用程式，例如 IBM SPSS Modeler。停用 SMS 將更正此狀況。請造訪 Microsoft 網站以獲取有關 SMS 的進一步資訊。

## **JAWS 和 Java**

不同版本的 JAWS 將為基於 Java 的軟體應用程式提供不同層次的支援。雖然 IBM SPSS Modeler 可以使用所有 JAWS 的最新版本，但某些版本在應用到基於 Java 的系統時還是會出現一些小問題。請造訪 JAWS for Windows 的網站 <http://www.FreedomScientific.com>。

## **在 IBM SPSS Modeler 中使用圖形**

資訊的可視化顯示（例如，直方圖、評估圖表、多重繪圖和散佈圖）很難使用螢幕閱讀器進行說明。但是請注意，可以使用輸出視窗中提供的文字摘要檢視 Web 圖形和分佈。

---

## 第 16 章 Unicode 支援。

---

### IBM SPSS Modeler 中的 Unicode 支援

IBM SPSS Modeler 10.0 IBM SPSS Modeler 和 IBM SPSS Modeler Server 均完全支持 Unicode。這樣就可以與其他支援 Unicode 的應用程式（包含多語言資料庫）交換資料，同時不會由於針對特定語言環境的編碼轉換而導致資訊丟失。

- IBM SPSS Modeler 將 Unicode 資料儲存庫在內部，並且可以讀取寫入資料庫中以 Unicode 儲存的多語言資料而不會發生資料丟失。
- IBM SPSS Modeler 可以讀取寫入用 UTF-8 編碼的文字檔。文字檔的匯入並匯出將預設使用語言環境編碼，但也支援 UTF-8 作為替代編碼。此設定可以在檔案匯入並匯出節點中指定，也可以在串流內容對話框中變更預設編碼設定。請參閱第 37 頁的『設定串流的一般選項』主題，以取得更多資訊。
- 以語言環境編碼儲存的 Statistics、SAS 和文字資料檔案在匯入時將轉換為 UTF-8 格式，匯出時又轉換為原來的格式。寫入任何檔案時，如果存在語言環境字集中不存在的 Unicode 字元，則這些字元將被替代並顯示警告資訊。僅當從支援 Unicode 並且包含來自另一語言環境或來自多個語言環境或字集的字元的資料來源（資料庫或 UTF-8 文字檔）匯入資料時，才會發生此種情況。
- IBM SPSS Modeler Solution Publisher 影像使用 UTF-8 編碼，且真正實現了平台間和語言環境間的可移植。

#### 關於 Unicode

Unicode 標準的目標是提供一種多語言文字編碼的統一方法，以便可以輕鬆地跨地界、語言環境和應用程式共用文字。Unicode 標準現已發展到版本 4.0.1，它定義了一個作為當今世界上所有一般字集的超級集合的字集，並為每個字元分配一個唯一名稱和碼點。這些字元及其碼點與 ISO-10646 定義的廣用字集 (UCS) 中的字元及其碼點完全相同。詳細資訊請參閱 Unicode 首頁。



---

## 注意事項

本資訊係針對 IBM 在美國所提供之產品與服務所開發。IBM 可能會提供此資料的其他語言版本。不過，您可以要求擁有一份該語言的產品或產品版本以取用它。

在其他國家，IBM 不見得有提供本文件所提及之各項產品、服務或功能。請洽詢當地的 IBM 業務代表，以取得當地目前提供的產品和服務之相關資訊。本文件在提及 IBM 的產品、程式或服務時，不表示或暗示只能使用 IBM 的產品、程式或服務。只要未侵犯 IBM 之智慧財產權，任何功能相當之產品、程式或服務皆可取代 IBM 之產品、程式或服務。不過，任何非 IBM 之產品、程式或服務，使用者必須自行負責作業之評估和驗證責任。

本文件所說明之主題內容，IBM 可能擁有其專利或專利申請案。提供本文件不代表提供這些專利的授權。您可以書面提出授權查詢，來函請寄到：

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785US*

如果是有關雙位元組 (DBCS) 資訊的授權查詢，請洽詢所在國的 IBM 智慧財產部門，或書面提出授權查詢，來函請寄到：

*Intellectual Property Licensing  
Legal and Intellectual Property Law  
IBM Japan Ltd.  
19-21, Nihonbashi-Hakozakicho, Chuo-ku  
Tokyo 103-8510, Japan*

International Business Machines Corporation 只依「現況」提供本出版品，不提供任何明示或默示之保證，其中包括且不限於不侵權、可商用性或特定目的之適用性的隱含保證。有些轄區在特定交易上，不允許排除明示或暗示的保證，因此，這項聲明不一定適合您。

本資訊中可能會有技術上或排版印刷上的訛誤。因此，IBM 會定期修訂；並將修訂後的內容納入新版中。IBM 隨時會改進及/或變更本出版品所提及的產品及/或程式，不另行通知。

本資訊中任何對非 IBM 網站的敘述僅供參考，IBM 對該網站並不提供任何保證。這些網站所提供的資料不是 IBM 本產品的資料內容，如果要使用這些網站的資料，您必須自行承擔風險。

IBM 得以各種 IBM 認為適當的方式使用或散布 貴客戶提供的任何資訊，而無需對 貴客戶負責。

如果本程式之獲授權人為了 (i) 在個別建立的程式和其他程式（包括本程式）之間交換資訊，以及 (ii) 相互使用所交換的資訊，因而需要相關的資訊，請洽詢：

*IBM Director of Licensing  
IBM Corporation  
North Castle Drive, MD-NC119  
Armonk, NY 10504-1785US*

