



---

## ハイライト

IBM SPSS Categories を使用して次のことを実行できます。

- 複雑なカテゴリ型データと数値型データ、および高次元のデータを視覚化して探索
  - 大規模な2次元表や多次元表に含まれている情報を理解
  - バイプロット、トリプロット、および知覚マップを使用してデータ内の関係を把握
- 

# IBM SPSS Categories

結果を予測して、カテゴリ型データの関係性を明らかに

IBM SPSS Categories では、予測分析、知覚マップ、選好尺度法、および次元縮小法を通じてデータを最大限に活用できます。SPSS Categories は、複雑なカテゴリ型データと数値型データ、および高次元データについて明確に洞察するために必要なツールを提供します。

例えば、SPSS Categories を使用すると、消費者が製品やブランドにどのような特徴を関連付けているのか、あるいは競合製品と比較して製品がどのように認識されているのかを明らかにできます。

SPSS Categories では、予測変数と目的変数の両方が数値型、順序型、または名義型である場合は回帰分析を行うことが可能です。また、スコア、度数、格付け、順位、または類似性を表した大きなテーブル内の行と列の関係性を見ることができ、データを視覚的に把握することができます。これにより、以下のことが可能になります。

- 従来の回帰分析、主成分分析、および正準相関分析に類する順序型と名義型データを扱い、理解する
- 数値型データ内の非標準残差や予測変数と目的変数の値との間の非線型関係を扱う。数値型とカテゴリ型のデータに対して Ridge 回帰、Lasso、Elastic Net、変数選択、モデル選択のオプションを使用
- バイプロットとトリプロットを使用して、相関分析のオブジェクト (ケース)間、カテゴリ間、および変数間の関係を表す
- 1 組または 2 組のオブジェクトの類似点を知覚マップ上の距離として表現



### 質的変数を量的変数に変換

SPSS Categories の高度な手続きを使用すると、カテゴリ型データに対してさまざまな統計手法を実行できます。

---

## 多変量データを簡単に分析および解釈

SPSS Categories の機能を使用すると、多変量データとその関係性を従来より高いレベルで簡単に分析および解釈できます。

- カテゴリ回帰を使用して、数値、(非)順序付きカテゴリ型予測変数の組み合わせに基づいて、名義型、順序型、または数値型の目的変数の値を予測可能
- 最適尺度法によって、Multiple R が最大化されるように変数を数量化
- 次元縮小手法によって、わかりやすい知覚マップとパイプロットを使用してデータ内の関係性を明確に把握
- 要約図により、2 つ以上の変数間の関係を把握できるように互いに類似する変数やカテゴリを表示

---

SPSS Categories の最適尺度法を使用して、カテゴリ型データに対して測定の単位とゼロの位置を設定できます。これにより、異なる測定尺度 (名義、順序、数値型) が混在する変数を分析するための、新しい統計機能が使用できます。

最適尺度法を使用して重回帰を実行すると、数値型、順序型、および名義型の予測変数と目的変数の値が混在している場合に、回帰を適用できます。最新バージョンの SPSS Categories には、モデルの選択と正則化手法が搭載されています。コレスポンデンス分析や多重コレスポンデンス分析を実行して、2 つ以上の名義変数の類似性を数値的に評価できます。また、コレスポンデンス分析を使用して、非負の項目があるどのような表でも分析ができます。

さらに、主成分分析を行い、重要な成分に基づいてデータを低減できます。変数、カテゴリ、およびオブジェクトのパイプロットとトリプロットでは、それらの関係を表示します。これらのオプションは、数値型データも扱えます。最適尺度法によって、順序と名義尺度の数量化に基づいた相関行列を表示します。または、変数を複数のセットに分割してから、非線型正準相関分析を使用してそれらのセット間の関係性を分析できます。

### 隠れた関係性を視覚的に表現

(市場区分、医学的診断、サブカルチャー、政党、生物学的種など、研究の対象となるカテゴリがどのようなものであっても、最適尺度法を使用すれば、2 次元テーブルの制約から解放され、変数間の関係をより大きな枠組みで捉えることができます。統計的なレポートだけでなく、データのマップも表示できます。

次元縮小法では、知覚マップとパイプロットを使用することで、テーブルでは把握できないデータの関係性を明らかにできます。

- 知覚マップは、高精度の要約グラフで、お互いに類似する変数やカテゴリを視覚的に表示し、高解像度の要約グラフです。2 つ以上のカテゴリ変数間の関係について独自の洞察が得られます。
- 一方、パイプロットとトリプロットを使用すると、ケース、変数、およびカテゴリ間の関係を把握できます。例えば、商品、顧客、および特徴等について検討することができます。

