

**LAPORAN PRAKTIK  
PEMOGRAMAN PYTHON**

**PRAKTIKUM 10  
SCRAPING DATA**



**Disusun oleh :**

Muhammad Asror Alva 'Izzi  
(V3922050)

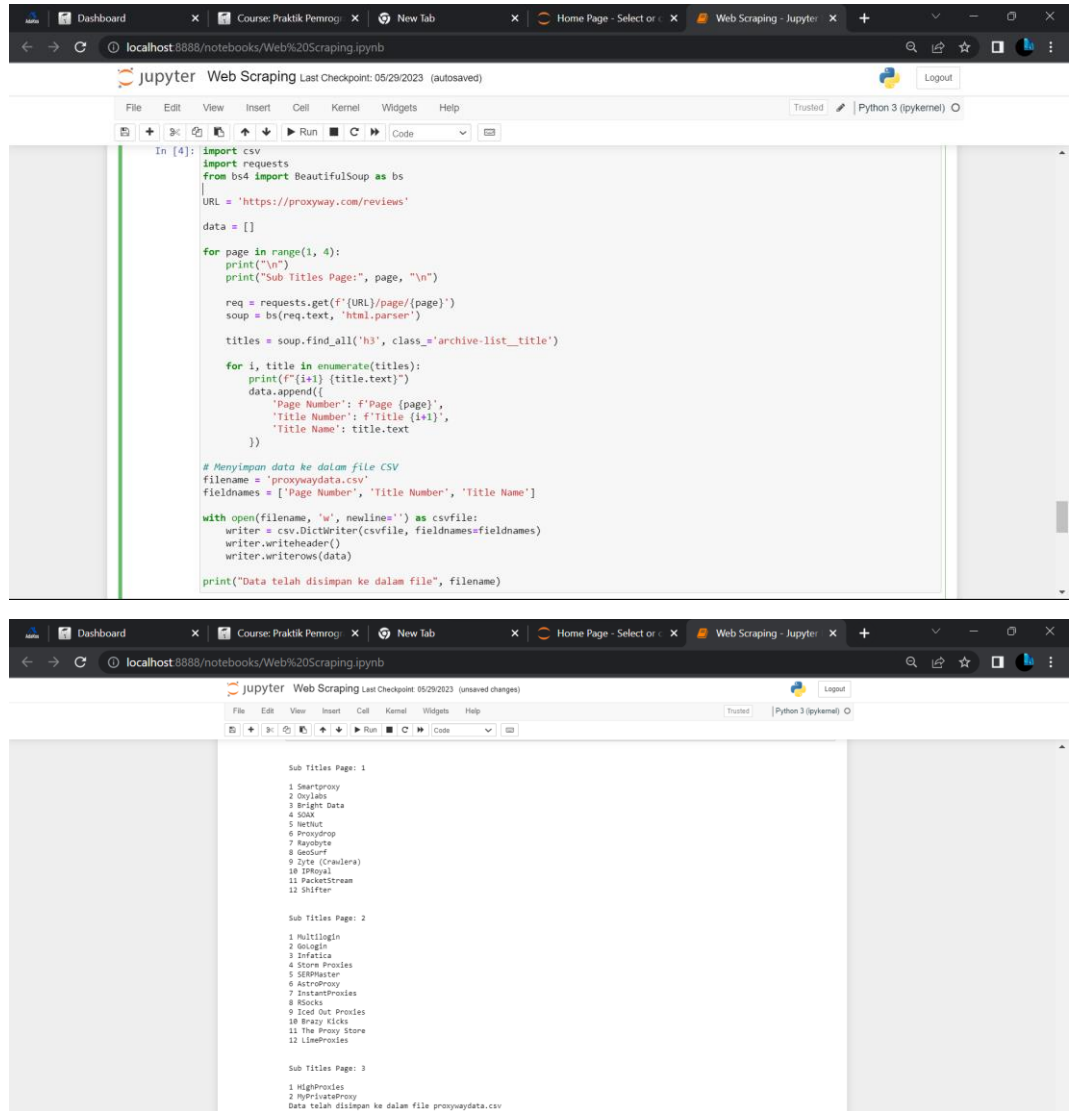
**Dosen**

Yusuf Fadila Rachman S.Kom,M.Kom

**PS D-III TEKNIK INFORMATIKA  
SEKOLAH VOKASI  
UNIVERSITAS SEBELAS MARET  
2023**

# HASIL

## 1. Hasil dan script scraping web



The first screenshot shows the Jupyter Notebook interface with a Python script for web scraping. The script imports the necessary libraries, sets the URL, and iterates through the first four pages of the website <https://proxyway.com/reviews>. It extracts the titles of the reviews and appends them to a list. The data is then saved to a CSV file named 'proxywaydata.csv'.

```
In [4]: import csv
import requests
from bs4 import BeautifulSoup as bs
|
| URL = 'https://proxyway.com/reviews'
|
| data = []
|
| for page in range(1, 4):
|     print("\n")
|     print("Sub Titles Page:", page, "\n")
|
|     req = requests.get(f'{URL}/page/{page}')
|     soup = bs(req.text, 'html.parser')
|
|     titles = soup.find_all('h3', class_='archive-list_title')
|
|     for i, title in enumerate(titles):
|         print(f'{i+1} {title.text}')
|         data.append({
|             'Page Number': f'Page {page}',
|             'Title Number': f'Title {i+1}',
|             'Title Name': title.text
|         })
|
| # Menyimpan data ke dalam file CSV
| filename = 'proxywaydata.csv'
| fieldnames = ['Page Number', 'Title Number', 'Title Name']
|
| with open(filename, 'w', newline='') as csvfile:
|     writer = csv.DictWriter(csvfile, fieldnames=fieldnames)
|     writer.writeheader()
|     writer.writerows(data)
|
