

LAPORAN KOMUNIKASI DATA

“Web scraping menggunakan python”



Nama: Fenny Anggraini

NIM: 09011182025019

Kelas : SK4A Indralaya

Dosen pengampuh : Adi Hermansyah, S.KOM., M.T.

Jurusan Sistem Komputer
Fakultas Ilmu Komputer
Universitas Sriwijaya
2022

I. JUDUL

“Melakukan Web Scraping dengan Menggunakan Python”

II. TUJUAN

- Memahami penggunaan python dan mengimplementasikannya
- Mampu memahami definisi dan fungsi dari metode web scraping
- Mampu melakukan proses web scraping terhadap website besar

III. ALAT & BAHAN

I. Alat

- Laptop atau PC

II. Bahan

- Visual Code Studio
- Microsoft Excel
- Ekstension Python

IV. DASAR TEORI

Web Scraping

Web Scraping atau yang dikenal sebagai ekstraksi web adalah teknik untuk mengekstraksi data dari World Wide Web (WWW) dan menyimpannya ke file sistem atau basis data untuk dijadikan analisis data. Web Scraping dapat dilakukan baik secara manual oleh seorang pengguna atau secara otomatis oleh bot atau crawler web. Proses Web Scraping dari Internet dapat dibagi menjadi dua langkah berurutan, yaitu mengakuisisi sumber daya web dan kemudian mengekstraksi informasi yang diinginkan dari data yang diperoleh. Secara khusus, program Web Scraping dimulai dengan meminta HTTP untuk memperoleh sumber daya dari yang ditargetkan oleh situs web. Permintaan ini dapat diformat kedalam URL yang berisi permintaan GET atau HTTP yang berisi POST. Setelah permintaan berhasil diterima dan diproses oleh situs web yang ditargetkan, sumber daya yang diminta akan diambil dari situs web dan kemudian dikirim kembali ke program Web Scraping. Sumber daya ini bisa dalam berbagai format, seperti halaman web yang dibangun dengan HTML, XML atau JSON, atau data multimedia seperti gambar, audio, atau video. Terdapat dua modul penting dari Web Scraping - modul untuk menulis permintaan HTTP, seperti Urllib2 atau selenium dan satu lagi untuk parsing dan mengekstraksi informasi dari kode HTML mentah, seperti BeautifulSoup atau Pyquerr.

V. PROSEDUR PRAKTIKUM DAN HASIL PRAKTIKUM

WEBSITE BUKALAPAK (API BUKALAPAK)

Source code

```
run.py  X
bukalapak > hasil > run.py > ...
1  import requests
2  import csv
3
4  key = input('masukkan keyword :')
5  write = csv.writer(open('hasil/{}.csv'.format(key), 'w', newline=''))
6  header = ['Nama', 'Harga', 'Stok']
7  write.writerow(header)
8
9  url = 'https://api.bukalapak.com/multistrategy-products'
10
11  count = 0
12  for page in range(1,11):
13      parameter = {
14          'keywords': key,
15          'limit': 50,
16          'offset': 0,
17          'facet': True,
18          'page': page,
19          'shouldUseSeoMultistrategy': False,
20          'access_token': 'jzCqLX8aWcaCp09pVz6UdThQgbKGrSAg8-UnQG8wGD8sLg'
21      }
22
23      r = requests.get(url, params=parameter).json()
24
25      products = r['data']
26      for p in products:
27          nama = p['name']
28          harga = p['price']
29          stok = p['stock']
30          count+=1
31          print('No :',count, 'nama :',nama, 'harga:',harga, 'stok :',stok)
32          write = csv.writer(open('hasil/{}.csv'.format(key), 'a', newline=''))
33          data = [nama, harga, stok]
34          write.writerow(data)
35
```

Code di atas merupakan code untuk melakukan proses webscraping bukalapak yang dilakukan menggunakan compiler dengan ekstensi python.

