

# 個案分析與實作一 (SAS EM 分群分析) :

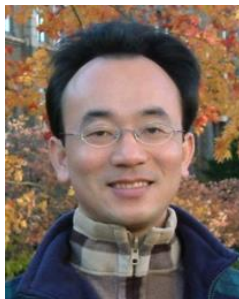
## Case Study 1

### (Cluster Analysis – K-Means using SAS EM)

1042DM06

MI4 (M2244) (3094)

Tue, 3, 4 (10:10-12:00) (B216)



Min-Yuh Day

戴敏育

Assistant Professor

專任助理教授

Dept. of Information Management, Tamkang University

淡江大學 資訊管理學系

<http://mail.tku.edu.tw/myday/>

2016-03-22



# 課程大綱 (Syllabus)

週次 (Week)	日期 (Date)	內容 (Subject/Topics)
1	2016/02/16	巨量資料探勘課程介紹 (Course Orientation for Big Data Mining)
2	2016/02/23	巨量資料基礎：MapReduce典範、Hadoop與Spark生態系統 (Fundamental Big Data: MapReduce Paradigm, Hadoop and Spark Ecosystem)
3	2016/03/01	關連分析 (Association Analysis)
4	2016/03/08	分類與預測 (Classification and Prediction)
5	2016/03/15	分群分析 (Cluster Analysis)
6	2016/03/22	個案分析與實作一 (SAS EM 分群分析)： Case Study 1 (Cluster Analysis – K-Means using SAS EM)
7	2016/03/29	個案分析與實作二 (SAS EM 關連分析)： Case Study 2 (Association Analysis using SAS EM)

# 課程大綱 (Syllabus)

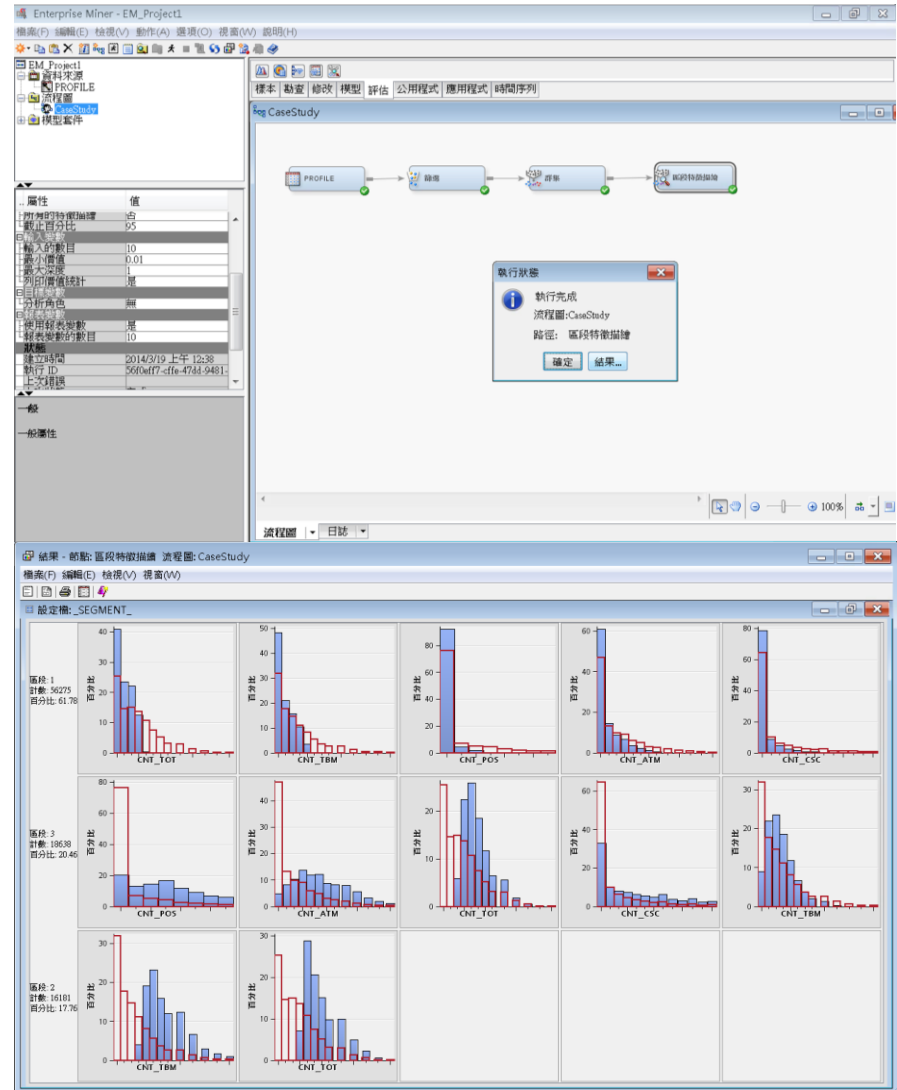
週次 (Week)	日期 (Date)	內容 (Subject/Topics)
8	2016/04/05	教學行政觀摩日 (Off-campus study)
9	2016/04/12	期中報告 (Midterm Project Presentation)
10	2016/04/19	期中考試週 (Midterm Exam)
11	2016/04/26	個案分析與實作三 (SAS EM 決策樹、模型評估) : Case Study 3 (Decision Tree, Model Evaluation using SAS EM)
12	2016/05/03	個案分析與實作四 (SAS EM 迴歸分析、類神經網路) : Case Study 4 (Regression Analysis, Artificial Neural Network using SAS EM)
13	2016/05/10	Google TensorFlow 深度學習 (Deep Learning with Google TensorFlow)
14	2016/05/17	期末報告 (Final Project Presentation)
15	2016/05/24	畢業班考試 (Final Exam)

# 個案分析與實作一 (SAS EM 分群分析) :

## Case Study 1 (Cluster Analysis – K-Means using SAS EM)

### Banking Segmentation

EM_LIB.PROFILE						
	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34.0	3.0	3.0	9.0	49.0
2	000052094	44.0	17.0	5.0	18.0	84.0
3	000057340	122.0	26.0	32.0	35.0	216.0
4	000076885	42.0	3.0	6.0	1.0	52.0
5	000089150	20.0	15.0	2.0	2.0	39.0
6	000094512	83.0	20.0	49.0	3.0	155.0
7	000096396	33.0	9.0	49.0	17.0	108.0
8	000098062	22.0	5.0	9.0	1.0	37.0
9	000122779	21.0	27.0	2.0	1.0	51.0
10	000125838	127.0	3.0	2.0	3.0	135.0
11	000137317	52.0	20.0	2.0	11.0	85.0
12	000147896	101.0	40.0	8.0	1.0	150.0
13	000150206	156.0	52.0	80.0	4.0	292.0
14	000160501	31.0	32.0	42.0	12.0	117.0
15	000176820	107.0	3.0	2.0	1.0	113.0
16	000187784	54.0	14.0	7.0	24.0	99.0
17	000189200	92.0	3.0	2.0	2.0	99.0
18	000198327	44.0	27.0	17.0	3.0	91.0
19	000200237	146.0	3.0	2.0	1.0	152.0
20	000207980	39.0	5.0	2.0	4.0	50.0
21	000227355	180.0	3.0	2.0	10.0	195.0
22	000228317	46.0	14.0	2.0	1.0	63.0
23	000228906	25.0	3.0	2.0	1.0	31.0
24	000229984	15.0	3.0	2.0	1.0	21.0
25	000241022	13.0	3.0	2.0	1.0	19.0
26	000253260	65.0	3.0	2.0	1.0	71.0
27	000257910	35.0	3.0	2.0	1.0	41.0
28	000261717	11.0	5.0	2.0	1.0	19.0
29	000266600	15.0	3.0	2.0	1.0	21.0



# 行銷客戶分群

# 案例情境

- ABC銀行的行銷部門想要針對該銀行客戶的使用行為，進行分群分析，以了解現行客戶對本行的往來方式，並進一步提供適宜的行銷接觸模式。
- 該銀行從有效戶(近三個月有交易者)，取出10萬筆樣本資料。  
依下列四種交易管道計算交易次數：
  - 傳統臨櫃交易(TBM)
  - 自動櫃員機交易(ATM)
  - 銀行專員服務(POS)
  - 電話客服(CSC)

# 資料欄位說明

- 資料集名稱： profile.sas7bdat

Name	Model Role	Measurement Level	Description
ID	ID	Nominal	Customer ID
CNT_TBM	Input	Interval	Traditional bank method transaction count
CNT_ATM	Input	Interval	ATM transaction count
CNT_POS	Input	Interval	Point-of-sale transaction count
CNT_CSC	Input	Interval	Customer service transaction count
CNT_TOT	Input	Interval	Total transaction count

ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	3.0	3.0	3.0	3.0	12.0
2	17.0	5.0	1.0	1.0	24.0
3	26.0	5.0	3.0	3.0	37.0
4	3.0	6.0	1.0	3.0	13.0
5	15.0	5.0	5.0	3.0	28.0
6	20.0	4.0	3.0	3.0	30.0
7	4.0	4.0	1.0	3.0	12.0
8	5.0	4.0	1.0	1.0	11.0
9	25.0	5.0	1.0	3.0	34.0
10	3.0	3.0	3.0	3.0	12.0
11	20.0	5.0	1.0	3.0	29.0
12	4.0	6.0	1.0	3.0	14.0
13	5.0	4.0	4.0	3.0	16.0
14	3.0	4.0	1.0	1.0	9.0
15	3.0	5.0	1.0	1.0	10.0
16	14.0	7.0	3.0	3.0	27.0
17	3.0	3.0	3.0	3.0	12.0
18	25.0	1.0	3.0	3.0	32.0
19	3.0	3.0	1.0	3.0	10.0
20	5.0	3.0	4.0	3.0	15.0
21	3.0	3.0	3.0	3.0	12.0
22	14.0	3.0	1.0	3.0	21.0
23	3.0	3.0	1.0	1.0	8.0
24	3.0	3.0	1.0	1.0	8.0
25	3.0	3.0	1.0	1.0	8.0
26	3.0	3.0	1.0	1.0	8.0
27	3.0	3.0	1.0	1.0	8.0
28	5.0	3.0	1.0	1.0	10.0
29	3.0	3.0	1.0	1.0	8.0

# 行銷客戶分群實機演練

分析目的

依據各往來交易管道 (TBM、ATM、POS、CSC)  
進行客戶分群分析。

演練重點：

- 極端值資料處理
- 分群變數選擇
- 衍生變數產出
- 分群參數調整與分群結果解釋



# SAS Enterprise Miner (SAS EM) Case Study

- SAS EM 資料匯入4步驟
  - Step 1. 新增專案 (New Project)
  - Step 2. 新增資料館 (New / Library)
  - Step 3. 建立資料來源 (Create Data Source)
  - Step 4. 建立流程圖 (Create Diagram)
- SAS EM SEMMA 建模流程

# Download **EM\_Data.zip** (SAS EM Datasets)

[http://mail.tku.edu.tw/myday/teaching/1042/BDM/Data/EM\\_Data.zip](http://mail.tku.edu.tw/myday/teaching/1042/BDM/Data/EM_Data.zip)

<http://mail.tku.edu.tw/myday/teaching.htm>

- [巨量資料探勘 \(Big Data Mining\) \(MI4P\) \(Spring 2016\)](#)  
(MI4) (2 Credit, Elective) (M2244) (3094)  
(1042) (資管四 MI4) (選修2學分) (2016.02-2016.06)  
(週二 Tue, 3, 4, 10:10-12:00) (B216)
  - 週次 (Week) 日期 (Date) 內容 (Subject/Topics)
    - 1 2016/02/16 [巨量資料探勘課程介紹 \(Course Orientation for Big Data Mining\)](#) [ppt] [pdf]
    - 2 2016/02/23 [巨量資料基礎：MapReduce典範、Hadoop與Spark生態系統 \(Fundamental Big Data: MapReduce Paradigm, Hadoop and Spark Ecosystem\)](#) [ppt] [pdf]
    - 3 2016/03/01 [關連分析 \(Association Analysis\)](#) [ppt] [pdf]
    - 4 2016/03/08 [分類與預測 \(Classification and Prediction\)](#) [ppt] [pdf]
    - 5 2016/03/15 [分群分析 \(Cluster Analysis\)](#) [ppt] [pdf]
    - 6 2016/03/22 [個案分析與實作一 \(SAS EM 分群分析\)：Case Study 1 \(Cluster Analysis – K-Means using SAS EM\)](#) [ppt] [pdf] **[EM\_Data]**
    - 7 2016/03/29 [個案分析與實作二 \(SAS EM 關連分析\)：Case Study 2 \(Association Analysis using SAS EM\)](#) [ppt] [pdf] [EM\_Data]
    - 8 2016/04/05 教學行政觀摩日 (Off-campus study)
    - 9 2016/04/12 期中報告 (Midterm Project Presentation)
    - 10 2016/04/19 期中考試週 (Midterm Exam)
    - 11 2016/04/26 [個案分析與實作三 \(SAS EM 決策樹、模型評估\)：Case Study 3 \(Decision Tree, Model Evaluation using SAS EM\)](#) [ppt] [pdf] [EM\_Data]
    - 12 2016/05/03 [個案分析與實作四 \(SAS EM 迴歸分析、類神經網路\)：Case Study 4 \(Regression Analysis, Artificial Neural Network using SAS EM\)](#) [ppt] [pdf] [EM\_Data]
    - 13 2016/05/10 Google TensorFlow 深度學習 (Deep Learning with Google TensorFlow)
    - 14 2016/05/17 期末報告 (Final Project Presentation)
    - 15 2016/05/24 畢業班考試 (Final Exam)
  - References:
    1. [SAS Big Data Champion 2016](#)
    2. [SAS Resource](#)
    3. SAS Enterprise Guide (SAS EG) [[Secured Data](#)]
    4. [SASMining2012 IMTKU Project Presentation 20120921](#) [ppt] [pdf]
    5. [SASMining2012 IMTKU Project Report 20120921](#) [pdf]
    6. [SASMining2012 Final List](#) [pdf]

# Upzip EM\_Data.zip to C:\DATA\EM\_Data

The screenshot shows a Windows Explorer window with the address bar set to "電腦 > System (C:) > DATA > EM\_Data". The search bar contains "搜尋 EM\_Data". The left sidebar shows the folder tree with "EM\_Data" selected. The main pane displays four SAS Data Set files:

File Name	Type	Modification Date	Size
credit.sas7bdat	SAS Data Set	2013/4/21 下午 05:22	361 KB
inq2006.sas7bdat	SAS Data Set	2013/5/29 下午 01:20	13.4 MB
profile.sas7bdat	SAS Data Set	2006/9/23 下午 10:08	5.39 MB
webstation.sas7bdat	SAS Data Set	2006/9/25 上午 01:45	49.5 MB

At the bottom of the window, a status bar indicates "4 個項目".

# Upzip EM\_Data.zip to C:\DATA\EM\_Data

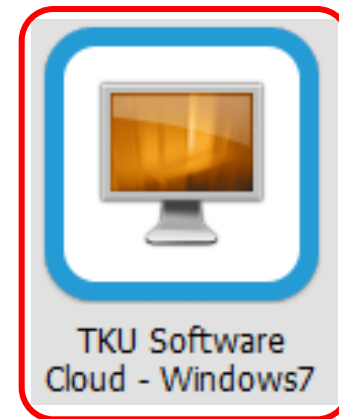
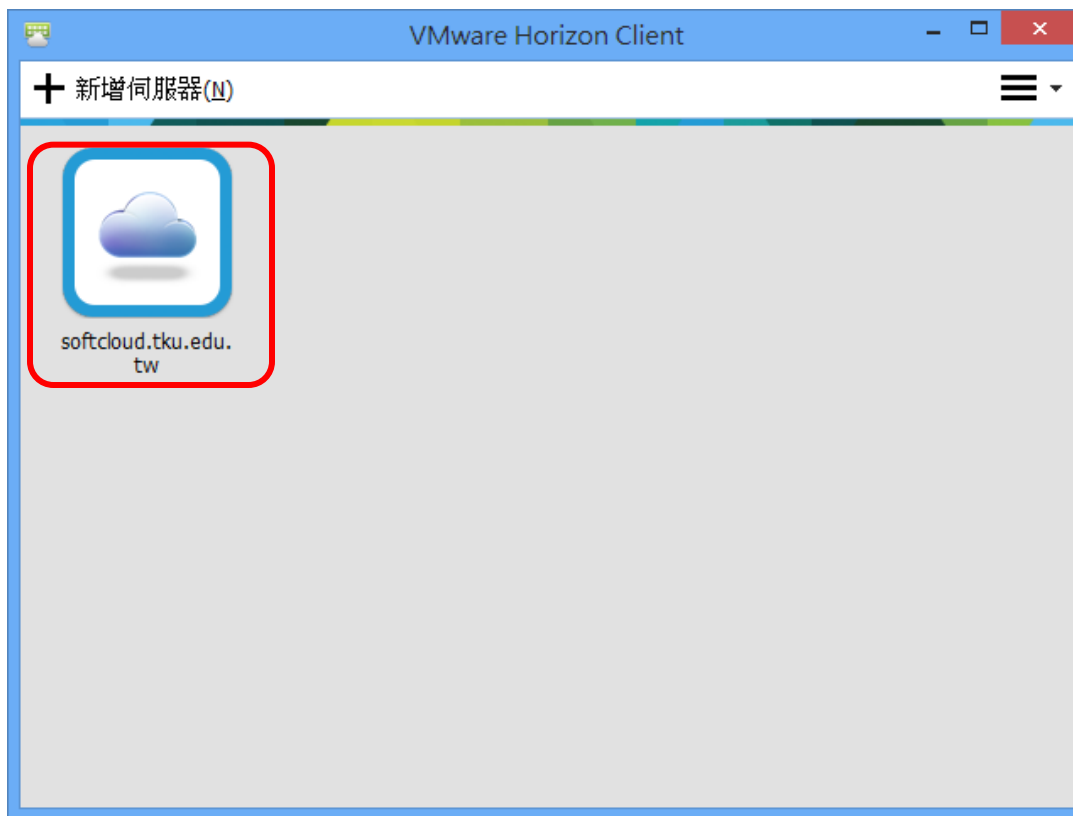
The screenshot shows a Windows Explorer window with the address bar set to 'System (C:) > DATA > EM\_Data'. The search bar contains '搜尋 EM\_Data'. The left sidebar shows the folder tree with 'EM\_Data' selected. The main pane displays a table of files:

名稱	修改日期	類型	大小
credits.sas7bdat	2013/4/21 下午 05:22	SAS Data Set	361 KB
inq2006.sas7bdat	2013/5/29 下午 01:20	SAS Data Set	13,729 KB
profile.sas7bdat	2006/9/23 下午 10:08	SAS Data Set	5,521 KB
webstation.sas7bdat	2006/9/25 上午 01:45	SAS Data Set	50,789 KB

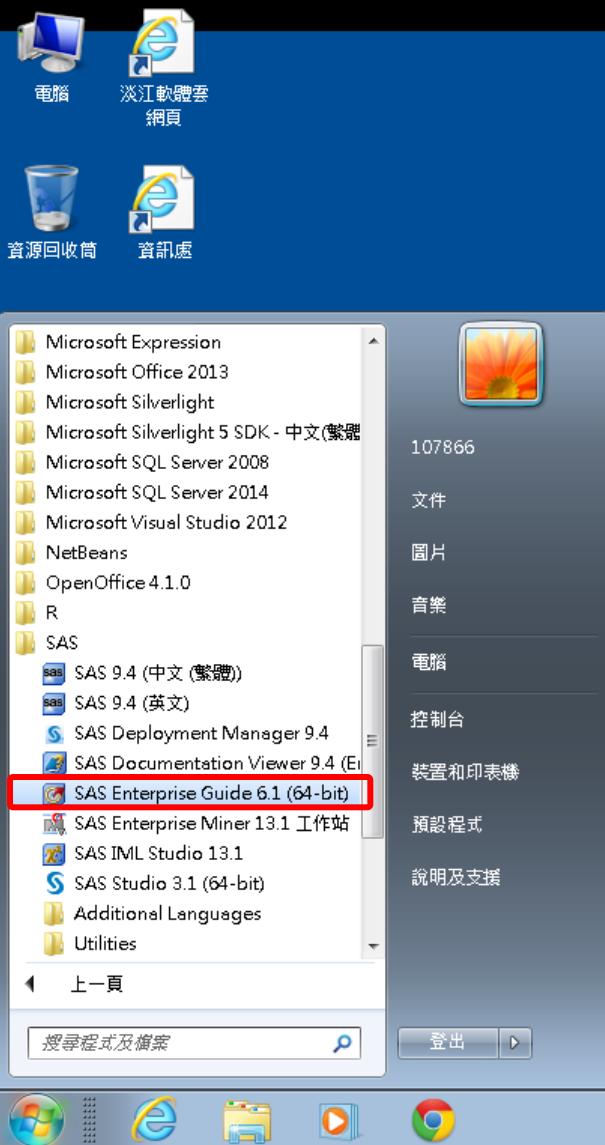
At the bottom of the window, a status bar indicates '4 個項目' (4 items).

# VMware Horizon View Client

[softcloud.tku.edu.tw](http://softcloud.tku.edu.tw)  
SAS Enterprise Miner



# SAS Enterprise Miner 13.1 (SAS EM)



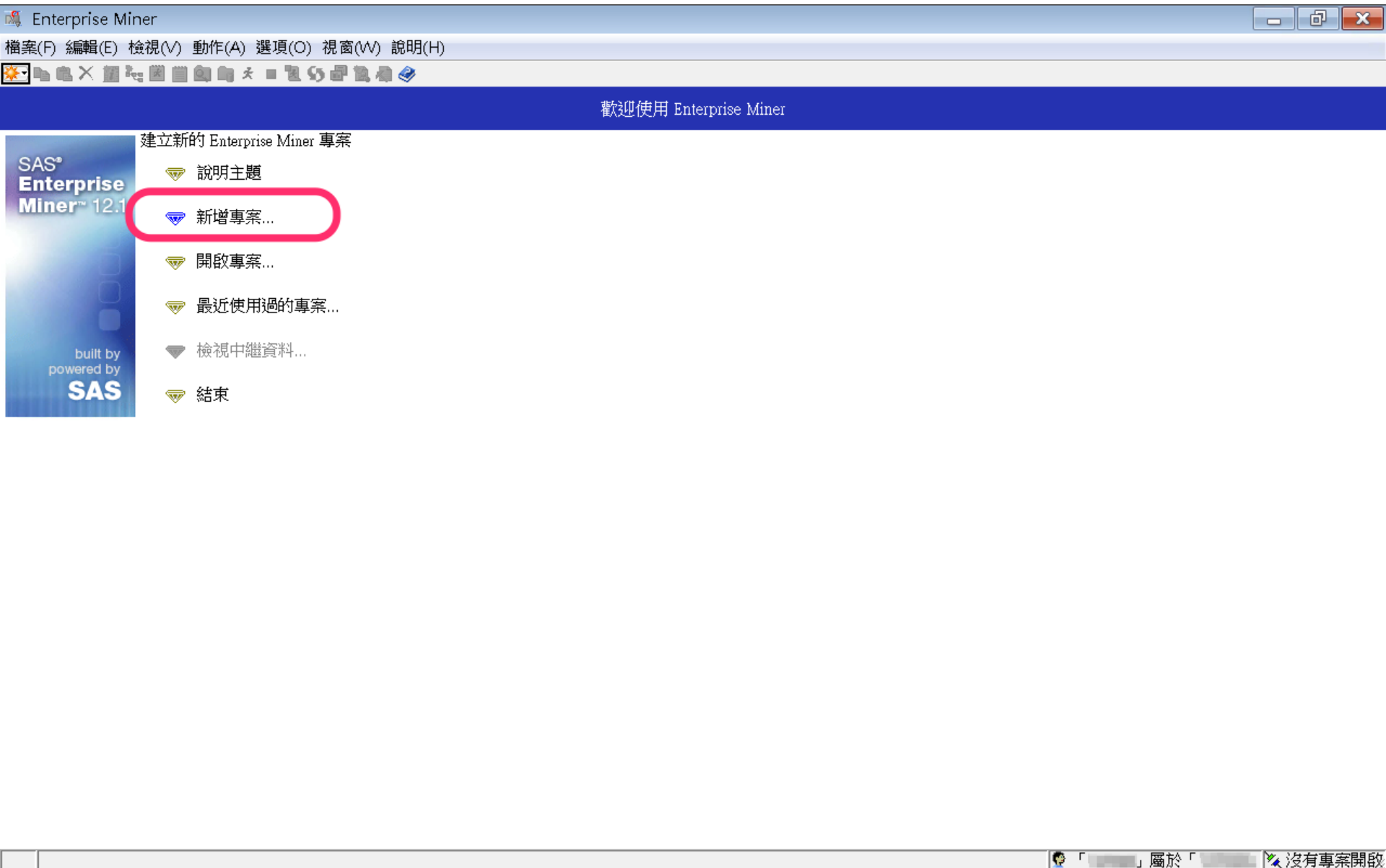
## 淡江軟體雲使用注意事項

- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室，或來電 26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。

# SAS EM 資料匯入4步驟

- Step 1. 新增專案 (New Project)
- Step 2. 新增資料館 (New / Library)
- Step 3. 建立資料來源 (Create Data Source)
- Step 4. 建立流程圖 (Create Diagram)

# Step 1. 新增專案 (New Project)





# Step 1. 新增專案 (New Project)

The sidebar on the left side of the application features the SAS Enterprise Miner 12.1 logo and the text "built by powered by SAS". Below the logo is a vertical list of menu items, each preceded by a small icon:

- 說明主題
- 新增專案...
- 開啟專案...
- 最近使用過的專案...
- 檢視中繼資料...
- 結束

A modal dialog box titled "建立新的專案 -- 第 1 / 2 步 指定專案名稱和伺服器目錄" is displayed. The dialog contains the following elements:

- A small SAS Enterprise Miner 12.1 logo on the left.
- Instructional text: "在 SAS 伺服器上為此專案指定專案名稱和目錄。所有的 SAS 資料集和檔案都將寫入此位置。"
- A text input field labeled "專案名稱" containing the text "EM\_Project1".
- A text input field labeled "SAS 伺服器目錄" containing the text "C:\DATA". To the right of this field is a "瀏覽(B)" button.
- Navigation buttons at the bottom: "< 上一步(B)", "下一步(N) >", and "取消(C)".

# Step 1. 新增專案 (New Project)



- 說明主題
- 新增專案...
- 開啟專案...
- 最近使用過的專案...
- 檢視中繼資料...
- 結束

建立新的專案 -- 第 2 / 2 步 新增專案資訊

新增專案資訊	
名稱	EM_Project1
伺服器目錄	C:\DATA

< 上一步(B)   完成(F)   取消(C)

# SAS Enterprise Miner (EM\_Project1)

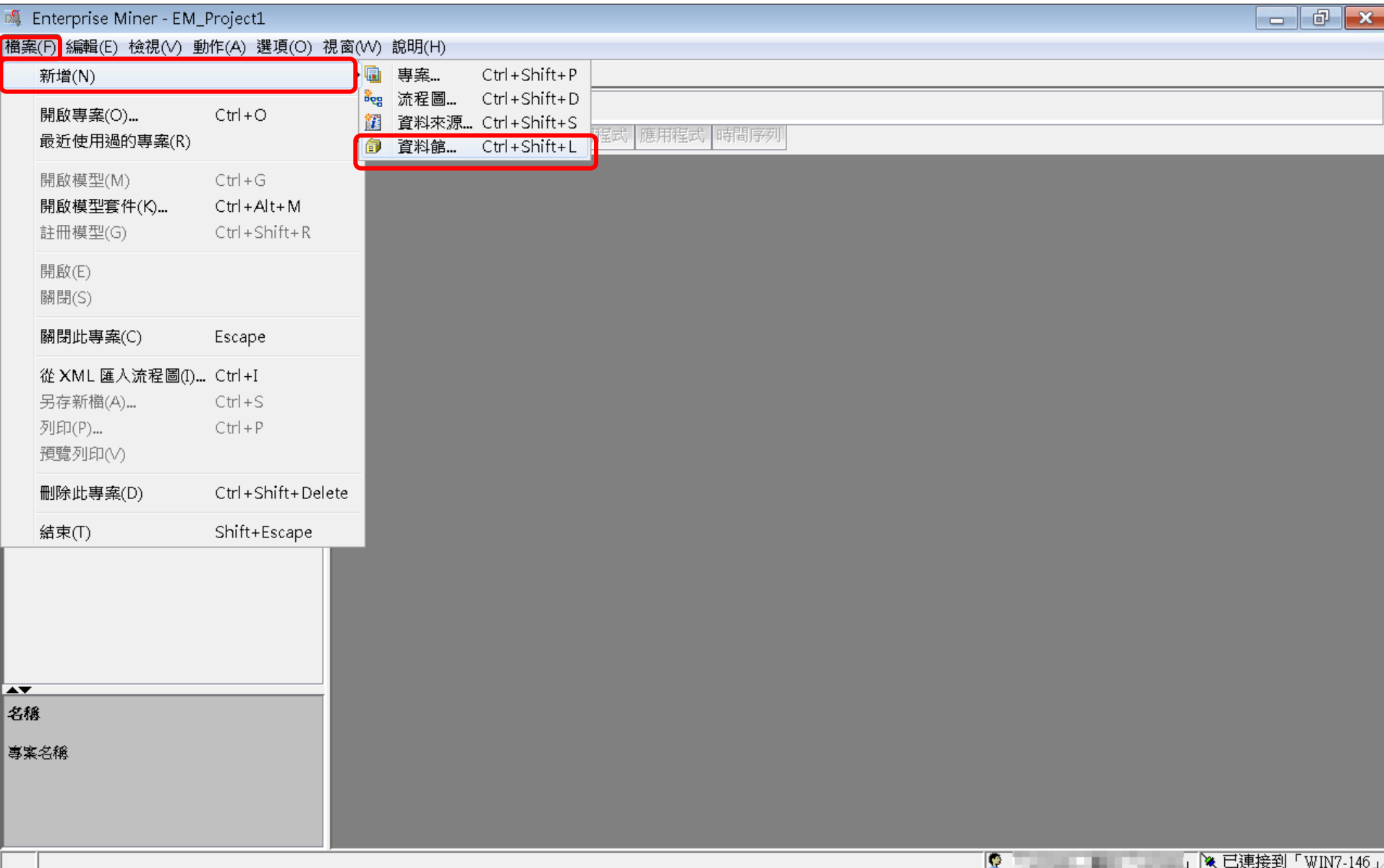
The screenshot displays the SAS Enterprise Miner application window titled "Enterprise Miner - EM\_Project1". The interface includes a menu bar with options: 檔案(F), 編輯(E), 檢視(V), 動作(A), 選項(O), 視窗(W), 說明(H). Below the menu is a toolbar with various icons. The main workspace is divided into several sections:

- Left Panel:** A tree view showing the project structure with folders for "EM\_Project1", "資料來源", "流程圖", and "模型套件". Below this is a table of project attributes.
- Attribute Table:** A table with two columns: "屬性" (Attribute) and "值" (Value).

