

Updated on 18/03/2024

Sign up

Statistical Modeling Training

2 days (14 hours)

Presentation

Statistical modeling is the process of applying statistical analysis to a data set. It involves using mathematical models and statistical assumptions to generate sample data and make predictions about the real world.

A **statistical model** is a collection of probability distributions representing all possible outcomes of an experiment. During our training course you will learn about different statistical models, such as linear models or ANOVA models.

Applying statistical modeling to raw data enables data scientists to approach data analysis strategically. It provides them with intuitive visualizations that help identify relationships between variables and predict their outcome.

Our Statistical Modeling Essentials course will introduce you to the basics of applied statistical analysis, statistical formulas and tests. You'll learn how to use R and Excel. You'll be able to design a fact-based analysis report, exploit statistical parameters and validate the accuracy of an estimate.

On completion of this course, you will be able to analyze statistical, descriptive and forecasting data.

Objectives

- Master the basics of applied statistical analysis
- Apply statistical formulas
- Using fundamental statistical tests
- Design a fact-based analysis report
- Understanding a data series using statistical parameters
- Discover R and Excel tools to implement the models studied
- Confirm the precision of an estimate with confidence intervals

- Be able to predict the following behaviors
- How to check suitability for a model

Target audience

- Database users and business managers
- Data Scientist
- Engineers
- Data Analysts
- Anyone interested in applied statistical analysis

Prerequisites

General knowledge of mathematics, statistical analysis and Excel

Program of our Statistical Modeling Essentials course

Introduction to statistical modeling

- What is statistical modeling?
- Mathematical modeling
- Define descriptive statistics
- Sampling methods
- Statistical inference

Parametric methods

- Simple/multiple linear regression
 - Model
 - Least squares
 - Estimates
 - Confidence intervals
 - Variable selection
- Analysis of variance
- Analysis of covariance
- Comparison with Anova
- Comparison with linear regression
- Generalized linear model
 - Poisson regression
 - Polytomous model
- Bayesian modeling
- Time series analysis

Non-parametric methods

- Spline regression
- Estimated by local averages
- Core estimation
- Local polynomial regression

Polytomous quantitative variables

- Ordered variables
- Unordered variables
- Nested choice variables
- Formalization and examples
- Variable applications

Binary quantitative variables

- Parameter estimation and interpretation
- Validate results
- Examples and formalization
- Case studies

Data cleansing

- Identify any superfluous codes
- Identify large quantities of missing data
- Deleting variables
- Univariate analysis
- Bivariate analysis

Confidence intervals

- Estimation and confidence intervals
- Confidence interval statistical laws
- Validate the accuracy of an estimate

Statistical parameters

- Use statistical parameters
- Understanding a series of data
- Statistical aspects of modeling
 - Model validation
 - Coefficient tests
 - Study of residues and influential points
- Model-fit verification

Modeling with Python and Excel

- Using regression methods
- Analyzing variance in Python
- Tackle a concrete question using a modeling approach
- Modeling with Python

Companies concerned

This training course is aimed at both individuals and companies, large or small, wishing to train their teams in a new advanced computer technology, or to acquire specific business knowledge or modern methods.

Positioning on entry to training

Positioning at the start of training complies with Qualiopi quality criteria. As soon as registration is finalized, the learner receives a self-assessment questionnaire which enables us to assess his or her estimated level of proficiency in different types of technology, as well as his or her expectations and personal objectives for the training to come, within the limits imposed by the selected format. This questionnaire also enables us to anticipate any connection or security difficulties within the company (intra-company or virtual classroom) which could be problematic for the follow-up and smooth running of the training session.

Teaching methods

Practical course: 60% Practical, 40% Theory. Training material distributed in digital format to all participants.

Organization

The course alternates theoretical input from the trainer, supported by examples, with brainstorming sessions and group work.

Validation

At the end of the session, a multiple-choice questionnaire verifies the correct acquisition of skills.

Sanction

A certificate will be issued to each trainee who completes the course.