

IBM SPSS Statistics 20 ブリー  
フ ガイド

**IBM**®

注：この情報とサポートされている製品をご使用になる前に、「注意事項」（p.176）の一般情報をお読みください。

本版は IBM® SPSS® Statistics 20 ,および新版で指示されるまで後続するすべてのリリースおよび変更に対して適用されます。

Adobe 製品のスクリーンショットは Adobe Systems Incorporated の許可を得て転載しています。

Microsoft 製品のスクリーンショットは Microsoft 社の許可を得て転載しています。

Licensed Materials - Property of IBM

© Copyright IBM Corporation 1989, 2011.

U.S. Government Users Restricted Rights - Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

---

# はじめに

『IBM SPSS Statistics 20 ブリーフ ガイド』には、IBM® SPSS® Statistics の各コンポーネントを理解するためのチュートリアルが含まれています。このガイドは、Windows、Macintosh、および Linux を含む、すべてのオペレーティング システム バージョンのソフトウェアを対象としています。チュートリアルでは、順を追って実行したり、追加情報が必要なトピックに戻ったりできます。このガイドは、SPSS Statistics Core システムに付属のオンライン チュートリアルを補足するものとして使用できます。または、オンライン チュートリアルを無視してこのガイドのチュートリアルから開始することもできます。

## IBM SPSS Statistics 20

IBM® SPSS® Statistics 20 は、データ分析のための統合システムです。SPSS Statistics では、ほとんどのファイル形式のデータをサポートし、取り込んだデータを使用して、グループ化されたレポートおよび図表を生成したり、分布とトレンド、記述統計、および複雑な統計の分析のプロットを作成できます。

SPSS Statisticsを使用すると、初心者には簡単に、熟練したユーザーにはより便利に統計分析ができます。コマンド シンタックスを 1 行も入力することなく、メニューとダイアログ ボックスを選択するという単純な操作だけで、複雑な分析が実行できます。データ エディタは、データの入力や作業データ ファイルの参照に最適な、シンプルかつ効率的なスプレッドシート形式の機能を備えています。

## インターネット リソース

IBM Corp. Web サイト (<http://www.ibm.com/support>) では、FAQ や、データ ファイルへのアクセス、その他の有用な情報を提供しています。

また、SPSS USENET ディスカッション グループ (IBM Corp. の主催ではありません) も開設されており、関心がある方なら誰でも利用できます。USENET のアドレスは、`comp.soft-sys.stat.spss` です。

また、USENET グループにゲートウェイされる電子メッセージ リストに加入することもできます。加入するには、`listserv@listserv.uga.edu` に電子メール メッセージを送ります。電子メールメッセージのテキストは「`subscribe SPSSX-L firstname lastname`」としてください。これで、`listserv@listserv.uga.edu` に電子メールを送信すれば、リストにメッセージを投稿できるようになります。

## 追加の出版物

The IBM SPSS Statistics Statistical Procedures Companion, by Marija Norušis, has been published by Prentice Hall. この本では、IBM® SPSS® Statistics Base、ロジスティック回帰モデル、一般線型モデルの概要が解説されています。また、Prentice Hall 社からは、『IBM SPSS Statistics Advanced Statistical Procedures Companion』という本も出版されています。この本には、高度なモジュールや回帰モジュールの中の手続きの概要があります。

## IBM SPSS Statistics オプション

次のオプションは、(Student 版ではなく) 完全版の IBM® SPSS® Statistics システムへのアドオンの拡張機能として使用できます。

**Statistics Base** では、度数、クロス表、記述統計量、OLAP Cube やコードブック レポートなど、基本的な分析やレポートに関する幅広い統計手順を提供しています。また、因子分析、クラスタ分析、最近隣分析、判別関数分析など、さまざまな次元縮小、分類、セグメント化の方法が用意されています。さらに、SPSS Statistics Base では、t 検定、分散分析、線型回帰、順序回帰など平均の比較や予測方法のアルゴリズムが数多く提供されています。

**Advanced Statistics** は、高度な実験科学や生物医学の研究でよく使用される手法に重点をおいています。一般線型モデル (GLM)、線型混合モデル、分散成分分析、対数線型分析、順序回帰分析、保険統計生命表、Kaplan-Meier の生存分析、Cox 回帰、標準および拡張 Cox 回帰の手続きが含まれます。

**ブートストラップ** とは、平均値、中央値、比率、オッズ比、相関係数、または回帰係数など、推定に対する標準語さおよび信頼区間の頑強な推定を取得する方法です。

**カテゴリ** は、コレスポンデンス分析を含む最適尺度法手続きを行います。

**コンプレックス サンプル** では、意識、市場、健康、世論などの調査を行う場合や、社会科学の分野でサンプル調査を使った研究を行う場合に、複雑なサンプル デザインをデータ分析に組み込むことができます。

**コンジョイント** により、個別の製品属性が消費者や一般市民の嗜好に対して与える影響の現実的な測定手法が提供されます。コンジョイントがあれば、購入決定時に消費者が行う、製品属性グループ中での、各製品属性のトレードオフ効果を簡単に測定することができます。

**カスタム テーブル** は、複合スタブ バナー テーブルや多重回答データの表示など、さまざまな高品質の表形式の報告書を作成します。

**データの準備** はデータを素早くスナップショットにして視覚化します。無効なデータ値を識別する検証規則を適用する機能が搭載されています。この機能を使うと、作成した規則に基づいて、範囲外の値、欠損値、空

白値などにフラグを付けることができます。また、個々の規則違反や、ケースあたりの規則違反の数を記録して変数に保存することができます。ユーザーがコピーしたり修正したりして利用できる、定義済みの規則の集合も用意されています。

**Decision Trees** では、ツリー ベースの分類モデルを作成します。ケースをグループに分類したり、独立（予測）変数の値を基に従属（目的）変数を予測するためのモデルです。この手続きには、分類を探索的、および確証的に分析するための検証ツールが用意されています。

**ダイレクト マーケティング** を使用して、ダイレクト マーケティング向けに特別に設計された方法を使用して、組織はマーケティングのプログラムをできる限り効果的なものにすることができます。

**Exact Tests** は、サンプルが少ないか分布が著しく偏っているために通常の検定では不正確になるような場合に、正確な統計検定の p 値を計算します。このオプションは、Windows オペレーティング システムでのみ使用できます。

**見込み** は、複数の曲線当てはめモデル、平滑法モデル、自己回帰関数の推定方法を使用した包括的な予測分析と時系列分析を行います。

**欠損値** は、欠損データのパターンの表示、平均値と他の統計量の推定、および欠損観測値への値の代入を行います。

**Neural Networks** を使用すると、価格とその他の変数の関数として製品の需要を予測することにより、また、購買傾向や人口統計上の特性に基づいて顧客を分類することにより、ビジネス意思決定を行うことができます。Neural Networks は、非線型のデータ モデリング ツールです。入力と出力の複雑な関係のモデリングや、データ内でのパターンの検索に使用できます。

**Regression** は、従来の線型統計モデルに適合しないデータを分析する手法を提供します。プロビット分析、ロジスティック回帰、重み付け推定、2 段階最小 2 乗回帰、および従来の非線型回帰の手続きが含まれています。

**Amos™** (analysis of moment structures) では、構造方程式モデリングを使用して、属性、認識、および、振る舞いを生じさせるその他の因子に関わる概念モデルの確認と説明を行ないます。

## トレーニング セミナー

IBM Corp. では一般公開およびオンサイトで IBM® SPSS® Statistics のトレーニング セミナーを実施しています。すべてのセミナーは米国およびヨーロッパの主要都市で定期的で開催されます。セミナーに関する詳細については、<http://www.ibm.com/software/analytics/spss/training/> を参照してください。

## テクニカル サポート

テクニカル サポートのサービスをご利用いただけます。IBM Corp. 製品の使用方法や、対応しているハードウェア環境へのインストールに関して問い合わせることもできます。テクニカル サポートの詳細については、IBM Corp. Web サイト (<http://www.ibm.com/support>) を参照してください。連絡の際は、所属団体名、サポート契約などを確認できるよう、あらかじめ手元にご用意ください。

## IBM SPSS Statistics 20 Student 版

IBM® SPSS® Statistics 20 Student 版は機能が限定されていますが、SPSS Statistics の十分に強力なバージョンといえます。

### 機能

Student 版には、次のような IBM® SPSS® Statisticsに含まれている重要なデータ分析ツールが数多く備わっています。

- スプレッドシート形式のデータ エディタ。データ ファイルの入力、変更、および表示に使用します。
- t 検定、分散分析、クロス表を始めとする統計的手続き。
- インタラクティブ グラフ。図表の構成要素を動的に変更および追加することができ、変更が即座に反映されます。
- 標準高解像度のグラフィック。分析およびプレゼンテーション用に多数の図表を配列することができます。

### 制限

Student 版は、大学などの授業で学生および教師が使用することを前提に作成されており、それ以外の目的に使用することはできません。IBM® SPSS® Statistics 20 Student Version には、次の制限が適用されます。

- 変数の個数が 50 を超えるデータ ファイルは使用不可。
- Student 版で SPSS Statistics アドオン モジュール (Regression や Advanced Statistics など) は使用不可。
- SPSS Statistics コマンド シンタックスが使用できない。IBM® SPSS® Statistics の Student 版では一連のコマンドをシンタックスまたは「ジョブ」ファイルに保存して分析を繰り返すことができません。ただし、完全バージョンでは、この操作は可能です。
- スクリプト機能およびオートメーション機能が使用できない。SPSS Statisticsの完全バージョンでは、頻繁に実行するタスクをスクリプトでオートメーション化することができますが、Student 版ではそれができません。

## 学生向けテクニカル サポート

IBM SPSS ソフトウェア製品の Student 版、アカデミック版、Grad パック版を使用している学生の場合、学生用の特別オンライン ページ、[Solutions for Education \(http://www.ibm.com/spss/rd/students/\)](http://www.ibm.com/spss/rd/students/) ページを参照してください。大学提供の IBM SPSS ソフトウェアのコピーを使用している場合、大学の IBM SPSS 製品コーディネータにお問い合わせください。

## 教官向けテクニカル サポート

授業で Student 版を使用している講師の方々は、テクニカル サポートを受けることができます。テクニカル サポートに連絡するには、当社 Web サイト (<http://www.ibm.com/support>) をご覧になるか、当社 (03-5466-5511) までお電話ください。

---

# 内容

<b>1</b>	<b>はじめに</b>	<b>1</b>
	サンプル ファイル . . . . .	1
	データ ファイルを開く . . . . .	1
	分析の実行 . . . . .	3
	結果の表示 . . . . .	7
	グラフの作成 . . . . .	8
<b>2</b>	<b>データの読み込み</b>	<b>11</b>
	IBM SPSS Statistics データファイルの基本構造 . . . . .	11
	IBM SPSS Statistics データファイルの読み込み . . . . .	11
	スプレッドシートからのデータの読み込み . . . . .	12
	データベースからのデータの読み込み . . . . .	14
	テキストファイルからのデータの読み込み . . . . .	21
<b>3</b>	<b>データ エディタの使用方法</b>	<b>29</b>
	数値データを入力する . . . . .	29
	文字型データの入力 . . . . .	32
	データの定義 . . . . .	34
	変数ラベルの追加 . . . . .	34
	変数の型と書式の変更 . . . . .	35
	数値型変数の値ラベルの追加 . . . . .	36
	文字型変数の値ラベルの追加 . . . . .	38
	データ入力での値ラベル使用方法 . . . . .	39
	欠損データの処理 . . . . .	40
	数値型変数の欠損値 . . . . .	41
	文字型変数の欠損値 . . . . .	43
	変数属性のコピーと貼り付け . . . . .	44
	カテゴリ変数の変数プロパティの定義 . . . . .	48



<b>4</b>	<b>多重データソースの使用方法</b>	<b>55</b>
	多重データソースの基本的な処理	56
	コマンドシンタックスの複合データセットを使用する作業	57
	データセット間での情報のコピーと貼り付け	58
	データセットの名前の変更	58
	複合データセットを抑制	59
<b>5</b>	<b>個々の変数に対する要約統計量の検討</b>	<b>60</b>
	尺度レベル	60
	カテゴリデータの要約統計量	60
	カテゴリデータのグラフ	62
	スケール変数の要約統計量	64
	スケール変数のヒストグラム	66
<b>6</b>	<b>グラフの作成と編集</b>	<b>68</b>
	図表作成の基本	68
	図表ビルダー ギャラリーの使用	69
	変数と統計量の定義	71
	テキストの追加	74
	図表を作成する	75
	図表編集の基本	76
	グラフ要素の選択	77
	[プロパティ] ウィンドウの使用	78
	棒の色の変更	78
	目盛りラベルに表示される数値の書式を設定する。	80
	テキストの編集	82
	データ値ラベルを表示する	83
	テンプレートの使用	84
	図表オプションの定義	89
<b>7</b>	<b>出力の作業</b>	<b>93</b>
	ビューアを使用する	93

ピボット テーブル エディタの使用方法	95
出力の定義へのアクセス	95
テーブルのピボット	96
層の作成と表示	98
テーブルの編集	99
行および列を隠す	100
データ表示書式の変更	100
テーブルルック	102
定義済み書式の使用方法	102
テーブルルック スタイルのカスタマイズ	103
デフォルト テーブル書式の変更	106
初期の表示設定のカスタマイズ	108
変数ラベルと値ラベルの表示	109
SPSS 出力を他のアプリケーションで使用する	111
結果を Word のテーブルとして貼り付ける	111
結果を Word のテーブルとして貼り付ける	112
Microsoft Word、PowerPoint および Excel ファイルへの結果のエクスポート	114
結果の PDF へのエクスポート	123
結果の HTML へのエクスポート	126

## 8 シンタックスの使用 127

シンタックスを貼り付ける	127
シンタックスを編集する	129
シンタックス ファイルを開いて実行する	131
エラー枠について	132
分割点の使用	132

## 9 データ値の変更 135

スケール変数からカテゴリ変数を作成する	135
新しい変数の計算	141
数式での関数の使用	143
条件式の使用	145
日付と時刻を使用した作業	147
2つの日付間の時間の長さの計算	148
日付への期間の加算	152

## 10 データの並べ替えと選択 156

データを並べ替える.....	156
ファイルの分割処理 .....	157
ファイルの分割処理用にケースを並べ替える.....	159
ファイルの分割処理のオンとオフを切り替える .....	159
ケースのサブグループを選択する .....	159
条件式に基づいたケースを選択する.....	160
無作為抽出を選択する.....	161
日付の範囲またはケース番号の範囲を選択する .....	162
選択されなかったケースの処理.....	163
ケースの選択状態 .....	164

## 付録

A サンプル ファイル	165
-------------	-----

B 注意事項	176
--------	-----

索引	179
----	-----



# はじめに

このガイドは、データの有効な分析を行えるように用意された一式のチュートリアルを提供します。チュートリアルでは、順を追って実行したり、追加情報が必要なトピックに戻ったりできます。

この章では、基本的な機能の概要を説明し、標準セッションについて具体的に説明します。すでに定義されている IBM® SPSS® Statistics のデータ ファイルを取得して、簡単な統計分析およびグラフを作成します。

この章で説明する多くのトピックの詳細は、後の章で説明します。ここでは、後続のチュートリアルを理解するための基本について習得します。

## サンプル ファイル

ここで表示される多くの例では、データ ファイル demo.sav を使用します。このデータ ファイルは数千人分の架空調査ファイルで、基本的な人口統計および消費者情報などを含んでいます。

Student 版を使用する場合、demo.sav のバージョンは、1500 人のケースに対応するように縮小された、元のデータ ファイルの代表的なサンプルになります。このデータ ファイルを使用して得た結果は、ここで表示される結果とは異なります。

製品とともにインストールされるサンプル ファイルは、インストールディレクトリの Samples サブディレクトリにあります。[サンプル] サブディレクトリ内に次の各言語の別のフォルダがあります。英語、フランス語、ドイツ語、イタリア語、日本語、韓国語、ポーランド語、ロシア語、簡体字中国語、スペイン語、そして繁体中国語です。

すべてのサンプル ファイルが、すべての言語で使用できるわけではありません。サンプル ファイルがある言語で使用できない場合、その言語のフォルダには、サンプル ファイルの英語バージョンが含まれています。

## データ ファイルを開く

データ ファイルを開くには、次の手順を実行します。

- ▶ メニューから次の項目を選択します。  
ファイル(F) > 開く(O) > データ(A)...

または、ツールバーの [ファイルを開く] ボタンを使用することもできます。

図 1-1  
[ファイルを開く] ツールバー ボタン



ファイルを開くためのダイアログ ボックスが表示されます。

デフォルトでは、IBM® SPSS® Statistics のデータ ファイル（拡張子は .sav）が表示されます。

この例では、ファイル demo.sav を使用します。

図 1-2  
データ エディタ中の demo.sav ファイル

	年齢	婚姻状況	住居年数	収入	収入カゴリ	車	車カゴリ	教育
1	55	1	12	72.00	3.00	36.20	3.00	
2	56	0	29	153.00	4.00	76.90	3.00	
3	28	1	9	28.00	2.00	13.70	1.00	
4	24	1	4	26.00	2.00	12.50	1.00	
5	25	0	2	23.00	1.00	11.30	1.00	
6	45	1	9	76.00	4.00	37.20	3.00	
7	42	0	19	40.00	2.00	19.80	2.00	
8	35	0	15	57.00	3.00	28.20	2.00	
9	46	0	26	24.00	1.00	12.20	1.00	
10	34	1	0	89.00	4.00	46.10	3.00	
11	55	1	17	72.00	3.00	35.50	3.00	

データ エディタにデータ ファイルが表示されます。データ エディタの中で、マウスのカーソルを変数名（列見出し）の上に移動すると、（その変数にラベルが定義されていれば）より詳しい変数ラベルが表示されます。

デフォルトで実際のデータ値が表示されます。ラベルを表示するには

- ▶ メニューから次の項目を選択します。

表示 > 値ラベル

または、ツールバーの [値ラベル] ボタンを使用することもできます。

図 1-3  
[値ラベル] ボタン



これで、値を説明する値ラベルが表示されるので、回答の解釈がより容易になります。

図 1-4  
[データ エディタ] に表示される値ラベル

	年齢	婚姻状況	住居年数	収入	収入カテゴリ	車	車カテゴリ	教育
1	55	既婚	12	72.00	50 - 74	36.20	高い	中
2	56	未婚	29	153.00	75以上	76.90	高い	中
3	28	既婚	9	28.00	25 - 49	13.70	安い	専門学
4	24	既婚	4	26.00	25 - 49	12.50	安い	大
5	25	未婚	2	23.00	25未満	11.30	安い	高
6	45	既婚	9	76.00	75以上	37.20	高い	専門学
7	42	未婚	19	40.00	25 - 49	19.80	普通	専門学
8	35	未婚	15	57.00	50 - 74	28.20	普通	高
9	46	未婚	26	24.00	25未満	12.20	安い	中
10	34	既婚	0	89.00	75以上	46.10	高い	専門学
11	55	既婚	17	72.00	50 - 74	35.50	高い	専門学

## 分析の実行

アドオン オプションがある場合、[分析] メニューには、レポートのリストおよび統計分析カテゴリが含まれます。

簡単な度数分布表（度数のテーブル）の作成から始めます。この例では Statistics Base オプションが必要です。

- ▶ メニューから次の項目を選択します。  
分析(A) > 記述統計 > 度数分布表...

[度数分布表] ダイアログ ボックスが表示されます。

図 1-5

[度数分布表] ダイアログ ボックス



各変数の隣にあるアイコンは、データ型、および、尺度に関する情報を提供します。

	数値	文字列	日付	時刻
スケール (連続)		利用不可		
順序				
名義				



- ▶ [世帯全体の収入カテゴリ (千ドル) [収入カテゴリ]] という変数をクリックします。

図 1-6

[度数分布表] ダイアログ ボックスの変数ラベルおよび変数名



リスト内の変数ラベルや変数名の一部が隠れて見えない場合は、そのラベルまたは名前の上にカーソルを合わせると、全体が表示されます。記述的な変数ラベルの後に、角かっこで囲まれた変数名（「収入カテゴリ」）が表示されます。[世帯全体の収入カテゴリ (千ドル)] が変数ラベルです。変数ラベルがない場合は、変数名だけがリスト ボックスに表示されます。

ダイアログ ボックスは、ウィンドウと同様に、外枠や角をクリックしてドラッグするとサイズ変更できます。たとえば、ダイアログ ボックスの幅を大きくすると、変数リストの幅も大きくなります。

図 1-7

サイズ変更したダイアログ ボックス

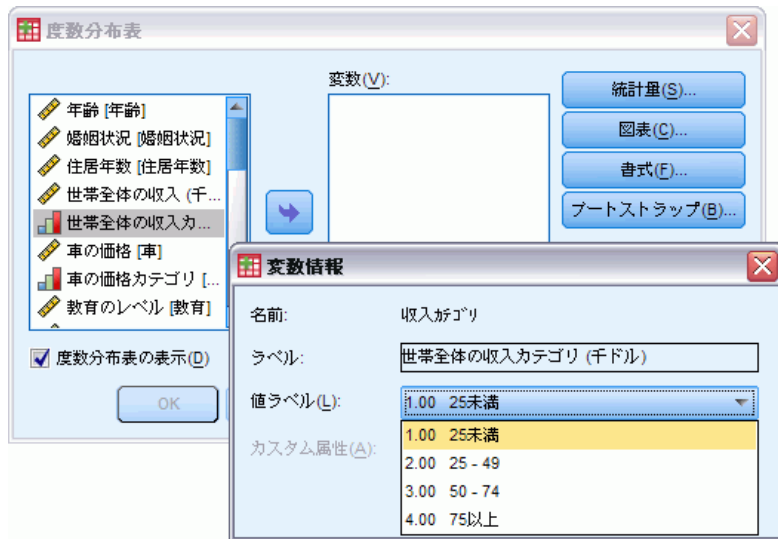


このダイアログ ボックスで、分析対象の変数を左側のソース リストから選択し、右側の [変数] リストにドラッグ アンド ドロップで移動します。分析を実行する [OK] ボタンは、[変数] リストに最低 1 つの変数がないと使用できません。

多くのダイアログで、リスト内の任意の変数名を右クリックすると、詳細情報を確認できます。

- ▶ [世帯全体の収入カテゴリ (千ドル) [収入カテゴリ]] を右クリックして、[変数情報] を選択します。
- ▶ [値ラベル] ドロップダウン リストの下矢印をクリックします。

図 1-8  
定義済みの「収入」変数のラベル

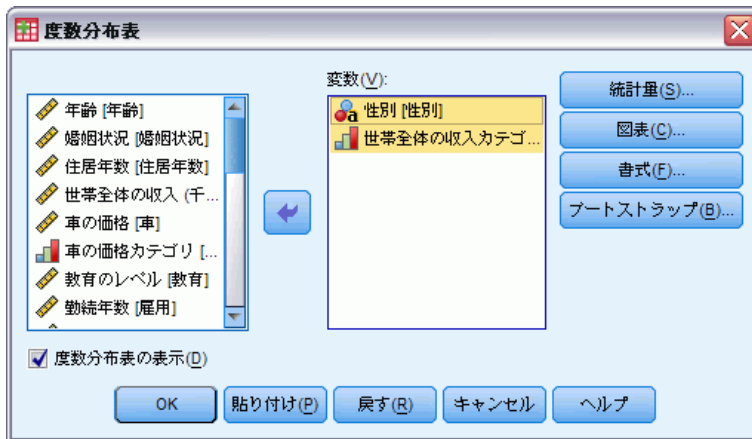


変数の定義済みの値ラベルがすべて表示されます。

- ▶ ソース変数リストで [性別 [性別]] をクリックして、移動先の [変数] リストにドラッグします。

- ▶ ソース リストで [世帯全体の収入カテゴリ (千ドル) [収入カテゴリ]] をクリックして、移動先のリストにドラッグします。

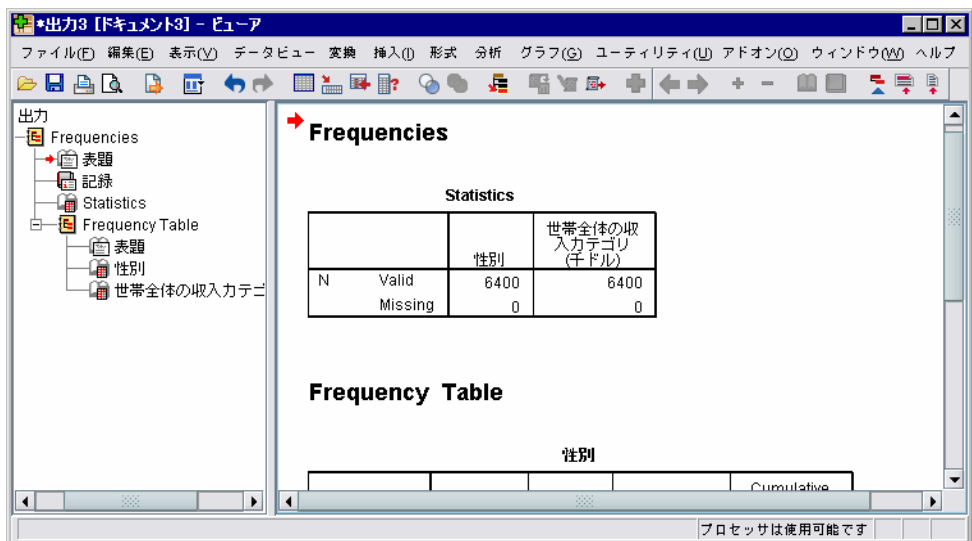
図 1-9  
分析するために選択された変数



- ▶ [OK] をクリックして手続きを実行します。

## 結果の表示

図 1-10  
[ビューア] ウィンドウ

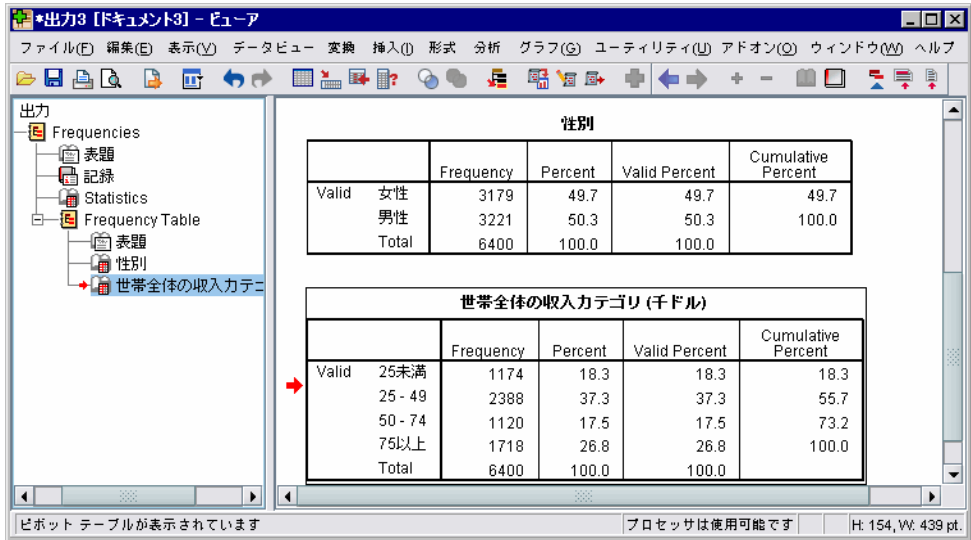


結果が [ビューア] ウィンドウに表示されます。

アウトライン ウィンドウ枠で項目を選択すれば、[ビューア] のどの項目にもすぐに移動できます。

- ▶ [世帯全体の収入カテゴリ(千ドル)[収入カテゴリ]] をクリックします。

図 1-11  
「世帯全体の収入カテゴリ(千ドル)」の度数分布表



「世帯全体の収入カテゴリ(千ドル)」の度数分布表が表示されます。この度数分布表は、各「世帯全体の収入カテゴリ(千ドル)」における人数とパーセンテージを表しています。

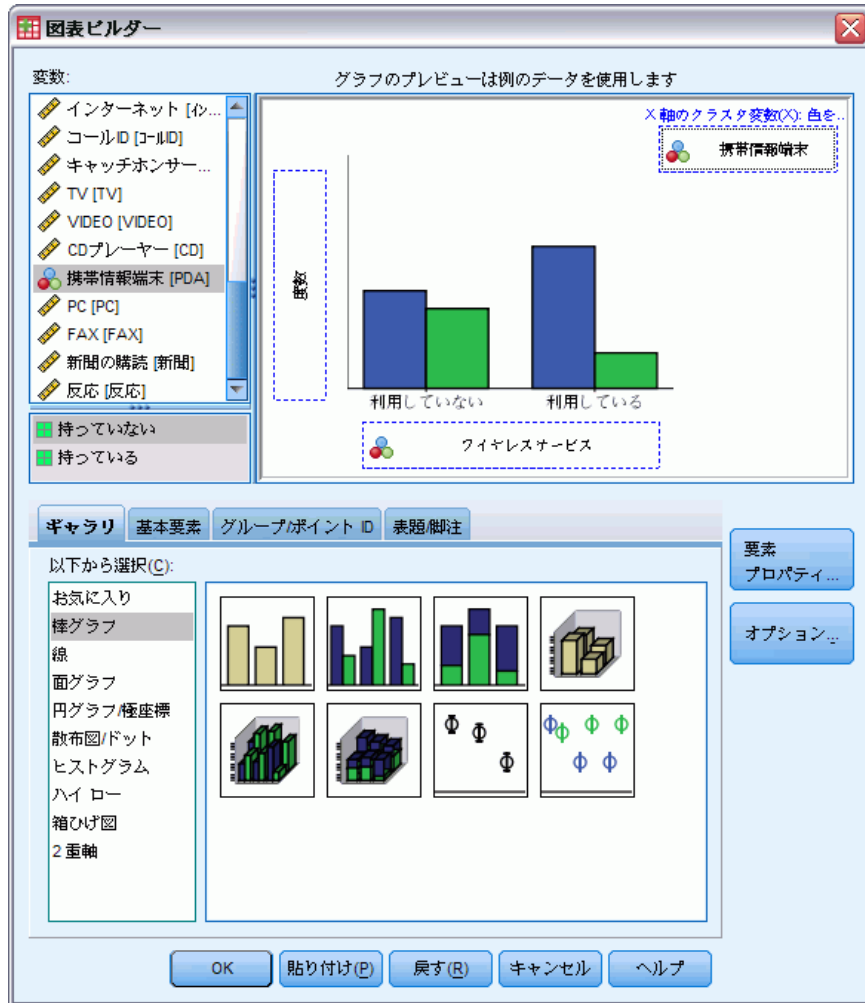
## グラフの作成

統計手続きによってはグラフを作成できますが、[グラフ]メニューを使用してグラフを作成することもできます。

たとえば、ワイヤレス サービスと携帯情報端末 (personal digital assistant:PDA) の所有の関係を示すグラフを作成するとします。

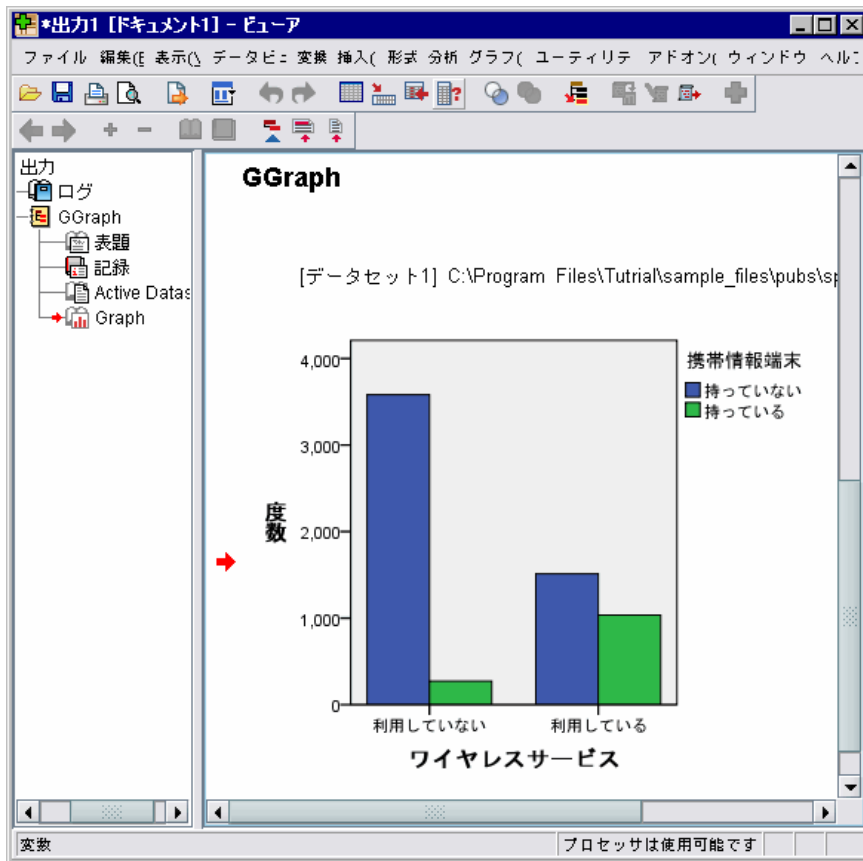
- ▶ メニューから次の項目を選択します。  
グラフ(G) > 図表ビルダー(C)...
- ▶ [ギャラリー] タブ (がまだ選択されていなければ) をクリックします。
- ▶ [棒グラフ] (がまだ選択されていなければ) をクリックします。
- ▶ [クラスタ棒グラフ] をキャンバスにドラッグします。キャンバスというのは、ギャラリーの上にある広い領域のことです。

図 1-12  
[図表ビルダー] ダイアログ ボックス



- ▶ 変数リストを下へスクロールして、[ワイヤレス サービス [ワイヤレス]] を右クリックし、尺度として [名義] を選択します。
- ▶ [ワイヤレス サービス [ワイヤレス]] を X 軸にドラッグします。
- ▶ [携帯情報端末 [PDA]] を右クリックして、測定レベルとして [名義] を選択します。
- ▶ キャンバスの右上隅にある、クラスター ドロップ領域に、[携帯情報端末 [PDA]] 変数をドラッグします。
- ▶ [OK] をクリックして、図表を作成します。

図 1-13  
[ビューア] ウィンドウに表示される棒グラフ



[ビューア] に棒グラフが表示されます。このグラフは、ワイヤレス サービスを使用する人は、使用しない人よりも、PDA を持っている人がかなり多いと見込まれることを表しています。

[ビューア] ウィンドウの内容枠のグラフやテーブル上をダブルクリックすると、グラフやテーブルを編集したり、結果を他のアプリケーションにコピーして貼り付けることができます。これらの操作については、後で説明します。

# データの読み込み

データは、直接入力することも、さまざまなソースからインポートすることもできます。この章では、IBM® SPSS® Statistics データ ファイル、Microsoft Excel などのスプレッドシート アプリケーション、Microsoft Access などのデータベース アプリケーション、およびテキスト ファイルに保存されたデータを読み込む手順について説明します。

## IBM SPSS Statistics データファイルの基本構造

図 2-1  
データ エディタ

	年齢	婚姻状況	住居年数	収入	収入カテゴリ	車	車カテゴリ	教育
1	55	1	12	72.00	3.00	36.20	3.00	
2	56	0	29	153.00	4.00	76.90	3.00	
3	28	1	9	28.00	2.00	13.70	1.00	
4	24	1	4	26.00	2.00	12.50	1.00	
5	25	0	2	23.00	1.00	11.30	1.00	
6	45	1	9	76.00	4.00	37.20	3.00	
7	42	0	19	40.00	2.00	19.80	2.00	
8	35	0	15	57.00	3.00	28.20	2.00	
9	46	0	26	24.00	1.00	12.20	1.00	
10	34	1	0	89.00	4.00	46.10	3.00	
11	55	1	17	72.00	3.00	35.50	3.00	

IBM® SPSS® Statistics データ ファイルは、ケース（行）と変数（列）で構成されています。ケースとは、調査の個々の回答を指します。変数は、調査に使用した個々の質問への回答を指します。

## IBM SPSS Statistics データファイルの読み込み

保存したデータは、IBM® SPSS® Statistics データ ファイル（拡張子は .sav）に格納されます。製品と共にインストールされている `demo.sav` ファイルを開くには、次の手順を実行します。

- ▶ メニューから次の項目を選択します。  
ファイル(F) > 開く(O) > データ(A)...
- ▶ demo.sav を参照して開きます。 [詳細は、A 付録 p.165 サンプル ファイル を参照してください。](#)  
データ エディタにデータが表示されます。

図 2-2  
開いたデータ ファイル

	年齢	婚姻状況	住居年数	収入	収入カテゴリ	車	車カテゴリ	教育
1	55	1	12	72.00	3.00	36.20	3.00	
2	56	0	29	153.00	4.00	76.90	3.00	
3	28	1	9	28.00	2.00	13.70	1.00	
4	24	1	4	26.00	2.00	12.50	1.00	
5	25	0	2	23.00	1.00	11.30	1.00	
6	45	1	9	76.00	4.00	37.20	3.00	
7	42	0	19	40.00	2.00	19.80	2.00	
8	35	0	15	57.00	3.00	28.20	2.00	
9	46	0	26	24.00	1.00	12.20	1.00	
10	34	1	0	89.00	4.00	46.10	3.00	
11	55	1	17	72.00	3.00	35.50	3.00	

## スプレッド シートからのデータの読み込み

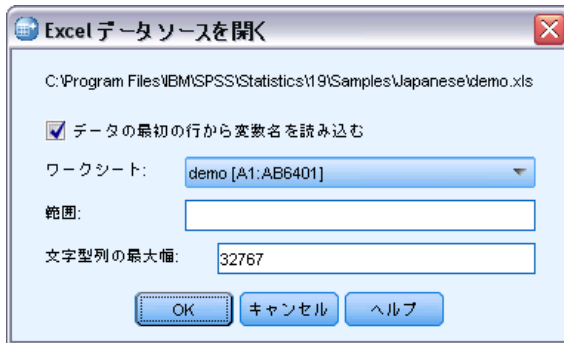
データ エディタにすべてのデータを直接入力する代わりに、Microsoft Excel などのアプリケーションからデータを読み込むことができます。列見出しを変数名として読み込むこともできます。

- ▶ メニューから次の項目を選択します。  
ファイル(F) > 開く(O) > データ(A)...
- ▶ 表示するファイルの種類として [Excel (\*.xls)] を選択します。
- ▶ demo.xls を開きます。 [詳細は、A 付録 p.165 サンプル ファイル を参照してください。](#)

[Excel データ ソースを開く] ダイアログ ボックスが開きます。ここでは、スプレッドシートに変数名を含めるかどうかと、どのセルをインポートするかを指定します。Excel 95 以降では、インポートするワークシートも指定できます。



図 2-3  
[Excel データソースを開く] ダイアログ ボックス



- ▶ [データの最初の行から変数名を読み込む] が選択されていることを確認します。これは、列見出しを変数名として読み込むオプションです。

IBM® SPSS® Statistics変数名の規則に合わない列見出しは、有効な変数名に変換され、元の列見出しは変数ラベルとして保存されます。 スプレッドシートの一部だけをインポートする場合は、[範囲] テキスト ボックスでインポートするセルの範囲を指定します。

- ▶ [OK] をクリックすると、Excel ファイルが読み込まれます。

データ エディタにデータが表示されます。列見出しは変数名として使用されています。 変数名にはスペースを含むことができないので、元の列見出しに含まれていたスペースは、取り除かれています。たとえば、Excel ファイルの [婚姻状況] は、「婚姻状況」という変数になります。元の列見出しは、変数ラベルとして保持されます。

図 2-4  
インポートした Excel データ

The screenshot shows the Microsoft Access interface with a table named '24: 収入カテゴリ' (24: Income Category). The table contains 9 rows of data. The columns are: 年齢 (Age), 婚姻状況 (Marital Status), 居住年数 (Years of Residence), 収入 (Income), and 収入カテゴリ (Income Category). The status bar at the bottom indicates 'プロセッサは使用可能です' (Processor is available).