選好尺度法を使用すると、オブジェクト間の関係性をさらに視覚化できます。この手法のベースとなっている画期的な展開アルゴリズムは、順序型データに関する非数値的分析を実行して、意味のある結果を得ることができます。近接性尺度法を使用すると、オブジェクト間の類似性を分析して、それらのオブジェクトの特性を同じ分析に組み込むことができます。

## IBM SPSS Categories の使用方法

以下の手法を使用すると、データ分析をより有意義なものにできます。

- カテゴリー回帰 (CATREG) は、複数の数値、(非) 順序カテゴリー型予測変数から、名義型、順序型、または数値型の目的変数の値を予測します。最適尺度法を使用した回帰を使用して、例えば、仕事の満足度を職種、居住地域、および出張の頻度から予測できます。
- 最適尺度法は、Multiple R が最大化される方法で変数を数値化します。残差が非標準の場合や、予測変数が目的変数と線型で関係していない場合に、選好尺度を数値に適用できます。Ridge 回帰、Lasso、および Elastic Net という 3 つの新しい正則化手法は、パラメーター推測値を固定することで予測精度を高めます。自動変数選択によって、大規模のデータ・セット (変数がオブジェクトより多い) を分析可能になります。そして、数値尺度レベルを使用することで、数値型データに対しても Lasso または Elastic Net を使用して回帰で正則化できます。また、CATREG を使用して、数値型データとカテゴリー型データの両方に対して特定の一般化加法モデル (GAM) を適用できます。
- コレスポネンス分析 (CORRESPONDENCE) を使用すると、行と列の間の対応関係 (コレスポネンス) がセル内に表示されている 2 次元表の分析と、マップ上に行と列をポイントで表示できます。代表的なコレスポネンス表は、2 つの名義変数の同時度数が含まれたセルで構成されるクロス集計です。SPSS Categories では、これらの名義変数のカテゴリー間の関係性を視覚的に表現します。
- 多重コレスポネンス分析 (MULTIPLE CORRESPONDENCE) は、分析されるすべての編数が名義尺度 (順序のないカテゴリー) である多変数カテゴリー型データを分析します。この分析は 3 つ以上の変数を使用できるという点でコレスポネンス分析と異なります。例えば、多重コレスポネンス分析を使用して、好きなテレビ番組、年齢層、および性別の間

の関係を検証できます。SPSS Categories で作成された低次元マップを検討することで、どのグループが得ている番組に関連しているのか、また各番組の類似性も把握できます。

- カテゴリー主成分分析 (CATPCA) では、最適尺度法を使用して主成分分析を一般化することで、異なる測定尺度が混合した変数を扱えるようにします。カテゴリー主成分分析は多重コレスポネンス分析と似ていますが、変数ごとに分析の尺度 (名義、順序、数値型) を指定できるという点が異なります。例えば、自動車のさまざまなブランドや、価格、重量、燃費などの特性の間を表示できます。または、自動車を車種 (小型、中型、コンバーチブル、SUV など) 別に記述し、これらの分類を使用して CATPCA で自動車のポイントをグループ分けします。分類変数に大きな重み付けをすることで、自動車はそれらのクラス・ポイントは高度にクラスタ化されます。SPSS Categories では、オブジェクト、グループ、および変数間の複雑な関係を低次元マップに表示できるため、これらの関係を簡単に理解できます。
- 非線型正準相関分析 (OVERALS) は、最適尺度法を使用して正準相関分析を一般化することで、異なる測定尺度が混合した変数を扱えるようにします。この分析では、セット内の相関を除去した後に、同じグラフ内で複数の変数セットを比較できます。例えば、味覚の研究におけるスープなどの商品の特性を分析できます。審査員はセット内の変数に相当し、スープはケースに相当します。OVERALS は、相関を除去した後に評価内容を平均して、各種の特性を組み合わせることで、スープ間の関係を表示します。または、各審査員がスープの特徴 (変数) に基づいて判断している場合があります。この場合は、各審査員は 1 つのセットを構成され、OVERALS は相関を除去した後に特徴を平均し、審査員ごとに得点を結合します。複数の目的変数の値を予測変数から動じに予測する場合に、重回帰を一般化することもできます。

