

# IBM SPSS Text Analytics for Surveys

*Easily make your survey text responses usable in quantitative analysis*



---

## Highlights

*With IBM SPSS Text Analytics for Surveys you can:*

- Identify major themes without reading responses word for word
  - Distinguish between positive and negative comments easily
  - Use pre-built categories for customer, product and employee satisfaction surveys\*\*
  - Create categories more quickly and categorize responses more consistently
  - Save money by eliminating or reducing your reliance on outside services
  - Save time – and make results more consistent – by reusing categories in ongoing or similar surveys
  - Increase satisfaction with survey results
- 

The words people use to answer a question tell you a lot about what they think and feel. That's why you include open-ended questions in your surveys. They provide more varied and detailed information than closed-ended questions do. Traditionally, categorizing survey text responses has been time-consuming, tedious and expensive, often limiting your ability to unlock the full value of these responses.

With IBM® SPSS® Text Analytics for Surveys\*, you can quantify text responses for analysis along with other survey data – thoroughly, reliably and quickly. And IBM SPSS Text Analytics for Surveys is easy to use – even if you haven't performed text analysis before. It automates the process, while still allowing you to intervene manually to refine your results. You never lose control of the categorization process.

Use IBM SPSS Text Analytics for Surveys to categorize text responses whether you conduct surveys to support decision-making in business, education or government or as part of your academic research. No matter your field, you now have a way to combine your qualitative and quantitative analyses using one advanced but easy-to-use desktop software program.

## **An efficient way to make text responses count**

IBM SPSS Text Analytics for Surveys is an ideal tool for categorizing text responses and quantifying customer, employee or student satisfaction, so you can integrate the new results for analysis with the rest of your data.

Specifically designed for survey text, this product is based on our automated natural language processing (NLP) software technologies. These technologies analyze text as a set of phrases and sentences whose grammatical structure creates a context for the meaning of the response. Without having to read responses word for word, you can identify major themes and distinguish between positive and negative statements. Using this software, you can automate the creation of categories and categorization of responses to transform unstructured survey data into quantitative data.

\* IBM SPSS Text Analytics for Surveys was formerly called PASW® Text Analytics for Surveys.

\*\* An exclusive feature of the English-language version.



---

## Business Benefits:

- Gain deeper insight into your customers' (or employees' or constituents') thoughts, feelings and preferences
  - Make better informed decisions based on greater analytical value from your text responses
  - Save time by automating much of the text analytics processes
  - Save money by eliminating or reducing the need for outside coding services
- 

IBM SPSS Text Analytics for Surveys is a stand-alone program that works well with IBM® SPSS® Statistics\*\*\* and other IBM SPSS products that you might already use, including IBM® SPSS® Custom Tables\*\*\* and software from the IBM® SPSS® Data Collection family\*\*\*. You can also use it with Microsoft® Excel.

Two versions of IBM SPSS Text Analytics for Surveys are available for separate purchase: one for analyzing English, Dutch, French, German and Spanish survey text; and a second for analyzing Japanese text. The English-language version also includes an interface that supports third party translation options purchased separately.

## Easy to use and control

IBM SPSS Text Analytics for Surveys is simple to use. First, you import your text responses. The Project Wizard walks you through set up, so you can quickly import your text responses, extract concepts, and obtain initial results. Extraction results are displayed as lists of terms, types and patterns.

- Terms are single or compound words
- Types are groups of similar terms – positive or negative comments, for example
- Patterns are a combination of either terms and types or types and types – such as qualifiers and adjectives that reveal opinions about a particular subject

Onscreen, you see the extracted terms, types or patterns, as well as the text responses. You can create categories in several ways:

- Automatically – Use one of the pre-built sets of categories (for customer, employee or product satisfaction) by using linguistic algorithms, a semantic network\*\* or frequency
- Manually – Drag terms, types, patterns or responses into particular categories
- By using a combination of these methods

It's easy to switch between methods, and it puts you in full control of the categorization process.

Categorizing all responses satisfactorily will probably require manual intervention. You may need to assign and reassign individual responses to categories, as well as create, combine and rename categories. IBM SPSS Text Analytics for Surveys enables you to do this quickly and easily – and at any time – using either drag-and-drop or context menus.

\*\*An exclusive feature of the English-language version.

\*\*\*IBM® SPSS® Statistics and IBM® SPSS® Data Collection were formerly called PASW® Statistics and PASW® Data Collection.

