



In [7]: *#First we should import all the libraries*
#smtp -- simple mail transfer protocol

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

In [8]: `url = 'https://www.amazon.in/s?k=playstation+5&crd=4XCP6JRIR7LH&sprefix=%2Caps%2C740&ref=nb_sb_ss_recent_2_0_recent'`

In [9]: *#Now we have to send a request to http of that link, to do that we have to include user agent and accept language as headers beco*
#Every link has a user agent and accept language as headers.

#User agents -- it is basically tells us that you are trying to access this website and you are a genuine user by identifying our
#and some other required info.

In [10]: *#Headers for request*

```
headers = ({'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/114.0.0.0 Safa
```

In [11]: *#HTTP request*

```
webpage = requests.get(url, headers=headers)
```

In [12]: `webpage`

Out[12]: `<Response [200]>`

In [13]: *#<Response [200]> means our http request was successfull*

In [14]: webpage.content

```
#This is in bytes formate
```

```

<!-- sp:feature:csm:head-open-part2 --><script type='text/javascript'\>\nwindow.ue_ihb = (window.ue_ihb || window.ueini
t || 0) + 1;\nif (window.ue_ihb === 1) {\n\nvar ue_csm = window,\n    ue_hob = +new Date();\n(function(d){var e=d.ue=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);\n\n\n    var ue_err_chan = '\jserr-rw';\n(function(d,e){functio
n h(f,b){if(!a.ec<a.mxe)&&f){a.ter.push(f);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!1}var k="FATAL",m="ERROR",n="WARN",p="DOWNGRADED",a={ec:0,ecf:0,\npec:0,t
s: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);\n\n\nvar ue_id =
'\6FWYSTZMDNS8PY1W0A58',\n    ue_url = '\rd/uedata',\n    ue_navtiming = 1,\n    ue_mid = '\A21TJRUN4KGV',\n    ue_sid
= '\259-4231219-9265649',\n    ue_sn = '\www.amazon.in',\n    ue_furl = '\fls-eu.amazon.in',\n    ue_surl = '\https://una
gi-eu.amazon.com/1/events/com.amazon.csm.nexusclient.prod',\n    ue_int = 0,\n    ue_fcsn = 1,\n    ue_urt = 3,\n    ue_rpl
_ns = '\cel-rpl',\n    ue_ddq = 1,\n    ue_fpf = '\fls-eu.amazon.in/1/batch/1/OP/A21TJRUN4KGV:259-4231219-9265649:6FWYST
ZMDNS8PY1W0A58$uedata=s:',\n    ue_sbuimp = 1,\n    ue_ibft = 0,\n    ue_sswmts = 0,\n    ue_jsmtf = 0,\n    ue_fnt = 0,\n
ue_lpsi = 6000,\n    ue_no_counters = 1,\n\n    ue_swi = 1;\nvar ue_viz=function(){function k(c){if(b.ue.vi
z.length<p&&!1){var a=c.type;c=c.originalEvent;/^focus./.test(a)&&c&&(c.toElement||c.fromElement||c.relatedTarget)||a=e[m]

```

In [15]: #Change from bytes formate into html

```
soup = BeautifulSoup(webpage.content, 'html.parser') #This means send your web content and parse it into html formate.
```

In [16]: print(soup)

```
<!DOCTYPE html>
<html class="a-no-js" data-19ax5a9jf="dingo" lang="en-in"><!-- 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 crossorigin="" href="https://images-eu.ssl-images-amazon.com" rel="dns-prefetch"/>
<link crossorigin="" href="https://images-eu.ssl-images-amazon.com" rel="preconnect"/>
<link crossorigin="" href="https://m.media-amazon.com" rel="dns-prefetch"/>
<link crossorigin="" href="https://m.media-amazon.com" rel="preconnect"/>
<link crossorigin="" href="https://completion.amazon.com" rel="dns-prefetch"/>
<link crossorigin="" href="https://completion.amazon.com" rel="preconnect"/>
<!-- sp:end-feature:cs-optimization -->
<!-- sp:feature:csm:head-open-part2 -->
<script type="text/javascript">
```

