
Web Scraper for real estate website

Last update on

Jan 12, 2026

Khaled Alruwita

(1) الفكرة

المشروع عبارة عن **سكريب (Web Scraper)** يسحب بيانات العقارات من موقع **عقار** بشكل آلي، ويخزنها في **داتابيس SQLite** تقدر تستخدمها لاحقاً في التحليل أو العرض أو أي مشروع آخر.

السكريب مصمم إنه:

- يشتغل بشكل مستمر
- يتعامل مع الصفحات الديناميكية (بفضل **Playwright**)
- يتعامل مع الانقطاعات
- ما يكرر البيانات

(2) الأدوات المستخدمة:

Playwright (1)

- يشغل متصفح حقيقي (Chromium)
- يفتح الصفحات كأنك مستخدم فعلي
- يتعامل مع JavaScript (تحميل، scroll، أزرار)
- مناسب لمواقع ما تنفع معها requests العادية

BeautifulSoup (2)

- بعد ما الصفحة تتحمل، نستخدمه لتحليل ال HTML

- استخراج:

- السعر
- العنوان
- نوع العقار
- المساحة
- الرابط
- التاريخ

SQLite (3)

- قاعدة بيانات خفيفة كوني بستخدمها على جهاز ضعيف مثل Raspberry pi
- ملف واحد (`aqar.sqlite`)

(3) كيف يمشي السكريبر ؟

(1) تشغيل السكريبر

أول ما تشغل الملف:

- يفتح المتصفح عن طريق Playwright
- يروح لصفحة العقارات الرئيسية
- يجزّز الاتصال بقاعدة البيانات

(2) تحميل الصفحة

- ينتظر لين الصفحة تكتمل (حق العناصر تظهر)
 - يسوي scroll عشان يحمل إعلانات أكثر
 - يتأكد إن المحتوى صار موجود فعلاً قبل التحليل
-

(3) قراءة الإعلانات

لكل إعلان عقار:

- يقرأ البيانات من ال HTML
 - ينظف النصوص (أرقام عربية → إنجليزية)
 - يتأكد إن القيم مو فاضية
-

(4) منع التكرار

قبل ما يحفظ الإعلان:

- يفحص هل الرابط موجود مسبقاً في SQLite
 - إذا موجود → يتجاهله
 - إذا جديد → يكمله
-

(5) التخزين في SQLite

يحفظ الإعلان في جدول مثل:

- id
- title
- price
- area
- city
- url
- created_at

كل إعلان جديد ينضاف مباشرة للقاعدة.

4) رابط المشروع في GIT

<https://github.com/KhaledAlruwita/realestate-scraper>