

## CURSO

# FUNDAMENTOS DE R

## Programación para la exploración, manipulación y análisis de datos

*Instructores: MSc Patricia Dell’Arciprete <sup>1,2</sup> y Dr Julio Lancelotti <sup>1</sup>*

*<sup>1</sup>CONICET – CENPAT (Centro Nacional Patagónico)  
<sup>2</sup>Universidad Nacional de la Patagonia San Juan Bosco*

### OBJETIVOS GENERALES

Este curso tiene por propósito enseñar el lenguaje de programación R. Mediante ejemplos y ejercicios específicamente diseñados, se enseña a aplicar en forma eficiente las principales herramientas para la manipulación y exploración, presentación gráfica y aplicaciones que dispone el programa

### OBJETIVOS PARTICULARES

- Presentar el lenguaje de programación R. Fuentes de descarga, de paquetes, instalación, instalación y uso de editores externos al R.
- Brindar fuentes de información resaltando la variedad de temas para los cuales se han desarrollado paquetes asociados al R para poder abarcarlos.
- Familiarizar a los participantes con la escritura, documentación y uso de los scripts.
- Desarrollar temas, ejemplos y dar ejercitación en temas variados.
- Presentar a los alumnos conceptos principales de programación en R, procedimientos de filtrado de información, resúmenes de datos, procesos iterativos, condicionales y creación de funciones propias.
- Dar ejemplos concretos de procesamiento de datos en R relacionados a temas de investigación en biología.

## CONTENIDOS MÍNIMOS

Instalación del programa y de librerías (módulos adicionales) a partir de los repositorios públicos. Instalación y uso de editores de texto para la escritura, edición y almacenamiento de código. Personalización del espacio de trabajo. Conceptos básicos del lenguaje de programación en R. Importación y exportación de datos. Tipos de objetos en R: generación y manipulación. Funciones integradas. Funciones personalizadas: creación y uso. Condicionales. Funciones estadísticas: uso y acceso a resultados parciales. Procesos iterativos. Funciones gráficas: tipos, diseño y formatos de gráficos, composición y exportación de figuras.

## ORGANIZACIÓN DE LOS CONTENIDOS - PROGRAMA ANALÍTICO

**Introducción al software gratuito R.** Definición de software gratuito y libre. Instalación y recomendaciones. Personalización del espacio de trabajo. Uso del editor del R. Instalación y personalización de un editor alternativo externo al R. Herramientas y funciones generales relacionadas a la sesión de R (identificación de directorio de trabajo, cerrar programa, etc.). Manejo de ventanas emergentes. Cargar y activar librerías. Ayudas en línea. Formas alternativas de acceder al espacio de trabajo. Definición de rutas relativas o absolutas

**Operaciones básicas.** ¿Cómo "funciona" el R? Re-tipeado automático. Escape (tecla Esc). Sensibilidad a las mayúsculas. Criterios para nombrar los objetos en R. Listar y remover objetos del espacio de trabajo. Cómo citar R y las librerías. Reciclado de código y de los objetos. Edición de los objetos interactivamente y por indexación.

**Tipos de objetos y su estructura.** Vectores, data.frame, matrices, listas, arrays. Tipos de vectores: numéricos, lógicos, texto, listas. Asignación de objetos. Valores especiales (no disponibles): NA, Inf, NaN. Factores. Generación de objetos mediante funciones específicas. Indexación por filas y/o columnas. Descripción básica y reconocimiento de atributos de objetos. Uso de expresiones lógicas como criterios de selección de elementos de un objeto. Filtrado: identificar elementos según criterio/s, uso de condiciones simultáneas o alternativas.

**Importar y exportar archivos y gráficos en R.** Uso de funciones específicas para la importación/exportación de archivos (e.j tablas) y figuras. Exploración de argumentos de las funciones de importación/exportación.

**Redistribuir y resumir información.** Resumir información (cálculo de promedio, desvío, etc) de acuerdo a múltiples criterios de clasificación. Extraer o dar otro formato a la información existente. Dividir o concatenar objetos (ej. vectores, tablas, matrices).

**Gráficos.** Funciones básicas para la generación de gráficos y sus argumentos, Funciones gráficas de alto y bajo nivel. Sistemas gráficos. Dispositivos gráficos (graphics devices). Librerías gráficas. Exportación de gráficos con formatos específicos (png, jpeg, etc). Especificación de parámetros gráficos (ejes, leyendas, colores, etc). Elaboración de figuras compuestas (varios paneles en una misma figura). Funciones gráficas interactivas. Gráficos 3D.

**Condicionales.** Tipos, estructura y aplicaciones.

**Procesos iterativos (loops).** Estructura y aplicaciones. Ejemplos de aplicaciones estadísticas.

**Funciones.** Estructura básica de las funciones. Funciones y argumentos. Crear y usar funciones. Argumentos y documentación de las funciones definidas por el usuario.

## **METODOLOGÍA DE TRABAJO. MODALIDAD DE CURSADA**

La modalidad es presencial. Los alumnos desarrollarán ejemplos ya resueltos a continuación de la presentación de cada tema. Periódicamente resolverán, en forma individual o grupal, ejercicios que involucren uno o más temas ya vistos.

El material del curso es exclusivamente digital y se entregará antes del inicio. Los archivos están organizados en carpetas de nombres sugerentes a los que se podrá acceder mediante la forma convencional del Explorador de Windows en un ambiente Windows (o el equivalente en otros sistemas) y, además, mediante enlaces a partir de una página Web local que se entregará como parte del material del curso.

Por correo electrónico, se enviarán a los participantes instrucciones para la instalación de software.

Las clases se proyectarán en pantalla mediante un cañón y se hará uso asiduo de pizarrón o pizarra. De estar disponible la conexión, también se usaría la Internet.

En la medida que el tiempo lo permita, se hará énfasis en aquellos temas o en aquellos ejemplos que se perciban sean de interés para la mayoría de los asistentes.

Como parte de la práctica, se destinará tiempo para que los alumnos importen al R sus propios datos e inicien / consulten sobre los siguientes pasos de manipulación y análisis.

Habrá un examen final que los alumnos deberán resolver individualmente en un período de 2 semanas. De esta forma tendrán tiempo de revisar todo el material e asentar conocimientos.

### **MATERIAL DIDÁCTICO**

- Scripts. Los distintos temas se presentan en una serie de “scripts”. En ellos se desarrolla el material conceptual e inmediatamente se presentan ejemplos con sus soluciones. Varios temas son expuestos en forma gráfica o esquemática a partir de archivos Powerpoint propios o a partir de contenidos provenientes de otras fuentes. Los scripts (y sus vínculos a material soporte con otros formatos) están diseñados para servir como material de estudio y referencia futura. Cada uno contiene un índice para la fácil ubicación de los temas.
- Ejercitación / práctica asociada a cada tema
- Referencia a distintas fuentes de información
- Bibliografía básica de referencia básica con formato pdf

### **CARGA HORARIA**

40 h: 20 % clases teóricas, 60% clases teórico-prácticas, 20% prácticas

### **CONDICIONES Y CONOCIMIENTOS PREVIOS**

Tener conocimientos básicos de computación: instalación de programas, carpetas y archivos. No se asume experiencia previa en programación.

## **CERTIFICACIÓN DEL CURSO**

Cada día de curso consta de dos módulos de 4 h cada uno, el de la mañana y el de la tarde. Se otorgarán certificados de asistencia cuando se esté presente a 8 de los 10 módulos y certificados de aprobación si, además, se aprueba un examen final (a entregar dos semanas después de terminado el curso).

## **BIBLIOGRAFÍA (selección reducida)**

Crawley, Michael J 2013. The R Book, Ed John Wiley and Sons Ltd, England

Salas, Christian. 2008. ¿Por qué comprar un programa estadístico si existe R?  
Ecología Austral 18: 223-231

The R Project for Statistical Computing. <http://www.r-project.org/>

Maindonald, J. H. (2008) .Using R for Data Analysis and Graphics. Introduction, code and commentary, pp. 96: Centre for Mathematics and Its Applications, Australian National University.