



---

## 產品特色

- 輕鬆存取、準備及整合結構化資料文字、網頁與調查資料
  - 迅速從文字識別及擷取情緒，適用 30 多種語言
  - 透過支援 IBM Classic Federation Server 和 zDB2® 資料儲存庫，將存取範圍擴大至大型主機等舊式系統儲存的文字
  - 可直接從介面匯出至 Cognos® 8 Business Intelligence，拓展可見度，提供完整的客戶視野
  - 透過 zLinux、SuSE Linux Enterprise Server 及 IBM Smart Analytics System for Power 系列產品，提供更多部署選項
- 

# 運用非結構化資料提升模型精確度

IBM SPSS Modeler 協助您根據預測智慧做出商業決策，從而實現卓越成果。預測智慧不僅可評估趨勢（評比成果、計畫和績效），還能預測未來（評估可能結果，並瞭解各種因素如何交互作用，影響結果），因此可協助組織擬定更有效的策略。

部署 Modeler，您可使用功能強大且業經實證的分析技術，更深入瞭解客戶或關係人，從而加快商業問題解決速度。然而，大部分資料都是以非結構化或文字化形式存在，例如附註、檔案或網頁，如果僅以結構化資料建模，可能無法得到完整的商業程序和成果視野。

IBM SPSS Modeler Premium 協助組織發掘各式資料隱藏的預測智慧，分析結構化的數字資料之外，更從網路活動、部落格內容、客戶意見回饋、電子郵件、文章等非結構化資料，蒐羅資訊，藉以探索概念和情感間的關係，盡可能建立最精確的預測模型。

## 提升生產力、品質及準確度

部署 Modeler Premium，您可以在互動式的視覺化環境內，同時執行文字分析與資料採礦。透過直觀化圖形介面，您可以將資料採礦程序的每個步驟，當成「串流」的一部分，輕鬆檢視。文字分析簡易又有效，並有互動圖形協助瀏覽及顯示文字資料和模式，便於進行即時分析，同時提供功能強大的分類技術，將文字轉換成分析資產。

最新版本提供強化功能，方便企業將文字分析功能擴展至整個組織。如欲進一步瞭解此等強化功能，請參閱第 2 頁。



透過 IBM SPSS Modeler 的視覺化介面，您可以輕鬆存取及整合許多不同來源的資料，包括幾乎任何類型資料庫的資料、試算表或純文字檔（例如 IBM® SPSS® Statistics、SAS® 和 Microsoft® Excel® 檔案）以及文字資料和來自 Web 2.0 的資料（例如 RSS 資訊餽送和 IBM® SPSS® Data Collection 產品檔案），現支援 Cognos® 8 Business Intelligence 以及建置在 IBM Classic Federation Server 和 zDB2® 上的舊式系統。沒有任何資料採礦解決方案能夠提供如此多用途。

IBM SPSS Modeler Premium 提供功能強大的自動化工具，包括自動資料準備與自動建模，使用者可以輕鬆準備分析資料、根據資料背後隱藏的模式找出最佳模型，並迅速產生一致且準確的結果。產業專屬的文字分析套件與樣板能夠加速分析程序，確保使用者獲得最準確的結果。強大的自然語言處理功能，可將文字分類整理成階層式結構，而這些分類又可自動整合至預測模型。

此解決方案支援「跨產業資料採礦標準程序」(CRISP-DM)，這可讓分析師專注於解決商業問題，而非程式設計。個別專案可利用 CRISP-DM 專案管理程式，有效分門別類。

## Modeler Premium 14.1 的新功能

最新版 Modeler Premium 的新功能可擴大預測智慧在組織內的部署範圍，讓員工更深入瞭解自己的公司、平時作業的環境以及客戶和利害關係人，更專注於自己的核心業務，更迅速完成日常決策。新功能包括：