Output

1. Saat program berhasil dijalankan, maka user akan diminta untuk memasukkan input keyword dari produk yang ingin dicari dan di export menjadi file csv.

```
PS D:\bukalapak> & 'C:\Users\asus\AppData\Local\Microsoft\WindowsApps\python3.7.exe' 'c:\Use
rs\asus\.vscode\extensions\ms-python.python-2022.6.0\pythonFiles\lib\python\debugpy\launcher'
'50979' '--' 'd:\bukalapak\hasil\run.py'
```

masukkan keyword :handuk

2. Program akan mengeksekusi input yang dimasukan oleh user, dan menampilkan semua produk yang tersedia pada beberapa page daripada website Bukalapak.

PROBLEMS	OUTPUT	DEBUG CONSOLE	TERMINAL
No : 1	nama : Handuk Mandi Dewasa Jumbo Maestro 70 x 140	TERLARIS	harga: 27500 stok : 796
No : 2	nama : Handuk Mandi Hotel Polos Putih 70 x 140	TERLARIS	harga: 27500 stok : 390
No : 3	nama : Handuk Mandi Katun Bambu Morning Whistle by Terry Palmer	MW02 70x135	harga: 749 00 stok : 18
No : 4	nama : Rak Kamar Mandi Dinding Rak Bumbu Dapur Stainless Steel Rak Tempel Gantung Rak H	anduk dapat dipasang tanpa harus dilubangin dibagian	harga: 69000 stok : 548
No : 5	nama : VONE RH-05 Towel Bar Rak Handuk Dinding Kamar Mandi Stainless	harga: 109812	stok : 68
No : 6	nama : Handuk Mandi Hotel Premium 70 x 140 cm Polos	harga: 45900	stok : 998629
No : 7	nama : Handuk Mandi Dewasa Jumbo Maestro 70 x 140	TERLARIS	harga: 27500 stok : 796
No : 8	nama : Handuk Terry Palmer Premium 70x140cm	harga: 65000	stok : 1573
No : 9	nama : Handuk Mandi Dewasa Motif Chanel 70x140	harga: 25500	stok : 2
No : 10	nama : Handuk Microfiber Polos Ukuran 70 x 140 cm	harga: 26500	stok : 88821
No : 11	nama : Handuk Mandi Katun Bambu Morning Whistle by Terry Palmer	MW02 70x135	harga: 74 900 stok : 18
No : 12	nama : Grosir Handuk Merah Putih Besar 70 x 140 by Terry Palmer - Grosir Handuk Dewasa Lembut Murah	harga: 42900	stok : 795
No : 13	nama : Handuk Terry Palmer Premium 50x100 cm / Towel / Handuk Mandi	harga: 34000	stok :

Dan seterusnya sampai baris terakhir,

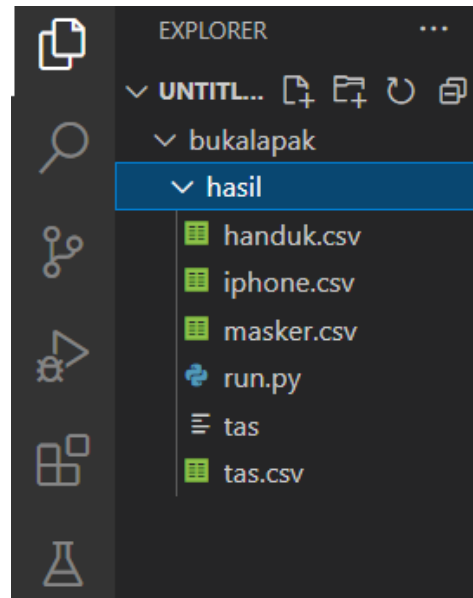
```

PROBLEMS      OUTPUT      DEBUG CONSOLE  TERMINAL
No : 489 nama : rak handuk stainless 5 baris plus kapstok 4kait PROMO TERMURAH harga: 165000 stok : 65
No : 490 nama : K90011 - Handuk Merah Putih - 135 x 70 - Murah Berkualitas harga: 67500 stok : 999
No : 491 nama : MOKHAMANO HANDUK MICROFIBER QUICKDRY HANDUK OLAHRAGA JNU0755 BLUE harga: 15200 stok : 995
No : 492 nama : gantungan handuk / rak handuk stainless fiorentino harga: 59000 stok : 1214
No : 493 nama : Washlap Handuk Chalmer Waslap Bayi Wash Cloth 13x17 cm harga: 6900 stok : 896
No : 494 nama : Sorex Handuk Mandi Dewasa Microfiber Towel Bright Lembut Daya Serap Tinggi HM 880 harga: 65900
No : 495 nama : Rak Handuk Gantung Dinding Aluminium Rak Dinding Kamar Mandi Atau Toilet harga: 51900 stok : 1
No : 496 nama : Rak handuk gantungan handuk dan baju harga: 120000 stok : 12
No : 497 nama : Handuk Hotel Polos Uk. 60x120cm Gramasi 600 GSM harga: 57600 stok : 183
No : 498 nama : Rak handuk gantungan baju kamar mandi harga: 165000 stok : 4
No : 499 nama : gantungan baju-capstok-kapstok-gantungan handuk harga: 55000 stok : 555
No : 500 nama : Handuk Mandi Motif Pelangi harga: 29000 stok : 592
PS D:\bukalapak>

