

## Web Scrapping

### Librerias

```
1 import requests
2 from bs4 import BeautifulSoup
3 import pandas as pd
```

### Funciones

```
1 url = 'https://books.toscrape.com/catalogue/page-1.html'
2
3 libros = []
4
5 def extraer_data_libros(url):
6     res = requests.get(url)
7     soup = BeautifulSoup(res.content, 'html.parser')
8
9     for libro in soup.find_all('article', class_ = 'product_pod'):
10         titulo = libro.h3.a['title']
11         #print(titulo)
12         precio = libro.find('p', class_ = 'price_color').text
13         #print(precio)
14         disponibilidad = libro.find('p', class_ = 'instock availability').text.strip()
15         #print(disponibilidad)
16         libros.append({
17             'Titulo': titulo,
18             'Precio': precio,
19             'Disponibilidad': disponibilidad
20         })
21     return libros
```


## Web Scrapping

```
1 url = 'https://books.toscrape.com/catalogue/page-{}.html'
2
3 datos_libros = []
4
5 for pag in range(1,50): # Podemos aumentar el rango a 50
6     url1 = url.format(pag)
7     #print(url1)
8     datos_libros.extend(extraer_data_libros(url1))
9     print(datos_libros)
```

🔗 [{"Titulo": "A Light in the Attic", "Precio": "£51.77", "Disponibilidad": "In stock"}, {"Titulo": "Tipping the Velvet", "Precio": "£53.7"}]

```
1 df_libros = pd.DataFrame(datos_libros)
```

```
1 df_libros
```



	Titulo	Precio	Disponibilidad
0	A Light in the Attic	£51.77	In stock
1	Tipping the Velvet	£53.74	In stock
2	Soumission	£50.10	In stock
3	Sharp Objects	£47.82	In stock
4	Sapiens: A Brief History of Humankind	£54.23	In stock
...	...	...	...
29395	Icing (Aces Hockey #2)	£40.44	In stock
29396	Hawkeye, Vol. 1: My Life as a Weapon (Hawkeye #1)	£45.24	In stock
29397	Having the Barbarian's Baby (Ice Planet Barbar...	£34.96	In stock
29398	Giant Days, Vol. 1 (Giant Days #1-4)	£56.76	In stock
29399	Fruits Basket, Vol. 1 (Fruits Basket #1)	£40.28	In stock

29400 rows × 3 columns

1 Start coding or [generate](#) with AI.