

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

#Q1) Write a python program to display all the header tags from wikipedia.org and make data frame.

```
page = requests.get('https://en.wikipedia.org/wiki/Main_Page')
# page
soup = BeautifulSoup(page.content)
# soup
# first_header = soup.find('header', class_="mw-body-header vector-page-titlebar vectblank")
# first_header
all_header = []
for i in soup.find_all(['h1' , 'h2' , 'h3' , 'h4' , 'h5' , 'h6']) :
    all_header.append(i.text)
# print(all_header)
df = pd.DataFrame({'Header' : all_header})
df
```

Header

- | | |
|---|-------------------------------|
| 0 | Main Page |
| 1 | Welcome to Wikipedia |
| 2 | From today's featured article |
| 3 | Did you know ... |
| 4 | In the news |
| 5 | On this day |
| 6 | Today's featured picture |
| 7 | Other areas of Wikipedia |
| 8 | Wikipedia's sister projects |
| 9 | Wikipedia languages |

"""Q2) Write a python program to display list of respected former presidents of India(i.e. Name , Term of office) from <https://presidentofindia.nic.in/former-presidents.htm> and make data frame."""

```
url = "https://presidentofindia.nic.in/former-presidents"
page = requests.get(url)
page #response < 200 , so i can scrap data from website easily
soup = BeautifulSoup(page.content)
# soup
# first_name = soup.find('div' , class_="desc-sec")
# first_name
presidents_name = soup.find_all('div' , class_="desc-sec")
list_of_presidents = []
# for i in soup.find_all('div' , class_="desc-sec") :
#     list_of_presidents.append(i)
# list_of_presidents
for president in presidents_name :
    name = president.find('h3').text.strip()
    term = president.find('h5').text.strip()
    # term_of_office = president.find('p')
    list_of_presidents.append({'Name' : name , 'Term' : term })
df = pd.DataFrame(list_of_presidents)
df
```

	Name	Term
0	Shri Ram Nath Kovind	14th President of India
1	Shri Pranab Mukherjee	13th President of India
2	Smt Pratibha Devisingh Patil	12th President of India
3	DR. A.P.J. Abdul Kalam	11th President of India
4	Shri K. R. Narayanan	10th President of India
5	Dr Shankar Dayal Sharma	9th President of India
6	Shri R Venkataraman	8th President of India
7	Giani Zail Singh	7th President of India
8	Shri Neelam Sanjiva Reddy	6th President of India
9	Dr. Fakhruddin Ali Ahmed	5th President of India
10	Shri Varahagiri Venkata Giri	4th President of India
11	Dr. Zakir Husain	3rd President of India
12	Dr. Sarvepalli Radhakrishnan	2nd President of India
13	Dr. Rajendra Prasad	1st President of India

```
# Q5) Write a python program to scrape mentioned news details from https://www.cnbc.com/world/?region=world and
# make data frame
# i) Headline
# ii) Time
# iii) News Link
```

```
page = requests.get('https://www.cnbc.com/world/?region=world')
# page
soup = BeautifulSoup(page.content)
# soup
# first_headline = soup.find('div' , class_ = "LatestNews-container")
# first_headline
data = soup.find_all('div' , class_ = "LatestNews-container")
# data
scraping_data = []
for i in data :
    headline = i.find('a' , class_ = "LatestNews-headline").text.strip()
# headline
    time = i.find('time' , class_ = "LatestNews-timestamp")
# time
    news_link = i.find('a' , class_ = "LatestNews-headline")['href']
# news_link
    scraping_data.append({'Headline' : headline , 'Time' : time , 'News_Link' : news_link})
# scraping_data
df = pd.DataFrame(scraping_data)
df
# time = soup.find('time' , class_ = "LatestNews-timestamp")
# time
# news_link = soup.find('a' , class_ = "InvestingClubPill-investingClubPillLink")
# news_link
```

