

LAPORAN TUGAS HARIAN 3

DATA SCRAPING

Disusun untuk Memenuhi Tugas Harian 3 Mata Kuliah Data Scraping
Dibimbing oleh Pelsri Ramadar N.S. M.Kom



Oleh:

Annisa Maharani (1123102127)

PROGRAM STUDI S1 TEKNIK INFORMATIKA
SEKOLAH TINGGI ILMU KOMPUTER PGRI BANYUWANGI
2022

1. Kode Program

```
from bs4 import BeautifulSoup
from selenium import webdriver
import time

URL = 'https://www.tokopedia.com/search'

params = {
    'q': 'batik banyuwangi'
}

driver = webdriver.Chrome()
fullURL = f"{URL}?q={params['q']}"
driver.get(fullURL)

time.sleep(5)

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

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

for i in range(len(data)):
    nama_produk = data[i].find('span', {'class': '_0T8-iGxMpV6NEsYEhwkqEg=='})
    harga = data[i].find('div', {'class': '_67d6E1xDKIzw+i2D2L0tjw=='})
    penjual = data[i].find('span', {'class': 'T0rpy-LEwYNQifsgB-3SQw==pC8DMVkBZGW7-egObcwMFQ== flip'})

    if nama_produk and harga and penjual :
        print("nama produk: "+nama_produk.text)
        print("harga: "+harga.text)
        print("penjual: "+penjual.text)

print(data)
driver.quit()
```

Penjelasan:

Kode Python di atas bertujuan untuk melakukan web scraping di situs e-commerce Tokopedia guna mencari informasi produk berdasarkan kata kunci "batik banyuwangi". Program menggunakan Selenium untuk membuka dan mengotomasi browser serta BeautifulSoup untuk memproses dan mengekstrak elemen HTML dari halaman yang telah dimuat. URL dasar pencarian Tokopedia disimpan dalam variabel URL. Kata kunci pencarian dimasukkan ke dalam parameter, yang merupakan URL lengkap dengan query pencarian "batik banyuwangi". Selenium membuka browser Chrome secara otomatis menggunakan `webdriver.Chrome()`, lalu mengarahkan browser ke fullURL menggunakan metode `get()`. Untuk memastikan semua elemen pada halaman selesai dimuat, program menunggu selama 5 detik menggunakan `time.sleep(5)`. Selanjutnya, sumber HTML halaman diambil menggunakan `driver.page_source`, kemudian diparsing menggunakan BeautifulSoup untuk mempermudah analisis dan pencarian elemen yang sesuai.

Program mencari semua elemen `<div>` dengan class `'css-5wh65g'`, yang berisi data produk. Setiap elemen ini diperiksa untuk menemukan sub-elemen yang berisi nama produk yang dicari menggunakan elemen `` dengan class tertentu, harga yang dicari dalam elemen `<div>` dengan class lain, dan nama penjual yang dicari dalam elemen `` dengan class berbeda. Jika ketiga informasi tersebut berhasil ditemukan, data akan dicetak ke layar. Setelah iterasi selesai, program mencetak semua elemen hasil pencarian untuk keperluan debugging, memastikan elemen yang berhasil ditemukan sesuai dengan yang diharapkan. dan kemudian menutup browser menggunakan `driver.quit()`.