



IBM SPSS Text Analytics for Surveys

調査の自由記述回答が量的分析で
容易に使用可能に

ハイライト

IBM SPSS Text Analytics for Surveys を使用して、以下を行います。

- 回答を一語一句読むことなく、主要テーマを特定
- ポジティブなコメントとネガティブなコメントを容易に区別
- 顧客、製品、従業員の満足度調査には、製品付属の 카테고리を使用
- カテゴリを素早く作成し、回答の 카테고리化の整合性を保持
- 分析のアウトソーシングを廃止または縮小することで費用を節約
- 進行中の調査や類似の調査で用いた カテゴリを再利用して、時間を短縮し、一貫性のある結果を出力
- アンケート結果を活用することで、顧客満足度を向上

アンケート回答者が使用する言葉は、回答者の感性や意見を知る上で大きなヒントとなります。そのため、調査においては自由記述式の質問をすることがあります。自由記述回答では、選択式回答に比べ、多岐にわたる詳細な情報が得られます。しかし、これまでは自由記述回答をカテゴリー化（コード化）するのに時間、費用、労力がかかるため、その価値が十分に生かされていたとは言えませんでした。

IBM SPSS Text Analytics for Surveys を利用することで、確実・迅速・詳細に自由記述回答を定量化し、他の調査データと合わせて解析することができます。また、SPSS Text Analytics for Surveys はこれまでテキスト分析を実行したことがない方でも、簡単にご利用いただけます。分析プロセスは自動化されていますが、手動で操作することも可能なため、より精度の高い結果を得るためのカテゴリー化プロセスを実現しました。

ビジネス、教育、行政分野での意思決定に役立てたり、教育機関での研究の一環として調査を実施するときなど、どのような状況においても、SPSS Text Analytics for Surveys を使用して、自由記述回答をカテゴリー化することができます。この使いやすいデスクトップ・ソフトウェアを使用すれば、適用分野に関係なくも定性分析と定量分析を結合できます。

テキスト回答を定量化する効果的な方法

SPSS Text Analytics for Surveys は、自由記述回答をカテゴリー化して、顧客、従業員、学生の満足度を定量的に評価するのに理想的なツールです。定量化した結果は、他のデータと統合して分析することができます。



自由記述回答から重要なキーワードを抽出し、それらを分類するために、SPSS Text Analytics for Surveysは、大規模な感性辞書と高度な自然言語処理技術を用いています。これらの技術では、テキストを句と文の集合体として解析し、文法上の構造から回答の意味を明らかにします。そのため、回答を一語一句読まずともキーワードを判別し、ポジティブやネガティブなコメントを区別することができます。

このソフトウェアを使用すると、カテゴリの作成および回答のカテゴリ化を自動的に実行し、構造化されていない調査データを定量データに変換できます。

IBM SPSS Text Analytics for Surveys はスタンドアロン製品ですが、IBM SPSS Statistics や他の IBM SPSS 製品と組み合わせて使用することもできます。IBM SPSS Custom Tables や、IBM SPSS Data Collection ファミリーのソフトウェアなどを組み合わせることができます。また、Microsoft® Excel とあわせて使用することもできます。

優れた操作性と管理

SPSS Text Analytics for Surveys の操作は簡単です。まず、テキスト回答をインポートし、次にマウスをクリックしてキーワード「抽出」を実行します。回答からキーワードを自動的に抽出し、このキーワードの属性である品詞・感性タイプを付与し、さらに語と語を関係性からまとめたパターンを出力します。

- キーワードとは、回答データから抽出された語や複合語です。
- タイプとは、[名詞]、[形容詞] などの品詞タイプ、および [良い - 効果が満足]、[悪い - 悲しい] などの意見や感性です。
- パターンとは、意見や感性と、それらが対象としている主題を“対”にしたもので、キーワード + [タイプ]、または [タイプ] + [タイプ] の組み合わせです。

「抽出」を実行すると、抽出されたキーワード、タイプ、またはパターン (語と語の関係性) が画面左下に、また対応する自由記述回答が画面右に表示されます。次にカテゴリを作成しますが、それにはいくつかの方法があります。

- 自動作成 – 言語学的手法に基づくカテゴリの作成、または、出現頻度に基づくカテゴリの作成を使用
- 手動作成 – キーワード、タイプ、パターン、回答 (原文) をドラッグ
- 上記の方法を組み合わせで作成