	年齢	婚姻状況	居住年数	収入	収入カテゴリ
1	55	1	12	72.00	3.00
2	56	0	29	153.00	4.00
3	28	1	9	28.00	2.00
4	24	1	4	26.00	2.00
5	25	1	2	23.00	1.00
6	45	0	9	76.00	4.00
7	44	1	17	144.00	4.00
8	46	1	20	75.00	4.00
9	41	0	10	26.00	2.00

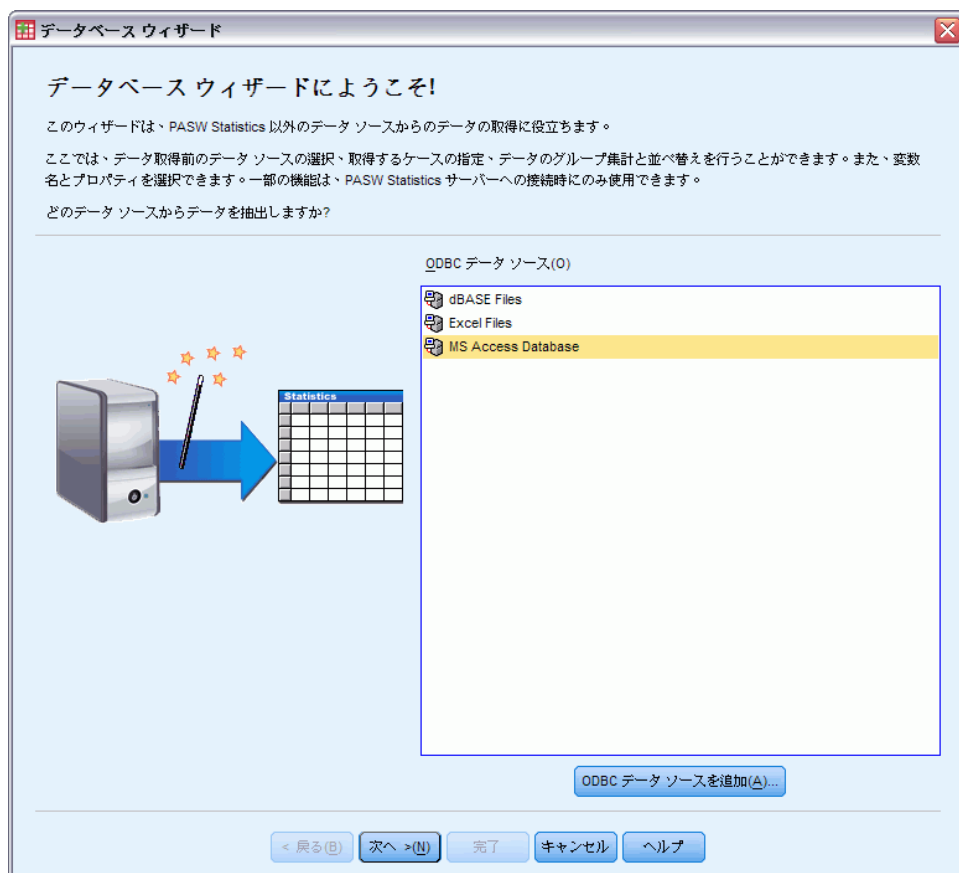
## データベースからのデータの読み込み

データベース ソースからのデータは、データベース ウィザードを使用すれば簡単にインポートできます。ODBC (Open Database Connectivity) ドライバを使用しているデータベースは、ドライバをインストールすれば直接読み込むことができます。インストール CD に、多数のデータベース形式に対応した ODBC ドライバが用意されています。その他のドライバは、それぞれのソフトウェア開発会社から入手できます。ここでは、広く普及しているデータベース アプリケーションの 1 つである Microsoft Access を例に取って説明します。

注:これは Microsoft Windows に固有の例であり、Access 用の ODBC ドライバが必要です。ステップは他のプラットフォームと同様ですが、サードパーティの Access 用 ODBC ドライバが必要な場合があります。

- ▶ メニューから次の項目を選択します。  
ファイル(F) > データベースを開く(B) > 新規クエリー(N)...

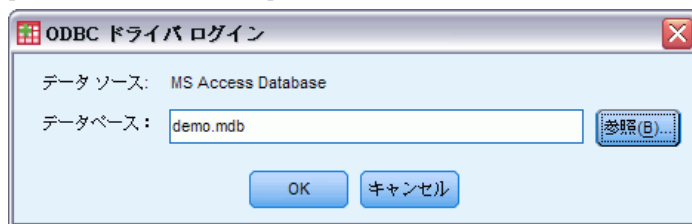
図 2-5  
[データベース ウィザードによるこそ!] ダイアログ ボックス



- ▶ データ ソースのリストから [MS Access Database] を選択し、[次へ] をクリックします。

注: また、インストールの状況によっては、ウィザードの左側に OLEDB データソースのリストが表示されることもあります (Windows オペレーティング システムのみ) が、この例では、右側に表示された ODBC データソースのリストを使います。

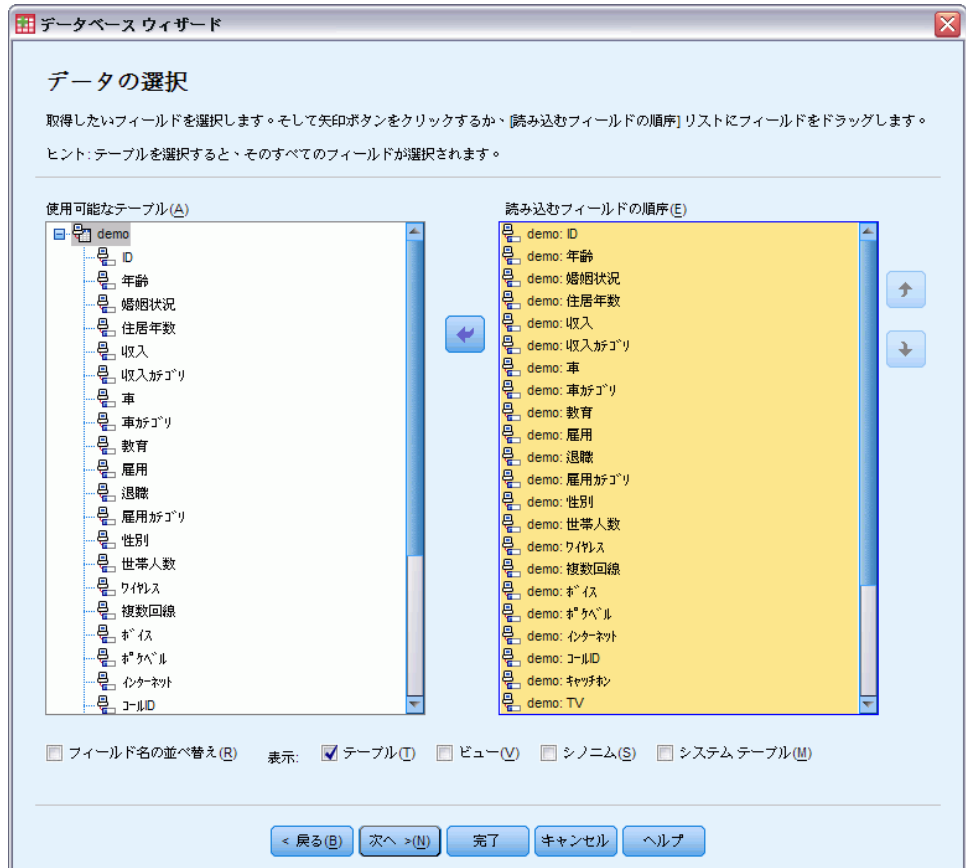
図 2-6  
[ODBC ドライバ ログイン] ダイアログ ボックス



- ▶ [参照] をクリックし、目的の Access データベース ファイルのあるフォルダに移動します。
- ▶ demo.mdb を開きます。詳細は、A 付録 p.165 サンプル ファイル を参照してください。
- ▶ [ODBC ドライバ ログイン] ダイアログ ボックスで [OK] をクリックします。

次のステップでは、インポートするテーブルと変数を指定します。

図 2-7  
[データの選択] ステップ



- ▶ [demo] テーブル全体を [読み込むフィールドの順序] リストへドラッグします。
- ▶ [次へ] をクリックします。

次のステップでは、インポートするレコード（ケース）を選択します。

図 2-8  
[抽出するケースの制限] ステップ

すべてのケースをインポートしたくない場合は、ケースのサブセット（たとえば、31 才以上の男性）をインポートしたり、データ ソースからケースを無作為抽出してインポートできます。データ ソースが大きい場合、小さなサンプルを代表とすれば処理時間を短縮することができます。

- ▶ [次へ] をクリックして先へ進みます。

フィールド名は、変数名を作成するために使用されます。フィールド名は、必要に応じて有効な名前に変換されます。その場合、元のフィールド名は、変数ラベルとして保持されます。また、データベースをインポートする前に、変数名を変更できます。

図 2-9  
[変数を定義します。] ステップ



- ▶ [性別] フィールドで [数値への値の再割り当て] セルをクリックします。このオプションを選択すると、文字型変数が整数型変数に変換され、元の値は変換後の変数の値ラベルとなります。
- ▶ [次へ] をクリックして先へ進みます。

データベース ウィザードでの選択を基に作成された SQL ステートメントが [結果] ステップに表示されます。このステートメントは、すぐに実行することも、ファイルに保存して後で実行することもできます。

図 2-10  
[結果] ステップ



- ▶ [完了] をクリックすると、データがインポートされます。



これで、インポート用に選択した Access データベース内のデータすべてが、データ エディタで使用可能になりました。

図 2-11  
Access データベースからインポートしたデータ

	年齢	婚姻状況	居住年数	収入	収入カテゴリ
1	55	1	12	72.00	3.00
2	56	0	29	153.00	4.00
3	28	1	9	28.00	2.00
4	24	1	4	26.00	2.00
5	25	1	2	23.00	1.00
6	45	0	9	76.00	4.00
7	44	1	17	144.00	4.00
8	46	1	20	75.00	4.00
9	41	0	10	26.00	2.00

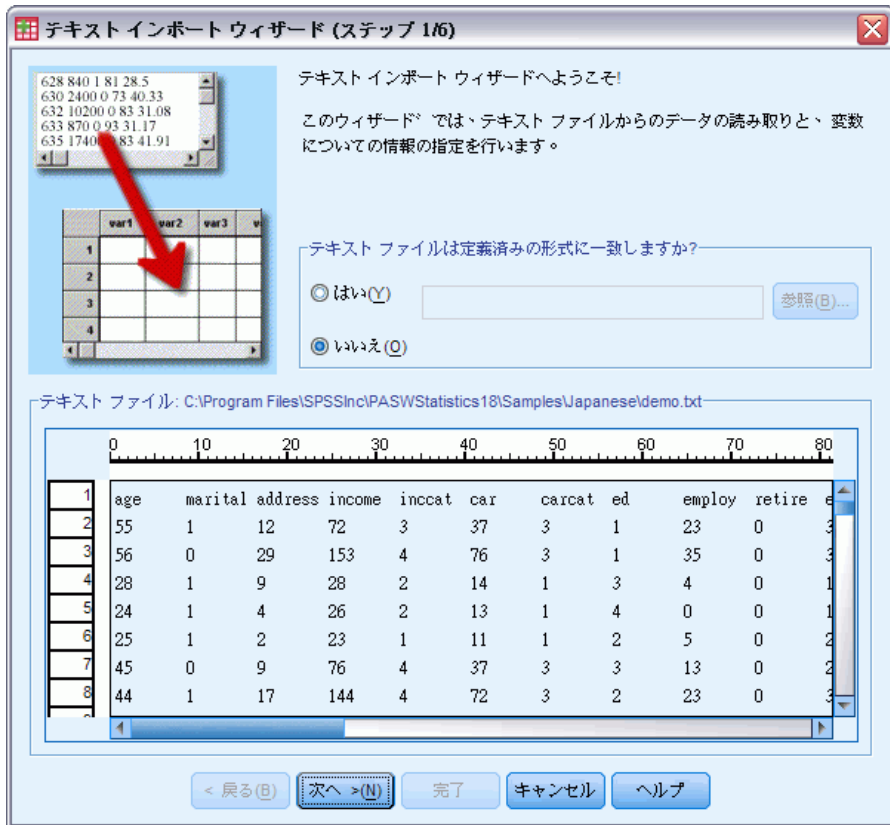
## テキスト ファイルからのデータの読み込み

広く普及しているデータ ソースのもう 1 つの形式が、テキスト ファイルです。表計算ソフトやデータベースでも、内容をいずれかのテキスト ファイル形式で保存できるのが一般的です。カンマ区切りまたはタブ区切りのファイルとは、各変数をカンマまたはタブで区切った一連のデータ行を指します。この例では、タブ区切りのデータを使用します。

- ▶ メニューから次の項目を選択します。  
ファイル(F) > テキスト データの読み込み(D)...
- ▶ 表示するファイルの種類として [Text (\*.txt)] を選択します。
- ▶ demo.txt を開きます。詳細は、A 付録 p.165 サンプル ファイル を参照してください。

テキスト インポート ウィザードのステップに従って、指定したテキストファイルの解釈方法を定義します。

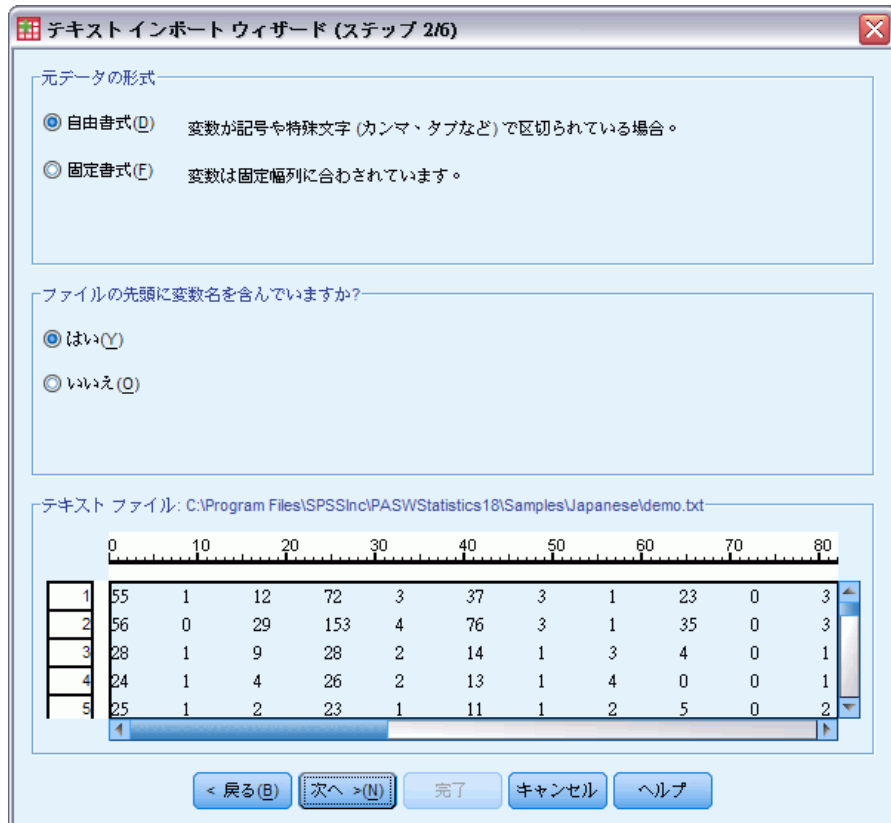
図 2-12  
テキスト インポート ウィザード (ステップ 1/6)



- ▶ ウィザードのステップ 1 では、定義済みの形式を選択するか、新しい形式を作成します。新しい形式を作成する場合は、[いいえ] を選択してください。
- ▶ [次へ] をクリックして先へ進みます。

前述のとおり、このファイルの形式はタブ区切りです。また、変数名はファイルの最初の行で定義されています。

図 2-13  
テキスト インポート ウィザード (ステップ 2/6)



- ▶ データが区切り構造になっているため、[自由書式] を選択します。
- ▶ 変数名がファイルの最初の行から読み込まれるようにするため、[はい] を選択します。
- ▶ [次へ] をクリックして先へ進みます。

- ▶ 次のダイアログ ボックスの最上部にあるセクションに、「2」を入力します。これは、データ行がテキスト ファイルの 2 行目から始まることを表します。

図 2-14  
テキスト インポート ウィザード (ステップ 3/6)

テキスト インポート ウィザード - 自由書式 (ステップ 3/6)

最初のケースの取り込み開始行番号(F)

ケースの表される方法

各行が 1 つのケースを表す(L)

特定の数の変数が 1 つのケースを表す(V)

インポートするケース数

すべてのケース(A)

最初の(T)  ケース

無作為抽出するケース割合 (概算)(P):  %

データ プレビュー

	0	10	20	30	40	50	60	70	80		
1	55	1	12	72	3	37	3	1	23	0	3
2	56	0	29	153	4	76	3	1	35	0	3
3	28	1	9	28	2	14	1	3	4	0	1
4	24	1	4	26	2	13	1	4	0	0	1

< 戻る(B) 

- ▶ ダイアログ ボックスにあるその他の値はデフォルトのままにし、[次へ] をクリックして先へ進みます。

ステップ 4 のデータ プレビューを見ると、データが適切な形で読み込まれているかどうかを即座に確認できます。

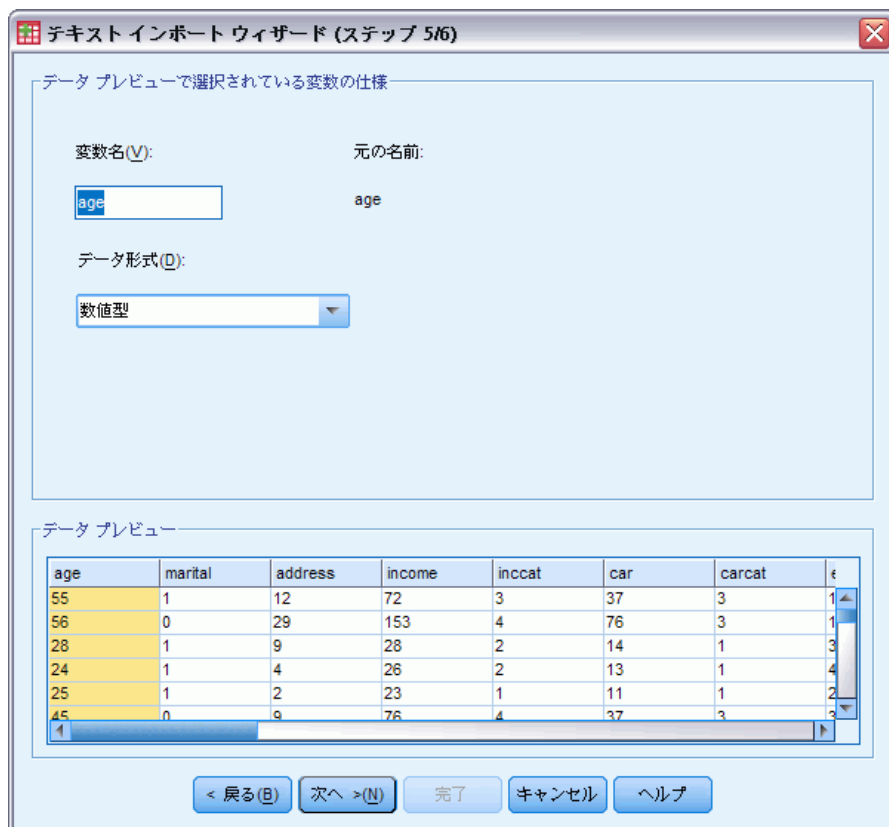
図 2-15  
テキスト インポート ウィザード (ステップ 4/6)



- ▶ [タブ] を選択し、その他のオプションの選択を解除します。
- ▶ [次へ] をクリックして先へ進みます。

書式条件に合わせて変数名が切り詰められ、不適切な名前になった場合は、このダイアログ ボックスで編集できます。

図 2-16  
テキスト インポート ウィザード (ステップ 5/6)



データの型は、ここでも定義できます。たとえば、収入変数であればドルを単位とした金額が含まれると考えられます。

データ型を変更するには、次の手順に従ってください。

- ▶ [データ プレビュー] で、変更する変数（この例では「収入」）を選択します。

- ▶ [データ形式] ドロップダウン リストから [ドル記号] を選択します。

図 2-17  
データ型の変更



- ▶ [次へ] をクリックして先へ進みます。

図 2-18  
テキスト インポート ウィザード (ステップ 6/6)



- ▶ ダイアログ ボックス内の値をデフォルトのままにし、[完了] をクリックしてデータのインポートを開始します。



# データ エディタの使用方法

データ エディタには、アクティブなデータ ファイルの内容が表示されます。データ エディタ内の情報は、変数とケースで構成されています。

- [データ ビュー] では、列が変数を、行がケース（観測値）を表します。
- [変数ビュー] では、各行が変数であり、各列がその変数に関連する属性です。

変数は、収集した異なるデータの種類を表すために使用されます。アンケート調査がその好例です。アンケート調査におけるそれぞれの質問に対する回答が変数に相当します。変数には、数値型、文字型、通貨型、日付型を含むさまざまな型があります。

## 数値データを入力する

データはデータ エディタに入力することができます。データ エディタは、小さいデータ ファイルを取り扱ったり、比較的大きなデータ ファイルに軽微な編集を加える際に便利です。

- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。

ここで、使用する変数を定義する必要があります。この場合、必要な変数は、年齢、婚姻、収入です。

図 3-1  
[変数ビュー] に表示された変数名

	名前	型	幅	小数桁数	ラベル	値	欠損値
1	年齢	数値型	8	2		なし	なし
2	婚姻状況	数値型	8	2		なし	なし
3	年収	数値型	8	2		なし	なし
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

- ▶ 1 列目の 1 行目に、「年齢」と入力します。
- ▶ 2 行目に、「婚姻状況」と入力します。
- ▶ 3 行目に、「年収」と入力します。

新しい変数は、自動的に数値型に設定されます。

変数名を入力しない場合、一意の名前が自動的に作成されます。ただし、これらの名前は記述的ではないので、データ ファイルが多数ある場合にはお勧めしません。

- ▶ [データビュー] タブをクリックし、データの入力を続けます。

これで、[変数ビュー] に入力した名前が、[データ ビュー] の最初の 3 列の見出しになります。

1 列目の 1 行目からデータを入力していきます。

図 3-2  
[データビュー] で入力された値

	年齢	婚姻状況	年収	var	var	var	var
1	55.00	1.00	72000.00				
2	53.00	0.00	153000.00				
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							

- ▶ [年齢] 列に、55 と入力します。
- ▶ [婚姻状況] 列に、1 と入力します。
- ▶ [年収] 列に、72000 と入力します。
- ▶ 次の被験者のデータを追加するために、1 列目の 2 行目にカーソルを移動します。
- ▶ [年齢] 列に、53 と入力します。
- ▶ [婚姻状況] 列に、0 と入力します。
- ▶ [年収] 列に、153000 と入力します。

現在、[年齢] と [婚姻状況] 列には、本来の整数ではなく、小数が表示されています。これらの変数の小数部を非表示にするには、次の操作を実行します。

- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。

- ▶ 小数部を非表示にするために、[年齢] 行の「小数桁数」列に「0」と入力します。
- ▶ 小数部を非表示にするために、[婚姻状況] 行の「小数桁数」列に「0」と入力します。

図 3-3  
更新された [年齢] および [婚姻状況] の [小数桁数] プロパティ

	名前	型	幅	小数桁数	ラベル	値	欠損値
1	年齢	数値型	8	0		なし	なし
2	婚姻状況	数値型	8	0		なし	なし
3	年収	数値型	8	2		なし	なし
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

## 文字型データの入力

データ エディタには、テキスト文字列等の非数値型データも入力できます。

- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。
- ▶ 最初の空白行の最初のセルに、変数名として「性別」と入力します。
- ▶ 入力したセルの隣にある [型] セルをクリックします。

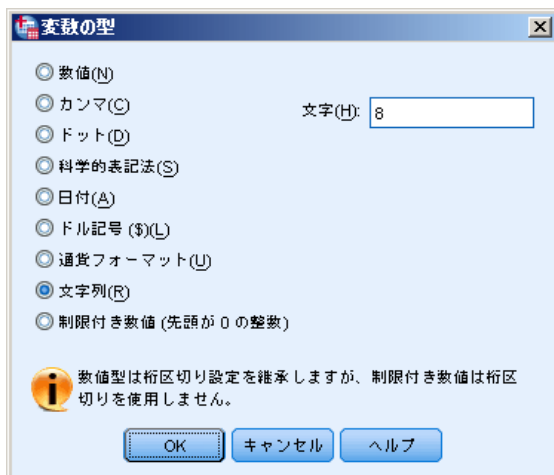
- ▶ [型] セルの右端にあるボタンをクリックし、[変数の型] ダイアログボックスを開きます。

図 3-4  
[性別] の [型] セルのボタン



- ▶ [文字型] を選択して変数の型を指定します。
- ▶ [OK] をクリックし、変更を保存してデータ エディタに戻ります。

図 3-5  
[変数の型] ダイアログボックス



## データの定義

データの型を定義するだけでなく、変数名とデータ値の記述的な変数ラベルと値ラベルも定義できます。このような記述的なラベルは、統計レポートおよび図表に使用されます。

## 変数ラベルの追加

ラベルには、変数についての説明が記述されています。この説明は、多くの場合、変数名よりも長くなります。ラベルの長さは、最高で 255 バイトです。ラベルは、出力の中で変数を識別するために使用されます。

- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。
- ▶ [年齢] 行の [ラベル] 列に、「回答者の年齢」と入力します。
- ▶ 「婚姻状況」行の [ラベル] 列に、「婚姻状況」と入力します。
- ▶ 「年収」行の [ラベル] 列に、「世帯全体の収入」と入力します。

- ▶ 「性別」 行の [ラベル] 列に、「性別」と入力します。

図 3-6  
[変数ビュー] に入力された変数ラベル

	名前	型	幅	小数桁数	ラベル	値	欠損値
1	年齢	数値型	8	0	回答者の年齢	なし	なし
2	婚姻状況	数値型	8	0	婚姻状況	なし	なし
3	年収	数値型	8	2	世帯全体の収入	なし	なし
4	性別	文字列	8	0	性別	なし	なし
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

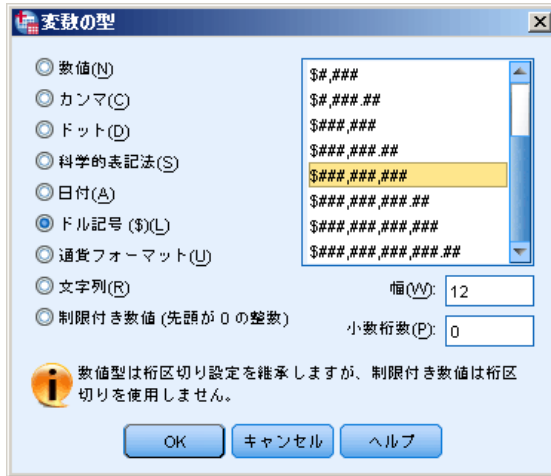
## 変数の型と書式の変更

[型] 列には、各変数について現在のデータの型が表示されます。最も一般的なデータ型は数値型と文字型ですが、これ以外にもさまざまな形式を使用することができます。現在のデータ ファイルでは、[年収] 変数が数値型として定義されています。

- ▶ 「年収」 行の [型] セルをクリックし、さらに、セルの右端のボタンをクリックして、[変数の型] ダイアログ ボックスを開きます。

- ▶ [ドル記号] を選択します。

図 3-7  
[変数の型] ダイアログ ボックス



現在選択されているデータの型の書式オプションが表示されます。

- ▶ この例では、通貨の書式として、[\$###,###,###] を選択します。
- ▶ [OK] をクリックし、変更を保存します。

## 数値型変数の値ラベルの追加

値ラベルにより、文字ラベルに変数値をマップすることができるようになります。この例では、「婚姻状況」変数に適切な値が 2 つあります。0 は被験者が独身であることを意味し、1 は既婚であることを意味します。

- ▶ 「婚姻状況」行の [値] セルをクリックし、さらに、セルの右端にあるボタンをクリックして、[値ラベル] ダイアログ ボックスを開きます。

[値] は、実際の数値です。

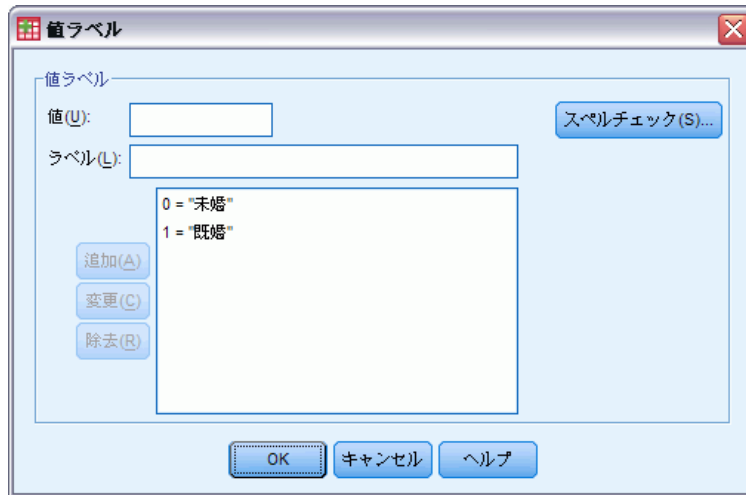
[値ラベル] は、指定された数値に適用される文字列ラベルです。

- ▶ [値] フィールドに「0」と入力します。
- ▶ [値ラベル] フィールドに「未婚」と入力します。



- ▶ [追加] をクリックし、このラベルをリストに追加します。

図 3-8  
[値ラベル] ダイアログ ボックス



- ▶ [値] フィールドに「1」と入力し、[値ラベル] フィールドに「既婚」と入力します。
- ▶ [追加] をクリックします。次に、[OK] をクリックし、変更を保存してデータ エディタに戻ります。  
これらのラベルは [データ ビュー] にも表示できるので、データをさらに読みやすくすることができます。
- ▶ [データ エディタ] ウィンドウの下端にある [データビュー] タブをクリックします。
- ▶ メニューから次の項目を選択します。  
表示 > 値ラベル

データ エディタで値を入力するときに、ラベルが表示されるようになります。この設定には、有効な回答を提案したり、より記述的な回答を提供できる利点があります。

すでに、[値ラベル] のメニュー項目がアクティブになっている（隣にチェック マークがついている）状態で、再度 [値ラベル] を選択すると、値ラベルの表示はオフになります。

図 3-9  
[データビュー] に表示された値ラベル

The screenshot shows a window titled '\*無題1 [データセット0] - データ エディタ(D)'. The menu bar includes 'ファイル(F)', '編集(E)', '表示(V)', 'データビュー', '変換', '分析', 'グラフ(G)', 'ユーティリティ(U)', 'アドオン(O)', 'ウィンドウ(W)', and 'ヘルプ'. The status bar at the bottom shows 'データ ビュー' and '変数 ビュー'. The main area displays a table with the following data:

	年齢	婚姻状況	年収	性別	var	var	var
1	55	既婚	\$72,000				
2	53	未婚	\$153,000				
3							
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							

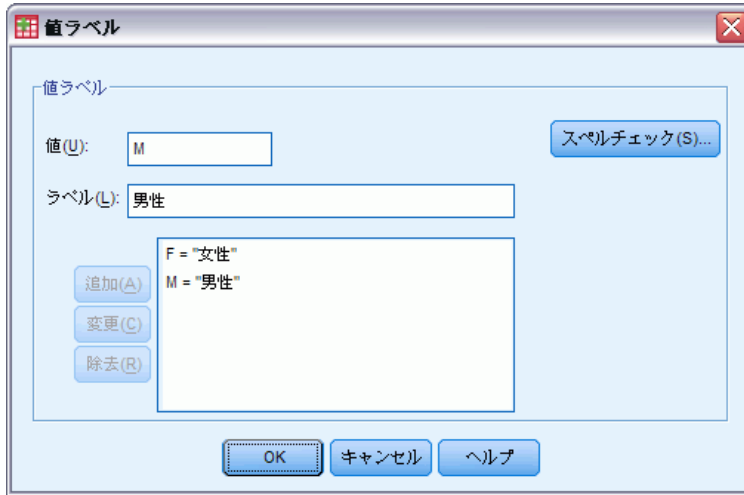
## 文字型変数の値ラベルの追加

文字型変数についても同様に値ラベルが必要になる場合があります。たとえば、被験者の性別を識別するために、1文字の M または F がデータとして使用することがあります。値ラベルを使用して、M が男性 (Male) を表し、F が女性 (Female) を表すことを指定できます。

- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。
- ▶ 「性別」行の [値] セルをクリックし、さらに、セルの右端にあるボタンをクリックして、[値ラベル] ダイアログ ボックスを開きます。
- ▶ [値] フィールドに「F」と入力し、[値ラベル] フィールドに「女性」と入力します。

- ▶ [追加] をクリックし、このラベルをデータ ファイルに追加します。

図 3-10  
[値ラベル] ダイアログ ボックス



- ▶ [値] フィールドに「M」と入力し、[値ラベル] フィールドに「男性」と入力します。
- ▶ [追加] をクリックします。次に、[OK] をクリックし、変更を保存してデータ エディタに戻ります。

文字型値では大文字と小文字が区別されるので、大文字と小文字を一致させる必要があります。小文字の m は大文字の M と同じではありません。

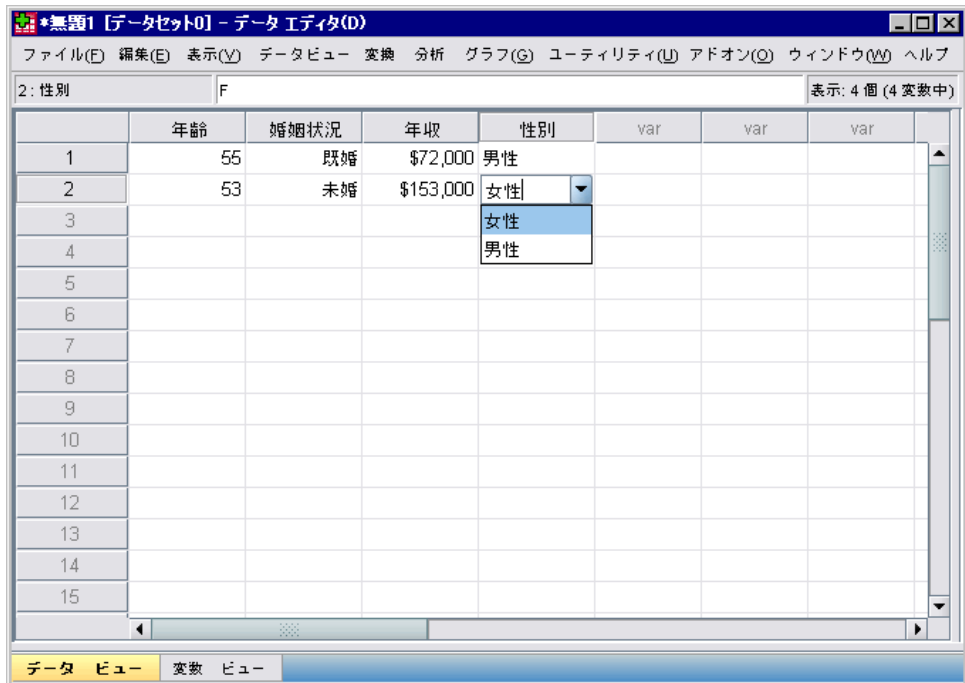
## データ入力での値ラベル使用方法

値ラベルは、データ入力に使用することができます。

- ▶ [データ エディタ] ウィンドウの下端にある [データビュー] タブをクリックします。
- ▶ 1 行目の「性別」のセルを選択します。
- ▶ セルの右端にあるボタンをクリックして、ドロップダウン リストから [男性] を選択します。
- ▶ 2 行目の「性別」のセルを選択します。

- ▶ セルの右端にあるボタンをクリックして、ドロップダウン リストから [女性] を選択します。

図 3-11  
変数ラベルを使用した値の選択



定義済みの値だけがリストに表示されるので、適切な書式のデータだけが入力されることが保証されます。

## 欠損データの処理

欠損または無効なデータは、無視するのが通例です。調査回答者が、ある種の質問には回答しなかったり、答えを知らなかったり、または、不適切な書式で回答することがあります。このようなデータを除外または識別しないと、分析結果が不正確になる可能性があります。

数値データでは、空白のデータ フィールドまたは無効な入力が含まれるデータ フィールドは、1 つのピリオドによって識別されるシステム欠損フィールドに変換されます。

図 3-12  
ピリオドとして表示される欠損値

The screenshot shows a window titled '\*無題1 [データセット0] - データ エディタ(D)'. The menu bar includes 'ファイル(F)', '編集(E)', '表示(V)', 'データビュー', '変換', '分析', 'グラフ(G)', 'ユーティリティ(U)', 'アドオン(O)', 'ウィンドウ(W)', and 'ヘルプ'. Below the menu bar, there is a search field containing '20:' and a status indicator '表示: 4 個 (4 変数中)'. The main area is a table with the following data:

	婚姻状況	年収	性別	var	var	var	var
1	既婚	\$72,000	男性				
2	未婚	\$153,000	女性				
3	.	.	.				
4							
5							
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							

At the bottom of the window, there are two tabs: 'データ ビュー' (Data View) and '変数 ビュー' (Variable View).

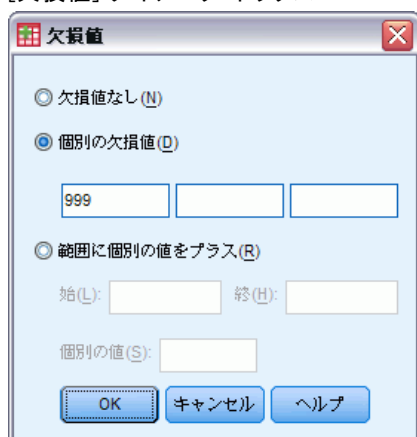
値が欠損する理由は、分析にとって重要である場合があります。たとえば、質問に答えることを拒否した回答者と、質問が該当しなかったため答えなかった回答者とを区別することは重要です。

## 数値型変数の欠損値

- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。
- ▶ 「年齢」行の [欠損値] セルをクリックし、さらに、セルの右端にあるボタンをクリックして、[欠損値] ダイアログ ボックスを開きます。

このダイアログでは、異なる欠損値を 3 つまで設定するか、ある値の範囲に個別の値を 1 つ追加して指定することができます。

図 3-13  
[欠損値] ダイアログ ボックス



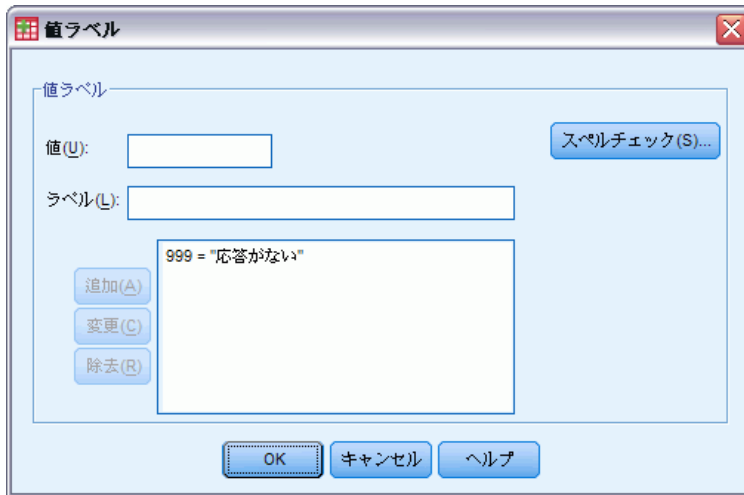
- ▶ [個別の欠損値] を選択します。
- ▶ 最初のテキスト ボックスに「999」と入力し、残りの 2 つのテキスト ボックスは空白のままにしておきます。
- ▶ [OK] をクリックし、変更を保存してデータ エディタに戻ります。

これで欠損データ値が追加されたので、その値にラベルを適用することができます。

- ▶ 「年齢」行の [値] セルをクリックし、さらに、セルの右端にあるボタンをクリックして、[値ラベル] ダイアログ ボックスを開きます。
- ▶ [値] フィールドに「999」と入力します。

- ▶ [ラベル] フィールドに「なし」と入力します。

図 3-14  
[値ラベル] ダイアログ ボックス



- ▶ [追加] をクリックし、このラベルをデータ ファイルに追加します。
- ▶ [OK] をクリックし、変更を保存してデータ エディタに戻ります。

## 文字型変数の欠損値

文字型変数の欠損値は、数値変数の欠損値と同様に処理します。ただし、文字型変数では、数値変数の場合とは異なり、空白のフィールドがシステム欠損として指定されません。その代わりに、空白文字として解釈されます。

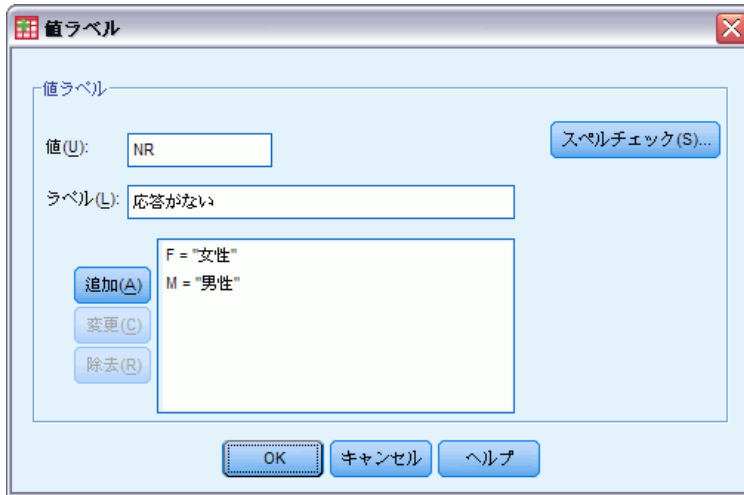
- ▶ [データ エディタ] ウィンドウの下端にある [変数ビュー] タブをクリックします。
- ▶ 「性別」行の [欠損値] セルをクリックし、さらに、セルの右端にあるボタンをクリックして、[欠損値] ダイアログ ボックスを開きます。
- ▶ [個別の欠損値] を選択します。
- ▶ 最初のテキスト ボックスに「NR」と入力します。

文字型変数の欠損値では、大文字と小文字が区別されます。したがって、nr は欠損値として扱われません。
- ▶ [OK] をクリックし、変更を保存してデータ エディタに戻ります。

次に、欠損値にラベルを付けることができます。

- ▶ 「性別」行の [値] セルをクリックし、さらに、セルの右端にあるボタンをクリックして、[値ラベル] ダイアログ ボックスを開きます。
- ▶ [値] フィールドに「NR」と入力します。
- ▶ [ラベル] フィールドに「なし」と入力します。

図 3-15  
[値ラベル] ダイアログ ボックス



- ▶ [追加] をクリックし、このラベルをプロジェクトに追加します。
- ▶ [OK] をクリックし、変更を保存してデータ エディタに戻ります。

## 変数属性のコピーと貼り付け

変数に対して変数属性を定義すれば、属性をコピーし、他の変数に適用することができます。



- ▶ [変数ビュー] で、最初の空白行の最初のセルに、「結婚年齢」と入力します。

図 3-16

[変数ビュー] に表示された [結婚年齢] 変数

	名前	型	幅	小数桁数	ラベル	値	欠損値
1	年齢	数値型	8	0	回答者の年齢	{999, なし}...	999
2	婚姻状況	数値型	8	0	婚姻状況	{0, 未婚}...	なし
3	年収	ドル記号 (\$)	12	0	世帯全体の収入	なし	なし
4	性別	文字列	8	0	性別	{F, 女性}...	NR
5	結婚年齢	数値型	8	2	結婚年齢	なし	なし
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

- ▶ [ラベル] 列に、「結婚年齢」と入力します。
- ▶ 「年齢」行の [値] セルをクリックします。
- ▶ メニューから次の項目を選択します。  
編集 > コピー
- ▶ 「結婚年齢」行の [値] セルをクリックします。
- ▶ メニューから次の項目を選択します。  
編集 > 貼り付け (X11 ARIMA / 度数分布表 / 記述統計)

これで、「年齢」変数の定義済みの値が、「結婚年齢」変数に適用されます。

複数のターゲットセルを単に選択する（列をクリックしてドラッグする）だけで、属性を複数の変数に適用することができます。

図 3-17  
複数のセルが選択されている状態

	型	幅	小数桁数	ラベル	値	欠損値	列
1	数値型	8	0	回答者の年齢	{999, なし}...	999	8
2	数値型	8	0	婚姻状況	{0, 未婚}...	なし	8
3	ドル記号 (\$)	12	0	世帯全体の収入	なし	なし	8
4	文字列	8	0	性別	{F, 女性}...	NR	8
5	数値型	8	2	結婚年齢	{999.00, なし}...	なし	8
6							
7							
8							
9							
10							
11							
12							
13							
14							
15							
16							
17							

属性の貼り付けを実行すると、その属性は選択されているすべてのセルに適用されます。

空白行に値を貼り付けると、新しい変数が自動的に作成されます。

特定の変数の属性を、すべて、他の変数にコピーするには

- ▶ 「婚姻状況」 行の行番号をクリックします。

図 3-18  
選択された行

	名前	型	幅	小数桁数	ラベル	値	欠損値
1	年齢	数値型	8	0	回答者の年齢	{999, なし}...	999
2	婚姻状況	数値型	8	0	婚姻状況	{0, 未婚}...	なし
3	年収	ドル記号 (\$)	12	0	世帯全体の収入	なし	なし
4	性別	文字列	8	0	性別	{F, 女性}...	NR
5	結婚年齢	数値型	8	2	結婚年齢	{999.00, な...	なし
6	VAR00001	数値型	8	2		{999.00, な...	なし
7	VAR00002	数値型	8	2		{999.00, な...	なし
8	VAR00003	数値型	8	2		{999.00, な...	なし
9							
10							
11							
12							
13							
14							
15							
16							
17							

- ▶ メニューから次の項目を選択します。  
編集 > コピー
- ▶ 最初の空白行の行番号をクリックします。
- ▶ メニューから次の項目を選択します。  
編集 > 貼り付け (X11 ARIMA / 度数分布表 / 記述統計)

「婚姻状況」変数のすべての属性が、新しい変数に適用されます。

図 3-19  
値がすべて同じ行に貼り付けられた

	名前	型	幅	小数桁数	ラベル	値	欠損値
1	年齢	数値型	8	0	回答者の年齢	{999, なし}...	999
2	婚姻状況	数値型	8	0	婚姻状況	{0, 未婚}...	なし
3	年収	ドル記号 (\$)	12	0	世帯全体の収入	なし	なし
4	性別	文字列	8	0	性別	{F, 女性}...	NR
5	結婚年齢	数値型	8	2	結婚年齢	{999.00, なし}...	なし
6	VAR00001	数値型	8	2		{999.00, なし}...	なし
7	VAR00002	数値型	8	2		{999.00, なし}...	なし
8	VAR00003	数値型	8	2		{999.00, なし}...	なし
9	VAR00004	数値型	8	0	婚姻状況	{0, 未婚}...	なし
10							
11							
12							
13							
14							
15							
16							
17							

## カテゴリ変数の変数プロパティの定義

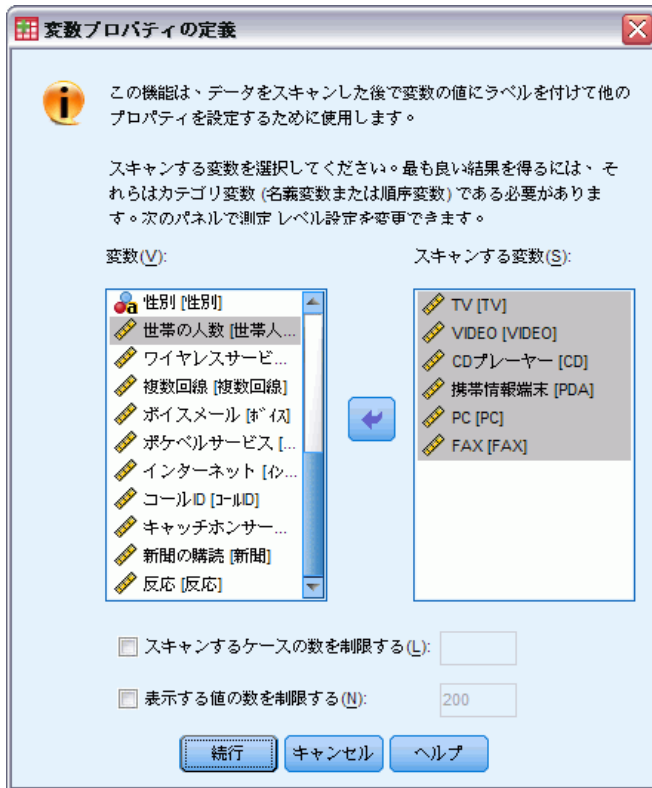
カテゴリ（名義、順序）データに対しては、[変数プロパティの定義] を使用して、値ラベルやその他の変数プロパティを定義できます。[変数プロパティの定義] 処理：

- 実際のデータ値をスキャンして、選択した各変数に対して一意的な全データ値をリストします。
- ラベルのない値を認識して、「自動ラベル」機能を提供します。
- 他の変数から選択した変数へ、あるいは選択した変数から追加的な変数へ、定義済みの値ラベルをコピーできます。

この例では、データ ファイル demo.sav を使用します。詳細は、A 付録 p.165 サンプル ファイル を参照してください。このデータ ファイルには値ラベルがすでに定義されているので、定義済みの値ラベルを持たない値を入力します。

- ▶ データ エディタの [データ ビュー] で、「パソコン」変数（必要に応じて右にスクロールしてください）の最初のデータ セルをクリックし、「99」と入力します。
- ▶ メニューから次の項目を選択します。  
データ > 変数プロパティの定義(V)...

