



Ciclo 1 Fundamentos de programación

Reto 5

Descripción del problema: Una institución educativa tiene un registro de notas obtenidas por los estudiantes en varias asignaturas en un determinado curso. A partir de estos datos se desea tener un informe del porcentaje de hombres y de mujeres que están en el registro y el promedio de cada asignatura (entre las cuales están: matemáticas, escritura y lectura). Como dato importante también se desea saber el promedio de cada asignatura por género y la media aritmética obtenida por preparación, es decir, si realizó o no el curso de preparación de los exámenes.

Para realizar esta tarea los datos se encuentran en un archivo llamado **StudentsPerformance.csv**, el cual se encuentra disponible en la ruta siguiente que usted puede descargar:

<https://github.com/pyplusplusMF/reto5grupo8/blob/main/StudentsPerformance.csv?raw=true>

Escriba una función que reciba como parámetro la **url del archivo (str)**, y lo lea desde el **directorio de trabajo actual**. A partir de estos datos, construya un **DataFrame** sobre el cual, realizará los análisis y retorne un diccionario con los datos solicitados.

La función debe retornar un **diccionario** que tenga como llaves:

- **'percentByGender' : (dict)**, donde el valor corresponde a otro diccionario cuyas llaves de segundo nivel son: **'percentMale' : (float)** y **'percentFemale' : (float)**, donde cada llave contiene un valor de tipo real, redondeado a dos cifras decimales que corresponde al porcentaje hombres y mujeres respectivamente.
- **'meanByCourse' : (dict)**, el valor corresponde a otro diccionario cuyas llaves de segundo nivel son los nombres de las asignaturas del DataFrame, es decir: **'math score' : (float)**, **'writing score' : (float)** y **'reading score' : (float)**, y el valor de cada asignatura es el promedio obtenido en ellas, redondeado a dos cifras decimales.
- **'meanByGender' : (dict)**, el valor corresponde a otro diccionario cuyas *llaves de segundo nivel* son los nombres de las asignaturas del DataFrame, es decir: **'math score' : (float)**, **'writing score' : (float)** y **'reading score' : (float)**, y el valor de cada asignatura es otro diccionario cuyas *llaves de tercer nivel* son el género **'female' : (float)** y **'male' :**



(float) y el valor es el promedio por género de la asignatura, redondeado a dos cifras decimales.

- **'meanByPreparation' : (dict)**, el valor es otro diccionario cuyas *llaves de segundo nivel* son los estados de preparación del examen que son **'completed' : (dict)** y **'none' : (dict)**, en donde las *llaves de tercer nivel* son: **'math score' : (float)**, **'writing score' : (float)** y **'reading score' : (float)**, cuyo valor es el promedio obtenido por estado de preparación, redondeado a dos cifras decimales.

Nota: Recuerde que el fichero debe estar en la misma carpeta donde desarrolle su programa de Python. El archivo debe ser un archivo .csv separado por comas. El encabezado de cada serie está en inglés, por tal razón usted debe escribir las llaves como se describieron en el enunciado. Para mas información sobre las columnas del dataset puede consultar en <https://www.kaggle.com/spscientist/students-performance-in-exams?select=StudentsPerformance.csv>

Nota: Utilice un bloque **try except** para abrir el archivo, en caso de error, retorne la siguiente cadena: 'Opps algo salio mal, revise nuevamente'

Ejemplo:

Entrada	Salida
nombre del archivo .csv	return
Nombre del archivo en formato .csv, el archivo se debe cargar directamente desde esta dirección: https://github.com/pyplusplusMF/reto5grupo8/blob/main/StudentsPerformance.csv?raw=true	<pre>{'percentByGender': {'percentMale': XXX, 'percentFemale': XXX}, 'meanByCourse': {'math score': XXX, 'writing score': XXX, 'reading score': XXX}, 'meanByGender': {'math score': {'female': XXX, 'male': XXX}, 'writing score': {'female': XXX, 'male': XXX}, 'reading score': {'female': XXX, 'male': XXX}}, 'meanByPreparation': {'completed': {'math score': XXX, 'writing score': XXX, 'reading score': XXX}, 'none': {'math score': XXX, 'writing score': XXX, 'reading score': XXX}}}</pre> <p>Nota: las XXX corresponden a los datos numéricos que usted debe obtener, son números reales redondeados a dos cifras decimales, puede usar round.</p>

Entrada:

rutaDelArchivo =

<https://github.com/pyplusplusMF/reto5grupo8/blob/main/StudentsPerformance.csv?raw=true>

Nombre	Tipo	Descripción
StudentsPerformance.csv	str	dataset de la información, todos los datos se encuentran en inglés, debe poner las llaves como se describen en el enunciado.



Salida: Retorno de un texto en caso de error de apertura del archivo o un diccionario que contiene las llaves descritas a continuación:

Nombre	Tipo	Descripción
percentByGender	dict	El valor corresponde a otro diccionario cuyas <i>llaves de segundo nivel</i> son: 'percentMale' : (float) y 'percentFemale' : (float) , donde cada llave contiene un valor de tipo real, redondeado a dos cifras decimales que corresponde al porcentaje hombres y mujeres respectivamente.
meanByCourse	dict	El valor corresponde a otro diccionario cuyas <i>llaves de segundo nivel</i> son los nombres de las asignaturas del DataFrame, es decir: 'math score' : (float) , 'writing score' : (float) y 'reading score' : (float) , y el valor de cada asignatura es el promedio obtenido en ellas, redondeado a dos cifras decimales.
meanByGender	dict	El valor corresponde a otro diccionario cuyas <i>llaves de segundo nivel</i> son los nombres de las asignaturas del DataFrame, es decir: 'math score' : (float) , 'writing score' : (float) y 'reading score' : (float) , y el valor de cada asignatura es otro diccionario cuyas <i>llaves de tercer nivel</i> son el género 'female' : (float) y 'male' : (float) y el valor es el promedio por género de la asignatura, redondeado a dos cifras decimales.
meanByPreparation	dict	El valor es otro diccionario cuyas <i>llaves de segundo nivel</i> son los estados de preparación del examen que son 'completed' : (dict) y 'none' : (dict) , en donde las <i>llaves de tercer nivel</i> son: 'math score' : (float) , 'writing score' : (float) y 'reading score' : (float) , cuyo valor es el promedio obtenido por estado de preparación, redondeado a dos cifras decimales.
Opps algo salio mal, revise nuevamente	str	Mensaje que retorna la función, si se usa try, except y se valida que en realidad exista el archivo



Esqueleto:

rutaDelArchivo =

<https://github.com/pyplusplusMF/reto5grupo8/blob/main/StudentsPerformance.csv?raw=true>

```
import pandas as pd

def studentPerformance (rutaDelArchivo:str) -> dict:

    pass
```

Enlace de descarga del archivo :

<https://github.com/pyplusplusMF/reto5grupo8/archive/refs/heads/main.zip>