

## Introduction to IBM SPSS Text Analytics (v18)

### Course Outline

#### Unit 1 - Introduction to text mining

- Describe text mining and its relationship to data mining
- Explain CRISP-DM methodology as it applies to text mining
- Describe the steps in a text mining project

#### Unit 2 - An overview of text mining

- Describe the nodes that were specifically developed for text mining
- Complete a typical text mining modeling session

#### Unit 3 - Reading text data

- Reading text from multiple files
- Reading text from Web Feeds
- Viewing text from documents within Modeler

#### Unit 4 - Linguistic analysis and text mining

- Describe linguistic analysis
- Describe Templates and Libraries
- Describe the process of text extraction
- Describe Text Analysis Packages
- Describe categorization of terms and concepts

#### Unit 5 - Creating a text mining concept model

- Develop a text mining concept model
- Score model data
- Compare models based on using different Resource Templates
- Merge the results with a file containing the customer's demographics
- Analyze model results

#### Unit 6 - Reviewing types and concepts in the Interactive Workbench

- Use the Interactive Workbench
- Update the modeling node
- Review extracted concepts

#### Unit 7 - Editing linguistic resources

- Describe the resource template
- Review dictionaries
- Review libraries
- Manage libraries

#### Unit 8 - Fine tuning resources

- Review Advanced Resources
- Extracting non-linguistic entities
- Adding fuzzy grouping exceptions
- Forcing a word to take a particular Part of Speech
- Adding non-Linguistic entities

#### Unit 9 - Performing Text Link Analysis

- Use Text Link Analysis interactively
- Create categories from a pattern
- Use the visualization pane
- Create text link rules
- Use the Text Link Analysis node

#### Unit 10 - Clustering concepts

- Create Clusters
- Creating categories from cluster concepts
- Fine tuning Cluster Analysis settings

#### Unit 11 - Categorization techniques

- Describe approaches to categorization
- Use Frequency Based Categorization
- Use Text Analysis Packages to Categorize data

- Import pre-existing categories from a Microsoft Excel file
- Use Automated Categorization with Linguistic-based Techniques

#### Unit 12 - Creating categories

- Develop categorization strategy
- Fine tuning the categories
- Importing pre-existing categories
- Creating a Text Analysis Package
- Assess category overlap
- Using a Text Analysis Package to categorize a new set of data
- Using Linguistic Categorization techniques to Creating Categories

#### Unit 13 - Managing Linguistic Resources

- Use the Template Editor
- Share Libraries
- Save resource templates
- Share Templates
- Describe local and public libraries
- Backup Resources
- Publishing libraries

#### Unit 14 - Using text mining models

- Explore text mining models
- Develop a model with quantitative and qualitative data
- Score new data

#### Appendix A - The process of text mining

- Explain the steps that are involved in performing a text mining project