

Q1. What is Web Scraping? Why is it Used? Give three areas where Web Scraping is used to get data.

Web Scraping:

Web scraping is the process of extracting data from websites. It involves fetching the web page and then extracting the required information from the HTML. The extracted data can be saved, analyzed, or used for various purposes.

Why it is Used:

Data Extraction: Extract specific information from websites.

Automation: Automate repetitive tasks like data entry or monitoring.

Research: Gather data for research and analysis.

Three Areas of Use:

E-commerce: Scraping product details, prices, and reviews from e-commerce websites.

Real Estate: Extracting property details, prices, and locations from real estate websites.

Job Market: Scraping job listings, requirements, and salaries from job portals.

In [ ]: Q2. What are the Different Methods Used for Web Scraping?

Manual Copy-Paste: Manually copying and pasting data.

Regular Expressions: Pattern-based extraction.

DOM Parsing: Parsing HTML using libraries like BeautifulSoup.

XPath: Navigating XML/HTML documents.

APIs: Using APIs provided by websites for data retrieval.

Q3. What is Beautiful Soup? Why is it Used?

Beautiful Soup:

Beautiful Soup is a Python library for pulling data out of HTML and XML files. It provides Pythonic idioms for iterating, searching, and modifying the parse tree. It sits on top of an HTML or XML parser and provides Pythonic ways of searching and navigating the parse tree.

Why it is Used:

Simplifies the process of pulling data from HTML or XML.

Provides a convenient way to navigate the parse tree.

Works well with various parsers, such as lxml and html5lib.

Q4. Why is Flask Used in This Web Scraping Project?

Flask is likely used to create a web application or API that exposes the scraped data to users. It provides a lightweight and easy-to-use framework for building web applications in Python. Flask can handle HTTP requests, allowing users to interact with the scraped data through a web interface or API.

Q5. Write the Names of AWS Services Used in This Project. Also, Explain the Use of Each Service.

The provided information doesn't mention specific AWS services used in the project. If you have details about the AWS services used, please provide them, and I can help explain their use. If not, you may want to check the project documentation or source code for details on AWS services integration.