Leverage your Python (above version 3) and JavaScript skills to complete tasks. **We have 2 optional tasks, but you have to complete at least 3 tasks to get passed.**

Task 1: Parse data from internet and save to files by Python

Here are URLs for tourist spots in Taipei provided by Taipei City Government:
https://padax.github.io/taipei-day-trip-resources/taipei-attractions-assignment-1
https://padax.github.io/taipei-day-trip-resources/taipei-attractions-assignment-2

Write a Python program without third-party libraries to parse data from above URLs and output 2 files in CSV format:

1. Output one spot information per line into spot.csv. If there are many image URLs in source data, we only output the first one.
2. Grouping spot titles by MRT station, output 1 MRT station and nearby spot titles per line into mrt.csv.

Put spot.csv and mrt.csv files in your weekly task folder.

Note: You can extract the district from the address column in source data, which should be 3 chinese words, from one of 中正區、萬華區、中山區、大同區、大安區、松山區、信義區、士林區、文山區、北投區、內湖區、南港區

Format of spot.csv

SpotTitle,District,Longitude,Latitude,ImageURL
SpotTitle,District,Longitude,Latitude,ImageURL
...

Example of spot.csv

新北投溫泉區,北投區,123.5446,24.5312,https://www.travel.taipei/pic/11000848.jpg
士林官邸,士林區,122.4564,23.546,https://www.travel.taipei/pic/11000999.jpg
...

Foramt of mrt.csv

StationName,AttractionTitle1,AttractionTitle2,AttractionTitle3,...

StationName,AttractionTitle1,AttractionTitle2,AttractionTitle3,...

...

Example of mrt.csv

新北投,新北投溫泉區,北投圖書館,地熱谷,...

劍潭,國立臺灣科學教育館,社子島環島與二重疏洪道自行車道,國立臺灣科學教育館,...

...

Task 2: Parse web page data and save to files by Python (Optional)

Here is URL for PTT lottery board:
https://www.ptt.cc/bbs/Lottery/index.html

Write a Python program to parse article title, like/dislike count, and publish time for every article in PTT lottery board. Format of publish time should be <u>EEE MMM DD HH:MM:SS YYYY</u>, for example: Fri Jul 14 23:14:36 2023. We can get publish time by parsing article pages. If there is no publish time data, fill it with an empty string.

Parse every article data in the first 3 pages, excluding deleted ones.

In this task, we can use the BeautifulSoup library to parse HTML code.

Output one article information per line into <u>article.csv</u>, putting in your weekly task folder.

Format of article.csv

ArticleTitle,Like/DislikeCount,PublishTime

ArticleTitle,Like/DislikeCount,PublishTime

…

Example of article.csv

[問題] 享受輸的感覺539,4,Fri Jul 14 23:34:43 2023

[報牌] 39樂合彩,0,Sat Jul 15 00:01:14 2023

…

## Task 3: Parse data from internet and render as HTML by JavaScript

Here is URL for tourist spots in Taipei provided by Taipei City Government:
https://padax.github.io/taipei-day-trip-resources/taipei-attractions-assignment-1

We have to write JavaScript code to parse spot data from the above URL, and render them in the page completed in week 1, without third-party libraries.

Steps and Details:

1. Connect to the URL to retrieve data by JavaScript after the page loads.

2. Render retrieved data as spot list in the page, following the same layout design as assignment in week 1. We only render the title and first image for each spot. Use at least document.createElement() and appendChild() methods to complete rendering. Don't use innerHTML.