```

Pada gambar diatas, terdapat 500 produk yang ditampilkan.

- Semua produk yang telah dicari melalui input user tersebut akan secara langsung di export menjadi file csv dan masuk kedalam folder ‘hasil’.



- Berikut ini merupakan tampilan dari salah satu file csv yang telah diexport.

	A	B	C
1	Nama	Harga	Stok
2	Handuk Mandi Dewasa Jumbo Maestro 70 x 140 TERLARIS	27500	796
3	Handuk Mandi Hotel Polos Putih 70 x 140 TERLARIS	27500	390
4	Handuk Mandi Katun Bambu Morning Whistle by Terry Palmer MW02 70x135	74900	18
5	Rak Kamar Mandi Dinding Rak Bumbu Dapur Stainless Steel Rak Tempel Gantung Rak Handuk dapat dipasang tanpa harus dilubangi dibagian	69000	548
6	VONE RH-05 Towel Bar Rak Handuk Dinding Kamar Mandi Stainless	109812	68
7	Handuk Mandi Hotel Premium 70 x 140 cm Polos	45900	998629
8	Handuk Mandi Dewasa Jumbo Maestro 70 x 140 TERLARIS	27500	796
9	Handuk Terry Palmer Premium 70x140cm	65000	1573
10	Handuk Mandi Dewasa Motif Chanel 70x140	25500	2
11	Handuk Microfiber Polos Ukuran 70 x 140 cm	26500	88821
12	Handuk Mandi Katun Bambu Morning Whistle by Terry Palmer MW02 70x135	74900	18
13	Grosir Handuk Merah Putih Besar 70 x 140 by Terry Palmer - Grosir Handuk Dewasa Lembut Murah	42900	795
14	Handuk Terry Palmer Premium 50x100 cm / Towel / Handuk Mandi	34000	937
15	Handuk Mandi Hotel Polos Putih 70 x 140 TERLARIS	27500	390
16	Handuk Mandi Dewasa Gucci Polos 70 x 140 TERLARIS	27500	243
17	Kintakun Handuk Tartan Katun 70 x 140 cm - Maroon	48950	8
18	Kintakun Handuk Premium Beglance Katun Navy 70 x 140 cm	120150	15
19	VONE TS-D504 Towel Shelf Rak Handuk Dinding Kamar Mandi Chrome	267600	32
20	Kintakun Handuk Butterfly Katun 70 x 140 cm - Yellow	48950	17
21	Kintakun Handuk Tartan Katun 70 x 140 cm - Tosca	48950	4
22	Handuk Mutia Dewasa - Handuk Mandi - High Quality Towel - Handuk Besar	43000	130
23	Rak Handuk Dinding Toilet Aluminium A345	49000	126
24	Handuk Mandi Dewasa SUNNY Morning Whistle by Terry Palmer uk.70x135cm Katun Bambu	74900	7
25	Handuk microfiber 70x140cm	69000	23
26	Handuk Mandi Dewasa Polos Ukuran 70x140 Lembut	27500	522
27	Handuk Mandi Bayi Jumbo Microfiber 70 x 140 cm Handuk Mandi Anak	14500	199282
28	Handuk Travel - Salon - Hotel Premium 50 x 100	19500	997603