- 多次元尺度法 (PROXSCAL) は、類似要素または非類似要素 (近接要素) が含まれた 1 つまたは複数の行列の多次元尺度化を実行します。または、PROXSCAL に入力された、多変量データに含まれるケース間の距離を計算できます。PROXSCAL は、近接度をマップ内の距離として表示して、オブジェクト間の関係を空間的に理解できるようにします。多重近接行列の場合は、PROXSCAL はそれらの類似性と差異を分析します。例えば、PROXSCAL を使用して、さまざまな年齢層の消費者が好む各種コーラ味の間の類似性を表示できます。若年層は従来の味と新しい味の違いをより意識する一方で、成人はダイエット・コーラとノンダイエット・コーラの対比を重視するという結果が得られる可能性があります。
- 選好尺度法 (PREFSCAL) は、例えば消費者と商品といった 2 組のオブジェクト間の関係を視覚的に検証します。選好尺度法は、多次元展開を実行することで、これら 2 組のオブジェクト間の関係を 2 組のポイント間の距離として表現するマップを見つけ出します。例えば、ドライバー達が 26 種類のモデルの自動車を 10 種類の特性について 6 段階で評価した場合、どのモデルが類似し、どの人達がこれらのモデルを最も好むのかを、クラスターで表示されている含マップを得ることができます。このマップは、10 種類の特性を折衷し、これら 10 種類の異なる特性のプロットは、どのようにそれぞれがマップの次元を重み付けしているかを表します。

IBM SPSS Categories は、クライアント専用ソフトウェアですが、より高度なパフォーマンスと拡張性が必要な場合は、サーバー版をご利用ください。

## より高い価値を創出するコラボレーション

IBM SPSS Collaboration and Deployment Services と統合して使用することで、分析資産の共有および分散を効率的に行い、社内外のコンプライアンス要件を満たす方法で保護し、分析結果を公開して、より多くのビジネス・ユーザーが閲覧および利用することが可能です。

[ibm.com/spss/cds](https://ibm.com/spss/cds)

## を参照してください。機能

### 統計

#### CATREG

- 最適尺度法を使用したカテゴリ回帰分析
  - 各変数の変換の種類を指定。単調スプライン順序、非単調スプライン、順序型、名義型、多重名義型、数値型
  - 連続変数の離散化オプション。最適分布 (正規または一様) に基づいて事前に設定された個数のカテゴリにグループ分け。あるいは固定幅 (等間隔) で元のスコアをグループ分け。
  - 欠損値の処理法の指定。リスト毎に除外、あるいは最頻値で置き換え/カテゴリを追加。
  - 補足として扱われるオブジェクトを指定
  - 初期布置を計算方法を指定
  - 反復数の指定
  - 収束基準を指定
  - 以下の方法で結果をプロット (作図)
    - 変換プロット (カテゴリ指標に対する最適尺度のカテゴリ数量化)
    - 残差プロット

- 使用中のデータ・ファイルに変換済み変数、予測値、および残差を追加
- 結果の出力
  - 多重 R、R<sup>2</sup> 乗、および調整済み R<sup>2</sup> 乗
  - 標準化回帰係数、標準誤差、ゼロ次の相関、部分相関、偏相関、変換された予測変数の Pratt の相対重要度、変換前後の許容度、F 統計量
  - 記述統計量の表: 周辺度数、変換の種類、欠損値の個数、最頻値
  - 反復の履歴
  - 適合とモデルのパラメーターの表: 最適尺度水準に基づく自由度が含む分散分析表、調整済み R<sup>2</sup> のモデル要約表 (最適尺度、t 値、有意確率)。ゼロ次相関、部分相関、偏相関および変換前後の許容度と重要度を別の表で表示
  - 相関行列の固有値と変換後の予測変数の相関
  - 相関行列の固有値と元の予測変数の相関
  - カテゴリーの数量化
    - 離散化および変換されたデータを外部データ・ファイルに書き込み
- 3 種類の新しい正則化手法: Ridge 回帰、Lasso、Elastic Net
  - パラメーター推定値を固定することで予測精度を向上
  - 大規模データの分析 (変数がオブジェクトより多い)
  - 予測セットから自動変数選択を取得
  - 後で使用するために、正規化されたモデルと係数を新しいデータ・セットに書き込み
- 2 種類の新しいモデル選択と予測精度評価手法: .632 ブートストラップおよび相互検証 (CV)
  - .632(+) ブートストラップと相互検証 のオプションを使用して予測に最適なモデルを探索
  - ブートストラップを使用して係数の標準誤差のノンパラメトリック推定値を取得
- Systematic multiple start
  - 単調変換を伴う場合に、全体の最適ソリューションを見つけ出す
  - 再利用するために回帰係数を新しいデータ・セットに書き込み