	Headline	Time	News_Link
0	Nikki Haley slams Trump for trying to torpedo ...	[14 Min Ago]	https://www.cnbc.com/2024/01/28/nikki-haley-sl...
1	Unions, with power and popularity rising, are ...	[38 Min Ago]	https://www.cnbc.com/2024/01/28/unions-with-po...
2	Online sports betting stocks are primed for a ...	[41 Min Ago]	https://www.cnbc.com/2024/01/28/online-sports-...
3	Here's what we think about when deciding how m...	[2 Hours Ago]	https://www.cnbc.com/2024/01/28/heres-what-we-...
4	I've spent 20 years studying how to raise succ...	[2 Hours Ago]	https://www.cnbc.com/2024/01/28/i-spent-20-yea...
5	Mark Cuban: Here's the moment I knew I wouldn't...	[2 Hours Ago]	https://www.cnbc.com/2024/01/28/mark-cuban-the...
6	The 7 most popular countries for U.S. workers ...	[3 Hours Ago]	https://www.cnbc.com/2024/01/28/the-7-most-pop...
7	36-year-old quit her job to write novels—how s...	[3 Hours Ago]	https://www.cnbc.com/2024/01/28/how-36-year-ol...
8	Workers are paying to get wages early. It's 'p...	[3 Hours Ago]	https://www.cnbc.com/2024/01/28/why-one-expert...
9	Big pharma is preparing to lose revenue from b...	[4 Hours Ago]	https://www.cnbc.com/2024/01/28/big-pharma-mer...
10	Chinese stocks are starting to turn around. Ho...	[4 Hours Ago]	https://www.cnbc.com/2024/01/28/chinese-stocks...
11	Earnings playbook: Your guide to the busiest w...	[5 Hours Ago]	https://www.cnbc.com/2024/01/28/earnings-playb...
12	Wall Street touted 5 portfolio stocks this wee...	[22 Hours Ago]	https://www.cnbc.com/2024/01/27/wall-street-to...
13	Beijing intensifies military pressure on Taiwa...	[23 Hours Ago]	https://www.cnbc.com/2024/01/27/beijing-intens...
14	How the Apple iPhone became one of the best-se...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/how-the-apple-...
15	Mark Ruffalo 'couldn't afford a car' when he f...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/mark-ruffalo-c...
16	10 of our stocks report earnings next week. He...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/10-of-our-stoc...
17	Florida is the No. 1 state to retire in 2024—N...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/best-states-to...
18	The future of American amusement park in Disne...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/six-flags-futu...
19	2 surprising reasons why skipping breakfast is...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/2-reasons-why-...
20	28-year-old worth more than \$500,000: 5 things...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/28-year-old-wi...
21	Gen Z vs. their parents: How the generations s...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/gen-z-vs-their...
22	How Guyana's big oil boom turned it into the w...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/how-guyanas-bi...
23	Why kids TV content is vital to streaming subs...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/kids-streaming...
24	These energy stocks may rally after oil prices...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/energy-stocks-...
25	Buy these top tech stocks into earnings, Goldm...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/goldmans-top-t...
26	Quality stocks are dominating the market, up 2...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/quality-stocks...
27	A handful of space companies are running out o...	[January 27, 2024]	https://www.cnbc.com/2024/01/27/three-space-co...
28	WWE founder Vince McMahon resigns from TKO Gro...	[January 26, 2024]	https://www.cnbc.com/2024/01/26/wwe-founder-vi...
29	Why weakness in small caps may be a short-term...	[January 26, 2024]	https://www.cnbc.com/2024/01/26/why-weakness-i...

```

url = "https://www.journals.elsevier.com/artificial-intelligence/most-downloaded-articles"
page = requests.get(url)
soup = BeautifulSoup(page.content, 'html.parser')

# The class 'sc-orwwe2-3 jOMrrY' might change, inspect the HTML to find the correct class
articles = soup.find_all('li', class_='sc-9zxyh7-1 sc-9zxyh7-2 kOEIEO hvoVxs')

scraping_data = []

for item in articles:
    # Use find() instead of find_all() to get a single element
    paper_title = item.find('a', class_='sc-5smygv-0 fIXTHm').text.strip()
    # scraping_data.append({'Paper Title' : paper_title})
    # scraping_data

    authors = item.find('span', class_='sc-1w3fpd7-0 dnCnAO').text.strip()
    date = item.find('span', class_='sc-1thf9ly-2 dvvggWt').text.strip()
    paper_url = item.find('a', class_='sc-5smygv-0 fIXTHm')['href']

    # paper_url = item.find('a', class_='anchor doi anchor-default')

    scraping_data.append({'Paper Title': paper_title, 'Authors': authors, 'Published Date': date, 'Paper URL': paper_url})
scraping_data

df = pd.DataFrame(scraping_data)
df
# Print the scraped data
# for data in scraping_data:
#     print(data)