すべての回答を十分にカテゴリ化するには、手動による作業が必要になることもあります。例えば、個々の回答をカテゴリに振り分けてから、再度振り分け直したり、カテゴリを作成し、組み合わせから、新しいカテゴリ名を付ける、といった作業です。SPSS Text Analytics for Surveys では、ドラッグ & ドロップまたはポップアップ・メニューによって、手動作業をいつでも迅速かつ簡単に進めることができます。

既にカテゴリが存在する場合は、ルールを手動で作成する代わりに、ルールを正確に作成して自動化することができます。そのためには、抽出結果とブール型演算子を使用して、カテゴリの条件規則を作成します。そうすることで、より複雑な条件に基づいて回答をカテゴリ化したり、誤った回答を取り除いたりできます。

製品の視覚化パネル機能を使用して、表示された内容を参照しながら、カテゴリを手作業で改良することができます。例えば、棒グラフ、Web グラフ、Web テーブルを使用すると、共起する回答を含むカテゴリが一目でわかります。それを見ながら、特定のカテゴリを結合するか、それとも共有されている回答にふさわしいカテゴリを新たに作成するかを判断できます。

SPSS Text Analytics for Surveys を学習させる場合には、対象テキスト回答の一部をカテゴリ化してから、データ・セット全体をインポートし、抽出処理と分類処理を再実行します。進行中のアンケートの場合は、いったんカテゴリを作成してから、より新しいバージョンのデータをソフトウェアにインポートすると、信頼度が高く、一貫性のあるカテゴリ化処理を短時間で実行できます。

回答のカテゴリ化が終了すると、構造化されていなかった調査データが量的データに変換されています。このデータを 2 値型、またはカテゴリ型としてエクスポートし、SPSS Statistics、SPSS Data Collection、Excel などのソフトウェアで、他の量的データと合わせて解析できます。

SPSS Text Analytics for Surveys では、分析結果を他の分析者と容易に共有できます。抽出結果、カテゴリ、言語リソース (辞書) を含むプロジェクト・ファイルを組織内で共有できます。このほか、カテゴリを XML ファイルとしてインポートまたはエクスポートして、カテゴリを共有することもできます。こうした機能により、他の分析者は、新規プロジェクトでも既存のカテゴリを簡単に再利用できます。

言語リソース (辞書) の重要性

SPSS Text Analytics for Surveys は、特別なカスタマイズをしなくてもキーワードを抽出し、自由記述回答をカテゴリ化できます。ただし、製品の辞書機能をカスタマイズして、抽出結果の精度を高めることもできます。そうすることで、カテゴリ化プロセスをよりスムーズに進めることができます。次のような活用方法があります。

- 同じ製品ラインに属する製品名を指定する場合は、タイプ (キーワード) 辞書を修正します
- 業界や分野の専門用語を抽出したり、グループ化したりする場合は、その用語をタイプ (キーワード) 辞書に追加します
- 化学物質、遺伝子、医薬品などで同じものを異なる形式の名前で参照する場合は、類義語辞書を修正します
- 自社名などの「ノイズ」用語によって抽出結果やカテゴリ結果に生じる混乱を防止する場合は、不要語辞書を修正します

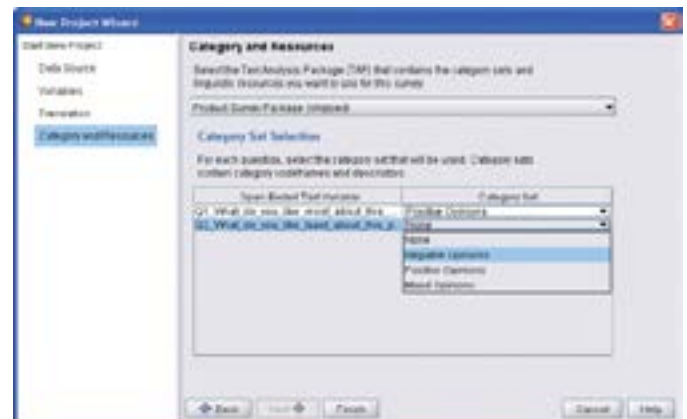
新機能

SPSS Text Analytics for Surveys の最新版では、階層カテゴリ構造 (サブカテゴリを含むカテゴリ) を構築して使用できます。これにより、キーワードの組み立てをより詳しく、より現実的に表現できます。

