

```
In [ ]: #kartik
```

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

```
In [2]: url = 'https://www.moneycontrol.com/'
        response = requests.get(url)
        html = response.content
```

```
In [3]: soup= BeautifulSoup(html, 'html.parser')
        soup
```

```
Out[3]: <!DOCTYPE html>

<html lang="en">
<head>
<meta content="text/html; charset=utf-8" http-equiv="Content-Type"/>
<meta content="600" http-equiv="refresh"/>
<meta content="ie=edge" http-equiv="x-ua-compatible"/>
<meta content="width=device-width, initial-scale=1.0, maximum-scale=5.0" name="viewport"/>
<title>Business News Today: Stock and Share Market News, Economy and Finance News, Sensex, Nifty, Global Market, NSE, BSE Live IPO News</title>
<link href="https://images.moneycontrol.com/images/responsive/common/apple-touch-icon.png" rel="apple-touch-icon"/>
<meta content="Business News Today: Read the latest business news on the Indian economy, global market, upcoming IPOs and more. Get Live Stock Price, Stock and Share market news, Finance News, Sensex, Nifty Live, Commodity Market, IPO news, economy news, and personal finance news today only at Money control" name="description"/>
<meta content="Business News, Business News Today, Market News, Share Market News, Share Market news, Finance news, Live Stock Market, Commodity Market, IPO news, economy news, and personal finance news today only at Money control" name="description"/>
```

```
In [4]: links = soup.find_all('a')
for link in links:
    print(link.get('href'))
javascript:void(0)
javascript;;
javascript;;
javascript;;
javascript;;
javascript;;
javascript;;
javascript;;
javascript;;
javascript;;
https://www.moneycontrol.com/portfolio-management/user/update_profile (http
s://www.moneycontrol.com/portfolio-management/user/update_profile)
https://www.moneycontrol.com/promos/pro.php (https://www.moneycontrol.com/p
romos/pro.php)
https://www.moneycontrol.com/india/bestportfoliomanager/investment-tool (ht
tps://www.moneycontrol.com/india/bestportfoliomanager/investment-tool)
https://www.moneycontrol.com/bestportfolio/wealth-management-tool/stock_wat
chlist (https://www.moneycontrol.com/bestportfolio/wealth-management-tool/s
tock_watchlist)
https://mmb.moneycontrol.com/ (https://mmb.moneycontrol.com/)
```

```
In [5]: import re

newsHyperLinks = []

title= []
VALID_ET_NEWS_LINK_REGEX = re.compile(r"www.moneycontrol.com/news/(business|co

links =soup.find_all('a' , href=True)

for link in links:
    hrefVal = link.get('href')
    r1=VALID_ET_NEWS_LINK_REGEX.search(hrefVal)

    if r1:
        newsHyperLinks.append(hrefVal)
        title_data=link.get_text()
        title.append(title_data)
print(len(newsHyperLinks))
print(newsHyperLinks[1])
print(title[2])
```

247

<https://www.moneycontrol.com/news/business/markets/wall-street-ends-down-on-nagging-uncertainty-about-fed-rate-path-10523101.html> (https://www.moneycontro
l.com/news/business/markets/wall-street-ends-down-on-nagging-uncertainty-abou
t-fed-rate-path-10523101.html)

World Bank appoints Indian-origin Ajay Banga as president

```
In [6]: import pandas as pd

data = {'News Title': title, 'News Link': newsHyperLinks}
df = pd.DataFrame(data)
```

```
In [8]: df.head()
```

	News Title	News Link
0	Govt exp...	https://www.moneycontrol.com/news/business/eco...
1	Wall Str...	https://www.moneycontrol.com/news/business/mar...
2	World Ba...	https://www.moneycontrol.com/news/business/ban...
3	US FTC t...	https://www.moneycontrol.com/news/world/us-ftc...
4	US sprin...	https://www.moneycontrol.com/news/world/us-spr...

In []:

In []: