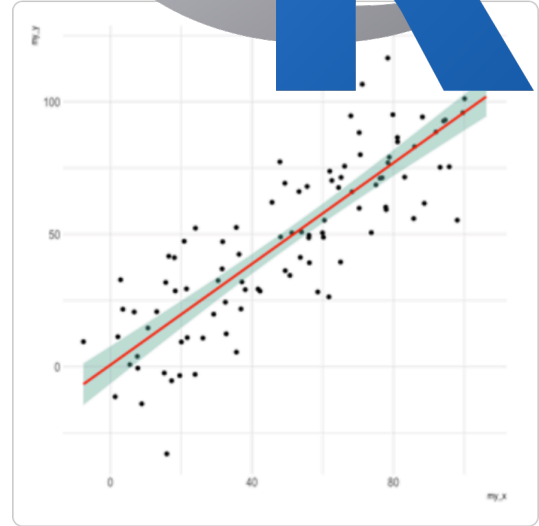
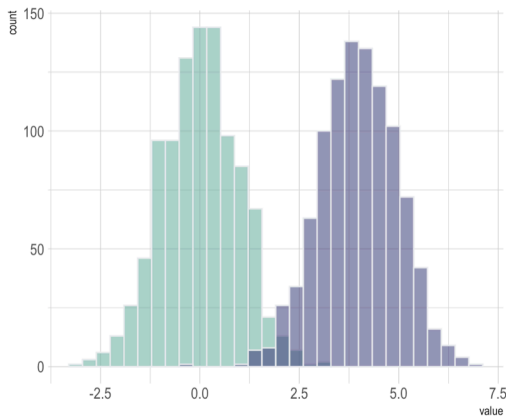


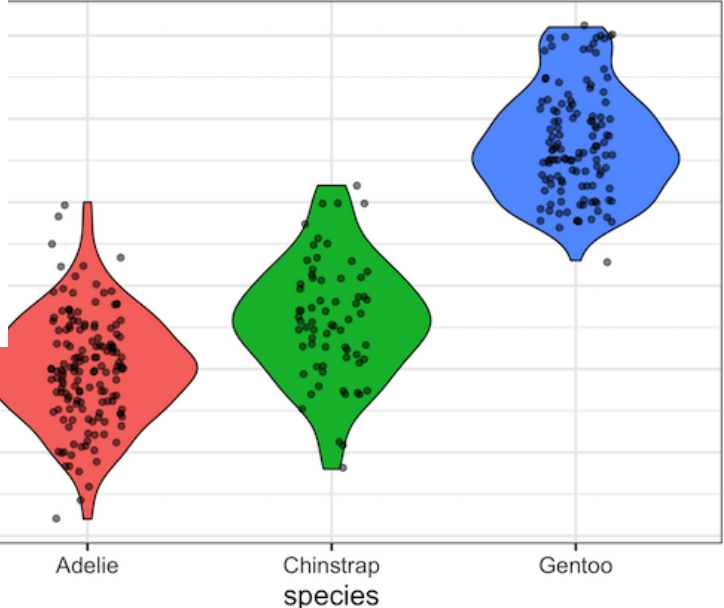
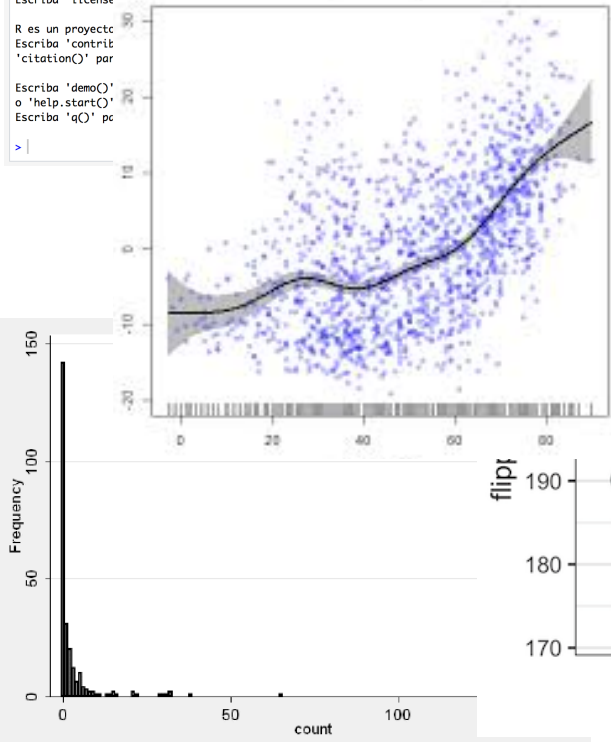
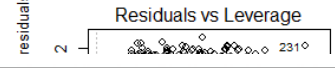
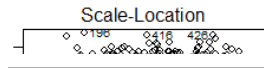
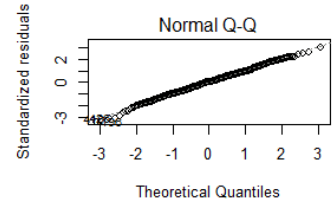
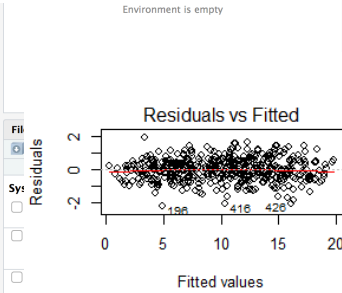
Introduction to R language

