

# Exercícios BeautifulSoup4 - Quotes to Scrape

**Prof. Thiago Bluhm**

**Site:** <http://quotes.toscrape.com>

**Objetivo:** Aprender web scraping com BeautifulSoup4 através de exercícios práticos passo a passo.

---

## Preparação Inicial


```
from urllib.request import urlopen
from bs4 import BeautifulSoup

# Abrir conexão com o site
html = urlopen("http://quotes.toscrape.com")

# Instanciar BeautifulSoup
soup = BeautifulSoup(html, 'html.parser')
```

---

## EXERCÍCIO 1: Configuração Básica

 **Questão:** Abra o site <http://quotes.toscrape.com> usando `urlopen` e crie uma instância do BeautifulSoup. Imprima o título da página.

### Código:


```
from urllib.request import urlopen
from bs4 import BeautifulSoup

# Seu código aqui
```

 **Resultado esperado:** Título da página: Quotes to Scrape


---

## EXERCÍCIO 2: Primeira Quote

 **Questão:** Encontre e imprima o texto da primeira quote (citação class 'text') da página.


### Código:

```
# Seu código aqui (use soup.find)
```

 **Resultado esperado:** "The world as we have created it is a process of our thinking..."

---

## EXERCÍCIO 3: Autor da Primeira Quote

 **Questão:** Encontre e imprima o nome do autor da primeira quote.


### Código:

```
# Seu código aqui (procure por small e class author)
```

 **Resultado esperado:** Autor: Albert Einstein

---

## EXERCÍCIO 4: Tags da Primeira Quote

 **Questão:** Encontre e imprima todas as tags (palavras-chave) da primeira quote.

### Código:


```
# Seu código aqui (procure por div com class tags)
```

 **Resultado esperado:**

```
Tags da primeira quote:
- change
- deep-thoughts
- thinking
- world
```


---

## EXERCÍCIO 5: Contar Total de Quotes

 **Questão:** Conte quantas quotes existem na primeira página.


### Código:

```
# Seu código aqui (use find_all)
```

 **Resultado esperado:** Total de quotes na página: 10

---

## EXERCÍCIO 6: Lista de Todos os Autores

 **Questão:** Crie uma lista com todos os autores únicos da página (sem repetições).


 **Código:**

```
# Seu código aqui (use set para eliminar duplicatas)
```

 **Resultado esperado:** Lista de 10 autores únicos ordenados alfabeticamente.

---


## EXERCÍCIO 7: Quote Mais Longa

 **Questão:** Encontre qual é a quote mais longa (maior número de caracteres) e imprima ela com seu autor.

 **Código:**

```
# Seu código aqui
```


## EXERCÍCIO 8: Buscar por Autor Específico

 **Questão:** Encontre todas as quotes de "Albert Einstein" na página.

 **Código:**

```
# Seu código aqui
```


## EXERCÍCIO 9: Tags Mais Populares

 **Questão:** Conte quantas vezes cada tag aparece na página e mostre as 5 mais populares.

 **Código:**

```
# Seu código aqui (use Counter ou dicionário)
```


## EXERCÍCIO 10: Navegação - Links do Menu

 **Questão:** Encontre todos os links de navegação no menu superior da página.

 **Código:**

```
# Seu código aqui (procure por nav ou ul)
```


## **EXERCÍCIO 11: Extrair Informações Completas**

 **Questão:** Crie um dicionário com todas as informações de cada quote (texto, autor, tags).

 **Código:**

```
# Seu código aqui
```


## **EXERCÍCIO 12: Buscar Quote por Palavra-chave**

 **Questão:** Encontre todas as quotes que contêm a palavra "life" (vida).

 **Código:**

```
# Seu código aqui
```


## **EXERCÍCIO 13: Verificar Estrutura da Página**

 **Questão:** Encontre e imprima a estrutura básica da página (header, main, footer).

 **Código:**

```
# Seu código aqui
```

## **EXERCÍCIO 14: CSS Selectors**

 **Questão:** Use CSS selectors para encontrar todas as quotes e seus autores.

 **Código:**

```
# Seu código aqui (use soup.select)
```


## **EXERCÍCIO 15: Relatório Final Completo**

 **Questão:** Crie um relatório completo com estatísticas da página.











 **Código:**

```
# Seu código aqui - crie um relatório completo
```

## EXERCÍCIO BÔNUS: Salvar em CSV

 **Questão:** Salve todas as quotes em um arquivo CSV.

### Conceitos Aprendidos

-  `urlopen()` - Abrir conexões web
-  `BeautifulSoup()` - Parser HTML
-  `find()` - Buscar primeiro elemento
-  `find_all()` - Buscar todos os elementos
-  `select()` - CSS selectors
-  Navegação por atributos e classes
-  Extração de texto e atributos
-  Manipulação de dados extraídos
-  Análise estatística dos dados
-  Exportação para CSV

 **Parabéns!**