| print("Data telah disimpan ke dalam file", filename)
```

The second screenshot shows the output of the script. It displays the titles of the reviews for the first three pages. The output is as follows:

```
Sub Titles Page: 1
1 Smartproxy
2 Oxylabs
3 Bright Data
4 SOAX
5 NetNut
6 Proxydrop
7 Proxybyte
8 GeckSurf
9 Oryx (Crawlers)
10 IPRoyal
11 PacketStream
12 Shifter

Sub Titles Page: 2
1 Multilogin
2 Oologin
3 Infatica
4 Store Proxies
5 SERMaster
6 AstroProxy
7 InstantProxies
8 RSocks
9 Iced Out Proxies
10 Brany Clicks
11 The Proxy Store
12 LineProxies

Sub Titles Page: 3
1 HighProxies
2 MyPrivateProxy
Data telah disimpan ke dalam file proxywaydata.csv
```

## 2. Penjelasan dari script diatas

Script ini menggunakan library **csv**, **requests**, dan **BeautifulSoup** untuk melakukan scraping data dari situs web <https://proxyway.com/reviews> dan menyimpannya ke dalam file CSV.

- `URL` berisi URL situs web yang akan di-scrape.
- `data` adalah list kosong untuk menyimpan data yang di-scrape.

- Dilakukan loop dari 1 hingga 3 untuk mengakses halaman-halaman yang berbeda.
- Dalam setiap loop, dilakukan permintaan GET ke URL halaman yang sesuai.
- Hasil respons dari permintaan GET diparsing menggunakan BeautifulSoup.
- Dalam parsing, diambil judul-judul yang ada di halaman tersebut.
- Setiap judul ditambahkan ke list `data` dengan informasi nomor halaman dan nomor judul.
- Setelah semua data diambil, list `data` disimpan ke dalam file CSV dengan nama 'proxywaydata.csv'.
- File CSV memiliki tiga kolom: 'Page Number', 'Title Number', dan 'Title Name'.
- Setelah data disimpan, pesan "Data telah disimpan ke dalam file 'proxywaydata.csv'" ditampilkan sebagai konfirmasi.