## CORRESPONDENCE

- コレスポンデンス分析
  - データをケース・ファイルとして、または直接テーブル入力として入力
  - 解の次元数を指定
  - 2 つの距離尺度から選択: カイ 2 乗距離 (コレスポンデンス分析) またはユークリッド距離 (バイプロット分析)
  - 5 種類の標準化から選択: 行の平均値の除去、列の平均値の除去、行と列の平均値の除去、行の合計を均等化、列の合計の均等化
  - 5 種類の正規化: 対称的、主成分、行主成分、列主成分、ユーザー指定
  - 結果の出力
- コレスポンデンス表
  - 要約表: 特異値、イナーシャ、次元によって説明されるイナーシャの割合、次元によって説明されるイナーシャの累積、次元の最大数に関する信頼統計量、行プロファイル、列プロファイル
  - 行ポイントと列ポイントの概要: マス、得点、イナーシャ、次元のイナーシャに対するポイントの寄与率、ポイントのイナーシャに対する次元の寄与率
  - 行と列の信頼統計量: アクティブな行ポイントと列ポイントの標準偏差と相関

## MULTIPLE CORRESPONDENCE

- 多重コレスポンデンス分析 (SPSS Categories 13.0 より前のバージョンに含まれていた HOMALS の後継)
  - 変数の重みを指定
  - 連続変数の離散化または文字型変数を整数値に変換。乗算 (積算)、順位化 (ランク付け)、オプションである分布 (正規分布または一様分布) に基づいて事前に選択された個数のカテゴリーに値をグループ分け。あるいは事前に指定した感覚でカテゴリーにグループ分け。カテゴリー型データに値を割り当てる際には、ランク付け/グループ化オプションも使用できる。

- 欠損データの扱い方法を指定データ行列のセルで有効値を持たないものを除外、変数の最頻値として、あるいは別個のカテゴリとして欠損値を算出、あるいはリスト毎に除外。
  - 補助として使用されるオブジェクトと変数を指定 (補助オブジェクトにしか含まれないカテゴリのために、完全な出力を含む)
  - 解の次元数を指定
  - 初期布置の座標を含むファイルを指定。また変数を固定された (一定の) 初期布置に適用
  - 5 種類の正規化オプションから選択: 変数主成分 (変数間の連関を最適化)、オブジェクト主成分 (オブジェクト間の距離を最適化)、対称的 (オブジェクトと変数間の関係性を最適化)、独立、ユーザー指定 (変数主成分正規化とオブジェクト主成分正規化の値をユーザーが指定)
  - 反復回数の指定
  - 収束基準の指定
  - 結果の出力
    - モデルの要約
    - 反復統計量と反復履歴
    - 記述統計量 (度数、欠損値、および最頻値)
    - 変数ごとまたは次元ごとに判別測定
    - カテゴリ数量化 (重心座標)、マス、カテゴリのイナーシャ、次元のイナーシャに対するカテゴリの寄与率、カテゴリのイナーシャに対する次元の寄与率
    - 各次元の相関行列の固有値および変換後の変数間の相関
    - 相関行列の固有値および元の変数間の相関
    - オブジェクト・スコア
    - オブジェクトの寄与率: マス、イナーシャ、次元のイナーシャに対するオブジェクトの寄与率、オブジェクトのイナーシャに対する次元の寄与率
  - 結果をプロット (作図)
    - カテゴリのプロット: カテゴリ・ポイント、変換 (カテゴリ指標に対する最適カテゴリ数量化)、選択された変数の残差、選択された変数のカテゴリ・ポイントのジョイント・プロット
    - オブジェクト・スコアのプロット
    - 判別測定のプロット
    - 選択された変数のオブジェクトと重心バイプロットの作図
  - 変換された変数とオブジェクト・スコアを作業中のデータ・ファイルに書き込み
  - 離散化されたデータ、変換されたデータ、およびオブジェクト・スコアを外部データ・ファイルに書き込み
- ### CATPCA
- 最適尺度法を使用したカテゴリカル主成分分析
    - 各変数を分析するための最適尺度水準を指定。スプライン順序、スプライン名義、順序型、名義型、多重名義型、数値型
    - 変数の重みを指定
    - 連続変数を離散化あるいは文字型変数を整数値に変換。乗算 (積算)、順位化 (ランク付け)、オプションである分布 (正規分布あるいは一様分布) に基づいて事前に設定された個数のカテゴリに値をグループ分け。あるいは事前に指定した間隔でカテゴリにグループ分け。カテゴリ型データに値を割り当てる際には、ランク付け/グループ化オプションも使用可
    - 欠損データの扱い方を指定。データ行列のセルで有効値を持たないものを除外、変数の最頻値としてあるいは別個のカテゴリとして欠損値を算出、あるいはリストごとに除外
    - 結果の出力
      - モデルの要約
      - 反復の統計量と履歴統計量
      - 記述統計量 (度数、欠損値、および最頻値)
      - 変数と次元によって説明される分散