屬性	值
名稱	EM_Project1
專案啟動程式碼	
專案巨集變數	
建立時間	2014/3/18 下午 10:38
伺服器	
網格可用	否
路徑	C:\DATA
中繼資料資料夾路徑	
最大並行工作數目	預設
- Bottom Left Panel:** A section labeled "名稱" (Name) with a sub-label "專案名稱" (Project Name).
- Top Right Panel:** A set of navigation tabs: 樣本, 動查, 修改, 模型, 評估, 公用程式, 應用程式, 時間序列.

The main workspace area is currently empty, showing a large grey rectangle.

# Step 2. 新增資料館 (New / Library)



# Step 2. 新增資料館 (New / Library)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM Project1

- 資料來源
- 流程圖
- 模型套件

樣本 調查 修改 模型 評估 公用程式 應用程式 時間序列

資料館精靈 -- 第 1 / 3 步 選取動作

請選取動作

- 建立新的資料館(R)
- 修改資料館(M)
- 刪除資料館(D)

< 上一步(B) 下一步(N) > 取消(C)

屬性	值
名稱	EM_Project1
專案啟動程式碼	
專案巨集變數	
建立時間	2014/3/18 下午 10:3
伺服器	
網格可用	否
路徑	C:\DATA
中繼資料資料夾路徑	
最大並行工作數目	預設

名稱

專案名稱

已連接到「WIN7-146」

# Step 2. 新增資料館 (New / Library)

The screenshot shows the Enterprise Miner interface with a 'Library Wizard' dialog box open. The dialog is titled '資料館精靈 -- 第 2 / 3 步 建立或修改' (Library Wizard -- Step 2 / 3 Create or Modify). It contains the following fields and controls:

- 名稱 (Name):** A text box containing 'EM\_Lib'.
- 引擎 (Engine):** A dropdown menu set to 'BASE'.
- 資料館資訊 (Library Information):**
  - 路徑 (Path):** A text box containing 'C:\DATA\EM\_Data' with a '瀏覽(R)...' (Browse...) button to its right.
  - 選項 (Options):** An empty text box.
- Navigation:** Buttons for '< 上一步(B)' (Previous Step), '下一步(N) >' (Next Step), and '取消(C)' (Cancel).

In the background, the Enterprise Miner main window is visible, showing a project tree on the left with 'EM\_Project1' selected. The main workspace contains a toolbar with icons for '樣本' (Sample), '勘查' (Explore), '修改' (Modify), '模型' (Model), '評估' (Evaluate), '公用程式' (Utilities), '應用程式' (Applications), and '時間序列' (Time Series). A status bar at the bottom right indicates '已連接到「WIN7-146」' (Connected to 'WIN7-146').

# Step 2. 新增資料館 (New / Library)

The screenshot shows the Enterprise Miner interface with a file selection dialog box open. The dialog box is titled "開啟" (Open) and shows the contents of the "EM\_Data" folder. The files listed are:

- credits.sas7bdat
- inq2006.sas7bdat
- profile.sas7bdat
- webstation.sas7bdat

The dialog box also shows the file name "C:\DATA\EM\_Data" and the file type "所有檔案" (All files). The "開啟(O)" button is highlighted, and a tooltip "開啟選取的檔案" (Open selected file) is visible over it.

The background interface shows the "Enterprise Miner - EM\_Project1" window with a menu bar (檔案(F), 編輯(E), 檢視(V), 動作(A), 選項(O), 視窗(W), 說明(H)) and a toolbar. The left sidebar shows a tree view with "EM Project1" and its sub-items: "資料來源", "流程圖", and "模型套件". The bottom left shows a table with project properties.

屬性	值
名稱	EM_Project1
專案啟動程式碼	
專案巨集變數	
建立時間	2014/3/18 下午 10:30
伺服器	
網絡可用	否
路徑	C:\DATA
中繼資料資料夾路徑	
最大並行工作數目	預設

# Step 2. 新增資料館 (New / Library)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM Project1  
資料來源  
流程圖  
模型套件

樣本 動查 修改 模型 評估 公用程式 應用程式 時間序列

資料館精靈 -- 第 3 / 3 步 確認動作

屬性	值
動作	建立新的
名稱	EM_Lib
引擎	BASE
路徑	C:\DATA\EM_Data
選項	

狀態  
動作成功!  
資料館「EM\_Lib」已建立。

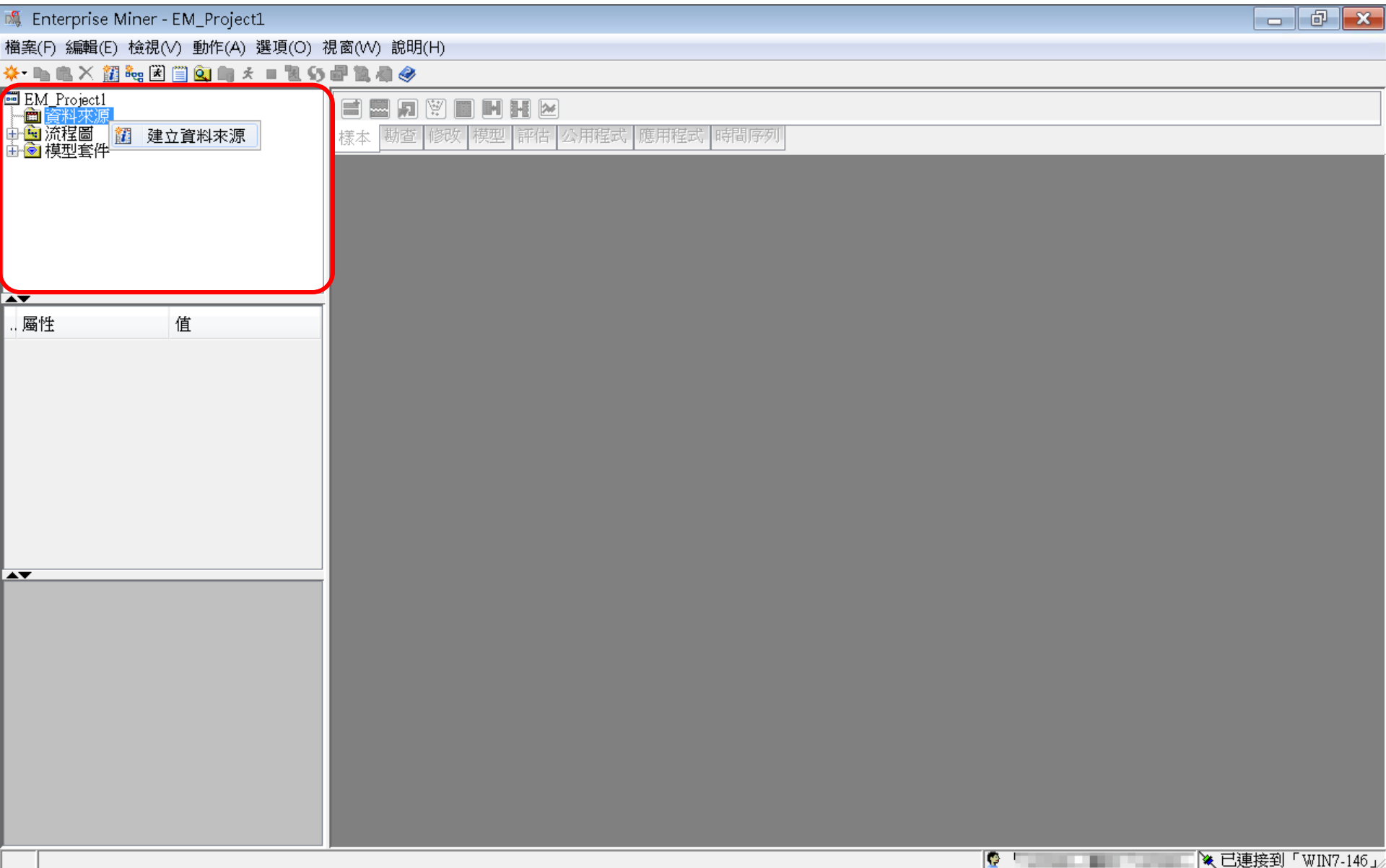
< 上一步(B) 完成(F)

名稱  
專案名稱

已連接到「WIN7-146」



# Step 3. 建立資料來源 (Create Data Source)



# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 1 / 8 步 中繼資料來源

選取中繼資料來源

來源(S): SAS 表格

< 上一步(B) 下一步(N) > 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 2 / 8 步 選擇 SAS 表格

選擇 SAS 表格

表格(T):  瀏覽(R)...

< 上一步(B) 下一步(N) > 取消(C)

.. 屬性 值

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM Project1

- 資料來源
- 流程圖
- 模型套件

樣本 調查 修改 模型 評估 公用程式 應用程式 時間序列

選擇 SAS 表格

SAS 資料館	名稱	引擎	路徑
Em_lib	Em_lib	BASE	C:\DATA\EM_Data
Maps	Maps	V9	C:\Program Files\SASHome\SASFo...
Mapsgfk	Mapsgfk	V9	C:\Program Files\SASHome\SASFo...
Mapssas	Mapssas	V9	C:\Program Files\SASHome\SASFo...
Sampsio	Sampsio	V9	C:\Program Files\SASHome\SASFo...
Sashelp	Sashelp	V9	C:\Program Files\SASHome\SASFo...
Sasuser	Sasuser	V9	C:\Users\..... Documents\My SA...
Work	Work	V9	C:\Users\..... AppData\Local\Te...

取得詳細資料(D) 重新整理(R) 屬性(P)... 確定(O) 取消(C)

瀏覽(R)...

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 2 / 8 步 選取 SAS 表格

選取 SAS 表格

表格(T):

選取 SAS 表格

名稱	類型
Credit	表格
Inq2006	表格
Profile	表格
Webstation	表格

取得詳細資料(D) 重新整理(R) 屬性(P)... 確定(O) 取消(C)

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 2 / 8 步 選擇 SAS 表格

**DatabaseName.TableName**

選擇 SAS 表格

**LibraryName.TableName**

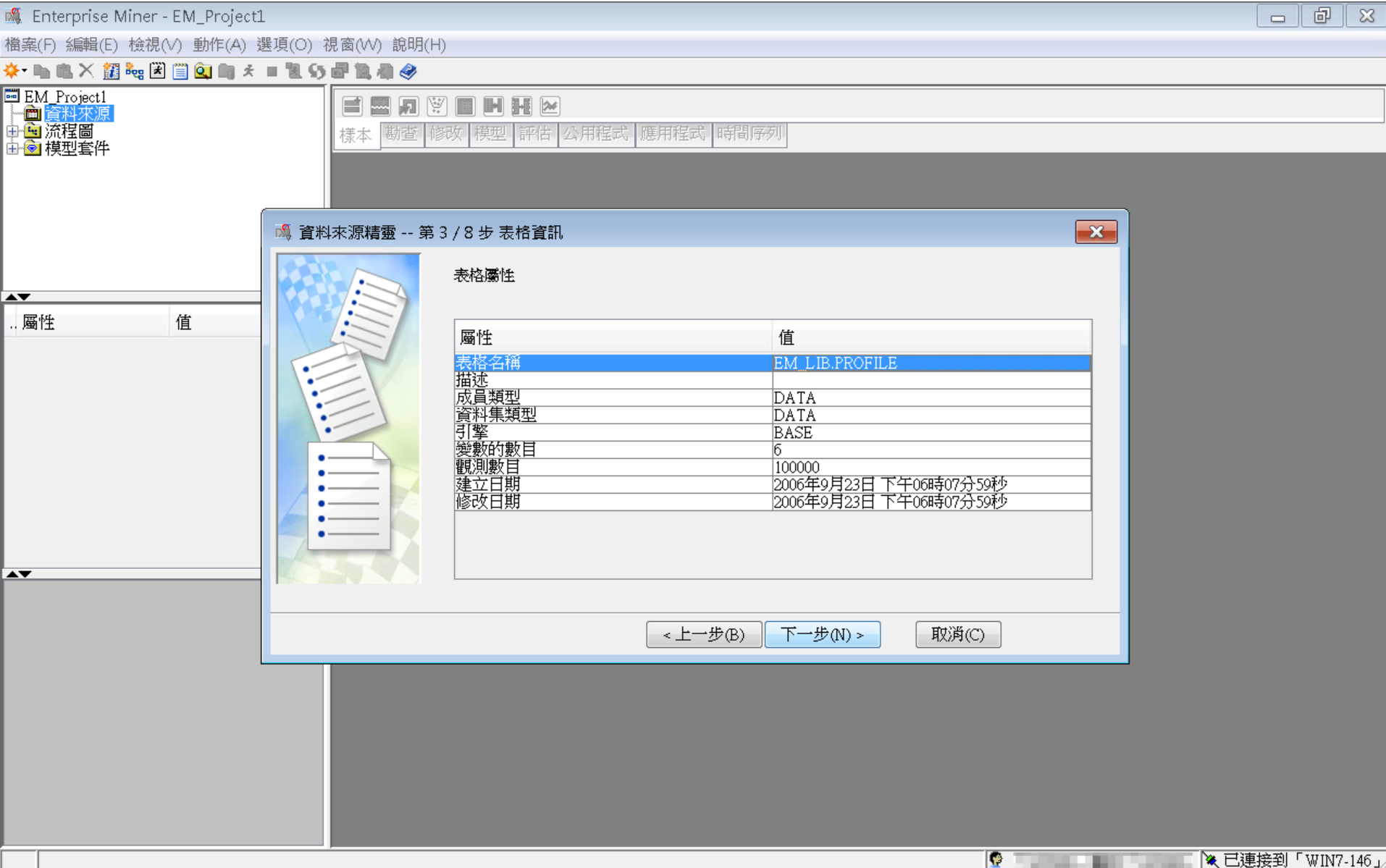
**EM\_LIB.PROFILE**

表格(T)  瀏覽(R)...

< 上一步(B) **下一步(N) >** 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)



Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 3 / 8 步 表格資訊

表格屬性

屬性	值
表格名稱	EM_LIB.PROFILE
描述	
成員類型	DATA
資料集類型	DATA
引擎	BASE
變數的數目	6
觀測數目	100000
建立日期	2006年9月23日 下午06時07分59秒
修改日期	2006年9月23日 下午06時07分59秒

< 上一步(B)    下一步(N) >    取消(C)

.. 屬性    值

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 4 / 8 步 中繼資料顧問選項

中繼資料顧問選項

使用「基本」設定時，以變數特性為基礎，設定初始量值層級和角色。

使用「進階」設定時，以變數特性和分配為基礎，設定初始量值層級和角色。

基本(S)  進階(A) 自訂(U)...

< 上一步(B) **下一步(N) >** 取消(C)

已連接到「WIN7-146」



# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
流程圖  
模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 5 / 8 步 欄中繼資料

(無)  非 等於

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	角色	層級	報表	順序	刪除	下限	上限
CNT_ATM	輸入	間隔	否		否	.	
CNT_CSC	輸入	間隔	否		否	.	
CNT_POS	輸入	間隔	否		否	.	
CNT_TBM	輸入	間隔	否		否	.	
CNT_TOT	輸入	間隔	否		否	.	
ID	ID	名目	否		否	.	

顯示程式碼(S) 勘查(E) 計算摘要(O) < 上一步(B) 下一步(N) > 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
流程圖  
模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 5 / 8 步 欄中繼資料

(無)  非 等於

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	角色	層級	報表	順序	刪除	下限	上限
CNT_ATM	輸入	間隔	否		否	.	
CNT_CSC	輸入	間隔	否		否	.	
CNT_POS	輸入				否	.	
CNT_TBM	輸入				否	.	
CNT_TOT	輸入				否	.	
ID	ID				否	.	

摘要統計

是否要計算摘要統計?

是(Y) 否(N)

顯示程式碼(S) 勘查(E) 計算摘要(O) < 上一步(B) 下一步(N) > 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM Project1

- 資料來源
- 流程圖
- 模型套件

樣本 調查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 5 / 8 步 欄中繼資料

(無)  非 等於

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	最小值	最大值	平均值	標準差	偏態	峰度
CNT_ATM	3	628	19.49971	20.8561	2.357293	15.33353
CNT_CSC	1	607	6.68411	12.12856	6.236494	108.4633
CNT_POS	2	345	11.9233	20.73384	3.343805	15.43941
CNT_TBM	10	14934	68.13696	101.1542	53.05219	5925.782
CNT_TOT	17	15225	106.2441	113.3704	39.2061	3976.09
ID	.	.	.	.	.	.

顯示程式碼(S) 調查(E) 重新整理摘要(U) < 上一步(B) 下一步(N) > 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 6 / 8 步 建立樣本

是否要建立樣本資料集?  
 否(O)  是(Y)

表格資訊

欄	6
列	100000

樣本大小

類型	百分比
百分比	20
列	

< 上一步(B) 下一步(N) > 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- 流程圖
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 7 / 8 步 資料來源特性

您可以變更名稱和角色，還可以為要建立的資料來源指定母體區段識別碼。

名稱(E): PROFILE

角色(R): 原始

區段(S):

附註(O):

< 上一步(B) 下一步(N) > 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM Project1

- 資料來源
- 流程圖
- 模型套件

樣本 調查 修改 模型 評估 公用程式 應用程式 時間序列

資料來源精靈 -- 第 8 / 8 步 摘要

中繼資料已完成。

資料館: EM\_LIB  
資料來源: PROFILE  
角色: 原始

角色	層級	計數
ID	名目	1
輸入	間隔	5

< 上一步(B) 完成(F) 取消(C)

已連接到「WIN7-146」

# Step 3. 建立資料來源 (Create Data Source)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
模型套件

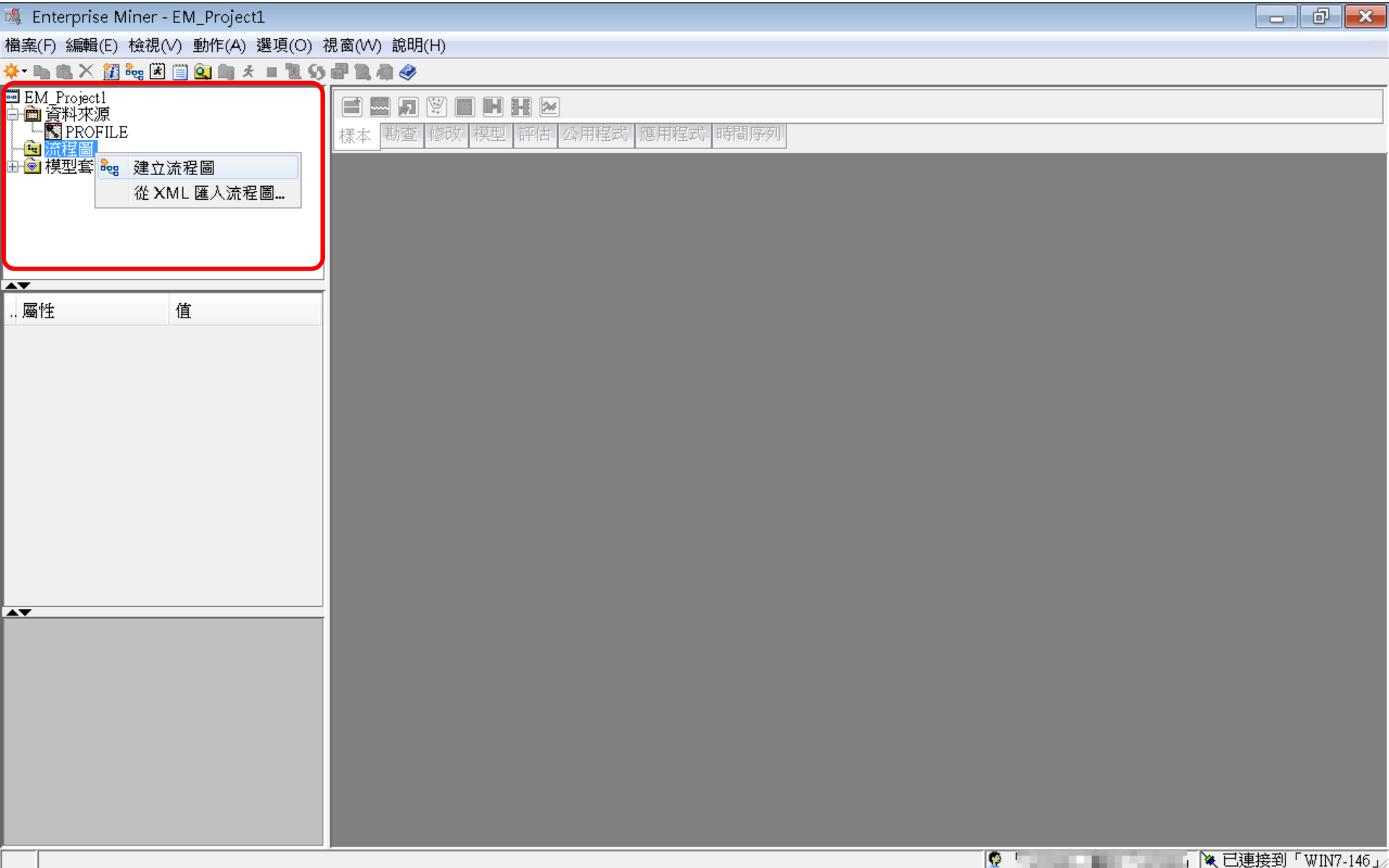
屬性	值
ID	profile
名稱	PROFILE
變數	
決策	
角色	原始
附註	
資料館	EM_LIB
表格	PROFILE
樣本資料集	
大小類型	
樣本大小	
類型	DATA
觀測數目	100000
欄數	6
位元組數	5653504
區段	

**ID**  
資料來源識別碼。與資料來源關聯的中繼資料表格儲存在 EMDS SAS 資料館中，並使用識別碼作為這些表格名稱的前置碼。

樣本 動查 修改 模型 評估 公用程式 應用程式 時間序列

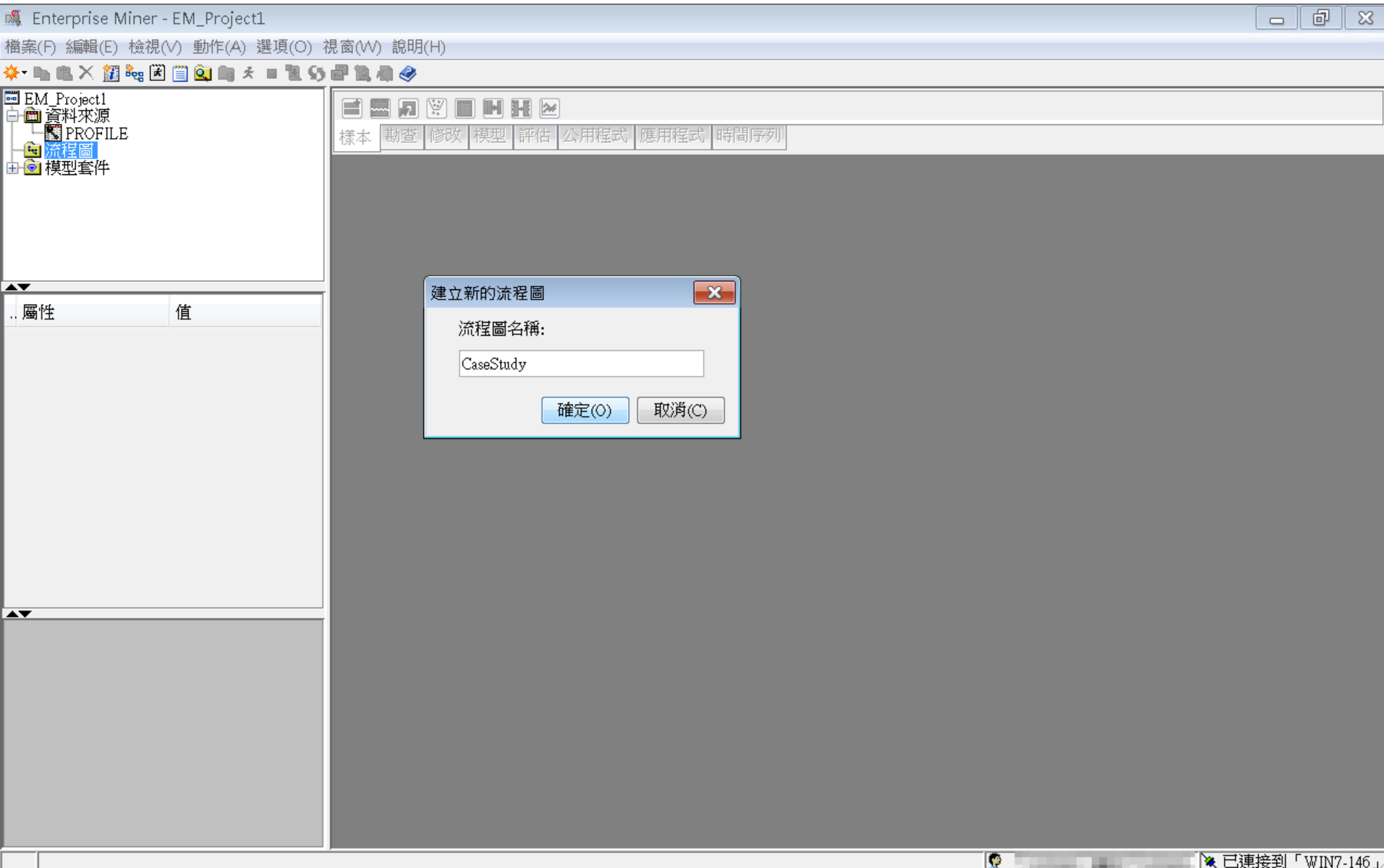
已連接到「WIN7-146」

# Step 4. 建立流程圖 (Create Diagram)





# Step 4. 建立流程圖 (Create Diagram)



# Step 4. 建立流程圖 (Create Diagram)

The screenshot displays the SAS Enterprise Miner interface. The main window is titled 'Enterprise Miner - EM\_Project1'. The menu bar includes '檔案(F)', '編輯(E)', '檢視(V)', '動作(A)', '選項(O)', '視窗(W)', and '說明(H)'. The toolbar contains various icons for file operations and editing. The left sidebar shows a project tree with 'EM\_Project1' expanded, containing '資料來源', 'PROFILE', '流程圖', 'CaseStudy', and '模型套件'. Below the tree is a table of properties for the selected 'CaseStudy' object.

