

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

```
In [2]: url="https://www.amazon.in/s?k=tops+for+women+western&i=apparel&crd=E7U7JPHM57AK&srefix=tops+%2Capparel%2C260"
```

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

```
Out[3]: <Response [503]>
```

```
In [4]: headers={      'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Ch
        }
```

```
In [5]: response=requests.get(url,headers=headers)
        response
```

```
Out[5]: <Response [200]>
```

```
In [6]: if response.status_code == 200:
        print("Successfully fetched the page!")
        else:
            print(f"Failed to retrieve the page. Status code: {response.status_code}")
```

Successfully fetched the page!

```
In [7]: soup= BeautifulSoup(response.content,"html.parser")
        #soup
```

```
In [8]: # print(soup.prettify())
```

```
In [9]: #tops=soup.find_all("div",class_="sg-col-inner")
        #tops
```

```
In [10]: tops=soup.find_all("div",class_="a-section a-spacing-base")
        #tops
```

```
In [11]: len(tops)
```

```
Out[11]: 65
```

```
In [12]: print(len(tops))
```

65

```
In [13]: top1=[]
```

```
In [14]: for top in tops:
        top_name=top.find("span",class_="a-size-base-plus a-color-base")
        if top_name:
            Title=top_name.get_text(strip=True)
            Total_price=top.find("span",class_="a-price-whole")
            if Total_price:
                Price=Total_price.get_text(strip=True)
            top1.append(
                {'Title':Title,
                'Price':Price,
                }
            )
```

```
In [15]: print(top1)
```

```
[{'Title': 'Leriya Fashion', 'Price': '459'}, {'Title': 'DTR FASHION', 'Price': '319'}, {'Title': 'TTFPL FASHION STORE', 'Price': '389'}, {'Title': 'W for Woman', 'Price': '1,589'}, {'Title': 'Gufrina', 'Price': '399'}, {'Title': 'Dream Beauty Fashion', 'Price': '199'}, {'Title': 'Toochki', 'Price': '490'}, {'Title': 'GRECIILOOKS', 'Price': '399'}, {'Title': 'FERY LONDON', 'Price': '539'}, {'Title': 'Label RITU KUMAR', 'Price': '690'}, {'Title': 'SheWill', 'Price': '368'}, {'Title': 'JUNEBERRY', 'Price': '349'}, {'Title': 'TOPLOT', 'Price': '399'}, {'Title': 'KOTTY', 'Price': '273'}, {'Title': 'rytras', 'Price': '499'}, {'Title': 'DEIDAD', 'Price': '294'}, {'Title': 'SheWill', 'Price': '368'}, {'Title': 'LITZO', 'Price': '499'}, {'Title': 'rytras', 'Price': '479'}, {'Title': 'NAINVISH', 'Price': '349'}, {'Title': 'Sheetal Associates', 'Price': '281'}, {'Title': 'Dream Beauty Fashion', 'Price': '205'}, {'Title': 'Max', 'Price': '719'}, {'Title': 'rytras', 'Price': '499'}, {'Title': 'Leriya Fashion', 'Price': '459'}, {'Title': 'HIGH STAR', 'Price': '864'}, {'Title': 'GRECIILOOKS', 'Price': '549'}, {'Title': 'HIGH STAR', 'Price': '712'}, {'Title': 'HIGH STAR', 'Price': '768'}, {'Title': 'Max', 'Price': '879'}, {'Title': 'POPWINGS', 'Price': '299'}, {'Title': 'VastraVibes', 'Price': '449'}, {'Title': 'Carlton London', 'Price': '499'}, {'Title': 'rytras', 'Price': '499'}, {'Title': 'Istyle Can', 'Price': '299'}, {'Title': 'PLADO', 'Price': '499'}, {'Title': 'Gufrina', 'Price': '399'}, {'Title': 'KERI PERRY', 'Price': '594'}, {'Title': 'TADKEE', 'Price': '499'}, {'Title': 'Leriya Fashion', 'Price': '399'}, {'Title': 'Sugathari', 'Price': '344'}, {'Title': 'Max', 'Price': '399'}, {'Title': 'VERO MODA', 'Price': '1,624'}, {'Title': 'Shasmi', 'Price': '399'}, {'Title': 'KOTTY', 'Price': '266'}, {'Title': 'Maaesa', 'Price': '649'}, {'Title': 'Max', 'Price': '599'}, {'Title': 'Max', 'Price': '359'}, {'Title': 'Max', 'Price': '269'}, {'Title': 'Generic', 'Price': '299'}, {'Title': 'Mack Jonney', 'Price': '299'}, {'Title': 'Marie Claire', 'Price': '499'}, {'Title': 'Max', 'Price': '349'}, {'Title': 'GRECIILOOKS', 'Price': '399'}, {'Title': 'Style Quotient', 'Price': '730'}, {'Title': 'BIBA', 'Price': '800'}, {'Title': 'GRECIILOOKS', 'Price': '449'}, {'Title': 'FIORRA', 'Price': '499'}, {'Title': 'Arayna', 'Price': '499'}, {'Title': 'PERFETTE WOLLO', 'Price': '399'}, {'Title': 'Zilcon', 'Price': '399'}, {'Title': 'HIGH STAR', 'Price': '789'}, {'Title': 'High Star Clothing', 'Price': '1,039'}, {'Title': 'Janasya', 'Price': '449'}, {'Title': 'KBZ', 'Price': '716'}]
```

