1) Write a python program to display all the header tags from wikipedia.org

```
from urllib.request import urlopen
from bs4 import BeautifulSoup
import requests
import pandas as pd

page=requests.get('https://en.wikipedia.org/wiki/Main_Page')

soup=BeautifulSoup(page.content)
print(soup.prettify())

header=["h1", "h2", "h3"]
for tags in soup.find_all(header):
    print(tags.name + ' -> ' + tags.text.strip())
```

output is:

```
h1 -> Main Page
h1 -> Welcome to Wikipedia
h2 -> From today's featured article
h2 -> Did you know ...
h2 -> In the news
h2 -> On this day
h2 -> Today's featured picture
h2 -> Other areas of Wikipedia
h2 -> Wikipedia's sister projects
h2 -> Wikipedia languages
h2 -> Navigation menu
h3 -> Personal tools
h3 -> Namespaces
h3 -> Views
h3 -> Search
h3 -> Navigation
h3 -> Contribute
h3 -> Tools
h3 -> Print/export
h3 -> In other projects
h3 -> Languages
```

----

2) Write a python program to display IMDB's Top rated 100 movies' data (i.e. name, rating, year of release) and make data frame.

url_link=requests.get('https://www.imdb.com/chart/top/?ref_=nv_mv_250')

movie_list=BeautifulSoup(url_link.content,"html.parser")

movie_list

scrapped_movies=movie_list.find_all('td',class_='titleColumn')

scrapped_movies

```python
movies_list=[]
for movie in scrapped_movies:
    movie=movie.get_text().replace("\n","")
    movie=movie.strip(" ")
    movies_list.append(movie)
movies_list


scrapped_ratings=movie_list.find_all('td',class_='ratingColumn imdbRating')
scrapped_ratings


ratings=[]
for rating in scrapped_ratings:
    rating=rating.get_text("\n"," ")
    ratings.append(rating)
ratings


data=pd.DataFrame()
data['Movies Name']=movies_list
data['Ratings']=ratings
data.head(100)
```

output is:

	Movies Name	Ratings
0	1. The Shawshank Redemption(1994)	9.2
1	2. The Godfather(1972)	9.2
2	3. The Dark Knight(2008)	9.0
3	4. The Godfather Part II(1974)	9.0
4	5. 12 Angry Men(1957)	8.9
...	...	...
95	96. Jagten(2012)	8.3
96	97. M - Eine Stadt sucht einen Mörder(1931)	8.3

	Movies Name	Ratings
97	98. North by Northwest(1959)	8.3
98	99. Vertigo(1958)	8.2
99	100. Idi i smotri(1985)	8.2

100 rows × 2 columns

3) Write a python program to display IMDB's Top rated 100 Indian movies' data (i.e. name, rating, year of release) and make data frame

```
page=requests.get('https://www.imdb.com/india/top-rated-indian-
movies/?pf_rd_m=A2FGELUUNOQJNL&pf_rd_p=461131e5-5af0-4e50-bee2-
223fad1e00ca&pf_rd_r=1Y5PAT9634QHA08P693G&pf_rd_s=center-
1&pf_rd_t=60601&pf_rd_i=india.toprated&ref_=fea_india_ss_toprated_india_tr_india250_s
m')

indian_movie_list=BeautifulSoup(page.content,"html.parser")
indian_movie_list

scrapped_indian_movies=indian_movie_list.find_all('td',class_='titleColumn')
scrapped_indian_movies

indian_movies_list=[]

for movie in scrapped_indian_movies:
    movie=movie.get_text().replace("\n","")
    movie=movie.strip(" ")
    indian_movies_list.append(movie)
indian_movies_list

scrapped_ratings=movie_list.find_all('td',class_='ratingColumn imdbRating')
scrapped_ratings

ratings=[]
for rating in scrapped_ratings:
    rating=rating.get_text("\n"," ")
    ratings.append(rating)
ratings

data=pd.DataFrame()
data['Indian Movies Name']=indian_movies_list
data['Ratings']=ratings
data.head(100)
```

output is:

	Indian Movies Name	Ratings
0	1. Rocketry: The Nambi Effect (2022)	9.2
1	2. Anbe Sivam (2003)	9.2
2	3. Golmaal (1979)	9.0
3	4. Nayakan (1987)	9.0
4	5. Jai Bhim (2021)	8.9
...	...	...
95	96. Rang De Basanti (2006)	8.3
96	97. Baasha (1995)	8.3
97	98. Baahubali 2: The Conclusion (2017)	8.3
98	99. Masaan (2015)	8.2
99	100. Kahaani (2012)	8.2

100 rows × 2 columns

4) Write s python program to display list of respected former presidents of India(i.e. Name , Term of office) from https://presidentofindia.nic.in/former-presidents.ht

```python
page=requests.get('https://presidentofindia.nic.in/former-presidents.htm')

soup=BeautifulSoup(page.content,"html.parser")
print(soup.prettify())

 title=[]

for i in soup.find_all('div',class_="presidentListing"):
  i=i.get_text().replace("\n","")
  i=i.strip(" ")
  title.append(i)
title
```

