

```
# import module
import requests
from bs4 import BeautifulSoup
```

```
HEADERS = ({'User-Agent':
             'Mozilla/5.0 (Windows NT 10.0; Win64; x64) \
             AppleWebKit/537.36 (KHTML, like Gecko) \
             Chrome/90.0.4430.212 Safari/537.36',
             'Accept-Language': 'en-US, en;q=0.5'})
```

```
# user define function
# Scrape the data
def getdata(url):
    r = requests.get(url, headers=HEADERS)
    return r.text
```

```
def html_code(url):

    # pass the url
    # into getdata function
    htmldata = getdata(url)
    soup = BeautifulSoup(htmldata, 'html.parser')

    # display html code
    return (soup)
```

```
url = "https://www.amazon.in/HP-245-G8-AMD-Ryzen/dp/B093L8FSP2?th=1"
```

```
soup = html_code(url)
print(soup)
```

```
<!-- DOCTYPE doctype html>
<html class="a-no-js" data-19ax5a9jf="dingo" lang="en-in"><!-- sp:feature:head-sta
<head><script>var aPageStart = (new Date()).getTime();</script><meta charset="utf-
<!-- sp:end-feature:head-start -->
<!-- sp:feature:argonath-data -->
<script data-argonath="t2">document.write("<div id='argonath_loader' style='positi
<!-- sp:end-feature:argonath-data -->
<script type="text/javascript">var ue_t0=ue_t0||new Date();</script>
<!-- sp:feature:cs-optimization -->
<meta content="on" http-equiv="x-dns-prefetch-control"/>
<link href="https://images-eu.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 -->
<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=d.ue||{},f=Date.now||function(){return+new Date};e.d=funct
```

```

    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.
    pec:0,ts:0,erl:[],ter:[],mxe:50,startTimer:function(){a.ts++;setInterval(function(

var ue_id = '32EJ7RQW08T3JGGDH8AE',
    ue_url = '/rd/uedata',
    ue_navtiming = 1,
    ue_mid = 'A21TJRUUN4KGV',
    ue_sid = '257-6432916-4637862',
    ue_sn = 'www.amazon.in',
    ue_furl = 'fls-eu.amazon.in',
    ue_surl = 'https://unagi-eu.amazon.com/1/events/com.amazon.csm.nexusclient.prc
    ue_int = 0,
    ue_fcsn = 1,
    ue_urt = 3,
    ue_rpl_ns = 'cel-rpl',
    ue_ddq = 1,
    ue_fpf = '//fls-eu.amazon.in/1/batch/1/OP/A21TJRUUN4KGV:257-6432916-4637862:32
    ue_sbuimp = 1,
    ue_cel_lclia = 1,
    ue_ibft = 0,
    ue_sswmts = 1,
    ue_jsmtf = 0,
    ue_fnt = 0,

    ue_swi = 1;
var ue_viz=function(){(function(b,e,a){function k(c){if(b.ue.viz.length<p&&!l){var
"isibilityState";k({});d&&e.addEventListener(g,k,0);b.ue&&d&&(b.ue.pageViz={event:

(function(d,h,N){function H(a){return a&&a.replace&&a.replace(/^\s+|\s+$/g,"")}fun
c){return c+b}}),d.ue_sid=b),c&&a.tag("page-source:"+c),d.ue_fpf=w}function P(){va
function(a){f?(e(a),l(a)):(l(a),e(a))};b[c]&&(b[c].isUeh=1)}function S(m,b,c,q){fu

```

#Scrape customer name

```

def cus_data(soup):
    # find the Html tag
    # with find()
    # and convert into string
    data_str = ""
    cus_list = []

    for item in soup.find_all("span", class_="a-profile-name"):
        data_str = data_str + item.get_text()
        cus_list.append(data_str)
        data_str = ""
    return cus_list

```

```

cus_res = cus_data(soup)
print(cus_res)

```

```

['Amazon Customer', 'Naveen', 'JoKrsna', 'JoKrsna', 'Jacob Sundar Raj', 'Vibha', 'nil

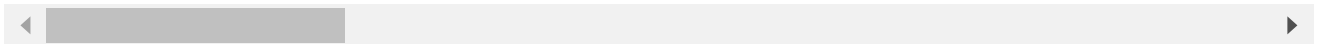
```

```
#Scrape User Review
def cus_rev(soup):
    # find the Html tag
    # with find()
    # and convert into string
    data_str = ""

    for item in soup.find_all("div", class_="a-expander-content reviewText review-text-con
        data_str = data_str + item.get_text()
    result = data_str.split("\n")
    return (result)

rev_data = cus_rev(soup)
rev_result = []
for i in rev_data:
    if i is "":
        pass
    else:
        rev_result.append(i)
rev_result

[' Good product it has AMD processor hence no lag at all. Plenty of storage, boot ti
' The charging wire is damaged...request to replace the same immediately,',
' "Camera not found" issue there for last 3 weeks: Worst experience in contacting c
' Good laptop for home use and office. Low graphics game can be played. Battery goo
' Its been a month I am using it. Classy look with good performance. Initially i do
' Reviewing after 2 months of use -0 - 10Speed - 0Display - 5Battery - 7Connectivit
' unexpectedly slow with 8 cores and 8gb ram with a measly 720p screen, that is jus
' Battery back up only 2.30 hrs.screen is not down.. charger cable length is very s
```



```
#Scraping Production information
def product_info(soup):

    # find the Html tag
    # with find()
    # and convert into string
    data_str = ""
    pro_info = []

    for item in soup.find_all("ul", class_="a-unordered-list a-vertical a-spacing-mini"):
        data_str = data_str + item.get_text()
        pro_info.append(data_str.split("\n"))
        data_str = ""
    return pro_info

pro_result = product_info(soup)

# Filter the required data
for item in pro_result:
    for j in item:
        if j is "":
```

```
    pass
else:
    print(j)
```

AMD Ryzen 5 3500U APU with Radeon Vega 8 Graphics (2.1 GHz base clock, up to 3.7 GHz boost)

#Scraping Review Image:

```
def rev_img(soup):
```

```
    # find the Html tag
    # with find()
    # and convert into string
    data_str = ""
    cus_list = []
    images = []
    for img in soup.findAll('img', class_="review-image-tile"):
        images.append(img.get('src'))
    return images
```

```
img_result = rev_img(soup)
```

```
img_result
```

```
['https://images-na.ssl-images-amazon.com/images/G/01/x-locale/common/grey-pixel.gif',
 'https://images-na.ssl-images-amazon.com/images/G/01/x-locale/common/grey-pixel.gif',
 'https://images-na.ssl-images-amazon.com/images/G/01/x-locale/common/grey-pixel.gif']
```