

IBM SPSS Statistics 28 簡易手冊



附註

使用此資訊和支援的產品之前，請先閱讀 [第 79 頁的『注意事項』](#) 中的資訊。

產品資訊

若新版本未聲明，則此版本便適用於 IBM® SPSS® Statistics 28.0.0 版以及之後發行的所有版本和修正。

© Copyright International Business Machines Corporation .

目錄

第 1 章簡介.....	1
範例檔案.....	1
開啟資料檔.....	1
執行分析.....	2
建立圖表.....	4
第 2 章讀取資料.....	7
IBM SPSS Statistics 資料檔的基本架構.....	7
讀取 IBM SPSS Statistics 資料檔.....	7
讀取 Excel 資料.....	8
從資料庫讀取資料.....	10
從文字檔讀取資料.....	12
第 3 章使用資料編輯器.....	17
輸入數字資料.....	17
輸入字串資料.....	18
定義資料.....	19
新增變數標籤.....	19
變更變數類型和格式.....	20
新增值標籤.....	20
處理遺漏的資料.....	21
數字變數的遺漏值.....	21
字串變數的遺漏值.....	22
第 4 章檢查個別變數的摘要統計量.....	23
測量層級.....	23
類別資料的摘要量數.....	23
類別資料的圖表.....	24
尺度變數的摘要量數.....	25
尺度變數的直方圖.....	26
第 5 章建立並編輯圖表.....	27
圖表建立基礎.....	27
使用圖表建置器展示區.....	27
定義變數與統計量.....	28
新增文字.....	29
建立圖表.....	30
第 6 章輸出的處理.....	33
使用檢視器.....	33
使用「樞紐表編輯器」.....	34
存取輸出定義.....	34
樞紐表.....	35
建立和顯示階層.....	36
編輯表格.....	36
隱藏列和欄.....	37
變更資料顯示格式.....	37
表格格式集.....	38
使用預先定義的格式.....	38
自訂表格格式集樣式.....	39

變更預設的表格格式.....	41
自訂初始顯示設定.....	42
顯示變數和值標籤.....	43
使用其他應用程式中的結果.....	44
將結果貼上為 Word 表格.....	44
將結果以文字形式貼上.....	45
將結果匯出至 Microsoft Word、PowerPoint 和 Excel 檔.....	45
匯出結果至 PDF.....	51
匯出結果至 HTML.....	53
第 7 章使用語法.....	55
貼上語法.....	55
編輯語法.....	56
開啟與執行語法檔.....	57
使用中斷點.....	57
第 8 章修改資料值.....	59
從尺度變數建立類別變數.....	59
計算新變數.....	60
使用表示式中的函數.....	61
使用條件表示式.....	62
日期與時間的操作.....	62
計算兩個日期之間的時間長度.....	63
在日期上新增期間.....	64
第 9 章排序與選取資料.....	67
排序資料.....	67
分割檔案處理.....	67
排序觀察值以進行分割檔案處理.....	69
開啟與關閉分割檔案處理.....	69
選擇觀察值子集。.....	69
選取以條件表示式為基礎的觀察值子集.....	70
選取隨機樣本.....	71
選取時間範圍或觀察值範圍.....	71
未選擇之觀察值的處理方式.....	72
觀察值選取狀態.....	72
第 10 章範例檔案.....	73
注意事項.....	79
商標.....	80

第 1 章 簡介

本手冊將介紹許多功能的操作方式。本手冊的設計是以步驟式與實際操作的方式提供說明。範例中的所有檔案會連同應用程式一起安裝，因此您只要按照步驟執行分析，便可得到相同的結果。

若您想要多種統計量分析技巧的詳細範例，請嘗試逐步說明的案例研究（在「說明」功能表上提供）。

範例檔案

這裡的範例大部分使用資料檔 *demo.sav*。這個資料檔是一份虛擬的調查，樣本數為數千人，包含基本的人口統計資料和消費者資料。

如果您使用學生版，您的 *demo.sav* 版本是原始資料檔的代表性樣本，樣本數減少至 1,500 的觀察值限制。您使用該資料檔所分析的結果，會與此處顯示的結果不同。

與產品同時安裝的範例檔存放在安裝目錄的範例子目錄中。在下列每種語言的「範例」子目錄中存有個別資料夾：英文、法文、德文、義大利文、日文、韓文、波蘭文、俄文、簡體中文、西班牙文和繁體中文。

並非所有範例檔案皆提供各種語言。如果範例檔案沒提供您需要的語言，語言資料夾有英文版的範例檔案。

開啟資料檔

開啟資料檔：

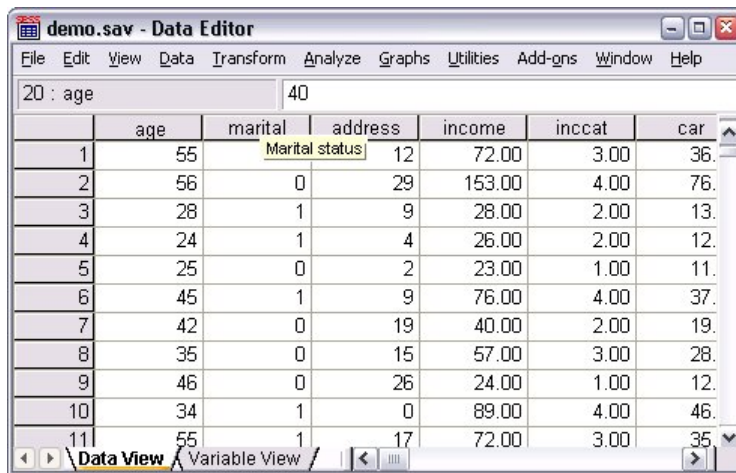
1. 在功能表上，選擇：

檔案 > 開啟 > 資料...

顯示一個開啟檔案的對話框。

依照預設，會顯示 IBM SPSS Statistics 資料檔（副檔名為 *sav*）。

本範例使用檔案 *demo.sav*。



	age	marital	address	income	inccat	car
1	55	Marital status	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

圖 1. 資料編輯器中的 *demo.sav* 檔案

資料檔會顯示在「資料編輯器」中。在「資料檢視」中，若變數定義了變數標籤，則當您將滑鼠游標放在變數名稱（直欄標題）上時，會顯示更詳細的變數標籤。

依預設，也會顯示實際資料數值。若要顯示標記：

2. 在功能表上，選擇：

檢視 > 值標籤



圖 2. 值標籤按鈕

或者，可使用工具列的「值標籤」按鈕。

	age	marital	address	income	inccat	car
1	55	Married	12	72.00	\$50 - \$74	36.
2	56	Unmarried	29	153.00	\$75+	76.
3	28	Married	9	28.00	\$25 - \$49	13.
4	24	Married	4	26.00	\$25 - \$49	12.
5	25	Unmarried	2	23.00	Under \$25	11.
6	45	Married	9	76.00	\$75+	37.
7	42	Unmarried	19	40.00	\$25 - \$49	19.
8	35	Unmarried	15	57.00	\$50 - \$74	28.
9	46	Unmarried	26	24.00	Under \$25	12.
10	34	Married	0	89.00	\$75+	46.
11	55	Married	17	72.00	\$50 - \$74	35.

圖 3. 資料編輯器中顯示的值標籤

會顯示描述性值標籤，讓您較易解讀回應值。

執行分析

若您有任何附加選項，「分析」功能表會包含報告和統計分析類別的清單。

我們從建立一個簡單的頻次表（個數表）開始。本範例需要 Statistics Base 選項。

1. 在功能表上，選擇：

分析 > 敘述性統計量 > 頻次...

會顯示「頻次」對話框。

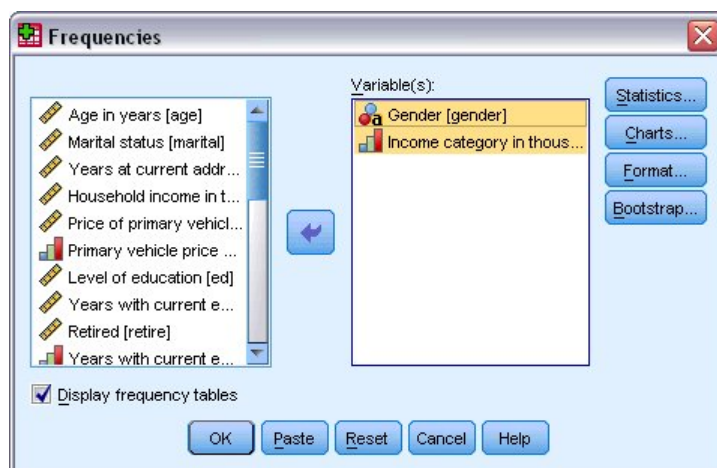













圖 4. 「頻次」對話框

各變數旁的圖示會提供關於資料類型與測量值水平的資訊。

	數值型(N)	字串	日期	時間(M)
尺度 (連續)		無		
序數(O)				
名義(M)				

如果清單中的變數標籤及/或名稱被截斷，則當游標置於它上方時，會顯示完整的標記/名稱。變數名稱 *inccat* 會接在描述性統計量變數標籤之後，顯示在方括號中。收入類別（千元） [*inccat*] 是變數標籤。如果沒有定義變數標籤，則在清單方框中僅會顯示變數名稱。

您可以按一下並拖曳外框或角落，便可像視窗一樣調整對話框的大小。例如，如果您將對話框變寬，變數清單也會變寬。

在對話框中，從左側的來源清單選取您想要分析的變數，再將它們拖放到在右側的「變數」清單中。**確定**按鈕的功能為執行分析；「變數」清單內至少要有一個變數，此按鈕才能啟用。

在許多對話框中，您可在清單中任一變數名稱上按一下右鍵，然後從蹦現功能表選取**變數資訊**。

2. 在來源變數清單中按一下「性別 [*gender*]，再將該變數拖放到目標「變數」清單中。
3. 按一下來源清單中的「收入類別（千元） [*inccat*]，再將其拖放到目標清單中。

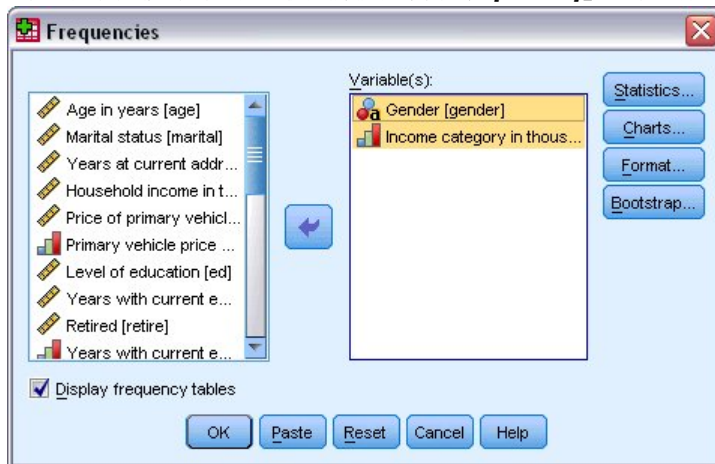


圖 5. 選取分析的變數

4. 按一下「**確定**」執行程序。

結果會顯示在「檢視器」視窗中。

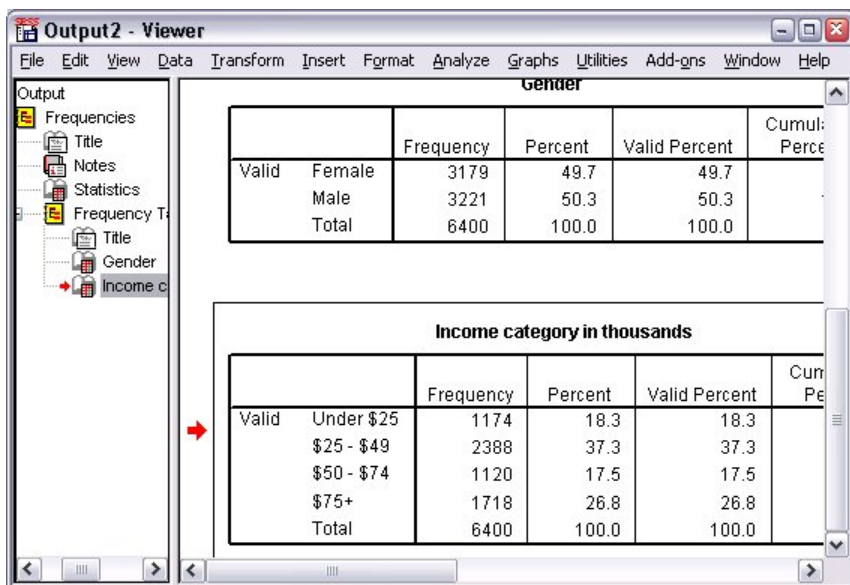


圖 6. 收入類別的頻次表

建立圖表

雖然有些統計程序可以建立圖表，您也可以使用「統計圖」功能表建立圖表。

例如，您可以建立一個圖表，顯示無線電話服務和擁有 PDA（個人數位助理）間的關係。

1. 在功能表上，選擇：

圖形 > 圖表建置器...

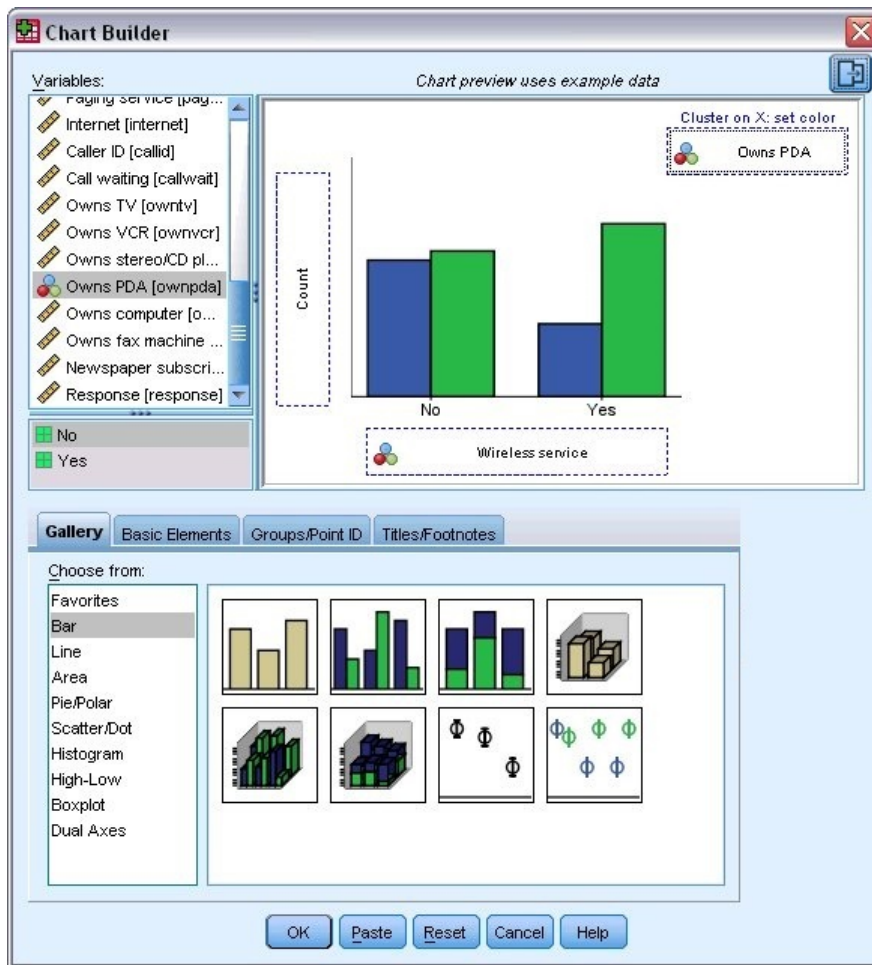


圖 7. 有完整放置區域的圖表建置器對話框

2. 若未選取的話，請按一下「展示區」標籤。
3. 若未選取的話，請按一下「長條圖」。
4. 將「集群長條」圖示拖曳至畫布，也就是「展示區」上的大型區域。
5. 向下捲動「變數」清單，在「無線服務 [wireless]」上按一下右鍵，並選擇「名義」作為其測量層級。
6. 將無線服務 [wireless] 變數拖曳至 x 軸。
7. 在「擁有 PDA [ownpda]」上按一下右鍵，並選擇「名義」作為其測量層級。
8. 將擁有 PDA [ownpda] 變數拖曳到畫布右上角的集群放置區域。
9. 按一下「確定」建立圖表。

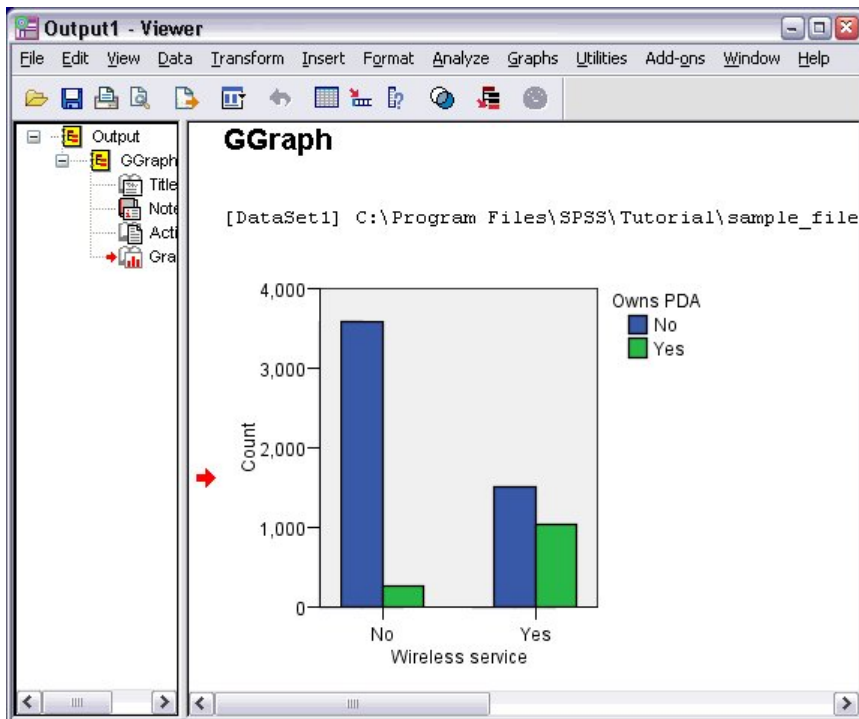


圖 8. 檢視器視窗中顯示的長條圖

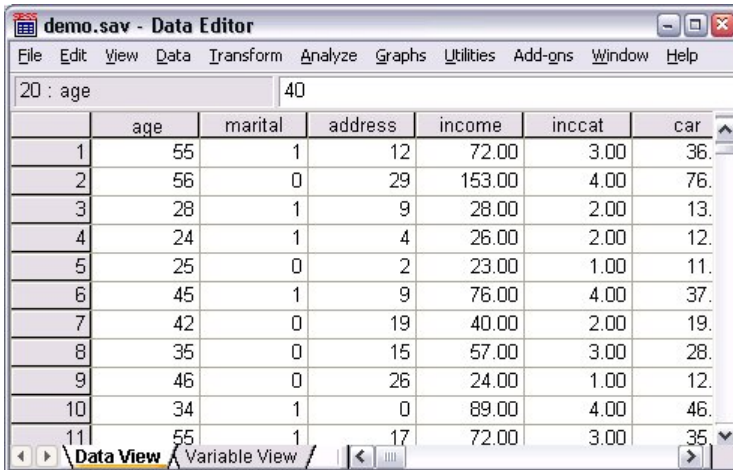
「檢視器」中會顯示此長條圖。此圖表顯示有無線電話服務的人擁有 PDA 的可能性遠超過沒有無線電話服務的人。

您可以在「檢視器」視窗目錄框中的圖表和表格上連按兩下，便可編輯圖表和表格，您也可以將結果複製並貼至其他的應用程式中。稍後會說明這些主題。

第 2 章 讀取資料

您可以直接輸入資料，或從多個不同的來源匯入資料。本章將討論以下處理程序，包括讀取儲存在 IBM SPSS Statistics 資料檔、試算表應用程式 (如 Microsoft Excel)、資料庫應用程式 (如 Microsoft Access) 以及文字檔中的資料。

IBM SPSS Statistics 資料檔的基本架構



	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

圖 9. 資料編輯器

IBM SPSS Statistics 資料檔是根據觀察值 (列) 和變數 (欄) 來分類。在此資料檔中，觀察值代表研究中的個別應答者，而變數則代表對研究中所詢問每個問題的回應。

讀取 IBM SPSS Statistics 資料檔

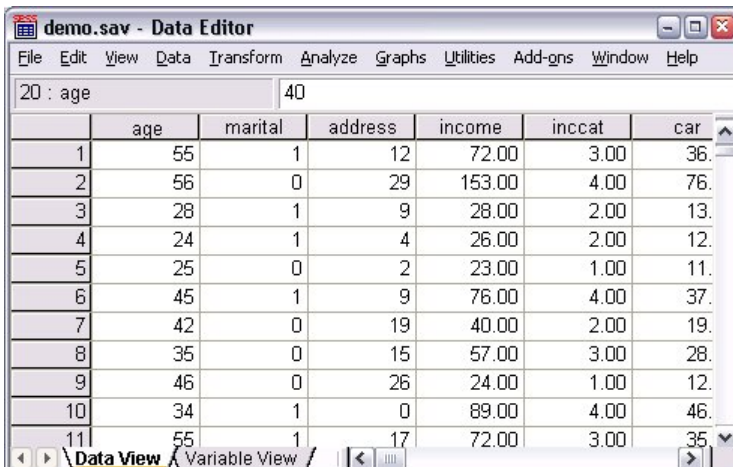
IBM SPSS Statistics 資料檔的副檔名為 *.sav*，其中包含您儲存的資料。

1. 在功能表上，選擇：

檔案 > 開啟 > 資料...

2. 瀏覽並開啟 *demo.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

現在資料會顯示在「資料編輯器」中。



	age	marital	address	income	inccat	car
1	55	1	12	72.00	3.00	36.
2	56	0	29	153.00	4.00	76.
3	28	1	9	28.00	2.00	13.
4	24	1	4	26.00	2.00	12.
5	25	0	2	23.00	1.00	11.
6	45	1	9	76.00	4.00	37.
7	42	0	19	40.00	2.00	19.
8	35	0	15	57.00	3.00	28.
9	46	0	26	24.00	1.00	12.
10	34	1	0	89.00	4.00	46.
11	55	1	17	72.00	3.00	35.

圖 10. 已開啟的資料檔

讀取 Excel 資料

您可以從應用程式（如 Microsoft Excel）讀取資料，而不用直接將所有資料輸入「資料編輯器」中。您也可以讀取直欄標題作為變數名稱。

1. 在功能表上，選擇：

檔案 > 匯入資料 > Excel

2. 跳至 Samples\English 資料夾，並選取 demo.xlsx。

「讀取 Excel 檔案」對話框顯示資料檔的預覽。顯示檔案中第一張樣式表的內容。如果有多張樣式表，則可以從清單選取樣式表。

您可以看到性別的部分字串值有前導空格。MaritalStatus 的部分值顯示為逗點 (.)。

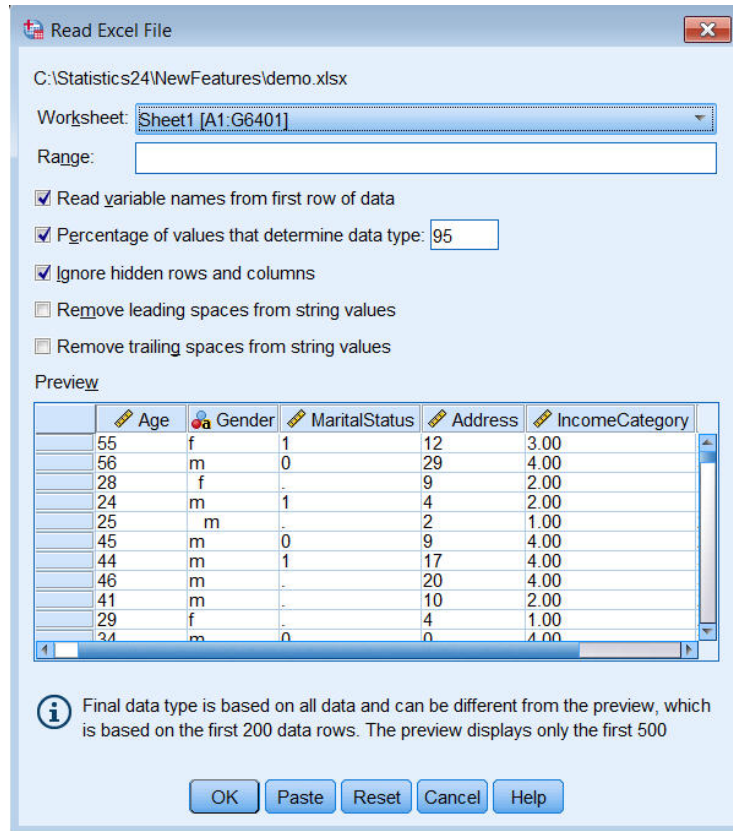
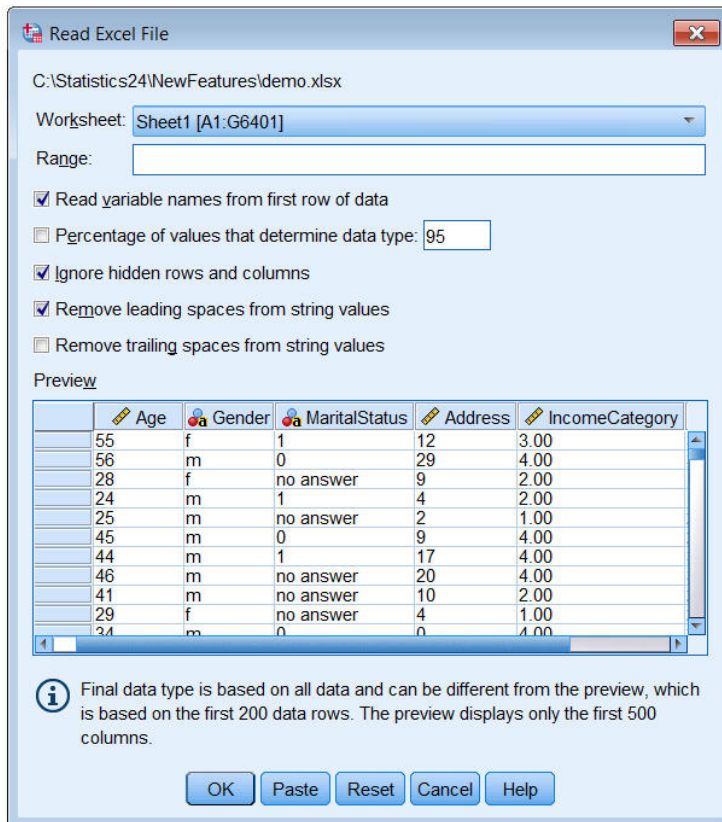


圖 11. 「讀取 Excel 檔案」對話框

3. 確定已選取從資料第一列資料讀取變數名稱。如果直欄標題不符合變數命名規則，它們會轉換成有效的變數名稱。而原始的直欄標題則存成變數標籤。
4. 選取移除字串值的前導空格。
5. 取消選取判定資料類型的值的百分比。



現在，系統遺失的資料格中顯示字串值「無回答」。如果沒有值參數的百分比，且直欄包含混合資料類型，則變數讀作字串資料類型。保留所有值，但是將數值作為字串值來處理。

6. 選取（勾選）判定資料類型的值百分比，以將 *MaritalStatus* 看作數值變數。
7. 按一下「確定」以讀取 Excel 檔案。

現在資料會顯示在「資料編輯器」中，並用直欄標題作為變數名稱。由於變數名稱不能包含空格，故會將原始欄標題中的空格移除。例如，直欄標題 "*Marital Status*" 會轉換為變數 *MaritalStatus*。原來的欄位標題保留為變數標籤。

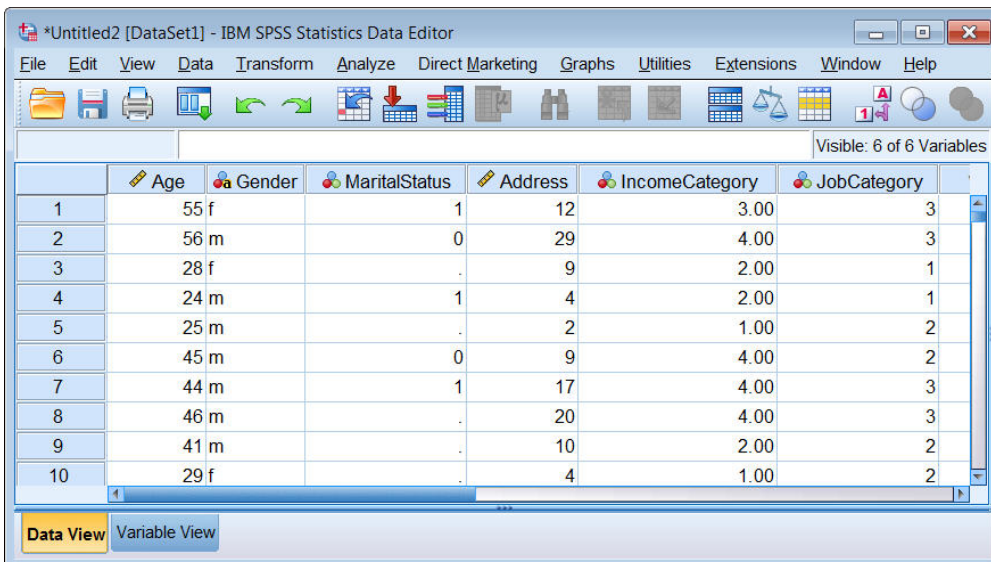


圖 12. 匯入的 Excel 資料

相關資訊

第 73 頁的『範例檔案』

從資料庫讀取資料

使用「資料庫精靈」輕鬆從資料庫來源匯入資料。在安裝 ODBC（開放式資料庫連結）驅動程式後，便可以直接讀取使用此驅動程式的任何資料庫。安裝光碟提供許多資料庫格式所用的 ODBC 驅動程式。其他驅動程式則可以向協力廠商索取。此範例將討論最常用的資料庫應用程式之一，Microsoft Access。

附註：本範例僅適用 Microsoft Windows，且需要 Access 的 ODBC 驅動程式才能使用。Microsoft Access ODBC 驅動程式僅使用 IBM SPSS Statistics 32 位元版本。範例中的步驟與其他平台類似，但需要協力廠商的 Access ODBC 驅動程式才能使用。

1. 在功能表上，選擇：

檔案 > 匯入資料 > 資料庫 > 新建查詢...

