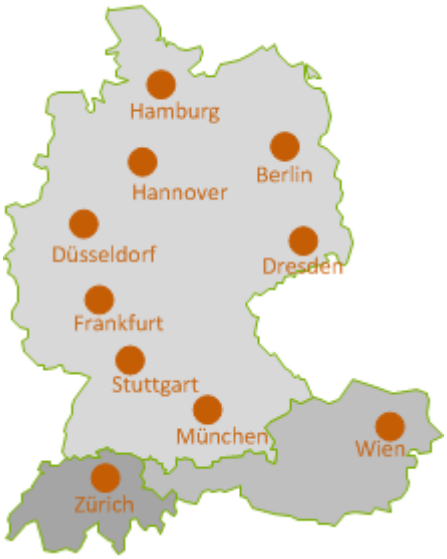
Course Catalog

Table Of Contents

a. Locations



Our trainings take place at various locations in the German-speaking countries.

Public trainings:

You can enroll for public trainings at our training centers across Germany like in Berlin, Dresden, Hamburg, München / Munich, Düsseldorf, Frankfurt, and Stuttgart. Not all public trainings will be organized in all cities but you can still book a particular training for your team in one of our training and conference centers.

In Austria you can attend seminars and trainings in Wien / Vienna while we offer training dates in Switzerland in Zürich / Zurich.

On-site trainings:

We have mobile and flexible trainers / lecturers who like to visit you and your team for an on-site training or a training in a conference center or hotel near you.

USA

Chicago	Tel: Fax:
Miami	Tel:+1.305.395.7962 Fax:+1.305.395.7964
New York	Tel:+1.212.380.1181 Fax:+1.305.395.7964

1. R Foundation

A. R

(i) Bayesian Statistics using R

Overview

Course ID	1000031
Language	en
Duration	3 D ys
Delivery mode	Classroom
Course Type	
Target Group	Data Analysts
Prerequisites	Basics in R and Statistics
Method	Presentation with examples and hands-on labs.
Course level	Beginning

Course Dates

Chicago	Miami	New York
2,050.00 USD	1,950.00 USD	2,050.00 USD
07-09 Sep 02-04 Nov 28-30 Dec	24-26 Aug 19-21 Oct 14-16 Dec	31 Aug - 02 Sep 26-28 Oct 21-23 Dec

Prices plus local taxes.

Course Description

Bayesian statistics is a subset of the field of statistics in which the evidence about the true state of the world is expressed in terms of degrees of belief or, more specifically, Bayesian probabilities. The general set of statistical techniques can be divided into a number of activities, many of which have special Bayesian versions. This training shows how to use Bayesian and probabilistic thinking to analyze data, to make predictions, and to fit models. In a first part, you will see the differences between the frequentist and probabilistic approach and see how you can use R for Bayesian statistics. In a second part, you will see how you can apply Bayesian inference as a method of statistical inference in which Bayes' rule is used to update the probability for a hypothesis as evidence is acquired. A third part focuses on the formulation of statistical models where the unique feature of Bayesian statistics consists in requiring the specification of prior distributions for any unknown parameters. The training closes with a part on machine learning / Data Mining for classification. The examples and hands-on labs are carried out using both R and OpenBUGS. OpenBUGS is a software for the Bayesian analysis of complex statistical models using Markov Chain Monte Carlo (MCMC) methods.

Course Outline

A. Bayesian Statistics

(0.5 Days) Introduction: Quantifying Uncertainty Using Probabilities, Models and Prior Probabilities, Likelihoods and Posterior Probabilities, Bayesian Sequential Analysis - Review of Probability: Events and Sample Spaces, Unions - Intersections, Complements - Marginal and Conditional Probabilities - Bayes' Rule - Addition and Multiplication Rules

B. One-Parameter Models

(0.5 Days) Bayesian Models - Prior Probability and Prior Distributions - The Posterior Distribution - Conjugate Priors - Inference for a Population Proportion: Frequentist Approach, Bayesian Inference, Bayesian Point Estimates - R for Bayesian Analysis - Inference Using Nonconjugate Priors on Mean and Variance - Noninformative Priors

