

SOC4001 Procesamiento avanzado de bases de datos en R

Tarea 3

Ponderación: 12% de la nota final del curso

Entrega: Desde el momento de entrega, los estudiantes tienen plazo hasta el domingo 1 de Noviembre a las 23:59pm para completar esta tarea.

Formato: Desarrollar esta tarea en un RScript, agregando comentarios cuando sea necesario.

- 1) Carga la base de datos “Chile” del paquete `carData` y crea un objeto que los contenga los datos. Llama tal objeto “datos_chile”. Carga la librería `tidyverse` y ejecuta la siguientes operaciones usando las herramientas contenidas de `tidyverse`:
- 2) Añade a “datos_chile” un variable llamada “year” con valor 1988 en todas las filas
- 3) Calcula el año de nacimiento de cada individuo. Añade a “datos_chile” un variable llamada “birthyear” que contenga esta información
- 4) Usando la función `case_when()` añade a “datos_chile” un variable llamada “cohort73” que tome valor 1 si la persona tenía 15 año o más el año del golpe de estado (1973) y valor 0 si tenía menos de 15. Trata las observaciones que no cumplan ninguna de estas condiciones como valores perdidos.
- 5) Usando la función `group_by()` añade a “datos_chile” un variable llamada “income_region_educ” que contenga el promedio de la variable “income” por región y nivel educacional.