

# LAPORAN DATA SCRAPING

Disusun untuk Memenuhi Matakuliah Data Scraping  
Dibimbing oleh Pelsri Ramadar N.S., M.Kom



Oleh:

**FAVIAN NURRUDDIN**

**1123102138**

**PROGRAM STUDI S1 TEKNIK INFORMATIKA  
SEKOLAH TINGGI ILMU KOMPUTER PGRI BANYUWANGI  
2024**

## LAPORAN MATERI 8

### Percobaan Code

Kode Python ini adalah script web scraping yang menggunakan Selenium dan BeautifulSoup untuk mengekstrak informasi produk kopi Banyuwangi dari halaman website Tokopedia.

#### a. Script / Setting Program

```
from selenium import webdriver
from selenium.webdriver.chrome.options import Options
from bs4 import BeautifulSoup
import time

URL = 'https://www.tokopedia.com/search?st=product&q=kopi+banyuwangi'

options = Options()
options.add_argument("user-agent=Mozilla/5.0 (Windows NT 10.0; Win64; x64)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/128.0.6613.186 Safari/537.36")

driver = webdriver.Chrome(options=options)
driver.get(URL)

time.sleep(50)

html = driver.page_source

soup = BeautifulSoup(html, 'html.parser')

data = soup.find_all('div', {'class': 'css-5wh65g'})

for item in data:
    nama_produk = item.find('span', {'class': '_OT8-iGxMpV6NEsYEhwkqEg=='})
    harga_produk = item.find('div', {'class': '_67d6E1xDKlzw+i2D2L0tjw=='})
    penjual = item.find('span', {'class': 'T0rpy-LEwYNQifsgB-3SQw=='})

    if nama_produk and harga_produk and penjual:
        print(f>Nama Produk : {nama_produk.text.strip()} if nama_produk else 'Tidak tersedia')
        print(f>Harga Produk: {harga_produk.text.strip()} if harga_produk else 'Tidak tersedia')
        print(f>Nama Penjual: {penjual.text.strip()} if penjual else 'Tidak tersedia')
        print("-" * 50)

driver.quit()
```

## b. Langkah Uji Coba

### 1. Import Libraries :

- **selenium.webdriver** : Mengimpor Selenium untuk mengontrol browser secara otomatis
- **selenium.webdriver.chrome.options** : Mengimpor kelas Options dari pustaka Selenium.
- **bs4** : Mengimpor BeautifulSoup dari bs4 untuk memproses HTML.
- **Time** : Mengimpor modul time untuk memberikan jeda selama eksekusi.

### 2. URL Target :

- **URL** pencarian Tokopedia dengan kata kunci kopi banyuwangi.

### 3. Konfigurasi Browser:

- **Options()** : Membuat opsi untuk Chrome dengan Options().
- **options.add\_argument()** : Menambahkan argumen user-agent agar tampak seperti permintaan dari browser sebenarnya.

### 4. Inisialisasi Web Driver:

- **driver = webdriver.Chrome(options=options)** : Membuka Chrome dengan opsi yang sudah diatur.
- **driver.get(URL)** : Mengunjungi URL target menggunakan Selenium.

### 5. Memberikan Jeda dan Mengambil HTML :

- **time.sleep(50)** : Memberikan jeda waktu selama 50 detik untuk memastikan halaman selesai dimuat.
- **html = driver.page\_source** ; Mendapatkan HTML dari halaman yang telah dimuat.

### 6. Proses HTML dengan BeautifulSoup :

- **soup = BeautifulSoup(html, 'html.parser')** : Mengurai HTML dengan BeautifulSoup menggunakan parser html.parser..

### 7. Mencari Data Produk :

- **data = soup.find\_all('div', {'class': 'css-5wh65g'})** : Mencari elemen div dengan kelas spesifik (css-5wh65g) yang berisi informasi produk.

## 8. Iterasi dan Ekstraksi Data Memakao Loop :

- Iterasi elemen data yang berisi daftar produk.
- Mencari nama produk, harga, dan penjual dengan find() menggunakan kelas tertentu.
- Mengecek apakah ketiga elemen ditemukan.
- Jika ditemukan, menampilkan data produk dengan print():
  - text.strip() digunakan untuk membersihkan spasi tambahan.
  - Tampilkan garis pemisah untuk setiap produk.