上述資料之取得有其特殊要件，在某些情況下必須付費方得使用。

IBM 基於 IBM 客戶合約、IBM 國際程式授權合約或雙方之任何同等合約的條款，提供本文件所提及的授權程式與其所有適用的授權資料。

所引用的效能資料和客戶範例僅為說明用途呈現。實際效能結果可能會依據特定配置和作業條件而有所不同。

本文件所提及之非 IBM 產品資訊，係取自產品供應商，或其發佈的聲明或其他公開管道。IBM 並未測試過這些產品，也無法確認這些非 IBM 產品的執行效能、相容性或任何對產品的其他主張是否完全無誤。如果您對非 IBM 產品的性能有任何的疑問，請逕向該產品的供應商查詢。

關於 IBM 未來方針或目的之聲明，隨時可能更改或撤銷，不必另行通知，且僅代表目標與主旨。

本資訊中含有日常商業活動所用的資料及報告範例。為了盡可能地加以完整說明，範例中含有個人、公司、品牌及產品名稱。所有這些名稱全為虛構，任何與實際個人或商場企業類似之處，純屬巧合。

---

## 商標

IBM、IBM 標誌及 [ibm.com](http://ibm.com) 是 International Business Machines Corp. 在世界許多管轄區註冊的商標或註冊商標。其他產品及服務名稱可能是 IBM 或其他公司的商標。IBM 商標的最新清單可在 Web 的 "Copyright and trademark information" 中找到，網址為 [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)。

Adobe、Adobe 標誌、PostScript 及 PostScript 標誌是 Adobe Systems Incorporated 在美國及（或）其他國家或地區的註冊商標或商標。

Intel、Intel 標誌、Intel Inside、Intel Inside 標誌、Intel Centrino、Intel Centrino 標誌、Celeron、Intel Xeon、Intel SpeedStep、Itanium 及 Pentium 是 Intel Corporation 或其子公司在美國及其他國家或地區的商標或註冊商標。

Linux 是 Linus Torvalds 在美國及（或）其他國家或地區的註冊商標。

Microsoft、Windows、Windows NT 及 Windows 標誌是 Microsoft Corporation 在美國及/或其他國家或地區的商標。

UNIX 是 The Open Group 在美國及其他國家或地區的註冊商標。

Java 和所有以 Java 為基礎的商標及標誌是 Oracle 及（或）其子公司的商標或註冊商標。

---

## 產品說明文件的條款

這些出版品的使用授權受到下列條款的限制。

### 適用性

這些條款係附加於 IBM 網站的任何使用條款上。

### 個人使用

貴客戶可以為了非商務性的私人用途而複製這些出版品，但必須保留所有專利注意事項。如果沒有 IBM 明文同意，貴客戶不能散佈、顯示或衍生這些出版品或其中的任何部分。

### 商業使用

貴客戶只能在貴企業內重製、散布和顯示這些出版品，但必須保留所有專利聲明。未經 IBM 明確同意，不得在您的企業外散布、展示或衍生這些出版品或其任何部分。



## 權利

除了本項許可權所明確授予者之外，並未明示或暗示授予出版品或任何資訊、資料、軟體或其中的其他智慧財產的任何其他許可權、授權或權利。

IBM 保留在判定出版品的使用將損害其利益或判定未適當遵守上述指示時，撤銷此處所授予之許可權的權利。

除非完全符合所有適用的法規（包括所有美國輸出法規），否則 貴客戶不能下載、輸出或重複輸出這項資訊。

IBM 不提供這些出版品內容的任何保證。這些出版品是依「現狀」提供，不含任何明示或默示之保證（包括但不限於可售性、未涉侵權及符合特定效用的保證）。



# 索引

索引順序以中文字，英文字，及特殊符號之次序排列。

## 〔一劃〕

一般應用程式 21

## 〔三劃〕

三角函數 119

大小 61

在概要中 61

大於運算子 116

子選用區

在選用區標籤上顯示 175

建立 175

從選用區標籤中刪除 175

小數位數

顯示格式 39

小數點符號

數字顯示格式 37

工具列 15

工具提示

標註節點 53

已產生的模型選用區 13

## 〔四劃〕

不等運算子 116

中間滑鼠按鈕

模擬 19, 32

互動式輸出 64

內容 79

表格 79

專案資料夾 160

報告階段 161

資料串流 36

樞紐表 79

分支、建模及評分 48, 150, 152

分佈函數 120

分組列或行 75

分組符號

數字顯示格式 37

分組節點

效能 178

分割表格 84

控制表格分頁 84

少於運算子 116

文件 3

文字 62, 65, 69

文字 (繼續)