図 3-20  
最初の [変数プロパティの定義] ダイアログ ボックス



最初の [変数プロパティの定義] ダイアログ ボックスでは、値ラベルまたはその他のプロパティを定義する名義または順序変数を選択します。

- ▶ [TV を所有 [owntv]] から [FAX を所有 [ownfax]] を [スキャンする変数] リストにドラッグ アンド ドロップします。

選択した変数のすべての尺度アイコンにより、これらの変数がカテゴリ変数ではなくスケール変数であることが示されます。この例で選択した変数はすべて、実際は数値 0 と 1 を使用してそれぞれ [いいえ] と [はい] を表すカテゴリ変数です。そして、[変数プロパティの定義] で変更するプロパティの 1 つは、尺度です。

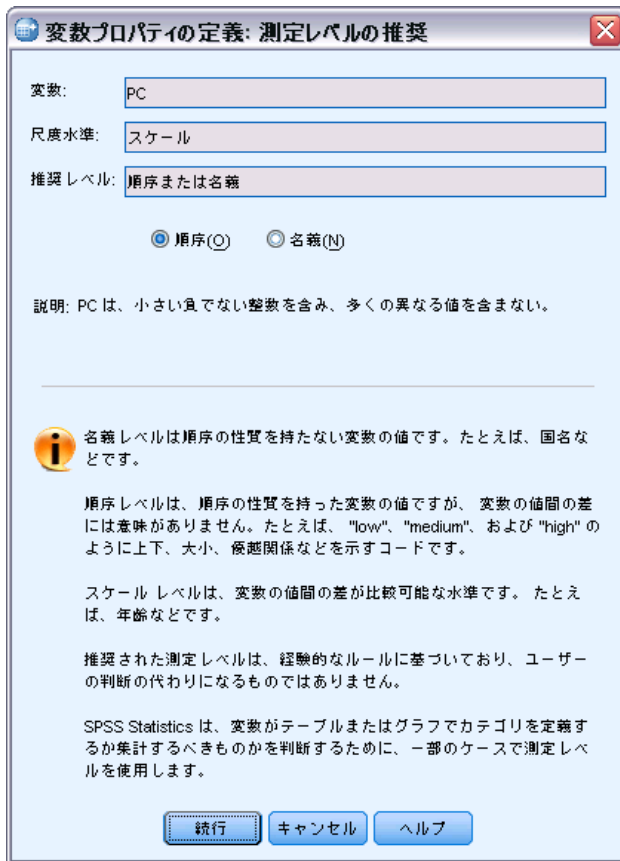
- ▶ [続行] をクリックします。

図 3-21  
[変数プロパティの定義] メイン ダイアログ ボックス



- ▶ [スキャンされた変数のリスト] で、「PC」を選択します。  
選択した変数の現在の尺度は [スケール] です。尺度はドロップダウン リストから選択できます。さらに、推奨される尺度を調べることもできます。
- ▶ [推奨] をクリックします。  
[尺度の示唆] ダイアログ ボックスが表示されます。

図 3-22  
[測定レベルの推奨] ダイアログ ボックス



この変数には多数の異なる値が含まれておらず、スキャンされたケースのすべてに整数値が含まれているので、順序または名義が即低水準として適切であることが示唆されます。

- ▶ [順序] を選択し、[続行] をクリックします。

選択した変数の尺度が [順序] に設定されました。

[変数ラベル グリッド] には、選択されている変数のすべての一意なデータ値、値に対して定義されている値ラベル、スキャンされたケース内でそれぞれの値が発生する回数（度数）が表示されます。

データ ビューに入力した値 99 もグリッドに表示されています。値を変更したのは 1 つのケースに対してだけなので、その度数は 1 と表示されています。さらに、値 99 に対しては値ラベルを定義していないので、[ラベル] 列は空白となっています。[スキャンされた変数のリスト] の 1 列目の x は、選択されている変数に、定義済みの値ラベルを持たない観測値が 1 つ以上含まれていることを示します。

- ▶ 値 99 の [ラベル] 列に、「回答なし」と入力します。
- ▶ 値 99 の [欠損値] 列のチェック ボックスを選択して、値 99 が「ユーザー欠損」であることを示します。  
ユーザー欠損値として指定されたデータ値は、特別に扱われるために区別され、ほとんどの計算から除外されます。

図 3-23  
[パソコン] に定義された新しい変数プロパティ

変数プロパティの定義

スキャンされた変数のリスト (C) 現在の変数: PC ラベル(L): PC

測定(M): 順序(O) 推奨(S) 型(T): 数値型

役割: 入力 幅(W): 4 小数桁(D): 0

ラベルなしの値: 0 属性(B)...

変数ラベルグリッド(V): グリッドでラベルの追加または編集をしてください。下部で追加する値を入力できます。

	変更済	欠損値	度数	値	ラベル
1	<input type="checkbox"/>	<input type="checkbox"/>	3588	0	持っていない
2	<input type="checkbox"/>	<input type="checkbox"/>	2811	1	持っている
3	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	1	99	回答がない
4	<input type="checkbox"/>	<input type="checkbox"/>			

スキャンされたケース: 6400 プロパティをコピー

値リストの制限: なし 別の変数から(F)... 他の変数へ(O)... ラベルなし変数 自動ラベル(A)

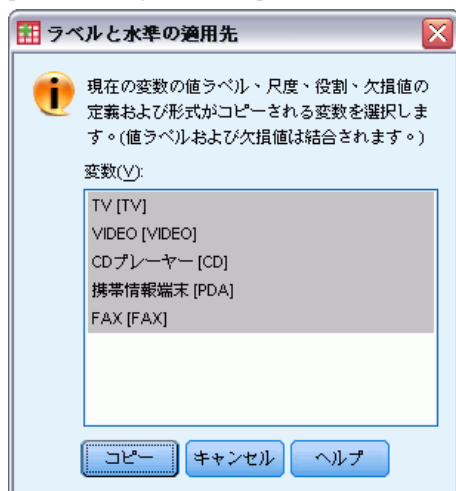
OK 貼り付け(P) 戻す(R) キャンセル ヘルプ

「パソコン」変数プロパティの変更作業を終わりにする前に、同じ尺度、値ラベル、欠損値の定義をリストの他の変数にも適用しましょう。

- ▶ [プロパティをコピー] 領域の [他の変数へ] をクリックします。



図 3-24  
[ラベルと水準の適用先] ダイアログ ボックス



- ▶ [ラベルと水準の適用先] ダイアログ ボックスで、リスト内の変数をすべて選択し、[コピー] をクリックします。

ここで、[変数プロパティの定義] メイン ダイアログ ボックスの [スキャンされた変数リスト] で他の変数を選択してみてください。これらはすべて順序変数で、値 99 がユーザー欠損値として設定され、値ラベルが [回答なし] と定義されているはずです。

図 3-25  
[fax] に定義された新しい変数プロパティ



- ▶ [OK] をクリックして、定義したすべての変数プロパティを保存します。

# 多重データソースの使用方法

バージョン 14.0 を開始すると、多重データ ソースを同時に開くことにより、以下のことが簡単に実行できます

- データ ソース間を前後に切り替え。
- 異なるデータ ソースの内容の比較。
- データ ソース間のデータのコピーと貼り付け。
- 分析用のケースまたは変数、あるいはその両方の複数のサブセットを作成。
- 最初に各データを保存することなく、様々なデータ形式（たとえば、スプレッドシート、データベース、テキスト データ）から多重データ ソースを結合。

## 多重データソースの基本的な処理

図 4-1  
同時に開いている 2 つのデータソース

The image shows two overlapping SPSS Data Editor windows. The top window displays data for 'demo.sav' with the following table:

	年齢	婚姻状況	住居年数	収入	収入カテゴリ	車	車カテゴリ
1	55	1	12	72.00	3.00	36.20	3.0
2	56	0	29	153.00	4.00	76.90	3.0

The bottom window displays data for 'survey\_sample.sav' with the following table:

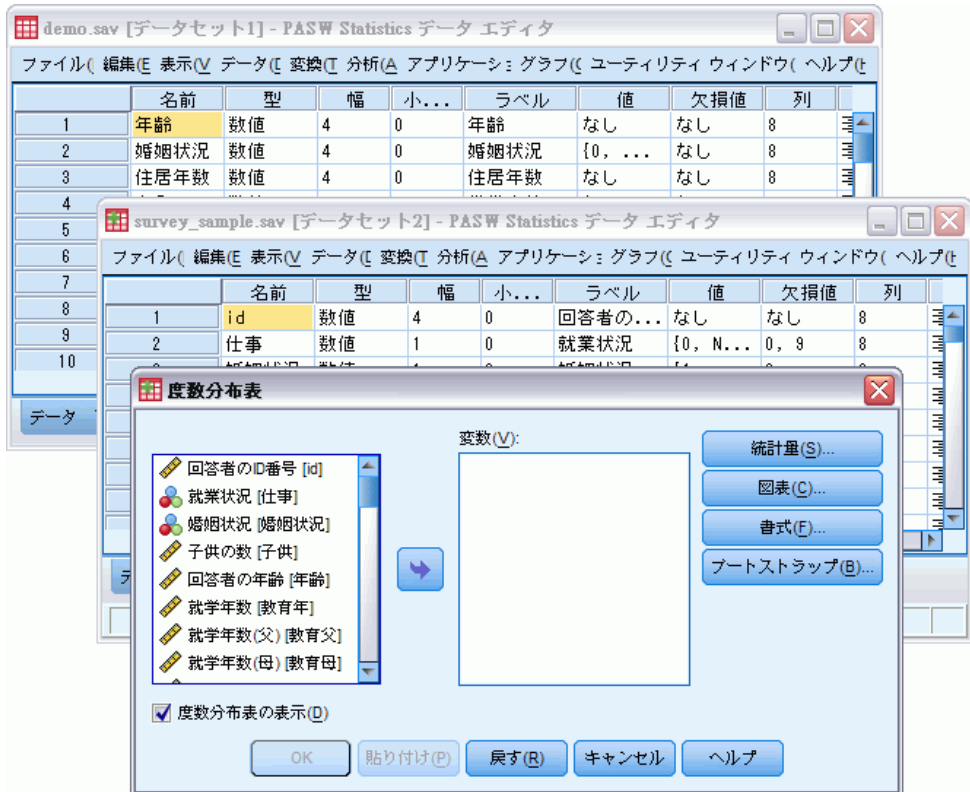
	id	仕事	婚姻状況	子供	年齢	教育年	教育父
1	1	1	3	2	60	12	1
2	2	2	5	0	27	17	2
3	3	1	1	2	36	12	1
4	4	1	5	0	21	13	9
5	5	1	5	0	35	16	9
6	6	1	3	1	33	16	
7	7	1	4	0	43	12	1
8	8	1	5	0	29	13	1
9	9	2	1	2	39	18	1

デフォルトでは、開いたデータソースがそれぞれ新しい [データ エディタ] ウィンドウに表示されます。

- すでに開いているデータソースは、開いたままさらに使用できます。
- データソースは、最初に開いた時点で、自動的に **アクティブなデータセット** に変わります。
- 使用するデータソースの [データ エディタ] ウィンドウの任意の場所をクリックするだけで、または [ウィンドウ] メニューからそのデータソースの [データ エディタ] ウィンドウを選択することにより、作業データセットを変更することができます。

- 作業データセットの変数のみが分析に使用できます。

図 4-2  
作業データセットの変数を含む変数リスト



- データにアクセスするダイアログ ボックスが開いているときは、作業データセットを変更することはできません(変数リストを表示するすべてのダイアログ ボックスを含む)。
- セッション中には、少なくとも 1 つの [データ エディタ] ウィンドウが開いていなければなりません。開いている最後の [データ エディタ] ウィンドウを閉じると、IBM® SPSS® Statistics は自動的にシャットダウンし、最初に変更の保存を要求するダイアログ ボックスが表示されます。

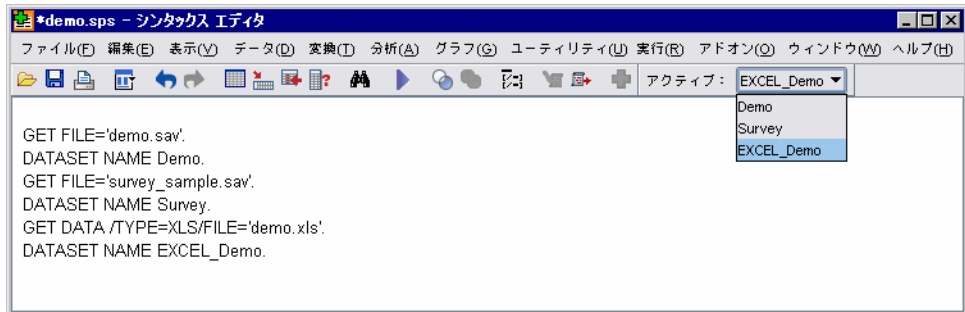
## コマンド シンタックスの複合データセットを使用する作業

コマンド シンタックスを使ってデータ ソースを開く場合 (たとえば、GET FILE、GET DATA)、複数のデータ ソースを同時に開くためには各データセットに対して明示的に名前を指定する DATASET NAME を使用する必要があります。

コマンド シンタックスを使用すると、アクティブなデータセット名がシンタックス ウィンドウのツールバーに表示されます。次のアクションでアクティブ データセットが変更します。

- DATASET ACTIVATE コマンドの使用。
- データセットの [データ エディタ] のどこでもクリック。
- シンタックス ウィンドウのツールバーからデータセット名を選択。

図 4-3  
シンタックス ウィンドウのツールバーに表示されたデータセットを開く



## データセット間での情報のコピーと貼り付け

単独のデータ ファイル内で情報をコピーして貼り付けるのと基本的に同じ方法でデータと変数定義属性の両方を 1 つのデータセットから別のデータセットにコピーすることができます。

- 選択したデータ セルを [データ ビュー] にコピーして貼り付けると、データ値だけが貼り付けられ、変数定義属性は貼り付けられません。
- 列の 1 番上の変数の名前を選択して [データ ビュー] の変数全体をコピーして貼り付けると、すべてのデータおよびその変数のすべての変数定義属性が貼り付けられます。
- 変数定義属性または [変数ビュー] 内の変数全体をコピーして貼り付けると、選択された属性（すなわち変数定義全体）は貼り付けられますが、データ値が貼り付けられることはありません。

## データセットの名前の変更

メニューやダイアログ ボックスを使用してデータ ソースを開くと、それぞれのデータ ソースには自動的に DataSetn というデータセット名が割り当てられます。ここで、n は、連続する整数値 を表します。ただし、コマンド シンタックスを使ってデータ ソースを開くときは、DATASET NAME でデータセット名を明示的に指定しないと、データセット名は割

り当てられません。よりわかりやすいデータセット名を付けるには、以下のようにします。

- ▶ 名前を変更するデータセットの [データ エディタ] ウィンドウのメニューから、以下を選択します。  
ファイル > データセットの名前の変更...
- ▶ 変数命名規則に従った新しいデータセット名を入力します。

## 複合データセットを抑制

複合データセット機能を抑制し、一度に 1 個のデータセットだけを利用したいときは次のようにします。

- ▶ メニューから次の項目を選択します。  
編集 > オプション...
- ▶ [全般] タブをクリックします。

[一度に 1 個のデータセットを開く] を選択 (クリック) します。

# 個々の変数に対する要約統計量の検討

この章では、単純な要約測度、および使用する統計量の種類に対する変数の尺度の影響について説明します。ここでは、データ ファイル `demo.sav` を使用します。詳細は、A 付録 p.165 サンプル ファイル を参照してください。

## 尺度レベル

データの種類が異なると、尺度に応じて、適切な要約統計量も異なります。

**カテゴリ。** 一定の数の異なる値またはカテゴリ（たとえば、性別や配偶者の有無）から成るデータです。**質的データ**とも呼ばれます。カテゴリ変数は、文字型データまたはカテゴリ（たとえば、0 = 未婚および 1 = 既婚）を表す数値コードを使用する数値型変数です。カテゴリ データには、次に示す 2 つの基本的な種類があります。

- **名義。** カテゴリに固有の順序がない場合のカテゴリ データです。たとえば、[営業] の職種カテゴリは、[マーケティング] または [研究] の職種カテゴリより上位でも下位でもありません。
- **順序。** カテゴリに意味のある順序がある場合のカテゴリ データです。ただし、カテゴリ間には測定可能な距離はありません。たとえば、高、中、低という値には順序がありますが、これらの値の間の「距離」は計算できません。

**スケール。** 間隔または比率尺度について測定したデータです。ただし、データ値は、値の順序と値の間の距離を示します。たとえば、\$72,195 の給料は、\$52,398 の給料より高く、2 つの値の距離は \$19,797 です。また、**量的データ**、**連続型データ**とも呼ばれます。

## カテゴリ データの要約統計量

カテゴリ データについて、最も典型的な要約統計量は、各カテゴリのケース数またはケースのパーセントです。**最頻値**は、最大のケース数を含むカテゴリです。順序データについて、カテゴリが多数ある場合は、**中央値**（この値を中心にケースが上下半分に分かれます）も、有用な要約統計量になる場合があります。



[度数分布表] 手続きにより、変数の各観測値についてのケース数とケースのパーセントの両方を示す度数分布表が作成されます。

- ▶ メニューから次の項目を選択します。  
分析(A) > 記述統計 > 度数分布表...

注：この機能は Statistics Base オプションが必要です。

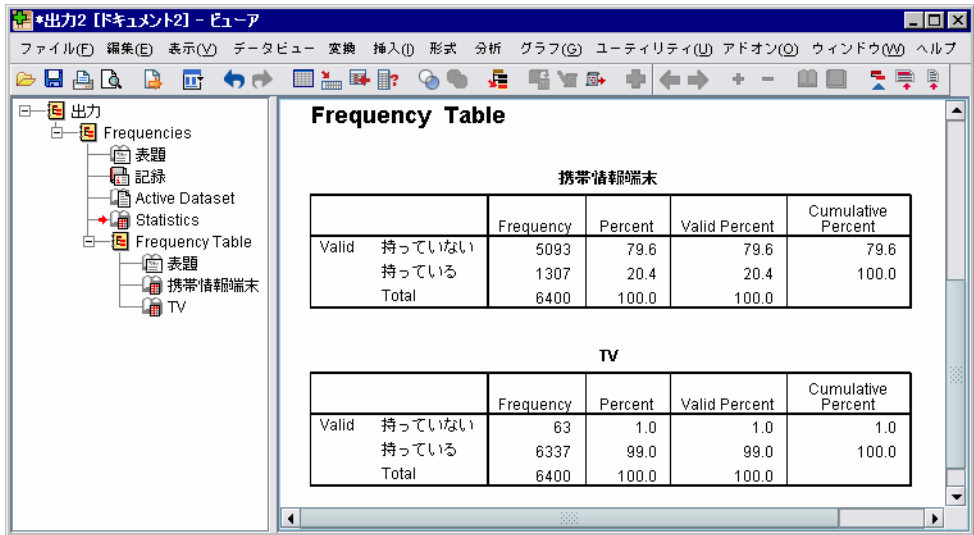
- ▶ [携帯情報端末 [PDA]] と [TV [TV]] を選択し、[変数] リストに移動します。

図 5-1  
分析のために選択したカテゴリ変数



- ▶ [OK] をクリックして手続きを実行します。

図 5-2  
度数分布表



度数分布表は、[ビューア] ウィンドウに表示されます。度数分布表では、わずか 20.4% の人だけが PDA を所有している一方で、ほとんどすべての人 (99.0%) が TV を所有していることが明らかです。これは、それほど興味深い発見ではないかもしれませんが、むしろ、テレビを所有していない人の小グループについて、より詳しく調べるの方が興味深いかもしれません。

## カテゴリ データのグラフ

棒グラフまたは円グラフとともに度数分布表に情報を図示することができます。

- ▶ [度数分布表] ダイアログ ボックスを再び開きます。(変数は、ここでも 2 つ選択されている必要があります)。

ツールバーの [ダイアログ リコール] ボタンを使用すると、最近使用した手続きにすばやく戻ることができます。

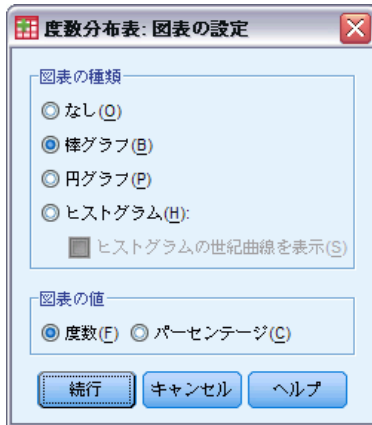
図 5-3  
[ダイアログ リコール] ボタン



- ▶ [図表] をクリックします。

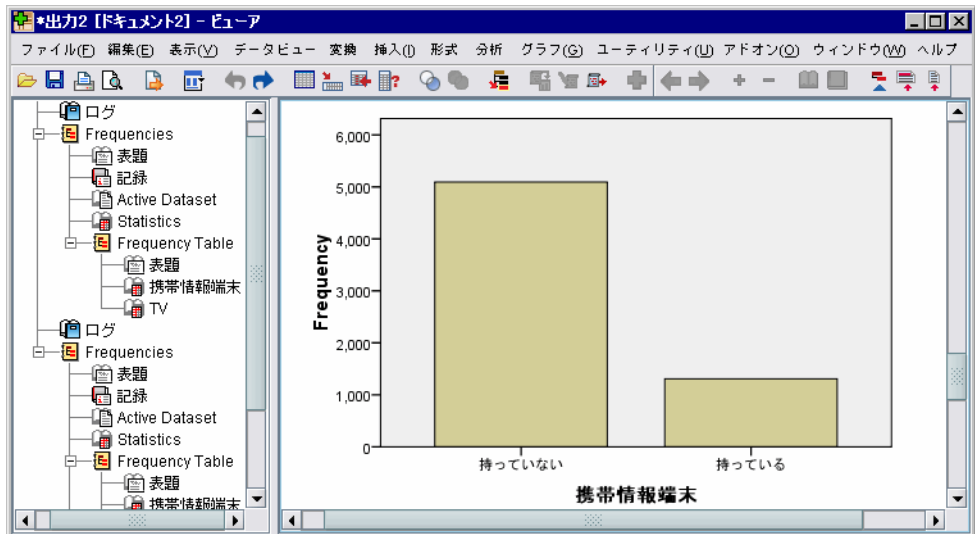
- ▶ [棒グラフ] を選択し、[続行] をクリックします。

図 5-4  
[度数分布表: 図表の設定] ダイアログ ボックス



- ▶ メイン ダイアログ ボックスの [OK] をクリックし、手続きを実行します。

図 5-5  
棒グラフ



これで、度数分布表の他に、同じ情報が棒グラフの形で表示され、ほぼすべての人が TV を所有しているのに対し、ほとんどの人が PDA を所有していないことが分かります。

## スケール変数の要約統計量

スケール変数には、多くの要約統計量を使用することができます。たとえば、次のものが含まれます。

- **中心傾向の測度。** 最も一般的な中心傾向の測度は、**平均値**（算術平均）と **中央値**（この値を中心にケースが上下半分に分かれます）です。
- **散らばりの測定。** データの変化量または広がり量を測定する統計としては、**標準偏差**、**最小値**、**最大値**があります。

- ▶ [度数分布表] ダイアログ ボックスを再び開きます。
- ▶ [戻す] をクリックし、これまでの設定をすべてクリアします。
- ▶ [世帯全体の収入カテゴリ (千ドル) [収入カテゴリ]] を選択し、[変数] リストに移動します。

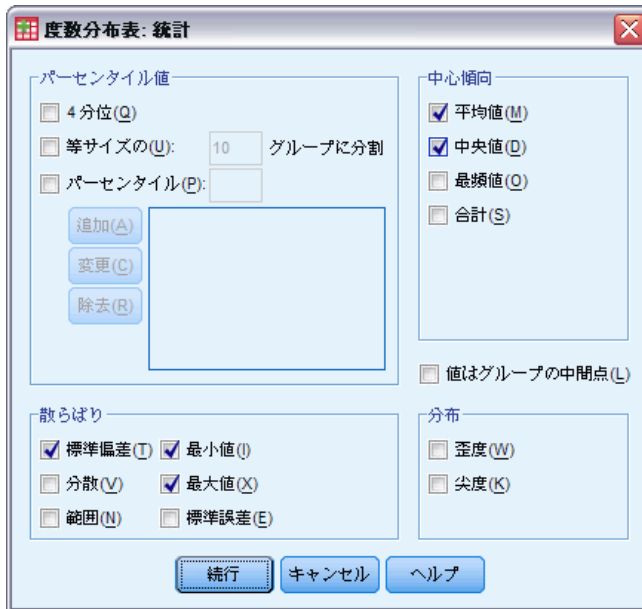
図 5-6  
分析のために選択したスケール変数



- ▶ [統計量] をクリックします。

- ▶ [平均値]、[中央値]、[標準偏差]、[最小値]、[最大値] を選択します。

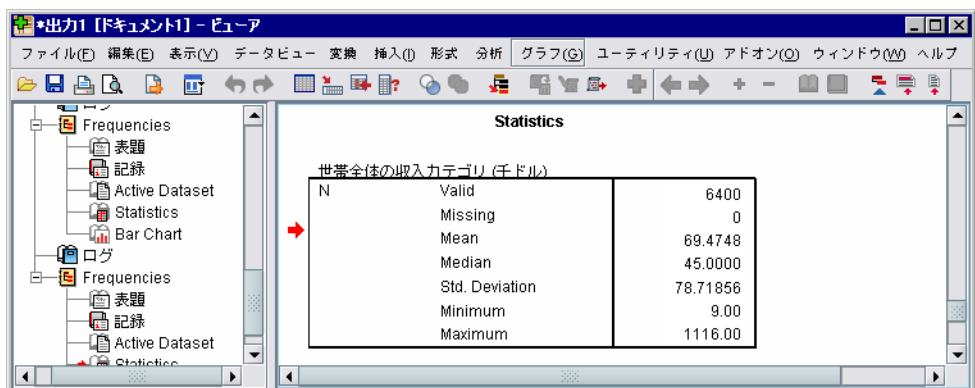
図 5-7  
[度数分布表: 統計] ダイアログ ボックス



- ▶ [続行] をクリックします。
- ▶ メイン ダイアログ ボックスで [度数分布表の表示] の選択を解除します。  
(異なる値が、データ ファイルにあるケースの数とほぼ同数存在することがあるので、度数分布表は、通常、スケール変数に対して特に有用なわけではありません)。
- ▶ [OK] をクリックして手続きを実行します。

度数分布表の統計量テーブルが、[ビューア] ウィンドウに表示されます。

図 5-8  
度数分布表の統計テーブル



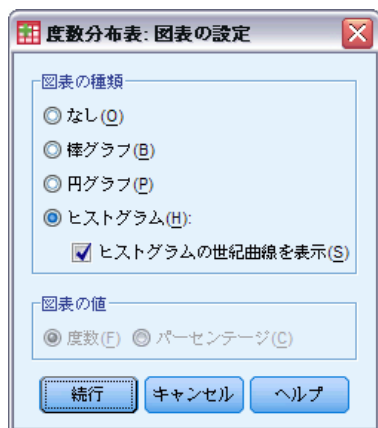
この例では、平均値と中央値との間に、大きな差があります。平均値は中央値より、25,000 近くも大きく、このことは、値が正規分布ではないことを示しています。ヒストグラムを使用すると、分布を視覚的に確認することができます。

## スケール変数のヒストグラム

- ▶ [度数分布表] ダイアログ ボックスを再び開きます。
- ▶ [図表] をクリックします。
- ▶ [ヒストグラム] と [正規曲線付き] を選択します。

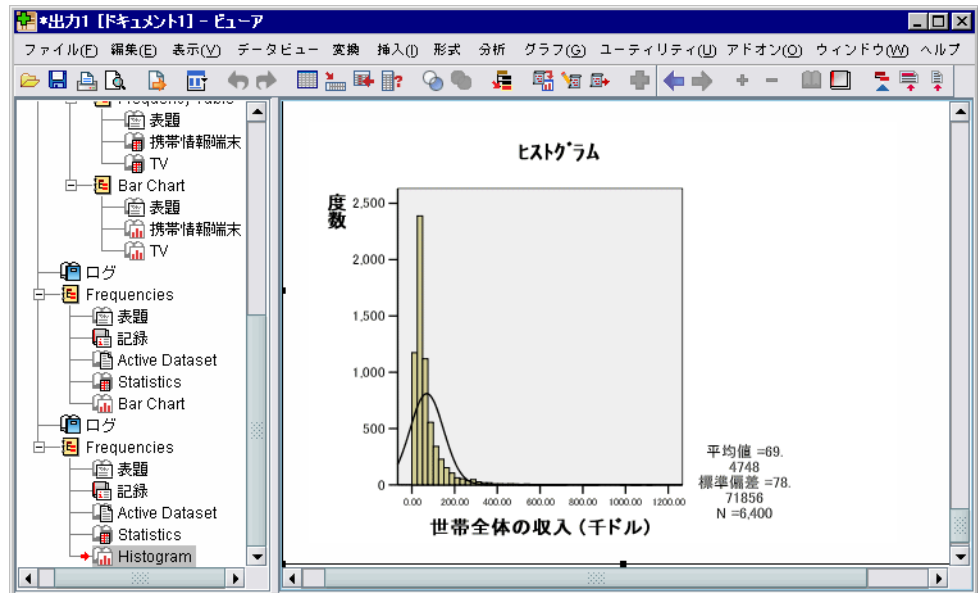
図 5-9

[度数分布表: 図表の設定] ダイアログ ボックス



- ▶ [続行] をクリックし、メイン ダイアログ ボックスの [OK] をクリックして、手続きを実行します。

図 5-10  
ヒストグラム(S)



ケースの大部分は、スケールの下限でクラスタ化され、多くは 100,000 未満となります。ただし、500,000 の範囲およびそれ以上にあるケースが少数あります（あまりに少数であるため、表示するにはヒストグラムの修正が必要です）。これらの非常に高い値は少数のケースに限られていても、平均値には有意な効果を持ちますが、中央値にはほとんど、あるいはまったく効果を持ちません。そのため、この例での中心傾向の指標としては、中央値がより適切になります。

# グラフの作成と編集

さまざまな種類の図表を作成および編集できます。この章では、棒グラフの作成と編集を行います。その他のタイプのグラフについても、ここで説明する要領で作成と編集を行えます。

## 図表作成の基本

図表作成の基本を具体的に説明するために、さまざまなレベルの仕事満足度に対応した平均収入の棒グラフを作成します。この例では、データファイル `demo.sav` を使用します。詳細は、[A 付録 p.165 サンプルファイル](#) を参照してください。

- ▶ メニューから次の項目を選択します。  
グラフ(G) > 図表ビルダー(C)...



