DIGS-BB **Training for PhD students WS 2021**

Applied statistical data analysis with R

Lecturer in charge:	Dr Michael Seifert (Group Leader Bioinformatics Core Unit, IMB TU Dresden)
Date:	Winter Semester (1 – 3 Dec and 7 Dec 2021)
Time:	9:00 - 13:00
Location:	House 105, Blasewitzer Str. 86, Computer pool (2. Floor, Room 2.350)
Target audience:	DIPP PhD Students
No of participants:	10
Registration deadline:	12 November 2021. A letter of motivation needs to be provided (half a page).
Prerequisites:	The target audience are beginners. No specific basic knowledge about R is required. It would be helpful if you already have basic knowledge about statistical testing.
Course requirements:	Full attendance and active participation during the course and exercises. Corona 3G rules and mask for every day.
Course conclusion:	After the course, the participants should be able to work with R and Rstudio, to transfer learned methods and concepts to the analysis of own data sets, to realize a basic statistical analysis in R, and to interpret the results. There will be no exam, but there are exercises during the course that you have to do on your own to apply the learned basics.

Introduction to the course (what does the course cover)

Course aim: The introductory course provides a basic understanding of how to use the statistics software R in combination with the user interface Rstudio. Starting with an introduction to the basic functionality of R, practical data sets are analyzed in the course in order to learn how to use different statistical methods and how to create different types of graphics.

TECHNISCHE UNIVERSITÄT DRESDEN

dresden international phd program

Training for PhD students WS 2021

Learning goals

- R as calculator, comfortable user interface RStudio
- Introduction to R syntax and basic data structures (vector, dataframe)
- Data sets: load, check, modify, and save
- Graphics: histogram, box plot, scatter plot, Kaplan-Meier curves
- Statistical methods and their assumptions: t-test, Wilcoxon test, chi-square test, linear regression, Kaplan-Meier estimator, logrank test

Course Content:

01 Dec 2021: Introduction: R and Rstudio; Basics Part 1: Data types, Operations, Data structures

02 Dec 2021: Basics Part 2: Work with data sets; Visualization: Histogram, Box plot, Scatter plot

03 Dec 2021: Visualization: Histogram, Box plot, Scatter plot; Statistical Tests Part 1: Basics and assumptions, t-test

07 Dec 2021: Statistical Tests Part 2: Chi-Square; Regression; Survival Analysis: Kaplan-Meier-Estimator, Logrank-Test

Preparation of Course:

It would be helpful if you would already have basic knowledge about statistical testing (e.g. null hypothesis, alternative hypothesis, p-value, test decision). Wikipedia could be a starting point for this (https://en.wikipedia.org/wiki/Statistical_hypothesis_testing). You can further have a look at the R website (https://www.r-project.org/) and at Rstudio (https://www.rstudio.com/) to have a basic idea about the software tools that we will use during the course.