

Python: Key projects

SCRAPING

- `requests.post()`
- Es un metodo para extraer datos de la web
- [Link reference](#)

STEPS:

1. Extraer el archivo HTML de la pagina
2. Determinar la ruta de dato a extraer
3. Configurar beautifulsoup `pip install bs4`
4. Extraer dato requerido

ERROR más frecuente

- Cuando la página realiza cambios en su estructura; puede romper la estructura del programa.

SUGGESTIONS

- Antes de usar SCRAPING, se sugiere usar una API (COMO PRIMERA OPCIÓN).
- Los ERRORES más peligrosos son los ERRORES SILENCIOSOS
- Usar dos rutas para validar el SCRAPING.
- Se puede usar SCRAPING cuando deseas ofrecer un valor agregado y no afecta el funcionamiento del sistema.

References

- [How To Use *args and **kwargs in Python 3](#)

FUTURE Topics

- `requests.post()` para peticiones POST
- Selenion
- Romper captcha
- Impedir scraping a tu pagina