[図表ビルダー] ダイアログ ボックスは、グラフの作成時に、完成状態をプレビューすることができる、インタラクティブなウィンドウです。

図 6-1  
[図表ビルダー] ダイアログ ボックス



## 図表ビルダー ギャラリーの使用

- ▶ [ギャラリー] タブ (がまだ選択されていなければ) をクリックします。

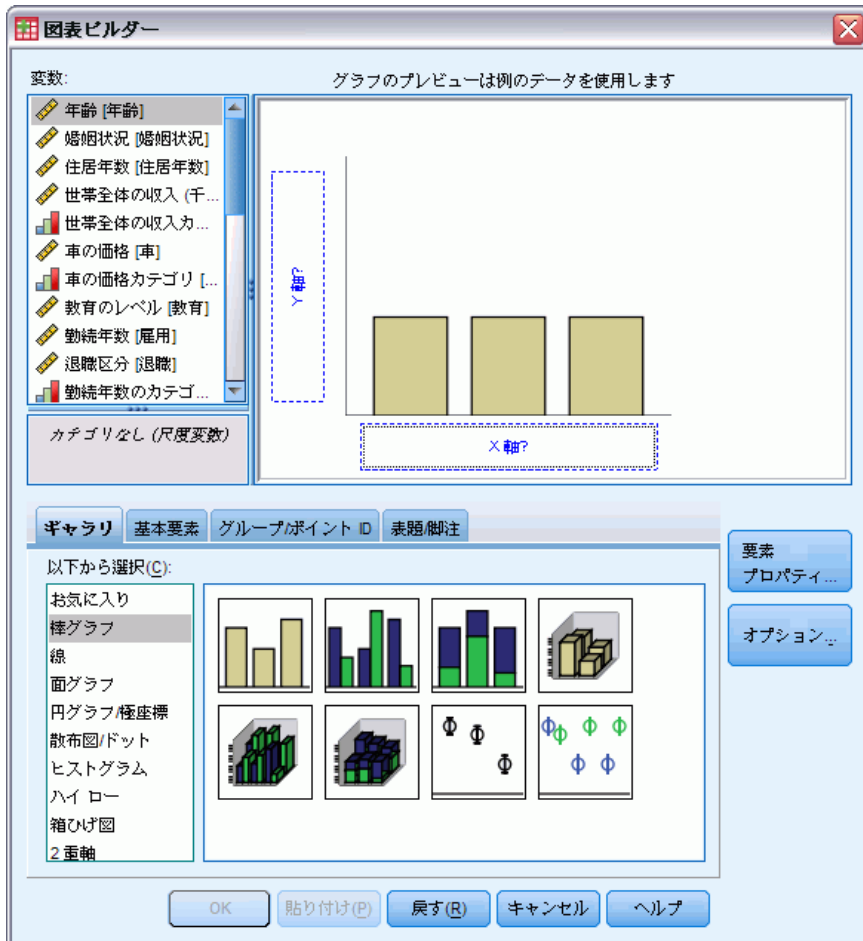
ギャラリーには、さまざまな定義済みのグラフが含まれており、グラフの種類ごとに分類されています。[基本要素] タブには、図表を作成するための基本的な要素 (軸やグラフ要素など) が用意されていますが、[ギャラリー] を使用するほうがより簡単です。

- ▶ [棒グラフ] (がまだ選択されていないならば) をクリックします。

ギャラリー内の利用可能な棒グラフを示すアイコンがダイアログボックスに表示されます。このアイコンの絵には、グラフの種類を特定するのに十分な情報が含まれているはずですが、さらに情報が必要な場合には、アイコンの上にカーソルを移動して、グラフのツールヒントを表示することもできます。

- ▶ 単純棒グラフのアイコンを「キャンバス」にドラッグします。「キャンバス」というのは、ギャラリーの上方にある広い領域のことです。[図表ビルダー] のキャンバスにグラフのプレビューが表示されます。このグラフの描画に使われるデータは、実際のデータではないことに注意してください。これはサンプルデータです。

図 6-2  
[図表ビルダー] のキャンバス上の棒グラフ



## 変数と統計量の定義

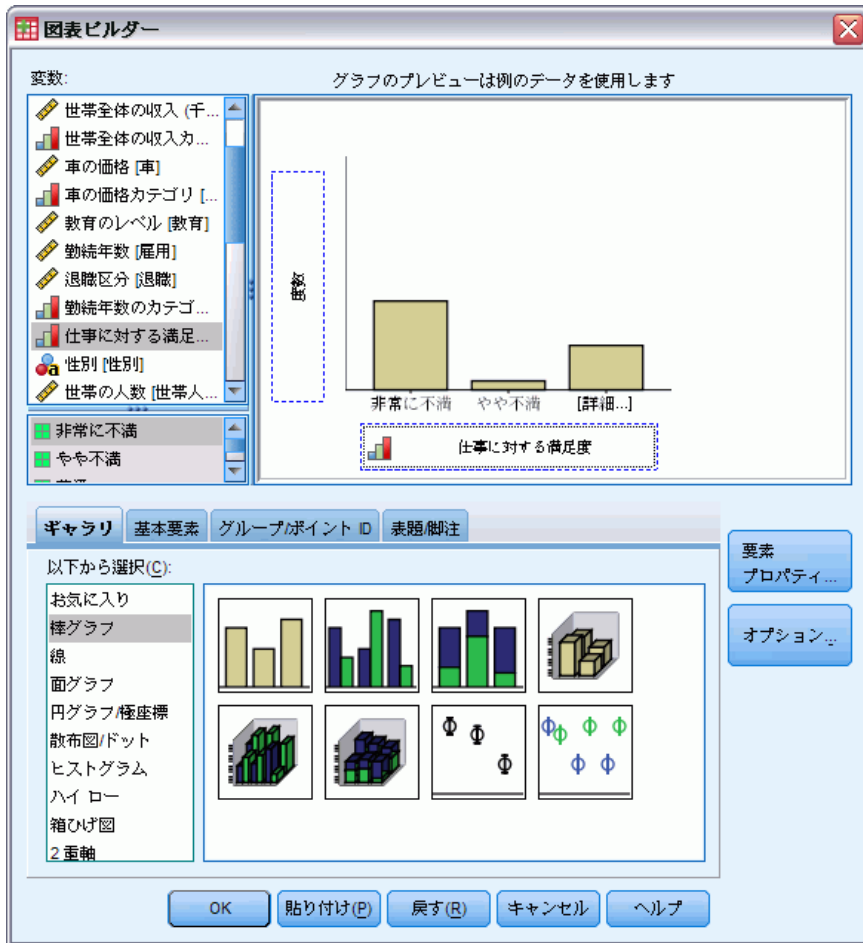
キャンバス上にはグラフが表示されていますが、棒の高さを制御したり、各棒に対応する変数カテゴリを指定する変数や統計量がないので、このグラフはまだ完全ではありません。変数や統計量がない場合は、グラフは作成できません。変数は、キャンバスの左側にある変数リストからドラッグすることにより追加できます。

図表ビルダーでは、変数の尺度が重要です。ここでは、x 軸に対し、[仕事に対する満足度] 変数を使おうとしています。けれども、変数の隣の(定規のような)アイコンは、この変数の尺度がスケールに指定されていることを示しています。正しいグラフを作成するには、尺度がカテゴリであるものを使う必要があります。変数ビューに戻って尺度を変更するかわりに、図表ビルダーで一時的に尺度を変更することもできます。

- ▶ 変数リストの中の [仕事に対する満足度] を右クリックして、[順序] を選択します。[仕事に対する満足度] のカテゴリは、満足度のレベルによって、順位付けが可能なので、[順序] という尺度は適切です。尺度を変更すると、アイコンも変化するということに注意します。

- ▶ 今度は、変数リストから [仕事に対する満足度] を x 軸のドロップ領域にドラッグします。

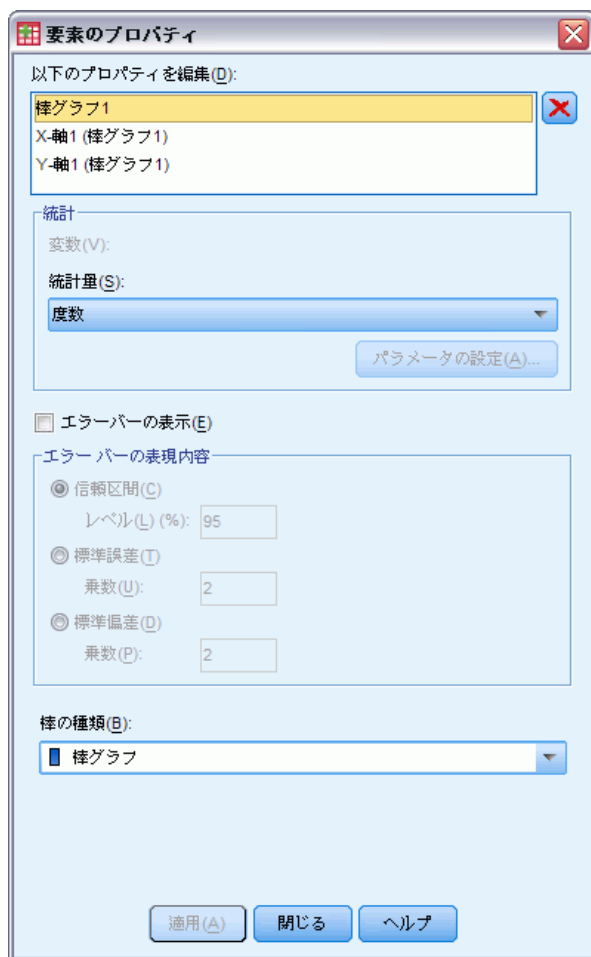
図 6-3  
x 軸のドロップ領域に入った [仕事に対する満足度]



デフォルトでは、y 軸のドロップ領域は、[度数] 統計量になっています。これ以外の統計量（パーセンテージや平均）を利用したい場合には、簡単に変更することができます。この例ではどちらの統計量も使わないのですが、統計量を変更する必要が生じたときのために、その手順を確認しておきます。

- ▶ [要素のプロパティ] をクリックして、[要素のプロパティ] ウィンドウを表示します。

図 6-4  
[要素のプロパティ] ウィンドウ



[要素のプロパティ] ウィンドウでは、さまざまなグラフ要素のプロパティを変更することができます。グラフ要素（棒グラフの棒）やグラフの軸などがこうした要素に含まれます。[プロパティの編集] から要素を選択して、その要素に関連するプロパティを変更します。また、リストの右側にある赤い X に注意します。このボタンを押すと、キャンバスからグラフ要素が削除されます。Bar1 が選択されているため、表示されるプロパティは、グラフ要素、具体的には棒グラフ要素に適用されます。

[統計量] ドロップダウン リストには、利用できる統計量が表示されます。同じような統計量が、通常どの種類のグラフでも利用できます。ただし一部の統計量については、y 軸のドロップ領域に変数が存在する必要があります。

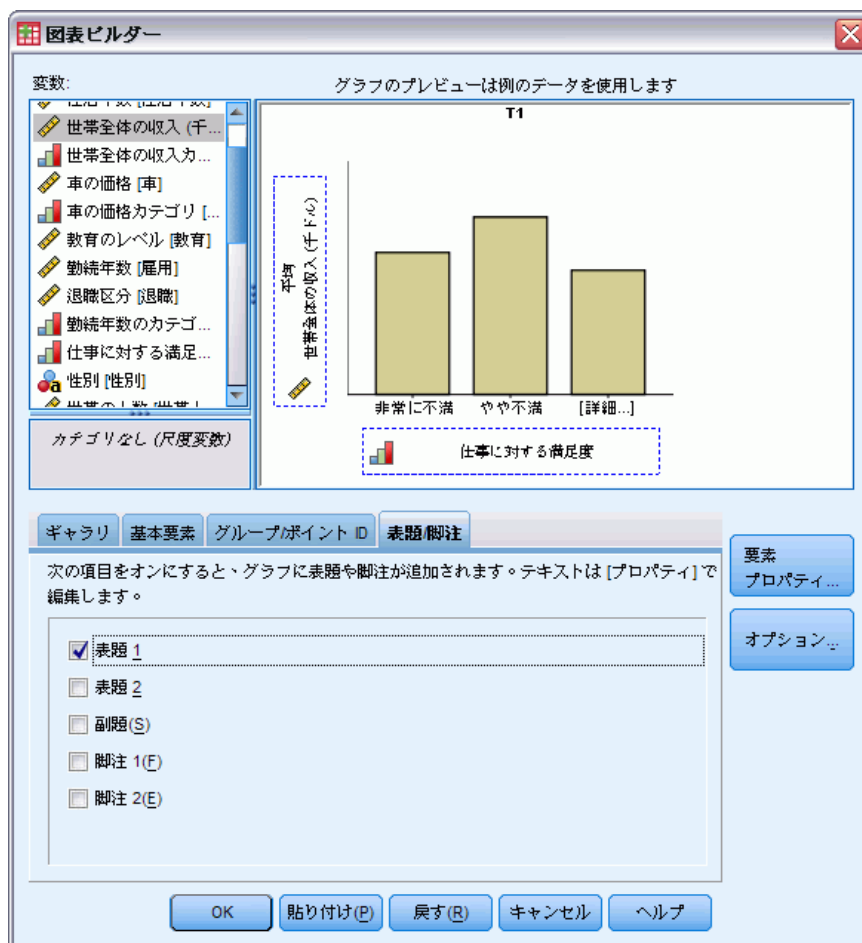
- ▶ [図表ビルダー] ダイアログ ボックスに戻って、[世帯全体の収入] を変数リストから y 軸のドロップ領域にドラッグします。y 軸の変数はスカラーで、x 軸の変数はカテゴリ変数（順序はカテゴリ測定レベルの一種）なので、デフォルトでは、y 軸のドロップ領域は [平均] 統計量になります。変数や統計量はこのままでよいので、要素のプロパティを変更する必要はありません。

## テキストの追加

グラフには、表題や脚注を追加することもできます。

- ▶ [表題/脚注] タブをクリックします。
- ▶ [表題 1] を選択します。

図 6-5  
キャンバスに表示された [表題 1]



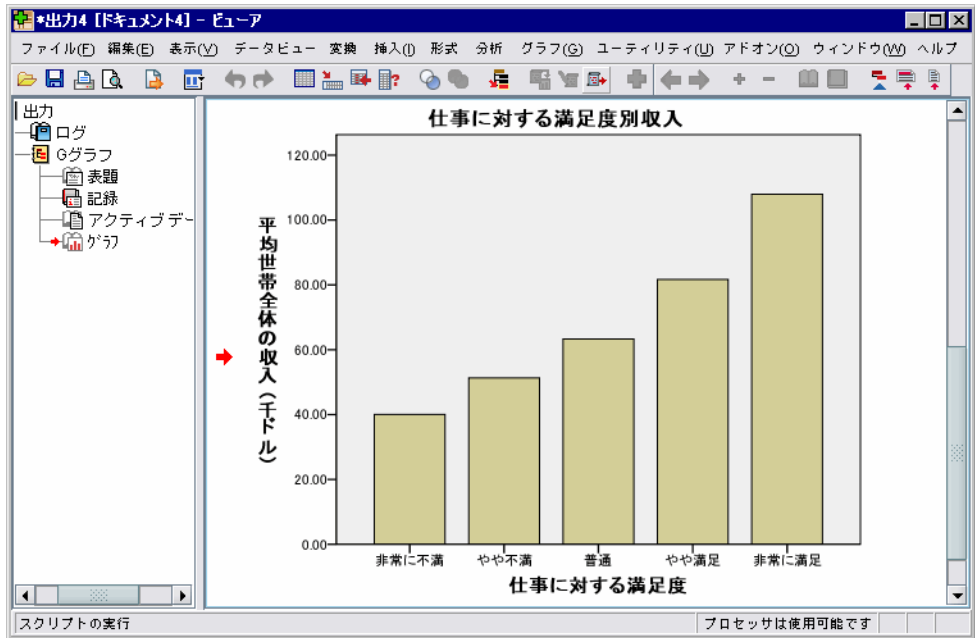
[T1] というラベルのついた表題がキャンバスに表示されます。

- ▶ [要素のプロパティ] ウィンドウの、[プロパティの編集] から、[表題 1] を選択します。
- ▶ [内容] テキストボックスに、「仕事に対する満足度別収入」と入力します。これは、表題として表示されるテキストです。
- ▶ [適用] をクリックしてテキストを保存します。このテキストは [図表ビルダー] には表示されませんが、グラフの生成時には表示されます。

## 図表を作成する

- ▶ [OK] をクリックして、棒グラフを作成します。

図 6-6  
棒グラフ



この棒グラフでは、仕事に対する満足度の高い回答者ほど、世帯収入も高くなる傾向にあることがわかります。

## 図表編集の基本

さまざまな方法でグラフを編集することができます。作成した棒グラフのサンプルに対して、次の操作を行います。

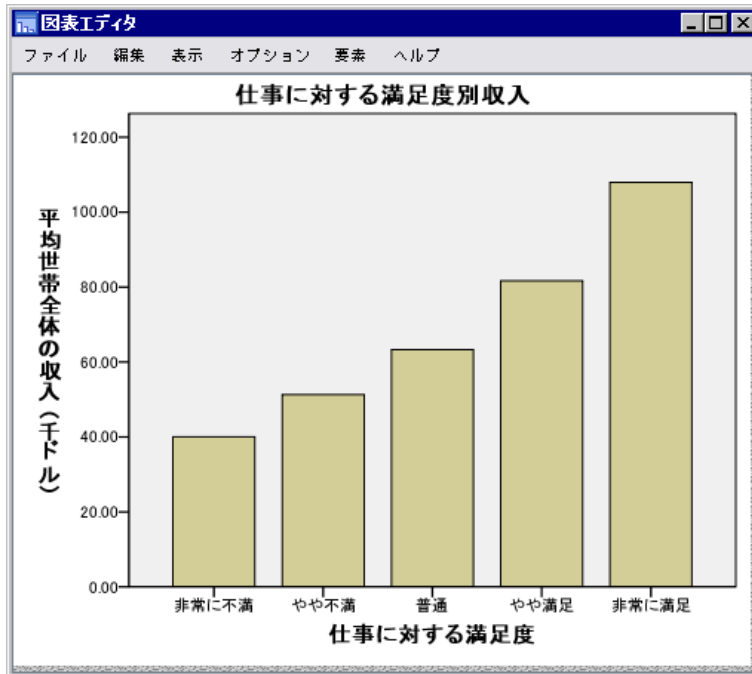
- 色を変更する。
- 目盛りラベルに表示される数値の書式を設定する。
- テキストを編集する。
- データ値ラベルを表示する。
- 図表テンプレートを使用する。

グラフを編集するには、図表エディタでグラフを開きます。



- ▶ 棒グラフをダブルクリックして、図表エディタで開きます。

図 6-7  
図表エディタで開いた棒グラフ



## グラフ要素の選択

図表要素を編集するには、最初にその要素を選択します。

- ▶ いずれか 1 つの棒をクリックします。棒を囲む四角形は、その棒が選択されていることを示します。

単純なグラフで要素を選択する際には一般的な規則があります。

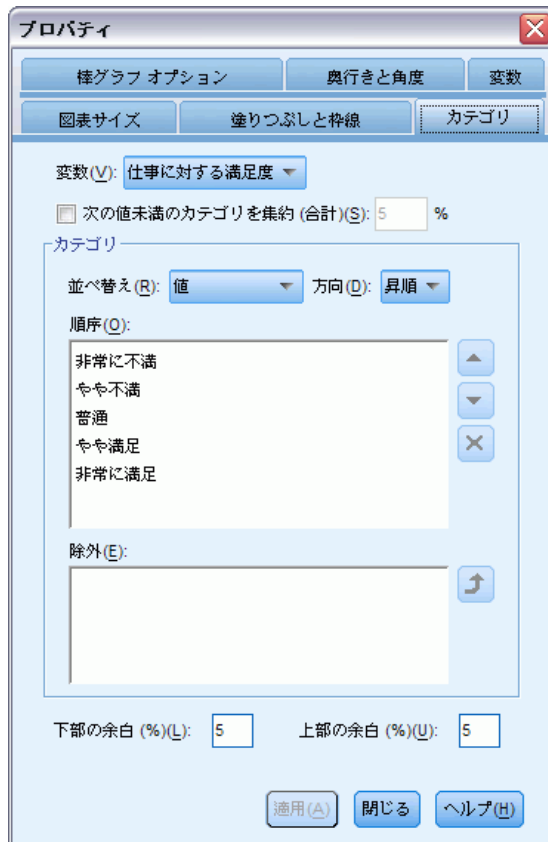
- グラフ要素が選択されていない場合、いずれかのグラフ要素をクリックすれば、すべてのグラフ要素を選択できます。
- すべてのグラフ要素が選択されている場合は、クリックしたグラフ要素だけを選択できます。別のグラフ要素を選択する場合は、そのグラフ要素をクリックします。複数のグラフ要素を選択するには、Ctrl キーを押しながら要素を 1 つずつクリックします。
- ▶ すべての要素の選択を解除するには、Esc キーを押します。
- ▶ 再びすべての棒を選択するには、任意の棒をクリックします。

## [プロパティ] ウィンドウの使用

- ▶ 図表エディタのメニューから次の項目を選択します。  
編集 > プロパティ

[プロパティ] ウィンドウが開き、選択した棒に適用するタブが表示されます。これらのタブは、図表エディタで選択する図表要素によって異なります。たとえば、棒ではなく、テキスト枠を選択した場合、[プロパティ] ウィンドウには別のタブが表示されます。これらのタブを使用して、図表の編集のほとんどを行います。

図 6-8  
[プロパティ] ウィンドウ



## 棒の色の変更

最初に、棒の色を変更します。[塗りつぶしと枠線] タブで、グラフ要素（線とマーカーを除く）の色属性を指定します。

- ▶ [塗りつぶしと枠線] タブをクリックします。
- ▶ [塗りつぶし] の横にある色見本をクリックして、棒の塗りつぶし色の変更を指定します。色見本の下にある数字は、現在の色に対応する赤、緑、および青の設定を示しています。
- ▶ 下から 2 行目の左から 2 番目の空色をクリックします。

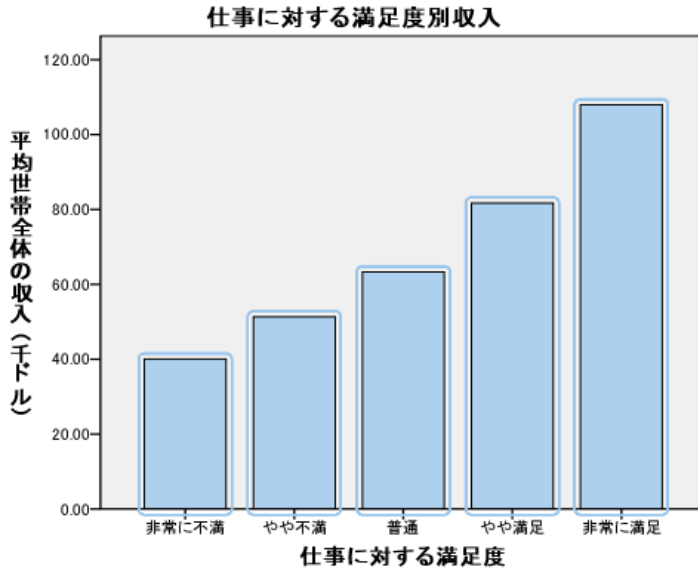
図 6-9  
[塗りつぶしと枠線] タブ



- ▶ [適用] をクリックします。

この時点で、このグラフ内の棒は空色になります。

図 6-10  
青い棒を表示する編集された棒グラフ



## 目盛りラベルに表示される数値の書式を設定する。

y 軸の数値の尺度が千単位であることに注意してください。グラフをより解釈しやすく優れたものにするには、目盛りラベルの数値の書式を変更して、軸の表題を編集します。

- ▶ y 軸の目盛りラベルのいずれかをクリックして、その目盛りラベルを選択します。
- ▶ [プロパティ] ウィンドウをすでに閉じている場合、再び開くにはメニューから次の項目を選択します。

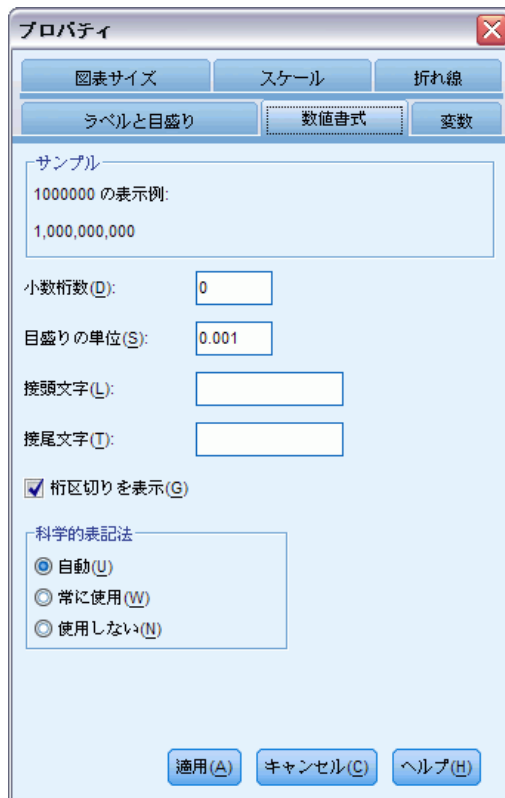
編集 > プロパティ

注：これ以降は、[プロパティ] ウィンドウが開いていることを前提とします。[プロパティ] ウィンドウを閉じた場合は、上記の手順に従って開いてください。また、キーボードのショートカット (Ctrl キーを押しながら T) を使用して、ウィンドウを再び開くこともできます。

- ▶ [数値書式] タブをクリックします。
- ▶ 目盛りラベルに小数桁数が表示されないように、[小数桁数] テキストボックスに「0」と入力します。

- ▶ [目盛りの単位] テキスト ボックスに「0.001」と入力します。目盛りの単位は、図表エディタで、表示される数値を除算する数値です。0.001 は小数なので、この数で除算すると、目盛りラベルに表示される数値は、1,000 ずつ増加 します。したがって、数値は千単位で表示されなくなります。つまり、数字は尺度化されません。
- ▶ [桁区切りを表示] を選択します。桁区切りでは、コンピュータのロケールで指定された文字を使用して、数値の 3 桁ごとにマークを付けます。

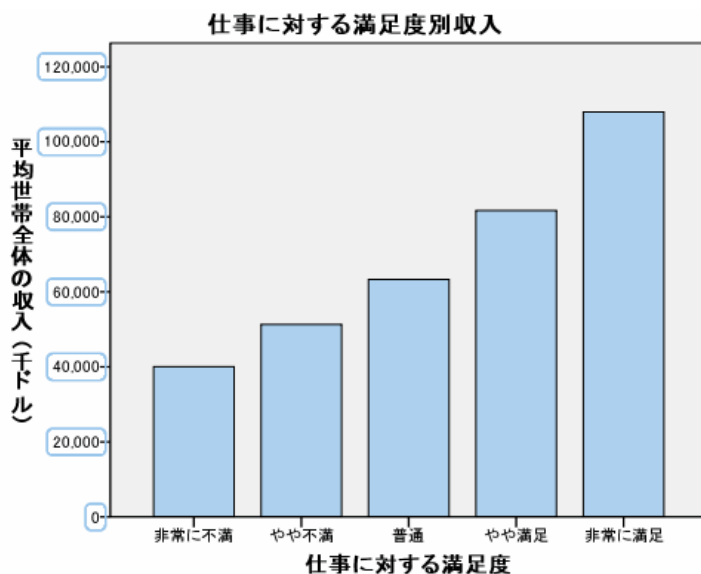
図 6-11  
[数値書式] タブ



- ▶ [適用] をクリックします。

目盛りラベルには、新しい数値の書式が反映されます。小数部分はなく、数値に目盛りは付かず、また 1000 ごとに文字が指定されています。

図 6-12  
新しい数値書式を表示する編集された棒グラフ



## テキストの編集

目盛りラベルの数値の書式を変更したので、軸の表題が正確ではなくなります。次に、新しい数値書式を反映するために、軸の表題を変更します。

注：テキストを編集するために、[プロパティ] ウィンドウを開く必要はありません。グラフ上で直接テキストを編集できます。

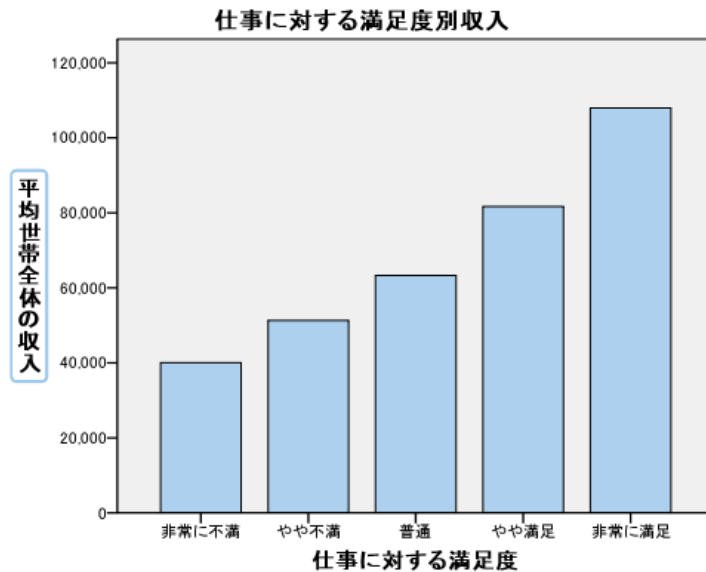
- ▶ y 軸の表題をクリックして選択します。
- ▶ 編集モードに入るには、軸の表題をもう一度クリックします。編集モードの間、図表エディタでは、テキストが回転して水平に配置されます。また、赤く点滅する棒カーソルも表示されます（例では表示されていません）。

- ▶ 次の手順を実行して、以下のテキストを削除します。

千ドル

- ▶ Enter キーを押して編集モードを終了し、軸の表題を更新します。この時点で、軸の表題は、目盛りラベルの内容を正確に示します。

図 6-13  
編集された y 軸の表題を表示する棒グラフ

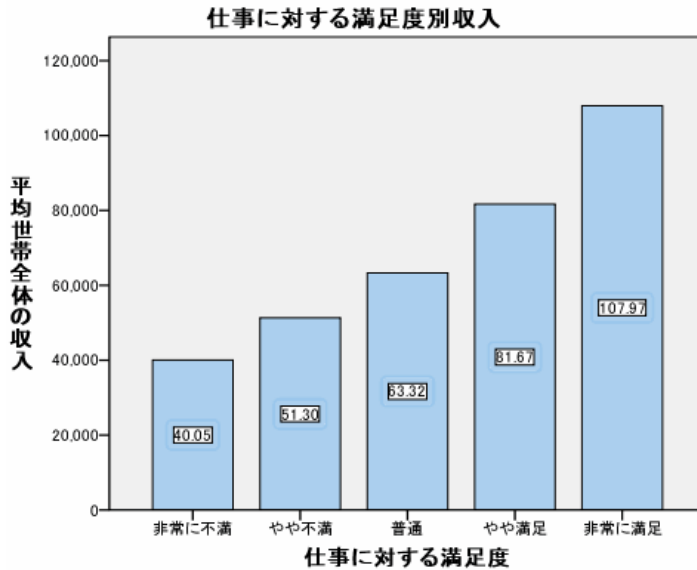


## データ値ラベルを表示する

通常行う作業としてはこれ以外に、正確な値をグラフ要素（この例では棒）に関連付けて表示するという作業があります。これらの値は、データラベルに表示されます。

- ▶ 図表エディタのメニューから次の項目を選択します。  
要素(M) > データラベルの表示

図 6-14  
データ値ラベルを表示する棒グラフ



グラフ内の各棒は、世帯収入の平均を正確に表示しています。単位は千単位であることに注意してください。再び [数値書式] タブを使用すれば、目盛りの単位を変更できます。

## テンプレートの使用

図表を何度も繰り返し変更する場合は、図表テンプレートを使用して、図表の作成および編集にかかる時間を短縮できます。図表テンプレートには、特定の図表の属性が保存されます。その後、図表を作成または編集する場合にそのテンプレートを適用できます。

現在の図表をテンプレートとして保存し、新しい図表を作成する際にそのテンプレートを適用できます。

- ▶ メニューから次の項目を選択します。  
ファイル(F) > 図表テンプレートの保存...(S)

[図表テンプレートの保存] ダイアログ ボックスでは、テンプレートに含める図表属性を指定できます。

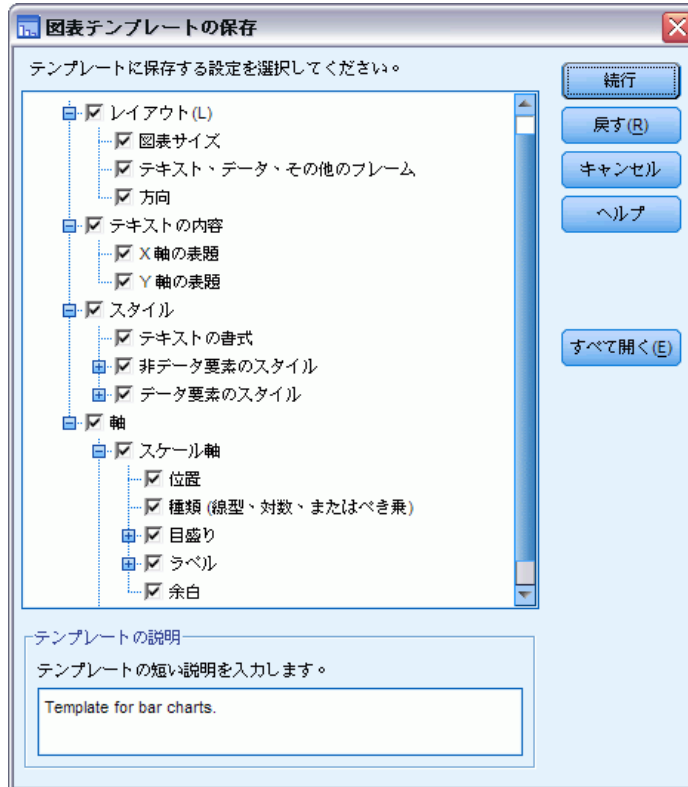
ツリー表示で項目のいずれかを開くと、図表と一緒に保存できる特定の属性を確認できます。たとえば、ツリーの [スケール軸] 部分を開くと、テンプレートに含めるデータ値ラベルのすべての属性を表示できます。任意の属性を選択して、テンプレートに含めることができます。



- ▶ 使用可能なすべての図表属性をテンプレートに含めるには、[すべての設定] を選択します。

テンプレートの説明を入力することもできます。この説明は、テンプレートを適用するときに表示されます。

図 6-15  
[図表テンプレートの保存] ダイアログ ボックス



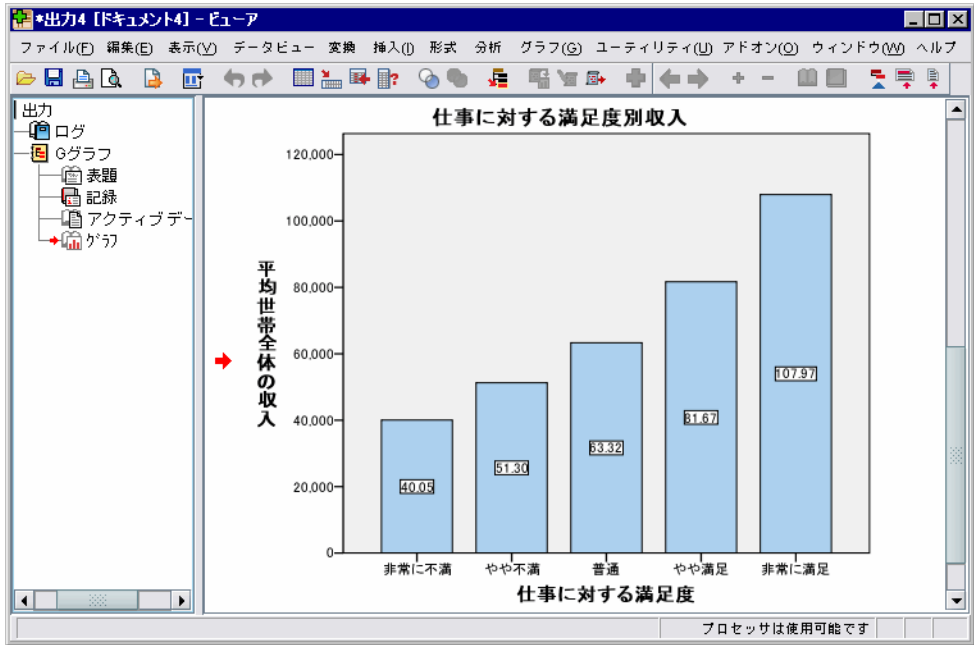
- ▶ [続行] をクリックします。
- ▶ [テンプレートの保存] ダイアログ ボックスで、テンプレートの場所とファイル名を指定します。
- ▶ 操作が終了したら、[保存] をクリックします。

図表の作成時、または図表エディタで、テンプレートを適用できます。次の例では、図表の作成中にテンプレートを適用します。

- ▶ 図表エディタを閉じます。更新された棒グラフがビューアに表示されます。

図 6-16

ビューアに表示される更新された棒グラフ



- ▶ ビューアのメニューから、次の項目を選択します。  
グラフ(G) > 図表ビルダー(C)...

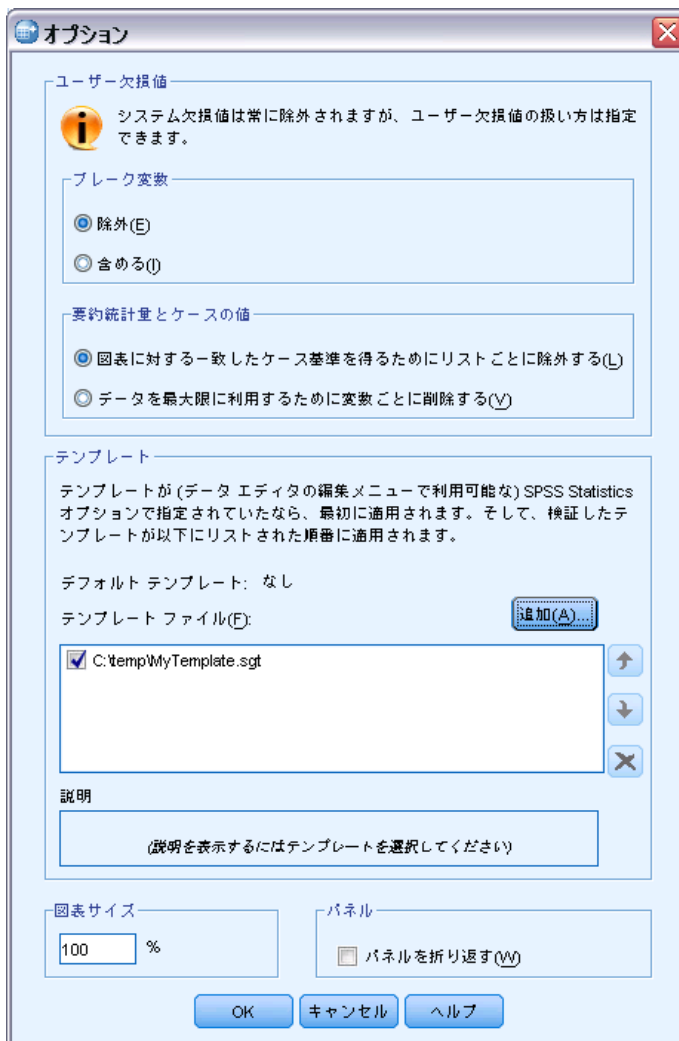
図表ビルダーは、元の図表の作成時に入力した変数を「記憶」しています。ただし、ここでは、テンプレートを適用したときに図表の書式がどのように設定されるかを確認するため、少し異なる図表を作成します。

- ▶ [仕事に対する満足度] をドロップ領域から変数リストにドラッグして、x 軸から削除します。ドロップ領域をクリックして、Delete キーを押してもかまいません。
- ▶ 変数リストの中の [教育のレベル] を右クリックして、[順序] を選択します。
- ▶ 変数リストから [教育のレベル] を x 軸のドロップ領域にドラッグします。  
今度は、表題が不正確になったので、削除します。
- ▶ [表題/脚注] タブで、[表題 1] の選択を解除します。  
今度は、新しいグラフに適用するテンプレートを指定します。
- ▶ [オプション] をクリックします。

- ▶ [オプション] ダイアログ ボックスの [テンプレート] グループで、[追加] をクリックします。
- ▶ [テンプレート ファイルの参照] ダイアログ ボックスで、すでに [図表テンプレートの保存] ダイアログ ボックスを使用して保存したテンプレート ファイルを指定します。
- ▶ そのファイルを選択して、[開く] をクリックします。

図 6-17

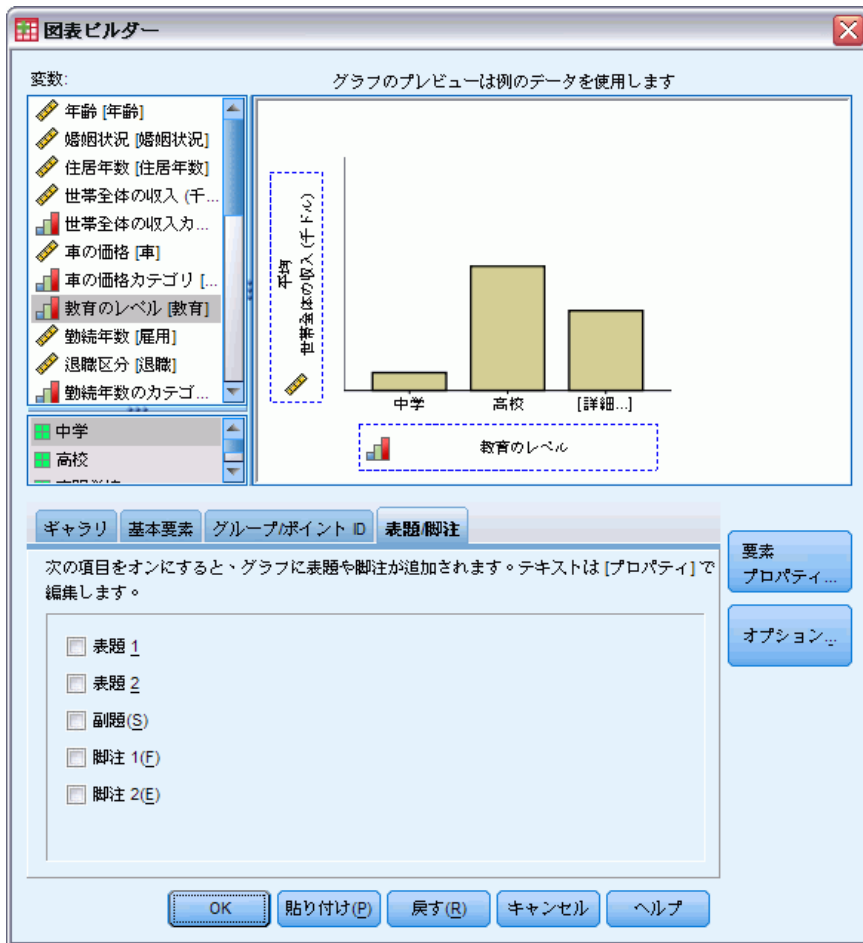
## [オプション] ダイアログ ボックスとテンプレート



[オプション] ダイアログ ボックスに、選択したテンプレートのファイルパスが表示されます。

- ▶ [OK] をクリックして [オプション] ダイアログ ボックスを閉じます。

図 6-18  
ドロップ領域がすべて設定された [図表ビルダー]

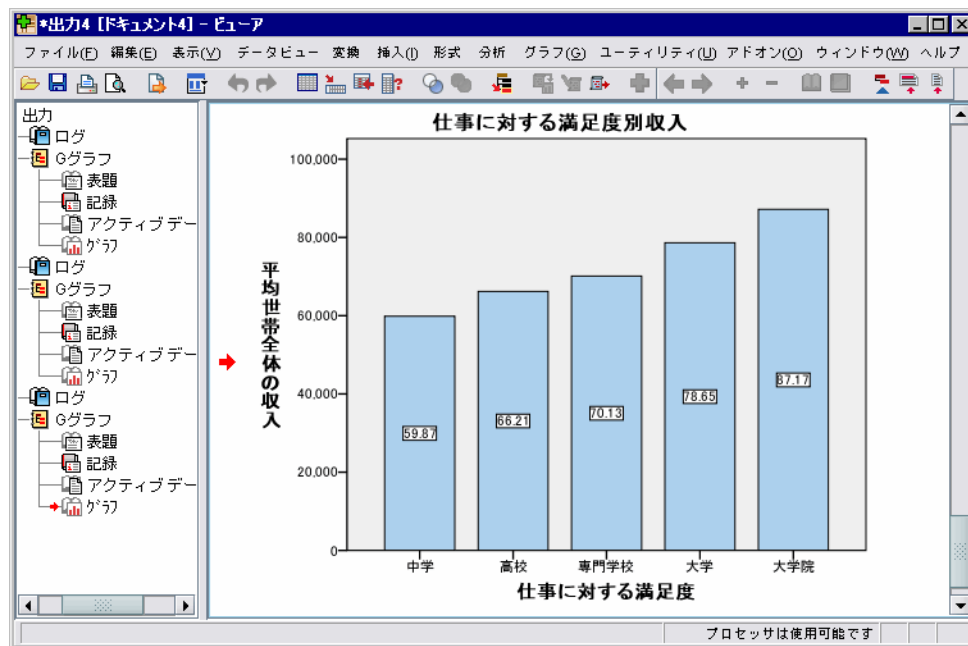


- ▶ [図表ビルダー] ダイアログ ボックスで [OK] をクリックして、グラフを作成し、テンプレートを適用します。

新しい図表の書式設定は、以前に作成および編集した図表の書式設定と一致します。x 軸上の変数は異なりますが、それ以外の点に関して、両方の図表は似ています。前のグラフの表題は、図表ビルダーから削除しても、テンプレートに保存されるということに注意してください。

図表の作成後にテンプレートを適用する場合は、図表エディタでその操作を実行できます（[ファイル]メニューの[図表テンプレートの適用]を選択します）。

図 6-19  
ビューアに表示される更新された棒グラフ



## 図表オプションの定義

テンプレートを使用して図表の書式を設定するだけでなく、オプションを使用して、図表の作成方法をさまざまな角度から管理できます。

- ▶ データ エディタまたはビューアのメニューから、次の項目を選択します。  
編集 > オプション(O)...

