



TRATAMIENTO DE CONJUNTOS DE DATOS Y PROGRAMACIÓN ESTADÍSTICA CON R

Docentes:

Isabel Padilla Sánchez
Joaquin Planelles Romero

Objetivos:

Conocer la existencia y utilidad de varias librerías de R especialmente diseñadas para tratar conjuntos de datos.

Conocer los principales comandos y saber utilizar las funciones que tienen estas librerías.

Saber programar rutinas de R, que nos permitan filtrar casos, crear nuevas variables, conectar a una base de datos externa y otras muchas operaciones habituales en el tratamiento de conjuntos de datos.

Programa del curso:

1. Preliminares

- En el curso se trabajará con R, utilizando el entorno RStudio
- El curso se centrará en el uso de tidyverse. Se trata de un conjunto de librerías de R especialmente diseñadas para tratamiento de datos.

2. tidyverse

Introducción a tidyverse: librerías y gramática. El núcleo de tidyverse y otras librerías de tidyverse.

3. Librería dplyr.

En esta unidad aprenderemos una sintaxis sencilla para filtrar casos, crear nuevas variables, cruzar tablas y otras operaciones habituales en el trabajo que desarrollamos en Estadística Pública.

3.1 Manipulación de columnas. Comandos `select()` y `mutate()`.

3.2 Manipulación de filas. Comandos `filter()` y `arrange()`.

3.3 Funciones de sumarización. Comandos `group_by()` y `summarise()`.

3.4 Combinar tablas. Comandos `bind_cols()`, `bind_rows()` y familia de comandos `join_`.

4. Librería tidyr

La librería tidyr facilita un conjunto de funciones que te ayudan a organizar/ordenar tus datos.

4.1 Tablas dinámicas. Comandos `pivot_longer()` y `pivot_wider()`.

4.2 Otras funciones de tidyr. Comandos `separate()` y `expand()`.



Junta de Andalucía

5. Librería lubridate

La librería lubridate es una librería de tidyverse, especialmente diseñada para trabajar con fechas y duraciones.

5.1 Tratamiento de fechas con el paquete base de R.

5.2 Lectura de fechas con lubridate.

5.3 Aritmética de fechas con lubridate.

6. Librería dbplyr

La librería dbplyr funciona como un traductor a SQL. Esto te permite, por ejemplo, conectarte desde R a una base de datos externa y pedirle que realice determinadas acciones mediante sentencias SQL.

6.1 R y la memoria volátil de tu equipo.

6.2 Conexión a bases de datos desde R.

6.3 Lanzar consultas desde R a la base de datos.

7. Otra alternativa para tratar datos desde R. Librería data.table

7.1 Introducción a data.table.

7.2 Sintaxis data.table.

7.3 Librería dtplyr.

Observaciones:

Tipo: Semipresencial (Teleformación)

Duración: 40 horas (teóricas y prácticas)

Fechas: Del 19 de mayo al 24 de junio

Horario: Libre (Excepto sesión/es presencial/es que se comunicaran previamente y se llevarán a cabo Telemáticamente)