屬性	值
ID	EMWS1
名稱	CaseStudy
狀態	開啟
附註	
歷程記錄	
編碼	ms-950 Traditional Chi
資料表示法	WINDOWS_32
原生作業系統	是

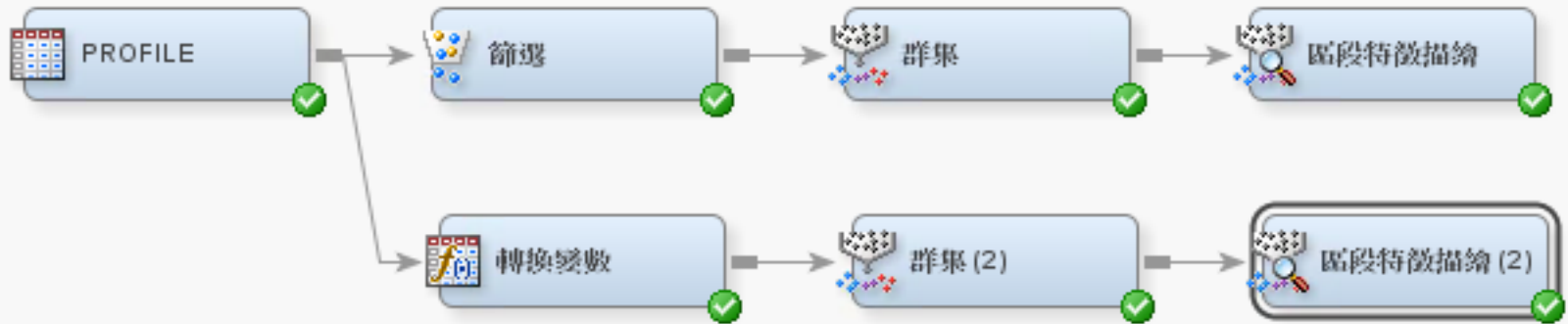
Below the table, the 'ID' section provides a description: '流程圖識別碼。此識別碼和 SAS 資料館名稱對應，SAS 資料館名稱是用於識別此流程圖內容在伺服器上的實體位置。'

The main workspace is titled 'CaseStudy' and contains a blank area for creating a flowchart. The toolbar above the workspace includes icons for '樣本', '勘查', '修改', '模型', '評估', '公用程式', '應用程式', and '時間序列'. The status bar at the bottom shows '流程圖' and '日誌' dropdown menus, and a system tray icon indicating '已連接到「WIN7-146」'.

# SAS Enterprise Miner (SAS EM) Case Study

- SAS EM 資料匯入4步驟
  - Step 1. 新增專案 (New Project)
  - Step 2. 新增資料館 (New / Library)
  - Step 3. 建立資料來源 (Create Data Source)
  - Step 4. 建立流程圖 (Create Diagram)
- SAS EM SEMMA 建模流程

# 案例情境模型流程





EM\_Project1

- 資料來源
- PROFILE**
- 流程圖
- CaseStudy
- 模型套件

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

PROFILE

屬性	值
ID	profile
名稱	PROFILE
變數	
決策	
角色	原始
附註	
資料館	EM_LIB
表格	PROFILE
樣本資料集	
大小類型	
樣本大小	
類型	DATA
觀測數目	100000
欄數	6
位元組數	5653504
區段	

**ID**

資料來源識別碼。與資料來源關聯的中繼資料表格儲存在 EMDS SAS 資料館中，並使用識別碼作為這些表格名稱的前置碼。

# EM\_Lib.Profile

	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34.0	3.0	3.0	9.0	49.0
2	000052094	44.0	17.0	5.0	18.0	84.0
3	000057340	122.0	26.0	32.0	36.0	216.0
4	000076885	42.0	3.0	6.0	1.0	52.0
5	000089150	20.0	15.0	2.0	2.0	39.0
6	000094512	83.0	20.0	49.0	3.0	155.0
7	000096396	33.0	9.0	49.0	17.0	108.0
8	000098062	22.0	5.0	9.0	1.0	37.0
9	000122779	21.0	27.0	2.0	1.0	51.0
10	000125838	127.0	3.0	2.0	3.0	135.0
11	000137317	52.0	20.0	2.0	11.0	85.0
12	000147896	101.0	40.0	8.0	1.0	150.0
13	000150206	156.0	52.0	80.0	4.0	292.0
14	000160501	31.0	32.0	42.0	12.0	117.0
15	000176820	107.0	3.0	2.0	1.0	113.0
16	000187784	54.0	14.0	7.0	24.0	99.0
17	000189200	92.0	3.0	2.0	2.0	99.0
18	000198327	44.0	27.0	17.0	3.0	91.0
19	000200237	146.0	3.0	2.0	1.0	152.0
20	000207980	39.0	5.0	2.0	4.0	50.0
21	000227355	180.0	3.0	2.0	10.0	195.0
22	000228317	46.0	14.0	2.0	1.0	63.0
23	000228906	25.0	3.0	2.0	1.0	31.0
24	000229984	15.0	3.0	2.0	1.0	21.0
25	000241022	13.0	3.0	2.0	1.0	19.0
26	000253260	65.0	3.0	2.0	1.0	71.0
27	000257910	35.0	3.0	2.0	1.0	41.0
28	000261717	11.0	5.0	2.0	1.0	19.0
29	000266600	15.0	3.0	2.0	1.0	21.0



EM\_Project1  
 資料來源  
 PROFILE  
 流程圖  
 CaseStudy  
 模型套件

.. 屬性	值
ID	profile
名稱	PROFILE
變數	
決策	
角色	原始
附註	
資料館	EM_LIB
表格	PROFILE
樣本資料集	
大小類型	
樣本大小	
類型	DATA
觀測數目	100000
欄數	6
位元組數	5653504
區段	

## ID

資料來源識別碼。與資料來源關聯的中繼資料表格儲存在 EMDS SAS 資料館中，並使用識別碼作為這些表格名稱的前置碼。



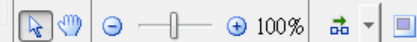
樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

篩選

根據指定的準則將資料從觀測中移除。



PROFILE



流程圖 | 日誌



EM\_Project1

- 資料來源
- PROFILE
- 流程圖
- CaseStudy
- 模型套件

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列



屬性	值
<b>一般</b>	
節點 ID	Filter
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
<b>類別變數</b>	
類別變數	
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01

一般

一般屬性

Navigation and zoom controls: back, forward, zoom in, zoom out, 100%, and other utility icons.

流程圖 日誌





- EM\_Project1
  - 資料來源
    - PROFILE
  - 流程圖
    - CaseStudy
  - 模型套件

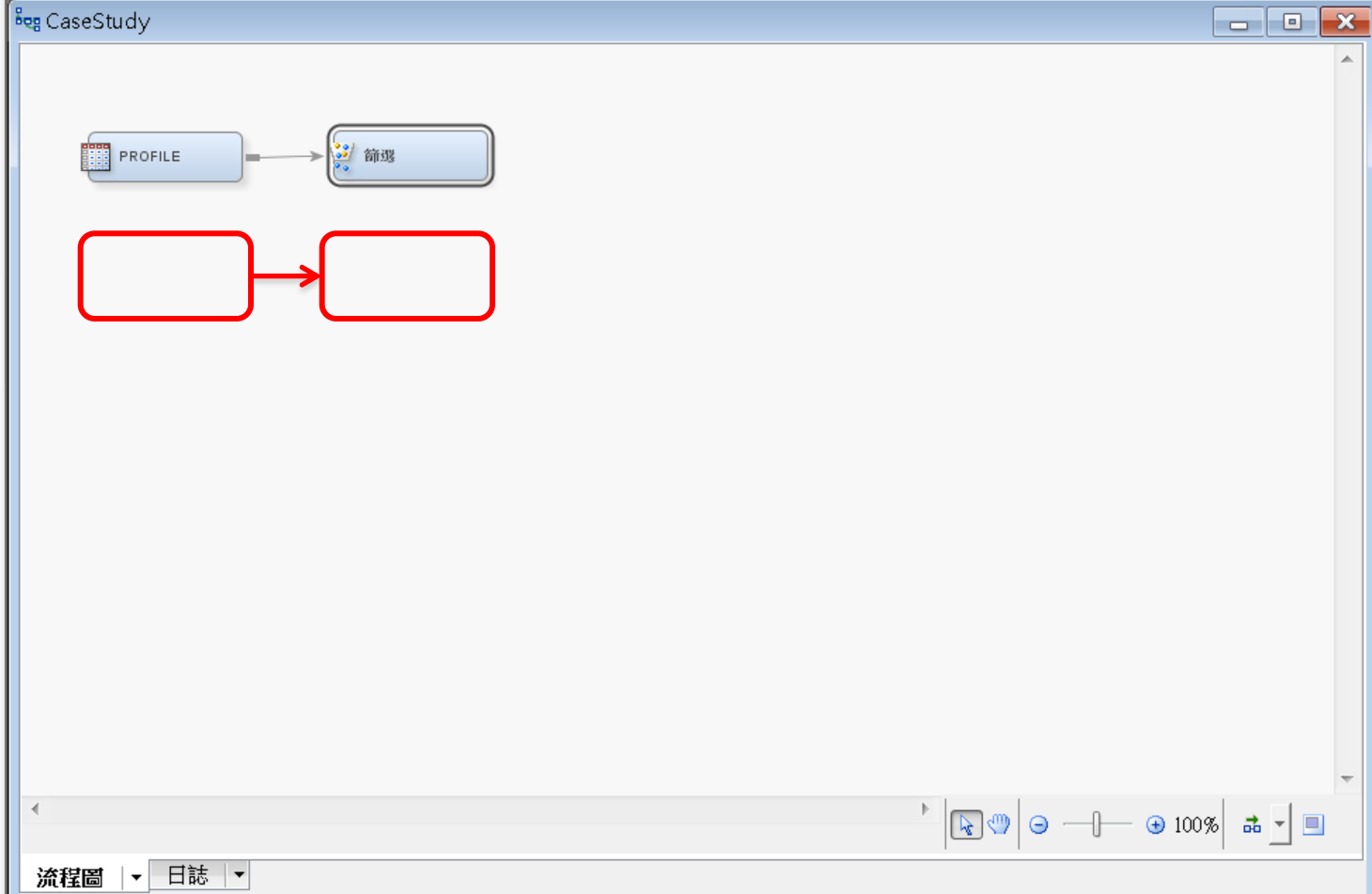
屬性	值
<b>一般</b>	
節點 ID	Filter
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
<b>類別變數</b>	
類別變數	
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01

一般

一般屬性

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列





EM\_Project1

- 資料來源
- PROFILE
- 流程圖
- CaseStudy
- 模型套件

屬性	值
<b>一般</b>	
節點 ID	Filter
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
<b>類別變數</b>	
類別變數	
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01
最大層級數目截止值	25
<b>間隔變數</b>	
間隔變數	
預設篩選方法	平均值的標準差
保留遺漏值	是
調整參數	
<b>評分</b>	
建立評分程式碼	是
更新異值層級	否

一般
一般屬性

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

更新  
執行  
建立模型套件...  
結果...  
將路徑匯出作為 SAS 程式

剪下  
複製(C)

刪除  
重新命名

全選  
選取節點  
連接節點  
中斷連接節點

100%

流程圖 | 日誌



EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy



屬性	值
<b>一般</b>	
節點 ID	Filter
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
<b>類別變數</b>	
類別變數	
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01
最大層級數目截止值	25
<b>間隔變數</b>	
間隔變數	
預設篩選方法	平均值的標準差
保留遺漏值	是
調整參數	
<b>評分</b>	
建立評分程式碼	是
更新具值層級	否

**執行狀態**

執行完成  
流程圖: CaseStudy  
路徑: 篩選

確定 結果...

一般  
一般屬性

流程图 日誌 100%

# 篩選一間隔變數 ...

The screenshot shows the Enterprise Miner interface. The workflow diagram in the center features a 'PROFILE' task connected to a '篩選' (Filter) task, both marked with green checkmarks. The '篩選' task is highlighted with a red box. Below the workflow, a configuration table for the '篩選' task is displayed, with a red dashed box around it. The table lists various parameters and their values. The '調整參數' (Adjust Parameters) row is highlighted with a red box, and its configuration options are also highlighted with a red dashed box in the left-hand pane.

屬性	值
一般	
節點 ID	Filter
匯入的資料	
匯出的資料	
附註	
訓練	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
類別變數	
類別變數	
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01
最大層級數目截止值	25
間隔變數	
間隔變數	
預設篩選方法	平均值的標準差
保留遺漏值	是
調整參數	
評分	
建立評分程式碼	是
更新量值層級	否
狀態	
一般	
一般屬性	

執行完成



EM\_Project1  
 資料來源  
 PROFILE  
 流程圖  
 CaseStudy  
 模型套件

屬性	值
<b>一般</b>	
節點 ID	Filter
匯入的資料	...
匯出的資料	...
附註	...
<b>訓練</b>	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
<b>類別變數</b>	
類別變數	...
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01
最大層級數目截止值	25
<b>間隔變數</b>	
間隔變數	...
預設篩選方法	平均值的標準差
保留遺漏值	是
調整參數	...
<b>評分</b>	
建立評分程式碼	是
更新量值層級	否

一般

一般屬性

執行完成

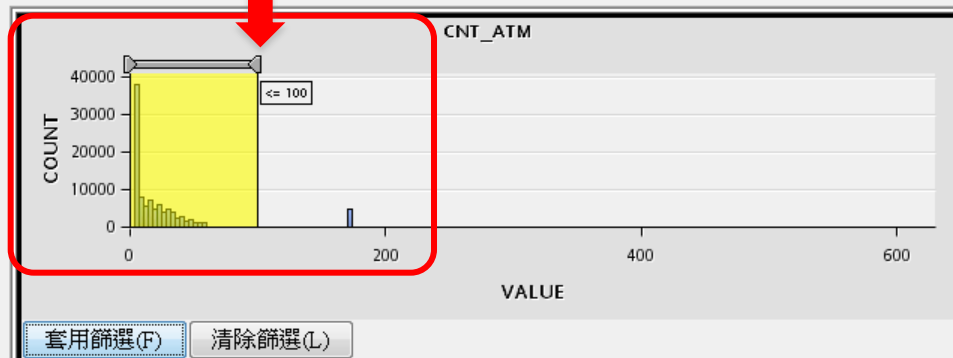


樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy



互動式間隔篩選



套用篩選(F)

清除篩選(L)

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	報表	篩選方法	保留遺漏值	篩選下限	篩選上限	角色	層級
CNT_ATM	否	使用者指定	預設	0	100	輸入	間隔
CNT_CSC	否	使用者指定	預設	0	30	輸入	間隔
CNT_POS	否	使用者指定	預設	0	60	輸入	間隔
CNT_TBM	否	使用者指定	預設	0	321	輸入	間隔
CNT_TOT	否	使用者指定	預設	0	441	輸入	間隔

重新整理摘要(R)

確定(O)

取消(C)

# 篩選一匯入的資料 ... [勘查]

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy

PROFILE → 篩選

匯入的資料 - 篩選

連接埠	來源	表格	角色	資料存在
DATA		EMWS1.Ids_DATA	原始	是
VALIDATE				否
TEST				否

瀏覽(B)... 勘查(X)... 屬性(P)... 確定(O)

屬性 值

一般

節點 ID Filter

匯入的資料

匯出的資料

附註

訓練

匯出表格 已篩選

篩選表格 訓練資料

分配資料集 是

類別變數

類別變數

預設篩選方法 稀有值 (百分比)

保留遺漏值 是

正規化值 是

最小次數截止值 1

最小百分比截止值 0.01

最大層級數目截止值 25

間隔變數

間隔變數

預設篩選方法 平均值的標準差

保留遺漏值 是

調整參數

評分

建立評分程式碼 是

更新量值層級 否

狀態

匯入的資料

節點匯入的表格集

執行完成

流程圖 日誌

# 篩選一匯入的資料

EMWS1Ids\_DATA

檔案(F) 檢視(V) 動作(A) 視窗(W)

樣本屬性

屬性	值
列	未知
欄	6
資料館	EMWS1
成員	IDS_DATA
類型	VIEW
抽樣方法	前 N 列
提取大小	最大
已擷取的列	6000
隨機種子	12345

樣本統計資料

觀測 #	變數名稱	標籤	類型	遺漏值...	最小值	最大值	平均值	層級數目	眾數百...	眾數
1	ID		CLASS	0				128+	0.775194	000041360
2	CNT_ATM		VAR	0	3	250	19.02667			
3	CNT_CSC		VAR	0	1	328	6.6545			
4	CNT_POS		VAR	0	2	212	11.83617			
5	CNT_TBM		VAR	0	10	2741	68.1285			
6	CNT_TOT		VAR	0	17	2747	105.6458			

EMWS1Ids\_DATA

觀測 #	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	00041360	34	3	3	9	49
2	00052094	44	17	5	18	84
3	00057340	122	26	32	36	216
4	00076885	42	3	6	1	52
5	00089150	20	15	2	2	39
6	00094512	83	20	49	3	155
7	00096396	33	9	49	17	108
8	00098062	22	5	9	1	37
9	00122779	21	27	2	1	51
10	00125838	127	3	2	3	135
11	00137317	52	20	2	11	85
12	00147896	101	40	8	1	150
13	00150206	156	52	80	4	292
14	00160501	31	32	42	12	117
15	00176820	107	3	2	1	113
16	00187784	54	14	7	24	99
17	00189200	92	3	2	2	99
18	00198327	44	27	17	3	91

# 篩選一匯出的資料... [勘查]

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
- PROFILE
- 流程圖
- CaseStudy
- 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy

PROFILE → 篩選

匯出的資料 - 篩選

連接埠	表格	角色	資料存在
TRAIN	EMWS1.Filter_TRAIN	訓練	是
VALIDATE	EMWS1.Filter_VALIDATE	驗證	否
TEST	EMWS1.Filter_TEST	測試	否

瀏覽(B)... 勘查(X)... 屬性(P)... 確定(O)

屬性 值

一般

- 節點 ID: Filter
- 匯入的資料
- 匯出的資料
- 附註
- 訓練
- 匯出表格: 已篩選
- 篩選表格: 訓練資料
- 分配資料集: 是
- 類別變數
- 類別變數
- 預設篩選方法: 稀有值 (百分比)
- 保留遺漏值: 是
- 正規化值: 是
- 最小次數截止值: 1
- 最小百分比截止值: 0.01
- 最大層級數目截止值: 25
- 間隔變數
- 間隔變數
- 預設篩選方法: 平均值的標準差
- 保留遺漏值: 是
- 調整參數
- 評分
- 建立評分程式碼: 是
- 更新量值層級: 否
- 狀態

匯出的資料

節點匯出的表格彙集

執行完成

流程圖 日誌

100%



# 篩選一匯出的資料

EMWS1.Filter\_TRAIN

檔案(F) 檢視(V) 動作(A) 視窗(W)

樣本屬性

屬性	值
列	未知
欄	6
資料館	EMWS1
成員	FILTER_TRAIN
類型	VIEW
抽樣方法	前 N 列
提取大小	最大
已擷取的列	60000
隨機種子	12345

樣本統計資料

觀測 #	變數名稱	標籤	類型	遺漏值...	最小值	最大值	平均值	層級數目	眾數百...	眾數
1	ID		CLASS	0				128+	0.775194	000041360
2	CNT_ATM		VAR	0	3	82	17.48542			
3	CNT_CSC		VAR	0	1	43	5.217783			
4	CNT_POS		VAR	0	2	74	9.300717			
5	CNT_TBM		VAR	0	10	371	62.84018			
6	CNT_TOT		VAR	0	17	445	94.8441			

EMWS1.Filter\_TRAIN

觀測 #	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34	3	3	9	49
2	000052094	44	17	5	18	84
3	000057340	122	26	32	36	216
4	000076885	42	3	6	1	52
5	000089150	20	15	2	2	39
6	000094512	83	20	49	3	155
7	000096396	33	9	49	17	108
8	000098062	22	5	9	1	37
9	000122779	21	27	2	1	51
10	000125838	127	3	2	3	135
11	000137317	52	20	2	11	85
12	000147896	101	40	8	1	150
13	000160501	31	32	42	12	117
14	000176820	107	3	2	1	113
15	000187784	54	14	7	24	99
16	000189200	92	3	2	2	99
17	000198327	44	27	17	3	91
18	000200237	146	3	2	1	152

# 勘查 (Explore) — 群集 (Cluster)

The screenshot displays the Enterprise Miner software interface. The main window shows a workflow diagram with a 'PROFILE' node connected to a '篩選' (Filter) node. The 'Explore' (勘查) step is highlighted with a red box in the toolbar. A tooltip for the 'Cluster' (群集) icon indicates it is used for segmenting data for observation.

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性 值

屬性	值
一般	
節點 ID	Filter
匯入的資料	
匯出的資料	
附註	
訓練	
匯出表格	已篩選
篩選表格	訓練資料
分配資料集	是
類別變數	
類別變數	
預設篩選方法	稀有值 (百分比)
保留遺漏值	是
正規化值	是
最小次數截止值	1
最小百分比截止值	0.01
最大層級數目截止值	25
間隔變數	
間隔變數	
預設篩選方法	平均值的標準差
保留遺漏值	是
調整參數	
評分	
建立評分程式碼	是
更新量值層級	否
狀態	

匯出的資料  
節點匯出的表格案。

執行完成

樣本 修改 模型 評估 公用程式 應用程式 時間序列

群集  
執行分段資料庫可以使用的觀測群集。

樣本 勘查 修改 模型 評估 公用程式

流程圖 日誌



- EM\_Project1
  - 資料來源
  - PROFILE
  - 流程圖
  - CaseStudy
  - 模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy

```
graph LR; PROFILE[PROFILE] --> 篩選[篩選]; 篩選 --> 群集[群集];
```

- 編輯變數...
- 更新
- 執行**
- 建立模型套件...
- 結果...
- 將路徑匯出作為 SAS 程式
- 剪下
- 複製(C)
- 刪除
- 重新命名
- 全選
- 選取節點
- 連接節點
- 中斷連接節點

100%

流程圖 日誌

屬性	值
<b>一般</b>	
節點 ID	Clus
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	自動
最大的群集數目	10
選取準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
<b>類別變數編碼</b>	
序數編碼	排名
名目編碼	GLM
<b>初始群集種子</b>	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	

一般
一般屬性



EM\_Project1  
 資料來源  
 PROFILE  
 流程圖  
 CaseStudy  
 模型套件

屬性	值
<b>一般</b>	
節點 ID	Clus
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	自動
最大的群集數目	10
選取準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	
序數編碼	排名
名目編碼	GLM
初始群集種子	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	

一般

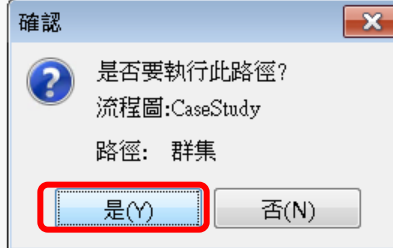
一般屬性

執行完成



樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy



流程圖 | 日誌





- EM\_Project1
  - 資料來源
    - PROFILE
  - 流程圖
    - CaseStudy
  - 模型套件

屬性	值
<b>一般</b>	
節點 ID	Clus
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	自動
最大的群集數目	10
選取準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	
序數編碼	排名
名目編碼	GLM
初始群集種子	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	

一般  
一般屬性

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy



Navigation icons: back, forward, zoom in, zoom out, 100%, refresh, print.

流程圖 | 日誌



EM\_Project1  
 資料來源  
 PROFILE  
 流程圖  
 CaseStudy  
 模型套件

屬性	值
<b>一般</b>	
節點 ID	Clus
匯入的資料	...
匯出的資料	...
附註	...
<b>訓練</b>	
變數	...
群集變數角色	區段
內部標準化	標準化
<b>群集的數目</b>	
指定方法	自動
最大的群集數目	10
<b>選取準則</b>	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
<b>類別變數編碼</b>	
序數編碼	排名
名目編碼	GLM
<b>初始群集種子</b>	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	

一般

一般屬性

執行完成



樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy



執行狀態



執行完成

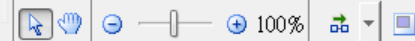
流程圖:CaseStudy

路徑: 群集

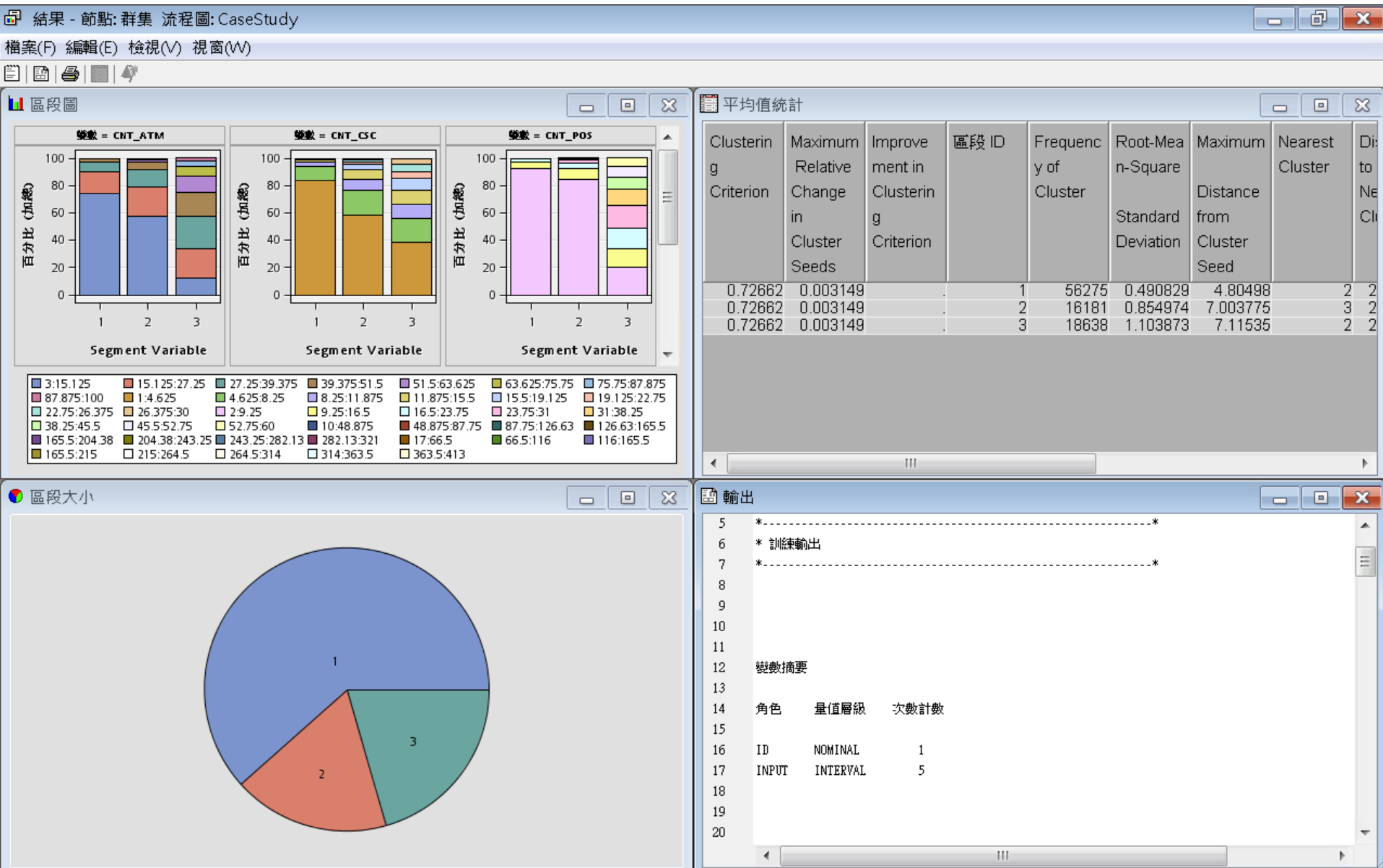
確定

結果...

流程圖 | 日誌



# 群集 (Cluster) 結果



# 評估 (Assess) — 區段特徵描繪 (Segment Profile)

The screenshot displays the Enterprise Miner software interface. The main window shows a workflow diagram with three steps: 'PROFILE', '篩選' (Filter), and '群集' (Cluster). The 'Assess' (評估) step is highlighted in red. A tooltip for the 'Assess' step reads: '區段特徵描繪 使用「區段特徵描繪」節點檢查分段或群集的资料，並識別母體中資料區段不同的因子。' (Segment Profile: Use the 'Segment Profile' node to check segmented or clustered data, and identify factors that differentiate data segments in the population.)

Below the workflow diagram, a toolbar contains several icons. The 'Assess' icon, which depicts a cluster of data points, is circled in red. Below the toolbar, a row of buttons labeled '樣本' (Sample), '勘查' (Explore), '修改' (Modify), '模型' (Model), and '評估' (Assess) is shown. The '評估' button is also circled in red.

On the left side of the interface, a table lists various properties and their values:

屬性	值
一般	
節點 ID	Clus
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	自動
最大的群集數目	10
選取準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	
序數編碼	排名
名目編碼	GLM
初始群集種子	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
訓練選項	
一般	
一般屬性	

At the bottom left, the status bar indicates '執行完成' (Execution completed).



# 評估 (Assess) — 區段特徵描繪 (Segment Profile)

The screenshot displays the Enterprise Miner software interface. The main window is titled "Enterprise Miner - EM\_Project1". The menu bar includes "檔案(F)", "編輯(E)", "檢視(V)", "動作(A)", "選項(O)", "視窗(W)", and "說明(H)". The toolbar contains various icons for file operations and analysis. The left sidebar shows a project tree with "EM\_Project1" containing "資料來源", "PROFILE", "流程圖", "CaseStudy", and "模型套件". Below the tree is a properties table.

