import requests

from bs4 import BeautifulSoup

import csv

def get\_html(url):

headers = {'User-Agent': 'Mozilla/5.0'}

response = requests.get(url, headers=headers)

response.raise\_for\_status() # Raise an HTTPError if the HTTP request returned an unsuccessful status code

return response.text

def scrape\_pages(base\_url, num\_pages):

data = []

for page in range(1, num\_pages + 1):

print(f"Scraping page: {page}")

url = f"{base\_url}/page/{page}/"

try:

html = get\_html(url)

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

quotes = soup.find\_all('div', class\_='quote')

for quote in quotes:

text = quote.find('span', class\_='text').get\_text(strip=True)

author = quote.find('small', class\_='author').get\_text(strip=True)

tags = [tag.get\_text(strip=True) for tag in quote.find\_all('a', class\_='tag')]

data.append({'text': text, 'author': author, 'tags': ', '.join(tags)})

except requests.exceptions.RequestException as e:

print(f"Failed to retrieve data from page {page}: {e}")

return data

def save\_to\_csv(data, filename):

with open(filename, mode='w', newline='', encoding='utf-8') as file:

writer = csv.DictWriter(file, fieldnames=['text', 'author', 'tags'])

writer.writeheader()

writer.writerows(data)

def main():

base\_url = 'http://quotes.toscrape.com/'

num\_pages\_to\_scrape = 5

scraped\_data = scrape\_pages(base\_url, num\_pages\_to\_scrape)

print("Scraped data:", scraped\_data)

save\_to\_csv(scraped\_data, 'quotes.csv')

if \_\_name\_\_ == "\_\_main\_\_":

main()