將輸出匯成文字 65, 69

新增文字檔到檢視器 62

新增至檢視器 62

文字資料檔

編碼 191

文字編碼 37

日期

使用 129

轉換 129

日期函數 111, 112

date\_before 116, 126

date\_days\_difference 126

date\_in\_days 126

date\_in\_months 126

date\_in\_weeks 126

date\_in\_years 126

date\_months\_difference 126

date\_weeks\_difference 126

date\_years\_difference 126

@TODAY 函數 126

日期時間函數

datetime\_date 126

datetime\_day 126

datetime\_day\_name 126

datetime\_day\_short\_name 126

datetime\_hour 126

datetime\_in\_seconds 126

datetime\_minute 126

datetime\_month 126

datetime\_month\_name 126

datetime\_month\_short\_name 126

datetime\_now datetime\_second 126

datetime\_time 126

datetime\_timestamp 126

datetime\_weekday 126

datetime\_year 126

日期格式 38, 111, 112

日期/時間值 99

日誌檔

顯示產生的 SQL 40

比較函數 116

片段 48

已定義 13

## 〔五劃〕

主視窗 12

主機名稱

IBM SPSS Modeler Server 8, 9

功能選擇節點

遺漏值 90

卡方測試

機率函數 120

可視化程式設計 11

句號 37

必要欄位 56, 57

必備欄位 57

用戶端

預設目錄 166

目錄

預設值 166

## 〔六劃〕

交互樹狀結構視窗

協助工具 189

列 84

在樞紐表中選取 84

列出串流的所有備註 52

列印 20, 71, 72, 79, 81, 84

文字輸出 71

串流 18, 34

頁首/頁底 71

控制表格分頁 84

階層 71, 79, 81

預覽列印 71

圖表 71

圖表大小 72

樞紐表 71

編頁碼 72

調整表格 79, 81

輸出項目間之空間 72

合併節點

效能 178

因子 189

地理空間座標

選取坐標系 42

顯示格式 41

地理空間座標系統選擇 42

地理空間座標格式 41

在 CLEM 表示式中的反斜線字元 110

在 IBM SPSS Collaboration and Deployment Services 儲存庫中搜尋物件 144

在 IBM SPSS Collaboration and

Deployment Services 儲存庫中儲存物件 139

多二分集

在 CLEM 表示式中 101

- 多回應集
  - 在 CLEM 表示式中 105
- 多重回應集
  - 在 CLEM 表示式中 101
- 多重開啟資料檔 86
- 多個 IBM SPSS Modeler 階段作業 11
- 多種類集
  - 在 CLEM 表示式中 101
- 字元 109, 110
- 字串 109, 110
  - 比對 98
  - 在 CLEM 表示式中操作 98
  - 取代 98
- 字串函數 122
- 字典檔 188
- 字型 61, 81, 167, 168
  - (大綱窗格中) 61
- 自訂選用區建立 174
  - 子選用區建立 175
- 自動化 95
- 色板 12
  - 自訂 174

## (七劃)

- 串流 7, 12, 148
  - 地理空間座標系統 42
  - 更名 47, 53
  - 建置 31
  - 停用節點 33
  - 執行 48
  - 連接節點 32
  - 部署選項 149
  - 備份檔 54
  - 新增至專案 157, 158
  - 新增備註 49
  - 新增節點 32, 33
  - 跳過節點 33
  - 載入 55
  - 標註 49, 53
  - 調整大小以檢視 18
  - 選項 36, 37, 38, 39, 40, 41
  - 儲存 54
  - 儲存在 IBM SPSS 協同作業及部署服務儲存庫中 141
  - 檢視執行時間 43
- 串流內容
  - Analytic Server 41
- 串流名稱 53
- 串流參數 44
- 串流畫布
  - 設定 41
- 串流預設編碼 37
- 串流說明 47, 48
- 位元函數 121

- 伺服器
  - 通過 COP 搜尋伺服器 9
  - 登入 8
  - 新增連線 9
  - 預設目錄 166
- 刪除群組標記 75
- 刪除輸出 60
- 呈示視圖時出錯
  - 記憶體不足 165
- 序列函數 129
- 快取
  - 啟用 165
  - 清除 34, 37
  - 設定快取 34
  - 儲存 34
- 快取檔節點
  - 載入 55
- 快速樞紐表 86
- 快速鍵 19
  - 一般用法 58
  - 鍵盤 19, 182, 184, 186
- 更名
  - 串流 47
  - 節點 53
- 更新模型 151
- 求反函數 118
- 決策樹
  - 協助工具 189
- 系統
  - 選項 165
- 系統遺漏值
  - 在記錄中 90
- 角度
  - 測量單位 39

## (八劃)

- 使用者 ID
  - IBM SPSS Modeler Server 8
- 使用者定義函數 (UDF) 102, 103
- 使用者選項 166
- 來源節點 31
  - 重新整理 37
  - 資料對映 56
- 函數 111, 112, 114, 115, 129
  - 在 CLEM 表示式中 102
  - 使用者定義函數 (UDF) 102
  - 處理遺漏值 93
  - 資料庫 102, 103
  - 範例 95
  - @BLANK 92
  - @FIELD 102, 134
  - @GLOBAL\_MAX 133
  - @GLOBAL\_MIN 133
  - @GLOBAL\_SDEV 133
  - @GLOBAL\_SUM 133

- 函數 (繼續)
  - @GLOBAL\_平均值 133
  - @PARTITION 134
  - @PREDICTED 102, 134
  - @TARGET 102, 134
- 協助工具 181, 190
  - 範例 187, 188
  - IBM SPSS Modeler 中的功能 181
  - IBM SPSS Modeler 中的提示 189
- 取代模型 166
- 取消欄位對映 56
- 取整函數 118
- 弧度
  - 測量單位 39
- 或運算子 118
- 版本標記, IBM SPSS Collaboration and Deployment Services Repository 物件 148
- 版面設定 71, 72
  - 頁首/頁底 71
  - 圖表大小 72
- 版面調整 60, 86
  - 輸出 60, 86
- 物件
  - 內容 160
- 物件內容, IBM SPSS Collaboration and Deployment Services Repository 147
- 狀態
  - 載入 55
  - 儲存 54
- 直欄 83, 84
  - 在樞紐表中選取 84
  - 變更樞紐表中的寬度 83
- 直欄寬度 79, 83, 86
  - 控制最大寬度 79
  - 控制換行文字寬度 79
  - 控制預設寬度 86
  - 樞紐表 83
- 知識探索 21
- 空白 89, 99
  - 從字串中移除 98, 122
- 空白值處理
  - CLEM 函數 134
- 空值 99
- 空間
  - 從字串中移除 98, 122
- 空間函數 120
- 表示式 109
- 表示式建置器 184
  - 存取 102
  - 使用 102
  - 概述 102
- 表格 84, 184
  - 字型 81
  - 背景顏色 81
  - 控制表格分頁 84