---

*“With the introduction and inspired development of IBM SPSS Text Analytics for Surveys, IBM has succeeded where other software vendors have failed. Finally, there is a powerful, intuitive tool that reduces the time needed to analyze large amounts of qualitative data and opens the door to new methods of targeted analysis.”*

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

---

If your organization has pre-existing categories, you can recreate their rules with precision and automate their creation. You can do so by creating conditional rules, using extraction results and Boolean operators. This enables you to categorize responses based on more complex information or filter erroneous responses.

Use the product’s visualization capabilities to help manually refine categories. For example, use a bar chart, web graph or web table to quickly reveal which categories contain co-occurring responses. Then you can decide whether to combine certain categories or create new ones that better account for shared responses.

You can train IBM SPSS Text Analytics for Surveys by categorizing a subset of your text responses, and then importing the entire dataset and re-running the extraction and categorization. For ongoing surveys, you create categories once and then import the newer version of the data into the software for fast, reliable, and consistent categorization.

When you have finished categorizing your responses, your unstructured survey data has been transformed into quantitative data. You can export this data as dichotomies or categories and analyze it with other quantitative data in programs such as IBM® SPSS® Statistics, IBM SPSS Data Collection, or Excel®.

IBM SPSS Text Analytics for Surveys makes it easy to work with others. You can share project files – which include extracted results, categories and linguistic resources – across your organization. Additionally, you can share categories by importing and exporting them as an XML file. These features allow others to easily reuse category work in new projects.

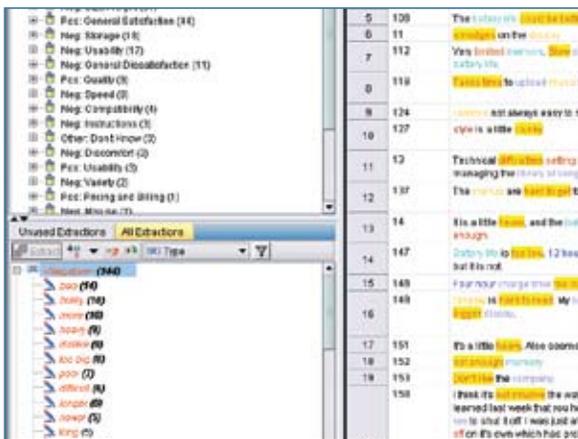
### **Linguistic technologies make the difference**

IBM SPSS Text Analytics for Surveys extracts concepts and categorizes text responses without any customization. You can, however, customize some of the product’s dictionaries to refine extraction results. This makes the categorization process operate more smoothly. For example:

- To indicate that certain product names belong in the same product line, modify the Type Dictionary
- To extract and group industry- and domain-specific terms, add these terms to the Type Dictionary
- To indicate that variant forms of the name of a chemical, gene or medicine, for example, refer to the same thing, modify the Substitution Dictionary
- To prevent a “noise” term, such as your organization’s name, from cluttering extraction and category results, modify the Exclude Dictionary



Creating a project is easy using the Project Wizard. The Wizard steps you through the process of selecting a data source, variables and Text Analysis Packages (TAPs) containing pre-built categories and linguistic resources.



IBM SPSS Text Analytics for Surveys makes the process easy to understand by displaying the text responses in the right-hand pane and extracted concepts in the pane at lower left. Automatic color-coding shows which terms have been extracted and identifies their type.

With the English-language version of IBM SPSS Text Analytics for Surveys, you can choose to create categories using a semantic network. The software uses this method in combination with others to optimize the quality of results. For instance, the semantic network will automatically understand that “apples” and “oranges” are types of fruit.

You can get better out-of-the-box results in the English-language version by using pre-built Text Analysis Packages (TAPs). Use TAPs to categorize responses in customer, product or employee satisfaction surveys.

### New features in IBM SPSS Text Analytics for Surveys 4

- Additional TAPs are available to help analyze thoughts and feelings about advertising, about banking services and about brands\*\*
- A new semantic network (along with linguistic algorithms) lets you automatically create categories and categorize responses
- An IBM SPSS semantic network supports hundreds of general themes for enhanced category building\*\*
- Additional libraries are available for banking, emoticons, finance and slang

### How IBM SPSS Text Analytics for Surveys works

Get reliable results faster with these automated features

#### 1. Extract key concepts and opinions

Extract key concepts automatically from responses to an open-ended question. The software creates a list of terms, types and patterns. Automatic color coding identifies which terms have been extracted and identifies their type.

