

Actividades a entregar

1. Los ficheros en formato FASTA¹ se utilizan ampliamente en Bioinformática. Consisten en una o más secuencias de bases o aminoácidos divididas en líneas de tamaño razonable (normalmente 70 caracteres), y cada una de ellas



precedida por una línea que comienza con un carácter ">". Esta línea es una descripción de la secuencia y contiene varios identificadores y comentarios que pertenecen a la secuencia que sigue. Escribir un programa que:

- Pida un nombre de fichero
- Compruebe si es correcto
- Si no es correcto, imprima un mensaje de error y vuelva al comienzo.
- Abra el fichero para lectura.
- Cuente el número de secuencias en un fichero FASTA
- Imprima el total de cada base y el número total de bases encontrado en todas las secuencias.



La salida del programa será así:

```
Fichero FASTA:  
5 secuencias  
100 bases (A: 34; C:17; T: 25; G:24)
```

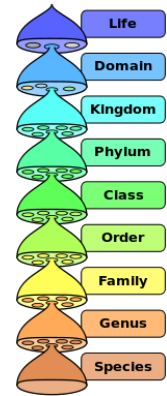
Recomendaciones:

- Usad la librería `os` para funciones de manejo del sistema de ficheros.
- Se valorará la modularidad del código (uso de funciones para pedir el nombre y la ruta del fichero, comprobar si existe y para efectuar el recuento).
- Para leer el fichero usad cualquier aproximación (columnas de ancho fijo, línea dividida con `split`, `csv`, ...).

¹ https://es.wikipedia.org/wiki/Formato_FASTA

2. El fichero `mamiferos.txt` contiene datos de historia natural de los mamíferos. Su estructura es como sigue:

- Primera línea (comienza con >): campos del fichero.
- Resto: valores de orden, familia, género, especie, peso del adulto (gr.) y de la cría (gr.), separados por comas.
- Los datos estarán **ordenados** alfabéticamente por orden, familia, género y especie.



Codificar un programa para:

- Abrir el fichero para lectura.²
- Pedir por teclado una *familia* concreta.
- Localizar la familia en el fichero.
- Localizar especies **cuya cría supere el 15% del peso de su madre**.
- Imprimir el resultado en pantalla (o mensaje de error).
- Cerrar el fichero.

Ejemplos para *Bovidae* y *Macroscelididae*:

Familia Bovidae			
Género	Especie	Peso madre (g)	Peso cría (g)

Gazella	leptoceros	23600.00	4640.00
Gazella	thomsonii	16300.00	2553.75
Madoqua	saltiana	3250.00	690.00
Neotragus	batesi	2500.00	500.00

Familia Macroscelididae			
Género	Especie	Peso madre (g)	Peso cría (g)

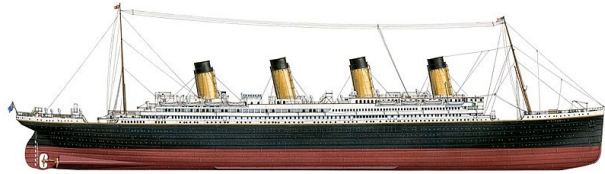
Elephantulus	intufi	51.95	10.00
Elephantulus	myurus	51.94	8.10
Elephantulus	rufescens	58.00	10.73
Macroscelides	proboscideus	40.00	7.00

Recomendaciones:

- Se valorará la modularidad del código (uso de funciones).
- Para leer el fichero se puede usar cualquier aproximación (columnas de ancho fijo, línea dividida con split, csv, ...).

² Es obligatorio controlar si existe el fichero. Si no existe imprimir un mensaje de error y terminar.

3. El fichero `titanic.csv` tiene los datos de algunos pasajeros del trágico viaje inaugural, todos separados por punto y coma (carácter `;`). En concreto, si sobrevivieron o



no, clase (1, 2 o 3), nombre, sexo, edad y tarifa del billete. El ejercicio consiste en procesar el fichero e imprimir una tabla como la siguiente:

	Clase			
Superv.	1	2	3	Total
No	80	97	368	545
Si	136	87	119	342
Total	216	184	487	887



Además, se pide crear dos ficheros en una subcarpeta `resultados`, que deberá crear el programa (si ésta no existe):

- Un fichero de texto en formato CSV con los datos de supervivientes agrupados por tramos de edad. Constará de una cabecera con los nombres de los campos (`'RANGO_EDAD'`, `'NUMERO'`) y una línea por cada rango de edad (p.e. 0-10, 10-50, 50-100). Los datos de cada línea estarán separados por puntos y comas (carácter `;`).
- Un fichero de texto `'plano'` con los nombres de los fallecidos que pagaron más de una cantidad determinada (introducida por teclado).



Recomendaciones:

- Se valorará la modularidad del código (uso de funciones).
- Para leer el fichero utilizad la aproximación que creáis mejor (columnas de ancho fijo, línea dividida con `split`, `csv`, ...).
- Para efectuar el recuento utilizad cualquier tipo de colección (lista, diccionario, ...)
- Para simplificar el ejercicio el fichero existe y es correcto.