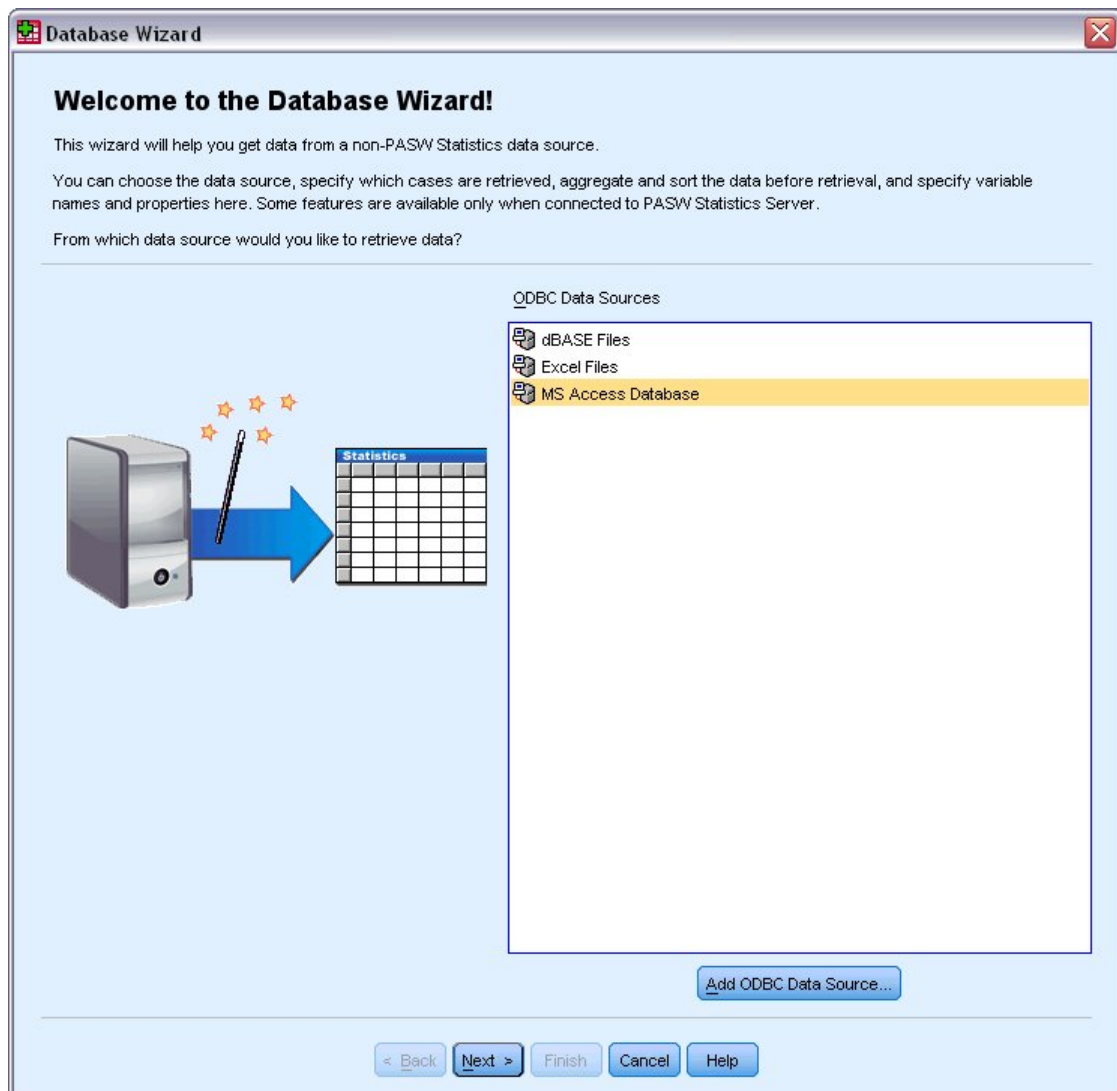


圖 13. 「歡迎使用資料庫精靈」對話框

2. 從可用資料來源清單中選取「MS Access 資料庫」，並按一下「下一步」。

附註：根據您的安裝，您可能也會在精靈（僅適用於 Windows 作業系統）的左側看到 OLEDB 資料來源清單，但此範例使用顯示在右側的 ODBC 資料來源清單。

3. 按一下「瀏覽」以瀏覽至您要開啟的 Access 資料庫檔案。
4. 開啟 *demo.mdb*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。
5. 按一下登入對話框中的「確定」。

在接下來的步驟中，您可以指定要匯入的表格和變數。

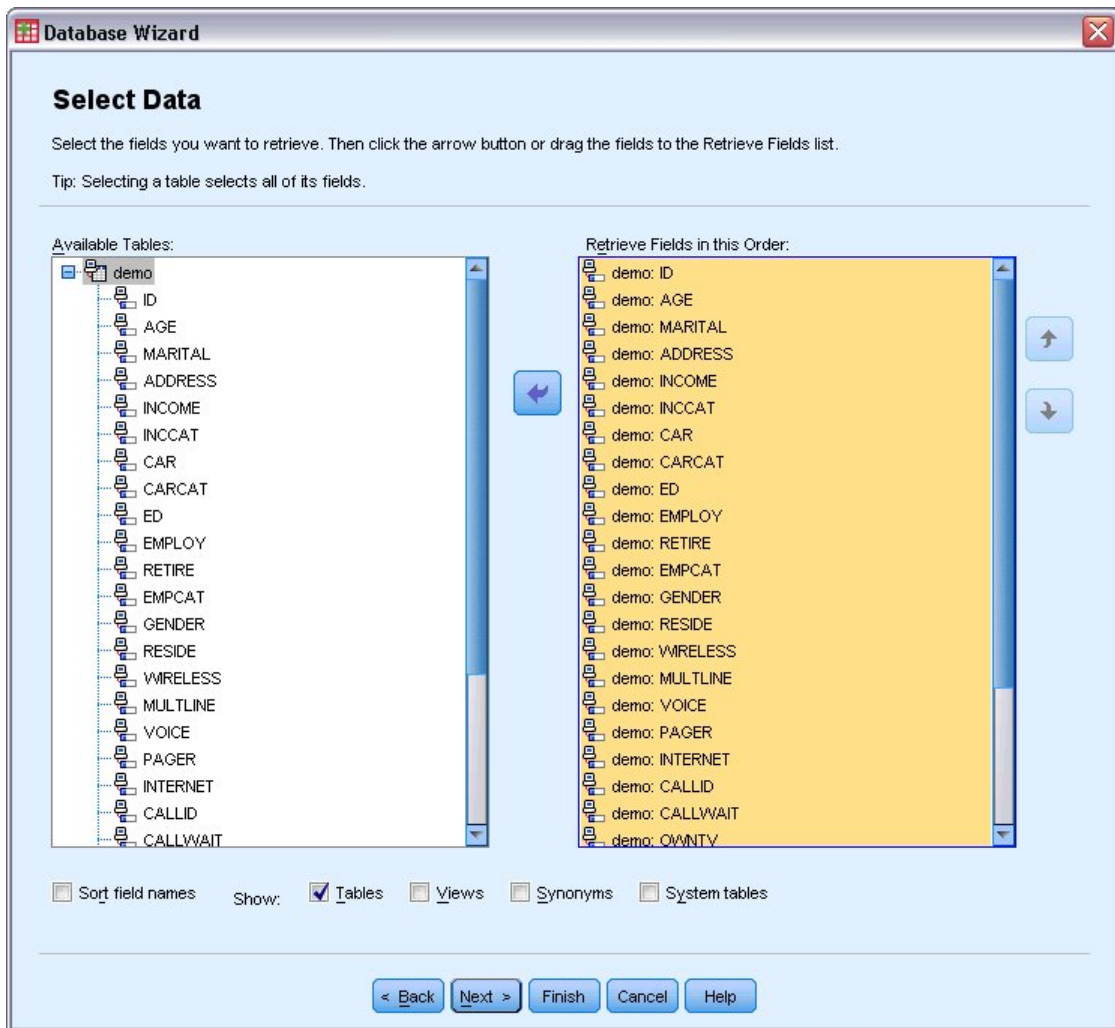


圖 14. 「選擇資料」步驟

- 將整個示範表格拖曳到「按此次序抽取欄位」清單中。
- 按一下「**下一步**」。

在接下來的步驟中，可以選取要匯入哪些記錄（觀察值）。

如果您不要匯入所有觀察值，則您可以匯入觀察值子集（如 30 歲以上的男性），或從資料來源匯入隨機的觀察值樣本。如果是大型的資料來源，則您可以將觀察值個數限制為小型的代表樣本，以減少處理的時間。

- 按一下「**下一步**」以繼續進行。

使用欄位名稱來建立變數名稱。必要時，該名稱會轉換成有效的變數名稱。保留原始的欄位名稱作為變數標籤。您也可以匯入資料庫之前變更變數名稱。

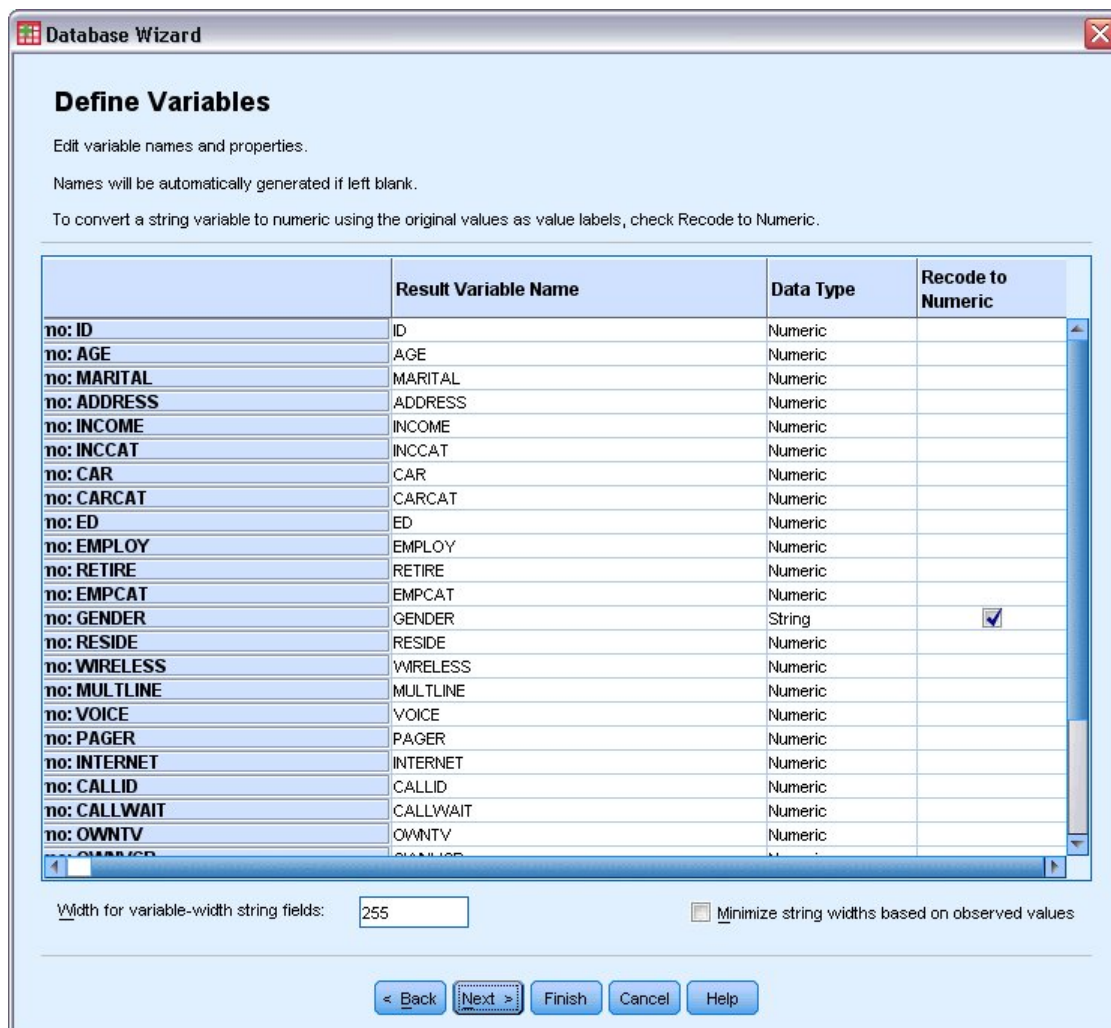


圖 15. 「定義變數」步驟

9. 按一下「性別」欄位中的**重新編碼為數值**資格。此選項會將字串變數轉換成整數變數，並保留原始值作為新變數的值標籤。
10. 按一下「**下一步**」以繼續進行。

「結果」步驟中會顯示您在「資料庫精靈」中選擇建立的 SQL 陳述式。您可以立即執行此陳述式，或儲存到檔案中於日後使用。

11. 按一下「**完成**」以匯入資料。

現在「資料編輯器」中可以使用您選取匯入之 Access 資料庫中的所有資料。

從文字檔讀取資料

文字檔代表另一個一般資料來源。許多試算表程式和資料庫可以將其內容存成多種文字檔格式。以逗點或 Tab 鍵分隔的檔案，參考使用逗點或 Tab 鍵的資料列來表示每個變數。在此範例中，資料由 Tab 鍵區隔。

1. 在功能表上，選擇：

檔案 > 匯入資料 > 文字資料

2. 跳至 Samples\English 資料夾，並選取 demo.txt。

「文字匯入」精靈會引導您完成如何解譯指定文字檔的定義程序。

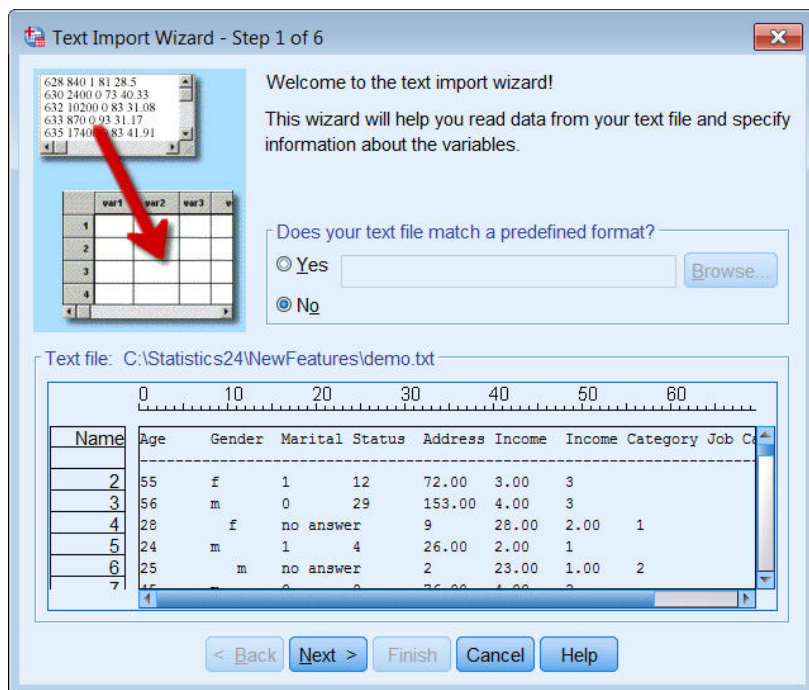


圖 16. 文字匯入精靈：第 1 步（共 6 步）

3. 在步驟 1 中，您可以在精靈中選擇預先定義的格式，或者建立新的格式。選取否。

4. 按一下「下一步」以繼續進行。

如稍早所言，此檔案使用以 Tab 鍵分隔的格式。並且此檔案的第一行也已定義變數名稱。

5. 在精靈的步驟 2 中，選取**分隔**指示檔案使用分隔格式化結構。

6. 選取**是**指示檔案頂端包含變數名稱。

7. 按一下「下一步」以繼續進行。

8. 在步驟 3 中，輸入 2 表示第一個資料觀察值的起始行號（因為變數名稱位於第一行）。

9. 保留此步驟其餘的預設值，並按一下「下一步」以繼續進行。

步驟 4 中的「資料」預覽為您提供確保正確讀取檔案的便捷方式。

10. 選取「**標籤**」，然後取消選取其他定界字元。依預設，選取**空格**，因為檔案包含空格。對於此檔案，空格是資料值的一部分，而不是定界字元。您需要取消選取**空格**來正確讀取檔案。

11. 選取**移除字串值的前導空格**。字串值開頭的空格會影響表示式中字串值的求值方法。在此檔案中，性別的部分值含有非值部分的前導空格。如果您不移除這些空格，則值 " f" 會被看作與值 "f" 不同。

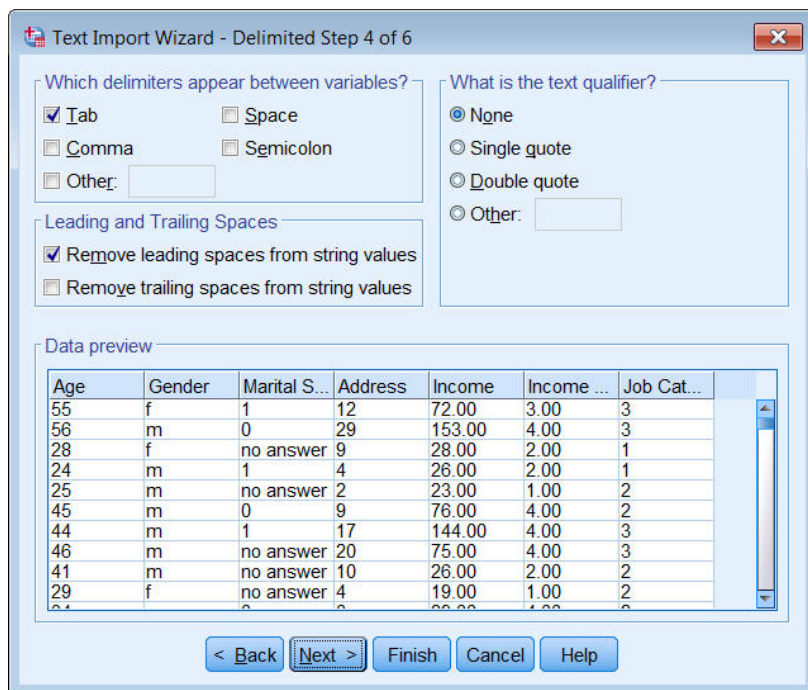


圖 17. 文字匯入精靈：第 4 步（共 6 步）

12. 按一下「**下一步**」以繼續進行。

因為變數名稱已修改為符合命名規則，第 5 步允許您編輯任何不需要的名稱。您也可以在這裡定義資料類型。例如，您可以將 *Income* 變更為元貨幣格式。若要變更資料類型：

13. 在**資料預覽**中，選取 *Income*。
 14. 從「資料格式」下拉清單中選取「**金額**」。

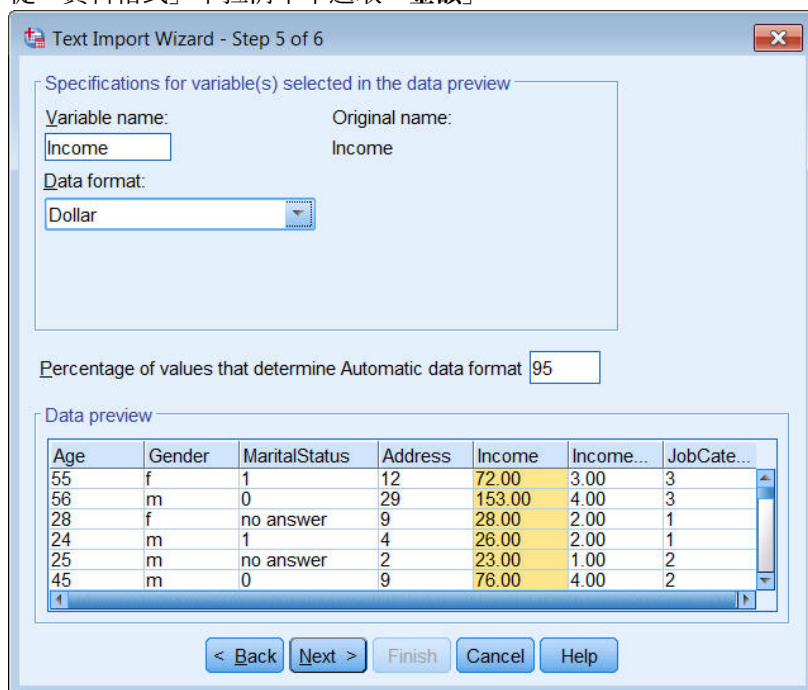


圖 18. 變更資料類型

變數 *MaritalStatus* 同時包含字串和數值。小於 5% 的值為字串。根據判定自動資料格式的值百分比的預設值 95%，變數將被看作的數值，而字串值會設定為系統遺漏值。如果沒有任何資料格式符合百分比

值，則變數會被看作字串變數。如果將此設定變更為 100，則保留所有值，但所有數值都會被看作字串。

15. 按一下「**下一步**」以繼續進行。
16. 在最後一個步驟中保留預設選擇，並按一下「**完成**」匯入資料。

第 3 章 使用資料編輯器

「資料編輯器」會顯示作用中資料檔的內容。「資料編輯器」中的資訊包含變數和觀察值。

- 在「資料視圖」中，欄代表變數，列代表觀察值。
- 在「變數視圖」中，每一列為一變數，每一欄為與該變數相關的屬性。

變數用來代表不同類型的編譯資料。一種常見的類比便是調查。調查內每一個問題的答案都等於一個變數。變數有許多不同類型，包括數字、字串、貨幣和日期。

輸入數字資料

可在「資料編輯器」中輸入資料，這對於小的資料檔或是要在較大資料檔中進行小規模修改時非常實用。

1. 按一下「資料編輯器」視窗中的「變數視圖」標籤。

您必須定義要使用的變數。在此情況下，只需要三個變數：年齡、婚姻狀態和收入。

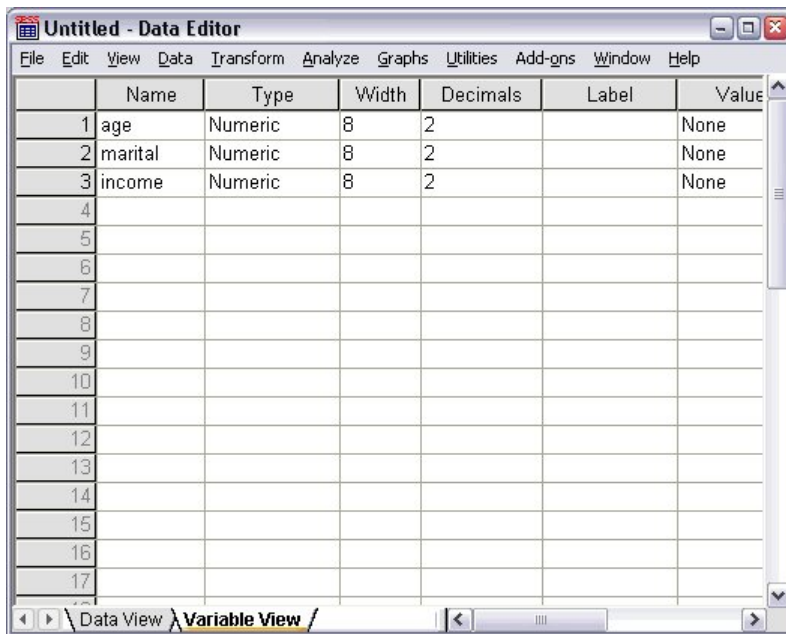


圖 19. 「變數視圖」中的變數名稱

2. 在第一欄的第一列中，輸入年齡。
3. 在第二列中，輸入婚姻。
4. 在第三列中，輸入收入。

新變數會自動被指定為數值資料類型。

如果您不輸入變數名稱，則系統會自動建立唯一名稱。然而，這些名稱並不具描述性，因此不建議使用在大的資料檔上。

5. 按一下「資料視圖」標籤繼續輸入資料。

您在「變數視圖」輸入的名稱現在會成為「資料視圖」前三個欄的標題。

於第一欄起始的第一列內，開始輸入資料。

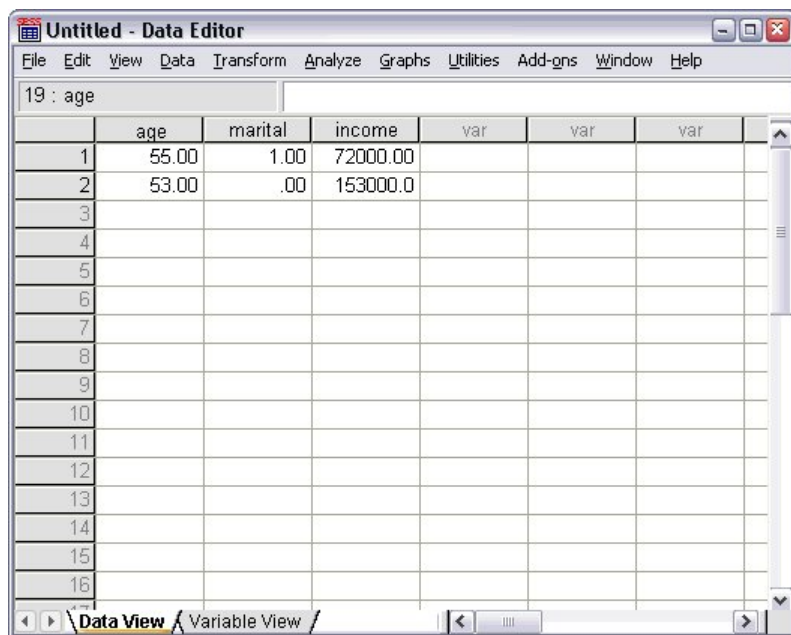


圖 20. 在「資料視圖」中輸入的數值

6. 在年齡欄中，輸入 55。
7. 在婚姻欄中，輸入 1。
8. 在收入欄中，輸入 72000。
9. 將游標移到第一欄第二列的位置，新增下一筆主題資料。
10. 在年齡欄中，輸入 53。
11. 在婚姻欄中，輸入 0。
12. 在收入欄中，輸入 153000。

目前雖然年齡和婚姻欄中的數值應為整數，但還是會顯示小數點。若要隱藏這些變數的小數點：

13. 按一下「資料編輯器」視窗中的「變數視圖」標籤。
14. 在年齡列的小數點欄中，輸入 0 以隱藏小數點。
15. 在婚姻列的小數點欄中，輸入 0 以隱藏小數點。

輸入字串資料

文字字串等非數值資料，也可在「資料編輯器」中輸入。

1. 按一下「資料編輯器」視窗中的「變數視圖」標籤。
2. 在第一個空白列的第一個資料格中，輸入 sex 作為變數名稱。
3. 按一下輸入項旁的類型資料格。
4. 按一下類型資料格右側的按鈕，開啟「變數類型」對話框。
5. 選擇字串指定變數類型。
6. 按一下確定儲存您的選擇，並回到「資料編輯器」。

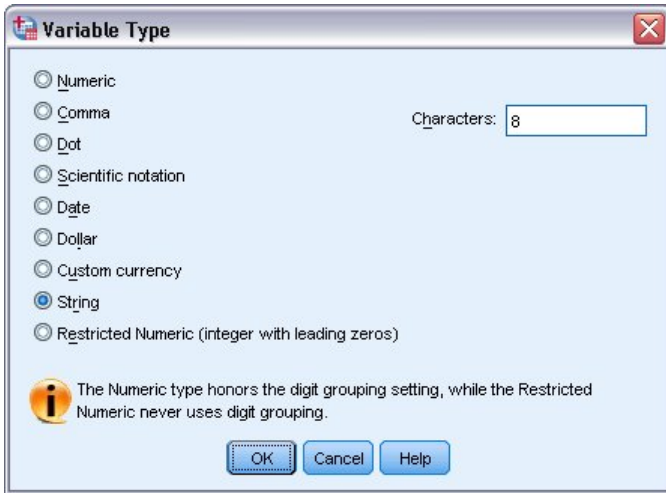


圖 21. 「變數類型」對話框

定義資料

除了定義資料類型外，您也可以為變數名稱和資料值定義描述性的變數標籤和值標籤。統計報表和圖表會用到這些描述性的標記。

新增變數標籤

標記用來提供變數的說明。這些描述通常可視為變數名稱的加長版。標記最多可有 255 個字元。在您的輸出使用這些標記來識別不同的變數。

1. 按一下「資料編輯器」視窗中的「變數視圖」標籤。
2. 在年齡列的標記欄中，輸入 Respondent's Age。
3. 在婚姻列的標記欄中，輸入 Marital Status。
4. 在收入列的標記欄中，輸入 Household Income。
5. 在性別列的標記欄中，輸入 Gender。