表格 (繼續)  
 新增至專案 157  
 對齊 81  
 儲存格內容 81  
 儲存輸出 54  
 邊緣 81  
 表格分頁 84  
 表格外觀 78  
 建立 78  
 套用 78  
 表格圖表 85  
 非運算子 118  
 冠軍挑戰者分析 137, 149  
 品質節點  
 遺漏值 90  
 垂直標記文字 75

## 〔九劃〕

建立規則節點  
 載入 55  
 建模  
 分支 48  
 建模節點 31  
 建模選用區標籤自訂 176  
 效能 179  
 待續文字 81  
 用於樞紐表 81  
 指令行  
 啟動 IBM SPSS Modeler 8  
 指數函數 118  
 為節點和串流命名 53  
 相等運算子 116  
 科學記號 86  
 不顯示於輸出 86  
 顯示格式 39  
 背景顏色 81  
 重新排列列和行 74  
 重新整理  
 來源節點 37  
 頁底 71  
 頁首 71

## 〔十劃〕

值 96  
 通過資料審核檢視 105  
 新增到 CLEM 表示式中 105  
 差異節點  
 效能 178  
 效能  
 建模節點的 179  
 處理節點的 178  
 CLEM 表示式 180  
 時間函數 111, 112

時間函數 (繼續)  
 time\_before 116, 126  
 time\_hours\_difference 126  
 time\_in\_hours 126  
 time\_in\_mins 126  
 time\_in\_secs 126  
 time\_mins\_difference 126  
 time\_secs\_difference 126  
 時間格式 38, 111, 112  
 時間與日期函數 111, 112  
 時間欄位  
 轉換 129  
 案例 21  
 特殊字元  
 從字串中移除 98  
 特殊函數 134  
 租戶  
 IBM SPSS Analytic Server 10  
 記憶體 86  
 堆疊溢位錯誤 165  
 管理 165  
 記憶體不足 165  
 記錄 21  
 系統遺漏值 90  
 遺漏值 90  
 訊息  
 顯示產生的 SQL 40  
 迴歸 189

## 〔十一劃〕

停止執行 15  
 停用節點 33, 34  
 剪下 15  
 區域圖  
 空間函數 120  
 參數  
 在 CLEM 表示式中 105  
 串流 44  
 執行時期提示 44  
 評分 150  
 階段作業 44  
 模型建置 150  
 類型 44  
 堆疊溢位錯誤 165  
 埠號  
 IBM SPSS Modeler Server 8, 9  
 基於樹狀結構的分析  
 一般應用程式 21  
 執行串流 48  
 執行時間, 檢視 43  
 密碼  
 IBM SPSS Analytic Server 10  
 IBM SPSS Modeler Server 8  
 專案 15, 157

專案 (繼續)  
 在 IBM SPSS Collaboration and  
 Deployment Services 儲存庫中 159  
 物件內容 160  
 建立新的項目 158  
 建置 158  
 產生報告 161  
 設定內容 159  
 設定預設資料夾 157  
 新增物件 158  
 資料夾內容 160  
 標註 160  
 儲存在 IBM SPSS 協同作業及部署服  
 務儲存庫中 141  
 關閉 161  
 「類別」視圖 157  
 CRISP-DM 視圖 157  
 將輸出貼入到其他應用程式 63  
 將輸出複製並貼上至其他應用程式 63  
 常態分配  
 機率函數 120  
 從 IBM SPSS Collaboration and  
 Deployment Services 儲存庫中擷取物件  
 143  
 控制要顯示的列數 79  
 捲動  
 設定選項 41  
 排序  
 樞紐表列 75  
 排序節點  
 效能 178  
 啟用節點 33  
 啟動對話框 167  
 旋轉標記 75  
 條件 97  
 清單 109, 110  
 移動列和行 74  
 符合 508 要求 181  
 符號函數 118  
 統計資料檔案  
 編碼 191  
 終端節點 31  
 處理程序協調器 9  
 處理程序節點 31  
 效能 178  
 規則集  
 評估 37  
 貨幣顯示格式 39  
 通知  
 設定選項 166  
 通過 COP 搜尋連線 9  
 逗點 37  
 連接字串 115

- 連線
  - 至 IBM SPSS Collaboration 和 Deployment Services Repository 138
  - 至 IBM SPSS Modeler Server 8, 9
  - 伺服器叢集 9
  - 到 IBM SPSS Analytic Server 10
- 部署 137
- 部署選項 149
- 部署類型 149

## (十二劃)

- 備份 (n.)串流檔案
  - 還原 54
- 最小化 17
- 單一登入 8
- 單一登入、IBM SPSS Collaboration 和 Deployment Services Repository 137, 138
- 報告
  - 產生 161
  - 設定內容 161
  - 新增至專案 157
  - 儲存輸出 54
- 尋找和取代
  - 檢視器文件 62
- 復原 15
- 插入群組標記 75
- 提示
  - 一般用法 58
  - 協助工具 189
- 提示，執行時期 44
- 換行 79
  - 控制換行文字的行寬 79
- 替代的列色彩
  - 樞紐表 80
- 測量系統 86
- 畫布 12
- 登入 IBM SPSS Modeler Server 8
- 註腳 80, 82, 83
  - 重新編號 83
  - 標記 80
- 註解
  - 列出串流中的所有註釋 52
  - 的註釋 49
  - 專案 160
  - 鍵盤捷徑 186
  - 轉換為備註 52
- 註釋
  - 資料夾 160
- 評分
  - 分支 48, 150, 152
- 評估節點
  - 效能 178
- 貼上 15