また、最新版では、日本語版用に 64 ビットのオペレーティング・システムをサポートしており、大規模なデータ・ファイルを使用する際のパフォーマンスが向上します。

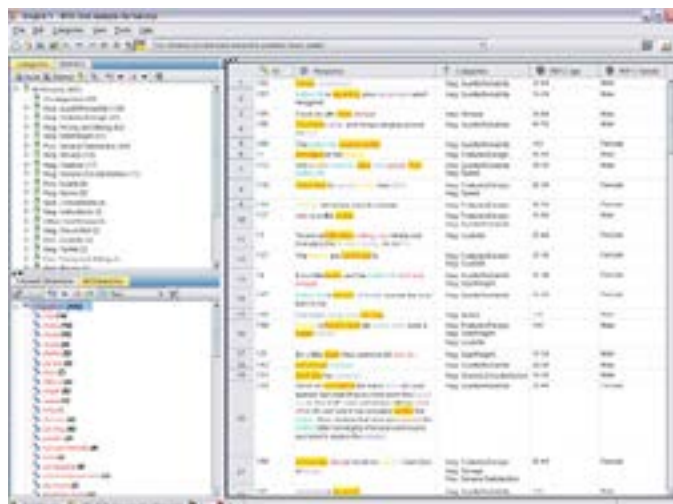
さらに、SPSS Text Analytics for Surveys 4 では、以下の新機能が追加されました。

- 新たな TAP が加わり、広告、銀行サービス、ブランドなどについての感想や意見を分析に使用できます*。
- セマンティック・ネットワークを (言語学的アルゴリズムに加えて) 使用して、カテゴリを自動的に作成し、回答をカテゴリ化します。
- IBM SPSS セマンティック・ネットワークは、カテゴリ作成の強化のために、数百の一般的なテーマをサポートします*。
- 銀行、顔文字、金融、スラングのライブラリーが追加されています。

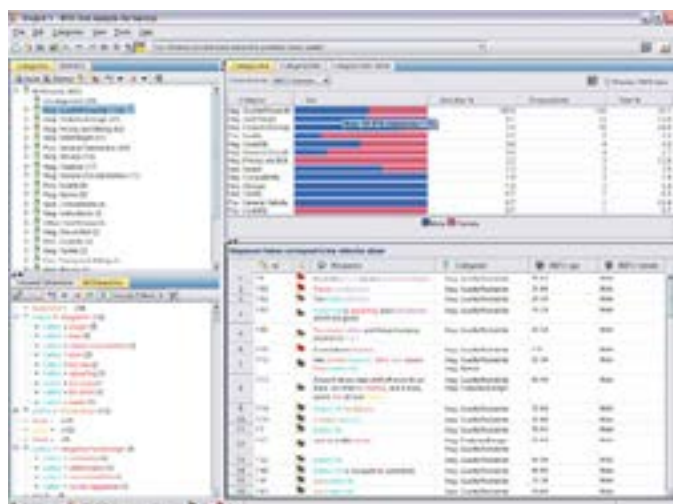


プロジェクト・ウィザードを使用すると、プロジェクトの作成を簡単に行えます。このウィザードは、データ・ソース、変数、および製品付属のカテゴリや言語リソースを含むテキスト分析パッケージ (TAP) を選択するためのプロセスをガイドします。

* 英語版のみの限定機能。



SPSS Text Analytics for Surveys では、自由記述回答を右パネルに表示し、抽出した概念をパネルの左下に表示しているため、プロセスを簡単に理解できます。自動的に設定された色から、抽出されたキーワードとそのタイプがわかります。



SPSS Text Analytics for Surveys の視覚化機能により、参照変数 (性別など) を棒グラフに重ね合わせて、カテゴリをプロファイル化できます。この例の棒グラフの青の「Male」部分をクリックすると、カテゴリに含まれる男性のすべての回答がデータ・パネルに表示されます。

IBM SPSS Text Analytics for Surveys の仕組み

信頼できる結果をより高速に取得する自動化機能

1. キー・コンセプトと意見の抽出

自由記述回答から重要なコンセプトを自動的に抽出します。キーワード、タイプ、パターン (語と語の関係性) のリストが作成されます。自動的に設定された色から、抽出されたキーワードとそのタイプがわかります。

