

# SOC4001 Procesamiento avanzado de bases de datos en R

## Tarea 2

Ponderación: 12% de la nota final del curso

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 `if_else()` añade a “datos\_chile” un variable llamada “vote\_no” que tome valor 1 si la persona declara que votará por el No y valor 0 en cualquier otra caso.
- 5) Usando la función `case_when()` añade a “datos\_chile” un variable llamada “cohort73” que tome valor 1 si la persona tenía 18 años o más el año del golpe de estado (1973) y valor 0 si tenía menos de 18. Trata las observaciones que no cumplan ninguna de estas condiciones como valores perdidos.
- 6) Usando la función `group_by()` añade a “datos\_chile” un variable llamada “no\_by\_groups” que contenga el promedio de la variable “vote\_no” por región, nivel educacional y cohorte (cohort73).
- 7) Usando la funciones `summarise()` y `group_by()` calcula el promedio de la variable “vote\_no” por región, nivel educacional y cohorte (cohort73) y almacénalo en una variable llamada “datos\_chile”. Asigna el resultado a un nuevo objeto llamado `resultados`.
- 8) Usando la funciones `summarise()`, `across()` y `group_by()` calcula el promedio y la desviación estándar de las variables “vote\_no” e “income” por región, nivel educacional y cohorte (cohort73). Asigna el resultado a un nuevo objeto llamado `resultados`.