

```
!pip install bs4
```

```
Collecting bs4
  Downloading bs4-0.0.2-py2.py3-none-any.whl (1.2 kB)
Requirement already satisfied: beautifulsoup4 in /usr/local/lib/python3.10/dist-packages (from bs4) (4.12.3)
Requirement already satisfied: soupsieve>1.2 in /usr/local/lib/python3.10/dist-packages (from beautifulsoup4->bs4) (2.5)
Installing collected packages: bs4
Successfully installed bs4-0.0.2
```

```
!pip install requests
```

```
Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (2.31.0)
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests) (3.6)
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests) (2.0.7)
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests) (2024.2.2)
```

```
from bs4 import BeautifulSoup
import requests
import pandas as pd
```

```
URL="https://www.amazon.com/GoPro-HERO12-Black-Waterproof-Stabilization/dp/B0CDDY8F9L/ref=sr_1_3?dib=eyJ2IjojMSJ9.tdrNeFgZKWOI_mMeamikyl"
```

```
HEADERS="Mozilla/5.0 (X11; Ubuntu; Linux x86_64; rv:123.0) Gecko/20100101 Firefox/123.0"
```

```
webpage = requests.get(URL,HEADERS)
```

```
webpage
```

```
<Response [200]>
```

```
soup=BeautifulSoup(webpage.content,"html.parser")
```

```
soup

<!DOCTYPE html>
<html class="a-no-js" data-19ax5a9jf="dingo" lang="en-us"><!-- sp:feature:head-start -->
<head><script>var aPageStart = (new Date()).getTime();</script><meta charset="utf-8"/>
<!-- sp:end-feature:head-start -->
<!-- sp:feature:csm:head-open-part1 -->
<script type="text/javascript">var ue_t0=ue_t0||new Date();</script>
<!-- sp:end-feature:csm:head-open-part1 -->
<!-- sp:feature:cs-optimization -->
<meta content="on" http-equiv="x-dns-prefetch-control"/>
<link href="https://images-na.ssl-images-amazon.com" rel="dns-prefetch"/>
<link href="https://m.media-amazon.com" rel="dns-prefetch"/>
<link href="https://completion.amazon.com" rel="dns-prefetch"/>
<!-- sp:end-feature:cs-optimization -->
<!-- sp:feature:csm:head-open-part2 -->
<script type="text/javascript">
window.ue_ihb = (window.ue_ihb || window.ueinit || 0) + 1;
if (window.ue_ihb === 1) {

var ue_csm = window,
    ue_hob = +new Date();
(function(d){var e=d.ue||{};f=Date.now||function(){return+new Date};e.d=function(b){return f()-(b?
0:d.ue_t0)};e.stub=function(b,a){if(!b[a]){var c=[];b[a]=function()
{c.push([c.slice.call(arguments),e.d(),d.ue_id])};b[a].replay=function(b){for(var
a;a=c.shift();){b(a[0],a[1],a[2])};b[a].isStub=1}};e.exec=function(b,a){return function(){try{return
b.apply(this,arguments)}catch(c){ueLogError(c,{attribution:a||"undefined",logLevel:"WARN"})}}}})(ue_csm);

var ue_err_chan = 'jserr-rw';
(function(d,e){function h(f,b){if(!(a.ec>a.mxe)&&f){a.ter.push(f);b=b||{};var
c=f.logLevel||b.logLevel;c&&c!=k&&c!=m&&c!=n&&c!=p||a.ec++;c&&c!=k||a.ecf++;b.pageURL=""+(e.location?
e.location.href:"");b.logLevel=c;b.attribution=f.attribution||b.attribution;a.erl.push({ex:f,info:b})}function l(a,b,c,e,g)
{d.ueLogError({m:a,f:b,l:c,c:""+e,err:g,fromOnError:1,args:arguments),g?{attribution:g.attribution,logLevel:g.logLevel}:void
0);return l}var k="FATAL",m="ERROR",n="WARN",p="DOWNGRADED",a={ec:0,ecf:0,
pec:0,ts:0,erl:[],ter:[],buffer:[],mxe:50,startTimer:function(){a.ts++;setInterval(function()
{d.ue&&a.pec<a.ec&&d.ue("at");a.pec=a.ec},1E4)};l.skipTrace=1;h.skipTrace=1;h.isStub=1;d.ueLogError=h;d.ue_err=a;e.onerror=1}}
(ue_csm>window);

var ue_id = 'PK8T5YB1D8DC8A4YHH1Q',
    ue_url = '/rd/uedata',
    ue_navtiming = 1,
    ue_mid = 'ATVPDKIKX0DER',
    ue_sid = '000-0000000-8675309',
```

```

ue_sn = 'www.amazon.com',
ue_furl = 'fls-na.amazon.com',
ue_surl = 'https://unagi-na.amazon.com/1/events/com.amazon.csm.nexusclient.prod',
ue_int = 0,
ue_fcsn = 1,
ue_urt = 3,
ue_rpl_ns = 'cel-rpl',
ue_ddq = 1,
ue_fpf = '//fls-na.amazon.com/1/batch/1/OP/ATVPDKIKX0DER:000-0000000-8675309:PK8T5YB1D8DC8A4YHH1Q$uedata=s:',
ue_sbuimp = 1,
ue_ibft = 0,
ue_sswmts = 0,
ue_jsmtf = 0,
ue_fnt = 0.

```

```
links=soup.find_all("a")
```

```
links[3]
```

```

<a aria-label="Choose a language for shopping." class="nav-a nav-a-2 icp-link-style-2" href="/customer-preferences/edit?
ie=UTF8&preferencesReturnUrl=%2F&ref=topnav_lang" id="icp-nav-flyout">
<span class="icp-nav-link-inner">
<span class="nav-line-1">
</span>
<span class="nav-line-2">
<span class="icp-nav-flag icp-nav-flag-us icp-nav-flag-lop"></span>
<div>EN</div>
<span class="nav-icon nav-arrow"></span>
</span>
</span>
</a>

```

```
link=links[3].get('href')
```

```
link
```

```
 '/customer-preferences/edit?ie=UTF8&preferencesReturnUrl=%2F&ref=topnav_lang '
```

```
product_list="https://amazon.com"+link
```

```
product_list
```

```
 'https://amazon.com/customer-preferences/edit?ie=UTF8&preferencesReturnUrl=%2F&ref=topnav_lang '
```

```
new_webpage=requests.get(product_list,HEADERS)
```

```
new_webpage
```

```
<Response [200]>
```