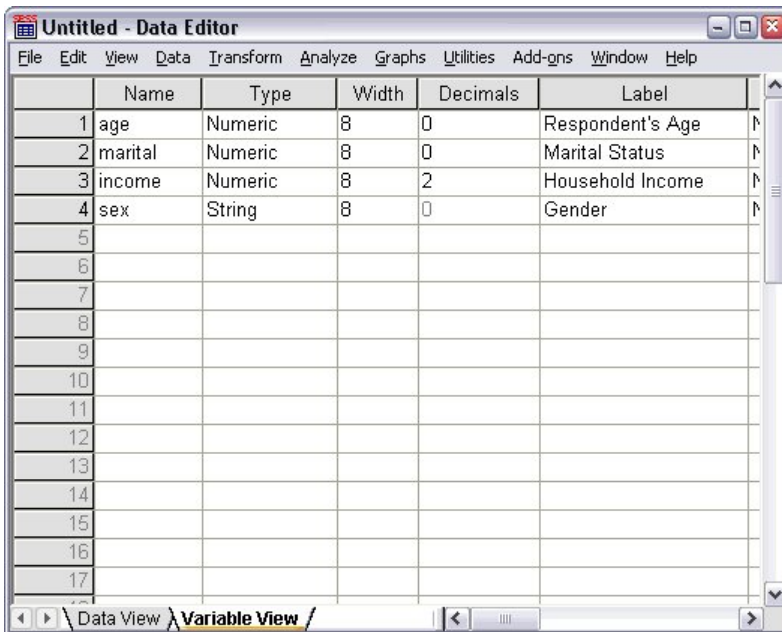


圖 22. 在「變數視圖」中輸入的變數標籤

變更變數類型和格式

類型欄顯示每一個變數目前的資料類型。最常見的資料類型是數值和字串，但亦支援許多其他的格式。在目前資料檔案中，已定義收入變數為數值類型。

1. 按一下「收入」列的「類型」資料格，然後按一下資料格右側的按鈕開啟「變數類型」對話框。
2. 選擇「貨幣」。

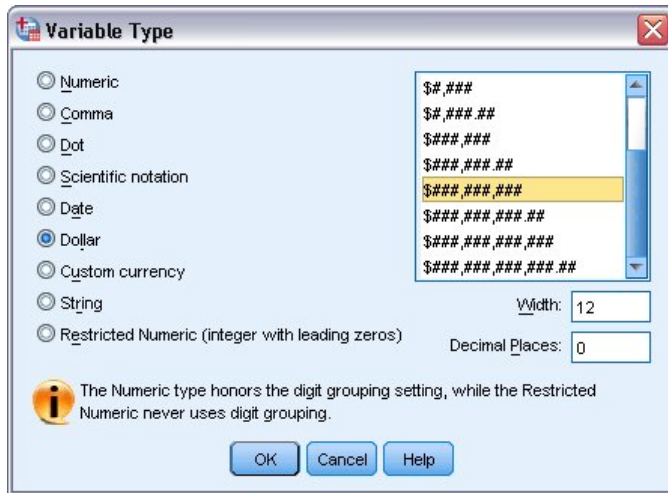


圖 23. 「變數類型」對話框

顯示目前所選資料類型的格式選項。

3. 對於此範例的貨幣格式，選擇「**\$###,###,###**」。
4. 按一下**確定**儲存您的變更。

新增值標籤

值標籤可讓您將變數數值對應至字串標記。在本範例中，婚姻變數有兩個可接受的數值。數值 0 表示受試者為單身，數值 1 表示受試者已婚。

1. 按一下「婚姻」列的「數值」資料格，然後按一下資料格右側的按鈕開啟「變數標籤」對話框。

數值是真正的數字。

值標籤是套用於特定的數值的字串標記。

2. 在「數值」欄位中輸入 0。
3. 在「標記」欄位中輸入單身。
4. 按一下**新增**，將此標記新增至清單中。

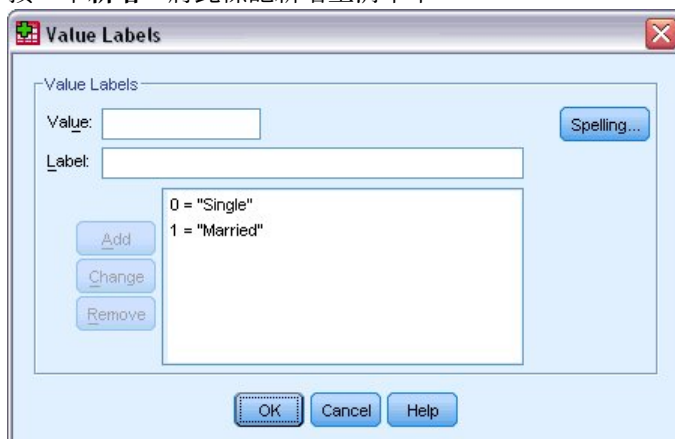


圖 24. 「值標籤」對話框

5. 在「數值」欄位中輸入 1，然後在「標記」欄位中輸入已婚。
6. 按一下**新增**，再按一下**確定**儲存您的變更，回到「資料編輯器」。

這些標記也會在「資料視圖」中顯示，讓資料更具可讀性。

7. 按一下「資料編輯器」視窗下的「**資料視圖**」標籤。
8. 在功能表上，選擇：

檢視 > 值標籤

在「資料編輯器」中輸入數值後，清單中會顯示標記。此步驟有助於建議有效的回應值，並提供更具描述性的答案。

如果「值標籤」功能表項目已是作用中 (在旁有勾號)，則再次選取「**值標籤**」會將值標籤顯示關閉。

處理遺漏的資料

遺漏或無效資料太常見而無法忽略。調查的應答者可能拒絕回答特定的問題、不知道答案或是以預料之外的格式回答問題。如果您不過濾或識別出這些資料，您的分析結果可能會不正確。

對於數值資料，會將空白資料欄位或是包含無效輸入的欄位，轉換成以單句點表示的系統遺漏。

發生遺漏值的原因對您的分析可能很重要。例如，您可能會發現區別拒絕作答的應答者，與因問題設計不當而無法作答的應答者，是很有用的。

數字變數的遺漏值

1. 按一下「資料編輯器」視窗中的「**變數視圖**」標籤。
2. 按一下年齡列的遺漏資料格，然後按一下資料格右側的按鈕來開啟「遺漏值」對話框。

在此對話框中，您最多可以指定三個不同的遺漏值，或者您可以指定一個數值範圍外加一個額外的離散值。



圖 25. 「遺漏值」對話框

3. 選擇**離散遺漏值**。
 4. 在第一個文字方框中輸入 999，另外兩個文字方框則保持空白。
 5. 按一下**確定**儲存您的變更，並回到「資料編輯器」。
- 既然已新增遺漏資料值，那麼可將標記套用至該值。
6. 按一下年齡列的數值資料格，然後按一下資料格右側的按鈕來開啟「**值標籤**」對話框。
 7. 在「值」欄位中鍵入 999。
 8. 在「標記」欄位中輸入 未作答。
 9. 按一下**新增**，將此標記新增至資料檔中。
 10. 按一下**確定**儲存您的變更，並回到「資料編輯器」。

字串變數的遺漏值

處理字串變數遺漏值的方式，與處理數值變數遺漏值的方式類似。但是與數值變數不同的是，字串變數的空白欄位並不指定為系統遺漏值。而是解譯為空白字串。

1. 按一下「資料編輯器」視窗中的「變數視圖」標籤。
2. 按一下「性別」列的「遺漏」資料格，然後按一下資料格右側的按鈕來開啟「遺漏值」對話框。
3. 選擇**離散遺漏值**。
4. 在第一個文字方框中，輸入 NR。

字串的遺漏值是區分大小寫的。因此，*nr* 值不會視為遺漏值。

5. 按一下**確定**儲存您的變更，並回到「資料編輯器」。
- 現在您可以新增遺漏值的標記。
6. 按一下「性別」列的「數值」資料格，然後按一下資料格右側的按鈕來開啟「值標籤」對話框。
 7. 在「數值」欄位中輸入 NR。
 8. 在「標記」欄位中輸入 未作答。
 9. 按一下**新增**，將此標記新增至您的專案中。
 10. 按一下**確定**儲存您的變更，並回到「資料編輯器」。

第 4 章 檢查個別變數的摘要統計量

本小節討論簡單的摘要量數，以及變數的測量層級如何影響應使用的統計量類型。我們將使用資料檔 *demo.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

測量層級

不同的摘要量數適用於不同類型的資料，這需依照測量的層級而定：

類別。 限定數值或類別的資料（例如，性別或婚姻狀態）。也稱為**定性資料**。類別變數可以是字串（文數字）資料或數值變數，這類變數使用數字編碼來代表類別（例如，0 = 未婚而 1 = 已婚）。類別變數有兩種基本類型：

- **名義。** 未繼承類別順序的類別資料。例如，銷售工作類別並不高於或低於行銷或研究工作類別。
- **序數。** 類別順序具某種意義的類別資料，但各類別間不具任何可測量的距離。例如，值高、中、低存在在一種順序，但這些值之間的「距離」是無法計算的。

尺度。 資料以區間或比例尺度測量，其資料值同時表示數值順序及數值之間的距離。例如，薪資 \$72,195 比 \$52,398 高，而兩數值之間的距離為 \$19,797。同時這也稱為**數值**或**連續資料**。

類別資料的摘要量數

類別資料最典型的摘要量數，為各類別中觀察值的個數或百分比。**眾數**是指觀察值個數最多的類別。如果類別的數量很大，則對序數資料而言，**中位數**（在此值以上和以下的觀察值各半）可能也是一項有用的摘要量數。

「頻次」程序會產生頻次，表格中會同時顯示觀察值的個數和百分比，而這些觀察值代表所觀察到的每個變數值。

1. 在功能表上，選擇：

分析 > 敘述性統計量 > 頻次...

附註：本功能需要 Statistics Base 選項。

2. 選取「擁有 PDA [ownpda]」和「擁有電視機 [owntv]」，然後將它們移至「變數」清單中。

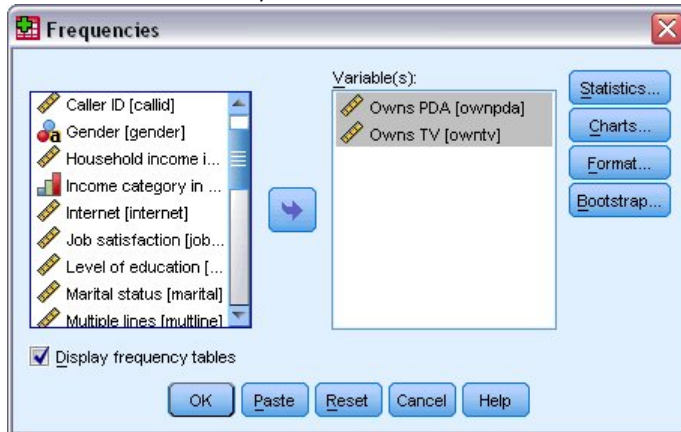


圖 26. 選取類別變數進行分析

3. 按一下「確定」執执行程序。

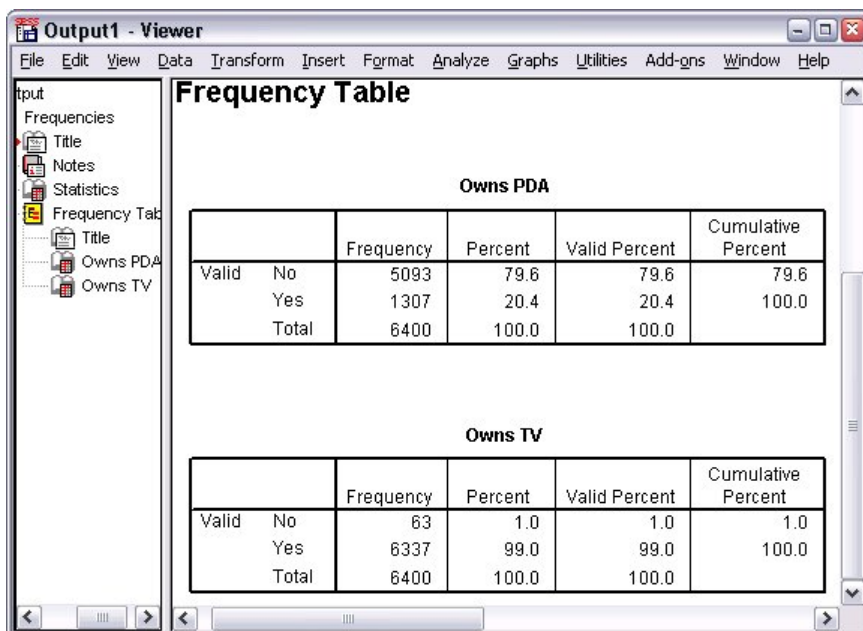


圖 27. 頻次

「檢視器」視窗中顯示頻次。頻次中顯示，只有 20.4% 的人有 PDA，但幾乎每個人都有電視機 (99.0%)。這可能不是一項有趣的發現，但如果我們能多瞭解一下這一小部分沒有電視的人或許會更有趣。

類別資料的圖表

您可以使用長條圖或圓餅圖，以圖形的方式來顯示頻次中的資訊。

1. 再次開啟「頻次」對話框。（這兩個變數仍然應保持選定）。

您可以使用工具列上的「叫回對話」按鈕，快速回到最近使用的程序。



圖 28. 「叫回對話」按鈕

2. 按一下「圖表」。
3. 選取「長條圖」，然後按一下「繼續」。
4. 在主對話框中按一下「確定」以執行程序。

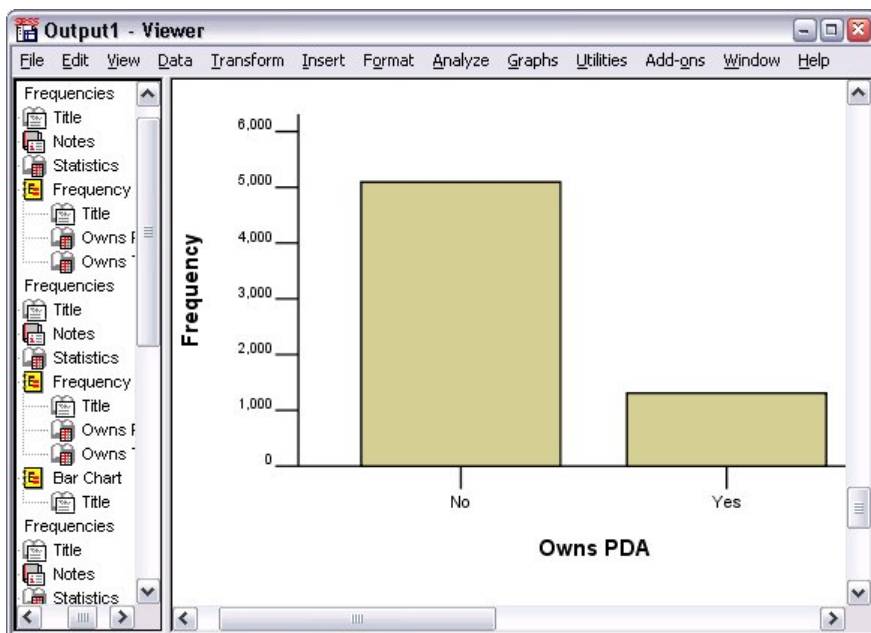


圖 29. 長條圖

除了頻次外，現在以長條圖的形式顯示相同資訊，更容易看出大部分的人並沒有 PDA，但幾乎每個人都有電視機。

尺度變數的摘要量數

尺度變數可以使用多種摘要量數，包括：

- **集中趨勢量數**。最常見的集中趨勢量數為**平均數**（算數平均數）和**中位數**（在此值以上和以下的觀察值各半）。
- **離散量數**。此為測量資料中變異或分散程度的統計量，包括標準差、最小值和最大值。

1. 再次開啟「頻次」對話框。
2. 按一下**重設**，清除之前的任何設定。
3. 選取「家庭收入（千元） [income]」，並將之移到「變數」清單中。
4. 按一下「**統計量**」。
5. 選取「**平均數**」、「**中位數**」、「**標準差**」、「**最小值**」和「**最大值**」。
6. 按一下「**繼續**」。
7. 在主對話框中取消選取「**顯示頻次**」。（對尺度變數而言，頻次通常不是非常有用，因為頻次中不同數值的數目，幾乎和資料檔中的觀察值個數一樣多。）
8. 按一下「**確定**」執执行程序。

「檢視器」視窗中顯示「頻次統計量」表。

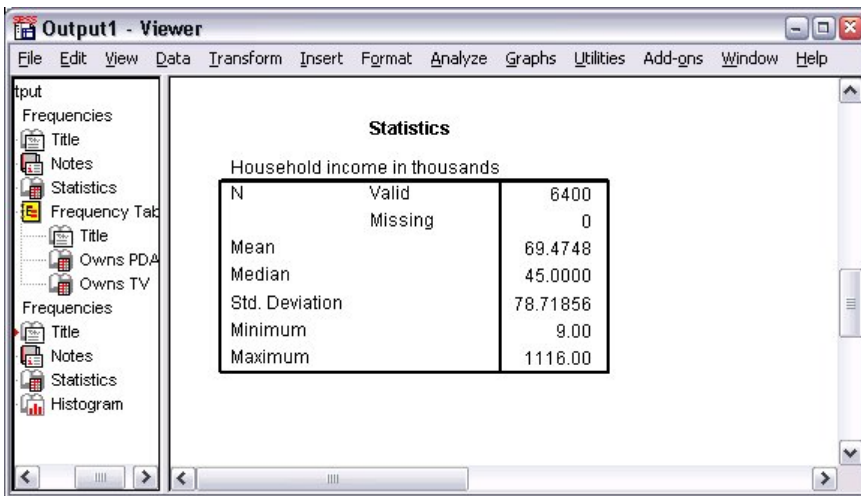


圖 30. 「頻次統計量」表

在此範例中，平均數與中位數有很大的差異。平均數幾乎比中位數大 25,000，表示數值並非常態分佈。您可在視覺上用直方圖來檢查分配狀態。

尺度變數的直方圖

1. 再次開啟「頻次」對話框。
2. 按一下「圖表」。
3. 選取「直方圖」和「有常態曲線」。
4. 按一下「繼續」，再按一下主對話框中的「確定」來執程序。

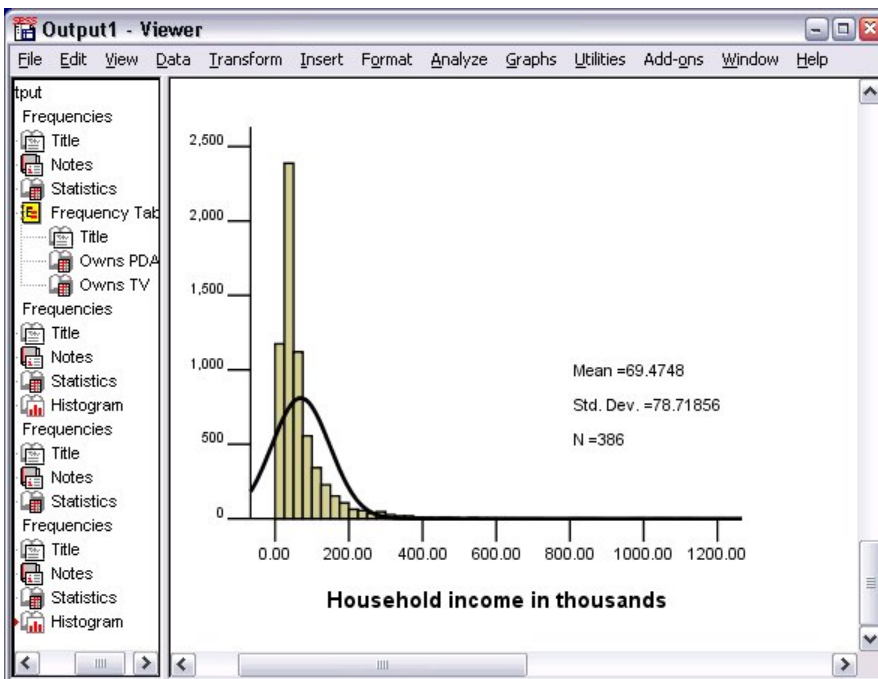


圖 31. 直方圖

大部分觀察值都聚集在尺度下方的末端，而其中大多落在 100,000 以下，但有少部分觀察值位於 500,000 和以上的範圍（數量太少，不修改直方圖幾乎看不到）。這些數值較高的少部分觀察值，對平均數有極大的影響，但對中位數的影響很小或甚至沒影響，因此在這個範例中，中位數是較佳的集中趨勢指標。

第 5 章 建立並編輯圖表

您可以建立並編輯多種類型的圖表。在本章，我們會建立並編輯長條圖。您可以將原則套用至任何圖表類型。

圖表建立基礎

為了示範建立圖表的基礎，我們將建立不同工作滿意度層級的平均收入長條圖。本範例使用資料檔 *demo.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

1. 在功能表上，選擇：

圖形 > 圖表建置器...

「圖表建置器」對話框為互動式視窗，可讓您預覽您所建立的圖表外觀。

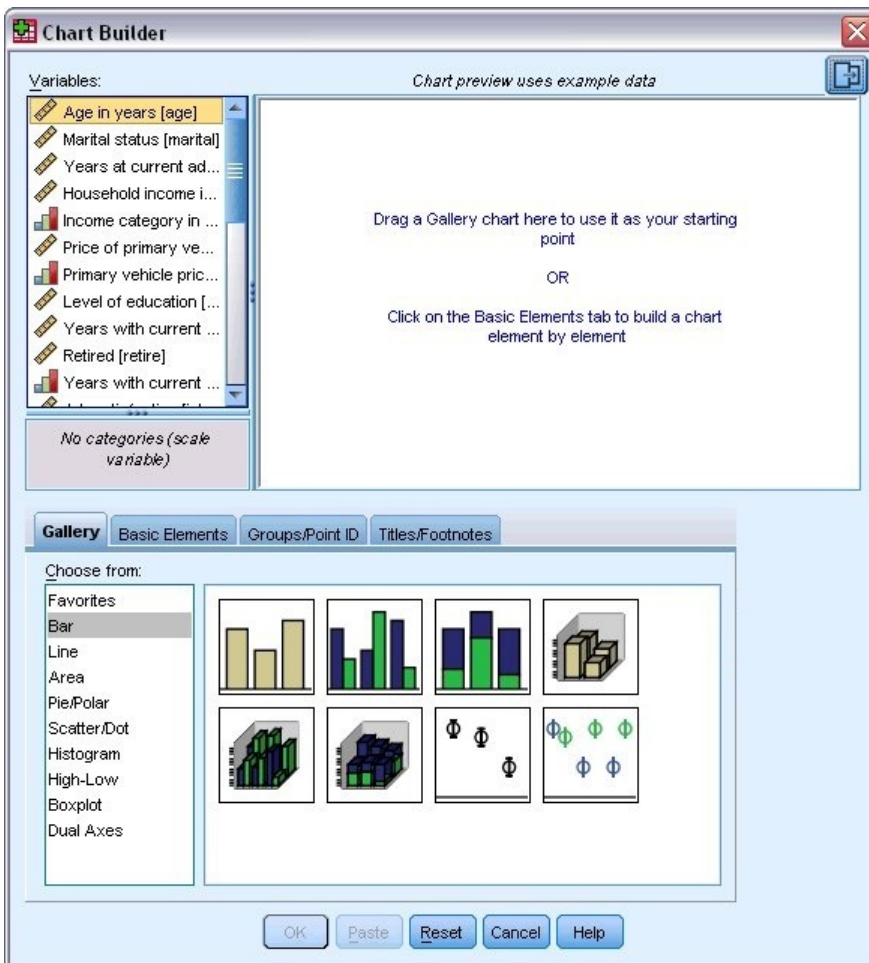


圖 32. 「圖表建置器」對話框

使用圖表建置器展示區

1. 如果尚未選取，請按一下「展示區」標籤。

「展示區」包含預先定義的許多不同圖表，根據圖表類型編排。「基本元素」標籤也提供一些基本元素（如軸與圖形元素）可用於從頭開始建立圖表，但使用「展示區」會較為容易。

2. 如果尚未選取，請按一下「長條圖」。

代表「展示區」中可用長條圖的圖示會出現在對話框中。圖片應提供足夠資訊以識別特定圖表類型。若您需要更多資訊，亦可將滑鼠游標停留在圖示上，以顯示圖表的「工具提示」說明。

3. 將簡單長條圖的圖示拖曳至「畫布」，也就是展示區上的大型區域。「圖表建置器」會在畫布顯示圖表預覽。請注意，用於繪製圖表的資料並非您的實際資料。這只是範例資料。

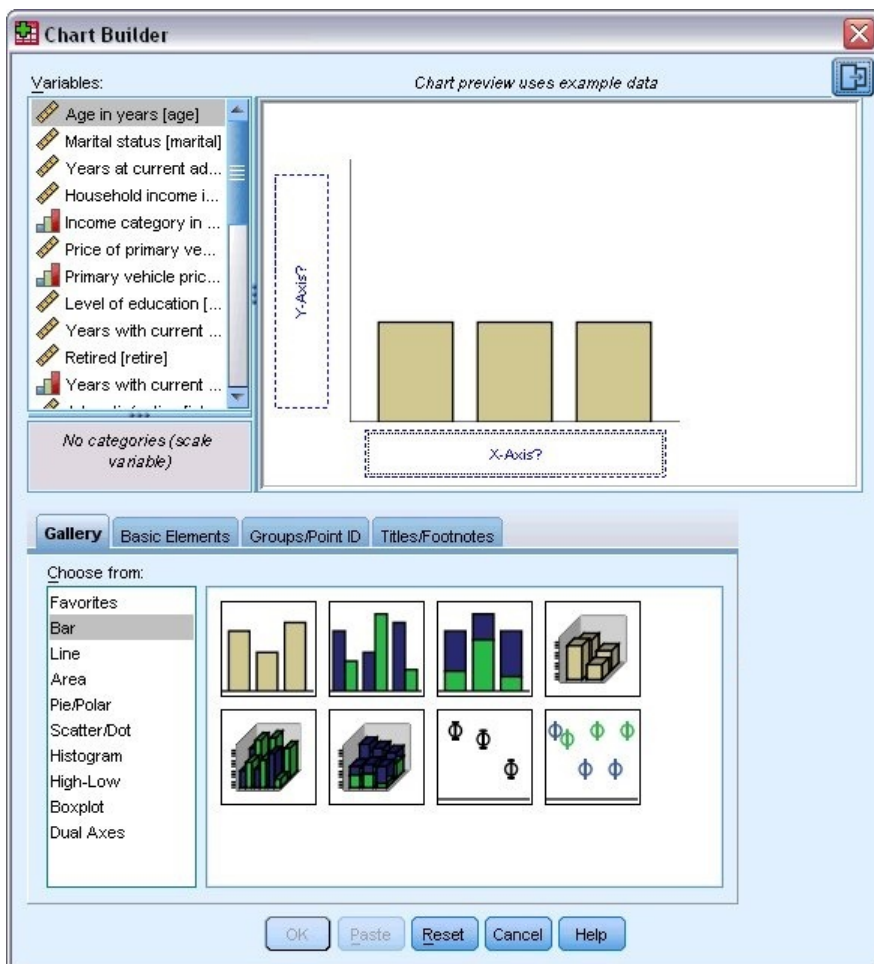


圖 33. 「圖表建置器」畫布上的長條圖

定義變數與統計量

雖然畫布上有圖表，但仍未完成，因為沒有變數或統計量來控制長條的高度，並指定各長條對應的變數類別。圖表不得沒有變數與統計量。您可從畫布左邊的「變數」清單中拖曳變數以新增。

變數的測量層級在「圖表建置器」中相當重要。您即將使用 x 軸上的「工作滿意度」變數。然而，變數旁的圖示（看起來像尺）表示變數的測量層級已定義為尺度。若要建立正確的圖表，您必須使用類別測量層級。您可在「圖表建置器」中暫時變更測量層級，而不用回到上一步變更「變數視圖」中的測量層級。

1. 在「變數」清單上的「工作滿意度」按一下滑鼠右鍵，並選擇「序數」。序數是適當的測量層級，因為「工作滿意度」中的類別可利用滿意度層級加以分級。請注意，當您變更測量層級後，圖示會變更。
2. 現在，將「工作滿意度」從「變數」清單拖曳到 x 軸放置區域。

y 軸放置區域預設為「個數」統計量。若您希望使用其他統計量（如百分比或平均數），可輕易變更。本範例中不會使用這些統計量，但我們會複習這些程序，以因應您其他時候變更統計量的需求。

3. 按一下「圖表建置器」側邊列中的**元素內容**標籤。（如果側邊列未顯示，請按一下「圖表建置器」右上角中的按鈕以顯示側邊列。）



圖 34. 元素內容

「元素內容」可讓您變更各種圖表元素的內容。這些元素包括圖形元素（如長條圖中的長條）和圖表上的軸。從「編輯內容」清單中選取一個元素，以變更與此元素關聯的內容。亦請注意，清單右邊會有紅色的 X。此按鈕會從畫布刪除圖形元素。因為選取了「**Bar1**」，所顯示的性質會套用至圖形元素，特別是長條圖形元素。

此「統計量」下拉清單會顯示可用的特定統計量。每個圖表類型通常都可使用相同的統計量。請注意，有些統計量需要 y 軸放置區域包含變數。

4. 將家庭收入（以千為單位）從「變數」清單拖曳至 y 軸放下區域。由於 y 軸上的變數是尺度變數，而 x 軸上的變數是類別變數（序數為類別測量層級的一種），因此 y 軸放置區域預設為「平均數」統計量。這些已是您需要的變數與統計量，因此無須變更元素性質。

新增文字

您也可以在圖表上新增標題與註腳。

1. 按一下「標題/註腳」標籤。
2. 選取「標題 1」。

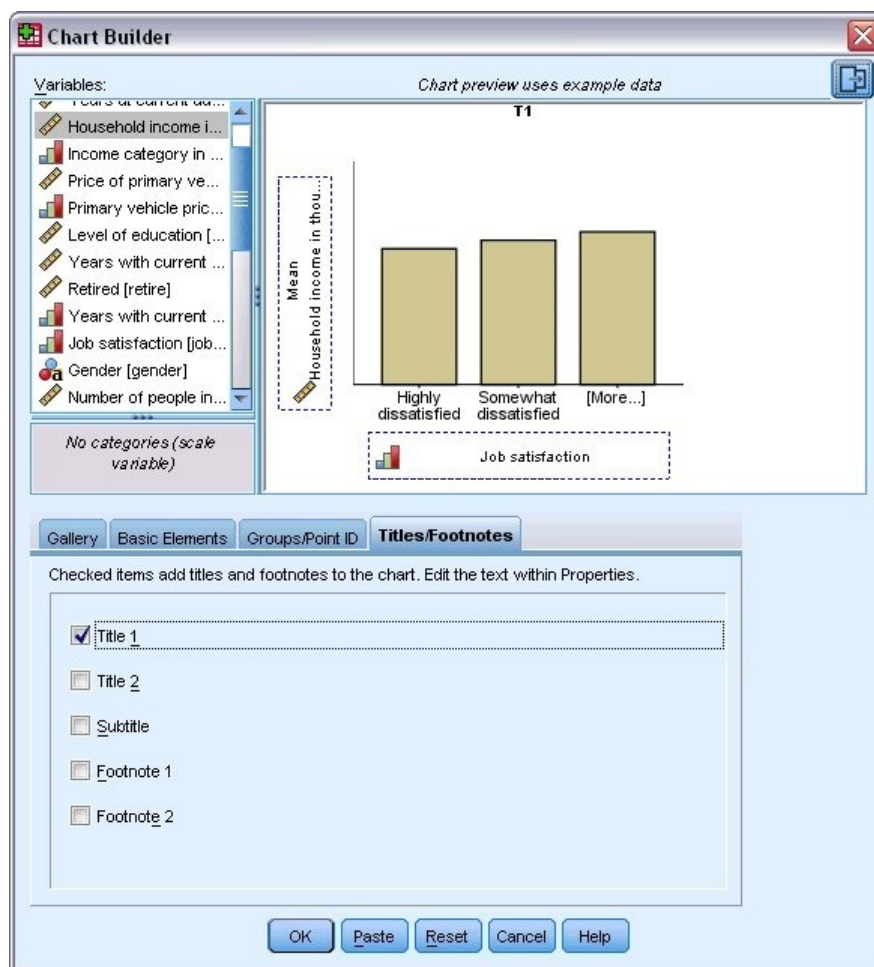


圖 35. 畫布內顯示的「標題 1」

此標題會在畫布內以標記 **T1** 顯示。

3. 在「元素內容」標籤的「編輯內容」清單中，選取「**標題 1**」。
4. 在「內容」文字方框中，輸入收入（依工作滿意度分）。這就是會顯示出來的標題文字。

建立圖表

1. 按一下「**確定**」以建立長條圖。

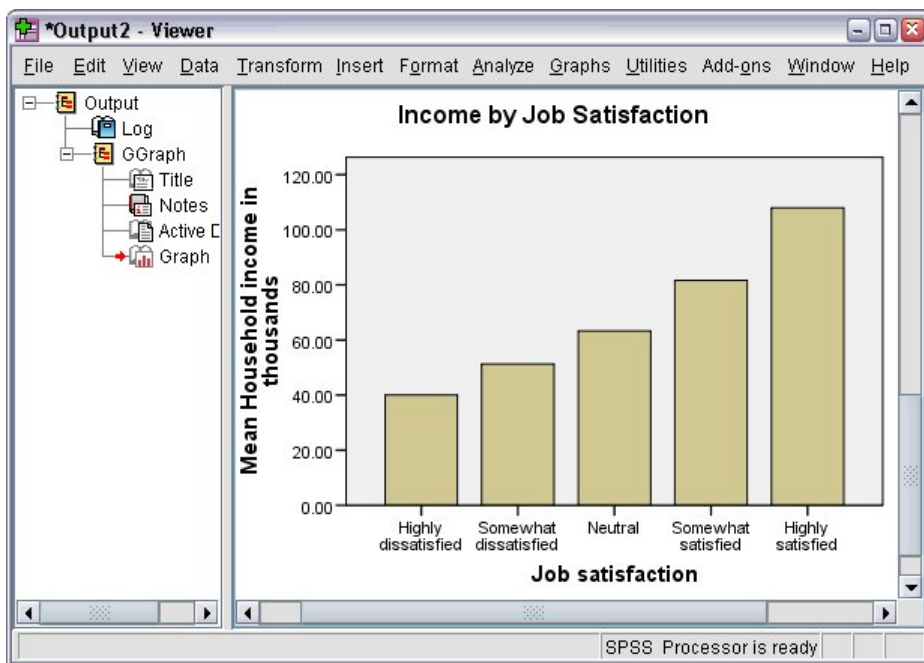


圖 36. 長條圖

長條圖顯示，對自己工作較滿意者，傾向擁有較高的家庭收入。

第 6 章 輸出的處理

在「檢視器」中顯示執行統計程序的結果。產生的輸出可以使用統計表、圖表、統計圖或文字的方式呈現，這視您執行程序時的選擇而定。本節使用 *viewertut.spv* 和 *demo.sav* 檔。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

