

# Título:

Introducción a R y RStudio para la investigación social: aplicaciones de análisis de datos y visualización.

## Características:

Virtual Sincrónico, 8 horas, 4 clases semanales de 2 horas.

## Días y Horarios:

Jueves 6, 13, 20 y 27 de agosto de 2026, de 19 a 21 hs.

## Breve Fundamentación:

La aplicación y uso de estadísticas y el análisis de datos en las ciencias sociales requiere en forma creciente el manejo y la aplicación de lenguajes de programación y entornos estadísticos dinámicos para el procesamiento eficiente y transparente de la información.

Entre las herramientas existentes, R es un entorno de software libre líder en el análisis de datos, especialmente valorado por profesionales de las ciencias sociales debido a su versatilidad, su activa comunidad académica y su capacidad para realizar análisis reproducibles y visualizaciones de alta calidad. En combinación con la interfaz visual de RStudio, ofrece una plataforma potente que permite una amplia aplicación para usos académicos, investigación aplicada, estudios de opinión pública, medios de comunicación y consultoría profesional.

Se espera que, a la finalización del curso, los cursantes sean capaces de manipular bases de datos, importar y limpiar registros de información social, generar tablas de contingencia y estadísticos descriptivos, construir gráficos bivariados bajo la gramática de gráficos (ggplot2), y al mismo tiempo dar una adecuada interpretación sociológica de los resultados obtenidos con código limpio y reproducible.

## Objetivos generales:

- Que el cursante comprenda y aplique aspectos fundamentales del manejo, del análisis de datos estadísticos y del uso de bases de datos mediante el entorno y lenguaje de programación estadística R y RStudio.
- Que el cursante logre mediante el uso del software y paquetes específicos (ecosistema *tidyverse*), obtener estadísticas descriptivas, realizar cruces de variables, diseñar representaciones gráficas de nivel profesional, y comprenda las diversas potencialidades del análisis reproducible en ciencias sociales.
- Que el cursante conozca herramientas de inteligencia artificial que pueden ayudar a automatizar tareas, documentar procesos y depurar scripts de programación en R.

# Contenidos:

El curso está organizado en 4 clases, de 2 horas cada una, en modalidad virtual sincrónica.

Se ponen a disposición del cursante bases de datos de carácter social reales o simuladas (por ejemplo, bases de estadísticas vitales u otras bases de acceso público) a modo de ejemplo y para su uso didáctico en los prácticos guiados.

Se espera que parte de los materiales, códigos de R y análisis generados puedan ser utilizados para investigaciones o trabajos en curso. Los cursantes pueden proponer y poner a disposición bases de datos propias de uso actual en sus investigaciones en curso para resolver las dinámicas prácticas.

## Clase 1:

- Entorno e interfaz de RStudio, consola, paneles de control y archivos de script (.R).
- Concepto de análisis reproducible y buenas prácticas en la estructuración de proyectos de trabajo.
- Creación y manejo de variables, objetos y funciones. Tipos y clases de datos (numéricos, caracteres, lógicos y factores).
- Estructuras de datos fundamentales en R: vectores y hojas de datos (*data frames*).
- Organización de los datos, operadores lógicos, aritméticos y relacionales simples.
- Procedimientos iniciales para el filtrado, manipulación básica y selección de elementos en vectores y data frames.
- Primeros pasos en estadística descriptiva: obtención de medidas de tendencia central y dispersión mediante funciones base.

## Clase 2:

- Obtención de frecuencias y porcentajes a partir de la importación y lectura de datos en formatos externos comunes (archivos CSV, Excel y paquetes estadísticos mediante librerías como *readr*, *readxl* y *haven*).
- Introducción al ecosistema *tidyverse* para manipulación estructurada de bases de datos de ciencias sociales.
- Uso de comandos específicos del paquete *dplyr* para la transformación de bases de datos: selección de variables (*select*), filtrado de registros (*filter*) y ordenamiento de filas (*arrange*).
- Creación y modificación de variables cuantitativas y cualitativas mediante funciones de cálculo (*mutate*).
- Procesamientos específicos para la limpieza de datos: manejo de valores perdidos (valores NA) y recodificación de factores de nivel nominal u ordinal.
- Criterios metodológicos para la depuración de bases de datos antes de iniciar análisis estadísticos.
- Uso estratégico de Inteligencia Artificial para la corrección de errores, depuración de código y optimización de scripts en R.

### Clase 3:

- Técnicas avanzadas de agregación y resumen de datos: agrupación por variables de control, segmentación y estadísticas grupales utilizando *group\_by* y *summarize*.
- Obtención de tablas de contingencia bivariadas y análisis de relaciones entre variables cualitativas (porcentajes de fila, columna y totales).
- Análisis exploratorio bivariado: comparación de promedios grupales a partir de variables sociodemográficas cruzadas.
- Análisis de asociación cuantitativa para variables de escala continua.
- Interpretación sociológica y metodológica de tablas descriptivas cruzadas e índices de asociación.

### Clase 4:

- Introducción a la gramática de gráficos y principios del diseño visual de datos utilizando el paquete *ggplot2*.
- Construcción, personalización y edición estética de gráficos descriptivos esenciales: gráficos de barras para variables cualitativas, histogramas, diagramas de caja (*boxplots*) y gráficos de dispersión (*scatterplots*).
- Exportación estructurada de resultados: almacenamiento automatizado de tablas descriptivas procesadas y exportación de visualizaciones en formatos de alta calidad para informes académicos o profesionales.
- Introducción conceptual a la generación de reportes dinámicos utilizando R Markdown para integrar código y texto.

### Instructor:

**Eduardo Marcelo Pietrafesa.** Sociólogo, especialista en análisis cuantitativo, procesamiento de datos y el uso de tecnologías de información y lenguajes de programación estadística (R y RStudio) aplicados a las ciencias sociales.