- **整合 Cognos**。分析師現可透過 Cognos 8 Business Intelligence 軟體，直接存取結構化資料。Cognos 軟體整理並提供完整一致的資訊視野，協助企業做出決策，若再加入 Modeler 的分析功能，組織更可快速可靠地評估特定結果的可能性。此外，由於 Modeler 可將結果寫入 Cognos 8 Business Intelligence，因此企業使用者和所有資訊利害關係人，都可透過 Cognos 儀表板，存取整合結構化和非結構化資料得出客戶視野後，所建立的預測智慧。
- **強化 InfoSphere™ Warehouse 及 IBM DB2® 功能**。部署 InfoSphere Data Warehouse 的組織，現可從 Modeler 直接使用各種資料採礦演算法，這可讓使用者享受直觀化圖形介面的便利，卻不影響資料採礦的效能。目前支援的演算法包括邏輯回歸、Naïve Bayes、時間序列和徑向基函數 (RBF)。

- **大型主機資料採礦功能**。大型主機電腦通常儲存許多有關組織營運歷程的資料。由於 Modeler 能透過 IBM System z<sup>®</sup> 作業系統直接存取系統上的資料，因此這些組織可以檢視過往事件對未來有何影響，有助更明確瞭解現行活動，掌握瞬息萬變的商場情勢，並且根據可靠的預測智慧進行規劃工作。
- **支援在 System z 運行的 Linux<sup>®</sup>**。在 System z 上運行 Linux 具有許多優勢，可協助組織簡化複雜的資訊系統，同時滿足現今企業對安全性、透明度和成本控制的需求。IBM SPSS Modeler Professional 現已開始支援這類環境。

除上述之外，Modeler Premium 14 新增下列強化功能，亦有效加強其從文字資料擷取見解的能力：

- 使用產業專屬文字分析套件與樣板，分析銀行，保險或廣告文字、情感和俚語，結果更快、更準確
- 建立階層式分類架構，以更合邏輯、更詳細的方式整理概念
- 匯入階層式種類、註解和關鍵字描述子等預先定義的分類，以及將其匯出至 Excel
- 儲存階層式種類以供重複使用，使用加強型語義網路分組技術建立種類
- 運用可識別不同產業差異的語義網路，同時以更有效的方式使用硬體，藉此提高擷取文字的速度和準確度，即使是大型資料集，亦能輕鬆處理
- 使用加強型文字連結規則編輯器，定義及測試範例文字的規則，再將此規則應用至實際資料

### 擴大資料探索範圍

只有 Modeler Premium 可讓您直接存取文字、網頁和調查資料，並將這些特殊類型的資料整合至預測模型，以便提供更有用的建議，改善成果。

互動式文字採礦工作台可節省時間和精力，同時協助您取得競爭優勢。不需要語言背景就能使用，明顯優於其他的文字分析工具。您可以使用資源編輯器，一套管理文字擷取過程的整合工具，針對特定領域輕鬆自訂概念字典，加快您尋找相關概念與關聯的速度。

您還可以直接從 **Modeler** 的主要工具列，為特定商業應用程式建立自訂樣板與儲存庫，再將這些寶貴資源重複使用於其他產品和應用程式，包括 **IBM® SPSS® Text Analytics for Surveys**。Modeler Premium 的語言資源支援眾多產業和應用，包括情緒分析、客戶關係管理、安全情報、市場情報、生命科學（基因組學和 MESH）和 IT。

### 提供最多技術選擇

**IBM SPSS Modeler** 提供一系列先進的資料採礦技術，旨在滿足需要每種資料採礦應用的需求，包括下列資料採礦演算法。

- 分類演算法：使用決策樹狀結構 (Decision Tree)、類神經網路 (Neural Network)、邏輯回歸 (Logistic Regression)、時間序列 (Time-Series)、支援向量機 (Support Vector Machines)、Cox 回歸 (Cox regression) 等技術，根據歷程資料做出預測。運用自動分類建模技術提供二進位和數字結果，簡化模型建立程序。
- 分割演算法：運用自動叢集、異常偵測及叢集類神經網路技術，分群人員或偵測異常模式。使用自動分類，可在單一步驟應用多種演算法，不必猜測就能正確選出適當的技術。
- 關聯演算法：使用 Apriori、CARMA 和循序關聯規則，探索關聯、連結或序列。

