



23%

ATIVIDADES
8 de 10DISCORD
ALURAFÓRUM DO
CURSOVOLTAR
PARA
DASHBOARDMODO
NOTURNO

16.1k xp

a

☰ 08

Separando apenas palavras.

Vimos que é necessário criar uma lista apenas com as palavras do nosso **corpus** para “treinar” o corretor ortográfico. Como no processo de tokenização geramos uma lista de sequência de caracteres e não apenas palavras, é preciso gerá-la.

```
def separa_palavras(lista_tokens):  
    lista_palavras = []  
    for token in lista_tokens:  
        if token.isalpha():  
            lista_palavras.append(token)  
    return lista_palavras
```

[COPIAR CÓDIGO](#)

Considerando o trecho de código acima, assinale a alternativa com a devida entrada `lista_tokens` da função `separa_palavras()` para gerar a seguinte lista de saída:

```
['Olá', 'tudo', 'bem']
```

A

Para gerar `['Olá', 'tudo', 'bem']` a entrada deve ser `['Olá', ',', 'tudo12', 'bem', '?']`





23%

ATIVIDADES
8 de 10

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



16.1k xp

a



Para gerar ['Olá', 'tudo', 'bem'] a entrada deve ser ['Olá', ',',
'tudo', 'bem', '?']



Parabéns, isso mesmo! A função `isalpha()` verifica se todos os caracteres do token são alfabéticos, se o retorno for verdadeiro ele é armazenado na variável `lista_palavras`.

C

Para gerar ['Olá', 'tudo', 'bem'] a entrada deve ser ['Olá', ',
'tudo', 'bem?']



PRÓXIMA ATIVIDADE