

## HW Session 13 – Web Scrapping

- Please create a scraping script that giving the result of csv file with the product and price name

notes : you could choose your own method and website, but if not please use this

: [https://shopee.co.id/api/v4/search/search\\_items?by=relevancy&limit=60&match\\_id=1028154&newest=0&order=desc&page\\_type=collection&scenario=PAGE\\_COLLECTION\\_SEARCH&version=2](https://shopee.co.id/api/v4/search/search_items?by=relevancy&limit=60&match_id=1028154&newest=0&order=desc&page_type=collection&scenario=PAGE_COLLECTION_SEARCH&version=2) using urllib library in python

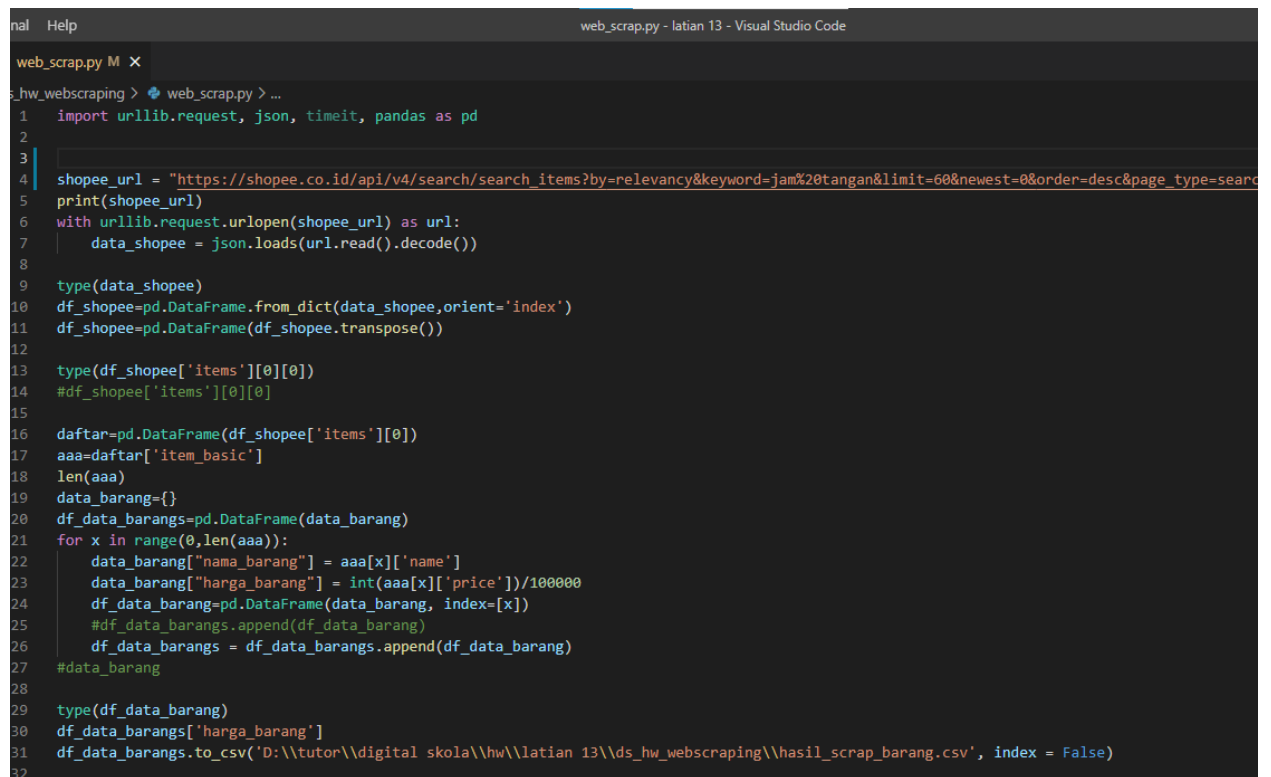
### 1. Link Github

[https://github.com/tianjol/ds\\_hw\\_webscraping.git](https://github.com/tianjol/ds_hw_webscraping.git)

### 2. Python Script (detail script on spare file or github)

```
shopee_url =  
https://shopee.co.id/api/v4/search/search\_items?by=relevancy&keyword=jam%20tangan&limit=60&newest=0&order=desc&page\_type=search&scenario=PAGE\_GLOBAL\_SEARCH&version=2
```

```
>>> scrap "jam tangan from shopee"
```



```
web_scrap.py X  
ds_hw_webscraping > web_scrap.py > ...  
1 import urllib.request, json, timeit, pandas as pd  
2  
3  
4 shopee_url = "https://shopee.co.id/api/v4/search/search_items?by=relevancy&keyword=jam%20tangan&limit=60&newest=0&order=desc&page_type=search  
5 print(shopee_url)  
6 with urllib.request.urlopen(shopee_url) as url:  
7     data_shopee = json.loads(url.read().decode())  
8  
9 type(data_shopee)  
10 df_shopee=pd.DataFrame.from_dict(data_shopee,orient='index')  
11 df_shopee=pd.DataFrame(df_shopee.transpose())  
12  
13 type(df_shopee['items'][0][0])  
14 #df_shopee['items'][0][0]  
15  
16 daftar=pd.DataFrame(df_shopee['items'][0])  
17 aaa=daftar['item_basic']  
18 len(aaa)  
19 data_barang={}  
20 df_data_barangs=pd.DataFrame(data_barang)  
21 for x in range(0,len(aaa)):  
22     data_barang["nama_barang"] = aaa[x]['name']  
23     data_barang["harga_barang"] = int(aaa[x]['price']/100000  
24     df_data_barang=pd.DataFrame(data_barang, index=[x])  
25     #df_data_barangs.append(df_data_barang)  
26     df_data_barangs = df_data_barangs.append(df_data_barang)  
27 #data_barang  
28  
29 type(df_data_barang)  
30 df_data_barangs['harga_barang']  
31 df_data_barangs.to_csv('D:\\tutor\\digital skola\\hw\\latian 13\\ds_hw_webscraping\\hasil_scrap_barang.csv', index = False)  
32
```

