

# Data Mining Using SAS Enterprise Miner

This course covers the skills required to assemble analysis flow diagrams using the rich tool set of SAS Enterprise Miner for both pattern discovery (segmentation, association, and sequence analyses) and predictive modeling (decision tree, regression, and neural network models).

## Learn how to

- define a SAS Enterprise Miner 5 project and explore data graphically
- apply association and sequence discovery to transaction data
- modify data for better analysis results
- understand predictive modeling results and regression modeling output
- compare and explain complex models
- generate and use score code.

**Who should attend:** Data analysts, qualitative experts, and others who want an introduction to SAS Enterprise Miner

## Prerequisites

Before attending this course, you should be acquainted with Microsoft Windows and Windows-based software. In addition, you should have at least an introductory-level familiarity with basic statistics and regression modeling.

## Course Contents

### Introduction

- touring SAS Enterprise Miner 5.2
- placing SAS Enterprise Miner in the analysis workflow
- application examples and case studies
- pattern discovery and predictive modeling

### Accessing and Assaying Prepared Data

- SAS Enterprise Miner system architecture
- defining a SAS Enterprise Miner 5.2 project
- defining a data source
- validating source data

### Introduction to Pattern Discovery

- clustering and segmenting data
- using the Transform Variables node

- using the Clustering node for *k*-means cluster analysis
- applying association and sequence discovery
- using the Associations node in a consumer banking example
- quantifying the associations among items
- exploring sequences among items

### Introduction to Predictive Modeling with Decision Trees

- defining a modeling data source
- partitioning data for model development
- growing a decision tree with the Desktop Tree Application
- running the Decision Tree node
- using Decision Tree node options
- understanding predictive modeling results data

### Introduction to Predictive Modeling with Regressions

- comparing linear and logistic regression
- using the Regression node
- imputing missing values with the Impute node
- replacing data values with the Replacement node
- performing input selection
- understanding regression modeling output
- extending regression models with polynomial and interaction terms

### Introduction to Predictive Modeling with Neural Networks and Other Modeling Tools

- introduction to neural network (multilayer perceptron) models
- using the Neural Network node
- performing model selection with the Auto Neural node
- other SAS Enterprise Miner 5.2 modeling tools

### Model Assessment

- defining a prior vector
- effects of prior vectors on model development
- changing model selection criteria
- defining a profit matrix
- comparing models with the Model Comparison node
- introduction to model assessment statistics

### Model Implementation

- defining a Score data set
- scoring a data set with the Score node

- using a SAS Code node to export scored data
- generating and using score code

### Special Topics

- using the Variable Selection node
- combining models with the Ensemble node
- consolidating categorical inputs with the Decision Tree node
- explaining complex models with the Decision Tree node

### Case Studies

- segmenting bank customer transaction histories
- association analysis on Web services data
- creating a simple risk model from consumer loan data
- predicting university enrollment
- creating a response model for an insurance product

### **Software Addressed**

This course addresses the following software product(s): SAS Enterprise Miner.  
This course is appropriate for students who are using SAS Enterprise Miner Release 5.2.