



XPLOR12

DOCUMENTATION

Using BeautifulSoup, SQLite

Prepared by
Müller István
@codeatn8

Introduction:

This Python code is a web scraper developed by Stefan Müller to scrape prices and direct links of items available on the website Fuego.sk. It uses the requests and BeautifulSoup libraries to make HTTP requests and parse HTML content respectively. It also uses SQLite3 to create a database and store the scraped data.

Functionality:

The code scrapes prices and direct links of items from the website Fuego.sk for over 25 pages and 250+ items. It then stores the scraped data in a SQLite3 database called fuego_data.db. The code iterates over multiple pages of the website and for each item on each page, it extracts the name, price and URL of the item. It then inserts this data into the SQLite3 database.

Usage:

To use this code, save it as a Python file with a .py extension, and run it in a Python environment. The following libraries should be installed:

- sqlite3
- time
- datetime
- requests
- BeautifulSoup
- colorama

When the code is executed, it will print a welcome message, connect to the SQLite3 database and create a new table to store the items and prices. It will then iterate over multiple pages of the website and for each item, it will extract the name, price and URL and insert this data into the SQLite3 database. Finally, it will commit the changes to the database and close the connection. It will also print the time taken for the entire scraping process and the timestamp when the scraping was completed.

Limitations:

This code is designed to scrape data from the website Fuego.sk only. It may not work for other

websites due to differences in their HTML structure. Also, scraping large amounts of data may result in IP blocking or other issues. Therefore, this code is intended for personal use only and not for commercial purposes.

Welcome XPLORER™!

XPLOR™ is a webscraper Developed by Stefan Müller @codeatn8!
ONLY FOR PERSONAL USE! All items are available for public!
XPLOR™ Scrapes prices and direct links from website Fuego.sk
Over 25 pages and 250+ items & stores it in a .DB file.

LETS ROCK BABY!

XPLOR™ >> Page 1 was succesfully scraped!
XPLOR™ >> Page 2 was succesfully scraped!
XPLOR™ >> Page 3 was succesfully scraped!
XPLOR™ >> Page 4 was succesfully scraped!
XPLOR™ >> Page 5 was succesfully scraped!
XPLOR™ >> Page 6 was succesfully scraped!
XPLOR™ >> Page 7 was succesfully scraped!
XPLOR™ >> Page 8 was succesfully scraped!
XPLOR™ >> Page 9 was succesfully scraped!
XPLOR™ >> Page 10 was succesfully scraped!
XPLOR™ >> Page 11 was succesfully scraped!
XPLOR™ >> Page 12 was succesfully scraped!
XPLOR™ >> Page 13 was succesfully scraped!
XPLOR™ >> Page 14 was succesfully scraped!
XPLOR™ >> Page 15 was succesfully scraped!
XPLOR™ >> Page 16 was succesfully scraped!
XPLOR™ >> Page 17 was succesfully scraped!

QNA

Do I need to install XPLO12 application?

No, there is a XPLO12.exe RUN file.

What contains XPLO12 folder?

XPLO12.exe & assets file folder which contains the icons, logos and project files.

Is internet connection necessary for XPLO12?

Yes, it is. The speed of scraping is based on multiple aspects:

- internet connection stability
- connection speed
- size of data

How to give an INPUT for XPLO12?

Application XPLO12 is a web scraper for one specific website www.fuego.sk/elektronika. There's no need to give an input because the scraping is automatic until you run the application. The application at the end of the process (when it reaches 25 pages) will make an output file .db (database - SQLite) in your XPLO12 Folder (where the application .exe is).

What is the OUTPUT of XPLO12?

It's a .DB database file (id, name, price, url). Id, name, url from the website, prices in Euro (€).

Application will give you two output files. The "fuego_data.db-journal" and "fuego_data.db". You will need to open "fuego_data.db" using SQL Viewer.

What if I had to exit from the application?

There's no problem. The fuego_data.db stays clean until your process ends successfully.

HOWEVER, IF YOU REPEAT THE PROCESS: XPLO12 will repeat the whole process, so you will get the same 25 pages scraped under your previous results. If the items on the page get refreshed, you will see a difference between the results. XPLO12 will not overwrite any of your previous results.

What if it is not working?

You will need to check your internet connection and speed, firewall, access to applications.

What about projectsource_folder?

Folder "projectsource_folder" contains [xplorasset.pdn](#) and [xplorico.pdn](#).