[オプション] ダイアログ ボックスには、多くの構成設定が含まれています。[グラフ] タブをクリックして、使用可能なオプションを確認します。

図 6-20  
[オプション] ダイアログ ボックスの [グラフ] タブ



オプションを使用して、図表の作成方法を管理します。新しい図表ごとに、次のことを指定できます。

- 現在の設定とテンプレートのどちらを使用するか。
- 幅対高さの比（縦横比）。
- テンプレートを使用しない場合に書式に使用する設定。
- グラフ要素のスタイル変更。

スタイルの変更により、新しい図表内のグラフ要素のスタイルを指定できます。この例では、色のスタイル変更の詳細について確認します。

- ▶ [色] をクリックして、[データ要素 カラー] ダイアログ ボックスを開きます。

単純なグラフの場合、指定した 1 つのスタイルが図表エディタで使用されます。グループ化されたグラフの場合、そのグラフ内のグループ（カテゴリ）ごとに変更した一連のスタイルが図表エディタで使用されます。

- ▶ [単純なグラフ] を選択します。
- ▶ 下から 2 行目の右から 3 番目にある、薄緑色を選択します。

図 6-21  
[データ要素 カラー] ダイアログ ボックス



- ▶ [続行] をクリックします。
- ▶ [オプション] ダイアログ ボックスで、[OK] をクリックして、色のスタイルの変更を保存します。

新たに作った単純なグラフのグラフ要素の色は薄緑色になります。

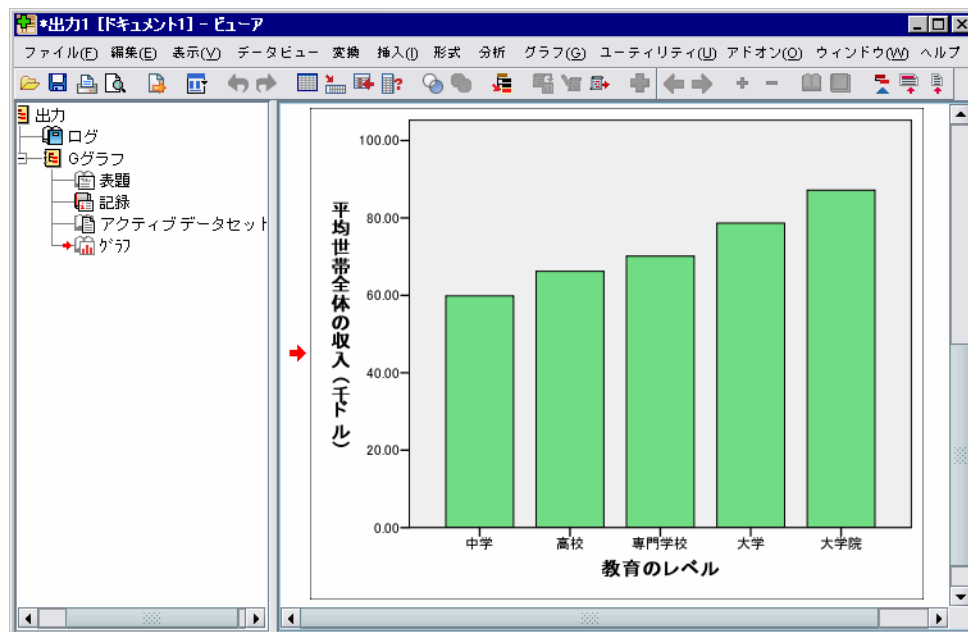
- ▶ データ エディタまたはビューアのメニューから、次の項目を選択します。  
グラフ(G) > 図表ビルダー(C)...

図表ビルダーには、最後に作成したグラフが表示されます。このグラフには、テンプレートが関連付けられていたということを思い出してください。そのテンプレートは、もう使われていません。

- ▶ [オプション] をクリックします。
- ▶ 先に追加したテンプレートの選択を解除（チェックマークをはずす）します。テンプレートは、赤い X をクリックして削除することもできます。削除する代わりにで選択解除することにより、そのテンプレートは別の機会に利用することができます。
- ▶ [OK] をクリックして、図表を作成します。

新しいグラフ内の棒は薄緑色です。この図表が、先ほどの図表と異なる点は他にもあります。表題がなく、軸ラベルは多数あり、データラベルはありません。違いが発生した理由は、テンプレートをこの図表に対して適用しなかったためです。

図 6-22  
ビューアに表示される更新された棒グラフ



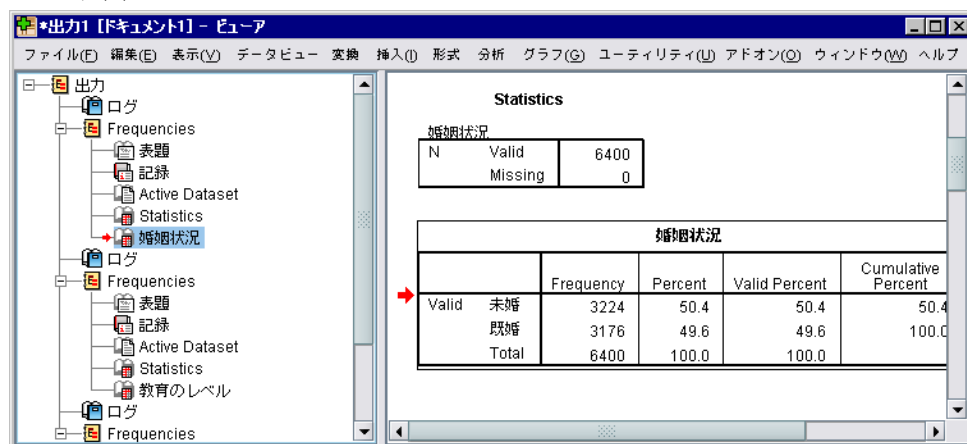


# 出力の作業

統計手続きの実行結果は、ビューアに表示されます。手続きを実行するときに選択した内容に応じて、統計テーブル、グラフ、またはテキストが出力として生成されます。この項では、ファイル viewertut.spv および demo.sav を使用します。詳細は、A 付録 p.165 サンプル ファイル を参照してください。

## ビューアを使用する

図 7-1  
ビューア(V)



ビューア ウィンドウは、2 つのウィンドウ枠に分かれています。アウトラインのウィンドウ枠には、ビューアに保存されている情報のアウトラインが表示されます。出力内容のウィンドウ枠には、統計テーブル、グラフ、およびテキストの出力が表示されます。

ウィンドウ内で縦方向または横方向に移動するには、スクロールバーを使います。より簡単に移動するには、アウトラインのウィンドウ枠内で項目をクリックし、それを出力内容のウィンドウ枠に表示させる方法があります。

ビューアにテーブル全体が入りきらないときや、アウトラインのウィンドウ枠が狭すぎる場合は、ウィンドウのサイズを簡単に変更できます。

- ▶ アウトラインのウィンドウ枠の右枠をクリックしてドラッグすれば、幅が変わります。

アウトラインのウィンドウ枠にある開いた本のアイコンは、その項目がビューアに表示されていることを示します。ただし、出力内容のウィンドウ枠の画面に見えている部分にあるとは限りません。

- ▶ テーブルや図表を隠すには、アウトラインのウィンドウ枠で本のアイコンをダブルクリックします。

データを隠すと、開いた本のアイコンが閉じた本のアイコンに変わります。

- ▶ 隠されている出力を再び表示するには、閉じた本のアイコンをダブルクリックします。

特定の統計手続きの出力をすべて隠したり、ビューア内のすべての出力を隠すことができます。

- ▶ 特定の手続きの結果を隠すには、手続きの左にあるマイナス符号 (-) 付きのボックスをクリックします。すべての出力を隠すには、アウトラインのウィンドウ枠で一番上にある項目のボックスをクリックします。

アウトラインが閉じ、結果が隠れたことがわかります。

出力の表示順を変更することもできます。

- ▶ アウトラインのウィンドウ枠で、移動したい項目をクリックします。
- ▶ 選択された項目をアウトライン内の新しい場所へドラッグし、マウスボタンを離します。

図 7-2  
ビューアでの出力を並べ替えた状態

The screenshot shows a software window titled '\*出力1 [ドキュメント1] - ビューア'. The left pane shows a tree view of outputs, with 'クロス集計表' selected. The main pane displays a table titled 'クロス集計表'.

**処理したケースの要約**

	ケース				合計	
	有効数		欠損			
	N	%-セット	N	%-セット	N	%-セット
携帯情報端末 * 性別 * インターネット	6145	96.0%	255	4.0%	6400	100.0%

**携帯情報端末と性別とインターネットのクロス表**

インターネット		性別	
		女性	男性
利用していない	携帯情報端末 持っていない	1897	1962
	期待度数	1903.4	1955.6
	携帯情報端末の %	49.2%	50.8%
持っている	度数	327	323
	期待度数	320.6	329.4
	携帯情報端末の %	50.3%	49.7%
合計		2224	2285

出力項目は、出力内容のウィンドウ枠でもドラッグ アンド ドロップで移動できます。

## ピボット テーブル エディタの使用法

統計手続きの結果のほとんどは、ピボット テーブルとして表示されます。

### 出力の定義へのアクセス

出力には、さまざまな統計用語が表示されます。これらの統計用語の定義を、ビューアに直接表示させることができます。

- ▶ [情報端末と性別とインターネットのクロス表] テーブルをダブルクリックします。
- ▶ [期待度数] を右クリックし、ポップアップ コンテキスト メニューから [用語のヘルプ] を選択します。

定義がポップアップ ウィンドウに表示されます。

図 7-3  
ポップアップに表示された定義

The screenshot shows a window titled '\*viewer.tut.spv [ドキュメント1] - ビューア'. The main area displays a pivot table with the following data:

			携帯情報端末と性別とインターネットのクロス表		
			性別		合計
			女性	男性	
インターネット	携帯情報端末	利用していない	1897	1962	3859
		持っている	1903.4	1955.6	3859.0
			50.8%	50.8%	100.0%
			323	329.4	650.0
			50.3%	49.7%	100.0%
合計			2224	2285	4509
			2224.0	2285.0	4509.0
			49.3%	50.7%	100.0%
利用している	携帯情報端末	持っていない	513	512	1025
		持っている	512.5	512.5	1025.0
			50.0%	50.0%	100.0%
			305	306	611
			305.5	305.5	611.0
			49.9%	50.1%	100.0%

The context menu is open over the '期待度数' cell, displaying the following text:

The number of cases that would be expected in the cell if the row and column variables are statistically independent or unrelated to one another.

## テーブルのピボット

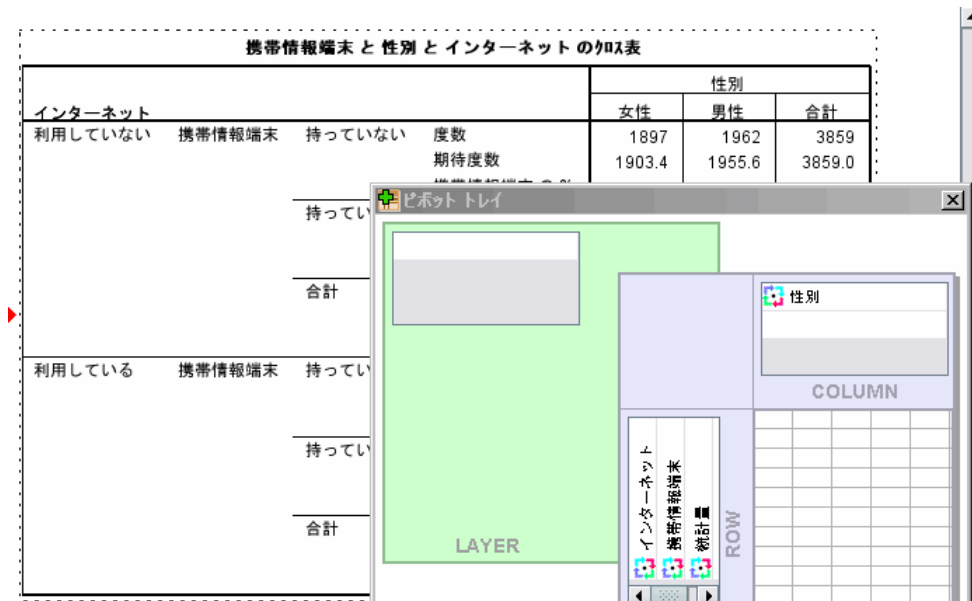
デフォルトで生成されるテーブルには、期待するほど整然と情報が表示されるとは限りません。ピボットテーブルでは、行と列の入れ替えやデータの順序変更などのさまざまな方法でテーブルを調整できます。たとえば、横長のテーブルの場合、行と列を入れ替えれば縦長のテーブルになります。テーブルのレイアウトが変わっても、内容は変わりません。これは、データを別のより好ましい形で表示するための技法です。

- ▶ アクティブになっていない場合は、[情報端末と性別とインターネットのクロス表] テーブルをダブルクリックしてアクティブにします。
- ▶ [ピボット トレイ] ウィンドウが表示されていない場合には、メニューから次の項目を選択します。

ピボット(P) > ピボットトレイ

[ピボット トレイ] を使って、データを列、行、層の間で移動することができます。

図 7-4  
ピボットトレイ



- ▶ 行次元の [統計量] 要素を [性別] の下にある列次元にドラッグします。変更を反映してテーブルが即座に再構成されます。

図 7-5  
行を列に移動する

			性別						合計		
			女性			男性					
インターネット			度数	期待度数	携帯情報端末の%	度数	期待度数	携帯情報端末の%	度数	期待度数	携帯情報端末の%
利用していない	携帯情報端末	持っていない	1897						0.0	100.0%	
		持っている	327					0.0	100.0%		
合計			2224					0.0	100.0%		
利用している	携帯情報端末	持っていない	513					0.0	100.0%		
		持っている	305					0.0	100.0%		
合計			818					0.0	100.0%		

ピボット トレイ内の要素の順序は、テーブル内の要素の順序を反映しています。

- ▶ 行次元の [携帯情報端末] 要素を [インターネット] 要素の前にドラッグアンド ドロップで移動して、2 つの行の順序を入れ替えます。

図 7-6  
行を入れ替える

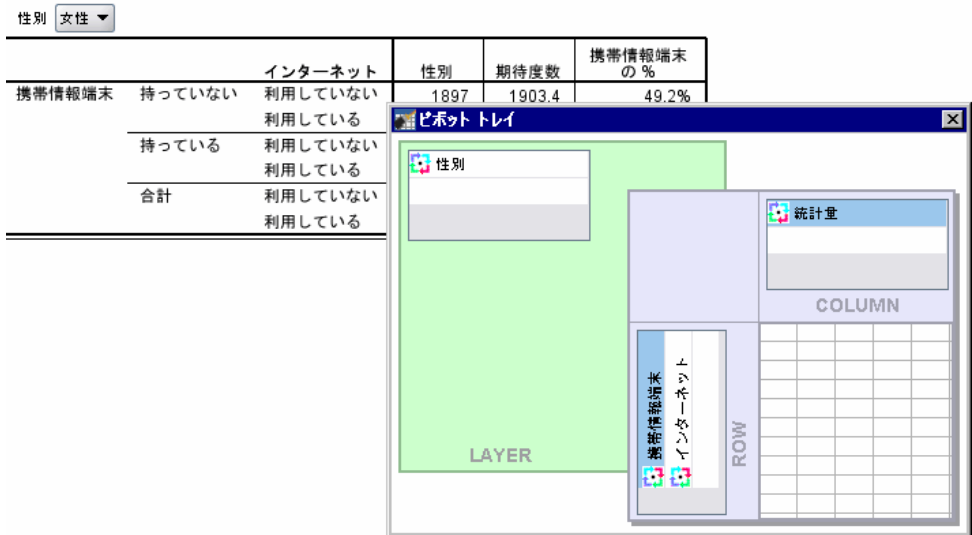
			性別						合計		
			女性			男性					
インターネット			度数	期待度数	携帯情報端末の%	度数	期待度数	携帯情報端末の%	度数	期待度数	携帯情報端末の%
携帯情報端末	持っていない	利用していない	1897					59	3859.0	100.0%	
		利用している	513					25	1025.0	100.0%	
	持っている	利用していない	327					50	650.0	100.0%	
		利用している	305					1	611.0	100.0%	
合計			2224					99	4509.0	100.0%	
利用している			818					86	1636.0	100.0%	

## 層の作成と表示

テーブルが大きく、情報のカテゴリが入れ子になっている場合は、層を使うと便利です。層を作成すると、テーブルの構造が単純になり、読みやすくなります。

- ▶ [性別] 要素を、列次元から層次元にドラッグします。

図 7-7  
層次元の [性別] ピボットアイコン



別の層を表示するには、テーブル内のドロップダウン リストからカテゴリを選択します。

図 7-8  
層を選択する

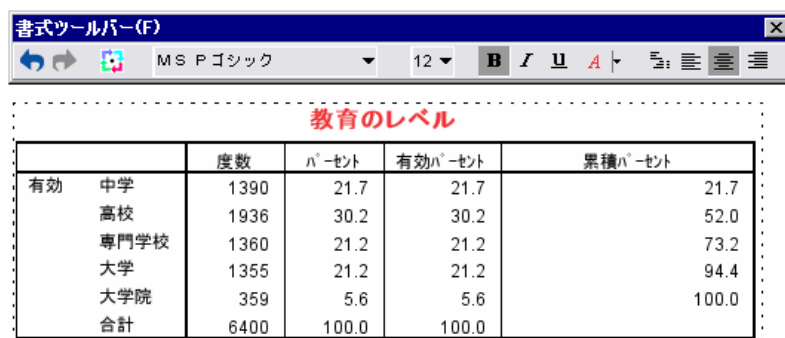
		インターネット	性別	期待度数	携帯情報端末の %
携帯情報端末	持っていない	利用していない	1897	1903.4	49.2%
	持っている	利用していない	513	512.5	50.0%
		利用している	327	320.6	50.3%
		利用している	305	305.5	49.9%
合計	持っていない	利用していない	2224	2224.0	49.3%
	持っている	利用している	818	818.0	50.0%

## テーブルの編集

ピボット テーブルは、ユーザー指定のテーブルルックを作成してある場合を除き、標準書式を使って作成されます。テーブル内のテキストはすべて、書式を変更できます。変更できる書式には、フォント、サイズ、フォントのスタイル（太字または斜体）、および色があります。

- ▶ [教育のレベル] テーブルをダブルクリックします。
- ▶ [書式] ツールバーが表示されていない場合には、メニューから次の項目を選択します。  
表示 > ツールバー(B)
- ▶ 表題のテキスト [教育のレベル] をクリックします。
- ▶ ツールバーのフォント サイズのドロップダウン リストから、[12] を選択します。
- ▶ 表題テキストの色を変更するには、[文字の色] ツールをクリックし、新しい色を選択します。

図 7-9  
書式が設定し直されたピボット テーブルの表題テキスト



		度数	パーセント	有効パーセント	累積パーセント
有効	中学	1390	21.7	21.7	21.7
	高校	1936	30.2	30.2	52.0
	専門学校	1360	21.2	21.2	73.2
	大学	1355	21.2	21.2	94.4
	大学院	359	5.6	5.6	100.0
	合計	6400	100.0	100.0	

テーブルやラベルの内容を編集することもできます。たとえば、このテーブルの表題を変更できます。

- ▶ 表題をダブルクリックします。
- ▶ 新しいラベルとして、「**学歴**」と入力します。

注：テーブル内の値を変更しても、合計やその他の統計量は再計算されません。

## 行および列を隠す

テーブルに表示されるデータのいくつかが役に立つものでなかったり、テーブルを不必要に複雑にする場合があります。そのような場合でも、データを削除する代わりに、行や列全体を隠すことができます。

- ▶ アクティブになっていない場合は、[学歴] テーブルをダブルクリックしてアクティブにします。
- ▶ [有効パーセント] 列のラベルをクリックして選択します。
- ▶ [編集] メニューまたはマウス右ボタン コンテキスト メニューから次の項目を選択します。

選択(C) > データとラベルのセル

- ▶ [表示] メニューから [隠す] を選択するか、マウス右ボタン コンテキスト メニューから [カテゴリを隠す] を選択します。

選択された列は、隠れていますが削除はされていません。

図 7-10

[有効パーセント] 列が隠されているテーブル

教育レベル		度数	パーセント	累積パーセント
有効	中学	1390	21.7	21.7
	高校	1936	30.2	52.0
	専門学校	1360	21.2	73.2
	大学	1355	21.2	94.4
	大学院	359	5.6	100.0
	合計	6400	100.0	

列を再び表示するには

- ▶ メニューから次の項目を選択します。

表示 > すべて表示

行も、列と同じ方法で隠したり表示したりできます。

## データ表示書式の変更

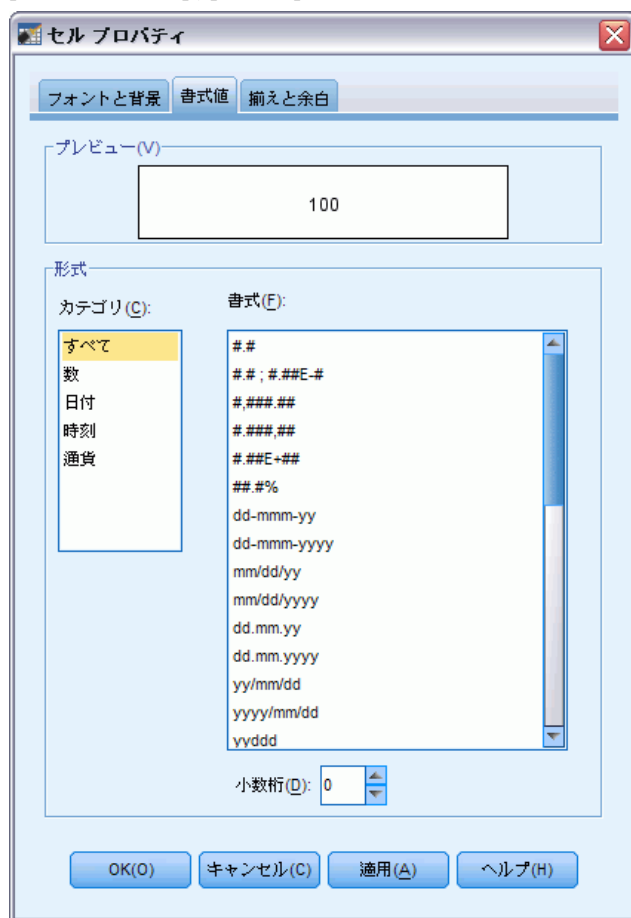
ピボット テーブルのデータの表示書式は簡単に変更できます。

- ▶ アクティブになっていない場合は、[学歴] テーブルをダブルクリックしてアクティブにします。
- ▶ [パーセント] 列のラベルをクリックして選択します。



- ▶ [編集] メニューまたはマウス右ボタン コンテキスト メニューから次の項目を選択します。  
選択(C) > データセル
- ▶ [書式] メニューまたはマウス右ボタン コンテキスト メニューから [セルプロパティ] を選択します。
- ▶ [書式値] タブをクリックします。
- ▶ この列の小数を隠すために、[小数桁数] フィールドに「0」を入力します。

図 7-11  
[セルプロパティ]、[書式値] タブ



データ型と書式もこのダイアログ ボックスで変更できます。

- ▶ [カテゴリ] リストから目的の型を選択し、[書式] リストでその型に適用する書式を選択します。

- ▶ 変更を適用するには、[OK] または [適用] をクリックします。

図 7-12  
[パーセント] 列の小数が隠されている状態

### 教育レベル

	度数	パーセント	累積パーセント
有効 中学	1390	22	21.7
高校	1936	30	52.0
専門学校	1360	21	73.2
大学	1355	21	94.4
大学院	359	6	100.0
合計	6400	100	

[パーセント] 列の小数は表示されなくなっています。

## テーブルルック

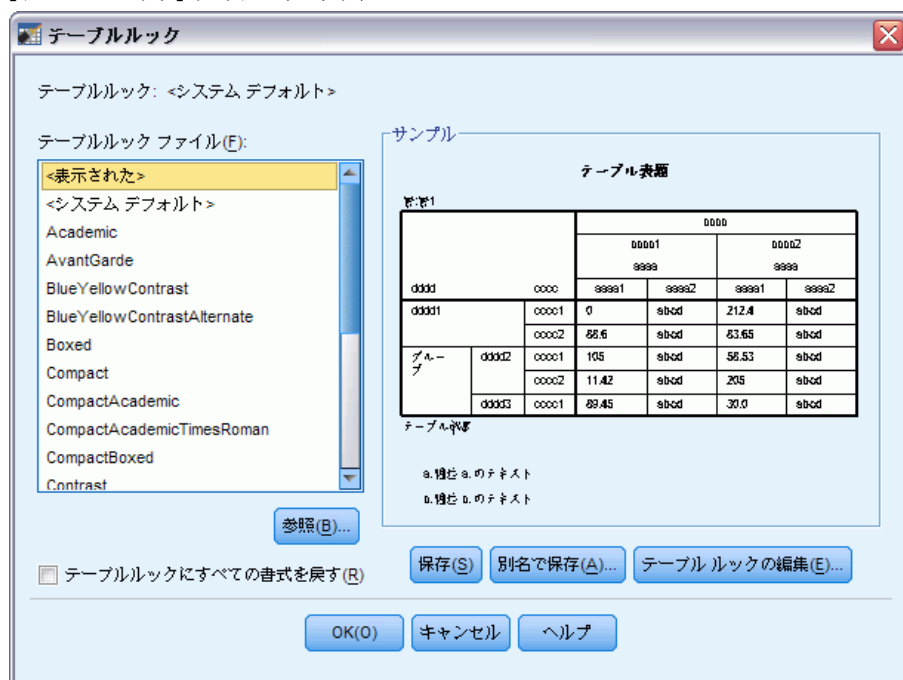
明確で簡潔、かつ意味のある結果を引き出せるかどうかは、テーブルの形式によって決まります。テーブルが読みにくいと、情報がなかなか把握できないことがあります。

### 定義済み書式の使用方法

- ▶ [婚姻状況] テーブルをダブルクリックします。
- ▶ メニューから次の項目を選択します。  
書式 > テーブルルック(L)...

[テーブルルック] ダイアログ ボックスに、各種の定義済みスタイルが表示されます。リストの中からスタイルを選択すると、右側の [サンプル] ウィンドウにプレビューが表示されます。

図 7-13  
[テーブルルック] ダイアログ ボックス



スタイルはそのまま使うことも、必要に応じて編集することもできます。

- ▶ 既存のスタイルをそのまま使う場合は、スタイルを選択し、[OK] をクリックします。

## テーブルルック スタイルのカスタマイズ

書式は、必要に応じてカスタマイズすることができます。背景の色から枠線のスタイルまで、テーブルのほぼすべての要素がカスタマイズ可能です。

- ▶ [婚姻状況] テーブルをダブルクリックします。
- ▶ メニューから次の項目を選択します。  
書式 > テーブルルック(L)...
- ▶ 目的の書式に最も近いスタイルを選択し、[テーブルルックの編集] をクリックします。

- ▶ [セル書式] タブをクリックすると、書式のオプションが表示されます。

図 7-14  
[テーブル プロパティ] ダイアログ ボックス



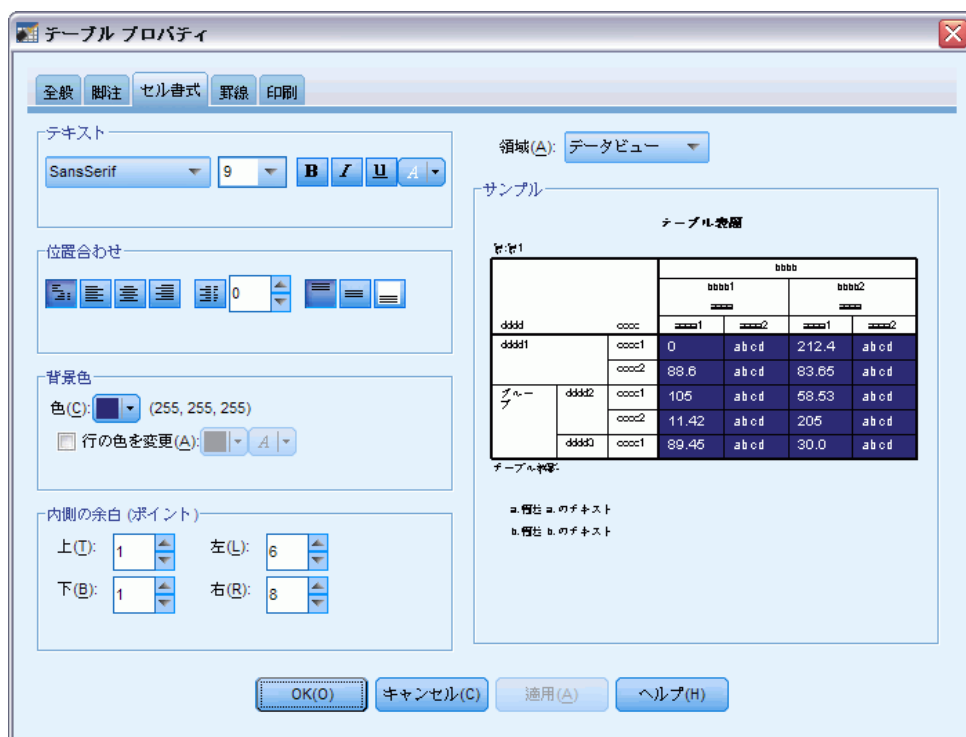
書式オプションには、フォント名、フォント サイズ、スタイル、および色があります。さらに、配置、テキストと背景の色、および余白のサイズも指定できます。

右側の [サンプル] ウィンドウに、書式を変更したときのテーブルのプレビューが表示されます。テーブルの各領域に、異なる書式スタイルを適用することができます。たとえば、表題とデータには違う書式を使うのが一般的です。編集するテーブル領域を選択するには、[領域] ドロップダウン リストで領域の名前を選択するか、または [サンプル] ウィンドウで目的の領域をクリックします。

- ▶ [領域] ドロップダウン リストから [データ] を選択します。
- ▶ [背景] ドロップダウン パレットから新しい色を選択します。
- ▶ 次に、新しいテキストの色を選択します。

[サンプル] ウィンドウに新しいスタイルが表示されます。

図 7-15  
テーブル セル書式の変更



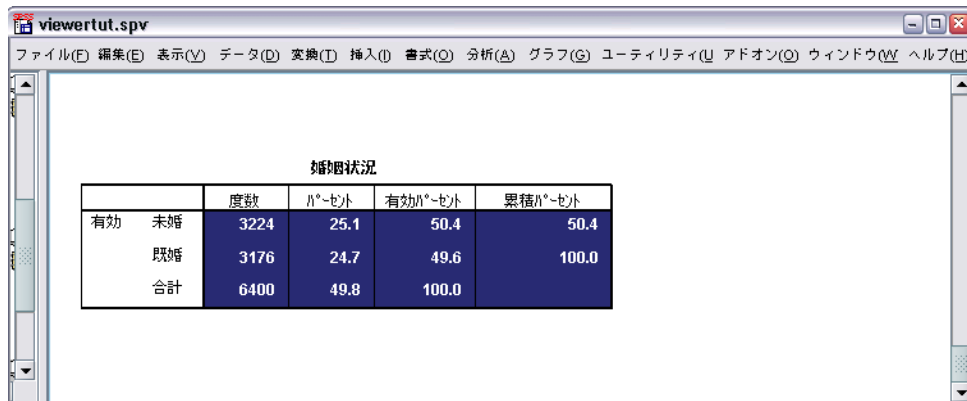
- ▶ [OK] をクリックすると、[テーブルルック] ダイアログ ボックスに戻ります。

新しく作成したスタイルを保存しておけば、後でテーブルに簡単に適用することができます。

- ▶ [名前を付けて保存] をクリックします。
- ▶ 目的の保存先ディレクトリまで移動し、[ファイル名] テキスト ボックスに新しいスタイルの名前を入力します。
- ▶ [保存] をクリックします。
- ▶ [OK] をクリックすると変更が適用され、ビューアに戻ります。

テーブルに、ユーザー指定の書式が含まれているのがわかります。

図 7-16  
ユーザー指定のテーブルルック



		度数	割合	有効割合	累積割合
有効	未婚	3224	25.1	50.4	50.4
	既婚	3176	24.7	49.6	100.0
	合計	6400	49.8	100.0	

## デフォルト テーブル書式の変更

テーブルの作成後でも書式を変更できますが、デフォルトのテーブルルックを変更しておけばテーブルを作成するごとに書式を変更する必要がなくなり、効率的に作業を行えます。

ピボット テーブルに対してデフォルトのテーブルルック スタイルを変更するには、メニューから次の項目を選択します。

編集 > オプション(O)...

- ▶ [オプション] ダイアログ ボックスの [ピボットテーブル] タブをクリックします。

図 7-17  
[オプション] ダイアログ ボックス



- ▶ 新しく作成するテーブルのすべてに適用するテーブルルック スタイルを選択します。

右側にある [サンプル] ウィンドウには、各テーブルルックのプレビューが表示されます。

- ▶ [OK] をクリックすると、設定が保存され、ダイアログ ボックスが閉じます。

デフォルトのテーブルルックを変更した後で作成したテーブルはすべて、自動的に新しい書式設定になっています。

## 初期の表示設定のカスタマイズ

初期の表示設定には、ビューア内でのオブジェクトの配置や、デフォルトでのオブジェクトの表示または非表示、ビューア ウィンドウの幅があります。これらの設定を変更するには、次の手順に従ってください。

- ▶ メニューから次の項目を選択します。  
編集 > オプション(O)...
- ▶ [ビューア] タブをクリックします。

図 7-18  
[ビューア] オプション



設定は、オブジェクト別に適用されます。たとえば、図表の表示設定をカスタマイズしても、テーブルの表示設定には影響しません。カスタマイズしたいオブジェクトだけを選択し、変更を加えます。

- ▶ [表題] アイコンをクリックすると、表題の設定が表示されます。
- ▶ [中央] をクリックすると、ビューアに表示される表題がすべて横方向の中央に配置されるようになります。



ログメッセージや警告メッセージのように、出力を乱雑にしがちな要素を隠すこともできます。アイコンをダブルクリックするだけで、そのオブジェクトの表示プロパティが切り替わります。

- ▶ [警告] アイコンをダブルクリックして、出力内の警告メッセージを隠します。
- ▶ [OK] をクリックすると変更が保存され、ダイアログボックスが閉じます。

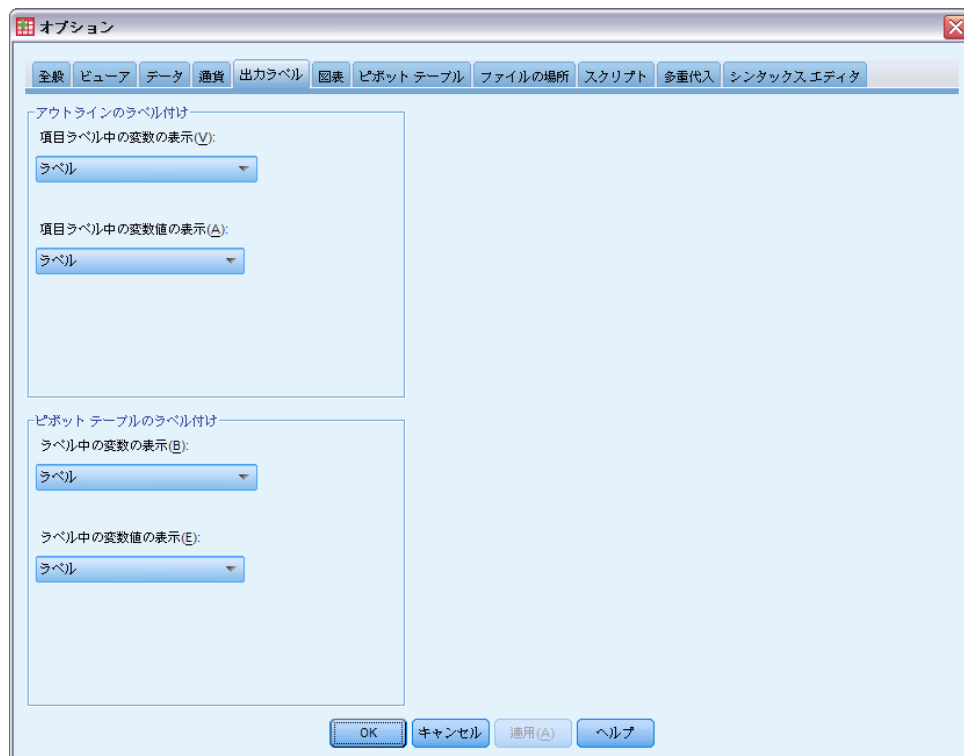
## 変数ラベルと値ラベルの表示

多くの場合、変数名や実際のデータ値より、変数および値のラベルを表示の方が効果的です。ときには、名前とラベルの両方を表示したい場合もあります。

- ▶ メニューから次の項目を選択します。  
編集 > オプション(O)...
- ▶ [出カラベル] タブをクリックします。

図 7-19

[出カラベル] のオプション



アウトラインのウィンドウと出力内容のウィンドウ枠に対して別の設定を指定できます。たとえば、アウトラインのウィンドウ枠にはラベル、出力内容のウィンドウ枠には変数名とデータ値を表示すると仮定します。

- ▶ ラベルの代わりに変数名を表示するために、[ピボット テーブルのラベル付け] グループで、[ラベル中の変数の表示] ドロップダウン リストから [名前] を選択します。
- ▶ 次に、ラベルの代わりにデータ値を表示するために、[ラベル中の変数値の表示] ドロップダウン リストから [値] を選択します。

図 7-20  
[ピボット テーブルのラベル付け] の設定



セッション中に作成された後続のテーブルには、これらの変更が反映されます。

図 7-21  
変数名と値の表示

婚姻状況					
		度数	パーセント	有効パーセント	累積パーセント
有効	0	3224	25.1	50.4	50.4
	1	3176	24.7	49.6	100.0
	合計	6400	49.8	100.0	

## SPSS 出力を他のアプリケーションで使用する

SPSS 出力は、さまざまなアプリケーションで使用できます。たとえば、プレゼンテーションやレポートにテーブルや図表を組み込むことができます。

次の例では Microsoft Word を使いますが、他のワープロソフトでも作業内容はほぼ同じです。

### 結果を Word のテーブルとして貼り付ける

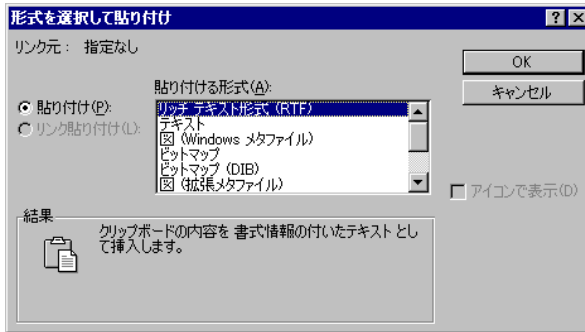
ピボットテーブルは、Word の表として Word 内に貼り付けることができます。フォントサイズや色などのテーブル属性は、すべて保持されます。テーブルは、Word のテーブル形式で貼り付けられるので、Word の中の他のテーブルと同じように編集することができます。

- ▶ ビューアで [婚姻状況] テーブルをクリックします。
- ▶ メニューから次の項目を選択します。  
編集 > コピー
- ▶ 使用するワープロソフトを開きます。
- ▶ ワープロソフトのメニューから、次の項目を選択します。  
編集 > 形式を選択して貼り付け...

- ▶ [形式を選択して貼り付け] ダイアログ ボックスで [リッチ テキスト形式 (RTF)] を選択します。

図 7-22

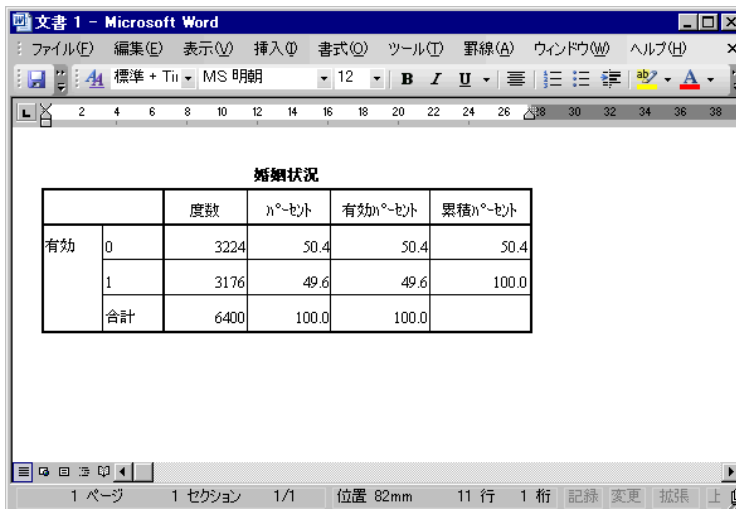
[形式を選択して貼り付け] ダイアログ ボックス



- ▶ [OK] をクリックして、結果を現在の文書に貼り付けます。

図 7-23

Word に表示されたピボット テーブル



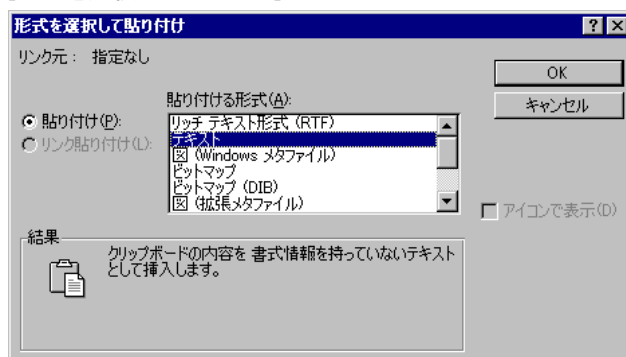
テーブルが文書内に表示されています。必要に応じてユーザー指定の書式を適用したり、データを編集したり、テーブルのサイズを変更できます。

## 結果を Word のテーブルとして貼り付ける

ピボット テーブルは、テキスト形式で他のアプリケーションにコピーできます。この形式で他のアプリケーションへコピーした場合、書式スタイルは失われますが、コピーしたテーブル データを編集できるという利点があります。

- ▶ ビューアで [婚姻状況] テーブルをクリックします。
- ▶ メニューから次の項目を選択します。  
編集 > コピー
- ▶ 使用するワープロ ソフトを開きます。
- ▶ ワープロ ソフトのメニューから、次の項目を選択します。  
編集 > 形式を選択して貼り付け...
- ▶ [形式を選択して貼り付け] ダイアログ ボックスで [テキスト]を選択します。

図 7-24  
[形式を選択して貼り付け] ダイアログ ボックス



- ▶ [OK] をクリックして、結果を現在の文書に貼り付けます。

図 7-25  
Word に表示されたピボット テーブル

		婚姻状況				
		度数	パーセント	有効パーセント	累積パーセント	
有効	0	3224	50.4	50.4	50.4	
	1	3176	49.6	49.6	100.0	
	合計	6400	100.0	100.0		

テーブルの列は、タブで区切られています。ワープロ ソフトのタブ位置を調整することで、列の幅を変更できます。

## Microsoft Word、PowerPoint および Excel ファイルへの結果のエクスポート

結果を Microsoft Word、PowerPoint、または Excel ファイルにエクスポートできます。ビューア内の選択した項目またはすべての項目をエクスポートできます。この項では、ファイル msouttut.spv および demo.sav を使用します。詳細は、A 付録 p.165 サンプル ファイル を参照してください。

注 :PowerPoint へのエクスポートは、Windows オペレーティング システム上のみで利用でき、Student 版では利用できません。

[ビューアの概要] ウィンドウ枠で、エクスポートしたい特定の項目を選択したり、すべての項目またはすべての表示項目をエクスポートしたりすることができます。

- ▶ ビューアのメニューから、次の項目を選択します。

ファイル(F) > エクスポート(E)...

