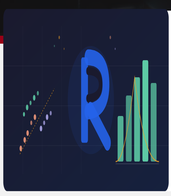
Medical Data Analysis for Young Scientists

Summer School 2026

If the word "statistics" makes you want to close this page—don't! This course isn't about memorizing formulas or doing hand calculations. It's about learning to ask meaningful questions of real medical data and getting clear, actionable answers. Whether you're tracking disease outbreaks, evaluating treatment effects, or exploring health trends, data skills are becoming as essential as a stethoscope in modern medicine.

Learn R | Manage and Load Real Data | Explorative Data Analysis | Data Quality | Regression and Confidence



Learn R

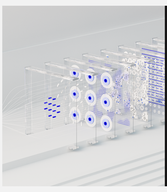
R is a premier language specifically designed for statistical computing and data analysis. Its strength is powered by a vibrant, global community that continuously develops new methods and packages, making it an indispensable tool for turning data into insight.

Load Real Data

You likely use Excel to organize results—it's a familiar starting point. However, for analyzing larger, complex datasets, Excel has limits. This module will show you how to import your data into RStudio, a user-friendly environment for more powerful, reproducible, and efficient analysis.

Explorative Data Analysis

Exploratory Data Analysis (EDA) is a dynamic process of investigating datasets. By leveraging visual and statistical techniques, it focuses on understanding data structure, uncovering key insights, and identifying potential issues—setting a clear, informed direction for deeper analysis.



Data Quality

Garbage in, garbage out. Data management is critical from the very beginning of your study. Assessing data quality—verifying accuracy, completeness, consistency, and anomalies—isn't a chore; it protects the integrity of your results. This module will equip you to build these essential checks into your research workflow.

Regression & Confidence

Regression is a core technique in both statistics and machine learning that quantifies relationships between variables. It not only estimates the effect of predictors but also provides confidence intervals—a crucial measure of uncertainty that is essential for building transparent and trustworthy AI models.

Course Structure

Day 01

Introduction to Medical Data Analysis - How to load data for analytics (and why)

Welcome and introduction to the workshop
Introduction to R and RStudio
Practice: Live coding, Hands-on writing basic scripts

Day 02

Explorative Data Analysis - Visualisation, Trends & Distributions

Statistical Foundations, Explorative Data Analysis with data visualisation, descriptive statistics
Practice: Hands-on Creating Basic Statistical Plots and Summaries in R

Day 03

Dive into your Data - Analyse Data Quality: Detect outliers and missing values

Understand core data quality principles and cleaning techniques
Master essential data preparation skills using R.
Practice: Live coding, hands-on lab by cleaning a real medical dataset.

Day 04

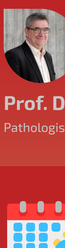
Detect Releations and Analyse Confidence

Foundations of correlation, regression, and time series analysis.
Implementing statistical models using R.
Practice: Live coding, hands-on lab to visualize and interpret relationships in a clinical dataset.

Day 05

The Big Picture: Applying the Complete Data Analysis Pipeline

Development of a general medical data analysis pipeline
Collaborative Project Work: Applying Skills to Real-World Medical Data
Closing Summer School: Project Presentations, Closing Remarks and Certificate Distribution



Your Instructor

"Learning is an exciting adventure: enjoy discovering and understanding by doing and experimenting!"

Prof. Dr. med. Thomas Schrader

Pathologist, Computer Scientist, Clinical Risk Manager

Technische Hochschule Brandenburg
University of Applied Sciences
Fachbereich Informatik und Medien

14. September - 18. September 2026
Mo - Fr
09:00 - 16:00 Uhr

Medizinische Hochschule Brandenburg Theodor Fontane
Lehrgebäude am Nikolaiplatz
Brandenburg an der Havel

Summer School Sprache: Deutsch
Präsentationen & Online-Material: Englisch