

TEST

Cada pregunta del test **sumará 1 punto** en caso de contestarse de forma correcta y **restará 0.3 puntos** en caso de contestarse de forma errónea. Las preguntas en blanco no restarán. Para superar esta parte, será necesario obtener **al menos 2 puntos sobre 4**.

1. ¿Cuál es el resultado de ejecutar el siguiente código?

```
def fun(a=2,b=3):  
    return a*b  
print(fun(b=5))
```

- a) Se imprime el valor 10
- b) Se produce un error en tiempo de ejecución
- c) Se imprime el valor 6
- d) Se imprime el valor 15

2. Dado el siguiente fragmento de código:

```
import pandas as pd  
import numpy as np  
data = pd.DataFrame(np.arange(20).reshape((5, 4)),  
                    index=['X', 'Y', 'Z', 'W', 'V'],  
                    columns=['X', 'Y', 'Z', 'W'])
```

¿Qué ocurre si ejecutamos `data=data.drop('Y')`?

- a) Se elimina la columna Y de data
- b) Se elimina la fila Y de data
- c) Se eliminan tanto la columna Y como la fila Y de data
- d) Se genera un error al no poder decidir si hay que eliminar por filas o por columnas

3. Dado el siguiente fragmento de código, cuál será el contenido de `l` tras ejecutarse el código:

```
a = (x ** 2 for x in range(10) if x % 2 == 0)  
l = [x for x in a if x % 3 == 0]
```

- a) `[0, 66]`
- b) `[0, 18, 36]`

- c) []
- d) [0, 36]

4. ¿Cuál es el resultado de ejecutar el siguiente código?

```
tup = (1, [] , "a")  
tup[1].append(4)  
print (tup[1])
```

- a) Se produce un error en tiempo de ejecución
- b) Se imprime 1
- c) Se imprime [4]
- d) Se imprime []

Problemas

Cada uno de estos problemas se puntuará sobre 3 puntos. Para superar esta parte será necesario obtener **al menos 2 puntos sobre 6**. Se recuerda a los estudiantes que los pequeños errores en el código (como nombres exactos de funciones, orden de los parámetros y similares) no se penalizarán.

Problema 1. El director de una empresa está interesado en contratarnos como científicos de datos y en la entrevista nos informa de que tienen un conjunto de datos guardado en una hoja de cálculo con el que quieren trabajar.

El director nos informa de que en primer lugar solo quieren trabajar con datos numéricos. Además, como su conjunto de datos contiene muchas filas y su empresa tiene recursos de computación limitados, nos pide que si podemos hacer algo para que ocupe lo mínimo posible en memoria.

Para realizar el ejercicio se pide explicar a alto nivel qué acciones tomaríamos para realizar lo que nos pide el director. Posteriormente, se pide escribir el código asociado a ello. Finalmente, añadir código para ver qué tamaño (en cuanto a filas y columnas) tenía la colección inicial y qué tamaño tenía la colección final.

Problema 2. Un portal de internet que ofrece una aplicación similar a spotify (streaming de música) nos contrata como expertos en datos para ayudarles a gestionar su base de datos de canciones. Esta base de datos se encuentra almacenada en un fichero CSV con la siguiente estructura:

- En cada línea se encuentran los campos de cada canción, separados por comas.
- En la primera línea se encuentran los nombres de los campos, separados por comas, a modo de cabecera.

- De cada canción se dispone de los siguientes datos: nombre, grupo, género y duración. Los tres primeros almacenados como texto y el último almacenado como un entero que indica el número de segundos de duración de la canción.

También nos indican que, tras analizar el comportamiento de sus usuarios, las tres acciones más demandadas son las siguientes:

- (a) Crear una lista de canciones con la discografía completa de un grupo (todas las canciones de dicho grupo).
- (b) Crear una lista aleatoria de canciones de un determinado género que, conjuntamente, tenga una duración aproximada de X segundos. Es decir, que si al añadir una canción llegamos a X segundos en total (o nos pasamos), dejamos de añadir canciones a la lista.
- (c) Encontrar la canción más larga de un determinado género.

Para realizar el ejercicio se pide explicar a alto nivel qué acciones tomaríamos para preparar los datos de manera que estas operaciones se ejecuten en el menor tiempo posible y, adicionalmente, escribir el código asociado a ello. Por último, se pide el código necesario para resolver las tres acciones que más demandan los usuarios.