```

	Paper Title	Authors	Published Date	Paper URL
0	Reward is enough	David Silver, Satinder Singh, Doina Precup, Ri...	October 2021	https://www.sciencedirect.com/science/article/...
1	Explanation in artificial intelligence: Insign...	Tim Miller	February 2019	https://www.sciencedirect.com/science/article/...
2	Creativity and artificial intelligence	Margaret A. Boden	August 1998	https://www.sciencedirect.com/science/article/...
3	Conflict-based search for optimal multi-agent ...	Guni Sharon, Roni Stern, Ariel Felner, Nathan ...	February 2015	https://www.sciencedirect.com/science/article/...
4	Knowledge graphs as tools for explainable mach...	Ilaria Tiddi, Stefan Schlobach	January 2022	https://www.sciencedirect.com/science/article/...
5	Law and logic: A review from an argumentation ...	Henry Prakken, Giovanni Sartor	October 2015	https://www.sciencedirect.com/science/article/...
6	Between MDPs and semi-MDPs: A framework for te...	Richard S. Sutton, Doina Precup, Satinder Singh	August 1999	https://www.sciencedirect.com/science/article/...
7	Explaining individual predictions when feature...	Kjersti Aas, Martin Jullum, Anders Løland	September 2021	https://www.sciencedirect.com/science/article/...
8	Multiple object tracking: A literature review	Wenhan Luo, Junliang Xing and 4 more	April 2021	https://www.sciencedirect.com/science/article/...
9	A survey of inverse reinforcement learning: Ch...	Saurabh Arora, Prashant Doshi	August 2021	https://www.sciencedirect.com/science/article/...
10	Evaluating XAI: A comparison of rule-based and...	Jasper van der Waa, Elisabeth Nieuwburg, Anita...	February 2021	https://www.sciencedirect.com/science/article/...
11	Explainable AI tools for legal reasoning about...	Joe Collenette, Katie Atkinson, Trevor Bench-C...	April 2023	https://www.sciencedirect.com/science/article/...
12	Hard choices in artificial intelligence	Roel Dobbe, Thomas Krendl Gilbert, Yonatan Mintz	November 2021	https://www.sciencedirect.com/science/article/...
13	Assessing the communication gap between AI mod...	Oskar Wysocki, Jessica Katharine Davies and 5 ...	March 2023	https://www.sciencedirect.com/science/article/...
14	Explaining black-box classifiers using post-ho...	Eoin M. Kenny, Courtney Ford, Molly Quinn, Mar...	May 2021	https://www.sciencedirect.com/science/article/...
15	The Hanabi challenge: A new frontier for AI re...	Nolan Bard, Jakob N. Foerster and 13 more	March 2020	https://www.sciencedirect.com/science/article/...
16	Wrappers for feature subset selection	Ron Kohavi, George H. John	December 1997	https://www.sciencedirect.com/science/article/...
17	Artificial cognition for social human–robot in...	Séverin Lemaignan, Mathieu Warnier and 3 more	June 2017	https://www.sciencedirect.com/science/article/...
18	A review of possible effects of cognitive bias...	Tomáš Kliegr, Štěpán Bahník, Johannes Färnkranz	June 2021	https://www.sciencedirect.com/science/article/...
19	The multifaceted impact of Ada Lovelace in the...	Luigia Carlucci Aiello	June 2016	https://www.sciencedirect.com/science/article/...
20	Robot ethics: Mapping the issues for a mechani...	Patrick Lin, Keith Abney, George Bekey	April 2011	https://www.sciencedirect.com/science/article/...
21	Reward (Mis)design for autonomous driving	W. Bradley Knox, Alessandro Allievi and 3 more	March 2023	https://www.sciencedirect.com/science/article/...
22	Planning and acting in partially observable st...	Leslie Pack Kaelbling, Michael L. Littman, Ant...	May 1998	https://www.sciencedirect.com/science/article/...
23	What do we want from Explainable Artificial In...	Markus Langer, Daniel Oster and 6 more	July 2021	https://www.sciencedirect.com/science/article/...


```
# Q7)Write a python program to scrape mentioned details from dineout.co.inand make data frame.  
# i) Restaurant name  
# ii) Cuisine  
# iii) Location  
# iv) Ratings  
# v) Image URL
```

```
page = requests.get('https://www.dineout.co.in/delhi-restaurants/buffet-special')  
# page  
soup = BeautifulSoup(page.content)  
# soup  
# rest_details = soup.find_all('div' , class_ = 'restnt-main-wrap clearfix')  
# rest_details  
# rest_details = soup.find_all('div' , class_ = 'restnt-card-wrap-new')  
restaurants = soup.find_all('div', class_='restnt-card restaurant')  
# restaurants  
details = []  
for i in restaurants :  
    restaurant_name = i.find('a' , class_ = 'restnt-name ellipsis').text.strip()  
    cuisine_element = i.find('span' , class_ = 'double-line-ellipsis')  
    cuisine = cuisine_element.text.strip() if cuisine_element else None  
    location = i.find('div' , class_ = 'restnt-loc ellipsis').text.strip()  
    rating = i.find('div' , class_ = 'restnt-rating rating-4').text.strip()  
    image = i.find('img' , class_ = 'no-img')  
    image_url = image['data-src']  
    details.append({'Restaurant Name' : restaurant_name , 'Cuisine' :cuisine , 'Location' : location , 'Rating' : rating , 'Image URL' : image_url})  
#details  
df = pd.DataFrame(details)  
df
```