使用檢視器

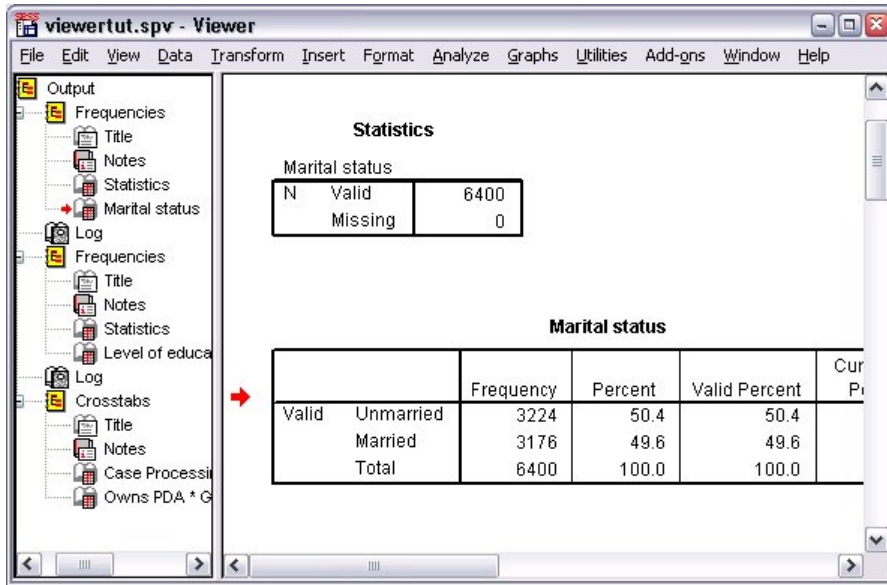


圖 37. 檢視器

「檢視器」視窗內有兩個框架：**大綱窗格**內包含「檢視器」中所有資訊的概要。**目錄框**則包含統計表、圖表和文字輸出。

使用捲軸，便可以水平和垂直方向來瀏覽視窗內容。若要以更簡單的方式瀏覽，按一下大綱窗格中的項目，目錄框內便會顯示該項目的內容。

1. 按一下大綱窗格的右邊界並拖曳，便可改變大綱窗格的寬度。

在大綱窗格內有一個開啟書本的圖像，表示該項目雖然目前在目錄框內看不到，但在「檢視器」內可以看見該項目。

2. 若要隱藏表格或圖表，可連按兩下大綱窗格中該項目的書本圖像。

開啟的書本圖像會闔起來，代表已隱藏與該項目的相關資訊。

3. 若要重新顯示隱藏的輸出結果，連按兩下關閉書本圖像即可。

您也可以隱藏特定統計程序所有的輸出結果，或是「檢視器」內所有的輸出結果。

4. 若於您要隱藏程序的輸出結果，可按一下程序旁內含減號 (-) 的方框，或是按一下大綱窗格中最上層項目旁的方框，便可隱藏所有的輸出。

概要收合起來，在視覺上表示這些結果已經隱藏。

您也可以變更顯示輸出的順序。

5. 在大綱窗格中，按一下您要移動的項目。
6. 將所選項目拖曳到大綱窗格內的新位置。

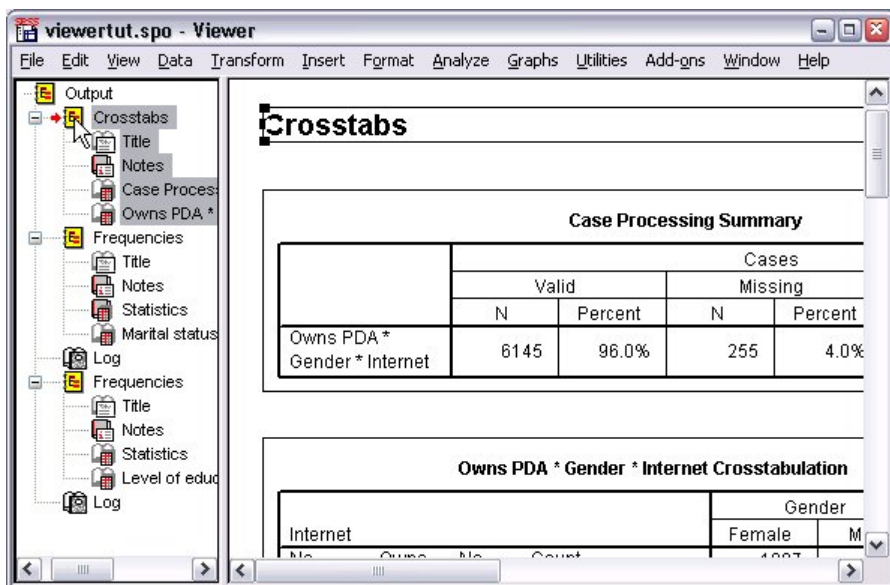


圖 38. 將「檢視器」中的輸出結果重新排序

您也可以在目錄框中按一下並拖曳輸出結果，來移動輸出項目。

使用「樞紐表編輯器」

大部分的統計程序結果會以樞紐表顯示。

存取輸出定義

在輸出結果中會顯示許多統計辭彙。這些辭彙的定義可直接從「檢視器」中存取。

1. 連按兩下「Owns PDA * Gender * Internet Crosstabulation」表格。
2. 在「期望個數」按一下滑鼠右鍵，再從蹦現功能表中，選擇「這是什麼？」。

在即現式視窗中會出現詞彙的定義。

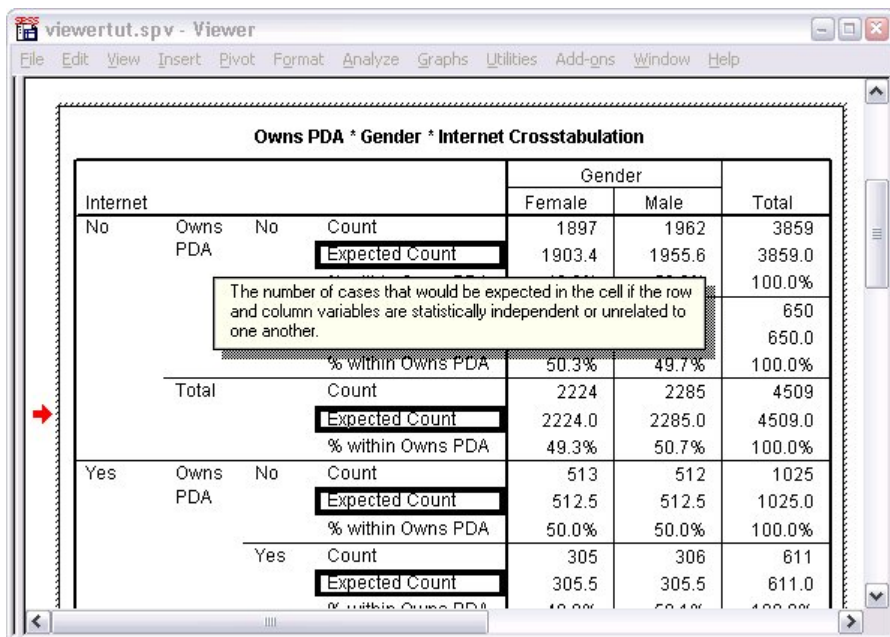


圖 39. 蹦現定義

樞紐表

產生出預設表格，其中顯示的資訊可能不夠整齊、清楚。在樞紐表上，有許多方法可以轉置列和欄（翻轉表格）、調整表格內的資料順序和修改表格。例如，您可以藉由轉置列與欄，將短的寬表格變成長的窄表格。變更表格配置不會影響結果。而是以不同或較符合需求的方式，來顯示資訊。

1. 如果尚未啟動，請連按兩下「擁有 PDA * 性別 * 網際網路交叉列表」表格將其啟動。
2. 如果看不到「樞紐分析匣」視窗，請從功能表選擇：

樞紐 > 樞紐分析匣

「樞紐分析匣」功能可在直欄、列和階層之間移動資料。

Internet			Gender		Total
			Female	Male	
Owens PDA	No	Count	1897	1962	3859
		Expected Count	1903.4	1955.6	3859.0
	Yes	% within Owens PDA	49.2%	50.8%	100.0%
		Count	327	323	650

圖 40. 樞紐分析匣

3. 將「統計量」元素從「列」維度拖曳至「性別下的「直欄」維度。表格立即重新配置，以反映您的變更。

元素在樞紐分析匣中的順序，代表元素在表格中的順序。

4. 將「擁有 PDA」元素拖放至列維度中的「網際網路」元素之前，以反轉這兩列的順序。

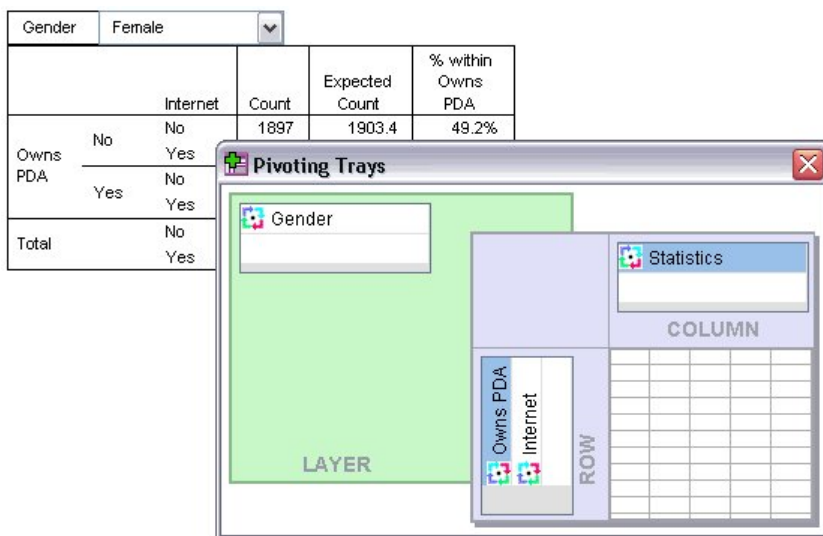
			Gender				Total		
			Female		Male		Count	Expected Count	% within Owens PDA
Internet	Count	Expected Count	% within Owens PDA	Count	Expected Count	% within Owens PDA			
Owens PDA	No	1897	1903.4	49.2%	1962	1955.6	50.8%	3859.0	100.0%
	Yes	327	323	650	100.0%	611.0	100.0%	4509.0	100.0%
Total	No	2124	2127	4251	2184	2184	4368	1636.0	100.0%
	Yes	327	323	650	611.0	100.0%	1636.0	100.0%	

圖 41. 交換列

建立和顯示階層

對於具有巢狀類別資訊的大型表格而言，分層是很有用的功能。藉由建立階層，您可以簡化表格的外觀，讓表格更容易檢視。

1. 將「性別」元素從「欄」維度拖曳至「階層」維度。



Gender	Female				
		Internet	Count	Expected Count	% within Owns PDA
		No	1897	1903.4	49.2%
Owns PDA	No	Yes			
	Yes	No			
		Yes			
Total	No				
	Yes				

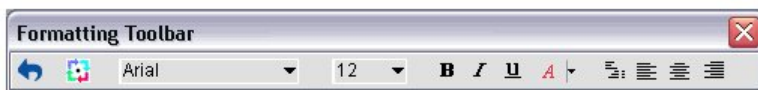
圖 42. 於「層級」維度中的「性別」樞紐圖像

若要顯示不同的分層，請從表格的下拉清單中選取類別。

編輯表格

除非您另行自訂「表格格式集」，否則一般會以標準格式來建立樞紐表。您可以變更表格中任何文字的格式。您可以變更的格式包括字型名稱、字型大小、字型樣式（粗體或斜體）和色彩。

1. 連按兩下「教育程度」表格。
2. 如果看不到「格式」工具列，請從功能表選擇：
檢視 > 工具列
3. 按一下標題文字，「教育程度」。
4. 在工具列上的字型大小下拉清單上，選擇「**12**」。
5. 若要變更標題文字的色彩，請按一下文字色彩工具，並選擇新色彩。



Level of education					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7	21.7
	High school degree	1936	30.3	30.3	52.0
	Some college	1360	21.3	21.3	73.2
	College degree	1355	21.2	21.2	94.4
	Post-undergraduate degree	359	5.6	5.6	100.0
	Total	6400	100.0	100.0	

圖 43. 重設樞紐表內標題文字的格式

您也可以編輯表格目錄和標記。例如，您可以變更此表格的標題。

6. 連按兩下標題。
7. 輸入教育程度作為新標記。

附註：如果您變更表格內的數值，並不會重新計算總計和其他統計值。

隱藏列和欄

表格內的有些資料可能沒有用處，或者讓表格變得不必要的複雜。幸好您可以隱藏整個列和欄，但不會遺失任何的資料。

1. 如果尚未啟動，請連按兩下「教育程度表」表格將其啟動。
2. 按一下「有效百分比」欄標籤以選取該欄。
3. 從「編輯」功能表選擇，或用滑鼠右鍵按一下蹦現功能表選擇：

選取 > 資料及標籤資料格

4. 從「檢視」功能表中，選擇「隱藏」，或用滑鼠右鍵按一下蹦現功能表，選擇「隱藏類別」。
該欄會被隱藏起來，但不會被刪掉。

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	21.7	21.7
	High school degree	1936	30.3	52.0
	Some college	1360	21.3	73.2
	College degree	1355	21.2	94.4
	Post-undergraduate degree	359	5.6	100.0
	Total	6400	100.0	

圖 44. 表格中「有效百分比」一欄已隱藏

若要重新顯示該欄：

5. 在功能表上，選擇：

檢視 > 全部顯示

顯示和隱藏列的方法與欄相同。

變更資料顯示格式

若要改變樞紐表資料的顯示格式，其實很容易。

1. 如果尚未啟動，請連按兩下「教育程度表」表格將其啟動。
2. 按一下「百分比」直欄標題以選取該欄。
3. 從「編輯」功能表選擇，或用滑鼠右鍵按一下蹦現功能表選擇：

選取 > 資料格

4. 從「格式」功能表選擇，或用滑鼠右鍵按一下蹦現功能表選擇「資料格內容」。
5. 按一下「格式化數值」標籤。
6. 在「小數點」欄位中輸入 0，隱藏此欄中所有的小數點。

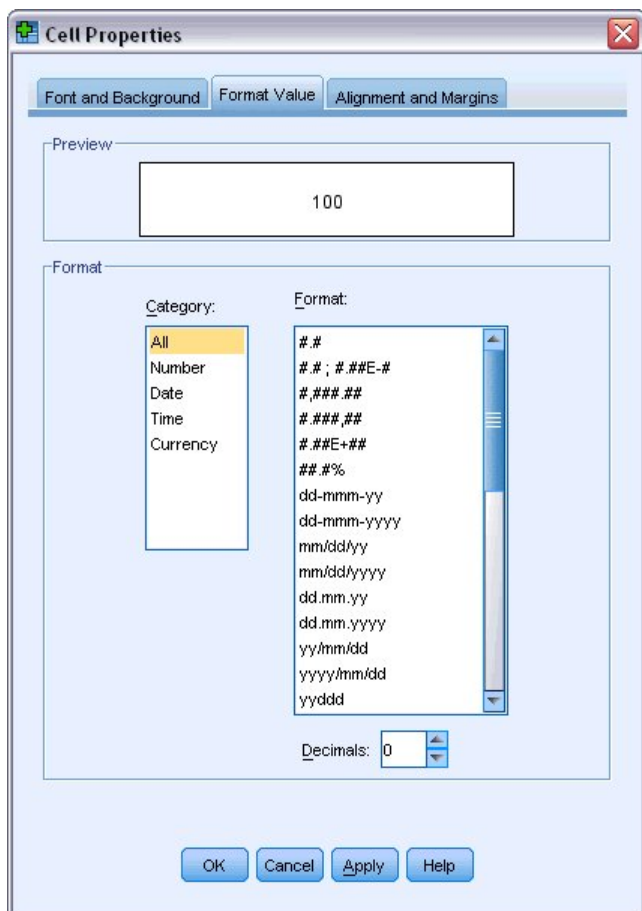


圖 45. 「資料格性質」的「格式化值」標籤

在此對話框中也可以變更資料類型和格式。

7. 在「類別」清單中，選取您要的類型，再從「格式」清單中選取類型的格式。
8. 按一下「確定」或「套用」來套用變更。

Education Level

		Frequency	Percent	Cumulative Percent
Valid	Did not complete high school	1390	22	21.7
	High school degree	1936	30	52.0
	Some college	1360	21	73.2
	College degree	1355	21	94.4
	Post-undergraduate degree	359	6	100.0
	Total	6400	100	

圖 46. 「百分比」一欄隱藏了小數點

百分比一欄的小數點現在已隱藏。

表格格式集

為了提供清楚、簡明、有意義的結果，那麼表格的格式是關鍵。如果您的表格很難檢視，則可能會不容易了解表格內的資訊。

使用預先定義的格式

1. 連按兩下「婚姻狀態」表格。
2. 在功能表上，選擇：

格式 > 表格格式集...

「表格格式集」對話框列出各種預先設定的樣式。選取清單中的樣式，可從右邊的「範例」視窗中預覽該樣式。

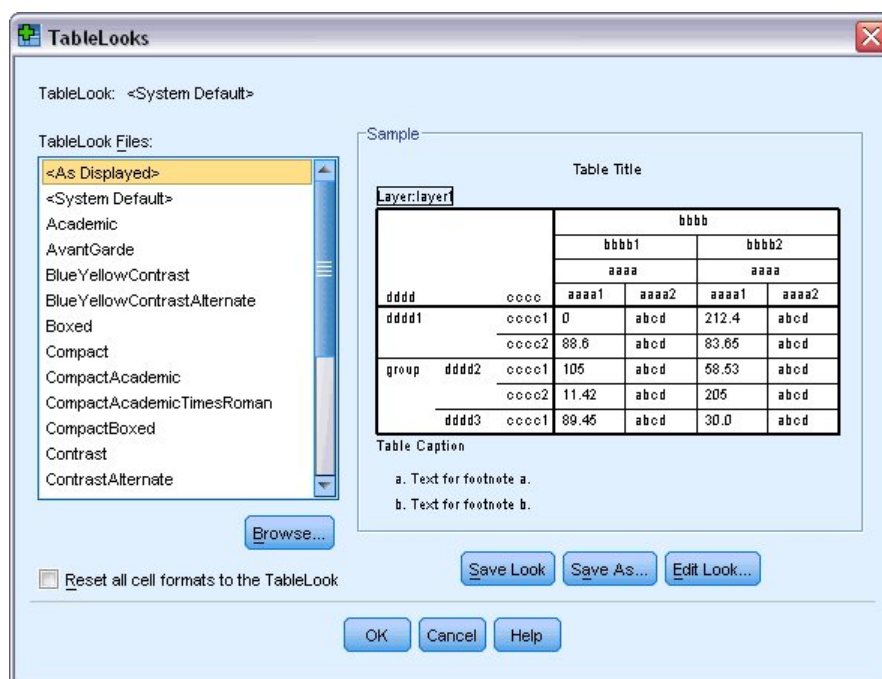


圖 47. 表格格式集對話框

您可以使用原有的樣式，或將現有樣式編輯成符合您需要的樣式。

3. 若要使用現有樣式，請選取一個樣式，並按一下「確定」。

自訂表格格式集樣式

您可以自訂格式，以符合您特定的需求。從背景色彩到框線樣式，幾乎表格的所有層面都可以自訂。

1. 連按兩下「婚姻狀態」表格。
2. 在功能表上，選擇：

格式 > 表格格式集...

3. 選取最接近您所需格式的樣式，按一下「編輯表格設定」。
4. 按一下「資料格格式」標籤，檢視格式選項。

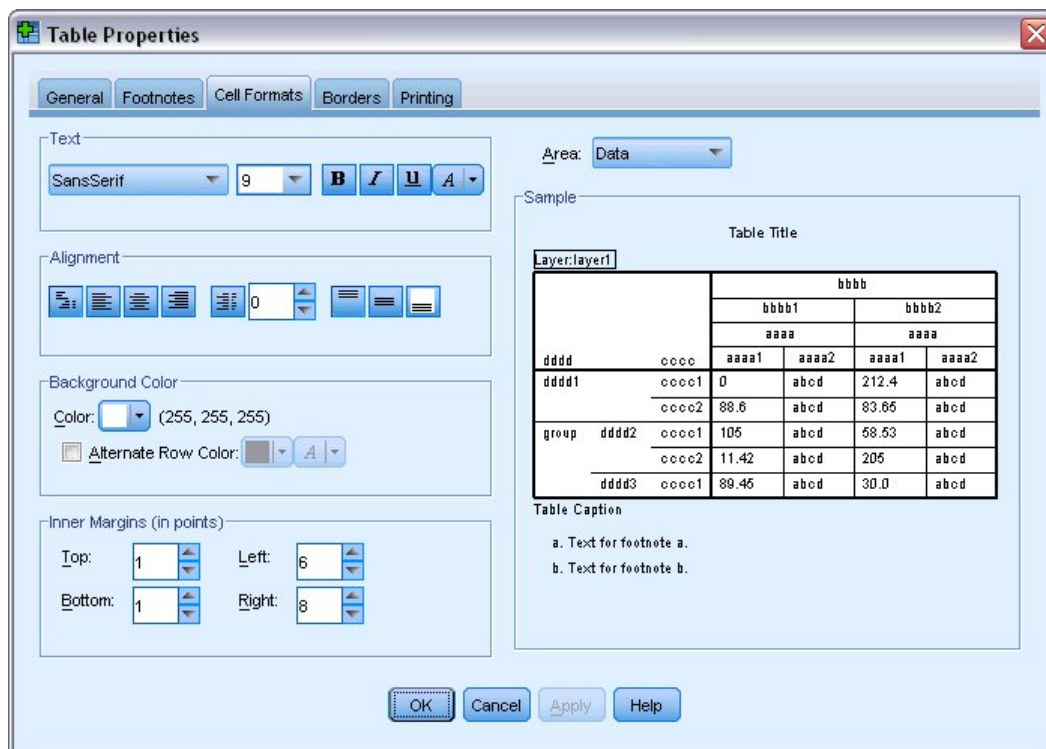


圖 48. 「表格內容」對話框

格式選項包括字型名稱、字型大小、樣式和色彩。其他選項還包括對齊、文字和背景色彩、以及邊界大小。

右側的「範例」視窗中，可看到表格在格式變更後的結果。表格的每一個區域都可以有不同的格式樣式。例如，標題的樣式可以與資料樣式不同。若要選擇欲編輯的表格區域，您可以選擇「區域」下拉清單內的區域名稱，或者在「範例」視窗中按一下您要變更的區域。

5. 選取「區域」下拉清單中的「資料」。
6. 從「背景」下拉式面板中，選取一個新色彩。
7. 接著選取新的文字色彩。

「範例」視窗顯示新的樣式。

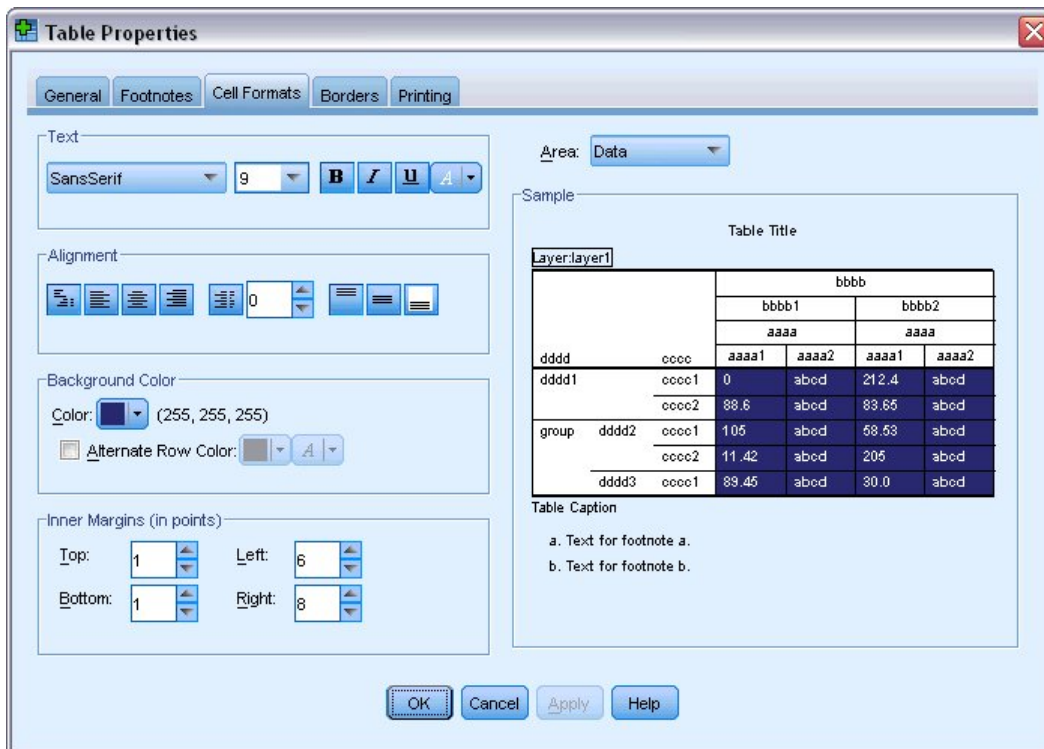


圖 49. 變更表格資料格式

8. 按一下「確定」，回到「表格格式集」對話框。

您可以儲存您的新樣式，讓您在日後可輕易地套用樣式。

9. 按一下「另存新檔」。
10. 導覽至目標目錄，在「檔名」文字框中輸入新樣式的名稱。
11. 按一下「儲存」。
12. 按一下「確定」套用您的變更，並返回「檢視器」。

表格現在包含您指定的自訂格式。

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unmarried	3224	50.4	50.4	50.4
	Married	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

圖 50. 自訂表格格式集

變更預設的表格格式

雖然表格格式在建立後仍可以變更，但變更預設的「表格格式集」可能會更有效率，如此您不用每次建立表格時都要變更格式。

若要變更樞紐表的預設「表格格式集」樣式，請從功能表選擇：

編輯 > 選項...

1. 請在「選項」對話框中，按一下「樞紐表」標籤。

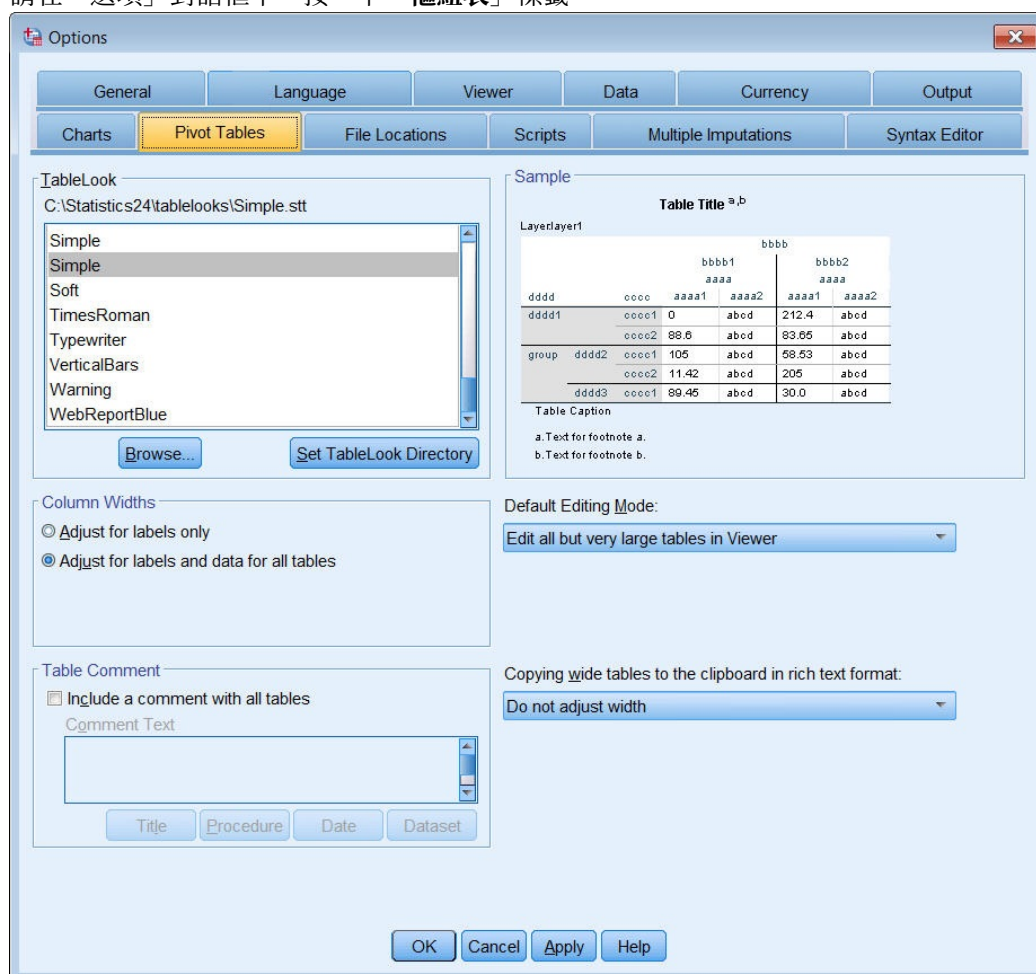


圖 51. 「選項」對話框

2. 選取要用於所有新表格上的「表格格式集」樣式。

右側的「範例」視窗顯示每個「表格格式集」的預覽。

3. 按一下「確定」儲存您的設定，並關閉對話框。

所有在變更預設表格格式集後建立的表格，會自動符合新的格式規則。

自訂初始顯示設定

初始顯示設定包括「檢視器」內物件的對齊方式、物件是否依預設顯示或隱藏、及「檢視器」視窗的寬度。變更這些設定：

1. 在功能表上，選擇：

編輯 > 選項...

