

```
In [ ]: pip install BeautifulSoup
```

```
In [ ]: pip install bs4
```

```
In [1]: from bs4 import BeautifulSoup
import requests
import pandas as pd
```

```
In [2]: url="https://www.bewakoof.com/men-clothing"
```

```
In [3]: soup=requests.get(url)
soup
```

```
Out[3]: <Response [200]>
```

```
In [4]: page=requests.get(url)
page.text
```

```

1.1.1;text=1;time=1;url=1;week=1;text=0;1;select=1;password=0;},re
urn{proxy:{assets:void 0,beacon:void 0},privacy:{cookies_enabled:!0},ajax:
{deny_list:void 0,block_internal:!0,enabled:!0,harvestTimeSeconds:10,autoSt
art:!0},distributed_tracing:{enabled:void 0,exclude_newrelic_header:void 0,
cors_use_newrelic_header:void 0,cors_use_tracecontext_headers:void 0,allowe
d_origins:void 0},session:{domain:void 0,expiresMs:d.oD,inactiveMs:d.Hb},ss
l:void 0,obfuscate:void 0,jserrors:{enabled:!0,harvestTimeSeconds:10,autoSt
art:!0},metrics:{enabled:!0,autoStart:!0},page_action:{enabled:!0,harvestTi
meSeconds:30,autoStart:!0},page_view_event:{enabled:!0,autoStart:!0},page_v
iew_timing:{enabled:!0,harvestTimeSeconds:30,long_task:!1,autoStart:!0},ses
sion_trace:{enabled:!0,harvestTimeSeconds:10,autoStart:!0},harvest:{tooMany
RequestsDelay:60},session_replay:{autoStart:!0,enabled:!1,harvestTimeSecond
s:60,sampling_rate:50,error_sampling_rate:50,collect_fonts:!1,inline_image
s:!1,inline_stylesheet:!0,mask_all_inputs:!0,get mask_text_selector(){retur
n e.mask_selector},set mask_text_selector(t){u(t)?e.mask_selector=t+", [data
-nr-mask]":null===t?e.mask_selector=t:(0,1.Z)("An invalid session_replay.ma
sk_selector was provided and will not be used",t)},get block_class(){retur
n"nr-block"},get ignore_class(){return"nr-ignore"},get mask_text_class(){re
turn"nr-mask"},get block_selector(){return e.block_selector},set block_sele
ctor(t){u(t)?e.block_selector+=", ".concat(t):""!=t&&(0,1.Z)("An invalid se
ssion_replay_block_selector was provided and will not be used",t)},get mask

```

```
In [5]: soup=BeautifulSoup(page.text)
        soup
```

```
Out[5]: <!DOCTYPE html>
<html><head>
<meta content="4d8KRpUCzf-ghquF_HLNphibw_xRvUS3kAvzpnF2KlA" name="google-si
te-verification"/>
<meta content="-TTsyPFCOB4bQlMHawiO4oZGcGX9pxqenB_NZNhhhgA" name="google-si
te-verification"/>
<meta content="wQa0CQvt7fd2VIoa9qhkevhwuBvhRwiaB3hVizI90Y" name="google-si
te-verification"/>
<meta content="o7KS3-fVt15m4Cc1HeYh5HIpv532A4FvvYQzIf-LnsQ" name="google-si
te-verification"/>
<link href="https://images.bewakoof.com" rel="preconnect"/>
<link href="https://images.bewakoof.com" rel="dns-prefetch"/>
<link href="https://static.bewakoof.com" rel="preconnect"/>
<link href="https://static.bewakoof.com" rel="dns-prefetch"/>
<link href="https://cdn.polyfill.io" rel="preconnect"/>
<link href="https://cdn.polyfill.io" rel="dns-prefetch"/>
<link href="https://www.googletagmanager.com" rel="preconnect"/>
<link href="https://www.googletagmanager.com" rel="dns-prefetch"/>
<link href="https://www.google-analytics.com" rel="dns-prefetch"/>
<link href="https://www.google-analytics.com" rel="dns-prefetch"/>
```

```
In [6]: soup.h1.text
```

```
Out[6]: 'Men Clothing'
```



```
In [10]: catogory=soup.find_all('div',class_="tag_label_plp")
```

```
In [12]: amount=[]
brand=[]
material=[]
```

```
In [13]: for i in price:
price=i.text
amount.append(price)
for i in brand_name:
brand_name=i.text
brand.append(brand_name)
for i in catogory:
catogory=i.text
material.append(catogory)
```

```
In [15]: a=pd.DataFrame({"amount":amount,"brand":brand,"material":material})
a
```

Out[15]:

	amount	brand	material
0	₹749	bewakoof x marvel	100% COTTON
1	₹699	Bewakoof®	100% COTTON
2	₹599	Bewakoof®	100% COTTON
3	₹449	Bewakoof®	100% COTTON
4	₹1199	Bewakoof®	SOFT TERRY COTTON
5	₹699	Bewakoof®	100% COTTON
6	₹649	bewakoof x garfield	100% COTTON
7	₹699	Bewakoof®	100% COTTON
8	₹699	Bewakoof®	100% COTTON
9	₹699	Bewakoof®	100% COTTON

```
In [ ]: sc
```

```
In [16]: excel=a.to_csv("bewakoof.csv")
```

```
In [ ]:
```