

09_scrape_02: Web scraping

Objective: Basic Scraping a Web Page HTML by BeautifulSoup

Problems

Q1 Function:

Write function Q1(file_path). that counts the number of national holidays in Thailand each year provided. The user will first enter a file path HTML for which it wants to count national holidays. The function will count the national holidays in each specified year and return a list that contains the number of national holidays for each year, from the first Monday (index 0) until the following Sunday.

Input:

file_path HTML (e.g. "Holidays_2023.html")

Output:

A list that contains the number of national holidays for each year, counted from the first Monday until the following Sunday (e.g., [9, 2, 0, 3, 5, 3, 4])

Q2 Function:

Write a function Q2(file_path) that finds Visakha Bucha Day for each year provided in the HTML file. The user will input a file path to an HTML document, as in Q1. The function will return the day of the week of Visakha Bucha Day that year.

Input:

file_path HTML (e.g. "Holidays_2023.html")

Output:

A string containing the Visakha Bucha Day for each year
(e.g., "Saturday").

Expected Results

Input	Output
<code>print(Q1("Holidays_2023.html"))</code>	<code>[9, 2, 0, 3, 5, 3, 4]</code>
<code>print(Q2("Holidays_2023.html"))</code>	<code>Saturday</code>

Template codes

```
import ...

def Q1(file_path): # DO NOT modify this line
    # remove pass and replace with you code
    Pass

def Q2(file_path): # DO NOT modify this line
    # remove pass and replace with you code
    Pass

exec(input().strip()) # do not delete this line
```

Example HTML file attached

```
<!DOCTYPE html>
<!--[if IE 8 ]><html class="ie ie8" lang="en"> <![endif]--><!--[if (gte IE
9)|!(IE)]><!--><html lang="en" xmlns="http://www.w3.org/1999/xhtml"> <!--<![endif]-->
<head>
<!-- InMobi Choice. Consent Manager Tag v3.0 (for TCF 2.2) -->
<script async="true" type="text/javascript">
(function() {

...

```