**Basic Web Scraping Questions (1-15):**

1. **What is web scraping?**  
   Can you explain the purpose of web scraping and give an example of how it's used?
2. **What are the legal and ethical considerations of web scraping?**  
   How do you ensure compliance when scraping websites?
3. **What is the requests library in Python?**  
   How do you use it to send an HTTP request to a website?
4. **What is BeautifulSoup?**  
   How do you use BeautifulSoup to parse HTML and extract data?
5. **What is the difference between HTML parsing and JSON parsing?**  
   When would you use each method?
6. **How do you handle websites that do not allow scraping through their robots.txt file?**
7. **What are some common status codes in HTTP?**  
   Explain what 200, 404, and 500 status codes mean in the context of web scraping.
8. **How would you handle an HTTP error while scraping data from a website?**  
   What steps would you take to retry or manage failed requests?
9. **How do you extract specific information, like headlines or product prices, from a website?**  
   Can you show how to extract data using HTML tags (e.g., <div>, <h1>) with BeautifulSoup?
10. **What is XPath?**  
    How is it used to select elements on a webpage in web scraping?
11. **How do you deal with websites that use pagination?**  
    Can you explain how to scrape data from multiple pages?
12. **How would you scrape data from a table on a webpage?**  
    What steps would you take to extract the rows and columns of the table?
13. **What is the role of headers in HTTP requests for web scraping?**  
    Why is it important to set headers like "User-Agent"?
14. **What is a "User-Agent" string, and how does it help in web scraping?**
15. **How would you save the data you scraped into a CSV file using Python?**

**Intermediate Web Scraping Questions (16-25):**

1. **What are some common Python libraries used for web scraping?**  
   Compare BeautifulSoup, Scrapy, and Selenium.
2. **What are the differences between requests and urllib in Python?**  
   When would you choose one over the other for web scraping?
3. **How do you scrape dynamic content (e.g., JavaScript-generated pages) using Python?**  
   How would you approach scraping a website that relies on JavaScript to load data?
4. **What is Selenium, and how is it used for web scraping?**  
   Give an example of automating a browser using Selenium.
5. **What are some common anti-scraping techniques used by websites?**  
   How would you bypass protections like CAPTCHAs, IP blocking, or dynamic content loading?
6. **How do you implement delays between requests to avoid getting blocked?**  
   Why is it important to introduce random delays in your scraping scripts?
7. **What is multithreading, and how can it improve the speed of your web scraping?**  
   Can you explain how to scrape multiple pages simultaneously using Python?
8. **How do you handle large datasets efficiently when scraping?**  
   What techniques or libraries would you use to store and process large amounts of scraped data?
9. **Can you explain how to scrape data from APIs in Python?**  
   What’s the difference between scraping an API and scraping a webpage?
10. **What would you do if a website you’re scraping suddenly changes its HTML structure?**  
    How would you modify your scraper to adapt to the changes?

What is regular expression?

What is key component in regex python?

What are the most common methods in regex?

Key component in web scraping.

Html structure

Libraries for web scraping:- bs4 , requests, urllib, scrapy, selenium.

What is json format?

What is api?

Find and findall in bs4?

What is the use of requests library in python?

How to handle dynamic content while web scraping?

For dynamic content use selenium, scrapy which can intract with js and load dynamic content.