

Introduction to IBM SPSS Modeler and Data Mining (v18) SPVC

Course Outline

1: Introduction to data mining

- List two applications of data mining
- Explain the stages of the CRISP-DM process model
- Describe successful data-mining projects and the reasons why projects fail
- Describe the skills needed for data mining

2: Working with IBM SPSS Modeler

- Describe the MODELER user-interface
- Work with nodes
- Run a stream or a part of a stream
- Open and save a stream
- Use the online Help

3: Creating a data-mining project

- Explain the basic framework of a data-mining project
- Build a model
- Deploy a model

4: Collecting initial data

- Explain the concepts "data structure", "unit of analysis", "field storage" and "field measurement level"
- Import Microsoft Excel files
- Import IBM SPSS Statistics files
- Import text files
- Import from databases
- Export data to various formats

5: Understanding the data

- Audit the data
- Explain how to check for invalid values
- Take action for invalid values
- Explain how to define blanks

6: Setting the unit of analysis

- Set the unit of analysis by removing duplicate records
- Set the unit of analysis by aggregating records
- Set the unit of analysis by expanding a categorical field into a series of flag fields

7: Integrating data

- Integrate data by appending records from multiple datasets
- Integrate data by merging fields from multiple datasets
- Sample records

8: Deriving and reclassifying fields

- Use the Control Language for Expression Manipulation (CLEM)
- Derive new fields
- Reclassify field values

9: Identifying relationships

- Examine the relationship between two categorical fields
- Examine the relationship between a categorical field and a continuous field
- Examine the relationship between two continuous fields

10: Introduction to modeling

- List three modeling objectives
- Use a classification model
- Use a segmentation model