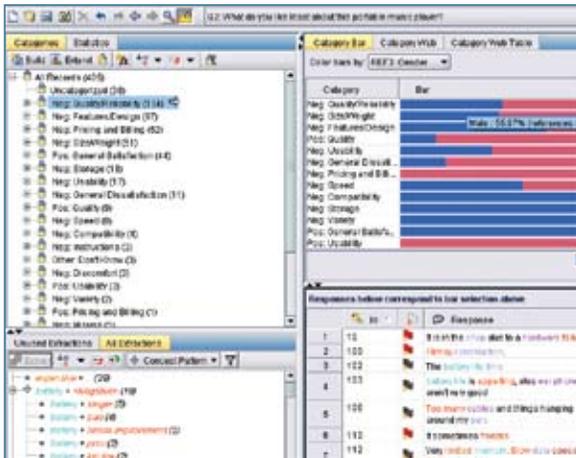
#### 2. Create categories and categorize text responses

Automatically create categories and categorize responses using term derivation, term inclusion, a semantic network or frequency. Also, categorize responses manually by dragging terms, types and responses within the interface.

#### 3. Refine categories

Visualization capabilities enable you to quickly see which categories share responses. This can help you to refine categories manually. A web graph showing which categories share responses enables the user to decide whether to combine certain categories or to create new ones that better account for shared responses.

\*\*An exclusive feature of the English-language version.



IBM SPSS Text Analytics for Surveys' visualization features enable you to overlay reference variables, such as gender, onto a bar chart to profile categories. When you click on the blue or "Male" portion of the bar in this example, all responses from males in the category are displayed in the Data pane.

#### 4. Summarize your findings

Quickly and easily create and export a summary bar chart at any point during your project. Use it to communicate to others your top categories. Results can be exported to clipboard or html pages to create graphs in your presentations that communicate survey findings.

#### 5. Export results for analysis and graphing

When you are satisfied with your categories, you can export results either as dichotomies or as categories. These can be used to create tables and graphs, either separately or in combination with other survey data. Export results to IBM SPSS Statistics Base to create crosstabs or whatever your analysis requires. Results can also be exported to IBM SPSS Statistics Base to create graphs that communicate survey findings.

### About IBM Business Analytics

IBM Business Analytics software delivers complete, consistent and accurate information that decision-makers trust to improve business performance. A comprehensive portfolio of business intelligence, predictive analytics, financial performance and strategy management, and analytic applications provides clear, immediate and actionable insights into current performance and the ability to predict future outcomes. Combined with rich industry solutions, proven practices and professional services, organizations of every size can drive the highest productivity, confidently automate decisions and deliver better results.

As part of this portfolio, IBM SPSS Predictive Analytics software helps organizations predict future events and proactively act upon that insight to drive better business outcomes. Commercial, government and academic customers worldwide rely on IBM SPSS technology as a competitive advantage in attracting, retaining and growing customers, while reducing fraud and mitigating risk. By incorporating IBM SPSS software into their daily operations, organizations become predictive enterprises – able to direct and automate decisions to meet business goals and achieve measurable competitive advantage.

## Features

### User interface

- Project Wizard guides business users through the coding process
- View data, create categories and categorize responses in the Text Analysis window:
  - View all open-ended questions and responses in the Project view
  - View responses to a single question in the Question view
- View all responses and reference variables in the Data pane
- View extraction results by term, type or pattern in the Extracted Results pane
- View categories and their contents in the Category pane
- Create conditional rules by using extraction results and Boolean operators in the
- Conditional Rule pane
- Category rules editor to create new Boolean rules (using wildcards and other syntactical enhancements)
- Profile categories and view response co-occurrence by using a category bar chart, web graph or web table in the Visualization pane
- Use flags in the Data pane to mark which responses are complete and which ones require follow-up
- View and customize linguistic resources in the Directory Editor window

### Import data from:

- IBM SPSS Statistics (SAV)
- IBM SPSS Data Collection (MDD)
- Excel (XLS)
- Excel (XLSX) for Office 2007
- ODBC-compliant databases

### Translate non-English languages

- Translate non-English languages into English with Language Weaver (requires a Language Weaver license).