2. カテゴリの作成と自由記述回答の分類

派生関係のキーワード、内包関係のキーワード、セマンティック・ネットワーク、頻度を使用して、カテゴリの生成、回答の分類を自動的に実行します。また、インターフェース上で、キーワード、タイプ、回答をドラッグすることにより、回答を手動でカテゴリ化します。

3. カテゴリの調整

視覚化機能を使用すると、回答を共有しているカテゴリがすぐにわかります。これにより、カテゴリを手動で容易に調整できます。Web グラフには、回答を共有しているカテゴリが表示され、特定のカテゴリを結合するか、それとも共有されている回答にふさわしいカテゴリを新たに作成するかを判断できます。

4. 検索結果の要約

プロジェクトとして分析をしている間は、いつでも要約した棒グラフを短時間で簡単に作成し、エクスポートできます。それを使用して、優先度の高いカテゴリを関係者に伝えます。結果をクリップボードや HTML ページにエクスポートして、アンケート結果を示すプレゼンテーション用のグラフを作成します。

5. 分析やグラフ作成のための結果のエクスポート

作成したカテゴリに問題がなければ、2 値型またはカテゴリ型として結果をエクスポートできます。表やグラフを作成する場合、このデータだけを使用することもできますし、他のアンケート・データと組み合わせて使用することもできます。結果を SPSS Statistics Base にエクスポートすると、クロス集計表など解析に必要なものをすべて作成できます。また、SPSS Statistics Base にエクスポートされた結果から、アンケート結果を伝えるためのグラフを作成できます。

「IBM は、インスピレーションにあふれた IBM SPSS Text Analytics for Surveys を開発および投入したことで、他のソフトウェア・ベンダーが失敗した領域で成功を収めました。ついに登場した強力な直観的なツールによって、膨大な定性データの分析に必要な時間が短縮されるだけでなく、的を絞って分析するという新たな手法への扉が開かれたといえます」

– Karl Buchholz
氏、Executive Vice President, Business Development
Data Specialists Inc.

機能 ユーザー・インターフェース

- プロジェクト・ウィザードの案内に従ってコード化処理を実行
- テキスト分析ウィンドウで、データの表示、カテゴリの作成、回答のカテゴリ化を実行
 - プロジェクト全体ビューで、すべての自由記述式の質問と回答を表示
 - 質問ビューで、特定の質問に対する回答を表示
- データ・パネルに、すべての回答データと参照変数データを表示
- 抽出結果パネルに、キーワード、タイプ、パターン (語と語の関係性) ごとに抽出結果を表示
- カテゴリ・パネルに、カテゴリ名とその内容を表示
- 条件規則エディターで、抽出結果とブール型演算子を使用して、条件規則を作成
- 新たなブール型演算子を作成するカテゴリ条件規則エディター (ワイルドカードや他のシンタックスの機能強化を使用)
- 視覚化パネルで、別棒グラフ、Web グラフ、Web テーブルを使用して、カテゴリをプロファイル化し、共起する回答を表示
- データ・パネルで、完了している回答とフォローアップが必要な回答へのフラグ付け
- 辞書エディター・ウィンドウで、言語リソース (辞書) を表示して、カスタマイズ

インポート可能なファイル形式:

- IBM SPSS Statistics (SAV)
- IBM SPSS Data Collection (MDD)
- Excel (XLS)
- Excel (XLSX) for Office 2007
- ODBC 準拠データベース

言語が英語以外の場合の翻訳

- 言語が英語以外の場合は、SDL の自動翻訳ソリューションを使用して英語に翻訳します。その場合、SDL のライセンスが必要です。

キー・コンセプトと意見の抽出

- プロジェクト・ウィザードにより簡単にセットアップ
- 言語リソース (辞書) を利用して、キーワード、タイプ (品詞・感性)、パターン (語と語の関係性) を自動抽出
- カテゴリの自動生成 (レビューと微調整が必要な場合、手動で作成することも可能)
- 言語学的アルゴリズムと新たなセマンティック・ネットワークを使用して、カテゴリを自動生成し、回答をカテゴリ化
- IBM SPSS のセマンティック・ネットワークは、カテゴリ作成の強化のために、数百の一般的トピックをサポート
- 抽出結果とブール型演算子を使用して、回答をカテゴリ化する条件規則を作成
- カテゴリ定義を変えずに、カテゴリに回答を「含む」/「含まない」の選択を「強制投入」/「強制排除」機能で実行
- 抽出結果の保存
- 他のプログラムで作成したカテゴリの再利用をサポート

