

# Beautiful Soup

---

# Beautiful Soup



Es una biblioteca de Python para analizar documentos HTML.

```
1 <html>
2   <head>
3     <meta charset="utf-8"/>
4     <title>Esto es un ejemplo</title>
5   </head>
6   <body>
7     <h1>Este es el título 1</h1>
8     <h2>Este es el título 2</h2>
9     <h3>Este es el título 3</h3>
10  </body>
11 </html>
```

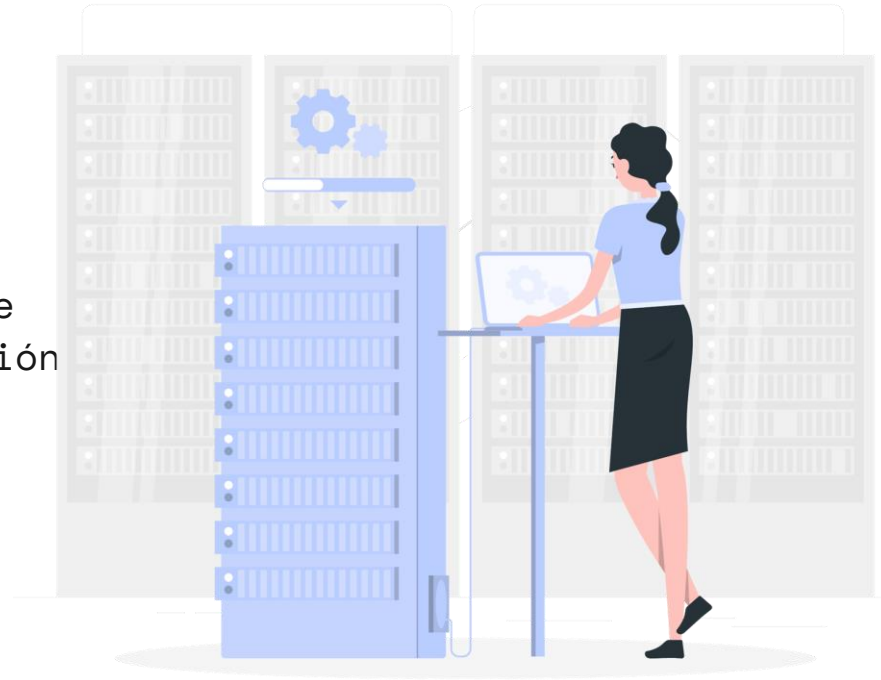
**Este es el título 1**

**Este es el título 2**

**Este es el título 3**

# Web Scraping

o raspado web, es una técnica utilizada mediante programas de software para extraer información de sitios web.

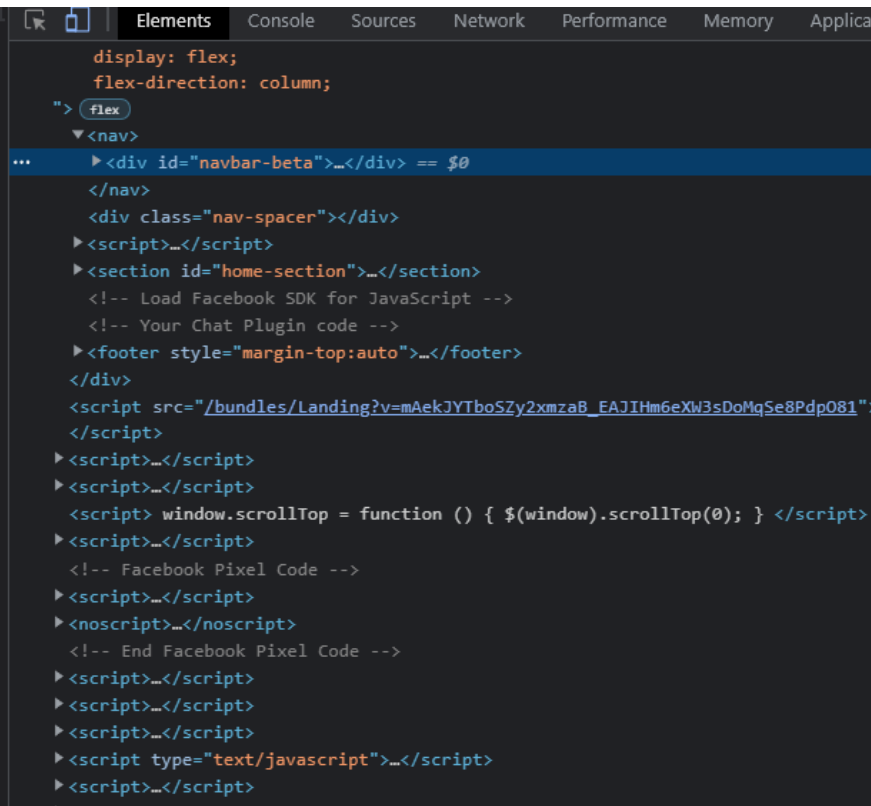
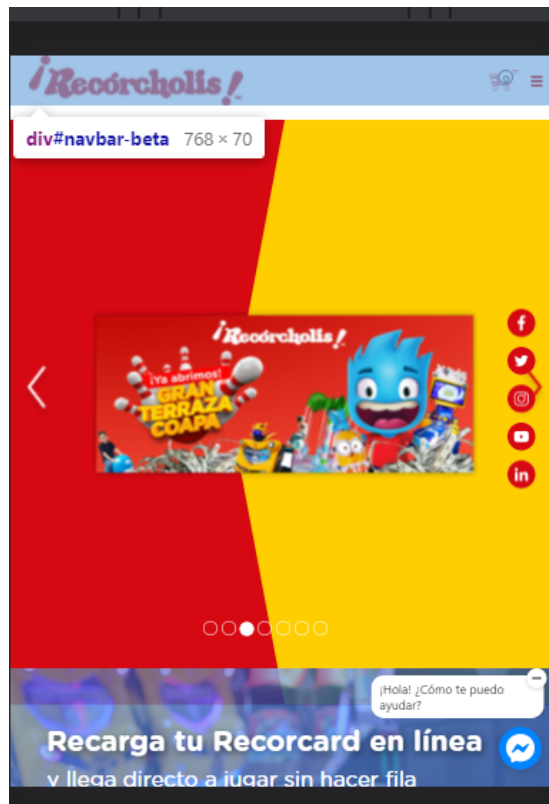