屬性	值
一般	
節點 ID	Prof
匯入的資料	...
匯出的資料	...
附註	...
訓練	
變數	...
一般	
中間點數目	8
所有的特徵描繪	否
截止百分比	95
輸入變數	
輸入的數目	10
最小價值	0.01
最大深度	1
列印價值統計	是
目標變數	
分析角色	無
報表變數	
使用報表變數	是
報表變數的數目	10
狀態	
一般	
一般屬性	

The main workspace shows a workflow diagram with four nodes: "PROFILE", "篩選", "群集", and "區段特徵描繪". A red arrow points from a red box around the "評估" (Assess) icon in the toolbar to the "區段特徵描繪" node in the workflow. The status bar at the bottom indicates "執行完成" (Execution Complete).

# 評估 (Assess) — 區段特徵描繪 (Segment Profile)

The screenshot shows the Enterprise Miner interface. The main workspace displays a workflow diagram with four nodes: 'PROFILE', '篩選', '群集', and '區段特徵描繪'. The '區段特徵描繪' node is highlighted with a red box, and its context menu is open, with the '執行' (Execute) option also highlighted in red. The context menu includes options like '更新', '建立模型套件...', '結果...', '將路徑匯出作為 SAS 程式', '剪下', '複製(C)', '刪除', '重新命名', '全選', '選取節點', '連接節點', and '中斷連接節點'.

On the left, the '屬性' (Properties) pane shows the following table:

屬性	值
一般	
節點 ID	Prof
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
一般	
中間點數目	8
所有的特徵描繪	否
截止百分比	95
輸入變數	
輸入的數目	10
最小價值	0.01
最大深度	1
列印價值統計	是
目標變數	
分析角色	無
報表變數	
使用報表變數	是
報表變數的數目	10
狀態	
一般	
一般屬性	

At the bottom left, the status bar indicates '執行完成' (Execution completed).

# 評估 (Assess) — 區段特徵描繪 (Segment Profile)

The screenshot displays the Enterprise Miner software interface. The main window shows a workflow diagram with four steps: PROFILE, 篩選 (Filter), 群集 (Cluster), and 區段特徵描繪 (Segment Profile). The '區段特徵描繪' step is highlighted with a green border. Below the workflow, a table lists various properties and their values.

屬性	值
<b>一般</b>	
節點 ID	Prof
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
<b>一般</b>	
中間點數目	8
所有的特徵描繪	否
截止百分比	95
<b>輸入變數</b>	
輸入的數目	10
最小價值	0.01
最大深度	1
列印價值統計	是
<b>目標變數</b>	
分析角色	無
<b>報表變數</b>	
使用報表變數	是
報表變數的數目	10
<b>狀態</b>	
<b>一般</b>	
一般屬性	

正在執行「區段特徵描繪」

# 評估 (Assess) — 區段特徵描繪 (Segment Profile)

The screenshot displays the Enterprise Miner interface. On the left, a tree view shows the project structure: EM\_Project1, 資料來源 (Data Sources), PROFILE, 流程圖 (Flowchart), CaseStudy, and 模型套件 (Model Kits). Below this is a table of properties for the selected node.

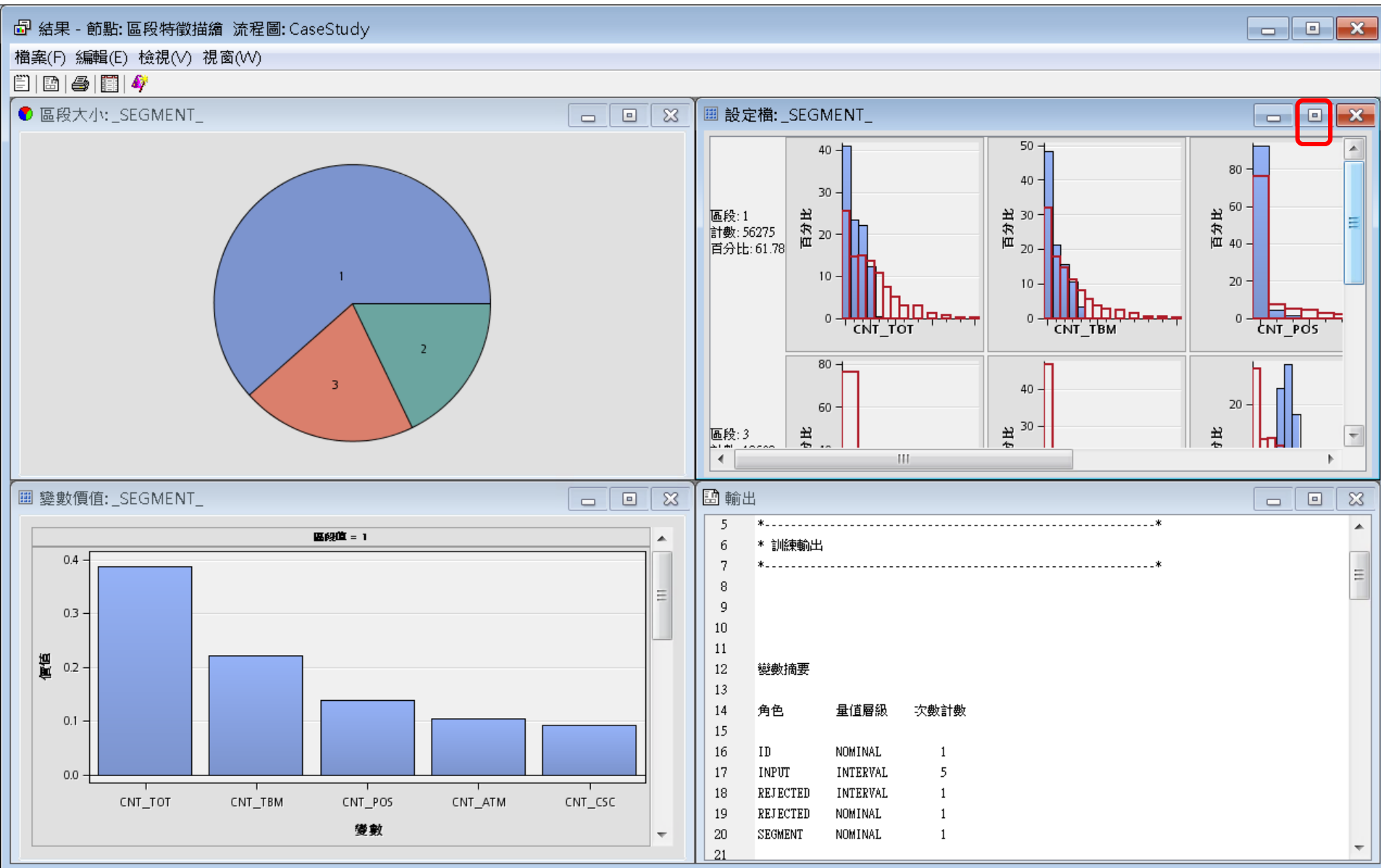
屬性	值
一般	
節點 ID	Prof
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
一般	
中間點數目	8
所有的特徵描繪	否
截止百分比	95
輸入變數	
輸入的數目	10
最小價值	0.01
最大深度	1
列印價值統計	是
目標變數	
分析角色	無
報表變數	
使用報表變數	是
報表變數的數目	10
狀態	
一般	
一般屬性	

The main workspace shows a workflow diagram with four steps: PROFILE, 篩選 (Filter), 群集 (Cluster), and 區段特徵描繪 (Segment Profile). All steps have green checkmarks. A dialog box titled "執行狀態" (Execution Status) is open, displaying the following information:

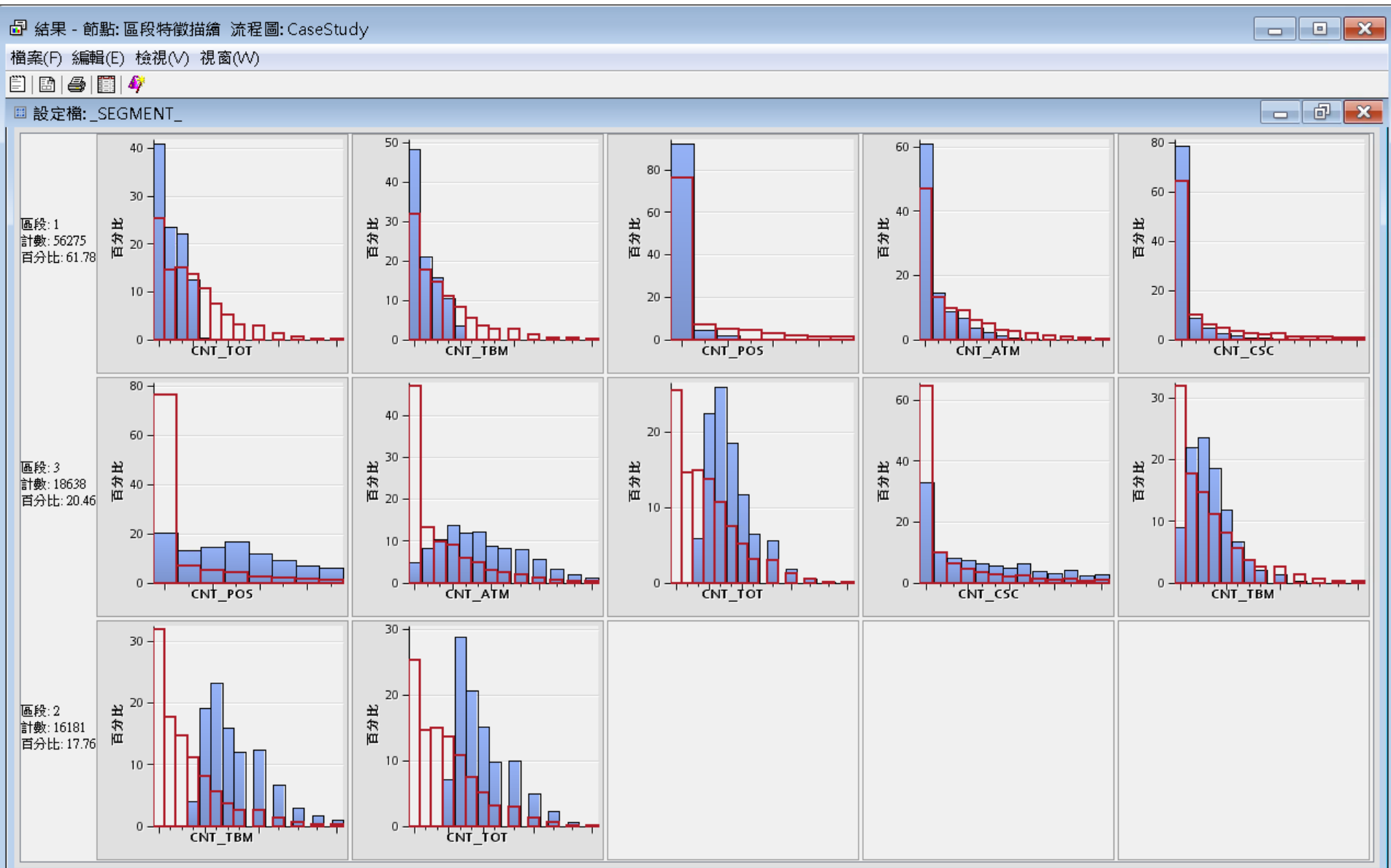
- 執行完成 (Execution Complete)
- 流程圖: CaseStudy (Flowchart: CaseStudy)
- 路徑: 區段特徵描繪 (Path: Segment Profile)
- Buttons: 確定 (OK), 結果... (Results...)

The "結果..." button is highlighted with a red rectangle. The status bar at the bottom indicates "執行完成" (Execution Complete).

# 區段特徵描繪 (Segment Profile)



# 區段特徵描繪 (Segment Profile)



# 轉換變數 (Transform Variables)

The screenshot displays the Enterprise Miner software interface. The main window shows a workflow diagram with four steps: PROFILE, 篩選 (Filter), 群集 (Cluster), and 區段特徵描繪 (Segment Feature Description). The 'Transform Variables' step is highlighted in red in the top toolbar. A tooltip for this step reads '將轉換套用至資料集變數。' (Apply transformation to dataset variables).

Below the workflow diagram, the text '轉換變數 (Transform Variables)' is written in red. Underneath, a row of icons represents different workflow stages: 樣本 (Sample), 勘查 (Explore), 修改 (Modify), 模型 (Model), and 評估 (Evaluate). The '修改' (Modify) icon is highlighted with a red box.

On the left side, a table lists various properties and their values:

屬性	值
一般	
節點 ID	Prof
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
一般	
中間點數目	8
所有的特徵描繪	否
截止百分比	95
輸入變數	
輸入的數目	10
最小價值	0.01
最大深度	1
列印價值統計	是
目標變數	
分析角色	無
報表變數	
使用報表變數	是
報表變數的數目	10
狀態	
一般	
一般屬性	

At the bottom of the interface, the status bar shows '執行完成' (Execution completed).

# 轉換變數 (Transform Variables)

## 訓練/公式 ...

The screenshot displays the Enterprise Miner interface. The main workflow window shows a sequence of steps: PROFILE, 篩選 (Filter), 群集 (Cluster), and 區段特徵描繪 (Segment Feature Description). The '轉換變數' (Transform Variables) step is highlighted with a red box, and a red arrow points to its icon in the top toolbar.

The configuration panel for the '訓練' (Train) step is shown below, enclosed in a red dashed box. It lists various options with their current values:

屬性	值
一般	
節點 ID	Trans
匯入的資料	...
匯出的資料	...
附註	...
訓練	
變數	...
公式	...
交互作用	...
SAS 程式碼	...
預設方法	
間隔輸入	無
間隔目標	無
類別輸入	無
類別目標	無
將遺漏值處理為層級	否
樣本屬性	
方法	前 N 個
大小	預設
隨機種子	12345
最佳等距分箱	
等距箱數的數目	4
遺漏值	搜尋時使用
群組方法	
截止值	0.1
群組遺漏	否
等距箱數的數目	變數
將最小值增加到位移是	
位移值	1



# 轉換變數 (Transform Variables)

## 訓練/公式 ... [建立]

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性 值

一般  
節點 ID Trans  
匯入的資料  
匯出的資料  
附註  
訓練  
變數  
公式  
交互作用  
SAS 程式碼  
預設方法  
間隔輸入 無  
間隔目標 無  
類別輸入 無  
類別目標 無  
將遺漏值處理為層級否  
樣本屬性  
方法 前 N 個  
大小 預設  
隨機種子 12345  
最佳等距分箱  
等距箱數的數目 4  
遺漏值 搜尋時使用  
群組方法  
截止值 0.1  
群組遺漏 否  
等距箱數的數目 變數  
將最小值增加到位移是  
位移值 1

執行完成

公式

建立

次數

CNT\_ATM

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
----	----	----	----	----	----	----	----	----

輸出 樣本 日誌

預覽(P) 確定(O) 取消(C)

# 轉換變數 (Transform Variables)

## 訓練/公式 ... [建立] 增加轉換 (衍生變數)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

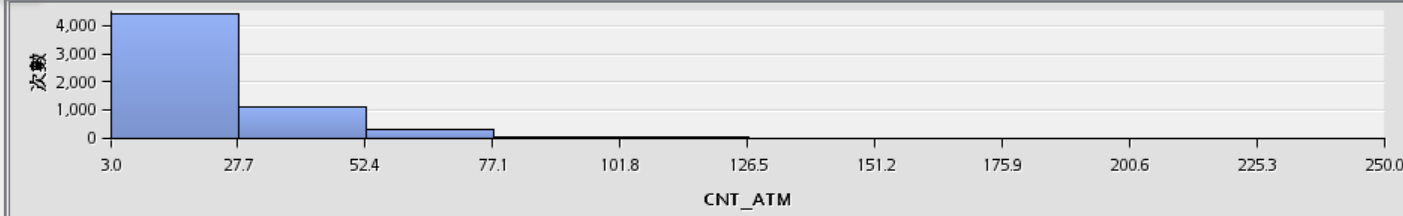
EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性 值

一般  
節點 ID Trans  
匯入的資料  
匯出的資料  
附註  
訓練  
變數  
公式  
交互作用  
SAS 程式碼  
預設方法  
間隔輸入 無  
間隔目標 無  
類別輸入 無  
類別目標 無  
將遺漏值處理為層級否  
樣本屬性  
方法 前 N 個  
大小 預設  
隨機種子 12345  
最佳等距分箱  
等距箱數的數目 4  
遺漏值 搜尋時使用  
群組方法  
截止值 0.1  
群組遺漏 否  
等距箱數的數目 變數  
將最小值增加到位移是  
位移值 1

執行完成

公式



次數

CNT\_ATM

欄:  標籤(A)  採礦(M)  基本(B)  統計(T)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

輸入

名稱	類型	長度	格式	層級
----	----	----	----	----

輸出 樣本 日誌

預覽(P)

增加轉換

屬性 值

名稱	TRANS_0
類型	數值
長度	8
格式	
層級	間隔
標籤	

公式:

TRANS\_0 =

建置... 確定(O) 取消(C)

確定(O) 取消(C)

# 轉換變數 (Transform Variables)

將 偏態 (Skewed input) →

取 LOG 值轉換變數 → 常態 (Normal Distribution)

*Category Logit Score =*

$$\log \left( \frac{\text{transaction count}_{\text{in category}}}{\text{transaction count}_{\text{out of category}}} \right)$$

LGT\_ATM =

LOG (CNT\_ATM /

(CNT\_TOT - CNT\_ATM))

# 轉換變數 (Transform Variables)

## 訓練/公式 ... [建立] 增加轉換 (衍生變數)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

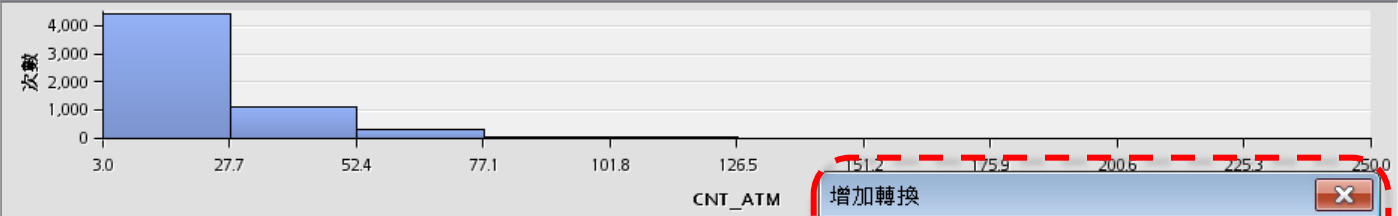
EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性 值

一般  
節點 ID Trans  
匯入的資料  
匯出的資料  
附註  
訓練  
變數  
公式  
交互作用  
SAS 程式碼  
預設方法  
間隔輸入 無  
間隔目標 無  
類別輸入 無  
類別目標 無  
將遺漏值處理為層級否  
樣本屬性  
方法 前 N 個  
大小 預設  
隨機種子 12345  
最佳等距分箱  
等距箱數的數目 4  
遺漏值 搜尋時使用  
群組方法  
截止值 0.1  
群組遺漏 否  
等距箱數的數目 變數  
將最小值增加到位移是  
位移值 1

執行完成

公式



次數

CNT\_ATM

欄:  標籤(A)  採礦(M)  基本(B)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

輸入

增加轉換

屬性	值
名稱	LGT_ATM
類型	數值
長度	8
格式	
層級	間隔
標籤	
角色	

公式:

```
LGT_ATM =  
LOG(CNT_ATM / (CNT_TOT - CNT_ATM))
```

建置... 確定(O) 取消(C)

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
LGT_ATM	數值	8		間隔	$LGT\_ATM = \log(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$			

輸出 樣本 日誌

預覽(P) 確定(O) 取消(C)

**$LGT\_ATM = \log(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$**

# 轉換變數 (Transform Variables)

## 運算式產生器

The screenshot displays the Enterprise Miner interface with the 'Transform Variables' dialog box open. The 'Formula Generator' window is also visible, showing the formula  $LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$  being applied to the variable 'LGT\_ATM'. The 'LGT\_ATM' variable is highlighted in the '編輯轉換' (Edit Transformation) window, and its properties are shown in the table below.

名稱	值
名稱	LGT_ATM
類型	數值
長度	8
格式	
層級	間隔
標籤	

The '公式' (Formula) field in the '編輯轉換' window contains the following formula:

$$LGT\_ATM = LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$$

The '變數清單' (Variable List) window shows the following table:

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0

The 'LGT\_ATM' variable is highlighted in the table. The '公式' (Formula) field in the '編輯轉換' window contains the following formula:

$$LGT\_ATM = LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$$

The '屬性' (Properties) window shows the following table:

名稱	值
名稱	LGT_ATM
類型	數值
長度	8
格式	
層級	間隔
標籤	

The '公式' (Formula) field in the '屬性' window contains the following formula:

$$LGT\_ATM = LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$$

The '輸出' (Output) window shows the following table:

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
LGT_ATM	數值	8		間隔	$LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$		輸入	否

The '輸出' (Output) window also shows the following formula:

$$LGT\_ATM = LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))$$

# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性 值

屬性	值
一般	
節點 ID	Trans
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
公式	
交互作用	
SAS 程式碼	
預設方法	

公式

次數

CNT\_ATM

欄:  標籤(A)  採礦(M)  基本(I)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

編輯轉換

屬性	值
名稱	LGT_TBM
類型	數值
長度	8
格式	
層級	間隔
標籤	
角色	

公式:

LGT\_TBM =  
LOG(CNT\_TBM / (CNT\_TOT - CNT\_TBM))

建置... 確定(O) 取消(C)

遺漏值 搜尋時使用

詳細方法

截止值 0.1

詳細遺漏 否

等距箱數的數目 變數

將最小值增加到位移是

位移值 1

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
LGT_ATM	數值	8		間隔	LOG(CNT_ATM / (CNT_TOT - CNT_ATM))		輸入	否
LGT_CSC	數值	8		間隔	LOG(CNT_CSC / (CNT_TOT - CNT_CSC))		輸入	否
LGT_POS	數值	8		間隔	LOG(CNT_POS / (CNT_TOT - CNT_POS))		輸入	否
LGT_TBM	數值	8		間隔	LOG(CNT_TBM / (CNT_TOT - CNT_TBM))		輸入	否

輸出 樣本 日誌

預覽(P) 確定(O) 取消(C)

執行完成

# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

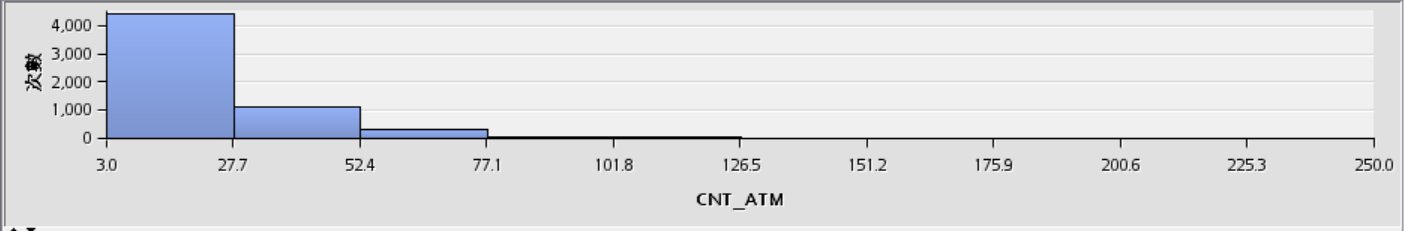
Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性	值
一般	
節點 ID	Trans
匯入的資料	
匯出的資料	
附註	
訓練變數	
公式	
交互作用	
SAS 程式碼	
預設方法	
間隔輸入	無
間隔目標	無
類別輸入	無
類別目標	無
將遺漏值處理為層級否	
樣本屬性	
方法	前 N 個
大小	預設
隨機種子	12345
最佳等距分箱	
等距箱數的數目	4
遺漏值	搜尋時使用
群組方法	
截止值	0.1
群組遺漏	否
等距箱數的數目	變數
將最小值增加到位移是	
位移值	1

公式



欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

輸入

預覽選取變數「LGT\_ATM」的繪圖。

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
LGT_ATM	數值	8		間隔	$\text{LOG}(\text{CNT\_ATM} / (\text{CNT\_TOT} - \text{CNT\_ATM}))$		輸入	否
LGT_CSC	數值	8		間隔	$\text{LOG}(\text{CNT\_CSC} / (\text{CNT\_TOT} - \text{CNT\_CSC}))$		輸入	否
LGT_POS	數值	8		間隔	$\text{LOG}(\text{CNT\_POS} / (\text{CNT\_TOT} - \text{CNT\_POS}))$		輸入	否
LGT_TBM	數值	8		間隔	$\text{LOG}(\text{CNT\_TBM} / (\text{CNT\_TOT} - \text{CNT\_TBM}))$		輸入	否

輸出 樣本 日誌

預覽(P) 確定(O) 取消(C)

執行完成

# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

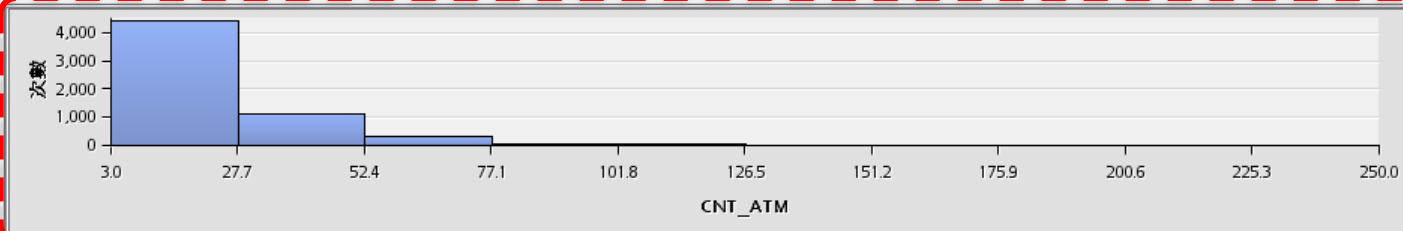
屬性 值

一般  
節點 ID Trans  
匯入的資料  
匯出的資料  
附註  
訓練  
變數  
公式  
交互作用  
SAS 程式碼  
預設方法  
間隔輸入 無  
間隔目標 無  
類別輸入 無  
類別目標 無  
將遺漏值處理為層級否  
樣本屬性  
方法 前 N 個  
大小 預設  
隨機種子 12345  
最佳等距分箱  
等距箱數的數目 4  
遺漏值 搜尋時使用  
群組方法  
截止值 0.1  
群組遺漏 否  
等距箱數的數目 變數  
將最小值增加到位移是  
位移值 1

執行完成

公式

偏態 (Skewed input)



次數

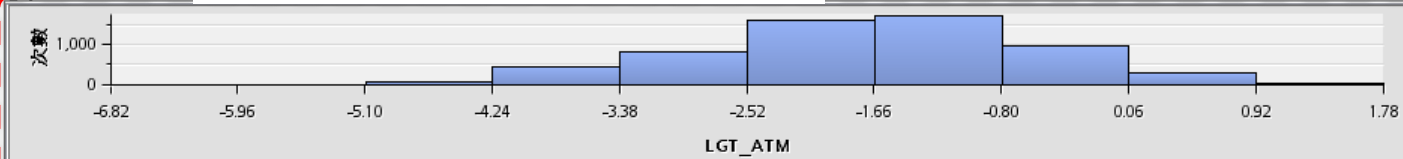
CNT\_ATM

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

LGT\_ATM = LOG  
(CNT\_ATM /  
(CNT\_TOT - CNT\_ATM))

常態 (Normal Distribution)



次數

LGT\_ATM

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
LGT_ATM	數值	8		間隔	LOG(CNT_ATM / (CNT_TOT - CNT_ATM))		輸入	否
LGT_CSC	數值	8		間隔	LOG(CNT_CSC / (CNT_TOT - CNT_CSC))		輸入	否
LGT_POS	數值	8		間隔	LOG(CNT_POS / (CNT_TOT - CNT_POS))		輸入	否
LGT_TBM	數值	8		間隔	LOG(CNT_TBM / (CNT_TOT - CNT_TBM))		輸入	否

輸出 樣本 日誌

預覽(P) 確定(O) 取消(C)



# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

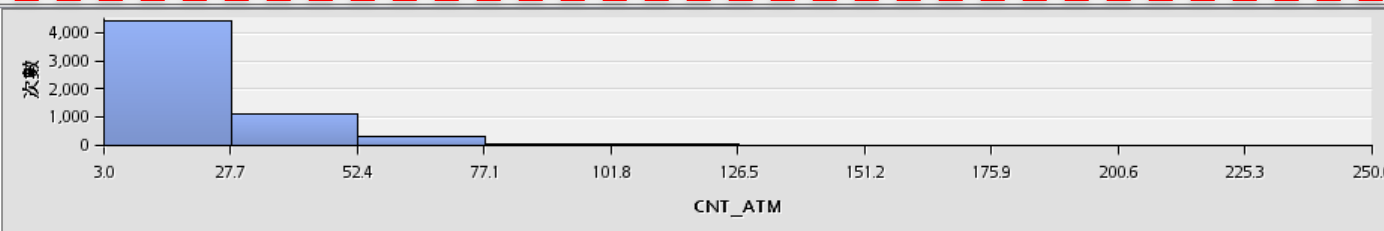
EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性 值

一般  
節點 ID Trans  
匯入的資料  
匯出的資料  
附註  
訓練  
變數  
公式  
交互作用  
SAS 程式碼  
預設方法  
間隔輸入 無  
間隔目標 無  
類別輸入 無  
類別目標 無  
將遺漏值處理為層級否  
樣本屬性  
方法 前 N 個  
大小 預設  
隨機種子 12345  
最佳等距分箱  
等距箱數的數目 4  
遺漏值 搜尋時使用  
群組方法  
截止值 0.1  
群組遺漏 否  
等距箱數的數目 變數  
將最小值增加到位移是  
位移值 1

公式

偏態 (Skewed input)



次數

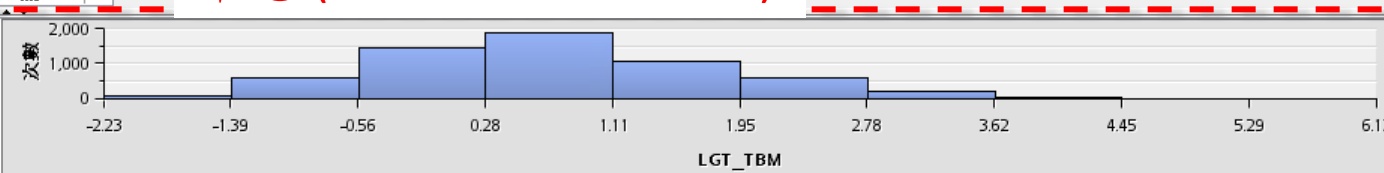
CNT\_ATM

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	角色	層級	方法	等距箱數的數目
CNT_ATM	輸入	間隔	DEFAULT	4.0
CNT_CSC	輸入	間隔	DEFAULT	4.0
CNT_POS	輸入	間隔	DEFAULT	4.0
CNT_TBM	輸入	間隔	DEFAULT	4.0
CNT_TOT	輸入	間隔	DEFAULT	4.0

輸入

常態 (Normal Distribution)



次數

LGT\_TBM

名稱	類型	長度	格式	層級	公式	標籤	角色	報表
LGT_ATM	數值	8		間隔	$\text{LOG}(\text{CNT\_ATM} / (\text{CNT\_TOT} - \text{CNT\_ATM}))$		輸入	否
LGT_CSC	數值	8		間隔	$\text{LOG}(\text{CNT\_CSC} / (\text{CNT\_TOT} - \text{CNT\_CSC}))$		輸入	否
LGT_POS	數值	8		間隔	$\text{LOG}(\text{CNT\_POS} / (\text{CNT\_TOT} - \text{CNT\_POS}))$		輸入	否
LGT_TBM	數值	8		間隔	$\text{LOG}(\text{CNT\_TBM} / (\text{CNT\_TOT} - \text{CNT\_TBM}))$		輸入	否

輸出 樣本 日誌

預覽(P) 確定(O) 取消(C)

執行完成

# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

The screenshot shows the Enterprise Miner interface. On the left, a tree view shows the project structure: EM\_Project1, 資料來源, PROFILE, 流程圖, CaseStudy, and 模型套件. Below the tree is a properties table for the selected 'CaseStudy' node.