ビューア内のすべてのオブジェクトをエクスポートする代わりに、表示されているオブジェクト (アウトラインのウィンドウ枠の開いた本のアイコン) だけをエクスポートしたり、アウトラインのウィンドウ枠で選択したオブジェクトだけをエクスポートできます。アウトラインのウィンドウ



生成されたファイルを Word で開くと、結果がどのようにエクスポートされたかがわかります。ここではすべてのオブジェクトのエクスポートを選択しているため、表示オブジェクトではない記録オブジェクトが Word に表示されます。

図 7-27  
Word で開いた output.doc

記録	
出力の作成日付	18-10-2007 21:51:13
コメント	
入力	データ
	C:\Program Files\SPSS\Tutorial\sample_files\pubs\spss\datajobs\demo.sav
	アクティブ データセット
	データセット1
	フィルタ
	<なし>
	重み付け
	<なし>
	分割ファイル
	<なし>
	作業データファイル内の行数
	6400



ピボット テーブルは、元のピボット テーブルのフォント、色、罫線等の書式が保たれた Word の表になります。

図 7-28  
Word で開いたピボット テーブル

The screenshot shows a Microsoft Word window titled 'output.doc - Microsoft Word'. The window contains two pivot tables. The first table is titled '性別' (Gender) and the second is '世帯全体の収入カテゴリ (千ドル)' (Household Income Category in thousands of dollars). Both tables show counts, percentages, and cumulative percentages.

		度数	パーセント	有効パーセント	累積パーセント
有効	女性	3179	49.7	49.7	49.7
	男性	3221	50.3	50.3	100.0
	合計	6400	100.0	100.0	

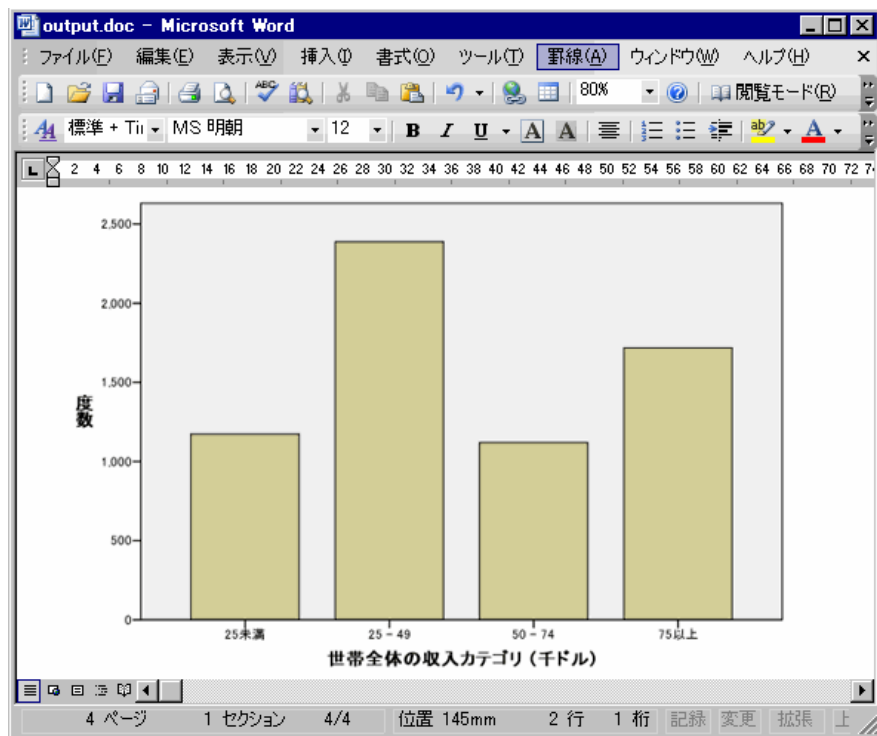
  

		度数	パーセント	有効パーセント	累積パーセント
有効	25未満	1174	18.3	18.3	18.3
	25 - 49	2388	37.3	37.3	55.7
	50 - 74	1120	17.5	17.5	73.2

図表は、グラフィック イメージとして Word 文書に挿入されます。

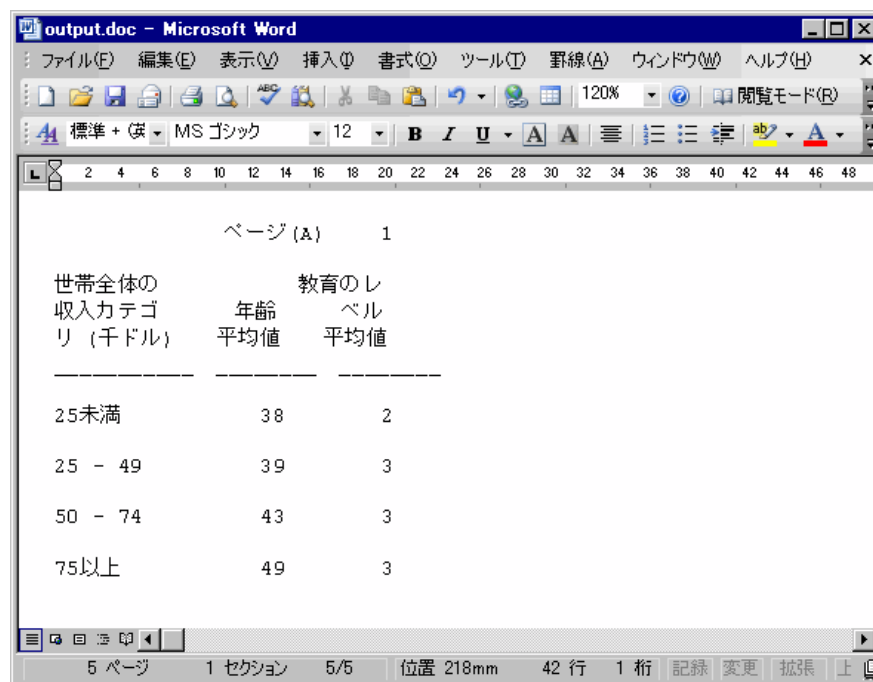
図 7-29

Word で開いた図表



テキスト出力は、ビューア内のテキスト オブジェクトに対して使用されるフォントと同じフォントで表示されます。正しい配置を行うためには、テキスト出力で固定ピッチ（等幅）フォントを使用する必要があります。

図 7-30  
Word で開いたテキスト出力



The screenshot shows a Microsoft Word window titled "output.doc - Microsoft Word". The window displays a table with the following content:

世帯全体の 収入カテ ゴリ (千ドル)	年齢 平均値	教育のレ ベル 平均値
25未満	38	2
25 - 49	39	3
50 - 74	43	3
75以上	49	3

The status bar at the bottom of the window indicates: 5 ページ, 1 セクション, 5/5, 位置 218mm, 42 行, 1 桁, 記録, 変更, 拡張, 上.

PowerPoint ファイルへのエクスポートでは、エクスポートされた各項目が個別のスライドに配置されます。PowerPoint にエクスポートされたピボットテーブルは、元のピボットテーブルのフォント、色、罫線等の書式が保たれた Word の表になります。

図 7-31  
PowerPoint に表示されたピボットテーブル

The screenshot shows a PowerPoint slide with the following pivot table:

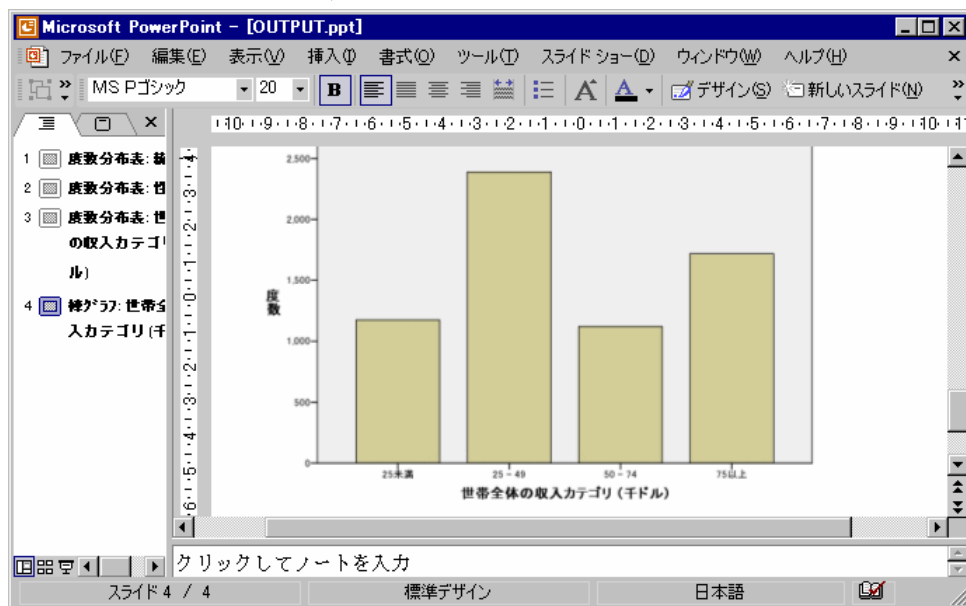
		性別			
		度数	割合	有効割合	累積割合
有効	女性	3179	49.7	49.7	49.7
	男性	3221	50.3	50.3	100.0
合計		6400	100.0	100.0	

The slide title is '度数分布表: 性別'. The status bar at the bottom indicates 'スライド 2 / 4', '標準デザイン', and '日本語'.

PowerPoint にエクスポートするために選択された図表は、PowerPoint ファイルに埋め込まれます。

図 7-32

PowerPoint に埋め込まれた図表



Excel ファイルへのエクスポートでは、結果は Word へのエクスポートとは異なる方法でエクスポートされます。

図 7-33

Excel で開いた output.xls

性別				
		度数	パーセント	有効パーセント
有効	女性	3178	49.7	49.7
	男性	3221	50.3	50.3
	合計	6400	100.0	100.0
世帯全体の収入カテゴリ (千ドル)				
		度数	パーセント	有効パーセント
有効	25未満	1174	18.3	18.3
	25 - 49	2388	37.3	37.3
	50 - 74	1120	17.5	17.5
	75以上	1718	26.8	26.8
	合計	6400	100.0	100.0

ピボット テーブルの行、列、およびセルは、Excel の行、列、およびセルとなります。

テキスト出力の各行は、Excel ファイルの行になります。この場合、1 行の内容はすべて 1 つのセルに入れられます。

図 7-34  
Excel で開いたテキスト出力

	A	B	C	D	E	F
72	世帯全体の	教育のレ				
73	収入カテゴ	年齢	ベル			
74	リ(千ドル)	平均値	平均値			
75	_____	_____				
76						
77	25未満	38	2			
78						
79	25 - 49	39	3			
80						
81	50 - 74	43	3			
82						
83	75以上	49	3			
84						
85						
86						

## 結果の PDF へのエクスポート

ビューア内のすべてのオブジェクト、または選択したオブジェクトを PDF (portable document format) ファイルにエクスポートできます。

- ▶ PDF にエクスポートする結果が表示されている [ビューア] ウィンドウで、次のメニューを選択します。  
ファイル(F) > エクスポート(E)...

- ▶ [出力のエクスポート] ダイアログ ボックスで、[エクスポート書式のファイルの種類] ドロップダウン リストから [Portable Document Format] を選択します。

図 7-35

## [出力のエクスポート] ダイアログ ボックス

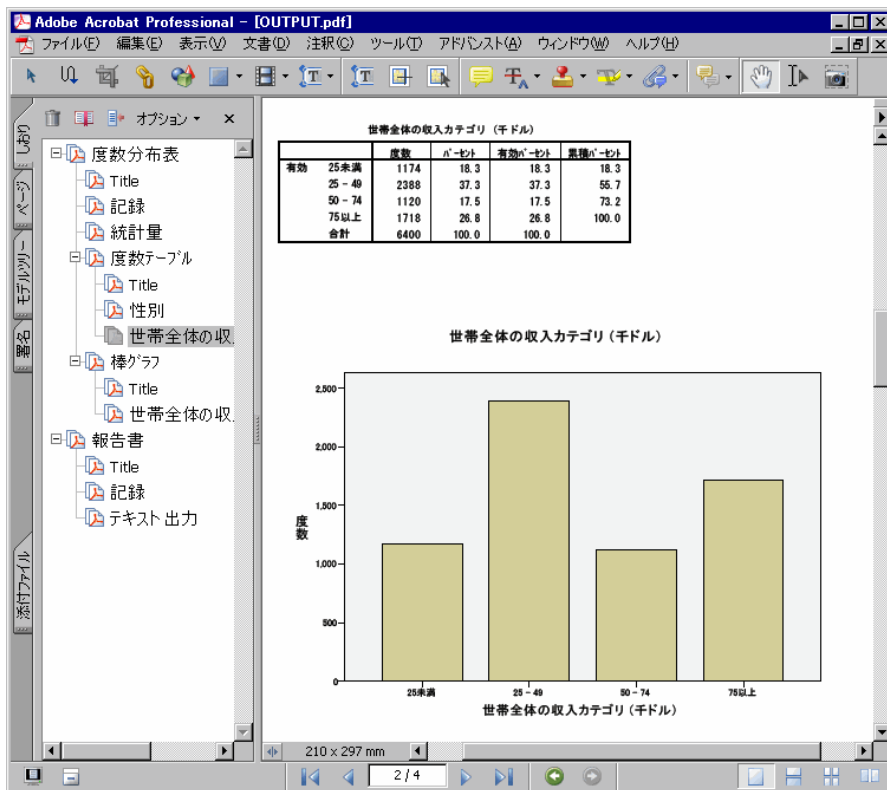


- 文書内の移動が簡単に行えるように、ビューア文書のアウトラインウィンドウ枠が PDF ファイルのブックマークに変換されます。
- PDF 文書のページ サイズ、方向、余白、ページのヘッダーやフッターの内容と表示、および図表の印刷サイズは、各種ページ設定オプション ([ビューア] ウィンドウの [ファイル] メニューの [ページ設定]) で制御します。
- PDF 文書の解像度 (DPI) は、デフォルトの解像度設定または現在選択されているプリンタに適用される現在の解像度設定となります ([ページ設定] で変更できます)。最大解像度は 1200 DPI です。プリンタ設



定の方が解像度が高い場合、PDF 文書の解像度は 1200 DPI になります。注：高解像度の文書を低解像度プリンタで印刷すると、最適な印刷結果が得られないことがあります。

図 7-36  
ブックマーク付きの PDF ファイル



## 結果の HTML へのエクスポート

結果は、HTML（ハイパーテキスト マークアップ言語）形式にエクスポートすることもできます。HTML として保存すると、非グラフィック出力は、すべて、単一の HTML ファイルにエクスポートされます。

図 7-37  
Web ブラウザで開いた output.htm

The screenshot shows a web browser window with the following content:

**性別**

	度数	パーセント	有効パーセント	累積パーセント
女性	3179	49.7	49.7	49.7
有効 男性	3221	50.3	50.3	100.0
合計	6400	100.0	100.0	

**世帯全体の収入カテゴリ (千ドル)**

	度数	パーセント	有効パーセント	累積パーセント
25未満	1174	18.3	18.3	18.3
有効 25 - 49	2388	37.3	37.3	55.7
有効 50 - 74	1120	17.5	17.5	73.2
有効 75以上	1718	26.8	26.8	100.0
合計	6400	100.0	100.0	

HTML へのエクスポートの際には、図表をエクスポートすることも可能ですが、この場合、単一のファイルにエクスポートされるわけではありません。

各図表はユーザーが指定した形式のファイルとして保存され、このようなグラフィック ファイルへの参照を、作成された HTML 文書に含めることができます。また、すべての図表、または選択した図表を個別のグラフィック ファイルにエクスポートするオプションも用意されています。

# シンタックスの使用

強力なコマンド言語を使用して、多くの一般的なタスクを保存して自動化できます。コマンド言語を使用すると、メニューやダイアログ ボックスにはない機能も実行できます。コマンドのほとんどは、メニューやダイアログ ボックスから利用できます。ただし、コマンド言語を使用しなければ利用できないコマンドやオプションもあります。コマンド言語を使用すると、シンタックス ファイルにジョブを保存することもできるので、日を改めて分析を繰り返すことができます。

コマンド シンタックス ファイルは、IBM® SPSS® Statistics シンタックス コマンドを含む単なるテキスト ファイルです。ユーザーは、シンタックス ウィンドウを開いて直接コマンドを入力することもできますが、作業の一部もしくは全部をダイアログ ボックスに行わせた方が簡単なことが多いです。

この章の例では、データ ファイル demo.sav を使用します。詳細は、[A 付録 p.165 サンプル ファイル](#) を参照してください。

注: コマンド シンタックスは、Student 版では利用できません。

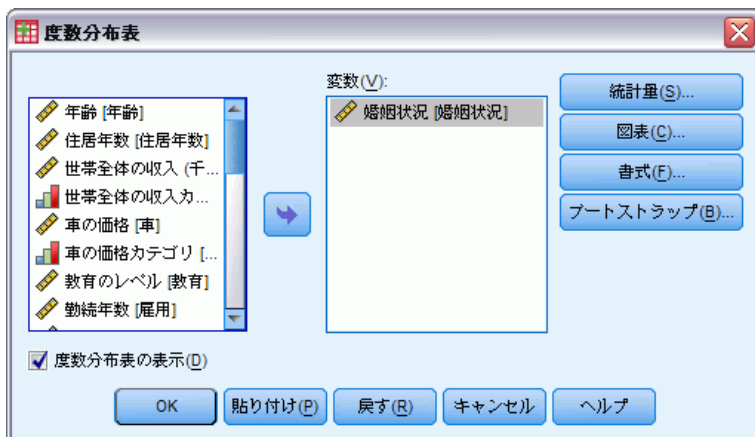
## シンタックスを貼り付ける

シンタックスを作成する最も簡単な方法は、ダイアログ ボックスにある [貼り付け] ボタンを使用する方法です。

- ▶ データ ファイル demo.sav を開きます。詳細は、[A 付録 p.165 サンプル ファイル](#) を参照してください。
- ▶ メニューから次の項目を選択します。  
分析(A) > 記述統計 > 度数分布表...

[度数分布表] ダイアログ ボックスが開きます。

図 8-1  
[度数分布表] ダイアログ ボックス



- ▶ [婚姻状況 [婚姻状況]] を選択し、[変数] リストに移動します。
- ▶ [図表] をクリックします。
- ▶ [図表の設定] ダイアログ ボックスで、[棒グラフ] を選択します。
- ▶ [図表の値] グループで、[パーセント] を選択します。
- ▶ [続行] をクリックします。
- ▶ [貼り付け] をクリックすると、ダイアログ ボックスでの指定によって作成されたシンタックスが、シンタックス エディタへコピーされます。

図 8-2  
度数分布表シンタックス



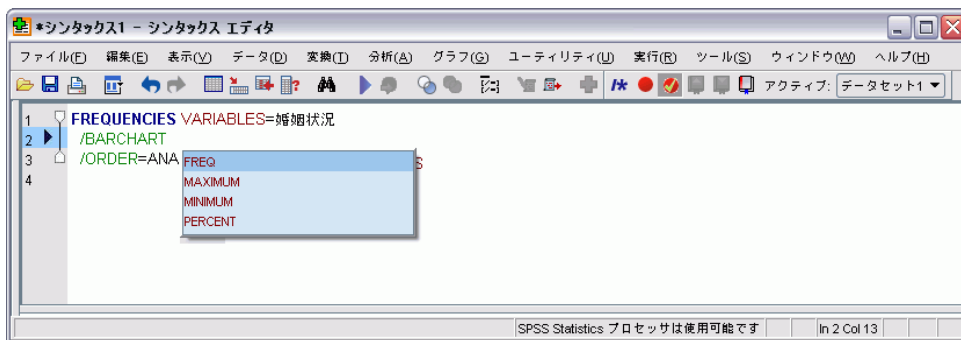
- ▶ 現在表示されているシンタックスを実行するには、メニューから次の項目を選択します。

実行(R) > 選択

## シンタックスを編集する

シンタックス ウィンドウでシンタックスを編集することができます。たとえば、サブコマンド `/BARCHART` を変更して、パーセントではなく度数の値を表示することができます。(サブコマンドはスラッシュで示されます)。度数を表示するキーワードがわかっている場合は、それを直接入力できます。キーワードがわからない場合、サブコマンド名の前にカーソルを置き、`Ctrl + スペースキー`を押して、サブコマンドに利用できるキーワードのリストを取得できます。サブコマンドの自動完了コントロールを表示します。

図 8-3  
キーワードを表示する自動完了コントロール



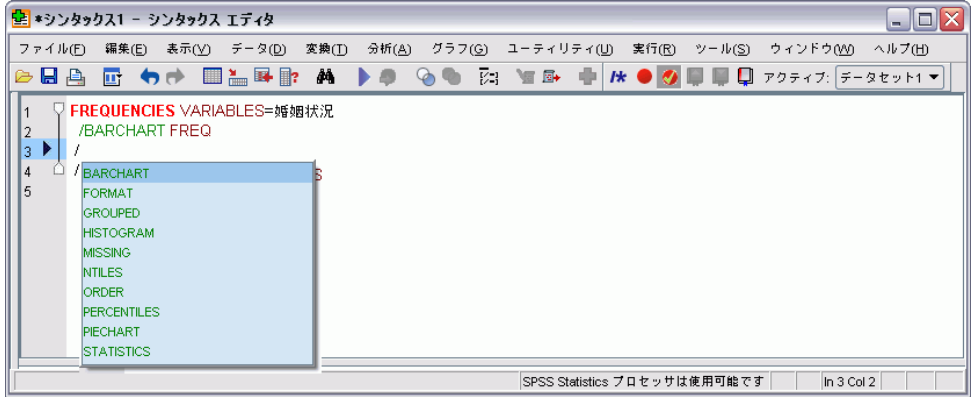
- ▶ 度数の `[FREQ]` というラベルの付いた項目をクリックします。自動完了コントロールの項目をクリックすると、現在のカーソルの場所にその項目が挿入されます (元の `PERCENT` キーワードは手動で削除されています)。

デフォルトでは、自動完了コントロールで、使用できる条件のリストを要求されます。たとえば、棒グラフと同時に円グラフも使用したいとします。円グラフは、別のサブコマンドで指定します。

- ▶ `FREQ` キーワードの後にスラッシュを入力して `Enter` キーを押し、サブコマンドの開始を指示します。

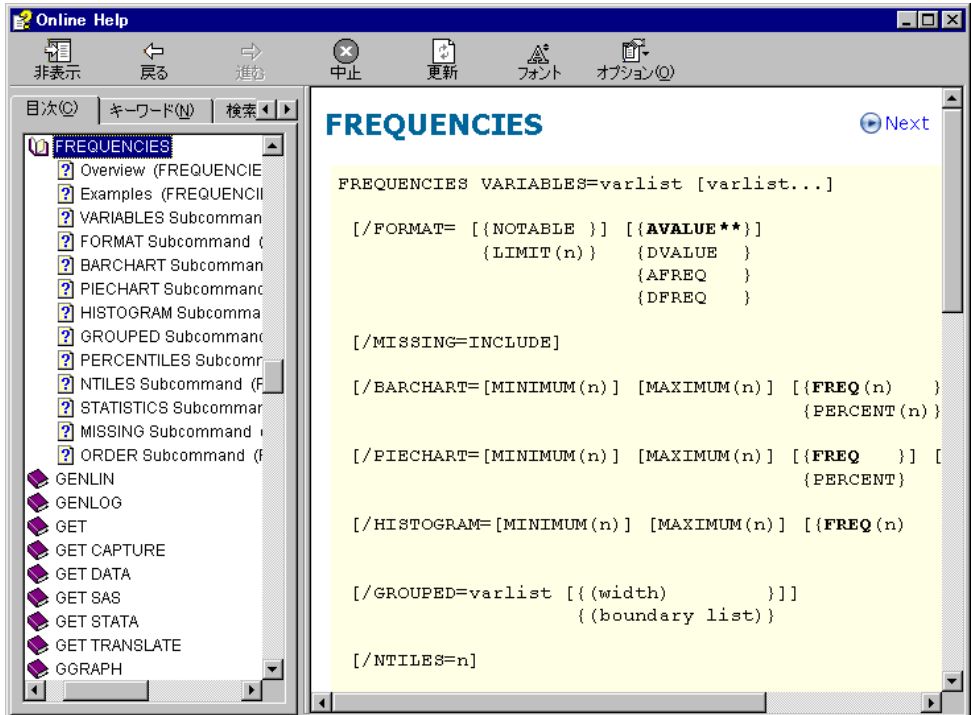
シンタックス エディタは、現在のコマンドに対するサブコマンドのリストを要求します。

図 8-4  
サブコマンドを表示する自動完了コントロール



現在のコマンドの詳細なヘルプを表示するには、F1 キーを押します。これにより、現在のコマンドのコマンド シンタックスのリファレンス情報に直接アクセスすることができます。

図 8-5  
FREQUENCIES コマンドのシンタックス ヘルプ



シンタックス ウィンドウで表示されるテキストは、色付きで表示されます。カラー コード化によって、認識された言葉のみに色がつくため、認識されていない言葉を迅速に特定できます。たとえば、FORMAT サブコマンドの綴りを誤って FRMAT と記述したとします。サブコマンドは、デフォルトでは緑色ですが、テキスト FRMAT は認識されないため、色なしで表示されます。

## シンタックス ファイルを開いて実行する

- ▶ 保存済みのシンタックス ファイルを開くには、メニューから次の項目を選択します。

ファイル(F) > 開く(O) > シンタックス(S)...

ファイルを開くための標準のダイアログ ボックスが表示されます。

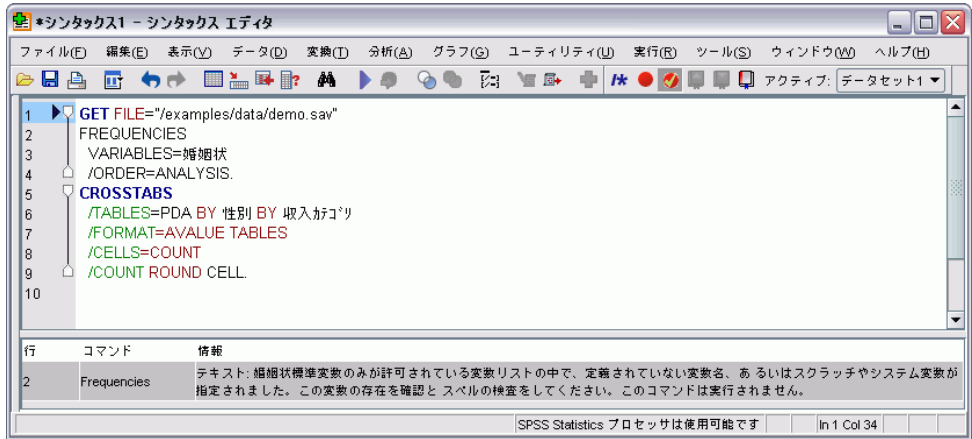
- ▶ シンタックス ファイルを選択します。シンタックス ファイルが表示されない場合は、表示するファイルの種類として、[シンタックス (\*.sps)] が選択されていることを確認します。
- ▶ [開く] をクリックします。
- ▶ コマンドを実行するには、シンタックス ウィンドウの [実行] メニューを使います。

コマンドを特定のデータ ファイルに適用する場合は、コマンドを実行する前にそのデータ ファイルを開いておくか、そのデータ ファイルを開くコマンドを追加する必要があります。この種のコマンドは、データ ファイルを開くダイアログ ボックスから貼り付けることができます。

## エラー枠について

エラー枠では、現在の実行のランタイム エラーを表示します。エラー枠には、エラーが発生した行番号のほか、各エラーの詳細が含まれています。

図 8-6  
シンタックス エディタに表示されるエラー枠



エラーが発生したコマンドの最初の行にカーソルを置いてエラーの場所のエントリをクリック。

## 分割点の使用

分割点を使用して、コマンド シンタックスの実行をシンタックスの特定のポイントで停止させ、準備ができた際に実行を続行できます。これにより、シンタックス ジョブの中間点で出力またはデータを表示、または、データの現在の状態に関する情報を表示する FREQUENCIES などのコマンド シンタックスを実行できます。分割点は、コマンド内の特定行でなく、コマンドのレベルに設定できます。

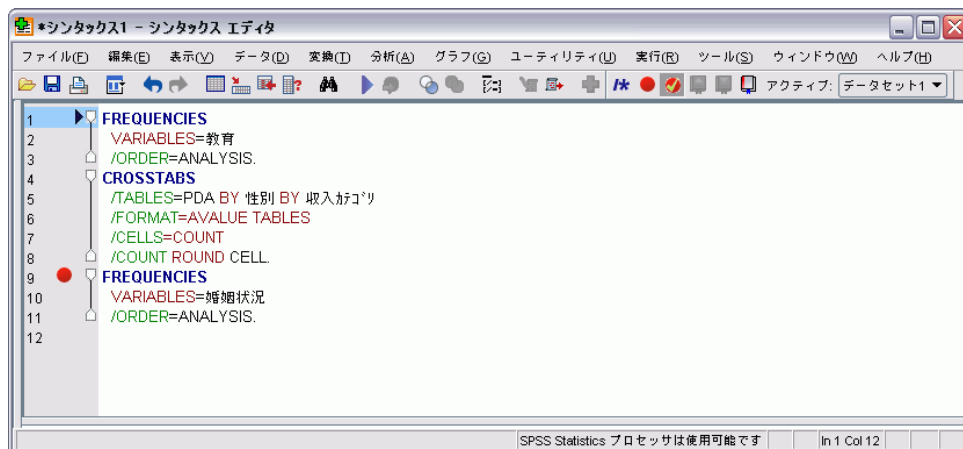
コマンドに分割点を挿入するには

- ▶ コマンドに関連するテキストの左側の任意の場所をクリックします。

分割点は、コマンド テキストの左側にあるガッターで赤い円で表示され、クリックした場所に関係なくコマンド名と同じ行内に表示されます。

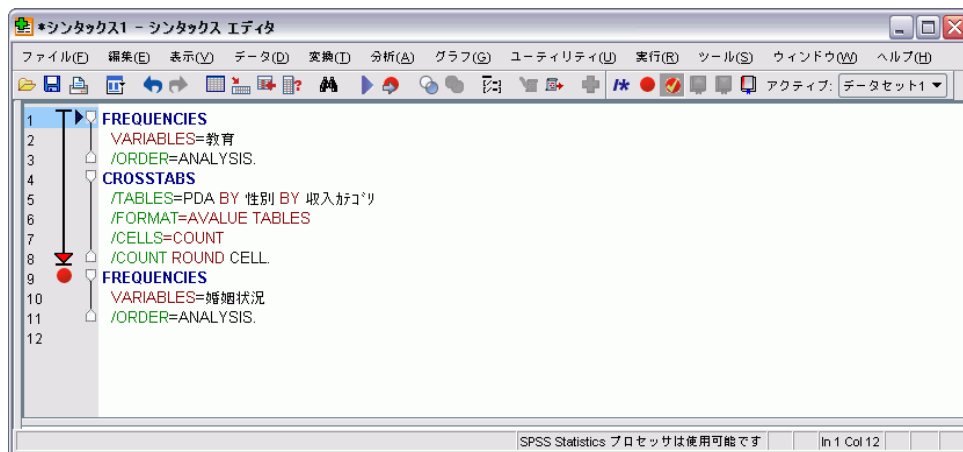


図 8-7  
[シンタックス エディタ] ウィンドウに表示される分割点



分割点を含むコマンド シンタックスを実行する場合、分割点を含む各コマンドの前で実行は停止します。

図 8-8  
分割点で停止した実行



コマンド テキストの左側にある下向き矢印は、シンタックス実行の進捗状況を示します。範囲は、最初のコマンド実行から最後のコマンド実行にまでわたり、分割点を含むコマンド シンタックスの実行時に特に役立ちます。

分割点の後で実行を再開するには

- ▶ シンタックス エディタのメニューから、次の項目を選択します。  
実行(R) > 続行



# データ値の変更

開始時に、データが、必要な分析やレポートに適した最も有用な方法で構成されているとは限りません。たとえば、次の操作を実行する場合があります。

- スケール変数からカテゴリ変数を作成する。
- いくつかの応答カテゴリを 1 つのカテゴリ内に結合する。
- 既存の 2 つの変数の差分を計算した新しい変数を作成する。
- 2 つの日付間の時間の長さを計算する。

この章では、データ ファイル demo.sav を使用します。詳細は、A 付録 p. 165 サンプル ファイル を参照してください。

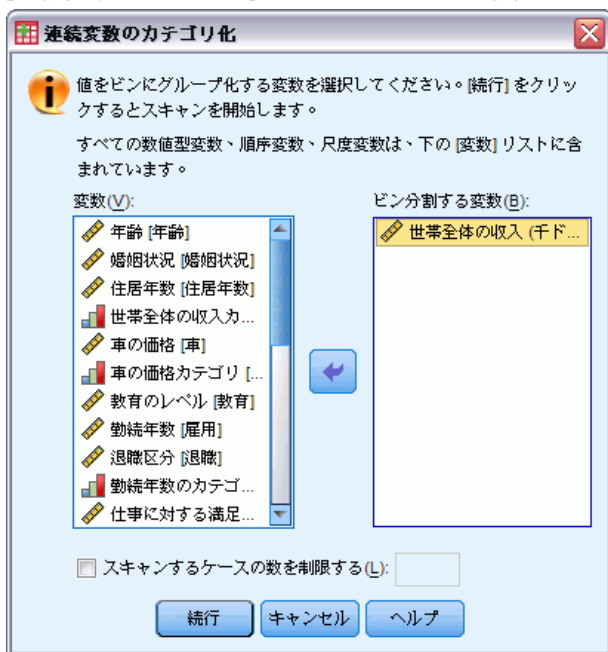
## スケール変数からカテゴリ変数を作成する

実際には、データ ファイル demo.sav にあるいくつかのカテゴリ変数は、そのデータ ファイル内のスケール変数から導かれます。たとえば、変数「収入」は、4 つのカテゴリにグループ化された [年収] です。このカテゴリ変数は、1-4 の整数値を使用し、\$25 未満、\$25-\$49、\$50-\$74、\$75 以上の収入カテゴリを示します（単位： 1000）。

カテゴリ変数「収入」を作成するには、次の操作を実行します。

- ▶ [データ エディタ] ウィンドウのメニューから、次の項目を選択します。  
変換(T) > 連続変数のカテゴリ化(B)...

