

# ADVANCED DATA PREPARATION WITH PASW MODELER

■ Duration: 1 day    ■ Level: Intermediate

## Target Audience

This follow-up course to “Introduction to PASW Modeler and Data Mining” (also known as Introduction to Clementine and Data Mining) is designed for anyone who wishes to become familiar with the full range of techniques available in PASW Modeler for data and file manipulation.

## Prerequisites

General computer literacy. Some experience using PASW Modeler (formerly Clementine) including familiarity with the PASW environment, creating streams, reading data files, and doing simple data exploration and manipulation. Prior completion of “Introduction to PASW Modeler and Data Mining” is strongly encouraged. No statistical background is necessary.

## Overview

In this course, you’ll examine additional topics to aid in the preparation of data for a successful data mining project. You’ll learn how to sample records from files, handle missing data, modify fields and create new fields, and work with dates, strings and sequence data.

## Course Content

- Sampling data
- Working with dates
- Data transformations on continuous fields
- Working with string data
- Working with sequence data
- Writing out data files
- Efficiency in PASW Modeler
- Database joins with PASW Modeler
- Statistics transform node

**Products used: PASW Modeler Base**