C. Multiparameter Models

(0.25 Days) Informative Priors for Mean and Variance - Conjugate Joint Prior Density for Mean and Variance

D. Model Fit using Markov Chain Monte Carlo (MCMC)

(0.5 Days) Sampling-Based Methods - Markov Chain Monte Carlo (MCMC) Methods - Bayesian Models - Hierarchical Models: Fitting Bayesian Hierarchical Models, Estimation Based on Hierarchical Models - Software OpenBUGS

E. Regression and Hierarchical Regression Models

(0.5 Days) Review of Linear Regression - Introduction to Bayesian Simple Linear Regression - Generalized Linear Models - Hierarchical Normal Linear Models - Model Comparison, Model Checking, and Hypothesis Testing - Bayes Factors for Model Comparison and Hypothesis Testing - Bayes Factors and Bayesian Hypothesis Testing

F. Data Mining and Classification in Bayesian Statistics

(0.75 Days) Statistics for Machine Learning - Learning as Inference - Principal Components Analysis - Naive Bayes - Nearest Neighbour Classification - Gaussian Processes

(ii) Exploratory Data Analysis using R

Overview

Course ID	1000032
Language	en
Duration	3 D ys
Delivery mode	Classroom
Course Type	
Target Group	Data Analysts
Prerequisites	Basics in R and Statistics
Method	Presentation with examples and hands-on labs.
Course level	Beginning

Course Dates

Chicago	Miami	New York
2,050.00 USD	1,950.00 USD	2,050.00 USD
07-09 Sep 02-04 Nov 28-30 Dec	24-26 Aug 19-21 Oct 14-16 Dec	31 Aug - 02 Sep 26-28 Oct 21-23 Dec

Prices plus local taxes.

Course Description

Exploratory Data Analysis (EDA) is a statistical approach to analyzing data sets to summarize their main characteristics. This training primarily focuses on four main techniques of EDA: Principal Component Analysis (PCA) for quantitative variables, Correspondence Analysis (CA) and Multiple Correspondence Analysis (MCA) for categorical variables and finally (hierarchical and partitioning) clustering methods. As an umbrealla technique, this training also shows Factor Analysis (FA) and Multiple Factor Analysis (MFA). For the hands-on labs and practical examples the participants will use R and esp. FactoMineR - a special R package for the exploratory data analysis.

Course Outline

A. Principal Component Analysis (PCA)

(0.75 Days) Objectives of PCA and Introduction to PCA - Studying Individuals: The Cloud of Individuals, Fitting the Cloud of Individuals - Variables: The Cloud of Variables, Fitting the Cloud of Variables - Relationships - Interpreting the Data - Testing the Significance of the Components - Implementation with R and FactoMineR

B. Correspondence Analysis (CA)

(0.25 Days) Objectives and the Independence Model - Fitting the Clouds: Row and Column Profiles - Interpreting the Data - Implementation with R and FactoMineR

C. Multiple Correspondence Analysis (MCA)

(0.25 Days) Objectives: Studying Individuals, Variables, and Categories - Defining Distances between Individuals and Distances between Categories - CA on the Indicator Matrix: Relationship between MCA and CA, The Cloud of Individuals, Variables, and Categories - Implementation with R and FactoMineR

D. Clustering

(0.75 Days) Concepts of Similarity and Distance: Similarity between Individuals and Groups - Ward's Method - Partitioning and Hierarchical Clustering - Direct Search for Partitions: K-means Algorithm - Clustering and Principal Component Methods - Implementation with R and FactoMineR

E. Multiple Factor Analysis (MFA)

(0.75 Days) Factorial Analysis of Mixed Data - Weighting Groups of Variables - Comparing Groups of Variables and Indscal Model - Qualitative and Mixed Data - Multiple Factor Analysis and Procrustes Analysis - Hierarchical Multiple Factor Analysis - Implementation with R and FactoMineR

(iii) Geostatistics and the Analyis of Spatial Data

Overview

Course ID	1000029
Language	en
Duration	2 D ys
Delivery mode	Classroom
Course Type	
Target Group	Data Analysts
Prerequisites	Basics in R and Statistics
Method	Presentation with examples and hands-on labs.
Course level	Beginning

Course Dates

Chicago	Miami	New York
1,750.00 USD	1,700.00 USD	1,750.00 USD
03-04 Sep 29-30 Oct 24-25 Dec	20-21 Aug 15-16 Oct 10-11 Dec	27-28 Aug 22-23 Oct 17-18 Dec

Prices plus local taxes.