Output is:

```
'Shri Ram Nath Kovind (birth - 1945)Term of Office: 25 July, 2017 to 25
July, 2022 https://ramnathkovind.nic.in',
 'Shri Pranab Mukherjee (1935-2020)Term of Office: 25 July, 2012 to 25
July, 2017 http://pranabmukherjee.nic.in',
 'Smt Pratibha Devisingh Patil (birth - 1934)Term of Office: 25 July, 2
007 to 25 July, 2012 http://pratibhapatil.nic.in',
 'DR. A.P.J. Abdul Kalam (1931-2015)Term of Office: 25 July, 2002 to 25
July, 2007 http://abdulkalam.nic.in',
 'Shri K. R. Narayanan (1920 - 2005)Term of Office: 25 July, 1997 to 25
July, 2002',
```

```
 'Dr Shankar Dayal Sharma (1918-1999)Term of Office: 25 July, 1992 to 2
5 July, 1997',
 'Shri R Venkataraman (1910-2009)Term of Office: 25 July, 1987 to 25 Ju
ly, 1992',
 'Giani Zail Singh (1916-1994)Term of Office: 25 July, 1982 to 25 July,
1987',
 'Shri Neelam Sanjiva Reddy (1913-1996)Term of Office: 25 July, 1977 to
25 July, 1982',
 'Dr. Fakhruddin Ali Ahmed (1905-1977)Term of Office: 24 August, 1974 t
o 11 February, 1977',
 'Shri Varahagiri Venkata Giri (1894-1980)Term of Office: 3 May, 1969 t
o 20 July, 1969 and 24 August, 1969 to 24 August, 1974',
 'Dr. Zakir Husain (1897-1969)Term of Office: 13 May, 1967 to 3 May, 19
69',
 'Dr. Sarvepalli Radhakrishnan (1888-1975)Term of Office: 13 May, 1962
to 13 May, 1967',
 'Dr. Rajendra Prasad (1884-1963) Term of Office: 26 January, 1950 to 1
3 May, 1962']
```

9) Write a python program to scrape mentioned details from dineout.co.in :
 i) Restaurant name ii) Cuisine iii) Location iv) Ratings v) Image URL

```python
page=requests.get('https://www.dineout.co.in/delhi-restaurants/buffet-special')

soup=BeautifulSoup(page.content,"html.parser")
print(soup.prettify())

title=[]
for i in soup.find_all('div',class_="restnt-info cursor"):
    title.append(i.text)
title

loc=[]
for i in soup.find_all('div',class_="restnt-loc ellipsis"):
    loc.append(i.text)
loc

sta=[]
for i in soup.find_all('span',class_="double-line-ellipsis"):
    sta.append(i.text.split()[6:7])
sta

image_url=[]
for i in soup.find_all("img",class_="no-img"):
    image_url.append(i["data-src"])
image_url

ratings=[]
for i in soup.find_all('div',class_="restnt-rating rating-4"):
    ratings.append(i.text)
ratings
```

```
print(len(title),len(ratings),len(loc),len(sta),len(image_url))

df=pd.DataFrame({'Title':title,'Cusine':sta,'Location':loc,'Ratings':ratings,'Image_URL':image_url})
df
```

output is:

	Title	Cusine	Location	Ratings	Image_URL
0	Castle BarbequeConnaught Place, Central Delhi	[Chinese,]	Connaught Place, Central Delhi	4.1	https://im1.dineout.co.in/images/uploads/resta...
1	Jungle Jamboree3CS Mall,Lajpat Nagar - 3, Sout...	[North]	3CS Mall,Lajpat Nagar - 3, South Delhi	3.9	https://im1.dineout.co.in/images/uploads/resta...
2	Cafe KnoshThe Leela Ambience Convention Hotel,...	[Italian,]	The Leela Ambience Convention Hotel,Shahdara, ...	4.3	https://im1.dineout.co.in/images/uploads/resta...
3	Castle BarbequePacific Mall,Tagore Garden, Wes...	[Chinese,]	Pacific Mall,Tagore Garden, West Delhi	3.9	https://im1.dineout.co.in/images/uploads/resta...
4	The Barbeque CompanyGardens Galleria,Sector 38...	[North]	Gardens Galleria,Sector 38A, Noida	4	https://im1.dineout.co.in/images/uploads/resta...
5	India GrillHilton Garden Inn,Saket, South Delhi	[North]	Hilton Garden Inn,Saket, South Delhi	3.9	https://im1.dineout.co.in/images/uploads/resta...
6	Delhi BarbequeTaurus Sarovar Portico,Mahipalpu ...	[North]	Taurus Sarovar Portico,Mahipalpur, South Delhi	3.7	https://im1.dineout.co.in/images/uploads/resta...
7	The Monarch - Bar Be Que	[North]	Indirapuram Habitat	3.8	https://im1.dineout.co.in/images/uploads/resta...

	Title	Cusine	Location	Ratings	Image_URL
	VillageIndirapura m Ha...		Centre,Indirapura m, Ghaziabad		
8	Indian Grill RoomSuncity Business Tower,Golf C...	[North]	Suncity Business Tower,Golf Course Road, Gurgaon	4.3	https://im1.dineout.co.in/images/uploads/r esta...