屬性	值
一般	
節點 ID	Trans
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
公式	
交互作用	
SAS 程式碼	
預設方法	
間隔輸入	無
間隔目標	無
類別輸入	無
類別目標	無
將遺漏值處理為層級	否
樣本屬性	
方法	前 N 個
大小	預設
隨機種子	12345
最佳等距分箱	
等距箱數的數目	4
遺漏值	搜尋時使用
群組方法	
截止值	0.1
群組遺漏	否
等距箱數的數目	變數
將最小值增加到位移是	
位移值	1

The main workspace shows a workflow diagram with nodes: PROFILE, 篩選, 群集, and 區段特徵描繪. A '轉換變數' node is added below the '篩選' node. A context menu is open over the '轉換變數' node, with the '執行' (Execute) option highlighted in red.

- 編輯變數...
- 更新
- 執行**
- 建立模型套件...
- 結果...
- 將路徑匯出作為 SAS 程式
- 剪下
- 複製(C)
- 刪除
- 重新命名
- 全選
- 選取節點
- 連接節點
- 中斷連接節點

At the bottom, the status bar shows '執行完成' (Execution completed).

# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy

PROFILE → 篩選 → 群集 → 區段特徵描繪

轉換變數

執行狀態

執行完成  
流程圖: CaseStudy  
路徑: 轉換變數

確定 結果...

屬性	值
一般	
節點 ID	Trans
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
公式	
交互作用	
SAS 程式碼	
預設方法	
間隔輸入	無
間隔目標	無
類別輸入	無
類別目標	無
將遺漏值處理為層級	否
樣本屬性	
方法	前 N 個
大小	預設
隨機種子	12345
最佳等距分箱	
等距箱數的數目	4
遺漏值	搜尋時使用
群組方法	
截止值	0.1
群組遺漏	否
等距箱數的數目	變數
將最小值增加到位移是	
位移值	1

公式

執行完成

# 轉換變數 (Transform Variables)

## 增加轉換 (衍生變數)

結果 - 節點: 轉換變數 流程圖: CaseStudy

檔案(F) 編輯(E) 檢視(V) 視窗(W)

轉換統計

來源	方法	變數名稱	公式	層級數目	非遺漏值	遺漏	最小值	最大值	平均值	標準差	偏態	峰度	標籤
Input	Original	CNT_ATM			100000	0	3	628	19.49971	20.8561	2.357293	15.33353	
Input	Original	CNT_CSC			100000	0	1	607	6.68411	12.12856	6.236494	108.4633	
Input	Original	CNT_POS			100000	0	2	345	11.9233	20.73384	3.343805	15.43941	
Input	Original	CNT_TBM			100000	0	10	14934	68.13696	101.1542	53.05219	5925.782	
Input	Original	CNT_TOT			100000	0	17	15225	106.2441	113.3704	39.2061	3976.09	
Output	Formula	LGT_ATM	LOG(CNT_ATM / (CNT_TOT - CNT_ATM))		100000	0	-7.93272	1.956639	-1.705	1.146881	-0.26459	0.001649	
Output	Formula	LGT_CSC	LOG(CNT_CSC / (CNT_TOT - CNT_CSC))		100000	0	-7.9179	1.110041	-3.22027	1.096677	0.203771	-0.36907	
Output	Formula	LGT_POS	LOG(CNT_POS / (CNT_TOT - CNT_POS))		100000	0	-8.93742	2.061423	-2.68836	1.144301	0.193406	-0.33408	
Output	Formula	LGT_TBM	LOG(CNT_TBM / (CNT_TOT - CNT_TBM))		100000	0	-3.46934	6.124319	0.709437	1.083986	0.350694	-0.12867	

輸出

```
5 *-----*
6 * 訓練輸出
7 *-----*
8
9
10
11
12 變數摘要
13
14 角色    量值層級    次數計數
15
16 INPUT    INTERVAL    5
17
18
19 *-----*
20 * 評分輸出
21 *-----*
22
23
24
25 公式轉換
26 (最多列印 500 個觀測)
27
```

LGT\_ATM = LOG(CNT\_ATM / (CNT\_TOT - CNT\_ATM))  
LGT\_CSC = LOG(CNT\_CSC / (CNT\_TOT - CNT\_CSC))  
LGT\_POS = LOG(CNT\_POS / (CNT\_TOT - CNT\_POS))  
LGT\_TBM = LOG(CNT\_TBM / (CNT\_TOT - CNT\_TBM))



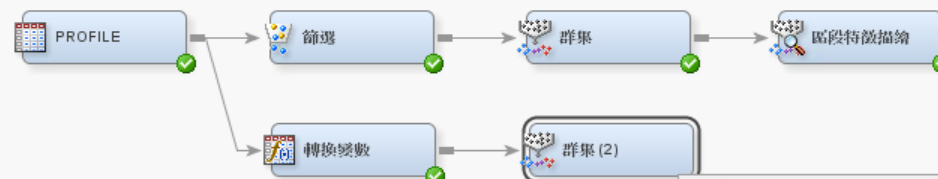
EM\_Project1  
 資料來源  
 PROFILE  
 流程圖  
 CaseStudy  
 模型套件

屬性	值
<b>一般</b>	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	自動
最大的群集數目	10
選擇準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
<b>類別變數編碼</b>	
序數編碼	排名
名目編碼	GLM
<b>初始群集種子</b>	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	
使用預設值	是
設定	
遺漏值	



樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy



編輯變數...

更新

執行

建立模型套件...

結果...

將路徑匯出作為 SAS 程式

剪下

複製(C)

刪除

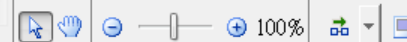
重新命名

全選

選取節點

連接節點

中斷連接節點



流程圖 | 日誌



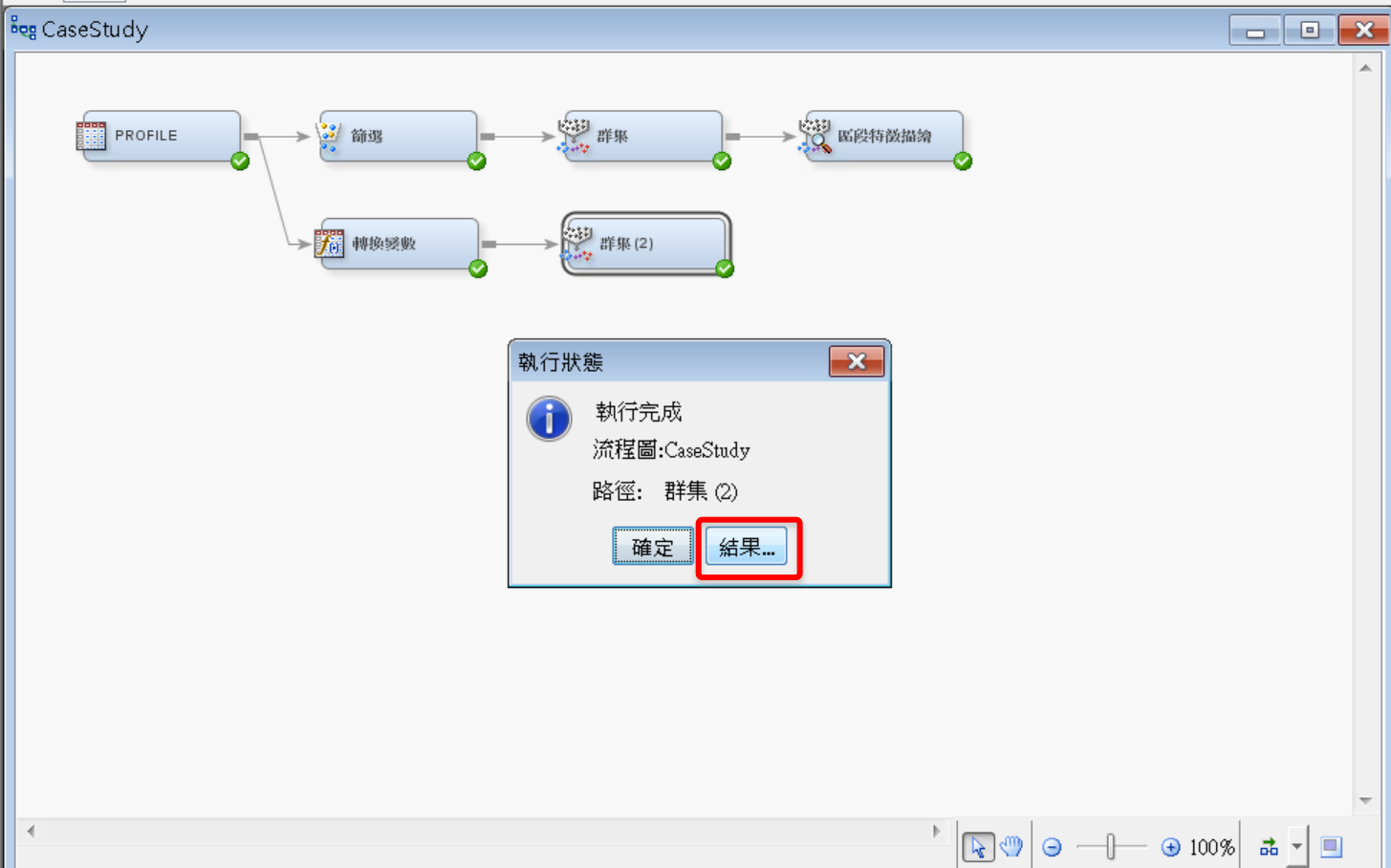
EM\_Project1

- 資料來源
- PROFILE
- 流程圖
- CaseStudy
- 模型套件

屬性	值
<b>一般</b>	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
<b>群集的數目</b>	
指定方法	自動
最大的群集數目	10
<b>選取準則</b>	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
<b>類別變數編碼</b>	
序數編碼	排名
名目編碼	GLM
<b>初始群集種子</b>	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	
使用預設值	是
設定	
遺漏值	

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列



執行狀態

執行完成

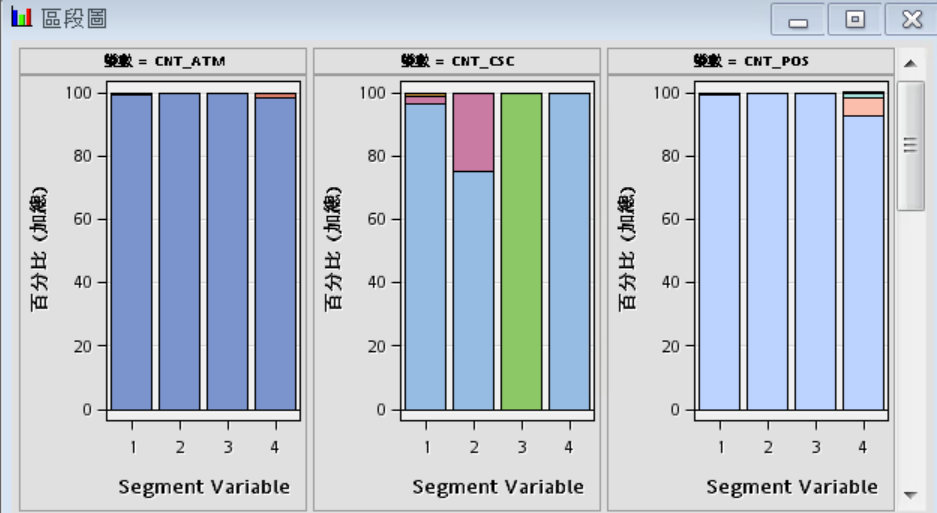
流程圖: CaseStudy

路徑: 群集 (2)

確定 結果...

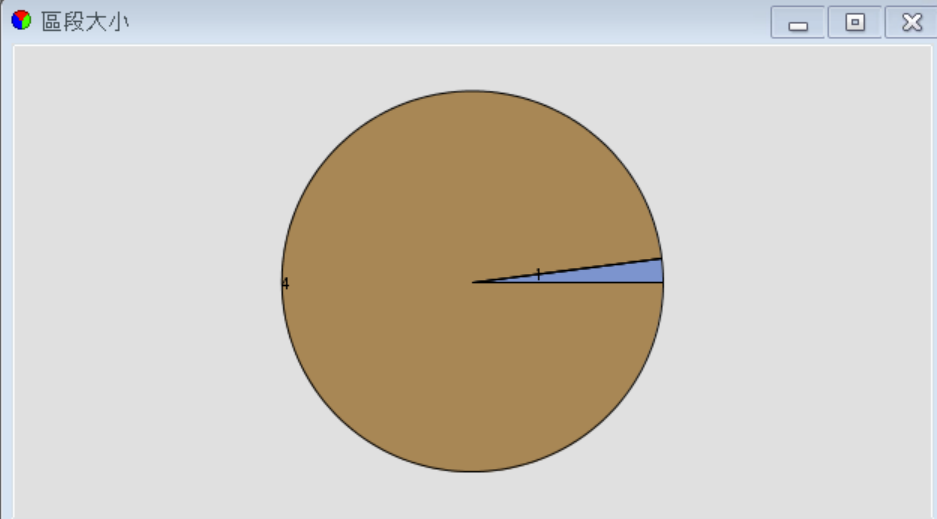
流程圖 | 日誌

100%



### 平均值統計

Clusterin g Criterion	Maximum Relative Change in Cluster Seeds	Improve ment in Clusterin g Criterion	區段 ID	Frequenc y of Cluster	Root-Mea n-Square Standard Deviation	Maximum Distance from Cluster Seed	Nearest Cluster
0.91879	0.028016	.	1	1889	1.608709	49.01384	4
0.91879	0.028016	.	2	8	7.213317	38.14711	1
0.91879	0.028016	.	3	1	0	0	2
0.91879	0.028016	.	4	98102	0.895164	31.40487	1



### 輸出

```

5 *-----*
6 * 訓練輸出
7 *-----*
8
9
10
11
12 變數摘要
13
14 角色    量值層級    次數計數
15
16 ID      NOMINAL      1
17 INPUT   INTERVAL     9
18
19
20
    
```



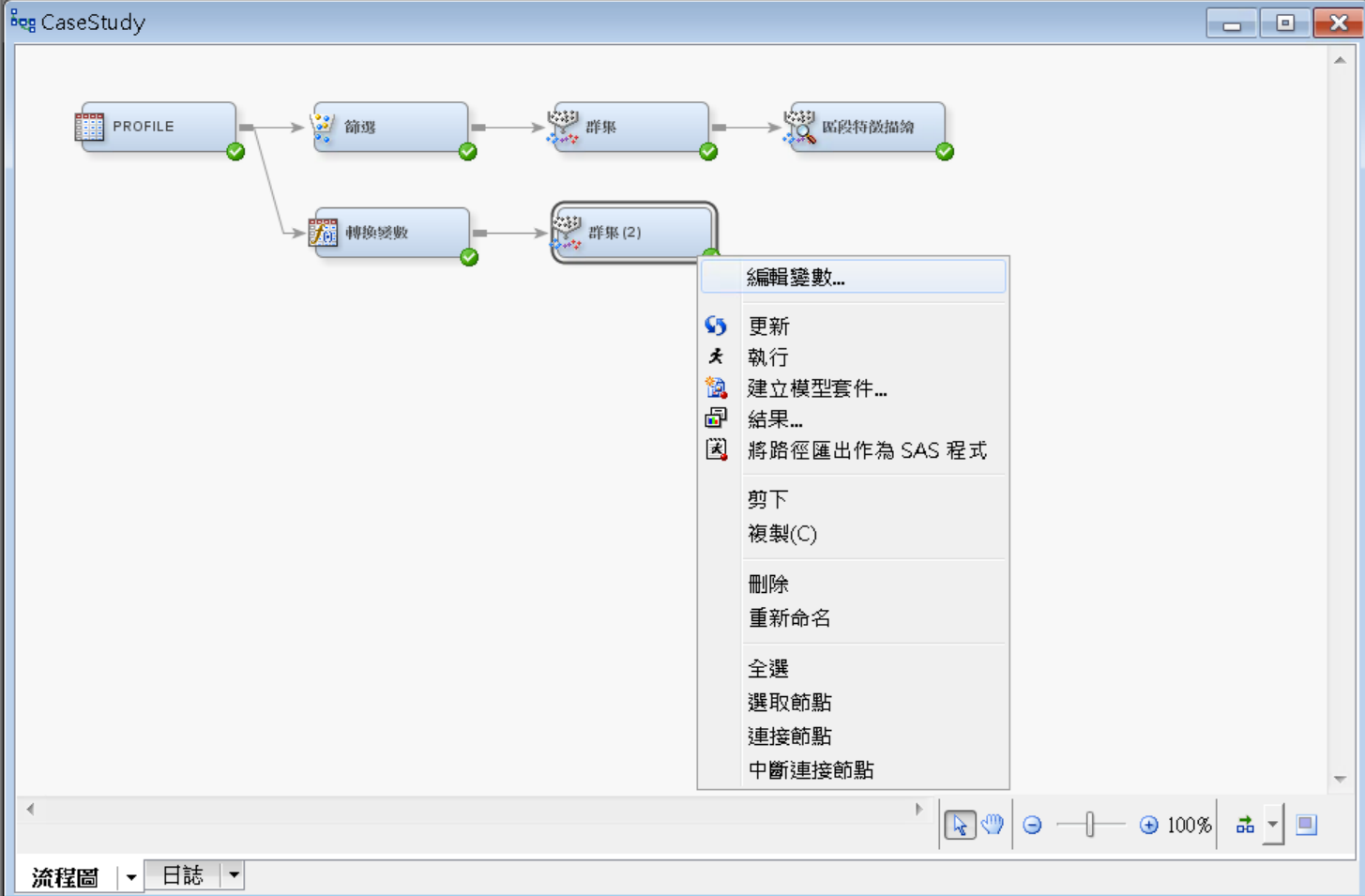
EM\_Project1

- 資料來源
- PROFILE
- 流程圖
- CaseStudy
- 模型套件

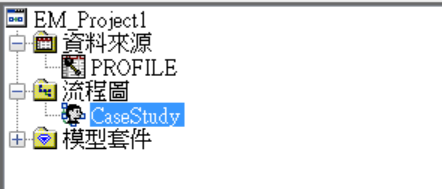
屬性	值
<b>一般</b>	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
<b>群集的數目</b>	
指定方法	自動
最大的群集數目	10
<b>選取準則</b>	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
<b>類別變數編碼</b>	
序數編碼	排名
名目編碼	GLM
<b>初始群集種子</b>	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	
使用預設值	是
設定	
遺漏值	

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列







屬性	值
<b>一般</b>	
節點 ID	Clus2
匯入的資料	...
匯出的資料	...
附註	...
<b>訓練</b>	
變數	...
群集變數角色	區段
內部標準化	標準化
群集的數目	...
指定方法	自動
最大的群集數目	10
選擇準則	...
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	...
序數編碼	排名
名目編碼	GLM
初始群集種子	...
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	
使用預設值	是
設定	...
遺漏值	...

一般

執行完成



樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy

變數 - Clus2

 (無)  非 等於

套用

重新設定

 標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	使用	報表	角色	層級
CNT_ATM	預設	否	輸入	間隔
CNT_CSC	預設	否	輸入	間隔
CNT_POS	預設	否	輸入	間隔
CNT_TBM	預設	否	輸入	間隔
CNT_TOT	預設	否	輸入	間隔
ID	是	否	ID	名目
LGT_ATM	預設	否	輸入	間隔
LGT_CSC	預設	否	輸入	間隔
LGT_POS	預設	否	輸入	間隔
LGT_TBM	預設	否	輸入	間隔

勘查(O)...

更新路徑(U)

確定(O)

取消(C)



EM\_Project1

- 資料來源
  - PROFILE
  - 流程圖
  - CaseStudy
- 模型套件

屬性	值
<b>一般</b>	
節點 ID	Clus2
匯入的資料	...
匯出的資料	...
附註	...
<b>訓練</b>	
變數	...
群集變數角色	區段
內部標準化	標準化
群集的數目	...
指定方法	自動
最大的群集數目	10
選擇準則	...
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	...
序數編碼	排名
名目編碼	GLM
初始群集種子	...
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	
使用預設值	是
設定	...
遺漏值	...

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

變數 - Clus2

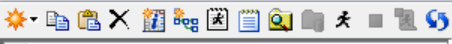
(無)  非 等於

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	使用	報表	角色	層級
CNT_ATM	否	否	輸入	間隔
CNT_CSC	否	否	輸入	間隔
CNT_POS	否	否	輸入	間隔
CNT_TBM	否	否	輸入	間隔
CNT_TOT	否	否	輸入	間隔
ID	是	否	ID	名目
LGT_ATM	預設	否	輸入	間隔
LGT_CSC	預設	否	輸入	間隔
LGT_POS	預設	否	輸入	間隔
LGT_TBM	預設	否	輸入	間隔

套用 重新設定

勘查(X)... 更新路徑(U) 確定(O) 取消(C)



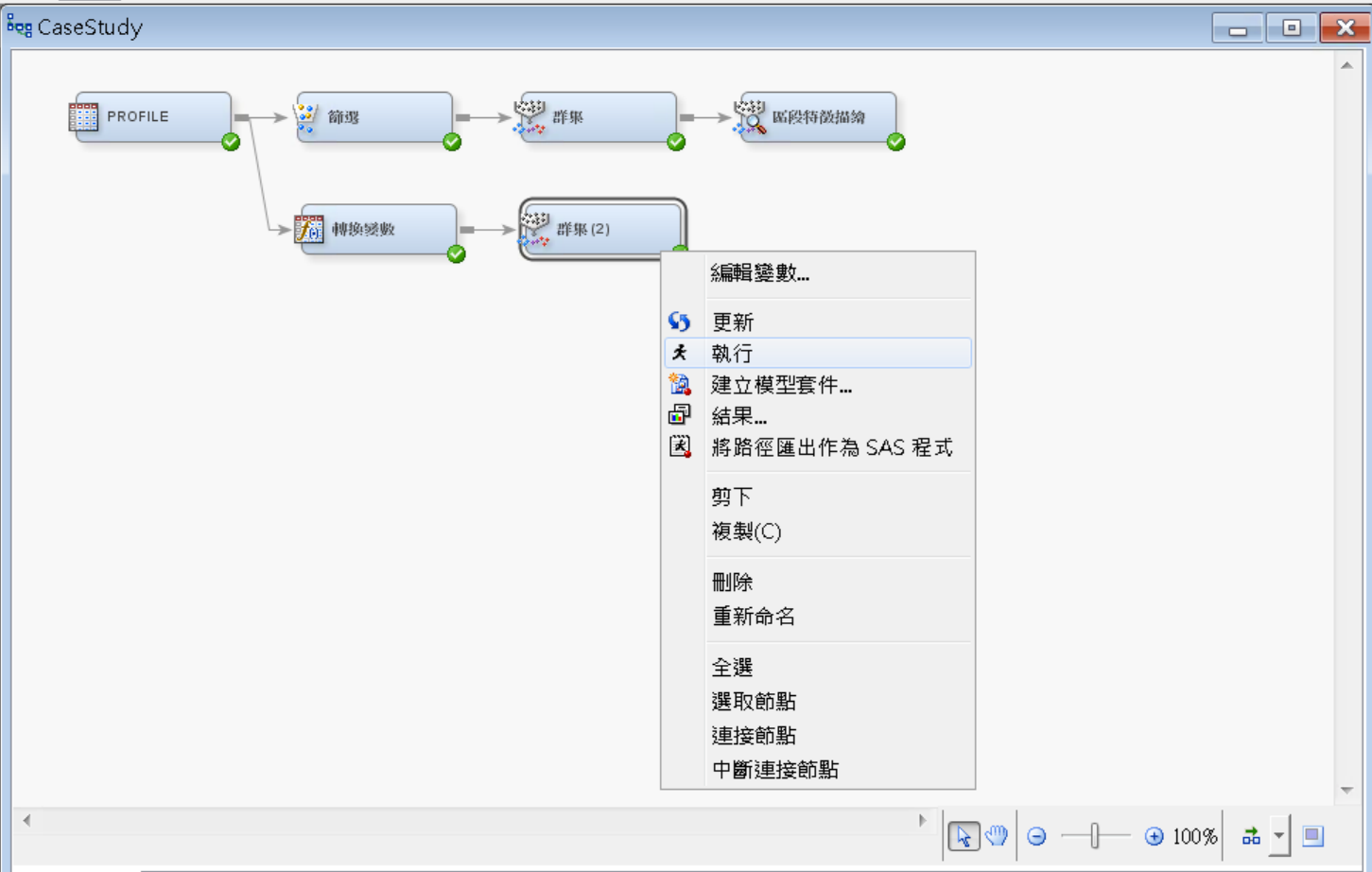
EM\_Project1

- 資料來源
- PROFILE
- 流程圖
- CaseStudy
- 模型套件

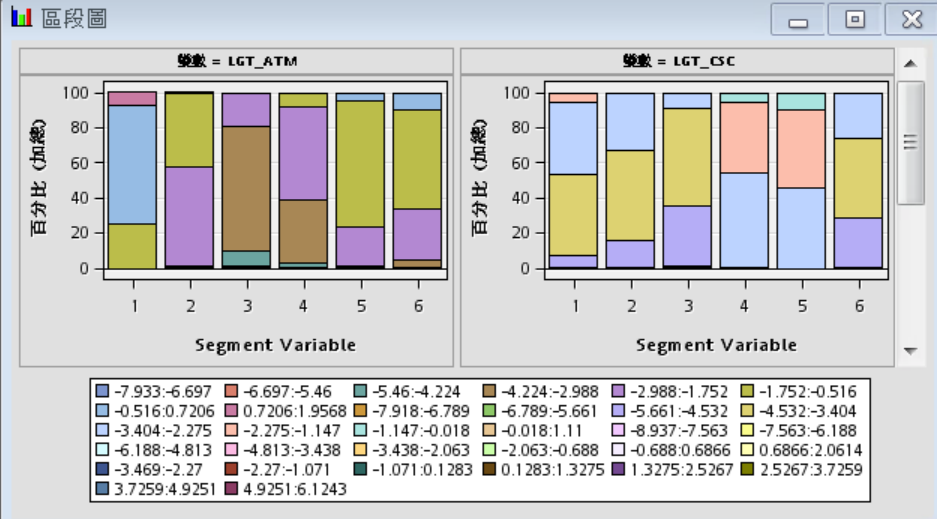
屬性	值
<b>一般</b>	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
<b>訓練</b>	
變數	
群集變數角色	區段
內部標準化	標準化
<b>群集的數目</b>	
指定方法	自動
最大的群集數目	10
<b>選取準則</b>	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
<b>類別變數編碼</b>	
序數編碼	排名
名目編碼	GLM
<b>初始群集種子</b>	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
<b>訓練選項</b>	
使用預設值	是
設定	
遺漏值	

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

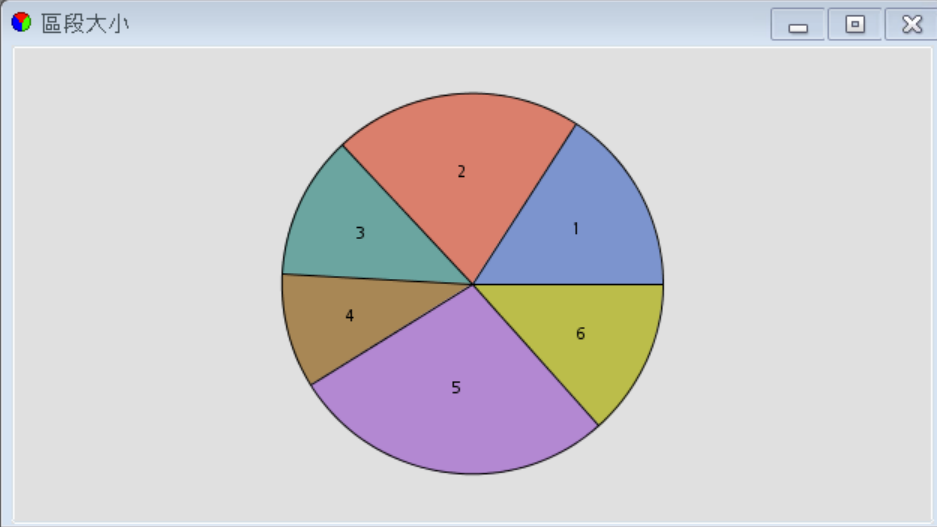


- 編輯變數...
- 更新
- 執行**
- 建立模型套件...
- 結果...
- 將路徑匯出作為 SAS 程式
- 剪下
- 複製(C)
- 刪除
- 重新命名
- 全選
- 選取節點
- 連接節點
- 中斷連接節點



### 平均值統計

Clusterin g Criterion	Maximum Relative Change in Cluster Seeds	Improve ment in Clusterin g Criterion	區段 ID	Frequenc y of Cluster	Root-Mea n-Square Standard Deviation	Maximum Distance from Cluster Seed	Nearest Cluster
0.560591	0.019301		1	15919	0.563905	3.458123	5
0.560591	0.019301		2	20907	0.516724	2.951991	4
0.560591	0.019301		3	12277	0.519155	6.096798	2
0.560591	0.019301		4	9675	0.563037	4.830533	2
0.560591	0.019301		5	27899	0.570554	4.069428	1
0.560591	0.019301		6	13323	0.630681	3.947949	5



### 輸出

```

5 *-----*
6 * 訓練輸出
7 *-----*
8
9
10
11
12 變數摘要
13
14 角色    量值層級    次數計數
15
16 ID      NOMINAL      1
17 INPUT   INTERVAL     9
18
19
20
    
```

# 使用者指定最大的群集數目

The screenshot displays the Enterprise Miner software interface. On the left, a tree view shows the project structure: EM\_Project1, 資料來源, PROFILE, 流程圖, CaseStudy, and 模型套件. Below this is a configuration table for the 'CaseStudy' workflow.

