



Objetivo

Ejercitar las distintas maneras de visualizar datos.

Ejercicios

1. Palmer Penguins es un dataset real obtenido de investigaciones realizadas en tres islas diferentes de la Antártida. ¡Felicitaciones! Debido a su entusiasmo por los pingüinos y sus capacidades, ha sido contratado para analizar los datos. Su trabajo consiste en lo siguiente.
2. Cargar el dataset con las siguientes líneas:

```
import seaborn
data_ping = seaborn.load_dataset('penguins')
```

y responder las siguientes preguntas:
 - a. ¿Qué representa cada línea del dataframe?
 - b. ¿Cuántas muestras hay en total?
 - c. ¿Cuáles son las especies de pingüinos consideradas?
 - d. ¿Cuáles son las islas estudiadas?
 - e. Para cada pingüino, ¿con qué datos contamos?
3. Averiguar si las islas están pobladas mayormente por alguna especie en particular, o si éstas coexisten, y en ambos casos deberá notificar en qué proporciones. Es importante mencionar que deberá reportar sus descubrimientos de manera resumida a través de gráficos de barra y de torta.
4. Realizar un histograma de la variable grosor del pico. Repetir separando por especies (con el mismo rango de valores en los ejes, para poder comparar).
5. Realizar lo mismo con las demás variables corporales de los pingüinos. A partir de estos gráficos, responder:
 - a. ¿Se puede determinar la especie de un pingüino a partir de una sola característica?
 - b. ¿Hay alguna característica que permita discernir entre especies mejor que otras?
6. Realizar ahora histogramas de las variables observadas, separadas por sexo (f/m). De manera análoga, considerar si hay alguna variable que permita deducir el sexo de un pingüino.
7. Realizar scatterplots de pares de variables corporales, separadas por sexo. A partir de los gráficos, responder:
 - a. ¿Hay algún par de variables que permita deducir el sexo?
 - b. ¿Y si se fija una especie en particular?



8. Realizar un scatterplot de las variables largo y grosor del pico agregando colores para distinguir las especies. Responder:
 - a. ¿Qué especie muestra mayor dispersión en estas variables?
 - b. La relación entre estas variables, ¿es similar en las 3 especies?
9. Para los pingüinos hembra, a partir de una de sus característica en particular (obviamente que no sea a partir del campo especie), ¿se puede predecir a qué especie pertenece? ¿Y combinando más de una característica?
10. Repetir el punto anterior pero para pingüinos macho.
11. Realizar un gráfico (de línea) donde se vea la relación entre la variable largo del pico y masa corporal, para la especie Adelie. Sugerencia: reordenar el subconjunto de pingüinos Adelie por la variable largo del pico, y utilizarlo para graficar.
12. Realizar un gráfico de barras apiladas para visualizar la cantidad de pingüinos de cada sexo dentro de cada especie.
13. Realizar un boxplot de la variable ancho del pico, separado por sexo. ¿Qué se observa?
14. Realizar un boxplot de la variable largo de aleta, separado por especie. ¿Qué se observa?
15. Realizar un violinplot de la variable largo de aleta, separado por sexo.
16. Realizar un violinplot de la variable masa corporal, separado por especie.