- 距離函數 120
- 開啟
  - 串流 55
  - 狀態 55
  - 專案 158
  - 節點 55
  - 模型 55
  - 輸出 55
- 階段作業參數 44
- 階層 71, 77, 79, 81
  - 列印 71, 79, 81
  - 在樞紐表上 77
  - 建立 77
  - 顯示 77
- 集 37

## (十三劃)

- 匯入
  - PMML 154
- 匯出
  - 串流說明 48
  - PMML 154
- 匯出圖表 65, 70
- 匯出輸出 65, 68, 69
  - Excel 格式 65, 67
  - HTML 66
  - HTML 格式 65
  - PDF 格式 65, 68
  - PowerPoint 格式 65
  - Web 報告 66
  - Word 格式 65, 67
- 填入節點
  - 遺漏值 92
- 搜尋
  - 串流中的節點 46
- 搜尋和取代
  - 檢視器文件 62
- 新功能 5
- 新增
  - 到專案 158
- 新增 IBM SPSS Modeler Server 連線 9
- 新增群組標記 75
- 會員函數 116
- 概要 61
  - 在檢視器中 61
  - 收合 61
  - 展開 61
  - 變更層級 61
- 滑鼠
  - 在 IBM SPSS Modeler 中使用 19, 32
- 節點 7
  - 在串流中取消 33
  - 在串流中連接 32
  - 在串流中繞過 33

- 節點 (繼續)
  - 在選用區上顯示 175
  - 自訂子選用區建立 175
  - 自訂選用區建立 174
  - 刪除 32
  - 刪除連線 34
  - 的順序 177
  - 效能 178, 179
  - 停用 33, 34
  - 執行時間 43
  - 將備註新增到 49
  - 從選用區中刪除 175
  - 啟用 33
  - 設定選項 34
  - 搜尋 46
  - 新增 32, 33
  - 新增至專案 157, 158
  - 資料預覽 36
  - 載入 55
  - 預覽資料 36
  - 標註 49, 53
  - 編輯 34
  - 複製 34
  - 選用區標籤自訂作業 176
  - 儲存 54
  - 儲存在 IBM SPSS 協同作業及部署服務儲存庫中 142
  - 簡介 31
  - 鎖定 36
- 節點名稱 53
- 節點選用區選擇 175
- 群組標記 75
- 解除鎖定 IBM SPSS Collaboration 和 Deployment Services Repository 物件 145
- 資料
  - 預覽 36
- 資料串流
  - 建置 31
- 資料夾、IBM SPSS Collaboration 和 Deployment Services Repository 145, 146
- 資料庫
  - 函數 102, 103
- 資料庫函數
  - 在 CLEM 表示式中 103
  - 使用者定義函數 (UDF) 103
- 資料採礦 21
  - 策略 23
  - 應用程式範例 29
- 資料採礦的應用程式 22
- 資料對映工具 56
- 資料審核節點
  - 用於探索 21
  - 用於資料採礦 22

- 資料編輯器
  - 多重開啟資料檔 86
  - 描述性統計量選項 87
- 資料檔
  - 多重開啟資料檔 86
- 資料類型 96
  - 在參數中 44
- 資訊函數 115
- 載入
  - 狀態 55
  - 節點 55
- 運算子
  - 在 CLEM 表示式中 102
  - 結合字串 115
- 運算子優先級 112
- 預設
  - 專案階段 157
- 預覽
  - 節點資料 36

## (十四劃)

- 圖
  - 新增至專案 157
- 圖形
  - 儲存輸出 54
- 圖表 59, 65, 85
  - 從樞紐表建立 85
  - 匯出 65
  - 儲存輸出 54
  - 隱藏 59
- 圖像
  - 設定選項 18, 41
- 實數 109
- 對中輸出 60, 86
- 對映資料 57
- 對映欄位 56
- 對話框 86
  - 變數顯示順序 86
  - 顯示變數名稱 86
  - 顯示變數標籤 86
- 對齊 60, 86
  - 輸出 60, 86
- 對數函數 118
- 慣例 114
- 管理程式 13
- 網格線 84
  - 樞紐表 84
- 網域名稱 (Windows)
  - IBM SPSS Modeler Server 8
- 聚集節點
  - 效能 178
- 與運算子 118
- 製作公用程式 86
  - 從日誌登載檔使用指令語法 86

- 語言
  - 選項 165
  - 變更輸出語言 76
- 語言環境
  - 選項 165
- 認證
  - 用於 IBM SPSS Collaboration and Deployment Services Repository 138

