

Tarea 1 - Modulo 2

La siguiente tarea consiste en **12 preguntas cortas** que consideran el material cubierto en las tres últimas clases.

Recuerde que si no tiene instalada alguna de las librerías que se requieran para esta tarea deberá hacerlo utilizando `install.packages("NombreDeLaLibrería")`. De forma similar, y teniendo ya las librerías instaladas, cuando quiera utilizar alguna de estas deberá cargarlas cada vez que inicie una sesión de R utilizando `library(NombreDeLaLibrería)`.

Recuerde que si quiere saber más sobre una función en particular puede utilizar `?NombreFuncion` o `help(NombreFuncion)`.

Finalmente, la tarea debe ser entregada en **UN SOLO ARCHIVO RMARKDOWN** cuyo nombre deberá ser **Apellido_Nombre.Rmd**

1. ¿Qué saldrá en la consola al ejecutar el siguiente código? O en otras palabras, ¿qué se imprime cuando se ejecuta la variable resultado?

```
x <- 2
y <- -1
b <- c(x, y)
resultado <- b[1]
resultado
```

2. Se tiene un objeto de la siguiente forma:

```
objeto <- c(1,2,3,4,5,'a')
```

¿Cuál de las siguientes opciones es correcta para que solo se imprima (se vea) un objeto con valores de 1, 2, 3, 4, 5? (sin la letra a).

- I) objeto[-a]
- II) objeto[1:5]
- III) objeto[-6]

- a) Solo I
- b) Solo II
- c) Solo III
- d) Solo II y III
- e) Todas

3. Cargue el archivo `datos_tarea.xlsx` y asígnelo a la variable `datos`. Utilice la función `read_xlsx` de la librería `readxl`.

4. A través del uso de funciones, describa el objeto `datos`. ¿A qué tipo de objeto corresponde `datos`? ¿Cuántas observaciones y variables tiene? ¿De qué tipo son sus variables? ¿Cuál es el valor máximo y mínimo de la variable `obesity_rate`?

5. Replique el resultado de la siguiente línea de código utilizando la función `filter()` de la librería `dplyr`.

```
datos[datos$year == 2010,]
```

6. Replique el resultado de la siguiente línea de código utilizando la función `select()` de la librería `dplyr`.

```
datos[, "Country"]
#o
datos[, 1]
```

7. Replique el resultado de la siguiente línea de código utilizando las funciones `filter()` y `select()` de la librería `dplyr`. Asigne el resultado a `datos2`.

```
datos[datos$year == 2005 | datos$year == 2015, c("Country", "year", "obesity_rate")]
#o
datos[datos$year == 2005 | datos$year == 2015, c(1, 4, 5)]
```

8. Utilizando la función `spread()` de la librería `tidyr` convierta los valores de la variable `year` dentro `datos2` en columnas. Los valores (filas) de estas nuevas (2) columnas deben ser los correspondientes a `obesity_rate`. Asigne el resultado a `datos3`.

9. Convierta `datos3` nuevamente a la forma de `datos2` (columnas de años como valores dentro de una columna `year`) utilizando la función `gather()` de la librería `tidyr`.

10. Replique el resultado de las siguientes líneas de código utilizando la función `rename()` de la librería `dplyr`. Asigne el resultado a `datos4`.

```
colnames(datos2) <- c("País", "Año", "tasa_obesidad")
```

11. Vuelva a escribir el código descrito a continuación utilizando `%>%`.

```
summarize(
  group_by(
    filter(datos, year == 1995),
    Region),
  promedio_region = mean(obesity_rate, na.rm = TRUE))
```

12. Tomando como base el siguiente código:

```
x <- rnorm(100)
y <- 2*(x^2) + 10 + rnorm(100)
d <- as.data.frame(cbind(x,y))
```

Genere un gráfico con la variable `x` en el eje horizontal e `y` en el eje vertical. Luego agregue dos `geom` uno con puntos de color azul y luego una línea de tendencia (`smooth`) de color rojo. Considere como base `ggplot(d, aes(,))`. Utilice la librería `ggplot2`.