# Course : Data mining in practice

*Practical course - 3d - 21h00 - Ref. DMP*
*Price : 2010 € E.T.*

★★★★★ 5 / 5

Data mining involves discovering patterns, correspondences and motifs in a set of numerical or qualitative data. This activity is based on an algorithmic toolkit that will be presented in this course. The data mining approach will be illustrated on several projects, using R.

## Teaching objectives

**At the end of the training, the participant will be able to:**

- Understand the benefits of the data mining approach
- Translating and responding to a problem
- Learn about the main data mining methods
- Identify and use data mining tools
- Pose a data mining problem and find the appropriate method
- Ability to report results

## Intended audience

Researchers, data analysis project managers, data center, marketing or quality managers, database users and business managers, future data scientists.

## Prerequisites

Basic knowledge of statistics or knowledge equivalent to that acquired in the course "Descriptive statistics, introduction" (ref. UES).

## Course schedule

### 1 The data mining project

- The data scientist's problem: from data to information.
- Vocabulary and concepts.
- Descriptive exploration of the dataset.
- Metadata for monitoring data mining projects.
- Reminders about R software.

**Hands-on work**

Using R. Descriptive characterization, definition and entry of metadata for a dataset.

**PARTICIPANTS**

Researchers, data analysis project managers, data center, marketing or quality managers, database users and business managers, future data scientists.

**PREREQUISITES**

Basic knowledge of statistics or knowledge equivalent to that acquired in the course "Descriptive statistics, introduction" (ref. UES).

**TRAINER QUALIFICATIONS**

The experts leading the training are specialists in the covered subjects. They have been approved by our instructional teams for both their professional knowledge and their teaching ability, for each course they teach. They have at least five to ten years of experience in their field and hold (or have held) decision-making positions in companies.

**ASSESSMENT TERMS**

The trainer evaluates each participant's academic progress throughout the training using multiple choice, scenarios, hands-on work and more.
Participants also complete a placement test before and after the course to measure the skills they've developed.

## 2 Data mining techniques

- Classification-based method: identification of statistical groups of individuals.
- Association method: highlighting a cause and a consequence.
- Estimation method: addition of a number or frequency to a data set.
- The benefits of data mining for processing large volumes of data.
- Segmentation method: definition of criteria, extension of the classification method and k-means principle.
- Forecasting methods: the importance of timing and assumptions.

**Hands-on work**
Understand the different methods depending on the needs expressed.

## 3 Statistical tools

- Descriptive methods: correlation, classification, Kohonen networks, association rules.
- Predictive methods: regression, decision trees, neural networks, K-nearest neighbors.
- Implementation of k-means classification and CAH (Classification Ascendante Hiérarchique).
- Principle of supervised methods.

**Hands-on work**
Put the different methods into practice in R.

## 4 Data visualization

- Data visualization objectives.
- The different types of quantitative data representation.
- Design dashboards.

**Hands-on work**
Creating a dashboard with R using quantitative data. Represent quantitative and qualitative data with R.

## 5 Analysis of qualitative and textual data

- Specifics of the problem and alternatives (factorial correspondence analysis, contingency table).
- Introduction to instantiation, pattern, vector and heuristics.
- How to use a vector, indexing and scoring space.
- Different types of transformation and processing of a text document.

**Hands-on work**
Qualitative and textual data processing in R.

**TEACHING AIDS AND TECHNICAL RESOURCES**

• The main teaching aids and instructional methods used in the training are audiovisual aids, documentation and course material, hands-on application exercises and corrected exercises for practical training courses, case studies and coverage of real cases for training seminars.
• At the end of each course or seminar, ORSYS provides participants with a course evaluation questionnaire that is analysed by our instructional teams.
• A check-in sheet for each half-day of attendance is provided at the end of the training, along with a course completion certificate if the trainee attended the entire session.

**TERMS AND DEADLINES**

Registration must be completed 24 hours before the start of the training.

**ACCESSIBILITY FOR PEOPLE WITH DISABILITIES**

Do you need special accessibility accommodations? Contact Mrs. Fosse, Disability Manager, at psh-accueil@orsys.fr to review your request and its feasibility.

## Dates and locations

**REMOTE CLASS**
2026 : 8 June, 19 Oct.

**PARIS LA DÉFENSE**
2026 : 8 June, 19 Oct.