Course Description

Geostatistics is a branch of statistics focusing on spatial or spatiotemporal datasets. Such spatial and spatio-temporal data are everywhere. Beyond creating and viewing maps, spatial data analysis is concerned with questions not directly answered by looking at the data themselves. These questions refer to hypothetical processes that generate the observed data. Statistical inference for such spatial processes can be done using the statistical programming language and environment R. This training show beginners in geostatistics and participants working in the various domains of geoscience how to use R for their geostatistical analyses, visualization and plotting, model fitting, and inferential statistics. The first part of the training covers diverse techniques for handling spatial data in R, functions for import and exports of spatial data and creating diagrams and maps. The second part introduces time as a second dimension for spatio-temporal data. The third part shows you how to analyze spatial data and presents methods and functions for the anylsis of spatial point patterns and spatial point processes, interpolation, spatial prediction, the analysis of correlation, the variogram analysis as well as kriging, filtering or smoothing. This part also deals with modeling areal data and the analysis of spatial autocorrelation or fitting models.

Course Outline

A. Handling Spatial Data in R

(0.5 Days) Classes for Spatial Data in R - Visualising Spatial Data: The Traditional Plot System, Trellis/Lattice Plots, Interactive Plots, Colour Palettes and Class Intervals - Spatial Data Import and Export: Coordinate Reference Systems, Vector File Formats, Raster File Formats, Google Earth, Google Maps, Geographical Resources Analysis Support System (GRASS) - Map Overlay or Spatial Join - R-Packages: rdgal, spplot and ggplot, latticeExtra, raster, rgeos

B. Spatio-Temporal Data

(0.25 Days) Types of Spatio-Temporal Data - Handling Time Series Data - Selection, Addition, and Replacement of Attributes - Overlay and Aggregation - Visualization: Multi-panel Plots, Space-Time Plots, Animated Plots, Time Series Plots - R-Packages: xts, spacetime

C. Analyzing Spatial Data

(0.5 Days) Preliminary Analysis of a Point Pattern: G Function (Distance to the Nearest Event), F Function (Distance from a Point to the Nearest Event) - Statistical Analysis of Spatial Point Processes: Homogeneous and Inhomogeneous Poisson Processes, Estimation of the Intensity, Likelihood of an Inhomogeneous Poisson Process - Applications in Spatial Epidemiology: Case–Control Studies, Binary Regression, Accounting for Confounding and Covariates - R-Packages for the Statistical Analysis of Spatial Data: spatial, maptools, splancs, spatstat,

D. Interpolation and Geostatistics

(0.5 Days) Exploratory Data Analysis - Non-geostatistical Interpolation Methods - Estimating Spatial Correlation using the Variogram: Exploratory Variogram Analysis, Cutoff, Lag Width, Direction Dependence, Variogram Modelling, Multivariable Variogram Modelling - Spatial Prediction: Universal, Ordinary, and Simple Kriging, Kriging in a Local Neighbourhood, Multivariable Prediction: Cokriging, Trend Functions and Their Coefficients, - Kriging, Filtering, Smoothing - Model Diagnostics: Cross Validation Residuals, Cross Validation z-Scores, Multivariable Cross Validation - Geostatistical Simulation

E. Modelling Areal Data

(0.25 Days) Spatial Neighbours and Spatial Weights - Testing for Spatial Autocorrelation - Fitting Models of Areal Data

(iv) Graphical analysis of spatiotemporal data

Overview

Course ID	1000027
Language	en
Duration	2 D ys
Delivery mode	Classroom
Course Type	
Target Group	Data Analysts
Prerequisites	Basics in R and Statistics
Method	Presentation with examples and hands-on labs.
Course level	Beginning

Course Dates

Chicago	Miami	New York
1,850.00 USD	1,800.00 USD	1,850.00 USD
20-21 Aug 08-09 Oct 26-27 Nov	06-07 Aug 24-25 Sep 12-13 Nov	13-14 Aug 01-02 Oct 19-20 Nov

Prices plus local taxes.

