

Introducción al análisis de datos con R

Conceptos básicos de R y RStudio

Ejercicios - Sesión 1

5 - 9 junio de 2023

Introducción

En los ejercicios de la primera sesión vamos a trabajar sobre cómo realizar un primer acercamiento exploratorio a los datos con R. Antes de empezar:

- Abre un nuevo script y limpia el espacio de datos ejecutando `rm(list = ls())`. Este comando eliminará todos los datos (objetos) del espacio de trabajo, evitando posibles confusiones.
- Carga los paquetes que necesitas para realizar el ejercicio, ejecutando las líneas de `library()`. En caso de que alguno de ellos no esté instalado, instalalo utilizando `install.packages("package")`. Fundamentalmente, en este ejercicio se utilizará la librería **tidyverse**.
- Establece el directorio de trabajo donde se encuentran los datos.
- Carga el archivo de datos en formato CSV utilizando `read_csv()` y asígnalo a un objeto llamado **data**.

Observando la estructura de los datos

- Explora la base de datos que acabas de cargar:
 - Muestra las primeras 6 filas utilizando `head()`.
 - Explora la estructura de los datos utilizando `glimpse()`. ¿Cuántas filas tiene? ¿Cuántas columnas tiene? ¿De qué tipo son las variables?
 - Verifica si existen valores perdidos en alguna columna utilizando `colSums(is.na())`.

Análisis exploratorio de los datos

- Calcula la media de años de educación que tienen las madres utilizando `mean()`.
- Calcula la mediana de los ingresos del hogar utilizando `median()`.
- Calcula la media de años de educación de la madre para cada itinerario educativo cursado por los alumnos. Utiliza `aggregate()`.
- Calcula la media de ingresos del hogar para cada itinerario educativo cursado por los alumnos. Utiliza `aggregate()`.
- Realiza un análisis para mostrar, en términos porcentuales, si la elección de los itinerarios educativos (school) difiere según el estado civil de la madre (marital). Utiliza `tabyl()`.