

Resumen del módulo Introducción a R

Diplomado Data Science UC 2022

Sección Johnson

Natalie Julian

Link para descargar R: <https://cran.r-project.org/>

Link para descargar RStudio: <https://www.rstudio.com/products/rstudio/download/#download>

Link para usar Rstudio Cloud: <https://rstudio.cloud/>

Atajos en el teclado

Ctrl + L	Cmd + L	Limpia la consola
Ctrl + Shift + N	Cmd + Shift + N	Crea un nuevo script
Ctrl + Enter	Cmd + Enter	Corre la línea de código actual
Ctrl + Shift + Enter	Cmd + Shift + Enter	Corre todo el código
Ctrl + S	Cmd + S	Guarda el script
Ctrl + Shift + C	Cmd + Shift + C	Convierte en comentario las líneas de código seleccionadas
Ctrl + Shift + R	Cmd + Shift + R	Crea secciones nuestro código
Alt + -	Option + -	Crea la flechita <-
Ctrl + Shift + M	Cmd + Shift + M	Crea el operador pipe %>%

Manejo del Workspace

getwd()	Retorna el directorio (carpeta) en la que R está trabajando actualmente
setwd()	Setea tu directorio en otra ruta
File>>New Project	Crea un nuevo proyecto
dir.create()	Crea una nueva carpeta en el proyecto
sys.date()	Entrega la fecha actual del sistema
remove()	Elimina objetos creados
install.packages()	Instala paquetes/librerías
library()	Carga paquetes/librerías a la sesión

Funciones básicas

print()	Printea/Muestra el input en la consola
c()	Para crear vectores
log(), sqrt(), exp(), pi	Funciones matemáticas básicas
min(), max(), mean(), median(), sum()	Funciones para el análisis descriptivo
class()	Indica la naturaleza del objeto
args()	Indica los argumentos de la función evaluada
paste()	Concatena dos objetos
summary()	Entrega un resumen descriptivo del objeto

Vectores

<code>seq()</code>	Crea secuencias de números en un rango
<code>rep()</code>	Crea un vector de repeticiones con cierto patrón
<code>is.vector()</code> , <code>is.matrix()</code>	Entrega TRUE si el objeto es vector o matriz
<code>as.numeric()</code> , <code>as.character()</code>	Transforma el objeto al formato numérico o character
<code>length()</code>	Entrega el largo del vector
<code>%in%</code>	Busca si los elementos pertenecen a un vector dado
<code>ifelse()</code>	Si se cumple una condición retorna 'x' si no, retorna 'y'

Matrices

<code>matrix()</code>	Crea una matriz con las dimensiones deseadas
<code>nrow()</code> , <code>ncol()</code> , <code>dim()</code>	Indica las dimensiones de la matriz o dataframe
<code>cbind()</code> , <code>rbind()</code>	Crea una matriz a partir de dos vectores
<code>rownames()</code>	Entrega el vector de nombres de las filas del objeto

Dataframes y Tibbles (Tidyverse)

<code>data.frame()</code>	Crea una dataframe
<code>str()</code>	Indica el formato de las columnas
<code>names()</code>	Entrega el vector de nombres de las columnas
<code>as_tibble()</code>	Convierte el objeto a tibble
<code>tibble()</code>	Crea un tibble
<code>is.na()</code>	Retorna un vector de TRUE/FALSE indicando si es NA
<code>na.omit()</code>	Elimina todas las filas con algún NA
<code>mtcars\$mpg</code> , <code>mtcars[,1]</code> , <code>mtcars[["mpg"]]</code>	Distintas formas de extraer la variable mpg de una dataframe o tibble
<code>apply()</code>	Para aplicar funciones a las columnas o filas de una dataframe

Listas

<code>list()</code>	Crea una lista
<code>names()</code>	Entrega los nombres de los elementos de la lista

Importación de Datos (*readxl, openxlsx, rio, janitor*)

<code>read_excel()</code>	Para importar archivos en formato Excel
<code>clean_names()</code>	Limpia los nombres de las columnas de un objeto
<code>read_xlsx()</code>	Para importar archivos Excel con celdas agrupadas
<code>import_list()</code>	Para importar todas las hojas de un archivo excel
<code>load()</code>	Para importar archivos en formato RData

Manipulación de datos (*tidyverse, dplyr*)

<code>select()</code>	Extrae determinadas columnas
<code>filter()</code>	Para filtrar registros
<code>mutate()</code>	Para crear variables nuevas o redefinir variables
<code>rename()</code>	Para cambiar el nombre de una variable
<code>group_by()</code>	Para agrupar respecto a una o más variables
<code>summarise()</code> ó <code>summarize()</code>	Para calcular estadísticas de resumen de los datos
<code>count()</code>	Para contar los registros
<code>slice()</code>	Para seleccionar algunos registros
<code>arrange()</code>	Para ordenar en base a una variable
<code>case_when()</code>	Para crear categorías nuevas a partir de una variable

Gráficos (*ggplot2*)

<code>ggplot()</code>	Base para crear un gráfico en ggplot
<code>geom_point()</code>	Gráfico de puntos
<code>geom_histogram()</code>	Gráfico de histograma
<code>geom_boxplot()</code>	Gráfico de boxplot
<code>geom_bar()</code>	Gráfico de barras
<code>geom_line()</code>	Gráfico de línea
<code>scale_x_continuous()</code> , <code>scale_y_continuous()</code>	Para modificar los ejes del gráfico
<code>ylab()</code> , <code>xlab()</code> , <code>labs()</code>	Para cambiar el título de los ejes y agregar subtítulo
<code>ggtitle()</code>	Para definir el título del gráfico
<code>theme()</code>	Para modificar otros aspectos del gráfico

Texto (*stringr*)

<code>toupper()</code> , <code>tolower()</code>	Transforma a mayúscula o minúscula todo el texto
<code>grepl()</code>	Busca determinado patrón en un vector
<code>str_detect()</code>	Indica si se encuentra determinado patrón de texto o no
<code>str_remove_all()</code>	Remueve determinado patrón de todas las coincidencias
<code>str_replace()</code>	Reemplaza cierto patrón por el deseado

Fechas (*lubridate*)

<code>as.Date()</code>	Convierte a formato fecha determinado vector
<code>parse_date_time()</code>	Convierte a formato fecha determinado vector pero es más flexible y sirve si hay dos o más patrones de fecha
<code>day()</code> , <code>year()</code> , <code>month()</code> , <code>yday()</code>	Extrae día, año, mes, día del año, de la fecha

Cruces de tablas (*dplyr*)

<code>left_join()</code>	Cruce por la izquierda, se complementan los registros de A con los registros de B
<code>inner_join()</code>	Cruce intersección, se seleccionan los elementos en A y B
<code>full_join()</code>	Cruce unión, se incluyen todos los elementos de A y B (independiente de si cruzan o no)

<code>right_join()</code>	Cruce por la derecha, se complementan los registros de B con los registros de A
---------------------------	--