```
# -*- coding: utf-8 -*-

"""Alaiba_Nawaz_Day4.ipynb


Automatically generated by Colaboratory.


Original file is located at

    https://colab.research.google.com/drive/1d096QnUJZp5KWshhIN4cmNO5UomlUuCg

"""


from bs4 import BeautifulSoup

import requests

import pandas as pd

import seaborn as sns


import warnings

warnings.filterwarnings("ignore")


URL = "https://sunnah.com/"

response = requests.get(URL)

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


table = soup.find('div', class_="collections")

hadith_links = table.find_all('a')


df = pd.DataFrame(columns=['Hadith Collection', 'Book Name', 'Book Range', 'Narrated By', 'Translation',
'Arabic'])
```

```
for link in hadith_links:
    hadith_link = link['href']
    hadith_book = link.find('div', class_="english_collection_title")

    response = requests.get(URL + hadith_link)
    soup = BeautifulSoup(response.content, 'html.parser')
    temp = soup.find('div', class_="book_titles titles")

    if temp:
        hadith_books_links = temp.find_all('a')
        book_range = temp.find_all('div', class_="book_range")

        if hadith_books_links:
            for book, range_value in zip(hadith_books_links, book_range):
                book_name = book.find('div', class_="english english_book_name").text.strip()
                book_range_value = range_value.text.strip() if range_value else 'NaN'
                hadith_text_link = book['href']

                response = requests.get(URL + hadith_text_link)
                soup = BeautifulSoup(response.content, 'html.parser')
                temp2 = soup.find_all('div', class_="hadithTextContainers")

                if temp2:
                    for narration in temp2:
                        narrated_by = narration.find('div', class_="hadith_narrated")
```

```
                translation = narration.find('div', class_="text_details")
                arabic = narration.find('div', class_="arabic_hadith_full arabic")

                data = data.append({
                    'Hadith Collection': hadith_book.text.strip(),
                    'Book Name': book_name,
                    'Book Range': book_range_value,
                    'Narrated By': narrated_by.text.strip() if narrated_by else '',
                    'Translation': translation.text.strip() if translation else '',
                    'Arabic': arabic.text.strip() if arabic else ''
                }, ignore_index=True)

data.to_csv("Hadiths.csv", index=False)

data

import sqlite3

# Connect to the SQLite database
conn = sqlite3.connect('hadiths.db')
cursor = conn.cursor()

# Create the table in the database
cursor.execute('''CREATE TABLE IF NOT EXISTS hadiths (
                id INTEGER PRIMARY KEY,
                hadith_collection TEXT,
```

```
                book_name TEXT,
                book_range TEXT,
                narrated_by TEXT,
                translation TEXT,
                arabic TEXT
            )''')

# Iterate over the scraped data and insert into the database
for index, row in df.iterrows():
    hadith_collection = row['Hadith Collection']
    book_name = row['Book Name']
    book_range = row['Book Range']
    narrated_by = row['Narrated By']
    translation = row['Translation']
    arabic = row['Arabic']

    # Insert the data into the database
    cursor.execute("INSERT INTO hadiths (hadith_collection, book_name, book_range, narrated_by, translation, arabic) VALUES (?, ?, ?, ?, ?, ?)",
                   (hadith_collection, book_name, book_range, narrated_by, translation, arabic))

# Commit the changes and close the connection
conn.commit()
conn.close()
```