カテゴリの作成

- 顧客/従業員/製品の満足度調査に同梱のカテゴリとリソース (TAP: テキスト解析パッケージ) を使用
- TAP は、広告、銀行サービス、ブランドについての感想や意見の分析に使用可能
- 新機能! 階層カテゴリ構造 (サブカテゴリを含むカテゴリ) を構築して使用可能
- 他のプログラムで作成したカテゴリを再利用
- 既存のカスタマー・コード・フレームをインポート (特定の Excel 形式が必要)
- 言語学的アルゴリズムとセマンティック・ネットワークを使用して、カテゴリを自動生成し、回答をカテゴリ化
- カテゴリの構築/分類の強化により、カテゴリを最初から作成するか、既存のカテゴリの定義を改善
- 手動でのレビューおよび微調整をサポート

- ・ 視覚化機能を使用して、共起カテゴリを表示
- ・ カテゴリ化した回答を関連性別にソート
- ・ 未抽出の語句をカテゴリ定義に「強制投入」し、その語句を含む回答を当該カテゴリに自動的に振り分け
- ・ カテゴリ・リストと視覚化の内容の一部を印刷
- ・ カテゴリを今後の調査で再利用

結果の微調整

- ・ インターフェース上で、キーワード、タイプ、回答をドラッグして回答を手動でカテゴリ化
- ・ カテゴリ化した回答を関連性別にソート
- ・ カテゴリの棒グラフ、Web グラフ、Web テーブルを使用して、共起する回答を表示
- ・ 完了している回答とフォローアップが必要な回答に「フラグ」付け
- ・ 参照変数を棒グラフに重ね合わせて、カテゴリをプロファイル化
- ・ カテゴリ・リストと視覚化の内容の一部を印刷

2 値型またはカテゴリとして結果をエクスポート (以下のファイル形式を使用)

- ・ IBM SPSS Statistics (SAV)
- ・ Excel (XLS)
- ・ Excel (XLSX) for Office 2007

リソースと結果の共有

- ・ 抽出結果、カテゴリ、言語リソース (辞書) を含むプロジェクト・ファイルを共有
- ・ 新規プロジェクトで使用するためにカテゴリとカテゴリ定義をエクスポート
- ・ カテゴリの要約棒グラフを作成し、エクスポート
- ・ プロジェクト・ファイルの一部または別のファイルとして、カスタム・ライブラリーを共有

辞書

- ・ 専門ユーザー向けに、固有表現 (電話番号、日付、金額など) のような拡張リソースを編集可能
- ・ キーワード辞書: キーワードの品詞タイプ、感性タイプの (再) 定義をサポート (カスタマイズ可能)
- ・ 類義語辞書: 1 つの代表語のもとに複数の類義語をグループ化するための類義語をグルーピング (カスタマイズ可能)
- ・ 不要語辞書: 抽出中に無視すべき「ノイズ」語を登録 (カスタマイズ可能)

ライブラリー

- ・ Survey ライブラリー: パターン・ルールやタイプに関係するリソースのほか、定義済みの類義語や不要語リストを含む (変更不可)
- ・ Project ライブラリー: 特定のプロジェクト向けの辞書の変更点が格納
- ・ Core ライブラリー: 人名、地名、商品名、組織名用のキーワード辞書を含む
- ・ Budget ライブラリー: 修飾句や形容詞句で表された語句の定義済みタイプを含む
- ・ Opinions ライブラリー: 修飾句や形容詞句用のキーワードをグループ化する定義済みタイプを 7 つ含む
- ・ English Variation ライブラリー: 正しくグループ化するために、特定の英語バリエーションに特化した類義語定義を含む
- ・ 銀行、顔文字、金融、スラングのライブラリーを追加

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。
- ・ 新機能! 日本語版は、64 ビット Windows をサポート

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19-21

IBM のホーム・ページはこちらからご覧になれます。
ibm.com

IBM、IBM ロゴ、ibm.com および SPSS は、世界の多くの国で登録された International Business Machines Corporation の商標です。*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 をご覧ください。

Microsoft、Windows、Windows NT および Windows ロゴは、Microsoft Corporation の米国およびその他の国における商標です。

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