Web Scraping Script Report

Overview

This Python script is used to scrape the latest hot deals from the RedFlagDeals forums.

It automates the process of gathering deal-related data such as store names, item details,

votes, usernames, timestamps, categories, and more. The primary purpose of the script is

to help users quickly access and analyze deal information without manually browsing through

multiple forum pages.

The script uses two main libraries:

- requests: To send HTTP requests and fetch web content.

- BeautifulSoup: To parse and extract data from the HTML content.

The script is divided into several key functions, each with specific tasks, and includes

an interactive menu system for the user to choose the desired functionality.

How to Use

1. Installation:

   - Install the required Python libraries by running:

     pip install requests beautifulsoup4

   - Ensure you have Python 3.x installed.

2. Running the Script:

   - Execute the script from the command line or any Python IDE.

- After execution, you will be presented with a menu containing five options:

  1. Display Latest Deals

  2. Analyze Deals by Category

  3. Find Top Stores

  4. Log Deal Information

  5. Exit

3. Menu Options:

  - Choose an option by entering the corresponding number.

  - For example, selecting "1" will display the latest deals, while "3" will allow

    you to input the number of top stores you want to display.

  - You can continue using the script until you choose the "Exit" option by selecting "5."

4. Output:

  - Deal information is displayed directly in the terminal.

  - If you choose the "Log Deal Information" option, URLs of deals are logged into a text file.

Important Functions

1. display_latest_deals():

  - This function fetches the latest deals from the RedFlagDeals forum page using the requests

library.

  - It parses the HTML content using BeautifulSoup and extracts details such as:

   * Store Name

   * Item Name

   * Votes

* Username of the post author

* Timestamp of the deal post

* Category of the deal

* Number of replies and views

* URL of the deal

- The function prints all the extracted details in a well-formatted manner.

2. analyze_deals_by_category():

- This function analyzes the deals based on their category.

- It fetches the same forum page as the previous function, but instead of displaying all deals,

  it counts how many deals belong to each category.

- The categories and their respective deal counts are printed, helping the user identify popular

  categories on the forum.

3. find_top_stores():

- This function identifies the top stores based on the number of deals they have posted.

- The user is prompted to input the number of top stores they wish to see.

- The function extracts store names from the forum page, counts the deals per store, and displays

  the top stores with the most deals.

4. log_deal_information():

- This function allows users to log the URLs of deals from a specific category to a text file.

- After scraping the forum, the function displays a list of categories to the user.

- The user selects a category, and the function logs all the deal URLs from that category into a

  file named 'log.txt'.

- This feature is useful for saving deals for later review.

5. exit_program():

   - This function provides a clean exit from the program when the user selects the "Exit" option

    from the menu.