In [16]: `df=pd.DataFrame(top1)`

In [17]: `df`

Out[17]:

	Title	Price
0	Leriya Fashion	459
1	DTR FASHION	319
2	TTFPL FASHION STORE	389
3	W for Woman	1,589
4	Gufrina	399
...
60	Zilcon	399
61	HIGH STAR	789
62	High Star Clothing	1,039
63	Janasya	449
64	KBZ	716

65 rows × 2 columns

In [18]: `df.dropna()`

Out[18]:

	Title	Price
0	Leriya Fashion	459
1	DTR FASHION	319
2	TTFPL FASHION STORE	389
3	W for Woman	1,589
4	Gufrina	399
...
60	Zilcon	399
61	HIGH STAR	789
62	High Star Clothing	1,039
63	Janasya	449
64	KBZ	716

65 rows × 2 columns

In [19]: `df.duplicated().sum()`

Out[19]: `np.int64(6)`

In [20]: `df.drop_duplicates()`

Out[20]:

	Title	Price
0	Leriya Fashion	459
1	DTR FASHION	319
2	TTFPL FASHION STORE	389
3	W for Woman	1,589
4	Gufrina	399
5	Dream Beauty Fashion	199
6	Toochki	490
7	GRECILOOKS	399
8	FERY LONDON	539
9	Label RITU KUMAR	690
10	SheWill	368
11	JUNEBERRY	349
12	TOPLOT	399
13	KOTTY	273
14	rytras	499
15	DEIDAD	294
17	LITZO	499
18	rytras	479
19	NAINVISH	349
20	Sheetal Associates	281
21	Dream Beauty Fashion	205
22	Max	719
25	HIGH STAR	864
26	GRECILOOKS	549
27	HIGH STAR	712
28	HIGH STAR	768
29	Max	879
30	POPWINGS	299
31	VastraVibes	449
32	Carlton London	499
34	Istyle Can	299
35	PLADO	499
37	KERI PERRY	594
38	TADKEE	499
39	Leriya Fashion	399
40	Sugathari	344
41	Max	399
42	VERO MODA	1,624
43	Shasmi	399
44	KOTTY	266
45	Maaesa	649
46	Max	599
47	Max	359
48	Max	269
49	Generic	299
50	Mack Jonney	299
51	Marie Claire	499
52	Max	349
54	Style Quotient	730
55	BIBA	800

56	GRECIILOOKS	449
57	FIORRA	499
58	Arayna	499
59	PERFETTE WOLLO	399
60	Zilcon	399
61	HIGH STAR	789
62	High Star Clothing	1,039
63	Janasya	449
64	KBZ	716

In [21]: `df.isnull().sum()`

Out[21]: Title 0
Price 0
dtype: int64

In [22]: `df.describe()`

Out[22]:

	Title	Price
count	65	65
unique	44	38
top	Max	399
freq	7	10

In [23]: `df.tail()`

Out[23]:

	Title	Price
60	Zilcon	399
61	HIGH STAR	789
62	High Star Clothing	1,039
63	Janasya	449
64	KBZ	716

In [24]: `df.head()`

Out[24]:

	Title	Price
0	Leriya Fashion	459
1	DTR FASHION	319
2	TTFPL FASHION STORE	389
3	W for Woman	1,589
4	Gufrina	399

In [25]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 65 entries, 0 to 64
Data columns (total 2 columns):
#   Column  Non-Null Count  Dtype
---  -
0    Title   65 non-null      object
1    Price   65 non-null      object
dtypes: object(2)
memory usage: 1.1+ KB
```

In [26]: `df.sample()`

Out[26]:

	Title	Price
8	FERY LONDON	539