

# Bi7550 Analysis of community ecology data in R program (AnaDatR)

Teacher: [David Zelený](#), Masaryk University Brno, Czech Republic

Spring semester 2015, every second Friday 8:30 - 11:30 in [A9-316](#); link to IS: [Bi7550](#)

## Class schedule

## Course history (scripts created during the course)

- 2015-02-20
  - [history of commands](#)
  - [random walk script](#)
- 2015-03-06
  - [history of commands](#)
  - [the script from the class](#)
  - [HOMEWORK 1 \(Exercise 2 from CA+DCA\)](#)
- 2015-03-20
  - [the script from the class](#)
  - [HOMEWORK 2 \(Exercise 2 from Suppl variables\)](#)
- 2015-04-03
  - [the script from the class](#)
  - NO homework :)
- 2015-04-17
  - [the script from the class](#)
  - [HOMEWORK 3 \(Exercise 1 from Forward selection\)](#)
- 2015-04-24
  - [the script from the class](#)

## Final report

### Summary

- **Language:** Czech, Slovak or English;
- **Length:** 5-10 manuscript pages, including everything;
- **Deadline:** one day before oral exam, upload to Odevzdávárna repository on IS.
- **Please, bring one printed copy of the report for the exam - we will use it for discussion!**

## Details

Use your own multivariate data, or try to borrow some data from friends or colleagues, or try to find some from freely available online sources (some are listed in the section *Sources of published or freely available ecological datasets* in [Links](#)). If you can't get any data, write me, I will send you some. Data should be multidimensional - community matrix and matrix of environmental variables. Use these data to conduct any three analyses we discussed at the course (e.g. indirect ordination, direct ordination and forward selection).

The report should contain the following items:

1. brief description of the dataset (what kind of organism, sampling design, how many samples or species or environmental variables etc.);
2. question(s) you tried to answer by your analyses, and which analyses you have chosen for it;
3. description of results and what is your opinion about that;
4. appended R script.

Don't forget to include also figures!

The number of pages is not limited, but please, keep it short, I need to read all of them (mainly do not copy there parts of your theses with details not relevant for the data or analysis). Recommended length is 5-10 manuscript pages, including everything.

The report **should contain R script** you used for analysing the data. By script I DO NOT mean the history of all commands you used for analysis (including things like searching in the help or unsuccessful trials to use some commands) - I mean the script which you can directly use to re-analyze the whole study, i.e. containing data import, uploading libraries, analysis and drawing of figures. Feel free to add comments inside to make it easier to navigate for you and me during the exam. **Please, append the script to the end of the report, and format the font using Courier New.**

Various deviations from these rules are possible and welcome. If you have your own dataset and you want to play with it in R, you are welcome to do it in the form of this final report. If you are not sure if the data are suitable, just mail me to ask.

To hand the report, please **upload it at least one day before the exam into Odevzdávárna repository**. I need to see it one day before (optimally 24 hours) to be able to read all reports before the day of exam. If you cannot access Odevzdávárna in IS, send me the report by email. If your report is for joint exam together with exam for Bi7540 Data analysis in community ecology, it's enough to deposit it only once into Bi7540 repository.

**Bring one printed copy of the report for the exam!**

## Details about exam

Language of exam: Czech, Slovak or English.

**To register for the exam date, please use the IS system - there you will find all the dates available for exam.** The evening before the examination day I will send email to all registered participants with the time schedule for examinations (exact time you should come, so as you don't

have to wait the whole day for it). The order will be derived from the time of registration for the exam in IS.

The exam will take ~30 minutes. You can bring whatever you want, notes from the class or some books, and you are free to use them. We will discuss the analyses you applied on the data in your final report. I will first ask some details about your data, what kind of question you tried to answer and what analysis you have chosen for it (potentially also what other methods could have been used instead). I won't require deep understanding of the theory behind (but this is required at the exam for the parallel class Bi7540) or that you should remember names of all arguments in functions - I will be happy if you can show me that you have idea about what function can be used for what, or how to find it out. Expect that I will point to the appended R script in your report to ask what the specific line of the script is actually doing, to be sure that it is written by you.

## Joined exam from Data analysis in community ecology (Bi7540) Analysis of community ecology data in R (Bi7550)

Information about rules for this joined exam can be found at [the website of Data analysis in community ecology \(Bi7540\)](#).

---

### About the class

The class is focused on the use of R for analysis of multivariate ecological data. It is meant to be a continuation of [Bi7540 Data analysis in community ecology](#) (both classes could be taken in one semester, although for a formal reason you need to apply for exception - this will be automatically granted), but oriented more practically and limited purely on the use of R (I expect, at least partially, students to gain their theoretical knowledge in Bi7540 or other courses). Before signing for this course, student should have also elementary experience with R program (gained by self-study or by attending other classes, such as Bi7560 Introduction to R, Bi8190 Visualization of biological data, Bi7920 Analysis of biological data etc.). In case that enrolled students have no experience with using of R program, I will at the beginning insert one class devoted to "simple guide to R", which should ensure that everybody will have basic ability to master R program for the purpose of this class.

### Syllabus

- Introduction, data import, vegan library, recommended references, details about final examination.
- Unconstrained ordination (CA, PCA, DCA, PCoA, NMDS, drawing ordination diagrams, environmental variables in unconstrained ordination).
- Constrained ordination (RDA, CCA, db-RDA).
- Monte Carlo permutation test, forward selection, variance partitioning.
- Numerical classification (hierarchical and non-hierarchical classification, dendrogram, evaluation of classification results, indicator species).
- *Optional*: Trait vs environment analysis (community-weighted mean, fourth-corner problem, RLQ).

- *Optional:* Analysis of diversity (alpha and beta diversity, rarefaction curves).

*(Note: optional topics will be inserted in case that participants are interested and there is enough time for it - in any case, teaching materials for these topics will be provided)*

## What to prepare for class

- Computer with installed R program and connection to internet. You may preferably install also some editor, e.g. RStudio (see [About R program](#) link for instruction how to install)
- If you won't be connected to internet during the class, install also these libraries: vegan, cluster, maptree, rgl, MASS, labdsv, BiodiversityR, packfor. To do this, open R (Rgui.exe), get connected to internet, and to command line copy the following:

```
install.packages (c('vegan', 'cluster', 'maptree', 'rgl', 'MASS', 'labdsv',  
'BiodiversityR'))  
install.packages('packfor', repos='http://R-Forge.R-project.org')
```

From:

<https://www.davidzeleny.net/anadat-r/> - **Analysis of community ecology data in R**

Permanent link:

<https://www.davidzeleny.net/anadat-r/doku.php/en:anadatr-2015>

Last update: **2017/10/11 20:36**