

Módulo 7 - Strings

Aula 7.4 - Arquivos de Texto - Exercícios

Instruções para submissão: Para cada questão, crie no VSCode um arquivo chamado `aula4_questaoX.py`, sendo `X` o número da questão. Faça commit de todos os arquivos para a pasta `modulo7` do seu repositório da disciplina no GitHub.

- 1) Escreva um script Python que solicita uma frase do usuário e a salve em um arquivo chamado "frase.txt" no mesmo local do seu script. Imprima em seguida o caminho completo do arquivo salvo.

```
Digite uma frase: Bom dia, meu nome é Davi.  
Frase salva em /Users/laranjeira/python-basico/frase.txt
```

- 2) Escreva um script que leia o arquivo salvo no exercício anterior e salve em um novo arquivo "palavras.txt", removendo todos os espaços em branco e caracteres não alfabéticos, e separando cada palavra em uma linha. Ao final, imprima o conteúdo do arquivo "palavras.txt".

```
Bom  
dia  
meu  
nome  
é  
Davi
```

- 3) Baixe o arquivo contendo o roteiro do filme brasileiro "Estômago" e salve em seu computador com o nome "estomago.txt". Em seguida crie um script em Python que abra o arquivo para leitura e imprima:
 - O texto das primeiras 25 linhas
 - O número de linhas do arquivo
 - A linha com maior número de caracteres
 - O número de menções aos nomes dos personagens "Nonato" e "Íria" (inclua todas as variações de maiúsculas e minúsculas e atenção para não incluir a substring "iria" se ela fizer parte de outras palavras).
- 4) Vamos fazer o jogo da forca! Antes de programar:
 - Crie um arquivo no seu computador chamado "gabarito_forca.txt" com uma lista de 10 palavras de sua escolha (separadas por quebras de linha, "\n"). Essas serão as opções de palavra do jogo.
 - Crie um arquivo chamado "gabarito_enforcado.txt" com o conteúdo apresentado ao final dessa questão.

Escreva um programa em Python para executar o jogo, de acordo com as definições:

- Abra o arquivo "gabarito_forca.txt" e escolha aleatoriamente uma palavra;
- Com o arquivo "gabarito_enforcado.txt", crie uma lista de strings com os estágios do enforcado;
- No início exiba o número de letras na palavra como underscores;
- Permita que o jogador insira letras para adivinhar a palavra;
- Em caso de acerto, mostre o progresso do jogador substituindo os underscores correspondentes à letra digitada;
- Em caso de erro, crie a função "imprime_enforcado()" que recebe um inteiro indicando o número de erros do jogador e imprime o enforcado correspondente;
- Limite o número de tentativas para 6 (as partes do enforcado).

gabarito_enforcado.txt

```
|---|
|
|
=====
```

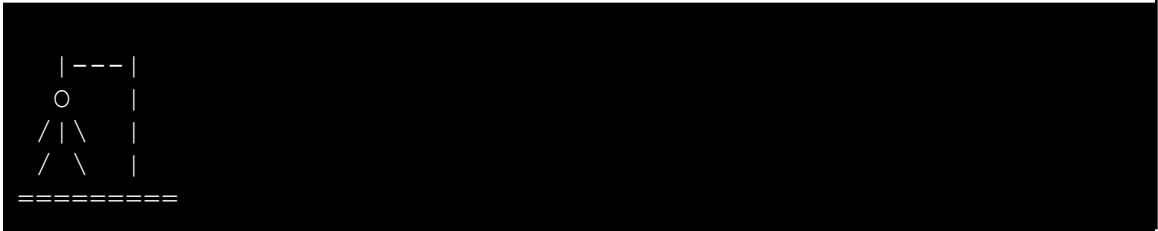
```
|---|
o  |
|
=====
```

```
|---|
o  |
|  |
=====
```

```
|---|
o  |
/  |
=====
```

```
|---|
o  |
/|\ |
=====
```

```
|---|
o  |
/|\ |
/  |
=====
```



5) A extensão ".csv" significa "*comma-separated values*" ou "valores separados por vírgula". É a extensão utilizada por sistemas de gerência de tabelas como o Microsoft Excel ou Google Sheets. Nesse exercício vamos criar uma planilha com dados sobre livros que você já leu ou gostaria de ler. Siga as instruções.

