

## Introducción a *Tidyverse* para Ciencia de Datos en R Estación Lastarria – 2022

### Descripción

Este curso de 5 sesiones aborda conceptos de estadística y ciencias de datos usando el lenguaje de programación R y el megapaquete Tidyverse. Aprenderás a utilizar funciones para administrar, analizar y visualizar datos de forma reproducible y fácil de comunicar.

### Responsable

#### **Sara Acevedo**

Química de la Pontificia Universidad Católica de Chile, MSc. Soil and Biogeochemistry en UC Davis y estudiante Doctorado Ciencias de la Ingeniería en PUC.

Actualmente su investigación se centra en estudios de laboratorio y de campo relacionados con la ciencia del suelo y los estudios ambientales, específicamente en la Física del suelo. Además se especializa en el lenguaje R enfocado en estadísticas y modelos predictivos. Cuenta con experiencia impartiendo talleres en este tema.

### Horario y realización

Las sesiones se realizarán los días miércoles de 19:00 a 21:00 horas (Chile) a partir del miércoles 15 de junio. Las sesiones se llevarán a cabo vía Google Meet, serán grabadas y posteriormente entregadas a las personas participantes vía Drive y YouTube para ser consultadas en el momento que requieran.

### **Público objetivo**

Personas que estudien o ejerzan profesionalmente en campos de las Ciencias Sociales y las Humanidades, interesados en el perfeccionamiento de técnicas computacionales para el análisis de información social.

### **Objetivos de aprendizaje**

Usando dataframes de distintos tipos (data de ecología y ciencias sociales), el curso se enfoca en:

- Manejar de forma general las herramientas más usadas en análisis exploratorio de datos
- Construir datasets informativos y reproducibles con datos limpios
- Comunicar las hipótesis generadas por los datos por medio de visualizaciones y reportes

### **Requisitos mínimos**

- Ganas de aprender y tolerancia a la frustración
- Inglés básico-intermedio
- Computador con RStudio y R instalado (pedir instructivo si no lo tienes instalado).
- Conexión estable a internet

### **Metodología**

- Clases expositivas
- Entrega de material en RMarkdown
- Ejercicios a resolver en RMarkdown
- Clases de 2 horas cronológicas de duración



ESTACIÓN  
LASTARRIA

### Contenidos por sesión

Sesión	Contenidos	Fecha
1. Comprender los elementos básicos del lenguaje de programación en R (1/2)	<ul style="list-style-type: none"><li>● Importancia de la reproducibilidad de los datos e información</li><li>● Aprender buenas prácticas</li><li>● Familiarizarse con el IDE RStudio</li><li>● Conocer operadores y funciones (1/2)</li></ul>	15-06-2022
2. Comprender los elementos básicos del lenguaje de programación R (2/2)	<ul style="list-style-type: none"><li>● Conocer operadores y funciones (2/2)</li><li>● Tipos de datos</li><li>● Importar/exportar datos desde .xlsx y .csv</li><li>● Test estadísticos básicos</li></ul>	22-06-2022
3. Paquetes <i>dplyr</i> y <i>tidyr</i> para limpieza de datos	<ul style="list-style-type: none"><li>● Funciones básicas</li><li>● Formato de datos</li><li>● Ejercicios data wrangling</li></ul>	29-06-2022
4. Paquete <i>ggplot2</i> para visualización de datos	<ul style="list-style-type: none"><li>● Componentes ggplot2</li><li>● geom básicos</li><li>● Estilos y colores.</li></ul>	06-07-2022
5. Paquete <i>purrr</i> para bucles	<ul style="list-style-type: none"><li>● Visualización de información cualitativa.</li><li>● Redes y visualización del contenido.</li><li>● Exportar a XLS, SPSS y CSV.</li></ul>	13-07-2022