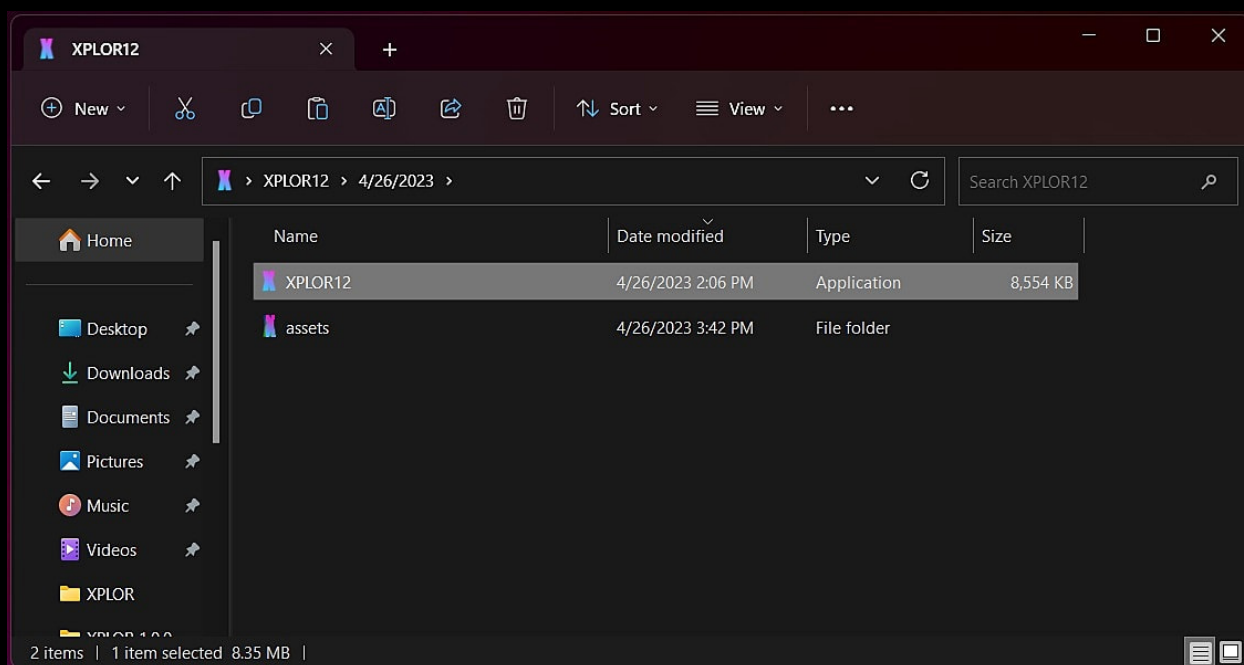
[.PDN \(Paint.NET\)](#) - open source software with two types of logos.

[xplorasset.pdn](#) contains the asset folder elements.

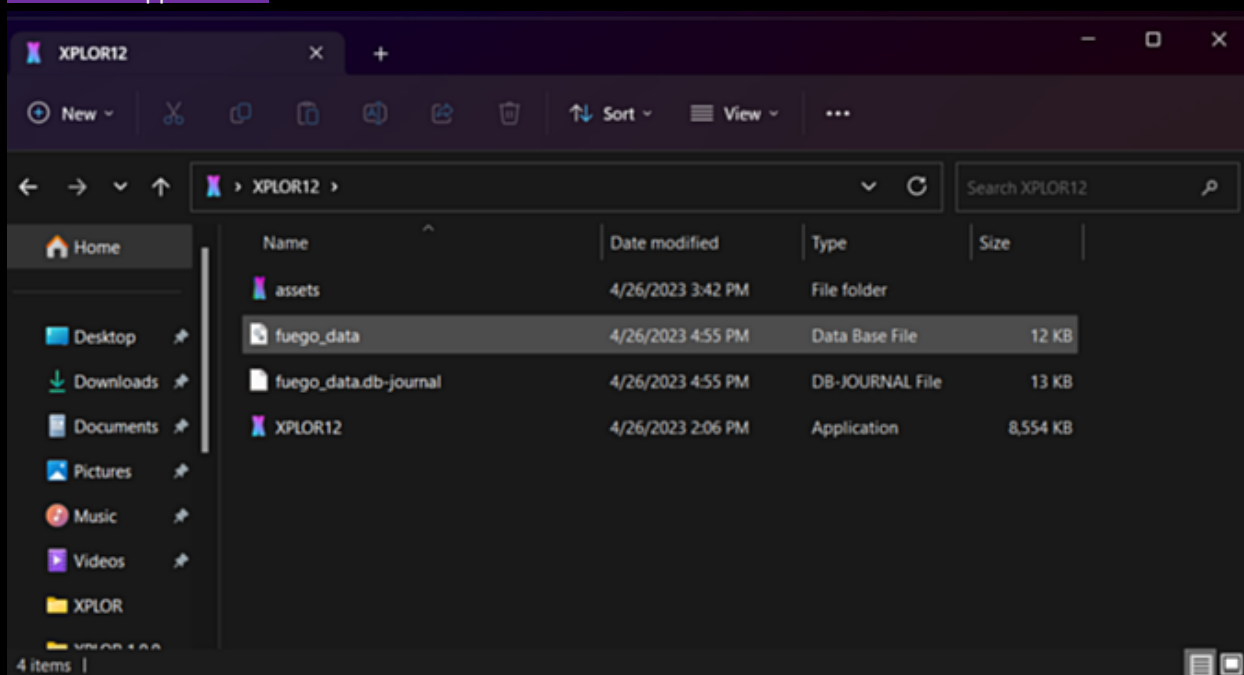
[xplorico.pdn](#) contains the application logo.

What if the application stuck?

You'll need to check your internet connection stability and speed, then restart. No files will be corrupted.



Picture: Unzipped folder



Picture: Output files