	A	B	C
475	Handuk Terry Palmer Motif-1 70x140 Halus Premium Tebal Anti-bakteri	84500	48
476	handuk sport souvenir 30x70 polos katun halus	15000	4645
477	Handuk Terry Palmer Signature Tebal 550gram 70x140cm	84000	37
478	Rak Gantungan Handuk Kamar Mandi Bathroom Towel Holder - TK4	51600	70
479	TOTO Rak Handuk - Combination Towel Shelf & Towel Bar TX726AES	1785000	87
480	Gantungan handuk kamar mandi Aluminium A345	58000	5
481	BEST SELLER HANDUK MUKA SAPU TANGAN CHALMER 30X30	8900	974
482	HANDUK KIMONO DEWASA POLOS HITAM Size XL	235000	3
483	Kapstok Dinding Gantungan Handuk Baju Kamar Mandi Toilet WC Dapur S685	33600	932
484	VONE TS-D601 Towel Shelf Rak Gantungan Handuk Dinding Kamar Mandi	165000	138
485	VONE HK-M03 Gantungan Baju Handuk Satuan Kapstok Dinding Stainless	18600	264
486	Rak handuk lipat-gantungan handuk sterilis-gantungan handuk kamar mandi	155000	27
487	handuk sapu tangan oshibori corak kotak 30x30 tebal serap	8000	1181
488	Gantungan Handuk Kamar Mandi Bathroom Towel Holder TKU004	51800	25
489	Handuk Lexus Tanggung by Toyo Terry	38500	46
490	rak handuk stainless 5 baris plus kapstok 4kait PROMO TERMURAH	165000	65
491	K90011 - Handuk Merah Putih - 135 x 70 - Murah Berkualitas	67500	999
492	MOKHAMANO HANDUK MICROFIBER QUICKDRY HANDUK OLAHRAGA JNU0755 BLUE	15200	995
493	gantungan handuk / rak handuk stainless fiorentino	59000	1214
494	Washlap Handuk Chalmer Waslap Bayi Wash Cloth 13x17 cm	6900	896
495	Sorex Handuk Mandi Dewasa Microfiber Towel Bright Lembut Daya Serap Tinggi HM 880	65900	57
496	Rak Handuk Gantung Dinding Aluminium Rak Dinding Kamar Mandi Atau Toilet	51900	195
497	Rak handuk gantungan handuk dan baju	120000	12
498	Handuk Hotel Polos Uk. 60x120cm Gramasi 600 GSM	57600	183
499	Rak handuk gantungan baju kamar mandi	165000	4
500	gantungan baju-kapstok-kapstok-gantungan handuk	55000	555
501	Handuk Mandi Motif Pelangi	29000	592
502			

VI. KESIMPULAN

Berdasarkan hasil implementasi webscraping pada website bukalapak dapat ditarik kesimpulan sebagai berikut:

- 1) Website bukalapak ini yang dihasilkan dengan menerapkan teknik web scraping ini berhasil mengekstrak informasi yang di simpan dalam bentuk news list.
- 2) Website berhasil menyimpan otomatis data hasil scraping dan history pada database.
- 3) Web scraping merupakan tindakan legal karena pada penelitian ini data diambil dari pihak yang sudah memberikan izin demi tujuan yang lain yakni meningkatkan promosi dan pemasaran produk. Web scraping di sini tidak melakukan pencurian data, manipulasi informasi, dsb. Bahkan web scraping dapat memberikan mutual simbiosis dengan meningkatnya trafik atas sumber asli link yang di-scrape.

LINK GITHUB

https://github.com/fennyanggraini17/Web-Scraping-Fenny_Anggraini_SK4A_Reg.git