

Question 4 (10 points):

In this task, you are required to identify and utilize online tools for web scraping data from websites without the need for coding, with a specific focus on Parsehub. The objective is to gather data and save it in formats like CSV, Excel, or any other suitable file format.

You have to mention an introduction to the tool which ever you prefer to use, steps to follow for web scrapping and the final output of the data collected.

Upload a document (Word or PDF File) in the same repository and you can add the link in the ipynb file.

For the given task I have used ParseHub tool for web scrapping to get some information from the below website.

[Amazon.com](https://www.amazon.com)

<https://www.amazon.com › Literature-Fiction-Books>

Introduction to ParseHub:

ParseHub is a robust online scraping tool that allows users to extract data from websites in a systematic and automated manner. It is a user-friendly, cloud-based platform that does not need substantial programming experience. ParseHub has a point-and-click interface, making it accessible to users with varied degrees of technical ability.

Steps to Follow for Web Scrapping with ParseHub:

- **Install ParseHub:** Begin by creating a ParseHub account at parsehub.com. We need to download and install the ParseHub desktop program in our PC once we have enrolled.
- **Start a New Project:** After installing ParseHub, we need to launch it and select "New Project.and need to " Enter the URL of the website to be scrapped.

- **select elements:** Using the ParseHub point-and-click interface, we need to pick the individual components (text, photos, links, etc.) from the web page that we wish to scrape. ParseHub will generate a "template" of the data to be extracted.
- **Create Rules:** Create rules for the selected components to personalize our scrape. To improve the data extraction process, we may rename fields, add filters, and construct conditional logic.
- **Run the Project:** To begin the scraping process, we need to click the "Run" button. Based on our instructions, ParseHub will explore the website and collect data.
- **evaluate and Modify:** After the initial scrape, we can evaluate the data and make any changes to the template and rules that are required if the results are not as expected.
- **plan and Automate:** For continuing scraping activities, ParseHub let us plan automated runs at certain intervals, ensuring that our data is constantly up to date.
- **Export Data:** Once we've finished scraping the data, we can export it in a variety of forms, including CSV, Excel, JSON, and API integration.
- **Data Collection Final Output:** The ultimate output of data obtained with ParseHub will be determined by our settings. It usually contains structured data in the format given during the scraping process. Through ParseHub's API, this data may be exported to a file format of our choosing or incorporated into other apps.

Results:

The below picture shows the results of Amazon data providing a comprehensive snapshot of product and pricing available on the platform. This data typically includes product, prices, ratings, author information. It also offers insights into customer reviews and ratings, allowing users to gauge product popularity.

Here product indicates “Book”

```
In [1]: import pandas as pd
```

```
In [2]: data = pd.read_csv("Web scrapping.csv")
```

```
In [3]: data
```

Out[3]:

	Product_Author	Product_Author_url	Product_Ratings	Product_Price	Product_Price_url
0	Stephen King	https://www.amazon.com/Stephen-King/e/B000AQ08...	4.4 out of 5 stars	17.0	https://www.amazon.com/Holly-Stephen-King/dp/1...
1	Book 1 of 2: The Emphyrean	https://www.amazon.com/dp/B0BV9WPS4M?binding=h...	4.8 out of 5 stars	19.0	https://www.amazon.com/Fourth-Wing-Emphyrean-Re...
2	Book 3 of 3: Knockemout	https://www.amazon.com/dp/B0BHF4KZ8N?binding=p...	4.7 out of 5 stars	13.0	https://www.amazon.com/Things-We-Left-Behind-K...
3	Bonnie Garmus	https://www.amazon.com/Bonnie-Garmus/e/B09964C...	4.6 out of 5 stars	15.0	https://www.amazon.com/Lessons-Chemistry-Novel...
4	Ann Patchett	https://www.amazon.com/Ann-Patchett/e/B000AQ6Q...	4.4 out of 5 stars	17.0	https://www.amazon.com/Tom-Lake-Novel-Ann-Patc...
5	Abraham Verghese	https://www.amazon.com/Abraham-Verghese/e/B000...	4.7 out of 5 stars	19.0	https://www.amazon.com/Covenant-Water-Abraham-...
6	Barbara Kingsolver	https://www.amazon.com/Barbara-Kingsolver/e/B0...	4.6 out of 5 stars	19.0	https://www.amazon.com/Demon-Copperhead-Novel-...
7	Book 1 of 2: The Housemaid	https://www.amazon.com/dp/B0CGXVCWY9?binding=p...	4.5 out of 5 stars	9.0	https://www.amazon.com/Housemaid-Freida-McFadd...
8	James McBride	https://www.amazon.com/James-McBride/e/B000AP9...	4.6 out of 5 stars	19.0	https://www.amazon.com/Heaven-Earth-Grocery-St...
9	Book 1 of 5: A Court of Thorns and Roses	https://www.amazon.com/dp/B07F7JQTCR?binding=p...	4.6 out of 5 stars	10.0	https://www.amazon.com/Court-Thorns-Roses-Sara...