Using R as continuous learning support in Sea Sciences Degree

Jose Zubcoff*1, Francisco Gomariz-Castillo1,2, Jose Vicente Guardiola1 and Aitor Forcada1

1. Departamento de Ciencias del Mar y Biología Aplicada, Universidad de Alicante, 03080, Alicante, España
2. Instituto Euromediterráneo del Agua, Campus de Espinardo, s/n, 30001 Murcia, España
*Contact author: jose.zubcoff@ua.es

Keywords: bioinformatics, biostatistics, statistics learning, education research

This work presents how R is used in the development of skills within the Sea Sciences Degree at the University of Alicante. R is used as an aid in the gradual development of the student curriculum: from basics statistics in the first year, to complex experimental design, multivariate analysis or application to other fields such as geostatistics, in the latter years of the degree. Usually, the use of different software along a degree hinders the student’s acquisition of knowledge, hence it is desirable to homogenize the use of tools, in order to focus efforts on learning concepts/methodologies. In this regard, we strongly recommend the use of R as a tool for support learning due to its great versatility.

The use of R along the course is justified by several advantages: i) it is free and open source software; ii) it is cross-platform, which facilitates its use by students; iii) it has great potential for data manipulation; and iv) the large and dynamic community that gives support to R and also the amount of documentation that is easily accessible. The supporting community has focused, not only on the implementation of new treatment techniques and data analysis, also on the development of packages designed to facilitate the access and consultation to some specific information, such as rfishbase, a client access to the specialized database FishBase, widely used in Marine Sciences.

We introduce R in the first year Statistics subject. R is used to initiate students into basic concepts of descriptive statistics and univariate and bivariate statistical inference (goodness of fit contrasts, contrasts normal populations, one-way ANOVA, contrasts for proportions, measures of association and linear regression) to allow the student to provide a basis for dealing with the resolution of problems in subsequent courses. In this course, students are also introduced into aspects of importing data or connection to databases using rodbc. Later, in the third year of the degree, within the subject "Statistics applied to marine resources", R and specifics packages are used to learn about experimental design, multifactorial ANOVA (using package GAD), and several multivariate analysis such as multiple linear regression, cluster, MDS, PCA, … (using packages faraway, vegan, scatterplot3d, rgl and bpca). Finally, as an optional subject, in "Geographic Information Systems and Remote Sensing" R is used for support both, GIS and remote sensing, due to its interoperability with other Open Source softwares like GRASS and QGIS. In this case, R is used for teaching geostatistics, through the wide variety of packages available for management and analysis of spatial data (maptools, rgdal or spgrass6 to work with spatial information and sp, gstat or automaps for spatial interpolation and geostatics). In this course, R is also used for the spatial distribution of species by glm models, and as a complementary tool for analysis by specific libraries to support several tasks, such as the analysis of spectral signatures or supervised and unsupervised classification in remote sensing.

As a result of the formative development, at the end of the degree, the student must have acquired among other competences: be able to apply mathematical and statistical knowledge to biology and the different application areas (resources, planning and management, modeling, etc.), understand and apply basic aspects of sampling and data processing, get the ability to recognize and solve different problems using these tools and, be able to communicate the results of the experiments. Furthermore, the use of R will be also a valuable skill acquired by students.