

LABORATORIO2 BIOINFORMÁTICA SERGIO SUALDEA Y JUAN JARÉN

Crear un nuevo dataframe con todas las especies menos los Droides

```
starwars_nodroids <- starwars %>% filter(species != "Droid")
```

¿Cuántos registros cumplen las condiciones finales?

```
#77 personajes que no son especie Droide
```

¿Cómo calcularías la desviación estándar (sd) de esos parámetros?

```
tabla_estadistica <- starwars %>% group_by(species) %>% summarise(mean_height =  
mean(height, na.rm=TRUE), mean_mass = mean(mass, na.rm = TRUE), sd_height =  
sd(height, na.rm=TRUE), sd_mass = sd(mass, na.rm = TRUE))
```

Al crear los gráficos puedes observar que hay un punto que corresponde a un personaje con una masa muy grande. Inspecciona el dataset, filtra usando las funciones de tidyverse, crea un nuevo dataframe sin ese personaje y crea de nuevo el gráfico final.

```
#Descubrimos quien es el que mas pesa
```

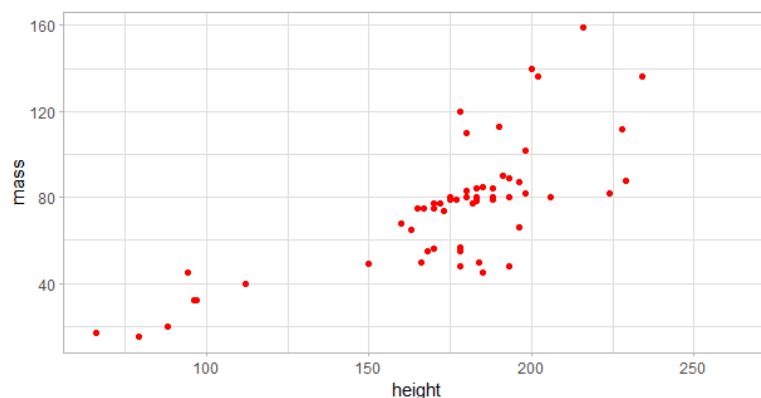
```
masas <- starwars %>% select(name, mass)
```

```
#Ahora creamos un dataframe sin este personaje
```

```
starwars_sin_gordo <- starwars %>% filter(name != "Jabba Desilijic Tiure")
```

```
#Creamos el plot del nuevo dataframe
```

```
ggplot(starwars_sin_gordo, aes(height, mass)) + geom_point(colour = "red") + theme_light()
```



EJERCICIO

Cargamos en R studio el dataset toy.csv

Inspecciona el dataset, haz un resumen de la media (mean) de las variables (Peso, Altura, IMC, IAS, CCintura). Agrupando por sexo.

```
media_variables_sexo <- toy %>% group_by(Sex) %>% summarise(mean_peso =  
mean(Weight_Kg, na.rm=TRUE), mean_altura = mean(Height_cm, na.rm=TRUE),  
mean_IMC = mean(IMC_clas, na.rm=TRUE), mean_IAS = mean(IAS, na.rm=TRUE),  
mean_Ccintura = mean(Ccintura, na.rm=TRUE))
```

Haz una tabla sólo con los pacientes femeninos

```
pacientes_femenino <- toy %>% filter(Sex != "Men")
```

¿Cuántos registros cumplen las condiciones?

Hay 58 pacientes femeninos

¿De estos cuantos tienen Sobrepeso (Overweight)?

```
pacientes_femenino %>% filter(IMC_clas == "Overweight") %>% tally()
```

Hay 9 con sobrepeso

Haz un gráfico usando ggplot relacionando el IMC (Índice de masa corporal) con el peso (Weight_Kg) de todos los pacientes.

```
ggplot(toy, aes(IMC, Weight_Kg)) + geom_point()
```

Repítelo filtrando sólo los pacientes categorizados como "Overweight" y "Obesity".

```
pacientes_sobrepeso_obesidad <- toy %>% filter(IMC_clas != "Normal")
```

```
ggplot(pacientes_sobrepeso_obesidad, aes(IMC, Weight_Kg)) + geom_point()
```