

Tarea: Listas y Dataframes en R - Unidad 5

Prof. Rafael Puche & Fernando Hernández

Julio 2025

Instrucciones

1. Abre un nuevo script de R o un archivo `.R` en tu editor de código en `Rstudio`.
2. Realiza los ejercicios propuestos a continuación.
3. Guarda tu script con el nombre `tarea_ls_df_nombre.R`.
4. Cuando hayas completado todos los ejercicios, sube ambos archivos (`tarea_ls_df_nombre.R`) por correo electrónico indicando en el asunto `Tarea - Clase 3 - Nombre`

Ejercicios

Parte 1: Listas

1. Crea una lista llamada

```
AgendaTe1f
```

con los siguientes elementos:

- Nombres
 - Teléfono
 - Correo Electrónico
 - Dirección
 - Tiene laptop
2. Revisa la estructura de la `AgendaTe1f`, donde se visualice los tipos de vectores en la lista. Tip: usa las funciones empleadas hoy en clase.

Parte 2: Dataframes

2.1 Crear un data frame

1. Crea un dataframe llamado `pacientes` con las siguientes columnas: nombre, edad, genero, diagnostico y tratamiento. Ingresar los datos de al menos 5 pacientes diferentes.
2. Agrega una nueva columna llamada `fecha_ingreso` al dataframe `pacientes` utilizando la función `cbind`.
3. Crea un nuevo dataframe llamado `nuevos_pacientes` con las columnas nombre, edad, género y diagnostico. Ingresar los datos de al menos dos pacientes nuevos.
4. Combina los dataframes `pacientes` y `nuevos_pacientes` utilizando la función `rbind`.

2.2 Analisis de un data frame

En esta parte, utilizaremos el dataframe `penguins` del paquete `palmerpenguins` de R, que contiene datos biológicos reales de varias especies de pingüinos del género *Pygoscelis* que coexisten en las islas del archipiélago Palmer, en la Antártida.

1. Instala el paquete de datos empleando el comando: `install.packages("palmerpenguins")`
2. Carga el dataframe `penguins` usando `library(palmerpenguins)`.
3. Guarda el dataframe en un objeto llamado `pinguinos <- penguins`. Nota: Este dataframe contiene varios `NA` por lo cual usa la función `na.omit` para limpiar tus datos
4. Realiza la exploración del dataframe usando `dim()`, `str()`, `summary()` según corresponda:
 1. ¿Cuáles son sus dimensiones (filas y columnas)?
 2. ¿Qué tipos de datos (`factor`, `numeric`, `integer`, etc.) contienen las columnas?
 3. ¿Cuál es la **mediana** para la longitud de las aletas y para la masa corporal?
 4. ¿Cuales y cuántas especies diferentes de pingüinos hay en el conjunto de datos?
5. Realice subconjuntos de datos de `pinguinos` teniendo en cuenta lo siguiente:
 1. Divida el dataframe en dos grandes grupos de pollo utilizando los operadores `[]` `$` y guarda en objetos separados `pinguinos_machos` y `pinguinos_hembras` **Tip:** Encuentra la columna correcta que te permita realizar el subconjunto.
 2. Cuantos pinguinos pesan mas de 5000 gramos?
 3. Cuantos `pinguinos_hembras` hay mayores a 4000 gramos?

Bonus Track ¿Existen diferencias significativas en la masa corporal entre machos y hembras?
¿Cuál de los dos sexos tiende a ser más pesado?