### 3. Result

hasil_scrap_barang.csv - Excel		
File Home Insert Page Layout Formulas Data Review View Developer Help Tell me what you want to do		
Clipboard Font Alignment Number Styles		
A10 X ✓ f JAM TANGAN DIGITAL RUBBER PRIA / WANITA / ANAK ANAK / REMAJA SEGI DW-5600 GROSIR KARET RUBBER SPORT LED KOREA SPORT ELEKTRONIK MOTIF BUNGA DAN POLOS STYLE COUPLE WATCHBRANDS MURAH JF0		
	A	B C
1	nama_barang	harga_barang
2	JAM TANGAN DIGITAL RUBBER PRIA / WANITA / ANAK ANAK / REMAJA SEGI DW-5600 GROSIR KARET RUBBER SPORT LED KOREA SPORT ELEKTRONIK MOTIF BUNGA DAN POLOS STYLE COUPLE WATCHBRANDS MURAH JF052 / C8	11300
3	CODâ€¦ Jam Tangan Digital Wanita Pria LED Rubber Watch Anti Air Ready Warna Harga Grosir S1192	16890
4	C ( C00023 ) JAM TANGAN LED DEWASA MODEL KOREA / HARGA GROSIR	3999
5	WBS JAM TANGAN COUPLE FREE BATERAI SEIKO 5 MURAH RANTAI ANALOG CASUAL PERMIUM IMPORT GROSIR PRIA WANITA â€¦COD ANAK REMAJA DEWASA SK002	18300
6	Jam Tangan Analog Pria Stainless Steel Quartz Fashion Casual Pria Import	17200
7	Jam Tangan Quartz Strap PU Analog Casual Gaya Retro Untuk Wanita Watch Korean Fashion	16800
8	WBS â€¦COD JAM TANGAN COUPLE ALBA RANTAI GROSIR IMPORT PRIA WANITA PASANGAN MURAH PREMIUM AL013	18999
9	SmartWatch Y68 Sport Tahan Air Bluetooth Smart Watch Pelacak Kebugaran Gelang Pedometer Heart Rate	88000
10	JAM TANGAN DIGITAL RUBBER PRIA / WANITA / ANAK ANAK / REMAJA SEGI DW-5600 GROSIR KARET RUBBER SPORT LED KOREA SPORT ELEKTRONIK MOTIF BUNGA DAN POLOS STYLE COUPLE WATCHBRANDS MURAH JF052 / C8	11300
11	CODâ€¦ Jam tangan LED karet digital rubber electronic fashion jam wanita pria anak-anak watch import S1166	13700
12	Jam Tangan GENEVA STAINLESS STEEL 3 TITIK - New Watch Analog Stainless Steel - Fashion Wanita	13280
13	JAM TANGAN PRIA WANITA DIGITAL RUBBER SPORT BULAT LED MOTIF BUNGA DAN POLOS COUPLE ANAK REMAJA DAN DEWASA GROSIR MURAH IMPORT JF063	12750
14	JAM TANGAN LED PRIA/WANITA APPLE	26000
15	Jam Tangan Elektronik Motif Bunga Daisy Kecil Anti Air Untuk Wanita / Jam Tangan Motif Bunga	11000
16	Jam Tangan Pria Gelang Set Box ( FREE 1pcs Gelang random ) - Jam Tangan Pria GNV90075L	35040
17	Jam Tangan SHSHD RANTAI - Jam Tangan Strap Stainless Steel/Jam Tangan Pria Terbaru/Fashion Pria	13950
18	WBS JAM TANGAN PRIA SPORT ANALOG KULIT QUICKSILVER SURF TANGGAL HARI MURAH â€¦COD WANITA COUPLE QS003	20400
19	â€¦COD MESIN JAPAN !! JAM TANGAN LED WANITA PRIA RUBBER DIGITAL WATCH MURAH BERKUALITAS JF037	5460
20	WBS â€¦COD Jam Tangan Pria Wanita Sport Rantai Geneva Stainless Steel Korea Fashion Grosir GN042/C7	13250
21	â€¦COD "Penjualan Terbaikâ€¦ 100%ORI JAM TANGAN PRIA +FREE BOX Baterai KULIT PREMIUM DIAMETER 4.2CM   JAMTANGAN FASHION PRIA	89888
22	JAM TANGAN LED DIGITAL FASHION KISD&TEENAGER(REMAJA) DISPLAY WATCH ELEKTRONIK IMPORT LED WANITA PRIA	5460
23	CODâ€¦ Jam Tangan Wanita Pria LED MI Jam Digital Sport Fashion Watch Rubber Tali Warna Kontras Grosir Murah S1181	15800
24	â€¦ JIMALL â€¦ JT002 Jam Tangan Led Digital Watch Pria Wanita Fashion Sport Keren Souvenir Import Grosir	4199
25	WBS Jam Tangan SHSHD Pria dan Wanita Sport Chrono Rariasi Rantai Geneva Jam Tangan Stainless Steel Jam Tangan Pria Fashion Analog Grosir Model Terbaru SH003	13950
26	Jam Tangan Casual Gaya Retro Korea CandyCat Untuk Wanita Strap PU / Jam Tangan Wanita Casual	14500
27	â€¦COD JAM TANGAN RIURRFR S-SPORT S-1790 DIGITAL FASHION I FD IF038 IMPORT GROSIR MURAH PRIA WANITA	7250