5-9 June 2023

Aldo Barreiro Felpeto



```
1 ##Section
2
3 ##Load data
4 data<-read.csv("data.csv", sep=";", as.is=T)
5 head(data)
6
7 ##Linear model
8 mod1<-lm(Chol.Blood~Gender*Physical.condition*Age,data=data)
9 summary(mod1)
10 plot(mod1)
11 res<-residuals(mod1)
12 shapiro.test(res)
13 require(car)
14 leveneTest(res~Gender*Physical.condition,data=data)
15 #Analysis of variance
16 Anova(mod1)
17
18 #Interaction plot
19 interaction.plot(data$Gender,data$Physical.condition,data$Chol.Blood)
20
```



INTRODUCTION TO R LANGUAGE

CIIMAR, 5-9 June 2023

Session 1. 10:00 – 14:00 h, 5 June 2023

- 1.1 History of R. Features of R as a programming language. Advantages and disadvantages of using R
- 1.2 Graphical users' interfaces
- 1.3 Objects
- 1.4 exercises

Session 2. 10:00 – 14:00 h 6 June 2023

- 2.1 Functions
- 2.2 Indexing vectors, matrices, arrays, lists and data frames
- 2.3 Exercises
- 2.4 Loops and the family of functions “apply”

Session 3. 10:00 – 14:00 h 7 June 2023

- 3.1 Conditional statements
- 3.2 Exercises
- 3.3 Importing and exporting data
- 3.4 Exercises

Session 4. 10:00 – 14:00 h 8 June 2023

- 4.1 Introduction to *tidyr*, *Rmarkdown* and *R notebook*
- 4.2 Basic R graphics packages
- 4.3 Construction of basic plots I
- 4.4 Construction of basic plots II

Session 5. 10:00 – 14:00 h 9 June 2023

- 5.1 Exercises
- 5.2 *ggplot2* package I
- 5.3 *ggplot2* package II
- 5.4 Exercises

Docent: Aldo Barreiro Felpeto. CIIMAR.

Price: 200 € (150 € for CIIMAR/UP members)

Registration: after announcement, up to fill 25 available positions.

Registration, together with the payment information, is available in the CIIMAR website, through the link that is sent with the e-mail announcing the course.

Proof of payment required to book the place (send proof to abarreiro@ciimar.up.pt).

After sending the proof of payment, a confirmatory e-mail for the registration will be sent.

Important information:

- All the course will be taught through a zoom platform.
- The course will be taught in English.
- No previous background with R is needed.
- All the information and materials necessary for the development of the course (instructions to install R and R packages, pdf with lessons content, scripts with examples and exercises, data for examples and exercises) will be made available for all the participants in the course through a link to the *Open Science Framework* platform.