2. 按一下「檢視器」標籤。

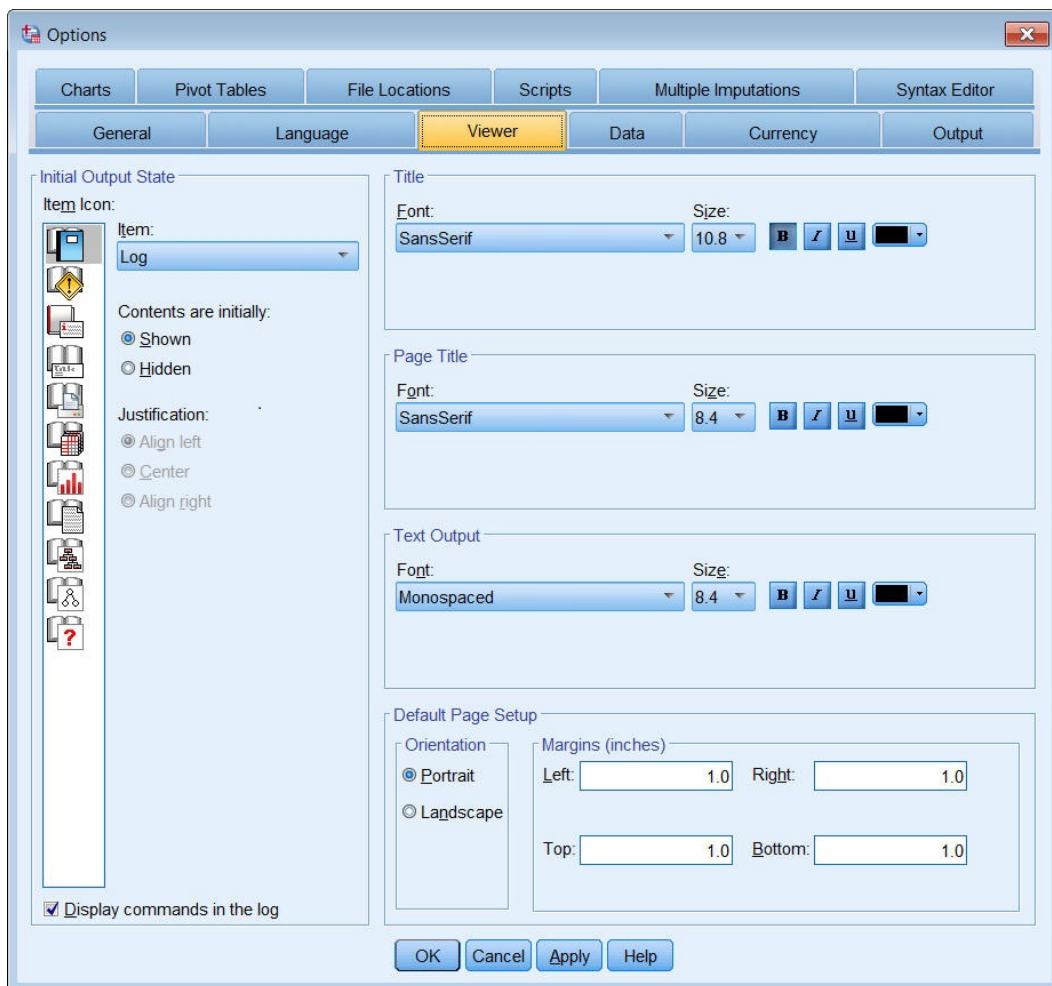


圖 52. 「檢視器」選項

逐一將物件套用設定。例如，您可以自訂圖表的顯示方式，而不需對表格的顯示方式做任何變更。只要選取您要自訂的物件，再加以變更。

3. 按一下「標題」圖像以顯示其設定。
4. 按一下「對中」，讓所有標題在「檢視器」的中央（水平）顯示。

您也可以隱藏可能會使輸出變得擁擠的元素，如日誌和警告訊息在圖像上連接兩下，可自動變更物件的顯示性質。

5. 連接兩下「警告」圖像，隱藏輸出中的警告訊息。
6. 按一下「確定」儲存您的設定，並關閉對話框。

顯示變數和值標籤

在大部分的情況下，顯示變數和數值標籤比顯示變數名稱或實際的資料值還有效率。然而在某些情況下，您可能要同時顯示名稱和標記。

1. 在功能表上，選擇：
編輯 > 選項...
2. 按一下「輸出標記」標籤。

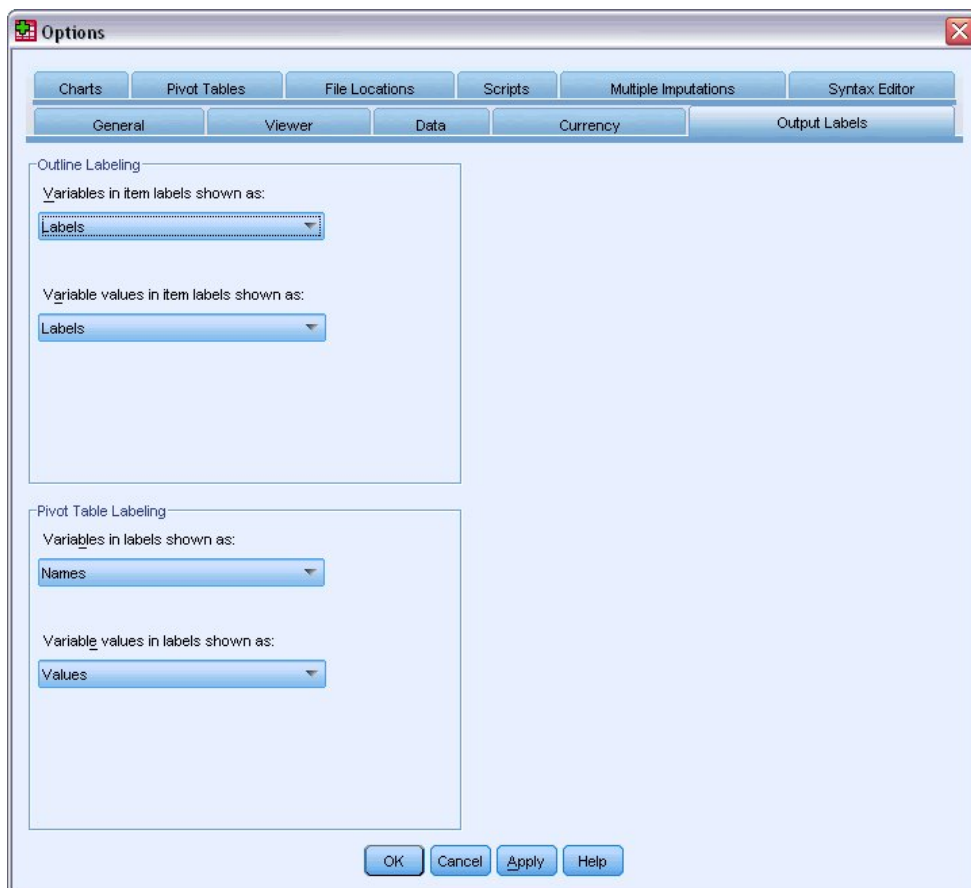


圖 53. 「樞紐表標記」設定

您可以為大綱窗格和目錄框指定不同的設定。例如，在大綱窗格中顯示標記，在目錄框中顯示資料值：

3. 在樞紐表標記群組中，選取「標籤中的變數」下拉清單中的「名稱」，以顯示變數的名稱而非標記。
4. 接著，選取「變數值標籤」下拉清單的「數值」，以顯示資料值而非標記。

後續在階段作業中產生的表格會反映這些變更。

marital

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0	3224	50.4	50.4	50.4
	1	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

圖 54. 顯示變數名稱和數值

使用其他應用程式中的結果

許多應用程式都可使用您的結果。例如，您可以在簡報或報告中包含表格或圖表。

下列範例是以 Microsoft Word 為例，但它的作業方式與其他文書處理應用程式很類似。

將結果貼上為 Word 表格

您可以將樞紐表貼到 Word 中，如同 Word 的原生表格。所有表格屬性，例如字型大小和色彩，都會保留。因為表格是以 Word 表格格式貼上，所以您可以使用 Word 編輯，就如同任何其他表格。

1. 按一下「檢視器」中的表格以選取該表。
2. 在功能表上，選擇：

編輯 > 複製

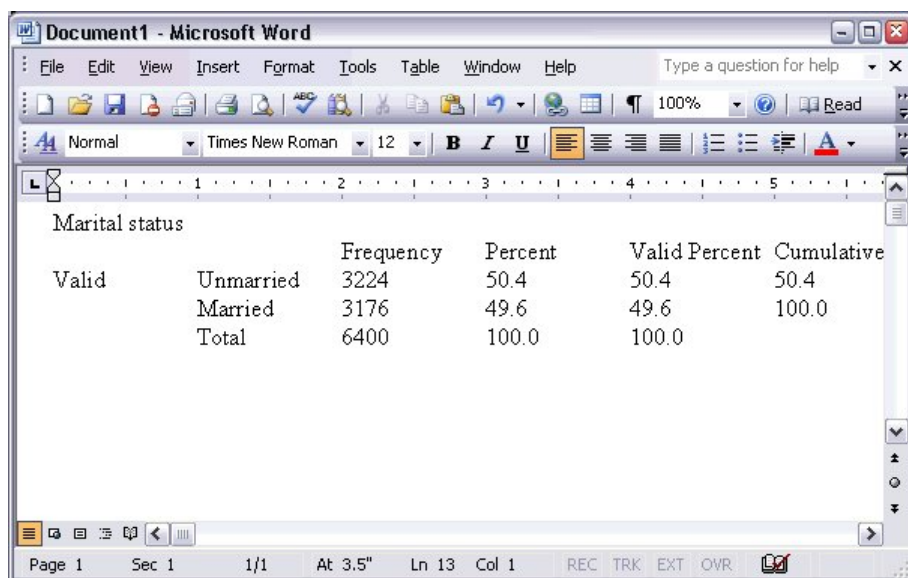
3. 開啟您的文書處理應用程式。
4. 在文書處理器功能表中選擇：
編輯 > 選擇性貼上...
5. 在「選擇性貼上」對話框中，選取「格式化文字 (RTF)」。
6. 按一下「確定」，將結果貼到現有的文件上。

您的文件現在顯示一個表格。您可以視需要套用自訂的格式、編輯資料並改變表格大小。

將結果以文字形式貼上

樞紐表可以純文字格式複製到其他應用程式中。使用這種方法不會保留表格的格式樣式，但在表格貼入目標應用程式後，您仍能夠編輯表格資料。

1. 按一下「檢視器」中的表格以選取該表。
2. 在功能表上，選擇：
編輯 > 複製
3. 開啟您的文書處理應用程式。
4. 在文書處理器功能表中選擇：
編輯 > 選擇性貼上...
5. 在「選擇性貼上」對話框中，選取「未設定格式的文字」。
6. 按一下「確定」，將結果貼到現有的文件上。



The screenshot shows a Microsoft Word window titled 'Document1 - Microsoft Word'. The ribbon is set to 'Table' and the 'Table' menu is open. The table displayed is a pivot table with the following data:

Marital status		Frequency	Percent	Valid Percent	Cumulative
Valid	Unmarried	3224	50.4	50.4	50.4
	Married	3176	49.6	49.6	100.0
	Total	6400	100.0	100.0	

圖 55. 在 Word 中顯示樞紐表

表格中的每一個欄以定位點區隔。您可以調整文書處理應用程式中的定位點，來變更欄的寬度。

將結果匯出至 Microsoft Word、PowerPoint 和 Excel 檔

您可以將結果匯出到 Microsoft Word、PowerPoint 或 Excel 檔案。您可以匯出「檢視器」中的選取項目或所有項目。本節使用 *msoutut.spv* 和 *demo.sav* 檔。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

附註：「匯出至 PowerPoint」功能僅適用於 Windows 作業系統，學生版不提供此功能。

在「檢視器」概要框中，您可選取要匯出的特定項目，或匯出所有項目或所有可見項目。

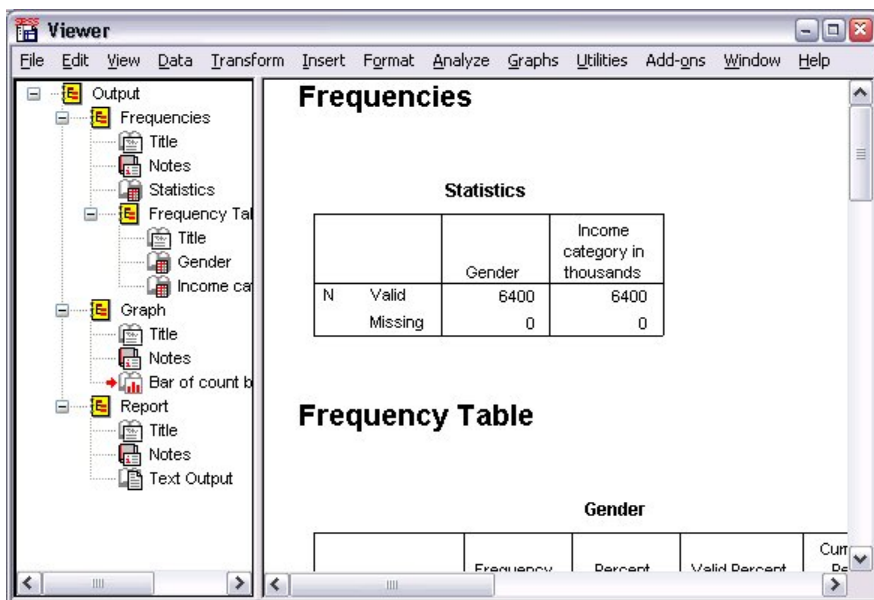


圖 56. 檢視器

1. 從「檢視器」功能表選擇：

檔案 > 匯出...

除了將「檢視器」內所有物件匯出，您也可以選擇僅匯出可見的物件（大綱窗格中開啟的書本），或是您在大綱窗格中選取的物件。如果您在大綱窗格中沒有選取任何項目，您便無法匯出選取的物件。

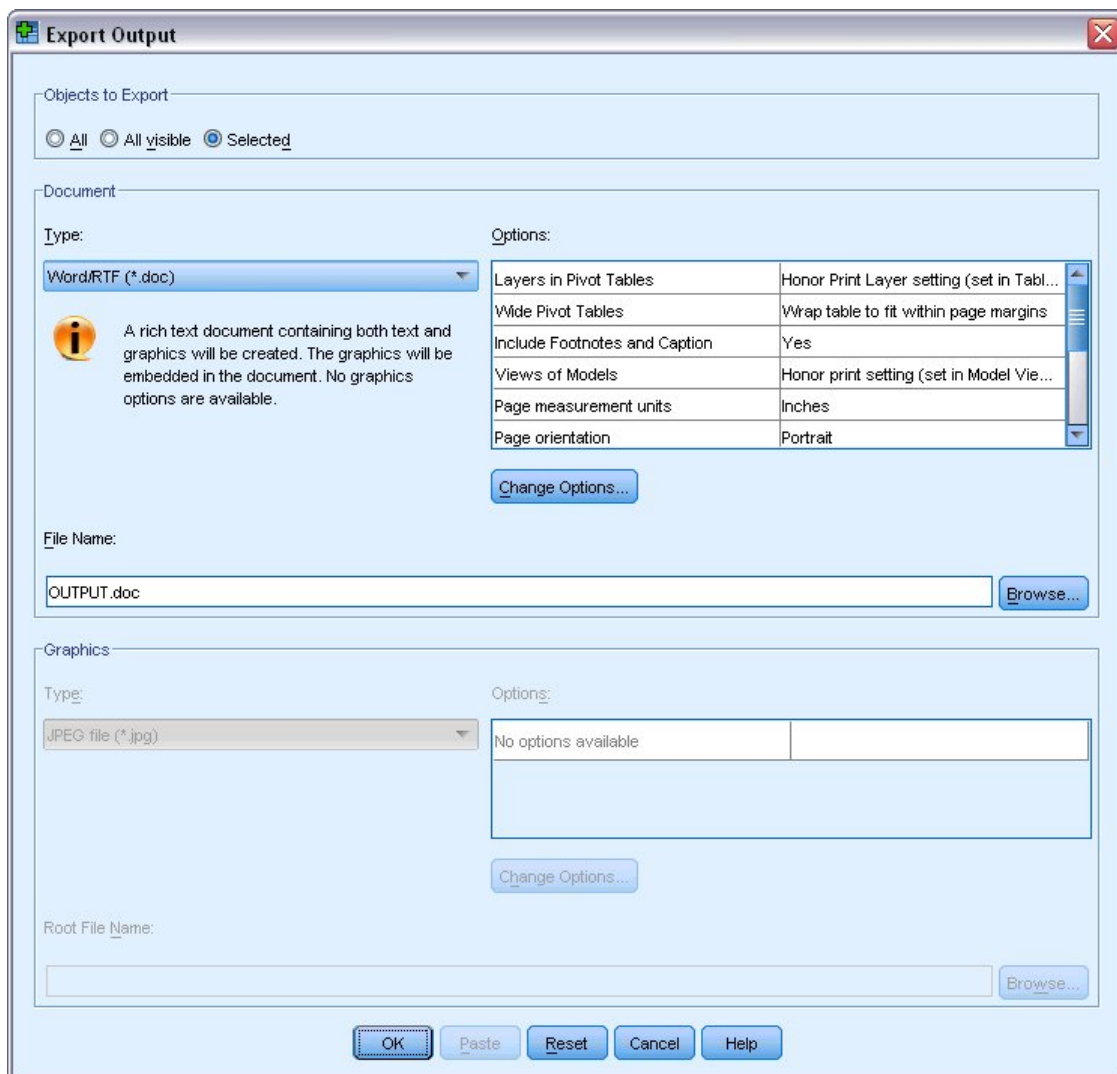


圖 57. 「匯出輸出」對話框

2. 在「要匯出的物件」群組中，選取「全部」。
3. 從「類型」下拉清單中，選取「**Word/RTF 檔案 (*.doc)**」。
4. 按一下「確定」，產生 Word 檔案。

當您以 Word 開啟結果檔案，您可以看到結果匯出的方式。請注意，因為您選擇匯出所有物件，看不到的物件也會出現在 Word 中。

樞紐表變成 Word 表，保留原始樞紐表的所有格式，包括字型、色彩、框線等等。

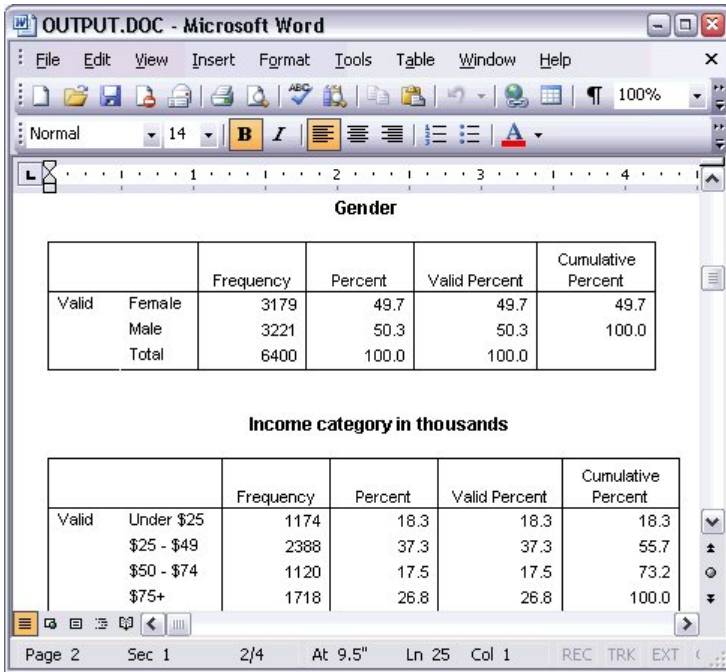


圖 58. Word 中的樞紐表

圖表會以圖形影像的形式包含在 Word 文件中。

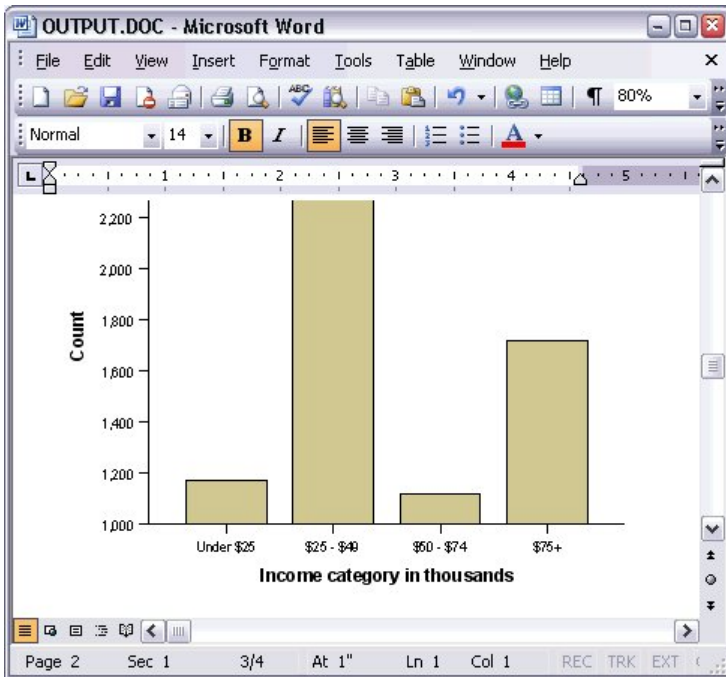


圖 59. Word 中的圖表

系統以「檢視器」中文字物件所使用的同一種字型顯示文字輸出。為了能適當的對齊，文字輸出應使用固定點數（單一空間）字型。

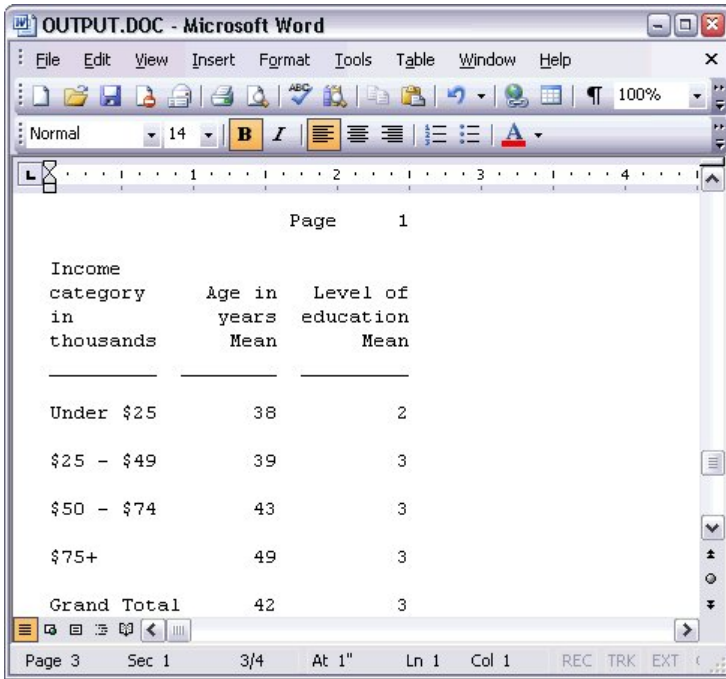


圖 60. Word 中的文字輸出

如果匯出為 PowerPoint 檔案，每個匯出項目會放置在不同的投影片中。匯出至 PowerPoint 的樞紐表會變成 Word 表，具有原始樞紐表的所有格式，包括字型、色彩、框線等等。

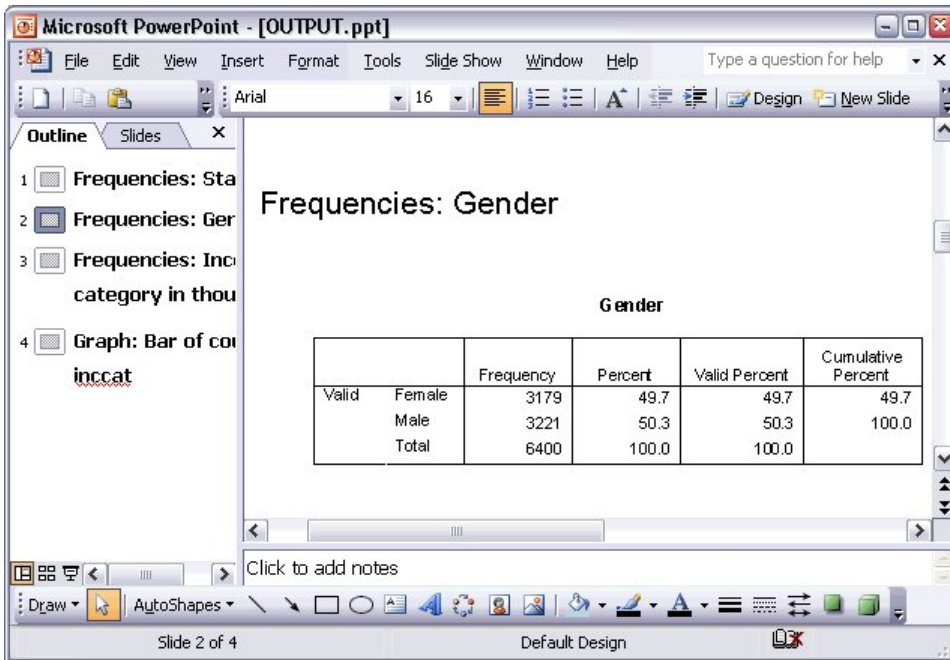


圖 61. PowerPoint 中的樞紐表

選取匯出至 PowerPoint 的圖表會嵌入至 PowerPoint 檔中。

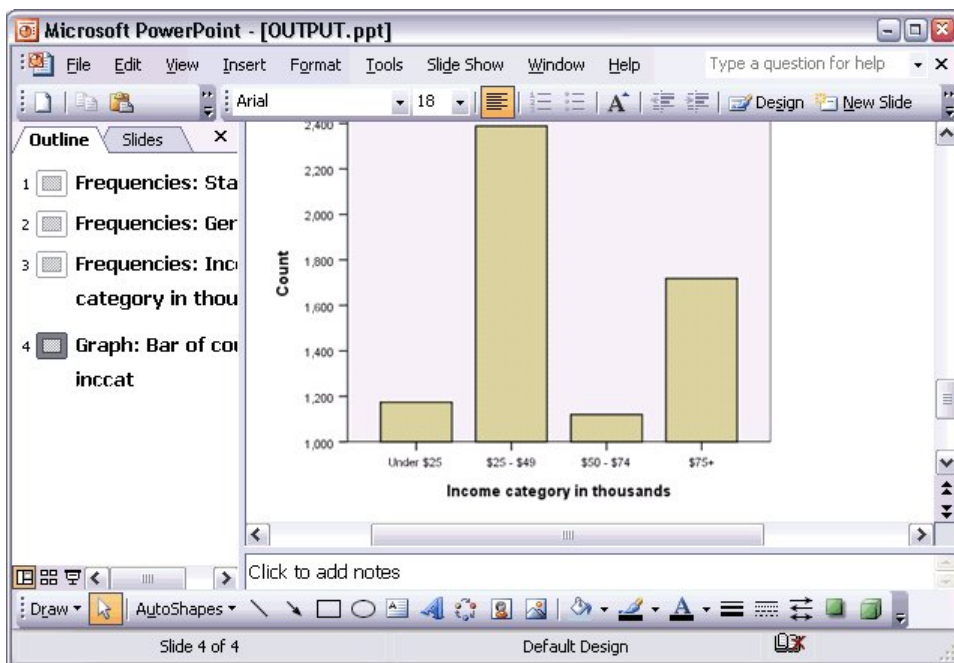


圖 62. PowerPoint 中的圖表

附註：「匯出至 PowerPoint」功能僅適用於 Windows 作業系統，學生版不提供此功能。如果您匯出 Excel 檔，匯出的結果也不同。

Gender		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Female	3,179	49.7	49.7	49.7
	Male	3,221	50.3	50.3	100.0
	Total	6,400	100.0	100.0	

Income category in thousands		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Under \$25	1,174	18.3	18.3	18.3
	\$25 - \$49	2,388	37.3	37.3	55.7
	\$50 - \$74	1,120	17.5	17.5	73.2
	\$75+	1,718	26.8	26.8	100.0
	Total	6,400	100.0	100.0	

圖 63. Excel 中的 Output.xls

樞紐表的列、欄和資料格，變成了 Excel 的列、欄和資料格。

文字輸出中的每一行在 Excel 檔中都是一列，且資料格內容為所含行之全部內容。

	A	B	C	D	E
68	Income				
69	category	Age in	Level of		
70	in	years	education		
71	thousands	Mean	Mean		
72					
73					
74	Under \$25	38	2		
75					
76	\$25 - \$49	39	3		
77					
78	\$50 - \$74	43	3		
79					
80	\$75+	49	3		
81					
82	Grand Total	42	3		

圖 64. Excel 中的文字輸出

匯出結果至 PDF

您可以將「檢視器」中所有的項目或選取的項目匯出至 PDF（可攜式文件格式）檔案。

1. 從含有您要匯出至 PDF 的結果的「檢視器」視窗中，選擇以下功能表：

檔案 > 匯出...