屬性	值
一般	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	自動
最大的群集數目	使用者指定
選取準則	自動
群集方法	自動
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	
序數編碼	排名
名目編碼	GLM
初始群集種子	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
訓練選項	
使用預設值	是
設定	
遺漏值	
指定方法	

The main workspace shows a workflow diagram for 'CaseStudy'. The workflow starts with a 'PROFILE' node, which branches into two paths: one leading to '篩選' (Filter) and another to '轉換變數' (Transform Variables). Both paths converge into a '群集' (Cluster) node, which then leads to '區段特徵描繪' (Segment Feature Plot). A second '群集 (2)' (Cluster 2) node is highlighted with a red box, indicating it is the current focus of the configuration.

At the bottom of the interface, there are tabs for '流程圖' (Flowchart) and '日誌' (Log), and a status bar indicating '執行完成' (Execution Complete).

# 使用者指定最大的群集數目：5

The screenshot displays the Enterprise Miner interface. On the left, a tree view shows the project structure: EM\_Project1, 資料來源, PROFILE, 流程圖, CaseStudy, and 模型套件. Below this, a table lists various attributes and their values. A red dashed box highlights the '訓練' (Training) section of this table, specifically the '最大的群集數目' (Maximum number of clusters) row, which is set to 5. The main workspace shows a workflow diagram for 'CaseStudy' with nodes: PROFILE, 篩選 (Filter), 轉換變數 (Transform Variables), 群集 (Cluster), 群集 (2), and 區段特徵描繪 (Segment Feature Description). The bottom status bar indicates '執行完成' (Execution completed).

屬性	值
一般	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	使用者指定
最大的群集數目	5
選取準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	
序數編碼	排名
名目編碼	GLM
初始群集種子	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
訓練選項	
使用預設值	是
設定	
遺漏值	

# 群集 / 執行

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

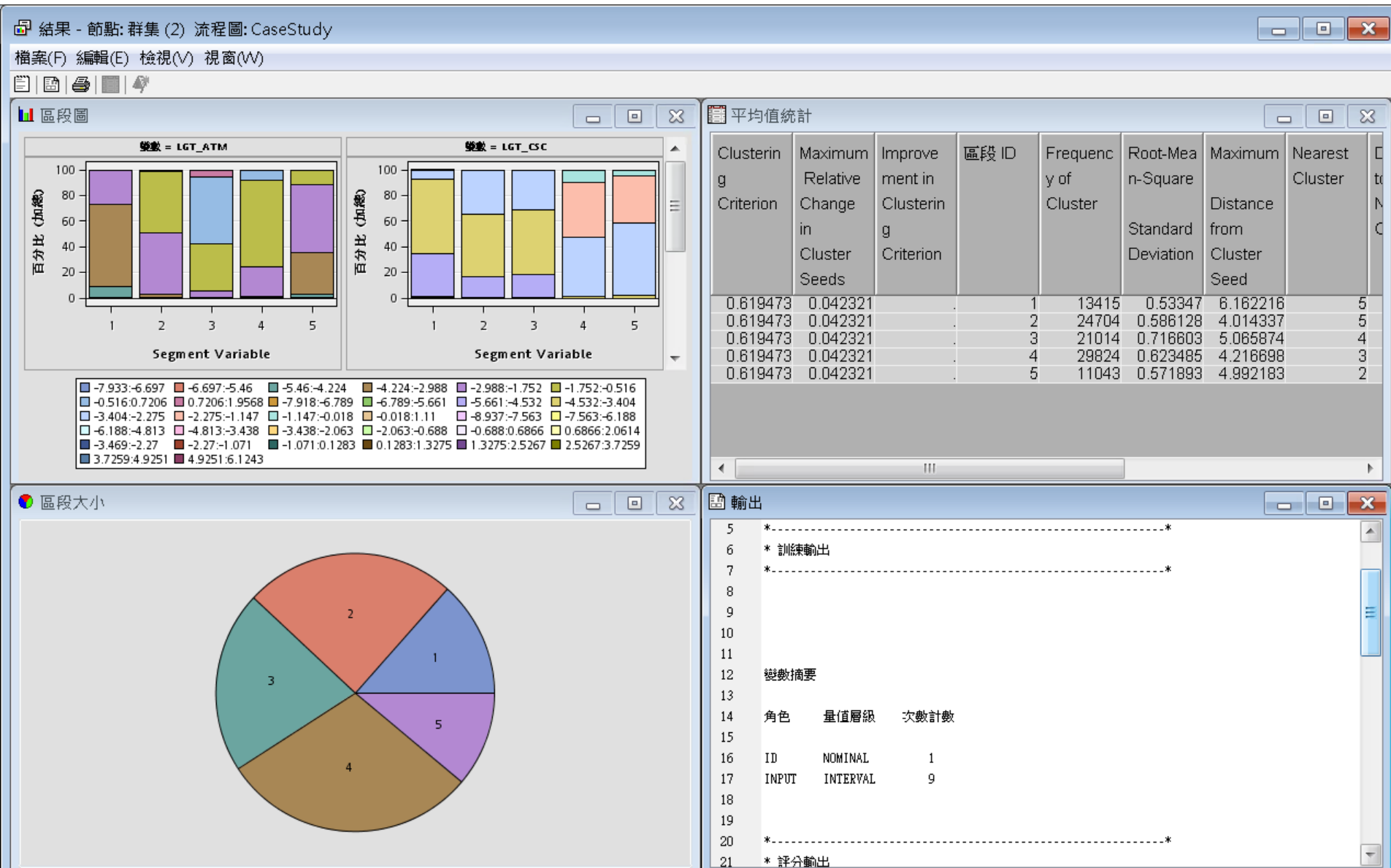
CaseStudy

```
graph LR; PROFILE --> 篩選; PROFILE --> 轉換變數; 篩選 --> 群集; 轉換變數 --> 群集(2); 群集 --> 區段特徵描繪;
```

屬性	值
一般	
節點 ID	Clus2
匯入的資料	
匯出的資料	
附註	
訓練	
變數	
群集變數角色	區段
內部標準化	標準化
群集的數目	
指定方法	使用者指定
最大的群集數目	5
選擇準則	
群集方法	Ward
初步最大值	50
最小值	2
最終最大值	20
CCC 截止值	3
類別變數編碼	
序數編碼	排名
名目編碼	GLM
初始群集種子	
種子初始化方法	預設
最小半徑	0.0
在訓練期間漂移	否
訓練選項	
使用預設值	是
設定	
遺漏值	
一般	

執行完成

# 群集執行結果





# 區段特徵描繪 (Segment Profile)

## 編輯變數

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1  
資料來源  
PROFILE  
流程圖  
CaseStudy  
模型套件

屬性	值
一般	
節點 ID	Prof2
匯入的資料	
匯出的資料	
附註	
訓練變數	
一般	
中間點數目	8
所有的特徵描繪	否
截止百分比	95
輸入變數	
輸入的數目	10
最小價值	0.01
最大深度	1
列印價值統計	是
目標變數	
分析角色	無
報表變數	
使用報表變數	是
報表變數的數目	10
狀態	
一般	
一般屬性	

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列

CaseStudy

```
graph LR; PROFILE[PROFILE] --> 篩選[篩選]; 篩選 --> 群集[群集]; 群集 --> 區段特徵描繪[區段特徵描繪]; 篩選 --> 轉換變數[轉換變數]; 轉換變數 --> 群集2[群集 (2)]; 群集2 --> 區段特徵描繪2[區段特徵描繪 (2)];
```

編輯變數...

- 更新
- 執行
- 建立模型套件...
- 結果...
- 將路徑匯出作為 SAS 程式
- 剪下
- 複製(C)
- 刪除
- 重新命名
- 全選
- 選取節點
- 連接節點
- 中斷連接節點

100%

執行完成

# 區段特徵描繪 (Segment Profile)

## 編輯變數

Enterprise Miner - EM\_Project1

檔案(F) 編輯(E) 檢視(V) 動作(A) 選項(O) 視窗(W) 說明(H)

EM\_Project1

- 資料來源
  - PROFILE
  - 流程圖
  - CaseStudy
- 模型套件

變數 - Prof2

欄:  標籤(A)  採礦(M)  基本(I)  統計(T)

名稱	使用	報表	角色	層級
CNT_ATM	否	否	輸入	間隔
CNT_CSC	否	否	輸入	間隔
CNT_POS	否	否	輸入	間隔
CNT_TBM	否	否	輸入	間隔
CNT_TOT	否	否	輸入	間隔
Distance	預設	否	拒絕	間隔
ID	預設	否	ID	名目
LGT_ATM	是	否	輸入	間隔
LGT_CSC	是	否	輸入	間隔
LGT_POS	是	否	輸入	間隔
LGT_TBM	是	否	輸入	間隔
SEGMENT	預設	否	區段	名目
SEGMENT	預設	否	拒絕	名目

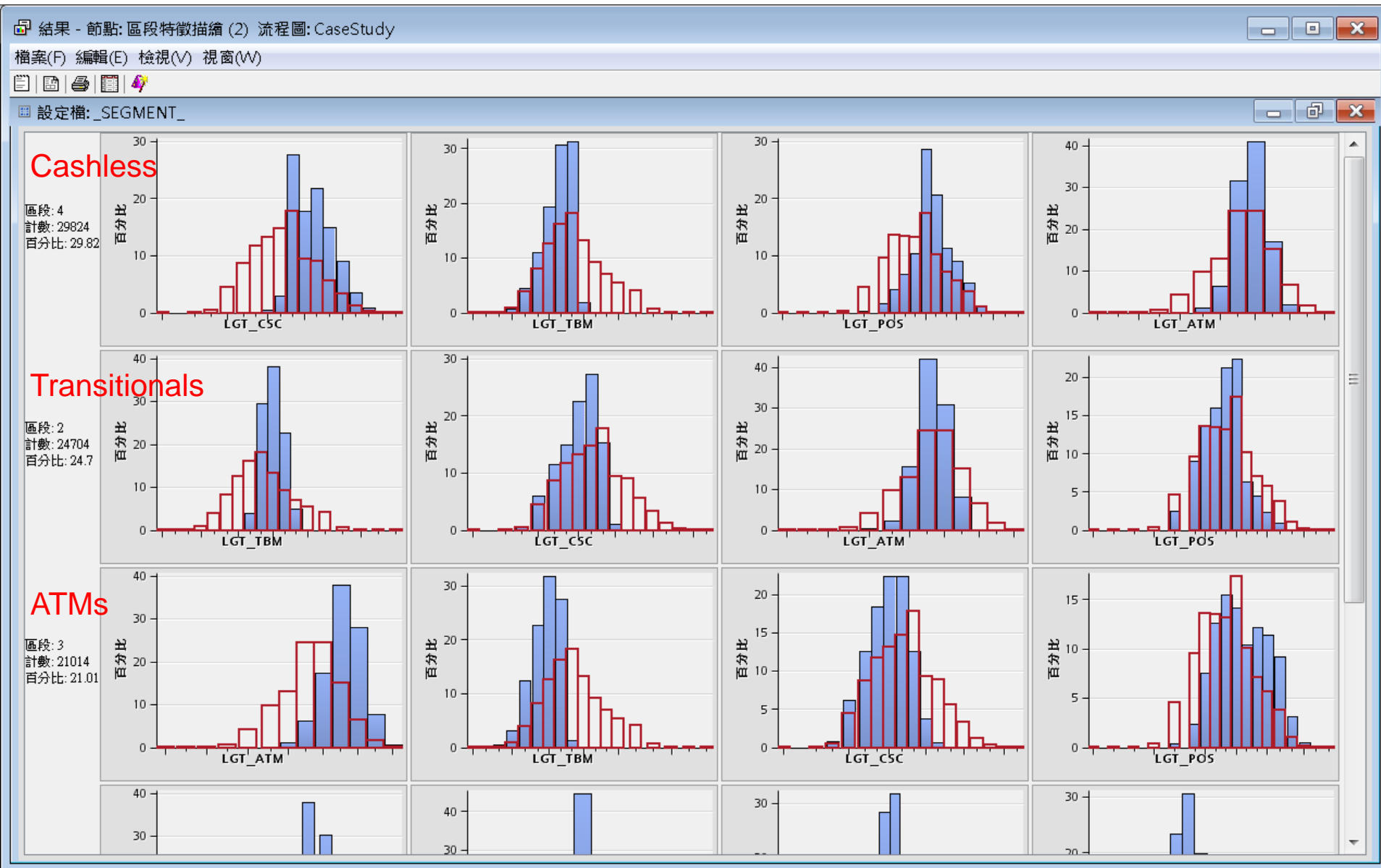
執行完成

勘查(X)... 更新路徑(U) 確定(O) 取消(C)

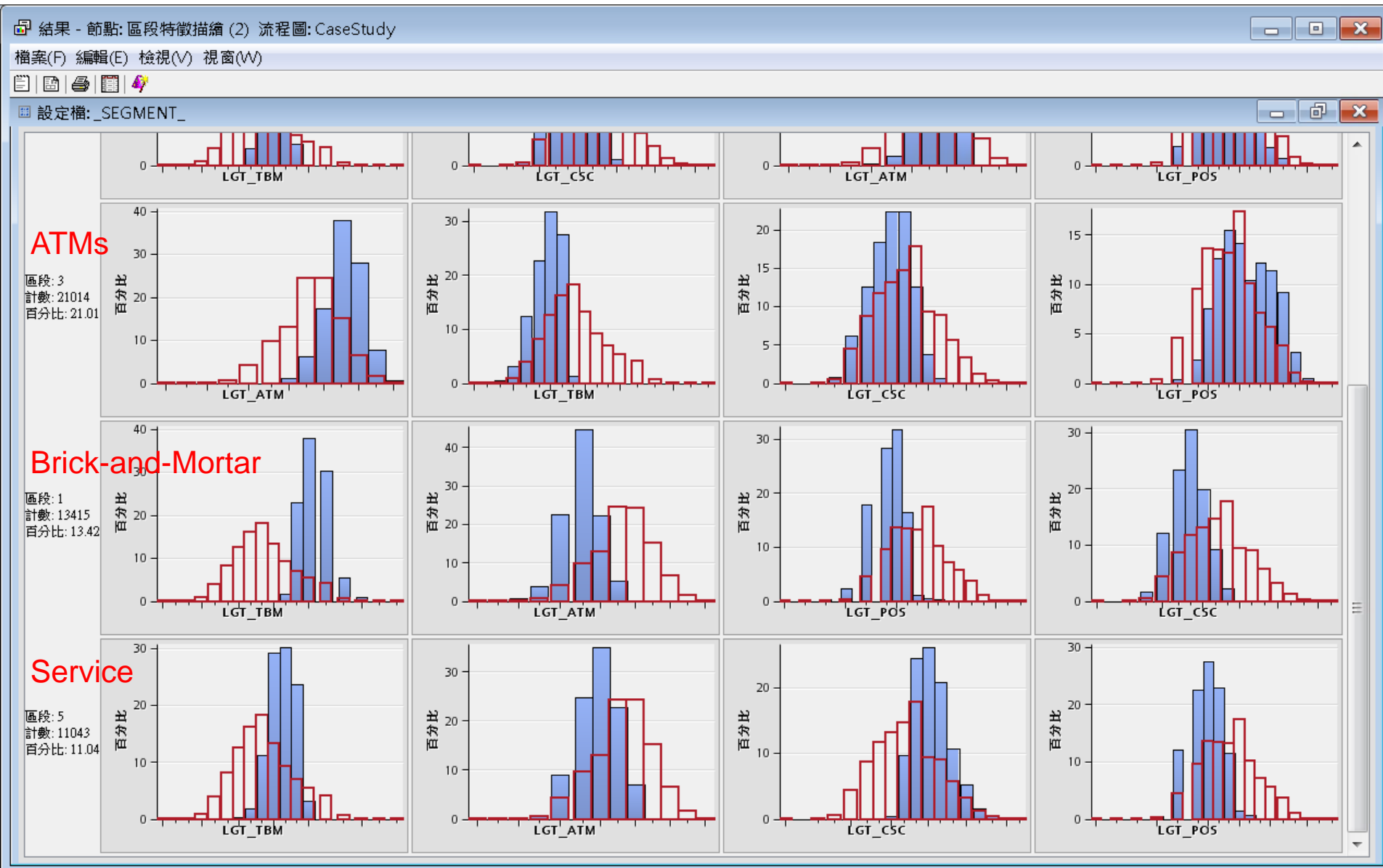
# 區段特徵描繪 (Segment Profile)



# 區段特徵描繪 (Segment Profile)



# 區段特徵描繪 (Segment Profile)



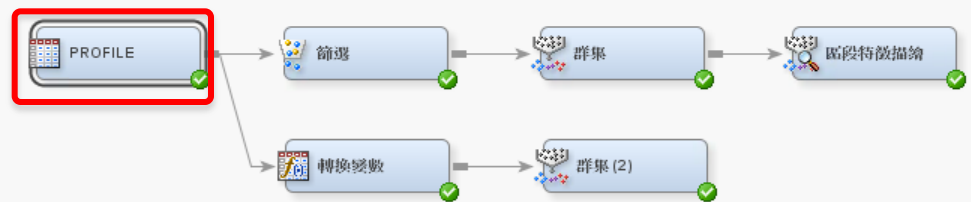


- EM\_Project1
  - 資料來源
    - PROFILE
  - 流程圖
  - CaseStudy
  - 模型套件

屬性	值
一般	
節點 ID	Ids
匯入的資料	...
匯出的資料	
附註	
訓練	
輸出類型	檢視
角色	原始
重新執行	否
彙總	否
刪除對應變數	是
變數	
決策	
重新整理中繼資料	
顧問	基本
進階選項	
資料	
資料選取	資料來源
樣本	預設
樣本選項	
資料來源	
資料來源	PROFILE
資料來源屬性	
新增表格	
表格名稱	
變數驗證	嚴格
新增變數角色	拒絕
中繼資料	

CaseStudy

樣本 勘查 修改 模型 評估 公用程式 應用程式 時間序列



匯入的資料 - PROFILE

連接埠	來源	表格	角色	資料存在
DATA		EM_LIB.PROFILE		是

瀏覽(B)... 勘查(X)... 屬性(P)... 確定(O)

# 分群分析 (Cluster Analysis)

## EM\_Lib.Profile

The screenshot displays the Enterprise Miner interface. On the left, a tree view shows the project structure: EM\_Project1, 資料來源 (Data Sources), PROFILE, 流程圖 (Flowchart), CaseStudy, and 模型套件 (Model Kits). Below this is a properties pane for the 'PROFILE' node, showing various attributes and their values, such as '節點 ID' (Node ID) as 'Ids' and '資料來源' (Data Source) as 'PROFILE'. The main workspace shows a flowchart with a 'PROFILE' node connected to a '匯入的資料 - PROFILE' (Import Data - PROFILE) node. Below the flowchart is a table for the imported data, titled 'EM\_LIB.PROFILE'. The table has 7 columns: ID, CNT\_TBM, CNT\_ATM, CNT\_POS, CNT\_CSC, and CNT\_TOT. It contains 23 rows of data. At the bottom of the interface, the status bar indicates '執行完成' (Execution Complete).

	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34.0	3.0	3.0	9.0	49.0
2	000052094	44.0	17.0	5.0	18.0	84.0
3	000057340	122.0	26.0	32.0	36.0	216.0
4	000076885	42.0	3.0	6.0	1.0	52.0
5	000089150	20.0	15.0	2.0	2.0	39.0
6	000094512	83.0	20.0	49.0	3.0	155.0
7	000096396	33.0	9.0	49.0	17.0	108.0
8	000098062	22.0	5.0	9.0	1.0	37.0
9	000122779	21.0	27.0	2.0	1.0	51.0
10	000125838	127.0	3.0	2.0	3.0	135.0
11	000137317	52.0	20.0	2.0	11.0	85.0
12	000147896	101.0	40.0	8.0	1.0	150.0
13	000150206	156.0	52.0	80.0	4.0	292.0
14	000160501	31.0	32.0	42.0	12.0	117.0
15	000176820	107.0	3.0	2.0	1.0	113.0
16	000187784	54.0	14.0	7.0	24.0	99.0
17	000189200	92.0	3.0	2.0	2.0	99.0
18	000198327	44.0	27.0	17.0	3.0	91.0
19	000200237	146.0	3.0	2.0	1.0	152.0
20	000207980	39.0	5.0	2.0	4.0	50.0
21	000227355	180.0	3.0	2.0	10.0	195.0
22	000228317	46.0	14.0	2.0	1.0	63.0
23	000228906	25.0	3.0	2.0	1.0	31.0

# **SAS Enterprise Guide (SAS EG)**

## **Excel and SAS Data Conversion**



# SAS Enterprise Guide (SAS EG)

**Open SAS .sas7bdat File**

**Export to Excel .xlsx File**

**Import Excel .xlsx File**

**Export to SAS .sas7bdat File**

# SAS Enterprise Guide 5.1 (SAS EG)



## 淡江軟體雲使用注意事項

- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室，或來電 26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。

# New Project

SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P) 工具(T) 說明(H) | 處理流程

專案樹狀結構 | 處理流程

執行(R) | 停止(S) | 匯出(X) | 排程(D) | 縮放(Z) | 專案記錄檔(L) | 屬性(I)

歡迎使用 SAS Enterprise Guide

選取這些選項的其中一個來開始:

開啟專案

- 更多專案...

新增

- 新增專案**
- 新增 SAS 程式
- 新增資料

輔助

- 教學課程: SAS Enterprise Guide 入門

不要再顯示此視窗

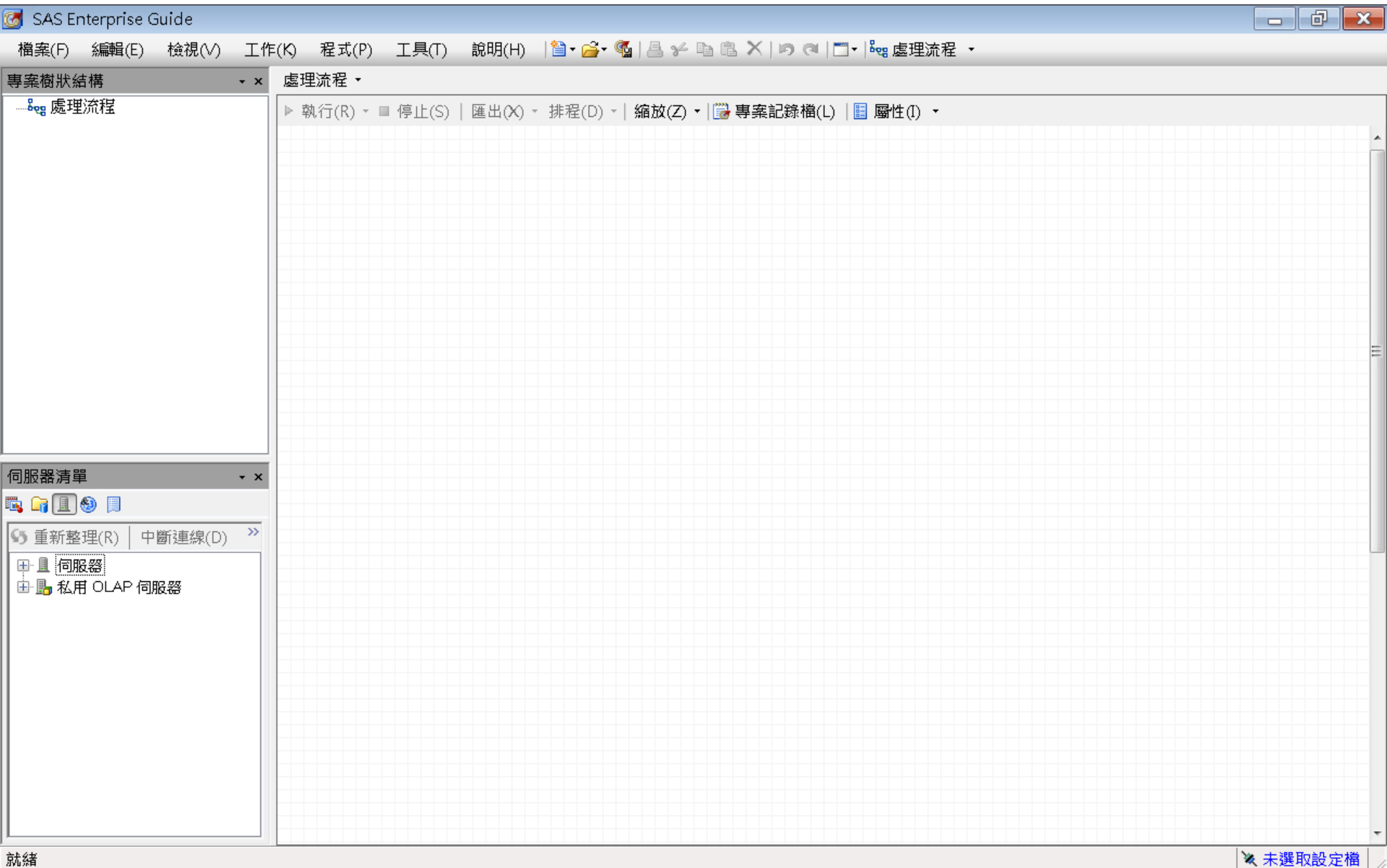
伺服器清單

重新整理(R) | 中斷連線(D)

- 伺服器
- 私有 OLAP 伺服器

就緒 | 未選取設定檔

# Open SAS Data File: Profile.sas7bdat



- 新增(N) ▾
- 開啟(O) ▾
- 關閉專案(C)
- 儲存「專案」(S) Ctrl+S
- 將「專案」另存為(A)...
- 儲存(V) Ctrl+Shift+S
- 另存新檔(E)...
- 匯入資料(I)
- 執行「專案」(U)
- 排程「專案」(H)
- 匯出(R) ▾
- 傳送至(D) ▾
- 發行(L)
- 版面設定(G)...
- 預覽列印(W)
- 列印(P)...
- Ctrl+P
- 專案屬性(J)
- 最近使用的專案(T) ▾
- 最近使用的程式(M) ▾
- 結束(X)

- 處理流程 ▾
- 專案(P) Ctrl+O
- 資料(D)...
- 資料勘查(X)...
- 程式(G)...
- 報表(T)...
- 預存程式(S)...
- OLAP Cube(U)...
- Information Map(I)...
- Exchange(E)...
- OLEDB(O)...
- ODBC(B)...
- 其他(R)...

SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P)

專案樹狀結構

處理流程

處理流程

執行(R)

---

伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
- 私有 OLAP 伺服器

就緒

開啟「資料」

瀏覽 搜尋 SAS 資料夾

查詢(I): EM\_Data

名稱	類型	大小	修改日期
credit.sas7bdat	SAS Data Set	361 KB	2013/4/21 下午 0
inq2006.sas7bdat	SAS Data Set	13,729 KB	2013/5/29 下午 0
profile.sas7bdat	SAS Data Set	5,521 KB	2006/9/23 下午 3
webstation.sas7bdat	SAS Data Set	50,789 KB	2006/9/25 上午 0

檔案名稱(N): profile.sas7bdat

檔案類型(T): 所有已知的資料檔案 (\*.sas7bdat;\*.sas7bview;\*.sd2;\*.mdb;\*.accdb;\*.xls;\*.xlsx;\*.xlsm;\*.xlsb;\*.t

開啟(O) 取消

專案樹狀結構

處理流程  
 profile

伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
      - EC100001
      - MAPS
      - MAPSGFK
      - MAPSSAS
      - SASHELP
      - SASUSER
      - WORK
    - 檔案
  - 私用 OLAP 伺服器

profile

篩選和排序(L) 查詢產生器(Q) | 資料(D) 描述(B) 圖形(G) 分析(Z) | 匯出(X) 傳送至(N)

	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34	3	3	9	49
2	000052094	44	17	5	18	84
3	000057340	122	26	32	36	216
4	000076885	42	3	6	1	52
5	000089150	20	15	2	2	39
6	000094512	83	20	49	3	155
7	000096396	33	9	49	17	108
8	000098062	22	5	9	1	37
9	000122779	21	27	2	1	51
10	000125838	127	3	2	3	135
11	000137317	52	20	2	11	85
12	000147896	101	40	8	1	150
13	000150206	156	52	80	4	292
14	000160501	31	32	42	12	117
15	000176820	107	3	2	1	113
16	000187784	54	14	7	24	99
17	000189200	92	3	2	2	99
18	000198327	44	27	17	3	91
19	000200237	146	3	2	1	152
20	000207980	39	5	2	4	50
21	000227355	180	3	2	10	195
22	000228317	46	14	2	1	63
23	000228906	25	3	2	1	31
24	000229984	15	3	2	1	21
25	000241022	13	3	2	1	19
26	000253260	65	3	2	1	71
27	000257910	35	3	2	1	41
28	000261717	11	5	2	1	19
29	000266600	15	3	2	1	21
30	000269176	107	78	22	7	214
31	000273510	110	15	6	21	152

專案樹狀結構 profile

- 處理流程
- profile
- 開啟「profile」(O)
- 匯出(X)
- 篩選和排序(F)...
- 查詢產生器(Q)...
- 發行(L)...
- 傳送至(N)
- 複製(C)
- 貼上(P)
- 刪除(D)
- 重新命名(M)
- 關閉資料(S)
- 屬性(I)

ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
000041360	34	3	3	9	49
000052094	44	17	5	18	84
000057340	122	26	32	36	216
000076885	42	3	6	1	52
000089150	20	15	2	2	39
000094512	83	20	49	3	155
000096396	33	9	49	17	108
000098062	22	5	9	1	37
000122779	21	27	2	1	51
000125838	127	3	2	3	135
000137317	52	20	2	11	85
000147896	101	40	8	1	150
000150206	156	52	80	4	292
000160501	31	32	42	12	117
000176820	107	3	2	1	113
000187784	54	14	7	24	99
000189200	92	3	2	2	99
000198327	44	27	17	3	91
000200237	146	3	2	1	152
000207980	39	5	2	4	50
000227355	180	3	2	10	195
000228317	46	14	2	1	63
000228906	25	3	2	1	31
000229984	15	3	2	1	21
000241022	13	3	2	1	19
000253260	65	3	2	1	71
000257910	35	3	2	1	41
000261717	11	5	2	1	19
000266600	15	3	2	1	21
000269176	107	78	22	7	214
000273510	110	15	6	21	152

伺服器清單

- 重新整理(R) | 中斷連線(D)
- 伺服器
  - 本機
    - 資料館
      - EC100001
      - MAPS
      - MAPSGFK
      - MAPSSAS
      - SASHELP
      - SASUSER
      - WORK
    - 檔案
  - 私用 OLAP 伺服器



專案樹狀結構

- 處理流程
  - profile

伺服器清單

重新整理(R) | 中斷連

- 伺服器
  - 本機
    - 資料館
      - EC10000
      - MAPS
      - MAPSGF
      - MAPSSA
      - SASHELP
      - SASUSER
      - WORK
    - 檔案
  - 私用 OLAP 伺服器

「profile」的屬性

一般

欄  
進階  
摘要

profile

檔案屬性

檔案名稱: C:\DATA\EM\_Data\profile.sas7bdat 變更(C)...

建立日期: 2006年9月23日 下午 06:07

上次修改日期: 2006年9月23日 下午 06:07

資料類型: SAS 資料集

標籤:

其他 (F1)...

確定 取消

30	000269176	107	78	22	7	214
31	000273510	110	15	6	21	152

專案樹狀結構

處理流程  
profile

伺服器清單

重新整理(R) | 中斷連

伺服器  
本機  
資料館  
EC10000  
MAPS  
MAPSGF  
MAPSSA  
SASHELP  
SASUSER  
WORK  
檔案  
私用 OLAP 伺服器

「profile」的屬性

一般  
欄  
進階  
摘要

欄

名稱	類型	長度	格式	輸入格式	標籤
ID	字元	9			
CNT_TBM	數值	8			
CNT_ATM	數值	8			
CNT_POS	數值	8			
CNT_CSC	數值	8			
CNT_TOT	數值	8			

複製到剪貼簿(B)

顯示資料項目的欄屬性 (名稱、類型、長度、格式、輸入格式及標籤)。您可以使用特定欄的 [屬性] 視窗來修改該欄的這些屬性。

其他 (F1)...

確定 取消

30	000269176	107	78	22	7	214
31	000273510	110	15	6	21	152

專案樹狀結構

處理流程

- 開啟「profile」(O)
- 匯出(X)
  - 匯出「profile」(X)...
  - 匯出「profile」作為專案中的步驟(S)...
- 篩選和排序(F)...
- 查詢產生器(Q)...
- 發行(L)...
- 傳送至(N)
- 複製(C)
- 貼上(P)
- 刪除(D)
- 重新命名(M)
- 關閉資料(S)
- 屬性(I)

ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
000041260	34	3	3	9	49
000076885	42	3	6	1	52
000089150	20	15	2	2	39
000094512	83	20	49	3	155
000096396	33	9	49	17	108
000098062	22	5	9	1	37
000122779	21	27	2	1	51
000125838	127	3	2	3	135
000137317	52	20	2	11	85
000147896	101	40	8	1	150
000150206	156	52	80	4	292
000160501	31	32	42	12	117
000176820	107	3	2	1	113
000187784	54	14	7	24	99
000189200	92	3	2	2	99
000198327	44	27	17	3	91
000200237	146	3	2	1	152
000207980	39	5	2	4	50
000227355	180	3	2	10	195
000228317	46	14	2	1	63
000228906	25	3	2	1	31
000229984	15	3	2	1	21
000241022	13	3	2	1	19
000253260	65	3	2	1	71
000257910	35	3	2	1	41
000261717	11	5	2	1	19
000266600	15	3	2	1	21
000269176	107	78	22	7	214
000273510	110	15	6	21	152

伺服器清單

重新整理(R) | 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
      - EC100001
      - MAPS
      - MAPSGFK
      - MAPSSAS
      - SASHELP
      - SASUSER
      - WORK
    - 檔案
  - 私用 OLAP 伺服器

匯出

儲存位置(I): EM\_Data

名稱	修改日期	類型
credit.sas7bdat	2013/4/21 下午 05:22	SAS
inq2006.sas7bdat	2013/5/29 下午 01:20	SAS
profile.sas7bdat	2006/9/23 下午 10:08	SAS
webstation.sas7bdat	2006/9/25 上午 01:45	SAS

檔案名稱(N): profile

檔案類型(T): SAS 資料檔案 (V7 完整名稱) (\*.sas7bdat)

儲存(S) 取消

送至(N) |

CSC	CNT_TOT
9	49
18	84
36	216
1	52
2	39
3	155
17	108
1	37
1	51
3	135
11	85
1	150
4	292
12	117
1	113
24	99
2	99
3	91
1	152
4	50
10	195
1	63
1	31
1	21
1	19
1	71
1	41
1	19
1	21
7	214
21	152

未選取設定檔

SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P) 工具

專案樹狀結構

- 處理流程
  - profile

篩選和排序(L)

1	000041360
2	000052094
3	000057340
4	000076885
5	000089150
6	000094512
7	000096396
8	000098062
9	000122779
10	000125838
11	000137317
12	000147896
13	000150206
14	000160501
15	000176820
16	000187784
17	000189200
18	000198327
19	000200237
20	000207980
21	000227355
22	000228317
23	000228906
24	000229984
25	000241022
26	000253260
27	000257910
28	000261717
29	000266600
30	000269176
31	000273510

伺服器清單

- 重新整理(R) | 中斷連線(D)
- 伺服器
  - 本機
    - 資料館
      - EC100001
      - MAPS
      - MAPSGFK
      - MAPSSAS
      - SASHELP
      - SASUSER
      - WORK
    - 檔案
  - 私有 OLAP 伺服器

就緒

匯出

儲存位置(I): EM\_Data

名稱

名稱	修改日期	類型
credit.sas7bdat	2013/4/21 下午 05:22	SAS
inq2006.sas7bdat	2013/5/29 下午 01:20	SAS
profile.sas7bdat	2006/9/23 下午 10:08	SAS
webstation.sas7bdat	2006/9/25 上午 01:45	SAS

檔案名稱(N):

檔案類型(T): SAS 資料檔案 (V7 完整名稱) (\*.sas7bdat)

Microsoft Excel 活頁簿 (\*.xlsx)

Microsoft Excel 97-2003 活頁簿 (\*.xls)

Microsoft Access 2002-2003 資料庫 (\*.mdb)

文字檔 (以逗號分隔) (\*.csv)

文字檔 (\*.txt)

文字檔 (以 Tab 分隔) (\*.tab)

文字檔 (以空格分隔) (\*.txt)

所有 HTML 檔 (\*.htm;\*.html)

儲存(S) 取消

SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P) 工具

專案樹狀結構

- 處理流程
  - profile

篩選和排序(L)

1	000041360
2	000052094
3	000057340
4	000076885
5	000089150
6	000094512
7	000096396
8	000098062
9	000122779
10	000125838
11	000137317
12	000147896
13	000150206
14	000160501
15	000176820
16	000187784
17	000189200
18	000198327
19	000200237
20	000207980
21	000227355
22	000228317
23	000228906
24	000229984
25	000241022
26	000253260
27	000257910
28	000261717
29	000266600
30	000269176
31	000273510

伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
      - EC100001
      - MAPS
      - MAPSGFK
      - MAPSSAS
      - SASHELP
      - SASUSER
      - WORK
    - 檔案
  - 私用 OLAP 伺服器

就緒

匯出

儲存位置(I): EM\_Data

名稱

修改日期

類型

這個資料夾是空的。

檔案名稱(N): profile\_excel

檔案類型(T): Microsoft Excel 活頁簿 (\*.xlsx)

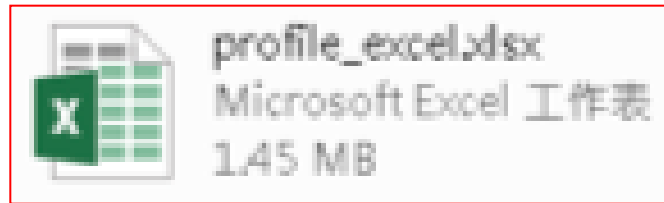
儲存(S) 取消

# Export SAS .sas7bdat to to Excel .xlsx File

The screenshot displays the SAS Enterprise Guide interface. The main window shows a data table with columns: ID, CNT\_TBM, CNT\_ATM, CNT\_POS, CNT\_CSC, and CNT\_TOT. A dialog box titled "Enterprise Guide" is overlaid on the table, stating "正在匯出「profile」..." (Exporting 'profile'...). On the right side, a file list shows two files: "profile.sas7bdat SAS Data Set 5.39 MB" and "profile\_excel.xlsx Microsoft Excel 工作表 1.45 MB". A red arrow points from the first file to the second, indicating the export operation. The status bar at the bottom indicates "正在執行: 匯出工作 - 正在執行..." (Running: Export job - running...).

ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	34	3	3	9	49
2	44	17	5	18	84
3	122	26	32	36	216
4	42	3	6	1	52
5	20	15	2	2	39
6	83	20	49	3	155
7	33	9	49	17	108
8	22	5	9	1	37
9	21	27	2	1	51
10	127	3	2	3	135
11	50	20	2	11	85
12	000147896		8	1	150
13	000150206		80	4	292
14	000160501		42	12	117
15	000176820		2	1	113
16	000187784	54	14	7	99
17	000189200	92	3	2	99
18	000198327	44	27	17	91
19	000200237	146	3	2	152
20	000207980	39	5	2	50
21	000227355	180	3	2	195
22	000228317	46	14	2	63
23	000228906	25	3	2	31
24	000229984	15	3	2	21
25	000241022	13	3	2	19
26	000253260	65	3	2	71
27	000257910	35	3	2	41
28	000261717	11	5	2	19
29	000266600	15	3	2	21
30	000269176	107	78	22	214
31	000273510	110	15	6	212

# Import Excel File to SAS EG





# Import Excel File to SAS EG

The screenshot shows the SAS Enterprise Guide interface. On the left, there is a '專案樹狀結構' (Project Tree) pane with '處理流程' (Process) and '伺服器清單' (Server List) sections. The '伺服器清單' section shows a tree structure: 伺服器 (Servers) > 本機 (Local) > 資料館 (Data Warehouse) > 檔案 (Files) > 私人 OLAP 伺服器 (Private OLAP Server). The main window is a file explorer titled '開啟' (Open) showing the 'EM\_Data' folder. A table lists the contents:

名稱	類型	大小	修改日期
profile_excel.xlsx	Microsoft Excel 工作表	1,491 KB	2014/3/

At the bottom of the file explorer, the '檔案名稱(N):' field contains 'profile\_excel.xlsx' and the '檔案類型(T):' dropdown is set to '可匯入的資料檔案 (\*.xls;\*.xlsx;\*.xlsm;\*.xlsb;\*.mdb;\*.accdb;\*.txt;\*.csv;\*.asc;\*.tab;\*.htm;\*.html)'. The status bar at the bottom left shows '就緒' (Ready) and the bottom right shows '未選取設定檔' (No profile selected).

## 專案樹狀結構

處理流程

- profile\_excel.xlsx
  - 匯入資料 (profile\_excel.xlsx)

## 處理流程

執行(R) 停止(S) 匯出(O)



## 從 profile\_excel.xlsx 匯入資料



1 / 4 指定資料

對於其他需要使用 SAS 資料檔案以進行資料分析和報表的工作，「匯入資料」精靈可將非 SAS 資料轉換成 SAS 資料檔案。

## 來源資料檔案

位置(L): 本機檔案系統  
 檔案路徑(P): C:\DATA\EM\_Data\profile\_excel.xlsx  
 資料類型(T): Excel 活頁簿

## 輸出 SAS 資料集

SAS 伺服器(S): Local  
 資料館(R): WORK  
 資料集(A): profile\_excel







## 伺服器清單

重新整理(R) 中斷連線(D)

伺服器

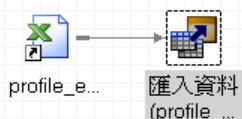
- 本機
  - 資料館
  - 檔案
- 私用 OLAP 伺服器

## 專案樹狀結構

- 處理流程
- profile\_excel.xlsx
  - 匯入資料 (profile\_excel.xlsx)

## 處理流程

▶ 執行(R) | ■ 停止(S) | 匯出(O)



## 從 profile\_excel.xlsx 匯入資料



2 / 4 選取資料來源



## 選取範圍

 使用工作表(W)

profile

 範圍的第一列包含欄位名稱(M) 重新命名欄以符合 SAS 命名規則(R)。 使用工作表內特定範圍的儲存格(S)

左上方儲存格(T):

右下方儲存格(L):

 必要時展開列範圍(X)

重設範圍(A)

 使用預先定義的命名範圍(P)

&lt; 上一步(B)

下一步(N) &gt;

完成(F)

取消

說明

## 伺服器清單



重新整理(R) | 中斷連線(D) &gt;&gt;

- 伺服器
- 本機
  - 資料館
  - 檔案
- 私有 OLAP 伺服器

專案樹狀結構

處理流程

- 處理流程
  - profile\_excel.xlsx
    - 匯入資料 (profile\_excel.xls)

處理流程

執行(R) 停止(S) 匯出(O)

profile\_e... → 匯入資料 (profile\_...)

伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
    - 檔案
  - 私用 OLAP 伺服器

從 profile\_excel.xlsx 匯入資料

3 / 4 定義欄位特性

sas

選取欄並定義特性(S):

Inc	來源名稱	名稱	標籤	類型	來源格式	長度	輸出格式	輸出輸入格式
<input checked="" type="checkbox"/>	ID	ID	ID	字串	\$CHAR9.	9	\$CHAR9.	\$CHAR9.
<input checked="" type="checkbox"/>	CNT_TBM	CNT_TBM	CNT_TBM	數字	BEST12.	8	BEST12.	BEST12.
<input checked="" type="checkbox"/>	CNT_ATM	CNT_ATM	CNT_ATM	數字	BEST12.	8	BEST12.	BEST12.
<input checked="" type="checkbox"/>	CNT_POS	CNT_POS	CNT_POS	數字	BEST12.	8	BEST12.	BEST12.
<input checked="" type="checkbox"/>	CNT_CSC	CNT_CSC	CNT_CSC	數字	BEST12.	8	BEST12.	BEST12.
<input checked="" type="checkbox"/>	CNT_TOT	CNT_TOT	CNT_TOT	數字	BEST12.	8	BEST12.	BEST12.

全選(A) 全部清除(C) 修改(M)...

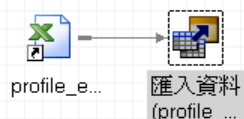
<上一步(B) 下一步(N) > 完成(F) 取消 說明

## 專案樹狀結構

- 處理流程
- profile\_excel.xlsx
  - 匯入資料 (profile\_excel.xlsx)

## 處理流程

▶ 執行(R) | ■ 停止(S) | 匯出(O)



## 從 profile\_excel.xlsx 匯入資料



4 / 4 進階選項

- 將資料內嵌在產生的 SAS 程式碼中(E)。
- 盡可能使 SAS/ACCESS Interface to PC Files 匯入資料(P)。
- 從文字型資料檔案移除可能導致傳輸錯誤的字元(R)。

&lt; 上一步(B)

下一步(N) &gt;

完成(F)

取消

說明

## 伺服器清單

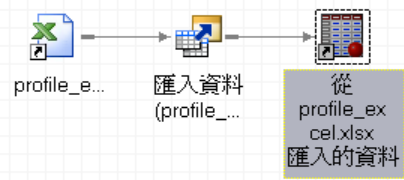
重新整理(R) | 中斷連線(D) &gt;&gt;

- 伺服器
  - 本機
    - 資料館
    - 檔案
  - 私有 OLAP 伺服器

專案樹狀結構

- 處理流程
  - profile\_excel.xlsx
    - 匯入資料 (profile\_excel.xlsx)

處理流程



伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
    - 檔案
  - 私有 OLAP 伺服器

專案樹狀結構

匯入資料 (profile\_excel.xlsx[profile])

處理流程

- profile\_excel.xlsx
  - 匯入資料 (profile\_excel.xlsx)

程式碼 記錄檔 輸出資料

修改工作(Y) 篩選和排序(L) 查詢產生器(Q) | 資料(D) 描述(B) 圖形(G) 分析(Z) | 匯出(X) 傳送至(N)

	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34	3	3	9	49
2	000052094	44	17	5	18	84
3	000057340	122	26	32	36	216
4	000076885	42	3	6	1	52
5	000089150	20	15	2	2	39
6	000094512	83	20	49	3	155
7	000096396	33	9	49	17	108
8	000098062	22	5	9	1	37
9	000122779	21	27	2	1	51
10	000125838	127	3	2	3	135
11	000137317	52	20	2	11	85
12	000147896	101	40	8	1	150
13	000150206	156	52	80	4	292
14	000160501	31	32	42	12	117
15	000176820	107	3	2	1	113
16	000187784	54	14	7	24	99
17	000189200	92	3	2	2	99
18	000198327	44	27	17	3	91
19	000200237	146	3	2	1	152
20	000207980	39	5	2	4	50
21	000227355	180	3	2	10	195
22	000228317	46	14	2	1	63
23	000228906	25	3	2	1	31
24	000229984	15	3	2	1	21
25	000241022	13	3	2	1	19
26	000253260	65	3	2	1	71
27	000257910	35	3	2	1	41
28	000261717	11	5	2	1	19
29	000266600	15	3	2	1	21

伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
    - 檔案
  - 私用 OLAP 伺服器

專案樹狀結構

匯入資料 (profile\_excel.xlsx[profile])

 處理流程  
 profile\_excel.xlsx  
 匯入資料 (profile\_excel.xlsx)

程式碼 記錄檔 輸出資料

修改工作(Y) 篩選和排序(L) 查詢產生器(Q) | 資料(D) 描述(B) 圖形(G) 分析(Z) | 匯出(X) 傳送至(N)

	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOI
99972	875475072	81	37	2	1	121
99973	875478517	56	15	10	3	84
99974	875481914	49	28	91	87	255
99975	875484829	233	10	5	1	249
99976	875488178	104	3	2	1	110
99977	875505005	27	47	37	53	164
99978	875508051	28	3	2	1	34
99979	875511290	13	5	2	1	21
99980	875514158	54	8	18	4	84
99981	875528729	71	4	14	1	90
99982	875536451	109	3	2	11	125
99983	875555772	26	22	2	1	51
99984	875568109	124	15	2	12	153
99985	875577833	115	80	9	5	209
99986	875589010	48	3	82	8	141
99987	875593202	143	9	2	4	158
99988	875594835	51	3	2	1	57
99989	875599149	58	15	27	17	117
99990	875606621	25	3	2	1	31
99991	875619290	103	6	2	10	121
99992	875625622	10	4	2	1	17
99993	875625635	41	3	2	1	47
99994	875628606	20	14	23	1	58
99995	875636141	132	3	2	11	148
99996	875638608	24	6	2	1	33
99997	875656523	79	9	2	7	97
99998	875663288	44	111	57	9	221
99999	875670226	10	5	2	1	18
100000	875674475	128	18	20	1	167

伺服器清單

重新整理(R) 中斷連線(D)

 伺服器  
 本機  
 資料館  
 檔案  
 私有 OLAP 伺服器



# Export Excel .xlsx File to SAS .sas7bdat File

The screenshot displays the SAS Enterprise Guide interface. The main workspace shows a workflow with three steps: 'profile\_e...', '匯入資料 (profile\_...)', and '從 profile\_excel.xlsx 匯入的資料'. A context menu is open over the third step, with '匯出(X)' selected. A secondary menu is also visible, showing options like '匯出「從 profile\_excel.xlsx 匯入的資料」(X)...' and '匯出「從 profile\_excel.xlsx 匯入的資料」作為專案中的步驟(S)...'. The left sidebar shows the '專案樹狀結構' (Project Tree) and '伺服器清單' (Server List). The status bar at the bottom indicates '就緒' (Ready) and '未選取設定檔' (No profile selected).

SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P) 工具(T) 說明(H) | 處理流程

專案樹狀結構

處理流程

profile\_excel.xlsx

匯入資料 (profile\_excel.xlsx[profi

執行(R) 停止(S) | 匯出(X) 排程(D) 縮放(Z) | 專案記錄檔(L) 屬性(I)

profile\_e... 匯入資料 (profile\_... 從 profile\_excel.xlsx 匯入的資料

匯出「從 profile\_excel.xlsx 匯入的資料」(X)...

匯出「從 profile\_excel.xlsx 匯入的資料」作為專案中的步驟(S)...

開啟(O)

匯出(X)

篩選和排序(F)...

查詢產生器(Q)...

從「從 profile\_excel.xlsx 匯入的資料」執行分支(B)

發行(L)...

傳送到(S)

將「從 profile\_excel.xlsx 匯入的資料」連結至(K)...

複製(C)

貼上(P)

刪除(D)

重新命名(M)

屬性(I)

伺服器清單

重新整理(R) 中斷連線(D)

伺服器

本機

資料館

檔案

私有 OLAP 伺服器

就緒

未選取設定檔

SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P) 工...

專案樹狀結構

處理流程

處理流程

- profile\_excel.xlsx
- 匯入資料 (profile\_excel.xlsx[profi

伺服器清單

重新整理(R) 中斷連線(D)

- 伺服器
  - 本機
    - 資料館
    - 檔案
  - 私用 OLAP 伺服器

匯出

儲存位置(I): EM\_Data

名稱	修改日期	類型
credit.sas7bdat	2013/4/21 下午 05:22	SA...
inq2006.sas7bdat	2013/5/29 下午 01:20	SA...
profile.sas7bdat	2006/9/23 下午 10:08	SA...
webstation.sas7bdat	2006/9/25 上午 01:45	SA...

檔案名稱(N): Profile\_SAS

檔案類型(T): SAS 資料檔案 (V7 完整名稱) (\*.sas7bdat)

儲存(S) 取消

就緒

電腦 > System (C:) > DATA > EM\_Data

搜尋 EM\_Data

組合管理 使用 SAS Enterprise Guide 5.1 開啟 新增資料夾

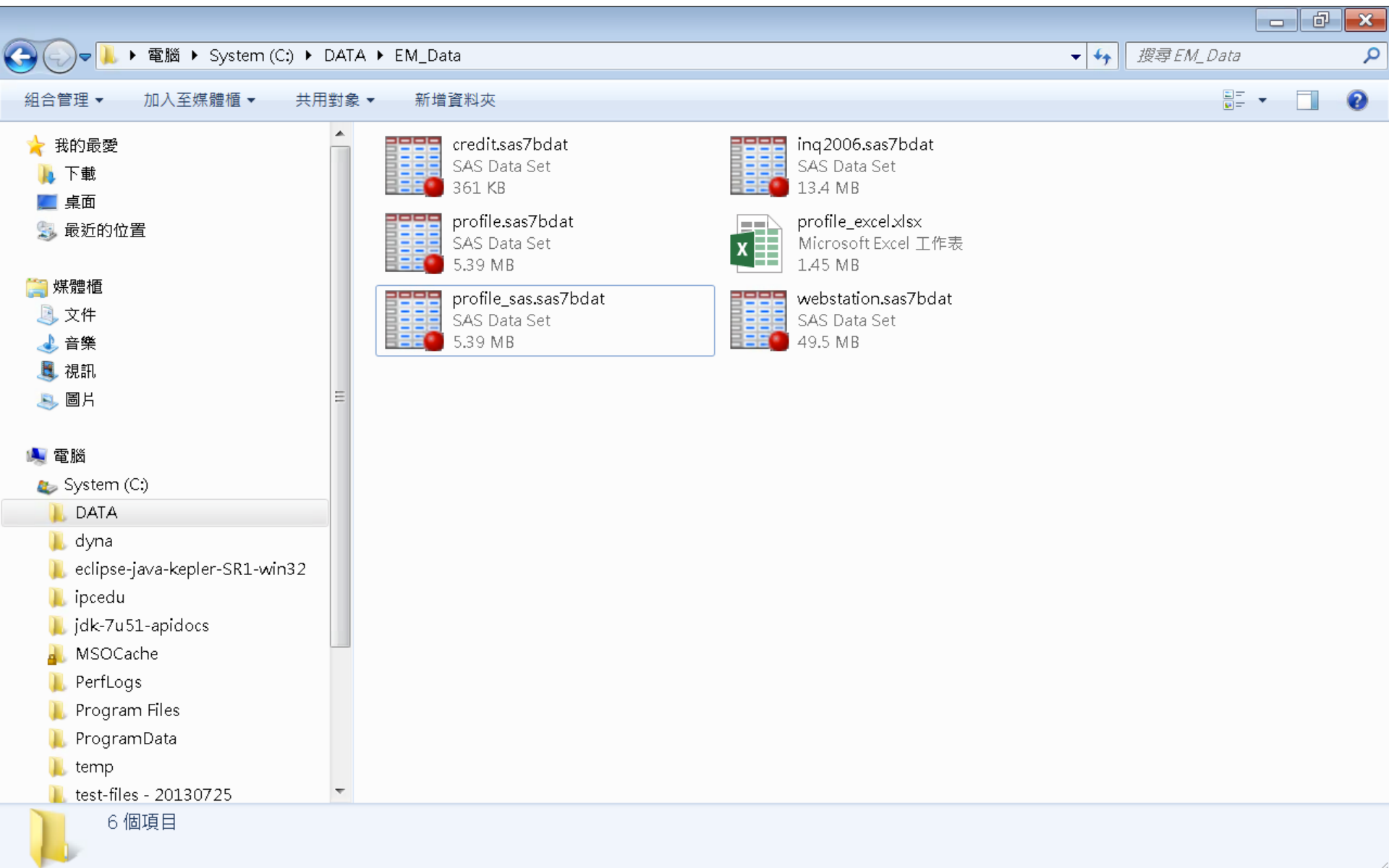
我的最愛  
 下載  
 桌面  
 最近的位置

媒體櫃  
 文件  
 音樂  
 視訊  
 圖片

電腦  
 System (C:)  
 DATA  
 dyna  
 eclipse-java-kepler-SR1-win32  
 ipcedu  
 jdk-7u51-apidocs  
 MSOCache  
 PerfLogs  
 Program Files  
 ProgramData  
 temp  
 test-files - 20130725

名稱	類型	大小	修改日期
credits.sas7bdat	SAS Data Set	361 KB	2013/4/2...
inq2006.sas7bdat	SAS Data Set	13,729 KB	2013/5/2...
profile.sas7bdat	SAS Data Set	5,521 KB	2006/9/2...
profile_excel.xlsx	Microsoft Excel 工作表	1,491 KB	2014/3/1...
profile_sas.sas7bdat	SAS Data Set	5,521 KB	2014/3/1...
webstation.sas7bdat	SAS Data Set	50,789 KB	2006/9/2...