3. Render the first 3 spots in the small boxes, following 10 spots in the big boxes, ignoring others.
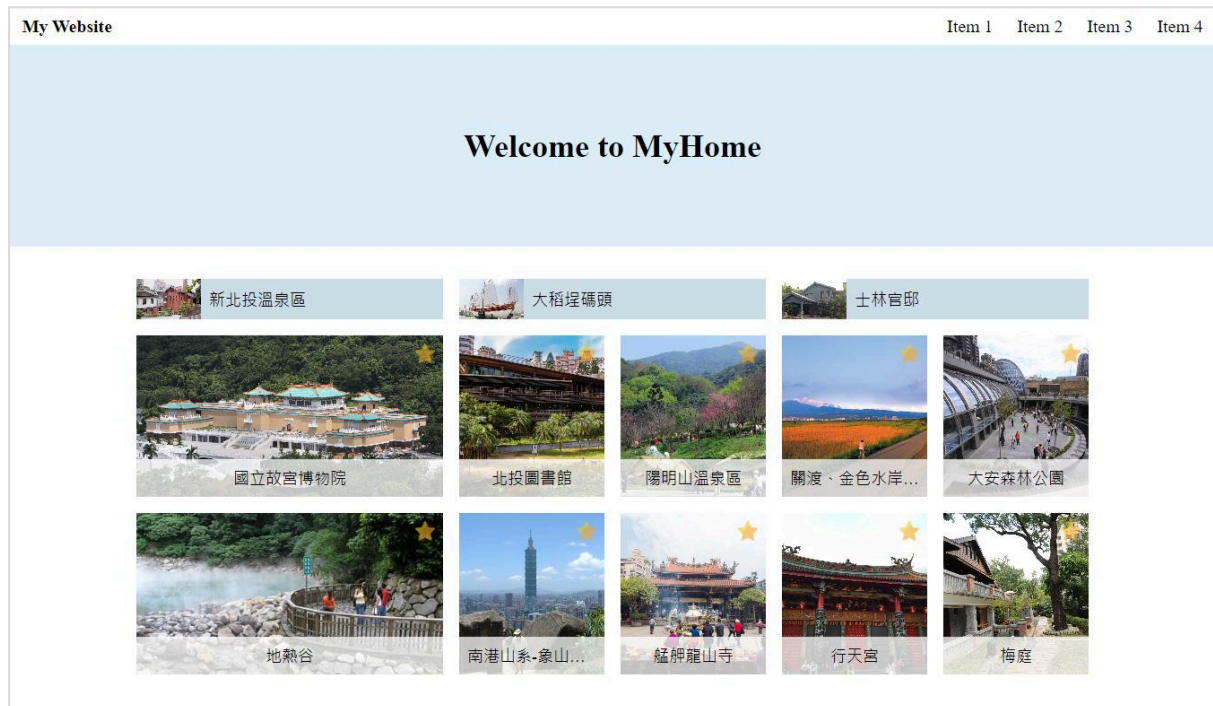
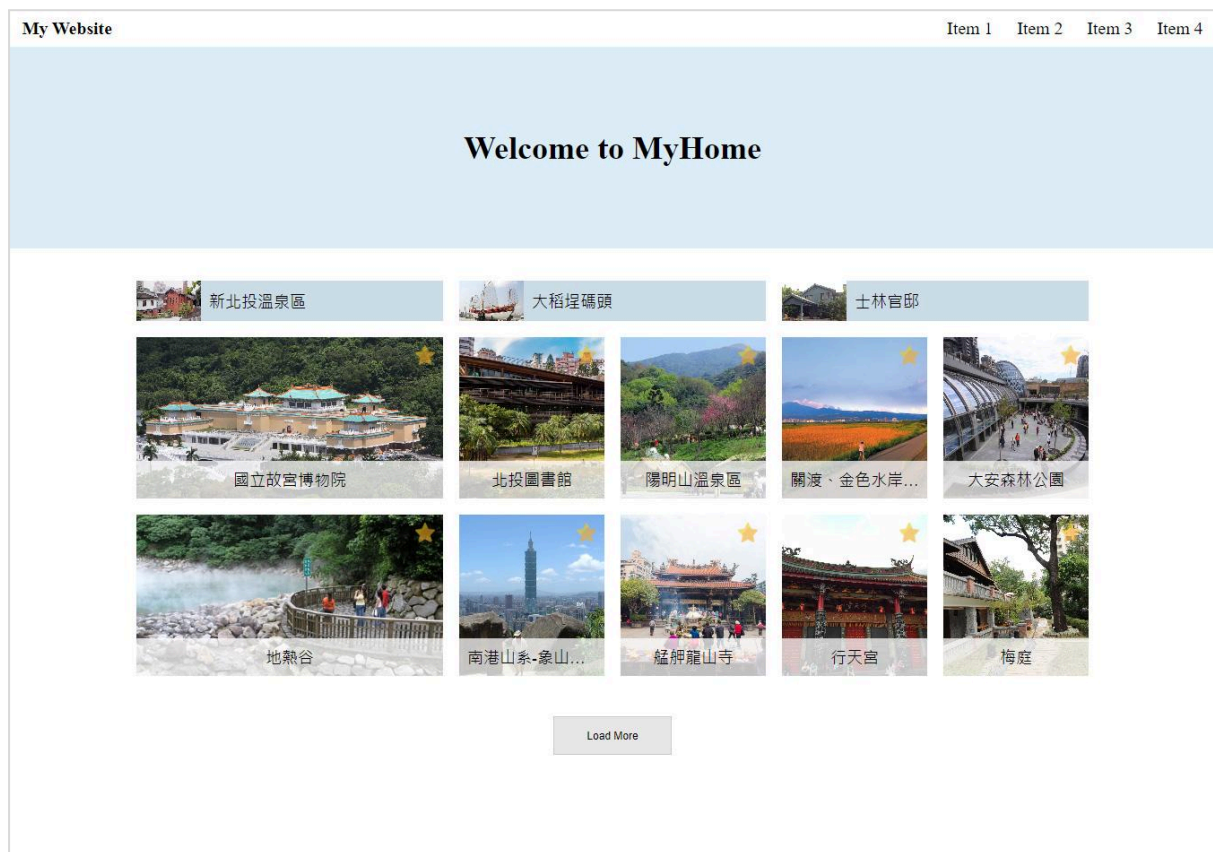## Task 4: Build paging mechanism by JavaScript (Optional)

Continue the previous task, build a load more button at the bottom of the spot list. Render the next 10 spots if we clicked on this button, until the end of data.

Refer to screenshots below, click load more button in the first screenshot, we should get the second one, and so on.
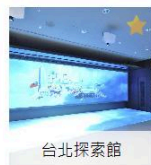
My Website

Welcome to MyHome

新北投溫泉區
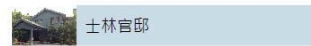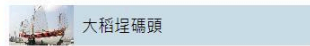
大稻埕碼頭

士林官邸

國立故宮博物院

北投圖書館

陽明山溫泉區

關渡、金色水岸...

大安森林公園

地熱谷

南港山系-象山...

艋舺龍山寺

行天宮

梅庭

中正紀念堂

台北探索館

二格山系-指南...

北投溫泉博物館

琉園水晶博物館...

華山1914文化創意產業園區

長榮海事博物館

台北101

袖珍博物館

新店溪、大漢溪...

Load More