2. 在「匯出輸出」對話框中，從「匯出格式檔案類型」下拉清單中選擇「可攜式文件格式」。

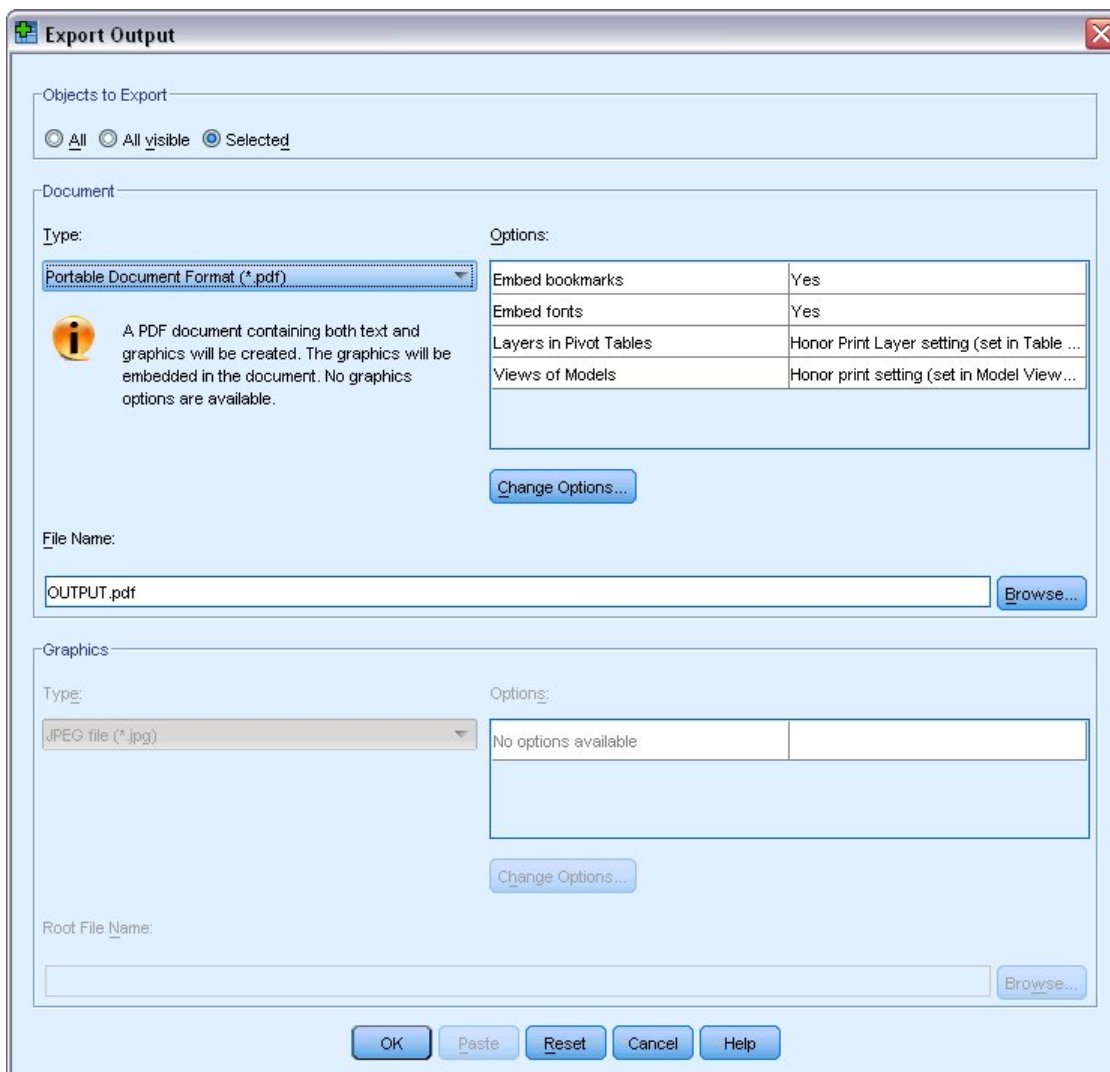


圖 65. 「匯出輸出」對話框

- 「檢視器」文件的概要框在 PDF 檔案中會轉換成為書籤，以便於巡覽。
- 頁面大小、方向、邊緣、頁首及頁底的内容與顯示、PDF 文件的印出圖表大小等，都是由頁面設定選項（「檢視器」視窗其「檔案」功能表中的「頁面設定」）來控制。
- PDF 文件的解析度 (DPI) 是預設印表機或目前所選印表機（可使用「頁面設定」進行變更）的目前解析度設定。最高解析度為 1200 DPI。如果印表機的設定較高，PDF 文件解析度仍然是 1200 DPI。附註：以較低解析度的印表機進行列印時，高解析度文件的品質會降低。

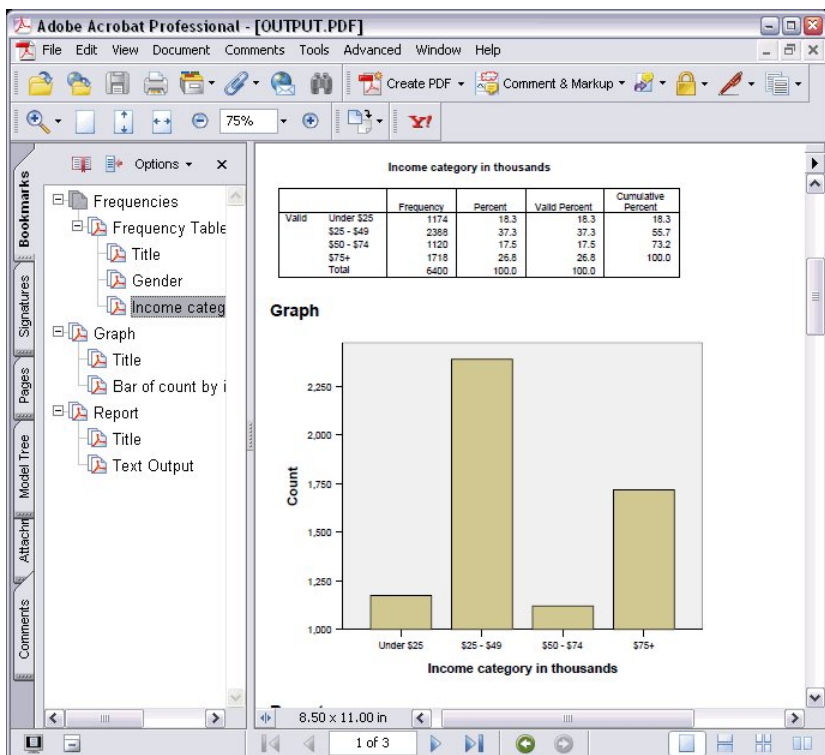


圖 66. 含有書籤的 PDF 檔案

匯出結果至 HTML

您也可以將結果匯出至 HTML（超文字標記語言）。儲存為 HTML 時，所有非圖形輸出結果會匯出至單一的 HTML 檔案。

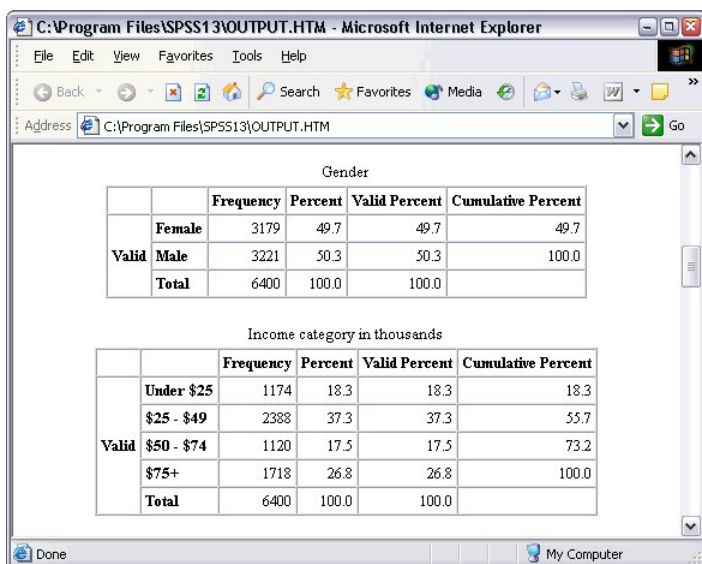


圖 67. Web 瀏覽器內的 Output.htm

當您匯出至 HTML，也會匯出圖表，但不是匯至單一的檔案。

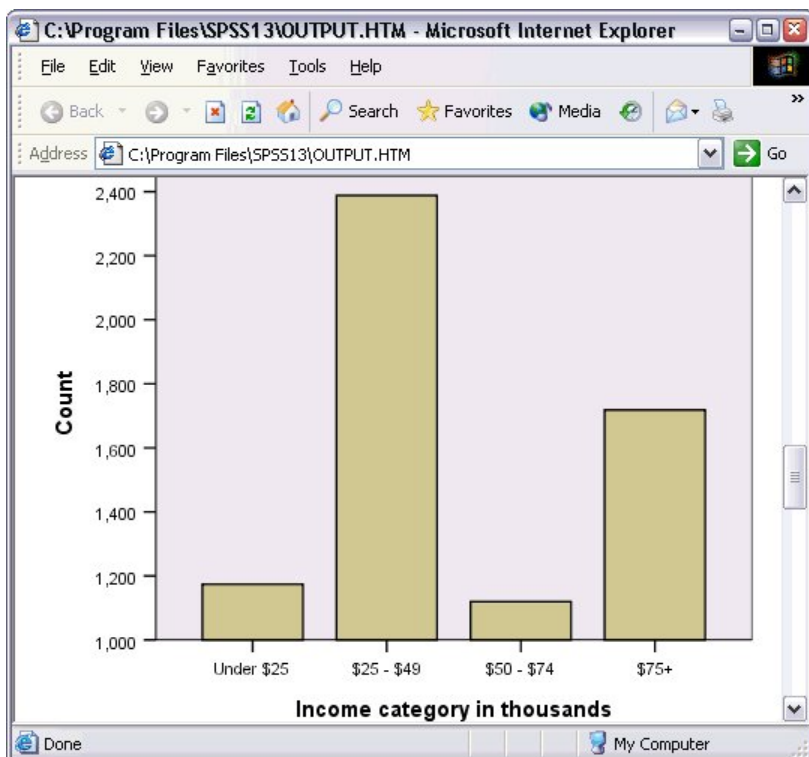


圖 68. HTML 的圖表

每一個圖表都可以儲存成您指定格式的檔案，這些圖形檔的參照會放在 HTML 中。也可匯出所有圖表（或選取的圖表）到不同的圖形檔中。

第 7 章 使用語法

您可以使用功能強大的指令語言，儲存並自動化許多共通的工作。它也提供部分在功能表和對話框中沒有的功能。大部分的指令，都是透過功能表和對話框來存取的。然而有些指令和選項，卻只能透過指令語言才能使用。您也可以指令語言將工作儲存到語法檔中，以便您可以在日後重複分析。

指令語法檔只是一個文字檔，其內包含 IBM SPSS Statistics 語法指令。您可以開啟語法視窗，並直接輸入指令，但是讓對話框為您執行部分或所有的作業，會比較容易。

本章範例使用資料檔 *demo.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

附註：學生版中無法使用指令語法。

貼上語法

建立語法最簡單的方式，就是使用大部分對話框中的「貼上」按鈕。

1. 開啟資料檔 *demo.sav*。如需相關資訊，請參閱主題 第 73 頁的『第 10 章 範例檔案』。
2. 在功能表上，選擇：
 分析 > 敘述性統計量 > 頻次...
3. 選取「婚姻狀態 [*marital*]，並將它移到「變數」清單中。
4. 按一下「圖表」。
5. 在「圖表」對話框中，選擇「長條圖」。
6. 在「圖表值」組別中，選取「百分比」。
7. 按一下「繼續」。按一下「貼上」，將建立為對話框選擇結果的語法，複製到「語法編輯器」中。

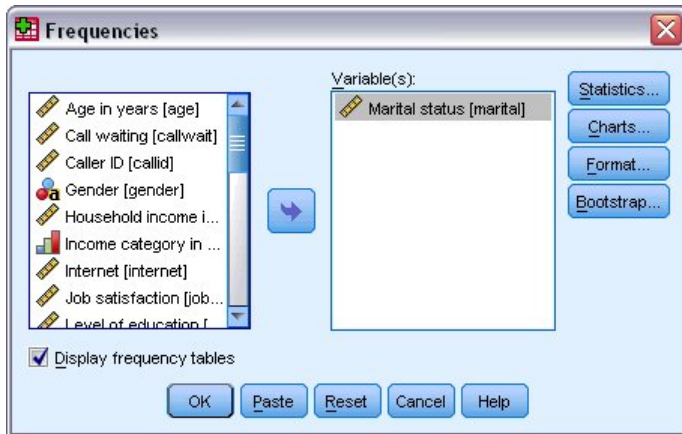


圖 69. 「頻次」對話框

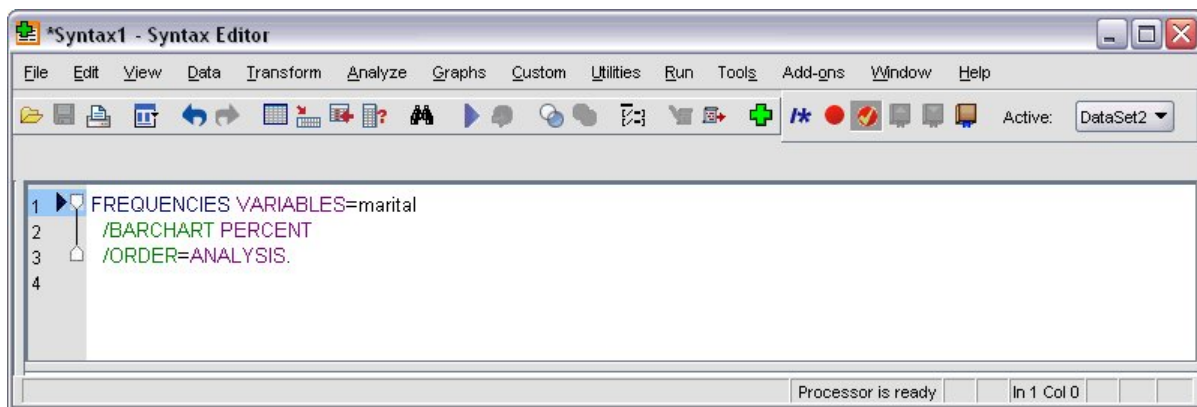


圖 70. 頻次語法

8. 若要執行目前顯示的語法，請從功能表選擇：

執行 > 選擇

編輯語法

您可以在語法視窗中編輯語法。例如，您可以變更次指令 `/BARCHART`，以顯示頻次而不是百分比。（以斜線來表示次指令。）如果您知道應使用哪一個關鍵字以顯示頻次，您也可以直接輸入此關鍵字。如果您不知道應該使用哪一個關鍵字，您可以將滑鼠游標置於任何次指令名稱處，按下 `Ctrl + 空白鍵`，就可以取得可用的關鍵字清單。這會顯示次指令的自動完成控制。

從 `BARCHART` 次指令刪除關鍵字 `PERCENT`。

按 `Ctrl` 鍵加空格。

1. 按一下頻次中標記為 **FREQ** 的項目。在自動完成控制中按一下項目，將會在現行游標位置插入該項目。

依預設，在您進行輸入時，自動完成控制會顯示可用的項目清單來提示您。例如，您也許希望將圓餅圖和長條圖一同併入。這個圓餅圖是由另一個次指令所指定。

2. 請在 **FREQ** 關鍵字後輸入一個正斜線以指示次指令的開頭處，並按下 `Enter`。

「語法編輯器」會顯示目前指令的次指令清單來提示您。

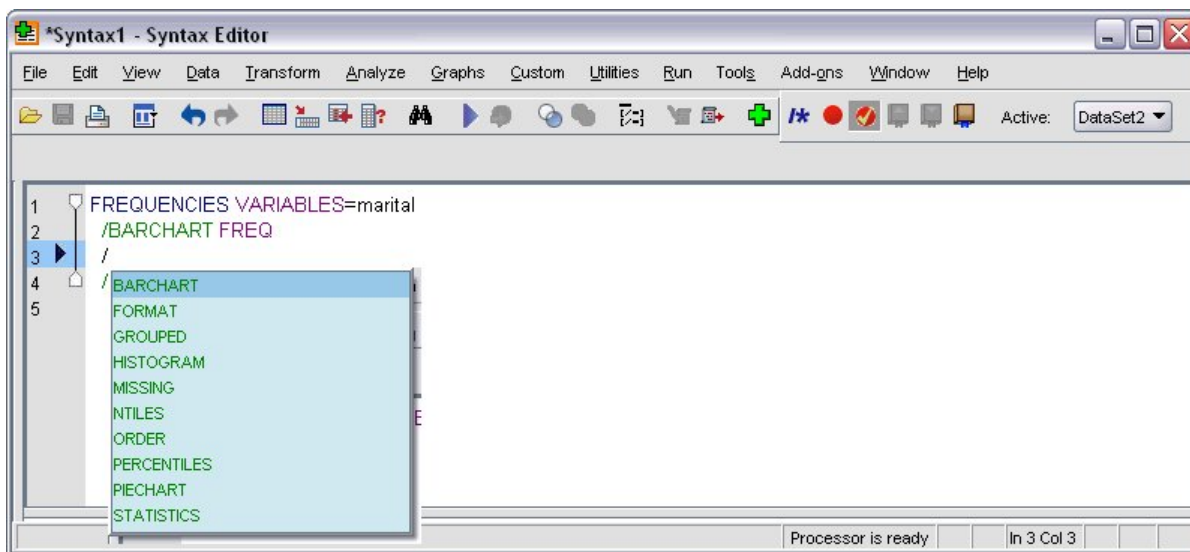


圖 71. 自動完成控制可顯示次指令

若想要取得更多有關現有指令的詳細說明，請按 `F1` 按鍵。這會直接顯示目前指令的指令語法參考資訊。

您可能已經注意到了，語法視窗中顯示的文字是有顏色的。彩色編碼可讓您快速辨認出未辨識的項目，因為只有已辨識的項目才有顏色。例如，您將 FORMAT 次指令誤拼為 FRMAT。次指令依預設會變成綠色，但 FRMAT 這個字則顯示未著色，因為這個字無法被辨識。

開啟與執行語法檔

1. 若要開啟儲存的語法檔，請從功能表選擇：

檔案 > 開啟 > 語法...

顯示用於開啟檔案的標準對話框。

2. 選擇語法檔。如果沒有顯示語法檔，請確定已選取「語法 (*.sps)」為您要檢視的檔案類型。

3. 按一下「開啟」。

4. 使用語法視窗中的「執行」功能表執行指令。

如果將指令套用至特定的資料檔，則在執行指令之前必須先開啟該資料檔，或者包含可開啟該資料檔的指令。您可以從開啟資料檔的對話框中貼上這類指令。

使用中斷點

中斷點可讓您在語法中的指定點停止執行指令語法，並在就緒時繼續執行。這樣可讓您在語法工作的中間點檢視輸出或資料，或是執行可顯示有關資料目前狀態的語法，例如 FREQUENCIES。您只能在指令階層中設定中斷點，在指令中的特定行處是無法進行這種設定的。

若要在指令中插入中斷點：

1. 在與指令關連之文字的左邊範圍中任意點一下。

中斷點在指令文字左邊的範圍中會以紅色圓圈表示，不論您點選的位置在何處，它都會和指令名稱位於同一行。

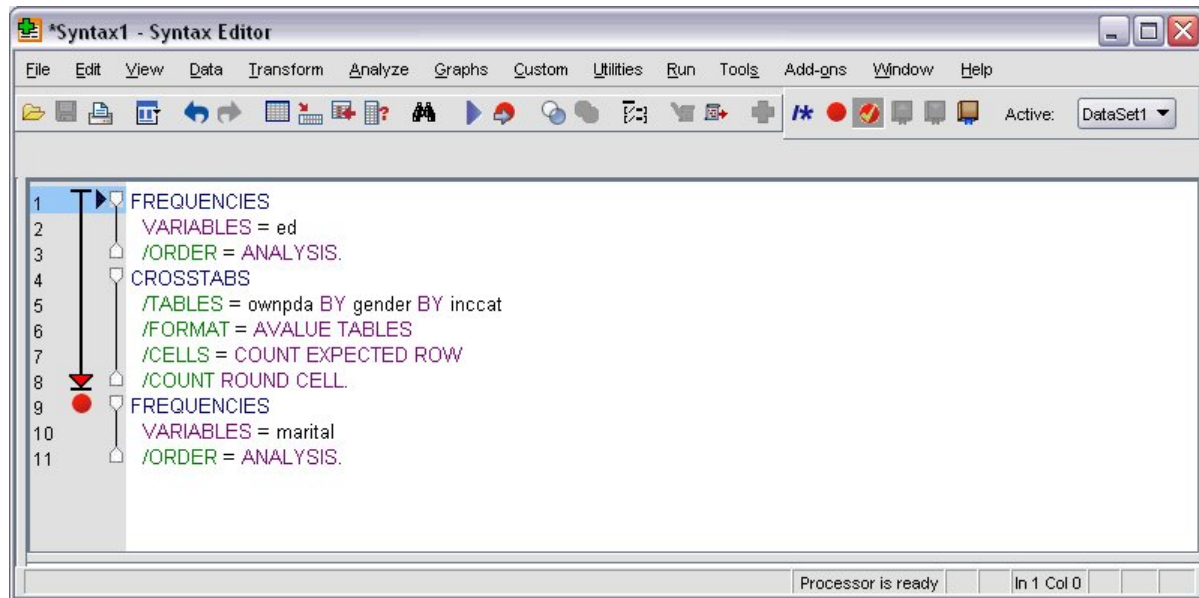


圖 72. 執行在中斷點處中止

執行包含中斷點的指令語法時，執行會在包含中斷點的指令前中止。

指令文字左邊的向下箭頭會顯示語法執行的進度。它會畫出從第一個指令執行到最後一個指令執行的範圍，在執行包含中斷點的指令語法時特別有幫助。

若要繼續位於中斷點後的執行：

2. 從「語法編輯器」中的功能表選擇：

執行 > 繼續

第 8 章 修改資料值

您剛開始使用的資料，可能並非總是依照對分析或報告需求最有利的方式來編排。例如，您可能希望：

- 從尺度變數建立類別變數。
- 將數個回應類別合併成一個類別。
- 建立一新變數，以計算兩個現有變數之間的差異。
- 計算兩個日期之間的時間長度。

本章使用資料檔 *demo.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

從尺度變數建立類別變數

資料檔 *demo.sav* 中的數個類別變數，事實上是該資料檔所衍生的尺度變數。例如，變數 *inccat* 只是分組成四種類別的 *income*。此類別變數使用整數 1 到 4 來代表下列收入類別（以千元為單位）：\$25、\$25–\$49、\$50–\$74 和 \$75 或以上。

若要建立類別變數 *inccat*：

1. 從「資料編輯器」視窗中的功能表選擇：

轉換 > Visual Binning...

在初始的 Visual Binning 對話框中，選擇您要為其建立新 bin 變數的尺度和（或）序數變數。**Binning** 是指將兩個或多個連續數值群組成相同的類別。

因為 Visual Binning 需依靠資料檔中的實際值來幫助您做出正確的 bin 選擇，所以它需要先讀取資料。如果您的資料檔中含有大量的觀察值個數，讀取資料可能需要一點時間，並且此初始對話框也能讓您限制讀取（「掃描」）的觀察值個數。我們的樣本資料檔並不需要進行限制。雖然此檔案含有 6,000 個以上的觀察值，但掃描此數目的觀察值並不需要太多時間。

2. 將「變數」清單中的「家庭收入（千元）[*income*]」拖放到「變數至 Bin」清單中，然後按一下「繼續」。

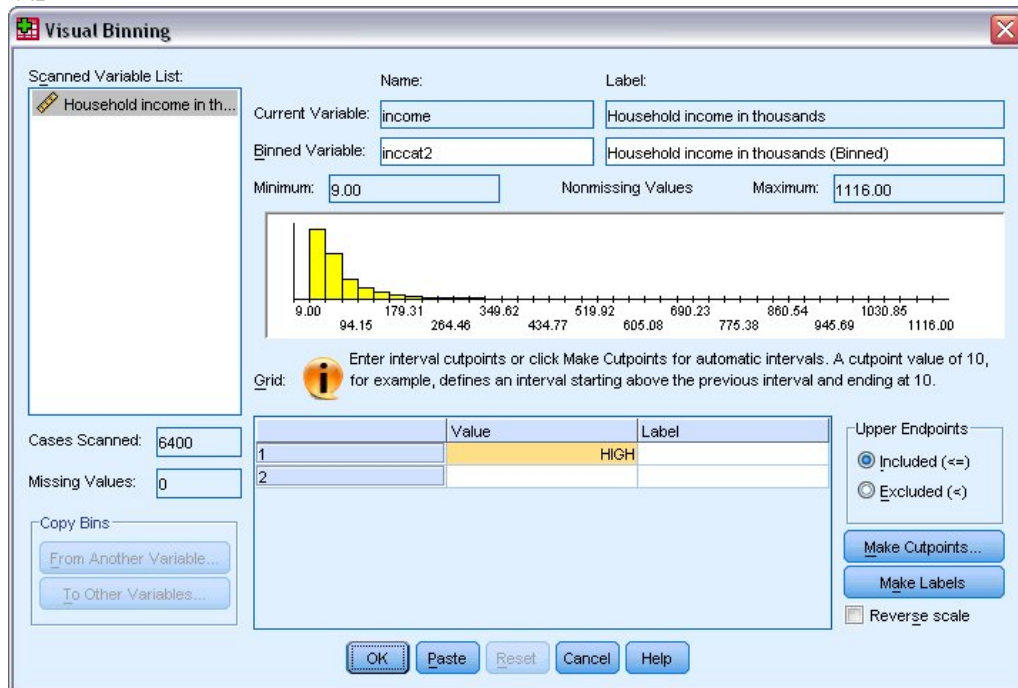


圖 73. 主要視覺化 Binning 對話框

3. 在主要視覺化 Binning 對話框中，選擇「掃描變數」清單中的「家庭收入（千元）[*income*]」。

直方圖顯示選擇變數的分配（此觀察值中的分配嚴重歪斜）。

4. 輸入 `inccat2` 作為新的 bin 變數名稱，然後輸入收入類別 [千元] 作為變數標籤。
5. 按一下**製作分割點**。
6. 選取**等寬間隔**。
7. 輸入 25 作為第一個分割點位置，輸入 3 作為分割點的數目，然後輸入 25 作為寬度。

bin 類別的數目為分割點數目加一。因此在這個範例中，新的 bin 變數將擁有四種類別，前三個類別各包含 25（千元）的範圍，而最後一個則包含最高分割點數值 75（千元）以上的所有數值。

8. 按一下「**套用**」。

現在網格中所顯示的數值代表已定義的分割點，為各類別的上端點。直方圖中的垂直線亦代表分割點的位置。

依預設，這些分割點的數值包含在對應的類別中。例如，25 的第一個數值應包含小於或等於 25 的所有數值。但在此範例中，我們要的是對應於 25 以下、25 到 49、50 到 74 和 75 或以上的類別。

9. 在「上限點」組別中，選取**排除 (<)**。
10. 然後按一下**製作標記**。

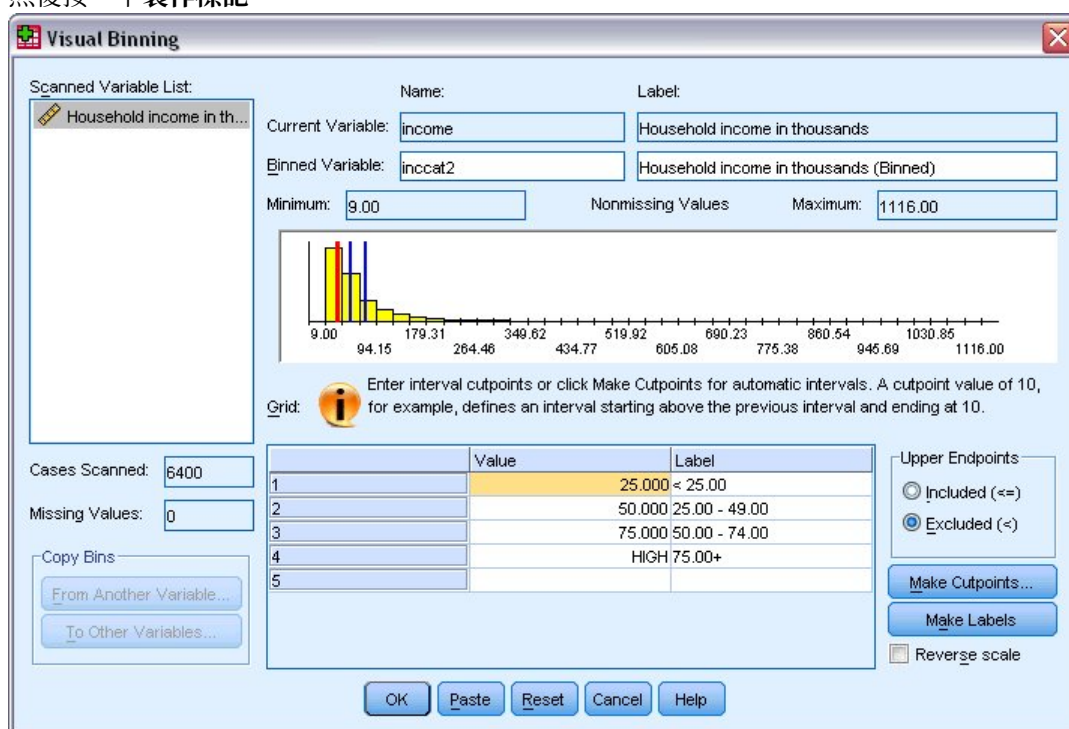


圖 74. 自動產生值標籤

這會自動為每個類別產生描述性的值標籤。因為指定給新 bin 變數的實際數值，不過是從 1 開始的整數，因此值標籤非常實用。

您也可以手動輸入或變更網格中的分割點和標記、拖放直方圖中的分割點線條來變更分割點位置，以及將分割點線條拖曳到直方圖以外的地方以刪除分割點。

11. 按一下「**確定**」以建立新的 bin 變數。

「資料編輯器」中會顯示新的變數。因為變數是加到檔案的結尾處，因此新變數會顯示在「資料視圖」中最右邊一欄，以及「變數視圖」中的最後一列。

計算新變數

您可使用多種數學函數，根據高度複雜的等式計算新變數。但在此範例中，我們只計算代表兩現有變數值之間差異的新變數。

資料檔 *demo.sav* 包含兩變數，一個代表應答者目前的年齡，另一個代表任職目前工作的年數。但如果沒有，則包含代表應答者在開始該工作時年齡的變數。我們可以建立新變數，計算目前年齡與任職目前工作年數之間的差異，這應該是應答者開始該工作時的大約年齡。

1. 從「資料編輯器」視窗中的功能表選擇：

轉換 > 計算變數...