## (十五劃)

- 寬表格
  - 貼入 Microsoft Word 63
- 廣域功能 133
- 廣域值
  - 在 CLEM 表示式中 105
- 數字 99, 109
- 數字顯示格式 39
- 數值型函數 118
- 數值標籤 76, 87
  - 在樞紐表上 87
  - 數值標籤 87
- 暫存目錄 10
- 樞紐表 59, 63, 65, 71, 74, 75, 76, 77, 78, 79, 80, 81, 83, 84, 85, 86
  - 一般內容 79
  - 內容 79
  - 分組列或行 75
  - 列印大型表格 84
  - 列印階層 71
  - 字型 81
  - 刪除群組標記 75
  - 呈現表格更快速 86
  - 快速樞紐表 86
  - 使用 74
  - 使用圖示 74
  - 待續文字 81
  - 背景顏色 81
  - 格寬 83
  - 從表格建立圖表 85
  - 控制表格分頁 84
  - 控制要顯示的列數 79
  - 排序列 75
  - 旋轉標記 75
  - 移動列和行 74
  - 復原變更 76
  - 插入列和欄 75
  - 插入群組標記 75
  - 替代的列色彩 80
  - 註腳 82, 83
  - 註腳內容 80
  - 貼入到其他應用程式中 63
  - 貼成表格 63
  - 階層 77
  - 匯出成 HTML 65

- 樞紐表 (繼續)
  - 新表格的預設設定 86
  - 解散列或行的群組 75
  - 預設的行寬調整 86
  - 對齊 81
  - 網格線 84
  - 語言 76
  - 數值標籤 76
  - 樞紐表 74
  - 標題 82
  - 編輯 74
  - 調整以配合頁面 79, 81
  - 選取列與行 84
  - 儲存格內容 81
  - 儲存格格式 80
  - 隱藏 59
  - 舊版表格 85
  - 轉置列與行 74
  - 邊框 80
  - 邊緣 81
  - 變更外觀 78
  - 變更顯示次序 74
  - 變數標籤 76
  - 顯示和隱藏儲存格 77
  - 顯示隱藏的邊框 84
- 樞紐表中的儲存格 77, 80, 83
  - 格式 80
  - 寬度 83
  - 選擇 84
  - 隱藏 77
  - 顯示 77
- 樞紐表中的顏色 80
  - 邊框 80
- 標記 75
  - 刪除 75
  - 插入群組標記 75
  - 顯示 37
- 標註
  - 串流 49, 53
  - 節點 49, 53
- 標題 62, 82
  - 新增至檢視器 62
- 標籤
  - 值 154
  - 變數 154
- 標籤, IBM SPSS Collaboration and Deployment Services Repository 物件 148
- 模型 48
  - 更新 151
  - 取代 166
  - 匯出 168
  - 新增至專案 157
  - 儲存在 IBM SPSS 協同作業及部署服務儲存庫中 142
  - 模型更新 149

- 模型塊 48
- 模型選用區 142
- 範本 56
- 範本欄位 57
- 範例
  - 概述 4
  - 應用程式手冊 3
- 編頁碼 72
- 編碼 37, 191
- 線性迴歸
  - 匯出為 PMML 168
- 複製 15
- 調整大小 17
- 調整大小串流以檢視 18

## 〔十六劃〕

- 幕次 (指數) 函數 118
- 導覽
  - 鍵盤捷徑 182
- 整數 109
- 樹狀結構型分析
  - 一般應用程式 21
- 機率函數 120
- 機器學習 21
- 螢幕閱讀軟體 182, 184, 188, 189
  - 範例 187, 188
- 輸出 13, 59, 60, 63, 65, 72, 86
  - 互動式 64
  - 加密 72
  - 刪除 60
  - 移動 60
  - 貼入到其他應用程式中 63
  - 匯出 65
  - 對中 60, 86
  - 對齊 60, 86
  - 複製 60
  - 儲存 72
  - 檢視器 59
  - 隱藏 59
  - 變更輸出語言 76
  - 顯示 59
- 輸出物件
  - 儲存在 IBM SPSS 協同作業及部署服務儲存庫中 142
- 輸出節點 31
- 輸出檔
  - 儲存 54
- 選用區標籤自訂作業 176
- 選取方法 84
  - 選取樞紐表中的列與行 84
- 選項 86, 87, 165
  - 一般 86
  - 串流內容 36, 37, 38, 39, 40, 41, 42, 43
  - 使用者 166

- 選項 (繼續)
  - 「資料編輯器」中的描述性統計量 87
  - 語法 168
  - 樞紐表設定 86
  - 輸出標記 87
  - 檢視器 86
  - 顯示 167
  - IBM SPSS Modeler 的 165
  - PMML 168
- 選擇性複製 63
- 遺漏值 89, 90, 99
  - 在記錄中 90
  - 系統 90
  - CLEM 表示式 93
- 錯誤訊息 42

## 〔十七劃〕

- 優先級 112
- 儲存
  - 多個物件 54
  - 串流 54
  - 狀態 54
  - 節點 54
  - 輸出物件 54
- 儲存格內容 81
- 儲存圖表 65, 70
  - BMP 檔 65, 70
  - EMF 檔 65
  - EPS 檔 65, 70
  - JPEG 檔 65, 70
  - meta 檔 65
  - PICT 檔 65
  - PNG 檔 70
  - PostScript 檔 70
  - TIFF 檔 70
- 儲存輸出 65, 68, 69
  - 文字格式 65, 69
  - Excel 格式 65, 67
  - HTML 65, 66
  - HTML 格式 65
  - PDF 格式 65, 68
  - PowerPoint 格式 65, 68
  - Web 報告 66
  - Word 格式 65, 67
- 應用程式 21
- 應用程式範例 3
- 檔案 62
  - 新增文字檔到檢視器 62
- 檢查 CLEM 表示式 106
- 檢視器 59, 60, 61, 72, 86, 87
  - 大綱窗格 59
  - 收合概要 61
  - 刪除輸出 60
  - 展開概要 61
  - 移動輸出 60