	Restaurant Name		Cuisine		Location	Rating	Image URL
0	Castle Barbeque	₹ 2,000 for 2 (approx)	Chinese, North Indian		Connaught Place, Central Delhi	4	https://im1.dineout.co.in/images/uploads/resta...
1	Cafe Knosh	₹ 3,000 for 2 (approx)	Italian, Continental	The Leela Ambience Convention Hotel,Shahdara, ...		4.3	https://im1.dineout.co.in/images/uploads/resta...
2	India Grill	₹ 2,400 for 2 (approx)	North Indian, Italian		Hilton Garden Inn,Saket, South Delhi	3.9	https://im1.dineout.co.in/images/uploads/resta...
3	The Barbeque Company	₹ 1,700 for 2 (approx)	North Indian, Chinese		Gardens Galleria,Sector 38A, Noida	3.9	https://im1.dineout.co.in/images/uploads/resta...
4	Delhi Barbeque	₹ 1,800 for 2 (approx)	North Indian	Taurus Sarovar Portico,Mahipalpur, South Delhi		3.7	https://im1.dineout.co.in/images/uploads/resta...
5	The Monarch - Bar Be Que Village	₹ 1,900 for 2 (approx)	North Indian	Indirapuram Habitat Centre,Indirapuram, Ghaziabad		3.8	https://im1.dineout.co.in/images/uploads/resta...
6	The Barbeque Times	₹ 1,500 for 2 (approx)	North Indian, Contine...		M2K Corporate Park,Sector 51, Gurgaon	4.1	https://im1.dineout.co.in/images/uploads/resta...