import requests

from bs4 import BeautifulSoup

import pandas as pd

# -------------------------------------------

# Step 1: Define the URL to scrape

# -------------------------------------------

url = "https://quotes.toscrape.com/"

try:

    # -------------------------------------------

    # Step 2: Send a GET request to the website

    # -------------------------------------------

    response = requests.get(url, timeout=10)  # Request webpage with a timeout

    response.raise\_for\_status()  # Raise an error for HTTP failures (4xx, 5xx)

    # -------------------------------------------

    # Step 3: Parse the webpage content using BeautifulSoup

    # -------------------------------------------

    soup = BeautifulSoup(response.text, 'html.parser')

    # -------------------------------------------

    # Step 4: Extract quotes and authors

    # -------------------------------------------

    data\_list = []  # List to store scraped data

    quotes = soup.find\_all('div', class\_='quote')  # Find all quote blocks

    for quote in quotes:

        # Extract quote text safely

        text\_element = quote.find('span', class\_='text')

        text = text\_element.get\_text(strip=True) if text\_element else "N/A"

        # Extract author name safely

        author\_element = quote.find('small', class\_='author')

        author = author\_element.get\_text(strip=True) if author\_element else "N/A"

        # Add to list

        data\_list.append([text, author])

    # -------------------------------------------

    # Step 5: Create a DataFrame

    # -------------------------------------------

    df = pd.DataFrame(data\_list, columns=["Quote", "Author"])

    # -------------------------------------------

    # Step 6: Save DataFrame to CSV file

    # -------------------------------------------

    df.to\_csv("scraped\_quotes.csv", index=False, encoding='utf-8')

    print("✅ Web Scraping Successful! Data saved to scraped\_quotes.csv")

    # -------------------------------------------

    # Step 7: Print the first 5 rows as output

    # -------------------------------------------

    print(df.head())

except requests.exceptions.RequestException as e:

    print(f"❌ Failed to retrieve the webpage. Error: {e}")