



亞洲大學
ASIA UNIVERSITY

**Midterm Project Report
Advanced Computer Programming**

Web Scrapping with Python

Student Name : Nathanael Firgo

Student ID : 112021185

Teacher : DINH-TRUNG VU

2024-04

Chapter 1 Introduction

1.1 Github

- 1) **Personal Github Account:** <https://github.com/112021185>
- 2) **Group Github Account:** <https://github.com/112021179>
- 3) **Group Project Repository:** <https://github.com/112021179/1>
- 4) **List of submitted files:**
 - 112021185.py

1.2 Topic

Build a web scraper to extract information about the title, price, and availability of books from a website :

<https://books.toscrape.com/>

1.3 Project Overview

This program have used request, pandas, and beautifulsoup to scrape a website. It then creates an CSV file with the price, title, information, and availability of the books from that certain website

Chapter 2 Implementation

2.1 Function 1

The first function, which is the while loop, is used to create a loop that keeps on retrieving information. It retrieves the information using request library to send an http get request to the generated url, then it starts to scrape the page. Beautiful soup is used to parse the html content, and the find method is what is used to get the certain information that we want

2.2 Function 2

The second function, is to save the information that is already scraped, to be then saved into a csv file.

Chapter 3 Results

3.1 Result 1

In Visual Studio Code, the program will return the number of page from the website that is being scraped. The end result is a csv file of a list of books title, price, information, and also the availability of the books

```
PS C:\Users\ASUS> & C:/Users/ASUS/AppData/Local/Microsoft/WindowsApps/python3.11.exe "c:/Users/ASUS/Downloads/tgs_kuliah/fundamental computer programming/midterm project(web scraping).py"
scraping page number: 1
scraping page number: 2
scraping page number: 3
scraping page number: 4
scraping page number: 5
```

	title,link,price,avail
2	0,A Light in the Attic,a-light-in-the-attic_1000/index.html,Â€51.77,In stock
3	1,Tipping the Velvet,tipping-the-velvet_999/index.html,Â€53.74,In stock
4	2,Soumission,soumission_998/index.html,Â€50.10,In stock
5	3,Sharp Objects,sharp-objects_997/index.html,Â€47.82,In stock
6	4,Sapiens: A Brief History of Humankind,sapiens-a-brief-history-of-humankind_996/index.html,Â€54.23,In stock
7	5,The Requiem Red,the-requiem-red_995/index.html,Â€22.65,In stock
8	6,The Dirty Little Secrets of Getting Your Dream Job,the-dirty-little-secrets-of-getting-your-dream-job_994/index.html,Â€33.34,In stock
9	7,"The Coming Woman: A Novel Based on the Life of the Infamous Feminist, Victoria Woodhull",the-coming-woman-a-novel-based-on-the-life-of-the-infamous-feminist-victoria-woodhull_993/index.html
10	8,The Boys in the Boat: Nine Americans and Their Epic Quest for Gold at the 1936 Berlin Olympics,the-boys-in-the-boat-nine-americans-and-their-epic-quest-for-gold-at-the-1936-berlin-olympics_992/index.html

Chapter 4 Conclusions

This program can scrape and take information of books from a website, to then be given to the user in a csv file form.