

## CASO PRÁCTICO: ESTUDIOS DESCRIPTIVOS

Vamos a hacer un estudio estadístico utilizando una base de datos que contiene las siguientes variables: edad, raza, peso, relación con el tabaco e hipertensión de un total de 189 mujeres (n=189).

Nuestras cinco variables son diferentes y, por tanto, para estudiarlas no vamos a poder utilizar ni la misma representación gráfica ni los mismos parámetros de medida.

La edad es una variable que toma un valor numérico y, por tanto, se trata de una **variable cuantitativa continua**.

En el caso de la raza, estamos ante una **variable cualitativa**, ya que como hemos explicado, la raza no se puede medir numéricamente y tomará valores en forma de categorías (negras, blanca y otras). Además, al tratarse de tres categorías o modalidades, se trata de una variable politómica o no dicotómica, porque toma más de dos.

En cuanto al peso, ya sí que lo podemos medir numéricamente, por lo que sí la trataremos como una **variable cuantitativa continua**.

Por último, nos encontramos con la relación que presenta cada mujer con el tabaco, es decir, si ésta es fumadora o no, y si padece hipertensión o, de lo contrario, no es hipertensa. Estas variables, como podemos observar, no se pueden medir numéricamente, por lo que las consideramos cualitativas y, además, solo pueden tomar dos modalidades: fumadora o no fumadora e hipertensa o no hipertensa, que además son valores excluyentes, por lo que podemos concluir que se tratan de **variables cualitativas dicotómicas**.

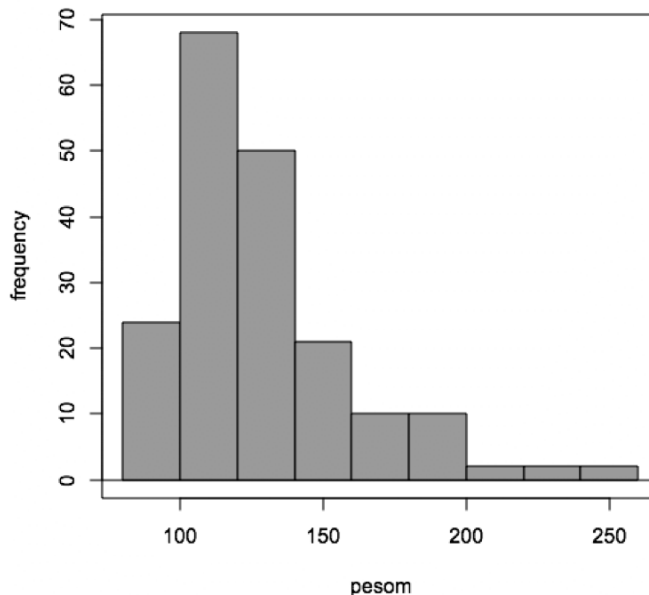
Ahora que sabemos con qué tipos de variables contamos, vamos a hacer un pequeño estudio estadístico en el que las relacionamos, teniendo en cuenta los tipos de variables y eligiendo en función del mismo, la representación gráfica que más se ajuste a las necesidades que queremos.

En primer lugar vamos a cargar el archivo "mujeres". Abrimos el programa de R, cargamos el Rcmdr, bien con el comando `>library(Rcmdr)` o bien utilizando las pestañas del R: [paquetes -> cargar paquete -> Rcmdr](#) y nos aparecerá la pantalla del R commander.



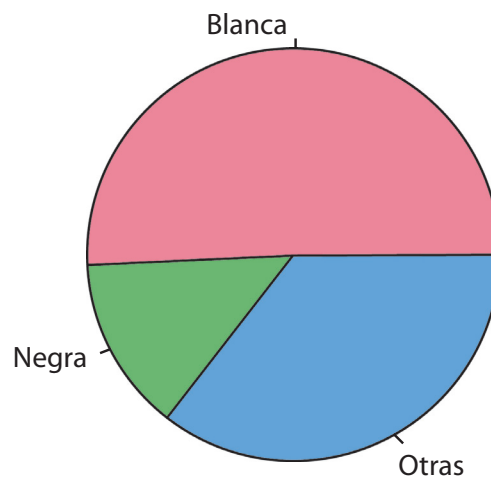
Ahora tenemos que cargar los datos en el programa. En las pestañas del R commander seleccionamos **Datos -> Importar datos desde Excel**, y damos nombre a nuestro archivo ("muj"). Cuando nos aparezca el explorador, buscamos nuestro archivo ("mujeres") y seleccionamos la hoja a la que lo queremos importar (Hoja 1). Si todo ha salido bien, podremos observar el nombre de nuestro archivo en la ventana de "Conjunto de datos" y estaremos listos para trabajar.

