

Q1.

""" Web scraping is the process of automatically extracting data from websites. It involves sending HTTP requests to web pages, parsing the HTML or other structured data on those pages, and then extracting specific information.

Uses :

- Data Collections
- Business Intelligence
- Research and Analysis
- Lead Generation
- Price Comparison

Specific areas :

- E-commerce and Price Comparison
- Real Estate and Property Listings
- Social Media and Sentiment Analysis

"""

Q2.

"""

- Manual Copy-Paste
- HTML Parsing
- XPath and CSS Selectors
- Regex
- Scrapping Frameworks and tools

"""

Q3.

""" BeautifulSoup is a Python library that is commonly used for web scraping and parsing HTML or XML documents. It provides a convenient way to navigate and manipulate the elements of a web page's Document Object Model (DOM).

- Parsing HTML and XML
- Simplifies Web Scrapping
- Element Navigation
- Data Extraction
- Open source and Widely Adopted

""

Q4.

""

- Data Representations
- Data Storage and Retrieval
- Automation
- Integration
- Visualization and Analytics
- API Endpoints

""

Q5.

""

- EC2 (Elastic Compute Cloud):

Use: EC2 instances can be used for running web scraping scripts or applications. These virtual machines provide compute capacity in the cloud, allowing you to run your scraping code on scalable and customizable instances.

- S3 (Simple Storage Service):

Use: S3 is an object storage service that can be used to store scraped data. You can save the data files, such as HTML, JSON, or CSV, in S3 buckets for easy access, sharing, and further processing.

- DynamoDB:

Use: DynamoDB is a NoSQL database service. You can use it to store structured data from your web scraping, such as metadata, URLs, or other data that requires fast and scalable database storage.

""