```python
import requests
from bs4 import BeautifulSoup
import json

# 建立空的電影清單
movies = []

for i in range(1, 15068):
    url = f"https://movies.yahoo.com.tw/movieinfo_main/{i}"
    response = requests.get(url)
    soup = BeautifulSoup(response.text, 'html.parser')

    # 取得電影ID
    movie_id = soup.find('span', {'class': 'num'})
    if movie_id:
        movie_id = movie_id.text

    # 取得中文片名
    ch_name = soup.find('h1')
    if ch_name:
        ch_name = ch_name.text

    # 取得英文片名
    en_name = soup.find('h3')
    if en_name:
        en_name = en_name.text

    # 取得標籤
    tags = []
    tag_elements = soup.find_all('div', {'class': 'level_name_box'})
    for tag_element in tag_elements:
        tag = tag_element.text.strip()
        tags.append(tag)

    # 取得電影簡介
    intro = soup.find('div', {'class': 'gray_infobox_inner'})
    if intro:
        intro = intro.text.strip()

    # 取得上映日期
    release_date = soup.find('div', {'class': 'movie_intro_info_r'})
    if release_date:
        release_date = release_date.span.text.strip()

    # 將每筆電影資料存成一個字典
    movie = {
        "電影ID": movie_id,
        "中文片名": ch_name,
        "英文片名": en_name,
        "標籤": tags,
        "電影簡介": intro,
        "上映日期": release_date
    }

    # 將這筆電影資料加入清單中
    movies.append(movie)

# 將整個清單存成一個 JSON 檔案
with open('hw2.json', 'w', encoding='utf-8') as f:
    json.dump(movies, f, ensure_ascii=False, indent=2)
```

⎋

```
-------------------------------------------------------------------------
ValueError                               Traceback (most recent call last)
/usr/local/lib/python3.9/dist-packages/urllib3/response.py in
_update_chunk_length(self)
    760             try:
--> 761                 self.chunk_left = int(line, 16)
    762             except ValueError:

ValueError: invalid literal for int() with base 16: b''

During handling of the above exception, another exception occurred:

InvalidChunkLength                       Traceback (most recent call last)
```
⬍ 14 frames
```
InvalidChunkLength: InvalidChunkLength(got length b'', 0 bytes read)

During handling of the above exception, another exception occurred:

ProtocolError                            Traceback (most recent call last)
ProtocolError: ("Connection broken: InvalidChunkLength(got length b'', 0 bytes
read)", InvalidChunkLength(got length b'', 0 bytes read))

During handling of the above exception, another exception occurred:

ChunkedEncodingError                     Traceback (most recent call last)
/usr/local/lib/python3.9/dist-packages/requests/models.py in generate()
    761                     yield chunk
    762             except ProtocolError as e:
--> 763                 raise ChunkedEncodingError(e)
    764             except DecodeError as e:
    765                 raise ContentDecodingError(e)

ChunkedEncodingError: ("Connection broken: InvalidChunkLength(got length b'', 0
bytes read)", InvalidChunkLength(got length b'', 0 bytes read))
```

Colab 付費產品 - 按這裡取消合約

● ✕