透過業經實證的自然語言處理 (NLP) 語言擷取程序，Modeler Premium 可從多種非結構化資料取出主要概念，並加以分類。而後，再將擷取的概念、意見和種類結合結構化資料，應用於預測模型，藉以發掘行動、行為、模式和關聯背後的寶貴見解。

文字連結分析 (TLA) 技術可協助從多種語言識別及擷取情緒及意見，包括荷蘭文、英文、法文、德文、義大利文、日文、葡萄牙文和西班牙文。英文版介面亦支援第三方翻譯軟體。

## 最佳化現行資訊技術

IBM SPSS Modeler Premium 可擴充的開放式架構，能夠充分發揮現有 IT 基礎架構的價值。其與現有系統整合，無論存取資料或部署結果，均無需改動資料的格式。資料庫內建模、多執行緒作業、叢集和內嵌演算法等技術，有助節省資源、加快達效速度及降低整體 IT 成本。

## 支援整個企業的資料採礦

IBM SPSS Modeler Premium 可有效分析一般中小型組織產生的資料量，而存在大量或複雜資料採礦需求的組織，則適用 IBM SPSS Modeler Premium Server。Modeler Server 採用主從式架構，可讓許多分析師同時工作，運算資源不會因此吃緊。企業版支援主流資訊平台執行資料庫內採礦功能，可有效處理大量資料。此外，Modeler Server 還提供其他部署選項，協助您跨地域或跨部門擴展資料和文字採礦功能，並快速提供決策者所需的結果。

為提升 IT 效率，Modeler Premium Server 現可安裝運行 Linux 作業系統的 IBM System z 硬體。這樣的 IBM System z 環境可以提供更強運算能力，同時處理多種工作量，因此有助加快速度分析的數量與速度，滿足現今企業對於效能、可擴充性和可靠性的需求。

此外，IBM SPSS Modeler Server 還提供額外的部署選項，協助您跨地域或跨部門擴展資料和文字採礦功能，快速提供決策者所需的結果。例如，如欲有效管理分析資產及自動化分析程序，可以使用 Modeler Premium 搭配 IBM® SPSS® Collaboration and Deployment Services。您也可以透過 IBM SPSS Decision Management 應用程式，使用文字資料產生的見解，取得更準確的結果，例如，改善即時和批次評分、提供來電客戶即時建議，或加快保險理賠處理。

## 功能

### 資料理解

- 運用自動輔助器建立各種互動式圖形
- 使用視覺化連結分析，檢視資料間的關聯
- 藉由選取圖形上的區域或項目及檢視選定的資訊，來與資料互動；或選取重要資料作為分析之用
- 直接從 **Modeler** 介面存取 **IBM SPSS Statistics** 圖形和報告工具

### 資料準備

- 從 **Cognos® 8 Business Intelligence**、**IBM DB2®**、**Oracle®**、**Microsoft SQL Server™**、**Informix®**、**Neoview**、**Netezza**、**mySQL (Sun)** 和 **Teradata** 資料來源存取作業資料，以及透過支援 **zDB2** 和 **IBM Classic Federation Server**，存取大型主機資料
- 匯入定界和固定寬度的文字檔、**Statistics** 檔案、**SAS**、**Data Collection** 資料來源或 **XML**
- 選用 **Modeler** 提供的多個資料清理選項，包括移除或取代無效資料、自動插補遺漏值以及調和異常值和極端值
- 運用自動化的資料準備程序，一個步驟完成資料的詢問、調節及分析準備
- 將資料匯出為定界的文字檔、**Excel**、**Statistics**、**SAS**、**Cognos 8 Business Intelligence** 套件和作業資料庫
- 使用欄位篩選、命名、衍生、區間、重新分類、價值重置及欄位重新排序功能
- 運用記錄選項、取樣（包括叢集和分層取樣）、合併（包括內部結合、全部外部結合、局部外部結合和反結合）以及連結；排序、集成和平衡

