**Comics Price Guide Scraper**

**Overview**

This Python script automatically gathers comic book data from [comicspriceguide.com](https://www.comicspriceguide.com). It logs into the website, scrapes comic series and issue details, and saves the collected information into an Excel file.

**Key Features**

* **Login:** Automatically logs into the site using your credentials.
* **Scraping Levels:**
  + **2-Level Mode (Grid → Detail):**  
    The script goes directly from the grid (list of issues) to each detail page.
  + **3-Level Mode (Main Series → Grid → Detail):**  
    The script first gathers a list of series from a main page, then visits each grid page for that series, and finally extracts detailed data for each issue.
* **Data Collected:**  
  Comic title, issue number, publisher, publication date, "Years" (e.g., "2021-Current"), volume, country, variant, edition, and image URL.

**How to Use It**

**1. Configuration**

* **SCRAPE\_DEPTH:**
  + Set to **2** if you want the script to go directly from the grid page to the detail pages.
  + Set to **3** if you want the script to first collect series links from a main page, then the grid, then the details.

You can change this option near the top of the script:

python

Copy

SCRAPE\_DEPTH = 3 # Use 2 for grid→detail; 3 for main series→grid→detail.

* **TARGET\_URLS:**  
  Add one or more URLs to the list. For example:

python

Copy

TARGET\_URLS = [

"https://www.comicspriceguide.com/genres/erotica/un",

"https://www.comicspriceguide.com/genres/satire/um"

]

This allows you to scrape multiple sections of the website.

* **ENABLE\_VARIANTS\_CHECK:**  
  Set this to 1 if you need to click the "Show Variants" checkbox on grid pages. Otherwise, leave it as 0.
* **Login Credentials:**  
  The script uses hard-coded credentials in the login\_to\_site function. Update the username and password in the script if needed.

**2. Running the Script**

1. **Install Requirements:**  
   Make sure you have Python 3 and install the required packages:

bash

Copy

pip install selenium pandas webdriver-manager

1. **Run the Script:**  
   Open your terminal, navigate to the directory with the script, and run:

bash

Copy

python your\_script\_name.py

1. **Output:**
   * The script logs its progress to the console and to scraped\_comics.log.
   * Once finished, it produces an Excel file named scraped\_comics.xlsx with all the collected data.
   * A live update file (scraped\_comics\_live.xlsx) is also created periodically.

**How It Works (In Simple Terms)**

* **Login:** The script opens the website, clicks the "Login" button, enters your username and password, and waits until the login popup disappears.
* **Scraping Process:**
  + In **3-level mode**, the script first visits a main page to collect series links.
  + For each series (or if using 2-level mode), it visits a grid page that lists comic issues.
  + The grid page is scanned to get links and extra information like the publication date and "Years" value.
  + Then, the script visits each detail page for an issue to grab more specific details.
* **Data Storage:** All information is compiled into a table (using Pandas) and saved as an Excel file.

**Final Notes**

* **Editing:**  
  You can change the target URLs, choose between 2-level or 3-level scraping, and adjust other settings at the top of the script.
* **Use Cases:**  
  This tool is useful if you want to build a database of comic book issues for analysis, price tracking, or research.
* **Troubleshooting:**  
  If you experience issues, check the log file (scraped\_comics.log) for error messages and adjust wait times or configuration options as needed.