- 檢視器 (繼續)
  - 尋找和取代資訊 62
  - 結果框 59
  - 搜尋和取代資訊 62
  - 概要 61
  - 輸出項目間之空間 72
  - 儲存文件 72
  - 隱藏結果 59
  - 變更概要字型 61
  - 變更概要的大小 61
  - 變更概要層級 61
  - 顯示資料數值 87
  - 顯示數值標籤 87
  - 顯示選項 86
  - 顯示變數名稱 87
  - 顯示變數標籤 87
- 縮放 15
  - 樞紐表 79, 81
- 鍵盤捷徑 182, 184, 186
- 隱藏 59, 60, 77, 78
  - 列與行。 77
  - 程序的結果 60
  - 註腳 82
  - 維度標記 78
  - 標題 78, 82

## 〔十八劃〕

- 簡介 109
  - IBM SPSS Modeler 7, 165
- 翻轉日 38
- 舊版表格 85
- 轉換函數 115
- 轉置列與行 74
- 鎖定 IBM SPSS Collaboration and Deployment Services Repository 物件 145
- 鎖定節點 36
- 雜誌資料 22
- 顏色
  - 設定 167

## 〔十九劃〕

- 羅吉斯迴歸 189
  - 匯出為 PMML 168
- 邊框 80, 84
  - 顯示隱藏的邊框 84
- 關鍵字
  - 標註節點 53
- 類別 15, 157
- 類型節點
  - 效能 178
  - 遺漏值 92



類神經網路節點  
大型集合 37  
效能 179

## (二十劃)

警告 42  
設定選項 166

## (二十一劃)

屬性 21  
欄位 21, 109, 110  
在 CLEM 表示式中 105  
檢視值 105

## (二十二劃)

歡迎使用對話框 167

## (二十三劃)

變數 21, 86  
對話框中的顯示順序 86  
變數名稱 86  
滑鼠右鍵按鈕「說明」 86  
變數標籤 76, 86, 87  
在樞紐表上 87  
滑鼠右鍵按鈕「說明」 86  
數值標籤 87  
邏輯函數 118  
顯示 59, 77, 78  
列或行 77  
結果 59  
註腳 82  
維度標記 78  
標題 78, 82  
顯示次序 74  
顯示格式  
小數位數 39  
分組符號 39  
地理空間座標 41, 42  
科學 39  
貨幣 39  
數字 39

## A

abs 函數 118  
allbutfirst 函數 122  
allbutlast 函數 122  
alphabefore 函數 122  
arccos 函數 119  
arccosh 函數 119

arcsin 函數 119  
arcsinh 函數 119  
arctan 函數 119  
arctan2 函數 119  
arctanh 函數 119  
area 函數 120

## B

BMP 檔 65, 70  
匯出圖表 65, 70

## C

cdf\_chisq 函數 120  
cdf\_f 函數 120  
cdf\_normal 函數 120  
cdf\_t 函數 120  
CLEM 102  
函數 102, 103  
表示式 97, 109  
建立表示式 102  
資料類型 109, 110  
語言 109  
範例 95  
檢查表示式 106  
簡介 20, 95  
CLEM 函數  
三角的 119  
日期時間 126  
比較 116  
可用清單 114  
字串 122  
序列 129  
空白值和空白置 134  
空間 120  
特殊函數 134  
參考 115  
逐位元 121  
廣域 133  
數值的 118  
遺漏值 93  
隨機 122  
轉換 115  
邏輯 118  
probability 120  
CLEM 表示式  
效能 180  
close\_to  
空間函數 120  
close\_to 函數 120  
Cognos Active Report 66  
COP 9  
cos 函數 119  
cosh 函數 119

count\_equal 函數 100, 116  
count\_greater\_than 函數 100, 116  
count\_less\_than 函數 100, 116  
count\_non\_nulls 函數 116  
count\_not\_equal 函數 100, 116  
count\_nulls 函數 93, 100, 116  
count\_substring 函數 122  
CRISP-DM 15, 157  
專案視圖 157  
CRISP-DM 程序模型 23  
crosses  
空間函數 120  
crosses 函數 120

## D

datetime\_date 函數 115  
date\_before 函數 116  
DIFF 函數 129  
distance  
空間函數 120  
div 函數 118  
DTD 154

## E

endstring 函數 122  
EPS 檔 65, 70  
匯出圖表 65, 70  
Excel 格式  
匯出輸出 65, 67

## F

f 分佈  
機率函數 120  
first\_index 函數 101, 116  
first\_non\_null 函數 101, 116  
first\_non\_null\_index 函數 101, 116  
fracof 函數 118

## H

hasendstring 函數 122  
hasmidstring 函數 122  
hasstartstring 函數 122  
hassubstring 函數 122  
HTML 65, 66  
匯出輸出 65, 66  
HTML 輸出  
螢幕閱讀軟體 189

## I

- IBM SPSS Analytic Server
  - 多個連線 10
  - 連線 10
- IBM SPSS Collaboration and Deployment Services 137
- IBM SPSS Modeler 1, 11
  - 文件 3
  - 協助工具特性 181
  - 從指令行執行 8
  - 啟動使用 7
  - 提示和快捷鍵 58
  - 概述 7, 165
  - 選項 165
- IBM SPSS Modeler Advantage 137
- IBM SPSS Modeler Server 1
  - 主機名稱 8, 9
  - 使用者 ID 8
  - 埠號 8, 9
  - 密碼 8
  - 網域名稱 (Windows) 8
- IBM SPSS 協同作業及部署服務儲存庫 137
  - 刪除物件和版本 146
  - 物件內容 147
  - 將專案轉移至 159
  - 連接至 138
  - 單一登入 138
  - 搜尋範圍 144
  - 資料夾 145, 146
  - 認證 138
  - 儲存物件 139
  - 擷取物件 143
  - 瀏覽 139
  - 鎖定和解除鎖定物件 145
- if、then、else 函數 118
- INDEX 函數 129
- integer\_bitcount 函數 121
- integer\_leastbit 函數 121
- integer\_length 函數 121
- intof 函數 118
- isalphacode 函數 122
- isendstring 函數 122
- islowercode 函數 122
- ismidstring 函數 122
- isnumbercode 函數 122
- isstartstring 函數 122
- issubstring 函數 122
- issubstring\_count 函數 122
- issubstring\_lim 函數 122
- isuppercode 函數 122
- is\_date 函數 115
- is\_datetime 函數 115
- is\_integer 函數 115
- is\_number 函數 115

