

WEB SCRAPING WITH PANDAS

Level: Easy

Time: Fast

Tools: Python, Pandas

Know how to choose a website

Web scraping with Pandas is very basic. It can be done using the `pd.read_html()` function, which actually uses beautiful soup, one of the most known libraries for web scraping.

The function will look for tables within the content to bring you the result, but as it is ready for use, there is only so much it will be able to do, thus the method won't do well with complex tables or custom content to be retrieved.

You should aim for simple tables embedded in websites, like this one where you have a content (text), then a table and more content.

← → ↺ 🏠 pt.wikipedia.org/wiki/Fundo_Monetário_Internacional

Cotas

Cada país membro detém no FMI uma cota a ser determinada com base em seus indicadores econômicos, entre eles o PIB. Quanto maior a contribuição ao FMI, maior é o peso do voto nas decisões. Há uma revisão geral das cotas a cada cinco anos. O Fundo pode propor um aumento nas cotas de determinado país, mas é necessária a aprovação por 85 % dos votos para qualquer modificação. Os membros que queiram aumentar sua cota devem pagar ao Fundo a mesma quantia em DES correspondente ao aumento.

Os cinco maiores acionistas são: Estados Unidos, Alemanha, Japão, França e Reino Unido.

Cada país pode retirar 25 % de sua cota correspondente. Acima deste percentual, é preciso assinar um termo (carta de intenções, atrelada geralmente a um memorando técnico de entendimento) onde se compromete a reduzir o déficit fiscal e promover a estabilização monetária.

A partir de 1980, o FMI passa a funcionar como supervisor da dívida externa.

Recentemente, o combate à pobreza mundial vem-se tornando uma preocupação central.

• Dez maiores cotistas

| Posição | País Membro | Cotas (milhões DES) | % das cotas |
|---------|----------------|---------------------|-------------|
| 1º | Estados Unidos | 37 149,30 | 17,46 |
| 2º | Japão | 13 312,80 | 6,26 |
| 3º | Alemanha | 13 008,20 | 6,11 |
| 4º | Reino Unido | 10 738,50 | 5,05 |
| 5º | França | 10 738,50 | 5,05 |
| 6º | Itália | 7 055,50 | 3,32 |
| 7º | Arábia Saudita | 6 985,50 | 3,28 |
| 8º | China | 6 369,20 | 2,99 |
| 9º | Canadá | 6 369,20 | 2,99 |
| 10º | Rússia | 5 945,40 | 2,79 |
| 17º | Brasil | 3 036,10 | 2,46 |
| 42º | Portugal | 1 029,70 | 0,43 |

• Juntos, os 10 primeiros possuem 55,3% da capacidade total de votos.

Formas de financiamento

- SBA - Acordo de crédito contingente ou acordo *stand-by* (*Stand-by agreement*) - é a política mais comum de empréstimos do FMI. É utilizada desde 1952 em países com problemas de curto prazo na balança de pagamentos. Essa política envolve apenas o financiamento direto de 12 a 18 meses. O prazo de pagamento vai de três a cinco anos. São cobrados juros fixos de 2,22% mais uma taxa variável que pode chegar a 2%.
- ESF - Programa de Contenção de choques externos (*Exogenous Shocks Facility*) - Crises e/ou conflitos temporários vinculadas a outros países e que influem no comércio, flutuações no preço de commodities, desastres naturais. Duraam de 1 a 2 anos. Força apenas nas causas do choque.

The code

The code is really simple.

```
# Web scraping: just use pd.read_html('web address')
page = pd.read_html('https://pt.wikipedia.org/wiki/Fundo_Monet%C3%A1rio_Interna
cional')
```

It will return a list for you

```
type(page)
[OUT]: list
```

If you print it on the screen, you will see that we brought all the content from the page.

```
page
0 NaN Esta página cita fontes, mas estas não cobrem ..., Unnamed: 0 NaN
1 Mapa-múndi destacando os países-membros do FMI... Mapa-múndi destacando os países-membros do FMI... Unnamed: 1
2 Organização internacional
3 22 de junho de 1944 (77 anos)
4 Sede Washington, D.C. Estados Unidos
5 Membros Estados da ONU e Kosovo
6 Línguas oficiais Inglês, francês e espanhol
7 Filiação Organização das Nações Unidas (ONU)
8 Diretora-geral Kristalina Gueorguieva
9 Sítio oficial www.imf.org,

Membros eleitos Grupos
0 Bélgica Áustria, Bielorrússia, Bélgica, Hungria, Cazaq...
1 Países Baixos Bósnia e Herzegovina, Bulgária, Croácia, Chipr...
2 México Costa Rica, El Salvador, Guatemala, Honduras, ...
3 Itália Albânia, Grécia, Itália, Malta, Portugal, San ...
4 Canadá Antígua e Barbuda, Bahamas, Barbados, Belize, ...
5 Finlândia Dinamarca, Estônia, Finlândia, Islândia, Letôn...
6 Coreia do Sul Austrália, Kiribati, Coreia do Sul, Ilhas Mars...
7 Egito Bahrein, Egito, Iraque, Jordânia, Kuwait, Líba...
8 Malásia Brunei, Camboja, Ilhas Fiji, Indonésia, Laos, ...
9 Tanzânia Angola, Botswana, Burundi, Eritreia, Etiópia, ...
10 Suíça Azerbaijão, Quirguistão, Polónia, Sérvia e Mon...
11 Irão Afeganistão, Argélia, Gana, Irã, Marrocos, Paq...
12 Brasil Brasil, Colômbia, República Dominicana, Equado...
13 Índia Bangladesh, Butão, Índia, Sri Lanka
14 Argentina Argentina, Bolívia, Chile, Paraguai, Peru, Uru...
15 Guiné Equatorial Benim, Burquina Fasso, Camarões, Cabo Verde, C...,

Posição País Membro Cotas (milhões DES) % das cotas
0 1º Estados Unidos 37 149,30 1746
1 2º Japão 13 312,80 626
2 3º Alemanha 13 008,20 611
3 4º Reino Unido 10 720,50 505
```

Now we need to find the table we want: ten biggest shareholders from IMF.

To do that, we will look at each item. And it happens to be on the index 3.



```
# Shareholders: index 3  
page[3]
```

| | Posição | País Membro | Cotas (milhões DES) | % das cotas |
|----|---------|----------------|---------------------|-------------|
| 0 | 1º | Estados Unidos | 37 149,30 | 1746 |
| 1 | 2º | Japão | 13 312,80 | 626 |
| 2 | 3º | Alemanha | 13 008,20 | 611 |
| 3 | 4º | Reino Unido | 10 738,50 | 505 |
| 4 | 5º | França | 10 738,50 | 505 |
| 5 | 6º | Itália | 7 055,50 | 332 |
| 6 | 7º | Arábia Saudita | 6 985,50 | 328 |
| 7 | 8º | China | 6 369,20 | 299 |
| 8 | 9º | Canadá | 6 369,20 | 299 |
| 9 | 10º | Rússia | 5 945,40 | 279 |
| 10 | 17º | Brasil | 3 036,10 | 246 |
| 11 | 42º | Portugal | 1 029,70 | 43 |

```
# Now, just save it in a variable and use for your project  
df = page[3]  
  
type(df)  
pandas.core.frame.DataFrame
```