## 9. Menutup Browser :

- **driver.quit()** : Menutup browser setelah semua operasi selesai.

### c. Hasil Uji Coba

```
Nama Produk : Kopi Osing Blend - Kopai Osing Blend - Kemiren, Banyuwangi
Harga Produk: Rp125.000
Nama Penjual: Kopai Osing
-----
Nama Produk : Kopi Osing Arabica - Kopai Osing Arabica - Kemiren, Banyuwangi
Harga Produk: Rp150.000
Nama Penjual: Kopai Osing
-----
Nama Produk : Kopi Robusta Asli Banyuwangi 200 Gram (Kopi Bubuk Murni)
Harga Produk: Rp25.000
Nama Penjual: Teras Kopi
-----
Nama Produk : Kopi Lanang Asli Banyuwangi 200 Gram (Kopi Bubuk 100% Murni)
Harga Produk: Rp30.000
Nama Penjual: Teras Kopi
-----
Nama Produk : KOPI ROBUSTA / kopi kemiren / kopi banyuwangi / kopi osing
Harga Produk: Rp45.500
Nama Penjual: Merizart
-----
Nama Produk : Kopi Lanang Khas Banyuwangi
Harga Produk: Rp35.000
Nama Penjual: kopi Lanang Banyuwangi
-----
```

```

Nama Produk : Kopi Robusta Blend Osing Banyuwangi 250gr Omahkopi
Harga Produk: Rp87.000
Nama Penjual: Omahkopi
-----
Nama Produk : Kopi Lanang Robusta Ijen | Kopi Lanang Robusta Banyuwangi 100gr
Harga Produk: Rp32.000
Nama Penjual: Toko Madu&Kopi
-----
Nama Produk : PAKET KOPI LANANG DAN KOPI ROBUSTA BANYUWANGI (BUBUK)
Harga Produk: Rp50.000
Nama Penjual: Teras Kopi
-----
Nama Produk : Kopi Bubuk Robusta Jaran Goyang Banyuwangi 250 gram
Harga Produk: Rp38.000
Nama Penjual: Toko Bsart
-----

```

#### d. Analisa Hasil

##### 1. Pengaturan dan Inisialisasi :

- Kode mengimpor pustaka yang diperlukan: selenium untuk mengontrol browser, BeautifulSoup untuk parsing HTML, dan time untuk menunggu.
- URL yang dituju adalah halaman pencarian produk kopi Banyuwangi di Tokopedia.

##### 2. Pengaturan Browser :

- Opsi untuk Chrome diatur dengan menambahkan user-agent agar tampak seperti permintaan dari browser biasa.

##### 3. Mengakses Halaman :

- Skrip membuka halaman web yang ditentukan dan menunggu selama 50 detik untuk memastikan semua elemen halaman dimuat sepenuhnya.

##### 4. Pengambilan Data :

- Setelah halaman dimuat, kode mengambil sumber HTML dan mem-parsing-nya menggunakan BeautifulSoup.
- Kode mencari semua elemen dengan kelas tertentu yang berisi informasi produk.

##### 5. Ekstraksi Informasi :

- Untuk setiap item produk yang ditemukan, kode mencoba mengekstrak nama produk, harga, dan nama penjual.
- Jika semua informasi tersedia, data tersebut dicetak ke terminal.

## 6. Penutupan :

Setelah selesai, driver browser ditutup dengan `driver.quit()`.

**Mengetahui:**

Dosen Pengampu Mata Kuliah

**Pelsri Ramadar N.S., M.Kom**

NIKP. 0704068408

## PENUTUP

### Kesimpulan

Script ini efektif untuk mengumpulkan informasi tentang produk kopi Banyuwangi dari Tokopedia, termasuk nama produk, harga, dan penjual. Namun, penggunaan web scraping tetap harus mematuhi kebijakan situs web yang bersangkutan dan hukum yang berlaku.