- 有資料重組、分割和調換選項可供選擇
- 有眾多字串函數可供選擇：字串建立、替代、搜尋和比對、空格移除以及截斷
- 直接從 **Modeler** 存取 **Statistics** 提供的資料管理和轉換功能
- 運用 **RFM** 評分：彙整客戶交易以供近期性、頻率和金額評分，並結合這些資料產生完整的 **RFM** 分析

### 文字理解與準備功能

- 從檔案、作業資料庫和 **RSS** 資訊饋送（亦即部落格、網頁資訊饋送），擷取文字資料
- 選取國家語言擷取程式選項，包括荷蘭文、英文、法文、德文、義大利文、葡萄牙文、西班牙文或日文，亦可使用第三方翻譯軟體，幾乎所有語言都能翻譯
- 擷取特定領域的概念，例如單項、表情、縮寫、縮略詞等
- 使用更準確的語言演算法及內嵌或使用者指定的語言資源，計算同義字
- 按人員、組織、期限、產品、地點和其他使用者定義類型，命名概念
- 擷取非語言項目，例如地址、貨幣、時間、電話號碼和社會安全號碼
- 使用及自訂預先建置的樣板和儲存庫，支援情緒分析、客戶關係管理、安全情報、市場情報、生命科學和 **IT**
- 善用預先套裝的文字分析套件 (**TAP**)，這適用於主流的商业應用程式，您亦可建立自己的文字分析套件

- 使用概念叢集演算法，根據詞彙共現性建立叢集，提供相關主題的概要檢視
- 使用文字分類演算法，根據內容智慧分組文字文件和記錄
- 啟用進階概念選取和取消選取功能，供預測模型之用
- 使用文字型和視覺化報告，詢問概念關係、共時性、頻率與類型

### 文字連結分析

- 從荷蘭文、英文、法文、德文和西班牙文文字，識別及擷取情緒（例如好惡）
- 識別人員與事件、疾病與基因等項目之間的關聯
- 從部落格內的 **URL** 識別及擷取內容
- 將意見、語義關係和相連事件納入可部署的預測模型
- 透過互動式圖表，顯示兩概念間複雜的多種語義連結

### 建模及評估

- 運用各種資料採礦算法和多種進階功能，讓資料發揮最大效用。
- 使用互動模式與方程式瀏覽器，並且檢視進階的統計輸出資料
  - 透過變數重要性圖形，顯示資料屬性對預測結果的相對影響
  - 結合多種模型（系集模型）或使用單一模型，分析第二個模型
  - 使用自動分類（二進位和數字）和叢集，無需選取個別演算法

- 使用 Modeler 的元件級延伸架構 (CLEF)，整合自訂演算法
- 透過整合 Statistics，使用 R 擴充分析選項

### 建模演算法包括

- C&RT、C5.0、CHAID 及 QUEST：決策樹狀結構演算法，包括互動式樹狀結構建構
- Decision List：互動式規則建構演算法
- K-Means、Kohonen、Two Step、Discriminant、Support Vector Machine (SVM)：叢集與分割演算法
- Factor/PCA、Feature Selection：資料縮減演算法
- Regression、Linear、GenLin (GLM)：線性方程式建模
- 自我學習反應模型 (SLRM)：具增量學習功能的貝氏定理模型
- 時間序列：產生並自動選取時間序列預測模型
- Neural Networks：具倒傳遞學習功能的多層感知器，及徑向基函數網路
- Support Vector Machine：進階演算法，效能精準，適用於各種資料集
- Bayesian Networks：圖形概率模型
- Cox 迴歸：計算事件的可能時間
- Anomaly Detection：透過使用叢集式演算法，偵測異常的記錄
- KNN：最近鄰接項建模與評分演算法
- Apriori：熱門關聯探索演算法，具進階評估功能
- CARMA：關聯演算法，支援多種結果
- Sequence：序列關聯演算法，適用於注重順序的分析
- 支援 IBM InfoSphere 資料庫內採礦演算法：Decision Tree、Association、Sequence、Regression、Logistic Regression、Clustering、Naïve Bayes、Time-Series 及 Radial Basis Function (RBF)
- 支援 Microsoft SQL Server 資料庫內採礦演算法：Decision Tree、Association Rules、Linear Regression、Clustering、Sequence Clustering、Naïve Bayes、Time-Series 及 Neural Network
- 支援 Oracle 資料庫內採礦演算法：Decision Tree、General Linear Model (GLM)、O-Cluster (Orthogonal Partitioning Clustering)、KMeans、Apriori、Minimum Description Length (MDL)、Support Vector Machine、Naïve Bayes、Adaptive Bayes、Non-Negative Matrix nFactorization 及 Artificial Intelligence (AI)

