

RESUMEN

APUNTES IMPORTANTES SOBRE INTRODUCCIÓN AL LENGUAJE R APRENDIDOS EN COURSEARA

- Puede utilizar el lenguaje de programación R para realizar cálculos estadísticos, visualización de datos y análisis predictivos.
- Los cuatro tipos de datos más comunes en R son entero, numérico, carácter y lógico.
- Puede utilizar la función `class()` y las funciones `is.integer()`, `is.numeric()`, `is.character()` y `is.logical()` para determinar el tipo de datos.
- Puede convertir algunos tipos de datos en otros tipos de datos utilizando las funciones `as.integer()`, `as.numeric()`, `as.character()` y `as.logical()`.
- R proporciona operadores matemáticos que puede utilizar para realizar cálculos sobre sus datos.
- Utilizar variables en sus cálculos y darles nombres descriptivos puede ayudarle a acortar su código y facilitar su lectura.
- Puede controlar el orden de las operaciones utilizando paréntesis.
- El entorno de desarrollo de R incluye la consola de R, archivos de script de R y espacios de trabajo.
- Dos herramientas importantes para trabajar con código R son RStudio y Jupyter.
- Las funciones de RStudio, como el resaltado de sintaxis y el autocompletado de código, facilitan la escritura de código.
- Los componentes principales de RStudio incluyen el editor de archivos, la consola, el espacio de trabajo y el explorador de archivos, parcelas y paquetes.
- Un cuaderno Jupyter está formado por celdas que pueden contener código, archivos markdown o texto sin formato.
- Un cuaderno Jupyter contiene narración, código, datos y gráficos, imágenes o vídeos.
- Un vector es una cadena de números, caracteres o datos lógicos.
- Los factores (también conocidos como variables categóricas) son variables que adoptan un número limitado de valores diferentes que pueden ser nominales u ordinales.
- Puede utilizar R para realizar operaciones sobre un vector, como ordenar los elementos, encontrar el número más pequeño o el más grande, o realizar operaciones aritméticas sobre sus valores.
- Las listas pueden almacenar distintos tipos de datos, a diferencia de los vectores, que sólo pueden almacenar datos de un único tipo.
- Una matriz es una estructura unidimensional o multidimensional que contiene datos del mismo tipo (cadenas, caracteres o enteros).
- Una matriz es como un array, pero debe ser bidimensional y puede organizarse por columnas o filas.

- La principal diferencia entre un marco de datos y otras estructuras de datos, como una lista, es que cada variable tiene un vector de elementos del mismo tipo.
- Las sentencias If utilizan operadores de comparación y lógicos para probar condiciones en el código.
- Los bucles For realizan una operación para cada elemento de una columna de lista, vector o marco de datos.
- Los bucles While realizan una operación hasta que una condición deja de ser cierta.
- Las funciones pueden ser predefinidas o definidas por el usuario.
- En las funciones definidas por el usuario, puede controlar el valor de retorno de una función, añadir lógica mediante sentencias if y llamar a otras funciones.
- Puede definir variables globales utilizando el operador de asignación de variables <<-.
- Puede utilizar las funciones de R para manipular los caracteres de una cadena, dividir una cadena en un vector y recuperar subcadenas específicas dentro de una cadena.
- Las expresiones regulares se utilizan para hacer coincidir patrones en cadenas y texto.
- Puede convertir fechas de un formato a otro utilizando las funciones as.POSIXct() y as.Date().
- Puede realizar operaciones con objetos de datos utilizando funciones como Sys.Date(), Sys.Time(), date() y as.Date().
- Puede interceptar errores en el código R y proporcionar un manejo personalizado de errores y advertencias utilizando sentencias tryCatch().
- Puede leer un archivo CSV o un archivo Excel en un conjunto de datos utilizando las funciones read_csv() y read_excel() y, a continuación, acceder a sus filas, columnas y puntos de datos individuales.
- La función readLines() lee un archivo de texto en un vector de caracteres.
- Las funciones length() y nchar() devuelven información sobre un vector de caracteres.
- La función scan() lee un archivo de texto en un vector de caracteres con cada palabra individual como elemento.
- La función write() exporta un conjunto de datos como un archivo de texto.
- Las funciones write.csv() y write.table() exportan un conjunto de datos como un archivo .csv.
- La función write.xlsx() exporta un conjunto de datos como un archivo Excel.
- La función save() guarda objetos R en archivos .RData.
- HTTP es un protocolo de comunicación para obtener recursos web para clientes desde servidores en Internet y que cada instancia se compone de una solicitud y una respuesta.
- La API REST es un servicio web que utiliza la arquitectura REST para gestionar una solicitud en un servicio web frontend.

- El paquete httr en R tiene funciones que realizan operaciones HTTP y REST comunes.
- El paquete rvest en R tiene funciones que puede utilizar para realizar tareas comunes de web scraping, como leer HTML de una variable de caracteres, leer HTML de una URL, descargar una página web y leerla sin conexión, extraer datos de nodos de una página web y convertir una tabla de una página web en un marco de datos.