### Extract key concepts and opinions

- A Project Wizard makes setup easier
- Extract terms, types and patterns automatically using linguistic resources
- Create categories automatically (or manually if desired for review and refinement)
- *New!* Use linguistic algorithms and a new semantic network to automatically create categories and categorize responses
- *New!* IBM SPSS Semantic Network supports hundreds of general themes for enhanced category building\*\*
- Create conditional rules to categorize responses by using extraction results and Boolean operators
- Allows the “force-in”/“force-out” of responses into/out of categories without changing the category definition
- Allows extraction results to be saved
- Supports the reuse of categories created in other programs

### Create categories

- Use pre-built categories and resources (TAPs) for customer/employee/product satisfaction surveys\*\*
- *New!* TAPs are available for analyzing thoughts and feelings about advertising, about banking services and about brands\*\*

- Re-use categories created in other programs
- Import pre-existing customer code frames (specific Excel format is required)
- Use linguistic algorithms and a semantic network to automatically create categories and categorize responses
- Category building/classification enhancements to create categories from scratch or to better define existing ones
- Supports manual review and refinement
- Shows response co-occurrence in categories using visualizations
- Sort your categorized responses by relevance
- “Force-in” an unextracted word or phrase into a category definition and automatically assign responses containing it to that category
- Print category lists and some visualizations
- Re-use categories in future surveys

### Refine results

- Categorize responses manually by dragging terms, types and responses within the interface
- Sort categories by relevance
- View response co-occurrence by using a category bar chart, web graph or web table
- Use “flags” to mark responses as completed or to follow-up
- Profile categories by overlaying reference variables onto bar charts
- Print category lists and some visualizations

\*\* An exclusive feature of the English-language version.

**Export results as dichotomies or categories (in the following file formats):**

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

**Share resources and results**

- Share project files that contain extracted results, categories and linguistic resources
- Export categories and category definition for use in new projects
- Create and export summary bar graphs of your categories
- Share custom libraries as part of a project file or as a separate file

**Dictionaries**

- More advanced resources such as non-linguistic entities (phone numbers, dates, amounts of money) are editable for advanced users
- Type Dictionary: Supports the grouping of similar terms (customizable)
- Substitution Dictionary: Contains synonyms to group similar terms under a single target name (customizable)
- Exclude Dictionary: Contains “noise” terms to be ignored during extraction (customizable)

**Libraries**

- Survey Library: Contains resources related to pattern rules and types, as well as a predefined list of synonyms and excluded terms (proprietary)
- Project Library: Stores dictionary changes for a particular project
- Core Library: Contains reserved Type Dictionaries for Person, Location, Product and Organization
- Budget Library: Contains a built-in type for words or phrases that represent qualifiers and adjectives
- Opinions Library: Contains seven built-in types that group terms for qualifiers and adjectives
- English Variation Library: Contains cases in which certain English-language variations require synonym definitions to group them properly
- *New!* Additional libraries are available for banking, emoticons, finance and slang\*\*

**System requirements**

- Operating system: Microsoft Windows 7 Ultimate, Professional and Home Premium (32-bit and 64-bit); Windows Vista Ultimate, Enterprise, Business, Home Premium and Home Basic (32- and 64-bit); Windows XP Professional, Service Pack 2 or later (32-bit and 64-bit). Japanese language version supports 32-bit Windows only.
- Hardware:
  - Processor: Intel Pentium-class; 3 GHz recommended
  - Monitor: 1024 x 768 (SVGA) resolution
  - Memory: 512 MB RAM minimum (1 GB for Japanese language version); 1 GB or more for large datasets (2 GB for Japanese language version)
  - Minimum free space: 800 MB minimum (1GB for Japanese language version); more recommended for larger datasets

*\*\* An exclusive feature in the English-language version.  
Features subject to change based on final product release.*



---

© Copyright IBM Corporation 2010

IBM Corporation  
Route 100  
Somers, NY 10589

US Government Users Restricted Rights - Use, duplication of disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Produced in the United States of America  
May 2010  
All Rights Reserved

IBM, the IBM logo, ibm.com, WebSphere, InfoSphere and Cognos are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. If these and other IBM trademarked terms are marked on their first occurrence in this information with a trademark symbol (® or ™), these symbols indicate U.S. registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at "Copyright and trademark information" at [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml).

SPSS is a trademark of SPSS, Inc., an IBM Company, registered in many jurisdictions worldwide.

Other company, product or service names may be trademarks or service marks of others.



Please Recycle

---