図 9-1  
[連続変数のカテゴリ化] ダイアログ ボックス (最初のページ)

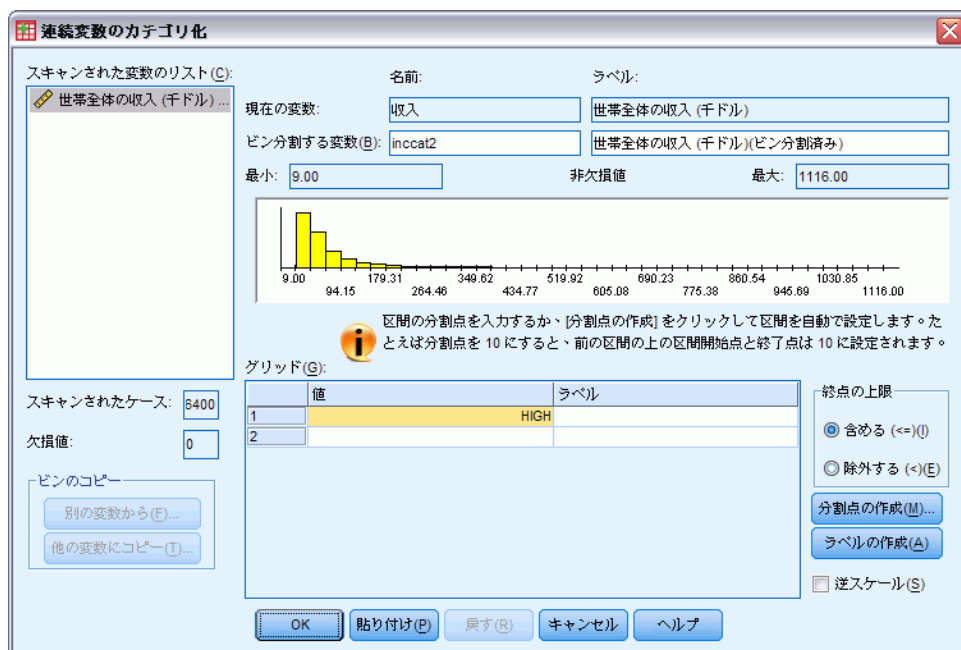


[連続変数のカテゴリ化] ダイアログ ボックスの最初のページでは、新しいビン分割を作成するスケール変数または順序変数、あるいはその両方を選択します。**ビン分割**とは、複数の連続値を同一のカテゴリにグループ化することです。

連続変数のカテゴリ化では、適切なビン分割を行えるよう、データ ファイル内の実際の値を使用します。そのため、最初にデータ ファイルを読み込む必要があります。データ ファイルに大量のケースが含まれている場合、この処理に時間がかかることがあります。そこで、この初期ダイアログ ボックスでは、読み込む（または「スキャン」する）ケースの数を制限できます。このサンプル データ ファイルでは、この処理は必要ありません。サンプル データ ファイルには 6,000 を超えるケースが含まれていますが、これをスキャンするのに時間はそれほどかかりません。

- ▶ [変数] リストから [世帯全体の収入 [収入]] を [ビン (分割)] リストにドラッグ アンド ドロップし、[続行] をクリックします。

図 9-2  
[連続変数のカテゴリ化] ダイアログ ボックス (メインのページ)



- ▶ [連続変数のカテゴリ化] ダイアログ ボックス (メインのページ) の [スキャンされた変数のリスト] で、[世帯全体の収入 [収入]] を選択します。選択した変数の分布がヒストグラムに表示されます (このケースでは、かなりの歪みが見られます)。
- ▶ ビン分割後の新しい変数名として「収入2」と入力し、変数ラベルとして「収入のランク」と入力します。
- ▶ [分割点の作成] をクリックします。

図 9-3  
連続変数のカテゴリ化の [分割点の作成] ダイアログ ボックス

**分割点の作成**

等幅の区間(E)

区間 - 2つ以上のフィールドに入力してください

最初の分割点の位置(F): 25.00

分割点の数(N): 3

幅(W): 25

最後の分割点の位置: 75.00

スキャンされたケースに基づく、等しいパーセンタイル(U)

区間 - どちらかのフィールドに入力してください

分割点の数(N):

幅 (%)(W):

スキャンされたケースに基づく、平均値と選択された標準偏差にある分割点(C)

標準偏差 +/- 1(1)

標準偏差 +/- 2(2)

標準偏差 +/- 3(3)

**i** [適用] をクリックすると、現在の分割点定義がこの設定に置き換わります。  
残りの値はすべて最後の区間に含まれます。区間の数は、分割点の数 + 1 になります。

- ▶ [等幅の区間] を選択します。
- ▶ [最初の分割点の位置] に 25 を、[分割点の数] に 3 を、[幅] に 25 を入力します。

ビン分割後のカテゴリの数は、分割点の数に 1 を加えた数となります。したがってこの例の場合、新しくビン分割した変数のカテゴリは、範囲が 25 (千ドル) の 3 つのカテゴリと、分割点の最大値である 75 (千ドル) を超えるすべての値が含まれる最後のカテゴリの、合計 4 つとなります。

- ▶ [適用] をクリックします。

図 9-4  
[連続変数のカテゴリ化] ダイアログ ボックス (メインのページ) での分割点の定義



定義された分割点がグリッドに値として表示されます。これらの値は、各カテゴリの終点の上限を表します。ヒストグラムに表示されている垂直線は、分割点の位置を示します。

デフォルトでは、これらの分割点の値は、対応するカテゴリに含まれます。たとえば、最初の値 25 には、25 以下のすべての値が含まれます。ただし、この例では、25 未満、25 ~ 49、50 ~ 74、および 75 以上に対応するカテゴリを定義します。

- ▶ [終点の上限] グループで、[除外する (<)] を選択します。
- ▶ 次に、[ラベルの作成] をクリックします。

図 9-5  
自動生成された値ラベル



それぞれのカテゴリに対して、記述的な値ラベルが自動生成されます。新しくビン分割した変数に割り当てられる実際の値は 1 から始まる連続した整数なので、値ラベルを効果的に使用できます。

グリッド内の分割点やラベルは手動で入力または変更できます。さらに、ヒストグラムに表示されている分割点の線をドラッグ アンド ドロップして分割点の位置を変更したり、ヒストグラムの外に分割点の線をドラッグして分割点を削除することもできます。

- ▶ [OK] をクリックして、新しいビン分割を作成します。



データ エディタに新しい変数が表示されます。変数はファイルの最後に追加されるので、[データ ビュー] の右端列および [変数ビュー] の最終行に表示されます。

図 9-6  
データ エディタに表示された新しい変数

	PC	FAX	新聞	反応	収入カテゴリ2	var
1	持っていない	持っていない	購読している	なし	50.00 - 74.00	
2	持っていない	持っていない	購読している	あり	75.00+	
3	持っている	持っていない	購読していない	なし	25.00 - 49.00	
4	持っている	持っている	購読していない	なし	25.00 - 49.00	
5	持っていない	持っていない	購読していない	なし	< 25.00	
6	持っている	持っていない	購読している	なし	75.00+	
7	持っていない	持っていない	購読している	なし	25.00 - 49.00	
8	持っている	持っていない	購読している	なし	50.00 - 74.00	
9	持っていない	持っていない	購読していない	なし	< 25.00	
10	持っていない	持っている	購読していない	あり	75.00+	
11	持っている	持っていない	購読していない	なし	50.00 - 74.00	
12	持っている	持っている	購読していない	なし	< 25.00	
13	持っていない	持っていない	購読していない	なし	25.00 - 49.00	
14	持っている	持っていない	購読している	あり	75.00+	

## 新しい変数の計算

さまざまな数学的機能を使用すると、高度で複雑な方程式に基づいて新しい変数を計算できます。しかし、この例では、簡単に既存の 2 つの変数の値の差分である新しい変数を計算します。

データ ファイル demo.sav には、回答者の現在の年齢に対する変数および現在の仕事の就業年数に対する変数が含まれています。しかし、回答者が仕事を始めたときの年齢に対する変数は含まれていません。現在の年齢と、現在の仕事の就業年数の差分、つまり回答者がその仕事を始めたおおよその年齢を計算して新しい変数を作成することができます。

- ▶ [データ エディタ] ウィンドウのメニューから、次の項目を選択します。  
変換(T) > 変数の計算(C)...
- ▶ [目標変数] に、「就業年齢」と入力します。
- ▶ ソース変数リストの [年齢 [年齢]] を選択し、矢印ボタンをクリックして [数式] テキスト ボックスにコピーします。

- ▶ ダイアログ ボックスの計算パッドでマイナス (-) ボタンをクリックします (またはキーボードでマイナス (-) キーを押します)。
- ▶ [勤続年数 [雇用]] を選択し、矢印ボタンをクリックして式にコピーします。

図 9-7  
[変数の計算] ダイアログ ボックス



注: 現在の職種の変数を選択するときには注意が必要です。不要 な変数を再割り当てしたカテゴリ バージョンもあります。数値式は「年齢 - 勤続カテゴリ」ではなく、[年齢 - 勤続年数] にする必要があります。

- ▶ [OK] をクリックして、新しい変数を計算します。

データ エディタに新しい変数が表示されます。変数はファイルの最後に追加されるので、[データ ビュー] の右端列および [変数ビュー] の最終行に表示されます。

図 9-8  
データ エディタに表示された新しい変数

	FAX	新聞	反応	収入カテゴリ2	就業年齢	var
1	持っていない	購読している	なし	50.00 - 74.00	52.00	
2	持っていない	購読している	あり	75.00+	53.00	
3	持っていない	購読していない	なし	25.00 - 49.00	27.00	
4	持っている	購読していない	なし	25.00 - 49.00	23.00	
5	持っていない	購読していない	なし	< 25.00	23.00	
6	持っていない	購読している	なし	75.00+	43.00	
7	持っていない	購読している	なし	25.00 - 49.00	40.00	
8	持っていない	購読している	なし	50.00 - 74.00	34.00	
9	持っていない	購読していない	なし	< 25.00	44.00	
10	持っている	購読していない	あり	75.00+	32.00	
11	持っていない	購読していない	なし	50.00 - 74.00	54.00	
12	持っている	購読していない	なし	< 25.00	27.00	
13	持っていない	購読していない	なし	25.00 - 49.00	30.00	
14	持っていない	購読している	あり	75.00+	41.00	

## 数式での関数の使用

あらかじめ定義されている関数を数式で使用できます。次の関数を含む、70 個以上の組み込み関数が用意されています。

- 算術関数
- 統計関数
- 分布関数
- 論理関数
- 日付と時刻の集計および抽出関数
- 欠損値関数

- ケース間関数
- 文字型関数

図 9-9  
関数の分類を示す [変数の計算] ダイアログ ボックス



関数は、算術演算を行うグループ、統計的な指標を計算するグループなど、論理的に異なるグループから構成されています。便宜的に、\$TIMEなどの一般に使用される多くのシステム変数も、適切な関数グループに含まれています。現在選択されている関数（この場合は SUM）、またはシステム変数に関する簡単な説明が [変数の計算] ダイアログ ボックスの特定の領域に表示されます。

### 式に関数を貼り付ける

関数を式に貼り付けるには、次の手順を実行します。

- ▶ 式の中で、関数を表示させたい場所にカーソルを置きます。

- ▶ [関数グループ] リストから適切なグループを選択します。[すべて] というグループでは、使用可能な関数およびシステム変数がすべてリスト表示されます。
- ▶ [関数と特殊変数] リスト内の関数をダブルクリックします（または、関数を選択し、[関数グループ] リストの隣にある矢印をクリックします）。  
式に関数が挿入されます。式の一部を強調表示して関数を挿入すると、強調表示した部分が関数の最初の引数として使用されます。

### 式の関数を編集する

関数は、引数を入力するまで完成しません。引数は、貼り付けられた関数にクエスチョン マークで表されます。クエスチョン マークの数は、関数を完成させるために必要な引数の最小数を示します。

- ▶ 貼り付けられた関数のクエスチョン マークを強調表示します。
- ▶ 引数を入力します。引数に変数名の場合は、変数リストから貼り付けることができます。

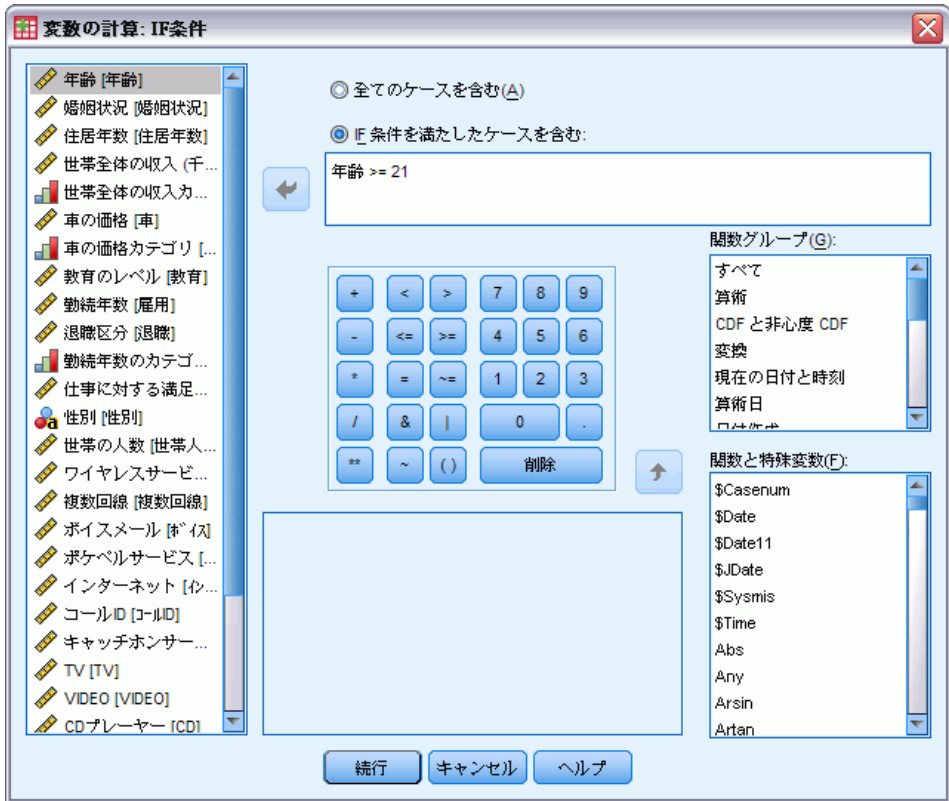
## 条件式の使用

条件式（論理式とも呼ばれます）を使用して、選択したケースのサブセットだけにデータの変換を適用することができます。条件式は、各ケースに真の値、偽の値、または欠損値のいずれかの結果を返します。条件式の結果が「真」であれば、そのケースに変換が適用されます。結果が「偽」または「欠損」の場合、ケースは変換されません。

条件式を指定するには、次の手順を実行します。

- ▶ [変数の計算] ダイアログ ボックスで [f] をクリックします。[変数の計算: IF 条件] ダイアログ ボックスが開きます。

図 9-10  
[変数の計算: IF 条件] ダイアログ ボックス



- ▶ [If 条件を満たしたケースを含む] チェック ボックスをオンにします。
- ▶ 条件式を入力します。

次に示すように、ほとんどの条件式には、関係演算子が少なくとも 1 つ含まれています。

年齢 >= 21

または

収入 \* 3 < 100

最初の例では、[年齢 [年齢]] が 21 以上のケースだけが選択されます。2 つ目の例では、[世帯全体の収入 [収入]] \* 3 の結果が 100 未満となるケースが選択されます。

また、次に示すように、論理演算子を使用して 2 つ以上の条件式をリンクさせることもできます。

年齢 >= 21 | 学歴 >= 4

または

収入\*3<100 & 学歴=5

最初の例では、[年齢 [年齢]] 条件または [教育のレベル [教育]] 条件のいずれかを満たすケースが選択されます。2 つ目の例では、[世帯全体の収入 [収入]] と [教育のレベル [教育]] の両方の条件を満たすケースが選択されます。

## 日付と時刻を使用した作業

日付と時刻を使用して実行する多くの一般的な作業は、日付と時刻ウィザードを使用して実行できます。このウィザードを使用すると、次の操作を実行できます。

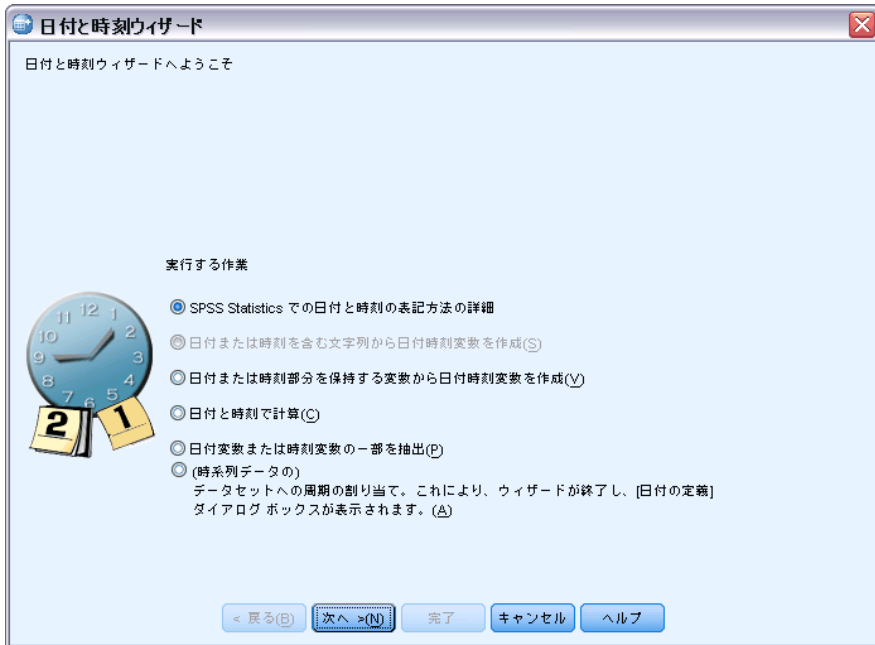
- 日付または時刻を含む文字列から日付時刻変数を作成。
- 日付または時刻の異なる部分が含まれた変数を結合することにより、日付時刻変数を作成する。
- 日付時刻変数に対する加算または減算を行う。2 つの日付時刻変数の加算または減算も含まれます。
- 日付変数または時刻変数の一部を抽出する。たとえば、mm/dd/yyyy 形式の日付時刻変数から、月初からの日数を抽出できます。

この項の例では、データ ファイル `upgrade.sav` を使用します。詳細は、[A 付録 p. 165 サンプル ファイル](#) を参照してください。

日付と時刻ウィザードを使用するには、次の手順を実行します。

- ▶ メニューから次の項目を選択します。  
変換(T) > 日付と時刻ウィザード(D)...

図 9-11  
日付と時刻ウィザードの開始画面



日付と時刻ウィザードの開始画面では、一般的なタスクのセットが提供されます。現在のデータに当てはまらないタスクは無効です。たとえば、データ ファイル「upgrade.sav」には文字型変数が含まれていないので、文字列から日付変数を作成するタスクは無効です。

IBM® SPSS® Statistics の日付と時刻に慣れていない場合は、[日付と時刻の表記方法の詳細] を選択して、[次へ] をクリックできます。これにより、日付時刻変数の簡単な概要が記述された画面が表示されます。[ヘルプ] ボタンをクリックするとリンクが表示され、より詳しい情報を参照できます。

## 2 つの日付間の時間の長さの計算

日付に関する最も一般的なタスクは、2 つの日付間の時間の長さを計算することです。例として、あるソフトウェア企業が、各顧客が最後にアップグレードを行ってから経過した年数を特定することによりアップグレードライセンスの購入を分析することに関心を持っているとします。データ ファイル「upgrade.sav」には、各顧客が最後にアップグレードを購入した日付が含まれていますが、購入日からの経過年数は含まれていません。最後のアップグレードの日付と次の製品リリースの日付間の時間の長さを年単位で表す新しい変数を作成することにより、この年数を測定します。

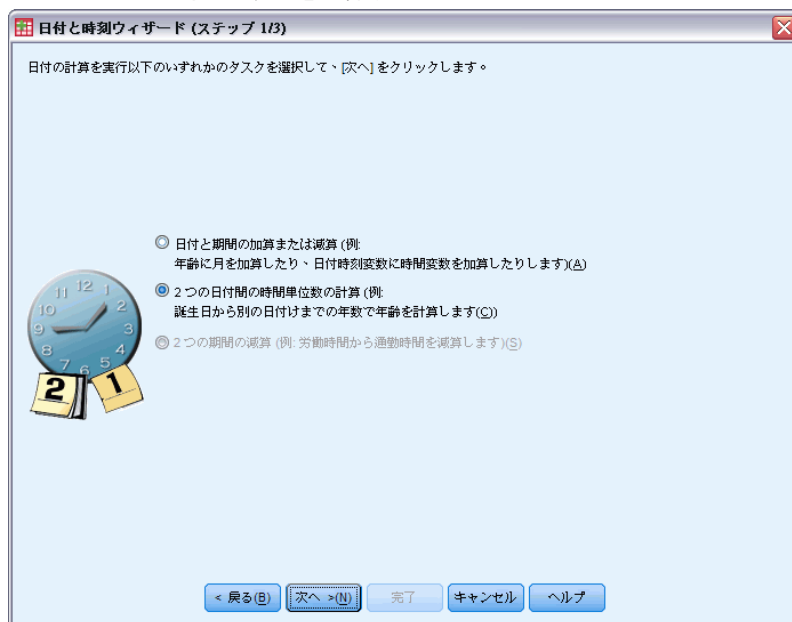
2 つの日付間の時間の長さを計算するには、次の手順を実行します。



- ▶ 日付と時刻ウィザードの開始画面で [日付と時刻で計算] を選択し、[次へ] をクリックします。

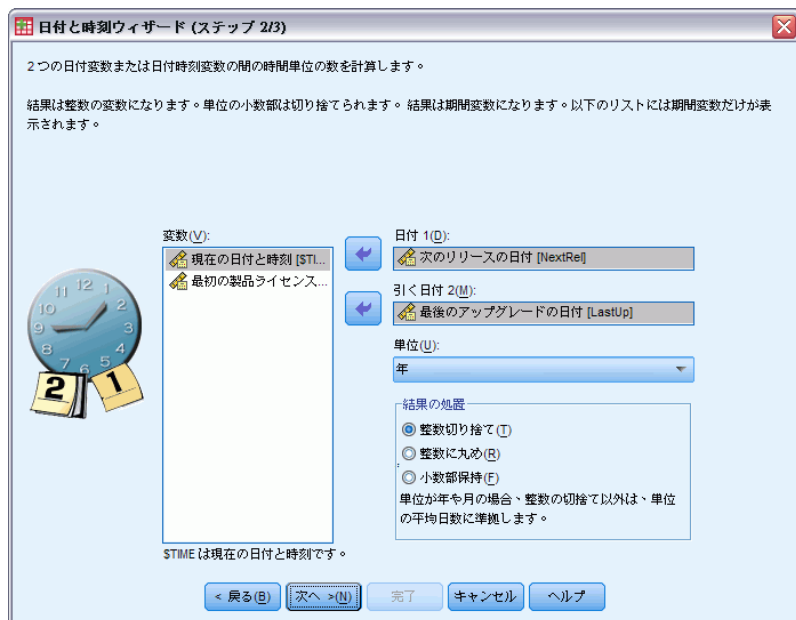
図 9-12

## 2つの日付間の時間の長さを計算する: ステップ 1



- ▶ [2つの日付間の時間単位数の計算] を選択し、[次へ] をクリックします。

図 9-13  
2つの日付間の時間の長さを計算する: ステップ 2



- ▶ [日付] で [次のリリースの日付 [NextRel]] を選択します。
- ▶ [期間変数] で [最後のアップグレードの日付 [LastUp]] を選択します。
- ▶ 単位は [年] を、結果の処理方法は [整数に切り捨て] を選択します。(これらはデフォルトで選択されています)。
- ▶ [次へ] をクリックします。

図 9-14  
2つの日付間の時間の長さを計算する: ステップ 3



- ▶ 結果変数の名前として「YearsLastUp」と入力します。結果変数に既存の変数と同じ名前を付けることはできません。
- ▶ 結果変数のラベルとして「最後のアップデートからの年数」と入力します。結果変数の変数ラベルはオプションです。
- ▶ [今すぐ変数を作成する] の選択内容をデフォルトのままにし、[完了] をクリックして新しい変数を作成します。

データ エディタに表示される新しい変数 YearsLastUp は、2つの日付間の年数を表す整数です。年の小数点以下は切り捨てられています。

図 9-15  
データ エディタに表示された新しい変数

	custid	PurDate	Support	LastUp	NextRel	YearLastUp	
1	1	12/30/1998	4	02/28/2002	06/01/2004	2	
2	2	06/28/2001	2	09/28/2002	06/01/2004	1	
3	3	08/27/1999	2	09/27/2001	06/01/2004	2	
4	4	02/22/2000	4	01/22/2003	06/01/2004	1	
5	5	01/26/2000	2	08/26/2001	06/01/2004	2	
6	6	07/10/1999	3	07/10/2003	06/01/2004	0	
7	7	01/24/2003	2	07/24/2003	06/01/2004	0	
8	8	06/15/1999	2	09/15/2003	06/01/2004	0	
9	9	01/18/2003	5	07/18/2003	06/01/2004	0	
10	10	12/02/2002	4	06/02/2003	06/01/2004	0	
11	11	08/10/2000	1	10/10/2002	06/01/2004	1	
12	12	05/27/1999	2	07/27/2000	06/01/2004	3	
13	13	02/28/1999	4	10/28/2002	06/01/2004	1	

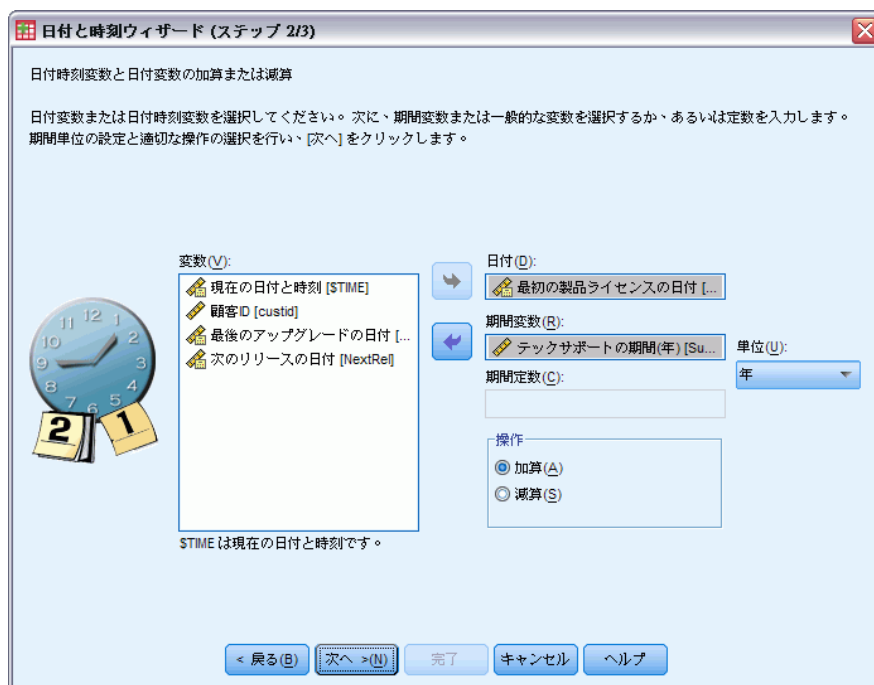
## 日付への期間の加算

日付に対して 10 日または 12 か月などの期間を加算したり減算したりすることができます。引き続き、前の項で示したソフトウェア企業の例を使用して、各顧客の初期技術サポート契約が終了する日付の特定について考えます。データ ファイル「upgrade.sav」には、契約したサポート年数が格納された変数と、最初の購入日付が格納された変数が含まれています。サポート年数を購入日付に加算することにより、初期サポートの終了日付を決定できます。

日付に期間を加算するには、次の手順を実行します。

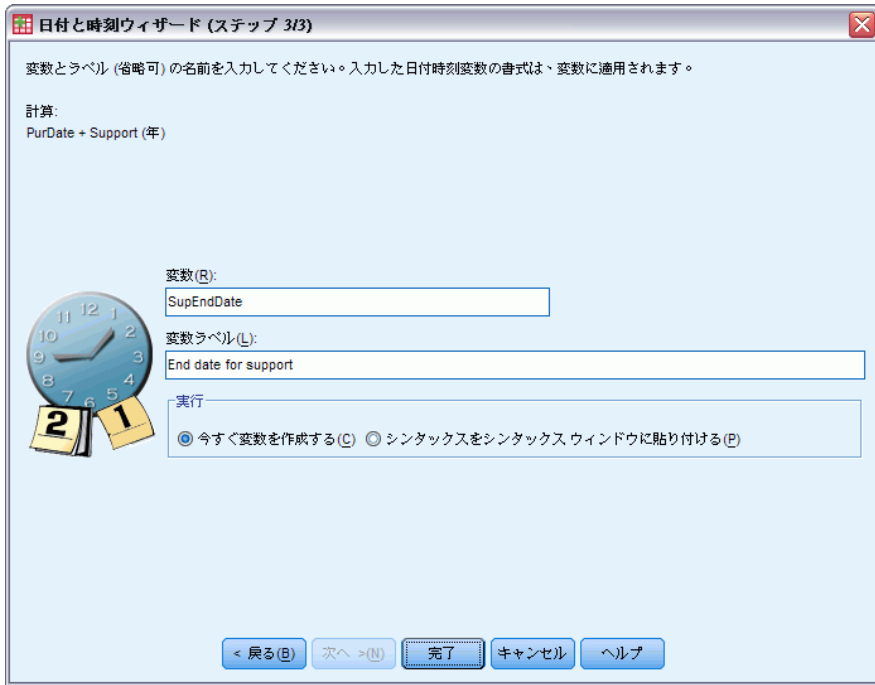
- ▶ 日付と時刻ウィザードの開始画面で [日付と時刻で計算] を選択し、[次へ] をクリックします。
- ▶ [日付と期間の加算または減算] を選択し、[次へ] をクリックします。

図 9-16  
日付への期間の加算: ステップ 2



- ▶ [日付] で [最初の製品ライセンスの日付 [PurDate]] を選択します。
- ▶ [期間変数] で [テックサポートの期間(年) [Support]] を選択します。  
テックサポートの期間(年) [Support] は単なる数値型変数なので、この変数を期間として加算する際に使用する単位を指定する必要があります。
- ▶ [単位] ドロップダウン リストから、[年] を選択します。
- ▶ [次へ] をクリックします。

図 9-17  
日付への期間の加算: ステップ 3



- ▶ 結果変数の名前として「SupEndDate」と入力します。結果変数に既存の変数と同じ名前を付けることはできません。
- ▶ 結果変数のラベルとして「サポートの終了日」と入力します。結果変数の変数ラベルはオプションです。
- ▶ [完了] をクリックして新しい変数を作成します。

データ エディタに新しい変数が表示されます。

図 9-18  
データ エディタに表示された新しい変数

The screenshot shows the SPSS Data Editor window for a file named \*upgrade.sav. The window title is "データセット1 - データ エディタ(D)". The menu bar includes options like "ファイル(F)", "編集(E)", "表示(V)", "データ(D)", "変換(T)", "分析(A)", "グラフ(G)", "ユーティリティ(U)", "アドオン(O)", "ウィンドウ(W)", and "ヘルプ(H)". The main area displays a data table with the following columns: "Date", "Support", "LastUp", "NextRel", "YearLastUp", and "SupEndDate". The first column is labeled "1: custid" and contains values from 1 to 14. The "Date" column contains dates from 1998 to 2003. The "Support" column contains values from 1 to 5. The "LastUp" column contains dates from 2001 to 2003. The "NextRel" column contains dates from 2001 to 2004. The "YearLastUp" column contains values from 0 to 3. The "SupEndDate" column contains dates from 2001 to 2008. The status bar at the bottom shows "データ ビュー(D)" and "変数 ビュー(V)".

	Date	Support	LastUp	NextRel	YearLastUp	SupEndDate
1	2/30/1998	4	02/28/2002	06/01/2004	2	12/30/2002
2	5/28/2001	2	09/28/2002	06/01/2004	1	06/28/2003
3	3/27/1999	2	09/27/2001	06/01/2004	2	08/27/2001
4	2/22/2000	4	01/22/2003	06/01/2004	1	02/22/2004
5	1/26/2000	2	08/26/2001	06/01/2004	2	01/26/2002
6	7/10/1999	3	07/10/2003	06/01/2004	0	07/10/2002
7	1/24/2003	2	07/24/2003	06/01/2004	0	01/24/2005
8	5/15/1999	2	09/15/2003	06/01/2004	0	06/15/2001
9	1/18/2003	5	07/18/2003	06/01/2004	0	01/18/2008
10	2/02/2002	4	06/02/2003	06/01/2004	0	12/02/2006
11	3/10/2000	1	10/10/2002	06/01/2004	1	08/10/2001
12	5/27/1999	2	07/27/2000	06/01/2004	3	05/27/2001
13	2/28/1999	4	10/28/2002	06/01/2004	1	02/28/2003
14	1/22/2004	5	07/02/2004	06/01/2004	2	01/22/2006

# データの並べ替えと選択

データ ファイルは、特定の要求に合った理想的な形式で構成されているとは限りません。分析用のデータを準備するために、次に示す機能をはじめとする、さまざまなファイル変換機能を選択することができます。

- **データの並べ替え。**1 つ以上の変数値に基づいてケースの並べ替えができます。
- **ケースのサブグループの選択。**分析をケースのサブグループに制限することもできますが、異なるサブグループに対して同時に分析を実行することもできます。

この章の例では、データ ファイル demo.sav を使用します。 [詳細は、A 付録 p.165 サンプル ファイル を参照してください。](#)

## データを並べ替える

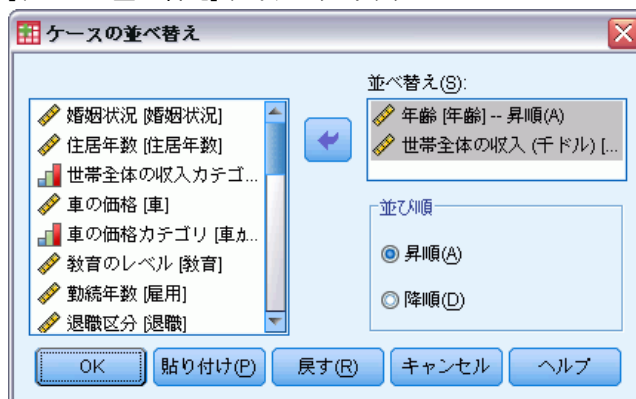
ケースの並べ替え（データ ファイルの行の並べ替え）は、ある種の分析で役立つもので、ときには必ず必要になります。

1 つ以上の並べ替え変数の値に基づいてデータ ファイルのケースを並べ替えるには、次の操作を行います。

- ▶ メニューから次の項目を選択します。  
データ > ケースの並べ替え...

[ケースの並べ替え] ダイアログ ボックスが表示されます。

図 10-1  
[ケースの並べ替え] ダイアログ ボックス





- ▶ 変数 [年齢 [年齢]] および [世帯全体の収入 (千ドル) [収入]] を、[並べ替え] リストに追加します。

複数の並べ替え変数を選択した場合、ケースの並べ替えの順序は、[並べ替え] リストに表示されている変数の順序で決まります。この例では [並べ替え] リストのエントリに基づいて、ケースが [年齢 [年齢]] のカテゴリ内で [世帯全体の収入 (千ドル) [収入]] の値によって並べ替えられます。文字型変数では、値がアルファベットの場合、並べ替え順序として大文字が小文字よりも優先されます (たとえば、文字型値「Yes」は [yes] よりも優先されます)。

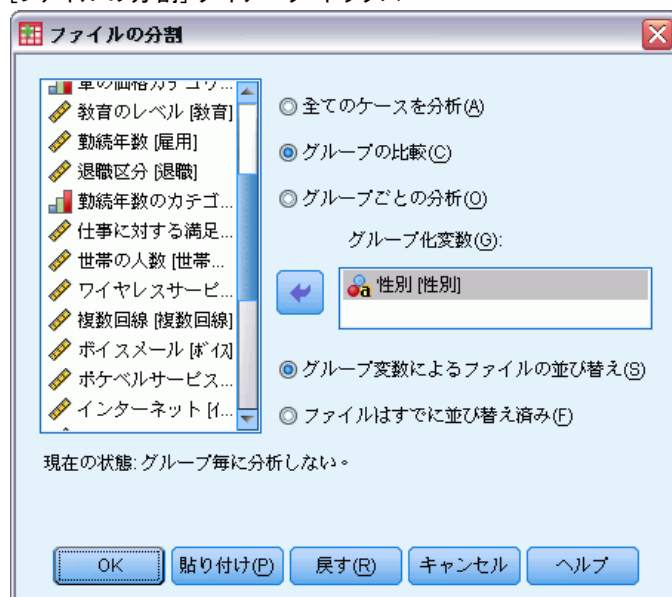
## ファイルの分割処理

分析用にデータ ファイルを複数のグループに分割するには、次の手順を実行します。

- ▶ メニューから次の項目を選択します。  
データ > ファイルの分割...

[ファイルの分割] ダイアログ ボックスが表示されます。

図 10-2  
[ファイルの分割] ダイアログ ボックス



- ▶ [グループの比較] または [グループごとの分析] を選択します。(この手順の後の例で、この 2 つのオプションの違いを説明します)。

- ▶ ファイルを別々のグループに分割する変数として、[性別 [性別]] を選択します。

グループ化変数として、数値型、短い文字型、および長い文字型の変数を使用できます。グループ化変数によって定義された各サブグループに対して、別々に分析が実行されます。複数のグループ化変数を選択した場合、ケースがグループ化される方法は、[グループ化変数] リストに表示されている変数の順序で決まります。

[グループの比較] を選択すると、[度数分布表] 手続きによって生成された以下の要約統計量のテーブルのように、ファイルの分割グループのすべてが、同じテーブルに入ります。

図 10-3

1 つのピボット テーブルが表示された [ファイルの分割] 出力

統計量			
世帯全体の収入 (千ドル)			
女性	度数	有効	3179
		欠損値	0
	平均値		68.7798
	中央値		44.0000
	標準偏差		75.73510
男性	度数	有効	3221
		欠損値	0
	平均値		70.1608
	中央値		45.0000
	標準偏差		81.56216

[グループごとの分析] を選択して [度数分布表] 手続きを実行した場合、男性と女性の 2 つのピボット テーブルが作成されます。

図 10-4

女性のデータを表すピボット テーブルが表示された [ファイルの分割] 出力

統計量 <sup>a</sup>		
世帯全体の収入 (千ドル)		
度数	有効	3179
	欠損値	0
平均値		68.7798
中央値		44.0000
標準偏差		75.73510

<sup>a</sup> 性別 = 女性

図 10-5

男性のデータを表すピボット テーブルが表示された [ファイルの分割] 出力

統計量 <sup>a</sup>		
世帯全体の収入 (千ドル)		
度数	有効	3221
	欠損値	0
平均値		70.1608
中央値		45.0000
標準偏差		81.56216

<sup>a</sup> 性別 = 男性

## ファイルの分割処理用にケースを並べ替える

ファイルの分割手続きでは、グループ化変数のいずれかに異なる値が出現するたびに、新しいサブグループが作成されます。そのため、ファイルの分割処理を起動する前に、グループ化変数の値に基づいてケースを並べ替えておくことが重要です。

ファイルの分割では、デフォルトで、グループ化変数の値に基づいてデータ ファイルが自動的に並べ替えられます。ファイルがすでに正しい順序に並べ替えられている場合は、[ファイルは既に並べ替え済み] を選択すると、処理時間を節約できます。

## ファイルの分割処理のオンとオフを切り替える

起動したファイルの分割処理は、オフにしない限り、そのセッションの間中有効のままです。

- **全てのケースを分析。**このオプションは、ファイルの分割処理をオフにします。
- **グループの比較またはグループごとの分析。**このオプションは、ファイルの分割処理をオンにします。

ファイルの分割処理が有効になっている場合は、アプリケーション ウィンドウの下部のステータス バーに、「分割ファイル オン」というメッセージが表示されます。

## ケースのサブグループを選択する

変数および複雑な式を使った基準に基づいて、分析を特定のサブグループに制限することができます。ケースの無作為抽出を選択することもできます。サブグループを定義するときに使用する基準には、次のものを含めることができます。

- 変数の値と範囲
- 日付と時刻の範囲
- ケース（行）番号
- 算術式
- 論理式
- 関数

ケースのサブセットを分析のために選択するには、次の手順を実行します。

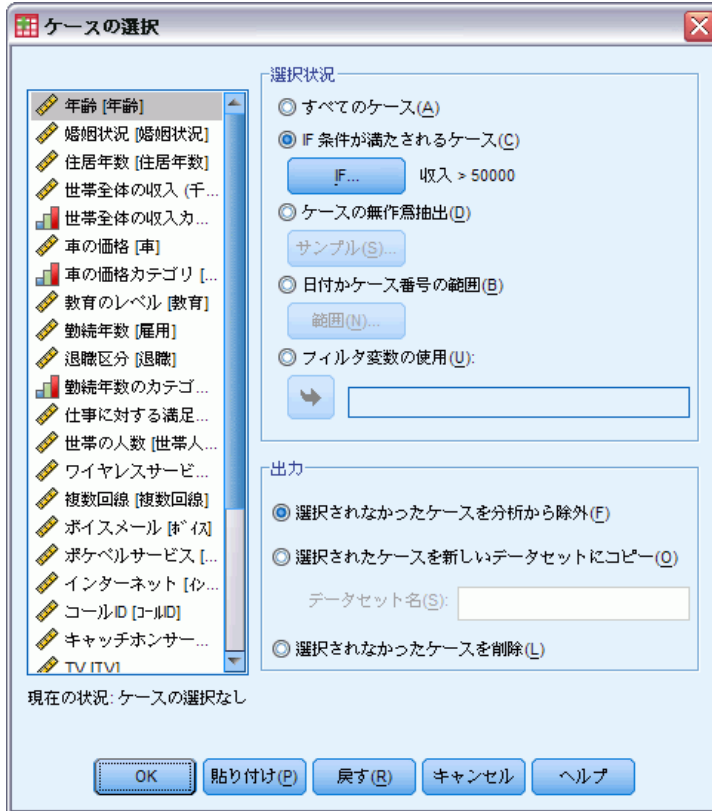
- ▶ メニューから次の項目を選択します。

データ > ケースの選択...

[ケースの選択] ダイアログ ボックスが開きます。

図 10-6

[ケースの選択] ダイアログ ボックス



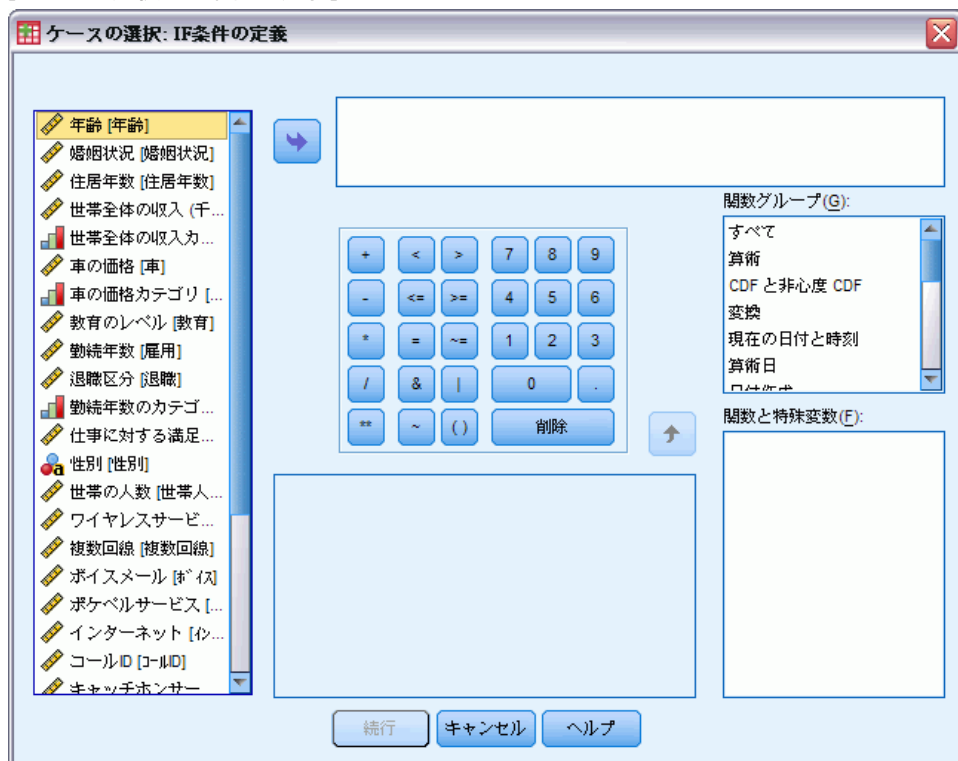
## 条件式に基づいたケースを選択する

条件式に基づいてケースを選択するには、次の手順を実行します。

- ▶ [ケースの選択] ダイアログ ボックスで、[IF 条件が満たされるケース] を選択して [IF] をクリックします。

[ケースの選択: IF 条件の定義] ダイアログ ボックスが開きます。

図 10-7  
[ケースの選択: IF 条件の定義] ダイアログ ボックス



条件式には、既存の変数名、定数、算術演算子、論理演算子、関係演算子、関数を使うことができます。テキスト ボックスでは、出力ウィンドウのテキストと同様に、式を入力したり編集することができます。また、計算パッド、変数リスト、および関数リストを使って式に要素を貼り付けることもできます。 [詳細は、9 章 p.145 条件式の使用 を参照してください。](#)

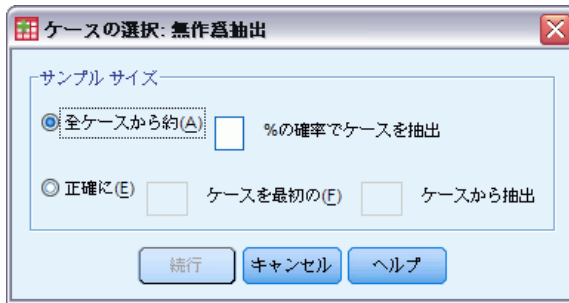
## 無作為抽出を選択する

無作為抽出を行うには、次の手順を実行します。

- ▶ [ケースの選択] ダイアログ ボックスで、[ケースの無作為抽出] を選択します。
- ▶ [サンプル] をクリックします。

[ケースの選択: 無作為抽出] ダイアログ ボックスが開きます。

図 10-8  
[ケースの選択: 無作為抽出] ダイアログ ボックス



サンプル サイズには、次のオプションのどちらかを選択できます。

- **全ケースの。**ユーザーがパーセントを指定します。このオプションは、指定されたパーセントとほぼ同じ割合でケースから任意のサンプルを選択します。
- **正確に。**ケースの数をユーザーが指定します。いくつかのケースからサンプルを生成するかも指定しなければなりません。この 2 番目の数はデータ ファイル内にあるケースの総数以下でなければなりません。この数がデータ ファイル内にあるケースの総数よりも多いと、その数に比例してサンプルのケースは要求された数よりも少なくなってしまう。

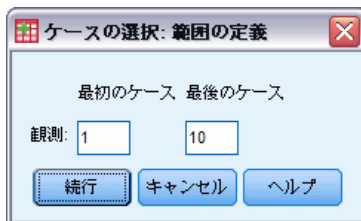
## 日付の範囲またはケース番号の範囲を選択する

日付、時刻、観測番号（行番号）に基づいてケースの範囲を選択するには、次の手順を実行します。

- ▶ [ケースの選択] ダイアログ ボックスで、[日付かケース番号の範囲] を選択して [範囲] をクリックします。

[ケースの選択: 範囲の定義] ダイアログ ボックスが開きます。ここで、観測番号（行番号）の範囲を選択できます。

図 10-9  
[ケースの選択: 範囲の定義] ダイアログ ボックス



- **最初のケース。**範囲の開始日付または時刻の値、あるいは両方を入力します。日付関数が定義されていない場合は、開始観測番号（データ エディタの行番号。ただし、ファイルの分割がオンになっていない場合）

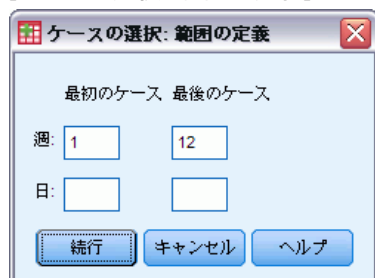
を入力します。[最後のケース]の値を指定しないと、開始日付と時刻から時系列の終了までのすべてのケースが選択されます。

- **最後のケース。**範囲の終了日付または時刻、あるいは両方の値を入力します。日付関数が定義されていない場合は、終了観測番号（データ エディタの行番号。ただし、ファイルの分割がオンになっていない場合）を入力します。[最初のケース]の値を指定しないと、時系列の先頭から終了日付と時刻までのすべてのケースが選択されます。

日付変数が定義された時系列データの場合、定義されている日付変数に基づいて日付と時刻の範囲を選択できます。各ケースが異なる時刻の観測数を表し、ファイルは時系列順に並べ替えられます。

図 10-10

[ケースの選択: 範囲の定義] ダイアログ ボックス (時系列)



時系列データの日付変数を生成するには、次の手順を実行します。

- ▶ メニューから次の項目を選択します。  
データ > 日付の定義...

