



UNIVERSIDAD NACIONAL DEL SUR  
BAHIA BLANCA - ARGENTINA

DEPARTAMENTO DE: BIOLOGÍA , BIOQUÍMICA Y FARMACIA

CURSO DE POSGRADO: Curso-Taller Introducción a R	CODIGO : <hr/>
--	-------------------

HORAS CLASE		PROFESOR RESPONSABLE
TEORICAS 14	PRACTICAS 18	Martín Amodeo

**OBJETIVO GENERAL**

En este curso teórico-práctico veremos el uso básico del software R para el análisis y procesamiento de datos en las ciencias biológicas y ambientales. El curso está destinado a becarios/as doctorales y post-doctorales, así como investigadores/as, profesionales y docentes de instituciones científicas públicas.

- OBJETIVOS PARTICULARES**
- Comprender la importancia del software basado en código, como *R*, en el marco de la ciencia abierta y reproducible.
  - Familiarizarse con la estructura de archivos y flujo de trabajo típicos en un proyecto de análisis de datos con aplicaciones en las ciencias biológicas y ambientales.
  - Experimentar con herramientas de manipulación y re-organización de los datos comúnmente utilizadas.
  - Desarrollar competencias y hábitos deseables en el contexto de un tratamiento ordenado de los datos (*tidy data*) y el flujo de trabajo con *R*.
  - Aplicar visualizaciones básicas de los datos que permitan explorar gráficamente y comprender su estructura, así como reportar resultados básicos.
  - Ejemplificar el uso del software en su aplicación a análisis básicos típicos de las ciencias biológicas y ambientales (análisis de correlación, modelos lineales).

**REQUISITOS**

Para realizar el curso-taller se requiere disponer de una computadora con acceso a internet, y tener instalados los software *R* y *Rstudio* (instructivos a disposición) y Zoom. Es recomendado tener nociones generales de estadística descriptiva y análisis básicos (análisis de correlación, regresión, varianza).

PROGRAMA SINTÉTICO

- Software libre, programación y análisis de datos
- *Software R* y *Rstudio* como interfaz de uso
- Funciones y paquetes. Interacción entre software base y librerías
- Objetos. Estructura de archivos, rutas y proyectos en *Rstudio*. Carga de datos
- Tipos de datos, tipos de variables. Organización y manipulación de datos (operaciones típicas, filtros, resúmenes)
- Variables y operaciones estadísticas básicas. Visualización de datos con *ggplot2*. Funcionamiento básico con una, dos y tres variables
- Aplicación a un proyecto de análisis de datos: Análisis de Correlación y Modelos Lineales. Estructura y flujo de trabajo

PROGRAMA ANALÍTICO (incluir actividades, bibliografía y tipo de evaluación)

El dictado del curso es de modalidad virtual, con una carga horaria total de 32 horas distribuidas en siete días. Los temas propuestos se organizan en distintos módulos que abarcan el uso básico del software, la estructura y flujo de trabajo en un proyecto de análisis de datos, herramientas de manipulación, visualización con aplicaciones a análisis comúnmente utilizados en las ciencias biológicas y ambientales. En cada unidad se expondrán los fundamentos teóricos básicos mediante clases expositivas, demostraciones prácticas y se llevarán a cabo actividades prácticas con un énfasis especial en desarrollar competencias y hábitos deseables en el flujo de trabajo con R. Las clases y demostraciones serán dictadas en encuentros virtuales sincrónicos diagramadas en secciones de 25 minutos con 5 minutos de descanso, totalizando 2 horas por día durante 7 días. Las actividades prácticas de aplicación se plantean de forma asincrónica involucrando 3 horas por día durante 6 días. Para cada práctico, los participantes tendrán instancias para recibir asistencia y evacuar dudas, una a modo de consulta (dirigida a quienes requieran) y otra incluida en las secciones expositivas del día siguiente a modo de recapitulación (dirigida a todos los participantes).

La evaluación se realizará mediante la participación y desempeño en seis actividades prácticas que involucran la entrega de documentos de código, los cuales serán evaluados según criterios de factibilidad, reproducibilidad y originalidad. Mediante la integración de estos índices de evaluación se verá plasmado si el participante logró aplicar las herramientas incorporando los conceptos planteados.

Unidad 1. Ciencia abierta y reproducible. Software libre, programación y análisis de datos. Uso del software, estructura y flujo de trabajo en un proyecto de análisis de datos. *Rstudio*, interfaz de uso.

Unidad 2. Funciones y paquetes. Objetos. Estructura de archivos, rutas y proyectos en *Rstudio*. Carga y manipulación de datos, conversiones y operaciones típicas en la estructura de datos (filtros, resúmenes).

Unidad 3. Tipos de datos, tipos de variables. Organización y manipulación de datos (operaciones típicas, filtros, resúmenes). Variables y operaciones estadísticas básicas.

Unidad 4. Visualización de datos con *ggplot2*. Funcionamiento básico con una, dos y tres variables.

Unidad 5. Aplicación a un proyecto de análisis de datos: Análisis de Correlación y Modelos Lineales. Estructura y flujo de trabajo según los objetivos de investigación.

## BIBLIOGRAFÍA

- Crawley, M. J. (2007). The R book. Wiley. Chichester, England.
- Kabacoff, R. (2022). R in action: Data analysis and graphics with *R* and *Tidyverse* (Third edition). Manning Publications. Los Ángeles, USA.
- Rodríguez-Sánchez, F., Pérez-Luque, A. J., Bartomeus, I., & Varela, S. (2016). Ciencia reproducible: ¿qué, por qué, cómo? *Ecosistemas*, 25(2), 83–92.  
<https://doi.org/10.7818/ECOS.2016.25-2.11>
- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer. New York, USA.  
<https://doi.org/10.1007/978-0-387-98141-3>
- Wickham, H., Çetinkaya-Rundel, M., & Grolemund, G. (2023). *R for data science: Import, tidy, transform, visualize, and model data* (2nd edition). O'Reilly Media, Inc. Sebastopol, USA.
- Zuur, A. F., Ieno, E. N., & Meesters, E. (2009). *A Beginner's Guide to R*. Springer New York, USA.  
<https://doi.org/10.1007/978-0-387-93837-0>