2. 在「目標變數」中輸入 *jobstart*。

3. 在來源變數清單中選擇「年齡 [*age*]」，並按一下箭頭按鈕將之複製到「數值表示式」文字方框中。

4. 在對話框中按一下計算機面板上的減號 (-) 按鈕（或按鍵盤上的減號鍵）。

5. 選擇「現職年數 [*employ*]」，並按一下箭頭按鈕將之複製到表示式。

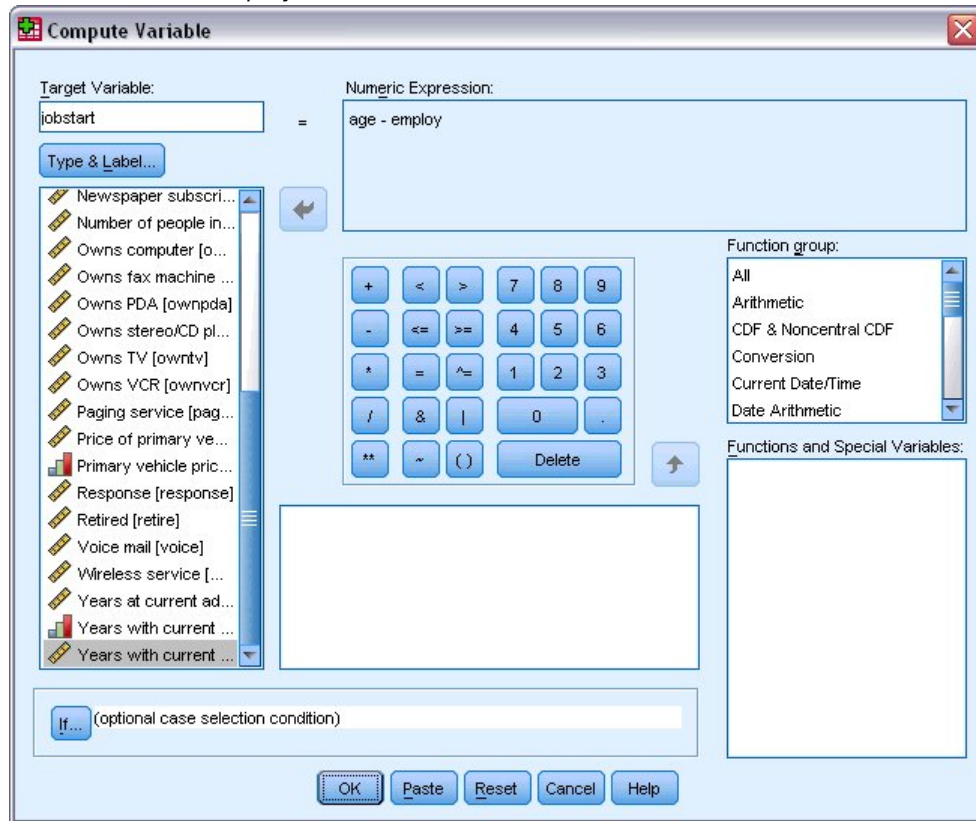


圖 75. 計算變數對話框

附註：請小心選擇正確的職業變數。變數也有重新編碼的類別版本，但這不是您所需要的。數值表示式應該是 *age-employ*，而不是 *age-empcat*。

6. 按一下**確定**計算新變數。

「資料編輯器」中會顯示新的變數。因為變數是加到檔案的結尾處，因此新變數會顯示在「資料視圖」中最右邊一欄，以及「變數視圖」中的最後一列。

使用表示式中的函數

您也可以使用表示式中預先定義的函數。有 70 種以上的內建函數可用，包括：

- 算術函數
- 統計函數
- 分配函數
- 邏輯函數
- 日期和時間聚集與擷取函數
- 遺漏值函數

- 交叉觀察值函數
- 字串函數

函數會整理為依邏輯區分的群組，如算數運算或其他計算統計多項式。為方便起見，一些常用的系統變數，如 $\$TIME$ （目前日期與時間）亦包含於適當的函數群組中。

將函數貼到表示式中

若要將函數貼到表示式中：

1. 將游標放在表示式中您要顯示函數的位置。
2. 從「函數」群組清單中選擇適當群組。標籤為「全部」的函數群組會列出所有可用函數與系統變數。
3. 按兩下「函數與特殊變數」清單中的函數（或選擇函數並按一下「函數」群組清單旁的箭頭）。

函數會插入表示式中。如果您將部分表示式反白，然後插入函數，則表示式的反白部分將作為函數中的第一個引數。

編輯表示式中的函數

您要輸入引數後函數才算完整，而所貼上的函數中以問號來表示引數。問號的數目表示完成函數所需的最小引數數目。

1. 將所貼上之函數中的問號反白。
2. 輸入引數。如果引數是變數名稱，則您可以用變數清單來貼上其名稱。

使用條件表示式

您可以使用條件表示式（也稱為邏輯表示式）指定選定觀察值子集的轉換方式。條件表示式可將各觀察值之值（true、false 或 missing），傳回給系統。如果條件表示式的結果為 true，則轉換會套用至該觀察值。如果結果為 false 或 missing，則轉換將不會套用至該觀察值。

若要指定條件表示式：

1. 在「計算變數」對話框中，按一下 **If**。如此就會開啟「觀察值選擇條件」對話框。
2. 選取「包含滿足條件的觀察值」。
3. 輸入條件表示式。

大部分的條件表示式至少會包含一個相關運算子，如：

$age \geq 21$

或

$income * 3 < 100$

在第一個範例中，只選擇年齡 $[age]$ 值在 21 或以上的觀察值。在第二個範例中，觀察值的家庭收入（千元） $[income]$ 乘以 3 必須小於 100 才會被選取。

您也可以使用邏輯運算子來連結兩個或多個條件表示式，如：

$age \geq 21 \mid ed \geq 4$

或

$income * 3 < 100 \ \& \ ed = 5$

在第一個範例中，符合年齡 $[age]$ 條件或教育層級 $[ed]$ 條件的觀察值呈現已選。在第二個範例中，觀察值必須同時符合家庭收入（千元） $[income]$ 和教育層級 $[ed]$ 的條件才會被選取。

日期與時間的操作

許多與日期和時間有關的常用工作可使用「日期與時間」精靈執行。使用此精靈，您可：

- 從包含日期或時間的字串變數中建立日期/時間變數。
- 結合日期或時間變數的不同部分，建構日期/時間變數。

· 新增或刪減日期/時間變數的數值，包括新增或刪減兩個日期/時間變數。
· 擷取日期或時間變數的一部分；例如，從格式為 mm/dd/yyyy 的日期/時間變數中取得月中天數。
本節範例使用資料檔 *upgrade.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。
若要使用「日期和時間精靈」：

1. 在功能表上，選擇：

轉換 > 日期和時間精靈...



圖 76. 「日期和時間精靈」簡介畫面

「日期與時間精靈」的簡介畫面為您呈現一組常用任務。無法套用至目前日期的任務會被停用。例如，資料檔 *upgrade.sav* 未包含任何字串變數，因此從字串建立資料變數的任務將會停用。

如果您是初次使用 IBM SPSS Statistics 的日期與時間，可選取「了解日期和時間的表示方法」並按一下「下一步」。會前往一畫面，提供日期/時間變數的簡要說明，並有一連結，可透過「說明」按鈕存取更詳細的資訊。

計算兩個日期之間的時間長度

最常用的任務之一包含計算兩日期之間的日數長度。舉例而言，假設軟體公司希望藉從客戶上次購買更新至今的年數，分析更新授權的購買情形。資料檔 *upgrade.sav* 包含各客戶上次購買更新日期的變數，但沒有購買至今的年數。就可利用一時間長度新變數來計算上次購買更新到下次發行產品的年數，以提供此量的測量值。

若要計算兩個日期之間的時間長度：

1. 在「日期與時間精靈」介紹畫面上選擇「以日期和時間計算」，並按一下「下一步」。
2. 選擇「計算介於兩個日期之間的時間單位數」並按一下「下一步」。

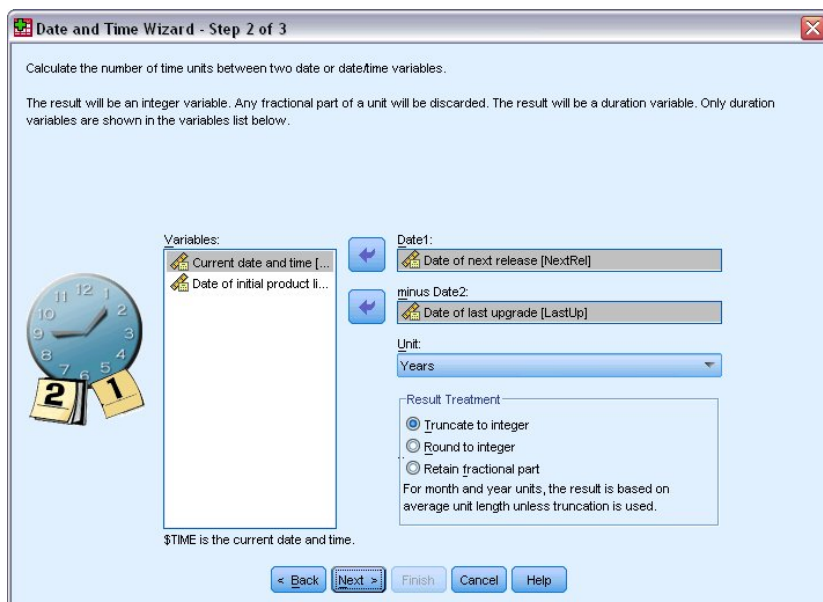


圖 77. 計算兩個日期之間的時間長度：步驟 2

3. 在步驟 2 中，將下次發佈日期選作日期 1。
4. 將 *Date of last upgrade* 選為日期 2。
5. 選取「年數」作為「單位」，選取「截斷為整數」作為「結果處理」。（這些為預設選項。）
6. 按一下「下一步」。
7. 在步驟 3 中，輸入 *YearsLastUp* 作為結果變數的名稱。結果變數不得與現有變數具備相同名稱。
8. 輸入「上次更新至今年數」作為結果變數的標籤。結果變數的變數標籤是選用的。
9. 保留「立即建立變數」的預設選項，並按一下「完成」建立新變數。

在「資料編輯器」中顯示的新變數 *YearsLastUp*，為兩日期間的整數年數。年數的分數部分會被截斷。

在日期上新增期間

您可在日期上新增或刪減期間，如 10 天或 12 個月。延續上節軟體公司的範例，假設該公司要計算各客戶初始技術支援合約結束的日期。資料檔 *upgrade.sav* 包含支援合約的年數變數，以及初始購買日期變數。您可將支援年數增加至購買日期，決定支援結束的日期。

若要在日期上新增期間：

1. 在「日期與時間精靈」介紹畫面上選擇「以日期和時間計算」，並按一下「下一步」。
2. 在第 1 步中，選取「從日期加上或減去持續時間」，並按「下一步」。

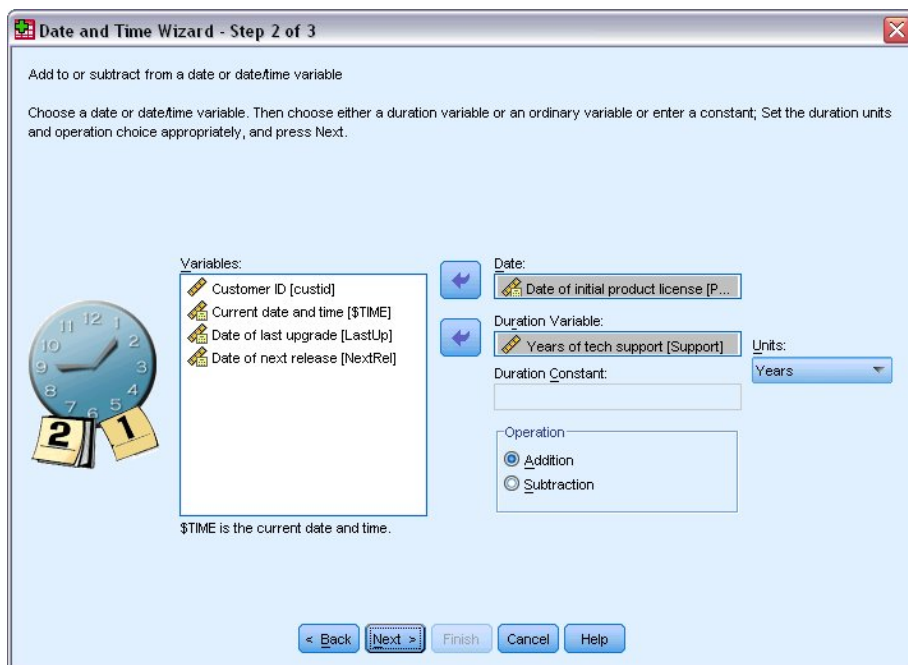


圖 78. 在日期上新增持續時間：步驟 2

3. 將起始產品授權日期 選為日期。
 4. 在步驟 2 中，選取技術支援年限 作為「持續時間變數」。
- 由於技術支援年份 只是數值變數，您必須指定新增此期間時要使用的單位。
5. 從「單位」下拉清單中選擇「年數」。
 6. 按一下「下一步」。
 7. 在步驟 3 中，輸入 *SupEndDate* 作為結果變數的名稱。結果變數不得與現有變數具備相同名稱。
 8. 輸入支援結束日期作為結果變數標籤。結果變數的變數標籤是選用的。
 9. 按一下「完成」建立新變數。
- 「資料編輯器」中會顯示新的變數。

第 9 章 排序與選取資料

資料檔的組成方式，無法永遠符合您的需求。若要準備資料以供分析，您可從多種轉換功能中選取，包括：

- **資料排序**。您可以依一個或多個變數，將觀察值排序。
- **選取觀察值子集**。您可以只分析某個觀察值子集，也可以同時分析多個不同的子集。

本章範例使用資料檔 *demo.sav*。如需相關資訊，請參閱主題第 73 頁的『第 10 章 範例檔案』。

排序資料

排序觀察值 (排序資料檔的列) 通常非常實用，且對特定分析類別而言是必須的。

若要根據一或多個排序變數值，重新排列資料檔中觀察值的順序：

1. 在功能表上，選擇：

資料 > 排序觀察值...

顯示「排序觀察值」對話框。

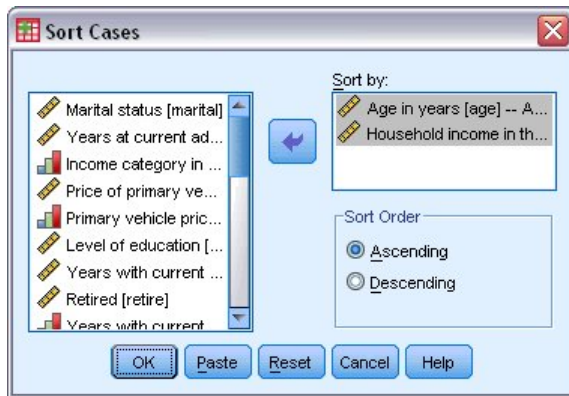


圖 79. 觀察值排序對話框

2. 將「年齡 (年為單位) [age]」和「家庭收入 (千元) [income]」變數新增至「依...排序」清單。

若您選取多個排序變數，會以出現於「依...排序」清單中的順序決定要排序的觀察值。在此範例中，根據「依...排序」清單的項目，會先根據「年齡 (年為單位) [age]」類別中的「家庭收入 (千元) [income]」數值排序觀察值。對於字串變數，大寫字母的排序順序會在對等的小寫字母之前 (例如，字串值 *Yes* 在排序清單中的順序會在 *yes* 之前)。

分割檔案處理

若要將您的資料檔分割為不同群組以進行分析：

1. 在功能表上，選擇：

資料 > 分割檔案...

會顯示「分割檔案」對話框。

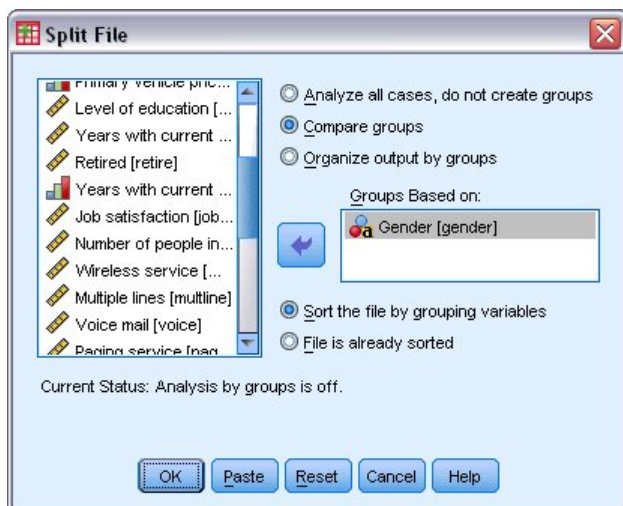


圖 80. 分割檔案對話框

2. 選取「比較組別」或「依組別安排輸出」。(這些步驟下的範例顯示兩種選項間的差異。)
3. 選取「性別 [gender]」以根據這些變數將檔案分割為不同群組。

您可使用數值、短字串和長字串變數，作為分組變數。會針對分組變數所定義的各個子群組執行個別分析。若您選取多個分組變數，會以出現於「以組別為準」清單中的順序決定觀察值分組的方式。

若您選取「比較組別」，所有分割檔案組別的結果都會包含於相同表格，如「頻次」程序所產生的以下摘要統計量表格。

Statistics

Household income in thousands

Female	N	Valid	3179
		Missing	0
	Mean		68.7798
	Median		44.0000
	Std. Deviation		75.73510
Male	N	Valid	3221
		Missing	0
	Mean		70.1608
	Median		45.0000
	Std. Deviation		81.56216

圖 81. 以單一樞紐表輸出分割檔案

若您選取「依組別安排輸出」並執行「頻次」程序，會建立兩個樞紐表：女性樞紐表和男性樞紐表。

Statistics^a

Household income in thousands

N	Valid	3179
	Missing	0
Mean		68.7798
Median		44.0000
Std. Deviation		75.73510

a. Gender = Female

圖 82. 以女性樞紐表輸出分割檔案

Statistics^a

Household income in thousands

N	Valid	3221
	Missing	0
Mean		70.1608
Median		45.0000
Std. Deviation		81.56216

a. Gender = Male

圖 83. 以男性樞紐表輸出分割檔案

排序觀察值以進行分割檔案處理

「分割檔案」程序會在每次遇到分組變數的不同值時，建立一個新的子群組。因此，請務必根據分組變數的值排序觀察值，再呼叫分割檔案處理。

依照預設，「分割檔案」會根據分組變數的值，自動排序資料檔。若檔案已排序為正確順序，您可選取「檔案已排序」，以節省處理時間。

開啟與關閉分割檔案處理

呼叫分割檔案處理後，除非您關閉，否則在剩餘作業階段皆有效。

- 分析所有觀察值。此選項會關閉分割檔案處理。
- 「比較組別」和「依組別整理輸出」。此選項會開啟分割檔案處理。

若分割檔案處理為有效狀態，應用程式視窗下方的狀態列會出現「分割檔案開啟」訊息。

選擇觀察值子集。

您可根據包含變數和複雜表示式的條件，將您的分析限制於特定子群組。或者，您也可以選擇隨機的觀察值樣本。定義子群組所需之準則包括：

- 變數值和範圍
- 日期和時間範圍
- 觀察值（列）個數
- 算術表示式
- 邏輯表示式
- 函數

若要選取觀察值子集以進行分析：

1. 在功能表上，選擇：

資料 > 選取觀察值...

這樣會開啟「選擇觀察值」對話框。

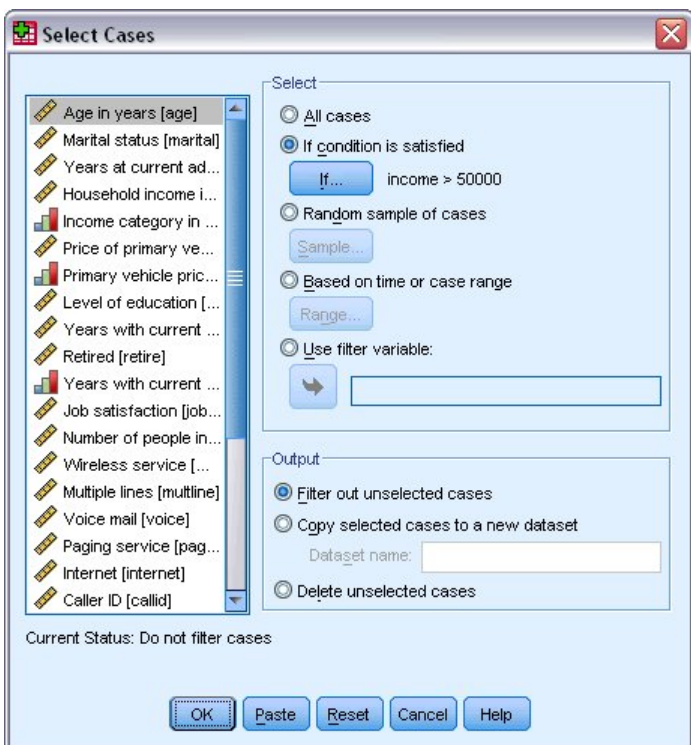


圖 84. 「選擇觀察值」對話框

選取以條件表示式為基礎的觀察值子集

若要根據條件表示式選擇觀察值：

1. 在「選擇觀察值」對話框中，選取「如果滿足設定條件」，然後按一下「若」。這樣會開啟「選取觀察值條件」對話框。

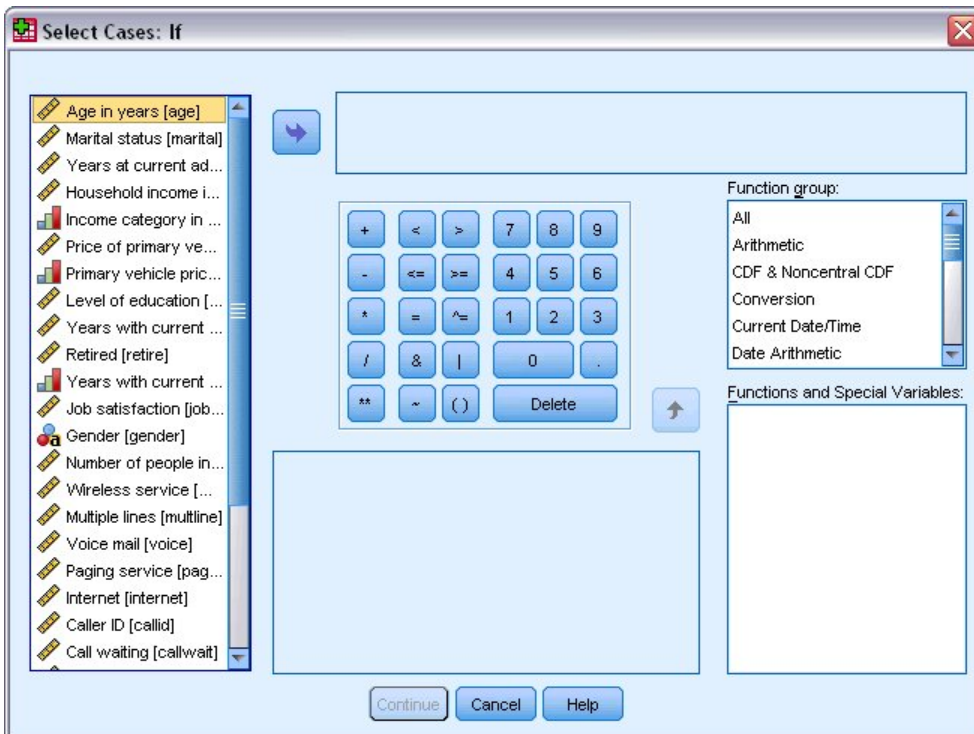


圖 85. 選取觀察值條件對話框

條件表示式可使用現有變數名稱、常數、算術運算子、邏輯運算子、關係運算子和函數。您可在文字框中，以在輸出視窗中處理文字的方式，輸入與編輯表示式。您亦可使用計算機面板、變數清單和函數清單，將元素貼上至表示式。如需相關資訊，請參閱主題第 62 頁的『使用條件表示式』。

選取隨機樣本

若要取得隨機樣本：

1. 在「選擇觀察值」對話框中選取「**觀察值的隨機樣本**」。
2. 按一下「**樣本**」。

這樣會開啟「選擇觀察值隨機樣本」對話框。

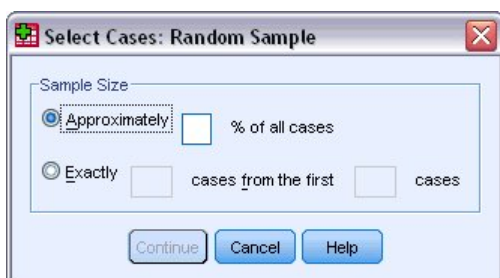


圖 86. 選擇觀察值隨機樣本對話框

您可選擇下列替代選項之一，作為樣本大小：

- **近似於**。使用者指定的百分比。此選項會產生近似指定觀察值百分比的隨機樣本。
- **恰好**。除了使用者指定的觀察值個數之外。您還必須指定產生樣本的觀察值個數。而第二個數字，應該小於或等於資料檔中的觀察值總個數。如果這個數字超過資料檔中的觀察值總個數，則樣本中所含的觀察值個數，就會按照比例，逐漸地低於指定個數。

選取時間範圍或觀察值範圍

若要根據日期、時間或觀察值 (列) 編號選取觀察值範圍：

1. 在「選擇觀察值」對話框中選取「**以時間或觀察值範圍為準**」，並按一下「**範圍**」。

這樣會開啟「選擇觀察值範圍」對話框，讓您從中選取觀察值 (列) 編號範圍。

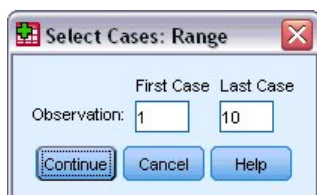


圖 87. 「選擇觀察值範圍」對話框

- **第一個觀察值**。輸入範圍的起始日期和/或時間值。若未定義日期變數，輸入起始觀察值編號 (除非分割檔案模式開啟，否則為「資料編輯器」中的列號)。若您未指定最後一個觀察值數值，會選取起始日期/時間到結束時間序列中的所有觀察值。
- **最後一個觀察值**。輸入範圍的結束日期和/或時間值。若未定義日期變數，輸入結束觀察值編號 (除非分割檔案模式開啟，否則為「資料編輯器」中的列號)。若您未指定第一個觀察值數值，會選取時間序列開始到結束日期/時間中的所有觀察值。

對於已定義日期變數的時間序列資料，您可根據定義的日期變數，選取日期和/或時間的範圍。各個觀察值都代表不同時間的觀察值，檔案會依時間順序排序。



圖 88. 「選擇觀察值範圍」對話框 (時間序列)

若要產生時間序列資料的日期變數：

2. 在功能表上，選擇：

資料 > 定義日期...

未選擇之觀察值的處理方式

您可以任選下列其中一個選項作為未選擇之觀察值的處理方式：

- **過濾掉未選取的觀察值。** 未選擇的觀察值不列入分析但保留於資料集內。如果您關掉過濾功能，您可稍後於作業階段中使用未選擇的觀察值。如果您選擇一個隨機樣本，或如果您以條件表示式為基礎選擇觀察值，將產生名為 *filter_\$* 的變數，其包含數值 1 的選擇觀察值及數值 0 的未選擇觀察值。
- **複製已選擇觀察值至新資料集。** 選擇的觀察值被複製到新的資料集，不影響原始資料集。未選擇的觀察值不列入新的資料集，並以原始狀態留置於原始資料集中。
- **過濾掉未選取的觀察值。** 自資料集刪除未選擇的觀察值。如果要復原已刪除的觀察值，必須離開檔案且不儲存任何變更，再重新開啟檔案。如果您將變更儲存於資料檔，觀察值便永久刪除。

附註：如果您刪除未選擇的觀察值並儲存檔案，將無法復原觀察值。

觀察值選取狀態

如果您已選擇觀察值子集，但仍未捨棄未選擇的觀察值，那麼在「資料編輯器」中，那些未選擇觀察值的列號，會被標上一條對角線。

1
2
3
4
5
6
7
8
9
10
11

圖 89. 觀察值選取狀態

第 10 章 範例檔案

與產品同時安裝的範例檔存放在安裝目錄的範例子目錄中。在下列每種語言的「範例」子目錄中存有個別資料夾：英文、法文、德文、義大利文、日文、韓文、波蘭文、俄文、簡體中文、西班牙文和繁體中文。

並非所有範例檔案皆提供各種語言。如果範例檔案沒提供您需要的語言，語言資料夾有英文版的範例檔案說明。

以下是使用於本檔案中不同範例的範例檔之簡要說明。

- **accidents.sav**。這是有關某保險公司研究年齡和性別風險因素對給定地區汽車意外事件的假設資料檔。每一個觀察值對應至一個年齡類別和性別的交叉分類。
- **adl.sav**。這是有關致力於確定一個建議中風病患治療類型之效益的假設資料檔。醫師隨機指定女性中風病患至兩個組別之一。第一組接受標準的物理治療，而第二組則接受額外的情緒治療。在治療了三個月後，將每一個病患進行日常活動的能力記分為序數變數。
- **advert.sav**。這是有關一家零售商致力於調查廣告費與廣告後銷售情形之間的關係的假設資料檔。為了這個目的，他們收集了過往銷售數字和相關的廣告費用。
- **aflatoxin.sav**。這是有關檢定玉米作物是否有黃麴毒素（一種毒物，其濃度在介於和處於作物產量中都有很大的差異）的假設資料檔。一名穀物加工者收到來自 8 個作物產量各 16 個樣本，並以十億當量 (PPB) 來測量黃麴毒素的層級。
- **anorectic.sav**。在研究厭食/暴食行為的標準症狀參照時，研究人員¹對 55 名已知存在進食障礙的青少年進行了調查。每個病患在四年之中被訪問四個回合，所以得到總數為 220 的觀察值。在每次觀察中，為病患在 16 種症狀上逐一評分。目前遺漏了第二次訪察的病患 71，第二次訪察的病患 76，以及第三次訪察的病患 47 的症狀分數，因此只剩下 217 個有效觀察值。
- **bankloan.sav**。這是有關一家銀行致力於減少放款利率預設值的假設資料檔。本檔包含 850 位以前的客戶與現在的準客戶的財務和人口資料。前 700 個觀察值為以前有借貸的客戶。最後 150 個觀察值是銀行需要作信用風險優良與不良分類的準客戶。
- **bankloan_binning.sav**。這是包含 500 位以前客戶的財務和人口資料的假設資料檔。
- **behavior.sav**。在一個經典範例²中，52 名學生被要求以 10 點的標準對 15 種狀況和 15 種行為的組合進行評價，該 10 點的標準從 0 = 「非常合適」到 9 = 「非常不合適」。平均值超過個別值，值會被視為相異性。
- **behavior_ini.sav**。本資料檔包含 *behavior.sav* 之二維解的起始配置。
- **brakes.sav**。這是有關一間生產高效能汽車碟型煞車片工廠中品質管制的假設資料檔。資料檔包含由 8 個生產機器分別取得 16 個碟片的直徑測量。煞車的目標直徑是 322 公釐。
- **breakfast.sav**。在一項經典研究³中，21 名 Wharton School MBA 學生及其配偶被要求按照喜好程度順序對 15 種早餐食品進行評價，從 1 = 「最喜歡」到 15 = 「最不喜歡」。他們的喜愛程度分六種不同情況記錄，從「整體喜愛」到「點心，僅配飲料」。
- **breakfast-overall.sav**。本資料檔只包含第一種情況—「整體喜愛」—所喜愛的早餐項目。
- **broadband_1.sav**。這是包含全國性寬頻服務地區用戶數目的假設資料檔。本資料檔包含四年期間 85 個地區每月的用戶數目。
- **broadband_2.sav**。本資料檔與 *broadband_1.sav* 相同，但多了三個月的資料。
- **car_insurance_claims.sav**。在別處被提出和分析的⁴關於汽車損壞賠償的資料集。理賠金額的平均數可建模為具有伽瑪分佈，使用反鏈結函數將應變數的平均數相關至一被保險人年齡、車輛類型和車齡的線性組合。提出理賠的數量可以用作尺度加權。

¹ Van der Ham, T., J. J. Meulman, D. C. Van Strien, and H. Van Engeland. 1997. Empirically based subgrouping of eating disorders in adolescents: A longitudinal perspective. *British Journal of Psychiatry*, 170, 363-368.