Fetch links as list of tag objects

```
In [17]: links = soup.find_all('a', attrs={'class': 'a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal'})
```

```
In [18]: links
```

```
Out[18]: [<a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Sony-PS5-PlayStation-Consol
e/dp/B0BRC72X8/ref=sr_1_1?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-1"
target="_blank"><span class="a-size-medium a-color-base a-text-normal">Sony PS5 PlayStation Console</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Sony-PS5-Digital-Standalone/
dp/B0BSNHFVL4/ref=sr_1_2?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-2" t
arget="_blank"><span class="a-size-medium a-color-base a-text-normal">Sony PS5 Digital Standalone</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Sony-PlayStation-Console-Rag
narok-Standard/dp/B0BXR5NB6Z/ref=sr_1_3?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C74
0&sr=8-3" target="_blank"><span class="a-size-medium a-color-base a-text-normal">Sony PS5 PlayStation Console+God Of War Ra
gnarok | Standard Edition | PS5 Game (PlayStation 5)</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/F1-23-Standard-PlayStation-
5/dp/B0C4T3CYW7/ref=sr_1_4?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-4"
target="_blank"><span class="a-size-medium a-color-base a-text-normal">F1 23 | Standard Edition | PlayStation 5</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Moto-GP-23-Standard-PlayStat
ion/dp/B0C4TGYJC2/ref=sr_1_5?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-
5" target="_blank"><span class="a-size-medium a-color-base a-text-normal">Moto GP 23 | Standard Edition | PlayStation 5</span>
</a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Hogwarts-Legacy-Standard-Pla
yStation-PS5/dp/B0BBV2KB8R/ref=sr_1_6?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&
sr=8-6" target="_blank"><span class="a-size-medium a-color-base a-text-normal">Hogwarts Legacy | Standard Edition| PlayStat
ion 5 (PS5)</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Sony-Ragnarok-Standard-Game-
PlayStation/dp/B0B6FGSKCQ/ref=sr_1_7?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&a
mp;sr=8-7" target="_blank"><span class="a-size-medium a-color-base a-text-normal">Sony God Of War Ragnarok | Standard Edition |
PS5 Game (PlayStation 5)</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/PowerWash-Simulator-Standard
-PlayStation-5/dp/B0C4TG2C54/ref=sr_1_8?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C74
0&sr=8-8" target="_blank"><span class="a-size-medium a-color-base a-text-normal">PowerWash Simulator | Standard Edition | P
layStation 5</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/PS5-Grand-Theft-Auto-V/dp/B0
9XJ8FGVP/ref=sr_1_9?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-9" target
="_blank"><span class="a-size-medium a-color-base a-text-normal">PS5 Grand Theft Auto V</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/PS5-Horizon-Forbidden-West-s
tandard/dp/B09FJX6JSZ/ref=sr_1_10?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-10" target="_blank"><span class="a-size-medium a-color-base a-text-normal">Sony Horizon Forbidden West | Standard Edition
| PS5 Game (PlayStation 5)</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/Sony-PS5-CoD-MWII/dp/B0C7CJF
SZR/ref=sr_1_11?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C740&sr=8-11" target="_
blank"><span class="a-size-medium a-color-base a-text-normal">Sony PS5 CoD MWII</span> </a>,
  <a class="a-link-normal s-underline-text s-underline-link-text s-link-style a-text-normal" href="/PlayStation-5-DualSense-Wire
less-Controller/dp/B08H99BPJN/ref=sr_1_12?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&prefix=%2Caps%2C
740&sr=8-12" target="_blank"><span class="a-size-medium a-color-base a-text-normal">PlayStation 5 DualSense Wireless Controller</span> </a>]
```

-PlayStation 5/>dp/B0C4TG2C54/ref=sr_1_8?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&sprefix=%2Caps%2C740&sr=8-8" target="_blank">PowerWash Simulator | Standard Edition | PlayStation 5/>

PS5 Grand Theft Auto V

Sony Horizon Forbidden West | Standard Edition | PS5 Game (PlayStation 5)

Sony PS5 CoD MWII

PlayStation 5 DualSense Wireless Controller

PS5 Marvel's Avengers

PS5 Assassin's Creed Valhalla

Dorca Horizontal Dust Cover for PS5 Console Dual Guard with Back Cable Port for Sony Playstation 5 Console Digital & Disk Version

TCOS TECH Mobile Phone Clip Mount Holder for PS5 Controller Mobile Phone Stand Mount Clamp Bracket with Adjustable Stand for Playstation 5 Wireless Controller

In [19]: `links[0]`

Out[19]: `Sony PS5 PlayStation Console `

In [20]: `#pull href link`

`link = links[0].get('href')`

In [21]: *#Append product*

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

In [22]: `product_list`

Out[22]: 'https://amazon.in/Sony-PS5-PlayStation-Console/dp/B08RCP72X8/ref=sr_1_1?crid=4XCP6JRIR7LH&keywords=playstation+5&qid=1689344738&srefix=%2Caps%2C740&sr=8-1'

In [23]: *#Now we need to connect our system link to the above link, to do that,*

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

In [24]: `new_webpage`

Out[24]: <Response [200]>

In [25]: *#Parse HTML to pull the link into html formate*

```
new_soup = BeautifulSoup(new_webpage.content, 'html.parser')
```

In [26]: `new_soup`

Out[26]: <!DOCTYPE html>

```
<!--[if lt IE 7]> <html lang="en-us" class="a-no-js a-lt-ie9 a-lt-ie8 a-lt-ie7"> <![endif]-->
<!--[if IE 7]> <html lang="en-us" class="a-no-js a-lt-ie9 a-lt-ie8"> <![endif]-->
<!--[if IE 8]> <html lang="en-us" class="a-no-js a-lt-ie9"> <![endif]-->
<!--[if gt IE 8]><!-->
<html class="a-no-js" lang="en-us"><!--<![endif]--><head>
<meta content="text/html; charset=utf-8" http-equiv="content-type"/>
<meta charset="utf-8"/>
<meta content="IE=edge,chrome=1" http-equiv="X-UA-Compatible"/>
<title dir="ltr">Amazon.in</title>
<meta content="width=device-width" name="viewport"/>
<link href="https://images-na.ssl-images-amazon.com/images/G/01/AUIClients/AmazonUI-3c913031596ca78a3768f4e934b1cc02ce23810
1.secure.min._V1_.css" rel="stylesheet"/>
<script>

if (true === true) {
  var ue_t0 = (+ new Date()),
      ue_csm = window,
```

In [31]: `product_name = soup.find("span", attrs={"class": 'a-size-medium a-color-base a-text-normal'}).text`


```
In [31]: product_name = soup.find("span", attrs={"class": 'a-size-medium a-color-base a-text-normal'}).text
```

```
In [32]: #new_soup.find("span", attrs={"id": 'ProductTitle'})
```

```
In [33]: #Pull the price of specific product
```

```
product_price = soup.find("span", attrs={"class": 'a-price'}).find('span', attrs={'class': 'a-offscreen'}).text
```

```
In [34]: #Pull Ratings
```

```
product_rating = soup.find("span", attrs={"class": 'a-icon-alt'}).text
```

```
In [37]: print(product_name, ",", product_price, ",", product_rating)
```

```
Sony PS5 PlayStation Console , ₹54,300 , 4.4 out of 5 stars
```