

Curso de Introducción a R (Opción A)

Pablo Cabrera Álvarez

17/6/2019

Objetivos

Este curso tiene como objetivo **introducir la dinámica de trabajo en R** para investigadoras que previamente usaban programas de tipo GUI, como SPSS o Stata. Este curso está pensado para hacer una transición hacia trabajar con código y escribir *scripts* básicos en R. Al final de este curso se conocerá el funcionamiento básico de R con el que se podrá hacer manipulaciones sencillas de datos y tabulaciones.

Estructura

El curso consta de **cinco módulos básicos** (MB) en los que se expondrán los principios generales del lenguaje R. Existe la posibilidad de ampliar los módulos básicos con otros optativos (MO) de carácter más específico sobre temas como la visualización de datos o el trabajo con datos de encuestas. Cada módulo tiene una **duración de 4 horas** e incluyen aproximadamente una hora y media de exposición y dos horas de práctica.

Introducción: si yo vivía bien con SPSS... (MB)

Paquetes: base

Datos: mtcars

- ¿Por qué R?
- Cómo funciona R
- R-Studio o cómo evitar que la gente salga corriendo
- Cómo evitar la frustración: pedir ayuda `help()`, `?` y otras fuentes
- El lenguaje orientado a objetos `IMOP <- c("Ana", "Sara")`
- ¿Dónde se almacenan los objetos? El *workspace* en R
- ¿Qué tipos de objetos hay en R?

Práctica: Generación e interacción con los diferentes tipos de objetos en R. Búsqueda de ayuda y lectura de la documentación.

Operaciones básicas o cómo sobrevivir a una pelea con R (MB)

Paquetes: base, forcats, lubridate, hms, stringr

Datos: mtcars, titanic

- El poder de transformar objetos `class()` y `as.*()`
- Además de sumar, ¿qué puedo hacer en R? Los operadores
- Sí, pero en las encuestas hay códigos y etiquetas: factores `factor()`, `levels()`, `ordered()` y `fct_relevel()`
- Cadenas de texto `str_locate()`, `str_remove()`, `str_replace()` y `str_sub()`
- Fechas `dmy_hms()` y `parse_date_time()`
- Valores perdidos NA

Práctica: Trabajar con diferentes tipos de vectores y tratar los valores perdidos

Manipulación de datos (lo básico, pero en R) (MB)

Paquetes: base, haven, readr, dplyr, tidyr

Datos: mtcars, iris, Barómetro CIS

- ¿Cómo abro los datos en R? `read_csv()`, `read_spss()`, `read_rds()` y `load()`
- Echar un vistazo a los datos: `View()`, `head()`, `tail()`, `str()` y `colnames()`
- Ordenar con `arrange()`
- Seleccionar variables `select()`
- Filtrar casos con `filter()`
- ¿Cómo se llama aquí el *compute*? `mutate()`
- Recodificar con `recode()`

Práctica: Manipular un conjunto de datos de encuesta.

Manipulación de datos (algo más que lo básico) (MB)

Paquetes: base, dplyr, tidyr

Datos: Base de datos electorales, Barómetro CIS, iris

- Un toque de elegancia: el uso de *pipes* `%>%`
- Agrupar `group_by()`
- Resumir `summarise()`
- Combinar bases de datos: filas `rbind()` y `bind_rows()`
- Combinar bases de datos: columnas `cbind()`, `bind_cols()`, `left_join()` y `right_join()`
- Cambiar de formato largo a ancho `spread()`
- Cambiar de formato ancho a largo `gather()`

Práctica: Manipular varios conjuntos de datos para crear una base de datos conjunta.

¿Y todo esto para hacer una tabla? (MB)

Paquetes: base, dplyr, tidyr, readr, haven, sjmisc, expss

Datos: Barómetro CIS

- Tablas de frecuencias `frq()`
- Estadísticos descriptivos `descr()`
- Tablas de contingencia `crosstab()`
- tablas para preguntas de respuesta múltiple `tab_cells()`
- Tablas personalizadas `tab_*`
- Exportar resultados `openxlsx`

Práctica: Crear tabulaciones para análisis y tablas para ser exportadas.

Metodología

Cada sesión tendrá una duración aproximada de cuatro horas y serán de carácter práctico. Las explicaciones irán mostrando código que se podrá ir ejecutando en paralelo. Cada sesión tendrá asociada una práctica que consistirá en utilizar las funciones en casos prácticos, un *script* comentado para facilitar la resolución de los ejercicios y otro *script* con las soluciones.