Syllabus

Course description

The course begins with an introduction to the R language from a historical perspective, highlighting the features that make it different from other languages and how these contribute to its success. They are also exposed the advantages of disadvantages of learning R, which simply depend on the user needs.

Then, a brief overview of graphical user's interfaces available in R. Here will be highlighted the use of *RStudio*, due to its increased use, although the course material is not developed with this interface.

The rest of the sessions 1-3 consist in a hierarchical overview of the language components, with examples and exercises to train. During these sessions, students of any level are expected to understand what is an R object and its classes, control structures (loops, conditional statements), what is a function, how to operate with them and the basis to program them. It will be shown the variety of functions available in R, classified by families and/or tasks for which they were developed.

At the end of these sessions will be shown examples and performed exercises of operationally simple but conceptually advanced programming tasks.

At the beginning of session 4, it will be made a short introduction with examples to the new trend in R language represented by *tydir*, as well as the applications of *Rmarckdown* and *R notebook*, which are strongly increasing its use for the development of reproducible documents and workflows.

The rest of session 4 and session 5 are devoted to the construction of plots, with an overview of the nowadays extremely wide variety of possibilities offered by R. Also, for beginner students, it will be shown how to construct plots from its more basic elements. There will be a special emphasis on the package *ggplot2*, due to its increase in popularity.

The course is open to any level from undergraduate students to senior researchers.

Course methodology

The course contents are short theoretical introductions to each topic with worked examples in R language, throughout which all features of that specific topic are shown. At the end of each section, the students will perform exercises that will be revised and corrected during the lessons.

The theoretical explanations, as well as the worked examples are fully developed in a pdf, with R scripts with the examples as additional support. Access to all this material will be provided days before the course through a link to the *Open Science Framework* platform. Solutions to the exercises will also be delivered to the students by the end of the course.

General aim of the course

Learning the elementals of R programming. Getting an overview of features that R offers for more advanced users. Provide reference material (explanations, examples) for further development of more advanced skills in R programming.

Specific aims of the course

- Understand the features that make R different from other software.
- Learn the essential operations that will need to be performed in most of our R sessions (setting working directory, loading data, exporting data and plots).
- Getting an overview of the variety of R functions.
- Being introduced to trendy applications in R (*tidyr*, *Rmarckdown*, *R notebook*, *RStudio*, *ggplot2*).
- Getting an overview of the potential of R for graphics.
- Constructing graphics with the most important elements in the *graphics* and *ggplot2* packages.
- Acquire a clear perspective about how R could be useful for us and improve our work.



Aldo Barreiro Felpeto is a researcher at Centro Interdisciplinar de Investigação Marinha e Ambiental (CIIMAR) associated to the University of Porto (Porto, Portugal). His research career has focused in plankton ecology. He defended his Ph.D. dissertation in 2007 in the Department of Ecology at the University of Vigo (Vigo, Spain) about interactions between zooplankton and toxic phytoplankton species from the Spanish NW Atlantic coast, southern Baltic sea and southern Tirreno coast. In 2008-2010, he performed a post-doctorate in the

Department of Ecology and Evolutionary Biology at Cornell University (Ithaca, New York, USA). Since 2011 he is a researcher at CIIMAR.

He developed a strong background in statistics and dynamic modelling with R software, attending 10 courses in the period 2006-2018 and since 2013, organizing 14 editions of courses about different aspects of statistics and programming with R, mostly in CIIMAR, but also in the University of Vigo (Spain) and the University of Magallanes (Chile). He co-authored two books about statistics and programming: *Tratamiento de Datos* (Ed. Díaz de Santos, Madrid, 2006) and *Tratamiento de Datos con R, SPSS y ESTADÍSTICA* (Ed. Díaz de Santos, Madrid, 2010).

Due to his expertise in statistics and programming, he has developed collaborations in different fields of ecology, but also environmental sciences and molecular biology. He has published 53 articles, accounting for an *h* index of 22 and an *i10* index of 38.