### 部署

- 使用 SQL 或 PMML (XML 格式，預測模型的標準格式) 匯出模型
- 運用 Collaboration and Deployment Services 的創新分析管理、程序自動化和部署功能

### Modeler 伺服器 (選配)

- 使用資料庫內採礦技術和先進的資料庫技術，在資料庫中建立模型，並充分利用高效能的資料庫實作
- 使用 SQL-pushback 技術促使資料轉換，並選取建模演算法直接納入作業資料庫
- 善用高效能硬體 (包括 IBM System z 大型主機)，縮短解決方案推出時間，並透過平行實行串流和多個模型，提高投資報酬率
- 透過安全套接層 (SSL) 加密方法，在 Modeler Client 和 Modeler Server 之間安全傳輸機密資料

## 關於 IBM 商業智慧分析

IBM 商業智慧分析軟體提供完整、一致及正確的資訊，值得決策者的信賴，提升企業效能。綜合性的產品組合，包括商業智慧、預測分析、財務績效及策略管理，並以分析應用程式提供明確且立即可行的深入見解，讓您瞭解目前績效，預測未來成果。結合豐富的業界解決方案，以及深獲肯定的實務與專業服務，讓任何規模的組織都能帶動最優異的生產力，安心地自動制定決策，創造更優異的成果。

IBM SPSS Predictive Analytics 軟體身為本產品組合的一部分，可協助組織預測未來事件，並可依據見解主動行事，帶動更優異的商業成果。全球各地的企業、政府及學術單位都使用 IBM SPSS 技術，作為吸引、保留、擴大客戶群並減少詐騙案件和規避風險的競爭優勢。組織將 IBM SPSS 軟體納入日常營運，就能成為具備預測能力的企業，能夠引導及自動化各種決策滿足商業目標，並達成可衡量的競爭優勢。如需進一步資訊，或要聯絡業務代表，請造訪 [www.ibm.com/spss](http://www.ibm.com/spss)。



---

© Copyright IBM Corporation 2010

110 台北市松仁路 7 號 3 樓  
技術諮詢熱線：0800-000-700  
台北市松仁路 7 號 3 樓

美國政府使用者的注意事項 - 使用、複製及公開權依 GSA ADP Schedule Contract 與 IBM Corp. 所提出的限制而定。

台灣印製  
2010 年 5 月  
版權所有

IBM、IBM 標誌、ibm.com、WebSphere、InfoSphere 及 Cognos 是國際商業機器股份有限公司在美國及/或其他國家的商標或註冊商標。如果這些或其他 IBM 商標在本文首次出現時，帶有商標符號

(® 或 ™)，代表這類商標在本文發佈時已於美國註冊，或為 IBM 擁有的普通法商標。這類商標可能已在其他國家註冊或屬於普通法商標。IBM 最新的商標清單，請造訪 IBM 網站的「版權及商標資訊」：[www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)。

SPSS 是 IBM 旗下 SPSS, Inc. 的商標，已於全球許多地區登記註冊。

其他公司、產品或服務名稱可能是其所屬公司的商標或服務標誌。



請回收

---