

**CURSO**

# **Introducción al análisis de datos con R**

---

**Mayo 2023**

## Presentación

En la actualidad, cada vez son más las áreas de las ciencias sociales en las que el análisis de datos contribuye al abordaje de preguntas de investigación y al estudio de problemáticas de índole social, tanto desde una perspectiva exploratoria, como explicativa.

En este sentido, existen distintas técnicas y herramientas posibles de ser empleadas en el análisis de datos sociales. Sin embargo, el presente curso se centra en el uso del lenguaje R a través del entorno de desarrollo Rstudio.

## Destinatarios

El curso se orienta a diversos perfiles que busquen incorporar conocimientos en las técnicas, herramientas y conceptos del análisis de datos mediante el lenguaje de programación R. Dado que se trabajará con datos vinculados a ámbitos de las ciencias sociales, es posible que resulte más atractivo para aquellas personas que se desempeñen académica o profesionalmente en dichas áreas; aunque la propuesta se encuentra abierta a profesionales y/o perfiles de todos los ámbitos.

## Módulo I: Introducción al análisis de datos con R

### Objetivos

Al finalizar el primer módulo del curso, se espera que los participantes sean capaces de:

- Trabajar de manera fluída con R y Rstudio.
- Utilizar herramientas de desarrollo colaborativo en la nube para proyectos de análisis de datos.
- Identificar fuentes de datos, extraer, transformar y desarrollar data para el abordaje de preguntas o problemas de índole social .
- Desarrollar un análisis exploratorio de los datos mediante técnicas estadísticas y herramientas introductorias de visualización.

## Requisitos

Este Módulo del curso no requiere conocimientos previos en programación, ni experiencia previa en R.

## Cursada

**Fecha de inicio:** Jueves 18 de mayo

**Cantidad de clases:** 6 clases (9 horas totales)

**Modalidad de cursada:** sincrónica virtual

**Días y horarios de cursada:** jueves de 19 a 20:30h

**Precio del Módulo I:** \$ 8.500

## Contenidos del Módulo I

### **Clase 1: Introducción a la Ciencia de Datos - Programación en R**

Introducción al análisis de datos mediante R. Explicación de software libre, lenguaje de programación e IDLE. Instalación de R y RStudio. Descripción de un entorno de trabajo. Instalación e introducción a Github con el objetivo de poder crear un repositorio del trabajo en clase. Primeros pasos en la herramienta. Instalación de paquetes y operadores lógicos.

### **Clase 2: Procesos ETL: Extract, Transform and Load**

Operaciones básicas en R. Identificación de elementos: vector, lista, dataframe, matriz, funciones. Importación y exportación de bases de datos (.csv, .xlsx, .sav, entre otros). Desarrollo de procesos de extracción, limpieza y transformación de datos con Tidyverse.

### **Clase 3: Análisis exploratorio I**

Profundización de técnicas de limpieza y transformación de datos con Tidyverse. Exploración de datos mediante técnicas estadísticas para conocer la estructura del dataset, casos atípicos, y casos perdidos o nulos. Introducción a la visualización para el análisis exploratorio (visualizaciones estáticas).

#### **Clase 4: Análisis exploratorio II**

Profundización de técnicas de limpieza y transformación de datos con Tidyverse. Exploración de datos mediante técnicas estadísticas para conocer la estructura del dataset, casos atípicos, y casos perdidos o nulos. Introducción a la visualización para el análisis exploratorio (visualizaciones estáticas).

#### **Clase 5 y 6: Data viz. Técnicas de visualización en R.**

Desarrollo de recursos y técnicas de visualización de datos. Visualizaciones en capas y customización del gráfico. Desarrollo de gráficos interactivos en HTML. Paquetes específicos para distintos tipos de datos y casos de uso: plotly, ganimate, highcharter. Quarto: Comunicación y visualización de resultados en formato interactivo.

## **Aprobación**

Para la aprobación se exigirá la entrega de un trabajo final en el que se apliquen las técnicas desarrolladas durante el curso. Los participantes contarán con 2 semanas para realizar el trabajo, una vez finalizada la cursada.

El trabajo deberá partir de una pregunta de investigación o situarse en un área de interés a ser indagada mediante el análisis de datos. Se espera que los participantes exploren, manipulen y transformen los datos, desarrollen visualizaciones y elaboren conclusiones sobre la información procesada.

**Fecha de presentación del trabajo final:** 15 de julio