

INTRODUCTION TO CLEMENTINE AND DATA MINING

■ Platform: Windows ■ Duration: 3 days ■ Beginner

TARGET AUDIENCE

Anyone with little or no experience of using SPSS Clementine or of data mining in general.

PREREQUISITES

General computer literacy. Attendees will also greatly benefit if they have an understanding of their organisation's data and knowledge of their 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 Clementine. The principles and practice of data mining are illustrated using the CRISP-DM methodology. You'll follow the stages of a typical data mining project, from reading data, to data exploration, data transformation, modelling, and effective interpretation of results. You'll also learn how to read, explore, and manipulate data while emphasizing some standard data preparation techniques, and then create and use successful models.

COURSE CONTENT

Following an overview of the main features of Clementine and an introduction to essential terminology, you will proceed logically through the following topics:

- Introduction to data mining
- The CRISP-DM methodology
- Best practices for data mining
- The basics of using Clementine
- 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 modelling
- Modelling techniques in Clementine
- Automatic modelling for binary outcomes
- Evaluating and comparing model performance
- Deploying and using models
- Running SPSS commands from Clementine