選取了 3 個項目 修改日期: 2006/9/23 下午 10:08 ... 建立日期: 2014/3/19 上午 08:43  
 大小: 12.2 MB



# Profile\_Excel.xlsx

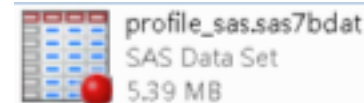


Microsoft Excel interface showing the ribbon (File, Home, Insert, Layout, Formulas, Data, Review, View, Help) and the spreadsheet data.

The spreadsheet contains the following data:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	C
1	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT									
2	000041360	34	3	3	9	49									
3	000052094	44	17	5	18	84									
4	000057340	122	26	32	36	216									
5	000076885	42	3	6	1	52									
6	000089150	20	15	2	2	39									
7	000094512	83	20	49	3	155									
8	000096396	33	9	49	17	108									
9	000098062	22	5	9	1	37									
10	000122779	21	27	2	1	51									
11	000125838	127	3	2	3	135									
12	000137317	52	20	2	11	85									
13	000147896	101	40	8	1	150									
14	000150206	156	52	80	4	292									
15	000160501	31	32	42	12	117									
16	000176820	107	3	2	1	113									
17	000187784	54	14	7	24	99									
18	000189200	92	3	2	2	99									
19	000198327	44	27	17	3	91									
20	000200237	146	3	2	1	152									
21	000207000	20	5	2	4	50									

# Profile\_SAS.sas7bdat



SAS Enterprise Guide

檔案(F) 編輯(E) 檢視(V) 工作(K) 程式(P) 工具(T) 說明(H) | 處理流程

專案樹狀結構

處理流程  
profile\_sas

伺服器清單

重新整理(R) | 中斷連線(D)

伺服器  
私有 OLAP 伺服器

篩選和排序(L) | 查詢產生器(Q) | 資料(D) | 描述(B) | 圖形(G) | 分析(Z) | 匯出(X) | 傳送至(N)

	ID	CNT_TBM	CNT_ATM	CNT_POS	CNT_CSC	CNT_TOT
1	000041360	34	3	3	9	49
2	000052094	44	17	5	18	84
3	000057340	122	26	32	36	216
4	000076885	42	3	6	1	52
5	000089150	20	15	2	2	39
6	000094512	83	20	49	3	155
7	000096396	33	9	49	17	108
8	000098062	22	5	9	1	37
9	000122779	21	27	2	1	51
10	000125838	127	3	2	3	135
11	000137317	52	20	2	11	85
12	000147896	101	40	8	1	150
13	000150206	156	52	80	4	292
14	000160501	31	32	42	12	117
15	000176820	107	3	2	1	113
16	000187784	54	14	7	24	99
17	000189200	92	3	2	2	99
18	000198327	44	27	17	3	91
19	000200237	146	3	2	1	152
20	000207980	39	5	2	4	50
21	000227355	180	3	2	10	195
22	000228317	46	14	2	1	63
23	000228906	25	3	2	1	31
24	000229984	15	3	2	1	21
25	000241022	13	3	2	1	19
26	000253260	65	3	2	1	71
27	000257910	35	3	2	1	41
28	000261717	11	5	2	1	19
29	000266600	15	3	2	1	21
30	000269176	107	78	22	7	214
31	000273510	110	15	6	21	152

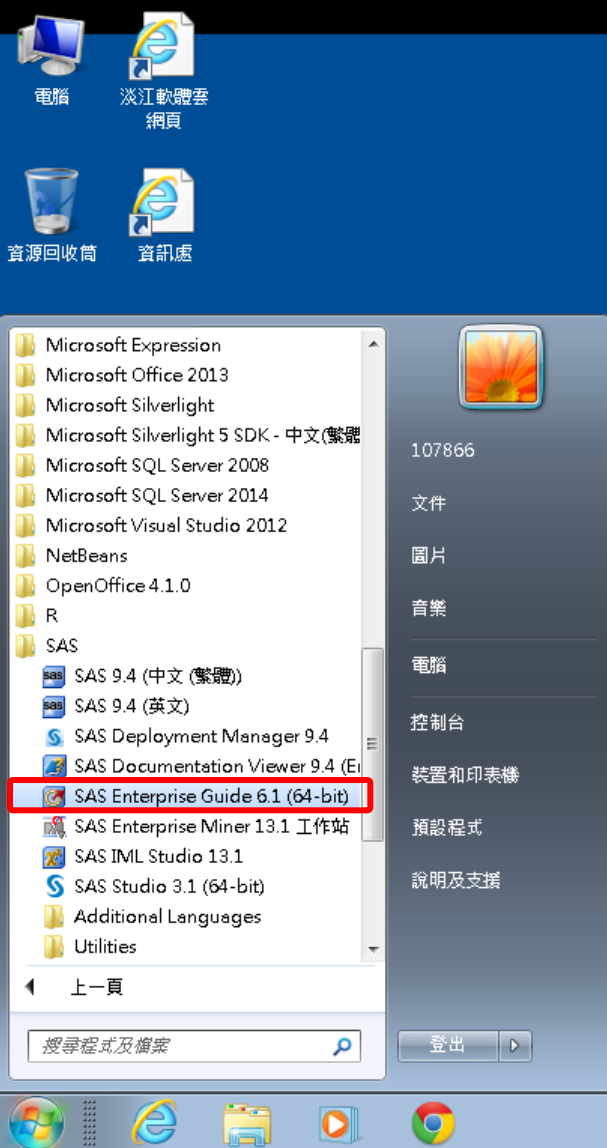
就緒

未選取設定檔

# **SAS Locale Setup Manager**

## **→ English UI**

# SAS Enterprise Guide (SAS EG)

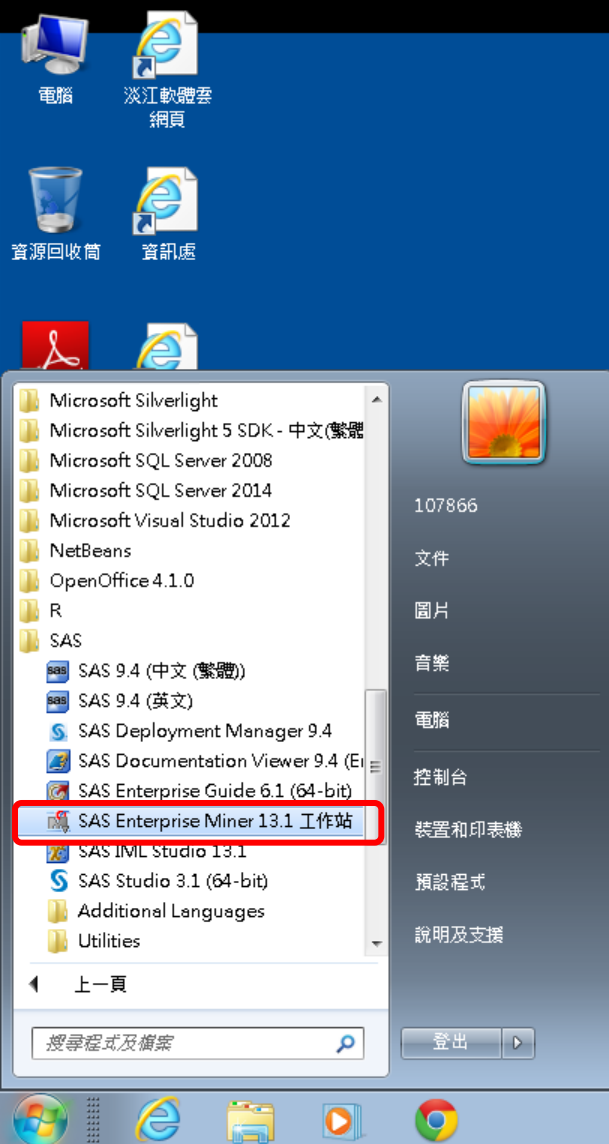


## 淡江軟體雲使用注意事項

- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室，或來電 26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。

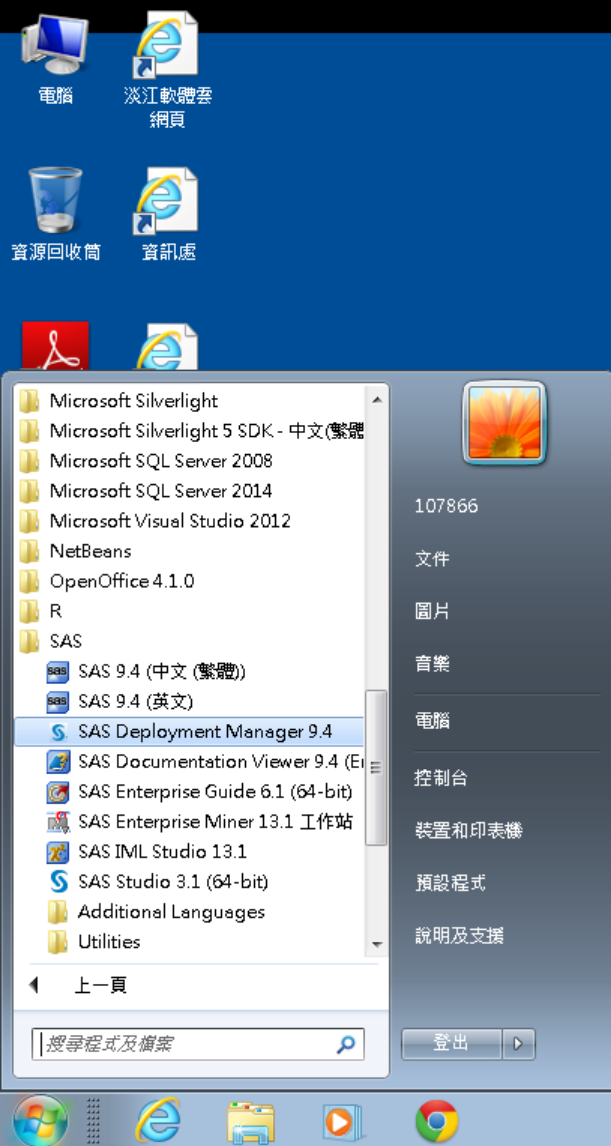


# SAS Enterprise Miner 13.1 (SAS EM)



## 淡江軟體雲使用注意事項

- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室,或來電 26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。



## 淡江軟體雲使用注意事項

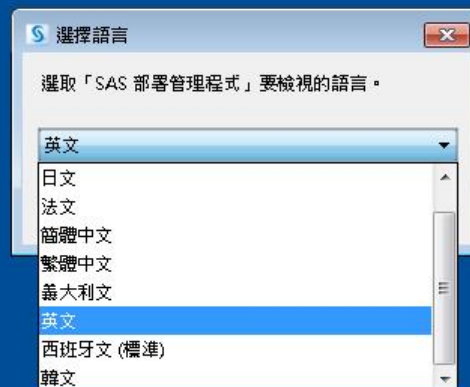
- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室,或來電 26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。



## 淡江軟體雲使用注意事項

- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室,或來電 26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。



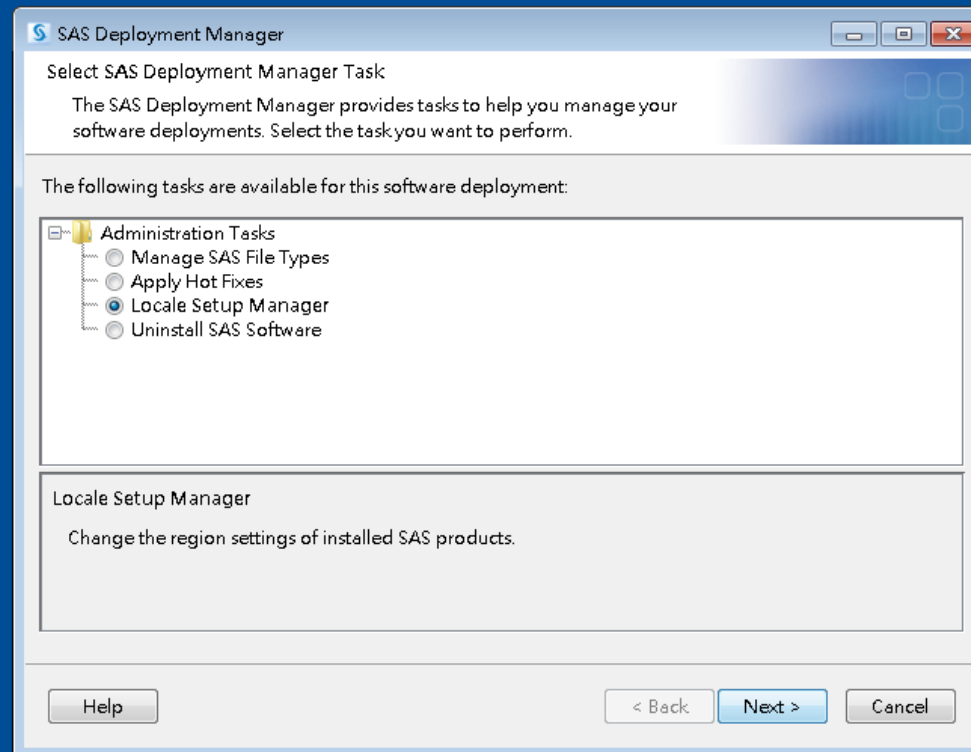


## 淡江軟體雲使用注意事項

- 登入淡江軟體雲時，會自動產生一部新的雲端主機。登出雲端主機前，請先將檔案儲存於您攜帶的隨身碟、淡江大學網路隨身碟、Google雲端硬碟...等，否則登出後雲端主機的所有資料將會被刪除，並回到您的個人主機畫面。
- 為免雲端主機帳號被冒用，每次使用完請登出。
- 軟體雲相關操作說明，請點選雲端主機桌面上『操作手冊』圖示。
- 軟體雲使用問題請至B212辦公室，或來電26215656-2129 為您服務。
- 請尊重智慧財產權，嚴禁安裝不法軟體，違法者自負相關法律責任。



## 淡江軟體雲使用注意事項



淡江軟體雲時，會自動產生一部新  
端主機。登出雲端主機前，請先將  
儲存於您攜帶的隨身碟、淡江大學  
隨身碟、Google雲端硬碟...等，  
登出後雲端主機的所有資料將會被  
，並回到您的個人主機畫面。

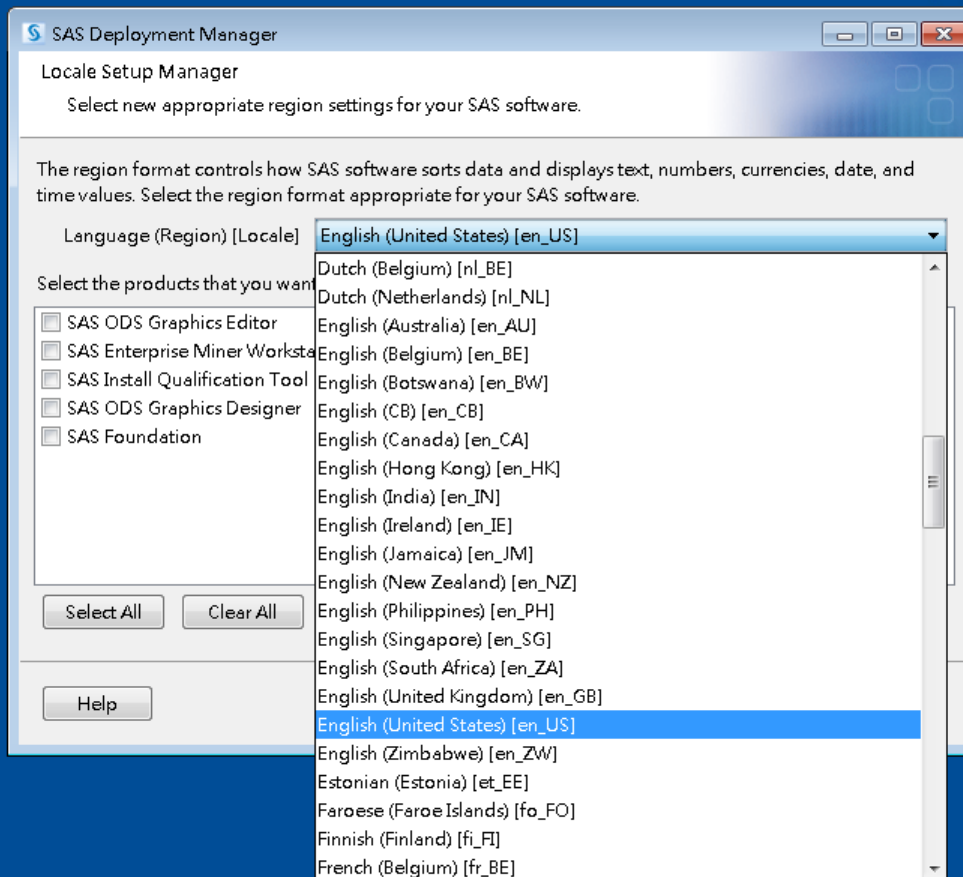
雲端主機帳號被冒用，每次使用完  
出。

查相關操作說明，請點選雲端主機  
上『操作手冊』圖示。

查使用問題請至B212辦公室，或來電  
L5656-2129 為您服務。

重智慧財產權，嚴禁安裝不法軟  
違法者自負相關法律責任。

## 淡江軟體雲使用注意事項



淡江軟體雲時，會自動產生一部新  
端主機。登出雲端主機前，請先將  
儲存於您攜帶的隨身碟、淡江大學  
隨身碟、Google雲端硬碟...等，  
登出後雲端主機的所有資料將會被  
，並回到您的個人主機畫面。

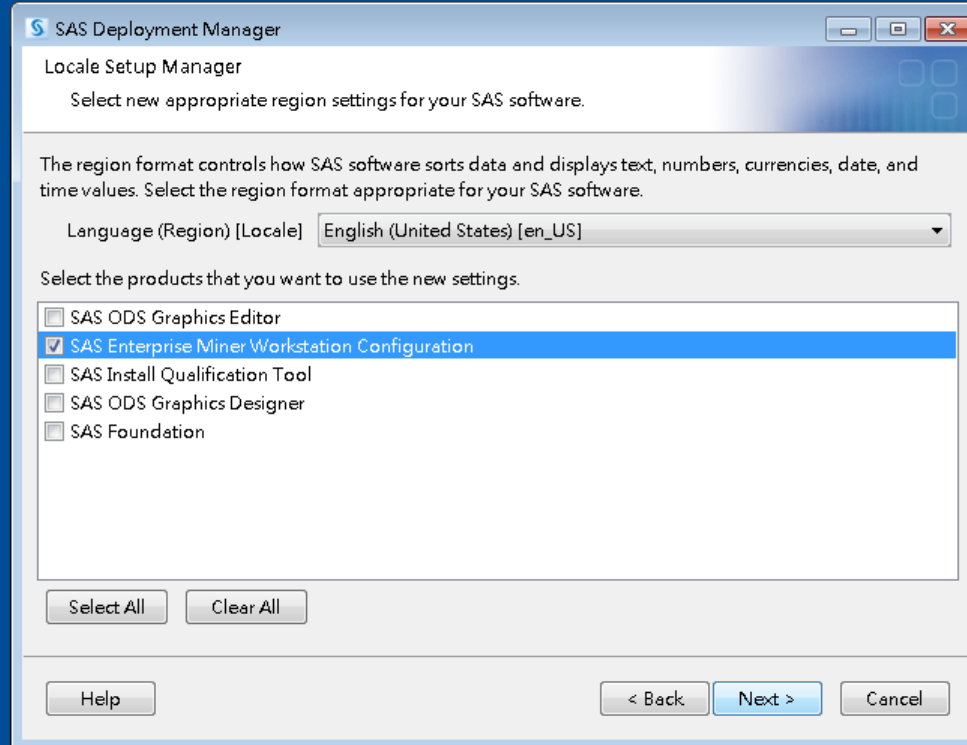
雲端主機帳號被冒用，每次使用完  
出。

與相關操作說明，請點選雲端主機  
上『操作手冊』圖示。

與使用問題請至B212辦公室，或來電  
15656-2129 為您服務。

與智慧財產權，嚴禁安裝不法軟  
違法者自負相關法律責任。

## 淡江軟體雲使用注意事項



淡江軟體雲時，會自動產生一部新  
端主機。登出雲端主機前，請先將  
儲存於您攜帶的隨身碟、淡江大學  
隨身碟、Google雲端硬碟...等，  
登出後雲端主機的所有資料將會被  
，並回到您的個人主機畫面。

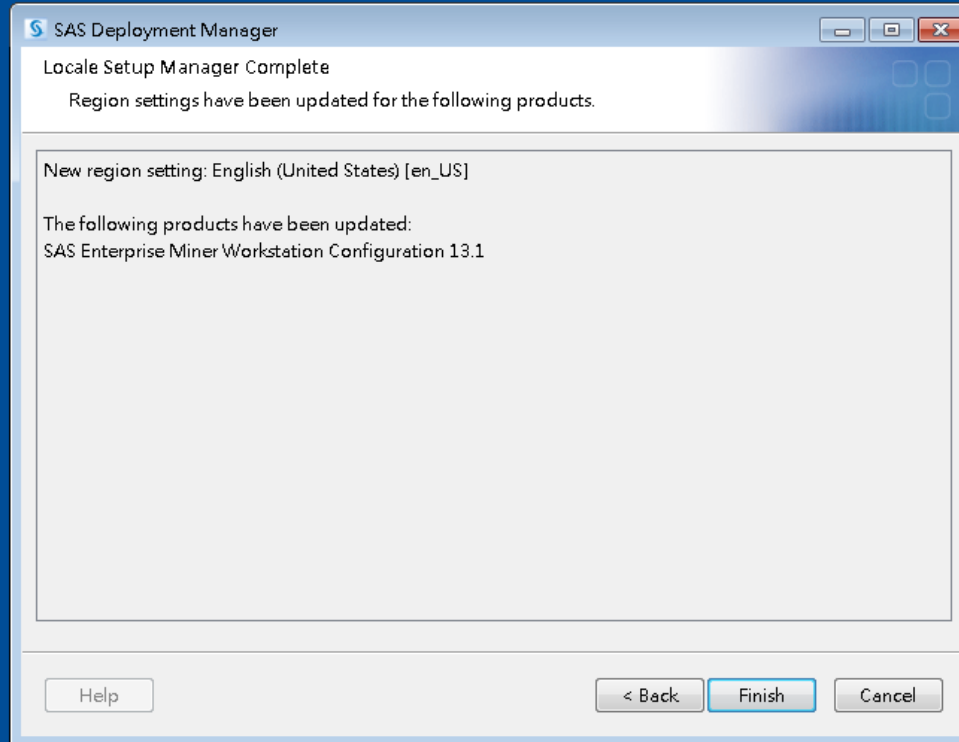
雲端主機帳號被冒用，每次使用完  
出。

與相關操作說明，請點選雲端主機  
上『操作手冊』圖示。

與使用問題請至B212辦公室，或來電  
L5656-2129 為您服務。

與智慧財產權，嚴禁安裝不法軟  
違法者自負相關法律責任。

## 淡江軟體雲使用注意事項



淡江軟體雲時，會自動產生一部新  
端主機。登出雲端主機前，請先將  
儲存於您攜帶的隨身碟、淡江大學  
隨身碟、Google雲端硬碟...等，  
登出後雲端主機的所有資料將會被  
，並回到您的個人主機畫面。

雲端主機帳號被冒用，每次使用完  
出。

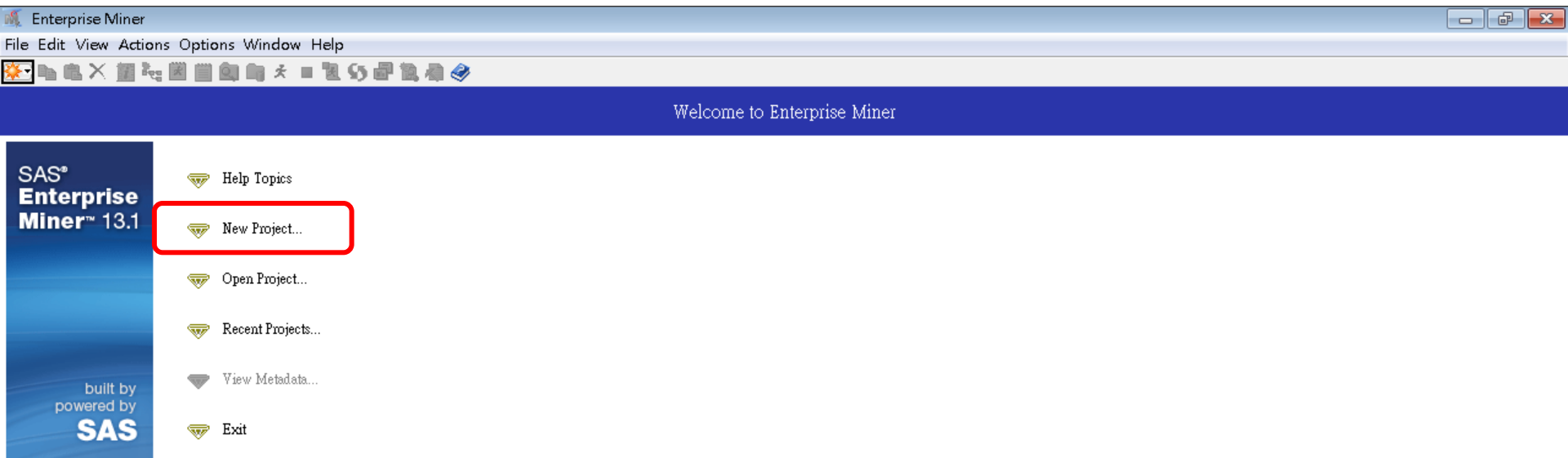
與相關操作說明，請點選雲端主機  
上『操作手冊』圖示。

與使用問題請至B212辦公室，或來電  
15656-2129 為您服務。

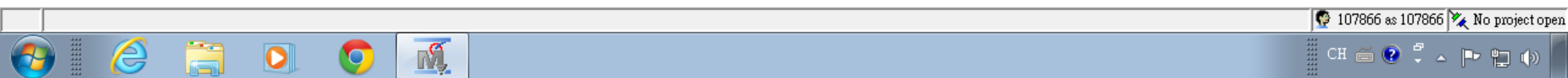
重智慧財產權，嚴禁安裝不法軟  
違法者自負相關法律責任。



# SAS Enterprise Miner (SAS EM)



## SAS Enterprise Miner (SAS EM) English UI



# Reference

- 資料採礦運用：以SAS Enterprise Miner為工具，李淑娟，2015，SAS賽仕電腦軟體
- Jim Georges, Jeff Thompson and Chip Wells, Applied Analytics Using SAS Enterprise Miner, SAS, 2010
- SAS Enterprise Miner Course Notes, 2014, SAS
- SAS Enterprise Miner Training Course, 2014, SAS
- SAS Enterprise Guide Training Course, 2014, SAS