

# Taller de R: Estadística y Programación

## Taller 3

24/11/2023

En este taller se evalúan los temas vistos en las clases 9 y 10 del curso. Lea atentamente las instrucciones del taller.

### Instrucciones

- Este taller representa el **33.3%** de la nota total del curso y podrá responderlo de manera individual o en grupo de hasta 3 personas. En las primeras líneas del script debe escribir su nombre, código y la versión de R sobre la que está trabajando. Además, al inicio del código debe llamar/installar las librerías que va a usar en la sesión. Por ejemplo: `pacman`, `rio`, `data.table` y `tidyverse` (a lo menos).
- Asegúrate de descargar las bases de datos del repositorio <https://github.com/taller-r-202302/problem-sets> y crear un nuevo repositorio en tu cuenta de GitHub. Si va a trabajar en grupo, solo 1 de los integrantes del grupo debe crear el repositorio y compartir el acceso a los demás integrantes. El repositorio debe ser público para que se pueda acceder desde cualquier cuenta de GitHub. Adicionalmente, este repositorio debe incluir a lo menos tres carpetas: `input` (datos originales), `output` (datos procesados) y `code` (script con la respuesta del taller).
- Todos los integrantes del grupo deberán colgar el link al repositorio de GitHub en la actividad **Problem-set-3** de Bloque Neón antes de las 23:59 horas del 8 de diciembre de 2023.
- Por favor sea lo más organizado posible y comente paso a paso cada línea de código, pero recuerden **NO** usar ningún acento o carácter especial dentro del código para evitar problemas al abrir los scripts en los diferentes sistemas operativos.
- No seguir las instrucciones tiene una penalización del **20%** de la nota final.

### Problem set 3

Deberá usar un loop o una función para importar los archivos de la Gran Encuesta Integrada de Hogares que se encuentran en la carpeta `input`.

#### 1. Bucle (50%)

- **1.1 Lista de archivos en input**

Cree un objeto que almacene el vector de nombres de los archivos dentro de la carpeta `input`. Asegúrese de que cada archivo contenga la ruta con la ubicación de cada archivo.

**Hint:** Para este punto, puede usar la función `list.files()` con el argumento `recursive = T`.

- **1.2 Importar archivos: Resto - Características generales (Personas)**

Usa el objeto creado en el punto anterior como insumo de una función que te permita importar los archivos **Resto - Características generales (Personas)** para todos los meses y años. Asegúrate de agregar a cada dataframe una variable que te permita identificar el año al que pertenece el archivo.

**Hint:** Para este punto, puedes crear una función que importe un archivo y combinarla con la función `lapply`. O puedes crear un bucle y almacenar los dataframes en un objeto tipo lista. Para añadir a cada dataframe el año al que pertenece el archivo, puedes agregar a cada dataframe una variable con la ruta del archivo y después extraer el año de esa variable.

- **1.3 Combinar conjuntos de datos**

Combina todos los `data.frame` que importaste en el punto anterior en un solo `data.frame` llamado `cg`.

**Hint:** Para este punto, puede usar la función `rbindlist` de la librería `data.table`.

## **2. Visualizaciones (50%)**

Usa la librería `ggplot2` para generar al menos dos visualizaciones de las variables que desees visualizar. Cada visualización debe incluir un título, un tema y etiquetas para los ejes `x` e `y`.