1.What is a GET request?

A **GET request** is an HTTP method used to **retrieve data** from a server.
When you visit a webpage or use `requests.get(url)` in Python, you're sending a GET request asking the server to return that page's content (like HTML, JSON, etc.).

response = requests.get("https://www.bbc.com/news")

2.How do you install external packages in Python?

You use **`pip`**, Python's package manager.
pip install requests beautifulsoup4

3.What is a User-Agent in HTTP?

A **User-Agent** is a string sent in HTTP headers that identifies your client (browser, bot, or script).
Some websites block requests from unknown agents, so you include a fake browser User-Agent to mimic a real browser.
headers = {'User-Agent': 'Mozilla/5.0'}
requests.get(url, headers=headers)

4.What is soup.find_all() used for?

`find_all()` in **BeautifulSoup** is used to **search for all tags of a specific type** in HTML.
soup.find_all('h2') # Finds all <h2> headline tags

5.What are the risks of web scraping?

Legal issues — Some websites prohibit scraping in their *Terms of Service*.

Ethical issues — Overloading a website with frequent requests can cause performance problems.

Structural changes — Websites can change their layout, breaking your scraper.

IP blocking — Sites may block you if you scrape too aggressively.

6.What's the difference between id and class in HTML?

Id --- Unique identifier for one element ---
class --- Can be shared by multiple elements ---
7.What is an HTML tag?

An **HTML tag** defines the structure and content of a webpage.
Tags are enclosed in `< >` brackets and usually come in pairs.
<h2>Breaking News</h2>

8.What does .text return in BeautifulSoup?

`.text` (or `.get_text()`) extracts and returns the **visible text content** inside an HTML tag — removing all markup.
headline = soup.find('h2').text
9.What is a try-except block?

A **try-except** block is used for **error handling** in Python.
It lets your program continue running even if an error occurs.

```
try:
    response = requests.get(url)
except Exception as e:
    print("Error:", e)
```

10.What are HTTP status codes?

HTTP status codes indicate the **result of a web request**.

Code	Meaning	Description
200	Ok	Request succeeded
301/302	Redirect	Resource moved to another URL
403	Forbidden	Access denied
404	Not found	Resource doesn’t exist
500	Server error	Internal server issue