Datamining technique with R modeling and data representation

★★★★★

MR-42 **3 Days (21 Hours)**

Description

Data mining is based on the mastery of fundamental data exploration techniques: descriptive, predictive or exploratory statistics. This practical course will introduce you to methods such as regressions and PCA and teach you how to implement them with R software.

Who is this training for ?

For whom

Infocentre / Datamining / Marketing / Quality managers, users and business database managers.

Prerequisites

Aucune

Training objectives

- Understand the principle of statistical modeling Choose regression type based on data type

  Making predictions Create selections and rankings in large volumes of data to identify trends

Training program

Introduction à la modélisation

- Modeling: regression.
- Statistical modeling: reminders of statistical tests.
- Data analysis.
- Introduction to R software.
- Practical work Presentation of several modeling examples.
- Installation of R and the packages to be used.
- Applications on R, tests and interpretations on examples .

Analyse de régression linéaire

- Principle of linear regression.
- Simple regression, when the model has a single parameter for continuous data.
- Multiple regression, when there are more than 'a parameter.
- Other types of models for continuous data.
- Practical work Practical application in R.
- Case of simple regression and regression multiple.

Analyse de régression logistique

- Presentation of the different types of logistic regression.
- Binary logistic regression.
- Ordinal logistic regression.
- Multinomial logistic regression.
- Practical work Application on R with practical cases for cases of non-continuous data.
- Processing on data with two modalities, then with ordinal modalities, then nominal modalities.

Analyse en composantes

- Presentation of the different types of analyzes and selection.
- Principal Component Analysis (PCA).
- Multiple Correspondence Analysis (MCA).
- Hierarchical Classification on Principal Components (CHCP).
- Practical work The principal components make it possible to understand the covariance structure of the initial variables and/or to create a smaller number of variables to using this structure.
- Applications on R.

Analyse factorielle des données

- Understand the principle of factor analysis: summarize the structure of data into a fewer number of dimensions.
- Factor Correspondence Analysis (CFA).
- Analysis Multiple Factor Analysis (AFM).
- Factor Analysis for Mixed Data (AFDM).
- Practical work Factor analysis exercises on R.
- Identification underlying "factors" of dimensions associated with significant variability.