Course Description

Space-time datasets are indexed both in space and in time. Their one- or two-dimensional analysis will typically start displaying the data in diagrams revealing the inner nature and relationships of the underlying variables. This training is organized into three parts, each devoted to different types of data. Each part comprises several topics and hands-on labs according to the various visualization methods or data characteristics. In the first part of the training, you will see how you can visualize time series data by using packages like zoo and xts for the analysis of time series data and packages like ggplot2, latticeExtra, and googleVis for their presentation. The next part of the training focuses on visualisation techniques for spatial data and presents packages like raster, rasterVis, maps, and googleVis. The third and last part finally combines variables which measure time and spatial data and teaches you how to create diagrams for such complex datasets.

Course Outline

A. Visualization of Time Series

(0.75 Days) Introduction to Displaying Time Series - Time on the Horizontal Axis - Time as a Conditioning or Grouping Variable - Time as a Complementary Variable - R Packages for time series data: zoo and xts - R Packages for visualization: ggplot2, latticeExtra, and googleVis

B. Visualization of Spatial Data

(0.75 Days) Introduction to Displaying Spatial Data - Thematic Maps: Proportional Symbol Mapping, Choropleth Maps, Raster Maps, Vector Fields - Reference and Physical Maps - Packages for working with OpenStreetMap - R Packages for spatial data: sp, maptools, gstat, and rgdal - R Packages for visualization: raster, rasterVis, maps, and googleVis

C. Visualization of Space-Time Data

(0.5 Days) Introduction to Displaying Spatiotemporal Data - Spatiotemporal Raster Data - Spatiotemporal Point Observations - R Package spacetime for spatiotemporal data

(v) Statistical Analysis with Graphics using R

Overview

Course ID	1000024
Language	en
Duration	4 D ys
Delivery mode	Classroom
Course Type	
Target Group	Data Analysts
Prerequisites	Basic knowledge of statistics
Method	Lecture with examples and exercises.
Course level	Advanced

Course Dates

Chicago	Miami	New York
2,400.00 USD	2,250.00 USD	2,400.00 USD
27-30 Jul 07-10 Sep 19-22 Oct 30 Nov - 03 Dec	24-27 Aug 05-08 Oct 16-19 Nov 28-31 Dec	31 Aug - 03 Sep 12-15 Oct 23-26 Nov

Prices plus local taxes.

Course Description

Graphics can effectively complement statistical data analysis in various ways. Successful graphics arise from a combination of good design and good implementation. This training explores mainly two R packages for statistical graphics: lattice and ggplot2. The lattice package extends the R language by providing a coherent set of tools to produce statistical graphics with an emphasis on multivariate data. ggplot2 is an R package for producing statistical, or data, graphics, but it is unlike most other graphics packages because it has a deep underlying grammar. This makes ggplot2 very powerful, because you are not limited to a set of pre-specified graphics, but you can create new graphics that are precisely tailored for your problem. The training is divided into two parts, with the first being an introduction to the development of graphics using the lattice package and the second one using the ggplot2 package for similar visualizations but also far more complex and more sophisticated visual analyses.

Course Outline

A. Introduction to graphics in R

(0.25 Days) Introduction to the lattice package - Multipanel conditioning - The "trellis"" object and its properties: the formula, data, conditioning and various plots/tiles in one diagram - Dimension and physical layout - Grouped displays - Annotation: Captions, labels, and legends

B. Graphics for Univariate Distributions

(0.75 Days) Density Plot - Histograms - Normal Q–Q plots - The empirical CDF (Cumulative Distribution Function) - Box-and-whisker plots - Strip plots - Working with small and large datasets

C. Graphics for Multivariate Distributions

(0.5 Days) Displaying Multiway Tables: Dot plots, Bar charts, Visualizing categorical data - Scatter Plots and Extensions - Trivariate Displays: Three-dimensional scatter plots, Surfaces and two-way tables