Para empezar, vamos a representar gráficamente el peso de las mujeres estudiadas para poder observar la frecuencia. Al tratarse el peso de una variable cuantitativa, la representación gráfica que más se va a ajustar a lo que queremos conseguir es el histograma. Para ello, seleccionamos en las pestañas del R commander **Gráficas ->Histograma** y en la ventana que nos aparece seleccionamos la variable que queremos estudiar, que en nuestro caso es el peso.



Esta sería la gráfica que nos permitirá observar la frecuencia del peso en las mujeres estudiadas:

### Raza

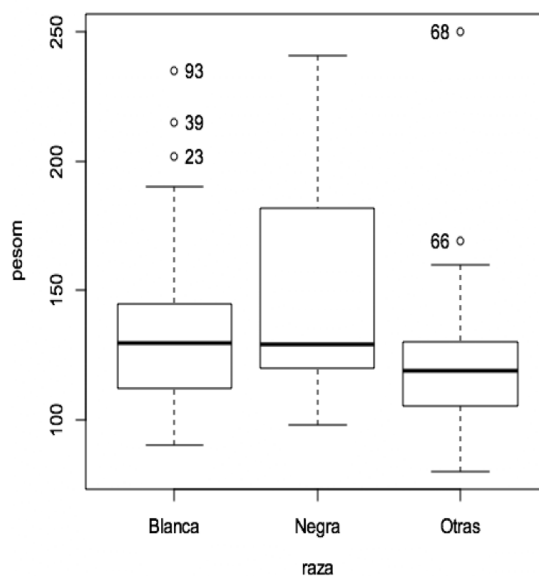


Ahora vamos a probar con una variable cualitativa como puede ser la raza. La raza, como ya hemos dicho, es politómica (tiene más de dos categorías), por lo que para observar su distribución vamos a utilizar el diagrama de sectores. Para elaborarlo, procederemos igual que en el caso anterior, pero ahora, en lugar de seleccionar "Histograma", seleccionamos "Diagrama de sectores" y como variable seleccionamos en lugar del peso, la raza, y el resultado lo podéis observar en la gráfica:

Muchas veces es de gran utilidad la elaboración de una tabla que nos muestre las variables con las que contamos, ya que nos permite relacionarlas entre si. Para ello seleccionamos la pestaña del R commander [Estadísticos-> Tablas de contingencia -> Tablas de doble entrada](#) y seleccionamos las dos variables que queramos relacionar. Un ejemplo sería crear una tabla para ver el número de mujeres que fuman según su raza. Para ello, las variables que seleccionaremos serán "raza" y "tabaco" y este sería el resultado:

Tabaco		
Raza	No Fuma	Fuma
Blanca	44	52
Negra	16	10
Otra	55	12

Por último, vamos a elaborar un tipo de gráfica que nos muestre la distribución de las variables a estudiar. Si queremos ver si el peso de las mujeres estudiadas en función de la raza se desvía sustancialmente de la distribución normal o, de lo contrario, se ajusta a ésta, la representación gráfica más apropiada sería el diagrama de cajas. Para elaborarlo, procederemos de la misma manera que las gráficas anteriores salvo que a la hora de seleccionar el tipo de gráfica, elegiremos "Diagrama de caja". Cuando nos aparezca la ventana para seleccionar la variable, seleccionaremos la opción "Gráfica por grupos" y una vez allí, las dos variables de las que queramos conocer la distribución, que en nuestro caso son el peso y la raza. Este sería el resultado:



Ahora os propongo elaborar gráficas como las anteriores utilizando variables diferentes que aparezcan en la base de datos, así como un par de tablas.

---