

INTRODUCTION TO MODELER AND DATA MINING

■ Duration: 3 days ■ Beginner

Target Audience

Anyone with little or no experience of using PASW Modeler or of data mining in general.

Prerequisites

General computer literacy. It would be helpful if you had an understanding of your organisation's data, as well as any of your organisation's business issues that are relevant to the use of data mining. No statistical background is necessary.

Overview

This course provides you with an overview of data mining and the fundamentals of using PASW Modeler. Using the CRISP-DM methodology, the principles and practice of data mining are illustrated. The course structure follows the stages of a typical data mining project, from reading data, to data exploration, data transformation, modeling, and effective interpretation of results. The course provides training in the basics of how to read, explore, and manipulate data with PASW Modeler, and then create and use successful models.

Course Content

- Introduction to data mining
- The CRISP-DM methodology
- Best practices for data mining
- The basics of using PASW Modeler
- Reading data files
- Working with dates
- Auditing and exploring data quality
- Searching for anomalous data and outliers
- Data manipulation
- Searching for relationships among fields
- Combining data files by appending and/or merging
- Restructuring data files with aggregate
- Sampling data
- Partitioning data for modeling
- Modeling techniques in PASW Modeler
- Automatic Modeling for Categorical Targets
- Evaluating and comparing model performance
- Deploying and using models
- Running PASW Statistics commands from PASW Modeler

Products used: PASW Modeler Base and PASW Classification

