

R in statistical analysis for geneticists

Organization and carrying out: **prof. Idzi Siatkowski** - Head of the Department of Bioinformatics in the Chair of Mathematical and Statistical Methods, Poznań University of Natural Sciences

20 hours: lectures + computer workshops

Courses will take place in a lecture room in the building of IPG PAS, all PhD students are asked to bring laptops.

Book for classes: Siatkowski, Zyprech-Walczak. R w analizie statystycznej dla przyrodników. Poznań: Publisher University of Life Sciences in Poznan, 2018. ISBN 978-83-7160-899-5. [in Polish]. You can buy the book during classes from prof. Idzi Siatkowski.

Topics:

1. Introduction (1 hour)

- 1.1. What is R?
- 1.2. The advantages of R
- 1.3. Installing R and RStudio
- 1.4. Packages
- 1.5. Help and documentation

2. Calculations in R (3 hours)

- 2.1. Simple mathematical calculations
- 2.2. Variables
- 2.3. Numeric and character vectors, matrices and data frames

3. Data preparation (2 hours)

- 3.1. Data frames
- 3.2. Text files
- 3.3. Excel files
- 3.4. Data of various formats

4. Visualizations (3 hours)

- 4.1. Graphic presentation of data
- 4.2. Function graphs

5. Testing (4 hours)

- 5.1. Introduction
- 5.2. Two populations - normal distributions
- 5.3. Two populations - any distributions
- 5.4. Analysis of variance — ANOVA
- 5.5. Multiple comparison tests

6. Dependencies (1 hour)

- 6.1. Correlations
- 6.2. Contingency tables

7. Linear and multiple linear regression (4 hours)

- 7.1. Linear regression
- 7.2. Multiple linear regression
- 7.3. Selection of variables

8. Exam (2 hours)