

Datos Solidarios con R: Introducción

Análisis y visualización de datos para el sector solidario en Colombia

Este minicurso es una breve introducción al análisis de datos usando lenguaje R, direccionado para jóvenes vinculados al sector solidario en Colombia. Mediante tutoriales temáticos, este curso guiará los participantes en la realización de las siguientes actividades:

1. Como crear un script (códigos)
2. Cargar Paquetes
3. Abrir una base de datos
4. Limpiar y procesar los datos
5. Calcular estadísticas
6. Generar gráficos

Principios Cruciales en el Análisis de Datos

El presente curso enfatizará los siguientes principios que ayudan a mantener la integridad y confiabilidad de nuestro análisis:

Simplicidad: Escribimos el código de la manera más simple y clara posible para que sea fácil de leer y corregir.

Reproducibilidad: Las bases de datos y el script de análisis pueden compartirse con un colega que pueda ejecutar el análisis y obtener el mismo resultado.

Transparencia: Todo el análisis puede inspeccionarse para rastrear el origen de cada resultado e identificar errores.

Comprensión de los datos: El análisis debe ayudarnos a comprender, no a ocultar, el contenido de nuestros datos.

¿Qué es R?

R es un lenguaje de programación y entorno de software gratuito y de código abierto, diseñado específicamente para el cálculo estadístico y la creación de gráficos. Es una potente herramienta para el análisis, la manipulación y la visualización de datos.

El lenguaje R suele ser el vehículo predilecto para la investigación en metodología estadística, y R ofrece una vía de código abierto para participar en dicha actividad.

Interfase del R Studio

R Studio es un entorno de desarrollo integrado (Integrated Development Environment -IDE), es decir, es una aplicación de software que proporciona facilidades completas para el desarrollo de software usando el lenguaje R. Así, R Studio una interfaz amigable permite la edición de códigos, el análisis de datos y la creación de herramientas, como páginas webs, de forma más intuitiva.

A continuación, veremos una imagen que ilustra los componentes principales de esta interfase.

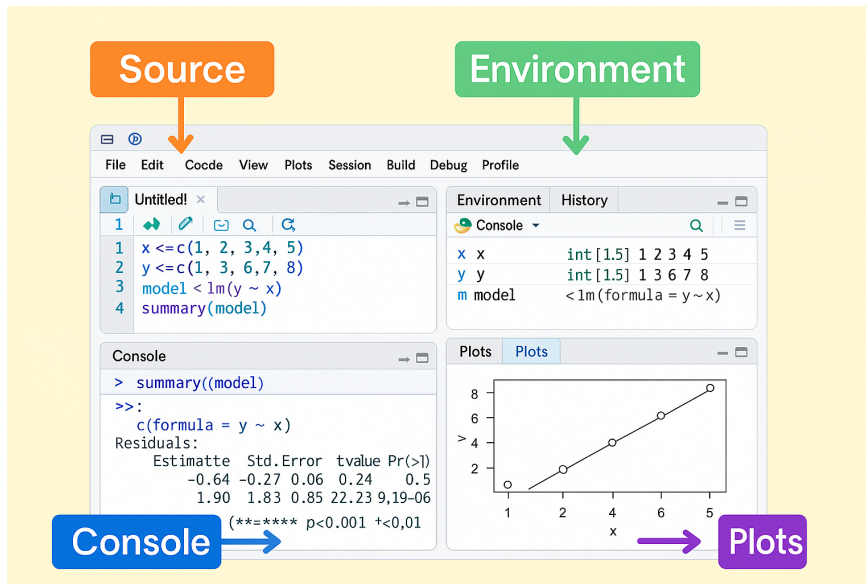


Figure 1: Fuente: Elaboración Propia

R Studio está compuesto por 4 paneles principales (como se muestra en la imagen anterior), aunque su posición puede variar a criterio de los usuarios. A continuación, se especifican los paneles y su función:

1. Panel de Fuente (Source): Es el espacio para escribir y editar código en un archivo de script o para cargar uno previamente guardado.
2. Panel de Consola (Console): Aquí se ejecutan los comandos que se escriben, recibiendo la respuesta del programa.
3. Panel de Entorno (Environment): Este panel central es donde se muestran las variables, objetos y datos cargados, junto con pestañas para el historial de comandos, conexiones y sistema de control de versiones (VCS).
4. Panel de Archivos (Files/ Plots): Permite ver los archivos en el disco, explorar la estructura de directorios, buscar y abrir archivos, y visualizar gráficos.

Próximos pasos

En el próximo tutorial aprenderemos:

- Cómo crear nuestro primer script,
- Las funcionalidades básicas de R,
- Qué son paquetes cómo usarlos, y
- Cómo podemos leer bases de datos externas en R Studio.

Licencia Material educativo diseñado para el curso **Datos Solidarios con R**.
© 2025 — Yulieth Martínez Villalba.