

Taller introductorio de R para el análisis de datos en ciencias ambientales
Posgrados en Ciencias Ambientales – Facultad de Ciencias – UDELAR
Docente a cargo: Marcos Texeira

Carga horaria: 30 horas de docencia directa

Créditos: 4

Unidad Análisis e integración

Horario:

09:00 a 12:00

13:00 a 16:00

Programa

INTRODUCCIÓN:

R es un poderoso lenguaje de programación y entorno “dedicado” para el análisis de datos. Este taller pretende ser una introducción a R, dirigido a usuarios principiantes e intermedios provenientes de las ciencias ambientales y disciplinas relacionadas. La primera parte del taller estará dedicada a los fundamentos y rudimentos del lenguaje, mientras que la segunda parte introduce al análisis de datos con diferentes técnicas (desde las básicas y clásicas a algunas modernas) y a la interpretación de los resultados de dichos análisis.

OBJETIVOS:

El objetivo del taller es familiarizar al estudiante con el entorno de programación R y con sus funcionalidades básicas, así también como con algunas de las funcionalidades avanzadas. Para ello se proponen una serie de actividades de análisis basadas en ejemplos.

CONTENIDO:

Día 1-2

- Fundamentos de programación en R
- Instalación de R y editores
- Línea de comandos / scripts
- Manejo de directorios y estructura de datos
- Control de flujo – loops
- Funciones empotradas (built-in) y definidas por el usuario

- Paquetes
- Gráficos básicos y mediante el ggplot

Día 3

Estadística clásica con R

- Pruebas clásicas (test de t, correlación, no paramétricos clásicos)
- Modelo lineal I: regresión lineal simple
- Modelo lineal II: análisis de experimentos factoriales – ANOVA
- Modelo lineal III: Regresión lineal múltiple
- Estadística multivariada básica: análisis de componentes principales y análisis de agrupamientos (cluster analysis).

Día 4-5

- Extensiones a la estadística clásica en R
- Modelos no lineales
- Modelos lineales generalizados: regresión logística / regresión Poisson
- Modelos lineales con estructuras de autocorrelación (espacial y temporal)
- Árboles de regresión y clasificación / bosques aleatorios (Random Forests)
- Redes neuronales artificiales: preceptrones multicapa y redes recurrentes.

APROBACIÓN DEL CURSO

La aprobación del taller se consigue mediante la realización de una actividad práctica de proceso y análisis de datos en la cual se evaluarán las habilidades adquiridas por los estudiantes. Los mismos contarán con un plazo de un mes a partir de la finalización del taller.

BIBLIOGRAFÍA:

- Crawley, M.J. 2007. The R Book. Wiley.
- Faraway, J. 2006. Practical Regression and ANOVA using R. Chapman and Hall.
- Faraway, J. 2007. Extending the linear model with R. Chapman and Hall.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A.A y Smith, G.M., 2009. Mixed Effects Models and Extensions in Ecology with R. Springer.
- Borcard, D., Gillet, F. y Legendre, P. 2011. Numerical Ecology with R. Springer.
- Wickham, H y Grolemund, G. 2017. R for data Science. O'Reilly.