- Selecione pelo menos 10 livros que você leu ou gostaria de ler. Você deve reunir as seguintes informações: título, autor, ano de publicação e número de páginas.
- No Python, crie um arquivo chamado "meus_livros.csv", aberto para escrita.
- Na primeira linha escreva os títulos da planilha separados por vírgula (sem espaço em branco). Os títulos são: "Título", "Autor", "Ano de publicação" e "Número de páginas". Lembre de finalizar a linha com uma quebra de linha.
- A partir da segunda linha escreva as informações de cada livro que você levantou, separando cada informação por uma vírgula (sem espaço em branco). Lembre de finalizar cada linha com uma quebra de linha.
- Feche o arquivo para salvá-lo e abra com a ferramenta de planilhas de sua escolha. Como você já tem conta no Google, sugiro abrir com o Google Sheets.

Seu arquivo deve ser aberto como uma planilha parecida com essa:

Título	Autor	Ano de publicação	Número de páginas
O Caçador de Pipas	Khaled Hosseini	2003	368
Torto Arado	Itamar Vieira Junior	2019	264

6) Vamos descobrir as músicas mais populares do Spotify nos últimos 10 anos!

- Crie uma conta no [Kaggle](#), uma das principais plataformas de ciência de dados e aprendizado de máquina. Em disciplinas avançadas vamos trabalhar com bases de dados provenientes de lá!
- Baixe o arquivo spotify-2023.csv no final da [página que apresenta os dados](#).
- No Python, abra o arquivo para leitura e imprima as cinco primeiras linhas
 - Para abrir o arquivo, defina o parâmetro `encoding='latin-1'`
- Após compreender a estrutura do arquivo (divisão em colunas, caracter separador de coluna, etc.) passamos para a etapa de extração de informações.

O arquivo está estruturado da seguinte forma: cada linha representa uma música e contém as seguintes informações separadas por vírgula (CSV):

track_name,artist(s)_name,artist_count,released_year,released_month,released_day,in_spotify_playlists,in_spotify_charts,streams,in_apple_playlists

Usaremos apenas informações das colunas:

track_name	Nome da música
artist(s)_name	Nome do artista
artist_count	Número de artistas listados em artist(s)_name
released_year	Ano de lançamento
streams	Número de vezes que a música foi tocada no Spotify

Você deve criar um script Python para processar esse arquivo e gerar uma lista com 10 elementos, cada qual representando a música mais tocada de cada ano no intervalo de 2012 a 2022. Considere somente músicas dentro do intervalo solicitado. Cada elemento da lista produzida deve conter as seguintes informações:

```
[track_name, artist(s)_name, released_year, streams]
```

Essa atividade tem alguns desafios. Assim como as colunas da tabela são separadas por vírgulas, músicas com mais de um artista (`artist_count>1`) terá o campo `artist(s)_name` entre aspas com o nome dos artistas separado por vírgulas. Ex:

Seven (feat. Latto) (Explicit Ver.),**"Latto, Jung Kook"**,2,2023, ...

Há também nomes de músicas entre aspas por conter caracteres especiais como vírgulas ou aspas. Ex:

"Peso Pluma: Bzrp Music Sessions,Vol.55",**"Bizarrap,Peso Pluma"**,2,2023,

Você deve ignorar essas linhas, e terá portanto que propor uma verificação para identificá-las.

Ao final imprima a lista produzida. Ex:

```
[[ 'When I Was Your Man', 'Bruno Mars', 2012, 1661187319],  
 [ 'I Wanna Be Yours', 'Arctic Monkeys', 2013, 1297026226],  
 ...,  
 [ 'As It Was', 'Harry Styles', 2022, 2513188493]]
```