

Ejercicio práctico: Introducción a R Commander

Sesión 1. Curso: ‘Análisis estadístico aplicado con R Commander’

Ejercicio 1

Se proporcionan dos bases de datos denominadas `predimed.xlsx` y `predimed.sav` (recuperadas desde el paquete de R `compareGroups`), las cuales están relacionadas con el estudio PREDIMED.

De forma resumida, el estudio PREDIMED fue un ensayo clínico multicéntrico donde los participantes fueron asignados aleatoriamente (...) a tres grupos de intervención (control con dieta mediterránea, dieta mediterránea con suplemento de frutos secos, y dieta mediterránea con suplemento de aceite de oliva), para observar el desarrollo de eventos cardiovasculares durante el tiempo de seguimiento (más de 7 años).

El contenido de ambas bases de datos es:

- `predimed.xlsx`:
 - `id`: identificador único de cada sujeto.
 - `group`: grupo de exposición, variable categórica con tres categorías (grupo control - `Control`, suplemento de frutos secos - `MedDiet + Nuts` y suplemento de aceite de oliva - `MedDiet + V00`).
 - `sex`: sexo, variable categórica con dos categorías (`Male` y `Female`).
 - `age`: edad, variable numérica con la edad en años.
 - `smoke`: tabaquismo, variable categórica con tres categorías (`Never`, `Current` y `Former`).
 - `htn`: hipertensión, variable categórica con dos categorías (`No` y `Yes`).
 - `diab`: diabetes tipo II, variable categórica con dos categorías (`No` y `Yes`).
 - `hyperchol`: dislipemia, variable categórica con dos categorías (`No` y `Yes`).
 - `famhist`: historia familiar de enfermedad coronaria, variable categórica con dos categorías (`No` y `Yes`).
 - `hormo`: terapia hormonal sustitutiva, variable categórica con dos categorías (`No` y `Yes`).
 - `toevent`: tiempo de seguimiento (años) hasta el evento cardiovascular, variable numérica.
 - `event`: desarrollo de enfermedad coronaria, variable categórica con dos categorías (`No` y `Yes`).
- `predimed.sav`
 - `id`: identificador único de cada sujeto.
 - `bmi`: índice de masa corporal, variable numérica ($\frac{Kg}{m^2}$).
 - `waist`: circunferencia de la cintura en cm, variable numérica.
 - `wth`: razón cintura/altura, variable numérica.
 - `p14`: puntuación de la escala de adherencia a dieta mediterránea, variable numérica.

Utilizando R Commander se ha de realizar:

1. Importar ambas bases de datos.
2. Unir ambas bases de datos empleando la variable `id`.
3. Recodificar el IMC numérico en cuatro categorías:
 - `bajopeso`: cuando $IMC < 18$.
 - `normopeso`: cuando $18 \leq IMC < 25$.
 - `sobrepeso`: cuando $25 \leq IMC < 30$.
 - `obesidad`: cuando $IMC \geq 30$.
4. Calcular la altura y el peso de cada sujeto con la información disponible.