² Price, R. H., and D. L. Bouffard. 1974. Behavioral appropriateness and situational constraints as dimensions of social behavior. *Journal of Personality and Social Psychology*, 30, 579-586.

³ Green, P. E., and V. Rao. 1972. *Applied multidimensional scaling*. Hinsdale, Ill.: Dryden Press.

⁴ McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

- **car_sales.sav**。本資料檔包含假設性的銷售估計、定價、和不同的品牌與車輛型式的實體規格。定價和實體規格是由 *edmunds.com* 和製造商處輪流取得。
- **car_sales_uprepared.sav**。這是 *car_sales.sav* 的修改版本，其中不包含任何欄位的轉換版本。
- **carpet.sav**。在常用範例⁵中，有一家公司想要銷售一種新型的地毯清潔用品。該公司希望調查以下會對消費者偏好產生影響的五個因素：包裝設計、品牌名稱、價格、優質家用品標誌、以及退貨保證。包裝設計有三個因素層級，每個層級中的清潔刷位置都不相同；三個品牌名稱（*K2R*、*Glory*、及 *Bissell*）；三個價格層級；且最後兩個因素各有兩個層級（無論無或有）。十名消費者將這些因素所定義的 22 種組合分級。「偏好」變數包含每個組合平均排名的等級。排名數值較小者會對應高偏好程度。這個變數反映每個組合偏好的整體量數。
- **carpet_prefs.sav**。本資料檔是根據 *carpet.sav* 所描述的可相同範例，但它包含 10 個消費者每一個人的實際等級。消費者被要求將 22 個產品組合從最喜歡排列到最不喜歡。變數 *PREF1* 到 *PREF22* 包含相關組合的 ID，如 *carpet_plan.sav* 中所定義。
- **catalog.sav**。本資料檔包含郵購公司銷售三項產品的每月假設銷售數字。也包含五個可能預測變數的資料。
- **catalog_seasfac.sav**。本資料檔與 *catalog.sav* 相同，不過多了一組由「週期性分解」程序所計算的週期性因素以及隨附的資料變數。
- **cellular.sav**。這是有關一家手機公司致力於減少客戶流失的假設資料檔。客戶流失傾向分數套用於帳戶，範圍由 0 至 100。帳戶分數 50 或以上有可能正尋求變更供應商。
- **ceramics.sav**。這是有關一家製造商致力於確定一種新的優良合金是否較標準的合金有較大的耐熱性的假設資料檔。每一個觀察值代表對合金之一的不同檢定；記錄了讓軸承失效的溫度。
- **cereal.sav**。這是有關對 880 人的早餐喜好進行訪談的假設資料檔，也記下他們的年齡、性別、婚姻狀態、和是否有活躍的生活型態（根據他們是否一週運動兩次）。每一個觀察值代表一位不同的應答者。
- **clothing_defects.sav**。這是有關一家服裝工廠品質管制過程的假設資料檔。由該工廠所生產的每一批產品中，檢查員取出一件服裝的樣本並計算不合格的服裝個數。
- **coffee.sav**。這是關於六種冰咖啡的認知品牌形象⁶的資料檔案。對 23 種冰咖啡中每一種的印象屬性，由群眾來選取依其屬性描述的所有品牌。該六種品牌已標示為 AA、BB、CC、DD、EE、和 FF，以保持機密。
- **contacts.sav**。這是有關一群公司電腦銷售代表聯絡清單的假設資料檔。每一個聯絡人依他們在公司所服務的部門及其公司的等級而分類。最後一次銷售的金額、到最後一次銷售的時間、和該聯絡人公司的規模也都被列入記錄。
- **creditpromo.sav**。這是有關一家百貨公司致力於評估近期信用卡促銷活動效果的假設資料檔。為達此目標，隨機選取了 500 位持卡人。有半數收到廣告，促銷在未來三個月購買將獲得降低利率的優惠。半數收到標準的週期性廣告。
- **customer_dbase.sav**。這是有關一家公司致力於使用其資料倉庫的資訊來對最有可能回應的客戶提供優惠的假設資料檔。隨機選取客戶庫的子集，提供優惠，再將他們的回應記錄下來。
- **customer_information.sav**。本檔案是包含客戶郵寄資訊的假設資料檔，例如姓名和地址。
- **customer_subset.sav**。80 個 *customer_dbase.sav* 的觀察值子集。
- **debate.sav**。這是有關一項政治辯論會參與者辯論前和辯論後接受調查之成對回應的假設資料檔。每一個觀察值對應至一位不同的應答者。
- **debate_aggregate.sav**。這是將 *debate.sav* 中之回應作整合的假設資料檔。每一個觀察值對應至辯論前和辯論後對偏好之交叉分類的反應。
- **demo.sav**。這是有關提供郵寄每月優惠之購買客戶資料庫的假設資料檔。記錄了客戶是否對該優惠回應，以及各種的人口資訊。
- **demo_cs_1.sav**。這是有關一家公司致力於匯編調查資訊資料庫之第一步的假設資料檔。每一個觀察值對應至一個不同的城市，也記錄了其地區、省、區、和城市識別。

⁵ Green, P. E., and Y. Wind. 1973. *Multiattribute decisions in marketing: A measurement approach*. Hinsdale, Ill.: Dryden Press.

⁶ Kennedy, R., C. Riquier, and B. Sharp. 1996. Practical applications of correspondence analysis to categorical data in market research. *Journal of Targeting, Measurement, and Analysis for Marketing*, 5, 56-70.

- **demo_cs_2.sav**。這是有關一家公司致力於匯編調查資訊資料庫之第二步的假設資料檔。每一個觀察值對應至在第一步中選取的城市中的一個不同的家庭單位，也記錄了其地區、省、區、分區、和單位識別。也納入了由該設計的前兩階段所得之取樣資訊。
- **demo_cs.sav**。這是包含以複合取樣設計所收集之調查資訊的假設資料檔。每一個觀察值對應至一個不同的家庭單位，也記錄了各種的人口和取樣資訊。
- **diabetes_costs.sav**。這是一個假設資料檔，包含保險公司維護之有關患有糖尿病的投保人的資訊。每一個觀察值對應於一個不同的投保人。
- **dietstudy.sav**。該假設資料檔案包含對 "Stillman diet"⁷ 的研究結果。每一個觀察值對應至一個不同的受試者，並記錄下他或她飲食法前、後之體重（磅）和三酸甘油酯層級（毫克/100 毫升）。
- **dmdata.sav**。這是一個假設資料檔，包含了直效行銷公司的統計人口資訊與購買資訊。**dmdata2.sav** 包含收到測試郵件的連絡人子集資訊，而 **dmdata3.sav** 則包含剩餘未收到測試郵件的連絡人資訊。
- **dvdplayer.sav**。這是有關新 DVD 播放器開發的假設資料檔。市場行銷團隊使用原型收集了焦點組別資料。每一個觀察值對應至不同調查到的使用者，並記錄下一些有關他們的人口資訊和他們對有關原型問題的回應。
- **german_credit.sav**。該資料檔案取自加州大學歐文分校的 Repository of Machine Learning Databases⁸ 中的 "German credit" 資料集。
- **grocery_1month.sav**。本假設資料檔是將 **grocery_coupons.sav** 資料檔和每週購買的「彙總」，因此每一個觀察值對應至一個不同的客戶。結果部分每週變更的變數消失了，而目前所記錄的銷售量是在研究的四週期間銷售量之總和。
- **grocery_coupons.sav**。這是包含某連鎖雜貨店想要知道他們客戶購買習慣所收集之調查資料的假設資料檔。每一個客戶被追蹤了四週，每一個觀察值對應至一個不同的客戶-週，並記錄有關客戶在何處及如何購物的資訊，包含那一週在雜貨店花了多少錢。
- **guttman.sav**。Bell⁹ 提供了一個表格，用來說明可能的社交群體。Guttman¹⁰ 使用此表的一部分，其中將此類事項描述為群體歸屬感、成員的親近感以及關係的形式化的 5 個變數與 7 個理論社交團隊交叉，包括人群（例如，參加足球比賽的人員）、觀眾（例如，劇院或教室的人）、公眾（例如，報紙或電視觀眾）、群體（例如，具有更強烈互動的觀眾）、主要群體（親密）、次要群體（誌願）和現代社群（由於緊密距離和專業服務需求產生的鬆散聯盟）。
- **health_funding.sav**。這是包含健保基金（每 100 人的金額）、疾病率（每 10,000 人的比率）、造訪健保機構的比例（每 10,000 人的比率）的假設資料檔。每一個觀察值代表一個不同的城市。
- **hivassay.sav**。這是有關一家製藥實驗室致力於開發一種偵測 HIV 感染快速檢驗的假設資料檔。檢驗結果是八個紅色加深的陰影，陰影愈深表示感染的可能性愈大。進行了一項實驗室的試驗，在 2,000 個血液樣本中，有半數遭到 HIV 的感染，而半數則未感染。
- **hourlywagedata.sav**。這是有關在辦公室和醫院任職的護士依經驗層級不同之鐘點費的假設資料檔。
- **insurance_claims.sav**。這是有關某保險公司希望建立標示可疑潛在詐欺理賠的模型的假設資料檔。每一個觀察值代表個不同的理賠。
- **insure.sav**。這是有關一家保險公司正在研究表示客戶是否必定理賠 10 年壽險合約之風險因素的假設資料檔。在資料檔中的每一個觀察值代表二份合約，其一記錄了理賠而另一則否，二者的年齡和性別相符。
- **judges.sav**。這是有關受過訓練的裁判（加上一位熱心人士）為 300 個體操表演評分的假設資料檔。每一列代表一個不同的表演；裁判們觀看相同的表演。
- **kinship_dat.sav**。Rosenberg 和 Kim¹¹ 開始分析 15 個親屬關係項（伯母、兄弟、表兄妹、女兒、父親、孫女、祖父、祖母、孫子、母親、侄子或外甥、姪女或外甥女、姐妹、兒子和叔叔）。他們請四組大學生

⁷ Rickman, R., N. Mitchell, J. Dingman, and J. E. Dalen. 1974. Changes in serum cholesterol during the Stillman Diet. *Journal of the American Medical Association*, 228:, 54-58.

⁸ Blake, C. L., and C. J. Merz. 1998. "UCI Repository of machine learning databases." Available at <http://www.ics.uci.edu/~mlearn/MLRepository.html>.

⁹ Bell, E. H. 1961. *Social foundations of human behavior: Introduction to the study of sociology*. New York: Harper & Row.

¹⁰ Guttman, L. 1968. A general nonmetric technique for finding the smallest coordinate space for configurations of points. *Psychometrika*, 33, 469-506.

¹¹ Rosenberg, S., and M. P. Kim. 1975. The method of sorting as a data-gathering procedure in multivariate research. *Multivariate Behavioral Research*, 10, 489-502.

(兩組女性、兩組男性) 根據其相似性來分類整理這些稱謂。他們請其中兩組 (一組女性、一組男性) 作兩次分類整理, 第二次要根據與第一次不同的準則進行分類整理。因此, 取得總計六個「來源」。每一個來源對應於一個 15 x 15 的近似性矩陣, 其資料格等於來源中人數減去物件在該來源中分割在一起的次數。

- **kinship_ini.sav**。本資料檔包含 *kinship_dat.sav* 之三維解的起始配置。
- **kinship_var.sav**。本資料檔包含自變數「性別」、「世代」、和可用來解讀 *kinship_dat.sav* 解答維度的(分離)「度」。尤其, 它們可用來將解答空間限制為這些變數的線性組合。
- **marketvalues.sav**。本資料檔有關於一項在伊立諾州阿爾岡京 (Algonquin, Ill.) 的新房屋開發在 1999–2000 年期間之房屋銷售情況。這些銷售與公共記錄有關。
- **nhis2000_subset.sav**。「國民健康訪問調查 (NHIS)」為美國民間人口的一大型民眾調查。其以具全國代表性的家庭為樣本, 面對面的完成訪問。而取得各家庭中成員的人口統計學資訊及健康行為、健康狀態方面等觀察報告。本資料檔包含一個 2000 年調查資訊的子集。國家衛生統計中心。2000 年「國民健康訪問調查 (NHIS)」。公用資料檔和檔案。 ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/。2003 年曾存取。
- **ozone.sav**。本資料包含對六個氣象變數所作的 330 個觀察值, 以自其餘的變數中預測臭氧濃度。在以前的研究人員中,¹² 和 ¹³ 發現了這些變數之間的非線性關係, 這妨礙了標準迴歸方法。
- **pain_medication.sav**。本假設資料檔包含治療慢性關節炎疼痛之消炎藥物臨床試驗的結果。特別關注於藥物發生作用的時間以及它是如何與現用藥物作比較。
- **patient_los.sav**。本假設資料檔包含對因可能為心肌梗塞 (MI, 或「心臟病») 入院病患的治療記錄。每一個觀察值對應至一個不同的病患並記錄許多與其留院期間有關的變數。
- **patlos_sample.sav**。本假設資料檔包含病患在為心肌梗塞 (MI, 或「心臟病») 治療期間接受血栓溶解治療的治療記錄樣本。每一個觀察值對應至一個不同的病患並記錄許多與其留院期間有關的變數。
- **poll_cs.sav**。這是有關民意測驗專家致力於確定交付立法之前公眾對法案支援層級的假設資料檔。觀察值對應至登記選民。每一個觀察值記錄下選民的郡、鎮、和他居住的鄰近範圍。
- **poll_cs_sample.sav**。本假設資料檔包含列於 *poll_cs.sav* 中的選民樣本。樣本是根據在 *poll_csplan* 計劃檔中指定的設計來取得, 而本資料檔記錄了包含機率和樣本加權。但是請注意, 由於取樣計劃採用到機率 - 比例 - 大小 (PPS) 方法, 也用到一個包含聯合選擇機率的檔案 (*poll_jointprob.sav*)。其他與選民人口及其對提議法案之意見有關的變數都在取樣後收集並加入資料檔中。
- **property_assess.sav**。這是有關郡財產估價人員致力於對限定資源保持財產價值評估維持最新的假設資料檔。觀察值對應至郡內過去一年銷售的財產。資料檔中的每一個觀察值記錄了財產所在的鎮、上次訪查該財產的估價人員、自那次評估後經過的時間、當時定的估價、和該財產銷售價值。
- **property_assess_cs.sav**。這是有關州財產估價人員致力於對限定資源保持財產價值評估維持最新的假設資料檔。觀察值對應至州中的財產。資料檔中的每一個觀察值記錄了郡、鎮、和財產所在的鄰近範圍、自最後一次評估後經過的時間、和當時定的估價。
- **property_assess_cs_sample.sav**。本假設資料檔包含列於 *property_assess_cs.sav* 中的財產樣本。樣本是根據在 *property_assess_csplan* 計劃檔中指定的設計來取得, 而本資料檔記錄了包含機率和樣本加權。另外的變數「目前價值」是在取樣後收集並加入資料檔中。
- **recidivism.sav**。這是有關政府法令執行機構致力於瞭解其轄區內之再犯率的假設資料檔。每一個觀察值對應至一個先前的違法者並記錄其人口資訊、第一次犯罪的一些細節、然後是直到第二次被捕的時間 (如果它發生在第一次被捕的兩年之內)。
- **recidivism_cs_sample.sav**。這是有關政府法令執行機構致力於瞭解其轄區內之再犯率的假設資料檔。每一個觀察值都對應於前一個違法者 (在 2003 年 6 月期間第一次被捕並獲釋放的違法者) 並記錄其個人背景資訊、第一次犯罪的部分詳細資料以及第二次被捕的資料 (如果它發生在 2006 年 6 月結束之前)。違法者是根據 *recidivism_cs.csplan* 中指定的取樣計劃從取樣部門選取的; 因為取樣計劃採用機率 - 比例 - 大小 (PPS) 方法, 還用到一個包含聯合選擇機率的檔案 (*recidivism_cs_jointprob.sav*)。
- **rfm_transactions.sav**。本檔案是包含購買交易資料的假設資料檔, 包括購買日期、購買項目及每一項交易的金額。

¹² Breiman, L., and J. H. Friedman. 1985. Estimating optimal transformations for multiple regression and correlation. *Journal of the American Statistical Association*, 80, 580-598.

¹³ Hastie, T., and R. Tibshirani. 1990. *Generalized additive models*. London: Chapman and Hall.

- **salesperformance.sav**。這是有關評估兩個新售貨員訓練課程的假設資料檔。六十個員工，分成三個組別，全部接受標準訓練。此外，組別二得到技術訓練；組別三則是實務輔導簡介。每一個員工在訓練課程結束時接受測驗並記錄他們的分數。在資料檔中每一個觀察值代表一個不同的訓員，並記錄他們所分派的組別和他們在測驗中得到的分數。
- **satisf.sav**。這是有關一家零售公司在 4 個商店位置所作之滿意度調查的假設資料檔。總共有 582 位客戶接受調查，每一個觀察值代表一位客戶的回應。
- **screws.sav**。該資料檔案包含關於螺釘、螺栓、螺母和圖釘¹⁴的性質的資訊。
- **shampoo_ph.sav**。這是有關一家美髮產品工廠品質管制過程的假設資料檔。在固定的時間間隔，記錄下六個不同輸出批次的測量和它們的 pH 值。目標範圍是 4.5 – 5.5。
- **ships.sav**。在別處被提出和分析的¹⁵關於波浪對貨船造成的損壞的資料集。事件次數可建模為以卜瓦松分佈比率發生，給定船型、建造期間和服務期間。以因素交叉分類所形成的表格的每一個資料格服務月數的整合，提供了暴露於風險之值。
- **site.sav**。這是有關一家公司致力於為事業擴展選擇新地點的假設資料檔。他們僱請兩位顧問分別評估該地點，除了一份廣泛的報告之外，他們還要將每個地點摘要為前景「佳」、「可」、或「差」。
- **smokers.sav**。本資料檔是由「1998 年全國家庭毒品濫用調查」中摘錄，且是美國家庭的機率樣本。
(<http://dx.doi.org/10.3886/ICPSR02934>) 因此，分析本資料檔的第一步應該是將資料加權以反映母體趨勢。
- **stocks.sav** 本假設資料檔包含一年的股票價格和數量。
- **stroke_clean.sav**。本假設資料檔包含一個醫療資料庫，其在以「資料準備」選項中的程序清理之後的狀態。
- **stroke_invalid.sav**。本假設資料檔包含一個醫療資料庫的起始狀態並包含幾個資料輸入錯誤。
- **stroke_survival**。本假設資料檔是有關缺血性中風的病患，其在結束康復計劃後存活時間方面，面臨許多挑戰。中風後，記載了心肌梗塞、缺血性中風、或出血性中風的發生，以及事件記錄的時間。由於它只包含在康復計劃所管制的中風存活的病患，此樣本的左側被截斷。
- **stroke_valid.sav**。本假設資料檔包含一個醫療資料庫，在其值以「驗證資料」程序檢查之後的狀態。它仍包含可能的異常觀察值。
- **survey_sample.sav**。本資料檔包含調查資料，包括人口資料和各種態度測量。雖然已修改一些資料數值，且為人口資料之目的新增了一些額外的虛構變數，但是資料仍是以「1998 NORC 基本社會調查」的變數子集為基礎。
- **tcm_kpi.sav**。這是一個假設資料檔，包含某個公司的每週關鍵績效指標值。它還包含同一時段內，許多可控制度量值的每週資料。
- **tcm_kpi_upd.sav**。此資料檔與 *tcm_kpi.sav* 相同，但是包含四個額外週的資料。
- **telco.sav**。這是有關一家電信公司致力於在客戶庫中減少客戶流失的假設資料檔。每一個觀察值對應至一位不同的客戶並記錄不同的人口資料和服務使用方式資訊。
- **telco_extra.sav**。本資料檔類似於 *telco.sav* 資料檔，但「任期」的對數轉換客戶花費變數已予刪除，並更換為標準的對數轉換客戶花費變數。
- **telco_missing.sav**。本資料檔是 *telco.sav* 資料檔的子集，不過某些人口資料值已更換為遺漏值。
- **testmarket.sav**。本假設資料檔有關於一家速食連鎖店計劃在功能表中加入新的項目。有三個可能的活動來促銷此新產品，所以該新項目在幾個隨機選取市場中的地點作介紹。在每一個地點使用不同的促銷，並記錄該新項目前四週的每週銷售量。每一個觀察值對應至一個不同的地點-週。
- **testmarket_1month.sav**。本假設資料檔是將 *testmarket.sav* 資料檔和每週購買的「彙總」，因此每一個觀察值對應至一個不同的客戶。結果部分每週變更的變數消失了，而目前所記錄的銷售量是在研究的四週期間銷售量之總和。
- **tree_car.sav**。這是包含人口資料和車輛購買價格資料的假設資料檔。
- **tree_credit.sav**。這是包含人口資料和銀行放款歷史資料的假設資料檔。
- **tree_missing_data.sav** 這是包含有大量遺漏值的人口資料和銀行放款歷史資料的假設資料檔。
- **tree_score_car.sav**。這是包含人口資料和車輛購買價格資料的假設資料檔。

¹⁴ Hartigan, J. A. 1975. *Clustering algorithms*. New York: John Wiley and Sons.

¹⁵ McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd ed. London: Chapman & Hall.

- **tree_textdata.sav**。一個只有兩個變數的簡單資料檔，主要目的在顯示變數預設狀態（在指定測量層級和值標籤之前）。
- **tv-survey.sav**。這是有關一家電視製片廠考量是否要延長一個成功節目的播送所作之調查的假設資料檔。有 906 位應答者被問到在不同的狀況下他們是否願意觀看這個節目。每一列代表一個不同的應答者；每一直欄為一個不同的狀況。
- **ulcer_recurrence.sav**。本檔案包含一項用來比較兩種防止潰瘍復發治療法功效之研究的部分資訊。它提供了區間資料的優秀範例並且已在別處被提出和分析¹⁶。
- **ulcer_recurrence_recoded.sav**。本檔案是將 *ulcer_recurrence.sav* 的資訊重新組織，以讓您為此研究的每一個區間事件機率而非只是研究目的事件機率建模。已在別處被提出和分析¹⁷。
- **verd1985.sav**。此資料檔案涉及一項意見調查¹⁸。在調查中記錄了來自 15 個受訪者對 8 個變數的回應。所需的變數被分成三組。集 1 包括 *age* 和 *marital*，集 2 包括 *pet* 和 *news*，集 3 包括 *music* 和 *live*。Pet 調整為多重名義量數，*age* 調整為序數量數，其他的變數調整為單一名義量數。
- **virus.sav**。這是有關一家網際網路服務提供者致力於在其網路上判斷病毒之影響的假設資料檔。他們在其網路上追蹤從發現病毒直到控制威脅的這段時間，被病毒感染之電子郵件的流量（約略）百分比。
- **wheeze_steubenville.sav**。這是關於空氣污染對兒童¹⁹ 健康影響的縱向研究的一個子集。本資料包含來自俄亥俄州 Steubenville，年齡 7、8、9 和 10 歲兒童的氣喘聲狀態之重複二元測量，以及其母親在本研究的第一年是否抽煙的固定記錄。
- **workprog.sav**。這是有關一項政府職業計劃，設法將弱勢民眾安置到較好之工作的假設資料檔。一個樣本的可能計劃參與者被追蹤，他們之中某些被選取加入本計劃，而其他的則否。每一個觀察值代表一位不同的計劃參與者。
- **worldsales.sav** 本假設資料檔包含依洲和產品分類之銷貨收益。

¹⁶ Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

¹⁷ Collett, D. 2003. *Modelling survival data in medical research*, 2 ed. Boca Raton: Chapman & Hall/CRC.

¹⁸ Verdegaal, R. 1985. *Meer sets analyse voor kwalitatieve gegevens (in Dutch)*. Leiden: Department of Data Theory, University of Leiden.

¹⁹ Ware, J. H., D. W. Dockery, A. Spiro III, F. E. Speizer, and B. G. Ferris Jr.. 1984. Passive smoking, gas cooking, and respiratory health of children living in six cities. *American Review of Respiratory Diseases*, 129, 366-374.

注意事項

本資訊係針對 IBM 在美國所提供之產品與服務所開發。IBM 可能會以其他語言提供本資料。但是，您可能需要具有該語言的產品或產品版本，才能存取該產品。

IBM 可能並未在其他國家提供在本文件中討論到的產品、服務或功能。有關目前在貴地區可供使用的產品與服務相關資訊，請洽您當地的 IBM 服務代表。對於 IBM 產品、程式或服務的任何參考，目的並不是要陳述或暗示只能使用 IBM 產品、程式或服務。任何功能相等且未侵犯 IBM 智慧財產權的產品、程式或服務皆可使用。但是，評估及確認任何非 IBM 產品、程式或服務的操作之責任應由使用者承擔。

IBM 可能有一些擁有專利或專利申請中的項目包含本文件所描述的內容。本文件的提供並不表示授與您對於這些專利的權利。您可以將書面的授權查詢寄至：

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
美國

對於與雙位元組 (DBCS) 資訊相關的授權查詢，請與貴國的 IBM 智慧財產部門聯絡，或將查詢郵寄至：

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
19-21, Nihonbashi-Hakozakicho, Chuo-ku
Tokyo 103-8510, Japan

International Business Machines Corporation 只依「現況」提供本出版品，不提供任何明示或默示之保證，其中包括且不限於不侵權、可商用性或特定目的之適用性的隱含保證。有些地區不允許特定交易中明示或默示的保固聲明，因此，此聲明或許對您不適用。

此資訊內容可能包含技術失準或排版印刷錯誤。此處資訊會定期變更，這些變更將會納入新版的聲明中。IBM 可能會隨時改善和 / 或變更此聲明中所述的產品和 / 或程式，恕不另行通知。

本資訊中任何對非 IBM 網站的敘述僅供參考，IBM 對該網站並不提供任何保證。該「網站」的內容並非此 IBM 產品的部分內容，使用該「網站」需自行承擔風險。

IBM 可能會以任何其認為適當的方式使用或散佈您提供的任何資訊，無需對您負責。

意欲針對達成以下目的而擁有本程式相關資訊之程式被授權人：(i) 在獨立建立的程式與其他程式 (包括本程式) 之間交換資訊及 (ii) 共用已交換的資訊，應聯絡：

IBM Director of Licensing
IBM Corporation
North Castle Drive, MD-NC119
Armonk, NY 10504-1785
美國

在適當條款與條件之下，包括某些情況下 (支付費用)，或可使用此類資訊。

在本文件中描述的授權程式及其適用之所有授權材料皆由 IBM 在與我方簽訂之 IBM 客戶合約、IBM 國際程式授權合約或任何相等等效合約中提供。

本文件中引用的效能資料及用戶範例僅供敘述之目的。特定配置及作業條件下的實際效能結果可能不同。

本文件所提及之非 IBM 產品資訊，取自產品的供應商，或其發佈的聲明或其他公開管道。IBM 並未測試過這些產品，也無法確認這些非 IBM 產品的執行效能、相容性或任何對產品的其他主張是否完全無誤。有關非 IBM 產品的功能問題應直接洽詢該產品供應商。

關於 IBM 未來方針或意圖的所有聲明僅代表目標或目的，得依規定未另行通知即變更或撤銷。

此資訊包含用於日常企業運作的資料和報表範例。為了儘可能提供完整說明，範例中包含了人名、公司名稱、品牌名稱和產品名稱。這些名稱全為虛構，如與實際人員或企業之名稱有所雷同，純屬巧合。

著作權授權：

本資訊含有原始語言之範例應用程式，用以說明各作業平台中之程式設計技術。貴客戶可以為了研發、使用、銷售或散布符合範例應用程式所適用的作業平台之應用程式介面的應用程式，以任何形式複製、修改及散布這些範例程式，不必向 IBM 付費。這些範例並未在所有情況下完整測試。故 IBM 不保證或默示保證這些樣本程式之可靠性、服務性或功能。這些程式範例以「現狀」提供，且無任何保證。IBM 對因使用這些程式範例而產生的任何損害概不負責。

這些範例程式或任何衍生成果的每份複本或任何部分，都必須依照下列方式併入著作權聲明：

© Copyright IBM Corp. 2021. 本程式之若干部分係衍生自 IBM 公司的範例程式。

© Copyright IBM Corp. 1989 - 2021. All rights reserved.

商標

IBM、IBM 標誌及 ibm.com 是 International Business Machines Corp. 在世界許多管轄區註冊的商標或註冊商標。其他產品及服務名稱可能是 IBM 或其他公司的商標。IBM 商標的最新清單可在 Web 的 "Copyright and trademark information" 中找到，網址為 www.ibm.com/legal/copytrade.shtml。

Adobe、Adobe 標誌、PostScript 以及 PostScript 標誌為 Adobe Systems Incorporated 於美國和 / 或其他國家的註冊商標或商標。

Intel、Intel 標誌、Intel Inside、Intel Inside 標誌、Intel Centrino、Intel Centrino 標誌、Celeron、Intel Xeon、Intel SpeedStep、Itanium 和 Pentium 為 Intel Corporation 或其分公司於美國和其他國家的商標或註冊商標。

Linux 為 Linus Torvalds 於美國和 / 或其他國家的註冊商標。

Microsoft、Windows、Windows NT 和 Windows 標誌為 Microsoft Corporation 於美國和 / 或其他國家的商標。

UNIX 為 The Open Group 於美國和其他國家的註冊商標。

Java 和所有以 Java 為基礎的商標及標誌是 Oracle 及（或）其子公司的商標或註冊商標。