## 選択されなかったケースの処理

選択されなかったケースの処理には、次のオプションのどちらかを選択できます。

- **選択されなかったケースを分析から除外。**選択されなかったケースは分析から除外されますが、データセットには残っています。[分析から除外]をオフにすると、選択されなかったケースをこのセッションの後に使用できます。無作為抽出を選択するか、条件式に基づいてケースを選択した場合、「filter\_\$」という名前の変数が生成されます。選択されたケースではこの変数に 1 が、選択されなかったケースでは 0 が含まれます。

- **選択されたケースを新しいデータセットにコピー。**新しいデータセットに、選択したケースがコピーされます。元のデータセットには、影響はありません。選択していないケースは、新しいデータセットにはコピーされず、そのまま元のデータセットの中に残っています。
- **選択されなかったケースの削除。**選択されなかったケースは、データセットから削除されます。削除されたケースを回復するには、変更を保存しないでファイルを終了し、再びファイルを開くしかありません。データファイルに変更を保存すると、ケースは永久に削除されます。

**注:** 選択されなかったケースを削除してファイルを保存した場合は、そのケースを回復することはできません。

## ケースの選択状態

あるケースのグループが選択されていて、かつ選択されていないケースが削除されていない場合、選択されなかったケースには、データ エディタの行番号の上に斜めの線が引かれます。

図 10-11  
ケースの選択状態

	年齢	婚姻状況	住居年数	収入	収入カテゴリ	車	車カテゴリ
<del>5</del>	25	0	2	23.00	1.00	11.30	1.00
6	45	1	9	76.00	4.00	37.20	3.00
7	42	0	19	40.00	2.00	19.80	2.00
8	35	0	15	57.00	3.00	28.20	2.00
9	46	0	26	24.00	1.00	12.20	1.00
<del>10</del>	34	1	0	89.00	4.00	46.10	3.00
11	55	1	17	72.00	3.00	35.50	3.00
12	28	0	3	24.00	1.00	11.80	1.00
<del>13</del>	31	1	9	40.00	2.00	21.30	2.00



# サンプル ファイル

製品とともにインストールされるサンプル ファイルは、インストールディレクトリの Samples サブディレクトリにあります。[サンプル] サブディレクトリ内に次の各言語の別のフォルダがあります。英語、フランス語、ドイツ語、イタリア語、日本語、韓国語、ポーランド語、ロシア語、簡体字中国語、スペイン語、そして繁体中国語です。

すべてのサンプル ファイルが、すべての言語で使用できるわけではありません。サンプル ファイルがある言語で使用できない場合、その言語のフォルダには、サンプル ファイルの英語バージョンが含まれています。

## 説明

以下は、このドキュメントのさまざまな例で使用されているサンプル ファイルの簡単な説明です。

- **accidents.sav**。与えられた地域での自動車事故の危険因子を年齢および性別ごとに調べている保険会社に関する架空のデータ ファイルです。各ケースが、年齢カテゴリと性別のクロス分類に対応します。
- **adl.sav**。脳卒中患者に提案される治療の効果を特定するための取り組みに関する架空のデータ ファイルです。医師団は、女性の脳卒中患者たちを、2 つのグループのいずれかにランダムに割り当てました。一方のグループは標準的な理学療法を受け、もう一方のグループは感情面の治療も追加で受けました。治療の 3 か月後に、各患者が日常生活の一般的な行動をどの程度とることができるかを、順序変数として得点付けしました。
- **advert.sav**。広告費とその売上成果の関係を調べるための小売業者の取り組みに関する架空のデータ ファイルです。この小売業者は、そのために、過去の売上と、それに関する広告費のデータを収集しました。
- **aflatoxin.sav**。収穫物によって濃度が大きく異なる毒物であるアフラトキシンを、トウモロコシの収穫物に関して検定することに関する架空のデータ ファイルです。ある穀物加工業者は、8 つそれぞれの収穫物から 16 のサンプルを受け取って、10 億分の 1 単位でアフラトキシン レベルを測定しました。
- **anorectic.sav**。拒食行動または過食行動の標準的な症状の特定を目指して、調査員 が、摂食障害を持つ大人 55 人の調査を行いました。各患者が 4 年間で 4 回診察を受けたので、観測値は合計で 220 になりました。観測値ごとに、16 種類の症状に関して患者の得点が記録

されました。患者 71 (2 回目)、患者 76 (2 回目)、患者 47 (3 回目) の症状の得点が見つからなかったので、残っている 217 回分の観測値が有効です。

- **bankloan.sav.** 債務不履行率を低減させるための銀行の取り組みに関する架空のデータ ファイルです。このファイルには、過去の顧客および見込み客 850 人に関する財務情報と人口統計情報が含まれています。最初の 700 ケースは、以前に貸付を行った顧客です。残りの 150 ケースは見込み顧客で、これらの顧客に関して銀行は信用リスクの良し悪しを分類する必要があります。
- **bankloan\_binning.sav.** 過去の顧客 5,000 人に関する財務情報と人口統計情報を含む架空のデータ ファイルです。
- **behavior.sav.** 52 人の学生に 15 の状況と 15 の行動の組み合わせについて、0 = 「非常に適切」から 9 = 「非常に不適切」までの 10 段階でランク付けするよう依頼した研究があります。個人間の平均を取ったため、値は非類似度としてみなされます。
- **behavior\_ini.sav.** このデータ ファイルには、behavior.sav の 2 次元の解の初期配置が含まれています。
- **brakes.sav.** 高性能自動車のディスク ブレーキを生産している工場での品質管理に関する架空のデータ ファイルです。このデータ ファイルには、8 台の機械で生産した 16 個のディスクの直径測定値が含まれています。ブレーキの目標の直径は 322 ミリメートルです。
- **breakfast.sav.** 21 人の Wharton School MBA の学生およびその配偶者に、15 種類の朝食を好みの順に (1 = 「最も好き」から 15 = 「最も嫌い」まで) ランク付けするよう依頼した研究があります。調査対象者の嗜好は、「すべて」から「スナックとドリンクのみ」まで、6 つの異なるシナリオに基づいて記録されました。
- **breakfast-overall.sav.** このデータ ファイルには、最初のシナリオ (「すべて」) のみの朝食の好みが含まれています。
- **broadband\_1.sav.** 全国規模のブロードバンド サービスの地域ごとの契約者数を含む架空のデータ ファイルです。このデータ ファイルには、85 地域の月々の契約者数が 4 年間分含まれています。
- **broadband\_2.sav.** このデータ ファイルは broadband\_1.sav と同じですが、データが 3 か月分追加されています。
- **car\_insurance\_claims.sav.** 他の場所 で表示および分析される、自動車の損害請求に関するデータセットです。逆リンク関数を使用して従属変数の平均値を保険契約者の年齢、車種、製造年の線型結合と関連付けることにより、平均請求数はガンマ分布としてモデリングできます。申請された請求の数は、尺度重み付けとして使用できます。
- **car\_sales.sav.** このデータ ファイルには、自動車のさまざまな車種やモデルの架空の売上推定値、定価、仕様が含まれています。定価と仕様はそれぞれ、edmunds.com と製造元のサイトから入手しました。

- **car\_sales\_upprepared.sav**。変換したバージョンのフィールドを含まない car\_sales.sav の修正したバージョンです。
- **carpet.sav**。一般的な例 としては、新しいカーペット専用洗剤を市販することに関心のある企業が消費者の嗜好に関する 5 種類の因子（パッケージのデザイン、ブランド名、価格、サービスシール、料金の払い戻し）の影響について調べたい場合があります。パッケージのデザインには、3 つの因子レベルがあります。それぞれ塗布用ブラシの位置が異なります。また、3 つのブランド名（K2R、Glory、および Bissell）、3 つの価格水準があり、最後の 2 つの因子のそれぞれに対しては 2 つのレベル（「なし」または「あり」）があります。10 人の消費者が、これらの因子により定義された 22 個のプロファイルに順位を付けます。変数「嗜好」には、各プロファイルの平均順位の序列が含まれています。順位が低いほど、嗜好度は高くなります。この変数には、各プロファイルの嗜好測定値がすべて反映されます。
- **carpet\_prefs.sav**。このデータ ファイルは carpet.sav と同じ例に基づいていますが、10 人の消費者それぞれから収集した実際のランキングが含まれています。消費者は、22 種類の製品プロファイルを、一番好きなものから一番嫌いなものまで順位付けすることを依頼されています。変数 PREF1 から PREF22 には、carpet\_plan.sav で定義されている、関連するプロファイルの ID が含まれています。
- **catalog.sav**。このデータ ファイルには、あるカタログ会社が販売した 3 つの製品の、架空の月間売上高が含まれています。5 つの予測変数のデータも含まれています。
- **catalog\_seasfac.sav**。このデータ ファイルは catalog.sav と同じですが、季節性の分解手続きとそれに付随する日付変数から計算した一連の季節因子が追加されています。
- **cellular.sav**。解約率を削減するための携帯電話会社の取り組みに関する架空のデータ ファイルです。解約の傾向スコアは、0 ~ 100 の範囲でアカウントに適用されます。スコアリングが 50 以上のアカウントはプロバイダの変更を考えている場合があります。
- **ceramics.sav**。新しい上質の合金に標準的な合金より高い耐熱性があるかどうかを特定するための、ある製造業者の取り組みに関する架空のデータ ファイルです。各ケースが 1 つの合金の別々のテストを表し、軸受けの耐熱温度が記録されます。
- **cereal.sav**。880 人を対象に、朝食の好みについて、年齢、性別、婚姻状況、ライフスタイルが活動的かどうか（週 2 回以上運動するか）を含めて調査した、架空のデータ ファイルです。各ケースが別々の回答者を表します。
- **clothing\_defects.sav**。ある衣料品工場での品質管理工程に関する架空のデータ ファイルです。工場で生産される各ロットから、調査員が衣料品のサンプルを取り出し、不良品の数を数えます。

- **coffee.sav.** このデータ ファイルは、6 つのアイスコーヒー ブランド について受けた印象に関連しています。回答者は、アイス コーヒーに対する 23 の各印象属性に対して、その属性が言い表していると思われるすべてのブランドを選択しました。機密保持のため、6 つのブランドを AA、BB、CC、DD、EE、および FF で表しています。
- **contacts.sav.** 企業のコンピュータ営業グループの担当者リストに関する架空のデータ ファイルです。各担当者は、所属する会社の部門および会社のランクによって分類されています。また、最新の販売金額、最後の販売以降の経過時間、担当者の会社の規模も記録されています。
- **creditpromo.sav.** 最近のクレジット カード プロモーションの有効性を評価するための、あるデパートの取り組みに関する架空のデータ ファイルです。このために、500 人のカード所有者がランダムに選択されました。そのうち半分には、今後 3 か月間の買い物に関して利率を下げることをプロモーションする広告を送付しました。残り半分には、通常どおりの定期的な広告を送付しました。
- **customer\_dbase.sav.** 自社のデータ ウェアハウスにある情報を使用して、反応がありそうな顧客に対して特典を提供するための、ある会社の取り組みに関する架空のデータ ファイルです。顧客ベースのサブセットをランダムに選択して特典を提供し、顧客の反応が記録されています。
- **customer\_information.sav.** 名前や住所など、顧客の連絡先情報を含む架空のデータ ファイルです。
- **customer\_subset.sav.** customer\_dbase.sav の 80 件のケースのサブセット。
- **debate.sav.** 政治討論の出席者に対して行った調査の、討論の前後それぞれの回答に関する架空のデータ ファイルです。各ケースが別々の回答者に対応します。
- **debate\_aggregate.sav.** debate.sav 内の回答を集計する、架空のデータ ファイルです。各ケースが、討論前後の好みのクロス分類に対応しています。
- **demo.sav.** 月々の特典を送付することを目的とした、購入顧客のデータベースに関する架空のデータ ファイルです。顧客が特典に反応したかどうか、さまざまな人口統計情報と共に記録されています。
- **demo\_cs\_1.sav.** 調査情報のデータベースをコンパイルするための、ある会社の取り組みの最初のステップに関する架空のデータ ファイルです。各ケースが別々の都市に対応し、地域、地方、地区、および都市の ID が記録されています。
- **demo\_cs\_2.sav.** 調査情報のデータベースをコンパイルするための、ある会社の取り組みの第 2 のステップに関する架空のデータ ファイルです。各ケースが、最初のステップで選択した都市の別々の世帯単位に対応し、地域、地方、地区、都市、区画、および単位の ID が記録されます。計画の最初の 2 つの段階からの抽出情報も含まれています。

- **demo\_cs.sav**。コンプレックス サンプル計画を使用して収集された調査情報を含む架空のデータ ファイルです。各ケースが別々の世帯単位に対応し、さまざまな人口統計情報および抽出情報が記録されています。
- **dmdata.sav**。これは、ダイレクト マーケティング企業の人口統計情報および購入情報を含む架空のデータです。dmdata2.sav には、テストメールを受け取った連絡先のサブセットの情報を含み、dmdata3.sav には、テストメールを受け取らなかった残りの連絡先に関する情報を含みます。
- **dietstudy.sav**。この架空のデータ ファイルには、“Stillman diet” の研究結果が含まれています。各ケースが別々の被験者に対応し、被験者のダイエット前後の体重（ポンド単位）と、トリグルセリドレベル（mg/100 ml 単位）が記録されています。
- **dvdplayer.sav**。新しい DVD プレーヤーの開発に関する架空のデータ ファイルです。プロトタイプを使用して、マーケティング チームはフォーカス グループ データを収集しました。各ケースが別々の調査対象ユーザーに対応し、ユーザーの人口統計情報と、プロトタイプに関する質問への回答が記録されています。
- **german\_credit.sav**。このデータ ファイルは、カリフォルニア大学アーバイン校の Repository of Machine Learning Databases にある “German credit” データセットから取ったものです。
- **grocery\_1month.sav**。この架空のデータ ファイルは、grocery\_coupons.sav データ ファイルの週ごとの購入を「ロールアップ」して、各ケースが別々の顧客に対応するようにしたものです。その結果、週ごとに変っていた変数の一部が表示されなくなり、買物の総額が、調査を行った 4 週間の買物額の合計になっています。
- **grocery\_coupons.sav**。顧客の購買習慣に関心を持っている食料雑貨店チェーンが収集した調査データを含む架空のデータ ファイルです。各顧客を 4 週間に渡って追跡し、各ケースが別々の顧客の週に対応しています。その週に食料品に費やした金額も含め、顧客がいつどこで買物をするかに関する情報が記録されています。
- **guttman.sav**。Bell は、予想される社会グループを示す表を作成しました。Guttman は、この表の一部を使用しました。この表では、社会相互作用、グループへの帰属感、メンバとの物理的な近接性、関係の形式化などを表す 5 個の変数が、理論上の 7 つの社会グループと交差しています。このグループには、観衆（例、フットボールの試合の観戦者）、視聴者（例、映画館または授業の参加者）、公衆（例、新聞やテレビの視聴者）、暴徒（観衆に似ているが、より強い相互作用がある）、第一次集団（親密な関係）、第二次集団（自発的な集団）、および近代コミュニティ（物理的により密接した近接性と特化されたサービスの必要性によるゆるい同盟関係）があります。

- **health\_funding.sav.** 医療用資金（人口 100 人あたりの金額）、罹患率（人口 10,000 人あたりの人数）、医療サービス機関への訪問率（人口 10,000 人あたりの人数）のデータを含む、架空のデータ ファイルです。各ケースが別々の都市を表します。
- **hivassay.sav.** HIV 感染を発見する迅速な分析方法を開発するための、ある製薬研究所の取り組みに関する架空のデータ ファイルです。分析の結果は、8 段階の濃さの赤で表現され、色が濃いほど感染の可能性が高くなります。研究所では 2,000 件の血液サンプルに関して試験を行い、その半数が HIV に感染しており、半分は感染していませんでした。
- **hourlywagedata.sav.** 管理職から現場担当まで、またさまざまな経験レベルの看護師の時給に関する架空のデータ ファイルです。
- **insurance\_claims.sav.** 不正請求の恐れがある、疑いを区別するためにモデルを作成する必要がある保険会社の仮説データ ファイルです。各ケースがそれぞれの請求を表します。
- **insure.sav.** 10 年満期の生命保険契約に対し、顧客が請求を行うかどうかを示す危険因子を調査している保険会社に関する架空のデータ ファイルです。データ ファイルの各ケースは、年齢と性別が一致する、請求を行った契約と行わなかった契約のペアを表します。
- **judges.sav.** 訓練を受けた審判（および 1 人のファン）が 300 件の体操の演技に対して付けた得点に関する架空のデータ ファイルです。各行が別々の演技を表し、審判たちは同じ演技を見ました。
- **kinship\_dat.sav.** Rosenberg と Kim は、15 種類の親族関係用語（祖父、祖母、父、母、叔父、叔母、兄弟、姉妹、いとこ、息子、娘、甥、姪、孫息子、孫娘）の分析を行いました。Rosenberg と Kim は、大学生の 4 つのグループ（女性 2 組、男性 2 組）に、類似性に基づいて上記の用語を並べ替えるよう依頼しました。2 つのグループ（女性 1 組、男性 1 組）には、1 回目と違う条件に基づいて、2 回目の並べ替えをするように頼みました。このようにして、合計で 6 つの「ソース」が取得できました。各ソースは、 $15 \times 15$  の近接行列に対応します。この近接行列のセルの数は、ソースの人数から、ソース内でオブジェクトを分割した回数を引いたものです。
- **kinship\_ini.sav.** このデータ ファイルには、kinship\_dat.sav の 3 次元の解の初期配置が含まれています。
- **kinship\_var.sav.** このデータ ファイルには、kinship\_dat.sav の解の次元の解釈に使用できる独立変数である性別、世代、および(ation), and 親等が含まれています。特に、解の空間をこれらの変数の線型結合に制限するために使用できます。
- **marketvalues.sav.** 1999 ~ 2000 年の間の、イリノイ州アルゴンキンの新興住宅地での住宅売上に関するデータ ファイルです。これらの売上は、公開レコードの問題となります。

- **nhis2000\_subset.sav**。National Health Interview Survey (NHIS) は、米国民を対象とした人口ベースの大規模な調査です。全国の代表的な世帯サンプルについて対面式で調査が行われます。各世帯のメンバーに関して、人口統計情報、健康に関する行動および状態の観測値が得られます。このデータ ファイルには、2000 年の調査から得られた情報のサブセットが含まれています。National Center for Health Statistics。National Health Interview Survey, 2000。一般使用データおよびドキュメント。[ftp://ftp.cdc.gov/pub/Health\\_Statistics/NCHS/Datasets/NHIS/2000/](ftp://ftp.cdc.gov/pub/Health_Statistics/NCHS/Datasets/NHIS/2000/)。2003 年にアクセス。
- **ozone.sav**。データには、残りの変数からオゾン濃度を予測するための、6 個の気象変数に対する 330 個の観測値が含まれています。それまでの研究者、が、他の研究者と共に、これらの変数間に非線型性を確認しています。この場合、標準的な回帰アプローチは使用できません。
- **pain\_medication.sav**。この架空のデータ ファイルには、慢性関節炎を治療する抗炎症薬の臨床試験の結果が含まれています。特に興味深いことは、薬の効果が出るまでの時間と、既存の薬剤との比較です。
- **patient\_los.sav**。この架空のデータ ファイルには、心筋梗塞 (MI、または「心臓発作」) の疑いで入院した患者の治療記録が含まれています。各ケースが別々の患者に対応し、入院に関連する多くの変数が記録されています。
- **patlos\_sample.sav**。この架空のデータ ファイルには、心筋梗塞 (MI、または「心臓発作」) の治療中に血栓溶解剤を投薬された患者のサンプルの治療記録が含まれています。各ケースが別々の患者に対応し、入院に関連する多くの変数が記録されています。
- **poll\_cs.sav**。市民の法案支持率を議会開会前に特定するための、世論調査員の取り組みに関する架空のデータ ファイルです。各ケースは登録有権者に対応しています。ケースごとに、有権者が居住している郡、町、区域が記録されています。
- **poll\_cs\_sample.sav**。この架空のデータ ファイルには、poll\_cs.sav の有権者のサンプルが含まれています。サンプルは、poll\_csplan 計画ファイルで指定されている計画に従って抽出され、このデータ ファイルには包含確率およびサンプル重み付けが記録されています。ただし、抽出計画では確率比例 (PPS) 法を使用するため、結合選択確率を含むファイル (poll\_jointprob.sav) もあります。サンプル抽出後、有権者の人口統計および法案に関する意見に対応する追加の変数が収集され、データ ファイルに追加されました。
- **property\_assess.sav**。限られたリソースで資産価値評価を最新に保つための、郡の評価担当者の取り組みに関する架空のデータ ファイルです。各ケースは、前年に郡内で売却された資産に対応します。データ ファイル内の各ケースでは、資産が存在する町、最後に訪問した評価担当者、その評価からの経過時間、当時行われた評価、および資産の売却価値が記録されています。

- **property\_assess\_cs.sav**。限られたリソースで資産価値評価を最新に保つための、州の評価担当者の取り組みに関する架空のデータ ファイルです。各ケースは州内の資産に対応します。データ ファイル内の各ケースでは、資産が存在する郡、町、および区域、最後の評価からの経過時間、および当時行われた評価が記録されています。
- **property\_assess\_cs\_sample.sav**。この架空のデータ ファイルには、property\_assess\_cs.sav の資産のサンプルが含まれています。サンプルは、property\_assess\_csplan 計画ファイルで指定されている計画に従って抽出され、このデータ ファイルには包含確率およびサンプル重み付けが記録されています。サンプル抽出後、現在の価値変数が収集され、データ ファイルに追加されました。
- **recidivism.sav**。管轄地域での累犯率を把握するための、政府の法執行機関の取り組みに関する架空のデータ ファイルです。各ケースは元犯罪者に対応し、人口統計情報、最初の犯罪の詳細、初犯から 2 年以内の場合は 2 回目の逮捕までの期間が記録されています。
- **recidivism\_cs\_sample.sav**。管轄地域での累犯率を把握するための、政府の法執行機関の取り組みに関する架空のデータ ファイルです。各ケースは 2003 年の 7 月に最初の逮捕から釈放された元犯罪者に対応し、人口統計情報、最初の犯罪の詳細、2006 年 7 月までの 2 回目の逮捕のデータが記録されています。犯罪者は recidivism\_cs.csplan で指定された抽出計画に従って抽出された部門から選択されます。調査では確率比例 (PPS) 法を採用したため、結合選択確率を保持したファイル (recidivism\_cs\_jointprob.sav) も用意されています。
- **rfm\_transactions.sav**。購入日、購入品目、各取引のマネタリー量など、購買取引データを含む架空のデータ ファイルです。
- **salesperformance.sav**。2 つの新しい販売トレーニング コースの評価に関する架空のデータ ファイルです。60 人の従業員が 3 つのグループに分けられ、全員が標準のトレーニングを受けます。さらに、グループ 2 は技術トレーニングを、グループ 3 は実践的なチュートリアルを受けます。トレーニング コースの最後に各従業員がテストを受け、得点が記録されました。データ ファイルの各ケースは別々の訓練生を表し、割り当てられたグループと、テストの得点が記録されています。
- **satisf.sav**。ある小売業者が 4 箇所の店舗で行った満足度調査に関する架空のデータ ファイルです。合計で 582 人の顧客を調査し、各ケースは 1 人の顧客からの回答を表します。
- **screws.sav**。このデータ ファイルには、ねじ、ボルト、ナット、鋸 (びょう) の特性に関する情報が含まれています。
- **shampoo\_ph.sav**。あるヘアケア製品工場での品質管理に関する架空のデータ ファイルです。定期的に、6 つの異なる製品が測定され、pH が記録されます。目標範囲は 4.5 ~ 5.5 です。



- **ships.sav.** 他の場所 で表示および分析される、波による貨物船への損害に関するデータセットです。件数は、船舶の種類、建造期間、およびサービス期間によって、ポワゾン率で発生するものとしてモデリングできます。因子のクロス分類によって形成されたテーブルの各セルのサービスマ月の集計によって、危険にさらされる確率の値が得られます。
- **site.sav.** 業務拡大に向けて新たな用地を選択するための、ある会社の取り組みに関する架空のデータ ファイルです。2 人のコンサルタントを雇って、用地を別々に評価させました。広範囲のレポートに加えて、各用地を「良い」、「普通」、「悪い」のいずれかで集計しました。
- **smokers.sav.** このデータ ファイルは、1998 年の National Household Survey of Drug Abuse から抜粋したものであり、アメリカの世帯の確率サンプルです。(<http://dx.doi.org/10.3886/ICPSR02934>) したがって、このデータ ファイルを分析する場合は、まず人口の傾向を反映させてデータを重み付けする必要があります。
- **stocks.sav** このデータ ファイルには、1 年あたりの在庫価格、量が含まれています。
- **stroke\_clean.sav.** この架空のデータ ファイルには、[データの準備] オプションの手続きを使用して整理した後の、医療データベースの状態が含まれています。
- **stroke\_invalid.sav.** この架空のデータ ファイルには、医療データベースの初期状態が含まれており、データ入力にいくつかエラーがあります。
- **stroke\_survival.** この架空のデータ ファイルは、虚血性脳卒中で数回の困難に直面した後リハビリ プログラムを終えた患者の生存時間に関するものです。脳卒中後、心筋梗塞の発生、虚血性脳卒中、または出血性脳卒中が注意され、イベントの時間が記録されます。脳卒中後に実施されたリハビリ プログラムの最後まで生存した患者のみが含まれるため、サンプルは左側が切り捨てられます。
- **stroke\_valid.sav.** この架空のデータ ファイルには、[データの検証] 手続きを使用して確認した後の、医療データベースの状態が含まれています。異常である可能性のあるケースが含まれています。
- **survey\_sample.sav.** このデータ ファイルには、人口統計データおよびさまざまな態度指標などの調査データが含まれています。これは「1998 NORC General Social Survey」の変数のサブセットに基づいていますが、いくつかのデータ値が変更され、追加の架空変数がデモの目的で追加されています。
- **telco.sav.** 顧客ベースにおける解約率を削減するための電気通信会社の取り組みに関する架空のデータ ファイルです。各ケースが別々の顧客に対応し、人口統計やサービス利用状況などのさまざまな情報が記録されています。
- **telco\_extra.sav.** このデータ ファイルは telco.sav データ ファイルに似ていますが、「期間」および対数変換された顧客支出の属性が削除され、標準化された対数変換顧客支出の変数に置き換えられています。

- **telco\_missing.sav.** このデータ ファイルは telco.sav データ ファイルのサブセットですが、一部の人口統計データ値が欠損値に置き換えられています。
- **testmarket.sav.** この架空のデータ ファイルは、新しいメニューを追加しようというファースト フード チェーンの計画に関連しています。新製品をプロモーションするためのキャンペーンには 3 つの候補があるため、新メニューはいくつかのランダムに選択した市場にある場所で紹介されます。場所ごとに別々のプロモーションを使用し、最初の 4 週間の新メニューの週間売上高が記録されます。各ケースが場所と週に対応します。
- **testmarket\_1month.sav.** この架空のデータ ファイルは、testmarket.sav データ ファイルの週ごとの売上を「ロールアップ」して、各ケースが別々の場所に対応するようにしたものです。その結果、週ごとに変わっていた変数の一部が表示されなくなり、売上高が、調査を行った 4 週間の売上高の合計になっています。
- **tree\_car.sav.** これは、人口統計および自動車購入価格のデータを含む架空のデータ ファイルです。
- **tree\_credit.sav.** これは、人口統計および銀行ローン履歴のデータを含む架空のデータ ファイルです。
- **tree\_missing\_data.sav.** これは、人口統計および銀行ローン履歴のデータと、多数の欠損値を含む架空のデータ ファイルです。
- **tree\_score\_car.sav.** これは、人口統計および自動車購入価格のデータを含む架空のデータ ファイルです。
- **tree\_textdata.sav.** 尺度および値ラベルを割り当てる前の、変数のデフォルトの状態を示すことを主な目的とする、変数を 2 つだけ含む単純なデータ ファイルです。
- **tv-survey.sav.** テレビ スタジオで実施された、ヒットした番組の放送期間を延長するかどうかを検討する調査に関する架空のデータ ファイルです。906 人の回答者に、さまざまな条件下でこの番組を視聴するかどうかを質問しました。各行は別々の回答者を表し、各列は別々の条件を表します。
- **ulcer\_recurrence.sav.** このファイルには、潰瘍の再発を防ぐための 2 つの治療の有効性を比較するように計画された調査の情報の一部が含まれています。これは区間調査の良い例であり、他の場所 で表示および分析されています。
- **ulcer\_recurrence\_recoded.sav.** このファイルでは、ulcer\_recurrence.sav の情報が、単に調査終了時のイベント確率ではなく調査の区間ごとのイベント確率をモデリングできるように再編成されています。これは他の場所 で表示および分析されています。
- **verd1985.sav.** このデータ ファイルは調査に関連しています。8 つの変数に対する 15 人の被験者の回答を記録しました。対象となる変数が 3 つのグループに分類されます。グループ 1 には「年齢」と「婚姻」、

グループ 2 には「ペット」と「新聞」、グループ 3 には「音楽」と「居住地域」がそれぞれ含まれます。「ペット」は多重名義として尺度化され、「年齢」は順序として尺度化されます。また、その他のすべての変数は単一名義として尺度化されます。

- **virus.sav**。自社のネットワーク上のウィルスの影響を特定するための、インターネット サービス プロバイダ (ISP) の取り組みに関する架空のデータ ファイルです。この ISP は、ネットワーク上の感染した E メール トラフィックの (およその) パーセンテージを、発見の瞬間から脅威が阻止されるまで追跡しました。
- **wheeze\_steubenville.sav**。これは、子供 に対する大気汚染の健康上の影響の長期調査から得られたサブセットです。このデータには、オハイオ州 スビューベンビル の 7 歳、8 歳、9 歳、10 歳の子供を対象に行った、喘鳴の状態の反復 2 値測定と、調査の初年に母親が喫煙していたかどうかの固定記録が含まれています。
- **workprog.sav**。体の不自由な人をより良い仕事に就かせようとする政府の事業プログラムに関する架空のデータ ファイルです。プログラムの参加者候補のサンプルが追跡されました。その中には、ランダムに選ばれてプログラムに登録された人と、そうでない人がいました。各ケースが別々のプログラム参加者を表します。
- **worldsales.sav** このデータ ファイルには、大陸および製品ごとの販売収益が含まれています。

# 注意事項

この情報は、世界各国で提供される製品およびサービス向けに作成されています。

IBMはこのドキュメントで説明する製品、サービス、機能は他の国では提供していない場合があります。現在お住まいの地域で利用可能な製品、サービス、および、情報については、お近くの IBM の担当者にお問い合わせください。IBM 製品、プログラム、またはサービスに対する参照は、IBM 製品、プログラム、またはサービスのみが使用することができることを説明したり意味するものではありません。IBM の知的所有権を侵害しない機能的に同等の製品、プログラム、またはサービスを代わりに使用することができます。ただし、IBM 以外の製品、プログラム、またはサービスの動作を評価および確認するのはユーザーの責任によるものです。

IBMは、本ドキュメントに記載されている内容に関し、特許または特許出願中の可能性があります。本ドキュメントの提供によって、これらの特許に関するいかなる権利も使用者に付与するものではありません。ライセンスのお問い合わせは、書面にて、下記住所に送ることができます。

IBM Director of Licensing, IBM Corporation, North Castle Drive,  
Armonk, NY 10504-1785, U. S. A.

2 バイト文字セット (DBCS) 情報についてのライセンスに関するお問い合わせは、お住まいの国の IBM Intellectual Property Department に連絡するか、書面にて下記宛先にお送りください。

神奈川県大和市下鶴間1623番14号 日本アイ・ビー・エム株式会社 法務・知的財産 知的財産権ライセンス渉外

**以下の条項は、イギリスまたはこのような条項が法律に反する他の国では適用されません。** International Business Machines は、明示的または黙示的に関わらず、第三者の権利の侵害しない、商品性または特定の目的に対する適合性の暗黙の保証を含むがこれに限定されない、いかなる保証なく、本出版物を「そのまま」提供します一部の州では、特定の取引の明示的または暗示的な保証の免責を許可していないため、この文が適用されない場合があります。

この情報には、技術的に不適切な記述や誤植を含む場合があります。情報については変更が定期的に行われます。これらの変更は本書の新版に追加されます。IBM は、本書に記載されている製品およびプログラムについて、事前の告知なくいつでも改善および変更を行う場合があります。

IBM 以外の Web サイトに対するこの情報内のすべての参照は、便宜上提供されているものであり、決してそれらの Web サイトを推奨するものではありません。これらの Web サイトの資料はこの IBM 製品の資料に含まれるものではなく、これらの Web サイトの使用はお客様の責任によるものとします。

IBM はお客様に対する一切の義務を負うことなく、自ら適切と考える方法で、情報を使用または配布することができるものとします。

本プログラムのライセンス取得者が (i) 別途作成されたプログラムと他のプログラム（本プログラムを含む）との間の情報交換および (ii) 交換された情報の相互利用を目的とした本プログラムに関する情報の所有を希望する場合、下記住所にお問い合わせください。

IBM Software Group, Attention:Licensing, 233 S. Wacker Dr., Chicago, IL 60606, USA.

上記のような情報は、該当する条項および条件に従い、有料で利用できるものとします。

本ドキュメントに記載されている許可されたプログラムおよびそのプログラムに使用できるすべてのライセンス認証された資料は、IBM Customer Agreement、IBM International Program License Agreement、および当社とかわした同等の契約の条件に基づき、IBM によって提供されます。

IBM 以外の製品に関する情報は、それらの製品の供給業者、公開済みの発表、または公開で使用できるソースから取得しています。IBM は、それらの製品のテストは行っておらず、IBM 以外の製品に関連する性能、互換性、またはその他の要求については確認できません。IBM 以外の製品の性能に関する質問は、それらの製品の供給業者に通知する必要があります。

この情報には、日常の業務処理で用いられるデータや報告書の例が含まれています。できる限り詳細に説明するため、例には、個人、企業、ブランド、製品などの名前が使用されています。これらの名称はすべて架空のものであり、実際の企業で使用される名称および住所とは一切関係ありません。

この情報をソフトコピーでご覧になっている場合は、写真やカラーのイラストが表示されない場合があります。

## 商標

IBM、IBM ロゴ、および [ibm.com](http://www.ibm.com)、SPSS は、世界の多くの国で登録された IBM Corporation の商標です。IBM の商標の現在のリストは、<http://www.ibm.com/legal/copytrade.shtml> を参照してください。

Adobe, the Adobe logo, PostScript, and the PostScript logo are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States, and/or other countries.

Intel、Intel のロゴ、Intel Inside、Intel Inside のロゴ、Intel Centrino、Intel Centrino のロゴ、Celeron、Intel Xeon、Intel SpeedStep、Itanium、および Pentium は、米国およびその他の国の Intel Corporation または関連会社の商標または登録商標です。

Java およびすべての Java ベースの商標およびロゴは、米国およびその他の国の Sun Microsystems, Inc. の商標です。

Linux は、米国およびその他の国における Linus Torvalds の登録商標です。

Microsoft、Windows、Windows NT、および Windows のロゴは、米国およびその他の国における Microsoft 社の商標です。

UNIX は、米国およびその他の国における The Open Group の登録商標です。

この製品は、WinWrap Basic (Copyright 1993-2007, Polar Engineering and Consulting, <http://www.winwrap.com>) を使用します。

その他の製品名およびサービス名等は、IBM または他の会社の商標です。

Adobe 製品のスクリーンショットは Adobe Systems Incorporated の許可を得て転載しています。

Microsoft 製品のスクリーンショットは Microsoft 社の許可を得て転載しています。



# 索引

- 度数分布表, 60
- 要約統計量
  - カテゴリ データ, 60
  - スケール変数, 64
- 条件式, 145
- 欠損値
  - 非数値型変数, 43
  - 数値型変数, 41
  - システム欠損, 40
- 商標, 177
- 図表
  - 棒, 62, 68
  - 図表オプション, 89
  - テンプレート, 84
  - 図表の作成, 68
  - 図表の編集, 76
  - ヒストグラム, 66
- 変数, 29
  - データの型, 35
  - ラベル, 34
- 度数
  - 度数のテーブル, 60
- 移動
  - ピボット テーブルの要素, 96
  - ビューア内の項目, 93
- 層
  - ピボット テーブルでの作成, 98
- Access (Microsoft), 14
- Excel (Microsoft)
  - 結果のエクスポート, 114
- Excel ファイル
  - 読み込み, 12
- HTML
  - 結果のエクスポート, 126
- PowerPoint (Microsoft)
  - 結果のエクスポート, 114
- syntax, 127
- Word (Microsoft)
  - 結果のエクスポート, 114
- カテゴリ データ, 60
  - 要約統計量, 60
- グラフ
  - 棒, 68
  - 図表オプション, 89
  - グラフの作成, 68
  - グラフの編集, 76
  - テンプレート, 84
  - 棒グラフ, 62
- ケース
  - 選択, 159
    - 並べ替え, 156, 159
  - ケースのサブセット
    - IF 条件が満たされるケース, 160
    - 条件式, 160
    - 選択, 159
    - 選択されなかったケースを分析から除外, 163
    - 選択されなかったケースの削除, 164
    - 日付および時刻に基づく, 162
    - 無作為抽出, 161
  - ケースの選択, 159
  - ケースの並べ替え, 156
- サンプル ファイル
  - 位置, 165
- 新しい変数の計算, 141
- システム欠損値, 40
- シンタックス ウィンドウ
  - 自動完了, 129
  - 分割点, 132
  - カラー コード化, 131
  - コマンドの実行, 128, 131
  - コマンドの編集, 129
  - コマンドの貼り付け, 127
- シンタックス ファイル
  - 開く, 131
- [シンタックス ヘルプ] ツール, 131
- シンタックスの貼り付け
  - ダイアログ ボックスから, 127
- スケール変数
  - 要約統計量, 64
- スケール データ, 60
- スプレッドシート ファイル
  - 変数名の読み込み, 12
  - 読み込み, 12
- データ エディタ
  - 複数のオープン データ ファイル, 55
- データセット
  - 名前の変更, 58
  - データセットの名前の変更, 58
- データ ファイル
  - 複数のオープン データ ファイル, 55
- テキスト インポート ウィザード, 21
- テキスト データ ファイル
  - 読み込み, 21

## 索引

- 数式での関数, 143
- データ入力, 29, 32
- 文字型データ
  - データの入力, 32
- 名義データ, 60
- 数値データ, 29
- 比率データ, 60
- 質的データ, 60
- 連続データ, 60
- 量的データ, 60
- 間隔データ, 60
- 順序データ, 60
- データ エディタ
  - 非数値データの入力, 32
  - 数値データの入力, 29
- データの入力
  - 非数値型, 32
  - 数値型, 29
- データの型
  - 変数, 35
- データベース ウィザード, 14
- データベース ファイル
  - 読み込み, 14
- ピボット テーブルの編集, 99
- ピボット テーブル内の行および列を隠す, 100
- ピボット テーブルの行と列の入れ替え, 96
- ビューア
  - 出力の移動, 93
  - 出力を隠す、または表示する, 93
- ファイルの分割処理, 157
- 複数のオープン データ ファイル, 55
  - 抑制, 59
- 変数ラベル
  - 作成, 34
- 値ラベル
  - 非数値型変数, 38
  - 数値型変数, 36
  - 定義, 36, 38
  - [ビューア] 内の表示の制御, 36, 38
- 変数ラベルを作成, 34
- 測定レベル, 60
- 日付変数と時刻変数, 147
- 日付と時刻ウィザード, 147
- 法律に関する注意事項, 176
- 変数の属性
  - 再使用, 44
- 結果のエクスポート
  - Excel, 114
  - HTML, 126
  - PowerPoint, 114
  - Word, 114
- 変数属性のコピー, 44
- 値の再割り当て, 135
- ヒストグラム, 66
- ピボット テーブル
  - 書式化, 99
  - 編集, 99
  - 層, 98
  - 行および列を隠す, 100
  - セルの書式, 100
  - セルのデータ型, 100
  - 行と列の入れ替え, 96
  - ピボット トレイ, 96
  - 定義へのアクセス, 95
  - 小数を隠す, 100