D. Advanced Graphical Parameters of the lattice Package

(0.5 Days) The parameter system: Themes and devices - Plot Coordinates and Axis Annotation: Axis annotations (ticks and labels), Limits and aspect ratio, Scale components and the axis function, Labels and Legends - Data Manipulation: Combining data sources, Subsetting, Ordering levels of categorical variables, Manipulating the "trellis" Object

E. Introduction to graphics using ggplot2

(0.5 Days) Datasets - Basic use - Colour, size, shape and other aesthetic attributes - Plot geometries - Components of the layered ggplot2 grammar: Layers, Scales, Coordinate system, Faceting

F. Visualization in ggplot2

(1 Day) Layers - Overall layering strategy - Aesthetic mappings - Creating a plot - Basic plot types - Displaying distributions - Dealing with overplotting - Surface plots - Drawing maps - Revealing uncertainty - Statistical summaries - Annotating a plot

G. Optimizing plots for publication and presentation

(0.5 Days) Themes - Customising scales and geoms - Multiple plots on the same page

(vi) Statistical analysis using Bayesian Networks

Overview

Course ID	1000019
Language	en
Duration	2 D ys
Delivery mode	Classroom
Course Type	
Target Group	Data Analysts
Prerequisites	Basics in Statistics
Method	Presentation with examples and hands-on labs.
Course level	Beginning

Course Dates

Chicago	Miami	New York
1,750.00 USD	1,700.00 USD	1,750.00 USD
03-04 Sep 29-30 Oct 24-25 Dec	20-21 Aug 15-16 Oct 10-11 Dec	27-28 Aug 22-23 Oct 17-18 Dec

Prices plus local taxes.

Course Description

A Bayesian network, Bayes network, belief network, Bayes(ian) model or probabilistic directed acyclic graphical model is a probabilistic graphical model that represents a set of random variables and their conditional dependencies via a directed acyclic graph. There are three main inference tasks for Bayesian networks: Structure learning, inferring unobserved variables, and parameter learning. This training presents the diverse techniques of statistical data analysis using Bayesian networks and shows in hands-on labs using R how to implement the techniques and algorithms. You will become familiar with R packages like bnlearn, deal, pcalg, and catnet for structure learning, and you will get to know packages like gRbase and gRain for inferential analysis. Time series data will be analyzed using packages like vars, lars, simone, and GeneNet.

Course Outline

A. Introduction

(0.25 Days) Introduction to Graph Theory: Graphs, Nodes, and Arcs - Bayesian Networks

B. Bayesian Networks and Static Data

(0.75 Days) Bayesian Networks: Essential Definitions and Properties: Graph Structure and Probability Factorization, Fundamental Connections, Equivalent Structures, Markov Blankets - Static Bayesian Networks Modeling: Constraint-Based Structure Learning Algorithms, Score-Based Structure Learning Algorithms, Hybrid Structure Learning Algorithms, Parameter Learning

C. Bayesian Networks and Time Series Data

(0.5 Days) Time Series and Vector Auto-Regressive Processes (VAR) - Dynamic Bayesian Networks: Essential Definitions and Properties, Dynamic Bayesian Network Representation of a VAR Process - Algorithms: Least Absolute Shrinkage and Selection Operator (LASSO), James–Stein Shrinkage, First-Order Conditional Dependencies Approximation

D. Bayesian Network Inference Algorithms

(0.25 Days) Reasoning Under Uncertainty: Probabilistic Reasoning and Evidence, Algorithms for Belief Updating: Exact and Approximate Inference, Causal Inference - Inference in Static Bayesian Networks: Exact Inference, Approximate Inference - Inference in Dynamic Bayesian Networks

b. Disclaimer

Comelio GmbH
Goethestr. 34
13086 Berlin
Germany

- Tel: +49.30.8145622.00
- Fax: +49.30.8145622.10

- www.comelio.com | .de | .at | .ch
- www.comelio-seminare.com
- info@comelio.com
- https://www.facebook.com/comeliogroup
- https://twitter.com/Comelio