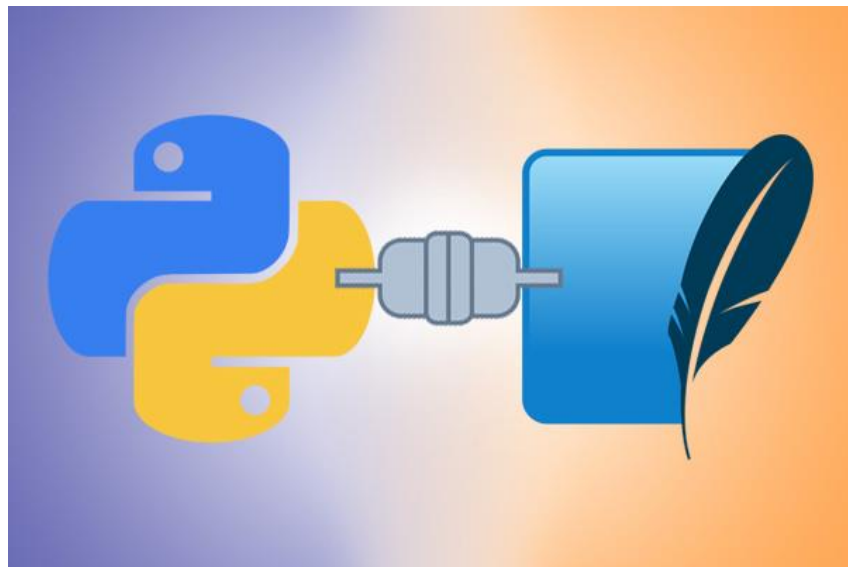
# HTML

HTML, siglas en inglés de HyperText Markup Language, hace referencia al lenguaje de marcado para la elaboración de páginas web.



# EJEMPLOS





# Base de datos

---

# Base de datos

es un conjunto de información que se relaciona entre sí, que está almacenada y organizada de forma sistemática para facilitar su preservación, búsqueda y uso.



# Conceptos fundamentales

Emp_Id	Last_Name	First_Name	Gender	Title
1000	Torbati	Yolanda	F	Programmer
1001	Kleinn	Joel	M	Programmer
1002	Ginsburg	Laura	F	President
1003	Cox	Jennifer	F	Programmer
1005	Ziada	Mauri	M	Product Designer
1006	Keyser	Cara	F	Account Executive
1010	Smith	Roxie	M	Programmer
1011	Nelson	Robert	M	Programmer
1012	Sachsen	Lars	M	Support Technician
1013	Shannon	Don	M	Product Designer

## Tabla

Ente en el cual se  
almacenaran  
datos(registros)

## Registros

Cada registro es  
representado como  
una fila de nuestra  
tabla

## Campos

Se representan como  
las columnas que  
fraccionan la info  
de mi registro



# SQL vs. NoSQL



## NoSQL



Gaming



Social



IoT



Web



Mobile



Enterprise



Key/value store



Document database



Column family store

## SQL



Web



Mobile



Enterprise



Data mart



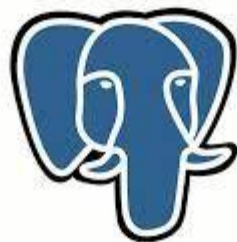
Relational table storage



Relationships use joins

# Bases de datos

**ORACLE®**  
DATABASE



PostgreSQL



Microsoft®  
SQL Server®



SQLite

# Sqlite con Python

---



Sqlite es una herramienta de software libre, que permite almacenar información en dispositivos con pocas capacidades de hardware de una forma sencilla.

# Características

- La base de datos completa se encuentra en un solo archivo.
- Es totalmente autocontenida (sin dependencias externas).
- Cuenta con librerías de acceso para muchos lenguajes de programación.
- El código fuente es de dominio público y se encuentra muy bien documentado.
- Multiplataforma

# Aplicaciones

- Cuando se requiere una base de datos integrada dentro de una aplicación.
- Como cache local de un manejador de base de datos empresarial.
- Para aplicaciones en dispositivos móviles que manejan una BD local
- Almacenamiento persistente de objetos, configuraciones y preferencias de usuario.



# CRUD

Es un acrónimo que en informática significa:

