



AlwaysR, Módulo 2: Manejo de datos y figuras en R

Descripción

La mayor parte del tiempo de trabajo invertido en un proyecto se relaciona a la lectura, filtrado y preprocesamiento de la información de entrada. Así mismo, el manejo de datos y la generación de gráficos constituye una sección fundamental en todo proyecto y precisamente R dispone de una amplia variedad de herramientas que facilitan estas tareas y permiten exportar resultados en múltiples formatos.

El Módulo 2 del programa AlwaysR está orientado a afianzar el conocimiento en labores de filtrado de tablas de datos, extracción de subconjuntos de información, validación de variables y combinación de tablas, así como en la generación de figuras complejas en los dos entornos más populares dentro de R: **graphics** y **ggplot**.

Objetivos de aprendizaje

- Aprender los conceptos ligados al manejo avanzado de tablas de datos en R.
- Comprender y ejecutar las principales funciones para el manejo e interrelación de tablas de datos, tanto en entorno tradicional como en el ofrecido por las librerías **dplyr** y **tidyr**.
- Mejorar las habilidades del participante en la generación de gráficos exploratorios, incluyendo la capacidad de incluir múltiples elementos en una figura (líneas, puntos, polígonos), así como de generar gráficos de múltiples figuras.
- Introducir al participante en la generación de mapas simples y en la inclusión de variables ambientales dentro de un mapa.
- Introducir a la sintaxis y generación de figuras en el entorno **ggplot**.

Material

- Una computadora por participante con sistema operativo MS Windows, MacOS o alguna distribución de Linux con conexión a internet.
- Software: R, RStudio y Rtools (este último, solo en entorno MS Windows).

Prerrequisitos

Para este módulo, el participante debe conocer de antemano conceptos básicos de manejo en R, tales como clases, métodos, principales objetos (vectores, matrices, arreglos, *data frames*, listas); así como de conceptos, comandos y funciones para realizar indexación y ordenamiento básico de datos.

Contenido

El curso está dividido en cinco sesiones de 4 horas académicas cada una, haciendo un total de **20 horas**.

-
- Día 1 Funciones de agrupación y ejecución de comandos por grupos (familia apply). Uso de funciones de reestructuración de tablas de datos en paquete **base** y **dplyr**.
- Día 2 Funciones para fusión de tablas (*data merging*) y manejo de datos agrupados (*data grouping*) en paquete **base** y **dplyr**. Uso de bucles y condicionales para validación de tablas de datos.
- Día 3 Teoría: Definición de elementos en la interfaz base de gráficos en R. Descripción y manipulación de los principales parámetros gráficos (par).
Lab: Figuras básicas (puntos, líneas, barras, boxplot, histogramas) y cómo añadir elementos sobre figuras (puntos, líneas, polígonos, flechas).
- Día 4 Figuras múltiples (arreglo de figuras). Creación de figuras con enfoque espacial (mapas) en R.
- Día 5 Introducción al paquete **ggplot2** (sintaxis y principales funciones). Generación de figuras múltiples y manipulación de elementos en **ggplot2**.
-

Certificación y sistema de calificación

Al culminar su participación, cada participante recibirá un certificado de Asistencia o de Asistencia y aprobación, dependiendo del grado de cumplimiento de los objetivos del programa. Los detalles de los criterios y sistema de calificación están detallados en nuestro portal web (cousteau-group.com/tyc/) y para este módulo constará de una evaluación de las participaciones en las sesiones, así como de un examen al final del curso. Para la calificación final, se tomará en cuenta las siguientes proporciones:

15%	Participación durante el curso
30%	Asignaciones durante el curso
55%	Examen al final del curso

La calificación final se evaluará en una escala de 0-100, que luego se convertirá a una categoría de A (entre 90 y 100), B (entre 80 y 89,9), C (entre 70 y 79,9) o D (70 o menos). Para obtener un certificado de **Asistencia y aprobación**, el participante deberá obtener una calificación de **C o superior**.

Condiciones generales

Todo participante tiene la responsabilidad de leer y conocer los **Términos y condiciones** estipulados en el portal web de Cousteau Consultant Group visitando el siguiente enlace: cousteau-group.com/tyc/. Cada participante está en la obligación de respetar y cumplir las normas e indicaciones estipuladas. Cousteau Consultant Group se reserva el derecho de restringir la continuidad de un participante en las sesiones del curso si es que se comprueba una violación del código de conducta interno.