

# dyplr: Ejercicios

Bioinformática - Pasto 2017

Guillermo Torres

## Ejercicios

Para estos ejercicios vamos a utilizar el dataset sobre el sueño de los mamíferos. Descargar el archivo **msleep\_ggplot2.csv** de la siguiente manera:

```
library(downloader)
url <- "https://raw.githubusercontent.com/guigotoe/Workshop-Pasto-2017/master/extdata/"
filename <- "msleep_ggplot2.csv"
download(paste0(url,filename), destfile=filename)
msleep <- read.csv(filename)
## o
dir <- "extdata/"
filename <- "msleep_ggplot2.csv"
msleep <- read.csv(paste0(dir,filename))
```

Vamos a leer estos datos y repasar las funciones de selección y filtraje de dplyr. También vamos a revisar dos clases diferentes: data frames y vectores.

1. Leer el archivo **msleep\_ggplot2.csv**. Luego usar la función **class()** para determinar que tipo de objeto es.
2. Ahora usar la función **filter()** (asegurarse de cargar el paquete dyplr) para seleccionar solamente los primates ( *Primates* en columna *order*). Cuantos animales en la tabla son primates? *Pista*: la función **nrow()** devuelve el número de filas de un data frame o matriz y la función **ncol()** devuelve el número de columnas de un data frame o matriz.
3. Cual es la clase de el objeto que se obtiene despues de aplicar la función **filter()**?
4. Ahora usar la función **select()** para extraer las horas de sueño total de los primates. Qué clase es este objeto? *Pista*: usar **%>%** para concatenar el resultado de la función filter para la selección.
5. Ahora deseamos calcular la cantidad promedio de sueño de los primates (el promedio de los numeros previamente obtenidos). Tener en cuenta que la función media (mean) requiere un vector, por lo tanto si la aplicamos sobre la salida del punto anterior obtendremos un error. *Pista*: revisar el archivo de ayuda de la función **unlist()** para calcular el promedio deseado.
6. Para el último ejercicio, nosotros podemos usar la función **summarize()** de dplyr. Aunque no se ha introducido esta función, pueden leer el archivo de ayuda y repetir el ejercicio 5, esta vez usando filter y summarize para obtener la respuesta.