- 成分負荷
- 各次元のカテゴリータン量化とカテゴリータン座標 (ベクトル座標や重心座標)
- 相関行列の固有値および変換済み変数の相関
- 相関行列の固有値および元の変数の相関
- オブジェクト (成分) のスコア
- 結果のプロット (作図)
  - カテゴリータンプロット: カテゴリータン・ポイント、変換 (カテゴリータン指標に対する最適カテゴリータン量化)、選択された変数の残差、選択された変数のカテゴリータン・ポイントを同時プロット
- オブジェクト (成分) のスコアのプロット
- 成分負荷のプロット

### PROXSCAL

- 多次元尺度分析
  - 1 つ以上の近隣 (対称または非対称) の正行行列の読み込み
  - 重み、初期構成、固定座標、および独立変数の読み込み
  - 近接を順序型 (非行列) または数値型 (行列) として扱う。順序変換は結合観測を離散値または連続値として扱える
  - 3 種類の個人差モデルおよび識別モデルとともに多次元尺度法を指定
  - 固定座標または独立変数を指定して布置を制限。さらに、独立変数の変換 (数値型、名義型、順序型、スプレイン型) を指定

### PREFSCAL

- 2 組のオブジェクトに含まれる変数間の関係性を視覚的に検討して共通の定量的尺度を見つけ出す
  - 1 つ以上の近接の矩形行列の読み込み
  - 重み、初期構成、および固定座標の読み込み
  - 線型関数、順序関数、スムーズ順序関数、またはスプレイン関数を使用して近接要素を任意で変換
  - 同一、重み付けユークリッド・モデル、または一般化ユークリッド・モデルを使用した多次元展開法の指定
  - 固定された行・列座標を指定して布置を制限

### システム要件

要件はプラットフォームによって異なります。

[ibm.com/spss/requirements](http://ibm.com/spss/requirements)

### IBM ビジネス・アナリティクスについて

IBM Business Analytics ソフトウェアは、業績改善に取り組む意思決定者に対し、実践的な洞察を提供します。IBM は、ビジネス・インテリジェンス、予測分析と高度な分析、財務パフォーマンスと戦略の管理、ガバナンス、リスクおよびコンプライアンス (GRC)、そしてアナリティック・アプリケーションからなる包括的なポートフォリオを用意しています。

IBM ソフトウェアは、ビジネスの傾向やパターンあるいは異常の発見、仮定に基づくシナリオの比較、潜在的な脅威や機会の予測、重要なビジネス・リスクの特定および管理、さらには経営資源に関する計画、予算および予測を実現します。IBM の世界中のお客様は、この充実したアナリティクスを使うことで、業績への理解を深める一方、成果への予測を高め、目標への確かな道筋をつけることができます。



---

日本アイ・ビー・エム株式会社  
〒103-8510  
東京都中央区日本橋箱崎町 19 番 21 号

IBM のホーム・ページはこちらからご覧になれます。  
[ibm.com](http://ibm.com)

IBM、IBM ロゴおよび [ibm.com](http://ibm.com) は、世界の多くの国で登録された International Business Machines Corporation の商標です。他の製品名およびサービス名等は、それぞれ IBM または各社の商標である場合があります。このような商標は、その他の国においても登録商標またはコモン・ロー上の商標である可能性があります。現時点での IBM の商標リストについては、

[ibm.com/legal/copytrade.shtml](http://ibm.com/legal/copytrade.shtml) をご覧ください。

他の会社名、製品名およびサービス名等はそれぞれ各社の商標です。

© Copyright IBM Corporation 2012



リサイクル可能

**Business Analytics** software