Ecological data analysis in R (OCEAN 5098)

Information

Time: Tuesday 13:20- 16:20 (Spring Semester - odd year)

Lecturers: Vianney DENIS

Credits: 3

Description

This course is mainly designed to introduce students with the use of R language. It will introduce R environment but also popular functions to sort, visualize and analyze data in ecology. Broad topics covered will include: introduction to R language and basic functions, data manipulation, graphics, maps, linear model, parametric and non-parametric analysis, multivariate analysis, etc.

Aim and objectives

The aim of this course is for the students to get familiar with R language. This course explores the multiple advantages of using R language for formatting reports, preparing presentation, exploring and analyzing ecological, data editing and sharing project. The course assumes that students have no prior knowledge in R or other programming languages. It will start from scratch, **i.e.** downloading R and installing it on your computer. Along the semester, you will be introduced to the use of RStudio/Posit GUI, Rmarkdown, and GitHub repositories that complete harmoniously your R experience.

After taking this course, a student should:

- be aware of the multiples advantages of using R;
- not be afraid of using code to organize, visualize and analyze data;
- become a self-learner, able to explore and solve problem;
- know several statistical tools for the analysis of ecological data in particular;
- be able to analyze and valorize scientifically any simple ecological data sets;
- be able to help and advices on the use of R.

Tentative Schedule

Schedule is susceptible to changes according to progress along the semester

Week	Tentative Topic	Content (main)
1	Introduction	Presentation course
2	Environment	R, R Studio, R Markdown, Git/GitHub

Week	Tentative Topic	Content (main)
3	Data manipulation	Formatting and converting data set
4	Graphics	Visualization functions and customization, ggplot2
5	Statistics I	Generate and test simple hypotheses
6-7	Statistics II	Generalized Linear Models
8-9	Similarities	Ecological resemblance and matrix
10	Classification	Cluster analysis
11-12	Ordination: unconstrained	PCA, PCoA, nMDS
13-14	Ordination: constrained	Redundancy and canonical analysis
15-16	Functional analysis	From trait to ecological functions

Evaluation & Final Assignment

(1) Along the semester, you will be assigned with a couple of more or less simple exercise to practice what was learn in class. There is no secret, practice will make you better and I need to make sure you are not lost.

(2) At the beginning of the class, one student may be randomly pick up in order to summarize his homework (*.Rmd* and/or *.html*) or knowledge he got form previous class (5-10 minutes).

(3) DATATHON (**NEW**) as final assignement. Students will be provided with a data set to analyze (type 'Scientific Data': https://www.nature.com/sdata/). Groups (2-3 students) will identify relevant research questions based on this data set and will prepare a short scientific report (max. 20 pages + references) while emphasizing methodology and results (scientific format). Introduction and discussion/conclusion could be presented as bullet points. Report should be sent as a *.Rmd* **AND** *.html* or share through a GitHub repository.

Final score is typically composed of participation/assiduity in class (**10%**), homework/review (**20%**), final group report (**70%**)

Requirements

Basic knowledge and interest in Biology/Ecology/Informatics. The course requires students to bring/use their own personal computer.

References & Readings

Borcard, D., Gillet, F. and Legendre, P. (2018) Numerical Ecology with R. Springer. DOI: 10.1007/978-3-319-71404-2

Paradis, E. (2005) R for beginner. Available among many others documentation at https://cran.r-project.org/other-docs.html (page frozen and no longer actively maintained)

Ramette, A. (2007) Multivariate analyses in microbial ecology. FEMS Microbiology Ecology, 62: 142-160. DOI: 10.1111/j.1574-6941.2007.00375.x

Xie, Y., Allaire, J.J.,Grolemund, G. (2020) R Markdown: The Definitive Guide. Chapman & Hall/CRC. The online version of this book is free to read here.

Zuur, A. F., Ieno E. N., Smith, G. M. (2007) Analysing Ecological Data. Springer. DOI: 10.1007/978-0-387-45972-1

Zuur, A. F. (2010) A protocol for data exploration to avoid common statistical problems. Methods in Ecology and Evolution, 1: 3-14. DOI: 10.1111/j.2041-210X.2009.00001.x