- is\_real 函數 115
- is\_string 函數 115
- is\_time 函數 115
- is\_timestamp 函數 115

## J

- Java 190
- JAWS 181, 188, 190
- JPEG 檔 65, 70
  - 匯出圖表 65, 70

## K

- Kohonen 節點
  - 大型集合 37
  - 效能 179
- K-Means 節點
  - 大型集合 37
  - 效能 179

## L

- last\_index 函數 101, 116
- LAST\_NON\_BLANK 函數 129
- last\_non\_null 函數 101, 116
- last\_non\_null\_index 函數 101, 116
- length 函數 122
- locchar 函數 122
- locchar\_back 函數 122
- log10 函數 118
- lowertoupper 函數 122

## M

- matches 函數 122
- MAX 函數 129
- max 函數 116
- max\_index 函數 101, 116
- max\_n 函數 100, 116
- MEAN 函數 129
- mean\_n 函數 100, 118
- meta 檔 65
  - 匯出圖表 65
- MIN 函數 129
- min 函數 116
- min\_index 函數 101, 116
- min\_n 函數 100, 116
- mod 函數 118

## N

- num\_points
  - 空間函數 120

- num\_points 函數 120

## O

- OFFSET 函數 129
- oneof 函數 122
- overlap
  - 空間函數 120
- overlap 函數 120

## P

- PDF
  - 匯出輸出 65, 68
- pi 函數 119
- PMML
  - 匯入模型 154
  - 匯出模型 154
  - 匯出選項 168
- PMML 模型
  - 線性迴歸 168
  - 羅吉斯迴歸 168
- PNG 檔 65, 70
  - 匯出圖表 65, 70
- PostScript 檔 (密封式) 65, 70
  - 匯出圖表 65, 70
- PowerPoint 68
  - 將輸出匯出為 PowerPoint 檔 68
- PowerPoint 格式
  - 匯出輸出 65
- PowerPoint 檔案 157

## R

- random 函數 122
- random0 函數 122
- rem 函數 118
- replace 函數 122
- replicate 函數 122

## S

- SAS 檔案
  - 編碼 191
- Script 顏色
  - 設定 168
- Scripting 20, 95
- SDEV 函數 129
- sdev\_n 函數 100, 118
- sin 函數 119
- SINCE 函數 129
- sinh 函數 119
- skipchar 函數 122
- skipchar\_back 函數 122
- SoundEx 函數 125

soundex\_difference 函數 125  
SQL 產生  
    記載 40  
    預覽 40  
sqrt 函數 118  
startstring 函數 122  
stripchar 函數 122  
strmember 函數 122  
subscrs 函數 122  
substring 函數 122  
substring\_between 函數 122  
SUM 函數 129  
sum\_n 函數 100, 118

## T

t 分佈  
    機率函數 120  
tan 函數 119  
tanh 函數 119  
testbit 函數 121  
THIS 函數 129  
TIFF 檔 70  
    匯出圖表 65, 70  
time\_before 函數 116  
to\_date 函數 115, 126  
to\_dateline 函數 126  
to\_datetime 函數 115  
to\_integer 函數 115  
to\_number 函數 115  
to\_real 函數 115  
to\_string 函數 115  
to\_time 函數 115, 126  
to\_timestamp 函數 115, 126  
trim 函數 122  
trimend 函數 122  
trim\_start 函數 122

## U

undef 函數 134  
Unicode 支援 191  
unicode\_char 函數 122  
unicode\_value 函數 122  
uppertolower 函數 122  
URL  
    IBM SPSS Analytic Server 10  
UTF-8 編碼 37, 191

## V

value\_at 函數 101, 116

## W

Web 報告 66  
    匯出輸出 66  
within  
    空間函數 120  
within 函數 120  
Word 格式  
    匯出輸出 65, 67  
    寬表格 65

## 〔特殊字元〕

@BLANK 函數 93, 115, 134  
@DIFF 函數 129  
@FIELD 函數 93, 134  
@FIELDS\_BETWEEN 函數 93, 100, 134  
@FIELDS\_MATCHING 函數 93, 100,  
    134  
@INDEX 函數 129  
@LAST\_NON\_BLANK 函數 129, 134  
@MAX 函數 129  
@MEAN 函數 129  
@MIN 函數 129  
@MULTI\_RESPONSE\_SET 函數 101,  
    134  
@NULL 函數 93, 115, 134  
@OFFSET 函數 129  
    效能考量 180  
@PARTITION\_欄位 函數 134  
@PREDICTED 函數 134  
@SDEV 函數 129  
@SINCE 函數 129  
@SUM 函數 129  
@TARGET 函數 134  
@TESTING\_PARTITION 函數 134  
@THIS 函數 129  
@TODAY 函數 126  
@TRAINING\_PARTITION 函數 134  
@VALIDATION\_PARTITION 函數 134







Printed in Taiwan